

UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS
SECCION DE POST GRADO



**EL DATAMART EN UNA EMPRESA
INDUSTRIAL DEL SECTOR PLÁSTICO**

TESIS

Para optar el Grado Académico de:

MAESTRO EN CIENCIAS
MENCIÓN EN INGENIERIA DE SISTEMAS

RUBEN ROBERTO MELLADO FLORES

Lima - Perú
2002.

AGRADECIMIENTOS

A mis padres que en paz descansen, a mi esposa y mis hijas por entender y comprender el hecho de dedicarle un tiempo importante a éste trabajo, a mi asesora por su persistencia y motivación para iniciar y concluir el presente trabajo, a todos los profesores de postgrado de la UNI, por sus enseñanzas que contribuyeron a preparar el presente trabajo y a mis ex compañeros de aula por los momentos que compartimos nuestras experiencias profesionales, a todos ellos les dedico el esfuerzo realizado en la presente.

INDICE

| | Página |
|--------------------------|--------|
| Agradecimientos | II |
| Resumen ejecutivo | III-IV |
| Introducción | 1-5 |

CAPITULO I: Marco Teórico Referencial

| | |
|---|-------|
| 1. Administración de las Relaciones con los Clientes ARC ó Customer Relationship Management | 6-14 |
| 2. ¿Que es un Datawarehousing? | 15-31 |
| 3. El Proceso de maduració de un Datawarehousing | 32-34 |
| 4. Teoria de las Etapas | 34-38 |
| 5. Sistema de Soporte a las Decisiones | 38-50 |
| 6. DataWarehousing vs DataMart | 51-56 |
| 7. Metodologías para Implementar Datawarehousig | 56-86 |

CAPITULO II: Propuesta de Metodología de Desarrollo de Prototipo

| | |
|---|-------|
| 2.1 Diagnostico de la Situación Actual de la Organización | 87-88 |
| 2.2 Diagnostico de la infraestructura de la Tecnología de la Información de la Organización | 88-89 |
| 2.3 Evaluación de Metodologías | 90-93 |

CAPITULO III: Desarrollo de Prototipo

| | |
|---|---------|
| 3.1 Diagnóstico de la situación actual de la organización | 94-105 |
| 3.2 Diagnóstico de la infraestructura en TI | |
| 3.2.1 Infraestructura de Hardware y Redes | 106 |
| 3.2.2 Infraestructura de Software | 106 |
| 3.2.3 Infraestructura de Personal de TI | 107-108 |
| 3.3 Desarrollo de DataMart | |

DESCRIPTORES TEMATICOS

| | |
|---|---------|
| 3.3.1 Acuerdos del Proyecto | 109-119 |
| 3.3.2. Proceso Back End | 120-141 |
| 3.3.3 Proceso Front End | 142-147 |
| 3.3.4. Proceso de Establecimiento de Plataforma Tecnológica | 148 |

Conclusiones y recomendaciones

| | |
|--------------------|---------|
| 1. Conclusiones | 150-151 |
| 2. Recomendaciones | 152 |

Bibliografía y anexos

| | |
|-----------------|---------|
| 1. Bibliografía | 153-155 |
| 2. Anexos | 156-194 |

DESCRIPTORES TEMATICOS

Datamart

Metodologías de implementación de Datamart

Sistemas de apoyo a las decisiones

Datawarehousing

Administración de las Relaciones con los Clientes ó Customer Relationship Management CRM

RESUMEN EJECUTIVO

Pensamos que una empresa que quiera ser competitiva en los actuales momentos debe adoptar los conceptos de la administración digital y en lo relacionado al Front-office establecer sistemas del tipo CRM, Administración de las Relaciones con los Clientes.

Parte central de éstos sistemas son los Datamart.

En el presente trabajo queremos demostrar que con el uso de la Tecnología de la Información podemos obtener ventaja competitiva, para lo cuál planteamos el uso del Datamart.

Primero planteamos una metodología de implementación, para luego desarrollar un prototipo de Datamart en el área comercial de CIDELSA, empresa del sector industrial plástico.

Para la parte de metodología de implementación planteamos primero el Diagnóstico de la Organización, para lo cuál propusimos:

Análisis estratégico, apoyados en la teoría desarrollada por Michael Porter y las matrices metodológicas.

Análisis de la Tecnología de Información de la organización, apoyados con la Curva de Nolan y los Tipos de Sistemas de Soporte a las Decisiones

La conclusión de éste diagnóstico fue que CIDELSA como una opción estratégica para obtener ventaja competitiva, debe implementar un sistema de información de ayuda a las decisiones si es que quiere conservar su mercado actual y mejorar su participación en el mismo. Así mismo de acuerdo a éste diagnóstico la empresa CIDELSA debe comenzar su sistema de información con el tipo I, es decir registrando la información para luego dar respuesta a la pregunta estratégica: ¿Qué sucedió?.

Para la parte del desarrollo en sí del Datamart y luego de revisar diversas metodologías decidimos optar por la planteada por Larisa Mos y Sid Adelman que plantea lo siguiente:

- Se dividió el desarrollo en tres grandes etapas: Front End, Back End e Infraestructura tecnológica.
- Estas tres grandes etapas se desarrollaron en forma simultánea en loops, finalizado cada loop había resultados controlables.
- Antes de iniciar los loops se preparó un documento denominado Acuerdos del Proyecto, donde se limitó los alcances del Datamart. Participaron en preparar dicho documento: los sponsors, personal de tecnología de información y los usuarios.

INTRODUCCION

A través de los años hemos observado como las organizaciones han usado los Sistemas de Información. Desde la mecanización de las transacciones operativas hasta el apoyo a la toma de decisiones a nivel gerencial. En ésta última década el uso de los Sistemas de información tiene dos características adicionales: dichos sistemas se están utilizando para dar ventaja competitiva a las organizaciones y el desarrollo tecnológico de la Tecnología de Información ha sido enorme.

La Administración Digital de nuestro tiempo consiste en adoptar sus dos columnas: para el Back-office, sistemas de optimización de costos, teniendo un enfoque hacia dentro de la empresa, enfocado al manejo eficiente de los recursos: Sistemas ERP, es decir Sistemas de Planeamiento de Recursos de la Empresa, ejemplo Sistemas SAP. Por otro lado en el Front-office, utilizando el enfoque de optimizar las interacciones con los clientes, con un manejo adecuado de las relaciones con los mismos, ejemplo Sistemas CRM, es decir Sistemas de Administración de las Relaciones con nuestros Clientes, cuya parte neurálgica es el Datawarehouse. Las empresas para mantenerse competitivas tendrán que adoptar éstos conceptos de la Administración Digital.

Planteamiento del problema

Existen diversidad de datos que se encuentran dispersos en las organizaciones y que se utilizan principalmente para el apoyo de las transacciones operativas de las empresas, es decir no se las utiliza, a pesar que existen, en el apoyo para la toma de decisiones a nivel gerencial, en forma sistemática, eficaz y oportunamente.

Por otro lado ante un entorno cada vez más competitivo y globalizado, las empresas tendrán que adoptar las herramientas que nos proporciona la administración digital, si no quieren quedarse con mercados marginales.

Objetivos :

Objetivo de la Investigación.-

Desarrollar un sistema de apoyo a las decisiones que permita incrementar la participación en el mercado de una empresa del sector industrial plástico, mediante el uso del Datamart.

Objetivos Específicos.-

Desarrollar una metodología que permita el uso del Datamart en empresas del sector.

Desarrollar un Prototipo:

Diseñar tablas de hecho y tablas de dimensión

Diseñar modelo starnet

Diseñar sistema de extracción, depuración y carga del Datamart

Diseñar interfases para usuarios

Justificación :

A través de los años hemos observado como las organizaciones han usado los Sistemas de Información. Desde la mecanización de las transacciones operativas hasta el apoyo de las decisiones a nivel gerencial. En ésta última década el uso de los Sistemas de Información tiene dos características adicionales: dichos sistemas se están utilizando para dar ventaja competitiva a las organizaciones y el desarrollo de la tecnología ha sido enorme.

El presente trabajo trata precisamente del desarrollo de un Sistema de Apoyo a las Decisiones, en una Empresa Industrial, teniendo como característica principal el uso de un Datamart.

Es posible que la razón más importante de construir un Datamart sea lo inadecuado de los sistemas actuales y la falta de información gerencial, incluso cuando la empresa está inundada de datos. Muchos sistemas de producción no satisfacen las necesidades del usuario empresaria. Por lo regular, los datos son inaccesibles e inconsistentes, tanto en forma como en significado. Por ejemplo, debido a la inconsistencia de datos, no coincide la información de ventas en diferentes reportes, la empresa carece de una imagen precisa de su ingreso. La falta de medidas comunes significa que los administradores no tienen una imagen clara del desempeño del negocio.

Por otro lado para implementar éste tipo de tecnologías de información, que demandan gran cantidad de dinero y tiempo se requiere de una metodología de desarrollo de lo contrario el proyecto está condenado al fracaso.

Alcances y Limitaciones:

El presente trabajo se desarrollará para una empresa industrial-comercial de mediana envergadura.

El prototipo se desarrollará en el área comercial de la empresa.

Para el modelado de datos utilizaremos el modelador ERWIN.

Para el diseño de interfases para los usuarios utilizaremos el Visual Basic.

El presente trabajo trata precisamente del desarrollo de un Sistema de Apoyo a las Decisiones, en una Empresa Industrial del sector plástico, teniendo como característica principal el uso de un Data Mart, cuyo prototipo es desarrollado en el área Comercial. Nuestra idea inicial fue el desarrollo completo de un Data Warehouse, pero las limitaciones de tiempo y costo fueron los motivos que nos llevaron a plantear el presente trabajo como sólo el desarrollo de un prototipo aplicado al área comercial.

Otro tema importante que se desarrolla en este trabajo es la metodología de implementación del Datamart. Muchos proyectos de este tipo han fracasado

porque no se han tomado el trabajo de seguir rigurosamente una metodología que analice y controle todos los parámetros que puedan afectar el éxito durante la implementación de éste tipo de tecnologías.

Luego de investigar y analizar el estado del arte referido a implementación de éste tipo de sistemas de apoyo a las decisiones, se ha desarrollado una metodología que aporta lo siguiente:

- Enlazar el desarrollo de un determinado tipo de sistema de apoyo a las decisiones con la estrategia de la empresa, para así lograr ventaja competitiva.

Luego de realizar el análisis estratégico de la empresa se podrá comprobar si el sistema de apoyo a las decisiones realmente dará ventaja competitiva a la empresa.

- Determinar el tipo de sistema de apoyo a las decisiones y los alcances teniendo en cuenta la madurez organizacional y tecnológica de la empresa.

Las empresas tienen diferente cultura organizacional e informática, por lo tanto los sistemas de apoyo a las decisiones que se quieran desarrollar deberán contemplar un análisis que diagnostique la situación de la empresa en éstos aspectos. Esta metodología propone tres tipos de sistema de apoyo a las decisiones cuya implementación dependerá del diagnóstico cultural y tecnológico de la empresa.

- Esta metodología recomienda que antes de implementar el primer Datamart en una organización se debe desarrollar un modelo de datos de la empresa para evitar futuras complicaciones. Es decir se deben detallar todas las entidades y sus posibles relaciones, aunque luego se tome una parte de éste análisis para implementar un Datamart. Esto ayudará a implementar en el futuro reportes y respuestas a consultas no desarrolladas inicialmente y a implementar otros Datamart sin tener que modificar ó corregir lo desarrollado en el modelo de datos inicial.

El presente trabajo tiene los siguientes capítulos:

Marco Teórico Referencial, donde se muestra el estado actual de éste tipo de tecnologías de información, metodologías de implementación, herramientas de desarrollo, etc. Diagnóstico de Situación Actual, donde en forma resumida se muestra la situación actual de la empresa Cidelsa, su problemática, sus oportunidades de negocio, sus amenazas, sus fortalezas y debilidades.

Propuesta de Metodología de Desarrollo, donde después de revisar una serie de metodologías aplicables a la implementación de este tipo de tecnologías, se plantea la que a nuestro entender se obtendrán resultados inmediatos y controlables.

Desarrollo de Prototipo, como explicáramos líneas arriba desarrollar todo un Data Warehousing era un objetivo inalcanzable por tiempo y dinero, por lo que nos planteamos el desarrollo de un prototipo de Datamart en el área comercial.

Finalmente encontramos el capítulo de Conclusiones y Recomendaciones, parte final de éste trabajo y donde se muestra los resultados encontrados en la aplicación del prototipo en el área comercial

CAPITULO I : MARCO TEORICO REFERENCIAL

1. Administración de las Relaciones con los Clientes ARC ó Customer Relationship Management CRM.

Durante el siglo veinte, muchas compañías perdieron el contacto con el cliente. En los días “pasados”, el carnicero, el bodeguero, ó en general el propietario de una tienda, no necesitaba las computadoras para conocer su mercado, como vender, como dar servicio a sus clientes, que productos debería tener en stock y en que cantidades. El comerciante obtenía dicho conocimiento interactuando diariamente con sus clientes. La mayoría de ellos se encontraba cerca del establecimiento comercial y acudía personalmente para la satisfacción de sus requerimientos.

Las economías de escala han transformado las pequeñas tiendas en grandes almacenes con miles de productos. La Web es la parte física de la entrada a la tienda, dejando solamente como evidencia de la interacción con los clientes, una serie de movimientos digitales.

Solamente las compañías modernas pueden capturar el conocimiento de los clientes utilizando la tecnología de la información. Son muchos los desafíos y las compañías se están moviendo a implementar nuevos sistemas y estrategias para la administración de las relaciones con los clientes (ARC).

1.1. Las necesidades urgentes de los negocios

Muchas compañías están descubriendo que un pequeño porcentaje de sus clientes representa la mayoría de sus ingresos. Parafraseando la Ley de Pareto, un pequeño porcentaje de los clientes representa el mayor porcentaje de ingresos por ventas.

En adición, notamos que la información de los clientes se ha convertido en una nueva herramienta competitiva, con ella las organizaciones deben mejorar su posición en el mercado.

En las pasadas dos décadas muchas compañías redujeron sus costos a través de la reingeniería, y crearon nuevas sinergias de los negocios a través de adquisiciones y fusiones. La competencia se ha intensificado y muchas empresas están observando nuevas fuentes de ingresos y beneficios. Las compañías reconocen que su base instalada de clientes es una mina de oro, que le puede dar la oportunidad de mejorar sus ingresos y beneficios.

Entonces no es en vano que ciertos consultores estén exhortando a sus clientes para que recompensen a sus clientes leales, practicando un marketing one-to-one, e implementando una vista para observar 360 grados las actividades de los clientes.

El objetivo es anticipar las necesidades de los clientes a través de un focalizado esfuerzo de marketing y ventas, asegurándonos los clientes y beneficios por haber proporcionado un excelente servicio a nuestros clientes.

1.2. ¿Qué es un ARC (CRM)?

Simplemente es una estrategia de negocios para incrementar los ingresos y los beneficios, focalizándose en los clientes. Esto es un cambio radical, ya que muchas compañías se focalizaban en sus productos, procesos, calidad ó finanzas como las claves para conducir las iniciativas corporativas, estructuras organizacionales y las medidas del performance. En esencia, ARC significa que las compañías utilizan al “cliente” como un lente para mirar a través de él y entender y definir sus mercados y ellos mismos.

En las estrategias del ARC convergen varias funciones de negocios y sistemas de información. Incluye marketing, ventas, y servicios al cliente. Estas

disciplinas están conducidas e integradas por una base de datos de los clientes y herramientas de análisis para entender y predecir el comportamiento de los clientes. (Ver Figura1).



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 1

Reconociendo la potencia de la visión del mercado centralizada al cliente, muchas compañías están ahora implementando sistemas ARC. De acuerdo a una reciente encuesta, más de la mitad de las respuestas (53%) habían desarrollado sistemas para optimizar las interacciones con sus clientes ó potenciales clientes ó estaban en el proceso de implementarlo. Otro 20% tenía presupuestado construir un sistema de administración centralizada en el cliente.

Dichas entrevistas fueron conducidas por el Instituto de Datawarehousing en Julio de 1998. Estuvieron basadas en 141 respuestas de administradores de

tecnología y usuarios de negocios de las más grandes corporaciones que son miembros del TDWI.

1.3 El Rol del Marketing

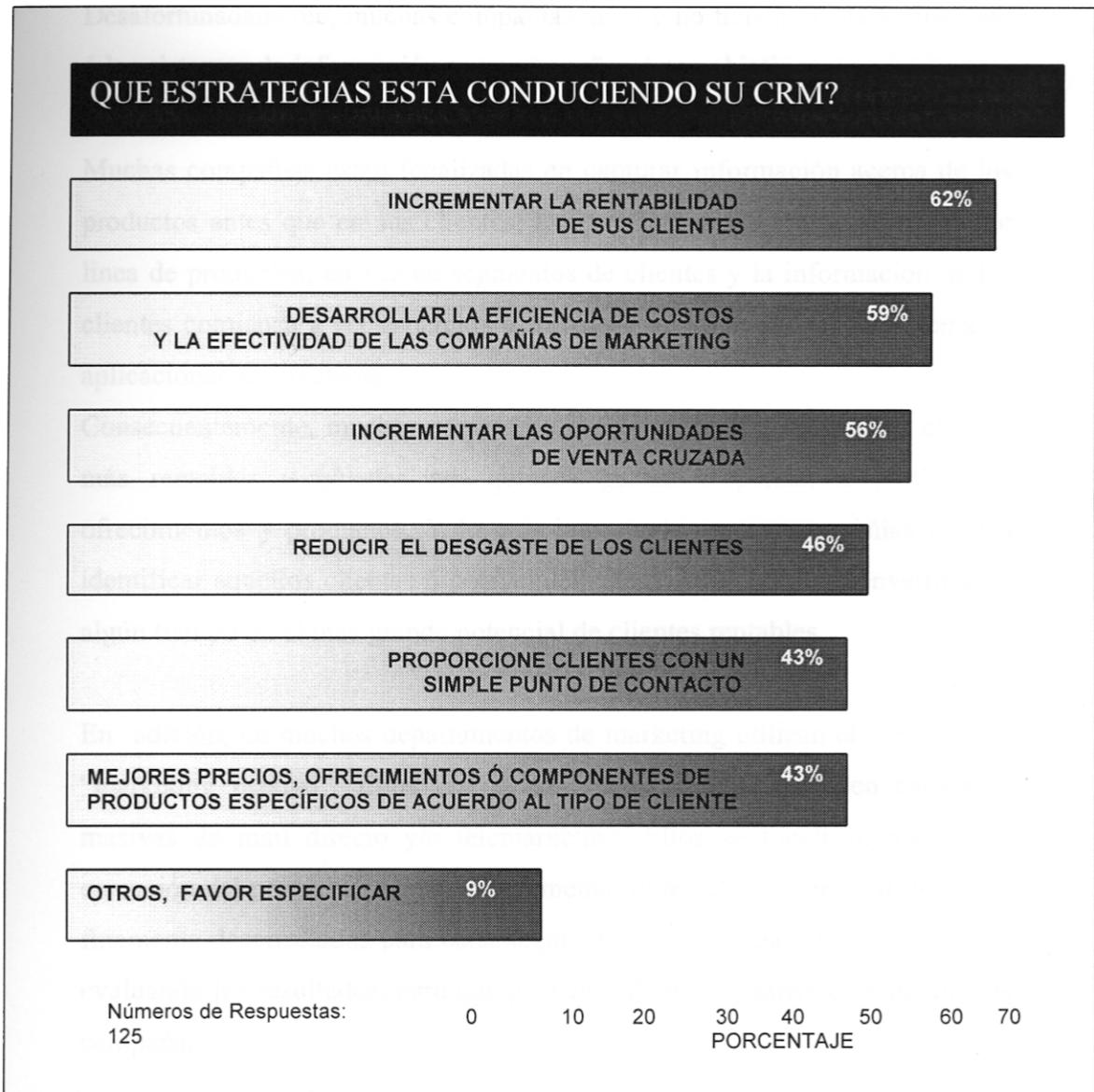
Una de las más fértiles áreas para el ARC es el departamento de Marketing, el cuál es el responsable de capturar las preferencias y actitudes de los clientes. Marketing habitualmente mantiene una base de datos de clientes y potenciales clientes de la compañía. El departamento de Marketing juega un rol crítico en la ayuda a la empresa para entender y optimizar todas las interacciones con sus clientes. Es el nexo de la empresa, proporcionando el contexto para que las actividades como las ventas, servicio al cliente, facturación, manufactura y otra, estén orientadas al cliente.

El Departamento de Marketing ayuda a encontrar y ejecutar estrategias para:

1. - Obtener nuevos clientes a menor costo, a través de campañas orientadas.
2. - Identificar y retener clientes rentables.
3. - Moverse de clientes no rentables a productos, servicios y canales más rentables.
4. - Incrementar el nivel de gasto por cliente a través de las campañas de cross-selling ó upselling.
5. - Monitorear la calidad de los servicios al cliente y sugerir cambios para mejorar su satisfacción.

En una encuesta que se realizó a diferentes empresas en el área de marketing, cuando se preguntó que estrategias están conduciendo las iniciativas de ARC de las empresas, casi la tercera parte (62%) dijo "incrementar la rentabilidad de los clientes leales" (ver Figura 2). Otro 59% dijo desarrollar al mínimo costo la eficiencia y efectividad de las campañas de marketing, e-mail, un 56% mencionó incrementar sus oportunidades de cross-selling. (Los participantes podían dar más de una respuesta). El 46% dijo "reducir el desgasto de los clientes", como su estrategia crítica. Un 43% dijo que

proporcionar un cliente con un simple punto de contacto. Otro 43% mencionó ofrecer a la medida, precios, productos y componentes que atraigan grupos específicos.



Traducido de *Journal of Data Warehousing*
 Vol. 5, # 4, 2000

Figura 2

1.4 La administración de las informaciones es la clave.

En modernas economías como la de hoy, la clave para conseguir las estrategias antes mencionadas es una efectiva administración de la información. Los departamentos de Marketing necesitan desarrollar sistemas

de base de datos de clientes que capturen y consoliden los millones de transacciones de los clientes en formatos digeribles para la toma de decisiones.

Desafortunadamente, muchas compañías de hoy no tienen el Marketing listo ó los sistemas de información que acompañen éstos objetivos.

Muchas compañías están focalizadas en capturar información acerca de los productos antes que en sus clientes. Los trabajadores están organizados por línea de productos, en vez de segmentos de clientes y la información de los clientes comienza a ser esparcida a través de la organización en sistemas y aplicaciones no efectivas.

Consecuentemente, muchas compañías no conocen quienes son sus clientes más rentables y quienes no, quienes gustan responder a particulares ofrecimientos y productos. Y lo más importante, pocas compañías pueden identificar aquellos clientes ó potenciales clientes que pueden convertirse en algún tiempo en el más grande potencial de clientes rentables.

En adición, en muchos departamentos de marketing utilizan el concepto de “marketing masivo”. Ellos una ó dos veces al año conducen campañas masivas de mail directo y/o telemarketing. Ellos se focalizan más en la ejecución del programa, en vez de segmentar el mercado y preparar mensajes finamente desarrollados para cada segmento. También pasan algo de tiempo evaluando los resultados, para que puedan refinar el desarrollo de las futuras campañas.

Parte del problema es que el departamento de marketing, tradicionalmente ha sido el menos automatizado de la compañía. El Marketing ha sido generalmente considerado más un arte antes que una ciencia, se confiaba en la intuición y experiencia de la gente de marketing para preparar campañas de mail directo, estrategias de productos y programas de incentivos a los clientes.

En adición, cuando marketing ha utilizado sistemas de información, ellos han sido sistemas propietarios costosos.

Afortunadamente, una nueva generación ha abierto tecnologías de información que han dado la posibilidad a la gente de marketing para automatizar sus actividades. Muchas compañías están administrando la data de sus clientes utilizando poderosas bases de datos relacionales, a un costo efectivo basado en Unix ó Windows NT. También está a disposición paquetes de aplicaciones y herramientas para asistir y analizar la data de los clientes y ejecutar estrategias de ARC.

1.5 Factores críticos de éxito en la construcción de sistemas ARC.

Para identificar los factores críticos de éxito en la construcción de sistemas ARC, TDWI¹ reclutó más de una docena de ejecutivos y administradores de las más grandes corporaciones y conversar en un worksshop en cómo desarrollar con éxito soluciones ARC.

Cada una de éstas personas tiene varios años de experiencia en ARC y vienen de compañías como: Fleet Financial Services, Federal Express, Eddie Bauer, HR Block, PRIMESTAR, Knoxville Utilities Board, and US West.

Sus tips y técnicas recayeron en tres categorías: el negocio, la gente y la tecnología.

Los siguientes son los resultados:

1.5.1 Negocio

1.5.2 Personal

1.5.3 Tecnología

¹ TDWI The Data Warehousing Institute

1.5.1 Negocio

1. - Obtener un ejecutivo sponsor.

Es imperativo que un ejecutivo, preferentemente el CEO ó CFO, defienda el proyecto.

2.- Crear un comité para administrar el proyecto.

La empresa debería formar un comité compuesto por ejecutivos, usuarios, tecnólogos de la información y administradores del marketing. El propósito de éste comité es definir, guiar y revisar el proyecto ARC.

Este comité debe valorar las capacidades actuales y la cultura de la compañía. Luego identificará las habilidades y tecnologías que deben ser adquiridas. Más importante, es que el comité identifique y priorize los proyectos basados en una cuidadosa valoración de la estrategia de negocio y el ratio riesgo / retorno de cada proyecto.

3. - Crear un plan de negocios

Debe existir un ejecutivo que administre el proyecto y cree un plan de negocios que distribuya los objetivos, costos, retornos, etc. El ejecutivo debe definir la métrica para evaluar el avance del proyecto, así como las tácticas para construir las capacidades requeridas.

Por ejemplo el ejecutivo debe decidir si :

- 1) si utiliza los sistemas y procesos existentes
- 2) si hace la reingeniería de ellos ó
- 3) comienza desde el inicio.

El ejecutivo debe proporcionar el liderazgo al proyecto.

Debe crear los equipos polifuncionales requeridos para definir los procesos, construir los sistemas y entrenar a los usuarios. Ellos deben marketear el proyecto internamente, monitorear los programas, solucionar los problemas que se pudieran presentar y celebrar los objetivos intermedios que se obtengan en el camino.

1.5.2 El Personal

1. - Actualizar el conjunto de habilidades internas.

Pocas compañías tienen la suficiente gente en el marketing ó en el departamento de los sistemas de información con las apropiadas habilidades y experiencia para llevar a cabo por primera vez proyectos ARC. Algunas compañías han implementado sistemas ARC algunas veces olvidándose de los sistemas de marketing y de la gente y otras empresas se han entusiasmado y contrataron expertos de ARC a costos elevados.

La mejor forma es mezclar el equipo existente con uno ó dos consultores externos con experiencia en ARC. Los proyectos deberían ser desarrollados por equipos múltiples de no más de cuatro personas. Estos equipos deberían ser potenciados para tomar decisiones en su dominio y mantener el desarrollo del proyecto en forma eficiente.

2. - Escoger proveedores confiables

Para proyectos grandes es deseable escoger sistemas integradores con experiencia previa en la administración de proyectos ARC.

3. -Publicite el esfuerzo

Muchos proyectos de ARC envuelven considerables cambios organizacionales, por lo tanto es importante publicitar el proyecto periódicamente. Esto incluye revistas internas, boletines e-mail, presentación de los ejecutivos. Algunas compañías usan fuentes externas, como revistas, periódicos, etc.

1.5.3 Tecnología

1. - Implementar una infraestructura escalable.

Ultimamente los sistemas ARC envuelven la reintegración de la empresa alrededor de la información del cliente. Esta es una formidable tarea y requiere la integración de varios sistemas y procesos de negocios, incluyendo las disciplinas del ARC mencionadas líneas arriba.

2. ¿Qué es un Data warehouse?

De acuerdo con W. H. Inmon, quien es considerado como el padre del data warehouse: "Un data warehouse es un conjunto de datos integrados orientados a una materia, que varían con el tiempo y que no son transitorios, los cuales soportan el proceso de toma de decisiones de una administración."

De acuerdo con algunas organizaciones, el data warehouse es una arquitectura. Para otras, es un depósito semánticamente consistente en datos (separados y que no interfieren con los sistemas operativos y de producción existentes) que llenan por completo los diferentes requerimientos de acceso y reporte de datos. Para algunos otros, el data warehouse es un proceso continuo que mezcla los datos de varias fuentes heterogéneas. incluyendo datos históricos y adquiridos para soportar la constante necesidad de consultas estructuradas y/o ad hoc, reportes analíticos y soporte de decisiones.

Así como hay gran divergencia para establecer una definición precisa de un data warehouse, hay un claro consenso de que la tecnología del data warehouse es un ingrediente esencial en el conjunto de soluciones para el soporte de decisiones en una empresa.

Conforme se incrementa el interés en el tema del data warehouse, muchas personas declaran haber construido uno. Con más frecuencia, la gente ha construido un recurso de acceso de último usuario para navegar, reportar y unirse a un Sistema de Soporte de Decisiones (SSD).

Otros han transferido información de sus bases operacionales de datos a una base de datos multidimensional y han unido herramientas de análisis a las bases de datos multidimensionales.

La Software Magazine, en su editorial de Diciembre del 2000 titulado "El Data warehousing merece algo mejor", hace la pregunta: "¿En qué se

convertirá el movimiento del Data Warehousing? Sus metas valen la pena: obtener la información correcta para las personas adecuadas en el momento conveniente para que tomen decisiones que pueden valer millones de dólares. Con eso es suficiente. Son escasas y lejanas las aplicaciones actuales para un almacenamiento más allá de los soportes de decisión rudimentarios. Incluso los campeones en el data warehousing como Aaron Zornes, del grupo META, están de acuerdo en que hace falta todavía mucho trabajo para que el data warehousing llegue a su verdadero potencial."

Muchos de los problemas fundamentales implícitos todavía no se han solventado en forma adecuada -, por ejemplo, los siguientes retos:

- La integración de datos y metadatos de varias fuentes,
- Calidad de información: limpiar y refinar.
- Condensación y adición de datos.
- Sincronización de las fuentes con el data warehouse para asegurar una actualización constante del data warehouse conforme se crean nuevos datos dentro de las fuentes.
- Aspectos del desempeño relacionados con compartir la misma computadora y plataformas RDBMS, tales como la base de datos y las herramientas del data warehouse.
- Administración de metadatos.

Entonces, ¿qué constituye una data warehouse completo? ¿Existe un examen como aquel tan popular para la literatura prohibida? ¿Cómo se anticipa uno a los retos tecnológicos y al correcto diseño de soluciones dentro de la arquitectura del data warehouse? Y ¿Cómo se construye uno en forma sistemática? Todas estas son preguntas válidas.

Un data warehouse es diferente de las bases de datos operacionales que soportan las aplicaciones de un Procesamiento de Transacción en Línea (OLTP, On-Line Transaction Processing). El data warehouse es lo siguiente:

- **Está orientado a una materia.-** Organiza y orienta los datos desde la perspectiva del último usuario. Muchos sistemas operativos organizan sus datos desde la perspectiva de la aplicación, de modo que el acceso de la aplicación a los datos tenga la mayor eficiencia posible. Con frecuencia, la información que está organizada para que una aplicación del negocio la recupere y actualice con facilidad no está organizada necesariamente de modo que un analista con herramientas gráficas inteligentes de consulta puede formular las preguntas empresariales correctas. Esto se debe al enfoque del diseño de la base de datos –la eficiencia de recuperación y actualización de la aplicación- al momento en que se implementó por primera vez.

- **Administra grandes cantidades de información.-** La mayoría de los data warehouse contiene información histórica que se retira con frecuencia de los sistemas operativos porque ya no es necesaria para las aplicaciones operacionales y de producción. Por el volumen de información que un data warehouse debe manejar, también debe ofrecer opciones para la adición y la condensación que clasifican de esta inmensa cantidad de datos. En resumen , los actuales usuarios de un data warehouse buscan “árboles en el bosque”. Por lo tanto, un data warehouse maneja información a diferentes niveles de granularidad. Por la necesidad de administrar toda la información histórica y además los datos actuales, un data warehouse es mucho mayor que las bases de datos operacionales.

- **Guarda información en diversos medios de almacenamiento.-** Por los volúmenes de información que deben manejarse, un data warehouse frecuentemente guarda información en diferentes medios de almacenamiento.

- **Comprende múltiples versiones de un esquema de base de datos.-** Debido a que el data warehouse tiene que guardar información histórica y administrarla, y como la información histórica ha sido manejada en distintos momentos por diferentes versiones de esquemas de bases de datos, en ocasiones el data warehouse tiene que controlar información originada en organizaciones de bases de datos diferentes.

- **Condensa y agrega información.-** Con frecuencia, es muy alto el nivel de detalle de la información guardada por bases de datos operacionales para cualquier toma de decisiones sensata. Un data warehouse condensa y agrega la información para presentarla en forma comprensible a las personas. La condensación y adición es esencial para retroceder y entender la imagen global.

- **Integra y asocia información de muchas fuentes de información.-** Debido a que las organizaciones han administrado históricamente sus operaciones utilizando numerosas aplicaciones de software y múltiples bases de datos, se requiere de data warehouse para recopilar y organizar en un solo lugar la información que estas aplicaciones han acumulado al paso de los años. Esta es una tarea desafiante por la diversidad de tecnologías de almacenamiento, de técnicas de administración de bases de datos y de la semántica de los datos.

2.1 Los requerimientos de un Data Warehouse

¿Cómo se empiezan a formular los requerimientos para un data warehouse? Existen muchas formas de considerar un data warehouse. Las siguientes son algunos ejemplos:

- Un data warehouse es simplemente un sistema de aplicación empresarial con su propia base de datos. Esta base de datos se genera a partir de otras bases de datos operacionales, no de información inicial que se introduce. El

data warehouse ofrece una serie de características y funciones para implementar procesos empresariales y enlazarlos con otros procesos fuera del ámbito del data warehouse. En forma muy similar a otros sistemas de aplicaciones empresariales, se requiere que el data warehouse proporcione al usuario final estos conjuntos prescritos de características y funciones del modo más eficiente posible.

- Un data warehouse es una capacidad latente. Almacena información resumida que se organiza de acuerdo con temas empresariales, tales como clientes y productos, para analizar la información con más facilidad. La carga de mostrar, organizar y reportar la información que guarda el data warehouse corresponde a las herramientas que deben incorporarse en el data warehouse. En esta visión, un data warehouse tiene una capacidad latente que sólo se vuelve útil cuando las herramientas de análisis y reporte se aplican con inteligencia a los datos que conserva el data warehouse. Se requiere que el data warehouse sustente un extenso rango de herramientas de acceso, operadas por una extensa gama de usuarios finales. El data warehouse también debe guardar y administrar un ámbito de información grande para servir a una extensa clientela.

- El data-warehouse es una base de datos históricos, la cual es una acumulación de muchos años de información transaccional en línea, organizada para hacer eficiente el almacenamiento y facilitar la recuperación. Es necesario que el data warehouse organice grandes cantidades de información de manera compacta y eficiente. También se requiere que proporcione técnicas para resumir, a fin de que los usuarios finales comprendan las lecciones de los antecedentes con más facilidad.

- En ocasiones, el data warehouse es una tienda de datos operacionales. Entrega información operacional a un amplio rango de usuarios copiando información de los sistemas operacionales de bases de datos. En este caso, se requiere que el data warehouse distribuya información operacional de manera

eficiente a un gran rango de usuarios. También se necesita que el data warehouse haga los cambios tecnológicos necesarios para mover la información de su base de datos operacional a la tecnología de almacenamiento que se emplea en el data warehouse.

