From Performance to Dependability Benchmarking:

A Mandatory Path

TPC

Marco Vieira

University of Coimbra – Portugal

mvieira@dei.uc.pt

Who am I?

- Marco Vieira
- Assistant Professor – University of Coimbra
- Adjunct Associate Teaching Professor – CMU
- Expert on experimental dependability evaluation and benchmarking
- Participated in many EU research projects
  - Lot work on dependability benchmarking

University of Coimbra

- Oldest University in Portugal
  - One of the oldest in the world
  - 800 years of history and tradition
- Internationally ranked as the best University in Portugal
- Software and systems Engineering Research Group
  - Long tradition on experimental dependability assessment

I'm here to challenge you!!!

- Is computer benchmarking only about performance?
  - NO!!!
- e.g. nowadays most systems need to guarantee high availability and reliability
  - It is mandatory to shift the focus from measuring performance to the measurement of both performance and dependability attributes
- Don't computers fail?
  - What is the impact of failures into the system?

Why do computers fail?

Too many reasons…

- Hardware problems

Too many reasons…

- Environment problems

Too many reasons…

- Bad configuration

Too many reasons…

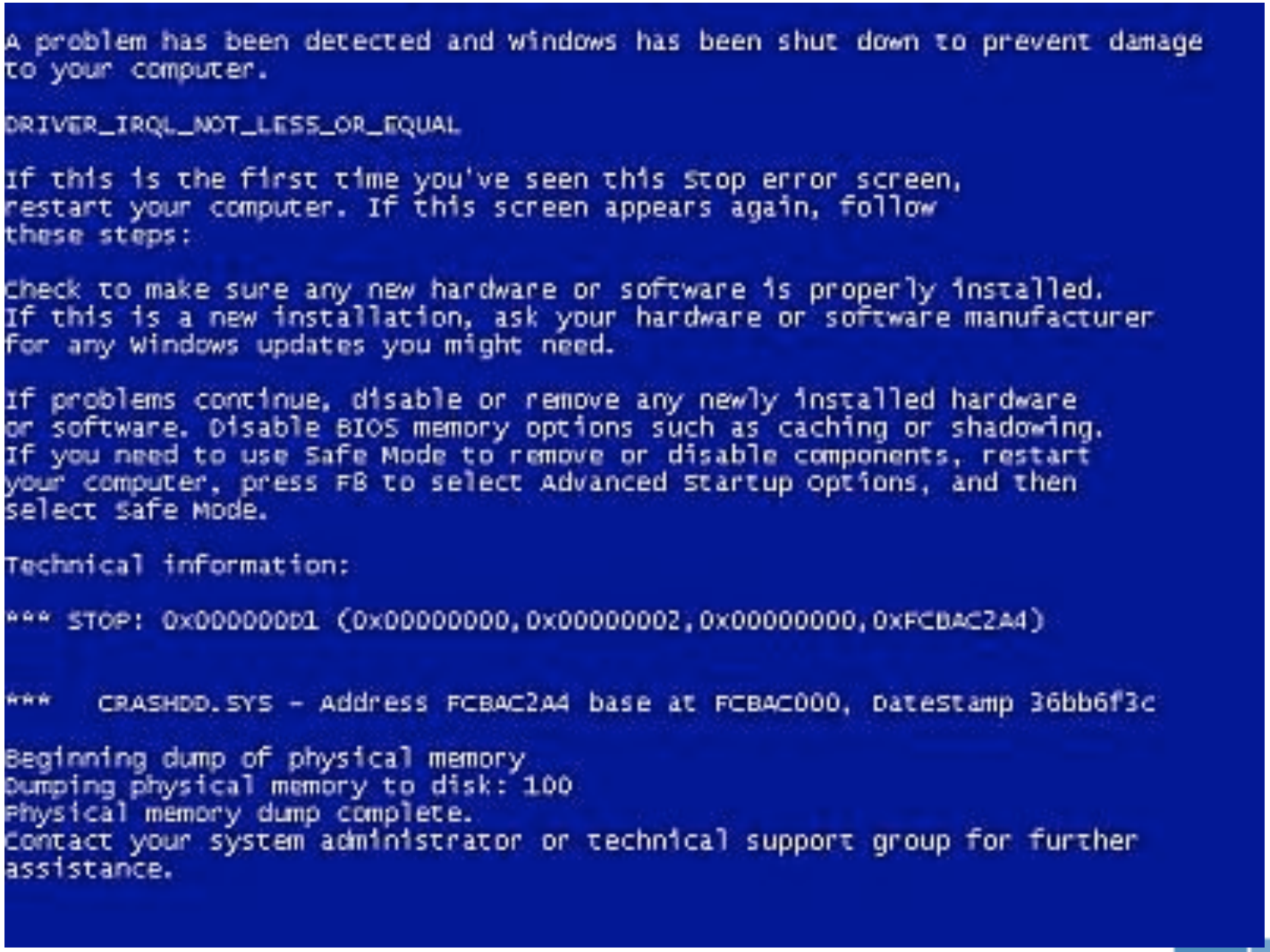
- Misuse

Too many reasons…

- Not proven design

But most of the times it is due to…

- Software problems



Now seriously…

- AT&T
  - 9 hrs crash (long distance calls)
- AOL
  - 6 million users affected
- AMAZON.COM
  - 12 hrs Crash
- …

http://www5.in.tum.de/~huckle/bugse.html

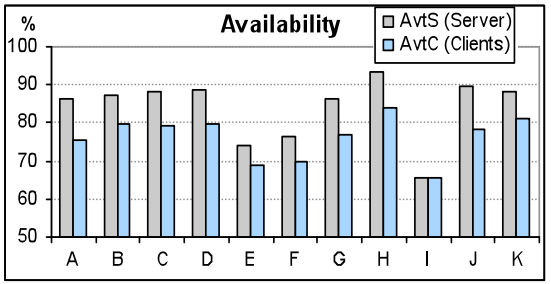
Are users asking for dependability?

- e.g., very common to find CFP asking for systems with a given level of availability
  - Data center for PT
- e.g., EU is putting huge money of dependability benchmarking research
  - Large scale, dynamic, complex, service-based systems

