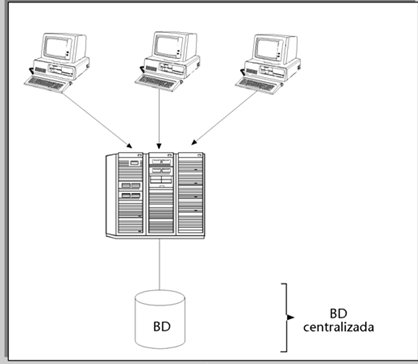
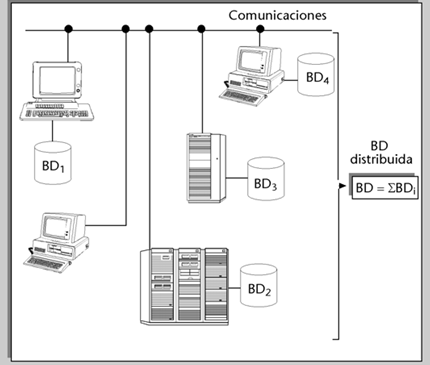
BASES DE DATOS CENTRALIZADAS Y DISTRIBUIDAS.



CENTRALIZADAS

Una base de datos centralizada es una base de datos almacenada en su totalidad en un solo lugar físico, es decir, es una base de datos almacenada en una sola máquina y en una sola CPU. Los sistemas de bases de datos centralizadas son aquellos que se ejecutan en un único sistema informático sin interaccionar con ninguna otra computadora.

DISTRIBUIDAS

Cuando una BD está controlada por más de un servidor se dice que está distribuida. Estos servidores están interconectados mediante una red de datos y los SGBD proporcionan mecanismos para poder consultar la información independientemente del servidor y de la ubicación de los datos.

Ventajas de las BD distribuidas

Mejora del rendimiento: Los sitios no se ven congestionados por muchas transacciones, ya que estas se dispersan entre varios sistemas. Así cuando una transacción requiere acceso a más de un sitio este acceso puede hacerse en paralelo.

Fiabilidad: Aunque algún nodo de la BD sufra algún percance, el resto del sistema seguirá funcionando con normalidad.

Acceso más flexible: el número de usuarios conectado que puede mantener el sistema se incrementa al haber más capacidad de cómputo y, gracias a la red de datos, los usuarios pueden acceder desde más ubicaciones.

Escalabilidad: Al distribuir la BD en distintas ubicaciones se permite incrementar de forma muy sencilla los recursos en cualquier momento para acomodar el sistema a las demandas de los usuarios y aplicaciones.

Inconvenientes de las BD distribuidas

Costosos de implementar (duplicidad de recursos), dificultad en la administración (datos ubicados en distintas localizaciones) y además se tiene una dependencia de las redes de datos para el correcto funcionamiento del intercambio de datos

Componentes de las BD distribuidas

Un sistema de gestión de bases de datos distribuidas (SGBDD) incorpora, además de los componentes habituales en un SGBD centralizado, las siguientes funcionalidades adicionales:

* Sistemas gestores con capacidades distribuidas.
* Un conjunto de bases de datos locales a cada uno de los servidores que albergarán la BBDD
* Una red de comunicaciones: accede a sitios remotos y transmite consultas y datos a través de varios sitios mediante [una red](http://www.monografias.com/Computacion/Redes/) de comunicación.
* Diccionario de datos global. Almacena el esquema de distribución y replicación de los datos en el catálogo del sistema. Así se conoce la ubicación de cada uno de los datos que componen la BD distribuida.
* Enlaces entre las bases de datos locales. Establece las estrategias de ejecución de las consultas y las transacciones que acceden a los datos en más de un sitio.
* Mantiene la integridad en los permisos sobre datos replicados.
* Mantiene la consistencia de las copias de los datos replicados.
* Garantiza la recuperación del sistema en una caída.

Tipos de [bases de datos](http://www.monografias.com/trabajos11/basda/basda.shtml) distribuidas

Las BDD pueden ser:

* Homogéneas: Todos los sitios tienen el mismo SGBD, son conscientes de la existencia de los demás sitios y cooperan en el procesamiento de las solicitudes. Los sitios locales mantienen un mismo esquema y SGBD.
* Heterogéneas: Cada sitio puede tener un SGBD distinto así como esquemas diferentes. En este caso, las aplicaciones que consultan la base de datos distribuida juegan un papel fundamental como mediador.

