E-Infrastructures
H2020-EINFRA-2016-2017

EINFRA-11-2016: Support to the next implementation phase of Pan-European High Performance Computing Infrastructure and Services (PRACE)

PRACE-5IP

PRACE Fifth Implementation Phase Project

Grant Agreement Number: EINFRA-730913

D6.2

Final Report on the PRACE Operational Services
Final

Version: 1.0
Author(s): Filip Stanek, IT4Innovations
Date: 12.04.2019

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: EINFRA-730913**	
	Project Title: PRACE Fifth Implementation Phase Project	
	Project Web Site: http://www.prace-project.eu	
	Deliverable ID: D6.2	
	Deliverable Nature: DOC_TYPE: Report	
	Dissemination Level: PU	**Contractual Date of Delivery:** 30 / 04 / 2019
		Actual Date of Delivery: 30 / 04 / 2019
	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 Operational Services**	
	ID: D6.2	
	Version: 1.0	**Status: Final**
	Available at: http://www.prace-project.eu	
	Software Tool: Microsoft Word 2013	
	File(s): D6.2.docx	
Authorship	**Written by:**	Filip Stanek, IT4Innovations
	Contributors:	Fabio Affinito, CINECA Eoin McHugh, ICHEC Ralph Niederbeger, JUELICH Xavier Delaruelle, CEA Martin Bidner, HLRS Cristiano Padrin, CINECA Javier Bartolome, BSC Zoltan Kiss, KIFÜ Lukasz Flis, Cyfronet Miroslaw Kupczyk, PSNC
	Reviewed by:	Manuel Fiolhais, UC-LCA Florian Berberich, JUELICH
	Approved by:	MB/TB

Document Status Sheet

Version	Date	Status	Comments
0.1	27/03/2019	Draft	First draft and structure, User Services, AAI chapters
0.2	29/03/2019	Draft	Integration of new sites, Compute services, Data Services
0.3	01/04/2019	Draft	Generic services, Monitoring
0.3	03/04/2019	Draft	Network services, Operation security
0.4	04/04/2019	Draft	Wording, polishing of the text itself, Summaries, Version for internal review
0.5	08/04/2019	Draft	Revision after internal review
0.6	10/04/2019	Draft	Revision after internal review
1.0	12/04/2019	Final version	

Document Keywords

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

Table of Contents

List of Figures

List of Tables

References and Applicable Documents

[1] http://www.prace-project.eu
[2] Grid Community Forum: https://gridcf.org
[3] RFC 281 - Upgrade to Grid Community Toolkiy 6.2: https://prace-wiki.fz-juelich.de/bin/view/PRACE/Operations/RFC281
[4] GridFTP installation status: https://prace-wiki.fz-juelich.de/bin/view/PRACE/Operations/GridFTP
[5] GitHub repository of gtransfer tool: https://github.com/fr4nk5ch31n3r/gtransfer.

[6] AARC Project: https://aarc-project.eu/
[7] RFC 5280, Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile: https://tools.ietf.org/html/rfc5280
[8] EUGridPMA: https://www.eugridpma.org/
[9] IGTF: http://www.gridpma.org/
[10] GSI-OpenSSH: https://gridcf.org/gct-docs/latest/gsiopenssh/index.html
[11] Grid Community Toolkit:https://gridcf.org/gct-docs/
[12] GSISSH-Term: https://www.lrz.de/services/compute/grid_en/software_en/gsisshterm_en/
[13] DART: http://prace-ri.eu/Accounting-Report-Tool
[14] PRACE Helpdesk: https://tts.prace-ri.eu/
[15] Best Practical Request Tracker: https://www.bestpractical.com/rt-and-rtir
[16] PRACE SVN repository: https://prace.osd.surfsara.nl/svn/
[17] MarkDown: https://en.wikipedia.org/wiki/Markdown
[18] Environment Modules: http://modules.sourceforge.net/
[19] Environment Modules development forge: https://github.com/cea-hpc/modules

List of Acronyms and Abbreviations

AAA	Authorization, Authentication, Accounting
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
ESFRI	European Strategy Forum on Research Infrastructures
EUDAT	European Data Infrastructure
EUGridPMA	European Grid Policy Management Authority
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
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
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.
OS	Operating System
PA	Preparatory Access (to PRACE resources)
PATC	PRACE Advanced Training Centres
PCPE	PRACE Common Production Environment
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
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
PRACE BoD	PRACE Board of Directors
PRACE TB	PRACE 5IP Technical Board (group of Work Package leaders)
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
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 UYBHM)
BSC	Barcelona Supercomputing Center - Centro Nacional de Supercomputacion, Spain
CaSToRC	Computation-based Science and Technology Research Center, 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)
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)

EPCC	EPCC at The University of Edinburgh, UK
ETHZurich (CSCS)	Eidgenössische Technische Hochschule Zürich – CSCS, Switzerland
FIS	FACULTY OF INFORMATION STUDIES, Slovenia (3rd Party to ULFME)
GCS	Gauss Centre for Supercomputing e.V., Germany
GENCI	Grand Equipement National de Calcul Intensiv, France
GRNET	Greek Research and Technology Network, Greece
INRIA	Institut National de Recherche en Informatique et Automatique, France (3rd Party to GENCI)
IST	Instituto Superior Técnico, Portugal (3rd Party to UC-LCA)
IT4Innovations	IT4Innovations National supercomputing centre at VŠB-Technical University of Ostrava, Czech Republic
IUCC	INTER UNIVERSITY COMPUTATION CENTRE, Israel
JUELICH	Forschungszentrum Juelich GmbH, Germany
KIFÜ (NIIFI)	Governmental Information Technology Development Agency, Hungary
KTH	Royal Institute of Technology, Sweden (3rd Party to SNIC)
LiU	Linkoping University, Sweden (3rd Party to SNIC)
NCSA	NATIONAL CENTRE FOR SUPERCOMPUTING APPLICATIONS, Bulgaria
NTNU	The Norwegian University of Science and Technology, Norway (3rd Party to SIGMA)
NUI-Galway	National University of Ireland Galway, Ireland
PRACE	Partnership for Advanced Computing in Europe aisbl, Belgium
PSNC	Poznan Supercomputing and Networking Center, Poland
RISCSW	RISC Software GmbH
RZG	Max Planck Gesellschaft zur Förderung der Wissenschaften e.V., Germany (3 rd Party to GCS)
SIGMA2	UNINETT Sigma2 AS, Norway
SNIC	Swedish National Infrastructure for Computing (within the Swedish Science Council), Sweden
STFC	Science and Technology Facilities Council, UK (3rd Party to EPSRC)
SURFsara	Dutch national high-performance computing and e-Science support center, part of the SURF cooperative, Netherlands
UC-LCA	Universidade de Coimbra, Laboratório de Computação Avançada, Portugal
UCPH	Københavns Universitet, Denmark
UHEM	Istanbul Technical University, Ayazaga Campus, Turkey
UiO	University of Oslo, Norway (3rd Party to SIGMA)
ULFME	UNIVERZA V LJUBLJANI, Slovenia
UmU	Umea University, Sweden (3rd Party to SNIC)
UnivEvora	Universidade de Évora, Portugal (3rd Party to UC-LCA)
UPC	Universitat Politècnica de Catalunya, Spain (3rd Party to BSC)
UPM/CeSViMa	Madrid Supercomputing and Visualization Center, Spain (3rd Party to BSC)
USTUTT-HLRS	Universitaet Stuttgart – HLRS, Germany (3rd Party to GCS)
WCNS	Politechnika Wroclawska, Poland (3rd Party to PNSC)

Executive Summary

The objective of this deliverable is to present the activity done in the reporting period (May 2018 - April 2019) to operate and coordinate the common PRACE Operational services, foreseen by task 6.1 of WP6 in PRACE-5IP project. The operation of the 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, providing services for Tier-0, in a "single" pan-european HPC infrastructure.

This work is the continuation of the work done by task 6.1 in the previous PRACE-IP projects to give continuity to the PRACE Operational services for the HPC Eco-system.

Eight Tier-0 systems were operational in the second year of the PRACE-5IP project period:

- JUWELS at GCS@JUELICH;
- IRENE SKL at GENCI@CEA;
- IRENE KNL at GENCI@CEA;
- HAZELHEN at GCS@HLRS;
- SuperMUC phase2 at GCS@LRZ;
- MARCONI (BDW & KNL) at CINECA;
- MareNostrum4 at BSC;
- PizDaint at ETHZurich@CSCS

Furthermore, operational support has been provided to 23 national Tier-1 systems that provide services for Tier-0 (i.e. used from SMEs for the SHAPE activity, or as stepping stone towards Tier-0 systems, or to prototypes and asses new operational services). These Tier-1 systems are distributed among 14 different countries, ensuring a wide distribution of the European HPC eco-system.

