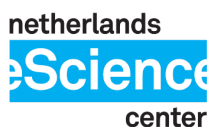


Call for Proposals

Collaborations to prepare Europe's Weather and Climate Models for pre-exascale systems (ESiWACE3-S1-2024)

2024-2025

A joint call for proposals by the Netherlands eScience Center, Eviden, Barcelona Supercomputing Center and Deutsches Klimarechenzentrum



EVIDEN



Part of the Services provided by the ESiWACE3 Center of Excellence
in Simulation of Weather and Climate in Europe

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Spain, Netherlands, Germany, Sweden, Finland, Italy and France, under grant agreement No 1010930.



Co-funded by
the European Union



Contents

1 Introduction	3
1.1 Background	3
1.2 Purpose of this call	4
1.3 Available budget	5
1.4 Validity of this call	5
2 Guidelines for applicants	5
2.1 Who can apply?	5
2.2 What can be applied for?	6
2.3 When can applications be submitted?	6
2.4 Preparing an application	7
2.5 Specific conditions	7
2.6 Submitting an application	8
3 Assessment procedure	9
3.1 Procedure	9
3.2 Admissibility and assessment criteria	9
4 Contact details	11
4.1 Contact	11
5 Acknowledgement	11
6 Disclaimer	11

1 Introduction

1.1 Background

Earth system models (ESMs) are typically large applications that are subject to ever-increasing complexity, with the aim to resolve more physical processes and represent the physical world with higher fidelity. At the same time, there are significant investments to adapt the software to the latest hardware architectures. Generally, these improvements go hand in hand, as higher accuracy often only can be achieved through increased resolution or larger ensemble sizes, and thus increasing the required computational resources. Where the model development is driven by scientific expertise, the performance optimization is driven by the desire to use compute resources efficiently and at sufficient scale.

After a long period of relative stability, the computing infrastructure landscape is rapidly evolving and diversifying, with more change on the horizon. Adapting existing software to the latest computing platforms and exploiting the full capability of the architecture requires specific expertise of performance engineering, hardware accelerators and programming models. As such, continued development of weather and climate models requires new collaborations between experts from different fields.

This call for proposals aims to support the exascale preparations for the ESM community in Europe. We create short collaboration projects that provide engineering time, guidance and knowledge transfer to improve model efficiency or port models to existing and upcoming computing infrastructures. All groups developing and maintaining ESM codes - not only the ESIWACE3 partners - can apply. Proposals will be subject to a review regarding technical feasibility by the service providing partners –being the Netherlands eScience Center, Eviden, BSC and DKRZ– and undergo a review more focussed on scientific impact by an external committee. When found eligible, the project will be granted in-kind contribution by one of these partners (or a combination thereof).

The **Netherlands eScience Center** is the Dutch national center of excellence for the development and application of research software to advance academic research. The eScience center employs a large team of eScience Research Engineers, who are experts in state-of-the-art computational methods and technologies and have a keen interest and experience in developing research software. The center harbors unique expertise on programming for heterogeneous hardware architectures, optimizing software for supercomputing platforms and deploying technologies in the often extensive fortran codes that govern the domain of atmospheric physics and oceanography.

The **Eviden** team in charge of the ESIWACE3 services, part of the Center for Excellence in Performance Programming (CEPP) team, is composed of experienced HPC experts involved in collaborative European projects, such as ESCAPE and ESIWACE in the weather and climate domain, but also in other domains. The core expertise of the team is HPC applications profiling and tuning on a variety of high-end HPC systems including heterogeneous systems, memory accelerators, etc. Thanks to Eviden's position as European leader in HPC the team has access to a wide set of state-of-the-art HPC hardware, software, tools and knowledge.

The **Deutsches Klimarechenzentrum (DKRZ)** is the German national high-performance computing and data management center dedicated to climate and Earth system research. DKRZ provides state-of-the-art

HPC resources, and provides support for users to optimize their code and utilize these resources efficiently. The institute also has a broad portfolio of maintained tools and applications in the weather and climate domain and manages and provides access to a vast amount of climate and Earth system data, including observational data, model output, and reanalysis data. DKRZ develops and offers tools and services for data management, analysis, visualization, and publication, to support scientists in their research.

