

**Universidad de las Ciencias Informáticas  
Facultad 3**



**Título: Mecanismo de gestión de datos para los  
procesos de agregación y consolidación en el  
subsistema Contabilidad de Cedrux.**

Trabajo de Diploma para optar por el título de  
Ingeniero en Ciencias Informática

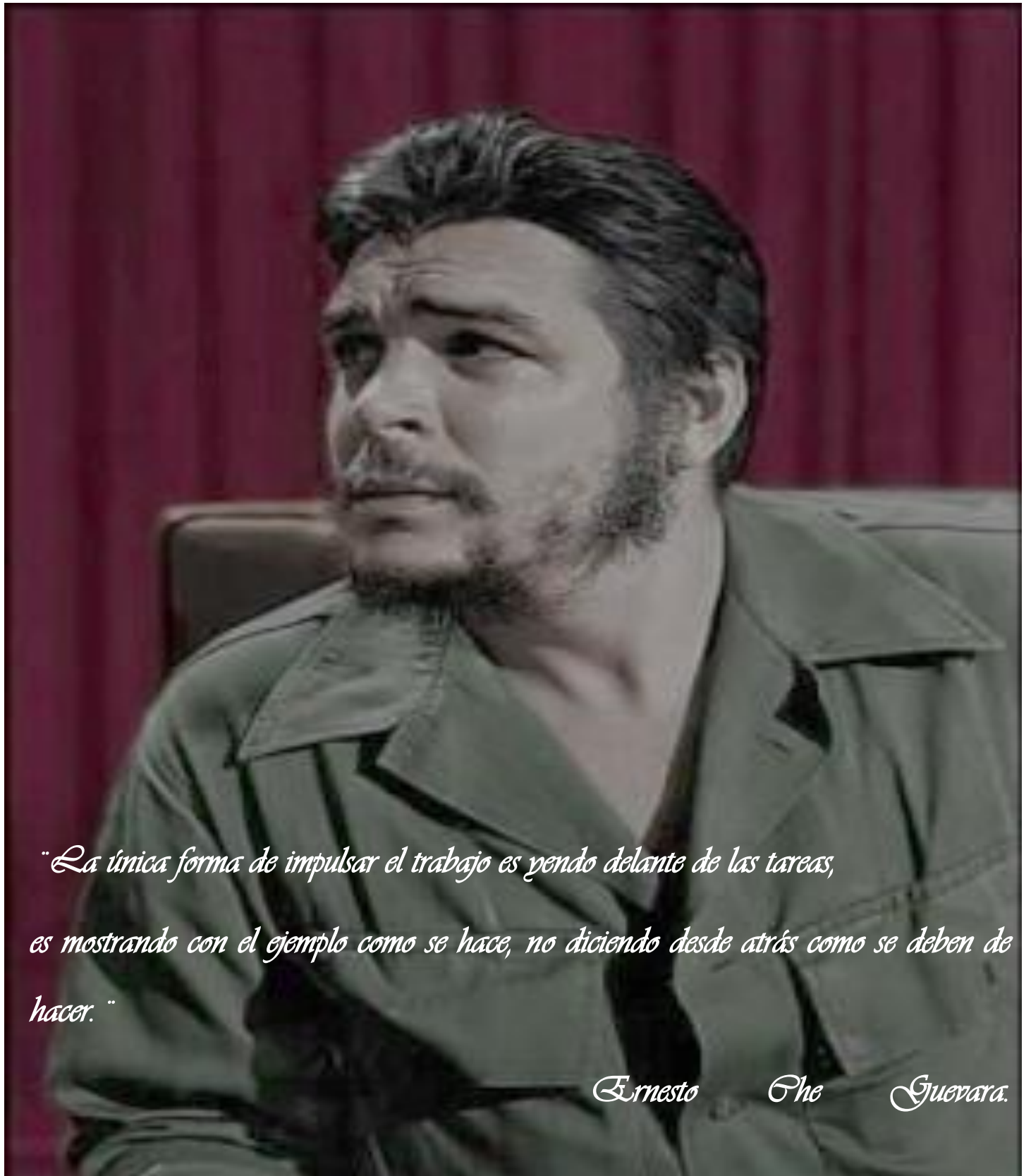
**Autor:** Yamira Antonia Roque Vázquez.

**Tutores:**

Luis Edgardo Mena Rodríguez.

Didier Roque Ginebra.

La Habana, Junio 2012.



*“La única forma de impulsar el trabajo es yendo delante de las tareas,  
es mostrando con el ejemplo como se hace, no diciendo desde atrás como se deben de  
hacer.”*

*Ernesto Che Guevara.*

## DECLARACIÓN DE AUTORÍA

Declaro ser autor de la presente tesis y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los 29 días del mes de junio del año 2012.

### **Autor:**

---

Yamira Antonia Roque Vázquez

### **Tutores:**

---

Ing. Luis Edgardo Mena Rodríguez

---

Ing. Didier Roque Ginebra

## **Tutores:**

✚ **Nombre y apellidos:** Ing. Luis Edgardo Mena Rodríguez.

✚ **Correo electrónico:** [lemena@uci.cu](mailto:lemena@uci.cu)

✚ **Situación Laboral:** Profesor.

✚ **Institución:** Universidad de las Ciencias Informáticas (UCI).

✚ **Dirección:** Carretera a San Antonio de los Baños, Km 2 1/2, Reparto Torrens, Boyeros.

✚ **Rol:** Arquitecto de Datos.

✚ **Nombre y apellidos:** Ing. Didier Roque Ginebra.

✚ **Correo electrónico:** [dginebra@uci.cu](mailto:dginebra@uci.cu)

✚ **Situación Laboral:** Profesor.

✚ **Institución:** Universidad de las Ciencias Informáticas (UCI).

✚ **Dirección:** Carretera a San Antonio de los Baños, Km 2 1/2, Reparto Torrens, Boyeros.

✚ **Rol:** Arquitecto del Sistema.

## **AGRADECIMIENTOS**

Agradecerle a la Revolución por permitirme estudiar en esta maravillosa Universidad.

Agradecer a una de las personas que más admiro y quiero, a nuestro Comandante Fidel Castro, por constituir un guía, ejemplo de lucha y perseverancia para todos los cubanos.

A mis tutores Mena y Didier, en especial a Mena que fue como el compañero de tesis que no tuve, el cual me apoyo en todo momento.

A todos los amigos y compañeros que me apoyaron durante estos inolvidables 5 años.

A Zaida pues fue mi amiga, mi familia y mi mayor apoyo en esta inmensa casa. Gracias por siempre escucharme, aconsejarme y soportar mis malcriadeces.

A Noel, Yosney, Collado, Ariel, Adrian y otras personitas que no dejaron de preocuparse y estar presente en el momento que los necesité.

A mi familia por siempre seguir mis pasos, apoyarme y quererme tanto.

Agradecer principalmente a mis padres, en especial a mi mama y mi nanito por nunca dejar de confiar en mí y tenerme en un pedestal, a quienes les debo todos mis triunfos. Por su inigualable amor, ternura y confianza: GRACIAS MIS AMORES.

## **DEDICATORIA**

A mi mamá: por ser mi musa, mi heroína, lo que más quiero en esta vida; gracias a ella he cruzado tantos obstáculos y luchado incansablemente por ser cada día mejor para que se sienta tan orgullosa de mi, como yo estoy de ella.

Al hermano que más quiero en la vida, a mi Nanito lindo por ser el mejor hermano, compañero y amigo del universo, por amarme y apoyarme tanto. Por quien trabajo tanto para ser un faro que él desee seguir. Por enorgullecerme cada día más.

A mi papá por confiarme en mí. Por mostrarme el valor del sacrificio y el esfuerzo en las condiciones más adversas.

A mi abuelita Olga por siempre estar cerquita cuidándome y apoyándome. Gracias por no abandonarme nunca, TE AMO.

A mi abuelo Negrito por ser como un padre, por cuidarme, aconsejarme y confiar tanto en mí.

A mi tía Rosi, Olguita, a mis primitas, a la niña más linda de la familia (Leidy), a mis tíos, a mi primo y su niñita y a mi abuela Puri.

A Zaida, Evelyn, Osleidys, Aralis y Martica por ser tan especiales y buenas amigas.

A todos mis amigos por estar en los momentos difíciles apoyándome y cuidándome.

## **RESUMEN**

Cuba implementa importantes transformaciones encaminadas al desarrollo de la economía. En este proceso la contabilidad juega un papel protagónico ya que garantiza el control interno administrativo, factor primordial en la economía.

En la Universidad de las Ciencias Informáticas (UCI) se está desarrollando el Sistema Integral de Gestión Cedrux el cual cuenta con un subsistema de Contabilidad. El mismo puede influir en el buen seguimiento y funcionamiento del modelo económico del país, permitiendo tener un conocimiento pasado y presente de las realidades económicas de las entidades; lográndose esto a través de los procesos contables agregación y consolidación de estados financieros.

Actualmente la solución que existe en dicho sistema informático para el proceso de agregación no satisface las necesidades de los usuarios ni permite la consolidación de los datos registrados. En el presente trabajo se propone un mecanismo de gestión de datos que garantice el correcto funcionamiento de ambos procesos, permitiendo a las entidades cubanas mejorar la toma de decisiones.

**PALABRAS CLAVES:** Proceso de agregación, proceso de consolidación, subsistema de Contabilidad.

## TABLA DE CONTENIDOS

INTRODUCCIÓN.....	6
CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA.....	9
INTRODUCCIÓN.....	9
1.1 Estudio de los procesos contables agregación y consolidación. ....	9
1.1.1 Contabilidad.....	9
1.1.2 Sistemas Contables.....	11
1.2 Data Mart.....	13
1.2.1 Procesos realizados para la obtención del Data Mart. ....	14
1.2.2 Modelo.....	16
1.2.3 Herramientas para el rediseño del Data Mart.....	26
1.2.4 Metodologías de desarrollo.....	33
1.3 Replicación de datos.....	38
1.3.1 Herramienta para replicar datos.....	39
1.4 Herramienta de prueba. ....	40
CONCLUSIONES PARCIALES.....	41
CAPÍTULO 2: ANÁLISIS Y DISEÑO DE LA SOLUCIÓN.....	42
INTRODUCCIÓN.....	42
2.1 Análisis, diseño e implementación del Data Mart. ....	42
Fase 1: Análisis de los requerimientos para el Data Mart.....	42
Fase 2: Análisis de los OLTP. ....	46
Fase3: Modelo físico del DM. ....	48
Fase 4: Proceso ETL.....	60
2.2 Solución para la replicación de los datos.....	64
CONCLUSIONES PARCIALES:.....	67
CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA DE LA SOLUCIÓN.....	68
INTRODUCCIÓN.....	68
3.1 Desarrollo del proceso de agregación.....	68
3.2 Validación del mecanismo de datos.....	70
3.2.1 Validación de las transformaciones realizadas al Data Mart del subsistema de Contabilidad. 70	
3.2.2 Validación del proceso de agregación. ....	72



3.2.3 Validación del correcto funcionamiento del entorno de replicación. ....	73
CONCLUSIONES PARCIALES.....	77
CONCLUSIONES .....	78
RECOMENDACIONES .....	79
REFERENCIAS BIBLIOGRÁFICAS.....	80
ANEXOS.....	84
GLOSARIO DE TÉRMINOS .....	95

## INTRODUCCIÓN

En la actualidad los sistemas informatizados han alcanzado un creciente y significativo auge, los mismos permiten que las empresas mejoren linealmente la productividad, calidad de los productos y servicios brindados al cliente. El desarrollo tecnológico marcha de forma ascendente y junto a este la competencia, por lo que cada institución se esfuerza para que su resultado sea el mejor y el más actualizado.

Cuba implementa importantes transformaciones en todos los ámbitos. Los cambios más importantes van encaminados al desarrollo de la economía. En este proceso la contabilidad juega un papel protagónico pues garantiza el control interno administrativo, factor primordial en la economía, además permite disponer de una información financiera veraz, confiable y oportuna, agente clave para dirigir económicamente una entidad.

En la Universidad de las Ciencias Informáticas se está desarrollando el Sistema Integral de Gestión Cedrux. El mismo es un paquete de soluciones integrales de gestión para las entidades presupuestadas y empresariales que respalda los lineamientos nacionales de desarrollo de soluciones bajo los principios de independencia tecnológica y con funcionalidades generales de los procesos y las particularidades de la economía cubana. Garantiza la seguridad, oportunidad y fiabilidad en la gestión de la información de las entidades nacionales y de los diferentes niveles de gobierno. Cedrux provee facilidades para la integración de las diferentes áreas productivas y departamentos administrativos en las entidades, a través de las potencialidades de la tecnología informática, siendo el subsistema Contabilidad General un ejemplo de lo descrito. Este subsistema presenta como objetivo principal: gestionar la información contable, configurar las cuentas, emitir comprobantes de operaciones en moneda contable y original, registrar los asientos de todos los subsistemas de forma tal que permitan agilizar los procesos contables y la emisión de estados financieros. El mismo puede influir en el buen seguimiento y funcionamiento del modelo económico que está siendo implementado en el país, permitiendo tener un conocimiento pasado y presente de las realidades económicas de las entidades; lográndose esto a través de los procesos contables agregación y consolidación de estados financieros. En dicho sistema informático la solución que existe para el proceso de agregación no garantiza la disponibilidad de los datos, pues realiza el mismo de forma desconectada y no guarda el historial de las acciones, además no se efectúa la consolidación de la información existente, dificultando la comparación en el tiempo de su situación financiera, disminuyendo la medición de los servicios que se ofrecen y de la correspondiente evaluación de políticas públicas.

Por lo antes planteado se presenta el siguiente **problema a resolver**: El subsistema de Contabilidad del Sistema Integral de Gestión Cedrux carece de un mecanismo que garantice la disponibilidad y el análisis tendencial de los datos históricos asociados a los procesos de agregación y consolidación.

Para dar solución al problema existente se tiene como **objeto de estudio**: procesos contables de la Contabilidad General, siendo el **campo de acción** los procesos de agregación y consolidación de la Contabilidad General en el Sistema Integral de Gestión Cedrux.

Debido a todos los inconvenientes antes presentados, se tiene como **idea a defender**: realizando un mecanismo de gestión que permita agregar y consolidar los datos del subsistema Contabilidad del Sistema Integral de Gestión Cedrux se garantizará la disponibilidad y los análisis tendenciales de los mismos.

Partiendo del análisis de las posibles soluciones al problema se plantea como **objetivo general**: Desarrollar un mecanismo de gestión de datos para los procesos de agregación y consolidación en el subsistema Contabilidad de Cedrux que garantice la disponibilidad de los mismos. Para darle solución al objetivo general planteado se desglosó los siguientes **objetivos específicos**:

- Elaboración del estado del arte.
- Análisis de la solución.
- Diseño de la solución del mecanismo de datos.
- Implementación de la solución del mecanismo de datos.
- Validar el funcionamiento del mecanismo de datos.

Para el cumplimiento de los objetivos trazados se definieron un conjunto de **tareas** a desarrollar:

1. Realizar un estudio de los procesos contables: agregación y consolidación; a través de los métodos teóricos: analítico-sintético e inductivo-deductivo.
2. Realizar un estudio de las herramientas, tecnologías y metodologías para desarrollar el mecanismo de gestión de datos para los procesos agregación y consolidación en el subsistema Contabilidad del Sistema Integral de Gestión Cedrux; a través de los métodos teóricos: analítico-sintético e inductivo-deductivo.
3. Realizar un estudio del Data Mart, desarrollado para el subsistema Contabilidad; a través de los métodos teóricos: analítico-sintético e inductivo-deductivo.
4. Realizar el análisis de los requerimientos del Data Mart del subsistema Contabilidad; a través de los métodos teóricos: analítico-sintético e inductivo-deductivo.

5. Realizar los modelos lógico y físico a la estructura de la base de datos, para efectuar los procesos de agregación y consolidación; usando el método teórico: modelación.
6. Identificar y especificar los requisitos de replicación del Data Mart; a través de los métodos teóricos: analítico-sintético e inductivo-deductivo.
7. Implementar las transformaciones al Data Mart, desarrollado para el subsistema Contabilidad, para efectuar los procesos de agregación y consolidación.
8. Realizar pruebas al mecanismo de gestión de datos desarrollado, a través del método empírico: experimento.

El presente trabajo está constituido por tres capítulos, conformados de la siguiente forma:

**Capítulo #1 Fundamentación Teórica:** Se centra en el análisis de las principales herramientas, tecnologías y software a utilizar para lograr los objetivos trazados en el trabajo, así como en abordar los conceptos que estén relacionados o puedan influir en la solución del mismo.

**Capítulo #2 Análisis y diseño de la solución:** En este capítulo se realiza el análisis y diseño de las transformaciones pertinentes a efectuar sobre el Data Mart del subsistema de Contabilidad, con el propósito de efectuar los procesos de agregación y consolidación, pasando para esto por cada una de las fases de la metodología a seleccionar. Además se plantea una configuración del software de replicación, para un entorno identificado previamente y cumpliendo con los requisitos trazados.

**Capítulo #3 Implementación y prueba de la solución:** Se realiza la implementación del proceso de agregación. Se ejecutan pruebas a la solución de réplica en el entorno identificado con el propósito de mantener actualizado en los nodos los procesos de agregación y consolidación, gestionándose de esta forma posibles errores en los mismos. Además se efectúan diferentes tipos de pruebas al Data Mart y al proceso de agregación desarrollado, para garantizar que funcionen correctamente.

# CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA

## INTRODUCCIÓN

En el presente capítulo se describe los procesos: agregación y consolidación, se estudian varios sistemas contables y se describen las principales características del proceso de réplica de datos y diseño de Data Mart. Se realiza la investigación y selección de las tecnologías, herramientas y metodologías a utilizar para lograr el diseño y la construcción del mecanismo de datos.

### 1.1 Estudio de los procesos contables agregación y consolidación.

#### 1.1.1 Contabilidad.

Existen varios conceptos de la contabilidad que han sido definidos por diversos autores y cuerpos colegiados de la profesión contable. A continuación se brinda algunos de estos:

*“La ciencia que tiene por objeto el registro de las operaciones económicas efectuadas por una persona o entidad, con el fin de conocer sus resultados y la situación de la misma.”[1]*

*“Ciencia y rama de las matemáticas, que tiene por objeto llevar cuenta y razón del movimiento de las riquezas públicas y privadas con el fin de conocer sus resultados.”[2]*

*“La contabilidad es el arte de registrar, clasificar y resumir en forma significativa y en términos de dinero, las operaciones y los hechos que son cuando menos de carácter financiero, así como el de interpretar sus resultados.”[3]*

*“La contabilidad como polo opuesto a la información empírica, es el medio que, por sus métodos y técnicas, permite controlar y visualizar, a través de estados financieros, información clara, veraz y oportuna de todos los recursos de la entidad y sólo así poder tomar decisiones, conscientes de sus consecuencias.”[4]*

A partir de las definiciones expuestas se puede concluir que la contabilidad se encarga de estudiar, medir y analizar el patrimonio de las entidades, con el propósito de servir a la toma de decisiones y control en las mismas. Esta ciencia social recoge dos procesos claves para evaluar el comportamiento financiero en las empresas, los mismos son: el proceso de agregación y consolidación de la información.

