Diseño y Evaluación de Configuraciones

I Panorámica sobre análisis y diseño de la prestaciones temporales de los sistemas informáticos.

(System Performance Analysis and Design)

J.M. Drake

Notas:

Algunos de los comentarios de las páginas se han extraido de los textos:

Raj Jain: “The art of computer Systems Performance Analysis”
Connie U. Smith: “ Performance Enegineering of software Systems”

Software performance Engineering

"Software Performance Engineering (SPE) represents the entire collection of software engineering activities and related analyses used throughout the software development cycle, which are directed to meeting performance requirements."

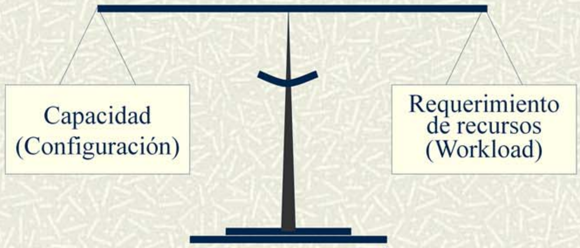
- Por performance hacemos referencia a características como productividad (throughput), respuesta temporal (timeliness), fiabilidad, consumo de potencia, necesidad de memoria, etc. de los sistemas informáticos.
- Son características no funcionales de naturaleza transversal que afecta a cada aspecto del sistema: Arquitectura, código, sistema operativo, middleware, servicios de comunicaciones, hardware,etc..
- Es de gran relevancia en el proceso de desarrollo de los sistemas por ser causas de retrasos, sobrecostos, fallos y fracasos.

Notas:

Computer system users, administrators, and designers are all interested in performance evaluation since their goal is to obtain or provide the highest performance at the lowest cost. This goal has resulted in continuing evolution of higher performance and lower cost systems leading to today's proliferation of workstations and personal computers, many of which have better performance than earlier supercomputers. As the field of computer design matures, the computer industry is becoming more competitive, and it is more important than ever to ensure that the alternative selected provides the best cost-performance trade-off.

Performance evaluation is required at every stage in the life cycle of a computer system, including its design, manufacturing, sales/purchase, use, upgrade, and so on. A performance evaluation is required when a computer system designer wants to compare a number of alternative designs and find the best design. It is required when a system administrator wants to compare a number of systems and wants to decide which system is best for a given set of applications. Even if there are no alternatives, performance evaluation of the current system helps in determining how well it is performing and whether any improvements need to be made. Unfortunately, the types of applications of computers are so numerous that it is not possible to have a standard measure of performance, a standard measurement environment (application), or a standard technique for all cases. The first step in performance evaluation is to select the right measures of performance, the right measurement environments, and the right techniques.

Resulta del balance entre capacidad y requerimientos



- Las características de performance resulta de un balance entre los requerimientos de recursos que tiene la aplicación y la capacidad y disponibilidad de recursos que tiene el sistema.
- Los objetivos de performance se pueden conseguir bien incrementando la capacidad como reduciendo los requerimientos.

Notas:

The performance balance despicts a system that fail to meet performance objetives because resources requirements exceed computer capacity. With SPE, analystsdetect these problems early in development and use quantitative methods to suppport cost-benefit analysis of hardware solutions versus software requirements or design solutions, versus a combination of the two. Developers implements software solution before problems are manifested in code.

The performance balance: The resource requirements outweigh the computer configuration capacity, thus causing performance degradation. Solutions either increase capacity or decrease demand.

Is SPE necessary?. Isn’t hardware fast enough and cheap to resolve performance problems? Surprisingly , the use of state-of-art hardware and software technology dramatically increases the sik of performance failure. This seems counter.intuitive, one would expect increased performance , but the newness of the products combined with developers inesxperience with the new enviroment leads to problems.

Carga, métricas y objetivos

- Tres aspectos que deben definir en cualquier estudios de performance son:
  - La carga (workload): La carga caracteriza la cantidad y la naturaleza de de los requerimientos que se realizan sobre el sistema.
    - Intensidad de la carga: Nº requerimientos/s
    - Patrón de carga: Modificaciones de las intensidades de la carga.
    - Benchmark: Define modelos de cargas normalizados para comparar.
  - Magnitudes a medir (métricas): Son las magnitudes que se utilizan para cuantificar lo que queremos medir:
    - Throughput (Número de tareas por unidad de tiempo que realiza).
    - Power consumption (energía eléctrica por unidad de tiempo consumida)
    - Response time (Tiempo utilizado para realizar un servicio)
    - ....
  - Objetivos del análisis (goals):Los objetivos de un análisis de performance pueden ser muy diversos:
    - Comparación: Evaluar si una alternativa de diseño o de implementación es mejor que otra.
    - Dimensionado del sistema: Determinar los recursos o el tamaño que debe establecerse en el sistema para que satisfaga ciertos requisitos de comportamiento.
    - ...

Notas:

An important feature of computer or communication systems is that their performance depends dramatically on the *workload* (or simply *load*) they are subjected to. The load characterizes the quantity and the nature of requests submitted to the system. Consider for example the problem of quantifying the performance of a web server. We could characterize the load by a simple concept such as the number of requests per second. This is called the *intensity of the workload*. In general, the performance deteriorates when the intensity increases, but often the deterioration is sudden; this is due to the non-linearity of queuing systems

The performance of a system depends not only on the intensity of the workload, but also its nature; for example, on a web server, all requests are not equivalent: some web server softwares might perform well with *get* requests for frequently used objects, and less well with requests that require database access; for other web servers, things might be different. This is addressed by using standardized mixes of web server requests. They are generated by a *benchmark*, defined as a load generation process that intends to mimic a typical user behaviour.

A performance *metric* is a measurable quantity that precisely captures what we want to measure, it can take many forms. There is no general definition of a performance metric: it is systemdependent, and its definition requires understanding the system and its users well. We will often mention examples where the metric is throughput (number of tasks completed per time unit), power consumption (integral of the electrical energy consumed by the system, per time unit), or response time (time elapsed between a start and an end events). For each performance metric, we may be interested in average, 95-percentile, worst-case, etc,

The goal of a performance evaluation may either be a *comparison* of design alternatives, i.e. quantify the improvement brought by a design option, or *system dimensioning*, i.e. determine the size of all system components for a given planned utilization. Comparison of designs requires a well-defined load model; however, the exact value of its intensity does not have to be identified. In contrast, system dimensioning requires a detailed estimation of the load intensity. Like any prediction exercise, this is very hazardous. For any performance evaluation, it is important to know whether the results depend on a workload prediction or not.

