DiRAC Resource Allocation Committee

Guidance Notes for Applications

Contents

1. Introduction

DiRAC is the national HPC resource for the UK astronomy, cosmology, particle physics and nuclear physics communities. Information on the HPC services that DiRAC offers can be found at www.dirac.ac.uk. This document explains the process for the allocation of time on the DiRAC facilities.

2. DiRAC Resource Allocation Committee

The DiRAC Resource Allocation Committee (RAC) has been established to oversee the time allocation for DiRAC. The RAC has two sub-panels, one for Particle Physics and Nuclear Physics and one for Astronomy and Cosmology, which will consider the proposals within their respective remits. An RAC meeting with representatives from both sub-panels will then determine the overall allocation of DiRAC time across the whole portfolio of proposals. The membership of the RAC can be seen at www.dirac.ac.uk.

3. DiRAC RAC Calls for Proposals

The RAC will issue calls to apply for time on the DiRAC facility twice a year.

Closing Date	**Allocations Start**	**Proposal Types**
3 September 2015	1 January 2016	All proposal types
3 March 2016	1 July 2016	All except Long Projects
Anytime	As soon as possible	Discretionary and Seedcorn proposals

To maximise the quality of the scientific output of DiRAC, the allocation of time for Long Projects and Short Projects will be determined via robust, transparent peer review. Discretionary and Seedcorn proposals will be appropriate for much smaller allocations of time and can be submitted at any time. They will not be peer reviewed.

4. DiRAC RAC 7th Call for Proposals (closing date 3 September 2015)

Proposals to the RAC 7th Call for Proposals should be written on the assumption that the current DiRAC 2 machines continue for the duration of the projects, including 3-year Long Projects. Due to the nature of the current funding for DiRAC, only allocations up to the end of March 2017 can be guaranteed. All allocations with a duration extending beyond March 2017 will be awarded with a firm allocation for the first year and provisional allocations for subsequent years, as has happened previously. Allocations for year 2 will be decided in 2016.

At present there is no funding confirmed for an upgrade/replacement of the DiRAC 2 machines. Should funding become available, the implications for projects with allocations

on the DiRAC 2 machines will be communicated to the project Principal Investigators as early as possible.

5. Enquiries

For assistance with making a proposal to the RAC, please contact the Project Director, Dr Jeremy Yates, jyates@star.ucl.ac.uk, or the Chair of the Technical Working Group, Dr Peter Boyle, paboyle@ph.ed.ac.uk.

6. Proposal Types

The categories of proposals are:

a. Long Projects
b. Short Projects
c. Discretionary/Seedcorn

6.1 Long Projects

It is anticipated that the majority of DiRAC compute time will be allocated to a relatively small number of Long Projects.

A Long Project is a clearly defined research programme of outstanding scientific merit which requires significant HPC resources over a 2-3 year timescale. Long Projects should cover a period longer than 12 months and up to 36 months duration. The proposed research should be world-leading, with the expectation of making step changes in knowledge through the use of DiRAC resources. Applicants must demonstrate a track record of the productive use of HPC. Long Projects may consist of a single project or a number of sub-projects, and may span multiple research groups and institutions.

6.2 Short Projects

A Short Project is a self-contained research problem typically lasting 6 months, up to a maximum of 12 months. The category will include proposals intended to develop exploratory study by users new to HPC or to DiRAC. Investigators on Long Projects may use the Short Project route to apply for resources for new sub-projects that fall outside the scope of the original, peer-reviewed Long Project proposal, provided this will not impact on their ability to deliver the research programme associated with the Long Project.

Resources will normally be allocated for a maximum of two allocation periods (i.e. a maximum of 6 months if the start date is not delayed), but applicants may exceptionally apply for up to 12 months if a strong case is made.

There is no limit to the maximum number of core hours that can be requested on a Short Project.

