

University Large Allocation Request Preparation Instructions

The next university deadline for submitting Large Allocation Requests is March 12, 2024.

Note: In addition to Large Allocation Requests, CISL offers opportunities for NSF awardees, graduate students, and postdocs to request smaller allocations. See the [University allocations](#) page for eligibility requirements and other information.

Page contents

- [Submitting a Large Allocation Request](#)
- [Request Summary document](#)
 - [Computational experiments](#)
 - [Resource requirements](#)
 - [Additional supporting information](#)

Submitting a Large Allocation Request

Requesters must prepare and submit a Request Summary document and attach it to the **Large Allocation Request** form as a PDF file that is no larger than 5 MB.

The instructions below explain how to prepare the document. When your document is ready, submit your request via the [Large Allocation Request](#) form.

Request Summary document

The Request Summary document must provide a self-contained description of your project and allocation request. The document may be no more than five (5) pages for Sections A–D below; Sections E–G should be included in the same uploaded document and are permitted an additional five (5) pages. The **five-page limit is mandatory** for all requests, and it is **strongly recommended** that you follow the template below to help the panel locate required information in your request.

The [CHAP Allocation Review Criteria](#) provide further detail on considerations the review panel uses to identify meritorious requests. Recent successful requests are available: Click to download [Proposal Sample 1](#) and [Proposal Sample 2](#).

Due to the rapidly growing scale of data generated by many university projects and constraints on the available storage within the CISL environment, the CHAP is scrutinizing user requests for storage resources much more closely. As with poorly justified requests for computing time, poorly justified requests for storage resources will result in reduced storage allocation awards. Review the guidance below for Section D and ensure that your submission addresses the relevant points supporting your efficient, effective, and appropriate use of CISL storage resources. Keep in mind that CISL storage resources are typically available to university projects only for the lifetime of the allocated project and associated NSF award; see [Campaign Storage](#).

A. Project information

- Title of project
- Title of NSF award (if different from project title) supporting the computational experiments. **Important:** The NSF award must explicitly support the computational experiments being proposed.
- NSF award number. Due to high demand for resources, the CHAP no longer reviews requests for pending awards. Submit your request only after the NSF has awarded a grant for the research.
- Name of Project Lead and his or her institution
- Submission date

B. Project overview

The overview of the project should typically be less than half a page and include:

- A summary of the science question and computational plan.
- The relationship of the proposed work to atmospheric and closely related sciences.
- The explicit linkage between the NSF award and the computational experiments being proposed. This is especially important if the published NSF award abstract does not clearly describe the computational component of the work funded.

C. Science objectives

The science objectives should be described *briefly*. This section should give sufficient information for understanding the computational plan in section D; it is not necessary to justify the science objectives as they must have already passed NSF review.

Advice for preparing your request. "Brief" is the operative word for your science description. The panel is not judging your science, only trying to understand how and if your computational experiments (described in Section D) will help you find answers to your science questions. This section should be between 1/2 and 1 page long.

D. Computational experiments and resource requirements

The bulk of the Request Summary should focus on Section D. Discuss your planned computational experiments and the resources needed to conduct the work in this section. Please cover these topics and follow the advice for how to best address each topic.

Computational experiments

- **Numerical approach.** Briefly describe the numerical approach(es) or model(s). If a standard community model is being used, simply explain why it is appropriate for the scientific objectives and include a reference to a published description of the model or method. If a community model is being modified, include a description of the modification sufficient to explain any changes in the computational cost of the model and explain why modifications are necessary for the scientific objectives. For a non-standard or non-community model, the numerical description should briefly describe the approximations and other methods proposed to obtain valid solutions to the problem.
- **Computational experiments.** Describe the computational experiments needed to address the scientific objectives. Clearly indicate the number and type of experiments and how they relate to the objectives. Be sure to justify the model configuration choices that are relevant to the experiment, such as domain, grid size, time steps, simulated time span, ensemble size, and parameter choices. References on the selection of the ensemble size are strongly encouraged. Without an adequate justification of the model configuration, the panel may reduce or deny your computing request.
- **Code performance.** Include documentation on program code performance (for example, timings, performance monitoring tools). You may refer to a web page detailing code performance. Describe how flexible your code is in the number of processors it can use and why you may choose a particular number. The CHAP will evaluate the likelihood that a request can scale up its production runs based on this information. Information on the portability of the code to other platforms may also be useful to the CHAP; requesters are strongly encouraged to provide this information about the code and the team's knowledge/use of HPC systems similar to Cheyenne.
- **Output and analysis.** Describe the key variables to be output, plans for analysis, and any steps taken to use storage resources efficiently and effectively. Projects with large-scale data output will be scrutinized in greater detail. In addition, the section should include a brief description of the post-processing calculations to be carried out and what output needs to be retained beyond the post-processing stage.

Resource requirements

Provide a table summarizing the resources required for each experimental configuration and the complete set of computational experiments. This must include the number of core-hours needed and the data output volume. The table should be accompanied by a narrative that elaborates on how you arrived at the numbers in the table and describes how you will use appropriate storage options or data analysis and visualization resources as detailed below.

With the petascale systems' ability to generate vast amounts of data, CISL allocation requests require users to consider their data workflows and to justify their storage resource use. Requesters should note that the CHAP Allocation Review Criteria also apply to requests for allocated storage resources – that is, GLADE project spaces and Campaign Storage archive use. The CHAP reviewers do not expect lengthy justifications; in most cases, a paragraph that demonstrates forethought commensurate with the scale of anticipated need will suffice. For requests over 50 TB, longer justifications become appropriate.

In addressing the CHAP review criteria with respect to storage resources, justify the project's rationale for which data will be stored, the project's work plan for managing data, and the project's need to retain the data to be stored long term. While some statement of storage resource needs is expected in Section D, you may choose to provide details related to longer-term storage needs in Section E (Data Management Plan) to keep sections A-D within the five-page limit. *While use of scratch disk space need not be formally justified in the project's allocation request, demonstrating effective use of scratch space within the overall work plan can help reviewers understand your needs for project space or Campaign Storage space.*

1. HPC. The table should give the core-hours per simulated year or appropriate time period and the total core-hours needed for each experimental configuration as well as the total core-hours for the request. Describe how you arrived at the number of core-hours for each proposed computational experiment. If not provided elsewhere, details on how HPC resource requirements are estimated **MUST** be included to help reviewers evaluate whether the resources sought are justified and will be used efficiently.

Estimating Derecho resource needs. Now that Derecho is available to all users, we strongly encourage researchers to request a Small allocation and use the project to capture actual benchmark performance data to support calculating the cost of your large-scale runs. For a limited time, we will also allow users to estimate from known Cheyenne costs: Derecho users can expect to see a 1.3x improvement over the Cheyenne system's performance on a core-for-core basis. Therefore, to estimate how many CPU core-hours will be needed for a project on Derecho, you can multiply the total for a Cheyenne project by 0.77. The core-hours used for a job are calculated by multiplying the number of processor cores used by the wall-clock duration in hours. Derecho core-hour calculations should assume that all jobs will run in the regular queue and that they are charged for use of all 128 cores on each node.

Estimating Derecho GPU resource needs. When requesting an allocation for Derecho GPU nodes, please make your request in terms of GPU-hours (number of GPU nodes x 4 GPUs per node x wallclock hours). Derecho GPU-hour estimates can be based on any reasonable GPU performance estimate from another system, including Casper. For projects that use fewer than four GPUs simultaneously, please request an allocation on Casper GPU.

2. Campaign Storage. In the table showing core-hours, include a column for the final, post-processed, post-analysis amount of data intended for Campaign Storage for each experimental configuration and include the total terabytes planned. Any plans to store raw, unprocessed or temporary data should be justified carefully. Projects with more than 50 TB planned for Campaign Storage are expected to provide more detailed justifications.

Justifying archive space requests. A successful justification for archive space would describe, for example, that the data have a meaningful purpose beyond the post-processing phase; that variables or time steps not needed for planned analyses will be filtered out; and that HPC and analysis stages are interleaved where practical to eliminate the need for short-term use of archive space. *A simple summation of all bytes generated by all proposed HPC runs may set an upper limit on archive space needs but will not automatically constitute a sufficient justification in and of itself.* In most cases, it is not necessary to archive all output; only the subset of data that is critical to the science. Where appropriate, the justification should also describe how the project will reduce archival holdings in subsequent years (for example, a project may have a larger need during peak activity that will decrease in out-years). As with computational justifications, the detail for storage justifications should grow commensurately with the project's anticipated need.

3. Data analysis and visualization (DAV). Describe any planned need for CISL's DAV clusters to analyze or visualize your results. *For standard interactive access to these clusters, we will provide up to 10,000 core-hours with no special justification.* Projects with more extensive plans for use of the clusters should justify their needs in a manner similar to their HPC requests based on benchmarks and cost estimates from sample runs.

Multi-year plan. While the CHAP makes an effort to award each project its full resource need, very large requests should consider providing a breakdown showing the projected use of core-hours and, if applicable, post-processed data during each year of the allocation. Tie this to the planned computational experiments completed or partially completed each year.

Special requirements. Please specify any resource requirements that fall outside of the default environment limits, such as the need for longer job runtime limits, that may affect your ability to complete the proposed computational experiments.

Additional supporting information

Sections E through G together must be no more than an additional five pages.

E. Data management plan. Consistent with NSF's requirement that all proposals include a data management plan, summarize your plan for managing the data resulting from this computational work throughout and beyond the period of performance for the NSF award. This section can be used to provide additional details or justification for the storage resource needs, to describe plans for sharing the project's data, and to summarize any anticipated long-term storage needs. *A well-justified data management plan is critical because of the potential for large-scale projects to produce extensive data output.*

F. Accomplishment report. The accomplishment report should encompass computations performed using CISL resources by the PI or Project Lead. *Clearly distinguish accomplishments on this CISL project* (that is, for prior CISL use associated with the same NSF award) and accomplishments from all past use of CISL resources. Related work performed on CISL resources by other members of a larger research group may be described, if relevant to this request. Briefly describe the calculations and scientific accomplishments that were completed. Include publications submitted or published that resulted from use of CISL resources. List graduate students who used these computational resources and indicate if these resources supported their thesis or dissertation research. If so, please include the thesis or dissertation title(s).

G. References. Please limit to those directly related to the proposed project and referenced in your Request Summary document.

H. Figures and captions. *Optional.* Figures may be embedded within the main body of the Request Summary; embedded figures count against the five-page limit. Figures and charts at the end of the Request Summary will not count against the five-page limit.