Es evidente que los requerimientos de un data warehouse son tan variados y diversos como las clases de usuarios que lo utilizan para obtener beneficios empresariales. Por lo tanto, es necesario clasificar los requerimientos del data warehouse utilizando técnicas clásicas.

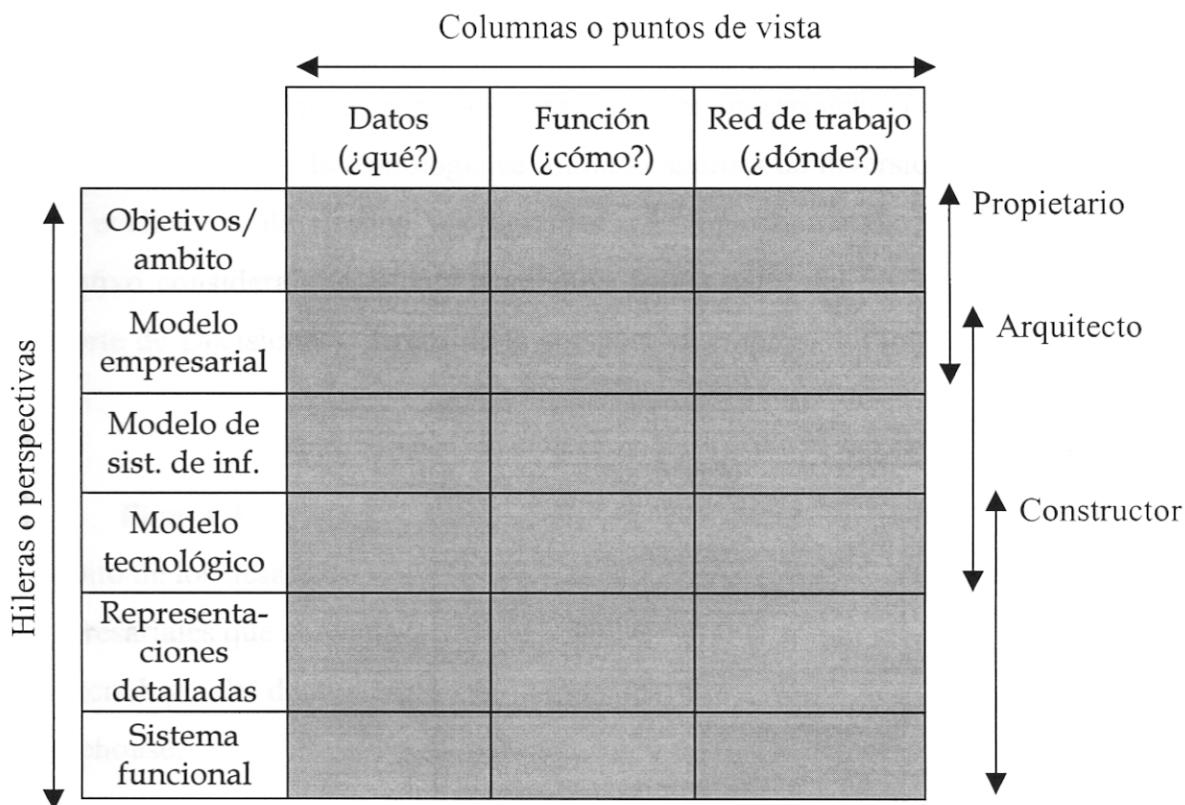


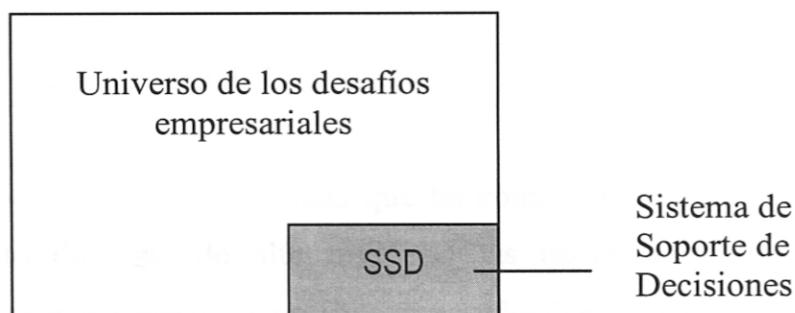
Figura 3a

El diagrama de Zachman (Figura 3a), es una de las formas más eficaces de visualizar un sistema desde muchas perspectivas. Los requerimientos que son visibles desde cada una de las perspectivas son los que proponen las personas que tienen esas perspectivas. Sin embargo, los requerimientos exactos para un data warehouse se analizan en secciones posteriores, ahora revisaremos las bases necesarias para entender las distintas perspectivas de un data warehouse.

En su documento, Zachman se refirió a las siguientes personas que tenían interés en un Sistema de Información:

- El dueño, o inversionista en el Sistema de Información, que intenta resolver un problema empresarial utilizando la Tecnología de la Información. Por lo general, el punto de vista del dueño se mide en dinero y la perspectiva que tiene del negocio es periodos y procesos empresariales. En general, el interés del dueño en la tecnología se limita al análisis de inversiones y pagos, y la evaluación de riesgos y programas. El inversionista o patrocinador ejecutivo considera que el data warehouse forma parte del SSD (Sistema de Soporte de Decisiones), dentro de la perspectiva empresarial total (ver figura 3).

Figura 3
Ámbito de los desafíos
Empresariales que solventa
La tecnología del data
Warehouse.



En el SSD, el dueño confía que al data warehouse ofrece análisis y razones para tomar decisiones empresariales. Gran parte de la capacidad de análisis con la que cuenta el data warehouse se basa en la información histórica y la extrapolación de las tendencias. Las áreas donde el inversionista requiere

Soporte de Decisiones son las que afecta el negocio en forma integral (Ventas, Mercadotecnia, Finanzas, Administración estratégica, Planeación y desarrollo de productos, Atención a clientes, Recursos humanos, etcétera).

- El usuario, que debe utilizar el data warehouse con regularidad para realizar funciones empresariales. La perspectiva que tiene este usuario del data warehouse está muy bien definida y prescrita como un conjunto de procesos y pasos que debe seguir para llevar a cabo el análisis y la función empresarial. Con frecuencia, al usuario empresarial no le interesa cómo se forma el data warehouse o como está estructurado. Sólo necesita que el data warehouse realice las funciones señaladas correcta y eficientemente.

- El arquitecto, que debe entender tanto las necesidades empresariales del data warehouse, como la tecnología de implementación que se necesita para solventarlas. El arquitecto debe:
 - Conocer las necesidades empresariales a corto y a largo plazo.
 - Entender las inversiones en tecnología actual y planeada en las plataformas computacionales, y los sistemas de administración de base de datos.
 - Asimilar la disponibilidad de los datos que proporcionan los sistemas operacionales.
 - Comprender los elementos de seguridad de información.

El arquitecto es el negociador entre el inversionista que ha comisionado un sistema basado en un punto de vista de alto nivel de las necesidades empresariales y el constructor que necesita especificaciones detalladas para preparar un data warehouse.

Desde el lado empresarial, el arquitecto busca las entidades y procesos del negocio capturados como información en un data warehouse. También observa los distintos modos en los que debe organizarse la información en el

data warehouse para sustentar las necesidades del usuario final. Desde el aspecto tecnológico, el arquitecto observa cómo interactuará el data warehouse con los sistemas operacionales existentes y las aplicaciones para cargar el data warehouse con información y mantenerla actualizada y consistente. Muchas veces, también es responsabilidad del arquitecto que la implementación del data warehouse vaya de acuerdo con las inversiones actuales en tecnología.

- El constructor, quien es el responsable de la instalación e integración física de diversos componentes del data warehouse. El constructor incorpora los requerimientos operacionales y de despliegue en el data warehouse, como lo son la disponibilidad, la capacidad de recuperación, y la seguridad. Un ejemplo de la perspectiva que tiene el constructor del data warehouse es el del sistema de Misión Crítica cliente / servidor, el cual se basa en la tecnología relacionar, las redes de área local y se implementa en C y C++.

El diagrama de Zachman también describe las distintas dimensiones que pueden adquirir estas perspectivas. Zachman describe las perspectivas de datos como el ¿Qué?, las perspectivas de proceso como el ¿Cómo?, la perspectiva geográfica o de red como el ¿Dónde?, los fundamentos del negocio como el ¿Por qué?, y después añade la dimensión del ¿Cuándo?

Si recuerda que el diagrama de Zachman representa dos dimensiones de un sistema, las hileras representan las distintas perspectivas de los depositarios (el dueño, el arquitecto y el constructor), y las columnas representan la visión del sistema (los datos, el proceso y la red).

2.2 La necesidad de una arquitectura común

Así como el diagrama de Zachman sirve para separar las diferentes perspectivas del data warehouse, se requiere un diagrama que separe los distintos componentes de una solución del data warehouse. Este diagrama

común (Ver figura 4), llamado la arquitectura de referencia permitirá integrar en una clasificación común los distintos tipos de información necesarios para construir un data warehouse. Esta clasificación permitirá comparar las opciones de distribuidores competidores, evaluar ventajas y desventajas, y descubrir brechas que deben llenarse con un desarrollo de software interno. La palabra referencia significa que la arquitectura es independiente de un distribuidor y caracteriza la naturaleza "genérica" de todos los data warehouses.

¿Por qué es importante la arquitectura de referencia? La arquitectura de referencia aporta lo siguiente:

- Ofrece un diagrama de un “anteproyecto” común. Es necesario que el inversionista, el arquitecto y el constructor tengan una terminología y un marco de referencia comunes cuando se discuten costos, estimados, riesgos, y el avance en el programa de implementación.

- Crea una base duradera para implementar la visión de la empresa. El diagrama es un método duradero de capturar un punto de vista. Con frecuencia, una visión es grande y de largo alcance, pero las posibilidades de recursos y el gasto son limitadas. La arquitectura de referencia permite apreciar la visión global en fases, sin alterarla.

- Proporciona alternativas en la implementación. Por ejemplo, muchas organizaciones han contemplado los mercados de datos y los data warehouses como instalaciones separadas. Los mercados de datos sirven a un grupo pequeño de clientes con necesidades departamentales muy específicas, mientras que el data warehouse sirve para grandes empresas. Es muy conveniente un diagrama que separe bien estos dos servicios, y permita que una organización construya una serie de mercados de datos y, mas tarde, un data warehouse.

- Permite ubicar las ofertas de distribuidores en el diagrama de arquitectura de referencia. Debido a que la tecnología del data warehouse aparece primero en el ciclo de la maduración de la tecnología, las soluciones las aporta el distribuidor, y no el cliente. Cada distribuidor utiliza una terminología distinta para sus componentes de solución. Un diagrama de referencia común permite al cliente evaluar y seleccionar, incluso cuando tenga que comparar peras con manzanas.

- Destaca los componentes de una solución que son valiosos para la producción. Conforme se despliega la solución de un data warehouse, surgen los elementos relacionados con el uso diario y de producción de la solución. Esto debe resolverse añadiendo componentes a la solución inicial. Incluir la arquitectura de referencia como parte de la implementación y el despliegue permite al constructor o al personal de despliegue identificar los componentes y añadirlos al diagrama de soluciones.

Utilizar la arquitectura de referencia como base para construir un data warehouse produce los siguientes beneficios:

- **Ventajas y desventajas arquitectónicas** - La arquitectura de referencia divide la solución en componentes bien definidos. Cada uno de los componentes puede construirse o comprarse completo, adquirirse en partes e integrarse en una solución. Por ejemplo, es posible comprar un administrador de data warehouse completo en Prism Solutions, construir un data warehouse propio utilizando un sistema de administración de base de datos relacionar Sybase, o integrar un Red Brick Warehouse como componente en una aplicación existente para el manejo de metadatos. Por lo tanto, la arquitectura de referencia ofrece una base para el análisis de ventajas y desventajas.

- **Estimados de costo / beneficio para el análisis de inversión** - La arquitectura de referencia también divide la solución en costos identificables o componentes que se pagan. Por ejemplo, tener información dentro del data

warehouse es un costo que se compensa con las herramientas de soporte de decisiones que usan esta información. Los inversionistas tienen que balancear el costo de cargar y sincronizar los componentes del data warehouse contra la retribución potencial de consultas y análisis.

▪ **Determinación de la capacidad de reuso de las inversiones existentes-**

La arquitectura de referencia contiene componentes que se relacionan con la infraestructura de la tecnología. A muchas organizaciones les gustaría preservar sus inversiones actuales en tecnología. Con la arquitectura de referencia, es posible para estas organizaciones ubiquen sus inversiones actuales en los componentes de la arquitectura y sólo integren y compren los componentes que no están disponibles.

▪ **Valoración de los riesgos tecnológicos** - Cada uno de los componentes en la Arquitectura de referencia está en una curva diferente de la madurez de la tecnología. Esto se debe a que la complejidad y sofisticación de cada uno de los componentes y al estado de la práctica en la industria para construir estos componentes. Con la arquitectura de referencia, es posible separar los componentes de alto y bajo riesgo basándose de su posición en la curva de madurez de la tecnología.

▪ **Valoración de los "aspectos de producción"** - La arquitectura de referencia muestra los componentes que se usan a diario. Por lo tanto, estos componentes deben activarse para ayudar a la producción, y no para un uso casual u ocasional. Deben listarse las cualidades de la fuerza de producción, y evaluarse y reforzarse los componentes que las requieran. La lista de las cualidades de producción también se usa para que el distribuidor evalúe los componentes.

▪ **Selección de componentes arquitectónicos compatibles para la facilidad en la integración** - La naturaleza compleja de las soluciones de data warehouse, y el hecho de que ningún distribuidor ofrezca una solución

satisfactoria que cumpla con las necesidades del cliente, determina la combinación de componentes compatibles y entrega oportuna que integran la solución.

▪ **Valoración de la compatibilidad de las inversiones existentes y los ofrecimientos de los distribuidores** - Las inversiones existentes son una realidad empresarial. Cualquier propuesta de un distribuidor para reemplazar una inversión debe compararse con las propuestas que complementan y extienden las inversiones actuales o planeadas.

▪ **Selección del ámbito de implementación** - Un modo de administrar el presupuesto de un proyecto de data warehouse es manejar el ámbito de su implementación. Esto se maneja de muchas formas, como las siguientes:

- Selección de metadatos.
- Selección de fuentes de datos.
- Implementación de un solo mercado de datos.
- Selección de técnicas que pueda manejar un data warehouse (v.c.,, sólo sistemas de base de datos relacional).

▪ **Integración de un grupo de personal y planeación del mejoramiento de las técnicas** - Como se mencionó antes, la planeación, implementación, despliegue y uso de un data warehouse requiere de un amplio rango de técnicas. La arquitectura de referencia ubica cada una de estas técnicas en el componente arquitectónico apropiado. Esto permite una distribución planeada del personal del proyecto, con base en la identificación de las técnicas y cuando se requerirán.

▪ **Decidir entre construir y comprar** - La arquitectura de referencia progresivamente separa la solución de data warehouse en niveles con cada vez más detalles. A cualquier nivel, puede presentarse la disyuntiva de construir o comprar cada componente. Esto permite a una organización

integrar los componentes que necesita y sólo construir los componentes que no existen en el mercado.

- **Creación de un programa completo y un plan del proyecto** - La arquitectura de referencia aporta una base para desarrollar planes del programa y del proyecto que dependan de los componentes que deben comprarse, desarrollarse o integrarse. Esta separación de componentes también permite secuenciar el aprovisionamiento del distribuidor y apoyar su incorporación en el programa global y el plan del proyecto.

2.3 Arquitectura de referencia

La arquitectura de referencia se describe primero desde un punto de vista abstracto y simplificado a alto nivel, del modo siguiente:

- Un conjunto de datos extraídos de bases de datos operacionales.
- Un software que prepara los datos para que los accedan los usuarios.
- Un conjunto de aplicaciones y herramientas que ejecutan un conjunto de consultas y análisis complejos.

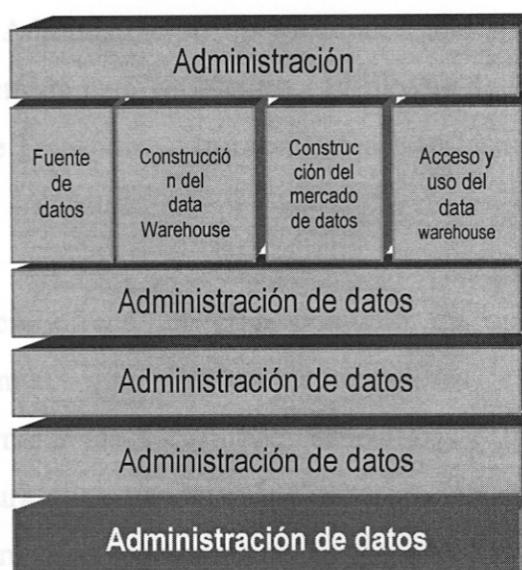
Después, descompondremos sistemáticamente la arquitectura de referencia en mayores detalles. Esto ayuda a comprender un diagrama conceptual simple, al mismo tiempo que expone los detalles necesarios para una implementación ordenada y correcta. Donde sea posible, se usará un lenguaje popular y inteligible. La industria sigue en su infancia y los términos y definiciones no han sido estandarizadas. En la actualidad, varios distribuidores están promoviendo diferentes términos en su literatura de mercadotecnia.

La figura 4 muestra los componentes de nivel superior de la arquitectura de referencia. La arquitectura de referencia del data warehouse divide los componentes del data warehouse en bloques como la Fuente de datos, el Mercado de datos, el data warehouse, Acceso y uso, y capas tales como

Administración de datos, Administración de metadatos, Transporte, e Infraestructura. Los bloques se relacionan con la funcionalidad específica del data warehouse. Las capas representan el ambiente necesario para implementar los bloques.

Figura 4

Un panorama de la arquitectura de referencia del data warehouse



La mezcla de los enfoques del data warehouse y del mercado de datos se refleja en la arquitectura de referencia. Por lo tanto, la arquitectura de referencia ofrece la mayor cantidad de opciones de implementación por un arquitecto. Las capas horizontales de Administración de Datos y Administración de Metadatos corresponden a las actividades que se relacionan con la extracción, carga y actualización que se requiere para mantener e abasto de información para el data warehouse y el mercado de datos.

Las otras capas horizontales corresponden a los servicios comunes que son esenciales para los bloques del data warehouse y Mercado de Datos. Estos servicios comunes son la Infraestructura y el Transporte. En la arquitectura de referencia, también se representa una capa adicional denominada Herramientas, Tecnologías y Funciones. Esta capa se usa para apoyar la metodología del data warehouse, la cual permite que la construcción y el

despliegue sean controlados por el proceso, las herramientas y la metodología.

Los bloques de construcción del data warehouse, construcción del Mercado de Datos, Acceso al data warehouse y Utilización, junto con las capas de Administración de Datos y Administración de Metadatos son las áreas de inversión nueva en el proyecto de data warehouse. El bloque de Fuentes de Datos, junto con las capas de Transporte e Infraestructura son las inversiones existentes en esa empresa.

La arquitectura de referencia ofrece un gran conjunto de opciones de implementación. Es común que los implementadores escojan la implementación que cubre mejor las necesidades específicas e ignoren los componentes de la arquitectura que consideran innecesarios. La arquitectura de referencia proporciona muchas opciones de implementación para un arquitecto. Es posible hacer lo siguiente:

- Proceder a construir un data warehouse de amplitud empresarial, un conjunto de mercados de datos, o ambos, mientras se reserva el derecho de posponer la decisión de data warehouse/mercado de datos.
- Seleccionar pocas fuentes de información para construir el data warehouse, mientras se reserva el derecho de admitir más en el futuro.
- Definir una plataforma múltiple, con arquitectura de varios niveles basada en que se usan diferentes tipos de plataformas para los distintos componentes arquitectónicos (Fuentes de Datos basadas en macrocomputadoras, Data warehouse basado en macrocomputadoras, Mercados de Datos basados en un servidor, Acceso y uso de usuario final basados en una estación de trabajo).
- Seleccionar los componentes tecnológicos para el Transporte (Middleware) y la Infraestructura (RDBMS y tecnología de base de datos IMS) que coincidan con la elección arquitectura de plataforma múltiple y varios niveles para las Fuentes de Datos, el Data warehouse, Mercados de Datos, y el Acceso y Uso de usuario final.

El siguiente grupo de figuras (5 a 12) muestra los componentes asociados con cada uno de los bloques arquitectónicos principales en la figura anterior. Los componentes que se muestran aquí corresponden a la tecnología que se conoce y usa actualmente. El diagrama arquitectónico permite agregar componentes adicionales cuando se cuenta con ellos y se usan en todas partes, En estas figuras, los bloques sombreados representan actividades relacionadas con la Administración de Metadatos. Los metadatos provienen de casi todas las actividades que operan y cambian la información dentro del data warehouse, en su recorrido desde la Fuente de Datos hasta el Data warehouse o el Mercado de Datos.

2.4 El Proceso de Maduración de un Data warehousing

La administración de las relaciones con el cliente (ARC), representa un nuevo recurso para administrar los negocios de la empresa y sus clientes. Lo básico del ARC es “el conocimiento del cliente”, usualmente contenido y obtenido en un Data Warehouse ó Data Mart.

Un Data Warehouse es una mezcla sofisticada de tecnología de base de datos, inteligencia de negocios y análisis detallado, con relación al marketing y los procesos de negocios. La implementación exitosa de un Data Warehouse depende de un proceso de aprendizaje colectivo y creativo del uso de esta tecnología.

Muchas empresas engloban éste concepto a través de las tres distintas fases evolucionarias de los Sistemas de Soporte a las Decisiones (SSD): reporte y organización de la data de los clientes, análisis y evaluación de esta data y el uso de los resultados del análisis para predecir patrones de comportamiento y identificar nuevas oportunidades de negocio. El resultado del proceso de maduración del Data Warehouse es altamente usado para centrarse en el conocimiento del cliente para encontrar oportunidades de negocio que interesen a los clientes, basados en su comportamiento y requerimientos.

Esto crea un ambiente para una alta retención de los clientes, avanzada captación de clientes y una extendida lealtad de los mismos.

2.5 El Proceso de Maduración en el Data Warehouse

El mundo se está transformando de sistemas de información basados en redes a sistemas basados en tecnologías centradas a los clientes. Transacciones inteligentes proporcionan la parte fundamental de la inteligencia en los negocios. El proceso de maduración del Data Warehousing , permite a las empresas transformarse ellas mismas de organizaciones centradas en los datos a organizaciones centradas a los clientes.

El Futuro

Tecnologías relacionales se convertirán en la norma para el uso de la información y la base para el conocimiento del cliente.

Nos estamos moviendo rápidamente de una tecnología de información a una tecnología relacional, y el efecto positivo será altos ingresos y beneficios para las compañías que lo adopten.

La información y el conocimiento le permiten a Ud. definir nuevas relaciones con sus clientes y reinventar sus organizaciones.

Las relaciones con sus clientes es un resultado esencial de usar la data en interacciones y transacciones combinadas con el conocimiento de los clientes a través del uso de tecnologías relacionales.

2.6 Definiendo actividades en un ambiente Data Warehouse

La creación de las bases del conocimiento, algunas veces llamado Data Warehouse ó Info-Estructura, proporciona oportunidades rentables para administrar los negocios y definir y analizar el comportamiento de los

clientes, con el objetivo de desarrollar y mejorar las relaciones a mediano y largo plazo.

| Definición de Actividades en un ambiente de Data Warehouse | |
|---|--|
| 1. Almacenar | Debe manejar un amplio monto de data detallada: cada transacción de negocios, cada llamada telefónica, cada llamada solicitando ayuda, cada compra, cada factura, etc. Nuevas oportunidades y ventajas competitivas aparecen en la noche: una data detallada permitirá una reacción inmediata. |
| 2. Actualizar | El datawarehouse es continuamente actualizado con los negocios y de marketing. No cada año ni cada mes, debe ser actualizado en forma diaria y continua |
| 3. Uso | Es usado por una gran cantidad de gente en marketing y la administración de otros departamentos. |
| 4. Consultas | La base de datos debe ser consultada y evaluada para identificar nuevos patrones. |
| 5. Operación | El sistema debe estar disponible para su uso en cualquier momento porque es el corazón operacional del marketing y la administración de la compañía. |
| 6. Escala | El Datawarehouse debe ser escalable. Debe poder crecer con el éxito de la compañía y con las demandas cada vez más sofisticadas del Departamento de Marketing. |
| 7. Seguridad | Debe poder proporcionar una apropiada protección a la data. El público general esta muy preocupado porque los detalles de su vida privada sean hechos públicos sin su consentimiento |

Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 5

3. Utilizando el ARC y el Data Warehouse a través de las Etapas de Aprendizaje

El Data Warehouse es una mina de oro de información, usada no solamente para realizar predicciones como las discutidas anteriormente, sino también para la administración de los negocios: ¿Cómo fue el último mes en comparación con los años anteriores?¿, Cómo le fue a la región sur compara con la región norte?¿, Quiénes son mis diez principales clientes?¿, Qué sucedería si bajo el costo de las tarifas de las llamadas locales un 50%, como la competencia lo viene haciendo?.

El uso de tecnologías relacionales es un proceso continuo de aprendizaje. Las etapas que conforman el proceso de maduración son interdependientes y

interrelacionadas. Estos períodos de tiempo pueden ser acortados, como veremos más adelante, a través de un temprano entendimiento de las futuras acciones de maduración, experiencias y administración del aprendizaje. Conocer el futuro, basado en las experiencias históricas del pasado, puede acelerar el proceso de maduración y posicionarnos mucho mejor para el aprovechamiento de las oportunidades.

4. Teoría de las Etapas: Una Estructura para el Desarrollo

4.1 Definiendo la Estructura

Durante sus dos décadas en IBM, como señala el autor, se encontró que la Teoría de las Etapas ha sido la piedra angular para la educación de los ejecutivos y los administradores acerca de los clientes y potenciales clientes para asegurar un claro entendimiento de las habilidades y cambios de la organización en lo referido a TI.

La Teoría de las Etapas ayuda a identificar las fuentes y conjunto apropiado de prioridades para el equipo de administración, aceptando los cambios que ocurrirán en función del tiempo. Conociendo que cambios están por venir, haremos más fácil dichos cambios. Conociendo donde está Ud. , actualmente, le permitirá donde Ud. quiere posicionarse.. “Facilitar el cambio” y “Administrar la maduración” acelera las oportunidades, objetivos y sucesos.

4.2 Las seis Etapas del Desarrollo

La Teoría de las Etapas define las siguientes seis etapas del desarrollo. Cuando una organización desarrolla un data warehouse, y los procesos de negocios relacionados, estructuras de organización, y las habilidades de negocios toman forma, cada etapa representa nuevos niveles de capacidad de usar la información inteligentemente y estratégicamente.

1. - Iniciación.- Comenzar: Construir y Aprender
2. - Desarrollo.- Aplicación, Tecnología, la Data expandida
3. - Control.- Intervención de la administración/ Posicionamiento

- 4. - Integración.- La Info-Estructura de la Empresa
- 5. - Distribución.- Compartir funciones, Usos y Mercados
- 6. - Estrategia.- Compartir tácticas, Planeamiento y Estrategia.

Usando la escala del tiempo, como en la Fig.6, facilita una apropiada perspectiva.



Figura 6

SEIS ETAPAS EN EL DESARROLLO DEL PROCESAMIENTO DE DATOS

| Procesos de Desarrollo | Etapa I Iniciación | Etapa II Contagio | Etapa III Control | Etapa IV Integración | Etapa V Administración de la Información | Etapa VI Madurez |
|--------------------------------------|---|-------------------------------------|---|--|--|---|
| Cartera de Aplicaciones | Aplicaciones para la reducción de los costos de las funciones | Proliferación | Mejor documentación y reestructuración de las aplicaciones existentes | Reajuste de las aplicaciones existentes empleando la tecnología de base de datos | Organización-Integración de las aplicaciones | Integración de la aplicación "reflejando" los flujos de información |
| Organización del PD | Especialización para el aprendizaje tecnológico | Programadores orientados al usuario | Administración media | Infraestructura de los servicios de computación y las cuentas de los usuarios | Administración de la información | Administración de las fuentes de información |
| Planeamiento y Control del PD | Laxos | Más laxos | Planeación y Control formalizados | Control y planeación hechos a la medida | Sistemas comunes y de datos compartidos | Planeación estratégica de las fuentes de información |
| Posición del Usuario | "No intervención" | Entusiasmo superficialmente | Se le hace responder arbitrariamente por ciertos costos | Concientización de la responsabilidad de ciertos costos | Verdaderamente responde por ciertos costos | Aceptación de responsabilidad conjunta del usuario y del PD |

Cuadro 1

Teoría de las Etapas ó Curva de Nolan se aplica como una ayuda para planear los sistemas de información. El presente análisis se realiza utilizando como marco el último artículo publicado por Richard L. Nolan, titulado “Cómo administrar las crisis en el procesamiento de datos” publicado en la colección de Biblioteca Harvard. En dicho artículo se presentan seis etapas de desarrollo. Dicho esquema que viene a reemplazar al concepto de cuatro etapas (ver anexo 2.4) que se describió en 1,974. Las cuatro etapas que se describió entonces continúan siendo válidas, empero la experiencia que se ha obtenido recientemente revela un panorama más amplio e intrigante.

Gracias a que hoy en día es posible analizar las experiencias que muchas compañías han obtenido con los sistemas de procesamiento de datos, los ejecutivos pueden basarse en observaciones recientes e importantes para el proceso de toma de decisiones. En primer término, podemos contemplar hechos pasados y futuros entre los que se destacan las seis etapas del desarrollo de la sistematización. Aún cuando ninguna compañía ha entrado todavía a la sexta etapa, algunas se aproximan a ella y un gran número se encuentra en las etapas intermedias. En la tercera se produce un salto notable en los costos de computación que mostraban ya una tendencia ascendente; la cuarta presenta la necesidad de los usuarios por adquirir un mayor control sobre los programas de procesamiento de datos; y la quinta y la sexta muestran el desarrollo y la madurez de un nuevo concepto de la administración de datos. Para los gerentes de sistemas y para los usuarios de los programas, ésta evolución tiene importantes implicaciones ya que la planeación, el control, las operaciones, la tecnología y los costos se ven afectados profundamente. Al emplear los lineamientos descritos, los directivos podrán determinar en qué punto del proceso evolutivo se encuentra su organización y serán capaces de enfrentarse de una manera mejor a las perspectivas de crecimiento que se abren ante sus ojos.

4.3 Aplicaciones de la Teoría de la Etapas

Existen varios métodos de usar la Teoría de las Etapas en la visualización y mejorar la comunicación de las experiencias, cambios, recursos y criterios de éxito. Algunas empresas lo han utilizado para entender el presente y posicionarse tecnológicamente en el futuro.

Otras lo han utilizado para aprender el uso de la tecnología, ó en la formulación de sus bases de datos, ó en la visión general de la inversión en tecnología ó en entender los cambios con la introducción y puesta en marcha de tecnologías.

5. Sistemas de Soporte a las Decisiones (SSD)

5.1 Las Tres Fases del SSD

5.2 Tipos de SSD

5.1 Las Tres Fases del SSD

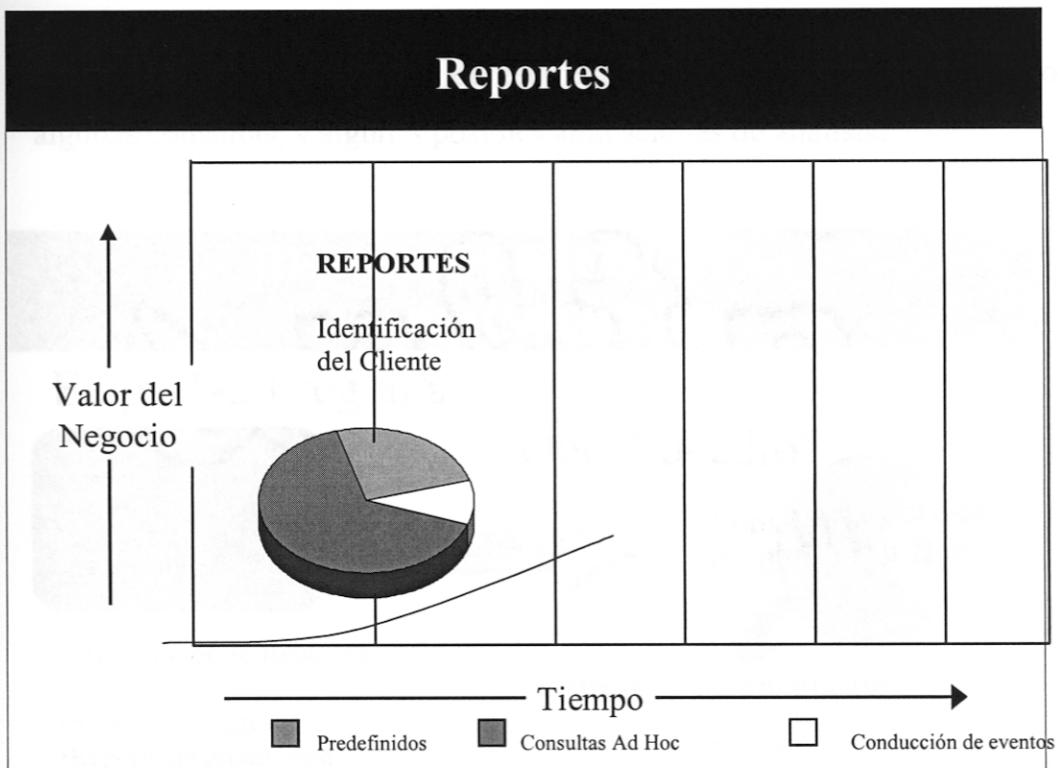
La Teoría de las Etapas puede ser usada para definir las tres mayores fases del desarrollo de los sistemas de soporte a las decisiones. Cada una de éstas fases representa una concertada evolución organizacional que trasformará no solamente los sistemas de procesamiento de la información, también están relacionados a los procedimientos de negocios, estructura organizacional, y las interacciones de la administración que conducen los negocios de la compañía.

5.2 Tipos de SSD

5.2.1 Data Warehouse/Sistema de Soporte a las Decisiones. Tipo uno : Reportes

El primer tipo de SSD es el caracterizado por una gran cantidad de consultas predefinidas. Este tipo de sistema de “reportes”usualmente proporciona un completo conjunto de diagramas, gráficos, cubos referidos a un área específica de negocios, respondiendo las más frecuentes preguntas referidas a

la empresa, los mercados y sus clientes. En la figura 7 se puede apreciar el uso de éste tipo y el valor para el negocio.