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

The Security Forum, responsible for all security related activities, is also coordinated by task 6.1, with periodic teleconferences to monitor the infrastructure and prevent possible incidents which could cause vulnerability on the PRACE RI. The PRACE CSIRT team has been established and is now in status of listed – a prerequisite to be fully Accredited to GÉANT's Trusted Introducer program. The full accreditation is ongoing.

1 Introduction

This deliverable describes the activities done in task 6.1 ”Operation and coordination of the comprehensive common PRACE operational services” of WP6 “Operational Services for the HPC eco-system” in PRACE-5IP. This task is responsible for the operations of the set of common services, which presents the PRACE Tier-0 systems as an integrated pan-European HPC eco-system [1]. The operational services are extended also to national Tier-1 systems, essential as a stepping-stone towards Tier-0 systems (Tier-1 for Tier-0). Examples of Tier-1 for Tier-0 activities are the SHAPE activity toward SMEs, the prototyping and assessment of new operational services investigated in tasks 6.2 and 6.3 of PRACE-5IP WP6, the testing and utilization of specific architectures and technologies, which are only available in specific countries.

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

This report mainly focuses on the activities done in the second reporting period (May 2018-April 2019) and complements the activity undertaken during the first year of PRACE-5IP.

In this reporting period, the operation of the common PRACE operational services has been coordinated and monitored constantly, by means of biweekly teleconferences, attended by all the partners involved in the task 6.1 activity. Furthermore, a successful WP6 Face-to-Face meeting has been organised, in Santiago de Compostela (Spain) on 6-7 of November 2018. This meeting has been very useful to discuss the status of the operational activity, to plan the activity for the subsequent periods and to agree on the teams involved in the different activities. The meeting had around 30 attendees and most of the different sites involved in WP6 have been represented.

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 eco-system. Section 3 gives a status overview of the common services for the different service areas:

- Network services (MD-VPN based network provided by GÉANT connecting Tier-0 and major Tier-1 centres);
- Data services;
- Compute services;
- Authorization, Authentication and Accounting;
- Operational security;
- User services;
- Monitoring services for operations;
- Generic services.

Finally, Section 4 presents the conclusions.

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

The first part of this section presents the changes implemented during the current reporting period (May 2018 – April 2019) concerning the status of Tier-0 sites, and the Tier-1 national sites providing Tier-1 for Tier-0 services. In the second part, the chronology and the status of the performed system upgrades is being detailed.

2.1 Maintaining the service

In the reporting period, few minor changes have been applied to improve the operational procedures; the main tasks of the operational procedures used to offer the PRACE services has been consolidated. For this reason, the activities concerning the managing of procedures and operability resulted simplified and well tested as was the continuation and evolution of the activity already in progress since the previous PRACE-4IP project.

The operational procedures are well established. Furthermore, efforts have been kept to maintain the documentation up-to-date and feedback has been periodically inquired to the participant sites. All participants can introduce, find and update information of their PRACE HPC systems more easily and in a standard way.

2.2 Status of Tier-0 & Tier-1 sites

The Tier-0 and Tier-1 systems constitute of a HPC eco-system offering high level services to the European computational community. At the end of the reporting period, according to the information recorded by all partners, 8 Tier-0 systems and 23 Tier-1 systems are integrated and in production. All systems are continuously monitored and the operational quality is assured by employing a specific regular activity provided daily by the members of PRACE-5IP WP6 task 6.1. This On-duty Activity is described below.

2.2.1 On-duty Activity

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

A weekly report on the On-duty Activity is produced, where the operator have 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 is constantly updated, using the report template agreed at the beginning of PRACE-5IP.

Starting from 1st of January 2019, a schedule planning has been defined for the 23 PRACE partners involved in the On-duty Activity. Each of them is in charge for monitoring the infrastructure, and reporting the related issues using the Trouble Ticketing tool (TTS). The 23 partners involved in the schedule are reported in Table 1 below.

1 BSC	13 NCSA
2 CASTORC	14 KIFU
3 CEA	15 PDC
4 CINECA	16 PSNC
5 CINES	17 RZG
6 CSC	18 SURFSARA
7 CYFRONET	19 VSB-TUO
8 EPCC	20 UIO
9 JUELICH	21 UHEM
10 HLRS	22 ETH
11 ICHEC	23 MPCDF
12 IDRIS	

Table 1: PRACE partners involved in the on duty activity

A given partner is involved in the shift one week out of 23, i.e. every 5 months. The weekly reports of the On-duty Activity are stored on the PRACE Operation wiki.

The TTS tool used by the On-duty Activity staff is 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 is hosted by CINECA on a specific Virtual Machine, and maintained since its deployment during the PRACE-3IP project.

In the next PRACE-6IP project the list of partners involved in the service will be modified according to the guaranteed availability by each partner. An internal review of the service is started in the last months of the reporting period, and it will be evaluated at beginning of PRACE 6IP.

During the reporting period, the following KPIs have been applied to better supervise the On-duty Activity:

Tickets resolving efficiency	
Description:	Number of tickets resolved by the Service Desk
Calculation:	(R/(R+O))*100
Inputs:	Number of open tickets reported (O) Number of resolved tickets reported (R)
Outputs:	Total number of worked tickets Percentage of resolved tickets respect worked tickets (%)
Time-interval:	Monthly, every first working day (update during the first PRACE Operations meeting of the following month) Annual report
Threshold:	Percentage of resolved tickets respect worked tickets > 90% (Annual threshold)

Tools:	RT
ITIL Category:	Service Operation – Incident Management
'KPI Lead':	Cristiano Padrin (CINECA)
Implementation plan:	Request Tracker (RT) provides all data necessary for computing this KPI. The necessary input data can be extracted using the Query Builder of RT.

Table 2: Tickets resolving efficiency

Average Initial Response	
Description:	Average time taken between the time a user creates a ticket and the time that the Service Desk opens the ticket
Calculation:	for i=1,(O+R) { T(i) = ST(i)-CT(i) } maximum value of vector T [sum of T(i)] / (O+R)
Inputs:	Number of open tickets (O) Number of resolved tickets (R) Vector of Creation times (CT) Vector of Starting times (ST) Vector of Initial Response times (T)
Outputs:	Maximum Initial Response Average Initial Response
Time-interval:	Monthly, every first working day (update during the first PRACE Operations meeting of the following month)
Threshold:	< 1 working day without a motivation - unlimited with motivation
Tools:	RT
ITIL Category:	Service Operation – Incident Management
'KPI Lead':	Cristiano Padrin (CINECA)
Implementation plan:	Request Tracker (RT) provides all data necessary for computing this KPI. The necessary output data can be extracted using the Query Builder of RT and managing the chart.

Table 3: Average initial response

In the last twelve months 45 tickets have been created: 2 of them had been rejected, 2 of them are stalled, 1 is still open and 40 had been resolved. This means that a percentage of resolved tickets respect to the worked tickets equals to 93,3%.

It's important to underline that, in principle, the activity is 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.

2.2.2 Production systems

At present (April 2019), according to the shared information by all partners, the Tier-0 ecosystem is made up of eight systems, distributed in seven sites, operated by six different partners, in the five Hosting Members countries (France, Germany, Italy, Spain and Switzerland) as reported in Table 4.

The peak performance of the Tier-0 systems ranges from more than 2 PFlop/s up to more than 21 PFlop/s for the PizDaint system in Switzerland. Three Tier-0 are accelerated: JUELICH/JUWELS with Nvidia V100, CSCS/PizDaint with Nvidia Tesla P100 and BSC/MareNostrum with Intel Xeon Phi. There is not a dominant vendor: three systems are IBM/Lenovo and other three systems are ATOS/Bull Sequana.

Most of Tier-0 systems are ranked in the Top500 (November 2018) and all ranked systems are in the first 30 positions. Since February 2019 CEA (France) had integrated two new Tier-0 systems: Irene (SkyLake) and Irene (Knights Landing), both with ATOS/Bull Sequana architecture. These systems are not yet ranked in the Top 500.

As far as the Tier-1 ecosystem is concerned, 23 systems are operating as Tier-1 for Tier-0 services. These Tier-1 systems are distributed in 19 different PRACE sites, operated by 16 partners, in 16 different European Countries. Table 5 presents the list of the Tier-1 systems.

The peak performance ranges from very small system partitions (<10 TFlop/s) up to large systems in excess of 2.5 PFlop/s. Nine systems deliver more than 1 PFlop/s.

About half of the Tier-1 systems are accelerated, either with Intel Xeon Phi or Nvidia accelerators. Several different vendors and architectures are present and this is a real advantage for the PRACE HPC infrastructure: HPE, SGI, ATOS/Bull, Cray, IBM and Lenovo are nearly equally represented; BlueGene, iDataPlex/NeXtScale, Bullx are the most popular architectures.

Partner	**Country**	**Tier-0**	**Architecture** - **CPU** - **Accelerator**	**Rpeak (TFlop/s)**
GCS-JUELICH	Germany	JUWELS	ATOS Sequana - Dual Intel Xeon Platinum 8168; 2x24-cores; 2.7GHz - Nvidia V100	9891.1
GCS-LRZ		SuperMUC phase 2	Lenovo NeXtScale nx360M5 WCT - Intel Haswell-EP 14 core; 2.6GHz	3580.0
GCS-HLRS		Hazelhen	Cray XC40 -	7420.0

			Intel Haswell E5 2680v3; 12-core; 2.5GHz	
GENCI-CEA	France	Irene (SKL)	ATOS/Bull Sequana - Intel SkyLake 8168; 24-core; 2.70GHz	6635.5
		Irene (KNL)	ATOS/Bull Sequana - Intel Knights Landing; 68-core; 2.70GHz	2339.6
BSC	Spain	Mare Nostrum4	Lenovo SD530 - SkyLake Intel Xeon Platinum 8160; 2x24-core; 2.10GHz - Intel Phi	11150.0
ETH-CSCS	Switzerland	Piz Daint	Cray XC50 - Intel Xeon E5-2609 v3; 12-core; 2.60GHz - Nvidia P100	15988.0
CINECA	Italy	MARCONI	Lenovo Adam Pass (Intel OmniPath Cluster) Intel Xeon Phi 7250 (Knights Landing) 1x68-core; 1.40 GHz	11000.0
			Lenovo Stark (Intel OmniPath Cluster) Intel Xeon 8160 (SkyLake); 2x24-core; 2.10GHz	9000.0

Table 4: PRACE Tier-0 systems

Partner	Country	Tier-1	Architecture - CPU - Accelerator	Rpeak (TFlop/s)
GCS-MPCDF	Germany	Draco	Lenovo - Intel Haswell Xeon E5-2698; 16-core; 2.3GHz - Nvidia GTX980	1463.2
		Cobra	Intel/Bull/Atos - Intel SkyLake; 20-core; 2.4 GHz - Nvidia V100	9793.5
GENCI-IDRIS	France	Turing	IBM BlueGene/Q - IBM PowerPC A2; 16-core; 1.60GHz	1258.0
EPCC	UK	Archer	Cray XC30 - Intel Ivy Bridge; 12-core; 2.7GHz	2550.5
BSC	Spain	Minotauro	Bull Bullx B505/R421-E4 - Intel Westmere-EP; 6-core; 2.53GHz + Intel Haswell E5-2630 v3; 8-core; 2.60GHz - Nvidia/M2090 and K80	339.54
CSC	Finland	Sisu	Cray XC40 - Intel Haswell; 12-core; 2.60GHz	1688.0
SURFsara	Netherlands	Cartesius	Bull Bullx B720/B710 - Intel Haswell; 12-core; 2.6GHz + Intel Ivy Bridge; 12-core; 2.4GHz + Intel Sandy Bridge; 8-core; 2.7GHz	1349.0

			Bull Bullx B515 - Intel Ivy Bridge 8 core; 2.5GHz - 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
SNIC-KTH	Sweden	Beskow	Cray XC40 - Intel Haswell; 32-core; 2.3GHz	1973.0
PSNC	Poland	Zeus BigMem	HP BL685c G7 AMD - AMD Interlagos; 16-core; 2.3GHz	61.2
		Zeus (GPGPU)	HP SL390s - Intel Westmere-EP; 6-core; 2.45GHz - Nvidia/M2090	136.8
		Prometheus	HP Apollo 8000 - Intel Xeon E5-2680 v3; 12-core; 2.5GHz	2400.0
		Eagle	Intel Cluster - Intel Haswell E5-2697 v3; 14-core; 2.60GHz	1380.0
		Bem	Intel Cluster - Intel Xeon E5-2670 v3; 2x12-core; 2.30GHz	640.0
SIGMA	Norway	Abel	MEGWARE MiriQuid - Intel Sandy Bridge-EP; 8-core; 2.6GHz	178.6
NUI Galway	Ireland	Kay	Intel/Penguin Computing - Intel SkyLake; 2x20-core; 2.4GHz	665.0
CaSToRC	Cyprus	Cy-Tera	IBM IDataPlex -	35.0

			Intel Westmere; 12-core; 2.67GHz - Nvidia M2070	
IT4I-VSB	Czech Republic	Anselm	Bull Bullx B510/B515 - Intel Sandy Bridge-EP; 8-core; 2.4GHz	66.0
		Solomon	SGI ICE-X - Intel Xeon E5-2680v3; 12-core; 2.5GHz - Intel PHI	2000.0
KIFU	Hungary	NIIFI-SC	HP Cluster Platform 4000SL - AMD Magny-Cours; 12-core; 2.2GHz	5.48
		Seged	HP Cluster Platform 4000SL - AMD Magny-Cours; 12-core; 2.2GHz - Nvidia M2070	14.0
		Leo	HP SL250s - Intel Xeon E5-2650 v2; 2.60GHz - Nvidia K20, K 40	254.0
		PHItagoras	HP SL250s - Intel Xeon E5-2680 v2; 2.80GHz - Intel/PHI 7120	27.0
CCSAS	Slovakia	Aurel	IBM; Power 775 - IBM Power7; 8-core; 3.84 GHz	128

Table 5: PRACE Tier-1 systems

2.3 System Upgrades

This section describes the activity related to the system upgrades and integration in the PRACE eco-system, being Tier-0 or Tier-1 systems involved in the Tier-1 for Tier-0 services.

In the current period this activity has been included under WP6 subtask 6.1.2 and is led by Barcelona Supercomputer Centre.

2.3.1 Operational procedures for new systems and system upgrades

Current operational procedures are well documented in the PRACE WIKI and BSCW including:

- Integration of new Tier-0/Tier-1 sites
  - Procedure/Template:
    - WIKI template that new sites must fulfil in order to complete their integration into PRACE infrastructure
  - Information for new Tier-1 sites
    - BSCW presentation about basic concepts that new sites should know before starting their integration into PRACE infrastructure
- 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 for the completion of the upgrade procedure

This documentation ensures that all systems follow equal procedures and are in line with the Service Catalogue requirements.

2.3.2 Overview of System Upgrades in PRACE-5IP

The Tier-0 system Irene from CEA and the Tier-1 system Kay from ICHEC are currently still in upgrading state:

System	Site	Tier	Status	Last Check
Irene	CEA	Tier-0	Upgrading	10/04/2019
Kay	ICHEC	Tier-1	Upgrading	15/01/2019

Table 6: System upgrades

2.3.3 Details of upgraded sites

The full list of upgrades during this last year is reported below:

1. Irene – CEA – Tier-0

   Irene is replacing Curie as Tier-0 system for PRACE and includes two compute partitions, a SKL and a KNL partition with a total peak performance of 9 PF which will be in production in CEA's Very Large Computing Centre (TGCC).

2. Kay – ICHEC – Tier-1

   Kay is replacing Fionn as Tier-1 system for PRACE and includes a cluster of 336 nodes with Intel Xeon Gold 6148 (Skylake) processors with a total peak performance of 665 TF.

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

This section presents the new sites and systems that are going to be integrated in the PRACE infrastructure.

2.4.1 Overview of new Tier-0/Tier-1 sites & systems on PRACE-5IP

In the first two years of PRACE5-IP, eleven new systems have been planned to be included in the PRACE infrastructure and to start providing services to it, as shown in the table below. Four system already finished their integration and seven systems are still in the integration phase.

Site	System Name	System Details URL	Tier	Integration Start Date	Status	Last Check
GRNET (Greece)	ARIS	http://doc.aris.grnet.gr	Tier-1	01/09/2015	**Completed** (12/2017)	08/02/2018
Cyfronet (Poland)	Prometheus	https://kdm.cyfronet.pl/portal/Prometheus	Tier-1	aprox. 09/2016	**Completed**	29/03/2019
CCSAS (Slovakia)	Aurel	http://vs.sav.sk/?lang=en&section=departments&sub=vvt&sub2=config	Tier-1	01/05/2017	In progress	13/08/2018
UL (Slovenia)	HPCFS	http://hpc.fs.uni-lj.si/hardware	Tier-1	aprox. 07/2017	In progress	07/11/2018
CESGA (Spain)	FiniTerrae	https://www.cesga.es/en/infraestructuras/computacion/FinisTerrae2	Tier-1	01/04/2016	In progress	08/05/2018
ICHEC (Ireland)	Kay	https://www.ichec.ie/about/infrastructure/kay	Tier-1	01/06/2018	In progress	15/01/2019
IUCC (Israel)	Moab	https://public-wiki.iucc.ac.il/index.php/PRACE	Tier-1	1/2/2017	**Completed**	11/06/2018
UHEM (Turkey)	Sariyer	http://wiki.uhem.itu.edu.tr/w/index.php/English	Tier-1	n/d	In progress	29/03/2019
CEA (France)	Irene	http://www-hpc.cea.fr/en/complexe/tgcc-Irene.htm	Tier-0	21/03/2018	In progress	21/03/2018
JSC (Germany)	JUWELS	http://www.fz-juelich.de/ias/jsc/EN/Expertise/Supercomputers/JUWELS/JUWELS_news.html	Tier-0	09/2018	**Completed**	28/03/2019
UC-LCA (Portugal)	Navigator	https://www.uc.pt/lca/ClusterResources/Navigator/description	Tier-1	10/2018	In progress	07/11/2018

