DiRAC Resource Allocation Committee

<u>Guidance Notes for Applications</u>

Table of 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) was established to oversee the time allocation for DiRAC project proposals. 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. A 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 is currently being refreshed and will be posted on www.dirac.ac.uk as soon as that process is complete.

3   DiRAC RAC Calls for Proposals

The RAC will issue one call to apply for time on the DiRAC facility per year.

Closing Date	Allocations Start	Proposal Types
14 November 2017	01 May 2018	All proposal types
Anytime	As soon as possible	Discretionary and Seedcorn proposals

To maximise the quality of the scientific output of DiRAC, the allocation of time for Thematic Projects (formerly known as 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 10th Call for Proposals (closing date 14 November 2017)

Proposals to the RAC 10th Call for Proposals should be written on the assumption that the current DiRAC 2 machines continue for the duration of the projects. All allocations will be awarded with a firm allocation for the first year and provisional allocations for any subsequent years. Proposals should be written in a clear and focused manner to aid the DiRAC RAC panel in managing the allocation of compute time.

5   Enquiries

For assistance with writing and submitting a proposal to the RAC, please contact the Project Director, Dr Mark Wilkinson ([miw6@leicester.ac.uk)](miw6@leicester.ac.uk) or the Chair of the Technical Working Group, Prof Peter Boyle ([paboyle@ph.ed.ac.uk)](paboyle@ph.ed.ac.uk).

6   Proposal Types

The categories of proposals are:

1. Short Projects
2. Thematic Projects (formerly known as Long Projects)
3. Research Software Engineer support
4. Discretionary/Seedcorn

6.1   Thematic Projects (formerly known as Long Projects)

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

A Thematic Project is a clearly defined research programme of outstanding scientific merit which requires significant HPC resources over 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. Thematic projects must be centred on a singular scientific theme, but can contain a small number of sub-project activities, and can span multiple ROs. They can consist of a number of institutions or groups working on a large project activity.

Applicants should consider the range of material held within submissions as very large proposals can have a detrimental effect to the peer review system due to effective justification and the need for reviewer diversity. There is no limit to the maximum number of resource hours that can be requested on Thematic projects.

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 Thematic 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 Thematic Project proposal, provided this will not impact on their ability to deliver the research programme associated with the Thematic Project.

Resources will normally be allocated for a maximum of 6 months, 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 resource 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 resource 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 Thematic Project.

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

6.4   Research Software Engineer Support

All approved projects will be given support for Code profiling and Simple Optimisation. For the 10th Call, existing and new projects, can apply for additional resources in order to undertake more extensive research software engineering support.

6.4.1   Simple Optimisation and Profiling Service

- New Projects will be required to complete the simple optimisation step within 2 months of the award.

- Existing projects will be contacted so that a similar service can be performed.
- This service will provide a profile of your code during runtime and simple optimisations will be applied to the workflow. We envisage that very few lines of code will be changed during this process.
- We will ensure that you are using an optimised version of the appropriate community code if this is applicable.

This support does NOT have to be requested in your application.

6.4.2 Requests for Research Software Engineer (RSE) Resources

You will be able to ask for further assistance from a research software engineer to carry out the tasks listed below. Please note:

- Research Software Engineer Support can only be requested for software which is relevant to existing projects which are currently running on DiRAC facilities, or for projects where computing time is requested in parallel with the research software engineer support.
- This effort is technical in nature and is NOT research support. In particular:
  - These resources are not meant to be a replacement for PDRA/Post-grad student activity. The construction of a piece of scientifically valid code is the project's responsibility and is NOT the role of research software engineers.
  - RSE's role to help to produce fast optimised code based upon work that you have already done and demonstrated is scientifically valid.
- In total 2.0 FTE of effort per year will be available. We expect a typical award to be 1-2 months of staff effort, but up to 12 months per year can be awarded if the RAC assess this to be well justified.
- Note that resource requests and science goals in DiRAC computing time applications should not depend on progress or goals you aim to achieve with the support of DiRAC Research Software Engineer Resources.
- Research Software Engineering Support will be distributed according to ranking and availability by the RAC.
- Any general or non-project specific code optimisation to community codes are not awarded through this mechanism; these requests should go directly through DiRAC and is outside the 2 FTE discussed here.

