

#### IDENTIFIER INFORMATION:

*The Project Title;* The UK Grid for Particle Physics Collaboration, GridPP2  
*The Principal Investigator and Institution;* Tony Doyle, Glasgow  
*Your name and Institution.* Removed

Information removed

#### THE AREA OF RESEARCH:

*What is the level of worldwide activity in this area?*

Grids for Particle Physics are major research areas in the US and Europe and activity is increasing rapidly in Japan, Taiwan and Korea and in South America. Computer Science funding for Grids for Particle Physics in the US is currently at a level of about \$8M per year with a matching level of effort from the Particle Physics community.

*What is the level and international standing of UK activity in this area?*

The UK has established a position as probably the strongest collaborative contributor to the international effort to deploy Grids for particle physics. This is largely due to the success of the GridPP project.

*Will the proposed project support an important area of research?*

The proposed project supports simultaneously the world's leading experiments at the frontiers of physics and the international effort to use the demands of particle physics to advance the science of data-intensive Grids. These are very important areas of research.

#### THE PROPOSAL:

*Is the science supported by the proposed project high quality/worthwhile?*

The physics supported by this project is widely recognized to be of the highest quality (the LHC program, BaBar, D0, CDF ...) and the Grid development promises to be the foundation of major economic benefits in the future.

*Are the objectives of the proposal clearly stated and soundly based?*

The proposal clearly identifies objectives in the ultimate provision of Grid-based resources for the UK particle physics program, plus more detailed objectives in the development of Grid middleware and operational procedures. The former are aligned with the needs of the major particle physics experiments in which the UK is participating. The latter are a well-chosen set of objectives that build on UK strengths and will contribute to the international effort to deploy Grid technology for particle physics. I was particularly pleased to see security and authentication as one identified focus.

*Does the proposed project have a sensible, clearly stated management and programme plan with time scales and milestones?*

The management plan has already been validated by the success of the GridPP project and will no doubt serve the new project well. The program plan focuses on the key middleware areas where the UK is likely to be able to contribute most, plus areas such as security where the need for close involvement of the computing centers requires a significant effort in each nation. The milestones for delivery of hardware, software and services are well matched to the needs of the LHC program. The proposal avoids dangerous over optimism about the rate at which Grid-based services will become mature and will be adopted by physicists.

*Novelty, originality, timeliness and likely significance to astronomy or particle physics?*

This proposal builds on the 3-year GridPP project that has already made major progress in Grid computing for particle physics. As a follow-on proposal it does not explicitly exhibit novelty or originality, but it does form part of an international research effort that is taking an energetic and novel approach to data-intensive science. The proposed work is well timed to create the Grid needed for the LHC – the world’s leading particle physics program – and will bring continued and vital benefits to currently running particle physics experiments such as BaBar.

*Do the applicants have appropriate relationships with other work in the UK and abroad?*

The project contact persons and the faculty and staff at UK universities who will be involved in the project are integrated in the UK particle physics community that is itself integrated into the worldwide particle physics community. Through their involvement in European projects (LCG, EGEE ...) and their contacts with US projects (PPDG, GriPhyN, iVDGL) the applicant have excellent contact with the international effort to deploy Grids for particle physics.

*Are reliable methods and techniques proposed?*

The Grid will involve a level of complexity much greater than that currently faced by individual computer centers. The proposed approach is sound in that it has identified the middleware components to be developed and deployed and places stress on testing and operational support. However, even though the proposal avoids excessive optimism about the rate at which Grid technology will become the normal approach to computing for particle physics, it will be a challenge to surmount complexity and reliability issues and keep the deployment fully on schedule.

*Is there any industrial relevance or potential for exploitation?*

Grid technology in general promises the creation of a huge market for services in the longer term. It is thus not surprising that businesses such as IBM are already involved. However, there is only a partial overlap between the needs of particle physics and the development of easily marketable commercial Grid services.

THE APPLICANTS:

*How well do you know the applicant(s) and by what route?*

Information removed.