Table 7: New systems integration

2.4.2 Details of new Tier-0/Tier-1 sites

The full list of the new systems integrated in the first two years of PRACE-5IP time-frame is reported below:

1. **ARIS – GRNET – Tier-1**

This system is composed of four islands of nodes with different architecture, it has completed its integration in February 2017. Its technical characteristics are specified below:

- 532 nodes, 4 islands:
  - 426 thin nodes – 64G RAM
    - Ivy Bridge – Xeon E5-2680v2
    - IBM NeXtScale nx360
  - 44 GPU (2xNvidia Tesla K40)
    - Dell PowerEdge R730
    - Haswell – Xeon E5-2660v3
  - 18 Phi: (2xIntel Xeon Phi 7210P)
    - Dell PowerEdge R760
    - Haswell – Xeon E5-2660v3
  - 44 Fat nodes – 512G RAM
    - Dell PowerEdge R820
    - Ivy Bridge – Xeon E5-4650v2
- InfiniBand FDR 56Gbps
- IBM GPFS 750TB

During PRACE-5IP, two services were missing to be integrated: UNICORE and iperf monitoring. UNICORE was dropped as core service and has not been installed, and iperf was installed in December 2017. The system is fully integrated and operational for PRACE users.

2. **Prometheus – Cryfronet – Tier-1**

Prometheus, a supercomputer with 2.4 PFlop/s of theoretical performance, has been listed at 131st position in the Top500 list (November 2018). Hewlett-Packard has built this top-efficient supercomputer located in Poland, according to requirements and partial design provided by Cyfronet.

Prometheus consists of more than 2,200 servers based on the HP Apollo 8000 platform, combined with super-fast InfiniBand network with 56 Gb/s capacity. Its energy saving and high-performance Intel Haswell latest-generation processors offer more than 53,000 cores. These are accompanied by 279 TB RAM in total, and by two storage file systems of 10 PB total capacity, and 180 GB/s access speed. Prometheus has been also equipped with 144 NVidia Tesla GPGPUs.

The integration process of the Prometheus system has finished and Prometheus is already awarding resources for PRACE.

3. **Aurel - CCSAS – Tier-1 regarding**

Aurel is a supercomputer composed by 2 frames, 16 drawers, 128 physical nodes (130 logical nodes), with a total of 4096 cores, which provides a total of 128 TFlops. Each compute node holds 256 GB of main memory. The cluster is using AIX 7.1 as operating system.

There are still pending tasks regarding the integration of the Aurel system.

4. **HPCFS – UL – Tier-1**

HPCFS is a supercomputer with 768 cores, 3TB of memory and 20TB of disk space. The configuration of the computing part of a supercomputer consists of 64 nodes composed of 16 frames, each of which contains 4 computational nodes.

Most of the tasks are pending to be done for the integration.

5. **FinisTerrae – CESGA – Tier-1**

FinisTerrae is a linux based heterogeneous cluster, with an Infiniband FDR low latency network interconnecting 317 computing nodes based on Intel Xeon Hasswell processors. Together, these nodes are able to provide a computing power of 328 TFLOPS, 44,8 TB of RAM and 1,5 PB of disk capacity. The system includes a high performance parallel storage system able to achieve a speed of more than 20 GB/s.

Although there are still some minor tasks missing, like the integration of GSI-SSH, FinisTerrae already is awarding resources for PRACE through DECI.

6. **Kay - ICHEC – Tier-1**

The integration process of the Kay system of the ICHEC started in June 2018 and currently only the network monitoring service is missing for a complete integration.

7. **Moab - IUCC – Tier-1**

The integration process of the Moab system of the IUCC has finished in June 2018 and the system is fully operational.

8. **Sariyer – UHEM – Tier-1**

Sariyer includes a Cluster of mixed Intel processors with a total of 108 compute nodes and 2964 cores, using a internconnect of FDR network.

The Sariyer system should be fully operational for PRACE users in June 2019.

9. **Irene – CEA – Tier-0**

Irene will be the next Tier-0 cluster for CEA, it will be a 9 PFlop system with two partitions, one with 1656 compute nodes based on Intel SkyLake processors and 666 compute nodes based on Intel KNL processors.

SkyLake partition will be connected using Infiniband EDR network, meanwhile KNL partition will be using new Bull eXascale Interconnect (BXI).

At the moment Irene still is not fully integrated.

10. JUWELS – GSC-JUELICH – Tier-0

The new Tier-0 system JUWELS of the JSC (Jülich Supercomputing Centre) already substituted the old system JUQEEN and is already awarding resources to PRACE.

- 2271 Standard notes, 2x24 cores (Xeon Platinium 8168), 96 GB
- 240 Large memory nodes, 2x24 cores (Xeon Platinium 8168), 192 GB
- 48 Accelerated nodes, 2x20 cores (Xeon Gold 6148), 192 GB, 4 GPUs (Nvidia V100)
- 12 Login nodes, 2x20 cores (Xeon Gold 6148), 768 GB
- 4 Visualization login nodes, 2x20 cores (Xeon Gold 6148), 768 GB, 1 GPU (Nvidia Pascal P100)

11. Navigator – UC-LCA – Tier-1

Navigator will be the new Tier-1 system of the LCA (Laboratory for Advanced Computing) of the UC (University of Coimbra):

- 164 Computing nodes, 2x12 cores (Xeon E5-2630v2), 96 GB
- 1 Login node, 2x6 cores (Xeon E5-2630v2), 64 GB
- 1 Lustre MDS, 6 cores (Xeon E5-2630v2), 128 GB
- 5 Lustre OSS, 6 cores (Xeon E5-2630v2), 32 GB

UC-LCA is currently focusing on connecting the system to PRACE MD-VPN as the first step of the PRACE integration. Delay of the integration into PRACE was caused by the transportation into a new datacentre in Portugal.

3 Operational Services

Common services are divided into thematic categories: Network, Data, Compute, 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.

Selection of common services is 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 within the current reporting period.

3.1 Network services

The main task within network services handled in the second reporting period of PRACE-5IP again has been the general operation of the PRACE-VPN network including integration of new and removal of old Tier-0 and Tier-1 PRACE HPC systems into the network infrastructure. Furthermore, user support concerning optimal network usage has been a main task.

During this second reporting period, the remaining PRACE partners needing connectivity to the PRACE-MDVPN-network have been integrated.

PRACE Site	System	HPC site or test IP address
	PRACE-Mon (Germany)	monctl.net.prace.fz-juelich.de. 134.94.115.218
	NONE.cy (Cyprus)	{CastoRC-NONE.cy} 82.116.198.225
IT4I	Anselm (Czech Republic)	gridftp-prace.anselm.it4i.cz. 195.113.250.164
IT4I	Salomon (Czech Republic)	gridftp-prace.salomon.it4i.cz. 195.113.250.173
csc	Sisu (Finland)	gridftp-prace1.csc.fi. 128.214.250.34
cea	Curie (France)	curie-prace.ccc.cea.fr
cnrs	Turing (France)	zahir135-dgiga0.idris.fr. 130.84.240.135
	JuDACsrv (Germany)	judac05p.zam.kfa-juelich.de. 134.94.115.217
lrz	SuperMuc (Germany)	supermuc-fat-prace.lrz.de. 195.37.7.9 195.37.7.8
MPCDF	Pracempcdf (Germany)	mon-prace.mpcdf.mpg.de. 130.183.230.67
grnet	Aris (Greece)	gssh-prace.aris.grnet.gr 195.251.114.114 195.251.114.115
	Seged (Hungary)	loginnode-vlan907.szeged.hpc.niif.hu. 193.224.66.197
	Leo (Hungary)	login-vlan907.debrecen2.hpc.niif.hu. 193.224.66.200
	PHItagoras (Hungary)	login-vlan907.budapest2.hpc.niif.hu. 193.224.66.196
ICHEC	Fionn (Irland)	{ICHECK-Fionn} 193.1.201.17
CINECA	Marconi (Italy)	gssh-prace.marconi.cineca.it. 130.186.26.9 130.186.26.23 130.186.26.8 130.186.26.7
PSNC	Eagle (Poland)	eagle-prace.man.poznan.pl 150.254.128.4
CYFRONET	Zeus (Poland)	prace-int.cyfronet.pl 150.254.128.65
	Supernova (Poland)	prace-bem-int.wcss.pl. 150.254.128.34
	Aurel (Slovakia)	{CCSAS-Aurel} 147.213.81.5
	LECAD (Slovenia)	{LECAD-LECAD} 193.2.78.225
BSC	Marenostrum (Spain)	gftp.prace.bsc.es 212.128.224.7
BSC	Minotauro (Spain)	gftp.prace.bsc.es 212.128.224.7
	Piz_Daint (Switzerland)	{CSCS-Piz_Daint} 148.187.128.41
SURF	Cartesius (The Netherlands)	int1-prace.cartesius.surfsara.nl. 145.100.18.14
uhem	PRACEGW (Turkey)	pracegw.uhem.itu.edu.tr 160.75.120.180

Figure 1: Overview of the PRACE Network by sites connected via MD-VPN (04.04.2019)

The Figure 1 describes the network monitoring status of the connected systems as of April, 4th 2019.

