E-Infrastructures
H2020- INFRAEDI-2018-2020

INFRAEDI-01-2018: Pan-European High Performance Computing infrastructure and services (PRACE)

PRACE-6IP

PRACE Sixth Implementation Phase Project

Grant Agreement Number: INFRAEDI-823767

D6.6

Final report on the PRACE and EuroHPC Operational Services

Final

Version: 1.1
Author(s): Filip Stanek, IT4Innovations
Date: 21.12.2022

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: INFRAEDI-823767**	
	Project Title: PRACE Sixth Implementation Phase Project	
	Project Web Site: https://www.prace-ri.eu/about/ip-projects/	
	Deliverable ID: D6.6	
	Deliverable Nature: Report	
	Dissemination Level: PU*	**Contractual Date of Delivery:** 31 / 12 / 2022
		Actual Date of Delivery: 19 / 12 / 2022
	EC Project Officer: Leonardo Flores Añover	

* - The dissemination level are indicated as follows: **PU** – Public, **CO** – Confidential, only for members of the consortium (including the Commission Services) **CL** – Classified, as referred to in Commission Decision 2005/444/EC.

Document Control Sheet

Document	**Title: Final report on the PRACE and EuroHPC Operational Services**	
	ID: D6.6	
	Version: 1.1	**Status:** ***Final***
	Available at: https://www.prace-ri.eu/about/ip-projects/	
	Software Tool: Microsoft Word 2016	
	File(s): D6.6-v1.1.docx	
Authorship	**Written by:**	Filip Stanek, IT4Innovations
	Contributors:	Ralph Niederberger, JUELICH Miroslav Puskaric, HLRS Zoltán Kiss, KIFÜ Alessandro Marani, CINECA Michael Mucciardi, BSC Mirosław Kupczyk, PSNC Lukasz Flis, CYFRONET Abdulrahman Azab, UiO Cristiano Padrin, CINECA
	Reviewed by:	Dirk Brömmel, JUELICH Lilit Axner, SNIC
	Approved by:	MB/TB

Document Status Sheet

Version	**Date**	**Status**	**Comments**
0.1	27/04/2022	Draft	First draft of Operational Services
0.2	23/06/2022	Draft	Inputs added for Security and Network services
0.3	24/06/2022	Draft	Inputs added for Data service, Generic services

Version	Date	Status	Comments
			and Status of TIER-0 & TIER-1 sites
0.4	30/06/2022	Draft	Inputs added for Integration of new and upgrade of Tier-0/Tier-1 systems
0.5	15/07/2022	Draft	Inputs added for Monitoring service
0.6	02/08/2022	Draft	Input added for Lightweight virtualization service
0.7	20/10/2022	Draft	Input added for AAA service
0.8	21/11/2022	Draft	Input added for Energy monitoring and consumption
0.9	23/11/2022	Draft	Draft version ready for internal review process, submitted for MB/TB review
1.0	12/12/2022	Draft	Final DRAFT for MB/TB approval
1.1	21/12/2022	Final version	Final version ready for EC submission

Document Keywords

Keywords:	PRACE, HPC, Research Infrastructure, Operational Services, Tier-0 Systems, Tier-1 Systems

Table of Contents

List of Figures

List of Tables

References and Applicable Documents

[1] PRACE-1IP Deliverable D1.6, Assessment of PRACE operational structure, procedures and policies, PRACE, June 2011, https://prace-ri.eu/wp-content/uploads/1IP-D6.1.pdf
[2] PRACE-6IP Deliverable D6.4, Final report on new prototypal services, PRACE, December 2021, https://prace-ri.eu/wp-content/uploads/PRACE-6IP-D6.4.pdf
[3] D. Cesarini, A. Bartolini, P. Bonfà, C. Cavazzoni, and L. Benini, “COUNTDOWN: a run-time library for application-agnostic energy saving in MPI communication primitives,” in Proceedings of the 2nd Workshop on AutotuniNg and aDaptivity AppRoaches for Energy efficient HPC Systems (ANDARE '18). Association for Computing Machinery, New York, NY, USA, Article 2, 1–6, November 2018, doi:10.1145/3295816.3295818.
[4] D. Cesarini, A. Bartolini, A. Borghesi, C. Cavazzoni, M. Luisier and L. Benini, "Countdown Slack: A Run-Time Library to Reduce Energy Footprint in Large-Scale MPI Applications," in IEEE Transactions on Parallel and Distributed Systems, vol. 31, no. 11, pp. 2696-2709, 1 Nov. 2020, doi: 10.1109/TPDS.2020.3000418.
[5] D. Cesarini, A. Bartolini, P. Bonfà, C. Cavazzoni and L. Benini, "COUNTDOWN: A Run-Time Library for Performance-Neutral Energy Saving in MPI Applications," in IEEE Transactions on Computers, vol. 70, no. 5, pp. 682-695, 1 May 2021, doi: 10.1109/TC.2020.2995269.

List of Acronyms and Abbreviations

AAA	Authorisation, Authentication, Accounting
AAI	Authorisation and Authentication Infrastructure
AARC	Authentication and Authorisation for Research and Collaboration
AEGIS	AARC Engagement Group for Infrastructures
aisbl	Association International Sans But Lucratif (legal form of the PRACE-RI)
APGridPMA	The Asian Pacific Grid Policy Management Authority
BGP	Border Gateway Protocol
BSS	Batch Scheduling System
BDW	Intel Broadwell CPU family
CA	Certificate Authority
CLI	Command Line Interfaces
CoE	Center of Excellence
CPU	Central Processing Unit
CP/CPS	Certificate Policy/Certification Practice Statement
CSIRT	Computer Security Incident Response Team
CUDA	Compute Unified Device Architecture (NVIDIA)
DARPA	Defense Advanced Research Projects Agency
DART	Distributed Accounting Reporting Tool
DEISA	Distributed European Infrastructure for Supercomputing Applications EU project by leading national HPC centres
DoA	Description of Action (formerly known as DoW)
EC	European Commission
EESI	European Exascale Software Initiative
EoI	Expression of Interest
EOSC	European Open Science Cloud
ESFRI	European Strategy Forum on Research Infrastructures
EUDAT	European Data Infrastructure
EUGridPMA	European Grid Policy Management Authority
Fenix	Fenix Research Infrastructure
GB	Giga (= 2^{30} ~ 10^9) Bytes (= 8 bits), also GByte
Gb/s	Giga (= 10^9) bits per second, also Gbit/s
GB/s	Giga (= 10^9) Bytes (= 8 bits) per second, also GByte/s
GCT	Grid Community Toolkit
GÉANT	Collaboration between National Research and Education Networks to build a multi-gigabit pan-European network. The current EC-funded project as of 2015 is GN4.
GFlop/s	Giga (= 10^9) Floating point operations (usually in 64-bit, i.e. DP) per second, also GF/s
GHz	Giga (= 10^9) Hertz, frequency =10^9 periods or clock cycles per second
GPU	Graphic Processing Unit
GridCF	Grid Community Forum
GSI	Grid Security Infrastructure
GT	Globus Toolkit
HET	High Performance Computing in Europe Taskforce. Taskforce by representatives from European HPC community to shape the European HPC Research Infrastructure. Produced the scientific case and valuable groundwork for the PRACE project.

HPC	High Performance Computing; Computing at a high performance level at any given time; often used synonym with Supercomputing
HPL	High Performance LINPACK
HTML	HyperText Markup Language
IdP	Identity Provider
IGTF	Interoperable Global Trust Federation
ISC	International Supercomputing Conference; European equivalent to the US based SCxx conference. Held annually in Germany.
KB	Kilo (= 2^{10} ~10^3) Bytes (= 8 bits), also Kbyte
KPI	Key Performance Indicator
KNL	Intel Knights Landing CPU family
LoA	Level of Assurance
LDAP	Lightweight Directory Access Protocol
LINPACK	Software library for Linear Algebra
MB	Management Board (highest decision making body of the project)
MB	Mega (= 2^{20} ~ 10^6) Bytes (= 8 bits), also MByte
MB/s	Mega (= 10^6) Bytes (= 8 bits) per second, also MByte/s
MD-VPN	Multi Domain Virtual Private Network
MFlop/s	Mega (= 10^6) Floating point operations (usually in 64-bit, i.e. DP) per second, also MF/s
MOOC	Massively open online Course
MoU	Memorandum of Understanding.
MPI	Message Passing Interface
NDA	Non-Disclosure Agreement. Typically signed between vendors and customers working together on products prior to their general availability or announcement.
NeIC	Nordic e-Infrastructure Collaboration
OoD	Operator on Duty
OS	Operating System
PA	Preparatory Access (to PRACE resources)
PATC	PRACE Advanced Training Centres
PCPE	PRACE Common Production Environment
PFlop/s	Peta (= 10^{15}) Floating-point operations (usually in 64-bit, i.e. DP) per second, also PF/s
PKI	Public Key Infrastructure
PMA	Policy Management Authority
PRACE	Partnership for Advanced Computing in Europe; Project Acronym
PRACE 2	The upcoming next phase of the PRACE Research Infrastructure following the initial five year period.
PTC	PRACE Training Centres
RHEL	Red Hat Enterprise Linux
RI	Research Infrastructure
RSA	rsa.com, Company specialized on authentication services based on the RSA algorithm
RT	Request Tracker, same as TTS
SCI	Security for Collaborating Infrastructures
SSH	Secure Shell
SVN	SubVersioN: software versioning and revision system
TAGPMA	The Americas Grid PMA
PPRT	PRACE Project Proposal Review Tool
PRACE BoD	PRACE Board of Directors

