

INTRODUCCION

Las tecnologías de la información experimentan fundamentales y profundos impactos en la forma en que los productos y servicios son ofertados, distribuidos, vendidos y consumidos, porque el negocio real detrás de ellos es simple y llanamente información.

Las megatendencias están rigiendo el futuro de los negocios hoy en día, así tenemos:

- Internet / WWW
- Datawarehousing
- Computación móvil
- Comercio electrónico
- Año 2000

Pero las megatendencias también afectan el mercado, puesto que cada vez más empresas están ingresando al WWW. Además se evidencia la necesidad de invertir en tecnologías muchas veces de alto costo.

El temor de la globalización, la apertura de mercados y la alta competitividad están haciendo que las organizaciones vean con mayor frecuencia a la tecnología como un aliado estratégico. La necesidad de información en las organizaciones crece geométricamente.

Para toda empresa, obtener información cada día es más importante, puesto que la información (y con ella la técnica informática) ya es un factor de la producción y es determinante para el éxito que una empresa pueda tener en su mercado.

Ordenadores cada vez más rápidos y capacidades han llevado consigo una ola de información, cuya evaluación se hace cada vez más dificultosa. Las empresas poseen multitud de datos, pero muy pocas son capaces de aprovecharlos con eficacia.

Hoy en día las empresas cuentan en su mayoría con la automatización de sus procesos, manejando gran cantidad de datos en forma centralizada y manteniendo sus sistemas en línea. Esta información es el pilar primordial de la empresa, constituyendo un recurso corporativo primario y parte importante de su patrimonio.

El nivel competitivo alcanzado en las empresas les ha exigido desarrollar nuevas estrategias de gestión. En el pasado, las organizaciones fueron típicamente estructuradas en forma piramidal con información generada en su base fluyendo hacia lo alto; y era en el estrato de la pirámide más alto donde se tomaban decisiones a partir de la información proporcionada por la base, con un bajo aprovechamiento del potencial de esta información.

Estas empresas, han reestructurado y eliminado estratos de estas pirámides y han autorizado a los usuarios de todos los niveles a tomar mayores decisiones y responsabilidades. Sin embargo, sin información sólida para influenciar y apoyar las decisiones, la autorización no tiene sentido.

Esta necesidad de obtener información para una amplia variedad de individuos es la principal razón de negocios que conduce al concepto de *DATAWAREHOUSE*.

En el *DATAWAREHOUSE* el énfasis no está sólo en llevar la información hacia lo alto, sino que a través de la organización, para que todos los empleados que la necesiten la tengan a su disposición.

La revolución del *DATA WAREHOUSE (DW)* está impulsada por la esperanza de que esta aproximación será capaz de ofrecer a las personas que toman decisiones en la organización un acceso integrado, consistente, fiable y rápido a los datos, que les permite tomar decisiones basadas en una mejor información.

I. DATA WAREHOUSE

DATA WAREHOUSE es un concepto relativamente nuevo en Panamá y para muchos desconocido, que viene a resolver problemas de manejo y uso adecuado de grandes fuentes de datos y de diversos tipos, para apoyar tomas de decisiones oportunas.

El ambiente competitivo en las empresas de los 90s así como el avance tecnológico en materia de sistemas de información, han provocado un nuevo enfoque en el tratamiento y proceso de la información ejecutiva, la cual es un elemento vital hoy en día como soporte en el proceso de toma de decisiones.

El concepto *DATA WAREHOUSING*, o el proceso de contar con la información más importante de la empresa (incluyendo la histórica), en un solo lugar, ha logrado convertirse en una valiosa herramienta y clave desde el punto de vista tecnológico.

La explotación del *DATA WAREHOUSING* mediante información de gestión, se fundamenta básicamente en los niveles agrupados o calculados de información.

Este tipo de sistemas ha existido desde hace tiempo, en el mundo de la informática bajo distintas denominaciones como: cuadros de mando, MIS, EIS, etc.

Su realización fuera del entorno del *DATA WAREHOUSE*, puede repercutir sobre estos sistemas en una mayor rigidez, dificultad de actualización y mantenimiento, malos tiempos de respuestas, incoherencias de la información, falta del dato agregado, etc.

En un mundo cada vez más acelerado y competitivo, el acceso a la información operacional es una necesidad esencial en el proceso decisivo de los negocios. Una estructura adecuada de almacenamiento de datos va a satisfacer esta necesidad. También le va a proporcionar el acceso a la información actual y los datos históricos a través de diversas bases de datos corporativas y va a permitir una mejor toma de decisiones basadas en la información. Sustentada a menudo con información externa, tal como datos económicos y demográficos, las soluciones de almacenamiento de datos se convierten en el depósito central de los datos. Le da a los usuarios a través de toda la organización la capacidad de repasar, analizar, modificar e imprimir información.

1. Definición de *DATA WAREHOUSE*

Hoy en día se puede mencionar que prácticamente no existe negocio el cual implícitamente tenga asociado un Sistema de Información y que un sistema de información de una u otra forma termina siempre relacionado a un negocio.

La definición de *DATA WAREHOUSE* tiene múltiples vertientes, según W.H. Inmon, uno de los precursores del concepto de *DATA WAREHOUSE*:

Data Warehouse es un sistema orientado al usuario final, integrado, con variaciones de tiempo y sobre todo una colección de datos como soporte al proceso de toma de decisiones.

