Industry Benchmarks and Proof of Concepts
Contents

1 Overview
2 SPEC Benchmarks
3 TPC and SAP Benchmarks
4 I/O Benchmark Tools
5 RPE2
6 Proof of Concept
7 Conclusion
Overview

- CPU and Server benchmarks
  - SPEC  [www.spec.org](http://www.spec.org)

- Application specific database benchmarks
  - TPC-C, TPC-E  [www.tpc.org](http://www.tpc.org)
  - TPC-H  [www.tpc.org](http://www.tpc.org)
  - SAP  [www.sap.com](http://www.sap.com)

- I/O benchmark tools
  - Iometer  [www.iometer.org](http://www.iometer.org)
  - Vdbench  [www.oracle.com](http://www.oracle.com)
  - Orion  [www.oracle.com](http://www.oracle.com)
  - SPC  [www.storageperformance.org](http://www.storageperformance.org)
Contents

1 Overview

2 SPEC Benchmarks

3 TPC and SAP Benchmarks

4 I/O Benchmark Tools

5 RPE2

6 Proof of Concept

7 Conclusion
SPEC Benchmarks

- Well documented [1]
  - 12 integer benchmark tests
  - 18 floating point benchmark tests

- Programming languages
  - C
  - C++
  - Fortran
  - Pearl

---
SPEC Benchmarks

- SPEC does not use Oracle data types

- Following Oracle data types use hardware arithmetic
  - PLS_INTEGER, SIMPLE_INTEGER
  - BINARY_FLOAT, BINARY_DOUBLE
  - SIMPLE_FLOAT, SIMPLE_DOUBLE
  - Performance of these data types may correlate with SPEC benchmark results
SPEC Benchmarks

- Following most used data types in Oracle applications need software library for basic operations and SQL built in functions
  - NUMBER 1)
  - VARCHAR2, CHAR, NCHAR
  - DATE, TIMESTAMP
  - BLOB, CLOB, BFILE
  - Performance of these data types may not correlate with SPEC benchmark results

1) This very unique numerical data type uses a binary code decimal (BCD) implementation. This data type can not use hardware arithmetic.
Contents

1 Overview
2 SPEC Benchmarks
3 TPC and SAP Benchmarks
4 I/O Benchmark Tools
5 RPE2
6 Proof of Concept
7 Conclusion
TPC and SAP Benchmarks

No benchmark results for YOUR platform

- Customer wants to know the key performance metrics of his platform
- Huge effort to run TPC on customer platform
  - Complexity
  - Rules
- Impossible to run SAP benchmarks on customer platform
  - Lack of documentation
TPC and SAP Benchmarks

Strict TPC rules regarding setup, sizing and usable features

- A benchmark should be operational on each customer system with any database size
- Customers would like to use ALL database features which may be helpful
  - to increase overall efficiency
  - to get most performance out of each license dollar
TPC and SAP Benchmarks

Unrealistic hardware configuration

- For a 10 TByte TPC-H benchmark, vendors use about 3’000 disks
- Most customers use by far less disks for a single 10 TByte database
TPC and SAP Benchmarks

Only one load profile: complete CPU saturation

- TPC-C and SAP measures the transaction rate at highest system load with > 95% CPU utilization
- Every very system administrator wants to avoid this situation in OLTP Systems
- Such a high load is an exception in OLTP Systems, not the rule
TPC and SAP Benchmarks

Unpractical performance metrics

- The TPC performance metric is just one number, e.g.
  - TPC-H: Qph@Size (Query per hour at given database size)
  - What does it mean?

- TPC and SAP do not make any predictions about
  - CPU performance
  - Server performance
  - Storage performance
  - Database performance for data load, data scan, backup, etc.

- TPC benchmark metrics do not relate to something that can be readily understood by users
TPC and SAP Benchmarks

Unpractical performance metrics

- TPC and SAP provide no real metrics for capacity planning
  - No values about best-case behavior (cached) and worst-case behavior (non cached)
  - No values about system behavior with increasing load from 1 process to n processes up to system saturation
Some quotes:

- *In short, TPC has become vendor-dominated, and it is time for TPC to reinvent itself to serve its customer community.*

- *At the present time, most TPC’s benchmarks have been politically engineered through vendor negotiation, and therefore lack relevance to any real world problem.*

- *In other words, TPC should become customer-focused and not vendor-focused.*

---

SAP Benchmarks

SAP Sales and Distribution (SD) Benchmark

- SAP does not document storage system configuration for its benchmarks
  - Storage system is not the limiting factor in SAP benchmarks
  - What is about price of such a storage system?

