

Aplicaciones de la supercomputación

Félix García Merayo
Dr. y Ldo. en Informática
Profesor Titular de Universidad-UPM
Vicepresidente de ACTA

Jeremiah P. Ostriker, astrofísico en la Universidad de Princeton, ha dicho que “no hace aún muchos años, la cosmología era más teología que ciencia”. Y ha añadido, “los supercomputadores han ayudado a cambiar esa opinión”.

En resumen, la investigación científica y el desarrollo técnico están soportados, hoy día, en un alto grado, por técnicas de simulación computacional o simulación numérica; con ello es posible la experimentación, el contraste de hipótesis y, como consecuencia, la verificación de las teorías, todo lo cual evita el experimento físico directo que consume mucho tiempo e incluso, en muchos casos, es imposible de realizar, por razones económicas, políticas o medioambientales.

1. INTRODUCCIÓN Y OBJETIVOS

La simulación es una de las técnicas más importantes utilizadas en el ámbito de la computación. Puede verse la huella de esa técnica en muchas áreas del campo científico como la previsión del cambio climático, monitorización de los reactores nucleares o la mejora de la eficiencia de los productos de automoción. La simulación por ordenador juega hoy día un papel indiscutible al facilitar soluciones a la gran mayoría de los retos y problemas técnicos y científicos planteados.

Los profesionales que utilizan la simulación como herramienta básica disponen de una potencia de cálculo muy notable debido a los avances en la tecnología del hardware de los computadores. Todavía más si se tiene en cuenta la utilización de las redes y del procesamiento masivamente paralelo.

La meta de hoy, en un mercado que evoluciona muy rápidamente, es la de crear una fabricación ágil de todo tipo de productos y con una necesidad limitada de prototipos. Ello conlleva la realización de simulaciones grandes, rápidas y complejas.

Estas ideas y tendencias no son nuevas. Llevamos una década manejando la misma filosofía para justificar con ello el procesamiento paralelo mediante supercomputadores como única solución a los problemas técnicos actuales, solución basada en técnicas de simulación. Estas son las opiniones de profesores, investigadores y técnicos relevantes como W. J. Camp y S. J. Plimpton de Sandia Nacional Lab, en Albuquerque o de A. S. Jacobson y A. L. Berkin de Jet Propulsion Lab, en California. [Communications of the ACM, Vol 37, Num 4].

Desde las primeras etapas del proceso electrónico de la información, siempre se han utilizado las máqui-

nas más potentes para resolver las aplicaciones científicas y las relativas a la ingeniería de cualquier tipo. Y esta tendencia todavía persiste. Nos centraremos en el área de desarrollo de problemas científicos basados en arquitecturas MIMD.

Una reseña de las aplicaciones más significativas que se ejecutan o simulan en arquitecturas reales incluiría:

- Circulación atmosférica global o previsión del tiempo en 3D
- Dinámica de fluidos
- Cromodinámica cuántica
- Flujo de la sangre en el corazón
- Dinámica molecular
- Evolución de las galaxias

La publicación *Computer* de IEEE, febrero 2004, en un artículo relativo a las tendencias de la industria, indicaba una aplicación que será operativa en el 2005.

La compañía IBM ha introducido una pequeña máquina prototipo (del tamaño de una lavadora) de altas prestaciones a la que le ha dado el nombre de *Blue Gene/L*, máquina con un bajo consumo de energía y, por tanto, económica y con un bajo nivel de disipación calorífica, características que los supercomputadores no siempre satisfacen: muchos de ellos poseen una alta cantidad de procesadores muy rápidos y próximos entre sí generadores de altas temperaturas y que deben trabajar por períodos limitados de tiempo.

IBM ha balanceado en este *super*, rendimiento y consumo de energía mediante la utilización de un gran número de procesadores todos ellos de baja potencia. En este sentido, *Blue Gene/L* alcanza los 1,4 Tflops, situándolo así en el puesto 73 de la lista Top500 de los superordenadores más rápidos del mundo. Esta relación se compila los meses de junio y noviembre de cada año interviniendo investigadores de las universidades de Mannheim y de Tennessee junto con el Lawrence Berkeley National Laboratory.

Blue Gene/L utiliza 512 unidades de proceso de la serie PowerPC 440 a 700 MHz pero con una frecuencia real de solo 500 MHz, precisamente para, de este modo, minimizar la disipación de calor. Cada CPU contiene dos PowerPC, cada una de ellas dotada con dos unidades de aritmética de coma flotante.

Este supercomputador es un prototipo para un sistema gubernamental estadounidense, el *ASCI Purple*, Accelerated Strategic Computing Initiative, conocido

recientemente con el nombre de Advanced Simulation and Computing Program, que podríamos traducir por sistema de Simulación Avanzada y Programa de Computación.

El prototipo, una vez consolidado, será utilizado por la institución Lawrence Livermore National Laboratory en el 2005 para modelar el desdoblamiento de proteínas humanas en la investigación médica.

2. ETAPAS QUE CONDUCEN A LA PARALELIZACIÓN

Veamos de cerca las etapas del proceso de creación de programas paralelos a partir de un algoritmo o programa secuencial ya existente. En relación con lo dicho, en muchas ocasiones, un buen algoritmo secuencial de un problema contribuye y apoya una paralelización posterior sencilla; en otras, se hace necesario el diseño y creación de un algoritmo nuevo.

El trabajo de paralelización [Supercomputación, Manual Formativo 30] conlleva: la identificación de aquellas partes de un algoritmo que pueden procesarse a la vez, en paralelo; determinar cómo distribuir el trabajo y los datos del problema entre los elementos de proceso de que se disponga; gestionar el acceso a esos datos, comunicar éstos entre sí y sincronizar todo el trabajo de forma global. Cuando hablamos de trabajo, nos referimos a la computación, al acceso a los datos y a la entrada y salida de información.

Las etapas de un proceso para la creación de un programa paralelo las puede realizar el programador o el propio software de sistemas, enlace entre el programador y la arquitectura de la máquina. Este software se compone de: compilador, programa en código y sistema operativo. La meta de la investigación actual se centra en el diseño y producción de compiladores paralelizantes y en los lenguajes paralelos, aunque el ideal sería encontrar una paralelización automática e inteligente con la mínima intervención del hombre.

Antes de examinar con más detalle el proceso de paralelización de un algoritmo, de un programa, vamos a establecer formalmente algunos conceptos informáticos que consideramos previos y necesarios.

- **Tarea.** Parte cualquiera de un trabajo realizado por un programa de ordenador. Es la unidad más pequeña que puede explotar, en concurrencia con otras, un programa paralelo. O lo que es lo

mismo: una tarea concreta la ejecuta un solo procesador del sistema de computación y, como consecuencia, la concurrencia entre procesadores se explota mediante la ejecución de varias tareas al mismo tiempo.

