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

D2.4.2
Initial Impact Assessment of the Research Infrastructure

Final

Version: 1.2
Author(s): Carlos Merida-Campos, BSC; Ana Bela Dias, NCF

Date: 24.10.2011

Project and Deliverable Information Sheet

PRACE Project	**Project Ref. №: RI-261557**	
	Project Title: PRACE First Implementation Project	
	Project Web Site: http://www.prace-project.eu	
	Deliverable ID: < **D2.4.2**>	
	Deliverable Nature: <DOC_TYPE: Report>	
	Deliverable Level: PU / PP / RE / CO *	**Contractual Date of Delivery:** 31 / October / 2011
		Actual Date of Delivery: 31 / October / 2011
	EC Project Officer: Bernhard Fabianek	

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

Document Control Sheet

Document	**Title: Initial Impact Assessment of the Research Infrastructure**	
	ID: D2.4.2	
	Version: 1.2	**Status:** Final version
	Available at: http://www.prace-project.eu	
	Software Tool: Microsoft Word 2007	
	File(s): D2.4.2.docx	
Authorship	**Written by:**	Carlos Merida-Campos, BSC Ana Bela Dias, NCF
	Contributors:	Sergio Bernardi, CINECA Virginie Mahdi, GENCI Georgi Prangov, NCSA Michael Browne, ICHEC Franziska Schaerli, CSCS Dominik Ulmer, CSCS Michele de Lorenzi, CSCS Manuel Fiolhais, UC-LCA Jussi Heikonen, CSC Pekka Manninen, CSC Marcin Ostasz, BSC Arnaud Delimoges, GENCI Nelly Stoyanova, NCSA Pascal Moussier, GENCI
	Reviewed by:	Florian Berberich, JSC, Hans Eide, USIT
	Approved by:	MB/TB

Document Status Sheet

Version	Date	Status	Comments
0.1	8/July/2011	Guidelines	Guidelines of content and initial materials to be further developed
0.2	1/September/2011	Scaffold	Sectioned deliverable with preliminary content on it
0.3	15/September/2011	First draft with populated sections	First input gathered and integrated From NCF, CSCS and BSC.
0.4	23/September/2011	Second draft with further populated sections	With additional input from BSC, CSC and CINECA
0.5	27/September/2011	Third draft with further populated sections	With additional input from BSC, ICHEC, UC-LCA, NCSA. Reviewed for consistency
0.6	3/October/2011	Fourth draft	Improved and structured introductory sections
0.7	4/October/2011	Fifth draft	Integrated comments from CSC
0.8	5/October/2011	Sixth draft	First draft of executive summary and conclusions integrated. Updated B3, D4, D5 and D9 comments on the different variables
0.9	8/October/2011	Pre-final	Completed conclusion, added refinements by CINECA (exec summary), UC (C1) and CSC (B10), minor corrections by NCF until section 3.3
1.0	10/October/2011	Final for internal review	Formatted and completed missing input and unaddressed comments
1.1	10/October/2011	Final for internal review	Reviewed and completed the two missing variables.
1.2	24/October/2011	Final version	All comments of reviewers addressed

Document Keywords

Keywords:	PRACE, HPC, Research Infrastructure, Monitoring, Reporting, Impact Assessment

Table of Contents

List of Figures

List of Tables

References and Applicable Documents

[1] High Performance Computing Facility Operational Assessment, FY 2010, Oak Ridge Leadership Computing Facility, Arthur S. Bland, James J. Hack, Ann E. Baker, Ashley D. Barker, Kathlyn J. Boudwin, Ricky A. Kendall, Bronson Messer, James H. Rogers, Galen M. Shipman and Julia C. White, U.S. Department of Energy, Office of Science, 2010, http://info.ornl.gov/sites/publications/files/Pub25273.pdf
[2] PRACE Statutes, Moniteur Belge Annex, 0098184, 5.7.2010
[3] STAR METRICS Program, https://www.starmetrics.nih.gov/
[4] Initiative aims to quantify the paybacks from basic research, David Kramer, Physics Today, August 2010.
[5] A note on the economic impact assessment of the Research Councils, André de Campos, http://www.rcuk.ac.uk/Publications/policy/Pages/impactassessment.aspx
[6] Deliverable D2.3.1, Report on in-kind contributions, PRACE-1IP project.
[7] Deliverable D6.1, Assessment of PRACE operational structure, procedures and policies, PRACE-1IP project.
[8] Deliverable D4.4, Global Collaboration Opportunities: a plan for collaboration with non-European stakeholders for strengthening HPC usage, of the PRACE-1IP project.
[9] Research Excellence Framework impact pilot exercise: Findings on the expert panels. A report to the UK higher education funding bodies by the chairs of the impact pilot panels.
[10] http://www.hefce.ac.uk/research/ref/pubs/other/re01_10/
[11] Deliverable D2.4.1, Monitoring and Reporting Procedures, PRACE 1IP project.
[12] TeraGrid, https://www.teragrid.org
[13] https://www.teragrid.org/web/eot/
[14] INCITE, http://science.energy.gov/ascr/facilities/incite/about-incite/
[15] NSF, http://www.nsf.gov/x
[16] NSF OCI, http://www.nsf.gov/dir/index.jsp?org=OCI
[17] Initiative aims to quantify the paybacks from basic research, David Kramer, Physics Today, August 2010.

List of Acronyms and Abbreviations

AC	Access Committee
BoD	Board of Directors
BSC	Barcelona Supercomputing Center (Spain)
CINECA	Consorzio Interuniversitario, the largest Italian computing centre (Italy)
CINES	Centre Informatique National de l'Enseignement Supérieur (represented in PRACE by GENCI, France)
CPU	Central Processing Unit
CSC	Finnish IT Centre for Science (Finland)
CSCS	The Swiss National Supercomputing Centre (represented in PRACE by ETHZ, Switzerland)
EC	European Community
ETHZ	Eidgenössische Technische Hochschule Zuerich, ETH Zurich (Switzerland)
GENCI	Grand Equipement National de Calcul Intensif (France)
GPU	Graphic Processing Unit
HM	Hosting Member
HPC	High Performance Computing
ICHEC	Irish Centre for High-End Computing
JSC	Jülich Supercomputing Centre (FZJ, Germany)
LINPACK	Software library for Linear Algebra

NCF	National Computing Facilities (Netherlands)
NCSA	National Centre for Supercomputing Applications (Bulgaria)
NIH	National Institutes of Health, USA
NSF	National Science Foundation, USA
OS	Operating System
PI	Principal Investigator
PRACE	Partnership for Advanced Computing in Europe; Project Acronym
PRACE-1IP	PRACE 1st Implementation Phase project
PRACE-2IP	PRACE 2nd Implementation Phase project
PRACE AISBL	PRACE Association International sans But Lucrative
R&D	Research and Development
SLA	Service Level Agreement
SME	Small and Medium Enterprise
SSC	Scientific Steering Committee
TCO	Total Cost of Ownership. Includes the costs (personnel, power, cooling, maintenance, ...) in addition to the purchase cost of a system.
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
UC-LCA	Universidade Coimbra-Laboratório Computação Avançada (Portugal)
USIT	Research Computing Services, University of Oslo (Norway)

Executive Summary

Good management is one of the key factors that guarantee the healthy status of an organization and its sustainability as well as its development and evolution. At present, the notion of good management for organizations with not-for-profit mission and objectives like PRACE AISBL is widely accepted. It is therefore important that the notion of impact assessment of the actions undertaken by the organization is understood and properly addressed by the governance and management bodies. Of course the non-profit character of PRACE has to be taken into consideration when implementing the necessary processes, tools and procedures for assessing the impact. The organisation drivers have to be different from those considered by commercial and business oriented organizations that focus almost exclusively on how the actions impact revenues and costs reduction. Moreover, impact evaluation is addressed not to improve the operation of the infrastructure or the performance of plans, but to know the Return of Investment of the organisation, the equivalent to the profit made for a profit-oriented organisation.

The objective of this deliverable is to propose a model for assessing the impact of PRACE by means of a selected group of variables whose measures can provide a picture of how well PRACE is doing.

The results of D2.4.1, Monitoring and Reporting Procedures, provided the necessary preliminary identification and analysis of a large set of variables where those related to impact and applicable to the specificity and peculiarity of the PRACE AISBL organization are a subset.

Impact assessment plays an important role in the overall PRACE workflow and evaluation process that provides feedback on how the actual results compare against the planning. We considered three types of evaluation: efficiency, effectiveness and impact.

Based on the PRACE workflow model, the monitoring targets and the derived variables are classified accordingly to four groups:

- **Input**: what resources PRACE uses in order to create and run its services such as monetary funding and budget, personnel resources and equipment;
- **Delivery**: the services provided by PRACE;
- **Output**: the concrete results obtained with allocated compute resources;
- **Environment**: the market environment in which PRACE delivers its services.

The decision on the types of impact to be considered for PRACE AISBL and how to elaborate on them is the first crucial task. The impact of PRACE AISBL can be divided in four types:

- **Scientific:** By allowing users to solve important queries in their scientific fields and introducing HPC to new fields of science, PRACE AISBL directly contributes to improving scientific knowledge in all fields of science for which supercomputing is essential;
- **Economic:** by contributing to the increase of scientific knowledge PRACE AISBL directly or indirectly contributes to the economic development of Europe;
- **Social:** by increasing scientific knowledge in research fields such as climate, life sciences, etc. PRACE AISBL contributes to improvement of quality of life of European population and European social development;

- **Environmental:** PRACE AISBL can have important contributions in the fields of:

- Climate;
- Energy;
- Water management.

Due to the characteristics of HPC, impact assessment can be:

- **Quantitative:** based on defined quantitative targets or compared to quantitative baselines that need to be properly identified and need to be analysed on a regular basis to decide if they are still valid or need to be revised;
- **Qualitative:** based on case studies or success stories that are properly identified and analysed. The Scientific Steering Committee of PRACE AISBL can have an important role for identifying and validating case studies and success stories.

Although qualitative assessment could prove cumbersome given its nature based not on measurement but rather on some subjective evaluation, it is felt that it could not be neglected because of its relevance when addressing economic, social and environmental impact. The deliverable recommends the following to be implemented by PRACE AISBL:

- The development of quantitative monitoring on the strategic level via Balanced Scorecard that combines between 6 and 12 strategic indicators for steering purpose. They are meant to represent the four dimensions “Input”, “Delivery”, “Output” and “Environment” highlighted by D2.4.1 outcome.
- Regular peer review of the PRACE organisation in order to reflect the state of the organisation at strategic level. The standard tool for this process in the academic realm is the peer review that should carried out by a set of senior experts outside the inner circle of PRACE stakeholders.

In deliverable D2.4.1 the monitoring variables were classified and described in detail. In the present deliverable the focus goes into impact and its assessment. Impact assessment has been investigated from the theoretical perspective to gain a deep understanding on the concept, and best practices have been analysed from different publicly available documents, interviews with managers and directors of non-profit organisations, and experience from the different centers of PRACE members.

1 Introduction

Business oriented organisations evaluate their activity in economic terms: measuring their revenues and setting up their strategies to maximise it. Companies usually set their strategic objectives taking into consideration short, medium or long term revenues, for example offering low prices to clear their stock, lowering their profit margin for avoiding the emergence of competitors or abandoning a profitable market for consolidating their presence in a more sustainable one. All these activities are profit-oriented and money-driven, and their impact is measured by the short, medium or long term economic benefit.

For Research Organisations such as PRACE AISBL, the mission is not profit-driven, and the impact of the actions is not measured in revenue terms, but in terms of how global and significant this impact is. Hence impact assessment is a key for a not-for profit organisation to know the actual results of the strategic decisions taken. However assessing impact is not as easy as checking the balance sheets of a company. Nevertheless, there are suitable mechanisms for impact analysis of any type of organisation. This deliverable shows a potential mechanism to assess the impact of a selected set of monitoring variables explained previously in D2.4.1 [11].

1.1 Impact Assessment in the workflow of an organisation

The fundamental purpose of any company is established through its vision and its mission. The strategic plan to implement the mission (business plan) encompasses the analysis of the context (market research, and analysis of competitors and stakeholders) and also the planning for the different relevant business areas such as financial, marketing and operation planning. The business planning might be up to the level of detail of the general institutional values of the workers. This planning and detailed description of the activities sets a solid base to start operating, but independently of the type of business area or the size of the business, the activity needs to be continuously monitored and the plans adjusted accordingly with the analysis and assessment of the observations. These adjustments will be oriented at improving the efficiency, effectiveness and impact of the organisation, hence fulfilling its mission while achieving higher customer satisfaction, improving processes and outcomes, and consolidating its sustainability.

