

Una vez hecha toda la labor de llenado de documentación, como parte de la metodología se creó el siguiente mapa de configuración con el que planteó un desarrollo de procesos el cual se presenta en el ANEXO E

4.4 Mejoramiento en procesos de Validación

Actualmente no existe en el mercado una herramienta que permita realizar las operaciones de validación requeridas para la fábrica de información por lo que se optó por llevarla a cabo mediante herramientas base como son el mismo sistema operativo desde donde se corre el proceso principal (UNIX) por medio de la utilización de los procesos denominados "shells", la explotación por medio de consultas SQL "structured query lenguaje" a los sistemas tanto origen como destino y finalmente la explotación de la propia metadata de la herramienta ETL para obtener información por medio de consultas SQL a sus tablas propietarias y conocer características de las sesiones como tiempos de proceso, horas de corrida, etcétera.

4.4.1 Bitácora de proceso general

La llamada a todos los procesos principales de la carga de la fábrica de información corporativa se realiza mediante un programa principal en UNIX.

La bitácora desarrollada tendrá las características mostradas en la tabla

4.7

| Comentario de proceso | Fecha y hora | Explicación |
|--|--------------|---|
| >>> INICIA PROCESO, | | Bandera de proceso. Marca el inicio. |
| >>> Revisando Servicios de XXX, | | Revisión de servicios (donde XXX puede ser el servicio de la base de datos o de la herramienta Powermart) |
| >>> Servicios de XXX OK. ó >>> Levantar Servicio XXX. | | Bandera de estatus de servicios. Si el servicio está arriba entonces OK sino, un mensaje será enviado a la bitácora y a los operadores. |
| >>> Inicia Carga de YYY, | | Hora de inicio de carga YYY. Donde YYY puede ser de catálogos, facturación, cuotas o demanda o cualquier concepto de negocio. |
| >>> Cargando YYY, | | Tiempo en que inicia la carga específica del módulo. |
| >>> Finaliza Carga de YYY, ó >>> Problema en Carga de YYY, | | Bandera de estatus de carga. Finaliza indica que todo estuvo correcto y si aparece algún error la palabra problema se le asignará a la bandera. |
| >>> PROCESO TOTALMENTE TERMINADO, | | Bandera de finalización de proceso |

Tabla 4.7 Cuerpo de bitácora principal

Los comandos necesarios dentro del proceso principal serán:

```
echo ">>> INICIA PROCESO, $(date)<<<" > /tmp/maintrade.stat
```

donde:

echo = Comando de UNIX que da como salida la información que necesitamos.

\$(date) = La fecha en un formato "Día Mes #día HH:MI:SS CST YYYY"

ejemplo: "Sun Mar 3 05:00:00 CST 2002"

" = Encierran el comentario a agregar a la bitácora

> = Símbolo de redireccionamiento a un tipo de salida específica, en este caso se direcciona al archivo de bitácora maintrade.stat

4.4.2 Bitácora de datos

En esta bitácora se utilizan sentencias "SQL" para generar la información a un nivel de detalle muy pequeño agrupando por totales por día sobre los conceptos importantes a validar.

La estrategia planteada es:

- Creación de tabla temporal MON_TRD_CARGADIARIA. Los valores a almacenar se muestran en la tabla 4.8.

| Campo | Tipo | Explicación |
|--------------------------|--------------|--|
| FUENTE | VARCHAR2(7) | La fuente pudiendo ser en este caso ESS, DW, ODS |
| CONCEPTO | VARCHAR2(11) | El concepto sobre el cual actúan los rubros a mostrar. Por ejemplo: FACTURAC que equivale a facturación, PCAPTADO que equivale a pedido captado, etcétera. |
| FECHA | DATE | El día a validar |
| UNIDADES_BRUTAS | NUMBER | Sumarizado por día del concepto |
| UNIDADES_NETAS | NUMBER | Sumarizado por día del concepto |
| KILOGRAMOS_BRUTOS | NUMBER | Sumarizado por día del concepto |
| KILOGRAMOS_NETOS | NUMBER | Sumarizado por día del concepto |
| PESOS_BRUTOS | NUMBER | Sumarizado por día del concepto |
| PESOS_BRUTOS_INTERMEDIOS | NUMBER | Sumarizado por día del concepto |
| PESOS_NETOS | NUMBER | Sumarizado por día del concepto |
| PESOS_ESPECIAL | NUMBER | Sumarizado por día del concepto |
| PESOS_PROMOCIONAL | NUMBER | Sumarizado por día del concepto |
| PESOS_LINEA | NUMBER | Sumarizado por día del concepto |

Tabla 4.8 Conceptos a almacenar para validación de la carga diaria

Es importante mencionar que este tipo de tablas se deberán crear dependiendo de cada desarrollo específico por iteración de la fábrica de información corporativa.

- Creación de programa "SQL" especial para extraer la información de los sistemas tanto fuente (el transaccional comúnmente) como de los destinos (comúnmente la FIC). Esta información se escribirá en la tabla anteriormente mencionada.

Al igual que en el punto anterior la forma en que se realice la explotación de información dependerá del sistema origen y por consiguiente de la iteración específica.

- Una vez llena la tabla se procedió a explotar la información realizando un formateo con los campos presentados en la tabla 4.9.

| Campo | Explicación |
|------------|---|
| FECHA | Es la fecha de carga |
| CONCEPTO | Es el concepto de negocio a validar (Facturación, Demanda, Cuota) |
| RUBRO | Es el rubro específico a validar (Kilos, Pesos, Unidades tanto brutas como netas) |
| ORIGEN | Sistema origen contra el que estamos comparando |
| DESTINO | Sistema destino |
| DIFERENCIA | Total de diferencia. |

Tabla 4.9 Conceptos a presentar en bitácora de carga

4.4.3 Bitácora de procesos individuales

Este tipo de bitácora es muy específica para el tipo de herramienta que estamos utilizando ya que se utiliza principalmente gracias a la funcionalidad de metadata que guarda la herramienta de ETL en una base de datos definida y que nos permite tomar la información de las corridas de procesos.

