PRACE Preparatory Access

Terms of Reference

1   Terms of Reference

PRACE (*Partnership for Advanced Computing in Europe*), the European research infrastructure on High Performance Computing (HPC), makes it possible for researchers from public and private institutions from across Europe and the world to apply for resources on high-end Tier-0 HPC systems via a centralised peer review process.

The objective of **PRACE Preparatory Access** is to allow PRACE users to optimise, scale and test codes on PRACE Tier-0 systems before applying to PRACE calls for *Project Access.* Production runs are not allowed as part of PRACE Preparatory Access.

Currently, PRACE offers the following schemes for Preparatory Access:

- **Type A:** this scheme is intended to produce scalability plots of the performance of the codes on PRACE HPC systems, as well as other parameters that may be relevant to apply for PRACE calls for *Project Access.* The maximum duration of Type A projects is 2 months.

- **Type B:** the objective of this scheme is to undertake code development and optimisation. Applicants will need[1] to describe the proposed work plan in detail, including the human resources and expertise available to implement the project. The maximum duration of Type B projects is 6 months.

- **Type C:** in this scheme, PRACE experts are requested to provide the necessary support to undertake adaptations (development and optimisation) to the codes of PRACE users. The maximum support that can be requested is the equivalent to **6 person-months of effort**. The maximum duration of Type C projects is 6 months.

- **Type D:** this scheme allows PRACE users to start a code adaption and optimisation process on a **PRACE Tier-1 system**. PRACE experts help in the system selection process. In addition to Tier-1 computing time, the PRACE user will also receive Tier-0 computing time towards the end of the project (in the form of Type A scheme) to test the scalability improvements. The work is supported by PRACE experts. The maximum support that can be requested is the equivalent to **6 person-months of effort**. The maximum duration of Type D projects is 12 months.

---

[1] Applicants requesting computing time on Piz Daint, at ETH Zurich/CSCS, will be asked to provide additional information on development milestones.

The Call for Proposals for PRACE Preparatory Access Type A and Type B schemes is a **continuously open** call, with a maximum time-to-resources-access (start-date) of two weeks after the date of submission.

Requests for **Type C** and **D** scheme are evaluated and granted on a **quarterly schedule**, with cut-offs for evaluations to be set at 11:00 AM CEST / CET of the **first working day of March, June, September and December**. Awarded proposals for Type C and D will have a maximum time-to-resources-access (start-date) of two months after the relevant cut-off date.

Before the end of Preparatory Access projects, users are requested to provide a final report of their project, using the corresponding PRACE template form available here.

The following table shows PRACE Tier-0 HPC systems and their current availability for Preparatory Access:

Tier-0 Systems	Architecture	Type A	Type B	Type C
HAWK, hosted by GCS at HLRS (Germany)	HPE Apollo	✓	✓	✓
Joliot-Curie KNL, hosted by GENCI at CEA (France)	BULL Sequana X1000	✓	✓	✓
Joliot-Curie Rome, hosted by GENCI at CEA (France)	BULL Sequana XH2000	✓	✓	✓
Joliot-Curie SKL, hosted by GENCI at CEA (France)	BULL Sequana X1000	✓	✓	✓
JUWELS Booster, hosted by GCS at JSC (Germany)	BULL Sequana XH2000	✓	✓	✓
JUWELS Cluster, hosted by GCS at JSC (Germany)	BULL Sequana X1000	✓	✓	✓
Marconi100, hosted by CINECA (Italy)	IBM Power 9 AC922 Whiterspoon	✓	✓	✓
MareNostrum, hosted by BSC (Spain)	Lenovo System	✓	✓	✓
Piz Daint[2], hosted by ETH Zurich/ CSCS (Switzerland)	Cray XC50 System	✓	✓	Not available
SuperMUC-NG, hosted by GCS at LRZ (Germany)	Lenovo ThinkSystem	✓	✓	✓

---

[2] Applications requesting Piz Daint must use GPU accelerated codes.