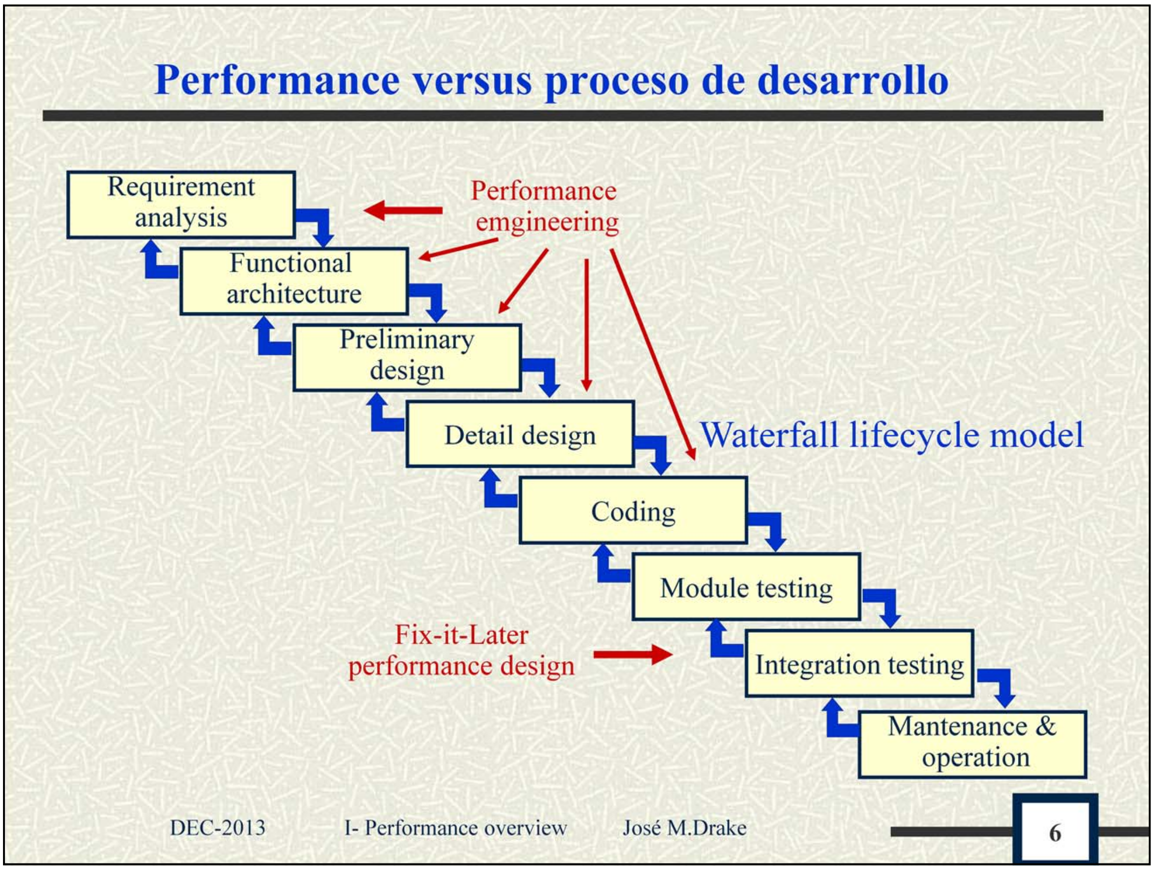
Estrategias de abordar el performance

- Fix-it-Later: El desarrollo se realiza sin considerar los objetivos de performance. En la fase final de integración se miden las características de performance y se configura el sistema.
  - Se da mayor importancia a las técnicas de medida y análisis.
- Performance engineering: En la fases de especificación de requerimiento se establecen la características de performance. y en todas las fases del desarrollo se condiciona las opciones de diseño a los requisitos.
  - Se da mayor importancia a las técnicas de modelado y diseño.

Notas:

Systems without critical performance equirements adopted the “fix-it-latter” methods. It advocates concentrating on software corretness, defers performance consideration to the integration-testing phase, and (if performance problems are detected rhen) corrects problems qith additional hardware, with systems and software tuning or both.

SPE adopts quantitative methods successfully used in engineering and other fields. It views performance as an integral part of software development, along with functionality, correctness, usability, maintainability and other quality factors.



Notas:

The most serious flaw in the fit-it-later approach is that developers detect performance problems late in the software livecycle. Performance problems become aparent in the system integration or maintenance stages. Performance problems that surfaces during integration can significantly delay software delivery.

Aspectos de análisis de performance (1)

1. Selección de la técnica de evaluación, métrica de medida de performance y workload.
   - Técnicas de evaluación: Medida experimental, simulación y predicción mediante modelos
   - Métrica representa el criterio usado para evaluar las características de performance del sistema: p.e. latencia, throughput, etc.
   - Workload representa los patrones con los que se requieren los servicios: p.e. instrucciones/s de la CPU, mensajes/s por la red, etc.

p.e.: ¿Que métrica se debe utilizar para comparar la performance de los siguientes sistemas?:

a) Dos gestores de transacciones
b) Dos discos de almacenamiento de datos
c) Dos conmutadores de paquetes

Notas:

Later, the terms “evaluation techniques.” “metrics,” and “workload” are explained in detail. Briefly, the techniques that may be used for performance evaluation are measurement, simulation, and analytical modeling. The term **metrics** refers to the criteria used to evaluate the performance of the system. For example, **response time**—the time to service a request—could be used as a metric to compare two timesharing systems. Similarly, two transaction processing systems may be compared on the basis of their throughputs, which may be specified in transactions per second (**TPS**).

The requests made by the users of the system are called **workloads.** For example, the CPU workload would consist of the instructions it is asked to execute. The workload of a database system would consist of queries and other requests it executes for users.

Aspectos de análisis de performance (2)

2. Medida correcta de las características de performance
   - Para medir las características de performance se necesitan dos herramientas:
     - El generador de carga del sistema
     - El monitor que mide los resultados que se requieren.

p.e.:¿Que tipo de monitor (software o hardware) es mas adecuado para medir las siguientes magnitudes

a) Número de instrucciones ejecutadas por una CPU.
b) Número de tareas en un sistema de trabajo compartido.
c) Tiempo de transmisión de un paquete por una red.

Notas:

To measure the performance of a computer system, you need at least two tools—a tool to load the system (**load generator**) and a tool to measure the results (**monitor**). There are several types of load generators and monitors. For example, to emulate

several users of a timesharing system, one would use a load generator called a **remote terminal emulator** (RTE).

Aspectos de análisis de performance (3)

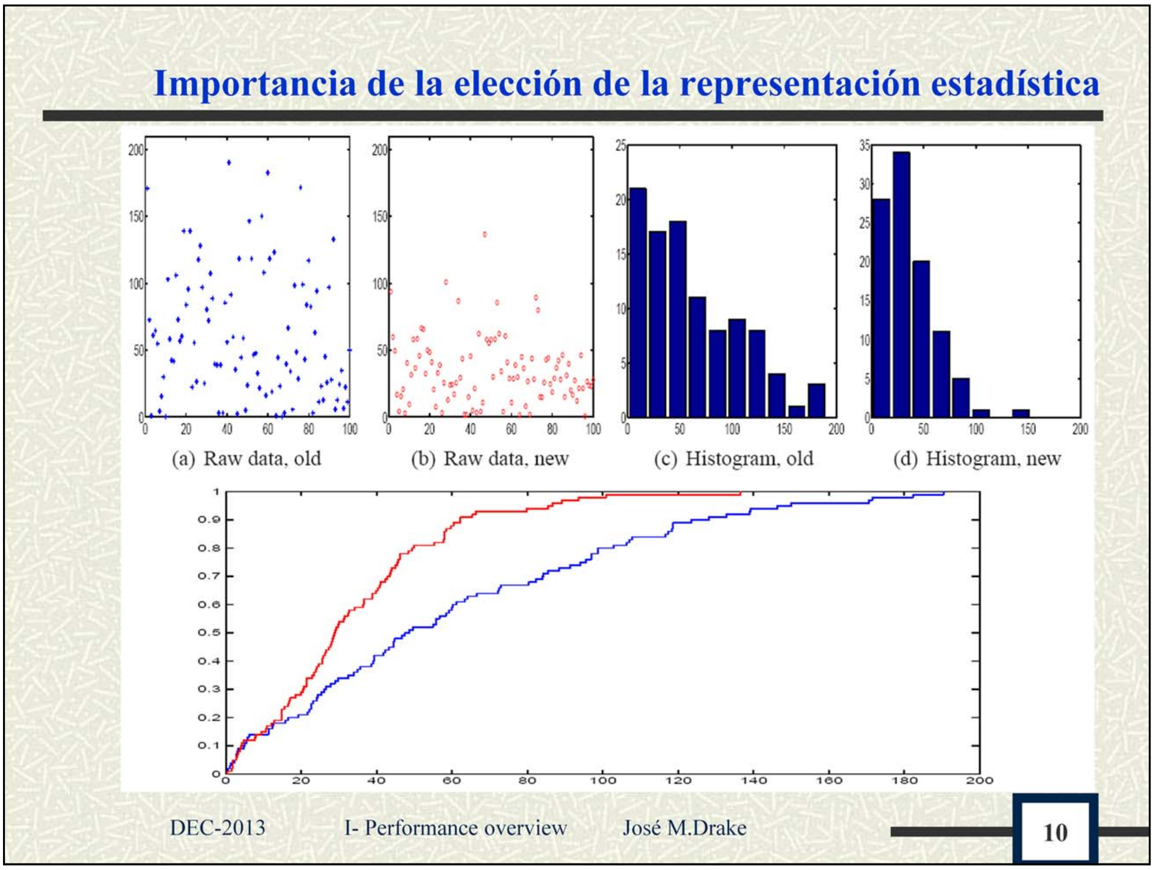
3. Comparación de las alternativas con la estadística correcta
   - La evaluación de la performance de un sistema consiste en la comparación de diferentes alternativas.
   - La comparación debe realizarse a magnitudes estadísticas apropiadas, (medias, variabilidad, tipo de distribución, etc..)

   p.e.:Si el número de paquetes que se pierde en una red es medida para cuatro longitudes de mensajes ¿Que red es mejor, la A o la B?

Msg size	Link A	Link B
1000	5	10
1200	7	3
1300	3	0
50	0	1

Notas:

Most performance evaluation problems basically consist of finding the best among a number of alternatives. If a measurement or a simulation is repeated several times, generally the results would be slightly different each time. Simply comparing the average result of a number of repeated trials does not lead to correct conclusions, particularly if the variability of the result is high. The statistical techniques used to compare several alternatives are discussed in this course.



Notas:

COMPARISON OF TWO OPTIONS. An operating system vendor claims that the new version of the database management code significantly improves the performance. We measured the execution times of a series of commonly used programs with both options. The data are displayed in teh Figure. The raw displays and histograms show that both options have the same range, but it seems (graphically) that the new system more often provides a smaller execution time. The box plots are more suggestive; they show that the average and the range are about half for the new system.

Upper figure: Measured execution times, in ms, for 100 transactions with the old and new code, with histograms.

Lower figure: Data of Example 2.1. Empirical distribution functions for the old code (right curve) and the new one (left curve). The new outperforms the old, the improvement is significant at the tail of the distribution.

