



## SEVENTH FRAMEWORK PROGRAMME

FP7-ICT-2013-10



**DEEP-ER**

**DEEP Extended Reach**

Grant Agreement Number: 610476

**D1.7**

**Midterm management report at month 36**

***Approved***

**Version:** 2.0

**Author(s):** E.Suarez (JUELICH)

**Contributor(s):** S.Eisenreich (BADW-LRZ), H.Ch.Hoppe (Intel), K.Thust (JUELICH), V.Beltran (BSC), A.Zitz (JUELICH), I.Zacharov (Eurotech)

**Date:** 04.05.2017

## Project and Deliverable Information Sheet

<b>DEEP-ER Project</b>	<b>Project Ref. №:</b> 610476	
	<b>Project Title:</b> DEEP Extended Reach	
	<b>Project Web Site:</b> <a href="http://www.deep-er.eu">http://www.deep-er.eu</a>	
	<b>Deliverable ID:</b> D1.7	
	<b>Deliverable Nature:</b> Report	
	<b>Deliverable Level:</b> CO*	<b>Contractual Date of Delivery:</b> 30 / September / 2016
		<b>Actual Date of Delivery:</b> 30 / September / 2016
<b>EC Project Officer:</b> Luis Carlos Busquets Pérez		

\* - 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

<b>Document</b>	<b>Title:</b> Midterm management report at month 36	
	<b>ID:</b> D1.7	
	<b>Version:</b> 2.0	<b>Status:</b> Approved
	<b>Available at:</b> Publishable part at: <a href="http://www.deep-er.eu">http://www.deep-er.eu</a>	
	<b>Software Tool:</b> Microsoft Word	
	<b>File(s):</b> DEEP-ER_D1.7_Midterm_management_report_M36_v2.0-PublishablePart	
<b>Authorship</b>	<b>Written by:</b>	E.Suarez (JUELICH)
	<b>Contributors:</b>	S.Eisenreich (BADW-LRZ), H.Ch.Hoppe (Intel), K.Thust (JUELICH), V.Beltran (BSC), A.Zitz (JUELICH), I.Zacharov (Eurotech)
	<b>Reviewed by:</b>	R.Léger (Inria), I.Schmitz (ParTec)
	<b>Approved by:</b>	BoP/PMT

**Document Status Sheet**

<b>Version</b>	<b>Date</b>	<b>Status</b>	<b>Comments</b>
1.0	30/September/2016	Final	EC submission
2.0	04/May/2017	Approved	EC approved

## Document Keywords

<b>Keywords:</b>	DEEP-ER, HPC, Exascale, midterm report, month 36
------------------	--

### Copyright notice:

© 2013-2016 DEEP-ER Consortium Partners. All rights reserved. This document is a project document of the DEEP-ER project. All contents are reserved by default and may not be disclosed to third parties without the written consent of the DEEP-ER partners, except as mandated by the European Commission contract 610476 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

Project and Deliverable Information Sheet .....	1
Document Control Sheet .....	1
Document Status Sheet .....	2
Document Keywords.....	3
Table of Contents .....	4
List of Figures.....	5
List of Tables .....	5
Executive Summary .....	6
1 Publishable summary .....	8
1.1 Project objectives.....	8
1.2 Work performed and main results .....	12
1.3 Expected final results .....	20
Annex A.....	21
A.1 Listing of dissemination activities.....	21
List of Acronyms and Abbreviations.....	25

**List of Figures**

Figure 1: Joint EEP booth at SC15. .... 15

Figure 2: High-level view of the DEEP-ER Prototype. BN=Booster Node; CN=Cluster Node; NVM=Non-Volatile Memory; NAM=Network Attached Memory ..... 16

Figure 3: Sketch of DEEP-ER I/O software layers. .... 18

Figure 4: Sketch of DEEP-ER resiliency layers..... 19

**List of Tables**

No table of figures entries found.

## Executive Summary

The DEEP – Extended Reach (DEEP-ER) project started on 1<sup>st</sup> October 2013 and will last 42 months. The project addresses two significant Exascale challenges: the growing gap between I/O bandwidth and compute speed, and the need to significantly improve system resiliency. DEEP-ER extends the Cluster-Booster Architecture first realised in the DEEP project by a highly scalable I/O system. Additionally, an efficient mechanism to recover application tasks that fail due to hardware errors will be implemented. The project will build a hardware prototype including new memory technologies to provide increased performance and power efficiency. As a result, I/O parts of HPC codes will run faster and scale up better. Furthermore, HPC applications will be able to profit from checkpoint and task restart on large systems reducing overhead seen today. To demonstrate it a set of seven applications with high societal impact are ported to the DEEP-ER prototype and make use of the I/O and resiliency capabilities available therein.

This report describes the objectives, work performed, resources used, and results achieved during **months 25 to 36** of the DEEP-ER project. The main achievements in the reporting period are enumerated below:

- Co-design effort stepping up: continuous discussions between hardware, software, and application developers to assure a coherent development that addresses all requirements.
- Update by partner Eurotech of the “Aurora Blade” architecture, with a new design of the KNL-based node board, now with one KNL per board. Commitment by partner Eurotech to produce this design and complete the DEEP-ER prototype with in-kind resources.
- Completion of the installation of the Software Development Vehicle (SDV) a hardware platform for software and application developments, which will also become the Cluster part of the DEEP-ER prototype. Completion of SDV with 8 pre-production KNL boards.
- DEEP-EST Booster in advance state of construction: first samples of the three key elements of the system (KNL blade, backplane and Root card) tests almost completed and production of all the devices needed for the full DEEP-ER system triggered. Rack and chassis mechanics available. Room infrastructure ready for the installation of the DEEP-ER Booster.
- First results of I/O benchmarks and application mock-ups running on the Software Development Vehicle.
- Network Attached Memory (NAM) prototype available. Functionality and performance established. LibNAM implementation ongoing. Integration of libNAM with SIONlib for checkpointing functionality ongoing.
- Development of the DEEP-ER I/O software stack – containing BeeGFS, SIONlib, and E10 – completed, taking into account the requirements from resiliency software and applications. Verification with applications and benchmarks ongoing.

- Implementation of resiliency software layer completed. Tests with applications and benchmarks ongoing.
- Benchmarks integrated in JUBE environment and periodic test procedure established. GERSHWIN application from Inria integrated in JUBE, integration of further applications ongoing. Additionally, benchmarks and mini-apps for I/O and resiliency benchmarking also included.
- Application adaptations and improvements ongoing: optimisation to take benefit from the second generation Intel® Xeon Phi™, code partition between Cluster and Booster parts of the DEEP architecture, integration with the I/O and resiliency tools developed in DEEP-ER. Porting to the SDV completed, benchmarking ongoing.
- Dissemination of project goals and status in various workshops and conferences, amongst others the Supercomputing Conference (SC15) in the US and the International Supercomputing Conference (ISC 2016) in Germany.
- Coordination and co-organisation of joint dissemination activities with other European Exascale Projects, i.e. for the joint booth at SC15 and ISC 2016, as well as the European HPC Summit Week. Preparation for a joint presence at SC 2016 ongoing, this time under the leadership of EXDCI.



## 1 Publishable summary

The DEEP-ER project tackles two important Exascale challenges. Firstly, the increasing gap in the growth rate of compute power with respect to the amount and performance of memory and storage available in HPC systems. Secondly, the high failure rates expected in Exascale systems as a consequence of the increased number of components and the need to take their performance and energy efficiency to the limits. To address these issues, DEEP-ER will extend the heterogeneous Cluster-Booster Architecture implemented by the DEEP<sup>1</sup> project by additional I/O and resiliency functionalities.

