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

D6.1

Interim report on the PRACE Operational Services
Final

Version: 1.0
Author(s): Filip Stanek, IT4Innovations, Cristiano Padrin, CINECA
Date: 26.08.2020

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: INFRAEDI-823767**	
	Project Title: PRACE Sixth Implementation Phase Project	
	Project Web Site: http://www.prace-project.eu	
	Deliverable ID: D6.1	
	Deliverable Nature: Report	
	Dissemination Level: PU*	**Contractual Date of Delivery:** 31 / 08 / 2020
		Actual Date of Delivery: 31 / 08 / 2020
	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: Interim report on the PRACE Operational Services**	
	ID: D6.1	
	Version: 1.0	**Status:** ***Final***
	Available at: http://www.prace-project.eu	
	Software Tool: Microsoft Word 2016	
	File(s): D6.1	
Authorship	**Written by:**	Filip Stanek, IT4Innovations, Cristiano Padrin, CINECA,
	Contributors:	Miroslav Puskaric, HLRS Ralph Niederbeger, JUELICH Zoltán Kiss, KIFÜ Alessandro Marani, CINECA Michael Mucciardi, BSC Lukasz Flis, CYFRONET
	Reviewed by:	Dirk Brömmel, JUELICH Dimitris Dellis, GRNET
	Approved by:	MB/TB

Document Status Sheet

Version	**Date**	**Status**	**Comments**
0.1	02/07/2020	Draft	First draft of Operational Services
0.2	12/07/2020	Draft	Inputs added for Data and Compute services
0.3	13/07/2020	Draft	Inputs added for AAA, Security, Network and User Services

0.4	29/07/2020	Draft	Inputs added for Generic services, Status of TIER-0 & TIER-1 sites
0.5	30/07/2020	Draft	Inputs added for System upgrades, Monitoring
0.6	31/07/2020	Draft	Inputs added for User services
0.7	03/08/2020	Draft	Draft final version, submitted for internal WP6 review
0.8	05/08/2020	Draft	Draft version ready for internal review process, submitted for MB/TB review
0.9	25/08/2020	Draft	Final DRAFT for MB/TB approval
1.0	26/08/2020	Final version	Final version ready for EC submission

Document Keywords

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

Disclaimer

This deliverable has been prepared by the responsible Work Package of the Project in accordance with the Consortium Agreement and the Grant Agreement n° INFRAEDI-823767. It solely reflects the opinion of the parties to such agreements on a collective basis in the context of the Project and to the extent foreseen in such agreements. Please note that even though all participants to the Project are members of PRACE AISBL, this deliverable has not been approved by the Council of PRACE AISBL and therefore does not emanate from it nor should it be considered to reflect PRACE AISBL's individual opinion.

Table of Contents

List of Figures

List of Tables

References and Applicable Documents

[1] http://www.prace-project.eu
[2] PRACE-5IP Deliverable D6.2, Final Report on the PRACE Operational Services, PRACE, April 2019
[3] AARC Project: https://aarc-project.eu/
[4] RFC 5280, Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile: https://tools.ietf.org/html/rfc5280
[5] EUGridPMA: https://www.eugridpma.org/
[6] IGTF: http://www.gridpma.org/
[7] DART: http://prace-ri.eu/Accounting-Report-Tool
[8] PRACE Helpdesk: https://tts.prace-ri.eu/
[9] Best Practical Request Tracker: https://www.bestpractical.com/rt-and-rtir

List of Acronyms and Abbreviations

AAA	Authorization, Authentication, Accounting
AAI	Authorization 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
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.
OoD	Operator on Duty
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 Intensif, 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 Jülich 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)
UNILUX	University of Luxembourg, Luxembourg
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	Universität 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 2019 - July 2020) to operate and coordinate the common PRACE Operational Services, foreseen by Task 6.1 of WP6 within the PRACE-6IP 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 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 provide continuity to the PRACE Operational Services for the HPC ecosystem.

Eight Tier-0 systems in five countries were operational in the first year of the PRACE-6IP project period.

Furthermore, operational support has been provided to twenty 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 prototype and asses new Operational Services). These Tier-1 systems are distributed among fifteen different countries, ensuring a wide distribution of the European HPC ecosystem and serving the DECI calls.

A new version of the PRACE Service Catalogue, which describes the PRACE common services, has been created and approved by the PRACE BoD to further guide the operational activity. The PRACE Service Catalogue represents a living document reflecting also the requirement to develop new user-friendly services in the context of PRACE.

Based on the procedures for incident and change management, the complete set of PRACE common services as defined in the PRACE 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. This is achieved with periodic teleconferences to monitor the infrastructure and prevent possible incidents which could cause vulnerabilities on the PRACE RI.

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 ecosystem” in PRACE-6IP. This task is responsible for the operation of the set of common services, which present the PRACE Tier-0 systems as an integrated pan-European HPC ecosystem [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 DECI calls, the SHAPE activity towards SMEs, the prototyping and assessment of new Operational Services investigated in Tasks 6.2 and 6.3 of PRACE-6IP WP6, and the testing and utilisation 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 continue 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 first reporting period (May 2019-July 2020) and complements the activity undertaken during the second year of PRACE-5IP and documented in PRACE-5IP Deliverable 6.2 [2].

In the 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. Two successful WP6 Face-to-Face meetings have been organised, in Bled (Slovenia) on 22th–23rd of October 2019 and on-line on 12th and 23rd of March 2020 due to the COVID-19 pandemic restrictions. These meetings aimed 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. Both meetings had around thirty attendees and most of the sites involved in WP6 have been represented. Aiming for new user-friendly services as asked for by the management board and the accompanying update of the PRACE Service Catalogue, an additional four half day long on-line meetings of all the partners were organized on 19th, 20th, 25th, and 26th March.

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 gives a status overview of the common services for the different service areas:

- Network services
- Data services
- Compute services
- Authorization, Authentication and Accounting
- Operational security
- User services
- Monitoring services for operations
- Generic services

Finally, Section 4 represents conclusions.

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

This section presents the changes implemented during the reporting period (May 2019 – July 2020) 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 being detailed. In the reporting 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 PRACE-5IP Project. The work was the continuation and evolution of the activity already in progress. Efforts have been made to keep the documentation up-to-date and feedback has been periodically inquired from the participant sites. After the approval of the new PRACE Service Catalogue (June 2020), the OoD Activity has been added to the list of core services and it's being managed as part of the system status overview.

2.1 Status of Tier-0 & Tier-1 sites

The Tier-0 and Tier-1 systems constitute the HPC ecosystem offering high level services to several European computational communities (both public and private researchers). At the end of the reporting period, according to the information recorded by all partners, eight Tier-0 systems and twenty Tier-1 systems are fully or partially integrated and in production. All systems are continuously monitored and the operational quality is assured by employing the specific daily On-duty Activity provided by the members of PRACE-6IP WP6 task 6.1 as described below.

2.1.1 On-duty Activity

The On-duty Activity is guaranteed daily by all partners who provide effort and/or systems in WP6 and is assigned 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 redirected to the local help desk of each individual site.

A weekly report on the On-duty Activity is produced, where the operators 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 weekly reports are kept on the PRACE wiki being weekly updated, using the report template agreed at the beginning of PRACE-5IP and confirmed for PRACE-6IP.

Starting from the 1st of January 2020, a schedule has been defined for the twenty five 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 list of partners involved in the schedule are reported in Table 1 below.

1 BSC	14 NCSA
2 CASTORC	15 KIFU
3 CEA	16 PDC
4 CINECA	17 PSNC
5 GRNET	18 WCNS
6 CSC	19 SURFSARA
7 CYFRONET	20 VSB-TUO
8 EPCC	21 UC-LCA
9 JUELICH	22 UHEM
10 HLRS	23 ETH-CSCS
11 ICHEC	24 UNILUX
12 CESGA	25 CCSAS
13 UL	

Table 1: PRACE partners involved in on duty activity

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.

The TTS system is organised in queues where every site is responsible for one queue, apart from the "Generic" queue that is looked after by the Operator on Duty. Table 2lists the sites that have a specific queue defined for themselves and that they are asked to keep monitored.

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

Table 2: PRACE partners involved in the monitoring of their own TTS queue

There are some partner sites that do not yet have to monitor their TTS queue. This is a temporary situation caused by the integration status of the sites that are new to the service. In addition, there are still queues kept in place for some partners who were active in the previous Implementation Phases of PRACE, although not having an integrated system in PRACE-6IP. Those are kept to ensure continuation for such sites that potentially re-join PRACE again.

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

Tickets resolving efficiency	
Description:	Ratio of resolved tickets 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':	Alessandro Marani and Francesco Cola (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 3: 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) (Note: Vector T holds the Initial Response times)
Inputs:	Number of open tickets (O) Number of resolved tickets (R) Vector of Creation times (CT) Vector of Starting times (ST)
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':	Alessandro Marani and Francesco Cola (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 4: Average initial response

In the last twelve months 115 tickets have been created: 1 of them is stalled, 2 are still open and 112 had been resolved. This means that the percentage of resolved tickets with respect to the created tickets is 97,4%.

It is important to underline that, in principle, the monitored KPIs and reported values are 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

At present (July 2020), according to the shared information by all project partners including their 3rd parties, the Tier-0 ecosystem is made up of eight systems, distributed in seven sites and operated by six different sites in five Hosting Members countries (France, Germany, Italy, Spain and Switzerland) as reported in Table 5.

The peak performance of the Tier-0 systems ranges from more than 2 PFlop/s up to around 16 PFlop/s for the PizDaint system in Switzerland. Peak performance for Hawk system in Germany is unannounced but is supposed to be about 26 PFlop/s. Two Tier-0 are accelerated: JUELICH/JUWELS with Nvidia V100 and CSCS/PizDaint with Nvidia Tesla P100. There is not a single dominant vendor: three systems are from IBM/Lenovo and three other systems are from ATOS/Bull Sequana.

Most of Tier-0 systems are ranked in the Top500 (June 2020) and the highest ranked system (PizDaint) is in 10th position. Since March 2020 HLRS (Germany) has integrated a new Tier-0 system, “Hawk”, with HPE Apollo architecture, while the previous Tier-0 system “Hazel Hen” has been decommissioned. This new system has not yet ranked in the Top 500.

As far as the Tier-1 ecosystem is concerned, twenty systems are operating as Tier-1 for Tier-0 services. These Tier-1 systems are distributed in seventeen different PRACE sites, operated by fifteen partners, in fourteen 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 2.5 PFlop/s. Seven systems deliver more than 1 PFlop/s.

About half of the Tier-1 systems is accelerated, either with Intel Xeon Phi or Nvidia accelerators. Several different vendors and system architectures (platforms) are present, presenting a real advantage for the PRACE HPC infrastructure: HPE, SGI, ATOS/Bull, Cray, IBM and Lenovo are all represented. Most typical architecture is an x86 cluster, popular platforms include ATOS Bullx/Sequana and Lenovo (former IBM) NeXtScale.

In the last twelve months, four systems were decommissioned (SISU, SEGED, HYDRA and TURING), and four new systems have been integrated or are in the process of being integrated (FINISTERRAE, GALILEO, HPCFS-U and SARIYER).

The whole infrastructure is constantly in operational status and incidents are infrequent. In the last twelve months, only four unscheduled maintenances have been reported. They all effected one site at a time and were resolved in a matter of hours.

Partner	Country	Tier-0	Architecture/Platform - CPU - Accelerator	Rpeak (TFlop/s)
GCS-JUELICH	Germany	JUWELS	ATOS Sequana - Dual Intel Xeon Platinum 8168; 2x24-cores; 2.7 GHz - Nvidia V100	9891.1
GCS-LRZ		SuperMUC phase 2	Lenovo NeXtScale nx360M5 WCT - Intel Haswell-EP 14 core; 2.6 GHz	3580.0
GCS-HLRS		Hawk	HPE Apollo - AMD Epyc Rome 7742; 64-core; 2.25 GHz	TBA
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; 2.70 GHz	2339.6
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	Lenovo Adam Pass (Intel OmniPath Cluster) Intel Xeon Phi 7250 (Knights Landing) 1x68-core; 1.40 GHz	11000.0

Partner	Country	Tier-0	Architecture/Platform - CPU - Accelerator	Rpeak (TFlop/s)
			Lenovo Stark (Intel OmniPath Cluster) Intel Xeon 8160 (SkyLake); 2x24-core; 2.10 GHz	9000.0

Table 5: PRACE Tier-0 systems

Partner	Country	Tier-1	Architecture - CPU - Accelerator	Rpeak (TFlop/s)
EPCC	UK	Archer	Cray XC30 - Intel Ivy Bridge; 12-core; 2.7 GHz	2550.5
BSC	Spain	Minotauro	Bull Bullx B505/R421-E4 - Intel Westmere-EP; 6-core; 2.53 GHz + Intel Haswell E5-2630 v3; 8-core; 2.60 GHz - Nvidia/M2090 and K80	250.9
SURFsara	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
PDC-KTH	Sweden	Beskow	Cray XC40 - Intel Haswell; 32-core; 2.3 GHz	1973.0
PSNC-CYFRONET	Poland	Zeus BigMem	HP BL685c G7 AMD - AMD Interlagos; 16-core; 2.3 GHz	61.2

Partner	**Country**	**Tier-1**	**Architecture - CPU - Accelerator**	**Rpeak (TFlop/s)**
		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
UIO - SIGMA	Norway	Abel	MEGWARE MiriQuid - Intel Sandy Bridge-EP; 8-core; 2.6 GHz	178.6
ICHEC	Ireland	Kay	Intel/Penguin Computing - Intel SkyLake; 2x20-core; 2.4GHz	665.0
IT4I-VSB	Czech Republic	Anselm	Bull Bullx B510/B515 - Intel Sandy Bridge-EP; 8-core; 2.4 GHz	66.0
		Solomon	SGI ICE-X - Intel Xeon E5-2680v3; 12-core; 2.5 GHz - Intel PHI	2000.0
KIFU	Hungary	Leo	HP SL250s - Intel Xeon E5-2650 v2; 2.60 GHz - Nvidia K20, K 40	254.0
		PHItagoras	HP SL250s - Intel Xeon E5-2680 v2; 2.80 GHz - Intel/PHI 7120	27.0
CCSAS	Slovakia	Aurel	IBM; Power 775 - IBM Power7; 8-core; 3.84 GHz	128

Partner	Country	Tier-1	Architecture - CPU - Accelerator	Rpeak (TFlop/s)
GRNET	Greece	Aris	IBM NeXtScale nx360 M4 - Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80 GHz * 2	190.8
			Dell PowerEdge R820 - Intel(R) Xeon(R) CPU E5-4650 v2 @ 2.40 GHz * 4	36.6
			Dell PowerEdge R730 - Intel(R) Xeon(R) CPU E5-2660 v3 @ 2.60 GHz * 2 - Nvidia Tesla K40m and Intel Xeon Phi 7120	244.3
			Supermicro SYS-4028GR-TVRT - Intel(R) Xeon(R) CPU E5-2698 v4 @ 2.20 GHz * 2 - Nvidia Tesla V100-SXM2	63.1
CESGA	Spain	Finisterrae	Bull Bullx B505/R424-E4 - Intel Haswell 12-core E5-2680v3; 2.50 GHz - NVidia Kepler K80	328.0
CINECA	Italy	Galileo	Lenovo NextScale - Intel Xeon E5-2697 (Broadwell); 2.30 GHz - NVidia Kepler K80	1536.0
UL	Slovenia	HPCFS-U	Intel x64 HPE + Supermicro blades - Intel Xeon E5-2680V3; 2.50 GHz	24.3

Partner	Country	Tier-1	Architecture - CPU - Accelerator	Rpeak (TFlop/s)
			- NVIDIA Tesla K80	
UHEM	Turkey	Sariyer	Intel x64 (Super Micro / Huawei / DELL) - Intel Xeon (E5-2680V3;2.50 GHz, E5-2680V4;2.40 GHz, Gold 6148; 2.40 GHz) - NVIDIA Tesla K20m	n/a

Table 6: PRACE Tier-1 systems

2.2 System Upgrades

This section describes the activity related to the system upgrades and integration in the PRACE ecosystem. After the approval of the new PRACE Service Catalogue (June 2020), the activity was added to the list of core services. In the reporting period this activity has been included under WP6 subtask 6.1.2 and is led by Barcelona Supercomputer Centre.

2.2.1 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 PRACE Service Catalogue requirements.

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

This section presents the new sites and systems that are planned to be integrated in the PRACE infrastructure during PRACE-6IP.

2.3.1 Overview of new Tier-0/Tier-1 sites & systems on PRACE-6IP

During PRACE-6IP, eleven systems are planned to be integrated into the PRACE infrastructure and to start providing services to it, as shown in the table below. Six of these systems began their integration process in PRACE-5IP and their status is currently in progress. Five systems are new and are scheduled to be fully integrated during PRACE-6IP. One system (Cy-Tera) has not yet been acquired and is scheduled to begin the process of systems integration once it has been deployed by the end of Q3/Q4 2020.

