Tier 2 requirements of UKQCD for GridPP3

Chris Maynard *c.maynard@ed.ac.uk*
EPCC, University of Edinburgh

September 14, 2007

UKQCD is a collaboration of around 60 physicists from 7 British Universities, Cambridge, Edinburgh, Glasgow, Oxford, Liverpool, Southampton and Swansea. It has a broad science programme exploiting lattice field theory and numerical simulations applied to QCD. It is focused on problems of central interest in particle physics phenomenology: reducing theoretical uncertainties in the determination of Standard Model parameters, constraining models of New Physics beyond and deepening our understanding of QCD itself. The distinctive feature of current lattice QCD simulations is the inclusion of the u, d and s quarks. The outstanding systematic issue now being addressed is to simulate with the lightest quark masses close to their physical values, the aim being to calculate with comparable precision to that achieved hitherto in quenched simulations.

UKQCD has submitted a bid to the STFC call for high performance computing for support to continue running the special purpose QCDOC machine at Edinburgh, substantial BlueGene P resources and other linux cluster type resources. These machines will produce UKQCD primary data, called gauge configurations. UKQCD uses its distributed data storage grid system QCDgrid (in future DiGS) to store its gauge configurations.

Secondary data, known as hadronic correlators are then computed on the primary data, either on the same systems described above, or on cluster type machines. The secondary data is also stored on the DiGS system. The DiGS software is built on top of globus to replicate the data so that multiple copies are stored. Data access is via the metadata catalogue.

UKQCD has requested 300TB of tape initially rising to 1200TB in the fifth year at the tier one centre. In addition, disk space has been requested at 10% of the tape capacity at the tier one, and a further 25% at the tier two centres. In the DiGS development plan is a work package to alter the software so that it functions with SRM so that it can access the tape storage.

DiGS also has a job submission element. This could be used or adapted to work with job submission /work flow systems on the tier 2 centres.

UKQCD has requested 200 kSI2k unit of CPU in the first year rising to 300 by the fifth year. For the hadron correlator calculations the memory per core, and the interconnect are also very important. UKQCD has identified two different requirements for these calculations.

If memory per core is large enough, then the data structures (specifically objects called quark propagators) will fit in memory, and the calculation can be done in serial. Different processors can perform the same calculation on different gauge configurations, i.e. a task farm. For such calculations a minimum of 2Gbytes per core is required but 4 Gbytes would be highly desirable.

In the second type of calculation massively parallel processing using a message passing protocol such as MPI is required. For these calculations the memory requirements per core are less stringent as the quark propagators are spread out across many cores. To be effective the message passing needs to be done over a fast, low latency interconnect such as infinband, involving many (in this context 64-128) cores.

To access UKQCD data, the tier two centres would need to have a globus installation, and have the DiGS software set up as a client. The latter requirement is fairly lightweight. Gauge configurations could then be pulled from the grid, the hadron correlators calculated and submitted back to the DiGS system. This could be done using whatever job submission systems are installed at the tier 2 centres directly, or possibly indirectly using DiGS job submission package. This would use globus, to talk to the tier 2 centre job submission system. As a last resort direct logins could be used. UKQCD expects to be producing some files greater in size than the 2Gbyte limit imposed by 32 bit operating systems, but increasingly many systems are 64 bit this is unlikely to be an issue.

In summary UKQCD's storage requirements are detailed above, and its compute model outlined. To make effective use of GridPP3 compute nodes, UKQCD would require access to two different types of cluster. Firstly, a high memory per core cluster, with no real interconnect requirements, and secondly a group of cores with fast, low latency, interconnect, with no large memory requirements. Both sorts of cluster would need to be able to read and write to the QCDgrid system.