SEVENTH FRAMEWORK PROGRAMME

FP7-ICT-2013-10

DEEP-ER

DEEP Extended Reach

Grant Agreement Number: 610476

D1.5

Periodic progress report at month 24

Approved

Version:	2.0
Author(s):	M. Foszczynski (JUELICH), J.Kreutz (JUELICH)
Contributor(s):	S.Eisenreich (BADW-LRZ), H.Ch.Hoppe (Intel), K.Thust (JUELICH), V.Beltran (BSC), E.Suarez (JUELICH), A.Wolf (JUELICH), I.Zacharov (Eurotech)
Date:	09.12.2015

Project and Deliverable Information Sheet

<table>
<tr><td rowspan="6">DEEP-ER Project</td><td colspan="2">Project Ref. №: 610476</td></tr>
<tr><td colspan="2">Project Title: DEEP Extended Reach</td></tr>
<tr><td colspan="2">Project Web Site: http://www.deep-er.eu</td></tr>
<tr><td colspan="2">Deliverable ID: D1.5</td></tr>
<tr><td colspan="2">Deliverable Nature: Report</td></tr>
<tr><td>Deliverable Level:
CO*
(the present document contains only the public part of the deliverable)</td><td>Contractual Date of Delivery:
30 / September / 2015
Actual Date of Delivery:
30 / September / 2015</td></tr>
<tr><td colspan="2">EC Project Officer: Panagiotis Tsarchopoulos</td></tr>
</table>

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

Document Control Sheet

<table>
<tr><td rowspan="6">Document</td><td colspan="2">Title: Periodic progress report at month 24</td></tr>
<tr><td colspan="2">ID: D1.5</td></tr>
<tr><td>Version: 2.0</td><td>Status: Approved</td></tr>
<tr><td colspan="2">Available at: Publishable part at: http://www.deep-er.eu</td></tr>
<tr><td colspan="2">Software Tool: Microsoft Word</td></tr>
<tr><td colspan="2">File(s):DEEP-ER_D1.5_Periodic_progress_report_M24_v2.0-ECapproved-PublishablePart</td></tr>
<tr><td rowspan="4">Authorship</td><td>Written by:</td><td>M. Foszczynski (JUELICH), J.Kreutz (JUELICH)</td></tr>
<tr><td>Contributors:</td><td>S.Eisenreich (BADW-LRZ), H.Ch.Hoppe (Intel), K.Thust (JUELICH), V.Beltran (BSC), E.Suarez (JUELICH), A.Wolf (JUELICH), I.Zacharov (Eurotech)</td></tr>
<tr><td>Reviewed by:</td><td>I.Schmitz (ParTec), E.Suarez (JUELICH)</td></tr>
<tr><td>Approved by:</td><td>BoP/PMT</td></tr>
</table>

Document Status Sheet

Version	Date	Status	Comments
1.0	30/September/2015	Final	EC submission
2.0	09/December/2015	Approved	EC approved

Document Keywords

Keywords:	DEEP-ER, HPC, Exascale, midterm report, month 24

Table of Contents

List of Figures

List of Tables

No table of figures entries found.

Executive Summary

The DEEP – Extended Reach (DEEP-ER) project started on 1st 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 13 to 24** 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.
- Detailed feasibility study by partner Eurotech on the “Aurora Blade” architecture, including a preliminary design for a KNL-based node board. Commitment by partner Eurotech to produce this design, expressed at the interim review, and decision for DEEP-ER to follow this route. Board design phase ongoing.
- First results of I/O benchmarks and application mock-ups running on the selected Non-Volatile Memory devices.
- Functionality of Hybrid Memory Cube (HMC) controller (required for Network Attached Memory (NAM) prototype) established. The NAM architecture, the design of the first NAM prototype and a first version of the functional specification for the NAM have been completed. The NAM prototype is currently in production. Simultaneously UHEI is working on the design of the EXTOLL link implementation for the FPGA which is required to attach the NAM directly to the EXTOLL network
- Development of the DEEP-ER I/O software stack – containing BeeGFS, SIONlib, and E10 – ongoing, taking into account the requirements from resiliency software and applications. Implementation of node-local I/O caches for BeeGFS.
- Implementation of resiliency software layer progressing well. Discussions between I/O and resiliency software tasks take place to define interfaces and guarantee a coherent development of the full software stack. Models for the benefits and performance impact of task-based resiliency have been created and evaluated.
- Benchmarks integrated in JUBE environment and tests performed: identification of parameters required by BeeGFS to achieve high-speed in metadata handling. Integration of full applications ongoing. Integration of applications in JUBE ongoing.

- Application adaptations and improvements ongoing: optimisation to take benefit from 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.
- Dissemination of project goals and status in various workshops and conferences, amongst others the Supercomputing Conference (SC14) in the US and the International Supercomputing Conference (ISC 2015) in Europe.
- Coordination and co-organisation of joint dissemination activities with other European Exascale Projects, i.e. for the joint booth at SC14 and ISC 2015, and the industry-focussed Exascale workshop at the PRACEdays15.

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 tackle these issues, DEEP-ER will extend the heterogeneous Cluster-Booster Architecture implemented by the DEEP[1] 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 design of the DEEP-ER I/O software layer has been completed taking the application and resiliency software requirements into account (documented in deliverable D4.1). The interfaces between the three I/O APIs - BeeGFS, SIONlib, and Exascale10 (E10) - have been defined and documented in D4.2.

   → The resiliency software layer has been designed (documented in D5.1) taking the application requirements into account. First sketch of the abstraction layer below the user-level checkpoint/restart software is finished.

   → The interplay between I/O and resiliency software and between application-based checkpoint/restart and task-based resiliency software has been agreed.

   → Implementation of I/O and resiliency software is ongoing. Cross-Work Package discussions take place to guarantee a coherent development.

[1] www.deep-project.eu

→ The DEEP-ER Prototype will include high-performance NVM devices attached to the compute nodes that provide opportunities to store checkpoints and buffer I/O data. The parallel I/O implementations and the resiliency software layer will take full advantage of this novel hardware feature.

2. Develop a prototype system of the extended DEEP Architecture that leverages advances in hardware components (Intel's second generation Intel® Xeon Phi™ processors, high-speed interconnects and non-volatile memory devices) to further improve the performance and efficiency of the DEEP-ER Prototype and realise the novel I/O system and resiliency improvements. This prototype will allow proving the viability of the concept for 500 PFlop/s-class of supercomputers.

   → Initially, it was planned to build the DEEP-ER Prototype based on Eurotech's Aurora HiVe concept, defined and documented as the "Brick architecture" in Deliverables D3.1 and D3.2. Detailed investigations into the technical and market risks associated with developing a PCIe add-in card form factor KNL board (self-bootable and with the amount of memory required for the project), led to the project's proposition to adopt a simplified architecture based on an air-cooled, readily available compute board and PCI Express attached NVM and NIC devices. However, this proposal was considered too conservative at the M12 review.

   → Several alternative architectures have been studied in response to the reviewers' recommendations to fully address energy efficiency, density, and cooling. In conclusion to this analysis, in the interim review at M18 it was decided to adopt Eurotech's Aurora Blade architecture, extended by a custom-designed KNL node board.

   → Two important design choices were 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. In addition, a line of SSD replacement devices from Intel was selected as o n-node NVM devices.

   → The Software Development Vehicle (SDV), a hardware platform for development of system (I/O and resiliency) and application software, has been installed at JUELICH to allow WP4, WP5 and WP6 to continue their work until the Aurora Blade Prototype is available.

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, and as described above, 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 with two samples of these devices at Juelich, and a wide range of measurements with I/O benchmarks and application mock-ups are available. These clearly show substantial

performance increases over best-of-breed SSDs, in particular for scenarios with many parallel I/O requests. The SDV recently deployed contains 16 of these NVMe devices for further software development and evaluation. The DEEP-ER Prototype will either use the current product implementation as explained in D3.2, or the recently disclosed Intel® Optane™ follow-on product generation, which will bring further performance improvements.

→ The NAM will use partner UHEI's hybrid HMC controller implementation, which has been completed and functionally validated. Reliability issued noticed during validation could be tracked to early silicon versions of Micron's HMC devices and the FPGA used – they will not apply to the NAM prototype, which is based on a later stepping. Architecture and design of the NAM prototype have been fixed – a state-of-the-art Xilinx FPGA will implement the HMC controller, NAM functional logic and one EXTOLL link compatible with the full EXTOLL TOURMALET fabric speed. Implementation of the NAM prototype is progressing, with the PCB ready and FPGA firmware development underway on a simulation infrastructure.

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 Exascale10 (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. The synchronous version is already available and tested, the asynchronous is under implementation.

   → The needed preparatory phase for the further development of SIONlib is completed. Amongst other improvements, its communication layers have been refactored to eliminate code replications and increase the overall modularity and manageability of the library.

   → Partner Seagate has integrated the new BeeGFS functionalities for cache handling and user-level strip-size definition into E10. Exascale10 integration is now ongoing with a new driver and extensions developed and tested, now supporting caching functionalities for other file systems. A new support library

has been developed to make the integration of the new E10 functionality transparent to applications.

→ 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 already used to regularly monitor the overall I/O performance to measure the impact of the various developments done in the DEEP-ER project. This activity has resulted in the identification of the right BeeGFS parameters to achieve optimal performance for metadata handling. JUBE framework has been further extended by additional applications. A series of MAXW-DGTD I/O experiments have been run in the DEEP Cluster. Currently, work is ongoing regarding their analysis.

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 user-level and the task-based resiliency functionalities, and the interfaces between them have been defined.

   → The Scalable Checkpoint/Restart (SCR) library is being 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). DEEP-ER modifications to the SCR code have been re-organised and cleaned. The code was updated to reflect recent changes in the BeeGFS API. Autotools build process of SCR has been extended by automatic discovery of BeeGFS. Everything has been tested on the Cluster-Part of the current DEEP system.

   → OmpSs adaptations for task-based resiliency are under implementation. In order to extend the task-based implementation to support offloaded tasks, interaction with the ParaStation MPI layer has been thoroughly investigated.

   → Integration between the CP/RS framework, OmpSs and ParaStation MPI has been a subject of debate and two conceivable approaches have been identified. Further discussions within WP5 are ongoing.

   → In close collaboration with WP3, an event Monte Carlo failure model has been designed and implemented. Goal is 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 failure model and closed formula are aligned.

