SEVENTH FRAMEWORK PROGRAMME
Research Infrastructures

INFRA-2010-2.3.1 – First Implementation Phase of the European High Performance Computing (HPC) service PRACE

PRACE-1IP

PRACE First Implementation Project

Grant Agreement Number: RI-261557

D4.2
Tier-0 and Tier-1 Providers Relationship
Final

Version:	1.0
Author(s):	Michael Browne, ICHEC
Date:	26.06.2011

Project and Deliverable Information Sheet

<table>
<tr><td rowspan="8">PRACE Project</td><td colspan="2">Project Ref. №: RI-261557</td></tr>
<tr><td colspan="2">Project Title: PRACE First Implementation Project</td></tr>
<tr><td colspan="2">Project Web Site: http://www.prace-project.eu</td></tr>
<tr><td colspan="2">Deliverable ID: < D4.2></td></tr>
<tr><td colspan="2">Deliverable Nature: Report & Recommendations</td></tr>
<tr><td rowspan="2">Deliverable Level:
PU / PP / RE / CO *</td><td>Contractual Date of Delivery:
30/06/2011</td></tr>
<tr><td>Actual Date of Delivery:
30/06/2011</td></tr>
<tr><td colspan="2">EC Project Officer: Bernhard Fabianek</td></tr>
</table>

* - 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

<table>
<tr><td rowspan="6">Document</td><td colspan="2">Title: < Tier-0 and Tier-1 Providers Relationship></td></tr>
<tr><td colspan="2">ID: <D4.2></td></tr>
<tr><td>Version: <1.0></td><td>Status: Final</td></tr>
<tr><td colspan="2">Available at: http://www.prace-project.eu</td></tr>
<tr><td colspan="2">Software Tool: Microsoft Word 2007</td></tr>
<tr><td colspan="2">File(s): D4.2.docx</td></tr>
<tr><td rowspan="4">Authorship</td><td>Written by:</td><td>Michael Browne, ICHEC</td></tr>
<tr><td>Contributors:</td><td>Michael Browne & JC Desplat Irish Centre for High-End Computing (ICHEC) / NUI Galway, Jacko Koster UNINETT Sigma, Stephane Requena GENCI Giovanni Erbacci, Sergio Bernardi & Andrew Emerson CINECA.</td></tr>
<tr><td>Reviewed by:</td><td>D. Erwin, FZJ
Vit Vondrak, VSB-TU Ostrava</td></tr>
<tr><td>Approved by:</td><td>MB/TB</td></tr>
</table>

Document Status Sheet

Version	Date	Status	Comments
0.1	18/February/2011	Draft	Initial Version
0.2	03/June/2011	Draft	Circulated to all WP4 partners
0.4	10/June/2011	Draft	Passed to internal review
0.5	21/June/2011	Draft	Revised following internal review, circulated to all WP4 partners
0.6	22/June/2011	Draft	Minor corrections
1.0	26/June/2011	Final version	For approval

Document Keywords

Keywords:	PRACE, HPC, Research Infrastructure

List of Figures

List of Tables

References and Applicable Documents

[1] From Desktop To Teraflop: Exploiting the U.S. Lead in High Performance Computing, 1993
[2] PRACE-1IP Deliverable 4.3.1
[3] Top500 list, http://www.top500.org
[4] DEISA2 2.1.3 Deliverable http://www.deisa.eu/publications/fp7-deliverables/pm24/DEISA2-D2-1.3.pdf

List of Acronyms and Abbreviations

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)
CPU	Central Processing Unit
DEISA	Distributed European Infrastructure for Supercomputing Applications. EU project by leading national HPC centres.
EC	European Community
EESI	European Exascale Software Initiative
ESFRI	European Strategy Forum on Research Infrastructures; created roadmap for pan-European Research Infrastructure.
FTE	Full Time Equivalent
FZJ	Forschungszentrum Jülich (Germany)
GB	Giga (= 2^{30} ~ 10^9) Bytes (= 8 bits), also GByte
GCS	Gauss Centre for Supercomputing (Germany)
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
GPGPU	General Purpose GPU
GPU	Graphic Processing Unit
HET	High Performance Computing in Europe Taskforce. Taskforce by representatives from European HPC community to shape the European HPC Research Infrastructure. Produced the scientific case and valuable groundwork for the PRACE project.
HLRS	Hoechstleistungsrechenzentrum Stuttgart, Germany
HPC	High Performance Computing; Computing at a high performance level at any given time; often used synonym with Supercomputing
ICT	Information and Communication Technologies
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique (represented in PRACE by GENCI, France)
IoT	Institute of Technology, 3rd level Irish Education Institute which is not a university and generally focuses on technical and vocational disciplines often to degree level
JSC	Jülich Supercomputing Centre (FZJ, Germany)
LINPACK	Software library for Linear Algebra
LRZ	Leibniz Supercomputing Centre (Garching, Germany)
MHz	Mega (= 10^6) Hertz, frequency =10^6 periods or clock cycles per second
MPI	Message Passing Interface
NREN	National Research and Education Network
OpenMP	Open Multi-Processing
PFlop/s	Peta (= 10^{15}) Floating-point operations (usually in 64-bit, i.e. DP) per second, also PF/s
PRACE	Partnership for Advanced Computing in Europe; Project Acronym

SME	Small and Medium sized Enterprise
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
WP	Work Package a unit of defined work with the main project which may consist of several subtasks.
x86	Intel x86 or compatible processor or core, typically 64 bit (x86-64)

Executive Summary

The HPC ecosystem can be classified by means of tiers, corresponding to levels of performance and operational policy and the defining characteristics and roles of tiers 0,1 and 2 are largely agreed. The development of this ecosystem and the promotion of Tier-0 usage is a formidable challenge both organisationally and technically. The delivery of the PRACE-2IP project beginning in Q3 2012 will expand the PRACE RI user community to include Tier-1 users, drawn from a large body of researchers already active at Tier-1. This reinforces the need for development of the ecosystem with respect to both the users and providers.

The PRACE AISBL is now very representative with 21 members. This broad membership has meant that the consultations undertaken in the course of its work by this task with most other Work Packages particularly 2, 3, 5 and 7 have resulted in contacts with the membership as a whole. The expansion of the HPC ecosystem in parallel with the progression of the PRACE Implementation Phase projects will bring with it new external relationships e.g. joint training with LinkSCEEM-2. The US TeraGrid programme has been identified as a potential partner in relationships that go beyond training. This task's promotion of inter-tier migration to the HPC-Europa2 community addresses not only a wide geographical area but also a generally youthful demographic who will form part the HPC community in the future.

Together with Work Package 7 (Task 7.4.1) this task has participated in conducting an extensive user community survey and clearly established the potential for a well supported community to effect change in the codes they use due to their high level of involvement in code, demand for greater compute resources and understanding of the problem at hand. Greater availability of Tier-1 resources was identified as a key enabler of subsequent Tier-0 use, further validating a key aim of the PRACE-2IP project.

Case studies conducted by the task at the levels of a Tier-0 hosting country and two smaller countries with Tier-1 resources have provided insights into both ecosystem development and Tier-0 promotion. In each case dissemination and promotion of the opportunities offered by PRACE up to and including Tier-0 access is undertaken at national level. Assistance with the application process is readily available to researchers. Particularly in the case of the smaller nations the notion of the integration of Tier-0, and in time alternative Tier-1, access as part of a continuum of service available to their respective communities albeit with alternative access mechanisms is seen as crucial. Even in France, a Tier-0 hosting country, it is recognised that enablement work must be undertaken at the Tier-2 level if one is to foster a Tier-0 community to the extent that part of the response to Tier-0 provision has been to strengthen efforts at Tier-2 nationally.

Relationships that pertain to the ecosystem are key to its development. Within the PRACE project and RI relationships promote coordination of support, technical and organisational activity serving end users and service providers. The DEISA programme while not primarily concerned with inter-tier migration in effect faced the problem because of the performance differential that emerged between the systems involved, which were made up of a range of architectures. Consequently resource providers, scientific support staff and users developed a body of technical experience in this domain. The imperative to provide for a consistent experience for users across tiers and systems will primarily develop relationships between providers. PRACE has well developed plans underway for community involvement e.g. community code petascaling and a user forum. Recommendations have been made regarding the attributes of relevance to PRACE when considering an external relationship.

This task's contact with the HPCWorld project has informed the guidelines it is developing which will in turn feed into future allocation processes through which users will access resources. Users too must recognise that Tier-0 represents not only a shift in what can be achieved relative to before but also a commensurate shift in competition and expectation. Thus there should be openness to approaches needed to building the best possible project that may not have been required in the context of Tier-1. Collaborations should be considered at least at European level and a strategic approach to ensuring that the necessary software requirements can be met within realistic time scales is important.

1 Introduction

This document endeavours to capture activities undertaken to date by task 4.2 and actions that remain on going. The remit of this task two key actions are identified:

- Develop a set of best practices to prepare the user to access Tier-0.
- Expand the relations between Tier-0 and Tier-1 resource providers.

Best practices that have been distilled through consultation and representative case studies are outlined. The expansion of relationships comes in two main ways; through expanded relationships between PRACE partners especially in the context PRACE-2IP and through relationship with groups external to the PRACE projects. Both scenarios are currently on going and the relevant benefits, motivations and limits are described.

Over the course of the design of the task 4.2 description and subsequent delivery of this document, the PRACE-2IP programme has progressed rapidly from a proposal which would follow logically from PRACE-1IP to an approved project due to get underway formally in Quarter 3 of 2011. With this in mind, where there are related and imminent developments that are strictly part of the PRACE-2IP programme they will be referred to, recognising that in time both projects and indeed the Preparatory Phase will be seen as a continuum rather than distinct projects, especially by those external to the projects. The principle effect of PRACE-2IP with regard to this document will be the integration of Tier-1 resources into the PRACE ecosystem within the timeframe of PRACE-1IP. On going work in WP2 task 4 is looking in detail at the formidable challenges that exist in the assessment of metrics in the short term in relation to research infrastructures, thus care has been taken not to duplicate that effort.

1.1 Guide to the document

In accordance with the approach taken in framing the task description this document can be seen to address two main themes that of relations between Tier-0 and Tier-1 providers and the best practises for the preparation of users.

Section 2 focuses on the former, relationships. The defining attributes of Tiers-0 and 1 are discussed along with their existing roles to provide a contextual basis. New and expanded relationships are also discussed. Section 2.5 gives recommendations for the attributes of relevance to PRACE when considering an external relationship including examples.

Section 3 draws on the substantial efforts undertaken by both DEISA and PRACE-1IP WP7 to survey their user communities to assess their needs, preferences and work patterns. The results of which are analysed here from the perspective of the task description.

Three case studies have been undertaken to address both the areas of inter-tier relationships and best practices, as they exist in three European countries. One example of a large country with both Tier-0 and Tier-1 facilities is examined in Section 4 which describes the ecosystem and practices in France. Section 5 examines the situation faced by two smaller countries

(Ireland & Norway) which notably do not currently possess or plan to procure Tier-0 facilities.

Section 6 defines best practices for the preparation of Tier-0 usage. It should be noted that the preparation of users is understood to encompass not just the needs of specific individuals or research groups but those of the wider peer group from which they come and also technical staff of existing HPC service providers with whom they most likely collaborate on a day to day basis.

Section 7, draws together the conclusions and key findings reached in the previous sections. Where appropriate recommendations are made in both sections 2 6 and 7.

1.2 Other related or connected material

When considering the content of this document the reader may benefit from referring to the following Work Packages and tasks that are of relevance. Additional references to specific external material are provided throughout the document. The following table is a quick guide to related work in PRACE-1IP.

1IP	WP2	Deliverable 2.4 examines assessment and reporting metrics in detail.
	WP3	WP3 extensively covers training as it relates to PRACE system usage.
	WP4	Deliverable 4.3.1 defines Tier-1 and relates to the role of new European partners.
	WP5	Deliverable 5.2.1 discusses industrial relationships and related US activities.
	WP6	WP6 relates to technical and operational aspects of tier integration.
	WP7	Task 7.1 provides support to applicants in the form of compute resources and expertise in improving their codes for later production work. Task 7.2 is focused on helping communities as a whole.