PRACE TB	PRACE 5IP Technical Board (group of Work Package leaders)
PUHURI	Puhuri Project, NeIC
REFEDS	Research and Education FEDerations group
SP	Service Provider
TB	Tera (= 2^{40} ~ 10^{12}) Bytes (= 8 bits), also TByte
TCO	Total Cost of Ownership. Includes recurring costs (e.g. personnel, power, cooling, maintenance) in addition to the purchase cost.
TDP	Thermal Design Power
TFlop/s	Tera (= 10^{12}) Floating-point operations (usually in 64-bit, i.e. DP) per second, also TF/s
Tier-0	Denotes the apex of a conceptual pyramid of HPC systems. In this context the Supercomputing Research Infrastructure would host the Tier-0 systems; national or topical HPC centres would constitute Tier-1
TTS	Trouble Ticket System, same as RT
UEABS	Unified European Application Benchmark Suite
UNICORE	Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.
VPN	Virtual Private Network
WISE	Wise Information Security for collaborating E-infrastructures
WP	PRACE Work Package

List of Project Partner Acronyms

BADW-LRZ	Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften, Germany (3rd Party to GCS)
BILKENT	Bilkent University, Turkey (3rd Party to UHEM)
BSC	Barcelona Supercomputing Center - Centro Nacional de Supercomputacion, Spain
CaSToRC	The Computation-based Science and Technology Research Center (CaSToRC), The Cyprus Institute, Cyprus
CCSAS	Computing Centre of the Slovak Academy of Sciences, Slovakia
CEA	Commissariat à l'Energie Atomique et aux Energies Alternatives, France (3rd Party to GENCI)
CENAERO	Centre de Recherche en Aéronautique ASBL, Belgium (3rd Party to UANTWERPEN)
CESGA	Fundacion Publica Gallega Centro Tecnológico de Supercomputación de Galicia, Spain, (3rd Party to BSC)
CINECA	CINECA Consorzio Interuniversitario, Italy
CINES	Centre Informatique National de l'Enseignement Supérieur, France (3 rd Party to GENCI)
CNRS	Centre National de la Recherche Scientifique, France (3 rd Party to GENCI)
CSC	CSC Scientific Computing Ltd., Finland
CSIC	Spanish Council for Scientific Research (3rd Party to BSC)
CYFRONET	Academic Computing Centre CYFRONET AGH, Poland (3rd Party to PNSC)
DTU	Technical University of Denmark (3rd Party of UCPH)
EPCC	EPCC at The University of Edinburgh, UK
EUDAT CDI	EUDAT Collaborative Data Infrastructure
ETH Zurich (CSCS)	Eidgenössische Technische Hochschule Zürich – CSCS, Switzerland

GCS	Gauss Centre for Supercomputing e.V., Germany
GÉANT	GÉANT Vereniging
GENCI	Grand Equipement National de Calcul Intensif, France
GRNET	National Infrastructures for Research and Technology, Greece
ICREA	Catalan Institution for Research and Advanced Studies (3rd Party to BSC)
INRIA	Institut National de Recherche en Informatique et Automatique, France (3rd Party to GENCI)
IST-ID	Instituto Superior Técnico for Research and Development, Portugal (3rd Party to UC-LCA)
IT4I/VSB-TUO	Vysoka Skola Banska - Technicka Univerzita Ostrava, Czech Republic
IUCC	Machba - Inter University Computation Centre, Israel
JUELICH	Forschungszentrum Jülich GmbH, Germany
KIFÜ (NIIFI)	Governmental Information Technology Development Agency, Hungary
KTH	Royal Institute of Technology, Sweden (3rd Party to SNIC-UU), also PDC-KTH
KULEUVEN	Katholieke Universiteit Leuven, Belgium (3rd Party to UANTWERPEN)
LiU	Linkoping University, Sweden (3rd Party to SNIC-UU)
MPCDF	Max Planck Gesellschaft zur Förderung der Wissenschaften e.V., Germany (3rd Party to GCS)
NCSA	National Centre for Supercomputing Applications, Bulgaria
NTNU	The Norwegian University of Science and Technology, Norway (3rd Party to SIGMA2)
NUI-Galway	National University of Ireland Galway, Ireland
PRACE	Partnership for Advanced Computing in Europe aisbl, Belgium
PSNC	Poznan Supercomputing and Networking Center, Poland
SDU	University of Southern Denmark (3rd Party to UCPH)
SIGMA2	UNINETT Sigma2 AS, Norway
SNIC-UU	Uppsala Universitet, Sweden
STFC	Science and Technology Facilities Council, UK (3rd Party to UEDIN)
SURF	SURF is the collaborative organisation for ICT in Dutch education and research
TASK	Politechnika Gdańska (3rd Party to PNSC)
TU Wien	Technische Universität Wien, Austria
UANTWERPEN	Universiteit Antwerpen, Belgium
UC-LCA	Universidade de Coimbra, Labotatório de Computação Avançada, Portugal
UCPH	Københavns Universitet, Denmark
UEDIN	The University of Edinburgh
UHEM	The National Center for High Performance Computing, Turkey
UIBK	Universität Innsbruck, Austria (3rd Party to TU Wien)
UiO	University of Oslo, Norway (3rd Party to SIGMA2)
UL	Univerza V Ljubljani, Slovenia
ULIEGE	Université de Liège; Belgium (3rd Party to UANTWERPEN)
U Luxembourg	University of Luxembourg
UM	Universidade do Minho, Portugal, (3rd Party to UC-LCA)
UmU	Umea University, Sweden (3rd Party to SNIC-UU)
UnivEvora	Universidade de Évora, Portugal (3rd Party to UC-LCA)
UnivPorto	Universidade do Porto, Portugal (3rd Party to UC-LCA)
UPC	Universitat Politècnica de Catalunya, Spain (3rd Party to BSC)

USTUTT-HLRS	Universitaet Stuttgart – HLRS, Germany (3rd Party to GCS)
WCSS	Politechnika Wroclawska, Poland (3rd Party to PNSC)

Executive Summary

The deliverable presents the activities done in the 1st extension (January 2022 – June 2022) and the 2nd extension (July 2022 – December 2022) of PRACE-6IP covering the operation and coordination of the PRACE Operational Services. The operation of PRACE distributed HPC infrastructure involves the coordination of a set of services which integrate the Tier-0 systems and a number of national Tier-1 systems into a single pan-European HPC infrastructure. It needs to be mentioned that the 1st extension was reduced to nineteen partners with funded effort, while the 2nd extension reduced the funding further to three partners only to maintain the Generic services. These are of most importance for the general public and PRACE users. The rest of activities in second extension was cancelled with the exemption of services needed for smooth transition from PRACE to new European or national activities like EuroHPC. Such services were provided in a light-weight mode, in-kind by the individual partners. This work is the continuation of the work done by Task 6.1 in the previous period of PRACE-6IP as well as previous PRACE-IP projects to provide continuity to the PRACE Operational Services for the HPC ecosystem.

Twelve Tier-0 systems in five countries were operational in the first half of the year 2022. Furthermore, operational support has been provided to nineteen national Tier-1 systems that provide services for Tier-0 (i.e. as stepping stone towards Tier-0 systems, or to prototype and asses new operational services). These Tier-1 systems are distributed among thirteen different countries, ensuring a wide distribution of the European HPC ecosystem and hosting DECI project call awardees.

Based on the procedures for incident and change management, a set of common services as defined in the PRACE Service Catalogue (Networking, Data, AAA, Monitoring and Generic) has been operated and monitored on a day-by-day basis to assure continuity and integrity of the services.

A new activity from Task 6.2, called "Lightweight virtualization", was introduced to Task 6.1 after the completion of Task 6.2. It resulted in a three-day long workshop in Ljublana (Slovenia), held by PRACE and EuroHPC hosting site representatives.

Further development to establish a new federated approach to PRACE AAI was pursued with work addressing the architecture both from a technical as well as a procedural point of view. A pilot of federated access based on the architecture was done.

The Security Forum, responsible for all security related activities, was also coordinated by Task 6.1. All activities were coordinated with bi-weekly videoconferences of all participating partners to monitor the infrastructure and prevent possible incidents which could cause vulnerabilities on the PRACE RI. A continuous operator on duty service was provided by partners overseeing the state of systems, network, and help desk.

A structured plan to decommission PRACE services after the end of PRACE-6IP, keeping in mind the archival of valued knowledge and data of all the PRACE–IP projects, was setup and executed.

