DOE REVIEW REPORT

of the

RHIC COMPUTING FACILITY
and
RHIC OFF-LINE COMPUTING PLANS

held on
JULY 29-31, 1997

Daniel R. Lehman
Chairman
DOE Review Committee

RHIC Off Line Computing Review
July 29 - 31, 1997

EXECUTIVE SUMMARY

A Department of Energy (DOE) Office of Energy Research (ER) review of the Relativistic Heavy Ion Collider (RHIC) Off-Line Computing was held at Brookhaven National Laboratory (BNL) July 29-31, 1997. The review was conducted by a committee chaired by Mr. Daniel Lehman, Division of Construction Management Support. The committee, which was composed of scientific computing and management experts, was charged to evaluate the status and plans for meeting the off-line computing requirements of the RHIC experimental program.

The computing and data handling capacities required for RHIC are large when compared to previous detector systems in either High Energy or Nuclear Physics. The approach for meeting the RHIC off-line requirements includes a dedicated facility located at BNL, the RHIC Computing Facility (RCF), and facilities located at remote (i.e., non-BNL) locations. The RCF will primarily handle raw data from the experiments and contain the main data store for RHIC. However, substantial amounts of data processing and software development are planned to take place at remote sites.

The committee found that the overall strategy for the RCF to be fundamentally sound. The system architecture is reasonable, and the task partitioning is well matched to the physics needs. The planned schedule, which would achieve nominal capacity in 2001, is consistent with RHIC startup plans. The capital equipment costs for the main RCF components are well understood. Contingency is not budgeted separately as there are several ways to scale down capabilities if costs are greater than expected. The Committee did find that the experimental computing requirements need more study, especially those requirements that have profound effects on the RCF system architecture and on the size and scope of the remote facilities. The committee emphasized that unrecoverable architectural choices should not be made on the basis of unjustified and indefensible requirements. The committee also noted that it is important that the RCF and experimental collaborations undertake continuous exchange of information on hardware and software, and that RCF make hardware and software choices broadly aligned with those made at other labs such as Fermilab, SLAC, Jefferson Lab, CERN, and DESY. They also recommended that a "mock data challenge" be scheduled approximately one year before the arrival of real data to establish confidence that the RCF would perform as expected. The most challenging issue facing the RHIC Computing Head is the acquisition of capable personnel. The staffing plan for the RCF is lean and depends on RHIC direct hire staff, contributed staff from the BNL Computing & Communications Division (CCD), and contributed personnel from the experiments. Actual staffing is below desired levels and schedule success will depend on a dramatic increase in personnel, particularly in the area of software development and maintenance.

Collaborators for the RHIC experiments presented plans for remote facilities at MIT (PHOBOS), Japan-RIKEN (PHENIX), and LBNL/NERSC (STAR). Progress at these remote facilities is important to the general success of RHIC off-line computing. The MIT proposal is well underway and a proposal has been formally submitted to the Japanese government for the PHENIX/Japan Computing Center. The STAR proposal at LBNL/NERSC raised concerns about the relative advantage to the overall RHIC computing effort from leveraging non-RHIC managed resources against possible diminished returns due to the complexities involved in sharing resources and compatibility issues. The committee felt that the RCF should have priority for new RHIC computing capital equipment funds, since the RCF is at the core of the RHIC off-line computing effort.

The Head of the RCF has the responsibility for constructing the RCF and providing guidance and direction on the entire RHIC off-line computing effort. It is important that the technical strength of this individual be augmented by additional assistance in project management in order to ensure that near term milestones are met on schedule. In addition, additional mechanisms for strengthening communication with the experiments and with the non-RHIC NP and HEP computing communities need to be established.

The committee had numerous specific technical recommendations and recommendations directed at the need to develop a detailed implementation plan for the RHIC off-line computing effort. These recommendations are provided in the body of the report. It was agreed that RHIC off-line computing should be reviewed again in six months to review progress during this critical period of development.

1.0 Introduction

At the request of Dr. David Hendrie, Director of the Division of Nuclear Physics, an Office of Energy Research Review of Relativistic Heavy Ion Collider (RHIC) Off-line Computing was held on July 29-31, 1997 at Brookhaven National Laboratory (BNL). The review team was chaired by Mr. Daniel Lehman, Construction Management Support Division. The review committee consisted of scientific, technical, and management experts from the Department of Energy community. The names and affiliations of the team members are listed in Appendix B. The review agenda is contained in Appendix C. The purpose of the review was to evaluate: (1) the technical design of the BNL RHIC Computing Facility (RCF); (2) the advisability, adequacy and status of plans to use other computing centers to perform RHIC required computing tasks (primarily simulation and modeling); (3) the status of and plans for adopting and/or developing data handling and resource management software for the RCF; (4) capital and operating cost estimates for the BNL RCF and for non-BNL RHIC-related computing resources supported by the Nuclear Physics program; (5) the RHIC-wide strategy for meeting computing requirements while minimizing cost; (6) the schedule; and (7) the management structure.

