

Sesión 3: Modelo de Eficiencia y Escalabilidad para Clusters Heterogéneos

Índice

1. Introducción.
2. Modelo de Isoeficiencia.
3. Función de Isoeficiencia Heterogénea.
4. Resultados Experimentales.
5. Conclusiones y Trabajos Futuros.

Introducción

- Eficiencia y escalabilidad son dos características muy importantes de los sistemas paralelos, particularmente de los clusters.
- La **eficiencia** es un índice del grado de rendimiento real, comparado con el máximo valor alcanzable.
- La **escalabilidad** nos da una idea del comportamiento del sistema cuando se incrementa el número de procesadores.

16 de mayo de 2006

Jose Luis Bosque

3

Introducción

- El análisis de escalabilidad nos da una idea teórica del comportamiento del sistema al incrementar su capacidad de cómputo:
 - Comparar el comportamiento de diferentes algoritmos.
 - Predecir efectos en el rendimiento del sistema al cambiar alguno de sus parámetros.
 - Optimizar la implementación paralela.

16 de mayo de 2006

Jose Luis Bosque

4

Introducción

- Una cuestión importante a tener en cuenta es la heterogeneidad del sistema.
- Hay dos razones importantes que han estimulado la introducción de la heterogeneidad en los clusters:
 - La flexibilidad de estos sistemas que permite sustituir nodos antiguos o introducir nuevos nodos de forma muy sencilla.
 - La tremenda tasa de mejora de los PC's y estaciones de trabajo (Ley de Moore), por lo que los nuevos nodos que se introduzcan tendrán capacidades de cómputo muy diferentes.

16 de mayo de 2006

Jose Luis Bosque

5

Introducción

- Todos las métricas y modelos para eficiencia y escalabilidad asumen sistemas homogéneos.
- Estos modelos incluyen "N" el número de procesadores en sus expresiones.
- En sistemas heterogéneos la potencia de cómputo es también importante.
- Hay una fuerte necesidad para desarrollar herramientas y técnicas para analizar el comportamiento del entornos heterogéneos.

16 de mayo de 2006

Jose Luis Bosque

6

Introducción

- Nuevas Contribuciones en el campo de la escalabilidad de sistemas heterogéneos:
 - *Una nueva definición de eficiencia que puede aplicarse tanto a sistemas homogéneos como a heterogéneos.*
 - *Una técnica para analizar la escalabilidad de sistemas heterogéneos, basada en el concepto de isoeficiencia heterogénea.*
 - *Un conjunto de teoremas de escalabilidad que permiten predecir la escalabilidad de conjuntos de algoritmos de forma teórica.*

16 de mayo de 2006

Jose Luis Bosque

7

Estado del Arte

- Métrica de isovelocidad:
 - La velocidad de ejecución de un algoritmo se define como el trabajo dividido por el tiempo de respuesta.
 - La velocidad debería incrementarse de forma lineal con el tamaño del sistema.
 - Esto no es así, debido a las sobrecargas de comunicación y sincronización de los sistemas paralelos y que se incrementan al aumentar el número de procesadores.
 - Un sistema paralelo es escalable si la velocidad media alcanzada por el algoritmo sobre esta arquitectura permanece constante al aumentar el número de procesadores, permitiendo que el tamaño del problema crezca de la misma forma:

$$\psi(N, N') = \frac{N' \cdot W}{N \cdot W'}$$

16 de mayo de 2006

Jose Luis Bosque

8

Estado del Arte

■ Speedup asintótico:

- Para un algoritmo y un tamaño de problema dados, se define el paralelismo inherente como el cociente entre el tiempo serie y el tiempo en una máquina paralela ideal.
- Para el mismo algoritmo y tamaño del problema se obtiene el máximo speedup alcanzable por cualquier máquina con esa arquitectura y el paralelismo inherente es su speedup asintótico en la máquina ideal.
- Se define el speedup asintótico, $S(N)$, para una arquitectura, algoritmo y tamaño del problema dados como el mejor speedup que se puede alcanzar variando sólo el número de procesadores del sistema:
- Se define la función de escalabilidad, de una arquitectura y algoritmo concretos como el cociente de los speedups asintóticos de la máquina real y la ideal.

