Common Mistakes in Performance Evaluation

The Art of Computer Systems Performance Analysis
By Raj Jain

Adel Nadjaran Toosi

Wise men learn by other men's mistake,
fools by their own.

H. G. Wells

No Goals

- Many performance efforts are started without any clear goal.
- Metrics, workloads, and methodology all depend upon the goal.

- Setting goal is not trivial exercise.
- One the problem is clear and the goal have been written down, finding the solution is often easier.

Biased Goal

- Example:
  - To show that OUR system is better than THEIRS
- Mistake: Finding the metrics and workload such that OUR system turns out better rather than finding the right metrics and workloads.

Unsystematic Approach

- Selecting system parameters , factors, metrics, and workload arbitrarily.

Analysis without understanding the problem

- In experienced analysts feel that nothing really achieved until the model has been constructed and some numerical results obtained.
- With experiences, they learn that a large share goes in to defining the problem.(40%)
  - The problem well stated is half solved.

Incorrect Performance Metrics

- Choice of correct performance metrics depends upon the service provided subsystem being modeled.
- Comparing Two CPUS based on the throughput (MIPS)
  - CISCS
  - RISKS

Unrepresentative workload

- Workload used to compare two systems should be representative of the actual usage of the systems in the field.
- Example
  - If packet in the network are generally mixture of long and shot, workload should consist of short and long packet sizes.

Wrong Evaluation Technique

- Three Evaluation Technique:
  - Measurement
  - Simulation
  - Analytical Modeling
- Example:
  - Those who are proficient in queuing theory tends to change every performance problem to queuing theory, even if the system is too complex and easily available for measurement

Overlooking Important Parameters

- Good idea: Make a complete list of system parameters, e.g. Number of users, request size, request arrival pattern.
- Overlooking one or more important parameter may render the results useless.

Ignoring significant factors

- Parameters that are varies in the study are called **Factors**.
- Request size
- Not all factors have equal effect on the performance.
- You should find those which have significant impact on the performance.

Ignoring significant factors(Cont.)

- Factors that are under the control of the end user (or decision maker) are more important.
- It is important to understand the randomness of various system and workload parameters.
- Analyst may know the distribution for page references, but have no idea of the distribution of disk references.
  - Mistake: using page reference distribution and ignore the disk references even though the disk is may be the bottle neck of the system.
- **Sensitivity analysis** is a good practice

Sensitivity analysis

- Let us give an example: in any budgeting process there are always variables that are uncertain.
- Future tax rates, interest rates, inflation rates, headcount, operating expenses and other variables may not be known with great precision.
- Sensitivity analysis answers the question, "if these variables deviate from expectations, what will the effect be (on the business, model, system, or whatever is being analyzed)

Inappropriate Experimental Design

- Experimental design relates to the number of measurement or simulation experiments to be conducted.
- Proper selection => more information
- Improper selection => waste of the analyst's time and resources
- Each factor changed one by one, Simple Design, => wrong conclusion
- Full factorial design is a solution

Full factorial design

- A full factorial experiment is an experiment whose design consists of two or more factors, each with discrete possible values or "levels", and whose experimental units take on all possible combinations of these levels across all such factors.

Inappropriate level of detail

- Avoid formulation that are too narrow or too broad.
- For slight variations of common approach, a detailed model
- For comparing alternatives that are very different high-level models

No analysis

- Performance analyst good in measurement technique but lack of data analysis expertise.
- Enormous amount of collected data, but do not know how to analyze or interpret it.

Erroneous Analysis

- Taking average of ratios and too short simulations

No Sensitivity Analysis

- Putting too much emphasis on the results of the analysis, presenting it as fact rather than evidence.
- However, results may be sensitive to workload and system parameters.

Ignoring Errors in Input

- Sometimes, parameter of interest cannot be measured, instead another variable can be measured to estimated the parameter.
- Would this cause an insignificant change in the result?

Improper Treatment of Outliers

- Too high, too low values….
- If outliers is not caused by the a real system phenomenon, it should be ignored.
- If the outlier is a possible occurrence in a real system, it should be appropriately included in the model.
- Careful understanding the system being modeled is required.

Assuming No change in the Future

- It is often assumed that future will be the same as the past.
- Model based on the workload and performance observed in the past is used to predict performance in future.
- Analyst and decision maker should discuss this assumption and limit the amount of time into the future that predication are made.

Ignoring Variability

- Common to analyze only the mean performance
- Since determining the variability is often difficult, if it is not impossible.
- If the variability is high the mean alone is misleading.

Too complex Analysis

- Given that two analyses leading to the same conclusion, simpler and easier to explain is preferable.
- It is better to start with simple model, and introduce the complications.
- Models published in the literature, are generally complex.
  - Trivial models even when they are illuminating are not generally accepted for publication.
  - The ability to develop and solve complex model is valued more in academic circles.
  - However, in industry, the decision maker are rarely interested in complex models.
  - Frustrating for new graduates well trained in complex modeling

Improper Presentation of Results

- The eventual aim of every performance study is to help in decision making.
- The right metric to measure the performance of an analyst is not the number of analysis performed but the number of analyses that help the decision makers.
- This requires the proper use of words, pictures, and graphs to explain the result and analysis.

Ignoring Social Aspects

- Social and substantive
- Writing and speaking are social skills
- Modeling and data analysis are substantive
- Beginning analysts often fail to understand social skills are often more important than substantive skills.
- Weak presentation leads to rejection of the high-quality analyses

Omitting Assumptions and limitations

- Assumption and limitations of the analysis are often omitted in the final report.
- This may lead the user to apply the analysis to another context where assumption will not be valid anymore.

1. Is the system correctly defined and the goals clearly stated?
2. Are the goals stated in an unbiased manner
3. Have all the steps of the analysis followed systematically?
4. Is the problem clearly understood before analyzing it?
5. Are the performance metrics relevant for this problem?

6. Is the workload correct for this problem?

7. Is the evaluation technique appropriate?

8. Is the list of parameters that affect performance complete?

9. Have all parameters that affect performance been chosen as factors to be varied?

10. Is the experimental design efficient in terms of time and results?

11. Is the level of detail proper?
12. Is the measured data presented with analysis and interpretation?
13. Is the analysis statistically correct?
14. Has the sensitivity analysis been done?
15. Would errors in the input cause an insignificant change in the results?
16. Have the outliers in the input or output been treated properly?
17. Have the future changes in the system and workload been modeled?
18. Has the variance of input been taken into account?

19. Has the variance of the result been analyzed?
20. Is the analysis easy to explain?
21. Is the presentation style suitable for its audience?
22. Have the results been presented graphically as much as possible?
23. Are the assumptions and limitations of the analysis clearly documented?

Thank You