The **Barcelona Supercomputing Center (BSC)** is a well-known leading supercomputer facility and software consultancy agency and host of Mare Nostrum, one of the most advanced supercomputers in the world. The BSC Earth Science department focuses on developing and implementing global and regional models and data solutions for air quality, climate predictions and their applications. Its HPC team has unique knowledge of profiling and optimizing multi-component MPI applications, removing I/O bottlenecks and accelerating algorithms using GPU's.

1.2 Purpose of this call

This call for proposals aims to support the ESM-developing communities in Europe with preparing their simulation codes for (pre-)exascale computing platforms. The goal is to improve model efficiency and readying the software to enable model execution on existing and near-future hardware architectures. The goal of this effort can be enabling simulation experiments at unprecedented grid resolutions or ensemble sizes, or include computationally expensive physical processes that were previously unfeasible. It could also be to achieve a smaller time-to-solution or energy-to-solution for current state-of-the-art simulations.

Within this service, we expect to establish collaborative projects in which the team of ESIWACE experts provides advice, knowledge transfer and engineering effort to the applicant. A competitive proposal aims to have a long-lasting impact on (the main branch of) the source code of an established atmospheric or ocean model, or any other ESM component, to progress its capability of efficiently and effectively using modern computing hardware. To ensure such impact, there is a strong preference for models that are developed under sustainable and re-usable software engineering practices and are developed as open source projects with a permissive open source license.

We encourage proposals to target accelerator hardware such as GPUs, which will most likely constitute the bulk of the capability of future exascale systems. In such cases, the proposal should aim at porting the critical path of the code to the accelerator and tuning the resulting code for maximal throughput and/or higher energy efficiency on the device. We also welcome proposals on improving the performance of existing GPU-enabled codes, for example through exploring the use of mixed-precision. Together with the model developers, we will seek a balance between achievable speedup and sustainability of the solution.

Proposals for this call may also target traditional CPU architectures or many-core variants. In such cases, the effort may be focussed on e.g. improving cache efficiency, stimulating the use of SIMD instructions, increasing floating-point instruction rate by moving to lower precision numerical representation, implementing fine-grained or task-level parallelism, detection and elimination of MPI bottlenecks or reducing the pressure on the storage system.

Apart from offering limited engineering time and implementing code modifications, the projects principally aim to provide guidance and establish knowledge transfer between HPC experts and modeling groups. We believe that code transformations can only have a lasting impact if the software maintainers have sufficient expertise in the technologies and techniques that are being used.

1.3 Available budget

In this call, each grant consists of in-kind support in the form of a team of skilled Research Software Engineers (RSEs) employed by either the Netherlands eScience Center, Eviden, DKRZ or BSC. The RSEs will work remotely on the project for **up to 6 person months (PMs)**, to be spent over a period of **maximally one year** after the project start. The hours will be planned flexibly and in mutual agreement with the applicant. Included in this time is the dissemination of the work, both internally in e.g. ESIWACE3 project meetings and externally through conference and workshop presentations.

Given the total budget, up to six proposals can be granted as part of this call. After submission, the first round of grant decisions will be announced in **July 2024** and the applicants of awarded proposals will be contacted.

1.4 Validity of this call

This call is open from **March 4th 2024, 14:00 CET** until **June 3rd 2024, 14:00 CET**. After submission, a technical review is carried out focussing on technical feasibility, state-of-the art technology and expected sustainability of the proposed work. This review is done by the service providers, and applicants will receive a notification of the result within a month after their submission. The proposals that have passed the technical review are sent to the external scientific review committee for review, where the final selection is made based on the scientific merit and impact. Proposals submitted before **June 3rd 2024, 14:00 CET** will be considered for funding starting in September 2024 and be fully reviewed by **July 13th 2024**. Note that the ESIWACE3 consortium reserves the right to update or change this call for proposals, including all key dates, after each review moment.

2 Guidelines for applicants

2.1 Who can apply?

This call addresses all European ESM developing groups, including atmospheric sciences, oceanography, and climate-related domains such as land or sea ice modeling, land-surface modeling, atmospheric or ocean (bio-)chemistry, etc. The call includes, but is not restricted to global models, and therefore regional models or small-scale very high-resolution applications such as atmospheric mesoscale models are eligible as well. Members of the ESIWACE3 consortium are allowed to apply for this call, but external groups have a strong preference. The applicant's institution needs to be located in one of the countries participating in the EuroHPC Joint Undertaking.