DEEP-ER targets a seamless integration of a high-performance I/O subsystem into the Cluster-Booster Architecture. New memory technologies will be used to provide a multi-level I/O infrastructure capable of supporting data-intensive applications. Additionally, an efficient and user-friendly resiliency concept combining user-level checkpoints with transparent task-based application restart will be developed, which enables applications to cope with the higher failure rates expected in Exascale systems.

The DEEP-ER prototype system and the I/O and resiliency concepts will be evaluated using seven HPC applications from fields that have proven their need for Exascale resources. These applications will be ported and optimised to demonstrate the usability, performance and resiliency of the DEEP-ER Prototype. Systems that leverage the DEEP-ER results will be able to run more applications at the same time, thereby increasing scientific throughput. This is due to improved computational efficiency, better and more scalable I/O performance and a substantial reduction in the loss of computational work through system failures.

### 1.1 Project objectives

The specific objectives of the DEEP-ER project and the results already achieved towards them are:

- 1. Address two main Exascale challenges: I/O and resiliency. DEEP-ER will extend the DEEP Architecture by: i) a highly scalable, efficient and easy-to-use parallel I/O system; ii) providing a combination of low-overhead user-level checkpoint/restart and automatic task recovery.*
  - The implementation of the DEEP-ER I/O software – including BeeGFS, SIONlib, and E10 – has been completed taking the application and resiliency software requirements into account (see D4.1, D4.2, and D4.4). Optimised BeeGFS support for EXTOLL network under implementation.
  - The resiliency software has been implemented (see D5.1, D5.2, and D5.3) taking the application requirements into account.
  - Integration of resiliency and I/O software layers with each other completed. Benchmarks, tests, and user support are ongoing.
- 2. Develop a prototype system of the extended DEEP Architecture that leverages advances in hardware components (Intel's second generation Intel<sup>®</sup> Xeon Phi<sup>™</sup> processors, high-speed interconnects and non-volatile memory devices) to further improve the performance and efficiency of the DEEP-ER Prototype and realise the*

---

<sup>1</sup> [www.deep-project.eu](http://www.deep-project.eu)

*novel I/O system and resiliency improvements. This prototype will allow proving the viability of the concept for 500 PFlop/s-class of supercomputers.*

→ After initial exploratory studies of several architecture alternatives, in M18 it was decided to implement the DEEP-ER Prototype using the Aurora Blade architecture concept from Eurotech. After a first design phase in which a specially developed KNL-blade containing two fully independent Booster Nodes was considered and finally rejected due to its high complexity and high risk associated, the system design was completed for an Aurora KNL blade hosting only one KNL, in Adams Pass form factor. With this new approach the overall development effort on Eurotech side shifted from a complex board design into the mechanical and thermal integration of the various components needed to adapt the standard KNL board from Intel into the Aurora chassis and rack infrastructure.

→ Already before this decision on the KNL board, two additional design choices had been made: in light of the good progress with the EXTOLL TOURMALET implementation (based on an ASIC and performed outside the project), this interconnect was selected for the DEEP-ER Prototype. The exact network topology (an 3D torus open in one direction) was then settled through co-design discussions with the application developers. In addition, a line of SSD replacement devices from Intel was selected as on-node NVM devices.

→ The Software Development Vehicle (SDV), a hardware platform for development of system (I/O and resiliency) and application software, was installed at JUELICH, which will later become the Cluster side of the DEEP-ER prototype. The EXTOLL interconnect of the SDV has been very recently upgraded by substituting the version A2 Tourmalet NICs by its final implementation A3, which reaches a link bandwidth of over 100 GB/s.

→ A further upgrade of the SDV was the integration of 8 Intel Xeon Phi (KNL generation) devices, which are used to port and optimise the parts of the applications that will later run on the DEEP-ER Booster.

→ The room infrastructure has been prepared for the arrival of the DEEP-ER Booster, including electrical and water connections. The Booster rack has been delivered and connected to the cooling loop and is ready to soon receive the first chassis of four that will constitute the whole system.

3. *Explore the potential of new storage technologies (non-volatile and network attached memory) for use in HPC systems, with a focus on parallel I/O and system resiliency by integrating them with the DEEP-ER Prototype.*

→ NVM technology options for integration with the DEEP-ER prototype have been evaluated. A series of Intel SSD replacement devices was selected (see Deliverables D3.1 and D3.2). These devices implement the NVM Express (NVMe) interface and use PCI Express generation 3 links to connect to the compute nodes.

→ Extensive experiments were undertaken first with two samples of these devices at Juelich and later with the SDV and production NVMe devices, and a wide range of measurements with I/O benchmarks and application mock-ups

are available (see D3.3). These clearly show substantial performance increases over best-of-breed SSDs, in particular for scenarios with many parallel I/O requests.

→ Two NAM devices have been integrated into two of the Intel Xeon Phi (KNL) nodes of the SDV and the full functionality and performance of these has been verified.

→ The NAM uses partner UHEI's hybrid HMC controller implementation, which has been completed and validated. One NAM prototype device has been integrated to the SDV: a state-of-the-art Virtex 7 FPGA from Xilinx implements the HMC controller, NAM functional logic and two EXTOLL links compatible with the full TOURMALET EXTOLL fabric speed.

4. *Develop a highly scalable, efficient and user-friendly parallel I/O system tailored to HPC applications. The system will exploit innovative hardware features, optimise I/O routes to maximise data reuse, and expose a user friendly interface to applications. Its design will meet the requirements of traditional, simulation-based as well as emerging data-intensive HPC applications.*

→ The design of the DEEP-ER I/O system has been completed taking into account the outcome of the discussions with the experts from WP3 – to guarantee that the hardware provides the needed functionality – and with WP5 and WP6 to gather all their requirements on the I/O infrastructure.

→ The functionalities that each of the three I/O software packages – the BeeGFS file system of Fraunhofer, the parallel I/O library SIONlib, and the E10 software stack – must provide to the project, the interplay between them, and their interfaces have been described in deliverables D4.1 and D4.2.

→ In BeeGFS two new functionalities have been implemented: cache domain handling and user-level stripe-size definition. The cache domain will be executed on the node-level NVM devices and can be executed synchronous or asynchronously. Both the synchronous and asynchronous versions have been implemented.

→ SIONlib has been refactored and improved to eliminate code replications and increase the overall modularity and manageability of the library. The functionality required for buddy-checkpointing has been implemented and tested. SIONlib is used also as basis for buddy-checkpointing and supported the use of NAM.

→ Partner Seagate has completed the E10 implementation, new driver and extensions have been developed and tested, now supporting caching functionalities for BeeGFS and other file systems. A new support library has been developed to make the integration of the new E10 functionality transparent to applications.

→ On the SDV, access to the local very fast NVMe devices is possible for applications – either by using Posix I/O with a particular directory path or by relying on the “BeeGFS on demand” (BeeOND) functionality which provides the full BeeGFS interface for accessing local storage. The same setup will be used for the DEEP-ER Booster nodes.

→ A list of benchmarks to be used for the evaluation of the DEEP-ER I/O software has been identified (see D4.3) and they have been integrated into the JUBE benchmark environment. They are used to regularly monitor the overall I/O performance and to measure the impact of the various developments done in the DEEP-ER project. This activity has allowed finding bugs and identifying the right configuration parameters in various software packages. Some applications have been integrated in JUBE and a user-guide has been provided to the developers to facilitate the integration of the remaining ones.

