

Construcción y poblamiento de un datawarehouse basado en el paradigma de bases de datos objeto relacional

Construction and population of a datawarehouse based on the paradigm Of databases relational object

Juan Carlos Calabria Sarmiento¹

1 Mg Ingeniería de Sistemas y Computación, Especialista en desarrollo de Procesos Cognoscitivos. Docente Tiempo Completo Universidad Autónoma del Caribe. Grupo de Investigación SINT. Programa de Ingeniería de Sistemas Universidad Autónoma del Caribe. jcalabria@uac.edu.co

Recibido 09/05/2011, Aceptado 26/06/2011

RESUMEN

El presente artículo presenta la construcción de un sistema de información administrativo desarrollado bajo la tecnología data warehousing dirigida hacia la inteligencia de negocios. Implementando una metodología de diseño de bodega de datos denominada modelo de hechos dimensionales, la cual permite adaptar la funcionalidad de una base de datos multidimensional de soporte para usos de control, independiente de las fuentes de datos que suministran la información histórica.

A partir del modelado de la bodega de datos se sigue un patrón para suministrar la información del sistema mediante la aplicación de los procesos ETL, implementados por medio de la herramienta de diseño gráfico de transformaciones conocida como Kettle. Además, se hace uso de tres motores distintos (Access, PostgreSQL, MySQL) denominadas fuentes aisladas de origen de datos que serán objeto de integración y extracción para nuestra bodega desarrollada en un motor diferente (Oracle 10g Express) bajo el paradigma Objeto-Relacional. Los resultados expuestos servirán como fundamentos para encargados de administrar sistemas de información, así como para altos gerentes y administrativos que deseen implementar estrategias en nuevas tecnologías con el objetivo de lograr mayor eficiencia de acuerdo a las decisiones gerenciales que se dirijan a la obtención de resultados óptimos para las organizaciones.

Palabras clave: Bodega de Datos, Origen de Datos, Procesos ETL, Base de Datos Objeto-Relacional.

ABSTRACT

The present article present the construction of an administrative system of information developed under the technology Data Warehousing directed to the business Intelligence. It is implementing a methodology of design of Data warehouse named dimention fact model, which allows to adapt the functionality of a multidimensional Data Base of support for uses of control, independent of the data sources which give the historical information.

From the shaped one of the data warehouse it is still a boss to give the information of the system by means of the application of the processes ETL, implemented by means of the tool of graphical design of transformations known as Kettle. However, we use three different engines (Access, PostgreSQL, MySQL) named sources isolated of data origin that will be an object of integration and Extraction for our warehouse developed in a different engine (Oracle 10g Express) under the Object-Relational paradigm. The results exposed will serve as foundations for entrusted to administer systems of information, as well as for high managers and clerical that they wish implement strategies on new technologies with the aim to achieve major efficiency of agreement to the managerial decisions that go to the obtaining of ideal results for the organizations.

Key words: Data warehouse, Data Origin, Process ETL, Object-Relational Data Base.

1. Introducción

El crecimiento acelerado del volumen de datos que generan las organizaciones; exige a los administradores de los sistemas de información actuales desarrollar estrategias de análisis dirigidas a la toma de decisiones que beneficien la compañía. Esta actividad se complica aún más, si tenemos en cuenta que en la actualidad los volúmenes de datos que generan los procesos productivos, comerciales y operacionales de una organización sobrepasan el alcance de las estructuras que soportan los sistemas de información convencionales.

Para los encargados de tomar decisiones fundamentales se convierte en una necesidad complementar sus habilidades humanas con nuevas herramientas gerenciales que surgen para la administración de estos robustos sistemas de información; en la medida que se le dé importancia a la información que se ha suministrado respecto a cada uno de los procesos en el pasado, podrán hacerse análisis cuidadosos que darán como resultados estrategias para el óptimo desarrollo de las organizaciones en el futuro.

Pero es aun más exhaustiva esta labor y en muchos casos delicados, cuando la proveniencia de los datos que hacen parte del sistema son inclusive fuentes de información completamente aisladas y hasta obsoletas pero que las organizaciones por distintas razones se ven forzadas a depender de ellas, de acuerdo a los distintos casos citados anteriormente se ha hecho más fuerte el concepto de las llamadas bodegas de datos.

Se define bodega de datos como la integración de datos consolidados, almacenados en un dispositivo de memoria no volátil, proveniente de múltiples y posiblemente diferentes fuentes de datos. Con el propósito del análisis y a partir de este tomar decisiones en función de mejorar la gestión del negocio [1]. Es decir, que el objetivo principal de la data warehousing es construir un conjunto integrado de bases de datos, con orientación temática donde cada unidad de datos es relevante en algún momento del tiempo [2].

En cuanto a los procesos de transformación, extracción y carga de los datos de estas distintas fuentes que serán integradas en nuestra data warehouse se ha venido consolidado en un ordenado grupo de procesos estandarizados denominados ETL (Extracción, Transformación y Carga de datos), por medio de los cuales podemos llevar a cabo los procedimientos necesarios para la adecuada alimentación de los datos históricos de una bodega. El modelo citado sobre la bodega descrita en el presente artículo se denomina Objeto-Relacional el cual se usa para describir una base de datos que ha evolucionado desde el modelo relacional hasta una base de datos híbrido, que contiene ambas tecnologías: relacional y de objetos [3] cuyas ven-

tajas en rendimiento, velocidad de búsqueda y referencia directa de los objetos la tildan de la mejor alternativa para un sistema cuyo crecimiento es de valor continuamente exponencial.

A continuación, se presenta un método optimizado para la construcción y poblamiento de un datawarehouse, cuyo objetivo además, es permitir a las organizaciones la integración de los procesos soportados en sistemas de información de distintas proveniencias que sirvan como soporte de diligencias de control y de toma de decisiones para las personas responsables la actividad gerencial.

2. Metodología

La metodología de construcción de bodegas de datos propuesta se centra en una estructura tecnológica diseñada para que la organización disponga, en forma integrada y estandarizada, de la información correspondiente a la operación de la empresa, así como, proporciona a los usuarios, que tienen a su cargo la toma de decisiones, las herramientas adecuadas, para que a través de consultas rápidas, ellos mismos accedan la información requerida [4].

Para su adecuada estructuración es necesario reconocer un conjunto de operaciones y convertirlas en un sistema orientado al sujeto donde el instante de tiempo en que ocurre la acción es de importancia vital para su análisis, el enfoque inicial se dirige hacia el conocimiento de los distintos orígenes de datos (motores administradores de base de datos, sistemas de archivos etc.) de los cuales serán extraídos los datos que en nuestra bodega se convertirán en información, de estas bases de conocimiento se deben excluir un número de procesos relevantes y entender de estas sus variables principales, así como la interrelación entre sus componentes, a partir de esto se pueden crear una serie de estructuras denominadas dimensiones que serán relacionadas en una tabla de hechos dimensionales que dará soporte a la toma de decisiones.

Además, es primordial que esas dimensiones abstraídas se diseñen bajo el paradigma Objeto-Relacional propuesto para bodegas en este método, luego de su construcción conviene definir las reglas de uso para las transformaciones (Organización, Filtros etc.) que darán lugar a manera conveniente para suministrar la información hacia el data warehouse y será necesario el uso de consultas rápidas para verificar la correcta alimentación del sistema, a fin de dar mejor claridad al proceso, se muestra en la Figura 2 un esquema gráfico de los pasos correspondientes al método propuesto.

2.1 Identificar fuentes de Origen de Datos

En este paso el objetivo fundamental es la identificación formal de las distintas fuentes de datos que servirán de

Figura 1: Sistema desarrollado bajo Data Warehouse.
Figure 1: System developed under the Data Warehouse.

