The SciDAC LGT Project:

Current Status and Future Directions

Outline

- Hardware and Software Status
- Future Plans
- Educational Efforts
- Current Funding and 2003 Proposal
- Outlook

Two Hardware Tracks

- **QCD On a Chip (QCDOC)**
  - **Columbia**
  - **BNL**
- **Optimized Clusters**
  - **FNAL**
  - **JLab**
- **‘It is important to pursue both special purpose machines and commodity clusters. The two tracks have different risks, and it would be reckless to cut off one.’**

  The DOE Review Panel

2003 Plans for QCDOC

- **Construct 128 Node Prototype**
  - **Initial chips expected from IBM in May**
  - **Test hardware and software**
- **Construct 1.5 Tflops Development Machine**
  - **Construction in late summer, early fall**
  - **Verify suitability to be a major community computational resource**
  - **State of the art physics calculations**
  - **To be located at Columbia University**
  - **Operated by Columbia and BNL as a national facility**
  - **Funding requested in 2003 proposal**

2003 Plans for Clusters

- **Put existing clusters into production use**
  - **48 node dual Pentium 4–Myrinet cluster at FNAL**
  - **128 node single Pentium 4–Myrinet cluster at JLab**
- **Construct next generation clusters and bring them into production**
  - **128 node dual Pentium 4–Myrinet cluster at FNAL**
  - **256 node single Pentium 4–Gigabit Ethernet cluster at JLab**
- **Research on Processors and Networks**
  - **Itanium and AMD processors**
  - **Gigabit Ethernet**
  - **Field programmable gate arrays**

ORNL Computing Resources

- **IBM SP (Eagle)**
  - **704 Power3 Processors (375 MHz)**
  - **4 Processors/Node**
  - **1.0 Tflops Peak**
- **Compaq AlphaServer SC (Falcon)**
  - **256 EV67 Processors (667 MHz)**
  - **4 Processors/Node**
  - **0.3 Tflops Peak**
- **IBM Power4 (Cheetah)**
  - **864 Power4 Processors (1.3 GHz)**
  - **32 Processors/Node**
  - **4.5 Tflops Peak**

- **We obtained over 1M cpu–hours during the past year**

QCD Applications Program Interface

- **Unified programming environment to enable high efficiency use of targeted architectures**
  - **QCDOC**
  - **Optimized clusters**

- **Three layer structure**
  - **Level 3: Highly optimized, computationally intensive subroutines**
  - **Level 2: Data parallel language to enable rapid production of efficient code**
  - **Level 1: Message passing and linear algebra routines**

2003 Software Plans

- **Bring the applications codes and execution environment to the stage where the QCDOC and clusters can be used efficiently on the core problems by all collaboration members**
- **Advance work on the QCD API**
- **Benchmark applications codes on the QCDOC and clusters**

Hardware Deployment Plans

- **2003:**
  - **1.5 Tflops QCDOC (Columbia)**
  - **Prototype clusters (FNAL/JLab)**
- **2004:**
  - **10+ Tflops QCDOC (BNL**
  - **Prototype clusters (FNAL/JLab)**
- **2005:**
  - **Of order 10 Tflops Clusters (FNAL/JLab)**
- **2006:**
  - **10+ Tflops Clusters (FNAL/JLab)**
- **2007 and Beyond:**
  - **Expand scientific reach by constant investments**

Educational Activities

- Presentation to HENP Staff
- Presentations to HEPAP and NSAC
- Presentation to Ray Orbach
- Talking points
  - Importance of our calculations to the experimental programs in high energy and nuclear physics
  - United community
  - Major partners and collaborators
  - Need long term, steady state funding

Ray Orbach's Reaction

Orbach also called for a major effort to improve US computing capabilities... "To find ourselves second on an international scale is a national disaster," he said.

The study of [QCD] on a lattice is an excellent way to develop leading–edge computers, Orbach maintained. "QCD simulation leads the way," he said, calling it "as fundamental an exercise as experiment and theory."

APS News – July, 2002

Current Grant and Pending Proposal

- **SciDAC Grant**
  - $2M per year for three years, starting July, 2001
  - Supports software development and cluster prototyping
- **2003 Proposal**

Item	Cost
1.5 Teraflops QCDOC	$1500.0
QCDOC Support Staff	500.0
Project Manager	211.2
Other Staff	135.6
Total	$2346.8

All numbers are given in thousands of dollars.

