

# H2020-FETHPC-2014

# **Coordination of the HPC strategy**



# **EXDCI**

# **European eXtreme Data and Computing Initiative**

**Grant Agreement Number: FETHPC-671558** 

# D7.1 Initial methodology and monitoring tool-set

# **Final**

Version: 1.1

Editor: Guy Lonsdale Date: 12.08.2016

# **Project and Deliverable Information Sheet**

<b>EXDCI Project</b>	Project Ref. №: FETHI	PC-671558
	<b>Project Title: European</b>	eXtreme Data and Computing Initiative
	Project Web Site:	

<sup>\* -</sup> PU - Public, as referred to in Commission Decision 2991/844/EC.

# **Document Control Sheet**

	Title: Initial methodology and monitoring tool-set					
Document	<b>ID:</b> D7.1					
	Version: 1.0 Status: Draft					
	Available at: http://www.exdci.eu					
	Software Tool: Microso	ft Word 2013				
	File(s): D7.1.docx	<b>K</b>				
	Edited by: Guy Lonsdale, SCAPOS					
Authorship	<b>Contributors:</b>	Jean-Philippe Nominé, ETP4HPC				
		Maike Gilliot, ETP4HPC				
		Valeria Bartsch, Fraunhofer/ITWM				
		Kai Kruger, Fraunhofer/ITWM				
		Franz-Josef Pfreundt, Fraunhofer/ITWM				
		Guy Lonsdale, SCAPOS				
	Reviewed by:	Yvon Fournier, EDF				
		François Bodin, U. Rennes & EXDCI TB				
	Approved by:	MB/TB				

### **Document Status Sheet**

Version	Date	Status	Comments
0.1	05/07/2016	Draft	Creation of structure
0.2	22/07/2016	Draft	Consolidation of inputs
0.3	25/07/2016	Draft	Editorial improvements
0.4	26/07/2016	Draft	Full text for all sections
1.0	29/07/2016	Draft for internal	Various improvements
		review	and corrections
1.1	12/08/2016	Final for approval	Minor corrections

# **Document Keywords**

<b>Keywords:</b> HPC cPPP, Monitoring, Impact Assessment, KPI, Methodology
--

#### **Copyright notices**

© 2016 EXDCI Consortium Partners. All rights reserved. This document is a project document of the EXDCI project. All contents are reserved by default and may not be disclosed to third parties without the written consent of the EXDCI partners, except as mandated by the European Commission contract FETHPC-671558 for reviewing and dissemination purposes.

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

# **Table of Contents**

Proje	ect ai	nd Deliverable Information Sheet	i
Doci	umen	t Control Sheet	i
Doci	umen	t Status Sheet	ii
Doci	umen	ıt Keywords	ii
Tabl	e of	Contents	.iii
List	of Fi	gures	.iv
List	of Ta	ables	.iv
Refe	rence	es and Applicable Documents	v
List	of A	cronyms and Abbreviations	.vi
Exec	cutive	Summary	1
1	I	ntroduction	2
2	N	Methodology Background	3
	2.1	Balanced scorecard concept	
	2.2	The HPC Ecosystem BSC	4
3	I	nitial methodology elements and tools	4
	3.1	Survey for ETP4HPC Membership	
	3.2	EXDCI Survey for FET-HPC Projects and CoEs	5
	3.3	Industrial Competitiveness and Socio-Economy impact	5
	3.4	Operational Aspects of the Programme	6
	3.5	Management Aspects of the programme	7
4	C	Conclusion and next steps	8
5	A	Annexes	9
	5.1	HPC cPPP current KPIs	9
	5.2	ETP4HPC cPPP Annual Survey 2015	
	5.3	EXDCI survey 2016	
	5.4	Data sources	. 27

# **List of Figures**

Figure 1: Process model for the development of a balanced scorecard
List of Tables
Table 1: Perspectives and Goals of the BSC

# **References and Applicable Documents**

- [1] <a href="https://ec.europa.eu/digital-agenda/en/high-performance-computing-contractual-public-private-partnership-hpc-cppp/">https://ec.europa.eu/digital-agenda/en/high-performance-computing-contractual-public-private-partnership-hpc-cppp/</a>
- [2] Kaplan, Robert S; Norton, D. P. (1992). "The Balanced Scorecard Measures That Drive Performance". Harvard Business Review (January–February): 71–79.
- [3] Performance Measurement with a Balanced Scorecard for Non-University Research Institutes, Kramer and Pfeffer, Journal of Business and Economics, ISSN 2155-7950, USA, February 2014, Volume 5, No. 2, pp. 162-177
- [4] H2020 HPC R&I Activities https://ec.europa.eu/programmes/horizon2020/en/h2020-section/high-performance-computing-hpc
- [5] http://www.etp4hpc.eu/strategic-research-agenda/
- [6] Report prepared by IDC on: High Performance Computing in the EU: Progress on the Implementation of the European HPC strategy

  <a href="https://ec.europa.eu/digital-single-market/en/news/study-high-performance-computing-eu-progress-implementation-european-hpc-strategy-final-report">https://ec.europa.eu/digital-single-market/en/news/study-high-performance-computing-eu-progress-implementation-european-hpc-strategy-final-report</a>
- [7] <a href="http://ec.europa.eu/research/participants/portal/desktop/en/home.html">http://ec.europa.eu/research/participants/portal/desktop/en/home.html</a>
- [8] http://cordis.europa.eu
- [9] PRACE Annual Report 2015, pp 32-38 on Key Performance Indicators <a href="http://www.prace-ri.eu/IMG/pdf/Prace-annual-report2015">http://www.prace-ri.eu/IMG/pdf/Prace-annual-report2015</a> lowres.pdf
- [10] PRACE KPIs http://www.prace-ri.eu/prace-kpi/
- [11] https://www.limesurvey.org/
- [12] https://exdci.eu/activities/questionnaire-coe-fet-hpc

# **List of Acronyms and Abbreviations**

BSC Balanced Scorecard

CoE Centres of Excellence (for Computing Applications)

cPPP contractual Public Private Partnership

DX.Y Deliverable Number X.Y (Number Y of Work Package X)

EC European Commission

EU European Union

EXDCI European eXtreme Data and Computing Initiative

FET Future and Emerging Technologies

FET-HPC HPC component of FET programme in H2020

FP7 Framework Programme 7

H2020 Horizon 2020 – The EC Research and Innovation Programme in

Europe following FP7

HPC High Performance Computing ISV Independent Software Vendor KPI Key Performance Indicator R&D Research and Development R&I Research and Innovation SRA Strategic Research Agenda

WP Work Package

# **Executive Summary**

This report comprises the EXDCI Project Month 12 deliverable D7.1 "Initial methodology and monitoring tool-set". EXDCI WP7 addresses impact monitoring of the H2020 R&I activities linked to the HPC contractual Public Private Partnership strategy: focusing on the development on monitoring methodology and corresponding tools, it also operates the HPC Ecosystem Balanced Scorecard – a central component of the developed methodology that underpins the design of the monitoring tools developed in WP7 - in order to gather and issue data for the monitoring requirements of the HPC contractual Public Private Partnership.

Deliverable D7.1 reports on the motivation for the initial methodology and tool-set and on its realisation. Following an introduction, the subsequent sections explain: the background and balanced scorecard approach; the alignment with the pre-existing progress monitoring guidelines of the HPC contractual Public Private Partnership; the surveys deployed and data sources used; initial feedback relating to the operational use of the tools in gathering data for the 2015 HPC contractual Public Private Partnership report.