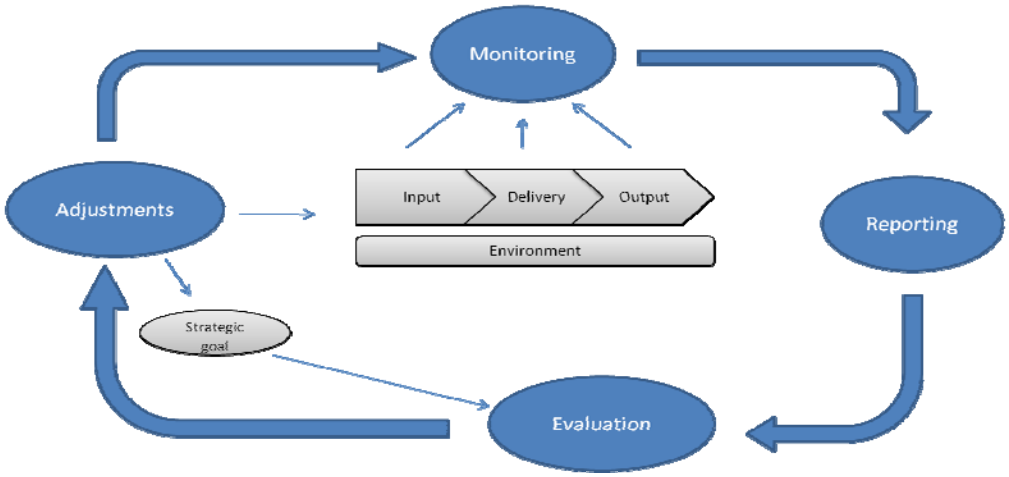

Figure 1: Management cycle of PRACE AISBL

Before initiating evaluation of its processes and procedures, the organisation has to select variables (monitoring variables) and set up mechanisms to continuously observe the selected variables, and also report them in the appropriate format to the appropriate stakeholders. As can be seen in Figure 1, evaluation is preceded of by two phases: monitoring and reporting. In Deliverable 2.4.1 [11], these mechanisms have been explained and a complete set of variables was identified, classified and defined.

Data for the variables identified can be captured at different stages of the workflow. Also in D2.4.1, the stages identified for classifying data were: input, delivery, output and environment.

Figure 2: Stages of the workflow of PRACE

1.2 Impact Assessment as an Evaluation process

Aside of the classification of the monitoring variables according to the stages of the workflow, these can be also classified according to the type of evaluation resulting from them. Evaluation is the comparison of actual results against the agreed strategic planning. It looks at what is set out to be done, at what has been accomplished, and how it has been accomplished. It can be formative (taking place during the life of a project or organisation, with the intention of improving the strategy or way of functioning of the project or organisation). It can also be summative (drawing lessons from a completed project or an organisation that is no longer functioning).

What monitoring and evaluation have in common is that they are geared towards learning from what has been done and how it has been done, but evaluation in particular, involves:

- Looking at what the project or organisation intended to achieve – what difference did it want to make? What impact did it want to make?
- Assessing its progress towards the intended achievements, i.e. its impact targets.
- Looking at the strategy of the project or organisation. Was the strategy adequate? Was the strategy followed effectively? Did the strategy work? If not, why not? What should be changed/adopted to reach the proposed strategy?
- Looking at implementation of the strategy. Were the resources efficiently used? How sustainable is the way in which the project or organisation works? What are the implications of the way the organisation works for the various stakeholders?

Three different types of evaluation have been identified [10]:

- ***Efficiency***: monitors if the input into the workflow is appropriate in terms of the output. In the case of PRACE, at present operating according with the Cycles model, the input is the type and amount of computing resources, the necessary support staff and the budget. In the near future, if PRACE evolves to a different model, the budget necessary to run the organisation will become much more relevant. These parameters

are very important for the day-to-day business of PRACE and can be of utmost importance for scaling up the organization;

- ***Effectiveness***: is a measure of the extent to which a development programme or project achieves the specific objectives it set. For PRACE, it will be a measure of the goals defined by the PRACE mission;
- ***Impact***: Indicates whether or not what has been done made a difference to the problem situation or status quo that the organization is trying to address. In the case of PRACE it is important to track and verify the impact of the mission defined for PRACE. In other words, is the strategy defined reaching the desired goals?

1.3 Importance of Impact Assessment

While taking wrong decisions and failing may be acceptable, for an organization it is essential to be able to learn from mistakes and act upon them. An inefficient and ineffective organisation will have difficulties in achieving its mission, and thus the organisation workflow needs to be carefully considered and evaluated to act upon it. However, assessing impact should be the most important evaluation objective for PRACE. Impact evaluation is addressed not only to improve the operation of the infrastructure or the planning performance, but, most importantly, to know the Return of Investment of the organisation, i.e. the equivalent to the profit for a profit oriented organisation. This evaluation is essential to justify the existence of the organisation, i.e. the final reason why the organisation operates and invest significant amounts of monetary and human resources for providing its services. Moreover, the results of this assessment will provide key information to the management of PRACE to decide on different levels of strategic changes.

In deliverable D2.4.1 [11] the monitoring variables were classified and described in detail. In the present deliverable the focus goes into impact and its assessment. Impact assessment has been investigated from the theoretical perspective to gain a deep understanding on the concept, and best practices have been analysed from different publicly available documents, interviews with representatives of not-for profit organisations, and the experience from the different computing centers of PRACE members.

Section 2 illustrates the theoretical impact assessment framework applicable to PRACE and how to build up on this framework.

Section 3 shows a proposal for a Balance Scorecard mechanism as tool to be used by the management of PRACE for the results of the assessment, and lists all the impact related variables described in D2.4.1. The description previously made covered the characterisation of the data to be captured and proposals for a mechanism to automatically capture and report this data. In this deliverable, for each of these variables, ideas for an impact assessment mechanism are presented. This section also provides guidelines for qualitative assessment in PRACE.

Finally, Section 4 provides the conclusions for this work and indicates how it will be pursued.

2 Theoretical framework for Impact Assessment

2.1 Types of impact

As discussed in D2.4.1 [11], the sources of data used for the monitoring variables are the following:

- Case studies and success stories;
- Recorded observation;
- Diaries;
- Recording and analysis of important incidents;
- Structured questionnaires;
- One-on-one interviews;
- Focus groups;
- Surveys;
- Systematic review of relevant official statistics.

The information to be collected must be meaningful, must be collected in a homogeneous way (this is very important if the information is collected by various parties) and should be stored taking accessibility into account. Usually it is possible to use reports, minutes, attendance registers, financial statements, and other day-to-day material as a source of monitoring. Nevertheless, sometimes other external information sources are required. The quality and validity of impact assessment depends on how trustful and reliable the sources of data are.

Impact, as previously defined in 1.2, indicates whether or not the organisation has made a difference to the problem situation or status quo that the organization is trying to address. In the case of PRACE it is important to track and verify the impact of the mission defined for PRACE. In other words, the impact reached by making HPC resources available to European researchers from academia and industry needs to be measured or inferred in order to validate the performance of the organisation and justify the investments made by the members and in particular by the Hosting Members. Impact assessment can also be used to create awareness of the importance of the mission of the PRACE Research Infrastructure and can be relevant for informing stakeholders, i.e. funding organisations, HPC users, other HPC organisations and research infrastructures, EC, and the public in general.

One of the first tasks is to decide on the types of impact that need to be considered for PRACE AISBL and elaborate on them. The impact of PRACE AISBL can be divided in four types:

- **Scientific:** users from academia and industry use the HPC resources made available by PRACE AISBL for solving or improving the knowledge of important questions that arise in their scientific fields. By allowing users to solve important problems in their scientific fields, PRACE AISBL directly contributes to improving scientific knowledge in all fields of science for which supercomputing is essential;
- **Economic:** by contributing to the increase of scientific knowledge PRACE AISBL directly or indirectly contributes to the economic development of Europe by:
  - Promoting multi-national collaborations inside Europe and/or with partners outside Europe;
  - Promoting collaborations with industry with potential knowledge transfer;
  - Creation of jobs, either through contributing to scientific work for PhD theses or to scientific work that can be applied in industrial applications;

  - Decreasing loss of property by e.g. contributing to the development of climate and earth-quake forecast, flooding prevention, etc.;
  - Decreasing energy dependence and costs by promoting development in the fields of fluid dynamics (combustion, aerodynamics), fusion energy, life sciences, etc.;
  - Contributing to green computing by pursuing computing solutions that are more energy efficient;
  - Increasing European competitiveness by attracting new industries to Europe, creating new businesses and increasing the number of jobs.
- **Social:** the mission of PRACE AISBL contributes also to the European social development in several ways, including:
  - Contributing to the creation of new jobs;
  - Decreasing impact of natural disasters produced by e.g. earth quakes and climate extremes;
  - Improving health and quality of life mainly by contributing to developments in the fields of medicine and life sciences. This may include where applicable, the assessment of lives saved, better health outcomes, changes to clinical guidelines and practices, and improved public awareness and other changes to healthcare policies;
- **Environmental:** PRACE AISBL important contributions in the fields of:
  - Climate;
  - Energy;
  - Water management.

This may include where applicable, the assessment of the influence on policy development and implementation relating to the environment and climate change, development of processes, services and technologies relating to conservation, environmental management, monitoring and risk assessment and product and service development (particularly in the oil, energy and utilities sectors).

These types of impact are not easy to measure and in some cases quantification is rather complicated or almost impossible due to the characteristics of HPC. In most cases HPC is essential for modeling various phenomena, either because they are very complex or are too dangerous and/or expensive to be tested under laboratory conditions or even because they are so poorly understood that several possible models need to be tested theoretically to draw conclusions regarding the best hints to pursue meaningful research work. Another characteristic of HPC is that although it is essential for advancing of science, it is in some cases seen simply as a tool, and researchers in their publications refer more to the theoretical models used for their simulations than to the computer resources and the software used. The role of computer resources is even less visible in some industrial applications. Though they were not possible without access to HPC resources, they enter the production circuit in some cases years after the initial research computer simulations and the role of HPC appears rather diluted. One of the best examples is the development of new drugs to cure or relief symptoms of diseases that afflict mankind. At present these developments are not possible without usage of HPC resources for simulation of viruses, proteins or other biological components, but have such a long time to market, i.e. till reaching the consumer, that it is very difficult to know how much computer resources (hardware and software) were necessary to develop the final drug. Similar examples, can be found e.g. in climate and earth-quake forecasts, or for understanding how to address coronary disease, obesity, ageing, etc. Some of the most important discoveries

in science were only possible not by using one single supercomputer, but on the contrary by using various generations of supercomputers that go far back in time.

Due to the typical characteristics of HPC, impact assessment can be:

- **Quantitative:** based on defined quantitative targets or compared to quantitative baselines that need to be properly identified and need to be analysed on a regular basis to decide if they are still valid or need to be revised.
- **Qualitative:** based on case studies or success stories that are properly identified and analysed. In some cases, the identification of case studies should be a distributed task and a specific committee should be in charge of evaluating them.

Though quantitative assessment seems, in general, to be more trusted, in the case of HPC qualitative impact assessment cannot be neglected and may be more relevant when addressing economic, social and environmental impact.

Scientific impact can seem to be easily addressed in a quantitative way by e.g. taking into account the impact factor of the scientific journals where research results are published. However, it should be noted that the impact factors of the top journals are dependent on the scientific field, i.e. what can be considered a high impact factor for one scientific field may be a lower one for another scientific field and vice versa. The same applies for the "h" and "g" indexes of researchers working in very different scientific fields. The conclusion is that even quantitative values need to be analysed with care in the case of HPC, especially if they apply to different scientific fields. There are of course some monitoring variables for which quantification is straightforward, e.g. success ratio of proposals for access to PRACE AISBL, geographic distribution of the projects supported by PRACE, etc.

