SEVENTH FRAMEWORK PROGRAMME
Research Infrastructures

INFRA-2012-2.3.1 – Third Implementation Phase of the European High Performance Computing (HPC) service PRACE

PRACE-3IP

PRACE Third Implementation Phase Project

Grant Agreement Number: RI-312763

D6.1.3
Final Operations Report including XSEDE Collaborative Services

Final

Version: 1.0
Author(s): Jules Wolfrat, SURFsara
Michael Rambadt, JUELICH
Date: 24.01.2015

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: RI-312763**	
	Project Title: PRACE Third Implementation Phase Project	
	Project Web Site: http://www.prace-project.eu	
	Deliverable ID: < **D6.1.3**>	
	Deliverable Nature: <DOC_TYPE: Report >	
	Deliverable Level: PU	**Contractual Date of Delivery:** 31 / 01 / 2015
		Actual Date of Delivery: 31 / 01 / 2015
	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: Final Operations Report including XSEDE Collaborative Services**	
	ID: D6.1.3	
	Version: <1.0 >	**Status:** Final
	Available at: http://www.prace-project.eu	
	Software Tool: Microsoft Word 2007	
	File(s): D6.1.3.docx	
Authorship	**Written by:**	Jules Wolfrat, SURFsara Michael Rambadt, JUELICH
	Contributors:	Morris Riedel, JUELICH Shantenu Jha, Rutgers University, US Sandra Gesing, University of Notre Dame, Indiana, US
	Reviewed by:	Mark Bull, EPCC Florian Berberich, JUELICH
	Approved by:	MB/TB

Document Status Sheet

Version	**Date**	**Status**	**Comments**
0.1	11/12/2014	Draft	Index proposed
0.2	07/01/2015	Draft	Sections for XSEDE collaborations have to be added
0.3	09/01/2015	Draft for internal review	Additional input for XSEDE collaborations
0.9	20/01/2015	Final draft	Accepted all updates of reviewers
1.0	21/01/2015	Final version	For approval

Document Keywords

Keywords:	PRACE, HPC, Research Infrastructure, Operations Report, XSEDE, Collaboration

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-312763**. 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

Table of Contents

Project and Deliverable Information Sheet i

Document Control Sheet i

Document Status Sheet i

Document Keywords ii

Table of Contents iii

List of Figures iii

List of Tables iii

References and Applicable Documents iv

List of Acronyms and Abbreviations v

Executive Summary 1

1 Introduction 2

2 PRACE Operational Common Services 3

- **2.1 Introduction** 3
- **2.2 Status of services and infrastructure** 3

3 XSEDE Collaborative Services 7

- **3.1 Introduction** 7
- **3.2 Smart Data Analytics for Earth Sciences across XSEDE and PRACE** 7
  - 3.2.1 *Objectives* *7*
  - 3.2.2 *Results* *7*
- **3.3 Interoperable High Throughput Binding Affinity Calculator for Personalised Medicine** 8
  - 3.3.1 *Objectives* *8*
  - 3.3.2 *Results* *9*
- **3.4 UNICORE Use Case Integration for XSEDE and PRACE** 9
  - 3.4.1 *Objectives* *9*
  - 3.4.2 *Results* *10*

4 Conclusions 11

List of Figures

Figure 1 Comparisons of different implementations of the parallel BDSCAN algorithm 8

List of Tables

Table 1 Available Tier-0 systems January 2015 4

Table 2 Overview of Tier-1 systems in January 2015 6

References and Applicable Documents

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

[2] D6.1.2 Second Annual Operations Report http://www.prace-ri.eu/IMG/pdf/d6.1.2_3ip.pdf

[3] http://www.prace-ri.eu/prace-resources/

[4] Shellshock vulnerability: https://www.us-cert.gov/ncas/alerts/TA14-268A

[5] POODLE vulnerability: https://www.openssl.org/~bodo/ssl-poodle.pdf

[6] PANGAEA Data Publisher for Earth & Environmental Sciences, http://www.pangaea.de/

[7] IAGOS In-Service Aircraft for a Global Observing System, http://www.iagos.org/

[8] Spatial Computing 2020, http://www.cra.org/ccc/files/docs/Spatial_Computing_White_Papers.pdf

[9] INGV – National Institute of Geophysics and Volcanology, http://www.ingv.it/en/

[10] https://www.xsede.org/software/genesis

[11] http://genesis2.virginia.edu/wiki/Main/GFFS

[12] http://www.scidb.org/

[13] S.K. Sadiq, D. Wright, S.J. Watson, S.J. Zasada, I. Stoica, Ileana, and P.V. Coveney. Automated Molecular Simulation-Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. Journal of Chemical Information and Modeling, 48,(9): 1909-1919, 2008.