EPCC (UK) is still discussing with JISC, the UK NREN, about possible technical integration into the PRACE MD-VPN. The HPC systems of EA-ECNIS (Bulgaria), HLRS (Germany), UIO (Norway) and KTH (Sweden) have and will presumably not been integrated into the PRACE dedicated virtual network infrastructure. Reasons for this are mostly technical (not needed or not possible).

The discussion which started in the first reporting period, led to the conclusion to set up a test environment for a monitoring solution via perfSONAR. A detailed description of this activity can be found in PRACE deliverable D6.3 chapter "Prototypal activities (perfSONAR network performance monitoring) ", since it is a joined activity/collaboration with GÉANT. It is mentioned here also, since it interacts deeply with the operation services work described about in this deliverable. The participating partners for the test installation have been BSC, CINECA, CSCS, IT4I, KIFU and JUELICH.

In the final state the monitoring solution will include perfSONAR client systems at each PRACE partner site, which will get their configuration from a central MadDash configuration server. The client will do iperf performance tests and store the results in a central Esmond database. A monitoring server (Apache) will allow users to see the results via a web interface.

All client systems, at the testing sites mentioned above, as well as the Esmond database server, the MadDash configuration server and monitoring server have been installed. Work on setting up the software environment is ongoing and will be worked on further in PRACE-6IP.

3.2 Data Services

The second year of the PRACE-5IP project started with the end-of-support of the Globus toolkit, an open source implementation of the gridFTP protocol. As a result, Grid Community Forum [2] was founded with the goal of establishing further support of the discontinued toolkit. Grid Community Toolkit (GCT) is an open source for of the discontinued toolkit and is available either as a source code or precompiled package for the following Linux distributions: Debian, Ubuntu, Red Hat Enterprise Linux (RHEL) compatible operating systems, Fedora, SLES and OpenSUSE. As a result, all PRACE sites are being advised to upgrade their installations through the request for comments #281 issued on December 13th 2018 [2]. Status can also be seen in Table 8.

Site	Status	Comment
BSC (es)		
CASTORC (cy)	Not applicable	System decommissioned
CEA (fr)		
CINECA (it)	Done	EPEL
CINES (fr)		
CSC (fi)		
CSCS (ch)		
CYFRONET (pl)		
EPCC (uk)		
JUELICH (de)	Done	EPEL
HLRS (de)	Done	
ICHEC (ie)	Done	2019/01/14. EPEL.
IDRIS (fr)		
IUCC (il)		
LRZ (de)	Done	18/12/2018
NCSA (bg)		

KIFU (hu)		
PDC (se)		
PSNC (pl)	Done	Changed from Globus-Toolkit-6-Stable to EPEL
RZG (de)		
SURFSARA (nl)	Done	3/12/2018, using EPEL repository
UHEM (tr)		
UIO (no)		
UL (si)		
VSB-TUO (cz)	Done	10/01/2019, EPEL, Anselm & Salomon
WCSS (pl)		

Table 8: Status of RFC #281

GridFTP installation status available at [4] provide overview of the gridFTP systems across Tier-0 and Tier-1 sites. Most installations are configured with separated front-end and back-end servers. TCP port range for data transfers is set between 20000 and 25000. Gtransfer tool that makes gridFTP transfers easier for the user is supported and available for download [5]. PRACE is active in the development and support of the GCF mostly thorough the Data Services activity of WP6 carried out by HLRS.

3.3 Compute Services

Compute services provide interfaces between users and computing capabilities. The target of this activity is to find out what services can be commonly provided in PRACE and try to maintain the uniformity among all sites providing computational hours.

In the PRACE-5IP time-frame the work consisted on maintaining the different components that are provided in the service catalogue, mainly the batch scheduling systems and UNICORE. An update of the activity done in the reporting period for these services is provided in the following.

One of the major changes compared to previous reporting periods has been the movement of the UNICORE service as an optional service inside the PRACE Service Catalogue.

3.3.1 Local batch scheduling

Local batch scheduling systems identify the software tools responsible for managing user jobs. The service consists on maintaining a list of different software components that current supercomputers deploy now, thus providing a statistic of which are the most used ones. In the next table it is possible to see the relationship between supercomputer and batch system. One of the most interesting systems is SLURM, that is step by step overtaking others. SLURM is an Open Source software maintained by SchedMD and one of its strong points is its modularity and flexibility, making relatively easy to adapt the source code to the site needs. These characteristics ensures a rapid growth and adaption to new systems, platforms and architectures and we think this is the reason that many sites are using it. The code of this software is open so it is possible to see that it has a very high quality. Sites are getting support from the community but also directly from the developers and, moreover, there is the possibility of having a contract with SchedMD for enterprise

support. All these features and other ones make SLURM the preferred solution among different sites. Another responsibility of Compute Services in PRACE-5IP has been the maintenance of a set of scripts to wrap the different inputs that different systems could take as parameters, motivated by the differences in interaction with batch systems for different products. A scripted file must be edited with a full description of a job, e.g. resources like core/cpu-hours, number of cores/cpu, I/O staging, service level, and all needed actions for preparing the execution environment.

Some IBM systems have migrated their batch scheduling system from Loadleveler to IBM LSF, as development was stopped some years ago, and now only one system keeps LoadLeveler as batch scheduling system.

These differences made it interesting to develop a set of wrappers with a common syntax for describing a job around the different batch systems and platforms. Such wrappers should make it easy for users to migrate their work from one system to another. Nowadays, given the complexity of systems and architectures, with different features, the work should be enhanced and reviewed during actual project to better enable the wrappers and adapt to different types of systems.

Partner	Site	Machine Name	Tier	Batch System
1. GCS (de)	JUELICH	JUWELS	Tier-0	SLURM
2. GCS (de)	LRZ	SuperMUC	Tier-0	IBM LSF
2. GCS (de)	LRZ	SuperMUC phase 2	Tier-0	IBM LSF
2. GCS (de)	HLRS	Hazelhen	Tier-0	Torque/Moab
2. GCS (de)	RZG	Hydra	Tier-1	IBM LSF
3. GENCI (fr)	CEA	Curie (Fat)	Tier-0	SLURM
3. GENCI (fr)	CEA	Curie (Hybrid)	Tier-0	SLURM
3. GENCI (fr)	CEA	Curie (Thin)	Tier-0	SLURM
3. GENCI (fr)	IDRIS	Turing	Tier-1	LoadLeveler
4. EPCC (uk)	EPCC	ARCHER	Tier-1	PBSPro
5. BSC (es)	BSC	Marenostrum4	Tier-0	SLURM
5. BSC (es)	BSC	Minotauro	Tier-1	SLURM
6. CSC (fi)	CSC	Sisu	Tier-1	SLURM
7. ETH (ch)	CSCS	Piz-Daint	Tier-0	SLURM
8. SURFSARA (nl)	SURFsara	Cartesius	Tier-1	SLURM
10. SNIC (se)	PDC	Beskow	Tier-1	SLURM
11. CINECA (it)	CINECA	Marconi	Tier-0	SLURM
11. CINECA (it)	CINECA	Galileo	Tier-1	PBS Pro
12. PSNC (pl)	CYFRONET	Zeus (BigMem)	Tier-1	SLURM
12. PSNC (pl)	CYFRONET	Zeus (GPGPU)	Tier-1	SLURM
12. PSNC (pl)	PSNC	Eagle	Tier-1	SLURM
12. PSNC (pl)	WCSS	Bem	Tier-1	PBS Pro
13. Sigma (no)	UIO	Abel	Tier-1	SLURM
16. NUI (ie)	ICHEC	Fionn	Tier-1	Torque/Moab
18. CaSToRC (cy)	CaSToRC	CY-TERA	Tier-1	SLURM
19. NCSA (bg)	NCSA	EA-ECNIS	Tier-1	IBM LSF
20. IT4I (cz)	IT4I/VSB-TUO	Anselm	Tier-1	PBS Pro
20. IT4I (cz)	IT4I/VSB-TUO	Salomon	Tier-1	PBS Pro

22. KIFU (hu)	KIFU	NIIFI-SC	Tier-1	SLURM
22. KIFU (hu)	KIFU	Seged	Tier-1	SLURM
22. KIFU (hu)	KIFU	Leo	Tier-1	SLURM
22. KIFU (hu)	KIFU	Phitagoras	Tier-1	SLURM
24. GRNET(gr)	GRNET	Aris	Tier-1	SLURM
25. CCSAS (sk)	CCSAS	Aurel	Tier-1	IBM LSF

Table 9: Local batch schedulers

3.3.2 UNICORE

Another way to create and operate on a common layer is through an even higher abstraction level. The UNICORE software, which is part of the PRACE software portfolio since the beginning, and which was originally adopted in the DEISA projects, accomplishes this. It allowed a user to manage single jobs as well as a workflow of jobs remotely through a Java-based graphical interface (URC) and/or a command line client (UCC). Access to end-systems relies on account information stored in the PRACE central LDAP, while resources are published on a central directory service by sites or resource providers.