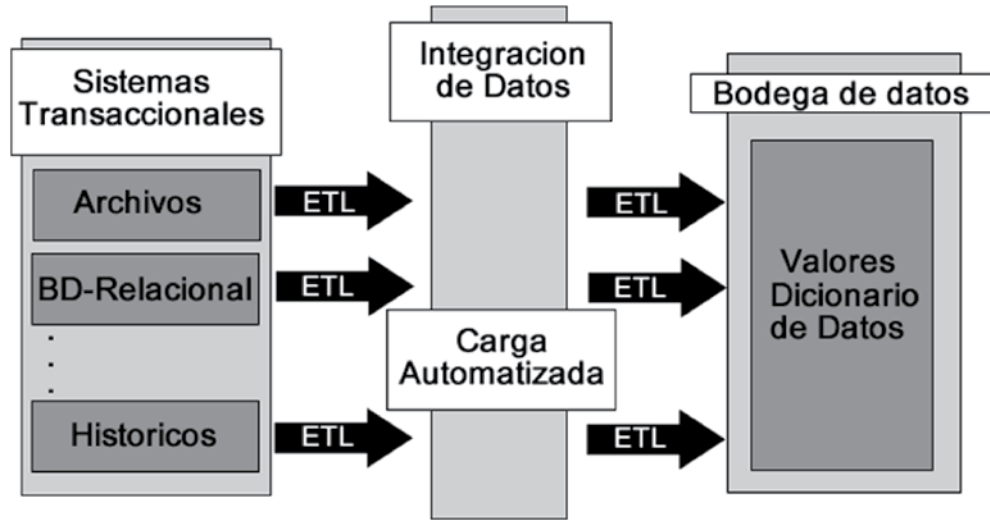
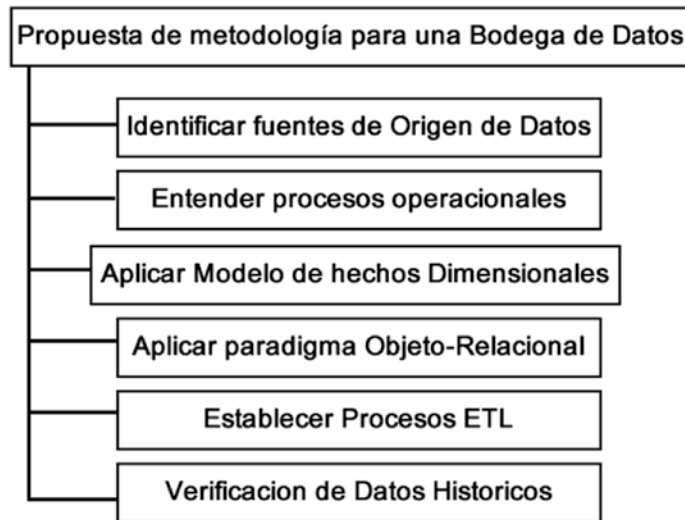


Figura 2: Propuesta de metodología para una Bodega de Datos.
Figure 2: Proposed methodology for a Data Warehouse.



origen para la alimentación y el suministro de la bodega, estas fuentes pueden ser de estructuras tan complejas como motores avanzados de administración de bases de datos, así como sistemas obsoletos y en cierto sentido reemplazados, limitados y de menor eficiencia como serían sistemas de archivos almacenados. Cada organización deberá ofrecer la información necesaria sobre los distintos sistemas de almacenamiento que hacen uso concerniente de las operaciones primordiales de sus procesos.

Para cuestión de comprensión del presente artículo se han identificado tres (3) fuentes de datos para el desarrollo de este caso guía citado, son las siguientes:

1. Access
2. PostgreSQL
3. MySQL

El primero es un programa, utilizado en los sistemas operativos Microsoft Windows, para la gestión de bases de datos creado y modificado por Microsoft y orientado a ser usado en entornos personales o en pequeñas organizaciones [5] por medio de sistemas de archivos.

El segundo es un sistema de gestión de bases de datos que trabaja bajo el paradigma Objeto-Relacional (ORDBMS) basado en el proyecto POSTGRES, de la universidad de Berkeley [6].

El tercero es un sistema de gestión de bases de datos relacional, licenciado bajo la GPL de la GNU. Su diseño multihilo le permite soportar una gran carga de forma muy eficiente [6].

Para incitar a mostrar la diferencia entre las distintas fuentes de datos que pueden usarse en la metodología propuesta se ha hecho necesario la utilización de tres orígenes de distintas procedencias y características incluyendo un limitado sistema de archivos.

2.2 Entender procesos operacionales

Este paso necesita del uso de capacidad de análisis e interpretación para entender la lógica operacional de las relaciones existentes entre los procesos de las bases de datos almacenadas en cada fuente. Es muy necesario que las personas encargadas de dar soporte a cada base de información suministren los modelos adecuados para el entendimiento de los distintos procesos de negocios que hacen

parte de cada sistema. Podrían para la disminución de la complejidad de este paso, recurrir a esquemas gráficos como modelos relacionales que representen claramente la lógica aplicada a cada sistema, esto con el objetivo de abstraer de manera formal las estructuras lógicas que pasarán a ser componentes vitales dentro la jerarquía de la bodega de datos. En la Figura 3 se muestra como el esquema relacional que representa la lógica de la base de datos almacenada en las tres fuentes mencionadas en el paso anterior correspondientes al caso en cuestión.

El esquema permite realizar una descripción de la lógica del negocio de esta base de datos, cuando se pueda describir cada proceso, sus variables y las relaciones que tiene con otros procesos se prepara para definir que estructuras son vitales y cuáles pueden ser derivadas de estas. Tener claro esto será fundamental para la ejecución correcta del siguiente paso.