#### 1 Introduction

The purpose of EXDCI WP7 is to generate and gather data and create the necessary analysis tools to support informed decision-making in relation to the development of the European HPC Ecosystem and the impact of the R&I activities linked to the HPC cPPP (contractual Public Private Partnership) strategy [1]. The workpackage comprises two tasks, which are aligned with the two central objectives of WP7: providing a set of methodologies and processes to be used in the measurement of Ecosystem development and progress; perform the periodic monitoring of the implementation of the HPC cPPP strategy. The two tasks are:

- Task 7.1: Methodology and establishing data capture procedures and tools
- Task 7.2: Data capture and analysis

This report comprises the first deliverable from Task 7.1 and covers the initial methodology and tool-set for data gathering. A key component of the approach taken in EXDCI WP7 is the use of the balanced scorecard concept (which will be described in Section 2.1). One key output from Task 7.1 will be a template for the operation of an HPC Ecosystem Balanced Scorecard, which will then be operated in Task 7.2 to monitor the development of the European HPC ecosystem and the impact of the actions relating to the HPC cPPP. The output of that monitoring will flow into the regular cPPP meetings, the annual cPPP reports and into the mid-term assessment of the HPC cPPP in 2017.

The close interaction between EXDCI WP7 and the reporting and monitoring activities within the HPC cPPP ensures on the one hand that the WP7 outputs will have a clear impact. On the other hand, it also means that WP7 cannot operate in isolation and without boundary conditions. Thus, we will see that the HPC Ecosystem Balanced Scorecard takes a pragmatic view of building from the pre-existing cPPP KPIs that were defined during the process of establishing the HPC cPPP.

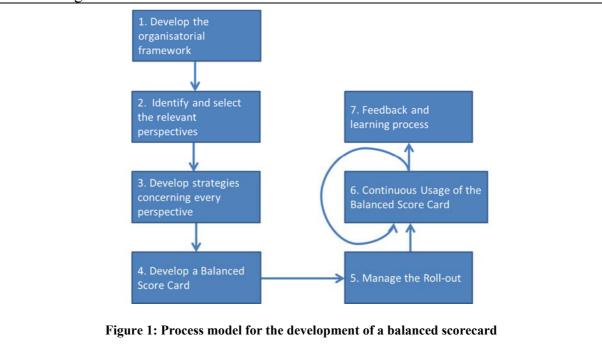
Section 2 of this report summarises the balanced scorecard concept and explains how this was used to structure key perspectives and to incorporate the pre-existing cPPP KPIs (Key Performance Indicators, detailed in Annex 1 to this report). Section 3 reports on the initial toolset arising from the selected methodology, covering surveys developed (included in Annexes 2 and 3) and data sources selected and employed. It also provides feedback on the initial operation, linked to data provision for the 2015 HPC cPPP report (that report itself is not within the scope of this deliverable, and its production is outside the scope of the EXDCI project, but EXDCI WP7 efforts support its elaboration and it will be further related to Deliverable 7.2). Section 4 provides concluding remarks and an outlook on future steps within WP7 Task 7.1.

# 2 Methodology Background

In this Section we provide a summary of the balanced scorecard concept and then the use of the concept, taking into account the needs, and existing contractual guidelines, in creating the first HPC Ecosystem Balance Scorecard. The methodology is then transferred to the tool-set (surveys and selected data sources) described in Section 3.

#### 2.1 Balanced scorecard concept

The balanced scorecard (BSC) [2] is a strategic planning and management system that is used extensively in business and industry, government, and non-profit organizations worldwide to align business activities to the vision and strategy of the organization, improve internal and external communications, and monitor organization performance against strategic goals. It was originally developed by Robert Kaplan (Harvard Business School) and David Norton as a performance measurement framework that added strategic non-financial performance measures to traditional financial metrics to give managers and executives a more 'balanced' view of organizational performance. With the help of the view of perspectives it is ensured that selected goals are pursued. The goals are assigned to a perspective. The perspectives can be chosen so that they fit to the problem. The process model for the development of a balanced scorecard is shown in Figure 1.

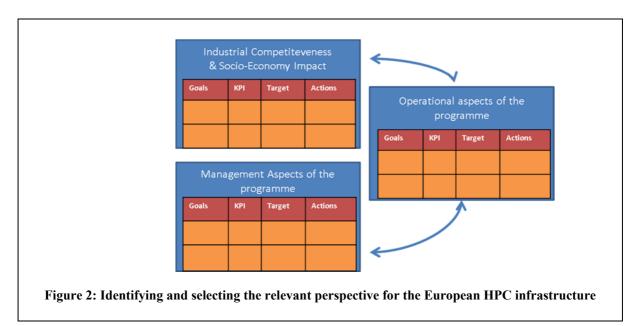


In general for businesses the following perspectives of a BSC are chosen: financial perspective, customer perspective, internal process perspective and learning & development perspective. Since functional goals dominate non-profit organisations, their ideal mission-oriented goal system differs greatly from the prevalent goal system existing at profit-oriented companies. This requires an adaptation of the perspectives of the BSC. Different suggestions to change the perspective have been made. E.g. an adaptation of the BSC with an impact perspective, an enabler and resources perspective, an external stakeholder perspective and a main operating activities perspective was proposed [3].

### 2.2 The HPC Ecosystem BSC

In order to adopt the BSC concept for the purposes of EXDCI and its input for the HPC cPPP, i.e. to create the European HPC Ecosystem BSC, a number of aspects and boundary conditions need to be taken into account. The initial steps needed are the selection or appropriate perspectives and the definition of goals and KPIs per perspective.

In line with the goals of EXDCI WP7, as described above, the overall target is to support informed decision-making in relation to the development of the European HPC Ecosystem and the impact of the R&I activities linked to the HPC cPPP strategy. Thus, the BSC must look beyond the internal view of the EXDCI participating organisations and it must also look beyond the activities of the projects funded under the H2020 HPC Program [4] – the FET-HPC Projects and Centres of Excellence. However, we are able to build on the visions and strategies developed in the ETP4HPC Strategic Research Agenda [5], which links to the multi-annual roadmap of the HPC cPPP. Furthermore, the HPC cPPP performance and impact monitoring guidelines, defined during its establishment, already categorised their set of KPIs (included in the Annex to the report, Section 5.1). We have adopted that categorisation for the definition of the HPC Ecosystem BSC as illustrated in Figure 2. The selected perspectives are Industrial Competitiveness & Socio-Economy Impact, Operational Aspects of the Programme and Management Aspect of the Programme.



In line with the standard procedure for taking existing strategies, substantiating them and then developing goals that are mapped to perspectives, the ETP4HPC SRA visions and strategies combined with the cPPP strategic goals are used to create the goals per BSC perspective presented in Table 1.

Perspective	Goal
Industrial Competitiveness and Socio-Economy Impact	<ul> <li>Increase market share</li> <li>Create innovation environment in HPC (exploited patents and standards)</li> <li>Increase employment</li> </ul>
	Support growth of SMEs
Operational aspects of the programme	<ul> <li>Effective research programme and coverage</li> <li>Develop performance of HPC technologies</li> <li>Provide education, training, skills development</li> <li>Increase use of HPC</li> <li>Develop a HPC software ecosystem</li> <li>Generate patent, inventions and contributions to standards</li> </ul>
Management aspects of the programme	<ul><li>Dissemination and Awareness</li><li>Effective execution</li></ul>

Table 1: Perspectives and Goals of the BSC

In order to be able to understand if a goal measures a target which has already been reached or a change (for example, in an identified company) will in future affect the goals, we need to determine cause and effect. As an example the goals for the perspective "Industrial competitiveness and socio-economy impact" are analysed concerning cause and effect: the creation of an innovation environment and support of growth of SMEs (cause) has the potential to lead to an increase in market share (effect).

Subsequently, and again building on the prior developments of the ETP4HPC SRA and the monitoring guidelines of the HPC cPPP, key performance indicators (KPIs) and corresponding targets are defined for all perspectives. In total 12 KPIs (with sub-KPIs) have been defined in the initial HPC cPPP BSC. These KPIs are presented in Table 2.

