Computational Resources for Lattice Gauge Theory

The goals of our research are to understand the physical phenomena encompassed by quantum chromodynamics (QCD) and to make precise calculations of the theory's predictions. This requires large scale numerical simulations within the framework of lattice gauge theory. Such simulations are necessary to solve the fundamental problems in high energy and nuclear physics that are at the heart of the Department of Energy's large experimental efforts in these fields. Computational facilities capable of sustaining many tens of teraflops are needed to achieve our near term scientific goals. Most members of the United States lattice gauge theory community have been working together for several years to develop unified plans for creating such facilities. The hardware we propose will consist of a mix of the QCDOC, a specially designed computer that combines computation and communication capabilities on a single chip; and commodity clusters, optimized for the study of QCD. We are targeting a price/performance of $1M per sustained teraflops. The software infrastructure needed to achieve very high efficiency on these platforms is under development with support from the Department of Energy's SciDAC program.

Major goals of the DOE's experimental program in high energy and nuclear physics are to: 1) verify the the Standard Model of High Energy Physics, or discover its limits, 2) determine the properties of hadronic matter under extreme conditions, and 3) understand the structure of nucleons and other hadrons. Lattice QCD calculations are essential to research in all of these areas. The terascale computing facilities we propose are required if these calculations are to reach the needed accuracy in step with the experiments they support. Our scientific goals were set out in detail in the proposal we submitted to the Department of Energy for funding in fiscal year 2003. This proposal can be found at the URL www.physics.ucsb.edu/ sugar/proposal.pdf, and the report of the panel that reviewed it is attached to this document. In the Appendix of the proposal we provided a description of our scientific objectives, identified projects that will receive high priority for early use of the proposed facilities, and estimated the resources that will be required to make substantial progress on these projects. The advent of terascale computing facilities coupled with experiments currently in progress, offers the prospect of major advances in our understanding of QCD.

The size of the computing resources we seek to construct is driven by our scientific objectives. In order to provide support to the experimental programs in high energy and nuclear physics in a timely fashion, and to keep pace with the ambitious plans of our colleagues in Europe and Japan, the U.S. lattice QCD community requires computing resources capable of sustaining tens of teraflops within the next few years. Resources of this magnitude are not currently available at national supercomputer centers, and would be extremely expensive to provide on general purpose machines in the required time frame. However, by taking advantage of simplifying features of lattice QCD calculations, such as regular grids and uniform, predictable communications between processors, it is possible to construct computers for lattice QCD that are far more cost effective than general purpose supercomputers, which must perform well for a wide variety of problems including those requiring irregular or adaptive grids, non-uniform communication patterns, and massive input/output capabilities. In addition, lattice gauge theory calculations require significantly less memory than most large scale applications, which also serves to reduce the cost of computers dedicated to this field relative to those that must serve a broad range of disciplines.

We have identified two computer architectures that promise to meet the needs of lattice QCD. One is the QCDOC, the latest generation of highly successful Columbia/Riken/Brookhaven National Laboratory (BNL) special purpose computers, which is being developed at Columbia University in partnership with IBM. The other consists of commodity clusters, which are being specially optimized for lattice QCD at Fermi National Accelerator Laboratory (FNAL) and Thomas Jefferson National Accelerator Facility (JLab). This two track approach will position us to exploit future technological advances, and enable us to retain the flexibility to invest in the hardware that will maximize the scientific output at each stage of the project. Furthermore, it ensures a robust national research effort in the face of unforeseeable circumstances in either track. Each architecture has its own strengths, and either may prove optimal for different aspects of our work. The QCDOC project is expected to provide very powerful computing platforms this year at a cost of approximately $1 per sustained Mflops. This platform is likely to be particularly effective in generating the computationally expensive dynamical quark lattices that are crucial for our research. The clusters will allow us to take advantage of the rapid advances in the commodity computing market to build increasingly cost-effective machines over time. The well developed software packages and flexible communications systems of clusters make them particularly advantageous for more complex physics applications and for the development of new algorithms and computational techniques crucial for the advancement of our field. During 2003 and 2004 we will focus on the QCDOC, because it is expected to provide the most capable and cost-effective hardware during this period. Extrapolations indicate that clusters may surpass the QCDOC in cost effectiveness thereafter, so we anticipate switching our focus to them for 2005 and 2006.