Table 1: Related works

2 Inter-Tier Relationships

2.1 Introduction

A strong HPC ecosystem and its further evolution is a matter of primary importance for PRACE RI. The PRACE Preparatory Phase project has conducted a wide analysis of the HPC ecosystem and pointed out that enablers for sustained high performance computing include scalable code development, integration with national infrastructures, sufficient data repositories, high capacity networking and competent people. Thus the only way to succeed in deploying the European centres is to address the whole HPC ecosystem by linking the Tier-0 systems and related services closely to the existing infrastructures.

Moreover the involvement of the national supercomputer centres is mandatory as well as the foundation for a sustainable and persistent infrastructure that goes beyond the usual duration of a project. PRACE AISBL and the PRACE projects are clearly addressing both of these aspects.

The promotion of the usage of Tier-0 resources needs focused action aimed at steering the development of sound and successful relationships among all the entities that play a central role in the provision of HPC resources. In order to further expand and evolve such an ecosystem PRACE has identified two specific actions that aim to expand inter-tier relationships, starting with an in depth analysis of what it is actually in place followed by a proposal for establishing new relationships.

Presently, the inter-tier relationships among resource providers are mostly accommodated within the PRACE AISBL and the PRACE implementation projects. PRACE-1IP and 2IP projects are addressing the integration of Tier-1 and Tier-0 resources. That will not only carry on the successful experience of DEISA but will also expand it including new national Tier-1 entities through PRACE membership.

Today the relationships between Tier-0 and Tier-1 providers are basically arbitrated by PRACE AISBL and might be seen as *intra-PRACE* relationships where PRACE members are actively defining the policies, the guidelines and the procedural and operational models. It has to be noted that even if PRACE does not include all the European countries nevertheless it embraces almost all of the national Tier-1 centres and providers.

A further category of inter-tier relationships comes from Tier-1 providers that are not themselves national centres but are region based, cross-country entities instead. As notable examples for such entities are LinkSCEEM-2 and HP-SEE, both FP7 funded projects, whose objective is the development of HPC regional infrastructures. Other examples of Tier-1 resource providers include HPC infrastructures that address the needs of very specific communities and are international in nature and scope.

2.1.1 *System Tiers*

It has become common to discuss the various levels of performance available across the spectrum of HPC systems in terms of the so-called Branscomb pyramid. The term derives from an influential report[1] compiled for the NSF by a panel lead by Lewis Branscomb. A pyramid of resources from desktop-based systems to the very fastest supercomputers at the uppermost tier, Tier-0, are described.

In the context of PRACE, of greatest interest are the uppermost tiers 1 and 0, and these are discussed in greater depth in the remainder of this Section. It is important to note that as Branscomb’s report covers the US, as such all tiers are represented within that single nation. The situation in Europe is quite different in that only a small number of countries will have all tiers in place. The majority of countries will have only tiers up to 1 in operation. Indeed in the smaller countries or countries with a nascent HPC community Tier-1 may also be absent or if present then only at a scale at the extreme end of what constitutes Tier-1 (e.g. performance of a few tens of TFlops).

2.2 Tier-0

2.2.1 *Tier-0 Definition*

Tier-0 is the term used to describe the category of systems which provide the highest levels of performance and scale currently available. Such systems are very costly and only a handful of systems which are operational at a given time will be considered to belong to this Tier. The following is the formal definition of Tier-0 as used by PRACE. This document bases its use of the term on this definition.

[1] From Desktop To Teraflop: Exploiting the U.S. Lead in High Performance Computing, 1993

Definition of what constitutes Tier-0 as used by PRACE heretofore:

PRACE has a working definition of a Tier-0 system, it was drawn up during the PRACE-PP project and has subsequently been cited in the deliverable produced by Work Package 4.3.1. This definition is quite short and is as follows:

- The cycles of the system are allocated through the PRACE peer review process (in other words: the governance over the system is in the hands of PRACE)
- The system is part of a 100 Mio Euro installation, and is of capability-class (meaning that jobs can use the full machine)

2.2.2 *Existing Tier-0 systems*

In addition to their role in the PRACE Research Infrastructure (RI) these systems also play central roles in their respective national HPC services. It is also important to note that the rate of advancement of the very fastest systems, as documented by the Top500[1] list, is extremely rapid. It is generally accepted that they increase in speed by a factor of 1.8 per annum, which is in excess of the widely accepted rate of advancement predicted by Moore's law. Currently there are ten systems worldwide listed in the Top500 with a peak performance of over one Petaflop/s.

Jugene

Jugene is an IBM Blue Gene/P system hosted by the Gauss Centre for Supercomputing at Forschungzentrum Juelich (FZJ), Jülich, Germany. Significantly it was the first PRACE system and has been available for successful PRACE resource applications since July 1st 2010. It is made up of 294,912 PowerPC 450 cores, with each node having 4 cores along with 2GB of memory yielding a total of 147TB and has a peak performance of one Petaflop/s.

CURIE

The CURIE system is hosted by the CEA at Bruyères-Le-Châtel in France and is funded by GENCI. CURIE is a BULL system based on Intel x86 processors spread across a mix of fat and thin nodes interconnected via Infiniband (QDR). In its final phase, end of 2011, the peak performance will be over 1.7 Petaflop/s. There are 92,160 cores each with 4GB of memory.

Hermit

Hermit is a CRAY XE6 system hosted by the Gauss Centre for Supercomputing at Hoechstleistungsrechenzentrum Stuttgart (HLRS), Stuttgart, Germany. Once fully operational in November 2011 it will be the 3rd PRACE system. It is made up of 113,664 cores based on AMD Interlagos processors. A small test and development system is available at the time of writing. Each node will have either 32GB or 64GB of memory. The system will have a peak performance of 1 petaflop.

2.2.3 *Planned Tier-0 systems*

Further Tier-0 systems are currently in the planning stages. Given that these systems are yet to be procured, there is considerable uncertainty regarding their technical details and timeframe for their operational availability. The system formally announced to date is the SuperMUC system which will be hosted by the Gauss Centre for Supercomputing at Leibniz-Rechenzentrum (LRZ) and will be operational in 2012. It will have 110,000 Intel x86 cores

[1] http://www.top500.org

and a peak performance of roughly 3 Petaflop/s. It will be based on the IBM System x iDataPlex server platform.

It is also worth noting that as systems at Tier-0 and 1 progress through their overlapping lifecycles of typically 3-4 years or possibly undergo upgrades the difference in compute performance between them will vary dramatically. For instance, it likely that in the near future several European Tier-1 systems will have a performance of roughly 0.8 Petaflop/s. However operational policies and access requirements will still differ.

2.3 Tier-1

2.3.1 *Tier-1 Definition*

Tier-1 systems exist in far greater number than Tier-0, largely due to their more modest size and cost and widespread use for the provision of national HPC services. The majority of countries represented by PRACE possess systems of this scale. The nature of such machines and the support processes associated with them can vary somewhat. In recognition of this diversity a working definition of Tier-1 has been prepared by PRACE-1IP WP4.3. This document bases its use of the term on this definition, as outlined below.

Definition of what constitutes Tier-1 as put forward by the Work Package 4.3 deliverable (D4.3.1 [https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/d573070/D4.3.1.pdf]):

In order to participate in the PRACE Tier-1 resource exchange model, partners must be able to offer resources which meet the **Tier-1 system** criteria and an operational model which meets the **Tier-1 centre** criteria. (Note that we expect a **Tier-1 system** to be a single resource i.e. not a grid, but that the services provided by the **Tier-1 centre** may be distributed, if this is the usual way that HPC services are provided in that country). Partners must also provide a **Tier-1 contribution** of sufficient size to make their participation in a resource exchange scheme viable.

Tier-1 centre:

- offers user training in HPC programming;
- offers technical and user support via a help desk;
- offers applications support (code development and optimization);
- offers advanced development platforms for HPC (i.e. pre-cursors of future Tier-0 architectures);
- offers services to support Grand Challenges and provide a ramp to allow scientists to move to Tier-0 systems;
- is willing to implement necessary middleware for integrating the site into the PRACE Tier-1 infrastructure;
- is willing to sign a SLA with PRACE (i.e. to address a list of pre-requisite requirements).

Tier-1 system:

- cycles or science projects allocated via external (national) peer review;
- supports the work of researchers nationally and/or within a region of Europe (if such arrangements are in place from the Tier-1 centre);
- has a hardware and software configuration able to meet the computational needs of a range of science areas;

- has a high enough performance to meet most of the national requirements below Tier-0 (will typically be no more than one order of magnitude lower than the Tier-0 systems);
- serves as a development platform for the most scalable codes (which will be directed towards the Tier-0 systems);
- has batch queues configured to support capability computing (use of maximum available computing power to solve a problem in the shortest amount of time) although the system may support both capability and capacity computing;
- is connected to the PRACE infrastructure with sufficient network/bandwidth to facilitate the exchange of large datasets (with other Tier-1 and Tier-0 systems) and participates in PRACE Tier-1 resource exchange.

Tier-1 Contribution:

- usually 5% (or more) of the resources available on a Tier-1 system annually (the typical DEISA contribution);
- resources committed should be sufficient for at least two DECI projects per year (typical DEISA contribution was for ~5 DECI projects per year);
- for smaller centres where 5% of the total resources is too small to support two DECI projects per call, there is the possibility of banking resources and contributing a larger amount of resources to alternate calls e.g. 10% to every second call (the feasibility of this approach will be investigated during the resource exchange pilots).

2.3.2 *Tier-1 systems*

Systems in this tier can share a great deal of technical expertise with those in Tier-0, differing significantly only in scale. However the role they play in their respective HPC ecosystems is quite different. Furthermore this role will differ from country to country and is likely to be highly dependant on the size and maturity of the HPC community in a given country. For example in the case of Ireland, the Stokes system (c.40TFlops peak) forms the highest performance layer of the national service and serves a diverse user community. However in the context of other European Tier-1 systems its peak performance is modest. Further specific examples of this are cited in the three case studies discussed in section 5.

Typically these systems will provide a mix of capability and capacity service. They will be built using well proven technology and are expected to be highly reliable. They may be hosted by dedicated national centres or academic institutions. The above definition is at pains to emphasise that these systems are not merely defined by their associated hardware but importantly also the services and support that exist around them which enable the maximum benefit to derived from them, e.g. user software support and a robust online helpdesk.

Table 2 lists the systems which are currently planned for applicants to the PRACE DECI calls from 11/2011 or in some cases from 12/2012.

Country	PRACE Partner	System Name	Architecture	Performance LINPACK [Tflops]
Germany	JUELICH	Juropa	Bull Nehalem Cluster	183
	GCS/RZG	vip	IBM Power6	98
	GCS/RZG	genius	IBM BG/P	48
	GCS/LRZ	LRB-II	SGI Altix 4700	56.5
	GCS/LRZ		SuperMUC Migration System	60
	GCS/HLRS	Laki	NEC Nehalem Cluster	112.8
France	GENCI/IDRIS	BABEL	IBM BG/P	139
	GENCI/CINES	JADE	SGI ICE 8200	267
UK	EPCC	HECToR	Cray XE6	360
Spain	BSC	MareNostrum	Cluster	63.8
Finland	CSC	Louhi	Cray XT4/5	76.5
Switzerland	ETH Zurich	Rosa	Cray XT5	211.51
Netherlands	NCF/SARA	Huygens	IBM p575 HydroCluster	50
Sweden	SNIC/KTH	Lindgren	Cray XE6	305
Italy	CINECA		IBM SP6	78.7
Poland	PSNC		Bull MESCA Cluster	20
Norway	SIGMA			80
Ireland	NUI Galway	Stokes	SGI ICE 8200	36.56
Turkey	UYBHM	Karadeniz	Nehalem Cluster	
Cyprus	CASToRC			
Bulgaria	NCSA	EA”ECNIS”	IBM BG/P	23.42
Czech Republic	VSB-TUO	Ostrava	x86_64 cluster	54
Serbia	IPV	Blue Danube	Nehalem Cluster	36

Table 2: Non Tier-0 systems to be available in the PRACE DECI Call 11/2011 – 12/2012

2.4 Progression of existing relationships

2.4.1 *Introduction*

The inter-tier relationships among resource providers at intra-PRACE level address specific topics related to different types of interaction:

- Support to end users and their communities (code enabling, application porting, development, training and education, services).
- Technical and technological activities (operations, evolution, interoperability, integration).
- Organizational activities (strategies, common/shared resource allocation policies, evolution plans, long term/global collaboration programs).

The PRACE-PP proved to be a very effective means for establishing relationships that afterward were formalized within the PRACE AISBL. It set the basic strategy for the HPC ecosystem build up and provided the necessary background for the PRACE-1IP plan to focus on the expansion of the ecosystem.

It is rather clear that the majority of the relations are actually happening at PRACE implementation phase project level. There the members contribute with their efforts to the development of the ecosystem. At AISBL level most of the initiatives seem to aim mostly at consolidating those relations by means of appropriate council decisions and governance actions.

2.4.2 *User and community migration between tiers*

Most researchers are not familiar with computer resources available at the Tier-0 level. This is understandable given that on daily basis users more frequently use personal workstations or departmental clusters. Researchers may also have access to larger facilities at regional (also called Tier-2) or national computer centres, but there is considerable competition for these resources and grants obtained may be well below Tier-0 levels. The application software available to the communities as one might expect reflects the computer hardware that users commonly have access to, i.e. very often not suitable for Tier-0 with low parallel scalability and perhaps also limited hardware portability. Clearly there are number of challenges from both the user and resource provider perspective. Throughout this document issues relating to the migration of users and their communities from Tier-1 or below to the Tier-0 level available from the PRACE-RI are referred to. Previous experience of the DEISA project is drawn on as are early experience from PRACE itself, particularly from the preparatory calls and direct access calls, and the application enabling activities currently underway as part of WP7 in PRACE-1IP.

Previous role of DEISA

The DEISA project started in 2002 and its latest incarnation as DEISA2 officially finished on April 30th 2011, although most partners promised to maintain some aspects of the infrastructure (e.g. user support) for some months after this date to facilitate the transition to Tier-1 access for PRACE-2IP. Computational resources from DEISA were made available to individual researchers mainly via the DEISA Extreme Computing Initiative (DECI) calls, although some resources were also set-aside for user communities via the Virtual Community[1] model. Although not a Tier-0 project, the experience of DEISA is important both for users and resource providers.

DEISA developed from a collection of peer machines into a HPC ecosystem in its own right as the performance gap between the largest and smallest machines grew over time. This provided valuable experience about the benefits of having access to a range of architectures and sizes of machines within the same ecosystem so that both scientists and applications support staff could gain first hand experience of the role of smaller machines in providing a ramp of larger machines and in identifying and solving bottlenecks in scaling codes.

The inclusion of Tier-0 in PRACE represents a big difference with respect to DEISA, but a number of the practices introduced by DEISA have been adopted for all users by many computer centres so its role in tier migration should not be underestimated. A final important point is that many ex-DEISA users will be, or already are, PRACE project applicants and project holders.

Findings from PRACE application process

Insights on the migration process may also be obtained directly from PRACE itself, in particular by analysing the results of calls for PRACE resources such as the regular Tier-0 applications or preparatory access. For instance, in the second Regular Call which closed in January 2011, 46 proposals were received asking for a total of 167 million hours on CURIE

[1] http://www.deisa.eu/publications/fp7-deliverables/pm24/DEISA2-D2-1.3.pdf

and about a billion core hours on Jugene (both Tier-0 computers); the core hours available for the platforms were 40 million and 360 million, respectively. From these 46 proposals, 17 projects were granted, of which 7 were awarded to CURIE and 10 to Jugene. For the preparatory access calls, designed to prepare projects for Tier-0 and for which applications are accepted at any time, 24 projects have been accepted so far from a total of 30 submitted. All rejections were based on unsupported technical requirements or a misunderstanding of the nature of the call e.g. production work.

Operations and evolution of the infrastructure

The PRACE distributed research infrastructure will be operated and presented to the users as a single research infrastructure that provides seamless access to Tier-0 and Tier-1 resources and services. All the related operational activities have to be carried out in a coordinate fashion that require Tier-0 and Tier-1 providers to adopt a common framework of services and related organizational structure. They will have to work closely in order to synchronise the provision of services and their deployment and take into account the layer of PRACE services aimed to integrate the infrastructure.

All the technical and organizational aspects are specifically addressed in PRACE-1IP WP6 and will be also developed in PRACE-2IP with more specific focus on inter-tier aspects.

The scope of the infrastructure development and evolution is a further expression of the relationships between inter-tier resource providers. The scope of the prototype activity in the PRACE projects encourages levels of involvement and common focused effort in technology evaluation by PRACE members.

Although the objective of prototypes' work was aimed at providing support to the evolving Tier-0 architectures, the benefits of this work can be extended to Tier-1 level as is foreseen in the work plan of the PRACE-2IP project, where the integration of Tier-1 and Tier-0 infrastructures is specifically addressed.

Development at the community level

Although some researchers write and develop their own codes, many others use applications written and maintained by other researchers within their own user communities. Consequently, migration to Tier-0 of such codes enables an entire community to meet the technical requirements of Tier-0 access. It is important to collaborate not only with individual users but also at the community level to ensure that codes produced by, and available for, the community can be migrated to Tier-0. Community-level collaboration was present in DEISA, as mentioned above, and is included in PRACE-1IP as part of task 2 in WP7. Thus, in the latter activity a number of user communities have been chosen (e.g. materials science, astrophysics, engineering) and some codes important in these fields have been selected for optimization and enabling to Tier-0 level. For community work it is important to collaborate closely with both the code developers, in order to optimize the application code, and users so as to have representative input for providing test cases and benchmark data. PRACE-2IP's focus on applications will further aid this.

Existing Tier-0 centre perspective

From a Tier-0 centre perspective one of the key issues is allowing European researchers the potential to fully benefit from the resources and the services offered by the Tier-0 level of the PRACE RI. This means that important effort in promotion and training of European scientists by the PRACE RI must be followed up by the PRACE members and not only in a user centric approach but also by enrolling in European structured users communities.

Current activities in PRACE-1IP such as the creation of a PRACE Users Forum, the seasonal schools and the petascaling of communities' codes are addressing these issues and will need to be amplified in PRACE-2IP in order to raise their goals. From a user perspective it will also be important to have a common set of services and policies between the different Tier-0 centres (as well as the soon to be available Tier-1) in PRACE in order to simplify the daily access to the distributed infrastructure. These issues are already benefiting from the results of DEISA and PRACE-PP projects and are currently addressed by PRACE-1IP WP6 and soon by the PRACE-2IP project.

The integration of DEISA into PRACE is a real challenge especially setting up the needed ramp-up mechanisms to allow users to climb, if needed, the HPC pyramid from Tier-1 to Tier-0 resources. This will also allow the enlargement of the Tier-0 user base by gathering new countries, new scientific domains, new users communities (academic and industrial), and strengthening European collaborations in a context of global competitiveness.

Finally beyond the ramp-up consideration this integration will also allow PRACE to more efficiently cover the mix between capacity and capability workloads. As ensemble (or uncertainties studies) will be more and more important in the future these issues will need to be addressed in relation with users communities. As an example some actions have started in the umbrella of PRACE-1IP WP7 like the one with IS-ENES (InfraStructure for the European Network for the Earth System Modelling) for porting the climate code chain into CURIE, the French Tier-0 system.

These relations with users communities are also essential in order to be able to select the most pertinent scientific codes which may benefit from the optimisation services provided by PRACE.

As the software efforts for addressing the mid-term grand challenge of the Exascale are now quite known it is crucial to have with these communities a joint action for identifying the key codes which will be optimised or rewritten with the help of PRACE (or other FP7/8 projects). Provision of such services by the Tier-0 centres means that they will need to lean on their users support teams as well as all the expertise gathered by the others PRACE partners.

2.5 Expanding the ecosystem

As highlighted inter-tier relationships which are *intra-PRACE* relationships established since the PRACE-PP project have evolved under the PRACE AISBL umbrella. Further expanding those relationships may take advantage of work done in deliverables D4.3.1 [https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/d573070/D4.3.1.pdf] and D4.4 [https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/d565633/D4.4.pdf]. These deliverables focused on studying and proposing models to accommodate two important aspects:

- Cross national exchange of resources
- Collaboration frameworks

Additional HPC resource provision can come in the form of:

- Regional centres either in a single country or cross national
- Topical centres that offer dedicated facilities for a specific research domains which may include specialized large scale HPC resources
- Centres located outside the EU that offer access on a global scale

The objective of expanding the ecosystem can also rely on the goals set in the PRACE-2IP concerning the integration of the Tier-1 resource providers in PRACE. That will provide a powerful instrument for inter-tier relationship expansion by means of including new national providers. It is an open question whether this model, once implemented, can prove useful for the inclusion of providers that do not qualify as Adhering Members of PRACE.

One can answer positively only if the resource provision models show similarities to that being developed by PRACE. Therefore actions aimed at establishing new relationships based on resource exchange should take into account this constraint.

Nevertheless, regional and topical centres are good opportunities for expanding the entire ecosystem. To date the only examples that are of interest to PRACE are two on going projects:

- LinkSCEEM-2
- HP-SEE

PRACE has already established relationships with them in the area of joint training and joint events. There is potential to further develop these relationships, extending them to resource exchange although the resource exchange models differ from that of PRACE.

Beside the resource exchange model, the access model implemented by resource providers is a further area of relationship. For example the new DECI[1] call launched by PRACE implements the DEISA model that relies on national peer review systems. The inclusion of new resource providers enhances the possibility developing deeper relationships in this regard by means of creating a permanent forum for discussion and exchange of contributions.

Looking at the provision of resources from a global perspective, there is a growing interest on potential resource exchange with such providers. So far US TeraGrid[2] (and the follow on eXtreme Digital[3] (XD)) provider shows a promising potential to accommodate new relationships that go beyond the on going joint effort on the HPC training school.

The on going work of the HPCWorld[4] consortium will shortly provide guidelines relating to resource allocation. Where adopted as a common framework, these guidelines will ease the cross provision of resources.

A preliminary set of recommendations can be taken as follows:

1. Expanding the ecosystem to include all the present EU countries (as well as those countries due to formally enter the EU during the project lifetime) requires specific actions from the PRACE AISBL to establish official and formal contacts with those countries to begin common work plans, though it is likely that none of them currently have the potential to be a Tier-1 provider, e.g.: Denmark, Slovenia, Malta, Croatia, etc.
2. Expanding the ecosystem to include regional and topical centres requires the PRACE AISBL to promote formal relationships and collaboration plans. Interoperability more than integration might be the key for starting a resource exchange relationship.
3. Expanding the ecosystem to include non-EU providers requires of the PRACE AISBL focused actions to establish relationships at political/strategic level. That should include the setting up of a common, persistent steering committee to address the issues of resource exchange policies, interoperability and access programmes to enable cross exchange at the level of large common research activities that involve EU and overseas communities (climate change research initiatives, etc.)

[1] Distributed Extreme Computing Initiative
[2] https://www.teragrid.org/
[3] https://www.teragrid.org/web/about/xdtransition
[4] http://www.hpcworld.eu/

3 Questionnaire/Survey

3.1.1 *Introduction*

In this section we describe the key findings of a user survey performed as an objective of subtask 7.4.1 in the PRACE-1IP project. The survey was web based. It was promoted by European HPC centres to their own users. As it was anticipated the data would be of value to task 4.2 there was extensive collaboration at the drafting stage between tasks 4.2 and 7.4.1. The survey was implemented as an extensive questionnaire with questions requesting information on the usage of HPC systems. Thus, researchers were asked about:

- The computer architectures and systems they used.
- Software, such as applications and libraries.
- Current experience and expectations of PRACE and Tier-0 architectures.

