



## Editorial

Kai-Uwe Sattler<sup>1</sup> · Alfons Kemper<sup>2</sup> · Theo Härder<sup>3</sup>

© Der/die Autor(en) 2020, korrigierte Publikation 2021

### 1 Schwerpunktthema: Data Management for Future Hardware

Seit 2017 werden im DFG-Schwerpunktprogramm SPP 2037 „Scalable Data Management for Future Hardware“ Methoden und Techniken erforscht, die neue und zukünftige Hardware-Entwicklungen für Datenmanagementsysteme nutzbar machen sollen. Die Bandbreite moderner Hardwaretechnologien, die dabei betrachtet werden, reicht von Co-Prozessoren als Beschleuniger wie GPUs oder FGAs über Highspeed-Netzwerktechnologien wie RDMA bis hin zu modernen Speichertechnologien wie SSDs und persistentem Hauptspeicher. Diese werden im Kontext von klassischen Datenmanagementlösungen wie SQL-Datenbanken, aber auch für Nicht-Standard-Anwendungen wie Event Stores oder Datenstromsysteme eingesetzt. Arbeiten aus diesem SPP bilden das Schwerpunktthema dieser Ausgabe.

Im ersten Beitrag *RDMA Communication Patterns: A Systematic Evaluation* untersuchen Tobias Ziegler (TU Darmstadt), Viktor Leis (Uni Jena) und Carsten Binnig (TU Darmstadt), wie Remote Direct Memory Access (RDMA) effizient für Scale-Out-Datenbanksysteme genutzt werden kann. Dazu betrachten sie den Lösungsraum aus Kommunikationsprimitiven, Kommunikationsmodell und -topologie und präsentieren Ergebnisse einer praktischen Evaluation.

---

✉ Kai-Uwe Sattler  
kus@tu-ilmenau.de

Alfons Kemper  
alfons.kemper@in.tum.de

Theo Härder  
haerder@cs.uni-kl.de

- <sup>1</sup> FG Datenbanken und Informationssysteme, TU Ilmenau, 98684 Ilmenau, Deutschland
- <sup>2</sup> Institut für Informatik, TU München, 85748 Garching, Deutschland
- <sup>3</sup> AG Datenbanken und Informationssysteme, TU Kaiserslautern, 67663 Kaiserslautern, Deutschland

Moderne Speichertechnologien wie SSDs und persistenter RAM werden im Beitrag *Designing an Event Store for a Modern Three-layer Storage Hierarchy* von Nikolaus Glombiewski (Uni Marburg), Philipp Götze (TU Ilmenau), Michael Körber, Andreas Morgen und Bernhard Seeger (Uni Marburg) genutzt, um einen effizienten Event Store – einen persistenten Datenspeicher für temporale Datenströme – zu realisieren. Zur gleichzeitigen Unterstützung von Einfügeoperationen und Ad-hoc-Anfragen werden verschiedene Architekturen für mehrstufige Speicherhierarchien betrachtet.

Im Beitrag *MXKERNEL: A Novel System Software Stack for Data Processing on Modern Hardware* stellen Jan Mühlig (TU Dortmund), Michael Müller, Olaf Spinczyk (Uni Osnabrück) und Jens Teubner (TU Dortmund) eine Softwareplattform vor, die eine Task-basierte Parallelisierung unterstützt. Hierbei können Tasks mit Metadaten annotiert werden, die Speicherzugriffsverhalten und Abhängigkeiten zwischen Tasks beschreiben. Am Beispiel einer B<sup>link</sup>-Baum-Implementierung demonstrieren sie, wie Aspekte wie Concurrency Control und Scheduling unter Nutzung dieser Abstraktion realisiert werden können.

Mit dem Problem der Replikation von Zustandsmaschinen zur Erhöhung der Fehlersicherheit großer verteilter Systeme beschäftigt sich der Beitrag *Towards Log-Less, Fine-Granular State Machine Replication* von Jan Skrzypczak und Florian Schintke (Zuse-Institut Berlin). Die Autoren stellen ihre Erweiterung des Konsensprotokolls Paxos und die Nutzung von persistentem Speicher zur Realisierung von In-Place-Änderungen als Alternative zum klassischen Logging vor.

### 2 Fachbeitrag

In den letzten Jahren nutzen mehr und mehr relationale Datenbanksysteme Just-In-Time (JIT) Compilation zur Generierung von Maschinencode für optimierte Anfragepläne und führen diese direkt auf der Hardware aus. JIT Compi-

lation bietet jedoch auch Vorteile für andere Datenmodelle und Verarbeitungssysteme.