De acuerdo con algunas otras organizaciones, *DATA WAREHOUSE*, es una *arquitectura*. Para otros, es

simplemente un *almacén de datos* (separados y que no interfiere con los sistemas operativos actuales de una empresa) , para satisfacer las diversas consultas y requerimientos de información. Para algunos otros, *DW* es un proceso que agrupa datos desde múltiples fuentes heterogéneas, incluyendo datos históricos para soportar la continua necesidad de consultas, reportes analíticos y soporte de decisiones.

En definitiva *DATA WAREHOUSE* no es ni un producto de software ni una máquina, o tecnología de bases de datos en particular, sino una serie de componentes y procesos que en conjunto forman la arquitectura llamada *DATA WAREHOUSE*.

John Edwards en un artículo reciente de CIO Magazine, define *DATA WAREHOUSE* como:

un DW toma información de múltiples sistemas y bases de datos y la almacena de una manera que está diseñada para dar a los usuarios acceso más rápido, más fácil y más flexible a los aspectos claves

DATA WAREHOUSE, difiere de las bases de datos operacionales que soportan aplicaciones con el Proceso Transaccional en Línea (OLTP) de diversas formas. *DATA WAREHOUSE* incluye:

- Una orientación al usuario final.
- Administra y maneja un gran volumen de información
- Información somatizada y agregada
- Integra y asocia información desde múltiples fuentes y orígenes

Algunos consideran a *DATA WAREHOUSE* como un sistema de misión crítica. Debido a que las empresas actualmente demandan mayor información que sea oportuna, confiable y completa, desde un acceso directo y rápido a la información más importante, esto hace considerar a este tipo de sistemas como de Misión Crítica.

Los requerimientos principales de un sistema de Misión Crítica a considerar son:

- Disponibilidad
- Confiabilidad y consistencia
- Robustez
- Estándar
- Basado en los requerimientos del negocio
- Compatibilidad con tecnología actual y una infraestructura sólida
- Utilizado sobre bases diarias, es decir, para el proceso día a día
- Amigable
- Auditable
- Seguro

2. Propósito del DATA WAREHOUSE

El *propósito del DATA WAREHOUSE* es asistir al ejecutivo en el entendimiento del pasado y contar con los elementos para la planeación del futuro de corto, mediano y largo plazo.

Los ejecutivos y administradores buscan respuestas a preguntas como:

- ¿ Qué están comprando nuestros clientes? ¿Qué no están comprando?
- ¿ Qué está haciendo la competencia?
- ¿ Cómo están los costos por cada línea de producto, comparados con los últimos tres años?
- ¿ Qué factores causan incrementos en los costos?

Los ejecutivos quieren tener respuestas a preguntas cruciales para poder tomar decisiones.

Un **DATA WAREHOUSE** ayuda a resolver estas preguntas de forma eficiente y aporta elementos valiosos de toma de decisiones al personal encargado de las mismas, los cuales pueden ser analistas, ejecutivos o directores, y en forma inmediata, estos es, consultas en línea, sin necesidad de que el usuario final (el analista) solicite que se elaboren y ejecuten procesos especiales. Por lo tanto, es importante asegurar que los datos de la bodega sean adecuados, suficientes y seleccionados de acuerdo a las necesidades de decisión del negocio que se pretende soportar.

El objetivo principal de **DATA WAREHOUSE** es el uso adecuado de datos para obtener información útil para el soporte a la toma de decisiones, lo que es difícil de lograr con los sistemas operacionales (1,2,3) o sistema en línea.

El **DW** es el lugar donde la gente puede acceder sus datos

Con **DATA WAREHOUSE** se puede lograr:

Consultas más fáciles
Facilita la toma de decisiones
• Práctica de negocios inteligentes
• Obtener información valiosa y oportuna
• Impacto positivo sobre los procesos empresariales
• Ejecutivos mejor informados

Mediante las aplicaciones de **DW** se concentra e integra la información más importante de las organizaciones a través del tiempo, logrando hacer consultas de la información más fáciles y mucho más accesibles. **DW** convierte la información en **UTILIDADES**.

En primer lugar **DW**, no es un producto que pueda ser comprado en el mercado, sino más bien construido. **DW** es una combinación de conceptos y tecnología que cambian significativamente la manera en que es entregada la información a la gente de negocios. El objetivo principal satisfacer los requerimientos de la información interna de la empresa para una mejor gestión, con eficiencia y facilidad de accesos.

El **DW** puede verse como una bodega donde están almacenados todos los datos necesarios para realizar funciones de gestión de la empresa, de manera que puedan utilizarse fácilmente según se necesiten.

Los almacenes de datos (o **DATA WAREHOUSE**) generan bases de datos tangibles con una perspectiva histórica, utilizando datos de múltiples fuentes que se fusionan en forma congruente. Estos datos se mantienen actualizados, pero no cambian el ritmo de los sistemas transaccionales.

Muchos **DW** se diseñan para contener un nivel de detalle hasta el nivel de transacción, con la intención de hacer disponible todo tipo de datos y características, para reportar y analizar. Así un **DW** resulta ser un recipiente de datos transaccionales para realizar consultas operacionales. De esta forma, dentro de un almacén de datos existen dos tecnologías complementarias, una relacional para consultas y una multidimensional para análisis.

Si queremos extraer información de los datos estos deben ser:

- accesibles
- el acceso debe ser simple
- La información debe ser consistente

Si se desea que los datos tengan un significado, se deben hacer preguntas básicas como:

- ¿Cuál es la situación actual?
- ¿Cómo llegue aquí?
- ¿Cómo predecir el futuro?
- ¿Qué implica esta tendencia?