The questionnaire contained 50 questions (of varying complexity) and received over 300 responses from 7 countries, although it should be pointed out that two thirds of the responses came from just two countries, namely Italy and Germany. Since the Italian researchers who replied to the survey quoted CINECA as their HPC site, then some bias in the data might be present, although this is not likely to affect the findings that are of interest to us in this section (the German researchers instead were spread over four separate sites). Further information on the survey and its results can be found in deliverable D7.4.1 [https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/d533224/D7.4.1.pdf].

3.1.2 *Key Findings*

The survey described above was fairly comprehensive and not all of it is relevant to our discussion. We shall therefore concentrate on the key findings pertinent to the focus of this document, i.e. how the results can be interpreted to help the promotion of Tier-0 usage and in particular study both relationships between the two tiers and the best practices involved in their utilisation. We start by first looking at the current usage of Tier-1 systems, particularly with regard to applications and parallel scalability and we find that:

- A high percentage of users (nearly 50%), are either developers or actively contribute to the code they use.
- Few of the applications used scale up to Tier-0 levels, although there is a general requirement for more scalability.
- Despite the problems in scalability a majority of users are aware of the bottlenecks in their code and believe that they can be solved with low to medium effort.

For Tier-0 promotion these results are quite encouraging, despite the significant architectural differences between the two sets of systems.

After these detailed questions, the survey then queried whether users would be prepared to use PRACE resources and what factors would ease the migration. From the replies obtained we can make the following observations:

- About half the researchers hadn’t considered applying for PRACE, but primarily because they were not aware of the possibility rather than due to technical reasons. However, about 30% of respondents did point out they had sufficient resources.
- For migrating to Tier-0, the availability of Tier-1 machines would appear to be a key enabler, as would the availability of different architectures.

These considerations regard Tier-1 users but we also need to look at the current usage at the Tier-0 level and assess the user experience. Thus, we find that the following features are important:

- long term projects;
- large file system;
- variety of architectures in the infrastructure;
- technical support;
- the possibility of moving the accounting quota from system to system;
- the need to access accounting information, preferably in real-time;
- infrastructure for evaluation;

Long term, persistent disk space would appear to be of critical importance as requests of at least 1 terabyte are common. Naturally, the need for large and non-temporary disk storage has technical and financial consequences, and also implies investment in tools for archival and transfer. We notice also that the survey results did not indicate significant familiarity with data transfer technologies beyond simple UNIX copy commands (e.g. scp), despite the importance given to gridftp, IRODS, etc. in projects such as DEISA. Since data storage must be a feature of any Tier-0 promotion plan, this is another challenge which needs to be addressed.

We emphasise that the issues raised above are the main ones we consider relevant in our present discussion and there are many other findings, such as the possibility of having access to access Wikis and a desire for a fast and transparent review process, for example, which we have not discussed but can be found from the original reference for the survey.

4 French HPC ecosystem case study

4.1 Rational

Section 1 briefly alluded to the rational for the choice of examining the French HPC establishment and its approaches in this case study. In more detail the rational is as follows: France is one of only two European countries currently hosting a Tier-0 system, the other being Germany. France was deemed more representative than Germany because the scale of HPC operations in Germany coupled with its federal system which incurs a complexity that is less relevant to other nations than the more centralised French approach with, for now, a single Tier-0 system. Furthermore elements of Germany's Tier-0 infrastructure, i.e. Jugene and its operation are already well known due to their relatively long history at the core of the PRACE Tier-0 service. HPC services at other levels have existed it both countries for long periods. The French situation is also interesting in that there is some integration of civil and classified HPC activity. This study is complimentary to those in section 5.

4.2 The French organisation

In France, GENCI (Grand Equipement National de Calcul Intensif) is in charge of the coordination of the major French academic equipment in high performance computing. GENCI is a French legal entity taking the form of a "société civile" with five shareholders: French Ministry of Higher Research and Education, CEA, CNRS, French universities represented by the CPU "Conférence des Présidents d'Universités" and INRIA.

Created in January 2007, GENCI has the following missions:

- To define the national strategy and to coordinate the major computer equipment for the French computer centres for civilian research, by providing for their financing and assuming ownership of the equipment;
- To participate in the creation of a European high performance computing ecosystem, GENCI is the French representative in PRACE AISBL, HPC Europa2 and EESI;
- To promote the use of modelling, simulation and high performance computing in both fundamental and industrial research;
- To promote and execute all research required for optimizing the usage of current HPC facilities and anticipating futures technologies;
- To open its national HPC facilities to all interested scientific communities, academic or industrial, national, European or international.

4.3 Tier-1 resources and ramp-up to Tier-0

The French resources are operated by three national Tier-1 centres: CCRT at Bruyères-Le-Châtel (for the civilian activities of CEA), CINES at Montpellier (for the French Universities) and IDRIS at Orsay (for CNRS).

The most powerful Tier-1 facilities include a 267 TFlops SGI system at CINES (called Jade), a 139 TFlops IBM Blue Gene/P system at IDRIS (called Babel), a 103 TFlops + 192 TFlops (single precision) hybrid Bull system at CCRT (called Titane) and a 67 TFlops IBM Power6 system at IDRIS (called Vargas). These resources will be augmented at the end of 2011 by 20% of the cycles of the CURIE system.

At the Tier-1 level, the support teams are providing regular services of training and user support to national users. Three of these Tier-1 systems are considered as ramp up to Tier-0 systems: Babel at IDRIS, Jade at CINES and, in to a lesser extend, Titane at CCRT. On these systems, a special operational procedure is setup in order to ensure to users that they can get access to a large amount of resources for developing and preparing their code for the Tier0 level. These support teams are also working with eligible Tier-0 users by providing them help on PRACE preparatory access type B and C calls[1].

At each installation of a new Tier-1 system and during its acceptance procedure, a special campaign of “Grands Challenges” is organised by GENCI. The latest campaign has been organised on the SGI Jade 267 TFlops system with 8 different scientific projects and open to PRACE partners with contribution of users coming from CINECA and ICHEC. This “Grands Challenges” campaigns allow few selected users to run full scale applications on the whole configuration of the machine, fostering scientific advances and pushing the limits of the architecture in a mode close to the production mode.

Another ramp up for Tier0 systems, or at least large Tier-1 simulations, was the involvement of IDRIS into the DEISA project thought the DECI yearly calls for proposals. Around 10% of the capacity of Babel (IBM BG/P 139 TFlops) has been allocated and shared by DEISA resource exchanges mechanisms to European researchers.

GENCI’s facilities are available to French scientists through a Peer Review process based on scientific excellence. Calls for proposals are open twice a year and application to resources of the three national centres and nine different systems can be done through an online unified portal (called DARI). All proposals are reviewed by an independent scientific college of

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

recognised scientists. Based on the results of this evaluation, resources are granted by GENCI following the amount of resources requested according to the availability and the recommendations of the scientific college. In 2011 more than 600 research projects from a wide panel of scientific domains have been granted by GENCI resources.

In addition, GENCI acts as the French representative into the PRACE AISBL and is one of the four hosting partners together with Germany, Italy and Spain. GENCI recently ran a large procurement resulting in the purchase of a Bull system, a 1.6 Petaflop/s x86-64 Tier-0, which will be delivered in two phases starting end of 2010 to the new computing centre (called TGCC) at Bruyères-Le-Châtel (in the south of Paris). The team which operates CURIE come from the CEA's military applications division (CEA/DAM) who also operate the Tera100 petascale system for designing and simulating nuclear weapons. This allows synergies and technological transfers from the defence field to the academic one by using expertise and open source software developed by CEA/DAM into CURIE. On the other side, the multipurpose usage of CURIE allows to extend widely the functionalities of such system software tools developed by CEA.

4.4 Strengthening of the regional level

In 2011, GENCI launched a new initiative called Equip@meso with nine regional centres in order to strengthen the French Tier-2 layer. GENCI is not in charge of the strategy of equipment of the regional centres (it's the role of the French universities or other regional entities) but GENCI need to rely on such regional centres for getting a first access to HPC resources to new users, provide training services and allow special usages that are not possible on production systems located on the Tier-1 or the Tier-0 levels.

This initiative coordinated by GENCI with a 10M€ budget will allow these centres to:

- Update their existing facilities with new HPC systems. These systems will be chosen in order to ensure a good coverage of the existing HPC architectures.
- Deploy standards services to users like training and user support services.
- Implement, in some cases, Peer Review procedures.
- Establish strong relations with Tier-1 and between Tier-2 centres.
- Amplify, at the regional level, the HPC-PME national initiative (see below).
- Work on scientific dissemination across all the regional centres with the help of CERFACS (see section 4.8) and Maison de la Simulation.

This Equip@meso initiative is vital for enlarging the basis of the HPC pyramid in France by attracting new users' communities and providing them coherent services to climb the ladder from Tier-2 to Tier-0.

4.5 Key actors in HPC field

At the national level, dissemination of HPC usage, best practises, training and key scientific or industrial achievements is performed by multiples entities:

- GENCI regularly organises seminars targeted at future Tier-0 users, providing them with good practice, advice and feedback from Tier-0 or large scale Tier-1 users. The latest meeting has been held in Paris in May 19th with 80 attendees. Special initiatives have been launched by GENCI and CEA to help future French scientists in applying to PRACE calls, a mailing list called appels-prace@genci.fr is available to provide guidelines to futures applicants.

- GENCI is partner in the BULL Joseph Fourier award which annually grants €15,000 and CPU (or GPU) hours to up to 3 French scientists for their key scientific achievements using HPC.
- GENCI is also partner of the C3I (Certificat de Compétences en Calcul Intensif) which mandates a multidisciplinary jury to give an endorsement (as a proof of expertise) to recent postdocs or PhD students who developed strong skills in HPC.
- Groupe Calcul is an active group of experts who organise technological exchanges, seminars, trainings and HPC dissemination through dedicated mailing lists.
- ORAP (Organisation de Recherche Associative en Parallélisme) is a 20 year old association formed by EDF, CEA, INRIA and CNRS to promote HPC by organising twice a year very popular seminars.
- TER@TEC (http://www.teratec.eu/gb/index.html) is an initiative of 70 stakeholders from industry and academia spread along the whole value chain of HPC (from technology providers, ISVs and research labs to end users). TER@TEC association, with the cooperation of local councils, is setting up a European technology park dedicated to high-performance simulation and computing, at Bruyères-le-Châtel, next to CEA and GENCI supercomputing facilities.

4.6 Conditions of access to French industrial users

French industrial users are eligible to use GENCI's high performance computing facilities mainly in the two following cases:

- For open R&D projects: The industrial user must cooperate with an academic partner into a joint research project. The academic partner is the PI (Principal Investigator) of the proposal, which is peer-reviewed using scientific excellence criteria. If the project is selected, the partners of the proposal must commit on publishing the results of the project at the end of the grant period (one year). This mechanism helps to foster technological transfers between academia and industry.
- For a "one shoot" experiment: In order to help industrial users (large companies as well as SMEs) to assess the potential of HPC in their daily business, industrial users may access once in specific conditions and in a limited time period to GENCI's HPC facilities.

4.6.1 *Focus on the "HPC-PME Initiative"*

This "one shot" experiment condition has been used by GENCI, INRIA and OSEO (a public French bank for developing innovation in industry and especially SMEs) for setting up a joint initiative called "HPC-PME".

Created at the end of 2010 and currently supported by five of the French biggest technological clusters (Aerospace Valley for aeronautics/aerospace, Axelera for Chemicals/Materials, Minalogic for microelectronics, SYSTEM@TIC for embedded systems) and CAP DIGITAL (for digital media) this initiative aims to allow French SMEs to assess the potential of using HPC in their innovation processes. This initiative is based on an integrated offer providing services like training/best practices, expertise from public research (in various scientific domains as well as in HPC), access to HPC systems (the ones from GENCI as well as regional centres) and funding facilities.

HPC-PME is accessible through a web portal (http://www.initiative-hpc-pme.org/). SMEs can apply online to the initiative using a permanent call for proposals and submitted proposals are reviewed monthly by representatives from GENCI, INRIA and OSEO.

Two specific categories of SMEs are targeted by this initiative:

- SMEs that discovered recently numerical simulation and need to be strongly convinced of the potential of HPC.
- SMEs already using HPC on desktops, workstations or enterprise servers and which want to bridge a new industrial gap by using larger HPC facilities.

