

Tutorial: An In-Depth Look at Modern Database Systems

C. Mohan

IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120, USA
+1 408 927 1733
cmohan@us.ibm.com

ABSTRACT

This tutorial on Modern Database Systems (MDS) is targeted at a broad set of database systems and applications people. It is intended to let the attendees better appreciate what is really behind the covers of many of the modern database systems (e.g., NoSQL and NewSQL systems), going beyond the hype associated with these open source, commercial and/or research systems. The capabilities and limitations of such systems will be addressed. Modern extensions to decades old relational DBMSs like DB2 and SQL Server will also be described. Some application case studies will also be presented.

Categories and Subject Descriptors

H.2.4 [Database Management]: Systems – *distributed databases, query processing, relational databases, transaction processing*

General Terms

Algorithms, Design, Documentation, Languages, Management, Measurement, Performance, Reliability, Standardization

Keywords

APIs, Data Models, DBMS, HBase, Hype, In-memory, JSON, DB2, Informix, Oracle, NoSQL, Optimization, RDBMS, NewSQL

1. OUTLINE

The following is the outline of this tutorial on Modern Database Systems (MDS):

- Goal: Broad Survey of Modern Database Systems (MDS)
- Drivers Behind Emergence of MDS, Benchmarks and Performance Studies
- Classes of MDS
 - Evolution of Classical DBMSs (DB2, SQL Server, Oracle, Informix, PostgreSQL)
 - Brand New Systems (NewSQL, NoSQL)

- Hybrid Systems (Hadoop + SQL)
- Overviews of Specific Systems
 - DB2 BLU, IDAA, MS Hekaton, Oracle Exadata, Informix Warehouse Accelerator
 - SAP Hana, Google F1, NuoDB, VoltDB
 - MongoDB, DB2/Informix NoSQL, Oracle NoSQL, Aerospike, Neo4j
- Deployments or Use Case Scenarios: Facebook

This is a revised version of a tutorial presented first at the 39th International Conference on Very Large Databases in Riva del Garda, Italy in August 2013. This is also a follow up to my EDBT2013 keynote talk “History Repeats Itself: Sensible and Nonsensical Aspects of the NoSQL Hoopla”. More information on this tutorial and the extended versions of it can be found at <http://bit.ly/CMnMDS>

2. SPEAKER

Dr. C. Mohan has been an IBM researcher for 32 years in the information management area, impacting numerous IBM and non-IBM products, the research and academic communities, and standards, especially with his invention of the ARIES family of locking and recovery algorithms, and the Presumed Abort commit protocol. This IBM, ACM and IEEE Fellow has also served as the IBM India Chief Scientist. In addition to receiving the ACM SIGMOD Innovation Award, the VLDB 10 Year Best Paper Award and numerous IBM awards, he has been elected to the US and Indian National Academies of Engineering, and has been named an IBM Master Inventor. This distinguished alumnus of IIT Madras received his PhD at the University of Texas at Austin. He is an inventor of 40 patents. He has served on the advisory board of IEEE Spectrum and on the IBM Software Group Architecture Board’s Council. Mohan is a frequent speaker in North America, Western Europe and India, and has given talks in 40 countries. More information can be found in his home page at <http://bit.ly/CMohan>

(c) 2014, Copyright is with the authors. Published in Proc. 17th International Conference on Extending Database Technology (EDBT), March 24-28, 2014, Athens, Greece: ISBN 978-3-89318065-3, on OpenProceedings.org. Distribution of this paper is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0