

Does the TPC still have relevance?

H. Reza Taheri
HPTS 2017, 9-Oct-2017

vmware®

© 2016 VMware Inc. All rights reserved.

Outline

- History of the TPC
- Where things stand today
- Why the decline?
- The way forward
- Not gonna get into how representative the workloads are!!

How did we end up with the TPC?

- 1980s was the era of the Wild West of database benchmarking
- Performance claims based on own benchmarks
- Vendors published TP1/Debit-Credit and Wisconsin results
 - Standard benchmark definitions, but wildly-varying implementations, and frequent cutting of corners
 - Did not result in peace
 - Check the DeWitt clause of your software

How did we ensure benchmarking was fair and valid in the days before TPC?



AT&T
Information Systems

1100 East Warrenville Road
Naperville, IL 60566

December 12, 1985

Mr. James M. Gray
Tandem Computers
19333 Vallco Parkway
Cupertino, CA 95014

Dear Jim,

I am sending you a copy of my implementation of the Debit-Credit (TP1) benchmark as we discussed on the phone. I would appreciate your reviewing the code to see if it is a faithful CICS implementation of TP1 as it is described in the Tandem Technical Report and the Datamation article.

Thank you again for taking the time to help.

Very truly yours,

A handwritten signature in cursive script that reads "H. R. Taheri".

H. R. Taheri
Member of Technical Staff

Att.
Debit-Credit benchmark

History of the TPC

- Brainchild of Jim Gray and Omri Serlin; Tom Sawyer co-founder
- Took root at HPTS 1987
- Transaction Processing Performance Council (TPC) was formed on August 10, 1988 with eight member companies
 - Control Data Corp.; Digital Equipment Corp.; ICL; Pyramid Technology; Stratus Computer; Sybase; Tandem Computers; and Wang Laboratories
- The Standard Performance Evaluation Corporation (SPEC) was formed at about the same time
 - Focus on CPU performance
 - Releases benchmark kits versus paper functional specification
 - Still going strong
- In the 1990s, you couldn't sell h/w or s/w without TPC results
- Publication rate today is a fraction of the mid 1990s level

The value of the TPC

- Thorough standard specifications
 - TPC-E specification is 287 pages; TPC-DS 137 pages; TPC-C 132 pages; TPCx-HCI 286 pages
- Detailed Full Disclosure Report
 - Describe the result in enough detail to allow another member to reproduce it
 - Typical disclosure 100-300+ pages
 - Much shorted for the new Express benchmarks
 - Moving towards shorter FDRs supplemented by downloadable files
- Independent audit
- Fair use rules
 - When can you claim TPC results?
 - How are you allowed to compare to a competitor's result

TPC-A/B

- Start with Debit-Credit/TP1, make it a strong standard
- First TPC-A result in July 1990
 - Codified Debit-Credit
 - ~40 companies published results
- First TPC-B result in mid-1991
 - Codified, standardized TP1
- Both obsoleted on 6/6/95

TPC-C

- Approved as a new benchmark in July, 1992
- OLTP
 - More complex than TPC-A
 - More detailed rules; harder to cheat
- ~26 companies have published 782 results

TPC-D

- Decision Support
- Approved in April, 1994
- By February 1998, 28 official TPC-D results
- Obsoleted in June, 1999