The request from a previous PRACE review to address the topic "Energy consumption monitoring and reduction" in HPC was taken up on top of the standard activities of this task.

1 Introduction

This deliverable describes the activities performed in Task 6.1 “Operation and coordination of the comprehensive common PRACE Operational Services” of WP6 “Operational Services for the HPC ecosystem” in PRACE-6IP. This task is responsible for the operation of the set of common services, which present the PRACE Tier-0 and Tier-1 systems as an integrated pan-European HPC ecosystem. The task supports the PRACE calls, the DECI calls, the SHAPE activity towards SMEs, and the prototyping and assessment of new operational services investigated in Tasks 6.2 and 6.3 of PRACE-6IP WP6. Other supported activities are the testing and utilisation of specific architectures and technologies, which are only available in specific countries. Finally, the task also supports the daily operations of PRACE itself and other work packages in the project including the PRACE Website, Training Portal, Events Portal, CodeVault, and User Repository Service.

The operation and coordination of the common PRACE Operational Services provided in Task 6.1 continued well-established management procedures and organisation as set up already since PRACE-1IP. The task further continued the implementation of the roadmap to a professional service level of sustainable services with a defined quality of service.

This report focuses on the activities done in the 1st extension (January 2022 – June 2022) and the 2nd extension (July 2022 – December 2022) of PRACE-6IP and follows-up the activities undertaken during the previous reporting periods of PRACE-6IP.

In the 1st extension , the operation of the common PRACE Operational Services has been coordinated and monitored constantly by means of bi-weekly videoconferences. A single face-to-face meeting was held as part of the PRACE -6IP All Hands Meeting in Vienna on May 5th-6th 2022. The face-to-face and videoconference meetings aimed to discuss the status of the operational activities and to plan the decommissioning activity after the end of PRACE-6IP project. A light-weight procedure of coordination mostly based on email communication between the task leader (in-kind) and the remaining three partners was used in the 2nd extension.

Section 2 describes the status of the Tier-0 systems and the Tier-1 systems involved in the Tier-1 for Tier-0 activity, composing the PRACE HPC ecosystem. Section 3 describes the activities done mostly in the 1st extension by the different service areas:

● Network services; ● Data services; ● Authorisation, Authentication and Accounting;	● Operational security; ● Monitoring services for operations; ● Generic services.

Table 1: WP6 activities in the first extension of PRACE-6IP

Section 4 describes the new Task 6.1 activity “Lightweight virtualization” coming from Task 6.2 (which was not part of the second extension of the project) and is followed by Section 5 addressing the request from previous PRACE-6IP review, to elaborate on the topic of energy consumption monitoring and reduction on HPC systems.

Section 6 then concludes the document.

2 PRACE HPC ecosystem: Tier-0 and Tier-1 sites, system upgrades and new systems

This section presents the changes implemented during the 1st extension (January 2022 – June 2022) concerning the status of Tier-0 sites and the Tier-1 national sites providing Tier-1 for Tier-0 services. The chronology and the status of the performed system upgrades is reported in detail. In this period, the main tasks of the operational procedures used to offer the PRACE services have been kept similar to what we had in the previous reporting period. The focus was on organising and overseeing the operator and helpdesk on duty shifts, operational status of the integrated systems and network monitoring from the systems access point of view. The work was the continuation and evolution of the activity already in progress. Efforts have been made to keep the documentation (covering new integrations/upgrades of systems, list of operational systems, list of decommissioned systems) up-to-date and feedback has been periodically inquired from the participant sites.

2.1 Status of Tier-0 & Tier-1 sites

The Tier-0 and Tier-1 systems constitute an HPC ecosystem offering high level services to the European computational community. At the end of the 1st extension , according to the information recorded by all partners, 12 Tier-0 systems and 19 Tier-1 systems were fully or partially integrated and in production. All systems were continuously monitored and the operational quality was assured by employing a specific regular activity provided daily by the members of PRACE-6IP WP6 task 6.1. This On-duty activity is described below.

2.1.1 On-duty Activity

The On-duty activity was guaranteed daily by all partners who provide effort and/or systems in the WP6; this activity was assigned to partners following a weekly schedule. The topics/incidents reported through this service were mainly related to operational issues and activities needed to maintain the distributed infrastructure in good shape. Specific requests from users were rare and normally redirected to the local help desk of each individual site.

A weekly report on the On-duty activity was produced, where the operator has to report about any change of the status of the infrastructure, all occurred problems, the status and any other notification regarding the core services.

The related documentation on the wiki site was constantly updated, using the report template agreed at the beginning of PRACE-5IP and confirmed for PRACE-6IP.

Starting from 1st of January 2022, a schedule was defined for the 19 PRACE partners involved in the On-duty activity. Each of them had been in charge for monitoring the infrastructure and reporting the related issues using the trouble ticketing tool (TTS). The 19 partners involved in the schedule are reported in Table 2 below.

PRACE-6IP Partners	
1 BSC	11 KIFU
2. LRZ	12. PDC
3 CEA	13 PSNC

PRACE-6IP Partners	
4 CINECA	14. ICHEC
5 GRNET	15 SURF
6 CSC	16 IT4I/VSB-TUO
7 CYFRONET	17 UC-LCA
8 EPCC	18 ETH-CSCS
9 JUELICH	19 UNILUX
10. HLRS	

Table 2: PRACE-6IP partners involved in on-duty activity for the 2022 extension period

The TTS tool used by the On-duty activity staff was the Best Practical RT 4.2.8, an enterprise-grade issue tracking system. It is freely available under the terms of version 2 of the GNU License. It was hosted by CINECA on a specific virtual machine, and maintained since its deployment during the PRACE-3IP project.

The TTS system was organized in queues where every site was responsible for one, apart from the "Generic" queue that is in charge of the Operator-On-Duty. Table 3 lists the sites that have a specific queue defined for themselves and that they are asked to keep monitored.

PRACE-6IP Partners	
1 BSC	17 IDRIS
2 CASTORC	18 IUCC
3 CCSAS	19 JUELICH
4 CEA	20 KIFU
5 CESGA	21 LRZ
6 CINECA	22 NCSA
7 CINES	23 PDC
8 CSC	24 PSNC
9 CYFRONET	25 IT4I/VSB-TUO
10 EPCC	26 RZG
11 ETH-CSCS	27 STFC
12 GRNET	28 SUR
13 HLRS	29 UHEM
14 ICHEC	30 UIO
15 ICM	31 WCSS
16 UC-LCA	32 UL

Table 3: PRACE-6IP partners involved in the monitoring of their own TTS queue

Starting from the 1st of July 2022, the service was provided in-kind: there were no other shifts of On-duty operators, and the queues were supervised by the relative sites without funding from PRACE. Many queues were de-activated with the remaining on service until December 31st 2022. Table 4 lists the sites that have requested to keep their queue for the additional extension period (updated at June 24th 2022):

PRACE-6IP Partners	
1 HLRS	5 WCNS
2 CINECA	6 UC-LCA
3 PSNC	7 ICHEC
4 CYFRONET	

Table 4: PRACE-6IP partners involved in the monitoring of their own TTS queue during the final extension period

In the monitored period 13 tickets have been created: all of them had been resolved, meaning that a percentage of resolved tickets respect to the opened tickets is 100%.

It is important to underline that, in principle, the activity was related to the traffic on the General Queue where the tickets are normally created; however, the operator is in charge to report if a ticket moved into a site queue is not updated for more than a week, and/or if the owner is missing.

2.1.2 Production systems

Since this activity ended with the 1st extension, the last update covers the state as of June 2022. According to the shared information by all partners, the Tier-0 ecosystem was made up of twelve systems, distributed in seven sites, operated by seven different partners, in five Hosting Members countries (France, Germany, Italy, Spain and Switzerland) as reported in Table 5. Since the last update, Irene (GENCI-CEA) has increased their Tier-0 power by adding three new modules: Rome, V100 and A64FX.

Some of the Tier-0 systems consist of partitions (or modules) featuring different compute elements (mostly accelerators like GPGPU, but also different CPU architectures). Aggregate benchmarks that combine partitions into a single unit prove challenging and do not reflect the peak performance properly. For this reason, the list of TIER-0 systems shows results individually for the different partitions. The aggregated (net sum of all their partitions) peak performance of the Tier-0 systems ranges from more than 11 PFlop/s up to more than 82 PFlop/s for the JUWELS system in Germany. Five Tier-0 are accelerated: JUELICH/JUWELS (Cluster module), CINECA/MARCONI100, and GENCI/IRENE (A64FX) with Nvidia V100, JUELICH/JUWELS (Booster module) with Nvidia A100 and CSCS/PizDaint with Nvidia Tesla P100. The most dominant vendors are ATOS/Bull Sequana (six systems) and Lenovo (two systems).

Most of Tier-0 systems are ranked in the Top500 (June 2022) and the highest ranked system (JUWELS Booster module) is in 11th position. Irene (KNL, V100, A64FX) are currently out of the Top500.