[14] S.K. Sadiq, D.W. Wright, O.A. Kenway, and P.V Coveney. Accurate ensemble molecular dynamics binding free energy ranking of multidrug-resistant HIV-1 proteases. Journal of Chemical Information and Modeling, 50(5):890–905, 2010.

[15] Computing Clinically Relevant Binding Free Energies of HIV-1 Protease Inhibitors, D Wright, BA. Hall, O Kenway, S Jha and P V Coveney ACS Journal of Chemical Theory and Computation (in press) http://dx.doi.org/10.1021/ct4007037

[16] http://radicalpilot.readthedocs.org/en/latest/

[17] A. Streit, P. Bala, A. Beck-Ratzka, K. Benedyczak, S. Bergmann, R. Breu, J. M. Daivandy, B. Demuth, A. Eifer, A. Giesler & others UNICORE 6 - recent and future advancements annals of telecommunications-annales des télécommunications, Springer, 2010, 65, 757-762

[18] S. Gesing, R. Grunzke, J. Krüger, G. Birkenheuer, M. Wewior, P. Schäfer, B. Schuller, J. Schuster, S. Herres-Pawlis, S. Breuers, Á. Balaskó, M. Kozlovszky, A. S. Fabri, L. Packschies, P. Kacsuk, D. Blunk, T. Steinke, A. Brinkmann, G. Fels, R. Muller- Pfefferkorn, R. Jäkel & O. Kohlbacher A Single Sign-On Infrastructure for Science Gateways on a Use Case for Structural Bioinformatics Journal of Grid Computing, 2012, 10, 769-790

[19] R. Grunzke, G. Birkenheuer, D. Blunk, S. Breuers, A. Brinkmann, S. Gesing, S. Herres-Pawlis, O. Kohlbacher, J. Krüger, M. Kruse, R. Muller-Pfefferkorn, P. Schäfer, B. Schuller, T. Steinke & A. Zink A Data Driven Science Gateway for Computational Workflows UNICORE Summit 2012 Proceedings, 2012, 15, 35-49

[20] S. Gesing, S. Herres-Pawlis, G. Birkenheuer, A. Brinkmann, R. Grunzke, Pé. Kacsuk, O. Kohlbacher, M. Kozlovszky, J. Krüger, R. Muller-Pfefferkorn, P. Schäfer & T. Steinke A Science Gateway Getting Ready for Serving the International Molecular Simulation Community Proceedings of Science, Proceedings of Science, 2012, PoS(EGICF12-EMITC2)050

List of Acronyms and Abbreviations

AAA	Authorization, Authentication, Accounting
AISBL	Association Internationale Sans but Lucratif (legal form of the PRACE-RI)
AMD	Advanced Micro Devices
BSC	Barcelona Supercomputing Center (Spain)
CA	Certificate Authority
CaSToRC	Computation-based Science and Technology Research Center (Cyprus)
CEA	Commissariat à l'énergie atomique et aux énergies alternatives
CGI	Common Gateway Interface
CINECA	Consorzio Interuniversitario, the largest Italian computing centre (Italy)
CINES	Centre Informatique National de l'Enseignement Supérieur (represented in PRACE by GENCI, France)
CPU	Central Processing Unit
CSC	Finnish IT Centre for Science (Finland)
CSCS	The Swiss National Supercomputing Centre (represented in PRACE by ETHZ, Switzerland)
CSIRT	Computer Security Incident Response Team
Cyfronet	Academic Computer Centre CYFRONET AGH (Krakow, Poland)
DDR	Double Data Rate
DECI	Distributed European Computing Initiative
DEISA	Distributed European Infrastructure for Supercomputing Applications. EU project by leading national HPC centres.
DP	Double Precision, usually 64-bit floating point numbers
EC	European Community
EGI	European Grid Initiative
EGCF	European Globus Community Forum
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)
EUDAT	European Data Infrastructure
EUGridPMA	European Grid Policy Management Authority
EX	Expandable, e.g., Nehalem-EX (Intel)
FDR	Fourteen Data Rate
FZJ	Forschungszentrum Jülich (Germany)
GB	Giga (= 2^{30} ~ 10^9) Bytes (= 8 bits), also GByte
Gb/s	Giga (= 10^9) bits per second, also Gbit/s
GB/s	Giga (= 10^9) Bytes (= 8 bits) per second, also GByte/s
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)
GFFS	Global Federated File System
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
GNU	GNU's not Unix, a free OS
GPU	Graphic Processing Unit
HDD	Hard Disk Drive
HLRS	High Performance Computing Center Stuttgart (Stuttgart, Germany)
HPC	High Performance Computing; Computing at a high performance level at any given time; often used synonym with Supercomputing
HPL	High Performance LINPACK
IB	InfiniBand
IBM	Formerly known as International Business Machines
ICM	Interdisciplinary Center for Mathematical and Computational modelling (Warsaw, Poland)
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique (represented in PRACE by GENCI, France)
IGTF	Interoperable Global Trust Federation
ITIL	Information Technology Infrastructure Library
I/O	Input/Output
JSC	Jülich Supercomputing Centre (FZJ, Germany)
KB	Kilo (= 2^{10} ~ 10^3) Bytes (= 8 bits), also KByte
KPI	Key Performance Indicator
LDAP	Lightweight Directory Access Protocol
LINPACK	Software library for Linear Algebra
LRZ	Leibniz Supercomputing Centre (Garching, Germany)
LSF	IBM® Platform™ Session Scheduler
MB	Mega (= 2^{20} ~ 10^6) Bytes (= 8 bits), also MByte
MPI	Message Passing Interface
MPP	Massively Parallel Processing (or Processor)
NCSA	National Center for Supercomputing Applications (Bulgaria, USA)
NIIF	National Information Infrastructure Development Institute (Hungary)
NOC	Network Operations Center
NUMA	Non-Uniform Memory Access or Architecture
OS	Operating System
PB	Peta (= 2^{50} ~ 10^{15}) Bytes (= 8 bits), also PByte
PCPE	PRACE Common Production Environment
PDC	PDC Center for High-Performance Computing at the KTH Royal Institute of Technology (Sweden)
PFlop/s	Peta (= 10^{15}) Floating-point operations (usually in 64-bit, i.e. DP) per second, also PF/s
PKI	Public Key Infrastructure
PPR	PRACE Peer Review
PRACE	Partnership for Advanced Computing in Europe; Project Acronym
PSNC	Poznan Supercomputing and Networking Centre (Poland)
QDR	Quad Data Rate
RAID	Redundant Array of Independent (or inexpensive) Disks
RAM	Random Access Memory
RI	Research Infrastructure
RZG	Rechenzentrum Garching (RZG) of the Max Planck Society and the IPP (Germany)
SSD	Solid State Disk or Drive
SSH	Secure Shell

