

Guidelines for Resource Allocation on the National E-infrastructure

(Allocation period 2021.1)

Abbreviations:

Administration: UNINETT Sigma2

Committee: Resource Allocation Committee (Ressursfordelingskomité, RFK)

Partners/Metacenter: UNINETT Sigma2, NTNU, UiB, UiO, UiT

Definitions:

Resources:

The national e-infrastructure includes a variety of compute and storage facilities. Advanced user support is also considered a resource.

Allocation:

An allocation on a compute facility is in the form of computing hours (processor core hours). An allocation on a storage facility is in the form of Terabytes. An allocation for advanced user support takes the form of person effort (hours).

Allocation Period:

The available compute and storage resources in the national e-infrastructure are allocated per period. Unless stated otherwise, an allocation period for project data and HPC resources is a 6-month period, starting 1 April or 1 October.

Metacenter:

A collaboration of five organisations to pool competencies, resources and services from UiB, UiT, UiO, NTNU and UNINETT Sigma2.

Background

The term e-infrastructure in this document denotes the infrastructure provided by the resources financed in whole or in part by UNINETT Sigma2 AS, in particular:

- HPC: Norwegian infrastructure for high-performance computing
- Data Storage: Norwegian infrastructure for scientific data (NIRD)
- HPC and Storage for sensitive data: Services for Sensitive Data (TSD)
- Platform for Data Analytics: NIRD Service Platform is a Kubernetes based data analytics platform allowing pre/post-processing, visualisation, and AI/ML workflows

The e-infrastructure provides a compute- and data infrastructure with related support services for research and education financed by the RCN or ministries. The infrastructure is available for all research and education supported by public funding, including individuals and groups in all scientific disciplines at Norwegian universities, university colleges, Norwegian hospitals and research organizations. In addition, the infrastructure may be available to commercial partners who contribute to the funding of the e-infrastructure.

UNINETT Sigma2 AS has the strategic and operational responsibility of the e-infrastructure.

UNINETT Sigma2 and the university partners in the Metacentre operate and provide end-user support for compute and storage facilities. Service Agreement define the rights and obligations between UNINETT Sigma2 and the university partners.

The Board of UNINETT Sigma2 appoints the Resource Allocation Committee¹. The Committee evaluates proposals based on scientific merit and the evaluation of science of the research activity and documented needs. The combined resources are allocated by the Committee to individuals, groups and projects.

The Committee reports each year to the board of UNINETT Sigma2. The report shall include statistics on allocations, usage and approval rate for proposals, as well as an assessment of resource access. The report shall also include a high-level evaluation of the quality of the proposals, project results and future needs.

The resources shall be open to all scientific disciplines and access to the resources is granted based on submitted applications. Large proposals might be subjected to application performance evaluation.

Organization

The Committee meets before the start of each allocation period to decide on the allocation of available resources to projects for the upcoming period.

A minimum of five Committee Representatives are required to make the meeting quorate.

The Administration takes care of the secretarial tasks for the Committee.

- The Administration maintains contacts with projects that have (or would like to have) access to the resources in the e-infrastructure.
- The Administration prepares the Committee meetings to enable the Committee to evaluate the proposals, and to allocate the available resources.

Working Group

Prior to the Committee-meeting, a working group, whose members have good knowledge of the characteristics and status of the resources and of the (existing) projects that use the e-infrastructure, meets. The aim of the working group is to support the Committee in the technical evaluation of the proposals. For each proposal, the working group gathers information about the usage of the facilities by the projects, in particular concerning efficiency, scalability and relevance of the applications.

The group comments on the combination of the requested resources and the software needs, the feasibility of using the resources, the efficiency of the software and any other technical aspects that could be relevant information for the evaluation. Regarding storage applications in particular, the group also suggests the location of the primary site over the two NIRD disks, on the basis of operational considerations.

¹ [Mandate for Research Allocation Committee for e-infrastructure](#)

Principles for proposal-based resource allocation

A primary task of the Committee is to evaluate the proposals and decide on the allocation of the available resources. This includes the distribution of research projects over the available facilities to best ensure optimal utilization of the resources.