The following table lists the PRACE Tier-1 HPC systems, which are available in the context of Preparatory Access **Type D**. The specific system will be selected by PRACE, but applicants can specify preferences:

Tier-1 Systems	Architecture
ARCHER, hosted by EPCC (United Kingdom)	Cray XC30
Aris, hosted by GRNET (Greece)	Lenovo NextScale
Bem, hosted by WCSS (Poland)	Haswell based Cluster
Beskow[3], hosted by KTH-PDC (Sweden)	Cray XC40
Cartesius, hosted by SURFsara (Netherlands)	Bull Bullx B720/B710
EAGLE, hosted by PSNC (Poland)	Intel Xeon E5-2697
GALILEO, hosted by CINECA (Italy)	Lenovo NextScale
Mahti, hosted by CSC (Finland)	AMD Rome 7H12
Salomon, hosted by VSB-TUO at IT4I (Czech Republic)	SGI ICE-X

1.1 How to Apply

All proposals must be submitted via the Online Application Form.

All mandatory fields must be filled in before the application form can be submitted. After the form has been saved, applicants can still access and update it before final submission. No changes can be made after the final submission.

1.2 Eligibility

Proposals can be submitted by researchers from both public and private institutions. Only proposals with a civilian purpose will be eligible to participate in PRACE calls for proposals. PRACE Hosting Members may have further restrictions on who is eligible to access their own systems. It is the responsibility of the applicant to ensure that they are eligible to access the system(s) they have applied for.

Experienced researchers from public research organisations are eligible to apply as long as the project leader has an employment contract with her/his institution valid for at least 3 months after the end of the allocation period.

---

[3] Beskow will be replaced by January 2021.

Researchers from <u>private companies</u> are eligible to apply when the following criteria are fulfilled:

    a) The company has its head office or substantial R&D activity in Europe.
    b) The project leader has an employment contract with the organisation valid for at least 3 months after the end of the allocation period.
    c) Resources awarded are devoted solely for open R&D research purposes.

1.3 Evaluation of proposals

The evaluation of proposals is done using a light-weight procedure based on the technical assessment of the projects submitted. Scientific merits of the proposal are not evaluated.

Technical evaluations are performed by recognised technical experts from the relevant PRACE Hosting Members. In the case of proposals for **Type C** and **D** scheme, appropriate software experts from PRACE will also be involved in the review. All reviewers are instructed to evaluate the proposals as submitted. No interaction between the reviewers and the applicants is expected during the evaluation of proposals.

In case of Type D the specific Tier-1 system will be selected during the review process by the PRACE review team, the PRACE expert and taking into account the applicant's preferences, if provided in the proposal.

The results of the evaluation of applications to Type A and B schemes are communicated to the applicants within two weeks of the submission. For Type C and D scheme, the results are communicated within two months of the cut-off date following the submission of the application. For those proposals accepted, either totally or partially, this communication will include the corresponding start date for the project.

The evaluation decisions for PRACE Preparatory Access, including timing, duration and resources allocated, are not open for negotiation.

1.4 Criteria for assessment

Proposals for Preparatory Access Types B and C must provide a detailed description of the issues hindering the scalability of the code to higher computing performances, so as to facilitate the technical assessment. In addition to this, a clear plan for code development must be proposed and the available human resources and expertise (available and/or required) to execute this planning must be quantified.

Proposals for Preparatory Access Type D must provide information about the system(s) used so far and a short description of scalability issues already discovered (if known). In addition, similar to Type B and C, a clear work plan for code development must be proposed and the available human resources and expertise (available and/or required) must be specified.

1.5 Award of resources

The awarding of Preparatory Access proposals is <u>limited by the amount of resources available</u>. The systems and schemes available will be regularly updated on the PRACE website. For those proposals with a positive technical evaluation, awarding will be based on a first-come, first-served basis, according to the final submission date.

For Type C and D, there are also limitations regarding the availability of the support from PRACE experts, in terms of staff hours, and the type of expertise required. The expertise is provided by PRACE Hosting Members, and may not be available for certain research fields. If the requested support is not available from PRACE experts, the proposal will be rejected.