STFC	Science and Technology Facilities Council (represented in PRACE by EPSRC, United Kingdom)
SURFsara	Dutch national High Performance Computing & e-Science Support Center
TB	Tera (= 2^{40} ~ 10^{12}) Bytes (= 8 bits), also TByte
TFlop/s	Tera (= 10^{12}) Floating-point operations (usually in 64-bit, i.e. DP) per second, also TF/s
Tier-0	Denotes the apex of a conceptual pyramid of HPC systems. In this context the Supercomputing Research Infrastructure would host the Tier-0 systems; national or topical HPC centres would constitute Tier-1
UCC	UNICORE Command line Client
UCL	University College London
UHeM	National Center for High Performance Computing (Turkey)
UiO	University of Oslo (Norway)
UNICORE	Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.
VPN	Virtual Private Network
VSB-TuO	Vysoká škola báňská – Technical University of Ostrava (Czech republic)
WCSS	Wrocławskie Centrum Sieciowo-Superkomputerowe (WCNS, Wroclaw Centre for Networking and Supercomputing, Poland)

Executive Summary

This report presents the work performed by work package 6 of the PRACE-3IP project on the operation of the PRACE distributed infrastructure for Tier-0 and Tier-1 systems for the work extension in period[1] July 2014 to end of January 2015.

The main activity for this period has been the operation and coordination of the PRACE common services, which enable the interoperation among different PRACE HPC systems and which present a common interface to the PRACE Tier-0 and PRACE Tier-1 systems.

Because of the limited available effort in this period for operational activities the number of new proposed changes has been limited to the inevitable ones, such as the removal of a system or a configuration change at a site.

On operational security of the infrastructure, information on major vulnerabilities in basic system software has been exchanged and one incident at a partner site has been reported.

Six Tier-0 systems have been operational for the project period under review:

- JUQUEEN at GCS@FZJ;
- CURIE at GENCI@CEA;
- SuperMUC at GCS@LRZ;
- Hermit at GCS@HLRS, upgraded to Hornet
- FERMI at CINECA;
- MareNostrum at BSC

Hermit at GCS@HLRS, has been upgraded to a new system, Hornet, also at GCS@HLRS. With a peak performance of 3.8 PFlop/s for the new system this brings the total Tier-0 peak performance above 18 PFlop/s.

Three Tier-1 systems have been upgraded in capacity, one has been replaced and one site has disappeared from the list. 23 Tier-1 sites are operational at the end of January 2015.

Support for the enhancement of interoperability between PRACE and XSEDE has been given to three collaborating teams with members from both Europe and the US.