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.

→ In the first months of the project the application requirements – in terms of hardware capabilities, I/O and resiliency functionalities – have been gathered through a questionnaire. DDG teleconferences and face-to-face meetings are used for further co-design discussions, 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.

→ Various improvements are being implemented in the applications. Important topics in the code optimisations are: vectorisation, optimising communication strategies and/or numbering schemes, improving I/O, implementing checkpointing, etc. Further benchmarking and integration is in progress in most applications, as well as practical implementation of Cluster/Booster division and revised checkpointing software.

→ SIONlib and OmpSs are being integrated with several applications.

→ In collaboration with WP3 and WP4, mock-ups from the applications are being prepared to use them for I/O benchmarking. Mock-ups from the space weather and seismic applications are already available.

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.

1.2 Work performed and main results

According to the amended DoW, three milestones were to be reached between **month 13 and month 24** of the DEEP-ER project:

- **MS6** – “SDV, Booster CPU evaluator, NAM functional evaluator and NVM evaluator available”
  - **SDV:** All its hardware components have been procured and installed at JUELICH. Software installation and bring-up is ongoing and should be completed by mid-October.
  - **Booster CPU evaluator**: it is realised by using Intel-supplied KNL Customer Reference Boards (CRB). One early (A0 stepping) CRB has been delivered to Eurotech to enable early firmware development and progress with detailed KNL node board design. A set of engineering samples are planned to be released to Juelich as soon as they become available. These will be of a stepping that allows meaningful SW development work, and they will include a

least 8 GB of on-package MCDRAM and thus be representative of the production CPUs to be used in the DEEP-ER prototype.

- **NAM functional evaluator**: The NAM functional evaluator is realized as a standard height PCI-Express card developed by UHEI. It incorporates a Xilinx Virtex7 FPGA and a Hybrid Memory Cube. The card is currently in production with first hardware tests expected in mid-October 2015. Simultaneously UHEI is working on the design of the EXTOLL link implementation for the FPGA which is required to attach the NAM directly to the EXTOLL network. A first demonstration of the NAM functional evaluator connected to an EXTOLL TOURMALET ASIC is planned for the Supercomputing Conference SC15 in mid- November, 2015.
- **NVM evaluator**: two Intel-supplied SSD replacement devices implementing the NVMe interface are installed at JUELICH since the first months of the project. They have been used for extensive technology evaluation and benchmarking. A further 16 such devices are integrated in the SDV. This will enable software development and also larger scale measurements and evaluation of the NVM technology with benchmarks and applications.

- **MS7** – "Performance extrapolation based on design decisions": the main design decisions taken for the Aurora Blade Prototype have been analysed and their impact on scaling up the DEEP-ER Prototype is discussed in Deliverable D7.1. Using the Dimemas simulation tool from BSC, the behaviour of three scientific applications on the DEEP-ER prototype have been predicted, and their performance has been extrapolated for larger configurations. These results are also included in D7.1, submitted at M24 simultaneously to this report.
- **MS8**: "Overall design of Aurora Blade prototype completed": the feasibility study presented during the review at M18 has been updated with the design decisions taken later on, and submitted in M24 as deliverable D8.1.

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.

The Project Management Team organised the agenda for the first review meeting (at month 12) which took place on the 21st October 2014 in Brussels (Belgium). To fulfil the internal quality policies a rehearsal meeting one day before the review was conducted. As a result of the first review, the project has been evaluated as doing "good progress". Additionally, all deliverables submitted in the first year of the project were approved. All public approved deliverables have been uploaded to the project website. The comments from the reviewers and their recommendations concerning future work are addressed in section 2.2.

Addressing the reviewer recommendations at M12, an important management activity was focused on driving confidential technical discussions with several HPC system manufacturers to identify potential architecture alternatives for the DEEP-ER Prototype. Additionally, the PMT also prepared the agenda and slides for the interim review meeting at M18, focused on

the Prototype architecture, which took place on 26th March 2015 in Brussels. There, the need for a project extension was discussed. An amendment of the Description of Work has been prepared to adapt the project plan and its duration. The new DoW text was discussed within the consortium and the feedback from the partners were taken into account. Its final version was then unanimously approved by the Board of Partners and sent to the Project Officer.

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

Monthly teleconferences of the Team of Work Package leaders (ToW) were organised to periodically discuss the progress in all Work Packages (WPs). Furthermore, two face-to-face consortium meetings took place in the reporting period: one on 29th-30th March 2015 in Jülich (Germany) and one on 8th-9th September 2015 in Garching (Germany). Responsible for the local organisation of the meeting were partners JUELICH and BADW-LRZ, respectively. The agenda minutes were prepared by the PMT, which also chaired the meetings.

Deliverables D4.3, D1.4, and D5.2 have been submitted in time to the European Commission after having passed through the mandatory DEEP-ER internal review process.

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. 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, a face-lift of the project website was performed to make it more appealing to the target audiences. Currently, work is being done on a content plan for the second half of the year to have a steady stream of content and have the website up-to-date for the next big conference SC15 in Austin, Texas in November.

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. Retweets 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 retweets and mentions. LinkedIn is still slower, but postings are more frequent now and also more colleagues engage in the group. The number of members is still rather low. However, 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 one of the most important events in the HPC community: the Supercomputing Conference (SC), which took place in New Orleans (USA) in November 2014, and the International Supercomputing Conference (ISC) in Frankfurt (Germany) in July 2015.