1.6 PRACE HPC Systems Available

The Tier-0 computer systems and their operations accessible through PRACE are provided by the following PRACE members: BSC representing Spain, CINECA representing Italy, ETH Zurich/CSCS representing Switzerland, GCS representing Germany and GENCI representing France. These members offer the following Tier-0 HPC systems:

1.6.1 HAWK

HAWK is the new HPC system at HLRS. HAWK provides 5632 nodes, each one equipped with next generation AMD processors, code name Rome, offering 128 cores and 256 GByte of main memory per node. The nodes are connected with a next generation Infiniband network.

1.6.2 Joliot-Curie KNL

Bull Bullx cluster - hosted by GENCI at CEA, Bruyères-Le-Châtel, France. Details and terms of usage can be found here.

The Joliot-Curie KNL partition is composed of 3 cells each containing 276 nodes with one Intel Knights Landing 68-core 7250 1.4 GHz manycore processor with 16 GB of high-speed memory (MCDRAM) and 96 GB of main memory. These 3 cells are interconnected by a BULL BXI 100 Gb/s high speed network. A KNL node provides 64 cores for user jobs and keeps 4 cores for the system. A node is configured in quadrant for the cluster node and in cache mode for the memory.

1.6.3 Joliot-Curie Rome

Bull Bullx cluster - hosted by GENCI at CEA, Bruyères-Le-Châtel, France. Details and terms of usage can be found here.

The Joliot-Curie Rome partition is a BULL Sequana XH2000 system and is composed of 5 cells containing 2292 AMD Rome (Epyc) 2.6 GHz bi-processor compute nodes with 64-core per processor, 293 376 compute cores for 11.75 PFlop/s peak power and 2 GB/core. These 5 cells are interconnected by an Infiniband HDR100.

1.6.4 Joliot-Curie SKL

Bull Bullx cluster - hosted by GENCI at CEA, Bruyères-Le-Châtel, France. Details and terms of usage can be found here.

The Joliot-Curie SKL partition is composed of 6 cells, each containing 272 compute nodes with two 24-core Intel Skylake 8168 processors 2.7 GHz, 4 GB/core (192 GB/node). These 6 cells are interconnected by an Infiniband EDR 100 Gb/s high speed network.

1.6.5 JUWELS Cluster

Bull *Sequana system hosted by GCS in JSC, Jülich, Germany.* Details and terms of usage can be found here.

JUWELS (Jülich Wizard for European Leadership Science) is designed as a modular system. The JUWELS Cluster module, supplied by Atos, based on its Sequana architecture, consists of about 2500 compute nodes, each with two Intel Xeon 24-core Skylake CPUs and 96 GiB of main memory. The compute nodes are interconnected with a Mellanox EDR InfiniBand interconnect. The peak performance of this CPU based cluster partition is 10.4 petaflops. A booster module, optimized for massively parallel workloads, is currently scheduled for 2020.

1.6.6 JUWELS Booster[4]

The JUWELS Booster bases on the Sequana platform by Atos and is optimized for massively parallel workloads. It offers 936 nodes, AMD EPYC host CPUs and the latest generation on NVIDIA Ampere A100 GPUs. The Cluster and Booster are tightly integrated in the same InfiniBand fabric and can be coupled together to perform modular workflows. The JUWELS Booster is installed in 2020, therefore proposals for PRACE preparatory access can only be accepted after its official start of production.

1.6.7 Marconi100

Marconi100 is an IBM machine with 980 nodes (+ 8 login). Each node is equipped with 2 IBM POWER9 AC922 at 3.1 GHz (32 cores per node), 4 NVIDIA Volta V100 GPUs, Nvlink 2.0, 16GB, 256 GB of DDR4 RAM and 1.6 TB of Optane Memory. The used network is a Mellanox Infiniband EDR with DragonFly+ topology. It should be ready for production in April 2020, therefore it will be available for Preparatory Access projects only after this date.

1.6.8 MareNostrum

Lenovo System SD530 Compute Racks – hosted by BSC in Barcelona, Spain. Details and terms of usage will be made available here.