### Perspective

### **Key Performance Indicator (KPI)**

1	Industrial Competitiveness and				
	Socio-Economy Impact	Global market share of European HPC			
2	Industrial Competitiveness and	HPC additional investments			
	Socio-Economy Impact				
3	Industrial Competitiveness and	<u>Jobs</u>			
	Socio-Economy Impact				
4	Industrial Competitiveness and	Innovation Environment in HPC			
	Socio-Economy Impact	European HPC start-ups			
		<ul> <li>Number of new SME start-up companies created out of HPC research programmes in the PPP</li> </ul>			
		Unsuccessful HPC start-ups			
		Growth of existing European HPC start-ups			
5	Operational aspects of the	Research programme effectiveness and coverage			
	programme	Coverage of the R&I roadmap by calls topics			
		Number of co-ordinated calls launched			
		Number of responses to calls			
		Number of active research projects			
		Geographical coverage of project participation			
		<ul> <li>Additional leverage and Impact in other related programmes (e.g. areas such as nano-electronics,</li> </ul>			
		photonics, microelectronics, software, storage in other parts of Horizon2020			
6	Operational aspects of the	Performance of HPC technologies developed			
	programme	• Cost per petaflop/s			
		<ul> <li>Cost of ownership (power, space, operation-manpower)</li> </ul>			
		<ul> <li>Percentage of HPC systems with at least 30% of European engineered componentry inside (in</li> </ul>			
Europe and worldwide)					
		<ul> <li>Number of European systems in Top500 and Green 500</li> </ul>			
		<ul> <li>Range of architectures available in Europe</li> </ul>			
<ul> <li>Number of new prototypes made available per year via the PPP</li> </ul>					

# D7.1

# Initial methodology and monitoring tool-set

7	Operational aspects of the programme	<ul> <li>People, education, training and skills development</li> <li>Statistics on number of days of training delivered, range of nationalities and countries of work of course attendees, gender balance of training participants</li> <li>Origin of training participants: SME, large company, academia (graduate, post-graduate, engineers, developers)</li> <li>New curricula and trainings created</li> <li>Quality and pertinence of the trainings through feedback on evaluation forms</li> <li>Number of infrastructure operators jobs in relation to the Programme</li> <li>Increase of the graduate/postgraduate, PhD and post-doctoral positions related to HPC subjects, including, but not limited to, those directly related to the PPP stakeholders projects</li> </ul>
8	Operational aspects of the programme	<ul> <li>HPC use</li> <li>Growth in investment in HPC systems;</li> <li>Growth in the use of external HPC services by SMEs;</li> <li>Growth in the availability of application software on HPC systems</li> <li>Publications crediting the use of the HPC resources (number, citation indicators, ranking of journals, etc.)</li> <li>European research communities using HPC (at Tier-2, Tier-1 and Tier-0 levels)</li> <li>Organisation and participation in multi-disciplinary or cross-disciplinary activities (e.g. number of workshops organised and number of participants)</li> <li>Size and number of structured communities in HPC applications engaging with the PPP (disciplinary and multidisciplinary)</li> <li>Co-design initiatives</li> </ul>
9	Operational aspects of the programme	<ul> <li>HPC Software ecosystem</li> <li>Development of next-generation software codes, libraries and algorithms.</li> <li>Number of application software adapted to work with the next generation of machines</li> <li>Usage of European developed codes, algorithms and libraries</li> <li>Number of European codes benefiting from PPP results: open source applications, system software, ISV codes, etc.</li> <li>Increase in user base of European codes</li> <li>New tools of world-class interest for the research and industry communities (e.g. measured by increase in market share of tools of European origin)</li> </ul>

10	Operational aspects of the programme	Patent, inventions and contributions to standards in HPC by H2020 funded project
11	Management aspects of the programme	<ul> <li>Efficiency, openness and transparency of the PPP Consultation Process</li> <li>number of participants contributing to the strategy and implementation workshops</li> <li>Analysis of ETP4HPC members</li> <li>Monitoring of the decision making process during the consultation</li> </ul>
12	Management aspects of the programme	<u>Dissemination and Awareness</u>

Table 2: Key Performance Indicators for the HPC Ecosystem BSC

# 3 Initial methodology elements and tools

With the BSC defined as discussed above, data sources, data capture tools and subsequent analysis approaches (to map from raw data to appropriate ecosystem estimates or indicators) are required to address the specific KPIs. The main tools developed within WP7 (partly in collaboration with EXDCI WP4) were two surveys issued to: the ETP4HPC membership; the H2020 projects falling under the umbrella of the HPC cPPP, namely the FETHPC projects, the Centres of Excellence and in addition, the FP7 Exa-scale technology projects. More details on the approach taken, and in the case of the former, the specific tool development involved, are given in the following sub-sections. Thereafter, the specific strategy for data capture and subsequent analysis is presented for each perspective of the HPC Ecosystem BSC. Across perspectives, the following external data sources were used: European Commission/DG CONNECT H2020 data and statistics relating to the prior open calls and projects selected [4, 7, 8]; the IDC report on HPC in the EU [6]; PRACE documents relating to its KPIs [9, 10]; other publicly available reports or web-sources, such as, but not limited to, the Top500 and Green500 lists. Annex 5.4 provides an overview of data sources used per KPI.

In addition, certain KPIs are addressed, or also addressed, by specific collaboration between the European Commission and ETP4HPC in the context of the operation of the HPC cPPP, through Partnership Board discussions or in the joint preparation of the HPC cPPP annual report. Those activities are mentioned here, but are outside the scope of EXDCI WP7 activities. Feedback on the participants' experiences with the above-mentioned surveys was obtained via: responses in the comments sections included in the surveys; direct discussions with survey participants, e.g. at the HPC summit in Prague, particularly in connection with the EXDCI workshop (9-10 May 2016) and at ISC (the International Supercomputing Conference, 20-22 June 2016).

#### 3.1 Survey for ETP4HPC Membership

In order to address the KPIs that relate to organisational commitments and contributions to the objectives of the HPC cPPP, it was decided to extend and adapt the 2014 survey sent to the ETP4HPC membership. The individual ETP4HPC member organisations are as members of the association (indirect) stakeholders in the HPC cPPP; the ETP4HPC association is the contractual partner.

The complete survey can be found in the Annex contained in Section 5.2.

The survey method (questions and specific tool for execution) has been selected taking into consideration, in particular, two aspects: the need for a secure system appropriate for collection of sensitive data; data preparation and reporting overheads for the organisation. It was decided, for ease of use for both survey participants and the EXDCI WP7 team, to use a web-based system for the survey. The particular requirements for the choice of web-based tool were:

- Secure data storage (given that sensitive data is collected).
- Possibility for users to save sessions with only partial completion.
- Flexibility in survey output to facilitate subsequent processing/analysis.
- Inclusion of a mechanism to manage/monitor multiple responses from a single organisation.

A number of web-based survey systems were reviewed but rejected (for reasons of security/data protection), since the data storage occurred at the site of the providers (these included Google Form, Survey Monkey and Type Form). The survey tool selected was LimeSurvey [11], a very good tool allowing survey workflow conditional management, easy export in different formats, and extra scripting for further data and graphics processing whenever necessary.

Since the survey asks for business-relevant information, special care has been taken to explain to participants that this sensitive data is handled with care. The survey starts with an introduction in which it is explained that:

While we expect all ETP4HPC members to provide accurate information, only anonymised, collated data and statistics will be disclosed in any subsequent report (Annual cPPP report and some EXDCI deliverables).

