Computational Resources for Lattice Gauge Theory

Lattice QCD Executive Committee

M. Creutz	BNL
N. Christ	Columbia
P. Mackenzie	Fermilab
J. Negele	MIT
C. Rebbi	Boston U.
S. Sharpe	Washington
R. Sugar	UCSB

Introduction

In the last two decades large scale computer simulations of quantum chromodynamics (QCD) have grown explosively from the initial demonstration of quark confinement into a major quantitative tool of the theoretical particle physicist. The simulations provide crucial information in support of the experimental programs in high energy and nuclear physics. However, rapid developments in Japan and Europe currently jeopardize U.S. leadership in this field. In this white paper we present a comprehensive plan for the computational resources required for the U.S. to remain a major player in lattice gauge theory.

Our committee was originally formed to develop a proposal on behalf of the U.S. lattice gauge community to the Department of Energy's Scientific Simulation Initiative. This effort was met with great enthusiasm. The thirty-nine theorists listed in the appendix have indicated an interest in working together to acquire the computational resources required by our field. They constitute almost all senior U.S. practitioners of the field, and collectively represent all major U.S. lattice gauge theory collaborations.

Scientific Objectives

The twentieth century has brought striking progress in our understanding of the fundamental structure of matter, beginning with quantum mechanics and atomic physics, and culminating with the "Standard Model" of particle interactions. The dramatic successes of quantum electrodynamics and of electroweak unification explain a vast array of physical phenomena. However traditional theoretical tools have proven frustratingly limiting for the study of quantum chromodynamics, the component of the Standard Model which describes strongly interacting particles (hadrons). The rich and complex structure of this theory arises from the interactions between quarks and gluons, which comprise most of the known mass of the universe.

Lattice gauge theory, by way of large scale numerical simulations, provides the only comprehensive method to extract, with controlled systematic errors, first-principles predictions from QCD for a wide range of important particle phenomena. In the last decade, refined numerical algorithms and dramatic progress in massively parallel computation have brought lattice simulations to a new level. Some fundamental quantities are now calculated with accuracy comparable or better than the precision attainable in experimental measurements. Prime examples are the value of the coupling constant that determines the strength of quark-gluon interactions, and the underlying masses of the quarks, most particularly the c and b quarks. We now believe we can estimate with confidence the computing resources necessary for accurate determinations of a broad range of fundamental quantities. The tera-scale facilities envisioned in our plan will enable calculations essential to make precision tests of the Standard Model, to understand nucleon structure, and to determine the properties of hadronic matter under extreme conditions.

Tests of the Standard Model are the central focus of experiments planned at all U.S. high energy physics facilities, and are essential to our search for new physics. An immediate goal is the extraction of the basic parameters of the Standard Model. Prime among these are the matrix elements describing how the weak interactions mix different species of quarks. For heavy quark species these matrix elements remain poorly known. Their determination requires combining experimental measurements with lattice calculations. In principle, experiments can provide redundant determinations of the weak interaction matrix elements, thus subjecting the Standard Model to stringent tests. If the results are not consistent, the Standard Model is incomplete and new physics remains to be uncovered. However, the possibility of performing such tests is limited by the magnitude of the experimental errors and the accuracy achieved by lattice calculations. As Table I below illustrates, the dominant uncertainties currently lie with the lattice calculations. The resources outlined below will lead to substantial improvements. We estimate that the resources proposed for 2000–2001 should reduce the lattice errors by roughly a factor of two. The tera-scale calculations envisaged for 2002–2005 should make the uncertainties in several key quantities competitive with the experimental errors.

Experimental and Lattice Errors in Selected Weak Matrix Elements

Quark Decay Process	Matrix Element	Experimental Error	Lattice Error	Computer Resources (Gigaflop/s-Years)
$t \to d$	$\|V_{td}\|^2$	4%	30–40%	11
$t \to d/t \to s$	$\|V_{td}\|^2/\|V_{ts}\|^2$	not measured	15–20%	11
CP violation in $t \to d$	ImV_{td}^2	10%	20%	6

Table I: Experimental and lattice errors in selected weak matrix elements involving the decay of a top quark (t) to a down quark (d) or a strange quark (s). The first column gives the quark decay process, the second the matrix element(s) involved in this process, the third the experimental error in the matrix element, and the fourth the lattice error. The fifth column gives the computing resources expended on the current state of the art calculation in gigaflop/s-years. (One gigaflop/s-year is the number of floating point operations that a computer which sustains one gigaflop/s will perform in one year). The resources proposed in this paper would allow the lattice errors to be reduced to the level of the experimental ones.