16 de mayo de 2006

Jose Luis Bosque

9

Modelo de Isoeficiencia

- El modelo isoeficiencia parte de un sistema paralelo compuesto por N procesadores idénticos.
- Este modelo se basa en dos consideraciones:
 - La eficiencia de un sistema paralelo decrece cuando se incrementa el número de procesadores.
 - Para un número de procesadores dado, instancias más grandes de un problema obtienen mejores valores de speedup y eficiencia.
- Si se incrementan simultáneamente estos dos factores se podrá mantener la eficiencia constante.

16 de mayo de 2006

Jose Luis Bosque

10

Modelo de Isoeficiencia

■ Definiciones previas:

- Tiempo de ejecución secuencial.
- Tiempo de ejecución paralelo.
- Sobrecarga.
- Tamaño del problema.
- Eficiencia.

$$E = \frac{1}{1 + \frac{T_o}{t_c \cdot W}}$$

Modelo de Isoeficiencia

- Algunas combinaciones de algoritmos y arquitecturas cumplen esta condición y se denominan sistemas paralelos escalables.
- La ley que gobierna cuanto se debe aumentar el tamaño del problema, cuando crece el número de procesadores, determina como de escalable es el sistema y se denomina función de isoeficiencia.

$$W = K \cdot T_o(N)$$

Modelo de Isoeficiencia Heterogéneo

- El sistemas heterogeneos la potencia de cómputo depende tanto del número de procesadores como de la potencia de cada nodo individual, que puede ser muy diferente de unos a otros.
- El modelo de isoeficiencia se basa en el número de procesadores como parámetro, lo cual no es válido en sistemas heterogéneos.
- La heterogeneidad nos fuerza a buscar una formulación alternativa tanto para la eficiencia como para la escalabilidad.

16 de mayo de 2006

Jose Luis Bosque

13

Modelo de Isoeficiencia Heterogéneo

- Justificación de la elección del modelo de isoeficiencia:
 - Modelo sólidamente asentado.
 - Método a priori.
 - Captura en una sola expresión los efectos de las características de la arquitectura y del algoritmo.
 - Es más adecuado para medir la escalabilidad de algoritmos frente a la escalabilidad de las arquitecturas.

16 de mayo de 2006

Jose Luis Bosque

14

Modelo de Isoeficiencia Heterogéneo

- Definiciones previas:
 - La potencia de cómputo media de un nodo “i” en un sistema heterogéneo (P_i), se define como la cantidad de trabajo ejecutado en una unidad de tiempo en dicho procesador.
 - La potencia de cómputo total de un sistema heterogéneo compuesto por N procesadores ($P_T(N)$) se define como la suma de la potencia de cómputo de cada uno de los procesadores que componen el sistema.
- La potencia de cómputo total depende del número de procesadores y de la potencia de cómputo de cada procesador.

16 de mayo de 2006

Jose Luis Bosque

15

Modelo de Isoeficiencia Heterogéneo

- La eficiencia de un sistema paralelo heterogéneo se define como el cociente entre el tiempo de respuesta óptimo para resolver el problema concreto en ese sistema, dividido entre el tiempo de respuesta real, obtenido al realizar la ejecución.
- Tiempo óptimo se obtiene si se reparte la carga de trabajo de forma equilibrada y proporcional a la potencia de cómputo de cada nodo.

$$EF = \frac{W}{T_R \cdot P_T(N)}$$

16 de mayo de 2006

Jose Luis Bosque

16

Modelo de Isoeficiencia Heterogéneo

- Esta definición presenta una serie de ventajas:
 - Esta definición no necesita utilizar ningún procesador de referencia, como otras propuestas en la bibliografía.
 - Si es posible estimar la magnitud de la carga de trabajo para tamaños de problema grandes, no es necesario obtener los tiempos de ejecución sobre un único procesador, para estos tamaños.

16 de mayo de 2006

Jose Luis Bosque

17

Modelo de Isoeficiencia Heterogéneo

- Sea un sistema paralelo $S(N, P_T, W)$ y un sistema escalado $S'(N', P'_T, W')$ con $P'_T > P_T$. Se dice que S es un sistema escalable al pasar de S a S' si se puede mantener la eficiencia constante para algún tamaño de trabajo W' .
- La función de isoeficiencia queda definida por la siguiente expresión:

$$W' = W \cdot \frac{P'_T \cdot T'_R}{P_T \cdot T_R}$$

16 de mayo de 2006

Jose Luis Bosque

18

Modelo de Isoeficiencia Heterogéneo

- Para minimizar el tiempo de respuesta todos los procesadores deben finalizar su trabajo al mismo tiempo, lo que requiere una distribución equilibrada de la carga de trabajo.
- Dependiendo de la distribución de la carga de trabajo y de sus sobrecargas asociadas una combinación algoritmo-arquitectura mostrará diferentes grados de escalabilidad.