UNICORE is currently listed as an optional service for the Tier-0 and Tier-1 systems in the PRACE Service Catalogue.

This is the actual installation of UNICORE components that is needed in the PRACE infrastructure to provide this service:

Component	Description	Deployment
REGISTRY	Directory service publishing Tier-0 resources.	JUELICH (Primary), CINECA (Backup)
UNICORE/X	Translate abstract jobs into concrete jobs for a specific target system. Job submission and monitoring	Tier-0 System
GATEWAY	Main entrance to each Tier-0 system. Client connections go through the gateway, which forwards them to internal components, and vice versa.	Tier-0 System
XUUDB	User Database for authentication and authorization. It is synchronized with the PRACE LDAP.	Tier-0 System
TSI	Interface with the local batch system and storage resources	Tier-0 System
URC	Graphical user client (based on Eclipse Java)	Client-side
UCC	Command-line user client	Client-side

Table 10: UNICORE setup overview

3.4 AAA Services

The AAA activity is responsible for services, which provide Authentication, Authorization and Accounting facilities on the infrastructure. These services play a fundamental role in the operation of other PRACE services and the provision of access to PRACE resources. As a result, changes made to these services are normally of a routine nature and follow pre-documented policy to ensure continued stability and reliability of AAA services.

While the operational AAA services continue to be managed in this stable fashion, PRACE also needs to look forward to the future, particularly where collaborations with other external projects and e-infrastructures may be involved. PRACE and GÉANT have previously discussed the feasibility of a more general federated authentication and authorization infrastructure. This idea was further pursued with the AARC Project [6] during PRACE 5IP to investigate how the AARC Blueprint Architecture might be applied to provide future PRACE AAA services. PRACE AAA services have a representative on the AARC Engagement Group for Infrastructures (AEGIS) to provide direct consultation into their project activities. While the second phase of the AARC project ends in April 2019, PRACE and GÉANT continue to collaborate directly to produce an implementation of the AARC Blueprint Architecture.

3.4.1 Public Key Infrastructure - PKI

Several PRACE services rely on X.509 certificates [7] for authentication and authorization. These certificates must be issued by entities trusted by the service providers. PRACE relies on the Certificate Authorities (CA) accredited as a member by the EUGridPMA, the European Policy Management Authority (PMA) [8], or by one of the two other organizations, the TAGPMA and the APGridPMA, all three federated in the IGTF [9]. These PMAs require a minimum set of requirements for the CP/CPS (Certificate Policy/Certification Practice Statement) of the member CAs, as published and maintained by the PMAs in profile documents. PRACE itself is a member of the EUGridPMA as Relying Partner and participates in its activities.

For PRACE partners CA information is maintained at a central repository. The information is based on the IGTF distribution and updates are made available to the PRACE partners for download in several formats. To guarantee a high level of availability, the distributions are stored on a virtualized webserver that can migrate between two physical hosts at SURFsara. New IGTF distributions are, and have been, provided on an almost monthly basis.

3.4.2 User Administration

Information about users and their accounts is maintained in a Lightweight Directory Access Protocol (LDAP) based repository. This facility is used to share among PRACE partners the authorization information needed by PRACE services and is used to retrieve information about users and their projects. Authorization information is provided among others for data transfers with GridFTP, interactive access through GSI-SSH, job submission with UNICORE, accounting services and access to the helpdesk facilities.

A single LDAP server is used for PRACE Tier-0 accounts. For almost all Tier-1 accounts the same server is used, however some partners host a local LDAP server for the registration of Tier-1 users. At present only IDRIS chose to make use of this external hosting option. Separate databases are used for Tier-0 and Tier-1 accounts. An overview of the LDAP namespace is shown in Figure 2. For Tier-0 accounts the “ou=ua,dc=prace-ri,dc=eu” part of the name space is used and for Tier-1 accounts “ou=ua,dc=deci,dc=org” is used as top part. The Tier-1 accounts registered by other LDAP servers can be accessed through referrals.

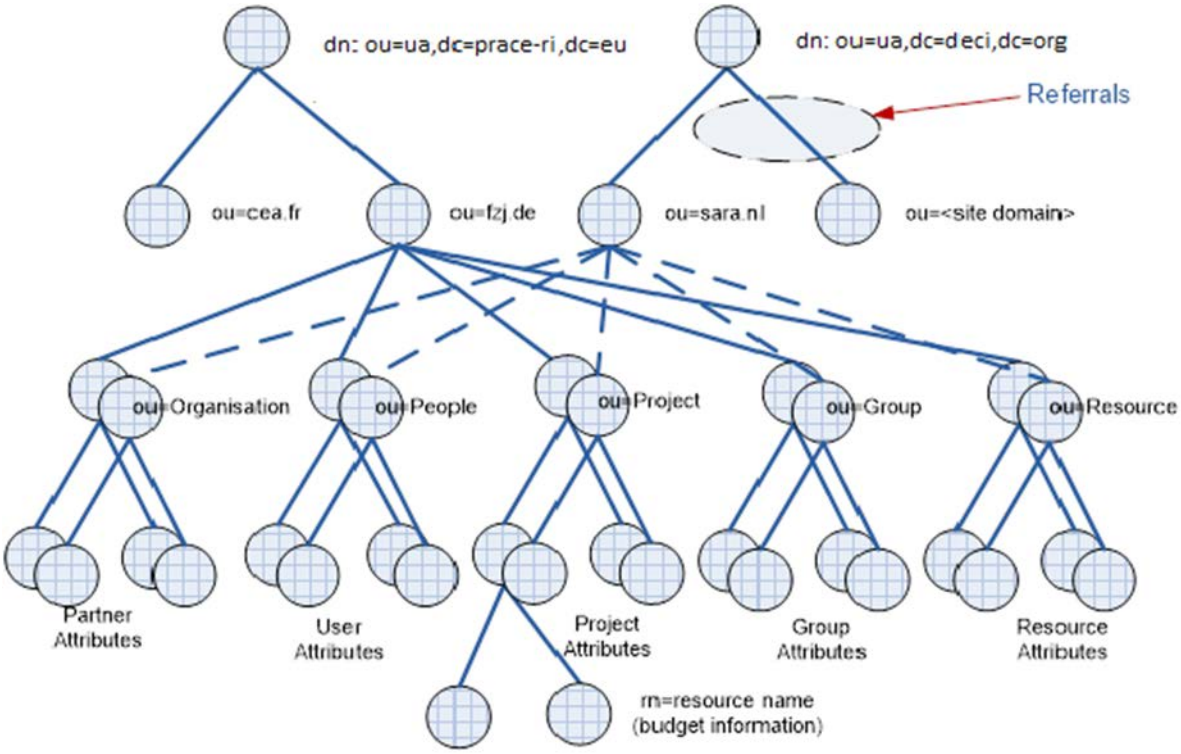

Figure 2: PRACE LDAP directory tree

All Tier-0 sites manage their own branch in the LDAP Directory Information Tree (DIT). They register the users who have access to their site in their LDAP branch, which starts at ou=<site domain>.

Each Tier-1 site also manages its own branch in the LDAP DIT, which is everything below the ou=<site domain> attribute. Each user of Tier-1 resources is assigned to a Home site, which is one of the Tier-1 partners and it's the Home site which creates and updates the account information of the user. The Home site in general is the site from the country where the user is working and not the site where Tier-1 resources have been granted to the user (the Exec site). The Home site mostly already knows the user and this makes the registration procedure easier. The Exec site retrieves the information from the LDAP server for the creation of a local account.

The main PRACE LDAP server and a replica server are operated by SURFsara. Both servers are virtualized and can be migrated between redundant physical servers. To avoid dependency on a single site for this core service, an additional remote replica server is operated by HLRS. The replica servers can be used if the primary server is unavailable for some reason. It is the distributed responsibility of the collective of Home sites to enter the correct information pertaining to their respective users in the central LDAP. It is the responsibility of each Exec site to retrieve the relevant information from the central LDAP and disperse it into their local user administration.

There is no specific prescription how to do this, merely that it must be done and that the delay should not be more than 24 hours. The Exec site may choose to fully automate this process in a way that fits its local user administration, or it may care of it entirely by a manual operation. This abstention from any specific definition of a required interoperability interface between the PRACE LDAP and a site-specific user administration in practice had made it easy to integrate new sites.

Like Exec sites, door nodes for interactive access, explained below, must process PRACE LDAP mutations in their local user administration. Unlike Exec sites, due to their function they cannot be selective and must always keep track of the complete active user community.

Occasionally, when a new partner site joins PRACE, a new branch to the LDAP is added. During PRACE 5IP two new Tier-1 LDAP branches have been added for IUCC(IL) and UL(SI).

No changes have been made to the LDAP schema.

A minor policy change has been made to allow re-allocation of a reserved but unused uidNumber / gidNumber range for PRACE Preparatory Access Type-D projects.

3.4.3 Interactive Access

Interactive access to the Tier-0 systems is provided either by the SSH (Secure Shell) facilities, provided by most distributions of operating systems, or X.509 based SSH facilities. The latter are mainly used for access from other PRACE sites using the dedicated PRACE network. On PRACE Tier-0 and Tier-1 sites the recommended X.509 based SSH facility is GSI-OpenSSH [10] (GSI-SSH for short). Historically, GSI-OpenSSH was distributed and supported by the Globus community but the Globus Toolkit project reached end-of-life in December 2018. Globus Toolkit has been replaced by the Grid Community Toolkit (GCT) [11] supported by the Grid Community Forum (GridCF) [2] with active participation by PRACE partners in the project. All sites that support GSI-SSH are required to use Grid Community Toolkit 6.2 with an up-to-date OpenSSH version from either their Linux distribution or source code.

For access from their workplace users can use GSISSH-Term, a GSI-SSH Java based client, which is supported by the PRACE partner LRZ [11].

Not all sites support GSI-SSH based access directly from the Internet. Therefore, two partners, IT4I (CZ) and SURFsara (NL), provide GSI-SSH based access for all PRACE users who can use these sites as door nodes to other PRACE sites using the PRACE network. Door node services were previously provided by CINECA but IT4I recently took over this responsibility from them.

3.4.4 Accounting Services

Users can get accounting information locally at the sites at which they consume resources. Each site runs a standard PRACE accounting web service that can be accessed easily using GSI authentication. In addition, PRACE users and staff can get accounting information from several sites in a uniform way with the DART tool [13]. With this tool users can display their accounting information for Tier-0 systems and Tier-1 sites.

3.5 Operational Security and Security Forum

The main task of operational security within PRACE-5IP includes the user support on IT security issues within the PRACE dedicated network as well as the operation of the PRACE CSIRT itself.

No major intrusions have been monitored. No services have been on risk.

One further task is the discussion on operational security within PRACE for potential future PRACE services. Here the PRACE Security Forum had several video conferences, discussing on these issues. This resulted into creation of defined mandatory requirements for security assessment

of a PRACE service, which were described in a policy document, together with a questionnaire and evaluation form to carry out the assessments. For four pilot services the input using the questionnaire was gathered and currently those are evaluated by a pair (per service) of PRACE Security Forum members to specify the security level of those services as well as to evaluate and further evolve the process of evaluation itself.

One major point is the Accreditation of the PRACE CSIRT team to GÉANT's Trusted Introducer program. The PRACE CSIRT team has become a listed team in the reporting period, which is a required prerequisite. Current activity is focussed on preparing for becoming accredited. Though assumed in period 1 of PRACE-5IP, accreditation could not be finished, because of the complexity of harmonizing security policies and procedures of the PRACE partners. This is more an administrative and political issue, than a technical one. The individual security procedures at the PRACE partner sites for incident handling and risk analysis are working without any problems. The harmonized security policy will be a very welcomed add on for the security of the PRACE infrastructure and a prerequisite for accreditation, but does not introduce real production status risks.

Furthermore, contributions have been made to the WISE group (Wise Information Security in collaborating Environments), a community of security experts working on security issues in multi organizational infrastructures. PRACE members are actively working together with staff members of other e-infrastructures in close collaboration. A Baseline AUP has been developed, in a final draft version currently, which will be published in the near future. It is envisioned that this AUP will be adopted by PRACE in the near future, too.

3.6 User Services

User services subtask is composed of the following four activities: the PRACE Common Production Environment (PCPE), the Trouble Ticket System (TTS), Helper scripts and PRACE User Documentation.

3.6.1 PRACE Common Production Environment

The PRACE Common Production Environment (PCPE) presents a common application development environment interface across all PRACE execution sites regardless of underlying architecture. It guarantees availability of a minimum set of software components to support the running of highly-scalable parallel simulation software. Thus it allows automated monitoring of key software components across all PRACE execution sites.

PCPE allows users to gain many benefits, such as improved efficiency in porting and running applications due to commonality of interface across PRACE; confidence that the required software for their research is in place at the start of their project; automated access to optimization best practice for local compilers, irrespective of vendor and hardware architecture.

Following the release of the updated specification of PCPE last year (v1.1), no update of the specification was achieved this year as no new need arise.

Like last year, a focus has also been put on the "module" tool command which is at the root of the PCPE to enable or disable parts of this software collection. The module command has different implementations and at the time of the DEISA project the implementation called Modules-Tcl was

the one chosen as the default to handle the DEISA Common Production Environment (DCPE). This Modules-Tcl implementation comes from the Environment Modules project [18], which also delivers a Modules implementation written in C language and widely available in Linux distributions. In 2017, CEA has taken over the leadership of the overall Environment Modules open source project. Work has been achieved to publish a combined release of "Modules" and "Modules-Tcl", making the later the default implementation provided by the project [19]. Thus clarifying choices available when deploying a module command at sites. This new major release of Modules is progressively adopted by Linux distributions. Some automated capabilities to handle dependencies between modulefiles have been added to Modules this year. This new feature could help sites to provide complex yet easy to use environments to their users.

3.6.2 Trouble Ticket System

The centralised Helpdesk [14] was deployed as part of the PRACE-1IP project. It's an important tool for the PRACE project staff to communicate among sites about problems. The Helpdesk is also available for users. However Tier-0 users are advised to contact the helpdesk of the Tier-0 sites directly.

TTS is based on the Request Tracker (RT) issue tracking system [15]. Since its deployment in PRACE-1IP, this service is hosted and operated at CINECA.

No relevant change was made on the PRACE TTS as no new need arise during this second year.

3.6.3 Helper scripts

User services subtask provides helper scripts for PRACE users and staff to easily access the PRACE services in operations. Among these scripts, the "prace_service" utility gives access to the directory of services helping users and staff to get information on server addresses and ports for the various services.

This year the prace_service configuration has been updated many times to include the new PRACE Tier-0 and Tier-1 systems or to remove the systems that have been decommissioned.

3.6.4 PRACE User Documentation

The User documentation for PRACE is available online on the PRACE-RI website [1]. Source document are held in the PRACE SVN repository [16] in the MarkDown [17] format. During the PRACE-4IP project a publication process was established to define how documentation changes made in the PRACE SVN repository are validated and uploaded on the PRACE-RI website.

This second year of activity some updates of documents happened mainly to add references to the new PRACE Tier-0 and Tier-1 systems and to remove descriptions of systems that have been decommissioned. In addition, the new PRACE repository service that has been put in operations this year has been included to the user documentation set.

3.7 Monitoring Services

Monitoring services delivered by task 6.1 are important part of PRACE infrastructure allowing insight into current state of services provided by Tier 0 and Tier 1 sites. The activity is being

realized by periodic execution of checks - scripts and programs written to verify proper functioning of services which are defined by sites in registry file called prace_service.config. Results of such executions are then aggregated, processed and presented via monitoring web interface available at mon.prace-ri.eu, see Figure 3.

