

LA COMPUTACION DISTRIBUIDA COMO  
UNA ALTERNATIVA DEL FUTURO\*

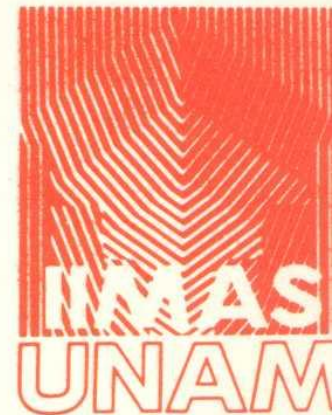
Adolfo Guzmán Arenas

*serie naranja: investigaciones*

INSTITUTO DE INVESTIGACIONES  
EN MATEMATICAS APLICADAS  
Y EN SISTEMAS

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

APARTADO POSTAL 20-726  
MEXICO 20, D. F.  
548-54-65



# comunicaciones técnicas

1979

Serie Naranja: Investigaciones

No. 216



## LA COMPUTACION DISTRIBUIDA COMO UNA ALTERNATIVA DEL FUTURO\*

Adolfo Guzmán Arenas

\* Informe Técnico- AHR-79-6

Recibida: 10 de octubre de 1979

INSTITUTO DE INVESTIGACIONES  
EN MATEMATICAS APLICADAS  
Y EN SISTEMAS

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

APARTADO POSTAL 20-726  
MEXICO 20, D. F.  
548-54-65



## LA COMPUTACION DISTRIBUIDA COMO UNA ALTERNATIVA DEL FUTURO

Adolfo Guzmán Arenas

Departamento de Computación, IIMAS  
Universidad Nal. A. de México

## RESUMEN

Después de haber fabricado enormes computadoras que son usadas a través del tiempo compartido y la multiprogramación, las economías de la integración en gran escala inducen al hombre a pensar en multitud de máquinas modestas (microcomputadoras) que colaboran para desarrollar las mismas tareas que hasta ahora han venido desarrollando sus hermanas mayores.

Es la división y repartición del trabajo en múltiples unidades de procesamiento; es la intercomunicación, colaboración y coordinación de tales unidades; es el abaratamiento de la transmisión de datos, al solo transmitirse información requerida; es el procesamiento local de la información local; éstos son los atributos principales de sistemas de procesamiento distribuido, los que son analizados en este artículo.

También se mencionan las redes de computadoras, antecesores ya clásicos de la computación distribuida.

El artículo concluye con la discusión de tres proyectos de computación distribuida que se están realizando en el Departamento de Computación en la UNAM: (1) la construcción de una computadora "grande" que posee LISP como lenguaje de máquina, y que internamente está formada por docenas de microprocesadores; (2) la construcción de una máquina especialmente diseñada para manejo eficiente de bases de datos y utilizable en el soporte de sistemas de información; (3) un sistema de procesamiento distribuido que, mediante una repartición juiciosa de la carga de procesamiento, y a través de una conexión (cable

coaxial) de alta velocidad, elimina la necesidad de tener una computadora "grande" en el sistema, el que únicamente consta de la red de "terminales" inteligentes.

#### PROBLEMAS PROPIOS DE COMPUTACION CENTRALIZADA

Para aplicaciones que no necesitan extensivamente capacidad de cómputo, la gran computadora puede ser costosa. Los costos de telecomunicaciones pueden ser excesivos en una organización con usuarios dispersos geográficamente. Una instalación local pudiera no tener todos los recursos requeridos para un trabajo específico. Los tiempos de respuesta para trabajos y procesos en tiempo real pueden tornarse impredecibles e incontrolables debido a la gran variación en carga típica de una instalación de múltiples usuarios. La gran complejidad del hardware y del software del sistema operativo necesarios en una gran instalación pueden fallar más de lo aceptable, frente a ciertos trabajos con tiempos críticos de entrega. Un usuario puede recibir poca atención y servicio deficiente de un gran centro de cómputo, debido a su gran dimensión y compleja estructura.

#### LA ECONOMIA DE ESCALAS AHORA FAVORECE AL PEQUEÑO COMPUTADOR

Los grandes y rápidos cambios en la tecnología de semiconductores han tenido un impacto fuerte en los costos relativos de las diferentes partes de un sistema de cómputo que posee comunicaciones. Mientras que los costos de lógica y de almacenamiento interno han disminuido muy rápidamente, los costos de almacenamiento (memoria) masivo han disminuido menos rápidamente, en tanto que los costos de comunicación han permanecido prácticamente constantes. Además, el software se ha converti-

do en un porcentaje creciente del costo total del sistema de cómputo. Ya no hay, pues, una penalidad económica severa en contra de distribuir la capacidad de procesamiento a las terminales individuales.