### **Proceso de agregación:**

La contabilidad es una ciencia constituida por un gran cúmulo de información, la misma atraviesa por procesos de: captación de los datos, cuantificación y representación de la misma. Para garantizar que la información obtenida sea útil para la toma de decisiones, es necesario realizar una síntesis de la misma, haciendo uso de los estados contables: balance y cuenta de resultados; que pongan de relieve magnitudes económicas de interés y que sean útiles en el proceso de adopción de decisiones. El anterior proceso descrito se le denomina agregación. [28]

Un ejemplo práctico en el que se describe el proceso que suma contablemente los balances y cuentas de está dado por una entidad "A" y una "B" donde la primera es padre de la segunda. Ver figura 1.

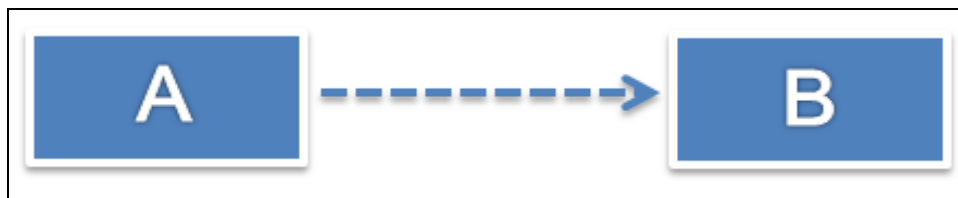


Figura 1: Representa la relación entre la entidad padre "A" y la entidad hija "B".

Descripción de las cuentas	Entidad A		Entidad B		Balance agregado	
	Activos	Pasivos	Activos	Pasivos	Activos	Pasivos
Existencias	450.00		310.00		760.00	
Efectivo en caja	700.00		650.50		1350.50	
Activos financieros	400.00		100.00		500.00	
Fondos para pagos mayores	230.00		100.01		330.01	
Ingreso	300.00		----		300.00	
Capital social		1000.00		400.40		1400.40
Efectivos y cheques por depositar		210.00		350.00		560.00
Fondos para pagos menores		870.00		410.11		1280.11
<b>TOTAL</b>	<b>2080.00</b>	<b>2080.00</b>	<b>1160.51</b>	<b>1160.51</b>	<b>3240.51</b>	<b>3240.51</b>

Tabla 1: Muestra la agregación de las cuentas para las entidades "A" y "B"

### **Proceso de consolidación:**

El proceso de consolidación es una técnica que mediante la agregación de cada una de las partidas y la realización de ajustes y eliminaciones que borren los efectos de las operaciones internas, permite obtener

los estados contables del grupo como si fuera una sola empresa. Los estados financieros consolidados, obtenidos a partir de la consolidación de los datos, son los que presentan la situación financiera y el resultado de la operación de una entidad integrada por la compañía tenedora, controladora o principal y sus subsidiarias. [28]

### **1.1.2 Sistemas Contables**

Con el fin de ganar en conocimiento sobre cómo se desarrollan los procesos de agregación y consolidación en sistemas empresariales, se realizó un estudio del comportamiento de estos en los principales sistemas contables cubanos:

#### **Versat Sarasola**

Sistema integrado de gestión económica, diseñado para ser utilizado por el sector empresarial cubano. Se ajusta a las características de cada entidad. Permite llevar el control y el registro contable individual de todos los hechos económicos que se originan en las estructuras internas de las entidades, obtener los estados financieros, análisis económicos y financieros en estos niveles. Facilita la consolidación de la información agrupando indistintamente las estructuras básicas de la empresa. Concibe la agregación contable a partir de un nivel informativo inferior donde se configuren los grupos de unidades para la exportación agregada. [21]

Presenta el inconveniente de que debe ser utilizado en Windows, que es una aplicación de escritorio y utiliza como sistema gestor de base de datos SQL, por lo que no cumple con el principio de independencia tecnológica que promueve el país.[38]

#### **RODAS XXI**

Sistema multiempresa y multiusuario creado por CITMATEL para la automatización de la gestión empresarial. Contiene diferentes módulos que pueden usarse integrados o independientes: Contabilidad, Efectivo Caja y Banco, Nóminas, Activos Fijos Tangibles, Inventarios, Cobros y Pagos, Facturación, Finanzas, Tele-cobranza. Trabajo solo sobre el sistema operativo Windows, utiliza como sistema gestor de base de datos SQL. Además de no contemplar la doble moneda ni la multimoneda y no presenta mecanismos estándares de integración con otras aplicaciones de su tipo. [39]

Consta de un subsistema de Contabilidad General en el que se realiza la consolidación de los datos, garantizando agrupar información de otras entidades en una sola contabilidad. Este proceso se realiza mediante comprobantes. [22]

## **Subsistema de Contabilidad General de Cedrux.**

Se encarga de gestionar los procesos contables, garantizando procesar la información, gestión de las cuentas, emisión de comprobantes de operaciones en moneda contable y original, registro de asientos, entre otros. Permite agilizar todos los procesos contables, recuperar los reportes mayores, submayores, balance de comprobación de saldos, utilidad acumulada, y comprobante de operaciones en la moneda original y en cualquier otra moneda. [23]

Cuenta con 6 módulos altamente relacionados:

### Configuración

Este proceso permite configurar los parámetros iniciales posibilitando un trabajo más cómodo con el subsistema. Se podrá configurar la estructura económica o entidad a la cual se va a asociar la contabilidad, así como los grupos y subgrupos contables y los contenidos económicos relacionados a estos. Permitirá configurar además el formato de los estados financieros, así como definir las cuentas para registrar los ingresos o gastos por redondeos en las diferencias de las tasas de cambio.

### Nomenclador de cuentas

Permite gestionar las cuentas contables y los niveles de análisis o aperturas necesarios para la entidad, así como los tipos de cuentas.

### Comprobante de operaciones

Garantizará la gestión de los comprobantes de operaciones donde se registran diariamente pases de las cuentas que son afectadas por operaciones contables, además de registrar y contabilizar los comprobantes de operaciones que son generados en los demás subsistemas. Se ordenan los comprobantes de operaciones de la entidad según el subsistema donde se crearon, pudiéndose se realizar las operaciones sobre los comprobantes.

### Análisis de consistencia

Este proceso permite realizar un análisis de cada cuenta contable incluyendo las cuentas de gastos asociadas a los Centros de costos y Elementos de gastos; validando que a las mismas no se le hayan registrado saldos contrarios a su naturaleza; además se podrán actualizar las tablas de resumen de los pases.

### Cierre



El proceso de cierre da por terminada las operaciones contables de un período contable o ejercicio económico en una entidad.

### Recuperaciones

*Mayor:* Muestra por fechas y comprobantes todas las afectaciones que ha tenido una cuenta en un período de tiempo determinado.

*Submayor:* Muestra por fechas y comprobantes todas las afectaciones que ha tenido una subcuenta en un período de tiempo determinado.

*Balance de comprobación de saldos:* Esta funcionalidad muestra un reporte donde se recogen todas las cuentas contables de las entidades y se verifica que el total del saldo debitado sea igual al acreditado, de no ser así existen errores en la contabilización de las operaciones. [19]

Este subsistema efectúa de forma defectuosa la agregación y no permite la realización del proceso de consolidación de los estados contables.

## **1.2 Data Mart**

Un Data Mart es una base de datos con funciones departamentales. Consta de las mismas características que un Data Warehouse (DW) y brinda las mismas facilidades, pero está orientado a una sola actividad y no a satisfacer las necesidades de toda la empresa. Un Data Mart (DM) puede ser alimentado desde los datos de un Almacén de Datos, o integrar por sí mismo un compendio de distintas fuentes de información. Un DM es una colección de datos: orientado a temas, integrado, no volátil y de tiempo variante. [26]

### Ventajas que proporciona:

- Brinda una herramienta para la toma de decisiones, basándose en información integrada y global del negocio.
- Supone una optimización tecnológica y económica en entornos de centro de información estadística.
- Facilita la aplicación de técnicas estadísticas de análisis y modelización para encontrar relaciones ocultas entre los datos del almacén; obteniendo un valor añadido para el negocio de dicha información.
- Permite reaccionar rápidamente a los cambios del mercado.

- Aumenta la competitividad en el mercado.
- Garantiza la capacidad de aprender de los datos del pasado y de predecir situaciones futuras en diversos escenarios.
- Permite al usuario adquirir mayor confianza acerca de sus propias decisiones y de las del resto, y lograr así, un mayor entendimiento de los impactos ocasionados. [26]

### **1.2.1 Procesos realizados para la obtención del Data Mart.**

La construcción de un Data Mart está orientada a la toma de decisiones por lo que un buen diseño de su base de datos (BD) favorece el análisis y la recuperación de datos para obtener una ventaja estratégica; todo esto unido al crecimiento de la complejidad de las BD de los sistemas operacionales y el aumento de los requisitos por parte de los clientes que utilizarán los reportes generados a partir del mismo torna complejo el procedimiento para su implementación, de ahí que sea necesaria la realización de procesos que contribuyan a la creación de este, dichos procesos son:

- Proceso de extracción, transformación y carga de datos (ETL por sus siglas en inglés).
- Proceso de explotación.
- Metadatos.

#### **Proceso de Extracción, Transformación y Carga.**

Este proceso se considera el más importante dentro del ciclo de vida del Data Mart. Es a través de él que se extrae los datos de las fuentes orígenes. Debido a la diversidad de procedencia, los datos contenidos en estas fuentes presentan en la mayoría de los casos distintos formatos y estructuras, entre las fuentes de origen más comunes se encuentran: los archivos de textos, hipertextos, hojas de cálculos, informes, bases de datos transaccionales, entre otros. Es por esto que los datos para poder ser almacenados en el Data Mart necesitan ser agregados, analizados, computados, procesados matemáticamente y pasados por transformaciones específicas. A través del proceso ETL se puede recuperar los datos de todos los sistemas operacionales, permitiendo que la información de los mismos pueda transformarse y moverse desde el sistema operacional u otros sistemas al Data Mart.

Para una mejor explicación del funcionamiento de este proceso, se dividió en 3 subprocesos:

*Proceso de extracción y limpieza de datos:*

En este proceso se procede a extraer la información de las diferentes fuentes orígenes para su posterior limpieza, posibilitando de esta manera eliminar redundancias y los datos inconsistentes, informando los errores cometidos en la extracción o en los sistemas fuentes.

Estas fuentes de origen son las encargadas de alimentar el Data Mart, estando en muchos casos diseñadas para registrar grandes cantidades de transacciones.

Una de las fuentes más comunes es la base de datos operacional (OLTP por sus siglas en inglés), estas son bases de datos que como requisito fundamental poseen el de soportar procesos transaccionales presentando algunas de las siguientes características:

- Son pobladas por usuarios finales.
- Se optimizan en función de los procesos transaccionales.
- Se actualizan constantemente.
- Contienen mucha información de detalle.
- Generalmente se encuentran normalizadas hasta segunda o tercera forma normal.

#### Proceso de transformación:

Luego de ser identificadas las diferentes fuentes de alimentación de datos es necesario realizar la unificación de los mismos, en este marco se define un único tipo de datos para cada uno de los campos al que se tendría que llevar la información de las distintas fuentes de alimentación externas. Para ello se realizan un grupo de funciones entre las que se encuentran: la estandarización de códigos, agregar nuevos atributos calculados, etc.

#### Proceso de carga:

Una vez realizado el proceso de extracción, limpieza y transformación de datos se procede a cargar los mismos en el Data Mart, una vez concluido este proceso la información ya está lista para ser analizada por los usuarios. [27]

#### **Proceso de explotación.**

Estando la información en el Data Mart se puede pasar al proceso de explotación. Mediante el empleo de diferentes herramientas de consulta realiza la extracción y análisis de la información en los distintos niveles de agrupación que se han definido o que el usuario final desee. [27]

## **Metadatos.**

Estos describen los tipos de datos, las definiciones físicas y lógicas de los mismos, las consultas e informes predefinidos, las reglas de validación y negocio, las definiciones de las fuentes de datos, las rutinas de transformación y de proceso. Se refieren a cualquier estructura que define un objeto del Data Mart. Es el punto central de referencia para el diseño, construcción, recolección y control de los datos del Data Mart. [27]

### **1.2.2 Modelo.**

Un modelo es una representación de la realidad que contiene las características generales de algo que se va a realizar. Es la representación en pequeña escala de algo, aunque también se le conoce como un esquema teórico. [29]

El modelado constituye una técnica de ingeniería probada y bien aceptada, que entre otros factores ayuda a:

- Visualizar a los usuarios el producto final.
- Comprender mejor el sistema.
- Comunicar las ideas a otros.

El modelado es útil tanto en pequeños como en grandes sistemas. Mientras más grande y complejo sea el sistema el modelado se hace más importante.

Los modelos se construyen para:

- Comunicar la estructura deseada y el comportamiento del sistema.
- Visualizar y controlar la arquitectura.
- Comprender qué se construye, muchas veces descubriendo oportunidades para la simplificación y reutilización.
- Controlar el riesgo. [30]

En sentido general, los modelos son construidos para comprender mejor el sistema que se está desarrollando. A través del modelado se puede reducir el problema que se está estudiando, posibilitando centrarse en un solo aspecto a la vez.

En la rama de la informática existen varias formas de enfocar un modelo durante el desarrollo de software, en el caso que nos ocupa el modelado de la Base de Datos (BD)<sup>1</sup> es el objetivo a cumplir, para la representación de la misma se utilizan varios de estos modelos.

### **1.2.2.1 Modelo de Datos.**

La descripción de la estructura de una base de datos, es el modelo de datos, es decir, una colección de herramientas conceptuales para describir datos, sus relaciones, semántica y restricciones. También se puede definir como una representación de los detalles relevantes de la realidad. [30]

Un modelo de datos se define como la combinación de tres componentes:

- Una colección de tipos de objetos de información, los cuales son las unidades básicas para construir cualquier base de datos.
- Una colección de reglas generales de integridad, las cuales limitan el conjunto de los tipos de objetos que pueden aparecer en forma legal en cualquier base de datos.
- Una colección de operadores, aplicables a los objetos para obtener información y para otros propósitos.

Los principales objetivos del proceso de modelado de datos es saber identificar cual es el problema y encontrar la forma de representarlo en un sistema. Esto significa saber de los datos, quiénes van a usarlos y cómo van a usarlos. El modelado de datos constituye uno de los elementos más importantes a la hora de iniciar el desarrollo de cualquier proyecto, pues constituye la estructura sobre la que reside la verdadera esencia de la aplicación. El modelado de datos es además una técnica independiente de la implementación de la base de datos.

Cuando un Modelo de Datos describe un conjunto de conceptos de un problema real en términos comprensibles por los expertos del dominio de aplicación, se le llama Modelo de Datos conceptual. Los modelos conceptuales se usan en las primeras fases del diseño de Bases de Datos. [30]

---

<sup>1</sup> *Una base de datos (BD) es una estructura que agrupa un conjunto de información organizada perteneciente a un mismo contexto, es un objeto estructurado. La mayoría de estas tienen formato electrónico, que ofrece un amplio rango de soluciones al problema de almacenar datos. [8]*

## **Modelo Relacional.**

Se basa en una estructura de datos simple y uniforme, la relación. El modelo relacional para la gestión de una BD es un modelo de datos basado en la lógica de predicados y en la teoría de conjuntos. Este Modelo representa la BD como una colección de relaciones. Cada relación representa una tabla. [31]

Características del modelo relacional:

- Son estructuras de datos simples. Consiste en tablas de dos dimensiones donde los elementos son ítem de datos. Esto permite un alto grado de independencia de la representación física de los datos.
- El modelo relacional provee una sólida fundamentación para la consistencia de los datos. El diseño de la BD es asistido por los procesos de normalización que eliminan las anomalías en los datos. Adicionalmente, el estado de consistencia de la BD puede ser uniformemente definido y mantenido a través de reglas de integridad.
- El modelo relacional permite la manipulación de las relaciones. Esta característica puede ser encargada a potentes lenguajes no procedimentales basados en la teoría (álgebra relacional) o en la lógica (cálculo relacional). [32]

### Ventajas

- Su objeto abstracto básico (es decir, la relación) tenga una representación (como tabla) muy sencilla.
- Evitan redundancia de datos.
- Es sencillo y fácil de mantener. [32]

### Desventajas

- Por presentar una visión casi normalizada totalmente, provoca un mayor tiempo de respuesta debido a la cantidad de uniones que se deben realizar entre las diferentes tablas para lograr reunir toda la información que el usuario desea obtener.
- Por la forma que se realiza su diseño resulta complejo establecer de forma clara alguna herencia que pueda existir.
- Divide los datos en muchas relaciones, cada una de las cuales se convierte en una tabla en la BD OLTP.
- Se torna en un modelo asombrosamente complejo generalmente constituido por cientos de tablas. [32]

## **Modelo Multidimensional.**

La modelación dimensional permite hacer simples y comprensibles BD, la cual puede ser visualizada como un “cubo” varias dimensiones, donde cualquier punto interior es una intersección de las coordenadas definidas por los ejes del cubo. [33]

Para poder entender en qué consiste el modelo multidimensional, se deben estudiar algunos conceptos importantes:

Un *cubo* es la unidad de representación de la información, equivalente a las tablas de las bases de datos relacionales, el cual representa la información en forma de matriz. A los ejes de la misma se les llama dimensiones, representando los criterios de análisis, y a los datos almacenados se les denominan medidas, y representan los indicadores o valores a analizar.

Las *dimensiones* son entidades del negocio respecto de la cual se deben calcular las métricas, caracterizan a una actividad o hecho, al cual le brindan una perspectiva adicional. Ejemplos: clientes, productos, tiempo, etc. [34]

Existen tres tipos de dimensiones:

- Dimensiones estándar.
- Dimensiones tiempo.
- Dimensiones Padre –Hijo.

### *Dimensiones estándar*

Es el tipo más sencillo de dimensión, tiene un conjunto fijo de jerarquías y niveles establecido por el diseñador.

### *Dimensiones tiempo*

El tiempo es parte implícita de la información que contiene el Data Mart. Se puede definir separándola en distintas jerarquías de tiempo:

- Año
- Semestre
- Mes

### *Dimensiones Padre –Hijo*

Se basan en relaciones recursivas existentes en una tabla de dimensión. Por ejemplo, un empleado tiene un jefe. Pero este jefe es también un empleado.

Se llama *hecho* a un evento específico u operación que se realiza en el negocio. Se encuentra estrechamente relacionado con el tiempo y constituye la unidad fundamental de análisis de datos para la toma de decisiones. Ejemplos: ventas, exportaciones, inversiones, etc. [34]

La estructura que forman los hechos y las dimensiones puede verse en la siguiente. Figura 2.

