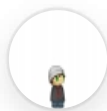


WUOLAH



Jiuxe

www.wuolah.com/student/Jiuxe



3430

Ejercicios grupos resueltos.pdf

Ejercicios Resueltos



2º Arquitectura de Computadores



Grado en Ingeniería Informática



Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación
UGR - Universidad de Granada



MÁSTER EN FINANZAS

¿Quieres alcanzar el **éxito profesional**?

■ Título Oficial

■ Prácticas Profesionales



Semana de Formación en Londres



www.cunef.edu

Ejercicio Voluntario 1. Arquitectura de Computadores.

La empresa *Freaky Datamining* estima que debe adquirir un nuevo servidor con una velocidad pico de 100 TFLOPS para alcanzar los niveles de tiempos de respuesta requeridos en su nueva generación de algoritmos para aplicaciones *Big Data*.

Para configurar la máquina se ha decidido utilizar, para cada nodo, un servidor HP ProLiant SL230s Gen8 (cuyas características se pueden consultar en http://h18000.www1.hp.com/products/quickspecs/14231_na/14231_na.pdf).

Concretamente, cada uno de estos servidores tiene dos procesadores Sandy Bridge Intel® Xeon® E5-2670 a 2.60GHz con 8-cores por procesador.

- (a) ¿Cuántos nodos (servidores HP ProLiant SL230s) se necesitan para configurar la máquina de 100 TFLOPS?
- (b) En la Lección 2 de AC se han presentado diferentes criterios de clasificación de computadores. Clasifique el nuevo servidor que se pretende adquirir, sus nodos, sus encapsulados y sus núcleos dentro de la clasificación de Flynn y dentro de la clasificación que usa como criterio el sistema de memoria. Razone su respuesta.
- (c) ¿Cuál es el número máximo de operaciones de coma flotante por ciclo de cada core del Intel Xeon E5-2670?

NOTA: La velocidad pico de cada core se puede determinar a partir de la lista TOP500 (<http://www.top500.org/>), en cuya última edición de Noviembre de 2012 se tiene que, por ejemplo, el equipo número 13 (el Yellowstone del National Center for Atmospheric Research), que utiliza el procesador Intel® Xeon® E5-2670 a 2.6 GHz, tiene una velocidad pico de 1503590 GFLOPS, y contiene 72288 cores.

SOLUCIÓN:

(a)

Como sabemos por la información que nos proporciona la lista TOP500 (y que se indica en la NOTA del problema) que una máquina con el procesador Intel® Xeon® E5-2670 a 2.6 GHz, tiene una velocidad pico de **1503590 GFLOPS**, y contiene **72288 cores**, la velocidad pico de un procesador para operaciones en coma flotante (número máximo de GFLOPS que puede proporcionar) es:

$$1503590 \text{ GFLOPS} / 72288 \text{ cores} = 20.8 \text{ GFLOPS/core}$$

Dado que necesitamos 100 TFLOPS = 100×10^3 GFLOPS, el número de cores que necesitamos es $(100 \times 10^3 \text{ GFLOPS}) / (20.8 \text{ GFLOPS/core}) = 4807.7$ cores, es decir 4808 cores

Como cada servidor HP ProLiant SL230s que constituirá un nodo tiene 16 cores (2 microprocesadores con 8 cores cada uno), el número de nodos es:

$$(4807.7 \text{ cores}) / (16 \text{ cores/nodo}) = 300.5 \rightarrow 301 \text{ nodos}$$

(b)

Desde el punto de vista de la taxonomía de Flynn es un computador MIMD. Cada uno de los microprocesadores que hay en el nodo tiene 8 núcleos cada uno que comparten la memoria local del microprocesador y por tanto son multiprocesadores UMA. Esos dos microprocesadores se interconectan en el nodo constituyendo un multiprocesador

¡Define tu sueño y alcánzalo!

NUMA dado que cada uno de ellos tiene su memoria principal local. Los nodos están interconectados a través de una red y configuran un computador NORMA (también suele denominárseles clusters).

(c)

Como la velocidad pico de un core es igual al producto de la frecuencia de reloj y el número de instrucciones por ciclo (en este caso operaciones en coma flotante por ciclo) tenemos que, en cada core:

$$20.8 \text{ GFLOPS} = (\text{Operac_Coma_Flotante/ciclo}) \times \text{Frecuencia (ciclos/s)}$$

$$20.8 \text{ GFLOPS} = (\text{Operac_Coma_Flotante/ciclo}) \times 2.6 \times 10^9 \text{ (ciclos/s)}$$

y por tanto

$$\begin{aligned} (\text{Operac_Coma_Flotante/ciclo}) = \\ (20.8 \times 10^9 \text{ Operac_Coma_Flotante/s}) / (2.6 \times 10^9 \text{ ciclos/s}) = 8 \end{aligned}$$

Por tanto, el número máximo de operaciones en coma flotante que puede terminar el core es **8**.