The members were also re-assured, via direct email communication, on the care taken to ensure confidentiality of the sensitive data:

"Let us remind you we are keeping data anonymous (on servers located in Germany), only accessible and processed by the core group of 4 people working on the KPIs and cPPP annual report elaboration – and we only deliver global estimates, trends and extrapolations to the EC."

#### 3.2 EXDCI Survey for FET-HPC Projects and CoEs

In order to avoid multiple survey requests being sent to the FET-HPC projects and CoEs within a short space of time, EXDCI workpackages 4 and 7 collaborated on the development of a global EXDCI survey addressing a broad scope of themes and information requests, for example, in order to stimulate cross-project collaboration or to gather feedback on the ETP4HPC SRA [5]. However, many aspects of the survey are related to impact assessment, directly or indirectly (such as identifying innovation via effective or anticipated outcomes of projects, like software development, potential patents filing, SME creation projects) and thus provide an additional tool for the HPC Ecosystem BSC. It should be pointed out, that the survey targets projects as opposed to individual organisations and does not include the types of commercially sensitive questions included in the ETP4HPC survey, discussed above. The survey was also web-based and was built on top of a pre-existing survey realised using SurveyMonkey.

#### 3.3 Industrial Competitiveness and Socio-Economy impact

The perspective "Industrial Competitiveness and Socio-Economy Impact" measures the overall performance of the European HPC market. The primary tool used for addressing this perspective was the survey for the ETP4HPC membership, since the membership covers a good selection of strategic actors in the European HPC Ecosystem (by mid-2016, the ETP4HPC membership had grown to 79 organisations in total). The survey itself includes questions which will support future extension of survey participants (snowball sampling) as discussed in Section 4.

The exception to the use of the ETP4HPC survey was the approach used for KPI 1 (addressing the global market share), for which the external, public data source (the IDC report, [6]) was used.

The sample sizes resulting from the survey preclude the use of formal statistical analysis techniques to relate the survey responses to statistics on the European HPC market; KPIs 2 to 4 target concrete (single number) statistics. However, the calculation of trends is made possible by a comparison of data samples from the two surveys addressing data for 2014 and 2015 (the latter relating to this report).

Key feedback on participant's experiences in handling the survey related to the following items: budget data - both the security/sensitivity & the overhead of calculating/identifying the data within the organisation; the overhead of generating details data relating to both standards activities and patents.

The two aspects mentioned regarding budget data are inter-related; while the sensitivity of data can be an issue, for the larger organisations there may often be challenges arising from inter-departmental communication and separation of responsibilities (e.g. for information security). That the issue is one for the larger organisations can be seen from the participation statistics: whereas only 60% of participating larger organisations answered the budget questions, all participating SMEs did so.

In order that organisations may be able to more easily estimate the overheads of survey completion in advance, it is suggested that a PDF version of the full survey is communicated with the initial request for participation in future surveys.

In order for the effort for providing feedback on standards activities and patents to be reduced, it is also suggested that in future only a reduced level of information on standards (supported by additional standards body searches for working items/themes) and only the number of patents be requested.

#### 3.4 Operational Aspects of the Programme

The next 6 KPIs are connected to the BSC perspective "Operational aspects of the programme". Table 3 presents for each of the KPIs the targeted sample and data acquisition method. For this BSC perspective, we have methods covering the complete sample to be measured. The tools deployed were both EXDCI and ETP4HC surveys, use of open data sources (European Commission sources: participant portal [7] and the CORDIS database [8]; Top500 & Green500 online data), collaboration with PRACE for data exchange.

In contrast to the Industrial Competitiveness and Socio-Economy impact perspective, the responses on the operational aspects of the programme are not sensitive.

KPI	Targeted sample	Method
5	FET-HPC projects and CoEs	Use European Commission data sources
6	European HPC systems	Use open data sources like the TOP500 and
		GREEN500 lists; exchanging data with PRACE
7	Training providers on HPC	Exchanging data with PRACE
8 &	European HPC communities: FET-	Web-based survey (EXDCI survey – c.f. Section 3.2
9	HPC projects and CoEs	& Annex 5.3); exchanging data with PRACE
10	HPC organisations	Web-based survey (ETP4HPC survey – c.f. Section
		3.1 & Annex 5.2)

Table 3: KPIs for Operational aspects of the programme: Targeted samples and data capture methods

We give here some comments relating to the EXDCI survey. While the level of participation is considered high - approx. 80% FP7 exascale projects, FET-HPC projects and CoEs responded - not all of the questions relating to the BSC were answered. This was in particular linked to

information on exploitation planning for project results and on use of the software and (remote) HPC systems (in particular, by SMEs). The fact that most of the projects were in their early development phases at the time of the operation of the survey certainly plays a major role. The limited number of end-user SMEs in the sample (projects) means that not all questions were applicable to individual projects.

#### 3.5 Management Aspects of the programme

The perspective on "Management aspects of the programme" includes 2 KPIs which relate to the internal management of the HPC cPPP (involving the European Commission and ETP4HPC), KPI 11, and to the dissemination activities of projects executed within the programme linked to the HPC cPPP, KPI 12. No specific tools or data acquisition methods are required of EXDCI. The HPC cPPP partners, and ETP4HPC in particular, are responsible directly for KPI 11. For KPI 12, EXDCI collaborates with the FETHPC projects and FP7 exascale projects, CoEs, as well as the EC, and PRACE, and indeed provides common information dissemination opportunities, such as the ISC 2016 exhibit.

### 4 Conclusion and next steps

The monitoring methodology for the current version of the HPC Ecosystems Balance Scorecard has been realised through a set of tools (two separate surveys, various data sources) and the initial operation of the BSC has been used to provide input to the HPC cPPP annual report for 2015. The experiences with the operation of the BSC and feedback from survey participants will be used to drive refinements to the methodology components. Of particular importance will be timely and appropriate reporting for the 2016 report, which will be an important part of the mid-term review of the HPC cPPP as a whole, providing feedback on the first 2 years of cPPP activities and lessons learned. Some specific directions for the BSC refinements can already be identified:

- Additional external resources are expected to be subcontracted to address quantitative socio-economic KPIs.
- Biases in the sampling of surveys could be addressed by statistical methods for the elaboration of quantitative KPIs, and/or, in some cases, cross-checking or correlation with other indicators
- The trend analysis needs to be established using the current BSC results and the initial monitoring data from the 2014 cPPP report as a baseline.
- The targeted sample for the "ETP4HPC survey", which was a key methodology component for the perspective "Industrial Competitiveness and Socio-Economy Impact", will be widened beyond the ETP4HPC membership. Organisations participating in the FET-HPC and CoE consortia and additional HPC organisations known to ETP4HPC members can provide a broader perspective participant pool. Indeed, the current ETP4HPC survey requests information about recently founded startups or spin-offs which is one example for the "snowball sampling" approach to be used.
- Based on response rates to specific areas of the EXDCI survey, adaptation of the question formulation and information requests can be made in order to facilitate increased depth of responses from the FET-HPC projects and CoEs, which will for the next annual survey be much further down their development and innovation paths.

#### 5 Annexes

#### 5.1 **HPC cPPP current KPIs**

Based on the Multi-Annual Roadmap of the HPC cPPP and Partnership Board discussions, the following detailed definition of HPC cPPP KPIs was validated during the cPPP Partnership Board of November 2014.

#### Key Performance Indicators (KPIs) for the HPC PPP

#### Reference:

These performance indicators are based in the section "Expected Impact" of the HPC PPP proposal (annex to the HPC PPP contractual agreement). The data strategy will use different sources of information to establish a solid basis for measurement, to validate the initial values and to stabilize the influencing factors for the measurements in the coming years. Projects selected in Horizon2020 Calls will be asked by the European Commission (EC) to provide specific additional information to the EC (units in charge of the projects) or the ETP4HPC in order to monitor specific indicators and metrics. The ETP4HPC will provide guidelines and suggest ideas on the specific information to be requested to projects. Additional data will be gathered in collaboration with PRACE, with the HPC Centres of Excellence, and other relevant initiatives (e.g. ECSEL, Photonics, etc).

