Computational Requirements for Lattice QCD

Steven Gottlieb, Indiana University

sg at indiana.edu

Capacity vs. Capability

- Capacity: total computational throughput (perhaps on many jobs)
- Capability: ability to produce leading performance on particular job
- Our research program is never limited to just one job, so we need total capacity at some reasonably large capability
- Our rate of scientific advancement is limited by available computational resources, so we must obtain the most cost effective hardware to maximize our scientific output

Analysis vs. Generation

Two types of jobs:
a) generation of configurations,
b) analysis of configurations

- Generation is a long stochastic evolution that is done in just a few streams (varying lattice spacing, quark mass). Computational need is sensitive to these parameters. Capability is limitation here.
- Analysis jobs are run on stored configurations. As a generation run proceeds, the configs are archived and **hundreds** of analysis jobs can be run in parallel (or quick succession).

Further Aspects of Analysis

- Multiple groups will analyze the same configurations for different physics projects. The coordination of configuration generation across the entire U.S. community is needed to make this possible. It will lead to major savings.
- The capability requirements for analysis vary widely. Some projects require capability similar to configuration generation, while others require much lower capability. The latter projects are best suited for clusters.
- Analysis projects tend to put more stress on I/O and storage than configuration generation.
- Multiple computers are valuable for analysis.

Computational Challenges

- Calculations must be carried out for several (small) lattice spacings in order to perform extrapolations to the continuum limit.
- It is too computationally expensive to perform simulations at the physical masses of the two lightest quarks ($\mathbf{u}$ and $\mathbf{d}$). So, we work with a range of light quark masses and perform extrapolations to their physical values using chiral perturbation theory.
- It is necessary to increase the physical size of the box in which the simulations are performed as the light quark masses are decreased to avoid finite size errors. Thus, as the lattice spacing and quark masses are decreased, the number of lattice points must be increased.

A Multiyear Program

- It will take several years to decrease the lattice spacing and to approach the chiral limit for each lattice spacing
- A capability of 5 TF is sufficient for all but the two most demanding runs below

a(fm)	m_l/m_s	Lattice	Traj.	TF-Yr
0.09	0.1	$40^3 \times 96$	3,000	0.54
0.09	0.05	$56^3 \times 96$	4,200	6.05
0.06	0.4	$48^3 \times 144$	3,000	0.45
0.06	0.2	$48^3 \times 144$	3,750	1.68
0.06	0.1	$60^3 \times 144$	4,500	7.98
0.06	0.05	$84^3 \times 144$	6,300	93.20
0.045	0.4	$56^3 \times 192$	4,000	2.25
0.045	0.2	$56^3 \times 192$	5,000	7.52
0.045	0.1	$80^3 \times 192$	6,000	54.80

Recent Allocations

- The Scientific Program Committee allocates time on the QCDOC and the prototype clusters built under the SciDAC grant. It will also allocate time on proposed computers.
- Requests for time in the most recent allocation process exceeded what is available, even given that people were aware of current capacity.
- Table shows requests in millions of node-hours

	QCDOC	FNAL	JLAB
available	91.5	4.8	5.7
requested	142.2	6.2	13.0

- Following slides contain more details of a limited number of large requests/allocations

Needs of Upcoming Runs: Asqtad

a(fm)	$\mathbf{m_l/m_s}$	Lattice	Traj.	TF-Yr	QCDOC $\mathbf{10^6}$ Node-Hr
0.09	0.1	$\mathbf{40^3 \times 96}$	2,000	0.36	11.3
0.06	0.4	$\mathbf{48^3 \times 144}$	3,000	0.45	14.1
0.06	0.2	$\mathbf{48^3 \times 144}$	1,875	0.84	26.4
Totals				1.65	51.8

Resources to generate gauge configurations with dynamical improved staggered quarks on the QCDOC in coming year. Assumed 280 MF/node achieved speed.