5. *Develop a unified user-level system that significantly reduces the checkpointing overhead by exploiting multiple levels of storage and new memory technologies. Extend the DEEP programming model to combine automatic re-execution of failed tasks and recovery of long-running tasks from multi-level checkpoints, and introduce easy-to-use annotations to control checkpointing.*

→ In a co-design effort, the overall resiliency software stack has been defined (see D5.1) taking into account the requirements from the WP6 application developers, the WP3 hardware capabilities, and the I/O functionality required from/provided by from WP4. Also the role to be played by the application-based and the task-based resiliency functionalities, and the interfaces between them have been defined.

→ The Scalable Checkpoint/Restart (SCR) library has been adapted to the needs of the DEEP-ER project including an API (abstraction layer) for the application users to apply SCR in their codes (see D5.2). The code has been adapted to reflect recent changes in the BeeGFS API, and the new SIONlib buddy-checkpointing functions have been implemented and tested.

→ OmpSs adaptations for task-based resiliency have been implemented, including the interaction with ParaStation MPI in order to extend the task-based implementation to support offloaded tasks.

→ Integration between the CP/RS framework, OmpSs and ParaStation MPI is completed.

→ In close collaboration with WP3, an event Monte Carlo failure model has been designed and implemented, in order to optimise policies that determine for each application the frequency, redundancy level and storage-location of each checkpoint. A closed formula has been also developed. Results from the failure model and the closed formula are aligned. The closed formula has been integrated with the SCR library, enabling the latter to provide the user indications on the required checkpointing frequency for each application.

→ Data from production machines accessible to DEEP-ER partners has been investigated to obtain an estimation of the MTBF in current HPC systems.

6. *Analyse the requirements of HPC codes carefully selected to represent the needs of future Exascale applications with regards to I/O and resiliency, guide the design of the DEEP-ER hardware and software components, optimise these applications for the extended DEEP Architecture and use them to evaluate the DEEP-ER Prototype. Selected applications cover the fields of Health, Earthquake Physics, Radio*

*Astronomy, Oil Exploration, Space Weather, Quantum Physics, and Superconductivity.*

→ Continuous co-design discussions have taken place to identify the application requirements – in terms of hardware capabilities, I/O and resiliency functionalities – and use them to define the hardware and software layers in the project. DDG teleconferences and face-to-face meetings have been used as the co-design platform, as applications evolve with time through code optimisations and implementation of new functionalities.

→ The structure of the applications has been analysed, performance and scaling tests have taken place. In several cases mock-ups of the applications have been developed, to easily implement code changes and analyse their impact on the overall performance.

→ These analyses have helped to improve the applications in various ways. Important topics in the code optimisations are: vectorisation, optimising communication strategies and/or numbering schemes, improving I/O, implementing checkpointing, etc.

→ SIONlib and OmpSs have been integrated in several applications. In collaboration with WP3 and WP4, mock-ups from some applications (e.g. space weather and seismic imaging) have been used for I/O benchmarking.

→ The applications have been ported and benchmarked on the SDV (see D6.2). Currently KNL-porting and optimisation is ongoing for the parts of the codes that will run on the DEEP-ER Booster.

→ The I/O benchmarking work is increasingly using the full application versions, which gives more realistic results and enables the project to extend the benchmarking to all applications.

7. *Demonstrate and validate the benefits of the extended DEEP Architecture and its first implementation (the DEEP-ER Prototype) with the DEEP-ER pilot applications and for applications that exploit generic multi-scale, adaptive grid and long-range force parallelisation models.*

→ First results, obtained by predicting the performance of three applications on the DEEP-ER Prototype with the Dimemas simulation tool by partner BSC and extrapolating the scaling characteristics have been obtained and documented in Deliverable D7.1. Further applications are being analysed and new modelling aspects, such as I/O performance, will be taken into account. Focus of work is now on I/O tracing and modelling.

## 1.2 Work performed and main results

According to the amended DoW, four milestones had to be reached between **month 25 and month 36** of the DEEP-ER project:

- **MS8:** “Overall design of Aurora Blade prototype completed”: Deliverable D8.1, originally submitted in M24, has been updated to reflect the final design, which bases the KNL-boards for the Aurora Blade architecture on commercially available Intel KNL

boards (S7200AP or “Adams Pass”). The document has been re-submitted in M29 (February 2016).

- **MS9:** “Applications ported to the SDV”: Deliverable D6.2 submitted in M30, with the results already achieved by the applications running on the Software Development Vehicle.
- **MS10:** “BNC evaluator (i.e. KNL-blade) and rest of Aurora Blade components available”.
  - KNL blade: samples were available at Eurotech in M33, comprising Intel’s Adam Pass board, and the three PCBs constituting the Blade Interface Board (BIB). Laboratory tests took place to verify its functionality. Thermal tests (with KNL blade attached to its cold plate) are currently ongoing.
  - Backplane samples have been also tested and its functionality verified, also in combination with the KNL blade.
  - Rootcard: design problems were found in the first samples and a re-spin of these components with the corrected design was needed. The second version of the Root card is currently being tested to verify that all issues have been solved.
- **MS11:** “First Aurora chassis delivered to Jülich”: due to difficulties during the test phase of the individual components, and the later arrival of some of their parts to Eurotech, the delivery of the first chassis had to be re-scheduled and is currently planned for the second half of October 2016 (M37).

#### Management, legal and administrative tasks

A large part of the management activities in the present reporting period were dedicated to monitor the progress of the project with regards to the achievement of all technical goals specified in the Description of Work (DoW) and the fulfilment of all commitments to the European Commission, as well as for addressing the recommendations issued by the reviewers in M24, including the preparation of the first DoW amendment.

The Project Management Team organised the agenda for the review meeting at month 24, which took place on December 9, 2015 in Brussels (Belgium), and the interim review at month 32, which took place on June 8, 2016 in Jülich (Germany). To fulfil the internal quality policies rehearsal meetings one day before the reviews were conducted. As a result of both reviews, the project has been evaluated as doing “good progress”. Additionally, all deliverables submitted between M13 and M32 have been approved. The comments from the reviewers in the interim review at M32 and their recommendations concerning future work are addressed in section 2.2.

An additional important management activity was focused on concluding the internal discussions and formal request for the first DoW amendment, which extended the project by 6 months until end of March 2017. The project is currently running following the amended DoW.

The financial statements from all partners were submitted to the NEF server after the end of the second project year and the financial data has been approved by the European Commission.

Monthly teleconferences of the Team of Work Package leaders (ToW) were organised to periodically discuss the progress in all Work Packages (WPs). Furthermore, bi-weekly teleconferences of the Design and Development Group (DDG) have been held to discuss the progress in the implementation of the different developments, and to drive co-design and cross-WP discussions.

Deliverables D1.6, D1.7, D3.3, D3.4, D4.4, D5.3, D6.2, D8.1, and D8.2 have been submitted to the European Commission after having passed through the mandatory DEEP-ER internal review process. The already approved public deliverables (the publishable part of D1.6, D3.3, and D3.4) have been uploaded to the project website.

### Dissemination, training and outreach

The DEEP-ER prototype breaks new ground in the combination of its principal components (KNL CPU, NVM devices, EXTOLL TOURMALET network), and the Aurora Blade architecture includes innovative ways to integrate, package and cool a highly efficient HPC system. In addition, the innovative DEEP-ER I/O and resiliency concepts will require the development of new techniques and tools never tested before. Access to the know-how achieved in this process shall not remain limited to the group of people directly involved in the project, but must be made available to a wider community to move the HPC field forward. For this reason, WP2 in DEEP-ER is entirely dedicated to the dissemination of the knowledge accumulated over the project's duration, and to train the users on its application.

