

Grid Deployment and Operations: EGEE, LCG and GridPP

Jeremy Coles

GridPP – CCLRC Rutherford Appleton Laboratory (RAL)

Abstract

Over the last year the Enabling Grids for E-science (EGEE) project has seen its deployed resources grow to over 12,000 CPUs in more than 30 countries. During this period the UK and Ireland Regional Operations Centre (UKI ROC) has played a significant role in the development and refinement of the presently used deployment and operations model. This paper reviews the background to the EGEE project and describes these models. As a large part of the EGEE resources are provided by sites of the Large Hadron computing Grid (LCG), and within the UK by the Grid for Particle Physics project (GridPP), this paper describes the deployment planning and operational activities surrounding these organisations and provides some insight into future plans and challenges.

1. EGEE background

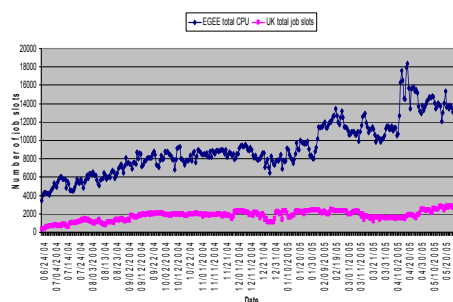
The EU funded Enabling Grids for E-science¹ (EGEE) project has developed a “production-level grid where user communities are able to do real work”. The infrastructure is available to support academic users 24 hours-a-day and integrates national and regional grid efforts across the globe². The project was founded on the successes of the European DataGrid Project³ (EDG) and DataTAG initiatives and continues to overlap strongly with the Large Hadron Computing Grid (LCG) production grid infrastructure. There are more than 900 registered users.

The current operational grid⁴ is based on LCG-2 middleware (see 2.1) and is deployed at about 130 resource centres (RCs) which are providing in excess of 12,000 CPUs and 4PB storage^a. There are some 2000 CPUs provided through LCG (i.e. non-EGEE) related sites in Taiwan, China, Japan, Canada, Pakistan, India and the USA. There are also two main industrial partners - HP and CMG. Expansion is continuing. A Memorandum of Understanding has been signed with South Korea and collaboration in the Baltic States has been proposed.

New regional grids integrate easily into the EGEE structure because of the establishment of Regional Operations Centres (ROCs) and Core Infrastructure Centres (CICs) which are described later. All federations are providing resources at a level close to or, in most cases, exceeding those committed to in the EGEE Technical Annex² which puts forward the ambitions and deliverables of the project.

^a These and other figures quoted are correct at the time of writing in July 2005.

Figure 1 shows the increase in computing resources as the EGEE project has grown. The data is based on the number of job slots being published (job slots and not CPUs because hyperthreading is sometimes enabled at sites and doubles their CPU count).



among the members of a group of partners. CERN is an exception in this scheme and takes responsibility for sites that do not fall within the control of an existing ROC. Taiwan is currently involved with CIC work and is expected to become a ROC for the Asia Pacific region.

Use of the deployed infrastructure is open to users who belong to a recognised Virtual Organization (VO). Usage follows organisational and operational agreements established by an Operational Advisory Group (OAG) which consists of members from SA1 and NA4 (the applications area). These agreements are in the form of MOUs which, for example, detail the essential requirements of the VO applications. VOs negotiate access to new site resources via the ROC and site managers on the basis of the MOUs. The need for an OAG was missed in the early stages of the project but has come to be seen as essential to provide a much needed link between the resource centres and VOs. It is hoped that this approach will encourage more sites to support VOs outside of their normal user community – many sites have been reluctant to fully open up their resources.

