UKQCD Grid Support Review 2005

Project Title: QCDgrid

Document Title: UKQCD Grid Support Review 2005

Document Identifier: QCDGRID2-SUPPORT2005-1.0.1

Document Filename: D1.5.2_support_review_2005.doc

Distribution Classification: Public

Authorship: MGB, DJB, JTP

Approval List: QCDgrid Development Team

Distribution List: UKQCD Collaboration

Document History:

Personnel	*Date*	*Summary*	*Version*
MGB, DJB, JTP	8/FEB/06	Minor revisions to original version.	1.0.1
MGB, DJB, JTP	3/FEB/06	First release.	1.0

Contents

1 Introduction 3
2 The UKQCD Grid 4
2.1 Hardware – the grid infrastructure 4
2.2 Management and client software 4
2.3 Grid users 4
3 Types of support 5
3.1 Routine operations 5
3.1.1 X.509 certificate management 5
3.1.2 New users and Grid Map file maintenance 5
3.1.3 Firewall configuration 5
3.1.4 Time synchronisation of hosting systems 5
3.1.5 Patching and software upgrades 6
3.2 Other maintenance tasks 6
3.2.1 Recovery from hosting system failure 6
3.2.2 Recovery from exhaustion of disk space 6
3.3 System monitoring 6
4 Review of 2004/2005 8
4.1 Improvements 8
4.1.1 Problem Report mechanism 8
4.1.2 Formal versioning of QCDgrid 8
4.1.3 The QCDgrid Test System 8
4.2 Notable problems and system failures 9
4.2.1 Exhaustion of the control node 9
4.2.2 Proxy certificate expiration for Control Node process 9
4.2.3 Hardware failure of the control node 9
4.2.4 Deployment of a new storage node (QCDOCx) 9
5 Future plans and proposals 10
5.1 Replacement of the control node 10
5.2 Relocation of the backup node 10
5.3 Delivery of a QCDgrid System Administrator course 10
5.4 EDG Virtual Organisation software 10
5.5 Time synchronisation across hosting systems 10
5.6 Automated monitoring of the UKQCD Grid 11
5.7 Better diagnosis of firewall issues 11
5.8 Upgrade of third-party software components 11
6 References 12

1 Introduction

The UKQCD Collaboration aims to “procure and jointly exploit computing facilities for lattice field theory calculations whose primary aim is to increase the predictive power of the Standard Model of elementary particle interactions through numerical simulation of Quantum Chromodynamics” [1]. Such numerical simulations produce large amounts of data in the form of binary files. The first QCDgrid project developed a software suite running on dedicated hardware (the UKQCD Grid system) that allows UKQCD members to store and share data amongst one another in a straightforward and reliable manner. Metadata in the form of XML files can also be stored on the system. A graphical data browser tool was developed to facilitate ease of use. The software suite also allows users to initiate simulations and post-processing jobs on machines on the grid.

This document provides a review of the support-related activities of the project, during the period from September 2004 to the end of 2005. It examines significant events and issues, proposing modifications to the support process in order to improve future performance and streamline the effort expenditure of the QCDgrid project.

2 The UKQCD Grid

Within this report, the UKQCD Grid is considered to consist of three different entities:

- Hardware – the physical infrastructure that underpins the grid.
- Software – the management and client software that provides grid functionalities.
- Users – the users that perform science using the grid.

Each of these categories is defined, in turn, below.

2.1 Hardware – the grid infrastructure

At the time of writing, the UKQCD Grid infrastructure consists of:

- six storage sites[1] (at Columbia University, the University of Edinburgh, the University of Liverpool, Rutherford Appleton Laboratories, the University of Southampton, and the University of Swansea);
- one control site, at Liverpool University;
- one backup node, at Swansea University;
- client machines distributed across the UK academic network and beyond.

Within this document, storage, control and backup nodes will be collectively referred to as hosting systems[2]. A list of UKQCD hosting systems is maintained at http://www.ph.ed.ac.uk/ukqcd/collaboration/Grid/QCDgrid_nodes.html.

The task of maintaining both hosting systems and client machines is generally the responsibility of local support staff at each individual site. Maintenance includes basic operation of the system (hardware, network connectivity/firewalls, operating system), as well as provision of core Grid software as required to operate the UKQCD Grid (for example: Globus and EDG tools).

2.2 Management and client software

The software that provides bespoke functionality to the UKQCD Grid is encapsulated within a software suite called QCDgrid. This software is the responsibility of the QCDgrid project. New versions of QCDgrid, containing extended functionality, feature enhancements, and bug fixes are periodically released by the project team and posted to the NeSCForge website https://forge.nesc.ac.uk/projects/qcdgrid/. Both user and administrator documentation is distributed along with project software.