Finally, it third dimension for classifying the impact needs to be considered, i.e. according to the timeframe for measuring the impact. According to this classification the following three impact categories apply:

- **Short Term timeframe:** the organization can have an impact in a short timeframe since the beginning of its operation. This impact has usually a short reach and significance; however it could be an indicator of the impact of the activity in the future. Hence the information provided by short timeframe impact is of relevance to the PRACE AISBL. In this category we classify the assessments made regarding awareness created of the ongoing work of the organization, the trend in attendance to events organized by PRACE, etc.
- **Medium Term timeframe:** in a medium term timeframe, an organization can start measuring the impact of its operations in a higher reach and significance context. In this timeframe, some socio-economic results can be fostered by the organization. In this category we classify the assessment made on the trends of job creation, publications, and any scientific impact in general.
- **Long Term timeframe:** the long term timeframe is when the real impact can reach society in general after the necessary and long stages of research and development and market adoption, a high reach and significance change could be made in society. This type of impact is the most important for an organization such as PRACE to keep track of. It is, however, also the most complex to analyse. In this category we classify the assessment made on the improvement of life quality, the improvement of practices in any discipline by the usage of improved tools, or the changes of social policies. Although it could happen that long timeframe impact would be reach in a short timeframe, this is very unlikely.

2.2 Impact assessment methods in institutions similar to PRACE

Most organisations similar to PRACE AISBL, i.e. organisations that allocate HPC resources through a peer review process, either in or outside Europe monitor the impact of their allocations.

As in the case of PRACE AISBL these organisations are interested in the quantitative and qualitative impact of their existence.

Quantitative impact

Quantitative impact mostly regards success rate of proposals, distribution of successful proposals per scientific field, percentage of allocation of projects to academia and industry and other quantitative measurements that mainly result from statistics regarding the calls for proposals. These quantitative measurements usually appear in the annual reports of the organisations and in other publications reporting on the annual activities. This is the case of most HPC European and worldwide organisations. These types of impact measurements are rather straightforward and can also easily be obtained for PRACE AISBL.

Qualitative impact

The best examples of qualitative impact appear in most European and worldwide websites of organisations that allocate HPC resources and are usually supported by case studies and/or success stories published in the respective websites. In these websites abstracts and participants in all projects supported are also usually published. This is not only important in terms of impact but also in terms of transparency of the peer review process used in that general information on allocated projects is public. PRACE AISBL is in line with this procedure and publishes on its website the abstract of projects, along with the name and institution of the participants in the project.

Case studies and success stories result usually from a selection of the most relevant projects allocated and usually include a description of the results obtained with HPC resources and also highlights of possible applications, either for future scientific or industrial development.

Outside Europe some examples of this type of qualitative impact can be found for institutions similar to PRACE AISBL:

TeraGrid [12]: TeraGrid is an open scientific discovery infrastructure combining leadership class resources at 11 partner sites (Indiana University, LONI, NCAR, NCSA, NICS, ORNL, PSC, Purdue University, SDSC, TACC and UC/ANL) to create an integrated, persistent computational resource. TeraGrid integrates high-performance computers, data resources and tools and high-end experimental facilities around USA. TeraGrid publishes every year a booklet entitled “TeraGrid Science Highlights” with science and engineering highlights of the results of projects allocated by TeraGrid. This booklet is also available at the TeraGrid website [12] in pdf format.

INCITE [14]: INCITE (Innovative and Novel Computational Impact on Theory and Experiment) is jointly managed by the Argonne Leadership Computing Facility and the Oak Ridge National Laboratory and is the HPC program of the Office of Science of the US Department of Energy (DoE). The INCITE program awards billions of processor hours on some of the higher-performance supercomputers of USA to researchers committed to accelerating scientific breakthroughs and bringing real-world benefits to USA citizens. INCITE publishes the abstract and the names and institutions of the applicants.

NSF [15]: NSF (National Science Foundation) OCI (Office of Cyberinfrastructure) [15] supports cyberinfrastructure resources, tools and related services such as supercomputers,

high-capacity mass-storage systems, system software suites and programming environments, scalable interactive visualization tools, productivity software libraries and tools, large-scale data repositories and digitized scientific data management systems, networks of various reach and granularity and an array of software tools and services that hide the complexities and heterogeneity of contemporary cyberinfrastructure while seeking to provide ubiquitous access and enhanced usability. OCI makes supercomputer resources available to users through a range of programs. The rubric "Discoveries" in the website lists case studies and success stories.

Efforts for measuring impact and justifying investments

As described above, ascertaining quantitative impact is not an easy task and most organisations focus on reporting qualitative impact mainly based in case studies and success stories. Nevertheless there are two efforts for developing a methodology for reporting quantitative impact, mainly economic and social, together with a tentative method for justifying the investments in HPC that are worthwhile some reflection.

One of these efforts was initiated by Research Councils in the UK and tries to understand the economic, social and environmental impacts emerging from the investments in science. This initiative results from the increase of the funding for science by the British Government in the last decade and tries to put emphasis on the accountability of the Research Centres in the UK. The initiative uses a combination of quantitative and qualitative methodologies. The quantifications regard mainly macroeconomic and microeconomic studies as well as survey data. The qualitative approaches regard mainly interviews and expert testimonies. Though this initiative focuses on science in general, it will be interesting to follow up on its developments to check if some best practices can be used for HPC and in particular for PRACE AISBL. Another important point that this study may also give some hints regards the relation between the investments and the impact in that the impact should be matched against the costs. This is not an easy task but should deserve some attention of PRACE AISBL.

Another effort, this time from USA, is the STAR METRICS (Science and Technology for America's Reinvestment: Measuring the Effect of Research on Innovation, Competitiveness, and Science) [3] program. This program is in line with the ambitions of the UK initiative and aims at measuring the impact of federal investments in science, particularly with respect to job creation and economic growth. STAR METRICS intends to use existing administrative data from federal agencies and their grantee institutions, and match them with existing research databases on economic, scientific and social outcomes. The initial goal of STAR METRICS is to provide mechanisms that will allow participating universities and federal agencies with a reliable and consistent means to account for the number of scientists and staff that are on research institution payrolls, supported by federal funds. In subsequent generations of the program, it is hoped that STAR METRICS will allow for measurement of science impact on economic outcomes (such as job creation), on knowledge generation (such as citations and patents) as well as on social and health outcomes. The STAR METRICS program is the follow up of a very successful pilot project tested in 2009 involving seven universities. STAR METRICS is a five-year program for which NIH and NSF have committed to providing a total of $1 million USD. STAR METRICS is seen as a very ambitious program [17] that has a long way to go before giving the planned information especially regarding societal benefits of research that only begins to show up years or even decades after the research funding occurred. Once again this program is directed to science in general, but it is interesting to follow up on these developments to analyse the methodology for deciding in best practices that can be applied for PRACE AISBL.

3 Recommendations towards PRACE impact assessment

As deliverable D2.4.1 has shown, a large number of variables can be used for monitoring the research infrastructure. Some of these variables, however, are either descriptive or very difficult to obtain. In particular in the field of impact assessment, qualitative assessments or assessments that require large and lengthy studies prevail. Furthermore, variables for impact and for operations work on very different timescales, which makes it difficult to construct a feedback cycle between input and operational principles on the one hand, and outcome and impact on the other. However, the PRACE research infrastructure has been set up with certain assumptions for best practices for its operation, which should lead to the intended impact of the organisation. Under these assumptions, it would be therefore sufficient in first approximation to measure the success of some of these assumed best practices. Whether the best practices really lead to the intended impact can then be verified in a second step. As the operational practices represent business practices, monitoring and steering methods in a business environment can be adapted to the needs of PRACE.

3.1 Quantitative impact assessment on the strategic level via Balanced Scorecard

For steering objectives, we therefore recommend to develop a Balanced Scorecard which combines between 6 and 12 strategic indicators. The indicators would represent the four dimensions “Input”, “Delivery”, “Output” and “Environment”. They should be easily obtainable and each of them is represented by a traffic light that can be red, yellow, or green. The research infrastructure, respectively the Council of PRACE, should formulate two threshold values for each parameter: one threshold that delimits the comfort zone (colour green) from the attention zone (yellow), and one threshold between the attention zone and the alarm zone (red). In order to be useful, only parameters that can be actively influenced by PRACE should be selected.

The Balanced Scorecard would represent the current status of the PRACE operation and could be regularly discussed within the PRACE Council. An example for a balance score-card of PRACE could look as follows (please note that this is not a proposal for individual variables or thresholds but only a demonstration of the principle, and that some of the parameters are not necessarily impact metrics):

Dimension	Parameter	Green zone	Yellow Zone	Red Zone	Assumption
Delivery	Success ratio of proposals as percentage of accepted vs. rejected proposals	<= 33%	33% < success ratio <= 50%	> 50%	PRACE has to attract the best science and has to be able to be selective
	Percentage of CPU time spent by jobs larger than or equal to 20% of the machine size	>= 40%	40% > percentage of CPU time on the target job size >= 25%	< 25%	PRACE addresses the needs in capability, not in capacity computing
	Overall satisfaction of users with the PRACE service measured on a scale between 0 and 10	>= 8	8 > overall satisfaction >= 6	< 6	The best scientists feel that the receive the best service
Output	Average journal impact factor of the 5 best publications enabled by PRACE	>= 30	30 > average impact factor of the 5 best publications >= 19	< 19	The best 5 publications should appear in recognised journals like “Science” or “Nature”
	Percentage of running projects with industry participation	>= 10 %	10% > percentage of projects with industry participation > 5%	< 5%	PRACE has identified service to industry as one of its success factors
Environment	Average percentage of HPC investments in the budget of the PRACE members	>= 35%	35% > average percentage >= 25%	< 25%	For PRACE being healthy, all PRACE members have to be able to continuously invest into infrastructure on their own level

It is important that in addition to the typical “traffic light” snapshot, the development of the indicators on the Balanced Scorecard are traced and presented over time.

3.2 Qualitative impact assessment on the strategic level via Case Studies

Although some of the impact variables monitored for PRACE need a qualitative assessment, this can be made via a simple analysis of data with the guidelines of the PRACE management, which will need to set-up the reference values to compare the result of the assessment. Using the terminology described in section 2.1, these are qualitative impact variables of a *short or medium term timeframe*, for which the relationship to PRACE can be stated directly and the relevance can be clearly assessed. In these “simple” cases, a 3-star mechanism or a traffic light as the one explained in the previous section is advisable.

Besides these qualitative assessment variables, there is an important case in the PRACE impact assessment framework represented by the *Social and Economic events* variable (see D11). This variable refers to the analysis of outstanding events or inventions, and the assessment of the potential influence of PRACE on them. Using our theoretical framework’s terminology, this variable refers to *long term timeframe* impact, and therefore, a deep analysis is required to assess not only the impact of the case study, but also the implication of PRACE

on it. For the complexity of this analysis, and the relevance of the conclusions to be extracted from it, a standard process to be followed is required. In [9] the Research Excellence Framework (UK) published the results of a recent pilot exercise of an impact assessment procedure, providing a set of findings and best practices that PRACE might take into consideration for running this type of task.

