

## Guidelines for Resource Allocation on the National E-infrastructure

(Allocation period 2018.2)

### Abbreviations:

Administration:	UNINETT Sigma2
Committee:	Resource Allocation Committee (Ressursfordelingskomité)
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 (or application 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.

## Background

The term e-infrastructure in this document denotes the infrastructure provided by the projects 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)
TSD:	Services for Sensitive Data

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 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 a set of large-scale compute and storage facilities. Service Agreements define the rights and obligations between UNINETT Sigma2 and the university partners.

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. The Committee evaluates proposals based on scientific merit and documented needs. Large proposals are subject to application performance evaluation.

## Organization

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.

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

## Working Group

Prior to the Committee-meeting, a working group meets whose members have good knowledge of the characteristics and status of the resources and of the (existing) projects that use the e-infrastructure. The aim of the working group is to support the Committee in the technical evaluation of the proposals. For each proposal, the group should comment 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.

The working group gathers information about the usage of the facilities by the projects, in particular concerning efficiency, scalability and relevance of the applications.

## 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 such that the (software) applications are assigned to the facilities best suited and most cost-efficient for these applications.

The following principals 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 prevail the previous applications.

The Administration maintains the communication with the applicants about the status of their proposals, and may request additional information to assure 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
- A brief project description
- Expected measurable output (PhDs, scientific publications).

The applicant must either have a permanent position or be a postdoctoral fellow 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., RCN, ministries, EU) to individuals, organizations or entities (e.g., universities, university colleges, research institutes).

Proposals for allocations can span multiple allocation periods. Projects that are granted access for multiple allocation periods must provide regular updates to the original (or last submitted) request for allocations.

#### 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.

#### 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](mailto:sigma2@uninett.no) describing purpose, anticipated resource needs, software needs and duration. The maximum small scale allocations for compute and storage resources are 10 000 CPU-hours and 5 TB, respectively. Small scale proposals might be submitted outside regular calls.

#### Proposals for Services for Sensitive Data (TSD)

Services for Sensitive Data (TSD) is a platform to store, compute and analyse on sensitive research data in compliance with Norwegian regulations regarding individuals' privacy. TSD is developed and is operated by UiO in partnership with UNINETT Sigma2, and is a part of the national infrastructure.

Research projects that have obtained an approval from the data protection authority of reference (REK, NSD, 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 Precondition 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 applies to the actual period regardless when the TSD-project has been activated.

Projects with granted resources from the Committee will be charged for only the utilized resources within the limit of the allocated CPU-hours and storage, according to the price list published [here](#). No administrative costs will be applied for the project.

## Evaluation and allocation

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

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

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

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 number of periods for which the project is granted allocations
- 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 as opposed 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) need 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](http://www.cristin.no)).

Registering publications is compulsory for all projects granted access to the national e-infrastructure. 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. Unused compute allocations at the end of an allocation period cannot be transferred to a future allocation period.
- b. The Committee might scale down applied computing time due to resource limitations, either generally or for specific project applications. In this case, part of the grant(s) might be made non-priority.
- c. 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. For applications for resources in the current allocation period, the administration might reduce the allocation based on available unallocated capacity in the period.
- d. 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), the Committee might transfer part of the allocation from priority to non-priority time.

### 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 50TB in the first allocation period.

- c. Projects that have not started using their allocation by the end of the second period will be discontinued.
- d. Under-usage (<50%) of allocations over two consecutive periods normally results in adjustment of the allocation quota to 75% of the current allocation. Further evaluation of potential reduction should be considered prior to each allocation.
- e. 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.
- f. All active storage allocations should be reviewed yearly (at a minimum) based on submitted reports, under (and over-) usage of allocated resources and the groups agility in making effective use of the granted allocations. 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 requested in the original application.

#### Specific guidelines for allocation of sensitive data services (TSD)

- a. Allocation of 10TiB storage and/or up to 10000 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. Further evaluation of potential reduction should be considered prior to each allocation.
- d. Extra-allocations between calls for applications can be processed and authorised by the Administration for requests up to 10TiB. 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.