---

[1] PRACE-3IP Extension denotes the period of M25-M31 extending the work of WP2 – WP7 by seven month in order to ensure a seamless and continuous support of the project for the PRACE RI prior to the planned start of the PRACE-4IP project in H2020

1 Introduction

This report describes the operations of the common services of the PRACE infrastructure for the period July 2014 through January 2015. It also describes the support given to enhance the interoperable facilities between the PRACE and XSEDE infrastructures for selected research teams, which started in April 2014. Originally the activities of WP6 were planned to finish at the end of June 2014, but because of the gap period with a possible follow-up project, some of the activities of WP6 have been extended for a period of seven additional months.

Because only effort for the coordination of operational activities for the seven additional months was available, it was agreed to minimize the effort needed to sustain the common PRACE services. As a result, major changes in services in this period should be scheduled as little as possible. Only minor changes, e.g. configuration changes due to changes in the infrastructure, and urgent changes because of security issues should be scheduled. In Section 2 details of the operational activities are given.

The support for enhancing the interoperability between PRACE and XSEDE started in April 2014 and was planned for at most six months, so it was decided that partners involved in the support should get effort after June 2014 for another three months. Because there have been delays in the start of the three selected projects the support period has been extended for another additional month. Section 3 describes the results of the support activities.

2 PRACE Operational Common Services

2.1 Introduction

The PRACE common services enable the interoperability among the systems providing PRACE resources. These services are described in the Service Catalogue, which is a living document, reflecting the evolution of available services. It is a collection of supported sustainable services for which user documentation is provided on the PRACE website, except for those services which are only used by PRACE staff, in which case the documentation is provided on the internal wiki. The user administration service is an example of an internal service only.

Although the effort to sustain the common services has been lowered for this period, the implemented processes to guarantee the availability and reliability of services have been supported. All proposed changes in the PRACE common services followed the Change Management procedure. The Operator on Duty and Helpdesk Duty, both tasks done in a weekly shift by partners, have been continued. These tasks guarantee a timely response to problems reported by monitoring services or users.

2.2 Status of services and infrastructure

The process of selection and deployment of a common set of services aims at presenting all Tier-0 and Tier-1 centres as a single distributed infrastructure, instead of a set of individual systems/computing facilities.

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

Selection of common services is published in the PRACE Service Catalogue and, once chosen, the responsibility for a service is taken by a service manager. Details about the Common Services can be found in the second year deliverable D6.1.2 [2]. No update on the status is given here because there have not been any major changes.

Proposed service changes are classified as major, minor or urgent, dependent on their impact and urgency. Three major changes have been completed in this period, all three related to the decommissioning of a Tier-1 system: at IDRIS, CEA and CINES, all three in France. Thirteen minor changes have been implemented in this period and no urgent changes have been scheduled.

The collaboration on operational security is very important. It is used to inform partners about possible vulnerabilities and about security incidents. Information about three vulnerabilities has been distributed and also discussed where needed: for MOAB (a workload management facility), relevant for three partners, for Shellshock [4], a bash shell vulnerability, and for POODLE [5], a SSL v3 vulnerability. For the latter, specific update information for UNICORE has been distributed too. One security incident has been reported by a partner. The partner provided detailed information that could be used by other partners to detect possible local impact.

Table 1 shows the six operational Tier-0 systems in January 2015. The system at HLRS@GCS (Germany) has been upgraded in this period from a CRAY XE6 to a CRAY XC40. The system went into production on the 15th of December 2014 and all PRACE user accounts have been migrated from the previous system HERMIT to the new system HORNET. With this upgrade the total theoretical peak performance of the Tier-0 resources is over 18 Pflop/s. More information on these resources can be found on the PRACE website [3].

Machine – Site	Architecture	TF/s	#Cores	Total Memory (TB)
MareNostrum – BSC (Spain)	IBM iDataPlex DX360M4 CPU: Intel Sandy Bridge-EP 8-core 2.7GHz Interconnect: InfiniBand FDR-10	1017	48448	96.6
	IBM iDataPlex DX360M4 CPU: Intel Sandy Bridge-EP 8-core 2.6GHz + Intel Xeon Phi 5110P Interconnect: InfiniBand FDR-10	100	672	2.6
CURIE – GENCI@CEA (France)	Bull Bullx BCS CPU: Intel Nehalem-EX 8-core 2.27GHz Interconnect: InfiniBand QDR / 2-plane Full Fat Tree	104	11520	45.0
	Bull Bullx B505 CPU: Intel Westmere-EP 4-core 2.67GHz + Nvidia Tesla M2090 Interconnect: InfiniBand QDR / 2-plane Full Fat Tree	198	1152	3.3
	Bull Bullx B510 CPU: Intel Sandy Bridge-EP 8-core 2.70GHz Interconnect: InfiniBand QDR / Full Fat Tree	1742	80640	315.0
FERMI – CINECA (Italy)	IBM BlueGene/Q CPU: IBM PowerPC A2 16-core 1.60GHz Interconnect: BG/Q / 5D Torus	2097	163840	160.0
JUQUEEN – GCS@FZJ (Germany)	IBM BlueGene/Q CPU: IBM PowerPC A2 16-core 1.60GHz Interconnect: BG/Q / 5D Torus	5872	458752	458.7
HORNET – GCS@HLRS (Germany)	Cray XC40 CPU: Intel Haswell E5 2680v3 12-core 2,5GHz Interconnect: Cray Aries	3790	94656	505.0
SuperMUC – GCS@LRZ (Germany)	IBM iDataPlex DX360M4 CPU: Intel Sandy Bridge-EP 8-core 2.7GHz Interconnect: InfiniBand FDR-10 / Pruned Fat Tree	3185	147456	288.0
	IBM BladeCenter HX5 CPU: Intel Westmere-EX 10-core 2.4 GHz Interconnect: InfiniBand QDR / Full Fat Tree	78	8200	52.5
Total		18183	1015336	1926.7

