Projects

DEEP Project successfully completed

The DEEP (Dynamical Exascale Entry Platform) project [1] has come to an end. In its final review, which took place in Jülich on the 2nd October 2015, DEEP was evaluated as a very successful project with a lasting impact in the European HPC landscape.

DEEP ran from December 2011 until August 2015. It is one of the first "European Exascale projects" selected by the European Commission through the FP7 funding programme, and received over 8 million Euro in funding. Led by the Jülich Supercomputing Centre (JSC), the consortium of 16 partners from European HPC industry also and academia included the Leibniz Supercomputing Centre (LRZ).

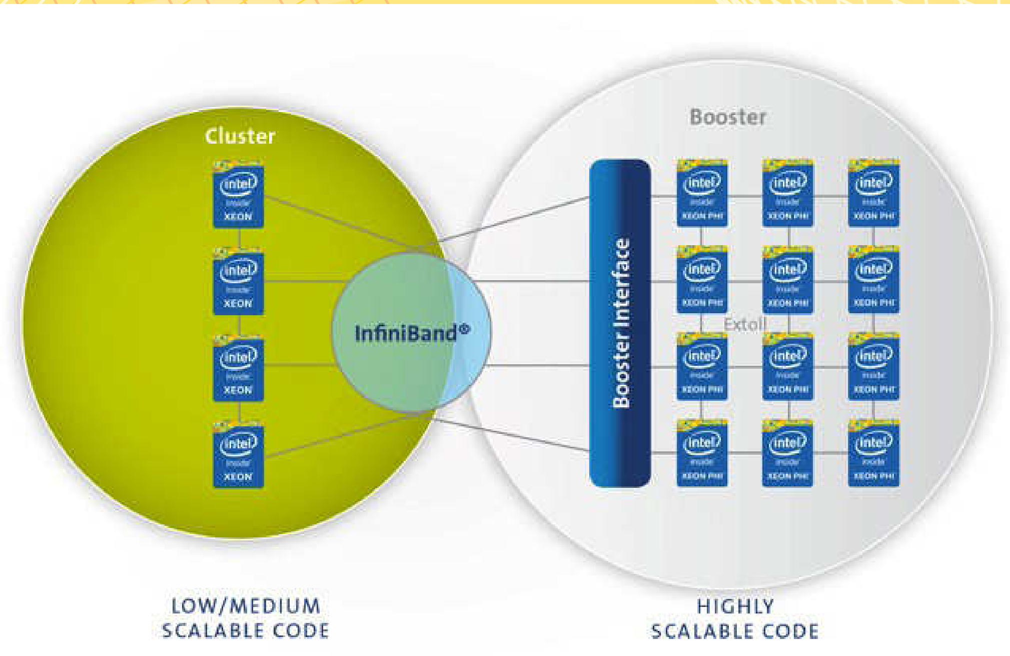

Figure 1: Cluster-Booster architecture in DEEP

The Cluster-Booster Architecture

Striving at pushing the applications' scalability to the limits, DEEP proposed a new approach to heterogeneous computing that best matches the different intrinsic concurrency levels in large simulation codes. The Cluster-Booster architecture [2] combines two distinct hardware components in a single platform:

- the Cluster – equipped with fast, general-purpose processors that show highest single thread performance but with a limited total number of (expensive) cores and less energy efficient

- and the Booster – composed of many-core Intel® Xeon Phi™ coprocessors connected by the EXTOLL [3] network, all together most energy efficient, highly scalable and massively parallel.

In DEEP, code parts of an application that can only be parallelized up to a limited concurrency level run on the Cluster profiting from its high single-thread performance, while the highly parallelizable parts of the simulation exploit the high scalability and energy efficiency of the Booster.

Hardware Prototypes

To demonstrate the Cluster-Booster concept, the DEEP project built hardware prototypes that fully leverage leading-edge multi-core and many-core processors, interconnects, packaging and cooling methods and monitoring/control approaches.

The DEEP Cluster consists of 128 dual-socket blade nodes integrated in Eurotech's proven and highly efficient off-the-shelf Aurora technology. Its direct liquid cooling technology enables year-round chiller-less cooling of the system.

