Dana Van Aken

Carnegie Mellon University
Department of Computer Science
Gates-Hillman Center 9223
Pittsburgh, PA 15213-3891 USA

Mobile: +1 (425) 283-8384
E-mail: dvanaken@cs.cmu.edu
Web: cs.cmu.edu/~dvanaken/

EDUCATION

Carnegie Mellon University — Pittsburgh, PA
Ph.D., Computer Science — Fall 2014 – present
Advisor: Andrew Pavlo
Thesis topic: Automatic DBMS Tuning Through Machine Learning

Carnegie Mellon University — Pittsburgh, PA
M.S., Computer Science — Spring 2018

University of Washington — Seattle, WA
B.S., Computer Engineering — Spring 2014

Western Washington University — Bellingham, WA
B.A., Business Finance — Spring 2008

RESEARCH

OtterTune: Automatic DBMS Tuning Through Machine Learning — 2014 – present
We are developing the foundation and corresponding practical techniques to automate the task of optimizing DBMS configurations for a broad class of application workloads. Our approach reduces the time and resources needed to tune the DBMS for each application by leveraging knowledge gained from previous tuning sessions.

Query-based Workload Forecasting for Self-Driving DBMSs — 2017 – 2018
QB5000 is a robust workload forecasting framework for self-driving DBMSs. The key advantage of our approach is that the data used to train our models is independent of the hardware and the database design, and thus it is not necessary to rebuild the models if the DBMS's hardware or configuration settings change.

Evaluating Concurrency Control in Distributed Databases — 2015 – 2017
To better understand the effect of concurrency control in a distributed database environment, we evaluate several concurrency control methods for on-line transaction processing (OLTP) workloads in an in-memory distributed database. Our analysis exposes the severe limitations of distributed DBMSs, including their lack of scalability.

Simplifying Deployment for Mobile/Cloud Applications — 2013 – 2014
Sapphire is a distributed programming platform providing customizable and extensible deployment of mobile/cloud applications. The key concept is an architecture that supports deployment managers, which solve complex distributed systems tasks, such as code-offloading and caching. Rather than writing distributed systems code, programmers compose a custom deployment to meet their application's needs.

PUBLICATIONS

Conference Publications

Lin Ma, Dana Van Aken, Ahmed Hefny, Gustavo Mezerhane, Andrew Pavlo, and Geoffrey J. Gordon. "Query-based Workload Forecasting for Self-Driving Database Management Systems". In: *Proceedings of the 2018 ACM International Conference on Management of Data.* SIGMOD'18, 2018.

Dana Van Aken, Andrew Pavlo, Geoffrey J. Gordon, and Bohan Zhang. "Automatic Database Management System Tuning Through Large-scale Machine Learning". In: *Proceedings of the 2017 ACM International Conference on Management of Data*. SIGMOD'17, 2017, pp. 1009–1024.

Rachael Harding, Dana Van Aken, Andrew Pavlo, and Michael Stonebraker. "An Evaluation of Distributed Concurrency Control". In: *Proc. VLDB Endow.* 10.5 (Jan. 2017), pp. 553 – 564.

Andrew Pavlo, Gustavo Angulo, Joy Arulraj, Haibin Lin, Jiexi Lin, Lin Ma, Prashanth Menon, Todd Mowry, Matthew Perron, Ian Quah, Siddharth Santurkar, Anthony Tomasic, Skye Toor, Dana Van Aken, Ziqi Wang, Yingjun Wu, Ran Xian, and Tieying Zhang. "Self-Driving Database Management Systems". In: *CIDR 2017, Conference on Innovative Data Systems Research*. 2017.

Irene Zhang, Adriana Szekeres, Dana Van Aken, Isaac Ackerman, Steven D. Gribble, Arvind Krishnamurthy, and Henry M. Levy. "Customizable and Extensible Deployment for Mobile/Cloud Applications." In OSDI, vol. 14, pp. 97-112. 2014.

Demonstrations

Bohan Zhang, Dana Van Aken, Justin Wang, Tao Dai, Shuli Jiang, Jacky Lao, Siyuan Sheng, Andrew Pavlo, and Geoffrey J. Gordon. "A Demonstration of the OtterTune Database Management System Tuning Service." *Under review.*

Dana Van Aken, Djellel E. Difallah, Andrew Pavlo, Carlo Curino, and Philippe Cudre-Mauroux. "BenchPress: Dynamic Workload Control in the OLTP-Bench Testbed." In: *Proceedings of the 2015 International Conference on Management of Data*. SIGMOD'15, 2015, pp. 1069-1073.

Invited Articles

Dana Van Aken, Andrew Pavlo, Geoffrey J. Gordon. "Tuning Your DBMS Automatically with Machine Learning". *AWS Machine Learning Blog*. June 2017.

AWARDS & HONORS

AWS Cloud Credits for Research 2015 – 2017

CRA-Women Grad Cohort Workshop Travel Grant 2017

National Science Foundation Graduate Research Fellowship 2016

Facebook Fellowship, Finalist 2016

Women in Research Lean In Event Travel Grant 2016

ACM SIGMOD Student Travel Grant 2015

CRA Outstanding Undergraduate Research Award, Nomination 2013

TEACHING

Head Teaching Assistant, Advanced Database Systems Spring 2017
Carnegie Mellon University, 15-721

Head Teaching Assistant, Database Applications Fall 2015
Carnegie Mellon University, 15-415/615

Teaching Assistant, Data Structures and Algorithms Winter 2013
University of Washington, CSE 373

Math Tutor, Academic Success Center Summer 2011 – Spring 2012
Bellevue College

MENTORING

Research Mentor, Bohan Zhang (M.S.) 2016 – present

Bohan and I are extending OtterTune's machine learning algorithms to support the optimization of more complex knob configurations. Bohan was a research intern when we started working together, and has continued to work with me after joining the M.S. program at CMU.

Research Mentor, Shuli Jiang (B.S.) 2017 – present

Shuli and I are exploring how to enable OtterTune to reuse information across different versions of the same DBMS in its models.

PROFESSIONAL EXPERIENCE

Microsoft Research Redmond, WA

Research Intern, Data Management, Exploration, and Mining Group Fall 2016

Google, Inc. Seattle, WA

Software Engineering Intern, Infrastructure Team Summer 2014

SERVICE

CMU Database Group Seminar Organizer 2017 – present

CSD Open House Graduate Student Panel 2016

SCS Ph.D. Sisters Program Mentor 2015