Panel Review

- **Scientific Merit:** Outstanding
  - There is intrinsic merit in understanding and validating QCD
  - The coupling with the experimental programs in high energy and nuclear physics is very important
- **Computational Science Interest:** Substantial
  - The hardware project will provide an important demonstration of the role of special purpose computers
  - The software project is an important example of making special purpose computers more broadly useful.
- **World Class Facility:** Yes
  - Connection to IBM and Blue Gene/L
  - Development of the QCD API

Additional Recommendations of the Panel

- **Leverage this project with a fellowship program to attract young people into the field**
- **Do not be too conservative in funding the project**
- **Think of current work as the early stage of a long term project**

Milestones for 1.5 Tflops QCDOC

- **Construction criteria for 10+ Tflops QCDOC**
  1. **Availability and performance of applications codes**
  2. **Availability and reliability of hardware**
  3. **Functionality of operating system**

- **Other Major Milestones**
  4. **Convenient and robust user environment**
  5. **High performance with level three inverters**

- **Achievement of the first three milestones is expected within two months of completion of the 1.5 Tflops machine**

- **Milestones 4. and 5. should be achieved three and six months after completion of the 1.5 Tflops machine, respectively**

Availability and Performance of Application Code

Programs implementing the Hybrid Monte Carlo algorithm to evolve full QCD lattices using staggered, asqtad, Wilson, clover Wilson and domain wall fermions should be available. With the full machine configured as a single partition, evolutions with each of these actions should achieve at least 20% efficiency (0.6 Tflops) with a local volume of 2^4 sites per node (12×16^3 lattice volume) for light quark masses. The ability to run general community software should be demonstrated by showing that standard MILC code using the QMP interface, but without specific optimization for QCDOC, achieves at least 10% efficiency (0.3 Tflops) on an 8^4 local volume (48×64^3 lattice volume). Similarly, MILC code modified to call level-3 inverters should run with at least 20% efficiency (0.6 Tflops) on a 2^4 local volume (12×16^3 lattice volume)

Availability and Reliability of Hardware

The full machine should run for a one week period with no more than one day lost in total for hardware maintenance and debugging on a combination of the application codes described above. (That is, at least an integrated 6 days of physics production running should be achieved. Uptime between the last checkpoint and a machine failure is not counted as production running.) Diagnostic support from the operating system should be sufficient that faults can be diagnosed and repaired without reference to the application code being run. During this week, 50% of the computer time should be spent on reproducibility checks, to verify that there are no undetected numerical errors.

Functionality of the Operating System

Applications running on a node of QCDOC should have a run time environment with support for the standard C/C++ library, standard UNIX I/O routines and QMP. In particular, applications must be able to read and write files on the host computer. The bandwidth to the host computer should be sufficient to permit a double precision $32^3 \times 64$ lattice (1.2 GBytes) to be loaded or unloaded from the host in less than 1 minute (20 Mbytes/sec). The host computer should provide a UNIX-like environment for users to load and run programs.

Convenient and Robust User Environment

A convenient and robust user environment must be provided. This will include the following functionality:

- **A new user can easily register, log on and move code and data to the facility. Procedures are web documented and yield a new account in 3 business days. The ssh and scp protocols are supported and 0.5 Gbyte of RAID disk space per user is routinely provided and backed up weekly. Much more space will be provided as required for particular projects.**

- **SciDAC specified file I/O is implemented allowing files created on a cluster to be used on QCDOC and vice-versa.**

- **MILC and SciDAC QDP (C and C++) codes compile easily.**

- **LQCD SciDAC batch protocols are supported and provide a uniform batch environment across both QCDOC and clusters. The batch queue and status of jobs can be viewed. Batch log files, stderr and stdout files are readily available and there is a defined method (automatic or administrative procedure) to ensure that time allocations work.**

- **An adequate level of personal, user support is provided. This support will be provided uniformly to both cluster and QCDOC users with the actual implementation and delivery shared between BNL, FNAL and JLab.**

Performance with Level Three Inverters

CPS, MILC and QDP++ software, using level-3 inverters should achieve 40% efficiency for full QCD hybrid Monte Carlo evolution on 2^4 local volumes per node using Wilson fermions and with somewhat reduced performance, 30% for staggered (naive or asqtad) fermions or 1.2 and 0.9 Tflops respectively.

Outlook

- Strong support from DOE
- Strong support from Laboratory partners
- Strong support from experimental community
- Excellent review of 2003 proposal
- DOE budget prospects for 2003 and 2004 are not strong
- Japanese Earth Simulator may improve prospects for computational initiatives
- Work already beginning on 2005 budget
- Imperative that we work together as a community