Status of LIGO Data Analysis System -- Hardware Procurement

Albert Lazzarini
LIGO Laboratory at Caltech

NSF Annual Review of LIGO Construction
30 April 2001
Caltech, Pasadena, CA

Outline for the Next Two Talks

- Hardware Procurement Update (Lazzarini)
  - » Actions taken in response to NSF recommendations from last review
  - » Procurement status
  - » Related activities within the Collaboration
- Software Development Update (Wiseman, Blackburn)
  - » LIGO Data Analysis System (LDAS) development & deployment
  - » LSC Activities
    - –Software development by LSC
    - –Mock Data Challenges

Progress Since the Last Review

Recommendation on Hardware Procurement Plan (HPP)

- Developed detailed Hardware Procurement Plan (HPP)
  - » Submitted to NSF for review 1 June 2000
  - » Began implementation after receipt of NSF Panel Report (02 Nov 2000)
  - » Defined three major procurement phases for the plan
    - –Submitted phasing to NSF for comment, approval 1Q2001
    - –Each phase of plan is approximately 1/3 of total system
    - –Executed Phase I 1Q2001
    - –Phase II planned for 3Q2001
    - –Phase III planned for 4Q2001

Progress Since the Last Review

Recommendation on LIGO Operations Plan for Data Analysis

- Developed a plan for operations of LIGO Data Analysis System during the LIGO I Science Run, FY2002-2006
  - » Plan submitted to NSF as part of the Renewal of Operations Proposal
  - » Plan reviewed, approved by NSF Review Panel 26 February 2001. Includes:
    - –Staff increase to cover administration of data, hardware both at Caltech and observatories
    - –Budget for WAN upgrade to higher bandwidth
    - –Budget for maintenance & upgrade of major LDAS hardware

Progress Since the Last Review

Recommendation on Data Archival & Access

- Data archive access tools under development
  - » Near-term: Client-based prototype toolkit for manipulating, querying metadata (DB2) and accessing raw frame data (HPSS)
    - –GUILD (GUI for LIGO Databases)
      - • Uses LDAS APIs for data access, queries
    - –LARS (LIGO Archive Retrieval System)
      - • Uses ftp from HPSS + Frame I/O Library for archive retrieval of data subsets
  - » By Science Run: LIGO Data Analysis System methods to obtain derived ("virtual") data products from the LIGO Data Center at Caltech
  - » LIGO + LSC are GriPhyN Collaboration members
    - –Developing applications for data mirroring, virtual data manipulation in a grid computing environment
    - –LIGO data will be mirrored to LSC Tier 2 centers

Progress Since the Last Review

Recommendation on Mock Data Challenges

- 3 Mock Data Challenges have been completed
  - » Data conditioning API integrated tests (Aug 2000)
  - » MPI API, inspiral coallescence and burst event search code integration into LIGO Data Analysis System test & validation (Jan 2001)
  - » Metadata API data insertion, retrieval tests (Apr 2001)
- Remaining MDCs later this year
  - » Test end-to-end pipelines
  - » Hierarchical inspiral search code
  - » Stochastic background search code
  - » Burst sources/unmodeled sources search code
  - » Continuous source (pulsar) search code
  - » Deep archive (HPSS) MDC

Progress Since the Last Review

Recommendation on Redundant Remote Data Archive

- Redundant, complete remote backup of LIGO data
  - » LIGO plans to archive 100% of data on high density tapes (HW compression) -- LIGO will have a dedicated silo
  - » Contacted SDSC to explore transfer of LIGO data from CACR to SDSC over high speed WAN link
    - –SDSC FTE, resource impact identified
    - –SRB toolkit could be used to re-stripe, catalog data in second archive
    - –Still exploring other options
  - » Distributed Terascale Facility
    - –LIGO has developed case use models for DT
    - –Data mirroring, backup model developed with CACR/SDSC as part of the proposal
  - » Fallback would be to store data tapes at sites

Progress Since the Last Review

Recommendation on Wide Area Network

- Upgrade of LIGO WAN to OC3
  - » Stream data to Caltech; avoid tape-to-tape transfers
  - » LIGO Laboratory granted sponsorship to Abilene/Internet2 through Caltech
  - » LIGO has identified high speed link options at each observatory site
    - –LLO: Bell South, Enron, Cogent, Neptune Communications (all commercial)
    - –LHO: ESnet (via existing DOE MOU), NaoNet (non-profit NW utilities net), Quest (commerical)
    - –Operational cost estimate included in FY2002-2006 operations proposal to NSF
      - Conservative estimate: $250k/site/yr
      - Costs vary widely; will negotiate once NSF approval received.

