8th TPC Technology Conference on Performance Evaluation & Benchmarking
TPCTC 2016

Industry Standards for the Analytics Era

Raghunath Nambiar
General Chair

@raghu_nambiar

42nd International Conference on
VERY LARGE DATA BASES

New Delhi, India • September 5 – 9, 2016

Very Large Data Bases
INDIA 2016

Announcement from VLDB

7:00 to 9:00 PM Today

Mr. Ravi Shankar Prasad, Government of India's Union Minister of Information Technology, How data is instrumental for making India Digital.

Prof. Deepak Phatak, IIT Bombay, The journey from VLDB 1996 till VLDB 2016 and vision beyond

TPCTC 2016

TPCTC Mission

- Established in 2009

- Keep pace with technological changes

- Explore new areas, ideas and methodologies in performance evaluating and benchmarking

- Enhancements to existing TPC benchmarks inline with technology and industry trends

- Foster collaboration between industry experts and research community

TPCTC Conferences

Co-located with VLDB since 2009

Year	Location	Keynote/Invited Talk
2009	Lyon	Michael Stonebraker
2010	Singapore	C Mohan
2011	Seattle	Umeshwar Dayal
2012	Istanbul	Michael J. Carey
2013	Riva del Grada	Raghu Ramakrishnan
2014	Hangzhou	Umeshwar Dayal (Invited)
2015	Kohala Coast	Jignesh Patel
2016	New Delhi	Anastasia Ailamaki

- Proceedings published by Springer

- Listed under TPC Technical articles portal, DBLP

Impact of TPCTC

- New members joining the TPC

- Formation of WBDB initiative (outside the TPC), merged with SPEC Research in '14

- New collaboration with universities

- Standards for virtualized database workloads

- Standard for data integration

- Standards for Big Data and Analytics

- Lessons learned in practice which has broad industry applicability

TPCTC 2016: Statistics

- Submissions 20 + 2

- Accepted 9

- 52 Authors from 11 countries

- 3-4 reviews per paper

Programme

08:30 - 09:00	**Opening Remarks, Welcome. Industry Standards for the Analytics Era** Raghunath Nambiar, Meikel Poess, Reza Taheri, Michael Majdalany
09:00 - 10:00	**Keynote** **Performance and Energy Analysis Using Transactional Workloads** Anastasia Ailamaki
10:00 - 10:30	**TPCx-HS on the Cloud** Nicholas Wakou, Michael Woodside, Arkady Kanevsky, Fazal E Rehman Khan and Mofassir Ul Islam Arif
10:30 - 11:00	**Tea Break**
11:00 - 11:30	**From BigBench to TPCx-BB: Standardization of a Big Data Benchmark** Bhaskar Gowda, Chinmayi Narasimhadevara, John Poelman, Meikel Poess, Patrick Nguyen, Paul Cao, Seetha Lakshmi and Tilmann Rabl
11:30 - 12:00	**Benchmarking Machine Learning using BigBench** Sweta Singh
12:00 - 12:30	**Benchmarking Exploratory OLAP** Mahfoud Djedaini, Pedro Furtado, Nicolas Labroche, Patrick Marcel and Verónika Peralta
12:30 - 14:00	**Lunch Break**

25 minutes for presentations
5 minutes for Q&A

Programme

14:00 - 14:30	**Lessons From OLTP Workload on Multi-Socket HPE Integrity Superdome X system** Rajesh Tadakamadla, Srinivasan Varadarajan Sahasranamam and Paul Cao
14:30 - 15:00	**Benchmarking Distributed Stream Processing Platforms for IoT Applications** Anshu Shukla and Yogesh Simmhan
15:00 - 15:30	**AdBench: A Complete Data Pipeline Benchmark for Modern Data Pipelines** Milind Bhandarkar
15:30 - 16:00	**Tea Break**
16:00 - 16:30	**Lessons Learned: Performance Tuning for Hadoop Systems** Manan Trivedi and Raghunath Nambiar
16:30 - 17:00	**Work-Energy Profiles: General Approach and In-Memory Database Application** Annett Ungethüm, Thomas Kissinger, Dirk Habich and Wolfgang Lehner
17:00 - 17:30	**Closing Remarks** Meikel Poess