Below there is a proposed guideline for an assessment process for this variable based on the findings of the cited work:

1. Case studies (events) should be collected in a distributed manner, having not only the PRACE RI office collecting them, but also PRACE Members. These studies should be reported with precise information of the event, and the references to the usage of HPC that led to the success of the case.
2. Upon reception of the events data, and with a 2 year periodicity, PRACE AISBL may perform a pre-assessment to identify the need to further analyze the events. This assessment will be based on the information provided, internet search and if necessary, communications with involved parties.
3. After the analysis of all the events received, PRACE AISBL should discard those that have not been considered as relevant. In the case that the event was suggested by a PRACE member, this member should be contacted and the decision should be communicated. The reasons why an event could be rejected are:
   - Lack of impact (the events should only be assessed in terms of impacts or benefits that have taken place already; the panel should not attempt to anticipate future or potential impact); any indication of future potential should be treated as contextual, for example to help explain the significance of what has been achieved, but not claimed as actual impact;
   - Lack of clarity about the link between the underpinning usage of HPC and the specific impact claimed; low probability of finding out relationship with HPC during the process or the involvement of HPC in the process has been found as clearly supplementary or irrelevant;
   - The event is covered in other impact assessment variables analysis (if that is the case, the information should be forwarded to the specific monitoring team). Events that have been reported previously may be accepted just when the event evolves over time to the successive evaluation exercises. However in each evaluation, they should only be assessed for the specific impact that has taken place during that assessment period;
   - Bad quality of the reported event. Reports should provide the necessary information to perform the pre-assessment, with focused and concise evidence, without generalised or overblown statements or unexplained lists of publications or references. The information should also incorporate specific and appropriate independent sources for corroborating information were appropriate.
4. Having decided the set of events to be further analysed, an Assessment Committee should be created. PRACE AISBL should decide the best composition of the committee involving persons related or knowledgeable on the events under assessment and personnel related or knowledgeable on the specific supercomputer infrastructure that provided access to the HPC resources involved in the event. Aside of evaluating the impact of the event, the Assessment Committee should judge how critical has been the usage of HPC or PRACE for the success of the event.
5. PRACE AISBL should create evaluation guidelines for the committee, and the committee should operate according to them. The guidelines may include principles such as the following ones:

- Assessment panels might develop more detailed guidance on what constitutes impact in the specific area to be assessed. This should include guidance about the types of impacts and indicators anticipated from research in the respective disciplines, expanding on the initial list. The guidance should be flexible enough to allow for a wide variety of impacts and indicators;
- The Assessment Committee should only assess the impact occurred during the assessment period and not attempt to anticipate future or potential impact;
- In case it is necessary, the committee should held a face to face meeting and provide a final report with the result of the assessment;
- Events may be rejected by the panel by any of the reasons listed in point number 3.
- The result of the assessment should provide sufficient evidence for cases in which the assessment is positive proving the influence of PRACE or other HPC institution in the event. Aside of this, the result should also evaluate the reach and significance of the event. The assessment shall include where possible an estimation of the economic impact that the event has reached. To integrate these results into the balance score card record, the number of PRACE influenced events could be counted and grouped into the impact type category (social, economic, scientific and breakthrough), and ranked using the 4 star classification described in the following.

Any qualitative impact assessment should be assessed taking into consideration two criteria:

- **Reach**
- **Significance**

These two elements are orthogonal and should be analysed separately. Once they are evaluated, for joining the assessments into the complete case study, it is necessary to detail the application of the criteria; for example, whether a high 'reach' necessarily meant international recognition or not, and whether a high impact case study would need to demonstrate extensive reach as well as major significance. Panels may make holistic judgements on the merits of each case, and as such case studies may be able to achieve the highest grade with either exceptional significance, or exceptional reach; it should be not essential to have both.

For scoring the case studies, the following guidelines of level of assessment are provided:

Four star	**Exceptional:** Ground-breaking or transformative impacts of major value or significance with wide-ranging relevance have been demonstrated
Three star	**Excellent**: Highly significant or innovative (but not quite ground-breaking) impacts relevant to several situations have been demonstrated
Two star	**Very good**: Substantial impacts of more than incremental significance or incremental improvements that are wide-ranging have been demonstrated
One star	**Good**: Impacts in the form of incremental improvements or process innovation of modest range have been demonstrated
Unclassified	The impacts are of little or no significance or reach; or the underpinning research was not of high quality; or research-based activity did not make a significant contribution to the impact.

3.3 Regular peer review of the PRACE organisation

The quantitative impact assessment on a short time scale described has to be complemented by qualitative and/or targeted studies on certain subjects. In addition to these studies, PRACE should have regular review processes in place, which reflect the state of the organisation at the strategic level and are distant from the daily operations. The standard tool in the academic domain for this process is peer review. In a peer review of the organisation, a set of senior experts who are not part of the inner circle of PRACE stakeholders (i.e. not from institutions of representatives of PRACE members, PRACE users, PRACE founders or governing bodies, or PRACE suppliers), receive information about PRACE, e.g. the reporting package and other reports, and carry out interviews with PRACE members and stakeholders. Typically, the panel of experts receives terms of reference for the peer review which contain questions about three different fields of interest:

- Is the organisation doing the right things?
- Does the organisation do these things accurately? Is the organisation on track with its planning?
- What future developments that are important for the organisation, and do the organisation intend to address these developments?

The expert panel is not bound to the questions formulated in the terms of reference but is free to explore other findings during the review process. Thus, they are able to address topics that PRACE may not have identified yet. The peers should write a summary report of their findings and make recommendations for the improvement of PRACE.

It is recommended that PRACE let such a peer review be carried out every four years and that the report of the review panel is made available to all PRACE members. As the report will contain confidential information, only excerpts or summaries of the report should be made available to external PRACE stakeholders.

3.4 Variables to assess the impact in PRACE

The different monitoring variables identified in [11] were classified according to the purpose of their analysis to know whether they provide *efficiency*, *effectiveness* or *impact* information. Annex A shows a summary of the variables that are described next.

The process to assess the impact with the information captured for each of the impact variables is described as follows:

Delivery Variables

Peer Review Process

B.1 Success ratio of proposals – How strict is the PRACE AISBL peer review procedure?

Monitoring variable: Success ratio of proposals

Description: The success ratio of proposals, i.e. the percentage of proposals supported by PRACE compared to the total number of proposals to the call, for each call.

Main objective: The main objective of the analysis of this variable is to assess the quality and quantity of the proposals submitted to PRACE AISBL. This will be an indicator on the impact that PRACE is doing attracting good scientist to the usage of HPC. The Access Committee should set a quality threshold above which a given proposal should be accepted (bearing in

mind that some proposals might not be accepted because of the lack of resources on the PRACE side).

Source of data: All necessary data can be extracted from the PRACE tool for peer review. All data regarding PRACE calls is registered and stored using this tool.

Impact analysis: The type of impact assessment inferred from this monitoring variable is mainly scientific impact.

Main stakeholders: HPC users, scientific reviewers, AC and SSC (to provide the quality threshold)

Description of impact assessment: Two main quantitative values can be determined:

Success ratio of proposals: Defined as the percentage of proposals supported by PRACE compared to the total number of proposals to the call, for each call. This parameter can be measured for all types of proposals, i.e. for preparatory, project and programme access. For preparatory access, due its associated goal of testing or development of software applications there is no need of defining a baseline and the quantitative values should be mainly used as support of the need of this type of access. For project and programme the calculated value should be compared to a baseline value. As one of the main criteria of the PRACE peer review process is scientific excellence, it is important that PRACE has a sufficiently large pool of proposals for supporting proposals that are considered to be scientifically excellent, innovative and are expected to have relevant impact (mainly scientific, but also potential for knowledge transfer into future applications). Usually it is considered that the pool of proposals is large enough if the success ratio is around 30%.
Proposals above the quality threshold: percentage of proposals that were not supported by PRACE but were considered to be above the quality threshold by the AC regarding the total number of proposals. These proposals are not supported usually because of lack of resources. The baseline value should be around 20% to 30% of the total number of proposals. This means that PRACE is indeed selecting excellent proposals and not good or average proposals.

Calculation of these values is rather straightforward and these parameters can give quantitative information of the scientific impact of PRACE AISBL. It is also interesting to see the temporal evolution of these parameters to infer the interest and needs of the users, i.e. are more users applying to PRACE resources, and the evolution of the success ratio, i.e. is PRACE peer review becoming more or less strict.

B.2 H-index of applicants – How good is the scientific track record of the researchers supported by PRACE AISBL?

Monitoring variable: h-index and g-index of applicants

Description of the variable: h-index and g-index are quantitative measurements of the scientific productivity of an individual or Institution, from the point of view of the recognition by the scientific community. They are based on the number of publications and on the number of citations to the published work.

Main Objective: Though a typical h- or g-index of leading scientists might vary from area to area, the variable provides a way to measure the scientific excellence of the applicants to the PRACE services. One index or the other (or both) is always considered in any type of scientific assessment of both individuals and Institutions. The increase of these indexes due to PRACE usage is a measurement of the scientific impact of the PRACE infrastructure.

Source of Data:

- Indexes are available from scientific data bases (such as ISI Web of Knowledge by Thomson Reuters);
- Information provided by the PI in the application forms.

Main stakeholders:

- PRACE users;
- Assessment Committee;
- Entities related to PRACE;
- Funding agencies of PRACE HPC centers.

Description of the impact assessment process:

- The variable (either h-index or g-index) has just one single input parameter: an integer number;
- The value of the h- index is equal to the number of papers (N) ranked in decreasing order of citations that have N or more citations; the value of the g-index is the unique largest integer number such that the top N articles, also ordered in number of citations, receive (together) at least N^2 citations;
- h/g- indexes are quantitative parameters;
- The impact is mainly scientific;
- Reference values: strongly dependent on the scientific area of the proposals. A list with reference values should be produced for the main scientific areas (Physics, Chemistry, Engineering, Earth Sciences, Life Sciences, Health Sciences, etc.). Since h/g-indexes are usually non-decreasing with age, these factors should be used with caution in any evaluation of impact because they are likely to be lower for young scientists who nevertheless can have a very relevant scientific career. Instead they are appropriate for an overall characterization of scientific proposals to PRACE, rather than being used (with large impact) in the peer review process;
- Assessment process: The person assigned to the general gathering of information should carry on the task of getting the data from application forms or from databases;
- Results should be presented in spreadsheets together with other information which might help to characterize the scientific profile of the user communities of PRACE facilities. Some averages obtained from the indexes (for each field) should also be provided in reports referring to the scientific profiles of PRACE users. This helps to assess the scientific quality of applicants. On the other hand, and most important for the assessment of the scientific impact of PRACE, the changes in h/g-indexes of applicants due to their usage of PRACE resources, if traceable, provides a direct quantitative indication of the impact in science due to PRACE. However, gathering the information with this objective should only occur at least some three years after the scientific publication originated from the PRACE related research work.

B.3 Resource allocation – How are the resources allocated distributed (project type, scientific field, geographically)?

Monitoring Variable: Resource allocation

Description: Resource allocation can give very important information about distribution of resource allocation per type of allocation (project, preparatory access, programme), per scientific field; geographic distribution of the institution of the principal investigator;

geographic distribution of the institutions of all researchers who collaborate in the project proposed (i.e. principal investigator plus collaborators); types of institutions, i.e. academia or industry, etc.

Main objective: The main objective is to understand who (users academia or industry), how (for testing, scalability or performance development of software applications or projects, where (country, type of institution and in which scientific field) are the PRACE resources being used. Knowing this distribution PRACE might be able to assess the impact of PRACE serving to a particular demand.

Source of data: PRACE AISBL tool for peer review

Impact analysis: Scientific and regarding distribution per country and industry has also an economic and social component.

Main stakeholders: Users, national governments, Members of PRACE AISBL, Council, AC and SSC

Description of impact assessment: The following parameters can be measured:

- **Distribution per allocation type:** The amount of proposals and resources allocated for each type of allocation, i.e. preparatory, project and programme access. These quantitative values are more informative and may only justify the need for these 3 types of access. The following parameters can be measured for each type of access;
- **Geographic distribution:** percentage of resources and amount of proposals proposed and supported by PRACE per country of the PI, and per country of all collaborators in the proposal, for each call. Here the baseline should follow the level of research development for each country. Indicators can be found at EUROSTAT;
- **Types of institutions:** mainly academia and industry of the PI and all collaborators in the project. A good target for industrial participation not only as PIs but also as collaborators will be around 3% for the initial years of activity of PRACE;
- **Distribution per scientific field:** percentage of resources and proposals proposed and supported for each scientific field. Here the target should be more or less the same percentage for each of the main scientific fields identified as being dependent of HPC resources for the progress of research.

These metrics can be performed for a single call and also for a group of calls to infer the time evolution and eventually draw conclusions that can be used to reflect on the usage of the HPC Research Infrastructure. As for all kind of time evolutions caution should be taken regarding the analysis and the conclusions that can be drawn, especially if the time lag is rather short and includes starting up conditions, i.e. the initial calls may show an initial peak of number of applications to the PRACE resources. This effect can be due either to an over-expectation from the initial calls and may result in some proposals not fulfilling the minimum requirements for Tier-0 machines due to low scalability of the codes or not being adequate to the architecture of the machine requested.

These values can be presented in a graphic way, e.g. pie chart. It will be important also to compare this parameter for PRACE with equivalent parameters for national HPC centres of at least the largest countries in Europe.

B.4 Technical specifications of available systems – What services are offered to users?

Monitoring Variable: Technical specifications of systems available through PRACE AISBL and also of other computer systems made available by PRACE Members

Description: Give detailed technical information on the characteristics of the resources made available to users. The technical specifications include the following parameters:

Main objective: Use of the information on PRACE technical specifications to assess the impact of PRACE in terms of its evolution in terms of machine performance and capability.