Si la cantidad de trabajo que la tarea lleva a cabo es pequeña, se dice que la tarea es de **grano fino**, *fine grained*; en otro caso, será de **grano grueso**, *coarse grained*.

- **Proceso.** Entidad abstracta que realiza tareas. Un programa paralelo se compone de varios procesos cooperativos, cada uno de los cuales ejecuta un subconjunto de tareas de un programa. Las tareas se identifican con los procesos mediante un mecanismo de asignación tarea/s-proceso.

En lenguaje coloquial, un proceso es un segmento de código que se procesa en un determinado procesador de forma concurrente con otros procesos.

- **Procesador.** Recurso físico donde se ejecutan procesos. Los programas paralelos se escriben en función de los procesos necesarios y no de los procesadores físicos de que se dispone; los procesos se aplican o asignan posteriormente a los procesadores disponibles del sistema de computación.
- **Red de interconexión.** Red física capaz de conectar entre sí los procesadores del sistema de computación paralela.

Dados estos conceptos, veamos cuáles son las etapas de creación de un programa paralelo, que se ilustran en la Figura 1:

1. **Descomposición** del algoritmo en tareas.
2. **Asignación** de tareas a los procesos.

3. **Coordinación** de: acceso a los datos, comunicación y sincronización entre procesos.
4. **Aplicación** o vinculación de los procesos con los procesadores.

Veamos un poco más de cerca cada una de esas etapas.

Descomposición. Separar en partes la computación total, de forma que las tareas estén disponibles dinámicamente a medida que el programa se va ejecutando. En un instante determinado del tiempo, cuanto mayor sea el número de tareas disponibles en la ejecución mayor será el número de procesos y, por tanto, de procesadores que entrarán en acción de forma efectiva.

En general, dado un problema de un cierto tamaño y su correspondiente descomposición en tareas, puede construirse un **perfil de concurrencia**, *concurrency profile*, indicador de cuántas operaciones o tareas de la aplicación se realizan a la vez, en un instante determinado, lo que nos dará el aprovechamiento del sistema total en cada momento, es decir, el equilibrio de ocupación de cada uno y todos los procesadores.

Asignación. Por asignación quiere significarse el mecanismo mediante el cual las tareas se distribuyen entre los procesos, es decir, qué proceso es responsable de qué parte de la computación. El objetivo en esta etapa debe ser equilibrar la carga entre procesos para reducir así la comunicación entre los mismos, operación que suele ser la más cara, sobre todo cuando procesos diferentes se ejecutan en procesadores también diferentes. Ese hecho se conoce con el nombre de **equilibrio de cargas**, *load balancing*. La carga de trabajo a equilibrar, comprende: la computación propiamente dicha, la entrada/salida y el acceso o comunicación a/entre los datos.

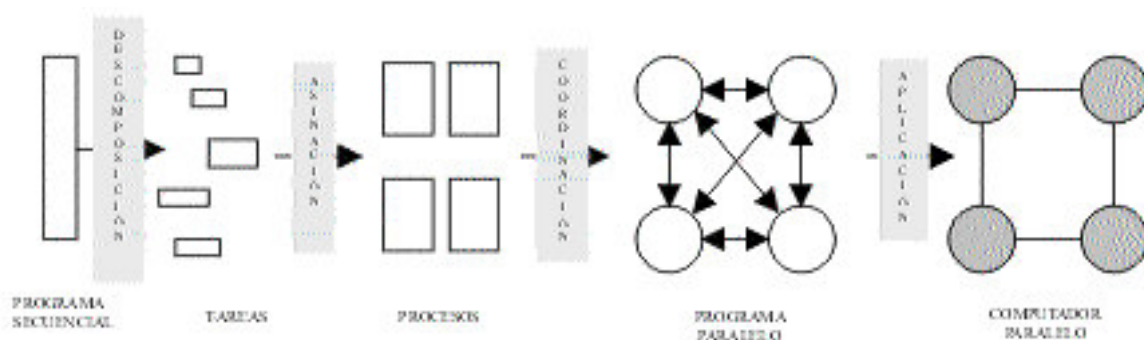


Figura 1. Etapas de la paralelización de una aplicación

Coordinación. Esta etapa se refiere, de forma particular, al momento en el que se escriben los programas mediante un lenguaje de programación concreto, teniendo en cuenta: cómo organizar las estructuras de los datos, cómo planificar las tareas asignadas a un proceso para explotar así la *localidad* de los datos y cómo organizar y expresar la comunicación y sincronización entre procesos. Por tanto, los dos agentes que juegan aquí el papel más importante son la arquitectura hardware del sistema paralelo y el modelo de programación adoptado.

Aplicación. Los procesos cooperativos resultantes de las tres etapas anteriores constituyen lo que se denomina un **programa paralelo** completo. Este programa puede controlar por sí mismo la **vinculación** o asociación entre procesos y procesadores o bien dejar esa misión para que sea realizada por el sistema operativo encargado de llevar a cabo una ejecución paralela de la aplicación.

2. Proceso simbólico. Incluyen los sistemas de base de datos, las aplicaciones de Inteligencia Artificial, IA, como, por ejemplo, los sistemas expertos o los sistemas de percepción y planificación.

Estas dos categorías se distinguen, entre otras cosas, por la relación entre los cálculos a ejecutar y el movimiento de los datos entre todos los procesadores que posea el sistema de computación.

3. CARACTERÍSTICAS GENERALES DE LAS APLICACIONES CIENTÍFICAS

En general, las aplicaciones científicas y, como consecuencia, también las pertenecientes al campo de la ingeniería, poseen un número de características comunes que se hacen patentes en los programas de ordenador con los que se da solución eficiente a unas y otras. Enumeraremos, en primer lugar, las **características de los cálculos** para después analizar las **características de los datos** con los que se realizan tales cálculos.

Aunque estas características son muy generales, se encuentran sin embargo en prácticamente todas las aplicaciones importantes ejecutadas sobre supercomputadores.

Características de los cálculos

- Una estructura básica y frecuente en los programas de ordenador es el **bucle** o **iteración**: ejecución repetida de una serie de instrucciones de máquina en las que sólo cambian algunas variables o índices, estos últimos representados por números enteros, manteniéndose constante el resto del cuerpo del programa.
- El **volumen** de los **cálculos** es **muy considerable**, es decir, los modelos que el usuario se propone ejecutar con el computador sobrepasan con frecuencia las posibilidades de los ordenadores más potentes, ya sea por consideraciones de memoria insuficiente como por velocidad no apropiada.
- Las dos características precedentes corresponden con mucha frecuencia a la resolución de **sistemas de ecuaciones** de todo tipo mediante el uso de algoritmos iterativos, abundantes en la bibliografía del Cálculo Numérico.