Acknowledgements

TPC Transaction Processing
Performance Council

VLDB 2016
Staffs and
Volunteers

Springer

52 authors from
11 countries

TPC – State of the Union

About the TPC

- The TPC is a non-profit, vendor-neutral organization, established in August 1988

- Reputation of providing the most credible performance results to the industry.

- Solid foundation for complete system-level performance

- Methodology for calculating total-system-price and price-performance

- Methodology for measuring energy efficiency

Mission

- Founded to define transaction processing benchmarks (1998)

- Define transaction processing benchmarks and database benchmarks (1999)

- Define data centric benchmarks (2015)

Major Activities of the TPC

- Development of new benchmarks relevant to the industry

- Maintenance and refinement of existing benchmarks

- Publication and promotion of TPC model and results

- Resolution of disputes and challenges

- Sponsorship of TPCTC to encourages researchers and industry experts
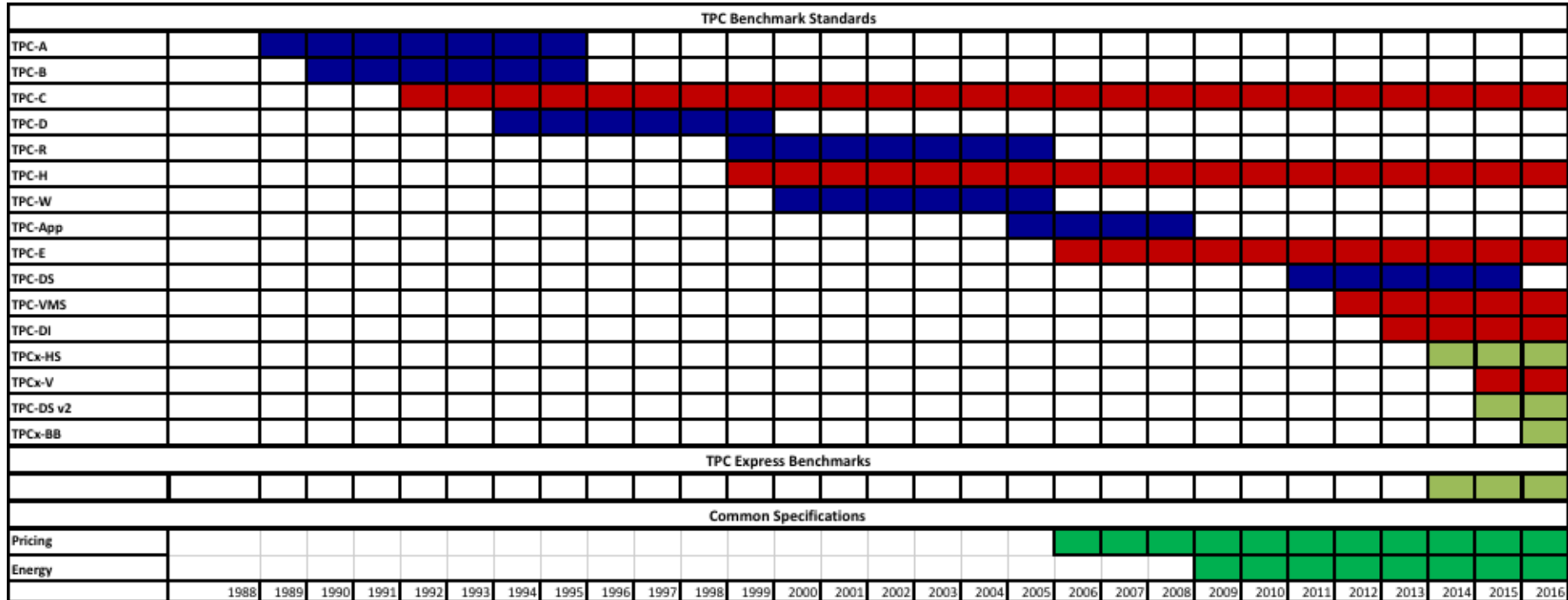
TPC Membership

20 full members and 5 associate members. 6 new members since 2013
http://www.tpc.org/information/about/join.asp

TPC Timeline

	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
TPC Benchmark Standards																													
TPC-A		■	■	■	■	■	■	■																					
TPC-B			■	■	■	■	■	■																					
TPC-C					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TPC-D						■	■	■	■	■	■																		
TPC-R									■	■	■	■	■	■	■														
TPC-H									■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TPC-W									■	■	■	■	■																
TPC-App																	■	■	■	■									
TPC-E														■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
TPC-DS																					■	■	■	■	■	■			
TPC-VMS																						■	■	■	■	■	■		■
TPC-DI																							■	■	■	■	■		■
TPCx-HS																											■	■	■
TPCx-V																												■	■
TPC-DS v2																											■		■
TPCx-BB																													■
TPC Express Benchmarks																													
																											■	■	■
Common Specifications																													
Pricing																			■	■	■	■	■	■	■	■	■	■	■
Energy																					■	■	■	■	■	■	■	■	■

Staying Relevant ! (and Ahead)
Reinventing the TPC

- TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC)

- TPC-Express Class Benchmark

- New Mission – focused on developing data centric benchmarks standards

- New Areas: Virtualization, Big Data, Internet of Things

TPC "Express" Benchmark Standards

- To keep pace with rapidly changing industry demands

- Easy to implement, run and publish, and less expensive

- Test sponsor is required to use the TPC provided kit

- The vendor may choose an independent audit or peer audit

- 60 day review/challenge window apply (as per TPC policy)

- Approved by super majority of the TPC General Council, No Mail Ballot

- All publications are required to follow the TPC Fair Use Policy
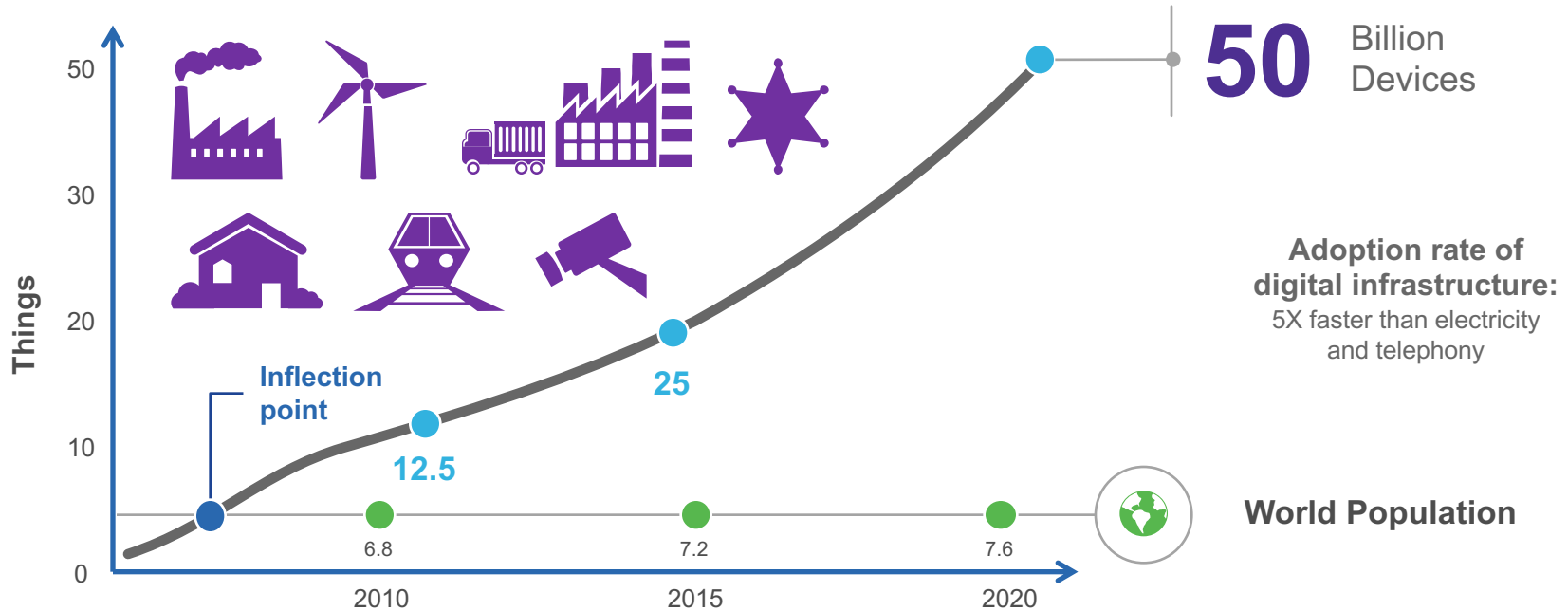
Virtualization