RHIC will be able to create matter at such extremely high temperatures and densities that scientists hope to observe phenomena that have not occurred in the natural universe since the "Big Bang." This will make possible studies of the fundamental properties of matter in a state in which the primordial quarks and gluons are not confined within the nucleons that make up the nucleus of the atom. This state of matter offers an exciting new area of scientific study for physicists. In RHIC, two beams of heavy ions will speed in opposite directions around a pair of rings in a tunnel approximately 2.3 miles in circumference. At the six different points around the accelerator where the beams collide, experiments will be carried out with an interaction energy of 100 billion electron-volts for each of the protons and neutrons contained in the colliding nuclei (100 GeV/u). At start-up, four of these interaction regions will have detectors to record information from the collisions. The computing and data handling capacities required for these detectors are large when compared to previous detector systems in either High Energy or Nuclear Physics. Thus it is imperative that unusual care be taken in the design and planning for RHIC-wide computing to make sure that the complex and massive data stream from the RHIC detectors can be properly handled.

As requested by the committee chairman, the BNL Project Head, Dr. Satoshi Ozaki, briefed the committee prior to the start of presentations by BNL staff, summarizing the status of the RHIC construction project. The RHIC project has been making steady progress toward completion in the third quarter of FY 1999. As of June 30, 1997 construction of the collider accelerator is at the 83% completion level, and the detector construction at the 60% completion level. The two major detectors, STAR and PHENIX, are closest to completion with their main magnets almost ready for power and water. Construction work on most detector components has started and detector roll-in is planned for mid-1999. Dr. Ozaki expects beam collisions during CY 1999 consistent with BNL’s goal of starting the full scale physics program as early as possible.

Following Dr. Ozaki’s overview presentation, the review committee received technical, cost and schedule presentations of the plans for RHIC off-line computing, and held discussions with its technical managers. The review committee shared its overall assessment of RHIC Off-line Computing with BNL at a closeout briefing on July 31st. This report contains the committee's findings, comments, and recommendations.

2.0 Computing Requirements and Overall Strategy

Findings

The RHIC Computing Facility Team presented the review committee with computing requirements prepared by the RHIC experiments, with a division of the computing tasks into several distinct functional areas, with an overall architecture to meet experimental needs, with a partition between tasks to be done at the RHIC Computing Facility and tasks that could be done elsewhere, with a schedule for bringing on computing systems with capacity growing each year, and with resource requirements to deliver these systems.

The committee feels that the overall strategy is fundamentally sound. The overall architecture makes sense, and the task partitioning is well matched to the physics needs. The planned schedule, which achieves nominal capacity only in 2001, appears to be adequate given the current understanding of RHIC startup plans. Tasks that must be performed locally at the RCF are correctly identified.

The justification for the experimental computing requirements was not adequate. In some cases (such as for event reconstruction) these requirements are reasonable estimates that still have considerable uncertainty, while in other cases (such as for analysis CPU capacity) the requirements have little justification and are insufficient to serve as a basis for system architecture decisions. The experimental estimates for storage requirements did not include any justification for what fraction of this storage needs to be available on disk or robotically as opposed to on the shelf.

Comments

The experimental requirements need more study. Those requirements that affect only the scale of the required computing systems can be left uncertain (as long as the level of uncertainty is understood), since the architecture is fundamentally scalable and can thus be adjusted (provided funding is available) to meet these requirements as they become more precise. However, other requirements have profound effects on the overall system architecture. Unrecoverable architectural choices should not be made on the basis of unjustified and indefensible requirements.

A large fraction of computing tasks is planned for outside of the RCF. The tasks so identified are clearly the ones most suited for such distribution, but the extent of this partitioning has been determined by financial rather than by scientific grounds. It would be beneficial if some of these tasks, particularly analysis of simulated events, could be brought into the RCF. In particular tasks that place great demands on the network should only be performed locally.

The proposed manpower appears barely adequate to meet the needs. Careful attention should be paid to make sure that the planned manpower levels are achieved on schedule. Both BNL management and DOE will need to cooperate with creative solutions to help meet the manpower targets.

