

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



Facultad de Ingeniería en Electricidad y Computación

**“IMPLEMENTACIÓN DE UN DATAMART COMERCIAL PARA
UNA EMPRESA DE TELECOMUNICACIONES”**

EXAMEN DE GRADO (COMPLEXIVO)

Previa a la obtención del grado de:

INGENIERA EN COMPUTACIÓN

ESPECIALIZACION SISTEMAS TECNOLÓGICOS

DIANA DORA CHONILLO LUCÍN

GUAYAQUIL – ECUADOR

AÑO: 2015

AGRADECIMIENTO

Mis más sinceros agradecimientos a mi abuelita y a mi madre por ser los pilares más importantes en mi formación personal. A todos mis profesores, desde la escuela hasta la universidad, gracias por sus enseñanzas y dedicación a la formación de nuestra sociedad.

DEDICATORIA

El presente proyecto lo dedico a Dios, a mi madre, y a mi esposo e hijos, por el gran significado en mi vida y por su constante apoyo.

TRIBUNAL DE SUSTENTACIÓN



M.Sc. Gonzalo Luzardo M.

PROFESOR DELEGADO



M. Sc. Vanessa Echeverría B.

PROFESOR DELEGADO

DECLARACIÓN EXPRESA

"La responsabilidad por los hechos, ideas y doctrinas expuestas en este Informe me corresponde exclusivamente; y, el patrimonio intelectual de la misma, a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL"

(Reglamento de Graduación de la ESPOL).



Diana Chonillo Lucín

RESUMEN

El crecimiento de la empresa de telecomunicaciones con mayor participación de mercado en el país ha generado un desafío en cuanto a la explotación de su información para el análisis de datos y la toma de decisiones.

Con el propósito de que el área Comercial y Marketing pueda tener una herramienta que le permita poder identificar a sus clientes y crear productos acorde a sus necesidades se implementó el proyecto Datamart Comercial.

Para llevar a cabo este proyecto se utilizó la metodología de Kimball, que tiene un enfoque de menor a mayor (bottom-up), esto es la creación de un datamart para un área específica como el área Comercial y Marketing para luego integrarlo en un gran almacén de datos que es el datawarehouse.

Al finalizar el proyecto los usuarios fueron capaces de contar con una base de datos sólida y accesible, además de contar con una aplicación que le permite realizar un análisis de datos multidimensional y generar informes ad-hoc.

ÍNDICE GENERAL

AGRADECIMIENTO	ii
DEDICATORIA	iii
TRIBUNAL DE SUSTENTACIÓN	iv
DECLARACIÓN EXPRESA	v
RESUMEN	vi
ÍNDICE GENERAL.....	vii
ÍNDICE DE FIGURAS.....	ix
INDICE DE TABLAS	x
INTRODUCCIÓN	xi
CAPÍTULO 1	1
1. METODOLOGÍA UTILIZADA PARA LA IMPLEMENTACION DEL DATAMART	1
1.1 Marco teórico	1
1.1.1 Sistemas de información	1
1.1.2 Datawarehouse	3
1.1.3 Datamart	3
1.1.4 Metodologías para la implementación de un datawarehouse	3
1.2 Selección de la metodología.....	5
1.2.1 Principios básicos del Ciclo de Vida Dimensional del Negocio.....	5
1.2.2 Tareas de la metodología de Kimball.....	7
1.2.2.1 Planificación.....	8
1.2.2.2 Análisis de requerimientos	8
1.2.2.3 Tareas en la línea tecnológica	8
1.2.2.4 Tareas en la línea de Datos	9
1.2.2.5 Tareas en la línea de aplicación	13
1.2.2.6 Despliegue.....	13
1.2.2.7 Mantenimiento y Crecimiento.....	14

CAPÍTULO 2.....	16
2. RESULTADOS OBTENIDOS	16
2.1 Planificación	16
2.2 Análisis de requerimientos	17
2.3 Tareas en la línea tecnológica	17
2.4 Tareas en la línea de datos	18
2.4.1 Modelado dimensional	18
2.1.1 Diseño físico	20
2.1.2 Diseño e implementación de ETL	20
2.2 Tareas en la línea de aplicación	21
2.3 Despliegue	22
CONCLUSIONES Y RECOMENDACIONES	25
BIBLIOGRAFÍA.....	26

ÍNDICE DE FIGURAS

Figura 1.1 Enfoque de Inmon	4
Figura 1.2 Enfoque de Kimball	5
Figura 1.3 Tareas de la Metodología de Kimball	7
Figura 1.4 Ejemplo de Modelo Dimensional	10
Figura 1.5 Matriz de Procesos / Dimensiones.....	11

INDICE DE TABLAS

Tabla 1 Matriz de procesos / dimensiones.....	19
Tabla 2 Convención de nombres a utilizar para la implementación	20

INTRODUCCIÓN

Una de las principales necesidades de una empresa en crecimiento es la administración de su información de las distintas bases de datos operacionales que posee, de manera consolidada y estandarizada, que le ayude en el análisis gerencial, y a la toma de decisiones.