16 de mayo de 2006

Jose Luis Bosque

19

Modelo de Isoeficiencia Heterogéneo

- Lema 1: Sea un sistema heterogéneo $S(N, P_T, W)$ y un sistema escalado, $S'(N', P'_T, W')$, con carga de trabajo infinitamente divisible. Si la carga de trabajo se reparte de forma equilibrada y proporcional a la potencia de cómputo de cada uno de los nodos y los tiempos de sobrecarga son constantes para todos los procesadores entonces los tiempos de respuesta de todos los nodos son iguales y el tiempo de respuesta del sistema se puede expresar como

$$T_1 = T_2 = \dots = T_N = T_R = \frac{W}{P_T(N)} + C_0$$

16 de mayo de 2006

Jose Luis Bosque

20

Modelo de Isoeficiencia Heterogéneo

- Teorema 1: Dado un sistema heterogéneo $S (N, P_T, W)$ y un sistema escalado, $S' (N', P'_T, W')$, con carga de trabajo infinitamente divisible. Si la carga de trabajo se reparte de forma equilibrada y proporcional a la potencia de cómputo de cada uno de los nodos y los tiempos de sobrecarga son constantes para todos los procesadores entonces ambos sistemas tienen la misma eficiencia (y por lo tanto son escalables) si y sólo si:

$$W' = W \cdot \frac{P'_T \cdot C'_0(N')}{P_T \cdot C_0(N)}$$

16 de mayo de 2006

Jose Luis Bosque

21

Modelo de Isoeficiencia Heterogéneo

- Lema 2: Para las mismas condiciones que en el Lema 1, si los tiempos de sobrecarga pueden expresarse como un término constante para todos los procesadores, más un segundo término proporcional a la potencia de cómputo de cada nodo, entonces los tiempos de respuesta de todos los nodos son iguales y el tiempo de respuesta del sistema se puede expresar como

$$T_1 = T_2 = \dots = T_N = T_R = \frac{W + C_0 \cdot P_T + C_1 \cdot \sum_{i=1}^N P_i^2}{P_T}$$

16 de mayo de 2006

Jose Luis Bosque

22

Modelo de Isoeficiencia Heterogéneo

- Teorema 2: Para las mismas condiciones que en el caso anterior, si los tiempos de sobrecarga pueden expresarse como un término constante para todos los procesadores, más un segundo término proporcional a la potencia de cómputo de cada nodo, entonces la función de isoeficiencia se puede expresar como:

$$W' = W \frac{C_0 P_T + C_1 \sum_{i=1}^{N'} (P'_i)^2}{C_0 P_T + C_1 \sum_{i=1}^N (P_i)^2}$$

16 de mayo de 2006

Jose Luis Bosque

23

Modelo de Isoeficiencia Heterogéneo

- Lema 3: Para las mismas condiciones que en el Lema 1, si los tiempos de sobrecarga pueden expresarse como un término constante para todos los procesadores, más un segundo término proporcional a la potencia de cómputo de cada nodo, entonces los tiempos de respuesta de todos los nodos son iguales y el tiempo de respuesta del sistema se puede expresar como:

$$T_1 = T_2 = \dots = T_N = T_R = \frac{W + C_0 \cdot P_T + C_1 \cdot \frac{W}{P_T} \cdot \sum_{i=1}^N P_i^2}{P_T}$$

16 de mayo de 2006

Jose Luis Bosque

24

Modelo de Isoeficiencia Heterogéneo

- Teorema 3: Para las mismas condiciones que en los casos anteriores, si los tiempos de sobrecarga se pueden expresar como la suma de un término constante para todos los procesadores, más un segundo término proporcional a la carga de trabajo de cada nodo, entonces la función de isoeficiencia heterogénea se puede expresar como:

$$W' = W \frac{C_0 P_T}{C_0 P_T + C_1 \frac{W}{P_T} \sum_{i=1}^N P_i^2 - C_1 \frac{W}{P_T} \sum_{i=1}^{N'} P_i^{12}}$$