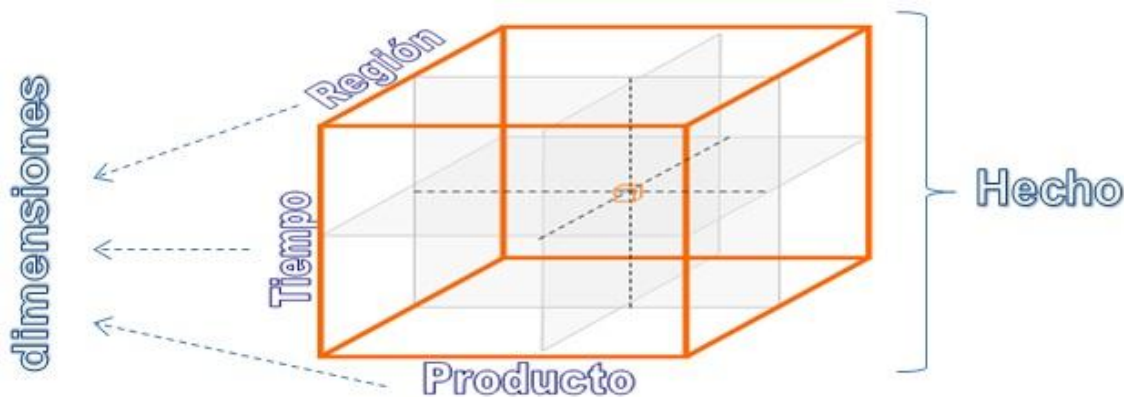


Figura 2: Cubo OLAP

La principal característica del modelo dimensional es su sencillez, permitiéndoles a los usuarios una fácil comprensión de las bases de datos. [35]

### Elementos del Modelo Multidimensional

Para una correcta comprensión y construcción de un modelo multidimensional se explicaran cada uno de los elementos teóricos relevantes que lo componen.

#### Tablas de dimensiones

Las dimensiones organizan los datos en función de un área de interés para los usuarios. Cada dimensión describe un aspecto del negocio y proporciona el acceso intuitivo y simple a datos. Una dimensión provee al usuario de un gran número de combinaciones e intersecciones para analizar datos.

Las tablas de dimensiones son las compañeras de las tablas de hechos. Cada dimensión se define por su clave primaria que sirve para mantener la integridad referencial en la tabla de hechos con la que se relaciona. Un cubo requiere que se defina al menos una dimensión en su esquema.



Cada nivel de una dimensión debe corresponderse con una columna en la tabla de la dimensión. Los niveles se ordenan por grado de detalle y se organizan en una estructura jerárquica. Cada nivel contiene miembros, los miembros son los valores de la columna que define el nivel.

Poseen una columna clave, comúnmente auto-generada y columnas que almacenan los identificadores de cada registro en sus sistemas de origen. [36]

### Tablas de hechos

Las tablas de hechos son las tablas primarias del modelo dimensional conteniendo los valores del negocio que se desea analizar así como cada una de las llaves de las dimensiones involucradas en el mismo. Todas estas columnas son valores numéricos calculables durante el proceso extracción, transformación y carga de datos.

### Medidas o métricas

El modelo dimensional divide el mundo de los datos en dos grandes tipos: las medidas y las dimensiones de estas medidas. Las medidas, siempre son numéricas, se almacenan en las tablas de hechos y las dimensiones que son textuales se almacenan en las tablas de dimensiones. Las medidas son los valores de datos que se analizan.

Una medida es una columna cuantitativa, numérica, en la tabla de hechos. Las medidas representan los valores que son analizados, constituyendo:

- Valores que permiten analizar los hechos.
- Bases a partir de las cuales el usuario puede realizar cálculos.

Las medidas pueden clasificarse en:

- Naturales
- Calculadas

*Medidas naturales:* Son el resultado de la aplicación de operaciones matemáticas sencillas a un solo campo existente en la tabla de hechos. Cuando se define una medida se debe tener en cuenta cual será la forma de agregación (agrupación de la misma) al subir por la estructura dimensional.

*Medidas calculadas*: son el resultado de las diferentes operaciones que se pueden realizar con los valores de las medidas naturales. Debe tenerse en cuenta que estas medidas calculadas se pueden obtener durante el proceso ETL y después del mismo.

En sentido general la expresión medidas calculadas es muy amplio y engloba a cualquier manipulación de las medidas naturales que faciliten el análisis de los hechos, pero por lo general en una medida calculada puede haber:

- Cálculos Matemáticos.
- Expresiones condicionales.
- Alertas.

Estos tres tipos (cálculos, condiciones y alertas) usualmente pueden existir juntos dentro de la misma medida calculada. [36]

### **Estilos de diseño de las Bases de Datos Multidimensionales**

Para un mayor análisis y comprensión de la información que se desea insertar en el Data Mart existen tres estilos de diseño para la construcción de los mismos. A continuación se explican en detalle cada uno de ellos:

#### Estilo de estrella

Esta estructura es considerada la más básica y sencilla de diseñar con respecto a las existentes. Está definida por una tabla central o tabla de hechos y un conjunto de tablas organizadas alrededor de ésta, denominadas tablas de dimensiones, brindando la sensación de una estrella.

En el estilo estrella cada dimensión está compuesta por una sola tabla. Cada tabla de dimensión se vincula con la tabla de hechos por un identificador. Dado que las tablas de dimensión no se encuentran normalizadas se logrará en el modelo del Data Mart, una menor cantidad de tablas. Ver Figura 3.

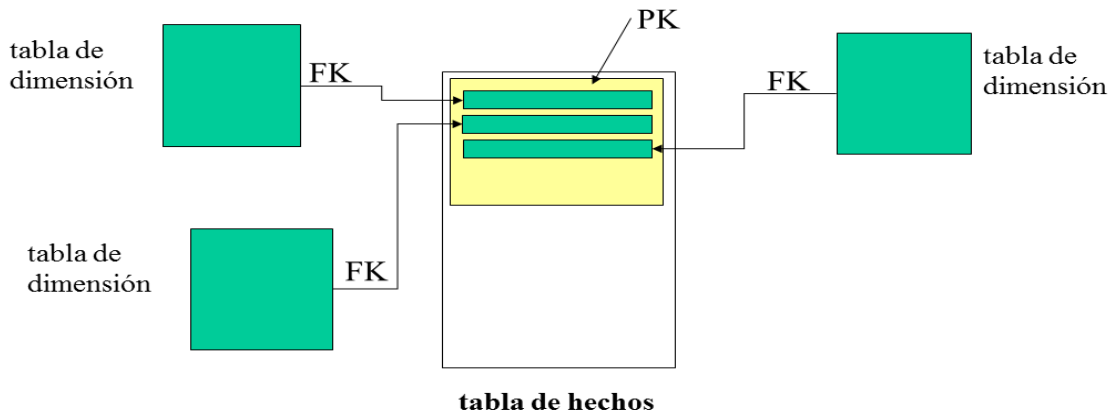


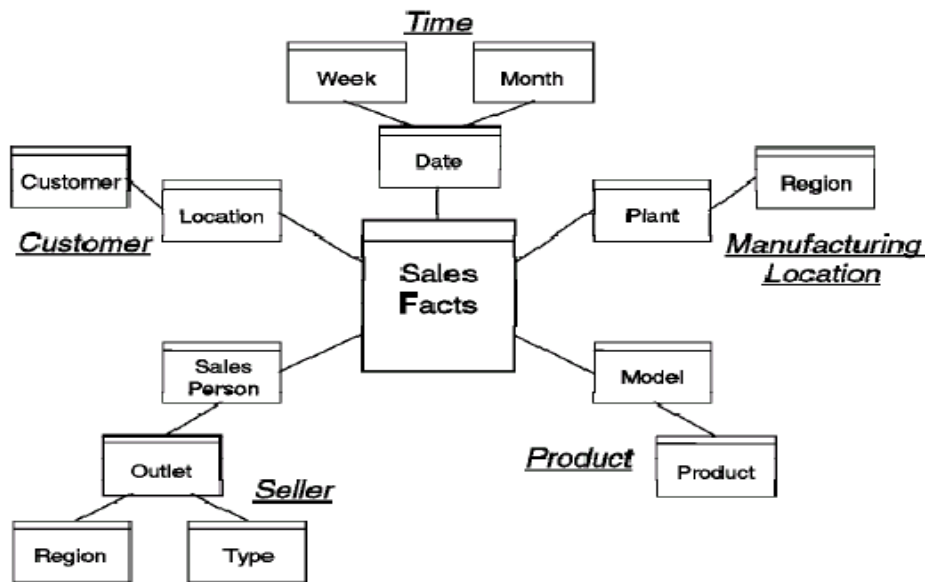
Figura 3: Estilo de estrella

Las principales características del estilo de estrella son:

- El centro de la estrella es la tabla de hechos.
- Los puntos de la estrella son las tablas de dimensiones.
- Cada esquema está compuesto por una sola tabla de hechos.
- Mayor facilidad de mantenimiento con relación a los otros estilos multidimensionales. [27]

#### Estilo copo de nieve

En este estilo, algunas de las tablas de dimensión se encuentran normalizadas para eliminar redundancia de datos. A diferencia del estilo estrella, los datos de las dimensiones se reparten en múltiples tablas. Como ventaja principal de este estilo es el ahorro de espacio de almacenamiento en disco, pero en contrapartida un aumento en la cantidad de tablas y disminución del rendimiento para la recuperación de información. Ver Figura 4.



*Figura 4: Estilo copo de nieve*

Las principales características del estilo copo de nieve son:

- Generalmente es un esquema parcialmente normalizado, aunque pudiera estarlo totalmente.
- Los distintos niveles se encuentran almacenados en tablas separadas.
- Se argumenta ahorro de espacio.

#### Estilo constelación de hechos

El estilo constelación de hechos, referencia las situaciones en que un único modelo multidimensional posee múltiples hechos, y por lo tanto, crea múltiples estilos estrellas, básicamente esta estructura es requerida cuando los hechos no comparten todas las dimensiones. [33]

Este esquema es más complejo que las otras arquitecturas debido a que contiene múltiples tablas de hechos. Con esta solución las tablas de dimensiones pueden estar compartidas para más de una tabla de hechos.

El estilo de constelación de hechos posee mucha flexibilidad y ésta es su gran virtud; sin embargo, el problema es que cuando el número de las tablas vinculadas aumenta, la arquitectura puede llegar a ser muy compleja y difícil de mantener.

En un estilo de constelación de hechos, las distintas tablas de hechos están asignadas a las dimensiones relevantes para cada uno de ellos. Esto puede ser útil cuando los hechos están asignados a un nivel de una dimensión y los otros hechos a otro nivel de detalle de otra dimensión. Ver Figura 5. [37]

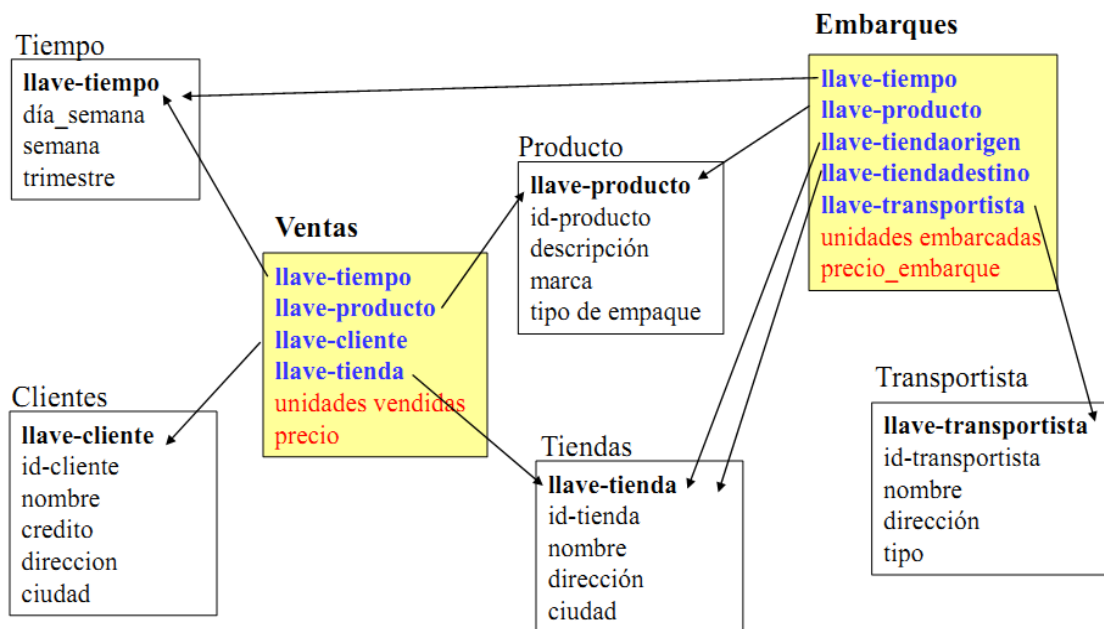


Figura 5: Estilo constelación de hechos

Ventajas del Modelo Multidimensional:

- Por presentar una visión de los datos no normalizada en su totalidad agiliza el proceso de recuperación de datos disminuyendo en gran medida el tiempo de respuesta del sistema.
- Agrupa los datos en pocas dimensiones, cada una de las cuales se convierte en una tabla en la BD.
- Se torna en un modelo simple, consistente de pocas tablas.
- No es simétrico, pudiéndose determinar cuál tabla es más importante.
- Se puede apreciar con facilidad qué/cuáles tablas del modelo contienen medidas numéricas.
- Para llegar a la información que el usuario desea obtener de una consulta donde se encuentren involucradas dos tablas solo existe un único camino de posibles conexiones entre ellas.
- Para consultas, que alcanzan muchos registros y muchas tablas, los modelos multidimensionales

son muy fáciles de comprender por los usuarios y muy sencillos de recorrer por el software. [41]

### **1.2.3 Herramientas para el rediseño del Data Mart.**

#### **Sistemas Gestores de Base de Datos.**

Un Sistema Gestor de Bases de Datos (SGBD) surge para el manejo de los datos en una BD, permitiendo almacenarlos y posteriormente acceder a estos de forma rápida y estructurada. Consta de una colección de programas cuyo objetivo es servir de interfaz entre la base de datos, el usuario y las aplicaciones. Se compone de un lenguaje de definición, manipulación y consulta de los datos. Un SGBD permite definir la información a distintos niveles de abstracción y manipular dichos valores, garantizando la seguridad e integridad de los mismos. Algunos ejemplos de SGBD son Oracle, DB2, PostgreSQL, MySQL, SQL Server. [9]

#### Características de los Sistema Gestor de Base de Datos:

- Abstracción de la información.
- Independencia.
- Redundancia mínima.
- Consistencia.
- Seguridad.
- Integridad.
- Respaldo y recuperación.
- Control de la concurrencia. [9]

#### **PostgesSQL 8.3**

PostgreSQL es un Sistema de Gestión de BD relacional orientada a objetos, se distribuye como software libre. Consta de un código extensible que permite a los grupos de desarrollo hacer mejoras dentro de él.

#### Características de PostgreSQL como herramienta:

- Soporte SQL92/SQL99: PostgreSQL implementa un subconjunto extendido de los estándares SQL92 y SQL99.

- Transacciones: Permiten el paso entre dos estados consistentes manteniendo la integridad de los datos.
- Integridad referencial: PostgreSQL soporta integridad referencial, la cual es utilizada para garantizar la validez de los datos de la base de datos.
- Bloqueos de tabla y filas: PostgreSQL ofrece varios modos de bloqueo para controlar el acceso concurrente a los datos en tablas. Algunos de estos modos de bloqueo los adquiere PostgreSQL automáticamente antes de la ejecución de una declaración, mientras que otros son proporcionados para ser usados por las aplicaciones.
- Soporte de tipos y funciones de usuario: PostgreSQL soporta operadores, funciones métodos de acceso y tipos de datos definidos por el usuario. Incorpora una estructura de datos: arreglos.
- Conectividad TCP/IP, JDBC y ODBC.
- Interfaz con diversos lenguajes: C, C++, Java, Delphi, Python, Perl, PHP, Bash.
- Restricciones y disparadores: Tienen la función de mantener la integridad y consistencia en la BD. Ejecución de acciones antes o después de un evento de BD.
- Múltiples tipos de datos predefinidos: Como todos los manejadores de bases de datos, PostgreSQL implementa los tipos de datos definidos para el estándar SQL3 y aumenta algunos otros.

### Ventajas

- Estable.
- Alto Rendimiento.
- Flexibilidad.
- Se puede extender su funcionalidad.
- Gran Compatibilidad.
- Permite crear o migrar aplicaciones desde Access, Visual Basic, Visual Fox Pro, Visual C/C++, Delphi para usar PostgreSQL como servidor de DB's.
- Varias Interfaces de Programación: ODBC, JDBC, C/C++, SQL Embebido, Tcl/Tk, Perl, Python, PHP.

### **Arquitectura cliente/servidor.PgAdmin III**

PgAdminIII es una aplicación gráfica para gestionar el gestor de base de datos PostgreSQL, siendo la más completa y popular. Programada en lenguaje C++ usando la librería gráfica multiplataforma wxWidgets, lo que permite que se pueda usar en Linux, FreeBSD, Solaris, Mac OS X y Windows. Incluye un editor SQL con resaltado de sintaxis, un editor de código de la parte del servidor, un agente para lanzar scripts programados, soporte para el motor de replicación Slony-I. La conexión al servidor puede hacerse mediante TCP/IP3 o Unix Domain Sockets (en plataformas \*nix), y proporciona la encriptación mediante SSL4 para mayor seguridad. Se distribuye libremente bajo licencia de tipo GNU5 separadamente de PostgreSQL. [14]

#### Características de la herramienta:

- Esquema de navegación de todos los objetos de PostgreSQL.
- Diálogos de creación y propiedades de objetos (usuarios, tablas, base de datos, disparadores, etc.).
- Herramienta de edición/visualización de tablas.
- Habilidad para navegar y conectarse a múltiples servidores a la vez.
- Interfaz de usuario intuitiva y traducida a más de 20 idiomas.
- Muestra la estructura de la base de datos y todos los detalles de los objetos contenidos en la misma.
- Se puede controlar o administrar los usuarios de las base de datos, manejando los privilegios, usuarios, grupos y contraseñas.
- Permite llevar un control sobre el estado del servidor de base de datos.
- Posee una herramienta avanzada para consultas.
- Tiene una herramienta de Mantenimiento que ejecuta tareas como: reconstruir las estadísticas de las base de datos y tablas, limpiar o eliminar los datos sin usar y reorganizar los índices.
- Permite sacar copias o respaldos de las base de datos y restaurarlas haciendo uso de las herramientas pg\_dump y pg\_restore de PostgreSQL.
- La ventana del “estado del servidor” muestra los usuarios actualmente conectados, los bloqueos y características del servidor seleccionado. [15]



## **Herramientas CASE.**

Las herramientas de Ingeniería de Software Asistida por Computadoras (CASE por sus siglas en inglés) están destinadas a aumentar la productividad en el desarrollo de software, reduciendo el costo del mismo en términos de tiempo y de dinero. [40]

### **Visual Paradigm 6.4**

Visual Paradigm es una herramienta UML<sup>2</sup> profesional que soporta el ciclo de vida completo del desarrollo de software: análisis y diseño orientados a objetos, construcción, pruebas y despliegue.

