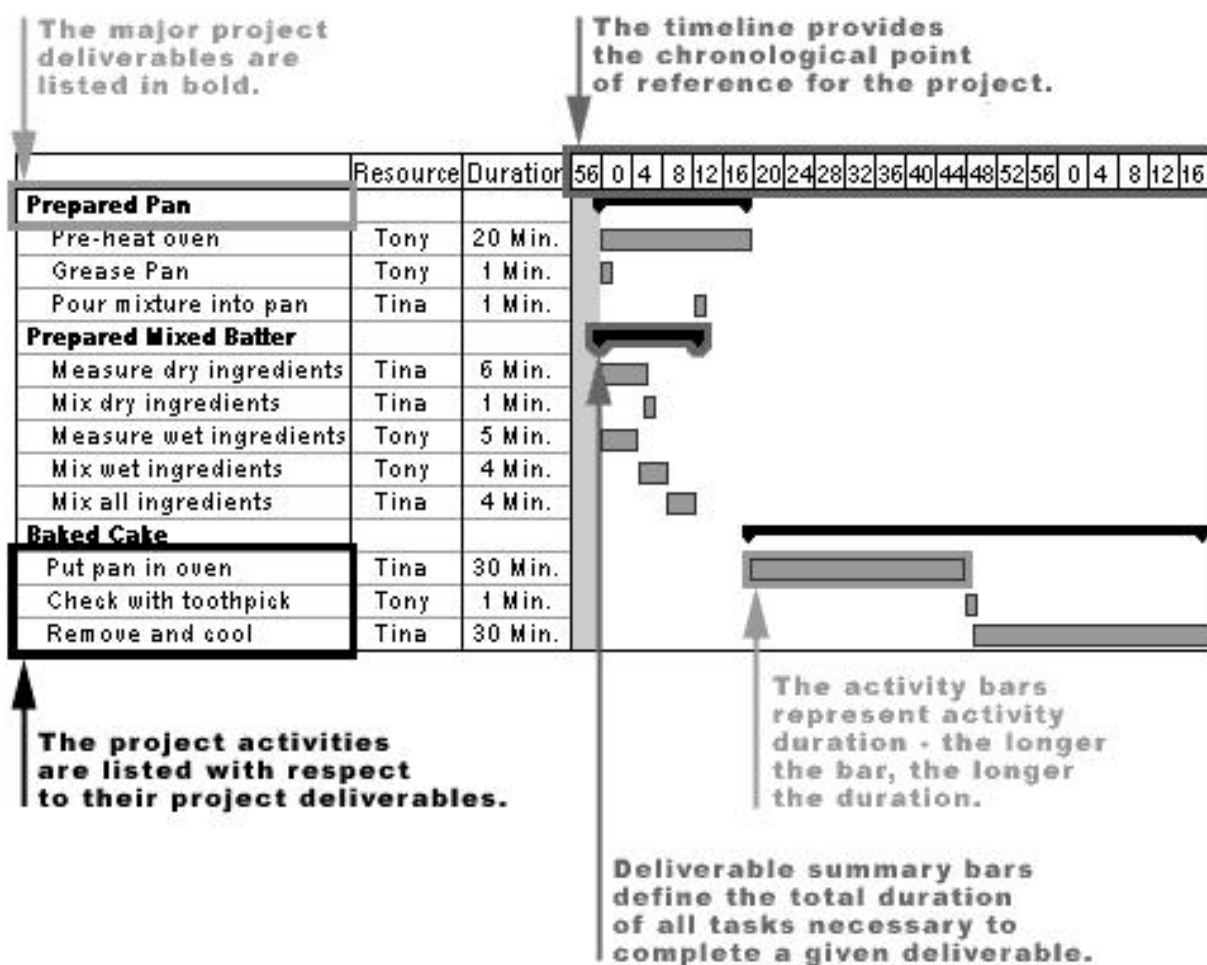


## 2. Representación gráfica aplicada a la evaluación de prestaciones

### 2.1 Presentación de los resultados

Para presentar y analizar los resultados obtenidos de la ejecución de un monitor sobre un sistema o una comparativa entre varios sistemas normalmente se usa algún tipo de gráfico, como por ejemplo el indicado en las ilustraciones del capítulo anterior, gráficos de barras de evolución temporal. Más habitualmente, para resumir las prestaciones de todo el sistema se suele usar un gráfico de Gantt , o bien un gráfico de Kiviat. Veremos éste con más detenimiento.