II. VENTAJAS DEL DATA WAREHOUSE

- Beneficios

Para los usuarios:

Confianza en el uso de sistemas: El incremento en la rapidez de las consultas afianzará el uso de los sistemas operativos de la empresa.

Reducción en tiempos de consultas: Se adquiere una solución que reduce tiempos de espera por parte de los usuarios.

Apoyo a la toma de decisiones: Los empresarios tendrán a la mano los resúmenes de información para armar reportes comparativos, utilizando los indicadores más comunes de la empresa, para apoyarse en la toma de decisiones.

Para el Area de Sistemas:

Facilidad de uso: La facilidad de uso de esta herramienta le permitirá tener un DW en minutos.

Acceso a cualquier base de datos: el soporte para acceder múltiples bases de datos amplía las áreas de aplicación de esta herramienta, para generar resúmenes operativos de su información.

Mayor rendimiento: el rendimiento de los servidores de su empresa se verá altamente beneficiado, debido a que el tiempo de consulta se verá disminuído.

Además *DATA WAREHOUSE* permite



:

Incluso durante la tramitación de un incidente comercial el administrativo puede sacar siempre informaciones de clientes, artículos y condiciones.

Permite información en línea (online) para la tramitación del negocio diario sin papeles.

La información está disponible rápidamente y pulsando una simple tecla.

Acceso directo, fácil y económico a todos los datos de la empresa.

Poseer una base de datos consistente en toda la empresa.

Visualizar en varias formas los resultados de sus consultas.

Redactar informes con la ayuda de módulos de textos e informes prefabricados.

Navegar rápida e intuitivamente por los datos de la empresa.

Identificar y analizar las tendencias en el tiempo más breve.

Capacidad para utilizar una variedad de herramientas para procesar y resumir los datos, así finalmente reagrupar toda la información.

Capacidad para explorar e identificar relaciones de datos a través de múltiples bases de datos.

Los datos históricos aumentan su valor a lo largo del tiempo.

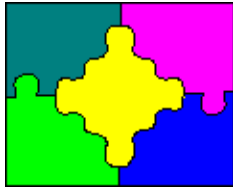
Crecimiento escalable.

Los logros son evidentes: mayor productividad por decisiones correctas en un tiempo más corto.

III. COMPONENTES Y ESTRUCTURA BASICA DEL DATA WAREHOUSE

Arquitectura DATA WAREHOUSE:

3.1. Componentes y estructura



Antes de describir la arquitectura DW vamos a señalar la siguiente consideración ya generalizada, presente en la literatura: el término DW se utiliza indistintamente para hablar de la arquitectura en sí como también para uno de los componentes que la conforman, específicamente el que tiene relación con el almacenamiento físico de datos.

La estructura básica de la arquitectura DW incluye:

- *Datos Operacionales*: un origen de datos para el componente de almacenamiento físico DW.
- *Extracción de Datos*: selección sistemática de datos operacionales usados para poblar el componente de almacenamiento físico DW.
- *Transformación de Datos*: procesos para sumar y realizar otros cambios en los datos operacionales para reunir los objetivos de orientación a temas e integración principalmente.
- *Carga de Datos*: inserción sistemática de datos en el componente de almacenamiento físico DW.
- *Herramienta de Acceso al componente de almacenamiento físico DW*: herramienta que proveen acceso a datos.

Fig.: ESTRUCTURA BASICA DEL DATA WAREHOUSE

IV. FASES DE IMPLEMENTACION DE UN DATA WAREHOUSE:

Tal y como aparecía en un artículo en ComputerWorld. un DATA WAREHOUSE no se puede comprar, se tiene que construir Como hemos mencionado con anterioridad, la construcción e implantación de un DW es un proceso evolutivo.

El proceso debe apoyarse en una metodología específica para este tipo de procesos. Si bien es importante la elección de la mejor de las metodologías, también es importante el realizar un control para asegurar el seguimiento de las mismas.

Las fases que se establezcan en el alcance de este proyecto son fundamentales. El seguir los pasos de la metodología y comenzar el DW por un área específica de la empresa, permitirá obtener resultados tangibles en un corto espacio de tiempo.

A continuación una metodología interactiva distribuída en cinco fases:

- *Definición de los objetivos*: los objetivos que se establezcan para el desarrollo de DW, juegan un papel preponderante. Una vez establecidos los objetivos, todo proyecto debe desarrollarse de forma clara y directa. El esclarecimiento de los objetivos son los pilares básicos para el desarrollo de todo proyecto.
- *Definición de los requerimientos de información*: tal como sucede en todo proyecto , sobre todo si involucra técnicas novedosas como son las relativas al DW, se deben analizar todas las necesidades y hacer comprender las ventajas que este sistema puede reportar. Es en este punto donde se debe detallar los pasos a seguir en proyectos de este tipo, donde el usuario juega un

papel preponderante.

- **Diseño y Modelización:** los requerimientos de información identificados durante la anterior fase proporcionarán las bases para realizar el diseño y la modelización del DW. En esta fase se identificarán las fuentes de los datos (sistema operacional, fuentes externas.) y las transformaciones necesarias para, a partir de dichas fuentes, obtener el modelo lógico de datos del DW. Este modelo estará formado por entidades y relaciones que permitirán resolver necesidades de negocio de la organización.

El modelo lógico se traducirá posteriormente en el modelo físico de datos que se almacenarán en el DW y que definirá la arquitectura de almacenamiento del DW adaptándose al tipo de explotación que se realice del mismo.

- **Implementación:** la implantación de un DW lleva implícito los siguientes pasos:
 - Extracción de los datos del sistema operacional y transformación de los mismos.
 - Carga de los datos validados en el DW. Esta carga deberá ser planificada con una periodicidad que se adaptará a las necesidades de refresco detectadas durante la fase de diseño del nuevo sistema.
 - Explotación del DW mediante diversas técnicas dependiendo del tipo de aplicación que se de a los datos:

*On-Line analytical processing (OLAP)

*Decision Support Systems ó Información de Gestión

*Visualización de la información

*Data Mining o minería de datos

La información necesaria para mantener el control sobre los datos se almacena en los metadatos técnicos (cuando describen las características físicas de los datos) y de negocios (cuando describen como usan esos datos). Dichos metadatos deberán ser accesibles por los usuarios finales que permitirán en todo momento tanto al usuario, como al administrador que deberá además tener la facultad de modificarlos según varíen las necesidades de información.

Con la finalización de esta fase se obtendrá un DW disponible para su uso por parte de los usuarios finales y el departamento de informática.

- **Revisión:** la construcción del DW no finaliza con la implantación del mismo, si no que es una tarea iterativa en la que se trata de incrementar su alcance aprendiendo de las experiencias anteriores.

Después de implantarse, debería realizarse una revisión del DW planteando preguntas que permitan, después de los seis o nueve meses posteriores a su puesta en marcha, definir cuáles serán los aspectos a mejorar o potenciar en función de la utilización que se haga del nuevo sistema.

- **Diseño de la estructura de cursos de información:** con la información obtenida de reuniones con los distintos usuarios se diseñaran una serie de cursos a medida, que tendrán como objetivo proporcionar la formación estadística necesaria para el mejor aprovechamiento de la funcionalidad incluida en la aplicación. Se realizarán prácticas sobre el desarrollo realizado, las cuales permitirán fijar conceptos adquiridos y servirán como formación a los usuarios.

4.1 . ESTRATEGIAS DE IMPLEMENTACION

En esta guía resaltaremos algunas consideraciones que recomendamos deben seguirse a la hora de querer implementar un proyecto de este tipo:

- La base de datos de riesgos debe estar separada de la base de datos operacionales con el propósito de no interferir en la actividad del día a día, disponiendo de esta manera de la información necesaria y un entorno orientado hacia la consulta y el análisis de los datos.
- Concepción del sistema como un conjunto de herramientas de análisis; esto es, porque las actividades de Análisis de Riesgo no se pueden automatizar, porque ellas requieren del análisis y decisión del usuario.
- Diseñar el sistema no orientado a procesos porque se debe disponer de un conjunto abierto de herramientas que se utilizan con propósitos determinados no relacionadas con las necesidades operativas del día a día.
- Abordar el sistema con un enfoque de desarrollo gradual, que debe comenzar con un esqueleto básico de funcionalidad y datos que produzcan resultados a corto plazo y permita aprender en la práctica y luego ir configurando progresivamente nuevas funcionalidades conforme la experiencia lo vaya requiriendo.

Una vez en marcha estas estrategias de implementación el esquema de Gestión y Explotación de nuestro *DATA WAREHOUSE* queda como se muestra en la siguiente figura que nos muestra un sistema DW ya implementado y listo para explotarlo.

V. DATA MINING Y DATA MART

1. Definición de *DATA MART*:

Pero, ¿ qué ocurre cuando DW crece y se hace más complejos ?

El rendimiento de las consultas vuelve a sufrir y el modelo centralizado deja de ser el más eficiente. En estos casos, la solución propuesta es crear *almacenes de datos especializados por áreas como Ventas o Compras*, que reciben los datos desde un almacén centralizado (DW) y que pueden residir en diferentes máquinas... estos almacenes se conocen como *DATA MARTS (DM)*.

Un *DATA MART* es una aplicación de un *DATA WAREHOUSE* construida rápidamente para soportar una línea de negocio simple. Los *DATA MARTS*, tienen las mismas características de integración, no volatilidad y orientación temática que el DW. Representan una estrategia de *divide y vencerás* para ámbitos muy genéricos de un *DATA WAREHOUSE*.

Esta estrategia es particularmente apropiada cuando el DW central crece muy rápido y los distintos departamentos requieren sólo una pequeña porción de los datos contenidos en él. La creación de los *DATA MART* requiere de algo más que una simple réplica de datos: se necesitarán tanto la segmentación como algunos métodos adicionales de consolidación.

La aproximación a una arquitectura descentralizada de *DATA MART*, podría ser originada de una situación como la descrita a continuación:

Un ejemplo de cuando aplicar *DATA MART*, se da en áreas específicas como el departamento de marketing o finanzas; el departamento de marketing, emprende el primer proyecto de *DATA WAREHOUSE*, como una solución departamental, creando el primer *DATA MART* de la empresa.

Visto el éxito del proyecto, otros departamentos como el de Riesgos, o el Financiero se lanzan a crear sus *DATA MARTS*. Marketing, comienza a usar otros datos que también los usan los *DATA MARTS* de Riesgos y Financieros, y estos hacen lo propio.

Esto parece ser una decisión normal, puesto que las necesidades de información de todos los *DATA MART* crecen conforme el tiempo avanza. Cuando esta situación evoluciona, el esquema general de integración entre *DATA MARTS* pasa ser, la del gráfico anterior.

