1. ¿Qué es big data?

la tendencia en el avance de la tecnología que ha abierto las puertas hacia un nuevo enfoque de entendimiento y toma de decisiones, la cual es utilizada para describir enormes cantidades de datos (estructurados, no estructurados y semi estructurados) que tomaría demasiado tiempo y sería muy costoso cargarlos a un base de datos relacional para su análisis.

<https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>

youtube: <https://www.youtube.com/watch?v=Ju2oDsHAL-o>

1. Rendimiento de las bases de datos
2. ¿Qué es Datawarehouse?

Un **Datawarehouse** es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de pespectivas y con grandes velocidades de respuesta.

<http://www.sinnexus.com/business_intelligence/datawarehouse.aspx>

youtube: <https://www.youtube.com/watch?v=jFsRdTcljeU>

1. Inteligencia de negocios y minería de datos

Datamining (Minería de datos)

El **datamining** (*minería de datos*), es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.

http://www.sinnexus.com/business\_intelligence/datamining.aspx

Youtube: <https://www.youtube.com/watch?v=kXEa6l2Nh50>

1. Datamart

Youtube: <https://www.youtube.com/watch?v=Pz_V1CQMvYQ>

SISTEMA DE BASE DE DATOS DISTRIBUIDA

Una base de datos distribuida (DDB) es aquella en la que sus datos se encuentran repartidos en varias computadoras que est´an interconectadas por una red de comunicaciones. Un sistema Administrador de Base de Datos Distribuidas (DDBMS) es el software que administra la DDB y proporciona un mecanismo de acceso que hace transparente esta distribuci´on a los usuarios.

* Los datos se almacenan en una serie de sitio. Cada sitio consta de su propio procesador
* Los procesadores de cada sitio estan interconectados por una red de comunicaci´on (no multiprocesadores). 1 Sistemas de base de datos paralelo.
* Una base de datos distribuida es una base de datos, no una coleccion de archivos (datos l´ogicamente relacionados que son presentados al usuario con un patr´on de acceso). 1 Modelo de datos relacional.
* Un DDBMS es un DBMS completo. 1 No es un sistema de archivos remotos, ni un sistema TP.

YOUTUBE: <https://www.youtube.com/watch?v=GrdcvdpVvvQ>

# Rendimiento de la base de datos

Otras versiones



Cuando diseñe una base de datos, debe asegurarse de que realiza todas las operaciones importantes de forma rápida y correcta. Algunos problemas de rendimiento se pueden resolver una vez que la base de datos se encuentra en producción. Sin embargo, otros pueden ser el resultado de un diseño inadecuado y se pueden solucionar mediante el cambio de la estructura y el diseño de la base de datos.

Cuando diseña e implementa una base de datos, debe identificar las tablas de gran tamaño y los procesos más complejos que realizará la base de datos. También debe prestar una atención especial al rendimiento cuando diseña estas tablas. Además, debe considerar los efectos que puede tener en el rendimiento el aumento del número de usuarios con acceso a la base de datos.

Los siguientes cambios de diseño, entre otros, pueden mejorar el rendimiento:

* Si una tabla que contiene cientos de miles de filas debe resumirse en un informe diario, puede agregar a la tabla una o varias columnas que contengan datos previamente agregados para utilizarlos sólo en dicho informe.
* Las bases de datos pueden normalizarse en exceso. Esto significa que la base de datos se define con un gran número de tablas pequeñas interrelacionadas. Cuando la base de datos procesa los datos de estas tablas, debe realizar muchas más operaciones para combinar los datos relacionados. Este procesamiento adicional puede repercutir negativamente en el rendimiento de la base de datos. En esos casos, una reducción de la normalización de la base de datos para simplificar procesos complejos puede mejorar el rendimiento.

YOUTUBE: