

1

Rendimiento

Objetivos

En los problemas de este capítulo se tratan aspectos relacionados con el rendimiento del computador y de sus componentes, desde la medición del rendimiento a la aplicación de la ley de Amdahl.

Problema 1. _____

Indica cuál o cuáles de las siguientes afirmaciones son CIERTAS. Contesta NINGUNA si crees que ninguna lo es.

- A) Si el tiempo de respuesta de una tarea en dos computadores A y B es T_a y T_b respectivamente, se dice que A proporciona una aceleración T_a/T_b con respecto a B.
- B) En general, el incremento del número de núcleos del procesador incrementa la productividad del computador.
- C) El rendimiento es importante solo referido a la CPU, pues el rendimiento de otros componentes del computador apenas influye sobre el rendimiento global del mismo.
- D) El tiempo de respuesta de una misma tarea en un mismo computador varía habitualmente entre diferentes ejecuciones de la misma.
- E) La desviación típica del tiempo de respuesta de una tarea es habitualmente cero.
- F) Aplicando la ley de Amdahl al desarrollo de programas, una buena estrategia de programación es optimizar el código ya desde el primer momento, antes de tener una idea clara sobre los tiempos de ejecución de las diferentes secciones que los componen.

| |
|-------|
| B y D |
|-------|

Problema 2. _____

Indica cuál o cuáles de las siguientes afirmaciones son CIERTAS. Contesta NINGUNA si crees que ninguna lo es.

- A) Aplicando la ley de Amdahl se deduce que una aceleración de valor 2 en el procesador nunca puede conseguir que el tiempo de respuesta de una tarea sea inferior a la mitad.
- B) Si un procesador tiene un número de MIPS superior a otro significa que su rendimiento es mayor.
- C) Si dos procesadores A y B incorporan el mismo núcleo y el procesador A trabaja a una frecuencia superior a la de B, se puede decir que el número de MIPS de A es mayor que el de B y que además el rendimiento del procesador A es mayor que el de B.
- D) El tiempo de respuesta de un programa viene dado única y exclusivamente por el tiempo de CPU del mismo.
- E) Si dos procesadores A y B implementan el mismo juego de instrucciones y la frecuencia de trabajo de A es mayor que la de B, esto no implica que el procesador B tenga un menor rendimiento, pues B puede tener un número de ciclos por instrucción menor que A.
- F) El tiempo de CPU de un programa indica el tiempo que dedica la CPU a ejecutarlo sin contar los tiempos de espera de la misma.

A, C, E y F

Problema 3. _____

Indica cuál o cuáles de las siguientes afirmaciones son CIERTAS. Contesta NINGUNA si crees que ninguna lo es.

- A) Los benchmarks empleados para medir el rendimiento en un computador deberían ser representativos de la carga empleada en el mismo.
- B) Comparando el rendimiento de dos computadores empleando benchmarks, no podría ocurrir que un benchmark indicase un rendimiento mayor de uno de los computadores y otro benchmark un rendimiento mayor del otro computador.
- C) Los benchmarks basados en carga real tienen el inconveniente de que son difíciles de reproducir las condiciones de prueba.
- D) No existen benchmarks para medir el rendimiento de diferentes partes del computador, como son los discos o las interfaces gráficas.
- E) Las condiciones de prueba son difícilmente reproducibles cuando se emplean benchmarks de carga sintética, pero los resultados son muy representativos.
- F) Los benchmarks basados en carga analítica emplean modelos matemáticos de los dispositivos a evaluar y de la carga de trabajo.

A, C y F

Problema 4.

Un sistema dispone de una placa base que incorpora la posibilidad de conectar uno o dos procesadores. Se pretende evaluar el rendimiento del mismo empleando uno o dos procesadores, de cuatro núcleos cada uno.