#### Dentro de sus características:

- Multiplataforma.
- Interoperabilidad.
- Modelado de los Requisitos.
- Colaboración de Equipo.
- Generación de Documentación.
- Editor de Detalles de Casos de Uso.
- Ingeniería de Código.
- Modelado de Procesos de Negocio.
- Integración con Entornos de Desarrollo.
- Modelamiento de base de datos.

#### Ventajas:

- Diseño centrado en casos de uso y enfocado al negocio que genera un software de mayor calidad.
- Uso de un lenguaje estándar común a todo el equipo de desarrollo que facilita la comunicación.
- Capacidades de ingeniería directa (versión profesional) e inversa.
- Modelo y código que permanece sincronizado en todo el ciclo de desarrollo.

---

<sup>2</sup> UML: Lenguaje Unificado de Modelado

- Disponibilidad de múltiples versiones, para cada necesidad.
- Disponibilidad en múltiples plataformas. [16]

Una desventaja de esta herramienta es que posee licencia privada, pero la Universidad de las Ciencias Informáticas posee dicha licencia. [17]

### **Plataforma de inteligencia de negocio.**

Las aplicaciones de inteligencia de negocio han evolucionado en muchas direcciones, debido al crecimiento exponencial de la información. Las mismas son herramientas de soporte de decisión, que permiten en tiempo real: acceso interactivo, análisis y manipulación de información crítica para las empresas. Un ejemplo de dichos software es:

#### **Pentaho.**

Pentaho es una plataforma de inteligencia de negocio, orientada a la solución y centrada en procesos. Iniciado por una comunidad de código libre. El proyecto abarca las siguientes áreas de reporte:

- Pentaho Reporting.
- Pentaho Analysis.
- Pentaho Dashboards.
- Pentaho Data Integration.
- Data Mining.
- Pentaho BI Platform.

Entre sus componentes fundamentales se encuentran:

- Plataforma 100% J2EE, asegurando la escalabilidad, integración y portabilidad.
- Servidor: puede correr en servidores compatibles con J2EE como JBOSS AS, WebSphere, Tomcat, WebLogic y Oracle AS.
- Base de datos: vía JDBC, IBM DB2, Microsoft SQL Server, MySQL, Oracle, PostgreSQL, NCR Teradata, Firebird.
- Sistema operativo: no hay dependencia. Lenguaje interpretado.

- Lenguaje de programación: Java, Javascript, JSP, XSL (XSLT/XPath/XSL-FO).
- Interfaz de desarrollo: Java SWT, Eclipse, Web-based.
- Repositorio de datos basado en XML.
- Todos los componentes están expuestos vía Web Services para facilitar la integración con Arquitecturas Orientadas a Servicios (SOA). [18]

## **Herramientas de integración de datos.**

### **Pentaho Data Integration**

Dentro de la suite de Pentaho, se encuentra la herramienta: Pentaho Data Integration (PDI por sus siglas en inglés). Muchas organizaciones tienen información disponible en aplicaciones y base de datos, pero de forma dispersa. La herramienta de integración limpia e integra esta información y la pone en manos del usuario. Provee una consistencia, una sola versión de todos los recursos de información. Facilita una poderosa extracción, transformación y carga de los datos. Además garantiza evitar grandes cargas de trabajo manual frecuentemente difícil de mantener y de desplegar. [10]

Las características básicas de esta herramienta de integración son:

- Entorno gráfico de desarrollo.
- Código Libre, sin coste de licencia.
- Uso de tecnologías estándar: Java, XML, JavaScript.
- Fácil de instalar y configurar.
- Permitir conexión a más de 18 gestores.
- Multiplataforma: Windows, Macintosh, Linux.
- Basado en dos tipos de objetos: Transformaciones (colección de pasos en un proceso de extracción, transformación y carga) y trabajos (colección de transformaciones). [18]

Pentaho Data Integration está formada por cuatro herramientas:

- Spoon: Interfaz gráfico que permite diseñar transformaciones.
- Pan: Programa que ejecuta transformaciones diseñadas por Spoon en XML o en un almacén.

Normalmente las transformaciones se programan en lotes para su ejecución automática a intervalos regulares.

- Chef: Interfaz gráfico que permite diseñar trabajos.
- Kitchen: Programa que ejecuta trabajos diseñados por Chef en XML o en un almacén. Normalmente los trabajos se programan en lotes para su ejecución automática a intervalos regulares. [10]

## **Lenguaje de Consulta.**

### **El Lenguaje de Consulta Estructurado.**

El Lenguaje de Consulta Estructurado (SQL por sus siglas en inglés) es el lenguaje estándar ANSI/ISO1 de definición, manipulación y control de base de datos relacionales. Es un lenguaje declarativo: sólo hay que indicar qué se quiere hacer. En cambio, en los lenguajes procedimentales es necesario especificar cómo hay que hacer cualquier acción sobre la base de datos. Es un lenguaje muy parecido al lenguaje natural.

#### Características:

Lenguaje de acceso a base de datos relacionales: Permite la especificación de distintos tipos de operaciones entre éstas. Posibilita la realización de consultas para recuperar información de las base de datos de forma sencilla.

Lenguaje declarativo de alto nivel: Maneja conjuntos de registros y no registros individuales, ofrece una elevada productividad en la codificación y en la orientación a objetos.

Lenguaje de manipulación de datos (DML): Contribuye a la gestión de la base de datos a través de tareas de consulta. [13]

#### Ventajas

- Independencia de los fabricantes.
- Portabilidad a cualquier tipo de plataforma.
- Estandarizado.
- Basado en el modelo relacional.

- Lenguaje de alto nivel.
- Múltiples vistas de los datos.
- Definición dinámica de datos.

#### **1.2.4 Metodologías de desarrollo.**

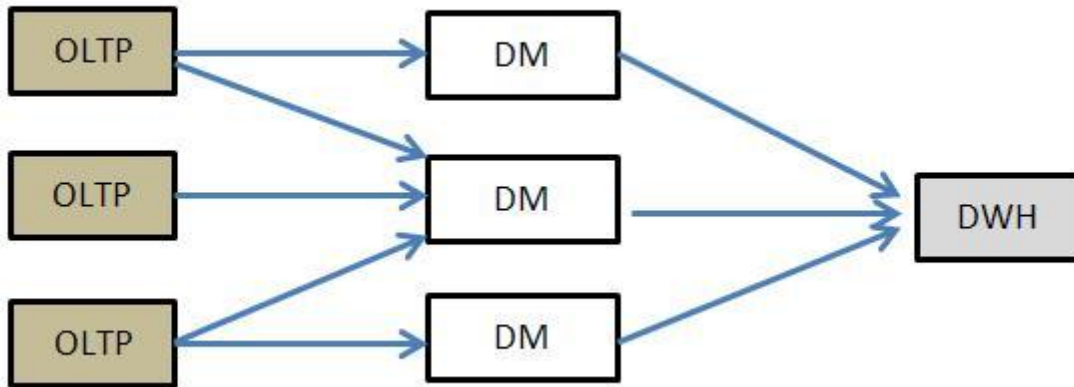
Para la construcción de un DM es necesario tener en cuenta dos tipos de metodologías, para el diseño y la implementación.

##### **Metodología para el diseño de la solución.**

Esta metodología se centra en definir la estructura y diseño del DM. La selección de una de estas en particular depende de las características del proyecto a desarrollar. Dentro de los máximos representantes de este tipo de metodología se encuentra: Ralph Kimball y Bill Inmon, los que plantean diferentes puntos de vistas para el diseño de los mismos.

##### **Metodología de Ralph Kimball**

Kimball señala que: “El Data Warehouse (DW) es la unión de todos los Data Marts de una entidad”, siendo además una copia de los datos transaccionales, estructurados de una forma especial para realizar su análisis, de acuerdo al modelo dimensional no normalizado. Este enfoque también es conocido como Bottom-up. Esta característica le hace más flexible y sencillo de implementar, pues se puede construir un Data Mart como primer elemento del sistema de análisis, y luego ir añadiendo otros que compartan las dimensiones ya definidas o incluyan otras nuevas. En este sistema, los procesos Extracción Transformación y Carga (ETL por sus siglas en inglés) extraen la información de los sistemas operacionales y los procesan, realizando posteriormente el llenado de cada uno de los Data Mart de una forma individual, aunque siempre respetando la estandarización de las dimensiones. Este enfoque conduce a una solución completa en período de tiempos cortos. (Figura 6)

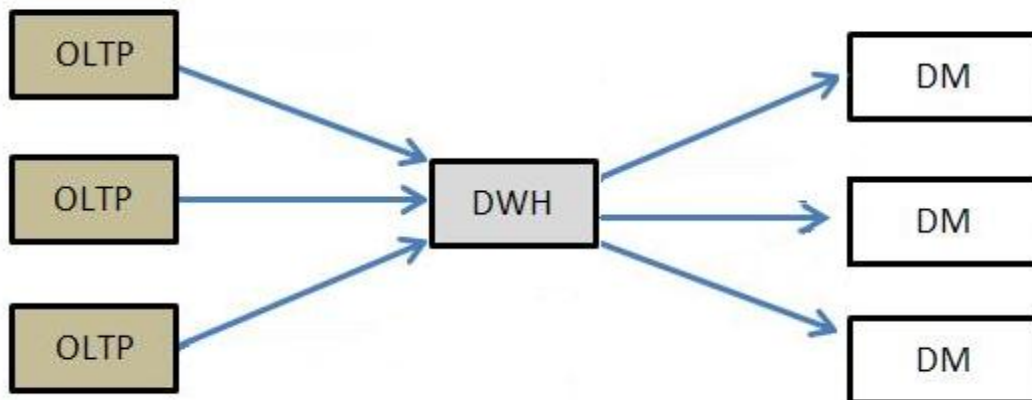


*Figura 6: Enfoque Kimball.*

### **Metodología de Bill Inmon**

Inmon plantea transferir la información de las diferentes base de datos operacional de las organizaciones a un DW centralizado, donde los datos puedan ser utilizados para su posterior análisis. La información ha de estar a los máximos niveles de detalle. Los Data Mart son tratados como subconjuntos de un DW corporativo, y son construidos para cubrir las necesidades individuales de análisis de cada departamento, siempre partiendo del DW central.

En el enfoque Inmon, referenciado normalmente como Top-down, los datos son extraídos de los sistemas operacionales por los procesos ETL donde son validados y consolidados para su posterior almacenamiento en el DW, donde además existen los llamados metadatos que documentan de una forma clara y precisa el contenido del DW. Una vez realizado este proceso, los procesos de actualización de los Data Mart departamentales obtienen la información de él, y con las consiguientes transformaciones, organizan los datos en las estructuras particulares requeridas por cada uno de ellos. Al tener este enfoque global, es más difícil de aplicar en un proyecto sencillo. (Figura 7). [43]



*Figura 7: Enfoque Inmon.*

### **Metodología para la implementación de la solución.**

Para la realización del DM del Subsistema de Contabilidad de Cedrux se estudiaron las siguientes metodologías: Hefesto, DATEC, UCID. La selección de una de estas depende de las características del negocio y de la organización para la cual se va a construir.

### **Metodología de Hefesto**

Hefesto es una metodología propia, cuya propuesta está fundamentada en una muy amplia investigación, comparación de metodologías existentes, y experiencias propias en procesos de confección de almacenes de datos. Es una de las más difundidas y utilizadas por su fácil implementación y aporte práctico, aunque no propone de forma explícita los artefactos y entregables a generar en cada fase.

Esta metodología cuenta con las siguientes características:

- Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- Se basa en los requerimientos del usuario, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio.
- Reduce la resistencia al cambio, ya que involucra al usuario final en cada etapa para que tome decisiones respecto al comportamiento y funciones del DW.

- Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
- Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.
- Es independiente de las herramientas que se utilicen para su implementación.
- Es independiente de las estructuras físicas que contengan el DW y de su respectiva distribución.
- Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.
- Se aplica tanto para el desarrollo de Data Mart como para DW. [42]

### **Metodología de DATEC**

Basada fundamentalmente en el enfoque Kimball, consta de 5 fases y un total de 43 artefactos. En esta se definen además los hitos de desarrollo así como los roles y sus responsabilidades en el mismo definiendo también las herramientas a emplear en cada una de sus fases. Su aporte radica en la integración de algunas prácticas de RUP con este enfoque. Si bien esto ayuda a obtener una documentación amplia para el proyecto, implica el empleo de mayores recursos como tiempo, esfuerzo y personas en la generación de estos artefactos para cada una de sus fases, estas son las siguientes:

- Fase de requerimientos.
- Fase de Arquitectura y Diseño.
- Fase de Implementación y Prueba.
- Fase de despliegue.
- Fase de soporte.

### **Metodología UCID**

Esta metodología está basada fundamentalmente en los principios planteados en la metodología Hefesto, presentando las mismas 4 fases de desarrollo, definiendo nuevos artefactos y detallando de forma más específica los planteados por la propia metodología. Cumple con las mismas características fundamentales de la metodología Hefesto.

Las fases y artefactos generados para cada una de ellas son los siguientes.

#### **1. Análisis de los requerimientos:**



- Planificar Entrevista
- Identificar Preguntas
- Identificar Perspectivas e Indicadores
- Construir Modelo Conceptual
- Aprobación del Modelo Conceptual por el cliente

## **2. Análisis de los OLTP**

- Definir Estado General de los sistemas fuentes
- Determinación de indicadores
- Establecer correspondencias
- Nivel de granularidad
- Construir Modelo Conceptual ampliado
- Definir reglas del negocio

## **3. Modelo físico del DW**

- Definir Tipo de Modelo Lógico del DW
- Definición de estándares para objetos físicos
- Identificar dimensiones
- Identificar hechos
- Realizar Uniones entre
- dimensiones y hechos
- Diseñar tablas y columnas físicas

## **4. Proceso ETL**

- Mapeo de datos fuente-a-destino.
- Establecer condiciones adicionales y restricciones.
- Cargas incrementales de datos.

- Diseño, construcción de la automatización del sistema de ETL. [27]

### **Conclusiones de las metodologías.**

La metodología de arquitectura, Kimball es más apropiado para sistemas pequeños donde se haga necesario asegurar la usabilidad de los usuarios con un sistema fácil de entender y garantizar un rápido desarrollo en la solución. La metodología DATEC y Hefesto posibilitan organizar de forma eficaz el proceso de creación de los DW; el empleo de la primera implica la creación de gran cantidad de artefactos, para la documentación, haciendo que el proceso de implementación del Data Mart se torne engorroso y tedioso, y la segunda, si bien por su aporte práctico es fácil de entender e implementar, no define de manera concisa los artefactos a generar en cada fase. Por las diversas razones antes planteadas se seleccionó la metodología UCID para el rediseño del DM del subsistema de Contabilidad.

### **1.3 Replicación de datos.**

La replicación de datos se entiende como el proceso de copiar y mantener objetos de una base de datos en múltiples base de datos, de tal forma que los cambios realizados localmente, son enviados y aplicados a las base de datos remotas. Proporciona una manera rápida y confiable de copiar y distribuir información de una base de datos a otra u otras, permitiendo mejorar el funcionamiento de las consultas. [5]. La replicación asegura que los datos estén siempre disponibles en el momento y en el lugar necesario. [6]

#### **Características de la replicación de Base de Datos:**

**Efectividad:** Depende de la forma en la que los datos sean distribuidos y almacenados. A mayor efectividad, mayor será la disponibilidad de datos para ejecutar procesos paralelos.

**Alta Disponibilidad:** Es la razón de tiempo prudente en la que un servicio puede ser accedido. En el mejor de los casos puede ser de un 100%, a pesar de los fallos que se puedan presentar en el servidor, debe existir un servidor adicional que posea alguna técnica de replicación que lo pueda suplantar en caso de ser necesario. Esta característica de replicación se refiere a la accesibilidad de los usuarios a servicios del sistema. Un sistema está disponible si es completamente operacional durante un período de tiempo. La misma se concentra en la duración del tiempo en el cual se espera que el sistema se mantenga en operación continua o en el caso de recuperación confiable.

**Tolerancia a fallos:** Garantiza un comportamiento correcto, donde efectivamente pueden existir un número finito de fallos y tipos de fallos.

**Coordinación:** Cada una de las partes que conforman la base de datos distribuida, acuerda un consenso para realizar las invocaciones de los servicios a los objetos, que al final de la transacción debe realizarse tal y como fue solicitada, para lo cual debe utilizar algún tipo de ordenamiento. [7]

### 1.3.1 Herramienta para replicar datos.

#### Reko

Reko pretende cubrir las principales necesidades relacionadas con la distribución de datos entre los gestores más populares como la protección, recuperación, sincronización de datos, transferencia de datos entre diversas localizaciones o la centralización de la información en una única localización. Facilita la representación de escenarios de replicación complejos al permitir la definición de localizaciones o nodos y la selección de los datos de replicación. Brinda una herramienta Web para la administración y monitoreo de los datos de replicación. [20]

#### Principales características:

*Tipos de replicación:* Soporta la replicación a través de disparadores<sup>3</sup>.

*Gestores soportados:* Se integra actualmente con los gestores de base de datos Oracle y PostgreSQL.

*Selección de datos:* Provee la facilidad de seleccionar el subconjunto de tablas y datos a ser replicados de la base de datos mediante la definición de filtros aplicables a las tablas y la selección de usuarios propios del gestor.

*Transmisión de datos:* La transferencia de datos de replicación puede realizarse a soporte para replicación mediante protocolos de transferencia de archivos (TCP/IP, FTP, HTTP) o por ficheros de forma manual. Los archivos de gran tamaño son enviados por Protocolo de Transferencia de Archivos FTP permitiendo resumir la transmisión en caso de interrupciones en la red.

*Configuración entre nodos:* El mecanismo de registro entre nodos de replicación se basa únicamente en un identificador (id) del nodo lo que permite la abstracción de los datos físicos de cada nodo como son el

---

<sup>3</sup> Un disparador en una Base de datos, es un procedimiento que se ejecuta cuando se cumple una condición establecida al realizar una operación de inserción, actualización o borrado. [12]

IP y el protocolo de comunicación. Este mecanismo permite que los nodos puedan moverse por distintas subredes y mantener sus datos sincronizados.

*Seguridad:* Los nodos de replicación manejan credenciales entre ellos para verificar la autenticidad de los datos transferidos. El envío de datos se realiza utilizando protocolos de comunicación seguros como son (HTTPS y SSL).

*Monitoreo:* Permite conocer el estado de los datos transmitidos, puede realizarse en tiempo real a través de la Web así como dar un seguimiento al funcionamiento interno del mecanismo.

*Interfaz visual Web:* La administración y configuración de la réplica se realiza a través de una interfaz Web, por lo que es administrable vía remota usando solamente un navegador.