En esta situación, es fácil observar cómo este esquema de integración de información de los *DATA MARTS*, pasa a convertirse en un rompecabezas en el que la gestión para obtener la información se ha complicado. No obstante, lo que ha fallado no es la integración de *DATA MARTS*, sino su forma de integración.

En efecto, un enfoque más adecuado sería la coordinación de la gestión de información de todos los *DATA MART* en un *DATA WAREHOUSE* centralizado, como lo muestra la figura anterior.

En esta situación los *DATA MARTS* obtendrían la información necesaria, ya previamente cargada y depurada en el DW corporativo, simplificando el crecimiento de una base de conocimientos a nivel de toda la empresa.

Esta simplificación proveería la descentralización de las labores de gestión de los *DATA MARTS* y el DW corporativo, generando economías de escala en la gestión de los *DATA MARTS* implicados.

Toda empresa puede ser vista en base al proceso productivo que la sustenta. El resultado de los costos y beneficios de este proceso productivo forman una cadena de valor, donde cada eslabón (proceso de negocios) adiciona valor a la empresa. De esta forma es claro que las empresas deben buscar optimizar cada uno de sus eslabones sin perder de vista la cadena total.

Dado que un *DATA MART* soporta menos usuarios que un *DATA WAREHOUSE* se puede optimizar para recuperar más rápidamente los datos que necesitan los usuarios. La arquitectura de un *DATA MART* es aconsejable porque:

- Menores cantidades de datos implican que se procesan antes, tanto las cargas de datos como las consultas.
- Las peticiones pueden acotarse al área o red que sirve esos datos, sin afectar al resto de los usuarios.
- La aplicación cliente, que pide la consulta es independiente del servidor que la procesa y del servidor de bases de datos que almacenan la información.

2. Definición de *DATA MINING*:

El análisis y procesamiento de datos en la bodega se puede apoyar y complementar con varias técnicas de análisis, entre las que se destaca la minería de datos (*data mining*), técnica de inteligencia artificial que permite buscar condiciones anormales o fuera de rango, ya sea de manera programada o automática, en los datos contenidos en la bodega, facilitando así el análisis que de otra manera sería difícil obtener con otros métodos. Un ejemplo claro de esto se puede aplicar en un banco y el sistema de tarjetas de crédito, fácilmente se puede programar un minero para que automáticamente avise y busque dentro de la cartera de clientes todos los tarjeta habientes que hayan vencido; y proyecte al siguiente mes el crecimiento de la cartera vencida y avisara en caso de sobrepasar el monto de cartera vencida definida como crítica.

Las herramientas de *DATA MINING* recorren los datos detallados de transacciones para desenterrar

patrones. Por lo general, los resultados generan extensos reportes o se les analiza con herramientas de visualización de datos descubiertos.

El procesamiento informático es excelente y rentable para el despliegue masivo de consultas, análisis y reportes de datos de dos o tres dimensiones. Las herramientas de procesamiento analítico permiten diversas visualizaciones con los datos, como ventas por marca, tienda, temporada y períodos de tiempo, las cuales se pueden definir, consultar y analizar. Las herramientas del *DATA MINING* son esenciales para comprender el comportamiento de los clientes.

Los usuarios clave en perspectiva del Data Mining son los analistas empresariales y los profesionales en tecnología de la información que auxilian a los usuarios empresariales. Quienes obtienen beneficios de los resultados del data Mining son los gerentes empresariales y los ejecutivos que desean entender los factores de éxito del negocio con base en datos completos del cliente y, utilizar luego, este conocimiento para afinar las estrategias de producción, precios y comercialización, mejorar el nivel de éxito y las estrategias, e impulsar el balance.

VI . OTROS SISTEMAS:

1. Sistema OLTP

Históricamente, los datos de la corporación suelen residir en *bases de datos* (BD) que se diseñaron principalmente para introducir y almacenar datos, mediante el llamado *Proceso de Transacciones On Line* (*OLTP*). Este método es idóneo para insertar, modificar o borrar registros, pero no lo es tanto para responder a complejas consultas. La relación entre los datos responde a las técnicas llamadas de *Entidad-Relación*.

Un sistema operacional o de procesamiento en línea es un sistema tal como el de administración de recursos humanos, de asignación de bancarios de recuperación y control de cartera o de control de seguros, y su función principal es dar el soporte a las necesidades del día; son sistemas normalmente optimizados para el manejo de un conjunto predefinido de transacciones.

Los sistemas operacionales de los cuales se transferirá la información seleccionada, pueden haber sido contruidos utilizando manejadores de datos relacionados, manejadores de archivos jerárquicos, de archivos planos u otro tipo de manejadores. Por ello, es necesario analizar y definir cuidadosamente de los sistemas operacionales aquellos datos que representen la esencia o filosofía del negocio que se pretenda manejar, para que al transferir los datos a la bodega, ese conocimiento primordial se capture en lo que se conoce como metadatos, que son precisamente, los que describen a los datos provenientes de los sistemas operacionales.

2. Diferencias: OLTP vs. DATA WAREHOUSE:

Fig. ESQUEMA DE RELACION ENTRE OLTP Y DW

Es de suma importancia comprender perfectamente estas diferencias para evitar caer en el diseño de un *DW* como si fuera una aplicación de transacción en línea (*OLTP*).

Las aplicaciones de *OLTP* están organizadas para ejecutar las transacciones para los cuales fueron hechas, como por ejemplo: mover dinero entre cuentas, un cargo o abono, una devolución de inventario, etc. Por otro lado, un *DW* 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 *DW* es

menor al de un *OLTP*. Es común encontrar que los sistemas transaccionales son accesados por cientos de usuarios simultáneamente, mientras que los *DW* sólo por decenas. Los sistemas *OLTP* realizan cientos de transacciones por segundo mientras que una sola consulta de un *DW* puede tomar minutos. Otro factor es que frecuentemente los sistemas transaccionales son menores en tamaño a lo *DW*, esto es debido a que un *DW* puede estar formado por información de varios *OLTP*.

Existe también diferencia en el diseño, mientras que un *OLTP* es extremadamente normalizado, un *DW* tiende a ser desnormalizado. El *OLTP* normalmente está formado por un número mayor de tablas, cada una con pocas columnas, mientras que un *DW* el número de tablas es menor, pero cada una de estas tiende a ser mayor en número de columnas.

Los *OLTP* son continuamente actualizados por los sistemas operacionales del día con día, mientras que los *DW* son actualizados en batch (back up) de manera periódica.

Las estructuras de los *OLTP* son muy estables, rara vez cambian, mientras que los *DW* sufren cambios constantes derivados de su evolución. Esto se debe a que los tipos de consultas a los cuales están sujetos, son muy variados y es imposible preverlos todos de antemano.

<u><i>OLTP</i></u>	<u><i>DATA WAREHOUSE</i></u>
Orientada a transacciones	Orientada a conceptos
Detallada	Sumarizada
Actualizada en línea	Representa valores a un tiempo
Usuarios a nivel operativo	Usuarios a nivel gerencial
Corre en base a repeticiones	Corre heurísticamente
Muy sensitivo al desempeño	Poco sensitivo al desempeño
Accesa unidades a la vez	Accesa conjuntos de unidades a la vez
Orientado a una operación	Orientado a análisis
Estructura estática	Estructura flexible
Sin redundancia	Con mucha redundancia
Alta probabilidad de acceso	Modesta probabilidad de acceso
Administrada como un todo	Administrada por partes
Información bruta (datos)	Información procesada (información)
Actualizada en línea	Actualizada en batch
Muchas tablas con pocas columnas	Pocas tablas con muchas columnas

Normalmente los sistemas operacionales de los cuales se extraen datos son muy diversos y cada uno de ellos aporta varios gigabytes, por lo que es normal que las bodegas de datos contengan del orden de 20, 50, 100, 200 o más gigabytes, aunque es posible, y en algunos casos puede ser deseable, construir un *DW* tan pequeño como en 200 o 500 megabytes.

Los *DW* surgen precisamente en respuesta a los problemas asociados a realizar análisis de datos sobre *Bases de Datos del tipo OLTP*. La solución propuesta por el *DW* es extraer los datos de una (o más) bases operacionales y moverlos a una Base de Datos independiente y orientada a consultas.

Los objetivos de los sistemas *OLTP* y *DATA WAREHOUSE* para toma de decisiones son muy diferentes: tratar de diseñar un *DATA WAREHOUSE* pensando en un sistema operacional es un fracaso. El éxito en la implementación de una arquitectura de *DW* en las empresas radica en parte en el éxito del diseño de los sistemas operacionales, ya que estos son los proveedores de los datos y los que se deben de adaptar rápida y flexiblemente a los cambios del negocio. Se deben tener ideas muy claras de lo

siguiente: Qué datos utilizar?, ¿cómo se deben transformar?, ¿cómo se deben transferir?, almacenar y organizar; y finalmente, cómo se deben acceder y analizar?.

Actualmente es innegable que los sistemas de información *OLTP*, contruidos utilizando manejadores de bases de datos relacionados son la norma. Es una tecnología madura que provee las facilidades necesarias. Cuando los usuarios de negocios empujados por las necesidades del mercado iniciaron con sus demandas de información actualizada, de proyecciones en el tiempo, de análisis comparativos entre regiones en diferentes períodos de tiempo, los desarrolladores utilizaban herramientas y manejadores de bases de datos que tenían el alcance.

3. Una máquina o dos?

Es mucho más habitual encontrar separado el *DW del OLTP*, debido a factores bien específicos y de considerable relevancia para su desempeño. Los argumentos que favorecen el tener máquinas separadas el *OLTP y el DW* son:

- El *DW* tiene una significativa y altamente variable demanda de recursos, por lo tanto puede entorpecer considerablemente el desempeño de *OLTP*.
- Los sistemas en cuestión son configurados de manera muy diferente
- A veces, los datos del *DW* son integrados de múltiples sistemas *OLTP* remotos, y por lo tanto el *DW* puede verse como un conjunto de recursos centralizados. Es obvio entonces que estén físicamente separados ambos sistemas.

La razón para tenerlos en la misma máquina está en el hecho de que al ser la estructura básica del *DW* distinta a la del *OLTP*, el dato tiene que ser copiado y reestructurado por el *DW*. Para ahorrar envíos de datos entre máquinas, es mejor realizar este proceso dentro de una sola.

• CUANDO IMPLEMENTAR DATA WAREHOUSE

7.1. ¿ Quién necesita un Data Warehouse y porqué ?

El *DW* implica el concepto de valor de la información. Datos sobre el negocio son útiles cuando el tiempo es un factor en contra. Las personas que toman las decisiones claves necesitan un acceso rápido a los datos actuales aportados por diferentes sistemas de producción y bases de datos que operan tanto de forma interna como externa para tomar decisiones claves sobre la dirección y operaciones del negocio.

- *Los usuarios de la información corporativa:* Normalmente el usuario de la información provista por un *DW* es un trabajador especializado en una concreta línea el negocio de un departamento. Este departamento normalmente trata con una gran cantidad de transacciones y clientes. Estas actividades se reflejan en la infraestructura como bases de datos que llegan a cientos de gigabytes y continúa creciendo.

La mayoría de las aplicaciones de *DW* se utilizan para alimentar proyectos de marketing que se dirigen a los clientes potenciales más adecuados con el menor costo de venta. Estos proyectos de crecimiento o supervivencia se llevan a cabo normalmente cuando:

- Los mercados son muy dinámicos y requieren una rápida respuesta ante el cambio en la demanda, regulaciones y costos.
- Hablamos de clientes que demandan más productos individualizados o servicios, que requieren gradualmente más conocimiento de:

¿ Quiénes son?

¿ Cómo compran ?

¿ Qué piensan de mi compañía ?

- El liderazgo competitivo requiere mejoras aceleradas en desarrollo de productos, claro entendimiento de las necesidades de los clientes para un ajuste del producto óptimo y un aumento de la habilidad para articular diferencias competitivas en las necesidades de los clientes.

Pero... ¿ cuáles son las necesidades más comunes de los clientes?

- Aumentar el acceso de datos
 - Aumentar la productividad del usuario
 - Soportar decisiones complejas
 - Mantener los datos como un recurso del negocio
 - Integrar información en la empresa
 - Mejorar los servicios al consumidor
 - Reducir tiempo de venta
 - Reducir los costes operativos
 - Reducir los costes de las ventas
- *Tipos de aplicaciones en las que utilizar las técnicas disponibles sobre DATA WAREHOUSE*
 - *DW y Sistemas de Marketing:* La aplicación de tecnologías de DW supone un nuevo enfoque en marketing, haciendo uso del marketing de base de datos. En efecto, un sistema de marketing Warehouse implica un marketing científico, analítico y experto, basado en el conocimiento exhaustivo de clientes, productos, canales y mercado.

Este conocimiento se deriva de la disposición de toda la información necesaria, tanto interna como externa, en un entorno DW, persiguiendo con toda esta información, la optimización de las variables controladas del Marketing Mix y el soporte a la predicción de las variables no controlables (mediante técnicas de Data Mining) Basándose en el conocimiento exhaustivo de los clientes se consigue un tratamiento personalizado de los mismos tanto en el día a día (atención comercial) como en acciones de promoción específica.

Las áreas en las que se puede aplicar las tecnologías de DW a marketing son, entre otras:

- Investigación comercial
 - Segmentación de mercados
 - Identificación de necesidades no cubiertas y generación de nuevos productos, o modificación de productos existentes
 - Fijación de precios y descuentos
 - Definición de la estrategia de canales de comercialización y distribución
 - Definición de la estrategia de promoción y atención al cliente
 - Relación con el cliente
 - Lanzamiento de nuevos productos
 - Campañas de ventas cruzadas, vinculación, fidelización, etc.
 - Apoyo al canal de venta con información cualificada.
- *DW y Análisis de Riesgo Financiero:* El DW aplicado al análisis de riesgos financieros ofrece capacidades avanzadas de desarrollo de aplicaciones para dar soporte a las diversas actividades de

gestión de riesgos. Es posible desarrollar cualquier herramienta utilizando las funciones que incorpora la plataforma, gracias a la potencialidad estadística aplicada al riesgo de crédito.

Así se puede usar para llevar a cabo las siguientes funcionalidades:

- Para la gestión de la posición: determinación de la posición, cálculo de sensibilidades, simulaciones, monitorización riesgos contra límites, etc.
- Para la medición de riesgos: simulación de escenarios históricos, simulación de Montecarlo, modelos de valoración, análisis de rentabilidad, establecimiento y seguimiento de límites, etc.

El uso del DW ofrece una gran flexibilidad para creación y modificación de modelos propios de valoración y medición de riesgos, tanto motivados por cambios en la regulación como en avances en la modelización de estos instrumentos financieros.

Ello por cuanto se puede almacenar y poner a disposición información histórica de mercado y el uso de técnicas de Data Mining nos simplifica la implantación de cualquier método estadístico.

7.2.3. DW y Análisis de Riesgo de Crédito : La información relativa a clientes y su entorno se ha convertido en fuente de prevención de Riesgos de Créditos. En efecto, existe una tendencia general en todos los sectores a recoger, almacenar y analizar información crediticia como soporte a la toma de decisiones de análisis de riesgo de créditos.

- **DW: Otras áreas de aplicación :** Otras áreas de la empresa han aplicado las soluciones que proporciona la tecnología *Data Warehouse* para mejorar gran parte de sus procesos actuales, entre ellos:
 - Control de gestión: sistema de presupuestación, análisis de desviaciones, etc.
 - Logística: mejora de la relación con los proveedores, racionalización de los procesos de control de inventarios, optimización de los niveles de producción, previsión de la demanda en infraestructura.
 - Recursos humanos: planificación de incorporaciones, gestión de carreras profesionales, asignación de recursos a proyectos alternativos, etc.