Site	System Name	System Details URL	Tier	Integration Start Date	Status	Last Check
CCSAS (Slovakia)	Aurel	http://vs.sav.sk/?lang=en&section=departments&sub=vvt&sub2=config	Tier-1	01/05/2017	In progress	29/07/2020
UL (Slovenia)	HPCFS	http://hpc.fs.uni-lj.si/hardware	Tier-1	approx. 07/2017	In progress	29/07/2020
CESGA (Spain)	FinisTerrae	https://www.cesga.es/en/infrastructures/computing/	Tier-1	01/04/2016	In progress	29/07/2020
UHEM (Turkey)	Sariyer	http://wiki.uhem.itu.edu.tr/w/index.php/English	Tier-1	03/2019	In progress	29/07/2020
CEA (France)	Irene	http://www-hpc.cea.fr/en/complexe/tgcc-Irene.htm	Tier-0	21/03/2018	In progress	29/07/2020
UC-LCA (Portugal)	Navigator	https://www.uc.pt/lca/ClusterResources/Navigator/description	Tier-1	10/2018	In progress	29/07/2020
IT4I-VSB-TUO (Czechia)	Barbora	https://docs.it4i.cz/barbora/introduction/	Tier-1	n/d	In progress	29/07/2020
CINECA (Italy)	Galileo	https://wiki.u-gov.it/confluence/display/SCAIUS/UG3.3%3A+GALILEO+UserGuide	Tier-1	8/1/2019	In progress	29/07/2020
CSC (Finland)	Puhti	https://research.csc.fi/csc-s-servers	Tier-1	14/01/2020	In progress	29/07/2020
HLRS (Germany)	Hawk	https://www.hlrs.de/systems/hpe-apollo-hawk/	Tier-0	2/1/2020	In progress	29/07/2020
CASTORC (Cyprus)	Cy-Tera	http://castorc.cyi.ac.cy/infrastructure	Tier-1	Q3/Q4 2020	In progress	29/07/2020

Table 7: New systems integration

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

The full list of the new systems integrated in the first two years of PRACE-6IP time-frame is reported below:

1. **Aurel - CCSAS – Tier-1**

Aurel is a supercomputer composed by 2 racks, 16 drawers, 128 physical nodes (130 logical nodes), with a total of 4096 cores, which provides a total of 128 TFlop/s. 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.

2. **HPCFS – UL – Tier-1**

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

- Control – 1 node
    - NX server DL160 G6: 2x Intel Xeon X5670 48 GB of memory 2x 500 GB SAS disk (system, mirror) remote management
- Compute – 64 nodes
    - SL2x170zG6 work nodes with Infiniband 64x - 64 nodes x 6 cores x 2 processors = 768 cores (1536 process threads) 2 x Intel Xeon X5670 / node 48 GB 2R DDR3 memory 500 GB disk Infiniband QDR HBA remote management
- Networking
    - Infiniband network 72 port QDR – 32 Gbit/s throughput
    - 6 x Voltaire 32-port bound in "fat-tree" topology with "constant bisection bandwidth"
    - 4 x 48 port ProCurve 1 Gbit switches with 4x10 Gbit uplink
- Storage
    - 25 TB NFS server for HOME
        - DL180 G6 and D2700 disk array E5504 processor 8 GB memory 3x2 GB 2Rx8 10 Gbit NFS 50 x SFF 500 GiB 2.5 "SATA2 disk (data)
    - 12 TB LUSTRE DL170 servers in DL1000 case for $ WORK
        - 6 x OSS servers 4x SFF 500 GB SATA disks in RAID5 E5504 processor 4 GB RAM

- MDS server 2x160 GB SFF SATA drives in RAID1 for the system 6xSATA 160 GB disks for OST E5504 processor 8 GB RAM

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

3. FinisTerrae – CESGA – Tier-1

FinisTerrae is a Linux based heterogeneous cluster, with an Infiniband FDR low latency network interconnecting 320 compute nodes based on Intel Xeon Haswell processors. Together, these nodes contain 7712 cores and are able to provide a peak computing power of 328 TFlop/s, 44.8 TB of RAM and 3.4 PB of disk capacity (local and network), and 1.36 PB of long term tape storage. The system includes a high performance parallel Lustre storage system which provides access to 750 TB of storage at 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.

4. Sariyer – UHEM – Tier-1

Sariyer includes a Cluster of mixed Intel processors with a total of 143 compute nodes and 4396 cores, using an InfiniBand FDR 56 Gb/s / EDR 100 Gb/s network. The Sariyer system PRACE integration process is in progress. Once this is complete, PRACE users will be provided with standard access to Sariyer resources.

- Configuration 1 nodes: 35 compute nodes, 349 TB Lustre parallel storage system
  - Intel XEON Gold 6148 V5, 192 GB Memory
- Configuration 2 nodes: 93 compute nodes, 349 TB Lustre parallel storage system
  - Configuration 2A, 15 compute nodes: Intel XEON Gold E5-2680 v4, 64 GB Memory
  - Configuration 2B, 83 compute nodes: Intel XEON Gold E5-2680 v4, 128 GB Memory
- Configuration 3 nodes: 15 compute nodes, 349 TB Lustre parallel storage system
  - Intel XEON Gold E5-2680 v4, 512 GB Memory

5. Irene – CEA – Tier-0

Irene is Tier-0 cluster for CEA, it is a 9.4 PFlop system with two partitions, one with 1656 compute nodes based on Intel SkyLake processors and 828 compute nodes based on Intel KNL processors. The SkyLake partition will be connected using an Infiniband EDR network, while the KNL partition will be using the new Bull eXascale Interconnect (BXI). An extension of Joliot-Curie was installed at the end of 2019 consisting of 2,292 dual processor AMD Rome (Epyc) compute nodes at a performance of 11.75 PFlop/s. Also an exploratory partition based on ARM processors will be installed at the end of 2020.

- SKL Irene (Skylake):
  - 1,656 dual-processor Intel Xeon 8168 (SKL) fine nodes at 2.7 GHz with 24 cores per processor, for a total of 79,488 computing cores and a power of 6.86 PFlop/s

    - 192 GB DDR4 memory / node
    - Infiniband EDR interconnection network.
    - Access to 500 GB/s to a multi-tiered Lustre parallel file system
- KNL Irene (Knight Landing):
    - 828 Intel Xeon Phi 7250 (KNL) manycore nodes at 1.4 GHz with 68 cores per processor, i.e. a total of 56,304 cores for a performance of 2.52 PFlop/s
    - 96 GB DDR4 + 16 GB MCDRAM / node
    - Atos-BULL BXI interconnection network.
    - Access to 500 GB/s to a multi-tiered Lustre parallel file system

An extension of Joliot-Curie was installed at the end of 2019, it consists of the following partitions:

- **AMD Irene ROME (Epyc)**
    - 2,292 dual processor AMD Rome (Epyc) compute nodes at 2.6 GHz with 64 cores per processor, for a total of 293,376 compute cores and a power of 11.75 Pflop/s
    - 256 GB DDR4 memory / node
    - Infiniband HDR100 interconnection network.
- **Irene V100 (Data processing partition / IA partition)**
    - 32 hybrid nodes with per node 2 Intel CascadeLake 20-core 2.1 GHz processors and 4 nVIDIA V100 GPUs, for a total of 128 GPUs for a power of 1.13 PFlop/s.

6. 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. The delay of the integration into PRACE was caused by the move into a new datacentre in Portugal.

7. Barbora – IT4I/VSB-TUO - Tier-1

The Barbora cluster consists of 201 computational nodes of which 192 are regular compute nodes, 8 are GPU Tesla V100 accelerated nodes and 1 is a fat node. Each node is a powerful x86-64 computer, equipped with 36/24/128 cores (18-core Intel Cascade Lake 6240 / 12-core Intel Skylake Gold 6126 / 16-core Intel Skylake 8153), with at least 192 GB of RAM. User access to the Barbora cluster is provided by two login nodes. The nodes are interlinked through high speed InfiniBand and Ethernet networks. The Fat node is equipped with a large amount (6144 GB) of memory.

Virtualization infrastructure provides resources to run long-term servers and services in virtual mode.

- 192 compute nodes without an accelerator
- 8 compute nodes with a GPU accelerator - 4x NVIDIA Tesla V100-SXM2
- 1 fat node - equipped with 6144 GB of RAM

8. **Galileo – CINECA - Tier-1**

The GALILEO supercomputer has been introduced first in January 2015 and it has been available to Italian public and industrial researchers until January 2018. It has been the national Tier-1 system for scientific research. Starting from January 2018 GALILEO has been reconfigured with Intel Xeon E5-2697 v4 (Broadwell) nodes, inherited from MARCONI system.

- 1022 36-core compute nodes. Each one contains 2 18-cores Intel Xeon E5-2697 v4 (Broadwell) at 2.30 GHz
- 60 compute nodes equipped with 1 nVidia K80 GPU
- 2 compute nodes equipped with 1 nVidia V100 GPU

9. **Puhti – CSC – Tier-1**

The Puhti supercomputer was launched on September 2, 2019. It is an Atos cluster system with a variety of different node types. It is targeted at a wide range of workloads and it can also be used to run larger simulations until Mahti becomes available.

Puhti has a total of 682 CPU nodes, with a theoretical peak performance of 1.8 PFlop/s. Each node is equipped with two latest generation Intel Xeon processors, code name Cascade Lake, with 20 cores each running at 2.1 GHz. The interconnect is based on Mellanox HDR InfiniBand. The nodes are connected with a 100 Gb/s HDR100 links a fat tree topology with a blocking factor of approximately 2:1.

