

---

---

INTEROFFICE MEMORANDUM

---

---

**TO:** LSC COMP COMM  
**FROM:** LEE SAMUEL FINN  
**SUBJECT:** FNAL MEETING MINUTES  
**DATE:** 3/10/02

---

In Attendance

- Anderson, Warren
  - Blackburn
  - Brady
  - Finn
  - Katsavounidis
  - Lazzarini
  - Nash
  - Wiseman
1. Making Grid Computing Work for LIGO (Nash; See [http://www.ligo.caltech.edu/~lazz/LSC\\_Computing/NashSlides20020302.pdf](http://www.ligo.caltech.edu/~lazz/LSC_Computing/NashSlides20020302.pdf))
    - a. Everything flows from this
      - i. Need to get data into resource list
    - b. Three separable general Issues
      - i. Management of Coordinated Resources
        - (1) Need to recognize that resources are also not uniformly dispersed: lack of balance
      - ii. Support: Demand Side Resources: hardware, software, data network
        - (1) Hardware usage
        - (2) Software categories
      - iii. Support: Supply side resources
    - c. Our task: manage resources. Look at where we are ... define where we want to get
  2. Overview of tasks and sites (Lazz; see [http://www.ligo.caltech.edu/~lazz/LSC\\_Computing/LSCComputerUsageModel.pdf](http://www.ligo.caltech.edu/~lazz/LSC_Computing/LSCComputerUsageModel.pdf))
    - a. Hidden support costs in many of the activities. In particular, there are both schedulable support activities, and individually driven support activities.
  3. Discussion
    - a. Exposing what data is where to ... who? User? Job scheduling policy / agent?
  4. iVDGL Communications:
    - a. We have representatives on all the iVDGL committees. Communication to iVDGL through these committees should go through the representative / liaison but only after being cleared by committee.
  5. Resource Allocation for Users
    - a. This is after the analysis proposal to the LSC is approved, which (among other things) *defines* needed resources. The purpose here is to determine how we get the resources to the user(s).

- b. Note that there are legitimate applications that do not need to have gone through a proposal process: e.g., algorithm development requires cycles, but not data.
- c. Possible concern: we are writing this policy for “LSC”. Do we need to be concerned that some Lab activities may not be “LSC” activities? Leave as footnote to be dealt with if it becomes a problem.
- d. It is understood that authority to manage resource allocation on a day-to-day basis will almost always be delegated
- 6. PM Agenda. (Note: in the end we did not cover all of these topics)
  - a. Resource allocation for users
  - b. Prioritized list of tasks & actions
  - c. Immediate issues for operating centers
  - d. Configuration of future sites
  - e. Access control for E7, S1, S2, etc.
  - f. Requirements for site and grid data catalog/ database
- 7. Lunch break
- 8. Resource allocation for users.
  - a. This is after the proposal process, which (among other things) identifies the anticipated call on resources. Here we are concerned with getting the resources to the users.
  - b. It is first necessary to identify the resources at the collaborations command.
    - i. These are identified in the MoUs established between the centers and the Lab Directorate.
      - (1) MoUs need to identify hardware, software, storage, data, configuration, FTEs for system and user support, availability, and (related to availability) some measure of the resources dedicated to collaboration work (e.g., like allocation units used in allocated resources at national supercomputing centers)
    - ii. Questions raised but not resolved:
      - (1) Frequency of MoU renegotiation
      - (2) Can resources committed in one MoU cycle be withdrawn? To what extent are resources, acquired in part of whole by virtue of association with LIGO/LSC, “owned” by the host institution?
      - (3) If resources can be withdrawn, how can there be any long-range planning of resources?
  - c. LSC Comp Coord Comm will establish a regular review of LSC user requests to access LSC resources
    - i. Users? Individuals or groups.
    - ii. Requests? A short statement explaining what needs to be done (including context), the required resources, etc.
    - iii. Requests originating from analysis groups should be communicated en masse by the group leader
    - iv. Requests not originating from analysis groups, which have passed through the proposal process, will be considered as appropriate.
      - (1) There are legitimate requests that can arise outside the proposal process:: for example, development activities that involve testing new analysis methods on simulated data, detector characterization activities, source simulation activities (if that becomes a recognized collaboration function)
    - v. Review will be on a monthly basis
      - (1) Urgent requests, that cannot wait until the next monthly review, can be forwarded to Comp Comm chair

