Incluso la pequeña empresa puede necesitar el procesamiento en paralelo

Antonio Gómiz

Junio, 2007

Resumen

En este ensayo, dirigido a cualquiera interesado en el cálculo científico y tecnológico, se presenta una justificación sencilla del procesamiento en paralelo.

Introducción

La competición en mercados cada vez más abiertos obliga a las empresas fabricantes de cualquier tipo de objeto, desde juguetes a equipos industriales, a mejorar continuamente sus diseños y sistemas de producción. Para ello se puede recurrir a un buen número de técnicas como el control estadístico de la calidad, los diseños experimentales y la simulación por ordenador.

Estos cálculos son cada vez más complejos e involucran un mayor número de datos, haciéndose necesario con más frecuencia el uso de ordenadores con arquitecturas avanzadas y de aplicaciones informáticas capaces de aprovechar sus prestaciones.

Uno de los problemas que puede sacar más partido de la computación en paralelo es la solución de sistemas de ecuaciones lineales. La primera sección coloca este problema en un marco conveniente de modo que se pueda apreciar su importancia.

La segunda sección muestra los resultados de un problema típico en Ingeniería producidos por un entorno de diseño que usa el método del elemento finito, una clase de algoritmos que requiere la solución de sistemas con muchos miles de ecuaciones.

1. Solución de problemas científicos y tecnológicos con ordenador

Un enfoque bastante adecuado para una gran cantidad de problemas consta de cinco etapas (véase por ejemplo la introducción de [Gustafson], la de [Jung], capítulos 7 y 8 del volumen 2 de [Foumeny], parte E de [Kreyszig]):

1. Definición del problema.
2. Selección de los modelos matemáticos.
3. Selección de los métodos numéricos.
4. Implementación en el ordenador.
5. Validación de los resultados.

En general, éste es un proceso iterativo en el que desde cualquier etapa puede ser necesario regresar a una anterior debido a una gran cantidad de circunstancias.

1.1. Definición del problema

En esta etapa se decide qué se requiere resolver tan detalladamente como sea posible. Esto no siempre se consigue debido a dificultades de diverso tipo, por ejemplo en una etapa posterior puede hacerse evidente que se ha recabado un número insuficiente de datos experimentales.

1.2. Selección de modelos matemáticos

En esta etapa se trata de encontrar una formulación concreta del proceso físico, químico, biológico, etc. que recoja los aspectos cuantitativos esenciales de dicho proceso. Por ejemplo, para obtener un modelo para el espacio recorrido en un intervalo de tiempo [t_0, t] por una partícula con movimiento uniformemente acelerado basta usar la definición de aceleración, la segunda derivada del espacio con respecto del tiempo, e igualarla a una constante a:

$$\frac{d^2e}{dt^2} = a \tag{1}$$

En este ejemplo, la ecuación (1) se puede integrar inmediatamente en dicho intervalo dando lugar al modelo cuadrático equivalente:

$$e(t) = \frac{1}{2}at^2 + v_0 t + e_0 \tag{2}$$

donde $v_0 = -at_0$ es la velocidad inicial y $e_0 = -\,at_0^2/2 - v_0 t_0$ es el espacio inicial.

1.3. Selección de métodos numéricos

Una vez fijado el modelo y sus parámetros, se deciden los algoritmos más adecuados para evaluar las incógnitas que se establecieron como objetivo en la primera etapa.

Siguiendo con el ejemplo, supongamos que los datos disponibles consisten en que la partícula es un automóvil parado a las $t_0 = 0$ horas ($e_0 = 0$) y que a las cuatro horas ha recorrido 400 km. Supongamos además que se requiere calcular la posición del automóvil en las tres primeras horas ($t = 1, 2, 3$). Para ello basta una ecuación de primer grado para calcular la aceleración a, de modo que todos los parámetros del modelo (2) son conocidos y entonces se puede calcular el espacio recorrido con una simple sustitución del tiempo t por su valor, obteniendo $e(1) = 25$, $e(2) = 100$ y $e(3) = 225$.

Desafortunadamente, no siempre es posible obtener un modelo que permita el cálculo directo como el de la ecuación (2) y en general resulta necesario operar con modelos basados en ecuaciones diferenciales, ecuación (1). El enfoque típico consiste en descomponer el espacio en el que se desarrolla el fenómeno físico, químico, etc. en partes más pequeñas de modo que se puedan aproximar razonablemente bien las funciones del modelo en esas partes.

En este caso nuestro espacio es el intervalo temporal de cuatro horas, que podemos dividir en subintervalos iguales de una hora. Llamemos e_i a los espacios recorridos por el automóvil en la i-ésima hora, con $i = 0, 1, 2, 3, 4$.

La velocidad, derivada primera del espacio con respecto del tiempo, se puede aproximar mediante una fórmula de diferencias hacia adelante (ver por ejemplo [Burden], pág. 163 y siguientes, [Engeln], página 355). En el caso de la velocidad inicial v_0 será

$$v_0 \cong \frac{-3e_0 + 4e_1 - e_2}{2h} \tag{3}$$

donde h = 1 hora es la longitud de cada subintervalo temporal.

Análogamente, la aceleración se puede aproximar con una fórmula de diferencias centradas. La aproximación para la aceleración a_i en la *i*-ésima hora, con *i* = 1, 2, 3, será

$$a_i \cong \frac{e_{i-1} - 2e_i + e_{i+1}}{h^2} \tag{4}$$

Usando los datos del problema

$$a_1 = a_2 = a_3 = a, \qquad e_0 = 0, \qquad e_4 = 400, \qquad v_0 = 0, \qquad h = 1 \tag{5}$$

empleando las ecuaciones (3) y (4) y simplificando resulta el siguiente sistema de ecuaciones lineales

$$\begin{aligned} 4e_1 - \ \ e_2 \qquad\quad &= 0 \\ -3e_1 + 3e_2 - \ \ e_3 &= 0 \\ e_1 - 3e_2 + 3e_3 &= 400 \end{aligned} \tag{6}$$

cuya solución es $e_1 = 25,\ e_2 = 100,\ e_3 = 225$. En este ejemplo, la solución es exacta porque las aproximaciones (3) y (4) son de orden 2 para un polinomio cuadrático, ecuación (2), y por tanto son exactas. Este no es el caso general por lo que puede ser necesario realizar una división del espacio suficientemente fina como para garantizar la validez de las aproximaciones. Esto suele dar lugar a sistemas con muchos miles de ecuaciones.

De esta forma llegamos al punto crucial en la argumentación de este ensayo: es muy frecuente que un cálculo complejo se reduzca a resolver un sistema enorme de ecuaciones lineales o una secuencia de ellos (ver por ejemplo [Ames], [Burden], [Chandrupatla], [Versteeg]). De ahí la importancia de acelerar los algoritmos para estos sistemas, en particular recurriendo a la computación paralela (ver por ejemplo [Dongarra]).

1.4. Implementación en el ordenador

En esta etapa se desarrolla o elige el software que realizará el cálculo. Este software debe tener en cuenta un número considerable de detalles relativos a la arquitectura del ordenador que ejecutará las operaciones. En el caso del procesamiento en paralelo es esencial prestar atención a cómo está organizada la memoria del ordenador, su gestión, el número y tipo de procesadores, la topología de interconexiones, etc. (ver por ejemplo [Dongarra]).

1.5. Validación de los resultados

Los resultados producidos por el software deben considerarse con espíritu crítico y no aceptarse directamente. Dos aspectos son fundamentales:

1. La calidad de los resultados. Evidentemente, no sirve de nada obtenerlos a toda velocidad si los valores son erróneos. La verificación de los resultados se realiza frecuentemente mediante baterías de pruebas para las que se conoce la solución exacta. Existen casos en los que se llevan a cabo millones de pruebas, de modo que el cálculo en paralelo puede ser muy ventajoso.
2. El tiempo necesario para calcular valores suficientemente aproximados y físicamente realistas. Tampoco sirve de mucho acabar el cálculo de un equipo industrial después del plazo de entrega.

2. Un ejemplo típico

El método del elemento finito (ver [Zienkiewicz]) es el estándar de facto para la solución de multitud de problemas en Ingeniería. De modo análogo al ejemplo anterior, este método supone que el problema original, normalmente muy complejo, puede descomponerse un muchos otros muy fáciles de resolver y que el ensamblaje de estos resultados intermedios permitirá obtener una solución suficientemente aproximada al problema inicial.

La figura 1 muestra un envase a presión en fase de diseño para el que se requiere la distribución de tensiones y así prevenir la posibilidad de fallo del material en algún punto. Esta clase de equipo industrial y el problema que plantea no sólo concierne a las grandes empresas, de ahí el título de este ensayo.

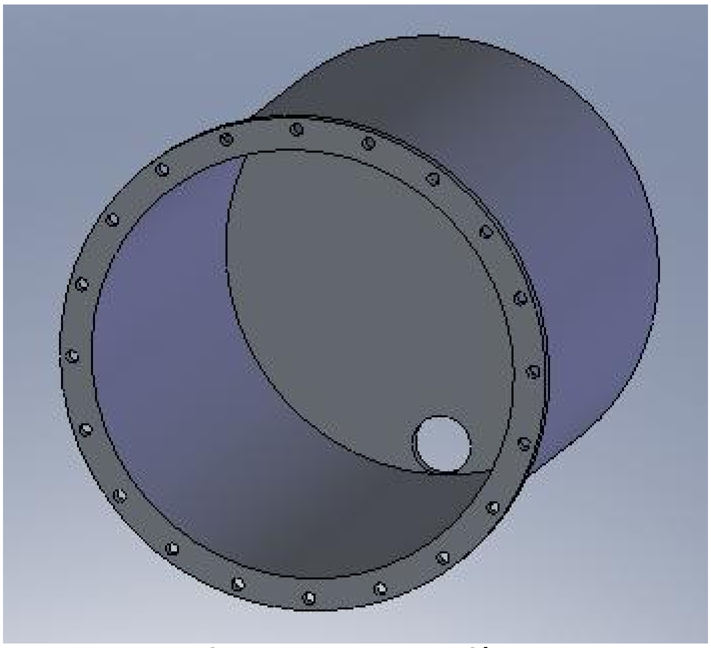

Figura 1. Envase a presión.

La figura 2 muestra la división del espacio a analizar, en este caso el cuerpo del envase a presión.

La figura 3 refleja el momento en el que se resuelve el sistema con más de 65000 ecuaciones: los dos núcleos del procesador trabajan en paralelo, tardando sólo unos segundos. El equipo incorpora un procesador Intel® Pentium® Dual Core a 3.2 GHz y 4 GB de memoria. Esta velocidad de respuesta para problemas no excesivamente complejos permite un diseño casi interactivo, con la posibilidad de realizar muchas simulaciones en un solo día de trabajo y producir diseños muy optimizados.

La figura 4 muestra los resultados, indicando el punto de máxima tensión que es unas seis veces menor que el límite elástico del material, un acero inoxidable, para las condiciones de presión fijadas. El diseño se considera entonces seguro.

Conclusiones

Las prestaciones del procesamiento en paralelo le hacen muy atractivo para un buen número de tareas de cálculo en Ingeniería y otros ámbitos científicos y tecnológicos. Su importancia se incrementa con la complejidad de los problemas a resolver.

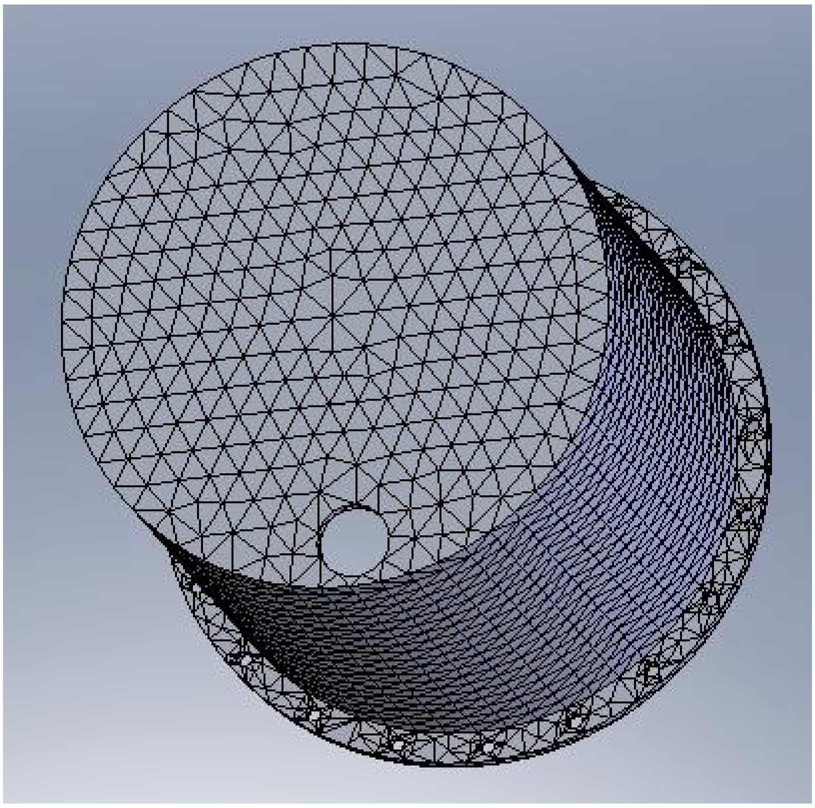

Figura 2. Mallado.

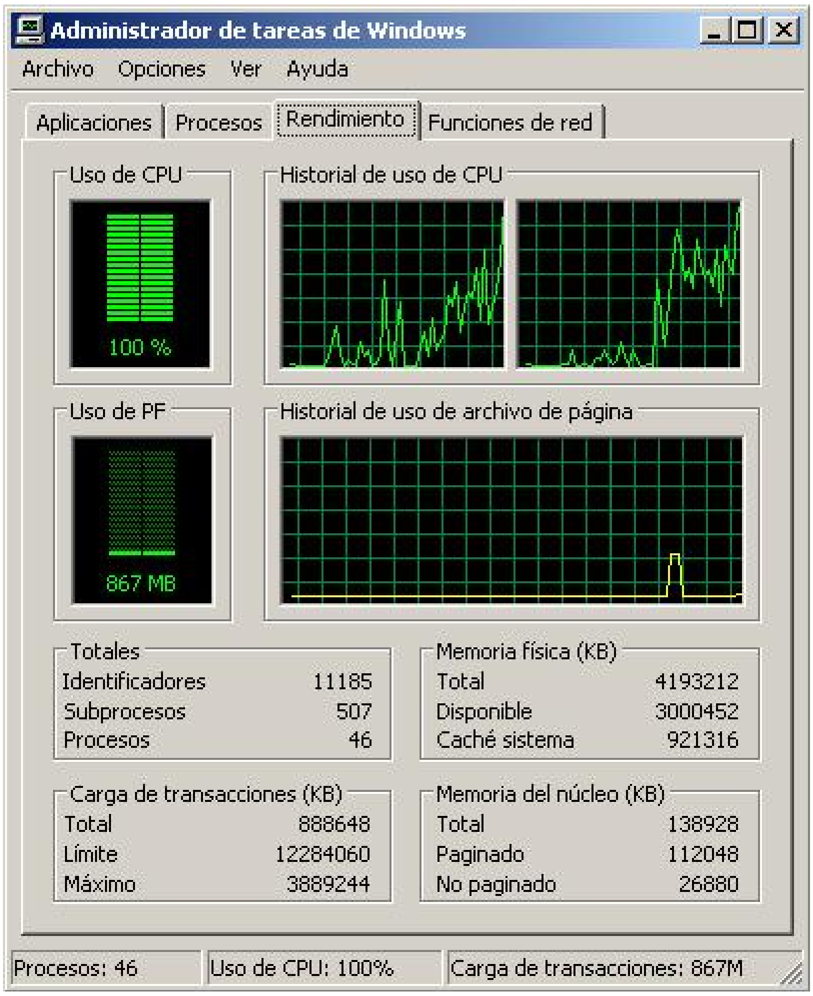

Figura 3. Momento en el que se resuelve el sistema de ecuaciones. Los dos núcleos del procesador trabajan al 100%.

