



**SEVENTH FRAMEWORK PROGRAMME
Research Infrastructures**

**INFRA-2011-2.3.5 – Second Implementation Phase of the European High
Performance Computing (HPC) service PRACE**



PRACE-2IP

PRACE Second Implementation Project

Grant Agreement Number: RI-283493

D6.1

First Annual Operations Report of the Tier-1 Service

Final

Version: 1.0
Author(s): Jules Wolfrat, SARA
Date: 27.08.2012

Project and Deliverable Information Sheet

PRACE Project	Project Ref. №: RI-283493	
	Project Title: PRACE Second Implementation Project	
	Project Web Site: http://www.prace-project.eu	
	Deliverable ID: < D6.1 >	
	Deliverable Nature: <DOC_TYPE: Report / Other>	
	Deliverable Level: PU *	Contractual Date of Delivery: 31/08/2012
		Actual Date of Delivery: 31/08/2012
EC Project Officer: Leonardo Flores Añover		

* - The dissemination level are indicated as follows: **PU** – Public, **PP** – Restricted to other participants (including the Commission Services), **RE** – Restricted to a group specified by the consortium (including the Commission Services). **CO** – Confidential, only for members of the consortium (including the Commission Services).

Document Control Sheet

Document	Title: First Annual Operations Report of the Tier-1 Service	
	ID: D6.1	
	Version: 1.0	Status: <i>Final</i>
	Available at: http://www.prace-project.eu	
	Software Tool: Microsoft Word 2007	
	File(s): D6.1.docx	
Authorship	Written by:	Jules Wolfrat, SARASARA
	Contributors:	Gerard Gil (CINES), Ralph Niederberger (FZJ), Hand Eide (UiO), Agnieszka Kwiecien (WCSS), Ender Güler (UHeM), Tilo Eissler (LRZ), Gabriele Carteni (BSC), Filip Stanek (VSB-TUO), Thomas Bönisch (HLRS), Andreas Schott (RZG), Eoin McHugh (ICHEC), Liz Sim (EPCC), Emanouil Atanassov (NCSA)
	Reviewed by:	Michael Browne, ICHEC Thomas Eickermann, FZJ
	Approved by:	MB/TB

Document Status Sheet

Version	Date	Status	Comments
0.1	07/June/2012	Draft	TOC
0.2	02/July/2012	Draft	Update Table 3, template for status services. Added contribution CINES

0.3	01/August/2012	Draft	Added contributions for sites and services
0.4	09/August/2012	Draft	Still waiting for some contributions
0.5	13/August/2012	Draft	For internal review
0.6	23/August/2012	Draft	After internal review, with some comments
1.0	24/August/2012	Final version	

Document Keywords

Keywords:	PRACE, HPC, Research Infrastructure
------------------	-------------------------------------

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° RI-283493. 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.

Copyright notices

© 2012 PRACE Consortium Partners. All rights reserved. This document is a project document of the PRACE project. All contents are reserved by default and may not be disclosed to third parties without the written consent of the PRACE partners, except as mandated by the European Commission contract RI-283493 for reviewing and dissemination purposes.

All trademarks and other rights on third party products mentioned in this document are acknowledged as own by the respective holders.

Table of Contents

Project and Deliverable Information Sheet	i
Document Control Sheet.....	i
Document Status Sheet	i
Document Keywords	iii
Table of Contents	iv
List of Figures	v
List of Tables.....	v
References and Applicable Documents	vi
List of Acronyms and Abbreviations.....	vi
Executive Summary	1
1 Introduction	2
2 Operation and coordination of comprehensive common Tier-1 services	3
2.1 Overview of Tier-1 infrastructure.....	4
2.2 Network Services	8
2.3 Data Services.....	9
2.3.1 <i>Status of Deployment</i>	9
2.3.2 <i>Documentation</i>	10
2.3.3 <i>Security</i>	10
2.3.4 <i>Advanced User Tools</i>	10
2.3.5 <i>Multi-cluster Shared File System – MC-GPFS</i>	10
2.4 Compute services	11
2.4.1 <i>Local Batch Systems</i>	12
2.4.2 <i>UNICORE</i>	13
2.4.3 <i>GRAM5</i>	14
2.5 AAA services	14
2.5.1 <i>Public Key Infrastructure - PKI</i>	15
2.5.2 <i>User Administration</i>	15
2.5.3 <i>Interactive access</i>	17
2.5.4 <i>Accounting services</i>	17
2.6 User Services	18
2.6.1 <i>PRACE Common Production Environment</i>	18
2.6.2 <i>User Documentation</i>	19
2.6.3 <i>PRACE Trouble Ticket System</i>	19
2.7 Monitoring Services.....	19
2.8 Internal services.....	20
3 Integration of new Tier-1 systems in the PRACE distributed RI	21
3.1 Introduction	21
3.2 Overview and status	22
3.2.1 <i>CINES</i>	22
3.2.2 <i>PSNC</i>	23
3.2.3 <i>ICHEC</i>	25
3.2.4 <i>NCSA</i>	26
3.2.5 <i>WCSS</i>	27
3.2.6 <i>UHeM</i>	28
3.2.7 <i>UiO</i>	30

3.2.8 VSB-TUO.....	31
4 Annexes	33
4.1 PRACE Service Catalogue.....	33
Introduction	33
Core services	34
Additional services	34
Optional services	34
PRACE Services.....	36
Uniform access to HPC.....	36
PRACE internal interactive command-line access to HPC	36
PRACE external (user) interactive command-line access to HPC	36
Project submission.....	37
Data transfer, storage and sharing	37
HPC Training.....	38
Documentation and Knowledge Base.....	38
Data Visualization	39
Authentication	39
Authorization	39
Accounting	40
Information Management.....	40
Network Management	41
Monitoring.....	42
Reporting	42
Software Management and Common Production Environment	42
First Level User Support	43
Advanced User Support.....	43

List of Figures

Figure 1 Structure of PRACE operational management	4
Figure 2 Overview PRACE network August 2012	8
Figure 3 Batch Systems distribution on PRACE Tier-1 systems	12
Figure 4 UNICORE architecture for PRACE Tier-1 systems.....	13
Figure 5 UNICORE deployment on PRACE Tier-1 systems.....	14
Figure 6 PRACE LDAP directory tree.....	16
Figure 7 Accounting architecture	18
Figure 8 PRACE Service provision scheme and contracts to its users.....	33

List of Tables

Table 1 Overview of sites and Tier-1 systems in use.....	6
Table 2 Provision of DECI resources, an “x” indicates participation	7
Table 3 Overview installed versions of GridFTP and Globus Toolkit.....	9
Table 4 MC-GPFS installations in PRACE – GPFS I/O clusters.....	11
Table 5 MC-GPFS installations in PRACE – systems with GPFS mounted file systems	11
Table 6 Provision of DECI resources by new partners, x = participation	21
Table 7 Overview of integration and planning for CINES.....	23

Table 8 Overview of integration and planning for PSNC	24
Table 9 Overview of integration and planning for ICHEC	26
Table 10 Overview of integration and planning for NCSA	27
Table 11 Overview of integration and planning for WCSS	28
Table 12 Overview of integration and planning for UHeM	30
Table 13 Overview of planning of the integration for UiO	31
Table 14 Overview of planning of the integration for VSB-TUO	32
Table 15 Classification of PRACE Services as part of the PRACE Service Catalogue.....	34
Table 16 Overview of PRACE services, categories and product classes	45

References and Applicable Documents

- [1] <http://www.prace-project.eu>
- [2] <http://jira.globus.org/browse/GT-195>
- [3] gtransfer: <https://github.com/fr4nk5ch31n3r/gtransfer>
- [4] IGE project: <http://www.ige-project.eu/>
- [5] Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile RFC 5280: <https://datatracker.ietf.org/doc/rfc5280>
- [6] EUGridPMA: <http://www.eugridpma.org>
- [7] IGTF: <http://www.igtf.net/>
- [8] PRACE distribution of CA information: <http://winnetou.sara.nl/deisa/certs>
- [9] Globus GSI-OpenSSH: <http://www.globus.org/toolkit/security/gsiopenssh>
- [10] Usage Record – Format recommendation: <http://www.ogf.org/documents/GFD.98.pdf>
- [11] Accounting Facilities in the European Supercomputing Grid DEISA, J. Reetz, T. Soddemann, B. Heupers, J. Wolfrat eScience Conference 2007 (GES 2007) http://www.ges2007.de/fileadmin/papers/jreetz/GES_paper105.pdf
- [12] DART: <http://www.deisa.eu/usersupport/user-documentation/deisa-accounting-report-tool>
- [13] PRACE User Documentation: <http://www.prace-ri.eu/Documentation-and-User-Support>
- [14] Inca home page: Inca: <http://inca.sdsc.edu/drupal/>
- [15] BSCW - Basic Support for Cooperative Work: <http://public.bscw.de/en/index.html>
- [16] TWiki® - the Open Source Enterprise Wiki and Web 2.0 Application Platform <http://twiki.org/>
- [17] TRAC <http://trac.edgewall.org/>
- [18] Subversion <http://subversion.tigris.org>

List of Acronyms and Abbreviations

AAA	Authorization, Authentication, Accounting.
AMD	Advanced Micro Devices
API	Application Programming Interface
BAdW	Bayerischen Akademie der Wissenschaften (Germany)
BSC	Barcelona Supercomputing Center (Spain)
BSCW	Basic Support for Cooperative Work
CA	Certificate Authority
CEA	Commissariat à l’Energie Atomique (represented in PRACE by GENCI, France)
CINECA	Consorzio Interuniversitario, the largest Italian computing centre (Italy)
CINES	Centre Informatique National de l’Enseignement Supérieur (represented in PRACE by GENCI, France)

CLE	Cray Linux Environment
CP/CPS	Certification Policy and Practice Statement
CPU	Central Processing Unit
CSC	Finnish IT Centre for Science (Finland)
CSCS	The Swiss National Supercomputing Centre (represented in PRACE by ETHZ, Switzerland)
DDR	Double Data Rate
DN	Distinguished Name (as used in X.500 style directories)
DECI	Distributed Extreme Computing Initiative
DEISA	Distributed European Infrastructure for Supercomputing Applications. EU project by leading national HPC centres.
DIT	Directory Information Tree (LDAP)
DMA	Direct Memory Access
DPMDB	DECI Project Management Database
DRAM	Dynamic Random Access memory
EC	European Community
EP	Efficient Performance, e.g., Nehalem-EP (Intel)
EPCC	Edinburg Parallel Computing Centre (represented in PRACE by EPSRC, United Kingdom)
EPSRC	The Engineering and Physical Sciences Research Council (United Kingdom)
ETHZ	Eidgenössische Technische Hochschule Zuerich, ETH Zurich (Switzerland)
ESFRI	European Strategy Forum on Research Infrastructures; created roadmap for pan-European Research Infrastructure.
EX	Expandable, e.g., Nehalem-EX (Intel)
FC	Fiber Channel
FP	Floating-Point
FZJ	Forschungszentrum Jülich (Germany)
GB	Giga (= $2^{30} \sim 10^9$) Bytes (= 8 bits), also GByte
Gbps	Giga (= 10^9) bits per second, also Gb/s and Gbit/s
GB/s	Giga (= 10^9) Bytes (= 8 bits) per second, also GByte/s
GCS	Gauss Centre for Supercomputing (Germany)
GÉANT	Collaboration between National Research and Education Networks to build a multi-gigabit pan-European network, managed by DANTE. GÉANT2 is the follow-up as of 2004.
GENCI	Grand Equipement National de Calcul Intensif (France)
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
GigE	Gigabit Ethernet, also GbE
HP	Hewlett-Packard
HPC	High Performance Computing; Computing at a high performance level at any given time; often used synonym with Supercomputing
HT	HyperTransport channel (AMD)
IB	InfiniBand
IBA	IB Architecture
IBM	Formerly known as International Business Machines
ICE	(SGI)
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique (represented in PRACE by GENCI, France)

