Reinventing the TPC: From Traditional to Big Data to Internet of Things

Raghunath Nambiar[1(✉)] and Meikel Poess[2]

[1] Cisco Systems, Inc., 275 East Tasman Drive, San Jose, CA 95134, USA
rnambiar@cisco.com
[2] Oracle Corporation, 500 Oracle Parkway, Redwood Shores, CA 94065, USA
meikel.poess@oracle.com

Abstract. The Transaction Processing Performance Council (TPC) has made significant contributions to the industry and research with standards that encourage fair competition to accelerate product development and enhancements. Technology disruptions are changing the industry landscape faster than ever. This paper provides a high level summary of the history of the TPC and recent initiatives to make sure that it is a relevant organization in the age of digital transformation fueled by Big Data and the Internet of Things.

Keywords: Database benchmarks · Big data · Internet of things

1 Reinventing the TPC

The Transaction Processing Performance Council (TPC) was formed in 1988, as a non-profit corporation focused on defining database processing benchmarks and disseminating objective, verifiable performance data to the IT industry. Over the years the TPC has gained the reputation of providing the most credible performance results to the industry often referenced in a role of consumer reports for the computing industry. The most critical contribution of the TPC has been providing the industry with a solid foundation for complete, system-level performance and methodologies for calculating the total price and price for performance.

Over the years the TPC has changed its mission – to define transaction-processing benchmarks (when founded in 1988), to defining transaction processing benchmarks and database benchmarks (1999) and to defining data-centric benchmarks (2015) inline with industry trends.

The first set of benchmark standards were focused on transaction processing. Later standards were developed for decision support systems addressing industry demands until the late 1990s, but the TPC was unable to address the rapidly changing industry landscape in the early part of the last decade with the emergence of virtualization, cloud, energy efficiency, etc. The main challenges were very long benchmark development cycles, unable to reach consensus on standards, failed benchmark standards due to lack of traction from the industry and industry consolidation of member companies [1, 2].

R. Nambiar and M. Poess (Eds.): TPCTC 2015, LNCS 9508, pp. 1–7, 2016.
DOI: 10.1007/978-3-319-31409-9_1

In 2009, the TPC initiated a set of initiatives to reinvent itself. The first major initiative was the International Technology Conference Series on Performance Evaluation and Benchmarking to bring industry experts and researcher to accelerate benchmark developments [3–8]. The second major initiative was the TPC Express benchmark initiative where benchmarks based on predefined, executable kits can be rapidly deployed and measured [8, 9]. The third major initiative is developing a set of benchmarks for emerging areas such as big data, analytics and Internet of Things [8, 22].

2 TPC Benchmark Timelines

To date, the TPC has approved a total of thirteen independent benchmarks. Of these benchmarks, TPC-C, TPC-H, TPC-E, TPC-DS, TPC-VMS, TPC-DI and TPCx-HS are currently active. TPC-VMC, TPCx-V, TPC-DS 2.0. TPCx-BB and TPC-IoTare under development. The timelines are shown in Fig. 1.