En sistemas de tiempo compartido y de multiacceso, hay varias razones sólidas funcionales para centralizar la base de datos, particularmente en aquéllos casos donde varios usuarios pueden alterarla y accederla. Sin embargo, no existen fuertes razones funcionales para centralizar las funciones mismas de procesamiento. La decisión de centralizar el procesamiento del usuario individual es una decisión económica. Con las tendencias actuales en los costos de equipo digital por una parte y los de comunicación por la otra, es económicamente atractivo efectuar tanto como sea posible de cómputo y procesamiento localmente, para así minimizar las necesidades de transmisión de datos y por ende los costos de comunicación.

#### VENTAJAS DE COMPUTACION DISTRIBUIDA

Puesto que el costo ya no es una barrera, se pueden aplicar las ventajas inherentes de los sistemas distribuidos a una gran cantidad de problemas de cómputo. Estas ventajas incluyen las siguientes:

1. Confiabilidad. Se puede tener redundancia de una manera relativamente barata, en un sistema distribuido. El sistema entero no tiene que ser duplicado, como es el caso de una sola computadora. Sólo un número pequeño

Capacidad de cómputo a bajo costo

Se presenta la capacidad de cómputo barata, en la forma de minicomputadores y microcomputadoras, como una alternativa atractiva frente a las computadoras medianas y grandes, para el desfogue de muchos trabajos de cómputo. Esta tendencia se incrementará en el futuro. Aún más, puede argüirse que la economía de escalas favorece ahora al computador pequeño. El costo por unidad de potencia de cómputo es menor en una línea de ensamble de un gran número de pequeñas computadoras, que en una línea de ensamble de un pequeño número de grandes computadoras.

de procesadores deben ser agregados para asegurar el grado requerido de disponibilidad. Se pueden conseguir estructuras de software más simples, y por lo mismo más confiables, en una colección de pequeñas computadoras.

2. Responsividad. El sistema distribuido puede ser más rápido en sus reacciones (más responsivo), puesto que se puede proporcionar acceso directo a una computadora, aún a pequeñas comunidades de usuarios. Esta responsividad toma la forma de un tiempo de ida-y-vuelta reducido en ambientes de procesamiento por lotes, y significa tiempos de respuesta más rápidos en un ambiente de teleproceso o tiempo real.
3. Crecimiento incremental. Un sistema distribuido próximo a sobrecargarse puede ser expandido incrementalmente a bajo costo, mediante la adición de más procesadores. Un sistema centralizado próximo a sobrecargarse puede asimismo ser descargado mediante la distribución de ciertas funciones hacia procesadores pequeños.
4. Correspondencia con patrones organizacionales. Muchas organizaciones poseen una naturaleza descentralizada. Podría presentar restricciones no naturales para la operación eficiente de la organización el tener instalaciones de cómputo centralizadas. Esta restricción se alivia mediante computación dispersa, pues permite a un grupo descentralizado su acceso a su propia computadora, la que puede optimizarse para las necesidades de tal grupo. Un arreglo en forma de red permitirá a la gerencia central tener acceso a información resumida de todos los grupos.
5. Se comparten recursos. Una red de sistemas de cómputo distribuidos permite a los usuarios en un sitio usar ventajosamente los recursos que se encuentran disponibles en otras localidades. Estos recursos pueden consistir de programas, bases de datos, y poder de cómputo. Las redes que compar-

ten recursos permiten balance de carga (de cómputo), respaldo, y disminuyen la duplicación de esfuerzos.

#### PROBLEMAS DE LA COMPUTACION DISTRIBUIDA

1. Algunos tipos de problemas se resuelven sobre todo en sistemas grandes, por ejemplo aplicaciones numéricas.
2. Un diseño pobre puede conducir a una pérdida de confiabilidad y de responsividad.
3. La descentralización de la capacidad de cómputo puede conducir a un crecimiento incontrolado, a imperios locales, y a la difusión y dispersión de personal de cómputo dentro de la organización.
4. Un sistema conteniendo muchas computadoras es bastante complejo de construir y validar. Aún no existen disponibles en el mercado una gran variedad de software que conduzca a la computación distribuida.

Antes de tomar decisión alguna, deben pesarse cuidadosamente las ventajas y desventajas de procesamiento centralizado versus procesamiento distribuido. Sin embargo, hay razones para creer que la tendencia hacia procesamiento distribuido crecerá, particularmente conforme el costo de poder de cómputo de cline.

#### DIFERENTES DEFINICIONES DE SISTEMAS DISTRIBUIDOS

- A. Para algunos usuarios, un sistema distribuido es un conjunto de terminales inteligentes, situadas en el lugar donde se requiere primariamente su uso, para dar a los elementos de la organización local más poder de cómputo y un soporte

de respuesta más rápida. Estas terminales llevan a cabo la mayoría de las funciones de cómputo para el grupo local. Cuando se torna necesario, las terminales se comunican entre sí y con grandes computadoras remotas, para mejor servicio.

Estos usuarios están preocupados por la optimización de sus funciones dentro del negocio, determinación de los mejores productos disponibles para la construcción del sistema, reducción de costos de comunicación, y tal vez con problemas severos de control asociados con facilidades de cómputo descentralizadas.

- B. Para otros usuarios, un sistema distribuido es una colección de múltiples computadoras o elementos de procesamiento, trabajando estrechamente vinculados en la solución de un solo problema. Un ejemplo podría ser un sistema en línea para Bancos, constituido por varias minicomputadoras ligadas entre sí por memoria compartida, por líneas de comunicación y/o por vías (buses) comunes. Cada minicomputadora procesa un subconjunto de las transacciones bancarias y actualiza una porción de una base de datos común.

Los usuarios de tales sistemas están preocupados con problemas de diseño de hardware y software, confiabilidad, sistemas operativos para multicomputadoras, y cómo descomponer óptimamente programas y bases de datos. También se preocupan de la integridad de la base de datos, y de la forma de poder seguir efec

#### Limitaciones en el Desarrollo de Sistemas de Cómputo

En el presente, los principales factores que limitan las mejoras ulteriores en todos tipos de sistemas de computación, incluyendo sistemas en línea, son:

1. Equipo de entrada y salida. Velocidad, costo y confiabilidad.
2. Memoria. Su velocidad y costo, tanto de memoria periférica como de memoria principal (interna).
3. Software. Su costo, su complejidad, el almacenamiento requerido, la dificultad de convertirlo a otra máquina (a otro hardware).



tuando transacciones (y brindando un servicio adecuado) frente a fallas de partes del sistema: terminales, líneas de transmisión, conmutadores de mensajes, etc.

- C. También hay usuarios para los cuales procesamiento distribuido significa el uso de pequeñas computadoras para desarrollar funciones que anteriormente habían sido desarrolla-

El denominador común de los sistemas distribuidos

El tema común que liga los diferentes tipos de sistemas distribuidos es el uso de múltiples elementos de cómputo que colaboran entre sí para realizar trabajos que hasta ahora habían sido desarrollados en grandes computadoras, o no habían sido desarrollados en ningún lado.

Los sistemas distribuidos se han desarrollado como alternativas a grandes instalaciones de cómputo porque, en muchas circunstancias, las instalaciones centrales no han sabido proveer soluciones óptimas o efectivas a problemas importantes.

res subalternos (backend) para dar soporte a funciones de administración de bases de datos. Los diseñadores de tales sistemas se preocupan de los métodos para acoplar máquinas pequeñas, a grandes, interacciones de los sistemas operativos, y el desarrollo de programas eficientes para los sistemas de soporte.

- D. Otros usuarios entenderán por sistema distribuido colecciones de centros de cómputo, geográficamente dispersos e independientes, ligados entre sí por líneas de transmisión de datos para permitir el uso común y compartido de recursos de hardware y de software. Estos sistemas se conocen como redes de cómputo, y son los predecesores de los verdaderos sistemas de procesamiento distribuido.

das por computadoras grandes. Estas pequeñas computadoras, al absorber tales funciones, releven de carga a las máquinas grandes y prolongan su vida. Por ejemplo, procesadores frontales de comunicaciones se han venido usando durante años para manejar comunicaciones de datos. Recientemente, se han desarrollado computado-

Un ejemplo es una red desarrollada por varias instituciones educativas, cada quien con su propia instalación de cómputo. La red permite a un usuario en una instalación el acceso a un programa o a un archivo residente en otro centro de cómputo. La red puede permitir asimismo la nivelación de las cargas y el respaldo cuando algún integrante de la red falla.

Los diseñadores de tales redes se preocupan por la seguridad de la red, el mantenimiento de la independencia local, algoritmos de cobranza apropiados, administración de la red, comunicaciones óptimas, y el eslabonamiento de computadoras heterogéneas.

#### PRINCIPALES CATEGORIAS DE SISTEMAS DE PROCESAMIENTO DISTRIBUIDO

En razón del grado de acoplamiento entre los elementos de procesamiento y el propósito para el cual se desarrolló el sistema, se distinguen tres categorías de sistemas de procesamiento distribuido:

- (1) Débilmente acoplados, cuando los procesadores están conectados por líneas seriales de ancho de banda relativamente bajo (inferior a 50 kilobits/seg). En su mayoría son redes de computadoras que permiten compartir recursos. Estas redes conectan centros de cómputo de propósito general que se encuentran dispersos geográficamente, siendo la conexión de acoplamiento débil. La red permite que los usuarios en una localidad usen ventajosamente los recur-sos disponibles en otras localidades.

Cuando las computadoras son heterogéneas y operan bajo diferentes estructuras de administración, surgen problemas significativos de software y de administración en este tipo de redes.

En este tipo de sistemas una computadora está restringida a su "medio ambiente" local (periféricos y memoria), pues los vínculos con otras máquinas son lentos. No se comparte memoria principal, ni se ejecutan trabajos por dos máquinas diferentes de la misma memoria de acceso aleatorio.

Cuando las computadoras son heterogéneas, no hay nivelación de carga, por regla general.

- (2) Moderadamente acoplados, cuando los procesadores están conectados por líneas seriales de alta velocidad (superior a 50 kilobits/seg), vías (buses) paralelas, y/o periféricos compartidos. Estos sistemas se conocen generalmente como sistemas de múltiples procesadores. Normalmente están formados por procesadores homogéneos (todos son del mismo tipo), y el sistema está dedicado a un conjunto de trabajos limitados y bien definidos. Los procesadores pueden estar en la misma localidad o geográficamente dispersos.

Este tipo de sistemas, cuando está formado por minicomputadoras, poseen el potencial de rebasar y superar a máquinas grandes de uso general, más caras, en una diversidad de aplicaciones dedicadas.

Algunos sistemas moderadamente acoplados constan de computadoras grandes auxiliadas por computadoras chicas para desfogar funciones críticas. Como ejemplo se puede citar el procesador subalterno (back-end) para manejo de bases de datos. Este es un proyecto que actualmente se encuentra en desarrollo en el Departamento de Computación (IIMAS) de la Universidad Nacional Autónoma de México: la manufactura de una máquina para manejar bases de datos eficientemente.

- (3) Fuertemente acoplados, cuando los procesadores están conectados por memoria principal compartida, vías de acceso

directo a memoria (DMA) y velocidades en los millones de bytes/seg.

Los multiprocesadores fuertemente acoplados proporcionan el potencial para una alta confiabilidad, alto desempeño, y computación geográficamente centralizada. Un acoplamiento fuerte entre procesadores siempre es un reto tecnológico en la coordinación de procesadores, y en software con imaginación.

Más adelante se describe el proyecto AHR: la construcción de una máquina que procesa LISP como lenguaje de máquina, y que internamente está constituida por varias docenas de microprocesadores Z-80. Este es un ejemplo de un sistema de multiprocesadores fuertemente acoplados, y está siendo llevado a cabo en el Departamento de Computación (IIMAS) de la UNAM.

#### EL DATOR COMO CELULA DEL PROCESAMIENTO DISTRIBUIDO

Un dator es la célula constituyente de un sistema de procesamiento distribuido, en especial de los débil y moderadamente acoplados.

Un dator es un sistema de procesamiento/captura/almacenamiento de la información que posee las siguientes características:

#### Bases de datos distribuidos

Una base de datos distribuida es tá formada por dos o más elementos de procesamiento, cada uno de los cuales tiene un conjunto de archivos asociado a él. Estos archivos pueden estar divididos en razón geográfica, o debi do a bases del funcionamiento de la compañía, o bien por razones de redundancia y confiabilidad.

Por regla general, una base de datos distribuida se construye sobre un sistema distribuido débil o moderadamente acoplado.

- 1) Capacidad de procesamiento propia;
- 2) Interacción con el usuario. Este interacciona el dator con la computadora central, cuando ésta existe aún;

- 3) Programas almacenados;
- 4) Parte del procesamiento es local y lo realiza el dator;
- 5) Comunicaciones en línea con otros datores, con la(s) computadora(s) central(es) y sus bases de datos correspondientes;
- 6) Entrada de datos orientada a la interacción con el usuario: teclado, lápiz luminoso, digitalizador de mapas, etc.;
- 7) Salida de datos orientada al usuario: impresora serial, tubo de rayos catódicos, sintetizador de voz, etc.;
- 8) Habilidad para compartir el procesamiento de trabajos, y cierto grado de coordinación entre datores;
- 9) Habilidad para compartir información con otros datores;
- 10) (opcional) Habilidad para balancear cargas con otros datores;
- 11) Almacenamiento local (v. gr., discos) de datos.

Las habilidades (1), (3) y (4) le dan al dator características de computadora; las habilidades (2), (6) y (7) le dan al dator características de terminal; las habilidades (4), (5), (8) y (9) le dan al dator propiedades para poder desarrollar computación distribuida; las habilidades (8), (9) y (10) hacen que los periféricos (con la información en ellos contenida) sean promiscuos; las habilidades (9) y (10) hacen que los datores (los elementos de procesamiento) sean promiscuos.

#### QUE ES UN DATOR

Un dator es al mismo tiempo computadora, terminal, nodo de una red, concentrador de información, adquisidor de datos y expulsor/exhibidor de resultados.

Los datores eliminan concentración de la capacidad de procesamiento en una computadora central, repartiéndosela entre ellos.

Los datores son terminales tan inteligentes que han eliminado (o reducido grandemente) la computadora central a la que estaban conectados.

Si existe, la computadora central de un sistema distribuido construido con datores tiene las siguientes funciones:

- a) Mantiene una base de datos central;
- b) Provee capacidad de cómputo elevada para problemas que no pueden descuartizarse satisfactoriamente entre los dadores constituyentes.

Si no existe, la computadora central de un sistema distribuido constituido con dadores es reemplazada, con elevadas economías, por un conjunto de dadores que colaboran en las diferentes tareas de la instalación: captura de datos, procesamiento de información, mantenimiento de una base de datos, generación y distribución de resultados.

#### CARACTERISTICAS DE LOS SISTEMAS DISTRIBUIDOS FORMADOS POR DATORES

Los sistemas distribuidos, cuando la distinción entre unidad central de proceso y terminal se desvanece, exhiben varias características deseables, además de las propias de los sistemas distribuidos (véase "Ventajas de Computación Distribuida", al principio de este artículo), entre las cuales aparecen:

Descentralización del procesamiento. Un sistema distribuido de dadores puede poseer o no una (o más) computadoras centrales. La eliminación de esta(s) computadora(s) ofrece ahorros económicos, puesto que varios computadores pequeños (dadores) son más baratos que uno grande (véase "La Economía de Escalas Ahora Favorece al Pequeño Computador", arriba).

También se ahorra el envío de información al lugar central de procesamiento, al procesarse la información localmente, en el lugar donde se produce, y por lo común por el dator mismo que la capta. Es decir, la computación o procesamiento está distribuida.

A veces, y especialmente frente a grandes problemas numéricos (v. gr., inversión de grandes matrices; problemas de pro-

gramación dinámica o no lineal), no se sabe el algoritmo para obtener eficazmente la solución del problema usando datores distribuidos, y tiene que seguirse el procedimiento convencional de usar una computadora grande central. Sin embargo, ésta a veces puede reducirse a un uso especializado de "tritador numérico", y aún pudiera ser sustituida por procesadores especializados (procesadores para arreglos, para Transformadas Rápidas de Fourier, procesadores de punto flotante, etc.) En estos últimos casos, se tendría un computador central ajustado a la labor especial, auxiliado por datores en una configuración distribuida.

Nivelación de cargas. Con esto se designa la distribución de los diferentes trabajos a los diferentes datores, de tal suerte que los recursos del sistema se usen equitativa o niveladamente.

Si consideramos que cada usuario está conectado a cada teclado activo, entonces la nivelación de cargas en un sistema distribuido puede tomar tres aspectos:

I. Hay un solo proceso (task) por usuario (es decir, por teclado).

En este caso, salvo que se sepa cómo efectuar paralelismo dentro de un mismo task (lo que ocurre por ejemplo en la Máquina de LISP del Proyecto AHR, a ser descrito después), no hay posibilidad de nivelar la carga, pues a cada usuario se le asigna cuando mucho una c.p.u. y no hay dos cpu's colaborando en el mismo proceso.

Nivelación de carga en redes

La habilidad de tomar una carga de trabajo dada y distribuirla entre las computadoras de una red o los datores de un sistema distribuido, para así hacer un uso igual de los recursos de la red, es un servicio común en un gran número de redes. Se ofrece en las redes CYBERNET, DCS y TSS, pero no es una característica básica de las redes ARPA, MERIT, OCTUPUS o TUCC. En éstos puede agregarse bajo ciertas restricciones.

- II. Hay varios procesos por usuario. Aún así, puede no haber nivelación de cargas, si (como es el caso de las computadoras PERQ de la compañía THREE RIVERS) se asigna una c.p.u. a todos los procesos del mismo usuario. En este caso, se procesan en la localidad donde el usuario reside los trabajos que él genera, por la c.p.u. local. El monitor local es de multiprogramación, ya que puede atender varios trabajos, todos sin embargo del mismo usuario.
- III. Las cpu's de los dadores son promiscuas. Esto quiere decir que los dadores pueden ejecutar tanto los trabajos emanados del usuario "local", como los otros usuarios, importados (los trabajos) de otros dadores a través de las facilidades de comunicación existentes.

#### Compartición de datos en redes

Todas las redes actuales, incluyendo las débilmente acopladas, permiten el envío de datos hacia el programa que los requiere.

Si las redes son homogéneas y los datos a transmitirse son grandes, a menudo resulta más conveniente mandar el programa a donde están los datos. Esto no puede hacerse en redes heterogéneas.

Generalmente, para que exista nivelación de carga, los sistemas distribuidos deben estar moderada o fuertemente acoplados, debido a la elevada ocupación de los canales de comunicación.

El caso es el mismo que en el caso II anterior. Un usuario puede generar varios trabajos, y los dadores poseen

sistemas operativos de multiprogramación. Sin embargo, ahora un sistema operativo puede exportar carga hacia otro dator que se encuentre menos ocupado. Para compartir carga eficazmente es necesario poder compartir datos también, debido a que un programa (trabajo) que emigra de un dator a otro necesita llevarse consigo los datos a que hace referencia, parte de los cuales estarán en el dator del emigrante. El sistema distribuido que se está diseñando en el



Departamento de Computación (IIMAS) de la UNAM poseerá entre sus características el de compartir cargas, es decir, el tener datores promiscuos.

Periféricos Promiscuos. Los datores poseen cierta capacidad de almacenamiento periférico. Estos dispositivos pueden ser accedidos por otros datores del sistema, por lo que la información que se posee se comparte entre todos los datores. Sin embargo, se entiende que la información local se accesa más rápidamente que la información foránea.

En sistemas fuertemente acoplados, los datores comparten memorias y dispositivos de almacenamiento masivo. En sistemas débilmente o moderadamente acoplados, a menudo los dispositivos periféricos pueden ser accedidos directamente sólo por el dator al que pertenecen, el que sin embargo tiene "la obligación" de leer de ellos datos y enviárselos a otros datores que los soliciten. De esta forma, cada dator "indirectamente" accesa todos los datos, asegurándose así su promiscuidad.

Al tener los datos descentralizados, en vez de tenerlos concentrados en la computadora central, se obtiene un mejor rendimiento de los sistemas de cómputo, ya que varios datos pueden estar procesándose separadamente (en paralelo, simultáneamente) en diferentes datores, y se evita asimismo el envío de datos locales a un lugar central de proceso. Esto último disminuye los costos de transmisión de datos. Se transmiten sólo excepciones (regiomontano que fue a Mérida a cambiar su cheque), resúmenes (control de inventarios, control de gastos) ó resultados.

Regla: poner la información dentro del dator que sirve al usuario que más usa esa información.

A veces se incluye esta otra regla: sólo el dator local puede hacer actualizaciones a su información local, pero todos

pueden consultarla. Viórase a veces, permitiéndose que varios dadores hagan actualizaciones a una información foránea: actualizar el número de zapatos existentes en la bodega, desde varios puntos de venta.

Captura de datos/interacción hombre-máquina. La captación de datos se realiza en forma local, donde cada dator provee la inteligencia (en forma de programas de validación y verificación) para que los errores e inconsistencias se corrijan tan pronto como se cometen y detectan. Es decir, se pone un filtro oportuno a los datos erróneos, los que no llegan muy adentro del sistema de información.

Si se tienen dadores promiscuos (con nivelación de cargas), entonces resulta adecuado tener un dator por cada usuario "consistente." En esta forma la verificación de errores de entrada es excelente. Y cuando el usuario se va y ya no ocupa su dator, la cpu de éste colabora a procesar trabajos de otros usuarios que han permanecido activos.

Como ejemplo de este sistema, en el laboratorio de Inteligencia Artificial del M.I.T. (Instituto Tecnológico de Massachusetts) han instalado un teclado y una pantalla por usuario, pero la memoria para el despliegue de caracteres y la capacidad de cómputo para tal despliegue está centralizada en una pdp-11 satélite de la pdp-10, que es la máquina que procesa la información.

Si las pantallas son de alta resolución, del tipo de raster-scan-displays, y el sistema operativo es de multiprogramación, entonces es posible designar parte de la pantalla para la salida de un trabajo, otra parte para la salida de otro, etc. En esta forma es posible que el mismo usuario esté trabajando en el mismo dator con dos ó más programas suyos, interactivos y

concurrentes. Tal sucede con las máquinas PERQ. Y si existe nivelación de carga, pudiera suceder que sus diferentes trabajos estuvieran siendo procesados en dadores distintos.

La idea de que los dadores desocupados contribuyan al desfogue de trabajos que les corresponden a dadores "ocupados" es buena, puesto que permite un uso más eficiente de la capacidad de cómputo instalada. Si yo soy el usuario o dueño de un dator, y no lo estoy usando, qué bueno que pueden usar de él otros usuarios más trabajadores que yo. Esto no ocurre con los teléfonos, por ejemplo: los aparatos telefónicos de las oficinas que están cerradas no los puede usar nadie. Sin embargo, esto sí ocurre con los selectores para cada teléfono: los teléfonos desocupados no usan selectores, y éstos quedan libres para su uso por los teléfonos ocupados. (Los selectores están centralizados en la central telefónica). Lo que aquí se propone es el uso promiscuo de recursos distribuidos.

#### Acceso a información remota.

Si imaginamos un sistema de dadores donde cada uno posee parte de la información total o base de datos, entonces hay dos formas substancialmente diferentes de usar la información:

1. Cada dator usa su información local o cualquier información foránea, sin mezclarlas o mezclándolas poco y de manera simple. Por ejemplo, un dator puede contestar una pregunta mediante un acceso a su base de datos

#### Nivelación estática de carga

En un sistema de bases de datos distribuidas, la carga total puede nivelarse preasignando funciones y segmentos de la base de datos a los dadores, en forma estática. Este método es usado a menudo en un sistema distribuido débilmente acoplado, el que puede estar formado por computadoras heterogéneas, pequeñas, medianas y microcomputadoras. Las funciones que se deben proveer son: (1) asegurar acceso remoto y local a las partes de la base de datos distribuida, y (2) trasladar las diferentes partes de la base de datos de uno a otro dator.

local, o a la base de datos foránea. O bien, la pregunta puede responderse con 95% de la información local y 5% de la foránea, o viceversa.

2. Para contestar una pregunta o resolver un problema, el dator debe acceder varias bases de datos y conjuntar, mezclar y entreverar las informaciones parciales obtenidas, de una manera compleja y por un número sustancial de veces.

#### Nivelación dinámica de carga

Un sistema distribuido puede usarse para desparramar dinámicamente la carga de procesamiento total, y la carga de acceso a la base de datos, a través de los dadores disponibles.

