Report: 5th Int'l Workshop on Self-Managing Database Systems (SMDB 2010)

Shivnath Babu
Duke University
shivnath@cs.duke.edu

Kai-Uwe Sattler
Ilmenau University of Technology
kus@tu-ilmenau.de

1 Introduction

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 autonomic data management systems. Previous workshops of the SMDB series focused on core topics in self-managing databases like physical design tuning, problem diagnosis and recovery, and database integration and protection. In addition to core topics, the 2010 workshop aimed to broaden the interest range by covering emerging research areas like Cloud computing, multitenant databases, large-scale storage systems, and datacenter administration.

2 Workshop Overview

The 2010 workshop took place on March 1st, 2010 in Long Beach, CA, on the day before ICDE. The workshop received generous sponsorships from Oracle, Sybase iAnywhere, IBM, and Ingres.

The program committee consisted of the members of the SMDB Workgroup's executive committee and other researchers well-known in the area. We received 15 submissions out of which 6 submissions were accepted as long papers and 3 submissions as short papers. In addition to the research paper sessions, the workshop included keynote and panel sessions. The average attendance over all sessions was 35 participants, with the attendance peaking to around 45 during the panel. The full workshop program as well as presentation slides can be found on the workshop's Web site at http://db.uwaterloo.ca/tcde-smdb/smdb10.

3 Keynote

Oliver Ratzesberger, Sr. Director Architecture & Operations at eBay Inc., delivered the keynote. Oliver has more than 12 years of experience in large-scale Data Warehousing and Business Intelligence, and is now responsible for the overall eBay site capacity planning and operations cost management – one of the world's largest database infrastructure with more than 20,000 servers.

Oliver's keynote presentation, titled "Agile Enterprise Analytics," gave insights into eBay's new approach of *Analytics as a Service (AaaS)*. AaaS is implemented as virtual data marts which are logical instances created on a shared physical infrastructure. A virtual data mart can be created by a user in less than five minutes by filling a Web form. A virtual data mart may be created for prototyping, development, or testing as well as for production use such as click stream, financial, or performance analytics. The virtual data marts run on a

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

single MPP infrastructure on a single copy of the data, exploiting workload management techniques such as queuing and resource partitioning. The talk concluded with the thesis that agile analytics enables users to try out new ideas much faster and lowers the time to market. Going forward, workload management, scalability, and virtualization will be key factors determining the success of AaaS.

4 Paper Sessions

The research paper program was organized into three paper sessions: Indexing and Workload Management, On Cloud and Column Stores, and Query Optimization and Workload Management.

The papers in the first session dealt with index and workload management in database systems. Graefe and Kuno [1] proposed an approach that enables self-selecting and self-tuning indexes for point (equality) queries. Previous work on adaptive indexing had focused on range queries. The new approach uses ideas from database cracking and adaptive merging. Simulation results show that the overhead (the number of records accessed beyond the number of records used in case of full indexes) can be reduced significantly compared to a pure database cracking approach.

Schmidt and Härder [2] presented their work on index selection in native XML databases. Specific challenges in XML indexing include support for different node types, flexible path expressions, and the indexing of both structure and content. After a discussion of storage structures and different XML index types, the authors proposed an approach for estimating the costs and benefits of indexes. Finally, they presented strategies for generating and selecting index candidates. The experimental results showed the benefits of self-tuning indexing as well as the ability to adapt to workload shifts.

Powley et al. [3] dealt with the problem of workload management. They proposed an approach for workload execution control based on query throttling. Here, problematic queries are slowed down to free resources for more important parts of the workload. The authors described three approaches for throttling: a simple controller which uses a diminishing step function, a controller based on a control theoretic approach where the database is treated as a black box, and a hybrid variant that combines these two techniques. The experimental results verified that by using the controllers the important workload can meet desired performance goals continuously.

The main topic of the papers in the second session was data management in the Cloud. Ganapathi et al. [4] introduced a framework for predicting resource requirements for Cloud computing applications. The goal of this work was to predict execution times of MapReduce jobs and Hive queries. For this purpose, the authors used a machine learning technique for correlation analysis between query plan features (such as job configuration parameters like the number of maps and reduces as well as input bytes) and performance metrics (execution time, map time, reduce time, output bytes). Then, these statistical models were used to predict performance metrics for a given setup or to identify workload features which affect performance. Based on these results, the authors discussed the design of a workload generator. At the end of the workshop, this paper was selected by the audience as the most popular paper of this workshop.