Para llevar a cabo una estimación de rendimiento se ejecutan una o varias instancias de un benchmark en cuatro escenarios diferentes:

- A) Se instala un único procesador y se ejecuta una sola instancia del benchmark. En este caso el tiempo medio de respuesta es de 10 ms.
- B) Se instala un único procesador y se ejecutan cuatro instancias del benchmark simultáneamente, una en cada núcleo. En este caso el tiempo medio de respuesta es de 10.5 ms.
- C) Se instalan dos procesadores y se ejecuta una instancia del benchmark. En este caso el tiempo medio de respuesta es de 10.1 ms.
- D) Se instalan dos procesadores y se ejecutan 8 instancias del benchmark simultáneamente, una en cada uno de los núcleos disponibles. En este caso el tiempo medio de respuesta es de 11.2 ms.

Cada experimento se repitió varias veces y el valor proporcionado es el tiempo de respuesta medio.

- **4.1** ¿Cuál crees que es la razón para que el tiempo medio de respuesta del benchmark incremente cuando se ejecutan varias instancias simultáneamente en varios núcleos?

Porque el hecho de emplear varios núcleos, pertenecientes a un único procesador o a varios procesadores, implica que se tienen varias unidades de ejecución, pero el resto de unidades del computador (memoria, dispositivos de E/S, interconexiones) no están replicados.

- **4.2** Si un usuario nunca tuviera que ejecutar varias tareas simultáneamente, ¿obtendría desde su punto de vista mayor rendimiento con un sistema monoprocesador o con uno multiprocesador? Justificar cuantitativamente la respuesta.

Obtendría más rendimiento con el sistema monoprocesador porque el tiempo de respuesta ejecutando una tarea (10ms) es menor que en un sistema multiprocesador (10.1ms). La productividad sería de $1/10 = 0.1$ tareas/ms = 100 tareas/s para el monoprocesador y de $1/10.1 = 0.09901$ tareas/ms = 99.01 tareas/s para el multiprocesador.

De los resultados anteriores podría extraerse la conclusión de que la opción de emplear dos procesadores proporciona un rendimiento menor que emplear un único procesador. El problema es que la métrica de rendimiento empleada es válida para comparar la velocidad de ejecución de una única tarea sin que haya más ejecutándose al mismo tiempo, que es lo que interesa en el escenario planteado, pero no para comparar la velocidad de ejecución de varias tareas simultáneamente, lo cual es habitual en muchos sistemas actuales.

4 Rendimiento

- 4.3 Suponiendo que el usuario tiene habitualmente que ejecutar ocho tareas simultáneamente, ¿es mejor utilizar un sistema monoprocesador o un sistema multiprocesador? Justificar cuantitativamente la respuesta.

En este caso interesa analizar la productividad:
Monoprocesador: $4/10.5 = 0.38095$ tareas/ms = 380.95 tareas/s
Multiprocesador: $8/11.2 = 0.71428$ tareas/ms = 714.28 tareas/s.
Por lo tanto, el sistema multiprocesador tendría más rendimiento.

Problema 5.

Para la mejora de rendimiento de un computador uno de los principios básicos que se deben seguir es la *Ley de Amdahl*, que determina que:

La mejora obtenida en el rendimiento de un sistema completo debido a la modificación de uno de sus componentes está limitada por la fracción de tiempo que se utiliza dicho componente.

Para ilustrar este principio supongamos que tenemos un programa en el que el 20 % del tiempo de ejecución (tiempo de respuesta) se realizan operaciones de entrada y salida con el disco, mientras que el 80 % restante se corresponde con tiempo de CPU. En esta situación, se plantean dos posibles mejoras:

- A) Sustituir el disco por otro el doble de rápido.
- B) Cambiar la CPU por otra el doble de rápida.

- 5.1 ¿Cuál es la aceleración obtenida con cada una de las opciones anteriores? Para su obtención debes emplear la fórmula de aceleración deducida de la ley de Amdahl.

$$A_a = 1/[(1 - 0,2) + (0,2/2)] = 1,11$$
$$A_b = 1/[(1 - 0,8) + (0,8/2)] = 1,67$$

- 5.2 Para entender la esencia de las aceleraciones, debes calcular ahora la aceleración correspondiente a la mejora (A) sin usar las fórmulas teóricas que proporcionan el tiempo de respuesta mejorado y la aceleración. Para ello puedes suponer que el tiempo de respuesta es T antes de la mejora y dividirlo por el tiempo de respuesta después de la mejora.