The Partnership Board of the HPC PPP will be the main mechanism for dialogue to monitor the objectives of the PPP. The ETP4HPC will establish a specific working group for monitoring the KPIs ("KPI-wg"). The KPI-wg will investigate mechanisms for anonymous data collection, which is believed to be extremely important if KPI monitoring via data collection is to be effective. It has to be noted that there are issues related to the delivery of data from HPC technology organisations in general but also within the ETP4HPC membership, as many aspects can be considered to be addressing business-sensitive data. While many of the KPIs can be addressed through surveys, the success of these is related to the willingness of individual organisations to provide data and the need for anonymisation of information.

The results of the KPIs monitoring exercise need to be available for the mid-term review (June 2017) and the final review of the HPC PPP (2021) respectively. Regarding HPC market, data will be acquired amongst other sources through the study on HPC Return On Investment (ROI) carried out by the EC, (available in mid-2015). In addition, the ETP4HPC will explore different possibilities to execute some of the data collection and analysis, including two studies performed by independent 3rd parties for the mid-term and final review of the HPC PPP.

#### A. Indicators for Industrial Competitiveness and Socio-Economy Impact

#### • KPI 1: Global market share of European HPC

<u>Metrics:</u> HPC systems, components and tools based on technologies developed and built in Europe, volume (in generated income) of HPC technology exported from Europe (European HPC technology developers) to the rest of the world.

**Baseline:** In 2011, the global market share of Europe was 5%

<u>Targets:</u> Reach a global market share of at least 7.5% by 2017, and 12.5% in 2020 <u>Data strategy:</u> Data to be acquired from studies (EC in 2015, and two studies from ETP4HPC for the mid-term and final review in 2017 and 2021)

#### • KPI 2: HPC additional investments

<u>Metrics:</u> The level of high-tech investment generated by the PPP, and the additional investments leveraged in the HPC value chain; relation to the investments made into European HPC companies by private investors and venture capital funds.

**Baseline:** No baseline

<u>Target:</u> In direct R&I activities, matching the Community funding in the PPP (~700 million€ by 2020). In leveraged investments, four-fold effect in industrial effort per pubic Euro in the PPP.

<u>Data Strategy:</u> Data to be acquired by the ETP4HPC from the studies and companies/organisations participating in HPC research programmes and other organisations in HPC in Europe

#### • KPI 3: Jobs

<u>Metrics:</u> Direct, sustainable jobs out of HPC research programmes recommended by the PPP, and indirect jobs in technology companies further downstream and in end-user organisations of HPC technologies and applications

Baseline: No baseline

*Target:* direct jobs: 400 by 2017, and 1000 by 2020. Indirect jobs, 10000 in technology companies and 100000 in HPC end-users organisations by 2020

<u>Data Strategy:</u> Data to be acquired by the ETP4HPC from the studies and PPP partners and companies/organisations participating in HPC research programmes and other organisations in HPC in Europe

#### • KPI 4: Innovation Environment in HPC

#### Metrics:

- European HPC start-ups (not just those arising from H2020 projects)
- Number of new SME start-up companies created out of HPC research programmes in the PPP (only successful SMEs with a sustainable business )
- Unsuccessful HPC start-ups
- Growth of existing European HPC start-ups

Baseline: No baseline

Target: 4 successful new SMEs in the PPP by 2017, and 10 by 2020

<u>Data Strategy:</u> Data to be obtained from HPC PPP Projects funded under H2020, project final reports, participating HPC vendors, etc. The ETP4HPC will put in place a dedicated workgroup supporting and monitoring the creation of SMEs in the European HPC business.

#### B. Indicators for the operational aspects of the programme

#### • KPI 5: Research programme effectiveness and coverage

Metrics: quality of the research programmes launched

- Coverage of the R&I roadmap by calls topics
- Number of co-ordinated calls launched
- Number of responses to calls
- Number of active research projects
- Geographical coverage of project participation
- Additional leverage and Impact in other related programmes (e.g. areas such as nanoelectronics, photonics, microelectronics, software, storage in other parts of Horizon2020 <u>Data strategy:</u> Data from proposals and projects to be provided by the European Commission Units in charge of the corresponding Horizon2020 Calls, and with ETP4HPC and other EC units regarding the impact in other programmes.

#### • KPI 6: Performance of HPC technologies developed

<u>Metrics:</u> Technological contribution of the initiative to the next generation of HPC in terms of the affordability and usability of the technologies developed in the PPP:

- Cost per petaflop/s
- Cost of ownership (power, space, operation-manpower)
- Percentage of HPC systems with at least 30% of European engineered componentry inside (in Europe and worldwide)
- Number of European systems in Top500 and Green 500
- Range of architectures available in Europe
- Number of new prototypes made available per year via the PPP

*Data Strategy:* Specific ETP4HPC working group.

#### • KPI 7: People, education, training and skills development

 $\underline{\textit{Metrics:}}$  Showing the on the European HPC knowledge base providing High-skilled HPC profiles and curricula developed in the PPP:

- Statistics on number of days of training delivered, range of nationalities and countries of work of course attendees, gender balance of training participants
- Origin of training participants: SME, large company, academia (graduate, post-graduate, engineers, developers)
- New curricula and trainings created
- Quality and pertinence of the trainings through feedback on evaluation forms
- Number of infrastructure operators jobs in relation to the Programme
- Increase of the graduate/postgraduate, PhD and post-doctoral positions related to HPC subjects, including, but not limited to, those directly related to the PPP stakeholders projects <u>Data Strategy:</u> The ETP4HPC will put in place a dedicated workgroup monitoring this KPI.

#### •KPI 8: HPC use

<u>Metrics:</u> Use of the HPC technologies developed in academia and industry (in particular SMEs):

- Growth in investment in HPC systems;
- Growth in the use of external HPC services by SMEs:
- Growth in the availability of application software on HPC systems
- Publications crediting the use of the HPC resources (number, citation indicators, ranking of journals, etc.)
- European research communities using HPC (at Tier-2, Tier-1 and Tier-0 levels)
- Organisation and participation in multi-disciplinary or cross-disciplinary activities (e.g. number of workshops organised and number of participants)

- Size and number of structured communities in HPC applications engaging with the PPP (disciplinary and multidisciplinary)
- Co-design initiatives

<u>Data strategy:</u> Data to be obtained from the studies and the HPC cPPP Projects funded under H2020, PRACE and HPC Centres of Excellence.

#### • KPI 9: HPC Software ecosystem

<u>Metrics:</u> Impact on software ecosystem (number of applications, number of users, etc). Large scale scientific and industrial applications adapted to the next computing generation addressing key economic areas and societal challenges

- Development of next-generation software codes, libraries and algorithms.
- Number of application software adapted to work with the next generation of machines
- Usage of European developed codes, algorithms and libraries
- Number of European codes benefiting from PPP results: open source applications, system software, ISV codes, etc.
- Increase in user base of European codes
- New tools of world-class interest for the research and industry communities (e.g. measured by increase in market share of tools of European origin)

<u>Data strategy:</u> Data to be obtained from the HPC PPP Projects funded under H2020, PRACE and HPC Centres of Excellence.

# • KPI 10: Patent, inventions and contributions to standards in HPC by H2020 funded projects

<u>Metrics:</u> Patent, direct contributions and activities leading to standardisation, and inventionsubmissions out of HPC research programmes recommended by the PPP

*Target:* 40 per year by 2017, 80 per year by 2020

<u>Data Strategy:</u> Data to be obtained from the HPC cPPP Projects funded under H2020

#### C. Indicators for management aspects of the programme

# • KPI 11: Efficiency, openness and transparency of the PPP Consultation Process *Metrics*:

- Monitoring the number of participants contributing to the strategy and implementation workshops
- Analysis of ETP4HPC members to provide evidence for representation of the HPC community
- Monitoring of the decision making process during the consultation

Data Strategy: ETP4HPC will collect and deliver this data on a bi-annual basis.

