National Computational Infrastructure for Lattice Gauge Theory
Progress Report: September 15, 2006 – February 1, 2007

Lattice QCD Executive Committee

R. Brower, (Boston U.) N. Christ (Columbia U.), M. Creutz (BNL),
P. Mackenzie (Fermilab), J. Negele (MIT), C. Rebbi (Boston U.),
D. Richards (JLab), S. Sharpe (U. Washington), and R. Sugar (UCSB)

I. Overview

This report describes progress in the SciDAC-2 Project *National Computational Infrastructure for Lattice Gauge Theory*, from the beginning of the project on September 15, 2006 through February 1, 2007. The objective of this project is to construct the software needed to study quantum chromodynamics (QCD), the theory of the strong interactions of subatomic physics. It builds upon the successful efforts of the SciDAC-1 project of the same name, in which a QCD Applications Programming Interface (QCD API) was developed that enables lattice gauge theorists to make effective use of a wide variety of massively parallel computers. In this SciDAC-2 Project optimized versions of the QCD API are being created for the IBM BlueGene/L and BlueGene/P, the Cray XT3 and its successors, and clusters based on multi-core processors and Infiniband communications networks. The QCD API is being used to enhance the performance of the major QCD community codes and to create new applications. A QCD physics toolbox is being constructed which will contain sharable software building blocks for inclusion in application codes, performance analysis and visualization tools, and software for automation of physics workflow. Software for the management of large clusters is being developed. New software tools are being designed for managing the large data sets generated in lattice QCD simulations, and for sharing them through the International Lattice Data Grid consortium. A common computing environment is being developed for the dedicated lattice QCD computers at BNL, FNAL, and JLab. Finally, research on multi-scale algorithms recently begun in collaboration with members of the Terascale Optimal PDE Simulations (TOPS) Center is being extended. Work is well underway in all of these areas.

This project is an initiative of the USQCD Collaboration, which consists of nearly all the high energy and nuclear physicists in the United States engaged in the numerical study of QCD. The Collaboration's web page *www.usqcd.org* contains information regarding its scientific objectives, membership, and initiatives. All software developed under this grant is publicly available, and can be downloaded from a link on the USQCD web site, or directly from the URL *usqcd.jlab.org/usqcd-software/*. The overall responsibility for the project is vested in the Lattice QCD Executive Committee, whose members are listed at the top of this report. They meet every other week via conference calls to discuss major issues. Robert Sugar, the Chair of the Executive Committee, is spokesperson for the effort, and contact person for the DOE. Ten universities and three national laboratories are funded to carry out work under this grant. Each of them has a local principal investigator who oversees the work at his site. A list of participating institutions and their principal investigators is attached to this report. Richard Brower serves as coordinator of the software effort, and Chair of the Software Coordinating Committee. This committee meets once a week via conference calls to coordinate work, discuss progress on individual projects, and to set priorities.

The Software Coordinating Committee held a workshop at Boston University on October 27 and 28, 2006 to discuss progress and plan the early stages of the SciDAC-2 work. The agenda for the workshop, minutes of it, and copies of view graphs from many of the talks can be found at the URL *super.bu.edu/~brower/workshop*.

II. Progress to Date

In our SciDAC-2 proposal we set out milestones to be achieved by each participating institution during the first two years of the project. We believe we are on schedule to meet all of them. Below we indicate the work currently being done at each institution.

BNL: At BNL, Chulwoo Jung is optimizing software and implementing new algorithms for the QCDOC, BG/L and BG/P computers. The Columbia Physics System (CPS) codes contain the new RHMC algorithm, in production use on QCDOC and BG/L, with both domain wall fermions and p4 staggered fermions. This algorithm has sped up the generation of ensembles by a factor of six or more and BNL is helping to optimize this algorithm for the ASQTAD action in the MILC code. Following work of Pavlos Vranas of IBM, BNL has also achieved 19% of peak speed on BG/L on the single mass sparse matrix inversion subroutine that makes up the bulk of floating point operations in QCD calculations. Enno Scholz has installed and tested all SciDAC software libraries (such as QIO, QMP) on QCDOC and has modified CPS to make use of them. Enno has archived lattices to our web site for public use and has now made it possible, via QIO, for CPS to store lattices into SciDAC format. Both he and Efstratios Efstathiadis are working to evolve QIO to allow light quark propagator storage and cataloging from CPS. This will involve defining the appropriate meta-data to store with the propagators and creating a catalog of stored propagators.

