COMPUTACIÓN PARALELA EN LAS GEOCIENCIAS

José Gómez-Valdés
CICESE, Km 107 Carret. Tijuana-Ensenada, Ensenada, B.C., México, A.P. 2732.

RESUMEN

Esta nota trata sobre conceptos básicos de computación paralela. Se hace una revisión de la evolución de las computadoras paralelas. Hoy día los sistemas a la vanguardía son: Cray T3D e IBM SP. Por su bajo costo y su escalabilidad, las redes de estaciones de trabajo en un ambiente paralelo son una opción factible para hacer computación paralela en la actualidad. La computación paralela se ha estado empleando sistemáticamente en las Geociencias desde principios de la presente década. En México no se esta promoviendo ni la investigación ni la enseñanza de esta nueva tecnología como se esta haciendo en la mayoría de los paises europeos y en los Estados Unidos, aunque ya se empieza a trabajar en este campo en algunas instituciones del país.

INTRODUCCIÓN

La computación paralela (CP) es un tipo de cálculo a la vanguardia en la simulación y predicción númerica en las geociencias. Consiste en usar más de una unidad de procesamiento (nodo) en los trabajos de computación, donde para algunos casos un nodo está formado por un **CPU** (Control Processing Unit), una unidad de memoria y una unidad de entrada-salida, de forma tal que, en ausencia de comunicación entre nodos, el rendimiento se incremente linealmente con el número de procesadores. La idea de paralelismo no es nueva. Se puede decir que desde la aparición del *Homo habilis*, la especie humana ha reunido fuerzas y habilidades para resolver problemas que para un sólo miembro de la especie resultaría imposible de resolver. Paralelismo por computadoras, sin embargo, aunque sus origenes están en el diseño de las primeras computadoras, es una práctica que sólo ha recibido aceptación ámplia en las comunidades académicas de los Estados Unidos y de Europa a partir de los 90's, en especial, grupos en las ciencias del espacio y de las geociencias han incursionado éxitosamente en este campo. Aquí hablaremos sobre los conceptos básicos de la CP, y de sus aplicaciones. Veremos la urgencia de promover ésta nueva tecnología entre los jóvenes mexicanos.

COMPUTADORAS PARALELAS Y COMPUTADORES EN PARALELO

Las supercomputadoras se definen como las máquinas más rápidas y más caras. Pueden hacer entre 100 y 200 **Megaflops** (Millones de operaciones de punto flotante) y costar decenas de millones de dólares. Actualmente las computadoras paralelas, de unos cuantos procesadores, compiten en rapidez con las supercomputadoras de un sólo procesador, y tienen como ventajas: escalabilidad, que significa la posibilidad de adherir más nodos, y bajos costos (la IBM SP puede alcanzar más de 100 **Megaflops** de rendimiento con 16 procesadores y cuesta del orden de un millón de dólares). En cambio, como no hay una manera automática o mecánica de convertir programas secuenciales a programas paralelos se tiene que aprender a programar en paralelo, lo cual implica consumo de tiempo y de energía. La CP usa dos, diez, cien o aun miles de unidades de procesamiento para resolver simultáneamente diferentes partes del mismo problema. Generalmente, los nodos se comunican entre sí para sincronizar las tareas asignadas a cada uno de ellos. De acuerdo a la operación de la unidad de control, las computadoras paralelas pueden ser **SIMD** (Single Instruction stream, Multiple Data streams) o **MIMD** (Multiple Instruction streams, Multiple Data streams). La CM-2, Connection Machine, que cuenta con 65,536 procesadores, es un ejemplo de una máquina **SIMD**, es decir simultáneamente se ejecuta la misma instrucción en todos los procesadores pero con diferentes datos. Por otro lado, la Intel iPSC/2 cuyo número máximo de procesadores es de 128, es un ejemplo de una máquina **MIMD**, es decir cada procesador puede ejecutar instrucciones diferentes en datos diferentes (más detalles sobre clasificaciones de computadoras se pueden encontrar en Flynn 1966). De acuerdo al modo de manejar el flujo de información a través de la memoria, las computadoras paralelas pueden ser de memoria compartida o de memoria distribuída (Hennessy y Patterson, 1990). Las computadoras Cray con más de un procesador son máquinas de memoria compartida y las Intel de memoria distribuida. Debido al bajo costo de los procesadores, y a su tendencia a la baja, una de las arquitecturas más prometedoras es la combinación **MIMD**-memoria distribuida. Hoy día los ejemplos más destacados de esta arquitectura son: IBM SP y Cray T3D, los **cluster** y las redes de estaciones de trabajo en un ambiente de pasar mensajes, e.g., **MPI** (Message-Passing Intarface) o **PVM** (Parallel Virtual Machine); éstos dos últimos ejemplos constituyen computadores en paralelo.