LIGO Data Analysis System

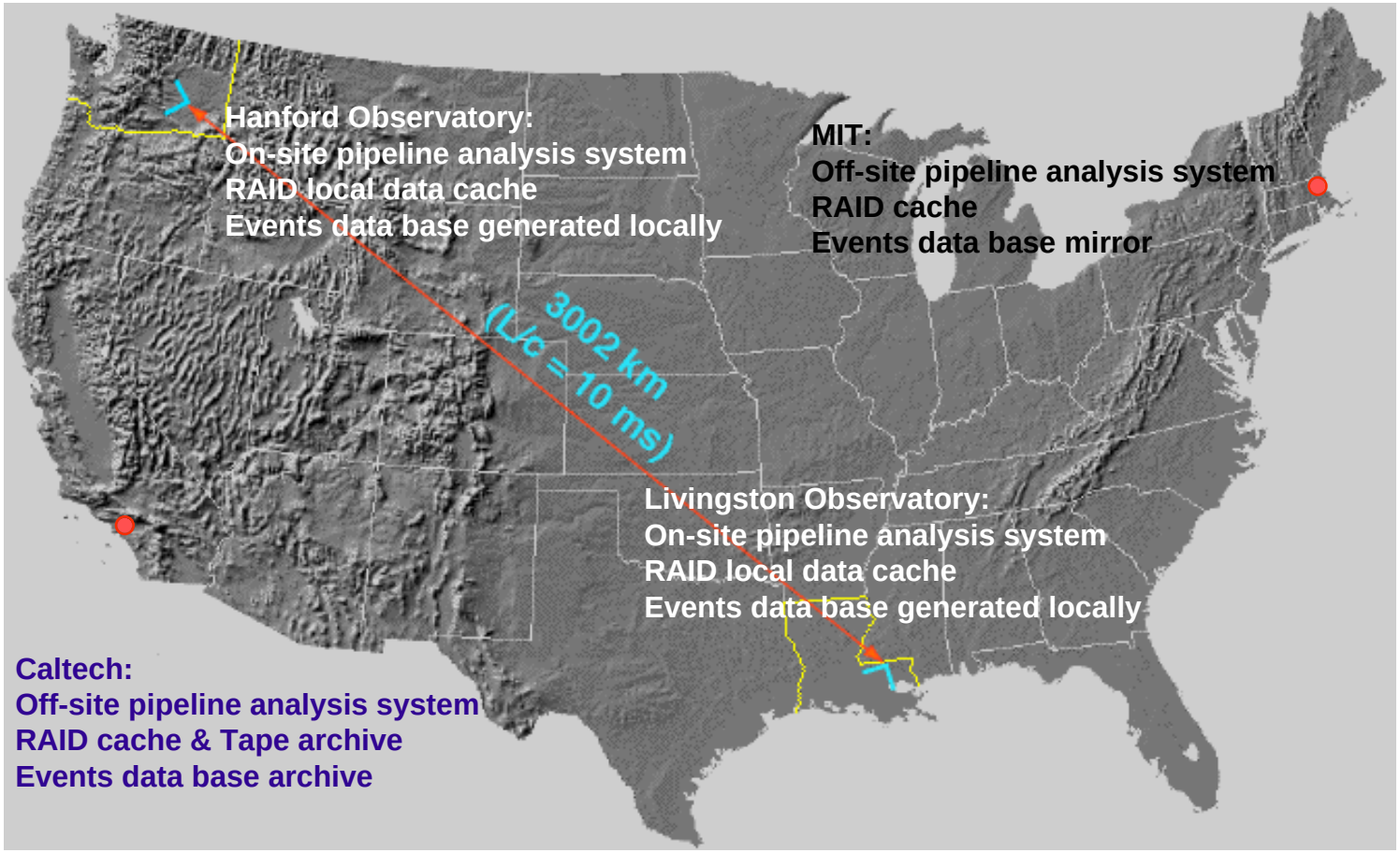
LDAS Design

- LDAS concept developed during 1997 - 1999;
  - » White paper for NSF, LSC (1997)
  - » Requirements review (1997)
  - » Design review (1999)
  - » WP, review panels included LSC members outside LIGO Laboratory
- Capabilities fulfill expected Laboratory, Scientific Collaboration needs through the LIGO I Science Run
- Mock Data Challenges provide final review & validation
- Hardware procurement plan developed during 1999 - 2000
  - » Revised plan sent to NSF after Annual Review in May 2000.
  - » Review panel approved proceeding with initial part of procurement

LDAS Hardware Procurement

- LIGO Data Analysis System (LDAS) is the last major construction component of LIGO I
  - » Computing, networking, data storage, and software systems to process, archive LIGO data stream
  - » 4 geographically isolated facilities
    - –Hanford & Livingston Observatories: on-site (near real-time) data mass storage (disk systems); pipeline analysis computing; data servers;
    - –Caltech: main archive -- tapes and disk systems; off-site pipeline analysis computing; data servers
    - –MIT: data mirror; off-site pipeline analysis computing; data servers

The LIGO Laboratory Sites



LIGO-G010185-00-E

LIGO Data Analysis System

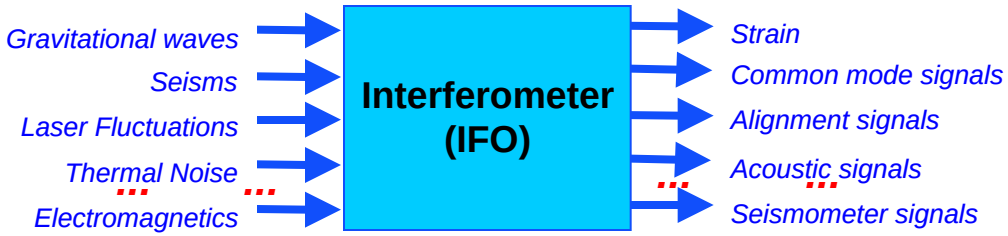
Primary Requirements

- Computationally intense parallel pipeline processes -- Linux clusters running Message Passing Interface (MPI)
  - » Detect astrophysical signals in the data stream
- Data preprocessing and conditioning -- SMP servers
  - » Provide reduced data products, prepare data for pipeline analysis
- Data archive -- RAID caches/Tape archive
  - » Disk caches at Observatories and MIT
    - –Diagnostics, data analysis
  - » Data Center at Caltech
    - –Archival of raw data and data products (metadata, events)

What are the LIGO Data?

- Continuous time series: 2^N samples/second, 2-byte ADC data & 4-byte calculated data
- Data analysis: digital signal processing -- matched optimal filters
  - » Computationally intense operations on data: ~10 MFLOPS/byte for inspiral searches with GW channel
  - » 90% of CPU time spent on f - t transformations
  - » Analysis performed in both domains
    - –Single channel, over a long time; many channels, over a short time
    - –How to cache, catalog, replicate, this *virtual data*
- Results of analysis: events, spectra, n-D representations (“images”)
  - » Environmental, instrumental “events”: vetoes
  - » Astrophysical events
  - » Time stamp, Process ID generating event, Parameters associated with event, ...
  - » Stored in a relational database for later retrieval, reanalysis:
    - –tables, “blobs”, links to data

Raw Data Stream Characteristics

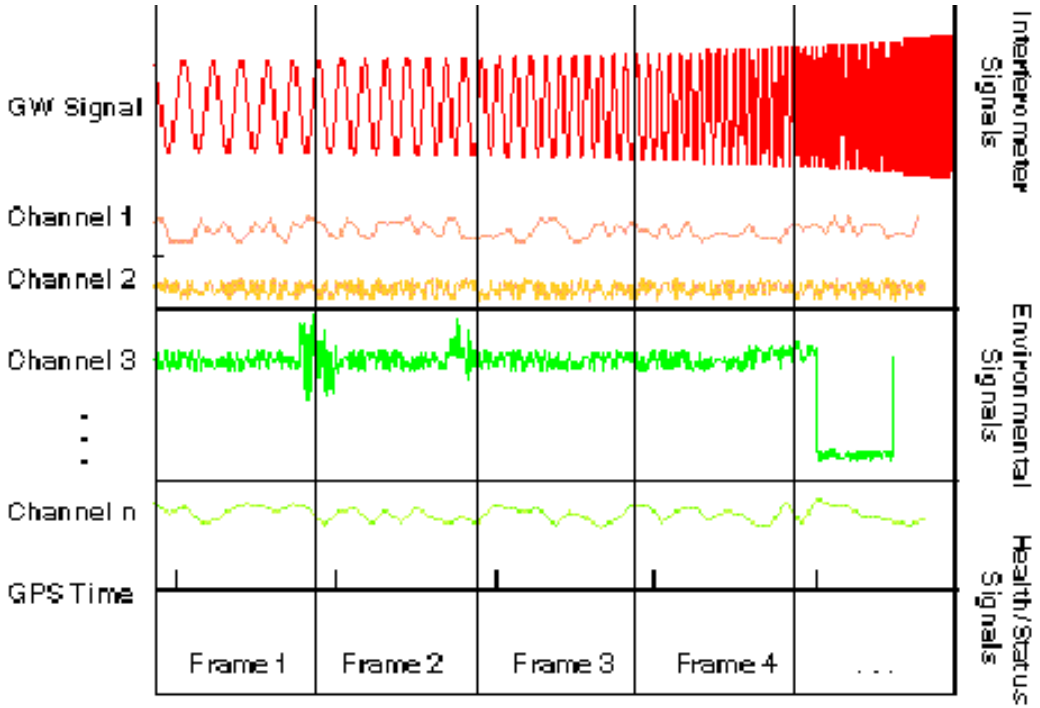


Science Channel (1%)

Interferometer Channels (86%)

Environmental Channels (13 %)

- **All interferometric detector projects have agreed on a standard data format**
- **Anticipates joint data analysis**
- **LIGO frames for 3 interferometer are ~ 9MB/s**
  - **96 kB/s strain (2 Bytes x 3 IFOs x 16 kSample/s)**
  - **~ 7.7 MB/s other interferometer signals**
  - **~ 1.2 MB/s environmental sensors**
  - ***Strain is ~1% of all data***