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 7

Estrategia del Data Warehouse

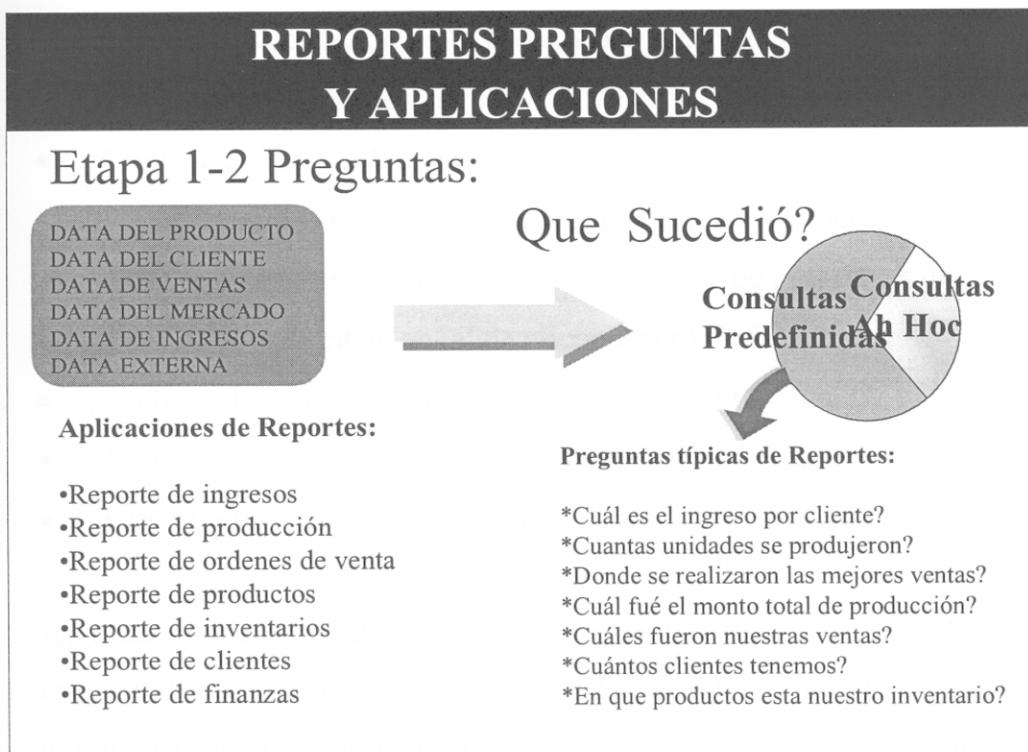
Las consultas son conocidas, y la data está mayormente sumariada y presentada rápidamente. Se requiere respuesta a la pregunta estratégica: ¿Qué sucedió?.

Algunos ejemplos de data warehouse son:

- Cuáles son el total de utilidades, ventas, gastos, volúmenes ó productos producidos?

- ¿Donde están las mayores ventas, utilidades, despachos, ó servicios ocurridos?
- ¿Cuáles son las diferencias ó comparaciones con el periodo pasado?
- ¿Cuál es nuestra mayor / menor recurso productivo (dinero, productos, transporte, gente, etc.)?

La Figura 8 muestra el uso de ésta etapa, Incluye la mayoría de reportes, algunas consultas, y algunas posibles aplicaciones de análisis.



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 8

SSD : Reportes- Preguntas Clave

La Etapa uno representa el punto central ya que proporciona la mayoría de data en forma inmediata de acuerdo a los requerimientos de los usuarios. En algunas implementaciones esta limitada a una data sumariada. Cuando se usa un DW para reportes en ARC, se focaliza en definir las características y

hábitos de los clientes. Algunas de las esenciales preguntas usualmente han sido:

- ¿Quiénes son nuestros clientes? (Edad, Ingresos, Sexo, Grupo)
- ¿Donde viven? (Geografía, Economía, Estilo,etc)
- ¿Qué han comprado en el pasado? (Visión histórica)
- ¿Cómo han comprado?(Información de transacciones financieras)
- ¿Cuáles son los más rentables? (márgenes)
- ¿Cuántas veces los hemos contactado?
- ¿Cuándo los contactamos? (Ciclo, Calendario, Eventos,etc)
- ¿Cuáles son las respuestas positivas y negativas en nuestros contactos?
- ¿Cuál es el costo de llegar a ellos vía el canal escogido?
- ¿Cuáles son los grupos de clientes que compran productos similares?
- ¿Cuál es promedio de ingresos por cliente y el de gastos?
- ¿Cuál es el porcentaje anual de clientes perdidos?
- ¿Cuál es el ingreso por producto, por cliente, por canal?
- ¿Cuál es el costo de adquisición anual por canal?
- ¿Con qué forma de pago y en que tiempo cancelan los clientes?

Algunas aplicaciones de Reportes proporcionan algunas respuestas y fueron encontradas en algunas bases de datos de negocios, pero el DW proporciona una nueva visión y una habilidad para usar combinadamente la data detallada a través de la organización para entender el pasado.

5.2.2 Data Warehouse/ Sistema de Soporte a las Decisión. Tipo dos : Análisis

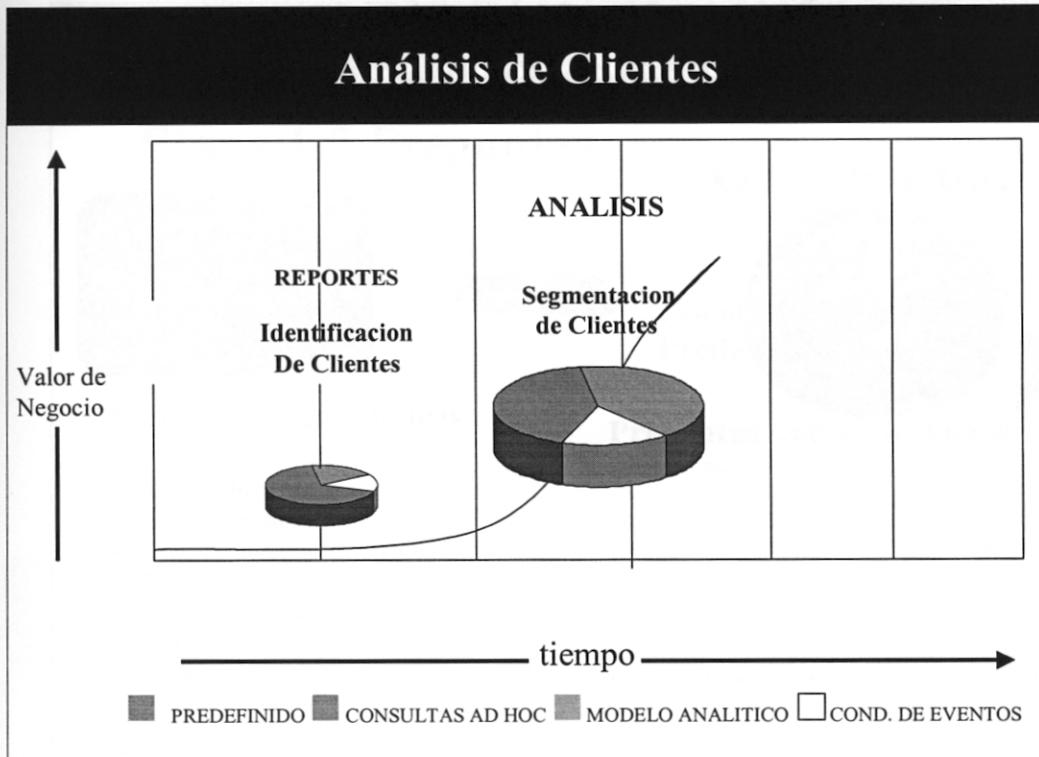
Una vez que hemos aprendido “que pasó” en el uso inicial del DW/SSD, nos movemos a uno más complejo de consultas específicas del segundo tipo de SSD: Análisis. Nos focalizamos en la pregunta: ¿Porqué sucedió?.

Esto es un proceso organizacional para entender los factores que incidieron acerca de los resultados “descubiertos” tempranamente.

Esto es una transformación importante para entender el valor del data warehouse.

Como se muestra en la Figura 9, el uso de la información de los clientes acelera la habilidad de segmentar y analizar los clientes y sus acciones. En adición los tipos de preguntas son más sofisticadas. Algunas consultas típicas al data warehouse son como sigue:

- ¿Porqué nuestro equipo no reunió ó excedió los pronósticos ó los objetivos?
- ¿Porqué los volúmenes fueron bajos ó los despachos fueron más tardes que lo esperado?
- ¿Qué causó la mayoría de resultados positivos ó los márgenes más altos?
- ¿Donde conseguimos actualmente nuestros mejores beneficios?
- ¿Porqué los inventarios ó los recursos no se mueven también?



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 9

SSD: Análisis – Preguntas Clave

Etaapa tres y cuatro del análisis ARC: Focalizándonos en Entender al Cliente.

- ¿Porqué el promedio de rentabilidad de los clientes está bajo?
- ¿Porqué la campaña no alcanzó las metas planeadas?
- ¿Porqué el porcentaje de respuesta es más bajo que antes?
- ¿Porqué la rentabilidad entre productos específicos varía mucho?
- ¿Porqué se ha incrementado el costo de adquisición de éste canal?

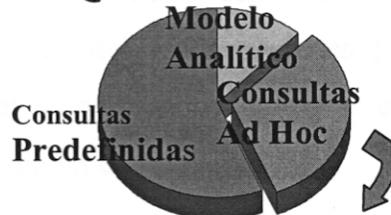
ANALISIS EN ETAPAS TRES Y CUATRO

Etapa 1-2 Preguntas

DATA DEL PRODUCTO
DATA DEL CLIENTE
DATA DE VENTAS
DATA DEL MERCADO
DATA DE INGRESOS
DATA EXTERNA



Que Sucedió?



Aplicaciones en Análisis

Análisis de Ingresos
Análisis de Producción
Análisis de Ordenes de Compra
Análisis de Productos
Análisis de Inventarios
Análisis de Clientes
Análisis de Finanzas

Preguntas típicas de Análisis:

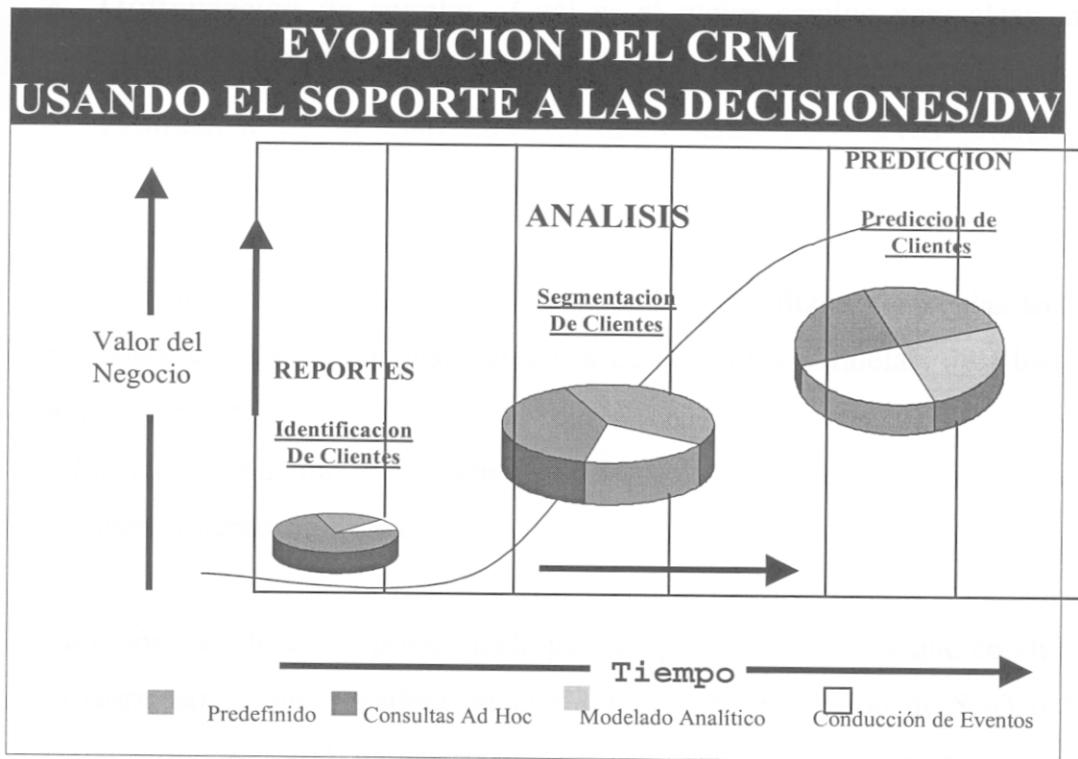
Porque el promedio de rentabilidad de los clientes esta bajo?
Porque la campaña no alcanzó las metas planeadas?
Porque las campañas no alcanzo las metas planeadas?
Porqué el porcentaje de respuesta es mas bajo que antes?
Porque la rentabilidad entre productos especificos varía mucho?
Porque se ha incrementado el costo de adquisición de este canal?

Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 10

5.2.3 Data Warehouse/Sistema de Soporte a la Decisión Tipo Tres: Pronosticando el Futuro

Un alto porcentaje de predicciones es una habilidad especializada que verdaderamente separa a las empresas líderes del resto. Quienes puedan anticipar tendencias y capitalizarlas, antes que se conviertan en un conocimiento común tendrán una obvia ventaja en el mercado



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 11

Una organización se mueve a través de una estructura definida en la Teoría de las Etapas, la más madura de las etapas proporciona la ruta que proporciona el más alto rendimiento. Algunas de las preguntas (con sus correspondientes aplicaciones predictivas) que la fase Tres puede ayudarnos a responder son:

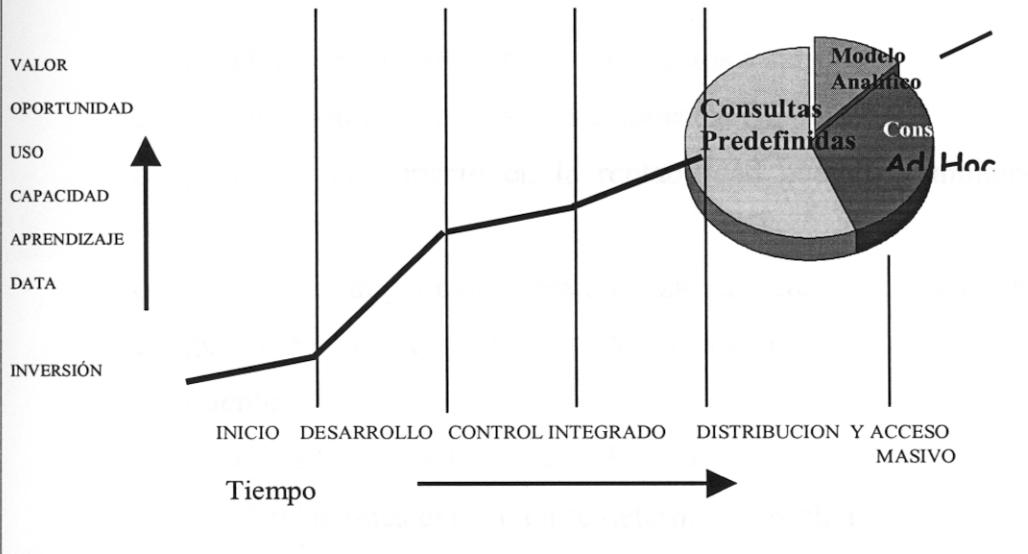
- **Aplicaciones de retención de clientes.**- ¿Qué clientes tienen alto riesgo de dejarnos?
- **Segmentación de Mercados.**- ¿Qué productos ó servicios comprarán nuestros clientes?
- **Optimización de canales.**-¿Cuál es el mejor camino para alcanzar un cliente?
- **Pronóstico de la demanda.**- ¿Cuáles serán los nuevos productos a vender?

Estas aplicaciones son ahora muy sofisticadas y utilizan avanzadas técnicas del soporte a las decisiones, funciones de consultas paralelas, data histórica masiva detallada, información funcional cruzada referida a los clientes, conocimiento del comportamiento finito, etc.

La empresa está ahora posicionada para responder preguntas que conducen a la respuesta: ¿Qué sucederá en el futuro?. Este es el tipo de SSD ó Data warehouse que está caracterizado por aplicaciones y preguntas como se muestra en las Figuras 12 y 13.

ETAPA DE QUE SUCEDERA Y CAPACIDAD DE PREDICCION

Etapa 5-6: Que Sucederá?

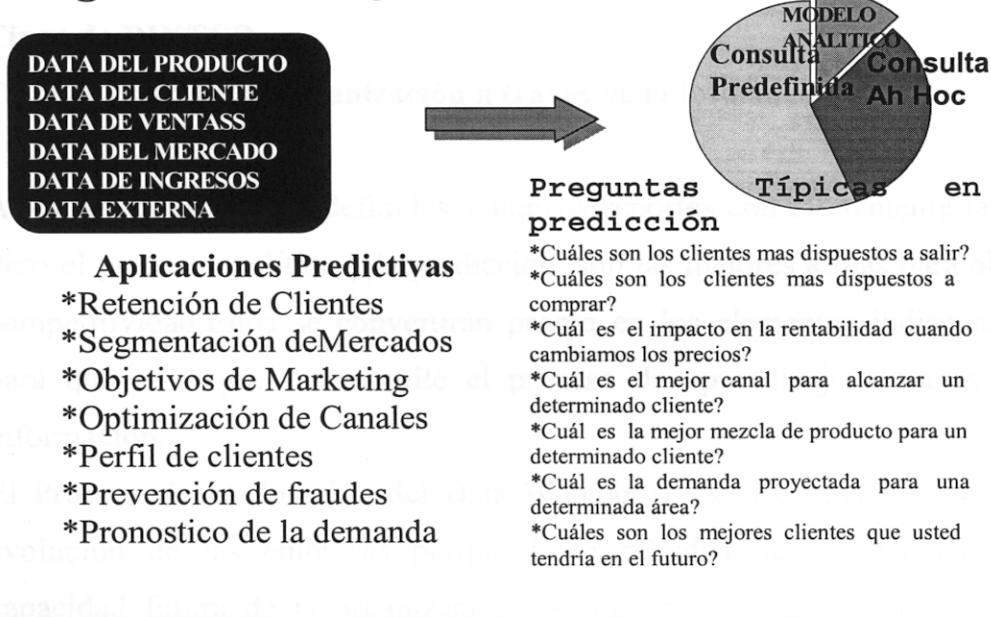


Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 12

APLICACIONES DE PREDICCION Y PREGUNTAS EN PREDICCION

Stage 5 & 6 - Que sucederá?



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 13

SSD: Predicción- Preguntas Clave

Tipo Tres SSD con ARC: “Pensando en los gustos de sus Clientes”

- ¿Cuáles son los clientes más dispuestos a salir?
- ¿Cuáles son los clientes más dispuestos a comprar?
- ¿Cuál es el impacto en la rentabilidad cuando cambiámos los precios?
- ¿Cuál es el mejor canal para alcanzar un determinado cliente?
- ¿Cuál es la mejor mezcla de productos para un determinado cliente?
- ¿Cuál es la demanda proyectada para una determinada área?
- ¿Qué relaciones existen entre determinados clientes?
- ¿Qué segmentos tienen un potencial de desarrollo/declinación?
- ¿Cuáles son los mejores clientes que Ud. quiere en el futuro?
- ¿Los clientes están gustosos de comprar nuevamente?
- ¿Si ellos están dispuestos a comprar, Qué comprarían?
- ¿Cuál es el impacto de nuevos productos para cada cliente?

Tipos de DW/SSD

Transformando la Organización a través de la información

Al inicio, consultas predefinidas y nuevos reportes con ciertamente la base. Pero el proceso analítico y la predicción, son las mejores armas para obtener competitividad. Ellos se convertirán pronto en los elementos indispensables para que la empresa desarrolle el proceso de aprendizaje a través de la información.

El Proceso de Maduración del Data Warehouse es el elemento clave en la evolución de las empresas porque los resultados de los negocios y la capacidad futura de la organización son mejorados, para obtener posición competitiva en el mercado a través de focalizarse en éstas tres fases de evolución:

| Tipos de usos de los Sistemas de Soporte a las decisiones | | |
|--|------------|---|
| Fase 1 | Reportes | Definición y captura de data de los clientes y construcción de una estructura del presente y actividad histórica en el datawarehouse. |
| Fase 2 | Análisis | Aplicación de la data para identificar patrones de actividad para clientes individuales y para mercados definidos. |
| Fase 3 | Predicción | Aplicación de los patrones definidos de la actividad actual de los negocios para ayudar a predecir el comportamiento de los clientes é identificar en el presente y el futuro los requerimientos de los clientes. |

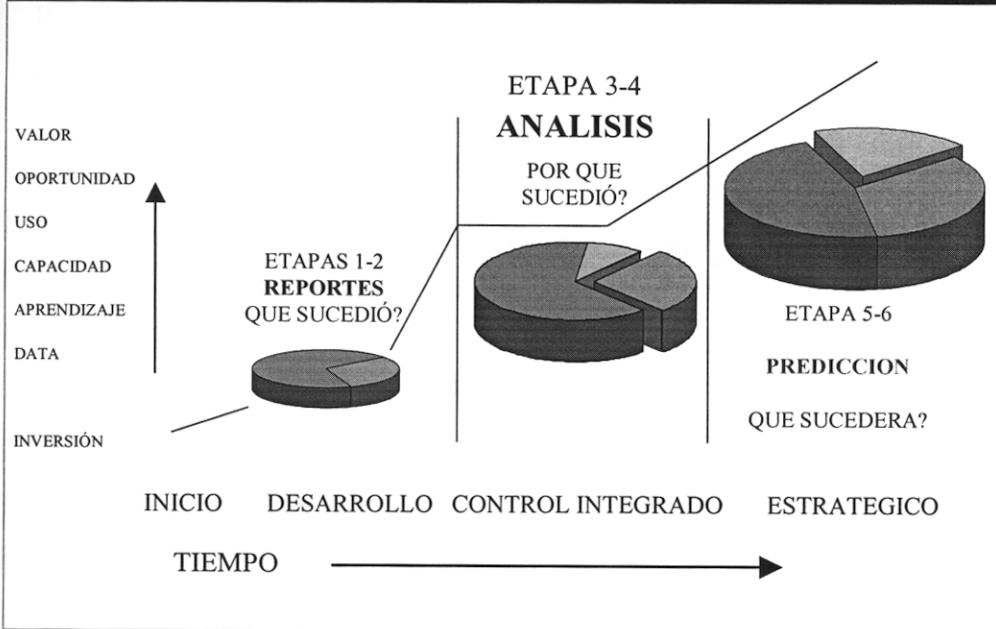
Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 14

Toda organización: privada ó gubernamental, debe conocer que para obtener un alto retorno de la inversión, debe hacerlo a través de un nuevo ambiente de aprendizaje: El Data Warehouse, en una Infoestructura para el nuevo Milenium. Una empresa que esté utilizando sucesivamente el ARC y la Infoestructura, la habilita para tener la triple visión, para conocer el pasado, analizar el presente y predecir el futuro.

Una visión general del cambio de rol del DW/SSD es mostrado en la figura 15: “Migración en una Infoestructura basada en el Conocimiento”. Note que existen dos etapas en cada Tipo de DW/SSD.

Migraciones en un Conocimiento Basado en la Infraestructura



Traducido de *Journal of Data Warehousing*
Vol. 5, # 4, 2000

Figura 15

La focalización de la información y la inversión en la Infoestructura debería ser en los clientes, no solamente en sus productos. La segunda focalización debe ser en finanzas, recursos, productos / servicios, canales, cadenas de distribución, inventarios, franquicias, proveedores, transportes, servicios, etc. Pero se encontrará fantásticos retornos de la inversión cuando el Data Warehouse se focalice en sus Clientes.

6. Data Warehousing vs. Data Mart

Por muchos años, los sistemas que extraían y almacenaban datos de diversas fuentes para que ayudaran a la toma de decisiones, se llamaron Data Warehouses. En fechas recientes se ha hecho una distinción entre los grandes sistemas para almacenar datos (data warehouses) y los sistemas más pequeños (data marts), aún cuando el concepto general sigue nombrándose Data Warehousing.

Información proveniente de la industria indica que aprox. el 75% del data warehouse actual es, de hecho, data mart. En la Conferencia Mundial de Data Warehousing de Meta Group/DCI 1997 en Febrero de 1997, se observó que “el objetivo de las empresas se ha desplazado de la, justificación del costo del data warehousing a la aplicación interna de la emisión de data marts”.

Los data marts se ajustan mejor a las necesidades que tiene una parte específica de un negocio, más que a las de toda una empresa. Optimizan la distribución de información útil para la toma de decisiones y se enfocan al manejo de datos resumidos ó de muestras, más que a la historia presentada en detalle. De igual forma, no necesitan ser administrados centralmente por el departamento de sistemas de una organización, sino que pueden estar a cargo de un grupo específico dentro del área de la empresa que los utilice.

La creciente popularidad de los data marts se basa en varias buenas razones. Por un lado, disminuyen significativamente el costo de creación y de operación, lo cuál los pone al alcance de muchas compañías. Con los data marts se pueden llegar a prototipos más rápidamente y obtener sistemas completamente desarrollados e implementados dentro de tres a seis meses. El problema es que tienen un alcance más limitado que el data warehousing, ya que se enfocan a un conjunto concreto de necesidades, por lo mismo, son ideales para trabajar con objetivos y equipos de trabajo precisos.

A menudo, las pequeñas empresas y los departamentos autónomos de una organización prefieren utilizar éstos para construir su propio mecanismo para toma de decisiones. Muchos departamentos de sistemas aprovechan la eficacia de esta aproximación y actualmente construyen un data warehouse para un solo tema ó un data mart cada vez que se necesita, lo que les permite ganar experiencia y el apoyo de los administradores, los cuales ven los beneficios constantemente. El hecho de comenzar con un plan modesto e ir creciendo conforme se aprende más sobre la fuente de datos y sobre las necesidades finales del usuario, permite que las organizaciones justifiquen el uso de los data marts conforme estos avanzan.

En ocasiones, los proyectos que comienzan como data warehouses evolucionan a data marts. Cuando las organizaciones acumulan grandes cantidades de datos históricos para el apoyo de las decisiones, que rara vez ó nunca usan, pueden reducir la información guardada y convertir su data warehouse en un data mart mejor enfocado. O bien, pueden dividir el data warehouse en diferentes data marts que ofrecen un tiempo de respuesta más rápido, un acceso más fácil y una menor complejidad para los usuarios finales.

Los abogados del data mart sugieren que pueden haber trayectorias alternas, mucho más fáciles para el éxito de un SSD, que construyendo un data warehousing. Una de éstas trayectorias es construir varios data marts y cuando se desarrollen y crezcan lo suficientemente grandes, se les denominará data warehousing. Los abogados del data mart proponen que el data mart es mucho más rápido y económico de construir que un data warehousing. Argumentan que cuando se construye un data mart no se requiere una gran disciplina de la organización y de ninguna preocupación por la arquitectura a largo plazo.

Desafortunadamente, la cosa no es tan sencilla, construyendo una arquitectura que consiste enteramente en data marts , probablemente se conduzca a la

organización a un nivel más grande. Definitivamente no habrá integración de datos y un SSD sin la integración es como un hombre sin un sistema esquelético.

Se ha descubierto que cuando se desarrolla SSD de la manera descrita líneas arriba, se tiene:

- Redundancia masiva de datos detallados e históricos a partir de un data mart a otro.
- Resultados contrarios e irreconciliables a partir de un data mart al siguiente.
- Una interfaz inmanejable entre los data marts y el ambiente de uso de la herencia.

Finalmente se descubrió que un SSD sin un data warehousing podría ser extremadamente insatisfactoria.

Un data mart y un data warehousing tienen esencialmente arquitecturas diferentes, aunque cuando están vistas de lejos y superficialmente, puedan parecer similares.

Los data warehousing son perceptiblemente diferentes de los datamarts. Los data warehousing se arreglan alrededor de temas corporativos encontrados en el modelo corporativo de los datos. El data warehousing es construido y poseído, por organizaciones centralmente coordinadas. El datawarehousing representa un esfuerzo verdaderamente corporativo.

El data mart se diseña para satisfacer las necesidades de un departamento. Muchos departamentos con objetivos muy diversos deben ser satisfechos. Esto es porque hay muchos data mart en la corporación, cada uno con sus propias características. El data warehousing se diseña para satisfacer las necesidades colectivas de la corporación entera. Un diseño dado puede ser óptimo para un solo departamento ó la corporación, pero para ambos. Los objetivos de diseño para la corporación son muy diferentes de los objetivos de diseño para un departamento dado.

A continuación detallamos algunas diferencias específicas:

- La granularidad de la data en un data mart es muy diferente a la granularidad de la data en un data warehouse. El data mart contiene data resumizada ó agregada. El data warehousing contiene la mayor data detallada que está en la organización. El data mart contiene mucha mayor granularidad que el data warehouse, no se puede fácilmente descomponer la granularidad del datamart en la granularidad para el data warehouse. Pero es más fácil ir en la dirección contraria, es decir sumarizar la data detallada.
- La estructura de la data en un data mart, comúnmente estrella, es ligeramente compatible con la estructura de la data en un data warehouse, estructura normalizada.
- La cantidad de data histórica encontrada en un data mart es muy diferente de la data histórica encontrada en el data warehouse. Los data warehouse contienen cantidades enormes de historia. Los data marts contienen solamente modestas cantidades de historia.
- Las áreas temas encontradas en un data mart son ligeramente relacionadas con los temas encontrados en un data warehousing.
- Las relaciones encontradas en el data mart no son las relaciones encontradas que se encuentran en un data warehousing.
- Los tipos de consulta que satisfacen un data mart son diferentes de las consultas encontradas en un data warehouse.
- La clase de usuarios que se encuentran en los data marts son diferentes al tipo de usuarios encontrados en los data warehouse.
- La clave de las estructuras encontradas en un data mart son significativamente diferentes de las claves de las estructuras encontradas en un data warehousing.

Porque Implementar primero un Datamart

No todas las organizaciones están preparadas para invertir inmediatamente en *grandes ambientes de warehouse integrados*. En muchos casos lo mejor puede ser comenzar con una *pequeña solución de datamart*. Soluciones de éste tipo

ofrecen algunas veces alto valor a los negocios que focalizarse inmediatamente en construir grandes soluciones técnicas orientadas a warehouse de la empresa.

Pero existen riesgos de construir exclusivamente estrategias de datamart:

- Estructuras independientes que contienen data diferente, resultan reportes desiguales
- Procesos de extracción redundantes que demandan recursos adicionales a los sistemas operacionales fuentes
- Requerimientos adicionales de soporte, para mantener el crecimiento de almacenes de datos y sistemas complejos.

En contraste algunas organizaciones han tratado de diseñar y construir grandes warehouse en lugar de crear una simple solución para la empresa. Proyectos de éste tipo algunas veces consumen vastos montos de tiempo, dinero y recursos y ofrecen decepcionantes beneficios para la empresa.

Los siguientes problemas son citados como las causas de falla de éste tipo de proyectos:

- Un alcance demasiado amplio y detallado, resulta en “parálisis de análisis”
- Carencia de un fuerte respaldo de la organización que inhabilita a resolver diferentes prioridades y necesidades en medio de las unidades de negocios
- Problemas de performance resultantes de grandes cantidades de data

Lo aprendido a la fecha es que lo mejor es comenzar con pequeños proyectos de datamart, pero antes de implementar el primer datamart, hay que crear un modelo de temas de la empresa que nos muestre los elementos de la data que atraviesan la organización y proporcionar un plan de cómo deben ser integrada para futuras interacciones. Este trabajo es creado a un alto nivel, permitiendo que en un mínimo de tiempo obtener la información crítica que nos permitirá desarrollar la arquitectura de información de la organización.

La data común de los datamart adicionales implementados son migrados al warehouse de la empresa.

Este tipo de soluciones tiene las siguientes ventajas:

- Minimiza dramáticamente la inversión en el warehouse de la empresa
- Elimina las extracciones redundantes de data y la administración de la data relacionada con los datamart “independientes”
- Asegura recursos centrales de TI que son usados para soportar la infraestructura del warehouse
- Construye componentes del warehouse de la empresa con sólo la data requerida que verdaderamente soporte los negocios.

7. Metodologías para Implementar Datawarehouse

7.1 Metodología para Implementar Datawarehouse desarrollada por Leveraged Solutions

La presente metodología fue desarrollada por la compañía Leveraged Solutions, Inc. (LSI). Ha desarrollado ésta metodología luego de años de refinamiento y testeo.