Im Fachbeitrag *Efficient Compilation of Regular Path Queries* untersuchen Frank Tetzl, Wolfgang Lehner (TU Dresden) und Romans Kasperovics (SAP SE) von verschiedenen Blickwinkeln Techniken für die JIT Compilation bei deklarativer Graphverarbeitung, insbesondere bei regulären Pfadanfragen. Dabei evaluieren und vergleichen sie mehrere Ansätze zur Verarbeitung von regulären Pfadanfragen für verschiedenartige Graph-Workloads.

### 3 Community-Beiträge

Die Rubrik „Dissertationen“ enthält in diesem Heft 11 Kurzfassungen von Dissertationen aus der deutschsprachigen DBIS-Community.

Die Rubrik „Community“ berichtet unter *News* über aktuelle Informationen (u. a. über die zeitliche Verschiebung der BTW 2021 an der TU Dresden), welche die DBIS-Gemeinde betreffen.

## 4 Künftige Schwerpunktthemen

### 4.1 Digitale Lehre im Fachgebiet Datenbanksysteme

Die Lehre im Bereich Datenbanken und Informationssysteme hat ihren festen Platz in den Curricula für Informatik-Studiengänge an Universitäten und Hochschulen. Neben klassischen Inhalten wie dem relationalen Modell oder SQL finden sich in den Lehrveranstaltungen auch stetig neue Themen, u. a. NoSQL und NewSQL. Der wachsenden Bedeutung von Big Data und Data Analytics wird auch durch eigene Profilierungen und Studiengänge im Bereich Data Science Rechnung getragen.

Neben diesen inhaltlichen Veränderungen macht die Digitalisierung natürlich auch vor der Durchführung der Lehre selbst nicht halt. Neue Lehrformen wie das Flipped-Classroom-Modell oder digitale Angebote wie Massive Open Online Courses (MOOCs) setzen mit Videos und Quizzes verstärkt auf digitale Lernmaterialien. Technische Innovationen, wie z. B. die Virtualisierung mit Docker oder die Verfügbarkeit großer Datensätze, ermöglichen Lernenden Zugriff auf komplexe Lernumgebungen für praxisnahe Übungen.

Dieses Themenheft des Datenbank-Spektrums soll einen Überblick über die Entwicklungen der digitalen Lehre im Bereich Datenbanken sowohl im Universitäts- und Hochschulkontext als auch in der betrieblichen Weiterbildung geben. Zu den relevanten Themenbereichen gehören unter anderem:

- Architekturen und Werkzeuge zur Durchführung praktischer Übungen u. a. im Bereich relationaler Datenbanksysteme oder Big-Data-Systeme
- Systeme zur (semi-)automatischen Bewertung typischer Aufgabenformate im Bereich Datenbanken
- Aufbau und Erfahrungsberichte zu neuartigen Curricula oder Lehr-Lern-Szenarien (z. B. Flipped Classroom, Blended Learning)
- Evaluationen zur Wirksamkeit digitaler Lehre.

Wir erbitten Einreichungen in Deutsch oder Englisch mit einem Umfang von 8 bis 10 Seiten (zweispaltig) gemäß den Layoutvorgaben (siehe [www.springer.com/13222](http://www.springer.com/13222)).

Frist zur Einreichung: 1. Okt. 2020

Erscheinen des Themenheftes: DASP-1-2021 (März 2021)

Gast-Editoren:

Stefanie Scherzinger, Uni Passau  
[stefanie.scherzinger@uni-passau.de](mailto:stefanie.scherzinger@uni-passau.de)  
 Andreas Thor, HTWK Leipzig  
[andreas.thor@htwk-leipzig.de](mailto:andreas.thor@htwk-leipzig.de)

### 4.2 Berlin Institute for the Foundations of Learning and Data (BIFOLD)

Das Berlin Institute for the Foundations of Learning and Data (BIFOLD) ist ein von BMBF und dem Land Berlin gefördertes Kompetenzzentrum, das aus der Fusion des Berlin Big Data Center (BBDC) und dem Berliner Zentrum für Maschinelles Lernen (BZML) hervorgegangen ist. BIFOLD hat sich zum Ziel gesetzt, hochinnovative Technologien zu entwickeln, die riesige Datenmengen organisieren und mit deren Hilfe fundierte Entscheidungen getroffen werden können, um wirtschaftlichen und gesellschaftlichen Mehrwert zu schaffen. Zu diesem Zweck werden die bislang isoliert voneinander existierenden Gebiete Datenmanagement und Maschinelles Lernen verschmolzen. Die Technologien des Zentrums sollen den Stand der Technik in der Erforschung von Methoden des Datenmanagements, des maschinellen Lernens und deren Schnittstelle vorantreiben und die führende Stellung Deutschlands in Wissenschaft und Wirtschaft im Bereich der KI ausbauen. Als Technologietreiber stehen mehrere wirtschaftlich, wissenschaftlich und gesellschaftlich relevante Anwendungsbereiche im Fokus: Fernerkundung, digitalisierte Geisteswissenschaften, die Medizin sowie Informationsmarktplätze.