FNAL, Illinois Institute of Technology and Vanderbilt University: FNAL is working to optimize performance on Opteron processors. It has modified the mvapich MPI software so that jobs are correctly bound to NUMA processor and memory nodes. The resulting job launch software has been used in production on the new Kaon cluster since September. This software avoids the 10% to 15% drop in performance which occurs when MPI processes migrate from one Opteron core to another in the course of job execution. The AMD processor simulator has been installed and is being used to analyze the SSE math kernels used in the LQCD SciDAC libraries.

FNAL and the Illinois Institute of Technology have worked closely together on the workflow sub-project since the start of the SciDAC-2 grant. Since September, they has surveyed existing active workflow projects. A workshop with the VDS (now called Swift) workflow project was held at FNAL on December 18. Video conferences were held in the last few weeks with the Askalon and Kepler workflow project teams. A detailed requirements document for an LQCD workflow system was prepared jointly by IIT and FNAL. This document has been used in discussions with other workflow projects and has served well as a learning tool so that the various parties can understand one another. The Karajan execution engine has been installed on one of the FNAL clusters, and one of the weak decay analysis workflows has been coded in the Karajan language as a test case.

FNAL and Vanderbilt have worked closely together on the cluster reliability sub-project since the start of the SciDAC-2 grant. A survey of existing management systems was performed, and the sub-project team decided to evaluate two existing packages (OpenNMS and AWARE) in depth. An initial set of requirements has been developed for the cluster reliability sub-system and is being

used as a reference to evaluate these existing systems. At Vanderbilt, this has involved installing the software on a small test stand, testing the components, and investigating how to extend the packages. At FNAL, a database has been established to track computing jobs executed on the LQCD production clusters, as well as monitoring data.

JLab: The software group at JLab is carrying out research aimed at improving algorithms and producing high performance code for the study of lattice QCD. In particular, they have implemented optimized code within the Chroma software suite for the generation of gauge configurations on anisotropic lattices using improved Wilson fermions, with additional improvements under development. A new optimized multi-threading library has been developed (QMT) for Intel and AMD platforms that has less overhead than than previous optimizations, including previous custom code, Posix threads and OpenMP implementations. This work is applicable not only to the USQCD clusters, but also to supercomputers, in particular the Opteron based Cray XT3. Exploratory work has been done on level 3 and toolbox modules to facilitate use in diverse application codes without jeopardizing the flexibility needed to obtain optimal performance. Work has been started on large data stores (database and other) to support physics data analysis. In addition to software developments, improvements have been made in testing, documentation and version control for released software modules. Software support, including on-site training, has been given to other labs and universities. The lab has taken the lead in preparing the USQCD Regional Grid to be interoperable with the International Lattice Data Grid (ILDG) by addressing issues in Metadata, VO Membership, and various file services.

Boston University: As software coordinator and Chair of the Software Coordinating Committee, Richard Brower plays a critical role in leading the overall project. In addition to chairing weekly meetings via conference calls, he organized the kick-off Software Workshop at Boston University on October 27-28. James Osborn at BU is the developer of the C implementation of QDP and has continued to work with collaborators at Arizona, Indiana and Utah to integrate it into the MILC code. He has also developed a BlueGene/L specific implementation of the message passing library QMP that provides a factor of two speedup for short messages, and he has been instrumental in providing an overall speedup of the MILC code on the BlueGene/L by the same factor. Brower, Rebbi and Mike Clark, a BU postdoctoral fellow, have been working in collaboration with TOPS to study multi-grid methods for lattice QCD. Recent progress includes the demonstration that the smooth aggregation adaptive multi-grid method, when restricted to regular coarse grids, still completely eliminates all ill-conditioning in the chiral limit of the 2-d Schwinger. Regular sub-grids allows efficient data parallel codes that are currently being developed for the SU(3) 4-d Wilson Dirac operator on top of QDP.

DePaul University: DePaul University PI Massimo De Pierro is leading the effort to develop a visualization toolkit for lattice QCD. During the first 4 months he conducted an exploratory analysis, and developed a prototype application. The prototype application generates Monte Carlo configurations for a 3D Ising model and displays, in real time, the Markov Chain evolution of domains in the lattice. Dr. Di Pierro is in the process of extending the system to deal with pure gauge observables in QCD such as topological charge and energy density.

University of Arizona, Indiana University and University of Utah: Arizona, Indiana and Utah are working together to carry out a major overhaul of the MILC code to exploit the advantages of the SciDAC software. They have added three new modules to the SciDAC Level 3 QOPQDP

package, which is published on the USQCD software site. These modules are needed to optimize production code for the IBM BlueGene and commodity clusters at FNAL and JLab. Specifically, the modules compute the gauge force, the optimized multi-source Asqtad fermion force, and the Asqtad link fattening. These modules complete the suite of four Level 3 routines needed for the improved Asqtad RHMC lattice generation. The RHMC algorithm is being tuned, and currently yields a speedup of a factor of four over the older R algorithm for the lowest quark masses currently being used in production runs.

The MILC group has also tested and debugged the BlueGene-specific version of the low-lying message-passing package QMP and the linear algebra QLA suite. The former was developed by lattice SciDAC participants at Boston University and the latter was developed jointly with lattice SciDAC participants at Boston University and MIT. These new QMP-BGL and QLA-BGL packages are now ready for incorporation in production code. On a half-rack BlueGene/L with the single-mass inverter, they now expect to be able to achieve 15% of peak performance in single precision at fine production granularity (problem size per processor), and to do better at coarse granularity. They have collaborated with lattice SciDAC participants at BNL in developing two new Level 3 modules for future production running on the DOE funded QCDOC at BNL. These modules should be ready for production in the next quarter. Finally, benchmarks for individual routines have been run on the Cray XT3 and XT4 as a first step in optimizing the MILC code for these platforms.

MIT: Andrew Pochinsky of MIT is leading an effort to optimize the QCD API for the Blue Gene series of computers. He has ported the level 3 domain wall fermion code to the Blue Gene/L, and it is now included in Chroma and is publicly available in both single and double precision. He has extended the gcc compiler to utilize the Double Hummer instruction set, which enables the SciDAC project to maintain a common domain wall fermion code for SSE, AltiVec and BG/L machines. A cooperative research agreement with IBM to provide him with proprietary information to enable optimization on the Blue Gene series of computers is nearly complete, and should be finalized as soon as the DOE intellectual property attorneys agree that MIT may place all research results arising from this research in the public domain.

University of North Carolina: During the first five months of SciDAC 2, computer scientists at the University of North Caroline focused on designing and developing a web based QCD performance database – High Performance Computing (HPC) Database, and on performance profiling for the latest releases of MILC code. The HPC Database is a web based infrastructure designed to store the performance data collected by our performance team at Renaissance Computing Institute (RENCI) for QCD applications on various high performance computing systems. It also provides web interfaces for users to browse the performance data, perform statistical analysis and conduct performance comparisons. The goal is to create a knowledge base for maintaining and sharing the performance analysis results among the QCD community. The UNC computer scientists have finished the database design and the initial implementation of the system. Currently, the database contains the performance files they collected by running SvPablo instrumented MILC on several HPC systems during the year of 2006. All the performance data can be browsed via the web interface. In addition, two simple web queries are provided for fine grained data search. Recently, with partial funding from SciDAC PERI (Performance Engineering Research Institute) project, they started to integrate the HPC Database with PERI Performance Database as part of PERI's Application Engagement effort.

For performance profiling, they have been conducting hardware performance counter profiling studies of the FOR_ALL_SITES loops and the contained SU3 operations in MILC to get definitive measurements of the effects of memory bandwidth and latencies of the on-node calculations. The goal is to measure the performance ”headroom” and to identify restructuring methods, either manual or compiler-based, to improve on-node performance and to reduce the performance ”drop off” as the local domain size increases beyond the size of cache. In addition, they also started performance studies for Chroma code and an efficiency assessment of the QDP++ library by utilizing the ROSE tool developed at Lawrence Livermore National Laboratory.

UCSB: As chair of the Lattice QCD Executive Committee Robert Sugar provides overall leadership and coordination of the project. UCSB administers funds for travel not covered by grants to other participating institutions. These trips have included visits of collaboration members to participating institutions for joint work, and attendance at meetings directly related to the project. UCSB also administers travel funds for Principal Investigators S. Sharpe and R. Sugar.

Participating Institutions and Principal Investigators:

Physics:

Boston University*, Richard Brower † ‡ and Claudio Rebbi †
Brookhaven National Laboratory*, Michael Creutz † ‡
Columbia University, Norman Christ †
Fermi National Accelerator Laboratory*, Paul Mackenzie † ‡
Indiana University*, Steven Gottlieb ‡
Massachusetts Institute of Technology*, John Negele † ‡
Thomas Jefferson National Accelerator Facility*, David Richards † and William (Chip) Watson ‡
University of Arizona*, Doug Toussaint ‡
University of California, Santa Barbara*, Robert Sugar † ‡
University of Utah*, Carleton DeTar ‡
University of Washington, Stephen Sharpe †

Computer Science:

DePaul University*, Massimo DiPierro ‡
Illinois Institute of Technology*, Xian-He Sun ‡
University of North Carolina*, Daniel Reed ‡
Vanderbilt University*, Theodore Bapty ‡

* Institution receiving funding
† Project Principal Investigator, Member of Lattice QCD Executive Committee
‡ Institution Principal Investigator