Energy Benchmarks: A Detailed Analysis

Meikel Poess, Oracle

Raghunath Nambiar, Hewlett Packard

Kushagra Vaid, Microsoft

John M. Stephens, Gradient Systems

Karl Huppler, IBM

Evan Haines, ICF International

Agenda

1. Motivation
2. Overview of benchmark consortia:
   - Transaction Processing Performance Council
   - Standard Performance Corporation
   - Storage Performance Council
   - US Environmental Protection Agency
3. Benchmark Comparisons
   - What hardware components are measured
   - What workload is run
   - What are the main characteristics of the energy metrics
   - What are the accuracy and calibration requirements
4. Conclusion

Motivation

- Energy efficiency is the top priority of today's data center mangers, because of:

1. Increased demand for high performance and high capacity

- Enterprise data warehouses double in size every three years, because business generate more data, store more data and retain data for longer time

- Petabyte data warehouses are increasingly common (AT&T, eBay, Yahoo, Walmart are >5 petabytes)

- Energy consumption of a typical data center is 20 times higher than that of a commercial building

- In 2006 data centers consumed about 60 Billion KWh equaling about 1.5 percent of the total US electricity consumption [1]

[1] EPA Report

E-Energy 2010 M. Poess et al.

Motivation

1. Increased energy cost

- Assuming 300GB disk drives, RAID10, 8W and $.11/kWh
- The disk drives alone for a 1 Petabyte data warehouse would cost about $53,000 a year
- Not counting:
  - Servers
  - Enclosures
  - Network gear
  - Performance

Motivation

- Consequently standard performance consortia started adding energy metrics into their benchmarks:
  - Transaction Processing Performance Council (TPC)
  - Standard Performance Evaluation Corporation (SPEC)
  - Storage Performance Council (SPC)
- US Environmental Protection Agency (US EPA)

Motivation

- Each standard and approach is slightly different:
  - Hardware components measured
  - Workload executed
  - Main characteristics of the energy metrics
  - Accuracy and calibration requirements

- In this paper we provide an overview of the four aforementioned approaches

Transaction Processing Performance Council (TPC)

- "Founded in 1988 by a group of companies, who were interested in delivering verifiable performance data to the industry"
- Non-profit organization
- Started publishing benchmarks in 1989
- Two OLTP benchmarks TPC-C and TPC-E
- One data warehouse benchmark TPC-H
- Is working on ETL and Virtualization benchmarks
- Benchmark tend to have a long shelf live

Standard Performance Evaluation Corporation (SPEC)

- "SPEC was established in 1988 to maintain and endorse a standardized set of relevant benchmarks that can be applied to the newest generation of high-performance computers"

- Non-profit corporation

- Starting with component type benchmarks such as SPEC92 it now offers higher level benchmarks, such as SPECjAppServer2004 or SPECjEnterprise2010

- Very large suite of benchmarks

- Benchmarks have a shelf live of a couple of years

Storage Performance Council (SPC)

- "Founded in 1997 to define, standardize, and promote storage subsystem benchmarks"
- Non-profit organization
- Started publishing benchmarks in 2001
- SPC has developed and publicized benchmarks focused on:
  - Storage subsystems SPC-1
  - Components of storage subsystems SPC-1C:
    - Adapters
    - Controllers
    - Storage area networks

United States Environmental Protection Agency (EPA)

- Established in 1970 to protect human health and to safeguard the natural environment—air, water, and land—upon which life depends

- EPA's Energy Star® Computer program promotes the manufacturing and purchasing of energy-efficient personal computers, monitors and printers

- Working on Energy Star Program for Servers and Data Centers

Hardware Components

- Modern computer systems are often split into multiple tiers:
  - Tier 1: User interface
  - Tier 2: functional process logic (application server)
  - Tier 3: Database Management System (DBMS)

- Each tier can be further split into components, e.g.
  - Servers, storage, controller etc.