*What is your assessment of the international and national standing of the applicant(s).*

The applicants are the internationally visible leaders of the UK effort in Grids for particle physics.

*What is your assessment of the past record of applicant(s)?*

The success of the current GridPP project confirms my assessment that the applicants record in planning and managing research, development and deployment involving information technology in particle physics is outstanding.

*Do the applicant(s) have a level of resources presently available to support credibility of proposal?*

This proposal clearly has the support of the entire UK particle physics community.

OTHER FACTORS:

*Please detail any other factors about the application to which you wish to draw attention.*

I was surprised to see tape-storage proposed without any reservations for volumes of data comparable to the disk-resident data. I expect the cost-effectiveness of tape storage in such cases to become very unclear in the near future. It may be possible to move towards a better-optimized storage system within the proposed funding envelope .

OVERALL SCIENTIFIC ASSESSMENT:

Alpha - 5

RESOURCES:

*Please state whether the resources requested have been justified and are appropriate to the proposed project, and proportionate to the likely scientific impact. What resource modifications would you recommend? Please pay particular attention to staffing and equipment. With any other resource modifications, which you recommend, to what extent do you regard the programme as cost-effective/value for money? Please note that applicants have not been asked to submit an RG2 form or highly detailed financial information at this stage.*

The resource planning and costing appears sound. Fortunately the "dot-com-bust" has made it once again possible to perform first-class work in computing at the proposed university and laboratory salary levels.

The proposed program is costly, but addresses unprecedented opportunities in physics and Grid computing. I regard it as excellent value for money.

SUMMARY:

*Quality of science:* Excellent

*Novelty and timeliness:* Excellent (of overall program of Grids for particle physics)

*Viability:* Excellent

*Planning:* Excellent

*Past effectiveness of applicant(s):* Excellent

*Suitability of applicant(s)/research team:* Excellent

*Suitability of department(s):* Excellent

*Cost-effectiveness/value for money (taking account of any resource modifications which you have recommended):* Excellent

COMMENTS FOR THE APPLICANT:

THE REFEREE:

*Please rate your confidence in your evaluation of the applicant and the research proposed (high/satisfactory/low), and indicate your areas of particular expertise.*  
High

IDENTIFIER INFORMATION:

\* The Project Title:

GridPP2

\* The Principal Investigator and Institution

David Britton, Imperial College

Tony Doyle, Glasgow

Steve Lloyd, Queen Mary

\* Your name and Institution.

Removed

THE AREA OF RESEARCH:

\* What is the level of worldwide activity in this area?

High: The needs of the large HEP projects such as LHC and its experiments, are clearly of huge importance to the community.

\* What is the level and international standing of UK activity in this area?

High: The UK community plays an active role in this area. GridPP in particular play a key role in the IT infrastructure projects that are focused on meeting the data and processing demands set by the new LHC at CERN.

\* Will the proposed project support an important area of research?

Yes: supporting the UK's commitment to the computing infrastructure demanded by the LHC is rated as the highest PPARC goal over the next 5 years (referring to the 2003 SC Strategy paper).

THE PROPOSAL:

\* Is the science supported by the proposed project high quality/worthwhile?

Yes.

\* Are the objectives of the proposal clearly stated and soundly based?

For the most part, yes. I do have some concerns that work in the 'applications/experiments interface area' is ill-defined, with a lack of deliverables, milestones, effort breakdown given in the application.

It is reassuring to note that GridPP2 will see the implementation of their middleware adopting OGSA standards as these develop.

\* Does the proposed project have a sensible, clearly stated management and programme plan with time scales and milestones?

For the most part, yes. Time scales and milestones will also need to be aligned with those of the emerging LCG and EGEE projects, once these

are better known (EGEE is currently being reviewed through the EU FP6 programme).

It would have been good to have seen a fuller estimate of what the situation will be post 2007, especially in terms of compute resource demands. Table 4 on page 44 provides some indication, but this only refers to the Tier 0/1 requirement at CERN.