16 de mayo de 2006

Jose Luis Bosque

25

Validación del Modelo

- El tiempo de sobrecarga tiene tres componentes:
 - Un primer término constante para todos los nodos.
 - Un segundo término, directamente proporcional al número de nodos en el sistema.
 - Un tercer término directamente proporcional al número de paquetes de trabajo procesados por cada nodo.

$$T_o = C_o + C_1 \cdot N + C_2 \cdot P_i \cdot \frac{W}{P_T}$$

16 de mayo de 2006

Jose Luis Bosque

26

Validación del Modelo

- Dado un sistema heterogéneo $S (N, P_T, W)$, un sistema escalado $S' (N', P'_T, W')$, y una carga de trabajo infinitamente divisible. Si la carga de trabajo se distribuye de forma equilibrada y proporcional a la potencia de cómputo de cada nodo y los tiempos de sobrecarga de la aplicación son los descritos anteriormente entonces la función de isoeficiencia se puede expresar como:

$$W' = W \frac{C_0 P'_T + C_1 N' P'_T}{C_0 P_T + C_1 N P_T + C_2 \frac{W}{P_T} \sum_{i=1}^N P_i^2 - C_2 \frac{W}{P_T} \sum_{i=1}^{N'} (P'_i)^2}$$

16 de mayo de 2006

Jose Luis Bosque

27

Validación del Modelo

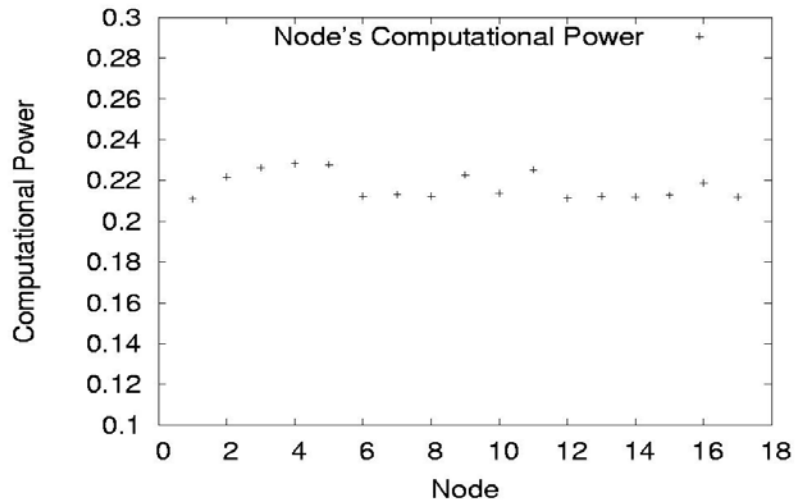
- Objetivos de los experimentos realizados:
 - Verificar experimentalmente que la definición de eficiencia heterogénea puede aplicarse tanto a sistemas homogéneos como a heterogéneos, e incluso se convierte en la definición de eficiencia clásica cuando se aplica a sistemas homogéneos.
 - Verificar que las funciones de isoeficiencia homogéneas y heterogéneas se pueden aplicar para modelar y predecir la escalabilidad de sistemas basados en clusters.
 - Mostrar que no es posible aplicar las definiciones de escalabilidad e isoeficiencia clásicas para clusters heterogéneos.