Source of data: Computer centres of HMs and other PRACE members

Impact analysis: Mainly related to scientific and economic impact related to performance optimisation.

Main stakeholders: Users, computer centres and PRACE BoD

Description of impact assessment: The following parameters should be used:

- **HPC systems:** Name, vendor, model, installation/last upgrade, node description (CPU, GPU, memory), number of nodes, interconnect type, interconnect bandwidth, peak performance, power consumption and share made available through PRACE.
- **Storage by tier:** size, bandwidth
- **Software environment:** compilers, debuggers, libraries, software packages, etc.

It is important to monitor the previous parameters of each computer system to asses the progress of PRACE in terms of position of its machines in the Top 500 list, as well as the evolution of the capability and performance of the machines made available to users by PRACE. PRACE should define an indicator for bi-annual improvement of the total capability of the machines available through PRACE.

B.5 Usage per job submitted to the systems – How are the machines used?

Monitoring Variable: Distribution per job size and duration

Description: Typology of the projects running on PRACE systems by number of cores used per run and job duration.

Main objective: In terms of impact assessment, the objective of the analysis of this variable is to identify the tendency on the type of usage to assess if the users are adapting to the availability of petaflop-systems and exploit the high parallelism made available by PRACE. In terms of efficiency this variable should also be used for deciding on corrective measures if it is considered that the machines are not used in the optimal way.

Source of data: Data acquired by the computing centres of the HMs. This data should be part of the standard data requested from each computing centre. WP6 of PRACE-1IP is working on these technical specifications.

Impact analysis: Scientific and technologic

Main stakeholders: Users, computer centres, SSC if corrective measures need to be extended to the peer review process.

Description of impact assessment: The aim of this indicator is to determine trends in the size and duration of jobs. These trends will in turn show if users' usage of systems is adapting in light of available resources, to make most effective use of the resources.

It must take account of the prevailing scheduling policy on the system. Ideally one might expect to see job size grow to meet the limits that are imposed on a given system as a user adapts to it. Furthermore if a user consistently uses the largest permitted job size, it raises the question: should the project or future projects using the same code use a larger system if available?

The duration of a job is a limit of productivity and ideally can be addressed by scaling the code in question up to use more cores to complete the job more quickly. If the duration of a job is long especially if this is relative to the prevailing scheduling limits on a given system,

this is an indication that this code is a candidate for further scaling work in the interest of job turn-around time and end user productivity, assuming that such efforts have not already been exhausted. Conversely if a significant fraction of job durations are very short, of the order of minutes, this can be indicative of an unreliable or misconfigured code which is failing during the start-up phase and thus the code in question would most likely benefit from debugging effort.

This variable must take into account the system architecture in question.

In summary the important analysis in terms of impact assessment is to identify a trend in the usage of the system that indicates the adaptation of users to the PRACE resources.

B.6 Training, seminars, and related statistics – How does PRACE AISBL engage in training?

Monitoring Variable: Training events

Description of the variable: Events organized by PRACE for user support. These can be directly or indirectly organized/sponsored training events worldwide. Support to user communities that use or intend to use the services of PRACE AISBL includes training in HPC programming and related issues in seasonal schools, seminars and workshops, PRACE training portal and most importantly in PRACE Advanced Training Centers (PATCs).

Main objective: Training is essential in realizing the full potential of the PRACE infrastructure. Increasing parallelism and evolving programming models make it more difficult to write scalable codes that efficiently utilize the latest HPC systems. Poorly written programs perform badly and waste resources. Training events should also catalyse the interaction and sharing of HPC-knowledge and tradition between various fields of science and technology. Moreover, training should support the introduction of HPC/eScience into new fields of science and technology. Increased knowledge is finally transferred into economical impact. To make sure that the knowledge transfer is as efficient as possible, training events need to be monitored.

Sources of data:

- Recorded observation (training portal hit rates, course participant statistics etc.)
- Structured questionnaires (feedback from participants)
- Case studies/success stories

Impact analysis: The impact is primarily scientific. Well performing scalable codes allow treating larger and more complex problems and produce results faster, with better accuracy, which gives a competitive advantage to a researcher. A secondary impact is economic and social: better science can lead to discoveries and eventually commercial products which affect everyday life.

Main stakeholders: PRACE users, scientific communities, computing centers, PRACE Advanced Training Centers (PATCs).

Description of the impact assessment process: Monitoring should include:

- Number and volume (hours/days) of training events: Quantitative;
- Number of persons trained in PRACE events and Advanced Training Centers: Quantitative;
- Feedback from participants: Qualitative;

- Feedback forms from participants;
- User activity in Training Portal: Quantitative;
- Number of visits/hits, feedback;
- Training material made available through the PRACE Training portal;
- Amount of material: Quantitative;
- User ranking of material: Qualitative;

All training activities should include a survey regarding the quality and adequacy of training as well as the quality and appropriateness of the materials distributed during the training.

The organisers of the events will compile the data from the event and the corresponding survey available to PRACE AISBL. It will be also important to monitor the amount of downloads of training material and ask visitors to fill in a survey. The BoD collects all info in a yearly report.

The results should be reported with as statistics based on the above quantitative criteria. Qualitative evaluations must be emphasized and conclusions drawn accordingly for further developing the training activity. The reference values for this variable will necessarily need to measure the incremental trend on PRACE training and numbers of interested attendees in order to show a positive impact.

Output Variables

C.1 Publications and success stories – What is the result of the resources allocated by PRACE in terms of publications?

Monitoring variable: Publications of any type (peer reviewed or not), PhD theses, success stories

Description of the variable: number of (peer reviewed) papers, academic theses, publications related to conferences, scientific success stories, any type of publications (including reports) resulting from contracts with industry, government or organization, etc.

Main Objective: The usage of HPC resources became critical in any field of science. The possibility to solve mathematically intricate problems allows for the consideration of more realistic (and complex) models describing real systems, therefore boosting progress in sciences with consequences in economy and society. Scientific publications are the way to report scientific discoveries and the measurement of its quantity and quality is also an indicator of the usefulness of the scientific tools, including supercomputers, used in the research process. One expects that the usage of PRACE resources leads to publications in the top journals in each area

Source of Data:

Final reports of applicants and scientific databases (such as ISI Web of Knowledge by Thomson Reuters)

To facilitate the search in data bases, all articles related to PRACE should compulsory contain an acknowledgement to PRACE AISBL in a way to be defined by the organization (via a User Agreement). This should also play the role of an appropriate identifier in any search for PRACE related publications in databases.

Main stakeholders:

- PRACE users
- Entities related to PRACE
- Funding agencies of PRACE HPC centers.

Description of the impact assessment process:

The variable has several parameters, each being concretized by an integer number. The parameters are:

- papers in top class journals (e.g. Nature, Science)
- papers the top 10% journals of the area
- Papers in refereed indexed journals
- Other papers
- books
- chapters of books
- PhD theses completed
- conference proceedings
- extended abstracts of conference presentations
- reports of industry contracts and government/organization research contracts
- any other publication resulting from scientific research
    - success stories
- The value for each of the above mentioned parameters should be provided by the PI of a granted project in the final report
- All are quantitative parameters; success stories are also qualitative
- The impact is mainly scientific
- Reference values: strongly dependent on the scientific area of the proposals. A list with reference values should be produced for the main scientific areas (Physics, Chemistry, Engineering, Earth Sciences, Life Sciences, Health Sciences, etc.). Regarding the first two parameters, one should expect three papers, irrespective of the area.
- Assessment process: The person assigned to the general gathering of information should carry on the task of getting the most relevant data from the final scientific reports or from databases.
- Results should be presented in spreadsheets together with other information that might contribute to characterize the science produced with PRACE tools, and provide information on its quality. The increase of the groups' productivity due to the PRACE supercomputers, if traceable, would give a direct quantitative indication of the scientific impact of PRACE in the different areas.

C.2 Typology of projects regarding additional funding – Are the projects supported by PRACE also being funded by other institutions?

Monitoring Variable: Project finance structure in terms of additional funding or private/public collaborations

Description of the variable: This variable provides information on the engaged different additional funding sources or public/private collaborations during the life cycle of a given project supported by PRACE. For the purposes of PRACE mission and goals, it would be essential to be picked up information how many PRACE granted projects have got support from EC and/or from collaboration with industry. Additional emphasis can be paid on the international partnerships involving parties outside Europe and source of funding. Additionally, an interesting dimension to the variable is regarding EC sources of funding for projects supported by PRACE. This will give a global perception of the funding involved in HPC development at EU level.

Main objective: Having a clear picture about additional funding or private/public collaborations gained by PRACE supported projects is a direct way to observe the impact of European HPC infrastructure on scientific and industrial competitiveness.

The main objective is to follow the strong connections between PRACE allocation process and the ability and chances of the granted projects to accumulate additional funding or create successful private/public collaborations.

Source of data:

Track of PRACE reported data: it can be included in the requested information by PRACE users including members and related organizations input about the project finance structure.

Main stakeholders:

- EC and national public authorities responsible for implementation of HPC policy and programs as well as are highly aware about PRACE activities and their results;
- Scientific and industrial entities that would like to obtain some understandings about success stories and practices;

Description of the impact assessment process:

- Quantitative parameters:
    - Amount of additional funding drawn by institutions in the fulfilment of PRACE granted resources;
    - Number of created public/private partnerships based on the fulfilment of PRACE granted resources;
    - Number of participating scientists and researchers by industrial institutions and non-PRACE members countries;
    - Ratio of amount of additional funding drawn by institutions in the fulfilment of PRACE granted resources and total amount of funding resources.
- Scientific and economic impact
- Reference value: PRACE should set medium to long-term objectives for this variable. Target value should be defined according to the recorded average data for previous period (at least 2 years). For instance: an increase in the total additional funding or private/public collaborations with more than 10 % annually should be placed in the green zone; the annual data in the range from 0 to 10 % should be placed in the yellow zone; and a negative growth is in the red zone;
- Assessment process: should be set based on numbers from previous period. After each PRACE call, the peer review staff produce a short analysis of the outcomes of the call including information of additional funding, trying to establish a project finance structure

regarding funding by other resources (public and/or private). The data can be easily obtained through the system for application to PRACE resources developed by CINES, because applicants should be asked to indicate other types of funding received for the project in the application forms.

– Results: should be presented as a 3 star evaluation according to the reference value.

Environmental Variables

D.1 Financial elements in the annual reports of PRACE members – How are the PRACE AISBL members doing?

Monitoring variable: Financial performance of PRACE members

Description of the variable: Financial data of the PRACE partners for general expenditures and income, including funding and revenues from activities.

Main objective: To provide an indication on the overall financial performance of the association members in order to highlight whether PRACE is impacting the specific activities of its members.

Source of data:

– Official publications of PRACE members and their organizations

– Financial reports of PRACE members and their organizations

– Annual public balance sheet of members

– Input from PRACE members and related organizations

Main stakeholders (in the assessment process): PRACE Members and related organizations

Description of the impact assessment process: The variable is decomposed into five parameters whose values will be derived form the sources mentioned above and their level of accuracy could vary depending on the availability of information. However despite the difference of precision measurement of the impact should not be affected since the objective is to show how the trend evolution.

- Consolidated Expenditures for operational activities
  - Quantitative parameter
  - Economic impact
  - Assessment process: the data are actively gathered by means of analysis of the financial documentation and public balance sheets and direct input from members.
  - Report format: a 3 star evaluation according to the reference value and historical trend graph.
- Personnel
  - Quantitative parameter;
  - Economic impact;
  - Assessment process: the data are actively gathered by means of analysis of the financial documentation and balance sheets and direct input from members;
  - Report format: a 3 star evaluation according to the reference value and historical trend graph.
- Investments

  - Quantitative parameter;
  - Economic impact;
  - Assessment process: the data are actively gathered by means of analysis of the financial documentation and balance sheets and direct input from members;
  - Report format: a 3 star evaluation according to the reference value and historical trend graph.
- Funding
  - Quantitative parameter;
  - Economic impact;
  - Assessment process: the data are actively gathered by means of analysis of the financial documentation and balance sheets and direct input from members;
  - Report format: a 3 star evaluation according to the reference value and historical trend graph.
- Income from activities
  - Quantitative parameter;
  - Economic impact;
  - Assessment process: the data are actively gathered by means of analysis of the financial documentation and balance sheets and direct input from members;
  - Report format: historical trend graph.

The reference values for these parameters could be defined accordingly to strategic targets and/or defined by general European and global targets, however it must be noted that not all PRACE members would be able to provide precise and homogeneous information for the previous parameters. Hence, a valid approach for this circumstance would be to assess the impact according to the increasing or decreasing general trends of the values reported.

D.2 Technology transfer – Did the allocations to PRACE AISBL resources result in technology transfer?

Monitoring variable: Patents and spin-offs

Description of the variable: A counter of the patents that are filled in relationship to PRACE (by PRACE, or thanks to PRACE), and a counter of the spin-offs that are created in relationship to PRACE (exploiting PRACE related patents or any other PRACE asset).

Main Objective: One of the means by which PRACE will directly revert into economy is via the transfer of the technology directly or indirectly created with HPC resources and the exploitation of intellectual property. One of the clearest indicators for these two facts is the filing of patents and other spin-offs.

Source of Data:

- Surveyed data (to the PRACE Users as follow up on their usage of PRACE)
- Track of reported data
- Case Studies/success stories

Main stakeholders:

- The PRACE users or entities related to PRACE users who create spin-offs or file patents will be reporting data to this variable through periodic surveys
- The computer centers that are more in touch with the PRACE users will help to identify case studies to further analyze and figure out the possible relationship of PRACE in the creation of spin-offs or patents

Description of the impact assessment process:

The variable is decomposed in two parameters:

- Number of PRACE related Patents:
  - Description: refers to the patents or any other type of IPR that have been registered for which at some point of the development process, PRACE resources have been used (economical, computational or human);
  - Type of parameter: quantitative;
  - Type of impact assessed: economic impact;
  - Reference value: Will be set according to the strategic objectives of the institution. It can be assumed that 1 patents would be a desirable outcome during the first five years, and increasingly more;
  - Assessment process: The person assigned to the general gathering of information should carry on the task of getting the data from surveys, and also conduct the identification of cases subject to study. This person should track the number of patents found and keep the necessary references to them;
  - Results should be presented as a 3 star evaluation according to the reference value. Accessory information could be provided in the format of an incremental graph over time.
- Number of PRACE related spin-offs:
  - Description: refers to the spin-offs/companies or any other business entitity that have been created, and that has a relationship with PRACE in any of the following manners: the activity involves usage of PRACE, the entity has been supported by PRACE or that the entity exploits a PRACE related IPR.
  - Type of parameter: quantitative parameter
  - Type of impact assessed: economic impact
  - Reference value: Will be set according to the strategic objectives of the institution. It can be assumed that one spin off would be a desirable outcome during the first five years, and increasingly more.
  - Assessment process: The person assigned to the general gathering of information should carry on the task of getting the data from surveys, and also conduct the identification of cases subject to study. This person should track the number of patents found and keep the necessary references to them.
  - Results should be presented as a 3 star evaluation according to the reference value. Accessory information could be provided in the format of an incremental graph over time.

D.3 Relationships with other European Institutions – How is PRACE AISBL collaborating with other institutions?

Monitoring Variable: Relationships with other Institutions and Projects

Description of the variable: This variable refers to all relationships and/or collaborations with other European, non-European and International organizations, initiatives and projects. It would be looked at institutions, initiatives and projects that require HPC resources as well as organisations representing scientific and research fields and/or are dependent of HPC usage for research development.

Main objective: The analysis of the data captured in this variable will provide an indication on the impact that PRACE has on the overall development and coherence of the European HPC ecosystem.

Source of data:

- Track of PRACE annual reported data
- PRACE Press releases;
- EU announcement of FPs Calls results

Main stakeholders:

- PRACE AISBL – BoD and Council;
- Public authorities both EU and national responsible for implementation of HPC policy and programs;
- PRACE members and related organizations;
- Research Institutions and communities that would like to establish relationships with PRACE AISBL.

Description of the impact assessment process:

- Quantitative parameters:
  - Number of established relationships between PRACE ASIBL and/or groups of PRACE members with other European, non-European and International Institutions and Projects;
  - Amount of EU funding dedicated for joint partnerships, initiatives and projects between PRACE AISBL or groups of PRACE members with other European, non-European and International Institutions and Projects;
  - Number of scientists and researchers by PRACE or its members involved in joint partnerships, initiatives and projects with other European, non-European and International Institutions and Projects;
  - Amount of the provided PRACE HPC resources to other European, non-European and International Institutions and Projects.
- Scientific and economic impact
- Reference value: PRACE should set medium to long term objectives for this variable. Target value is according to PRACE Collaboration Plan and regularly opened Calls under European Framework Programs;
- Assessment process: should be set based on numbers from previous period in the range from 2 to 3 years. It should be taken into consideration that most partnerships usually start up in an ad-hoc way and in some cases through personal contacts, therefore, it is important that partnerships reach maturity and become established at the level of PRACE or members' institutions with clear and defined processes and objectives commonly

accepted. It is advisable that the BoD will analyse all partnerships to be established by PRACE.

– Results: should be presented as a 3 star evaluation according to the reference value.

D4. HPC related jobs – Is PRACE AISBL fostering the creation of new jobs?

Monitoring variable: HPC related Jobs

Description of the variable: The number of jobs in the area of HPC in countries associated to PRACE and all related trends measured on an annual basis. The measurement (both absolute numerical values and yearly trends) should be arranged according to industrial sector type, geographical area, job-level and contract type.

Main objective: The analysis of this variable will provide insights on the PRACE impact showing whether the Europeans are gaining skills related to HPC and whether they are able to utilise those skills in the job market. It will also show the strength of the companies of the European HPC sector and their ability to both create jobs (and thus add value to the economy) and to attract HPC talent. This variable is also an indirect measure of the size of the European HPC market and sector. A growth in the number of jobs will tell the reader that European HPC is growing, while, for example, an increase in HPS Sales positions at non-European companies only will indicate that European HPC is not generating enough economic value.

Source of data:

– EU Employment, Social Affairs and Inclusion Statistics and Analysis Reports

– European Centre for the Development and Vocational Training (CEDEFOP)

– Sector studies available

– Sector studies, sample surveys and structured interviews commissioned for the purpose of this measurement

– Internet job portals (those featuring IT jobs)

– Case studies based on job fairs, market watch publications, press releases, etc.

Impact analysis:

Social impact – the number of HPC jobs indicates how HPC contributes to the welfare of the Europeans

Economic impact – the number of HPC jobs mirrors the competitiveness of the European HPC sector and its ability to create value to the economy

Main stakeholders (in the assessment process):

These data should be shared with and analysed by:

– EU industrial development resources – in order to adjust their industrial and skill development policies

– Governments at various levels – in order to assess their HPC sectors, job markets, skill bases and associated policies

– Industrial Development Organisations– in order to create policies supporting the growth of the HPC sector

– Employment, Educational and Skill Development Organisations – in order to cater for the development of HPC skills required

Description of the impact assessment process:

A. Absolute values – these values are volume-based and can serve as an illustration of the behaviour of the HPC job market e.g. to assess HPC's contribution to the European economy. They can also be applied in order to calculate the trends related to the various parameters.

a. Number of jobs by Skill Required
b. Number of jobs by Industrial Sector
c. Number of jobs by Geographical Area (country, region)
d. Number of jobs by Job Level (management, technical, etc.)
e. Number of jobs by Contract (tenure) Type (permanent, temporary, etc.)

- Measuring Europe's performance:

In general, these data will show how PRACE affects the development of high-tech skills in Europe. PRACE is an instrument of industrial development in the HPC area in Europe and its impact manifests itself through the presence of HPC-related jobs, which in turn have an affect on a number of other economic and social areas.

As an example, such a measurement will help provide a conclusion on which industrial sectors benefit from the implementation of PRACE, or which industries lack skills in order to develop their HPC competitiveness, possibly through PRACE-related initiatives. It will also help determine where industrial 'clusters' of competitiveness are being created using the geographical breakdown suggested above. Furthermore, this measurement will provide information on the value brought by the HPC jobs into the European economy and whether they are high-level, skilled jobs, or low-paid, unsophisticated positions, indirectly indicating which skills need to be addressed. It will also be interesting to know whether the occurrence of such jobs is permanent or temporary phenomenon.

As a reference value, a positive trend in job creation means that PRACE is achieving its desired impact.

These data are best presented using a bar diagram; however, verbal description or other graphical formats (e.g. multi-dimensional bar diagrams, pie-chart in order to show a geographical breakdown).

- Comparison with other regions and industries:

These values should be compared with those of other technologies and industries and also with those of other global regions. Europe's competitiveness should be measured against the progress made by other regions or nations involved in the HPC industry and the metrics above will help assess whether, for example, Europe generates a comparable number of HPC jobs per capita. As another example, it will also be worthwhile to assess whether there is a correlation between the number of jobs created by HPC and the number of (similar level, or related) jobs added by other, related industries.

These comparative data should be shown using bar diagrams, with text comments if needed.

B. Annual trends – these data should be compared with the trends in the overall job market in the various categories. In principle, such a comparison will help assess the performance of the European HPC sector in driving growth across various industrial areas.

a. Change in the Number of jobs by Skill Required
b. Change in Number of jobs by Industrial Sector
c. Change in Number of jobs by Geographical Area (country, region)
d. Change in Number of jobs by Job Level (management, technical, etc.)
e. Change in Number of jobs by Contract (tenure) Type (permanent, temporary, etc.)

- Measuring Europe's performance:

Trends help assess the behaviour of the parameters in question over a period of time and they lead to conclusion whether the growth or decline in the HPC sector. Various conclusions can be drawn from the observations of trends, also through their comparison with other variables. For example, decision makers might be interested whether SMEs create more or less HPC jobs than other types of jobs, or whether in general HPC job market grows faster or slower than the rest of the economy. As yet another example, it might be interesting to know whether the number of skilled HPC positions in increasing (and in what sector).

These data should be presented for a particular year, but also showing the previous three years in order to determine how permanent the growth is. Bar diagrams should be used for this purpose, however, their exact format and the amount of comments should be determined by the combination of trends and variables chosen.

- Comparison with other regions and industries:

It will be interesting to know how Europe's HPC job trends behave in comparison with other regions. For example, decision makers might require a conclusion on whether the growth in European permanent jobs is faster than that in other, competing regions.

The best way to show these data is through a set of bar diagrams.

C. Qualitative data
   a. Evidence based on case studies (e.g. the creation of new jobs through the opening of a new facility/department, the establishment of a new university course, a new research and development initiative, etc.,) – this can be used to illustrate the figures and trends above

In general, this evidence should be used to illustrate the measurements carried out in A and B. Other, independent data can also be included if they illustrate a phenomenon related to the topic of the measurement.

D5. European end-user companies in the HPC area – Is PRACE AISBL influencing European companies?

Monitoring variable: European end-user companies in the HPC area

Description of the variable: This variable presents a map of the European HPC end-user industry by showing what competencies European companies possess in this area.

Main objective: One of PRACE's objectives is the stimulation of industrial development in the HPC area. Its impact thus should also be measured by the presence of European end-user companies in the HPC area. This variable will also serve the purpose of identifying the competency level of the European companies in the area of HPC usage. It will show how many companies Europe has, and of what calibre/value these are, in the various areas of HPC.

Source of data:

- EU, national and regional level industry directories
- Databases maintained by various PRACE members
- Company rankings and other data available from professional media
- Survey and case study data collected during PRACE events

As there is no single source of information that could provide such data, it is recommended that PRACE AISBL develop and maintain its own database with all European companies that possess or are interested in possessing HPC manufacturing or application capabilities.

Impact analysis:

Economic – this variable will show the development of Europe's competitive position in this area

Social – in a way, skill levels can be used to demonstrate the advantages European society receives due to the progress of HPC

Main stakeholders (in the Assessment process):

The results should be reviewed by the organisations responsible for the development of the HPC industry in Europe. Such bodies define priorities that define which areas should require more focus and in which areas Europe should acquire leadership. Target values (e.g. in terms of having a certain skill in Europe) should also be set by those organisations.

Description of the impact assessment process:

The absolute values of this measurement (i.e. numbers) can be used to assess the overall condition of the European HPC end-user community as an impact of the operation of PRACE. It should be ascertained whether the competencies of the end-user ecosystem (as presented in the matrix suggested above) match the general competencies of the European economy. In other words, by targeting the areas where Europe possess competencies not exploited yet by HPC, the region can increase its competitiveness levels. Another example of a conclusion drawn from this measurement could be that of whether Europe is able to provide certain HPC end-user skills (i.e. professionals).

