FENIX

RESEARCH INFRASTRUCTURE

D5.1

Report on resource allocation mechanism specification

Work package:	WP5 - Resource allocation policies design and implementation	
Author(s):	Debora Testi	CINECA
Reviewer #1	Thomas Eickermann	JUELICH
Reviewer #2	Colin McMurtrie	CSCS
Dissemination Level	Public	
Nature	Report	

Date	Author	Comments	Version	Status
24.01.2019	D. Testi, C. Padrin	Initial version circulated	V0.1	Draft
21.02.2019	D. Testi	New version including answers to comments on the previous one	V1.0	Draft
19.03.2019	D. Testi	Merged inputs on some sections	V1.1	Draft
20.03.2019	D. Testi	Final revision and draft for internal review ready	V1.2	Draft
11.04.2019	Anne Carstensen	Update of Section 4.1	V1.3	Draft
12.04.2019	D. Testi	Consolidated document	V2.0	Final
27.11.2019	D. Testi	Update of Section 4.1.2 according to feedback from 3rd Intermediate ICEI Review (RV3), added diagram in Section 2	V2.1	Draft
28.11.2019	Anne Nahm	Editorial updates	V2.2	Final

The ICEI project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 800858.

Executive Summary

The ICEI partners are extending their respective infrastructures with ad hoc hardware components and they are federating services in order to allow research communities to execute their data processing workflows.

The consortium is going to provide these resources for free to users from the neuroscience community (via HBP, PRACE, and national grants) and to other researchers (via PRACE, and national grants) based on a peer-review process along the whole duration of the project.

Even if the hardware procured at the different sites is different, it will be presented in a coherent and transparent way to the users requesting access. For this reason, ICEI decided to rely on the definition and allocation of Fenix Credits for the provisioning of resources which will be the standard unit of measurement to let the users know how many resources will be made available in a certain period of time and monitor how these will be consumed.

This document describes how the Fenix Credits are defined and how the resources will be distributed to the users from the different communities via the Human Brain Project (HBP) and PRACE.

Contents

Acronyms

AAI	Authentication and Authorization Infrastructure
ACD	Active Data Repositories
ACL	Access Control List
API	Application Programming Interface
ARD	Archival Data Repositories
BSC	Barcelona Supercomputing Center
CapEx	Capital Expenditure
CDP	Co-design Project
CEA	Commissariat à l'énergie atomique et aux énergies alternatives
CINECA	Consorzio Interuniversitario
CLI	Command Line Interface
CSCS	Centro Svizzero di Calcolo Scientifico
DL	Data Location Service
DM	Data Mover Service
DT	Data Transfer Service
FPA	Framework Partnership Agreement
FURMS	Fenix User and Resource Management Services
GoP	Group of Procurers
GUI	Graphical User Interface
HBP	Human Brain Project
HPAC	High Performance Analytics and Computing
HPC	High Performance Computing
HPDA	High Performance Data Analytics
HPST	High-Performance Storage Tier
IaaS	Infrastructure as a Service
IAC	Interactive Computing Services
ICCP	Interactive Computing Cloud Platform
ICEI	Interactive Computing E-Infrastructure for the Human Brain Project
ICN	Interactive Computing Node
IdP	Identity Provider

IPR	Intellectual Property Rights
JP	Joint Platform
JSC	Jülich Supercomputing Centre
LCST	Large-Capacity Storage Tier
MS	Monitoring Services
NDA	Non-Disclosure Agreement
NETE	External Interconnect
NETI	Internal Interconnect
NMC	Neuromorphic Computing
NVM	Non-Volatile Memory
NVRAM	Non-Volatile Random Access Memory
OIDC	OpenID Connect
OpEx	Operational Expenditure
PaaS	Platform as a Service
PCP	Pre-Commercial Procurement
PI	Principal Investigator
PID	Persistent Identifier
PIE	Public Information Event
PRACE	Partnership for Advanced Computing in Europe
Q&A	Questions and Answers
QoS	Quality of Service
R&D	Research & Development
R&I	Research & Innovation
RBAC	Role-Based Access Control
RFI	Request For Information
SCC	Scalable Computing Services
SGA	Specific Grant Agreement
SIB	Science & Infrastructure Board
SLA	Service Level Agreement
SP	Subproject
TCO	Total Cost of Ownership
TGCC	Très Grand Centre de Calcul

UI	User Interface
US	User Support Services
VM	Virtual Machine Services

1. Introduction

The ICEI project aims to provide users with specific federated hardware solutions and services for users of the neuroscience community and to researchers at large.

This document describes, first of all, how the available ICEI resources will be presented to users (via Fenix Credits), then the deliverable presents how the consortium is going to provide these resources for free via HBP and/or PRACE based on a peer-review process along the whole duration of the project.