1.2 The UK and Ireland ROC

The UK and Ireland (UKI) ROC⁶ has contributed widely to the development of the models for deployment and operations described later. The ROC consists of three partners: Grid Ireland⁷, the UK National Grid Service⁸ (NGS) and the UK Grid for Particle

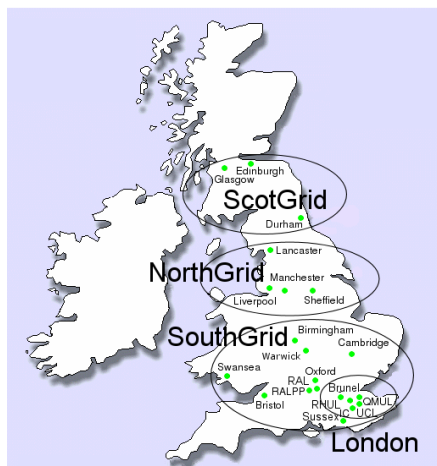


Figure 2: UK GridPP regional deployment structure

Physics⁹ (GridPP). At present GridPP provides most of the UKI resources visible in EGEE

largely because EGEE now runs operations for the Large Hadron Collider Grid¹⁰ (LCG). LCG has a primary goal of producing a grid infrastructure to cope with the vast data output from CERN LHC experiments from 2007 onwards. NGS (4 sites) and Grid Ireland (17 sites) resources in EGEE are expected to increase over the coming year.

2. Deployment

The UKI ROC team consists of a manager, deployment coordinators for each of the geographical regions (see Figure 2), a security officer, helpdesk (provided by the Grid Operations Support Centre at RAL) and a Core Infrastructure Centre (CIC) group. All activities are overseen by a UKI Deployment Board setup by GridPP with participation from the other UKI grids. In addition a number of experts are available to support deployment efforts in critical areas such as networking and storage. The UK Tier-2 Centres are organisational structures with their constituent sites appearing on the grid as separate entities but it is expected that eventually their resources will appear externally as a single collection of resources.

The deployment team communicate and provide regional help using email lists, web-page summaries, FAQs and more recently using Wikis which have the required dynamic characteristics to provide for a rapidly changing environment. Review meetings take place face-to-face every 3-4 months and monthly by teleconference. These are particularly useful to share technical knowledge following a new middleware release which contains new components and functionality.

2.1 Middleware overview

Each EGEE release (currently still tagged as an LCG2 release) comprises a set of profiles for the different grid elements:

- **Computing Elements:** computing resources at each site, composed of a front-end node and a set of Worker Nodes. Jobs are managed via a Local Resource Management System (e.g. PBS, LSF);
- **Worker Nodes:** the computing nodes where jobs are run;
- **Storage Elements:** classical disk storage nodes available via a GridFTP server and Storage Resource Management (SRM) based storage nodes which use an SRM

interface. There are three main SRM implementations currently in place: dCache, the Disk Pool Manager (DPM) and CastorSRM

- **Resource Broker (RB)**: finds the best Computing Element that matches the job requirements. Normally it provides also the Logging and Bookkeeping service, to keep track of the status and the history of the submitted jobs. RBs are run in most federations and at CERN;
- **User Interface**: node where the end user accesses the grid resources (typically to submit jobs to a Resource Broker);
- **Information Index** is provided by the Berkeley Database Information Index **BDII**: node which collects the status of Computing Elements and Storage Elements. It is accessed by other tools/services (like the Resource Broker) to find the available computing/storage resources. BDIIs are run by most federations and also by many VOs ;
- **Replica Location Service (RLS)**: node which maintains information about the physical locations of logical identifiers of data and provides access to this information. This currently runs at CERN;
- **MyProxy**: service provides a credential repository (needed to renew grid proxies for very long jobs);
- **Monitoring service**: server for a Relational-Grid Monitoring Architecture (R-GMA) based information system. GridICE information providers are included in the CE, SE, WN and RB;
- **VO server**: Lightweight Directory Access Protocol (LDAP) or VO Management Server (VOMS) based implementation that contains the list of authorized VO members.

An overview of the EGEE middleware with particular attention to the relationship with external grid middleware and other software packages is described in the LCG-2 Middleware Overview document¹¹. The EGEE source code is available from the CERN Concurrent Versions System (CVS) repository under the “lcgware” area¹².