The following principles form the basis for the evaluations and allocations by the Committee:

- **Transparency:** The criteria for assessing proposals and details of the evaluation process are published before applicants submit proposals. The criteria define how the Committee operates and manages the evaluation process.
- **Assessment:** The Committee may use external experts to (i) assess the scientific merit of proposals against the published criteria and to (ii) assess the technical feasibility of proposals.
- **Confidentiality:** The Committee and the Administration treat proposals in confidence. Those who advise the Committee are required to do the same.
- **Prioritization:** Proposals may be prioritized for access, by assessing the merit of each proposal against general or published criteria for assessment and against other proposals.
- **Managing Interests:** All participants in the evaluation and allocation process (including the Committee members) are required to declare interests so that any conflicts are identified and managed.
- **Reporting:** Applications might experience reduced allocations due to missing publication reporting. This is further described in the application form and allocation letters.
- **Fairness to Science:** The evaluations and allocations must be fair to the science proposed, rather than to an individual applicant, group or institution.

Proposals

The Committee issues calls for proposals twice a year:

- A call in January/February, with allocations effective 1 April
- A call in July/August, with allocations effective 1 October

Researchers may apply for allocations on one or more facilities and they can apply individually. Researchers working in the same unit or group (or on the same project) are encouraged to submit a single proposal. For research projects handling non-sensitive personal data, it is not required that researchers submit a separate proposal for each research project: multiple research projects with overlapping activity may submit a single proposal. In cases where projects are handling sensitive data, there must be an application per authorized research project. Special needs for resources or support must be stated in the proposal. If additional applications for the same period are received, the last application will always override the previous applications.

The Administration maintains communication with the applicants about the status of their proposals and may request additional information to ensure completeness and correctness of the proposals. The tasks of the Administration include:

- Preparation and publication of the calls for proposals
- Maintenance of web pages with information for users (and potential users) on the e-infrastructure
- Communication of the results of the evaluation and allocations to the applicants, plus communication of the allocations to the Partners that operate the facilities

Proposals must include sufficient information for the Committee to be able to judge the scientific quality of the proposed activity and judge the type and quantity of resources requested. Essential elements of each proposal are:

- Details regarding the applicant
- Details on how the (research) activity is funded
- The type and amount of resources requested
- Software needs
- Core count and core memory needs
- A brief project description
- Expected measurable output (PhD degrees, master degrees, scientific publications, patents, others).

The applicant must either have a permanent position or be a postdoctoral fellow or temporary employed researcher at his/her home institution. Research proposals are granted resources if they are financed by public funding, i.e., money from governmental sources (e.g., the Research Council of Norway, ministries, the European Union) to individuals, organizations or entities (e.g., universities, university colleges, research institutes).

Proposals for allocations can span multiple allocation periods. Such proposals require a verification from applicants in advance of each allocation period in order to be processed by the Committee for subsequent periods.

Additional conditions for proposals to TSD – Sensitive Data Services: Research projects that have obtained approval from the data protection authority of reference (Regional Ethical Committee, Norwegian Center for Research Data, Datatilsynet or similar) may apply for TSD resources if:

- the project needs compute resources on a HPC cluster and/or
- the project needs 10 TB or more storage capacity for sensitive data

A condition to use TSD services is to have a TSD-project (indicated by the alphanumeric code *pXXX*). Application for resources might be sent prior or after obtaining access to the TSD-project. Applications will be evaluated by the Committee and granted on a per period basis.

If the application is granted, the allocated resources will be assigned to the *pXXX* project indicated in the application form. If a *pXXX* project does not exist at the time the resources are allocated, it is responsibility of the applicant to contact the TSD operation, initiate a new TSD-project and communicate to the Administration the *pXXX* project assigned to the granted allocation. Checking of permits and formal approval of projects is the responsibility of TSD operations at UiO/USIT and not the responsibility of the Committee. Allocated quotas apply to the actual period regardless when the TSD-project has been activated.

Proposals to the NIRD Service Platform: Researchers that have a data storage project might apply for access to the resources on the NIRD Service Platform. Applications can be submitted at any time during the year by sending a mail to sigma2@uninett.no. Further guidance will then be provided.

