

# Component vs. End-to-End Benchmarks

Meikel Poess

Oracle Corporation

# Typical Elements of Today's Industry Standard Benchmarks

- Scope
  - Who can play in the benchmark
- Workload
  - Data
  - Operations on data
- Execution Rules
  - What operations to run when and on what data
- Measurement Methodology
  - What and How to report measurements
- Metric(s)
  - How to combine measurements into one or multiple numbers
  - Where to publish the metric and how to compare with others
- Referee
  - How to guarantee correctness of measurements (a.k.a. AUDIT)

# Component Benchmarks

- Limited in their scope
  - Measure performance of small number of system components (hardware or software)
  - Extreme case, only one system component
    - CPU, disk, Graphic Card
  - Example: SPEC CPU2006 benchmark
    - Benchmark to test performance of a system's processor, memory subsystem and compiler
    - Benchmark publication 6 pages
    - Single system, no disk (except for OS disk)

NovaScale T840 F3



<http://www.bull.com/catalogue/details.asp?tm=nsu-tower&opt=nst840f03&dt=dw>

# Component Benchmarks

- Easier to develop than an End-to-End benchmark
  - Scope is limited
- Easier to implement for benchmark sponsors than End-to-End benchmark
  - For components that expose standardized API's a benchmark can be developed to run as-is
- Easier to run
- Less expensive to run
- Does not express a customer environment
- Does not expose interaction problems of components

# End-to-End Benchmark

- Broad in scope:
  - Measure performance of entire systems
  - Extreme case, entire lifecycle of a system
    - E.g. TPC-C benchmark
    - Benchmark publication 741 pages (<http://c970058.r58.cf2.rackcdn.com/fdr/tpcc/Bull-Escala-PL6460-080615-FDR.pdf>)



## 128 Clients

IBM System x3550  
Dual-core 2.0GHz Intel® Xeon™  
4MB L2 Cache  
1GB Memory  
73.4GB SAS Drive  
Integrated dual-port 10/100/1000 Ethernet

## Escala PL6460R

32 Processor Chips with  
64 5.0GHz POWER6™ Cores  
32MB L3 Cache per chip  
4096GB Memory  
8 146GB Internal SCSI Drives  
68 4Gb dual-port Fibre Channel Adapters  
10 10/100/1000 Ethernet Adapters

## Storage

68 IBM System Storage DS4800  
784 IBM System Storage DS4000 EXP810  
10,992 73.4GB 15K RPM 4Gbps Drives

# End-to-End Benchmark

- More difficult to develop than component benchmark
  - Wide scope, which might include many components and vendors
- More difficult to implement than component benchmark
  - Benchmark tends to be specification driven, i.e. vendors need to develop a Benchmark KIT
  - Expresses real customer environment
  - Exposes component interaction problems

# Specification vs. KIT based

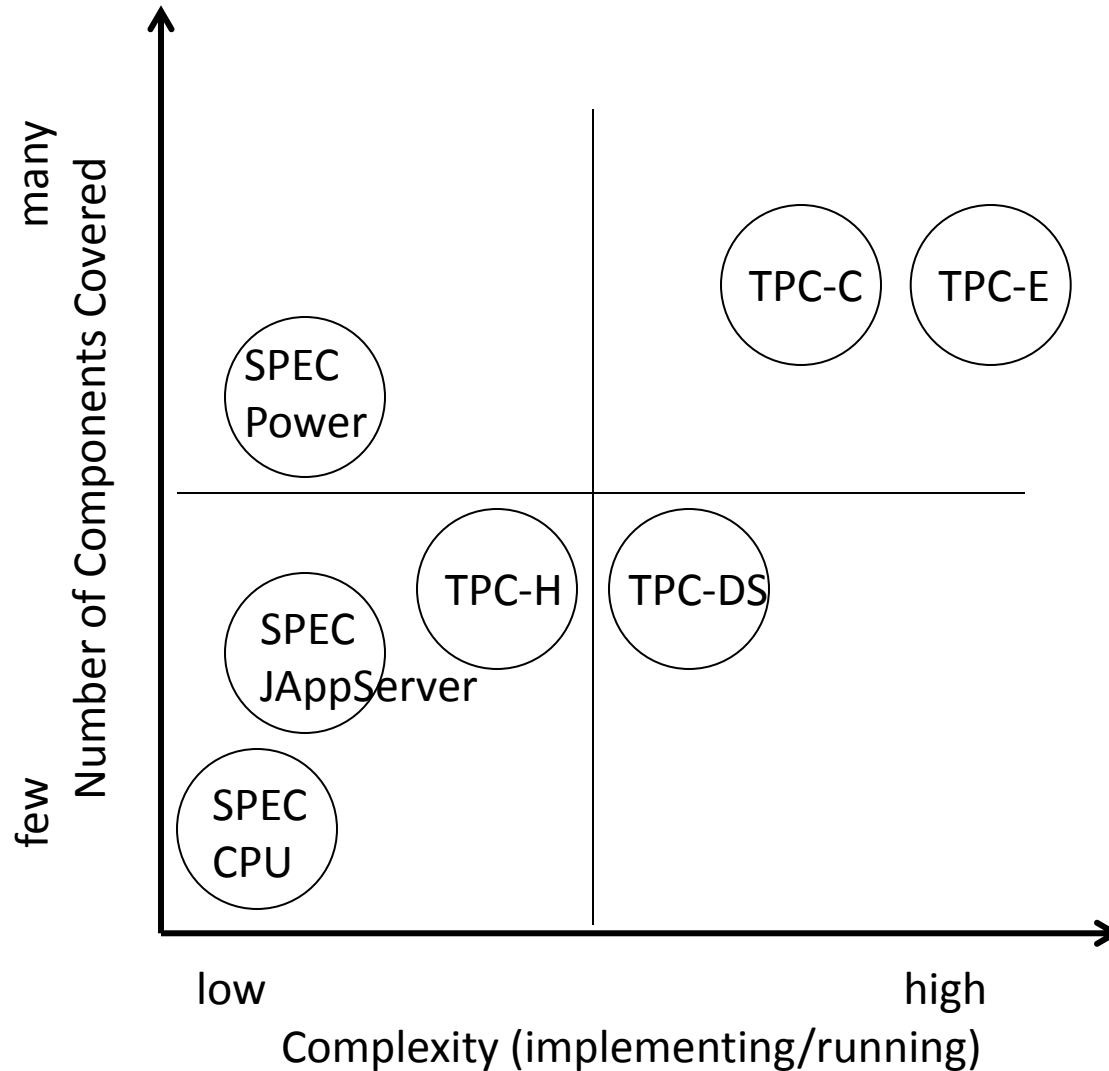
## **Specification based**

- Written in English prosa
- Needs to cover all possible cases
- Leaves room for interpretation
- Requires vendors to built Benchmark KIT, which is an interpretation of the spec
- Interpretation leads to problems

## **KIT based**

- Written to be executed as-is
- Very specific for an limited number of systems
- Leaves no room for interpretation
- Vendors can run Benchmarks without further investments

# Classification





# Questions to the Community

- Shall we develop a Big Data Benchmark as a Component Benchmark or as an End-to-End benchmark?
- Is it possible to develop a benchmark as a set of component benchmarks that could be combined into an end-to-end benchmark
- Shall we strictly adhere to a KIT based benchmark or a specification based benchmark?