Recent experiments have determined a crucial ratio, ϵ'/ϵ, involving two CP violating K meson decays. In principle, lattice calculations can relate this ratio to other Standard Model parameters, providing a stringent new test of the model. The Riken/Brookhaven/Columbia group is working in this direction using a novel computational technique referred to as "domain-wall fermions." Present computing resources should enable an important first step towards the calculation of ϵ'/ϵ, but definitive results will require computers with throughput of several teraflop/s.

Experimental measurements of quark and gluon distributions in hadrons have progressed tremendously over the last decade. Lattice calculations have the potential to confront these measurements with the predictions of QCD, although thus far such efforts have been hampered by inadequate computational resources. Detailed experimental information comes from deeply inelastic lepton scattering experiments at SLAC and CERN as well as Drell-Yan processes measured at FNAL. The upcoming RHIC spin program will give further data on these subtle distributions. Electron scattering experiments at CEBAF and Bates exploring parity violating effects are providing information on the strange quarks, and future experiments will explore transitions to excited nucleon states. Multi-teraflop/s computers will enable new lattice calculations of form factors and moments of structure functions, which are crucial for comparing these experiments with theory. Furthermore, QCD calculations enable first-principles exploration of conjectured states with unusual quantum numbers. Extracting experimental signals for such particles is often difficult, and accurate calculations of masses and branching ratios will significantly aid in their discovery. These searches represent an important subfield of high energy physics, with major roles played by the BNL and CEBAF programs.

Under normal conditions quarks and gluons are confined inside hadrons. However at sufficiently high temperatures and/or pressures a new state of matter should appear, a phase best described as a plasma of unconfined quarks and gluons. This quark-gluon plasma filled the entire universe up to roughly 100 microseconds after the big bang, and it may still play a role today in the cores of neutron stars. Its observation is the primary goal of RHIC. Lattice gauge theory simulations have already provided the best estimate of the temperature (~ 2 trillion ^{0}C) at which the plasma appears, and computing resources in the teraflop/s regime will enable detailed studies of the nature of the transition and the equation of state. These calculations will provide essential input to phenomenological studies of the complex data expected from RHIC experiments.

Resource Requirements

The resources required for a lattice QCD calculation depend significantly on the quantity under study. The calculation of the strong coupling constant mentioned earlier required approximately one gigaflop/s-year. That is, a computer capable of sustaining one gigaflop/s would take one year to complete this calculation. Correspondingly a machine capable of sustaining two gigaflop/s would require six months. The sustained performance of a computer on a definite application is smaller than its peak performance, by a factor which depends on the application and the architecture of the machine. For QCD calculations this factor is typically in the range of 2 to 4. We quote sustained performances in this document because this frees our estimates from machine dependent factors. Current state-of-the-art lattice calculations use of order fifty gigaflop/s-years, while many of the most important calculations alluded to above will require several teraflop/s-years for definitive results.

In view of these facts, it is worthwhile reviewing the resources presently available to lattice gauge theorists. The bulk of these come from special purpose computers designed and built specifically for the study of lattice QCD. Norman Christ and colleagues at Columbia University have produced a succession of such machines, the latest being the QCDSP. The current hardware at Columbia sustains 120 gigaflop/s for production code, and a sister machine in the Riken Center at Brookhaven sustains 180 gigaflop/s.

For comparison, the CP-PACS collaboration, including the majority of Japanese lattice theorists, has for several years been operating a computer which sustains 300 gigaflop/s. The machine was designed by physicists and built by the Japanese computer industry. The

KEK group, also in Japan, currently has a 50 gigaflop/s machine for its own use and has a 600 gigaflop/s version of the CP-PACS computer on order. In Italy, the APE group is building a number of special purpose machines sustaining 125 gigaflop/s. Four of these computers will be installed at Italian universities early next year.