IGE	Initiative for Globus in Europe
IGTF	International Grid Trust Federation
I/O	Input/Output
ISTP	Internal Specific Targeted Project
IPP	Max Planck Institute for Plasmaphysics
JSC	Jülich Supercomputing Centre (FZJ, Germany)
KB	Kilo ($= 2^{10} \sim 10^3$) Bytes ($= 8$ bits), also KByte
KTH	Kungliga Tekniska Högskolan (represented in PRACE by SNIC, Sweden)
LDAP	Lightweight Directory Access Protocol
LRZ	Leibniz Supercomputing Centre (Garching, Germany)
MB	Mega ($= 2^{20} \sim 10^6$) Bytes ($= 8$ bits), also MByte
MB/s	Mega ($= 10^6$) Bytes ($= 8$ bits) per second, also MByte/s
MDT	MetaData Target
MFC	Memory Flow Controller
MFlop/s	Mega ($= 10^6$) Floating point operations (usually in 64-bit, i.e. DP) per second, also MF/s
MHz	Mega ($= 10^6$) Hertz, frequency $= 10^6$ periods or clock cycles per second
MIPS	Originally Microprocessor without Interlocked Pipeline Stages; a RISC processor architecture developed by MIPS Technology
Mop/s	Mega ($= 10^6$) operations per second (usually integer or logic operations)
MoU	Memorandum of Understanding.
MPI	Message Passing Interface
MPP	Massively Parallel Processing (or Processor)
NFS	Network File System
NIC	Network Interface Controller
NUMA	Non-Uniform Memory Access or Architecture
OS	Operating System
PCPE	PRACE Common Production Environment
pNFS	Parallel Network File System
PMA	Policy Management Authority
POE	Parallel Operating Environment
PoP	Point of Presence
POSIX	Portable OS Interface for Unix
PRACE	Partnership for Advanced Computing in Europe; Project Acronym
PRACE-PP	PRACE Preparatory Phase Project
PRACE-RI	PRACE Research Infrastructure
PRACE-1IP	PRACE First Implementation Phase
PRACE-2IP	PRACE Second Implementation Phase
PSNC	Poznan Supercomputing and Networking Centre (Poland)
QDR	Quad Data Rate
RI	Research Infrastructure
RT	Request Tracker – software for trouble ticket handling
RZG	Rechenzentrum Garching (RZG) of the Max Planck Society and the IPP (Germany)
SAN	Storage Area Network
SARA	Stichting Academisch Rekencentrum Amsterdam (Netherlands)
SGI	Silicon Graphics, Inc.
SNIC	Swedish National Infrastructure for Computing (Sweden)
SSH	Secure Shell

STFC	Science and Technology Facilities Council (represented in PRACE by EPSRC, United Kingdom)
TB	Tera (= 240 ~ 1012) Bytes (= 8 bits), also TByte
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
UHeM	National Center for HPC of Turkey (Formerly UYBHM)
UNICORE	Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.
VSb-TUO	Vysoká škola báňská – Technical University of Ostrava (Czech republic)
WCSS	Wrocławskie Centrum Sieciowo-Superkomputerowe (WCNS, Wrocław Centre for Networking and Supercomputing, Poland)
X.509	IETF standard used for certificates in Public Key encryption, RFC 5280

Executive Summary

The main objective of WP6 is the operation of an infrastructure of national HPC Tier-1 systems in Europe, which provides resources to users through PRACE calls for proposals for Tier-1 access (the DECI calls). This Tier-1 infrastructure is the continuation of the infrastructure built by previous DEISA projects and is now tightly integrated with the PRACE Tier-0 infrastructure as a result of the efforts of both the PRACE-1IP and PRACE-2IP projects. The services operated by this work package are based on the services listed in the Service Catalogue, which has been defined by PRACE-1IP-WP6. These services are accepted by all Tier-1 sites.

The second objective of this activity is the integration of new Tier-1 sites. Ten new sites will also provide resources for PRACE calls in the two year period of PRACE-2IP. The integration of these sites is at different stages, because some already provided resources for the first call, which started allocation in November 2011, and some will only provide resources for the last call, starting November 2012, because the system is not available yet. The status of these sites is discussed in section 3.

The management structure of operations is the same as adopted for the Tier-0 infrastructure and the same persons are responsible for particular services on both the Tier-0 and Tier-1 infrastructure. In general there is a close collaboration between operations of the Tier-0 and Tier-1 infrastructure, which is important because of the integration between the two environments. Operational procedures and facilities are also shared. Examples are the change management procedure and the helpdesk facility.

In this project period the DEISA infrastructure is successfully adapted to the PRACE infrastructure. The dedicated 10Gbps network is now operated by PRACE in close collaboration with GÉANT and some of the new sites are already connected too. The data management facilities based on GridFTP are supported by all sites. These facilities enable both high throughput data transfers between PRACE sites as well as between PRACE sites and external sites. For running jobs on the systems users have access to local batch systems and can use UNICORE to submit jobs remotely. A common user administration facility enables the distribution of information needed for the authentication and authorisation for PRACE services and the user needs only one interface for the registration. The helpdesk facility based on the RT facilities is fully migrated to the PRACE facility and the PRACE Common Production Environment (PCPE) is implemented on all sites. The monitoring facilities provided with the Inca tools are successfully migrated to PRACE, with interfaces to the PRACE helpdesk system and the PRACE maintenance information pages. All information for the installation and configuration of services was migrated and adapted in the first months to the PRACE internal wiki pages. The status of services is discussed in detail in section 2.

1 Introduction

The objectives of WP6 are:

- Provision and operation of comprehensive common services at the Tier-1 level, and integration of those Tier-1 resources and services in the PRACE distributed Tier-0 infrastructure;
- Integration of new Tier-1 services and systems

For each of these objectives a dedicated task is defined. Task 6.1 is responsible for the operation of the Tier-1 services and the achievements and status are described in section 2. Section 3 describes the achievements and status of task 6.2, Integration of new Tier-1 systems in the PRACE distributed RI.

Two successful all-hands meetings have been organised, in October 2011 in Amsterdam and in April 2012 in Bologna. 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.

2 Operation and coordination of comprehensive common Tier-1 services

This section describes the work done by the task WP6.1. The objective of this task is to continue the coordinated and integrated operations of the comprehensive and common Tier-1 services which have been provided through the DEISA infrastructure. These services include:

- Network services – network connectivity through GÉANT and a dedicated network provided by GÉANT, monitoring of the network, and collaboration with the GÉANT NOC;
- Data services – GridFTP and GPFS and data management tools;
- Compute services – Local batch schedulers, UNICORE, GRAM;
- AAA – PKI, user administration, accounting and interactive access;
- User services – Common production environment based on modules, first and second level support and a help desk, monitoring of applications availability, user documentation and advanced application production assistance;
- Monitoring – Service availability with Inca, maintenance information;
- Internal or generic services – Collaborative workspaces, software versioning facility, DECI project management database.

Based on the Service Catalogue, prepared by WP6 of PRACE-1IP, the set of services and software tools used was assessed by the Tier-1 sites in the first year of the 2IP project. The Service Catalogue lists and describes the complete set of operational services that the PRACE RI is providing, both for Tier-0 and Tier-1 services.

The purpose of the PRACE Service Catalogue is:

- To describe all PRACE operational services
- To define PRACE service categories, and classify all PRACE services accordingly

Three service classes have been defined: Core services, Additional services and Optional services. The details of this classification can be found in the current version of the PRACE Service Catalogue in Annex 4.1.

The Service Catalogue also provides for every service the actual product or products used. These products have been classified too and separate classifications are given for the Tier-1 and Tier-0 infrastructure, as the classification can differ between the two infrastructures.

The operational organisation uses two different structures, where the first is the organisation along the different service categories and the second along the sites forming the infrastructure. These two structures together form the matrix of the operational organisation as shown in Figure 1. The horizontal structure is needed because the PRACE services are tightly integrated among sites, so one team of persons needs to be responsible for the service over the whole infrastructure. The vertical level is needed because a site is responsible for the status of the local services. This organisational structure is shared between the Tier-0 and Tier-1 infrastructure and guarantees a transparent integration of the Tier-0 and Tier-1 infrastructures.

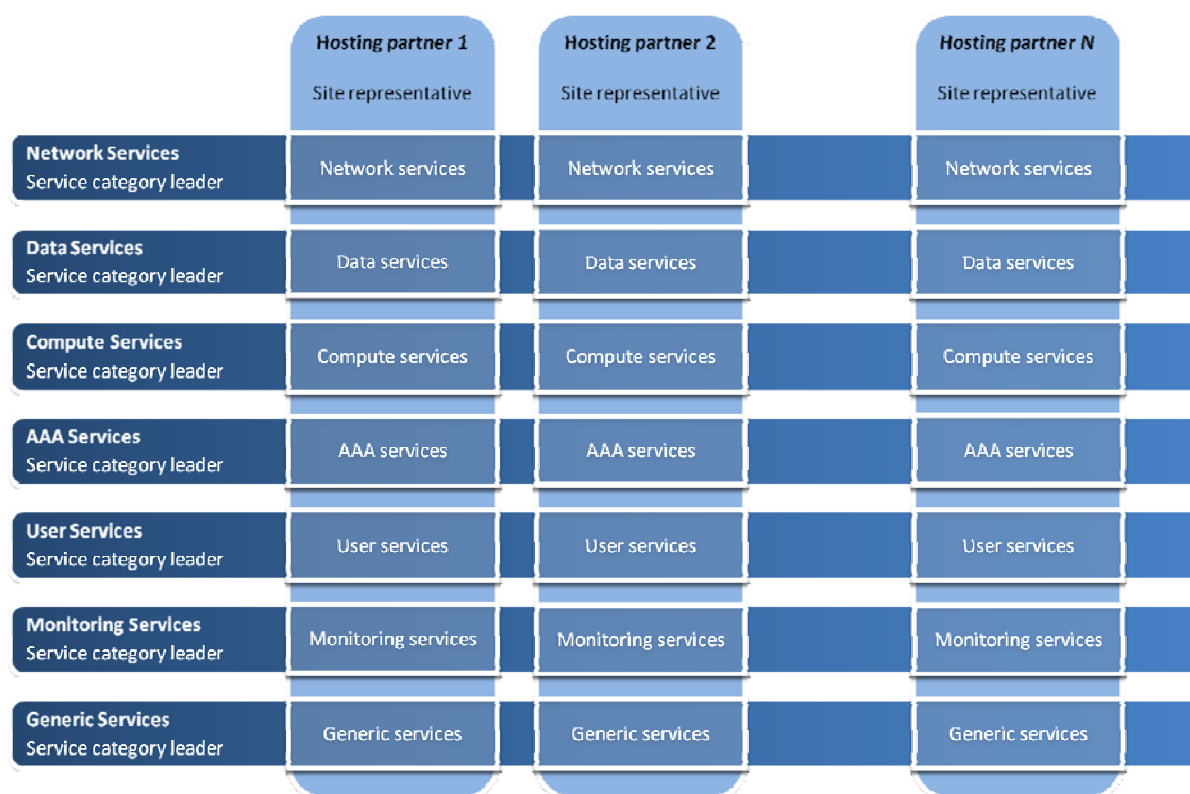


Figure 1 Structure of PRACE operational management

Bi-weekly video meetings are organized to discuss all operational issues: the status of the infrastructure, the status of proposed and planned changes and other issues. Attendance in general is very good; all operational sites and all service leaders are expected to attend. Minutes of all meetings are published on the BSCW pages.

In the next sections an overview of the available resources is given, followed by the status of the different services for this period.

2.1 Overview of Tier-1 infrastructure

Table 1 gives an overview of the available resources of the Tier-1 infrastructure by the end of the first project year, August 2012. Compared to the last available resources for DEISA, May 2011, the total compute capacity increased from 1822 to 3080 TF/s, an increase of almost 70 percent. Five or more percent of these resources are available for DECI calls on a yearly basis. Not all sites participate in each call, because of system upgrades and migrations, or systems that only will become available with DECI-9. Not all of the latter sites are yet included in Table 1. The number of sites increased from 12 in DEISA to 18 by the end of August 2012 and will increase to 22 by the end of 2012. Some sites also provide access to more than one system. The total amount of memory increased with 75 percent and the number of cores increased with more than 50 percent.