#### • KPI 12: Dissemination and Awareness

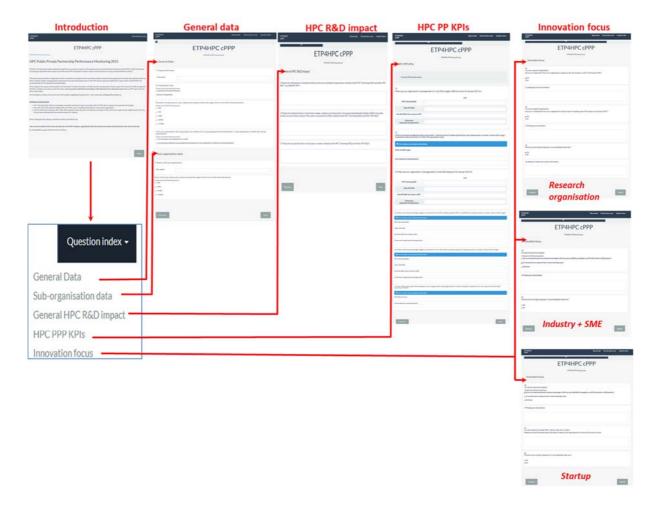
<u>Metrics</u>: Make HPC visible to the general public in Europe and to a broad range of stakeholders. Wide dissemination of information and tangible examples about how HPC solutions contribute to the day to day live of European citizens by using various communication channels like social media, print, video, etc. awareness and information actions held for promoting the PPP activities to a broad range of stakeholders (within and beyond the ones included in ETP4HPC) – this includes events, targeted Newsletters, social media, etc.

<u>Data Strategy:</u> Data to be obtained from the HPC Projects funded under H2020 and from the HPC PPP actors and other stakeholders (ETP4HPC, Centres of Excellence, PRACE, etc.), and to be collected by the European Commission units and the ETP4HPC. The ETP4HPC will be responsible for the dissemination of the information.

#### 5.2 ETP4HPC cPPP Annual Survey 2015

This survey has been prepared Q1 of 2016, so as to go on monitoring private side (research and industry) investments. It is based on the 2014 survey, improving some questions, and adds a specific section on innovation seen from the perspective of different stakeholders (research organisations, industrial companies, SMEs). The questionnaire targets the ETP4HPC members.

While the former sections are quite generic, the latter "Innovation focus" section varies according to the type of organisation (Research; Private company – large industrial company or SME; Startup).



<sup>&</sup>quot;General HPC R&D impact" section deals with patents and standardisation activities.

<sup>&</sup>quot;HPC PPP KPIs" section deals with R&D investments and jobs.

<sup>&</sup>quot;Innovation focus" sections deals with SME dynamics and life cycles (creation, start-up, scale-up, acquisition...) and more generally the acquisition and ceasing of technological activities by larger companies.

#### Introduction

ETP4HPC Exit and clear survey cPPP

#### ETP4HPC cPPP

ETP4HPC cPPP Annual Survey

#### HPC Public Private Partnership Performance Monitoring 2015

ETP4HPC will use the information collected through this survey as part of a report on the progress of the Contractual Public Private Partnership (PPP) for High\_Performance Computing as requested by the European Commission (each PPP is expected to produce a report of this kind and we are using a template defined by the EC).

This impact assessment effort is supported by EXDCI Coordination and Support Action and we will combine outcomes of this questionnaire with other data collected within the EXDCI activities. Besides using global data, stemming from the aforementioned sources, for the cPPP Annual report and related EXDCI reports, EXDCI and ETP4HPC will provide feedback to the respondents and stakeholders.

This is a high priority request and the quality of the information included in the Report will affect the future development of the European HPC Community. While we expect all ETP4HPC members to provide accurate information, only anonymised, collated data and statistics will be disclosed in any subsequent report (Annual cPPP report and some EXDCI deliverables).

Do not he sitate to contact us if you have any further question regarding this questionnaire - main contact: jean-philippe@office.etp4hpc.eu

#### Definitions in this document:

- . HPC Technolody R&D: R&D for technologies covered by the technical research priorities of the ETP4HPC SRA or related and comparable technologies.
- Other HPC R&D: R&D activities relating to the use of HPC, e.g. for modelling and simulation or data analytics applications.
- Non HPC R&D with an impact on HPC: Other R&D activities in areas other than HPC that have an impact on HPC and/oro the results of which might be used in HPC (e.g. microprocessor development that could also impact HPC systems).

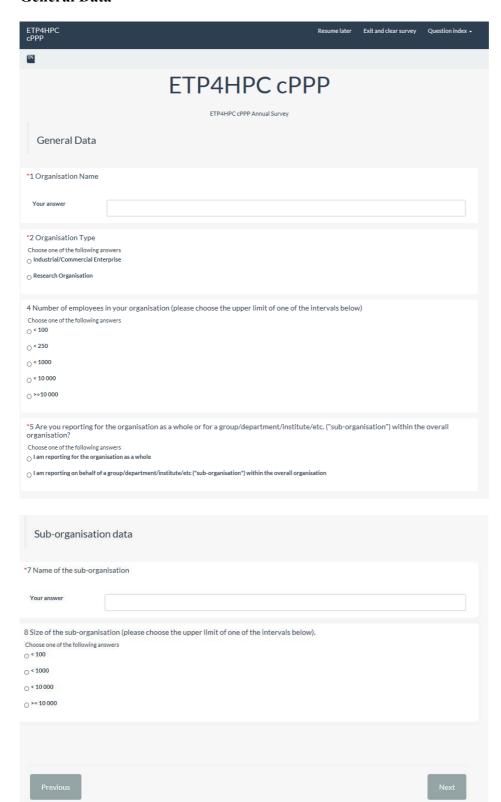
We are looking for data relating to activities carried out IN EUROPE only.

This survey is an update of the one we sent last year to ETP4HPC members, regarding the 2012-2014 period, and answers should therefore refer only to 2015 only.

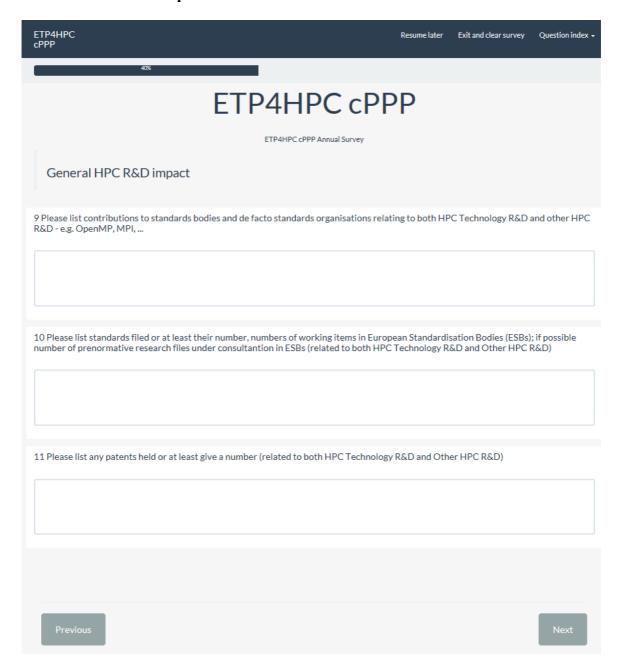
By clicking Next you agree with this terms of conditions.