Plans were described in the proposal to do detailed modeling and simulation of proposed architectures and to compare estimated performance with that achieved by prototype and early pre-production systems. No concrete work has been done yet here, nor is it clear that current manpower estimates include personnel for this task. This is an important verification of the planned architecture and should not be ignored.

Schedules and milestones for computing system installation should be presented in parallel with experimental milestones to help demonstrate the suitability of the installation plans. These schedules should be watched carefully as understanding of accelerator startup plans becomes more precise.

Recommendations

1. A joint task force including membership from RCF and the experiments should be formed immediately to study and document the computing requirements. RCF personnel need to communicate to the experiments which requirements will affect system architecture and when irreversible decisions must be made; experiment personnel must communicate the level of uncertainty in the requirement estimates and the cost to the physics program where requirements cannot be met (in particular for the reduction in bandwidth to the tape robot system). Revised requirements should be documented in roughly 3 months. Close collaboration of this nature between RCF and the experiments should be ongoing as RCF is assembled and brought into operation, and the experiments gain further understanding of their requirements, to identify problems and solutions as early as possible.

3.0 RHIC Computing Facility (Hardware at BNL) Overall

Findings

The RCF team has assimilated a great deal of input on the computing and data systems requirements of the RHIC project, and considered several logical alternatives in mapping those requirements onto current and projected hardware technologies. RCF has a complex and challenging computer facility architecture, and it is notable that the current designs fit together without inconsistencies, a mark of good design team communication.

Comments

The projected costs of the main components of the RCF are well understood. If costs are greater than expected, there are several ways to scale down capabilities, while maintaining functionality. Likewise, there are ways of scaling up the capabilities, should project requirements dictate them and funding become available. The planned four year replacement cycle is appropriate for this facility.

The RCF staffing plan is adequate, but lean, for choosing, procuring and installing the equipment in RCF. To meet requirements to process and store data in 1999, it is very important to obtain the planned new staff to perform engineering and integration work.

The computer room space assigned to RCF does not allow for growth of facilities beyond the initial complement. Also, space has not been identified for the large number of tapes to be stored on shelves over the coming years of RHIC operation.

When RHIC is fully operational, there will be circa 300 visiting scientists on site at a time. The vast majority of them will need computer access to RCF facilities. Work locations for their desktop computers/X-terminals have not been identified.

Recommendations

1. Computer room space planning for RCF computer and data systems growth should be done before equipment is installed. Plans for tape storage and accessibility need to be made.

2. Locations for visiting scientists need to be identified, providing space and connectivity for their computers/X-terminals. The scope of the desktop computer/Xterminal resources provided by RHIC needs to be determined.

3. Since staffing is lean, it is important to leverage knowledge and efforts at other sites. For example, collaboration on the development of RCF and the remote RHIC computing facilities (e.g., MIT, NERSC, RIKEN) could leverage effort and avoid unnecessary divergence of computing facilities.

3.1 General Computing Environment (GCE)

Findings

The General Computing Environment provides general purpose on site computing and file resources for RHIC users, and serves as the local and remote user gateway to the other RCF resources. It has the characteristics of a general user facility, with a potential user base of 300 on site users and at least that many off site users.

Comments

Because GCE will serve a large number of users, planning for services and operations needs to be carefully done. Much useful information is available from other facilities.

User demand for general file server space may ramp up quicker than demand for other components of the RCF. Therefore, the deployment of file servers may have to be moved up earlier than projected.

Recommendations

None

3.2 Central Reconstruction Server (CRS)

Findings

The Central Reconstruction Server performs the initial processing of raw experimental data into more compact and easily analyzable form. Data Logger systems take the stream from an experiment and partition it into a component for immediate reconstruction and a component directed toward storage for later reconstruction. An Intel processor farm carries out the initial analysis, using standard task farming methods and experiment specific algorithms.

Comments

The CRS is well thought out, and the remaining design decisions are straightforward.

Current architecture and capacity plans assume that raw data will only be reconstructed once. Given the nature of the experiments and the anticipated data flows, it is not possible to predict the fraction of data that must be reconstructed in the first pass, or the precise data volumes. For at least some experiments, later calibration adjustments may dictate that all data be reconstructed a second time.