We propose to construct a distributed computing facility with major hardware located at BNL, FNAL and JLab. Initially BNL will focus on the QCDOC, while FNAL and JLab will concentrate on clusters. This approach will allow us to take advantage of the very considerable expertise at each of the participating laboratories, while building platforms of the appropriate size to meet our research objectives. Tests run on the gate-level simulator for the QCDOC ASIC indicate that it will meet the design goal of $1M per sustained Tflops, and that its performance will scale linearly to tens of thousands of teraflops. Initial chips for the QCDOC ASIC are expected from IBM in May, 2003 We propose to build a 1.5 Tflops sustained QCDOC at Columbia University in the summer of 2003 to enable the machine designers to construct and debug it in house. We further propose to construct a 10 Tflops sustained QCDOC at BNL in 2004. The development work on the QCDOC, and plans for the construction and operation of the 1.5 Tflops machine are described in our 2003 proposal. The construction and operation of prototype clusters, and our plans for building terascale clusters in the future are also described there.

The construction and maintenance of computational infrastructure for lattice QCD will require a sustained, long term effort. Development work for the QCDOC is being funded through the HEP base program, the RIKEN Laboratory of Japan, and the UKQCD lattice gauge theory collaboration of Britain. The corresponding work for the clusters is funded through our SciDAC grant. The SciDAC grant provides funds to develop software that will enable the U.S. lattice QCD community to exploit both types of computers productively. The software effort is described in the 2003 proposal, with special emphasis on the relationship between this work and the hardware being developed The testing and evaluation of hardware and software is critical for this project. The software is being validated and benchmarked as it is produced. Before building any new, large computer platform, we will construct a smaller prototype to evaluate suitability for our research, and to determine the performance of production codes.

We will select machines (and their configurations) to maximize physics production, and will target software developments to areas with the greatest potential for further performance gains. Our plans for testing and evaluating hardware and software are set out in the 2003 proposal. The project to construct, staff and operate a distributed computing facility for lattice QCD will be managed under the administrative structure created for our current SciDAC grant. The Executive Committee will have overall responsibility for the project. The Scientific Program Committee will monitor the scientific progress of the project, and provide leadership in setting new directions. It will allocate the resources on all hardware built in this project. The Oversight Committee will review progress in implementing the plans of the collaboration, assist in the evaluation of hardware and software, review plans for new acquisitions, and make recommendations regarding alternative approaches or new directions. Members of the committees are listed in the 2003 proposal, where project management is discussed in greater detail. Virtually every senior member of the U. S. lattice gauge theory community is participating in this project at some level, as are a number of computer scientists and computer engineers.

There is a growing emphasis within the Office of Science on the advantages of optimizing the architectures of high performance computers for scientific research in general, just as we have emphasized the advantages of doing so for research in lattice QCD. For this reasons, we believe that our efforts may have impact beyond QCD. In particular, the QCDOC is on the direct design path of the IBM Blue Gene family of computers. We will make QCDOC hardware available for testing and benchmarking by computer and applications scientists to test this architecture before Blue Gene hardware becomes available.

We currently receive approximately $2M per year from the SciDAC program for software development and hardware prototyping. We have requested an additional $2.3M in fiscal year 2003 to build and operate the 1.5 Tflops sustained QCDOC, and to carry out development work not funded under our SciDAC grant. The 10 Tflops sustained QCDOC we propose to construct at BNL in 2004 2ill cost approximately $12M to build and operate as a national user facility for lattice QCD. Constant investments thereafter will enable us to expand our scientific reach to meet the exciting scientific opportunities presented to our field.