ETAPA	DEPENDENCIA ARQUITECTURA HW	OBJETIVO
DESCOMPOSICIÓN	NO	CONSEGUIR SUFICIENTE CONCURRENCIA
ASIGNACIÓN	NO	EQUILIBRAR CARGA DE TRABAJO REDUCIR INTERCOMUNICACIÓN
COORDINACIÓN	SI	REDUCIR COMUNICACIÓN ENTRE DATOS Y ENTRE PROCESADORES. PLANIFICACIÓN DE TAREAS
APLICACIÓN	SI	ASOCIAR PROCESOS RELACIONADOS CON EL MISMO PROCESADOR

Tabla 1. Etapas de la paralelización, recursos y objetivos

Trataremos ahora de identificar y describir en cierto modo las grandes familias de aplicaciones que se ejecutan en los supercomputadores de hoy día sin aspirar a conseguir una relación exhaustiva de las mismas ni de pormenorizar ninguna de ellas.

La mayor parte de las aplicaciones que pueden ser paralelizadas, es decir, que contienen un alto grado de paralelismo para poder ser subdivididas y distribuidas entre los distintos elementos de proceso de un supercomputador, pertenecen a dos categorías generales:

1. Proceso numérico. Nos referimos a las aplicaciones científicas y de ingeniería las cuales presentan grandes problemas de computación: cálculo de partículas en el plasma; dinámica de fluidos, como el modelo de previsión del tiempo o el diseño de aviones; diseño asistido por ordenador, CAD, y otras muchas.

Características de los datos de partida

- Los datos que intervienen en este tipo de aplicaciones son datos del *campo algebraico real* y por tanto, desde el punto de vista informático, vendrán expresados en **coma flotante**. Ello conlleva siempre la necesidad de manejar grandes precisiones, es decir, gran cantidad de cifras significativas en los mismos.
- En todo modelo, la **cantidad de datos** a manejar es **considerable** y, por tanto, se hace necesario disponer de memorias de gran capacidad.

LA NOTACIÓN CIENTÍFICA

Cuando se emplean cantidades del *campo real* en un sistema de computación, se opera con ellas siguiendo una técnica especial que se denomina **aritmética de la coma flotante**. Esos números reales se representan entonces, utilizando una notación conocida con el nombre de **notación científica** o también de **coma flotante**. Por ejemplo, la distancia de la Tierra al Sol es de unos 149500000000 metros, cantidad muy grande y con gran abundancia de ceros para ser manejada con sencillez. Haciendo uso de la notación científica en coma flotante, podríamos escribirla así: $1.495,0 \times 10^8$. De igual manera, un número extremadamente pequeño como sería 0,00000004321, podríamos escribirlo como $0,4321 \times 10^{-8}$. En estas representaciones, la *coma decimal* se desplaza o *flota* de acuerdo al valor del exponente de 10. Así, la distancia del primer ejemplo podría escribirse de varias formas:

$$1495,0 \times 10^8 = 149,5 \times 10^9 = 1,495 \times 10^{11} = \dots$$

La parte fraccionaria que precede a la potencia de 10, se denomina **mantisa**.

Con esta notación se facilita la *percepción* del número, es decir, de su orden de magnitud, siempre conservando el grado de *precisión* que no es más que el total de cifras significativas, cifras distintas de cero, que posee la mantisa. La precisión de las mantisas del ejemplo último es siempre de *cuatro cifras significativas*.

En las aplicaciones científicas o de ingeniería el uso de la aritmética de coma flotante es algo sin lo cual no sería posible la computación de las mismas.

4. ESTRUCTURA DE LAS GRANDES APLICACIONES CIENTÍFICAS

Nos vamos a apoyar en un ejemplo: la dinámica de fluidos aplicada a la aerodinámica. La dinámica de fluidos no es propiamente una aplicación sino más bien una técnica general empleada en un cierto número de aplicaciones: la aeronáutica, el automóvil, la meteorología, el estudio del flujo del aire a través de una turbina, el movimiento de un líquido a través de una red. Por tanto, una gran parte de las aplicaciones procesadas en los superordenadores, se basan en la teoría general de la dinámica de fluidos.

El ejemplo de la aerodinámica que vamos a analizar a continuación permite evidenciar la contribución del superordenador y los problemas derivados. De todas formas, la simulación digital que se aplica, muchas veces no excluye el que también haya que ejecutar una simulación analógica ya que ambas técnicas, en este caso, son complementarias.

El problema consiste en un análisis de los flujos aerodinámicos en torno a un objeto en movimiento. El objetivo es determinar en cada uno de los puntos del medio, y en cada instante del tiempo, el valor de los parámetros representativos como, la presión, la temperatura, etc. El algoritmo matemático que se aplica es el correspondiente a las ecuaciones de Navier-Stokes, las cuales expresan la conservación de la masa, del momento y de la energía y constituyen un sistema de ecuaciones diferenciales en derivadas parciales. El total de éstas sobrepasa las sesenta.

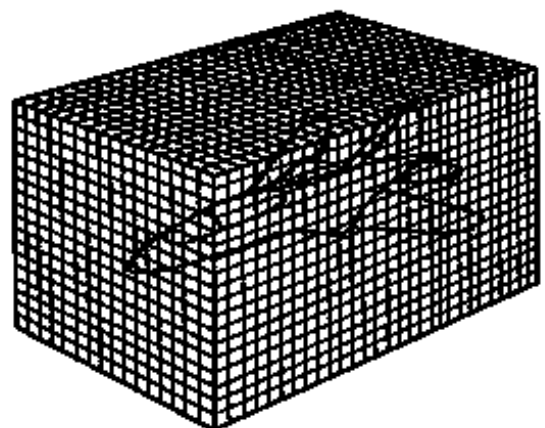


Figura 2. Discretización del espacio entorno a un avión

El problema con el que nos encontramos es que no es posible resolver tal sistema de ecuaciones de una

forma exacta; es necesario recurrir a métodos aproximados, es decir, utilizar métodos numéricos, conocidos globalmente como método de las *diferencias finitas*.

EL ALGORITMO

Las ecuaciones de Navier-Stokes aplicables al movimiento de un fluido viscoso son un caso particular de la ecuación del transporte de Boltzmann aplicable también al flujo de electrones en un semiconductor. Esta ecuación, en teoría de campos, se escribe de la forma vectorial siguiente:

$$\frac{d\vec{V}}{dt} = \frac{\partial \vec{V}}{\partial t} + (\vec{V} \cdot \nabla)\vec{V} = -\frac{1}{\rho} \nabla P + \vec{F} + \nu(\nabla \cdot \nabla)\vec{V} + \frac{\nu}{3} \nabla \nabla \cdot \vec{V}$$

siendo: (x, y, z) las coordenadas de posición de una partícula del fluido; $V(x, y, z, t)$ su velocidad correspondiente, de componentes respectivas, V_x, V_y, V_z ($= dx/dt$, etc.); t , la variable tiempo; ∇ es el operador gradiente; ρ la densidad, P la presión, \vec{F} el vector fuerza y ν la viscosidad.

La ecuación vectorial anterior equivale a tres ecuaciones escalares que se obtienen reemplazando sucesivamente el vector velocidad del fluido \vec{V} por sus respectivas componentes, V_x, V_y y V_z . Éstas, a su vez, conducen a la ecuación del movimiento o ecuación del *momento*, a dos ecuaciones de continuidad y a la ecuación del equilibrio de la energía, también conocida como *primera ley de la termodinámica* o *principio de la conservación de la energía*.

Existe otro método que lleva el nombre general de *elementos finitos*, muy empleado para el análisis de estructuras en la ingeniería civil. Es un método muy adecuado al estudio y simulación del medio continuo. El algoritmo matemático sobre el que descansa vuelve a ser un sistema de ecuaciones diferenciales en derivadas parciales no resoluble de forma exacta y si mediante métodos numéricos aproximados que conducen a ecuaciones algebraicas polinómicas. En lugar de discretizar el entorno en estudio por puntos, ese entorno se decompone en *elementos*. En la Figura 3 se muestra la diferencia entre ambos métodos. En (a) el entorno se ha discretizado mediante una malla a la que se aplicará las *diferencias finitas*; en (b) el medio se ha descompuesto en elementos superficiales que lo recubren totalmente y sobre ellos se aplicarán las técnicas de *elementos finitos*.

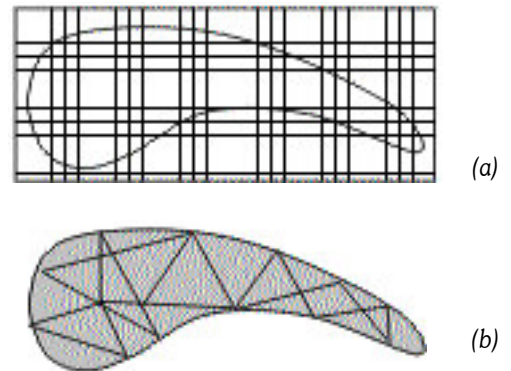


Figura 3. Rejillas para diferencias finitas (a) y para elementos finitos (b)

Volviendo a la aplicación de dinámica de fluidos, la técnica que se aplica es la de discretizar la región objeto del estudio sustituyéndola, como acaba de decirse, por una malla o rejilla de puntos, es decir, el número infinito de puntos que constituyen el medio continuo real se reemplaza por un número finito de puntos. En la Figura 2 se representa esta discretización en tres dimensiones. La discretización de un medio continuo permite, entonces, una simplificación de la formulación matemática del problema.

Esa operación transforma el sistema de ecuaciones en derivadas parciales citado en otro de ecuaciones algebraicas ordinarias, cuya resolución, como hemos dicho, conduce a la determinación de un valor aproximado de los parámetros del problema en cada uno de los puntos de la rejilla o malla. Ese sistema de ecuaciones algebraicas será, con toda seguridad, de un orden muy elevado, en correspondencia con el también elevado número de puntos de la rejilla donde se aplican las hipótesis. Los superordenadores permiten tratar esos sistemas, y sus algoritmos correspondientes, de forma muy rápida.

Hasta aquí habríamos aplicado técnicas y operaciones definidas más arriba: *descomposición* y *asignación*.

Cómo será la estructura de la programación o dicho de otra manera, una vez discretizado el medio, cómo procederá la computación.

1. Se adjudica un valor inicial, la *mejor estimación*, a los distintos parámetros.
2. Se aplica un algoritmo iterativo a esos valores iniciales. Tal algoritmo está constituido por una secuencia de operaciones cuyos datos de entrada son los valores iniciales y produce como salida resultados intermedios que son utilizados por la iteración siguiente.

3. Esos resultados intermedios constituyen una nueva entrada al algoritmo con el que se produce un nuevo conjunto de resultados intermedios que serán, sin duda, más cercanos a la solución definitiva que los intermedios previos.
4. Esta operación basada en iteraciones sucesivas se repite un elevado número de veces, hasta conseguir unos resultados que se consideren suficientemente próximos a la solución exacta. Técnicamente se dice que, hasta que el algoritmo converge.

Los pasos precedentes ponen de manifiesto la estructura de la programación con la que se consiguen resultados aproximados cada vez más próximos a los exactos y que constituyen la base de los algoritmos científicos: **el bucle** o iteración. Esta estructura repetitiva contenida en los programas de ordenador es relativamente sencilla de distribuir entre los procesadores de un sistema paralelo: cada procesador puede ejecutar una serie de iteraciones, al mismo tiempo que otro procesador ejecutará otras distintas.

En la Figura 4 se muestra el mismo móvil de la Figura 2 al que se le han añadido sólo algunos puntos de la rejilla en la que está inmerso. En la Tabla 2 se muestra de forma sintetizada la marcha del proceso descrito anteriormente: una vez adjudicado un valor inicial, *iteración 0*, al parámetro a investigar en cada punto de la malla, desde el 0 hasta el *m*, comienza la ejecución del algoritmo, produciendo resultados intermedios en cada iteración, *iteración 1, 2, 3, ..., n-1, n*. Puede observarse que, la tabla que acumula los cálculos, ©, se cumplimenta columna por columna, a medida que los cálculos se ejecutan.

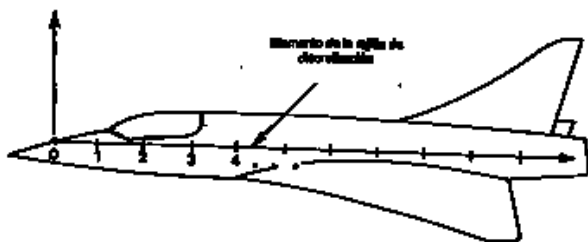


Figura 4. Simulación sobre un número discreto de valores

ITERACIÓN	PUNTOS					
	0	1	2	3	m
0	x	©	x	x	x
1	x	©	x	x	x
2	x	©	x	x	x
.
n-1	x	©	x	x	x
n	x	©	x	x	x

Tabla 2. Marcha de los cálculos sobre los puntos de una rejilla

De la anterior descripción podemos sacar las siguientes conclusiones:

- La representación discreta del medio o entorno conduce a resultados tanto más precisos cuanto más fina sea la discretización. O lo que es lo mismo: el número de puntos que forma la malla debe ser importante, es decir, elevado.
- Una simulación que tenga en cuenta la totalidad de la zona que rodea al móvil, debe ejecutarse sobre una rejilla de varias decenas de millones de puntos para obtener resultados con exactitud suficiente como para que puedan ser aprovechados por los ingenieros de diseño y producción.
- A cada punto de la rejilla se le asocia un cierto número de valores como, por ejemplo, los parámetros del problema, los elementos de la geometría, los resultados intermedios. En algunos problemas, el número de esos valores asciende a varias decenas.
- Si se considera una rejilla discreta de 10^7 puntos, con 20 valores por punto, el modelo de simulación considerado en este ejemplo comportaría 2×10^8 datos iniciales.
- Los procesos de discretización descritos sobrepasan en mucho las posibilidades de un computador convencional. De ahí, la necesidad de disponer de supercomputadores para llevar a cabo estas ejecuciones.

A la hora de entrar en la ejecución de los bucles habremos tenido que pensar y decidir en cómo coordinar esa ejecución que necesariamente será paralela y, como consecuencia, cuál será la vinculación o aplicación del total de bucles a realizar con el total de procesadores que constituyen el sistema paralelo de que se disponga.

5. UNA FAMILIA AMPLIA DE APLICACIONES

Hemos descrito un ejemplo de utilización de los supercomputadores en un problema de aerodinámica. Existen otras aplicaciones pertenecientes al campo de la dinámica de fluidos de las que vamos a describir algunas. Los usuarios son generalmente grandes compañías o bien organizaciones estatales que disponen de aplicaciones desarrolladas por ellas mismas. Estos son

los campos que someramente vamos a describir:

- Meteorología
- Termodinámica
- Oceanografía
- Industria del petróleo

La meteorología

Las aplicaciones en este campo comprenden el análisis de turbulencias y la simulación de la atmósfera, con resultados sobre la previsión del tiempo y la climatología en general. Para ello, la teoría de la dinámica de fluidos ya citada representa el agente principal sin olvidar otros fenómenos como, la evaporación, la condensación, el calor del sol, la física de la formación de las nubes, etcétera.

El usuario se ve forzado al empleo de los grandes ordenadores para conseguir modelos más y más completos donde poder integrar un gran número de parámetros.

El centro europeo de previsión meteorológica a medio plazo, siete días, está instalado en Gran Bretaña y la zona de estudio es el Noroeste de Europa, zona discretizada bajo una malla tridimensional que comprende quince capas horizontales cada una de 400 Km de lado. Los datos se toman diariamente y provienen de fuentes diferentes: estaciones terrestres, satélites meteorológicos, globos sonda, barcos y aviones de línea. Dada la diversidad de las fuentes de los datos, éstos, antes de entrar en el modelo, son depurados y homogeneizados.

El modelo simulado por la NASA tiene la geometría de una región comprendida entre dos esferas concéntricas: una correspondiente a la superficie de la corteza terrestre y la otra, la zona más alta de la atmósfera. Dado que las cantidades que intervienen en el modelo estarán tomadas y definidas en esa región tridimensional, y cada una de ellas variará con el tiempo, el resultado es que el problema a resolver es de cuatro dimensiones.

El modelo de simulación, utilizando los datos más recientes, se emplea para producir una nueva previsión más próxima al instante real y, como consecuencia, corrección de la inmediata anterior. El corazón del modelo son las leyes de la naturaleza expresadas, como ya se ha dicho, en ecuaciones diferenciales en derivadas parciales; el procedimiento consiste en calcular el estado del modelo en el instante $t + \Delta t$ a partir del t .

La previsión a largo plazo conlleva una modelización más fina empleando cálculos aún más voluminosos que el medio plazo. Actualmente, esta materia, estas aplicaciones, no tienen límite.

La termodinámica

La termodinámica va más allá incluso que la dinámica de fluidos, en el sentido que tiene en cuenta fenómenos suplementarios debidos a las reacciones químicas que se producen entre los propios fluidos que participan en el proceso en cuestión.

Actualmente existe un área en la que se dan esos procesos: la ingeniería del motor del automóvil que trata de optimizar, entre otras cosas, la forma y dimensiones de las cámaras de combustión para conseguir los motores más rentables posibles.

Otro área es la ingeniería nuclear. Para la refrigeración de los reactores nucleares se emplea un cierto número de fluidos, gases, líquidos y metales líquidos: en este entorno, las relaciones entre las transferencias de calor y el caudal de fluido son complejas dando lugar a cálculos de mucha intensidad.

La oceanografía

El estudio de los océanos presenta numerosos puntos de interés: desde la climatología hasta los diversos recursos del propio océano. Es muy importante la incidencia de estas masas líquidas sobre el clima, habida cuenta de la capacidad de agua almacenada, la transferencia de calor que se produce en el propio medio y con la atmósfera.

Para modelar el clima de la tierra es importante entender cómo la atmósfera interacciona con los océanos que ocupan las tres cuartas partes de la superficie terrestre. Desde este punto de vista, volvemos a encontrar aquí casi los mismos tipos de problemas que los ya descritos anteriormente con el estudio del modelo del tiempo con la diferencia que supone el hecho de que ahora los fenómenos se propagan más lentamente, horas en lugar de segundos, sobre distancias mucho más grandes. Además, estos estudios se hacen con un horizonte temporal de varios años y no para conocer la previsión del tiempo para el día siguiente.

Los modelos oceanográficos simulan el movimiento de las corrientes del agua del mar, corrientes que se desarrollan y evolucionan bajo la influencia de diversas

fuerzas físicas, incluyendo los efectos de la atmósfera, vientos y fricción con el fondo marino. En las cercanías de las costas se da además la fricción vertical, causa de corrientes circulares. El objetivo de estos modelos es diverso: desde simular esas corrientes circulares a lo largo del tiempo y llegar a entender sus interacciones con el flujo oceánico, hasta prever cuáles serán en el futuro los efectos de la transferencia de calor desde un punto del océano hacia los polos.

Una vez más la solución llega de la mano de la discretización de amplias zonas del océano mediante una malla o rejilla bidimensional en cada capa horizontal, añadiendo luego una sucesión de capas en profundidad. En cada punto de la rejilla intervienen parámetros como presión, velocidad, temperatura, etc. El modelo está basado numéricamente en la resolución de grandes sistemas de ecuaciones y la simulación conlleva un gasto de tiempo y potencia de máquina que no podrían conseguirse resultados eficientes sin la intervención de los superordenadores más potentes.

A mayor número de puntos por rejilla, mejor aproximación se obtendrá del modelo. Para el océano Atlántico, con una superficie de estudio de 2000x2000 Km², empleando una rejilla horizontal de 100x100 puntos, lo que implicaría una distancia entre puntos por cada dimensión, de 20 Km, no supondría una resolución demasiado fina, por lo que sería necesario utilizar una rejilla de más puntos. Para la cuarta dimensión, es decir, para el tiempo, intervalos cortos conducen a una simulación más aproximada. Así, para simular 5 años del movimiento del océano en intervalos de 8 horas, se necesitarían unos 5500 intervalos de tiempo. Todo lo anterior deja claro la necesidad de una computación de gran exactitud y de la utilización obligatoria de supercomputadores.