HPC-PME is a very upstream initiative and is not relevant to SMEs, which are already using production HPC clusters and are searching for On Demand Computing remote services.

One of the most important missions of HPC-PME is to foster technological transfer between expertise from public research and SMEs and the most valuable part consists in providing public research expertise and support for implementing new industrial methodologies using HPC.

During its process into HPC-PME, the SME may choose to use for experimenting, sizing their dataset, and implementing new methodologies either GENCI's Tier-1 facilities or regional facilities provided by Tier-2 centres partners of Equip@meso.

During this process, the experts dedicated to help the SME are also assessing future regular HPC usage by the SME after the end of their HPC-PME project. This regular usage may occur in two ways: either the SME buys and uses its own HPC facilities or uses a remote HPC infrastructure on which production usage is possible (which is not the case for GENCI's facilities).

HPC-PME is an open initiative, which aims to be extended sooner to other technological clusters and to existing similar local initiatives. Since the beginning of this initiative more than 10 SMEs have joined HPC-PME and the first entrants will finish the programme during the summer of 2011. The SMEs come from various locations in France, highlighting the need for a national initiative. They are from several industrial domains e.g. automotive, aerospace, medical, microelectronics, maritime, digital media, etc. and they also represent various kinds of SME i.e. subsidiaries of big companies, very small SMEs, Independent Software Vendors (ISVs) etc. This initiative may be strengthened in the future at the European level by interest from other European partners and could potentially form part of future PRACE Implementation Phase projects.

4.7 CCRT

CCRT is a large supercomputing centre operated by CEA and used jointly as Tier-1 in France by French research organizations and also by industrial companies. In order to address this last issue, a specific business model has been developed for industrial partnerships with CCRT (see below).

CCRT has been set up in 2003 by CEA to provide High Performance Computing resources for large scientific computations and to foster a real synergy between academia, other research organizations and industry by promoting exchanges and scientific collaborations between all partners.

The current CCRT industrial partners are EDF (French electricity provider), SAFRAN (Aeronautics Company), EADS (space industrial), Onera (French aeronautics agency), Areva (global nuclear industry leader) and Ineris (industrial risk assessment).

4.7.1 *CCRT business model for industrial partners*

Each industrial partner signs a partnership's contract with 3-4 years duration. This contract is renewable. Such a duration is necessary to set up a real collaboration between industrial users and the computing centre team on one hand and other academic users in the same scientific field on the other hand. It is especially important to improve usage of large simulations and to share experiment outcomes, including experience of open source or academic software. It also allows a fruitful balance and interaction between "regular production" usage and more experimental and research usage.

Each partner holds a share according to its contractual financial participation. The contribution is based on Total Cost of Ownership: exploitation costs and investment costs are taken into account as a whole and put in balance with the overall partner's needs.

4.7.2 *Current CCRT systems*

CCRT spans a diversity of supercomputer architectures to fulfil the various needs of the partners. The current systems available at CCRT for both Tier1 and dedicated industrial usage are:

- A hybrid computer, called Titane, installed by BULL in 2009, with 1068 Intel/Nehalem compute nodes (~100 Tflops) and 48 Nvidia/Tesla servers (200 Tflops SP) for an academic usage as French academic Tier1, upgraded in 2010 by 40 Tflops of Intel Nehalem based Bullx nodes for industrial usage.
- A Bull Novascale computer based on Intel/Itanium nodes for 48 Tflops.
- A vector system, provided by NEC with 8 SX8-R nodes and 3 SX9 nodes.

A post-processing cluster based on 38 Intel/Xeon nodes (8 cores) with 64 and 128 GB of memory and NVIDIA FX5800 graphic cards delivers remote visualization and post-processing services for users' data locally stored at CCRT. The storage system can host 1 PB of data on disks, provided by SGI, and 5 PB in SUN-SL 8500 robotics for longer term archiving.

4.8 CERFACS

CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique) is a research organization that aims to develop advanced methods for the numerical simulation and the algorithmic solution of large scientific and technological problems of interest for research as well as industry and that requires access to the most powerful computers presently available. CERFACS has 7 shareholders: CNES, the French Space Agency; EADS France, European Aeronautic and Defence Space Company; EDF, Electricité de France; Météo-France, the French meteorological service; ONERA, the French Aerospace Lab; SAFRAN, an international high-technology group and TOTAL, a multinational energy company.

CERFACS hosts interdisciplinary teams, both for research and advanced training that are comprised of: physicists, applied mathematicians, numerical analysts, and software engineers. Approximately 115 people work at CERFACS, including more than 95 researchers and engineers, coming from 10 different countries. They work on specific projects in 9 main research areas: parallel algorithms, code coupling, aerodynamics, gas turbines, combustion, climate, environmental impact, data assimilation, and electromagnetism & acoustics.

CERFACS has its own HPC facilities (approximately 20 TFlops) for code development and internal studies but also access to various HPC facilities from the French Tier-1 centres of GENCI to the PRACE and the DoE INCITE Tier-0 systems through regular calls.

5 Case studies of HPC ecosystems and practices in smaller European nations

5.1 Rational

In this Section we present the salient details of the HPC ecosystems that exist in the two smaller nations under examination, namely the Republic of Ireland and Norway. Of particular importance is the highlighting of strategies and policies that have proved valuable, which for various reasons that will be elaborated upon, would not be applicable or appropriate in the context of larger nations. Where lessons have been learnt that may apply to other nations they will be highlighted. Baseline conditions will be described detailing the size of the academic and industrial user communities. Taken together these countries are considered representative for the following reasons. They contrast in several interesting ways while still providing a HPC service with similar aims. The Irish HPC model is more centralised than that of Norway whilst Norway has a closer relationship between its NREN and HPC provision. The Irish service operates in an environment that is highly cost constrained by virtue of national economic issues. Both HPC services are relatively young by European standards and so have recently faced the challenges seen by emerging countries. Neither country has ambitions to host a Tier-0 system. Along with Poland and Bulgaria Ireland is one of the countries involved the PRACE DECI call for the first time this year. This section taken together with the French case study is believed to be a representative sample, covering both strategic and operational matters in significant detail thus allowing a third party to make a comparison with the situation in another country and or centre with which they are familiar.

5.2 Case study – The Irish HPC ecosystem

5.2.1 *Baseline academic and industrial conditions: the Irish ecosystem*

Industrial sector

The use of HPC in Ireland's industrial sector is limited though has significant potential for development. Pharmaceuticals and biomedical devices, Internet search, financial services and semiconductors, are well represented in Ireland. As yet the numbers employed in industry that have had the opportunity use HPC techniques on the national service at a postgraduate level is limited.

Academic sector

See Annex 8 for further details of their size and remit of 3rd level sector. The HPC environment has the following characteristics (bottom up):

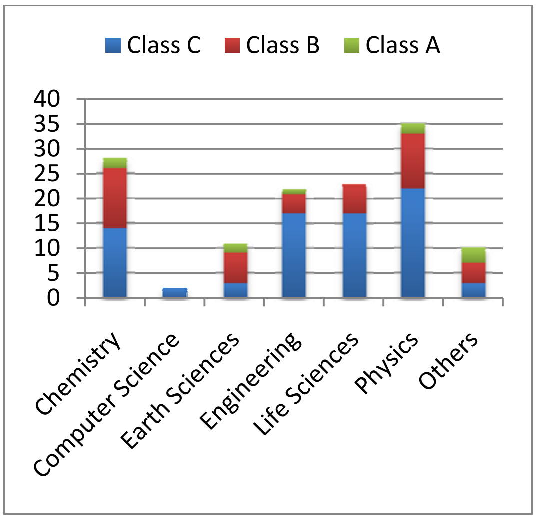
- A healthy Tier-2 layer, probably over-developed in relation to Tier-1;
- The presence of a so-called "Tier-1½", corresponding to resources from a Tier-1 computer (ICHEC operated National cluster "Stokes") utilised in a Tier-2 manner - the so-called Condominium cluster model.
- The presence of a single Tier-1 cluster, operated by ICHEC as a centralised compute resource (its performance is at the lower limit of what is considered Tier-1). Access to more powerful Tier-1 resources is provided through the DECI programme.

- The absence of Tier 0 resources with the exception of programmes such as PRACE.

Large European Nation		Ireland
Facilities provision	**Tier**	**Facilities provision**
Domestic PRACE centre(s)	0	International PRACE centres
National & regional centres, Grid collaborations	1	ICHEC & DECI
Local centres	2	Higher education institutes

Table 3: Per tier facility provision in large European nation versus Ireland

Figure 1 shows a break down of the current distribution of projects using ICHEC's systems (see also Section 5.2.4).



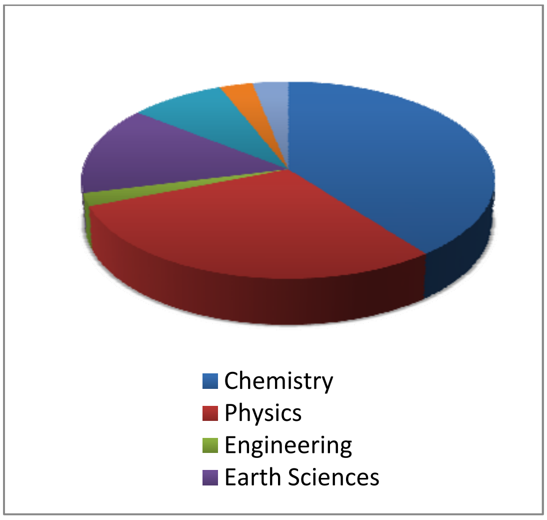


Figure 1: (left) Distribution of 131 active projects
Figure 2: (right) Utilisation by discipline in 2010

5.2.2 *National service ecosystem development goals*

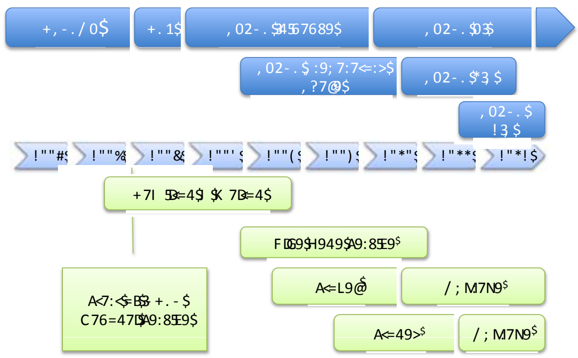
ICHEC was established as the national HPC provider in late 2005. Its policies are centred on research enablement, partnership with the user community and technology leadership e.g. the development of an emerging GPU national service.

General purpose Tier-1 service provision

Tier-1 provision is severely restricted by budgetary constraints, even before the advent of the current economic crisis. It became clear that ICHEC was becoming victim of its own success, with challenges of capacity. In response to this ICHEC has chosen not to orientate system procurement around maximising the absolute number of cores per Euro at the expense of potentially decreased system reliability or at the expense of reducing the human expertise component of the service. The current strategy is as follows:

- Maximise resources by adopting a 5 year life cycle with mid-term upgrade, providing better value for money than a simpler 3 year cycle.
- Avail of complementary Tier-1 resources through international access programmes e.g., DEISA DECI: three Irish successes at the last call.
- Avail of complementary Tier-1 resources through peer-to-peer agreement e.g., Grand Challenge demonstrator at CINES, etc.

- Early adoption of promising emerging technologies such as GPGPUs, which provides performance levels (per €) well beyond that which "conventional" HPC systems can deliver.
- Intensive training and partnership programme with the research community to ensure that researchers make an effective use of the infrastructure.