The current RCF plan calls for reconstruction of 50% of the events on fly, before the raw data are stored in the MDS. However, experience with existing collider detectors indicates a steady state delay of a few weeks between data taking and mature calibrated reconstruction. A more likely scenario is therefore that 10% or less of the events will be reconstructed on the fly for monitoring purposes while the remainder stream directly to the MDS, with the total to be reconstructed later.

Although the CPU capacity of the CRS is not scheduled to reach its nominal value until 2001, the data movement and logging capacity should already be at its nominal value by the start of the data taking.

Recommendations

1. The Data Logger systems and their associated data paths should have the capability to steer all of the data stream to the Managed Data Store. The path between the reconstruction processor farm and the MDS should be able to simultaneously handle the same size flow.

2. The reconstruction farm should be designed to be scalable to at least 4 times its currently planned capability.

3.3 Managed Data Store (MDS)

Findings

The Managed Data Store is the heart of the RCF. It holds data that must be readily available, in a disk cache and robotic tape archive. Key design factors included the disk and tape capacities, media cost and size, performance and aggregate throughput capabilities.

The choice of tape media is crucial, and the RCF team is well versed in the options.

Comments

The size and throughput requirements of the MDS make it a technical challenge. Its performance will largely establish the overall throughput limits for the RCF. The total data volume and the overall processing flow, i.e the patterns of usage of experimental data in its raw and derived forms, have a significant impact on the size and architecture of the MDS. These characteristics are not completely understood at this time. More information on the requirements and flow for data processing are needed from the experimental collaborations. This should be a two way interaction, with draft data system designs being shared with the experimental programs.

The capacity to create approximately 150 Tbytes of tape per year would be required to support export of data to the planned PHENIX Japan Computing Center.

The tradeoffs between RAID and non-RAID disk hardware have not been adequately considered.

It would be worthwhile to keep the option open to support more than one kind of tape medium.

Recommendations

1. The design of the MDS should take into consideration the possibility that all raw data may be reconstructed (again) after it has been first recorded to tape.

2. Carefully consider tradeoffs between reliability, capacity and performance when designing the disk cache for MDS.

3. Factor in job flow analyses and temporal data use patterns to determine the capacities and relative sizes of the disk and robotic tape store.

4. Evaluate requirements for data export capacity and associated export tape medium for all four experiments.

5. Evaluate the requirements for data import, e.g. simulated events from the PHENIX - J Center

3.4 Central Analysis Server (CAS)

Findings

The Central Analysis Server performs data mining operations on reconstructed data, analyzes elements of selected data subsets, and analyzes collections of data subsets. The current design is comprised of a processor farm with a high bandwidth data connection to the MDS.

Some uses of the CAS will be I/O intensive, e.g., data mining, while other uses may be CPU intensive.

Comments

The job mix and associated data volumes are not fully characterized at this time, and will have a substantial impact on design decisions.

CAS designs having two architectures should also be considered, e.g., targeting I/O and CPU intensive analyses.

The MDS and CAS need to be very tightly integrated to achieve the high performance and throughput required by RHIC analysis. MDS and CAS will also have many software components. Managing and limiting this complexity will be important in successfully achieving project deliverables and schedule. For example, in the aggregate CAS + MDS system, each logical function does not need to translate into a separate physical component.

Recommendations

1. The characteristics of the CAS job mix need to be better understood, in close association with the RHIC experiments, e.g., how much of the projected analysis load has high I/O (and low CPU) requirements?

2. Because there are limits to what can be known about the overall volume of CAS data

processing in advance of full operations, the CAS system should be scalable and flexible.

3. Exploit opportunities to leverage from and collaborate with other projects with similar architectures.

4.0 Data Handling and Resource Management Software

Findings

The RCF will be a complex system requiring data handling and resource management software in addition to the physicists' codes and the operating systems of the devices. Some components of this additional software can be bought commercially, for example the mass storage management system HPSS. Other components, for example reconstruction farm management software, can be based on the working systems already installed in other laboratories. Substantial fractions of the software needed to support data mining and data analysis do not exist anywhere. No firm decisions have been made to acquire any commercial software components of the production system.

Reconstruction farm software will either be cloned from existing working software or developed by RCF. A decision on the strategy will be taken in late fall 1997.

Data mining software is expected to be provided by the Grand Challenge project at LBL, but the RCF team favors simultaneous development of a fallback strategy of "coherent" analysis where the system makes serial passes through the data executing merged queries from several users.

Analysis software will be adapted, if possible, from a solution developed outside RCF. However no source of this software has been identified and the "personnel to even begin working on this problem do not exist at this time" at RCF.

