Towards flexible ExaScale Computing

Norbert Eicker
DEEP-EST project, JSC & BU Wuppertal

19 June 2019
The DEEP Projects

Research & innovation projects co-funded by the European Union
DEEP-EST (2017–2020): generalized modular supercomputing architecture, support for data analytics and ML
Co-design leading to novel, highly efficient, heterogeneous HPC architectures
One of only two Exascale project series funded by Brussel
Total funding: 30 M€
Motivation – One Size does not fit All

HPC systems come in two very different flavours

- General purpose Clusters with
  - High flexibility & reliable performance
  - Preferred by many applications since “good enough” performance is easy to achieve
  - Relatively high power consumption

- Dedicated, highly scalable HPC systems (MPPs)
  - Highest degree of parallelism, specialized fabrics
  - Few (highly parallel) codes can fully exploit them
  - Highly energy efficient

How to combine the best of these worlds into a single system?

Exploit system-level heterogeneity!
Motivation – Application View

Space Weather simulation (xPiC) from KU Leuven

Algorithm iterates between computing magnetic field and particle movement

- Field equations run best on Intel® Xeon® CPUs
- Particle solver & moment computation runs best on Intel® Xeon Phi™
DEEP-ER – Architecture & Results

Implemented & validated in DEEP and DEEP-ER

⇒ 35% performance gain
Combine HPC and ML in a Workflow

Data Analytics
Forecast solar wind conditions at L1 by remote image analysis

Space Weather Prediction
Detailed physics simulations of Earth's environment given solar wind conditions at L1

DL analysis triggers simulation

Observation satellites orbiting L1 point
Modular Supercomputing Architecture

**Cluster Module**
- High single thread performance
- Low/medium scalable code parts

**Extreme Scale Booster**
- Scalability and energy efficiency
- Highly scalable code parts

**Data Analytics Module**
- Large + fast memory + accelerators
- Data intensive and ML codes

**CN**: Cluster Node (general purpose processor)
**BN**: Booster Node (autonomous accelerator)
**DN**: Data Analytics Node (CPU + accelerators)
**NAM**: Network Attached Memory / **GCE**: Global Collective Engine
**SN**: Storage Node / **FE**: Front-End
Co-design Approach

Characterization and Requirements

DEEP-EST HW/SW architecture

Available Technologies

Analyze

Define

Feedback

Derive

Develop

Build

Evaluate

DEEP-EST prototype

DEEP-EST – Towards flexible ExaScale Computing

19 June 2019
European System Technology

**EXTOOLL TORMALET**
Network and Fabri³ nexus

**MEGWARE**
Packaging, cooling and power distribution

**ParaStation**
Cluster Middleware

- TOURMALET ASIC
- MEGWARE SlideSX-LC®
- MEGWARE ColdCon®
- MEGWARE ClustSafe®

ParaStation MPI / Resource Management
The DEEP journey

Starts with the Cluster-Booster architecture
- Enables flexible association and use of heterogeneous resources
- Accelerates heterogeneous applications
- Extends standard APIs

Continued with the DEEP-ER system
- Integrate local and network attached storage
- Provide scalable I/O and check pointing

Next destination: **Modular Supercomputer Architecture**
- Generalizes the Cluster-Booster concept
- Supports HPC, data analytics and machine learning workloads
- First implementation to be ready early 2020

**First production use of Cluster/Booster concept by JSC**
- JURECA Cluster+Booster in operation since 2018 / JUWELS Cluster started in 2018
The DEEP projects have received funding from the European Union’s Seventh Framework Programme (FP7) for research, technological development and demonstration and the Horion2020 (H2020) funding framework under grant agreement no. FP7-ICT-287530 (DEEP), FP7-ICT-610476 (DEEP-ER) and H2020-FETHPC-754304 (DEEP-EST).