Tiempo inicial = T
Tiempo mejorado = $0,8 \times T + (0,2 \times T)/2$
Dividiendo ambos: $A_a = 1/(0,8 + 0,1) = 1,11$

Problema 6.

Supongamos que tenemos un programa que se ejecuta en 100 segundos, 75 correspondientes a tiempo de CPU y el resto a E/S. El tiempo de CPU se reduce en un cuarto cada año (tarda un cuarto menos de tiempo en ejecutarse) durante los próximos 4 años, pero no así el tiempo de E/S.

- ❑ **6.1** Inicialmente, ¿qué porcentaje de tiempo representa la E/S respecto al tiempo total de ejecución del programa?

$$25/100 \times 100 = 25 \%$$

- ❑ **6.2** ¿Cuál será el tiempo de ejecución (tiempo de respuesta) del programa transcurridos los 4 años? Redondear a valores enteros.

$$T = T_{CPU} + T_{E/S} = (75 \times (3/4)^4) + 25 = 48.73 \text{ s}$$

- ❑ **6.3** Transcurridos los 4 años de mejora, ¿qué porcentaje de tiempo representa la E/S respecto al tiempo total de ejecución del programa?

$$25/48.73 \times 100 = 51.30 \%$$

- ❑ **6.4** ¿Cuál será el factor de aceleración de la CPU al cabo de los 4 años de mejora?

$$75/(48.73 - 25) = 3.16$$

- ❑ **6.5** ¿Cuál será el factor de aceleración del computador al final de los 4 años si se emplea el programa anterior como benchmark?

$$100/48.73 = 2.05$$

- ❑ **6.6** Los resultados de las dos preguntas anteriores son diferentes, ¿cómo explicas la diferencia?

La E/S no mejora y aumenta su peso porcentual en el tiempo de ejecución

Problema 7. _____

Las métricas de rendimiento se basan en la medición del tiempo de respuesta de una tarea, o del número de tareas que pueden ejecutarse en un periodo de tiempo. Por lo tanto, para comparar el rendimiento de dos computadores es fundamental decidir en primer lugar qué es la tarea. La decisión de lo que es la tarea viene dada por el tipo de trabajo a realizar por el computador y si se trata de un contexto de monoprocesamiento o multiprocesamiento. Por ejemplo, supongamos que el tipo de trabajo consiste en llevar a cabo conversiones de documentos OpenOffice a PDF. El contexto puede ser el de un simple usuario que pretende llevar a cabo la conversión de un documento a PDF, o el de una empresa que presta servicios de conversión de documentos a PDF a través de su web. En el primer caso estaríamos hablando de un contexto de monoprocesamiento, mientras que en el segundo se trataría de un contexto de multiprocesamiento.

Para ilustrar lo anterior se plantean dos alternativas: un computador personal con un procesador de 4 núcleos y 16 GiB de RAM y un servidor con 8 procesadores de 4 núcleos y 256 GiB de RAM. En un primer momento se ha ejecutado un *benchmark* monohilo que consiste en llevar a cabo la conversión de un documento tipo a PDF en ambos computadores y se han obtenido los siguientes tiempos de respuesta:

| Computador | Tiempo 1 | Tiempo 2 | Tiempo 3 | Tiempo 4 | Tiempo 5 |
|------------|----------|----------|----------|----------|----------|
| Personal | 1.10 | 1.30 | 0.95 | 1.05 | 1.20 |
| Servidor | 0.75 | 0.85 | 0.70 | 0.90 | 0.70 |

- ❑ **7.1** ¿Cuál es el rendimiento de cada computador empleando la inversa del tiempo de respuesta y la productividad?

Coinciden ambas métricas. La productividad del computador personal es $1/[(1.10+1.30+0.95+1.05+1.20)/5] = 0.89$ y la del servidor $1/[(0.75+0.85+0.70+0.90+0.70)/5] = 1.28$.

- ❑ **7.2** ¿Cuál es la aceleración del servidor frente al computador personal? ¿Crees que esa mejora justifica la diferencia de coste entre uno y otro en el contexto de monoprocesamiento considerado?

$A_{\text{Servidor/Personal}} = 1.28/0.89 = 1.44$
No lo justifica, pues el servidor será mucho más caro y no se saca provecho de sus múltiples procesadores.

Consideremos ahora un nuevo contexto de ejecución. Se trata de prestar servicios web de conversión de documentos a PDF, por lo que interesa atender el mayor número de peticiones posibles. En este caso se obtienen unos tiempos de respuesta para cada petición un 10 % superiores a los de la tabla anterior.

- ❑ **7.3** ¿Qué métrica de rendimiento emplearías en este contexto? ¿Por qué?

La productividad, pues se trata de un contexto de multiprocesamiento.

- ❑ **7.4** Indica en este caso la aceleración del servidor respecto al computador personal. ¿Crees que la mejora puede justificar el coste económico de uno frente a otro?

$A_{\text{Servidor/Personal}} = [1.28/1.1 * (8 * 4)]/[0.89/1.1 * (1 * 4)] = 11.5$
Sí, en este caso puede estar justificada.

Cuando la carga de trabajo consiste en una única tarea y además es monolítica, esto es, debe ejecutarse en una única CPU o núcleo, se desaprovechan las capacidades de multiprocesamiento de los computadores actuales. Una solución es rediseñar el software que implementa la tarea para que éste sea multihilo, lo que permite que cada núcleo pueda ejecutar una porción diferente de la tarea. Supongamos que la conversión de un único documento a PDF se realiza con una aplicación dividida en 8 hilos y que el tiempo de ejecución es el de la tabla anterior dividido por el número de hilos que se ejecutan concurrentemente.

- ❑ **7.5** ¿Qué métrica de rendimiento emplearías en este contexto? ¿Por qué?

Valdría tanto la productividad como la inversa del tiempo de respuesta, pues se ejecuta una única tarea.

- 7.6 Indica en este caso la aceleración del servidor respecto al computador personal.

$$A_{\text{Servidor/Personal}} = [1.28 * 8] / [0.89 * 4] = 2.88$$

Problema 8.

Se pretende comparar el rendimiento de dos computadores empleados para proporcionar servicios web: un computador personal que dispone de 1 procesador de 4 núcleos y un servidor que dispone de 4 procesadores de 6 núcleos cada uno. Para llevar a cabo esta comparativa se ejecuta un *benchmark* que solicita una única página web en ambos y se mide el tiempo de respuesta en milisegundos. El experimento se repite 4 veces con la misma página. Además, como el tiempo de respuesta depende de la complejidad de la página, se prueba con dos páginas diferentes que denominamos Página A y Página B. Los resultados del experimento se muestran a continuación:

| Computador | Página A | | | | Página B | | | |
|------------|----------|-----|-----|-----|----------|------|------|------|
| Personal | 2.5 | 1.7 | 1.9 | 2.3 | 16.4 | 12.5 | 13.3 | 12.1 |
| Servidor | 1.9 | 1.8 | 1.2 | 1.3 | 12.2 | 10.3 | 10.2 | 11.1 |

Considerando la operación de servir una página web,

- 8.1 ¿Cuál es la aceleración del servidor respecto del computador personal para cada tipo de página? ¿Cuál es la aceleración agregada del servidor respecto del computador personal teniendo en cuenta las páginas A y B?

$$A_{\text{servidor/personal(Página A)}} = 1.35 \quad A_{\text{servidor/personal(Página B)}} = 1.24$$

$$A_{\text{servidor/personal}} = 1.29 \text{ (media geométrica)}$$

Los servicios web son típicamente aplicaciones multihilo y servir una página web requiere habitualmente un hilo para su ejecución. Se lleva a cabo el mismo experimento, pero ahora se ejecuta un *benchmark* que lanza 120 peticiones simultáneas de la página A y 120 peticiones simultáneas de la página B. Se supondrá además que los tiempos necesarios para servir una página son los mismos de la tabla anterior, pero lógicamente al ser los procesadores multinúcleo se sirven varias peticiones de forma simultánea. En este nuevo contexto,

- 8.2 ¿Cuál es la aceleración del servidor respecto del computador personal para cada tipo de página? ¿Cuál es la aceleración agregada del servidor respecto del computador personal teniendo en cuenta las páginas A y B?

$$A_{\text{servidor/personal(Página A)}} = 8.13 \quad A_{\text{servidor/personal(Página B)}} = 7.44$$

$$A_{\text{servidor/personal}} = 7.78 \text{ (media geométrica)}$$

Problema 9.

Se pretende comparar el rendimiento de dos computadores, A y B. El computador A posee instrucciones para operar con elementos en coma flotante (FP) mientras que el computador B traduce dichas instrucciones a un conjunto de operaciones no flotantes mediante software. Ambos computadores trabajan a una frecuencia de 500 MHz. En ambos se ejecuta el mismo programa cuyas instrucciones se dividen de la siguiente forma:

| Tipo de instruccion | Porcentaje de ocurrencia | Duración de instrucción(ciclos) | |
|---------------------|--------------------------|---------------------------------|----|
| | | A | B |
| Suma FP | 16 % | 6 | 18 |
| Multipliación FP | 10 % | 8 | 36 |
| División FP | 8 % | 10 | 63 |
| Instrucción no-FP | 66 % | 3 | 3 |

- **9.1** Indica cuál es la productividad de estos computadores para este programa en Millones de Instrucciones Por Segundo (MIPS).

$$CPI_A = 0.16 \cdot 6 + 0.1 \cdot 8 + 0.08 \cdot 10 + 0.66 \cdot 3 = 4.54$$

$$MIPS_A = \frac{500 \cdot 10^6}{4.54 \cdot 10^6} = 110.13$$

$$CPI_B = 0.16 \cdot 18 + 0.1 \cdot 36 + 0.08 \cdot 63 + 0.66 \cdot 3 = 13.5$$

$$MIPS_B = \frac{500 \cdot 10^6}{13.5 \cdot 10^6} = 37.04$$

- **9.2** Indica el tiempo de ejecución del programa, en ambos computadores, teniendo en cuenta que tiene 12000 instrucciones.

$$T_A = \frac{12000}{110.13 \cdot 10^6} = 108.96 \cdot 10^{-6} = 108.96 \mu s$$

$$T_B = \frac{12000}{37.04 \cdot 10^6} = 323.97 \cdot 10^{-6} = 323.97 \mu s$$

- **9.3** ¿Qué porcentaje de ocurrencias deberían tener los distintos tipos de instrucciones para que ambos computadores obtuvieran el mismo tiempo de respuesta?

100% en instrucciones no-FP

- **9.4** ¿A qué frecuencia debería trabajar B para igualar la productividad de A? ¿Cuál sería la aceleración de B con esta nueva frecuencia de trabajo respecto de la antigua?

$$\text{Debería trabajar a } \frac{500 \cdot 13.5}{4.54} = 1486.78 \text{ MHz}$$

$$\text{La aceleración obtenida sería } A = \frac{P_A}{P_B} = 2.97$$

- **9.5** ¿Qué aceleración se obtendría al mejorar la unidad de ejecución en coma flotante de A a una el doble de rápida?

$$\text{Ciclos de programa: } 12000 \cdot 4.54 = 54480$$

$$\text{Ciclos de programa en coma flotante:}$$

$$12000 \cdot (0.08 \cdot 10 + 0.1 \cdot 8 + 0.16 \cdot 6) = 30720$$

$$\text{Porcentaje de tiempo utilizado por la unidad de coma flotante:}$$

$$\frac{30720 \cdot 100}{54480} = 56.39 \%$$

$$\text{La aceleración obtenida sería } A = \frac{1}{(1 - 0.5639) + \frac{0.5639}{2}} = 1.39$$

- **9.6** Se plantea ahora un computador con la misma arquitectura que A que posee 4 núcleos. ¿Cuál sería la productividad de este computador cuando se lancen 16 instancias simultáneas del programa anterior en tareas por segundo? (toma como tarea la ejecución completa de una instancia del programa).

$$P = \frac{4 \text{ tareas}}{108.96 \cdot 10^{-6} s} = 36710 \text{ tareas/s}$$