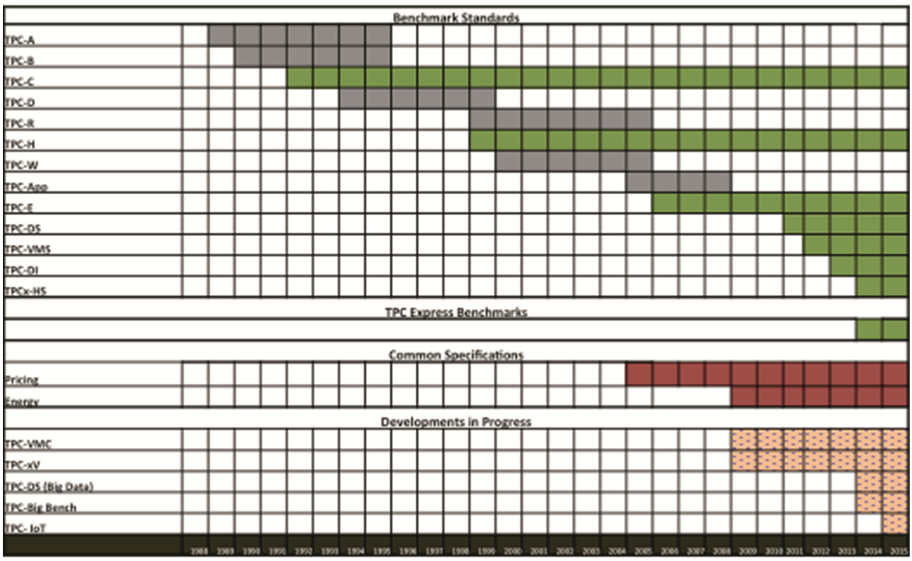

Fig. 1. TPC Benchmark Timelines

A high level summary of current active standards are listed below:

- TPC-C: An On-Line Transaction Processing (OLTP). It has been TPC's foundational and flagship benchmark with several hundreds of result publications across a variety of hardware and software systems. Historical analysis shows that the performance and price performance trend of TPC-C results has followed Moore's Law [10, 11].
- TPC-H: An ad-hoc, decision support benchmark widely popular in the industry and academia. Vendors continue to publish results on single node configurations as well as large scale-out configurations [12].

- TPC-E: An On-line Transaction Processing (OLTP) workload simulates the workload of a brokerage firm. It has been a popular benchmark but odate all publications have been on one database platform and on single node configurations [13].
- TPC-DS: A complex decision support benchmark representative of modern decision support systems. TPC took several years to develop this benchmark and reach consensus approving it as a standard. No official publications as of date. TPC-DS 2.0 is under development, one major change is loosening the relational database properties to support emerging platforms like Hadoop [14–16].
- TPC-VMS: A Single system virtualization benchmark leveraging 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 [17].
- TPC-DI: Data Integration (also known as ETL) benchmark combines and transforms data extracted from a brokerage firm's OLTP system along with other sources of data, and loads it into a data warehouse. No official publications as of date [18].
- TPCx-HS: Industry's first big data benchmark standard, also TPC's first benchmark in the TPC Express benchmark category. Modeled on a simple application, the standard is highly relevant to hardware and software dealing with Big Data systems in general. There have been over a dozen results publications till date [19, 20].

A high level summary of standards under development are listed below:

- TPC-VMC: A complex virtualization benchmarkfor database workloads [21].
- TPC-DS 2.0: A complex and comprehensive benchmark built on TPCDS 1.0 supporting emerging platforms like Hadoop. Industry's first standard for SQL based Big Data systems [14–16].
- TPCx-BB: A comprehensive Big Data benchmark in the TPC Express benchmark category leveraging existing workloads including TPC-DS, TPCx-HS, HiBench, etc. [20].
- TPC-IoT: A new committee with the mission of exploring standards for Internet of Things workloads [22].

3 TPCTC Conference Series

To keep pace with these rapid changes in technology, in 2006, the TPC initiated the conference series on performance analysis and benchmarking. The TPCTC has been challenging Industry experts and researchers to develop innovative techniques for performance evaluation, measurement, and characterization of hardware and software systems. Over the years it has emerged as a leading forum to present and debate the latest and greatest in the world of benchmarking. The topics of interest included:

- Big Data
- Data Analytics
- Cloud Computing
- In-memory databases
- Social media infrastructure
- Security

- Complex event processing
- Internet of Things
- Database Optimizations
- Disaster tolerance and recovery
- Energy and space efficiency
- Hardware innovations
- Hybrid workloads
- Virtualization
- Lessons learned in practice
- Enhancements to TPC workloads
- Data Integration

A short summary of TPCTC conference proceedings are listed below.