Los almacenes de datos o datawarehouse permiten pues almacenar información actual e histórica con el objetivo de enfrentar las necesidades de información oportuna, precisa y relevante para los encargados de tomar decisiones en la empresa. Según Ralph Kimball [3], un datawarehouse es una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis.

Para la implementación de este proyecto se escogió la metodología de Kimball, que tiene enfoque de menor a mayor, característica que le hace sencillo y flexible, debido a que permite construir un datamart como primer elemento del sistema de análisis; para luego ir añadiendo otros más que pueden compartir estructuras ya definidas o incluir nuevas. En el primer capítulo de la tesis se explicará la metodología seleccionada.

Los objetivos estratégicos obtenidos con este proyecto son: mejor calidad de datos para el análisis, creación o mejora de productos alineados a las necesidades de los clientes y captación de nuevos clientes. En el capítulo dos se detallará los resultados obtenidos de este proyecto.

En esta tesis también podemos se recalca la importancia de los sistemas de información como principal herramienta de las empresas contemporáneas que están ávidas de obtener una ventaja competitiva y mejorar la toma de decisiones.

CAPÍTULO 1

1. METODOLOGÍA UTILIZADA PARA LA IMPLEMENTACION DEL DATAMART

En este capítulo se presentan conceptos teóricos que son necesarios conocer para la implementación del datamart así también como la metodología utilizada en este proyecto.

1.1 Marco teórico

1.1.1 Sistemas de información

En su definición técnica, un sistema de información es un conjunto de componentes interrelacionados que recolectan, procesan, almacenan y distribuyen información para apoyar los procesos de toma de decisiones y de control en una organización [1].

Existen tres actividades principales en un sistema de información que son:

- Entrada: captura y recolección de datos del interior de la empresa como de su entorno.
- Procesamiento: conversión de datos a un formato significativo.

- Salida: transferencia de la información a los usuarios finales.

Desde la perspectiva de negocios, un sistema de información representa una solución organizacional y gerencial, basada en la tecnología de la información para un desafío o problema impuesto por el entorno [1].

Desde el punto de vista empresarial los sistemas de información se pueden clasificar en [1] [4]:

- **Sistema de procesamiento de transacciones (TPS):** procesa la información referente a las transacciones de la organización.
- **Sistema de información gerencial (MIS):** proveen reportes sobre el desempeño actual de la organización.
- **Sistema de soporte de decisiones (DSS):** proveen apoyo a la toma de decisiones no rutinaria analizando diferentes variables del negocio
- **Sistema de información ejecutiva (EIS):** proveen apoyo a la toma de decisiones no rutinaria que requieren además de juicio, evaluación y perspectiva gerencial. Además permite la incorporación de eventos externos.

Los sistemas gerenciales aquí descritos son sistemas para inteligencia de negocios (BIS).

La inteligencia de negocio, fue definida por Howard Dresner en 1989 como [5]: *“Un término que incluye las aplicaciones, infraestructura, herramientas y mejores prácticas que permiten el acceso al análisis de*

información para mejorar y optimizar el desempeño y la toma de decisiones”.

1.1.2 Datawarehouse

Un datawarehouse “*es una base de datos que almacena la información actual e histórica de interés potencial para los encargados de tomar decisiones en la compañía*” [1].

Su función entonces, es la de tener disponible información correcta, en el momento óptimo y en el formato adecuado para permitir a los usuarios realizar análisis de datos mediante el uso de herramientas de inteligencia de negocios.

1.1.3 Datamart

Kimball indica que la unión de todos los datamarts de una entidad forman un datawarehouse [3].

El datamart es el que almacena la información de un área específica de la empresa, con usuarios limitados, enfocados a un propósito específico y tienen una función de apoyo.

1.1.4 Metodologías para la implementación de un datawarehouse

Actualmente existen muchas metodologías de diseño y construcción de un datawarehouse, ya que cada fabricante de software de inteligencia de negocios busca imponer una metodología con sus productos. Sin

embargo, se imponen dos importantes metodologías: la de Kimball y la de Inmon [3].

Inmon es considerado el padre del “datawarehousing”, mientras que Kimball es considerado el líder del pensamiento del datawarehouse [6].

Metodología de Inmon.-

Esta metodología también se conoce como descendente (top-down). Indica que es necesario primero tener un lugar centralizado, el datawarehouse, para que los datos puedan ser utilizados para el análisis. Por lo tanto, los datamarts son creados después de haber implementado el datawarehouse. Ver Figura 1.1.



Figura 1.1 Enfoque de Inmon ¹

Metodología de Kimball.-

¹ <https://churriwifi.files.wordpress.com/2010/04/enfoque-inmon.jpg>

Concibe al datawarehouse corporativo como la unión de diferentes datamarts, estructurados comúnmente a través de una arquitectura de bus.

Esta metodología también conocida como ascendente (bottom-up), permite la implementación de pequeños datamarts para áreas específicas de la empresa, con pocos recursos y la integración paulatina en un gran almacén de datos. Ver Figura 1.2.

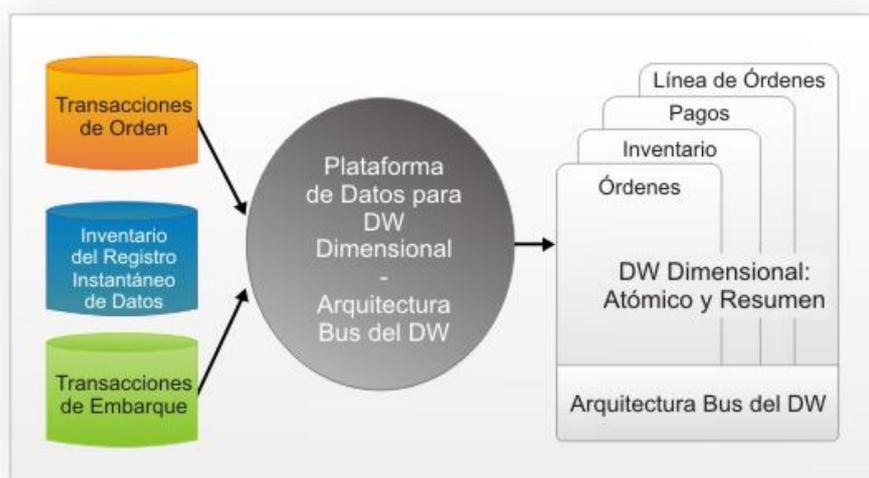


Figura 1.2 Enfoque de Kimball ²

1.2 Selección de la metodología

La metodología utilizada para la implementación de este datamart fue la de Kimball, la cual se basa en el Ciclo de Vida Dimensional del Negocio.

1.2.1 Principios básicos del Ciclo de Vida Dimensional del Negocio

² <https://churriwifi.files.wordpress.com/2010/04/enfoque-kimball.jpg>

Existen cuatro principios básicos en los que se basa esta metodología [2][3]:

- a) **Centrarse en el negocio**, se identifican los requerimientos del negocio y las partes involucradas. La probabilidad de éxito de un datamart es mucho mayor si existe una sólida comprensión del negocio de los usuarios y sus necesidades finales.
- b) **Construir una infraestructura de información adecuada**, esto es diseñar una base de información única, integrada, fácil de usar, de alto rendimiento que refleje los requerimientos de negocios de la empresa.
- c) **Realizar entregas en incrementos significativos**, esto es, crear el datawarehouse en incrementos entregables en plazos de 6 a 12 meses. Cada elemento identificado debe tener un valor de negocio que nos sirve para determinar el orden de los incrementos.
- d) **Ofrecer la solución completa**, esto es, proporcionar todos los elementos necesarios para entregar valor a los usuarios. Incluye también la entrega de herramientas de consulta ad hoc, aplicaciones para informes y análisis avanzado, capacitación, soporte, sitio web y documentación.