LIGO Data Products

Time series data

Mode	Raw and Derived Data for On-line Diagnostics	**Level 1** Full (100%) frame data for archiving	**Level 2** Strain and data summary, QA channels	**Level 3** Strain best estimate
Uncompressed Rate (MB/s)	LHO: 9.5 LLO: 5 Total: 14.5	LHO: 6 LLO: 3 Total: 9	Total: 0.300	Total: 0.006
w / 50% Hardware Compression MB/s onto tape media	-	LHO: 3 LLO: 1.5 Total:4.5	Total: 0.150	-
Data growth rate, per year of integrated running, ***TB/yr***.	-	LHO: 94.5 LLO: 47 Total:142	Total:9.5	***Total: 0.200***
Total including redundant 100% backup, ***TB/yr***.	-	LHO: 189 LLO: 94 ***Total:283***	***Total:19***	-
Purpose	For on-line monitoring of interferometers	Deep permanent archive	Science analysis, data exchange	Science analysis, data exchange
On-site look-back time	Must use real-time control and monitoring system (CDS) disk caches	LHO Disk cache: 28 d LLO Disk cache: 28 d	-	-
Off-site look-back time	-	As long as required	In perpetuity	In perpetuity

LIGO-G010185-00-E

Data products

Event database

Event source (processes or filters)

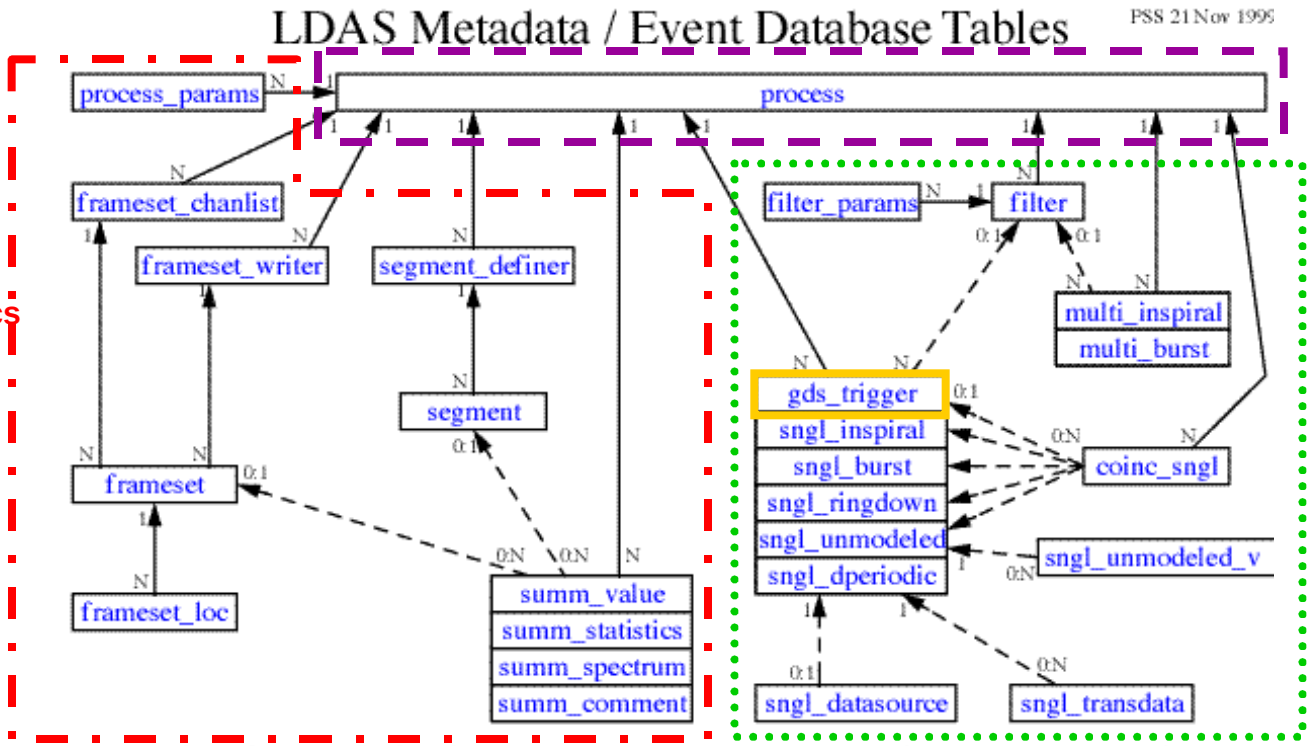
Raw data characteristics/location/statistics

Instrumental triggers (vetoes)

Astrophysical search triggers

Single interferometer

Multiple interfertometers



Arrows indicate "foreign key" referential integrity constraints. Values near the ends of the arrows (1, N, etc.) indicate the possible multiplicities. Dashed lines indicate optional relationships. Stacked tables (grouped by thick lines) have common relationships with other tables, except for relationship arrows connecting along the right edge. Examples: 1) Each segment is related to one segment_definer; 2) Each segment_definer is (generally) related to many segments; 3) A frameset is related to one frameset_chanlist entry and to one frameset_writer; 3) A summ_value (or summ_statistics, etc.) entry may or may not be related to a segment and/or a frameset; 4)A single-interferometer event (gds_trigger, sngl_inspiral, etc.) entry may be related to up to one sngl_datasource and/or any number of sngl_transdata entries.

LIGO Data Analysis System

System Capacities

LDAS Characteristics by Site			
LDAS Site	**Computational Power[a] GFLOPS**	**Storage Capacity TB**	**Networks[b] MB/s**
LIGO Hanford Observatory [LHO] On-Site LDAS	61 (114 nodes)	14 (RAID 5) 1.5 (Tape) 7 (IDEs in linux cluster)	100/1000 Mbps (LAN) 100 Mbps (linux cluster) T1 (LIGO WAN)
LIGO Livingston Observatory [LLO] On-Site LDAS	34 (66 nodes)	7 (RAID 5) 1.5 (Tape) 4 (IDEs in linux cluster)	100/1000 Mbps (LAN) 100 Mbps (linux cluster) T1 (LIGO WAN)
LIGO Laboratory at Caltech [CIT] Off-Site LDAS	110 (216 nodes)	6.5 (RAID) >360 (Tape)[c] 14 (IDEs in linux cluster) 10+ (IDE/SCSI RAID)	OC3 (LAN) 10/1000 Mbps (LAN) 100 Mbps (linux cluster) OC48 (University WAN)
LIGO Laboratory At MIT Off-Site LDAS	13 (28 nodes)	0.5 (RAID) 1.5 (Tape) 2 (IDEs in linux cluster)	100 Mbps (LAN, linux cluster) OC12 (University WAN)

a. Projections based on present technology extrapolated forward to the acquisition periods.

b. ATM networks represent existing infrastructure; no additional ATM infrastructure is planned as part of this procurement other than for switches providing interfaces to existing LANs.

c. This is based on acquisition of a separate robotic silo for LIGO at CACR and using a figure of 60GB/cassette media density; archive volume grows to 600TB with 100 GB cassettes.

LDAS Hardware Procurement

- Procurement divided into 3 Phases
  - » Phase I - 1Q2001: Hardware needed to support MDCs, engineering runs
  - » Phase II - 3Q2001: Hardware will be installed immediately after the “E6” engineering run to establish scientific upper limits with LIGO data
  - » Phase III - 4Q2001: finalize hardware configuration approximately 6 months before the beginning of the LIGO I Science Run.

