E-Infrastructures
H2020-EINFRA-2014-2015

EINFRA-4-2014: Pan-European High Performance Computing Infrastructure and Services

PRACE-4IP

PRACE Fourth Implementation Phase Project

Grant Agreement Number: EINFRA-653838

D6.1
First Annual Report on the PRACE Operational Services

Final

Version: 1.0
Author(s): Giovanni Erbacci, CINECA
Miroslaw Kupczyk, PSNC
Date: 22.04.2016

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: EINFRA-653838**	
	Project Title: PRACE Fourth 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:** 30 / April / 2016
		Actual Date of Delivery: 30 / April / 2016
	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 2991/844/EC.

Document Control Sheet

Document	**Title: First Annual 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 2010	
	File(s): D6.1.docx	
Authorship	**Written by:**	Giovanni Erbacci, CINECA Miroslaw Kupczyk, PSNC
	Contributors:	Martin Bidner, HLRS Thomas Bonisch, HLRS Xavier Delaruelle, CEA Lukasz Flis, CYFRONET Zoltan Kiss, NIIF Felip Moll, BSC Ralph Niederberger, FZJ Cristiano Padrin, CINECA Elda Rossi, CINECA Jules Wolfrat, SURFSara Oscar Yerpes, BSC
	Reviewed by:	Ioannis Liabotis, GRNET Florian Berberich, FZJ
	Approved by:	MB/TB

Document Status Sheet

Version	Date	Status	Comments
0.1	01/March 2016/	Draft	Structure and first elements of the document
0.2	15/March/2016	Draft	First draft of operational services
0.3	17/March/2016	Draft	Agreed the final structure of the document
0.4	31/March/2016	Draft	Added draft on user services
0.5	05/April/2016	Draft	Draft final version
0.6	07/April/2016	Draft	Final polishing and editing
0.7	08/April/2016	Draft	Draft version ready for internal review process
1.0	22/April/2016	Final version	Final version ready for EC submission

Document Keywords

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

Table of Contents

Project and Deliverable Information Sheet i
Document Control Sheet i
Document Status Sheet ii
Document Keywords iii
Table of Contents iv
List of Figures v
List of Tables v
References and Applicable Documents v
List of Acronyms and Abbreviations vi
List of Project Partner Acronyms viii
Executive Summary 1
1 Introduction 2
2 PRACE HPC Ecosystem: Tier-0 and Tier-1 sites, system upgrades and new systems 3
2.1 Maintaining the service 3
2.2 Status of Tier-0 & Tier-1 sites 3
2.2.1 *On-duty Activity* *3*
2.2.2 *Production systems* *4*
2.3 System Upgrades 7
2.3.1 *Overview of System Upgrades in 2015-2016 period* *7*
2.3.2 *Details of upgraded sites* *7*
2.4 New Tier-0/Tier-1 sites & systems 9
2.4.1 *Overview of new Tier-0/Tier-1 sites & systems* *9*
2.4.2 *Details of new Tier-0/Tier-1 sites* *9*
3 Operational Services 10
3.1 Network Services 10
3.2 Data Services 12
3.2.1 *GridFTP* *12*
3.2.2 *gtransfer* *14*
3.2.3 *MC-GPFS* *15*
3.3 Compute Services 15
3.4 AAA Services and Operational security 18
3.4.1 *Public Key Infrastructure - PKI* *18*
3.4.2 *User Administration* *18*
3.4.3 *Interactive Access* *20*
3.4.4 *Accounting Services* *20*
3.5 Security Forum 20
3.5.1 *Security Policies and Procedures* *20*
3.5.2 *Risk Reviews* *21*
3.5.3 *Operational Security* *21*
3.6 User Services 21
3.6.1 *PRACE Common Production Environment* *21*
3.6.2 *Trouble Ticket System* *22*
3.6.3 *Helper Scripts* *22*
3.6.4 *PRACE User Documentation* *22*

3.7 Monitoring Services......24
3.7.1 *Monitoring Middleware Change* 24
3.7.2 *The Transition* 25
3.7.3 *The Status* 26
3.8 Generic Services 27
3.8.1 *Catalogue* 27
3.8.2 *Roles and Actions* 28
4 Conclusions 30

List of Figures

Figure 1: *PRACE network topology* 12
Figure 2: *UNICORE Architecture* 17
Figure 3: *PRACE LDAP directory tree* 19
Figure 4: *Publication process of the PRACE User Documentation* 23
Figure 5: *Icinga2 site to site tests matrix* 26

List of Tables

Table 1: *PRACE partners involved in the on duty activity* 4
Table 2: *PRACE Tier-0 systems* 5
Table 3: *PRACE Tier-1 systems* 6
Table 4: *System upgrades from 02/2016 to 03/2016* 7
Table 5: *GridFTP deployment status on Tier-0 systems in March 2016* 13
Table 6: *GridFTP deployment status on Tier-1 systems in March 2016* 13
Table 7: *Local Batch Systems in PRACE* 16
Table 8: *UNICORE Software components* 17
Table 9: *Current status of Icinga integration* 27
Table 10: *Internal generic services* 28
Table 11: *External generic services* 28

References and Applicable Documents

[1] http://www.prace-project.eu

[2] PRACE-1IP Deliverable D6.3, Second Annual Report on the Technical Operation and Evolution: http://prace-ri.eu/IMG/pdf/d6.3_1ip.pdf

[3] PRACE-3IP Deliverable D6.1.1, First Annual Operation Report, PRACE, June 2013

[4] Beskow System - https://www.pdc.kth.se/resources/computers/beskow

[5] Hazel Hen System - https://wickie.hlrs.de/platforms/index.php/Cray_XC40

[6] ARCHER - http://www.archer.ac.uk/about-archer/news-events/news/2014-10-31-Phase2_upgrade.php

[7] Salomon System - https://docs.it4i.cz/salomon

[8] Leo System - https://wiki.niif.hu/HPC_en#NIIF_HPC_resources

[9] Galileo System - http://www.hpc.cineca.it/hardware/galileo

[10] RFC 5280, Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile: https://tools.ietf.org/html/rfc5280

[11] EUGridPMA: https://www.eugridpma.org/
[12] IGTF: http://www.gridpma.org/
[13] Globus GSI-OpenSSH: http://www.globus.org/toolkit/security/gsiopenssh
[14] GSISSH-Term: https://www.lrz.de/services/compute/grid_en/software_en/gsisshterm_en/
[15] DART: http://prace-ri.eu/Accounting-Report-Tool
[16] A Trust Framework for Security Collaboration among Infrastructures http://pos.sissa.it/archive/conferences/179/011/ISGC%202013_011.pdf
[17] PRACE Helpdesk: https://tts.prace-ri.eu/
[18] Best Practical Request Tracker: https://www.bestpractical.com/rt-and-rtir
[19] PRACE-RI Website: http://www.prace-ri.eu/
[20] PRACE SVN repository: https://prace.osd.surfsara.nl/svn/
[21] MarkDown: https://en.wikipedia.org/wiki/Markdown
[22] MarkDown's syntax: http://daringfireball.net/projects/markdown/syntax
[23] Pandoc: http://pandoc.org/
[24] PRACE Software Stack document.
[25] INCA Monitoring system, http://inca.sdsc.edu/
[26] ARGO Monitoring System, http://argoeu.github.io/
[27] Nagios, https://www.nagios.com/
[28] Zabbix, www.zabbix.org
[29] ICINGA, https://www.icinga.org/
[30] Opsview, https://www.opsview.com/

List of Acronyms and Abbreviations

AAA	Authorization, Authentication, Accounting
aisbl	Association International Sans But Lucratif (legal form of the PRACE-RI)
APGridPMA	The Asian Pacific Grid Policy Management Authority
BCO	Benchmark Code Owner
CA	Certificate Authority
CoE	Center of Excellence
CP/CPS	Certificate Policy/Certification Practice Statement
CPU	Central Processing Unit
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
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
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.
HMM	Hidden Markov Model
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
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
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.
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
PRIDE	Project Information and Dissemination Event
RI	Research Infrastructure
RT	Request Tracker
SCI	Security for Collaborating Infrastructures
SDSC	San Diego Supercomputer Center

SSH	Secure Shell
SVN	SubVersioN: software versioning and revision system
TAGPMA	The Americas Grid PMA
TB	Technical Board (group of Work Package leaders)
TB	Tera (= 240 ~ 1012) 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 (= 1012) 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
UNICORE	Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.

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 (3rd Party to GENCI)
CNRS	Centre National de la Recherche Scientifique, France (3rd 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.
GENCI	Grand Equipement National de Calcul Intensiv, France
GRNET	Greek Research and Technology Network, Greece
INRIA	Institut National de Recherche en Informatique et Automatique, France (3rd Party to GENCI)
IST	Instituto Superior Técnico, Portugal (3rd Party to UC-LCA)

IUCC	INTER UNIVERSITY COMPUTATION CENTRE, Israel
JKU	Institut fuer Graphische und Parallele Datenverarbeitung der Johannes Kepler Universitaet Linz, Austria
JUELICH	Forschungszentrum Juelich GmbH, Germany
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
NIIF	National Information Infrastructure Development Institute, Hungary
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, Austria
RZG	Max Planck Gesellschaft zur Förderung der Wissenschaften e.V., Germany (3rd 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, Labotatório de Computação Avançada, Portugal
UCPH	Københavns Universitet, Denmark
UHEM	Istanbul Technical University, Ayazaga Campus, Turkey
UiO	University of Oslo, Norway (3rd Party to SIGMA)
ULFME	UNIVERZA V LJUBLJANI, Slovenia
UmU	Umea University, Sweden (3rd Party to SNIC)
UnivEvora	Universidade de Évora, Portugal (3rd Party to UC-LCA)
UPC	Universitat Politècnica de Catalunya, Spain (3rd Party to BSC)
UPM/CeSViMa	Madrid Supercomputing and Visualization Center, Spain (3rd Party to BSC)
USTUTT-HLRS	Universitaet Stuttgart – HLRS, Germany (3rd Party to GCS)
VSB-TUO	VYSOKA SKOLA BANSKA - TECHNICKA UNIVERZITA OSTRAVA, Czech Republic
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 (February 2015 - March 2016) to operate and coordinate the common PRACE Operational services, foreseen by Task 6.1 of WP6 in PRACE-4IP project. The operation of the PRACE distributed HPC infrastructure involves the coordination of a set of services which integrate the Tier-0 systems and a number of national Tier-1 systems, providing services for Tier-0, in a "single" pan-european HPC infrastructure.