The DEEP Booster, on the other hand, has been designed and built from scratch within DEEP and is also based on Eurotech's Aurora technology. With 384 first-generation Intel® Xeon Phi™ coprocessors, the DEEP Booster is one of the largest Xeon Phi based systems in Europe, with a peak performance of up to 460 TFLOP/s. But even more importantly, this prototype is different from anything seen in the HPC landscape until now: it is the only platform world-wide in which the Xeon Phi processors do operate autonomously without being attached to a host. This provides full flexibility in configuring the right combination of Cluster and Booster nodes, to optimize the use of the hardware for each application.

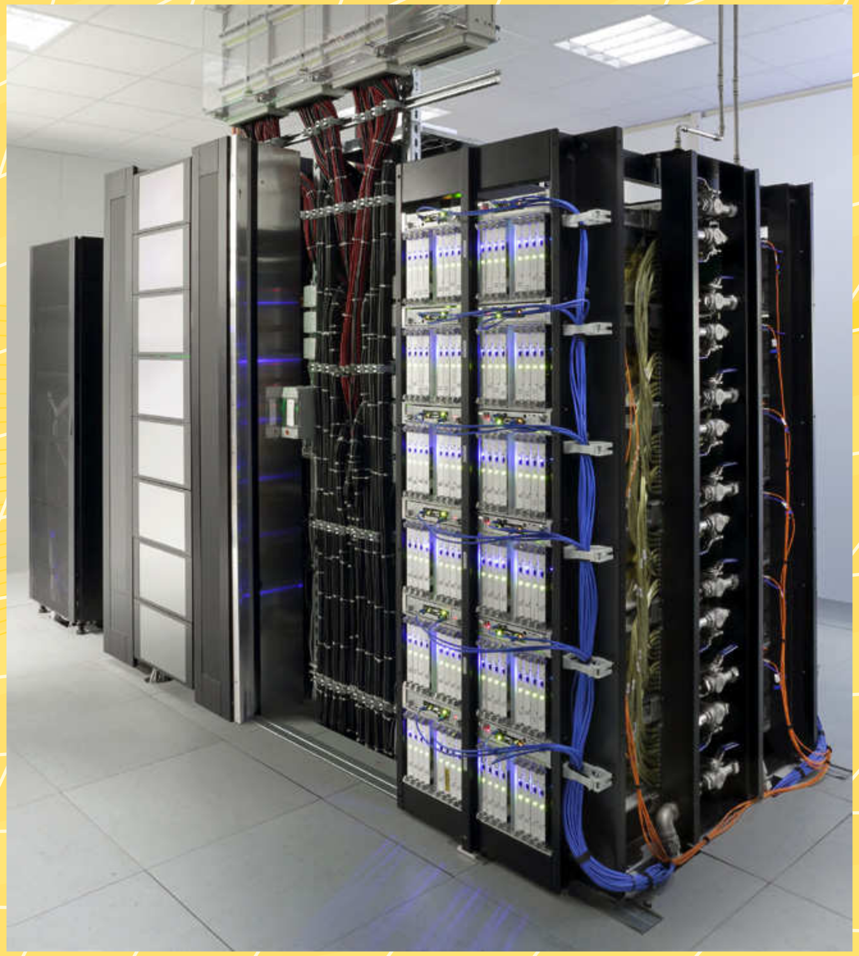

Figure 2: DEEP System. On the left side the Cluster; on the right side (with blue cabling): the Booster

Additionally, two smaller prototypes have been constructed in DEEP:

- The Energy Efficiency Evaluator: a smaller DEEP System (4 Cluster nodes + 8 Booster nodes) installed at LRZ to perform energy efficiency experiments

- The GreenICE Booster: a 32-node (Intel Xeon Phi) Booster prototype built by University of Heidelberg and Megware to test the ASIC implementation of the EXTOLL interconnect, and explore immersive cooling.

Energy Efficiency

The DEEP prototype is based on technologies that reduce the system's energy consumption and help users optimize and tune the system according to their needs. Partner LRZ developed in DEEP comprehensive monitoring capabilities that integrate system and infrastructure sensors, enabling an in-depth analysis of the machine's operating conditions.

The DEEP monitoring system delivers a wealth of voltage, current and temperature data for all system components at high frequency. The innovative monitoring and control hardware and software infrastructure prototyped in DEEP has created substantial progress in the field, showing how high-frequency sensor data can be collected and processed in a scalable way, and how it can effectively interact with the firmware of the system components to ensure safe and efficient operation.