- Each benchmark focuses on either a subset of the above tiers or the entire stack

Hardware Components Comparison

Measured Hardware Component		TPC-C	TPC-E	TPC-H	SPEC Power_ssj 2008	SPEC web2009	SPC-1/E	SPC-1C/E	EnergyStar for Server
Tier 1	System enclosure[1]								
	Secondary storage[2]							✓	
	Storage controller							✓	
	NIC							✓	
Tier 2	Server enclosure[1]	✓	✓	✓	✓	✓			✓
	Secondary storage[2]	✓	✓	✓		✓		✓	✓
	Storage controller	✓	✓	✓		✓		✓	
	Storage enclosure	✓	✓	✓		✓		✓	
	NIC	✓	✓	✓		✓		✓	
Tier 3	Server enclosure[1]	✓	✓				✓		
	Secondary storage[2]	✓	✓				✓	✓	
	Storage controller	✓	✓				✓	✓	
	Storage enclosure	✓	✓				✓	✓	
	NIC	✓	✓				✓	✓	
Network switch Tier 1 and 2		✓	✓			✓			
Network switch Tier 2 and 3		✓	✓						

[1] Includes board, fans, processor and DRAM
[2] Includes memory that retains information even when not powered (flash, hard disks etc.)

E-Energy 2010 M. Poess et al.

Workload

- In general workloads may vary from completely synthetic to particular customer scenarios

- The workload standard organizations employ tend to be real work like

  – Model scenarios that originate from the real world

  – Can be applied to a broad spectrum of industry applications

  – Are relatively easy to implement and control

Workload Comparison

Benchmark	Application	Workload description
TPC-C	3-Tier OLTP	OLTP system of an order-entry system
TPC-E	3-Tier OLTP	OLTP workload of a brokerage firm
TPC-H	DSS	Ad-hoc, decision support queries
SPECpower_ssj20 09	Server sideJAVA	Measures performance of server side Java applications
SPECweb2009	Web Server	Measures web server performance with simulated database tier
SPC-1/E	Simulated OLTP Database tier	Performance of a storage subsystem while performing the typical functions of a business critical application.
SPC-1C/E	Simulated OLTP Database tier	Performance of a storage subsystem while performing the typical functions of a business critical application.
Energy Star for Servers	TBD	Tools incorporated into future versions of the program will incorporate tests that mirror activity expected in multiple application scenarios.

Energy Metric Comparison

Metric Characteristic Comparison

Energy Metric Characteristics / The Energy Metric	TPC-C	TPC-E	TPC-H	SPEC Power ssj 2008	SPECWeb 2009	SPC-1/E	SPC-1C/E	Energy Star for Server[5]
Is mandatory (m) or optional (o)	o	o	o	m	m	o	o	m
Reports total energy consumed during benchmark operation	✓	✓	✓			✓	✓	
Reports energy consumption at peak performance levels	✓	✓	✓	✓	✓	✓	✓	✓
Reports energy consumption at idle state	✓	✓	✓	✓	✓	✓	✓	✓
Reports energy consumptions at intermediate performance levels				✓	✓	✓	✓	
Unite performance and energy				✓	✓	✓	✓	
Unite performance, price and energy						✓	✓	
Reports yearly energy cost						✓	✓	
Allows reporting of energy consumption of individual components	✓	✓	✓				✓	

Accuracy and Calibration Requirements

Requirement	TPC	SPEC	SPC	Energy Star for Server
Power Analyzer Accuracy [%]	2	2	2	2
Power Analyzer certification	NIST	NIST	NIST	NIST
Power Analyzer calibration	yearly	yearly	yearly	yearly
Ambient temperature [°C]	22-23	22-23	logged	18-27
Low end dew point [°C]	none	none	none	5.5
High end dew point [°C]	none	none	none	15
Atmospheric pressure [atm]	<1.1	none	none	none
Result certification	auditor	peer	auditor peer	self/audit

E-Energy 2010 M. Poess et al.

17

Conclusion

- Standards have some similarities
  - Peak and idle states
  - Power analyzer requirements
    - Accuracy
    - Certification
    - Calibration

- But there are also many differences
  - Workload
  - Mandatory vs. optional energy reporting
  - How energy consumption is reported
  - …

TPCTC 2010

- 2nd TPC Technology Conference on Performance Evaluation and Benchmarking

- Co-located with 36th International Conference on Very Large Data Bases

- 17 Sept 2010, Grand Copthorne Waterfront Hotel, Singapore

- More information: http://www.tpc.org/tpctc2010

VLDB SINGAPORE 2010

TPC Transaction Processing Performance Council

E-Energy 2010 M. Poess et al.