Table 1 Available Tier-0 systems January 2015

Table 2 shows the available Tier-1 systems for January 2015. In this reporting period three Tier-1 systems have been upgraded with additional capacity: at CSC (Finland), EPCC (UK) and SURFsara (Netherlands). The system at KTH (Sweden) has been replaced by a complete new system. CINES (France) disappeared from the list. The total number of sites is now 23 and the total peak performance of 13954 TF/s is an increase of more than 50% compared with that for June 2014.

Site	Architecture	TF/s	#Cores	Total Memory (TB)
BSC (Spain)	Bull Bullx B505 CPU: Intel Westmere-EP 6-core 2.53GHz + Nvidia Tesla M2090 Interconnect: InfiniBand QDR / 2-plane Full Fat Tree	185.8	1536	3.1
CaSToRC (Cyprus)	IBM IDataPlex Cluster CPU: Intel Westmere 12-core 2.67GHz + NVidia M2070 Interconnect: Infiniband QDR	35	1392	5.5
CINECA (Italy)	IBM iDataPlex DX360M3 CPU: Intel Westmere-EP 6-core 2.40GHz + Nvidia Tesla M2070 Interconnect: InfiniBand QDR / Fat Tree	293.1	3288	12.8
CSC (Finland)	Cray XC40 CPU: Intel Haswell 12-core 2.6GHz Interconnect: Cray Aries / Dragonfly	1688	40512	105.5
CSCS (Switzerland)	Cray XE6 CPU: AMD Interlagos 16-core 2.1GHz Interconnect: Cray Gemini / 3D Torus	400.0	47872	46.75
CYFRONET (Poland)	HP BL685c G7 CPU: AMD Interlagos 16-core 2.3GHz Interconnect: Infiniband QDR	61.2	6656	26
	HP SL390s CPU: Intel Westmere-EP 6-core 2.45GHz + Nvidia Tesla M2090 Interconnect: Infiniband QDR	136.8	528	3.6
EPCC (UK)	Cray XC30 CPU: Intel Ivy Bridge 12-core, 2.7GHz Interconnect: Cray Aries / Dragonfly	2550.5	118080	339
FZJ (Germany)	Bull NovaScale R422-E2 CPU: Intel Nehalem-EP 4-core 2.93 GHz Infiniband QDR Full Fat Tree topology	207.0	17664	53
ICHEC (Ireland)	SGI ICE X CPU: Intel Ivy Bridge 12-core 2.4 GHz Interconnect: Infiniband FDR	147.5	7680	20.4
ICM (Poland)	IBM Power 775 CPU: IBM Power7 8-core 3.836GH Interconnect: HFI / 2-level direct graph	78.4	2560	10.2
IDRIS (France)	IBM BlueGene/Q CPU: IBM PowerPC A2 16-core 1.60GHz Interconnect: BG/Q / 5D Torus	836.0	65536	65
IPB (Serbia)	HP Proliant SL250s CPU : Intel Sandy Bridge 8-core 2.60GHz + NVidia M2090 Interconnect: Infiniband QDR	106	1696	3.3
KTH (Sweden)	Cray XC40 CPU : Intel Xeon E5-2698v3 16 core * 2 Interconnect: Cray Aries / Dragonfly	1973	53632	104.7
NCSA (Bulgaria)	IBM BlueGene/ P CPU: IBM PowerPC 450, 850MHz Interconnect: IBM BG/P 3D Torus	27.8	8192	12
NIIF (Hungary)	HP Cluster Platform 4000BL CPU: AMD Magny-Cours 12-core 2.2GHz Interconnect: InfiniBand QDR	5.5	768	1.96