Si todos los dadores poseen la misma capacidad, los trabajos o transacciones pueden moverse libremente entre ellos. Sin embargo, puede ser deseable ejecutar un trabajo específico en una computadora particular, porque requiera acceso a una base de datos asociada a esa máquina. En otras ocasiones puede ser deseable mover trabajos, tal vez acompañados por sus datos, de un dator sobrecargado a otro que posee capacidad disponible.

desdoblarse, suponiendo que existe un dator por cada estado, en 30 preguntas, cada una por cada estado e, del siguiente tipo: "En tu estado e, y para la edad j variando de 6 a 10, dime cuántos niños hubo con tosferina con edad j el año pasado". La integración de las respuestas es simple, y consiste en sumar los arrosos parciales.

El desdoblamiento de preguntas requiere (si se realiza en forma automática) de conocimientos profundos, por

Para esto se necesita un desdoblador de preguntas, capaz de transformar una pregunta dirigida a toda la base total en una función lógica de preguntas parciales, las que al combinarse por un integrador de respuestas nos brinda la respuesta final a la pregunta inicial. Por ejemplo, la pregunta "¿cuántos niños varones entre 6 y 10 años tuvieron tosferina el año pasado, en toda la República?" puede

parte del desdoblador, de la constitución de las bases de datos parciales; se necesita introducir variables, como el estado y la edad en el ejemplo anterior, sobre las cuales cuantificar "para todas las edades j y todos los estados e, ...".

Acceso a procesamiento remoto. La nivelación de carga de la que ya se habló puede realizarse desde una manera muy sencilla hasta una forma muy compleja, como se esboza a continuación.

- a. Sólo un dator para un trabajo. Esta disciplina la siguen la mayor parte de las redes de procesamiento remoto. Una terminal en un extremo de la red usa una computadora en el otro extremo. La terminal (el proceso) ignora y no se beneficia de la gran cantidad de recursos que están también disponibles. Un trabajo entra completamente a un procesador (éste, sin embargo, estará por lo general ejecutando varios trabajos concurrentemente) a ejecutarse.
- b. Paralelismo entre trabajos. Cuando un usuario genera varios trabajos, éstos pueden emigrar hacia diversos dadores distintos. Cada dator posee su sistema operativo de multi programación, y ejecuta sus diversas tareas concurrentemente, o usando una disciplina de tiempo real.

Este esquema se usará en el Sistema de Procesamiento Distribuido que se está diseñando en el Departamento de Computación (IIMAS) de la UNAM.

- c. Paralelismo dentro de un mismo trabajo. Un solo trabajo (task) se subdivide en partes más pequeñas, las que se "exportan" a otros dadores para ser procesados por ellos, y cuyas respuestas se integran para constituir el resultado final. Este tipo de paralelismo es el más difícil de lograr, pues se requiere de la capacidad de segmentar automá

ticamente un trabajo en varios subtrabajos que se puedan después reintegrar (sus resultados) fácilmente.

En la máquina AHR que está siendo construida en el Departamento de Computación (IIMAS) de la UNAM se logra un paralelismo dentro de un mismo trabajo, ya que varias docenas de microcomputadoras procesan y evalúan un mismo programa escrito en LISP (un lenguaje de alto nivel).

Control del sistema distribuido de datos. Las labores de control y supervisión de una red o sistema distribuido de datos son, entre otras:

- nivelación de carga
- estadísticas de funcionamiento
- reconfiguración
- detección y aislamiento de datos y líneas defectuosas.

El control global (y cada una de sus funciones en particular) puede llevarse a cabo en una de las siguientes formas:

- 1) No existente.
- 2) Control centralizado. Cierta dator específico ejerce el control.
- 3) Control descentralizado. Cada dator efectúa ciertas labores de control, las que están diseñadas de tal suerte que se logra el efecto de control, sincronización, reconfiguración, etc., deseado.

### TRES TIPOS ESPECIFICOS DE SISTEMAS DISTRIBUIDOS

Para finalizar el artículo, se presentan tres máquinas de procesamiento distribuido que están siendo diseñadas y construidas en el Departamento de Computación (IIMAS) de la UNAM. Cabe mencionar que han surgido de ideas originales, de diseños

propios, no copiados de universidades o instituciones extranjeras.

Máquinas heterárquicas: Proyecto AHR. Esta arquitectura de cómputo consiste de varias docenas de microprocesadores, fuertemente conectados, que les permite ejecutar en paralelo un solo programa. El programa está escrito en LISP, por lo que es válido decir que la máquina AHR procesa LISP como lenguaje de máquina. Los datores comparten carga, la que es ejecutada en cada dator a nivel de un nodo. Básicamente un nodo es una instrucción primitiva de LISP. La arquitectura de hardware de la máquina es tal que sólo los nodos que pueden ser ejecutados en un momento dado son enviados a los datores para su ejecución. Todos los datores tienen acceso común a la memoria de variables (donde residen los diferentes valores de las variables de LISP), a la memoria pasiva (donde yacen los datos y los programas que no están activos) y a la parrilla (donde los programas activos están siendo evaluados o procesados).

Los datores no poseen memoria masiva ni teclados. Hay una conexión con una minicomputadora (llamada "caja inteligente") que es la que atiende a los usuarios y posee periféricos. Es decir, la máquina AHR puede considerarse como un periférico de esta minicomputadora. Este "periférico" tiene la habilidad de transformar programas (en LISP) en resultados.

Datores en un cable coaxial. El Proyecto de Sistemas Distribuidos (IIMAS) de la UNAM contempla la conexión a alta velocidad de varios datores homogéneos, cada uno de los cuales posee memoria masiva, teclado y pantalla. Se prevee que cada sistema operativo sea de multiprogramación. Se planea en la nivelación de cargas mediante la exportación de trabajos de los datores más ocupados a los menos atareados. Es decir, la arquitectura tendrá datores promiscuos, así como también periféricos

promiscuos. El sistema es fuertemente acoplado. Se usa paralelismo entre trabajos, no dentro de cada trabajo (como lo hace la máquina AHR). Probablemente exista un control descentralizado de el sistema de dadores.

Este proyecto, incipiente aún, promete ser un paradigma para el diseño y evaluación de diferentes ideas en el desarrollo de sistemas distribuidos comerciales y operacionales.

Máquina de bases de datos. Este proyecto contempla la construcción de un dator especializado, capaz de manejar eficientemente una base de datos mediana o grande. Se han analizado las operaciones principales que se efectúan en el manejo cotidiano de un sistema de información, y se ha propuesto la construcción de hardware especial (mergers, sorters, buscadores, memorias asociativas, direccionadores por contenido, memorias de caché, memorias de burbujas, dispositivos de cargas acopladas, etc.) y la modificación de equipos electrónicos ya existentes, para adecuarlos a las necesidades de la máquina de bases de datos.

Probablemente, la máquina de bases de datos se conectará como un periférico de un dator general, y funcionará como un procesador de "back-end", en tanto que la atención a los usuarios la realiza el dator general, ayudado y en coordinación con el sistema total de dadores. Es decir, la máquina de bases de datos se conectará al resto del sistema de dadores en forma análoga a la conexión de la máquina AHR a su "caja inteligente".

## CONCLUSIONES

Los sistemas de procesamiento distribuido ofrecen reducciones sustanciales en costo, así como ventajas en versatilidad, confiabilidad, crecimiento incremental, y disponibilidad,



por lo que se espera que su uso vaya en aumento constante y desplace a los sistemas tradicionales basados en una computadora central que comparte su tiempo entre usuarios que emplean terminales clásicas.

#### AGRADECIMIENTOS

El autor agradece la invitación del Dr. Héctor Nava Jaimes, Jefe del Departamento de Ingeniería Eléctrica del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional, a participar temporalmente como Profesor Visitante en su Departamento. Este trabajo se desarrolló en el CIEA del IPN, después de fructíferas discusiones con el Dr. Juan Milton Garduño, Profesor de Telecomunicaciones del mencionado Departamento, y con otros colaboradores del CIEA.

Los proyectos "Máquina AHR", "Computación Distribuida" y "Máquina para bases de datos" fueron originados por el autor en el Departamento de Computación del Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (UNAM), y deben su existencia a la asiduidad y constancia de sus integrantes, y al apoyo del Dr. Tomás Garza, Director del Instituto. El autor reconoce la valiosa ayuda prestada por los integrantes de estos proyectos (Luis Lyons, Luis Hugo Peñarrieta, Dora Gómez, Renato Barrera, David Rosenblueth, Kemer Norkin, Angel Kuri, entre otros), vertida a través de numerosas discusiones e innumerables tazas de café.

El intercambio de correspondencia con el Prof. Carl Hewitt, del Laboratory for Computer Science del Massachusetts Institute of Technology, quien ha popularizado el concepto de "actores", ha sido también influyente en las ideas de este artículo.

El trabajo aquí descrito ha sido patrocinado parcialmente por el Consejo Nacional de Ciencia y Tecnología (Proyecto 1632), México.

#### BIBLIOGRAFIA

1. Tutorial on Distributed Processing. Second Edition. Burt H. Liebowitz, John H. Carson. IEEE Computer Society. 5855 Naples Plaza. Long Beach, Cal. Publicación No. EHO 127-1. 1978.
2. Jon Roland. Microtecnología para las masas. Ciencia y Desarrollo 28, Sept.-Oct. 1979, 125-130.
3. Carl Hewitt. Design of the APIARY for knowledge-based systems. Working paper 186 B, July 1979. Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.
4. Velarde, C. y Rosenblueth, D. La máquina AHR para procesamiento en paralelo, etapa 1. Comunicación Técnica AHR-79-2. IIMAS-UNAM. 1979.
5. Guzmán, A. and Segovia, R. A parallel configurable LISP machine. Proceedings of the International Conference on Systems Sciences. Patras, Greece. 1976. Also available as: Comunicaciones Técnicas IIMAS No. 133, Vol. 7. 1976. IIMAS-UNAM. (Technical report AHR-76-1).