Relative to the special purpose machines, resources available to lattice gauge theorists through the national supercomputer centers are meager. The largest lattice gauge theory allocation at NERSC for FY 2000 is two gigaflop/s-years. The largest corresponding allocation at the NSF centers for the same period is ten gigaflop/s-years. The latter is for a large group with seven senior members; so, it is smaller on a per-faculty-member basis than the NERSC grant. It should be emphasized that lattice theorists are treated relatively well by the national centers. These facilities have a large number of users working in a wide variety of disciplines. All resources cannot go to a single field, nor can machines be optimized for one discipline.

These figures demonstrate that the only lattice gauge theory group in the U.S. with state of the art resources is the Columbia-Brookhaven-Riken collaboration. This group represents only 10% of the U.S. lattice community. The remaining 90% include some of the most highly respected individuals and strongest groups in the world. The U.S. cannot maintain its traditional position of leadership in this field unless this situation is corrected.

In Table II below we provide estimates, in sustained gigaflop/s, of the computing resources available to lattice theorists in various countries.

Computing Resources Available for Lattice QCD
in Sustained Gigaflop/s

Country	Current Resources	Expect Resources (In 2000)
Germany	134	264
Italy	154	554
Japan	380	980
United Kingdom	41	41
United States	320	320

Table II: Computing resources available to lattice gauge theorists in selected countries measured in sustained gigaflop/s for production QCD codes. The second column gives the resources currently available, and the third column includes machines expected to be installed in early 2000. Recall that of the 320 gigaflop/s in the United States 300 are provided by the two QCDSP machines at Brookhaven and Columbia.

Lattice gauge theorists in Europe and Japan are actively planning to acquire multi-teraflop/s-year capabilities within the next few years. The U.S. community must also do so if it is to capitalize on the exciting scientific opportunities outlined above, and thereby serve the needs of the high energy and nuclear experimental programs.

An important planning issue is whether to use commercial or self-built supercomputers. In Table III below we give estimates of the price performance ratio in dollars per megaflop/s for several commercial and specially built computers, including clusters built from commodity parts.

Cost Comparison of Computers for Lattice QCD

Computer	Cost ($/Mflops)
T3E	480
Origin 2000	190
CP-PACS	70
QCDSP	10
Indiana U. Cluster	10
Fermilab Cluster	25

Table III: The price/performance of various computers in dollars per sustained megaflop/s for QCD code. The last two lines are for clusters built during the past year from commodity parts. The costs are for hardware only, and do not include software, maintenance, or operational expenses, all of which can be considerable for commercial supercomputers. The performances were taken from benchmarks with production QCD codes. The costs for commercial machines come from a recent quotation which included a university discount. The cost of clusters varies significantly with the specific CPUs and communication hardware.

The cost effectiveness for lattice gauge theorists to build their own computers is striking. Hand-built computers take advantage of simplifying features of lattice QCD: they use regular grids and have uniform, predictable communications. In contrast, commercial machines must perform well for a wide variety of problems including those requiring irregular or adaptive grids, and non-uniform communication patterns. Thus, commercial supercomputers require considerably more expensive communication systems than are needed for lattice QCD.

Plan for 2000–2001

We divide our five-year plan into two phases, covering the years 2000–2001 and 2002–2005. The goal of the first phase is to enable three to five new 50 gigaflop/s-year calculations while simultaneously carrying out the research and development for the next phase, which aims at multiple teraflop/s-year projects.

There are three components to the hardware for this phase. First are the existing QCDSP machines, second is a cluster to be constructed at Fermilab, and third are clusters to be constructed at Jefferson Laboratory and MIT. We focus on commodity hardware and open source software to yield systems which are easily cloned and upgraded, allowing performance to follow the natural evolution of Moore's law. The current QCDSP is based on commodity processors from Texas Instruments, and the cluster projects are based on commodity processors for desktop computers.

The approximate cost of this two year program is $4,000,000 in addition to the funds already committed to the development of a follow-on machine to the QCDSP, the QCDOC. The bulk of the requested funds is for the cluster hardware, with the remainder for user support on the QCDSP and staff support for the clusters. We are convinced that this program is essential to the vitality of our field, and puts us in position for the next phase of our plan.