1.2.2 Tareas de la metodología de Kimball

En la figura 1.3 podemos observar que existe un rol central en la tarea de Definición de Requerimientos del Negocio. También podemos observar tres rutas que se enfocan en diferentes áreas:

- Tecnología (ruta superior): tareas relacionadas con software y hardware específico.
- Datos (ruta media): tareas del diseño e implementación del modelo dimensional-
- Aplicaciones (ruta inferior): tareas de diseño y desarrollo de aplicaciones de negocios para los usuarios finales

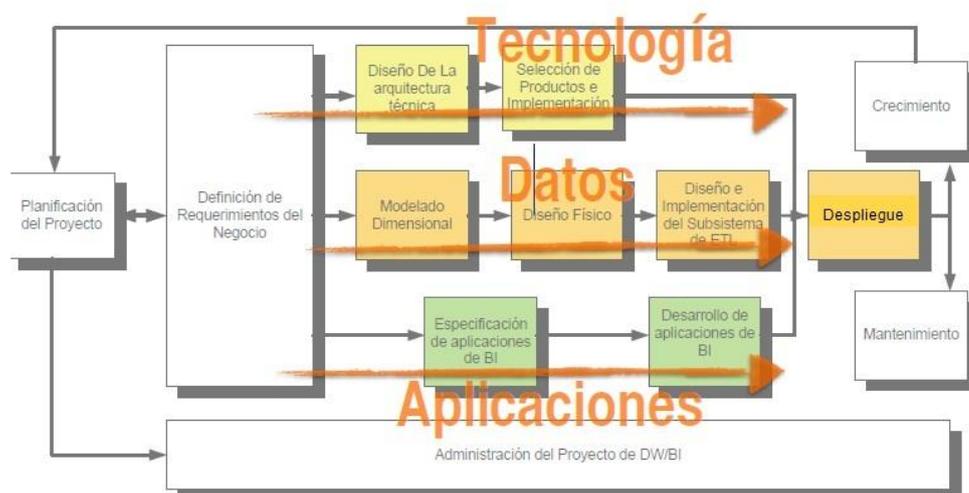


Figura 1.3 Tareas de la Metodología de Kimball ³

A continuación se detallan las tareas de la metodología de Kimball [5]:

³ http://www.interaktiv.cl/blog/wp-content/uploads/2012/04/4.-Metodologia_disegno_DW1.pdf

1.2.2.1 Planificación

- Definir el alcance, esto es entender los requerimientos del negocio e identificar riesgos.
- Identificar y programar las tareas.
- Planificar el uso de los recursos y asignar la carga de trabajo.
- Elaborar de un documento final que representa un plan del proyecto.

1.2.2.2 Análisis de requerimientos

- Contar con preparación previa sobre el negocio, los competidores, la industria y sus clientes.
- Examinar cuidadosamente el organigrama de la empresa e identificar el grupo de personas a entrevistar y programar las entrevistas.
- Llevar a cabo la entrevista y al final de la misma realizar un resumen de su contenido con los presentes.
- Revisar los resultados de la entrevista.
- Preparar y publicar los requerimientos entregables.

Luego de esta tarea, el proyecto se enfoca en las tres rutas que identificamos anteriormente en la Metodología de Kimball.

1.2.2.3 Tareas en la línea tecnológica

A continuación listamos las tareas en esta línea [4]:

- Identificar los elementos de la arquitectura técnica de los procesos ETL
- Desarrollar un modelo arquitectónico, diseñar y especificar los subsistemas para los proceso ETL
- Revisar, finalizar y documentar la arquitectura técnica para los procesos ETL
- Identificar los tipos de servicios que se necesitan para mostrar la información al usuario final. En este punto se recalca las herramientas de consultas y los servicios de escritorio.
- Identificar los tipos de servicios que se necesitan para la administración del ambiente del datawarehouse, su control y seguridad.
- Identificar y definir la infraestructura que sirve de base para todos los elementos arquitectónicos. Esto incluye hardware, red, sistema operativo y funciones de nivel inferior como la seguridad.

1.2.2.4 Tareas en la línea de Datos

En la línea de datos se identifican las siguientes tareas [3]:

- Implementar el modelo dimensional.
- Desarrollar el diseño físico.
- Diseñar e implementar los subsistemas ETL.

A continuación se explican cada una de estas tareas:

Implementar el modelo dimensional.-

Consiste en aplicar una técnica de diseño lógico que busca presentar los datos en un marco de trabajo estándar que sea intuitivo y con un acceso de alto rendimiento.