The Puhti AI artificial intelligence partition has a total of 80 GPU nodes with a total peak performance of 2.7 PFlop/s. Each node has two latest generation Intel Xeon processors, code name Cascade Lake, with 20 cores each running at 2.1 GHz. They also have four Nvidia Volta V100 GPUs with 32 GB of memory each. The nodes are equipped with 384 GB of main memory and 3.6 TB of fast local storage. This partition is engineered to allow GPU-intensive workloads to scale well across multiple nodes. The interconnect is based on a dual-rail HDR100 interconnect network connectivity providing 200 Gb/s of aggregate bandwidth in a non-blocking fat-tree topology.

- M nodes: 532 compute nodes, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 192 GB Memory
- L nodes: 92 compute nodes, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 384 GB Memory
- IO nodes: 40 compute nodes, 3600 GB local disk, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 384 GB Memory

- XL nodes: 12 compute nodes, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 768 GB Memory
- BM nodes: 12 compute nodes, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 1.5 TB Memory
- GPU nodes: 80 compute nodes, 3600 GB local disk, 4.8 PB Lustre parallel storage system
  - 2x20 cores @2.1 GHz (Intel XEON Gold 6230), 384 GB Memory
  - 4 GPUs(Nvidia V100) connected with NVLink, 4 x 32 GB Memory

10. Hawk – HLRS – Tier-0

Hawk is the flagship supercomputer of the High-Performance Computing Center Stuttgart (HLRS). At time of installation, it is amongst the fastest high-performance computers in the world operating at a peak system performance of 26 PFlop/s and the fastest general-purpose machine for industrial production in Europe.

- 5632 compute nodes, 2x64 cores @2.25 GHz (AMD EPYC 7742), 256 GB(16x16 GB) Memory
- Total System Memory ~1.44 PB
- Network:
  - Interconnect Topology: Enhanced 9D-Hypercube
  - Interconnect Bandwidth: 200 Gbits/s Infiniband
- Storage
  - DDN EXAScaler with IME ~25 PB
- Power Consumption
  - Max Power Consumption Per Rack (44 racks): ~90 kW
  - System Power Consumption (Normal): ~3.4 MW
  - System Power Consumption (LinPack Operation): ~4.1 MW

11. Cy-Tera – CASTORC – Tier-1

The Cy-Tera project aims at creating a research facility including a High-Performance Computing (HPC) infrastructure supporting cutting-edge scientific applications, including computational science research, associated HPC user support and training programs. It will constitute a major step in the development of the Computation-based Science and Technology Research Center (CaSToRC) of the Cyprus Institute (CyI). The Cy-Tera facility will be the first HPC facility at multi-Teraflop/s level in Cyprus, serving the needs of CyI and its partners (The National Center for Supercomputing Applications, located at the University of Illinois at Urbana-Champaign, the Jülich Supercomputing Centre, the University of Cyprus, and the Synchrotron-light for Experimental Science and Applications in the Middle East) for frontier research applications with scientific and social importance. The principal objectives of the Cy-Tera project are to:

- Set up a Tier-1 HPC facility based on multiple core architecture with scalable design.
- To provide computational resources and associated user support for the CyI research thrusts as well as for the research activities of its partner institutions.
- To develop a portfolio of research activities in computational science, including research at the interface of computer science and scientific computing, thus promoting cross-disciplinary research regarding new computer architectures and software engineering.

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. The 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 first reporting period of PRACE-6IP 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 reporting period, the PRACE partner EPCC has been integrated into the PRACE-MD-VPN. The partners UHeM (Turkey), and UC-LC (Portugal), are in the process to be connected soon. Besides the integration of new partners, several partners installed new HPC systems, which had to be integrated, sometimes in parallel to old systems still in production. A web page displays the reachability of relevant HPC systems, shown in Figure 1

BSC BSC:212.128.224.2 43.6/43.8/43.9 2020-07-07 20:51:01	**CASTORC** CASTORC:82.116.198.110 79.2/82.9/95.3 2020-06-30 18:41:01	**CCSAS** CCSAS:147.213.81.6 -/-/- 2019-11-03 02:31:01	**CCSAS** CCSAS:147.213.81.5 19.4/19.5/19.5 2020-06-30 15:01:01	**CEA-TGCC** irene-eu-prace-gw.ccc.cea.fr. 23.4/23.6/23.7 2020-07-07 20:51:01	**CINECA** r000u07l02-prace.marconi.cineca.it. 26.7/26.8/26.9 2020-06-23 00:31:02
CINECA r000u06l01-prace.marconi.cineca.it. 26.8/26.9/27.4 2020-06-23 00:31:02	**CINECA** r000u08l03-prace.marconi.cineca.it. 26.7/26.7/26.7 2020-06-26 18:06:01	**CSC** puhti-prace.csc.fi. -/-/- 2020-05-13 15:21:01	**CSCS** CSCS:148.187.128.41 19.1/19.4/20.1 2020-07-01 16:36:02	**EPCC** dtn01-prace.rdf.ac.uk. -/-/- 2020-05-11 15:21:01	**EPCC** dtn02-prace.rdf.ac.uk. -/-/- 2020-05-11 15:21:01
FZJ judac04p.zam.kfa-juelich.de. 0.18/0.22/0.26 2020-07-06 14:46:02	**FZJ** judac03p.zam.kfa-juelich.de. 0.20/0.26/0.41 2020-07-06 13:26:01	**FZJ** judac05.zam.kfa-juelich.de. 0.24/0.48/1.13 2020-07-07 20:36:01	**FZJ** monctl.net.prace.fz-juelich.de. 0.03/0.04/0.08 2015-12-17 12:58:42	**FZJ** judac05p.zam.kfa-juelich.de. 0.20/0.40/1.05 2020-07-07 20:36:01	**GRNET** gssh-prace-02.aris.grnet.gr. 47.9/48.1/48.6 2020-06-21 00:26:01
GRNET gssh-prace-01.aris.grnet.gr. 48.4/48.5/48.5 2020-06-22 02:21:01	**ICHEC** prace-gw.kay.ichec.ie. 34.1/34.1/34.3 2020-06-30 18:41:01	**ICHEC** prace-login.kay.ichec.ie. 34.3/34.4/34.7 2020-06-30 18:41:01	**ICHEC** prace-gridftp-fe.kay.ichec.ie. 34.2/34.4/34.6 2020-07-08 21:16:01	**IDRIS** zahir135-dgiga0.idris.fr. 22.6/22.9/23.4 2020-07-07 20:51:01	**LECAD** LECAD:193.2.78.225 24.9/25.9/29.4 2020-07-02 16:31:01
LRZ skx.supermuc-prace.lrz.de. 14.1/14.1/14.2 2020-06-25 23:21:01	**LRZ** skx.supermuc-prace.lrz.de. 14.0/14.0/14.2 2020-06-20 21:36:01	**LRZ** skx.supermuc-prace.lrz.de. 14.1/14.1/14.2 2020-07-10 11:41:01	**LRZ** skx.supermuc-prace.lrz.de. 14.0/14.1/14.1 2020-07-08 09:21:01	**NIIF** login-vlan907.debrecen2.hpc.niif.hu. 26.4/26.6/27.4 2020-07-07 13:36:01	**NIIF** login-vlan907.budapest2.hpc.niif.hu. -/-/- 2020-06-28 14:06:01
PSNC PSNC:150.254.128.1 23.1/23.3/23.5 2020-07-09 03:31:01	**RZG** mon-prace.mpcdf.mpg.de. -/-/- 2020-05-13 17:11:01	**SURFSARA** int1-prace.cartesius.surfsara.nl. -/-/- 2020-05-19 12:36:01	**UHEM** UHEM:160.75.120.180 37.2/37.2/37.3 2020-07-13 04:31:02	**VSB-TUO** gridftp-prace.anselm.it4i.cz. 18.8/18.9/19.1 2020-07-07 05:06:01	**VSB-TUO** gridftp-prace.salomon.it4i.cz. 18.8/18.9/19.1 2020-07-07 05:06:01
WCSS prace-bem-int.wcss.pl. 28.5/28.7/28.8 2020-07-13 08:46:01					

Figure 1: PRACE partition reachability

A further web page has been created, which displays in a five minute interval the network reachability of the different services, listed in the prace_service config file.

SITE	SERVICE	SERVERNAME	IP-ADDRESS	REACHABILITY	NET-LOCATION	LAST CHANGED
BSC	Iperf	212.128.224.2	212.128.224.2	43.8/43.9/44.1	PRACE-MDVPN	2020-06-25 19:02:02
BSC	internal gridftp	gftp.prace.bsc.es	212.128.224.7	43.7/43.8/43.9	PRACE-MDVPN	2020-06-30 12:17:01
CEA	external gridftp	garbin-prace.eole.ccc.cea.fr	132.167.142.97	not reachable	extern	2020-04-22 07:42:01
CEA	external gsissh	irene-amd-prace.ccc.cea.fr	132.167.142.120	not reachable	extern	2020-06-17 06:07:02
CEA	external gsissh	irene-eu-prace.ccc.cea.fr	132.167.142.112	23.6/23.6/23.7	extern	2020-06-30 00:47:01
CESGA	MyProxy	dtn.srv.cesga.es	193.144.32.30	not reachable	extern	2020-07-01 15:37:02
CINECA	Iperf	130.186.26.7	130.186.26.7	26.7/26.8/26.9	PRACE-MDVPN	2020-06-23 00:32:01
CINECA	MyProxy	grid.hpc.cineca.it	130.186.17.234	not reachable	extern	2020-06-12 06:27:01
CINECA	external gsissh	login.marconi.cineca.it	130.186.17.131	29.4/29.5/29.6	extern	2020-06-13 15:17:01
CINECA	external gsissh	login.marconi.cineca.it	130.186.17.132	29.4/29.5/29.6	extern	2020-06-15 11:57:01
CINECA	external gsissh	login.marconi.cineca.it	130.186.17.133	29.4/29.5/29.5	extern	2020-06-26 18:07:01
CINECA	internal gsissh	gssh-prace.marconi.cineca.it	130.186.26.8	26.6/26.7/26.8	PRACE-MDVPN	2020-06-23 00:32:01
CINECA	internal gsissh	gssh-prace.marconi.cineca.it	130.186.26.9	26.7/26.8/26.9	PRACE-MDVPN	2020-06-26 18:07:01
CSCS	Iperf	148.187.128.41	148.187.128.41	19.0/19.1/19.2	PRACE-MDVPN	2020-07-01 16:37:02
CYFRONET	external gridftp	prace-ui.cyfronet.pl	149.156.9.104	30.1/30.2/30.4	extern	2020-05-25 06:57:01
CYFRONET	external gridftp	prace.prometheus.cyfronet.pl	149.156.9.156	not reachable	extern	2020-04-07 13:11:46
CYFRONET	internal gridftp	prace-int.cyfronet.pl	150.254.128.65	30.8/30.9/30.9	PRACE-MDVPN	2020-06-21 00:02:02
CYFRONET	internal gridftp	prace-int.prometheus.cyfronet.pl	150.254.128.67	30.9/31.0/31.0	PRACE-MDVPN	2020-06-20 23:52:02
CaSToRC	Iperf	82.116.198.110	82.116.198.110	116/242/331	PRACE-MDVPN	2020-06-30 18:42:01
EPCC	external gridftp	dtn01.rdf.ac.uk	193.62.216.1	not reachable	extern	2020-05-11 15:17:01
EPCC	internal gridftp	dtn01-prace.rdf.ac.uk	193.62.216.70	not reachable	PRACE-MDVPN	2020-05-11 15:17:01
GRnet	Iperf	195.251.114.117	195.251.114.117	45.9/46.1/46.2	PRACE-MDVPN	2020-07-01 20:42:02
GRnet	external gridftp	gftp.aris.grnet.gr	195.251.23.89	47.2/47.3/47.4	extern	2020-07-02 00:47:01
GRnet	external gsissh	login.aris.grnet.gr	195.251.23.78	47.1/47.2/47.3	extern	2020-06-26 13:17:01
GRnet	external gsissh	login.aris.grnet.gr	195.251.23.79	44.7/44.8/44.9	extern	2020-06-26 13:17:01
GRnet	internal gridftp	gftp-prace.aris.grnet.gr	195.251.114.116	46.2/46.2/46.5	PRACE-MDVPN	2020-07-02 00:52:01
GRnet	internal gsissh	gssh-prace.aris.grnet.gr	195.251.114.114	48.3/48.4/48.6	PRACE-MDVPN	2020-06-22 02:22:01
GRnet	internal gsissh	gssh-prace.aris.grnet.gr	195.251.114.115	46.7/46.8/46.9	PRACE-MDVPN	2020-06-21 00:27:01
ICHEC	internal gridftp	prace-gridftp-fe.kay.ichec.ie	193.1.201.21	34.3/34.4/34.7	PRACE-MDVPN	2020-06-24 23:07:01
ICHEC	internal gsissh	prace-login.kay.ichec.ie	193.1.201.19	34.4/34.5/34.9	PRACE-MDVPN	2020-06-12 11:12:01
IT4I	external gridftp	gridftp.anselm.it4i.cz	195.113.250.84	16.9/16.9/17.0	extern	2020-06-16 13:27:02
IT4I	external gridftp	gridftp.salomon.it4i.cz	195.113.250.86	16.7/16.9/17.0	extern	2020-06-13 21:47:01
IT4I	external gsissh	anselm.it4i.cz	195.113.250.82	16.9/16.9/17.0	extern	2020-05-22 12:37:01
IT4I	external gsissh	anselm.it4i.cz	195.113.250.83	16.8/16.9/17.0	extern	2020-05-22 12:37:01
IT4I	external gsissh	salomon.it4i.cz	195.113.250.131	16.7/16.8/16.9	extern	2020-05-22 12:37:01
IT4I	external gsissh	salomon.it4i.cz	195.113.250.132	16.7/16.8/16.9	extern	2020-05-22 12:37:01
IT4I	external gsissh	salomon.it4i.cz	195.113.250.133	16.7/16.8/16.9	extern	2020-05-22 12:37:01
IT4I	external gsissh	salomon.it4i.cz	195.113.250.134	16.8/16.8/16.9	extern	2020-06-23 13:22:02
IT4I	internal gridftp	gridftp-prace.anselm.it4i.cz	195.113.250.164	18.8/18.9/19.0	PRACE-MDVPN	2020-06-20 23:17:01
IT4I	internal gridftp	gridftp-prace.salomon.it4i.cz	195.113.250.173	18.8/19.0/19.2	PRACE-MDVPN	2020-06-20 23:47:01
IT4I	internal gsissh	anselm-prace.it4i.cz	195.113.250.162	18.9/18.9/19.0	PRACE-MDVPN	2020-06-20 23:37:01
IT4I	internal gsissh	anselm-prace.it4i.cz	195.113.250.163	18.9/18.9/19.0	PRACE-MDVPN	2020-06-21 00:17:02
IT4I	internal gsissh	salomon-prace.it4i.cz	195.113.250.166	18.7/18.8/18.9	PRACE-MDVPN	2020-05-22 12:37:01
IT4I	internal gsissh	salomon-prace.it4i.cz	195.113.250.167	18.6/18.7/18.8	PRACE-MDVPN	2020-06-21 00:42:01
IT4I	internal gsissh	salomon-prace.it4i.cz	195.113.250.168	18.8/18.8/18.9	PRACE-MDVPN	2020-06-21 00:02:02
IT4I	internal gsissh	salomon-prace.it4i.cz	195.113.250.169	18.7/18.8/18.9	PRACE-MDVPN	2020-06-23 13:22:02
NIIF	internal gridftp	leo-login.sc.niif.hu	193.224.66.200	not reachable	PRACE-MDVPN	2020-06-29 22:17:01
NIIF	internal gridftp	phitagoras.sc.niif.hu	193.224.66.196	not reachable	PRACE-MDVPN	2020-06-28 14:07:01
PSNC	external gridftp	eagle.man.poznan.pl	150.254.160.193	21.0/21.1/21.4	extern	2020-05-25 06:57:01
PSNC	internal gridftp	eagle-prace.man.poznan.pl	150.254.128.4	23.0/23.1/23.1	PRACE-MDVPN	2020-06-21 00:22:02
SURFsara	GT5 GRAM5 internal	int2-prace.cartesius.surfsara.nl	145.100.18.15	not reachable	PRACE-MDVPN	2020-05-19 12:37:01
SURFsara	external gsissh	int1-bb.cartesius.surfsara.nl	145.100.200.14	11.1/11.3/11.4	extern	2020-06-30 10:42:01
SURFsara	internal gsissh	int1-prace.cartesius.surfsara.nl	145.100.18.14	not reachable	PRACE-MDVPN	2020-05-19 12:37:01
UiO	external gridftp	gridftp1.prace.uio.no	129.240.189.169	25.9/26.0/26.1	extern	2020-06-17 04:57:01
UiO	external gsissh	login1.prace.uio.no	129.240.189.160	25.8/25.9/26.0	extern	2020-06-25 19:47:01
WCSS	external gridftp	prace-bem.wcss.pl	156.17.5.146	26.1/26.2/26.3	extern	2020-07-02 16:12:02
WCSS	internal gridftp	prace-bem-int.wcss.pl	150.254.128.34	28.8/28.8/28.9	PRACE-MDVPN	2020-06-30 08:47:02

Figure 2: PRACE server/service reachability

Last but not least the "PRACE Path Discovery" web page lists the reachability of the different partners (ping to sample HPC system at every partner site) and showing additionally traceroute information.

PRACE Path Discovery

last page update: 2020-07-13 10:20:32

PRACE Site	System	HPC site or test IP address	Side internal	local router IP address	local link to PRACE-VPN via NREN	NREN Router IP address	PRACE VPN
	PRACE-Mon (Germany)	monctl.net.prace.fz-juelich.de. 134.94.115.218	Traceroute: 134.94.115.218				
	NONE.cy (Cyprus)	{CastoRC-NONE.cy} 82.116.198.110	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * 82.116.198.110				
	Anselm (Czech Republic)	gridftp-prace.anselm.it4i.cz. 195.113.250.164	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * 195.113.250.164				
	Salomon (Czech Republic)	gridftp-prace.salomon.it4i.cz. 195.113.250.173	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * 195.113.250.173				
	Puhti (Finland)	puhti-prace1.csc.fi. 128.214.250.21	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * 128.214.250.121 128.214.250.120 * *				
	Irene (France)	irene-eu-prace-gw.ccc.cea.fr. 132.167.142.112	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * * 132.167.142.112				
	Turing (France)	zahir135-dgiga0.idris.fr.					
	JuDACsrv (Germany)	judac05p.zam.kfa-juelich.de. 134.94.115.217	Traceroute: 134.94.115.217				
	SuperMuc (Germany)	skx.supermuc-prace.lrz.de. 195.37.7.1 195.37.7.2 195.37.7.3 195.37.7.4	Traceroute: 134.94.115.193 134.94.111.38 * * * * * 195.37.7.4				
	Pracempcdf (Germany)	mon-prace.mpcdf.mpg.de. 130.183.230.67	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * * * * *				
	Aris (Greece)	gssh-prace.aris.grnet.gr. 195.251.114.114 195.251.114.115	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * 195.251.114.115				
	Leo (Hungary)	login-vlan907.debrecen2.hpc.niif.hu. 193.224.66.200	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * 193.224.66.200				PRACE MD-VPN Realized on GÉANT infrastructure
	PHItagoras (Hungary)	login-vlan907.budapest2.hpc.niif.hu. 193.224.66.196					
	Fionn (Irland)	{ICHECK-Fionn} 193.1.201.17	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * * * 193.1.201.17				
	Marconi (Italy)	gssh-prace.marconi.cineca.it. 130.186.26.8 130.186.26.9	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * 130.186.26.61 130.186.26.9				
	Eagle (Poland)	eagle-prace.man.poznan.pl. 150.254.128.4	Traceroute: 134.94.115.193 134.94.111.38 * * * * * 150.254.128.30 150.254.128.4				
	Zeus (Poland)	prace-int.cyfronet.pl. 150.254.128.65	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * 150.254.128.94 150.254.128.65				
	Supernova (Poland)	prace-bem-int.wcss.pl. 150.254.128.34	Traceroute: 134.94.115.193 134.94.111.38 * * * * * 150.254.128.60 150.254.128.34				
	Aurel (Slovakia)	{CCSAS-Aurel} 147.213.81.5	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * 147.213.81.5				
	LECAD (Slovenia)	{LECAD-LECAD} 193.2.78.225	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * 193.2.78.225				
	Marenostrum (Spain)	gftp.prace.bsc.es. 212.128.224.7	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * * * 212.128.224.7				
	Minotauro (Spain)	gftp.prace.bsc.es. 212.128.224.7					
	Piz_Daint (Switzerland)	{CSCS-Piz_Daint} 148.187.128.41	Traceroute: 134.94.115.193 134.94.111.38 * * * * * * 148.187.128.41				

Figure 3: Overview of the PRACE Network by sites connected via MD-VPN

The Figure 3 describes the network monitoring status of the connected systems as of July, 13th 2020.

The HPC systems of 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.

A test environment for a monitoring solution via perfSONAR has been setup successfully between the partners BSC, CINECA, CSCS, IT4I, KIFU and JUELICH. 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.

The final monitoring solution including perfSONAR client systems at each PRACE partner site, which will get their configuration from a central MadDash configuration server, could not be setup globally. 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.

Preparation for this final monitoring solution will start in the second reporting period of PRACE-6IP.

3.2 Data Services

The service has been regularly checking the state of operation of the data transfer services at the corresponding PRACE sites. A major concern was the earlier announced end of life for the Globus Toolkit as was already mentioned in more detail in the deliverables of previous PRACE IPs. The sites were reminded to upgrade their installations to the Grid Community Toolkit. Other issues included expired server certificates and misconfiguration which have been resolved with the support provided. The overall availability of data transfer services has improved. It has also been noticed that some Tier-0 sites decided to use SSH based GridFTP instead of a GSI based one, to allow users without grid certificates to transfer their data. Furthermore, some sites implemented Globus which is a commercial spin-off of the discontinued Globus Toolkit.

In the meantime, one of the largest users and contributors to the Grid Community Toolkit announced the retirement timeline for GridFTP and GSI dependencies for their software after January 2022. Members of the Grid Community Forum are currently discussing the further support model. The data service leader has already updated the PRACE partners with this information and it will continue to do so in the future.

With the latest draft of the PRACE Service Catalogue, modifications to the data services have been proposed. The major difference is that the GridFTP has been replaced with the secure copy (SCP) tool as a core service for the data the transfer, which is part of the plan for replacing personal grid certificates with other more user-friendly means of authentication. Gsatellite has been removed from the catalogue due to the low interest from user side. The data service leader provided input to the document defining a plan for more user-friendly and forward-looking services for the future.

3.3 Compute Services

Compute services provide interfaces between users and computing capabilities. With the last change of the PRACE Service Catalogue, a major change for compute services has been introduced. After a survey among the partners reflecting the users' requirements for these services it has been agreed to keep the status of the Unicore service to optional, not requiring the installation on new and upgraded systems and make sure that there is a working, well documented and easy to use local batch scheduling available on the systems for the PRACE users as a mandatory requirement.

3.3.1 *Local* batch *scheduling*

Local batch scheduling in PRACE has converged over the past years to two dominant implementations: SLURM and PBS. There are three variants of PBS implemented across the sites, PBS Pro from Altair, PBS Pro Community Edition and Torque (former OpenPBS), where the last two are completely open source while the first is a commercial/licensed product with support from Altair. SLURM is open source as well with the possibility to acquire support either from different system vendors like HPE or ATOS or directly from the main developer SchedMD. At this point, the majority of the sites in PRACE are using SLURM naturally creating a unified environment for the users, where differences reflectmostly the hardware setup of the different systems. The sites which do not use SLURM are providing extensive documentation for the PRACE users to overcome the difference. This also is also the case for the last two legacy systems using LoadLeveler/LSF. As for PBS only five systems are using it, while SLURM is being deployed on the remaining twenty-one systems.

3.3.2 *Unicore*

As stated before, Unicore still remains an optional service and the main focus is to evaluate its benefits stemming from collaboration with other EU projects like FENIX, where it has been deployed. Another direction of the evaluation is the integration with the new proposed PRACE AAI solution based on federated authentication for which Unicore already has a good infrastructure in place, compared to traditional methods of access like SSH/GridFTP and others.

3.4 AAA Services

The AAA activity is responsible for services, which provide Authentication, Authorization and Accounting facilities on the infrastructure. In this reporting period, the AAA was operated in the same way as in the previous implementation phases of PRACE at the same time introducing changes to the existing parts based on the PRACE Service Catalogue update and also looking forward towards a new approach with working title “New PRACE AAI”. The new AAI aims to bring a more user-friendly approach towards the PRACE identity and usage of resources and will be based on a more general federated authentication and authorization infrastructure. The principles used for the design of such an AAI are coming from the previous activity in the AARC Project [3] with the goal to apply the AARC Blueprint Architecture to the PRACE AAA services. Another input into the design and implementation of this new AAI are active collaboration and discussion with other projects like the FENIX and PUHURI to guarantee best practice sharing and seamless interoperability. The work on this topic started already in this reporting period but the full design for all PRACE partners as well as a pilot for selected TIER-0 systems will be delivered in the second period.

As these changes were originally triggered by the request to simplify users’ identity management and access to PRACE services, the changes were agreed and manifested into the PRACE Service Catalogue affecting the components of the current PRACE AAA, namely PKI, User administration, Interactive access and Accounting services.

3.4.1 Public Key Infrastructure - PKI

Several PRACE services rely on X.509 certificates [4] for authentication and authorisation. All of these services are newly categorised as optional due to several constraints both on the users' as on the sites' sides. However, for services which are in place and are used, a working and reliable PKI needs to be available. The X.509 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) [5], or by one of the two other organizations, the TAGPMA and the APGridPMA, all three federated in the IGTF [6]. 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 virtualised 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 authorisation information needed by PRACE services and is used to retrieve information about users and their projects. Authorisation 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 4. 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.

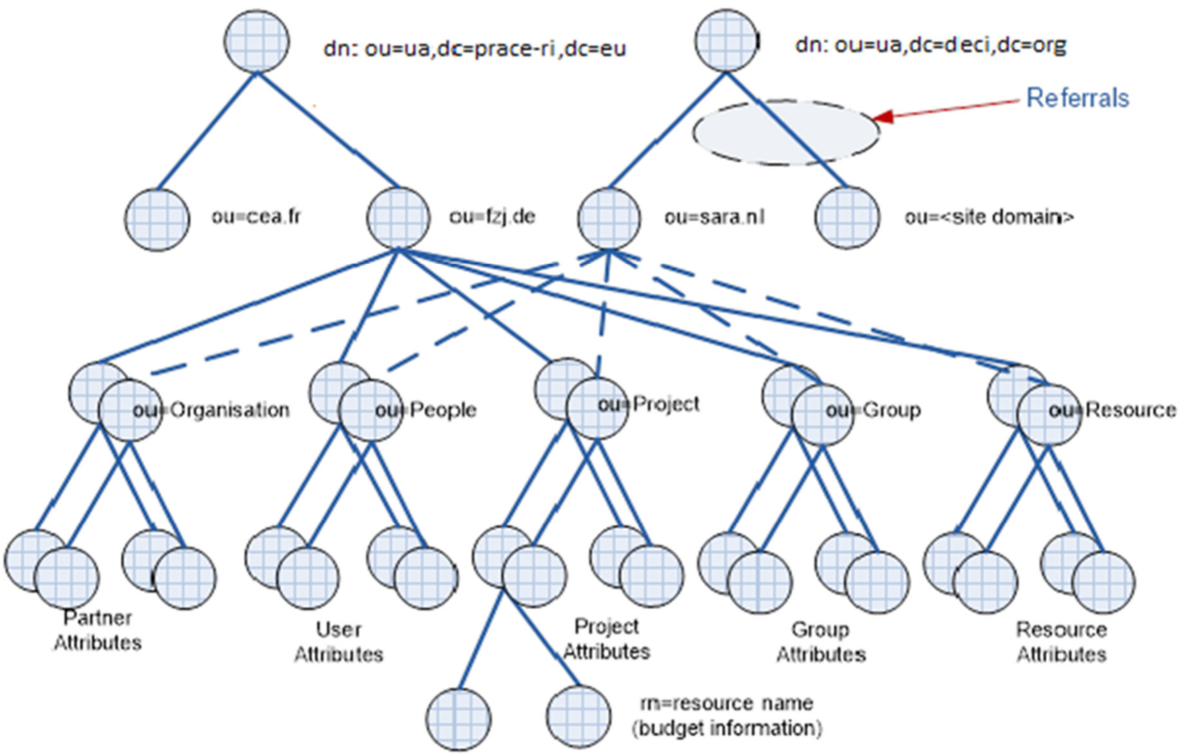

Figure 4: PRACE LDAP directory tree

All Tier-0 sites manage their own branch in the LDAP repository. 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 Directory Information Tree (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 is the Home site which creates and updates the account information of the user. In general, the Home site 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 virtualised and can be migrated if needed. 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 take care of it entirely by a manual operation. This absence of any specific definition of a required interoperability interface between the PRACE LDAP and a site-specific user administration has in practice 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. New Tier-1 LDAP branches have been added for CCSAS (Slovakia), UC-LCA (Portugal), and the University of Minho (Portugal). There were also branches removed, namely for IUCC (Israel) and IPB (Serbia). These changes caused an update of the internal tools which are used at the partner sites and which are distilling the information from LDAP, transform to different output formats required by the services and check for consistency and format of the LDAP entries.

No changes have been made to the LDAP schema.

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. In the last change of the PRACE Service Catalogue it was agreed that all the X.509 based access methods will be considered as optional and that a standard SSH access is mandatory for all PRACE sites' users. This created a requirement for the new PRACE AAI to enable federated access for users over SSH.

Not all sites support GSI-SSH based access directly from the Internet. Therefore, two partners, IT4I (Czech republic) and SURFsara (Netherlands), 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. Based on the recent changes SURFsara will cease this service in the next reporting period.

3.4.4 Accounting Services

Another result of the recent changes is that users must get accounting information locally at the sites at which they consume resources. Each site can optionally run a standard PRACE accounting web service that can be accessed easily using GSI (X.509 certificate) authentication. In addition, PRACE users and staff can get accounting information from several sites in a uniform way with the DART tool [7]. 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-6IP 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 and no services have been on risk in 2019. Unfortunately this changed in 2020.

On 5th February, the systems Galileo and Marconi at CINECA had to be taken offline because of compromised passwords and root access. Investigations showed that on 4th January 2020 the PRACE system Galileo at CINECA (Italy) got compromised. The attacker accessed the system with a stolen user password and managed to gain root privileges and recompile and replace sshd and ssh executables. On 18th January 2020 the PRACE system Marconi at CINECA also got compromised in a similar fashion. Investigations showed also that already in September 2019 first hack attempts had started. Users whose password has been stolen had a connection to and from login.hpc.icm.edu.pl, HPC centre of ICM University of Warsaw (Poland), in common. PRACE CSIRT have contacted them and they confirmed that their system has been compromised in a similar way, so it is likely that the user passwords were stolen on these machines already. In further investigations it turned out that further user/password combinations had been compromised, especially of systems at SURFsara (Netherlands), CSCS (Switzerland) and LRZ (Germany). All systems have been fixed and could go online again. Investigation on a second incident started on 11th May 2020 when the administrator of Juwels the PRACE HPC system at Juelich (Germany), got informed by administrators of systems at Baden Wuerttemberg (Germany) of backdoors in their systems. This time the incident spread over several systems Europe-wide including eight PRACE systems. The systems have been analysed and user access was cancelled. Investigations on this incident took several weeks before the systems could go online again. Until now the intention of the hackers could not be identified. It is not clear if industrial or national espionage, bitcoin mining or something else has been the intention. All potential software leaks have been closed. Unneeded services have been closed and systems hardened. During both incidents the PRACE CSIRT managed the information flow between the involved partners and exchanged information with other Europe-wide CSIRTs e.g. from EGI, EOSC-Hub, EUDAT, DFN, GÉANT and the GAUSS alliance in Germany.

One further task is the discussion of operational security within PRACE for potential future PRACE services. Within the reporting period no new services have gone into production.

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 still focussed on preparing for becoming accredited. The accreditation could not be started yet due to time constraints (the aforementioned incident described above) as well as because of the complexity of harmonising 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 harmonised 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.

3.6 User Services

The 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. Several changes were agreed by the partners to reflect the changes introduced by the PRACE Service Catalogue and by the transition to the new central WWW site.

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.

There were no changes towards this part, partners are free to use different implementation of the software modules which have to be syntax compatible with the original version allowing the users to use the same commands/working environment.

3.6.2 Trouble Ticket System

The centralised Helpdesk [8] was deployed as part of the PRACE-1IP project. It is an important tool for the PRACE project staff to communicate among sites about problems. Although the Helpdesk is also available for users the last years have shown that most of the users' issues are site specific and thus there is not much need to use the centralised one, but rather contact the local support staff using the site helpdesk directly.

TTS is based on the Request Tracker (RT) issue tracking system [9]. Since its deployment in PRACE-1IP, this service is hosted and operated at CINECA.

New site specific queues were created in PRACE TTS, namely CCSAS (Slovakia) and CESGA.

3.6.3 Helper scripts

The 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.

In this reporting period 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.

The *prace_service.config* file was extended with operational information for the network monitoring. Currently a change is made to extend the both the script and the config file format to distinguish and enable the SSH/SCP access differently from the GSI-SSH/GridFTP access as a result of the changes introduced by the new PRACE Service Catalogue.

3.6.4 PRACE User Documentation

With the transition to the new PRACE web site a change towards the documentation was agreed. The new approach tries to keep the new web site accessible and lightweight in comparison with the old design and not to be cluttered with all the possible content, especially if it concerns documentation of the systems already available at the partners' websites. The new approach aims to make sure that at the each partner site a specific PRACE user oriented documentation is available at least in English language, updated to current state and properly linked to the central web site. As already said, this makes the central site more lightweight and easy to navigate, does not create duplicates of the same information in just a different format/layout and makes sure it is updated or rather in full sync with the original one at each site.

3.7 Monitoring Services

Operating and supervising infrastructure is distributed among several project partners is a very challenging task. The goal of the monitoring service activity is to provide insight into current and past state of services hosted and delivered by 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. Infrastructure monitoring capabilities are provided by the solution initially developed and deployed during PRACE-5IP and extended over time. The monitoring system collects the health status data for hosts and services, provides alerting capabilities and collects all the data necessary to estimate service availability and quality for reporting and analytic purposes.