You must discuss any requests for Research Software Engineer with Prof Peter Boyle (paboyle@ed.ac.uk) and Dr Mark Wilkinson (miw6@leicester.ac.uk) before you apply for these resources. **Failure to do so will mean the request is not considered by the RAC.**

Below is a guide to the services on offer:

6.4.3 Making the code run faster
- Replacement of existing numerical libraries, functions and source code with faster variants
- Making use of SIMD vectorisation
- Elimination of cache misses and other code pipeline blockages
- If you have a workflow/code that you know works and want it to be re-written completely then that is a valid request
- If you have used symbolic logic resources to produce code, or produced a demonstrator code, to show an approach works and would like fast source code to be written, then that is a valid request. Examples of this are

(i) approaches that enable the code/workflow to run faster e.g. acceleration algorithms to explore parameter space more quickly, machine learning, accelerated lambda iteration method etc, or

(ii) codes that you have written to generate new science capability.

6.4.4 Making your code run on new architectures
- If your code needs re-writing to run on other architectures, then that is a valid request.
- Altering an OpenMP code to a MPI code is a valid request
- Altering a MPI code to become a hybrid MPI-OpenMP code is a valid request

But…

- **If you would like to see if a particular algorithmic approach would work or add new functionality and test it, that is out of scope.**

6.4.5 How to apply
Please complete a Research Software Engineer Support proposal form to request these resources. A short (half page A4) justification of the RSE effort must be provided along with a list of the activities requested to be undertaken. A Science Case for support and Technical Case should also be provided if a request for computing time is not being submitted in this current call. Please provide details of the existing award of computing time if this was allocated in a previous call and this request for RSE effort is in relation to that.

- **For the Technical Case your interaction with Prof Boyle and Dr Wilkinson will enable you to complete this section.**
- **Failure to discuss your resource request with Prof Boyle and Dr Wilkinson prior to your application will mean the request is not considered by the RAC.**
- Failure to meet the Technical Criteria will **mean that the RAC will not consider** the request for research software engineer resources.
- The Science Case should make it clear why this award will enable you to undertake internationally competitive research, advance the field significantly and generate new knowledge in the PPAN science areas.

The technical assessment criteria will be:

1. Has enough information been offered to make an assessment of whether the proposed work will achieve the required science outcomes?
2. Have all necessary pre-requisites, such as benchmarks, source code, description of work, been met?
3. Will the work plan and resource requirements deliver the required code/workflow outcomes? You will need to:-
   - Estimate how much speed up you will achieve and explain what new science you will achieve with the new resources. The latter will be in the form of smaller errors, more detailed physical resolution etc.
   - Demonstrate why you believe this can be achieved with the requested resource.

7 General guidance for applicants
- Proposals should be focused on scientifically coherent themes, but should hold sufficient technical and scientific detail.

- Proposals should include adequate detail to justify the requested allocation and should be written in a way that is accessible to the RAC panel, who may not be an expert in the given field (this especially applies to Thematic Projects proposals).
- Proposals with a large span of themes suffer as it becomes difficult to assign expert reviewers.

8   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 Thematic Projects a uniform resource usage profile will be assumed and Thematic 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 Thematic 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 of 1st November will have a start date of 1st April. However on this occasion, the closing date for the 10th call is 14th November 2017 and allocations will start on 1st May 2018. 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 DiRAC@Edinburgh and 76% for all other DiRAC systems. 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.

9   Assessment Criteria

Proposals will be assessed according to the following criteria:

Primary Criteria – Scientific Excellence (weight 0.7 for Thematic 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
- Justification for any research software engineering support requested
- Suitability of the investigator(s) for the proposed research
- Availability of sufficient researcher effort to carry out the proposed research
- Data management plan

Additional Criteria – Project Management (Thematic 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 research software engineering support 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).

10 Assessment Process

10.1 Documentation required
Applicants must submit the following documents as part of the application process:

- Application form
- Case for support
- Technical application form and Data Management Plan

The application form and case for support will be assessed by the relevant Sub-Panel. The technical application form will be assessed by a Technical Assessor and then collectively by the Technical Working Group who will assess the proposal from a technical perspective (to include criteria such as applicability to the machine) and will provide specific technical feedback for each proposal.

The Data Management Plan will be assessed by Dr Anthony Davenport, Programme Manager, and Dr Mark Wilkinson, DiRAC Director, in line with the STFC and DiRAC DMP guidelines (§15).

10.2 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.

10.3 Thematic Projects and Short Projects
The relevant RAC Sub-Panel Chair will assign an RAC Sub-Panel member to act as primary introducer for each project proposal received. For Short Projects the primary introducer will contact at least four expert referees to obtain a minimum of one report. For Thematic Projects the primary introducer will contact six 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 Thematic Projects a member of the Technical Working Group will be asked to provide a technical assessment of the proposal (as detailed above).

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.

11 How to Submit a Proposal

11.1 Submission process
All proposals, to include all of the documentation listed above, must be sent to STFC Swindon Office to the following email address: DiRACRAC@stfc.ac.uk. Proposals will then be forwarded to DiRAC.

11.2 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 point of contact 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.

11.3 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 STFC Programmes Directorate, DiRACRAC@stfc.ac.uk

11.4 Thematic Projects and Short Projects
Applicants should submit the application form together with a case for support and technical application form. Proposals should be sent to STFC Programmes Directorate, DiRACRAC@stfc.ac.uk

Before submitting a proposal the applicant(s) must contact the DiRAC Project Director, Dr Mark Wilkinson (miw6@leicester.ac.uk) or the Chair of the Technical Working Group, Prof 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 12 pages for Thematic Projects, and a maximum of 8 pages for Short Projects, font size 11 point, and contain the following sections:

1. Science justification and proposed research (maximum 8 pages Thematic Projects, 6 pages Short Projects)

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
- Justification for any research software engineering support requested
- 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 Thematic 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.)
- Data management plan

2. Project Management (Thematic 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 research software engineering support requested
- Efficiency of resource usage and how well code(s) vectorise

4. References (maximum 1 page)

12   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 research software engineer effort to assist with code optimisation. The RAC may also allocate research software engineer 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.

Requesting Exclusive Use of a DiRAC Machine

Some Thematic 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 Project proposals with overlapping scientific goals are received, if appropriate the RAC may invite the applicants to consider merging their proposals.

13   Data Management Plan

Projects are required to complete a short data management plan for all types of proposals. The completion of this section is <u>mandatory</u> and will be marked against in the review of the project by the DiRAC RAC.

PIs are requested to address all points which are found in the STFC Data Management Plan (DMP) guidelines (Link). This includes: which communities may have potential interest in the data being produced by their DiRAC project; how and where the data from the project will be stored; how long the data will be stored and what metadata processes will be used.

Applicants are encouraged to consult the DMPs of any collaborations they are a part of and to discuss with their universities data management units.

14   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 Thematic Projects will be required to complete annual progress reports. Report templates will be provided.

The annual progress report for Thematic Projects will be assessed by the appropriate RAC sub-panel against the original peer-reviewed milestones. Confirmation of resource allocations to a Thematic 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. Thematic 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 Thematic Project are proposed in the annual report. In the event the last annual report falls within six months of a final report deadline, the annual report can be integrated into the final report.

Failure to submit a satisfactory project annual report or project final report will result in subsequent proposals from the group being marked down.

15   Storage Policy

Quota - an amount of disk that you <u>cannot</u> 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. It is left to the users to manage their data within the limits of the set quota. However each DiRAC site reserve the right to sweep of files which have not been accessed for a given amount of time. This is at the discretion of each site."

/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.