The centre of the dissemination activities of DEEP-ER is its web site at [www.deep-er.eu](http://www.deep-er.eu). The web page is updated regularly and referred to in all other materials (articles, press releases, brochures, presentations, etc.). It is used to publish general information about the project, current activities, training opportunities, job vacancies, publications, tutorials, success stories, and achievements of the project.

Following previous recommendations, the content of the website is being further developed. In particular, more focus is being put on the applications and on their display as a global and collaborative effort, as well as on highlighting the outcome of the strong co-design performed in the project.

Two social media platforms have been chosen to disseminate DEEP-ER news amongst the HPC world and the general public: LinkedIn and Twitter. The already existing DEEP LinkedIn group has been extended to host also DEEP-ER. The strong link existing between both projects justifies the use of a single group. The same applies for Twitter. Updates are being regularly posted (at least at bi-weekly basis) and frequently re-posted by other Twitter users in the HPC community. The most recent Twitter posts are visible also at the main page of DEEP-ER's website. Continuous and steady increase in Twitter follower numbers has been observed. Re-tweets and interactions are in a solid state as well. @DEEPprojects Twitter account has been established as key player in the Twitter HPC community, providing impressions via re-tweets and mentions. LinkedIn is still slower, but postings are more frequent now and also more colleagues engage in the group. Although the total number of members is not too high, the number of project external members raises and a larger audience via likes and shares has been reached successfully.

Several high-profile dissemination activities have taken place in the reporting period. Partners from the DEEP-ER consortium presented the project's concept in conferences and

workshops, including two of the most important events in the HPC community the Supercomputing Conference (SC), which took place in Austin (USA) in November 2015, and the International Supercomputing Conference (ISC), which took place in Frankfurt (Germany) in June 2016.

At both SC2015 and ISC 2016, the DEEP-ER project co-organised a joint booth together with other European Exascale Projects (EEP) – Mont-Blanc (1 and 2), EPiGRAM, and EXA2CT. DEEP and DEEP-ER were jointly shown in a wall space describing the architecture and main goals of the projects. The HMC controller (the core component of DEEP-ER's NAM), was displayed as a demo at the booth. Additionally, a DEEP-ER flyer was prepared and distributed at the European Exascale Projects (EEP) booth and at the booths of other project partners. In addition, the DEEP-ER project was presented in the joint EEP BoF and at presentations and panel discussion at the Intel booth.



Figure 1: Joint EEP booth at SC15.

DEEP-ER was also presented at the European HPC Summit Week, which took place in Prague (Czech Republic) in May 2016, where FP7 and H2020 projects were presented. DEEP-ER members also participated in the ETP4HPC discussion in the Extreme Scale Demonstrators (EsDs) Workshop, which was organised at the last day of the summit week.

Additionally, several articles and publications on the project approach and results have been submitted. A list with all dissemination activities performed in the present reporting period is given in Annex A.1 of this report

Regarding contacts with industry, an action plan for industry and business cooperation has been prepared to define the activities and dissemination materials to be produced specifically for industrial contacts. Regarding those activities focused in the commercialisation of project results, the partners responsible for the different developments are already making significant efforts to increase the productising potential of their own and shared IP. Task 2.2 will first gather information on all the actions and plans from the individual partners. With this overview, potential synergies will be identified and strategies to increase the visibility of the existing and upcoming activities will be defined. Additionally, content is being created targeting potential users of a DEEP-ER system, that will be leveraged via the website, social media campaigns and potentially also mailing campaigns. Finally, the DEEP and DEEP-ER projects have been shown to an industry-focused audience at CeBIT16 in March 2016 at Hannover (Germany).



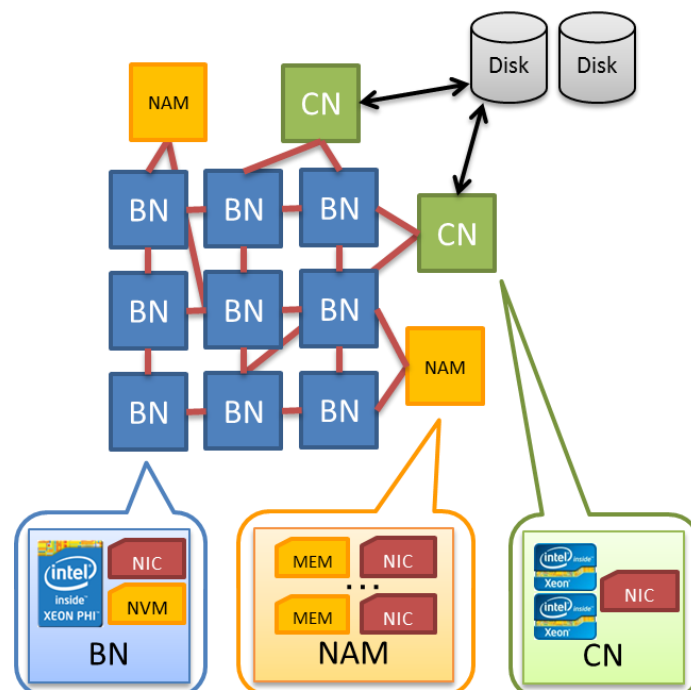
Training the community on how to use the software and hardware developed in DEEP-ER is an important part of the project. The main goal of the training events in DEEP-ER is to teach the application developers participating in the project on how to use the software tools and programming environment that will run on the DEEP-ER Prototype and other intermediate hardware evaluators. A hands-on training event, jointly organised with Mont-Blanc, for the application developers of both projects took place in March 2016 in Barcelona (Spain). A further training event was organised in the first days of September 2016 in Jülich, where Intel members instructed application developers on how to best prepare their codes to profit from the Intel Xeon Phi (KNL) architecture.

### Technical Work

The technical work in DEEP-ER is grouped into three main topics: system architecture and hardware, system software (including I/O and resiliency software), and applications.

#### *Overview*

The DEEP-ER project designs and builds a second-generation prototype (see Figure 2) of the Cluster-Booster Architecture. In the DEEP-ER Prototype the second generation Intel Xeon Phi processors (KNL) provides the compute power of the Booster Nodes (BN), while the most recent Intel Xeon processors populate the Cluster Nodes (CN). A uniform high-speed interconnect runs across Cluster and Booster, and network-attached memory (NAM) devices connected to it provide high-speed shared memory access. The Booster Nodes themselves also feature additional non-volatile memory (NVM) devices for efficiently buffering I/O and storing checkpoints.



**Figure 2: High-level view of the DEEP-ER Prototype. BN=Booster Node; CN=Cluster Node; NVM=Non-Volatile Memory; NAM=Network Attached Memory**

The DEEP-ER multi-level I/O infrastructure has been designed to support data-intensive applications and multi-level checkpointing/restart techniques. The project develops a scalable and efficient I/O software platform based on the BeeGFS parallel file system, the

parallel I/O library SIONlib, and the I/O software package Exascale10 (E10). It aims to enable an efficient and transparent use of the underlying hardware and to provide all functionality required by applications for standard I/O and checkpointing.

On top of this I/O infrastructure DEEP-ER develops an efficient and user-friendly resiliency concept combining user-level checkpoints with transparent task-based application restart. OmpSs is used to identify application's individual tasks and their interdependencies. The OmpSs runtime has been extended in DEEP-ER in order to automatically re-start tasks in the case of transient hardware failures. In combination with a multi-level user-based checkpoint infrastructure to recover from non-transient hardware-errors, applications are now able to cope with the higher failure rates expected in Exascale systems. DEEP-ER's I/O and resiliency concepts are being evaluated using seven HPC applications from fields that have proven their need for Exascale resources.

#### *System Architecture and New Technologies*

At the interim review at the M18 it was decided to build the DEEP-ER Prototype based on the Aurora Blade architecture to which development Eurotech did commit. The implementation foreseen then comprised the design by Eurotech of a new KNL-board hosting 2 KNL chips, which would constitute logically two fully independent nodes. Detailed analysis of the technical requirements and component placement led however to the conclusion that the complexity and risk of such development would be too high. Alternatively, in M24 Eurotech presented their decision to build a single-node KNL-blade for Aurora, which will integrate a commercially available KNL board (Intel Server board S7200AP or "Adam Pass"). The density of the overall DEEP-ER Prototype is kept by installing both the EXTOLL NICs and the NVM devices in the Root card. With this new approach the design risk is significantly reduced. Eurotech's effort is focused now to the mechanical integration and cooling of all the Aurora Blade components. The Eurotech Aurora cooling technology is being enhanced to support memory DIMMs (of the ultra-low profile or ULPDIMM variety) in addition to soldered-down memory. The best way to cool the ULPDIMMS that will populate the 6 memory slots in each KNL board is under investigation.

The EXTOLL TOURMALET ASIC-based NIC has progressed in the reporting period. A new ASIC stepping (A3) is now available and first tested by the project to integrate Intel Xeon Phi devices on the SDV. Later on, the full interconnect of the SDV was upgraded from A2 to the A3 version. The A3 EXTOLL TOURMALET delivers a stable 8.4 Gbit/s per lane network performance, achieving slightly over 100 Gbit/s raw bandwidth per link and fully matching DEEP-ER requirements.

Further synthetic, application and tools benchmarks have been conducted with the NVM devices installed in the Haswell and KNL nodes of the SDV in Juelich, improving the initial crude application mock-ups and extending the scope of applications considered. It has also become possible to conduct multi-node runs and transition to use the full applications.

Further work was done in close collaboration between UHEI and Micron on improving the HMC controller for current HMC silicon; and the controller design has been put into the open source domain. The NAM architecture had been fixed earlier and the first NAM prototype became available: it uses a state-of-the-art Xilinx FPGA to implement an HMC interface of 16 lanes, the NAM-specific logic that implements the RDMA operations and additional functionality to be provided by the NAM, and a single EXTOLL link of twelve lanes that is compatible with the lane speeds achieved by the TOURMALET ASIC implementation of

EXTOLL. The bring-up phase of the NAM prototype is completed and all the foreseen functions have been implemented and are being tested. Besides the NAM specific function blocks, the EXTOLL link to keep up with the much higher clocked ASIC NICs used by DEEP-ER has been implemented. The FPGA firmware implementation of the NAM is also progressing, taking into account the functionality requirements gathered from and agreed upon with the system and application work packages.

The installation of the Software Development Vehicle (SDV) has been completed: the system uses 16 dual-socket Intel Xeon E5 nodes (Haswell generation), have an Intel DC P3700 NVMe device attached to each node, and use EXTOLL TOURMALET as the interconnect (now upgraded to the A3 version). Pre-release KNL systems have been integrated into the SDV as they became available. The SDV nodes have been procured together with the external storage. It contains a RAID system with 24 hard disks with a total capacity of 144 TByte. A meta-data server and two storage servers orchestrate the system. All three servers host an EXTOLL TOURMALET card and are connected via cables to the EXTOLL NICs of the compute nodes. The SDV will later be connected with the DEEP-ER Prototype and act as the Cluster part of the Cluster-Booster architecture.

### System Software

On the software side, the reporting period focused on the implementation of the various components involved in the I/O and resiliency software stacks. Regular discussions take place between the developers involved to guarantee a coherent and consistent global picture, where all the software components fit together. An overview of the DEEP-ER I/O and resiliency software layers has been given in the DoW and is shown in Figure 3 and Figure 4, respectively.

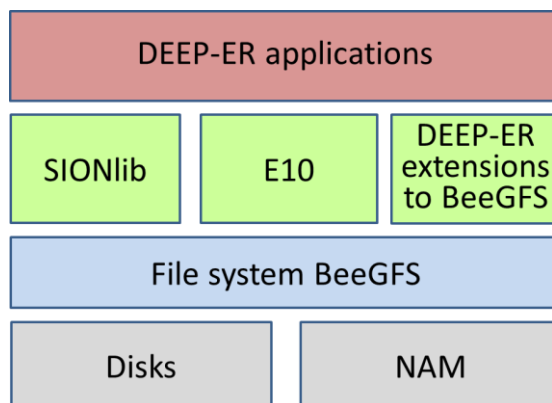


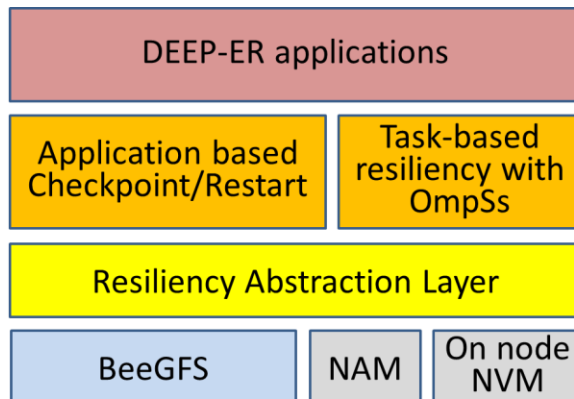
Figure 3: Sketch of DEEP-ER I/O software layers.

In particular, a close interrelation between BeeGFS, SIONlib, and the Scalable Checkpoint/Restart library (SCR) has been established. All three components cooperate to realise buddy checkpoints and the overall checkpointing functionality in an efficient way. Functionality on the SIONlib and BeeGFS sides has been implemented and their integration with SCR is completed.

BeeGFS has completed the synchronous and asynchronous version of its implementation. Furthermore, in tight collaboration with UHEI (WP3), FHG-ITWM is working on native support of EXTOLL by BeeGFS, to solve some network performance issues detected on the SDV.

The functionality required in SIONlib for the efficient implementation of buddy checkpointing is now available. The new features have been released to WP5 for their integration with SCR and verification of the overall usability and performance.

The implementation of the E10 API for I/O is completed and results of its evaluation have been published.



**Figure 4: Sketch of DEEP-ER resiliency layers**

The DEEP-ER resilience architecture is based on user-level checkpoint/restart techniques – which provide a high level of resiliency and are the most cost-effective in terms of I/O requirements– complemented with novel OmpSs task-based recovery techniques. With this combination, DEEP-ER develops new resiliency features to isolate partial failures of the system without requiring a full application restart, resulting in a more resilient, fine-grained and flexible architecture.

Additional to strong cooperation with the I/O software developers, the implementation in the failure recovery software packages themselves has been completed. The final version of the resiliency abstraction layer has been described in D5.2. The abstraction layer adds to SCR specific functions that allow efficiently exploiting DEEP-ER's I/O functionality for checkpoint/restart applications.

The SCR code makes use of the BeeGFS prefetch/flush functionality and already uses symlinks to keep path structures synchronous. Additionally, the buddy-checkpointing functions recently available in SIONlib have been implemented and tested. Currently, work is ongoing to enable the use of NAM for checkpointing combining SIONlib with SCR and libNAM.

Also the task-based resiliency software is complete. The needed adaptations in OmpSs have been done and several use cases have been tested. In particular, the seismic imaging application from BSC is being used to test the task-based resiliency functionality.

Beyond that, the ParaStation management daemon has been extended to provide an interface for querying resiliency-related status information from the MPI layer and thus also from the OmpSs runtime environment. Tight collaboration between ParTec and BSC ensured the complete verification of the combined software, which by now has been already tested and validated.

Finally, the implementation of the failure model is complete and the closed formula has been integrated with the SCR library, enabling the latter to provide the user indications on the required checkpointing frequency for each application.

The software developments in WP4 and WP5 are accompanied by benchmarking activities to document the progress in terms of performance and functionality. The Jülich Benchmarking Environment (JUBE) is used for this purpose. Benchmarks have been implemented in JUBE and are run frequently on the DEEP Cluster and SDV to monitor the I/O performance, as the software is being developed and updates are installed. With this activity some issues on the network performance on the SDV have been identified and finally solved.

### *Applications*

The application developers play a crucial role in the DEEP-ER project. Their work is two-folded: on the one hand they validate the work done by other technical work packages by porting their applications to the DEEP-ER Prototype; on the other hand, their input drives the development and future of both hardware and software architectures. Co-design discussions to gather more specific application requirements take place in the DDG and the consortium face-to-face meetings. Additionally, internal review meetings focused on the work done by WP6 take place during the consortium face-to-face meetings. In the two face-to-face meetings that took place in the reporting period, the developers have described their applications and presented the results that they had recently obtained, as well as the planned next steps. Other members of the consortium not involved in WP6 acted as internal reviewers and gave recommendations on the measures to be taken by each application team to achieve the results needed by the project. Additionally, questions on the specific requirements of each application were discussed, to continue with the co-design approach established in the DEEP-ER project.

In the reporting period the application developers have been performing modifications to adapt the codes to the DEEP-ER hardware and software architecture, as well as general code optimisations. Additionally, the codes have been ported and benchmarked on the SDV. The results achieved are reported in detail in D6.2. Currently focus is on testing the codes on KNL and optimising the parts that will later run on the DEEP-ER Booster.

### **1.3 Expected final results**

The DEEP-ER project will have installed the DEEP-ER Prototype in Jülich (Germany), containing the new generation of Intel Xeon Phi processors, non-volatile memory in the Booster Nodes, as well as additional memory connected to the network. A complete software stack based on ParaStation MPI and OmpSs will run on the machine, with DEEP-ER I/O layers providing parallel I/O functionality and an efficient infrastructure for failure recovery via easy-to-use application interfaces.

Porting and optimising applications on the DEEP-ER Prototype will have demonstrated the scalability and performance of the I/O and resiliency tools developed within the project. The experience gathered will have served to demonstrate that systems using the DEEP-ER results will be able to run more applications in the same time, thus increasing scientific throughput, and that the loss of computational work through system failures will be substantially reduced. Based on these results, performance models will have shown the expected performance at a larger scale system size, demonstrating the capabilities of an Exascale-ready DEEP-ER system.

## Annex A

### A.1 Listing of dissemination activities

This list reflects the dissemination activities performed **between months 13 and 24** of the DEEP-ER project.

#### 1.3.1.1 Conferences, workshops, and meetings:

- **LENS2015 International Workshop**, Oct 29 - 30, Akihabara, Japan
  - Eicker, N., "Taming Heterogeneity by Segregation – Taming Heterogeneity by Segregation -- The DEEP and DEEP-ER take on Heterogeneous Cluster Architectures" (presentation)
- **Supercomputing Conference SC15**, Austin, USA, November 16-19, 2015:
  - Joint booth of the European Exascale Projects (EEP). Booth #197. Participant projects: DEEP-ER, Mont-Blanc, EPIGRAM, and EXA2CT).
  - DEEP and DEEP-ER fliers distributed at the EEP and the partners' booths and on the attendees bag
  - DEEP+DEEP-ER video running at the booth of the European Exascale Projects
  - E.Suarez (JUELICH), presentation on DEEP-ER at the Intel Booth at a session called "An update on European HPC initiatives", November 19, 2015.
  - J. Schmidt (UHEI), "openHMC – Open Source Hybrid Memory Cube Controller" (presentation at the Emerging Technology Track)
  - S. Breuner (FHG-ITWM): BeeGFS presented at FHG-ITWM booth
  - W.Frings (JUELICH): SIONlib presented at JSC and DEEP-ER booths
- **The International Conference on RECONFIGurable Computing and FPGA**, Dec 07-09, Mayan, Mexico
  - J.Schmidt (UHEI), "openHMC – Open Source Hybrid Memory Cube Controller" (poster)
- **AGU Fall Meeting**, Dec 14 2015, San Francisco, USA
  - J. Amaya (KULeuven), "First-principle modeling of planetary magnetospheres: Mercury and the Earth" (poster)
- **The VSC Users Day**, Dec 14 2015, San Francisco, USA
  - J. Amaya (KULeuven), "Fully Kinetic 3D Simulations of the Interaction of the Solar Wind with Mercury" (poster)
- **HPC-LEAP Winter School 2016**, Jülich, Germany, January 15, 2016:
  - E.Suarez (JUELICH), "Implementing a new computing architecture paradigm" (presentation).
- **CeBIT 2016**, Hannover, Germany, March 14-18, 2016
  - DEEP + DEEP-ER topics presented at the booth of the North-Rhein Westfalia.
- **EGU General Assembly**, Vienna, Austria, April 17-22, 2016
  - J.Amaya (KULeuven), "Innovative HPC architectures for the study of planetary plasma environments" (accepted for presentation)
- **Parallel 2016 conference**, Heidelberg, Germany, April 07, 2016
  - C. Clauss & Th. Moschny (ParTec), "Verhalten von MPI Programmen im Fehlerfall (= Behaviors for MPI programmes when errors occur)" (presentation)
- **EGU conference**, Vienna, Austria, April 17-22, 2016
  - J.Amaya (KULeuven), "Innovative HPC Architectures for the Study of Planetary Plasma Environments" (poster presentation)
- **EASC 2016**, Stockholm, Sweden, May 10-14, 2016
  - J.Amaya (KULeuven), "Towards exascale simulations of space plasmas using the DEEP-ER architecture " (accepted for presentation)

- **European HPC Summit Week**, Prague, Czech Republic, May 10, 2016:
  - E.Suarez (JUELICH), “DEEP and DEEP-ER” (presentation).
- **BeeGFS User meeting 2016**, Kaiserslautern, Germany, May 18-19, 2016
  - C.Manzano (JUELICH), “BeeGFS in the DEEP/-ER Project” (presentation)
  - F.Kautz (FHG-ITWM), “BeeGFS User APIs” (presentation)
- **EMIT 2016**, Barcelona, Spain, June 3, 2016
  - E.Suarez (JUELICH), “Technology emerging from the DEEP & DEEP-ER projects” (presentation).
- **ISC 2016**, Frankfurt, Germany, June 20-23, 2016
  - J.Schmidt (JHEI), “Network Attached memory” (presentation at the PhD Forum)
  - N.Eicker (JUELICH), “Hardware Prototyping in DEEP and DEEP-ER” (presentation at Workshop ‘Developing Next-Gen HPC Architectures’)
  - R.Léger (Inria), “A feedback on approaching the DEEP-ER platform with a DGTD-based simulation software for Bioelectromagnetics applications” (presentation at Workshop ‘Form Follows Function’)
  - N.Eicker (JUELICH), (presentation at BoF ‘Exascale I/O: Challenges, Innovations & Solutions’)
  - J.Labarta (BSC), “The OmpSs Programming Model Vision” (presentation at BoF 11 ‘Programming Models for Exascale: Slow Transition or Complete Disruption’)
  - I.Zacharov (Eurotech), (presentation at BoF 15 ‘Monitoring Large-Scale HPC Systems: Data Analytics & Insights’)
  - I.Zacharov (Eurotech), “Aurora Tigon v4 with KNL, a system from research for research” (presentation at the Intel Collaboration Hub (booth #930))
  - E.Suarez (JUELICH), “System-level heterogeneity with Intel Xeon Phi processors” (presentation at Intel booth)
  - Video interview with insideHPC (to be published)
  - Written interview with Top500 blog (to be published)
- **JSC-KIT meeting**, Karlsruhe, Germany, June 30, July 1, 2016
  - E.Suarez (JUELICH), “DEEP/-ER cooling concept” (poster)
- **SAI Computing Conference**, July 15, 2016, London, UK
  - N. Eicker (JUELICH), “Taming Heterogeneity in HPC” (keynote presentation)
  - Th.Moschny (ParTec) and N.Eicker (JUELICH), presence at DEEP-ER booth
- **Kleine Nacht der Wissenschaft**, Jülich, Germany, September 2, 2016:
  - E.Suarez (JUELICH), “The future of supercomputing” (popular science colloquium).

### 1.3.1.2 Publications, proceedings, press-releases, and newsletters:

- **JCP**, “Exactly Energy Conserving Implicit Moment Particle in Cell Formulation.”, G.Lapenta (KULeuven) et al. (submitted)
  - arXiv preprint arXiv:1602.06326
- **CeBIT 2016 press release** (JUELICH), 07/03/2016, “DEEP Project Presents Next-Generation of Supercomputers”
  - <http://www.fz-juelich.de/SharedDocs/Pressemitteilungen/UK/DE/2016/16-03-07deep-cebit.html;jsessionid=A085C64D6693C3289B4ACAFDF4E3E9F5>
- **Primeur Magazine**: “Exascale Project DEEP-ER to present at CeBIT”, 01/03/2016, <http://primeurmagazine.com/flash/AE-PF-03-16-5.html>
- **Science Node**
  - “Boosting Science with the next generation of Supercomputers”

- <https://sciencenode.org/feature/boosting-science-with-the-next-generation-of-supercomputers.php>
- **insideHPC:**
  - Report about BeeGFS goes Open Source: <http://insidehpc.com/2016/02/beegfs-parallel-file-system-now-open-source/>
  - EXTOLL's network chip enables network attached accelerators of any kind, 17/06/2016 <http://insidehpc.com/2016/06/extolls-network-chip-enables-network-attached-accelerators-of-any-kind/>
  - "DEEP-ER Project Moves Europe Closer to Exascale", 04/07/2016 <http://insidehpc.com/2016/07/deep-er-project/>
  - "Taming Heterogeneity in HPC", 17/08/2016, <http://insidehpc.com/2016/08/taming-heterogeneity-in-hpc-the-deep-er-take/>
  - "New Report Looks at European Exascale Projects", 12/08/2016, <http://insidehpc.com/2016/08/european-exascale-projects/>
- **YouTube**
  - EMI 2016: Interview to Estela Suarez, Juelich, 26/07/2016, <https://www.youtube.com/watch?v=5KLORMYW4A4>
  - Taming Heterogeneity in HPC, 06/08/2016, <https://www.youtube.com/watch?v=aM9AkgG5ud4&feature=youtu.be>
- **Top500 Blog:**
  - "A Dive into DEEP-ER: Exascale Research with a distinctly European Flair ", 18/07/2016, <https://www.top500.org/news/a-dive-into-deep-er-exascale-research-with-a-distinctly-european-flair/>
  - "Extoll's network marches to the beat of a different drummer", 25/07/2016 <https://www.top500.org/news/extolls-network-marches-to-the-beat-of-a-different-drummer/>
- **DG Connect**, article on DEEP + DEEP-ER
- **Extoll:** "EXTOLL's network chip enables network attached accelerators of any kind", 16/06/2016, [http://www.deep-er.eu/images/EXTOLL\\_Tourmalet\\_ISC\\_v1.0.pdf](http://www.deep-er.eu/images/EXTOLL_Tourmalet_ISC_v1.0.pdf) (press release)
- **DEEP-ER Status Update:** DEEP-ER on the right path, 17/06/2016, <http://www.deep-er.eu/press-corner/news/184-deep-er-on-the-right-path.html> (press release)
- Eurotech: "Eurotech introduces the Aurora Tigon v4", 21/06/2016, <http://www.eurotech.com/en/press+room/news/?775> (press release)
- **European Exascale Projects:** A Lookback on 5 Years of European Exascale Research Collaboration, [http://exascale-projects.eu/EuroExaFinalBrochure\\_v1.0.pdf](http://exascale-projects.eu/EuroExaFinalBrochure_v1.0.pdf) 17/06/2016, (brochure)
- **Scientific Computing World**, Show preview (ISC 2016), 01/06/2016, print
- **INSIDE**, DEEP projects at CeBIT 2016, 15/06/2016, <http://inside.hlrs.de/#deep-project-at-cebit16>
- **Europa.eu/digitalsinglemarket**, "Europe towards Exascale", 02/09/2016 <https://ec.europa.eu/digital-single-market/en/news/europe-towards-exascale>
- 
- **IEEE Cluster Conference 2016**, September 13, 2016, Taipei, Taiwan
  - Congiu, G. (Seagate) "Improving Collective I/O Performance Using Non-Volatile Memory Devices" (paper and presentation)
  - [https://ssl.linklings.net/conferences/ieecluster/ieecluster2016\\_program/views/at\\_a\\_glance.html](https://ssl.linklings.net/conferences/ieecluster/ieecluster2016_program/views/at_a_glance.html)

### 1.3.1.3 Industry and business cooperation:



- Desktop research:
  - Industrial application fields of the project technology: 2 applications identified
    - Enhancing Oil Exploration (OE)
    - High temperature superconductivity (HTS)
  - Products/services that leverage the project technology and market targets:
    - Present. OE: Seismic analysis/reservoir simulations in frontier domains (Oil&Gas market). HTS: MRI-NMR (medical market).
    - Future. HTS: Magnetic levitation devices, fusion reactors, motors and generators, fault current (transportation, electronics, energy markets).
  - Potential recipients for dissemination activities
    - CAPEX purchase of DEEP-ER system. After benchmarks and proof of concept, for big companies (i.e. oil companies, MagLev trains companies...)
    - Cloud services for SMEs.
- Market analysis structure and set-up, work still on-going.

## List of Acronyms and Abbreviations

### A

**API:** Application Programming Interface.

### B

**BADW-LRZ:** Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften. Computing Centre, Garching, Germany

**BeeGFS:** The Fraunhofer Parallel Cluster File System (previously acronym FhGFS). A high-performance parallel file system to be adapted to the extended DEEP Architecture and optimised for the DEEP-ER Prototype.