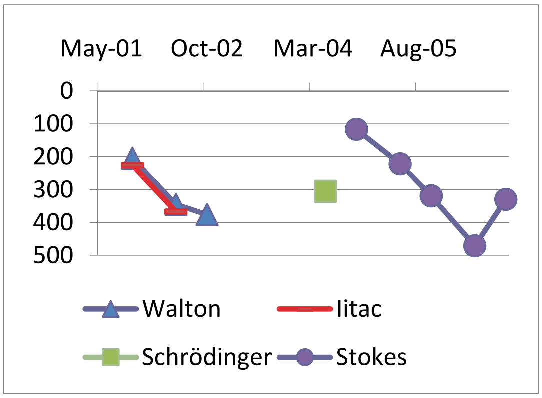


Figure 3 (left): Systems provision relative to PRACE (specifications detailed in annex 8.1)
Figure 4 (right): Irish systems in the Top500 (specifications detailed in annex 8.1)

Capability computing provision

The initiative to establish a Capability Computing service in 2008 was led by Prof Luke Drury of the Dublin Institute for Advanced Studies to address demand expressed through two user community surveys. These systems, a Blue Gene /L and /P, were operated by ICHEC. It was clear that systems of a limited size in a capability mode could only support a limited number of research projects, no more than 5 to 10 at any given time and were unsuitable for development without severely curtailing capability. Insufficient performance for the more demanding users was mitigated by the inclusion in the procurement of access to systems hosted in IBM's Rochester and Watson labs.

Integration with the existing ICHEC service was essential as was the necessity to engage in dissemination and outreach activity including co-hosted technical workshops with IBM's Dublin based HPC team. Two positions were funded to support the systems. The technical demands involved quickly showed this level of comprehensive support was crucial despite the small number of eligible users.

ICHEC does not foresee their replacement in the near to medium future. In conclusion while the Irish Blue Gene systems remained highly relevant up until the recent implementation phase of PRACE it is evident that the strength of the platform as deployed was as a stepping-stone to larger systems. The readiness of the user community to address large systems is now evident from the number of per capita and technical calibre of Irish applications made to PRACE, including a successful production application. Much as with Blue Gene class scalability, once demonstrated this avenue becomes increasingly interesting and relevant to the community.

ICHEC an official CUDA Research Centre

Ireland cannot realistically hope to have a leading role in terms of the scale of its systems or the size of its user base. It has decided to concentrate effort in niche areas where specialist expertise can make a significant impact. GPGPUs are one such area; ICHEC was designated as a CUDA Research Centre (CRC) by NVIDIA in June. A national GPGPU service is scheduled to open in Q3-2011. The intention is to provide users with assistance in migrating to large GPGPU enabled systems at Tier-1 and or Tier-0 and through WP7 and other work to enable the migration of community codes e.g. Quantum Espresso.

5.2.3 *Impact of external HPC access*

The role of Blue Gene systems

A significant benefit, as part of the contract with IBM, was the granting of access to larger (up to 4 cabinets) IBM hosted Blue Genes. The true value of this measure became increasingly evident later in the life of the systems as more users developed their codes sufficiently to benefit from the larger IBM systems located in the US. Ultimately the facility was used by virtually all of the Blue Gene users to access the larger IBM resources. In two cases access to the IBM hosted systems has acted as a useful intermediary stepping stone to JUGENE.

The role of HECToR

In 2009, ICHEC recognised the necessity to provide access to a subset of users to a system able to provide an alternative Tier-1 resources for users whose applications could not be ported or scale on the Blue Gene architecture. A bilateral agreement with EPCC to provide one million allocation units[1] of compute time on the HECToR system. Users had the opportunity to demonstrate if they could scale their applications or workflows beyond the limits imposed by the smaller size of ICHEC's system and to perform worthwhile science and ICHEC's computation scientist team gained valuable experience.

The role of DEISA

While ICHEC was aware of the value of DEISA it was felt that the priority was to focus on developing domestic capabilities to a point whereby the community would be in a position to make successful applications. In 2010 considerable effort was dedicated to promotion of the DECI-6 call including a workshop on application preparation, followed up by joint applications in several cases. The raised awareness of European programmes, particularly PRACE, amongst the user community coupled with a growing demand for resources contributed to the effort. Six applications were made three of which were successful (see Table 8). This strong result was an endorsement of the strategy in a heavily oversubscribed call.

The role of PRACE

To become an effective resource to domestic researchers PRACE access has to be seen as "pseudo-integrated" into our portfolio of access schemes, rather than as a stand-alone, disjointed programme. In other words, migration from the various tiers should be made as seamless as possible for end users. PRACE provides an excellent opportunity to address the Tier-0 deficit at a national level particularly for small countries where acquiring both sufficient general purpose and capability computing infrastructure is extremely difficult, without clearly favouring one at the expense of the other either. In light of this ICHEC has decided to prioritise the provision of general-purpose infrastructure domestically. Interest in both preparatory and production access is strong. See Table 8 for a list PRACE awards.

The role of United States HPC access programmes

ICHEC has actively promoted the Argonne Early Science Program and the annual INCITE[2] calls. ICHEC disseminates the calls through its website, community wide mailings and

[1] The Allocation Unit (AU) is a unit of computational work, equivalent to a 1 GFlop/s processor running for 1 hour, as assessed by the Rmax Linpack benchmark figure.
[2] http://www.doeleadershipcomputing.org/

publications, assistance is offered in the drafting of proposals and there is a stated openness to formal collaboration particularly where the addition of HPC expertise will augment the skills base required for the proposed work. Relative to PRACE calls the infrequency of US calls such as these which are both attractive and open to European researchers means that the turn around time required to produce tangible results is longer. However an application by an Irish researcher to the Argonne Early Science Program was very nearly successful and resulted in the award of discretionary time, see section 5.2.5. In 2010 two INCITE applications were made and the same or more is expected to be the case in 2011. It is worth noting that PRACE preparatory access results played a role strengthening the cases of two of the three US programme applications in 2010.

5.2.4 *Policies and procedures for user support and upward migration*

Research enablement

Helping researchers to tackle their challenges in the computational sphere is key part of ICHEC's mission. As a national facility, ICHEC has to cater for the needs of a diverse community of users and requirements. The procedure for access to the national compute clusters takes this diversity into account by providing three possible routes to apply for resources, ensuring the best compromise between a fast response time, fairness of the application process, and an efficient usage of the resources while complying with the strategic objectives of the national centre. The following approaches are taken on a routine basis:

- Training activities through the provision of a comprehensive range of training material and courses in various locations.
- Outreach activities in which ICHEC staff visit research groups.
- Consortium support activities in which specialist staff engage in a collaborative way, assisting groups with development, porting and optimisation of their parallel applications and workflows and in some cases become de facto members of the team. ICHEC currently has in its ranks 11 computational scientists supporting 131 projects and 249 users.
- Helpdesk support: to end-2010, a total of 2,700 queries have been logged and dealt with by ICHEC staff normally within 24 hours.

Having computational scientists "specialised" in certain research areas allows for the provision of in-depth support and has a number of benefits:

- A common terminology and understanding of the scientific requirements. Importantly this promotes optimisation and code development at algorithmic level, often a strict requirement in the move to Tier-0.
- Long-term relationships can be established through a consistent point of contact. Opportunities for joint scientific publications and progress are more likely to emerge.
- The specialised can act as a go-between to the rest of the ICHEC team and peers in other European centres e.g. Tier-0 centres.

Work with project holders drives migration activity from Tier-2 to Tier-1 and is in most cases mindful of the technical requirements or impediments that might later hinder the migration to Tier-0. In a proportion of cases upward migration is limited by external factors. For example is a group makes use of a 3rd party code for which source code is unavailable. In such cases involvement is more likely to investigation of alternative codes or techniques that might offer more opportunity for scaling e.g. checking if a PRACE supported applications can be used.

A corollary to this has been ICHEC's strategy of promoting key community applications for GPGPU enablement and 'petascaling' effort within PRACE. For example the Quantum Espresso package, a versatile package with an active community, is now GPGPU enabled to an extent, under the auspices of WP7 work undertaken by amongst others ICHEC staff. By default the work done here broadens the options of the domestic users of the package and strengthens the technical case they can present in subsequent applications.

Incremental improvement and building technical readiness in the community has been well validated by the success of Irish researchers in pursuing access through PRACE. At the technical evaluation stage a near 100% pass rate has been achieved across both preparatory and production applications and indicates that involvement with users prior to the submission of applications is beneficial.

There is reluctance by time pressed researchers to invest in the completion of applications for HPC access with no guarantee of success, be it domestic or international. There is a tendency and arguably a need for these applications to be technical and jargon heavy. Especially when moving between tiers there can be a change in emphasis and a demand for unfamiliar technical detail. This constitutes a relatively surmountable barrier to migration. It has long been ICHEC policy to actively assist potential applicants with the interpretation and completion of application forms. The role can involve simple proof reading, explaining or cross check technical details or acting as an intermediary with the awarding body to seek clarification of specific points. The computation skills of ICHEC staff and the domain specific scientific knowledge of the applicant are clearly complimentary in this regard.

5.2.5 *User migration example*

Dr. Turlough Downes of DCU/DIAS is an excellent example of an Irish researcher whose work has migrated through tiers. From the outset the code developed by Dr. Downes and his collaborators has been developed and run on desktop class systems. Under the auspices of the Cosmogrid project, a precursor to the present condominium model (see annex 8.1), several small-scale projects were undertaken by Dr. Stephen O'Sullivan a collaborator of Dr. Downes. These led in turn to a Class A and two Class B applications via the standard access mechanism (see annex 8.1). In cognisance of trends in system architecture the algorithm at the heart of the application was designed with scalability in mind. The objectives of these projects were met and the technical progress made throughout meant that scalability did not become an issue up to the scale of the single cabinet of Blue Gene/P available in Ireland.

Keen to tackle problems that could **only be solved by simulation at resolutions that require Tier-0 access** the research group enthusiastically engaged with ICHEC and pursued access to a number of systems outside Ireland i.e. Blue Gene in IBM's Rochester and Watson labs, PRACE prototypes (Jugene & Louhi), DECI 6 access and PRACE production access to Jugene.

The group have been able to demonstrate the dramatic result of near 70% linear scaling across the entire 294k core Jugene system. The group's tenacity is also noteworthy; in total three applications for PRACE production time have been made each refining the proposed science. The first two were both highly rated both technically and scientifically but failed at the scientific prioritisation stage. The third application was successful. **This level of commitment to a programme prior to securing access would not be typical at tiers 1 and 2**. The group has also pursued US programmes. While this has yet to result in a large-scale production award the promise shown in the applications has resulted in preliminary access being granted to perform porting and benchmarking on a discretionary basis.

An example of the support role played by ICHEC in this work was the provision of assistance in the configuration and placement of the largest jobs so as to best exploit the torus network to be found in Blue Gene systems. The lengthy development programme associated with the work outlined above has allowed the researchers to develop a high level of technical self-sufficiency. **Latterly the value of ICHEC's role has been in advising on developments in the access programmes** themselves which can be somewhat remote from the working scientist. Thus this group has exploited computing resources from the desktop through to Tier-0 and has tailored the problems they wish to address to the features of the programmes in question.

5.3 Case study – Norwegian HPC ecosystem

The present Norwegian national HPC infrastructure is implemented by the Notur II project. The consortium includes four university partners and one coordinating party:

- University of Oslo (UiO)
- University of Bergen (UiB)
- University of Tromsø (UiT)
- Norwegian University of Science and Technology (NTNU, Trondheim)
- UNINETT Sigma (coordinator, Trondheim)

Each of the four university partners host a HPC facility and carry out the necessary operational and user support tasks. UNINETT Sigma has entered into a 10-year agreement with the Research Council of Norway (2005-2014). The business between UNINETT Sigma and the universities is defined by a (10-year) Consortium Agreement and Service Level Agreements that define the obligations that the parties have towards each other.

5.3.1 *The Norwegian ecosystem*

Academic sector

The investment strategy for HPC systems in the national infrastructure can be summarized as follows: maintain a small set of systems that is able to process the variety of computational workloads that exist in the academic sector. In practice, this means that the systems have to be complementary in architecture. Presently, the national infrastructure includes two compute clusters (at UiO and UiT) that are used by a large number of communities, but in particular computational chemistry and life sciences. There are also two systems for latency-bound problems (at UiB and NTNU) that are used for climate and earth sciences, computational fluid dynamics and computational physics.