Aspectos de análisis de performance (4)

4. Diseño dirigido de los experimentos de medidas.
   - Cuando hay varios factores que afectan las performance del sistema, conviene diseñar experimentos que muestren independientemente las influencias.

p.e.:Si el comportamiento de un sistema depende de tres factores:

a) La técnica de recolección de basura utilizada: G1, G2 o ninguna.
b) Tipo de workload: Edición, procesamiento, inteligencia artificial.
c) El tipo de CPU: C1, C2 o C3

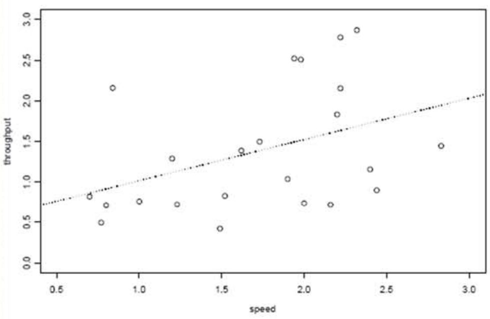
Que pruebas o experimentos conviene realizar.

Notas:

Given a number of factors that affect the system performance, it is useful to separate out the effects of individual factors. In Part IV on experimental design, techniques to organize experiments to obtain maximum information with a minimum number of experiments are presented.

Influencia de los factores ocultos

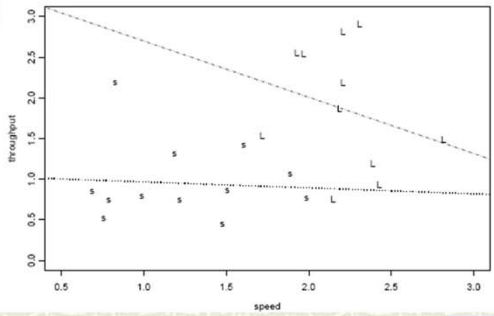
- Ignorar algunos factores ocultos puede invalidar los resultados del análisis de la performance de un sistema.



Left: plot of throughput (in Mb/s) versus speed (in m/s) for a mobile node.

Dercha: El mismo grafico, pero mostrando de forma diferenciada el tamaño del buffer del socket:;

s = small buffer, L = large buffer.



Notas:

TCP THROUGHPUT. Th figure, left, plots the throughput achieved by a mobile during a file transfer as a function of its velocity (speed). It suggests that throughput increases with mobility. The right plot shows the same data, but now the mobiles are separated in two groups: one group (‘s’) is using a small socket buffer (4K Bytes), whereas the second (‘L’) uses a larger socket buffer (16 K Bytes). The conclusion is now inverted: throughput decreases with mobility. The hidden factor influences the final result: all experiments with low speed are for small socket buffer sizes. The socket buffer size is a hidden factor.

Aspectos de análisis de performance (5)

5. Diseño correcto de los experimentos de simulación
   - Cuando se utiliza simulación, el modelo que se utiliza, el algoritmo de aleatoriedad, la simiente de inicio, el tiempo de simulación, las magnitudes que se observan, etc. debe ser adecuados al problema que se estudia:

p.e.: Si se quiere comparar dos algoritmos de actualización de memoria caché.

a) ¿Que tipo de modelo de simulación es adecuado?
b) ¿Cuanto tiempo debe durar la simulación?
c) ¿Se puede conseguir la misma resolución con una simulación mas corta?
d) ¿Como se puede decidir si el algoritmo de generación aleatoria es correcta?

Notas:

In designing a simulation model, one has to select a language for simulation, select seeds and algorithms for random-number generation, decide the length of simulation run, and analyze the simulation results.

Aspectos de análisis de performance (6)

6. Uso de modelos analíticos basados en modelos sencillos
   - Si el modelo es por ejemplo basado en colas, cual es distribución de servicio, longitud de cola, etc. que es apropiada al problema cuya performance se utiliza.

p.e.: Si los tiempos de respuesta medios de una base de datos son 3 segundos, y durante 1 minuto de observación se observa 10 segundos de cola vacía. Determinar:

a) Utilización del sistema.
b) Tiempo medio de servicio por cola.
c) Número de colas llenas durante el tiempo de observación.
d) Número medio de tareas en el sistema.
e) Percentil del 90% del tiempo de respuesta.
f) Percentil del 90% del tiempo de espera.

Notas:

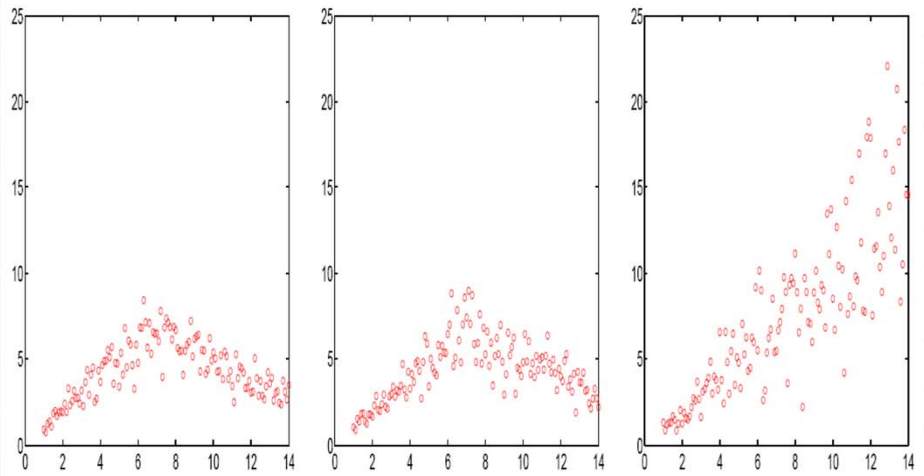
Queueing models are commonly used for analytical modeling of computer systems. Different types of queues and

networks of queues are discussed and and their use to answer commonly asked questions about system performance is described.

Aplicación del método científico

"Una conclusión no debe ser creída hasta que la conclusión haya sido específicamente demostrada."

- El olvido del método científico es una de las principales causas en el análisis de performance.



(a) Initially (b) With 2 additional base stations (c) With additional server memory

Notas:

JOE'S KIOSK. Joe's e-kiosk sells online videos to customers equipped with smartphones. The system is made of one servers and one 802.11 base station. Before deployment, performance evaluation tests are performed, as shown on Figure (a). We see that the throughput reaches a maximum at around 8 transactions per second.

