Computer Systems Performance Analysis and Benchmarking (37-235)

Analytic Modelling

Simulation

Measurements / Benchmarking

Lecture by:
Prof. Thomas Stricker

Assignments/Projects:
Christian Kurmann

Textbook:
Raj Jain, “The Art of Computer Systems Performance Analysis”, 1991 Wiley & Sons, New York

Topic of Today:

- Introduction
- Common mistakes ...
  ... and how to avoid them

Topics:

- **Evaluation technique, performance metrics, workloads.**
- **Measurements: How to do the properly.**
- **Statistical technics.**
- **Design of experiments (measured, simulated).**
- **Simulation.**
- **More on workloads.**
  - Survey of benchmarks.

Example (Techniques, Metrics, Workloads)

What performance metrics should be used to compare the performance of the following systems?

- **Two disk drives**
- **Two transaction processing systems**
- **Two packet retransmission algorithms**

Example: (Performance Measurements)

Which type of monitor (software and hardware) would be more suitable for measuring each of the following quantities?

- **Number of instructions executed by a processor**
- **Degree of multiprogramming on a timesharing system**
- **Response time of packets on a network**

Example: (Statistical Techniques)

- **The number of packets lost on two links was measured for four file sizes as shown below:**

File Size:	Link A	Link B
1000	5	10
1200	7	3
1300	3	0
50	0	1

Which link is better?

Example: (Design of Experiments)

The performance of a system depends on the following three factors:

- **Garbage collection technique used: (concurrent, stop and copy, none)**
- **Type of workload: (editing, search (AI), scientific computing)**
- **Type of CPU: (DEC Alpha, IBM Power 2, Intel P6)**

How many experiments are needed? How does one quantify the performance impact of each factor?

Example (Queueing Network Models)

The average. response time of a data-base system is 3 seconds. During a 1 min. observation interval, the idle time on the system was 1s. Determine the following:

- **System utilization**
- **Average service time per query**
- **Number of queries completed during the interval**
- **Average number of jobs in the system**
- **P[# of jobs in the system]>10**
- **90 percentile, response time or waiting time**

Why is Performance Evaluation an “Art” and not a “Science”?

Example:

Throughput table [Transactions per second]

System:	Workload 1	Workload 2
A	20	10
B	10	20

What is a fair comparison?

Absolute:

System:	Workload 1	Workload 2	Average
A	20	10	15
B	10	20	15

Relative to B:

System:	Workload 1	Workload 2	Average
A	2x	0.5x	1.25x
B	1x	1x	1x

Relative to A:

System:	Workload 1	Workload 2	Average
A	1x	1x	1x
B	2x	0.5x	1.25x

How to Learn System Performance Analysis?

- **Learn the mathematical basics (done best in a lecture)**
- **Read articles on real studies critically (see the first assignment in Uebung)**
- **Do a project... (see suggested list)**

Common mistakes... ...and how to avoid them

- **No goals**

  There is no such thing than a general purpose model. Models are developed best with a clear goal in mind.

  Example: model memory system performance with respect to communication operations in a parallel system.

- **Biased goals**

  Show that OUR system is better than THEIRS.

- **Unsystematic approach**

  Arbitrary selection of the parameter or thrive for cute theories rather than answers.

- **Analysis without understanding the problem**

  A problem well stated is half solved. 40% is defining the problem, 60% is looking into alternatives, getting results, interpret them and conclude. Customers want an answer not a model per se.

- **Incorrect performance metrics**

  Compare Mips for a CISC CPU vs. Mips for RISC CPU
  People go often for the easy metrics rather than the e.g. MHz clock frequency rather than the relevant ones.

- **Unrepresentative workload:**

  Workloads have a high impact on the results. It is very tempting to construct a favorable workload for an evaluation.

- **Wrong evaluation technique:**
  - Measurement
  - Simulation
  - Analytical modeling

  People tend to choose what they know best, rather than what works best for the problem. The three different techniques have different strengths and limitations.
- **Overlooking important parameters**

  List the characteristics of the workload and make sure you don't forget any characteristic that might be important. e.g. quantum sizes, working sets.
- **Ignoring significant factors**

  Factors are parameters varied by the user. It is important to understand what is fixed and what is random, sometime even how the randomness works.

- **Inappropriate design**

  Design is a question of coverage within the parameter space. Several techniques are given. Speciality of the author of our textbook.
- **Inappropriate level of detail**

  Shallow high level models allow many alternatives. Narrow and detailed model restrict investigation to a few alternative but give much insight about them.
- **No analysis**

  Study ends with a flood of data but lack of explanation.
- **Erroneous analysis**

  There are a few classic mistakes like average of ratios etc... Don't get caught at a conference with such a bug!

- **No sensitivity analysis**

  Computer scientists put too much emphasis on the result as a fact rather than on the evidence an how strong it is.
- **Improper treatment of outliers**

  Deciding which outliers are inherent in the system and which ones by are just unrelated bad luck is very difficult. One needs to understand the system to model it.
- **Assuming no changes over time**

  The future might be different, a good analysis can deal with change.
- **Ignoring variability**

  Focus on the mean, max or min, insufficient effort to get variability under control (statistically).

- **Too complex analysis**

  A simple analysis that leads to the same conclusion, with the same certainty, is always better than a complex one.
- **Improper presentation of results**

  One can go to a top systems conference like ASPLOS and still find bar charts about branch prediction with a range from 88% to 93%. Books have been written on that. e.g. "Wie lügt man mit Statistik".
- **Ignoring social aspects:**

  Modeling and analysis are substantive technical skills. Writing and speaking are social skills. Hint: combine them!
- **Omitting assumption and limitations**

  Those are often omitted, due to page limitations or because they are inconvenient to explain. This leads other user to apply the results out of context.

Textbook:

Raj Jain, “The Art of Computer Systems Performance Analysis”, 1991 Wiley & Sons, New York, ISBN 0471503363

Sources:

- Local:
  www.books.ch, Orell Fuessli, Fr. 185.--, in der Regel innert 4 Tagen versandt.
  www.freihofer.ch, Freihofer, Fr. 185.--, innert 2-4 Wochen lieferbar.
- Internet:
  www.books.com, US$ 84.95+$30+n*$6
  www.clbooks.com, US$ $84.95+ $19+n*$6
- Internet:
  www.amazon.de, Fr 143.65, 1 Woche