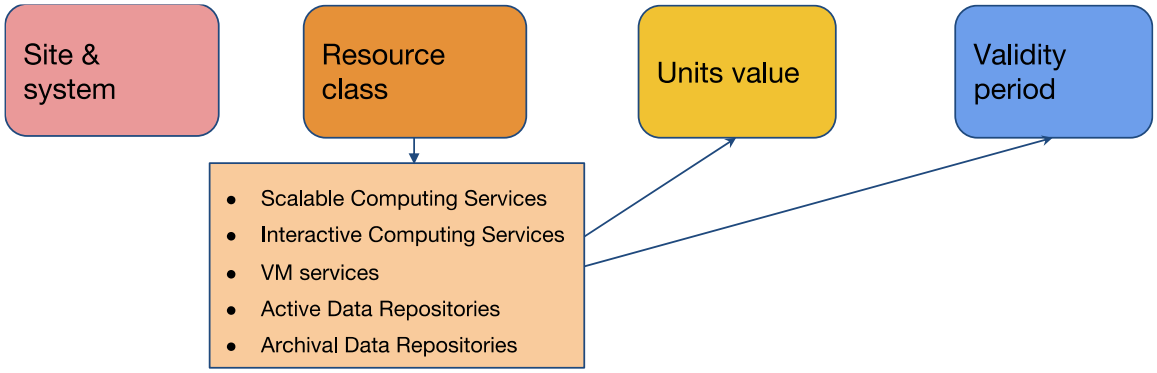
2. Fenix Credits

As previously said, the resources available via the ICEI project will be presented to the external users as credits.

> Fenix Credit is the token presented to external users and assigned/consumed to/by a research project

A Fenix Credit is defined by the following attributes:

- Resource class (which indicates the type of referred resource);
- Value in units (depending in the resource class above);
- Site and system for which the credit is provided (to allow the conversion from Fenix to local infrastructure information and vice versa);
- Validity period (period for which the resources will be available, depending in the resource class above).



Credits can be requested and used at different sites at the same time, i.e. a user can have data stored in a site, process them in a second different site and manage them in a third different site.

2.1 Credits: Resource Classes

As mentioned above, the credit definition is related to the specific type of requested resource (i.e. computing, storage, etc); we can identify the following five Resource classes associated with the different type of services that ICEI is providing (as already reported in detail in D3.1):

- Scalable computing services: massively parallel HPC systems for scalable applications.
- Interactive computing services, which provide quick access to single servers that provide access to data repositories and potentially allow to scale-out using scalable computing services.
- VM services: Service for deploying virtual machines in a stable and controlled environment that is, for example, suitable for deploying domain-specific platform services like the Collaboratory, HBP information catalogues, image services etc.
- Data Repositories: data storage can have different characteristics and has been classified as described in Annex 1:
    - Standard Archival Data Repository (ARD long) that continues to be available up to 6 months after the expiring date of the project.
    - Archival Data Repository (ARD 10y) that continues to be available up to 10 years after the expiring date of the project. The storage resources are provided by selected Fenix sites partially based on in-kind contributions.
    - Active Data Repository (ACD short) based on storage technologies optimised for high performance (high bandwidth and/or high IOPS rate).
    - Active Data Repository (ACD long) based on storage technologies optimised for capacity.

2.2 Credits: Value in Units

For each resource class, we identified: the unit of measurements that will be associated with the specific Resource Class so that all sites will present credits in the same way.

- Scalable computing services -> unit is the node*hour (i.e. the number of nodes times the period of time, expressed in hours, that the nodes are in use) budget for the requested project.
- Interactive computing services -> unit is the node*hour budget for the requested project.
- VM services -> different models have been defined to consider combinations of: Number of virtual CPUs (vCPU), Amount of memory (MEM), Number of virtual GPUs (vGPUs), Amount of capacity-optimised storage (HDD), Amount of performance-optimised storage (SSD), Number of virtual GPUs (vGPUs); details are provided in Annex 2.
- Archival data repository -> unit is referred to the storage occupation and is defined as TByte.
- Active data repository -> unit is referred to the storage occupation and is defined as TByte*days (i.e. the number of TBytes used times the period of time, expressed in days, that the data is present in the storage) for short ACD and in TByte for the long ACD. Notes: this is the storage provided on top of the standard one each centre provides when allocating a project on the site infrastructure.