Joe concludes that the bottleneck is the wireless LAN and decides to buy and install 2 more base stations. After installation, the results are on Figure (b). Surprisingly, there is no improvement.

The conclusion that the wireless LAN was the bottleneck was wrong. Joe scratches his head and decides to go more carefully about conclusions. Measurements are taken on the wireless LAN; the number of collisions is less than 0.1%, and the utilization is below 5%. This confirms that the wireless LAN is *not* a bottleneck. Joe makes the hypothesis that the bottleneck may be on the server side. After doubling the amount of real memory allocated to the server process, the results are as shown on Figure (c). This confirms that real memory was the limiting factor.

Errores típicos en el análisis de performance (1)

1. El análisis no hay objetivo: Cada modelo deben ser planteado para unos objetivos concretos. En función de ellos, debe elegirse la métrica, el workload, la metodología, etc.
2. Polarización a priori de los objetivos: El papel del analista debe ser imparcial, no debe incluir en el análisis ninguna preconcepción subjetiva
3. Estrategia no sistemática: Una estrategia en la que los parámetros del modelo, la métrica y el workload se elijan sin unos criterios bien establecidos, conduce a resultados no reproducibles ni comprensibles.
4. Abordar el análisis sin comprender el problema: Los modelos son el medio para obtener conclusiones y no el resultado final. Analistas que plantean modelos sin previamente haber definido los problema que pretende analizar no suele obtener resultados útiles.
5. Uso de una métrica inadecuada: La métrica debe elegirse en función de la naturaleza del servicio del sistema que es modelado. Un error común es elegir una métrica porque se sabe evaluar, con independencia de su relevancia respecto del problema que se analiza.

Notas:

No Goals: Goals are an important part of all endeavors. Any endeavor without goals is bound to fail.Performance evaluation projects are no exception. The need for a goal may sound obvious, but many performance efforts are started without any clear goals. A performance analyst, for example, is routinely hired along with the design team. The analyst may then start modeling or simulating the design. When asked about the goals, the analyst's answer typically is that the model will help answer a

design questions that may arise. Setting goals is not a trivial exercise. Since most performance problems are vague when first presented, understanding the problem sufficiently to write a set of goals is difficult. Once the problem is clear and the goals have been written down, finding the solution is often easier.

Biased Goals: Another common mistake is implicit or explicit bias in stating the goals. If, for example, the goal is "to show that OUR system is better than THEIRS," the problem becomes that of finding the metrics and workloads such that OUR system turns out better rather than that of finding the right metrics and workloads for comparing the two systems. One rule of professional etiquette for performance analysts is to be unbiased. *The performance analyst's role is like that of a jury*. Do not

have any preconceived biases and base all conclusions on the results of the analysis rather than on pure beliefs.

Unsystematic Approach: Often analysts adopt an unsystematic approach whereby they select system parameters, factors, metrics, and workloads arbitrarily. This leads to inaccurate conclusions. The systematic approach to solving a performance problem is to identify a complete set of goals, system parameters, factors, metrics, and workloads.

Analysis without Understanding the Problem: Inexperienced analysts feel that nothing really has been achieved until a model has been constructed and some numerical results have been obtained. With experience, they learn that a large share of the analysis effort goes in to defining a problem. This share often takes up to 40% of the total effort. This supports the old saying: *A problem well stated is half*

solved. Of the remaining 60%, a large share goes into designing alternatives, interpretation of the results, and presentation of conclusions. Development of the model itself is a small part of the problem-solving process.

Incorrect Performance Metrics: A metric refers to the criterion used to quantify the performance of the system. Examples of commonly used performance metrics are throughput and response time. The choice of correct performance metrics depends upon the services provided by the system or subsystem being modeled. By manipulating the metrics, it is possible to change the conclusions of a performance study. A common mistake in selecting metrics is that analysts often choose those that can be easily computed or measured rather than the ones that are relevant. Metrics that are difficult to compute are ignored.

Errores típicos en el análisis de performance (2)

6. Workload no representativo: La elección del workload tiene un impacto relevante sobre los resultados del análisis de un sistema y debe ser representativo del uso que se vaya a hacer del sistema.
7. Técnica de evaluación errónea: Hay tres técnicas de evaluación básicas: La basada en medida, la simulación y la basada en modelos analíticos. El uso de la técnica adecuada simplifica en gran medida el estudio del sistema.
8. Olvido de parámetros esenciales: Es una buena idea hacer una lista de las características del sistema y del workload que afectan el comportamiento del sistema, y en el análisis asignar a ellos valores que representen el caso que se estudia.
9. Olvido de factores significativos: Por factores entendemos aquellos que pueden ser controlados por el usuario final del sistema que se estudia, y que tienen un impacto sobre el comportamiento del sistema.
10. Diseño inadecuado de las pruebas: Aun cuando el número de parámetros sea pequeño su combinación factorial puede ser muy alta, e identificar el conjunto de casos que son relevante es importantes.

Notas:

Unrepresentative Workload: The workload used to compare two systems should be representative of the actual usage of the systems in the field. For example, if the packets in networks are generally a mixture of two sizes—short and long—the workload to compare two networks should consist of short and long packet sizes.The choice of the workload has a significant impact on the results of a performance study. The wrong workload will lead to inaccurate conclusions.

Wrong Evaluation Technique: There are three evaluation techniques: measurement, simulation, and analytical modeling. Analysts often have a preference for one evaluation technique that they use for every performance evaluation problem. For example, those proficient in queueing theory will tend to change every performance problem to a queueing problem even if the system is too complex and is easily available for measurement. Those proficient in programming will tend to solve every problem by simulation. This marriage to a single technique leads to a model that they can best solve rather than to a model that can best solve the problem. The problem with these transformations is that they may introduce phenomena into the model that were not present in the original system or they may leave out some important phenomena that were in the original system.

An analyst should have a basic knowledge of all three techniques. There are a number of factors that should be considered in selecting the right technique.

Overlooking Important Parameters: It is a good idea to make a complete list of system and workload characteristics that affect the performance of the system. These characteristics are called parameters. Workload parameters may include the number of users, request arrival patterns, priority, and so on. The analyst can choose a set of values for each of these parameters; the final outcome of the study depends heavily upon those choices. Overlooking one or more important parameters may render the results useless.

Ignoring Significant Factors: Parameters that are varied in the study are called **factors**. For example, among the workload parameters listed above only the number of users may be chosen as a factor; other parameters may be fixed at their typical values. Not all parameters have an equal effect on the performance. It is important to identify those parameters, which, if varied, will make a significant impact on the performance. Unless there is reason to believe otherwise, these parameters should be

used as factors in the performance study. Factors that are under the control of the end user (or decision maker) and can be easily changed by the end user should be given preference over those that cannot be changed. Do not waste time comparing alternatives that the end user cannot adopt either because they involve actions that are unacceptable to the decision makers or because they are beyond their sphere of influence.

Inappropriate Experimental Design: Experimental design relates to the number of measurement or simulation experiments to be conducted and the parameter values used in each experiment. Proper selection of these values can lead to more information from the same number of experiments. Improper selection can result in a waste of the analyst's time and resources. In naive experimental design, each factor is changed one by one. This "simple design" may lead to wrong conclusions if the parameters interact such that the effect of one parameter depends upon the values of other parameters.

Errores típicos en el análisis de performance (3)

11. **Nivel de detalle inadecuado:** Para comparar alternativas que son ligeras variantes de un modelo común puede requerirse un modelo específico, mientras que para comparar alternativas muy diferentes es mucho mas eficientes analizarlas con un modelo de mas alto nivel.
12. **No hay análisis:** Es común que el análisis sea realizado por expertos en medidas que obtienen una gran cantidad de información, pero no la somete a un análisis experto de datos del que se obtengan resultados
13. **No se estudian las sensibilidades de los análisis:** hay que comprobar la sensibilidad de los resultados del análisis respecto de la variación de los parámetros. El cual indica hasta que punto los resultados van a cambiar con pequeños cambios de los casos que se estudian.
14. **Errores en los datos de entrada:** Frecuentemente los datos que sirven de base al análisis poseen errores introducidos por la medida o por el método indirecto que se utiliza para obtenerlos. Estos errores pueden afectar de forma significativa los resultados del estudio.

Notas:

Inappropriate Level of Detail: The level of detail used in modeling a system has a significant impact on the problem formulation. Avoid formulations that are either too narrow or too broad. For comparing alternatives that are slight variations of a common approach, a detailed model that incorporates the variations may be more useful than a high-level model. On the other hand, for comparing alternatives that are very different, simple high-level models may allow several alternatives to be analyzed rapidly and inexpensively. A common mistake is to take the detailed approach when a high-level model will do and vice versa. It is clear that the goals of a study have a significant impact on what is modeled and how it is analyzed.

No Analysis: One of the common problems with measurement projects is that they are often run by performance analysts who are good in measurement techniques but lack data analysis expertise. They collect enormous amounts of data but do not know how to analyze or interpret it. The result is a set of magnetic tapes (or disks) full of data without any summary. At best, the analyst may produce a thick report full of raw data and graphs without any explanation of how one can use the results. Therefore, it is better to have a team of performance analysts with measurement as well as analysis background.

No Sensitivity Analysis: Often analysts put too much emphasis on the results of their analysis, presenting it as fact rather than evidence. The fact that the results may be sensitive to the workload and system parameters is often overlooked. Without a sensitivity analysis, one cannot be sure if the conclusions would change if the analysis was done in a slightly different setting. Also, without a sensitivity analysis, it is difficult to access the relative importance of various parameters.

Ignoring Errors in Input: Often the parameters of interest cannot be measured. Instead, another variable that can be measured is used to estimate the parameter. Such situations introduce additional uncertainties in the input data. The analyst needs to adjust the level of confidence on the model output obtained from such data.

Errores típicos en el análisis de performance (5)

15. Se supone que el sistema es invariable: No se considera que el workload puede cambiar en el tiempo, y resultados medidos en un sistema en un instante dado no es válido para casos futuros.
16. Se ignora la dispersión de los datos: Es mas fácil hacer cálculos basados en los valores medios, sin embargo la variabilidad de los resultados a veces es la causa de las características del sistema y tiene que ser considerada.
17. Análisis demasiado complejo: Es mejor empezar con análisis simples que describan las características básicas de un sistema, y posteriormente se refina y hace mas complejo en aquellos puntos que el primero no sea eficaz.
18. Presentación inadecuada de los datos: El objetivo de los análisis de performance son la base de las decisiones que deben ser realizadas por lo usuarios finales. La presentación de los resultados debe estar orientada a ellos. No incluir las suposiciones y limitaciones establecidas en el análisis pueden conducir a falasa interpretaciones de los resultados

Notas:

Assuming No Change in the Future: It is often assumed that the future will be the same as the past. A model based on the workload and performance observed in the past is used to predict performance in the future. The future workload and system behavior is assumed to be the same as that already measured. The analyst and the decision makers should discuss this assumption and limit the amount of time into the future that predictions are made.

Ignoring Variability: It is common to analyze only the mean performance since determining variability is often difficult, if not impossible. If the variability is high, the mean alone may be misleading to the decision makers. For example, decisions based on the daily averages of computer demands may not be useful if the load demand has large hourly peaks, which adversely impact user performance.

Too Complex Analysis: Given two analyses leading to the same conclusion, one that is simpler and easier to explain is obviously preferable. Performance analysts should convey final conclusions in as simple a manner as possible. Some analysts start with complex models that cannot be solved or a measurement or simulation project with very ambitious goals that are never achieved. It is better to start with simple models or experiments, get some results or insights, and then introduce the complications.

Improper Presentation of Results: The eventual aim of every performance study is to help in decision making. An analysis that does not produce any useful results is a failure, as is the analysis with results that cannot be understood by the decision makers. The decision makers could be the designers of a system, the purchasers of a system, or the sponsors of a project. Conveying (or selling) the results of the analysis to decision makers is the responsibility of the analyst. This requires the prudent use of

words, pictures, and graphs to explain the results and the analysis. *The right metric to measure the performance of an analyst is not the number of analyses performed but the number of analyses that helped the decision makers*.

Estrategia sistemática

- Para evita los problemas y errores señalados se debe construir las evaluaciones de performance siguiendo una estrategia sistemática:
  1. Definir el sistema y establecer los objetivos de la evaluación.
  2. Seleccionar los servicios y lo resultados del sistema que son de inetrés
  3. Seleccionar la métrica adecuada para comparar las alternativas.
  4. Listar los parámetros del sistema y del workload que afectan al comportamiento.
  5. Seleccionar los factores que pueden controlarse y afecta al comportamiento.
  6. Seleccionar la técnica de evaluación.
  7. Seleccionar el workload que corresponde al uso del sistema.
  8. Diseñar los experimentos que deben analizarse.
  9. Analizar e interpretar los datos para hacer evidentes los resultados.
  10. Generar una documentación orientada al usuario final.

Notas:

Most performance problems are unique. The metrics, workload, and evaluation techniques used for one problem generally cannot be used for the next problem. Nevertheless, there are steps common to all performance evaluation projects that help you avoid the common mistakes. These steps are as follows.

Historia del performance-oriented design

- En la época iniciales de la computación la pequeña capacidad y dimensión de los sistemas inducían tener siempre en cuenta el análisis de performance.
- La disponibilidad posterior de recursos abundante y la consecuente complejidad de los sistemas trajo la estrategia fix-it-later que actualmente impera. Las tópicos en los que se escudaba no realizar un diseño de performance eran:
  - Los problemas de performance son infrecuentes.
  - El hardware es rápido y barato.
  - Es desmasiado caro diseñar el software para que tenga mayor espuesta
  - Si es necesario, el software se puede ajustar a última hora.
  - Software eficiente equivale a código truculento.

Notas:

Analisis de los tópico de fix-it-later

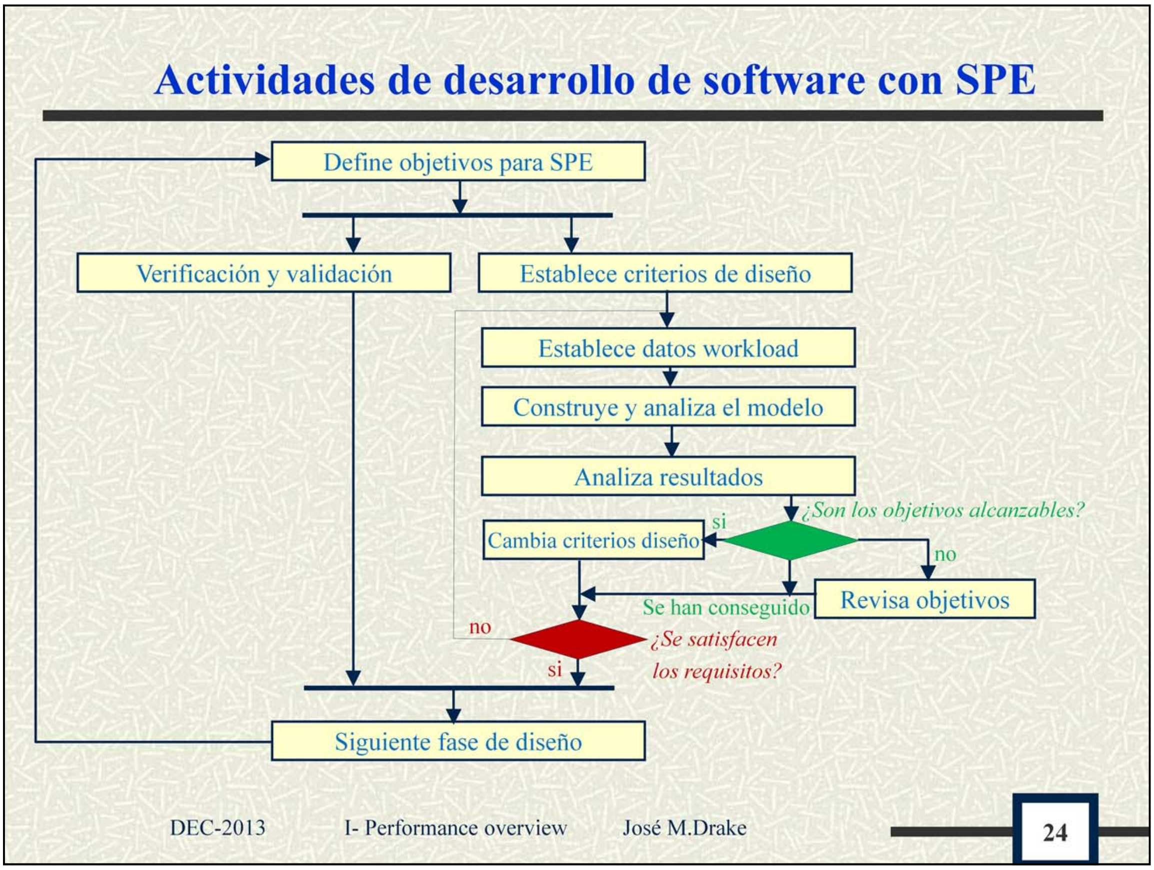
Tópico	Realidad	Consecuencias
“Los problemas de performance son raros”	- La capacidad induce mayor demanda y los problemas de performance son más graves. - Los diseñadores han perdido la intuición para incrementar la performance	-Hardware y ajustes no son soluciones inmediatas. - La baja respuesta del software produce una opción no siempre subsanable.
“El hardware es rápido y barato”	- El presupuesto no es ilimitado. -Incrementos de hardware requieren una planificación compleja. -La demanda del software puede superar cualquier capacidad hardware	-Se pierde el control de la planificación y presupuesto.
“Es demasiado caro diseñar software eficiente”	- Los problemas de performance al final, incrementa de forma imprevisible los plazos y los costos.	-Medir es mas caro si no estaba previsto. -Los costos de mantenimiento se disparan de forma incontrolada
“El ajuste puede hacer al final”	- Las causa de los problemas de performance son arquitecturales y de diseño, que son difíciles de corregir.	-El ajuste en el código tiene poco efecto sobre el comportamiento -E muy caro (o imposible) modificar las decisiones de diseño.
“Eficiencia equivale a código truculento”	-La capacidad de respuesta no es equivalente a eficiencia. -.Objetivos de performance requieren diseño inicial	-El código truculento puede ser la consecuencia de proveer la performance en la última fase.

