

Data Warehouse

1. Introducción:

Hoy en día toda empresa necesita depositar mucha confianza en la toma de decisiones sobre los negocios por lo que requerimos hechos y cifras, con las que tomaremos decisiones en nuestra empresa y estas deberán ser mas aceleradas; por este motivo se requieren herramientas que nos ayuden a minimizar el tiempo para analizar mucha información con mayor velocidad y precisión; El Componente de Bussines Intelligence que resuelve este caos de los datos es el **Data Warehouse** el cual es un conjunto de procesos y acciones, es una colección de datos orientados a un tema, integrados y no volátiles en el soporte al proceso de toma de decisiones de la gerencia

2. Data Warehouse:

Algunos lo llaman a Data Warehouse Business Intelligence or Decision Support en realidad es considerada la solución integral y oportuna para desarrollar negocio el Datawarehouse; es un proceso, no un producto y es una técnica para consolidar y administrar datos de variadas fuentes con el propósito de responder preguntas de negocios y tomar decisiones, de una forma que no era posible hasta ahora, se caracteriza por ser:

Integrado - Temático - Histórico - No volátil

Características:

1. Consolidar datos desde una variedad de fuentes. Dentro del marco conceptual de Data Warehousing los agruparemos dentro del proceso de Transformación de Datos.
2. Manejar grandes volúmenes de datos de una forma que no era posible, o no era costo efectiva. A estos medios los agruparemos en Procesamiento y Administración de Datos.
3. Acceder a los datos de una forma más directa, en "el lenguaje del negocio", y analizarlos para obtener relaciones complejas entre los mismos. Estos procesos se engloban en dos categorías que serán explicadas más adelante: Acceso a los Datos y Descubrimiento o Data Mining (Un sistema Datamining es una tecnología de soporte para usuario final, cuyo objetivo es extraer conocimiento útil y utilizable a partir de la información contenida en las bases de datos de las empresas; sus objetivos nos permitirán analizar factores de influencia en determinados procesos, predecir o estimar variables o comportamientos futuros, sementar o agrupar ítems similares, además de obtener secuencias de eventos que provocan comportamientos específicos).

Un Data Marts es un pequeño Data Warehouse, para un determinado numero de usuarios, para un área funcional, especifica de la compañía y un subconjunto de una bodega de datos para un propósito específico, su función es apoyar a otros sistemas para la toma de decisiones.

Los procesos que conforma el datawarehouse son: Extracción, Elaboración Carga y Explotación. El éxito de DW no está en su construcción, sino en usarlo para mejorar procesos empresariales, operaciones y decisiones. Posesionar un DW para que sea usado efectivamente, requiere entender los impactos de

implementación en los siguientes ámbitos: Impactos Humanos, impactos empresariales e impactos técnicos.

3. Impactos empresariales Se deben considerar los beneficios empresariales de los siguientes impactos:

- a) Los Procesos de Toma de Decisiones pueden ser mejorados mediante la disponibilidad de información.
- b) Los procesos empresariales pueden ser optimizados. El tiempo perdido esperando por información que finalmente es incorrecta o no encontrada, es eliminada.
- c) Conexiones y dependencias entre procesos empresariales se vuelven más claros y entendibles.
- d) Procesos y datos de los sistemas operacionales, así como los datos en el DW, son usados y examinados. Cuando los datos son organizados y estructurados para tener significado empresarial, la gente aprende mucho de los sistemas de información.

4. ¿Quiénes y para qué lo usan?

Dadas las características de un sistema de Data Warehousing, su aplicación puede tener variados fines, en una diversidad de industrias. No obstante, en términos generales, podemos decir que su aplicación más rica corresponde a entornos de empresas en los que se identifican grandes volúmenes de datos, asociados a: cantidad de clientes, variedad de productos y cantidad de transacciones. Veremos algunas aplicaciones típicas en distintas industrias.

Comercio Minorista: Utilizan grandes sistemas de Procesamiento Paralelo Masivo para acceder a meses o años de historia transaccional tomada directamente en los puntos de venta de cientos, o miles, de sucursales.

Manufactura de Bienes de Consumo Masivo: Las empresas de este sector necesitan hacer un manejo cada vez más ágil de la información para mantenerse competitivas en la industria. Los Data Warehouses se utilizan para predecir la cantidad de producto que se venderá a un determinado precio y, por consiguiente, producir la cantidad adecuada para una entrega "justo a tiempo".

Transporte de Cargas y Pasajeros: Se utilizan Data Warehouses para almacenar y acceder a meses o años de datos de clientes y sistemas de reservas para realizar actividades de marketing, planeamiento de capacidad, monitoreo de ganancias, proyecciones y análisis de ventas y costos, programas de calidad y servicio a clientes.

Telecomunicaciones: Estas empresas utilizan sus Data Warehouses para operar en un mercado crecientemente competitivo, desregulado y global que, a su vez, atraviesa profundos cambios tecnológicos. Se almacenan datos de millones de clientes

Muchas otras industrias y actividades utilizan actualmente, o están comenzando a instalar, Data Warehouses: entidades gubernamentales,

especialmente para el control impositivo, empresas de servicios públicos, de entretenimiento, editoriales, fabricantes de automóviles, empresas de petróleo y gas, laboratorios farmacéuticos, droguerías, etc.

5. Sistemas De Data Warehouse Y Oltp

Una base de datos para soportar procesos transaccionales en línea (OLTP), puede no ser adecuada para el Data Warehouse ya que ha sido diseñada para maximizar la capacidad transaccional de sus datos y típicamente tiene cientos de tablas la gran mayoría normalizadas. Su diseño también ha sido condicionado por los procesos operacionales que deberá soportar para la óptima actualización de sus datos, normalmente muchas de sus tablas en constantes y continuos cambios. Los sistemas Data Warehouse están orientados a procesos de consultas en contraposición con los procesos transaccionales.

Transforman los datos en información para su análisis. Perdurabilidad Los datos cambian constantemente, vistos globalmente en procesos de reporting sofisticados pueden perder consistencia, o bien, para no perder consistencia deben imponerse mecanismos de bloqueo de datos con un elevado consumo de recursos globales del sistema. Datos históricos con referencias temporales no sujetos a modificaciones.

6. Diferencias: Data Warehouse vs. OLTP

Los sistemas tradicionales de transacciones y las aplicaciones de Data Warehousing son polos opuestos en cuanto a sus requerimientos de diseño y sus características de operación. Es de suma importancia comprender perfectamente estas diferencias para evitar caer en el diseño de un Data Warehouse como si fuera una aplicación de transacciones en línea (OLTP).

Las aplicaciones de OLTP están organizadas para ejecutar las transacciones para los cuales fueron hechos, como por ejemplo: mover dinero entre cuentas un cargo o abono, una devolución de inventario, etc. Por otro lado, un Data Warehouse está organizado en base a conceptos, como por ejemplo: clientes, facturas, productos, etc.

Otra diferencia radica en el número de usuarios. Normalmente, el número de usuarios de un Data Warehouse es menor al de un OLTP. Es común encontrar que los sistemas transaccionales son accedidos por cientos de usuarios simultáneamente, mientras que los Data Warehouse sólo por decenas. Los sistemas de OLTP realizan cientos de transacciones por segundo mientras que una sola consulta de un Data Warehouse puede tomar minutos. Otro factor es que frecuentemente los sistemas transaccionales son menores en tamaño a los Data Warehouses, esto es debido a que un Data Warehouse puede estar formado por información de varios OLTP's.

Integrantes del Equipo 4:

Karina Gabriela Bringas García, Raquel Reza Sánchez, Alfonso Villegas Vergara.