Comments

Lack of manpower at RCF is a potentially fatal problem. The manpower shortfall affecting data handling and analysis software development corresponds closely to the difference between the current situation and the "unconstrained" need presented by the RCF head.

At present nobody is working on analysis and the principal strategy for data mining is to use the results of the Grand Challenge project. In its present form this strategy is not valid, since the declared aims of the Grand Challenge are to address only a fraction of the software needed for RHIC data mining. In addition, the Grand Challenge project has not yet accepted responsibility for providing any components of the system needed for RHIC.

The manpower needed at RCF should combine interest and skills in data management with experience in experimental data analysis. Such people are in limited supply. Hence it is essential to limit their tasks to those that cannot be achieved or avoided in any other way.

Part of the need for data handling and resource management software arises from the choice of loosely coupled farms as the principal computing resource. The software needs of a farm dedicated to serial event processing with moderate I/O are well understood in the HEP/NP community. The cost advantages of the farm solution will outweigh its software costs especially if care is taken to make maximum use of existing software solutions. The cost benefits of farms used for data mining and analysis are unclear and depend on the detailed nature of the tasks. It is likely that more monolithic SMP hardware would perform better for high I/O tasks and would reduce, but not eliminate the demands on the data handling and resource management software.

The use of commercial (or nearly commercial) software such as HPSS or Objectivity should not be restrained by its costs. Currently, either product would cost the equivalent of 2-3 FTE/years. If the product could play a valuable role it should be acquired and deployed rapidly so that the development of the software to glue it into the system can be appropriately focussed.

Although the scale of RHIC data handling needs exceeds that for experiments at Jefferson Lab, Fermilab or SLAC, ongoing work at these laboratories is likely to be very relevant to RHIC needs and there is genuine interest at the labs in the challenge presented by RHIC. It is important that RCF makes every effort to participate in collaborative information exchange or work. It is also important that RCF chooses hardware and software broadly aligned with those chosen by other labs. This logic mandates a choice of HPSS, strongly deprecates the choice of Solaris for Intel processors, and favors LSF as a batch management system.

It was noted that there is no definition of the data handling and resource management functions to be performed by RCF and hence no clear perception by the experiments of how they should plan to exploit these services. In addition, the RCF hardware acquisition plans were not complemented by plans from the experiments to stress test the developing analysis system. This lack of formal involvement of the experiments until the time of RHIC startup would make it impossible to have any confidence that the RCF hardware and software could perform the data mining and analysis functions.

To establish this confidence, a mock data challenge should be scheduled in which the experiments attempt to analyze a substantial simulated data sample using the software (experiment and RCF) and hardware planned to be ready at the time of the challenge. Planning should be adapted, if necessary, to ensure that a testable combination of hardware and software will be available.

Recommendations

1. The RCF personnel working on the data handling and resource management part of the RCF task should be augmented to reach the levels requested by the RCF head (5.5 in 1997, 8 in 1998, 9 in 1999). Skilled personnel are needed and must be recruited vigorously from within and outside the DOE/NP program.

2. The LBL Grand Challenge project should commit to providing defined components of the RCF software including prototype and pre-production versions. RCF hardware should be available to and should be used by the Grand Challenge.

3. HPSS should be acquired and installed as soon as possible.

4. RCF and the experiments should decide rapidly whether to adopt Objectivity. If the decision is positive, the software should be made available to the RHIC community as quickly as possible.

5. System design and product selection should take appropriate account of the additional manpower required to implement and exploit architectural solutions and particular products that are not mainstream in HEP/NP.

6. A "mock data challenge" should be scheduled approximately one year before the expected arrival of real data.

5.0 Site and Off-site Network Connectivity

5.1 On-site Network Connectivity

Findings

The Committee finds that plans for on-site connectivity seem well in hand. The RHIC team has followed an appropriate course of action during the planning phase of the project, taking advantage of recent developments in local area network technologies (full duplex Gigabit Ethernet). We believe the network design for moving data on-site will meet the needs of the project in a timely and cost-effective manner. The RHIC design team has prepared a reasonable fall-back option based on Fiber Channel which, although more costly, will meet all the technical needs of the project.

Comments

None

Recommendations

None

5.2 Off-site Network Connectivity

Findings

Like any contemporary large distributed scientific project, RHIC is at the mercy of the current state of the Internet. The committee feels that minimizing risk for the project will require a clear understanding on the part of both the project and the Department of Energy of the trade-offs to be made in the balance between local and remote computing, as well as in the philosophy of how the remote collaborative computing at DOE and non-DOE sites is handled.

