Taller 1

Ventajas y desventajas de los SGBDD

Sistemas de Bases de Datos, Thomas M. Connolly

Capítulo 22 Bases de datos distribuidas: conceptos y diseño

**Ventajas**

Refleja la estructura organizativa

Muchas organizaciones están distribuidas de forma natural en diversas ubicaciones. Por ejemplo, ACME tiene muchas oficinas en diversas ciudades. Resulta natural que las bases de datos utilizadas en dicho tipo de aplicaciones estén distribuidas entre esas diferentes ubicaciones. ACME (venta y alquiler de inmuebles) puede mantener una base de datos en cada sucursal donde se contengan los detalles relativos a cosas tales como el personal que trabaja en dicha oficina, los inmuebles que tiene en alquiler, los clientes que poseen o quieren alquilar dichos inmuebles. El personal que trabaje en una sucursal puede realizar consultas locales a la base de datos. El personal de las oficinas centrales puede necesitar realizar consultas globales que requieran acceder a los datos de todas las sucursales o de un número significativo de ellas.

Mejora la compartición de los datos y la autonomía local

La distribución geográfica de una organización puede reflejarse en la distribución de los datos; los usuarios de uno de los nodos pueden acceder a los datos almacenados en los otros nodos. Los datos pueden colocarse en el nodo más próximo a los usuarios que normalmente los usan. De esta forma, los usuarios tienen un control local de los datos y pueden, en consecuencia, establecer e imponer políticas locales relativas al uso de estos

datos. Un administrador global de la base de datos será responsable del sistema completo. Generalmente, parte de esta responsabilidad se delega en el nivel local, de modo que el DBA local puede gestionar el SGBD local

Mayor disponibilidad

En un SGBD centralizado, un fallo de la computadora hace que se detengan las operaciones del SGBD. Sin embargo, un fallo en uno de los nodos de un SGBDD, o un fallo de un enlace de comunicaciones que haga que algunos nodos sean inaccesibles, no hace que todo el sistema deje de operar. Los SGBD distribuidos están diseñados para continuar funcionando a pesar de tales fallos. Si se produce un fallo en un único nodo, puede que el sistema sea capaz de reencaminar las solicitudes dirigidas al nodo fallido hacia otro nodo.

Mayor fiabilidad

Ya que los datos pueden replicarse, de modo que estén almacenados en más de un nodo, el fallo de un nodo o de un enlace de comunicaciones no hace necesariamente que los datos dejen de estar accesibles.

Mayores prestaciones

Ya que los datos están localizados cerca del nodo de 'mayor demanda', y dado el inherente paralelismo de un SGBD distribuido, la velocidad del acceso a la base de datos puede ser superior que la que podría alcanzarse utilizando una base de datos centralizada remota. Además, dado que cada nodo gestiona únicamente una parte

de la base de datos completa, puede que no haya el mismo grado de contienda por el procesador y por los servicios de E/S que el que se experimenta en un SGBD centralizado.

Economía

En la década de 1960, la potencia de procesamiento se calculaba según el cuadrado del coste de los equipos: multiplicando el coste por tres, se incrementaba por nueve la potencia de procesamiento. Esto se conocía con el nombre de **Ley de Grosch**. Sin embargo, hoy en día se acepta con carácter general que resulta mucho más barato crear un sistema de pequeñas computadoras con la potencia equivalente de una única computadora de gran tamaño. Esto hace que sea más eficiente en términos económicos que las divisiones y departamentos corporativos adquieran computadoras independientes. También es mucho más barato añadir estaciones de trabajo a una red que actualizar un sistema mainframe. La segunda área potencial de ahorro de costes está relacionada con aquellas situaciones en las que las bases de datos son geográficamente remotas y las aplicaciones necesitan acceder a datos distribuidos. En tales casos, debido al alto coste de transmisión de los datos a través de la red (si lo comparamos con el coste del acceso local), puede ser mucho más económico particionar la aplicación y realizar el procesamiento localmente en cada nodo.