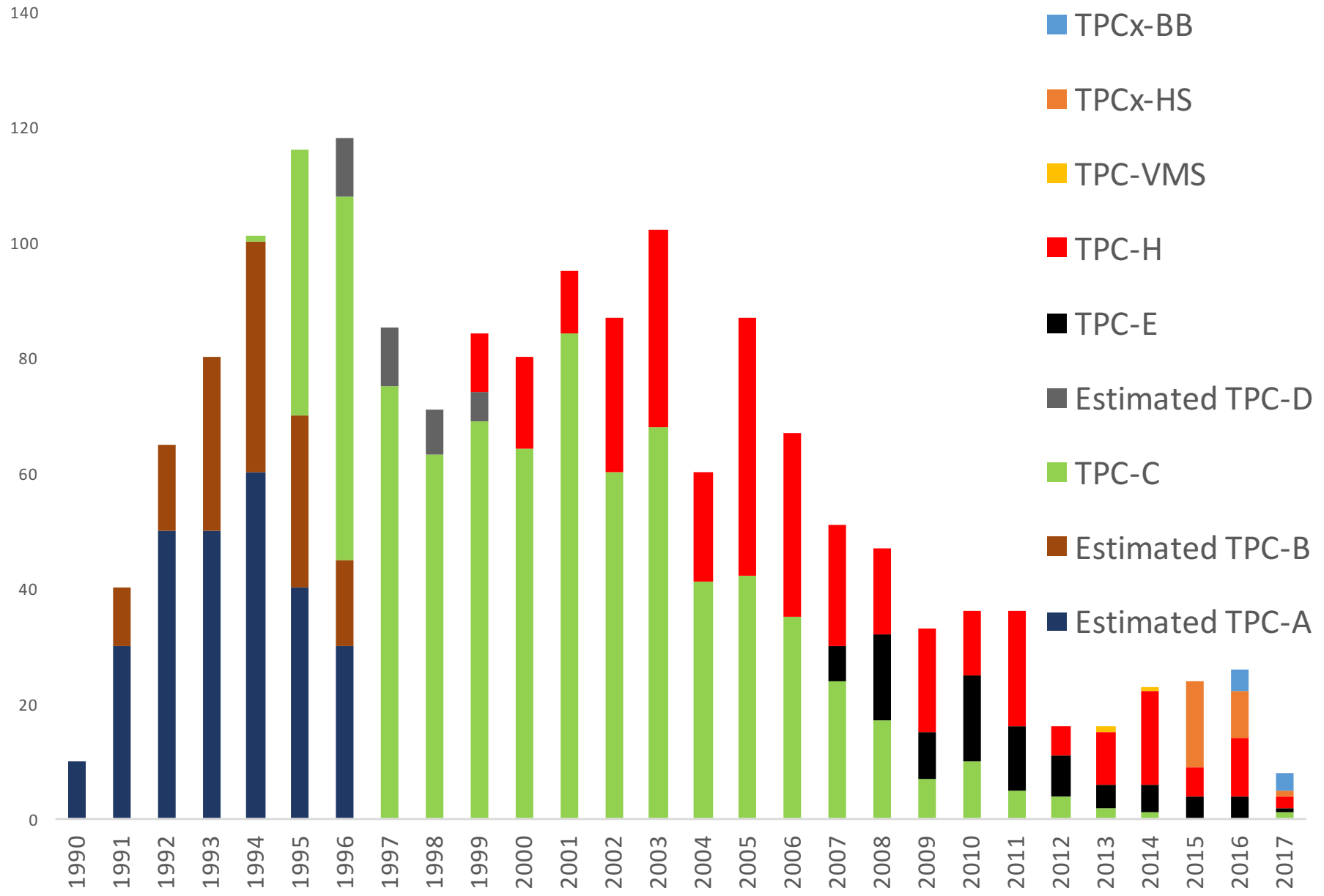
TPC-H

- The current DSS benchmark standard
- Major improvement over TPC-D
- 326 results
 - 10 different databases

TPC-E

- The current OLTP benchmark standard
 - More complex than TPC-C
 - 80 results
 - Compare to 782 for TPC-C
 - Only one DBMS has published
 - Compare to 9 for TPC-C
 - 9 companies have published
 - Compare to 26 for TPC-C

TPC publications by benchmark over the years



The hits and misses

- Hits
 - Rigor
 - Audit
 - Price/performance
 - Leveled playing field
 - Detailed spec and FDR
 - Scaling of DB size with throughput
 - Paper functional specification => vendor-neutral; anyone can implement
- Misses
 - Paper specification
 - Slow development cycle
 - High cost of entry
 - “representative”, up-to-date workloads???
 - Benchmark specials

Benchmark hits and misses

- Hits
 - TPC-A/B
 - TPC-C
 - TPC-E
 - TPC-H
- Misses
 - TPC-W
 - TPC-S
 - Original TPC-E
 - TPC-R
 - TPC-DI
- Jury is still out
 - TPC-DS
 - TPCx-V / TPCx-HCI

Why have benchmark publications dwindled?