Plantilla metodológica:

Descubrir el negocio

- Identificar la métrica y problemas del negocio
 - ❖ Recolectar benchmarks de la industria
- Prepare el descubrimiento del negocio
 - ❖ Reúnase con sus Clientes y defina a las personas a entrevistar
 - ❖ Programe el inicio y las fechas de las entrevistas
 - ❖ Envíe las correspondientes identificaciones
- Conduzca las reuniones
- Conduzca las entrevistas y capture los resultados

- ❖ Descubra las áreas problemáticas
- ❖ Descubra objetivos
- ❖ Descubra iniciativas
- ❖ Recolecte métricas del negocio y ratios
- ❖ Compare los resultados con los benchmarks de la industria
- Priorice los problemas, objetivos e iniciativas
 - ❖ Revise los resultados de las entrevistas con los entrevistados para validar y corregir
 - ❖ Priorice con las prioridades actuales del negocio
- Conduzca workshop
 - ❖ Seleccione los problemas del negocio de las entrevistas y priorice los pasos a seguir
 - ❖ Obtenga consenso en precisar la definición de los problemas del negocio
 - ❖ Determine el impacto en los departamentos involucrados y en toda la compañía.
 - ❖ Determine la información necesaria para las prioridades determinadas
 - ❖ Identifique el valor de la resolución de los problemas en los departamentos involucrados y en toda la compañía
 - ❖ Priorice los problemas de acuerdo a las necesidades de la organización
 - ❖ Obtenga consenso para luego estudiar los items prioritarios y trabajar en los departamentos afectados
- Conduzca entrevistas detalladas en los departamentos afectados
 - ❖ Defina las preguntas del negocio y las decisiones para las prioridades establecidas
 - ❖ Defina la información que podría soportar éstas preguntas y decisiones
 - ❖ Defina la información histórica necesaria para direccionar éstas preguntas

- ❖ Identifique cuando las preguntas del negocio podrían ser contestadas
- ❖ Identifique las personas, grupos, funciones las cuáles podría iniciar éstas preguntas
- ❖ Identifique las acciones que podrían ser tomadas basadas en éstas preguntas
- ❖ Identifique el personal problema en la implementación de la solución
- ❖ Identifique mediciones operacionales, financieros y del cambio de conducta
- ❖ Obtenga consenso en las necesidades de información para contestar las preguntas
- Identifique los elementos de data requeridos con la organización de TI
 - ❖ Defina los elementos de data para contestar las preguntas claves del negocio
 - ❖ Determine la disponibilidad de la data desde TI
 - ❖ Determine la ubicación de la data desde TI
 - ❖ Investigue la calidad de la data
 - ❖ Investigue la disponibilidad de la data histórica
 - ❖ Investigue alternativas
 - ❖ Identifique procesos y decisiones de manera diferente debido a la falta de organización
 - ❖ Documente disponibilidad y calidad de la data, cree el modelo elemental de la data y el correspondiente mapeo
- Determine el alcance del sistema para conseguir las prioridades del negocio
 - ❖ Determine el tamaño de la base de datos, los requerimientos y costos de la tecnología
 - ❖ Determine el alcance de las aplicaciones y el costo de su desarrollo
 - ❖ Determine los requerimientos y costos de entrenamiento
 - ❖ Determine la gente problema

- Determine el TIR
 - ❖ Asigne valor a las preguntas del negocio para determinar cuáles métricas podrían ser afectadas
 - ❖ Asigne TIR a las mayores áreas direccionadas
 - ❖ Consiga consenso con los actores de las mayores áreas
- Presente resultados a los ejecutivos
 - ❖ Prepare resultados
 - ❖ Prepare presentación
 - ❖ Prepare presentación con los ejecutivos clave
 - ❖ Presente resultados a ejecutivos
 - ❖ Obtenga retroalimentación y modifique resultados
 - ❖ Prepare resultados finales

Diseño del Datawarehouse

- Revisión de la arquitectura
 - ❖ Determinar arquitectura de clientes existentes
 - ❖ Determinar arquitectura deseada para clientes y plan para conseguirlo
- Revisar arquitectura de la data de la empresa
 - ❖ Revisar arquitectura de la data existente en la empresa
 - ❖ Revisar modelos lógicos de data existentes
 - ❖ Determinar estrategia a crear en la arquitectura de la data de la empresa en el contexto del datawarehouse
- Revisar la Administración de la Data
 - ❖ Revisar arquitectura de la administración de la data existente
 - ❖ Inventario de los sistemas de producción relevantes de almacenamiento de data
 - ❖ Investigar calidad de la data, redundancia, sistemas propietarios

- ❖ Determinar arquitectura deseada de administración de la data
- ❖ Determinar estrategia para la administración de la base de datos
- ❖ Determinar estrategia para administración de la metadata
- ❖ Determinar estrategia para la transformación de la data
- ❖ Determinar estrategia para seguridad de la data
- ❖ Determinar estrategia para acceso a la data
- ❖ Determinar estrategia para la administración del cambio
- ❖ Determinar estrategia para planificar la capacidad
- ❖ Determinar el conjunto de habilidades necesarias para implementar y administrar la estrategia de administración de la data
- Revisión de las Aplicaciones
 - ❖ Revisar el ambiente de las aplicaciones existentes
 - ❖ Determinar el ambiente de aplicaciones deseado para acceder al datawarehouse
 - ❖ Determinar las aplicaciones y desarrollar el ambiente adecuado
 - ❖ Determinar el conjunto de habilidades necesarias a implementar y la administración de las aplicaciones
- Revisión de las comunicaciones
 - ❖ Revisar las comunicaciones existente y la estrategia de redes
 - ❖ Evaluar la tecnología de redes, capacidad y posibilidades de crecimiento
 - ❖ Evaluar proceso de soporte de redes
 - ❖ Determinar medios de integrar warehouse “limpios” dentro de arquitecturas de redes de clientes
- Diseñar componentes con tecnología de warehouse “limpios”

- ❖ Determinar requerimientos para dimensionar la base de datos
- ❖ Determinar los requerimientos del procesador del servidor de base de datos
- ❖ Determinar requerimientos de redes
- ❖ Determinar requerimientos de desarrollo de aplicaciones
- ❖ Determinar requerimientos de software de redes
- ❖ Determinar requerimientos de workstation
- Revisar requerimientos de entrenamiento
 - ❖ Determinar habilidades de soporte de operaciones TI
 - ❖ Determinar habilidades de redes en TI
 - ❖ Determinar habilidades para la administración de la data
 - ❖ Determinar habilidades para administración de la base de datos
 - ❖ Determinar habilidades para el desarrollo de aplicaciones
 - ❖ Determinar habilidades para utilizar tecnología de usuario final
 - ❖ Crear plan de educación para TI
 - ❖ Crear entrenamiento del usuario final y plan de soporte

Diseño Lógico de la Data

- ❖ Examinar las preguntas del sistema de soporte de decisiones del negocio y las correspondientes respuestas
- Determinar los elementos de data necesarios
- Determinar las áreas temas de la data comprometida
- Por cada área tema:
 - ❖ Determinar las entidades primarias
 - ❖ Determinar los identificadores comunes de éstas entidades
 - ❖ Describir los atributos de la data de éstos identificadores
 - ❖ Note discrepancias en los atributos de la data lo cuál puede ocurrir por redundancia de fuentes. Las discrepancias serán

resueltas durante el mapeado de la herencia de la data en el paso de la base de datos

- ❖ Determine las entidades menores las cuáles serán usadas para describir las entidades mayores
 - ❖ Determine los atributos de la data de éstos identificadores de las entidades menores observando la data existente
 - ❖ Determine las relaciones entre entidades
 - ❖ Describa las llaves externas entre entidades. Describa los atributos de la data observando la data existente
 - ❖ Dibuje un diagrama de entidad-relación
 - ❖ Defina atributos los cuáles describen las claves primarias de las entidades
 - ❖ Documente todas las entidades, relaciones, atributos y discrepancias
 - ❖ Revise el modelo lógico con los expertos en el tema
- Presente el modelo lógico de datos al personal ejecutivo de TI

Mapeo de la Data

- Para cada área tema:
 - ❖ Determine las fuentes potenciales primarias para las entidades
 - ❖ Determine los conflictos referidos a fuentes correctas ó la integridad entre sistemas competentes
 - ❖ Determine los archivos de data como fuente de las entidades
 - ❖ Determine las fuentes de archivo y de los elementos de data de las claves primarias para las entidades
 - ❖ Determine el formato y dominio de la fuente de data
 - ❖ Determine el formato y dominio entre la fuente datos y el modelo lógico
 - ❖ Determine la frecuencia de actualización de las entidades
 - ❖ Determine el índice de las claves primarias y las entidades

- ❖ Determine las claves foráneas de las fuentes y el formato y dominio entre la fuente y el modelo lógico
 - ❖ También determine el formato y dominio entre la fuente y la entidad propietaria de la clave foránea
 - ❖ Determine el contador de registros desde el sistema fuente y el contador de filas para la base de datos objetivo
 - ❖ Determine la calidad de la integridad referencial
 - ❖ Determine los atributos de las entidades para la fuente de archivos y elementos de data
 - ❖ Determine el formato y dominio de los atributos entre la fuente y el modelo lógico
 - ❖ Resuelva los conflictos existentes entre los elementos de data de la fuente con los usuarios y el administrador de la data
 - ❖ Resuelva los conflictos en los índices entre los usuarios y la administración de la data
 - ❖ Resuelva los conflictos de formatos y dominio en los usuarios y la administración de la data
 - ❖ Determine la criticidad de pérdida de elementos de data
 - ❖ Presente el mapeo de datos a los usuarios, administración de la data de TI y la parte ejecutiva de TI
- Diseño General del Proceso de Transformación
 - ❖ Identificar archivos fuente
 - ❖ Diseñar carga inicial de data
 - ❖ Diseñar carga cíclica de la data para administrar cambios, nueva data, etc., desde el sistema fuente
 - ❖ Determinar la conversión de formatos y dominios. Determinar la edición y validación necesaria para la limpieza de la data
 - ❖ Determinar el contador de registros desde el sistema fuente y el contador de filas dentro de la base de datos objetivo

- ❖ Determinar la mezcla de fuente de datos
- ❖ Determinar la división de fuente de datos
- ❖ Determine la frecuencia de extracción
- ❖ Determine la salida del proceso de conversión y el proceso de extracción el cuál será la entrada para el datawarehouse
- ❖ Diseñe el proceso de carga en las bases de datos objetivo
- ❖ Determine el proceso de validación para la carga en el datawarehouse
- ❖ Determine el proceso de validación en la Metadata cuando el datawarehouse sea modificado
- ❖ Revise el diseño con el Administrador de la Data y los propietarios de las aplicaciones de los sistemas fuente
- Diseñe en forma detallada el proceso de transformación para cada módulo

Diseño Físico de la Base de Datos

- Por cada área tema:
 - Análisis de requerimientos de consultas
 - ❖ Revisar las consultas del negocio y el modelo lógico de datos
 - ❖ Convertir las consultas del negocio a SQL contra el modelo lógico de datos
- Revisar la frecuencia de actualización de los requerimientos desde los requerimientos del negocio y el resultado del mapeado de la data
- Determinar las tablas y las claves primarias de las tablas
 - ❖ Determine el formato y dominio de las claves primarias
- Determinar las características de las tablas
 - ❖ Determinar las claves foráneas, formatos y opciones de integridad referencial
 - ❖ Determinar el formato y dominio de los atributos
 - ❖ Determinar el orden de las columnas

- ❖ Determinar el contador de filas y el tamaño de las tablas
- ❖ Determinar si la lógica de las tablas debería ser combinada para reducir las conexiones
- ❖ Determinar si la lógica de las tablas debería ser dividida para reducir el tamaño, la mejora de la carga y la mejora del performance
- ❖ Determinar los índices secundarios para las tablas. Determinar el tamaño de los índices secundarios
- Determinar la distribución del procesador para administrar la base de datos
 - ❖ Determinar como distribuir la base de datos, espacios en tablas y
 - ❖ tablas en los procesadores
 - ❖ Determinar como distribuir los índices secundarios en los procesadores
- Determinar las vistas necesarias para la base de datos
 - ❖ Determinar las vistas para los usuarios por su seguridad, por la necesidad de conocer sus propósitos y hacer más fácil y comprensible la definición de la data para los usuarios
 - ❖ Determinar las vistas para esconder conexiones y otras complejidades del SQL
 - ❖ Determinar el control de uso y el proceso de seguimiento
 - ❖ Documentar el diseño físico de la base de datos
 - Definir consideraciones de diseño
 - Crear el Lenguaje de Definición de Data (DDL)
 - ❖ Revisar el diseño con el Administrador de la Data, aplicaciones y el Administrador de la Base de Datos
 - ❖ Revisar el diseño con los usuarios
 - ❖ Completar la documentación con los cambios y revisiones del diseño

Diseño de las Aplicaciones

- Entender las necesidades del negocio y sus requerimientos
 - ❖ Determinar el tipo de información requerida por los usuarios del negocio
 - ❖ Entender la terminología de los usuarios, alcances y secuencias para responder sus preguntas
 - ❖ Entender la información que el usuario necesita para contestar sus preguntas
 - ❖ Determinar las próximas preguntas y acciones que el usuario desearía conocer luego de recibida las primeras respuestas

- Determinar el control de uso y el proceso de seguimiento
- Determinar los medios de mostrar la preferencia y resultados para los usuarios
 - ❖ Determinar los mejores medios de mostrar las respuestas a las preguntas
 - ❖ Determinar los mejores caminos para presentar opciones para iniciar consultas para los usuarios
 - ❖ Diseñar opciones de menú
- Diseñar opciones de dimensionamiento, objetos y funciones
- Diseñar el display de salida
 - ❖ Diseñar medios para iniciar consultas basadas en resultados de consultas previas
 - ❖ Diseñar requisitos de SQL de la base de datos
 - ❖ Diseñar el inicio de la ventana
 - ❖ Diseñar el logotipo de la ventana y sus funciones
 - ❖ Diseñar las funciones de facilidades para la ayuda al usuario
 - ❖ Diseñar el entrenamiento y la documentación

- ❖ Revisar el diseño con los usuarios, el Administrador de la Base de Datos y el Desarrollador de Aplicaciones
 - ❖ Crear un plan para construir y examinar la aplicación
- Entender las preguntas del negocio a ser resueltas con el reporte en batch y las consultas en batch
 - ❖ Determinar el tipo de información requerida por las preguntas del negocio
 - ❖ Determinar la frecuencia y distribución de los requerimientos del negocio
- Diseñar el sistema de reporte en batch
 - ❖ Diseñar las consultas ad hoc de los procesos de sumisión
 - ❖ Diseñar las consultas recurrentes de los procesos de sumisión
 - ❖ Diseñar el layout de los reportes
 - ❖ Diseñar los resultados de los reportes y el proceso de distribución
 - ❖ Diseñar los procesos de resolución de problemas
 - ❖ Diseñar el entrenamiento para los usuarios
 - ❖ Revisar los diseños con los usuarios, Administrador de la Base de Datos y Desarrollador de Aplicaciones
 - ❖ Crear un plan para la construcción

Desarrollo de la Solución

Transformación de la Data

- Preparación
 - ❖ Revisar detalladamente las especificaciones de diseño del paso Diseño de la Transformación
 - ❖ Crear un plan para su construcción
 - ❖ Crear un proceso que chequee la validación, comprometa a los usuarios
 - ❖ Obtenga acceso al ambiente de desarrollo

- Construir y examinar la carga inicial por módulo
 - ❖ Construir los módulos de transformación para la carga inicial
 - ❖ Examinar la unidad de los módulos
 - ❖ Crear y examinar carga de módulos a la base de datos física
- Examinar sistema de módulos
- Examinar la carga de data a la base de datos física
 - ❖ Correr la carga y transformación en pequeños subconjuntos de data
 - ❖ Validar el subconjunto de datos con los propietarios y usuarios de la data
- Cargar toda la data inicial de la base de datos física
 - ❖ Correr la carga y transformación en completos conjuntos de datos
 - ❖ Validar en forma completa el conjunto de datos con los propietarios y usuarios de la data
- Construir y examinar los módulos en cargas cíclicas, por módulo
 - ❖ Construir la extracción y condicionantes para los módulos
 - ❖ Examinar la unidad de los módulos
- Examinar la carga de data a la base de datos física
 - ❖ Correr la carga y transformación en pequeños subconjuntos de data
 - ❖ Validar el subconjunto de datos con los propietarios y usuarios de la data
- Cargar toda la data inicial de la base de datos física
 - ❖ Correr la carga y transformación en completos conjuntos de datos
 - ❖ Validar en forma completa el conjunto de datos con los propietarios y usuarios de la data
- Examinar el sistema de carga cíclica de los módulos

7.2

7.2.1

7.2.1.1

7.2.1.2

• Cargar las tablas de producción

- ❖ Validar la carga inicial
- ❖ Validar con la Metadata
- ❖ Resolver problemas de performance en el proceso de carga
- ❖ Resolver problemas de performance en las consultas

7.2.1.3

7.2.1.4

7.2.1.5

7.2.1.6

• Crear los objetos de la base de datos para propósitos de chequeo

- ❖ Validar la carga inicial
- ❖ Validar con la Metadata
- ❖ Resolver problemas de performance en el proceso de carga
- ❖ Resolver problemas de performance en las consultas
- ❖ Crear las tablas y espacios de los índices secundarios de la base de datos
- ❖ Crear los accesos y las tablas correctas
- ❖ Crear los índices secundarios y los accesos correctos
- ❖ Crear los usuarios y los accesos correctos
- ❖ Crear las vistas y los accesos correctos

Desarrollo de la Aplicación

Crear el Front End y la Aplicación GUI

- Entender las preguntas del negocio
 - ❖ Determinar el tipo de información requerida por las preguntas del negocio
 - ❖ Entender la terminología de los usuarios, alcances y secuencias para responder sus preguntas
 - ❖ Entender la información que el usuario necesita para contestar sus preguntas
 - ❖ Determinar las próximas preguntas y acciones que el usuario desearía conocer luego de recibida las primeras respuestas

7.2

7.2.1

7.2.1.1

7.2.1.2

7.2.1.3 Cargar las tablas de producción

7.2.1.3.1

7.2.1.3.2

7.2.1.3.3

7.2.1.3.4

7.2.1.4 Crear los objetos de la base de datos para propósitos de chequeo

7.2.1.4.1

7.2.1.4.2

7.2.1.4.3

7.2.1.4.4

7.2.1.4.5

7.2.1.4.6

- ❖ Validar la carga inicial
 - ❖ Validar con la Metadata
 - ❖ Resolver problemas de performance en el proceso de carga
 - ❖ Resolver problemas de performance en las consultas
- Cargar las tablas de producción
 - ❖ Validar la carga inicial
 - ❖ Validar con la Metadata
 - ❖ Resolver problemas de performance en el proceso de carga
 - ❖ Resolver problemas de performance en las consultas
- Crear los objetos de la base de datos para propósitos de chequeo
 - ❖ Crear las tablas y espacios de los índices secundarios de la base de datos
 - ❖ Crear los accesos y las tablas correctas
 - ❖ Crear los índices secundarios y los accesos correctos
 - ❖ Crear los usuarios y los accesos correctos
 - ❖ Crear las vistas y los accesos correctos

Desarrollo de la Aplicación

Crear el Front End y la Aplicación GUI

- Entender las preguntas del negocio
 - ❖ Determinar el tipo de información requerida por las preguntas del negocio
 - ❖ Entender la terminología de los usuarios, alcances y secuencias para responder sus preguntas
 - ❖ Entender la información que el usuario necesita para contestar sus preguntas
 - ❖ Determinar las próximas preguntas y acciones que el usuario desearía conocer luego de recibida las primeras respuestas

7.2 Metodología de Desarrollo SAS Rapid Warehousing

Características :

Desarrollo interactivo .- Las arquitecturas de Data Warehouse soportan típicamente procesos de toma de decisiones estratégicas y las consultas a la data son algunas veces exploratorias y de naturaleza ad hoc. Como resultado es difícil determinar al inicio un conjunto completo de requerimientos que determinen las características del warehouse.

En adición la extracción de la data de los sistemas operacionales pueden tener discrepancias a través de los registros y sistemas. Por lo tanto es esencial que la metodología de datawarehousing sea flexible y se acomode a los cambios inesperados.

Con la aproximación interactiva de ésta metodología, las actividades del proyecto pueden ser conducidas con una tendencia cíclica, permitiendo a los administradores del proyecto programar las actividades en paralelo en todas las fases. Esta metodología facilita la implementación de grandes proyectos de datawarehousing en pequeños interactivos ó evolucionarios construcciones que permitirán una gradual implementación y posibilitan una respuesta inmediata a las demandas del negocio.

Arquitectura escalable.-

Esta metodología identifica los componentes esenciales que afectan la escalabilidad de una solución de warehousing. Proporciona una guía para planificar y construir soluciones de warehousing escalables. Estos componentes son identificados luego de implementaciones exitosas en diversas organizaciones en el mundo y fueron desarrolladas en respuesta a las diferentes necesidades de negocio.

En un proyecto escalable los componentes pueden ser omitidos ó el nivel de esfuerzo reducido.

Los componentes son los siguientes:

Sistemas operacionales: Los sistemas operacionales capturan las transacciones del día a día. En contraste con el data warehouse, éstos no retienen información por más de 90 días. La data es capturada para su uso por algún sistema específico y optimizada para consultas de registros individuales, por ejemplo una orden antes que una tabla. Estos proporcionan la fuente de data esencial para que trabaje el warehouse.

Data externa: Las fuentes de data externa, como el análisis del mercado y sistemas de abastecimiento, proporcionan una fuente de información esencial, la cuál puede ser incorporada al warehouse.

Almacén de datos operacionales: Un almacén de datos operacionales es un conjunto de datos integrados y orientados a un tema de los sistemas operacionales. Estos registros son actualizados cuando los sistemas operacionales son actualizados.

Data Warehouse: Es un almacén de datos que almacena información en un formato disponible para la inteligencia de negocios y los sistemas de soporte a las decisiones.

Data Warehouse de la empresa: El data warehouse de la empresa es la central de información de data común a toda la organización. Es una central de información diferente al sistema de almacenamiento del sistema operacional de la empresa.

Datamarts y Data Warehouse Departamentales: Los datawarehouses departamentales ó datamarts contienen información específica de una unidad de negocios ó departamento.

Los datamarts pueden ser contruidos basados en la data del data warehouse de la empresa, esto es, son dependientes de la data en el data warehouse, ó pueden ser creados separadamente y en forma independiente de ésta fuente de

datos. Los datamarts independientes no son recomendados porque pueden crear islas de información que el data warehouse puede evitar.

Datamarts Personales: Estos son conjuntos de datos que son extraídos de departamentos ó unidades de negocio y reúnen los requerimientos de usuarios individuales ó pequeños grupos de usuarios. Estos datamarts personales son actualizados cuando la información del warehouse es actualizada.

Aplicaciones de extracciones específicas: Contienen un subconjunto de datos que han sido extraídos de un departamento ó unidad de negocios denominado datamart. La extracción de aplicaciones específica no requiere de herramientas especiales, éstos son usados para proporcionar el input de otros procesos de negocios.

Portal de Web: Los portales inteligentes permiten la explotación del data warehouse a través de una simple interfase en la Web. Esta interfase proporciona una visión personalizada del data warehouse lo cuál permite a los individuos de la organización recibir la información que ellos quieren cuando quieran. Esta información para cada persona es centralmente controlada para asegurar que la data apropiada sea vista. Este mecanismo de portal permite que el personal comparta información sobre un problema del negocio ó sobre una oportunidad encontrada por una persona determinada y que rápidamente pueda ser accesada por otras interesadas por temas similares. Esta rapidez facilita el proceso de toma de decisiones.

Metadata : Es la información referida a la data en el warehouse, éstos procesos soportan al Warehouse, los negocios y la organización. Es usada para documentar el contenido del data warehouse, administrar los procesos y cambios, establecer seguridad y control y promover los estándares.

Reducción de
clasificación

Datamarts independientes: No todas las organizaciones están preparadas para invertir inmediatamente en grandes warehouse. En éstos casos ésta metodología puede ser usada para comenzar con una solución de un pequeño datamart. Estas soluciones pueden tener los siguientes riesgos:

- Estructuras independientes que contienen data diferente, resultan reportes diferentes
- Procesos de extracción redundantes que demandan recursos adicionales en sistema operacionales ya estresados
- Requerimientos de soporte adicionales para mantener sistemas de almacenamiento complejos

Focalización en el negocio.-

Dos características de la Metodología de Desarrollo SAS Rapid Warehousing es que se focaliza en los requerimientos del negocio y demanda la participación activa de toda la comunidad del negocio. Uno de los principales motivos de fallas de éste tipo de proyectos es que al usuario no se le involucra. Con éste tipo de metodología los usuarios del negocio son parte integral del equipo. Ellos contribuyen a los siguientes roles y funciones:

- Promover el proyecto
- Definir los requerimientos del negocio
- Revisar y validar
- Retroalimentación
- Participación en Workshops

Efectivo uso del Workshop.-

Un número de técnicas son usadas a través del ciclo de vida del proyecto para recolectar los requerimientos de alto nivel del negocio, entrevistas de feedback, surveys, prototipos, modelado, etc. Basado en las experiencias en la planificación y administración de proyectos de warehousing a través de todo

tipo de industrias, SAS ha encontrado que los workshops son la manera más efectiva para definir todo tipo de requerimientos.

En las fases tempranas del proyecto, los workshops establecen los elementos de alto nivel como los objetivos, alcances y prioridades. Esto es particularmente efectivo porque en ésta fase participan usuarios de diferentes áreas del negocio con diferentes prioridades para el warehouse. Los workshops también acortan el tiempo necesario para completar el proyecto porque reducen el número de interacciones para recolectar información durante las entrevistas.

Enfasis en la Calidad.-

La calidad a través de la administración efectiva es parte integral del éxito del data warehouse. Esta metodología establece todas las responsabilidades de todos los roles del proyecto. Para el administrador del proyecto, la metodología proporciona guías y procedimientos para cada paso. Cada fase de ésta metodología tiene definida claramente las actividades de planificación e implementación que aseguren la total calidad del warehouse y la explotación de las aplicaciones.

Los usuarios del negocio revisan y aprueban regularmente a través del desarrollo del proyecto y el administrador de la data es equipado con un protocolo para asegurar la integridad de la data cuando el warehouse es diseñado, construido y mantenido.

Responsabilidades del proyecto claramente definidas.-

Un importante aspecto de ésta metodología desarrolla los roles del proyecto para asegurar que las tareas sean realizadas para cumplir los objetivos del proyecto. Un rol del proyecto es un conjunto de especificaciones y actividades que requieren de ciertos conocimientos y habilidades.

Los roles del proyecto puede incluir:

- Ejecutivo auspiciador

tipo de industrias, SAS ha encontrado que los workshops son la manera más efectiva para definir todo tipo de requerimientos.

En las fases tempranas del proyecto, los workshops establecen los elementos de alto nivel como los objetivos, alcances y prioridades. Esto es particularmente efectivo porque en ésta fase participan usuarios de diferentes áreas del negocio con diferentes prioridades para el warehouse. Los workshops también acortan el tiempo necesario para completar el proyecto porque reducen el número de interacciones para recolectar información durante las entrevistas.

Enfasis en la Calidad.-

La calidad a través de la administración efectiva es parte integral del éxito del data warehouse. Esta metodología establece todas las responsabilidades de todos los roles del proyecto. Para el administrador del proyecto, la metodología proporciona guías y procedimientos para cada paso. Cada fase de ésta metodología tiene definida claramente las actividades de planificación e implementación que aseguren la total calidad del warehouse y la explotación de las aplicaciones.

Los usuarios del negocio revisan y aprueban regularmente a través del desarrollo del proyecto y el administrador de la data es equipado con un protocolo para asegurar la integridad de la data cuando el warehouse es diseñado, construido y mantenido.

Responsabilidades del proyecto claramente definidas.-

Un importante aspecto de ésta metodología desarrolla los roles del proyecto para asegurar que las tareas sean realizadas para cumplir los objetivos del proyecto. Un rol del proyecto es un conjunto de especificaciones y actividades que requieren de ciertos conocimientos y habilidades.

Los roles del proyecto puede incluir:

- Ejecutivo auspiciador

sistemas pueden ser usados para recopilar los hechos necesarios. El documento resultante donde se encuentran todos los requerimientos es revisado por todas las partes afectadas. Este documento incluye la identificación de los objetivos del negocio como el análisis de factibilidad técnica.

Diseño.-

Incluye las siguientes fases:

- Diseño lógico y físico en forma detallada del modelo de la data
- Especificación detallada del modelo del proceso de extracción, transformación y carga.
- Creación del modelo de aplicaciones ó selección de herramientas de explotación
- Diseño de aspectos adicionales como el de seguridad y modelos de metadata

El documento de requerimientos es usado como input para ésta fase, la cuál produce un documento de diseño detallado para la parte seleccionada.

Construcción.-

Durante ésta fase, el equipo de implementación codifica y puebla el warehouse con la data y desarrolla las aplicaciones de análisis y reportes para los usuarios. Los usuarios del negocio y el equipo de TI examina rigurosamente el warehouse y las aplicaciones para verificar que todos los criterios de aceptación son reunidos.

Examen final.-

Durante ésta fase final, un equipo de aseguramiento de la calidad independiente examina el sistema antes de entregarlo al cliente. El equipo revisa el performance del sistema y finaliza con la preparación de un test formal de aceptación. El test del sistema asegura que el sistema reúna los requerimientos funcionales. La preparación del test de aceptación involucra la verificación de todos casos de test específicos.

Despliegue.-

En ésta fase se asegura que los usuarios han sido entrenados para acceder rápidamente a la data y en el uso de las aplicaciones. De ésta manera se ayuda a promover la más amplia aceptación en todo el proyecto. Mientras más rápido, los usuarios puedan obtener beneficio del proyecto de warehouse, apoyarán con mayor esfuerzo al posterior desarrollo y se comprometerán mucho más.

Revisión.-

Tres acciones son ejecutadas en ésta fase:

- Seguir la fase desde Valoración hasta Construcción durante el proceso de implementación y aprender de los éxitos.
- Luego de 3 a 6 semanas revisar la fase de Despliegue y asegurar que los usuarios han sido correctamente soportados.
- Luego de 18 a 24 meses después de la construcción inicial medir los beneficios tangibles. Calcular el TIR y asegurarse que el warehouse está continuamente reuniendo los requerimientos del negocio.

7.3 Metodología para desarrollar un Data Mart

Por Larissa Moss y Sid Adelman

Luego de analizar diversas metodologías, creemos que la que más aconsejable es utilizar la metodología RAD con la siguientes modificaciones:

- Establecer tres grandes procesos a desarrollar, a los cuáles se le aplicará la metodología RAD, en forma simultánea.
- Antes de comenzar el punto anterior se deberá preparar un documento denominado: “Acuerdos del Proyecto”, donde se encontrarán los alcances, roles y responsabilidades, puntos de control, etc. concordados entre los auspiciadores del proyecto, el personal de TI y los usuarios. Este documento es la base y referencia para poder iniciar los loops simultáneos del punto anterior.

PASO 1 :

Acuerdos del Proyecto:

El objetivo principal de éste paso es producir un documento, el cuál es un mapa de ruta detallado para el proyecto y contiene las siguientes secciones:

Objetivos y Metas del Data Mart:

Es importante que los objetivos del Data Mart estén alineados con los objetivos y metas estratégicas de la organización. Ambos, los objetivos y metas del Data Mart y los objetivos y metas estratégicas deberán ser registrados en ésta sección.

Requerimientos específicos del negocio y su solución:

El Datamart es un complejo sistema de soporte a las decisiones que trabaja con diferentes tipos de bases de datos diseñados para resolver diferentes tipos de problemas en el negocio. En ésta sección se debería registrar la naturaleza de éstos requerimientos y categorizarlos en requerimientos de reportes operacionales ó requerimientos de análisis de negocios tácticos ó estratégicos. Un diagrama de alto nivel debería representar este sistema, mostrando el número de base de datos y su tipo, el potencial de files fuente y la propuesta de flujo para el ETL.

Análisis Costo / beneficio:

Se debe justificar económicamente la implementación del proyecto.

Alcance funcional de alto nivel:

Una vez que los problemas del negocio y las propuestas de alternativas de solución han sido mencionadas, es importante considerar el alcance del proyecto.

Dependerá de los siguientes factores:

Cuanta data deberá ser capturada desde los archivos fuentes

Cuanto esfuerzo habrá en el proceso de limpieza y transformación

Cuantos archivos fuente deberá ser integrados al proyecto deben haber

Disponibilidad de archivos fuentes externos analizar

Cuanta nueva tecnología deberá ser evaluada, instalada y aprendida

Conjunto de habilidades actuales de los equipos

Presupuesto anticipado y fecha de desembolso

Modelo de Data de alto nivel:

Introducción:

Una excelente técnica para determinar el alcance del proyecto es el modelado lógico de la data. Un modelo conceptual mostrará las áreas temas bajo consideración para:

Identificar los elementos de data que deben ser contemplados en el análisis de la data.

Capturar las reglas de negocio.

Mostrar las dependencias de la data y sus puntos de integración.

Coleccionar los metadata del negocio.

Infraestructura :

Data:

Existen dos tipos de infraestructura: la infraestructura tecnológica y la infraestructura de desarrollo. La infraestructura tecnológica puede incluir en el proyecto: nuevo procesador, un nuevo sistema de administración de la base de datos, nuevas herramientas de desarrollo, y comprar ó desarrollar el repositorio de metadatos. La infraestructura de Desarrollo del proyecto puede incluir: una metodología comprada ó desarrollada, nuevos estándares, nuevos roles y responsabilidades, etc.

Número de archivos fuentes:

Introducción: **Identificación de los archivos fuentes:**

Los usuarios y los miembros del equipo del proyecto deben determinar cuantos y cuáles serían los archivos fuentes a analizar.

Reportes de errores de Consulta:

Calidad de la data en los archivos fuente:

Introducción: **Identificación de nuevas excepciones:**

Una vez que han sido identificados los archivos fuentes, es importante observar los valores de la data que han sido identificados en el modelo lógico de datos y encontrar errores en la data, excepciones en las reglas de negocio, redefinir campos, etc. El número de excepciones encontradas y su severidad,

podrían indicarnos si es necesario un esfuerzo para la limpieza y transformación de la data.

Responsabilidad de los usuarios:

Los usuarios deben aprender sus nuevos roles y responsabilidades. El documento de Acuerdos del Proyecto es un excelente lugar para capturar dicha información.

Requerimientos de Disponibilidad:

Aquí se debe determinar cuantas horas al día, a la semana y al mes, el Datamart dará servicio a los usuarios.

Requerimientos de Seguridad:

Es erróneo asumir porque el Datamart es un sistema que de soporte a las decisiones, no es necesario trabajar requerimientos de seguridad. Esta presunción es peligrosa, ya que la data puede incluir trabajos con entorno de la Web.

Por ejemplo el detalle de algunos accesos podría ser definido. Algunos usuarios podrían tener acceso a alguna data ya sumariada, mientras que otros necesitan ver el detalle de ésta suma.

Reportes/Archivos de Consulta:

Muchos Datamart pueden proporcionar reportes y consultas pre-escritos y parametrizados. Estos necesitan ser almacenados y administrados sus versiones en una librería de consultas. Muchas herramientas de consulta proporcionan éstas librerías. El documento de Acuerdos del Proyecto debería

PASO 2:

Desarrollo paralelo de tres procesos:

Teniendo como base y referencia el documento “Acuerdos del Proyecto “, se procede a desarrollar la metodología RAD en tres grandes procesos:

Proceso Back End ó de Desarrollo de Base de Datos y Extracción, Transformación y Carga:

El documento de requerimientos del negocio del Acuerdos del Proyecto, debe ser analizado y de ser necesario actualizado. El modelo lógico de datos debe estar completamente terminado en éste paso. Los archivos fuente deben ser rigurosamente analizados para cada atributo definido. Durante éste paso las especificaciones para la programación de la transformación de la data de los archivos fuente deben concluirse ya sea con la ayuda de herramientas ó programas ETL ó código nativo.

Tan pronto el modelo lógico de la data ha sido terminado y las especificaciones de la programación ETL ha sido desarrolladas, los patrones de acceso son identificados del documento de requerimientos del negocio y la base de datos es diseñada.

Crea las bases de datos del Datamart.

El proceso de ETL es el más complicado de todos los procesos del Datamart. Este proceso puede consumir entre un 70% al 80% del tiempo de desarrollo.

El proceso ETL consiste en la extracción, ordenamiento, integración, limpieza y transformación de las fuentes de datos. También puede incluir: crear el metadata y los respectivos aplicativos.

Crea las funciones de ayuda para que los usuarios naveguen a través de las bases de datos del Datamart

Proceso Front End ó de Desarrollo de Consultas y Reportes:

El equipo del front end diseña y prototipea reportes y consultas. Usando el documento de los requerimientos del negocio de Acuerdos del Proyecto y los requerimientos de acceso de data que han sido identificados durante el proceso paralelo de Desarrollo de Base de Datos y ETL, un completo conjunto de especificaciones es desarrollado para despachar la data. Este conjunto incluye layout de reportes, ilustraciones de consultas, ejemplo de consultas ad hoc, displays de pantallas, etc.

Durante el desarrollo de éste paso, otro conjunto de actividades es realizado, como: desarrollo de help desk, producir manuales de entrenamiento, etc.

Este equipo se concentra en la presentación de la data al usuario final a través de una variedad de herramientas. Deben entender el ambiente cliente / servidor y todas las herramientas de reporte, OLAP, herramientas de consulta, herramientas WEB. Entender el diseño de bases de datos multidimensionales.

Proceso de Establecimiento de la Plataforma Tecnológica:

El propósito de éste documento es revisar y analizar la propuesta de los componentes de la plataforma tecnológica identificada durante el paso de Acuerdos del Proyecto, y determinar la plataforma tecnológica apropiada

Para el datawarehouse, lo cuál puede incluir:

- Unidad Central de Procesamiento
- Sistema de Administración de la base de datos
- Herramientas de extracción, transformación y carga
- Herramientas de limpieza
- Herramientas de Reportes y consultas
- Herramientas de acceso a la Web
- CASE
- Repositorios

Aquí está el equipo responsable por instalar y examinar todos los componentes de la infraestructura técnica, como: redes, hardware, software, herramientas, utilitarios, etc.

Notas: Este equipo debe estar conformado por personal con experiencia en el área de infraestructura técnica, con conocimientos en redes, hardware, software, herramientas, utilitarios, etc. Este equipo debe estar conformado por personal con experiencia en el área de infraestructura técnica, con conocimientos en redes, hardware, software, herramientas, utilitarios, etc.

El equipo de infraestructura técnica debe estar conformado por personal con experiencia en el área de infraestructura técnica, con conocimientos en redes, hardware, software, herramientas, utilitarios, etc.

Este equipo debe estar conformado por personal con experiencia en el área de infraestructura técnica, con conocimientos en redes, hardware, software, herramientas, utilitarios, etc.

Procedimientos: Este equipo debe estar conformado por personal con experiencia en el área de infraestructura técnica, con conocimientos en redes, hardware, software, herramientas, utilitarios, etc.

2.1 Diagnóstico de la Situación Actual del Proyecto
Este parte nos permitirá conocer la organización a la cual se le aplica la metodología propuesta, es decir conocer su posición competitiva, sus fortalezas y debilidades, sus oportunidades y amenazas, también su misión, sus visiones estratégicas y sus objetivos empresariales.