**BN:** Booster Node (functional entity)

**BNC:** Booster Node Card is a physical instantiation of the BN

**BoP:** Board of Partners for the DEEP-ER project

**BSC:** Barcelona Supercomputing Centre, Spain

**BSCW:** Basic Support for Cooperative Work, Software package developed by the Fraunhofer Society used to create a collaborative workspace for collaboration over the web

### C

**CINECA:** Consorzio Interuniversitario, Bologna, Italy

**CN:** Cluster Node (functional entity)

**Coordinator:** The contractual partner of the European Commission (EC) in the project

**CP/RS:** Checkpoint / Restart

**CPU:** Central Processing Unit

**CRB:** Customer Reference Board. An early version of a KNL board developed by Intel.

**CRESTA:** Collaborative Research into Exascale Systemware Tools & Applications: EU-funded Exascale project.

### D

**DDG:** Design and Developer Group of the DEEP-ER project

**DEEP:** Dynamical Exascale Entry Platform

**DEEP-ER:** DEEP Extended Reach: this project

**DEEP-ER Network:** high performance network connecting the DEEP-ER BN, CN and NAM; to be selected off the shelf at the start of DEEP-ER

**DEEP-ER Prototype:** Demonstrator system for the extended DEEP Architecture, based on second generation Intel® Xeon Phi™ CPUs, connecting BN and CN via a single, uniform network and introducing NVM and NAM resources for parallel I/O and multi-level checkpointing

**DEEP Architecture:** Functional architecture of DEEP (e.g. concept of an integrated Cluster Booster Architecture), to be extended in the DEEP-ER project

**DEEP System:** The prototype machine based on the DEEP Architecture developed and installed by the DEEP project

## E

- E10:** Exascale 10. Parallel I/O software developed by a consortium of partners around the EOFS community. Partner Xyratex is responsible for the development needed for the DEEP-ER project.
- EC:** European Commission
- EC-GA:** EC-Grant Agreement
- EEP:** European Exascale Projects
- EESI:** European Exascale Software Initiative (FP7)
- EOFS:** European Open File System.
- EU:** European Union
- Eurotech:** Eurotech S.p.A., Amaro, Italy
- Exaflop:**  $10^{18}$  Floating point operations per second
- Exascale:** Computer systems or Applications, which are able to run with a performance above  $10^{18}$  Floating point operations per second
- EXTOLL:** High speed interconnect technology for cluster computers developed by University of Heidelberg
- ETP4HPC:** European Technology Platform for High Performance Computing.

## F

- FhGFS:** Acronym previously used to refer to BeeGFS.
- FLOP:** Floating point Operation
- FP7:** European Commission 7th Framework Programme.
- FPGA:** Field-Programmable Gate Array, Integrated circuit to be configured by the customer or designer after manufacturing

## G

- GRS:** German Research School for Simulation Sciences GmbH, Aachen and Juelich, Germany

## H

- H5hut:** Library implementing several data models for particle-based simulations that encapsulates the complexity of parallel HDF5.
- HDF5:** Hierarchical Data Format: A set of file formats and libraries designed to store and organise large amounts of numerical data
- HMC:** Hybrid Memory Cube
- HPC:** High Performance Computing
- HW:** Hardware

## I

- ICT:** Information and Communication Technologies  
**IEEE:** Institute of Electrical and Electronics Engineers  
**Intel:** Intel Germany GmbH Feldkirchen,  
**IP:** Intellectual Property  
**iPic3D:** Programming code developed by the University of Leuven to simulate space weather  
**ISC:** International Supercomputing Conference, Yearly conference on supercomputing which has been held in Europe since 1986

## J

- JUBE:** Jülich Benchmarking Environment  
**JUDGE:** Juelich Dedicated GPU Environment: A cluster at the Juelich Supercomputing Centre  
**JUELICH:** Forschungszentrum Jülich GmbH, Jülich, Germany

## K

- KNC:** Knights Corner, Code name of a processor based on the MIC architecture. Its commercial name is Intel® Xeon Phi™.  
**KNL:** Knights Landing, second generation of Intel® Xeon Phi™  
**KULeuven:** Katholieke Universiteit Leuven, Belgium

## L

## M

- MIC:** Intel Many Integrated Core architecture  
**Mont-Blanc:** European scalable and power efficient HPC platform based on low-power embedded technology  
**Mont-Blanc 2:** Follow-up project of Mont-Blanc  
**MPI:** Message Passing Interface, API specification typically used in parallel programs that allows processes to communicate with one another by sending and receiving messages  
**MTBF:** Mean Time Between Failures.

## N

- NAM:** Network Attached Memory, nodes connected by the DEEP-ER network to the DEEP-ER BN and CN providing shared memory buffers/caches, one of the extensions to the DEEP Architecture proposed by DEEP-ER  
**NASA:** National Aeronautics and Space Administration, Washington, USA  
**NEF:** Network of European Foundations: name of server where financial data is uploaded to provide it to the EC.  
**NetCDF:** Network Common Data Form. A set of software libraries and data formats that support the creation, access, and sharing of array-oriented scientific data  
**NVM:** Non-Volatile Memory

**NVMe:** NVM Express. Specification for accessing solid-state drives attached through the PCIe bus.

## O

**OEM:** Original Equipment Manufacturer. Term used for a company that commercialises products out of components delivered by other companies.

**OmpSs:** BSC's Superscalar (Ss) for OpenMP

**OpenMP:** Open Multi-Processing, Application programming interface that support multiplatform shared memory multiprocessing

**OS:** Operating System

## P

**ParaStation Consortium:** Involved in research and development of solutions for high performance computing, especially for cluster computing

**ParaStationMPI:** Software for cluster management and control developed by ParTec

**Paraver:** Performance analysis tool developed by BSC

**Paraview:** Open Source multiple-platform application for interactive, scientific visualisation

**ParTec:** ParTec Cluster Competence Center GmbH, Munich, Germany

**PCI:** Peripheral Component Interconnect, Computer bus for attaching hardware devices in a computer

**PCIe:** PCI Express, Standard for peripheral interconnect developed to replace the old standards PCI, improving their performance

**PFlop/s:** Petaflop,  $10^{15}$  Floating point operations per second

**PM:** Person Month or Project Manager of the DEEP project (depending on the context)

**PMT:** Project Management Team of the DEEP-ER project

**PRACE:** Partnership for Advanced Computing in Europe (EU project, European HPC infrastructure)

**PROSPECT:** Promotion of Supercomputing Partnerships for European Competitiveness and Technology (registered association, Germany)

**PTC:** Persistent Task-based Checkpoint

## Q

**QCD:** Quantum Chromodynamics

**QPACE:** QCD Parallel Computing Engine. Specialised supercomputer for QCD Parallel Computing

## R

**R&D:** Research and Development

## S

- SC:** International Conference for High Performance Computing, Networking, Storage, and Analysis, organised in the USA by the Association for Computing Machinery (ACM) and the IEEE Computer Society
- Scalasca:** Performance analysis tool developed by JUELICH and GRS
- SCR:** Scalable Checkpoint/Restart library
- SDV:** Software Development Vehicle: a HW system to develop software in the time frame where the DEEP-ER Prototype is not yet available.
- SEO:** Search Engine Optimisation: the process of improving the visibility of a website or a web page in a search engine's results.
- SSD:** Solid State Disk
- SW:** Software

## T

- TFlop/s:** Teraflop,  $10^{12}$  Floating point operations per second
- ToW:** Team of Work Package leaders within the DEEP-ER project
- TP10:** Third Party under special clause 10.

## U

- UHEI:** University of Heidelberg, Germany
- UREG:** University of Regensburg, Germany

## V

- VI-HPS:** Virtual Institute for High Productivity Supercomputing
- VTune:** Commercial application for software performance analysis

## W

- WP:** Work Package

## X

## Y

## Z