2.2 Installation mechanisms

At the start of the EGEE project it was decided to support a preferred installation mechanism. LCFG (Local ConFiGuration) was adopted as a system for automatically installing and managing the large numbers of Unix systems. Despite the steep learning curve required to master it, most sites used it until a move away from Red Hat 7.3 to Scientific Linux at the beginning of 2005. The development effort (provided mainly by Edinburgh University and CERN) forked into updates to LCFG and a new product called Quattor (Quattor is an Administration ToolkiT for Optimising Resources). With the move away from LCFG the EGEE project made a decision not to mandate or support one installation mechanism. Today, Quattor components and configuration files (which are required for installing new LCG releases) are produced by a community of interested sites and not centrally by the SA1 deployment team. SA1 has encouraged the formation of independent support communities to reduce the workload on the central deployment team based at CERN who have at times struggled to support the growing needs of EGEE members.

The main problem with LCFG and its later version LCFGng (“ng” for next generation) was its modular nature. It was based on a rigid schema where components (e.g. network and hardware) ended up duplicating each other. Though the architecture of Quattor is similar to LCFG, there is a big difference in its global schema approach.

In addition to the automated installation mechanisms a manual installation method has always been provided with LCG releases. Since the end of 2004 a semi-manual installation mechanism known as YAIM (Yet Another Installation Method) has been developed. This allows middleware installation to be carried out using a series of scripts after the basic operating system has been installed by the sites’ preferred method. At present the majority of EGEE sites use a mixture of YAIM and Kickstart to install and configure their grid clusters, but interest in Quattor is growing especially for larger sites (where the advantages of less configuration problems and a more assured secure installation) are more evident. However, script based installation is much preferred by most system administrators due to the transparent nature of errors and ease of debugging. Until the middleware installs without problems there will continue to be conflict between difficult to learn

and use installation systems and those that are easy to grasp but prone to configuration problems. Other installation methods are being investigated. Indeed gLite (which uses a Web Services architecture to address shortcomings found with LCG2 functionality and stability), uses xml based configuration files. This move by the developers away from a solution the site administrators prefer highlights another issue – that there is not enough cross group collaboration in EGEE.

In addition to the middleware installation tools, mechanisms are also provided for Virtual Organizations to distribute and install application specific software on the EGEE infrastructure in user space.

2.3 Middleware releases

While the EGEE middleware developers (JRA1) provide new releases every few weeks for integration testing, the SA1 deployment team only release a main middleware tag to sites approximately every 3 months. Between integration testing and a release candidate being announced there is significant work to be done from packaging and testing to updating the release and support notes. Between the main releases are minor bug fix and security updates.

There has been some contention about how often releases should be made and what they should contain. Early releases were frequent and came with minimal advance warning. This was in line with the need to resolve major bugs and introduce functionality the users needed as soon as possible. However, as the number of sites (and resources) has grown it has become increasingly difficult to get sites to adopt new releases quickly partly due to uncertainty (why not wait for the next release?) and partly because additional smaller sites with less manpower have found it more difficult to schedule time to the activity. (Some sites installed the middleware in order to become part of the collaboration but have not kept up with releases). A compromise was reached whereby EGEE deployment has agreed to tag new major releases every 3 months and release on specific dates. In return the sites have been asked to install the release within 3 weeks. Security patches or minor updates (for example a simple RPM update) can be released at any time and sites are asked to install them as soon as possible.

While the approach has been largely agreed and followed it has proved difficult to get all sites to

upgrade quickly. Overall uptake of a release is improving as can be seen in Figure 2, however it is clear that for long periods multiple releases are going to be running in parallel. This is even more evident when considering sites not running the recommended Scientific Linux (SL). Often ports (undertaken by EGEE partners) are required to the non-standard operating systems or architectures (e.g. IA-64) and this introduces additional delays. Even with SL installed there are no “standard sites” (each having its own peculiarities) and the middleware has so far struggled with this situation.

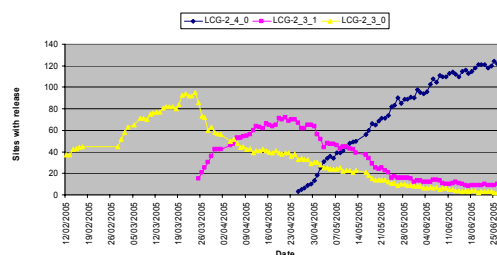


Figure 2 Middleware upgrade progress after a major release is almost linear despite frequent requests to upgrade quickly.