Site	Architecture	TF/s	#Cores	Total Memory (TB)
	HP Cluster Platform 4000BL CPU: AMD Magny-Cours 12-core 2.2GHz + Nvidia M2070 Interconnect : InfiniBand QDR	20	2304	5.6
PSNC (Poland)	SGI UV 1000 CPU: Intel Westmere-EX 8-core 2.67GHz Interconnect: NUMAlink® 5, 2D torus	21.8	2048	16
	SGI/Rackable C1103-G15 CPU: AMD Interlagos 12-core 2.40GHz + Nvidia Tesla M2050 Interconnect InfiniBand QDR, Fat Tree	224.3	5448	10.6
RZG (Germany)	IBM iDataPlex cluster CPU: Intel Sandy Bridge-EP 2x8-core 2.6GHz + Intel Ivy Bridge 2x10-core 2.8GHz partly with Nvidia K20X and XeonPhi Interconnect: InfiniBand 4xFDR14 / Fat Tree	1862.1	83848	260
STFC (UK)	IBM BlueGene/Q CPU: IBM PowerPC A2 16-core 1.60GHz Interconnect: BG/Q / 5D Torus	1258.3	98304	96
SURFsara (Netherlands)	Bull Bullx B720, B710 and R428 E3 CPU: Intel Haswell 2 × 12 2.6 GHz, Intel Ivy Bridge 2 × 12-core 2.4 GHz, and Intel Sandy Bridge 4 × 8-core 2.7 GHz Interconnect: InfiniBand 4 × FDR	1349.0	39904	110.7
	Bull Bullx B515 CPU: Intel Ivy Bridge 2 × 8-core 2.5 GHz + Nvdia Interconnect: InfiniBand 4 × FDR	210.0	1056	6.3
UHeM (Turkey)	HP Proliant BL460 Cluster CPU: Nehalem-EP 4-core 2.67GHz. Interconnect: InfiniBand DDR / Fat Tree	2.5	256	0.75
UiO (Norway)	MEGWARE MiriQuid CPU: Intel Sandy Bridge-EP 8-core 2.6GHz Interconnect: Infiniband FDR/ fat tree	178.6	9984	39.4
VSB-TUO (Czech Republic)	Bull Bullx B510 and B515 CPU: Intel Sandy Bridge-EP 8-core 2.4GHz (180 nodes) and 2.3GHz (27 nodes) Interconnect: Infiniband QDR / Fat Tree	66.0	3312	13.2
WCSS (Poland)	HP Cluster Platform 3000 BL2x220 CPU: Intel Westmere-EP 6-core 2.67GHz Interconnect: Infiniband DDR Full fat-tree	30.0	4848	9.6
Totals		14019	628904	1692

Table 2 Overview of Tier-1 systems in January 2015

3 XSEDE Collaborative Services

3.1 Introduction

PRACE and XSEDE provide peer-reviewed access to high-end HPC resources and services both in Europe and the US. At the end of 2013 PRACE and XSEDE published a common call for proposals to research teams who require interoperable facilities between PRACE and XSEDE. As a result, seven proposals were submitted and three have been selected for support for up to six months to enable their interoperable applications. The selected proposals are:

- Smart Data Analytics for Earth Sciences across XSEDE and PRACE
- Interoperable High Throughput Binding Affinity Calculator for Personalised Medicine
- UNICORE Use Case Integration for XSEDE and PRACE

For each of these projects the objectives and results are described.

3.2 Smart Data Analytics for Earth Sciences across XSEDE and PRACE

The collaborating team that applied for support:

European: Morris Riedel (JSC) and Guiseppe Fiameni (CINECA)

US: Andrew Grimshaw (University of Virginia), Geoffrey Fox (Indiana University), and Shaowen Wang (NCSA)

3.2.1 *Objectives*

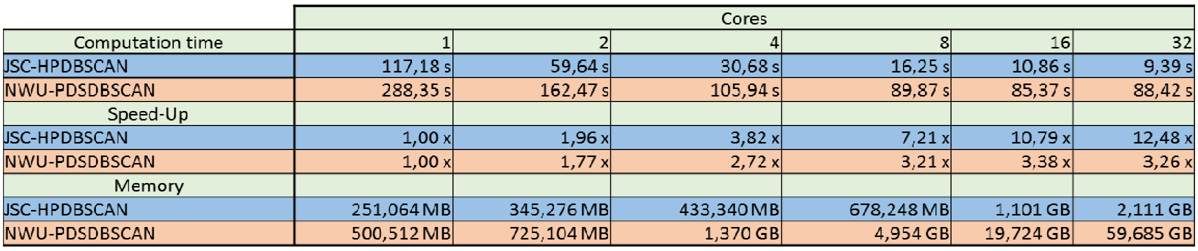
The ever-increasing amount of scientific data arising from measurements or computational simulations requires new ‘smart data analytics techniques’ capable of extracting meaningful findings from ‘pure big data’. XSEDE (including FutureGrid for Map-Reduce), as well as PRACE, already provide resources that could enable efficient and effective data analytics if several technical frameworks and data analysis packages were available. The key goal of this proposal is to make these packages available and to validate their usefulness in conjunction with earth science case studies. These are, for instance, quality control through outlier detection with PANGAEA data [6], longitude/latitude/altitude correlations with IAGOS data [7], event tracking analytics with NASA spatial computing datasets [8], continuous seismic waveforms analysis for earthquakes monitoring with INGV data [9], etc.