At both SC14 and ISC 2015 the DEEP-ER project co-organised a joint booth, together with other European Exascale Projects (EEP) – DEEP, Mont-Blanc (1 and 2), EPiGRAM, EXA2CT, NUMEXAS, and CRESTA). DEEP and DEEP-ER shared wall space describing the architecture and main goals of the projects, and the HMC controller (the core component of DEEP-ER's NAM), was displayed in the booth. Additionally, a DEEP-ER flyer was prepared and distributed at the EEP booth and at the booths of other project partners. . In addition, the DEEP-ER project was part of a demonstration at the Intel booth. Moreover, at ISC 2015 joint workshop has been organized together with EEP in the last day of the conference.

Figure 1: Joint EEP booth at SC14.

DEEP-ER is actively participating in the preparation of further joint activities with the European Exascale Projects (EEP) community, including the preparation of SC15..

A further highlight of the DEEP-ER dissemination activities achieved in this reporting period is the project video, a joint effort between DEEP and DEEP-ER, which explains in an attractive manner for the general public, the importance of HPC for society, the impact of the two projects, and the value of collaboration. The video is accessible through the project website and has been selected as "visual of the week" by iSGTW and featured prominently on insideHPC.

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, the dissemination work has focused on preparing a satellite event co-hosted at this year's PRACEdays15, taking place on May 26, 2015 in Dublin, Ireland. The event has been organised in collaboration with the other EEP members. During the event, a convincing program with a clear industry focus has been offered, including e. g. a keynote speaker from Airbus (Dr. Eric Chaput). DEEP/-ER's focus was put mainly on the applications that have been part of the project since the very beginning and want to convey the key message that the DEEP/-ER systems are not only ideal for academic but also for industrial users, and a corresponding flyer was distributed.

Currently, work is being done on a more detailed plan for industry and business co-operation.

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 for the DEEP-ER application developers took place in May 2015 in Barcelona (Spain).

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.

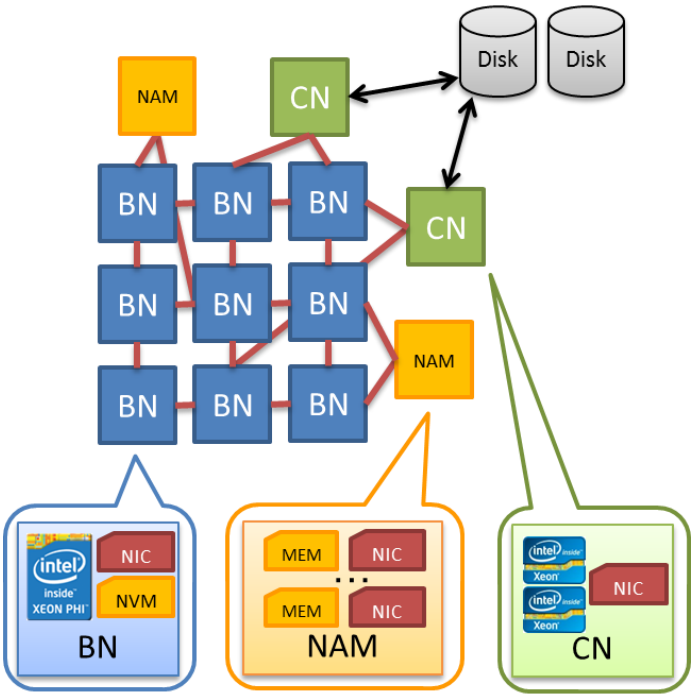

Figure 2: High-level view of the DEEP-ER Prototype. BN=Booster Node; CN=Cluster Node; NVM=Non-Volatile Memory; NAM=Network Attached Memory

The DEEP-ER multi-level I/O infrastructure has been designed to support data-intensive applications and multi-level checkpointing/restart techniques. The project will develop 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 will develop 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 will be 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 will be able to cope with the higher failure rates expected in Exascale systems. DEEP-ER's I/O and resiliency concepts will be evaluated using seven HPC applications from fields that have proven their need for Exascale resources.

System Architecture and New Technologies

Deliverable D3.1 submitted at M6 did specify the innovative "Brick architecture" (named Aurora HiVE by Eurotech), compared it to more conventional approaches and proposed to adopt it for the DEEP-ER Prototype. However, detailed investigations into the technical and market risks associated with developing a commensurate, PCIe add-in card form factor KNL board that satisfies the project requirements (in particular with regards to DRAM memory) led partner Eurotech to decide not to follow this path. As a consequence, Deliverable D3.2 proposed following a conventional approach based on air-cooled, readily available compute boards. The M12 review rejected this approach, and the reviewers recommended re-visiting the architecture: Eurotech should re-consider the decision of developing a KNL board for the Brick, and the project as a whole should search for alternatives.

After the review, Eurotech started a detailed feasibility study on developing a KNL board for DEEP-ER, assisted by partner Intel. Since a board for the Brick architecture commensurate with the DEEP-ER requirements was seen as too risky, the feasibility study proposed an extension of the Aurora Blade architecture by a KNL board and an integration of the NVM and EXTOLL PCI Express cards. This approach does increase the available board area, substantially reducing the technical risk, preserves direct liquid cooling with an option for hot-water cooling, enables flexibility in mixing and matching of KNL and regular Xeon blades, and can support both EXTOLL and InfiniBand NICs.