*Independiente de la plataforma:* El software de réplica puede ser instalado en cualquier sistema operativo y no necesita de un ambiente gráfico en el mismo para su funcionamiento.

*Tolerancia a fallos:* Detecta errores de conexión. Mantiene los datos de réplica en un estado estable en caso de desconexión. Al restablecerse la conexión, automáticamente sincroniza los datos entre las base de datos.

Reko permite que dos BD puedan replicar entre sí aunque se encuentren en subredes distintas y utilicen protocolos de transmisión diferentes, la única restricción es que todos deben estar conectados a un servidor central JMS o en caso de un clúster de servidores JMS, estar conectados a uno de ellos. Está certificado por CALISOFT, Calidad de Software, y aprobado por los Laboratorios de Pruebas por lo cual se considera confiable y de calidad. [12]

#### **1.4 Herramienta de prueba.**

Jmeter es una herramienta de software libre que ofrece la posibilidad de realizar pruebas de carga sobre diferentes aspectos de una aplicación desarrollada en Java. Esta herramienta posee dos tipos de generación de carga: indirecta, es decir, a través de una aplicación y directa que basa fundamentalmente su utilización en consultas grabadas en la traza o log del servidor de base de datos. Es fácilmente configurable y permite conocer los tiempos de respuesta experimentados por una aplicación cuando se tiene un número  $N$  de usuarios y el número real de transacciones procesadas por unidad de tiempo. Corre sobre la máquina virtual de java (JVM), genera carga para FTP, HTTP, HTTPS y SQL. [45]

### **Ventajas de la herramienta.**

- De las herramientas gratis, es la más completa y útil.
- Es una herramienta que sirve para realizar pruebas funcionales, pero también sirve para realizar pruebas de regresión en aplicaciones web.
- Tiene una estructura en árbol que le da potencia, permitiendo que sea la imaginación de quien la use la que ponga los límites a la hora de diseñar el plan de prueba.
- Brinda mayor cantidad de variantes para recoger los resultados obtenidos, que el resto de las herramientas gratis, lo que permiten hacer un análisis exhaustivo de las pruebas realizadas.[46]

### **CONCLUSIONES PARCIALES**

A lo largo del presente capítulo se ha hecho mención de los principales conceptos que se abordarán durante el desarrollo del trabajo de diploma. Se realizó una descripción de las tecnologías y herramientas que serán utilizadas para la construcción de la solución. Se hizo un estudio teórico de las principales metodologías existentes para la construcción de Data Mart, seleccionándose para el presente trabajo, la metodología UCID.

A partir de los conocimientos adquiridos y de la selección de herramientas, tecnologías y metodología seleccionada se da paso al desarrollo de la solución.

## CAPÍTULO 2: ANÁLISIS Y DISEÑO DE LA SOLUCIÓN

### **INTRODUCCIÓN**

En el presente capítulo se realiza el análisis, diseño e implementación del Data Mart del subsistema de Contabilidad, siguiendo las fases que define la metodología UCID, con el propósito de tener almacenado todos los datos necesarios para realizar los procesos contables agregación y consolidación. Además de definirse los requisitos funcionales y no funcionales para la replicación de los datos, se propone la configuración de replicación de BD necesaria para los procesos de agregación y consolidación de las entidades.

#### **2.1 Análisis, diseño e implementación del Data Mart.**

Las transformaciones hechas al Data Mart están en correspondencia con las etapas del proceso de desarrollo de software planteado en la metodología seleccionada en el capítulo anterior.

##### **Fase 1: Análisis de los requerimientos para el Data Mart.**

Esta etapa tiene como premisa que se formulen preguntas complejas sobre el negocio, que contenga variables de análisis relevantes para que se estudie la información a partir de diversos puntos de vista. Para esto se realizaron 3 entrevistas, el objetivo de las mismas se enfoca en la identificación de las dimensiones y medidas que se tomarán en cuenta para la transformación del DM.

##### **Cronograma de entrevistas.**

La tabla que se presenta a continuación muestra el cronograma de entrevistas, a partir del cual se obtendrá la información necesaria para efectuar las transformaciones pertinentes sobre el Data Mart del subsistema Contabilidad para garantizar la realización de los procesos contables: agregación y consolidación.

No.	Fecha	Hora	Lugar	Objetivo	Responsable	Aseguramientos	Participan
1	14/9/2011	9:30 am	Doc. Viejo( Lab. 6)	Conocer el negocio del subsistema Contabilidad.	Yamira Antonia Roque Vázquez.	Personal	<i>Est.</i> Yamira A. Roque Vázquez. <b>Esp. Funcional</b> Ing. Annilié Manresa
2	20/10/2011	9:30 am	Doc. Viejo( Lab. 6)	Identificar las dimensiones para realizar los procesos de agregación y consolidación.	Yamira Antonia Roque Vázquez	Personal	<i>Est.</i> Yamira A. Roque Vázquez. <b>Esp. F.</b> Ing. Annilié Manresa
3	29/10/2011	9:30 am	Doc. Viejo( Lab. 6)	Definir el grado de granularidad de cada una de las dimensiones identificadas.	Yamira Antonia Roque Vázquez	Personal	<i>Est.</i> Yamira A. Roque Vázquez. <b>Esp. F.</b> Ing. Annilié Manresa

Tabla 2: Cronograma de entrevistas.

### Descripción de las entrevistas.

### **Entrevista 1:**

La entrevista 1 se realizó con el objetivo de profundizar en el conocimiento del negocio del subsistema de Contabilidad. En la misma estuvo presente el especialista del subsistema junto al equipo de desarrollo. Se logró que el equipo de desarrollo entendiera el negocio del subsistema de Contabilidad garantizando el cumplimiento de los objetivos trazados y la base para el posterior desarrollo.

### **Entrevista 2:**

La entrevista 2 se realizó con el objetivo de identificar las dimensiones necesarias para realizar los procesos de agregación y consolidación en el subsistema de Contabilidad. En la misma estuvo presente el especialista del subsistema junto al equipo de desarrollo, llegando entre ellos a un acuerdo en cuanto a las principales funcionalidades a desarrollar. Con el cumplimiento de esta entrevista el equipo de desarrollo puede trabajar en la confección y modificación del modelo conceptual de los hechos: “cub\_hecho\_agregacion\_cuenta” y “cub\_hecho\_contabilidad\_cuenta” respectivamente.

### **Entrevista 3:**

La entrevista 3 se realizó con el objetivo de definir el grado de granularidad de cada una de las dimensiones identificadas, necesarias para realizar los procesos de agregación y consolidación en el subsistema de Contabilidad. En ésta participó el especialista funcional del subsistema en conjunto con el equipo de desarrollo. Se obtuvo en detalles el grado de granularidad de cada una de las perspectivas.

### **Identificar dimensiones.**

Para garantizar un buen desempeño en la fase de implementación del Data Mart es necesario un profundo estudio de las dimensiones a incluir, pues éstas constituirán las entidades de negocio de las cuales se obtendrán los filtros para el análisis de la información.

### **Dimensiones:**

#### **➤ Estructura Económica**

Esta perspectiva representa las estructuras económicas que en algún momento podrán ser objeto de análisis. Las mismas constituyen las divisiones económicas establecidas a nivel de país para agrupar a las distintas entidades contables.

#### **➤ Configuración**

Esta perspectiva representa la agregación que a realizar entre un conjunto de entidades.



## Estructuras en forma de árbol

La estructura en forma de árbol constituye una abstracción del negocio, su representación posee una relación con otro u otros objetos delimitados jerárquicamente entre ellos y sin niveles predefinidos. El Data Mart del subsistema de Contabilidad está constituido por 4 de estas tablas y surgió la necesidad de incluir una nueva:

- dim\_estructurae\_closure.

## Construcción del modelo conceptual

El modelo conceptual simboliza los objetos del negocio en términos de hechos, dimensiones, atributos de las dimensiones, cruces entre dimensiones, niveles de las dimensiones y medidas.

El modelo conceptual del hecho “cub\_hecho\_contabilidad\_cuenta” se cambió producto de la inserción de una nueva dimensión (estructura económica) y de las medidas: crédito acumulado y débito acumulado. Ver Figura 8.

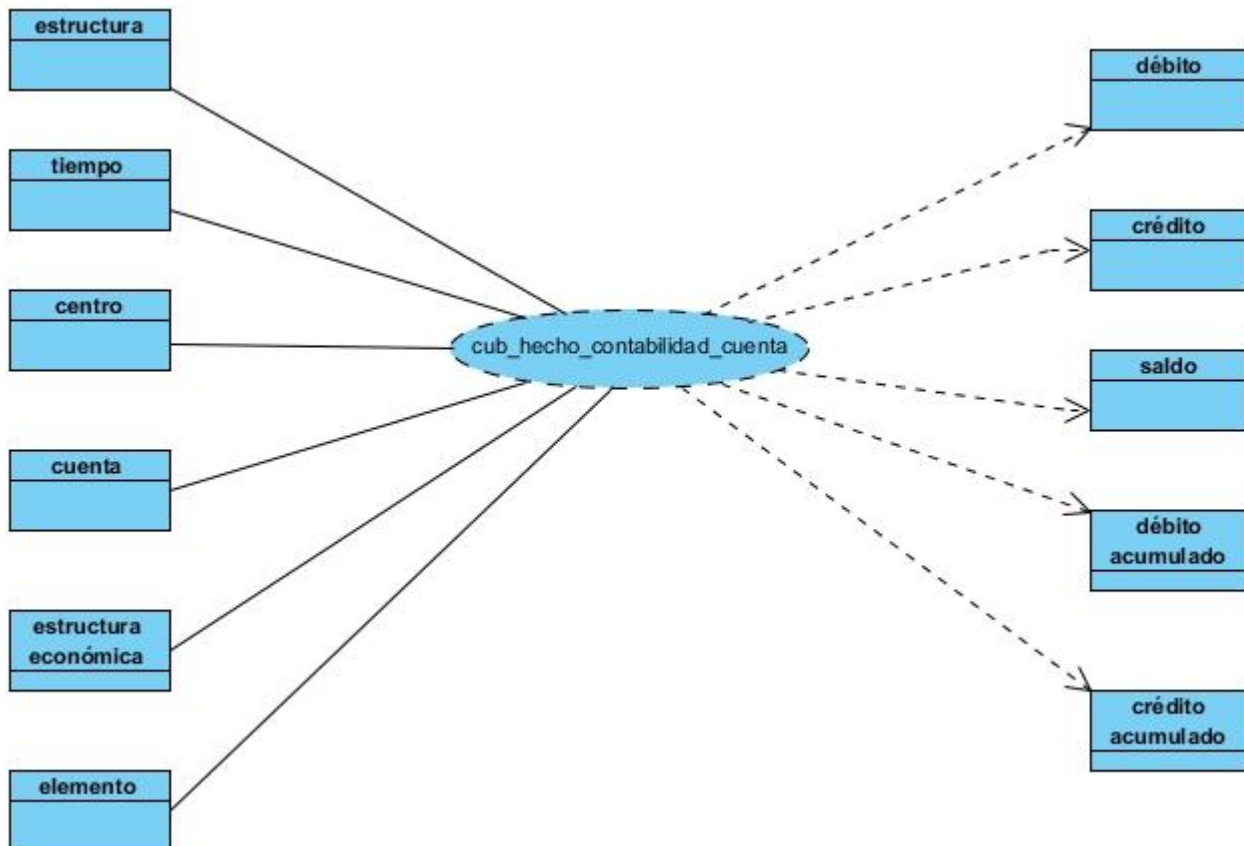


Figura 8: Modelo conceptual del hecho "cub\_hecho\_contabilidad\_cuenta".

La siguiente figura muestra el modelo conceptual definido para el hecho "cub\_hecho\_agregacion\_cuenta", el cual está formado por 5 dimensiones.

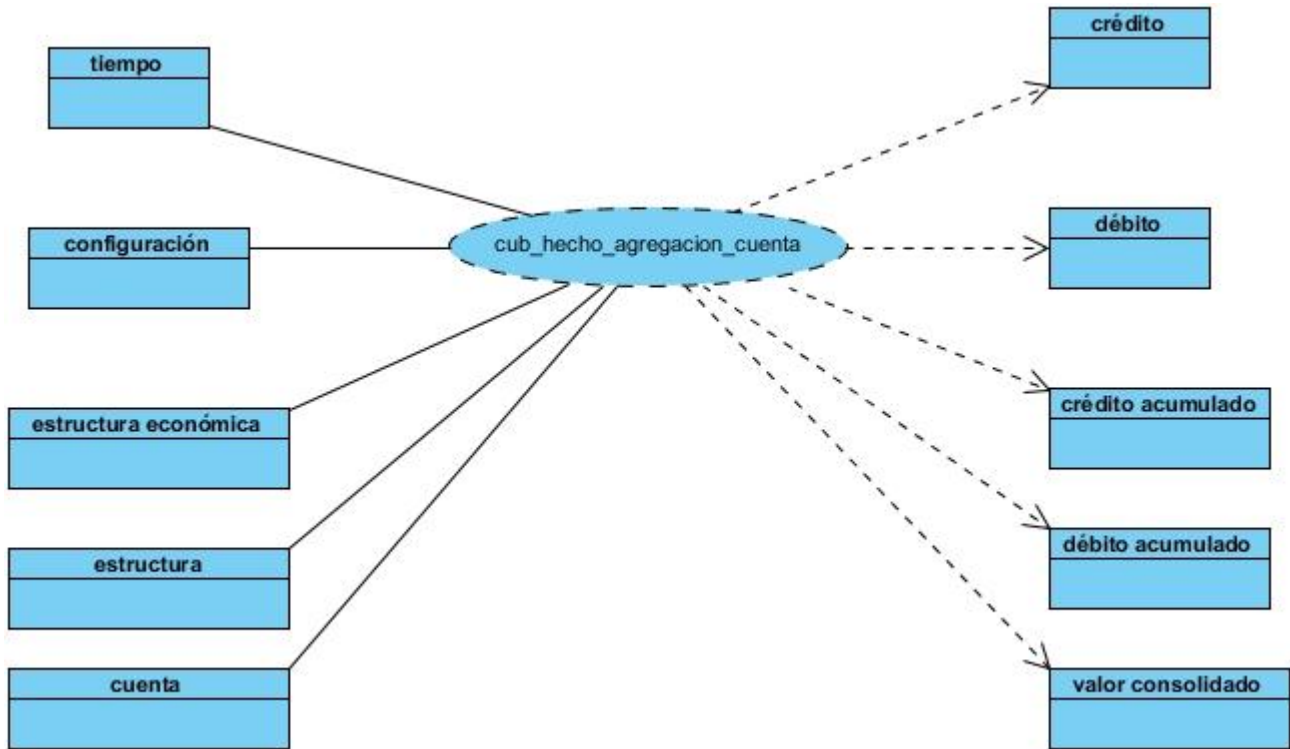


Figura 9: Modelo conceptual del hecho "cub\_hecho\_agregacion\_cuenta".

## Fase 2: Análisis de los OLTP.

### Nivel de granularidad.

La granularidad es considerada uno de los aspectos de más importancia en el diseño de un Data Mart. Se refiere al nivel de detalles o resumen de las unidades de datos, siendo este inversamente proporcional al nivel de granularidad existente.

La información de la dimensión *tiempo* es más conveniente almacenarla por mes y no diaria, para obtener los saldos agrupados en un período de tiempo que posibilite conocer la variación monetaria de las mismas. A continuación se muestra el nivel de granularidad definido para las dimensiones de los hechos "cub\_hecho\_contabilidad\_cuenta" y "cub\_hecho\_agregacion\_cuenta".

<b>Estructura</b>	<b>Cuentas</b>	<b>Estructura económica</b>	<b>Centros</b>	<b>Elementos</b>	<b>Tiempo</b>
Estructura	Grupo	Estructura económica	Centro	Elemento	Quinquenio
	Contenido				Año
	Cuenta				Semestre
					Trimestre
					mes

Tabla 3: Nivel de granularidad de las perspectivas para el hecho: "cub\_hecho\_contabilidad\_cuenta".

<b>Estructura económica</b>	<b>Consolidación</b>	<b>Tiempo</b>
Estructura económica	Consolidación	Quinquenio
		Año
		Semestre
		Trimestre
		mes

Tabla 4: Nivel de granularidad de las perspectivas para el hecho: "cub\_hecho\_agregacion\_cuenta".





### Fase3: Modelo físico del DM.

Para la realización de los procesos de agregación y consolidación en el subsistema de Contabilidad del Sistema Integral de Gestión Cedrux se hizo necesario utilizar campos que no se encontraban presente en ninguna de las tablas del Data Mart del subsistema Contabilidad, producto de esta problemática se transforma el diseño lógico de la base de datos existente.

La etapa del diseño lógico es independiente de los detalles de implementación y dependiente del SGBD PostgreSQL. La salida de esta etapa es el esquema lógico global y la documentación que lo describe, este artefacto recibe el nombre de: diccionario de datos.

### Diccionario de las nuevas tablas a introducir en el Data Mart del subsistema de Contabilidad.

Se confeccionó el diccionario de datos de las nuevas tablas a introducir, para garantizar un mayor conocimiento y control de las mismas durante la transformación del Data Mart y se cambió el mismo para el hecho: cub\_hecho\_contabilidad\_cuenta.

Nombre	Documentación
 dim_configuracion_agregacion	Guarda los identificadores de la estructura económica de las entidades que serán agregadas.
 dim_configuracion_agregacion_hijos	Agrupa las entidades (por el idestructurae) según la agregación a la que pertenezcan, (con el idconsolidacion).
 dim_estructurae	Guarda los identificadores de las estructuras económicas de las entidades y se especifica si esa entidad es la padre cuando el identificador estructura económica padre tenga valor 0.
 dim_estructurae_closure	Agrupa los identificadores de las estructuras económicas de las entidades, en forma de árbol, según los identificadores de la estructura económica padre.


 cub_hecho_agregacion_cuenta	Guarda los valores obtenido del proceso de agregación y presenta un campo (valor consolidado) a partir del cual se realiza la consolidación.
---	--

Tabla 5: Descripción del diccionario de datos de las nuevas tablas introducidas al Data Mart.

➤ **dim configuracion agregacion**

Nombre	Data type	Limitaciones	Anulable	Documentación
idconsolidacion	numeric(0,19)	PK	No	Es la llave primaria, será un número consecutivo generado a partir de una secuencia que incrementará de uno en uno. Es el identificador de la agregación que nace.
nombre	charactervarying(50)		No	Se plasman el nombre de la agregación que se crea.
idestructuraepadre	numeric(0,19)		No	Indica que entidad es el padre en esta agregación.

Tabla 6: Descripción de los campos de la tabla dim\_configuracion\_agregacion.

➤ **dim configuracion agregacion hijos**

Nombre	Date type	Limitaciones	Anulable	Documentación
idconsolidacion	numeric(0,19) )	FK	No	Es la llave foránea de la tabla dim_configuracion_agregacion.
idestructuraehijo	numeric(0,19) )		No	Se plasman el identificador de la estructura económica de las entidades que participa en la agregación, representada por el identificador de consolidación.