CAPITULO II : PROPUESTA DE METODOLOGIA DE DESARROLLO

Nuestra propuesta de metodología de desarrollo esta basado en dos premisas importantes:

a) Cualquier sistema de apoyo a las decisiones que queramos implementar en una organización debe ser una opción estratégica importante si es que deseamos que dicho sistema coadyuve a que la organización obtenga ventaja competitiva.

b) El tipo de sistema de apoyo a las decisiones a implementar en una organización depende del desarrollo en tecnología de información alcanzado por la misma.

Nuestro planteamiento reside en plantear para ambas premisas una metodología de análisis que permita obtener información relevante de ambos aspectos.

Presentamos a continuación las etapas de la metodología propuesta:

2.1 Diagnóstico de la Situación Actual de la Organización

Esta parte nos permitirá conocer la organización a la cuál se le aplicará la metodología propuesta, es decir conocer su posición competitiva, sus fortalezas y debilidades, sus oportunidades y amenazas, también su visión, misión, sus opciones estratégicas y sus objetivos empresariales:

Descripción de la Empresa

Planeamiento Estratégico

Visión y Misión de la Empresa

Análisis de las Cinco Fuerzas del Sector

* Análisis del entorno

Análisis de la Cadena de Valor

* Análisis interno

La conclusión importante de ésta parte es determinar si la implementación del Datamart Comercial contribuirá a que CIDELSA obtenga ventaja competitiva ó que simplemente ayude a la toma de decisiones a nivel gerencial.

2.2 Diagnóstico de la infraestructura de Tecnología de Información de la Organización

Este diagnóstico nos permitirá conocer el estado actual de la organización en lo referido a su madurez en tecnologías de información, es decir el hardware que utiliza, los programas de sus sistemas de producción, la preparación y cantidad del personal de Ti, etc.

Con ésta información podremos establecer los alcances y límites de la implementación:

Infraestructura de Hardware y Redes

Infraestructura de Software

Infraestructura de personal de TI

Aplicación de la Teoría de las Etapas

Tipo de Sistema de Soporte a las Decisiones

Todas las organizaciones tienen diferente grado de madurez en los aspectos relacionados a tecnología de información por lo tanto el tipo de Datamart a implementar estará en función de dichas características peculiares de cada empresa. La aplicación de la Teoría de las Etapas nos ayudará a entender dicha situación, es decir a entender el presente y posicionarnos tecnológicamente en el futuro. También nos ayudará a determinar el tipo de

Datamart a implementar, para aprender el uso de dicha tecnología, en la visión general de la inversión y en entender los cambios con la introducción y puesta en marcha de éste tipo de tecnologías.

2.3 Desarrollo de Datamart

Luego de realizar la evaluación de las metodologías presentadas en el Marco Teórico Referencial, mediante el cuadro 2, decidimos optar por la metodología desarrollada por Larissa Moss y Sid Adelman, pero con la atinencia que la parte de Requerimientos Específicos del Negocio del Paso 1, del Acuerdo del Proyecto, deberá finalizar con un modelo de datos de alto nivel de la empresa.

En la figura 16 vemos las actividades principales relacionadas a ésta metodología así mismo podemos observar en la figura 17 una de sus principales virtudes; el desarrollo en paralelo de los tres procesos a realizar: back end, front end y establecimiento de plataforma tecnológica y la aplicación del método RAD para cada una de ellas.

EVALUACIÓN DE METODOLOGÍAS

| METODOLOGÍA FACTOR DE EVALUACIÓN | LEVERAGED SOLUTIONS | SAS RAPID | LARISSA MOSS SID ADELMAN |
|---|------------------------|--------------|--------------------------------|
| ORIENTACIÓN A RESULTADOS | 6 | 8 | 10 |
| PROCESO FLEXIBLE | 8 | 10 | 10 |
| IDENTIFICACIÓN DE RESPONSABILIDAD | 9 | 9 | 9 |
| ADAPTABLE A ORGANIZACIÓN | 8 | 10 | 10 |
| MEDIBLE | 8 | 8 | 10 |
| ESCALABLE | 8 | 8 | 10 |
| FOCALIZADA EN EL NEGOCIO | 8 | 10 | 10 |
| FAMILIARIDAD DE TÉCNICAS | 8 | 10 | 10 |
| COSTO DE IMPLEMENTACIÓN | 6 | 6 | 8 |
| PUNTAJE TOTAL | 69 | 78 | 87 |

CUADRO 2

Orientación a resultados.-

En este factor lo que medimos es el nivel en que la metodología permite obtener resultados medibles en diferentes etapas de la implementación.

Proceso flexible.-

Aquí lo que queremos registrar es la capacidad de la metodología para permitir que durante la implementación poder regresar y corregir etapas anteriores y nuevamente continuar.

Identificación de la responsabilidad.-

Es importante que una metodología permita registrar para las diferentes etapas los responsables de las mismas, es decir que contemple la designación de equipos de trabajo y asignación de responsabilidades.

Adaptable a Organización.-

Cada organización tiene peculiaridades en cuanto a su desarrollo organizacional, nivel de TI, etc. Este factor mide como la metodología se adapta a nuestra empresa.

Medible.-

Todo proyecto de ésta envergadura deber ser susceptible de ser registrado, medido y controlado en sus diferentes etapas para poder tomar las medidas correctivas si fuera necesario. Este factor mide dicha capacidad.

Escalable.-

Toda organización crece con el tiempo y por ende su nivel de archivos, su desarrollo de aplicaciones y el requerimiento de recursos humanos. La metodología seleccionada debe contemplar dichos factores durante su implementación.

Focalizada en el negocio.-

Muchas metodologías se pierden en conceptos ajenos al tema principal de la organización. Este factor mide la capacidad de la metodología para no perder durante la implementación, a obtener resultados en benéfico directo del negocio.

Familiaridad de técnicas.-

Tratar de aprovechar el conocimiento desarrollado por otras organizaciones vecinas y el desarrollo de nuestros recursos humanos y de nuestros consultores confrontándolo con las exigencias técnicas de la metodología, nos permitirá conocer su grado de accesibilidad.

Costo de implementación.-

En este factor se mide el costo de implementar dicha metodología. A un mayor puntaje corresponde un costo menor.

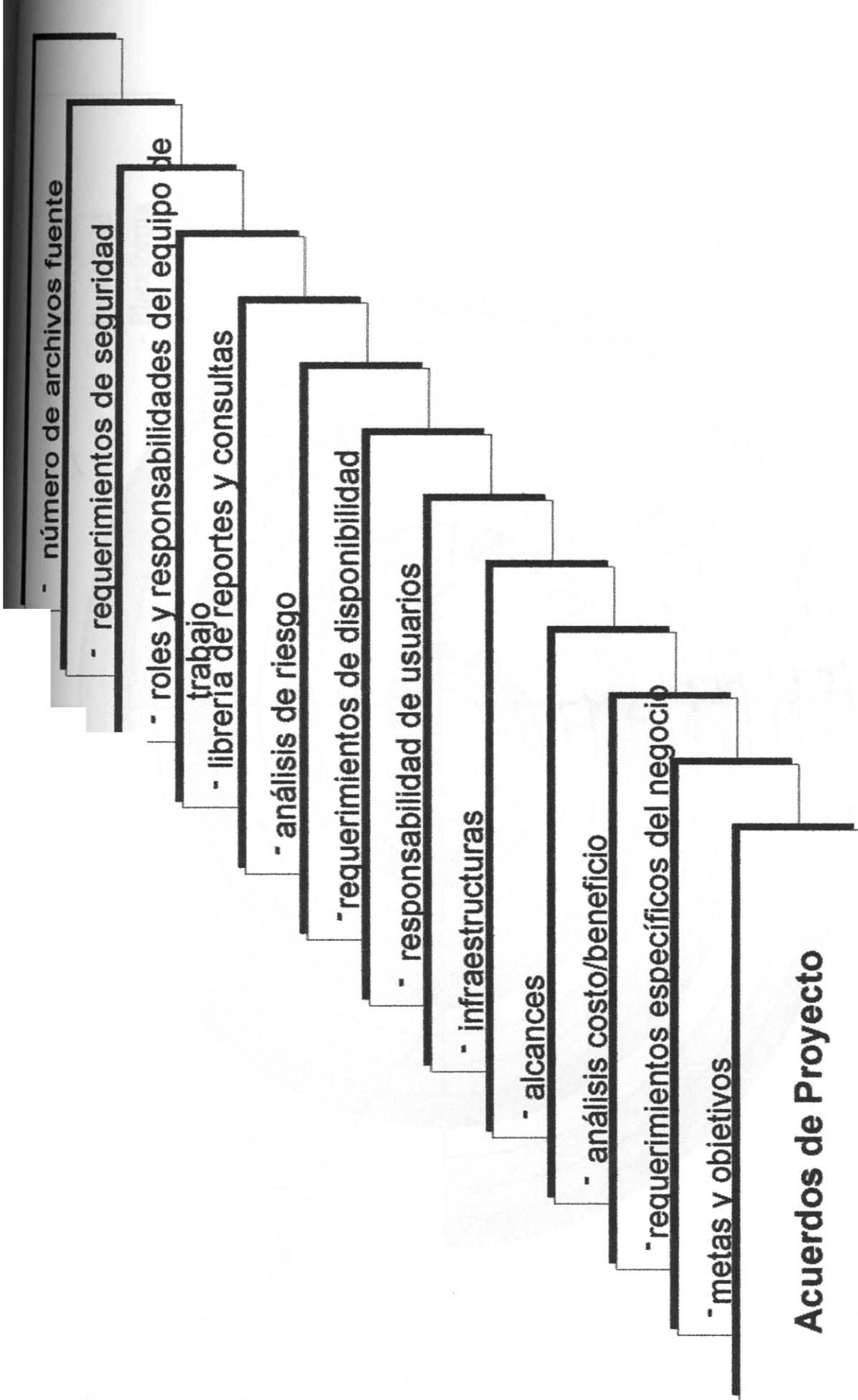


Figura 16

CAPITULO III: DESARROLLO DE PROTOTIPO

El Prototipo se basa en la visión y la misión de la empresa.(Ver pags. 98 y 99)

3.1 DIAGNOSTICO DE LA SITUACIÓN ACTUAL DE LA ORGANIZACION

3.1.1 Descripción de la empresa

CIDELSA (Comercial Industrial Delta S. A.) es una empresa industrial de tipo familiar, con más de 30 años de experiencia en el sector industrial de confecciones plásticas, constituida bajo la modalidad de sociedad anónima, mediante el aporte de capital de accionistas peruanos.

3.1.2 Actividad económica

Cidelsa inició sus operaciones comercializando artículos para pesca tales como redes y ropa impermeable. Luego empezó a fabricar ropa de seguridad para minería y pesca llegando a exportar a Bolivia. En 1984 formó Cidelsa Selva, iniciándose en la fabricación de toldos y coberturas, cuyas operaciones se llevaron a cabo en la ciudad de Pucallpa con la finalidad de aprovechar beneficios tributarios y arancelarios en la selva; en 1991 al eliminarse estos, se trasladó la planta industrial a Lima. Actualmente se fabrica y comercializa las líneas de productos mostrados en el Cuadro 3, destinados principalmente a los sectores minería y construcción.

CUADRO 3. Productos fabricados y Comercializados por CIDELSA

| PRODUCTOS | SECTORES INDUSTRIALES CONSUMIDORES Y APLICACIONES | MATERIALES UTILIZADOS Y PROCESADOS |
|--------------------------------|---|--|
| 1. Mangas de Ventilación | Minería: Socavones y Galerías. Construcción: Irrigación y Carreteras | Vinilona: Mediante sellado electrónico de alta frecuencia. Tela de Rafia de Polietileno: Mediante doble costura con hilos de nylon. |
| 2. Toldos, Mangas y Cobertizos | Minería. Construcción: viviendas. Comercio: Almacenes. Pesquería. Turismo y transporte. | Lona de Algodón: Mediante doble costura con hilos de nylon. Tejidos con recubrimiento de PVC: Mediante sellado electrónico ó electrosellado. |
| 3. Geomembranas | Minería: Impermeabilización de Procesos de extracción de minerales por solventes y almacenamiento de | Laminados en base a resinas de PVC. Laminados en base a Polietilenos de alta densidad (HDPE) (Nota: Cidelsa es representante de |

| | | |
|--|--|--|
| | relaves. Construcción: Irrigación y Saneamiento ambiental. | Serrot Corp, USA). |
| 4.Coberturas tensionadas y coberturas en general. | Comercio e Industrias en general: techos de almacenes y hangares. Turismo: Restaurantes, hoteles, etc. | Vinilona: Mediante sellado electrónico. |
| 5.Tanques Flexibles | Minería. Construcción. Pesquería. Turismo: Utilización para almacenamiento y transporte de agua. | Vinilona KPD3500 (tiene 2 capas de fibra poliéster intercalados con 3 capas de PVC). |
| 6. Geotextiles (*) | Construcción: drenes y subdrenes, rehabilitación de carreteras, control de erosión, refuerzos de túneles y terraplenes, defensas ribereñas, pavimentos asfálticos, etc. | Bidim Rhodia. |

(*) Productos sólo Comercializados

Las actividades industriales de Cidelsa pueden ser divididas en 4 rubros:

1. Transformación de laminados vinílicos (láminas flexibles de PVC) en diversos productos para aplicaciones de minería, construcción, industrias y agricultura.
2. Comercialización de Geosintéticos.
3. Comercialización de artículos de seguridad industrial.
4. Comercialización de estructuras metálicas utilizadas como soporte de los productos confeccionados con los laminados vinílicos.

PLANEAMIENTO ESTRATÉGICO

VISION :

Empresa líder en la confección y comercialización de productos plásticos, por su éxito con los clientes, innovación, tecnología y habilidad para competir exitosamente en los mercados internacionales. Para lo cual proveemos un ambiente en el cual sus empleados puedan innovar y sobresalir a través de su formación, capacitación y desarrollo. Además la empresa tiene un respeto integral por el medio ambiente y un compromiso con el desarrollo del país.

| AMBITO | ACTUAL |
|------------------------------|---|
| AMBITO DE PRODUCTOS | PRODUCCION Y COMERCIALIZACION DE CONFECCIONES PLASTICAS |
| AMBITO DE MERCADOS | SECTOR INDUSTRIAL, MANTENIMIENTO Y CONSTRUCCION |
| AMBITO GEOGRAFICO | EL PERU Y SUS PRINCIPALES CIUDADES |
| SINGULARIDADES DE LA EMPRESA | MEDIANA, FLEXIBLE, AGIL |

MISION :

| AMBITO | ACTUAL | FUTURO |
|------------------------------|---|--|
| AMBITO DE PRODUCTOS | PRODUCCION Y COMERCIALIZACION DE CONFECCIONES PLASTICAS | SATISFACCION DE NECESIDADES QUE TENGAN QUE VER CON EL USO DE PLASTICOS |
| AMBITO DE MERCADOS | SECTOR INDUSTRIAL, MINERO Y CONSTRUCCION | AGREGAR EL SECTOR: AGRÍCOLA PESQUERO Y PUBLICITARIO |
| AMBITO GEOGRAFICO | EL PERU Y SUS PRINCIPALES CIUDADES | TODA SUDAMERICA |
| SINGULARIDADES DE LA EMPRESA | MEDIANA, FLEXIBLE, AGIL | MEDIANA, FLEXIBLE, AGIL Y EFICIENTE CALIDAD EN EL SERVICIO AL CLIENTE |

3.1.3 ANALISIS DE LAS CINCO FUERZAS DEL SECTOR

1. - Nuevos Ingresos:

Existen empresas extranjeras deseosas de ingresar al mercado local como ARCOTEX de Chile y IKL de Colombia que aunque han exportaciones de algunos productos al mercado local todavía no se avizora la implementación de oficinas y planta en el país.

2. - Proveedores:

El poder de negociación que CIDELSA tiene con sus proveedores es bajo, ya que es una empresa cuyo insumo principal es la membrana plastificada, que es importada directamente de Corea, Brasil, EEUU, Canadá; para dichos proveedores el volumen de compra de CIDELSA no es significativo.

3. - Compradores:

Si hablamos por líneas de productos, podemos afirmar:

- Línea de Mangas de Ventilación

Bajo poder de negociación, debido a la competencia extranjera y la gran cantidad de empresas locales que fabrican y comercializan dicho producto.

- Línea de Geomembranas

Bajo poder de negociación, debido a la competencia extranjera y a los proveedores locales que fabrican y/o comercializan dicho producto.

- Línea de Geotextiles

Mediano poder de negociación, en nuestro país sólo existen representantes de compañías extranjeras, CIDELSA es representante de PAVCO de Colombia y gran parte del mercado local lo tiene dicha marca.

- Línea de Tanques Flexibles

Alto poder de negociación, debido a que no existe competencia local y la extranjera tiene mayores precios y tiempos de entrega más largos.

- **Línea de Coberturas**

Si hablamos de coberturas para galpones y coberturas comerciales, el poder de negociación es bajo debido a la alta competencia nacional con productos sustitutos: fibraforte, eternit, calamina, etc., si hablamos de coberturas tensionadas el poder de negociación es alto, ya que no existe competencia local de envergadura para proyectos medianos y grandes. La competencia extranjera está en desventaja por precios y tiempo de entrega.

4. -Sustitutos:

Recientemente se han desarrollado nuevos materiales para coberturas de galpones de bajo precio y de diversos colores como techolit, planchas acrílicas, etc. que compiten con nuestra línea de coberturas para galpones industriales.

En lo relacionado a geomembranas se tiene un nuevo producto: membrana asfáltica, que da mejores resultados en altas temperaturas ambientales.

5. -Competencia Actual en el Sector:

- **Línea de Mangas de Ventilación**

En mangas de ventilación de polietileno, se ha bajado la participación del mercado a un 40%, debido a la alta competencia con proveedores locales. También se han bajado los márgenes de contribución. En mangas de ventilación de PVC con tejido interno de poliéster, existe competencia internacional permanente proveniente de Noruega, Canadá, Chile, etc.

- **Línea de Geomembranas**

Para proyectos grandes, se compite con empresas de EEUU, Chile y Canadá. Para pequeños proyectos los competidores son empresas locales que tienen precios competitivos pero no tienen la infraestructura de CIDELSA, es decir infraestructura para mejorar tiempos de entrega y garantizar la calidad del producto.

- Línea de Geotextiles

Fuerte competencia con representantes de otras marcas de Brasil, EEUU y Canadá

- Línea de Coberturas Tensionadas y Coberturas Especiales

En la actualidad no existe una competencia local importante en éste rubro y aquella proveniente de EEUU, Canadá, etc., llega con precios muy elevados.

- Línea de Tanques Flexibles

No existe fabricante local de productos similares. Empresas chilenas y argentinas ofrecen sus productos, pero los clientes se encuentran dispersos y solicitan pequeñas cantidades, además por el corto tiempo de entrega que solicitan, los lleva a decidir por el producto de CIDELSA. Existen también representantes de compañías de Francia y EEUU, pero ofrecen productos con tiempos de entrega más largos y a mayores precios.

3.1.4 ANALISIS DEL ENTORNO

Luego de analizar las cinco fuerzas del entorno competitivo de CIDELSA y las matrices SEPT/DIR y de INCIDENCIA, detectamos lo siguiente:

Oportunidades :

- Megaproyectos en el sector minero y energético: proyectos como Antamina y Camisea van a generar requerimientos de Geomembranas, Geotextiles, Tanques Flexibles, Coberturas Especiales, etc.
- Exportar algunos productos como: Coberturas tensionadas, Tanques flexibles, Mangas de ventilación, etc.
- No existe competencia local en Coberturas Tensionadas y Tanques Flexibles por lo que es posible tener un gran poder de negociación con los clientes locales
- Aparte de los Tanques Flexibles para almacenar agua es posible desarrollar tanques para almacenar diesel ó gasolina

Amenazas :

- Ingreso de nuevos competidores extranjeros: colombianos y chilenos. Aunque todavía no instalándose con oficinas y planta sino más bien a través de representantes.
- Crecimiento de competidores nacionales: tenemos una empresa local que se ha especializado en la fabricación y comercialización de mangas de ventilación. Muchos extrabajadores de CIDELSA se han juntado y han formado una empresa que aunque es pequeña hay que tener mucho cuidado con su posible desarrollo, (Mayor información ver Matrices en Anexo 2.3).

3.1.5 ANALISIS DE LA CADENA DEL VALOR

Actividades Primarias:

Logística Interna.- El abastecimiento de insumos locales no es crítico, existen muchos proveedores locales y con buenos precios. Dentro de la estructura de costos los insumos importados representan entre el 60% al 70% del costo total del producto. La logística de éstos insumos es crítica por oportunidad y precios.

Operaciones.- Aunque se tiene una planta de fabricación, cuando hay que atender los picos del mercado local ó requerimientos de exportación se cuenta con el apoyo de plantas de Colombia y Brasil que nos brindan servicio de confección.

Logística Externa.- Por el tipo de producto, es decir producto no masivo de uso industrial, los canales de distribución no son complicados. Si se decidiera exportar con mayor fuerza tal vez sería necesario un mayor análisis de los costos de fletes marítimos, aranceles, etc.

Mercadotecnia y Ventas.- Esta organizado por sectores económicos, es decir: Construcción, Minería, Petróleo, Industria, Agricultura, Avícolas, Pesquería, Comercio, etc.

Los productos netamente técnicos se tratan de especificar primero con los consultores e ingenieros especialistas. Los representantes de venta que trabajan en un determinado sector económico pueden vender cualquier producto de CIDELSA. También se tienen representante de ventas en Provincias.

Servicio.- Aunque no existe un departamento que realice dicha función, es política de la empresa atender cualquier reclamo de cualquier cliente.

Actividades de Apoyo:

Infraestructura de la Empresa.- Dado el crecimiento de la empresa en los últimos años se requiere de herramientas de Planificación y Control financiero que actualmente no se posee. Los sistemas de información con que se cuentan apoyan la operatividad de la empresa no se cuenta con información que coadyuve a la toma de decisiones a nivel gerencial.

Administración de Recursos Humanos.- No esta asignado a una Gerencia en particular. Cada Gerencia recluta su propio personal requerido. Ante el crecimiento de la empresa en los últimos años, se ha contratado gerentes, empleados y obreros sin un formal proceso de selección. No existe un programa de capacitación del personal.

Desarrollo Tecnológico.- Se asiste permanentemente a ferias y exposiciones a nivel mundial, donde se aprecia la tecnología de punta en cuanto a maquinarias y materias primas. Se visitan periódicamente las plantas de fabricación más importantes de América Latina.

Abastecimiento.- La parte de importaciones es una función importante por el tipo de producto que comercializa CIDELSA. , tanto en sus insumos como en sus equipos. El departamento de Importaciones sólo realiza la parte operativa de la función. La planificación y control lo realiza cada gerencia implicada.

3.1.6 ANÁLISIS INTERNO

Luego de analizar las actividades primarias y las de apoyo de la cadena de valor de CIDELSA y sus interrelaciones podemos concluir:

Fortalezas :

- Al tener una Planta de fabricación en el país y contar con el apoyo de otras Plantas en Colombia y Brasil, le da a CIDELSA la oportunidad de ser bastante flexible con los requerimientos del mercado.
- Tener tecnología de equipos e insumos de última generación, por la permanente actualización en ferias, viajes al extranjero.
- Infraestructura de comercialización que tiene llegada a casi todos los clientes importantes de todos los sectores del país.

Debilidades :

- No existe una estrategia para la Administración de los Recursos Humanos, que permita realizar eficientemente el proceso de selección, capacitación y motivación del personal de la empresa.
- La gestión financiera de la empresa debe dotarse de las últimas herramientas para el apoyo de su función, es decir realizar las actividades de Planificación y Control Financiero eficientemente.
- CIDELSA no cuenta con un sistema de información para la ayuda en la toma de decisiones a nivel gerencial. El mercado cada vez más cambiante y presto a reacciones cada vez más rápidas por parte de las empresas, hace imprescindible contar con éste tipo de herramientas de gestión.

Luego de haber efectuado el análisis estratégico complementado con el desarrollo de las matrices para evaluar las opciones estratégicas, concluimos que la estrategia que mayor puntaje recibe es: **Desarrollar un sistema de apoyo a las decisiones que permita incrementar la participación en el mercado de los productos de Cidelsa.**

3.2 DIAGNOSTICO DE LA INFRAESTRUCTURA DE TECNOLOGÍA DE INFORMACIÓN DE LA ORGANIZACIÓN

3.2.1 Infraestructura de Hardware y Redes

Cidelsa cuenta actualmente con 30 computadoras conectadas en red LAN Fast Ethernet, mediante la tipología estrella, de 10/100 UTP, cableado estructurado categoría 5, todas las PC están por encima de procesadores Pentium 300 Mhz. Se tiene un servidor que opera con Windows 2000 Server, de las siguientes características :

Pentium III 600 Mhz, con 02 discos duros de 7 GB y 10 GB de tecnología SCSI, con 256 Ram. Se tiene un servidor de correos alojado en el servidor principal que tiene un Proxy server para compartir y aumentar la velocidad del acceso a internet. Además tiene un servidor de correos mail server de 32 bits denominado Mercury de Pegasus. Se tiene acceso a internet a través de una línea digital de 128 kbps.

3.2.2 Infraestructura de Software

Sistema de contabilidad, costos, ventas, almacenes que fue desarrollado en Fox Pro para DOS y que estuvo en operación desde 1,990 a 1998.

Sistema de Gestión Empresarial que contempla: contabilidad, costos, ventas, almacenes, producción, presupuestos, caja y bancos, cuentas por cobrar, cuentas por pagar, importaciones. Este sistema ha sido desarrollado en Visual Fox y está en operación desde 1998 hasta la fecha.

Adicionalmente a los sistemas de producción detallados líneas arriba detectamos que parte de la información pasada que se requiere para cargar el datawarehouse está en hoja de cálculo tipo Excel.

3.2.3 Infraestructura de personal de TI

Se cuenta con una Jefe de Sistemas que se encarga de dar mantenimiento a todos los sistemas de producción. El desarrollo de nuevas aplicaciones se encarga a terceros. El mantenimiento del hardware y redes se ha encargado a una empresa, que se encarga de dar el mantenimiento preventivo y correctivo. Esta empresa también colabora en la evaluación de toda adquisición de hardware.

3.2.4 Aplicación de la Teoría de las Etapas

De acuerdo a lo expuesto en el Capítulo I - Marco Teórico Referencial, en la parte de correspondiente a la Teoría de las Etapas ó Curva de Nolan, para el caso de la empresa CIDELSA, podemos concluir en lo siguiente en cuanto al desarrollo de los sistemas de información:

Cartera de Aplicaciones.-

Existen proliferación de aplicaciones para la reducción de los costos de las funciones (sistematizar los procedimientos administrativos).

Organización.-

Los programadores están orientados al usuario. Se entrevistan a los usuarios para obtener sus requerimientos.

Planeación y control del procesamiento de datos.-

Se compromete recursos necesarios para la implementación inicial de los sistemas de información, no existen muchas regulaciones para el uso del hardware y software.

Posición del usuario.-

El usuario se encuentra entusiasmado superficialmente con el procesamiento de datos.

Del análisis anterior podemos concluir que la empresa CIDELSA se encuentra en la etapa II CONTAGIO (ver cuadro 1), por lo tanto la

planeación de los sistemas de información deben tener presente este hecho, tratando de alentar una mayor participación de los usuarios en el uso de las herramientas de tecnología de información.

Los presupuestos en estas áreas tendrán que ser mayores y el control tendrá que ejercerse a un nivel general y no específico. Se tendrán que dar los primeros pasos para mejorar la infraestructura de los sistemas de información, por ejemplo para implementar una Base de Datos relacional.

Es importante señalar que un severo control y poca laxitud en ésta etapa puede obstaculizar el empleo de la tecnología de la información.

3.2.5 Tipo de Sistema de Soporte a las Decisiones

De acuerdo a lo expuesto en el Marco Teórico Referencial y a lo evaluado hasta el momento en CIDELSA en lo referido a tecnologías de información pensamos que el Sistema de Decisiones Tipo I es el que se encuadra mejor dentro de la estrategia a desarrollar (Ver figura 7). Dicho sistema se caracteriza por dar respuesta a la pregunta estratégica: ¿Qué sucedió?.

El datamart deberán generar una serie de reportes que nos dará una visión y una habilidad para usar combinadamente la data detallada a través de la organización para entender el pasado, en su primera fase.

3.3 DESARROLLO DEL DATAMART

3.3.1 ACUERDOS DEL PROYECTO.-

3.3.1.1 Objetivo Estratégico:

Desarrollar un sistema de apoyo a las decisiones que permita incrementar la participación en el mercado de los productos de Cidelsa.

3.3.1.2 Objetivos del Data Mart:

Comercialización dirigida, tanto para incorporar clientes nuevos como para conservar los actuales.

Productos de ventas cruzadas a clientes actuales

Mayoreo de productos

Incrementar la rotación de capital invertido en Materias Primas y Mercaderías

3.3.1.3 Requerimientos específicos del negocio:

Indicadores de Ventas

- Ventas en Unidades Físicas por Línea de Producto por Sector Económico en un tiempo determinado
- Ventas en Volumen Monetario (US\$) por Línea de Producto por Sector Económico en un tiempo determinado
- Ventas en Unidades Físicas por Cliente por Sector Económico por Línea de Producto en un tiempo determinado
- Ventas en Volumen Monetario (US\$) por Cliente por Sector Económico por Línea de Producto en un tiempo determinado

Indicadores de Participación en el Mercado

- Participación de mercado por Línea de Producto en un tiempo determinado

Indicadores de Gastos y Utilidad Bruta

- Gastos por Orden de Trabajo para una división específica en un tiempo determinado
- Utilidades brutas por sector económico por Línea de Producto en un tiempo determinado

Indicadores de Logística

- Consumo de materias primas y mercaderías por Línea de Producto en un tiempo determinado
- Rotación de materia primas y mercaderías en un tiempo determinado

3.3.1.4 Análisis Costo-Beneficio
Costos:

RECURSOS HUMANOS

| ACTIVIDADES | TIEMPO (semanas) | | | | | | | | | | PG1 | PG2 | | | | |
|--|------------------|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|--|---|
| | 1s | 2s | 3s | 4s | 5s | 6s | 7s | 8s | 9s | JFP | | | AS1 | AS2 | | |
| ACUERDOS DEL PROYECTO | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | | | | | | |
| Analista de Sistemas N° 01 | | | | | | | | | | | 1 | | | | | |
| Analista de Sistemas N° 02 | | | | | | | | | | | | 1 | | | | |
| MODELO DE DATOS | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |
| Analista de Sistemas N° 01 | | | | | | | | | | | | 1 | | | | |
| Analista de Sistemas N° 02 | | | | | | | | | | | | | 1 | | | |
| Procesos Back End: Modelo Starnet. | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |
| Analista de Sistemas N° 01 | | | | | | | | | | | | 1 | | | | |
| Procesos Back End: Tabla de Hechos y Dimensiones. | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |
| Analista de Sistemas N° 02 | | | | | | | | | | | | | 1 | | | |
| Procesos Back End: Mapeo con Fuentes de Datos. | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |
| Programador N° 01 | | | | | | | | | | | | | | | | 1 |
| Procesos Back End: Determinación de requerimientos de transformación de la data. | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |
| Programador N° 02 | | | | | | | | | | | | | | | | 1 |
| Procesos Back End: Extracción, limpieza y transformación de la data. | | | | | | | | | | | | | | | | |
| Jefe del proyecto | | | | | | | | | | | 1 | | | | | |

| Recursos de Hardware | Costos |
|------------------------|--------|
| Servidor IBM Netfinity | 2800 |
| Impresora Laser | 1500 |
| Suministros | 500 |

| | |
|---------------|-------------|
| Costos | 4800 |
|---------------|-------------|

| Recursos de Software Base | Costos |
|--------------------------------|--------|
| S.O. WINDOWS 2000 | 820 |
| MS OFFICE | 400 |
| ORACLE INTERNET DEVELOPER SUIT | 7245 |
| ORACLE STANDARD | 4365 |

| | |
|---------------|---------------|
| Costos | 12 830 |
|---------------|---------------|

| Recursos Humanos | Costos |
|------------------------------|--------|
| Jefe del Proyecto | 4500 |
| Analista de Sistemas Nro. 01 | 1800 |
| Analista de Sistemas Nro. 02 | 1200 |
| Programador Nro. 01 | 750 |
| Programador Nro. 02 | 500 |

| | |
|-----------------------|-------------|
| Costos de RRHH | 8750 |
|-----------------------|-------------|

| | |
|-----------------------|------------------|
| Costos Totales | \$ 26 380 |
|-----------------------|------------------|

Beneficios del Proyecto:

Acceso fácil a una data actualizada

Cuantitativos : productividad del usuario final

Eleva las Ventas:

Estas tácticas emplean el Datamart para determinar los segmentos del mercado y la predisposición de los clientes por productos determinados, el mayoreo de productos y opciones de paquetes. Incluye también aislar a clientes no rentables en segmentos rentables del mercado ó segmentos no rentables del micromercado en segmentos rentables de mercado, ó productos no rentables en líneas y bandas de producto rentables, y su “traspaso a la competencia”. En otras palabras el Marketing mejor direccionado.

Tenemos como promedio de Ventas al año: \$3,600,000.00

Si incrementamos al final del primer año 15 %, y de acuerdo a los índices de la empresa la utilidad neta se incrementaría en 3%, es decir tendríamos un ingreso adicional al final del primer año de: \$ 27,000.00

Incrementar la rotación de inventarios:

Se tiene como promedio los siguientes niveles de inventario como promedio:

Mercaderías : \$ 450,000.00

Materias Primas: \$ 350,000.00

El datamart nos segmentaría los mercados y clientes más rentables, por lo tanto nuestras inversiones en stocks en Mercaderías y Materias Primas estaría mejor diseñada.

Estimamos que reduciríamos en un 20% al final del primer año de implementado el datamart, es decir Mercaderías y Materias primas no rentables y sin movimiento sustantivo.

Si le aplicamos la tasa bancaria actual de 19.2%, tendríamos un ahorro de: \$30,720.00

Cualitativos :

Acceso fácil a una data actualizada

Mejora de la productividad del usuario final

Reportes más rápidos

Confiabilidad del Sistema

Mejora de la performance en las consultas

3.3.1.5 Alcance funcional de alto nivel:

De acuerdo a los resultados encontrados luego de realizar el análisis estratégico de Cidelsa vamos a circunscribir el presente Datamart al área Comercial.

A saber en CIDELSA, se encuentran los siguientes sistemas de producción:

Sistema de contabilidad, costos, ventas, almacenes que fue desarrollado en Fox Pro para DOS y que estuvo en operación desde 1,990 a 1998.

Sistema de Gestión Empresarial que contempla: contabilidad, costos, ventas, almacenes, producción, presupuestos, caja y bancos, cuentas por cobrar, cuentas por pagar, importaciones. Este sistema ha sido desarrollado en Visual Fox y está en operación desde 1998 hasta la fecha.

Adicionalmente a los sistemas de producción detallados líneas arriba detectamos que parte de la información pasada que se requiere para cargar el datawarehouse está en hoja de cálculo tipo Excel.

Por el volumen de información que se tiene y por lo especificado por los usuarios en cuanto a la frecuencia ó periodicidad de la información, creemos que lo más recomendable es cargar el datawarehouse en forma semanal.

3.3.1.7 Infraestructura :

Se ha decidido adquirir un Servidor para administrar toda la información relacionada al Datamart, porque el actual ya se encuentra sobrecargado con los sistemas de producción y la administración de las cuentas de correo.

Para la parte de software, teníamos 02 opciones: realizar el desarrollo nosotros mismos ó comprar herramientas ya existente en el mercado y con probados resultados. Se decidió escoger de las herramientas disponibles en el mercado por el tiempo de desarrollo y pruebas además de la confiabilidad del desarrollo que pudiéramos efectuar y de la calificación requerida de los equipos de trabajo disponibles..

A continuación enumeramos las herramientas que pudimos encontrar disponibles en el mercado:

3.3.1.8 Responsabilidad de los Usuarios:

A diferencia de otras metodologías el usuario deberá participar permanentemente en

Definición de requerimientos

Testeo de reportes y aplicaciones

Definición de alcances del Datamart

Educa a los usuarios en las capacidades del Datamart

3.3.1.9 Requerimientos de Disponibilidad:

El acceso a la información del datamart deberá estar disponible, inicialmente de Lunes a Viernes de 8 a.m. a 8 p.m.

Examina las aplicaciones

3.3.1.10 Requerimientos de Seguridad:

Implementar monitoreos e implementar posibles repuestas a incidentes

Implementar accesos a usuarios y seguridad

Evaluar y seleccionar plataformas para hardware

3.3.1.11 Roles y Responsabilidades del Equipo del Proyecto:

Administrador del Proyecto.-

Educa a los ejecutivos en las aplicaciones e impacto del Datamart

Obtiene soporte económico

Identifica y prioriza las aplicaciones
Define presupuestos y programas
Monitorea programas y presupuestos
Educa a los usuarios en capacidades del Datamart
Monitorea tendencias de la industria e identifica tecnologías que podrían ser adoptadas.

Administrador del Back End.-

Modela la data
Automatiza procesos de carga
Mantiene y actualiza los procesos de carga y actualización
Desarrolla sistemas de monitoreo para asegurar que la data ha sido actualizada y desarrolla programas de limpieza de data y implementación de la data adquirida.
Identifica fuentes de data
Programa procesos de adquisición y transformación de la data

Administrador del Front End.-

Desarrolla aplicaciones de soporte a las decisiones
Educa a los usuarios en las capacidades del Datamart
Entrena a los usuarios
Ayuda a los usuarios a encontrar la información correcta
Desarrolla aplicaciones tipo reportes
Examina las aplicaciones
Documenta las aplicaciones.
Administrador del Proyecto-Administrador del Back End-Administrador del Front End.-
Evaluar y seleccionar plataformas para hardware
Evaluar y seleccionar facilidades de redes
Evaluar y seleccionar software de administración de base de datos
Evaluar y seleccionar software de transformación y limpieza de data
Evaluar y seleccionar software de reportes, consultas y análisis

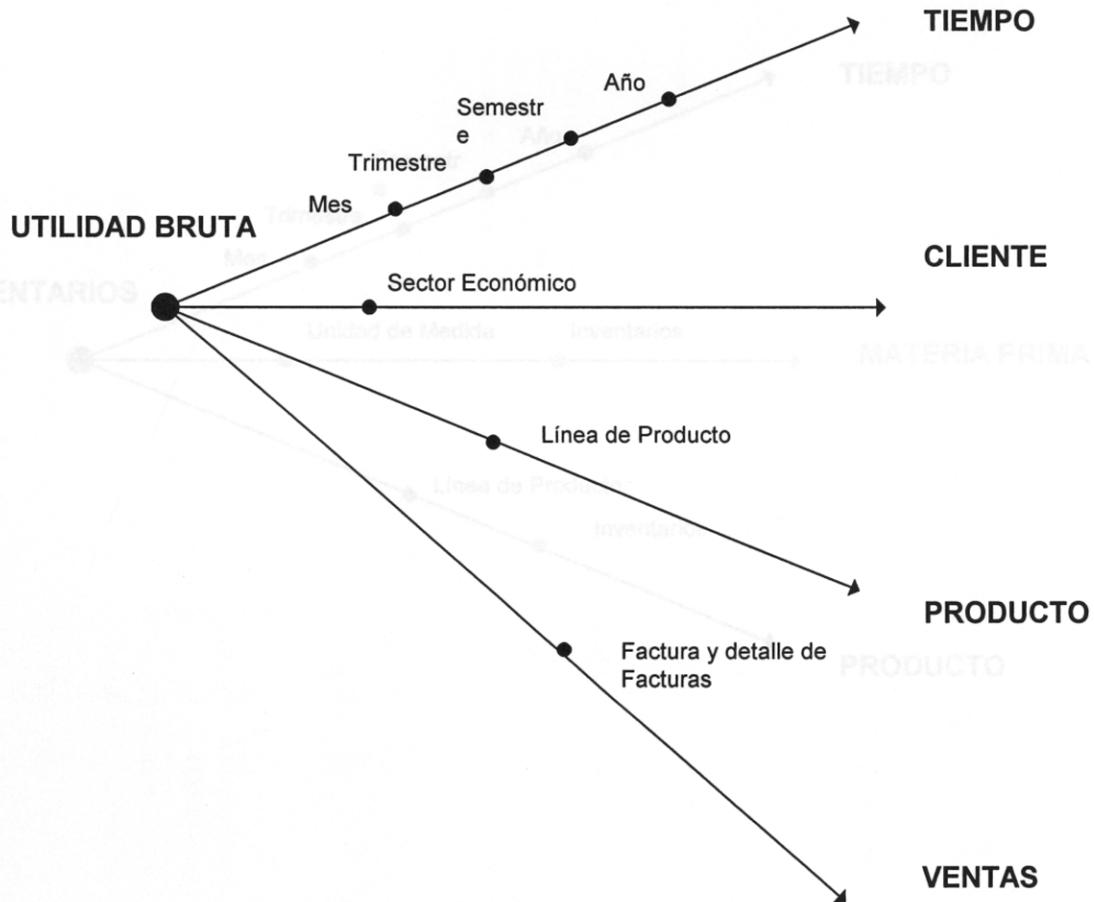
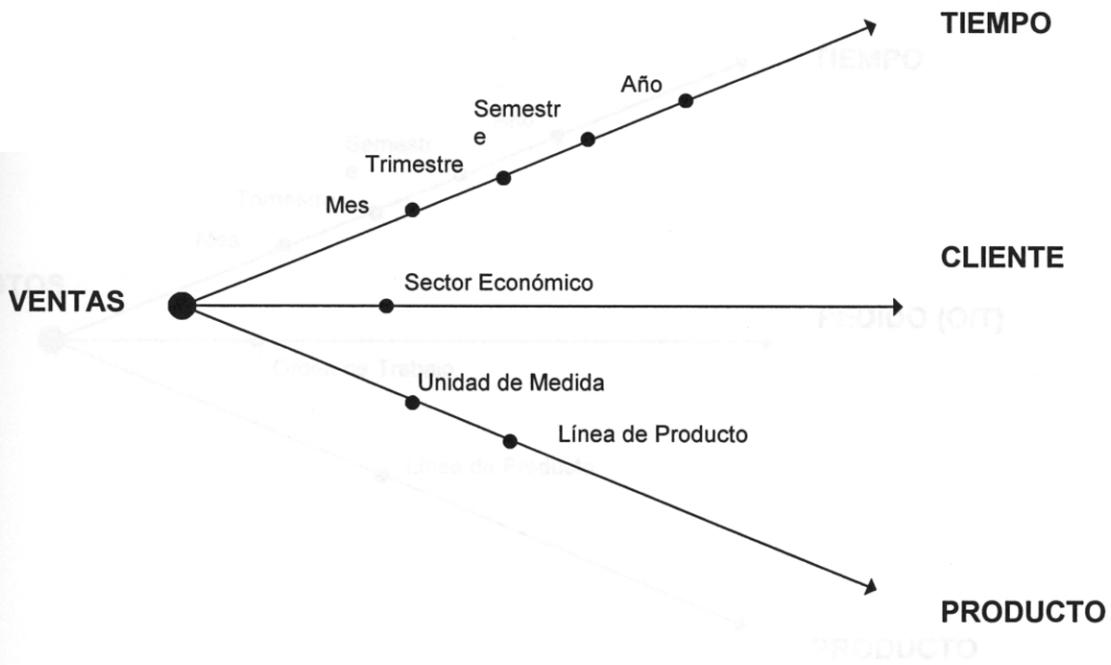
3.3.1.12 PLAN DE ALTO NIVEL

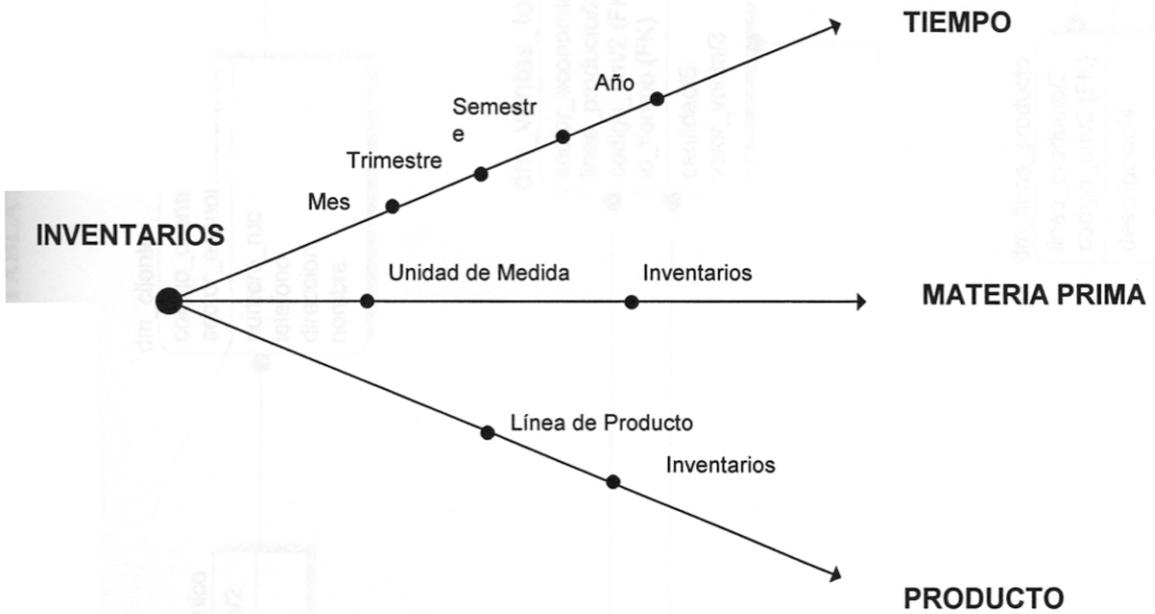
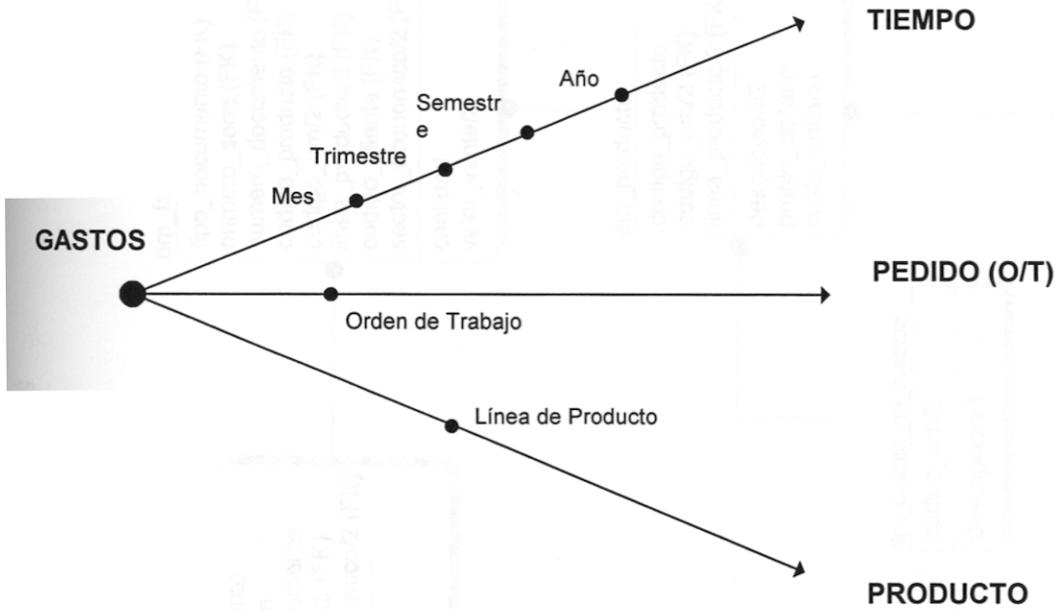
| ACTIVIDADES | TIEMPO (semanas) | | | | | | | | |
|--|------------------|----|----|----|----|----|----|----|----|
| | 1s | 2s | 3s | 4s | 5s | 6s | 7s | 8s | 9s |
| ACUERDOS DEL PROYECTO | ■ | | | | | | | | |
| MODELO DE DATOS | | ■ | | | | | | | |
| Procesos Back End: Modelo Starnet. | | | ■ | | | | | | |
| Procesos Back End: Tabla de Hechos y Dimensiones. | | | | ■ | | | | | |
| Procesos Back End: Mapeo con Fuentes de Datos. | | | | | ■ | | | | |
| Procesos Back End: Determinación de requerimientos de transformación de la data. | | | | | | ■ | | | |
| Procesos Back End: Extracción, limpieza y transformación de la data. | | | | | | | ■ | | |
| Procesos Back End: Carga de data. | | | | | | | | ■ | |
| Procesos Front End: Desarrollo de Reportes y Gráficos. | | | | | | | | | ■ |
| PROCESO ESTABLECIMIENTO PLATAFORMA TECNOLÓGICA | | | | | | | | | |
| IMPLANTACIÓN | | | | | | | | | ■ |

3.3.2 Proceso Back End

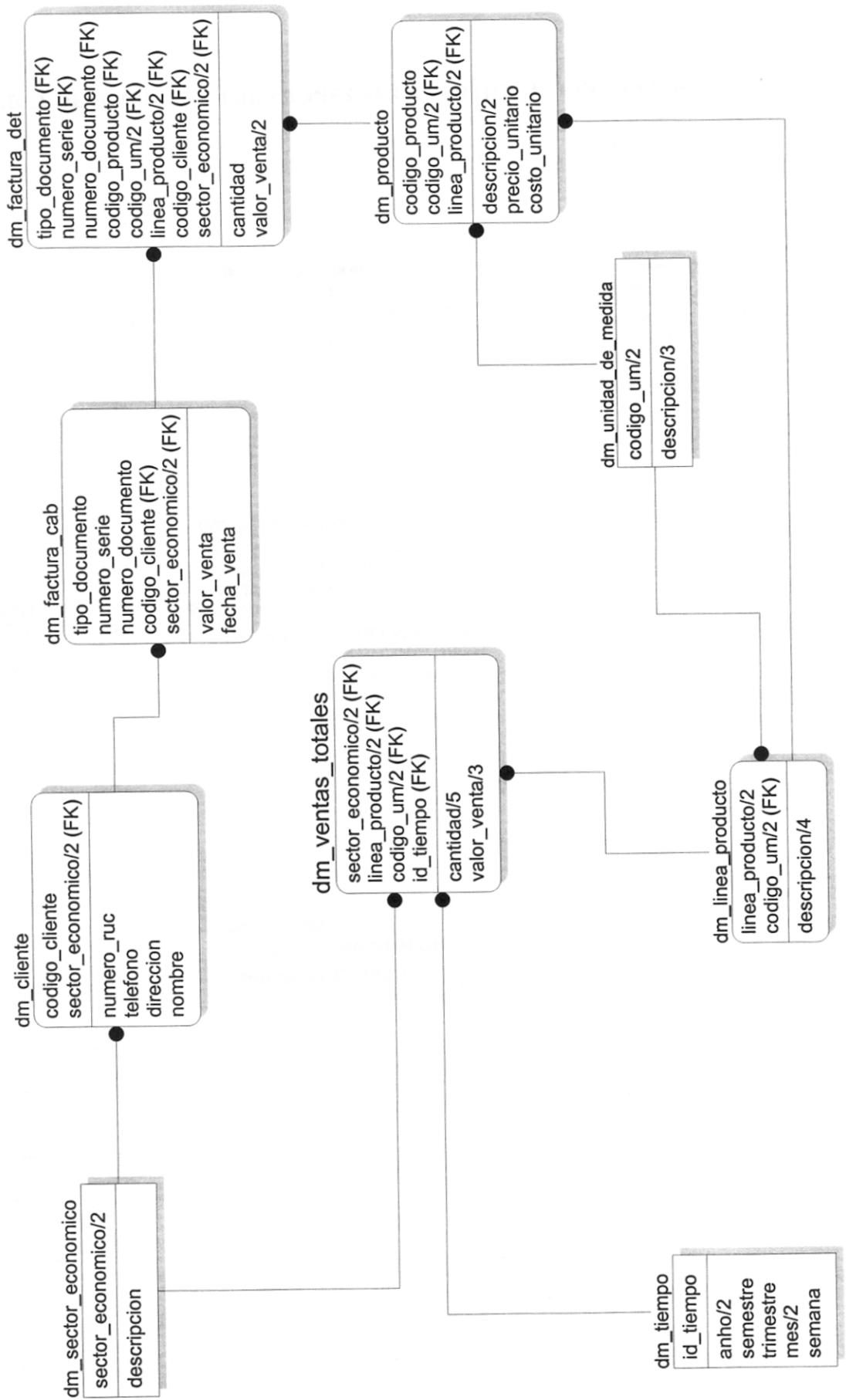
3.3.2.1 Modelo Starnet



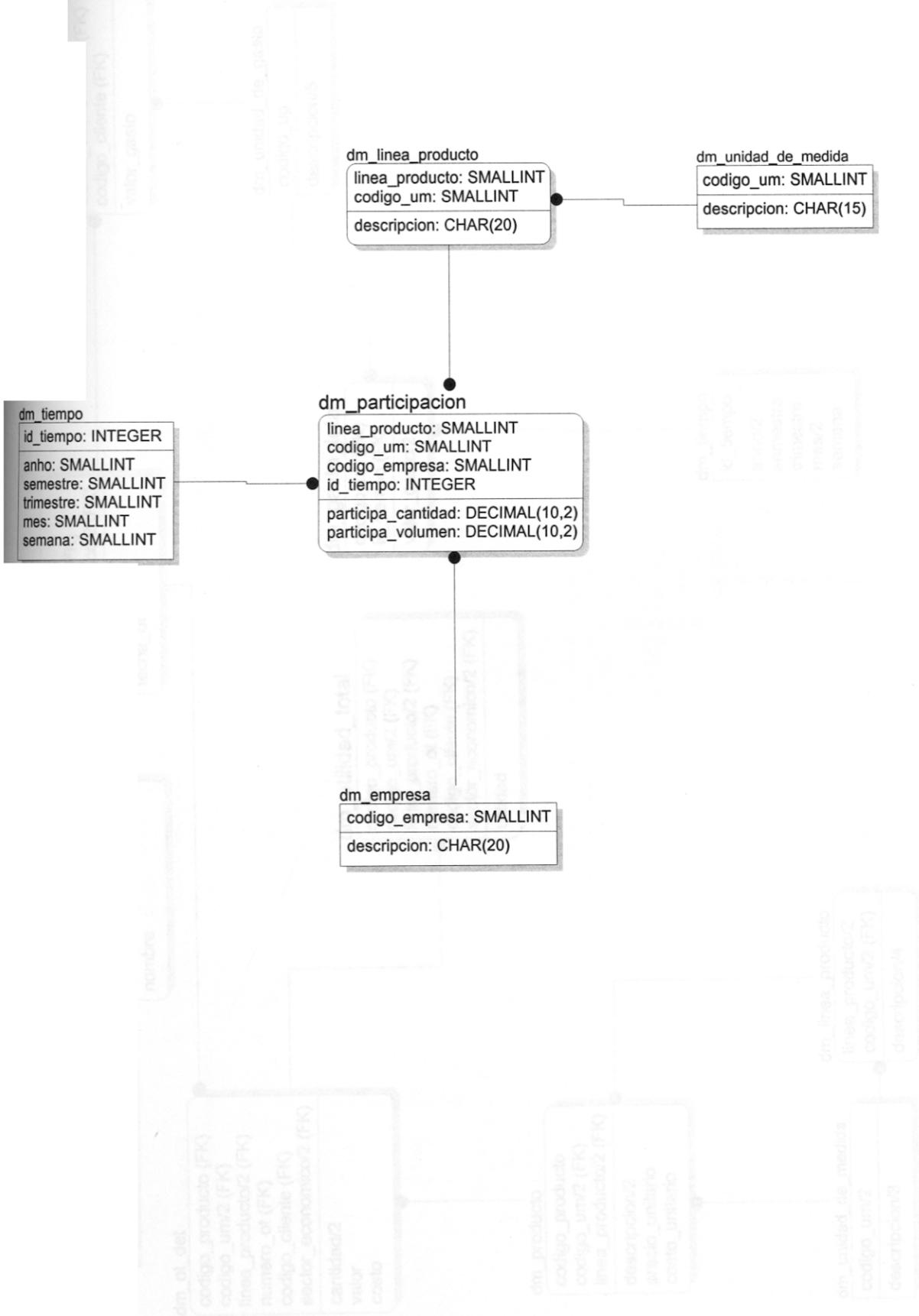




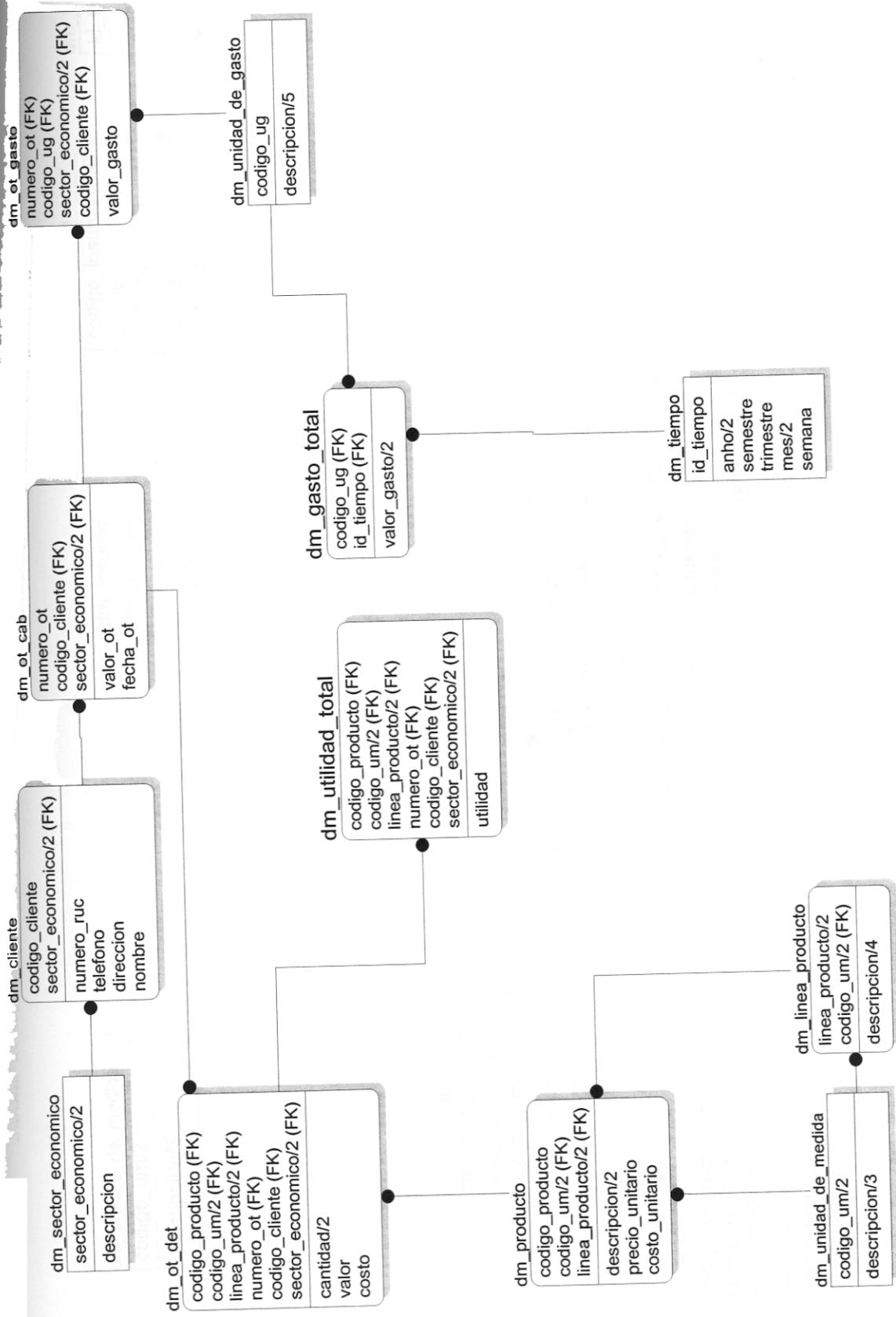
3.3.2.2 TABLA DE HECHOS Y DIMENSIONES POR LAS VENTAS



3.3.2.3 TABLA DE HECHOS Y DIMENSIONES DE LA PARTICIPACION EN EL MERCADO

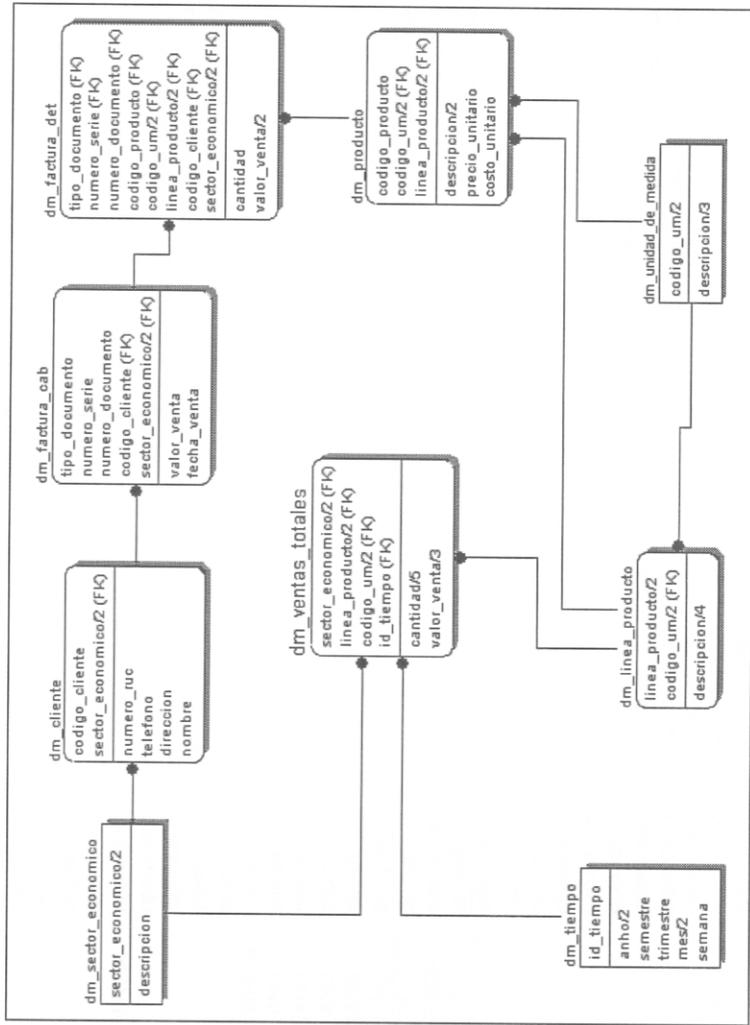
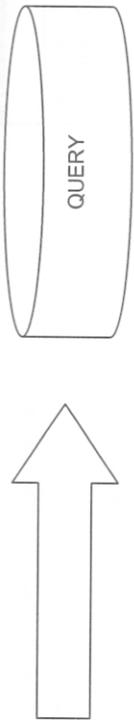
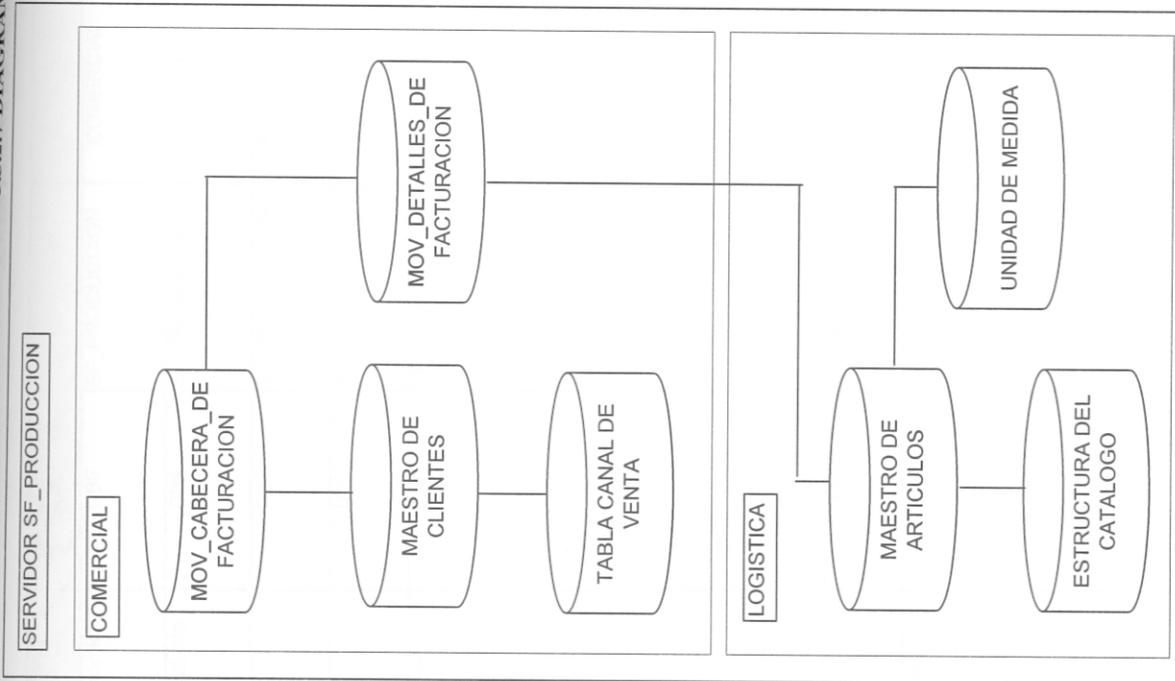


3.3.2.4 TABLA DE HECHOS Y DIMENSIONES POR EL GASTO Y LA UTILIDAD BRUTA



3.3.2.6 FUENTES DE DATOS PARA LOS INDICADORES POR VENTAS

| VARIABLES | TIPO DE BASE DE DATOS | SERVIDOR | BASE DE DATOS | SCRIPT | DESCRIPCION |
|-------------------|-----------------------|--------------|---------------|---|---|
| Tipo de Documento | DBF | SF_PRODUCION | COMERCIAL | SELECT fact_cod01c FROM comercial->mov_detalle_de_facturacion | Selecciona el tipo de documento: Factura o Boleta de Venta |
| Sector Económico | DBF | SF_PRODUCION | COMERCIAL | SELECT cli_cnl01c FROM comercial->maestro_cliente | Selecciona los campos de cli_cnl01c correspondiente al Sector Económico de la entidad |
| Línea de Producto | DBF | SF_PRODUCION | LOGISTICA | SELECT arti_str01c FROM logistica->maestro_articulos | Selecciona los campos de arti_str01c |
| Cantidad | DBF | SF_PRODUCION | COMERCIAL | SELECT fact_can01n FROM comercial->mov_detalle_de_facturacion WHERE fact_ind011 is TRUE | Selecciona el campo de cantidad facturada para un documento no anulado |
| Volumen de Ventas | DBF | SF_PRODUCION | COMERCIAL | SELECT fact_val05n FROM comercial->mov_detalle_de_facturacion WHERE fact_ind011 is TRUE | Selecciona el campo de valor de venta para un documento no anulado; el valor es el neto sin considerar el IGV |
| Ventas Totales | DBF | SF_PRODUCION | COMERCIAL | SELECT cab_fact_cod01c, cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.fact_can01n, det.fact_05n, cli.clie_cnl01c, art.arti_str01c, cab.fact_mnd01c, cab.fact_tca01n FROM comercial->mov_cabecera_de_facturacion cab, comercial->mov_detalle_de_facturacion det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.fact_ind011 is TRUE AND cab.fact_cod01c= det.fact_cod01c AND cab.fact_num01c= det.fact_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c | Con esta consulta es posible levantar la información por ventas desde la fuente de información de la base Comercial |