MareNostrum is based on Intel Xeon Platinum 8160 2.1 GHz (two 24-core CPUs, 48 cores per node), 2 GB/core (96 GB/node) and around 200 GB of local disk acting as local /tmp. A total of 48 racks, each with 72 compute nodes. All computer nodes are interconnected through an Intel Omni-Path network, with a full-fat tree network topology. MareNostrum has a peak performance of 11.15 Petaflops.

---

[4] Applications requesting JUWELS Booster hosted by GCS at JSC must use GPU accelerated codes.

1.6.9 Piz Daint[5]

Cray XC50 System – hosted by ETH Zurich/CSCS, Lugano, Switzerland. Details and terms of usage will be made available here.

Named after Piz Daint, a prominent peak in Grisons that overlooks the Fuorn pass, this supercomputer is a hybrid Cray XC50 system and is the flagship system for national HPC Service. The compute nodes are equipped with Intel® Xeon® E5-2690 v3 @ 2.60GHz (12 cores) and NVIDIA® Tesla® P100 16GB, and 64 GB of host memory.

The nodes are connected by the "Aries" proprietary interconnect from Cray, with a dragonfly network topology.

1.6.10   SuperMUC - NG

Lenovo ThinkSystem – hosted by GCS in LRZ, Garching, Germany. Details and terms of usage will be made available here.

SuperMUC-NG provides 6 480 Lenovo ThinkSystem dual socket nodes equipped with 24 core Intel Skylake EP processors and 96 GB of main memory. A subset of 144 fat nodes will be equipped with 768 GB of main memory. The nodes are connected via a fat-tree Omni-Path network. The Peak Performance will be at 26.7PF.

The maximum time allocations (in core hours) for scheme A, B and C for these systems are listed in the following table:

PRACE System	Type A	Type B	Type C
HAWK, hosted by GCS at HLRS (Germany)	50 000	250 000	250 000
Joliot-Curie KNL, hosted by GENCI at CEA (France)	50 000	200 000	200 000
Joliot-Curie Rome, hosted by GENCI at CEA (France)	50 000	200 000	200 000
Joliot-Curie SKL, hosted by GENCI at CEA (France)	50 000	200 000	200 000
JUWELS Booster, hosted by GCS at JSC (Germany)[6]	10 000	25 000	25 000
JUWELS Cluster, hosted by GCS at JSC (Germany)	20 000	50 000	50 000
Marconi100, hosted by CINECA (Italy)	12 500	25 000	25 000

---

[5] Applications requesting Piz Daint at ETH Zurich/CSCS must use GPU accelerated codes.

[6] The factor between the amount of core hours towards the available node hours on the JUWELS Booster is given by the number of host CPUs on each individual node (48 host CPUs per node).

MareNostrum, hosted by BSC (Spain)	50 000	100 000	100 000
Piz Daint, hosted by ETH Zurich/CSCS (Switzerland)	100 000	100 000	Not available
SuperMUC-NG, hosted by GCS at LRZ (Germany)	50 000	100 000	100 000

1.6.11 Tier-1 systems for Preparatory Access Type D

For scheme D Tier-1 computer systems and their operations accessible through PRACE are provided by the following PRACE members: CINECA representing Italy, CSC representing Finland, EPCC representing the United Kingdom, GRNET representing Greece, IT4I representing the Czech Republic, KTH-PDC representing Sweden, PSNC and WCSS representing Poland, and SURFsara representing the Netherlands.

The maximum time allocations (in core hours) for scheme D for Tier-1 systems and links to the details and terms of usage are listed in the following table:

Systems	Type D	Details
ARCHER, hosted by EPCC (United Kingdom)	150 000	Link
Aris, hosted by GRNET (Greece)	150 000	Link
Bem, hosted by WCSS (Poland)	150 000	Link
Beskow, hosted by KTH-PDC (Sweden)	150 000	Link
Cartesius, hosted by SURFsara (Netherlands)	150 000	Link
EAGLE, hosted by PSNC (Poland)	150 000	Link
GALILEO, hosted by CINECA (Italy)	150 000	Link
Mahti, hosted by CSC (Finland)	150 000	Link
Salomon, hosted by VSB-TUO at IT4I (Czech Republic)	150 000	Link

1.7 Terms of access

The Project Leader, also known as Principal Investigator (PI), shall lead the project and is expected to be an essential participant in its activities. The PI will have the overall responsibility for the management of the project and interactions with PRACE.