Site	Architecture	TF/s	#Cores	Total Memory (TB)
BSC	IBM PowerPC CPU: IBM/PPC970MP/2.3GHz Interconnect type : Myrinet	94.2	10240	20.0
CINECA	IBM Power6-575	101.0	5376	20.0
CINES	SGI Altix ICE 8200 CPU: -Intel/Xeon Harpertown Quad Core/3.0 GHz -Intel/Xeon Nehalem Quad Core/2.8 GHz Interconnect type: Infiniband DDR / Dual plane Hypercube	267.0	23040	91.0
CSC	Cray XT4 AMD-Opteron 4-core CPU: AMD / AMD Opteron quad-core / 2.3 GHz Interconnect: Cray SeaStar2 / 3D torus	32.3	4048	4.4
	Cray XT5 AMD-Opteron 8-core CPU: AMD / AMD Opteron 8-core / 2.3 GHz Interconnect: Cray SeaStar2+ / 3D torus	86.7	6816	6.1
CSCS	Cray XE6	400	47872	46.75
EPCC	Cray XE6	360	44544	59.4
	Cray XT4 CPU: AMD / Opteron /2.3GHz Interconnect: Cray SeaStar2 / 3D torus	114.4	12288	24
	Cray X2 CPU: Cray / X2 / 1.6 GHz Interconnect: Cray / fat tree + link to XT4 nodes using Cray SeaStar2 / 3D torus	2.9	112	0.9
FZJ	Bull Intel Nehalem-EP quad-core processors 2.93 GHz Infiniband QDR with non-blocking Fat Tree topology	207.0	17664	53.0
HLRS	NEC Intel Nehalem cluster dual socket quad-core @ 2.8GHz	62.0	5600	8.4
ICHEC	SGI Altix ICE 8200 CPU: Intel Xeon Woodcrest X5650 2.67GHz hex-core Interconnect: Infiniband DDR / dual plane Hypercube	41.0	3840	7.7
IDRIS	BlueGene/ P CPU: IBM, Blue Gene/P, 850MHz Interconnect: - 3 Dimensional BG/P-Torus	139.0	40960	20.0

Site	Architecture	TF/s	#Cores	Total Memory (TB)
	IBM Power6-575 CPU: IBM, Power6, 4.7GHz Interconnect: Infiniband/ flat tree	67.3	3584	10.8
KTH	Cray XE6 AMD Opteron 2354	305.6	36384	47.38
LRZ	IBM BladeCenter HX5 CPU: Intel / Xeon Westmere-EX / 2.4 Ghz Interconnect type: Infiniband QDR	78.0	8200	52.0
NCSA	BlueGene/ P CPU: IBM PowerPC 450, 850MHz Interconnect: IBM BG/P	27.85	8192	8.0
PSNC	SGI UV 1000 CPU: Intel Xeon E7-8837 Interconnect: NUMalink® 5, 2D torus	21.8	2048	16.0
	SGI/Rackable CPU: AMD Opteron 6234, Interlagos Interconnect InfiniBand QDR, Fat Tree	224.3	5448	10.6
RZG	BlueGene/ P CPU: IBM PowerPC 450 850MHz Interconnect: 3D Torus	54.0	16384	4.0
	IBM Power6-575 CPU: IBM Power6, 4.7GHz Interconnect: 8-link InfiniBand	120.0	6560	18.5
SARA	IBM Power6-575 CPU: IBM, Power6, 4.7GHz Interconnect: Infiniband/ flat tree	62.6	3328	15.3
UHeM	HP Proliant BL460 Cluster CPU: Intel Xeon 5550 – 2.67 GHz. Interconnect: Infiniband – 5 Stage CLOS	2.5	256	0.75
UiO	MEGWARE MiriQuid CPU: dual socket Xeon-2670 8C 2.6GHz Interconnect: Infiniband FDR/ fat tree	178.6	10080	39.37
VSB-TUO	System not installed yet	0	0	0
WCSS	HP Intel Westmere-EP cluster CPU: Intel Xeon X5650 six-core, 2.67 GHz Interconnect: Infiniband DDR fat-tree full-cbb	30.0	4848	9.6
Totals		3080	327712	594

Table 1 Overview of sites and Tier-1 systems in use.

Table 2 gives an overview of the DECI commitments of all Tier-1 sites from DECI-7 to DECI-9. For each call the start date of access to the resources is given.

Site (Country)	DECI-7 November 2011	DECI-8 May 2012	DECI-9 November 2012
BSC (Spain)	x	-	x
CINECA (Italy)	x	x	x
CINES (France)	x	x	x
CSC (Finland)	x	x	x
CSCS (Switzerland)	-	x	x
EPCC (UK)	x	x	x
FZJ (Germany)	x	x	x
HLRS (Germany)	x	x	-
Hartree (UK)	-	-	x
ICHEC (Ireland)	x	x	x
IDRIS (France)	x	x	x
IPB (Serbia)	-	-	x
KTH (Sweden)	x	x	x
NCSA (Bulgaria)	x	-	x
LRZ (Germany)	x	x	-
PSNC (Poland)	x	x	x
RZG (Germany)	x	x	x
SARA (Netherlands)	x	x	x
UHeM (Turkey)	-	x	x
UiO (Norway)	-	-	x
VSb-TUO (Czech Republic)	-	-	x
WCSS (Poland)	-	x	x

Table 2 Provision of DECI resources, an “x” indicates participation

2.2 Network Services

The PRACE network services are based on the developments done in the DEISA projects. Within the first year the main work has been to integrate the two worlds, Tier-0 systems and Tier-1 systems, into one homogeneous dedicated network infrastructure. The Tier-0 systems JUGENE at FZJ and CURIE at CEA, HERMIT at HLRS Stuttgart, and SuperMUC at LRZ Garching near Munich had to be connected to the PRACE backbone.

The backbone of PRACE has been changed slightly by sharing more links between different systems and/or sites/partners. Further additional Tier-1 systems have been connected, which are the CINES system at Montpellier, France, connected via RENATER, and Tier-1 facilities in Poland, connected through a 10 GE link via PIONEER to PSNC at Poznan. All those systems/ new connections were integrated into the multi domain network monitoring system of PRACE. Any relations to the former DEISA network were transferred / translated into PRACE, including network web pages, scripts, etc. An overview of the network status in August 2012 is shown in Figure 2.

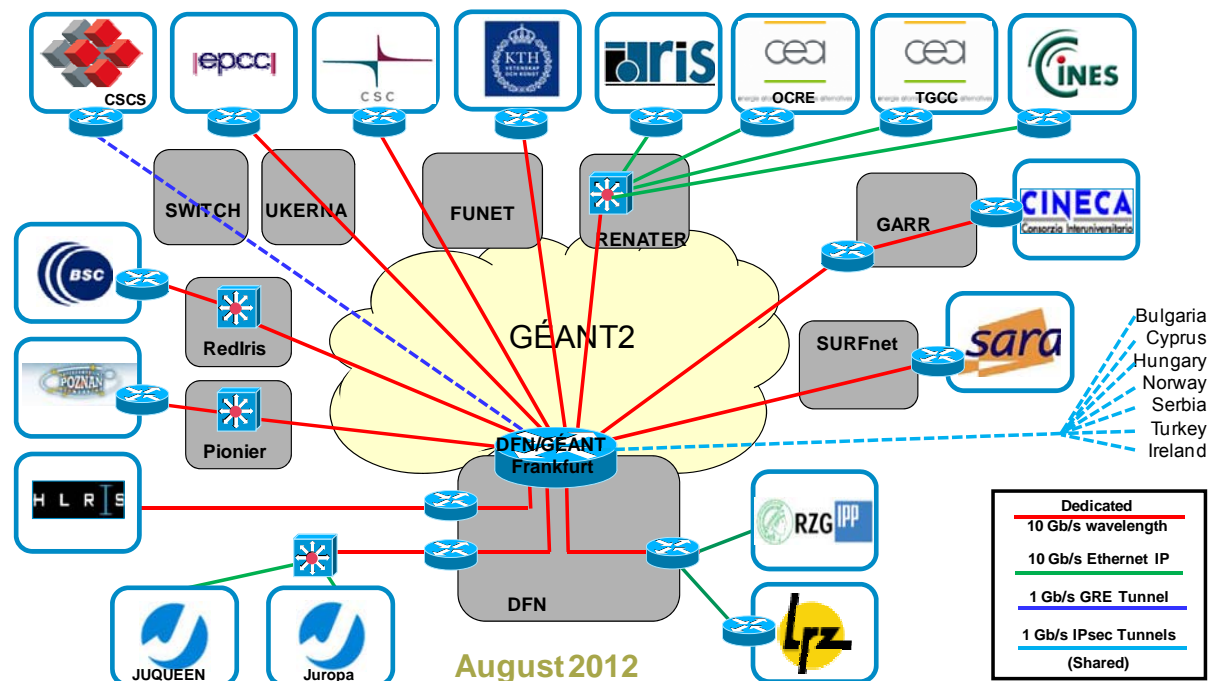


Figure 2 Overview PRACE network August 2012

The newly established network team needed a network operations mail list which was set up and maintained. Daily operational work had to be done, including coordination of a multi domain network staff working group. Network staff engaged in solving of network problems, i.e. interrupts because of network hardware, fibre cuts, interface errors, low throughput, routing problems, errors in network configurations, firewall settings, i.e. ports which had not been opened, security break ins as well as global network status monitoring and logging. Within the reporting period 25 network maintenance warnings have been received. Ten of these did not have any impact on the network itself since the links have been optically protected. 15 have been real maintenance sessions separating PRACE partners from the rest of the PRACE dedicated network; these maintenances all have been announced in advance and mostly been scheduled during night times (22:00 – 6:00 CET). Only three of those have been emergency maintenances during working hours. Additionally two real outages (fibre breaks) have occurred and which needed intensive error handling, finding the source of the outage. Both outages were solved within a 24 hour time frame. Thanks to the star topology of

the PRACE network only single sites did not have connectivity to the rest of the PRACE partners.

Because of a reduced budget the former design plan for the PRACE dedicated network could not be extended to all new sites, so that a new concept had to be established. This redesign required extensive discussions with all involved partners. The new design foresees generating 1 Gbit/s IPsec tunnel from partners with smaller systems to the central PoP of PRACE in Frankfurt providing access to the rest of the dedicated PRACE network infrastructure. In case of enhanced bandwidth needs in the future, some of these partners may change to the originally envisaged 10 Gbit/s dedicated wavelength solution. A detailed discussion of this new network set-up will be described in the M15 deliverable D6.2 "Network Capacity Planning" scheduled for November 2012.

2.3 Data Services

GridFTP is a data transfer protocol that can fully utilize the high bandwidths between the PRACE Computing centres, so it has been chosen as standard for data transfers between the sites.

GridFTP supports parallel TCP streams and multi-node transfers to achieve a high data rate via high bandwidth connections. Furthermore, third-party transfers can be established, which is very useful for the PRACE users.

In the PRACE-2IP project it has been agreed that GridFTP is a core service at all Tier-1 sites which have a direct connection to the dedicated PRACE network.

2.3.1 Status of Deployment

Table 3 gives the status in August 2012 for the sites have had deployed GridFTP Services for their Tier-1 system.

Site	Version (GridFTP/GT)
BSC	6.10/GT5.2.1
CEA (Inti)	6.05/GT5.2.0
CINECA	3.33/GT5.0.4
CINES	3.33/GT5.0.4
CSC	3.28/GT5.0.3
EPCC	3.28/GT5.0.3
FZJ (JuRoPa)	3.15/GT4.2.1
HLRS (Laki)	3.33/GT5.0.4
ICHEC	6.10/GT5.2.1
IDRIS	6.10/GT5.2.1
RZG (VIP)	2.7/GT4.0.7
SARA	6.10/GT5.2.1
UHeM	6.10/GT5.2.1
WCSS	6.05/GT5.2.0

Table 3 Overview installed versions of GridFTP and Globus Toolkit

The up-to-date status of the deployment is maintained on the inventory page on the PRACE Wiki.

During the first project year of PRACE-2IP, GridFTP has been successfully deployed on the following new Tier-1 sites: CINES, ICHEC and WCSS. On the Inti system at CEA the GridFTP service has been implemented for testing purposes. This system is not available for Tier-1 users.

For HLRS the current server structure has been replaced in July 2012. The GridFTP service will be split, which means that frontend and backend components of GridFTP will reside on distinct machines for improved security.

2.3.2 Documentation

The installation and configuration documentation at the PRACE Wiki has been updated for the new versions of the *GridFTP* Server (6.10) and the Globus Toolkit (5.2.1). Since there is a security bug in this version, discussed below, a reference to the corresponding Security Log information is provided.

2.3.3 Security

On 17 May 2012, a security bug [2] in GridFTP/GT was disclosed in the unpatched GT5.2.1 version of GridFTP: If a user's DN points to a non-existing or misspelled uid entry for the target system, then the user is mapped to the last entry of the */etc/passwd* file.

All partner sites have been informed to check this, a workaround has been proposed (introduce a dummy user with no home dir and no login shell) and also a patch is available in [2] that has to be implemented. In the PRACE Wiki the status of the vulnerability and of the implementation of the fix is logged.

2.3.4 Advanced User Tools

gtransfer [3] as wrapper-tool for *globus-url-copy* has been developed and undergone the service certification process in PRACE-1IP. Current state of implementation can be found in the PRACE wiki as well as the state of the evaluation by PRACE-1IP T6.3. The tool has been successfully deployed for data transfers between CSC in Finland and NERSC in the US.

The ISTP (Internal Specific Targeted Projects) process has been initiated to establish and deploy *gtransfer* as a reliable production service within PRACE-2IP Subtask T6.1.2.

2.3.5 Multi-cluster Shared File System – MC-GPFS

PRACE-2IP took over the GPFS-infrastructure from DEISA. The naming scheme for user-directories has been adjusted to reflect new PRACE-conventions. RZG is hosting file systems for all sites which don't provide their own servers. Several configuration-parameters still reflect the DEISA past, since they are more difficult to change, like the mount points. Due to upgrades and exchange of systems the currently available common file-system infrastructure is heavily reduced. Table 4 and Table 5 give an overview of the installed servers and clients, respectively.

