

Unidad I

Fundamentos de bases de datos distribuidas.

1.1 Conceptos básicos.

DEFINICION DE BASE DE DATOS DISTRIBUIDAS: Es una colección de datos que pertenecen lógicamente a un sólo sistema, pero se encuentra físicamente esparcido en varios sitios de la red. Un sistema de base de datos distribuidas se compone de un conjunto de sitios, conectados entre sí mediante algún tipo de red de comunicaciones, en el cual :

1. cada sitio es un sistema de base de datos en sí mismo, pero
2. los sitios han convenido en trabajar juntos (si es necesario) con el fin de que un usuario de cualquier sitio pueda obtener acceso a los datos de cualquier punto de la red tal como si todos los datos estuvieran almacenados en el sitio propio del usuario.

El procesamiento de bases de datos distribuidas es el proceso en el cual la ejecución de transacciones y la recuperación y actualización de los datos acontece a través de dos o más computadoras independientes, por lo general separadas geográficamente.

Las Bases de Datos Distribuidas , no son simplemente implementaciones distribuidas de bases de datos centralizadas, por que ellas permiten el diseño de sistemas que representan diferentes características de las tradicionales, de sistemas centralizados. Esto es por lo tanto útil para ver las características típicas de BDD. Los rasgos que caracterizan las BD tradicionales se aproximan al control centralizado, independencia de datos, reducción de redundancia, estructuras físicas complejas para acceso eficiente, integridad, recuperación control de concurrencia, privacidad y seguridad.

Los principales factores que distinguen un SBDD de un sistema centralizado son los siguientes:

- Hay múltiples computadores, llamados sitios o nodos.
- Estos sitios deben de estar comunicados por medio de algún tipo de red de comunicaciones para transmitir datos y órdenes entre los sitios

1.2 Objetivos de las B.D.D.

Transparencia de ubicación. Permite a los usuarios tener acceso a los datos sin que tenga conocimiento de la ubicación de éstos.

Transparencia de duplicación. Para que la transparencia de duplicación sea posible, los administradores de transacciones deben

traducir las solicitudes de procesamiento de transacción en acciones para el administrador de datos.

Transparencia de concurrencia. Cuando varias transacciones se ejecuten al mismo tiempo, los resultados de las transacciones no deberán afectarse.

Transparencia de fallas. Significa que a pesar de fallas las transacciones sean procesadas de un modo correcto. Frente a una falla, las transacciones deben ser atómicas, significa que se procesen todas o ninguna de ellas.

Localidad del procesamiento. Los datos se deben distribuir lo más cerca posible de las aplicaciones que los usan para maximizar la localidad del procesamiento, este principio responde a minimizar el acceso remoto a los datos.

Particionado de la Base de Datos. La base de datos se distribuye de modo que no haya solapamiento o duplicación de los datos mantenidos en las diferentes localidades, como no hay duplicaciones de los datos, se evitan los costos asociados con el almacenamiento y mantenimiento de datos redundantes.

Fragmentación de datos. Consiste en subdividir las relaciones y distribuir las entre los sitios de la red, tiene como objetivo buscar formas alternativas de dividir una las instancias de relaciones en otras más pequeñas.

1.3 Disciplinas de estudio.

Las principales disciplinas de estudio para conocer las bases son las siguientes:

A) INGENIERIA: Para conocer como se desarrollan y que forma tendran para implementarse

B) ALGEBRA: Busca establecer relaciones en base a funciones algebraicas

C) BASES DE DATOS: Buscan un adecuado funcionamiento de acuerdo a los

principios de estas

D) REDES: Implementado en adecuado sistema para su funcionamiento sin concurrencia

1.4 Arquitectura de bases de datos distribuidas

Podemos destacar tres niveles principales según la visión y la función que realice el usuario sobre la base de datos:

- Nivel Interno: es el nivel más cercano al almacenamiento físico de los datos. Permite escribirlos tal y como están almacenados en el ordenador. En este nivel se diseñan los archivos que contienen la información, la ubicación de los mismos y su organización, es decir se crean los archivos de configuración.
- Nivel conceptual: En este nivel se representan los datos que se van a utilizar sin tener en cuenta aspectos como lo que representamos en el nivel interno.
- Nivel externo: es el más cercano al usuario. En este nivel se describen los datos o parte de los datos que más interesan a los usuarios.

Estos tres niveles de visión de usuarios los proporcionan los sistemas gestores de base de datos.