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

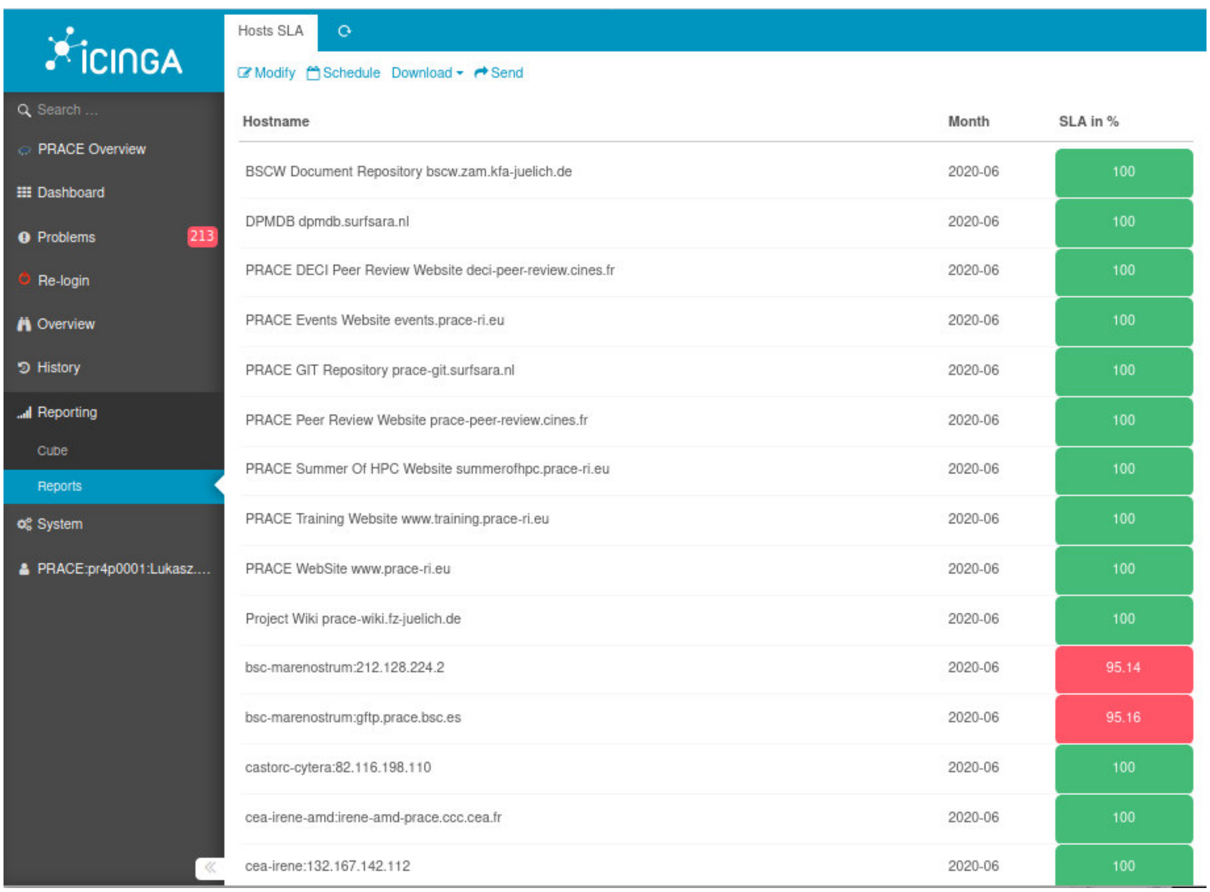

Figure 5: PRACE monitoring user interface

The number of hosts and services being subject for monitoring activities change over time with the machine commissioning/decommissioning cycles natural for this kind of environment. Configuration generators were created to keep monitoring configuration up-to-date and automatically adapt to infrastructure changes executed through RFCs and reflected in prace_service config file.

During the reporting period, the number of monitored entities has reduced due to number of machines being decommissioned whereas their replacements still being installed and enabled for PRACE. Currently 925 objects (services and hosts) are being actively monitored on given intervals. Below is a brief summary of changes regarding the computing sites.

Sites/machines removed from the system during M1-M12:

- idris
- lrz-supermuc
- lrz-supermuc-fat
- seged
- ncsa
- csc

New sites added to the monitoring system

- cea-irene
- cea-irene-amd
- cesga
- lrz-supermuc-ng

3.7.1 Task Activites during reporting period

Within the PRACE-6IP project – the monitoring system is in a stable production phase and most tasks related to the system were of maintenance nature, however, development of changes required by the project was being carried out and concluded with the deployment.

Maintenance and support activities performed regularly within the task include:

- Monitoring for operating system level vulnerabilities and installing necessary updates and mitigations
- Watching mailing lists and web for icinga2 and web interface updates and security issues
- Fixing bugs reported via TTS system related to monitoring service functioning
- 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, more significant changes were made to the system during reporting period:

- Upgrade of server operating system version to next LTS release in order to ensure access to current security updates and patches
- Completed rolling upgrade of icinga2 and web interface by several major versions to ensure newest functionality required by development of TTS integration and needed to solve object dependency issues
- Reorganization of object dependency and notification system configuration to avoid unnecessary actions and alerts as a preparation for TTS integration
- Rearrangement of e-mail notification mechanism

In addition to above changes and activities the service was focusing on the following development activities:

- Integration with TTS system - a functionality important from the OoD (Operator on Duty) activity perspective. This module allows for automatic TTS ticket creation and solution in case of change of service status in the monitoring.
- Redesigning the system to address the changes resulting from a PRACE Service Catalogue evolution
- Extending the system by adding new service checks to the monitoring

Among all above activities – the PRACE Service Catalogue imposed change is by far most interfering in terms of monitoring system configuration.

The PRACE Service Catalogue update consists of several changes to the list of mandatory and optional services. One fundamental change - critical from the monitoring system point of view - is transition of GSI-SSH service from mandatory to the optional requirement. The functionality of GSI-SSH will be covered by the new service - standard SSH protocol, meaning SSH without the GSI (X.509) legacy that turned out to be an obstacle for some less experienced users. From the monitoring perspective where GSI-SSH plays a significant role as major protocol for executing remote service checks, this change began with the process of adapting to its replacement.

Work has been done to allow for using both protocols in the transition period. It was successfully tested and confirmed that the functionality based on the new protocol is equivalent to its predecessor. The concept was confirmed by switching two sites in Poland to a new way of executing tests. The change did not affect already collected historical data. In terms of site-to-site (s2s) communication checks, a new test based on SSH and public key authentication has been developed and tested. For data transfers there is no need to create s2s GridFTP check equivalent as both SSH based data transfer protocols (SCP and SFTP) are using the SSH connection.

Despite required efforts on the monitoring service due to PRACE Service Catalogue update, migration from GSI-SSH to its less complex equivalent will be beneficial for users and administrators. From the monitoring perspective we predict significant improvements in the monitored services availability, as SSH compared to GSI-SSH requires drastically less maintenance and is less error prone.

Future development will focus on migrating all the sites from legacy GSI-SSH to SSH based executors. It is necessary to redesign prace_service config file to reflect changes in the catalogue in a way that allows better integration with the monitoring system.

Learning from recent HPC security incidents across the world, we are investigating more secure ways of handling and storing authentication keys (hardware based). In case of a positive outcome, the monitoring system will be modified in order to improve security levels.

3.8 Generic Services

This section describes the actions done during the reporting period within the Generic Services subtask. In general, all services that need an operational basis and a centralised distribution for the PRACE project (or a part of it) can be assumed as Generic Services.

The goal of this task is the provisioning of 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 liaison / consultant to consult with

- PRACE aisbl.
- Other WPs
- External providers

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

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

- Questions of another WP related to the operation of new or existing services
  - Finalised integration between Training, Events portals and PRACE Web site;
  - Negotiated and established the new training portal systems with WP4 (new Wordpress and ePrints instances) with WP4
  - Analysing and following implementation of introducing Single Sign On mechanism to Events Portal
  - Supporting piloting new Training Platform inclusion
  - Supporting implementation of introducing Events Portal to GÉANT edugain federation
  - Planning of Events Portal upgrade
  - Supporting git based configuration and document versioning system to go into production
- Questions of PRACE aisbl when negotiating with service providers offering service for PRACE
  - Consulting with PRACE aisbl IT service hosts upon request
  - Helping managing domain portfolio
  - Helping with managing certificates portfolio
  - 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, and in connection with HPC in Europe portal)
  - Helping WP3 with the usage of the hosting platform for the new prace-ri.eu website
  - Support for migration to new hosting platform, discussion and adjustment of technical requirements with provider

  - Detect problems of the new hosting platform, suggest possible resolutions
  - Support migration and hosting of the previous website containing data not available on the current website
  - Helping with GDPR update of new Training and PRACE web portal.
- Helping internal workflows of using a shared PRACE-owned resource
  - Handling *.prace-ri.eu wildcard certificate, helping with new certificate requests and renewal
  - Handling prace-ri.eu domain: subdomain requests, helping with resolving domain issues related to changes

One of the main tasks during this period was supporting the new prace-ri.eu and training.prace-ri.eu portal to finally go into production in early 2020. The task included introducing a new host for prace-ri.eu, requiring domain and certificate changes along with setting up the hosting and migrating the previous version of the site, defining requirements with developers and fixing initial hosting issues.

Two new services including new Wordpress based Training Portal and ePrints based material repository were put into production to support online learning, becoming especially important during COVID-19.

A pilot including a digital learning platform had been launched, requiring joining edugain federation and deploying integration of Events Portal to provide an uniform entry point and Single Sign On methodology for trainees to be used to participate video and hands-on sessions.

Prace-ri.eu wildcard certificates were renewed, requiring all subdomain certificates to be renewed as a duplicate containing eighteen certificates for twenty-two subdomains requiring coordination between eight different hosting providers to avoid downtime caused by invalid certificates.

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 25 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.

A new version of the PRACE Service Catalogue was approved by PRACE BoD to update the status of services related to the changes imposed by the design of the new user-friendly services.

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 PRACE Service Catalogue and was able to handle the security incidents that happened in 2020. 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 ecosystem as a service for science.