The trends of this measurement can be used to check whether the policies applied develop the end-user skills needed.

It will also be interesting to compare both absolute and trend values with those of other regions and industries.

The results of the measurement should be presented in the form of a map/matrix that shows the various technological areas of HPC usage such as automotive, fluid dynamics, visualisation, oil and gas, etc., together with company-related parameters such as company type, size/volume/turnover/value, location, ownership, etc.

D.6 Software development and industrial applications – How is PRACE AISBL influencing HPC applications?

Monitoring Variable: Software development for scalability development, industrial applications, creation of new collaborations

Description: This variable focus on the description of the existing European software and its development for usage by the European academia and industry.

Main objectives: assess the impact of PRACE in the development of HPC software.

Source of data: The data can be obtained from final reports of the projects supported by PRACE and by direct contacts with users from academia and industry.

Impact analysis: Scientific, economic and social.

Main stakeholders: Users from academia and industry

Description of the Impact Assessment process: The parameter to take into consideration is:

- Mapping of software functionalities per scientific/industrial area

Assessment of software development can be collected from projects (mainly preparatory access projects) allocated by PRACE. This information can be obtained from the reports filled by applicants after completion of those projects. Information on industrial applications can be obtained in a random manner from PRACE members and from projects run on the PRACE infrastructure by conducting market research or by receiving information from applicants in some cases long after the completion of the projects. First of all it is necessary to identify projects with potential industrial applications and follow up on their development long after the allocation of resources to the project. This assessment needs direct contact with the applicants of the projects identified and can be done by the PRACE members of the country of the institution of the applicants. The PRACE members shall collect the \necessary information and will present it to PRACE for further assessment. PRACE management should also have an active role in reviewing reports and monitoring the advances in software in general to complete the mapping.

The evolution of the mapping will permit the qualitative assessment of the influence of PRACE in software development.

D.7 Vendor companies – How is PRACE AISBL influencing vendor companies?

Monitoring variable: Vendor companies

Description of the variable: the variable provides a map of the vendor companies active in Europe. It's a composite variable that shows qualitative and quantitative parameters:

- Geographical distribution (location, number)
- Name, business sector, area of expertise
- Economics (size, turnover, value,)

Main objective: To provide an indicative value of the evolution of the European HPC companies. PRACE will be interested to know trends of the characteristics and the economics of software and hardware companies with some application or relationship with HPC that are active in Europe. The evolution of the specific business sector and area of expertise might provide insights of any impact of PRACE on the market.

Source of data:

- Surveys and analysis of SW and HW vendors from research companies (Gartner, IDC, CXP)
- Official EU statistics
- Official industrial statistics at country level
- PRACE members and related organizations input

Main stakeholders (in the assessment process): PRACE Members and related organizations

Description of the Impact assessment process: The aspects related to the social impact measurement are highlighted by means of the qualitative elements that contribute to indicate the geographical distribution throughout Europe of those companies, their area of expertise and business sector. The variable is composed of the following parameters:

- Number of vendor companies active in HPC related HW/SW
  - quantitative parameter
  - economic impact

  - the reference value could be defined accordingly to targets and forecast produced at EU strategic level for industrial development
  - Assessment process: the data gathering activity will be carried out by means of direct access to survey and analysis documents, by direct access to specialized data base collections.
  - Report format: a 3 star evaluation according to the reference value and associated historical trend graph and table listing name, sector and location of vendors

- Size, turnover, value of vendor companies active in HPC related HW/SW
  - quantitative parameters
  - economic impact
  - Reference value: the reference value could be defined accordingly to targets and forecast produced at EU strategic level for industrial development
  - Assessment process: the data gathering activity will be carried out by means of direct access to survey and analyse documents, direct access to specialized data base collections and further processing is necessary for extracting statistical information. This information should be assessed in terms of temporal evolution.
  - Report format: graphics showing a trends for different dimensions

- Name, location, business sector, area of expertise of vendor companies active in HPC related HW/SW
  - qualitative parameters
  - social and economic impact
  - Assessment process: the data gathering activity will be carried out by means of direct access to survey and analyse documents, direct access to specialized data base collections and further processing is necessary for extracting statistical information. This information should be assessed in terms of the expansion of the presence of companies in different geographic areas.
  - Report format: matrix table and summary descriptive reports

D.8 Investment of industry in HPC – How is PRACE AISBL influencing investment on HPC

Monitoring variable: Investment of industry in HPC

Description of the variable: The trend of HPC investments of industry in terms of the following indicators:
- Value of HPC purchases;
- Value of HPC purchases as percentage of total investment;
- Value of HPC purchases as percentage of investment in R&D.

It will be also interesting to follow the above parameters for the top 10 European R&D spenders.

Main objective: to measure the impact of PRACE on European industry, namely on HPC investments from industry. The monitoring of industry investments will contribute to evaluate whether PRACE is influencing (impacting) the HPC economy.

Source of data:
- Surveys and analysis of industry investments in R&D and HPC

- Official EU statistics
- Official industrial statistics at country level
- News reporting investments in acquisition
- PRACE members and related organizations involved with industry

Main stakeholders (in the assessment process):
- Members of PRACE
- Industrial organizations related to PRACE
- IAC (once established)

Description of the impact assessment process:
The variable is decomposed in the following three parameters that provide the economic dimension of the investments related to HPC only and as percentage of R&D industrial investment as well as the total investment.

- Value of HPC purchases
  - Quantitative parameter
  - Economic impact
  - The reference value could be defined accordingly to targets and forecast produced at EU strategic level for industrial development
  - Assessment process: data can be gathered by means of surveys and analysis published by research companies, by direct access to public data bases available at national and EU level. A further means is direct survey activity. Since the latter is a time consuming and effort intensive activity it should be foreseen for that a later implementation.
  - Report format: a 3 star evaluation according to the reference value, trend graphs
- Value of HPC purchases as percentage of total investment
  - Quantitative parameter
  - Economic impact
  - The reference value could be defined accordingly to targets and forecast produced at EU strategic level for industrial development
  - Assessment process: data can be gathered by means of surveys and analysis published by research companies, by direct access to public data bases available at national and EU level or derived from the Value of HPC purchases parameter.
  - Report format: a 3 star evaluation according to the reference value, trend graphs
- Value of HPC purchases as percentage of investment in R&D
  - Quantitative parameter
  - Economic impact
  - The reference value could be defined accordingly to targets and forecast produced at EU strategic level for industrial development
  - Assessment process: data can be gathered by means of surveys and analysis published by research companies, by direct access to public data bases available at national and EU level or derived from the Value of HPC purchases parameter.
  - Report format: a 3 star evaluation according to the reference value, trend graphs.

D.9. Industry participation in PRACE events – Is the dissemination of PRACE AISBL reaching industry?

Monitoring variable: Industry participation in PRACE events

Description of the variable: This variable shows what companies attend PRACE events (e.g. PRACE Industrial Seminars) compared to the overall structure of the European industry. These data should be shown using classification according to company type, size, value, turnover, geographical area, etc. The measurement should also distinguish between companies producing HPC solutions and those who take advantage of HPC as a tool.

Main objective: One of PRACE's desirable impact objectives is the dissemination of knowledge on HPC to the industry. This variable's aim is to assess whether PRACE events achieve the desired effect of attracting industrial organisations and transferring knowledge to them.

Source of data:

Data (surveys, interviews, meetings and discussions) collected during PRACE events

Impact analysis:

Economic – this variable helps to assess the industry's awareness of PRACE's work and provides an indication of the industry's skill levels.

Main stakeholders (in the Assessment process):

- EU industrial development organisations – in order to assess the maturity of the HPC industry
- National and regional governments – in order to assess the maturity of their HPC industrial base

Description of the Impact Assessment process:

A. Absolute values
    a. Number of companies by industrial sector
    b. Number of companies by geographical area
    c. Number of companies by size/value/turnover
    d. Number of companies by company type (corporate, SME, etc.)
B. Annual trends (measured between events)
    a. Change in Number of companies by industrial sector
    b. Change in Number of companies by geographical area
    c. Change in Number of companies by size/value/turnover
    d. Change in Number of companies by company type (corporate, SME, etc.)
C. Qualitative data
    a. Evidence based on case studies (e.g. presentation of new results, new applications, etc.)

These data should be assessed against the objectives set for the seminars. While in general, industry participation is a desirable outcome, it should also be determined whether the seminar programme is able to attract the targeted groups of participants, e.g. SMEs or software companies.

The absolute values and trends of this measurement should be correlated with other variables (e.g. the number of jobs, the number of vendors or end-users) in order to assess the general strength of the European HPC industry. For example, an increase in the number of companies participating in the PRACE events together with an increase of HPC jobs and HPC vendors would mean an improving position of the HPC industry.

As a reference, it should be assumed that an increasing number of companies participating means that PRACE is fulfilling its mission in this area. It is important to note that the companies participating should be of a profile matching the events' requirements and objectives (if such has been set, e.g. targeting a specific user group).

The data in A and B should be shown using bar diagrams accompanies by verbal comments including case studies. There should be an annual report on seminars together with other data such as case studies.

D.10 PRACE awareness – Is PRACE AISBL creating awareness of its mission and services?

Monitoring variable: PRACE raising awareness events and media coverage

Description of the variable: PRACE raising awareness events and media coverage is a synthetic variable that combines in itself 3 major groups of quantitative variables:

- Organization of special PRACE events: industrial events; scientific seminars/symposia; PRACE training events; collaboration activities with non-PRACE members and other RIs;
- Participation in international exhibitions and other events (for instance: ISC, SC, European ICT events);
- Public audience reached by the following channels: audio and video broadcasting; web streaming; printed materials (news in the press and other news channels, press releases) and give-aways.

PRACE success depends also on attracting the attention and interest of the public authorities and of different scientific and industrial communities that actively use and support the PRACE development.

Main objective: The visibility of the PRACE to both scientific communities and general public is a step leading to the creation of an informed overall perception and an increase and coherence understanding for the role and following impact of HPC on the science, society and economy.

In this sense, the main objective of this variable is to evaluate the implemented PRACE platform for dissemination and explanation of important HPC achievements including results obtained by PRACE allocated resources based on a set of channels utilized for reaching a wild array of interested parties.

Source of data:

Track of PRACE reported data

Main stakeholders:

- The PRACE users or entities that heavily rely on HPC in their working processes;
- Science and industrial users and community groups both in Europe and outside Europe that can become future potential users or partners of PRACE;
- EC and national public authorities responsible for implementation of HPC policy and are highly aware about PRACE activities and their results;
- General public highly interested in the implementation and development of European HPC ecosystem.

Description of the Impact assessment process:

The variable is decomposed in three subgroups including a set of parameters:

- Organization of special PRACE events:
  - Quantitative parameters:
    - Number of organized PRACE events – for instance: industrial seminars, scientific workshops, seasonal and other training events;
    - Number of attendees to the PRACE events – for instance: number of companies by industrial sector; number of companies by size/value/turnover; annual trends (measured between events)
  - Scientific and economic impact
  - Reference value: PRACE should set short to medium term objectives for this variable. Target value is according to the defined list of events included in the PRACE AISBL Dissemination Plan in collaboration with the PRACE AISBL Members;
  - Assessment process: should be set based on the trend analysis of the numbers from previous period.
  - Results: should be presented as a 3 star evaluation according to the reference value.
- Participation in international exhibitions and other events:
  - Quantitative parameters:
    - Number of PRACE presentations at international events.
  - Scientific impact
  - Reference value: PRACE should set short to medium term objectives for this variable. Target value is according to the PRACE AISBL Dissemination Plan;
  - Assessment process: should be set based on numbers from previous period. The overall assessment process should include:
    - Based on concrete pre-defined quantitative parameters preparing a report by the BoD to the Council (yearly) and final information published on yearly basis.
  - Results: should be presented as a 3 star evaluation according to the reference value.
- Public audience reached by different media channels:
  - Quantitative parameters:
    - Number of broadcasted PRACE video images on TV and radio interviews;
    - Number of web streaming events open to the general public- target value is according to the PRACE AISBL Dissemination Plan;
    - Number of press cuttings uploaded on PRACE media channels - this parameter depends on many specific circumstances and could not be targeted.
    - Annual budget for PRACE events;
    - Annual budget of PRACE for participation in international exhibitions and other events.
  - Science, economic and social impacts
  - Reference value: PRACE should set short to medium term objectives for this variable. Target values are according to the PRACE AISBL Dissemination Plan. For instance:

an increase in the total PRACE raising awareness events and media coverage_with more than 6% annually should be placed in the green zone; the annual results in the range from equals + 6% to equals – 6% should be placed in the yellow zone; and a decrease with more than – 6 % is in the red zone;

- Assessment process: should be set based on numbers from previous period and general investigation on practices implemented by similar infrastructures. The overall assessment process should include:
  - Requesting information from every PRACE Member (every 3 months) regarding PRACE media coverage. This should be done by the BoD;
  - Preparing a report by the BoD to the Council (yearly) and final information published on yearly basis.
- Results: should be presented as a 3 star evaluation according to the reference value.

