



GridPP Technical Board

---

# Rollout of the UK Grid for PP

---

Planning for EDG Testbed software deployment  
and support at participating UK sites

---

Document identifier:	<b>GridPP-TB-02-v4.2-UKrollout</b>
Date:	<b>30/07/2002</b>
Version:	<b>4.2</b>
Document status:	<b>FINAL</b>
Author	<b>Pete Clarke, John Gordon</b>

---

# 1. Introduction

## 1.1 Purpose of document

This document discusses issues pertinent to formulating a **planned and supported** deployment of the DataGRID testbed software to all participating sites in the UK.. Since the EDG software provides the base for GridPP developments, this is potentially of interest to all sites and experiments.

The goal is to define a scheme such that no institute system manager need be unsure about the what/when/how in respect of deploying EDG testbed software.

This document only addresses the issues related to Grid Testbed software deployment.

It does not address the issue of experimental production code distribution.

This document has been developed through several stages involving iteration with the GridPP community. It has finally been circulated to the entire UK Grid mailing list for comment. Following this consultation the status of this document is FROZEN.

New versions may be produced in the light of experience of the rollout of the imminent testbed releases.

## 1.2 Background

### 1.2.1 What is the UK Grid for PP?

The UK Grid for PP is the intersection and overlap of a number of grid projects both within the UK and extending into the UK from elsewhere. It includes a variety of computing resources and staff and aims to meet the UK commitments to a variety of PP projects. GridPP attempts to put in place an organisational structure to permit and encourage the sharing of resources and the mutual re-use of developments.

### 1.2.2 What does an Institute Gain?

By connecting some or all of its resources to the Grid, an institute raises its profile within the collaborations in which its physicists belong. It can more easily participate in world-wide activities like data challenges, reconstruction and thus raise the profile of the institute.

The institute retains ultimate control over its resources; both who and when they are used by others.

### 1.2.3 How does an Experiment Participate?

When its members join together in a Grid, an experiment gains an easily-managed pool of resources for use in major experiment-wide shared tasks like simulation or reconstruction or by sharing UK resources to give UK physicists a collective advantage over collaborators from other countries during analysis..

### 1.2.4 What does a user get?

Users will see a uniform access to resources between sites and experiments. With a common infrastructure, tools will be re-usable between experiments and are more

likely to be well documented. Expertise gained on one experiment will be useful to others.

### 1.2.5 The Community

Bigger than the advantages discussed above is the gain to the UK community as a whole by sharing resources and staff and developing computing together. Many of the tasks done in PP are common to all experiments; GridPP encourages collaboration between experiments, not just institutes. The resultant developments using grid technologies are immediately re-usable. Examples of such collaborations on Ganga and SAM are already visible and products of EDG are being re-used by BaBar and CDF.

### 1.2.6 Participation

There is a wide range of participation possible in GridPP.

At the very low end, users from an institute wishing to minimise local support will be able to login to a grid user interface machine at RAL and thus gain access to the wider resources available.

At the high end, an institute's JIF/JRE/SRF resources can play a major part in the computing of one or more experiments and achieve world-wide publicity.

GridPP aims to help all institutes and groups to participate in the UK grid for PP.

## 2. The Deployment

### 2.1 Types of site

Three types of site have been identified. These are defined below:

a) **Category-1: Core sites.**

These include at least the sites with allocated testbed effort (WP6 effort), i.e. CLRC-RAL, ICSTM, Manchester and Bristol. These sites are, per definition, fully participating in the work required to rollout testbed releases, in full coordination with EDG WorkPackage-6. A commitment to being a primary testbed site (as defined by EDG) is required to be within this group.

b) **Category-2: Friendly testers**

These are sites which typically do not have the effort, or otherwise cannot make the commitment to be a full category-1 testbed site, but nevertheless wish to be early adopters of testbed releases on an "as is" basis. Such sites are important for validation that the deployment instructions actually work. These sites will be supported through ad hoc phone meetings, email list ...etc. but do not expect a formal "turnkey" installation with full tracked backup.

c) **Category-3: User only sites**

These are the sites who only wish to see a testbed release when it is guaranteed to be working, and can be sure that it can be installed by following the canonical instructions. They should also have access to support through a single point of contact. These sites will include institutes who do not have the resources or inclination to be cat-1 or cat-2, or have

critical experiment production requirements, or even third party resource providers (e.g. eScience centres !)

***The only category where there is an entry requirement is cat-1, where the site would have to sign up to the expectation of an EDG testbed site. None of these categories is exclusive. Any site may, as it wishes, define itself to be in several categories. E.g. a site may chose to participate as cat-2 in respect of part of its resources (a development cluster), and also as cat-3 in respect of another part of its resources (some experiment production cluster). It is anticipated that most WP8 sites will wish to be cat-3 sites, however further discussion is needed to determine which category the sites upon which they may depend should be.***

**This document mainly concerns category-3 sites.**

It is expected that category-2 sites will nevertheless benefit from some sort of planning and order being imposed.

The document is not very relevant to category-1 sites.

