University of Connecticut
Office of the Vice Provost for
Research and Graduate Education
Research Advisory Council
Serving the Storrs and Regional Campuses

Application
1999-2000 Annual Equipment Grant Internal Award Program

Deadline: Monday, October 25, 1999, January activation; submit 1 original

---

1. **Item of Equipment:** Computer farm consisting of 1 server and 2 client nodes
[Specify the item of equipment or series of items that function as a system.]

2. **Amount Requested.** [Specify the requested information. Include a letter specifying amount and type of match from each source. Please include source FRS numbers.]

a. Amount requested via Equipment Competition: $ 20,000
[Minimum amount allowed is $5,000]

Amount Source

b. Matching funds from other sources: $ 13,500 department
[Specify amount(s) and source(s)]

c. Total purchase price [a + b]: $ 33,500

3. **Applicants' Names and Signatures.**

Richard Jones, U-44

Name, department, U-Box and **signature** Name, department, U-Box and **signature**

Robin Cote, U-44

Name, department, U-Box and **signature** Name, department, U-Box and **signature**

Juha Javanainen, U-44

Name, department, U-Box and **signature** Name, department, U-Box and **signature**

4. **Department Head Statement and Signature.** (This proposal is consistent with the program of my department and the funding of it would be an appropriate use of indirect cost recovery dollars.) *Important*: On an attached sheet, you must state the importance of this equipment to the total program of research and graduate training of the department(s). In instances of multiple requests from a department or program, it would be helpful if you could rank-order the proposals.

William Stwalley

Department head's name and **signature** Department head's name and **signature**

10. Department head's name and **signature** Department head's name and **signature**

A Seed for a Physics Simulation Farm using Commodity Processors and Internet 2 Communications

Project Abstract

Computer simulation is now a standard tool in most areas of physics research. Numerical models cannot replace the analytical approach to working out the consequences of a theory for experiment, but they are a useful tool in cases where the large number of variables or steps involved prohibits a solution by exact methods. An example of this is the calculation of the dynamics of atom-atom and atom-molecule collisions, where important progress is being made using numerical models to calculate the probabilities for different reaction channels to occur. Here a complete numerical solution to the dynamics in terms of the basic laws of physics would be prohibitive; the goal of this theoretical research is to establish models which are both numerically affordable and which successfully predict the important features of the scattering data. In the related area of many-atom dynamics, the goal is to discover the essential features of atom-atom interactions which are responsible for interesting collective effects which have been observed in many-atom systems. The selective inclusion of refined features of the atom-atom interaction in a simulation involving many atoms allows one to observe the connections between the simpler pair-wise interaction and more complex collective effects. Simulation is also an important tool in experimental research. High-energy physics experiments now spend a significant fraction of their budget and man-power on Monte Carlo simulation of the detailed response of the detector to the particles and radiation that are produced by the reactions they are studying. Because of the large number of resources and researchers involved, these collaborations rely on a distributed computing model, with clusters of workstations located at various institutes exchanging data and results over the internet. The Physics Department at the University of Connecticut is now undergoing significant growth in each of these areas of research. Two of the three P.I.'s on this proposal are junior faculty members who are initiating a new research effort in their respective fields. The facility we are proposing is the beginnings of a processor farm, an expandable cluster of low-cost processor nodes built around a central file server and connected over a high-performance conventional network. The farm will be equipped with a high quality-of-service internet connection to other research institutions. It will be used by all of us for producing results related to our research and for exchanging data with similar facilities at collaborating institutions. The initial investment requested in this proposal represents an increase by more than a factor of 10 over what is available to us at present in processing power alone, and a qualitative improvement in the internet service. More importantly, it is the seed of a facility that can grow and evolve with computer technology. The most rapid evolution in the industry is taking place in the speed of processors, which has more than doubled over the last two years. Purchasing independent compute nodes of a generic type and connecting them together with industry-standard 100Mb/s ethernet will allow future expansion of the farm to follow the most economical path as computer technology evolves. Expertise already exists between the P.I.'s on this proposal in setting up and managing a unix cluster. The team who set up and manages the 75-node Linux farm at Jefferson Lab have already been an important resource for advice, as have other university groups connected with that laboratory who are setting up seed clusters similar to this one. The computer center at the CERN laboratory in Geneva, Switzerland maintains an updated repository of scientific and public-domain software of relevance to scientific research for the major unix operating systems to high-energy groups around the world, which we will mirror on our file server.