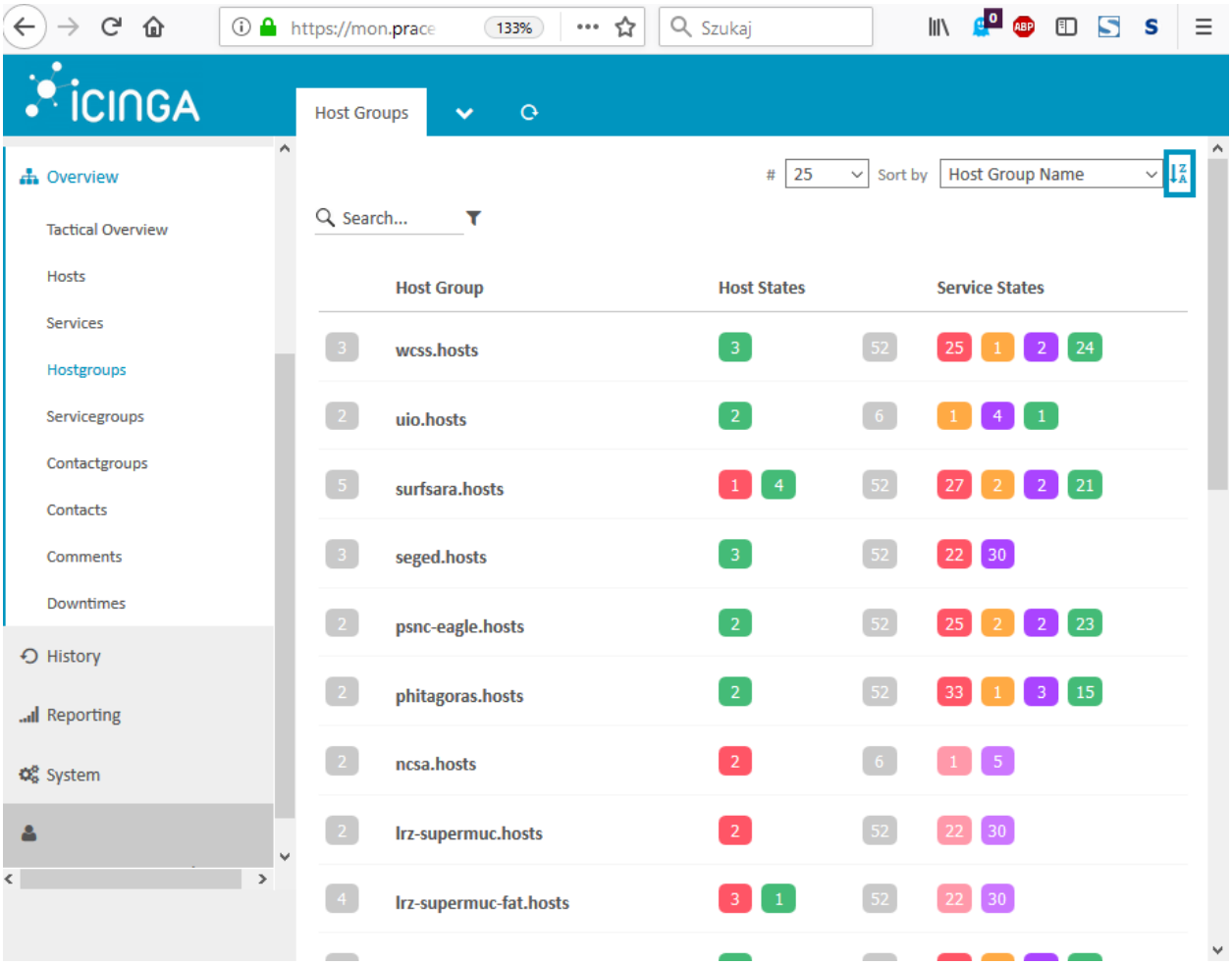

Figure 3: Monitoring overview

Based on the outcome of the tests various notifications can be send as well as they can be used for triggering of automated processes if desired.

The resource monitoring process is built on top of Icinga 2 - the 3rd party monitoring software - with addition of custom check scripts with configuration generators developed within the 6.1 task and which are constantly added and developed according to the project needs.

During the current reporting period the monitoring activity in task 6.1 was mainly focused on maintenance activities which include the following:

- providing the constant service access with minimal downtime by assuring stable infrastructure and up-to-date software
- assuring service security: reacting to security threats, addressing CVEs released by cyber security organizations by updating vulnerable software components

- configuration adjustments - adapting the developed monitoring eco system for project needs by fixing errors, adding new checks and addressing WP6 requests
- auto-deployment of client-side scripts and software updates

Task 6.1 providing monitoring activity was responsible for the service support by actively responding to and providing technical assistance with:

- site addition requests
- site removals as a result of decommissioning process
- user support requests related to the monitoring system registered in PRACE TTS
- general technical issues with monitoring

Apart from maintenance activities typical for production phase of services the monitoring activity performs development activities to extend functionality and improve usability of infrastructure.

The list of new functionalities include:

- user database integration with PRACE LDAP
- on-request: e-mail / jabber notifications
- dashboard improvements including integrated problem view - currently in testing and planned for release shortly after end of Q2/2019
- downloadable availability reports in PDF planned for release in April 2019
- redesigned service dependency to avoid unnecessary alerts
- IGTF monitoring adjustments

In terms of monitoring objects no significant expansion of the infrastructure was observed. Decrease in the following numbers results from decommissioning legacy systems and replacing many small units with larger systems.

Year	**2017**	**2019**
Monitored Sites	28	23
Monitored Generic Services	10	10
Monitored Host Objects	98	85
Monitored Service Objects	1612	1122

Table 11: Overview of monitored metrics

New systems that have been integrated with monitoring infrastructure include:

- CYFRONET-PROMETHEUS
- ICHEC Kay

The systems which have been removed while retaining availability data include:

- ICHEC Fionn
- NIIF-SC
- PSNC-Chimera

3.8 Generic Services

This section describes the actions done during the reporting period within Generic Services subtask. 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 of this task is the provisioning these services and the supervision of their operation, as they are crucial for the day-to-day work of the project.

3.8.1 Roles and actions

The leader of generic services is acting as a liason / consultant to consult with

- PRACE aisbl.
- Other WPs
- External providers

upon request when IT or operational issues or requests of these entities rise.

The following activities have been carried out in the reported period:

- Questions of another WP related to the operation of new or existing services
    - Suggesting operational aspects for new PRACE Events and Training portal with WP4;
    - Further integration plans between Training, Events portals and PRACE Web site;
    - Negotiated major upgrade options for Indico, committed upgrade.
    - Helped finalization of ‘Repository’ new Generic Service service (repository.prace-ri.eu) hosting and features with WP4/WP7
    - Negotiated and established the new training portal systems with WP4 (new Wordpress and ePrints instances) with WP4
    - Negotiated proposal of storage of web content in markdown format with WP3
    - Informed PRACE Security Forum of GS related incidents via reports (e.g PRACE Website incident)
    - Helped WP3 with the tendering process of the complete overhaul of prace-ri.eu portal

- Questions of PRACE aisbl when negotiating with service providers offering service for PRACE
    - Help with consulting mailing lists provider mitigating spam issues;
    - Consulting with service hosts;
    - Helping managing domain portfolio;
    - Help with certificate installation
    - Help detecting and mitigating security incident on website
    - Meeting with European E-Infrastructure Services Gateway (einfracentral.eu) to use their API, planning of scripts to submit PRACE data
    - Monitored KPIs related to generic services
    - Review/analysis of PRACE service portfolio for possible future integrations efforts with sites of other HPC stakeholders (with WP3 and WP4)
- Helping internal workflows of using a shared PRACE-owned resource
    - Handling *.prace-ri.eu wildcard certificate, general help, helping with new certificate requests;
    - Handling prace-ri.eu domain: subdomain requests, helping with resolving domain issues related to a change.

Participation to discussions of hosting centralisation plans of PRACE aisbl, and definition of potential alternatives of hosting e.g. websites has been done by the service.

In connection with the GDPR requirements, an action plan on preparation of implementing new data protection measures was designed and coordinated by PRACE aisbl, BoD and MB with the help of Generic Services leader. The process was started with service overview and by creating an assessment survey, review aspects based on results.

Migration from PRACE SVN to git with their web frontends has been initiated. The security review through a questionnaire about the service has been undergone. A test environment with WP6 related branches and commits migrated from the original PRACE SVN has been done.

4 Conclusions

In this reporting period, task 6.1 has continued the successful operation of the PRACE common services for the Tier-0 sites and the Tier-1 sites providing services for Tier-0. The operational procedures have continued to be successful in maintaining a reliable and available set of integrated services. The on-duty activity has a weekly schedule and continuously supervises the HPC infrastructure to assure a more concrete and punctual support on the day-by-day operation monitoring actions and support. This activity involves 23 PRACE partners with weekly shifts. Every week a report on the duty activity is produced, reporting the status of the infrastructure, of the core service and the problems occurred.

To ensure a seamless use of the PRACE distributed research infrastructure the PRACE common services are available on all PRACE Tier-0 systems and on the identified Tier-1 systems providing services for Tier-0. Through this common service provision, the Tier-0 and Tier-1 systems are presented as a single infrastructure, which enables a smooth interoperation of Tier-0 and Tier-1 systems.

Thanks to the procedures for incident and change management, task 6.1 operates and monitors on a day-by-day basis the complete set of PRACE common services, as defined in the Service Catalogue. This activity will continue and improve in PRACE-6IP in the context of the PRACE 2 programme, where new powerful Tier-0 and Tier-1 systems will further improve the infrastructure and enrich the whole European HPC eco-system at the service of Science.

The task has yet a lot of challenges to overcome PRACE-6IP, starting with the full Accreditation to GÉANT's Trusted Introducer program, further revisions of the PRACE Service Catalogue to incorporate the new demands of services asked by the PRACE user community and tested in task 6.2, security policy evaluation of all new and current services like transition to GCF and GIT, further evaluation of PCPE to reflect on the new architectures deployed on the upcoming PRACE systems and many more.