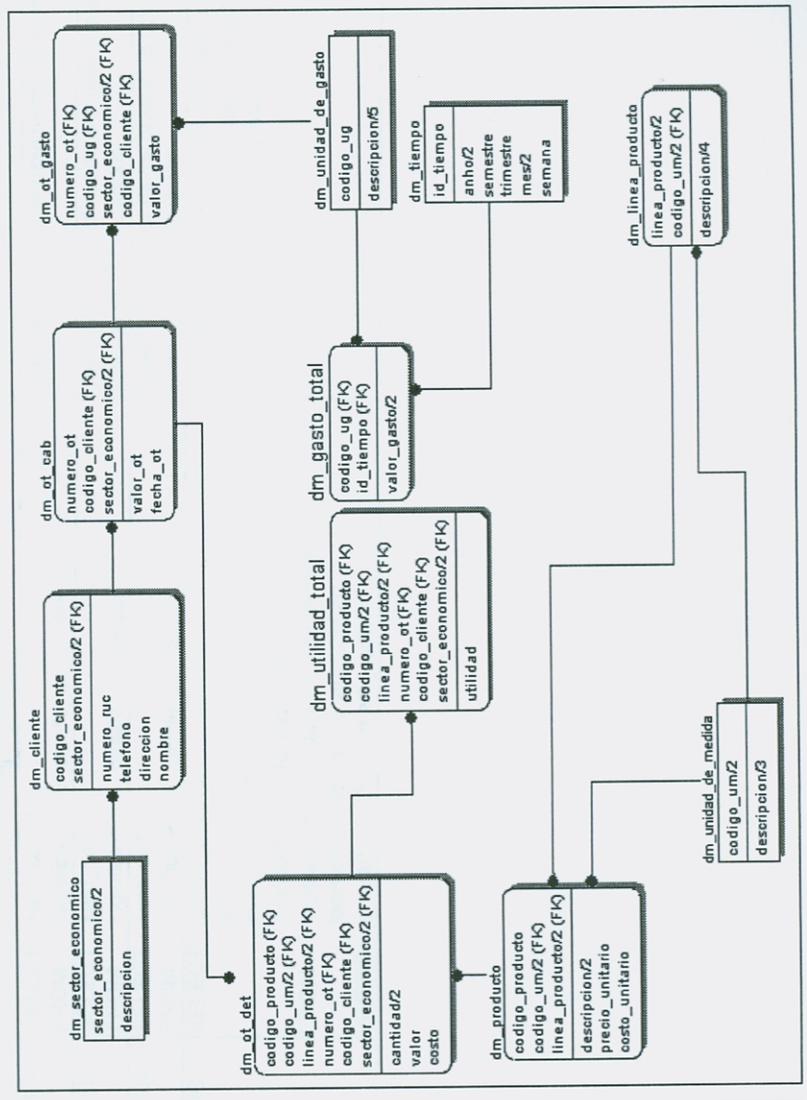
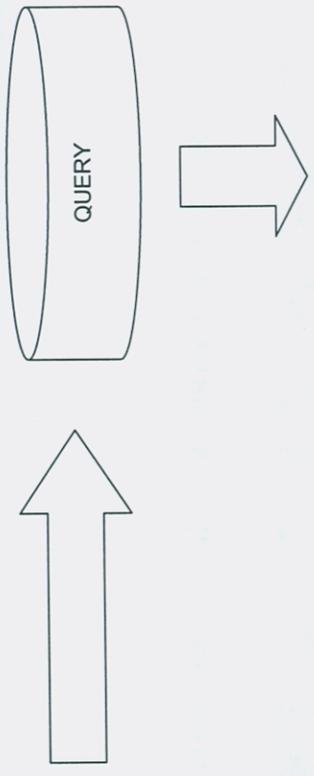
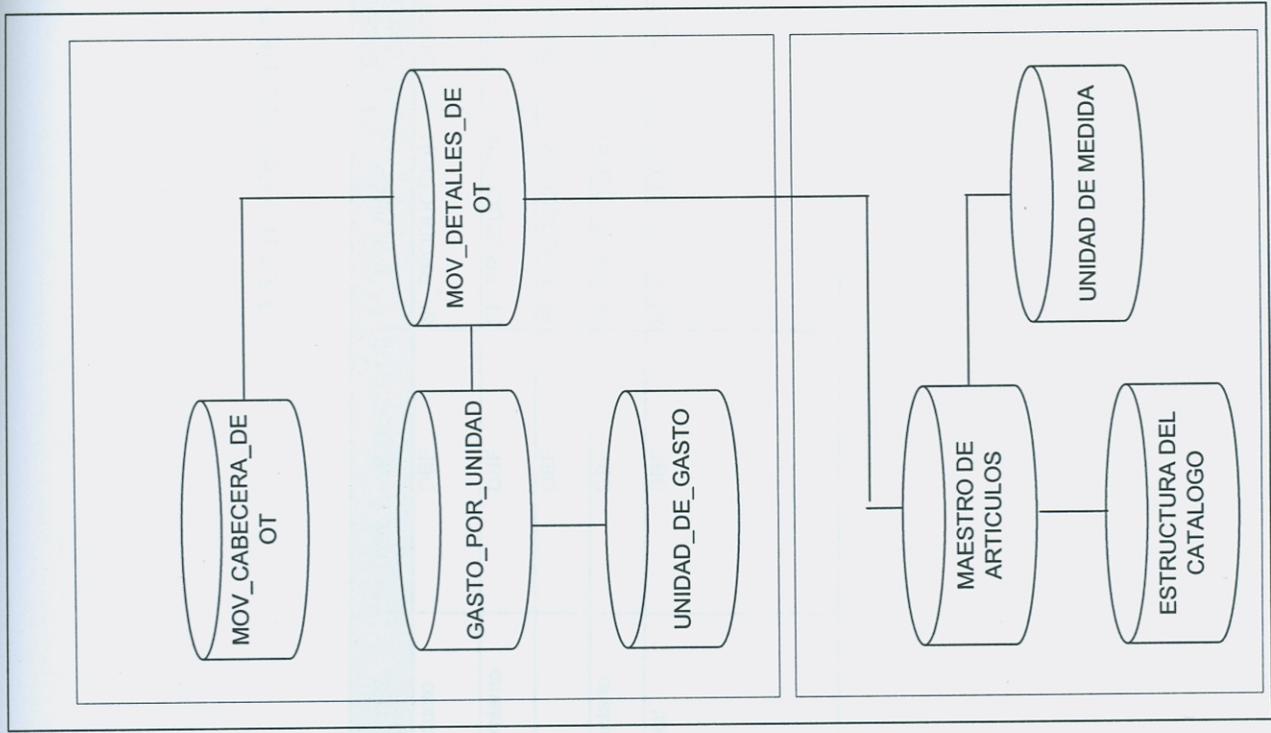


3.3.2.8 FUENTES DE DATOS PARA LOS INDICADORES POR VENTAS

| VARIABLES | TIPO DE BASE DE DATOS | SERVIDOR | BASE DE DATOS | SCRIPT | DESCRIPCION |
|-------------------|-----------------------|--------------|---------------|--|---|
| Sector Económico | DBF | SF_PRODUCION | COMERCIAL | SELECT cli_cnl01c FROM comercial->maestro_cliente SELECT arti_str01c | Selecciona los campos de cli_cnl01c correspondiente al Sector Económico de la entidad Cliente Selecciona los campos de arti_str01c correspondiente al Código de Clasificación de la entidad Artículo |
| Línea de Producto | DBF | SF_PRODUCION | LOGISTICA | FROM logistica->maestro_articulos SELECT otr_can01n FROM operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE | Selecciona el campo de cantidad por O/T para un documento no anulado |
| Cantidad | DBF | SF_PRODUCION | OPERACIONAL | SELECT ot_valor FROM operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE | Selecciona el campo de gasto para un documento no anulado |
| Valor por OT | DBF | SF_PRODUCION | OPERACIONAL | SELECT cab.num_01c, und.unidad_gasto, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, und.gasto_unidad, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli, operacional->gasto_unidad und WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c AND det.unidad_gasto= und.unidad_gasto | Con esta consulta es posible levantar la información por O/T desde la fuente de información de la base Operacional respecto del gasto por unidad de gasto |
| Gastos Totales | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c | Con esta consulta es posible levantar la información por O/T desde la fuente de información de la base Operacional respecto del costo pot detalle de O/T |
| Costo por OT | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c | Selecciona el campo de costo para un documento no anulado |
| Costos Totales | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c | Selecciona el campo de costo para un documento no anulado |

3.3.2.8 FUENTES DE DATOS PARA LOS INDICADORES POR VENTAS

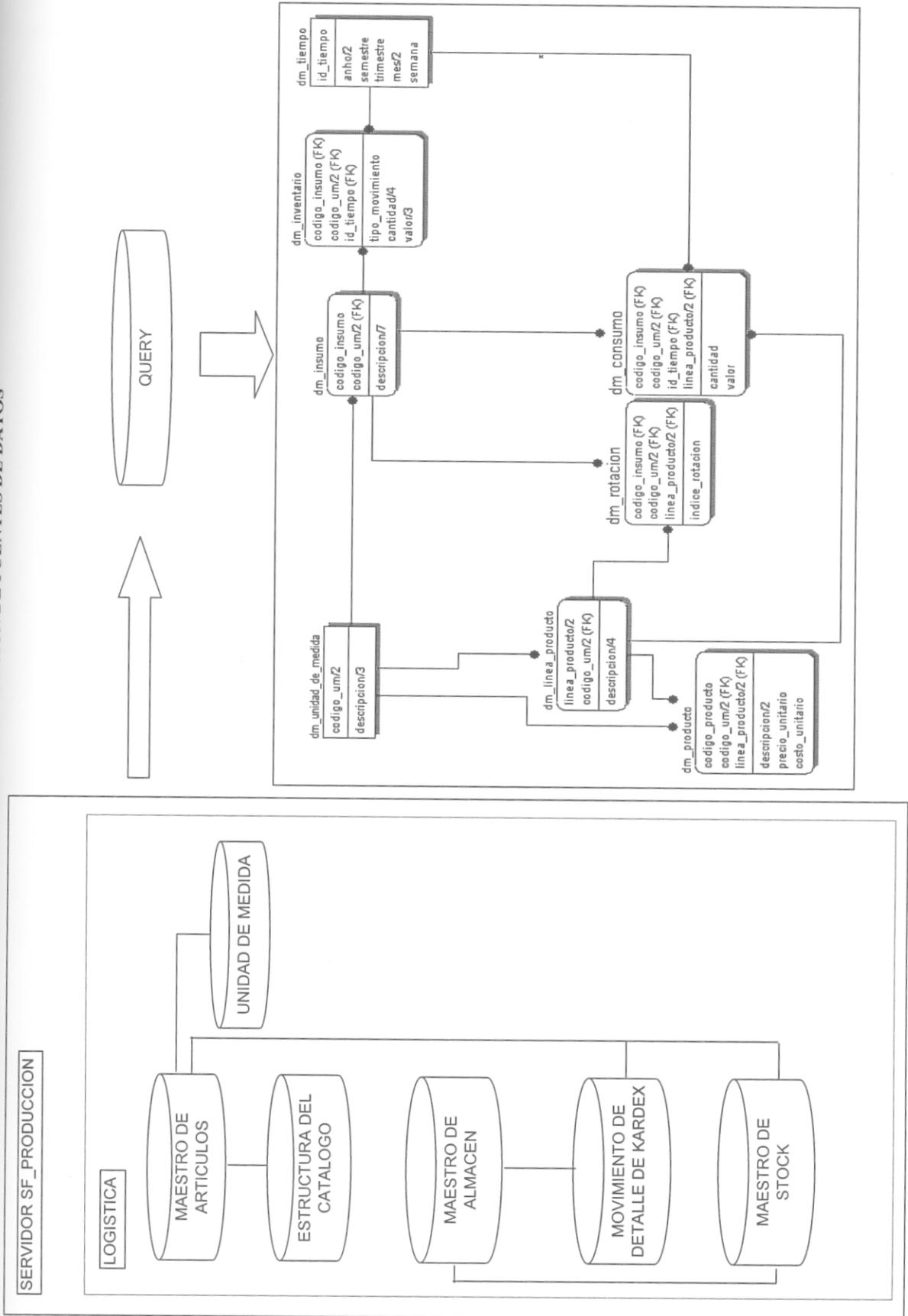
| VARIABLES | TIPO DE BASE DE DATOS | SERVIDOR | BASE DE DATOS | SCRIPT | DESCRIPCION |
|-------------------|-----------------------|--------------|---------------|--|---|
| Sector Económico | DBF | SF_PRODUCION | COMERCIAL | SELECT cli_cnl01c FROM comercial->maestro_cliente SELECT arti_str01c | Selecciona los campos de cli_cnl01c correspondiente al Sector Económico de la entidad Cliente Selecciona los campos de arti_str01c correspondiente al Código de Clasificación de la entidad Artículo |
| Línea de Producto | DBF | SF_PRODUCION | LOGISTICA | FROM logistica->maestro_articulos SELECT otr_can01n FROM operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE | Selecciona el campo de cantidad por O/T para un documento no anulado |
| Cantidad | DBF | SF_PRODUCION | OPERACIONAL | SELECT ot_valor FROM operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE | Selecciona el campo de gasto para un documento no anulado |
| Valor por OT | DBF | SF_PRODUCION | OPERACIONAL | SELECT cab.num_01c, und.unidad_gasto, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, und.gasto_unidad, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli, operacional->gasto_unidad und WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c AND det.unidad_gasto= und.unidad_gasto | Con esta consulta es posible levantar la información por O/T desde la fuente de información de la base Operacional respecto del gasto por unidad de gasto |
| Gastos Totales | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c AND det.unidad_gasto= und.unidad_gasto | Con esta consulta es posible levantar la información por O/T desde la fuente de información de la base Operacional respecto del costo pot detalle de O/T |
| Costo por OT | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c AND det.unidad_gasto= und.unidad_gasto | Selecciona el campo de costo para un documento no anulado |
| Costos Totales | DBF | SF_PRODUCION | OPERACIONAL | ot_costo operacional->mov_detalle_de_ot WHERE otr_ind011 is TRUE SELECT cab.num_01c, cab.fec_01f, cab.clie_cod01c, det.arti_cod01c, det.otr_can01n, det.otr_costo, cli.clie_cnl01c, art.arti_str01c, cab.otr_mnd01c, cab.otr_tca01n FROM operacional->mov_cabecera_de_ot cab, operacional->mov_detalle_de_ot det, logistica->maestro_articulos art, comercial->maestro_cliente cli WHERE cab.otr_ind011 is TRUE AND cab.otr_cod01c= det.otr_cod01c AND cab.otr_num01c= det.otr_num01c AND cab.clie_cod01c= cli.clie_cod01c AND det.arti_cod01c= art.arti_cod01c AND det.unidad_gasto= und.unidad_gasto | Selecciona el campo de costo para un documento no anulado |



3.3.2.10 FUENTES DE DATOS PARA LOS INDICADORES POR LOGISTICA

| VARIABLES | TIPO DE BASE DE DATOS | SERVIDOR | BASE DE DATOS | SCRIPT | DESCRIPCION |
|---------------------|-----------------------|--------------|---------------|--|---|
| Línea de Producto | DBF | SF_PRODUCION | LOGISTICA | SELECT arti_str01c FROM logistica->maestro_articulos | Selecciona arti_str01c correspondiente al Código de Clasificación de la entidad Artículo |
| Tipo de Movimiento | DBF | SF_PRODUCION | LOGISTICA | SELECT kard_cod01c FROM logistica->mov_detalle_de_kardex | Selecciona el tipo de movimiento: Entrada o Salida |
| Cantidad | DBF | SF_PRODUCION | LOGISTICA | SELECT kard_can01n FROM logistica->mov_detalle_de_kardex | Selecciona la cantidad de artículos que entraron o salieron en la transacción |
| Valor de Inventario | DBF | SF_PRODUCION | LOGISTICA | SELECT kard_val01n FROM logistica->mov_detalle_de_kardex | Selecciona el valor de los productos que entraron o salieron de la producción |
| Inventario Total | DBF | SF_PRODUCION | LOGISTICA | SELECT art.arti_cod01c, art.arti_str01c, mov.kard_cod01c, mov.kard_can01c, mov.kard_val01n FROM logistica->mov_detalle_de_kardex mov, logistica->maestro_de_articulos art WHERE mov.arti_cod01c= art.arti_cod01c | Con esta consulta es posible levantar la información por movimiento de inventarios desde la fuente de información de la base Logistical |

3.3.2.11 DIAGRAMA DE RELACION DE FUENTES DE DATOS



3.3.2.12 DETERMINACION DE REQUERIMIENTOS DE TRANSFORMACION DE DATA

| INDICADOR | BdD ORIGEN | TABLAS | DATO ORIGEN | TABLA DESTINO | DATOS DESTINO | FUNCION |
|-----------|------------|---|----------------------------|----------------------------------|----------------------------------|--|
| VENTAS | Comercial | mov_cabecera_de_facturacion mov_detalle_de_facturacion | fact_num01c | dm_factura_cab dm_factura_det | numero_serie numero_documento | SELECT LEFT(fact_num01c,3) FROM mov_cabecera_de_facturacion SELECT RIGHT(fact_num01c,6) FROM mov_cabecera_de_facturacion Esta función se implementa debido a que el número de documento origen es compuesto de la forma 999-999999, donde los tres primeros son el origen y los seis restantes son el número del documento |
| VENTAS | Comercial | mov_cabecera_de_facturacion | fact_mnd01c fact_tcm01n | dm_factura_cab dm_factura_det | valor_venta | SELECT fact_val05n * fact_tcm01n FROM mov_detalle_de_facturacion WHERE fact_mnd01c = "002" Esta función se implementa para efectuar el cambio para el valor de venta |
| LOGISTICA | Logística | mov_detalle_de_kardex | kard_cod01c | dm_inventario | tipo_movimiento | SELECT LEFT(kard_cod01c,2) FROM mov_detalle_de_kardex Esta función se implementará de modo que se relacionen los movimientos.- 001: Entrada 002: Salida |
| LOGISTICA | Logística | maestro_articulos maestro_estructura | arti_str01c | dm_producto | linea_producto | SELECT LEFT(kard_cod01c,3) FROM mov_detalle_de_kardex Esta función se implementa debido a que nos interesa los tres primeros dígitos que determinan la estructura del catálogo; los demás dígitos de la tabla origen se refieren a sub-clasificaciones los cuales tiene otras finalidades |

3.3.2.13 EXTRACCION DE LA DATA PARA EL INDICADOR POR VENTAS

VENTAS TOTALES

- 1 Leer la fecha inicial y final de proceso en :Fecha1 y :Fecha2
- 2 Conectarse a la Base Datos Comercial
- 3 Conectarse a la Base Datos DataMart
- 4 Crear el Cursor Factura_Total con el siguiente query.-

```
SELECT cab.fact_cod01c, cab.fact_num01c,
cab.fact_fec01f, cab.clie_cod01c, det.art_umd01c
det.arti_cod01c, det.fact_can01n, det.fact_val05n,
cli.clie_cnl01c, art.arti_str01c,
cab.fact_mnd01c, cab.fact_tca01n
INTO :Tipo_documento,
:Numero_documento,
:Fecha_documento,
:Codigo_cliente,
:Unidad_medida,
:Codigo_articulo,
:Cantidad_facturada,
:Valor_venta,
:Sector_economico,
:Codigo_estructura,
:Tipo_moneda,
:Valor_cambio
FROM comercial->mov_cabecera_de_facturacion cab,
comercial->mov_detalles_de_facturacion det,
logistica->maestro_articulos art,
comercial->maestro_cliente clie
WHERE cab.fact_ind011 is TRUE
AND cab.fact_cod01c= det.fact_cod01c
AND cab.fact_num01c= det.fact_num01c
AND cab.clie_cod01c= cli.clie_cod01c
AND det.arti_cod01c= art.arti_cod01c
AND cab.fact_fec01f BETWEEN :Fecha1 AND :Fecha2
```

- 5 Abrir el cursor Factura_Total
- 6 Mientras no sea Fin del cursor Factura_Total, aplicar pasos del 6 al 13, sino ir al 14
- 7 :Serie_documento = LEFT(:Numero_documento,3)
- 8 :Numero_documento = RIGHT(:Numero_documento,6)
- 9 :Codigo_estructura = LEFT(:Codigo_estructura,3)
- 10 Si :Tipo_moneda = "002", está en soles, entonces :Valor_venta = :Valor_venta/ :Valor_cambio
- 11 Ejecutar el siguiente Query

```
INSERT INTO dm_factura_det
( tipo_documento,
numero_serie,
numero_documento,
codigo_producto,
codigo_um,
linea_producto,
codigo_cliente,
sector_economico,
cantidad,
valor_venta)
VALUES
( :Tipo_documento,
:Serie_documento,
:Numero_documento,
:Codigo_articulo,
:Unidad_medida,
:Codigo_estructura,
:Codigo_cliente,
:Sector_economico,
:Cantidad_facturada,
:Valor_venta
)
```

- 12 Si la cabecera de la factura no esta creada, ejecutar el siguiente Query.-

```
INSERT INTO dm_factura_cab
( tipo_documento,
  numero_serie,
  numero_documento,
  codigo_cliente,
  sector_economico,
  valor_venta,
  fecha_venta)
VALUES
( :Tipo_documento,
  :Serie_documento,
  :Numero_documento,
  :Codigo_cliente,
  :Sector_economico,
  :Valor_venta,
  :Fecha_documento,
)
```

- 13 Retornar al paso 6
14 Procesar la venta total por sector economico y linea de producto

```
INSERT INTO dm_ventas_totales
SELECT a.sector_economico,
  b.linea_producto,
  sp_fecha(a.fecha_venta),
  sum(b.cantidad),
  sum(b.valor_venta),
FROM dm_factura_cab a,
  dm_factura_det b
WHERE a.fecha_venta BETWEEN :Fecha1 AND :Fecha2
AND a.tipo_documento= b.tipo_documento
AND a.serie_documento= b.serie_documento
AND a.numero_documento= b.numero_documento
GROUP BY 1,2,3
```

- 15 Procesar la venta total por cliente especifico y linea de producto

```
INSERT INTO dm_ventas_totales
SELECT a.codigo_cliente,
  b.linea_producto,
  sp_fecha(a.fecha_venta),
  sum(b.cantidad),
  sum(b.valor_venta),
FROM dm_factura_cab a,
  dm_factura_det b
WHERE a.fecha_venta BETWEEN :Fecha1 AND :Fecha2
AND a.tipo_documento= b.tipo_documento
AND a.serie_documento= b.serie_documento
AND a.numero_documento= b.numero_documento
GROUP BY 1,2,3
```

- 16 Actualizar la tabla de Sector Economico

```
INSERT INTO dm_sector_economico
SELECT clie_cn101c, cnal_vta01c
FROM tab_canal_venta
WHERE NOT EXISTS (SELECT sect.sector_economico
FROM dm_sector_economico sect
WHERE sect.sector_economico=
  tab_canal_venta.clie_cn101c)
```

17 Actualizar la tabla de Cliente

```
INSERT INTO dm_cliente
SELECT clie_cod01c,
      clie_cnl01c,
      clie_ruc01n,
      clie_tel01c,
      clie_dir01c,
      clie_nom01c
FROM maestro_cliente
WHERE NO EXIST (SELECT clie.codigo_cliente
FROM dm_cliente clie
WHERE clie.codigo_cliente=
maestro_cliente.clie_cod01c)
```

18 Actualizar la tabla de Productos

```
INSERT INTO dm_producto
SELECT arti_cod01c,
      arti_umd01c,
      arti_str01c,
      arti_des01c,
      arti_val01n,
      arti_cst01n
FROM maestro_articulos
WHERE NO EXIST (SELECT arti.codigo_producto
FROM dm_producto
WHERE arti.codigo_producto=
maestro_articulos.arti_cod01c)
```

19 Actualizar la linea de productos

```
INSERT INTO dm_linea_producto
SELECT LEFT(stru_cod01c,3),
      stru_des01c
FROM maestro_estructura
WHERE NO EXIST (SELECT stru.linea_producto
FROM dm_linea_producto
WHERE stru.linea_producto=
LEFT(maestro_estructura.stru_cod01c,3))
```

20 Cerra Base de Datos

21 Fin de Proceso

3.3.2.14 EXTRACCION DE LA DATA PARA EL INDICADOR POR GASTOS Y UTILIDAD

VENTAS TOTALES

- 1 Leer la fecha inicial y final de proceso en :Fecha1 y :Fecha2
- 2 Conectarse a la Base Datos Operacional
- 3 Conectarse a la Base Datos DataMart
- 4 Crear el Cursor OT_Total con el siguiente query.-

```
SELECT cab.otr_num01c,  
cab.otr_fec01f, cab.clie_cod01c, det.art_umd01c  
det.arti_cod01c, det.otr_can01n, det.otr_val05n,  
det.otr_val06, cli.clie_cn101c, art.arti_str01c,  
cab.otr_mnd01c, cab.otr_tca01n  
INTO :Numero_documento,  
:Fecha_documento,  
:Codigo_cliente,  
:Unidad_medida,  
:Codigo_articulo,  
:Cantidad_ot,  
:Valor_ot,  
:Costo_ot,  
:Sector_economico,  
:Codigo_estructura,  
:Tipo_moneda,  
:Valor_cambio  
FROM operacional->mov_cabecera_de_ot cab,  
operacional->mov_detalle_de_ot det,  
logistica->maestro_articulos art,  
comercial->maestro_cliente clie  
WHERE cab.otr_ind01l is TRUE  
AND cab.otr_num01c= det.otr_num01c  
AND cab.clie_cod01c= clie.clie_cod01c  
AND det.arti_cod01c= art.arti_cod01c  
AND cab.otr_fec01f BETWEEN :Fecha1 AND :Fecha2
```

- 5 Abrir el cursor OT_Total
- 6 Mientras no sea Fin del cursor OT_Total, aplicar pasos del 7 al 11, sino ir al 12
- 7 :Codigo_estructura = LEFT(:Codigo_estructura,3)
- 8 Si :Tipo_moneda = "002", está en soles, entonces :Valor_ot = :Valor_ot/ :Valor_cambio
- 9 Ejecutar el siguiente Query

```
INSERT INTO dm_ot_det  
( numero_ot,  
codigo_producto,  
codigo_um,  
linea_producto,  
codigo_cliente,  
sector_economico,  
cantidad,  
Valor_ot,  
Costo_ot)  
VALUES  
( :Numero_documento,  
:Codigo_articulo,  
:Unidad_medida,  
:Codigo_estructura,  
:Codigo_cliente,  
:Sector_economico,  
:Cantidad_ot,  
:Valor_ot,  
:Costo_ot  
)
```

10

Si la cabecera de la O/T no esta creada, ejecutar el siguiente Query.-

```

INSERT INTO dm_ot_cab
( numero_documento,
  codigo_cliente,
  sector_economico,
  Valor_ot,
  Costo_ot,
  fecha_ot)
VALUES
( :Numero_documento,
  :Codigo_cliente,
  :Sector_economico,
  :Valor_ot,
  :Costo_ot,
  :Fecha_documento,
)

```

11

Retornar al paso 6

12

Procesar el OT Gasto total

```

INSERT INTO dm_gasto_total
SELECT a.unidad_gasto,
  sp_fecha(a.fecha_ot),
  sum(b.Valor_ot)
FROM dm_ot_cab a,
  dm_ot_det b
WHERE a.fecha_ot BETWEEN :Fecha1 AND :Fecha2
AND a.numero_ot= b.numero_ot
GROUP BY 1

```

13

Procesar el OT Utilidad total

```

INSERT INTO dm_utilidad_total
SELECT a.sector_economico, a.linea_producto,
  sp_fecha(a.fecha_ot),
  sum(b.Valor_ot - b.Costo_ot)
FROM dm_ot_cab a,
  dm_ot_det b
WHERE a.fecha_ot BETWEEN :Fecha1 AND :Fecha2
AND a.numero_ot= b.numero_ot
GROUP BY 1

```

14

Actualizar la tabla de Sector Economico

```

INSERT INTO dm_sector_economico
SELECT clie_cnl01c, cnal_vta01c
FROM tab_canal_venta
WHERE NO EXIST (SELECT sect.sector_economico
  FROM dm_sector_economico sect
  WHERE sect.sector_economico=
  tab_canal_venta.clie_cnl01c)

```

15

Actualizar la tabla de Cliente

```

INSERT INTO dm_cliente
SELECT clie_cod01c,
  clie_cnl01c,
  clie_ruc01n,
  clie_tel01c,
  clie_dir01c,
  clie_nom01c
FROM maestro_cliente
WHERE NO EXIST (SELECT clie.codigo_cliente
  FROM dm_cliente clie
  WHERE clie.codigo_cliente=
  maestro_cliente.clie_cod01c)

```

16 Actualizar la tabla de Productos

```
INSERT INTO dm_producto
SELECT arti_cod01c,
      arti_umd01c,
      arti_str01c,
      arti_des01c,
      arti_val01n,
      arti_cst01n
FROM maestro_articulos
WHERE NO EXIST (SELECT arti.codigo_producto
                FROM dm_producto
                WHERE arti.codigo_producto=
                  maestro_articulos.arti_cod01c)
```

17 Actualizar la linea de productos

```
INSERT INTO dm_linea_producto
SELECT LEFT(stru_cod01c,3),
      stru_des01c
FROM maestro_estructura
WHERE NO EXIST (SELECT stru.linea_producto
                FROM dm_linea_producto
                WHERE stru.linea_producto=
                  LEFT(maestro_estructura.stru_cod01c,3))
```

18 Cerra Base de Datos

19 Fin de Proceso

3.2.15 EXTRACCION DE LA DATA PARA EL INDICADOR DE LOGISTICA

Leer la fecha inicial y final de proceso en :Fecha1 y :Fecha2
Conectarse a la Base Datos Logística
Conectarse a la Base Datos DataMart
Crear el Cursor Factura_Total con el siguiente query.-

```
SELECT art.arti_cod01c,  
       art.arti_umd01c,  
       art.arti_str01c,  
       mov.kard_cod01c,  
       mov.kard_can01c,  
       mov.kard_val01n  
INTO :codigo_articulo,  
     :unidad_medida,  
     :linea_producto,  
     :tipo_movimiento,  
     :cantidad,  
     :valor  
FROM logistica->mov_detalle_de_kardex mov,  
     logistica->maestro_de_articulos art  
WHERE mov.arti_cod01c= art.arti_cod01c  
AND mov.kard_fec01f BETWEEN :Fecha1 AND :Fecha2
```

- 5 Abrir el cursor Factura_Total
- 6 Mientras no sea Fin del cursor Factura_Total, aplicar pasos del 7 al 9
- 7 :Tipo_movimiento = LEFT(:Tipo_movimiento, 3)
- 8 Insetar el registro por movimiento de Kardex

```
INSERT INTO dm_inventario  
( codigo_insumo,  
  codigo_um,  
  tipo_movimiento,  
  cantidad,  
  valor)  
VALUES (  
  :codigo_articulo,  
  :unidad_medida,  
  :tipo_movimiento,  
  :cantidad,  
  :valor)
```

- 9 Retornar al paso 6
- 10 Actualizar la tabla de Productos

```
INSERT INTO dm_producto  
SELECT arti_cod01c,  
       arti_umd01c,  
       arti_str01c,  
       arti_des01c,  
       arti_val01n,  
       arti_cst01n  
FROM maestro_articulos  
WHERE NO EXIST (SELECT arti.codigo_producto  
                FROM dm_producto  
                WHERE arti.codigo_producto=  
                  maestro_articulos.arti_cod01c)
```

- 11 Actualizar la linea de productos

```
INSERT INTO dm_linea_producto  
SELECT LEFT(stru_cod01c,3),  
       stru_des01c  
FROM maestro_estructura  
WHERE NO EXIST (SELECT stru.linea_producto  
                FROM dm_linea_producto  
                WHERE stru.linea_producto=  
                  LEFT(maestro_estructura.stru_cod01c,3))
```

12 Actualizar la base de insumos

```
INSERT INTO dm_insumo
SELECT arti_cod01c,
       arti_umd01c,
       arti_des01c
FROM maestro_articulos
WHERE NO EXIST (SELECT arti.codigo_insumo
                FROM dm_insumo
                WHERE arti.codigo_insumo=
                  maestro_articulos.arti_cod01c)
AND maestro_articulos.tipo='INS'
```

13 Actualizar la base dm_consumo

```
INSERT INTO dm_consumo
SELECT a.codigo_insumo,
       b.linea_producto,
       a.cantidad,
       a.valor,
       sp_fecha(a.fecha_mov),
FROM dm_inventario a,
     dm_producto b
WHERE a.codigo_producto= b.codigo_producto
AND a.tipo_movimiento = 'SALDOFINAL' AND a.fecha_mov BETWEEN :Fecha1 and :Fecha2
```

14 Actualizar la dm_rotacion

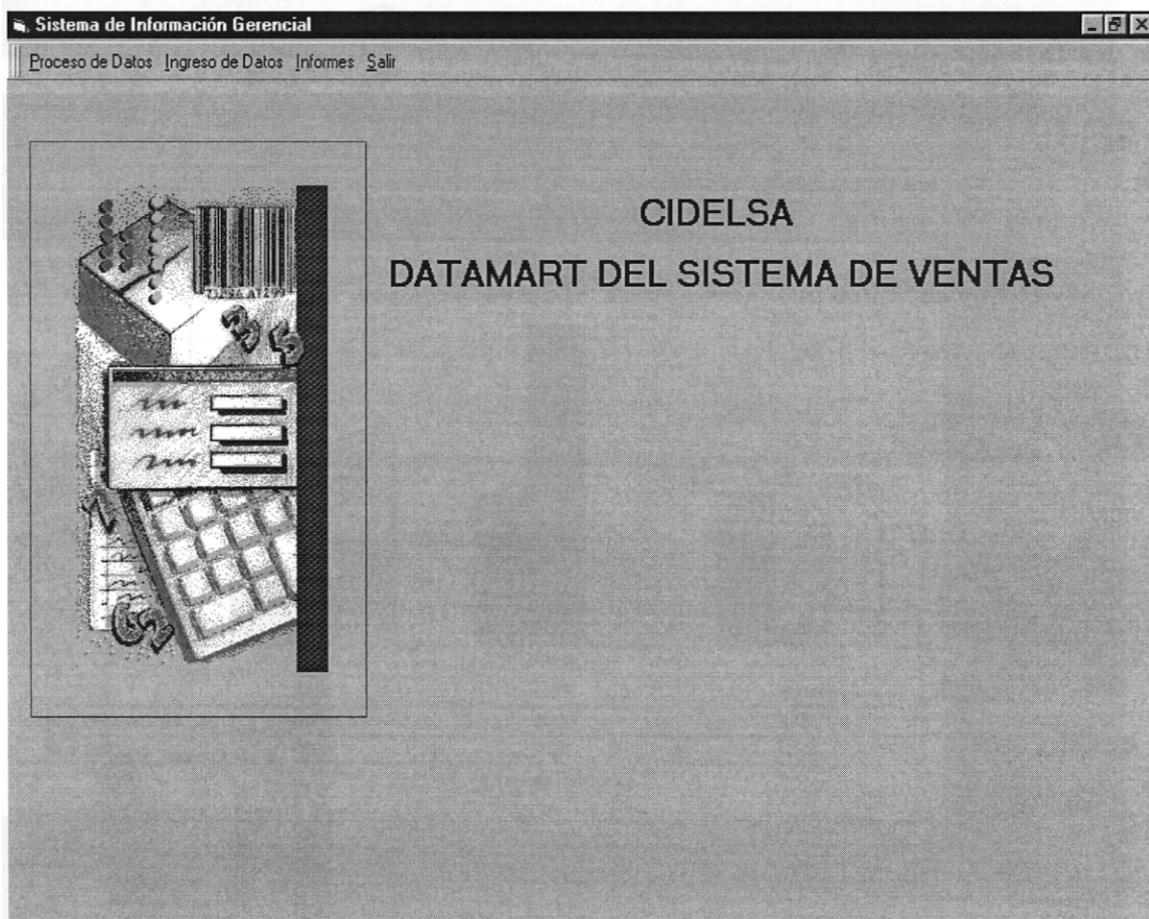
```
INSERT INTO dm_rotacion
SELECT sp_fecha(a.fecha_mov),
       sum(a.cantidad)/((fecha2-fecha1)/30)
FROM dm_inventario a
AND (a.tipo_movimiento = 'ENTRADA'
     OR a.tipo_movimiento = 'SALIDA')
AND a.fecha_mov BETWEEN :Fecha1 and :Fecha2
```