D.11 Social and Economic events – Is PRACE AISBL having socio-economic impact?

Monitoring variable: Social and Economic Events

Description of the variable: Identification of key social and economic events (inventions, economic turns, technology trends, etc.) that might be related to the HPC industry in general and with PRACE in particular.

Main objective: Usage of HPC in science and industry eventually has a direct impact in economy or society by permitting or by speeding up the discovery of new materials, design of better tools, discovering drugs, explaining phenomena or improving processes. These outcomes, usually covered in the mass media, emerge in the last stage of the research process, usually much later than the stages where HPC might have been involved. Hence this involvement might be concealed. The objective of this analysis is to assess the involvement of HPC in general and PRACE in particular on identified news on applications of scientific outcomes or new industrial advances. Other categories of news of economic or social nature might be also considered here, even if they are not related to scientific outcomes, but the assessment team estimate that HPC might have any implication.

Source of data:

Case studies/success stories

Main stakeholders:

- The identification of news to be analysed should be a distributed task. All members should be aware of this and have a permanent monitoring task for identifying and communicating relevant events, above all at national level. In selecting events for case studies, institutions should focus on those impacts that are more fully developed or significant ‘interim’ impacts.
- PRACE AISBL aside of collecting data, will be also monitoring events
- Assessment committee: if upon a preliminary analysis done by PRACE AISBL some events have been considered for further assessment, an assessment committee should be created. The committee should involve persons related or knowledgeable on the event under assessment and personnel related or knowledgeable on the specific infrastructure who provided access to HPC resources involved in the event.

Description of the impact assessment process[1]:

The variable has just one single input parameter: Identified events of high impact R&D outcomes or other social or economic events:

Description: This parameter registers the events identified for pre-assessment. The events reported should come documented with indicative reasons of why this event should be considered for assessment. Events should be related ideally to PRACE but if an HPC institution non-member of PRACE was involved, then it should be also reported.

Type of parameter: qualitative parameter

Type of Impact: mainly social, but also high impact cases could be found on economic, scientific or environmental areas

Reference Value: PRACE should set long term objectives for this variable. The existence of PRACE is only justified to make a high impact in society. However the time cycle of R&D until it reaches society could be very long (10 to 25 years). So the infrastructure could expect the first high impact results in the minimum scope of 5 years, mainly from other HPC infrastructures, and maybe some interim results from PRACE. The emergence of these events should continue from that point on, and any positive evaluation should be highly considered in the general impact evaluation for PRACE, however the evaluation should not be penalized for gaps of up to 2 years without any high impact event registered. A good indicator of success for PRACE after 10 years would be having at least one event rated with four stars (according the rating classification explained in Section 3.2).

The registering of high impact events where a non-PRACE HPC institution is involved should be used as benchmark, so that if PRACE is registering a higher rate of high impact events than other similar infrastructures in other continents, the conclusion should be positive, and inversely, if there are more high impact events where other infrastructures have been involved, then PRACE should consider the review of its strategy.

Assessment process: Described in Section 3.2

D.12 Environmental Impact – What is the impact of PRACE AISBL in the environment?

Monitoring variable: Ecological imprint

Description of the variable: The main directly and objectively measurable value described in this variable is power consumed by the computer systems concerned, both for computation and for cooling

Main objective: To assess the Ecological imprint of the PRACE infrastructure in terms of energy and efficiency gains.

Source of data:

- Recorded observation, electricity usage and practices as recorded by computer centres
- Recorded observation, reports and publications of any relevant research performed on the infrastructure

Main stakeholders (in the assessment process):

- Computer centres

[1] See also section 3.2

Description of the Impact assessment process:

The following parameters describe the metrics that will be collected twice yearly by computer centres and analyzed by the AISBL. All but one parameter are quantitative in configuration and nature and describe the energy usage of a system, which in turn reflects its environmental impact. The exception being the cost of the electricity per kWh, which is an economic impact parameter. This value will vary from one region to another.

Parameter description	
Centre	Name of the relevant centre
System name	Name of the system for clarity over time an in the case of centres which host multiple systems
Model	Basic description of the system, manufacturer model number etc.
Power usage	Typical power usage when loaded to the level given in the utilization profile
Cooling power usage	Typical power used for system cooling, where seasonally variable provide a figure adjusted for a complete year.
UPS or other electrical power overhead	Typical power overhead due to UPS or similar losses within the centre
Electricity cost	Electrical costs
Performance (Linpack)	Performance of the system as measure by the Linpack benchmark
Performance (peak)	Peak theoretical possible of the system based on manufacturer's specifications
Water requirement	Volume of water required for cooling, its source, typical operating temperatures and if it is a closed system
Utilisation profile	% of the total system which is typically in full use
Availability	% of time for which the system is operational
Hosting environment	Description of the physical hosting environment
Exception report	Any circumstance or deviation that is deemed relevant to these parameters
External reference(s)	Link or reference to an external source(s) of more detailed information regarding the system

It should be noted that there is little consistency regarding the issue of reporting ancillary power usage i.e. power used for cooling and provision of an uninterruptable power supply.

Thus it should be made clear when presenting or requesting data what a given power figure represents.

Climate change as a result of CO_2 production is frequently cited as the primary environmental impact of electricity production. However one cannot attribute a specific level of change to the production to levels of power concerned thus the power figures are themselves quoted, facilitating accurate comparison over time and with other areas of human activity. It is also important to analyse the "greenness" of the energy used by the different sites. This can be inferred from the energy certificates of the different sites.

PRACE aims to promote a high level of efficiency in the infrastructure. The results presented permit the assessment of system and hosting facility efficiency. The best figures in each case serve as the target for systems. It must be recognized that achieving the best figures is a typically a commercial compromise. The AISBL will periodically derive best in class benchmark figures for equivalent infrastructure outside of PRACE to be reported alongside results from within PRACE. This permits assessment relative to world-class infrastructure. Significant discrepancies will be assessed and reported upon.

Where scientific results with an environmental dimension are reported they will not be assessed quantitatively, as this is considered elsewhere.

The data is best reported in table format detailing the individual systems and key figures across multiple systems i.e. all systems surveyed of a given class.

The assessment should be made in terms of international benchmarking and trends with historical data.

4 Conclusions and future work

This study delivers a set of variables and related analysis methodologies in order to measure the impact of PRACE across different dimensions. This work builds on the previous research work on general monitoring and reporting methodology. It focuses on the variables related to impact and proposes methods for their analysis.

The document proposes a theoretical framework in order to show the importance of impact assessment in institutions such as PRACE, and provides a classification of various impact-related variables according to their interaction with the organisation and environment (input, delivery, output and environment), assessment type (qualitative, quantitative), impact area (scientific, social, environmental, and economic) and impact timeframe (short-, medium-, long-term).

The general conclusion is that impact assessment for PRACE is necessary and possible. The creation of the necessary framework and automation of the assessment process requires a significant amount of operational and organizational effort; however, this can be implemented according to a well-established methodology and clear procedures for the analysis of each type of information.

The time cycle of the impact of science on society or economy can be very long (10-25 years), thus the assessment of this type of impact of HPC in general, and PRACE in particular, needs to be carried out through qualitative assessment and review of case studies. Provided the importance of keeping track PRACE's involvement in high-impact events, the incorporation of such mechanism is advisable. This work provides guidelines for a mechanism to assess this type of impact.

Following the conclusions provided here, as a next step a pilot assessment of the short-term timeframe variables should be implemented. This will be provided in the next deliverable D2.4.3 ‘Second Impact Assessment of the Research Infrastructure’.

ANNEX A

The following table summarizes the impact variables that are relevant for the current analysis, showing also the classification according to the parameters defined in section 2.1.

Type of Variables	Impact Areas	Implementation Plan	Source of data	Timeframe of the impact assessed
Success ratio of proposals	Mainly scientific impact	Short term	Extracted from the PRACE tool for peer review	Short term timeframe
H-index and G-index of applicants	Mainly scientific impact	Short term	Indexes are available from scientific data bases; Information provided by the PI in the application forms.	Medium term timeframe
Resource allocation	Scientific regarding distribution per country and industry has also an economic and social component	Short term	PRACE AISBL tool for peer review	Medium term timeframe
Technical specifications of available systems		Short term		Medium to Long term timeframe
Distribution per job size and duration	Scientific and technologic impact	Short to medium term	Data acquired by the computing centres of the HMs	Short term timeframe
Training events	Primarily scientific impact; A secondary impact is economic and social	Short term	Recorded observation (training portal hit rates, course participant statistics etc.); Structured questionnaires (feedback from participants); Case studies/success stories	Short to medium term timeframe
Publications of any type (peer reviewed or not), PhD theses, success stories	Mainly scientific impact	Medium term	Final reports of applicants and scientific data bases (such as ISI Web of Knowledge by Thomson Reuteurs)	Medium term timeframe
Project finance structure in terms of additional funding or private/public collaborations	Scientific and economic impact	Medium term	Track of PRACE reported data: it can be included in the requested information by PRACE users including members and related organizations input about the project finance structure	Medium to Long term timeframe
Financial performance of PRACE members	Economic impact	Medium term	Official publications of PRACE members and their organizations Financial reports of PRACE members and their organizations Annual balance sheet of members Input from PRACE members and related organizations	Medium to Long term timeframe
Patents and spin-offs	Economic impact	Medium term	Surveyed data (to the PRACE Users as follow up on their usage of PRACE)	Medium to Long term timeframe

Type of Variables	Impact Areas	Implementation Plan	Source of data	Timeframe of the impact assessed
			Track of reported data Case Studies/success stories	
Relationships with other Institutions and Projects	Scientific and economic impact	Medium/long term	Track of PRACE reported data PRACE Press releases; EU announcement of FPs Calls results	Medium timeframe
HPC related job trend	Social and economic impact	Medium term	EU Employment, Social Affairs and Inclusion Statistics and Analysis Reports European Centre for the Development and Vocational Training (CEDEFOP) Sector studies available Sector studies, sample surveys and structured interviews Internet job portals Case studies	Medium to Long term timeframe
European end-user companies in the HPC area	Social and economic impact	Medium/long term	EU, national and regional level industry directories Databases maintained by various PRACE members Company rankings and other data available from professional media Survey and case study data collected during PRACE events	Medium to Long term timeframe
Software development for scalability increase, industrial applications, creation of new collaborations		Long term		Medium to Long term timeframe
Vendor companies	Economic impact	Medium term	Surveys and analysis of SW and HW vendors from research companies (Gartner, IDC, CXP) Official EU statistics Official industrial statistics at country level PRACE members and related organizations input	Medium to Long term timeframe
Investment of industry in HPC		Medium/long term	Surveys and analysis of industry investments in R&D and HPC; Official EU statistics; Official industrial statistics at country level; News reporting investments in acquisition; PRACE members and related organizations involved with industry	Long term timeframe
Industry participation in PRACE events	Economic impact	Short/medium term	Data (surveys, interviews, meetings and discussions) collected during PRACE events	Medium term timeframe
PRACE raising	Science and	Short term	Track of PRACE reported data	Short to

Type of Variables	Impact Areas	Implementation Plan	Source of data	Timeframe of the impact assessed
awareness events and media coverage	economic impact			Medium term timeframe
Social and Economic events	Mainly social, but also high impact cases could be found on economic, scientific or environmental areas	Long term	Case studies/success stories	Medium to Long term timeframe
Ecological imprint	Environmental impact	Short term	Recorded observation; electricity usage and practices as recorded by computer centres; Recorded observation, reports and publications of any relevant research performed on the infrastructure	Short to Medium term timeframe

Table 1: Monitoring variables and their main characteristics