

ARQUITECTURA AVANZADA DE PROCESADORES
ARQUITECTURA DE COMPUTADORES II
Curso 2014 - Segundo parcial, primera fecha 2/12/2014

APELLIDO, Nombre	
Nº de alumno	
Carrera	

Problema 1

A partir de una serie de corridas con un conjunto de datos reducido, se pudo estimar que la ejecución de una simulación con los datos completos tomaría 25 horas en un computador de primera gama. Se sabe que los primeros 60 minutos de la corrida consisten en la preparación y el particionado de tareas, de ejecución estrictamente serial e invariante con el tamaño del conjunto de datos. El resto del tiempo de ejecución es un algoritmo totalmente paralelizable con un costo de comunicación despreciable. Para acelerar los cálculos se alquilará tiempo en un cluster formado por nodos idénticos al ordenador utilizado en los ensayos, a un costo de \$10 por hora por cada computador utilizado.

- a. Si alquilo 12 nodos: ¿Cuánto tiempo necesito alquilar? ¿Cuál es la mejora respecto a la ejecución serial y cuál es el costo? Si alquilo 13hs: ¿Cuántos nodos son necesarios? ¿Cuál es la mejora respecto de la ejecución serial en este caso y cuál es el costo? Justifique: ¿Qué opción conviene si queremos minimizar la latencia? ¿Y si queremos minimizar el costo? ¿A qué modelo de medición de performance puede asimilar los cálculos anteriores?
- b. Expresé el costo como función de la eficiencia. Explique con sus palabras la diferencia entre los costos calculados en el inciso anterior, dado que la cantidad de trabajo realizada es la misma.
- c. Algún tiempo después, debido a una vacancia inesperada en la utilización del cluster, nos ofrecen un precio especial si alquilamos una ventana del cronograma formada por 5 nodos durante 9 horas. Suponiendo que aprovechamos la oportunidad y realizamos una simulación ajustando el tamaño del conjunto de datos para utilizar todas las computadoras durante todas las horas, ¿cuál es la mejora respecto del caso secuencial que tendríamos en este caso? Justifique: ¿A qué modelo de medición de performance paralela pertenece este último cálculo?

Problema 2

Un cluster de 64 nodos se utiliza para ejecutar una aplicación cuya sección serie es despreciable. La aplicación presenta una fracción comunicacional de 0.1, o sea que el tiempo de comunicaciones es el 10% del tiempo neto (total) de cómputo. Se ha observado que dicha fracción se mantiene constante, independientemente del número de nodos utilizados: si se agregan nodos el volumen de las comunicaciones aumenta proporcionalmente.

Se proponen dos opciones de mejora para el cluster, ambas con costos equivalentes de implementación:

- a. Duplicar el número de nodos.
- b. Duplicar el ancho de banda de la red.

¿Qué opción es más conveniente para la aplicación de interés? Justifique.

Problema 3

Reseñe brevemente las principales características de las arquitecturas más eficientes para la explotación del **paralelismo en los datos**. Remarque sus diferencias.

Problema 4

Compare ventajas y desventajas de la organización multiprocesador de acceso no uniforme a memoria (NUMA) respecto de una organización simétrica (SMP). Mencione cómo influye esta diferencia sobre el problema de la coherencia de los caches.