- Server virtualization has been a game-changing technology

- More virtual machines sold than physical machines

- TPC-VMS: The TPC Virtual Measurement Single System Specification leverages the TPC-C, TPC-E, TPC-H and TPC-DS Benchmarks by adding the methodology and requirements for running and reporting performance metrics for virtualized databases

- TPCx-V: The TPC Express Benchmark V measures the performance of a server running virtualized databases, models many properties of cloud servers, such as multiple VMs running at different load demand levels, and large fluctuations in the load level of each VM

Big Data

- Big data has become an integral part of enterprise IT ecosystem across major verticals. Big Data technology and services market is expected to grow at a 27% compound annual growth rate (CAGR) to $34 billion through 2017 - or about six times the growth rate of the overall Information and Communication Technology (ICT) market (Source: IDC)

- TPCx-HS: Developed to provide an objective measure of hardware, operating system and commercial Apache Hadoop File System API compatible software distributions

- TPC-DSv2: TPC-DS Version 2 enables emerging technologies, such as Big Data systems, to execute the benchmark. The major changes in Version 2 are in the area of ACID (Atomicity, Consistency, Isolation and Durability), data maintenance, metric calculation and execution rules

- TPCx-BB: Based on 30 frequently performed analytical queries in the context of retailers with physical and online store presence. The SQL queries can use Hive or Spark, while the machine learning algorithms use machine

IoT is Here Now

Things

- 50
- 30
- 20
- 10
- 0

Inflection point

12.5

25

50 Billion Devices

Adoption rate of digital infrastructure:
5X faster than electricity and telephony

World Population

6.8 7.2 7.6

2010 2015 2020

Internet of Things

- The Internet of Things (IoT) refers to the ever-growing network of physical objects that feature an IP address for internet connectivity, and the communication that occurs between these objects and other Internet-enabled devices and systems.

- IoT creates $14.4 Trillion of Value at Stake for Companies and Industries through 2020

- IoT = analytics. One of TPC's core area. But analytics can be relational or non-relational

- TPC IoT Working group was formed in 2015. TPC is exploring ideas and proposals

TPCTC 2018

9th TPC Technology Conference on
Performance Evaluation and Benchmarking
(TPCTC 2019)

VLDB
2017
Munich · Germany

43rd International Conference on
Very Large Data Bases

VLDB Approval is pending

TPC Benchmarks on Public Clouds

- Public cloud is becoming part of enterprise IT

- Customers are looking for performance and price-performance for instances on public clouds

- The TPC Pricing Specification 2.0 adds wording to support TPC benchmarking in cloud environments

Questions ?