7.2.5. Quienes han aplicado Data Warehouse:

España es uno de los países que ha implementado *DATA WAREHOUSE*, dando así como resultados casos de éxito: como la empresa *BANKINTER* quienes aplicaron DW en el mercado alternativo financiero. (*VER ANEXO*)

Otro caso de éxito ha sido el implementado por *UNION FENOSA*, empresa dedicada al negocio de sistemas eléctricos y no eléctricos, cuyo negocio enfrenta un mercado muy competitivo.

(*VER ANEXO*).

También el *METRO DE MADRID* ha puesto en práctica el sistema DW obteniendo óptimos resultados en el transporte de viajeros. DW fue implementado con planes de expansión para ofrecer a sus usuarios un mayor servicio. (*VER ANEXO*).

VIII. COSTOS DE UN DATA WAREHOUSE:

- Costos de Construcción
- Costos de Operación

- **Costos de construcción:** Los costos de construir un DW son similares para cualquier proyecto de tecnología de información. Estos pueden ser clasificados en tres categorías:
- **RRHH:** la gente necesita contar con un enfoque fuerte sobre el conocimiento del área de la empresa y de los procesos empresariales. Además es muy importante considerar las cualidades de la gente, ya que el desarrollo del DW requiere participación de la gente de negocios como de los especialistas tecnológicos ; estos dos grupos de gente deben trabajar juntos, compartiendo su conocimientos y destrezas en un espíritu de equipo de trabajo, para enfrentar los desafíos de desarrollo del DW.
- **TIEMPO:** se debe establecer el tiempo no tan solo para la construcción y entrega de resultados de DW, si no también para la planeación del proyecto y la definición de la arquitectura. La planeación y la arquitectura establecen un marco de referencia y un conjunto de estándares que son básicos para la eficacia de DW.
- **TECNOLOGIA:** muchas tecnologías nuevas son introducidas por el DW. El costo de esta nueva tecnología puede ser tan sólo la inversión inicial del proyecto.
- **Costos de Operación:** Una vez que está construido y entregado un DW debe ser soportado para que tenga un valor empresarial, son estas actividades de soporte, las fuentes de continuos costos operacionales para un DW. Se pueden distinguir tres costos de operación:
- **EVOLUTIVOS:** Ajustes continuos del DW a través del tiempo, como cambios de expectativas y, productos del aprendizaje del RRHH del proyecto mediante su experiencia usando el DW.
- **CRECIMIENTO:** incrementos en el tiempo en volúmenes de datos, del número de usuario del DW, lo cual con llevará a un incremento de los recursos necesarias como a la demanda de monitoreo, administración y sintonización del DW(evitando así, un incremento en los tiempos de respuesta y de recuperación de datos, principalmente).
- **CAMBIOS:** El DW requiere soportar cambios que ocurren tanto en el origen de los datos que éste usa, como en las necesidades de la información que éste soporta.

Los dos primeros tipos de costo de operación, son básicos en la mantención de cualquier sistema de información; sin embargo, se debe tener esencial cuidado en los costos ,de operación por cambio, ya que ellos consideran el impacto producto de la relación del OLTP y del Ambiente Empresarial, con el DW.

CONCLUSIONES

- Día a día aparecen en el mercado nuevos productos de software de ayuda a la implementación de cada una de las etapas que conforman la arquitectura DATA WAREHOUSE, la tecnología de apoyo, adicional a que aún no es madura, involucra tecnologías y conceptos nuevos y pretender ser experto de la noche a la mañana, es difícil, por lo que tratar de implementar DW pensando que ha de resolver los problemas de información podría llegar a ser contraproducente.
- Una recomendación para quienes decidan iniciar un modelo DW en sus empresas: deben apoyarse en consultores conocedores, serios y responsables que ayuden con bases sólidas a construir y mantener cada una de las etapas del proyecto.
- DATA WAREHOUSE es una colección de datos relacionados al oficio o a la actividad profesional, integrada y relacionada a la fecha en que se produjeron para poder analizarlos periódicamente. El enfoque está en la disposición eficiente de estos datos así como la ejecución de evaluaciones y análisis.
- DATA WAREHOUSE apoya a todos los ejecutivos que tienen que tomar decisiones con informaciones rápidas.
- Los logros son evidentes: MAYOR PRODUCTIVIDAD POR DECISIONES CORRECTAS EN UN TIEMPO MAS CORTO.
- DATA WAREHOUSE te permite desarrollar aplicaciones creando un resumen de tu base de

datos operativa con sólo la información realmente importante, logrando obtener consultas más eficientes y productivas.

BIBLIOGRAFIA

[www.intellicomp.cl /datawh.htm](http://www.intellicomp.cl/datawh.htm)

www.gcc.com.mx/soluciones/dwhouse.htm

www.bftsystems.com/sol_dataware.htm

www.wnet.es/Productos/infosagent1.htm

www.compag.es/soluciones/dwareymart/Secciones/dwsodwxx.htm

www.mtginc.com/spanish/whatwedo/technical/dws.html1

www.sun.es/success/warehouse/bankinter

www.194.174.14.5/Espanol/es_data_warehouse.htm

www.fcencias.unam.mx/revista/soluciones/30s/No34/dataware.htm1

www.business.carleton.ca/~aramirez/Espol/Modulo1/index.html1

www.datawarehouse.com

www.digital.com/alphaserver/solutions/dataware/dataware.htm1

www.intellicomp.cl/datawh.htm

www.pwp.starnetinc.com/larryg/defined.html

www.consisa.com/info-corp/preentaciones/nuevos-paradigmas/tsld012.tm

Entrevista realizada a Juan Palacios – Gerente de Capacitación de New Horizons