It is important that we follow more than a single track. The QCDSP has already demonstrated the ability to bring large computing resources to bear on specific lattice simulations. However, the number of users has been small, and it must be determined whether it and its planned successor, the QCDOC, have the flexibility to meet the needs of the entire U.S. lattice QCD community. Clusters can provide the power to generate the moderate dynamical quark lattices likely to be a focus of attention for the next two years, while providing the programming flexibility and large memories needed for the most complex calculations. The current uncertainty is whether they cost-effectively scale to the size required for the second phase of the plan.

A. The QCDSP

Over this period the Riken/Brookhaven/Columbia group will continue its exploitation of the QCDSP machines. Moreover, as a direct result of our planning deliberations, the Columbia group has formed a collaboration with the Seattle/Los Alamos/Ohio State group to calculate ϵ'/ϵ using a different algorithm, which will provide a crucial check on the calculation currently in progress. This project will provide a test of the suitability of the QCDSP and QCDOC to be used at a distance by outside groups. In a similar vein, the Riken/Brookhaven/Columbia group has made a 64 processor motherboard available to the MILC collaboration for their evaluation.

B. The Fermilab project

The construction of the cluster at Fermilab will be a joint project of the Cornell, Fermilab and MILC lattice QCD groups, and the Fermilab Computer Division, which has extensive experience in the construction and management of large PC farms. It will be used by the three lattice QCD groups to carry out a broad research program in which primary emphasis will be on B physics and other Standard Model phenomenology, but will also include the study of high temperature QCD, the mass spectrum of particles with exotic quantum numbers, and the development of improved actions. An eight-node prototype cluster has been constructed using PCI-based motherboards, the Linux operating system, Myrinet low-latency communications, and 500 MHz Pentium III CPU's. MILC collaboration code benchmarks at 80 megaflop/s per processor. A 64 node system, funded by Fermilab, is planned for early 2000. The system will be constructed to be as independent as possible of the detailed CPU and communication hardware, thus allowing rapid capitalization on future hardware developments.

Our plan recommends extending this project to a system of 512 nodes on dual processor motherboards. The cost will be \$2500 per node or less. Based on existing benchmarks, this gives a price-performance ratio of approximately \$16 per delivered megaflop/s. Pentium CPUs and Myrinet have been used as cost exemplars, but Athlon, Alpha, and G4 CPUs and VIA-based communications are also under consideration. This component of the plan is estimated to cost approximately \$1.5M for the hardware and the first year of operation.

C. The Jefferson Lab and MIT project

The Alpha-based cluster project at Jefferson Laboratory and MIT represents a collaboration involving Boston University, Jefferson Laboratory, MIT, and a few others. It also includes interactions with developers of a comparable cluster at the University of Wuppertal. These machines will concentrate on hadronic structure calculations. The project uses commodity processors and the Linux operating system. To facilitate scaling, it utilizes the highest performance nodes available. Based on exploratory studies on a four-node Alpha cluster, MIT recently purchased a 48 processor system comprised of 12 ES40 systems, each containing four 667 MHz Alpha EV6.7 processors, and connected with a 64-bit Myrinet network. Existing QCD code presently runs at 34% of the peak speed on a single node. The cluster should sustain 20 gigaflop/s when fully operational and optimized.

Our plan recommends extending this project to a 256-processor system at Jefferson Lab and an upgrade of the MIT cluster to 64 processors. This combination should provide approximately 140 gigaflop/s sustained. Although firm prices are not yet available, Compaq entered a collaboration with MIT with the expectation of offering, in the fall of 2000, a cluster at a cost of $10 per sustained megaflop/s for QCD. In addition, aggressive price competition is emerging from at least one other vendor of high performance Alpha boards. This component of the plan is estimated to cost approximately $1.5M for the hardware and the first year of operation.

D. Development of the QCDOC

The Columbia group has begun development of a follow-on to the QCDSP machines. The goal is an easily scalable massively parallel computer with a sustained performance in excess of 5 teraflop/s and a cost–performance ratio of under $1 per megaflop/s. The individual nodes will most likely be PowerPC processors combined with 4 Mbytes of memory and eight 0.5 Gbit/s communication ports, all on a single silicon chip. Thus a complete node would be a "system on a chip," providing the name "QCD On Chip" or QCDOC. The nodes will be manufactured by IBM, each being accompanied with a socket for additional off-chip memory, permitting as much as a further 1 Gbyte of memory per node. These 1–2 Watt, 1 gigaflop/s processors are joined into a four dimensional network scalable to tens of thousands of nodes. Already funded, the hardware development work on the QCDOC is an essential component of our plan.