[Proposals submitted outside the regular calls](#)

If a project exhausts its allocations before the end of the on-going allocation period, the project manager can submit a request for an extra allocation for the remainder of the period. There are no deadlines for such requests. Requests for extra allocations are subject to available capacity on the facilities. This also applies to proposals submitted outside regular calls. HPC applications submitted

after the deadline and extra applications granted in part or full will receive a lower scheduling priority than ordinary applications. This does not apply to applications for new projects.

Proposals for small scale exploratory work

To explore feasibility of the e-infrastructure, it is possible to apply for a small-scale allocation through a free-text email to sigma2@uninett.no describing purpose, anticipated resource needs, software needs and duration. The maximum small-scale allocations for compute and storage resources are 20 000 CPU-hours and 5 TB, respectively. Small scale proposals might be submitted outside regular calls.

Contribution Model

From 2018, Sigma2 has introduced a financial contribution model. Costs for storage and HPC resources will be applied to projects funded by RCN after July 2018 with either of the following needs and funding characteristics:

- a. Projects requiring more than 10 TB storage resources
- b. Projects requiring HPC-resources, RCN-funded with more than 15 MNOK which have budgeted for e-infrastructure
- c. Projects requiring TSD-resources, any type of funding.

Projects with granted resources from the Committee will be charged for only the utilized resources. The Administration has the responsibility to invoice projects for the utilised resources. Invoicing will be done once per year for the past two periods. Invoices will be sent to the project responsible unless differently specified. By starting using the resources it is assumed that the cost model has been read and approved by the project responsible

Evaluation and allocation

The Committee evaluates proposals for access to the e-infrastructure. The evaluation of proposals can include:

- The scientific merit of the proposed activity
- The need for access to national large-scale resources
- The feasibility of using the (requested) resources for the proposed activity
- Previous usage

When deemed necessary, the Committee may require and use additional information in the evaluation process, e.g., from experts in the scientific and technical area of a proposal, or experts in techniques for efficient resource utilization. Proposals for large allocations may be sent to external experts. Applicants should not reply to the experts' written assessments prior to the evaluation of the proposals.

The Committee assigns allocations based on the evaluation of the proposals:

- For continuation of a project that spans multiple allocation periods, resource usage in preceding allocation periods is considered.
- For the continuation of a project that spans multiple allocation periods, recent research output is considered.

After each allocation meeting, the Committee produces a list of all proposals, including allocations applied for, approved allocations and a brief motivation for the decisions.

Allocations are granted per allocation period, i.e., a project may apply for, but is not granted allocations for multiple periods. This limitation is needed to be able to take into account (i) changes in the total amount of allocations requested (by all proposals) for each new allocation period and (ii) regular changes in the available resources/facilities. Limiting the allocations per period, allows that projects are treated equally (or projects are prioritized) in the allocation decisions for each new period, independent of the time when the proposals for these projects were submitted.

The Committee has the right to modify the requested allocations by:

- Modifying the total allocation for the project in each allocation period
- Granting allocations to other resources than those applied for

Considerations in modifying the requested allocations in a proposal are:

- The quality of the proposal and outcome of the evaluation
- The total amount of allocations requested (for all proposals) versus the total amount of resources available
- The history of the applicant/project, e.g., in terms of requested allocations and actual usage in previous allocation periods, and scientific output
- How well justified is the request for access to specific platforms compared to more generic capacity resources.

The Administration can directly handle proposals for allocations on compute or storage facilities in case there are sufficient (unused) resources available on the requested facilities.

Proposals for small projects (e.g., for porting or testing purposes) will not be assessed by the Committee and are processed directly by the Administration.

Conditions on allocations

The following allocation principles are in effect for allocation of storage and compute resources:

- The Committee may at any time, revise the allocations for a project during an allocation period. In such case, the project responsible shall be informed and the reason stated.
- The allocations may only be used for the purposes described in the proposal. Abuse may result in a revision of allocations.
- The Committee and the Administration cannot be held responsible for any damage that results from usage of allocations granted by the Committee.
- Projects granted access on the compute and storage facilities are required to provide a list of all publications resulting from the use of these resources. Such reporting must be performed through the Norwegian research publication registry Cristin (www.cristin.no)
- Applications for storage must be complemented with a Data Management Plan, either created with easyDMP ([link](#)) or with any other approved tool.