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

Six Tier-0 systems are operational in the first year of the PRACE-4IP project period:

- JUQUEEN at GCS@FZJ;
- CURIE at GENCI@CEA;
- HAZELHEN at GCS@HLRS;
- SuperMUC at GCS@LRZ;
- FERMI at CINECA;
- MareNostrum at BSC

Furthermore, operational support has been provided to 26 national Tier-1 systems which 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).

A new version of the PRACE Service Catalogue, which describes the PRACE common services, has been approved from the PRACE BoD in March 2015, just after the start of PRACE-4IP, and is guiding the operational activity. The Service Catalogue represents a living document and, as a result of the successful evaluation in the first year of PRACE 4IP, it will be revised during the second year of PRACE-4IP. In the process towards establishing a PRACE Quality of Service and quality control the work on PRACE Operational Key Performance Indicators is being addressed and the measurement of the KPIs and their evaluation will be implemented during the second period of the project.

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

The Security Forum, responsible for all security related activities, is also coordinated by Task 6.1, with periodic teleconferences to monitor the infrastructure and prevent possible incidents which could cause vulnerability on the PRACE RI.

1 Introduction

This report describes the activity done in Task 6.1”*Operation and coordination of the comprehensive common PRACE operational services*” for the first project year of WP6” *Operational Services for the HPC Eco-system*“ in PRACE-4IP. This task is responsible for the operations 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 SHAPE activity toward SMEs, the prototyping and assessment of new operational services investigated in Tasks 6.2 and 6.3 of PRACE-4IP WP6, the testing and utilization of specific architectures and technologies which are only available in specific countries.
The operation and coordination of the common *PRACE* operational services provided in Task 6.1 have well established management procedures and organisation as set up already since PRACE-1IP, [2], [3]. The task further continues the implementation of the roadmap to a professional service level of sustainable services with a defined quality of service.

In the reporting period (February 2015 - March 2016) the operation of the common PRACE operational services has been constantly coordinated and monitored by biweekly teleconferences attended by all the partners involved in the Task 6.1 activity. Furthermore, two successful all-hands meetings have been organised, in September 2015 in Bologna (Italy) and in February 2016 in Edinburgh (UK). These meetings were used to discuss the status, to plan activities for the coming period and to define the teams for the different activities. Both meetings had around 40 attendees.

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 (dedicated network provided by GEANT connecting Tier-0 and major Tier-1 centres);
- data services;
- compute services;
- Authorization, Authentication and Accounting;
- operational security;
- user services;
- monitoring services for operations;
- generic services.

Finally Section 4 presents some plans for the next reporting period and draws some conclusions.

This deliverable is intended to support the PRACE partners to better identify and integrate the PRACE pan-European HPC infrastructure. The intended audience of this deliverable is primarily the PRACE project partners where the cooperation and interactions between different WPs can require access to the operational service of the HPC Infrastructure. In addition to the previous audience there are the members of the PRACE RI, the users and the stakeholders who consider access to the PRACE HPC resources an important instrument to enhance their competitiveness in science and technology.

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

The first part of this section will present the changes implemented during the current reporting period (February 2015 – April 2016) that concern the status of Tier-0 sites, and the Tier-1 national sites providing Tier-1 for Tier-0 services. In the second part, the chronology and the status of the performed system upgrades is being detailed. Finally, the new systems that have been integrated into the PRACE ecosystem are being presented.

2.1 Maintaining the service

In the current 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 periods, so the work from the point of view of managing the procedures and operability resulted simplified and well tested as was the continuation of the activity already in progress since the previous PRACE 3IP Project.

Nevertheless some minor changes to the operational procedures have been made and have been communicated to partners accordingly. Mainly these changes modified the way that data is organized in the PRACE Operations Wiki in order to improve productivity and reduce complexity. Also, an effort to remove or update deprecated information has been assessed.

As a consequence all participants can introduce, find and update information of their PRACE HPC systems more easily. The result is that they gave us good feedback on the procedures defined for keeping track of current, new and upgrades of systems.

One more important change has been the simplification of the operator on-duty and help desk on-duty activities, merged in a single more efficient action. Moreover, the Trouble Ticked System (TTS) environment has been updated to the most recent version.

2.2 Status of Tier-0 & Tier-1 sites

The Tier-0 and Tier-1 systems constitute a HPC eco-system offering high level services to the European computational community. At present 6 Tier-0 systems and 26 Tier-1 systems are in production. Their status is monitored and operational quality is assured by employing a specific regular activity provided day by day by members of Task 6.1 of PRACE-4IP WP6. This activity, called the on-duty activity, is described in the following passage.

2.2.1 *On-duty Activity*

The on-duty activity has been simplified. The simplification process was first discussed during the WP6 Face to Face meeting in Bologna in September 2015, and consists in concentrating the two previous on-duty activities HdD (Help Desk on duty) and OoD (Operator on duty) into a single one. Then the agreed process was formalised with RFC Nr. 210, approved by all the partners and implemented in December 2015. This change was proposed because usually the HdD and the OoD are operated by the same members of staff. Moreover the topics/incidents reported to this service are mainly related to operational issues and activities needed to maintain the distributed infrastructure in good shape. Requests from users are very rare and are normally redirected to the local Help Desk of each individual site.

With the introduction of this change, only one report was being produced every week, reporting the status of the infrastructure, all occurred problems, the status of the core services, and so on.

The related documentation on the wiki site has been updated, and a new report template has been produced.
Starting from 1st February 2015, a schedule has been defined for 20 PRACE partners involved in the on duty activity. Each of them is in charge of monitoring the infrastructure, and reporting issues using the Trouble Ticketing tool (TTS). The 20 partners involved in the schedule are reported in Table 1 below.

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

Table 1: *PRACE partners involved in the on duty activity*

A given partner is involved in the shift one week out of 20, i.e. every 4-5 months. The weekly reports of the OoD activity are stored on the PRACE Operation wiki.

The TTS tool used by the OoD 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 virtual machine where it is maintained since its deployment during the PRACE-3IP project. During the PRACE-4IP project, the TTS tool has been updated from version 3.8 to 4.2, mainly due to some security issues.

2.2.2 *Production systems*

At present (April 2016) the Tier-0 ecosystem is made up of 9 systems, distributed in 6 sites, operated by 4 partners, in the four core countries (France, Germany, Italy and Spain) as reported in Table 2.

The peak performance ranges from 100 TFlop/s up to more than 7 PFlop/s for the new Cray machine in Germany. Only few Tier-0 systems are accelerated, Curie/Hybrid with Nvidia cards and BSC/MareNostrum with Intel Phi.

All the major systems are ranked in the Top500 (November 2015), from position 8 to 93. Five of them are in the first 50 positions.

The dominant vendor is IBM/Lenovo with 5 systems.

Partner	Country	Tier-0	Architecture CPU	Rpeak (TFlop/s) Top500#(nov15)
GCS-JUELICH	Germany	Juqueen	IBM BlueGene/Q IBM PowerPC A2; 16-core; 1.60GHz	5802.0 11
GCS-LRZ	Germany	SuperMUC	IBM iDataPlex Intel Sandy Bridge-EP; 8-core; 2.7GHz	3185.0 23
GCS-LRZ	Germany	SuperMUC phase 2	Lenovo NeXtScale Intel Haswell-EP 14 core; 2.6GHz	3580.0 24
GCS-HLRS	Germany	Hazelhen	Cray XC40 ntel Haswell E5 2680v3; 12-core; 2.5GHz	7420.0 8
GENCI	France	Curie (Fat)	Bull Bullx BCS Intel Nehalem-EX; 8-core; 2.27GHz	104.4 --
GENCI	France	Curie (Hybrid)	Bull Bullx B505 Intel Westmere-EP; 4-core; 2.67GHz Nvidia/M2090	198.1 --
GENCI	France	Curie (Thin)	Bull Bullx B510 Intel Sandy Bridge-EP; 8-core; 2.70GHz	1741.8 53
BSC	Spain	Mare Nostrum	IBM iDataPlex Intel Sandy Bridge-EP; 8-core; 2.7GHz Intel Phi	1017.0 93
CINECA	Italy	FERMI	IBM BlueGene/Q IBM PowerPC A2; 16-core; 1.60GHz	2097.1 37

Table 2: *PRACE Tier-0 systems*

As far as the Tier-1 ecosystem is concerned, it is made up of 26 systems operating Tier-1 for Tier-0 services. These Tier-1 systems are distributed in 19 different PRACE sites, operated by 17 partners, in 17 different European Countries. Table 3 presents the list of the Tier-1 systems.

The peak performance ranges from very small system partitions (<10 TF/lop/s) up to large systems in excess of 2 PFlop/s. Nine systems deliver more than 1 PFlop/s.

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

Partner	Country	Tier-1	Architecture CPU	Rpeak (TFlop/s)
GCS / RZG	Germany	Hydra	IBM iDataPlex Intel Sandy Bridge-EP; 8core; 2.6GHz + Intel Ivy Bridge; 10-core; 2.8GHz, Nvidia GPU, Intel PHI	1501.7
GENCI	France	Turing	IBM BlueGene/Q IBM PowerPC A2; 16-core; 1.60GHz	1258.0
EPCC	UK	Archer BlueJoule	Cray XC30 Intel Ivy Bridge; 12-core; 2.7GHz IBM BlueGene/Q IBM PowerPC A2; 16-core; 1.60GHz	2550.5 1258.3
BSC	Spain	Minotauro	Bull Bullx B505 Intel Westmere-EP; 6-core; 2.53GHz, Nvidia/M2090	185.8
CSC	Finland	Sisu	Cray XC40 Intel Haswell; 12-core; 2.60GHz	1688.0
ETH	Switzerland	Rosa	Cray XE6 AMD Interlagos; 16-core; 2.1GHz	400.0
SURF sara	Netherlands	Cartesius1 Cartesius 2	Bull Bullx B720/B710 Intel Haswell; 12-core; 2.6GHz + Intel Ivy Bridge; 12-core; 2.4GHz+ Intel Sandy Bridge; 8-core; 2.7GHz Bull Bullx B515 Intel Ivy Bridge 8 core; 2.5GHz Nvidia K40	1349.0 210.0
SNIC	Sweden	Beskow	Cray XC40, Intel Haswell; 32-core; 2.3GHz	1973.0
CINECA	Italy	Galileo	IBM NeXtScale, Intel Haswell; 8-core; 2.40GHz Nvidia K80, Intel PHI	1000.0
PSNC	Poland	Zeus BigMem Zeus GPGPU Cane Chimera Supernova	HP BL685c G7 AMD Interlagos; 16-core; 2.3GHz HP SL390s Intel Westmere-EP; 6-core; 2.45GHz, Nvidia/M2090 SGI Rackable C1103-G15, AMD Interlagos; 12-core; 2.40GHz, Nvidia M2050 SGI UV 1000 Intel Westmere-EX; 8-core; 2.67GHz HP Cluster Platform 3000 BL2x220 Intel Westmere-EP; 6-core; 2.67GHz	61.2 136.8 224.3 21.8 30.0
SIGMA	Norway	Abel	MEGWARE MiriQuid Intel Sandy Bridge-EP; 8-core; 2.6GHz	178.6
NUI Galway	Ireland	Fionn	SGI ICE X Intel Ivy Bridge; 12-core; 2.4GHz	147.5
UHEM	Turkey	Karadeniz	HP Proliant BL460 Intel Nehalem-EP; 4-core; 2.67GHz	2.5
CaSToRC	Cyprus	CyTera	IBM IDataPlex Intel Westmere; 12-core; 2.67GHz, Nvidia M2070	35.0
NCSA	Bulgaria	EaEcnis	IBM BlueGene/P IBM PowerPC 450; 4-core; 850MHz	27.8
IT4I-VSB	Czech Rep	Anselm Solomon	Bull Bullx B510/B515, Intel Sandy Bridge-EP; 8-core; 2.4GHz SGI ICE-X Intel Xeon E5-2680v3; 12-core; 2.5GHz, Intel PHI	66.0 2000.0
NIIF	Hungary	NIIFI SC Seged Leo	HP Cluster Platform 4000SL AMD Magny-Cours; 12-core; 2.2GHz HP Cluster Platform 4000SL AMD Magny-Cours; 12-core; 2.2GHz, Nvidia M2070 HP SL250s Intel Xeon E5-2650 v2; 2.60GHz, Nvidia K20, K 40	5.48 14.0 254.0