LDAS Hardware Procurement

- Prototype hardware will be reused for certain components

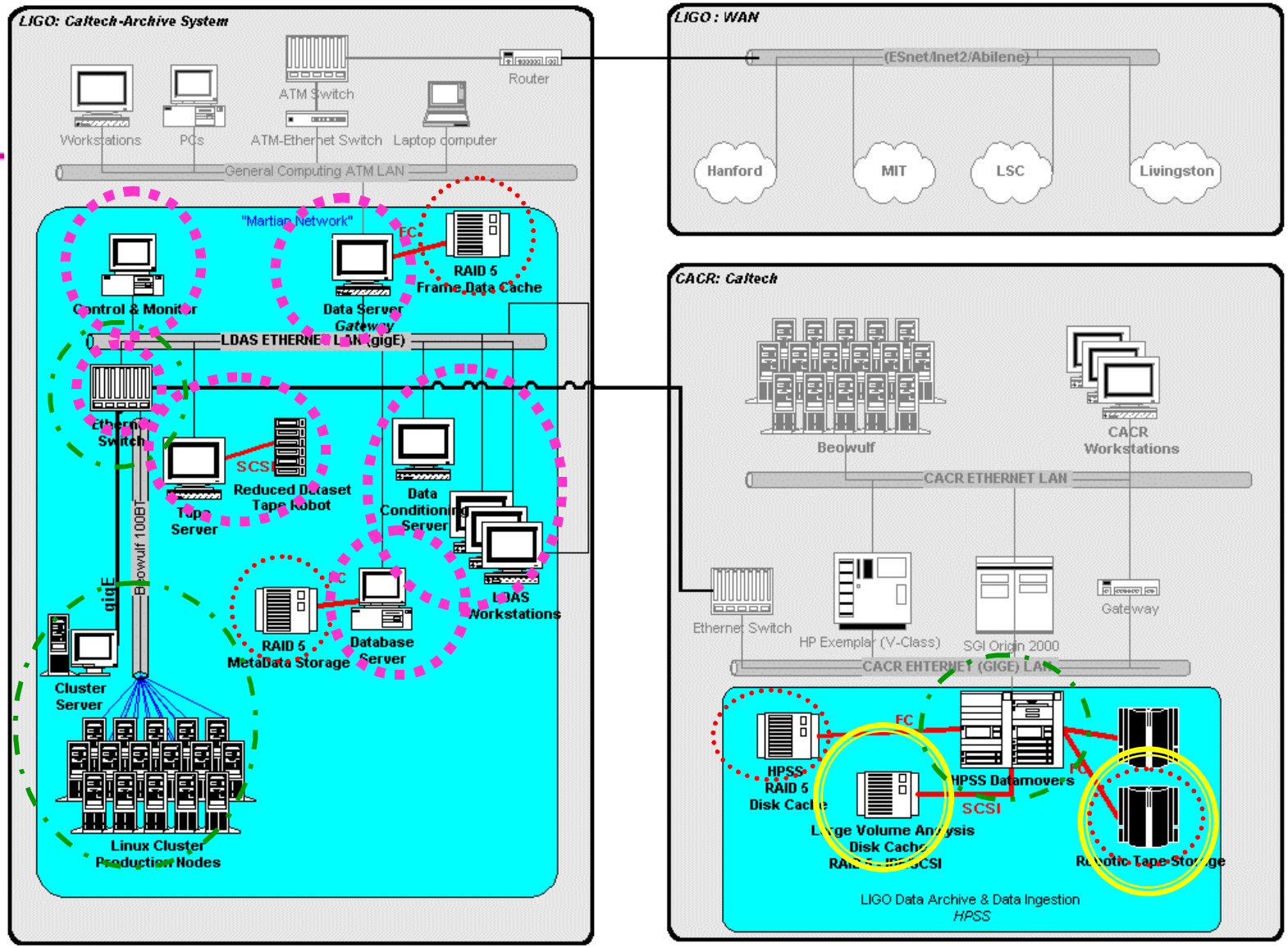
- Phase I: issued 9 Mar 2001 (slide 21)
  - » 28 TB of RAID disk systems for Observatories, Caltech, MIT
    - – Sites will have 1 month look-back capacity on spinning media @3 MB/s per interferometer (Hanford: 14 TB, Livingston: 7 TB)
    - – Caltech HPSS disk cache (6 .5 TB)
    - – MIT disk cache (0.5 TB)
  - » 6000 slot (500+TB capacity) robotic silo for HPSS at Caltech
  - » PCs, servers for E6 run (16 ea for Hanford, Livingston)

LDAS Hardware Procurement

- Phase II: -- after final beowulf, HPSS benchmark tests *(slide 21*
  - » Data movers for HPSS (support 5 continuous streams of data at full bandwidth)
  - » PCs for main clusters at Caltech, MIT, LLO, LHO (total of ~400 units)
- Phase III: complete HPSS for Science run *(slide 21)*
  - » High density tape drives for HPSS (baseline: STK model 9940)
  - » Tapes for Science Run
  - » Start build-up of large (inexpensive) disk farm in front of HPSS for data analysis
    - –Continue to grow farm throughout Science run
    - –Keep up with data growth
    - –Target keeping all commonly used, needed data on disk
    - –Use HPSS for backup, large data dumps to disk caches *(slide 23)*

Phased Procurement at Caltech

LDAS System Block Diagram for Caltech-Archive/Production System



Exist from Prototyping

Phase I

Phase II

Phase III

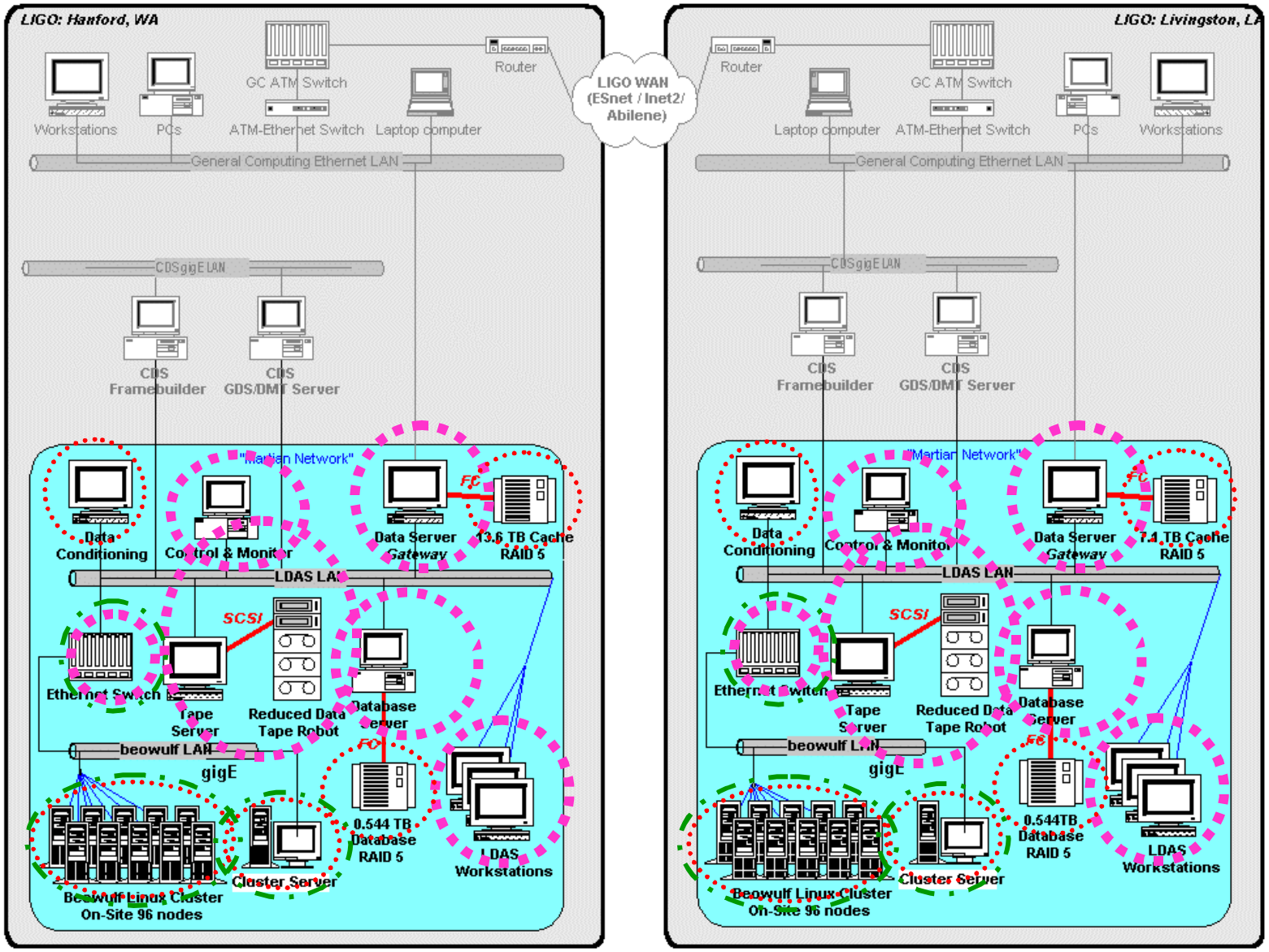
(slide 20)

