COMP 190-088: Systems Performance Analysis

Performance Metrics II

Jasleen Kaur

Department of Computer Science
The University of North Carolina at Chapel Hill

Spring 2005

Performance Metrics: Outline

- Common Performance Metrics
- Measuring Time
- **Metric Selection**
- Ratio Games

Common Mistakes In Metric Selection

- Arbitrary selection
  - Without considering goal of analysis
  - Leads to inaccurate conclusions
- Incorrect selection
  - Without considering system boundary and services provided
    - Eg, CPU comparison in MIPS
  - Selecting easy-to-compute, but irrelevant metrics
- Ratio games
  - Using ratios to define metrics to one's advantage

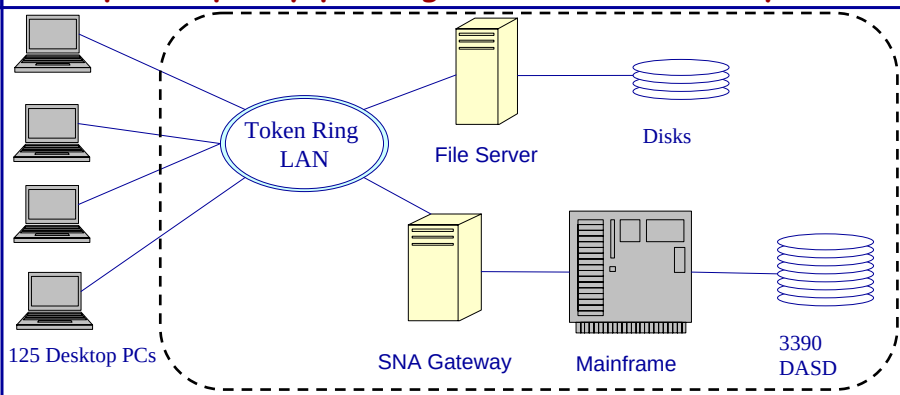
How to Select a Set of Metrics?

- A good set of metrics has:
  - Completeness:
    - All possible outcomes should be reflected in the set
  - Non-redundancy
    - Helps reduce the number of metrics
  - Low variability
    - Helps reduce the number of experiments needed to be confident of the value observed

Examples

- Single-CPU modeling:
  - Goal: to predict the achieved throughput of a processor in a system with compute-intensive tasks
  - Metrics:
    - MIPS
    - Number of tasks completed per second
    - Utilization / Idle time
- Multi-processor modeling:
  - Goal: to predict the achieved throughput of a multi-processor system with compute-intensive tasks
  - Metrics?

Example: Capacity planning of a client-server system



- Goal: which resources need to be upgraded to support an increase in number of clients from 125 to 200?
- Metrics:
  - Average Client Response time
  - Transactions per second
  - Utilization (network, file server, database)

Performance Metrics: Outline

- Common Performance Metrics
- Measuring Time
- Metric Selection
- **Ratio Games**

Ratio Games

- Fact:
  - Two performance ratios with different denominators (bases) are not comparable
- Ratio game:
  - The technique of using ratios with incomparable bases and combining them to one's advantage in a comparison test
- Many ways in which ratios can be used to win a comparison
  - Sometimes not obvious that a ratio has been taken

Analysts should be aware of ratio games others may play

Choosing an Appropriate Base System

- Simplest way of playing a ratio game in comparing two systems:
  - Present the performance of the systems on a variety of workloads
  - Take a ratio of performance of each workload, and then
  - Use the average ratio to show that one's own system is better.

Raw Measurements			With 6502 as a Base		With 8080 as a Base	
Benchmark	System		System		System	
	6502	8080	6502	8080	6502	8080
Block	41.16	51.50	1.00	1.25	0.80	1.00
Sieve	63.17	48.08	1.00	0.76	1.31	1.00
Sum	104.3	99.58	2.00	2.01	2.11	2.00
Average	52.17	49.79	1.00	1.01	1.06	1.00

- Three different conclusions:
  - Ratio of totals: 6502 is worse (4.7% more time).
  - With 6502 as Base: 6502 is better (1% less time).
  - With 8080 as Base: 6502 is worse (6% more time).

Using an Appropriate Ratio Metric

- Choose a performance metric which is a ratio of two different metrics

System	Throughput	Response	Power
A	10	2	5
B	4	1	4

- Example:
  - Network A has higher throughput, but also higher response time
    - Throughput is an HB metric; Response time is an LB metric
  - A suggests that use Power to compare
    - Power = Throughput/Response
  - Conclusion: Network A is better

Using Relative Performance Enhancement

- Use different machines/hardware to show that **relative** enhancement with one system is better than the other
  - Useful is performance metric is already specified

Alternative	Without	With	Ratio
A on X	2	4	2.00
B on Y	3	5	1.66

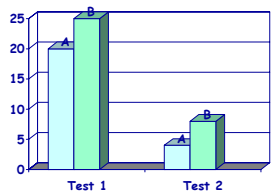
- Example: Comparison of two floating-point accelerators A and B
  - Measure MFLOPS on two different machines X and Y
    - Exploit the fact that A does not yet work on Y
  - Use the technique of incomparable bases
  - Conclusion: A gives higher improvement than B

Ratio Games With Percentages

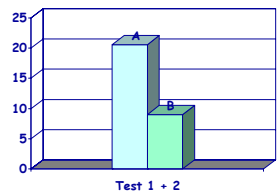
- Percentages are also ratios!
  - They allow playing ratio games in ways that do not look like ratios
- Example: Two experiments conducted repeatedly on two systems
  - Each experiment either passes or fails its performance goal

Test	System A			System B		
	Total	Pass	%Pass	Total	Pass	%Pass
1	300	60	20	32	8	25
2	50	2	4	500	40	8
Total	350	62	**20.6**	532	48	**9**

Both use incomparable bases!



Percent of tests passed



Percent of total tests passed

Playing with Percentages

- Percentages are also used to play psychological games
  - Large numbers attract attention
  - 1000% improvement vs. 11-time improvement
- Disguising small sample sizes
  - 83.3% of universities in town use our system X
  - 5 out of 6 universities use system X
- Base in percentage should be initial/earlier value
  - Some marketers ignore this
  - Memory prices have gone down by 400%
    - Are they now paying you to buy memory?

Wining Ratio Games

- How to win a ratio game?
- Under what conditions can conclusions be reversed (a different winner) by changing the base in a ratio game?

Winning ratio game: Rule 1

- If one system is better on all benchmarks, contradicting conclusions can't be drawn by any ratio game technique

	Raw Measurements		With A as a Base		With B as a Base	
	System		System		System	
Benchmark	A	B	A	B	A	B
I	0.5	1.0	1.0	2.0	0.5	1.0
J	1.0	1.5	1.0	1.5	0.67	1.0
Average	0.75	1.25	1.0	1.75	0.58	1.00

- System A is better in all 3 cases

Winning ratio game: Rule 2

- Even if one system is better on all benchmarks, a better relative performance can be shown by selecting appropriate base

	Raw Measurements		With A as a Base		With B as a Base	
	System		System		System	
Benchmark	A	B	A	B	A	B
I	0.5	1.0	1.0	2.0	0.5	1.0
J	1.0	1.5	1.0	1.5	0.67	1.0
Average	0.75	1.25	1.0	1.75	0.58	1.00
% Improvement	**40%**		**43%**		**42%**	

- System A designers may prefer using A as Base

More strategies

- If a system is better on some benchmarks and worse on others, contradicting conclusions can be drawn in some cases
  - Try all possible bases
  - Following rules may help select the base
- If the metric is a lower-is-better (LB) one, it is better to use your system as the base
  - LB metrics: Response time, Reaction time, Idle time, Cost/Performance ratio
- If metric is HB, use opponent as base.
  - HB metrics: Throughput, Efficiency, MIPS, MFLOPS
- Elongate benchmarks on which your system does well
  - If ratio-of-totals technique is used, you'll win
  - Shortening unfavorable benchmarks has same effect

Basis of Most Ratio Games

Approach of taking mean of ratios is wrong

- Ignores the fact that performance is affected by several factors and experimental errors
  - Example: "execution time of a workload on a system"
    - Factors: workload, system
- Need to isolate the impact of, and interaction between factors
  - Use Measurements, Modeling, Simulations
  - Requires knowledge of probabilistic, statistical, and experimental design concepts

Course Outline

- Selection of metrics
- **Performance Evaluation Methodologies**
- Workload selection
- Measurements tools
- Analysis of measured data
- Visualization of results
- System Modeling
- Simulations
- Case studies
- Distributed monitoring infrastructures
- PA in the Research and Industrial communities