	UiB	**UiO**	**UiT**	**NTNU**
System	Cray XT4	Sun cluster	HP cluster	IBM p575
Number of nodes	1388	304	704	186
Number of cores	5552	2528	5632	2976
Interconnect	SeaStar2	InfiniBand	InfiniBand	HPS
Tflops/s (peak)	51	25	60	21

Table 4: Current Norwegian HPC systems

The systems in the HPC infrastructure are general-purpose systems and are open to all individuals and research groups at the Norwegian universities, university colleges and research institutes. Access is by application. Applications are reviewed by a Resource Allocation Committee, appointed by the Research Council of Norway, whose mandate and working rules are similar to the Access Committee of the PRACE RI. In essence, all research projects with public funding are eligible to apply. All proposals also undergo a technical assessment that is carried out by support staff of the university partners.

The Norwegian Meteorological Institute (met.no) collaborates with NTNU. The IBM p575 is used a number of times per day for running the operational forecast models. This collaboration will continue in the coming years.

The total capacity that is available to the Research Council of Norway is ca. 80 million core hours per year. The four universities together have ca. 70 million hours (for local usage), while the Meteorological institute has access to ca. 7 million core hours per year.

It is planned that all four systems will be replaced in the period 4Q 2011 - 1Q 2013 by (at most) four systems of complementary architecture. The total computational capacity is expected to increase by a factor 5-10.

The base funding from the Research Council for the national HPC is presently ca. 2 MEuro per year. This is largely used for hardware investments and staffing for operations and user support. This is complemented by contributions from the university partners, in the form of in-kind (staffing, local infrastructure, electricity and cooling) and cash contributions (for hardware). This funding is complemented by one-time grants from the Research Council. E.g., in 2007, the infrastructure received a 10 million Euro grant for investments in new systems. A new one-time grant is expected to be made later in 2011 for the new investments that presently are being planned.

Norway participated in the creation of PRACE from the start in 2007 and became member of the PRACE RI in August 2009. Norway has not been a member of DEISA. NorGrid is the National Grid Initiative (NGI) that forms the Norwegian node in the European Grid Initiative.

Application support

The national infrastructure includes a number of FTEs for so-called advanced user support. Advanced user support aims at helping scientists to improve the performance and extend the capabilities of their applications. This can be in a number of ways, for example by code parallelization, porting, optimization, benchmarking, improving user-interfaces and software development. A main aim is also to achieve a better utilization of the (expensive) facilities in the national infrastructure.

A typical project for advanced support can provide:

- Support for applications and databases that are of importance for a number of years.
- Support for (existing) users with large needs for computation and storage.
- Support for (new) users that are not familiar with the supercomputing facilities.
- Support for the coordination of related applications and databases.
- Support for complex application enabling.

Advanced user support is considered to be a resource (similar to computing time). Users can apply for advanced support by submitting a proposal. Applications are evaluated by the Resource Allocation Committee that also decides on computing time allocations. The advanced support activity can also be used for scientists that need assistance in analysing the run-time behaviour of their applications and improving scalability to make the software suitable for PRACE Tier-1 and Tier-0 systems.

Industrial sector

The use of the national HPC infrastructure by the industrial sector is limited. The major HPC systems outside the national infrastructure are used in the petroleum sector, in particular by Statoil. In 2011, Statoil has installed HPC equipment with a total of 220 Tflops/s (peak performance). Companies that provide services for the petroleum sector (e.g., to process marine electromagnetic data) also operate their own HPC equipment. Examples of such companies are EMGS and Fugro that help their customers to increase their exploration success through modelling, integrating and interpreting these data.

The role of PRACE

Similar to the Irish case, PRACE access in Norway has to be seen as integrated into the portfolio of access schemes, rather than as a separate stand-alone programme. Migration from the national systems to PRACE Tier-1 and Tier-0 systems should be made as seamless as possible. Besides the application support provided by NTNU and UiO in tasks 7.1 and 7.2 of PRACE-1IP, the advanced user support activity of the national infrastructure can also be used for analysing and improving scalability of applications to make them suitable for PRACE Tier-1 and Tier-0 systems.

Norway has the ambition to provide a Tier-1 service in PRACE starting in the first half of 2012. This is somewhat later than other member countries and is due in part by constraints in funding in 2011.

The possibility of preparatory access and the calls for regular access to PRACE systems, as well as access to systems provided through HPC-Europa2[1] and INCITE, are communicated to the PIs that have access to the national systems on a regular basis.

5.4 Case study conclusions

While there are differences between the Irish and Norwegian model with respect to system operation, centralised versus distributed respectively, and levels of funding there are also many similarities. Neither national service provides or intends to provide a system that could be viewed as Tier-0. This recognises the costs involved would be simply too great for nations of circa 4M people. In the Irish case the investment in modest Blue Gene systems was not sustainable relative to the costs of general-purpose systems. However as this investment was made before the availability of PRACE infrastructure it played an important role in readying both the user community and the national service for the advent of access to PRACE hardware. Thus the role of PRACE in providing a further tier for researchers to aspire to if needed is clear.

Both smaller countries operate a centralised science case based application procedure for access to cycles and both state the importance of integrating the national process with that of PRACE in as seamless as fashion as possible. A major differentiating factor between Tier-0 and Tier-1 is the level of competition seen at the scientific peer review stage. Thus a robust peer review process at the Tier-1 level preferable with an international element, in addition to as ensuring good science at that level prepares users for that which is will be seen at Tier-0. The success rate for Tier-0 applications is typically in region of 20% which is significantly lower than would typically be the case at Tier-1 level. Common national service and PRACE application forms and or reviewing are not proposed but rather a common approach to user support and policy.

[1] http://www.hpc-europa.org/

Requests for assistance in preparing for and undertaking work on PRACE systems is in both cases seen as a very valid use of support personnel time. The vast majority of the awards listed in Table 8 Annex 8.1 would have adopted this model demonstrating its effectiveness.

The intention of both countries to provide PRACE Tier-1 access on national service systems in late 2011 and early 2012 will further the notion of integration and lower the barrier of entry to the user intent on migrating through the tiers.

Since the creation of GENCI one of the main goals both in France and in Europe has been to implement the so-called Branscomb pyramid with a seamless integration of multiple levels of resources. This approach will allow a researcher to use in a coherent and persistent way the different levels of resources according to their needs.

In 2011 GENCI started to work on strengthening the Tier-2 level (regional centres and universities) in order to increase training efforts and to increase the number of potential users.

Tier-2 centres very often represent the first step of learning HPC, enrolling new users which will benefit from locality, allow technology transfer between academia and industry and development and parallelizing of scientific codes. The Equip@meso will foster also inter-tier relations with Tier-1 and consequently in the future with Tier-0 centres.

This virtuous circle is the key issue that PRACE needs to address collectively in the future.

6 Best practises for user and community migration

Drawing from previous sections of the document it is possible to identify a number of key challenges and resulting approaches that have proved fruitful. The availability of Tier-0 resources in Europe via an open call is a new phenomenon for most researchers. Migration to Tier-0 is accompanied by specific challenges:

- The scientific case must be compelling, the notional value of resources required is so significant and access so competitive that only high-impact research will be given access.
- Services are tailored for so-called “Grand Challenges”, with requirements approaching the capability of the hardware.
- A high level of “technical readiness” is needed to achieve sufficient levels of scalability required to exploit the capability offered by Tier-0 systems.

The following recommendations and conclusions seek to address these challenges and outline what might be considered best practise in so doing.

6.1 Scientific case

Users must recognise that Tier-0 represents not only a shift in what can be achieved relative to before but also a commensurate shift in competition and expectation of results. To date Tier-0 production class projects represent a dramatic increase in compute power and decrease in the number of awarded projects relative to the DESIA DECI calls. This can be clearly seen in Figure 5 which shows all PRACE production awards to date (3 calls) and the DECI-6 awards normalised to Blue Gene/P core hours. The average DECI-6 award was equivalent to 5M Blue Gene/P hours, for the PRACE Production calls the figure is 35M.

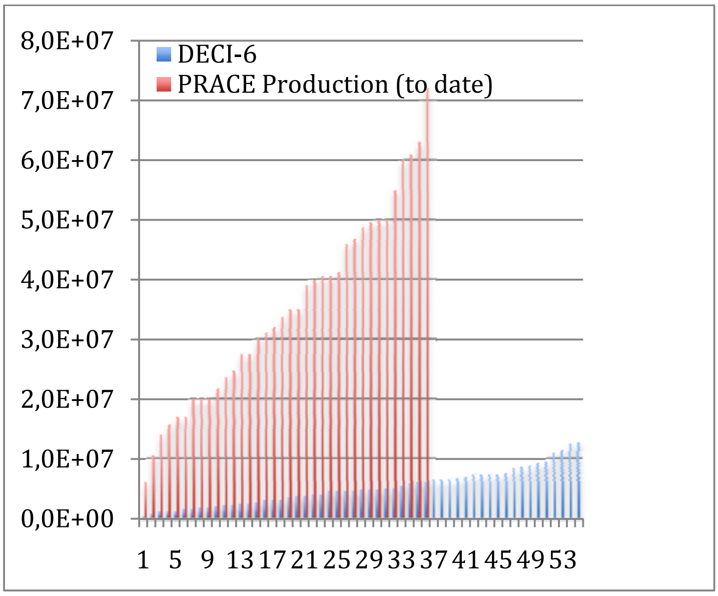


Figure 5: DECI-6 and PRACE production awards, normalised Blue Gene/P hours vs. number of awards

The unambiguous European remit of the infrastructure lends itself to transnational research groups or consortia thus enhancing the opportunity bring together the necessary “brainpower” to analyse simulation data and publish rapidly, a necessity when operating in such a highly competitive environment. It is an open question what role can be played in facilitating this process. This question is perhaps best considered by task 4.1 and or the nascent Users Council it is developing.

The applicant must give careful consideration to the review procedure especially at Tier-0 level. Section 6.2 provides recommendations on addressing the technical aspect of the review, however the challenge of defining and articulating the science case must by definition rest primarily with the applicant. National centres can assist here by aligning their application procedures to be complimentary to those of PRACE. The pending results of the HPCWorld project, see section 2.5, will be useful in this regard.

While the role of national centres in the science case is naturally limited in most cases as this is very much in the remit of the researcher the case studies have shown that there is a benefit to be had by centres engaging with researchers at the application stage to ensure the technical implementation is both feasible and well articulated, see the following paragraph. Where algorithmic changes are needed to achieve the required scalability the science case will normally be impacted. Consideration will have to be given to the numerical accuracy of results and their relation to the established literature.

The PRACE scientific steering committee has openly invited the proposition of new members to join its panel of reviewers. Membership of this panel by scientists of suitable standing from individual countries is strongly encouraged. From the PRACE perspective it promotes diversity and mitigates the possibility of apparent bias. From the member country perspective it serves a direct channel for dissemination to the local research community regarding the operation of the review process.

6.2 Engagement of HPC expertise

Where possible research groups can include computational specialists with scientific domain specific expertise most likely drawn form national centres (or equivalent). Experience as expressed in the case studies has shown that the involvement of HPC specialists as early as the project application preparation phase is best practise, greatly reducing the likelihood of rejection on the basis of technical misunderstanding. Exploiting Tier-0 systems is technically difficult work and can require system specific knowledge and experience that will be rare outside the setting of HPC centres. The emergence of technologies such as GPGPUs, which are technically difficult to use, further emphasises this role. On a case-by-case basis the benefit of additional computational expertise can be evaluated in terms of improved computational efficiency they may enable and what the financial value of this may be as a fraction of the notional value of the resource award in questions.

In advance of a “Grand Challenge” or Tier-0 resource application being made it is essential that appropriate preparations are undertaken this may involve development or at the very least benchmarking. Here too support should be extended, as this step is effectively essential to a successful Tier-0 application.

6.3 Outreach

Dissemination of the role of PRACE is crucial. National services or other Tier-1 and 2 providers should use the communications channels they have available to them to communicate the opportunities available to their user communities. Dissemination and promotion are aided by the complementary role PRACE access can play to existing services this allows the user community to structure their work and plan in advance their migration through tiers as required. Though the demands of Tier-0 usage should always be clear, i.e. it is not a merely a way to do more Tier-1 style work and the technical and scientific requirements reflect this.