(slide 22)

(slide 23)

Phased Procurement at Sites

LDAS System Block Diagram for Sites



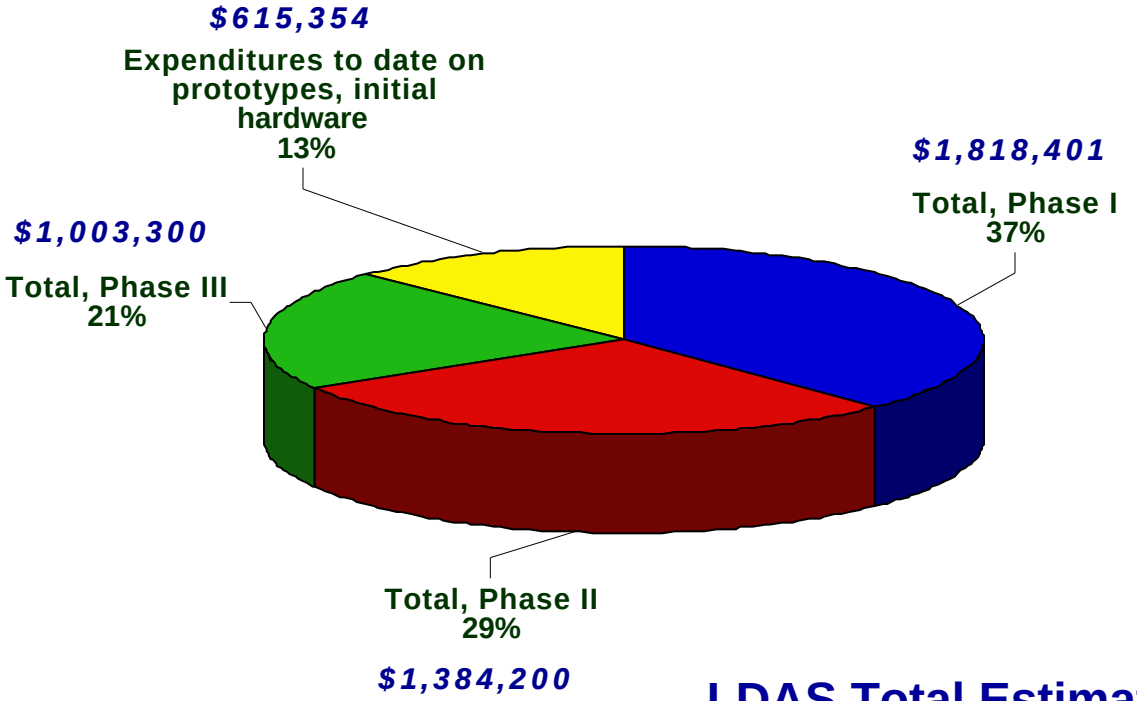
Exist from Prototyping

Phase I

Phase II

LDAS Procurements by Phase

Breakdown of LDAS Costs

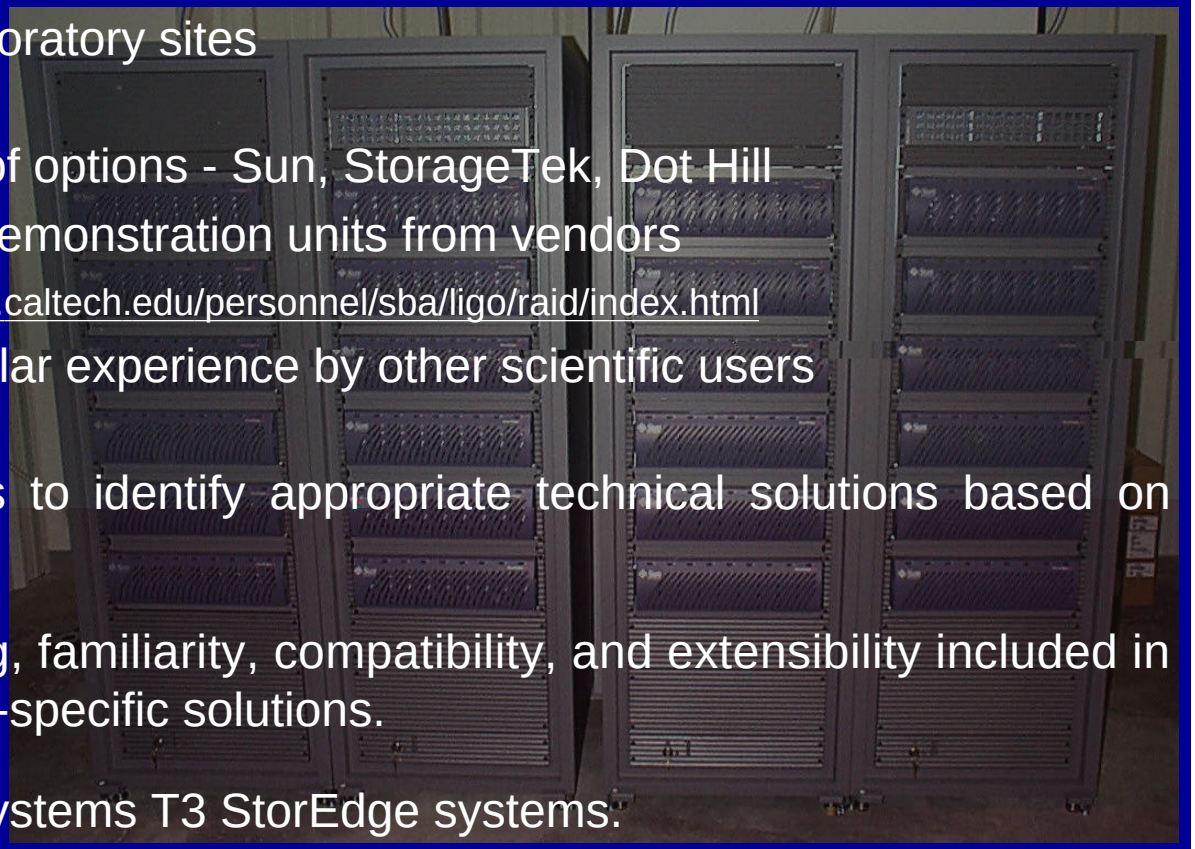


LDAS Total Estimated Cost: $4.8M

Annual Operations HW budget: 25% of plant cost

Selection of Disk Arrays
(Phase I)