- Three reasons (you may have more!)
 - Performance parity
 - Benchmarking kits are very expensive to develop
 - Nearly futile if you want to publish
 - Prohibitive cost if you want to use it internally or for a paper
 - There is no competition left!

HPE

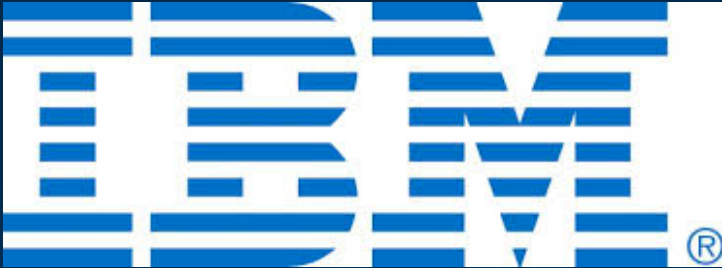


- Tandem one of the original 8
- DEC one of the original 8

Dell



IBM



Fujitsu



- ICL one of the original 8
- Pyramid one of the original 8

Oracle



Actian

INGRES



And the ones that faded away

- NCR
- AT&T
- AST research
- Olivetti
- SCO
- SGI
- MIPS
- Prime
- Wang (one of the original 8)
- CDC (one of the original 8)

Is there competitive pressure in the TPC?

- 6 database companies
 - DB2, SQL Server, Oracle
 - Actian, SAP, Teradata
- Little differentiation between various X86 servers
- 21 members
- There is little pressure to:
 - Publish
 - Contribute
 - Make the workloads more realistic

Express benchmarks to the rescue???

- TPC has been switching gears
 - Instead of paper functional specifications, release full benchmarking kits
 - Lower the cost of entry
 - This actually speeds up benchmark development
 - What SPEC has done for years
 - Simplifies the audit
 - Look for new market segments

Express benchmarks

- TPCx-HS
 - Does for Terasort what TPC-A did for Debit-Credit
 - Benchmarking kit based on open source Terasort
 - 24 results
- TPCx-BB
 - For Hadoop-based Big Data systems
 - Code came mainly from Intel
 - 7 results in a year
- TPCx-IOT
 - For IoT gateways; workload represents data injection into an IoT Gateway with continuous real-time analytic queries

TPCx-V and TPCx-HCI

- Started TPC-V in 2009 to measure performance of virtualized servers with a heavy TPC workload
 - Was badly needed back then
 - No so much when the benchmark was released in 2015
 - But we ended up with a complete, end-to-end, self-auditing/self-validating kit; 75K LOC written from scratch
 - PostgreSQL
 - Linux VMs
 - Rev 2.0.0 was just approved; several vendors waiting for an auditor to be certified

TPCx-HCI

- The TPCx-V benchmarking kit is a perfect tool for measuring performance of Hyper-Converged Infrastructure systems
 - Measures I/O performance of HCI, as well as its CPU cost of I/Os
 - Workload elasticity means you want frequent VM migrations to rebalance
 - Measures efficiency of VM migrations in the cluster
 - Needs uniformity of data access on all the nodes of an HCI cluster
 - Requires powering down a node, and reporting impact on performance and recovery time
- Strong interest in the benchmark
 - All the old-guard TPC members have a horse in the race
 - 2 new members have joined
 - The field is still young, and performance is a big unknown
 - The benchmark was approved last month
 - 7-month development cycle
- I am betting the farm on this benchmark

Wrap-up

- Not a rosy future with the status quo
- Need to recreate the excitement and the competition of the early days
- Express benchmarks – full end-to-end kits written by SMEs – are the future
 - Quick development cycles
 - Publically-available kits for anyone who wants to experiment with the benchmark
 - Cheaper process for anyone who wants to publish