Tabla 7: Descripción de los campos de la tabla dim\_configuracion\_agregacion\_hijos.

➤ **dim\_estructurae**

Nombre	Date type	Limitaciones	Anulable	Documentación
idestructurae	numeric(0,19)	PK	No	Es la llave primaria, el mismo será el identificador económico de una entidad.
codigoestructura	charactervarying(50)			Código que se asigna una entidad para ser identificada.
descripcion	charactervarying(255)			Indica el nombre de la entidad a la que

				pertenece el identificador económico.
idestructuraepadre	numeric(0,19)			Indica que entidad es padre de la que me encuentro posicionado, si el identificador posee valor 0, significa que no tiene padre, siendo la entidad predominante.

Tabla 8: Descripción de los campos de la tabla dim\_estructurae.

➤ **dim\_estructurae\_closure**

Nombre	Date type	Limitaciones	Anulable	Documentación
idestructurae	numeric(0,19)			Identificador económico de una entidad.
idestructuraepadre	numeric(0,19)			Indica que entidad es padre de la que me encuentro posicionado, si el identificador posee valor 0, significa que no tiene padre, siendo la entidad predominante.
distancia	numeric(0,19)			Constituye la distancia entre un identificador económico hijo hasta el

				identificador económico padre.
--	--	--	--	--------------------------------

Tabla 9: Descripción de los campos de la tabla dim\_estructurae\_closure.

➤ **cub hecho agregacion cuenta**

Nombre	Date type	Limitaciones	Anulable	Documentación
idhecho	numeric(0,19)	PK	No	Es la llave primaria, el mismo será un número consecutivo que será generado a partir de una secuencia que incrementará de uno en uno. Es el identificador del proceso de agregación que se realizó.
idestructura	numeric(0,19)			El id estructura de la entidad padre que participa en dicha agregación es almacenado en este campo.
idcuenta	numeric(0,19)			Guarda el identificador de la cuenta padre de



				todas las cuentas que son sumadas, sus débitos y créditos, en dicha agregación.
iddimtiempo	numeric(0,19)		No	Especifica el momento en que fue alterada una cuenta.
creditomb	numeric(0,19)			Crédito en moneda base. Constituye la sumatoria del crédito de la empresa, incluyendo el del período de apertura y de cierre.
debitomb	numeric(0,19)			Débito en moneda base. Constituye la sumatoria del débito de la empresa, incluyendo el del período de apertura y de cierre.
idestructurae	numeric(0,19)			El id estructura económica de la

				entidad padre que participa en dicha agregación es almacenado en este campo.
idconfiguracion	numeric(0,19)			Es el identificador de consolidación de la agregación.
valorconsolidacion	numeric(0,19)		No	Guarda el valor resultante del Comprobante de Operación, a partir del cual se realiza el proceso de agregación.
creditoacumulado	numeric(0,19)			Constituye la sumatoria del crédito de la empresa, exceptuando el del período de apertura y de cierre.
debiitoacumulado	numeric(0,19)			Constituye la sumatoria del débito de la empresa, exceptuando el del período de

				apertura y de cierre.
--	--	--	--	-----------------------

Tabla 10: Descripción de los campos de la tabla cub\_hecho\_agregacion\_cuenta.

➤ **cub hecho contabilidad cuenta**

Nombre	Date type	Limitaciones	Anulable	Documentación
idhecho	numeric(0,1 9)	PK	No	Es la llave primaria.
idestructura	numeric(0,1 9)			El id estructura de la entidad.
idcuenta	numeric(0,1 9)			Guarda el identificador de la cuenta.
idcentro	numeric(0,1 9)			Este campo representa el identificador del centro de costo. El valor que toma este campo es el correspondiente al propio centro en el sistema operacional.
idelemento	numeric(0,1 9)			Este campo representa el identificador del elemento de gasto.

				El valor que toma este campo es el correspondiente al propio elemento de gasto en el sistema operacional.
iddimtiempo	numeric(0,19)		No	Especifica el momento en que fue alterada una cuenta.
creditomb	numeric(0,19)			Crédito en moneda base. Constituye la sumatoria del crédito de la empresa, incluyendo el del período de apertura y de cierre.
debitomb	numeric(0,19)			Débito en moneda base. Constituye la sumatoria del débito de la empresa, incluyendo el del período de apertura y de cierre.
idestructurae	numeric(0,19)			El id estructura

	9)			económica de la entidad.
creditoacumulado	numeric(0,19)			Constituye la sumatoria del crédito de la empresa, exceptuando el del período de apertura y de cierre.
debiitoacumulado	numeric(0,19)			Constituye la sumatoria del débito de la empresa, exceptuando el del período de apertura y de cierre.

*Tabla 11: Descripción de los campos de la tabla cub\_hecho\_contabilidad\_cuenta.*

### **Modelo Lógico del Data Mart del subsistema de Contabilidad.**

Para el rediseño del Data Mart se optó por el modelo “constelación de hecho”, producto que las tablas de hechos, que se tienen, comparten en algunos casos las mismas dimensiones. Ver Figura 10.

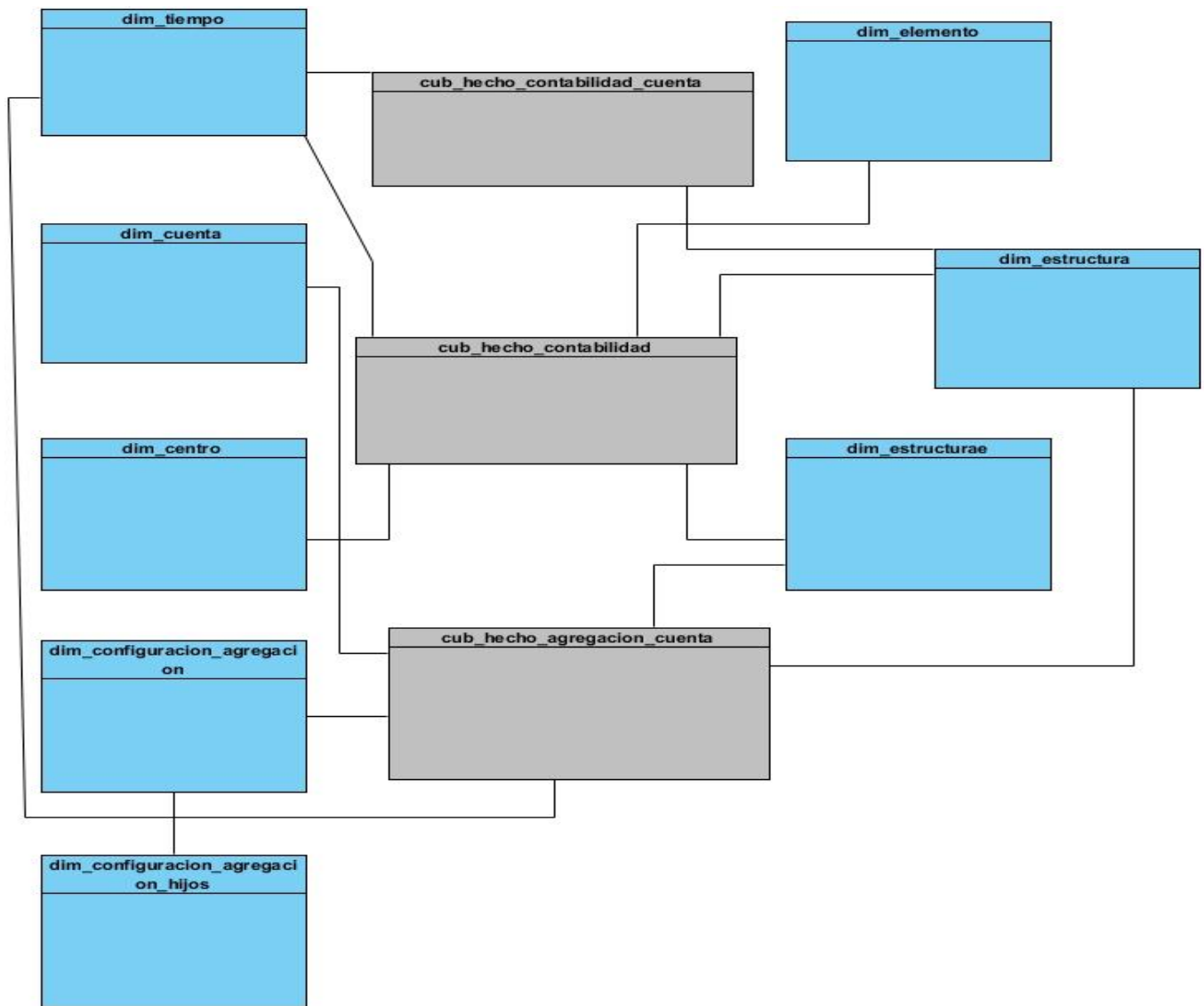


Figura 10: Modelo lógico del Data Mart del subsistema de Contabilidad.

### Diseño físico del Data Mart del subsistema de Contabilidad.

El modelo físico está compuesto por 3 tablas de hechos, 8 tablas de dimensiones y 5 tablas closure. De este total de tablas, fueron introducidas una tabla de hecho, 3 tablas de dimensiones y una tabla closure. Se emplearon tablas closure para gestionar el funcionamiento de las estructuras en forma de árbol. Ver Figura 11.

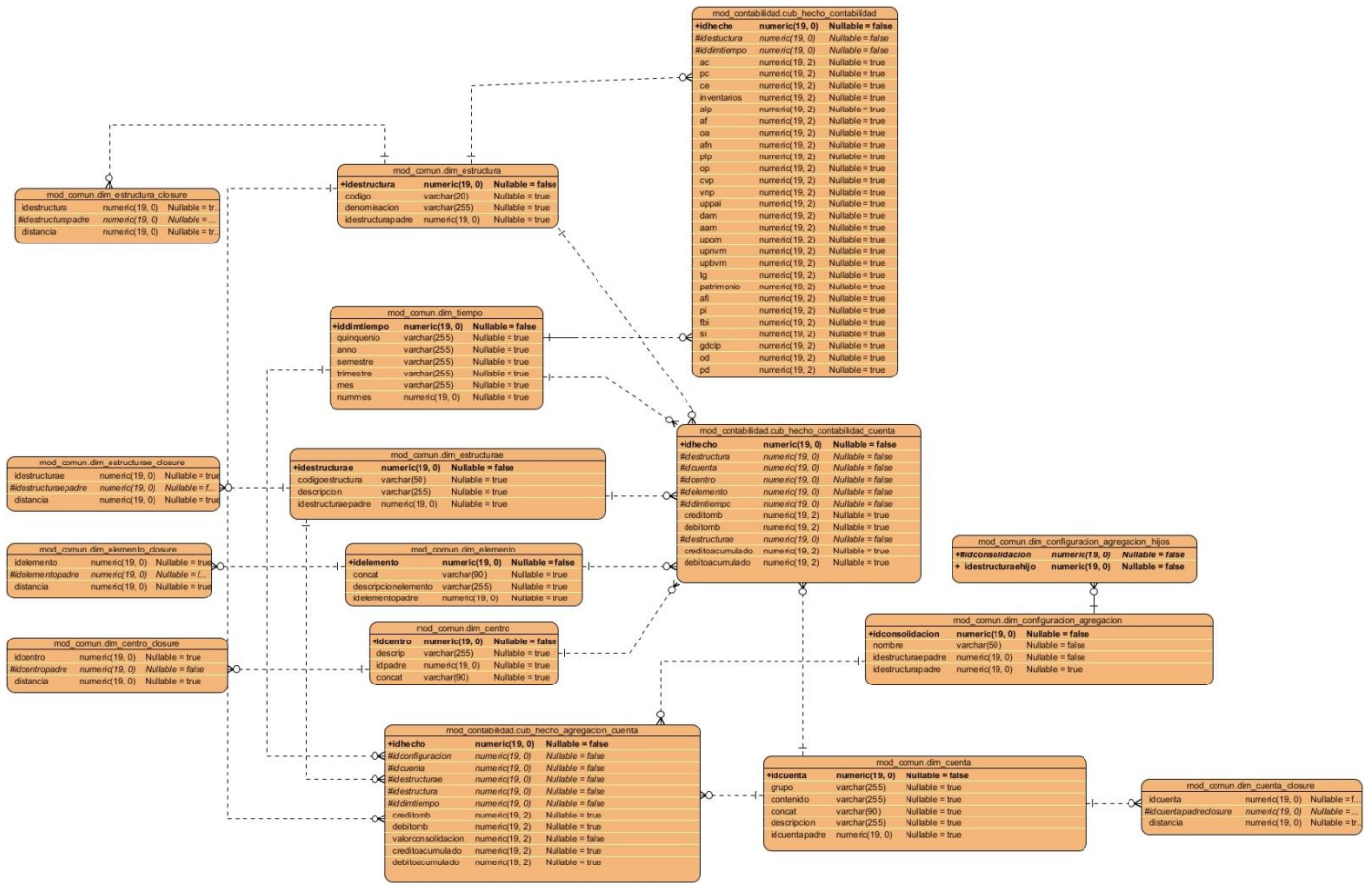


Figura 11: Modelo físico del Data Mart del subsistema de Contabilidad.

## Diseñar las tablas auxiliares físicas.

Producto a la alta complejidad que generan las consultas multidimensionales en tablas con relaciones recursivas, se crearon tablas auxiliares que permiten la prolongación de estas estructuras, soportando la creación de estas dimensiones en forma de árbol. Ver Figura 12.

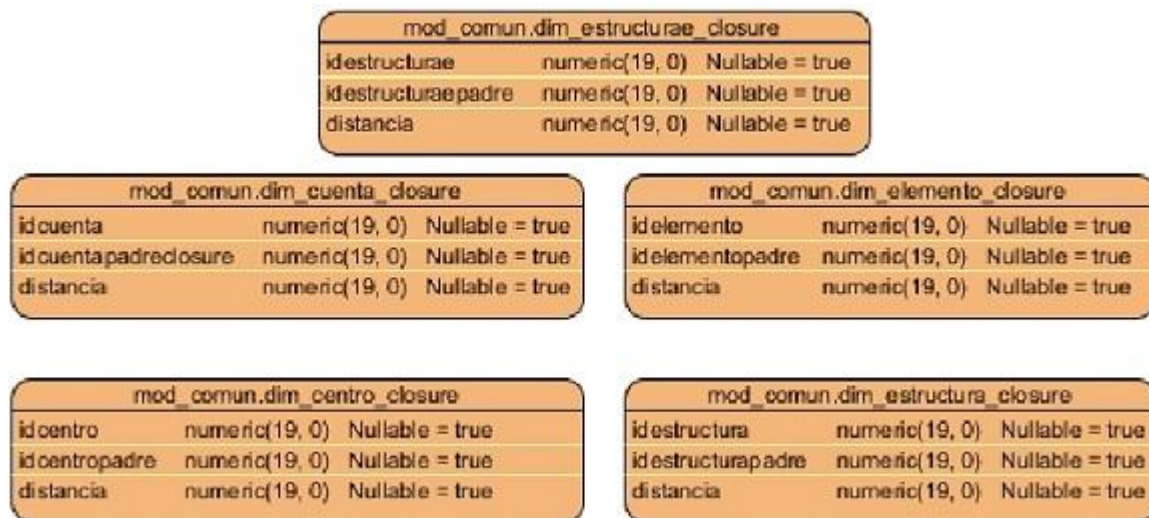


Figura 12: Modelo físico de las tablas closure del Data Mart del subsistema de Contabilidad.

#### Fase 4: Proceso ETL.

Un proceso de Extracción, Transformación y Carga organiza el flujo de los datos entre diferentes sistemas y aporta los métodos y herramientas necesarias para mover datos desde múltiples fuentes a un almacén de datos, reformatearlos, limpiarlos y cargarlos en otra base de datos. [24]

#### Mapeo de datos fuente a destino.

Con el propósito de un mayor control de los datos en el proceso ETL se realizó un mapeo de estos, documentando las relaciones entre los campos de la base de datos origen y su correspondiente en la base de datos destino. La siguiente tabla muestra una porción del mismo. El mapeo de las restantes tablas se observa en el Anexo 1.

Mapa lógico de datos					
Nombre BD fuente: Cedrux			Nombre BD destino: Data Mart		
Nombre de la tabla	Nombre del campo	Tipo de dato	Nombre de la tabla	Nombre del campo	Tipo de dato



dat_estructurae	idestructurae	numeric(19,0)	dim_estructurae	idestructurae	numeric(19,0)
dat_estructurae	codigoestructura	varchar(50)	dim_estructurae	codigoestructura	character varying(50)
dat_estructurae	descripcion	varchar(255)	dim_estructurae	descripcion	character varying(255)
dat_estructurae	idestructuraepadre	numeric(19,0)	dim_estructurae	idestructuraepadre	numeric(19,0)

Tabla 12: Mapa lógico de datos.

### Condiciones adicionales y restricciones

El conjunto de condiciones y restricciones que se listan a continuación garantizarán una mayor confidencialidad y disponibilidad de los datos en el Data Mart.

#### Restricciones:

1. Para el llenado de la tabla “cub\_hecho\_contabilidad\_cuenta” se hace necesario que las medidas sean calculadas por estructura, centro, cuenta, elemento, tiempo y estructura económica.
2. Para el llenado de la tabla “cub\_hecho\_agregacion\_cuenta” se hace necesario que las medidas sean calculadas por estructura, cuenta, estructura económica, tiempo y consolidación.
3. La división por 0 no está definida, por lo que en el caso de un denominador con valor 0, contablemente se divide por 1.

### Diseño e implementación del proceso de ETL.

El proceso de Extracción, Transformación y Carga se llenaron las nuevas tablas introducidas al Data Mart del subsistema de Contabilidad; utilizando la herramienta Pentaho Data Integration, con el propósito de obtener ganancia en términos de tiempo y total fiabilidad de los datos.

Durante el proceso de extracción se obtuvieron los datos desde los sistemas orígenes (tablas de la base de datos de CedruX, a través de consultas SQL; los mismos se convirtieron a un formato preparado para iniciar la próxima fase. Una parte intrínseca de este proceso es analizar los datos extraídos, de lo que resulta un chequeo que verifica si los datos cumplen las pautas o estructura que se esperaban. De no ser así los datos son rechazados. Posteriormente se realizó el proceso de transformación, donde se aplicaron una serie de reglas de negocio o funciones sobre los datos extraídos. Por último se efectuó el proceso de carga, momento en el cual los datos de la fase anterior (transformación) fueron cargados en el sistema destino (tablas del Data Mart del subsistema Contabilidad). Este proceso abarca una amplia variedad de acciones diferentes, tales como: mantener un historial de los registros de manera que se pueda hacer una auditoría de los mismos y disponer de un rastro de toda la historia de un valor a lo largo del tiempo almacenados en logs.

El orden en que se ejecuten el proceso, para las entidades identificadas, influye en el buen resultado del mismo. Para garantizar esto se definieron los procesos en paralelo, los independientes de otros trabajos para su ejecución, y los dependientes, los dependientes de la culminación de otros trabajos para su ejecución.