Notas:

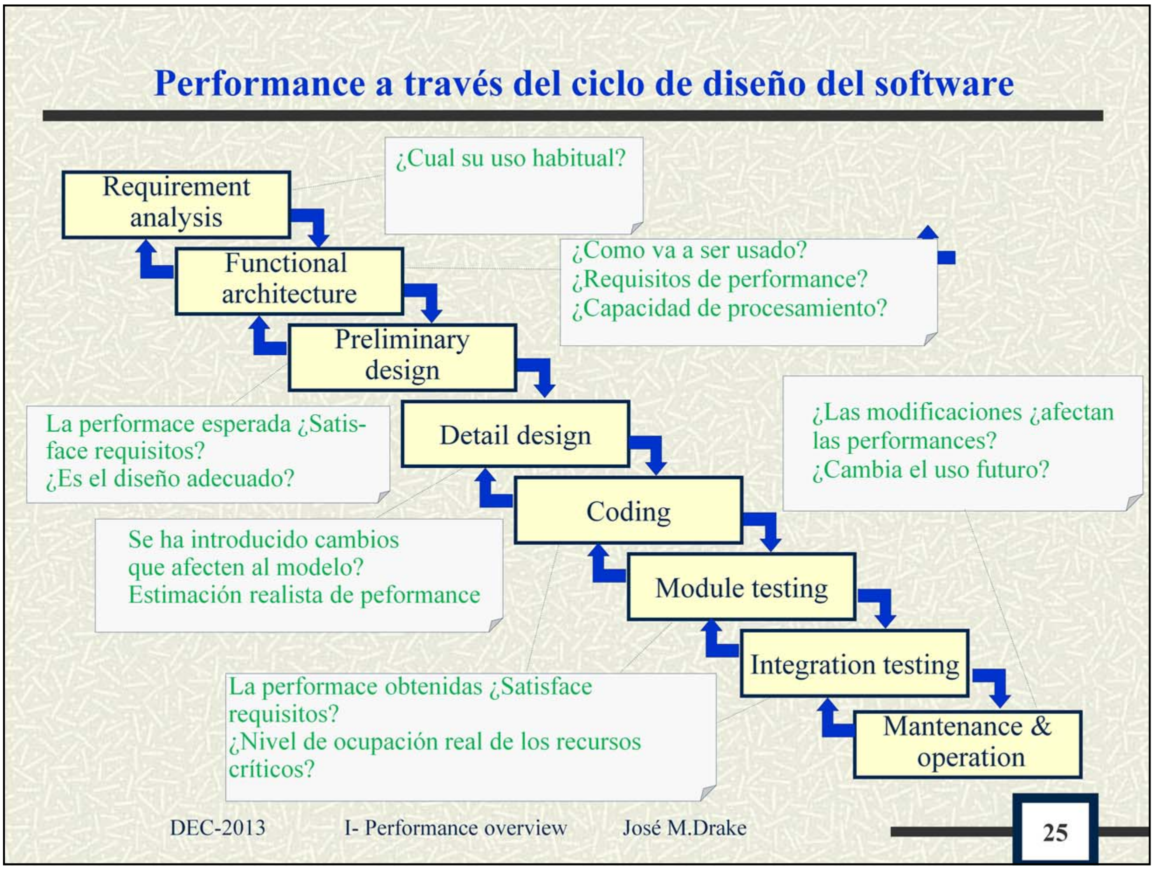
Efecto de SPE sobre el proceso de desarrollo del software

- La introducción del diseño de performance desde la fase de análisis de requerimientos tiene las siguientes ventajas:
  - Incrementa la productividad: Evita que los diseñadores trabajen en productos que posteriormente tienen que ser descartados. Obliga a los diseñadores a afrontar primero los temas críticos.
  - Incrementa la calidad y utilidad del software: Al inducir diseños adecuados y selección de la alternativa correcta, evita ajustes globales y parches de último momento.
  - Permite controlar el costo: Al prever desde la primera fase el hardware que se necesita y evitar integraciones posteriores no previstas
  - Mejora la eficiencia durante la implementación, prueba y la primera época de operación: Al proveer la capacidad de procesamiento que se necesita en el peor caso.

Notas:


Actividades de desarrollo de software con SPE
Define objetivos para SPE
Verificación y validación
Establece criterios de diseño
Establece datos workload
Construye y analiza el modelo
Analiza resultados
¿Son los objetivos alcanzables?
si
Cambia criterios diseño
no
Revisa objetivos
Se han conseguido
no
¿Se satisfacen
los requisitos?
si
Siguiente fase de diseño
DEC-2013
I- Performance overview
José M.Drake
24

Notas:


Performance a través del ciclo de diseño del software
Requirement analysis
Functional architecture
Preliminary design
Detail design
Coding
Module testing
Integration testing
Mantenance & operation
¿Cual su uso habitual?
¿Como va a ser usado?
¿Requisitos de performance?
¿Capacidad de procesamiento?
La performace esperada ¿Satis-face requisitos?
¿Es el diseño adecuado?
¿Las modificaciones ¿afectan las performances?
¿Cambia el uso futuro?
Se ha introducido cambios que afecten al modelo?
Estimación realista de peformance
La performace obtenidas ¿Satisface requisitos?
¿Nivel de ocupación real de los recursos críticos?
DEC-2013
I- Performance overview
José M.Drake
25

Notas:

Tendencias futuras

- "The future of Software Performance Enginnering" Murray Woodsite, 2008
- Requiere evolucionar:
  - Mayor fiabilidad de predicción.
  - Mayor nivel de abstracción
  - Mayor velocidad de análisis.
  - Mayor estandarización de las técnicas de instrumentación
- Integración de las técnicas de medidas con las técnicas basadas en modelos.
- Adaptación de la ingeniería SPE a las nuevas estrategias MDA, MDE.

Notas: