

PRÁCTICA 5: Arquitecturas paralelas

Bibliografía de referencia:

- [1] “Advanced Computer Architecture and Parallel Processing” H.El-Rewini, M.Abd-El-Barr.
- [2] Capítulo 4 “*Multiprocessors and Thread-Level Parallelism*” del libro “*Computer Architecture – A Quantitative Approach*” de J. Hennessy y D. Patterson, quinta edición.
- [3] Artículo “*Isoefficiency : Measuring the Scalability of Parallel Algorithms and Architecture*”. A.Y. Grama, A. Gupta, y V. Kumar

1) En [1] se proponen dos modelos iniciales para considerar la mejora en sistemas de cómputo paralelo: el modelo de subtareas de igual duración, y el de cálculo paralelo con sección serial.

- a) ¿Qué diferencia conceptual existe entre estos dos modelos? ¿Cuándo será más apropiado el primero y cuándo el segundo?
- b) Con sus palabras enuncie la Ley de Amdahl, deduzca la expresión de la cota a la mejora, y explique la consecuencia más inmediata de la ley referida a la relación costo/beneficio de incrementar la cantidad de nodos de procesamiento.
- c) Las conclusiones de la Ley de Amdahl son válidas en la medida en que se sostengan sus hipótesis previas. ¿Qué suposición clave hace la ley de Amdahl respecto de la fracción f ? Discuta la aplicabilidad de esta hipótesis si (a) mantengo el tamaño del problema constante y cambio la cantidad de nodos n , y (b) si permito que el tamaño del problema cambie mientras mantengo n constante.
- e) Un modelo más realista debe tener en cuenta el *overhead* de comunicación t_c . Determine la cota máxima que tiene la mejora alcanzable la mejora si se tiene en cuenta esta variable en ambos modelos (subtareas de igual duración y cálculo paralelo con sección serial). ¿Puede considerarse el costo de comunicación como parte del tiempo de ejecución serial?
- f) Defina con sus palabras la eficiencia referida a los sistemas de procesamiento paralelo, y dé su expresión para cada uno de los modelos obtenidos anteriormente.

2) Siguiendo con [1], respecto del modelo de Gustafson-Barsis:

- a) Señale la diferencia fundamental del modelo G-B respecto del modelo utilizado por la Ley de Amdahl.
- b) ¿Qué es el *scaled speedup* y cómo se compara con el *speedup* convencional? Determine su expresión en función de las fracciones serial s y paralela p de tiempo que insume la ejecución del problema en cada nodo.
- c) ¿Qué ocurre con el *scaled speedup* a medida que se incrementa n ? Contraste con Amdahl.
- d) Incorpore el tiempo de comunicación al modelo utilizado por G-B, suponiendo que este crece linealmente con la cantidad de nodos. ¿Qué ocurre ahora con el *scaled speedup* cuando n tiende a infinito?

3) Considere que cuenta con un clúster no homogéneo formado por 20 máquinas pertenecientes a diferentes generaciones tecnológicas, A, B y C. Originalmente el clúster estaba formado por 12 máquinas de generación A, cada una de las cuales es capaz de alcanzar un rendimiento sostenido de 200 Mflops. Sobre esta configuración del sistema se realizaron dos expansiones posteriores, primero con 6 máquinas de generación B (c/u de 250 Mflops) y luego con dos máquinas adicionales de generación C (400 Mflops).

Se debe ejecutar un programa cuya fracción paralelizable es del 70%, y del cual se sabe que ejecutado en la configuración inicial del clúster (cuando tenía 12 máquinas) insumía 12 horas. El tiempo de comunicación es despreciable.

- a) ¿Cuál será el tiempo de ejecución del programa en la configuración actual (20 máquinas) del clúster?
- b) Según Amdahl, ¿cuál es la cota mínima del tiempo de ejecución para este sistema?
- c) ¿Qué modificación debo hacer si quiero lograr un tiempo de ejecución por debajo de la cota anterior?

4) Considere tres tipos de computadores diferentes, A, B y C, con las siguientes capacidades de cálculo para un tipo de problema representativo:

- Máquinas A : 200 Mflops de capacidad sostenida por cada máquina.
- Máquinas B : 250 Mflops por cada máquina.
- Máquinas C : 400 Mflops por cada máquina.

Se quiere ejecutar una simulación de la cual se sabe que tiene una fracción no paralelizable del 1%, y que ejecutado en un clúster homogéneo de 120 máquinas tipo A insume 48 hs. Para ello existe la posibilidad de realizar la simulación en uno de tres clúster no homogéneos posibles, por un módico precio en cada caso:

	Conformación	Costo
Clúster X	100 máquinas A, 40 B, 10 C	\$4.0 por cada hora
Clúster Y	80 máquinas B, 10 C	\$4.5 por cada hora
Clúster Z	80 máquinas C	\$6.0 por cada hora

En todos los casos la parte serie se ejecuta en la mejor máquina disponible (una de tipo C).