- Benchware opinion
  - Storage system performance is essential for Oracle database environments
  - Storage system performance should be part of
    - platform evaluation
    - platform performance calibration
Contents

1 Overview
2 SPEC Benchmarks
3 TPC and SAP Benchmarks
4 I/O Benchmark Tools
5 RPE2
6 Proof of Concept
7 Conclusion
I/O Benchmark Tools

Comparison of I/O Benchmarks

- **Storage System I/O Performance**
  - Useful to test storage system performance at port level
  - Vendors data sheet numbers
I/O Benchmark Tools

Comparison of I/O Benchmarks

- **Server System I/O Performance**
  - Tools like vdbench, Iometer, Orion, etc. just generate I/O system calls, but no further I/O processing
  - Useful to analyze transfer performance between storage system and server system
  - Unable to benchmark storage grids
  - Unable to benchmark Oracle ASM infrastructure
I/O Benchmark Tools

Comparison of I/O Benchmarks

- **Database System I/O Performance**
  - Most complex I/O operation
  - Database buffer cache management
    - find a free slot
    - replace older blocks
    - synchronize access to buffer cache
    - database block consistency checks
  - Database I/O needs much more CPU resources than simple I/O generator
    - Rule of thumb: 25’000 IOPS per x86 core
    - Throughput does not scale linear
  - `dbms_resource_manager.calibrate_io` does not recognize hybrid storage systems and delivers wrong results
## Contents

1. Overview
2. SPEC Benchmarks
3. TPC and SAP Benchmarks
4. I/O Benchmark Tools
5. **RPE2**
6. Proof of Concept
7. Conclusion
Relative Performance Estimate 2 (RPE2)

- Previous methodology from Ideas International, since 2012 Gartner Inc.
- RPE2 is a theoretical performance estimate and not an actual observed measurement of server performance.
- It is based on published benchmark results and relative performance ratings from server manufacturers.
- The published or estimated performance points for each server processor option are aggregated by calculating a geometric mean value.
- The current RPE2 set includes the following six benchmark inputs in its calculation: SAP SD Two-Tier, TPC-C, TPC-H, SPECjbb2005, and two SPEC CPU2006 components.

Because RPE2 is build on industry benchmarks, it inherits all their shortcomings
Contents

1 Overview
2 SPEC Benchmarks
3 TPC and SAP Benchmarks
4 I/O Benchmark Tools
5 RPE2
6 Proof of Concept
7 Conclusion
Proof of Concept

Prototype in lab environment

- Time Consuming – Complicated – Expensive
  - Installation
  - Data Migration
  - Data masking
  - Simulation of large OLTP user populations
  - Simulation of SOA and ESB interfaces
  - ...
Proof of Concept

Prototype in lab environment

- **Application as load generator**
  - Measures performance of current software on new hardware technology
  - PoC provides a snapshot result - any change of application or data may change PoC result

- **PoC does not necessarily**
  - Identify performance capabilities and limitations of new hardware
  - Reflect correct price performance ratio of new hardware technology
Proof of Concept

Customer statement

“My applications and my data are so specific, I must run a proof-of-concept”

- Platform components like CPU, storage system, database system or network do not recognize neither any specific application nor any specific data
Contents

1  Overview
2  SPEC Benchmarks
3  TPC and SAP Benchmarks
4  I/O Benchmark Tools
5  RPE2
6  Proof of Concept
7  Conclusion
Conclusion

Weaknesses of industry benchmarks and proof of concepts

- Current industry benchmarks are inappropriate to calibrate performance of customers Oracle database platform

- Proof of Concept with application software
  - May be helpful to become familiar with new technologies
  - May be helpful for a draft capacity planning of new platforms
  - Are in general inappropriate to test performance limitations of new technologies and therefore are inappropriate for price performance considerations