Ejercicio Voluntario 2. Arquitectura de Computadores

La página web <http://zhanglab.ccmb.med.umich.edu/I-TASSER/> permite acceder al servidor *I-TASSER* online para obtener, a partir de una secuencia de aminoácidos, la estructura 3D de la proteína compuesta por dichos aminoácidos. Se ha desarrollado una aplicación de *cloud* que permite que un programa paralelo MPI que se ejecuta en un multicomputador pueda obtener estructuras 3D enviando a la página web la secuencia de aminoácidos que se precise en la aplicación correspondiente. El tiempo de respuesta para una solicitud de estructura 3D es del orden de 3 minutos.

Se está intentando solucionar un problema de bioinformática mediante dos programas paralelos alternativos:

- a) En uno de ellos, en cada iteración se solicita a la página web la estructura 3D de una secuencia de aminoácidos obtenida por un proceso que consume $T_a=60s$ en un procesador, y con la respuesta recibida se aplica un algoritmo en el que pueden trabajar en paralelo, y de forma independiente, todos los p procesadores del computador durante un tiempo $T_p=600s$ (si aumenta el número de procesadores que podemos poner a trabajar en el algoritmo consumen el mismo tiempo de 600s pero se consigue mejorar la calidad de la solución).
 - b) En el otro programa, igualmente en cada iteración se solicita a la página web la estructura 3D de una secuencia de aminoácidos obtenida por un proceso que consume $T_a=60s$ en un procesador, pero el trabajo a realizar con la respuesta recibida, que si se ejecutase en un solo procesador necesitaría 15 minutos, se distribuye entre los p procesadores utilizados en la implementación del algoritmo.
- ¿Cuál es la ganancia de velocidad en cada caso en función del número de procesadores utilizado, p ? ¿Qué ley, de Amdahl o Gustafson, sería la que se aplica en cada caso?
 - Para un valor del número de procesadores, p , exprese el valor de la f de la ley de Amdahl, llamémosle f_a , en función de la f de la ley de Gustafson, llamémosle f_g , (y del valor de p)?

(NOTA: Los tiempos de sincronización/comunicación entre los procesadores paralelos se pueden despreciar, y los restantes tiempos de comunicación se consideran incluidos en los 3 minutos de respuesta del servidor de la página web).

SOLUCIÓN:

¿Cuál es la ganancia de velocidad en cada caso en función del número de procesadores utilizado, p ? ¿Qué ley, de Amdahl o Gustafson, sería la que se aplica en cada caso?

En el programa (a) el tiempo de ejecución paralelo es constante aunque se utilicen más procesadores, a esos procesadores se les asigna una carga de trabajo que consume el mismo tiempo. Lo que sucede es que si ese mismo trabajo lo tuviera que hacer un solo procesador, el tiempo crecería (porque se hace más trabajo). En este caso se aplica la ley de Gustafson:

$$T_{\text{sec}} = 180 \text{ segundos} + 60 \text{ segundos} + p \cdot 600 \text{ segundos} = 240 + p \cdot 600 \text{ segundos}$$

$$T_{\text{par}} = 180 \text{ segundos} + 60 \text{ segundos} + 600 \text{ segundos} = 840 \text{ segundos}$$

$$S(p) = T_{\text{sec}} / T_{\text{par}} = (240 + p \cdot 600) / 840 = 0.285 + 0.715 \cdot p$$

Obsérvese que, según la ley de Gustafson, $f_g = 240/840 = 0.285$

En el programa (b), el tiempo de ejecución paralelizable de 15 minutos se distribuye entre los procesadores de la máquina. Por tanto,

$$T_{\text{sec}} = 180 \text{ segundos} + 60 \text{ segundos} + 900 \text{ segundos} = 1140 \text{ segundos}$$

$$T_{\text{par}} = 180 \text{ segundos} + 60 \text{ segundos} + 900/p \text{ segundos} = 240 + 900/p \text{ segundos}$$

$$S(p) = T_{\text{sec}} / T_{\text{par}} = 1140 / (240 + 900/p)$$

Obsérvese que, según la ley de Amdahl, $f_a = 240/1140 = 0.211$

Para un valor del número de procesadores, p , exprese el valor de la f de la ley de Amdahl, llamémosle f_a , en función de la f de la ley de Gustafson, llamémosle f_g , (y del valor de p)?

$$T_{\text{sec}} = T_{\text{par}} \cdot f_g + T_{\text{par}} \cdot (1 - f_g) \cdot p \quad (\text{teniendo en cuenta la propia definición de la } f_g)$$

Por otra parte, según la definición de la f de Amdahl (f_a),

$$f_a = (T_{\text{par}} \cdot f_g) / (T_{\text{par}} \cdot f_g + T_{\text{par}} \cdot (1 - f_g) \cdot p) \quad (\text{la parte no paralelizable del tiempo secuencial})$$

Eliminando T_{par} en el numerador y el denominador nos queda:

$$f_a = f_g / (f_g + (1 - f_g) \cdot p)$$

Es decir, para $p=1$, $f_a = f_g$, y para $p \rightarrow \infty$, $f_a \rightarrow 0$ (hay cada vez menos proporción de parte no paralelizable porque se añade trabajo a medida que se introducen más procesadores)