16 de mayo de 2006

Jose Luis Bosque

28

Validación del Modelo

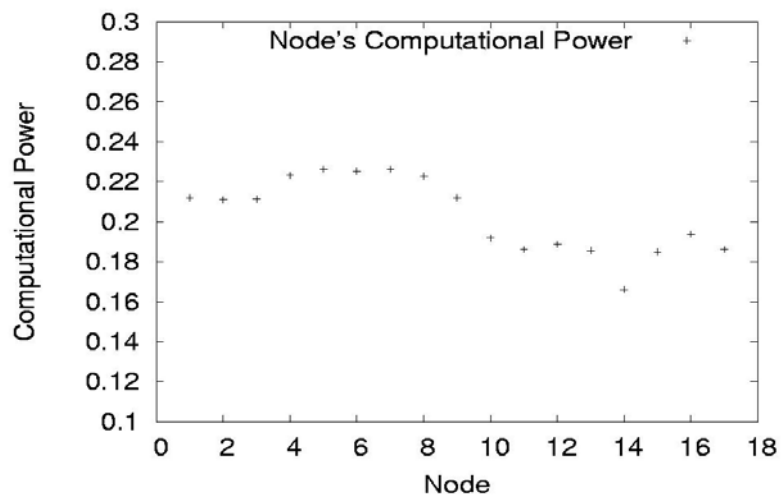


16 de mayo de 2006

Jose Luis Bosque

29

Validación del Modelo



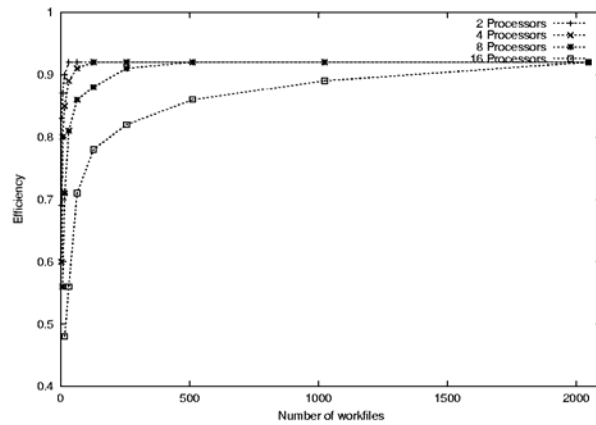
16 de mayo de 2006

Jose Luis Bosque

30

Validación del Modelo

Eficiencia clásica en un sistema homogéneo



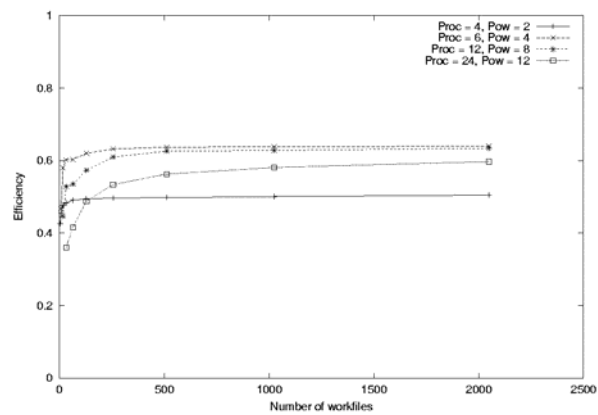
16 de mayo de 2006

Jose Luis Bosque

31

Validación del Modelo

Eficiencia clásica en un sistema heterogéneo



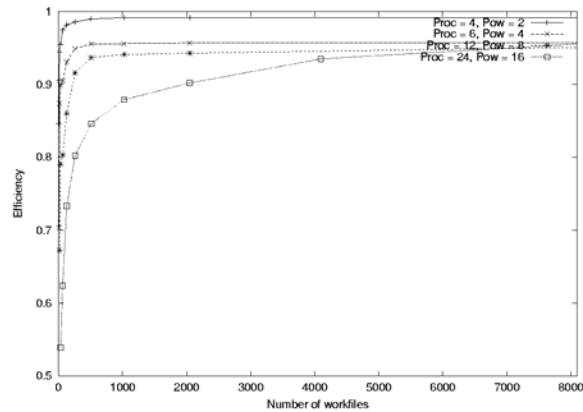
16 de mayo de 2006

Jose Luis Bosque

32

Validación del Modelo

Eficiencia propuesta en un sistema heterogéneo



16 de mayo de 2006

Jose Luis Bosque

33

Validación del Modelo

Cluster Homogéneo

Nodes	Files	Total Computational Power	Het. Eff	Class. Eff.	Estimated Workload
2	4	0.448	0.965	0.953	
4	16	0.904	0.949	0.969	16
8	64	1.765	0.95	0.947	61
16	256	3.483	0.952	0.937	250

16 de mayo de 2006

Jose Luis Bosque

34

Validación del Modelo

Cluster Heterogéneo

Files	Total Computational Power	Het. Eff	Class. Eff.	Estimated Workload
8	0.397	0.951	0.953	
32	0.797	0.953	0.834	32
128	1.598	0.954	0.837	129
512	3.241	0.953	0.848	526