It is also aimed to improve the scalability of key algorithms such as those of the Scalable GIS Analytics, for projecting and transforming geospatial big data into a common coordinate reference framework. As one result, future data-intensive scientific computing challenges shall be prepared by enabling data science groups that collaborate across the US and the EU to better utilize ‘smart data analytics’.

The underlying software integration infrastructure will consist of standards compliant software mechanisms already in use by both XSEDE and PRACE: UNICORE, GridFTP, and the Global Federated File System (GFFS provided by Genesis II) [10][11].

3.2.2 *Results*

Within this activity a parallel and scalable version of the Density-Based Scan (DBSCAN) algorithm has been developed. In the field of statistical data mining, DBSCAN is known for its effectiveness for automated outlier detection (i.e. JSC-HPDBSCAN). After evaluation of other parallel solutions (e.g. NWU-PDSDBSCAN) it became clear that, for the case study of automated quality control through outlier detection with PANGAEA datasets, a more scalable version for parallel HPC systems in XSEDE and PRACE is needed. The performance comparisons are shown in Figure 1 and as a normal MPI code it can take advantage of using UNICORE between XSEDE and PRACE. The results will be presented at the Research Data Alliance (RDA) Big Data Analytics (BDA) interest group session of the RDA 5th plenary in San Diego in March 2015 and the HPC Admin Workshop of the LinkSCEEM2 project in Cyprus in January 2015 in order to distribute the results more broadly.



	Cores					
Computation time	1	2	4	8	16	32
JSC-HPDBSCAN	117,18 s	59,64 s	30,68 s	16,25 s	10,86 s	9,39 s
NWU-PDSDBSCAN	288,35 s	162,47 s	105,94 s	89,87 s	85,37 s	88,42 s
Speed-Up						
JSC-HPDBSCAN	1,00 x	1,96 x	3,82 x	7,21 x	10,79 x	12,48 x
NWU-PDSDBSCAN	1,00 x	1,77 x	2,72 x	3,21 x	3,38 x	3,26 x
Memory						
JSC-HPDBSCAN	251,064 MB	345,276 MB	433,340 MB	678,248 MB	1,101 GB	2,111 GB
NWU-PDSDBSCAN	500,512 MB	725,104 MB	1,370 GB	4,954 GB	19,724 GB	59,685 GB

Figure 1 Comparisons of different implementations of the parallel BDSCAN algorithm

GFFS was available at the University of Virginia already when the project started. In the meantime the project team has installed GFFS also on the JUQUEEN system of Jülich Supercomputing Center (JSC). Since there was no GFFS version available before for Power architectures, like the one of JUQUEEN, this had to be developed first. During the installation configuration and security issues, e.g. firewall settings, had to be harmonized between PRACE and XSEDE. As a result GFFS now allows a user-friendly interoperable way to exchange data between XSEDE and PRACE. Additionally, a wrapper has been implemented to be able to use UNICORE on top of GFFS. Now it is possible to use GFFS for data exchange within UNICORE jobs. Amongst others this allows the continuous waveforms analysis for earthquake monitoring with INGV data.

During the project the data analytic tool SciDB [12] has also been evaluated. SciDB is an array-oriented database technology that is used for event tracking analytics on NASA spatial computing datasets. As a result of the evaluation, it turned out that SciDB can not be used on the PRACE and XSEDE resources in a reasonable way since it does not work well with the respective batch schedulers of the PRACE and XSEDE machines. As such, this technology is currently not suited for XSEDE-PRACE interoperability projects and alternatives will be explored (e.g. RASDAMAN).

3.3 Interoperable High Throughput Binding Affinity Calculator for Personalised Medicine

The collaborating team that applied for support:

European: Prof Peter V. Coveney (UCL) and Prof Dieter Kranzlmuller (LRZ/LMU)

US: Prof. Shantenu Jha (RADICAL, Rutgers)

3.3.1 *Objectives*

Personalized medicine is a unique medical approach in which personal medical information belonging to a patient is used to customize health treatment and healthcare. Simulation offers the possibility to assess the effectiveness of certain courses of treatment before they are administered using a patient specific model, in order to choose the best. This work will aim at cancer patients, for which there are many FDA approved drugs available; one needs to know which ones to select for any given patient or patient group.