Lo que se requiere para esta bitácora es:

- **Definición de tres tipos de apartados para la bitácora**
 - **Apartado 1:** Todos los procesos que estén incluidos dentro del mapa de configuración
 - **Apartado 2:** Procesos principales en donde se realice alguna carga al ODS o al DW
 - **Apartado 3:** Procesos con error tanto de BD o de cualquier otro tipo.

- **La información de los procesos se presentará con los campos presentados en la tabla 4.10.**

| Campo | Explicación |
|--------------|---|
| TABLA | Es la tabla de donde se escriben los registros |
| INSTANCIA | Si la misma tabla aparece en el mismo proceso, entonces se crean instancias de ésta para diferenciarla. |
| CARGADOS | Total de registros Cargados |
| RECHAZADOS | Total de registros Rechazados |
| SESION | Nombre de la sesión |
| INICIA | Fecha y hora en que inicia el proceso |
| FINALIZA | Fecha y hora en que finaliza el proceso |
| MINUTOS | Minutos transcurridos desde que inicial el proceso a que termina. |

Tabla 4.10 Conceptos a presentar en bitácora de procesos

- **Finalmente se creó un programa que accesa la información como se muestra en la tabla 4.10.**

4.5 Mejoramiento en la Explotación

La estrategia de explotación se plantea en la figura 4.6.

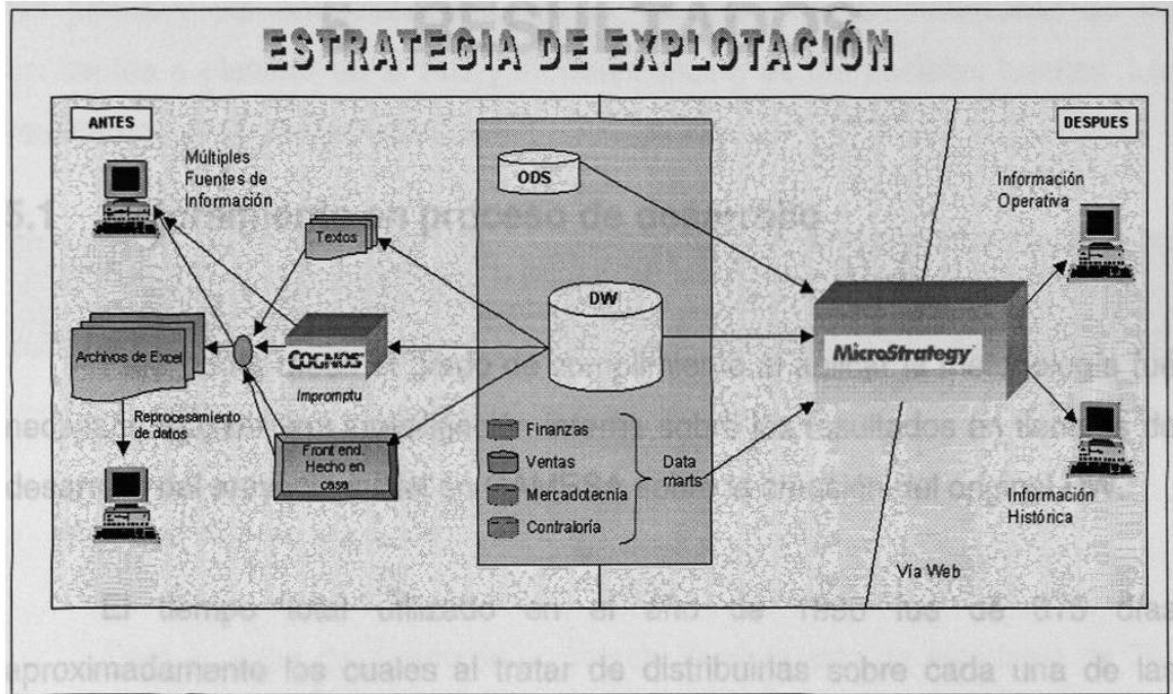


Figura 4.6. Estrategia de explotación

El objetivo principal de ésta arquitectura de Front End es el de buscar evitar los problemas actuales existentes en la distribución de información que se originan primordialmente con la existencia de las múltiples fuentes de datos del mismo DW en herramientas de presentación diversificadas y muchas veces con curvas de aprendizaje demasiado grandes, además evitaremos que los usuarios reprocesen los datos obtenidos por los diferentes medios con lo que se busca eliminar la cantidad de errores manuales que se pueden dar.

La estrategia se centra en la implementación de una plataforma denominada "Microstrategy" la cual apoyada sobre la base de metadata y del ODS y DW, permitirá el acceso a la información de una forma ordenada y con la disminución de tiempos de reproceso de los usuarios permitiéndoles dedicarse a su tarea o función que es principalmente la de analistas de información.

5 RESULTADOS

5.1 Mejoramiento en proceso de desarrollo

Para poder medir el grado de cumplimiento al aplicar la metodología fué necesario realizar una investigación interna sobre los resultados en tiempos de desarrollo del proyecto inicial en GAMESA sobre la creación del original DW.

El tiempo total utilizado en el año de 1996 fue de 375 días aproximadamente los cuales al tratar de distribuirlos sobre cada una de las actividades que se proponen en la metodología, esto permite obtener un tiempo de equivalencia el cual se aplica para medir el grado de cumplimiento. Cabe mencionar que de los 375 días sólo se tomarán 290 que son los tiempos comparativos de las fases de análisis, diseño y construcción.

Los tiempos del proyecto inicial fueron tomados a partir del momento en que se aceptó el proyecto de DW como parte de la solución de la compañía.

Los resultados obtenidos a raíz de la aplicación de la metodología en cuestión de tiempos, comparando los diferentes módulos se presentan a continuación.