2.3 Aplicar Modelo de hechos Dimensionales (Diagrama Estrella)

En este paso la importancia radica en construir a raíz de las bases de datos operacionales identificadas y comprendidas en los pasos anteriores nuestra bodega de datos, este proceso de construcción presenta algunos altibajos que podrían hacernos perder mas tiempo del dispuesto y podría ser un paso en falso si no se escoge un modelo adecuado para el diseño de nuestra data warehouse. El modelo de diseño escogido para la estructura de la bodega se denomina modelo de hechos dimensionales, a partir del

cual implementamos la extracción y poblamiento de datos provenientes de las fuentes antes mencionadas.

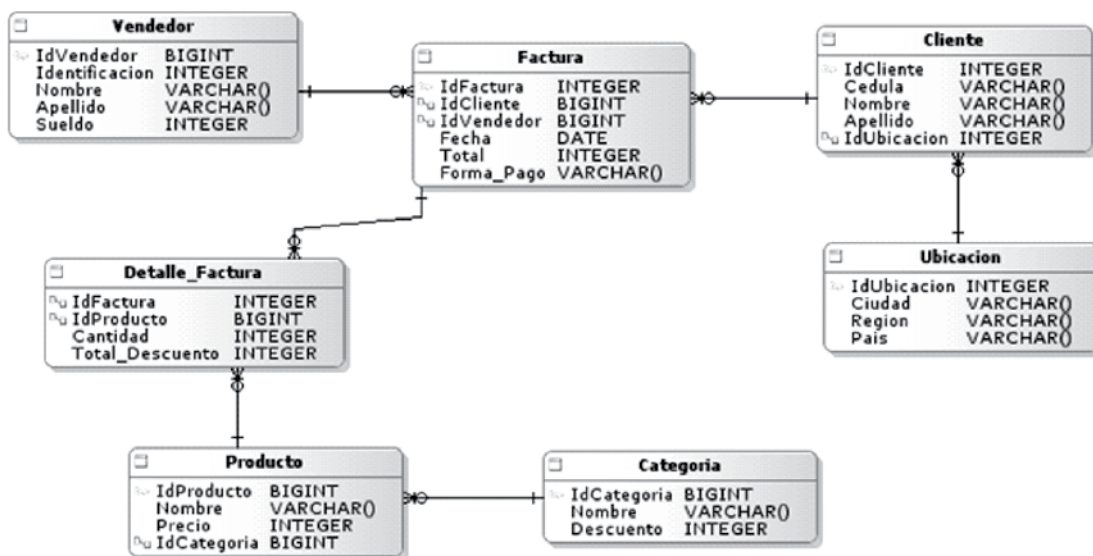
El uso del modelo de hechos dimensionales surge a partir del análisis de los datos en un tiempo finito, que ha traído consigo estudios sobre la mejor forma de almacenar y representar estos datos para que puedan ser consultados de una forma más rápida. El uso del modelo multidimensional es una de las aproximaciones más acertadas y seguidas por los especialistas en estos días. Este se basa en el estudio de los eventos del negocio analizados desde sus distintas dimensiones [7].

Este modelo exige para la elaboración de nuestra bodega de datos el uso distintos componentes como son los hechos, las medidas, las dimensiones cada una de las cuales se describen brevemente a continuación:

Medida es una propiedad de un hecho [7]. Una medida clásica y necesaria para este sistema podría ser el total de la factura, una medida por la cual podrían tomarse decisiones importantes sobre el sistema.

Hecho (Fact) se define como una operación que se realiza en el negocio en un tiempo determinado. Son objeto de análisis para la toma de decisiones. Se representan en una caja con su nombre y las medidas que lo caracterizan [7]. El hecho principal sobre un sistema de facturas y ventas, podría ser el hecho de facturar el cual podría contener las dimensiones del modelo.

Figura 3: Modelo Relacional de las fuentes de origen.
Figure 3: Relational model of the original sources.



Dimensión (Dim) es una característica de un hecho que permite su análisis posterior, en el proceso de toma de decisiones [7]. Entre las dimensiones destacadas para el modelo podemos tener, vendedor, cliente, producto y como todo modelo de DataWarehouse una siempre obligatoria dimensión tiempo que determine el momento en el cual fue desarrollado cada hecho.

2.3.1 Formulación de preguntas dirigidas a la toma de decisiones

La importancia en la creación de nuestra bodega y es lo que le da gran importancia a este artículo radica en el interés de las organizaciones de tomar decisiones sobre los hechos que acontecieron en algún instante de tiempo, para lograr ese objetivo se necesita entonces un buen análisis. Algunos interrogantes basados en el modelo construido sobre el caso citado en este artículo, sería que la compañía quisiera determinar cuáles son los productos más vendidos y los menos vendidos en temporadas especiales como la época de navidad, épocas vacacionales etc. Todo depende de los distintos sistemas para la formulación de preguntas que pueden causar decisiones importantes para una empresa comercial con un sistema como el expuesto en el presente caso:

- ¿Cuáles son los productos más vendidos y menos vendidos en el mes de diciembre?
- ¿Quiénes son los vendedores que mas producen en las épocas de temporadas críticas?

- ¿En qué ciudades se venden la mayor cantidad de cada categoría?

La respuesta a estas preguntas podría ocasionar para la compañía la incorporación de paquetes promocionales de ciertos productos para ciertas ciudades en épocas del año específicas. En este caso esto sería una medida de optimización de gastos de transporte, puesto que no generaría ganancias enviar productos a ciertas ciudades donde no se han vendido en los últimos años o buscar la forma de llamar la atención de ese grupo potencial de clientes para que levanten su atención sobre ese producto.

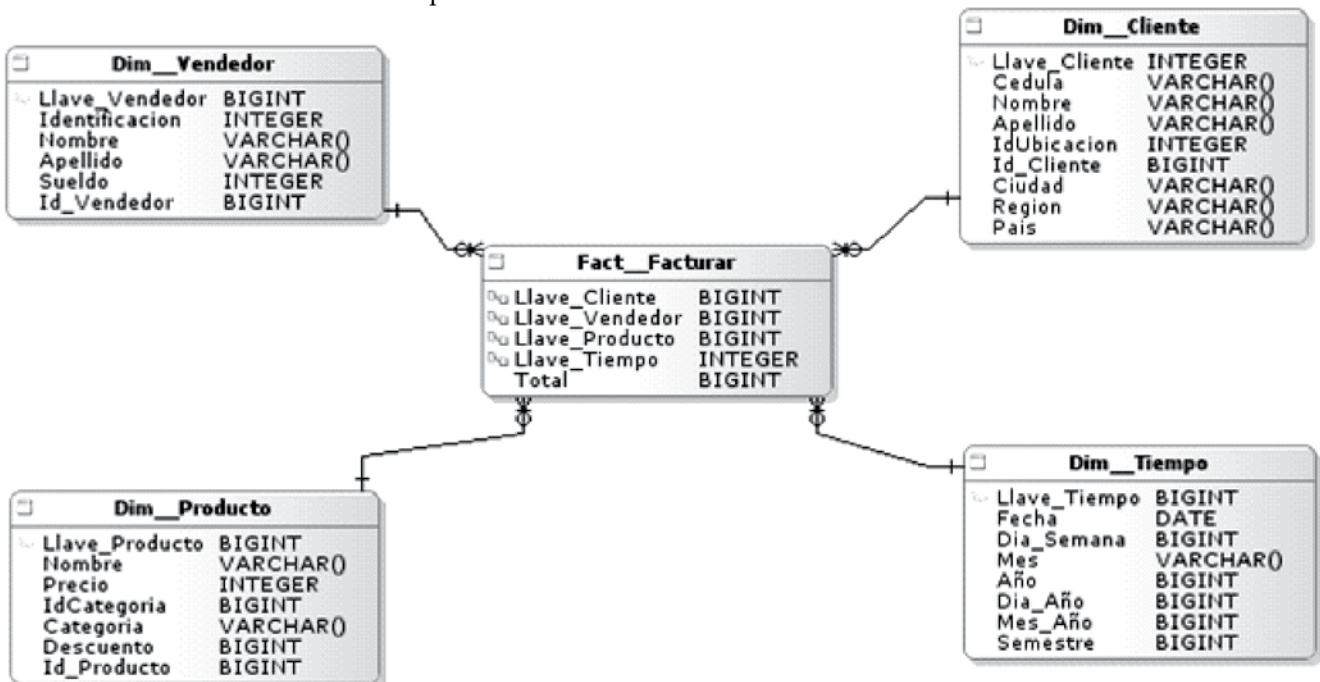
2.4. Aplicar paradigma objeto-relacional

En este paso el objetivo principal es definir la importancia de los modelos objeto-relacional para los nuevos sistemas de millones de datos a los que están siendo acostumbrados los diseñadores de sistemas de información modernos, una de las ventajas que presenta este modelo es que la integración de las representaciones relacional y orientada a objetos es semánticamente limpia y considerablemente más potente en rendimiento y velocidad de búsqueda [3].

En teoría el modelo objeto-relacional elimina las restricciones de llaves foráneas como en los reconocidos esquemas relacionales, en contraste se hacen al uso de referencias en memoria donde cada objeto guarda la dirección en memoria de sus objetos dependientes. El código del objeto que

Figura 4: Ejemplo Modelo de Hechos Dimensionales

Figure 4: Dimensional Fact Model Example



representa la tabla de hechos para la bodega citada en el presente artículo es la figura 5:

Figura 5: Creación del objeto correspondiente a la tabla de hechos de la Bodega.

Figure 5: Creation of the object corresponding to the fact table of the Data Warehouse.

```
Create or Replace Type Fact_Facturar As Object(  
Llave_Cliente Integer,  
Llave_Vendedor Integer,  
Llave_Producto Integer,  
Llave_Tiempo Integer,  
Total Integer  
);|
```

2.5 Establecer Procesos ETL (Extracción, Transformación, Carga)

Hacemos uso de la herramienta kettle, o también conocida como Spoon; implementada para hacer actividades de transformaciones a los datos, realizar trabajos completos sobre nuestras fuentes de origen incluyendo conexiones a distintos motores de base de datos y extracción de datos de archivos almacenados como por ejemplo Access, hojas de cálculo Excel; entre otras. Para el ejemplo citado en el presente artículo hemos de citar algunas de las características de esta herramienta entre las cuales se destacan las siguientes:

1. Las conexiones de base de datos describen el método mediante el cual Kettle se conectará a una base de datos. Se pueden crear conexiones específicas para un trabajo o transformación o guardarlas en el catálogo de Kettle para reutilizar las en múltiples transformaciones o trabajos [8].
2. Un salto conecta un paso de transformación o entrada de trabajo con otro. La dirección del flujo de datos del salto se indica con una flecha en el panel de vista gráfica. Un salto puede estar habilitado o deshabilitado [9].
3. El Access input permite la realización de una entrada de datos almacenados en un archivo creado con el programa Access; para la configuración de este paso es necesario especificar el archivo al cual se hará referencia para la extracción de los datos.
4. La entrada tabla permite disponer de una entrada de datos para una gran cantidad de motores de bases de datos; el asistente de configuración de este paso especifica el motor a conectar y a raíz de esto se especifican un conjunto de datos necesarios como por ejemplo el

nombre del servidor que contiene la base de datos, aumentaciones para el acceso a esta etc.

5. Seleccionar/Renombrar valores nos permite hacer ajustes a los datos extraídos en componentes de entrada de tabla, entre los cambios que deben ajustarse a los datos se encuentra el renombrar los campos, cambiar el tipo de datos para compactarlo de un motor a otro; cambio en la longitud máxima de los campos etc.
6. La salida tabla permite disponer de un componente donde se especifica el destino de los datos transformados o extraídos en pasos anteriores, el asistente de configuración define el motor destino de los datos; en este caso es el motor Oracle 10g donde se encuentra almacenada la bodega.

La calidad de los procesos de extracción, transformación y carga se llevan a cabo sobre los orígenes y destinos de datos; la facilidad para unificar datos de distintas fuentes, realizar modificaciones y suministrar estos datos a una bodega que es independiente de las fuentes[10]. Para las organizaciones una herramienta de base de conocimientos de inteligencia de negocios como esta es de vital importancia si no desean perder la información de archivos almacenados o de bases de datos cuyas capacidades han llegado a su límite superior, este método es una poderosa herramienta no solo para recuperar y administrar estos datos desde una nueva base de datos si no también que se pueden planificar estrategias para el futuro de la organización a base las decisiones que se tomen luego de responder los interrogantes de acuerdo al comportamiento de sus procesos y al resultado de sus datos históricos en cualquier instante de tiempo[11].

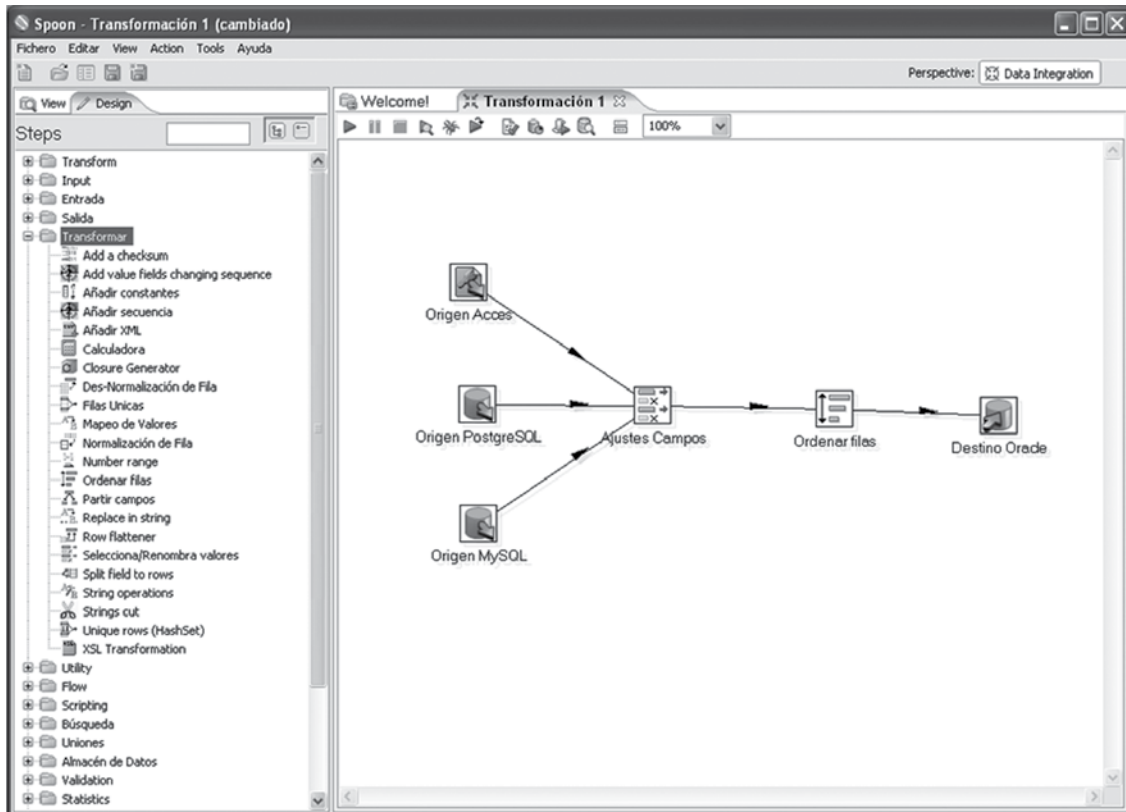
2.6 Verificación de datos históricos

Antes de realizar el análisis de resultados se debe hacer una verificación de los datos en cuestión de coherencia dentro de la DataWarehouse[12].; la herramienta Kettle luego de la ejecución de la transformación muestra una ventana de resultados donde entre otras columnas se muestran las cantidad de filas procesadas en cada uno de los saltos; estos valores deben coincidir con las consultas rápidas dentro de cada uno de las fuentes de origen de datos; además, la suma total debe coincidir con el numero de registros almacenados en nuestra bodega[13].

3. Resultados

Los resultados obtenidos de llevar a cabo un proceso de extracción, transformación y carga de datos como la expuesta en pasos anteriores toman validez en cuanto a dar respuestas concretas a las preguntas formuladas antes de realizar los procesos[14].; un análisis expuesto en tablas

Figura 6: Aplicación de procesos ETL con Kettle (Spoon).
Figure 6: Implementation of ETL processes with Kettle (Spoon).



por medio de consultas rápidas o por medio de la colaboración de una lenguaje de programación que le dé una vista mas amigable a los resultados pueden permitir para las personas encargadas de la toma de decisiones y llevar el control en las organizaciones sacar conclusiones median-

te las cuales fomentar estrategias, planes promocionales, reestructuración del los empleados etc.

Los resultados arrojados representados por medio de tablas del caso citado en el artículo se muestran a continuación:

Tabla 1: Resultados de la Bodega de Datos.
Table 1: Results of the Data Warehouse.

Mas Vendidos	Menos Vendidos	Mes
Sweter Hombre	Kit Escolar niño	Diciembre
Pantalon Hombre	Bolso Escolar	Diciembre
Blusa Niña	Lonchera niño	Diciembre
Blusa Mujer	Lonchera niña	Diciembre
Pantalon Mujer	Block rayado	Diciembre
Zapato Hombre	Carpeta Blanca	Diciembre
Sandalia Mujer	Boligrafo	Diciembre
Sandalia Niña	Kit Escolar niña	Diciembre
Correa Hombre	Resma sin raya	Diciembre
Buso Hombre	Agenda	Diciembre

Tablas 2 y 3: Resultado de la Bodega de Datos.
 Tables 2 and 3: Result of Data Warehouse.

Vendedor	Total Ventas	Mes
Edward Guerrero	\$ 2`600.000	Marzo
Jesus Peña	\$ 2`450.000	Marzo
Alexander Zapata	\$ 2`220.000	Marzo
Luis Henriquez	\$ 2`040.000	Marzo
Jader theran	\$ 1`950.000	Marzo
Vendedor	Total Ventas	Mes
Alexander Zapata	\$ 2`900.000	Agosto
Edward Guerrero	\$ 2`700.000	Agosto
Jader Theran	\$ 2`320.000	Agosto
Roberto Sanchez	\$ 2`150.000	Agosto
Luis Henriiquez	\$ 2`050.000	Agosto

Tabla 4: Resultado de la Bodega de Datos.
 Table 4: Results of the Data Warehouse.

Ciudad	Categoria
Barranquilla	Licores
Bucaramanga	Consumo masivo
Medellin	Consumo masivo
Cali	Consumo masivo
Bogota	Hogar

3.1. Análisis de resultados

De las tablas arrojadas por la bodega de datos pueden hacerse las siguientes conclusiones:

- El mes de diciembre no es rentable para la venta de útiles escolares
- El mes de diciembre es muy rentable para comercialización de ropa, accesorios y zapatería.
- Los vendedores más rentables en los meses críticos (Marzo y Agosto) son Edward Guerrero, Alexander Zapata, Jader Theran y Luis Barraza.
- La ciudad de barranquilla es la ciudad que se ha mantenido liderando la compra de licores.
- La ciudad de Bogotá lidera la compra de productos para el hogar.
- Para el suministro del mes de diciembre se duplicarán los productos de ropa y zapatería y se suministrará la mitad de productos escolares.
- Los vendedores que más venden en meses críticos tendrán mejor porcentaje de bonificación que los demás.
- Los suministros de la línea de alcohol se incrementarán para la ciudad de barranquilla
- Los suministros de la línea de hogar se incrementarán para la ciudad de Bogotá.

4. Conclusiones

- La implementación del modelo de hechos dimensionales facilitó la etapa de investigación y diseño; teniendo en cuenta que otros modelos podrían causar una complejidad mucho mayor en la estructura y administración de nuestra bodega de datos.
- El uso del paradigma objeto-relacional maximizó velocidad de búsqueda y rendimiento de la bodega, esta

ventaja se representa aun con más importancia cuando las bodegas sobrepasan en millares el número de datos almacenados; la cual es su especialidad.

- La herramienta de inteligencia de negocios Kettle fue de vital importancia durante el desarrollo de la investigación, por que sirvió como enlace mediador entre los sistemas de bases de datos origen y la bodega de datos destino, agilizando en gran medida la conexión con las fuentes donde se encontraba un obsoleto sistema de archivos como Access. La metodología expuesta en este artículo simplifica realizar trabajos sobre gran cantidad de bases de datos con esta herramienta.
- En cuanto a los resultados arrojados por el estudio se encuentran decisiones en las cuales interviene mucho dinero de por medio; como transporte de lotes de productos a ciudades donde casi no se venden en ciertas épocas del año; esto puede dar lugar a fortalecer por medio de estrategias promocionales la comercialización de dichos productos para esta parte del mercado.

Referencias

- [1] CHAUDHURI, Surajit. DAYAL, Umeshwar. An Overview of Data Warehousing and OLAP Technology. SIGMOD Record, 2007.
- [2] INMON, William.H. Building The Data Warehouse. QED Press / John Wiley, 2009.
- [3] MORGAN, Kaufmann. Data Mining: Practical Machine Learning Tools and Techniques (Second Edition), 2005.
- [4] WITTEN, Ian. FRANK, Eibe. Data Mining: Practical Machine Learning Tools and Techniques (Second Edition). Morgan Series in Data Management Systems. 2005.
- [5] PEREZ, Cesar. Data Mining, soluciones con Enterprise Miner. Alfaomega , 2006
- [6] STACKOWIAK, Robert. Oracle Data Warehousing, 2007.
- [7] WREMBEL. Robert. CONCILIA, Christian, Data Warehouse and OLAP Concepts, Architectures and Solutions, 2007.
- [8] INMON , William . The Father of Data Warehousing. Inmon Consulting Services. 2007.
- [9] KIMBALL, Ralph. The Data Warehouse Toolkit. John Wiley & Sons. 2006.
- [10] SANCHEZ, Jorge. Data Warehouse and OLAP Concepts, Architectures and Solutions. 2004.
- [11] MCGUFFEY, F. Designing the perfect Data Warehouse. 2008.
- [12] THOMSEN, Erik. Olap Solution. John Wiley & Sons. 2009.
- [13] HUMPHRIES, Mark. Hawkins, Michael. DATAWAREHOUSING Architecture and Implementation. Harris Keru's Enterprise Computing Institute. 2010.

ENLACES:

- [14] <http://wiki.pentaho.com/display/EAIes.Manual+del+Usuario+de+Spoon>