GPFS IO cluster	file system	GPFS version	Platform / OS
IDRIS	/deisa/idr	3.3.0.22	IBM Power 6 / AIX
RZG	/deisa/rzg	3.4.0.13	IBM Power 6 / AIX
SARA	/deisa/sar	3.3.0.20	IBM Power 6 / Linux

Table 4 MC-GPFS installations in PRACE – GPFS I/O clusters

Compute system	GPFS version	Platform/OS
CSC LOUHI	3.3.0.22	Cray XT4/XT5 / Linux
FZJ JuRoPa	3.4.0.14	Sun, Intel Xeon X5570 / Linux
IDRIS Babel	3.3.0.22	IBM BG/P / Linux
RZG VIP	3.4.0.13	IBM Power 6 / AIX
RZG Genius	3.3.0.12	IBM BG/P / Linux
SARA Huygens	3.3.0.20	IBM Power 6 / Linux

Table 5 MC-GPFS installations in PRACE – systems with GPFS mounted file systems

The future of the common file-system is uncertain, since the fraction of sites and systems able to participate is decreasing with more and more partners joining the Tier-1 infrastructure but not being connected to the 10Gbit/s dedicated backbone network. The full deployment of the VPN-tunnel, as described in section 2.2 will resolve this issue. Furthermore, the currently connected systems are going to be removed in the near future. Thus more and more, data-transfer tools must be used, since there is no storage provision for job output for a longer period than a short transfer time after the calculation. Although network capability is increasing this may become a critical issue in the future.

2.4 Compute services

All technical activities in this project period have addressed three main services that are responsible for the provision of job management features to PRACE DECI users. Two of them are considered as core services and one as optional:

- Local Batch Systems (core)
- UNICORE (core)
- Globus GRAM (optional)

Local Batch Systems allow users to interact directly with a specific job management subsystem when logged in on a particular system. Understanding the features provided by a specific batch system can require a considerable amount of time especially in the case where a user uses different resources controlled by different batch systems. In spite of this, it seems that users in most cases prefer the option to use the local batch system.

An alternative way of interaction is based on UNICORE, which offers a uniform interface to all systems belonging to the Tier-1 infrastructure. UNICORE is considered as a core service, this means that it must be available on all systems for providing users with an alternative option for managing their jobs. The importance of having a uniform interface like UNICORE becomes more relevant as more systems with different local batch environments are used.

Globus GRAM5 is considered as an optional service, i.e. each PRACE Tier-1 site optionally decides to provide it although service availability and long-term support is not guaranteed by PRACE. This facility mainly is provided by sites if needed for projects which also access other systems where Globus GRAM facilities are used.

Deployment of compute services involves 18 Tier-1 systems which have provided resources to users up to the DECI-8 call.

2.4.1 Local Batch Systems

A Local Batch System is a software product providing resource management and job scheduling features for an HPC system. These two features are typically implemented by separate, but tightly coupled, software components. They come built-in in a single software product, like IBM LoadLeveler and Altair PBS Pro, or separately as in the case for Torque (Resource Manager) and Maui (Job Scheduler).

Responsibility of the job scheduler is to act as front-end for the users and the resource manager, listening for job submission requests from users, putting submitted jobs in queues and compute a schedule to send to the Resource Manager when asked. The Resource Manager is responsible for managing the job execution, interacting with the underlying parallel environment (e.g. different MPI implementations, POE, etc...) monitoring resource usage and asking for a new schedule once new resources become available.

Figure 3 shows the distribution in PRACE of these two software components. It shows prevalence of bundled solutions with for the IBM LoadLeveler product a share of almost 40%. This means that seven systems, out of 18, have LoadLeveler installed. PBS Pro and Torque with Moab follow with four systems for each. Slurm is the only open source software solution providing resource allocation and scheduling capabilities; its share is about 10%.

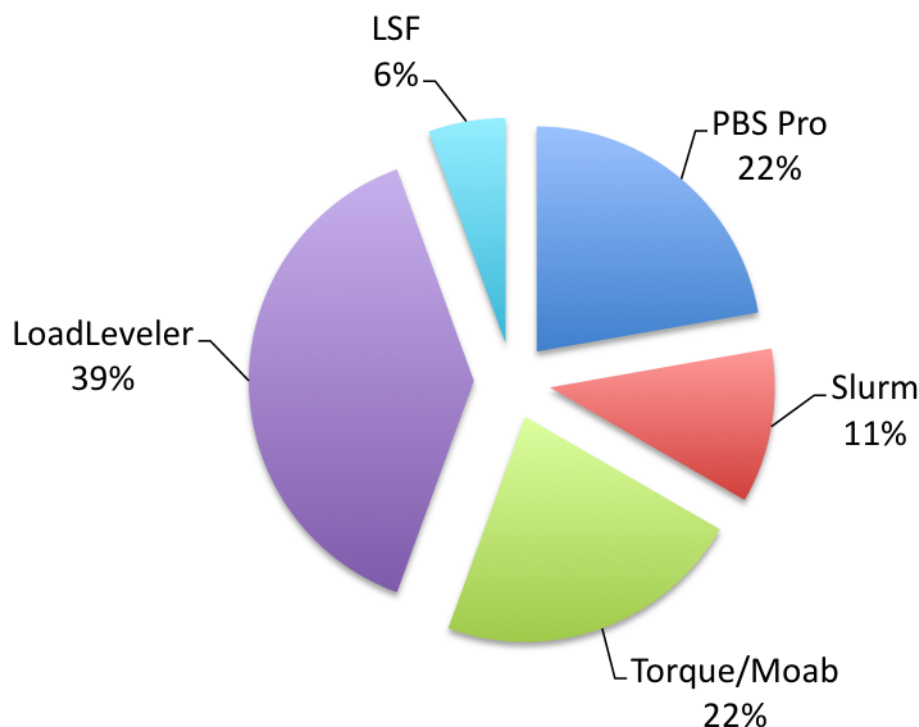


Figure 3 Batch Systems distribution on PRACE Tier-1 systems

The choice for a Batch System, especially for the Resource Manager component, depends on the underlying hardware layer and the parallel software stack. Sometimes specific software customisations (e.g. wrappers and/or software extensions) are also applied for improving, for example, the user interface and/or for fully exploiting the hardware features. This limits the coherency of the batch environments inside a supercomputing infrastructure like PRACE with systems with different hardware architectures.

Within DEISA many efforts were made to set up and maintain a batch system inventory with a mandatory set of features to be supported. The final result was a very limited set of basic features, mainly due to the heterogeneity of the batch systems and the customizations applied by single sites in choosing the right configuration for their job management system and the underlying hardware layer.

The objective is to have and maintain, first of all, a clear and up-to-date overview of the distribution of the batch systems. For the second year it is planned to set up communities of system managers around a specific batch system solution. This strategy allows partners to share their knowledge and get support from other partners about the setup of resource managers and schedulers, in particular for advanced configurations.

2.4.2 UNICORE

The UNICORE default version in PRACE is v6.4.2 and an upgrade to version 6.5, released on May 2012, is planned, to be completed in the second project year.

Deployment of UNICORE is based on a central registry hosted by FZJ and mirrored by CINECA to improve fault tolerance. Sites publish the services they provide (target systems, storage, applications, file transfer protocols, etc...) to the central registry. The central registry acts as a directory service for PRACE users by displaying all resources where they could have access. Users can get access to the central registry and then operate on different end systems through a graphical client (URC) and/or a command line client (UCC). A very basic design of the infrastructure with the UNICORE components is shown in Figure 4.

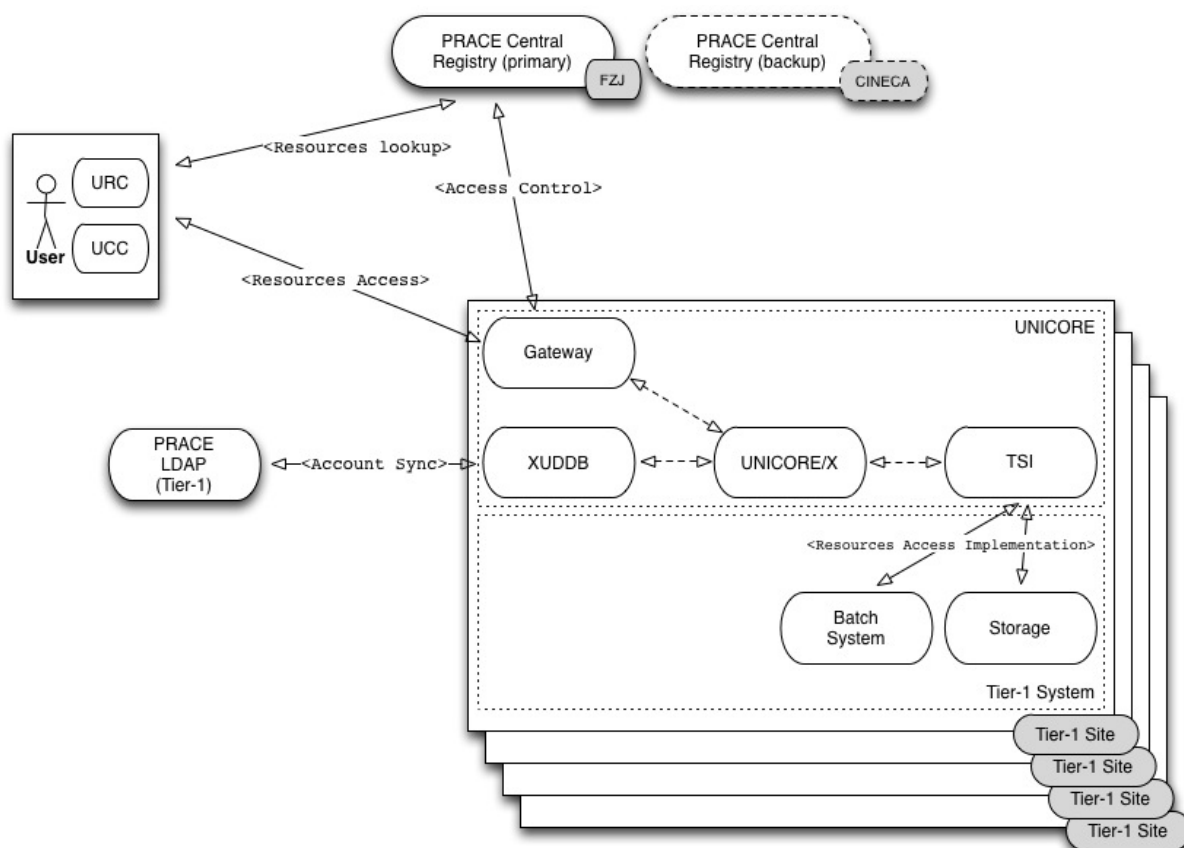


Figure 4 UNICORE architecture for PRACE Tier-1 systems

The deployment rate of UNICORE on Tier-1 sites is currently about 60%. This is mainly due to the integration of new Tier-1 systems, which is not yet fully completed. It is expected to

reach full deployment by the end of the 2012, one month after the start of the DECI-9 allocation period. Figure 5 shows the current distribution of UNICORE versions.

For the second year the maintenance of the infrastructure is planned – upgrades of the software, integration of new systems and phasing out of systems – and effort will be spent on the quality control and the assurance methodology, which was defined by WP6 of PRACE-1IP.

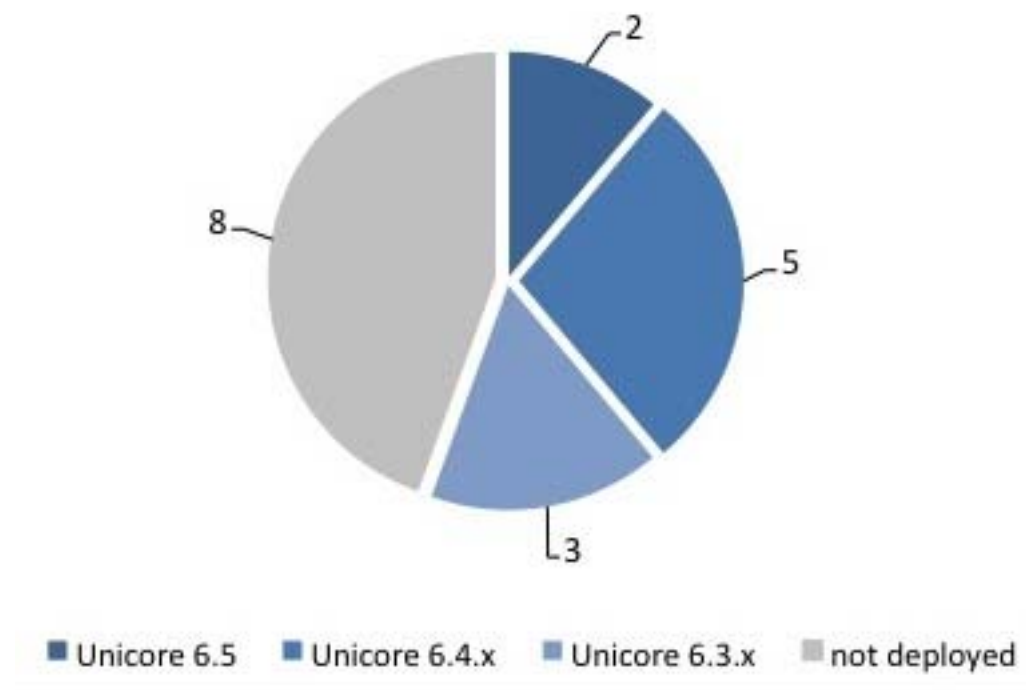


Figure 5 UNICORE deployment on PRACE Tier-1 systems

2.4.3 GRAM5

GRAM5 is considered an optional service for Tier-1 systems. Although it can work with the main batch systems its level of adoption suffers because LoadLeveler is not among the natively supported services. In the past, a solution has been developed by the IGE project [2] and this is planned too for new GRAM versions in the proposed MoU with IGE.

The main issue, especially from the performance point of view, is the efficient retrieval of the status of the running jobs. Interactions with different local Resource Managers, which are responsible to provide information about the status of jobs, have been tested and deployed except for LoadLeveler for which the parsing of the log file is still under investigation.

End of August three Tier-1 sites provide the GRAM5 service.

2.5 AAA services

The AAA activity is responsible for services which provide Authentication, Authorization and Accounting facilities on the infrastructure. This includes the support of a Public Key Infrastructure for authentication, a user administration service for authorisation and interactive access services. A general activity was the migration of information about the installation and configuration of services from the DEISA BSCW and wiki to the PRACE BSCW and wiki facilities and the adaptation of this information to the PRACE infrastructure.

2.5.1 *Public Key Infrastructure - PKI*

Several PRACE services rely on X.509 certificates [5] for the authentication and the authorization. These certificates must be issued by entities which are trusted by the service providers. PRACE relies on the Certificate Authorities (CA) accredited as a member by the EUGridPMA, the European Policy Management Authority [6], or by one of the two sister organizations TAGPMA and APGridPMA, all three federated in the IGTF [7]. These PMAs all require a minimum set of requirements for the CP/CPS of the member CAs, as published in a profile document. PRACE is a Relying Party member of EUGridPMA, which is important for the discussion of problems and requirements that we have with the deployment of the PKI.

For PRACE a distribution of CA information is maintained at a central repository [8]. The distribution is provided in several formats because services have different requirements for the presentation of the information. This distribution is based on the monthly release of a new distribution by the IGTF. Sites are informed about new updates and must update the local repositories used by services.

2.5.2 *User Administration*

Information about users and their accounts is maintained in an LDAP based repository. This facility is used to update the authorization information needed by services and can be used to retrieve information about users and the projects that they are affiliated to. Authorization information is provided among others for interactive access through GSI-SSH, job submission with UNICORE, accounting services and access to the helpdesk facilities.

At the start of the PRACE-2IP project the existing LDAP infrastructure for the former DEISA Tier-1 sites has been integrated with the existing LDAP infrastructure for Tier-0 sites. The integration was relatively easy because both facilities used the same LDAP schemas. The resulting LDAP infrastructure for Tier-0 and Tier-1 accounts is shown in Figure 6. The top part of the name space, which is called the suffix, for PRACE Tier-0 accounts is “ou=ua,dc=prace-project,dc=eu” and for Tier-1 accounts “ou=ua,dc=deisa,dc=org”. This means that two databases are used, however because the same LDAP schemas are used the same tools can be used to update and retrieve information. The difference in suffix only has an historical reason and doesn’t have any functional reason.

Each Tier-1 site 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 for the user.

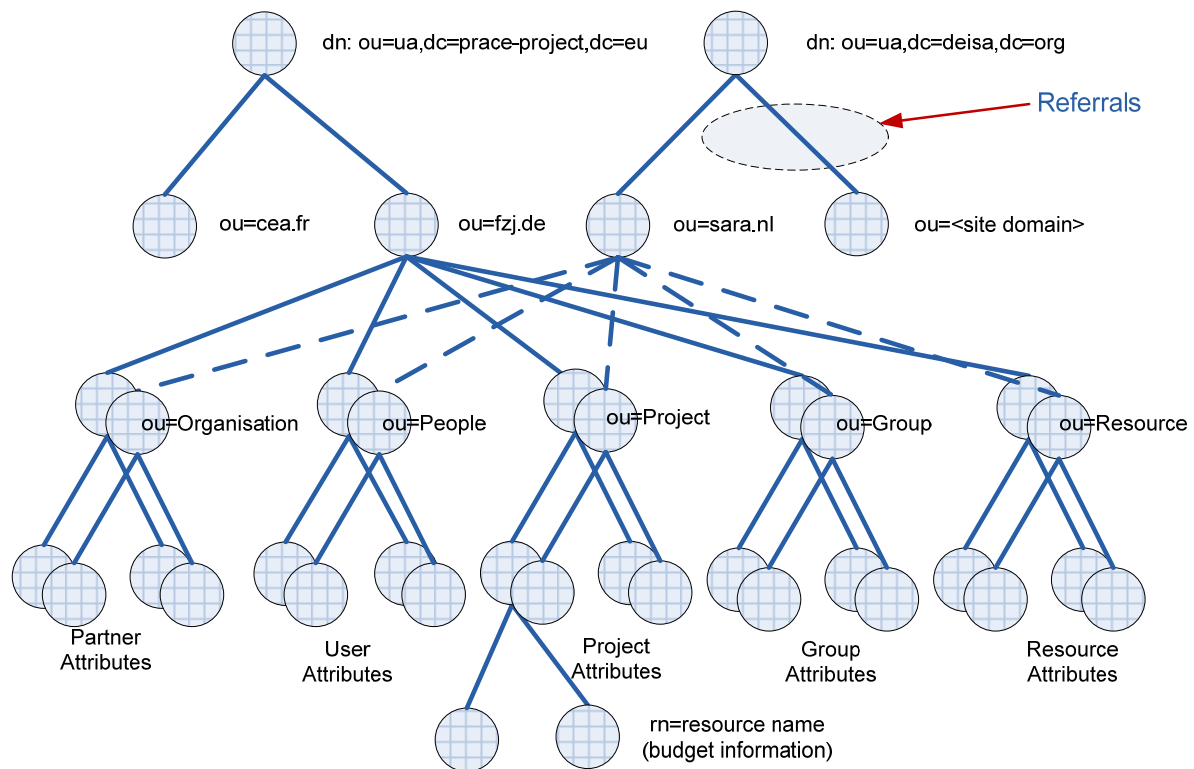


Figure 6 PRACE LDAP directory tree

In the DEISA set-up each Tier-1 site was operating its own server which hosted the branch for that partner. All information can be retrieved using LDAP referrals (see Figure 6). It was agreed at the start of PRACE-2IP that sites should migrate their branch to the main server, which also hosts the Tier-0 accounts. This doesn't change anything to the management structure of the accounts; each site is still responsible for its branch. The site can decide when to migrate, e.g. if a server has to be replaced anyway because of end of life. For new Tier-1 sites branches are created at the main server, so no additional servers have to be set up by these partners. In total 20 branches are operational, which includes seven sites that joined the Tier-1 infrastructure with the start of PRACE-2IP.

The main server is operated by SARA and for high availability a local replica server is operational too. Since early 2012 a backup server is operational at HLRS. This server contains all data from all PRACE LDAP servers and can be used in case the production servers at SARA cannot be reached.

Several changes in the LDAP information have been implemented in this project period. At the start of PRACE-2IP, with the expected addition of several new sites, it was realized that the agreement on the values to be used for the user names on the systems (the login names and numbers) could not be supported any longer. For this reason a new set of reserved character combinations was accepted, which allows an easy addition of new sites. All PRACE accounts start with the four character combination `pr[0-9][a-z]` and each site has a unique combination of the third and fourth character. All sites guarantee that these four character combinations are not used for local accounts. The only complication can be that a new site already uses such a combination. All new Tier-1 accounts have used the new values and existing accounts with the old values weren't used anymore for new DECI projects.

It was agreed to add to the LDAP repository information about authorisations for accessing the database which contains information for the DECI projects, the DPMDB. Two values

“deci-db-r” and “deci-db-w” are introduced for respectively giving read or read and write access to the database. These are values for an already existing attribute which can have multiple values, so no change in LDAP schemas was needed. Sites can now easily update this information for their staff and an automatic procedure can be used to update the authorisation information for the DPMDB.

The LDAP repository also contains personal information like e-mail address and telephone number of the user. For privacy reasons it was agreed that personal data can be removed after that an account is not used anymore. Basic account information must be kept because it's a requirement that an account may not be used again for another user. For the implementation of this policy a change of the DEISA schema attribute deisaDeactivated was accepted. The attribute deisaDeactivated is no Boolean attribute anymore but will have a string value of “active” for an active account, “closed” for an inactive account and “deleted” for an account for which personal data is removed. The Home site is responsible for the status of deisaDeactivated. The site must change status to “deleted” as required by legislation and internal rules. Other sites also must follow the local rules for the deletion of personal information after an account is deactivated, also if the status is not changed to “deleted” in LDAP.

The PRACE AAA Administration guide has been updated with the changes discussed above.

2.5.3 *Interactive access*

Interactive access to the Tier-1 systems is a basic requirement. For the home site this is in general provided using the SSH (Secure Shell) facilities or the GSI-OpenSSH facility of the Globus Toolkit [9], or GSI-SSH or gsissh for short. This is a core service and provided for all systems.

GSI-SSH is also used for interactive access between PRACE sites. GSI_SSH uses X.509 certificates for the authentication, which means that no exchange of passwords or ssh keys between the user and the site is needed before a user can have access. The authorisation for interactive access is based on information distributed through the LDAP facilities. Twelve sites provide access in this way from other PRACE sites. Three sites - LRZ, RZG and SARA – also provide interactive access from outside the PRACE infrastructure with GSI-SSH for all PRACE users. From these sites users can then access their final destination sites using the internal GSI-SSH service.

GSI-SSH_Term, a GSI-SSH JAVA based client for use on the end user systems, is supported by the PRACE partner LRZ.

2.5.4 *Accounting services*

Information about the usage of resources is important for users, Principal Investigators (PIs), partner staff and the resource managers, e.g. the DECI call team. PRACE provides facilities to publish and display usage with the following characteristics: 1) the usage of resources is published in a common format, which follows the recommendations of OGF's UR-WG (Usage Record Working Group) [10]; 2) access is based on the authorizations of the requestor, e.g. a normal user can only see his/her personal usage while the principal investigator of a project can see the usage of the whole project. Detailed information about the design considerations can be found in [11].

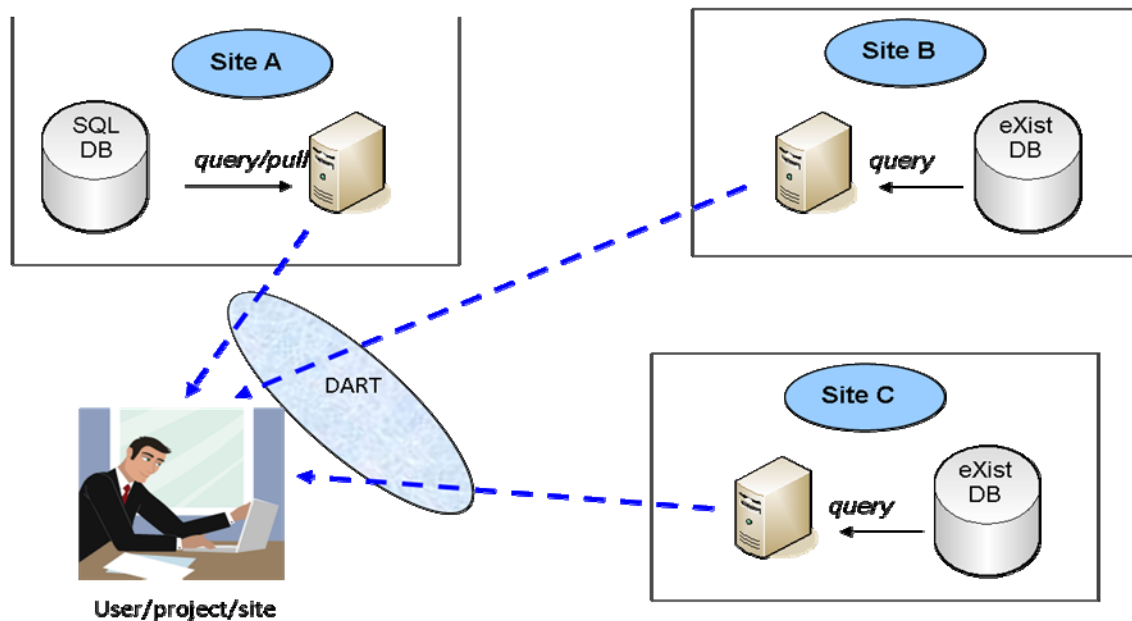


Figure 7 Accounting architecture

Figure 7 shows the basic setup of the facilities. Each site stores usage records for PRACE users in a local database, this can be an eXist database (sites B and C in the figure) or an SQL based database (site A). An Apache/CGI web interface is available which will publish data to authorized clients. The authorization is based on X.509 certificates and the access rights are given by the attribute `deisaAccountRole` of the user administration service. DART [12] is a Java Webstart tool, which can be used by a client to retrieve and to display the information from different sites.

11 sites are fully operational. Most of the new Tier-1 sites are still in the implementation phase of the service.

2.6 User Services

2.6.1 PRACE Common Production Environment

The PRACE Common Production Environment (PCPE) distribution has been updated to allow for installation on Tier-1 as well as Tier-0 sites. Due to the larger range of architectures present at the Tier-1 level compared to Tier-0 this has entailed introducing more flexibility for individual sites on deciding how the components of the PCPE are implemented locally.

The main activity in the period was the configuration of the PCPE on the new Tier-1 sites. Work was also completed in the period to integrate the monitoring of the PCPE status into the PRACE Inca monitoring tool.

In addition to deploying the PCPE on the new Tier-1 sites, additional user functionality was added to the PCPE. During the period a new service called ‘`prace_service`’ was implemented. The ‘`prace_service`’ command is of benefit to users when running any Globus based commands. The user may implement the Globus command without the need to know the details of the destination server or port numbers. For example, `gsissh `prace_service -i -s sara`` would invoke a `gsissh` session for the user on the Tier-1 system at SARA, automatically using the correct server and port details.

A second utility ‘`bolt`’ was developed and deployed, with the aim of simplifying the generation of user batch submission scripts. The idea behind `bolt` is that it will generate the correct job submission script for the execution site a user is working on with a common

syntax across all PRACE execution sites. The tool also has the potential to try to divide up the work in a pseudo-optimal way if a user decides to under-populate a compute node. Further functionality may be added to the PCPE as and when a need is identified, following the agreed work package Change Management processes.

2.6.2 *User Documentation*

The User documentation for PRACE is available online on the PRACE website [13]. The PRACE documentation is subject to change as new Tier-1 sites come online, or when existing sites deploy a new or modified architecture. All such changes to the documentation are managed following the agreed PRACE Documentation Review Panel procedures. Work will continue in PRACE-2IP to facilitate the generation of downloadable PDF versions of the documentation, in addition to the static web pages currently available.

2.6.3 *PRACE Trouble Ticket System*

The centralised PRACE Helpdesk was deployed initially for Tier-0 sites as part of PRACE - 1IP. Those Tier-1 sites providing compute resources for the DECI-7 and DECI-8 calls have now also been integrated to the helpdesk by WP6.

Further changes to the configuration will be completed to include the new sites involved in DECI-9. Additional queues within the helpdesk will be configured to support any further sites as and when required, usually in line with the announcement of each DECI call.

Support staffing effort for the Helpdesk is provided by the PRACE partners on a rotational basis, with each site manning the Helpdesk for one week at a time. A number of the new Tier-1 sites now form an integral part of the rota, providing user support for PRACE, with the support assistance from one of the original sites. The support site assistance mapping is detailed further in Section 3.1.

2.7 **Monitoring Services**

The monitoring subtask implements, deploys and operates tools for monitoring of PRACE infrastructure services. The subtask focuses on availability and functionality monitoring of services from infrastructure and user perspectives. For this PRACE utilizes the user-level monitoring application Inca [14].

The main activity was the integration of Tier-1 sites in the existing Tier-0 monitoring infrastructure as set up by PRACE-1IP. The configuration was extended by new Tier-1 sites and resources as well as removing Tier-1 resources which went out of business.

The implemented changes include the set-up of the PRACE Common Production Environment (PCPE) test suite for all sites. Furthermore, old resources, like HLRS NEC-SX-9 and CINECA SP6 have been removed from the Inca monitoring infrastructure.

In addition, new resources and middleware tests have been added, e.g. for the Globus services at EPCC (HECToR), CINECA, BSC (MARENOSTRUM), CINES and CEA as well as for UNICORE services at BSC (MARENOSTRUM), CEA (CURIE and INTI) and EPCC (HECToR).

Further adjustments to the existing configuration for different resources and test suites have been conducted to improve the Inca monitoring.

Details about all changes, including the affected sites and resources, are documented in Change Management documents in the internal PRACE wiki. The list of major changes is:

- Switch to new PRACE user in GSISSH test suite;
- Configure PRACE Common Production Environment (PCPE) tests (Java, Perl, Python, TCL, TK, gmake);
- Extension of UNICORE test suite (BSC (MARENOSTRUM), CEA, EPCC (HECToR));
- LDAP availability and audit series configured for NCSA, PSNC, UHeM and WCSS;
- Integration of FZJ JuRoPA in the monitoring infrastructure;
- LDAP tests: Fine-tuning of configuration (timeouts, test interval);
- Switch off monitoring for resources which went out of business (HLRS NEC-SX-9, CINECA SP6);
- Set-up Globus GSISSH and GridFTP service tests for various resources (EPCC (HECToR), CINECA, BSC (MARENOSTRUM), CINES and CEA);
- various modifications of existing configuration.

2.8 Internal services

For sharing documents among the WP6 partners the PRACE internal collaborative workspace based on BSCW [15] is used.

A wiki server, based on the TWiki® software [16], is used to share information among partners for the installation and configuration of the production services and for other operational activities which need regular updates:

- Change management;
- Maintenance information
- Operational security;
- Help-desk duty reports
- Operator on duty reports
- User documentation reviews

The service is hosted by FZJ and is shared between the Tier-0 and Tier-1 infrastructure. The service is only used for internal communication.

Based on the Subversion software [18] a software version control service is used for the distribution and management of software tools and configuration files that are maintained by PRACE partners. The service is based on the TRAC software [17], which integrates the Subversion software. The service is hosted by SARA and is shared between the Tier-0 and Tier-1 infrastructure.

WP6 is responsible for the support of the DPMDB facility, the database which is used to maintain information about the DECI projects. The DPMDB server is hosted by SARA. The software to manage the database was developed by the DEISA project.

On the request of some other work packages WP6 will take the responsibility for the hosting of some PRACE facilities, like the public web site, the training portal and the peer review database for Tier-0 projects. This only applies for the underlying server environment, including the software used to provide these services, but not the content provided through these services. The latter will be the responsibility of other activities. But basic requirements, like availability and security will be the responsibility of operations.

3 Integration of new Tier-1 systems in the PRACE distributed RI

3.1 Introduction

This task is responsible for the integration of new Tier-1 sites into the PRACE Tier-1 infrastructure which has been created through integration of the DEISA infrastructure. Ten new partner sites have committed resources for the first three DECI project calls issued by PRACE: DECI-7, started November 2011, DECI-8, started May 2012, and DECI-9, which will start November 2012. Table 6 gives an overview of the new sites and their commitment of resources for these three DECI calls. For each new site a support site is assigned. These act as a contact point for the integration process and are shown in the second column of the table.

Four new sites already provided resources for the DECI-7 call, by which access was provided already in November 2011. These four sites had to prepare in a very short period the facilities to give access to the users of this call. For instance the integration of sites in the PRACE dedicated network alone needs at least some months because of the coordination needed with external service providers (GÉANT, NRENs). It was clear that a full integration by the first of November was not possible. This did not prevent using the resources of these sites by DECI users, but not all PRACE services were available.

Site	Support Site	DECI-7	DECI-8	DECI-9
CINES (France)	IDRIS (France)	x	x	x
PSNC (Poland)	CSC (Finland)	x	x	x
ICHEC (Ireland)	EPCC (UK)	x	x	x
NCSA (Bulgaria)	CINECA (Italy)	x	-	x
WCSS (Poland)	CSC (Finland)	-	x	x
UHeM (Turkey) (Formerly UYBHM)	LRZ (Germany)	-	x	x
UiO (Norway)	FZJ (Germany)	-	-	x
IPB (Serbia)	CINECA (Italy)	-	-	x
VSb (Czech Republic)	HLRS (Germany)	-	-	x
Hartree (UK)	EPCC (UK)	-	-	x

Table 6 Provision of DECI resources by new partners, x = participation

Specific documentation is available to support new Tier-1 sites with their integration in the PRACE infrastructure. This information was prepared by the DEISA project and was

presented at the PRACE-2IP kick-off meeting in Barcelona in September 2011 and the WP6 all-hands meeting held early October in Amsterdam.

A problem was that most documentation on the installation and configuration of services only was available on the DEISA internal web servers, which were not accessible for the new Tier-1 partners. For this reason all documentation was migrated and adapted to the PRACE servers with high priority. The progress of the integration of sites is maintained in a dedicated internal wiki section and the status of the sites is discussed in the next section.

3.2 Overview and status

For each new site that participated in the DECI-7 or DECI-8 calls a short description is given and the status of the integration is provided using Table 16 from the Service Catalogue, which gives an overview of all services and products used. Services provided by just a single site have been deleted from the table, e.g. the PRACE help-desk. UiO and VSB-TUO, two of the sites which will start providing access with DECI-9 have provided their planning for the integration too.

3.2.1 CINES

CINES provides to the PRACE community a Tier-1 system called JADE. JADE is composed of two homogenous cluster partitions based on Intel processors. Some dedicated resources are provided to host several PRACE services and integrate the machine to the PRACE grid infrastructure. Table 7 shows the status of the integration for CINES.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Available (with a dedicated 10Gbe link)
Network management, Monitoring	Network monitoring	core	Available
Data	Data		
Data transfer, storage & sharing	MC-GPFS	optional	Not planned for now
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Planned for September 2012
Compute	Compute		
Uniform access to HPC	Local batch systems	core	Available
Uniform access to HPC	UNICORE	core	Planned for September 2012
Uniform access to HPC	Globus GRAM	optional	Planned for October 2012
AAA	AAA		
Authentication	PKI	core	Available

Service	Product	Service class	Status
Authorization, Accounting	User Administration (LDAP)	core	Available on the PRACE central server
Accounting, Reporting	Accounting/Apache/CGI DART	core	Available
PRACE internal interactive access	GSISsh	core	Available
PRACE internal interactive access	X.509-based SSH	optional	Planned for August 2012
PRACE external interactive access	at least one of SSH, GSISsh, X.509-based SSH	core	Available
User	User		
Software management & common production environment	Modules	additional	Available
Software management & common production environment	PCPE	core	Available
Data visualization	Data visualization services & tools	optional	Not planned
Monitoring	Monitoring		
Monitoring	Inca	core	Available

Table 7 Overview of integration and planning for CINES

3.2.2 PSNC

PSNC provides to the PRACE community Tier-1 systems called CANE and CHIMERA. The systems are connected to PRACE RI with the Polish optical network PIONIER. CHIMERA.MAN.POZNAN.PL - SGI UltraViolet, offering 16TB memory and 2048 cores is controlled by - Linux, with the inter-connectivity 15 GB/s in the ccNUMA architecture. CANE.MAN.POZNAN.PL – tightly coupled AMD Interlagos cluster with Tesla GPGPU, offering 10,6TB memory and 5448 cores is controlled by - Linux, with the inter-connectivity InfiniBand QDR - Fat tree, 32Gb/s.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Available (with a dedicated 10Gbe link)
Network management, Monitoring	Network monitoring	core	TBI (To be implemented)
Data	Data		

Service	Product	Service class	Status
Data transfer, storage & sharing	MC-GPFS	optional	Not planned
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Available
Compute	Compute		
Uniform access to HPC	Local batch systems	core	Available
Uniform access to HPC	UNICORE	core	Available
Uniform access to HPC	Globus GRAM	optional	Not planned
AAA	AAA		
Authentication	PKI	core	Available
Authorization, Accounting	User Administration (LDAP)	core	Available on the PRACE central server
Accounting, Reporting	Accounting/Apache/CGI DART	core	Installation in progress
PRACE internal interactive access	GSISSH	core	Available
PRACE internal interactive access	X.509-based SSH	optional	TBD
PRACE external interactive access	at least one of SSH, GSISSH, X.509-based SSH	core	Available
User	User		
Software management & common production environment	Modules	additional	TBI
Software management & common production environment	PCPE	core	Installation in progress
Data visualization	Data visualization services & tools	optional	Not planned
Monitoring	Monitoring		
Monitoring	Inca	core	TBI

Table 8 Overview of integration and planning for PSNC

3.2.3 ICHEC

ICHEC provides PRACE Tier-1 services on a system called Stokes. Stokes is the primary national service system in Ireland. Table 9 shows the status of the integration for ICHEC.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Planned (via shared 1Gbit link). Awaiting availability of the PRACE IPSEC hardware.
Network management, Monitoring	Network monitoring	core	Not available.
Data			
Data transfer, storage & sharing	MC-GPFS	optional	Not planned.
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Planned (October 2012)
Compute			
Uniform access to HPC	Local batch systems	core	Available.
Uniform access to HPC	UNICORE	core	Planned (October 2012)
Uniform access to HPC	Globus GRAM	optional	Not planned.
AAA			
Authentication	PKI	core	Available.
Authorization, Accounting	User Administration (LDAP)	core	Available. Provided via the central PRACE LDAP server.
Accounting, Reporting	Accounting/Apache/CGI DART	core	In progress. Completion in July 2012.
PRACE internal interactive access	GSISSH	core	In progress. Completion in July 2012.
PRACE internal interactive access	X.509-based SSH	optional	Not planned.
PRACE external interactive access	at least one of SSH, GSISSH, X.509-based SSH	core	Available
User			
Software management & common production environment	Modules	additional	Available.
Software management & common production	PCPE	core	Available.

Service	Product	Service class	Status
environment			
Data visualization	Various services & tools	optional	Not planned.
Monitoring			
Monitoring	Inca	core	Available

Table 9 Overview of integration and planning for ICHEC

3.2.4 NCSA

NCSA provides to the PRACE community a Tier-1 Blue Gene/P system. The system is composed of two racks with a total of 2048 PowerPC processors (8192 cores). Additional servers provide PRACE services and ensure access to 12 TB of disk storage. Table 10 shows the status of the integration for NCSA.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Planned via shared 1Gbit link. Waiting availability of the PRACE IPSEC hardware.
Network management, Monitoring	Network monitoring	core	TBI (To be implemented)
Data			
Data transfer, storage & sharing	MC-GPFS	optional	Not planned for now
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Planned (September 2012).
Compute			
Uniform access to HPC	Local batch systems	core	Available
Uniform access to HPC	UNICORE	core	Planned (September 2012).
Uniform access to HPC	Globus GRAM	optional	Available
AAA			
Authentication	PKI	core	Available

Service	Product	Service class	Status
Authorization, Accounting	User Administration (LDAP)	core	Available (on the central PRACE LDAP server)
Accounting, Reporting	Accounting/Apache/CGI DART	core	Planned (November 2012).
PRACE internal interactive access	GSISSH	core	Available
PRACE internal interactive access	X.509-based SSH	optional	Not planned for now
PRACE external interactive access	at least one of SSH, GSISSH, X.509-based SSH	core	Available
User		-	
Software management & common production environment	Modules	additional	Available
Software management & common production environment	PCPE	core	Available
Data visualization	Various services & tools	optional	Not planned for now
Monitoring			
Monitoring	Inca	core	Available

Table 10 Overview of integration and planning for NCSA

3.2.5 WCSS

WCSS provides to the PRACE community a Tier-1 system called Supernova. Supernova is a cluster composed of three homogenous partitions based on Intel Xeon processors, and one partition is available through DECI calls. This part of Supernova consists of 404 nodes based on Intel Westmere-EP architecture, connected via InfiniBand network. To integrate the cluster to the PRACE Tier-1 infrastructure some additional dedicated resources are provided. The system is connected to PRACE via the Polish optical network PIONIER. Table 11 shows the status of the integration for WCSS.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Available (10Gbps to PSNC, link from PSNC to PRACE Network shared with PSNC).

Service	Product	Service class	Status
Network management, Monitoring	Network monitoring	core	Available
Data			
Data transfer, storage & sharing	MC-GPFS	optional	Not planned for now
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Available
Compute			
Uniform access to HPC	Local batch systems	core	Available
Uniform access to HPC	UNICORE	core	Available
Uniform access to HPC	Globus GRAM	optional	Not planned for now
AAA			
Authentication	PKI	core	Available
Authorization, Accounting	User Administration (LDAP)	core	Available (on the central PRACE LDAP server)
Accounting, Reporting	Accounting/Apache/CGI DART	core	Available
PRACE internal interactive access	GSISsh	core	Available
PRACE internal interactive access	X.509-based SSH	optional	Not planned for now
PRACE external interactive access	at least one of SSH, GSISsh, X.509-based SSH	core	Available
User			
Software management & common production environment	Modules	additional	Available
Software management & common production environment	PCPE	core	Available
Data visualization	Various services & tools	optional	Not planned for now
Monitoring			
Monitoring	Inca	core	Available

Table 11 Overview of integration and planning for WCSS

3.2.6 UHeM

UHeM (listed with the old acronym UYBHM in the DoW) provides to the PRACE community a Tier-1 system called KARADENIZ. KARADENIZ is an Intel based Linux

cluster. Some dedicated resources are provided to integrate the machine to the PRACE grid infrastructure. Table 12 shows the status of the integration for UHeM.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Not available (There will be a tunnelled shared connection)
Network management, Monitoring	Network monitoring	core	Not available
Data	Data		
Data transfer, storage & sharing	MC-GPFS	optional	Not available
Data transfer, storage & sharing	GridFTP	core	Available
Data transfer, storage & sharing	UNICORE	additional	Planned for mid of September 2012
Compute	Compute		
Uniform access to HPC	Local batch systems	core	Available (Platform's LSF)
Uniform access to HPC	UNICORE	core	Planned for mid of August 2012
Uniform access to HPC	Globus GRAM	optional	Planned for October 2012
AAA	AAA		
Authentication	PKI	core	Available
Authorization, Accounting	User Administration (LDAP)	core	Available on the PRACE central server
Accounting, Reporting	Accounting/Apache/CGI DART	core	Planned for mid of August 2012
PRACE internal interactive access	GSISsh	core	Not available (No dedicated nor tunnelled connection to PRACE network is established)
PRACE internal interactive access	X.509-based SSH	optional	Not Available
PRACE external interactive access	at least one of SSH, GSISsh, X.509-based SSH	core	Available
User	User		
Software management & common production environment	Modules	additional	Available

Service	Product	Service class	Status
Software management & common production environment	PCPE	core	Available
Data visualization	Data visualization services & tools	optional	Planned for September 2012
Monitoring	Monitoring		
Monitoring	Inca	core	Planned for mid of August 2012

Table 12 Overview of integration and planning for UHeM

3.2.7 UiO

The University of Oslo (UiO) will provide Tier-1 services to the PRACE community starting with DECI-9. The services will be provided from a Linux cluster called Abel and supporting infrastructure. The Abel cluster is based on over 600+ Intel Sandy Bridge based nodes, with additional nodes providing GPU and large memory resources. All nodes are interconnected with FDR InfiniBand. Table 13 shows the planning of the integration for UiO.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Planned to be a dedicated 10Gbe link
Network management, Monitoring	Network monitoring	core	TBI (to be implemented)
Data	Data		
Data transfer, storage & sharing	MC-GPFS	optional	Not planned
Data transfer, storage & sharing	GridFTP	core	TBI
Data transfer, storage & sharing	UNICORE	additional	TBD (to be decided)
Compute	Compute		
Uniform access to HPC	Local batch systems	core	Available (SLURM)
Uniform access to HPC	UNICORE	core	TBD
Uniform access to HPC	Globus GRAM	optional	TBD
AAA	AAA		
Authentication	PKI	core	TBI
Authorization, Accounting	User Administration (LDAP)	core	Planned to be on the PRACE central server (TBI/TBD)
Accounting,	Accounting/Apache/CGI DART	core	TBI

Service	Product	Service class	Status
Reporting			
PRACE internal interactive access	GSISSH	core	TBI
PRACE internal interactive access	X.509-based SSH	optional	TBI
PRACE external interactive access	at least one of SSH, GSISSH, X.509-based SSH	core	TBI
User	User		
Software management & common production environment	Modules	additional	Available
Software management & common production environment	PCPE	core	TBI
Data visualization	Data visualization services & tools	optional	Not planned
Monitoring	Monitoring		
Monitoring	Inca	core	TBI

Table 13 Overview of planning of the integration for UiO

3.2.8 VSB-TUO

Most of the services will be tested and pre-deployed in 2nd half of 2012 before the installation of the final system. The final system should be available in January 2013. Table 14 shows the planning of the integration for VSB-TUO.

Service	Product	Service class	Status
Network Management, Monitoring			
Dedicated Network	PRACE Network	additional	Planned via shared 1Gbit link (January 2013). Waiting availability of the PRACE IPSEC hardware.
Network management, Monitoring	Network monitoring	core	Planned (January 2013)
Data			
Data transfer, storage & sharing	MC-GPFS	optional	Not planned.
Data transfer, storage & sharing	GridFTP	core	Planned (December 2012)
Data transfer, storage & sharing	UNICORE	additional	Planned (December 2012)
Compute			
Uniform access to	Local batch systems	core	Planned (January 2013)

Service	Product	Service class	Status
HPC			
Uniform access to HPC	UNICORE	core	Planned (December 2012)
Uniform access to HPC	Globus GRAM	optional	Planned (December 2012)
AAA			
Authentication	PKI	core	Planned (December 2012)
Authorization, Accounting	User Administration (LDAP)	core	Available. Provided via the central PRACE LDAP server.
Accounting, Reporting	Accounting/Apache/CGI DART	core	Planned (January 2013)
PRACE internal interactive access	GSISSH	core	Planned (December 2012)
PRACE internal interactive access	X.509-based SSH	optional	Not planned.
PRACE external interactive access	at least one of SSH, GSISSH, X.509-based SSH	core	Planned (December 2012)
User			
Software management & common production environment	Modules	additional	Planned (December 2012)
Software management & common production environment	PCPE	core	Planned (December 2012)
Data visualization	Various services & tools	optional	Not planned.
Monitoring			
Monitoring	Inca	core	Planned (December 2012)

Table 14 Overview of planning of the integration for VSB-TUO

4 Annexes

4.1 PRACE Service Catalogue

Introduction

The PRACE distributed research infrastructure is well on its path to provide a complete set of services to its users. Service provision to users is currently mainly done by the Tier-0 hosting partners, governed by the PRACE AISBL statutes and the Agreement for the Initial Period. Relations between Tier-0 sites and their users are typically managed through specific User Agreements between them. PRACE AISBL gives advice to the hosting sites on the allocation of compute resources based on the pan-European PRACE Peer Review. For the execution of the peer review and other services such as Web, the PRACE also uses services provided by third parties. Other important services such as user support and operation of the distributed infrastructure are provided by the PRACE-1IP project.

Tier-1 partners provide access to users, governed by the DECI commitments, currently within the Implementation Phase projects.

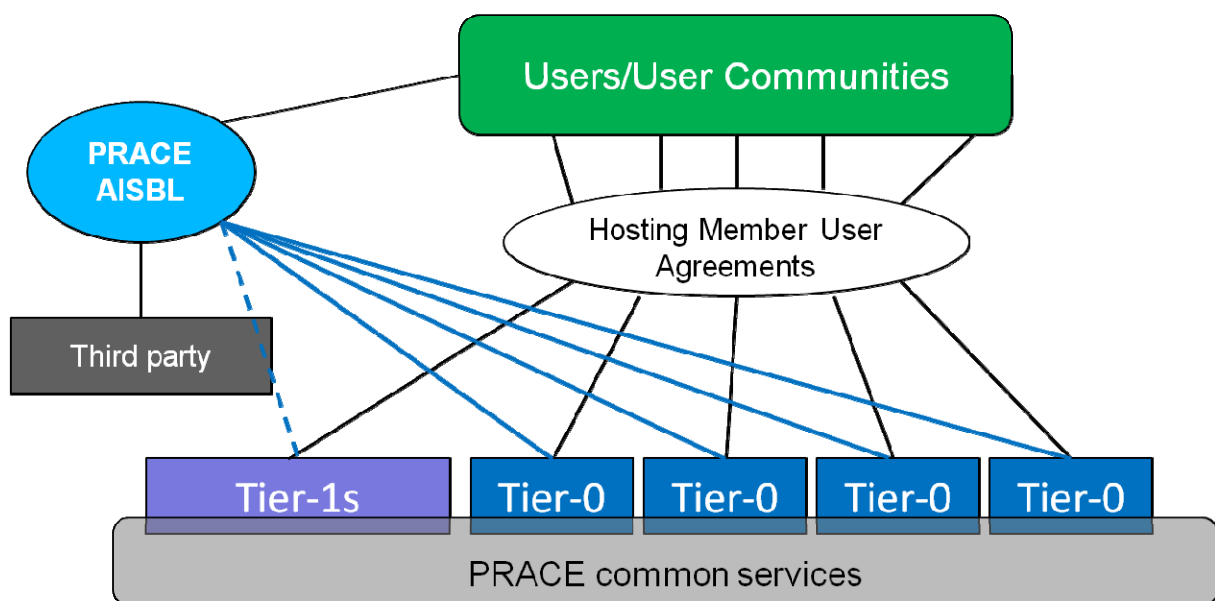


Figure 8 PRACE Service provision scheme and contracts to its users

To support a good and complete overview of all PRACE Operational Services, we have started to develop the PRACE Service Catalogue, which lists and describes the complete set of operational services that the PRACE AISBL is providing, from the point of view of PRACE as a service provider. In addition, Tier-1 services are added to this Service Catalogue to complete the picture of PRACE service provision.

The purpose of the PRACE Service Catalogue is:

- To describe all PRACE operational services
- To define PRACE service categories, and classify all PRACE services accordingly

In this way it describes the full PRACE service picture from hosting partners, other partners, the project and the PRACE AISBL.

An important aspect of the PRACE Service Catalogue is the classification of services. We have defined three service classes: Core services, Additional services and Optional services. The availability and support for each of these service classes is defined and described in Table 15.

Core services	
Availability:	Robust, reliable and persistent technologies that must be implemented and accessible at all PRACE Tier-0/1 sites, or provided centrally.
Support:	Support for these services is provided during support hours, i.e. the normal working hours according to the usual working arrangements of the particular Tier-0/1 site.
Additional services	
Availability:	Robust, reliable and persistent technologies that must be implemented and accessible at all PRACE Tier-0/1 sites where possible. Reasons for the service not being implemented at a Tier-0/1 site include technical, legal, financial and policy limitations, whenever an unreasonable effort is needed to provide the service.
Support:	If applicable, support for these services is provided during support hours.
Optional services	
Availability:	Implemented optionally by PRACE Tier-0/1 sites. Availability and long-term support are not guaranteed by PRACE.
Support:	PRACE AISBL and/or Tier-1 partners provide support for these services on a case by case basis, in addition to any support provided directly by the specific site.

Table 15 Classification of PRACE Services as part of the PRACE Service Catalogue

Every PRACE service will be classified according to this classification. It should be noted that the service classes define the availability of the services at the hosting sites, and are not related to service levels.

The definition of the services in the PRACE Service Catalogue is achieved through six criteria:

- **Description:** A brief summary of the service, indicating its value and a general overview of its implementation.
- **Class:** Services are arranged according to their expected availability and support across PRACE hosting partners. This classification is composed of three levels that indicate how essential a service is for the PRACE RI: Core, Additional, and Optional.
- **Provider:** The person(s), group(s), site(s), or team(s) involved in and responsible for the correct implementation and operation of the services.
- **Reference:** Documents and agreements that contain more specific details and information concerning the service provision.
- **Category:** Services are grouped into seven different categories, according to their

specific domain: Compute, User, Data, Generic, AAA, Network, and Monitoring.

- **Service:** Concrete services and/or software products that have been chosen to implement the service. For each service/product its Service Class (core, additional, optional) is indicated for Tier-0, Tier-1 and/or PRACE AISBL or a single partner.

The PRACE Service Catalogue will be regularly updated to document the actual status of all services and will be maintained as a living document, where all changes in services and their provision will be indicated. Status of services can change when new services are deployed, when levels of services are changed, when new service providers (i.e. new hosting partners) are integrated or when new software products are released. The document will at all times reflect the current situation of PRACE services, so that it can be used as the main reference document for service provision within PRACE.

The starting point for the list of services that are listed in the PRACE Service Catalogue was the list created by WP4 of the PRACE-PP in WP4.

PRACE Services

Uniform access to HPC				
Description:	Allows a user to execute code on PRACE Tier-0/1 systems, monitor its evolution and retrieve the results across Tier-0/1 systems.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (compute services representative of the PRACE Operational Team)			
Reference:	Draft User Agreement			
Category:	Compute			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	UNICORE	core	core	-
	Globus GRAM	optional	optional	-
	Local batch system	core	core	-
	DESHL	-	optional	-
Remarks:	-			

PRACE internal interactive command-line access to HPC				
Description:	Allows a employee of a PRACE partner to connect remotely to a Tier-0/1 system and execute command-line instructions.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (compute services representative of the PRACE Operational Team)			
Reference:	NA			
Category:	AAA			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	GSISSH	additional	core	-
	X.509-based SSH	optional	optional	-
Remarks:				

PRACE external (user) interactive command-line access to HPC	
Description:	Allows a user to connect remotely to a Tier-0/1 system and execute

	command-line instructions.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (compute services representative of the PRACE Operational Team)			
Reference:	Draft User Agreement			
Category:	AAA			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	At least one of SSH, GSISSH, X.509-based SSH	core	core	-
Remarks:				

Project submission

Description:	Provides Tier-0 users with a centralized point for submitting projects for Peer Review. In case of Tier-1 access, provision of DECI database for project registration.			
Class:	Core			
Provider:	PRACE Peer Review Team			
Reference:	PRACE PP D2.4.2			
Category:	User			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	PRACE peer review tool (for Tier-0 access)	-	-	core
	DECI database (for Tier-1 access)	-	-	core
Remarks:	-			

Data transfer, storage and sharing

Description:	Each PRACE User is provided a “home” directory and access to a project space shared with his User Group, at each of the assigned Tier-0/1 sites. The amount of space in each of these directories is indicated in Annex A of the User Agreement for Tier-0 sites. Data can be transferred to and from these directories.
Class:	Core

Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (data services representative of the PRACE Operational Team)			
Reference:	Draft User Agreement			
Category:	Data			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	MC-GPFS	optional	optional	-
	GridFTP	core	core	-
	UNICORE	additional	additional	-
Remarks:	GridFTP is a core service for Tier-1 only if a dedicated network is available			

HPC Training

Description:	Provides training sessions and workshops for topics and technologies in high-performance computing, as well as online and offline education material.			
Class:	Core			
Provider:	PRACE 1IP WP3, PRACE 2IP WP4, Tier-0/1 site, PRACE Advanced Training Centres			
Reference:				
Category:	User			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Training portal	-	-	core
Remarks:	-			

Documentation and Knowledge Base

Description:	User documentation in the form of an online knowledge base, including manuals and other information and tools that are indispensable for the users.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE AISBL + PRACE 1IP WP6, WP7, WP3 + PRACE 2IP WP6, WP7, WP3			
Reference:				
Category:	User			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	CMS	-	-	core
	Plone	-	-	core

	DocBook	optional	optional	-
Remarks:				

Data Visualization

Description:	Converts data into images as a tool to help users with analysis.			
Class:	Optional			
Provider:	Specific PRACE sites			
Reference:				
Category:	Generic			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Various services and tools	optional	optional	-
Remarks:				

Authentication

Description:	Confirm the identity of a user and bind that user to a new account. This involves identifying a user's certificate, creating a global PRACE RI account for the user on the central LDAP and making it available for distribution on all PRACE RI Resources.			
Class:	Core			
Provider:	Peer Review Team + Tier-0/1 site + PRACE 1IP/2IP WP6 (AAA services representative of the PRACE Operational Team)			
Reference:	PRACE Security Policy			
Category:	AAA			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	PKI	core	core	-
	MyProxy			core
Remarks:	My proxy is provided by multiple parties (e.g. as backup/disaster recovery).			

Authorization

Description:	Specifies access rights for each user account created based on the content of the specific User Agreement and the PRACE Security Policy. Ensures that security rules and access rights are obeyed, and manages changes to these
---------------------	---

	(based on new security policies or redefined User Agreements).			
Class:	Core			
Provider:	Peer Review Team + Security Forum + Tier-0/1 site + PRACE 1IP/2IP WP6 (AAA services representative of the PRACE Operational Team)			
Reference:	Draft User Agreement, PRACE Security Policy, PRACE Acceptable Use Policy			
Category:	AAA			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	LDAP (user administration)	core	core	-
Remarks:				

Accounting

Description:	Keeps track of resource usage linked to an account for analysis by users and management. Guarantees that users are not exceeding their limits, as specified by their User Agreement.			
Class:	Core			
Provider:	Peer Review Team + Tier-0/1 site + PRACE 1IP/2IP WP6 (AAA services representative of the PRACE Operational Team)			
Reference:				
Category:	AAA			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Apache/CGI DART	core	core	-
	LDAP (user administration)	core	core	-
Remarks:	-			

Information Management

Description:	Provides a common PRACE collaborative environment for sharing relevant information between PRACE sites (BSCW, wiki, subversion, ...).			
Class:	Core			
Provider:	WP6			
Reference:				
Category:	Generic			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or

				single partner
	TWiki	-	-	core
	SVN	-	-	core
	BSCW	-	-	core
	Prace-ri website	-	-	core
Remarks:	-			

Network Management

Description:	<p>Establishes and maintains network connections between all PRACE nodes (Tier-0 and Tier-1 systems). The PRACE Network Operations Centre (NOC) operates the PRACE backbone network and the corresponding network monitoring system. The PRACE NOC coordinates networking activities of PRACE partners, who are responsible for creation and management of network connection between the local resources and GÉANT (PRACE backbone).</p> <p>PRACE partner's local network specialists and the PRACE NOC should support PRACE users in using the PRACE network infrastructure.</p> <p>The PRACE backbone will be dedicated, whereas local site connectivity of HPC systems and PRACE servers to the global Internet are public.</p>			
Class:	Core			
Provider:	PRACE NOC and local NOCs of PRACE partners (at least one person per site should be also a network services representative of the PRACE Operational Team)			
Reference:	NA			
Category:	Network			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	PRACE dedicated network	core	additional	-
	DNS (PRACE RI domain management)	-	-	core
	PerfSonar framework	core	core	-
Remarks:	<p>Dedicated network is an additional service for Tier-1 partners. This means that a dedicated network is required unless unreasonable effort or funding is required.</p> <p>PerfSonar framework is only a service if a dedicated network is available.</p>			

Monitoring				
Description:	Periodically presents and analyzes up-to-date essential PRACE parameters and service availability to keep track of the situation of the distributed RI, for example: system uptime/downtime and usage levels, network connections, software and service availability.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (Monitoring services representative of the PRACE Operational Team)			
Reference:				
Category:	Monitoring			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Inca	core	core	-
Remarks:	-			

Reporting				
Description:	Periodic reports of system utilization from the Tier-0/1 hosting partner to the PRACE AISBL.			
Class:	Core			
Provider:	PRACE AISBL + Tier-0/1 Hosting Partner			
Reference:				
Category:	Monitoring			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Apache/CGI DART	core	core	-
Remarks:	-			

Software Management and Common Production Environment	
Description:	Provides software, tools, libraries, compilers, and uniform mechanisms for software and environment configuration. Presents users with a uniform environment across PRACE Tier-0/1 systems, hiding inessential details such as software installation paths.
Class:	Core
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 + WP7
Reference:	NA

Category:	Generic			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	Modules	additional	additional	-
	PCPE	core	core	-
Remarks:	-			

First Level User Support

Description:	Each PRACE User has access to a centrally managed Helpdesk. Issues raised to the Helpdesk are routed to the appropriate First Level Support team. First Level support is responsible for gathering the user's information and determining their issue by identifying what the user is trying to accomplish, analyzing the symptoms and figuring out the underlying problem.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 (User services representative of the PRACE Operational Team)			
Reference:	Draft User Agreement, PRACE 1IP D6.1			
Category:	User			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	RT-TTS	-	-	core
Remarks:	-			

Advanced User Support

Description:	Provision of support above and beyond basic problem analysis including but not limited to applications porting, performance tuning, pre-post processing, data access. Higher Level support receives issues that are escalated and routed from First Level User Support.			
Class:	Core			
Provider:	Tier-0/1 site + PRACE 1IP/2IP WP6 + WP7			
Reference:	Draft User Agreement, PRACE 1IP D6.1			
Category:	User			
Service:	Product/service	Tier-0	Tier-1	PRACE AISBL or single partner
	RT-TTS	-	-	core
Remarks:	-			

Service	Service class	Product	Tier-0	Tier-1	PRACE AISBL or single partner
Network management, Monitoring					
Dedicated network	core	PRACE Network	core	additional	
Dedicated network	core	DNS			core
Network management, Monitoring	core	PerfSonar framework	core	core	
Data	-				
Data transfer, storage & sharing	core	MC-GPFS	optional	optional	
Data transfer, storage & sharing	core	GridFTP	core	core	
Data transfer, storage & sharing	core	UNICORE	additional	additional	
Compute	-				
Uniform access to HPC	core	Local batch systems	core	core	
Uniform access to HPC	core	UNICORE	core	core	
Uniform access to HPC	core	Globus GRAM	optional	optional	
AAA	-			-	
Authentication	core	PKI	core	core	
Authentication	core	MyProxy			core
Authorization, Accounting	core	User Administration (LDAP)	core	core	
Accounting, Reporting	core	Apache/CGI DART	core	core	
PRACE internal interactive access	core	GSISsh	additional	core	
PRACE internal interactive access	core	X.509-based SSH	optional	optional	
PRACE external interactive access	core	at least one of SSH, GSISsh, X.509-based SSH	core	core	
User	-			-	
Software management & common production environment	core	Modules	additional	additional	
Software management & common production environment	core	PCPE	core	core	
First level user support, advanced user support	core	RT-TTS			core (tool)
Uniform access to HPC	core	DESHL		optional	
Documentation and knowledge base	core	CMS, Plone, DocBook			core
Project submission, Accounting	core	PRACE peer review tool (for Tier-0 access)			core

Project submission, Accounting	core	DECI database (for Tier-1 access)			core
HPC Training	core	Training portal			core
Data visualization	optional	Various services & tools	optional	optional	
Monitoring	-				
Monitoring	core	Inca	core	core	
Generic					
Information management	core	TWiki			core
	core	SVN			core
	core	BSCW			core
	core	prace-ri website			core

Table 16 Overview of PRACE services, categories and product classes