- **Tiempos en la etapa de análisis**

Dentro de esta etapa de análisis se consideraron los puntos importantes sobre el entendimiento de la situación actual, la programación y ejecución de las juntas y las entrevistas necesarias, así como el entendimiento de los conceptos a plasmar en la FIC y la identificación de las posibles fuentes. Los resultados se presentan en la tabla 5.1.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 28 días | 5 días |

Tabla 5.1 Comparativo Análisis Proyecto Original - FIC

- **Mejoramiento en etapa de diseño**

Para la etapa de diseño se estableció la propuesta de estrategia de FIC sobre la anterior de DW en la cual se definieron las formas de llevar a cabo el modelo, la aplicación de la metodología Kimball para el modelado dimensional y el mapeo campo a campo, se diseñó el modelo tanto en su forma lógica como física, el dimensionamiento y la parte de diseño de arquitectura de infraestructura y el diseño y la configuración de interfases.

- Para el Back End.

Los tiempos se presentan en la tabla 5.2.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 20 días | 19 días |

Tabla 5.2 Comparativo Diseño BE Proyecto Original - FIC

Aquí se presentó un pequeño desfase de tiempos con respecto al tentativamente estipulado debido principalmente al entendimiento formal del modelado dimensional y a su correcta aplicación.

Es importante mencionar que el planteamiento del modelo ya está realizado por lo que en sucesivas iteraciones ya no será necesario utilizar tanto tiempo.

- Para el Front End.

En esta parte se comparó solo la estrategia de creación de indicadores en la nueva herramienta y sólo 2 reportes para la explotación de la FIC. Los tiempos se presentan en la tabla 5.3.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 42 días | 6 días |

Tabla 5.3 Comparativo Diseño FE Proyecto Original - FIC

La diferencia en tiempos se presenta debido a que no existía en su momento una herramienta especializada, sino que tuvieron que diseñar una para poder presentar información al usuario.

- Mejoramiento en etapa de construcción

En esta etapa es donde se notó una gran mejoría que se vió reflejada principalmente por la aplicación de la herramienta especializada la cual disminuyó considerablemente los tiempos de construcción tanto para el Back end (mostrado en la tabla 5.4.) como para el Front End (tabla 5.5.) ya que todas las banderas, validaciones, alarmas y demás cuestiones tenían que ser creadas a través de métodos codificados los que dificultaron el rápido desarrollo originalmente.

- Para el Back End.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 120 días | 20 días |

Tabla 5.4 Comparativo Construcción BE Proyecto Original - FIC

- Para el Front End.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 80 días | 5 días |

Tabla 5.5 Comparativo Construcción FE Proyecto Original - FIC

- Etapa de pruebas:

Finalmente el establecimiento de una etapa de pruebas a diferentes niveles con respecto al desarrollo original representó una mejor aceptación en la futura implementación debido a la casi inexistencia de errores. En la tabla 5.6 se muestran los resultados para el Back End y en la tabla 5.7 para el Front End.

- Para el Back End

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 10 días | 9 días |

Tabla 5.6 Comparativo Pruebas BE Proyecto Original - FIC

- Para el Front End

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 15 días | 3 días |

Tabla 5.7 Comparativo Pruebas FE Proyecto Original - FIC

- Tiempo de implementación

Este punto, aún y cuando no era parte del objetivo de la tesis se plantea como comparativo en la cual se trabaja la implementación con el usuario aún y cuando ésta no se hizo a nivel nacional ya que solo se realizó con el superusuario de la información. Los resultados se muestran en la tabla 5.8.

| Tiempo de duración (Proyecto Original) | Tiempo de duración (FIC) |
|---|-----------------------------|
| 60 días | 3 días |

Tabla 5.8 Comparativo Implementación Proyecto Original - FIC

Uno de los puntos importantes obtenidos de aplicar la metodología es que todo lo desarrollado queda postrado sobre la base una metadata sólida y confiable que permitirá en un momento dado servir de armas para poder continuar con la construcción de la FIC en sentido incremental.

El comparativo de los tiempos del desarrollo actual contra el anterior se presenta en la figura 5.1.

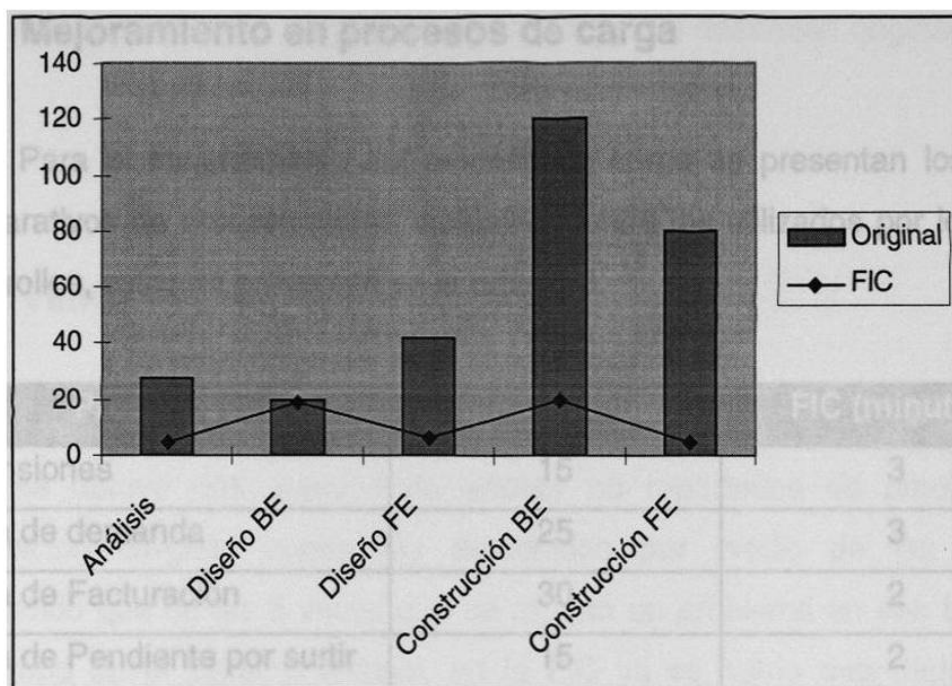


Figura 5.1. Gráfica comparativa de tiempos de proceso de desarrollo Original - FIC

Finalmente se resume que originalmente se utilizaron 290 días para las actividades de análisis, diseño y construcción las cuales se disminuyeron a 55 para un proyecto de las mismas dimensiones en una arquitectura de FIC.

5.2 Mejoramiento en procesos de carga

Para el mejoramiento del proceso de carga se presentan los tiempos comparativos de procesamiento actuales, contra los utilizados por los nuevos desarrollos, estos se presentan en la tabla 5.9.

| Procesos | Original (minutos) | FIC (minutos) |
|-------------------------------|--------------------|---------------|
| Dimensiones | 15 | 3 |
| Carga de demanda | 25 | 3 |
| Carga de Facturación | 30 | 2 |
| Carga de Pendiente por surtir | 15 | 2 |
| Carga de cuotas | 15 | 5 |

Tabla 5.9 Comparativo tiempos de proceso Proyecto Original - FIC

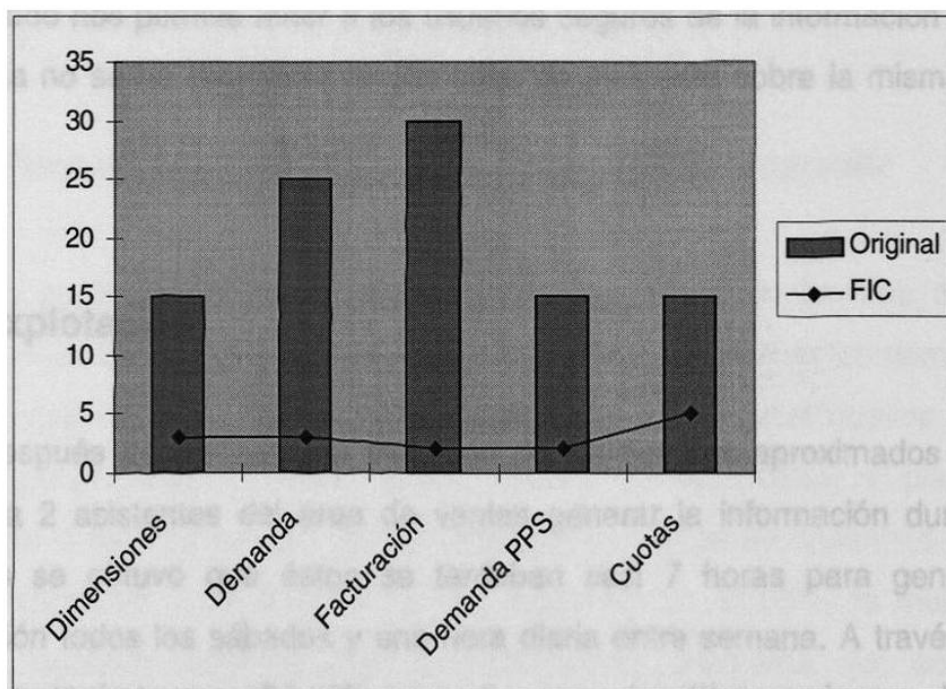


Figura 5.2. Gráfica comparativa de tiempos de proceso de carga Original - FIC

La información recopilada de los tiempos anteriores se obtuvo por observación de 30 días con respecto a los procesos originales y con un paralelo sobre la nueva arquitectura. En la figura 5.2 se puede observar la gráfica

comparativa de tiempos de proceso de la carga del desarrollo original contra el nuevo.

5.3 Validación

A través de una bitácora generada a partir de Agosto hasta Diciembre del 2001 se obtuvo una relación de errores no reportados de proceso y de información cargada contra los detectados por medio de las bitácoras obteniendo que de las 8 veces que se reportó un problema en ese transcurso de tiempo en el sistema original, en la FIC ya se había detectado antes y tomado las acciones correctivas necesarias.

Los resultados quizás no se presenten tan fácilmente en números pero lo desarrollado nos permite tener a los usuarios seguros de la información ya que a la fecha no se ha levantado ningún caso de problema sobre la misma en el área.

5.4 Explotación

Después de realizar una medición de los tiempos aproximados que le tomaba a 2 asistentes del área de ventas generar la información durante 6 semanas se obtuvo que éstos se tardaban casi 7 horas para generar la información todos los sábados y una hora diaria entre semana. A través de la nueva estrategia la consolidación se realiza en automático por lo que de las 7 horas que dedicaban anteriormente sólo se llevan 2 minutos en generar la información.

6 CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones

Se presentan las conclusiones identificándolas para cada uno de los objetivos buscados en el proyecto y finalmente un apartado de conclusiones generales.

6.1.1 Para el mejoramiento en el proceso de desarrollo

Un punto importante a discutir es el grado de cumplimiento de este objetivo. Algo que queda muy claro es que tomando en cuenta los tiempos de proyecto que tomó inicialmente con respecto a los tiempos nuevos con la metodología propuesta podemos llegar a concluir que la hipótesis H1 planteada al inicio de la tesis se cumple ya que a raíz de un tiempo de desarrollo de 290 se logró disminuir a 55 días lo cual se refleja en un 81.03% de mejora en tiempos logrando además una documentación adecuada (la cual era inexistente) y una metodología formal en la cual se plantea disminuir los tiempos de futuras iteraciones.

A raíz de este desarrollo además podemos concluir una serie de cuestiones que me gustaría comentar:

- Existe un indicador que no propuse para medir el grado de cumplimiento de este objetivo debido a que este consiste en algo no mensurable, este mide la ventaja que puede traernos el hecho de crear una documentación formal sobre todos y cada uno de los pasos dentro del desarrollo de una solución ya sea de DSS o de cualquier otra área de desarrollo de proyectos. Si nos remontamos a los problemas que existían en la compañía inicialmente y que formaban el porqué crear un concepto como lo es la FIC podemos verdaderamente pensar en que sin una base firmemente documentada no es posible realizar cambios de una manera sencilla, además de que dependeríamos de una persona ya que todo el conocimiento recaería en ella, si esa persona se va, entonces el proyecto pudiera desmoronarse. Soy muy enfático en este punto debido a que yo entré a esta área visualizando una creciente necesidad por parte de las diferentes áreas del negocio de tener un rápido crecimiento de lo que se podía decir se tenía como solución para el diseño de soporte a la toma de decisiones, pero debido a la falta de documentación, a la falta de una metodología y finalmente de una arquitectura robusta que permitiera este rápido desarrollo nos llevó casi 1 año empezar a trabajar sobre nuevos desarrollos. Es un hecho que después de implantar el concepto de FIC en la compañía el desarrollo de iteraciones ha ido en aumento.
- Actualmente utilizando esta metodología hemos empezado a crecer la FIC incrementando el alcance en el área de ventas, pero además incluyendo nuevas iteraciones como son indicadores de servicio, indicadores de análisis de mercado, indicadores de gastos, etcétera. La lista sigue creciendo. Lo más importante es que hemos incrementado la cantidad de indicadores necesarios para el negocio y la facilidad con la que lo hemos hecho ha dependido ahora únicamente del análisis ya que con la metodología aplicada el diseño y la construcción se vuelve muy sencilla.

- Un punto importante es que la metodología no es una solución mágica que va a dar resultados inmediatamente. Es necesario que la gente que entre al área tenga algunos conocimientos básicos de DW y manejo de bases de datos ya que ésta es la base para un rápido desarrollo.
- Los tiempos de desarrollo son relativos al tamaño del proyecto, pero lo que aseguramos es que por medio de la metodología los tiempos de desarrollo de nuevas iteraciones sean menores siempre y cuando se aplique correctamente.

6.1.2 Para el mejoramiento en el proceso de carga

A raíz de la organización del proceso de carga y de la realización de las interfases por medio de una herramienta especializada pudimos constatar que cumplimos con el objetivo de mejoramiento en estos procesos debido a que de los 100 minutos que se tardaba la carga de datos se disminuyó a 15 lo cual refleja un 85% de mejora con respecto al tiempo original.

Dentro de los puntos importantes cabe mencionar:

- El mejoramiento se obtuvo debido a varios factores, uno de ellos principalmente fue la herramienta utilizada. En la actualidad existen muchas herramientas en el mercado las cuales tienden a realizar las mismas funciones, algunas un poco mejor que otras, pero todas tienen un mismo concepto en común, evitamos la realización de codificación difícil de mantener por medio de una interfaz gráfica que permita organizar los procesos de tal forma que estos sean fácilmente entendibles y que guarden todo el desarrollo dentro de una metadata. Si una herramienta no posee estas características, entonces no es una herramienta útil para el objetivo

que estamos buscando ya que estaríamos cayendo en los problemas que se tenían originalmente:

- La misma herramienta da origen a una nueva forma de organizar los procesos. El mapa de configuración realizado no es más que una concepción de todos los procesos que de otra forma tendrían que estar codificados en programas a los cuales todos los servicios requeridos por el área de Back end tendrían que acoplárseles por medio de codificación especial. Una herramienta ETL nos permite concebir a los procesos como objetos encapsulados que en este caso se denominan mapeos.
- Uno de los puntos más importantes que quizás no se visualizaron dentro del concepto principal de la tesis fueron la optimización de los tiempos de construcción de interfases. Estos tiempos son dependientes de muchos factores, pero principalmente algo que los viene a mejorar enormemente es la creación de la documentación previa (mapeo campo a campo, entendimiento de las fuentes) que vienen a resumir el trabajo en realiza el análisis y después solo traducirlo en procesos. Aún así se observó que los tiempos originales de construcción con respecto a los nuevos se mejoraron en un 83.3%.
- Otra cuestión que también es importante mencionar es que el hecho de haber mejorado los tiempos de cargas en GAMESA puede aplicarse en otros lugares o en otras iteraciones, pero es importante tener en cuenta que un mejoramiento puede partir desde que se mejora la forma de estructurar los procesos (mapa de configuración), pero también depende del hardware en donde van a correr los procesos y muchas veces del sistema operativo que se tiene. Es posible tener la misma herramienta ETL en 2 sistemas operativos diferentes (NT y UNIX) por ejemplo y tener algunas mejoras considerables que van desde el mejor manejo de memoria, hasta la optimización de recursos internos. Es importante comentar que los tiempos obtenidos son con una infraestructura definida para las necesidades de GAMESA.

- Otro punto a considerar es el DBMS el cual juega un importante rol cuando se empiezan a manejar grandes volúmenes de información, y además la forma en que se diseñó la arquitectura de la base de datos cuando se utiliza la modelación dimensional en ambiente relacional como fue en este caso tiende a afectar los tiempos de carga y más aún los tiempos de explotación.
- La forma de estructurar los procesos con el mapa de configuración sirvió en la empresa para mejorar los tiempos, pero es importante recalcar que si hubiéramos tenido que partir del mismo desarrollo pero sin herramientas especializadas quizás también hubiéramos optimizado los tiempos de proceso, pero no hubiéramos creado el pilar para complementar el concepto de la FIC, que consiste en la metadata de todo esto y mejor aún, en la inclusión de todos los servicios necesarios para satisfacer todos los movimientos de datos.

6.1.3 Para el mejoramiento en el proceso de validación

Éste es quizás el punto más difícil de validar en esta tesis. El porqué es debido a que previamente el nivel de validación que se tenía era casi nulo y no existía una estrategia formal en la cual se pudieran apoyar para poder detectar errores. La forma en que se atacó, aunque suene un poco obvia, muchas veces no se aplica en las compañías por razones que van desde la imposibilidad de poder implementarlas debido a la forma en que organizan sus procesos, la falta de conocimiento, los desarrollos a vapor, etcétera.