El rendimiento y la eficiencia de las computadoras paralelas están aumentando a un ritmo acelerado. Por ejemplo, en 1989 los sistemas hipércubos Intel ofrecieron el computador Intel iPSC/2 con procesadores 80386/80387, con rendimiento sostenido de 1-2 **Megaflops** y 4 Megabytes (10^6 bytes) de memoria **RAM** (Random Access Memory) por nodo.

En 1991 estos sistemas cambiaron a Intel iPSC/860 que usa procesadores i860, con rendimiento sostenido de 3-4 **Megaflops** y 8 Megabytes de memoria **RAM** por nodo. La computadora Intel Paragón, la más reciente (1993) de las **MIMD** de memoria distribuida de la clase Intel, cuenta con procesadores i860 XP cuyo rendimiento téorico es de 75 **Megaflops** y 16 Megabytes o más de memoria **RAM** por nodo. Los nodos se encuentran ubicados en una malla, por lo que es relativamente fácil incrementar su número, ya que sólo se requiere insertar la nueva tarjeta de nodos. El rendimiento téorico de la Paragon del **San Diego Supercomputer Center** es de 40 Gigaflops (10^9) con 7.4 Gigabytes de memoria **RAM**. El uso de computadoras paralelas de muchos procesadores (masivas) permitiría manejar información a 10^{12} operaciones de punto flotante por segundo (Teraflops) y tener acceso a Terabytes de memoria a fines de la presente década (Hennessy y Patterson, 1990).

PROGRAMACIÓN EN EL MODELO DE PASAR MENSAJES

Debido a que cada aplicación científica tiene su propio grado de paralelismo, no hay reglas estándar para construir esquemas paralelos óptimos. Sin embargo, a la fecha, los principios básicos de programación paralela se han identificado muy bien, véase por ejemplo Ragsdale (1991). Estos principios son: partición, mapeo, estrategias de comunicación y granularidad.

La partición consiste en dividir un problema en partes que puedan trabajar en paralelo, tratando de colocar las partes en provecho de la escalabilidad y el balanceo de la información. Aquí escalabilidad significa hacer la partición independiente del número de nodos, con lo cual se podrá adherir más nodos sin tener que reprogamar. Al balancear la información se trata de hacer una distribución pareja del trabajo entre los nodos, con lo cual se mantendrá a todos los nodos ocupados y terminarán al mismo tiempo. Habrá ciclos pérdidos si alguno de los nodos tiene que esperar a otro para terminar.

Después de la partición, las partes deben ser mapeadas a los nodos para su ejecución, cuidando de colocar componentes del problema que intercambian información tan cerca como sea posible. El mapeo depende de la arquitectura del computador. Sin embargo, la mayoría de las computadoras parelalas pueden ser vistas por el programador como si estuvieran completamente conectados, si así se desea.

En el diseño de las estrategias de comunicación, es necesario tomar en cuenta la comunicación entre nodos de forma tal que los nodos gasten poco tiempo en comunicarse. Generalmente, el mecanismo de sincronización sirve tanto para el intercambio de información de datos como para lograr la coincidencia en tiempo de los procesos. Ya que para medir los costos totales de comunicación se toma en cuenta el costo de la inicialización del proceso, mensajes largos son preferidos sobre muchos mensajes cortos.

Se puede también estimar la granularidad, la cual es la razón entre tiempo de computación y tiempo de comunicación, del programa paralelo. Programas con granularidad alta gastan menos tiempo en comunicación que programas con granularidad baja. La meta del programador debe ser minimizar la razón entre comunicación y computación.

El rendimiento y la eficiencia de las computadoras paralelas, de agregados de computadores en paralelo, y de programas paralelos, se pueden cuantificar con ayuda de rutinas disponibles en esos sistemas y con principios cuantitativos de diseño de software y hardware.