Rogers et al. [5] presented a framework that optimizes operational costs of a DBMS on top of an Infrastructure as a Service (IaaS) based on a given query workload and the provider's pricing model while satisfying QoS expectations. They discussed two solutions: a white-box approach with fine-grained estimation of resource consumption and a black-box solution with coarse-grained profiling information. Both approaches were treated as constraint programming problems that were solved using a generic constraint solver. Experimental results using the TPC-H benchmark with PostgreSQL on Amazon EC2 showed the feasibility of the solutions.

Amossen [6] addressed the problem of vertical partitioning in OLTP databases. He presented a model for estimating the cost of executing a workload using a given partitioning schema. Based on the cost model, two algorithms for solving the partition design problem were presented. These algorithms preserve single-sitedness for read queries and allow column replication. The first algorithm uses a quadratic integer programming approach, while the second algorithm is based on simulated annealing. Experiments showed that the latter reduces

the computational effort by 37% for the TPC-C benchmark while returning solutions with costs close to the quadratic approach.

The short paper session contained one paper on query optimization and two papers on workload management. Dash et al. [7] addressed the problem of time-consuming calls to the query optimizer from automated physical design advisors. While the optimizer evaluates each configuration, it creates and evaluates several intermediate plans. Caching partial query plan costs can provide quick answers to subsequent calls to the optimizer. Based on this observation, the PostgreSQL optimizer was extended to support PINUM, an index selection tool for PostgreSQL. Experiments with a synthetic star-schema workload demonstrated 5 times reduction in overhead with only 3% loss in accuracy compared to direct optimizer calls.

The goal of the work presented by Holze et al. [8] was to predict shifts in workloads with a focus on periodic patterns. The proposed approach comprises three steps. The first step involves workload monitoring to collect OS and DBMS metrics as well as SQL statement characteristics. Next, a k-means-based clustering approach is used to derive statement classes. In the third step, the workload is modeled using n-gram models as an approximation of Markov models. Detecting that the current workload is similar to a workload W observed in the past allows the system to switch to a a pre-optimized configuration for W.

Finally, Abouzour et al. [9] described a multi-input single-output controller to determine the optimal multiprogramming level (MPL) in Sybase. The throughput level (number of server-side requests), the total number of outstanding requests, and the control interval are used as input. The output parameter is the MPL to use in the next control interval. The tuning task was implemented using different algorithms: hill climbing, global parabola approximation, as well as a hybrid approach that switches between the other two strategies. Experimental results showed that the hybrid approach was able to reach around 90% of the optimal (hand-tuned) value.

5 Panel

The workshop concluded with a panel on "Databases, MapReduce, and the Cloud–Oh My! What's in it for the Administrator?", with six distinguished panelists: Ashraf Aboulnaga from Univ. of Waterloo, Namit Jain from Facebook, Guy Lohman from IBM, Oliver Ratzesberger from eBay, Benjamin Reed from Yahoo!, and Jingren Zhou from Microsoft.

The computer science community has witnessed a recent debate on the performance advantages and disadvantages of database systems vs. MapReduce systems for large-scale analytics. The panelists were asked to discuss the system administration aspects of these two systems. Questions posed to the panelists included: (a) are there fundamental strengths or weaknesses that databases and MapReduce systems have regarding administration? (b) what does administering a MapReduce system involve? (c) does the emergence of Cloud computing shift the "cost of administration" equation in favor of either of these systems? (d) what lessons should MapReduce systems learn regarding administration from database systems? (e) will the roles of the data analyst, system developer, and system administrator increasingly overlap in the future?

Aboulnaga discussed how the number of software and hardware instances that have to be administered in a Cloud setting is very high, and the impact of an administrator mistake can be high as well. Administrative tasks like upgrades are hard on the Cloud. Aboulnaga also discussed how database and MapReduce systems can learn good practices from each other. For example, databases need to learn how to manage data that is not already loaded into the system, while MapReduce systems can learn how to auto-tune configuration parameters.