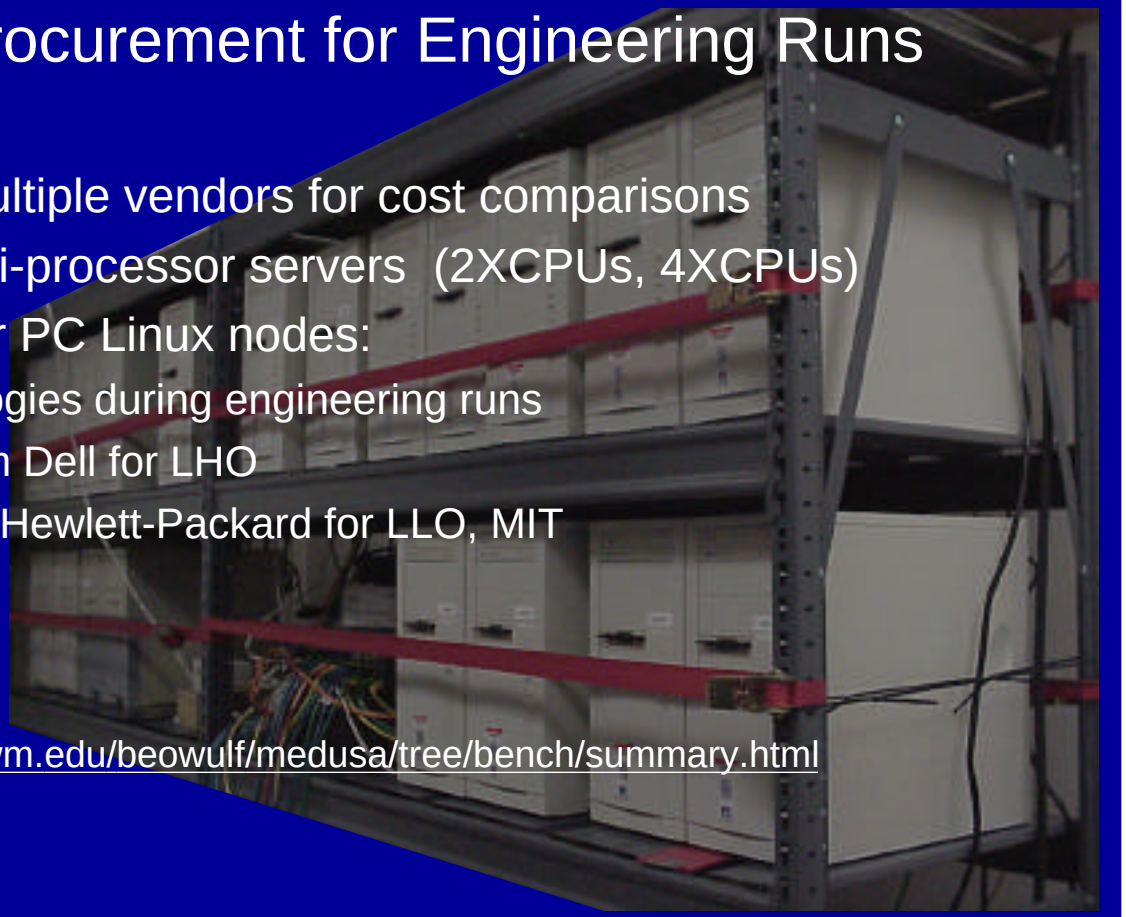
- Turnkey fibre channel RAID systems for primary storage at sites, archive
  - » Identical for all laboratory sites
- Evaluated a number of options - Sun, StorageTek, Dot Hill
  - » Benchmarks on demonstration units from vendors
    - » http://www.srl.caltech.edu/personnel/sba/ligo/raid/index.html
  - » Knowledge of similar experience by other scientific users
- Worked with vendors to identify appropriate technical solutions based on each product line.
- Other factors: training, familiarity, compatibility, and extensibility included in identifying the vendor-specific solutions.
- Selected Sun MicroSystems T3 StorEdge systems.



Selection of PC Clusters, Linux Servers

(Phase I)

- Phase I -- small lot procurement for Engineering Runs
  - » Obtained quotes from multiple vendors for cost comparisons
  - » Dell selected for the multi-processor servers (2XCPUs, 4XCPUs)
  - » Two vendors selected for PC Linux nodes:
    - –Use different PC technologies during engineering runs
    - –16 Pentium IV nodes from Dell for LHO
    - –16 Athlon K7 nodes from Hewlett-Packard for LLO, MIT
  - » Benchmarks
    - –Ref. Kent Blackburn's talk
    - –http://www.lsc-group.phys.uwm.edu/beowulf/medusa/tree/bench/summary.html



Selection of PC Clusters, Linux Servers

(Phase II)

- Phase II -- large procurement for Science Run
- Repeat cost, performance comparisons
- Install homogeneous clusters at each site
- O[400] nodes will be purchased

Selection of Robotic Silo
(Phase I)

- STK's 9310 Powderhorn, US-made system; rated at 350 cassette exchanges per hour.
- After-market OEM system from multiple vendors
  - » Dot Hill,
  - » Sun Microsystems
  - » StorageTek (the actual manufacturer)
- Refurbished units with a full service warranty offer significant savings
  - » RHIC/RFC experience
- Selected Dot Hill
  - » Sun Microsystems does not resell refurbished units

Selection of Networking Hardware
(Prototyping, Phase II)

- Prototyping:
  - » Move from ATM to GigE (1000 Mbps)
  - » Evaluated multiple manufacturers
    - –Extreme Networks, Foundry Systems, CISCO
  - » Confirmed interoperability of hardware with existing systems (e.g., CISCO)
  - » Extensible, non-blocking performance at lowest $/port
  - » Selected Foundry Systems for prototype (small) switches
- Phase II: Will re-evaluate market options when large beowulf switches need to be procured

Selection of Data Archival System

(Phase II)

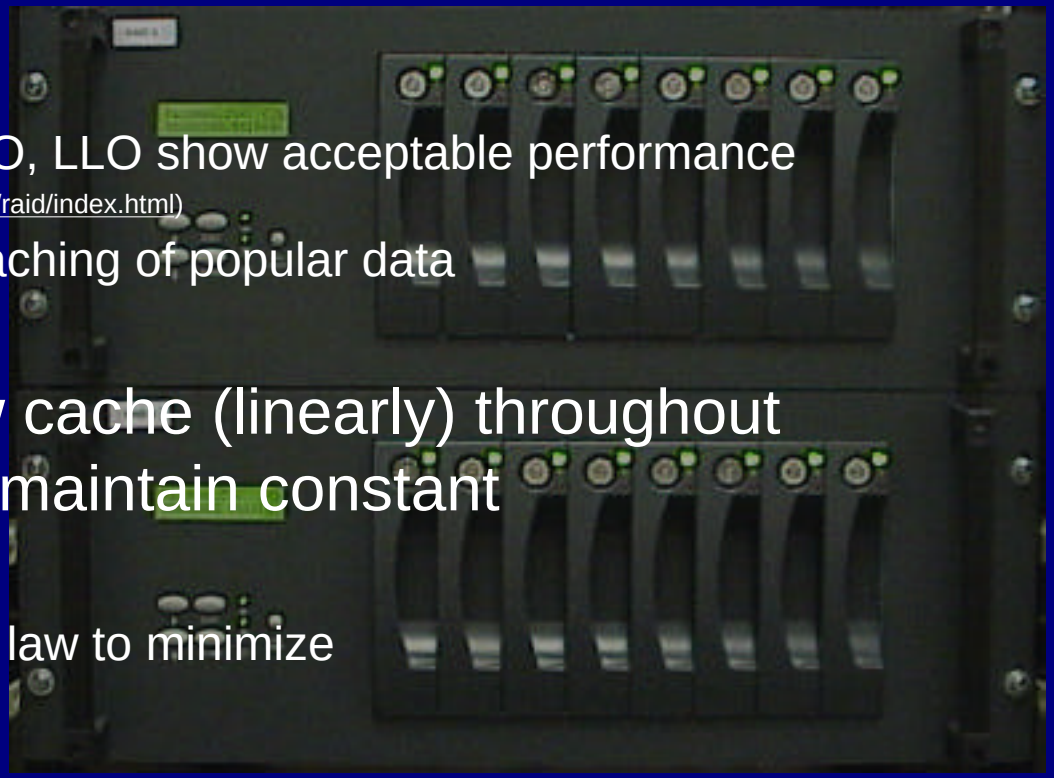
- LIGO has experience, familiarity with Caltech HPSS: 1998 - 2001
- Dec/2000: Visited RCF (RHIC/BNL) in December 2000, met with HPSS administrators, compared notes, exchanged plans
- Apr/2001: Evaluated IBM's port of HPSS to Sun Microsystem/Solaris servers as data movers for HPSS *(http://www.cacr.caltech.edu/~dkozak/beaverton/)*
  - » Battery of tests designed to identify/characterize known HPSS performance bottleneck -- metadata transactions
  - » Tested with 6 major types of Sun server systems, including Serengeti line
  - » Joint evaluation with IBM's Austin engineering
  - » Cost & performance with Sun are superior to IBM
- Extensible solution will go beyond the LIGO I Science Run
- Working with Sun to define configuration, negotiate costs

Selection of Data Archival System
(Phase III)

- Robot system tape drives
  - » Wait for STK release of high density 9940B line
  - » 100 GB/cassette
- Visit to RHIC/RCF confirmed LIGO decision to move to higher density drives
- Caltech/CACR will buy STK 9840 in near term for use with LIGO Silo
  - » Data migration off Redwood drives
  - » Test STK performance

Selection of Data Disk Cache
(Phase III)

- Large, inexpensive (e.g., IDE/SCSI) disk farm in front of HPSS
  - » Benchmarks on units at LHO, LLO show acceptable performance
    - – (http://www.srl.caltech.edu/personnel/sba/ligo/raid/index.html)
  - » Balance tape access vs. caching of popular data
- Buy 10TB initially, grow cache (linearly) throughout LIGO I Science Run to maintain constant HPSS:Farm data ratio
  - » Take advantage of Moore's law to minimize cost:data volume



Plan to Reach Science Run

Detector & Data Analysis

- Jan to mid-March
  - ***LHO 2k, continued work on improving robustness of lock, some work on sensitivity***
  - ***LLO 4k, Lock single arm, recombined Michelson with Fabry-Perot (F-P) arms, Power Recycled Michelson (PRM)***
  - *LHO 4k, installation*
  - » ***SW: Prepare LDAS release for E3***
  - » ***HW: Procure Phase I, final RAID configurations, HPSS tape silo, small beowulf clusters for E6***
- March 9-12
  - ***E3 (engineering run): coincidence run between LHO PEM and single F-P arm at LLO***
  - » ***SW+HW: Archive E3 data***

Plan to Reach Science Run

Detector & Data Analysis

- mid-March to mid-May
  - ***LHO 4k, complete installation, lock mode cleaner***
  - ***LHO 2k, repair, suspension sensor replacement, resurrect PRM studies***
  - ***LLO 4k, lock full interferometer, sensitivity/robustness***
  - » ***MDC: metaDataAPI (Caltech/LHO)***
  - » ***SW: Prepare LDAS release for E4***
  - » ***HW: Install Phase I hardware***
  - » ***HW: Benchmark HPSS on Sun hardware at Sun testbed facilities, Beaverton, OR***
- May
  - ***E4 run: LLO 4 km, operating in recombined mode (recycling?) + LHO PEM***
  - » ***MDC: MPI inspiral search (first of 4 MDCs tied to upper limits run)***
  - » ***SW+HW: Archive E4 data***

Plan to Reach Science Run

Detector & Data Analysis

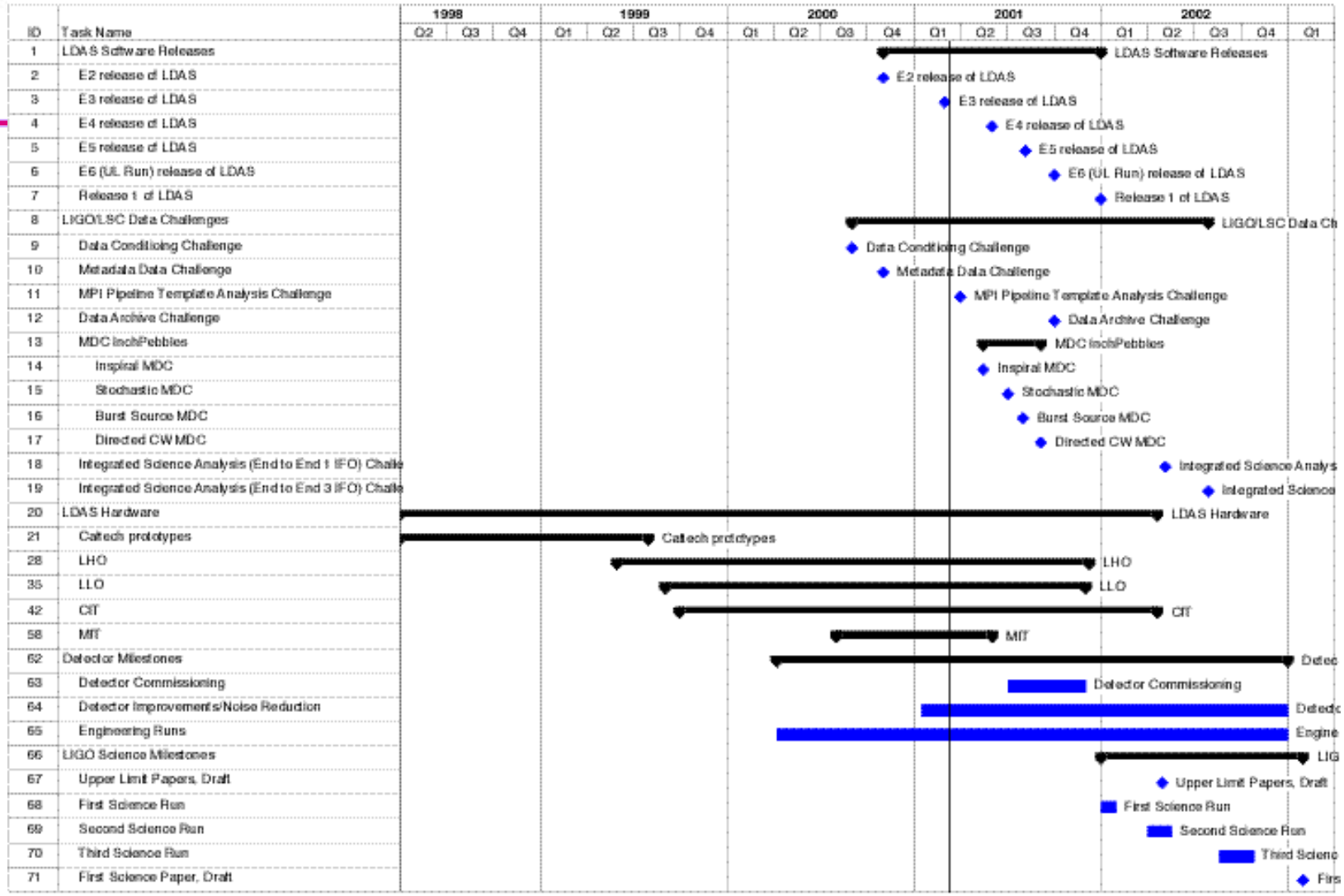
- May - June
  - *LHO 2k, bring full interferometer back on-line, sensitivity studies*
  - *LLO 4k, improve full interferometer lock, sensitivity studies*
  - *LHO 4k, PRM locking (no arms yet)*
  - » ***SW: Prepare LDAS release for E5***
  - » ***HW: Specify HPSS HW configuration for Phase II of procurement***
- late June - early July
  - *E5: LHO 2k in full recycled configuration, LLO 4k in full recycled configuration(?), LHO 4k in PRM mode*
  - » ***MDC: Stochastic background search***
  - » ***SW+HW: Archive E5 data***
- July - Sept
  - *LLO 4 k suspension sensor replacement, bring back on-line*
  - *LHO 2km sensitivity studies, 4k lock full interferometer*
  - » ***SW; Prepare LDAS release for E6***
  - » ***MDC : Burst search, CW (pulsar) MDC***
  - » ***HW: Benchmark PCs for large beowulf procurement, Procure Phase II***