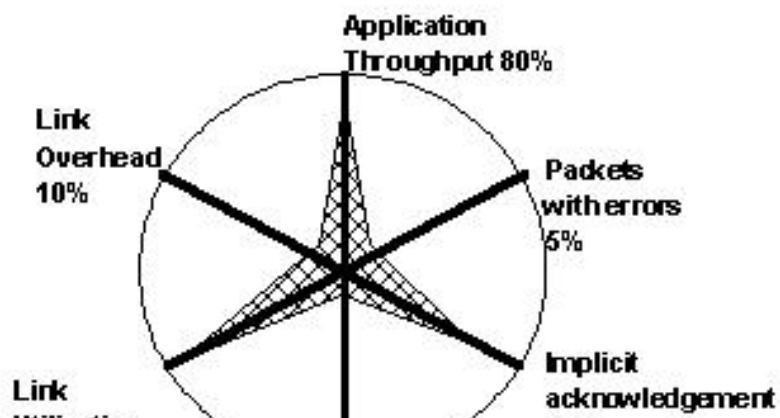


- En un gráfico de Gantt, se representa en abscisas el tiempo, y en ordenadas una línea que representa los instantes durante los cuales un

recurso ha estado ocupado. Se suelen usar gráficos similares en ingeniería del software. En [este sitio](#) se puede descargar el código fuente de un programa que traza un gráfico de Gantt; también se puede crear uno desde el mismo sitio.

- En un gráfico de Kiviat, se representan los porcentajes de uso y solapamiento de diferentes componentes del sistema como una figura geométrica que une diferentes puntos que representan esos porcentajes. Teóricamente, es posible ver de una sola vez el problema que tiene un sistema, con solo mirarlo. Habitualmente está dividido en 8 sectores; y en ellos se suele colocar alternativamente magnitudes del tipo “alto es mejor” con magnitudes del tipo “bajo es mejor”. Una de las distribuciones puede ser, en sentido e las agujas del reloj, CPU ocupada, sólo CPU, CPU y un canal, un canal sin CPU, un canal activo, CPU esperando, CPU usuario, CPU sistema, alternando índices de uso con índices de solapamiento, e índices del tipo HB (más alto es mejor) con otros del tipo LB (más bajo es mejor). Como “canal” normalmente se suele usar el o los discos, aunque también se podría representar la red o el subsistema gráfico, por ejemplo. Diferentes formas características indican diferentes "tipos" de sistemas, el sistema perfecto sería el denominado estrella de kiviat, una estrella de 4 puntas. En cualquier caso, hay que mirar siempre que los indicadores a las 12, 3, 6 y 9 sean altos, y los demás bajos, y que estos indicadores tengan más o menos el mismo valor. Valores altos de los indicadores impares indican un sistema que no funciona como debiera, y si los valores pares tienen valores muy diferentes, eso suele indicar un sistema desequilibrado, con demasiado uso de CPU o de alguno de los canales. Al gráfico de Kiviat se le puede añadir una tercera dimensión temporal, formando un tubo de Kiviat, que indica la progresión del gráfico de Kiviat a través del tiempo.

En principio, sería posible extraer una figura de mérito (FOM, figure of merit) del diagrama de Kiviat directamente, pero adolece de una serie de



fallos: por ejemplo, las variables están correlacionadas unas con otras; no todos los ejes tienen la misma escala; se consideran valores extremos buenos, y no necesariamente tiene por qué ser esto cierto.

En cuanto a representar la información sobre un sistema o un benchmark en un gráfico, el tipo del gráfico dependerá, sobre todo, del tipo de variable que se va a presentar: cualitativa o categórica, ordenada o no, que están definidas por una serie de clases mutuamente exclusivas, y cuantitativa discreta o continua, que es aquella variable cuyos niveles se expresan mejor numéricamente.

## 2.2 Reglas para representaciones gráficas

A la hora de representar gráficamente el resultado de un benchmark, o de cualquier tipo de medición de prestaciones o de carga de trabajo, es conveniente seguir las siguientes reglas:

- Minimización del esfuerzo por parte del lector: cuanto más fácil sea leer un gráfico, mejor. Por ejemplo, a la hora de poner una caja con la clave del gráfico o etiquetar las curvas, siempre será más fácil hacer lo último, puesto que no requiere comparación de colores por parte del lector.
- Maximizar información: debe de haber información suficiente en el gráfico para hacerlo autosuficiente; será mucho mejor poner la descripción de las variables que sus nombres, pues estos obligan a referirse al texto; y la descripción, cuanto más completa, mejor.
- Minimizar la tinta: presentar el máximo de información con el mínimo de tinta. Demasiada información hace el gráfico poco interesante; por ejemplo, las líneas de la rejilla no deberán de mostrarse a no ser que sean necesarias; deberán evitarse ejes superiores e inferiores, y habrá que ajustar la escala y las variables para que haya la máxima información posible con un mínimo número de cifras y rayas.
- Usar prácticas comúnmente aceptadas : como el eje x para la variable dependiente, el y para la dependiente, el origen en el 0,0, las abscisas y ordenadas en orden creciente. Si no se hace así, el lector tendrá que echar un rato hasta que entienda de qué va el tema.
- Evitar la ambigüedad: mostrar los ejes, las subdivisiones, las medidas, escalas, identificar cada una de las barras, hacerlo fácil de leer.

- No poner demasiada información ni demasiada poca: por ejemplo, traducir directamente una tabla a un gráfico sin ningún tipo de análisis suele ser absurdo, pues tiene la misma información, aunque la transmita de otra forma. Por otro lado, poner un gráfico con las medias sin indicar desviaciones estándar o rangos tampoco suele ser adecuado. Las medidas que se hagan deben de traducirse si es posible en un solo gráfico, o al menos un gráfico por cantidad a medir, evitando medidas que comuniquen poca información, por ejemplo.

## 2.3 Errores comunes en la representación gráfica de resultados

Algunos errores que se suelen cometer son los siguientes:

- Presentar demasiadas alternativas en un solo gráfico : no se deben de meter más de cinco o seis curvas en una sola gráfica, y en todo caso un gráfico de barras no debería de tener más de 10 barras, uno de tarta 8 componentes, uno de curvas 6. En un histograma, cada "cubo" debe de tener al menos 5 puntos de datos.
- Presentar muchas variables y en un solo gráfico : sobre todo si tienen diferentes escalas, ahorra espacio pero hace que la información se confunda.
- Usar símbolos en lugar de texto: en vez del nombre dado en una variable, debería de usarse lo que contiene una variable; por ejemplo, en vez de %CPU, se debería de usar Porcentaje de CPU dedicado al usuario, en un tipo de letra más pequeño.
- Seleccionar mal las escalas: como la mayoría de las veces lo hace el propio programa, uno tiene a quien echarle la culpa, pero es mejor hacerlo a mano, por ejemplo, cuando las variables tienen unos valores transitorios iniciales pequeños, y luego aumentan mucho.
- Usar un gráfico de líneas en vez de uno de barras: en variables categóricas, se debe de usar un gráfico de barras.

Por supuesto, a veces estas equivocaciones se suelen hacer a propósito, sobre todo si uno quiere demostrar algo con los gráficos. El saber este tipo de cosas ayuda también a identificarlas cuando uno se las encuentra:

- Usar orígenes no nulos para hacer énfasis de la diferencia: cuando la diferencia es pequeña, se usa un origen no nulo para que parezca más grande. Hay una

regla que dice que la altura del punto más alto debe de ser tres cuartos del offset horizontal del punto que haya más a la derecha.

- Trazar cantidades aleatorias sin los intervalos de confianza: la media transmite poca información si no se acompaña de la desviación típica, o rango o algún momento.
- Usar pictogramas escalados por altitud: a veces se usan barquitos, o cochecitos, o chuminaditas para trazar una información. Siempre hay que tener en cuenta que es la altura, y no la superficie, lo que representa la cantidad; si se mira solo la superficie, puede resultar que una cantidad parezca  $n^2$  veces mejor que otra, cuando solo es  $n$  veces mejor.

- Usar tamaños de célula no adecuados en histogramas, y usar escalas partidas en gráficos de columnas. En el primer caso, hay que aplicar tests estadísticos para comprobar que se haya hecho adecuadamente; siempre tendrá que haber al menos 5 puntos experimentales en cada cubo.

### 1 Ejercicios de autoevaluación

(Tomados del [Jain](#)) 1. ¿Qué tipo de gráfico (líneas o barras) se usaría para trazar:

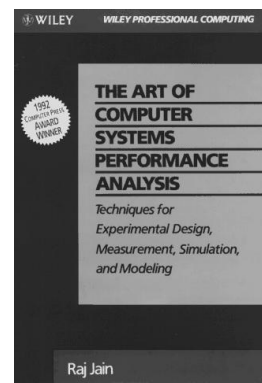
-Uso de CPU para 12 meses del año. -Uso de CPU en función del tiempo en meses. -Número de E/S a 3 discos duros: A, B, y C. -Número de E/S como una función del número de discos duros en el sistema. 2.

Para un sistema cualquiera, lista un número de magnitudes tipo "más alto es mejor" y "más bajo es mejor" y dibuja un gráfico de Kiviati típico usando valores cualesquiera.

## 2.4 Bibliografía y enlaces en internet

*Evaluación y explotación de sistemas informáticos* . R. Puigjaner, J. J. Serrano, A. Rubio.

*The art of computer systems performance analysis: Techniques for experimental design, measurement, simulation and modelling*, Raj Jain, Wiley, 1992. Por la fecha de publicación, se ve que tiene el mismo problema que el primero; sin embargo, tiene un enfoque ameno y completo al problema del análisis de prestaciones.



Representación visual de prestaciones de ordenadores paralelos, donde describe diferentes técnicas para visualizar carga de los sistemas, así como "camino críticos" y patrones de acceso a los datos.

Realizado por  
Juan Julián Merelo Guervós [jmerelo at geneura.ugr.es](mailto:jmerelo@geneura.ugr.es)  
Depto de Arquitectura y Tecnología de Computadores  
Universidad de Granada  
URL:<http://geneura.ugr.es/~jmerelo/DyEC/Tema2/DyEC-Tema2.html>