As far as the Tier-1 ecosystem is concerned, 19 systems are operating as Tier-1 for Tier-0 services. These Tier-1 systems are distributed in 15 different PRACE sites, operated by 14 partners, in 13 different European countries. Table 6 presents the list of the Tier-1 systems.

The peak performance ranges from very small system partitions (<30 TFlop/s) up to large systems in excess of 25 PFlop/s. Eight systems deliver more than 1 PFlop/s.

More than half of the Tier-1 systems are accelerated, either with Intel Xeon Phi or Nvidia accelerators. Several different vendors and architectures are present which this is a real advantage for the PRACE HPC infrastructure: HPE, SGI, ATOS/Bull, Cray, IBM and Lenovo are all represented.

In the last six months, 2 Tier-1 systems were decommissioned (SALOMON, PUHTI) and 1 new system has been integrated (DARDEL).

The whole infrastructure is constantly in operational status and incidents are infrequent. In the last 6 months, only one unscheduled maintenance has been reported.

Partner	**Country**	**Tier-0**	**Architecture** **-** **CPU** **-** **Accelerator**	**Rpeak (TFlop/s)**
FZJ-JUELICH	Germany	JUWELS (Booster Module)	Atos Sequana XH2000 - AMD EPYC Rome 24-core - Nvidia A100	73008.0
		JUWELS (Cluster Module)	ATOS Sequana X1000 - Dual Intel Xeon Platinum 8168 - Nvidia V100	9891.1
GCS-LRZ		SuperMUC - NG	Lenovo OceanCat SD650 DWC - Intel SkyLake 24 core; 2.7 GHz	26900.0
GCS-HLRS		Hawk	HPE Apollo - AMD Epyc Rome 7742; 64-core; 2.25 GHz	25159.7
GENCI-CEA	France	Irene (SKL)	ATOS/Bull Sequana - Intel SkyLake 8168; 24-core; 2.70 GHz	6635.5
		Irene (KNL)	ATOS/Bull Sequana - Intel Knights Landing; 68-core; 1.40GHz	2339.6
		Irene (Rome)	ATOS/Bull Sequana - Intel SkyLake 8168; 24-core; 2.70 GHz	12200.0
		Irene (V100)	ATOS/Bull Sequana - Intel Knights Landing; 68-core; 1.40 GHz	1130.0

		Irene (A64FX)	Fujitsu PRIMEHPC FX700 - A64FX Armv8.2-A SVE; 1.8 GHz - Nvidia V100	-
BSC	Spain	Mare Nostrum4	Lenovo SD530 - SkyLake Intel Xeon Platinum 8160; 2x24-core; 2.10 GHz	11150.0
ETH-CSCS	Switzerland	Piz Daint	Cray XC50 - Intel Xeon E5-2609 v3; 12-core; 2.60 GHz - Nvidia P100	15988.0
CINECA	Italy	MARCONI 100	IBM Power 9 AC922 (Witherspoon) - 2x16 cores; 2.6 GHz (3.1 GHz) - NVidia V100	32768.0

Table 5: PRACE Tier-0 systems

Partner	Country	Tier-1	Architecture - CPU - Accelerator	Rpeak (TFlop/s)
EPCC	UK	ARCHER-2	Cray Shasta - AMD EPYC Zen2 (Rome) 64 core; 2.2 GHz	25800.0
BSC	Spain	Minotauro	Bull Bullx B505/R421-E4 - Intel Haswell E5-2630 v3; 8-core; 2.60 GHz - Nvidia K80	250.9
SURF	Netherlands	Cartesius	Bull Bullx B720/B710 - Intel Haswell; 12-core; 2.6 GHz + Intel Ivy Bridge; 12-core; 2.4 GHz + Intel Sandy Bridge; 8-core; 2.7 GHz	1349.0
			Bull Bullx B515 - Intel Ivy Bridge 8-core; 2.5 GHz - Nvidia K40	210.0
			Bull Sequana X1110 - Intel Broadwell 16 core; 2.6 GHz	236.0
			Bull Sequana X1210 - Intel Knights Landing 64-core; 1.3 GHz	48.0
		Snellius	Lenovo - AMD Rome 7H12@2.6 GHz - Nvidia A100	CPU>1500.0 GPU>1500.0
PDC-KTH	Sweden	Beskow	Cray XC40 - Intel Haswell; 32-core; 2.3 GHz	1973.0
		Dardel	Cray EX - AMD Epyc; 2x64 cores; 2.25 GHz	2280.0

PSNC-CYFRONET	Poland	Zeus BigMem	HP BL685c G7 AMD - AMD Interlagos; 16-core; 2.3 GHz	61.2
		Prometheus	HP Apollo 8000 - Intel Xeon E5-2680 v3; 12-core; 2.5 GHz	2400.0
PSNC	Poland	Eagle	Intel Cluster - Intel Haswell E5-2697 v3; 14-core; 2.60 GHz	1380.0
PSNC - WCSS	Poland	Bem	Intel Cluster - Intel Xeon E5-2670 v3; 2x12-core; 2.30 GHz	860.0
ICHEC	Ireland	Kay	Intel/Penguin Computing - Intel SkyLake Xeon 6148 CPU; 2x20-core; 2.4 GHz	665.0
KIFU	Hungary	Leo	HP SL250s - Intel Xeon E5-2650 v2; 2.60 GHz - Nvidia K20, K40	254.0
		PHItagoras	HP SL250s - Intel Xeon E5-2680 v2; 2.80 GHz - Intel/PHI 7120	27.0
GRNET	Greece	ARIS	IBM NeXtScale nx360 M4 - Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80 GHz	190.8
			Dell PowerEdge R820 - Intel(R) Xeon(R) CPU E5-4650 v2 @ 2.40 GHz	36.6
			Dell PowerEdge R730 - Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60 GHz- Nvidia Tesla K40m and Intel Xeon Phi 7120	244.3
			Supermicro SYS-4028GR-TVRT -	63.1

			Intel(R) Xeon(R) CPU E5-2698 v4 @ 2.20 GHz- Nvidia Tesla V100-SXM2	
CESGA	Spain	Finisterrae	Bull Bullx B505/R424-E4 - Intel Haswell 12-core E5-2680v3; 2.50 GHz - NVidia K80	328.0
UL	Slovenia	HPCFS-U	Intel x64 HPE + Supermicro blades - Intel Xeon E5-2680V3 - NVIDIA Tesla K80	40.0
UHEM	Turkey	Sariyer	Intel x64 (Super Micro / Huawei / DELL) - Intel Xeon (E5-2680V3;2.50GHz, E5-2680V4; 2.40 GHz, Gold 6148; 2.40 GHz) - NVIDIA Tesla K20m	225.0
CSC	Finland	Mahti	Atos BullSequana XH2000 - EPYC 7H12	7060.0
UC-LCA	Portugal	Navigator	Fujitsu Primergy - Intel(R) Xeon(R) CPU E5-2697 v2 @ 2.70 GHz+ Intel(R) Xeon(R) Gold 6148 CPU @ 2.40 GHz+ Intel(R) Xeon(R) Gold 6154 CPU @ 3.00 GHz+ - NVIDIA V100	246.0

Table 6: PRACE Tier-1 systems

2.2 System Upgrades

This section describes the final activities related to the upgrades and integration of Tier-0 or Tier-1 systems into the PRACE ecosystem. These Tier-0 or Tier-1 systems provide the Tier-0 for Tier-1 services.Operational procedures for new systems and system upgrades

Operational procedures were well documented in the PRACE WIKI and BSCW including:

- Integration of new Tier-0/Tier-1 sites
  - Procedure/Template:
    - Updated WIKI template to differentiate core from optional services. The template accurately reflects the latest PRACE Service Catalogue that new sites must fulfil in order to complete their integration into the PRACE

infrastructure. This template includes links to help guide the user towards completion of core/optional tasks related to system integration.

- Information for new Tier-1 sites
  - BSCW presentation about basic concepts that new sites should know before starting their integration into PRACE infrastructure
  - Introduction Email outlining steps for obtaining access to relevant online resources needed for system integration and corresponding contacts that are responsible for individual PRACE service.

- Upgrade of systems (Tier-0/Tier-1)
  - Procedure for upgrade of systems:
    - WIKI guide with information to upgrade systems in the PRACE infrastructure
  - Report template for upgrades
    - WIKI template shared with site representatives so that they can self-document their system integration progress. Each unique system has its own tracking page which system owners are responsible for updating. This tracking page is used to keep track of the current integration status of each system.

This documentation ensures that all systems follow equal procedures and are in line with the service catalogue requirements. The service catalogue was introduced in PRACE-1IP [1] and updated since to reflect needed changes in the operational setup of PRACE.

2.2.1 Overview of System Upgrades in PRACE-6IP

The previous report mentioned two systems being upgraded or replaced, starting the reintegration process of those systems into PRACE services. Since the reintegration was still ongoing in the 1st extension, we mention them again here. A third system, Dardel from SNIC, requested an integration tracking page on April 4th 2022, but due to the ongoing installation that exceeded the 1st extension it has not completed full integration into PRACE services as of June 2022.

Here are the details of the three systems undergoing the "upgrade" and reintegration procedure:

- UK - EPCC – Archer2 replaced Archer

Created a new system integration page for this site which was shared with EPCC. As of June 2022 there were still some services missing.

- Sweden – Beskow

After the system had undergone an upgrade it needed to be re-integrated into PRACE services. Beskow had completed most but not all core tasks related to system re-integration into PRACE services, but was replaced by the installation of the Dardel system and thus its integration has been terminated.

- Sweden – Dardel.

This new system is currently undergoing installation at PDC/KTH-SNIC and began integration tracking on April 29th 2022.

More information about the systems Archer2, Beskow, and Dardel can be found in the Table 7.

2.3 New Tier-0/Tier-1 sites and systems

This section presents the systems that were decommissioned and gives an overview of new sites and systems that were planned to finish the integration in the PRACE infrastructure during PRACE-6IP.

During PRACE-6IP, sixteen systems had been planned to be integrated into the PRACE infrastructure and to start providing services to it, as can be seen in the table below. Of these 16 systems, one system began integration in PRACE-4IP and is still in progress, six systems began integration during 5IP, and nine systems began integration during PRACE-6IP. Additionally, three systems from PRACE 3IP were decommissioned.

See the below summary of systems that have either completed integration, been cancelled, decommissioned, or are still "in progress" for integration into PRACE services. Please note that systems pre-dating PRACE-4IP have been integrated and are only being tracked when they are decommissioned.

Updated status of systems that began integration during 3IP:

- Anselm (IT4I/VSB-TUO) - Decommissioned
- Archer (EPCC) - Decommissioned
- Abel (Sigma) – Decommissioned

Updated status of systems that began integration during 4IP:

- FinisTerrae (CESGA) – Partially integrated
    - Missing less than 10% of services integrated.

Updated status of systems that began integration during 5IP:

- Aurel (CCSAS) - Decommissioned before integration was complete
- Sariyer (UHeM)- Completed integration
- Navigator (UC-LCA)- Completed integration
- HPCFS (UL) - Partially integrated
    - Missing less than 10% of services integrated.
- Irene (GENCI/CEA)- Partially integrated
    - Missing less than 10% of services integrated.
- Barbora (IT4I)- Partially integrated
    - Missing less than 10% of services integrated.

Updated status of systems that began integration during 6IP:

- Cy-Tera (CaSToRC)- Cancelled
- Galileo (CINECA) - Decommissioned before integration was complete
- Snellius (SURF) - Completed integration

- Hawk (GCS)- Completed integration
- Puhti (CSC) – Decommissioned before integration was complete
- Mahti (CSC) - Partially integrated
  - Missing less than 10% of services integrated.
- Archer2 (EPCC) - Partially integrated
  - Missing integration of some services.
- Bob (MAC) - Partially integrated
  - Delayed installation prevented full integration during the 1st extension.
- Beskow (SNIC) – Partially integrated
  - Replaced by Dardel system before full integration.
- Dardel (SNIC) - Partially integrated
  - Installation of the new system was prolonged and exceeds the 1st extension.

Nr.	Site	System Name	System Details URL	Tier	Integration Start Date	Status
1	CASTORC (Cyprus)	Cy-Tera	https://www.cyi.ac.cy/index.php/castorc/research-information/castorc-completed-projects/cy-tera-cy-tera-high-performance-computing-facility-for-cyprus.html	Tier-1	04/2021	Cancelled
2	CCSAS (Slovakia)	Aurel	http://vs.sav.sk/?lang=en&section=departments&sub=vvt&sub2=config	Tier-1	05/2017	Decommissioned before integration
3	CINECA (Italy)	Galileo	https://wiki.u-gov.it/confluence/display/SCAIUS/UG3.3%3A+GALILEO+UserGuide	Tier-1	01/2019	Decommissioned before integration
4	IT4I/VSB-TUO (Czechia)	Anselm	https://docs.it4i.cz/anselm/hardware-overview/	Tier-1	12/2012	Decommissioned
5	EPCC (UK)	Archer	https://www.archer2.ac.uk/about/	Tier-1	12/2013	Decommissioned
6	UiO/Sigma2 (Norway)	Abel	https://www.uio.no/english/services/it/research/hpc/abel/	Tier-1	04/2013	Decommissioned
7	UHEM (Turkey)	Sariyer	http://wiki.uhem.itu.edu.tr/w/index.php/English	Tier-1	03/2019	Completed
9	UC-LCA (Portugal)	Navigator	https://www.uc.pt/lca/ClusterResources/Navigator/description	Tier-1	10/2018	Completed
10	SURF (NL)	Snellius	https://userinfo.surfsara.nl/systems/snellius	Tier-1	08/2021	Completed
11	HLRS (Germany)	Hawk	https://www.hlrs.de/systems/hpe-apollo-hawk/	Tier-0	01/2020	Completed
12	UL (Slovenia)	HPCFS	http://hpc.fs.uni-lj.si/hardware	Tier-1	07/2017	Partially integrated
13	CESGA (Spain)	FinisTerrae	https://www.cesga.es/en/infraestructuras/computacion/FinisTerrae2	Tier-1	04/2016	Partially integrated
14	CEA (France)	Irene	http://www-hpc.cea.fr/en/complexe/tgcc-Irene.htm	Tier-0	03/2018	Partially integrated
15	IT4I/VSB-TUO (Czechia)	Barbora	https://docs.it4i.cz/barbora/introduction/	Tier-1	05/2017	Partially integrated
16	CSC (Finland)	Puhti	https://research.csc.fi/csc-s-servers	Tier-1	01/2020	Decommissioned
17	CSC (Finland)	Mahti	https://research.csc.fi/-/mahti	Tier-1	10/2019	Partially integrated
18	EPCC (UK)	Archer2	https://www.archer2.ac.uk/about/	Tier-1	06/2021	Partially integrated
19	MACC (PT)	BOB	https://macc.fccn.pt/resources	Tier-1	04/2021	Partially integrated
20	KTH-SNIC/PDC (Sweden)	Beskow	https://www.pdc.kth.se/hpc-services/computing-systems/beskow-1.737436	Tier-1	09/2019	Partially integrated
21	KTH-SNIC/PDC (Sweden)	Dardel	https://www.pdc.kth.se/hpc-services/computing-systems/about-dardel-1.1053338	Tier-1	04/2022	Partially integrated

Table 7: System integration overview

3 Operational Services

Common services were divided into thematic categories: Network, Data, AAA, User, Monitoring and Generic. Each service category has a responsible person who is in charge of managing all the information and decisions related to a specific service area. The selection of common services was published in the PRACE Service Catalogue and once chosen, the responsibility for a service is taken by the respective service area. The following sections provide an update of the status of each service category and the main achievements mostly within the 1st extension.

3.1 Network services

The main task within network services handled in the last 1st extension of PRACE-6IP has been the continuation of the general operation of the PRACE-MDVPN network including integration of new and removal of old Tier-0 and Tier-1 PRACE HPC systems into the network infrastructure. Also, continued user support concerning network problems and optimal network usage have been important activities.

PRACE was asked in autumn 2021 by GÉANT to change the layout of the PRACE-MDVPN to a L3VPN, because of needed changes within the GÉANT infrastructure. So, the change of the infrastructure was planned and communicated to the PRACE partners in November 2021.

Since GÉANT scheduled the date for this transition to mid of April 2022 and because PRACE agreed to stop providing the PRACE-MDVPN functionality at the end of June 2022, it seemed in-efficient to proceed with the transition to the new L3VPN. The change would be in operation only for two and half months. So PRACE asked GÉANT to cancel the transition and extend the PRACE-MDVNP until end of June 2022.Fortunately, GÉANT was able to wait with the changes in their infrastructure until the shutdown of the PRACE-MDVPN, so that no extra work had to be done. The PRACE-MDVPN was shut down by July 1st 2022, both its operations as well as monitoring.

3.2 Data Services

During the 1st extension, the sites which were providing their infrastructure through PRACE calls were offering some of the data transfer tools defined by the PRACE service catalogue. The service leader was checking the state of operation of the data transfer infrastructure through PRACE monitoring tools and network information pages. With secure copy (SCP) being a core service for the data transfer, some sites were also offering GridFTP or UFTP for improved transfer performance. With both tools, there were at least development and maintenance activities to maintain software security and functionality. A new maintenance release of the Grid Community Toolkit (GCT), part of which is GridFTP, was published on May 24th 2022 and is available either as a source code or a precompiled package for the following Linux distributions: Debian, Ubuntu, RedHat, CentOS, Fedora, SLES and OpenSUSE. PRACE contributed to the maintenance of the GCT in the scope of data services activities carried out by HLRS. The service decommission is currently underway and the sites have decided to stop providing services by having the following PRACE events in mind: end of the 1st extension period (June 30th 2022), end of the 2nd extension period (December 31st 2022), or end of the allocation period of the PRACE Call 24 (March 31st 2023).

3.3 AAA Services

The main activity in the 1st and 2nd extension of PRACE-6IP in context of AAA, besides the continuation of maintenance and operation of the PKI, user administration (PRACE LDAP) and accounting services (DART), was the finalization of the new PRACE Federated AAI. After the agreements achieved during the half-day OIDC workshop reported in D6.5, and the implementation of the two drafts collecting technical requirements and policies description of the service, the discussion mainly progressed via e-mail, with one remote meeting per month planned until the end of the first quarter of 2022. The main topics of these meetings were the status of the implementation of the pilot, and the report about technical issues experienced in the integration of the OIDC token in the authentication mechanism. One of the key findings of the technical discussion was the aim to have the PRACE Federated AAI fully interoperable with other major European federated AAI mechanisms and thus usable in the context of existing and developed AAI, like EUDAT, Fenix, and PUHURI. For this reason, the work already done in Fenix RI was leveraged and the pilot was setup towards the existing Fenix infrastructure, consisting of the existing Identity Providers and Attribute server (FURMS).

3.3.1 Technical pilot for federated access mechanism

The pilot we developed shows how user credentials from one federated institution can access resources or services from another federated institution. As an example of such institutions CINECA and CSCS were chosen together with the OpenStack service. A CINECA user is successfully authenticated and authorized with CINECA credentials, to access the CSCS OpenStack service. Some technical issues delayed the first configuration test up to the last weeks of December 2021. Finally, in January 2022 a successful test was run between CINECA and CSCS, following the scheme in Figure 1.

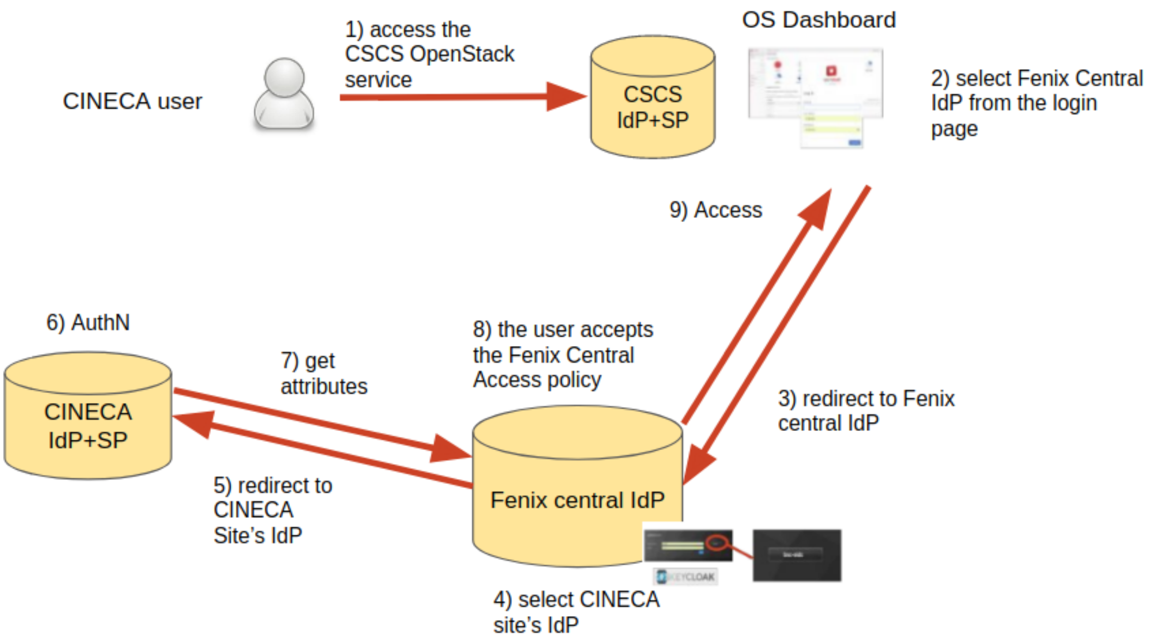

Figure 1: Federated PRACE AAA pilot scheme

Description of the pilot steps:

The CINECA user tries to access the CSCS OS dashboard (1) and is redirected to the Fenix central IdP (2-3) where he/she can select his/her own affiliation (CINECA) on a web page (4); the user is consequently redirected to the CINECA IdP web page (5) where authentication is done by username and password (6); the user is redirected to the Fenix Central IdP (7) where the user accepts the Fenix general Usage Agreement, i.e. by ticking a checkbox (8).

In case the user has not been provided with Fenix Credits on FURMS, or in case there are new updates of local site policies to be approved, the authentication flow stops here and the user is redirected to a predefined web page.

Finally, the user's profile is created in the Fenix Central IdP and redirected to the CSCS dashboard (9). The OpenStack Keystone SP takes care to map the user on the local user account and to apply local policies.

The main issue we experienced was solved after summer 2022, and it was related to the definition of attributes to map the user. This issue affected the roadmap of our activity, and it was agreed, at the end of 1st extension, to cancel the task of adding new sites as there was not much time left till the end of the PRACE 6IP project.

3.3.2 Definition of federated AAI Policies

The activity related to the policies definition produced a draft collecting all analyses and discussions occurred in the previous reporting period. We managed this draft in the first quarter of 2022 to report all agreements achieved by partners, with the aim to describe common rules to authorize users to PRACE services.

To define policies of PRACE Federated AAI, we took the inspiration from the work done in ICEI project by 5 PRACE hosting members, and by AARC Blueprint Architecture. We proceeded under eight principle aspects:

1. **Local site autonomy**: each partner providing a PRACE service can deny the authorization to a PRACE user if the local policies are not fully covered by policies of PRACE AAI. The same partner can ask the unauthorised PRACE user to accept missing policieswhich can be required every time the partner will need it, according to local policies.
2. **Federate local IdPs**: each PRACE partner can take the decision to federate or not the local IdP. All federated IdPs will be considered trustable to authorize the user to a PRACE centralized service (e.g. Peer Review, web-portal, etc.).
3. **New user subscription**: no new PRACE service will be needed to manage the user registration to PRACE. The user email is the criteria to identify the user and allow the mapping procedure in case of multiple local accounts linked to the same user (please, see item 4 below). A user already registered on local sites using the local IdP could become PRACE user after subscription to the PRACE mailing list accepting the PRACE Privacy Notice, the PRACE Data Protection Policy and the PRACE Website Legal Notice & Terms of Use, and after being provided with PRACE resources. PRACE users may also choose to leave the federation and only retain the local identity.

4. **Multiple accounts**: if a user has multiple accounts at different local sites, this user and all his site accounts must be mapped to a unique PRACE identity to avoid proliferation of identities.
5. **PRACE AAI concept**: the PRACE AAI is conceived as the collaboration of three services: a PRACE Central IdP which is responsible to proxy authentication requests among federated IdPs, an Attributes Provider (PRACE Peer Review process) and a Resources Provider (hosting site) managing users' authorization through budget allocation, groups and roles. All the services put in place by the PRACE AAI (PRACE Central IdP and PRACE Peer Review process), should be security audited by an external company before and during production phase.
6. **Federation**: each federated site needs to expose its user-base through its own IdP, supporting standard identification protocols, e.g. SAML v2, OIDC. The site is responsible for the operation of its IdP and free to decide how users should authenticate to it, e.g. username/password, X.509, Two Factor Authentication, RSA, etc. To be part of the federation, each federated IdP should be able to release a common set of user profile attributes in response to the authentication process. These attributes should meet the requirements of the REFEDS Research and Scholarship Entity Category (R&S).
7. **Central IdP**: the PRACE Central IdP needs to be able to proxy authentication requests coming from external IdPs, such as communities OIDC servers or the eduGain federation. When a site or community does not support any external IdP, the PRACE AAI should be capable of authenticating and validating users against specific data sources (e.g. a local site LDAP).
8. **Access to services and resources**: since PRACE delivered various classes of services, the PRACE AAI needs to cope with this heterogeneity and thus support various access interfaces, either based on Internet protocols, such as HTTP(S) (Web Portal, SWIFT, etc.) or SSH. In order to access resources, users need to accept the PRACE access policy and any extensions requested by local sites. For instance, a site may ask users to accept local policy besides the PRACE one. Policies should not include conflicting articles or create ambiguity. The fact that a user has accepted the PRACE general policy will be tracked at the PRACE Central IdP. Access to services and resources is granted through the creation of projects/budgets.

The main achievement we agreed on, was about the REFEDS Assurance Framework:

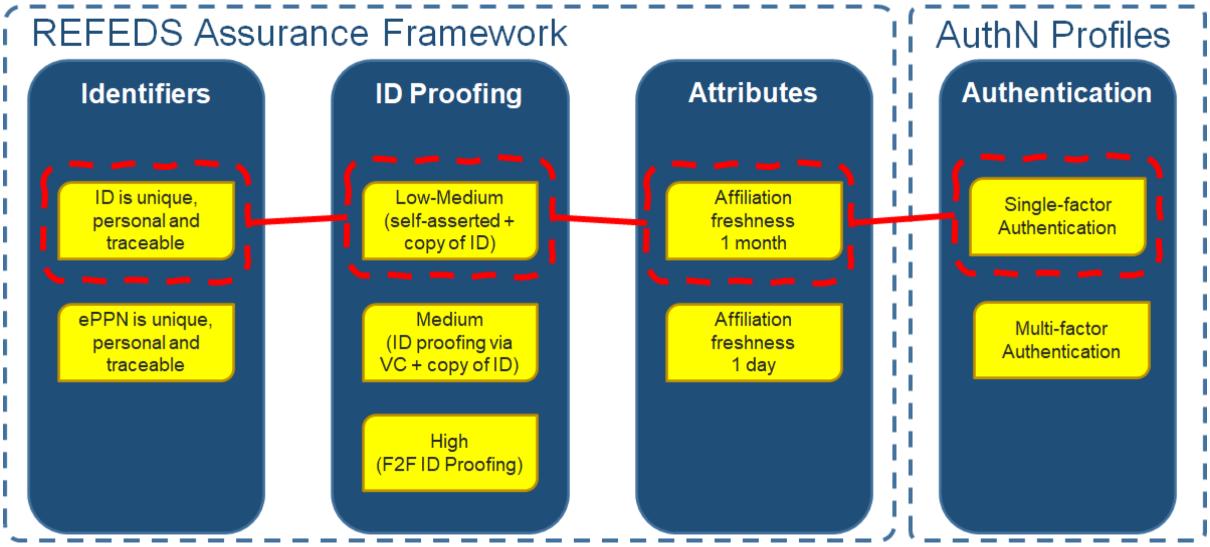

Figure 2: REFEDS Assurance Framework profile for PRACE AAI

The low level assigned to the PRACE AAI will guarantee to all partners a strong autonomy in the registration process of users, without preventing the authorization to a PRACE service.

In the second quarter of 2022, we agreed to merge the documents describing the policy and technical architecture of the PRACE Federated AAI into a single one, named PRACE AAI Strategy. The document was made available to the partners on BSCW and it was agreed to submit it to a review process to publish a white paper describing the whole activity and results achieved.

3.4 Operational Security and Security Forum

The main task of operational security within PRACE-6IP included the user support on IT security issues within the PRACE dedicated network as well as the operation of the PRACE CSIRT itself.

Within the 1st and 2nd extension of PRACE-6IP no major security incidents have been reported, that would have to be worked on by the PRACE CSIRT.

Discussions on standards concerning trust relationships between partners in collaborating infrastructures have been continued. The information exchanges on access methods like two-factor or multi-factor authentication have been intensified. Most partners are still evaluating which solution they want to use in the future.

The individual security procedures at the PRACE partner sites for incident handling and risk analysis are working without any problems.

3.5 Monitoring Services

The monitoring system (ICINGA) collects the health status data for PRACE-RI hosts and services, provides alerting capabilities and collects all the data necessary to estimate service availability and quality for reporting and analytic purposes. Operating and supervising infrastructure that is distributed among several project partners is a very challenging task. The goal of the monitoring service activity is to provide insight into the current and past state of services hosted and delivered by the project partners as well as monitor service interoperability in this complex setup. The task activity is also focused on developing the current monitoring system by addressing issues, deploying functional changes and adding functionalities based on the project requirements.

The solution used for PRACE systems monitoring is based on the Icinga2 software and incorporates several modifications and additions required by the project. The changes include: configuration generators based on the prace_service.config file, integration with the PRACE LDAP user database for authentication and finally a set of executors and checks - main scripts providing data of the host and service status information to the system for project specific services.

The service provides a centralized web interface as shown e.g. in Figure 3.

Until June 2022, the monitoring system kept track of 561 service instances and 62 hosts (end points) provided by 14 sites.

This number is subject to change over time as it heavily relies on the services provided by sites and registered in the prace_service.config file - the primary data source for the monitoring infrastructure.

Due to GSI-SSH service decommissioning the sites had to switch the communication protocol to SSH. Most of the PRACE sites have implemented the change.

GSISSH availability

Modify Schedule Download Send

Hostname	Service Name	SLA in %
bsc-marenostrum:gftp.prace.bsc.es	proxy.refresh	85.52
cea-irene-amd:irene-amd-prace.ccc.cea.fr	proxy.refresh	0
cea-irene:irene-eu-prace.ccc.cea.fr	proxy.refresh	0
cineca-marconi:gssh-prace.marconi.cineca.it	proxy.refresh	98.32
cyfronet-prometheus:prace-int.prometheus.cyfronet.pl	proxy.refresh	89.88
cyfronet:prace-int.cyfronet.pl	proxy.refresh	98.92
grnet:gssh-prace.aris.grnet.gr	proxy.refresh	99.29
ichec:prace-login.kay.ichec.ie	proxy.refresh	40.76
leo:leo-login.sc.niif.hu	proxy.refresh	98.46
phitagoras:phitagoras.sc.niif.hu	proxy.refresh	98.58
psnc-eagle:eagle-prace.man.poznan.pl	proxy.refresh	88.15
uc-lca:gsissh.prace.lca.uc.pt	proxy.refresh	73.46
uhem:pracegw.uhem.itu.edu.tr	proxy.refresh	0
wcss:prace-bem-int.wcss.pl	proxy.refresh	75.97
Total		67.66

Figure 3: PRACE monitoring user interface (availability of GSISSH service)

3.5.1 Task Activities during the 1st extension

The monitoring system is in a stable production phase and most tasks related to the system were of maintenance nature also during this period, however, development of changes required by the project evolution was being carried out and concluded with the deployment on the production instance.

Maintenance and support activities performed regularly within the task include:

- Watching for operating system level vulnerabilities and installing necessary updates and mitigations.
- Monitoring mailing lists and web for icinga2 and web interface updates and security issues.

- Addressing bugs and issues reported via TTS system related to monitoring service,
- Periodically reviewing system and service logs to ensure constant proper operation of the service.
- Providing help and expertise related to monitoring system operation to newly integrating sites and current ones facing operational issues.

Due to the production status of the service, maintenance tasks were performed with keeping in mind service availability on the highest level possible. Apart from the above, one significant change was made to the system during the 1st extension:

- Introduction of SSH based monitoring as a replacement for GSI-SSH based monitoring. This change is due to the service catalogue update. Required changes developed during the previous reporting period have finally been tested, confirmed to work properly and deployed in production. This allowed new site integrations to appear in the monitoring interface (as new systems no longer provide Globus based services).

The monitoring team has also prepared updated and revised documentation on the PRACE wiki, describing the monitoring system and providing the better and clearer integration guide for the new systems especially those without the GSI-SSH service.

The auto ticket creation mechanism has been introduced for selected services. The e-mail service to the responsible person is created when the monitoring detects a malfunction.

3.5.2 Task activities in the 2nd extension

Considering the agreements on shutting down the PRACE services, the monitoring service was maintained in-kind from July until the end of 2022. All systems available to PRACE were monitored as well as the PRACE generic services (address availability and credentials).

3.6 Generic Services

In general, all services that need an operational basis and a centralized distribution for the PRACE project (or a part of it) could be assumed as Generic Services. The goal is the supervision of their operation, as they are crucial for the day-by-day work of the project.

One of the main responsibilities of this task is to keep services up and running. That included handling the *.prace-ri.eu wildcard certificate, helping with new certificate requests and renewal and handling prace-ri.eu domain: subdomain requests, helping with resolving domain issues related to changes.

The decommissioning of existing Generic Services was one of the focus areas throughout the period. This included planning decommission of all Generic Services and gathering data for all WP leaders and service hosts on their ideas on decommissioning as well as negotiating with WP leaders and PMO about the requirements and schedule of decommission of Generic Services.

The services providing information to PRACE end users were chosen to be kept running at least until the end of the projects 2nd extension (December 2022).

A plan for secure and annotated archival of PRACE digital material legacy was created, including the selection of the right technical platform and deciding which content to migrate, supporting automated exports from current services.

PRACE Events, Material and CodeVault portals are among the services which need to be archived. CodeVault (including UEABS) was chosen to be archived into a free gitlab cloud based hosting, while content of the Events and Material Portal were chosen to be migrated into the Zenodo platform. The main PRACE communication points, e.g. the prace-ri.eu website will be operated further (at least during 2023).

PRACE data on EOSC HUB was kept updated with latest content.

4 Lightweight virtualization

Linux containerization is an operating system level virtualization technology that offers lightweight virtualization. An application that runs as a container has its own root file system, but shares the kernel with the host operating system. This has many advantages over virtual machines. First, containers are much less resource consuming since there is no guest OS. Second, a container process is visible on the host operating system, which gives the opportunity to system administrators for monitoring and controlling the behavior of container processes. Linux containers are monitored and managed by a container engine which is responsible for initiating, managing, and allocating containers. Docker is the most popular platform among users and IT centers for Linux containerization. A software tool can be packaged as a Docker image and pushed to the Docker public repository, Docker hub, for sharing. A Docker image can run as a container on any system that has a valid Linux kernel. HPC targeted platforms, e.g. Singularity and Sarus, make it possible to use Docker containers in production for HPC systems.

Unikernels are lightweight single application operating systems developed for the cloud, edge computing, Internet of Things, etc. They fit into small images, have low memory foot-print, and boot in less than one second. They are known for accelerating the execution of programs and improving throughput of network applications. This technology has proven its qualities for these domains with numerous unikernels and publications.

In PRACE-6IP project, this service focused on the evaluation of HPC container runtimes, Sarus and Charliecloud, in addition to the production support of Singularity. Unikernels have been evaluated for HPC workloads (mainly OpenMP). Containerized HPC using elastic composition of HTCondor clusters using Docker swarm has also been tested and evaluated. More information on the previous work done in PRACE-6IP can be found in [2].

4.1 Cooperation with EuroHPC

During the service extension, January – June 2022, the group work focused on the characteristics and challenges for supporting HPC containers on EuroHPC systems. The collaboration included EuroHPC LUMI, VEGA, and LEONARDO. A workshop has been held at the University of Ljubljana on June 30th – July 2nd where support team members from LUMI, VEGA, and LEONARDO have had a fruitful technical discussions on the different approaches, scenarios, and issues supporting HPC containers on the three EuroHPC systems.

LUMI is a pre-exascale EuroHPC supercomputer located in Kajaani, Finland. It is a Cray EX supercomputer supplied by Hewlett Packard Enterprise (HPE) and hosted by CSC – IT Center for Science. Leonardo is a pre-exascale EuroHPC supercomputer currently built in the Bologna Technopole, Italy. It is supplied by ATOS, based on a BullSequana XH2000 supercomputer and hosted by CINECA. VEGA is a petascale EuroHPC supercomputer located in Maribor, Slovenia. It is supplied by Atos, based on the BullSequana XH2000 supercomputer and hosted by IZUM.

Technologies supported by PRACE in this activity on the individual EuroHPC systems

- VEGA: Singularity is the main and default container runtime. There are plans to install NVIDIA enroot, SARUS, and also Redhat Podman for internal services
- LUMI: Only singularity is currently supported due to disabling user namespaces on compute nodes. All other container runtimes require enabling user namespaces to both run and build containers

- LEONARDO: Only singularity is currently supported enroot has been tested and can be enabled in the future.

5 Energy consumption monitoring and reduction

In September 2021, PRACE-6IP WP6 was tasked by the European Commission to produce a plan to monitor and reduce the energy consumption of HPC systems following the "European Green Deal" guidelines. This activity focused on developing a plan to adopt one or more tools already being developed by the hosting members of PRACE in several European projects to enable future wide adoption not only in PRACE-WP6 context, but for any system in the European HPC ecosystem.

5.1 The energy consumption issue

In the first months of the activity, we realized that the issue is now strictly focused on the energy wasted during computation. While vendors started to design processors to maximize performances, and to foster the end-user to control the power states obtained by Dynamic and Voltage Frequency Scaling, clock gating or throttling states, and idle states which switch off unused resources, the scientific and technician communities started to explore three approaches to reduce the energy wasted:

- to trade-off power consumption and performance to gain energy efficiency;
- to improve application performance under a power cap;
- to cut the IT energy waste by reducing the performance of the processing elements when the application is in a phase with communication slack available.

5.2 Analysis of implemented procedures in local sites

Our first action was to share the experience of each partner involved in PRACE-6IP WP6. This action was a follow-up on previous work done in PRACE-6IP. A first result of this analysis was that not all partners applied a monitoring of the energy consumed by users' jobs on their local systems, neither applied automatic or manual procedures to identify and extract metrics useful for monitoring, nor applied automatic or manual procedures to reduce the energy consumed during the computation. A second result was that those few partners who applied or were investigating a monitoring procedure and applied or were investigating a procedure to reduce the energy consumed by users' jobs, reported about a plug-in of the SLURM scheduler and licensed interfaces provided by vendors of HPC systems. A third and last result was that those few partners reporting their experience, remarked on the incoherency of most monitoring results and the inability of applying light procedures to measure the energy consumption of a system and/or of running jobs.

During the analysis we realized that only three partners were strongly involved in EU projects focused on investigating and solving the issue of energy consumption: BSC, CINECA, and LRZ.

5.3 Investigation of procedures and tools at CINECA

After a discussion between the most involved partners from BSC, LRZ, and CINECA it was agreed to leverage the work done by the REGALE project, where especially CINECA had strong presence.

The first meeting with REGALE representatives led to good awareness of the state of art monitoring of energy consumption on the HPC systems. We also identified possible solutions and procedures for the reduction of energy consumed by the HPC workload.

Based on the meeting the choice was done to foster the adoption of automatic procedures which are provided by customized tools responsible for measurement and automatic adjustment of the parameters of the compute elements, mostly CPU and GPU. The parameters monitored and adjusted are mainly frequency, temperature, and power consumed by the cores. The choice for the automatic adjustment rather than manual user orchestration is because such an approach is more effective allowing multiple adjustments through the runtime of the code. Also, the automated adjustment is more user friendly, as the user does not need to be involved in the process of monitoring, evaluation, and adjustment of the parameters.

CINECA installed such automated tools on the PRACE Tier-1 system Galileo100 some years ago to analyze the issue and investigate about a solution, as a task of the REGALE project. In these years, the system hosted the COUNTDOWN run-time library, developed by the University of Bologna and CINECA. In the follow-up meetings with REGALE, we were introduced to results of this tool, that aims mainly on the measurement and automatic reduction of frequency of the computing elements in order to save energy during communication and synchronization primitives. REGALE colleagues emphasized that on average COUNTDOWN reduces the energy consumption of 9.96 percent with an average overhead of 0.79 percent. When compared with the state of the art, COUNTDOWN is capable of achieving similar energy saving but with negligible impact on the performance of a MPI application. More information available at [3], [4], and [5].

5.4 Energy monitoring and reduction conclusions

Although the work done was based on a single tool, we believe that this is a viable and recommendable solution for the wide European HPC landscape. Additional work to reduce the power consumption on HPC systems is to be recommended, especially to target new compute elements like GPGPUs. Following the experiences gathered and the analysis performed, we strongly suggest adopting the COUNTDOWN run-time library. We also recommend fostering collaboration with main HPC system vendors to develop built-in interfaces useful to monitor in depth the power consumed by cores of the different compute elements used.

6 Conclusions

In both 1st and 2nd extension of PRACE-6IP, Task 6.1 has continued a reduced coordinated operation of the PRACE common services for the Tier-0 sites and the Tier-1 sites. The operational procedures have continued to deliver the compute resources to the users of the different PRACE activities including last PRACE calls, SHAPE calls, DECI calls, all training activities and others.

This has been heavily supported by the periodic bi-weekly meetings between the partners using a videoconference as well as by the on-duty activity with a weekly schedule of daily operation monitoring that regularly reported each week the status of the infrastructure, the core services and the issues identified and resolved.

Furthermore, the generic services continued to support the operations and extended the functionalities of the PRACE 'backend' including the PRACE website, Training platform, CodeVault, Events portal and others.

Thanks to the procedures for incident and change management, Task 6.1 operated and monitored on a day-by-day basis the complete set of PRACE common services, as defined in the PRACE Service Catalogue with zero security incidents reported.

Several new Tier-0 and Tier-1 systems were integrated while some older were decommissioned in favour for the new ones which increases the total compute power available for the European HPC users.

The federated approach for the new PRACE AAI was addressed with the aim to become the preferred approach to identities internally used by all PRACE partners and also to be interoperable with other major European computing endeavours with a successful pilot based on the agreed policy and architecture.

The Lightweight virtualization service used existing natural links between the persons in PRACE and the EuroHPC hosting sites and provided support on this topic, including a common face-to-face three-day workshop in Ljublana (Slovenia).

The need to shut down the PRACE services in manageable manner and without the loss of knowledge and data gathered throughout the all PRACE-IP projects a decommission plan for the services together with data archival plan was prepared, communicated to the stakeholders and executed.

Last but not least, PRACE attended to the request to address the energy consumption monitoring and reduction in HPC, following the "European Green Deal" guidelines. Leveraging and cooperating with the already established European project REGALE, PRACE identified a recommendable solution to drive down the energy cost of HPC calculations.

Most of these activities were a conclusion of more than 10 years history of PRACE operations which enabled a unified access and operation of Europe's most powerful systems to the European research, academia, and small business users. It also trained and impacted in a positive way all the administrators from the different partner sites. It enabled them to share and develop efficient and interoperable environments and techniques how to operate such large, complicated and demanding systems while maintaining a high level of security. These experiences will definitively prevail for the upcoming European HPC initiatives which will leverage them and thus keep the legacy of the effort spent.