Plan to Reach Science Run

Detector & Data Analysis

- late Sept
  - ***E6: triple coincidence run with all 3 interferometers in final optical configuration ("upper limit run")***
  - » ***SW+HW: Archive E6 data, on-site upper limit searches***
- Oct – early 2002
  - *Improve sensitivity and reliability*
  - *Alternate diagnostic testing with engineering runs*
  - » ***MDC : Data Archive***
  - » ***HW: Specify HPSS drives, tapes, IDE/SCSI disk cache for data at Caltech; Procure Phase III***
- Jan - July 2002
  - » ***SW+HW: Prepare Release 1 of LDAS for Science run***
  - » ***SW+HW : Integrated single interferometer running on-site***
  - » ***SW+HW : multiple interferometer running off-site***

LIGO/LSC Software Deployment Schedule



Beyond LIGO I

- LSC Data Analysis White Paper revision
  - » Develop strategic plan for placement, allocation of computing resources
  - » Methods for collaboration members to participate in new programs to augment resources within the LSC
- LIGO Laboratory:
  - » Worked with CACR, SDSC on Distributed Terascale Facility proposal
    - –Model for future LIGO Tier 1 Center upgrade
    - –Teraflops-scale computing for intense analysis
  - » MRE proposal to NSF for Advanced LIGO in 3Q2001;
    - –Identify & cost LDAS upgrades (LIGO Laboratory) for Advanced LIGO interferometers

Beyond LIGO I

- LSC Activities
  - » IT Research:
    - –ITR2000 -- GriPhyN -- SW/CS component R&D
      - Caltech, UWM, UTB members of collaboration
      - Research on, prototyping of virtual data catalogs, data mirrors for LIGO data
    - –International Virtual Data Grid Laboratory (iVDGL)
      - Develop 2 prototype LSC Tier2 centers for grid computing applications with LIGO data
      - European collaborators from Virgo, GEO in UK/EU Grid projects
        - » Use grid to exchange data
        - » Networked detector joint data analyses

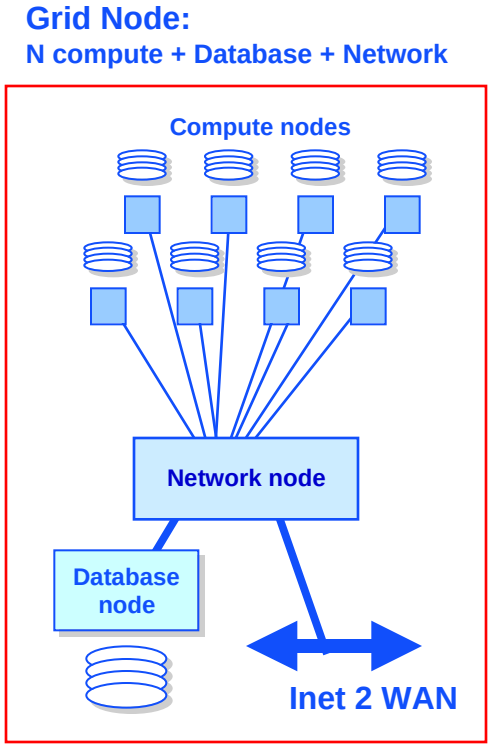
GriPhyN

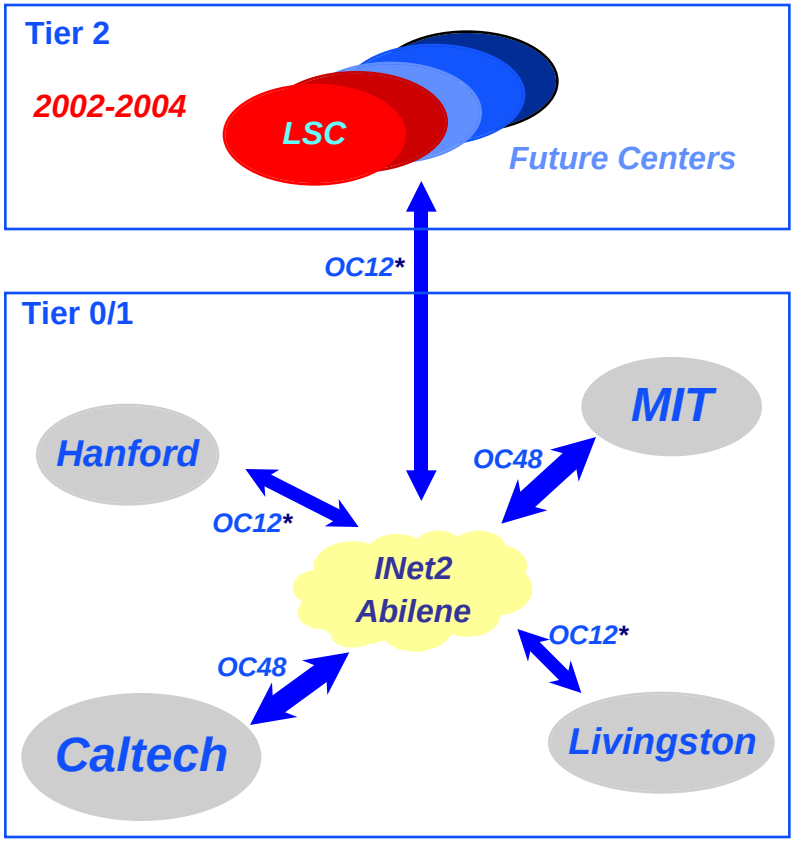
Grid Physics Network

- Caltech
  - » Virtual data models R&D with USC/Information Science Institute (Kesselman/ISI)
    - –Grid-enabling LDAS APIs for accessing LIGO data from the grid environment
      - Staging large computational tasks
      - Tracking data for delivery over a distributed grid environment
- UWM
  - » Porting of grid tools to the UWM beowulf
  - » Use of beowulf system to mirror large datasets (“datawulf”)
- UTB
  - » Educational Outreach coordination for GriPhyN

GriPhyN Tier 2 Hardware

Scenario for LSC Deployment





* Bandwidths reflect projected capacities in ~ 4+ yrs

Summary

- LDAS Procurement has begun
  - » Completion by time of Science Run
- Coordination with detector commissioning is driving deployment
- Future directions of growth for LIGO, LSC being developed in collaboration with other major NSF IT initiatives