Crecimiento modular

En un entorno distribuido, resulta mucho más sencillo expandir el sistema. Pueden añadirse nuevos nodos a la red sin afectar a las operaciones de otros nodos. Esta flexibilidad permite a una organización expandirse de forma relativamente simple. Para incrementar el tamaño de la base de datos, basta con añadir potencia de procesamiento

y espacio de almacenamiento a la red. En un SGBD centralizado, el crecimiento puede necesitar de cambios en el HW (la compra de un sistema más potente) y en el software (la adquisición de un SGBD más potente o más configurable)

Integración

La integración de los sistemas constituye un ejemplo concreto que demuestra cómo algunas organizaciones están forzadas a utilizar un procesamiento de datos distribuido para que sus sistemas heredados coexistan con los otros sistemas más modernos. Al mismo tiempo, ningún paquete software puede proporcionar toda la funcionalidad que una organización requiere hoy en día. Por tanto, es importante que las organizaciones puedan integrar componentes software de diferentes fabricantes para satisfacer sus requisitos específicos.

Capacidad de competir

Diversos desarrollos relativamente recientes dependen en buena medida de la tecnología de bases de datos distribuidas. Este es el caso, del e-Business, de la gestión de flujos de trabajo y de los sistemas de trabajo en colaboración basados en computadora. Muchas empresas han tenido que reorganizar sus operaciones y utilizar tecnología de bases de datos distribuidas para seguir siendo competitivas. Por ejemplo, aunque no es probable

que haya más personas que alquilen inmuebles simplemente debido a la existencia de Internet, ACME puede perder parte de su cuota de mercado si no permite a los clientes ver los inmuebles en línea.

**Desventajas**

Complejidad

Un SGBD distribuido que oculte la naturaleza distribuida de los datos a ojos del usuario y que proporcione un nivel aceptable de rendimiento, fiabilidad y disponibilidad es inherentemente más complejo que un SGBD centralizado. El hecho de que los datos puedan replicarse añade también un nivel adicional de complejidad al SGBD distribuido. Si el software no gestiona adecuadamente la replicación de los datos, habrá una degradación en la disponibilidad, la fiabilidad y las prestaciones por comparación con un sistema centralizado, y las ventajas que antes hemos citado se transformarán en desventajas.

Coste

La mayor complejidad implica que cabe esperar que los costes de adquisición y de mantenimiento de un SGBDD sean mayores que para un SGBD centralizado. Además, un SGBD distribuido requiere un hardware adicional para implantar la red de comunicaciones entre los distintos nodos. Con el uso de esta red, además,

existen una serie de costes de comunicación continuados. Por último, tenemos que tener en cuenta los costes adicionales de personal requeridos para gestionar y mantener los SGBD locales y la red subyacente.

Seguridad

En un sistema centralizado, el acceso a los datos puede controlarse fácilmente. Por el contrario, en un SGBD distribuido no sólo es necesario controlar el acceso a los datos replicados en múltiples ubicaciones, sino que también es necesario dotar de seguridad a la propia red. En el pasado, las redes se consideraban un medio de comunicación inseguro. Aunque esto continua siendo parcialmente cierto, diversos desarrollos de gran importancia han incrementado considerablemente el nivel de seguridad de las redes.

Control de integridad más complicado

La integridad de la base de datos es un término que hace referencia a la validez y coherencia de los datos almacenados. La integridad se expresa normalmente en términos de restricciones, que son reglas de coherencia que no se permite que la base de datos viole. Imponer las restricciones de integridad requiere, generalmente, acceder

a una gran cantidad de datos que definen la restricción pero que no están implicados en la propia operación de actualización. En un SGBD distribuido, los costes de comunicaciones y de procesamiento requeridos para imponer las restricciones de integridad pueden ser prohibitivos.

Carencia de estándares

Aunque los SGBD distribuidos dependen de que existan unas comunicaciones efectivas, sólo ahora estamos comenzando a ser testigos de la aparición de protocolos de comunicaciones y de acceso a datos estándar. Esta falta de estándares ha inundado considerablemente el potencial de los SGBD distribuidos. Asimismo, tampoco

hay herramientas o metodologías que ayuden a los usuarios a convertir un SGBD centralizado en un SGBD distribuido.

Falta de experiencia

Los SGBD distribuidos de propósito general no han sido aceptados ampliamente, aunque muchos de los protocolos y de los problemas se comprendan a la perfección. En consecuencia, todavía no tenemos el mismo nivel de experiencia en este sector que en el caso de los SGBD centralizados. Para alguien que esté pensando en adoptar esta tecnología, este factor puede ser uno de los obstáculos principales.

Diseño de la base de datos más complejo

Además de las dificultades normales de diseñar una base de datos centralizada, el diseño de una base de datos distribuida tiene que tener en cuenta la fragmentación de los datos, la asignación de fragmentos en los diferentes nodos y las cuestiones de replicación de los datos.