Table 3: *PRACE Tier-1 systems*

2.3 System Upgrades

System upgrades corresponds to the activity of keeping track of all the improvements done to the current computing resources that are integrated into the PRACE ecosystem, being them Tier-0 or Tier-1 systems.

In the current period, February 2015 – March 2016, this activity has been included under WP6 subtask 6.1.2 and lead by Felip Moll from Barcelona Supercomputer Centre with the co-leading of Javier Bartolomé from BSC as well.

2.3.1 *Overview of System Upgrades in 2015-2016 period*

A total of five sites have started upgrades on their respective systems and all five have finally completed the process. One of these sites was a Tier-0 site while the others were Tier-1 sites.

At the time of this report, March 2016, there is not on-going system upgrade activity.

In some cases, sites have decided not to include some services in their service catalogue due to a lack of physical resources (like the connection to GEANT network), to unavailability of operational services (like monitoring services), or to some other decision (many sites stated they do not provide Unicore services).

All the five sites have completed the upgrades following the standard procedures defined internally in the PRACE Operations wiki, with no major issues, providing a good feedback of the process to the subtask leaders.

In Table 4 we show a brief overview of the upgraded systems and the completion dates.

System Name	Site	Tier	Status
Beskow	KTH/PDC	Tier-1	Completed on 01/03/2015
Hazel Hen	HLRS	Tier-0	Completed on 16/10/2015
ARCHER	EPCC	Tier-1	Completed on 20/11/2015
Salomon	IT4I/VSB-TUO	Tier-1	Completed on 29/01/2016
Leo	NIIF	Tier-1	Completed on 05/02/2016

Table 4: *System upgrades from 02/2016 to 03/2016*

2.3.2 *Details of upgraded sites*

1. Beskow – KTH/PDC – Tier-1

Beskow has replaced the old Lindgren system that was fully dismantled after successful installation of the new system. Important points for Beskow are:

- Cray XC40 system architecture, based on Intel Haswell and Cray Aries interconnect, [4].
- Intended lifetime of 4 years, until 31/12/2018
- This system is not connected to the PRACE network, so the following services could not be provided:
  - PRACE Network
  - Data services (GridFTP)

  - Compute services (Unicore)
  - AAA Services (GSISSH).

2. Hazel Hen – HLRS – Tier-0

Hazel Hen is an upgrade/follow on project of the previous system known as Hornet. The Cray XC40 Hazel Hen is based on Intel Haswell processors and the Cray Aries network. Important points for Hazel Hen are:

- Cray XC40 system architecture, based on Intel Haswell and Cray Aries interconnect. See [5] for some more information.
- This is a Tier-0 system provided by the HLRS site as part of the GCS Gauss Centre.
- The system provides all the PRACE core services, except for:
  - UNICORE: Available only upon user request
  - LDAP: Not connected to PRACE LDAP
  - Accounting data: Available through their web interface
  - The optional service GSISSH is not available, substituted by the optional service SSH_X509.

3. ARCHER – EPCC – Tier-1

This system started the activity in November 2013 as the main UK National Supercomputing service. It has been expanded with the addition of a further 10 cabinets of Cray XC30 using the same Intel IvyBridge processor technology that they were using.

The main points of ARCHER system upgrade are:

- Cray XC30 system architecture, based on Intel Ivy Bridge and Cray Aries interconnect. See [6][5] for some more information.
- This partner provides all the PRACE services defined in the catalogue.

4. Salomon – IT4I/VSB-TUO – Tier-1

This system was categorized under a system upgrade because it was taken as an extension of the computing power of the centre, but it is, in fact, a new system. We maintained the system in System Upgrade section because the site was providing services to PRACE.

- The system consists of 1008 compute nodes equipped with Intel Xeon E5-2680v3 architecture and at least 128GB RAM. Interconnection is Infiniband based, [7].
- The site provides all the core services defined in the PRACE catalogue.

5. Leo – NIIF – Tier-1

Leo is the only Top500 HPC system in Hungary, installed in 2015 in the Debrecen city. The system upgrade has been performed with no big issues as all the services have been implemented without specific problems.

- The system is deployed under Intel Xeon E5-2650 v2 architecture using Infiniband FDR as interconnect. The GPUs model is Nvidia K20/K40. For more information see [8].
- The site provides all the core services defined in the PRACE catalogue.

6. Occigen – CINES – Tier-1

It is worth mention here that CINES started its upgrade procedure on June 2014 and installed the Occigen system on January 2015, but in February 2016 CINES decided to do not provide PRACE services. The reasons are mainly technical, as they are missing some front end and data transfer nodes and they are not going to connect the system to the PRACE network.

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

New Tier-0/Tier-1 sites & systems comprehends the activity of keeping track of all the new sites and possibly systems that are going to be integrated into the PRACE infrastructure.

In the current period 2015-2016, this activity has been included under WP6 subtask 6.1.2 and lead by Felip Moll from Barcelona Supercomputer Centre with the co-leading of Javier Bartolomé from BSC as well.

2.4.1 *Overview of new Tier-0/Tier-1 sites & systems*

In this reporting period (February 2015-March 2016) only one new system from CINECA has been integrated into PRACE and started to provide services for that. Currently there are three on-going integrations coming from NIIF, WCSS and GRNET, and two planned ones from UL and FIS. In this document we will comment only the finished integrations.

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

1. Galileo – CINECA - Tier-1

The integration of this system has started in November 2015 and was completed on January 2016. This new system is equipped with up-to date Intel accelerators (Intel Phi 7120p) and NVIDIA accelerators (NVIDIA Tesla K80). The processor architecture is based on Intel Haswell with 128GB RAM per each node.

Galileo is mainly used to develop and run applications targeted at hybrid architectures, leveraging software applications in the fields of computational fluid dynamics, material and life science, and geophysics.

The computing system has been made totally available to European researchers as a Tier-1 system of the PRACE infrastructure, providing all the services of the PRACE catalogue.

For more information on Galileo see [9].

3 Operational Services

Since the previous PRACE IP Projects we had defined a process of selection and deployment of a common set of services aiming at presenting all Tier-0 centres as a single distributed infrastructure, instead of a set of individual systems/computing facilities.

Common services are divided into thematic categories: Network, Data, Compute, AAA, User, Monitoring and Generic. Each service category has a responsible person who is in charge of managing all the information and decisions related to a specific service area.

Selection of common services is published in the PRACE service catalogue and once chosen, the responsibility for a service is taken by the respective service area.

The following sections provide the current status of each service category and the main achievements within the current reporting period.

3.1 Network Services

The PRACE network has constantly evolved over the last years, since its initial deployment by the former EU DEISA project.
The DEISA supercomputing environment has been strongly dependent on the deployment of an internal, dedicated high bandwidth network interconnect between computing platforms, and was strongly focused in HPC. It used Grid and multi-cluster technologies which have been relevant to achieve large scale, tight system integration of a super-cluster at the continental level. In this core infrastructure, state-of-the-art Grid technologies worked transparently in the background and have not been directly seen by the end users.
The installed network infrastructure has been adopted by the PRACE project years ago. For the internal communication between the HPC systems, the dedicated network interconnect infrastructure was chosen also. It allows unconstrained and undisturbed use of predefined communication bandwidth. This also simplifies the detection of bottlenecks and the evaluation of the interaction between the involved supercomputer systems, applications and the network itself. Furthermore a strict security policy had been introduced, only allowing systems to be connected to the network infrastructure that are PRACE resources. Starting with HPC systems only, later storage servers, gridftp and monitoring servers have been added. Each site agreed to connect no other systems to the PRACE backbone and not to route any other traffic via this infrastructure. Having this security policy in place, a “Net of Trust” has been established. Every site has provided a small IP address range out of its official address pool for PRACE usage. A system connected to PRACE may use these addresses for the PRACE dedicated network only.
The design plans and currently installed network infrastructure originate from the year 2005/2006. Since the GÉANT network, where the PRACE network is based on, has been advanced a lot over time, in 2014 GÉANT was asked to provide connectivity options for a PRACE network infrastructure for the next decade. Though the document produced by GÉANT provided a lot of interesting new features, an upgrade/rebuilt could not be realized because of different reasons, yet.
As of February 2016, the PRACE network connects HPC systems from 22 PRACE partners throughout Europe via a star topology network realized mostly by dedicated 10 Gb/s wavelength through the national NRENs and GÈANT infrastructure.
The main tasks performed have been optimizing, operating and monitoring the network infrastructure, attaching new and detaching old HPC systems to/from the network infrastructure. So, since last year, the JUGENE Tier-0 system in Juelich, as well as the supercomputing system at CINES have been decommissioned, JUQUEEN, FERMI and

MareNostrum Tier-0 systems have been newly configured into the network infrastructure and optimized for maximum throughput.

Within the last month several slight reconfigurations have been necessary, not because of PRACE requirements, but because of NREN requirements/optimization issues. Now Nordunet and PIONEER use their own direct fibers from Hamburg to Frankfurt instead of GÉANT fibres. Those changes led to short maintenance intervals only, having nearly no impact on PRACE productions.

Furthermore, the PRACE partner CSCS from Switzerland has been connected additionally providing CPU cycles in the future. The PRACE partner NIIF connected a third system (now NIIF-SC, NIIF-SEGED and NIIF-LEO) in December 2015. Additionally, first preparations have been made to connect GRNET in the near future.

In general, the PRACE network infrastructure has been stable for several years now. The PRACE network is still structured like a “star topology” with “Net of Trust” security policy. In principle it consists of three types of connectivity techniques:

a.) Tier-0 systems are connected by a dedicated 10 Gb/s wavelength from the home organization to a central PRACE router, located in a rack of the German NREN DFN at Interxion in Frankfurt. So the general layout of the PRACE network is a star topology. Over the time this strict direct wavelength access has been softened up. E.g. since several partners in France had to be connected, it was decided to have one 10 Gb/s wavelength between a Renater switch in Paris and the PRACE router in Frankfurt only. The partners in France again have their own connections to this Renater device in Paris, sharing the connection from Paris to Frankfurt. Similar changes have been made for the partners in Poland and partly in Germany.

b.) Since there are a lot of sites with lower bandwidth requirements, an IPSEC/GRE tunnel solution was introduced. Those sites have installed an IPsec/GRE tunnel endpoint at their home organisation and tunnel PRACE traffic through their local NREN and GÉANT to an IPsec/GRE tunnel endpoint at Frankfurt, which forwards the traffic to the central PRACE router locally. All tunnels are shared connections without bandwidth guarantees and end in one 1 Gb/s interface in this endpoint, i.e. shared usage of all those partners.

c.) Shortly after installation of this solution GÉANT announced the L2 VPN services. So, one PRACE partner, which had not been connected at this time, has been connected by this new technology (Partner NIIF in Hungary). This solution also provides no bandwidth guarantees.

The current topology is described in Figure 1 below.

Beneath operating the network infrastructure, upgrading network monitoring to a new version of iperf (iperf v3) at all sites as well as updating network monitoring information and operating the central PRACE network Web server, further plans for a future network layout have been made. Several additional discussions with GÉANT and corresponding NRENs have been started. As a result of all these activities valuable input for a Memorandum of Understanding (MoU) between PRACE and GÉANT have been provided. The PRACE-4IP GÉANT (GN4) MoU was signed on 22 March 2016 and will lead to further improvements of the network layout and within IT security configurations and co-operations.

A requirement the future PRACE network will provide allows new sites to be connected in short time frames and for short periods only. This requirement is related to the PRACE use case that through the PRACE calls European researchers get supercomputer cycles awarded, which they use to generate or manipulate huge data sets. Those Big Data cannot be exchanged via normal research networks without substantial delays. Therefore, a temporary connectivity to the dedicated PRACE network may be required for the start-up phase (some weeks), shutdown phase (some weeks) or through the lifetime of the whole granting period (typically 12 month, but there are also multi-year projects).

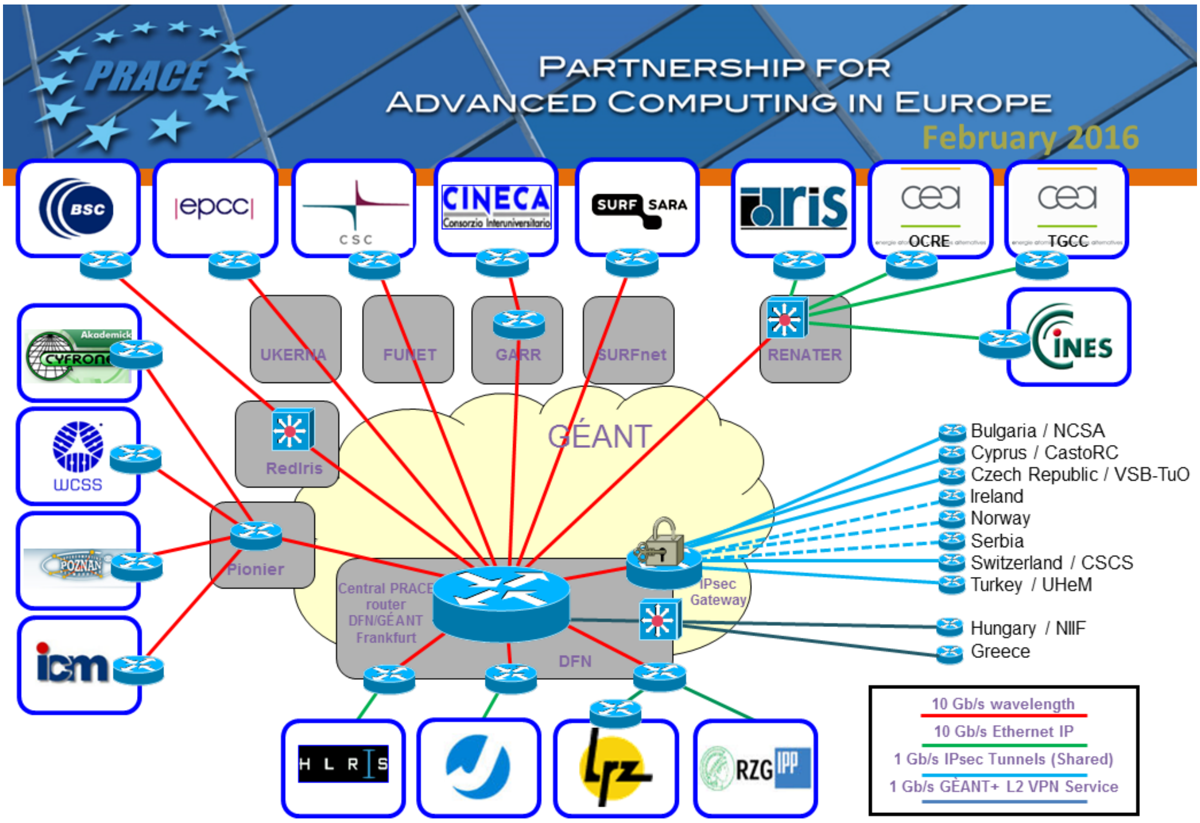

Figure 1: *PRACE network topology*

As already discussed between PRACE and GÉANT, it is assumed that the new "future" design will lift up the usability, flexibility and resiliency and security of the PRACE network infrastructure to a new level helping to solve the challenges arising in the future. After detailed preparation work in year one of PRACE-4IP, it seems to be the right time now to start over and use year two to build the new PRACE network infrastructure.

3.2 Data Services

Data services within PRACE 4IP are delivered by GridFTP as a core service, gtransfer as an additional service and MC-GPFS as an optional service.

3.2.1 *GridFTP*

GridFTP is a widely-used and state-of-the-art protocol for high performance file transfers. It can utilize high bandwidths among PRACE computing centres, therefore it has been chosen as the core service for data transfer within PRACE.

The GridFTP protocol enables high performance transfers by parallel streaming of data through the network. Furthermore, multiple processes are used to speed up the transfer process. This is achieved by configuring more than one backend process, which is responsible for the actual transfers. Additional functionality is provided by third-party transfers, which means that a user can delegate transfers between two remote servers.

Due to security reasons a split configuration is recommended: Frontend and backend parts of GridFTP reside on distinct hosts or machines and only the frontend is visible to a user. Hence the backend part, which usually has access to the supercomputers file systems, is shielded.

Moreover, PRACE internal documentation for GridFTP including an installation and tuning description as well as a special init script is available via the PRACE wiki.

Status of deployment

Table 5 gives the status of deployment on Tier-0 systems in March 2016. All Tier-0 sites have deployed GridFTP and use the split configuration as recommended.

Site / Tier-0 System	Version (GridFTP / Globus Toolkit)
BSC / MareNostrum	6.19 / GT 6.0.14
CEA / CURIE	7.25 / GT 6.0.0
CINECA / FERMI	8.7 / GT 6.0
FZJ / JUQUEEN	6.38 / GT 6.0
HLRS / Hazel Hen	7.26 / GT 6.0
LRZ / SuperMUC	9.4 / GT 6.0

Table 5: *GridFTP deployment status on Tier-0 systems in March 2016*

As shown in Table 6, which presents the status of GridFTP servers on Tier-1 systems, at the date of March 2016, fifteen Tier-1 sites successfully deployed GridFTP.

Site / Tier-1 System	Version (GridFTP / Globus Toolkit)
BSC / MinoTauro	6.19 / GT 6.0.14
CaSToRC / CY-TERA	9.1 / GT 6.0
CINECA / GALILEO	7.26 / GT 6.0
CSC / SISU	9.3 / GT 6.0
Cyfronet / ZEUS	7.26 / GT 6.0
EPCC	GT 6.0
ICHEC / Fionn	GT 6.0
IDRIS / Turing	8.7 / GT 6.0
NIIF / NIIFI SC	9.1 / GT 6.0
NIIF / SEGEDC	9.3 / GT 6.0
NIIF / Leo	9.3 / GT 6.0
RZG / Hydra	9.17 / GT 6.0
SURFsara / Cartesius	8.7 / GT 6.0
UHeM / Karadeniz	7.25 / GT 6.0
UiO / Abel	9.4 / GT 6.0
VSB-TUO / Anselm	8.7 / GT 6.0
VSB-TUO / Salomon	8.7 / GT 6.0
WCSS / Bem	9.4 / GT 6.0

Table 6: *GridFTP deployment status on Tier-1 systems in March 2016*

Security & Bugs

Globus Toolkit 5.2.5 reached end of life effective November 1st, 2015. Therefore, all sites have upgraded to Globus Toolkit 6.0 within the first project year of PRACE-4IP.

3.2.2 *gtransfer*

Gtransfer is an advanced toolkit for performing data transfers. It is a CLI toolkit based on GridFTP that makes use of other tools (like tgftp, globus-url-copy (guc) and uberftp) to provide an easy to use service for users. Gtransfer has been developed at HLRS and is supported by Frank Scheiner (HLRS). The software is available with version 0.4.1 meanwhile.

Gtransfer is an *additional service* in PRACE.

Functionalities of gtransfer

- Inter-network data transfers that can bridge private networks and public networks (e.g. the PRACE network and the Internet). This functionality has been e.g. used extensively by a project within PRACE to transfer data from Europe to the United States and vice versa.
- Optimized data transfer performance by automatically applying pre-optimized data transfer parameters depending on the source and destination of a data transfer or the size of the files transferred. Therefore users only need to provide the source and destination of a data transfer, additional expertise is not required.
- Reliability for data transfers by automatically restarting failed transfers.
- Interruptible data transfers that can be continued later from where they were interrupted.
- Command line completion for options, host addresses and remote directory browsing.
- Host aliases allow usage of easy to remember short names instead of lengthy host addresses. The gtransfer deployments at PRACE sites already include definitions for all available GridFTP hosts in PRACE.
- Persistent Identifiers (PIDs) as used and provided by EUDAT can be used as source of data transfers by users. This way users can easily access and make use of scientific data stored by EUDAT.
- Multipathing: Data transfers can be distributed over multiple paths. This way, users can benefit from the combined bandwidth of multiple paths.
- Data channel checksumming and data channel encryption is available.