### Procesos paralelos:

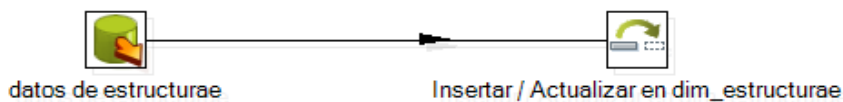


Figura 13: Proceso ETL para la tabla `dim_estructurae`.



Figura 14: Proceso ETL para la tabla `dim_configuracion_agregacion`.



Figura 15: Proceso ETL para la tabla `dim_configuracion_agregacion_hijos`.

## Proceso dependiente:

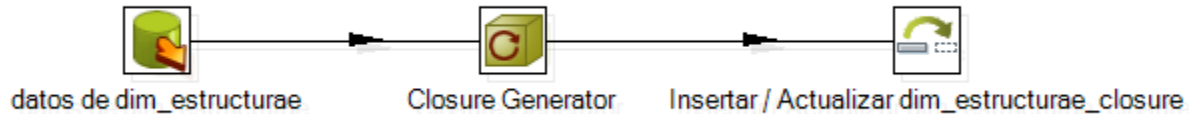


Figura 16: Proceso ETL para la tabla dim\_estructurae\_closure.



Figura 17: Proceso ETL para la tabla cub\_hecho\_contabilida\_cuenta.

Las consultas SQL introducidas para extraer los datos necesarios de la fuente origen y cargarlos en las tablas insertadas al Data Mart se observan en las siguientes imágenes:

```
SELECT mco.idconsolidacion,  
       mco.idestructurae as idestructuraepadre, mco.nombre,  
(SELECT mc.idestructuracomun FROM mod_contabilidad.dat_estructurae mc  
 WHERE idestructuraepadre = mc.idestructurae)
```

Figura 18: Consulta SQL para llenar la tabla dim\_configuracion\_agregacion.

```
SELECT dat_consolidacionentidades.idconsolidacion,  
       dat_consolidacionentidades.idestructurae as idestructuraehijo  
FROM mod_consolidacion.dat_consolidacionentidades, mod_consolidacion.conf_consolidacion  
WHERE conf_consolidacion.idconsolidacion = dat_consolidacionentidades.idconsolidacion;
```

Figura 19: Consulta SQL para llenar la tabla dim\_configuracion\_agregacion\_hijos.

```
SELECT de.idestructurae, de.codigoestructura, de.descripcion,  
(case when de.idestructurae = de.idestructuraepadre then 0 else de.idestructuraepadre end)
```

Figura 20: Consulta SQL para llenar la tabla dim\_estructurae.

```
SELECT idestructurae, idestructuraepadre FROM mod_comun.dim_estructurae
```

Figura 21: Consulta SQL para llenar la tabla dim\_estructurae\_closure.

A la implementación del proceso ETL realizado para la tabla `cub_hecho_contabilidad_cuenta`, se les adicionó otras subconsultas para cargar los nuevos campos: `creditoacumulado` y `debitoacumulado`. Ver Figura 22.

```
(SELECT sum(resa.creditomb) FROM mod_contabilidad.res_centrocuenta_elemento_periodo resa
WHERE res.idcuenta = resa.idcuenta and res.idestructuracomun = resa.idestructuracomun
and res.idestructurae = resa.idestructurae) as creditombacumulado

(SELECT sum(resa.debitomb) FROM mod_contabilidad.res_centrocuenta_elemento_periodo resa
WHERE res.idcuenta = resa.idcuenta and res.idestructuracomun = resa.idestructuracomun
and res.idestructurae = resa.idestructurae) as debitombacumulado
```

Figura 22: Consulta SQL para llenar la tabla `dim_estructurae_closure`.

## 2.2 Solución para la replicación de los datos

### Descripción del proceso de replicación en el entorno identificado.

CedruX presenta como perspectiva futura ser desplegado en todas las empresas del país. Con el propósito de mantener actualizada toda la información, que intervengan en los procesos de agregación y consolidación, en los nodos de las sedes donde se encuentre instalado el sistema, se utilizará el software de replicación Reko.

El ambiente de replicación donde se encuentre instalado el producto será Multi-maestro con una instancia en cada nodo. Los nodos estarán registrados de formas independientes y agrupadas por etiquetas, para permitir replicar, dado una configuración, a un nodo o a varios al mismo tiempo. El entorno que se definió para la réplica, para garantizar mantener actualizado los procesos de agregación y consolidación, se encuentra representado por la siguiente tabla; ubicada en el Data Mart del subsistema de Contabilidad del Sistema Integral de Gestión CedruX: `cub_hecho_contabilidad_cuenta`, la misma almacena todas las acciones realizadas sobre las cuentas, en las entidades. Para realizarse los procesos de agregación y consolidación la entidad padre necesita que la información guardada en cada una de estas tablas, de sus entidades hijas, se repliquen hacia él. La siguiente figura representa lo descrito anteriormente.

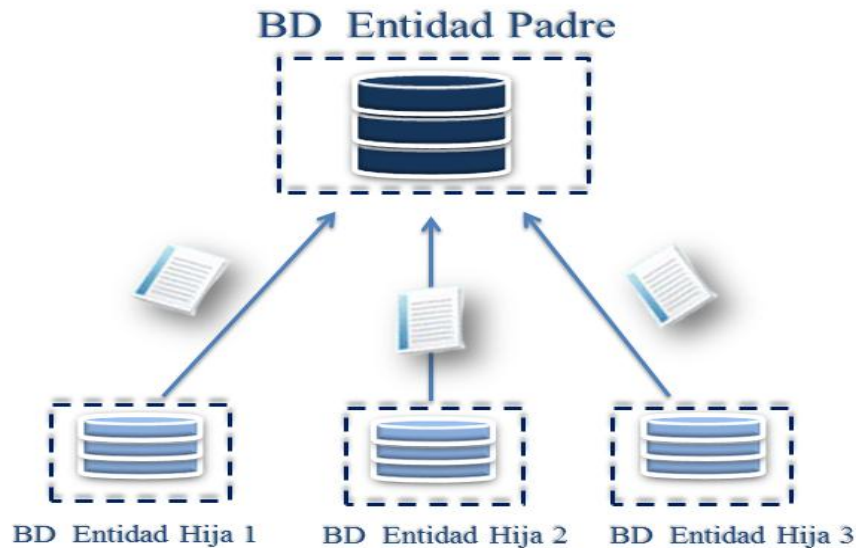


Figura 23: Representa el proceso de replicación de datos propuesto.

Se identificaron un conjunto de requisitos funcionales y no funcionales a cumplir por cada nodo a replicar.

Requisitos Funcionales:

**RF1:** La replicación debe ejecutarse a finalizar cada mes cuando se realice el cierre contable.

**RF2:** Se debe replicar todos los cambios realizados sobre el crédito, en moneda base, de las cuentas.

**RF3:** Se debe replicar todos los cambios realizados sobre el débito, en moneda base, de las cuentas.

**RF4:** Se debe replicar todos los cambios realizados sobre el crédito acumulado, en moneda base, de las cuentas.

**RF5:** Se debe replicar todos los cambios realizados sobre el débito acumulado, en moneda base, de las cuentas.

**RF6:** Se debe replicar todos los cambios realizados sobre las cuentas.

Requisitos No Funcionales:

**RFN1:** Todos los entornos deben tener instalado PostgreSQL 8.3

**RFN2:** Utilizar la versión del software Reko 2.0 en todos los nodos.

**RFN3:** Se debe garantizar la seguridad de los datos en todos los nodos.

**RFN4:** Se debe garantizar la integridad de los datos en todos los nodos.

Para garantizar el cumplimiento de la integridad y persistencia de los datos, el software escogido contiene mecanismos de chequeo de envío, confirmación de recepción y registro en una base de datos local entre los grupos replicables.

La seguridad de los datos se garantizará a través de las credenciales que manejan los nodos de replicación entre ellos, para verificar la autenticidad de los datos transferidos. Además el envío de los datos se realiza utilizando los protocolos de comunicación: HTTPS y SSL.

Para realizar el diseño de la configuración de réplica del entorno identificado, primeramente se selecciona el esquema donde se encuentran ubicada la tabla que se desea replicar, escogiéndose posteriormente la acción o las acciones que desencadenarán la replicación. La configuración que se propone depende de la acción que se realice sobre esta (inserción, actualización). Ver Figura 24.

Tabla	Inserción	Actualización	Eliminación
cub_hecho_agregacion_cuenta	No	No	No
cub_hecho_contabilidad	No	No	No
cub_hecho_contabilidad_cuenta	Sí <a href="#">Adicionar filtro Usuarios Adicionar columnas referencia</a>	Sí <a href="#">Adicionar filtro Usuarios Adicionar columnas referencia</a>	No

Figura 24: Configuración del entorno de replicación.

La replicación configurada garantizará la disponibilidad y persistencia de los datos ante posibles contingencias.

### **CONCLUSIONES PARCIALES:**

En el presente capítulo se identificaron y describieron nuevas dimensiones y hechos necesarios para ejecutar el proceso de agregación y consolidación en el Data Mart del subsistema de Contabilidad y consecuentemente se transformó el diseño del mismo desarrollándose para ello un nuevo proceso de Extracción, Transformación y Carga. Se identificó y describió un entorno de replicación que mantuviera disponibles, en la entidad padre, los datos necesarios para realizar el proceso de agregación y consolidación. Tomando como base los resultados del presente capítulo se puede dar paso a la implementación y prueba de la solución.

## CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA DE LA SOLUCIÓN.

### INTRODUCCIÓN

En el presente capítulo se implementa el proceso de agregación y se hace una descripción del mismo. Se realizan las validaciones de las transformaciones realizadas al Data Mart, del entorno de replicación propuesto y del mecanismo de datos realizado. Además se desarrollan los documentos: " *Plan de Prueba*"<sup>4</sup> y " *Resultados de Prueba*"<sup>5</sup>, con el fin de documentar el proceso y resultado de las pruebas realizadas al entorno de replicación definido.

#### 3.1 Desarrollo del proceso de agregación.

La función de agregación consiste en agrupar, según sus padres, las cuentas que se encuentren en un período de tiempo dado (*iddimtiempo*) y una configuración de una consolidación especificada (*idconsolidacion*), sumar sus créditos, débitos, créditos acumulado y débito acumulado, finalmente se insertan los valores resultantes en la tabla `cub_hecho_agregacion_cuenta`. La configuración de una consolidación representa la agregación que surge durante el proceso, la misma está relacionada por una o varias entidades que la conforman.

La siguiente figura muestra la implementación que garantizará el funcionamiento del proceso de agregación dado un tiempo determinado y una configuración específica. Ver Figura 25.

```
declare
idestructurae numeric;
idestructura numeric;
rec record;
begin
SELECT ca.idestructuraepadre, ca.idestructurapadre INTO idestructurae, idestructura
FROM mod_comun.dim_configuracion_agregacion ca WHERE ca.idconsolidacion = $1;
-- Se selecciona el identificador de la cuenta a almacenar y de esta la suma de su crédito y débito.
for rec in (SELECT
-- De la cuenta que se esta visitando se busca su cuenta padre.
(SELECT idcuentapadreclosure from mod_comun.dim_cuenta_closure where distancia = (select max(distancia)-1 as distancia)hastapadre
FROM mod_comun.dim_cuenta_closure where idcuenta = chc.idcuenta) and idcuenta = chc.idcuenta) cuentapadre,
-- se suma el crédito, débito, crédito acumulado y débito acumulado de la cuenta que se esta analizando.
SUM(chc.creditomb) as credito,
SUM (chc.debitomb) as debito,
SUM (chc.creditoacumulado) as creditoacumulado,
SUM (chc.debitoacumulado) as debitoacumulado
```

<sup>4</sup> Documento que surge como salida de la fase de diseño propuesta por la tesis de pregrado: "Propuesta de Procedimiento para la prestación del servicio de replicación de datos del Centro de Tecnologías y Gestión de Datos" (Ver en su anexo 8).

<sup>5</sup> Documento que surge como salida de la fase de diseño propuesta por la tesis de pregrado: "Propuesta de Procedimiento para la prestación del servicio de replicación de datos del Centro de Tecnologías y Gestión de Datos" (Ver en su anexo 9).



```

FROM
  mod_comun.dim_configuracion_agregacion_hijos cah,
  mod_comun.dim_configuracion_agregacion ca,
  mod_contabilidad.cub_hecho_contabilidad_cuenta chc
-- Deben cumplir con un tiempo y una consolidacion específica.
WHERE
  cah.idestructuraehijo = chc.idestructurae AND
  ca.idconsolidacion = cah.idconsolidacion AND
  ca.idconsolidacion = $1 AND
  chc.iddimtiempo = $2
-- Agrupo los resultados de cuenta repetidas en una.
GROUP BY
  cuentapadre)
loop
-- Inserta en la tabla hecho los valores de la agregación.
INSERT INTO mod_contabilidad.cub_hecho_agregacion_cuenta(
  idconfiguracion, idcuenta, idestructurae, idestructura,
  iddimtiempo, creditomb, debitomb, valorconsolidacion, creditoacumulado, debitoacumulado)
VALUES ($1, rec.cuentapadre, idestructurae, idestructura,
  $2, rec.credito, rec.debito, 0, rec.creditoacumulado, rec.debitoacumulado);
end loop;
end;

```

Figura 25: Representa el proceso de agregación para una consolidación específica.

La siguiente figura muestra la implementación que realizará el proceso de agregación para las configuraciones existentes en la tabla dim\_configuracion\_agregacion. Ver Figura 26.

```

declare
  idconsolidacion record;
  iddimtiempo record;

begin
  for idconsolidacion in (SELECT ca.idconsolidacion FROM mod_comun.dim_configuracion_agregacion ca)
  loop
    for iddimtiempo in (SELECT DISTINCT (cc.iddimtiempo) FROM mod_contabilidad.cub_hecho_contabilidad_cuenta cc)
    loop
      execute mod_comun.agregar_cuentas(idconsolidacion.idconsolidacion,iddimtiempo.iddimtiempo);
    end loop;
  end loop;
end;

```

Figura 26: Representa el proceso de agregación para todas las consolidación existentes.

### **3.2 Validación del mecanismo de datos.**

La validación tiene como objetivo comprobar que se está construyendo el producto correcto, examinar la validez de los resultados, constatar el cumplimiento de las necesidades definidas y de los requisitos del usuario. [44].

Para la validación de este trabajo se definieron tres momentos:

#### **3.2.1 Validación de las transformaciones realizadas al Data Mart del subsistema de Contabilidad.**

Las pruebas de rendimiento permiten determinar la rapidez con que realiza una tarea un sistema en condiciones particulares de trabajo. También puede servir para validar y verificar otros atributos de la calidad del sistema, tales como la escalabilidad, fiabilidad y uso de los recursos.

Existen varios tipos de pruebas de rendimiento, entre las que se encuentran:

- Prueba de carga.
- Prueba de estabilidad.
- Prueba de picos. [45]

Las escogidas para validar y probar las transformaciones realizadas al Data Mart fueron las pruebas de carga y estrés.

##### Pruebas de carga.

Son pruebas que se realizan para someter a la base de datos a un régimen de carga de trabajo similar al esperado en la explotación real del sistema. Permiten verificar y validar el desempeño de un elemento de un sistema bajo diferentes condiciones de carga:

- Número de usuarios.
- Número de transacciones.

Son importantes cuando los sistemas tienen que soportar un gran volumen de usuarios o transacciones concurrentes. [45]

##### Pruebas de estrés.

Somete a la aplicación y/o a la base de datos a un régimen de carga de trabajo muy superior al esperado en la explotación real. Permiten verificar la aceptabilidad del desempeño de la BD ante condiciones

anormales o extremas, dígame volumen de usuarios/transacciones extremadamente alto y escasos recursos (poco ancho de banda, memoria reducida, espacio en disco reducido).

En realidad las condiciones de estrés no se espera que sucedan, pero permiten documentar las condiciones bajo las cuales el sistema puede fallar. [45]

Es común que durante el proceso de las pruebas de carga y estrés se encuentren un conjunto de problemas, entre ellos: consultas mal diseñadas, necesidad de índices adicionales, consultas candidatas a optimización, código mal diseñado, tiempos de demora en las respuestas de magnitudes inaceptables, hardware insuficiente (RAM, CPU; tasa de acceso a disco, tasa de transferencia de red), problemas de control de concurrencia, etc. [45]

Para la realización de las pruebas a efectuar sobre Data Mart se usó la herramienta Jmeter, mediante esta se pueden simular diferentes situaciones y escenarios. Para el caso pertinente, las características del servidor de aplicaciones son las siguientes:

- El servidor de aplicaciones debe estar corriendo.
- Las características del servidor de aplicaciones es de 1GB de RAM con 160 GB de disco duro.
- Las características de la red 100 Mbps.

Los resultados arrojados por las pruebas fueron recogidos en la siguiente tabla:

Escenario general	Escenario de la sección	Carga de trabajo	Descripción	Resultado de la prueba
EC 1: Realizar la carga de todas las dimensiones insertadas.	EC 1.1: Realizar la carga de las peticiones realizadas, para una concurrencia de 10 usuarios.	10	Realizar la carga de las peticiones realizadas.	0.00% de errores.
	EC 1.2: Realizar la carga de las peticiones realizadas, para una concurrencia de 60 usuarios	60	Realizar la carga de las peticiones realizadas.	0.00% de errores

	EC 1.3: Realizar la carga de las peticiones realizadas, para una concurrencia de 120 usuarios	120	Realizar la carga de las peticiones realizadas.	0.00% de errores
	EC 1.4: Realizar la carga de las peticiones realizadas, para una concurrencia de 200 usuarios	200	Realizar la carga de las peticiones realizadas.	0.00% de errores

*Tabla 13: Pruebas de carga y estrés a la aplicación.*

Los resultados de las pruebas han sido satisfactorios, siendo el tiempo promedio de respuesta de las consultas insertadas de 2,215 segundos, para un total de 950 peticiones al servidor. El porcentaje de errores es de 0.00%, lo cual nos permite certificar que el Data Mart implementado funciona correctamente con un servidor de BD que cuente como mínimo con 1 GB de RAM, 160 GB de disco duro y que la red de datos cuente con una velocidad mínima de 100Mbits.

El proceso realizado a través de la aplicación Jmeter se puede observar en los Anexo 2, 3, 4.

### **3.2.2 Validación del proceso de agregación.**

Para validar el funcionamiento del proceso de agregación se realizaron pruebas de aceptación, las que consisten en revisar que la solución desarrollada cumpla con los requisitos identificados. Existen varios ejemplos de estas pero las realizadas fueron las pruebas alfa<sup>6</sup>. Para llevar a cabo el desarrollo de las mismas se creó un conjunto de datos con los cuales se realizó reiteradamente dicho proceso; los resultados obtenidos se sometieron a pruebas de observación, examinando que fuesen correctos.

