A Message from the SMDB 2020 Chairs

The SMDB Workshop series sponsored by the IEEE TCDE Workgroup on Self-Managing Database Systems brings together researchers and practitioners to exchange ideas related to self-managing data management systems. After the successful reincarnation of the workshop last year, we are organizing the SMDB Workshop once again given the renewed interest in self-managing databases and distributed systems (which have also been popularly called autonomous systems and self-driving systems in recent years). These systems are growing rapidly in scale and complexity, while skilled operational teams to run these systems have become rarer than ever. Furthermore, the agility with which enterprises want to move in the cloud can only be sustained via automation.

We are pleased to present a deep and diverse set of papers in this eleventh edition of the SMDB Workshop. Seven papers were submitted in total, which were subject to a rigorous review procedure; every paper had three to four independent reviews. Four papers (two regular and two short) were selected for publication in this proceedings volume. In addition, two invited regular papers were solicited from two research groups that have long experience with self-managing database systems: one from the Technical University of Munich and one from Aristotle University of Thessaloniki. Both invited papers were reviewed by the Program Committee Chairs to ensure their quality. The six papers were organized into two research sessions. We own our sincere gratitude to the members of the Technical Program Committee who contributed in assembling such a high-quality program, working within a very tight schedule.

The first paper in the first research session paper investigates the problem of rendering the Flink engine adaptive with a view to supporting a use case, in which the value-based partitioning of streaming data into keys needs to adapt to evolving data characteristics. A feedback-based architecture is proposed for getting statistical metadata from tasks downstream and passing them in upstream tasks that control the key assignment. The second paper presents a methodology of self-adapting data migration, which automatically adjusts migration strategies and its parameters accordingly with the intent of saving migration costs. The methodology considers various factors, including the query workload, the number and kind of changes in the data model caused by schema evolution, and application requirements. The third motivates the need to handle approximate constraints in self-managing databases and introduces the concept of PatchIndexes for nearly unique columns and nearly sorted columns, handling exceptions to constraints in a common data structure.

The first paper in the second research session introduces START, a self-tuning variant of the Adaptive Radix Tree (ART) that uses nodes spanning multiple key-bytes. START employs cost modeling and an optimizer for introducing these new node types and fine-tuning an existing ART, reducing its overall height, and improving performance. The second paper introduces cost-guided learning for cardinality estimation, a technique that helps learned cardinality

estimation models focus on predictions that have the largest impact on query plans, regardless of the accuracy of the cardinality estimates. The final paper proposes a Deep Reinforcement Learning (DRL) approach for index selection for replicas of a cluster database. Using DRL, an index advisor can improve its decision using the feedbacks of its decisions in a way that not only minimizes the workload processing cost but also considers the load balance among replicas.

In addition to the papers, the workshop also features two keynotes and one invited speech. The first keynote speech is titled “AIOps with the Oracle Autonomous Database" by Sandesh Rao, Vice President of the Autonomous Health and Machine Learning, Oracle Autonomous Database Group. The second keynote speech is titled “Software Hardware Co-Design for Cloud Native Database Systems” by Feifei Li, Vice President of Alibaba Group and Professor at University of Utah (on leave). The two keynote speeches are common with the HardDB/ActiveDB workshop. The invited speech is titled “AI-native Database” by Prof. Guoliang Li, Tsinghua University. We thank and appreciate deeply all invited speakers.

We would like to welcome all of you to attend SMDB 2020, which will take place online due to the directives involving COVID-19, and be part of the new resurgence we are seeing across academia, research, and industry in developing self-managing systems.

SMDB 2020 Workshop General Chair

Panos K. Chrysanthis
Director of Advance Data Management Lab
University of Pittsburgh

Meichun Hsu
Sr. Director R&D Database Server Technology
Oracle Corporation

SMDB 2020 Workshop Program Chairs

Herodotos Herodotou
Assistant Professor
Cyprus University of Technology

Yingjun Wu
Software Development Engineer
Redshift Team, Amazon Web Services