Once released, each new version of QCDgrid needs to be deployed onto the UKQCD infrastructure. This typically requires a software update/replacement to be performed on each storage site, the control node, the backup node, and any client machines that connect to/interact with the grid. Historically, software updates on hosting systems have been provided using a combination of effort from local administrative staff and QCDgrid project team members.

2.3 Grid users

Scientists from UKQCD typically connect to the grid via a hosting system or via a local installation of the QCDgrid client tools on individual workstation. Users who work with local copies of QCDgrid client tools are typically responsible for the maintenance and upkeep of this software. Users who connect via a hosting system rely on the particular arrangements of the hosting institution for maintenance of client software.

[1] A seventh storage node has been configured at the University of Edinburgh Advanced Computing Facility (the ACF). This is currently undergoing testing prior to being enabled as a real storage facility.

[2] Within this document, the University of Edinburgh ACF is considered as a storage node.

3 Types of support

3.1 Routine operations

As with any grid-based environment, there is a range of on-going maintenance that needs to be carried out to ensure the continued security, reliability, and integrity of the services provided.

3.1.1 X.509 certificate management

The UKQCD Grid requires X.509 certificate authentication for both users and systems that participate in the infrastructure. Only certificates are issued by the UK e-Science Certificate Authority (CA) are currently accepted. Typically, machine and user certificates issued by this CA have a one year lifetime. Towards the end of the lifetime of a certificate, the owner of the certificate is automatically warned of the approaching expiration date and provided with instructions on how to renew the certificate. If a certificate expires, then the user/machine that uses this certificate will no longer be able to contribute to the grid infrastructure.

3.1.2 New users and Grid Map file maintenance

New users are added to the UKQCD Grid via a semi-automated process. The majority of hosting sites run the EDG Virtual Organisation (VO) software. This allows a new user certificate *subject* to be added to a central VO server (hosted at University of Edinburgh) and then automatically disseminated to other sites running the EDG VO (within a 24 hour period).

The Rutherford Appleton Laboratories host does not currently run the EDG VO software. New user certificate subjects need to be manually added to this sites. This is typically initiated by a member of the QCDgrid team, who e-mails new certificate subjects to the local site administrator at the remote site.

3.1.3 Firewall configuration

Firewalls are put in place, generally by local system administrators, to protect their systems from unsolicited connections or attacks. A firewall typically restricts communications that pass through it to those originating from specific sets of IP addresses and machine port ranges. The UKQCD Grid uses Globus Toolkit Version 2 as middleware for the infrastructure. Within this environment, systems communicate over a range of non-standard ports, the use of which typically requires a non-default firewall configuration to be employed at all hosting sites.

To operate correctly, each hosting system on the UKQCD Grid must have access to all other hosting systems, through any firewalls, on the port range configured for Globus communications. A firewall that is configured incorrectly, in such a way as to block legitimate UKQCD Grid traffic, is likely to generate difficult to diagnose and unexpected behaviour.

3.1.4 Time synchronisation of hosting systems

Each individual host system has an internal clock. Due to manufacturing discrepancies between the calibration of different clocks, the local time recorded on different systems will tend to *drift* in relation to each other. For a correctly functioning grid, the hosting systems have to agree on the current time to a precision of approximately 1 minute[3]. For larger discrepancies, a “clock skew” problem is typically observed, manifest as a user or machine proxy certificate being rejected by a host.

Within the current UKQCD Grid infrastructure, significant drift has been observed over a timescale of weeks. This generates clock skew problems that require occasional, manual re-synchronisation of individual hosts. This is a manual step that can only be performed by a person with administrative privileges on the particular machine.

[3] This is the maximum time skew permitted by the Globus Toolkit specification.

3.1.5 Patching and software upgrades

In addition to the QCDgrid, the UKQCD Grid relies on a number of third-party software packages:

- Globus Toolkit;
- eXist database;
- EDG Virtual Organisation tools;
- Virtual Data Toolkit (VDT).

Periodically, updates to these packages are made available by the distributors, providing functionality updates and bug fixes. Currently, the UKQCD Grid has only an informal process for evaluating and adopting new versions of third-party tools. This is done on an "as needed" basis.

3.2 Other maintenance tasks

In addition to the above, routine tasks, other maintenance may be required from time-to-time in response to unpredictable events.

