

SECOND WORKSHOP ON BIG DATA BENCHMARKING

clds.sdsc.edu/wbdb2012.in

Center for Large-Scale Data Systems [CLDS]
San Diego Supercomputer Center, UC San Diego

Following the successful *First Workshop on Big Data Benchmarking* held in May 2012 in San Jose, CA, the *Second Workshop on Big Data Benchmarking* will be held December 17-18 in Pune, India, at the Hinjewadi Campus of Persistent Systems Ltd. The meeting will occur immediately after COMAD 2012, at the same venue, Dec 14-16.

Objectives

The workshop series on big data benchmarking fosters development of industry-standard benchmarks for evaluating hardware and software systems designed for big data applications. These benchmarks will facilitate the evaluation of alternative solutions to big data problems and enable comparisons of different big data solutions. For this they need to characterize new feature sets, enormous data sizes, large-scale and evolving system configurations, shifting loads, and heterogeneous technologies. There are new challenges and options in software for big data such as SQL, NoSQL, and the Hadoop software ecosystem; different modalities of big data, including graphs, streams, scientific data, and document collections, etc; and new server, storage and network options.

Meeting Registration / Attendance

There is no registration fee to attend the meeting. However, to ensure effective discussions and open and active participation by all attendees, the overall attendance will be capped at not more than about 60 attendees. Preference will be given to diverse representation across a wide variety of organizations and institutions – from hardware and software vendors to those representing end users. The workshop is limited to one attendee per company, except for sponsors who will be provided multiple workshop slots.

Participation

Workshop attendees are required to submit a short paper along with a brief biography (one paragraph), and also give a 5-minute lightning talk. The paper should be formatted according to the instructions described in the section on *Formatting Instructions*.

Work-in-Progress Papers

Authors are encouraged to submit original submissions of work in progress.

Use Case Descriptions

To facilitate development of a standardized big data benchmarking suite, authors working with big data systems and applications are encouraged to describe their use cases and system setups as well as their techniques for testing and monitoring.

Demos

Demos related to big data benchmarking are encouraged. Demo descriptions should be formatted according to the instructions described in the section on *Formatting Instructions*. A separate time slot will be provided for running the demos following the corresponding lightning talk.

Travel support

Limited travel support is available for academic participants. Please contact baru@sdsc.edu if you are interested in receiving some support.

This meeting is supported by a grant from the US National Science Foundation (NSF). Sponsorship for this meeting is also being provided by Persistent Systems Ltd., Seagate, Greenplum, Brocade, and Mellanox.

Themes for WBDB2012.in

Building on the discussions of the first meeting in May 2012, the second workshop will emphasize the following topics:

- Data generation, for data genres
- Benchmark metrics
- Benchmarking process
- Applications scenarios and use cases
- Benchmarking challenges

Except for invited talks, all presentations at the workshop will be in the form of 5-minute "lightning talks".

Formatting Instructions

Papers should be **between 3 to 6 pages**, covering one or more of the topics listed above. If the paper covers a topic that is not listed above, the authors must provide a detailed explanation how their topic is related to big data benchmarking and why it should be included. Papers should be formatted using the **Springer LNCS Proceedings** format.

Submission Instructions

Submission should be done via the CMT conference management service: cmt.research.microsoft.com/WBDB2_2012/

Important Dates:

December 1st, 2012:

All papers must be submitted using the CMT system.

December 7th, 2012:

Notification to authors about paper acceptance.

December 17-18, 2012:

WBDB2012.in meeting, Pune, India.

Organizing Committee:

Chaitan Baru – CLDS, UCSD
Milind Bhandarkar – Greenplum
Raghunath Nambiar – Cisco
Meikel Poess – Oracle
Tilman Rabl – MSR, U Toronto
S Sudarshan – IIT Bombay