The networking upgrades contained in this proposal are coordinated with the plans being formed by the UConn team in charge of the Internet 2 project, so that the two projects are mutually beneficial. The simulation farm project mates distributed computing with performance networking and quality-of-service internet connectivity in a configuration designed for incremental future growth along both of these two axes, for the benefit of physics research at the University of Connecticut.

High-Energy Physics Experimental Research

Research Objectives

The quest of high energy physics research is to discover the elementary particles of which everything is made and to understand their interactions. It is a law of physics that, in order to resolve structures at a progressively smaller scale, one must employ as probes particles of progressively higher energies. This has led to the modern picture of nature as a succession of scales, where an object that behaves like a simple particle at one scale is revealed as a complex of simpler objects at the next. A pre-eminent example of this is the succession atom-nucleus-proton-quark. Much of the research in physics is focused on the boundaries between scales, where two descriptions meet. The Thomas Jefferson National Accelerator Facility (Jefferson Lab) in Newport News, Virginia was built to explore the region between the proton and the quark scales. There is evidence that intermediate between the nuclear and quark scales there might exist a so-called hadronic scale with its own description that borrows particles and concepts from both but is equivalent to neither.

This research is more often classified as nuclear than high-energy physics, but from the experimental point of view the distinction is not important. The typical experiment consists of a beam from a particle accelerator striking a target surrounded by a large multi-function detector which records as much information as possible about the particles and radiation produced during the collision. Often the primary products of the collision cannot be observed directly in the detector, but decay into other particles, which themselves produce new particles through interactions in the detector, and so on until enough energy is deposited in a sensitive element of the detector to register a detectable signal in the electronics. The challenge of analyzing data from such an experiment is to put together enough information from these recorded signals to determine what were the original products of the reaction. Although the secondary interactions by which the energy from the primary particles is dispersed into the detector are usually well understood, there are many hundreds of them per collision, and they are stochastic in nature so that even if the primary reaction were identically repeatable, the pattern of hits in the detector would be different every time. The stochastic nature of the problem makes it ideal for simulation by Monte Carlo methods. Monte Carlo simulation of a high-energy physics experiment relies on pseudo-random number sequences to generate realistic examples of hit signatures in a numerical model of the detector, starting from a given set of particles from the primary interaction vertex and an extensive library that describes the interactions of known particles in materials. Knowing both the original reaction kinematics and the results that emerge from analysis of the recorded hits, one can estimate the average response of the detector, measure its resolution, develop better analysis algorithms, and determine the efficiency of the detector and the analysis package. Modern experiments rely on simulation at every stage in their life cycle, from design to calibration to analysis to efficiency corrections and the estimation of errors on the final results.

The University of Connecticut Physics Department has recently initiated a new research program in this area with the approval of two new junior faculty positions in a cooperative agreement with Jefferson Lab: Richard Jones (1996 hire) and a second experimentalist still under search. The agreement also included a theoretical position in a related area, filled by Michael Ramsey-Musolf in 1998. R. Jones is presently leading the UConn experimental effort at Jefferson Lab through participation in the Radphi (E94-016) experiment and the Hall D collaboration. Both the Radphi experiment, presently on the floor at Jefferson Lab, and the Hall D facility, still at the proposal stage, are devoted to the search for exotic mesons, novel states of matter which are predicted by many theories of hadronic structure. The farm entailed in this proposal is relevant to both of these research projects.

Strategic Plan

The Radphi experiment is being run by a collaboration of about 35 researchers from 10 institutions. The experiment was approved in 1994, has passed three successive commissioning runs during the last three summers, and expects to collect its major physics data during the summer of 2000. Together with his students, R. Jones has made two primary contributions to the Radphi experiment. In the area of hardware, they spearheaded a major upgrade of the Radphi detector using components of a detector borrowed from a prior experiment at CERN. In the area of software, the group developed the Monte

Carlo software for the Radphi apparatus. With this program they were able to demonstrate that the above-mentioned upgrade would improve the performance of the experiment by a factor close to 3 and enable improved rejection of background. The lack of a thorough Monte Carlo simulation was a weakness of the collaboration when we joined it, and our expertise in this area has contributed to our group's visibility at the laboratory.