- a) ¿Cuál es la cantidad de operaciones punto flotante que conforman la simulación?
- b) ¿Cuál es el límite inferior en el tiempo de ejecución de la simulación, según Amdahl? Indique claramente bajo qué condiciones en las que es válida su respuesta.
- c) ¿Qué clúster resolverá la simulación más rápidamente? ¿Cuál lo hará de forma más económica?

5) Se tiene un programa de reconocimiento de objetos en imágenes de radar SAR (Radar de Apertura Sintético) que se utiliza para procesar los archivos de imágenes de alta resolución recibidos de un satélite. En cada utilización se debe procesar un paquete 64 archivos que abarcan una cierta región geográfica, y el procesamiento de cada archivo consta de las siguientes etapas:

1. Carga del archivo de imagen en memoria.
2. Ejecución del algoritmo de detección.
3. Escritura de resultados en un nuevo archivo.

Los pasos (1) y (3) no son paralelizables, pero (2) sí lo es. Utilizando una única CPU se miden los siguientes tiempos totales de ejecución procesando un único archivo: para (1) 150ms, para (2) 800ms, y (3) 50ms.

Se proponen dos variantes del programa que se podrían desarrollar para reducir el tiempo de procesamiento:

- a) **Multithread.** Dividir el programa en *threads* para explotar los cuatro *cores* que tiene la CPU de la computadora que corre el programa. La comunicación de datos entre *cores* se realiza por memoria compartida, por lo que el costo de comunicación puede considerarse despreciable.
- b) **Multithread + Clúster.** Modificar el programa para que use un *clúster* de 16 computadoras idénticas a la primera, usando los cuatro *cores* que tiene cada una. Esto agrega un tiempo de comunicación de 200ms para trasladar los datos por red por *broadcast*.

Compare las mejoras obtenidas en cada caso en el tiempo total de ejecución que toma procesar cada paquete de imágenes, siempre tomando como referencia la configuración inicial del sistema (procesamiento secuencial). Determine la eficiencia en cada caso.

6) A partir de los resultados del ejercicio anterior se hace una serie de ensayos y se determina que la latencia de los accesos a disco está dominada por el movimiento mecánico de los cabezales para encontrar el inicio del archivo, por lo que para los tamaños de archivos utilizados en el paso (1) el tiempo de acceso no aumenta sustancialmente al incrementar su tamaño. De forma similar el tiempo de comunicación incluye la latencia de establecer la comunicación, la cual no depende de la cantidad de datos a transmitir.

Entusiasmados por estos hallazgos se emprende una nueva campaña de mediciones donde se ensaya aumentar en cuatro veces el tamaño de las imágenes individuales para reducir los costos fijos del procesamiento; en esta nueva configuración del sistema, para cubrir la misma región geográfica ahora es necesario procesar sólo 16 imágenes.

En esta nueva condición ahora los tiempos medidos por cada imagen individual procesada son: para (1) 200ms, para (2) 3200ms, y (3) 50ms; el tiempo de comunicación para transmitir los datos por red crece a 700ms.

Vuelva a calcular los *speedups* con los sistemas (a) y (b) respecto de la configuración original del sistema (procesamiento secuencial). Si para utilizar la opción (b) es necesario comprar las 15 máquinas adicionales que requiere el clúster, ¿le parece una inversión razonable?

7) Un programa que tarda 3 horas en ejecutarse en un procesador único, lo hace en 10 minutos en una red de 36 procesadores equivalentes. El tiempo de comunicación puede considerarse despreciable.

- a) Explique la utilidad de la métrica de Karp-Flatt.
- b) ¿Qué limitaciones tiene un análisis basado en esta métrica? ¿permite decir algo sobre cómo se modificará el *speedup* si cambia el tamaño del problema? ¿y si cambia la cantidad de procesadores?
- a) Calcule la métrica de Karp-Flatt para el programa. ¿Cuál sería la máxima mejora esperable si se pudiera incrementar indefinidamente el número de procesadores?

8) Suponga un programa de aproximadamente 10^{12} instrucciones que opera con matrices de números en punto flotante de doble precisión. La ejecución del programa en un procesador de 1 GFLOP demora $10^{12}/10^9=1000$ segundos. Suponga que se requiere reducir el tiempo de ejecución a 25 segundos o menos, y que para ello se dispone de una red de computadoras idénticas entre sí cada una con una capacidad de cómputo de 1 GFLOP y todas interconectadas con Ethernet de 100Mbps.

Luego de analizarlo se concluye que existe la posibilidad de paralelizar el programa para que sea ejecutado en múltiples nodos simultáneamente. Para ello se debe utilizar la siguiente estrategia: un único nodo principal debe enviar a cada uno de los n nodos secundarios, utilizando la red disponible, una matriz de 300×300 números. Esta será procesada por el nodo secundario utilizando $10^{12}/n$ instrucciones, resultando una nueva matriz de 300×300 que debe ser transmitida de regreso al nodo principal.