RENDIMIENTO Y EFICIENCIA

Alto rendimiento en CP implica al menos dos cosas: reducción del tiempo total de computación e incremento del trabajo de computación en un determinado tiempo, donde una unidad de tiempo de computación es el tiempo requerido para realizar una operación de punto flotante. Y una unidad de tiempo de comunicación se define como el lapso al pasar un valor de punto flotante entre dos procesadores, si el tiempo total de comunicación es mucho, se reduce el rendimiento del sistema.

Dos conceptos básicos en el análisis del rendimiento de los sistemas son el incremento de la rapidez de procesamiento y la eficiencia. Empecemos por el primero, el incremento de la rapidez se define como

$$S = \frac{T_{\text{sec}}}{T_{par}}, \tag{1}$$

donde T_{sec} es el tiempo total de computación de un procesador y T_{par} el tiempo total de computación de P procesadores en paralelo. Así, la eficiencia paralela se define como

$$\varepsilon = \frac{S}{P}, \tag{2}$$

donde S es el incremento de la rapidez y P el número de procesadores. El incremento de la rapidez es siempre menor que P (o, la $\varepsilon<1$).

Los costos totales de comunicación se pueden medir mediante la relación

$$Fc = \frac{1}{\varepsilon} - 1. \tag{3}$$

Supongamos que hacemos un algoritmo paralelo de un programa secuencial que se encuentra en plena producción, y encontramos que la eficiencia es muy pobre, digamos $\varepsilon<0.5$, esto implicaría que la eficiencia del algoritmo paralelo esta siendo afectada por alguno de, o todos, los factores siguientes (Fox et al.,1989):

- El problema no es paralelizable.
- El incremento de la rapidez esta limitado por la rapidez del nodo más lento (imbalance).
- Los costos de comunicación son muy altos.

Entonces se debe de trabajar más duro en el algoritmo paralelo antes de abandonarlo, proponiendo otras estrategias de paralelización.

Se puede obtener una estimación *a priori* del grado de paralelismo del problema a paralelizar con la **Ley de Amdahl**

$$A = 1.0 / (f_s + (1 - f_s) / P), \qquad (4)$$

donde f_s es la fracción de la parte no paralela del programa secuencial y P el número de procesadores.

Computación paralela en las Geociencias

Con la CP se pueden explorar problemas oceanográficos que requieran de un alto rendimiento computacional. Debido a las propiedades físicas del mar, el modelado océanico demanda alto rendimiento. La turbulencia, por ejemplo, se desarrolla con más facilidad en la capa de mezcla superficial de los océanos. Así, las escalas espaciales para resolver los procesos de transporte y de mezcla van del espesor de la capa de mezcla superficial, O(10^2 m), a las dimensiones de las cuencas océanicas, O(10^7 m), y las escalas temporales que son importantes en la circulación océanica van de los minutos de la fluctuación turbulenta a los cientos de años que toma la ventilación de las aguas del fondo.

Lee et. al. (1989) han usado con éxito una de las computadoras paralelas pioneras, la Intel iPSC/2 de 64 nodos, para modelar la propagación del sonido en los océanos, alcanzando un incremento en la rapidez de procesamiento de Cray X-MP/(iPSC/2)=2.8, usando 64 procesadores. Gómez-Valdés y Wang (1995), al hacer un algoritmo paralelo del modelo tridimensional de las ecuaciones primitivas de circulación costera de Wang (1985), han probado que la CP ofrece enormes ventajas sobre la computación con un sólo procesador, al lograr en una computadora Intel iPSC/2 de 32 nodos, un incremento en la rapidez de procesamiento de 3.9 usando 4 procesadores y de 29.3 usando 32 procesadores.

La modelación del transporte de solutos en medios heterogéneos es otro ejemplo de la necesidad de usar computación de alto rendimiento. Incluir en modelos de partículas del O (10^6) entidades sólo fue posible en supercomputadoras de 1993. Se prevé que para fines de la presente década se podrán usar del O (10^8) partículas en dichos modelos. Barry (1990) revisa el uso de supercomputadoras en la solución numérica de la ecuación tridimensional de advección-dispersión, cuya dinámica tiene también múltiples escalas temporales y espaciales, y recomienda a los hidrologistas tomar ventajas de la computación paralela.

