Plans for QCDgrid

Craig McNeile

mcneileamtp.liv.ac.uk

University of Liverpool

Why I “work” on QCDgrid?

The goal of studying the properties of quarks in HEP experiments requires non-perturbative QCD matrix elements to be computed. There are two types of paper that you want to be on:

- The first one on a subject. Flexible computing and archived data is important.
- The last one on a subject. This is an extensive calculation with many data sets.

One goal of the QCDgrid project is to correctly annotate and back up our data. Another aim is to create a common interface to running on the

High Precision Promises (SciDac)

Summary Table

Measurement	CKM Matrix Element	Hadronic Matrix Element	Non-Lattice Errors	Current Lattice Errors	Lattice Errors 0.6 TF-Yr MILC0	Lattice Errors 6.0 TF-Yr MILC1	Lattice Errors 60. TF-Yr MILC2/ DWF1
ε_K ($\bar{K}K$ mixing)	$\mathrm{Im}\, V_{td}^2$	$\hat{B}_K$	10%	20%	12%	5%	3%
ΔM_d ($\bar{B}B$ mixing)	$\lvert V_{td}\rvert^2$	$f_{B_d}^2 B_{B_d}$	6%	30%	16%–26%	8%–10%	6%–8%
$\Delta M_d / \Delta M_s$	$\lvert V_{td}/V_{ts}\rvert^2$	ξ^2	—	12%	8%	6%	3%–4%
$B \to \pi l \nu$	$\lvert V_{ub}\rvert^2$	$\langle \pi \rvert (V-A)_\mu \lvert B\rangle$	7%	15%	10%–13%	5.5%–6.5%	4%–5%
$B \to \binom{D}{D^*} l \nu$	$\lvert V_{cb}\rvert^2$	$\lvert \mathcal{F}_{B \to \binom{D}{D^*} l \nu}\rvert^2$	2%	4.4%	3%–4%	1.8%–2%	1%–1.4%

(Up to minor modifications, this table is same as that presented to HEPAP by Bob Sugar.)

What do QCD calculations involve?

In lattice QCD calculations QCD is solved by introducing a space-time lattice and putting the equations on a computer. The equations that need to be solved is a path integral, (essentially a (very) large multi-dimensional integral).

$$< F > \propto \int \mathcal{D}U F[M^{-1}] e^{-S_i} det(M)$$

A Monte Carlo process is used to solve this integral.

QCDgrid is used!

James Perry will talk about the technical aspects of QCDgrid later today.

When I run jobs on our cluster I use

```
system(‘‘cd $dir ; get-file-from-qcdgri
NF2/BETA52/CLOVER202/V16X32/KAPPA3550/G
‘‘.${tarfile}.’’ ‘‘);
```

(before I would have to manually copy data files).

The data is now backed up automatically at different locations.

Compute grid

The compute grid part of QCDgrid has been installed at the sites. The aim of this is to have a common interface to all platforms. The user chooses the machine and provides the executable.

This is an example I tried:

```
qcdgrid-job-submit doorstopper.epcc.ed.
input.xml stag_dyn
```

Data can be archived and obtained from the QCDgrid facility.

I don’t believe anyone is using this system at the moment for production runs.

Improving QCDgrid

- We are using globus 2.4. Things seem to move very fast in the globus world and backwards compatibility is not guaranteed.
- We are getting a new supercomputer (QCDOC). There have been some delays on this project, so we have not been generating new configurations. So the QCDgrid needs to be stress tested.

Use of XML

In the past we have used binary data files to store data with some metadata in the filename and directory structure. For example:

NF2/B52/C202/V16X32/K3500/Meson/
D52C202K3500U014010_LL3450X_FL3400X_CMe

We made a decision to use XML to store metadata and some data. (XML is not a good choice for large amounts of data).