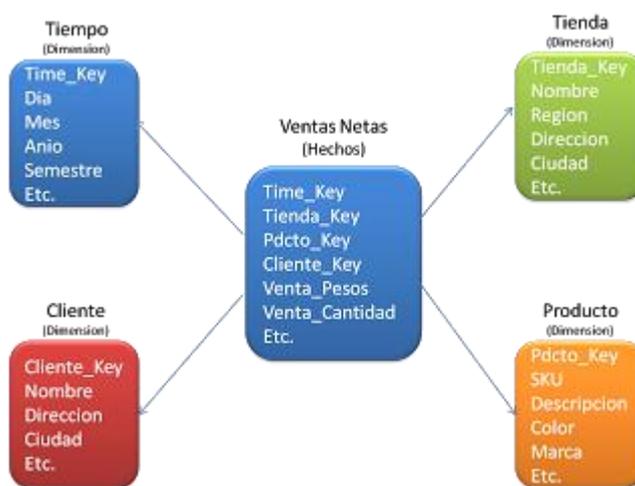


Figura 1.4 Ejemplo de Modelo Dimensional ⁴

Cada modelo dimensional se compone de una tabla con varias claves foráneas, llamada la tabla de hechos, y un conjunto de tablas más pequeñas llamadas tablas de dimensiones. Cada tabla de dimensiones tiene una clave principal única que corresponde exactamente a uno de los componentes de la clave foránea en la tabla de hechos. Además, las tablas de hecho típicamente incluyen otras columnas numéricas llamadas "métricas" o "hechos". Este modelo también se lo conoce como modelo estrella [7].

Para decidir qué modelo dimensional construir se puede hacer un planeamiento aproximado utilizando una matriz conocida como la matriz

⁴ http://3.bp.blogspot.com/_EUH3mbsliBw/TIRMDhD6doI/AAAAAAAAAFk/9-AZ9dzEnCY/s320/Modelo_Dimensional.png

de procesos /dimensiones (Enterprise Bus Matrix), pues obliga a listar y nombrar todos los datamarts que posiblemente se construyan y las dimensiones implicadas.

BUSINESS PROCESSES	COMMON DIMENSIONS						
	Date	Product	Warehouse	Store	Promotion	Customer	Employee
Issue Purchase Orders	X	X	X				
Receive Warehouse Deliveries	X	X	X				X
Warehouse Inventory	X	X	X				
Receive Store Deliveries	X	X	X	X			X
Store Inventory	X	X		X			
Retail Sales	X	X		X	X	X	X
Retail Sales Forecast	X	X		X			
Retail Promotion Tracking	X	X		X	X		
Customer Returns	X	X		X	X	X	X
Returns to Vendor	X	X		X			X
Frequent Shopper Sign-Ups	X			X		X	X

Figura 1.5 Matriz de Procesos / Dimensiones ⁵

En las filas de la matriz se identifican los procesos de negocio de la empresa, y en las columnas se identifican las dimensiones propuestas.

El proceso de diseño del modelo dimensional consiste en cuatro pasos [3]:

- a) **Elegir el proceso de negocio:** esta elección depende de la decisión gerencial, y de los requerimientos con prioridad alta o procesos de negocios más críticos.
- b) **Establecer el nivel de granularidad:** establece el nivel de detalle que se quiere presentar según los requerimientos y los datos actuales que se poseen.

⁵ <http://www.kimballgroup.com/wp-content/uploads/2013/08/Data-Warehouse-Bus-Matrix1.png>

- c) **Elegir las dimensiones:** se debe considerar que los atributos de las tablas dimensionales son posibles candidatos para ser encabezados de informes, tablas pivot, cubos u otra forma de visualización.
- d) **Identificar las tablas de hechos y sus medidas:** establece la tabla central del modelo dimensional, y contiene los valores de las medidas o indicadores de negocio. Las medidas de la tabla de hecho se pueden analizar, sumarizar o agrupar usando los criterios conocidos como dimensiones.

Desarrollar el diseño físico.-

Esta tarea se enfoca en la selección de las estructuras necesarias para soportar el diseño lógico las cuales comprenden:

- a) Definición de las convenciones estándares de nombres
- b) Afinamiento al ambiente de la base de datos
- c) Indexación y particionamiento de tablas

Desarrollar e Implementar los subsistemas ETL.-

Esta es una de las fases más importantes del proyecto que implica un 70% de su riesgo y esfuerzo. Las tareas relacionadas a esta fase son:

- a) **Extracción:** proceso que se encarga de obtener los datos que permiten efectuar la carga del modelo físico acordado.

- b) **Transformación:** proceso que convierte los datos fuentes aplicando diferentes reglas para aumentar su calidad y consistencia, consolidando la información de distintos sistemas.
- c) **Carga:** proceso que encarga de poblar el datawarehouse.

1.2.2.5 Tareas en la línea de aplicación

En la línea de aplicación se identifican las siguientes tareas [3]:

- Diseño de aplicaciones: aplicaciones candidatas e interfaces de navegación.
- Desarrollo de aplicaciones:
 - Configuración de metadatos del negocio
 - Infraestructura de herramientas: analíticas, operacionales y portal de navegación

1.2.2.6 Despliegue

Esta tarea se relaciona con el despliegue o liberación del datamart. En este punto convergen las líneas de tecnología, de datos y de aplicaciones. Antes del despliegue del marco de trabajo es necesario considerar ciertas tareas detalladas a continuación [2]:

- Desarrollar una estrategia de capacitación y soporte para el usuario final.
- Desarrollar la documentación para el usuario final y preparar el contenido para su capacitación.

- Determinar la preparación de instalación de las aplicaciones de escritorio.

Se recomienda además, llevar a cabo una prueba alfa interna, seguido de un período beta con un número limitado de clientes para luego hacer disponible el datamart al usuario final.

Existen tareas que deben realizarse como parte de la verificación después del despliegue del datamart como son:

- Instalar la herramienta de escritorio de la aplicación de inteligencia de negocio.
- Verificar y conciliar la información que obtiene el usuario final, y garantizar su calidad.
- Verificar el correcto funcionamiento de las aplicaciones que el usuario final utilizará.
- Capacitar al usuario en el uso de la herramienta y también explicar la estructura de los datos.
- Dar soporte al usuario final, esto es como parte de la etapa de estabilidad del datamart.

1.2.2.7 Mantenimiento y Crecimiento

Al entrar en la tarea de mantenimiento, se debe entender que el datawarehouse requiere de un enfoque diferente de los sistemas tradicionales. Es necesario seguir cuidando de la alimentación de datos,

y mantenimiento adecuado de la base de datos. Si el datawarehouse requiere desarrollar una tarea de crecimiento, es un índice de éxito, pues quiere decir que los usuarios están satisfechos y requieren del datawarehouse. A continuación se detallan algunas tareas.

- Continuar con el monitoreo de los procesos ETL
- Afinar el desempeño de la base de datos, debido a su crecimiento.
- Dar mantenimiento a los índices de la base de datos.
- Realizar respaldos diarios de la base de datos.
- Continuar brindado soporte al usuario.
- Dar prioridad a los nuevos requerimientos de nuevos datamarts.

CAPÍTULO 2

2. RESULTADOS OBTENIDOS

En este capítulo se presentan los resultados obtenidos en este proyecto aplicando la metodología de Kimball.

2.1 Planificación

En esta tarea se define el alcance de este proyecto y se identifican sus riesgos.

Alcance del proyecto:

- Proporcionar al usuario del área comercial una herramienta de análisis de datos que le permita la toma de decisiones y planear estrategias de mercado
- Consultar información de stock disponible, unidades pedidas, unidades vendidas, costo de equipos.
- Consultar información de cantidad de altas, bajas, migraciones de planes.
- Consultar información de cantidad de tarjetas vendidas, tarjetas activadas, dólares consumidos.
- Consultar información de minutos consumidos dentro de la operadora, con otras operadoras, minutos y dólares consumidos de minutos excedentes

Riesgos:

- Información no disponible para su obtención
- Inconsistencia de datos entre las diferentes base de datos
- No contar con recursos humanos que tenga el conocimiento adecuado
- Falta de involucramiento y compromiso de los usuarios gerentes
- Barreras por parte del usuario para utilizar nuevas herramientas de análisis

Se realizó la contratación de una empresa outsourcing para la implementación del proyecto "Datamart Comercial".

2.2 Análisis de requerimientos

En esta etapa se realizaron varias reuniones de trabajo, y se identificaron tres grupos involucrados: el gerente responsable de la toma de decisiones estratégicas, los analistas de mercado responsables de reportar a la gerencia, y el personal de sistemas que conoce los problemas informáticos y datos existentes.

En cada reunión se levantó un acta con los resultados y observaciones pudiendo clarificar los requerimientos del usuario.

La tarea más complicada en esta etapa fue programar las reuniones y comprometer a los usuarios.

2.3 Tareas en la línea tecnológica

La construcción del datawarehouse se hizo sobre una base de datos Oracle versión 9i debido a que es considerada como uno de los sistemas de base de datos más completos.

En relación con las definiciones de infraestructura tecnológicas éstas deben estar alineadas con los estandars empresariales.

2.4 Tareas en la línea de datos

2.4.1 Modelado dimensional

Se levantó la matriz de procesos/dimensiones, lo que ayudó a identificar las dimensiones en común, las tablas de hechos necesarias, además de los datamarts a construir. Ver Tabla 1.

Se pudo identificar, con ayuda de la matriz, los sistemas de los diversos procesos de negocios que se convertirían en fuentes de datos para el proceso de extracción.

Los siguientes datamarts fueron implementados en este proyecto. La lista se encuentra por orden de prioridad en base a la decisión de la Gerencia Líder del Proyecto de la empresa de telecomunicaciones:

- a. Datamart de Inventario.
- b. Datamart de Abonados.
- c. Datamart de Facturación.
- d. Datamart de Tarjetas.

PROCESOS DE NEGOCIO	DIMENSIONES COMUNES											
	fecha	producto	sub producto	plan	marca de equipo	modelo de equipo	bodega	operadora	nominación de tarjeta	region	localidad	clase de persona
Stock de Inventario	X				X	X	X			X	X	
Ventas de equipo en Inventario	X				X	X	X			X	X	
equipo en Inventario	X				X	X	X			X	X	
Activaciones de líneas nuevas	X	X	X	X						X	X	X
Renuncias de líneas	X	X	X	X						X	X	X
Cambios de planes	X	X	X	X						X	X	X
Tarjetas activadas	X	X	X	X					X	X	X	
Tarjetas vendidas consumidos de tarjetas	X	X	X	X					X	X	X	
Minutos de llamadas entrantes	X	X	X	X				X		X	X	X
Minutos de llamadas salientes	X	X	X	X				X		X	X	X
Dólares de llamadas salientes	X	X	X	X				X		X	X	X

Tabla 1 Matriz de procesos / dimensiones

2.1.1 Diseño físico

Se realizó un dimensionamiento del tamaño de las tablas y su crecimiento diario, quincenal y mensual. Además se trabajó junto al departamento de base de datos del “Área de Sistemas” para la proyección del crecimiento del datawarehouse. La tarea de afinamiento de la base datos fue realizada por el departamento de base de datos.

Se definió la convención estándar de nombres a utilizar para las tablas, columnas y constraints:

Elemento	nombre	prefijo	sufijo	significado
tabla	dw_XXX_fact	dw	fact	tabla de hechos del datawarehouse
tabla	dw_XXX_dim	dw	Dim	tabla de dimensiones del datawarehouse
columna	id_XXX	id		columna que contiene código identificador
columna	des_XXX	desc		columna que contiene descripción
constraints	pk_XXX	pk		clave primaria, donde xxx tabla
constraints	uk_XXX_yyy	uk		índice único, xxx tabla, yyy campo
constraints	fk_XXX_yyy	Fk		foreign key de la xxx tabla, yyy campo

Tabla 2 Convención de nombres a utilizar para la implementación

2.1.2 Diseño e implementación de ETL

Para la extracción de datos, se convirtió la información de las tablas a archivos planos para luego transferirla al servidor donde reside el

datawarehouse. Se crearon programas en shell que a través de sqlplus realizan la conexión de base de datos y con comandos spool extraen la información necesaria. La programación de la ejecución de estos procesos se la realiza a través de la herramienta Control-M [8], cuya función es ser un agente de automatización de procesos.

Para la transformación se creó un conjunto de procedimientos que se encarguen del reformato de datos, validación de datos, conversión de unidades, creación de nuevas columnas, y traducción de códigos.

Para la transformación, se utilizaron tablas externas para poder leer la información de los archivos planos provenientes de la extracción. Esta información se cargaba a estructuras temporales para la ejecución de los procedimientos que daban paso a la transformación.

Cuando la información ya se encuentra lista, entonces se da paso a su carga en las tablas de hechos y de dimensiones que alimentan cada uno de los datamarts.

Estos procesos también fueron programados para su ejecución diaria a través de la herramienta de Control-M.

Se realizaron las pruebas técnicas para garantizar la correcta ejecución de los procesos.

2.2 Tareas en la línea de aplicación

La aplicación que fue seleccionada como una herramienta de inteligencia de negocio fue Oracle Discoverer. Esta selección fue resultado de estudios y negociaciones con diferentes empresas, y fue una tarea que se llevó a cabo antes de iniciar con la primera tarea de planificación.

Oracle Discoverer provee los siguientes componentes:

- Oracle Discoverer Administrator.
- OracleAS Server Discoverer Plus.
- Oracle Discoverer Desktop.

La herramienta de Oracle Discoverer Administrator permite configurar y mantener las capas de usuario final y las áreas de negocio; controlar el acceso a la información; además de crear condiciones y cálculos que puedan ser incluidos por los usuarios finales de Discoverer en sus hojas de trabajo.