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 24th to August 28th, 2009. The keynote speaker was Michael Stonebraker. [1, 3]

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 13th to September 17th, 2010. The keynote speaker was C. Mohan. [4]

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 29th to September 3rd, 2011. The keynote speaker was Umeshwar Dayal. [5, 23]

The fourth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2012) was held in conjunction with the 38th International Conference on Very Large Data Bases (VLDB 2012) in Istanbul, August 27th to August 31st, 2012. The keynote speaker was Michael Carey. [6, 24]

The fifth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2013) was held in conjunction with the 39th International Conference on Very Large Data Bases (VLDB 2013) in Riva del Garda, Trento, Italy, August 26th to August 30st, 2013. The keynote speaker was Raghu Ramakrishnan. [7]

The sixth TPC Technology Conference on Performance Evaluation and Benchmarking (TPCTC 2014) was held in conjunction with the 40th International Conference on Very Large Data Bases (VLDB 2014) in Hangzhou, China, September 1st to September 5th, 2014. [8]

TPCTC has had significant positive impact to the TPC. TPC is able to attract new members from industry and academia to join the TPC. The formation of the Workshop Series on Big Data Benchmark (WBDB) was inspired by TPCTC [26]. TPCTC also triggered development of benchmark standards in virtualization, data integration and Big Data. The formation of a working group on IoT was a direct result of TPCTC conferences.

4 Big Data and the Internet of Things [25]

Industry and technology landscapes are changing rapidly. Two of the technologies that will change the world in the next decade years are expected to be Big Data and the Internet of things (IoT).

Big Data: Big Data is a popular term now that describes the exponential growth of data, often defined by the 5Vs, associated technologies like storage, and how to effectively process and drive business values. The Big Data technology and services market represents one of the fast-growing, multi-billion dollar, worldwide market that is expected to grow to a $60 billion marke, driving $300 billion in worldwide IT spending directly or indirectly by 2020.

Foreseeing the importance, in 2014 the TPC developed the TPC Express Benchmark HS (TPCx-HS) to provide the industry with verifiable performance, price-performance, and availability metrics of hardware and software systems dealing with Big Data. This standard can be used to assess a broad range of system topologies and implementation of Hadoop systems in a technically rigorous and directly comparable, and vendor-neutral manner. This is the first major step while the TPC continues to enhance and develop new standards in this area such as TPC-DS with support for Hadoop and TPC-Big Bench.

Internet of Things (IoT): IoT has emerged in the last few years, poised to transform virtually every major market segment which contains a complex mix of technologies and products, from data collection and data curation to complex analytics exploiting the data generated by an exploding number of connected devices. According to IDC, the global IoT market will grow from $665 billion in 2014 to $1.7 trillion in 2020. To put that in perspective, it's an absolutely enormous figure; only 16 economies in the world had gross domestic products exceeding $1 trillion in 2014.

As the IoT ecosystem evolves in the enterprises, it is eminent to have a set of standards that enable effective comparison of hardware and software systems and topologies in a technology and vendor-neutral manner. Continuing its commitment to bring relevant standards to the industry, today the TPC announced the formation of the TPC-IoT benchmark committee, tasked with developing industry standard benchmarks for benchmarking hardware and software platforms associated with IoT [25].

5 Outlook and Acknowledgements

Over the last few years the TPC has truly reinvented itself by investing in the Technology Conference Series, Express benchmark initiatives, and new areas including Big Data and the Internet of Things. TPC remains committed to developing relevant standards in collaboration with industry and research communities, and to continue to enable fair comparison of technologies and products in terms of performance, cost of ownership and energy efficiency.

Developing benchmark standards requires a huge effort to conceptualize, research, specify, review, prototype, and verify the benchmark. The authors acknowledge the work and contributions of past and present members of the TPC.