## 2.2 What Category-3 sites should see in terms of "turnkey" installation/update

Category-3 sites may expect a reliable "turnkey" installation/update procedure (within reason – let us not forget that this is an R&D project). That is to say, sitting in front of a web-browser and a telephone they should be able to rely upon:

1. A one stop set of instructions, which, insofar as possible, are proven to work.

This means that the installer should NOT have to have prior knowledge of many different GridPP and EDG instruction sets/ WWW pages etc. They MUST be able to start from a single GridPP page and from this be presented with all the instructions they need for to deploy the software for a User Interface (UI), Storage Element (SE) or Compute Element (CE) on a canonical system.

2. A single point of contact for support when problems arise.

We envisage such support going through the CLRC Grid eScience Support Centre (GSC). This will provide help through logged and ticketed telephone and/or email request. There should be some sort of guarantee on response time, and a well defined problem escalation procedure. The GSC have already agreed to this in principle.

### 2.2.1 Work/Tasks/Requirements to achieve this:

1. The proposed testbed release vetting procedure is described in an appendix. This should be formulated such that at the end of the procedure the instructions are known to work. The person identified above will be part of this vetting procedure, and therefore keep track of failures, subsequent fixes, and ensure appropriate documentation modification.
2. Support through the GSC should be constituted. This requires GridPP and the GSC to define and agree a workable "backend" to the help contact point. Where possible, each WP should identify two people who are formally prepared to field questions in good time.
3. A suite of "validation" tests must be made available. There will be run from the core sites to validate a successful installation at a remote site. Such would include visiting the information servers, submitting a job, etc..

4. An individual will be identified to take ownership of the Rollout support mechanisms described here and in the Appendices.

### 2.3 Third Party Sites

It is important that GridPP keeps in mind the possibility of using "third party sites" to provide resources. In the future this may be a commercial CPU cycle provider, but in the short term this will be willing eScience Regional centres not closely connected to the GridPP core sites.

For this to be feasible it will be mandatory that the installation/upgrade procedures are as "turnkey" as described above.

### 2.4 Production Requirements

In the case of LHC experiments: GridPP formally defines its policy on use of the Grid for production to follow that of the EDG and LCG projects.. In this way GridPP is not required to underwrite within the UK more or less than is expected internationally.

In the case of non-LHC experiments (BaBar, CDF, D0, Other) a policy needs to be developed in order to make it clear what may be reasonably expected of the GridPP testbed in the UK in respect of its use for production requirements in 2002/2003.

### 3. Appendix:

#### Suggested procedure for validating testbed releases and vetting the “public installation instructions”

1. Core sites to keep up with testbed releases as part of their normal remit. Members of core site will interact directly with each other and with WP6 and the Integration team to make each release work on the UK EDG Testbed sites.
2. Reports on problems and fixes found as part of core site activities should be posted to the tb-support email list (as well as bugzilla). In this way other “listening” sites can keep up to date with issues.
3. A bi-weekly testbed co-ordination virtual-meeting, lasting no more than 1.5 hours, is to be set up. Here progress and problems are to be reported, and any escalation actioned (e.g. anything which needs to go up to management).
4. As soon as possible/appropriate any willing category-2 sites (and at least two such sites must be prepared to do this) attempt to follow the most up to date installation/update instruction set – as specified in the document above. The bi-weekly virtual-meeting will determine when it is appropriate to recommend such attempts.
5. Any failures or problems encountered by cat-2 sites should be reported to the tb-support list (and Bugzilla where appropriate). When any problem is resolved the instructions should be updated and the cat-2 sites should re-start the installation from before the failure point (this may mean from the beginning, or from some other suitable point)
6. A suitable mechanism is needed to ensure timely response to problems reported on the tb-support email list. The point is that somehow outstanding problems need to be picked up on more than a best efforts basis. Until such systems are in place the recommendation must be to resubmit unanswered problems.
7. Continuing problems etc to be discussed at the bi-weekly virtual-meeting with participation from Core, cat-2 sites and TB.
8. When a cat-2 site claims to have successfully installed, it should be verified by a standard validation test suite run from CLRC-RAL
9. The above continues until at least 2 cat-2 sites can verify that they have successfully installed/upgraded from the canonical instructions alone, and passed the validation tests.

**At this point the testbed release is deemed to have been vetted for use in the UK by category-3 sites.**

10. These sites are then explicitly asked to install/upgrade
11. If they encounter problems then the official supported help route is through the GSC help desk (either phone or email list). This will have the advantage that the problem is logged, and its progress will be monitored.
12. These sites are also welcome to use more informal means (e.g. the tb-support email list, or participation in bi-weekly meetings) should they so wish.
13. Once an installation/upgrade is apparently successful, the validation tests are run to verify that the site is functioning correctly.

It is necessary to assign the management of this procedure to an identified individual. This person should be responsible for at least:

- integrity and verification of the instruction set,
- ensuring that the above steps take place in a timely manner

- ensuring that help requests from cat-3 sites are duly followed up, and any resulting experience is fed back into the installation procedure.

It has been agreed that this will be a person within the CLRC-RAL team.