Aufbauend auf weltweit anerkannten Forschungsergebnissen sollen eine automatische Optimierung, Parallelisierung sowie eine skalierbare und adaptive Verarbeitung von Algorithmen in heterogenen, verteilten Umgebungen unter Einsatz moderner Rechnerarchitekturen ermöglicht werden. Daneben stehen Erklärbarkeit, verantwortungsvol-

les Datenmanagement und innovative Anwendungen der Datenanalyse im Fokus. Behandelt werden dabei Bereiche des Datenmanagements, Maschinelles Lernen, der linearen Algebra, der Statistik, der Wahrscheinlichkeitstheorie, der Computerlinguistik sowie der Signalverarbeitung. Durch Entwicklung und Bereitstellung von Open-Source-Systemen sowie von Algorithmen und Methoden zur Datenanalyse wird das Zentrum die Ausbildung, Forschung, Entwicklung, Innovation und kommerzielle Nutzung von Big Data Analytics und KI-Anwendungen in Deutschland fördern und so deutschen Firmen einen Wettbewerbsvorteil sichern.

Wir erbitten Einreichungen in Deutsch oder Englisch mit einem Umfang von 8 bis 10 Seiten (zweispaltig) gemäß den Layoutvorgaben (siehe [www.springer.com/13222](http://www.springer.com/13222)).

Frist zur Einreichung: 1. Feb. 2021

Erscheinen des Themenheftes: DASP-2-2021 (Juli 2021)

Gasteditor:

Dr. Alexander Borusan, TU Berlin

[alexander.borusan@tu-berlin.de](mailto:alexander.borusan@tu-berlin.de)

### 4.3 Data Engineering for Data Science

Data engineering is a crucial part of any data science project: Data collection and metadata management are the prerequisite of any meaningful analysis and, in practice, take up the bulk of time spent in data science projects.

This special issue of *Datenbankspektrum* is an initiative of the newly founded DBIS working group “Data Engineering for Data Science”. We intend to provide a venue for discussions, interactions and collaborations on the potential of data management research to data science projects.

We call for articles that report on novel contributions in this area, such as:

- Interplay between data engineering and data science.
- Dedicated database and dataflow architectures.
- Managing data and event streams.
- Scalable data processing in data science.
- Managing metadata in data science projects.
- Data provenance in data science projects.
- Reproducibility and replicability of data analysis.
- Knowledge discovery in data science applications.
- Data and information visualization.
- Data and information flow engineering and management.
- Privacy preserving data, information, and information systems.
- Development of dedicated benchmarks for evaluating data engineering solutions.

Expected size of the paper: 8–10 pages, double-column (cf. the author guidelines at [www.springer.com/13222](http://www.springer.com/13222)).

Contributions either in German or in English are welcome.

Deadline for submissions: June 1st, 2021

Issue delivery: DASP-3-2021

It is planned that accepted papers will also be presented at LWDA 2021.

Guest editors:

Ralf Schenkel, Universität Trier

[schenkel@uni-trier.de](mailto:schenkel@uni-trier.de)

Stefanie Scherzinger, Universität Passau

[stefanie.scherzinger@uni-passau.de](mailto:stefanie.scherzinger@uni-passau.de)

Marina Tropmann-Frick, HAW Hamburg

[marina.tropmann-frick@haw-hamburg.de](mailto:marina.tropmann-frick@haw-hamburg.de)

### 4.4 Best Workshop Papers of BTW 2021

This special issue of the “*Datenbank-Spektrum*” is dedicated to the Best Papers of the Workshops running at the BTW 2021 at the TU Dresden. The selected Workshop contributions should be extended to match the format of regular DASP papers.

Paper format: 8-10 pages, double-column

Selection of the Best Papers by the Workshop chairs and the guest editor: Oct. 1st, 2021

Deadline for submissions: Nov. 1st, 2021

Issue delivery: DASP-1-2022 (March 2022)

Guest editor:

Theo Härder, University of Kaiserslautern

[haerder@cs.uni-kl.de](mailto:haerder@cs.uni-kl.de)

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.