2.3 Credits: Validity Period

This attribute of the credits indicates for how long the credit will be available for use; the period will depend on the resource class:

- Scalable computing services: When provisioning the credits, the validity period will be, in general, one year (i.e. 12 months). When consuming the credits, these will be allocated from the start date of the granted project for its whole duration. Site-local policies may further divide the 12-month period into quarters and allocate a budget per quarter, for consumption in said quarter.
- Interactive computing services: When provisioning the credits, the validity period will be, in general, one year (i.e. 12 months). When consuming the credits, these will be allocated from the granted project start date for its whole duration and we consider the resources available at the same time they are allocated. Site-local policies may further divide the 12-month period into quarters and allocate a budget per quarter, for consumption in said quarter.
- VM services: When provisioning the credits, the validity period will be, in general, one year (i.e. 12 months). When consuming the credits, these will be allocated from the granted project start date for its whole duration and we consider the resources available at the same time they are allocated.
- Active and archive data repositories: validity period for active storage will be established by each centre when creating the credits for the active storage. For the archival storage, validity period is from the credits' definition date to the end of ICEI resource provisioning. When consuming the credits:
    - in the active storage, data will be kept up to 2 months after the expiring date of the project;
    - in the archive storage up to 6 months for the long type and up to 10 years for the 10y type, after the expiring date of the project;
    - the backed-up data will be available up to the standard time offered by the single sites;
    - both storage validity can be extended by consumption of further credits.

2.4 Credits: Management

All credits and project information will be recorded in FURMS (see D4.15 technical specifications); every community will have a profile where to allocate credits with indications on the type of credits, where these credits can be consumed (i.e. the site name where credits are converted into a core-hour budget and memory occupancy), and for how long (validity period).

Every site can supervise the consumed credits allocated to a project on the specific infrastructure. This information has to be extracted periodically (suggestion is at least once per day and monitored). These values have to be recorded in the project profile, defined in FURMS.

A counter for available credits could be added in a project profile recorded in FURMS, in order to let Fenix staff and the principal investigator (PI) monitor the status of the project on a daily basis.

3. Provisioning of Fenix Credits

Fenix Credits are provided by ICEI Resource Providers who will periodically determine the available amount of credits over a certain period of time. The credits will be then provided to the Fenix communities based on the pre-defined share of the Fenix infrastructure (namely 25% will be provided to HBP, 15% to PRACE and the rest via national grants)[1]. From there, the Fenix Communities will be in charge of allocating their Fenix Credits to users via a transparent and peer-review Call for Projects (CFP). During each project's execution Fenix users will consume Fenix Credits. The ICEI Resource Providers will monitor the subsequent consumption of these credits.

More in detail the process will be the one described below:

- Conversion from resources to credits:
  a) Every site calculates the amount of credits that will be provided to Fenix during the period every quarter of year, distinguishing between the types defined above.
  b) ICEI PMO will allocate the credits available every quarter for the types of communities (HBP, PRACE and national ones).

Specific details on how the number of Fenix Credits provided for a given year of ICEI infrastructure operation have been computed for each of the sites will be documented in the deliverables D4.10, D4.11, D4.12, and D4.13.

Each community will follow the allocation mechanisms, based on a peer-review process, as reported in Section 4 and allocate the Fenix Credits to the approved projects. The Community Administrator will allocate the credits to the awarded projects.

When resources are consumed the amount of available credits will be updated accordingly in FURMS.

- Conversion from credits to resources:
  a) Every site receives the information from the FURMS about the number of Fenix Credits (per type) assigned to a project that will have to be defined in the site infrastructure.
  b) Every site converts site credits in resources (per type) following the definition given in the previous sections.
  c) Every site defines the project on the infrastructure and provides dedicated support and services when needed.

4. Allocation of Fenix Credits

A key aspect of the ICEI infrastructure is the separation of Resource Providers and User Communities. Resource Providers are those entities that provision ICEI infrastructure services and a User Community is a research community that consumes these resources. In this context HBP and PRACE are considered as User Communities.

[1] The amount of resources allocated to the different communities may temporarily change as defined in Grant Agreement.

HBP is given a programmatic access to 25% of the resources provided through this infrastructure, while another 15% will be provided to European researchers at large through PRACE.

Under the condition that a peer-review process is implemented, the final decision on which proposal is provided with what amount of resources remains within the User Community. The review process will have to comply to the following peer review principles that have been established by PRACE[2]:

- **Transparency**: The peer review process is transparent and clear to all relevant stakeholders including the involved funding agencies.
- **Fairness**: Proposals are evaluated on merit and potential high impact on European and international science and economy.
- **No parallel assessment**: The peer review process builds on the experiences and best practices of national and international institutions and constitutes a centralised peer review exercise recognised by the relevant scientific communities.
- **Reviews are done by experts in the scientific field of the proposal**, with no declared conflict of interest, based on criteria published in the Calls for Proposals and with a periodic reshuffling of reviewer's appointments.
- **Confidentiality**: Proposals will be treated with the needed confidentiality by all persons involved in the process. The identities of the peer reviewers shall not be disclosed.
- **Right to reply** to technical and scientific evaluations.

4.1 HBP Procedures and Policies

Currently the HBP core project is the EBRAINS User Community. In this section we explain the details of the mechanisms for allocating ICEI resources to the EBRAINS User Community. We start with introducing the actors, outline the decision-making procedure for accepting or rejecting proposals, provide details on the technical and scientific assessment of the proposals and the mechanisms for appointing experts. The below procedure has been discussed between ICEI and HBP and has been formally approved by the HBP Directorate during its meeting on February 26, 2019.