At the M18 interim review, Eurotech did commit to develop the proposed Aurora Blade architecture, and it was decided to use it for the construction of the DEEP-ER Prototype. Eurotech plans to deploy the DEEP-ER Prototype around M33 (June 2016).

Since the M18 review, excellent progress was made between Eurotech and Intel to discuss and resolve remaining technical issues with the Aurora Blade KNL board. The resulting design foresees to integrate two BNs per board, which means two KNL sockets, two “Wellsburg” chipsets from Intel, two times 96 GB of DDR4 DRAM, and two board management controllers (BMCs). Eurotech will stay close to the Intel reference design, with the prime exception being the use of soldered-on memory (required by liquid cooling and density objectives). This minimizes the technical risk involved in the design.

Work has started on the detailed design, to be carried on in the new WP8 as foreseen by the proposed new statement of work. Intel has supplied one CRB with A0 KNL CPUs to Eurotech, and analysis of the firmware design has started.

The EXTOLL TOUMALET ASIC-based NIC has progressed in the reporting period – a new PCB using more advanced technology was created to allow reliable operation at data rates of ≥ 8Gbit/s per lane, which matches the DEEP-ER requirement of 100 Gbit/s per link. This new EXTOLL TOURMALET NIC has now been validated for the full PCI Express generation 3 line speed (8 Gbit/s per lane), and validation of the target EXTOLL link speed is progressing.

For that reason, the project has decided to adopt EXTOLL TOURMALET as the uniform interconnect. The NICs will be connected via 16 lanes of PCI Express generation 3 to the KNL compute nodes (using Eurotech’s Aurora Blade backplane and root card concept), and standard Samtec cables with HDI6 connectors will be used to wire up the 3D torus topology of the DEEP-ER prototype. Initial work on mechanical and thermal integration of the EXTOLL TOURMALET NICs has started.

Further synthetic, application and tools benchmarks have been conducted with the two NVM devices installed at Juelich, improving the initial crude application mock-ups and extending the scope of applications considered. Once the SDV (see below) becomes available, these measurements can be extended to up to 16 nodes.

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 now the design of the first NAM prototype has been completed: it will use 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 EXTOLL TOURMALET ASIC implementation. PCB design is finished, and first boards are expected shortly.

The FPGA firmware implementation of the NAM is currently work in progress; a description of the initial set of functionality for the NAM prototype was discussed with the system and application SW work packages and agreed upon. Besides the NAM specific function blocks, focus is on enabling the EXTOLL link implementation to keep up with the much higher clocked ASIC NICs used by DEEP-ER.

Sizing and design of the Software Development Vehicle (SDV) was finalized: the system will use 16 dual-socket Intel Xeon E-5 nodes (Haswell generation), have an Intel DC P3700 device attached to each node, and use EXTOLL TOURMALET as the interconnect. KNL CRBs will be integrated into the SDV as they become available – two of these CRBs fit into a 1U 19”slot, and PCI Express add-in cards can be attached using a riser card. The SDV has been procured together with the external storage, which will be used for the SDV and later for the DEEP-ER Prototype. 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 SDV Haswell servers.

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 of the project. Regular discussions take place between the developers involved to guarantee a coherent and consistent global picture, where all the software components fit together. The overviews of the DEEP-ER I/O and resiliency software layers have been described in the DoW and are shown in Figure 3 and Figure 4, respectively.

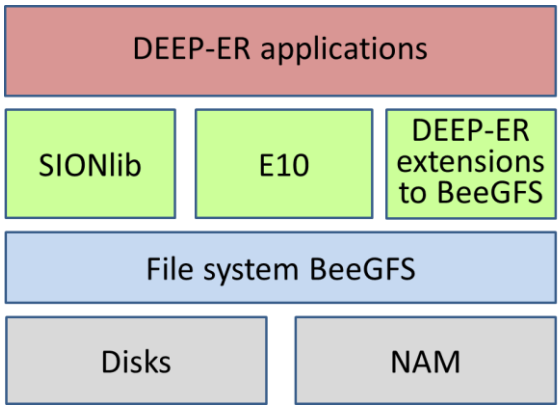

Figure 3: Sketch of DEEP-ER I/O software layers.

In particular, a close interrelation between BeeGFS, SIONlib, and the Scalable Checkcpoint/Restart library (SCR) has been established. All three components will cooperate to realise buddy checkpoints and the overall checkpointing functionality in an efficient way.

BeeGFS has implemented four new functions: a local cache layer in the file system, a capability to define the stripping size at user level, a follow-symlink feature, and a cyclic redundancy check (CRC) checksums calculation. A local cache layer will make use of the node-local NVMe devices to reduce the frequency of I/O to the global file system and increase therefore the overall scalability of the I/O system. Defining stripping size at user level will allow exploiting the knowledge of the user on the data layout in its application to more efficiently determine the physical distribution of the data in the storage. The new follow-

symlink feature allows the user to decide if a symlink should be created in the cache/global file system or if a copy of the destination file is required. The integration of the CRC checksum calculation helps to avoid additional reads from the cache/global file system.

The communication layers of SIONlib have been refactored to improve the modularity and manageability of the library. This preparatory phase is important to ease the integration of the functionality that will be implemented in SIONlib within DEEP-ER. For instance, SIONlib will be used to guarantee that checkpoints to buddy nodes will not generate a very large amount of small files but only a few large physical files, reducing the burden on the file system. Two different approaches to constructing an interface in SIONlib – exploiting BeeGFS's cache domains and local storage in general –, have been prepared. Both options will be implemented when the general buddy checkpointing implementation is finished.

The implementation of the E10 API for I/O has also progressed in the reporting period. The API already integrates the local-cache and user-level data-stripping functionality form BeeGFS. A new driver and extensions have been developed and tested, now supporting caching functionalities for other file systems. A new support library has been developed to make the integration of the new E10 functionality transparent to applications. First measurements performed on the DEEP Cluster show promising results.

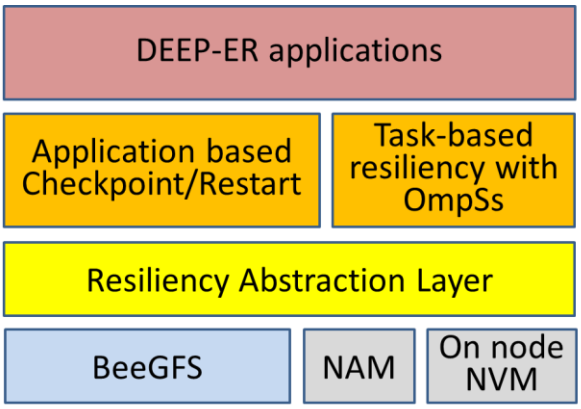

Figure 4: Sketch of DEEP-ER resiliency layers

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 the strong cooperation with the I/O software developers, also the implementation in the pure failure recovery software packages has progressed. A first version of the resiliency abstraction layer has been implemented and 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. A meeting with Adam Moody, main developer of the SCR library, took place during SC14. The ideas behind DEEP-ER and the extensions planned were discussed and Mr. Moody expressed his interest in future development of DEEP-ER's SCR modifications.

At the moment, DEEP-ER operates on a branch of the original SCR code. Significant effort has been spent in re-organising and cleaning the code to allow for a better integration of the

DEEP-ER modifications to the SCR into the original (branched) code. Recent changes to the BeeGFS API have been reflected in the SCR code which makes use of the BeeGFS prefetch/flush functionality and already uses symlinks to keep path structures synchronous.

Progress has been also done regarding the task-based resiliency software. Its design is completed and the implementation is ongoing, performing the needed adaptions in OmpSs and its runtime to support the failure recovery functionality.

Beyond that, the extension of the ParaStation management daemon by an interface for querying resiliency-related status information from the MPI layer and thus also from the OmpSs runtime environment has been envisaged in detail.

Finally, a failure model has been implemented in the form of an event based Monte Carlo simulation, to determine the frequency and location of checkpoints required by a given application. A closed formula has been developed and served to validate the model.

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 to monitor the I/O performance, as the software is being developed and updates are installed. With this activity the parameters required in BeeGFS for an optimal performance of metadata handling have been determined. Work will continue with the integration of mock-ups, proxy- and full-fledged applications in JUBE. The first application (the one from partner Inria) has been already integrated.

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 requierments 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. Some examples are: code refactoring to implement the Cluster-Booster division, multithreading, improvement of vectorisation, integration of SIONlib for I/O and/or checkpointing, porting to Xeon Phi, integration and/or optimisation of OmpSs, etc.

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, 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.

Annex A

A.1 Listing of dissemination activities

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

1.3.1.1 Conferences, workshops, and meetings:

- **Institute for Cyber-Enabled Research at Michigan State University**, 5 November 2014.
  - A.Johnson (KULeuven): "Experience in DEEP/ER with porting iPic3D to MIC, focusing on appropriate use of SoA and AoS representations" (presentation)

- **Supercomputing Conference SC'14**, New Orleans, USA, November 17-20, 2014:
  - Joint booth of the European Exascale Projects (EEP). Booth #1039. Participant projects: DEEP, DEEP-ER, Mont-Blanc, CRESTA, EPiGRAM, and EXA2CT).
  - S.Breuner (FHG-ITWM), W.Frings (JUELICH), K.Thust (JUELICH), N.Eicker (JUELICH), G.Congiu (Seagate), S.Narasimhamurthy (Seagatet). „DEEP-ER I/O: Addressing Exascale I/O problems" (poster at the Emerging Technologies Track)
  - DEEP and DEEP-ER flyers distributed at the EEP and the partners' booths and on the attendees bag
  - Inria's application and its work within DEEP-ER has been presented at the Inria booth
  - E.Suarez (JUELICH): DEEP-ER project, presentation at the Intel booth
  - V.Beltran (BSC): OmpSs Collective Offload, presentation at the Intel booth
  - J.Romein (ASTRON): Correlating Radio Telescope Data for the Square Kilometre Array, presentation at the Intel booth
  - DEEP+DEEP-ER video running at the booth of the European Exascale Projects
  - Soft News for Website**:** DEEP/-ER go DEEP South: Announcement SC14 http://www.deep-er.eu/press-corner/news/51-deep-er-go-deep-south
  - Soft News for Website**:** Gallery on SC14 pictures: http://www.deep-er.eu/press-corner/gallery.html
  -

- **JUELICH-JSC meeting (Visit C.Aubley),** Juelich, Germany, January 19, 2015:
  - E.Suarez (JUELICH), "DEEP and DEEP-ER" (presentation).