When accepting PRACE resources, awardees commit to:

a) Provide a final report to PRACE, for each Tier-0 system awarded, using the proper PRACE template, with the results obtained through the access to the PRACE Tier-0 system, as well as a qualitative feedback on the use of the resources. This report needs to be provided within the period established in the guide for applicants (see here). Type D proposals only have to provide one final report with the combined information of the granted Tier-0 and Tier-1 resources.

b) Acknowledge the role of the HPC centre(s) and of PRACE in all publications which include results from PRACE allocations. Users shall use the following wording in such acknowledgement in all papers and other publications:

*"**We acknowledge PRACE for awarding** us access to [resource-name hosted by at site] as part of PRACE Preparatory Access Type A/B/C during PERIOD OF TIME."*

Use as many instances of the pattern [resource-name hosted by at site] as the number of systems awarded via PRACE. Please follow these examples:

- HAWK at GCS@HLRS, Germany
- Joliot-Curie KNL at GENCI@CEA, France
- Joliot-Curie Rome at GENCI@CEA, France
- Joliot-Curie SKL at GENCI@CEA, France
- JUWELS Booster at GCS@JSC, Germany
- JUWELS Cluster at GCS@JSC, Germany
- Marconi100 at CINECA, Italy
- MareNostrum at Barcelona Supercomputing Center (BSC), Spain
- Piz Daint at ETH Zurich/CSCS, Switzerland
- SuperMUC-NG at GCS@LRZ, Germany

Respecting the words in bold above is very important since PRACE will use this word pattern when searching for bibliographic references in scientific articles.

c) Where technical support has been received, the following additional text should also be used:

"The support of [name of person/people] from [organisation name], [country] to the technical work is gratefully acknowledged."

d) Allow PRACE to publish the submitted report after one year from the termination of the allocation period.

e) **Contribute** to PRACE dissemination activities, including active participation in the annual PRACE Scientific and Industrial Conference (PRACEdays) and presentations at the PRACE booth at other HPC events. Selected awardees are expected to contribute to and attend such events at least once over the two year period starting from the end of the allocation period. Awardees will also be expected to reply favourably when asked to be interviewed for PRACE publications and/or send visualisations or other materials for promotional purposes.

Access will be free of charge but conditional on the fulfilment of the eligibility criteria and terms of access described in this document and in the Online Application Form. If this differs from the terms of access that the relevant Hosting Member may have in place, the terms of access of that Hosting Member will prevail.

Users will not hold PRACE and the relevant Members, including their Directors and staff, liable to any claim or expense arising out of the use of the awarded resources.

From the start to the end of the access period, the applicant should direct questions and requests for support to the user support of the Hosting Member(s) where resources have been awarded.

Applicants must inform promptly the Peer Review Team (via peer-review@prace-ri.eu) and the Hosting Member where the resources are allocated if resources cannot be consumed totally or partially during the allocation time. Any extension of the allocation has to be also requested through the Peer Review Team.

About PRACE
The Partnership for Advanced Computing in Europe (PRACE) is an international non-profit association with its seat in Brussels. The PRACE Research Infrastructure provides a persistent world-class high performance computing service for scientists and researchers from academia and industry in Europe. The computer systems and their operations accessible through PRACE are provided by 5 PRACE members (BSC representing Spain, CINECA representing Italy, ETH Zurich/CSCS representing Switzerland, GCS representing Germany and GENCI representing France). The Implementation Phase of PRACE receives funding from the EU's Horizon 2020 Research and Innovation Programme (2014-2020) under grant agreement 823767. For more information, see https://prace-ri.eu/.