4.1.1 Actors

Applicant(s)	Researchers jointly providing a proposal for use of ICEI resources that is in line with the mission of EBRAINS and the HBP. The researcher may or may not be member of the HBP project.
Principal Investigator	Applicant that leads a proposal and in case of approval the research project execution.
HBP DIR	Directorate of the Human Brain Project.

[2] http://www.prace-ri.eu/peer-review/

HBP SIB	Science and Infrastructure Board of the Human Brain Project.
EBRAINS IAC	Infrastructure Allocation Committee of EBRAINS[3], initially with members appointed by the HBP SIB.
Science Expert	Expert in the relevant scientific field who is capable of performing a scientific evaluation of a proposal as outlined in section "Scientific Assessment".
ICEI Expert	Expert from the concerned Resource Provider who is capable of performing a technical evaluation of a proposal as outlined in section "Technical Assessment".
ICEI PMO	Project Management Office of the ICEI project.
ICEI Resource Provider	Organisation that provides ICEI infrastructure services, i.e. BSC, CEA, CINECA, ETHZ/CSCS or JUELICH/JSC.

4.1.2 Process and Timeline

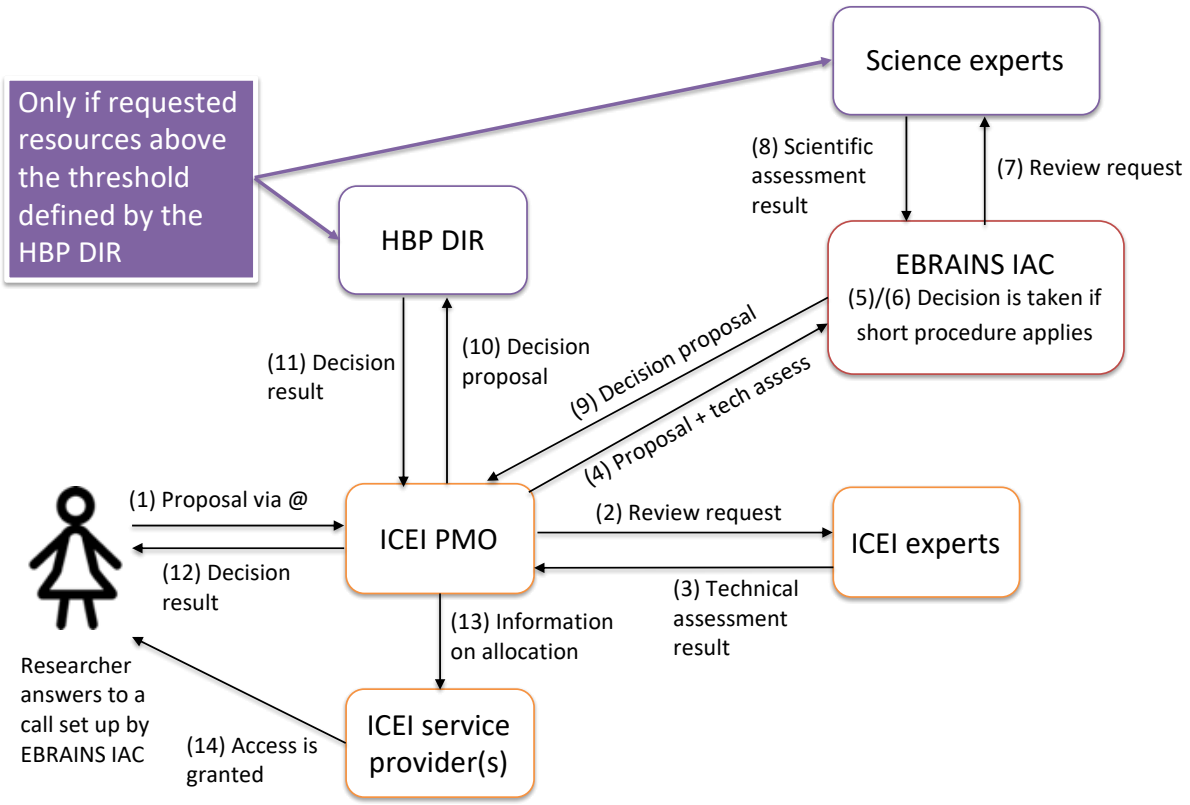
In the following, the process that leads to acceptance or rejection of proposals for use of ICEI resources is described. A graphical view of the process is shown in Figure 1.

1. The Applicants submit a proposal to ICEI PMO following the Application Template in response to a call prepared by the EBRAINS IAC and approved by the HBP DIR.
2. ICEI PMO selects from the pool of ICEI Technical Experts one person to perform a technical assessment of the submitted proposal and forwards the proposal to the expert with request for technical assessment within 5 working days after receiving the proposal.
3. ICEI Technical Expert returns the technical assessment report based on Technical Assessment Template within 5 working days after receiving the request.
4. ICEI PMO forwards the submitted proposal together with the technical assessment to EBRAINS IAC.
5. The EBRAINS IAC decides whether a shortened procedure for particularly small projects applies. A shortened procedure may only be considered if both of the following criteria are met:
   a) The amount of requested resources is below a threshold defined by the HBP DIR (currently 5% of the total offered resources); and
   b) The number of submitted proposals or granted and still active proposals, which have been submitted by the same team, is below a threshold defined by the HBP DIR (currently 4 proposals per team).
6. In case of the shortened procedure the EBRAINS IAC decides whether the proposal is accepted or rejected. The next step in the process is step #12. If the full procedure applies the process continues with the next step.

[3] An initial version of the terms of references for the IAC has been approved by the HBP DIR on its meeting on April 10, 2019.

7. EBRAINS IAC selects from the pool of Science Experts at least two persons to perform a scientific assessment of the submitted proposal and forwards the proposal to these experts within 5 working days after receiving the proposal. At least one Science Expert should come from outside of HBP.
8. Science Experts return the scientific assessment based on the Scientific Assessment Template within 15 working days.
9. EBRAINS IAC formulates on the basis of the scientific and technical assessment a decision proposal and sends it to ICEI PMO.
10. ICEI PMO forwards the decision proposal to HBP DIR for approval.
11. HBP DIR or, in case of the shortened procedure, EBRAINS IAC returns decision on approval or rejection to ICEI PMO.
12. ICEI PMO informs applicants about review outcome as well as the decision that was taken either by EBRAINS IAC or HBP DIR.
13. In case of a proposal being approved, ICEI PMO informs the relevant ICEI Resource Providers to provide access to the requested resources.[4]
14. ICEI Resource Providers grant access to the allocated resources and the support needed to execute the granted project.

Figure 1: Full ICEI Allocation Mechanism for HBP.



[4] In future this will become automatized through the FURMS service.

4.1.3 Technical Assessment[5]

All proposals will undergo a technical assessment. The technical assessment can result in three outcomes:

1. Accepted: The application fulfils all technical requirements use the selected ICEI services.
2. Conditionally accepted: The application does not meet the technical requirements to use the selected ICEI services, but the technical reviewers can identify the measures and time frame necessary to meet them.
3. Rejected: The application does not meet the minimum of technical requirements to use the selected ICEI services.

A proposal must convincingly address the following criteria:

- Are ICEI resources needed to reach the goals of the proposal?
- Are the requested ICEI infrastructure services and possibly necessary HBP platform tools and services available?
- Is other software, which the applicants intend to use, available on the systems which are planned to be used?
- Are the requested resources expected to be suitable for reaching the goals of the proposal?

These criteria should be fully addressed in the application. Reviewers will assess proposals against these criteria.

As a result of the technical assessment applicants may be redirected to a more appropriate approach for using ICEI resources by the ICEI Technical Experts.

4.1.4 Scientific Assessment

Scientific review is performed by internationally recognised experts in the field of research of the proposal. During the scientific assessment an increase or decrease in the requested resources can be recommended for consideration during resource allocation.

The proposals must address the following scientific criteria:

- **Scientific excellence**: The proposed research must demonstrate scientific excellence and a potential for high European and international impact.
- **Novelty and transformative qualities**: Proposals should be novel, develop an important scientific topic of major relevance to European research, describe possible transformative aspects, and expected advances.
- **Relevance to the call**: Proposal needs to address how the proposed project is addressing the scope of the call, if a specific scope is stated in the call.
- **Methodology**: The methodology used should be described and appropriate to achieve the goals of the project.

[5] Largely identical to what PRACE foresees for its peer review process.

- **Dissemination**: The planned channels and resources for dissemination and knowledge exchange should be described.
- **Management**: There must be a solid management structure which will ensure that the project will be successfully completed.

These criteria should be fully addressed in the application. Reviewers will assess proposals against these criteria.

4.1.5 Experts and Expert Pool

All Actors handling peer review are responsible for obtaining high quality non-conflicted technical and scientific reviews for each proposal. All Actors must respect the confidentiality rules and the HBP's current conflict of interest policy.[6]

Experts are selected from a pool of experts. This pool of Science and ICEI Experts is maintained by ICEI PMO and approved by SIB. Science Experts are proposed by EBRAINS IAC and involve both experts within as well as outside of HBP. ICEI Experts are proposed by the ICEI Resource Providers.

4.1.6 EBRAINS Infrastructure Allocation Committee

The EBRAINS IAC has the following tasks:

1. Prepare calls that are to be approved by the HBP DIR,
2. Decide on whether a scientific review is necessary for a given proposal based on criteria defined by HBP DIR,
3. Manage scientific review including assignment of Science Experts,
4. Prepare decision proposals based on the technical and scientific reviews for allocating the available resources,
5. Propose members for the pool of Science Experts,
6. Regularly report to HBP DIR about the proposals that have been accepted or rejected by the EBRAINS IAC as well as report to HBP DIR about issues with the resource allocation process or make suggestions for improving the process,
7. Regularly report to HBP SIB about the resource allocation process.

Its members are appointed by the HBP SIB and may include persons that are not part of HBP.

4.1.7 Current Procedure

While the above procedures were under discussion and being agreed, the following temporary solution was put in place to provide the resources already available in 2018:

- Applicants send a 2-3 pages proposal to icei-coord@fz-juelich.de;
- ICEI experts perform a technical review;
- HBP DIR makes decision to either approve, conditionally approve or reject the proposal;

[6] The version as of 25.09.2017 is available at https://sos-ch-dk-2.exo.io/public-website-production/filer_public/5b/41/5b410fd5-af19-4aa3-98e3-689499b9b992/hbp_sop_conflict_of_interest.pdf

- ICEI PMO informs the ICEI Resource Providers to make the approved resources available.

4.2 PRACE Access Policies

The provisioning of ICEI resources to European researchers has been discussed with the PRACE Board of Directors; different levels of provisioning have been discussed and agreed as described below.

For all types of calls, the calls are expected to be made public through PRACE channels so as to ensure broad dissemination.

4.2.1 Access Type A: Augmentation of PRACE Tier-0 call

While ICEI does not provide Tier-0 compute services, it provides services that would allow to enrich Tier-0 scalable compute services that have been made available to PRACE. Thus, the allocation of ICEI resources additionally to PRACE Tier-0 resources could help executing more complex workflows. It could be used for the following purposes (the list is not meant to be exhaustive):

- Use of ICEI interactive computing services for post-processing of data generated during calculations on Tier-0 systems;
- Speed-up of I/O bound Tier-0 computations through use of ICEI active data repositories that can act as I/O accelerators;
- Storing and sharing of data through ICEI archival data repositories;
- Managing access to Tier-0 resources and results generated using such resources via science domain specific platform deployed through ICEI VM services[7].

In this approach it is proposed to add in PRACE Tier-0 calls the possibility to express interest in the use of ICEI resources. This is being performed as a pilot phase in PRACE call #18 and call #19 which will also help to evaluate the interest and the impact on the current PRACE procedures.

Scientific and technical review

The scientific review of proposals asking for both, PRACE and ICEI resources, should be identical to reviews of proposals asking for PRACE resources, only[8].

Proposals asking for ICEI resources should undergo a technical review through the ICEI project. For this review, ICEI will adopt the same procedures as for the PRACE technical review. Additionally, feedback will be provided to PRACE about which ICEI sites would be capable to provision the necessary services. This information needs to be considered for the final decision on resource allocation as, in some specific cases, ICEI and PRACE resources might need to be provisioned by the same site for technical reasons. During the piloting phase, it will be avoided to change of the project template, and it is suggested to foresee applicants for additional ICEI resources to provide an additional section within the “Project Scope and Plan” document. This might be revised in the future.

Process and administrative aspects

The following process has been agreed:

[7] AiiDA services for material science projects could be a possible use case

[8] http://www.prace-ri.eu/peer-review/

- ICEI PMO informs PRACE peer-review office about the amount of ICEI resources available at sites on time to be published in the call documentation.
- A call text is published informing about the option of requesting ICEI resources in addition to the PRACE Tier-0.
- Proposals are submitted to PRACE without changing the procedure.
- For proposals asking for ICEI resources the relevant section of the "Project Scope and Plan" document is forwarded to ICEI PMO.
- A technical review of the request for ICEI resources is conducted by the ICEI Resource Providers and feedback is provided within the time limits defined by PRACE and based on a to be agreed template. PRACE conducts a scientific review independently of the ICEI project.
- PRACE decides on approval of projects taking the outcome of the technical review into account where necessary (a project can be approved for PRACE resources even if the ICEI resources request cannot be provided for some reasons).
- PRACE informs ICEI PMO about the amount of allocated resources.

Success criteria

With this being a pilot approach, the following success criteria are proposed:

- Availability of additional ICEI services is of interest for scientists submitting proposals for Tier-0 at PRACE (minimum 3 requests received)
- Successful allocation of ICEI resources through PRACE Tier-0 review process is demonstrated

Pilot phase in call #18

The described approach has been tested with success in PRACE Tier-0 call #18, which closed on the 4th September, 2018.

Among the applications, one has requested access to ICEI resources. The technical and scientific reviews have been performed. So far, no obstacles have been reported and we are replicating the pilot in call #19 (which will close on the 30th of April 2019).

4.2.2 Access Type B: Dedicated ICEI calls

ICEI resources can be used to realise smaller scale "Tier-1" projects (DECI-like mechanism).

DECI calls are managed by the PRACE implementation projects and a specific task has been defined in PRACE-6IP to manage the collaboration between PRACE and ICEI at Tier-1 level.

Details on the specific calls will be defined at the start of PRACE-6IP (April 2019), but some elements have been already agreed as reported below.

ICEI and PRACE are also in parallel discussing on the possibility to open dedicated Tier-1 call addressing some specific community needs (i.e. data intensive applications).

Scientific and technical review

The scientific review should follow the peer-review principles established by PRACE[9].

[9] http://www.prace-ri.eu/peer-review/

The technical review will be performed by the ICEI project following a procedure that is similar to the procedure established for PRACE Tier-1 calls.

Process and administrative aspects

The following process is proposed:

- ICEI PMO informs PRACE on available services.
- PRACE and ICEI agree on a call text.
- A call text and submission website are published by PRACE and are disseminated via ICEI, PRACE, and other channels.
- All proposals received until the earlier defined deadline are reviewed on a technical basis by ICEI Resource Providers.
- PRACE peer-review will follow the standard Tier-1 mechanism.
- Technical and scientific review reports are available to PRACE for resource allocation decisions.
- PRACE informs ICEI PMO about the amount of allocated resources.

Success criteria

With this being a pilot approach, the following success criteria are proposed:

- Scientists are attracted to submit proposals for ICEI resources (minimum 10 proposals, i.e. 2 per ICEI site);
- Successful allocation of ICEI resources through PRACE is demonstrated.

4.2.3 Access Type T: Augmentation of existing PRACE Tier-0 projects

ICEI started making resources available per April 1, 2018 at ETH Zuerich/CSCS, which is also providing PRACE Tier-0 compute resources. Part of the ICEI resources could be made available to projects, which have been granted access by PRACE to the Piz Daint system at CSCS based on an earlier Tier-0 call. This has allowed for fast and relatively informal access to already available ICEI resources for a temporary period of time (hence "T").

Scientific and technical review

Ongoing or just completed projects do not need any further review because they have already successfully passed review. Only the alignment of the extension with the original proposal needs to be confirmed. As the projects have already been scientifically reviewed, the scientific excellence of the projects has already been confirmed. Local expertise at the ICEI sites is being used to assess whether the proposed use of additional ICEI resources is in line with the approved project. In case specific science domain knowledge is required for this assessment coming from outside ICEI, a procedure will be discussed with PRACE to have ICEI involve necessary experts while respecting confidentiality commitments of PRACE.

Proposals asking for ICEI resources will have to undergo a technical review through the ICEI project. For this review, ICEI will adopt the same procedures as for the PRACE technical review.

Process and administrative aspects

The following process has been agreed:

- CSCS approaches projects that have been awarded TIer-0 resources at CSCS in calls 15, 16 and 17.
- Proposals for use of ICEI resources are reviewed by ICEI to assess whether they are scientifically in line with the original project and technically sound using the review template for Access Type T (see annex).
- PRACE BoD decides on approval of projects taking the outcome of the review performed by ICEI.
- PRACE informs ICEI about the amount of allocated resources.
- ICEI informs PRACE about the amount of used resources in a to be agreed way.

Success criteria

With this being a pilot approach, the following success criteria are proposed:

- Availability of additional ICEI services is endorsed by running projects (minimum 2 proposals);
- Successful allocation of ICEI resources through PRACE Tier-0 review process is demonstrated.

5. Consumption of Fenix Credits

The consumption of Fenix Credits should be monitored in order to be able to take corrective actions if needed:

- Every site makes available monthly the information on the amount of resources that have been allocated/consumed. This is a temporary solution until FURMS will be ready, then the monitoring would be possible in real time.
- ICEI Fenix administrator will update the information on the residual credits available.
- Quarterly the data are summarised in reports based on a pre-defined template for the ICEI PMO and other relevant bodies.

6. ANNEX 1: Fenix Storage Classes

In the ICEI Description of Work, Fenix did foresee two different storage classes:

- **Archival Data Repositories (ARD)**: Federated data store optimized for capacity, reliability and availability, used for long term storage of large data sets that cannot be easily regenerated. These repositories are accessed through a Swift interface to ensure coherent access control within a federated infrastructure.
- **Active Data Repositories (ACD)**: Site-local data repositories located close to computational and/or visualization resources at the same site, used for storing temporary slave replicas of data sets. In the near future they will typically be realised using parallel file systems, i.e. they are accessed through a POSIX (or POSIX-like) interface.

However, it became later evident, that, further differentiation of storage classes is needed for the following reasons:

- The Fenix architecture specifications mandate accessibility of ARDs through a Swift interface to ensure coherent access control mechanisms. As a consequence, however, storage systems, which do not offer a Swift interface, but otherwise would qualify as an ARD, will become available as an ACD. Because they are not necessarily optimised for extreme performance, the metric for resource consumption foreseen for ACDs are not appropriate in this case.
- From the main target science community, namely the brain research community, the need for being able to store data for a period of 10 years emerged.

To address these changes, we will from now on differentiate different types of ARDs and ACDs by introducing an additional attribute:

Class	Sub-class	Description	Credits Units
ARD	long	Standard Archival Data Repository that continues to be available up to 6 months after the expiring date of the project.	TByte
ARD	10y	Archival Data Repository that continues to be available up to 10 years after the expiring date of the project. The storage resources are provided by selected Fenix sites partially based on in-kind contributions.	TByte
ACD	short	Active Data Repository based on storage technologies optimised for high performance (high bandwidth and/or high IOPS rate).	TByte*day
ACD	long	Active Data Repository based on storage technologies optimised for capacity.	TByte

7. ANNEX 2: Fenix VM Models

7.1 Introduction

This document describes how Fenix VM Models could be defined in a similar manner as AWS. The choice of Fenix VM Models provided by the different ICEI Resource Providers depend on the locally available hardware resources. The way how hardware resources are allocated is assumed to be done in a similar way at the different sites to allow for performance of the Fenix VM Models at the different sites to be reasonably similar. Next the approach to announcing VM resources in user calls and finally the reporting of Fenix infrastructure resources is discussed.

7.2 VM Models

The different models differ (potentially) in terms of the following features:

- Number of virtual CPUs (vCPU)
- Amount of memory (MEM)
- Number of virtual GPUs (vGPUs)
- Amount of capacity-optimised storage (HDD)
- Amount of performance-optimised storage (SSD)

All other differences, e.g. single CPU performance, memory bandwidth etc. are assumed to be negligible or not relevant. It is furthermore assumed that applications can be deployed such that they can run on different processor architectures, i.e. also on non-x86 processors.

Currently, there is no clear need for provisioning different amounts of capacity-optimised storage space (HDD).

Model name	vCPU	MEM [GiByte]	vGPU	SSD [GiByte]	HDD [GiByte]
gpp.s	1	3			100
gpp.m	2	8			
gpp.l	4	16			
gpp.xl	16	64			
gpp-ssd.l	4	16		800	
gpp-ssd.xl	16	64		800	
gpu.m	2	8	1		
gpu.l	4	16	1		
gpu.xl	16	64	1		

7.3 Hardware resource allocation

The mapping of VM models onto hardware should follow similar rules at all Fenix sites to improve similarity in terms of performance. The following rules are proposed:

Allocate one vCPU per hardware thread

MEM refers to physically available memory minus space needed for OS as well as system services (including PFS). No overbooking of MEM.

SSD refers to the effectively available storage. No overbooking of SSD.

Moderate overbooking of GPUs, e.g. 2 vGPU per GPU (to be explored based on the use cases)

Let us consider the consequences of this policy for a specific example. Assume the following hardware resources being available:

Server type	Qty.	Physical resources per server				Virtual resources per server			
		#HW threads	Host memory [GiByte]	SSD [GiByte]	#GPUs	#vCPUs	MEM [GiByte]	SSD [GiByte]	#vGPU
A	8	40	192	0	0	40	176	0	0
B	8	40	192	1600	0	40	176	1600	0
C	4	40	192	0	1	40	176	0	2
Total virtual resources:						800	3520	12800	8

The maximum number of VMs per model would be:

Model name	Server A	Server B	Server C	Total
gpp.s	320	320	160	800
gpp.m	160	160	80	400
gpp.l	80	80	40	200
gpp.xl	16	16	8	40
gpp-ssd.l	0	16	0	16
gpp-ssd.xl	0	16	0	16
gpu.m	0	0	2	8
gpu.l	0	0	2	8
gpu.xl	0	0	2	8

However, different combinations would be possible. Some examples are:

- One Server C could host 2 VMs of model gpu.m and 36 VMs of model gpp.s.

7.4 Resource announcement

The following information is made available:

- The total number of vCPU, MEM, SSD, vGPU.
- The maximum number of VMs per model.

The information to be announced for the example presented in section 7.3 as already been provided in that section.

7.5 Resource reporting and KPI evaluation

The resources are reported as follows:

- The total amount of available resources is defined in terms of total number of vCPU, amount of MEM, amount of SSD and number of vGPU.
- The total amount of allocated resources is obtained by adding the number of vCPU, amount of MEM, amount of SSD and number of vGPUs for each of the allocated VMs. VMs are assumed to be running 24/7, i.e. the VM execution times are not reported.
- The KPI for infrastructure utilisation is evaluated for vCPU, MEM, SSD and vGPU separately.

Let us reuse the example of available hardware resources introduced in section 7.3. Let us assume that 64 gpp.s, 8 gpp.xl, 4 gpp-sdd.xl and 4 gpu.m VMs have been allocated. In total, this translates into the following resource consumption:

	Total resources	**Utilisation [%]**
#vCPU	264	33
MEM [GiByte]	992	28
SSD [GiByte]	3200	25
#vGPU	8	100