Deployment Status

Currently, gtransfer is deployed on all Tier-0 and several Tier-1 sites.

The deployment of gsatellite, a client toolkit with which a user can schedule data transfer tasks like batch jobs, as a PRACE service is currently under consideration.

The PRACE user docs for GridFTP and gtransfer have undergone a revision to reflect the changes and updates since PRACE-3IP.

3.2.3 *MC-GPFS*

This optional service is not provided by any site anymore. Practical experience over the years has shown that this service is not required by users anymore as the working methodology has changed. The process to remove the service from the service catalogue has been started.

3.3 Compute Services

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

Local Batch Systems are software responsible for managing user jobs. With "job" we not only consider sequential and/or parallel applications but also complex tasks like execution chains or workflow, data processing and data management. This software, which is usually part of a supercomputer software stack, plays a decisive role for finding out a common layer to access different systems and architectures.

The way to interact with batch systems is similar for different products. A scripted file must be edited with a full description of a job, e.g. resources like core/cpu-hours, number of cores/cpu, I/O staging, service level, and all needed action for preparing the execution environment. What differs is the way a job is described, basically made using predefined keywords. Another difference is about features that are provided, which depend on the underlying computing resources, e.g. availability of hardware accelerators, schedulers and application launchers, and shared or distributed memories.

Table 7 provides the list of Batch Systems which include features for resources management and job scheduling, deployed on all Tier-0 and Tier-1 systems currently available.

The overview shows the predominance of SLURM followed by IBM LSF® and then PBS Professional. SLURM has successfully been replacing IBM LSF® since last PRACE report. We think it is thanks to a very high effort from developers of improving code and more important to a huge quality contributors list, since SLURM is open source.

The similarities made it interesting to develop a set of wrappers with a common syntax for describing a job around the different batch systems and platforms. Such wrappers should make it easy for users to migrate their work from one system to another.

Work in this direction has been undertaken within Task 6.3 of WP6 in PRACE-3IP that evolved in the adoption of a set of standards that are asked to be implemented on the sites taking part in PRACE under the name of PRACE Common Production Environment.

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

Partner	Site	Machine Name	Tier	Batch System
1. FZJ (de)	FZJ	Juqueen	Tier-0	IBM LSF®
2. GCS (de)	LRZ	SuperMUC phase 1	Tier-0	IBM LSF®
2. GCS (de)	LRZ	SuperMUC phase 2	Tier-0	IBM LSF®
2. GCS (de)	HLRS	Hazelhen	Tier-0	Torque/Moab
2. GCS (de)	RZG	Hydra	Tier-1	IBM LSF®
3. GENCI (fr)	CEA	Curie (Fat)	Tier-0	SLURM
3. GENCI (fr)	CEA	Curie (Hybrid)	Tier-0	SLURM
3. GENCI (fr)	CEA	Curie (Thin)	Tier-0	SLURM
3. GENCI (fr)	IDRIS	Turing	Tier-1	LoadLeveler
4. EPCC (uk)	EPCC	ARCHER	Tier-1	PBSPro
4. EPCC (uk)	STFC	Blue Joule	Tier-1	IBM LSF®
5. BSC (es)	BSC	Marenostrum III	Tier-0	IBM LSF®
5. BSC (es)	BSC	Minotauro	Tier-1	SLURM
6. CSC (fi)	CSC	Sisu	Tier-1	SLURM
7. ETH (ch)	CSCS	Rosa	Tier-1	SLURM
8. SURFSARA (nl)	SURFsara	Cartesius	Tier-1	SLURM
8. SURFSARA (nl)	SURFsara	Cartesius (GPU)	Tier-1	SLURM
10. SNIC (se)	PDC	Beskow	Tier-1	SLURM
11. CINECA (it)	CINECA	FERMI	Tier-0	IBM LSF®
11. CINECA (it)	CINECA	GALILEO	Tier-1	PBS Pro
12. PSNC (pl)	CYFRONET	Zeus (BigMem)	Tier-1	PBS Pro
12. PSNC (pl)	CYFRONET	Zeus (GPGPU)	Tier-1	PBS Pro
12. PSNC (pl)	PSNC	Cane	Tier-1	SLURM
12. PSNC (pl)	PSNC	Chimera	Tier-1	SLURM
12. PSNC (pl)	WCSS	Supernova	Tier-1	PBS Pro
12. PSNC (pl)	WCSS	Bem	Tier-1	PBS Pro
13. Sigma (no)	UIO	Abel	Tier-1	SLURM
16. NUI (ie)	ICHEC	Fionn	Tier-1	Torque/Moab
17. UHEM (tr)	UHEM	Karadeniz	Tier-1	IBM LSF®
18. CaSToRC (cy)	CaSToRC	CY-TERA	Tier-1	SLURM
19. NCSA (bg)	NCSA	EA-ECNIS	Tier-1	IBM LSF®
20. IT4I-VSB (cz)	TUO	Anselm	Tier-1	PBS Pro
21. IPB (rs)	IPB	PARADOX	Tier-1	PBS Pro
22. NIIF (hu)	NIIF	NIIFI SC	Tier-1	SLURM
22. NIIF (hu)	NIIF	Seged	Tier-1	SLURM
22. NIIF (hu)	NIIF	Leo	Tier-1	SLURM

Table 7: *Local Batch Systems in PRACE*

Table 8 shows deployed components on Tier-0 Systems that allow having UNICORE working in PRACE.

Figure 2 gives a graphical overview of the deployment of the UNICORE architecture. The default version of UNICORE (7.3.0) is planned to be deployed in the near future, following a general upgrade plan for all systems connected to PRACE and providing UNICORE services.

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

Table 8: *UNICORE Software components*

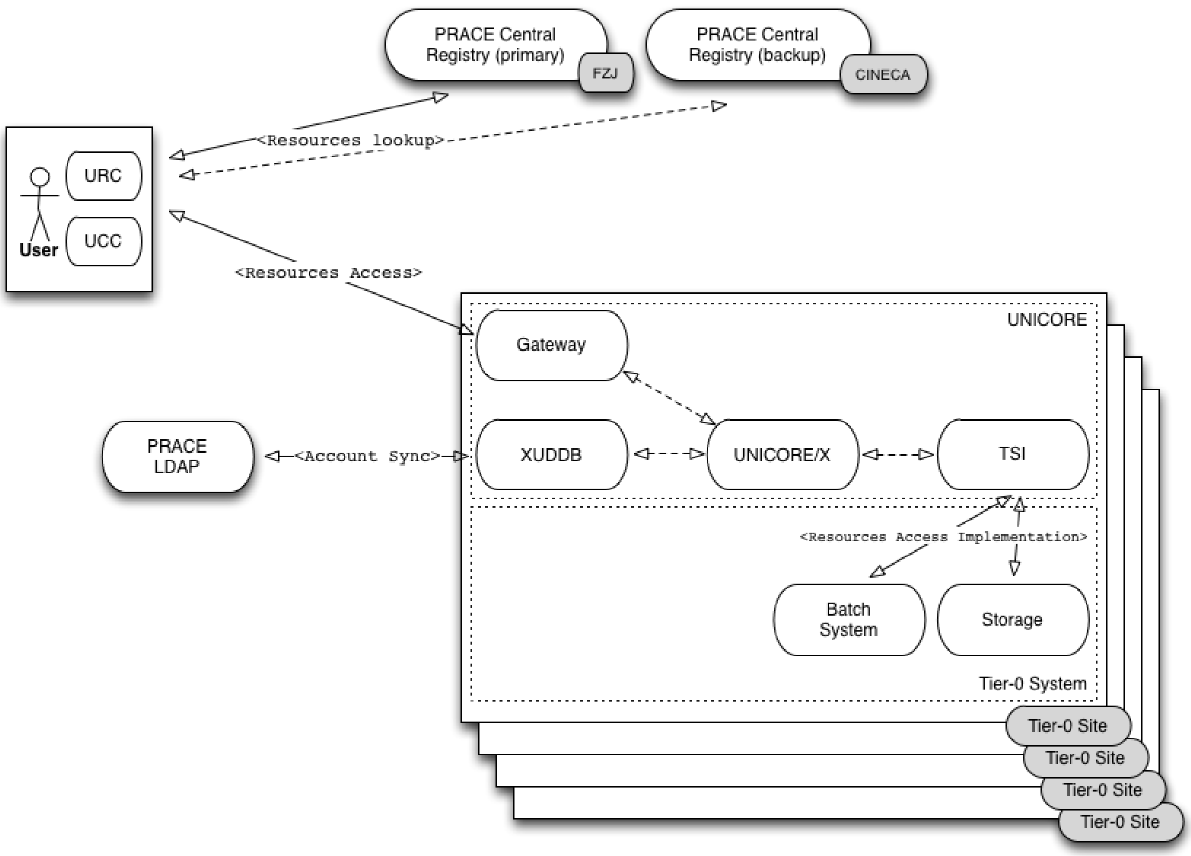

Figure 2: *UNICORE Architecture*

3.4 AAA Services and Operational security

The AAA activity is responsible for services, which provide Authentication, Authorization and Accounting facilities on the infrastructure. This includes the support for the services that provide interactive access to the PRACE systems.

3.4.1 *Public Key Infrastructure - PKI*

Several PRACE services rely on X.509 certificates [10] for authentication and authorization. These certificates must be issued by entities trusted by the service providers. PRACE relies on the Certificate Authorities (CA) accredited as a member by the EUGridPMA, the European Policy Management Authority (PMA) [11], or by one of the two other organizations, the TAGPMA and the APGridPMA, all three federated in the IGTF [12]. 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.

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. Ten updated IGTF distributions have been provided in this project period. PRACE is also member of the EUGridPMA as Relying Partner and participates in its activities.

3.4.2 *User Administration*

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

A single LDAP server is used for PRACE Tier-0 accounts. For most Tier-1 accounts the same server is used, however some partners host a local LDAP server for the registration of Tier-1 users. Separate databases are used for Tier-0 and Tier-1 accounts. An overview of the LDAP namespace is shown in Figure 3. For Tier-0 accounts the “ou=ua,dc=prace-project,dc=eu” part of the name space is used and for Tier-1 accounts “ou=ua,dc=deisa,dc=org” is used as top part. The Tier-1 accounts registered by other LDAP servers can be accessed through referrals.

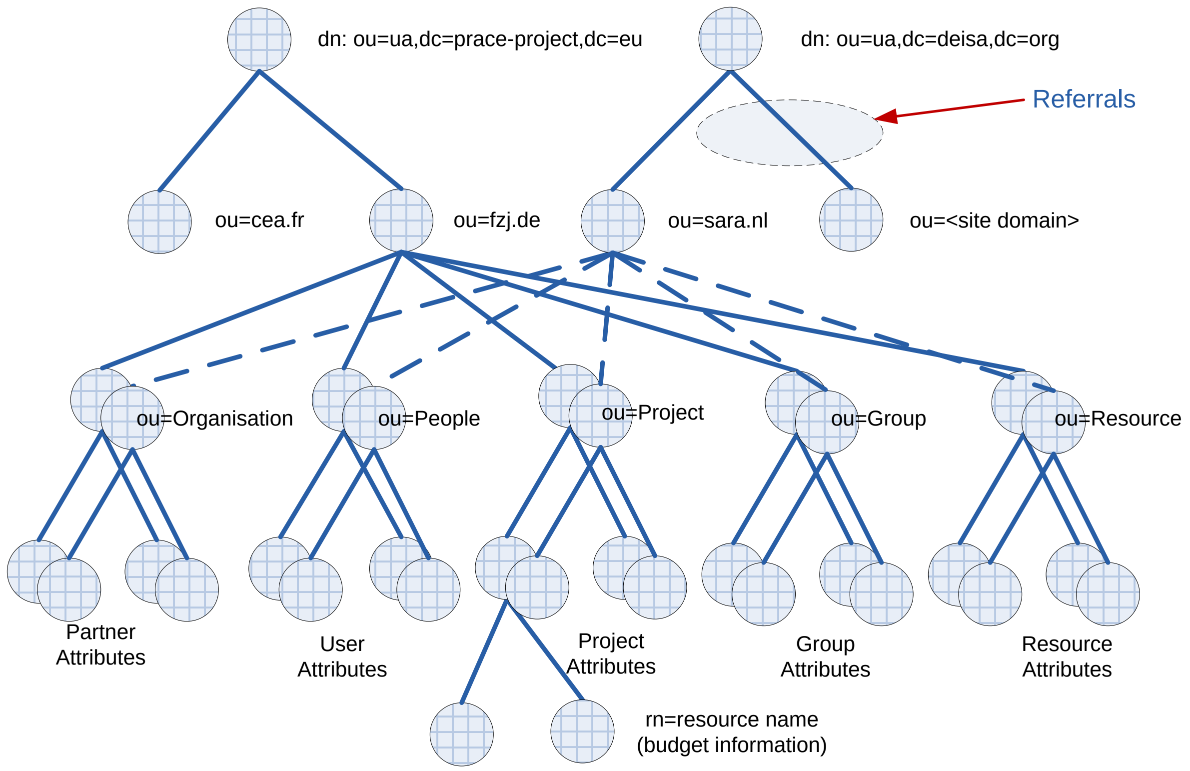

Figure 3: *PRACE LDAP directory tree*

All six operational 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's the Home site which creates and updates the account information of the user. The Home site in general is the site from the country where the user is working and not the site where Tier-1 resources have been granted to the user (the Exec site). The Home site mostly already knows the user and this makes the registration procedure easier. The Exec site retrieves the information from the LDAP server for the creation of a local account.

28 sites manage Tier-1 accounts through their branch. In this period one branch has been closed because the site, ICM from Poland, no longer is a PRACE partner site. Two new partners from Spain, CESGA and UPM, have been added as Tier-1 partners. The main PRACE LDAP server is operated by SURFsara and for high availability both a replica server at SURFsara and a remote replica server at HLRS are operational. The replica servers can be used if the primary server is unavailable for some reason.

Several minor configuration updates to the LDAP service have been applied. One update has been the addition of a new attribute value which will enable the exchange of user information with the EUDAT project. This exchange of user information is needed to enable access by PRACE users of data managed by EUDAT services. This collaboration with EUDAT is an activity of Task 6.3 of the PRACE-4IP project and details are reported by that task. The PRACE AAA Administration guide has been updated for this change.

3.4.3 *Interactive Access*

Interactive access to the Tier-0 systems is provided either by the SSH (Secure Shell) facilities, provided by most distributions of operating systems, or X.509 based SSH facilities. The latter are mainly used for access from other PRACE sites using the dedicated PRACE network. On PRACE Tier-0 and Tier-1 sites the recommended X.509 based SSH facility is GSI-OpenSSH [13] (GSI-SSH for short), distributed by the Globus community. All partners had to update these GSI-SSH facilities to the Globus Toolkit (GT) version 6 distribution by the end of 2015 if not done before, because support for the older version 5.25 ended November 2015. All sites that support GSI-SSH are now on GT version 6.

For access from their workplace users can use GSISSH-Term, a GSI-SSH Java based client, which is supported by the PRACE partner LRZ [14].

Not all sites support GSI-SSH based access from the internet. Therefore, two partners, CINECA (IT) and SURFsara (NL), provide GSI-SSH based access for all PRACE users who can use these sites as door nodes to other PRACE sites using the PRACE network.

3.4.4 *Accounting Services*

Users can get accounting information locally at the sites at which they consume resources. In addition users and PRACE staff can get accounting information from several sites in a uniform way with the DART tool [15]. With this tool users can display their accounting information for five Tier-0 systems and nine Tier-1 sites.

For configuration changes five updates of the tool have been distributed.

3.5 Security Forum

The PRACE Security Forum has three main tasks:

- Defining security related Policy and Procedures;
- The Risk Review of new services or service upgrades;
- The management of operational security.

All PRACE partners contributing services to the PRACE infrastructure are represented in the Forum and can vote on the decisions. In this section we report on the internal activities of the Security Forum. The security activities in collaboration with other organisations are described in PRACE-4IP Task 6.3 reports.

3.5.1 *Security Policies and Procedures*

The PRACE common services provide access among PRACE sites and because of this sites cannot be completely isolated from each other. So, sites are vulnerable for security incidents at one of the other sites. This means that sites must trust the security measures at other sites. For this reason, partners are asked to publish on the internal PRACE wiki pages the documents describing the local policies and procedures. In this way partners can access the security policies and procedures of each other. Some partners need more time for publication because the information is only available in the local language or management approval for external publication of the documents is needed.

Partners can also do a self-assessment using a document published by the Security for Collaborating Infrastructures (SCI) group [16]. This document provides a list of policies and procedures which should be implemented as a basis for trust among infrastructures, but equally can be used to build trust among partners within a collaboration.

On a more detailed level a document has been developed and discussed which provides a list of actions to be taken to set up a well-protected system, so minimizing the chance of incidents.

3.5.2 *Risk Reviews*

The Security Forum performs a risk assessment of a new service or the update of an existing service if there are changes in the security set-up of the latter.

In this reporting period a risk review has been done for the new monitoring facility of PRACE services, based on the ICINGA2 tool. This risk review helped to improve the description of the security aspects of the set-up of the new facility and enabled the acceptance of the service by partners.

3.5.3 *Operational Security*

All Tier-1 sites are member of the PRACE CSIRT team. Incidents are reported and discussed using an e-mail list and video/telcon facilities. Emergency phone numbers are available for all sites.

There were no security incidents reported by sites in this period. Information on six vulnerabilities in software has been shared among partners. Although information about vulnerabilities in general will be received by partners by other means too, it is regarded as helpful if specific information is also provided through internal channels. It will guarantee that attention will be given to those vulnerabilities. To further enhance this service for partners it is proposed that for sharing vulnerability information rules will be defined for when and how information should be shared. Also required is that the CSIRT team will assign the role for monitoring vulnerability information provided by other sources.

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. This year a special focus has been put on the User Documentation activity.

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 the underlying architecture. It guarantees availability of a minimum set of software components to support the running of highly-scalable parallel simulation software. Thus it allows automated monitoring of key software components across all PRACE execution sites.

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

This year PCPE has been installed on the newly integrated PRACE systems. Its specification has not evolved during the period but a specific focus will be put on this activity for the second year. Thanks to the newly deployed monitoring system, tests of the PCPE installation will be developed to help sites knowing their user environment is conform to the PCPE specification.

3.6.2 *Trouble Ticket System*

The centralised Helpdesk [17] was deployed as part of the PRACE-1IP project. It's an important tool for the PRACE project staff to communicate among sites about problems. The Helpdesk is also available for users. However Tier-0 users are advised to contact the helpdesk of the Tier-0 sites directly.

TTS is based on the Request Tracker (RT) issue tracking system [18]. Since its deployment in PRACE-1IP, this service is hosted and operated at CINECA. Version 3.8 of RT was running but this version was recently stated "end-of-life" by Best Practical, the editor of the RT open source software. Thus for security concern it was decided to upgrade the PRACE RT instance to the new stable version 4.2. RT upgrade operations were conducted by CINECA during the last quarter of 2015 and since the beginning of 2016 this new version is fully operational.

A survey has also been conducted this year to learn about the various solutions for issue management used at PRACE sites for their local operations. These site-specific TTS are often put in contact with the PRACE TTS (they receive the various ticket notifications sent by PRACE TTS). The goal of this survey was to know better these site-specific TTS in order to improve the interactions they have with the PRACE TTS. PRACE TTS sends automatic notifications at different user ticket stages. These automatic email notifications may sometimes conflict with automatic behaviours of site-specific TTS. With the information collected in this survey, global improvement of the PRACE TTS notifications will be studied.

3.6.3 *Helper Scripts*

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

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

The prace_service script itself has also been improved to support the Iperf3 service recently deployed at sites within the Network subtask. With this addition in prace_service, the network monitoring activity is now able to rely on this directory service rather than maintaining information of Iperf3 service addresses on its own.

3.6.4 *PRACE User Documentation*

The User documentation for PRACE is available online on the PRACE Research Infrastructure website [19]. A specific focus has been put on this activity this year. Firstly to synchronize documents written by WP6 staff with their published counterpart on prace-ri.eu website. Then clarification was added to the publication process and finally the edition and review work of documents was improved.

First actions made this year on the User Documentation was to synchronize the sources of the document hosted in the PRACE SVN repository [20] with their published counterpart on the WordPress instance that runs the prace-ri.eu website. After many years of documentation updates it was found that many user documents were different between their published version and their source version. Sometimes the published version was updated without the relative changes made on the source version of the document. Sometimes the source version was updated but not published. A global audit was conducted on all the documents on the

SVN repository and on the prace-ri.eu website. After that audit, everything was re-synchronized to get global coherency back.

After that global re-synchronization of the user documents, the publication process was clarified as follow (see Figure 4):

1. document leader or main author update source document and push this modification to the PRACE SVN;
2. author asks the User Documentation review panel to review the updated document;
3. the review panel validates or rejects the update (no answer from the panel means acceptance);
4. if accepted, the User Services subtask leader publishes the new version on the PRACE-RI Website.

Last evolution made on the user documentation was to change the markup language used to write WP6 user documentation. These docs were edited as raw HTML by the documentation review panel. Once committed to SVN and approved, these HTML contents were published by copy and paste to their relative WordPress post on prace-ri.eu. The goal of this change was to facilitate the edition work of document authors and document reviewers by using a simpler markup language: Markdown [21]. With HTML, documents are hard to write and maintain, especially when list or table elements are used. HTML implies to focus a lot on the markup whereas edition work should focus on the content and markup should be the work of the person who publishes the documents.

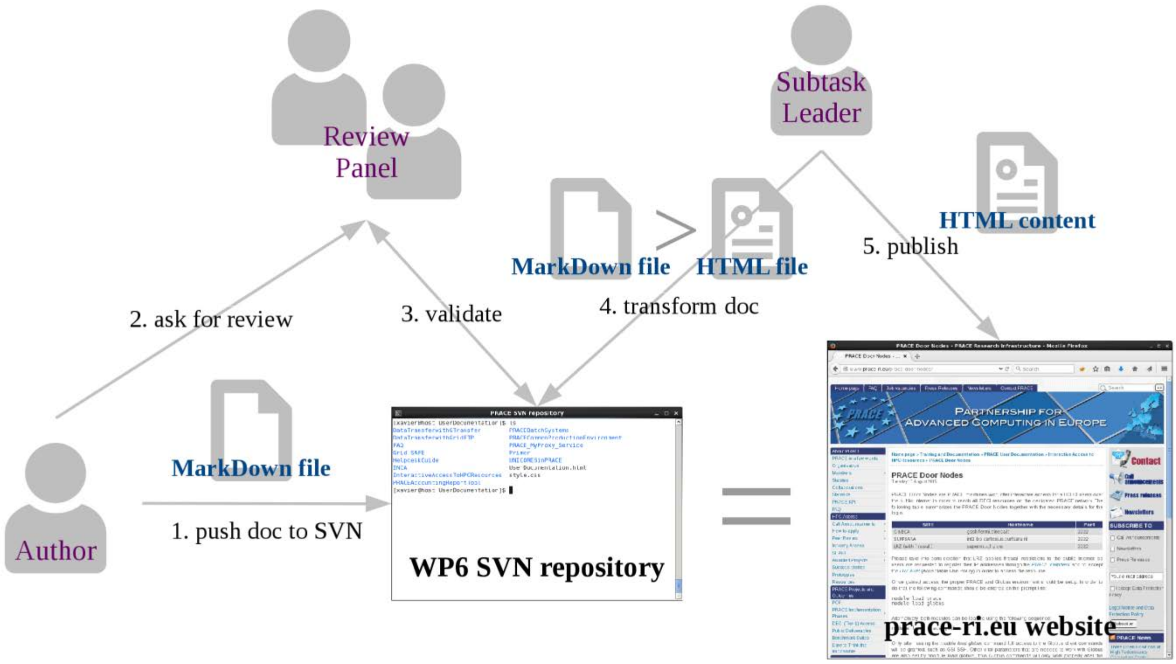

Figure 4: *Publication process of the PRACE User Documentation*

So to make easier the edition work it was proposed and accepted to move from writing HTML files to writing Markdown files. Markdown is lightweight markup language easy to write and easy to read. Markdown's syntax [22] is easy to use and much simpler than raw HTML, since for instance you do not have to enclose every text paragraph into tags. Markdown supports only a subset of what can be done with HTML, but this subset is sufficient to write user documentation.

Once this change was accepted, all the user documentation sources have been translated from HTML to MarkDown. Then, the publication process has been adapted to translate the relative MarkDown files to HTML, as the content of the PRACE-RI Website is managed with this markup language. To perform this translation, some scripts have been developed on top of the Pandoc open source tool [23].

3.7 Monitoring Services

Monitoring service is the crucial service for the entire operational PRACE-RI infrastructure. It gathers the states of the infrastructure services and tools among all sites connected through PRACE network. The processed data in the aggregated form, keeping the track of its changes, the service offers the feasible tool for site local-based operations in terms of the accessibility of the PRACE Software Stack [24]. This is also the complementary service for PRACE Network accessibility. The monitoring service features are also defined within the PRACE documentation as “core service”.

The main monitoring services activity during the current reporting period of the project was the transition to the new monitoring middleware and restoration of its functionality within the existing environment.

3.7.1 *Monitoring Middleware Change*

Up to 31 August 2015, INCA monitoring service was configured and available within PRACE RI. INCA [25] middleware was installed at LRZ but in PRACE-4IP, LRZ cannot maintain INCA any longer. This middleware is not supported by its developers SDSC (San Diego Supercomputer Center), the latest available version is: 2.5 released in 21 August 2009. The current state of INCA software might be described as ‘abandon-ware’ which implies the lack of updates and proper security maintenance. In such a situation it was considered as severe security risk for the PRACE infrastructure and required prompt and functional replacement.

We had started to investigate new possible monitoring middleware according to the following criteria:

- Licensing model;
- Active user community;
- Frequent live events available to the users;
- Blogs and other dissemination channels on the development status;
- Frequent software updates and fast vulnerability resolving;
- Popularity;
- PRACE-RI requirements.

The aforementioned criteria are important in case of the future possible dislocation of the Monitoring service installation to another site. In such a case, all monitored sites and hosts are affected by the configuration and even software change, which imply additional operational work. The new candidate, which was analysed, must address all the enumerated assumptions – mainly, the future change of the middleware software is highly undesirable.

Several monitoring candidates were considered: Argo [26], Nagios [27], Zabbix [28], Icinga [29], Opsview [30]. The Icinga ver.2 was chosen. Its monitoring paradigms are compatible with the Nagios and many host administrators know the basics of this one.

3.7.2 *The Transition*

The introduction of the new middleware determines the new communication approach. The main architecture assumptions taken under consideration are:

- Centralized repository of monitoring data on machine located in PSNC;
- Web portal available in public network;
- Monitoring data and Icinga communication runs in PRACE private address space only;
- Minimal engagement of each host administrator;
- Each monitored host equipped with 2 accounts.

Design goals:

- Keep probes/checks as light as possible;
- Get rid of all non-standard requirements for probes (libs, binaries, CPAN, PIP);
- Use bash and python (standard libs) for scripting;
- Use GIT for probes/checks development and (auto)deployment;
- Enable simple debug mode for each check.

All monitored sites have been requested to create two dedicated user-robot accounts. Manager account is dedicated for monitoring scripts and programs (i.e. used to run run **nrpe** daemon) or be used as remote execution account using **gsissh** access depending on the site choice. Of course, a site may decide to disable remote access to that account and maintain it on their own - by taking care of scripts deployment and configuration. The second account is used as destination for site-to-site tests (**gsissh** and **gridftp**). This account does not contain any scripts and monitoring components. It is used just for testing and may be restricted to allow restricted set of commands in the future. Figure 5 shows an example of the Icinga2 site to site tests matrix.

The authorization is based on the short lived proxy for site2site monitoring which is generated by the monitoring server. Unique proxy with different ID is generated for each of monitored hosts and logged for tractability purposes. **globus-url-copy** is used for uploading the proxy to the agent-account in each site. **globus-url-copy** will upload credentials only to the remote servers with the certificate issued by a trusted CA and with certificate subject defined in the configuration file. The creating and uploading process is deployed as Icinga2 check is, so it can be monitored easily. The monitoring (S2S) Proxy is stored in the home directory of agent-user.

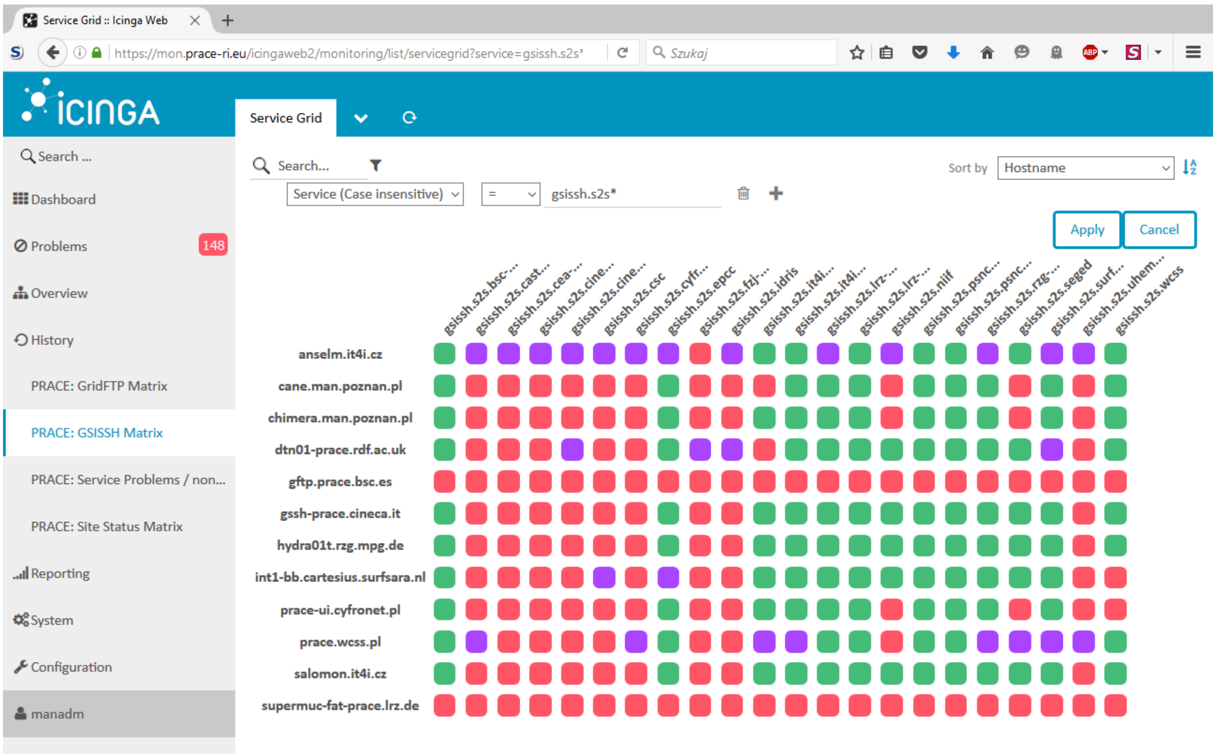

Figure 5: *Icinga2 site to site tests matrix*

3.7.3 *The Status*

The current configuration contains the monitoring of the following services:

- gsissh.port – Host availability check based on gsissh port state
- gsissh.s2s – site to site gsissh connection check
- gridftp.s2s – site to site gridftp connection check
- software.version.tools
- software.version.libraries
- software.version.compilers
- software.version.shells

The current status of integration is depicted in the Table 9. The integration means here the process of configuring remote infrastructure since the Icinga local middleware had been prepared much in advance.