Appendix document GridPP-PMB-24-Production\_v4.0 discusses metrics for success on page 6. These metrics are basically concerned with measuring the efficiency of the system - e.g. number of jobs submitted and processed as a function of possible jobs that could be run. There are no real metrics given that measure effectiveness. It would be reassuring to have some measure of how the system was performing in meeting the demands of those processes leading to the best science. Thus I would like to see a metric which perhaps shows the 'speed up of science output'. Perhaps some jobs leading to the best science gain massively from the use of GridPP2. Others might be only marginally improved. Effort should be focussed on the first case.

It might be useful to have a science advisory panel, made up of end users who are not necessarily core members of the GridPP2 team. It might be able to advise on issues concerning the scientific effectiveness of the project.

\* Novelty, originality, timeliness and likely significance to astronomy or particle physics?

A successful realisation of the goals of this proposal will be vital in ensuring that the UK community is fully able to scientifically exploit its investment in the LHC.

\* Do the applicants have appropriate relationships with other work in the UK and abroad?

Yes. The GridPP collaboration appears to be well integrated with appropriate activities both nationally and internationally.

\* Are reliable methods and techniques proposed?

Yes.

\* Is there any industrial relevance or potential for exploitation?

Probably. Aspects of the infrastructure that the GridPP2 consortium aim to deploy may be applicable to external groups. Additionally, GridPP members are playing a valuable role into inputting into external activities, such as the standards body at the Global Grid Forum, which in turn will enable industrial exploitation.

The proposed work with BT discussed in GridPP-PMB-18-Middleware\_v4.0 (p30 on) appears to be very valuable, especially in the area of Verification, Validation and Testing.

THE APPLICANTS:

\* How well do you know the applicant(s) and by what route?

Information removed

\* What is your assessment of the international and national standing of the applicant(s).

High. GridPP are effective in European wide initiatives such as EDG and LCG.

\* What is your assessment of the past record of applicant(s)?

Good.

\* Do the applicant(s) have a level of resources presently available to support credibility of proposal?

Yes.

OTHER FACTORS:

\* Please detail any other factors about the application to which you wish to draw attention.

OVERALL SCIENTIFIC ASSESSMENT:

Alpha 4.

RESOURCES:

a) Tier-1/ Tier-2:

A significant question concerns the relative merits of the Tier-1 and Tier-2 centres. It appears that in principle both service the same sort of jobs, i.e. it is not the case that Tier-1 has large shared memory machines for instance for specialised jobs, and Tier-2's have clusters for other jobs (see Appendix 1 of the supporting document GridPP-PMB-17-Tier2\_v4.0.).

In terms of servicing the HEP compute needs, it would appear that the Tier-2's more than cover the CPU requirement - indeed the 2007 figure (see table 13) more or less meets the demand predicted for CERN's Tier0/1 2009 demand (see Table 4).

Likewise they provide a fair fraction of the disk storage, perhaps only be weak in Tape storage. I note that the Tier-1 appears expensive c.f. the Tier-2, e.g. based on the 2007 figures I note a factor of ~x6 in cost for CPU's (noting that the Tier 1 provides 4030 CPU units c.f. 22961 from the Tier 2's at ~£2.5M c.f. ~£2.0M (if the SRIF etc funding is factored in to the Tier-2 h/w numbers)).

It maybe that the premium is reasonable considering the additional uptime commitments from the Tier-1 - but this could perhaps be further investigated.

b) Tier-2 for more general use.

Document GridPP-PMB-17-Tier2\_v4.0 - page 3 - notes that GridPP see a 'close collaboration between GridPP and the Core e-Science programme with efficient sharing of expertise and resources to the benefit of all.' This theme should be developed, shared Tier-2 resources could service a number of programmes external to just the HEP field. If resources were to be diverted from Tier-1 to Tier-2 I think there could be significant opportunities in this area (e.g. for Tier-2's to be resourced jointly with say astronomy to service some of that field's data centre storage activities.

In this context it would be useful to know how GridPP see the load usage on their compute/storage resources developing. Is it likely to be constant, or in long, short bursts? This could be compared with usage predictions from other disciplines. Perhaps an opportunity for a pilot study?

c) Applications.

The effort pencilled in for applications is fairly considerable (see table 18+19). The effort requested is an increase to 21 FTE/yr (up from 18) - and this for application interfacing only. I believe that this section is less well justified than some of the other sections (c.f. the tasks and milestones listed in the appendix document for the middleware area). Thus in Table 17 no indication of the effort request per area is given (nor is it spelt out in the associated document GridPP-PMB-19-applications\_v4.0). There is little indication of milestones and deliverables for the application interface area, with no real estimate or justification for effort per experiment. Further, it the prioritisation of application interface support should be linked to the Science Committee priorities. For instance, I see no mention of PhenoGrid and Calice in the current SC strategy paper.

On page 45, point 4, I note that the experiments will be submitting individual bids to PPARC to support applications development. It would be interesting to view these in the context of applications interface effort requested by GridPP2. Decisions should be linked, perhaps with a view to integration of those activities, and with an expectation that applications interfaces are only developed for the funded applications.

d) Middleware.

The main comment here is the increasing amount of resource that GridPP2 will be putting into middleware development. During GridPP1 significant middleware activity was funded via the EU DataGrid project - thus 27 FTEs were committed to middleware of which only 11 were PPARC funded. In GridPP the total middleware activity remains at the 28 FTE level, but now fully 19 FTEs will be PPARC funded. My concern here would then be as to why does the UK appear to have specific extra middleware requirements over and above the partners in DataGrid - who presumably will see their middleware requirements met by future EU funded projects such as EGEE? I believe that the answer to this is not well detailed in the proposal.

SUMMARY: You should rank the following Poor / Average / Good / Excellent / Unable to Judge:

- \* Quality of science will support excellent science
- \* Novelty and timeliness novelty: good  
timeliness: excellent
- \* Viability excellent
- \* Planning good
- \* Past effectiveness of applicant(s) good
- \* Suitability of applicant(s)/research team excellent
- \* Suitability of department(s) excellent
- \* Cost-effectiveness/value for money (taking account of any resource modifications which you have recommended) good

COMMENTS FOR THE APPLICANT:

- \* Are there any specific comments, issues or questions that you believe the applicant(s) should address in their response?

This is a well argued case. The applicants might wish to consider the following issues:

- a) The relative Tier-1/ Tier-2 costs - the Tier 1 and Tier 2's appear to service the experiments in the same fashion - i.e. a Tier 2 can take any job from an experiment that a Tier 1 can. The Tier 1 costs seem high compared to the cost of the Tier 2. Might it be reasonable to concentrate more resources into the Tier-2's, with the cost savings it would probably be possible to increase the response cover for some fraction of the Tier-2's and still be more economic.
- b) Is there a case to more fully develop the idea that Tier-2's could be used for PPARC/ science disciplines outside of HEP?
- c) Application Interfaces: the resource request is poorly specified, the applicants might wish to indicate where the resource (FTE's) would be deployed, perhaps with an emphasis on those experiments highly prioritised by the Science Committee.
- d) Middleware: the applicant should indicate why, even though the total PPARC + other + EU effort allocated to this area remains more or less constant from GridPP1 to GridPP2, the balance of PPARC funded effort has increased dramatically.
- e) Science use of the GridPP2 system: metrics measuring the scientific effectiveness of the system should be developed. Thus is the system optimised to speed the scientific productivity of the end user community, hopefully with an emphasis on high priority science areas. GridPP2 might see a role for a science advisory panel to help in this area.

THE REFEREE:

\* Please rate your confidence in your evaluation of the applicant and the research proposed (high/satisfactory/low), and indicate your areas of particular expertise.

I have a satisfactory-high level of confidence in my evaluation of both the applicants and the research proposed.

Information removed