Comments

The Internet today is a public arena with three overlapping user communities, with very different interests and agendas. First, there is the traditional university and government (funded) community. Second, there is the new community of commercial users building Intranets and Extranets for business purposes. Finally, there is the community of private citizens buying $19.95/mo connections through which to surf the web. All three user communities have created distortions in the fabric of the Internet based on artificial subsidies and usage patterns. The immediate result of this has been the call for creation of an Internet 2 by the university community, a call for connecting the largest of the research-1 universities to the NSF's v BNS (at OC3 speed or better), and pressure on the Department of Energy to disconnect its dedicated lines to university physics departments. Current DOE policy is to disconnect all dedicated university lines except where formal letters are drafted by university presidents or chief information officers explicitly requesting a direct connection to ESnet.

When RHIC is in operation, it is anticipated that the process of analyzing and reducing data will result in an overall reduction in data volume order of 10^8, from tens of terabytes of raw data to tens of killobytes in published results annually. Moving this analysis stream off-site as early in the reduction process as possible will off load a significant requirement from the RCF. The trade-off for doing this is that the earlier in the analysis process it is done, the greater the volume of data that will need to be moved and the greater the exposure to problems in the Internet. At one extreme, the analysis could all be performed at the RCF, with both local and remote users working through X-terminals. At the other extreme, data could be duplicated early in the analysis process and shipped to remote users by overnight air express or through the network.

ESnet, the DOE's mission-oriented network has traditionally provided an extremely high level of reliability and connectivity between the DOE labs. The committee expects this to continue, and that BNL's link to the ESnet will be increased to at least an OC3 (155 Mb/s) connection. However, a significant fraction of the RHIC collaborators will be located at non-DOE sites (university and foreign locations) where they will be directly competing with web-surfers and other users at their institutions.

Recommendations

RHIC experimental groups and BNL RHIC management should maintain a close working relationship with ESnet and the DNP program office. They should monitor the on-going needs for wide area network connectivity and be prepared to procure additional guaranteed network resources if needed.

In addition, the RHIC PI's located at university sites should work with their local network organizations to assure sufficient connectivity for their analysis needs.

6.0 Plans for non-BNL Computing Centers

Findings

Substantial aspects and amounts of RHIC computing resources, software development and data processing are planned to take place at non-BNL facilities. Of the approximately 1,000 physicists planning to activity participate, an estimated 700 will be at non-BNL locations. In terms of the four RHIC experiments:

BRAHMS has the lowest requirements for computing and storage. BRAHMS plans to primarily work at BNL plus some final analysis at non-BNL locations using desk top resources.

PHENIX plans to perform all of its simulations and a third of its processing of DSTs at the RIKEN BNL Center in Japan. This involves copying at BNL all of the PHENIX DSTs, estimated at 150 TB/yr, and transporting them to Japan. One third of the required analysis on the DSTs would take place using facilities at the Japanese Center which also would serve as a regional center for other Asian countries participating in RHIC. The other two thirds of the PHENIX analysis would take place partially at BNL and partially at other non-BNL facilities.

PHOBOS, while a modest size experiment in terms of construction costs, has substantial requirements for computing. The PHOBOS simulations and substantial fraction of the data analysis are planned to predominantly take place at MIT/LNS.

STAR is considering two options: perform the simulations and a substantial fraction of their analysis at a LBNL facility associated with NERSC, or do this at BNL. Both options require funds in addition to the $7.9M capital funds in the planning for the RCF start up.

For all four experiments four other aspects of the non-BNL computing were presented.

RCF includes 8 FTEs to come from non-BNL user groups.

When data analysis has proceeded to the point that the data is several GB to several tens of GB, the data will often be further analyzed at one of the 87 non-BNL institutions collaborating in RHIC.
The Grand Challenge Project on data storage and retrieval may be used in the RHIC data processing environment.

A major software development activity is required to integrate the software environment to be shared by the RCF at BNL and the four experiments.

Comments

General: Substantial uncertainties exist in the requirements for non-BNL resources for simulations and data analysis at this time. Many estimates have recently changed.

BRAHMS: None

PHENIX: The RIKEN facility in Japan is only a proposal at this time. No contingency plans were presented for the case that an alternative to the RIKEN proposal would be required.

PHOBOS: The MIT/LNS plans for PHOBOS appear adequate for the proposed tasks.