- a) ¿Cuántos nodos se requieren para obtener la mejora deseada?
- b) ¿Cuál es la eficiencia del sistema en ese caso?
- c) ¿Es posible reducir el tiempo de ejecución a 20 segundos?
- d) ¿Cuál es la máxima mejora que puede esperarse y cuántos nodos son necesarios en ese caso?
- e) Graficar la mejora y la eficiencia en función del número de nodos

9) Calcule el grado, el diámetro y el costo de interconectar N nodos utilizando cada una de las siguientes topologías estáticas: a) anillo; b) grilla; c) toroide 2D; d) toroide 3D; e) hipercubo. Compare la relación entre costo y cantidad de nodos de las topologías hipercubo y toroide 3D. Grafique esta relación para N en el rango 1 a 256; ¿qué tipo de conexión es más conveniente para un sistema de pocos nodos? ¿y para un super sistema de 4096 nodos?

10) Considere un sistema SMP basado en bus, construido con procesadores capaces de ejecutar 10^6 instrucciones/seg, conectados a su memoria principal a través de un bus de ancho de banda pico de 10^5 captaciones/seg. Los caches están diseñados para soportar una tasa de aciertos del 90%.

- a) ¿Cuál es el máximo número de procesadores que puede utilizar este sistema?
- b) ¿Cuál es la tasa de aciertos necesaria si deseo utilizar 20 procesadores?

11) Considere un sistema multiprocesador de tipo SMP formado por núcleos RISC de 32 bits, y cuya jerarquía de memoria se encuentra compuesta por dos niveles de cache y por la memoria principal.

- Cada núcleo ejecuta en promedio 600 MIPS y realiza 150×10^6 accesos a palabras en la jerarquía de memoria por segundo.

- La cache L1 es individual de cada núcleo, está unificada para datos e instrucciones, y fue diseñada con una tasa de acierto del 95%. La cache L2 es compartida por todos los núcleos y tiene una tasa de acierto del 92%.
- El ancho de banda del bus que une todas las cache L1 con la de nivel L2 es de 800 MB/seg, mientras que el bus que une la cache L2 con la memoria principal es de 50 MB/seg.

1. ¿Cuánto núcleos puede soportar este diseño de jerarquía de memoria?
2. Se propone incrementar el tamaño de la cache L2 para mejorar su tasa de aciertos. ¿Qué tasa de aciertos es necesaria para llevar la cantidad de núcleos a 6?

12) Analice la escalabilidad del siguiente conjunto respecto de la eficiencia y del costo.

El problema a resolver es una interacción entre nodos en tiempo real en que cada nodo debe enviar un mensaje cada 10^6 instrucciones a algún otro nodo de la red, con distribución uniforme para el destino (todos los nodos tienen la misma probabilidad de recibir cada mensaje).

Para ejecutar la aplicación se dispone de n nodos conectados en anillo, utilizando pasaje de mensajes. Cada mensaje insume un tiempo de comunicación igual a $T_m \times D$, donde $T_m = 1\mu s$ y D es la distancia entre el nodo origen y el nodo destino. ¿Depende la eficiencia de la cantidad de nodos n ? ¿Y del tamaño del problema?

13) Considere el siguiente problema a ser resuelto en un sistema paralelo.

Uno de los nodos se encarga de una sección serie inicial de aproximadamente 10^8 instrucciones. Luego debe transmitir 100KB a cada nodo para ponerlo a calcular. Cada nodo ejecuta independientemente 10^9 instrucciones. Al final de los cálculos, 100KB deben transmitirse desde cada nodo como respuesta y debe realizarse un número despreciable de instrucciones con los datos recibidos.

Suponga que se dispone de un clúster homogéneo con las siguientes características: nodos doble núcleo superescalares (HTx2) de 1GHz; red serie de 100MHz con topología de anillo con testigo que viaja en un único sentido (token ring).

- i) Dimensione la granularidad del problema. Respecto de la sección serie, qué modelo aplica mejor, Amdahl o Gustafson-Barsis?
- ii) Evalúe el impacto de las comunicaciones en el problema: ¿cuánto mejoraría si se utilizara un token ring de 1GHz?. ¿Y si se utilizara uno de 10MHz?
- iii) Evalúe la escalabilidad de conjunto respecto de la eficiencia y de su extensibilidad (facilidad para agregar nuevos nodos en la topología).
- iv) Considere el efecto de una sección serie que tenga una relación logarítmica con el número de nodos (por ejemplo $f = 10^7 \log_2(n)$).
- v) Estime el impacto respecto de la eficiencia y la extensibilidad si se utiliza una red con topología de toroide en dos dimensiones (cuatro enlaces por nodo en lugar de dos).