- **BDEC (Big Data and Extreme Scaling) Workshop,** Barcelona, Spain, January 28-30, 2015:
  - E.Suarez (JUELICH), "The DEEP (and DEEP-ER) projects" (presentation)

- **SUMA (Supermassive Computations in Theoretical Physics) Workshop**, Trento, Italy, February 11-13, 2015:
  - N. Eicker (JUELICH), "The European Supercomputer Projects DEEP & DEEP-ER" (presentation)

- **PARS Workshop**, **26. GI/ITG Workshop – Parallele Algorithmen, Rechnerstrukturen und Sytemsoftware**, Potsdam, Germany, May 7 – 8, 2015: http://www.cs.uni-otsdam.de/~schnor/potsdam/misc/workshops/2015/pars.html

  - A. Jakobs (JUELICH), “Particle-in-Cell algorithms on DEEP: The iPiC3D case study” (presentation)

- **Internal Training Workshop,** Barcelona, Spain, May 11 – 13, 2015

- **ECL meeting**, Jülich, Germany, May 13, 2015:
  - E.Suarez (JUELICH), “The DEEP and DEEP-ER status” (presentation)

- **JSC-LBL meeting (Visit S. Dosanjh)**, Jülich, Germany, May 22, 2015:
  - N.Eicker (JUELICH), “DEEP and DEEP-ER” (presentation)

- **PRACEdays15 – European Exascale Projects Satellite Event “Enabling Exascale in Europe for Industry”**, Dublin, Ireland, May 26, 2015:
  - E. Suarez, “The European Supercomputer Projects DEEP & DEEP-ER” (presentation with focus on application showcases)

- **2nd International HPC Forum**, 2015, Tianjin China, May 20, 2015:
  - B. Mohr (JUELICH), “Jülich on the Way to Exascale” (presentation)

- **Annual NorduGrid Conference 2015**, Bern, Switzerland, June 4 – 5, 2015: http://indico.hep.lu.se//contributionDisplay.py?contribId=8&sessionId=7&confId=1578
  - E. Suarez (JUELICH), "The DEEP-ER way of approaching Exascale I/O and resiliency" (invited talk)

- **Astronum 2015**, Avignon, France, June 6, 2015: http://irfu.cea.fr/ASTRONUM2015/
  - G. Lapenta (KU Leuven), “Using HPC Kinetic Simulations to Help the MMS Mission Find its Target: Reconnection Diffusion Regions” (presentation)

- **3rd JLESC Workshop**, Barcelona, Spain, 29 June – 1 July, 2015:
  - W. Frings (JUELICH), “Description of buddy checkpointing concepts” (presentation)

- **Nesus Working Group Meeting**, Leuven, Belgium, July 2, 2015
  - D. Alvarez Mallon (JUELICH), “DEEP & DEEP-ER: European HW and SW Innovations on the way to Exascale”

- **ISSS-12, 12th International School/Symposium for Space Simulations**, Prague, Czech Republic, July 3 – 10, 2015
  - J. Amaya (KULeuven), “Fully kinetic simulations of the Solar-Win-Magnetosphere-Interaction” (presentation)

- **ISC’15**, Frankfurt am Main, Germany, July 13 – 16, 2015
  - N. Eicker (JUELICH), “Taming Heterogeneity by Segregation – An Innovative Approach to Heterogeneous Exascale Architectures” presented at EEP workshop
  - E. Suarez (JUELICH), “Architecture Innovation with Intel Xeon PHI” presented at Intel booth
  - H.Ch. Hoppe (INTEL), “DEEP/-ER Applications” presented at Intel booth

- **Workshop – PIC methods for Emerging Architectures**, Santa Monica, USA, July 20 – 31, 2015:
  - G. Lapenta (LEUVEN), “A shootout between explicit and implicit approaches”, (presentation and discussion)

- **13th US National Conference on Computational Mechanics, USNCCM135**, San Diego, California, USA, July 26 – 30, 2015
  - V. Beltran (BSC): "Enabling Complex Applications on Heterogeneous Clusters with OmpSs MPI Offloading" (presentation)

- **CluStor – Workshop on Cluster Storage Technology**, Hamburg, Germany, July 30 – 31, 2015
  - W.Frings (JUELICH), "Buddy checkpointing concepts" (presentation)

- **International Conference on Numerical Simulation of Plasmas**, Golden, Colorado, USA, August 11, 2015:
  - G. Lapenta (LEUVEN), "Using HPC Kinetic Simulations to help the MMS mission find its target: reconnection diffusion regions" (presentation)

- **ParCo2015**, Edinburgh, UK, September 1 – 4, 2015
  - R. Leger (INRIA), "Assessing the DEEP-ER Cluster/Booster Architecture with a Finite-Element Type Solver for Bioelectromagnetics" (presentation)

- **CSP 2015**, Moscow, Russia, September 6 - 10, 2015
  N. Attig (JÜlICH), "Impacts of Current Hardware and Software Developments on Simulation Sciences" (presentation)

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

- **insideHPC:** Experimenting with innovative memory technology, Online at:
- http://insidehpc.com/2014/10/interview-experimenting-deep-er-memory-technology/ (first published on DEEP-ER website)

- **DEEP/-ER Video**, Online at:
  http://www.deep-er.eu/press-corner/news/52-video

- **insideHPC:** DEEP-ER project reaches for Exascale, Online at:
  http://insidehpc.com/2015/02/video-deep-er-project-reaches-for-exascale/

- **iSGTW:** DEEP-ER project reaches for Exascale (also featured as "visual of the week"), Online at:
  http://www.isgtw.org/feed-item/video-deep-er-project-reaches-exascale

- **Scientific Computing World:** Europe's Exascale on Display in May, Online at:
  http://www.scientific-computing.com/news/news_story.php?news_id=2650

- **InsideHPC:** Europe's Exascale on Display in May, Oline at:
  http://insidehpc.com/2015/04/european-exascale-display-may/05/04/2015

- **Primeur Magazine:** Exascale Project DEEP/-ER at ISC'15, Online at:
  http://primeurmagazine.com/flash/AE-PF-05-15-30.html

- **Press Release 1** – DEEP-ER mentioned: Eurotech delivers the Booster system to Jülich to complete the DEEP supercomputer, Online at:
  http://tinyurl.com/pq4qtqp

- **Coverage:**

insideHPC: http://insidehpc.com/2015/07/eurotech-delivers-booster-to-deep-project
HPCwire: http://www.hpcwire.com/off-the-wire/eurotech-delivers-booster-system-to-julich-to-complete-deep-supercomputer/
iSGTW: http://www.isgtw.org/feed-item/eurotech-delivers-booster-deep-project
idw online: https://idw-online.de/en/news634748

- **Press Release 2** – DEEP-ER mentioned: Extoll introduces the HPC network chip TOURMALET , Online at: http://tinyurl.com/nsg36wl

- **Coverage:**
  InsideHPC: http://insidehpc.com/2015/07/extoll-rolls-out-tourmalett-network-chip-at-isc-2015

  Inno Report: http://www.innovations-report.com/html/reports/information-technology/extoll-introduces-the-hpc-network-chip-tourmalet.html

- **Further Press Coverage:**

  Primeur Magazine: "Exascale projet DEEP/-ER at ISC in Frankfurt", 11 June, 2015, Online at: http://primeurmagazine.com/flash/AE-PF-05-15-30.html

  Scientific Computing World: "ISC High Performance Preview", 18 June, 2015, Online at: http://www.scientific-computing.com/show/show.php?show_id=28

  insideHPC: "DEEP moves towards Exascale at ISC'15", 3 August, 2015, Online at: http://insidehpc.com/2015/08/deep-project-moves-towards-exascale-at-isc-2015

  Primeur Magazine: "European Exascale Projects DEEP-ER and Mont-Blanc to Investigate New Exascale Technologies" (Video), 10 August, 2015, Online at: http://www.youtube.com/watch?t=16&v=tr_co6vu-4s

  Primeur Magazine: "European Exascale Projects DEEP-ER and Mont-Blanc to Investigae New Exascale Technologies" (Article), 10 August, 2015, Online at: http://primeurmagazine.com/weekly/AE-PR-09-15-48.html

  Primeur Magazine: "Demonstration: technology developments of the exascale project DEEP-ER" (Video), 10 August, 2015, Online at: http://www.youtube.com/watch?v=aA42Me4s-nI

  Mont-Blanc: "European Exascale Projects at ISC'15 – Day 1" (Video), Online at: http://www.youtube.com/watch?v=ZP4uUzG7GJc&feature=youtu.be

- **DEEP-ER Dissemination Material**
  - Complete re-worked version of project flyer for SC14 (slight adaptions and re-print for ISC'15 and SC15)
  - Created flyer directed at users of a DEEP/-ER system: "Enabling HPC Applications For Exascale" (available for download here: http://www.deep-er.eu/images/materials/DEEP_ER_User_web_062015.pdf)
  -
- **DEEP-ER Website**

- **Complete re-launch / facelift of DEEP-ER project website, including SEO components / optimisation. Go-live: July 14 2015**
- New content:
  Re-cap on DEEP-ER training session in Barcelona in May 2015
  - Re-cap on PRACEdays
  - ISC'15 highlights announcements
  - ISC'15 gallery
  - DEEP/-ER: From A Concept to Application Reality – featuring PREZI presentation
  - Re-writing / Updating of existing content in the course of the website re-launch

1.3.1.3 Participation at industry and business cooperation related events:

- ETP4HPC Working Group on Extreme Scale Demonstrators EsD, Brussels, Belgium, January 14, 2015
- ETP4HPC Steering Board, Teleconference, February 18, 2015
- ETP4HPC Steering Board, Barcelona, Spain, March 5, 2015
- ETP4HPC General Assembly, Barcelona, Spain, March 5, 2015
- PROSPECT General Assembly, Munich, Germany, March 25, 2015
- ETP4HPC Steering Board, Barcelona, Spain, September 7, 2015
- ETP4HPC General Assembly, Rome, Italy, September 29, 2015

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

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 1018 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

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
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)

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