Site	Site-Reported Status	Comment	Validated	Validation comment
BSC (es)	Yes	Added to grid-map file	OK	
CASTORC (cy)	Yes	Added to grid-map file. Accounts created locally.	OK	
CYFRONET (pl)	Yes		OK	
EPCC (uk)	Yes		OK	
LRZ (de)	Yes		OK	
NIIF (hu)	Yes	On NIIFI-SC, SEGED and LEO	OK	
PSNC (pl)	Yes		OK	
RZG (de)	Yes		OK	
SURFSARA (nl)	Yes		OK	
VSB-TUO (cz)	Yes	On both clusters (Anselm & Salomon)	OK	
WCSS (pl)	Yes		OK	
CEA (fr)			PROBLEM 20160322	gsissh: Perm. Denied - possible mapping issue
CINECA (it)	Yes		PROBLEM 20160322	problem with s2s mapping
CSC (fi)	Yes		PROBLEM 20160322	possible gridmapfile mapping problem
FZJ (de)	Yes	What will be executed?	PROBLEM 20160322	gsissh: Perm. Denied - possible mapping issue
ICHEC (ie)	Yes		PROBLEM 20160322	firewall issue: Connection Refused
IDRIS (fr)	Stopped	Waiting for additional information	PROBLEM 20160322	gsissh: Perm. Denied - possible mapping issue
NCSA (bg)			PROBLEM 20160322	Connection Timeout
UHEM (tr)			PROBLEM 20160322	gsissh: Perm. Denied - possible mapping issue
UIO (no)	Ongoing		PROBLEM 20160322	gsissh: Perm. Denied - possible mapping issue
CINES (fr)			NO-SITE-ENTRY	no site data in prace-service.config
CSCS (ch)			NO-SITE-ENTRY	no site data in prace-service.config
PDC (se)			NO-SITE-ENTRY	no site data in prace-service.config
HLRS (de)	Stopped	proxy solution like in INCA needed		

Table 9: *Current status of Icinga integration*

3.8 Generic Services

This section describes the actions done during the reporting period within Generic Services subtask. In general all services that need an operational basis and a centralized distribution for the PRACE project (or a part of it) could be assumed as Generic Services.
The goal of this task is the provisioning of these services and the supervision of their operation, as they are crucial for the day-to-day work of the project.
The Service leader of this task is Zoltan Kiss (NIIFI) since the beginning of PRACE-4IP.

3.8.1 *Catalogue*

When the project started, the catalogue of Generic Services available in PRACE wiki seemed to be out-dated, along with the services information in detail. From an analysis were discovered and identified several more PRACE services to be managed as Generic Services. It also identified that contacts, service responsibles, descriptions and hosting details of these services have to be documented. The services can be divided into two groups: internal and external services.

The services reported in Table 10 have been identified as Internal Services:

Service name	Purpose	Software	Host	Service Operator
BSCW	Document repository	Basic Support for Cooperative Work	FZJ	FZJ
DPMDB	DECI management interface	Custom	SURFsara	SURFsara
SVN	Version controlled repository	SVN + Trac	SURFsara	SURFsara
Wiki	Service similar to Wikipedia	TWiki	FZJ	FZJ
Mailing Lists	Mailing list service	Mailman	FZJ	FZJ

Table 10: *Internal generic services*

The services reported in Table 11 have been identified as External Services:

Service name	Purpose	Software	Host	Service Operator
WEB	Project Web	Wordpress	CINES	NIIF
Training Portal	Training Web	TYPO3	CSCS	CaSToRC
PRACE Events Service	Managing Events	Indico	NIIF	NIIF
CRM	Contact database	SugarCRM	CaSToRC	CaSToRC
PPR Tool	Project and DECI project review	Custom	CINES	CINES
Summer of HPC Webpage	SoHPC Web	Wordpress	ICHEC	UL

Table 11: *External generic services*

Service and infrastructure operational contacts were gathered for each service, created generic-services mailing list subscribing these contacts to make the work of this group easier. The address of the internal list is prace-generic-services@fz-juelich.de, and was used for announcements and organisation of activity.

3.8.2 *Roles and Actions*

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

- PRACE aisbl;
- Other WPs;
- External providers

upon request when IT or operational issues or requests of these entities arise. Examples from the reported period are:

- Questions of another WP related to the operation of new or existing services
    - Arranging hosting details of PRACE Events portal;
    - Integration between Training, Events portals and PRACE Web site;
    - Asking for PRACE website accounts for users responsible for documentation.

- Questions of PRACE aisbl when negotiating with service providers offering service for PRACE
    - Help with consulting mailing lists provider;
    - Consulting with service hosts;
    - Helping managing domain portfolio;
    - Help with certificate installation.
- Helping internal workflows by using a shared PRACE-owned resource
    - Handling *.prace-ri.eu wildcard certificates, general help, helping with new certificate requests;
    - Handling prace-ri.eu domain: subdomain requests, helping with resolving domain issues related to a change.

Reacting to the recent SSL related vulnerabilities and in general to maintain high level of security among generic services, an online SSL security analysis was made on the servers hosting these services.

Among several named vulnerabilities (Heartbleed, Heartbeat, POODLE, Logjam, BEAST, DROWN), the online tool at www.ssllabs.com was capable of analysing several other weaknesses and vulnerabilities.

In five cases, suggestions on improving the security of specific services were forwarded to service operators.

A request of structured data storage originating from WP3 and PMO was processed and presented to WP6, as potential services need to be analysed to offer detailed search on the structured data. The purpose of this service is to provide comfortable search functionality to reach results of PRACE activities faster e.g. by searching for author or title of a document directly.

The service needs to store different datasets produced by activities founded by PRACE, e.g. White Papers, Best Practice guides, Publications as a result of access and Project Deliverables.

The analysis of source data started to better understand the requirements to evaluate services capable of supporting WP3 to disseminate activities. The result of initial discussions has been that a repository service might be a good solution for the purpose. The activity will continue through the second year of the project.

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 services. The procedure for the on-duty activity has been simplified to assure a more concrete and punctual support on the day by day monitoring actions.

Six Tier-0 sites and 26 Tier-1 systems are actually in production and during the reporting period a total of six sites have upgraded their systems (Beskow at KTK/PDC, Hazel Hen at HLRS (Tier-0), Archer at EPCC, Salomon at IT4I-VSB, Leo at NIIF and Occigen at CINES). Furthermore, a new Tier-1 system has been installed and integrated in PRACE (Galileo in CINECA).

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

The new version of the PRACE Service Catalogue approved by PRACE BoD in March 2015 has guided the operational activity of the common services in the reporting period. Some of the major achievements are underlined here:

- The Network Services have evolved and actually the PRACE network connects HPC systems from 22 PRACE Partners. A new version of the network monitoring tools (iperf V 3) has been put in operation.
- A MoU has been signed between PRACE and GÉANT to improve the cooperation on the network layout and IT security. In the next year a new PRACE network infrastructure could be built.
- The Data services have mainly focused on GridFTP (file transfer) and Gtransfer. A split of GridFTP (frontend and backend parts of GridFTP resident on different hosts) has been recommended for security reasons; Furthermore the update of Globus Toolkit to version 6.0 has been adopted to all the sites.
- Gtransfer, an advanced toolkit for performing data transfer is deployed on all Tier-0 and several Tier-1 sites. The gsatellite client toolkit to schedule data transfer tasks is under evaluation.
- In the Compute Services activity different wrappers with a common syntax for describing a job around different batch systems and platforms have been analysed to adopt a set of standards to implement from different sites in the context of the PRACE Common Production Environment.
- PRACE continues to maintain the updated PRACE Partners information for the Certificate Authorities in a central repository, as different PRACE services rely on X.509 certificates for AAA procedures. PRACE is a member of the EUGridPMA (the European Policy Management Authority) as Relying Partner and participates in its activities.
- Information about PRACE users and their accounts is maintained and updated in the LDAP based repository. This facility is used to share among PRACE partners the authorisation information needed by PRACE services (i.e. data transfer via GridFTP, interactive access through GSI-SSH, Job submission via UNICORE, etc.).
- The PRACE Security Forum coordinated all security related issues. The reported security vulnerabilities have been managed efficiently. Furthermore, the Security

Forum has performed the risk review for the new monitoring facility of PRACE services based on the Icinga2 tool.

- The PRACE Common Production Environment (PCPE) has been installed on the newly integrated PRACE systems. In the second year of Task 6.1 activity, a focus will be put on the evolution of the service and its integration with the new monitoring service to help sites to know if the user environment conforms to the PCPE specifications.
- The Help Desk on Duty (HdD) and Operator on Duty (OoD) activities have been merged in a unique activity to simplify the process and make more effective the continuous on Duty activity. One report of the activity is produced every week, reporting the status of the infrastructure and the core service and the problems occurred.
- The PRACE Request Tracker instance for the TTS has been upgraded to the new stable version 4.2 for security reasons. The new version is fully operational in the TTS service hosted and operated by CINECA.
- The User documentation, available on-line on the PRACE-RI web site and on the SVN repository has been updated and synchronised, through a global audit conducted on all the documents. Furthermore, to simplify the editing work on the documents, it has been decided to adopt Markdown files instead of the older HTML files for the documentation.
- New tools for the monitoring service have been analysed to substitute the older INCA tool which is no longer fully supported. Icinga Version 2 has been selected and integrated to configure the different systems to monitor the PRACE infrastructure.
- Both internal (i.e. document repository, SVN, mailing list services, etc) and external (Web portal, web training portal, CRM, etc) generic services have been constantly monitored and managed.
- A request for a Structured Data Storage service (requested from WP3 to retrieve and access quickly and smartly to PRACE information and documents i.e. authors or titles of documents, etc) has been analysed and the requirements from WP3 will be evaluated for a future implementation of the service.

Thanks to the procedures for incident and change management the complete set of PRACE common services as defined in the Service Catalogue have been operated and monitored day by day. The Service Catalogue represents a living document and, as a result of the successful evaluation in the first year of PRACE 4IP, it will be revised during the second year of PRACE-4IP. In the process towards PRACE quality of service and quality control, a focused work on PRACE Operational Key Performance Indicators (KPI) is just started and will be better analysed. The measurement of the indicators and the evaluation will be fully addressed in year two of Task 6.1 activity.