La ingeniería del petróleo

Se trata de dos aspectos que, aunque diferentes, están íntimamente ligados. Por una parte, la búsqueda de yacimientos o *exploración* y, por otra, las técnicas de optimización para obtener una *producción* máxima y económica. El coste del petróleo, bajo tierra o bajo el mar, justifica la inversión en la elaboración de modelos de simulación cuya resolución última se lleva a cabo en computadores de altas prestaciones.

A medida que se avanza en profundidad, se comprueba que la Tierra está formada por capas distintas: arena, rocas, minerales varios, etc. Cada una de ellas

posee propiedades diferentes, como ocurre, por ejemplo, con la densidad. Al producirse una fuente de energía en la superficie, como la originada por una explosión controlada, la onda de esa energía se propaga hacia el interior de las sucesivas capas terrestres en profundidad. Y cuando esa onda se encuentra con la capa frontera entre dos diferentes, una parte de la energía se refleja y otra se refracta hacia el interior de la Tierra. Figura 5.

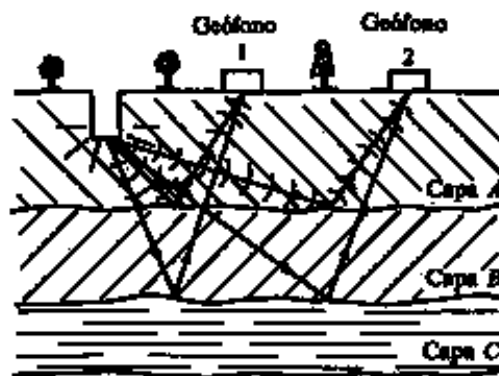


Figura 5. Recorridos de una onda expansiva

Mediante la colocación en la superficie de dispositivos detectores o geófonos, es posible registrar la energía reflejada. Empleando ecuaciones del tipo *espacio=velocidad x tiempo*, se llega a conocer la profundidad de la capa en cuestión. Después los geofísicos podrán levantar mapas con el perfil de cada capa detectada a partir de todos los datos sísmicos tomados.

Una vez recogidos los datos, en forma digital, de la zona en estudio, se procesarán en grandes computadores. Esta operativa se conoce con el nombre genérico de *tratamiento de la señal*. El objetivo de la exploración geofísica es la determinación de lugares con condiciones favorables para la acumulación de crudo. El proceso conlleva la ejecución de una enorme cantidad de cálculos matemáticos, entre los que se encuentran el filtrado y deconvolución para suprimir información parásita, transformaciones diversas, autocorrelación, etc. Se producen hasta cinco millones de cantidades en coma flotante, por cada kilómetro investigado. Téngase en cuenta que un registro típico de los datos tomados en campo, puede contener hasta 3.000 valores diferentes respecto a la variable tiempo, cada una correspondiente a los distintos lugares geográficos donde se hacen las mediciones.

En cuanto a la producción citada al principio, es necesario destacar que se trata tanto de la perforación como de la definición previa del propio depósito. Esta

última se lleva a cabo mediante un proceso de simulación con el que es posible estimar las reservas de crudo existentes así como las cantidades a extraer a lo largo del tiempo.

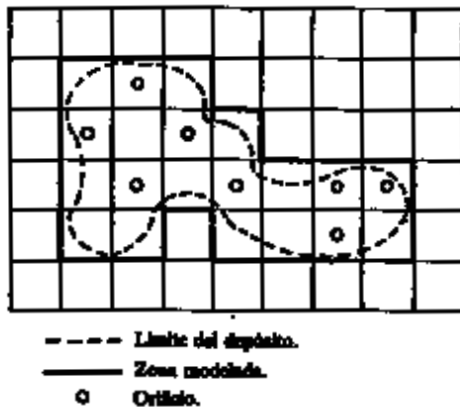


Figura 6. Rejilla 2D para el estudio de un depósito

Con el modelo de simulación se definen de una forma discreta, como se esquematiza en la Figura 6, el perfil plano que posee el depósito en estudio. A cada punto o nodo de los que constituyen la rejilla base se le asignan distintos parámetros como, permeabilidad, porosidad, grosor de la roca, presión, etc. Al fluido contenido en el depósito se le adjudican también otros parámetros como viscosidad, relación gas/crudo, densidad, etc. Y a la roca, su valor de compresibilidad.

Después de los consiguientes cálculos, relativos todos ellos a la mecánica de fluidos simplificados según la teoría de las diferencias finitas, será posible conocer y, en consecuencia, poder comparar las distintas políticas de explotación del depósito.

Realmente el modelo trabaja con esquemas tridimensionales lo que encarece el cálculo tanto en tiempo como en complejidad.

6. APLICACIONES BASADAS EN EL CONTINUO

Regresando al método de los *elementos finitos*, descrito someramente más arriba, añadir aquí que su campo de aplicación es muy amplio ya que los medios continuos constituyen una parte importante del entorno objeto de la simulación tanto en la industria como en la investigación. Por ello es común basarse en ese método para resolver problemas de estructuras en

ingeniería civil o dar solución a numerosos fenómenos físicos cuya formulación matemática conduce a sistemas de ecuaciones diferenciales en derivadas parciales que, como hemos advertido, pueden tratarse eficazmente mediante ese método.

En la aeronáutica. Citaremos a título de ejemplo solo algunas de las aplicaciones más generales. Análisis estático de estructuras de aeronaves, Figura 7, cohetes espaciales, vehículos terrestres, y ello para el estudio de, alas de los aviones, fuselaje, estudio de motores, reparto de temperaturas, elementos mecánicos de todo tipo, etc. También en el cálculo de las vibraciones y la estabilidad de las estructuras, así como respuesta a la dinámica de las cargas aleatorias, flujo de temperaturas transitorias en las cabezas de los cohetes, en las palas de las turbinas, etc.

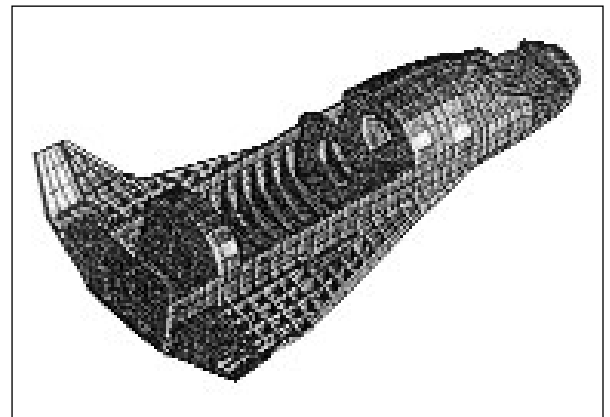


Figura 7. Modelización utilizando elementos finitos, Aviones Marcel Dassault

Las aplicaciones más voluminosas hacen relación al análisis no lineal de grandes deformaciones, estudio de las estructuras en vuelo, optimización del peso de las estructuras. También para la simulación de impactos.

El objetivo del ingeniero es trabajar con modelos que, aunque complejos, sea posible poder llegar tan cerca de la realidad como sea posible, lo que conlleva tener en cuenta grandes niveles de detalle.

En el automóvil. Muchas de las aplicaciones citadas anteriormente para la industria aeronáutica son también útiles en la del automóvil. La competencia empuja a los productores a optimizar sus modelos. Las principales áreas de aplicación son: cálculo de carrocerías, chasis, suspensión y numerosas piezas mecánicas. Se simula la estructura de los motores, se investiga obtener la mayor ligereza de los elementos constructivos que sea compatible con la rigidez y con la ausencia de vibraciones. También es posible estudiar la distribu-

ción de temperaturas en los motores y simular accidentes contribuyendo así a estudios de seguridad.

Las aplicaciones mencionadas también son aprovechables en la industria ferroviaria para la construcción de vagones, locomotoras y coches.

En la ingeniería civil. Cálculo de estructuras de hormigón armado y hormigón en masa; presas bóveda y gravedad; puentes, silos; edificios de estructura metálica y de hormigón.

Para todos esos problemas existen, desde hace ya algunos años, paquetes para el computador desarrollados fundamentalmente por compañías de software, basados en las técnicas de elementos finitos y que proporcionan excelentes soluciones cuando se emplea, como herramienta de base, el supercomputador.

7. EL CAMPO DE LA INVESTIGACIÓN

Los supercomputadores se aplican en numerosos dominios de la investigación pura y aplicada, en la física, en la química, etc.

Campo de la física. En esta ciencia existen numerosos dominios que demandan gran potencia de computación en cuanto a rapidez de cálculo y capacidad de memoria y en los que se hace necesario, por lo tanto, grandes supercomputadores. Los temas que se citan a continuación han sido identificados desde hace ya varios años por la NSF, *National Science Foundation*, en Estados Unidos.

- Teoría cuántica de campos, basada en cálculos que importan integrales de órdenes seis y ocho.
- Teoría de los quarks, tipo teórico de partículas elementales con las que se forman otras partículas, como el protón y el electrón.
- Estructura de la materia en los estados condensados.

En lo relacionado con la física nuclear, y especialmente con la física de las partículas, nuevas aproximaciones permiten, utilizando técnicas de Monte Carlo, determinar propiedades fundamentales como la masa de los *hadrons*, compuestos de multipletes de carga o isospín que son degenerados en masa cuando solamente se consideran interacciones fuertes con los constituyentes principales de los núcleos atómicos. Estas técnicas emplean rejillas de cuatro dimensiones con cientos de millares de nodos, en cada uno de los cuales

se realizan iteraciones con secuencias de operaciones muy complejas que conducen a un valor muy exacto de la propiedad estudiada. Los programas poseen un alto grado de paralelización.

Campos de la química y de la biología. La química cuántica utiliza como base las ecuaciones de *Schroedinger*, cuya solución permite prever las propiedades de los átomos y de las moléculas, como las propiedades eléctricas, su energía potencial y la estructura molecular. Como siempre, esas ecuaciones son complejas y su resolución pasa por aplicar métodos numéricos aproximados. Los tiempos de computación son muy grandes para poder obtener resultados suficientemente exactos sobre cada átomo y, por consiguiente, sobre la molécula, reunión de átomos.

Tanto la *mecánica* como la *dinámica moleculares* utilizan técnicas empíricas consistentes en la realización de *tests* para el cálculo de diversas estructuras moleculares con el fin de identificar aquellas que conducen a resultados más próximos a los datos experimentales obtenidos en los pasos precedentes.

En la *simulación de fluidos* se considera que éstos están constituidos por un gran número de moléculas cada una ejerciendo fuerza sobre las otras, y todas ellas permaneciendo en un espacio cerrado. Para ello se utilizan técnicas de Monte Carlo en las que son datos los movimientos aleatorios de los átomos, los cuales también se eligen de forma aleatoria. Se estudia la propagación de estos movimientos a través del fluido en el que cada átomo interacciona con sus vecinos más próximos. Del modelo se deducen las propiedades del fluido, propiedades que se comparan entonces con los resultados habidos de la experiencia.

Otras aplicaciones modernas en este campo son la química de los polímeros y la cristalografía. Hace muy pocos años era impensable abordar estos problemas con los supercomputadores existentes en ese momento, lo que dice mucho del avance tan espectacular que han logrado las técnicas del paralelismo y su hardware asociado.

8. TÉCNICAS DE DATA MINING: DESCUBRIR EL CONOCIMIENTO

Una *base de datos* es un conjunto de datos organizados de tal forma que se faciliten las operaciones de acceso y mantenimiento de los mismos. Las bases de

datos necesitan un soporte de almacenamiento exterior a la memoria principal RAM del ordenador. Por ello se hace imprescindible el disco magnético o disco duro. Una transacción a una base de datos consta de dos acciones: un proceso de pregunta o consulta, *query process*, para descifrar qué dato se necesita y una operación de *Entrada/Salida, E/S*, para obtener ese dato. El paralelismo trata de reducir los tiempos del proceso de consultas más que el de E/S.

El paralelismo relativo a la gestión de las bases de datos compartidas por varios usuarios y con millones de registros en su contenido relacionados entre sí, se debe contemplar desde dos vertientes:

1. Entre las transacciones propiamente dichas llevadas a cabo de forma *on-line*: se entiende transacciones diferentes que se hacen de forma concurrente.
2. Dentro de una transacción compleja: la transacción se descompone en partes que se ejecutan concurrentemente.

La mayor investigación llevada a cabo sobre la explotación del paralelismo en las bases de datos, se está haciendo sobre la segunda área enunciada. Incluso se han fabricado computadores específicos para la explotación paralela de transacciones sobre bases de datos comerciales.

DIVERSAS APLICACIONES SOBRE BASES DE DATOS

CLASE	CARACTERÍSTICAS
COMERCIAL Banca, líneas aéreas, ventas, fabricación	Transacciones simples y cortas Tráfico elevado Actualización moderada
INGENIERÍA CAD, diseño VLSI, Ensamblaje de aviones	Transacciones largas Diversas versiones de diseño Estructura jerárquica
ESTADÍSTICA Censo, tiempo	Relativamente estable Consultas estadísticas
TEXTO Bases de datos de bibliotecas	Esquemas de consulta complejos sobre datos-texto
SISTEMAS EXPERTOS Inteligencia Artificial, Data Mining	Consultas complejas Capacidad deductiva

Regresando a la aplicación planteada en este apartado, el proceso de la información en general, lo que hace algunos años se denominaba *aplicaciones de gestión*, se está convirtiendo en uno de los más importantes aliados del mercado del paralelismo. En los negocios se toman datos acerca de clientes y productos con el fin de extraer

de esos datos información útil o *conocimiento*. Esa es la base de la *minería de datos*, en inglés, *data mining*. Un ejemplo de esta técnica consiste en partir de una base de datos para determinar los esquemas o patrones de compra de ciertos grupos de población o bien segmentar o distribuir los clientes de una firma de acuerdo con las relaciones encontradas en sus esquemas de compras. Por ejemplo, no es *data mining* encontrar todos los clientes que han comprado una determinada marca de electrodoméstico la última semana; si lo será, sin embargo, segmentar los clientes de acuerdo con sus respectivos grupos de edad, sus ganancias mensuales y sus preferencias en electrodomésticos.

