

# CALL FOR PAPERS

Fifth Workshop on Big Data Benchmarking (WBDB 2014)  
August 5-6, 2014 - HPI, Potsdam, Germany

## Submission Deadline - May 30, 2014

The objective of the WBDB workshops is to make progress towards development of industry standard application-level benchmarks for evaluating hardware and software systems for big data applications.

Discussions from previous WBDB workshops have resulted in the formation of the TPC-BigData subcommittee, which is currently working on the TPCx-HS benchmark, for Hadoop Systems. There have also been presentations on a variety of proposed benchmarks including, BigBench, HiBench, BigDataBenchmark, BigDecision, and the Deep Analytics Pipeline.

### Meeting Topics:

To continue making progress towards big data benchmarking standards, the workshop will explore a range of issues including:

- Data features: New feature sets of data including, high-dimensional data, sparse data, event-based data, and enormous data sizes.
- System characteristics: System-level issues including, large-scale and evolving system configurations, shifting loads, and heterogeneous technologies for big data and cloud platforms.
- Implementation options: Different implementation options such as SQL, NoSQL, Hadoop software ecosystem, and different implementations of HDFS.
- Workloads: Representative big data business problems and corresponding benchmark implementations.
- Specification of benchmark applications that represent the different modalities of big data, including graphs, streams, scientific data, and document collections.
- Hardware options: Evaluation of new options in hardware including different types of HDD, SSD, and main memory, and large-memory systems, and new platform options that include dedicated commodity clusters and cloud platforms.
- Synthetic data generation: Models and procedures for generating large-scale synthetic data with requisite properties.
- Benchmark execution rules: E.g. data scale factors, benchmark versioning to account for rapidly evolving workloads and system configurations, benchmark metrics.
- Metrics for efficiency: Measuring the efficiency of the solution, e.g. based on costs of acquisition, ownership, energy and/or other factors, while encouraging innovation and avoiding benchmark escalations that favor large inefficient configuration over small efficient configurations.
- Evaluation frameworks: Tool chains, suites and frameworks for evaluating big data systems.
- Early implementations: E.g. of, say, BigBench, HiBench, BigDataBenchmark, BigDecision, and the Deep Analytics Pipeline, and lessons learned in benchmarking big data applications.

<http://clds.ucsd.edu/wbdb2014.de>

Premium Sponsor



# Fifth Workshop on Big Data Benchmarking (WBDB 2014) August 5-6, 2014 - HPI, Potsdam, Germany

## Important Dates:

- Submission Date: May 30
- Notification Date: June 20
- Camera-ready: August 30

## General Chairs:

Chaitan Baru, San Diego Supercomputer Center (SDSC) UC San Diego  
Tilmann Rabl, Middleware Services Research Group (MSRG), University of Toronto  
Kai Sachs, SAP AG

## Local Arrangements:

Matthias Uflacker, Hasso-Plattner-Institut

## Publicity Chair:

Henning Schmitz, SAP Innovation Center

## Publication Chair

Meikel Poess, Oracle

## Program Committee:

Milind Bhandarkar, Pivotal  
Anja Bog, SAP Labs  
Dhruva Borthakur, Facebook  
Joos-Hendrik Böse, Amazon  
Tobias Bürger, Payback  
Tyson Condie, UCLA  
Kshitij Doshi, Intel  
Pedro Furtado, University of Coimbra  
Bhaskar Gowda, Intel  
Goetz Graefe, HP  
Martin Grund, eXascale Infolab  
Alfons Kemper, TU München  
Donald Kossmann, ETH Zürich  
Tim Kraska, Brown University

Wolfgang Lehner, TU Dresden  
Christof Leng, UC Berkeley  
Stefan Manegold, CWI  
Raghu Nambiar, Cisco  
Manoj Karunakaran Nambiar, Tata Consulting Services  
Glenn Paulley, Conestoga College  
Scott Pearson, CLDS Industry Fellow  
Andreas Polze, Hasso-Plattner-Institut  
Alexander Reinefeld, Zuse Institute Berlin / HU Berlin  
Berni Schiefer, IBM Labs Toronto  
Saptak Sen, Hortonworks  
Florian Stegmaier, University of Passau  
Till Westmann, Oracle Labs  
Jianfeng Zhan, Chinese Academy of Science

## WBDB'14 - Venue Information

The Fifth Workshop on Big Data Benchmarking (WBDB 2014) will take place at the Hasso Plattner Institute (HPI) in Potsdam. Situated directly next to the Lake Griebnitzsee metropolitan train (S-Bahn) station, the Hasso Plattner Institute in Potsdam-Griebnitzsee is also situated in close proximity to Berlin and Potsdam's city center.

Potsdam's palaces and historical parks have been a part of the UNESCO World Heritage since 1990. The most well-known sights in the city are probably the Sanssouci Palace and the Cecilienhof Palace. The historical Russian colony, Alexandrovka, the Dutch Quarter, with its red brick houses, and the former quarter of Bohemian weavers bestow a flair to the city formed by the traditions of Europe. Also widely known are the film studios in Babelsberg and the many scientific institutions in Potsdam.



© Kay Herschelmann

## Contact Info:

For questions please contact Chaitan Baru, [baru@sdsc.edu](mailto:baru@sdsc.edu)

<http://clds.ucsd.edu/wbdb2014.de>

Sponsored by:



Pivotal



IT Systems Engineering | Universität Potsdam