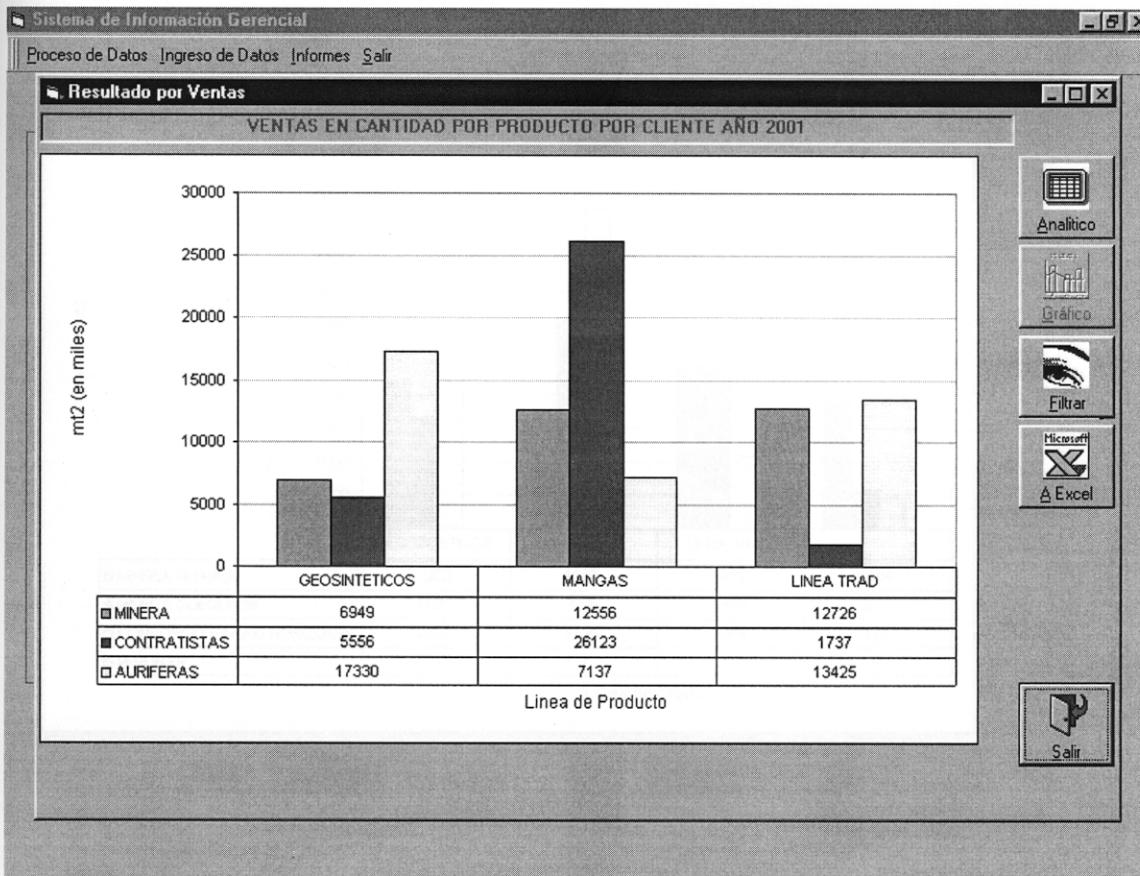
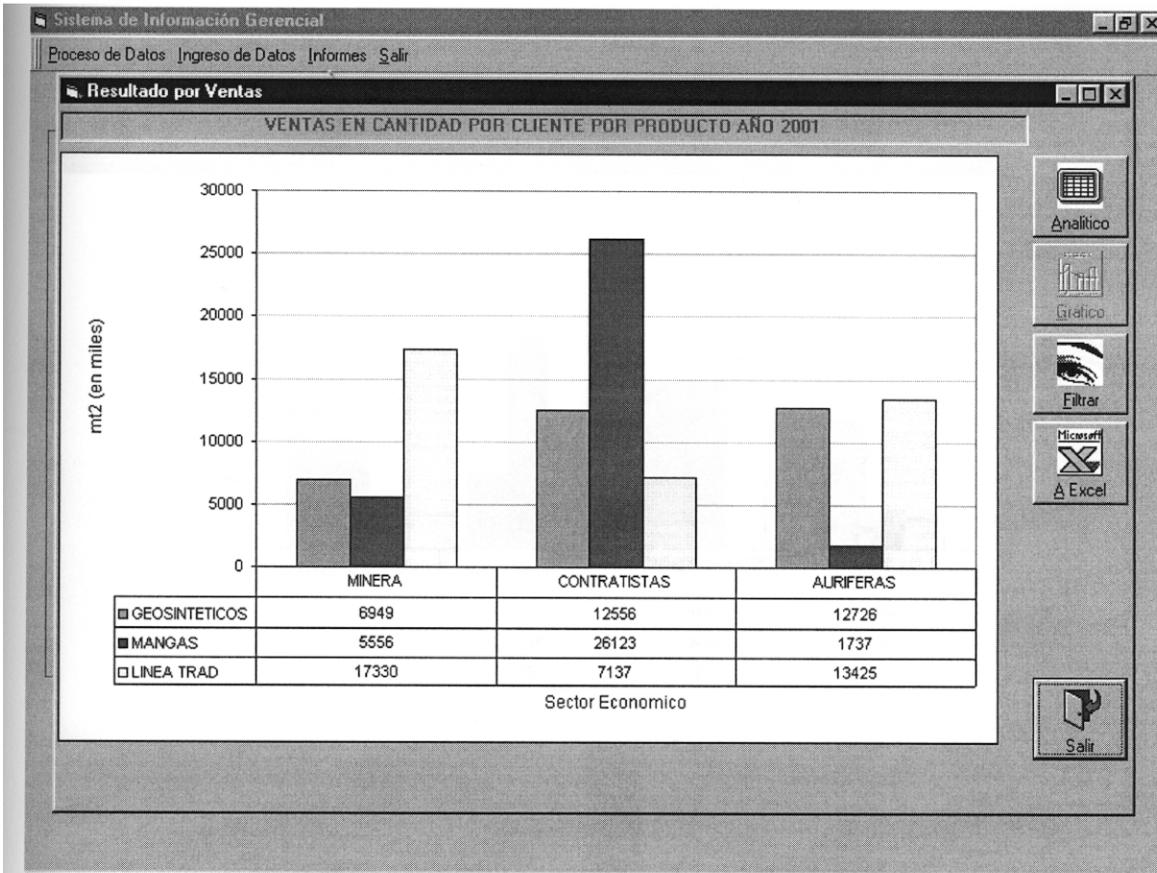
15 Cerrar las base de datos

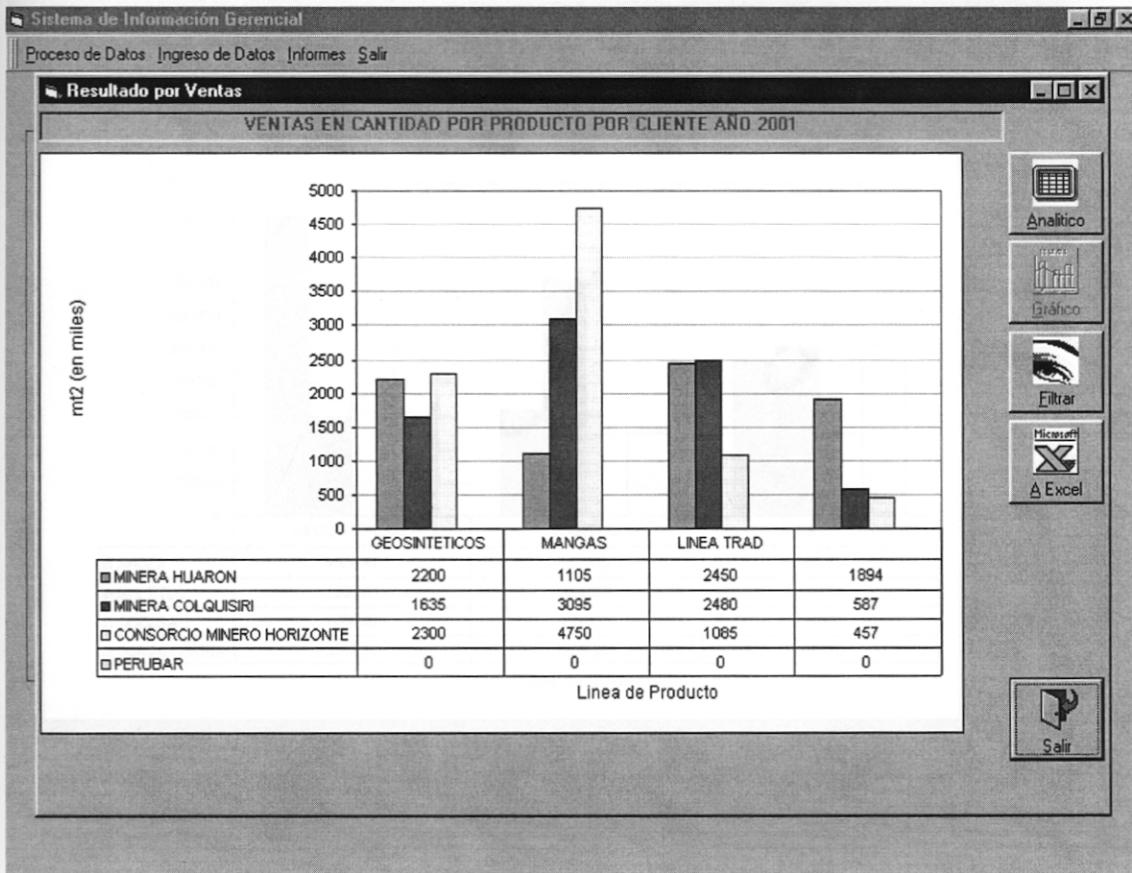
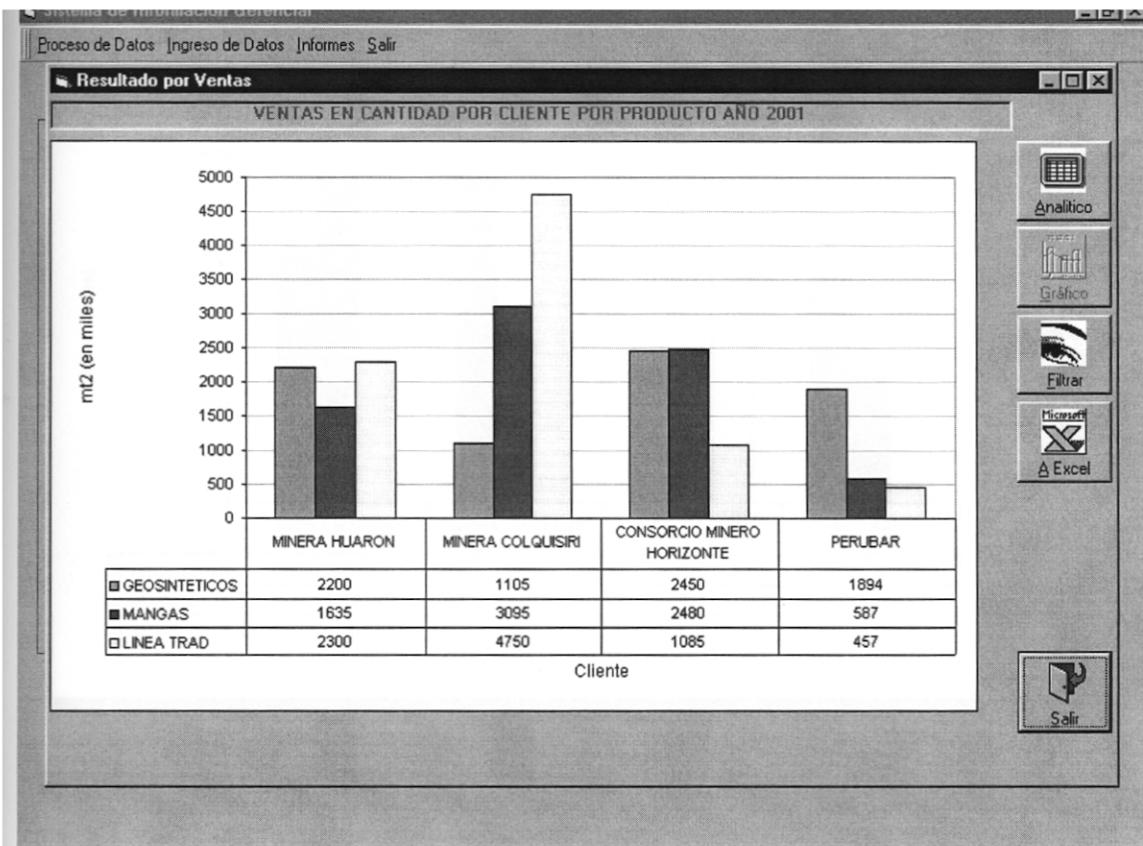
16 Fin de los procesos

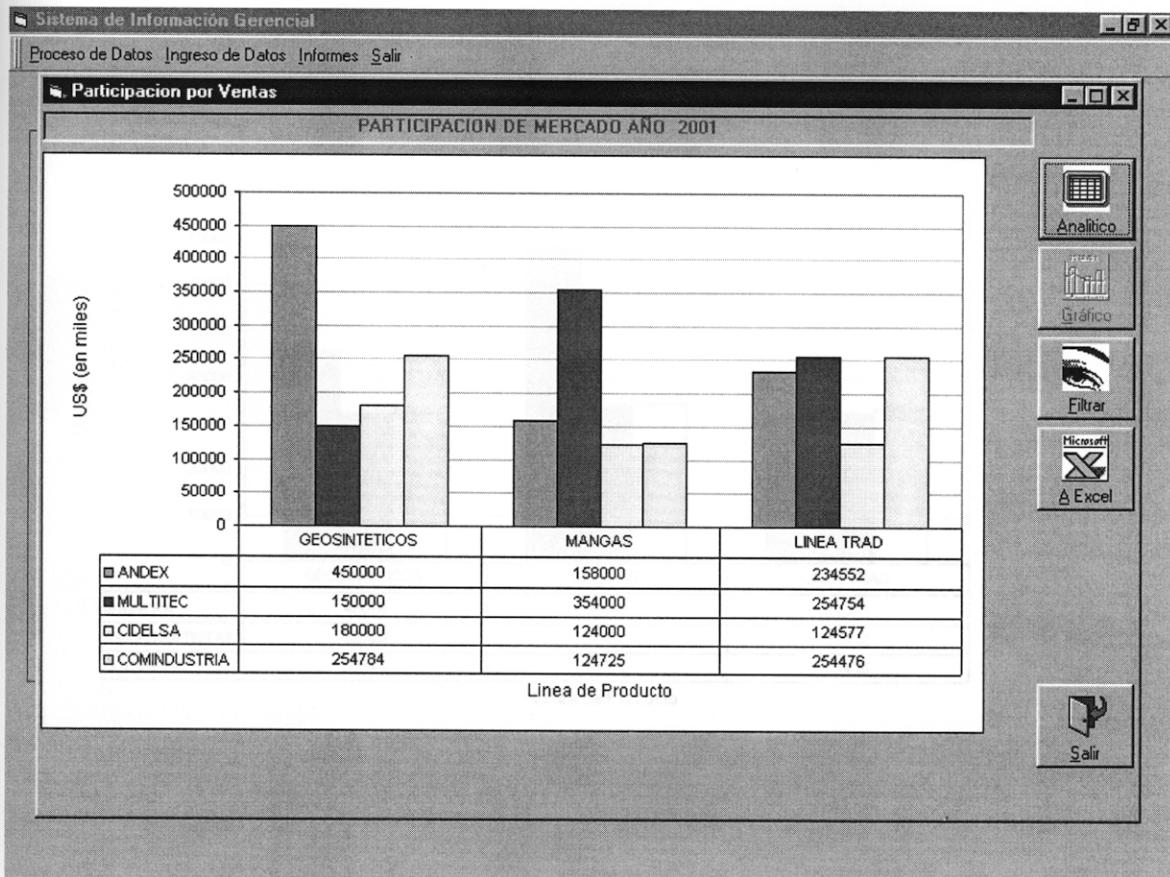
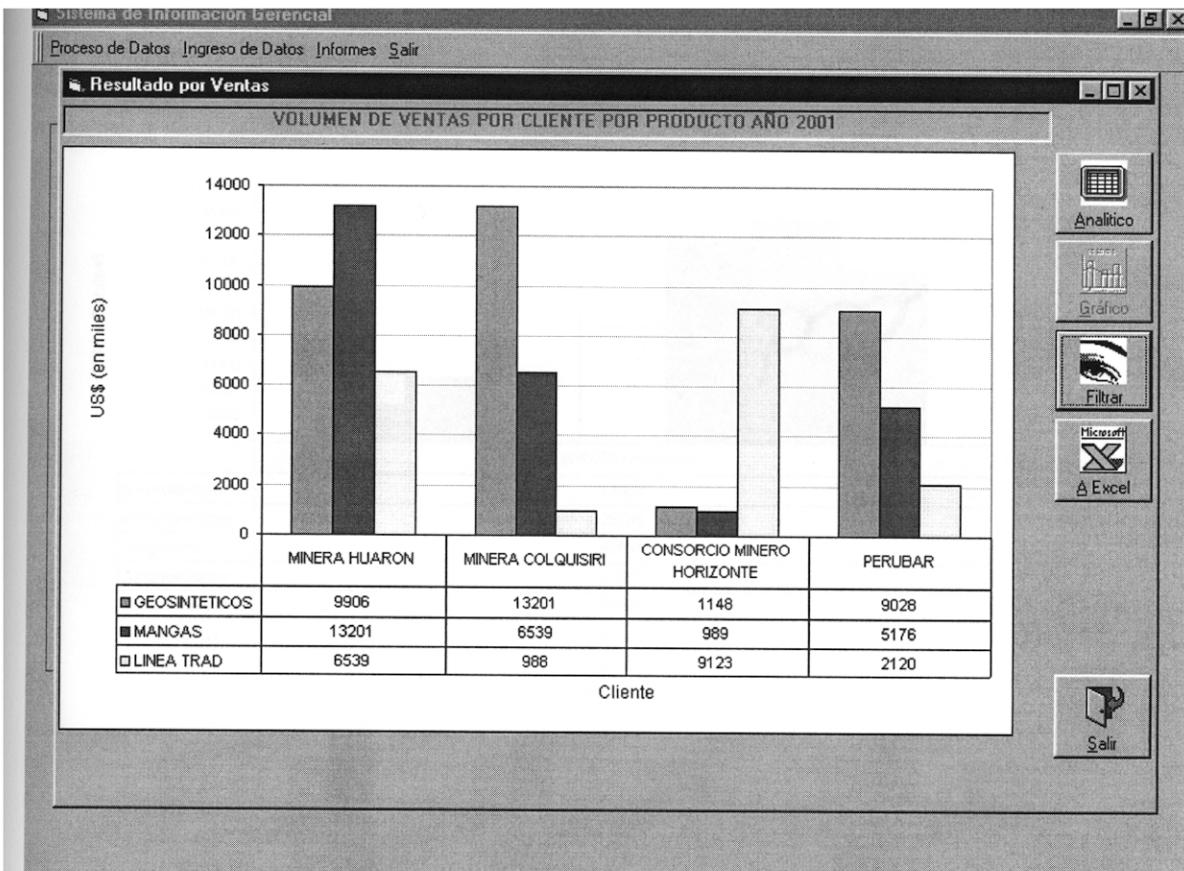
3.3.3 Proceso Front End

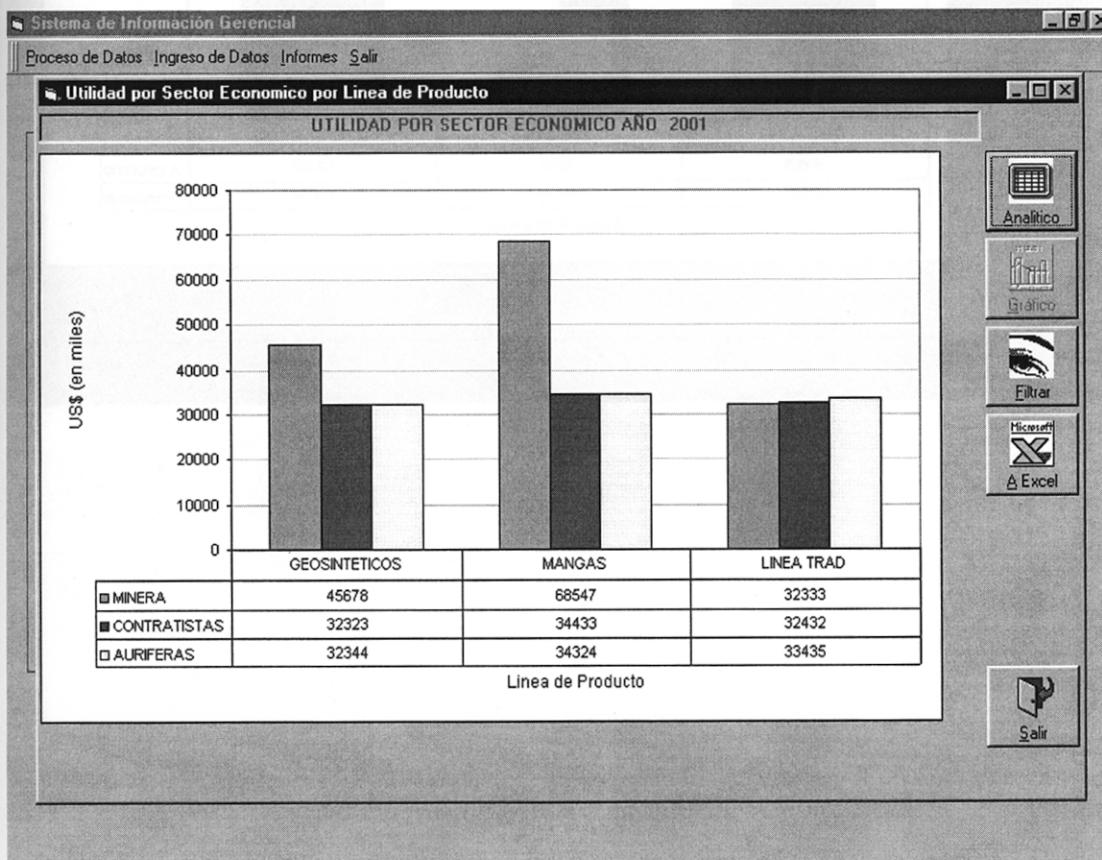
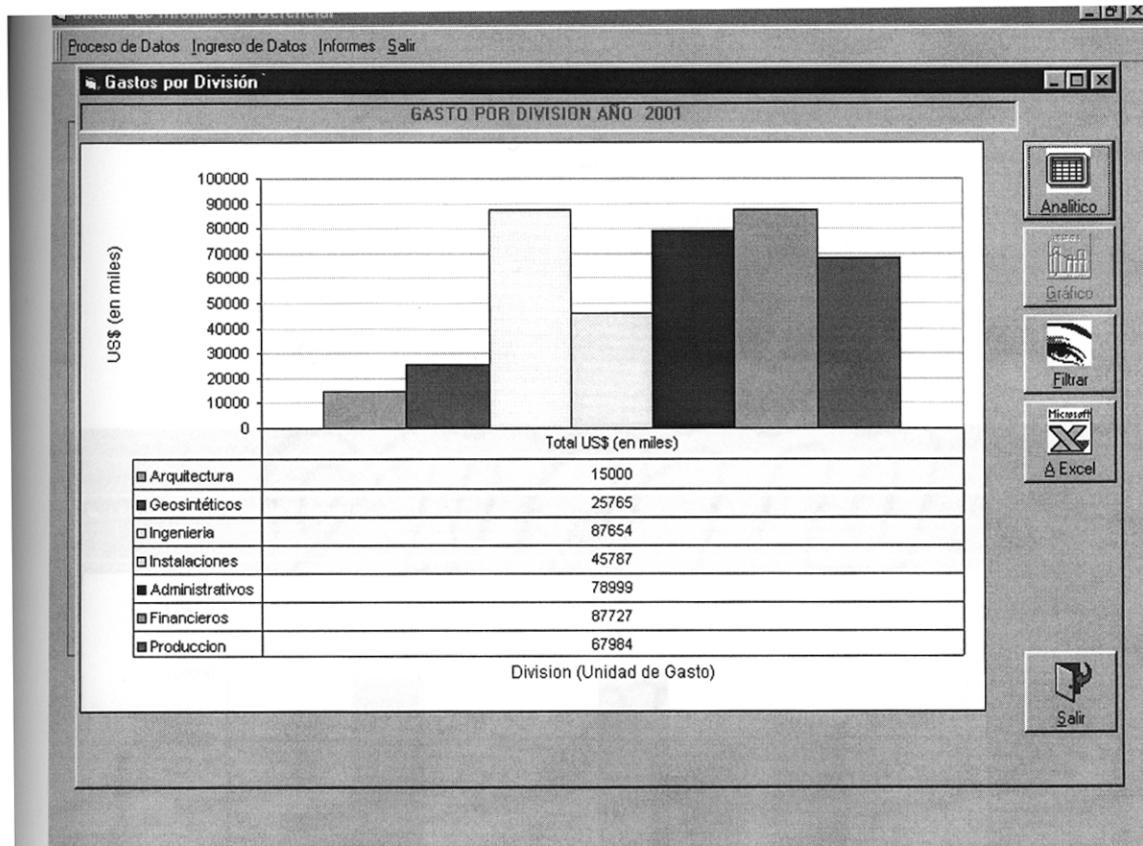
En este proceso desarrollaremos todas las pantallas por la cuál los usuarios del área comercial obtendrán información del Datamart. Con estas pantallas los usuarios tendrán la información definida en el punto 3.4 Acuerdos del Proyecto en la parte de Requerimientos Específicos del Negocio.

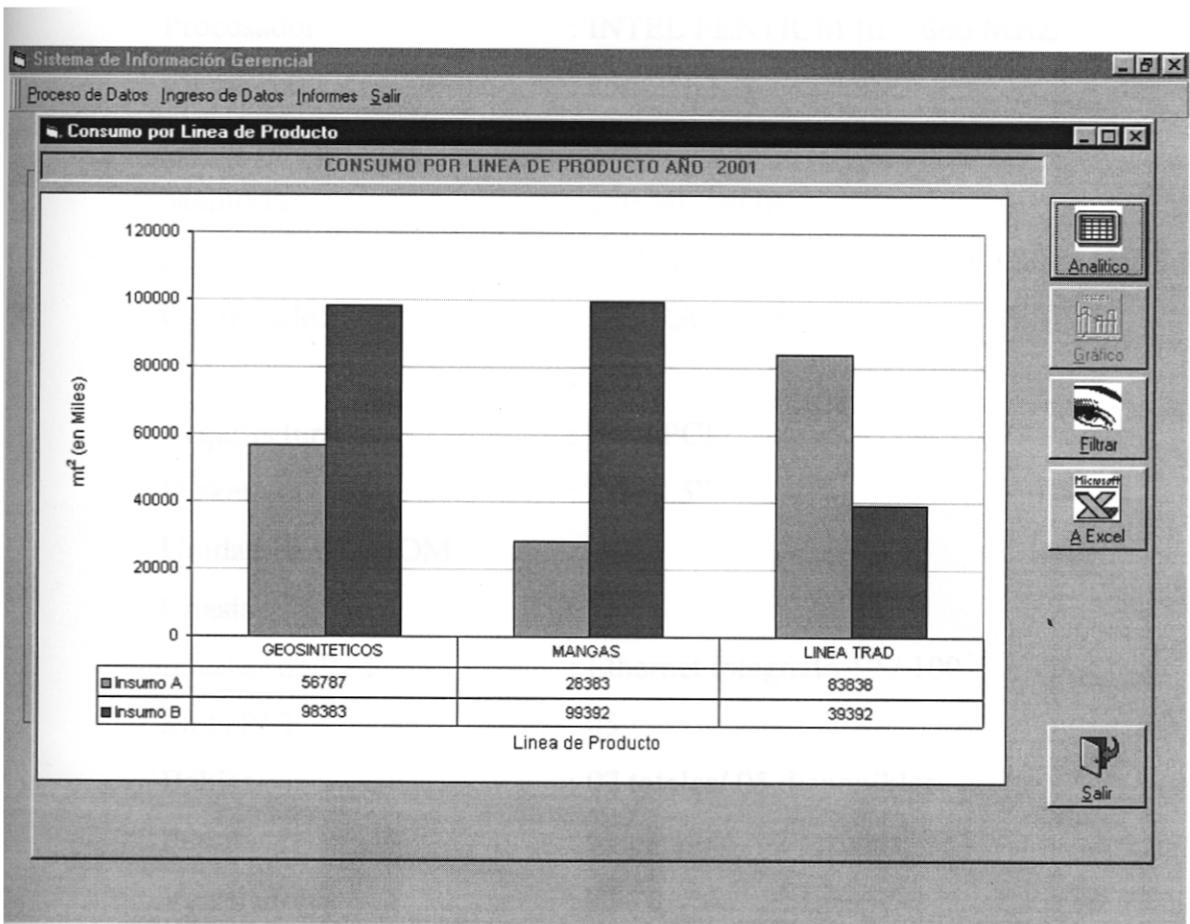












Para la parte de

Extracción, Transformación y Carga

Garantía

Microsoft SQL Server

Microsoft

Microsoft SQL Server

Microsoft

Microsoft

Microsoft

Microsoft SQL Server

Para la parte de Herramientas de Administración de Datos de Microsoft SQL Server

Extracción, Transformación y Carga, Agentes y Clusters de Microsoft SQL Server

Data Mart Suite, herramientas integradas para el desarrollo de DataMarts

construir DataMarts hoy.

3.3.4 Proceso de establecimiento de plataforma tecnológica

Se adquirirá un servidor con las siguientes características, para el trabajo del datamart:

Servidor IBM Netfinity-Serie X 200 MOD 8645-3AX

| | |
|-------------------|--|
| Procesador | : INTEL PENTIUM III – 866 MHZ |
| Cache Ext.L2 | : 256/256 KB |
| Cant.Procesadores | : 1 / 2 |
| Memoria | : 256 Mb./ 4GB |
| Disco Duro | : 18.2 GB10,000 rpm Ultra 160 SCSI Hot Swap |
| Controlador | : ULTRA 160 SCSI |
| Video | : 8 MB |
| Arquitectura | : BUS PCI |
| Diskettera | : 1 de 3.5” |
| Unidad de CD-ROM | : 48X |
| Chasis | : Torre |
| Interfaz de Red | : Ethernet integrado 10 / 100 |
| Slots PCI | : 05 |
| Bahías | : 07 totales/ 05 disponibles |
| Puertos | : 01 paralelo, 02 seriales, 02 USB |
| Ventiladores | : 03 |
| Fuente de poder | : 330 Watts, 110 ó 220v con switch automático |
| Teclado | : IBM |
| Software | : Server Guide, Norton Antivirus,IBM director. |
| Garantía | : 03 años |
| Monitor | : IBM SVGA 15” |

Para la parte de Herramientas de Administración de Base de Datos, Extracción, Transformación y Carga, reportes y Consultas usaremos el Oracle Data Mart Suite, herramienta que proporciona la más completa solución para construir DataMarts hoy.

Incluye el software de base de datos, el diseño, el movimiento de la data, las herramientas de transformación, documentación, y entrenamiento para una rápida y fácil implementación. En un paquete, Oracle DataMart Suite proporciona todo el software necesario para construir un DataMart, incluyendo: Oracle Data Mart Designer, una herramienta de fácil uso para diseñar un DataMart; Oracle DataMart Builder, una herramienta gráfica que extrae y transforma la data de los sistemas operacionales; Oracle8 Enterprise Edition, una base de datos escalable de alto performance; Oracle Web Application Server, para acceso a internet y el Oracle Discoverer, un generador de consultas, reportes y herramientas de análisis.

CONCLUSIONES Y RECOMENDACIONES

1. CONCLUSIONES

- La metodología de desarrollo propuesta nos brinda resultados inmediatos, en la implementación de Data Mart, además de comprometer a los usuarios, el personal de TI y los auspiciadores del proyecto, La definición de los Acuerdos del Proyecto es el documento donde se plasma dicho compromiso.
- La implementación de un Datamart Comercial en la empresa Cidelsa hará que consiga obtener ventaja competitiva, ya que la implementación de este sistema de apoyo a las decisiones es lo recomendado por el análisis estratégico realizado en el desarrollo del prototipo. Creemos que no vale la pena la inversión de tiempo y dinero si la implementación de un Datamart no coadyuva a dar ventaja competitiva a una organización.
- El modelo de Arquitectura de Referencia del Datawarehouse es un concepto muy valioso ya que nos ayuda a evaluar las herramientas disponibles, las herramientas necesarias, las herramientas que el mercado nos ofrece al momento de Diseñar un Datawarehouse.
- El modelo Starnet es una herramienta muy importante que nos ayuda a definir lo que el usuario realmente desea del Data Mart, como información para la toma de decisiones.
- La empresa CIDELSA se encuentra en la primera etapa de usar la información con valor agregado, es decir se encuentra en la etapa de analizar lo que sucedió.

Es un desafío continuar dándole valor agregado a la información y pasar luego a las etapas de analizar porqué sucedió y tomar medidas para el futuro.

- Para desarrollar un Data Mart se requiere de personal que entienda de datos como personal que entienda del negocio de la empresa.
- Los indicadores determinados en los Requerimientos Específicos del Negocio deben estar enlazados con los objetivos del Datamart y éstos a su vez con los objetivos estratégicos de la organización.

2. RECOMENDACIONES

- Evitar determinar las soluciones antes de entender completamente los requerimientos. Algunas veces la sensación de la urgencia causa que se escojan herramientas y técnicas de implementación del Data Mart sin entender la información que realmente requiere el negocio.
- Antes de pensar implementar un Datawarehousing es recomendable implementar primero Datamarts, para lo cual se deberá realizar un modelo de datos de la empresa antes de instalar el primer Datamart.
- Al momento de realizar el modelo de datos de la empresa se deberá considerar entidades que tal vez en ese momento no se registren datos en la organización pero que a futuro serán importantes. Sería recomendable realizar un modelo de datos contemplando como si la empresa estuviera en la etapa III de los tipos de sistema de apoyo a las decisiones.
- Cuando las organizaciones carecen del personal con las habilidades y la experiencia para construir un Data Mart, algunas veces deciden realizar outsourcing. El outsourcing puede ser una opción, pero ello requiere ser utilizado muy cuidadosamente. Es importante también contar en el equipo del proyecto con personal que conozca plenamente el negocio.
- En el proceso de Back End, la etapa de extracción, limpieza y transformación de la data es la más complicada, por lo cual se deberá contemplar el apoyo de asesores externos, como también el uso de herramientas CASE para simplificar el trabajo.

BIBLIOGRAFÍA Y ANEXOS

1. BIBLIOGRAFIA

- Sprague Ralph. Building Effective Decision Support Systems. Prentice Hall, 1982. Englewood Cliffs.
- Bodoly Samuel E. Modern Decision Making; a Guide to Modeling with Decision Support System. Mc Graw Hill, 1985. New York.
- Thierauf Robert J. User Oriented Decision Support Systems; Accent on Problem Finding. Prentice Hall, 1988. New Jersey.
- Alter S.L. Decision Support Systems: Current Practice and Continuing Challenge. Reading, MA Addison Wesley; 1980.
- Power D.J. DSS Glossary. World Wide Web, [Http://dss.cba.uni.edulglossaiy/dssglossary.htm](http://dss.cba.uni.edulglossaiy/dssglossary.htm) 1997.
- Power D.J. What is a DSS? DS, The On Line Executive Journal for Data. Intensive Decision Support, October 21, 1997. Vol 1, #3. University of Iowa.
- Albrecht Leon K. Organization and Management of Information and Management of Information Processing System. Macmillan 1993. New York

- Antony Robert N. Planning and Control Systems: A Framework for Analysis. Harvard University. 1975.
- . Curtice Robert M. Planning for Data Base Systems, Data base Management. Wellesley, Mass.: QED Information Sciences.
- . IBM, Business Systems Planning Guide, Application Manual. IBM Corporation, 1995. New York.
- Keen Peter T.W. and Michael Scott Morton, Decision Support Systems. Addison Wesley 1993.
- Turban Efrain. Decision Support and Expert Systems. Management Support Systems. Prentice Hall 1995.
- . Robert Codewell and Kirsten Aleo for Business Resource Software.
- . Hardjinder S. Gill y Prakash C. Rao. Data Warehousing, La integración de información para la mejor toma de decisiones. Prentice Hall Hispanoamérica S.A. 1996.
- . Jean Michel Franco y EDS – Instituto Prometheus. El Data Warehouse- El Data Mining. Ediciones Gestión 2000 S.A. 1997
- . The Data Warehousing Institute. Study 2000 Executive Summary.
- . The Data Warehousing Institute. Journal of Data Warehousing.

Volume 6. Number 1. Winter 2001.

- . Informix Software. Business Intelligence. Information Kit Way to Web.

- . The Data Warehousing Institute. Ten Mistakes to Avoid for Data Warehousing Managers.

- . The Data Warehousing Institute. 1999 Data Warehousing Buyers Guide. Volume 2, 1999.

- . The Data Warehousing Institute. Data Warehousing What Works? Volume 9. 2,000.

- . The Data Warehousing Institute. Journal of Data Warehousing. Volume 5. Number 3. Summer 2,000.

- . The Data Warehousing Institute. Data Warehousing What Works? Volume 10. 2,000.