The Radphi Monte Carlo program was written and tested by two students working over the summers of 1998 and 1999 on-site at Jefferson Lab. Now that the program has been developed, it must be put to work generating event samples. In production running, the Radphi experiment will collect data at a rate of 100 collisions per second, or about 8 million events per day. Only about 10% of these survive a pre-analysis selection pass. Thus over a run of 60 days we expect something on the order of 50 million analyzable events. Experiments typically budget for a factor of 3-5 more Monte Carlo events than real data, so that the errors on the results do not receive an appreciable contribution from the statistics of the simulation. Thus we expect to generate about 200 million simulated events to be analyzed in parallel with the data taken in 2000. Benchmarks taken on a Pentium-III processor at 500MHz show that a single processing node of this type can generate 5-10 Radphi events/s, depending on the primary reaction being generated. The exact mix of reactions to be generated will be decided as the analysis of the real data progresses. This leads to an estimate of 250-500 processor days for Monte Carlo production. Experience indicates that these figures should be doubled to take into account re-generation due to flaws in the numerical model that are revealed as the comparison between real and Monte Carlo data proceeds.

According to this estimate, only 60-120 days on the initial 8-node farm would be required to complete this task. However, the farm is to be shared between three research groups, and so some partitioning of this load with other institutes in the Radphi collaboration may be required to complete the task in a timely fashion. Once the Monte Carlo data are generated, they must be passed to the analysis group, typically at another institution. There the simulation data are passed through the same analysis chain as is used to process the real data, and the resulting distributions are used to measure the efficiency function of the detector and analysis package. The total size of the Monte Carlo event sample to be generated is about 5 terabytes, which is not a high data rate when spread out over the time it takes to do the generation. However, the shear size of the sample requires special facilities for archiving. These facilities exist at Jefferson Lab, and are intended to be accessed by outside institutes over Internet 2. Several university groups have already established Internet 2 connections between their home institute and the laboratory, and have begun working the bottle-necks out of the connection.

There is a 75-node Linux cluster already in operation at Jefferson Lab. However 80% of its capacity has been allocated to another experiment for the next few years, and for the remaining 15 nodes one must compete with another dozen or so experiments of comparable complexity to Radphi. Our collaboration plans to request additional computer time from the laboratory in order to do most of the real data analysis there, and to support the request by showing that home institutions are contributing significant resources for carrying out the Monte Carlo simulation there. The timely construction of a processor farm at UConn would proposal allow us to play a key role in making this case to the lab management. The establishment of an Internet 2 connection to our cluster will strengthen the linkage between the two processing streams.

The Hall D facility being planned for the coming decade, following the energy upgrade at Jefferson Lab foreseen for 2005, will have Monte Carlo needs that dwarf those of Radphi. However there is still significant time before major production must begin, during which computing and networking technologies will advance. The collaboration will also be much larger than Radphi, with more groups having developed facilities for distributed computing by that time. The network facilities being put forward in this proposal will put our group on the leading edge of this development of clusters coupled between institutes. It will leverage the experience in setting up and managing a unix cluster that exists among the P.I.'s on this proposal to increase the impact of UConn research on the nuclear physics community and strengthen the ties between our diverse research groups within the physics department.

Project Description

The equipment being requested in this proposal are the core components of a processor farm: a high-capacity file server and networking infrastructure. As a match, three P.I.'s on the proposal are together contributing a total of three dual-processor Pentium-550 compute nodes. One of these three nodes will be payed for out of a federal grant, which is excluded by the rules of this competition from being included in the total matching funds. Nevertheless the total package being put forward includes a dual-processor server plus three dual-processor nodes called clients in the description which follows.

