1st Workshop on Novel Data Management Ideas on Heterogeneous (Co-)Processors (NoDMC)

David Broneske[1], Dirk Habich[2]

The key objective of database systems is to reliably manage data, where high query throughput and low query latency are core requirements. To satisfy these requirements, database system constantly adapt to novel hardware features. Following that trend, the focus of this one-day workshop is to explore challenges and opportunities of data processing on existing and future heterogeneous hardware architectures. In detail, today's processors are no longer mainly bound by the density and frequency of transistors, but by their power and heat budgets. The so-called "power wall"forces hardware suppliers to rely more on the design of specialized devices optimized for certain types of calculations, which results in an increasingly heterogeneous hardware landscape. Therefore, to meet the above mentioned requirements in our data-driven world, tomorrow's database systems will have to exploit and embrace this increased heterogeneity.

The purpose of this workshop is to assist in the training and growth of a community of researchers and industry practitioners working on data (co-)processing issues on heterogeneous systems. To this end, we want to provide a forum to discuss challenges, progress and directions, while creating an environment for networking with people working on related topics and fostering future collaborations. Especially in the presence of the SPP 2037 on *Scalable Data Management for Future Hardware*, we want to strengthen collaborations beyond single SPP projects by bringing them into contact with other researchers. Moreover, the workshop is co-organized by the GI-Arbeitskreis *Data Management on Modern Hardware.*

The scope of the workshop includes, but is not limited to:

- Applications of modern hardware in
  - data mining
  - data-intensive machine learning
  - query processing
  - non-traditional applications (e.g. graph processing)

[1] University of Magdeburg, Germany, david.broneske@ovgu.de
[2] TU Dresden, Germany, dirk.habich@tu-dresden.de

doi:10.18420/btw2019-ws-01

- Algorithms and data structures for efficient data processing on and across different (co-)processors (e.g., GPUs, APUs, Accelerator cards, FPGAs)
- Exploitation of specialized ASICs
- Efficient memory management, data placement and data transfer strategies in heterogeneous systems
- Energy efficiency in heterogeneous (co-)processor environments
- Programming models and hardware abstraction mechanisms for writing data-intensive algorithms on heterogeneous hardware
- Query optimization, cost estimation and operator placement strategies for heterogeneous hardware
- Transaction processing in heterogeneous systems

With the given scope of the workshop, we are happy to announce a great program. The workshop starts with a keynote by Eric Focht from NEC Deutschland GmbH. He gives an insight into their new vector engine NEC Aurora TSUBASA and how to exploit the design for data management tasks. From the submissions, we were able to accept three technical papers (presented in Session 2) as well as four extended abstracts (presented in Session 3). In Session 2, the first talk by Pietrzyk et al. investigate the SX-Aurora TSUBASA processor for data intensive operations. Furthermore, Becher et al. present a query processing platform for a heterogeneous CPU/FPGA hardware system with a special focus on query placement strategies in such a heterogeneous system. The third technical paper by Götze et al. focuses on problems and possible design decisions for a transactional stream processing system on modern heterogeneous hardware. They especially consider many-core CPUs and non-volatile memory (NVM) and include some considerations on high bandwidth memory (HBM) and co-processors.

In Session 3, four extended abstracts are presented. The first (invited) extended abstract is by Carsten Binnig, who presents his CIDR 2019 paper about a new Data Processing Interface (DPI) for easy usage of RDMA in data intensive applications. Afterwards, Ziegler et al. present a scalable approach for query execution in distributed systems using RDMA under skewed workloads. The key idea is a clever data partitioning between storage and compute nodes and to enable work stealing between compute nodes. The third extended abstract is by Breßet al. presenting their VLDB Journal article. The main focus is on compiling queries for heterogeneous (co-)processors using processor-specific code optimizations to maximize the performance of these queries. The workshop program closes with the presentation by Schmidt et al. proposing an adapted data-placement algorithm for heterogeneous systems.

Last but not least, we like to thank everyone who contributed to this workshop, in particular, the authors, the reviewers, the BTW team, and all participants.

PC Chairs

- David Broneske (University of Magdeburg)
- Dirk Habich (TU Dresden)

Steering Committee

- Wolfgang Lehner (TU Dresden)
- Gunter Saake (University of Magdeburg)
- Kai-Uwe Sattler (TU Ilmenau)

Program Committee

- Carsten Binnig (TU Darmstadt)
- Sebastian Breß (DFKI Berlin)
- Matthias Böhm (IBM Almaden, Uni Graz)
- David Broneske (University of Magdeburg)
- Dirk Habich (TU Dresden)
- Constantin Pohl (TU Ilmenau)
- Hannes Rauhe (SAP SE)
- Knut Stolze (IBM Germany)
- Jens Teubner (TU Dortmund)