XML in production codes

Balint Joo (Edinburgh) and Robert Edwards (JLAB) have written an IO package based on libxml. http://www.jlab.org/ edwards/qdp/

```
int number;
read(fromxml,''/someData/lattice'',numb
QDPIO::cout << ''found number = '' << n
multi1d<int> arrayInt(3);
for(int i=0; i < arrayInt.size(); ++i)
  arrayInt[i] = i+37;
write(toxml,''arrayInt'',arrayInt);
```

Perhaps, data binding might be better.

Lattice collaborations are getting bigg

There was a round table meeting at lattice 2002 on the cost of lattice QCD calculations. To produce unquenched configurations requires at least 100 Tflop years. UKQCD's new supercomputer, the QCDOC, will have the performance of 5 Tflop sustained.

The MILC and other US collaborations have put their gauge configurations on a web server (http://qcd.nersc.gov/) at NERSC that is freely accessible and widely used.

A Japanese collaboration has just set up its own web server (http://www.lqa.rccp.tsukuba.ac.jp/).

Not everyone is so public spirited

(ILDG) International Lattice Data Gr

The ILDG was dreamt up by UKQCD, but has been enthusiastically supported by US and Japanese groups. The ILDG web site is http://www.lqcd.org/ildg There has been 4 video meetings on coordinating lattice QCD production runs since December 2002. There have been reports at the recent lattice conferences.
The ILDG vision

- To develop an international datagrid for the lattice field theory community.
- To develop an XML Schema suitable for describing the data generated by lattice

QCDML

A XML schema is like a struct in c or c++. It is one way of enforcing consistency of XML documents. (This is important as we wish to transfer data between collaborations). QCDML was developed, using the XML spy editor by Chris Maynard at Edinburgh, for metadata for gauge configurations. The QCDML proposal was then submitted to an international working group within ILDG. After two years a standard is almost agreed. The code to write out the QCDML from our production codes is mostly written. What we plan to do in gridpp2 is to catalogue the data from the production code in QCDgrid.

A fragment of QCDML

```
   <algorithm>
    <parameters>
       <name>targetResidue</name>
       <value>1e-07</value>
    </parameters>
    </algorithm>
    <precision>single</precision>
    <markovStep>
<markovChainLFN>www.lqcd.org/ildg/ukqcc
<series>1</series>
<update>010170</update>
<avePlaquette>0.53380336E+00</avePlaque
<dataLFN>D52C202K3500U010170</dataLFN>
```

The web form interface

The common parameters to a run.

```
<markovChainLFN>www.lqcd.org/ildg/ukqcc
<plaquetteGluonAction>
<glossary>www.lqcd.org/ildg/plaquetteGl
<gluonField>
<gaugeGroup>SU(3)</gaugeGroup>
<boundaryCondition>
<elem>periodic</elem>
</boundaryCondition>
</gluonField>
<couplings>
<beta>5.2</beta>
```

Bad things about XML

I believe that the majority of UKQCD supports the use of XML, but there are some critics (unbelievers).
We have only used XML in one physics run by the Liverpool graduate student Steve Miller. So not too much practical experience with huge data sets and lots of users.
Even though XML is so hyped the open source tool support is quite weak. Documents can become big without any increase in information.
Some curve fitting codes are in fortran (use a wrapper cfortran.h).

The ILDG lattice repository

There is a middleware working group with in the ILDG. The aim is to provide a single access point to different lattice collaborations public data. The access will be via a web page (and something I can call from a perl script).
The public repositories of the different lattice groups will be linked into a single virtual lattice data repository via web services.
This activity will be a major part of the gridpp2 effort. Work has already started, but many people involved have other commitments, so the dedicated effort will help a lot.

Conclusions

If you read the popular science press, then you will have seen claims, that lattice QCD is going to compute a small subset of quantities with high precision (errors of under 5 %) in unquenched QCD.

This requires better data handling facilities and cooperation between lattice collaborations. (The process of doing lattice QCD calculations needs to be reengineered). So we need QCDgrid facilities attached to a supercomputer.