Prácticamente en todas las ramas de las geociencias hay problemas que requieren de más de 200 **Megaflops** de rendimiento para su correcta solución. Ahí, casi todos los modelos en producción se han paralelizado o se ha intentado paralelizarlos, e.g., ya hay en el mercado las versiones paralelizadas del programa ENSOLV (Euler-Navier-Stokes flow SOLVer), del programa SMART que sirve para calcular convección en fluidos, entre otros. Además, Johnson et al. (1995) presentan un programa paralelo para la animación de una onda 3D en la zona de subducción, obteniendo un incremento en la rapidez de 0.5 en una KSR1. Puster y Jordan (1995) usan el modelo de pasar mensajes en el diseño de un programa computacional para estudiar métodos inversos en sismología. En el hilo de desarrollo de algoritmos en computadores en paralelo, se puede ver a Aranda-Álvarez et al. (1994) y Gómez-Valdés y Soto-Amador (1995), quienes implantan **PVM** en una red local heterogénea de estaciones de trabajo, prueban el sistema con modelos numéricos de acuíferos simples en elementos finitos, logrando reducir el tiempo de computación usando hasta 5 computadoras SUN.

Computación Paralela en México

Si bien uno puede encontrar programas de posgrado en los cuales se ofrece CP, por ejemplo, en el IPN, UNAM y Tecnológico de Monterrey, y en Ensenada se imparten cursos a nivel de maestría (CICESE) y de licenciatura (UABC) desde 1995, y se tienen proyectos de investigación aplicada y básica de esta disciplina desde 1994, cuando se implanta en CICESE **PVM** en la red principal, a la fecha este campo ha recibido poca atención en México. A nuestro conocimiento, no hay todavía en México un programa de métodos numéricos en los posgrados de las ciencias de la tierra en el que se discutan las versiones paralelas de los algoritmos clásicos, como se esta haciendo en universidades del extranjero.

En los centros de supercómputo de los Estados Unidos se cuenta con las computadoras paralelas más avanzadas, se promueve la CP a través de apoyos de tiempo de **CPU** y se hace labor de enseñanza y difusión. En Europa, en casi todos los paises se esta practicando la CP, destacando el Instituto Tecnológico de Munich (de donde se pueden obtener gratis los más modernos simuladores de computadoras) de Alemania y la University of Edinburgh de Inglaterra.

Como vimos las ofertas llamativas de la CP son la reducción de costos y la escalabilidad. Si bien la primera de ellas debería de ser suficiente para promover su uso en México, la segunda es también relevante, pues se puede incrementar el poder de cómputo de una institución académica en función del presupuesto disponible y de las necesidades reales de cómputo que, como sabemos, son rubros muy dinámicos.

La comunidad académica mexicana, particularmente la de geociencias en donde hay muchas aplicaciones, tiene ahora la oportunidad de practicar computación paralela en un tiempo

en el que ésta ha dejado de ser un paradigma en etapa de experimentación, ya que ha evolucionado de ese estadio a ser un modelo práctico. Se puede, como lo hemos señalado en esta nota, integrar computadores en paralelo de las arquitecturas más variadas, y con los avances recientes de la internet, disponer, en el mismo ambiente, de todas las computadoras (**UNIX, AIX o HP-UX**) del país. Los jóvenes mexicanos, sobre todo, deben de recibir de sus profesores los conocimientos nuevos al mismo tiempo que los jóvenes de otros paises.

AGRADECIMIENTOS

El autor desea agradecer la revisión y los comentarios al trabajo por parte de Raymundo Vega, Jesús Favela y Manuel Figueroa. Este trabajo fue financiado por CONACyT: referencia 225008-5-3523T.

REFERENCIAS

Aranda-Álvarez, J.J. Gómez-Valdés y E. Gómez-Treviño. 1994. Computación Paralela en la Solución Númerica de la Ecuación Nolíneal de Advección-Dispersión. GEOS, 14:5.

Barry, D.A. 1985. Supercomputers and Their Use in Modeling Subsurface Solute Transport. Reviews of Geophysics. 28, 277-295.

Flynn, M.J. 1966. Very high-speed computing systems. In Proceedings. IEEE. 54, 54-12.

Fox, G.M. Johnson, G. Lyzenga, S. Otto, J. Salmon and D. Walker. 1989. Solving Problems on Concurrent Processors. Volume I: General Techniques and Regular Problems. Prentice Hall, New Jersey. 592 pp.

