

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, Lancaster, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Zürich, Switzerland

John C. Mitchell

Stanford University, Stanford, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Dortmund, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/7408>

Raghunath Nambiar · Meikel Poess (Eds.)

Performance Characterization and Benchmarking

Traditional to Big Data

6th TPC Technology Conference, TPCTC 2014
Hangzhou, China, September 1–5, 2014
Revised Selected Papers



Springer

Editors

Raghunath Nambiar
Cisco Systems, Inc.
San Jose, CA
USA

Meikel Poess
Oracle Corporation
Redwood Shores, CA
USA

ISSN 0302-9743

ISSN 1611-3349 (electronic)

Lecture Notes in Computer Science

ISBN 978-3-319-15349-0

ISBN 978-3-319-15350-6 (eBook)

DOI 10.1007/978-3-319-15350-6

Library of Congress Control Number: 2015931067

LNCS Sublibrary: SL2 – Programming and Software Engineering

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media
(www.springer.com)

Preface

The Transaction Processing Performance Council (TPC) is a nonprofit organization established in August 1988. Over the years, the TPC has had a significant impact on the computing industry's use of industry-standard benchmarks. Vendors use TPC benchmarks to illustrate performance competitiveness for their existing products, and to improve and monitor the performance of their products under development. Many buyers use TPC benchmark results as points of comparison when purchasing new computing systems.

The information technology landscape is evolving at a rapid pace, challenging industry experts and researchers to develop innovative techniques for evaluation, measurement, and characterization of complex systems. The TPC remains committed to developing new benchmark standards to keep pace with these rapid changes in technology. One vehicle for achieving this objective is the TPC's sponsorship of the Technology Conference Series on Performance Evaluation and Benchmarking (TPCTC) established in 2009. With this conference series, the TPC encourages researchers and industry experts to present and debate novel ideas and methodologies in performance evaluation, measurement, and characterization.

The first TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2009) was held in conjunction with the 35th International Conference on Very Large Data Bases (VLDB 2009) in Lyon, France from August 24 to 28, 2009.

The second TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2010) was held in conjunction with the 36th International Conference on Very Large Data Bases (VLDB 2010) in Singapore from September 13 to 17, 2010.

The third TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2011) was held in conjunction with the 37th International Conference on Very Large Data Bases (VLDB 2011) in Seattle, Washington from August 29 to September 3, 2011.

The fourth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2011) was held in conjunction with the 38th International Conference on Very Large Data Bases (VLDB 2012) in Istanbul from August 27 to 31, 2012.

The fifth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2012) was held in conjunction with the 39th International Conference on Very Large Data Bases (VLDB 2012) in Riva del Garda, Trento, Italy from August 26 to 30, 2013.

This book contains the proceedings of the sixth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2014), held in conjunction with the 40th International Conference on Very Large Data Bases (VLDB 2014) in Hangzhou, China from September 1 to 5, 2014, including 12 selected peer-reviewed papers.

The hard work and close cooperation of a number of people have contributed to the success of this conference. We would like to thank the members of TPC and the organizers of VLDB 2014 for their sponsorship; the members of the Program Committee and Publicity Committee for their support; and the authors and the participants who are the primary reason for the success of this conference.

September 2014

Raghunath Nambiar
Meikel Poess

TPCTC 2014 Organization

General Chairs

Raghunath Nambiar
Meikel Poess

Cisco, USA
Oracle, USA

Program Committee

Chaitanya Baru
Daniel Bowers
Michael Brey
Wen-Chen Hu
Alain Crolotte
Akon Dey
Masaru Kitsuregawa
Harumi Kuno
Dhabaleswar Panda
Tilman Rabl
Reza Taheri
Xiaohua Tony Hu
Marco Vieira
Jianfeng Zhan

SDSC, USA
Gartner, USA
Oracle, USA
University of North Dakota, USA
Teradata, USA
University of Sydney, Australia
University of Tokyo, Japan
HP Labs, USA
The Ohio State University, USA
University of Toronto, Canada
VMware, USA
Drexel University, USA
University of Coimbra, Portugal
Chinese Academy of Sciences, China

Publicity Committee

Andrew Bond
Forrest Carman
Andreas Hotea
Michael Majdalany
Andrew Masland
Raghunath Nambiar
Meikel Poess
Da-Qi Ren
Reza Taheri

Red Hat, USA
Owen Media, USA
Hotea Solutions, USA
L&M Management Group, USA
NEC, USA
Cisco, USA
Oracle, USA
Huawei, China
VMware, USA

TPC 2014 Organization

Full Members

Actian
Cisco
Cloudera
Dell
Fujitsu
HP
Hitachi
Huawei
IBM
Intel
Inspur
MapR
Microsoft
NEC
Oracle
Pivotal
Red Hat
SAP
Teradata
Unisys
VMware

Associate Members

IDEAS International
ITOM International Co
San Diego Super Computing Center
Telecommunications Technology Association
University of Coimbra, Portugal

Steering Committee

Michael Brey	Oracle, USA
Paul Cao	HP, USA
Raghunath Nambiar	Cisco, USA
Jamie Reding	Microsoft, USA
Wayne Smith (Chair)	Intel, USA

Public Relations Committee

Andrew Bond	Red Hat, USA
Raghunath Nambiar (Chair)	Cisco, USA
Andrew Masland	NEC, USA
Meikel Poess	Oracle, USA
Reza Taheri	VMware, USA

Technical Advisory Board

Andrew Bond	Red Hat, USA
Paul Cao	HP, USA
Matthew Emmerton	IBM, USA
John Fowler	Oracle, USA
Andrew Masland	NEC, USA
Jamie Reding (Chair)	Microsoft, USA
Wayne Smith	Intel, USA

About the TPC

Introduction to the TPC

The Transaction Processing Performance Council (TPC) is a nonprofit organization that defines transaction processing and database benchmarks and distributes vendor-neutral performance data to the industry. Additional information is available at <http://www.tpc.org/>.

TPC Memberships

Full Members

Full Members of the TPC participate in all aspects of the TPC's work, including development of benchmark standards and setting strategic direction. The Full Member application can be found at <http://www.tpc.org/information/about/app-member.asp>.

Associate Members

Certain organizations may join the TPC as Associate Members. Associate Members may attend TPC meetings, but are not eligible to vote or hold office. Associate membership is available to nonprofit organizations, educational institutions, market researchers, publishers, consultants, governments, and businesses that do not create, market, or sell computer products or services. The Associate Member application can be found at <http://www.tpc.org/information/about/app-assoc.asp>.

Academic and Government Institutions

Academic and government institutions are invited to join the TPC and a special invitation can be found at <http://www.tpc.org/information/specialinvitation.asp>.

Contact the TPC

TPC

Presidio of San Francisco
Building 572B (surface)
P.O. Box 29920 (mail)
San Francisco, CA 94129-0920, USA
Voice: 415-561-6272
Fax: 415-561-6120
Email: info@tpc.org

How to Order TPC Materials

All of our materials are now posted free of charge on our web site. If you have any questions, please feel free to contact our office directly or by email at info@tpc.org.

Benchmark Status Report

The TPC Benchmark Status Report is a digest of the activities of the TPC and its technical subcommittees. Sign-up information can be found at the following URL: <http://www.tpc.org/information/about/email.asp>.

Contents

Introducing TPCx-HS: The First Industry Standard for Benchmarking Big Data Systems	1
<i>Raghunath Nambiar, Meikel Poess, Akon Dey, Paul Cao, Tariq Magdon-Ismael, Da Qi Ren, and Andrew Bond</i>	
An Evaluation of Alternative Physical Graph Data Designs for Processing Interactive Social Networking Actions	13
<i>Shahram Ghandeharizadeh, Reihane Boghrati, and Sumita Barahmand</i>	
On Characterizing the Performance of Distributed Graph Computation Platforms	29
<i>Ahmed Barnawi, Omar Batarfi, Seyed-Mehdi-Reza Behteshi, Radwa Elshawi, Ayman Fayoumi, Reza Nouri, and Sherif Sakr</i>	
Discussion of BigBench: A Proposed Industry Standard Performance Benchmark for Big Data	44
<i>Chaitanya Baru, Milind Bhandarkar, Carlo Curino, Manuel Danisch, Michael Frank, Bhaskar Gowda, Hans-Arno Jacobsen, Huang Jie, Dileep Kumar, Raghunath Nambiar, Meikel Poess, Francois Raab, Tilman Rabl, Nishkam Ravi, Kai Sachs, Saptak Sen, Lan Yi, and Choonhan Youn</i>	
A Scalable Framework for Universal Data Generation in Parallel	64
<i>Ling Gu, Minqi Zhou, Qiangqiang Kang, and Aoying Zhou</i>	
Towards an Extensible Middleware for Database Benchmarking	82
<i>David Bermbach, Jörn Kuhlenskamp, Akon Dey, Sherif Sakr, and Raghunath Nambiar</i>	
Scaling up Mixed Workloads: A Battle of Data Freshness, Flexibility, and Scheduling	97
<i>Iraklis Psaroudakis, Florian Wolf, Norman May, Thomas Neumann, Alexander Böhm, Anastasia Ailamaki, and Kai-Uwe Sattler</i>	
Parameter Curation for Benchmark Queries	113
<i>Andrey Gubichev and Peter Boncz</i>	
Downtime-Free Live Migration in a Multitenant Database	130
<i>Nicolas Michael and Yixiao Shen</i>	
Performance Analysis of Database Virtualization with the TPC-VMS Benchmark.	156
<i>Eric Deehr, Wen-Qi Fang, H. Reza Taheri, and Hai-Fang Yun</i>	

A Query, a Minute: Evaluating Performance Isolation in Cloud Databases . . . 173
Tim Kiefer, Hendrik Schön, Dirk Habich, and Wolfgang Lehner

Composite Key Generation on a Shared-Nothing Architecture. 188
*Marie Hoffmann, Alexander Alexandrov, Periklis Andritsos, Juan Soto,
and Volker Markl*

Author Index 205