6.3 Discretionary/Seedcorn Proposals

These are very small allocations of DiRAC resource (up to 50k core hours) for projects that fall into the following categories:

- Scientifically outstanding projects where DiRAC resources could enable a breakthrough to be made but where the impact of the research would be lost if the project were submitted according to the scheduled calls for proposals
- Projects aimed at exploring whether DiRAC could enhance a researcher's output, or where the researcher is not already a member of an existing Short Project or Long Project

Proposals may be submitted at any time. Decisions are normally made within two weeks.

7. Resource Allocations and Start Dates

The DiRAC resources are divided into four allocation periods, starting 1 January, 1 April, 1 July and 1 October. Successful applicants will be advised of the total amount of resources they have been allocated and the periods within which the allocations must be used. Resources must be used in the allocation period to which they were assigned; they cannot be carried over to the next allocation period.

For Long Projects a uniform resource usage profile will be assumed and Long Projects will therefore receive 1/12 of their total allocation during each three month allocation period (assuming a 36 month allocation). Significant deviations from uniform profiles may be requested in Long Project proposals where scientifically justified (e.g. due to timetabled availability of data sets), but it is not guaranteed that they can be accommodated.

Each call for proposals will announce the start date for resource allocations. The call with a closing date in early March will have a start date of 1 July and the call with a closing date in early September will have a start date of 1 January. It is not possible to request start dates earlier than these dates, but it is possible to request a delayed start date. This should be specified on the application form. Start dates must be on the first day of a month.

As the DiRAC machines are not available for 100% of the time, an Availability Factor will need to be applied to any awards. This is to take into account downtime due to maintenance and technical issues, and also the ability of the scheduling system to provision the system with jobs of different sizes and run times (the latter dominates the Availability Factor). Currently the Availability Factor is 79.2% for the BlueGene/Q and 76% for all other machines. Successful applicants can reasonably expect that their project will have available to it the project's resource award multiplied by the Availability Factor in each quarterly accounting period. The Availability Factors are under regular review and it is hoped that they will increase in time.

8. Assessment Criteria

Proposals will be assessed according to the following criteria:

Primary Criteria – Scientific Excellence (weight 0.7 for Long Projects, 0.9 for Short Projects)

- Significance of the proposed research goals with reference to the STFC Roadmap
- Appropriateness of the proposed methods/codes
- Appropriateness of the requested resources
- Suitability of the investigator(s) for the proposed research
- Availability of sufficient researcher effort to carry out the proposed research

Additional Criteria – Project Management (Long Projects only, weight 0.2)

- Appropriateness and transparency of the proposed allocation process between sub-projects
- Feasibility of project timeline given DiRAC resources requested and size of investigator team
- Publication plans

Additional Criteria – Technical Feasibility (weight 0.1)

- Appropriateness of the proposed architecture/machine selection
- Justification for any programmer time requested
- Efficiency of resource usage and how well code(s) vectorise

The technical feasibility will be assessed by the DiRAC Technical Working Group.

In cases of similarly ranked proposals, applicants who demonstrate more efficient use of DiRAC resources either in terms of actual code efficiency or more efficient operational strategies will be given preference (see Code Efficiency below).

9. Assessment Process

9.1 Discretionary/Seedcorn Proposals

Discretionary and Seedcorn proposals can be submitted at any time. They will be reviewed by the chairs of the sub-panels; they will not be sent out to referees. Applicants will be notified of the outcome as soon as possible, usually within two weeks.

9.2 Long Projects and Short Projects

Proposals for Short Projects must be submitted to one of the two calls for proposals each year. Long Project proposals must be submitted to the autumn call (closing date early September for 1 January start date).

The relevant RAC Sub-panel Chair will assign an RAC sub-panel member to act as primary assessor for each project proposal received. For Short Projects the primary assessor will

contact at least two expert referees to obtain a minimum of one report. For Long Projects the primary assessor will contact four expert referees to obtain a minimum of two reports, at least one of which should be from a non-UK based referee. For both Short and Long Projects a member of the Technical Working Group will be asked to provide a technical assessment of the proposal.

Project PIs will be given an opportunity to respond in writing to referees' comments and to the technical assessment. It is strongly recommended that PIs make use of this opportunity. Each proposal will be considered at a meeting of the relevant sub-panel. At the meeting the sub-panel will grade the proposal with reference to the proposal, the referee comments, the technical assessment and the PI's response to the referee reports and technical assessment.

Following the sub-panel meetings a meeting will be convened of the full RAC (or with the chairs of the sub-panels depending on the scale of the call). At this meeting an overall ranked list of the proposals will be produced and the resources to be assigned to each proposal will be agreed. Applicants will be notified of the outcome for their proposal and provided with feedback as soon as possible after the full RAC meeting.

10. How to Submit a Proposal

10.1 Principal Investigator

Each proposal must identify a Principal Investigator (PI) who has overall responsibility for the delivery of the proposed research and will act as the contact point for all DiRAC and/or RAC communications.

In line with STFC's research grant conditions the PI must be either a) resident in the UK, or b) be employed by an overseas Research Organisation approved by STFC as eligible to apply for research grant funding. The STFC eligibility criteria can be found via this link.

We welcome proposals that would be the UK's contribution to an International Research Programme. However it is expected the proposed research programme that will use the DiRAC facility will enhance the UK's research outputs and that the users fulfil the formal STFC research grant eligibility criteria.

10.2 Discretionary/Seedcorn Proposals

Applicants should submit the application form together with a science case (maximum one page) describing the work to be undertaken and explaining the reasons for wishing to use DiRAC. For proposals which are being submitted via this route due to their being time critical, the proposal should include a clear statement of why the impact of the work would be lost by delaying until the following RAC submission deadline. Proposals should be sent to the DiRAC helpdesk, dirac-support@epcc.ed.ac.uk.

10.3 Long and Short Projects

Applicants should submit the application form together with a case for support. Proposals should be sent to the DiRAC helpdesk, dirac-support@epcc.ed.ac.uk.

Before submitting a proposal the applicant(s) must contact the DiRAC Project Director, Dr Jeremy Yates (jyates@star.ucl.ac.uk), and/or the Chair of the Technical Working Group, Dr Peter Boyle (paboyle@ph.ed.ac.uk) to discuss which DiRAC machine(s) is best suited to the project's requirements. This should be entered on the application form, but please note that it may change when the RAC agrees the final allocations as the RAC aims to make optimal use of the DiRAC facility for the overall portfolio of projects.

The case for support must be a maximum of 10 pages, font size 11 point, and contain the following sections:

1. Science justification and proposed research (maximum 6 pages)

This should describe the proposed research programme and explain why the applicants wish to use DiRAC. It should address the following assessment criteria:

- Significance of the proposed research goals with reference to the STFC Roadmap
- Appropriateness of the proposed methods/codes
- Appropriateness of the requested resources
- Suitability of the investigator(s) for the proposed research
- Justification of any periods of machine use in exclusive mode
- Availability of sufficient researcher effort to carry out the proposed research (For Long Projects a table should be included showing the estimated effort, as a percentage, that the PI and each Co-I expect to contribute to the project.)

2. Project Management (Long Projects only, maximum 1 page)

This should cover:

- The project management structure
- A description of the internal allocation process for the allocation of time to sub-projects (if applicable)
- A work plan, with milestones against which the progress of the project will be measured
- Publication plans

3. Technical Feasibility (maximum 2 pages)

This should cover:

- Appropriateness of the proposed architecture/machine selection
- Justification for any programmer time requested

- Efficiency of resource usage and how well code(s) vectorise

4. References (maximum 1 page)

11. Code Efficiency

The aim of the RAC process is to maximise the output of high quality research by the DiRAC facility. Scientific excellence will be the primary driver for allocation decisions and the RAC will balance 'time to science' against reasonable requirements on the operational efficiency of approved projects and simulation codes. It is recognised that the cutting-edge and novel nature of research across the DiRAC community means that many DiRAC codes are under active development and may not be as efficient or scalable as more mature codes. Further, results obtained in a timely manner with a sub-optimal code will often have greater impact than results delayed by extended periods of code optimisation work.

However, while it is recognised that new HPC users, or users of new codes, may not have sufficient resources or experience to provide full details of code efficiency, applicants will be required to demonstrate that their operational plan is as efficient as possible (e.g. running multiple jobs concurrently on small numbers of cores versus running jobs sequentially on larger numbers of cores) and that the architecture requested is the most appropriate for the work.

In cases of similarly ranked proposals, applicants who demonstrate more efficient use of DiRAC resources either in terms of actual code efficiency or more efficient operational strategies will be given preference. Projects may request allocations of programmer effort to assist with code optimisation. RAC may also allocate programmer effort to particular projects where concerns about efficiency have been raised.

Where scientific projects intend to migrate between architectures, RAC particularly encourage codes that have been well-optimised and scaled to larger project sizes than would have otherwise been attainable. A development queue will be available for those preparing proposals who want to demonstrate code efficiency. This queue will have strict limits on resource usage. Requests for access to this queue for new users will be dealt with by the Technical Working Group.

12. Requesting Exclusive Use of a DiRAC Machine

Some Long Projects may include (or may consist entirely of) sub-projects, which require usage of an entire DiRAC machine, or significant fraction of a machine, for a period of longer than two days. This mode of use must be explicitly justified in the proposal, and a detailed timeline for the sub-project must be included. The technical assessment of such sub-projects will include an assessment of the efficiency of machine use.

If the request is approved, a fixed start date for the sub-project exclusive usage will be agreed to enable re-scheduling of other users to other machines during the period of unavailability. Time lost due to failure to meet the approved start deadline will not be compensated.

No more than two DiRAC machines will be operating in this mode at any one time. Where this mode of operation directly impacts on another project (e.g. due to technical requirements which mean it cannot be moved to another machine) it may be necessary to sub-divide periods of exclusive use or reserve a fraction of the cores for other projects. In cases where multiple Long Project proposals with overlapping scientific goals are received, if appropriate the RAC may invite the applicants to consider merging their proposals.

13. Project Reporting

Projects will be required to submit reports at the end of the project describing facility usage, progress against objectives, achievements and publications. In addition Long Projects will be required to complete annual progress reports. Report templates will be provided.

The annual progress report for Long Projects will be assessed by the appropriate RAC sub-panel against the original peer-reviewed milestones. Confirmation of resource allocations to a Long Project in subsequent allocation periods will be conditional on the approval of their progress report. The report should include explicit justification for any significant deviations from the science proposed in the original work plan, and any delays should be accounted for. Long Projects which are deemed not to be progressing satisfactorily may have their allocations in subsequent allocation periods reduced and/or may be required to submit interim reports at six month intervals. The RAC may solicit expert referees' reports in cases where scientifically significant changes to the original milestones for a Long Project are proposed in the annual report.

14. Storage Policy

Quota - an amount of disk that you cannot exceed.

Allocation - an amount of disk that you are guaranteed to have access to.

/home - this will be small and have quotas applied of, say, 10GB. This is for storing code, key input files etc but is not a working space.

/scratch or /work - this is the main working area available to all compute nodes and is quota'd by having separate volumes/file systems, one per project, thin provisioned on the storage. The amount of space allocated to scratch can be over-allocated, hence it is a quota rather than an allocation. Demand for capacity will be managed by automated sweeping of files which have not been accessed for more than XX days where XX may be, say, 60.

/data (medium term) - this is for storing results awaiting final post processing or on which more work may be required prior to publication and transfer back to the user's own institution. These will be quota-ed to reduce the risk of significant underutilisation. Sweeping will be used here as well, as it is not a permanent archive, but will be on a much longer access time, say 9 months.

Archive: This is tape storage. Applicants must specify what data products they wish to have backed up to tape.