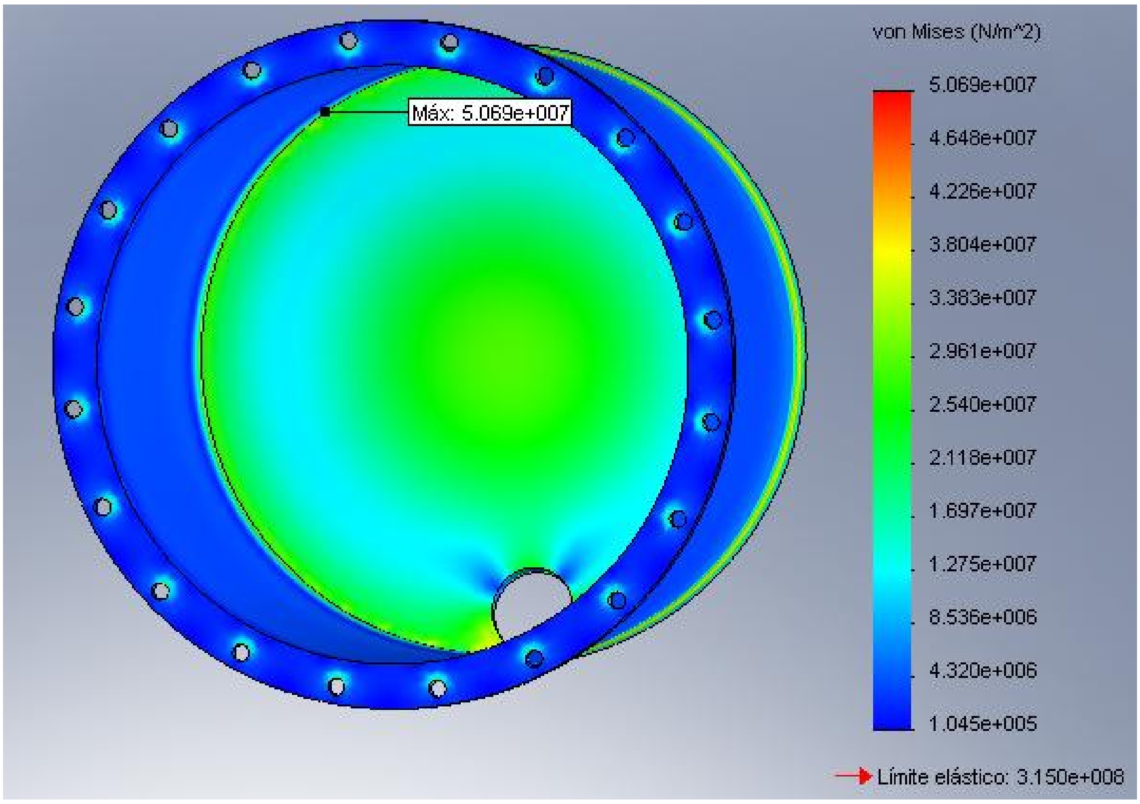

Figura 4. Resultados (tensiones de von Mises).

Referencias

[Ames] W. F. Ames. *Numerical Methods for Partial Differential Equations*. Academic Press, 1992.

[Burden] R. L. Burden, J. D. Faires. *Análisis Numérico. Segunda Edición*. Grupo Editorial Iberoamérica, 1996.

[Chandrupatla] T. R. Chandrupatla, A. D. Belegundu. *Introducción al Estudio del Elemento Finito en Ingeniería. Segunda Edición*. Prentice Hall, 1999.

[Dongarra] J. J. Dongarra, I. S. Duff, D. C. Sorensen, H. A. van der Vorst. *Numerical Linear Algebra for High-Performance Computers*. Society for Industrial and Applied Mathematics, 1998.

[Engeln] G. Engeln-Müllges, F. Uhlig. *Numerical Algorithms with Fortran*. Springer, 1996.

[Foumeny] E. A. Foumeny, P. J. Heggs. *Heat Exchange Engineering*. Ellis Horwood ltd., Chichester, 1991.

[Gustafson] G. B. Gustafson, C. H. Wilcox. *Analytical and Computational Methods of Advanced Engineering Mathematics*. Springer-Verlag, New York, 1998.

[Jung] M. Jung, U. Langer. *Finite-Elemente-Methode. Eine Einführung für Ingenieurstudenten*. Notas de clase, 1995.

[Kreyszig] E. Kreyszig. *Advanced Engineering Mathematics*. John Wiley & Sons, 1998.

[Versteeg] H. K. Versteeg, W. Malalasekera. *An Introduction to Computational Fluid Dynamics*. Addison Wesley Longman, Harlow, 1995

[Zienkiewicz] O. C. Zienkiewicz, R. L. Taylor. *The Finite Element Method. Fifth Edition*. Butterworth-Heinemann, 2000.