The Server

Roughly half of the funding requested in this proposal is for a central file server. This server will be used as a central repository for the scientific software and user files that must be shared by all nodes in the farm. In view of the possibility that the farm will include nodes of several different processor types (Intel, Alpha, PowerPC) 25GB of disk storage is foreseen for this purpose. The rest of the disk space on the server will be reserved for staging output from simulation jobs before they are written over the network for archival or immediate analysis. Efficient use of a high-bandwidth network connection requires a server with sufficient local disk capacity to sustain transfers and buffer data over a considerable period of time. As other applications of Internet 2 connectivity come into use (eg., collaborative applications) the server should have sufficient disk capacity to support the data buffering required for these applications. The 50GB of disk space set aside for staging of simulation data is not intended for data archival purposes. A high capacity tape drive is included for backup and occasional archive of data locally on tape. The server will run the Linux operating system, which takes effective advantage of the dual-processor design of this server and can exploit the high bandwidth of the 100Mb/s connection between the server and client nodes that is necessary to handle the requests from the clients for input data, and to merge the outputs streams from the many nodes working on a problem into a single output file. The server processors will also be able to execute simulation jobs in the background.

The Client Nodes

The client nodes chosen for the initial purchase are actually full-service high-end workstations with two processors each. Most farm implementations today are designed around clients that cost less than $1000 per processor. To reach this level requires careful optimization of the design around a particular simulation problem, which goes against our wish to share this resource for solving a variety of problems. As we gain experience with the resources required for each class of problem, new nodes can be added that are optimized to our particular needs for best performance/price ratio. Our choice for the initial clients also reflects the wish that these machines perform a second function as a work station for students carrying out research on the farm. This is why they come equipped with screens.

The Intranetwork

The nodes of the farm will be connected between each other and the server over local area network consisting of 100Mb/s ethernet links connecting each node to a central switch in a star configuration. The switch can connect up to 24 nodes at 100Mb/s apiece and a total of 1Gb/s between all nodes. The switch would also house the connection to the external network.

The Internetwork

This proposal is greatly helped by the fact that a collaboration of researchers at the University of Connecticut has already received federal funding for Internet 2 connectivity. The funded proposal actually does describe our Jefferson Laboratory research as one of the activities that would benefit from the new network, but directs the funding in its initial phase toward the specific research projects of the P.I.'s on the proposal. However once this service is in place at the university, any network traffic directed at Internet 2 sites, such as Jefferson Lab, will be automatically routed over Internet 2. To make effective use of this connection will require a fast link between our cluster switch and the computer center where the backbone of the campus network is housed. Also required for this plan to succeed is the coordination of our activities with other groups who are preparing to use Internet 2 resources at the university. Because of this consideration, instead of putting forward our own equipment list for Internet connectivity, we are requesting the equipment that was recommended to us by Robert Vietzke, UCC staff member and chief administrator of the Internet 2 grant. The approximate costs for each of the components listed in this category have been obtained from this source.

The software

The Linux operating system is planned to be used for the server and all nodes of the initial cluster. Some variant of unix, although not necessarily Linux, will be a requirement for all nodes to be added to the cluster in the future. Linux can be obtained essentially free of charge for the major processor types, and comes bundled with efficient compilers for Fortran 77, c and c++ together with their respective run-time libraries. The simulation programs will be custom codes written in-house, but will be able to take advantage of the extensive CERNLIB program library. The CERN laboratory in Geneva, Switzerland provides an important service to high-energy groups around the world by maintaining a repository of scientific and general-purpose software in the public domain, both in source code and binary form, for the dozen or so most popular variants of unix, including Linux. We have obtained approval from CERN to mirror the ASIS repository server, and provides many of the numerical tools needed for simulation. For parallelization of codes, Monte Carlo simulation is already parallel at the event level, and requires no additional work to adapt it for running on a farm. For the atomic physics calculations, some development may be required. Threaded compilers now exist that are able to automatically break up a program into parallel tasks and spread the load across the two processors in a single node, and this level of parallelization will be pursued initially. Sharing the load between processors on different nodes will require the manual restructuring of the code and use of the message-passing interface (MPI) or a special language like high-performance Fortran (HPF) to implement the interprocessor communications which must take place over the network.

Housing

Two possibilities exist for housing the new cluster. One possibility is to install the cluster in a corner of the physics department computer room. Adequate space in that room is currently available for this purpose, and housing the cluster there would allow easy access to the screens on the client nodes by students and post-docs. It already has the environmental controls that are necessary for safe operation of computers. The other possibility is to house it in existing lab space allocated to the research of one of us. A hybrid between these two is also possible if the proper networking cable is installed between the two areas.