Un tipo especial de minería de datos es la *minería por asociaciones*. El objetivo es descubrir relaciones especiales, denominadas *asociaciones*, dentro de la información disponible relativa a clientes diferentes y sus transacciones respectivas y generar luego reglas de inferencia sobre el comportamiento del cliente. Por ejemplo, la base de datos podría almacenar, para cada transacción, la lista de artículos adquiridos. El objetivo será determinar asociaciones entre conjuntos de artículos tales que la tendencia sea la de haberse adquirido todos juntos.

Veámoslo de forma más concreta. La base de datos contiene los registros correspondientes a las transacciones de compra de un cliente. Cada transacción tendrá su correspondiente identificador y un conjunto de atributos: los artículos comprados. El primer objetivo de la minería por asociaciones será, examinar la base de datos para poder determinar qué conjuntos de *a* artículos aparecen o se encuentran en ella, artículos que se presenten juntos y por encima de un determinado porcentaje de transacciones. El problema asociado con el objetivo enunciado es descubrir un conjunto de artículos de tamaño *a* adquiridos a la vez y que superen un porcentaje dado, así como la frecuencia o probabilidad de que eso ocurra.

Todo lo anterior conduce a un proceso en el que interviene el conteo de un número alto de transacciones de la base de datos que son candidato a los registros buscados, conteo que, por tanto, se ejecuta sobre cientos de miles de transacciones, ejecución que ha de llevarse a cabo en paralelo si se pretende obtener resultados reales en tiempos asequibles y satisfactorios.

9. PARA ESCRIBIR UN FINAL

Simulación de la evolución de las galaxias.
Se trata del estudio de la evolución de las estrellas en

un sistema de galaxias a través del tiempo para dar solución a cuestiones como, qué ocurre cuando las galaxias colisionan o cómo un conjunto aleatorio de estrellas se enlazan entre sí para dar lugar a una forma galáctica concreta. Este problema, *problema de los n-cuerpos*, conduce a una simulación del movimiento de un cierto número de cuerpos, estrellas en este caso, que se mueven bajo las fuerzas procedentes de los otros. El cálculo se discretiza en el espacio, considerando cada estrella como un cuerpo separado, y en el tiempo, simulando el movimiento de las galaxias a lo largo de pequeños intervalos de tiempo, en cada uno de los cuales se calculan las fuerzas gravitatorias sobre cada estrella debidas a las restantes y se actualiza su posición, velocidad y otros atributos.

El algoritmo utilizado se debe a Barnes-Hut y por eso, además, lleva su nombre. En este algoritmo se da una concurrencia muy amplia, aplicable a la multitud de estrellas del modelo y a los intervalos de tiempo. Pero dada la irregularidad y su naturaleza cambiante, se hace necesaria la explotación de esa concurrencia de manera eficiente sobre una arquitectura paralela.

Visualización de escenas complejas por medio de trazador de rayos. La visualización es un aspecto importante de la ingeniería de procesos: cuanto más real es una imagen, mejor será la evaluación de los resultados. El objetivo de los métodos gráficos por computador, es producir imágenes que no puedan distinguirse de las fotografías de escenas reales de tres dimensiones.

Una de las aplicaciones de gráficos por ordenador y su técnica asociada consiste en convertir las escenas en imágenes mediante trazador de rayos, *ray tracing*. La escena se representa como un conjunto de objetos en un espacio 3-D y la imagen a convertir se representa como un conjunto bidimensional de *pixels* de los que se calculan los valores de color, opacidad y brillo. La computación trata de simular el comportamiento físico de la luz produciendo, como resultado, una representación bidimensional sobre una imagen plana.

El paralelismo se explota sobre cada uno de los rayos y a través de los diversos *pixels*. Cada rayo es independiente de los demás y, por tanto, puede tratarse en paralelo con los otros. Se trata de una aplicación que requiere gran cantidad de datos tomados de un sistema paralelo con elementos de proceso remotos entre sí. Decía François Sillion, para poner de manifiesto la duración en tiempo que entraña este tipo de aplicaciones: *los resultados sobre fotos reales se consiguen en unidades de tiempo geológicas*.

Diseño VLSI, Very Large Scale Integration, asistido por ordenador. Cuando los computadores eran fabricados a base de transistores separados, los efectos de los errores en su diseño como, por ejemplo, los debidos a un error de conexión, podían solucionarse fácilmente. Los tiempos de diseño y fabricación eran menores que en las actuales técnicas para producir circuitos de gran integración VLSI. Localizar un problema en estos circuitos modernos se hace más complicado ya que no puede hacerse un test punto por punto entre los miles de circuitos contenidos en un chip VLSI.

Actualmente se hace una simulación de la lógica del circuito, simulación que representa una búsqueda probabilística de los errores. Esa simulación puede consumir miles de horas de CPU aunque se tratara de un gran computador. Es necesario, una vez más, acudir al computador paralelo ya que afortunadamente, la simulación aludida presenta un alto grado de paralelismo.

Sistemas de Inteligencia Artificial, IA. El paralelismo ha llamado a las puertas de las aplicaciones de IA debido a que las técnicas de *búsqueda*, *search* en inglés, juegan aquí un papel primordial y central. Los *sistemas expertos* realizan un razonamiento útil basado en conjuntos de reglas obtenidas de la propia experiencia humana. Pensemos, por ejemplo, en el diagnóstico médico. Los sistemas actuales contienen más de 10.000 reglas y su proceso en el computador presentan un alto grado de paralelismo potencial.

Las reglas se almacenan y procesan mediante programas denominados *sistemas de producción*. Actualmente se habla de sistemas de producción paralelos y, por tanto, se sustentan en arquitecturas superordenador.

Otros elementos de la Inteligencia Artificial además de los sistemas expertos: visión artificial, proceso del lenguaje natural, planificación y solución de problemas. Se trata, en todos los casos, de aplicaciones donde la concurrencia conseguida por el supercomputador convierten a éste en una herramienta fundamental.

El trabajo se expande hasta llenar el tiempo disponible para su terminación.

C. Northcote Parkinson

Es imposible producir un sistema útil cuando no se confía en ninguno de sus componentes.

Séneca