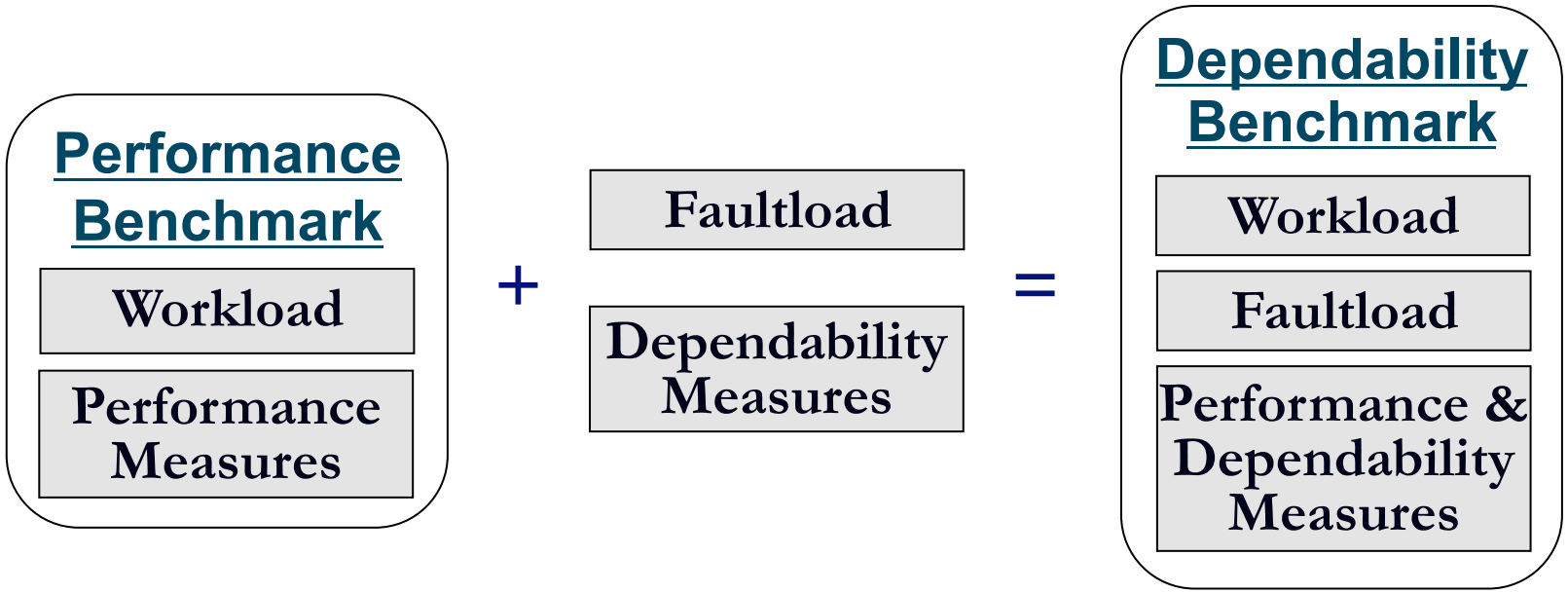
Dependability benchmarking

Procedures to measure both the dependability and performance of systems or components

- Compare systems or components from a dependability point-of-view:
  - Availability
  - Reliability
  - Safety
  - Confidentiality
  - Integrity
  - Maintainability



Components of a dependability benchmark



- Procedure and rules
- Experimental setup

Work on dependability benchmarking… (1)

- Special Interest Group on Dependability Benchmarking (SIGDeB)
  - Created by the International Federation for Information Processing (IFIP) Working Group 10.4 in 1999
  - Promote the research, practice, and adoption of benchmarks for computer systems dependability

Work on dependability benchmarking… (2)

- Assessing, Measuring and Benchmarking Resilience (AMBER)
  - Coordination Action supported by the European Commission in the 7th Framework Programme
  - Coordinating and advancing research in resilience measurement and benchmarking
  - Define what should be funded by the EU concerning research on resilience benchmarking
  - Jan 2008 – Dec 2009
  - Grant: € 1.05 M

Work on dependability benchmarking… (3)

- The DBench Project
  - Funded by the European Commission, under the Information Society Technologies Programme (IST), FP5
  - Devise benchmarks to evaluate and compare the dependability of COTS and COTS-based systems
    - General purpose operating systems
    - Real time kernels for onboard space systems
    - Engine control applications for automotive systems
    - On-line transaction processing systems
    - Web-servers
  - Jan 2002 – Mar 2004
  - € 1.5 M

Work on dependability benchmarking… (4)

- Berkeley University
  - Benchmarking the dependability of human-assisted recovery processes
- Sun Microsystems
  - Extremely active on dependability benchmarking!!!
  - High-level framework specifically dedicated to availability benchmarking of computer systems
  - Within this framework, two specific benchmarks have already been developed
    - System's robustness (degree of protection against outage events) in handling maintenance events
    - System recovery in a non-clustered standalone system

Work on dependability benchmarking… (5)

- IBM Autonomic Computing Initiative
  - Quantify a system's level of autonomic capability
    - Defined as the capacity of the system to react autonomously to problems and changes in the environment
  - The goal is to produce a suite of benchmarks covering the four autonomic capabilities
    - Self-configuration, self-healing, self-optimization, and self-protection
  - "Quantitative measurement of the autonomic capabilities of computing systems", IBM, United States Patent 7539904, May 2009.
    - Based on the work from DBench

Work on dependability benchmarking… (6)

- Intel Corporation
  - Focused on benchmarking semiconductor technology
  - E.g. impact of semiconductor technology scaling on neutron induced SER (soft error rate)

Approach to dependability benchmarking…

- Adding dependability metrics to performance benchmarks
- Define metric of interest
  - Focus on availability
- Investigate common failures
  - Based on real world scenarios
- Build on top of existing benchmarks
  - Did it using TPC-C
  - Can be applied to any benchmark
  - Similar to TPC-Energy model…

Benchmark example: DBench-OLTP

- Compare db-centric transactional systems
- Builds on top of TPC-C
  - (Customized) TPC-C workload
- Structured in Clauses:
  - Clause 1 – Preamble
  - Clause 2 – Benchmark Setup
  - Clause 3 – Benchmarking Procedure
  - Clause 4 – Measures
  - Clause 5 – Faultload
  - Clause 6 – Full Disclosure Report

Clause 4 – Measures (1)