3.2.1 Recovery from hosting system failure

On occasion, a hosting system may fail (for example, crash), leading to a range of predictable situations:

- For storage systems, failure will temporarily suspend the storage capability of the particular node. Recovery typically involves a reboot of the machine, since the QCDgrid software automatically validates and manages most data integrity issues. If a system persistently fails, then it may imply a more significant problem with the hardware/software make-up of the host. This will typically necessitate investigation by local administrative staff.

  If a storage system is down for a period of time, then data that is hosted on the facility will be duplicated by the control node to other storage facilities.
- For the control node, failure will switch the grid to read-only mode. The backup node should automatically take over from the control node, providing read-only access to data on the grid.
- For the backup node, failure will typically have no effect on the grid, providing that the control node continues to function correctly. However, one should bear in mind that the backup node will typically perform in another role, such as a storage node.

3.2.2 Recovery from exhaustion of disk space

The QCDgrid software selects a suitable storage node for new data based on a combination of proximity to the client and free capacity. If a particular storage site runs low on free capacity, then the QCDgrid software will attempt to release space by deleting redundant copies of data. Furthermore, the Control Node will e-mail a designated UKQCD Grid controller to warn them of the low capacity situation.

Storage capacity at an individual site should never be completely exhausted. However, the QCDgrid software cannot account for disk usage beyond its control. If capacity of a storage node is completely exhausted – due to the effects of a third-party process – the outcome is currently unclear.

3.3 System monitoring

A lightweight monitoring function has been implemented in the current QCDgrid software. This is performed by the control node, and involves a range of periodic checks:

- The control node will test hosting systems (using Globus GRAM) to confirm that they are responding to requests. If a node fails to respond to a request, then an e-mail is generated and sent to the designated UKQCD Grid controller.
- The control node will compare replica copies of individual files to check that they are in agreement (that is, have the same MD5 checksum value). If an inconsistency is found, then the UKQCD Grid is halted and an e-mail alert is sent to the UKQCD Grid controller.

- The control node will check that there is at least two copies of each files on the grid. If only one copy of a file is found, then a replica is created at another storage site.

Other monitoring processes are more ad hoc and depend on users to raise issues or behavioural anomalies that they observe using the NeSCForge Problem Report pages [2]. Details of how Problem Reports to NeSCForge are handled is provided in [3].

4 Review of 2004/2005

In this section, we look at specific events that have occurred during the review period, summarising the outcome of the event. The items noted in this section form the basis for the action proposed in Section 5.

4.1 Improvements

During 2005, a number of improvements have been made to the support process applied to the UKQCD Grid. These improvements have been implemented in order to sustain reliable operation of the UKQCD Grid and establish a predictable reaction to exceptional issues and failures. The improvements are described, in turn, below.

4.1.1 Problem Report mechanism

An issue tracking process has been established that leverages off the NeSCForge Problem Report mechanism [2] and follows a process defined in [3]. The process is designed:

- to improve the traceability of problem reports;
- to provide a single point of contact between users and the development/support team;
- to provide users with feedback on problems that they report;
- to document the process initiated by a problems report;
- to reduce the scope for duplicate reports.

Since installation, the PR mechanism has been generally successful. Both users and developers have responded positively to the system. Furthermore, no bugs/problems have been overlooked since the system came into use.

4.1.2 Formal versioning of QCDgrid

A formal versioning has been applied to the QCDgrid software suite. This is described in detail in [3]. The versioning system furnishes a number of important advantages:

- it allows the development team to isolate bugs/problems to particular versions of the software;
- it allows the UKQCD to confirm that each hosting system is running a current and supported version of QCDgrid software;
- it provides a coupling between the different documents and files that form a software release.

Since the introduction of the versioning system, two minor releases of the QCDgrid software have been made– the current version, at the time of writing, is QCDgrid Version 1.3.

The versioning system has been well-received by UKQCD. Furthermore, it has allowed the project team to bring the majority of hosting systems up to date with a supported version of the software. At the time of writing, all but one of the sites are using a supported version of QCDgrid (Version 1.2 or Version 1.3) – Rutherford Appleton Laboratories continue to use a pre-formal versioning release of the software which has tentatively been labelled as QCDgrid Version 1.1[4].

Excluding Rutherford Appleton Laboratories, the rollout of QCDgrid Version 1.2 took approximately three months to complete. This is longer than either originally anticipated (or desired).

4.1.3 The QCDgrid Test System

The QCDgrid project team have introduced a testing framework that is used to validate the functionality of new versions of the QCDgrid software. Tests are included to verify functional

[4] As part of the formal versioning process, a basic labelling was applied to early releases of the QCDgrid software. Specifically, issues of QCDgrid that interface to Globus Version 2.0 are identified as QCDgrid Version 1.0, and pre-versioning issues of QCDgrid that interface to Globus Version 2.4 are identified as QCDgrid Version 1.1.

compliance and ascertain the behaviour of the software under high load or common stress situations. The test framework is typically exercised to validate new releases of the QCDgrid software.