References

1. Stonebraker, M.: A new direction for TPC? In: Nambiar, R., Poess, M. (eds.) TPCTC 2009. LNCS, vol. 5895, pp. 11–17. Springer, Heidelberg (2009)
2. Nambiar, R., Poess, M.: Keeping the TPC relevant! PVLDB **6**(11), 1186–1187 (2013). Nambiar, R., Wakou, N., Masland, A., Thawley, P., Lanken, M., Carman, F., Majdalany, M.: Shaping the Landscape of Industry Standard Benchmarks: Contributions of the Transaction Processing Performance Council (TPC), TPCTC 2011, pp. 1–9
3. Nambiar, R., Poess, M.: Performance Evaluation and Benchmarking. LNCS, vol. 5895. Springer, Heidelberg (2009). ISBN 978-3-642-10423-7
4. Nambiar, R., Poess, M.: Performance Evaluation, Measurement and Characterization of Complex Systems. LNCS, vol. 6417. Springer, Heidelberg (2011). ISBN 978-3-642-18205-1
5. Nambiar, R., Poess, M.: Topics in Performance Evaluation, Measurement and Characterization. LNCS, vol. 7144. Springer, Heidelberg (2012). ISBN 978-3-642-32626-4
6. Nambiar, R., Poess, M.: Selected Topics in Performance Evaluation and Benchmarking. LNCS, vol. 7755. Springer, Heidelberg (2013). ISBN 978-3-642-36726-7
7. Nambiar, R., Poess, M.: Performance Characterization and Benchmarking. LNCS, vol. 8391. Springer, Heidelberg (2014). ISBN 978-3-319-04935-9
8. Nambiar, R., Poess, M.: Performance Characterization and Benchmarking. Traditional to Big Data. LNCS, vol. 8904. Springer, Heidelberg (2015). ISBN 978-3-319-15349-0
9. Huppler, K., Johnson, D.: TPC express – a new path for TPC benchmarks. In: Nambiar, R., Poess, M. (eds.) TPCTC 2013. LNCS, vol. 8391, pp. 48–60. Springer, Heidelberg (2014)
10. Nambiar, R., Poess, M.: Transaction performance vs. Moore's law: a trend analysis. In: Nambiar, R., Poess, M. (eds.) TPCTC 2010. LNCS, vol. 6417, pp. 110–120. Springer, Heidelberg (2011)
11. TPC-C Specification
12. TPC-H Specification
13. TPC-E Specification
14. TPC-DS Specification
15. Nambiar, R., Poess, M.: The making of TPC-DS. In: VLDB 2006, pp. 1049–1058 (2006)
16. Nambiar, R., Poess, M.: Why you should run TPC-DS: a workload analysis. In: VLDB 2007, pp. 1138–1149 (2007)
17. TPC-VMS Specification
18. TPC-xHS Specification
19. Nambiar, R., Poess, M., Dey, A., Cao, P., Magdon-Ismail, T., Qi Ren, D., Bond, A.: Introducing TPCx-HS: the first industry standard for benchmarking big data systems. In: Nambiar, R., Poess, M. (eds.) TPCTC 2014. LNCS, vol. 8904, pp. 1–12. Springer, Heidelberg (2015)
20. Chaitanya, K., et al.: Discussion of BigBench: a proposed industry standard performance benchmark for big Data. In: Nambiar, R., Poess, M. (eds.) TPCTC 2014. LNCS, vol. 8904, pp. 44–63. Springer, Heidelberg (2015)
21. Bond, A., Johnson, D., Kopczynski, G., Taheri, H.: Architecture and performance characteristics of a PostgreSQL implementation of the TPC-E and TPC-V workloads. In: Nambiar, R., Poess, M. (eds.) TPCTC 2013. LNCS, vol. 8391, pp. 77–92. Springer, Heidelberg (2014)
22. Nambiar, R.: Vendor-neutral benchmarks drive tech innovation. http://data-informed.com/vendor-neutral-benchmarks-drive-tech-innovation/

23. Dayal, U., Wilkinson, K., Simitsis, A., Castellanos, M., Paz, L.: Optimization of analytic data flows for next generation business intelligence applications. In: Nambiar, R., Poess, M. (eds.) TPCTC 2011. LNCS, vol. 7144, pp. 46–66. Springer, Heidelberg (2012)
24. Carey, M.J., Ling, L., Nicola, M., Shao, L.: EXRT: towards a simple benchmark for XML readiness testing. In: Nambiar, R., Poess, M. (eds.) TPCTC 2010. LNCS, vol. 6417, pp. 93–109. Springer, Heidelberg (2011)
25. Nambiar, R.: Benchmarking Internet of Things (IoT). http://blogs.cisco.com/datacenter/industry-standards-for-benchmarking-iot
26. Baru, C., Bhandarkar, M., Nambiar, R., Poess, M., Rabl, T.: Benchmarking big data systems and the big data top 100 list