In the UK, site administrators install the latest release (held in a central repository) with help from their Tier-2 coordinators and the UK support team. Support for administrators and users is also provided by a central EGEE group based in Karlsruhe, Germany. The Global Grid User Support¹³ (GGUS) facility is being built to interact directly with regional support desk applications and teams, and to provide a core into which each Virtual Organisation (VO) can link with user specific support. In addition GGUS provides links to useful tutorials (in conjunction with the National E-Science Centre in Edinburgh) and other support information such as the LCG Wiki¹⁴. With a rapidly growing repository of information and knowledge, and a frequently changing environment (both the middleware and deployment tools) many lessons have been learnt and captured in this Wiki and shared through a number of mailing lists and areas like the GridPP deployment web pages⁹. Mailing lists are becoming less favoured due to the large volume of messages they produce as the collaboration grows, yet it is proving difficult to change the behaviour of such a large group into using more appropriate communication mechanisms (like support desks).

GGUS is intended to provide a support system which is world wide in its organisation and which can deliver support with 24 hour cover every day of the year. This will include service centres in the US and Asia as well as Europe, with responsibility rotating around the globe in 8 hour shifts. There is a program of work in place to enhance the present system to bring its operations into alignment with the needs of EGEE and work more closely with other grid operations centres. This area is proving difficult because of the different support systems, languages and approaches employed across the EGEE federations.

2.4 Users

On the grid infrastructure, there are at present some 60 virtual organisations (VOs). EGEE only provides active support for some of these VOs. At the time of writing the number of EGEE supported VOs is 10 including the largest which are from the High Energy Physics area. It has become clear that the level of effort needed for user support within EGEE was underestimated. This has meant that the user support structure has not yet seen anything like the success experienced in the operations area. Second-level support experts tend to be those already in the project engaged with other work and the documentation produced is often inadequate for new comers.

The number of active grid users is climbing though perhaps not as quickly as if joining VOs had more experience in adapting their applications to work on the grid. More effort is required in this area and to train grid applications developers, especially as new and different groups have different priorities that need to be addressed – for instance HEP VOs have no requirement for Message Passing Interface (MPI) jobs while new VOs do. Another approach being pursued by GridPP is the development of web-portals¹⁵ which will simplify the process of preparing, submitting and reviewing grid jobs.

3. Operations

To maintain as much order as possible with such a widely distributed set of resources, the job of site certification (once undertaken centrally at CERN) is now the responsibility of the ROCs – this has been seen as necessary to ensure reliable and scalable operations as the grid continues to grow. When a new site is ready to join it is introduced as a candidate site. Various information is registered in the Grid

Operations Centre (GOC) database (GOCDB) including site contact email addresses and telephone numbers, a security contact, the proposed name of the site, its domain and the hours of local support. Additionally the site must also have agreed to the EGEE security policy.

When ready to become a certified site a comprehensive series of tests are run against the site (Site Functional Tests or SFTs). When these are passed to the satisfaction of the local ROC the site is certified and its monitoring enabled. When a site is certified it appears in the grid as a resource to which VOs can submit jobs.

3.1 Daily operations

The integrity of all the grid resources is constantly monitored by one of the (currently five) Core Infrastructure Centres. Each CIC takes responsibility for monitoring the daily-run job submission test results and grid statistics (gstat) pages¹⁶. Any problems which are found are recorded in a problem tracking tool and followed up using additional information provided by the Grid Operations Centre database (e.g. contact information, node and scheduled downtime details for the site)¹⁷ developed and supported at the Rutherford Appleton Laboratory in Oxfordshire. Changes in operational responsibility take place at weekly handover meetings where current issues and problems are discussed.

Each CIC is responsible for the service one week in five. Tools have been put in place to support this work, and operational procedures have been written to provide guidance. This service started at the beginning of November 2004 and led to a great improvement in the overall health of the grid. This move also freed up experts at CERN who had developed the monitoring tools and subsequently became caught up in the support effort as the scale of operations expanded.