- (2) Comp Comm will maintain a web site with current proposals and their disposition, including allocation, etc.
  - (3) Day to day resource allocation at a center may be delegated to a committee member. This is intended to avoid idle resources in cases where an allocation award already made cannot be used because, e.g., of software bugs
9. Requirements for data discovery and movement to support analysis activities at/ across sites
- a. It is necessary to keep track of what data is where, both for the purpose of synchronization of databases and data products and for the efficient allocation of jobs (e.g., when resources allow better to send jobs to data than data to jobs).
  - b. The information that must be available about the data at each site includes
    - i. Epochs (start time, duration)
    - ii. Data product level (i.e., level 1? 2? 3?) and instrument, other description
  - c. It must be possible to make the following queries
    - i. Has data X (already) been generated?
    - ii. What data is at Y [for use with/by Z]?
    - iii. Where is data X [for use with/by Z]?
    - iv. Here X is a description of the data (e.g., GPS time, duration, channels, rds designation, etc.), Y is a site (e.g., CIT, LHO, LLO, MIT, PSU, UTB, UWM), and Z is a “process” (e.g., LDAS, DMT, other?)
    - v. Responses should be available in both human and machine readable forms.
  - d. Data movement must be “automatic” through tools layered on existing analysis software (e.g., LDAS, DMT) and not requiring modification to those analysis environments. Here “automatic” means that data movement is a single command, which takes care of all the actions associated with the data movement
    - i. Middleware “broker” that uses grid tools to move data from/to, puts it in to place for use by Z, and makes appropriate changes to “databases”, etc.
  - e. The data discovery and movement tools, and all associated databases, etc., must be uniform across all LIGO/LSC centers.
  - f. Solution will be made in steps
    - i. V.1: level 1, 2, 3 data sets only.
    - ii. V.2: other data sets, tbd. This opens a can of worms re: how a data set is made self-describing, etc.
  - g. Access authority (LIGO-Apps working group in GriPhyN is working this problem):
    - i. Policy
    - ii. Implementation
  - h. These same issues need to be dealt with for metadata (i.e., relational databases associated with the data). The general requirements are the same (queries, uniform and automatic solution) but the details may be different owing to the differences in the data type.
  - i. Action for Patrick, Alan: at next LSC Comp Comm telecon get Scott Koranda to report on progress in working with ISI on catalog and middleware.
    - i. Include what fraction time Scott spending
    - ii. What fraction of ISI time is being spent on problem.
10. Resource Crises
- a. Personnel
    - i. Configuration control
    - ii. Implementation/porting/etc.

- iii. User support
  - b. Bandwidth
    - i. HPSS/tape
    - ii. WAN
11. HW & SW configuration management issues: dealing with heterogeneity
- a. All sites will be unavoidably different in certain ways
    - i. Purchases at different times; replacements and upgrades at different times; network bandwidths different; distinct missions lead to different resources and configurations.
  - b. Periodic software upgrades (O/S, LDAS, DMT, etc.) are driven by release of bug fixes, new or significantly improved functionality, and are not avoidable.
  - c. Nevertheless, there will be times when – by design! – we will be running different software on different clusters
    - i. Updates during engineering and/or science runs may only be on restricted set of resources
    - ii. Updates outside of engineering/science runs may “roll” on a schedule that depends on tasks going on at different sites
  - d. HW will be heterogeneous both within and across centers.
    - i. Different speeds, different processors, different memory, different nodes (mp vs. up) different network bandwidths, etc.
    - ii. Unavoidable: hardware purchases at different times, spread over time for both programmatic and budgetary reasons, replacement and upgrade hardware, network infrastructure not under center control (i.e., OC-3, OC-12 connectivity is generally a campus issue and beyond the present center budgets).
  - e. Dealing with heterogeneity:
    - i. O/S upgrades
      - (1) Decide when an upgrade is mandated in committee (including character of upgrade: e.g., rebuild vs. upgrade vs. rpm list vs. etc.)
      - (2) Schedule and react: but reaction time such that not out-of-sync for more than one allocation committee meeting
      - (3) Implies DC level of resources to support activity
        - (a) Each site must test and verify on test-bed that is appropriate to its hardware and software configuration(s) before propagating to cluster; etc.
    - ii. Application environments
      - (1) /ldcg, LAL, LDAS, DMT, etc.
    - iii. Reference platforms
      - (1) We never made it to this point
  - f. Configuration issues. Each site will have several partitions of its cluster: e.g.,
    - i. Local partition(s) for testing deployments (new/updated s/w, os, etc.)
    - ii. Local LDAS partition for development (i.e., LDAS development snapshot)
    - iii. Local DMT partition for development (i.e., DMT development snapshot)
    - iv. LDAS production partition
    - v. DMT production partition
    - vi. Source simulation activity partition (if this becomes a legitimate LSC activity)
    - vii. Other