OracleAS Server Discoverer Plus, esta herramienta permite acceder a la información de la base de datos de manera rápida en un formato fácil de leer y comprender; permite analizar la información para dar soporte a las decisiones de negocio, y además se pueden crear reportes; permite también el acceso vía web

2.3 Despliegue

Para tener un despliegue exitoso el primer paso a ejecutar es la preparación del ambiente para el uso de los datamarts:

- Configuración del datamart con la creación de cubos de información con la herramienta Oracle Discoverer Administrator en la base de datos del Datawarehouse de Producción.
- Despliegue de los procesos ETL para la alimentación del datamart.
- Configuración de procesos para ejecutar su programación diaria
- Levantamiento de la documentación para los escenarios de pruebas

Las siguientes tareas descritas se realizaron para cada uno de los datamarts que fueron parte del proyecto. El primer datamart desplegado fue el datamart de Inventario de Inventario:

Una vez que se cuenta con el ambiente listo, ejecutamos un periodo de pruebas internas, como parte de la recomendación de la metodología aplicada. Se listan las tareas realizadas:

- Despliegue de los procesos ETL para la alimentación del datamart en ambiente de producción.
- Generación de reportes.
- Cuadratura técnica de los datos.

Pasado el periodo alfa de pruebas internas, se ejecutaron las pruebas con usuario en ambiente producción, este periodo se denomina periodo beta. A continuación se listan las tareas realizadas:

- Preparación de manual de usuario sobre el nuevo aplicativo de análisis.
- Preparación de material para la capacitación del usuario.

- Capacitación al usuario designado por la gerencia Líder de Proyecto del área Comercial.
- Generación de reportes por el usuario.
- Cuadratura técnica de los datos para verificar la calidad y la consistencia de datos con el usuario.

Una vez que se finalizaron las pruebas betas se procedió con la liberación del Datamart de Inventario para los demás usuarios:

- Capacitación a usuarios finales para el uso de la herramienta Oracle Discoverer Plus Web.
- Entrega de documentación a usuarios, esto implica los manuales de funcionamiento del datamart en la interfaz del usuario.
- Entrega de documentación a operadores del departamento de Producción. Este departamento es responsable de velar por la ejecución de todos los procesos que se encuentren en ambiente de producción. Esto implica el manual del operador, es un documento donde se detalla los procesos involucrados, la programación de su ejecución, y el detalle de cada ambiente en el que se ejecuta; matriz de procesos, es un documento donde se resumen los procesos, nombre, tipo ejecución, periodicidad, acciones a tomar en caso de falla, entradas y salidas.
- Elaboración y aprobación de documentos de aceptación de cuadraturas.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

1. El área de Comercial de la empresa de telecomunicaciones pudo contar con un sistema de información gerencial que mejoró su efectividad en el análisis para poder crear productos estratégicos
2. La empresa de telecomunicaciones contó con su primer datawarehouse y se generó interés en otras áreas para implementar nuevos datamarts.

Recomendaciones

1. Actualmente existen herramientas de inteligencia de negocio en el mercado que tienen mayor ventaja que la seleccionada, por lo que se recomienda buscar nuevas opciones que puedan apoyar aún más al usuario.
2. Dado al crecimiento en volumen de la información se recomienda que la tecnología tanto de hardware como software sea de primera línea.

BIBLIOGRAFÍA

- [1] Kenneth C. Laudon Jane P. Laudon, Sistemas de Información Gerencial, Pearson Decimosegunda edición, 2012.
- [2] Ralph Kimball, The Data Warehouse Lifecycle Toolkit, Wiley Publishing inc. Segunda edición, 2008
- [3] Rivadera, La metodología de Kimball para diseño y almacenes de datos (datawarehouse).
<http://www.ucasal.edu.ar/htm/ingenieria/cuadernos/archivos/5-p56-rivadera-formateado.pdf>, fecha de consulta enero 2015
- [4] Sistema de Información,
http://es.wikipedia.org/wiki/Sistema_de_informaci%C3%B3n, fecha de consulta enero 2015
- [5] ¿Qué es Business Intelligence?,
<https://bimexico.wordpress.com/2013/07/11/que-es-business-intelligence/>, fecha de consulta enero 2015.
- [6] Fantasies in Data Warehousing: Kimball Vs Inmon Vs Me,
<http://it.toolbox.com/blogs/data-ruminations/fantasies-in-data-warehousing-kimball-vs-inmon-vs-me-55067>
Fecha de consulta febrero 2015
- [7] Base de Datos Estratégica. <http://basedatosestrategica.blogspot.com>. Fecha de consulta febrero 2015.
- [8] Control-M Workload Automation. <http://www.bmc.com/it-solutions/control-m.html> . Fecha de consulta febrero 2015.