3.2 Monitoring the Core services

In addition to monitoring overall site status information, the CICs are responsible for the core grid services. As such they monitor the status of the daemons (and host machine) of the Resource Broker; check job submission activity to verify the activities of a VO; advise the user community about service downtimes; monitor the BDII and MyProxy servers; maintain the RLS servers and back up the RMC/LRC databases.

Additionally, for the purposes of managing the grid the following are treated as core services and monitored by the CIC: Accounting; User registration; Security co-ordination; a Central User Interface (that any registered user can use) and the Grid Operations Centre Database (GOcdb).

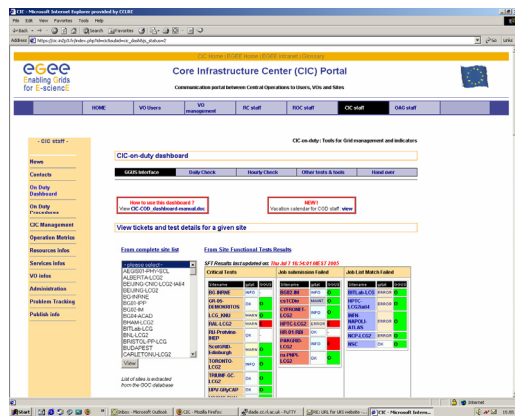


Figure 3 Screenshot of the CIC portal

To aid operational monitoring a number of real-time tools have been developed in addition to the SFTs and GOcdb already mentioned. There is a GIIS monitoring service, certificate lifetime monitor, live job monitor, a scheduled downtime service and a GridIce fabric sensors summary page. Some of this information is aggregated into weekly status reports for each region and passed on for discussion at a weekly operations meeting between representatives of the federations. A summary of the information is also presented to the CIC in a portal hosted by one of the partner sites.

There has been a lot of progress over the last 9 months and this is set to continue. A next step involves automating alarm systems to alert sites when problems occur as well as further developing interoperation with OSG and others.

3.3 Usage and metrics

Operational planning has been difficult without good data on the status of sites. With the introduction of standard tests being run against sites and the data being archived it has become possible to identify trends in performance – both for the whole grid, for regions and for individual sites. Data currently available traces such things as the number of concurrent running jobs, the number of available job slots, the amount of storage, downtime information, utilisation and more detailed information from middleware functionality tests (currently of the

order of 16 tests of things like copying and registering files and replication) together with measures of overall site stability (noting problems with configuration information and critical components failing).

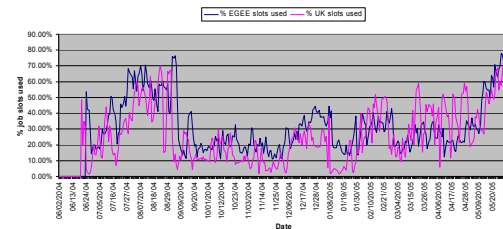


Figure 4 An indication of resource usage for EGEE and for UKI.

Figure 4 shows the variation in usage of EGEE and UKI resources as a percentage of the total available. It reveals that utilisation is still quite low but is improving. Peaks are seen at times corresponding to data challenges run by different VOs. While under utilisation is a concern for some it should be remembered that building a production grid requires a steady increase in provided resources even if current usage does not require it. SA1 had a deliverable to provide a system for the collection and dissemination of accounting information on the grid. The Accounting Processor using Event Logs (APEL) package was the result. It provides summary information on resource usage by VO per country or region. The customisable summary information can be accessed on the Accounting Portal pages¹⁷.

There is also a group within EGEE (JRA2) charged with developing a web site which displays information on the reliability of the grid¹⁸. At present the site is under development and requires a valid grid certificate to access the information. The data suggests that grid efficiency (in terms of jobs submitted that run successfully) is improving steadily. This has been verified independently by several VOs who see efficiencies between 65% and 98%.

3.3.1 Improving operational performance

One of the biggest operational problems has been in the area of site configuration and stability. As a result VOs have been selecting only specific sites that they consider reliable. In an attempt to generalise and ease this approach, the site functional tests are now run more regularly and used to populate a database. Through a web-browser the VOs are given the option to independently select those tests that they consider critical and also to white or

blacklist sites to always use or not use them. This approach, called “Freedom of Choice”, is expected to further improve job efficiencies because the information is updated every few hours and problematic sites are not used.

