Project Scope and Plan – Extreme Scale Access

|  |  |
| --- | --- |
| **Project title** |  |
| Research field |  |

**Principal Investigator (PI)**

|  |  |
| --- | --- |
| Title (Dr., Prof., etc.) |  |
| First (given) name |  |
| Last (family) name |  |
| Organisation name |  |
| Department |  |
| Group |  |
| Country |  |

**Co-PIs/Team members (same information)**

Provide the details of any Co-PIs and Team members in the project, including Title, First (given) name, Last (family) name, Organisation, Department, Group and Country.

**It is mandatory to include all team members on the online form.**

**Co-Principal Investigator (Co-PI)/Team member**

|  |  |
| --- | --- |
| Title (Dr., Prof., etc.) |  |
| First (given) name |  |
| Last (family) name |  |
| Organisation name |  |
| Department |  |
| Group |  |
| Country |  |

**IMPORTANT NOTICE**

All of the sections and subsections below **MUST BE COMPLETED** (unless stated otherwise). In case you wish to leave a section empty, please provide a reason. The Access Resource Committee (ARC) will not be able to evaluate proposals that neither provide the requested information nor a justification for the lack of such information for each section.

Applicants are strongly encouraged to **base their proposal on reliable benchmark data obtained on the target machine(s)** from previous calls or access programmes. Such data and support to properly collect these can be obtained from the EuroHPC Benchmark Access call. In order to have the necessary data on time, please submit your Benchmark proposals **at least 1 month before the submission deadline**. For additional questions, please contact the Peer-Review office at access@eurohpc-ju.europa.eu.

**The structure and formatting settings of this template must be preserved and respected** (change in font size or margin and spacing settings are not allowed). The maximum number of pages allowed is **20 pages**, including graphs, tables and references, but not counting the cover page. Reviewers will be instructed not to consider any pages out of the limit. **Instruction paragraphs and picture examples can be removed from the proposal text.**

**Upload a single document**, based on the present template, in PDF format **without exceeding** 8 MB.

**Proposals that do not follow the template or that are incomplete will be administratively rejected and will not be further evaluated.**

Key scientific/societal/technological contribution of the proposal (200 words max.)

Outline the scientific/societal/technological importance of your project, how High Performance Computing (HPC) will help you achieve your goals and what the major expected outcomes are. This section would typically be the same as the abstract of the proposal in the submission form.

<Enter your text here>

Detailed proposal information (Maximum 14 pages, graphs and tables included)

The information should be understandable to peer-review experts in your discipline. It must also have appropriate information for a broader audience as your proposal will be evaluated by a Panel and in parallel with proposals in other disciplines.

Justification for the importance of the scientific problem and the requested resources (~2 pages)

Describe the proposed research and the main scientific/technical advances you will achieve with the requested EuroHPC allocation. For industrial applications, proposals should demonstrate the innovation and industrial impact on the specific market and the broader socio-economic impact. For public sector applications, the proposal should demonstrate the innovative aspects of the applications, the expected societal impact, and how the application will contribute to the delivery of quality and efficient public sector services. The justification of the requested resources must be clearly linked to the software performance evaluation (Section 2.6).

<Enter your text here>

Overview of the project (~4 pages)

Describe the motivation, objectives and scientific challenges of the problem. Describe and justify the choice of computational methods. State the advances that will be enabled through the requested EuroHPC Extreme Scale Access award (e.g., impact on community paradigms, valuable insights or solving a long-standing challenge, new technology/therapy, etc.). Provide a list of expected outcomes of your proposal and, if relevant, the interdisciplinary value of your proposal.

<Enter your text here>

Validation, verification, state of the art (~2 pages)

Please describe the validity of the simulations and predictions made with this proposal. In case you provide references to relevant publications please include here also the key relevant results. Please address issues of reproducibility and highlight the predictive capabilities of your simulations.

Validation & Verification

Please summarize the validation of your model against experiments or other established reference data. Please also provide how the numerical consistency and stability of your computational method has been verified or provide evidence of existing verifications.

<Enter your text here>

Comparison with state of the art

Place the project in the context of competing work. Explain the relative advantages AND drawbacks of your approach.

<Enter your text here>

Software and Attributes (~2 pages)

(Please see also **Examples of Performance Reporting** in Section 2.6.2.1). Describe the software that will be used including a discussion of the state of the art in the field. The description should mention:

Software

Please describe all codes you are using in the proposal. Justify your choices and describe alternatives (if any).

<Enter your text here>

Particular libraries

Describe particular libraries required by the production and analysis software, algorithms and numerical techniques employed (e.g., finite element, iterative solver), programming languages. Please specify requirements for compilation or build environment (build system (e.g., cmake, python version), version control system (e.g., git, subversion) etc.).

<Enter your text here>

Parallel programming

Model(s) used (e.g., MPI, OpenMP/Pthreads, CUDA, OpenACC, etc.).

<Enter your text here>

I/O requirements

I/O requirements (e.g., amount, size, bandwidth, etc.) for execution, input files, restart and other output. Describe I/O strategy (number of files, frequency, read/write size) and I/O behaviour of your code during the period of calculations. Concerning the number of files, please indicate the largest associated number of files on the system at any given time also taking into consideration the temporary files. Pay attention to the system limits while indicating the number of files, refer to the Technical Guidelines document for exact values per system. Please specify the restart overhead, not only for I/O; (e.g., a code may have to perform a costly domain decomposition first).

<Enter your text here>

Data: Management Plan, Storage, Analysis and Visualization (~1 page)

Data Management Plan covering

Data Management Plan covering both short-term and long-term aspects, including needs for I/O bandwidth, number of files and input/output data volumes. Please specify for which system the data will be provided, how long the data must be stored at the computing centre after the termination of the project, how it will be moved from the centre, and how subsequent analysis will be performed. Please specify the availability of both code and data to other researchers, and how this will be handled. EuroHPC JU should be given credit for all data produced through EuroHPC JU allocations when publishing, and described in the provenance when depositing to other infrastructures.

<Enter your text here>

Project workflow

Project workflow including the role and timeline of data analysis and visualisation identify where the analysis will be done and any potential bottlenecks in the analysis process. Describe any analysis and/or data reduction tools used.

<Enter your text here>

Software workflow solution

Software workflow solution (e.g., pre- and post-processing scripts that automate run management and analysis) to facilitate this volume of work.

<Enter your text here>

I/O requirements

I/O requirements (e.g., amount, size, bandwidth, etc.) for data analysis and visualisation. Highlight any exceptional I/O needs. Please provide data for (one or several) precise systems that will be simulated.

<Enter your text here>

Performance of Software (Maximum 3 pages)

Testing of your code on the requested machine

It is strongly recommended that your production code is tested in the requested machine(s) (see also the text referring to Benchmark Access in the Important Notice at the top of page 1). Please specify the EuroHPC Benchmark Access project (if any) or other projects (previous PRACE Calls, national calls, etc.) used to prepare the Extreme Scale Access proposal. If the preparatory host machine is different from the target machine, please specify why you think the data is relevant. In the latter case, please report briefly the conversion factor (in terms of ratio of time to solution, flops or requested node hours) from the machine where the preparatory tests were performed to the requested system. Moreover, your proposal must account for all technical constraints and requirements of the targeted machine(s) as documented in the separate Technical Guidelines for Applicants document; failing to do so will significantly increase the risk that your project will be technically rejected.

<Enter your text here>

Quantify the HPC performance of your project

The presented data must be representative of the entire workflow of the project proposed and refer to the main application code you intend for the production work. The software scalability data (please see **Examples of** **Performance Reporting** below) must be used to choose the most efficient job size(s) for the simulations planned in production: the corresponding software performance must be clearly linked to the justification of the computing resources requested. The ARC will not accept estimates based on related codes and/or data related to parts of your production. All data must refer to the targeted systems in your production runs or a system with comparable size, software stack and with the same architecture, and network (e.g., a project can be accepted on LUMI-G if it was benchmarked on another GPU machine with the same AMD GPU). Please coordinate with the centres if in doubt about the portability of your code. Please specify that performance results are reported on the basis of one of the following: whole application including I/O; whole application except I/O; kernel only; other (please specify). More specifically you must include:

#####  Strong and weak scalability

Starting with the minimum size of the computer necessary to run the problem (usually 1 core or 1 node). Justify the minimum size for your scaling if it is larger than 1 core or 1 node (e.g., memory limitations). Please provide a justification in case that either the weak (e.g., study of one particular bio-molecule) or the strong (e.g., study of an ensemble) scalability metric is not considered relevant to your project. Please see **Examples of Performance Reporting** below for the requested format.

<Enter your text here>

**Examples of Performance Reporting**. For the weak and strong scaling please start with the minimum and finish with the maximum number of cores that are suitable for your application. Please mark the number of cores that you expect to perform the main load of your work. On the Y axis you may use time to solution (scaled or otherwise) or speedup with respect to the minimum number of cores. **The table with the timings is mandatory**.

The table should include the speedup and the parallel efficiency. Log/log plots are useful to span many orders of magnitude.



<Enter your text here>

#####  Precision reported

One of: single precision, double precision, mixed precision. Only the precision you use in the simulation is relevant.

<Enter your text here>

#####  Time-to-solution

Time-to-solution as normalised/averaged per iteration,

$$T\_{i}^{\*}=\frac{t\_{i}∙N\_{c}}{N\_{e}}$$

AND the normalized total time to solution

$$T\_{f}^{\*}=\frac{t\_{f}∙N\_{c}}{N\_{e}}$$

with $t\_{i}$ the time per iteration, $t\_{f}$ the total time to solution, $N\_{c}$ the number of cores and $N\_{e}$ the number of computational elements (size of the problem).

IMPORTANT: Justify the choice of your code (e.g., comparison with existing codes, methods or any other scientifically rigorous argumentation). See also the text referring to Benchmark Access in the Important Notice at the top of Page 1.

<Enter your text here>

#####  System scale

One of: results measured on full-scale system, projected from results of smaller system, other (please specify).

<Enter your text here>

#####  Measurement mechanism

One of: timers, FLOP count, static analysis tool, performance modelling, other (please specify).

<Enter your text here>

#####  Memory usage

Please specify requirements per node or core depending on the size of the problem.

<Enter your text here>

#####  OPTIONAL: Percentage of available peak performance

Please collaborate with the Centres on obtaining this information (see also the text referring to Benchmark Access in the Important Notice at the top of Page 1). Alternatively provide code specific metrics for the requested machine (FLOPS, etc.).

<Enter your text here>

Milestones (quarterly basis) (Maximum 1 page)

Goals and milestones should articulate simulation and developmental objectives and be sufficiently detailed to assess the progress of the project. It is especially important that you provide clear connections between the project’s overarching milestones, the planned production simulations, and the compute time expected to be required for these simulations. Please clarify any dependencies of milestones on other milestones. Please ensure that the node hour consumption is regular throughout the allocation or provide a requested schedule after consultation with the centres.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Run Type** | **Code(s)** | **No. of runs** | **No. of nodes** | **No. of steps per run** | **Time per step(s)** | **Total node hours** |
| A (init. condition prep.) | Code 1 |  |  |  |  |  |
| B (low resolution) | Code 2 |  |  |  |  |  |
| C (high resolution) | Code 2 |  |  |  |  |  |
| D (post processing) | Code 3 |  |  |  |  |  |
|  |  |  |  |  |  |  |

<Enter your text here>

Gantt Chart

Provide a Gantt Chart of the simulation plan in production indicating job sizes and scheduling of computing tasks highlighting the consumption peaks. The Gantt Chart should also include a data transfer plan and a communication plan, strategy and timeline for the dissemination of the results.

Please take into consideration that the homogenous consumption is encouraged.



<Enter your text here>

Personnel and Management Plan (0,5 page)

What personnel are already in place and what are their roles on the project? If applicable, describe (i) personnel that will be hired for the project in the future and their responsibilities and (ii) potential personnel turnover that may occur during the project and a strategy for replacing them. The EuroHPC Extreme Scale Access call welcome proposals from individual PIs or teams of collaborators. Outline the focus of each individual or subgroup and their interrelationships.

**It is mandatory to include all team members on the online form.**

<Enter your text here>

References (Maximum 30)

<Enter your text here>

Confidentiality (0,5 page)

* Is any part of the project covered by confidentiality? **yes/no**

If **YES**, please specify which aspect is confidential and justify (Maximum 500 words):

<Enter your text here>

* Does your project involve handling of personal data? **yes/no**

If **YES** please confirm to take care of the data controller’s responsibilities as described in applicable data protection legislation – and note that you will need to make a data processing agreement with the production site.

<Enter your text here>