Jain gave an overview of Facebook's 12 PetaByte data analytics infrastructure driven by Hadoop. This warehouse runs on 9600 cores, and absorbs 10 TeraBytes (compressed) of new data per day. Data analysts and engineers use a SQL-like language to run declarative queries in this massive warehouse. Jain mentioned the challenges in this setting which include ensuring performance isolation across jobs in a shared cluster, metadata discovery, cluster and job monitoring, and efficient query plan generation.

Lohman summarized the Cloud's main value propositions from a customer's perspective which include pay-as-you-go usage and reduction of human effort. A study done by IBM found that Cloud deployment solutions

can liberate considerable amounts of IT funds for new development. However, Lohman questioned whether these benefits will continue to be realized years from now when the Cloud becomes legacy. He warned of a future where the benefits of Cloud computing disappear due to improper management of mixed workloads, schema, service-level objectives, and security.

Ratzesberger developed on his keynote theme emphasizing the need for Analytics as a Service, and the accompanying management challenges. Virtual data marts have to be provisioned automatically on demand based on user specification, while addressing workload management challenges like performance isolation in a shared physical cluster.

Reed identified three current administrative roles in Hadoop at Yahoo!: developers who write Hadoop code with certain tuning knobs, operations personnel who configure Hadoop clusters but do not configure job-level tuning knobs, and release engineering personnel who prepare Hadoop images for release. Since interactions among people in these roles is limited, Reed opined that Hadoop stands to benefit immensely from auto-tuning techniques.

Zhou observed that while database systems run on many tens of machines, MapReduce systems run on tens of thousands of machines. While database administrators focus on issues like physical design tuning and system recovery, in large-scale systems running on the Cloud, most of an administrator's time today goes into keeping the system and data alive and available for use.

6 SMDB 2011

SMDB 2011 is being organized by Vivek Narasayya and Neoklis Polyzotis, and is scheduled for April 10th, 2011 in Hannover, Germany. Visit `http://db.uwaterloo.ca/tcde-smdb/smdb11` for more information.

References

[1] Goetz Graefe, Harumi Kuno: Adaptive Indexing for Relational Keys. In *Proc. ICDE Workshops 2010*, pp. 69-74, 2010.

[2] Karsten Schmidt, Theo Härder: On the Use of Query-Driven XML Auto-Indexing. In *Proc. ICDE Workshops 2010*, pp. 81-86, 2010.

[3] Wendy Powley, Patrick Martin, Mingyi Zhang, Paul Bird, Keith McDonald: Autonomic Workload Execution Control Using Throttling. In *Proc. ICDE Workshops 2010*, pp. 75-89, 2010.

[4] Archana Ganapathi, Yanpei Chen, Armando Fox, Randy Katz, David Patterson: Statistics-Driven Workload Modeling for the Cloud. In *Proc. ICDE Workshops 2010*, pp. 87-92, 2010.

[5] Jennie Rogers, Olga Papaemmanouil, Ugur Cetintemel: A Generic Auto-Provisioning Framework for Cloud Databases. In *Proc. ICDE Workshops 2010*, pp. 63-68, 2010.

[6] Rasmus Resen Amossen: Vertical Partitioning of Relational OLTP Databases Using Integer Programming. In *Proc. ICDE Workshops 2010*, pp. 93-98, 2010.

[7] Debabrata Dash, Ioannis Alagiannis, Cristina Maier, Anastasia Ailamaki: Caching All Plans with Just One Optimizer Call. In *Proc. ICDE Workshops 2010*, pp. 105-110, 2010.

[8] Marc Holze, Ali Haschimi, Norbert Ritter: Towards Workload-Aware Self-Management: Predicting Significant Workload Shifts. In *Proc. ICDE Workshops 2010*, pp. 111-116, 2010.

[9] Mohammed Abouzour, Kenneth Salem, Peter Bumbulis: Automatic Tuning of the Multiprogramming Level in Sybase SQL Anywhere. In *Proc. ICDE Workshops 2010*, pp. 99-104, 2010.