Gómez-Valdés, J. y R. Soto-Amador. 1995. Applications of PVM in Groundwater Modeling. Second UNAM-Cray Supercomputing Conference: Numerical Simulations in the Environmental and Earth Sciences, Mexico City, 21-24 june.

Gómez-Valdés, J. and D-P. Wang. 1995. Massively Parallel Processing in Coastal Ocean Circulation Model. Journal of Scientific Computing. 10, 305-323.

Hennessy, J. L. and D. A. Patterson. 1990. Computer Architecture A Quantitative Approach. Morgan Kaufmann Publishers, Inc. 594 pp.

Johnson, O.G., G. Chen and H-W. Zhou. 1995. 3D Seismic Study of Subduction Zones Using Parallel Wave Animation. (abstract), Eos Trans. AGU., 76 (17), April.

Puster, P. and T.H. Jordan. 1995. Parallel Programming in Global Seismology (abstract), Eos Trans. AGU., 76:17, April.

Ragsdale, S. 1991. Parallel Programming. McGraw-Hill, Inc. New York. 123 pp.

Wang, D-P. 1985. Numerical study of gravity curents in a channel. Journal of Physical Oceanography. 15, 299-305.

GLOSARIO

CPU (Central Processsing Unit): parte interna del computador que procesa la información. Incluye: el control de los circuitos, la unidad aritmética, la unidad de memoria y la unidad de control. El tiempo de **CPU** se puede medir con el comando ***UNIXTIME***.

Megaflops (Millones de operaciones de punto flotante): alternativa para medir el rendimiento de un computador, se define como el cociente entre el número de operaciones de punto flotante en el programa y su tiempo de ejecución multiplicado por 10^6. Ésta medida depende de la máquina y del programa.

SIMD (Single Instruction stream, Multiple Data streams): en un sólo ciclo opera una instrucción y simultáneamente varios flujos de información. En una computadora paralela SIMD la sincronización es automática.

MIMD (Multiple Instruction streams, Multiple Data streams): en cada ciclo pueden operar varias instrucciones y varios flujos de información a la vez. En computadoras paralelas **MIMD** la sincronización se debe de programar. Puede notarse que la arquitectura **MIMD** contiene a la arquitectura **SIMD**.

Cluster: un arreglo de computadoras coordinadas entre sí para ejecutar tareas, con los soportes necesarios y suficientes para formar un agregado eficiente sin llegar a contar con todo el hardware que se requiere para formar, por ejemplo, una computadora paralela. Los **clusters** que hay en el mercado son los servidores de la Digital VAX y agregados de estaciones de trabajo. (En la red de la Universidad de California de Los Angeles hay varios **cluster** de 4 IBM RS/6000).

MPI (Message Passing Interface): es un paquete de rutinas para hacer computación paralela en computadores en red o en computadoras paralelas. Se pretende que sea el estándar del modelo de pasar mensajes.

PVM (Parallel Virtual Machine): es un paquete de rutinas para hacer computación paralela en computadores en red o en computadoras paralelas. Es el paquete de dominio público más popular actualmente.

RAM (Random-Access Memory): memoria principal del computador en la cual se encuentran las instruciones de los programas y donde los datos del programa son almacenados; está conectada directamente al **CPU**.

San Diego Supercomputer Center: uno de los cuatro centros de supercómputo fundados por la National Science Foundation (NSF) de los Estados Unidos que usa la comunidad académica de ese país. A través de estos centros la NSF promueve el uso de nuevas tecnologías.

Ley de Amdahl: establece que el incremento del rendimiento que se obtiene al usar una mejora esta acotado por la fracción de tiempo en que se utiliza la mejora. Se aplica tanto a computadoras secuenciales como a paralelas.

UNIX: sistema operativo de la AT&T Bell Laboratories escrito en C, con el que se pueden hacer múltiples tareas y varios usuarios pueden trabajar simultáneamente en un sólo computador. Desde supercomputadoras hasta PC usan este sistema operativo.

AIX (Advanced Interactive Executive): sistema operativo de la IBM basado en el sistema operativo **UNIX**. Las estaciones de trabajo RS/6000 y las computadoras paralelas SP usan este sistema operativo.

HP-UX: sistema operativo de la HP basado en el sistema operativo **UNIX**. Las estaciones de trabajo HP 9000 y los **clusters** HP usan este sistema operativo.