STAR: The cost to the Nuclear Physics program are similar for the BNL and LBNL/NERSC proposals to provide additional computational resources to STAR. The issues are:

- Would placing resources at LBNL significantly help both the STAR and overall RHIC computing efforts via leveraging non-RHIC managed resources, particularly those associated with LBNL computational sciences manpower?

- Would functionality be diminished due to complexities involved with sharing resources and compatibility issues.

- What economies of scale are involved?

RCF NON-BNL FTES: History, the training of physicists and current market conditions indicate that it is unrealistic that non-BNL user groups could come up with 8 FTEs capable of performing useful computer science in support of the RCF.

NON-BNL DATA ANALYSIS: The requirements for supporting the 700 users located at non-BNL institutions appear to be considered as part of the desk top requirements for those groups.

GRAND CHALLENGE PROJECT: There appeared to be considerable uncertainty as to whether the Grand Challenge Project on data storage and retrieval will be used in the RHIC data processing environment.

SOFTWARE DEVELOPMENT: There appears to be a minimum of project management coordination of the software developments required to integrate the software for the computing environment of the RCF and the four experiments. Additionally, there is considerable missing manpower for this activity at both BNL and non-BNL locations.

Note, RFC plans on building up in the next two years a staff (34 FTEs) that is roughly half of what FNAL is currently using to provide essentially the same size data processing environment for their next Collider run, Run 2. RHIC experiment operations will begin at essentially the same time as Run 2. The FNAL group has a legacy of decades of experience. The BNL plus non-BNL efforts for RHIC computing are in their infancy and include a widely distributed software development activity.

RECOMMENDATIONS

1. Initiate an effective project management plan (schedule, manpower, milestones, cost) for the development of the software for the RHIC, i.e. BNL and non-BNL, data analysis/computational environment.

2. Make effective use of non-BNL resources for software development. This includes both working with non-BNL RHIC users and communicating with non-BNL laboratories such as CERN, DESY, FNAL, Jefferson Lab, and SLAC.

3. Perform a coherent global optimization for the manpower associated with providing the data analysis computational environment.

4. Actively support the RIKEN proposal.

5. Re-evaluate the optimal location of computational resources and the associated computational model following the re-evaluation of the computational requirements of the four experiments.

7.0 Cost, Schedule, and Funding (RHIC-Wide)

7.1 Cost

Findings

Cost estimates have been developed for the RHIC Computing Facility (RCF) and the non-BNL RHIC computing centers. The RCF cost estimate for capital expenditures, organized by the work breakdown structure, was presented in some detail from FY 1997 through FY 2001. Capital expenditures are estimated at $12 million and it is intended that the facility reach nominal year capacity levels at the end of this period. Contingency is not budgeted separately and the scope and efficiency of the facility will serve as contingency as necessary. A projection of the RCF operating costs through 2001 was also presented.

A staffing plan has been established that requires personnel to be supported by RCF construction and pre-operations funds, the BNL Computing & Communications Division, and the RHIC experiments. The plan requires that staffing levels increase from 12 to 34 in the next two years.

Summary level tables of the cost estimates and staffing plans are provided in Appendix D.

Comments

There has been a good effort to define the scope and estimate the costs for the construction and operation of the RCF. The cost estimate for capital equipment appears adequate and the strategy to use the facility scope as contingency is reasonable provided that the facility hardware is readily scalable. The attempt to benchmark the RCF costs and personnel estimates against the computer resources used for Collider Run 1 at Fermilab was instructive and indicates that RCF staffing levels are low, particularly in the area of software development and maintenance. As a result, the total cost estimate for the RCF is probably low and it will be a formidable challenge to deliver the projected capacity on schedule.

The BNL RCF is the core of the RHIC computing effort, with additional plans for significant off-site computing capability. A complete picture of the total resources, including labor, for the entire RHIC computing effort is needed. This estimate should identify the expected sources of funding and support for each piece of the program, and the projected costs for steady state operations and capital improvements.

Recommendations

1. Develop a comprehensive estimate of the total resources and sources of support required to the make the entire RHIC computing effort a success.

2. Evaluate the current plan for significant contributed manpower to RCF by the experiments and revise these plans as necessary to assure RCF staffing needs are met.

7.2 Schedule

Findings

The schedule for RHIC computing calls for the computing facilities to reach desired performance levels in 2001, the second full year of RHIC operations. This overall schedule appears to be consistent with the expectations of the experiments and the projected performance estimates for the collider. A milestone schedule for the RCF has been developed that identifies key technology decision points for major elements of the work breakdown structure. The RCF schedule is provided in Appendix E.