[Problemas](http://www.monografias.com/trabajos15/calidad-serv/calidad-serv.shtml#PLANT) fundamentales a resolver en las bases de [datos](http://www.monografias.com/trabajos11/basda/basda.shtml) distribuidas

* Diseño de bases de datos distribuidas
* Procesamiento y optimización de consultas
* Manejo de transacciones y [control](http://www.monografias.com/trabajos14/control/control.shtml) de concurrencia

DISEÑO

La etapa diferenciadora entre el [diseño](http://www.monografias.com/trabajos13/diseprod/diseprod.shtml) de una [base de datos](http://www.monografias.com/trabajos34/base-de-datos/base-de-datos.shtml) centralizada y una base de datos distribuida es el "Diseño de la [distribución](http://www.monografias.com/trabajos11/travent/travent.shtml)" que consta de dos actividades:

Fragmentación: Decidir "como" dividimos la BD y en "que" partes.

Asignación: Decidir "donde" ubicamos cada parte, así como si tendremos replicación de datos.

***Fragmentación***

El problema de fragmentación se refiere a cómo vamos a partir la [información](http://www.monografias.com/trabajos7/sisinf/sisinf.shtml) para distribuir cada parte a los diferentes sitios de la red

Una fragmentación óptima es aquella que produce un esquema de división que minimiza el [tiempo](http://www.monografias.com/trabajos901/evolucion-historica-concepciones-tiempo/evolucion-historica-concepciones-tiempo.shtml) de ejecución de las aplicaciones que emplean esos fragmentos.

***Tipos de fragmentación de datos***

Existen tres tipos de fragmentación

* Fragmentación horizontal: divide las tuplas de una tabla según una o varias condiciones sobre distintos atributos. La tabla original se reconstruye en base a una operación de unión de los fragmentos.
* Fragmentación vertical: Divide la información de una tabla en función por ejemplo de sus campos más accedidos y ligeros (los almacenaría en un servidor más potente) mientras que los otros campos más pesados se podrían almacenar en otro servidor menos potente pero con mayor capacidad de almacenamiento. Para recomponer la tabla original cada fragmento debe incluir la tabla primaria de la tabla.
* Fragmentación híbrida: incluye los dos casos anteriores.

La fragmentación será correcta si cumple:

* Debe ser completa, es decir, cada elemento de la tabla debe estar en algún fragmento.
* Debe ser reconstruible, se debe poder conseguir una tabla completa a partir de la combinación de todos los fragmentos.
* Un elemento que esté en un fragmento concreto no puede estar en ningún otro fragmento más que en aquél.

***Asignación de fragmentos***

Proceso mediante el cual se decide donde se ubicaran los fragmentos de la etapa anterior y si se harán réplicas de los mismos.

Hacer réplicas tiene sentido por:

* Confiabilidad : Mayor [seguridad](http://www.monografias.com/trabajos/seguinfo/seguinfo.shtml) ante perdida de datos
* Disponibilidad : Mayor [tolerancia](http://www.monografias.com/trabajos11/tole/tole.shtml) a fallos ante caídas de los computadores
* Aumento del paralelismo: Mayor [eficiencia](http://www.monografias.com/trabajos11/veref/veref.shtml) en las consultas de [lectura](http://www.monografias.com/trabajos14/textos-escrit/textos-escrit.shtml) al posibilitarse su descomposición en subconsultas y la paralelización de éstas.

Pero presenta el inconveniente de las consultas de [escritura](http://www.monografias.com/trabajos16/metodo-lecto-escritura/metodo-lecto-escritura.shtml), que conllevan las actualizaciones de todas las copias de la [red](http://www.monografias.com/Computacion/Redes/).

En la práctica: Sopesar lecturas vs escrituras

* Mas lecturas que escrituras : Replicación es buena idea
* Mas escrituras que lecturas: No replicamos.

Según el grado de replicación, distinguimos entre:

* BD fragmentada : Fragmentos disjuntos, cada uno en un nodo (no hay replicas)
* BD totalmente replicada : Se encuentra una copia de toda la BD en cada nodo
* BD parcialmente replicada: Mezcla las anteriores. Algunos fragmentos están replicados.

Como se ve las [técnicas](http://www.monografias.com/trabajos6/juti/juti.shtml) de fragmentación y replicación se combinan en la práctica.

***Replicación de fragmentos***

El problema de la replicación de segmentos asignación consiste en la determinación de que fragmentos se replicarán en diferentes sitios a pesar de los problemas que acarrea la actualización.