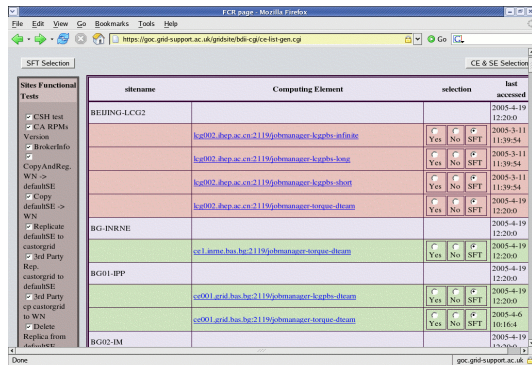


Figure 5 Screenshot of the Freedom of Choice tool that allows VOs to preselect sites having a specific operational test job results.

The “stability versus new features” conflict has been a problem and although a balance has almost been made, it is possibly about to be upset by the introduction of gLite. This would be less of a problem if the middleware was more faults tolerant such that small sites could not upset job distribution across the whole grid. Until that is the case priority for fixing problems needs to go to the larger sites with more resources.

4 LCG

The LCG project has the goal of prototyping and then deploying the computing environment required for the LHC experiments. Phase 1 has fulfilled the first objective in the period 2002 to 2005 and allowed experience to be gained in the running of a production grid service and the development of detailed technical designs for computing to meet the needs of the various experiments. Phase 2, 2006 to 2008, will build and commission the initial LHC computing environment. The project has been divided into six main areas – project management, applications (common physics application software), fabric (large cluster management and networking), distributed analysis, middleware and grid deployment. The last two areas are being done jointly with EGEE. The fabric area is benefiting from a research facility (LCG openlab – co-funded by industry and CERN) where experience is being gained with 64-bit programming, next generation I/O and new

installation and configuration tools. The other areas are dealt with in a GridPP paper¹⁹.

Due to the operating performance required for the LHC by 2007, current activities are constantly pushing the capabilities of the deployed infrastructure. They also require good interoperability and management with NorduGrid²⁰ and the US Open Science Grid²¹ (OSG). Early in 2005, test jobs were successfully submitted between the grids using test clusters and the changes required are now being incorporated into the standard EGEE release. The current goal is for full interoperability before the autumn of 2005 with collaboration on areas such as the new glue schema already taking place.

4.1 Baseline services

The Baseline Services group was setup between the experiments and regional centres to agree on basic needs to support the computing models for the start up of the LHC. It is defining services with targets for functionality and scalability taking account of site and other constraints. The discussions have among other areas covered storage management services, transfer services, grid catalogue and data management tools, as well as VO management services and workload management. This work continues to drive the direction of the EGEE middleware.

2.1 Service challenges

LCG progress is being enhanced by a decision of its strategic body, the Grid Deployment Board, to run a series of challenges between 2004 and 2007. Significant effort from all EGEE regions involved in LCG has been invested in the support of the challenges. This represents the first major significant use of a grid service of this size, and is an extremely valuable exercise, providing feedback on the middleware functionality, operations procedures, and debugging and management tools. During the challenges so far, significant numbers of jobs have been run, at peaks reaching over 10000 concurrent jobs.

Service Challenge 2 (SC2) dealt with sustained reliable data transfers to large LCG sites from CERN at a rate of 500 MB/s. SC3, which started in July 2005, extends this to include provision of a reliable base service with more sites and some basic experiment software. SC4 is scheduled for the Spring of 2006.

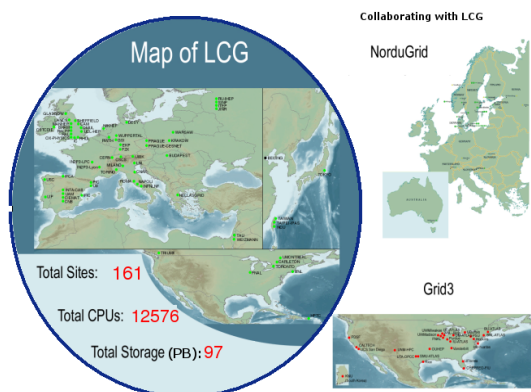


Figure 3: LCG has 3 major grid partners

It will include all major centres and intends to test the full LCG required software chain and sustain nominal final grid data throughput. The intention is to raise the service to full operational capacity by April 2007 – this must be capable of handling twice the nominal data throughput to allow for some periods of downtime.

Lessons learned are being used to improve the production service. Much of the benefit is coming through preparing sites for the challenges as this uncovers many hidden problems. In the UK the challenges prompted a re-evaluation of the High-Energy Physics network requirements and led to the first major use of the lightpaths setup under UKLIGHT²². Running and debugging the service challenge tests is being helped by the use of progress summary logs²³ which detail problems encountered and the resolution steps taken.

In addition to the main Service Challenges, EGEE is also conducting other challenges. The BioMedical VO have recently undertaken a job throughput challenge and the security group in EGEE have simulated incidents to test how well sites are able to respond.

3.2 Future activities

The EGEE grid is already at the scale of that required for the LHC in terms of number of sites, and is thus starting to reveal the full complexity of operations. However there is a major hurdle to be overcome with a move over to gLite middleware. Details of how the transition will happen are not yet final but it is likely that larger sites will start by deploying new components in parallel with existing implementations. As the new components demonstrate an improvement in performance

and an acceptable level of reliability more sites will adopt it. A pre-production service already exists within EGEE. Sites volunteer to be part of this limited testing environment but are too few to demonstrate some of the scaling issues that are likely to be encountered.

As the EGEE and LCG projects move towards a production service the middleware code will need further hardening. To aid this process the UKI region is developing a vulnerabilities database that aims to systematically capture details of security problems found and prioritise them. In parallel efforts to improve incident handling and site security practices will increase. It is not a trivial problem to coordinate security responses in such a large and diverse collaboration where one must consider legal responsibility globally. Tools are being developed to allow better general monitoring of cluster patching and logging and analysis of activities in preparation for the development of grid intrusion detection techniques.

GridPP has devised a deployment plan which is updated quarterly. As well as details of internal challenges to ensure site readiness, it contains milestones and associated metrics for monitoring the overall progress of deployment. High-level information is presented in the GridPP project map²⁴.

¹ <http://egee-intranet.web.cern.ch/>

² <https://edms.cern.ch/document/400278>

³ <http://eu-datagrid.web.cern.ch/eu-datagrid/>

⁴ <http://goc.grid.sinica.edu.tw/gstat/>

⁵ <http://egee-sal.web.cern.ch/egee%2Dsal/>

⁶ <http://goc.grid-support.ac.uk/ROC/>

⁷ <http://cagraidsvr06.cs.tcd.ie/>

⁸ <http://www.ngs.ac.uk/>

⁹ <http://www.gridpp.ac.uk/>

¹⁰ <http://lcg.web.cern.ch/LCG/>

¹¹ https://edms.cern.ch/file/498079/0.1/LCG-2_Overview.pdf

¹² <http://isscvcs.cern.ch:8180/cgi-bin/cvsweb.cgi/?hiddenonreadable=1&f=u&log>

¹³ <https://www.ggus.org>

¹⁴ <http://goc.grid.sinica.edu.tw/gocwiki/>

¹⁵ <http://www.gridpp.ac.uk/portal/>

¹⁶ <http://goc.grid-support.ac.uk/gridsite/monitoring/>

¹⁷ <http://goc.grid-support.ac.uk/gridsite/accounting/>

¹⁸ <http://ccjra2.in2p3.fr/EGEE-JRA2/OAmeasurement/showstats.php>

¹⁹ “GridPP: Meeting the particle physics computing challenge” – GridPP, AHM 2005

²⁰ <http://www.nordugrid.org/>

²¹ <http://www.opensciencegrid.org/>

²² <http://www.uklight.ac.uk/>

²³ <https://uimon.cern.ch/twiki/bin/view/LCG/ProgressLogs>

²⁴ <http://www.gridpp.ac.uk/pmb/>