Registering publications is compulsory for all projects granted access to the national e-infrastructure. Scientists having access to both storage and compute resources must register publications connected to both projects. It is the duty of each project responsible to ensure that relevant publications are reported to the Cristin system and tagged with the correct funding source. Projects are required to acknowledge the use of the national e-infrastructure in their publications.

Specific guidelines for allocation of compute resources

- a. Projects that have not registered publications in Cristin in the last three years before the period applied for, and which have received and used allocations before this three-year period, will not be granted resources.
- b. Unused compute allocations at the end of an allocation period cannot be transferred to a future allocation period.
- c. The Committee might scale down applied computing time due to resource limitations, either generally or for specific project applications.
- d. For projects applying for time after the deadline, the general rule is that they will be awarded non-priority CPU allocation only. This does not apply to new projects. The Committee may deviate from this general rule in specific cases.
- e. The Committee tacitly assumes resources being used evenly (approximate linear usage) throughout the allocation period. For projects using significantly less than this projected usage (underusing), part of the allocation might be transferred from priority to non-priority time.
- f. Allocations below 1% of a system's available RfK capacity might be decided by the Administration.

Specific guidelines for allocation of storage

- a. Allocation of resources below 25TB is delegated to the Administration (do not require the authorisation of the Committee).
- b. New projects / applicants are normally not granted with more than 10TB in the first allocation period. Projects that need more storage can apply for extra allocation during the allocation period (see c.)
- c. Extra-allocations between calls for applications can be processed and authorised by the Administration for requests up to 25TB. Requests above this limit should be forwarded to the Committee if urgency is justified.
- d. Prerequisite for accessing the allocated resources is to have a valid data management plan (DMP). (mail to indicating the project number (NSXXXXK) in the mail text or title. Received DMP are stored in Sigma2's internal database and are an updated version. It is the project owner's responsibility to submit updated versions. The Committee reserves the right to withdraw the quota if a valid DMP has not been provided within the first half of the current period. DMPs can be in free format or created through approved DMP tools (es. easyDMP, DMP Online, etc.)
- e. Projects that have not registered publications in Cristin in the last three years before the period applied for, and which have received and used allocations before this three-year period, might undergo a reduction of the allocated resources. It is the project owner's responsibility to record the publications.
- f. Projects that have not started using their allocation by the end of the second period will be discontinued.
- g. Under-usage (<50%) of allocations over two consecutive periods normally results in adjustment of the allocation quota to minimum 75% of the current allocation. Further evaluation of potential reduction should be considered prior to each allocation.

- h. Provided the usage from the previous allocation period is within the guidelines for allocation (no under-usage) the resources awarded for one allocation period are guaranteed for subsequent periods.
- i. Projects with no allocation for two consecutive periods will be decommissioned.

Specific guidelines for allocation of sensitive data services (TSD)

- a. Allocation of 10TB storage and/or up to 10 000 CPU-hours is delegated to the administration (do not require the authorisation of the Committee).
- b. New projects / applicants are normally not granted with more than 50TB in the first allocation period.
- c. Under-usage (<50%) of allocations over two consecutive periods normally results in adjustment of the allocation quota.
- d. Extra-allocations between calls for applications can be processed and authorised by the Administration for requests up to 10TB. Requests above this limit should be forwarded to the Committee if urgency is justified.
- e. Pre-condition for accessing the allocated resources is to have a valid data protection approval from the data protection authority of reference. Although projects can submit application and eventually be granted while they wait for the official notification of the issued approval, the access to the resources will happen only after such approval has been issued. The Administration reserves the right to withdraw the quota if the data protection approval has not been issued within the first half of the current period.
- f. Provided the usage from the previous allocation period is within the guidelines for allocation (no under-usage) the awarded resources for one allocation period are guaranteed for subsequent periods.
- g. Projects with no allocation for two consecutive periods will be decommissioned.