We aim to have a permanent impact on the source code of the application, and this is most likely achieved if the applicant is part of the core development team of the software; derived 'downstream'

versions or user-specific branches will have a smaller chance of being granted, as we wish not to contribute to the fragmentation of the European ESM landscape.

Applications that have already been awarded a project in calls for projects within ESIWACE2 Service 1 will have lower priority in this call.

2.2 What can be applied for?

The grant consists of **up to 6 PMs** of in kind consultancy, engineering and knowledge transfer time by one of the expert RSEs on high-performance and heterogeneous computing. These engineers will work remotely in small teams on the project and fulfill the hours within 12 months after the project starts. Each team will have a lead RSE who is the main contact point for the model developers. Throughout the service projects, the HPC experts are expected to work closely with the model development group to achieve ambitious goals in a short-lived project and establish the necessary transfer of knowledge to ensure maintainability of the developed code. Hence this requires active involvement from both sides. From the applicant we therefore expect

- assigning (or being) a contact point from the model development group that actively participates in the development process and display an open, collaborative attitude towards the ESIWACE team; knowledge transfer can be pursued by reviewing code, co-designing acceleration strategies, organizing virtual pair-programming sessions and helping with validation of the model output,
- guidance of the team of HPC experts by giving them access to the appropriate code repositories, providing representative test cases, describing validation procedures and coding standards, and helping getting access to specific hardware platforms (if requested in the proposal),
- co-leading a dissemination effort describing the work done. This can be a conference presentation or poster or (preferably) a journal paper,
- filling in a short evaluation of the provided service after the hours are finished.

Because the work happens on a project basis, there is no available budget for long-term maintenance of the contributed code. The long-term maintenance will be the responsibility of the applicant's modeling group. The contributed code will be documented according to standard good practice, or the modeling group internal coding standards.

Finally, the work will be disseminated in conferences and workshops by the RSEs, offering exposure for the applicant and promotion of the software. It is strongly encouraged that the applicant and the RSEs together promote and publish about the work carried out in the project. These efforts are considered part of the collaboration, and the time for RSEs to spend on these activities is part of the in-kind budget of the grant.

2.3 When can applications be submitted?

Proposals can be submitted in the period **March 4th 2024, 14:00 CET to June 3rd 2024, 14:00 CET**. Submitted proposals will be considered for funding starting in September 2024.

Further information about the assessment procedure is found in Section 3.

2.4 Preparing an application

A proposal can be submitted by completing a google form within the above time frame. The form consists of three sections addressing the following requested information:

1. A brief overview of the model, its scientific impact and societal relevance. An explanation and underlying motivation of the deployment of the model at the exascale.
2. A brief description of the targeted platform or hardware architecture. The limiting factors to the large-scale deployment of the application onto the target platform and some ideas about what performance gains can be expected (the latter may be somewhat speculative). A list of components or routines that need to be ported or optimized to achieve this speedup.
3. Plans for the sustainability, dissemination and maintenance of the developed software.

Applications must be completed in English.

2.5 Specific conditions

We expect the applicant to either provide or adopt basic infrastructure for remote software development such as a version control system, a platform to keep track of issues, progress and discussions (e.g. jira, github, ...) and to provide representative test cases for monitoring numerical results and performance of the program. In general, the applicant should display an open, collaborative mindset and is expected to give feedback to the RSEs during the development cycle. This means that the applicant is responsive to emails and announces periods of absence ahead of time.

The target hardware platform, together with all necessary compilers, firmware and performance measurement tools, has to be available to the engineer during the development process. Access to the target platform can be established either through (i) a joint external (pilot) project call to a European facility, (ii) an internal EuroHPC allocation of the ESIWACE3 project, (iii) permission for some internal facility given by the applicant to the RSEs, or (iv) internal access of the RSEs to private or national supercomputing facilities. There is no guarantee from the service provider that the latter option exists or has a sufficient budget for benchmarking. In case the applicant arranges access for the RSEs, we expect sufficient resources (computing budget, disk space) and permissions to install the required tools.

If there are any visits of the RSEs to the applicant's institute, the applicant is responsible for providing the RSE with a suitable, safe, and healthy working environment. This includes providing an environment that is harassment free and free of unwanted behavior, for more information see [code of conduct](#).

The applicant will write a short report to describe the progress that has been made by the collaboration, the impact on performance or portability of the code, and the impact on the scientific use cases of the code. This report is due in February 2025 for projects started in September 2024. The specific format for this short report will be distributed in advance.

When the above conditions are systematically violated, the ESIWACE3 consortium may choose at any moment to stop the project funding. In particular, the first two months of projects are considered to be a trial period, in which the effectiveness of the collaboration is evaluated and if deemed not effective the project can be terminated.



2.6 Submitting an application

The online application form will be made available on the ESIWACE website.

3 Assessment procedure

3.1 Procedure

First there will be a technical review of the proposal by one of the service providers. This review will not assess the scientific impact, but rather focus on the technical feasibility of the project: is the targeted performance improvement realistic, is the code practically portable to a new architecture within the timeline of the project, and can the underlying numerical algorithm be accelerated? We reserve the option to invite the applicant for a remote consultation meeting as part of the technical review process. The outcome of the technical review is binding, only a selection of applications will pass to the second round for scientific review.

The final decision on the grant will be made by an external scientific review committee of at least three renowned scientists in the fields of weather and climate, with affinity for HPC. We will avoid reviewers that are direct colleagues of any of the applicants.

3.2 Admissibility and assessment criteria

A proposal is considered admissible if the applicant is affiliated with a European non-commercial research institution, fulfills the description of 2.1, and there is no conflict of interests between the proposed project and the service providers. Furthermore, the code needs to have a license and the first-round reviewers (and eventual RSEs) need to be able to access the source code as part of the review process. Finally, the intended hardware platform has to exist within Europe, and obtaining access by the engineers to the hardware should be plausible.

Please note that applications that were already awarded a project within ESIWACE2 Service 1 will have lower priority in this call for projects in 2022.

The first round technical review will focus on the following criteria:

1. Technical feasibility: are the proposed code changes doable within the 6 PM time frame? Is it plausible that the resulting code will run faster on the proposed target platform?
2. Technological state-of-the-art: are the applicants aware of the current state-of-the-art algorithms and techniques in their domain? Is the target hardware likely to be part of future European exascale machines and does the proposed work bring the application closer towards exascale deployment?
3. Software sustainability and lateral impact: does the proposal provide a credible pathway towards sustainable software? Will the proposed changes be merged to the main branch of the model, and is this code sufficiently maintained by a community?

The second, decisive review round focuses on the scientific impact of the project, such as

1. Scientific quality: does the proposed work enhance the quality of the model output, does the increased performance of the model allow its users to reach unprecedented accuracy or enable scientific breakthroughs?
2. Scientific novelty: does the proposed work allow the model community to explore new territory, use the software in other scientific applications?



3. Lateral impact: does the proposed work have scientific impact in other areas or communities?
Are there possibilities for generalization of the contributed software?

4 Contact details

4.1 Contact

If you have specific questions about this call for proposals and the assessment procedure, please contact:

Open calls at eScience center:

Email: open-calls@esciencecenter.nl

For questions about the ESIWACE2 Centre of Excellence in Simulation of Weather and Climate in Europe, please contact:

ESIWACE3 Consortium

Email: esiwace3-communication@bsc.es

For questions about the Netherlands eScience Center in the context of this call, please contact:

Dr. Ben van Werkhoven, Senior Research Engineer

Tel.: +31 20 4604770

Email: b.vanwerkhoven@esciencecenter.nl

For questions about Eviden, please contact:

Dr. Erwan Raffin, distinguished HPC expert

Tel.: +33 2 56 85 04 24

Email: erwan.raffin@eviden.com

5 Acknowledgement

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Spain, Netherlands, Germany, Sweden, Finland, Italy and France, under grant agreement No 1010930

6 Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European High Performance Computing Joint Undertaking (JU), Spain, Netherlands, Germany, Sweden, Finland, Italy and France, . Neither the European Union nor the granting authority can be held responsible for them.