Next

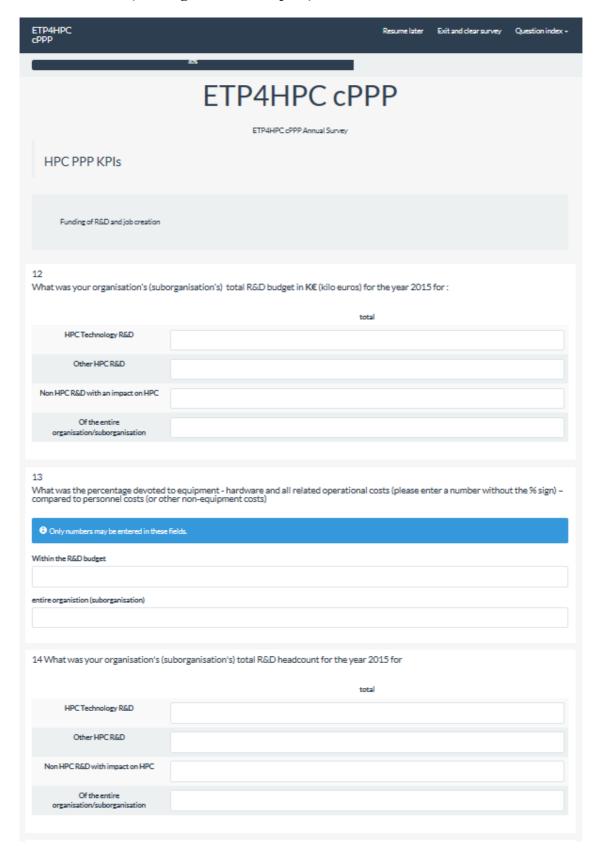
#### **General Data**

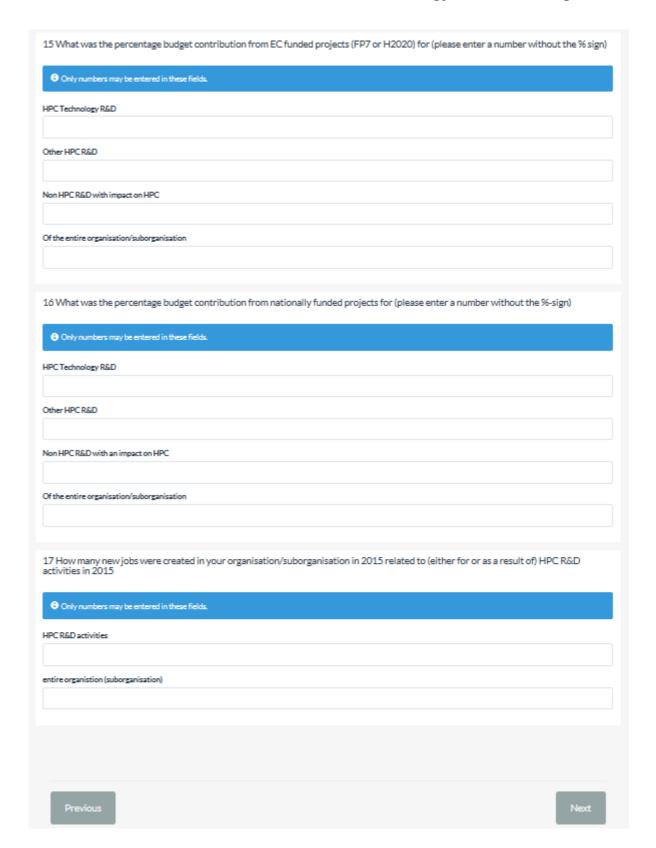


## General HPC R&D impact

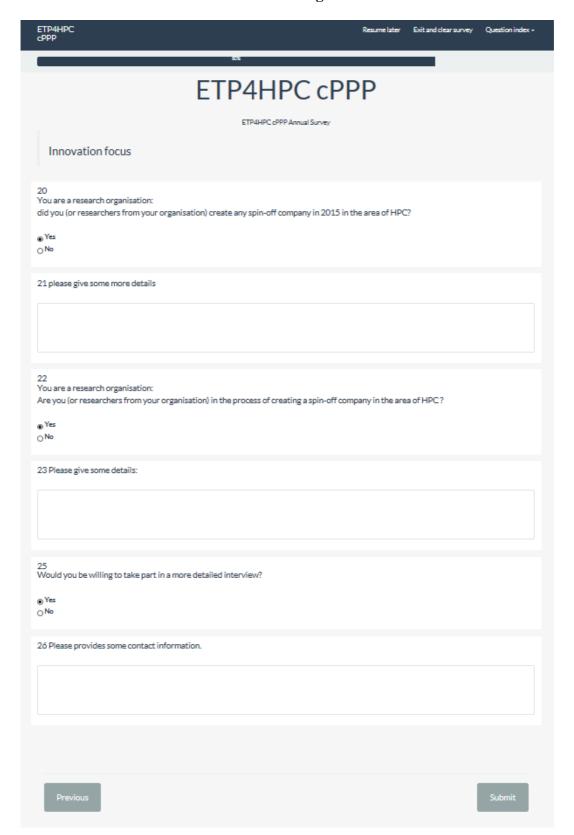


## HPC PPP KPIs (funding of R&D and jobs)

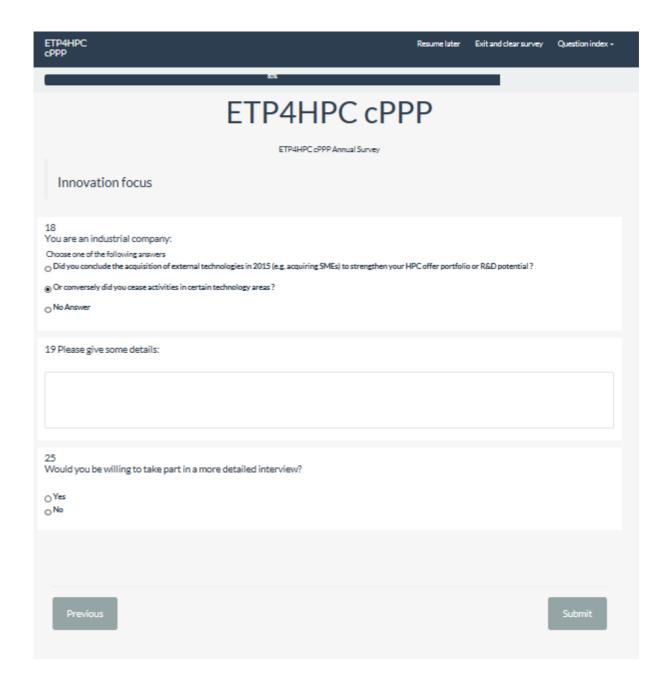




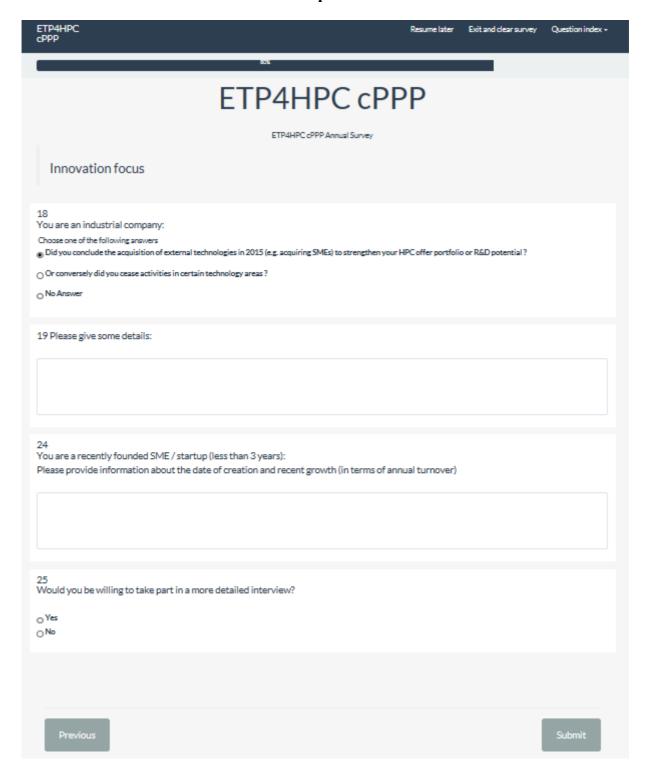
## Innovation focus / Variant 1 - Research Organisation



**Innovation focus / Variant 2 – Industrial company or SME (non startup)** 



## Innovation focus / Variant 2 - SME / startup

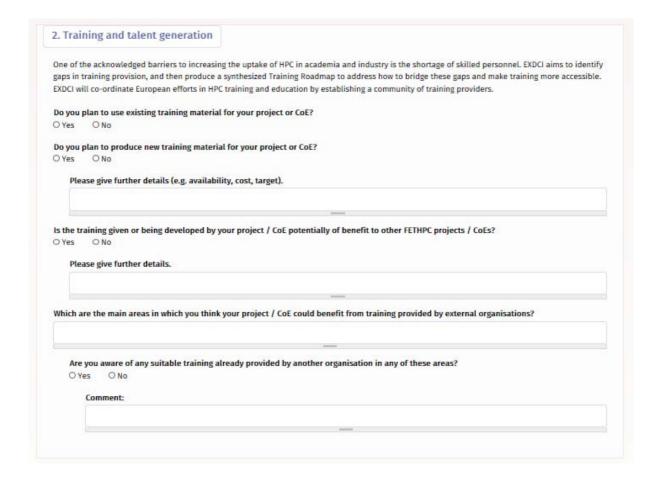


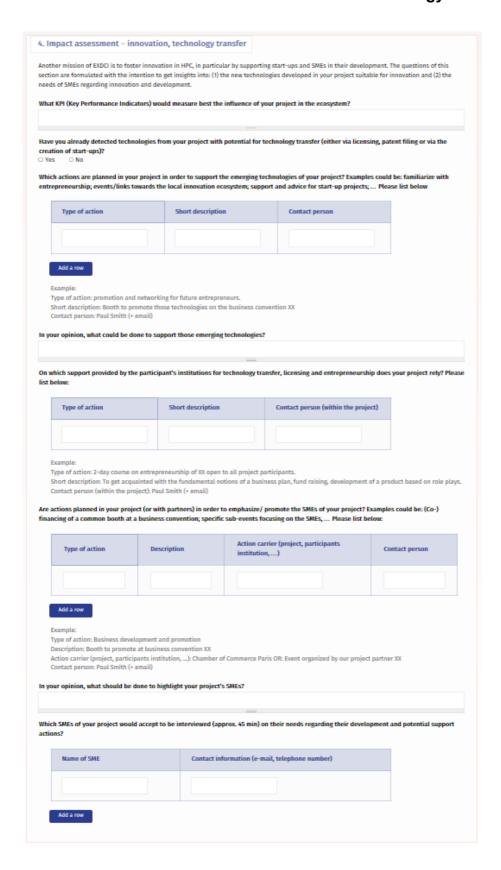
### 5.3 EXDCI survey 2016

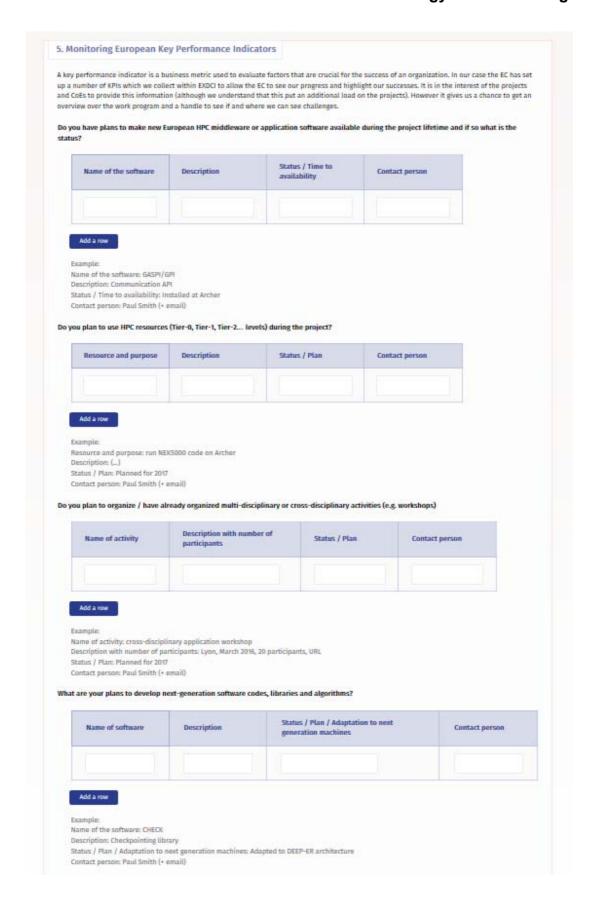
This Annex includes sections of the questionnaire [12] that was been prepared in the first quarter of 2016 under the umbrella of EXDCI (Technical Board, coordinated by the Project Scientific Director). WP7 actively participated in this effort.

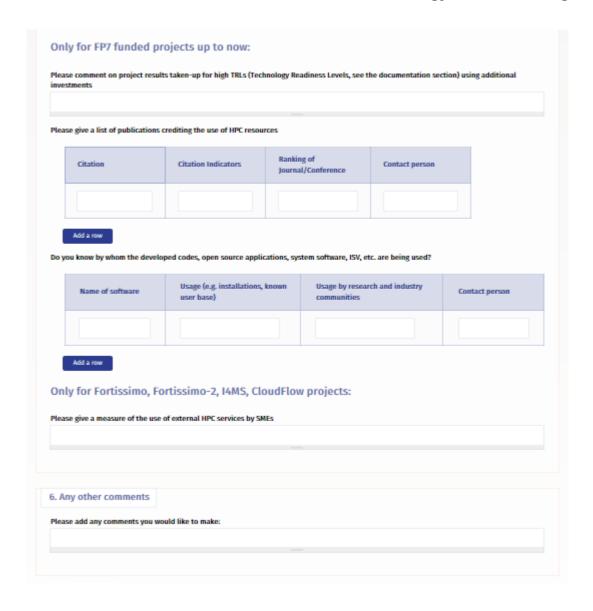
The sections of the survey which have either a direct or indirect relation to the HPC Ecosystem Balanced Scorecard are included below: Sections 2, 4 and 5.

#### D7.1









# 5.4 Data sources

	KPI	Key Performance Indicator (KPI)	EXDCI survey	ETPHPC survey	PRACE KPIs	EC H2020 stats	Analysts' reports	Public sources Web etc.
Industrial	1	Global market share of European HPC					***	*
Competitiveness and	2	HPC additional investments		***				
Socio-Economy	3	Jobs		***			*	
Impact	4	Innovation Environment in HPC: start-ups	**	***				*
	5	Research programme effectiveness and coverage: H2020 calls				***		
	6	Performance of HPC technologies developed	*		*			***
Operational aspects of the programme	7	People, education, training and skills development			***			
of the programme	8	HPC use	**		**			
	9	HPC Software ecosystem	**		**			
	10	Patent, inventions and contributions to standards in HPC by H2020 funded project	*	***				
Management aspects	11	Efficiency, openness and transparency of the PPP Consultation Process				***		
of the programme	12	Dissemination and Awareness	**	**	**			

	Not a data source
*	Complementary source
**	Important source
***	Main source