Needs of Upcoming Runs: DWF

a(fm)	m_l/m_s	Lattice	Traj.	TF-Yr	QCDOC 10^6 Node-Hr
0.11	0.4	$24^3 \times 64$	5,919	0.37	12.1
0.11	0.3	$24^3 \times 64$	6,835	0.81	26.5
0.11	0.2	$24^3 \times 64$	8,371	3.44	112.0
0.11	0.3	$32^3 \times 64$	6,835	3.19	104.0
0.083	0.5	$32^3 \times 64$	7,059	1.65	53.7
Totals				9.46	308

Resources to generate the DWF gauge configurations on the QCDOC in the coming two years. Assumed 270 MF/node achieved speed to get TF-yr.

Analysis Capacity Needed

a(fm)	m_l/m_s	Config.	QCDOC Node-Hr	FNAL P4E Node-Hr
0.09	0.1	333	500,000	175,000
0.06	0.4	500	840,000	294,000
0.06	0.2	250	840,000	294,000
Totals		1,083	2,180,000	763,000

Resources to run the spectrum code on the proposed lattices. The third column gives the number of configurations to be analyzed, the fourth column the node-hours required on the QCDOC, and the fifth column the node-hours required on the current FNAL P4 cluster with 2.8 GHz cpu and 800 MHz FSB.

a(fm)	$\mathbf{m_l/m_s}$	Config.	QCDOC Node-Hr	FNAL P4E Node-Hr
0.09	0.1	333	500,000	175,000
0.06	0.4	500	2,000,000	700,000
0.06	0.2	250	1,000,000	350,000
Totals		1,083	3,500,000	1,225,000

Resources to run the analysis code for light pseudoscalar mesons on the proposed lattices. The fourth column gives the node-hours required on the QCDOC, and the fifth column those required on the FNAL P4E cluster.

- 4.5 million node-hrs allocated at Fermilab for onium and heavy-light meson analysis
- 4.3 million node-hrs allocated at Jlab for three projects using Domain Wall quarks

I/O and storage requirements

estimate of storage and I/O for heavy-light analysis

size	gauge (GB)	light (GB)	heavy (GB)	job (GB)	config	ensemb (TB)
$20^3 \times 64$	0.15	0.04	0.59	15.6	2075	32
$28^3 \times 96$	0.61	0.15	2.43	64.3	1600	103
$40^3 \times 96$	1.77	0.44	7.08	187.6	500	94
$48^3 \times 144$	4.59	1.15	18.35	486.2	750	365
Totals						594

Disk storage requirement will be job size times (number of jobs running in parallel plus a few for staging)

Storage requirements at JLAB this year about 100 TB (estimated by G. Fleming).

International Resources for LQCD

- UKQCD: 5 TF QCDOC
- Japan:
  - Tsukuba: 2048 node cluster (early 2006) growing to 3072 or more nodes the next year. (2.8 GHz Xeon CPU; 12–17 TF peak)
  - KEK: New 20 TF (peak speed) system to replace Hitachi SR800F1, early 2006
  - Earth Simulator: Some fraction of this 40 TF machine will be used for LGT
- Germany:
  - Bielefeld: 5TF APEnext, end of this year
  - DESY: 2.4 TF APEnext, after July 2005
- Italy: 10.4 TF (peak) APEnext, July–Dec. 2005

Concluding Remarks

- The QCDOC is an excellent vehicle for creating state of the art configurations.
- Beginning with the large FNAL machine scheduled for FY 2006, clusters will play an increasingly important role in configuration generation.
- Some analysis jobs can easily be done with either the QCDOC or clusters.
- Some analysis jobs, particularly those that require large amounts of I/O, will be more easily done on clusters than on the QCDOC. Clusters already rival the QCDOC in cost effectiveness for analysis jobs, and will surpass it for such jobs in FY 2006.
- Investments by Italy, Germany and, particularly, Japan rival what is currently proposed by DOE and could result in systems that exceed the capability of our own.