Comments

The milestone schedule for the RCF is an appropriate first step in the process of developing a integrated schedule for the RHIC computing effort. The milestones were developed by the RCF technical leads for each of their respective areas. The RHIC Computing Head intends to review

the relationship between milestones to assure proper sequencing of these milestones.
In general, the schedule is ambitious given current staffing levels. It is important that another iteration of the schedule be developed that reflects the following information: 1) the RHIC computing effort through 2001; 2) a review of the interconnections between milestones; 3) milestones that describe the staged capacity upgrades of the RCF; 4) relevant milestones for the experiments, e.g., a mock data challenge, and including interconnection between RCF capabilities and experiment requirements as a function of time; and, 5) integration with off-site computing facilities.

Recommendations

1. Complete another iteration of the master milestone schedule that reflects the full scope of the RHIC computing effort and related milestones with the RHIC program.

7.3 Funding

Findings

The near term capital funding profile for the RCF is consistent with current work plans. This funding is provided by Nuclear Physics as part of the RHIC Additional Experimental Equipment program previously reviewed by the Nuclear Science Advisory Committee. Other funding sources for the RCF include RHIC pre-operations, RHIC operations, and contributed support from the BNL Computer and Communications Division. Support for significant off-site RHIC computing includes plans by the PHENIX collaboration for a computing facility in Japan, the PHOBOS collaboration's plans for use of the MIT Laboratory for Nuclear Science Scientific Computing Facility, and consideration by STAR for simulation computing at NERSC.

Comments

The contributed resource of BNL Computing and Communications Division staff to the RHIC Computing Facility is the result of negotiations and commitments made by BNL management. It is the view of the committee that these commitments are reasonable and appropriate given the central role the RHIC facility will have in the future of BNL. DOE and BNL management should work together to assure that these commitments are fulfilled.

Funding plans are not in place for computing capacity that satisfies the STAR request. Two options are under consideration: expansion of the RCF or computing at NERSC. Regardless of the option chosen, additional DOE funding will be required. A plan that captures all of the funding requirements for RHIC computing would be useful as a management tool.

Recommendations

1. A funding plan should be developed for the entire RHIC computing effort.

8.0 Project Management

Findings

Within the last year, RHIC management has been able to identify a capable, permanent leader of the RHIC Computing Facility, who brings with him both an extensive knowledge of HENP computing and the RHIC community. The director of RCF has been charged by DOE to also be responsible for plans for all of RHIC Computing, both on-site at RHIC and off-site at 3 regional centers for STAR, PHENIX and PHOBOS, and the many smaller groups at collaborating national laboratories and universities. This is a very challenging task which requires careful balancing between the central needs at RHIC and the regional needs of the experiments.

The manpower within RCF has increased dramatically within the last half year. The CCD division at BNL has contributed with 6 FTE's, which have been a very valuable addition. However, the total manpower associated with RCF is still behind schedule. Currently one open position has not been filled and only a fraction of 3 FTE's, which were supposed to be contributions from the RHIC experiments, have been recruited.

Capable technical leaders have been identified in many areas. These people seem to be well chosen and technically capable of leading the computing efforts in their respective areas.

The interaction with the RHIC experiments has taken place through bi-monthly meetings between RCF staff and the computing representatives of the experiments. Recently, weekly meetings between people resident at BNL have been initiated.

The committee was presented with a very preliminary task and milestone plan (WBS) and graphical outline of an organizational structure for RHIC Computing, based on proposals from the ROCOCO2 committee.

Comments

The current management has only had a few months to get organized and has already accomplished a lot. However, several management areas seem to need further attention. The current organizational structure, overall, as opposed to RCF management, seems rudimentary, and a more complete management plan needs to be implemented. In the area of RHIC Computing, this plan should be worked out in consultation with the RHIC experiments. A much more complete task list (WBS) with realistic milestones and resource allocations needs to be created. In particular, it is important to take the uncertain manpower situation into account in the resource allocation, and to work out a more detailed plan for the overall manpower needs that currently appear barely adequate.

The RCF director needs more administrative assistance.
The RCF management should be more vigorous in its attempt to find collaborative solutions with other HENP sites, in order not to duplicate efforts.

Recommendations

Complete an implementation plan for RHIC Computing that identifies the resources, tasks and schedules required to complete the project.

Appendix A - Review Letter and Charge to Committee

Appendix B - Review Participants

Appendix C - Review Agenda

Appendix D - Cost Tables

Appendix E - Project Schedule

Appendix G - Action Items