The testing module is archived in the CVS repository hosted on NeSCForge [2].

4.2 Notable problems and system failures

During the review period, the UKQCD Grid has encountered a number of problems. A description of each significant problem along with post-mortem information is provided below.

4.2.1 Exhaustion of the control node

During July/August 2005, the UKQCD Grid experienced a number of failures, caused by a reoccurring termination of the Control Node process. The source of the failures proved difficult to track down and involved both the QCDgrid project team and local support staff at the University of Edinburgh. The source was eventually traced to a load problem with the hosting system "edqcdgrid". Specifically, non-UKQCD Grid user activity was causing the memory on the system to be exhausted and the O/S to terminate processes such as the Control Node process.

The problem was resolved when users of the system were notified of the problem and consequently load on the system returned to previously experienced levels. However, the resolution does not prevent a future re-occurrence of a similar issue.

4.2.2 Proxy certificate expiration for Control Node process

During August 2005, the Control Node was, on several occasions, observed to enter a non-functioning state. This was caused directly by the expiration of the proxy certificate. Under normal operation, the proxy certificate should be renewed by the Control Node process if it is within one hour of expiring. Unfortunately, the volume of data held on the UKQCD Grid resulted in a single loop of the Control Node process taking in excess of an hour, blocking the proxy renewal. The problem was corrected by several performance enhancements to the validation code within the Control Node process.

The QCDgrid team are satisfied that this problem has now been resolved. The problem and resolution is documented on the NeSCForge Problem Report mechanism [2] under PR #144.

4.2.3 Hardware failure of the control node

At the end of October 2005, a hardware failure was experienced on the University of Edinburgh hosting system 'edqcdgrid'. This system was, at the time of writing, used both as a storage system and the control node. The hardware failure rendered the system unavailable throughout November.

When the Control Node fails, the Backup Node should automatically take over in read-only operating mode. In this particular instance, the Backup Node (hosted at the University of Swansea) did not take over control cleanly and users were denied access to the data catalogue. The failure of the backup node was traced to a faulty installation of Globus on the hosting system.

Currently, a temporary Control Node has been established at University of Liverpool. This has restored the UKQCD Grid to a reasonable level of operation since 24th November 2005, though several teething problems have been experienced.

4.2.4 Deployment of a new storage node (QCDOCx)

It was the intention of UKQCD to install a new storage node to hold data produced by QCDOC jobs. This storage node is intended to be deployed onto a standard, Linux-based compute cluster called QCDOCx. This deployment has experienced a number of significant delays, attributed to both hardware and software problems beyond the direct control of either UKQCD or the QCDgrid project team. At the time of writing, QCDOCx has QCDgrid Version 1.3 installed plus 8 Terabytes of disk capacity mapped for UKQCD Grid storage. However, the machine is operating as a client system only, with storage functionality suspended. This has been done to allow the machine to be system-tested and to confirm correct hardware/software configuration. Once this is done to the satisfaction of UKQCD and the QCDgrid team, storage functionality will be enabled.

5 Future plans and proposals

In this section, we describe the actions that have been proposed by the QCDgrid project team and UKQCD in response to the experiences documented in Section 4. Where possible, each action is attributed with a time-scale in order to allow the project team and UKQCD to gauge progress on the tasks.

5.1 Replacement of the control node

UKQCD and the QCDgrid project team have decided to transfer the control node process to a dedicated system to be hosted by the QCDgrid project team at University of Edinburgh.

As a dedicated system, the risks associated with third-party user activities affecting the operation of the UKQCD Grid (Section 4.2.1) should be eliminated. Furthermore, the new Control Node will be covered by a "one working day" maintenance policy, from Sun Microsystems, which should help to reduce downtime caused by hardware failures (Section 4.2.3).

It is hoped that the new Control Node will be in place by the end of February 2006.

5.2 Relocation of the backup node

Following on from the installation of a dedicated control node server (Section 5.1), the backup node function will be migrated from Swansea to Liverpool. This will be done, since the UKQCD have access to more local support time at Liverpool than at Swansea, simplifying the process of maintaining the machine. This should, in turn, reduce the likelihood of the backup node failing, as described in Section 4.2.3

It is hoped that the new Backup Node will be in place by the end of March 2006.

5.3 Delivery of a QCDgrid System Administrator course

The QCDgrid project, with the help of UKQCD Software Physics Programmer, aim to deliver a one day training course for system staff at hosting sites for the UKQCD Grid. A key motivation for the course is to reduce the downtime and project involvement in routine maintenance tasks, such as those described in Sections 3.1 and 3.2.