Software Developments

Programming a heterogeneous system like DEEP is a challenging task for developers of HPC applications. To minimize the effort of porting existing applications to the Cluster-Booster architecture, special emphasis was put in the DEEP project on developing a programming model that gives as much support as possible to the users. To achieve it, the team performed a tremendous co-design effort in collaborating with the hardware and the applications team.

While traditional supercomputers are either totally homogeneous or heterogeneous only at the node level, the DEEP system is heterogeneous at the cluster level, combining different kinds of nodes and networks. In order to hide this complexity from the application developers, the software stack implements two abstraction layers:

- ParTec's ParaStation MPI serves as the basic parallelization layer and was extended into a Global MPI covering both Cluster and Booster

- OmpSs from the Barcelona Supercomputing Centre (BSC) was chosen as programming model and extended to provide flexible and powerful offload features.

The DEEP system software and programming model allow computing tasks to be distributed dynamically to the most appropriate parts of the hardware to achieve highest computational efficiency. Based on existing standards and product quality solutions, extensions were made where necessary to make the unique DEEP features available or enhance the ease of programming. Today, it provides a solid base for increasing the circle of applications optimized for heterogeneous architectures in general, and in particular for the DEEP-ER project.

Furthermore, proven performance analysis and modelling tools from JSC and BSC were extended in the DEEP project to fully support the programming models; they were also used to predict the performance of scaled-up systems, establishing a precedent for full system performance projection in the scaling dimension without the need to first create analytical application models [4].

Applications

Six pilot applications were selected to investigate and demonstrate the benefits of combining hardware, system software and the programming model to leap beyond the limits of Amdahl's Law. During the project they were highly optimized, acted as drivers for co-design leading to the final realization of hardware and software in the project, and served to identify the main features of applications that most benefit from the DEEP concept [5].