Centres should not restrict their dissemination efforts to PRACE. There is much commonality in the aims and requirements of European and US Tier-0 activity thus centres should also promote US based calls. This can strengthen or prompt new intercontinental collaboration by researchers with little additional overhead.

Centres that work with users who have successfully migrated to Tier-0 should engage with those users to aid in the promotion of their work. Whilst acknowledging the achievement and research involved it also promotes the availability of Tier-0 access and associated services to a local audience.

6.4 Training and Education

Training plays a vital role and must not be confined to existing Tier-0 users. Indeed expansion of the ecosystem requires that the user community engage heavily with training programmes at the national or European level. Specifically to Tier-0 migration it should be recognised that advanced and focused training is required. Tailoring this training to a given system is reasonable given the small number of such systems. The importance of the provision of training is recognised as a requirement in the Tier-1 centre definition see section 2.3.1. This definition also highlights the need to offer advanced development platform access. In practice advanced platforms e.g. GPGPU enabled systems will fulfil a dual role that of development and training.

PRACE is frequently involved in the provision of advanced training through seasonal schools under the auspices of Work Package 3. Centres should promote this activity and users intending to use Tier-0 systems in the future should avail of the training opportunities made available. Of particular value to users migrating to Tier-0 are advanced courses covering topics such as scaling and debugging at Tier-0 scales. Doing so presents technical challenges not present at Tier-1 and so are unlikely to be well addressed by less specialised training material.

7 Conclusions

A window of opportunity exists for the promotion of Tier-0, DEISA through its DECI calls has established a pool of potential users and support personnel who are familiar with the challenges related to tier migration. PRACE activities such as community code petascaling and preparatory access also have a strong role in developing the skills base and the range of Tier-0 enabled software.

Soon to be commissioned large-scale Tier-1 systems will lessen the step change effect thus aiding the user community, though the user community must also be aware of the differences inherent in Tier-0 access.

It is clear that in each of the cases studied current policies for promotion of the PRACE RI at the national level are undertaken seriously. The opportunity to offer a coherent range of services that can be seen as integrated with existing services at the national level is seen as important by national level HPC providers in terms of its attractiveness to the user community. Both large and small countries are of the view that support at the level of Tier-2 and 1 are in long run required for the development of a Tier-0 community.

Relationships between PRACE partners are developing and well defined through both PRACE project and AISBL activities. The PRACE AISBL now incorporates 21 members i.e. countries. External relationships with communities are developing in response to community code enablement work. External relationships with other HPC programmes are emerging most particularly in the areas of joint training programmes, see sections 2.4 and 2.5, as such training activity plays a dual role of enabling the work of existing PRACE users but also being a first step to expanding the ecosystem. Section 2.5 outlines succinctly classes of relationship that are relevant to PRACE and what constraints should limit the expansion or number of such relationships.

The costs associated with procuring and operating Tier-0 systems are very significant. For small nations the costs are prohibitive and as seen in section 5.2.2 even relatively modest efforts in the area of capability computing can be unsustainable. Tier-0 class access is not a service which is commercially available, i.e. there are no commercial providers offering Petatflop/s class systems listed in the Top500. Thus in the absence of mechanisms such as those provided by PRACE, researchers in smaller countries have no means of migrating to Tier-0 class systems.

While costly projects such as a Tier-0 system are an option for only the largest nations smaller countries can invest in deploying human expertise to work on key and often emerging technologies such as GPGPUs which are having a disruptive effect. This type of investment is more sustainable and where the codes and hardware addressed have widespread application the benefits will spread through the ecosystem including to Tier-0 centres.

In summary this deliverable presents a body of detailed information on current operational practices in the context varied national circumstances along with findings and recommendations based on this and extensive consultation with relevant Work Packages. It recognises the differing circumstances that are to be found across both established and emerging countries in the area of HPC. Users of this document are invited to identify activities and practices relevant to them and examine the outcomes and recommendations that follow from them.

7.1 Future activity

This deliverable comes at month twelve, the midpoint of the project. Inline with the remaining duration of the project effort remains to be expended by the Work Package partners to further the aims of the deliverable. This process is timely as the advent of production Tier-1 activity within PRACE greatly enhances its relevance to emerging countries and the maturity and growing diversity of the Tier-0 service enhances its attractiveness to the existing Tier-1 user community. The following paragraphs describe activities that are now being planned to promote the involvement of new EU and potentially neighbouring countries.

A series of face-to-face meetings in Q3 and Q4 of 2011 commencing following the summer vacation period whereby the Work Package partners involved will seek to address both user communities and local HPC centre staff in a seminar format. The aim in the first instance is to disseminate the activities and services available through PRACE where necessary and then more specifically the findings and practice recommendations raised in this document. Initially Hungary which was recently made a PRACE AISBL member will be approached. On a phased basis meetings shall then follow with nations such as Denmark, Slovenia, Malta, Croatia and Romania. Several new EU countries are active at PRACE project level e.g. Serbia, the Czech Republic Bulgaria and Poland. Indeed Bulgaria and Poland are members of the new Tier-1 expanded DECI programme. In consultation with the relevant PRACE representatives in these countries the potential for dissemination will be discussed. This is a two-way exchange process whereby the PRACE project will also gather information. Though it reduces the opportunity for direct outreach the alternative of the reception of delegations is an option should it be necessary. This activity will be carried out in close consultation with the PRACE AISBL to ensure a common approach.

The HPC-Europa2 project largely targets researchers at the beginning of their careers. At the HPC-Europa2 Transnational Access Meeting (TAM) in early June 2011 a Work Package partner presented the opportunities available through PRACE with due emphasis on the inter-tier migration process. It is also planned to present at the next HPC-Europa2 TAM in order to maximise exposure at the European level to an audience whose makeup changes rapidly an might be considered open to new technologies an approaches.

The PRACE-2IP project clearly relates to PRACE-1IP and will involve 21 European partners from the outset and possible further full members or “observers” in the future. The will be considerable complimentary activities in PRACE-2IP. Once the project begins formal contact will be made with the leaders of Work Packages 3, 4 and 7 whose work will compliment this task. Collaborating with these Work Packages during the overlapping period will reinforce the work undertaken under the auspices of PRACE-1IP.

8 Annex

8.1 Irish Ecosystem supplementary data

8.1.1 3rd level education sector

The following statistics provide a broad outline of the scale of the 3rd level education sector in Ireland in terms of the number of students. Note the Republic of Ireland has 7 conventional universities, the statistics provided include in this number a further 5 small institutions which focus on dedicated fields e.g. The National College of Art and Design & The Royal College of Surgeons. While both offer extensive undergraduate programmes in ICT, the IoTs typically offer more certificate and diploma level places than the universities. At the postgraduate level the universities provide significantly more places than do the 15 IoTs

	Undergraduate awards in the academic year 2008/09		**Fulltime student enrolment in the academic year 2009/10**	
Field of Study, ISCED[1]	**IoTs**	**Universities**	**IoTs**	**Universities**
Broad Programmes	0	128	26	2
Education	96	1750	55	2243
Humanities and Arts	1564	4578	450	3340
Social Science Business and Law	6109	5599	1063	5242
Science	**1907**	**2387**	**725**	**3960**
Engineering, Manufacturing and Construction	**4096**	**1323**	**481**	**1740**
Agriculture and Veterinary	277	276	0	264
Health and Welfare	2685	3353	16	2527
Services	1499	735	123	162
Combined	37	0	0	0
Totals	**18270**	**20129**	**2939**	**19480**

Table 5: Headline 3rd level student number statistics[2]

8.1.2 Irish HPC Systems

Name	Walton	Iitac	Lanczos	Schrödinger	Stoney	Stokes
Operator	ICHEC	TCD	ICHEC	ICHEC	ICHEC	ICHEC
Manufacturer	IBM	IBM	IBM	IBM	Bull	SGI
Model	Cluster 1350	Cluster 1350	Blue Gene/L	Blue Gene/P	Novascale R422-E2	Altix ICE 8200EX
Processors	AMD Opteron 250 948 cores, 2.4GHz	AMD Opteron 250 712 cores, 2.4GHz	PowerPC 440 2048 cores, 700Mhz	PowerPC 450 4096 cores, 850MHz	Intel Xeon X5560 512 cores, 2.8 GHz	Intel Xeon E5462 2560 cores, 2.8GHz (upgraded to 3840 Xeon X5650 cores)
Interconnect	Gigabit	IB SDR	Proprietary	Proprietary	ConnectX IB	ConnectX IB

[1] Eurostat subject classification system.
[2] Source: Higher Education Authority.

Name	Walton	Iitac	Lanczos	Schrödinger	Stoney	Stokes
	Ethernet				DDR	DDR
Performance R_{max} (TFlops)	3.14	2.72	4.74	11.1	5.14	25.1 / 36.6
Tier	1 National	2 Project	1 National	1 National	2 Project	1 National
Currently Operational	No	No	No	No	Yes	Yes

Table 6: Irish HPC Systems

8.1.3 *ICHEC project classification*

- Class A applications ("Grand Challenge") are typically submitted by consortia concerned with "Grand Challenge" problems. These groups require resources representing a substantial fraction of the centre's resources over a long period. This type of application is expected to yield high-impact scientific publications.
- Class B applications ("Regular") meets the needs of the bulk of our user community, typically consisting of small research groups or individual researchers.
- Class C applications ("Discovery") support users with little or no prior experience of HPC, as well as more experienced users who would wish to gain a better understanding of their requirements before committing the resources to prepare an application for a Class A or B project. Their typical use consists in small-scale runs for the former, and code porting, optimising, and benchmarking for the latter. Significantly class applications are open to PhD students to make in there own name and without the involvement of their supervisors, thus lowering the administrative burden.

	Class A	**Class B**	**Class C**
CPU Hours	4,500,000	600,000	25,000
Duration	3 years	2 years	1 year
Review	International peer review	Scientific Council peer review	Internal review by ICHEC
Current number	10	43	78

Table 7: ICHEC Project class properties

Award Date	Call	Applicant(s)	Institution(s)	System(s)
Apr-11	PRACE production	Turlough Downes, Stephen O'Sullivan, Wayne O'Keeffe	DCU, DIAS, DIT	Jugene
	PRACE Preparetory Access Calls 1 & 2	Gareth Murphy	DIAS	CURIE, Jugene
	PRACE Preparetory Access Calls 1 & 2	Gareth O'Brien	UCD	CURIE, Jugene
	Discretionary Allocation	Turlough Downes, Stephen O'Sullivan, Wayne O'Keeffe	DCU, DIAS, DIT	Intrepid (DOE)
Feb-10	DECI 6	Simone Meloni	UCD	x86
	DECI 6	Turlough Downes	DCU, DIAS	Jugene
Jun-09	PRACE Prototype	Turlough Downes	DCU, DIAS	Jugene, Juropa, CSC-CSCS CrayXT5
Sep-09	PRACE Prototype	Jim Greer	Tyndall	Jugene, Huygens
	PRACE Prototype	Niall English	UCD	Jugene, Juropa CSC-CSCS CrayXT5
Jun-09	PRACE Prototype	Gary McGuire	UCD	Jugene, Juropa, CSC-CSCS CrayXT5, Huygens
	PRACE Prototype	Damien Thompson	Tyndall	Jugene
Mar-10	Grand challenge CINES	Niall English	UCD	Jade 2
Dec-08	IBM BG/P multi-cab access	Turlough Downes	UCD	Watson, Rochester
	IBM BG/P multi-cab access	Damien Thompson	Tyndall	Watson, Rochester
Nov-09	IBM BG/P multi-cab access	Niall English	UCD	Watson, Rochester
	IBM BG/P multi-cab access	Jiri Vala	NUIM	Watson, Rochester
Apr-09	ICHEC/EPCC agreement	Michael Nolan	Tyndall	Hector
	ICHEC/EPCC agreement	Graeme Watson	TCD	Hector
	ICHEC/EPCC agreement	Niall English	UCD	Hector

Table 8: ICHEC supported PI's who have secured resources externally