At the time of writing, this is in preparation and scheduled to run at the National e-Science Centre in Edinburgh on the 23rd March 2006.

5.4 EDG Virtual Organisation software

It remains the intention of UKQCD to deploy the EDG Virtual Organisation software at Rutherford Appleton Laboratories (RAL). Within the current system, RAL are required to manually add new certificates to their list of allowed users. This has been found, in the past, to lead to delays in new users becoming operational on the UKQCD Grid. However, looking to the future, the situation could be aggravated, once UKQCD becomes an ILDG service. A likely pre-requisite for this is an upgrade of the current EDG VO system to its successor – that is, EGEE VOMS [5]. Within ILDG, non-UK e-Science users will be attributed with restricted permissions to access data. It is not clear how these permissions can be queried by a host at RAL, without their usage of EGEE VOMS.

UKQCD continue to liaise with RAL staff (principally, Nick White) to resolve these problems. No timescale for completing the task is currently available.

5.5 Time synchronisation across hosting systems

The expectation is that the NTPD time synchronisation daemon will be deployed across the UKQCD Grid, in order to eliminate (or, at least, reduce the number of instances of) time skew problems (see Section 3.1.4). Dr Chris Maynard (at the University of Edinburgh) is currently investigating the most

appropriate method and technology for completing this deployment. It is intended that a time synchronisation mechanism will be in operation by the end of April 2006.

5.6 Automated monitoring of the UKQCD Grid

In addition to the current monitors and checks that run on the UKQCD Grid, the QCDgrid project team are – at the time of writing – exploring options for additional monitors. The intention is to establish a more concrete understanding of the state of data on the UKQCD Grid and, hopefully, provide a more rapid diagnosis of problems as they arise.

In particular, the project team are creating a prototype process that will compute the number of copies that exist for each file stored on the grid. In addition to providing an early indication of potential problems, it will provide new information for UKQCD and the project team, regarding the speed with which the control nodes adapts to change – for example, addition of new data to the grid or retirement of a storage node. It is hoped that the a prototype of this monitor will be evaluated during March 2006. Other specific elements will be developed and deployed during the year, based on need and effort availability.

5.7 Better diagnosis of firewall issues

The QCDgrid project team will prepare a sheet of instructions for local system administrative staff at hosting sites, to formalise the process used to diagnose firewall problems (Section 3.1.3). The intention is to reduce the scope and longevity of problems associated with mis-configured firewall. The instructions will be prepared for the QCDgrid System Administrator course (Section 5.3) and should be finalised after the course by the end of March 2006.

5.8 Upgrade of third-party software components

The QCDgrid software is currently built on top of Version 2.4 of the Globus Toolkit (GT 2.4). This version is no longer either developed or supported by the Globus Alliance. Recent evidence has suggested that an incompatibility gap is beginning to form between the requirements of current generation operating systems/services and the constraints imposed by GT2.4. For example, the GridFTP client which implements a secure file transfer model, required at a number of places in the QCDgrid design, is incompatible with the latest versions of the Linux kernel. The incompatibility gap is likely to increase as time progresses and lead to a reduction in the functionality of the QCDgrid.

The QCDgrid project has nine months of effort set aside for an upgrade to the third-party components on which the UKQCD Grid depends. The project team propose to invest this effort to upgrade the underlying Grid technology from Globus Version 2.4 to Globus Version 4.x. As part of this upgrade, the project team will evaluate new functionalities available in Version 4, with an intent to enhance the operation of QCDgrid. The project team also aim to upgrade the XML database (eXist) and EDG VO software to current versions. This will not only improve the service provided to UKQCD, but also simplify the integration of the UKQCD Grid into the International Lattice Data Grid (ILDG) [4]. It is intended that the overhaul of QCDgrid will be completed by the end of 2006.

6 References

[1] QCDGrid: Probing the building blocks of matter with the power of the Grid, QCDgrid Project Homepage available on-line at http://www.gridpp.ac.uk/qcdgrid/.

[2] NeSCForge QCDgrid Project Trackers, web-based reporting system, at http://forge.nesc.ac.uk/tracker/?group_id=13.

[3] M.G. Beckett, D.J. Byrne, *Work Package 1.5: QCDgrid Software Release Plan* (March 2005).

[4] The ILDG Wiki http://www.lqcd.org/ildg/tiki-index.php.

[5] EGEE, Activity JRA-3 Security, http://public.eu-egee.org/activities/jra3_details.html (December 2004).