

## **H2020-FETHPC-2014**

### **Coordination of the HPC strategy**



**EXDCI**

**European eXtreme Data and Computing Initiative**

**Grant Agreement Number: FETHPC-671558**

**D.6.1**

**Report on the BDEC workshop (China)**

*Final version*

Version: 1.5  
Author(s): Mark ASCH, Université de Picardie Jules Verne  
Thierry BIDOT, Neovia Innovation  
Date: 27.04.2017

## Project and Deliverable Information Sheet

<b>EXDCI Project</b>	<b>Project Ref. №: FETHPC-671558</b>	
	<b>Project Title: European eXtreme Data and Computing Initiative</b>	
	<b>Project Web Site: <a href="http://www.exdci.eu">http://www.exdci.eu</a></b>	
	<b>Deliverable ID: D6.1</b>	
	<b>Deliverable Nature: Report</b>	
	<b>Dissemination Level:</b> PU	<b>Contractual Date of Delivery:</b> 28 /04/ 2017
		<b>Actual Date of Delivery:</b> 27 / April / 2017
<b>EC Project Officer: Evangelia MARKIDOU</b>		

\* - The dissemination level are indicated as follows: **PU** – Public, **CO** – Confidential, only for members of the consortium (including the Commission Services) **CL** – Classified, as referred to in Commission Decision 2991/844/EC.

## Document Control Sheet

<b>Document</b>	<b>Title: Report on the BDEC Workshops (China)</b>	
	<b>ID: D6.1</b>	
	<b>Version: 1.5</b>	<b>Status: Final version</b>
	<b>Available at: <a href="http://www.exdci.eu">http://www.exdci.eu</a></b>	
	<b>Software Tool: Microsoft Word 2016</b>	
	<b>Files: EXDCI_D6.1_Report on the BDEC Workshops (CN)_V1.5.docx</b>	
<b>Authorship</b>	<b>Written by:</b>	Mark ASCH, Université de Picardie Jules Verne Thierry BIDOT, Neovia Innovation
	<b>Contributors:</b>	François BODIN, Université de Rennes Mool-Kyul KANG, Neovia Innovation Corentin LEFEVRE, Neovia Innovation
	<b>Reviewed by:</b>	Costas BEKAS, IBM Sergi GIRONA, Barcelona Supercomputing Centre
	<b>Approved by:</b>	MB/TB

## Document Status Sheet

Version	Date	Status	Comments
1.1	30/03/2017	Draft	Preliminary version
1.2	10/04/2017	Draft	First full version
1.3	15/04/2017	Draft	For internal review
1.4	22/04/2017	Final version	For approval
1.5	27/04/2017	Final version	Approved

## Document Keywords

<b>Keywords:</b>	BDEC, Big Data, Convergence, Data Analytic, ETP4HPC, Exascale, Extreme Computing, International Cooperation, Machine Learning, Numerical Laboratories, Simulation, PRACE, Research Infrastructure
------------------	---

### 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 own by the respective holders.

## Table of Contents

<b>Project and Deliverable Information Sheet .....</b>	<b>2</b>
<b>Document Control Sheet.....</b>	<b>2</b>
<b>Document Status Sheet .....</b>	<b>3</b>
<b>Document Keywords .....</b>	<b>3</b>
<b>Table of Contents .....</b>	<b>5</b>
<b>List of Figures .....</b>	<b>6</b>
<b>References and Applicable Documents .....</b>	<b>6</b>
<b>List of Acronyms and Abbreviations.....</b>	<b>6</b>
<b>Executive Summary .....</b>	<b>7</b>
<b>1 Introduction .....</b>	<b>8</b>
<b>2 Presentation of the BDEC initiative.....</b>	<b>9</b>
<b>3 Setting and Agenda .....</b>	<b>10</b>
<b>4 Report on Day 1.....</b>	<b>14</b>
<b>4.1 Keynotes .....</b>	<b>14</b>
<b>4.2 Roadmap presentations.....</b>	<b>15</b>
<b>4.3 Breakout sessions.....</b>	<b>16</b>
<b>4.4 Tour of Supercomputer Centre.....</b>	<b>17</b>
<b>5 Report on Day 2.....</b>	<b>18</b>
<b>5.1 Keynotes .....</b>	<b>18</b>
<b>5.2 Breakout sessions.....</b>	<b>18</b>
<b>6 Conclusion of the meeting.....</b>	<b>19</b>
<b>Annex – Workshop participants .....</b>	<b>20</b>

## List of Figures

Figure 1 - BDEC group photo .....	10
Figure 2 - Watson Yin introducing his keynote .....	14
Figure 3 - Zhiwei Xu explaining his keynote.....	14
Figure 4 - Rick Stevens explaining his keynote .....	15
Figure 5 - EU roadmap by Jean-Pierre Panziera .....	16
Figure 6 - Introduction to breakout session by Terry Moore and Mark Asch.....	17
Figure 7 - Tour of Supercomputer Centre by Haohuan Fu.....	17

## References and Applicable Documents

- [1] <http://exdci.eu>  
 [2] <http://www.exascale.org/bdec>

## List of Acronyms and Abbreviations

AIST	National Institute of Advanced Industrial Science and Technology (Japan)
BDEC	Big Data and Extreme-scale Computing
CN	People's Republic of China (China)
CNRS	Centre National de la Recherche Scientifique (National Centre for Scientific Research – France)
CSCS	Centro Svizzero di Calcolo Scientifico (Swiss National Computing Centre)
DoD	Department of Defense (USA)
DoE	Department of Energy (USA)
EC	European Commission
ECP	Exascale Computing Project
ETP4HPC	European Technology Platform for High Performance Computing
EU	European Union
EXDCI	European eXtreme Data and Computing Initiative
GPU	Graphics Processing Unit
GSIC	Global Scientific Information and Computing Centre (Japan)
HPC	High Performance Computing
ICT	Information and Communication Technology
INRIA	Institut National de la Recherche en Informatique et en Automatique (National Institute for Computer Sciences and Applied Mathematics – France)
IoT	Internet of Things
ISC	International Supercomputing Conference
KAUST	King Abdullah University of Science and Technology (Saudi Arabia)
NSF	National Science Foundation (USA)
PFlops	Peta Flops
PRACE	Partnership for Advanced Computing in Europe
RIKEN	Rikagaku Kenkyusho (Institute of Physical and Chemical Research – Japan)
SC16	Supercomputing 2016
SGI	Silicon Graphics International
SOA	Service Oriented Architecture
US/USA	United States of America

## Executive Summary

Since 2003, Big Data and Extreme-scale Computing workshops are organised as a worldwide international coordination.

In 2017, the workshop was organised in Wuxi (China) under the supervision of the National Supercomputing Centre in Wuxi which operates the Sunway TaihuLight supercomputer, the current 1<sup>st</sup> Top500 supercomputer.

The two-day workshop on Big Data and Extreme-scale Computing (BDEC) that began on March 9 in Wuxi was the latest instalment in a series of meetings in which eminent representatives of the scientific computing community are endeavouring to map out the ways in which Big Data challenges are intersecting with and affecting ongoing national plans for achieving exascale computing. With the help of the leadership of the National Supercomputing centre in Wuxi, which hosted the meeting, the 61 workshop participants included not only outstanding members of this community from the United States, the European Union, and Japan, but also 19 of their colleagues from China, the largest number ever to attend a BDEC meeting.

The workshop had two main goals: to update the group about ongoing national and regional development of Big Data and supercomputing infrastructure and applications; and to organize and focus the writing process that will produce the BDEC “Pathways to Convergence” report. The first objective was addressed by a set of keynotes and plenary talks that illuminated ongoing cyberinfrastructure efforts in different countries and described emerging applications of machine learning and artificial intelligence in traditional HPC areas of inquiry, such as cancer research and meteorology.

The second goal was the focus of the breakout sessions of three working groups, one for HPC and Cloud infrastructure, one for “Edge” computing environments (i.e., all the ICT lying outside centralized HPC/Cloud systems), and one for application paradigms that exploit one or both of these technological ecosystems. The rationale for the distinction between HPC/Cloud and Edge Computing environments lies in the problematic fact that, while the highest concentrations of computing and storage resources lie in the former, the vast majority of the rising flood of data is being generated in the latter.

The BDEC Wuxi meeting’s three breakout areas align with the key parts of “Pathways document” that is currently being completed. Building on the results of the workshop, the BDEC leadership plans to release the report late this spring.

Through global and inspirational talks (keynotes), state-of-the-art presentations (white papers on scientific cases, roadmaps on policies) and collective thinking (brainstorming sessions), the meeting achieved the following objectives: *(1) giving an overview of the current status of the convergence between HPC and Big Data, (2) determining the next challenges to tackle, (3) initiating the writing of a multi-authored, international document provisionally entitled “Pathways to Convergence”.*

## 1 Introduction

This report aims to give a description of the BDEC meeting that took place the 9<sup>th</sup> and 10<sup>th</sup> of March, 2017, in Wuxi, China under the supervision of the National Supercomputing Centre in Wuxi, which operates the Sunway TaihuLight, the current 1<sup>st</sup> Top500 supercomputer.

During two days, international experts from Asia, Europe and US, representing 7 countries (China, France, Germany, Japan, Spain and the USA) focused on “Pathways to Convergence” between Big Data and Exascale computing, which aim to develop and to complete the BDEC “Pathway” document. Machine learning and Artificial Intelligence, as related topics, were also addressed during this workshop, following the conclusions of the 2016 BDEC workshop held in Frankfurt (EU), which emphasized the necessity to include these issues in the BDEC subjects. During this workshop, there were overviews of international roadmaps with a limited number of invited keynotes and breakout groups.

China, the EU, Japan and the US gave an overview of the evolution of their HPC projects, strategies and ecosystem’s evolution at a high level.

The international cooperation was balanced with 25 experts coming from Asia (41%), 14 from Europe (23%) and 22 from the US (36%). The presentations were also balanced between the three continents with 5 presentations and keynotes coming from Asia, 3 for Europe and 3 for the US. The organisation was under the responsibility of an international board with equal participation from the three continents.

The EXDCI European project organized and aggregated the EU participation. Fourteen European stakeholders, representing a wide range of the EU ecosystem (science and industry representatives coming from 4 countries) attended the meeting. They were selected based on their participation and contribution to the EU HPC ecosystem and their knowledge related to the BDEC topics.

The report gives an overview of the BDEC meetings’ purposes and history, then gives a brief overview of the overall organisation and agenda of this meeting and will finally focus on the contents of the presentations, keynotes, white papers and breakout sessions that took place during these two days.



## 2 Presentation of the BDEC initiative

In the past years, the United States, the European Union, Japan and China have each moved aggressively to develop their own plans for achieving exascale computing in the next decade. Such concerted planning by the traditional leaders of HPC speaks eloquently about both the substantial rewards that await the success of such efforts, and about the unprecedented technical obstacles that block the path upward to get there.

While the exascale initiatives have understandably focused on the big challenges of exascale for hardware and software architecture, the relatively recent emergence of the phenomena of Big Data in a wide variety of scientific fields represents a tectonic shift that is transforming the entire research landscape on which all plans for exascale computing must play out.

Since 2013, the Big Data and Extreme-scale Computing (BDEC) workshops have been held to systematically map out and account for the ways in which the major issues associated with Big Data intersect with those of HPC and extreme scale computing.

These workshops have a triple role:

- **Coordination.** By sharing state of the art thoughts and projects, both within their scientific disciplines and the national or international frameworks, the participants aim to detect and/or initiate international coordination. As a side effect, this contributes to the dissemination of EU projects and policies.
- **Prospective.** Brainstorming is at the heart of the BDEC process. The goal is to identify the future challenges, breakthroughs and potential barriers between the HPC and Big Data communities. The result is to achieve a common view on scientific challenges to tackle.
- **Networking.** BDEC is a unique place where international researchers can meet and network.

Before the BDEC held in 2017 in Wuxi (China), the BDEC meetings have been held on a yearly basis, hosted alternatively in a concerned continent. The first BDEC meeting took place in Charleston (US) in 2013, followed by Fukuoka (Japan) in 2014, then in Barcelona (EU) in 2015 and Frankfurt (EU) in 2016. All details of these previous meetings can be found on a dedicated website: [www.exascale.org](http://www.exascale.org)

The costs of the BDEC meetings are shared between the host and the participants, applying the following rule: each host pays for the local costs (catering, meeting rooms) while travel costs are supported by the participants. As the meetings are held alternatively in each continent, the overall costs of the BDEC meetings are equally distributed.

For each meeting, the invited countries/continents are in charge to identify the proper experts who can give a strategic vision of the evolution of its continent/country, as well as to give high level scientific keynotes, reflecting the current state-of-the-art.

An executive board, composed of high level representatives coming from all the concerned parties, is in charge to set-up the agenda, follow the organisation of the meetings and validate the choices of experts and talks.

### 3 Setting and Agenda

The BDEC workshop took place in Wuxi from the 8<sup>th</sup> to the 10<sup>th</sup> of March 2017, at Grand Kingtown Hotel, which was chosen for its proximity with the Wuxi supercomputer centre.

The number of participants in the meeting was 61, of which 8 were women. The names and affiliations of the participants can be found in the annex.



Figure 1 - BDEC group photo

## Program of the workshop

The indicative program of the workshop is given below. Some minor modifications occurred during the course of the meeting.

<b>Wednesday, March 8th, 2017</b>	
7:00pm	Reception and Registration
<b>Thursday, March 9th, 2017</b>	
<b>INTRODUCTION</b>	
8:20am to 8:30am	<i>Welcome, Overview and Goals of the Meeting</i> Jack Dongarra, University of Tennessee Peter Beckman, Argonne National Laboratory
<b>KEYNOTES</b>	
8:30am to 8:50am	<i>Baidu ABC Platform</i> Watson Yin, Baidu
8:50am to 9:15am	<i>Low-Entropy Computing Systems A direction of energy efficient computing</i> Zhiwei Xu, Institute of Computing Technology, Chinese Academy of Sciences
9:15am to 9:45am	<i>Deep Learning in Cancer: Example for BDEC</i> Rick Stevens, Argonne National Laboratory
<b>OVERVIEWS</b>	
9:45am to 10:00am	<i>Big Data and Extreme Computing: Updates about China</i> Haohuan Fu, National Supercomputing Centre in Wuxi
10:00am to 10:15am	<i>US Overview/Update: Exascale Project Happenings</i> Robert Ross, Argonne National Laboratory
10:15am to 10:30am	<i>Updates on Europe strategy and plans</i> Jean-Pierre Panziera, Bull, ETP4HPC Sergio Girona, Barcelona Supercomputing Centre
10:30am to 10:45am	<i>BDEC Japan update for Open High Performance Computing and Big Data / Artificial Intelligence Infrastructure</i> Satoshi Matsuoka, Tokyo Institute of Technology
10:45am to 11:00am	<i>Overview of Pathways Document and Intro to Breakouts</i> Terry Moore, Mark Asch
11:00am to 11:15am	Coffee Break
<b>BREAKOUTS</b>	

11:15am to 12:30pm	<p><i>HPC/Cloud/Data Centre</i> Satoshi Matsuoka, Kate Keahey, Thomas Schulthess</p> <p><i>Data Workflows and Edge Computing</i> Ewa Deelman, Gabriel Antoniu, Geoffrey Fox</p> <p><i>Applications/Algorithms</i> David Keyes, Jean-Pierre Vilotte, Anshu Dubey</p>
12:30pm to 1:30pm	Lunch Break
<b>KEYNOTES</b>	
1:30pm to 2:00pm	<p><i>Big Data Assimilation for Extreme-scale Numerical Weather Prediction</i> Takemasa Miyoshi, RIKEN Advanced Institute for Computational Science</p>
2:00pm to 2:15pm	<p><i>NSF and Extreme Scale</i> William Gropp, National Centre for Supercomputing Applications</p>
<b>BREAKOUTS</b>	
2:15pm to 3:45pm	<p><i>HPC/Cloud/Data Centre</i> Satoshi Matsuoka, Kate Keahey, Thomas Schulthess</p> <p><i>Data Workflows and Edge Computing</i> Ewa Deelman, Gabriel Antoniu, Geoffrey Fox</p> <p><i>Applications/Algorithms</i> David Keyes, Jean-Pierre Vilotte, Anshu Dubey</p>
3:45pm to 4:15pm	<i>Reports from Breakouts and Synthesis of Document Status</i>
4:15pm	Tour of Wuxi Supercomputer Centre
5:00pm	Banquet
<b>Friday, March 10th, 2017</b>	
<b>KEYNOTES</b>	
8:30am to 9:00am	<p><i>Federated data services with abundant compute resources</i> Thomas Schulthess, ETH Zurich</p>
9:00am to 9:30am	<p><i>Scikit-learn: Machine learning for the small and the many</i> Gael Varoquaux, INRIA</p>
<b>BREAKOUTS</b>	
9:30am to 11:30am	<p><i>HPC/Cloud/Data Centre</i> Satoshi Matsuoka, Kate Keahey, Thomas Schulthess</p> <p><i>Data Workflows and Edge Computing</i> Ewa Deelman, Gabriel Antoniu, Geoffrey Fox</p> <p><i>Applications/Algorithms</i> David Keyes, Jean-Pierre Vilotte, Anshu Dubey</p>

11:30am to 12:30pm	Lunch Break
<b>FINAL</b>	
2:30pm to 3:30pm	<i>Reports from Breakouts and Synthesis of Document Status</i>
3:30pm to 4:00pm	<i>Conclusion of the Meeting</i> Jack Dongarra, University of Tennessee Peter Beckman, Argonne National Laboratory

The presentations were published online during the workshop on the dedicated website [www.exascale.org](http://www.exascale.org) thanks to the Argonne National Laboratory (ANL – US) team.

## 4 Report on Day 1

On the first day, the participants had a fresh view of the finished and upcoming projects in the BDEC field. Three keynotes opened the session followed by an overview session which was dedicated to have a look at the upcoming and accomplished work of the major players in the field: China, US, EU, and Japan.

Finally, breakout sessions were organized, which focused on the three following topics: HPC/Cloud/Data Centre, Data Workflows and Edge Computing, Applications / Algorithms.

### 4.1 Keynotes

Watson Yin, Zhiwei Xu and Rick Stevens presented the introductory keynotes. They introduced in this session their work and their prospective.

Watson Yin, vice-president of Baidu, started the meeting with a presentation of the Baidu ABC Platform, which combines Artificial Intelligence, Big Data and Cloud Computing. Baidu conducts several projects, among which self-driving cars and language recognition technologies. During the last years, in order to support these projects, the HPC needs of Baidu have grown rapidly and several calculators, cloud-oriented, have been deployed. Such platforms also anticipate the rise of the Internet of Things (IoT).



Figure 2 - Watson Yin introducing his keynote

Zhiwei Xu of the Chinese Academy of Sciences spoke about his research on low-entropy computing systems. His research aims at contributing to energy efficiency problems, working on the hypothesis that the lower a computing system entropy is, the more energy efficient it will be. This simple assertion has many potential applications in system architecture.



Figure 3 - Zhiwei Xu explaining his keynote

Rick Stevens, of Argonne National Lab, focused on Deep Learning in Cancer. He mentioned that cancer research can be significantly accelerated by developing new tools and methods, which combine Deep Learning, Big Data and HPC. US laboratories are currently working on this topic under three projects, addressing cancer biology, pre-clinical models and cancer surveillance. The first results of those projects are actually promising. And precisely by citing BDEC questions for Deep Learning, two ideas came up: (i) key frameworks, and workloads for Deep Learning and (ii) value of Deep Learning on its own software stack in the BDEC Universe.



**Figure 4 - Rick Stevens explaining his keynote**

In the afternoon, Takemasa Miyoshi, of the RIKEN Institute, presented the result of Big Data Assimilation for Extreme-scale Numerical Weather Prediction. Weather prediction has, since a long time be a sector in which HPC capacities were needed. Data Assimilation is the technique that combines predictive models with actual data from observations, e.g. Data Assimilation is at the forefront of a Big Data and HPC problem. The development of such an integrated approach has already produced very promising results, for example for sudden, severe rain previsions.

## 4.2 Roadmap presentations

The speakers respectively presented the current roadmaps of China, US, EU, and Japan.

Chinese projects are under a roadmap for the coming years (5-year plan) and the presenter, Haohuan Fu stressed the two exascale pilot systems which are Sunway TaihuLight and Tianhe-2 projects.

They had an improvement on their projects on Sunway TaihuLight, and were three Gordon Bell finalists at SC16, winning one Gordon Bell Prize. This was the first time that a Chinese computer entered the TOP500 list just by using their own hardware. The current peak performance of the Sunway TaihuLight is claimed at 125.4 PFlops. They intend to reinforce upcoming projects by combining HPC challenges, Deep Learning Applications, and Big Data as a long term plan.

The Chinese investments in 2016 were of 310 M¥ in HPC projects and of 389 M¥ in Cloud Computing and Big Data projects. Tianhe-2 super computer is the current hardware ongoing project with a 300 M¥ investment.

Rob Ross introduced the US focus on its exascale project by presenting ECP Software Technology overview. The US teams had built a comprehensive and coherent software stack by extending current technologies to exascale. They intend to work with several co-design centres to implement their roadmap.

The EU ambition is to become one of the world leaders in HPC by 2020. This ambition is implemented through the European Cloud Initiative, which supports the development of EU Data infrastructure, including European HPC. In January 2017, the European Commission released a communication on building a European Data Economy, and set-up three main objectives for the EU HPC ecosystem: (i) to possess two exascale machines by the years 2022-2023, at least one of which should be based on EU technology, (ii) to interconnect and mutualise EU and Member States HPC resources and (iii) to demonstrate scientific and industrial applications at the exascale level. The deployment of this strategy is based on the animation of a rich HPC ecosystem and no less than 5B€ of investments.



**Figure 5 - EU roadmap by Jean-Pierre Panziera**

To finish, as announced at BDEC Frankfurt which was held in June 2016, the Japanese roadmap focuses on the development of the Post-K era. They completed the functional design of system software and have been started its implementation.

Each initiative is not equal in terms of budget and global cooperation intensity is also different when looking at the players involved (research infrastructures, industrials...).

The review of the current roadmaps and objectives shows that the continuous efforts put in HPC are to continue in the following years, with a milestone on the first exascale pilots. The rise of the Big Data and Artificial Intelligence issues created many new scientific and industrial challenges and nurture promising applications in a wide variety of domains.

### **4.3 Breakout sessions**

On the first day, Terry Moore and Mark Asch introduced the overview of pathways by giving suggestions and possible questions and then there were two breakout sessions, which were based on three main pillars: HPC/Cloud Infrastructures, “Edge” Computing environments, and Applications paradigms.



Spokespersons were committed to each group with two functions:

- Organise the discussion and the capitalisation.
- Organise and present the results at the end of the process.



Figure 6 - Introduction to breakout session by Terry Moore and Mark Asch

#### 4.4 Tour of Supercomputer Centre

The end of the first day was dedicated to the Tour of Supercomputer Centre in Wuxi, presented by Haohuan Fu. During this visit, participants had an overall view of the Chinese Supercomputer.



Figure 7 - Tour of Supercomputer Centre by Haohuan Fu

## 5 Report on Day 2

The morning of the second day of the workshop was dedicated to two keynotes, followed by a breakout session. A synthesis concluded this workshop and prospected the next steps.

### 5.1 Keynotes

Thomas Schulthess opened the day by presenting Deep Learning tools on CSCS Cray system and showed the federated data pilot project with abundant compute resources. Among many, a specific constraint was the development of the interface between scientific teams and HPC resources. The solution found to keep the balance between a “user portal” and access to HPC resources was to design a Service Oriented Architecture (SOA), which is one of the only “scale-up” options.

Gaël Varoquaux of INRIA, represented “low performance computing” on his side, by explaining Scikit-learn, Machine Learning in Python. This program is based on Python which covers many users and developers to give a simple and efficient tool for data mining and data analysis. Scikit-learn has an ambition to reuse algorithms from laptops to data centres and they intend to improve from lessons of small-computer Machine Learning.

William Gropp presented the NSF overview and planning for the period 2017-2020. With the DoD and the DoE, the NSF is a major player for the development and deployment of the HPC ecosystem in the US. The overall implementation follows the US National Strategic Computing Initiative, as well as the NSF future directions for the Future of Advanced Computing. The NSF intends to implement a HPC software and algorithms ecosystem for the US scientists, and plays a role in the development of a HPC expert workforce. Such plans include research infrastructure dedicated to Open Science needs, such as the Blue Water instrument.

### 5.2 Breakout sessions

After an ultimate walkthrough of the ideas, people were asked to develop small chapters on the identified topics/challenges/clusters of ideas, in order to update and improve the “Pathways to Convergence” report.

## 6 Conclusion of the meeting

High Performance Computing, Machine Learning and Big Data are moving aggressively with strong interactions to new converged e-science paradigms.

Integrative models between HPC, Big Data and the newcomer Machine Learning require tackling new challenges and an integrative roadmap between each community. This is the goal of the “Pathways to Convergence” document. This document is expected to be delivered during Q3 2017 and presented at Supercomputing 2018 in Denver. It will be available on the BDEC dedicated website: [www.exascale.org](http://www.exascale.org). Within the frame of the EXDCI project, this document will be provided to the EC as a complementary document of the EXDCI deliverables.

The major developed countries are investing massively to help their communities build these paradigms and the associated eco-systems. For the European attendees of the workshop, these “exposés” were perceived as a sign of a dangerous stalling of the EU (except for Switzerland).

With the current plan, the first pre-exascale system available in Europe is planned for 2020 while such systems are already available in America (Titan, 27PF), China (Sunway TaihuLight, 125 PF) and Japan (Tsubame3.0, August 2017, 47.2 AI-PF, 12.1 PF, based on a new GPU compute blade co-designed by SGI and Tokyo Tech GSIC). Missing the deadline is not just being late; there are multiple negative consequences.

The full stack of software and hardware technologies must be addressed to create an efficient technological ecosystem. This set of resources won’t happen only with a high level, long-term effort. Lack of some part of technology, software or hardware, is very likely going to create larger issues.

Europe is significantly investing in HPC and Big Data but without the agility and speed that would allow us to catch-up with our competitors. There is a need to focus and organize the current efforts in a way that is closer to an industrial project rather than a set of research projects.

## Annex – Workshop participants

First Name	Last Name	Organization
Jack	Dongarra	University of Tennessee
Terry	Moore	University of Tennessee
Tracy	Rafferty	University of Tennessee
Geoffrey	Fox	Indiana University
William	Gropp	National Centre for Supercomputing Applications
Anshu	Dubey	Argonne National Laboratory
Franck	Cappello	Argonne National Laboratory
David	Keyes	KAUST & Columbia University
William	Tang	Princeton University/PPPL
Robert	Wisniewski	Intel Corporation
Alok	Choudhary	Northwestern University
Martin	Swany	Indiana University
Robert	Ross	Argonne National Laboratory
Ewa	Deelman	University of Southern California
Miron	Livny	University of Wisconsin-Madison
Peter	Beckman	Argonne National Laboratory
Rick	Stevens	Argonne National Laboratory
Sandor	Szalay	Johns Hopkins University
William	Kramer	NCSA/University of Illinois
Tarek	El-Ghazawi	The George Washington University
Barbara	Chapman	Stony Brook University and BNL
Katarzyna	Keahey	Argonne National Laboratory
Thomas	Schulthess	ETH Zurich / CSCS
Hugo	Falter	ParTec GmbH
Mark	Asch	University of Picardy Jules Verne
Gabriel	Antoniou	INRIA
François	Bodin	University of Rennes 1
Jean-Pierre	Panziera	Bull / Atos
Jean-Francois	Lavignon	Technology Strategy
Jean-Pierre	Vilotte	Institut de Physique du Globe de Paris, CNRS
Thierry	Bidot	Neovia Innovation
Gael	Varoquaux	INRIA
Jean-Claude	Andre	Jca Consultance et Analyse
Rosa	Badia	Barcelona Supercomputing Centre
Judit	Gimenez Lucas	Barcelona Supercomputing Centre
Sergi	Girona	Barcelona Supercomputing Centre
Hirotaaka	Ogawa	National Institute of Advanced Industrial Science and Technology (AIST)

First Name	Last Name	Organization
Yutaka	Ishikawa	RIKEN Advanced Institute for Computational Science
Hitoshi	Sato	National Institute of Advanced Industrial Science and Technology (AIST)
Takemasa	Miyoshi	RIKEN
Balazs	Gerofi	RIKEN Advanced Institute for Computational Science
Satoshi	Matsuoka	Tokyo Institute of Technology
Depei	Qian	Sun Yat-sen University
Yutong	Lu	National Supercomputing Centre in Guangzhou
Xuanhua	Shi	Huazhong University of Science and Technology
Guangrong	Gao	Beijing Institute of Technology, Shanghai Jiao Tong University
Zongben	Xu	Xi'an Jiaotong University
Minyi	Guo	Shanghai Jiao Tong University
Shimin	Yin	Baidu
Jian	Wang	Alibaba
Zhihui	Xu	Institute of Computing Technology Chinese Academy of Sciences
zhiwei	Xu	Institute of Computing Technology Chinese Academy of Sciences
Guoliang	Li	Tsinghua University
Kai	Lu	National University of Defense Technology
Long	Wang	Baidu
Yifeng	Chen	Peking University
Guangwen	Yang	Tsinghua University, National Supercomputing Centre in Wuxi
Haohuan	Fu	Tsinghua University, National Supercomputing Centre in Wuxi
Wei	Xue	Tsinghua University, National Supercomputing Centre in Wuxi
Lin	Gan	Tsinghua University, National Supercomputing Centre in Wuxi
Jinzhe	Yang	Tsinghua University, National Supercomputing Centre in Wuxi