La forma en que opté por medir el grado de cumplimiento fue comparando las veces que tuvimos problemas no detectados a tiempo con las veces que si se detectaron a partir de que se implementaron las bitácoras comprobándose en el corto tiempo que tienen funcionando que constituyen un importante indicador para la correcta y oportuna detección de errores. Muestran un mejoramiento sustancial el cual se puede comprobar por los cero reportes obtenidos hasta la fechas sobre los datos.

Puntos importantes a concluir son:

- **Aún y cuando este sea un punto básico a realizar en cualquier tipo de desarrollo, es increíble tener muchas veces tan poco desarrollado sobre este rubro, tal es el caso como en GAMESA.**
- **La forma en que se atacó este punto fue una estrategia definida e implementada basándonos en consejos de Inmon y Rudin [Inmon y Rudin, 98] pero lo más importante es que sí funciona y es un punto que no debe menospreciarse.**
- **Los puntos que se pudieron atacar se basan básicamente en la integridad de la información y en la confianza que los usuarios tengan sobre la misma. Casos como que un día se presentara un número y al día siguiente este mismo cambiara para la misma fecha generó problemas enormes en la confianza de los usuarios, además otro punto importante es que no se tenían medios para comprobarles si los errores existían debido a errores de carga, o errores en los datos fuente.**
- **Un concepto que se debe de tomar en cuenta que nos obligó a crear esta estrategia de validación el cual lo he vivido en carne propia es el enfrentamiento contra un sistema origen el cual, como todo sistema de información, permite a la gente meter mano sin siquiera informar o enterar a los sistemas involucrados de los cambios. En un concepto de ODS y DW es imprescindible estar enterado de estos cambios ya que la estrategia de carga generalmente está basada en criterios de fechas, por lo que fechas pasadas usualmente no se recargan al menos que se realice un requerimiento formal de esto. Si no se tienen las armas para poder demostrar los cambios en el sistema origen y a la hora que los usuarios empiezan a comparar información, es en este momento cuando se ven los problemas.**

6.1.4 Para el mejoramiento en el proceso explotación

Los resultados obtenidos en el mejoramiento de procesos por medio de la estrategia definida, verdaderamente trajeron cambios contundentes basados en los análisis de tiempos de ahorro presentados. Desde el hecho de que de 7 horas se baje a 2 minutos el proceso de explotación de información y además la posibilidad de que un Vicepresidente o un Director de la compañía tenga el acceso a la información sin necesidad de intermediarios es el principal beneficio obtenido de esta estrategia.

- Dentro de los puntos a considerar es que el proceso de explotación tiene diversos factores que permiten medirlo los cuales van desde los tiempos de respuesta dependiendo de la herramienta utilizada para esta explotación, la capacidad de navegación, uno muy importante que es la presentación de los datos 100% confiables, y el punto evaluado para esta tesis, que es la centralización de información que permita una explotación central y que disminuya los tiempos del proceso de explotación proponiendo una estrategia diferente de generación de información evitando la triangulación de la misma.
- Esta estrategia considero que debe ser utilizada en cualquier proyecto de DSS ya que si no, se seguiría teniendo el problema que se presentaba originalmente (la falta de integridad de la información) por posibles errores manuales o por la posible comparación de datos con criterios diferentes en diferentes sistemas o fuentes de datos (desde archivos textos generados, hojas de cálculo, etcétera).
- Lo difícil para lograr este objetivo llega cuando se requieren integrar las diferentes áreas con indicadores en común y que por consiguiente se requieren tener almacenados en la FIC. Un punto importante es que esto se logrará una vez que el concepto y el desarrollo de la FIC este muy robusto. Mientras tanto si es posible la integración pero tiene que utilizarse el

desarrollo incremental, tendiendo a atacar los indicadores principales del negocio y finalmente los adicionales.

6.1.5 Conclusiones en general

A raíz de todos los puntos trabajados podemos constatar que una metodología robusta, un buen conocimiento de lo que se tiene en mente pero además un ingrediente que no he mencionado y que considero sumamente importante y que viene a complementar lo antes mencionado es la visión que se tenga cuando se propone/diseña/construye una FIC.

Una FIC es un concepto que va más allá del simple almacenamiento de datos, este es el corazón y cerebro de la compañía el cual tiende a dirigir las diversas áreas del negocio las mismas que en conjunto hacen salir adelante a la compañía. Una compañía sin información en la actualidad, es una compañía con un problema serio que tarde que temprano será eliminada por los competidores que si tuvieron la visión y se amaron correctamente.

Una FIC es un concepto que también va más allá de la simple habilitación de información ya que debe de proporcionar información con valor agregado para la compañía y debe de proporcionar ventajas competitivas, y lo más importante, debe de ser lo suficientemente robusto para soportar las diversas áreas en conjunto.

Una FIC no se debe considerar como un sistema como muchos tienden a hacerlo, sino como una solución a la mayoría de los problemas que han surgido por el manejo de cantidades industriales de datos y su posterior transformación a información.

Como mencionaba anteriormente, esta metodología propuesta permitió atacar los problemas mencionados originalmente, podemos considerar que

ésta puede ser utilizada sin pensarse para diversas compañías y en otras áreas pero hay que considerar que el tamaño, la situación propia de la empresa, la forma en que se encuentra estructurada, etcétera, son los que rigen para poder repetir los mismos resultados obtenidos. Como también mencionaba anteriormente, este concepto actualmente está creciendo a pasos agigantados en la compañía, trayendo algunos nuevos problemas, pero algo que es importante recalcar es que, bien definidas las bases, lo demás es cuestión de acoplamiento mínimos para alcanzar el resultado esperado,

Un problema que ha presentado la mayoría de las iteraciones que hemos estado realizando y que hemos estado viviendo es que la metodología propone las bases, pero el punto crítico está en la integración de las fuentes debido a los diversos problemas internos que tienen cada uno de estos sistemas y que por cuestiones internas hacen difícil y conforman un escollo que puede poner en peligro la construcción de una FIC.

Finalmente como toda metodología siempre hay pasos adicionales que pueden acoplarse debido a que cada empresa es un mundo diferente pero pudimos constatar que teniendo como base los conceptos mencionados se tendrá un pilar (la metadata) que vendrá a soportar los futuros cambios que puedan avecinarse.

6.2 Recomendaciones

En el desarrollo de esta tesis se han tratado muchos aspectos relacionados con la implementación de la FIC, pero es importante recalcar que una FIC es una solución que soporta el desarrollo de soluciones para la toma de decisiones, pero el concepto como tal de FIC tiene muchos hallazgos, nuevas tendencias y nuevas formas de pensar que pueden complementar el desarrollo actual. Y más aún, la arquitectura de FIC no es la única solución

existente que puede solucionar los problemas del manejo de información en las empresas.

A continuación se listan algunos trabajos sobre los que se puede ahondar para futuras investigaciones:

- La construcción de un DW Distribuido y su impacto en el negocio.
- La implementación del concepto de minería de datos a diversos niveles en donde se debe buscar desde las herramientas especializadas para llevarla a cabo, la correcta integración de los datos y la búsqueda del conocimiento.
- El diseño de un Data Mart de Rentabilidad en donde se combine información de Ventas y Gastos para sacar las verdaderas ganancias
- El concepto de Data WebHouse
- El impacto de los medios de comunicación actuales sobre la distribución de información.
- La utilización de las tecnologías SAN (Storage Area Network) y NAS (Network Attached Storage) en el ambiente de la FIC.
- El diseño y la implementación de un ODS distributivo en las empresas para el manejo centralizado de la información y la correcta distribución de ésta.
- Un estudio en México sobre como las soluciones implementadas han tenido éxitos o fracasos, la determinación del porqué de éstos, y si fueron fracasos, entonces el cómo el concepto de FIC puede llegar a beneficiarlas.
- Como se robustece el concepto de metadata por medio de la aplicación de la metodología regida por el concepto de cadena de valor.
- El análisis del valor agregado que se obtiene por medio de la utilización del concepto de DW en la planeación estratégica.
- Como lograr una distribución eficiente del conocimiento de lo que existe dentro de una FIC permitiendo al usuario aprender del negocio a partir de ésta.

Además, algunos conceptos que por cuestiones de tiempo o porque no formaban parte del objetivo de esta tesis no fueron atacados con más detalle pero que corresponden a piezas del todo en el rompecabezas de la FIC y que podrían ser buenos temas para futuras investigaciones son:

- Optimización de la organización tecnológica para el mejoramiento en los tiempos de respuesta tanto en la carga, como en la distribución y explotación de información en una compañía en particular.
- Las ventajas que se pueden obtener por medio del diseño del ODS con una arquitectura relacional.
- La correcta implementación de mecanismos de detección de datos no muy utilizados y su administración.
- El diseño y la implementación de una metodología para la correcta evaluación de herramientas de ETL para el buen funcionamiento del DW.
- Estrategias avanzadas para la validación de la información por medio de la implementación del concepto "calidad de datos" y la utilización de herramientas especializadas para este fin.
- El diseño y la implementación de una estrategia más avanzada con herramientas especializadas para el control y manejo de la metadata.

Como esta tesis partía de un concepto que ya la compañía había aceptado y que se optó por mejorar no fué necesaria una estrategia de definición de costo/beneficio y la temática de venta del mismo, pero un buen tema de investigación para empresas que no conozcan o que no hayan tenido la visión de crear una FIC podría ser el establecimiento formal de los indicadores que pudieran establecer el porqué una solución como el DW o más aún de la FIC beneficiaría a la compañía.

BIBLIOGRAFÍA

- [Anahory y Murray, 97] Anahory, Sam; Murray, Dennis; *"Data warehousing in the real world : a practical guide for building decision support systems"*; Pie imprent Harlow, England ✆ ISBN:0201175193; Addison-Wesley, 1997
- [Ball, 99] Ball, Leslie; *"What is an ERP?"*; Artículo; <http://intra.som.umass.edu/somit/public/pages/ERP%20Final%20Presentation/index.htm>; 1999
- [CAI, 2001] <http://www.cai.com>
- [Chen, 76] Chen, Peter; *"The Entity-Relationship Model Toward a Unified view of data"*; Artículo; Instituto Tecnológico de Massachusetts; 1976
- [Corey y Abbey, 97] Corey, Michael; Abbey, Michael; *"Oracle Data Warehousing, Guía práctica para analizar, construir e implantar con éxito un sistema Data Warehouse"*; USA; ISBN:0078822424; Oracle Press, Mc Graw Hill, 1997
- [Dorsey, 00] Dr. Dorsey, Paul; *"Data Warehouses, ad hoc query tools and other ways to destroy your company"*; Dulcian Inc.; 2000
- [DWOL, 01] <http://www.datawarehousingonline.com>; The data Warehouse Institute on Line
- [Infoopt, 99] Informatica; *Conceptos de optimización*; 1999
- [INFORMATICA, 2001] <http://www.informatica.com>
- [Inmon e Imhoff y Sousa, 01] Inmon, W.H.; Imhoff, Claudia; Sousa, Ryan; *"Corporate Information Factory"*; Editorial Wiley; ISBN: 0-471-39961-2; USA; 2001
- [Inmon y Rudin, 98] Inmon, W.H.; Rudin, Ken; *"Data Warehouse Performance"*; Editorial Wiley; ISBN: 0471298085; USA; 1998
- [Inmon, 96] Inmon, W.H.; *"Building the Data Warehouse"*; Editorial Wiley; USA; 1996
- [Kimball y Reeves y Ross, 98] Kimball Ralph; Reeves, Laura; Ross, Margy; *"The Data Warehouse Lifecycle Toolkit"*; USA; ISBN: 0471255475; Wiley, 1998

- [Kimball, 96] Kimball Ralph; *"The Data Warehouse Toolkit"*; USA; ISBN: 0471153370; Wiley, 1996
- [MICROSOFT, 2001] <http://www.microsoft.com/office/project/default.asp>
- [MICROSTRATEGY, 2001] <http://www.microstrategy.com>
- [Ori y Santos, 96] Ori, R.;Santos, F., *"Data Extraction, Transformation, and Migration Tools"*; Basado en un documento del gobierno de E.U., 1996
- [Orr, 96] Orr, Ken; *"Data Warehousing Technology a White Paper by Ken Orr"*, The Ken Orr Institute; USA;Copyright 1996, edición revisada en el 2000.
- [Plaster, 97] Plaster, J. Donald; *"Enterprise Data Administration"*; Notas de sesión ICCM; 1997
- [Ruble, 1997] Ruble, David; *"Análisis y diseño práctico de sistemas"*; Editorial Prentice Hall; ISBN:9701701224; 1997
- [Sakaguchi y Frolick, 96] Sakaguchi, Toru; Frolick, Mark; *"A Review of the Data Warehousing Literature"*; University of Memphis; Abstract, 1996
- [TDWI, 00] The Data Warehouse Institute;*"Ten Mistakes to avoid"*; 2000
- [Tylor, 01] Tylor, Art; *"Data Warehouse and the ETL tool"*; Archivo CIBER, Inc.; <http://www.tdan.com>; 2001

LISTA DE TABLAS

| Tabla | Título | Página |
|-------|--|--------|
| 3.1 | Equivalencias de dimensionamiento generales | 74 |
| 3.2 | Resumen de dimensionamiento | 75 |
| 4.1 | Requerimientos | 86 |
| 4.2 | Matriz de bus | 90 |
| 4.3 | Desglose de conceptos en Matriz de bus | 91 |
| 4.4 | Sistemas identificados | 92 |
| 4.5 | Tipos de tablas del modelo | 95 |
| 4.6 | Dimensionamiento | 95 |
| 4.7 | Cuerpo de bitácora principal | 102 |
| 4.8 | Conceptos a almacenar para validación de la carga diaria | 103 |
| 4.9 | Conceptos a presentar en bitácora de carga | 104 |
| 4.10 | Conceptos a presentar en bitácora de procesos | 105 |
| 5.1 | Comparativo Análisis Proyecto Original - FIC | 108 |
| 5.2 | Comparativo Diseño BE Proyecto Original - FIC | 109 |
| 5.3 | Comparativo Diseño FE Proyecto Original - FIC | 109 |
| 5.4 | Comparativo Construcción BE Proyecto Original - FIC | 110 |
| 5.5 | Comparativo Construcción FE Proyecto Original - FIC | 110 |
| 5.6 | Comparativo Pruebas BE Proyecto Original - FIC | 110 |
| 5.7 | Comparativo Pruebas FE Proyecto Original - FIC | 111 |
| 5.8 | Comparativo Implementación Proyecto Original - FIC | 111 |
| 5.9 | Comparativo tiempos de proceso Proyecto Original - FIC | 113 |

LISTA DE FIGURAS

| Figura | Título | Página |
|--------|--|--------|
| 2.1 | La fábrica de información corporativa. | 19 |
| 2.2 | Arquitectura 1 (Mundo Externo, Aplicaciones, DW, Metadata) | 27 |
| 2.3 | Arquitectura 2 iniciando con Datamarts (Mundo Externo, Aplicaciones, Capa de IyT, DW, DM, Metadata) | 28 |
| 2.4 | Arquitectura 2 Datamarts y Data Warehouse en conjunto (Mundo Externo, Aplicaciones, Capa de IyT, DW, DM, Metadata) | 28 |
| 2.5 | Arquitectura 3 (Mundo Externo, Aplicaciones, Capa de IyT, DW, DM, Metadata y ODS) | 29 |
| 2.6 | Alineación de la FIC a las áreas competitivas del negocio | 32 |
| 2.7 | El ciclo de vida dimensional del negocio | 36 |
| 2.8 | Diagrama Estrella | 40 |
| 2.9 | Evolución de las configuraciones de la arquitectura | 42 |
| 2.10 | Arquitectura de Back End | 43 |
| 2.11 | Arquitectura de Front End | 45 |
| 2.12 | Metadata | 47 |
| 4.1 | Diagrama de situación actual | 82 |
| 4.2 | Plan de proyecto | 85 |
| 4.3 | FIC propuesta | 89 |
| 4.4 | Arquitectura propuesta | 99 |
| 4.5 | Infraestructura | 100 |
| 4.6 | Estrategia de Explotación | 106 |
| 5.1 | Gráfica comparativa de tiempos de procesos de desarrollo Original - FIC | 112 |
| 5.2 | Gráfica comparativa de tiempos de procesos de carga Original - FIC | 113 |

LISTA DE ABREVIATURAS

| Abreviatura | Significado |
|-------------|---|
| ADO | Almacén de datos Operativo |
| BD | Base de datos |
| CPU | Central Processing Unit |
| CVD | Ciclo de vida dimensional |
| CVDS | Ciclo de vida de desarrollo de Sistemas |
| DBMS | Database Management System |
| DM | Data Mart |
| DSS | Decision Support Systems o Solutions |
| DW | Data Warehouse |
| ERP | Enterprise Resource Planning |
| ETL | Extraction, Transformation and Loading |
| FAP | Filtrar, Agregar y Procesar |
| FIC | Fábrica de Información Corporativa |
| FTP | File Transfer Protocol |
| GB | Giga Bytes |
| IP | Internet Protocol |
| IyT | Integración y Transformación |
| MOLAP | Multidimensional on line analytical processing |
| NAS | Network Attached Storage |
| ODBC | Object DataBase Connectivity |
| ODS | Operational Data Store |
| OLAP | On line analytical Processing |
| PL | Procedural language |
| PPS | Pendiente por surtir |
| RAM | Random Access Memory |
| ROLAP | Relational on line Analytical Processing |
| SAN | Storage Area Network |
| SDVC | Utilizado para distinguir el CVDS pero en sentido inverso |
| SO | Sistema Operativo |
| SQL | Structured Query Language |
| TCP | Transfer communication protocol |
| TI | Tecnologías de información |
| TS | Table Space |
| UNIX | Siglas de sistema Operativo, familia de sistemas operativos |