To automate the selection of available drugs (via calculation of drug-binding affinities) – and ultimately to make it available for routine clinical use –the Binding Affinity Calculator (BAC) [13] has been developed to build the necessary patient specific models required to simulate drug performance, a process which requires a complex number of steps in order to customise a generic model with patient specific information, and then run the calculations. Often results are needed in short order, to meet timescales for clinical decision making; these are frequently needed within a timescale of 1-3 days. The predictions must be accurate, reliable and reproducible, thus imposing strong constraints on automation and throughput. Recently a protocol that can produce converged, robust and reproducible free energies for protein-ligand interactions has been released [14][15].

As part of novel scientific collaboration between CCS/UCL and RADICAL team at Rutgers, BAC is being redesigned and refactored to develop a High Throughput-BAC (HT-BAC) for greater throughput, scalability, and interoperability across XSEDE and PRACE. It will also provide support for more effective execution modes and application types and classes. In doing so, the tools that we are developing will exploit interoperability and support collective resource management. The project will employ and harden tools that will support data sharing across XSEDE and PRACE.

3.3.2 *Results*

The PRACE centers at LRZ (Munich, Germany), EPCC (Edinburgh, UK) and HARTREE, (Daresbury, UK), have provided access to their systems and enabled the interoperability among the systems needed for data movement and cross-system job submission capabilities [16].

3.4 UNICORE Use Case Integration for XSEDE and PRACE

The collaborating team that applied for support:

Sandra Gesing on behalf of the MoSGrid community and the computational radiation physics community.

European: Technische Universität Dresden, Helmholtz-Zentrum Dresden-Rossendorf and University of Tübingen

US: University of Notre Dame

3.4.1 *Objectives*

This project develops the UNICORE [17] integration of two use cases for the joint support of XSEDE and PRACE: the first one targets the molecular simulation community, the second one the computational radiation physics community.

The MoSGrid (Molecular Simulation Grid) project [18][19][20] offers a web-based science gateway supporting the community with various services for quantum chemistry, molecular modelling, and docking. The science gateway tackles the problem to increase the usability of compute-intensive and data-intensive tools and workflows in the three domains. Users gain access to distributed computing infrastructures (DCIs) via intuitive user interfaces for sophisticated tools, specialized workflows, and distributed repositories. The users can apply MPI versions of computational chemistry tools (e.g., NWChem, Gromacs). Currently, the MoSGrid community consists of about 200 users from several European countries.

The computational radiation physics community represented by the Helmholtz-Zentrum Dresden-Rossendorf (HZDR) focuses on the generation of advanced laser-driven sources of particle and X-ray beams. One of the major goals is to "simulate what is measured", aiming to reproduce experimental measurements and connecting them to the fundamental plasma processes on the single-particle scale. With optimized codes, it is now possible to simulate several steps per second. This has opened the door for direct comparison between experiment and simulation. Thus, highly automated organization of large-scale parameter surveys is essential to achieving timely feedback to experiments.

The technical challenges of the two use cases are quite different, but they complement each other and so are beneficial for other UNICORE projects in XSEDE and PRACE. The main technical challenge in MoSGrid is to extend the portal infrastructure for the use of the XSEDE and PRACE infrastructure via UNICORE and according credentials. In the radiation physics project, the main challenge is to make the corresponding tools on both XSEDE and PRACE systems available via UNICORE to allow for the exchange of common workflows, which can be applied on both infrastructures.

3.4.2 *Results*

The PRACE centre at FZJ (Jülich, Germany) has provided access to the JUQUEEN system and enabled UNICORE for both the MoSGrid and computational radiation community. Gromacs and NWChem are installed and can be reached via the provided UNICORE registry.

The MoSGrid portal has been enhanced to be able to use UNICORE for job submission to and between PRACE and XSEDE resources by using X.509 certificates for the job submission. Users can create a Proxy Certificate on XSEDE resources in case they own an XSEDE account but not a certificate. PRACE has provided the registries to allow for access to those resources where the Gromacs and NWChem tools are installed. Thus, workflows incorporating these tools can be used via the portal to PRACE and XSEDE.

4 Conclusions

In this period, WP6 has continued the successful operation of the PRACE common services for the Tier-0 and Tier-1 sites. The operational procedures have continued to be successful in maintaining a reliable and available set of services. The PRACE Security Forum coordinated all security related issues. The reported security vulnerabilities have been managed efficiently.

PRACE provided support to enable collaborative services with XSEDE. The results and experience can be used to provide support to other collaborating teams that use resources from both PRACE and XSEDE.

WP6 of PRACE-3IP, building on the work of previous PRACE projects, has laid the foundation for a successful continuation of the operation of the PRACE common services for the coming years.