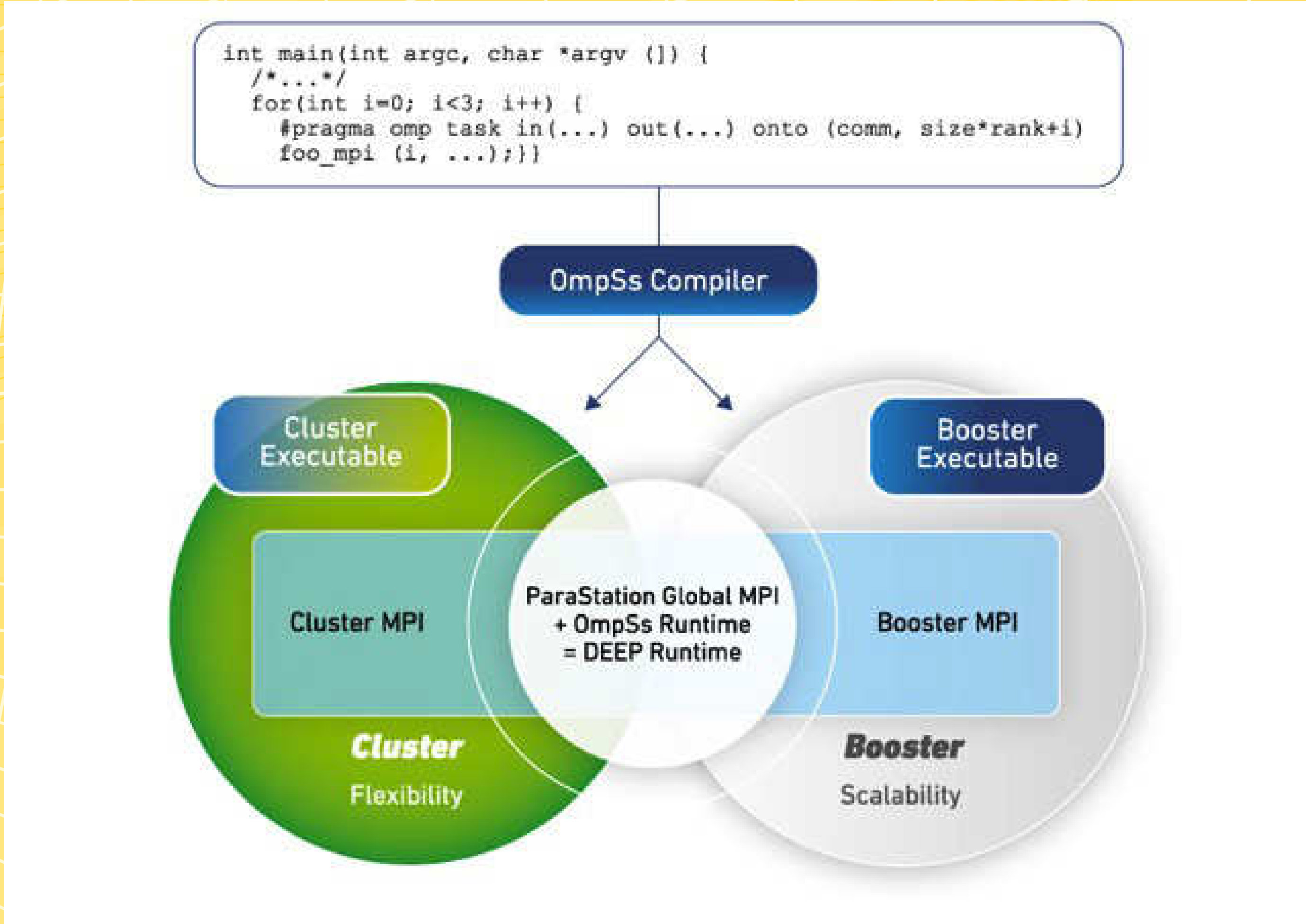

Figure 3: Application running on the DEEP programming environment

Every application is different and therefore needs to be considered as a different use case. However, the project delivered impressive evidence of the number of ways HPC applications can benefit from the flexibility of the DEEP hardware and software architecture. For instance, reverse offloading [6] (Booster to Cluster), I/O offloading and dynamic offloading of discrete tasks are all possible on a DEEP machine, and can easily be ported to other systems.

Across all applications it could be shown that only a limited amount of changes are necessary to benefit from the Cluster-Booster architecture. Furthermore, since the DEEP software interfaces are based on standards, the DEEP-enabled codes continue to run on conventional architectures, sometimes showing surprising performance and efficiency improvements compared to their old formulation.

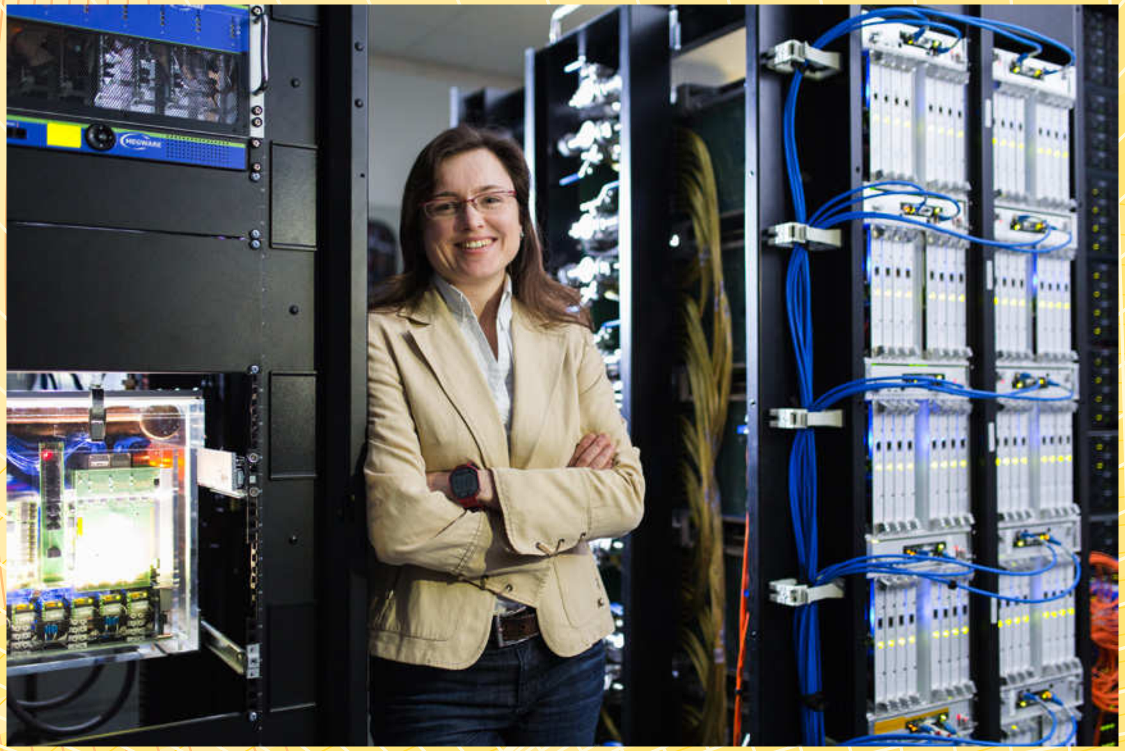

Figure 4: Project Manager Dr. Estela Suarez standing between the two DEEP Boosters. (Copyright: Forschungszentrum Jülich/Sascha Kreklau)

Even more importantly, the experience gathered in the application analysis and adaptation was distilled into "best-known methods", resulting in a playbook for tackling a wide range of additional applications and preparing them for DEEP-class systems. It is our hope that this will have a profound beneficial effect on the entire application ecosystem.

Impact

With its key achievements and the large body of expertise created, the DEEP project is poised to have a significant and lasting impact. Besides opening up new avenues for the architecture of efficient HPC systems, it has materially increased Europe's indigenous capabilities in HPC system design and production, and has produced a complete system software stack together with a programming environment for heterogeneous platforms. Six relevant applications in critical fields of the European Research Arena have been remodeled and adapted, and what is more, best-known methods have been established that will enable many more codes to reap the benefits of the DEEP software and hardware architecture.

The DEEP system has proven that the Cluster-Booster architecture concept of dynamically associating different kinds of computing resources to best match workload needs can be implemented with state-of-the-art multi-core and many-core technology, and that such a system can indeed provide a superior combination of scalability and efficiency. It has thereby opened up a new avenue towards affordable, highly efficient and adaptable extreme scale systems (up to Exascale-class), merging the hitherto separate lines of massively parallel and commodity Cluster systems. The sibling project DEEP-ER is already carrying the flag

further by integrating novel memory and storage concepts and providing scalable I/O and resiliency capabilities [7].

References

[1] The DEEP project, http://www.deep-project.eu

[2] Eicker, N., Lippert, T., Moschny, T., Suarez, E., The DEEP Project An alternative approach to heterogeneous cluster-computing in the many-core era, Concurrency and computation (2015) DOI: 10.1002/cpe.3562

[3] Neuwirth, S., Frey, D., Nuessle, M., Bruening, U., Scalable communication architecture for network-attached accelerators, High Performance Computer Architecture (HPCA), 2015 IEEE 21st International Symposium on; 7-11 February 2015, IEEE, 2015, pp. 627-638, DOI: 10.1109/HPCA.2015.7056068

[4] Rosas, C., Giménez, J., Labarta, J., Scalability prediction for fundamental performance factors, Supercomputing frontiers and innovations 1 (2) 2014, 4-19, DOI: 10.14529/jsfi140201

[5] Alvarez Mallon, D., Lippert, T., Beltran, V., Affinito, F., Jaure, S., Merx, H., Labarta, J., Staffelbach, G., Suarez, E., Eicker, N., Programming Model and Application Porting to the Dynamical Exascale Entry Platform (DEEP), Proceedings of the Exascale Applications and Software Conference, Edinburgh, Scotland, UK. 9-11 April 2013

[6] Jakobs, A., Zitz, A., Eicker, N., Lapenta, G., Particle-in-Cell algorithms on DEEP: The iPiC3D case study, Proceedings of the 26. GI/ITG Workshop Parallele Algorithmen, Rechnerstrukturen und Systemsoftware (PARS), Potsdam, 7-8 May 2015, PARS-Mitteilungen 32 (2015), p. 38-48

[7] The DEEP-ER Project, http://www.deep-er.eu

• Estela Suarez

Jülich Supercomputing Centre (JSC), Germany

contact: Estela Suarez, e.suarez@fz-juelich.de