Additional resources for user support are critical both to the new outside use of the current QCDSP machines, and to enable the larger lattice community to exploit the future QCDOC machine. Such outside use requires additional documentation for both hardware and software, as well as software support to provide the enhanced functionality and the library routines required for a broader range of physics projects. To this end our plan recommends funding for a computer scientist, software specialists, consultants and workshops for two years at an annual level of approximately $500,000.

We believe that the efforts just described form a coherent plan which will enable us to reach our goals for the period 2000-2001. The two clusters and two QCDSPs are needed to enable all the leading lattice QCD groups to undertake state of the art calculations. The two clusters will allow us to study scaling with different architectures, and the availability of the QCDSP to outside users will allow the full community to study its architecture.

Plan for 2002–2005

The objective of the second phase of our plan is for U.S. Lattice gauge theorists to carry out a broad range of precision calculations with direct bearing on the experimental programs in high energy and nuclear physics. Each such calculation will typically require multiple teraflop/s-years.

Hardware options for this phase will be some combination of the QCDOC and clusters. As indicated above, the QCDOC, currently under development at Columbia, is expected to cost less than $1 per megaflop/s, and it is anticipated that a machine which delivers at least 5 teraflop/s will be possible. Because of rapid developments, the cost and performance of clusters is difficult to predict three years in advance. Simply on the basis of Moore's law one would expect a cost of $2.50 per megaflop/s or less. However, some believe that the price-performance ratio of commodity interconnects, the main source of uncertainty in such estimates, may well fall more rapidly during the next few years than Moore's law would indicate.

Hardware decisions will be made on the basis of the experience gained in the first two years of the plan. This will include the outcome of the development work on the QCDOC and the clusters, and on developments in the commodity hardware market. We have discussed price-performance ratios throughout this document, and they are indeed important factors. Of course, what matters is the combined ability of the hardware and software to support the scientific program, and this will be the ultimate deciding factor.

We expect the total cost of our plan for the period 2002–2005 to be approximately \$8,000,000. We realize that funds of this magnitude are large on the scale of theoretical physics grants; however, as pointed out in the discussion following our presentation to HEPAP, they are very small compared to the cost of the experimental programs these calculations support. We are convinced that this level of funding is essential if the DOE is to reap the full benefits of its large investments in the experimental high energy and nuclear physics programs.

Finally, while we expect lattice calculations to produce major accomplishments over the years of our plan, we should also emphasize that anticipated progress in hardware and algorithms will make it possible to reap further important results in the outlying years. It is thus vital that the DOE plan on stable, long-term support for lattice QCD.

Members of the QCD Collaboration

T. Bhattacharya	LANL
C. Bernard	Washington U.
R. Brower	Boston U.
S. Chandrasekharan	Duke
N. Christ	Columbia
M. Creutz	Brookhaven
T. DeGrand	Colorado
C. DeTar	Utah
S.-J. Dong	Kentucky
T. Draper	Kentucky
T. Duncan	Pittsburgh
R. Edwards	Jefferson Lab
E. Eichten	FNAL
A. El Khadra	Illinois
S. Gottlieb	Indiana
R. Gupta	LANL
A. Hasenfratz	Colorado
U. Heller	Florida State
J. Hetrick	U. Pacific
G. Kilcup	OSU
J. Kuti	UC San Diego
A. Kronfeld	FNAL
F. Lee	George Washington
P. Lepage	Cornell
K.-F. Liu	Kentucky
P. Mackenzie	FNAL
R. Mawhinney	Columbia
C. Morningstar	Florida International
J. Negele	MIT
C. Rebbi	Boston U.
D. Richards	Jefferson Lab
S. Sharpe	U. Washington
J. Shigemitsu	Ohio State
D. Sinclair	Argonne
A. Soni	Brookhaven
R. Sugar	UC Santa Barbara
H. Thacker	Virginia
D. Toussaint	Arizona
U.-J. Wiese	MIT
W. Wilcox	Baylor