- Baseline performance measures:
  - tpmC
    - Number of transactions executed per minute
  - $/tpmC
    - Price per transaction
- Performance measures in the presence of the faultload:
  - Tf
    - Number of transactions executed per minute in the presence of the faultload
  - $/Tf
    - Price-per-transaction in the presence of faults

Clause 4 – Measures (2)

- Dependability measures:
  - Ne
    - Number of data errors detected
  - AvtS
    - Availability from the SUB point-of-view
  - AvtC
    - Availability from the end-users (terminals) point-of-view
- Availability:
  - During the benchmark execution
  - System is available when it is able to provide the service defined by the TPC-C transactions

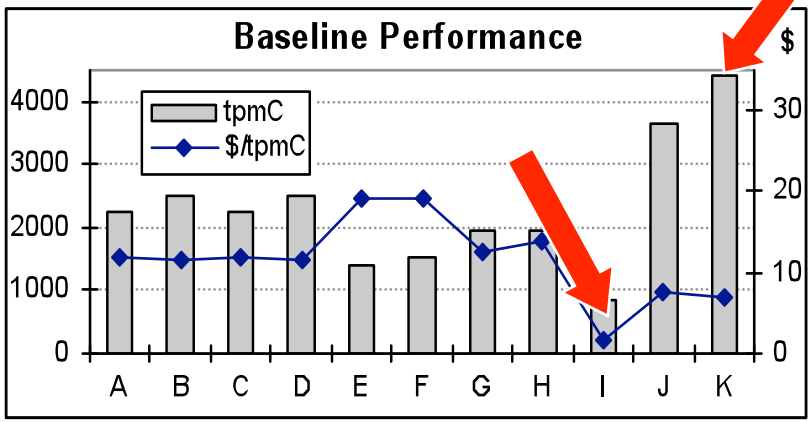
Clause 5 – Faultload

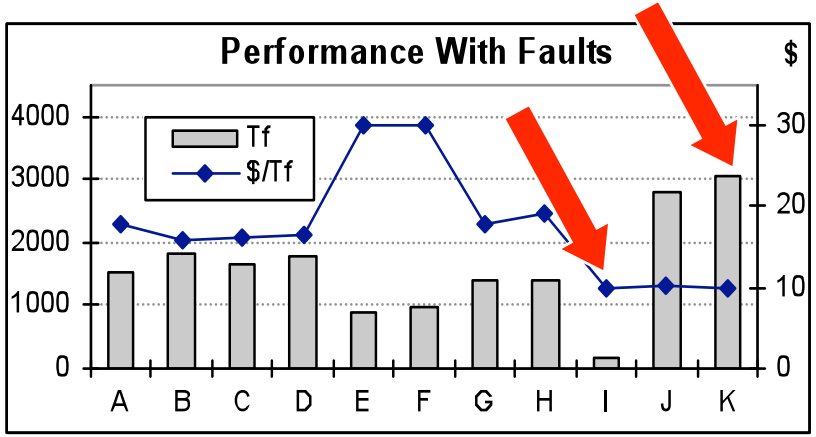
- Three different faultloads can be considered
- Operator faults
  - Emulate database administrator mistakes
- Software faults
  - Emulate software bugs in the operating system
- High-level Hardware failures
  - Emulates hardware component failures, such as:
    - Power failures
    - Disk failures
    - Corruption of storage media

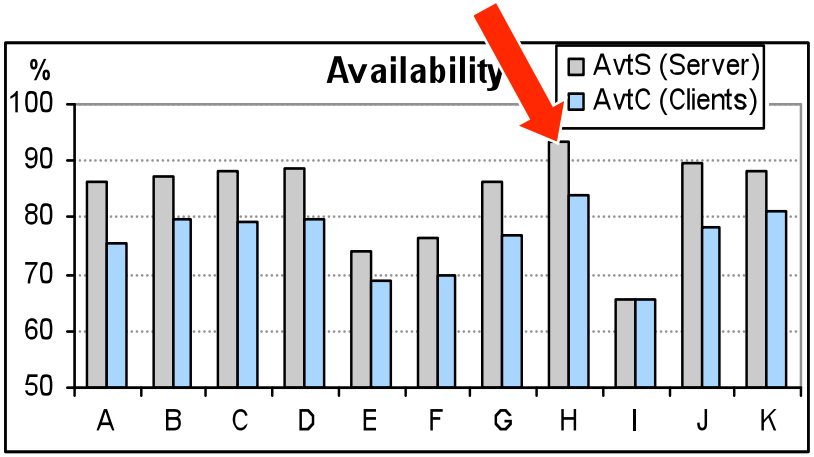
Benchmarking example

System	Hardware	Operating System	DBMS	DBMS Config.
A	• *Processor*: Intel Pentium III 800 MHz • *Memory:* 256MB • *Hard Disks:* Four 20GB/7200 rpm • *Network:* Fast Ethernet	Win2k Prof . SP 3	DB-1	Conf. A
B		Win2k Prof . SP 3	DB-2	Conf. A
C		WinXp Prof. SP 1	DB-1	Conf. A
D		WinXp Prof. SP 1	DB-2	Conf. A
E		Win2k Prof . SP 3	DB-1	Conf. B
F		Win2k Prof . SP 3	DB-2	Conf. B
G		SuSE Linux 7.3	DB-1	Conf. A
H		SuSE Linux 7.3	DB-2	Conf. A
I	• *Processor:* Intel Pentium IV 2 GHz • *Memory:* 512MB • *Hard Disks*: Four 20GB / 7200 rpm • *Network:* Fast Ethernet	Win2k Prof . SP 3	DB-1	Conf. A
J		Win2k Prof . SP 3	DB-2	Conf. A

DBench-OLTP Results summary







- Performance
- Availability
- Price

Vieira, M., Madeira, H.: A Dependability Benchmark for OLTP Application Environments, 29th Intl Conference on Very Large Data Bases, VLDB 2003, 2003

Conclusions

- Several works on dependability benchmarking have been conducted in the past
  - At both academia and industry
- No work has yet achieved the status of a real benchmark
  - Endorsed by a standardization body or corporation

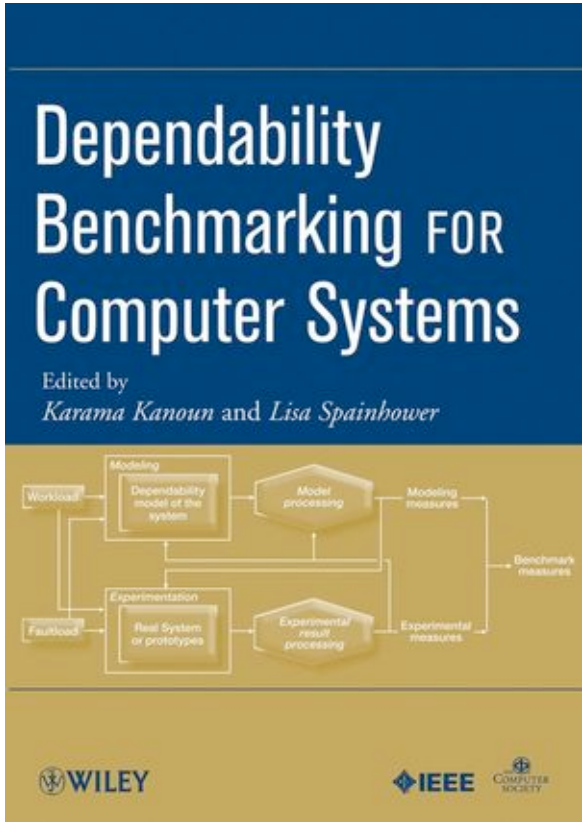
Dependability benchmarking seams to be a mandatory path!!!

- Isn't it time for TPC to start paying attention to it? ☺

The future…

- Investigate:
  - Including (optional) dependability metrics in performance benchmarks
  - Application across different domains/benchmarks
  - TPC-Energy model?
- Continue working on this topic
  - More detailed proposal in April?
- Join TPC as a non-profit organization

Bibliography



Dependability Benchmarking for Computer Systems
Karama Kanoun, Lisa Spainhower (Eds)
ISBN: 978-0-470-23055-8
Wiley-IEEE Computer Society Press
August 2008

Questions & Comments

Marco Vieira

University of Coimbra – Portugal

mvieira@dei.uc.pt