Los valores que se obtuvieron producto de las pruebas realizadas al proceso de agregación se pueden observar en la siguiente figura. El conjunto de datos que se utilizó se encuentra registrado en el Anexo 5.

---

<sup>6</sup> Realizadas por el usuario junto al desarrollador, como observador, en un entorno controlado (simulación de un entorno de producción). Es la primera versión del programa.

idhecho [PK] numeric	idconfiguracion numeric(19,0)	idcuenta numeric(19,0)	idestructurae numeric(19,0)	idestructura numeric(19,0)	iddintiempo numeric(19,0)	creditomb numeric(19,2)	debitomb numeric(19,2)	valorconsolid numeric(19,2)	creditoacumu numeric(19,2)	debitoacumu numeric(19,2)
194	900000000000000	90000	9000000	900000000	9900001001	86.00	85.00	0.00	186.00	185.00
195	900000000000000	90005	9000000	900000000	9900001001	19.00	12.00	0.00	27.00	20.00
196	900000000000001	90000	9000000	900000000	9900001001	86.00	85.00	0.00	186.00	185.00
197	900000000000001	90005	9000000	900000000	9900001001	19.00	12.00	0.00	27.00	20.00
198	900000000000002	90000	9000000	900000000	9900001001	86.00	85.00	0.00	186.00	185.00
199	900000000000002	90005	9000000	900000000	9900001001	19.00	12.00	0.00	27.00	20.00
200	900000000000003	90000	9000000	900000000	9900001001	86.00	85.00	0.00	186.00	185.00
201	900000000000003	90005	9000000	900000000	9900001001	16.00	9.00	0.00	22.00	15.00
202	900000000000004	90000	9000000	900000000	9900001001	4.00	4.00	0.00	4.00	4.00
203	900000000000004	90005	9000000	900000000	9900001001	3.00	3.00	0.00	5.00	5.00
204	900000000000003	90000	9000000	900000000	9900001001	90.00	89.00	0.00	190.00	189.00
205	900000000000003	90005	9000000	900000000	9900001001	16.00	9.00	0.00	22.00	15.00
206	900000000000003	90000	9000000	900000000	9900001001	90.00	89.00	0.00	190.00	189.00
207	900000000000003	90005	9000000	900000000	9900001001	16.00	9.00	0.00	22.00	15.00

Figura 27: Registra los resultados del proceso de agregación realizado.

A partir de la valoración de los resultados alcanzados se puede garantizar el buen funcionamiento del proceso de agregación desarrollado.

### 3.2.3 Validación del correcto funcionamiento del entorno de replicación.

Con el propósito de validar la solución de replicación se realizaron un conjunto de pruebas. Durante la fase de la preparación de las mismas se comprobó la disponibilidad de recursos humanos y técnicos necesarios. A partir del cumplimiento de lo establecido en el plan de descripción de la solución y se preparan las condiciones que permitirán representar las situaciones previstas para las pruebas, esto se recoge en el documento "Plan de Prueba". Ver el documento.

Para medir la calidad y garantizar el buen funcionamiento en el entorno identificado el tipo de pruebas realizadas fueron de carga y de almacenamiento, trazándose para esto las siguientes precondiciones y poscondiciones sobre un entorno de prueba, donde van a existir dos nodos, de donde sale la información y a donde se desea replicar la misma.

#### **Precondiciones:**

##### Replicación del nodo1 (nodo hijo) al nodo2 (nodo padre)

- La tabla cub\_hecho\_contabilidad\_cuenta del *nodo1* se le insertó tres nuevas tuplas. La figura 28 muestra el estado de la tabla antes de las inserciones y en la figura 29 luego de realizarse estas.

	idhecho [PK] numeric	idestructura numeric(19,0)	idcuenta numeric(19,0)	idcentro numeric(19,0)	idelemento numeric(19,0)	iddimtiempo numeric(19,0)	creditomb numeric(19,2)	debitomb numeric(19,2)	idestructurae numeric(19,0)	creditoacumu numeric(19,2)	debitoacumu numeric(19,2)
120	9900001455	900000000	90036	90000000006	90000000006	9900000993	0.00	0.00			
121	9900001456	900000000	90036	90000000006	90000000018	9900000977	0.00	100.00			
122	9900001457	900000000	90036	90000000006	90000000046	9900000992	0.00	0.00			
123	9900001458	900000000	90036	90000000006	90000000046	9900000993	0.00	20020.00			
124	9900001459	900000000	90036	90000000006	92233600000	9900000992	0.00	4204.85			
125	9900001460	900000000	90036	90000000006	92233600000	9900000993	0.00	9176.70			
126	9900001461	900000000	90212	90000000004	90000000006	9900001005	0.00	10.00			
127	9900001462	900000000	90281	90000000004	90000000137	9900000992	0.00	8098.26			
128	9900001463	900000000	90281	90000000004	90000000137	9900000993	0.00	23377.38			
129	9900001464	900000000	90281	90000000006	90000000137	9900000992	0.00	7605.90			
130	9900001465	900000000	90281	90000000006	90000000137	9900000993	0.00	24776.70			
131	9900001466	900000000	90281	90000001591	90000000137	9900000977	-20.00	0.00			
132	9900001467	900000000	90288	92233600000		9900000977	0.00	300.00			
133	9900001468	900000000	90288	92233600000		9900000981	0.00	20.00			
134	9900001469	900000000	90290	92233600000	92233600000	9900000977	-40.00	0.00			
135	9900001470	900000006	90024			9900000980	4895.00	3807.33	9000073	4895.00	3807.33
136	9900001471	900000006	90025			9900000980	0.00	1087.67	9000073	0.00	1087.67
137	9900001472	900000132	90149			9900000989	90.00	0.00	9000092	604.00	0.00
138	9900001473	900000132	90158			9900000989	0.00	90.00	9000092	0.00	90.00

Figura 28: Estado inicial de la tabla cub\_hecho\_contabilidad\_cuenta del nodo1.

	idhecho [PK] numeric	idestructura numeric(19,0)	idcuenta numeric(19,0)	idcentro numeric(19,0)	idelemento numeric(19,0)	iddimtiempo numeric(19,0)	creditomb numeric(19,2)	debitomb numeric(19,2)	idestructurae numeric(19,0)	creditoacumu numeric(19,2)	debitoacumu numeric(19,2)
121	9900001456	900000000	90036	90000000006	90000000018	9900000977	0.00	100.00			
122	9900001457	900000000	90036	90000000006	90000000046	9900000992	0.00	0.00			
123	9900001458	900000000	90036	90000000006	90000000046	9900000993	0.00	20020.00			
124	9900001459	900000000	90036	90000000006	92233600000	9900000992	0.00	4204.85			
125	9900001460	900000000	90036	90000000006	92233600000	9900000993	0.00	9176.70			
126	9900001461	900000000	90212	90000000004	90000000006	9900001005	0.00	10.00			
127	9900001462	900000000	90281	90000000004	90000000137	9900000992	0.00	8098.26			
128	9900001463	900000000	90281	90000000004	90000000137	9900000993	0.00	23377.38			
129	9900001464	900000000	90281	90000000006	90000000137	9900000992	0.00	7605.90			
130	9900001465	900000000	90281	90000000006	90000000137	9900000993	0.00	24776.70			
131	9900001466	900000000	90281	90000001591	90000000137	9900000977	-20.00	0.00			
132	9900001467	900000000	90288	92233600000		9900000977	0.00	300.00			
133	9900001468	900000000	90288	92233600000		9900000981	0.00	20.00			
134	9900001469	900000000	90290	92233600000	92233600000	9900000977	-40.00	0.00			
135	9900001470	900000006	90024			9900000980	4895.00	3807.33	9000073	4895.00	3807.33
136	9900001471	900000006	90025			9900000980	0.00	1087.67	9000073	0.00	1087.67
137	9900001472	900000132	90149			9900000989	90.00	0.00	9000092	604.00	0.00
138	9900001473	900000132	90158			9900000989	0.00	90.00	9000092	0.00	90.00
139	9900001474	900000132	90158			9900000989	2000.00	90.00	9000092	7.00	90.00
140	9900001475	900000132	90158			9900000989	2000.00	90.00	9000092	70.00	910.00
141	9900001476	900000000	90036			9900000989	100.00	1100.00	9000092	1100.00	100.00

Figura 29 Estado de la tabla cub\_hecho\_contabilidad\_cuenta del nodo1 luego de insertadas las tuplas.

- La tabla cub\_hecho\_contabilidad\_cuenta del nodo2 carece de las tres tuplas que se le insertaron a la tabla cub\_hecho\_contabilidad\_cuenta del nodo1. Ver figura 30.

	idhecho [PK] numeric(19,0)	idestructura numeric(19,0)	idcuenta numeric(19,0)	idcentro numeric(19,0)	idelemento numeric(19,0)	iddit tiempo numeric(19,0)	creditomb numeric(19,2)	debitomb numeric(19,2)	idestructurae numeric(19,0)	creditoacumu numeric(19,2)	debitoacumu numeric(19,2)
119	9900001454	900010000	90036	9000000006	9000000006	9900000993	0.00	0.00			
120	9900001455	900010000	90036	9000000006	9000000006	9900000993	0.00	0.00			
121	9900001456	900010000	90036	9000000006	9000000018	9900000977	0.00	100.00			
122	9900001457	900010000	90036	9000000006	9000000046	9900000992	0.00	0.00			
123	9900001458	900010000	90036	9000000006	9000000046	9900000993	0.00	20020.00			
124	9900001459	900010000	90036	9000000006	9213360000	9900000992	0.00	4204.35			
125	9900001460	900010000	90036	9000000006	9213360000	9900000993	0.00	9176.70			
126	9900001461	900010000	90212	9000000004	9000000006	9900001005	0.00	10.00			
127	9900001462	900010000	90231	9000000004	90000000137	9900000992	0.00	8098.26			
128	9900001463	900010000	90231	9000000004	90000000137	9900000993	0.00	23377.38			
129	9900001464	900010000	90231	9000000006	90000000137	9900000992	0.00	7605.90			
130	9900001465	900010000	90231	9000000006	90000000137	9900000993	0.00	24776.70			
131	9900001466	900010000	90231	90000001591	90000000137	9900000977	-20.00	0.00			
132	9900001467	900010000	90238	9223360000		9900000977	0.00	300.00			
133	9900001468	900010000	90238	9223360000		9900000981	0.00	20.00			
134	9900001469	900010000	90230	9223360000	9213360000	9900000977	-40.00	0.00			
135	9900001470	900010000	90024			9900000980	4895.00	3807.33	9000073	4895.00	3807.33
136	9900001471	900010000	90025			9900000980	0.00	1087.67	9000073	0.00	1087.67
137	9900001472	900010132	90149			9900000989	90.00	0.00	9000092	604.00	0.00
138	9900001473	900010132	90158			9900000989	0.00	90.00	9000092	0.00	90.00

Figura 30: Estado de la tabla cub\_hecho\_contabilidad\_cuenta del nodo2 antes de realizar la replicación.

**Poscondiciones:**

- La tabla cub\_hecho\_contabilidad\_cuenta del nodo1 no cambia.
- La tabla cub\_hecho\_contabilidad\_cuenta del nodo2 presenta las tres tuplas que se le insertaron a la tabla cub\_hecho\_contabilidad\_cuenta del nodo1. Ver figura 31.

	idhecho [PK] numeric	idestructura numeric(19,0)	idcuenta numeric(19,0)	idcentro numeric(19,0)	idelemento numeric(19,0)	iddimtiempo numeric(19,0)	creditomb numeric(19,2)	debitomb numeric(19,2)	idestructurae numeric(19,0)	creditoacumu numeric(19,2)	debitoacumu numeric(19,2)
121	9900001456	900000000	90036	90000000006	90000000018	9900000977	0.00	100.00			
122	9900001457	900000000	90036	90000000006	90000000046	9900000992	0.00	0.00			
123	9900001458	900000000	90036	90000000006	90000000046	9900000993	0.00	20020.00			
124	9900001459	900000000	90036	90000000006	92233600000	9900000992	0.00	4204.85			
125	9900001460	900000000	90036	90000000006	92233600000	9900000993	0.00	9176.70			
126	9900001461	900000000	90212	90000000004	90000000006	9900001005	0.00	10.00			
127	9900001462	900000000	90281	90000000004	90000000137	9900000992	0.00	8098.26			
128	9900001463	900000000	90281	90000000004	90000000137	9900000993	0.00	23377.38			
129	9900001464	900000000	90281	90000000006	90000000137	9900000992	0.00	7605.90			
130	9900001465	900000000	90281	90000000006	90000000137	9900000993	0.00	24776.70			
131	9900001466	900000000	90281	90000001591	90000000137	9900000977	-20.00	0.00			
132	9900001467	900000000	90288	92233600000		9900000977	0.00	300.00			
133	9900001468	900000000	90288	92233600000		9900000981	0.00	20.00			
134	9900001469	900000000	90290	92233600000	92233600000	9900000977	-40.00	0.00			
135	9900001470	900000006	90024			9900000980	4895.00	3807.33	9000073	4895.00	3807.33
136	9900001471	900000006	90025			9900000980	0.00	1087.67	9000073	0.00	1087.67
137	9900001472	900000132	90149			9900000989	90.00	0.00	9000092	604.00	0.00
138	9900001473	900000132	90158			9900000989	0.00	90.00	9000092	0.00	90.00
139	9900001474	900000132	90158			9900000989	2000.00	90.00	9000092	7.00	90.00
140	9900001475	900000132	90158			9900000989	2000.00	90.00	9000092	70.00	910.00
141	9900001476	900000000	90036			9900000989	100.00	1100.00	9000092	1100.00	100.00

Figura 31: Estado de la tabla cub\_hecho\_contabilidad\_cuenta del nodo2 luego de realizada la replicación.

La siguiente imagen muestra el monitoreo de cada una de las replicaciones realizadas desde el software Reko 2.0.

Monitor en tiempo real ( Iniciado 05/06/2012 03:55 p.m. )							
<input checked="" type="radio"/> Ambas direcciones		<input type="radio"/> Salida	<input type="radio"/> Entrada	<input type="radio"/> Detener Monitor	Acciones por mostrar: 0	Acciones recibidas: 0	Acciones enviadas: 4
Dirección	Fecha	Remitente	Destinos	SQL	Tabla	Tipo de acción	
	5/6/2012 15:55:23	nodo1	nodo2: Enviado	insert into "mod_contabilidad"."cub_hecho_contabilidad_cuenta" ("creditoacumulado", "idestructurae", "idcuenta", "idhecho", "iddimtiempo", "idelemento", "creditomb", "idcentro", "debitoacumulado", "debitomb", "idestructura") values (0.00, 9000092, 90036, 9900001482, 9900000977, null, 0.00, null, 0.00, 0.00, 900000000)	mod_contabilidad.cub_hecho_α	Inserción	
	5/6/2012 15:55:23	nodo1	nodo2: Enviado	insert into "mod_contabilidad"."cub_hecho_contabilidad_cuenta" ("creditoacumulado", "idestructurae", "idcuenta", "idhecho", "iddimtiempo", "idelemento", "creditomb", "idcentro", "debitoacumulado", "debitomb", "idestructura") values (19999090.00, 9000092, 90036, 9900001481, 9900000977, null, 1999990.00, null, 0.00, 0.00, 900000000)	mod_contabilidad.cub_hecho_α	Inserción	
	5/6/2012 15:55:23	nodo1	nodo2: Enviado	insert into "mod_contabilidad"."cub_hecho_contabilidad_cuenta" ("creditoacumulado", "idestructurae", "idcuenta", "idhecho", "iddimtiempo", "idelemento", "creditomb", "idcentro", "debitoacumulado", "debitomb", "idestructura") values (0.00, 9000092, 90036, 9900001480, 9900000977, null, 0.00, null, 0.00, 0.00, 900000000)	mod_contabilidad.cub_hecho_α	Inserción	
	5/6/2012 15:55:23	nodo1	nodo2: Enviado	insert into "mod_contabilidad"."cub_hecho_contabilidad_cuenta" ("creditoacumulado", "idestructurae", "idcuenta", "idhecho", "iddimtiempo", "idelemento", "creditomb", "idcentro", "debitoacumulado", "debitomb", "idestructura") values (0.00, 9000092, 90036, 9900001479, 9900000977, null, 0.00, null, 0.00, 0.00, 900000000)	mod_contabilidad.cub_hecho_α	Inserción	

Figura 32: Monitoreo de las operaciones de replicación realizadas.



A partir de los resultados concebidos en las pruebas realizadas anteriormente, se generó el documento "Resultados de la Prueba". Ver el documento.

Luego de la evaluación y valoración de estos se puede afirmar que la replicación entre los nodos descritos, para el entorno identificado, cumple con los resultados esperados y garantiza la disponibilidad y efectividad de la información, con un tiempo de respuesta satisfactorio.

### **CONCLUSIONES PARCIALES**

Se desarrolló la implementación del mecanismo de gestión de datos para el proceso agregación, garantizando el funcionamiento del mismo en el subsistema de Contabilidad. Para evaluar su funcionamiento se realizaron pruebas de observación al entorno de réplica planteado y a la solución del proceso de agregación desarrollado, siendo satisfactorio los resultados arrojados, además se demostró que las transformaciones realizadas al Data Mart cumplieran con las expectativas planteadas inicialmente.

## **CONCLUSIONES**

En el presente trabajo se hizo un estudio de los principales conceptos de la Contabilidad General mediante la cual se logró la comprensión de los procesos de agregación y consolidación, se seleccionaron las herramientas, tecnología y software que facilitaron el desarrollo de los mismos.

A partir de los pasos definidos por la metodología UCID, para el desarrollo de almacenes de datos, se realizó el rediseño al Data Mart del subsistema de Contabilidad y el proceso ETL para los cambios introducidos al mismo. Se identificó y configuró un entorno de replicación que mantuviera disponible, en las entidades padres, los datos necesarios para realizar el proceso de agregación y consolidación.

Se implementó el mecanismo de gestión de datos para los procesos de consolidación y agregación en el subsistema Contabilidad de Cedrux que garantiza la disponibilidad de los mismos.

Se realizaron pruebas de observación al entorno de réplica planteado, cumpliendo los objetivos específicos trazados. Además se validó la solución del proceso de agregación y las transformaciones realizadas al Data Mart, siendo satisfactorio los resultados arrojados.

A partir de lo antes expuesto se puede afirmar que se le dio cumplimiento al objetivo general planteado, resultado que permitirá a los usuarios tener un conocimiento exacto de las realidades económicas de las entidades.

## **RECOMENDACIONES**

Para un posterior desarrollo y perfeccionamiento del mecanismo de datos desarrollado se recomienda:

- Realizar el Comprobante de Operaciones para mostrar los resultados del proceso de consolidación.
- Continuar perfeccionando el Data Mart para el cálculo de otros indicadores financieros.
- Realizar la configuración de réplica para otros entornos dentro del Sistema Integral Cedrux.

## REFERENCIAS BIBLIOGRÁFICAS

- [1]. **CORTÉS, J.** *Contabilidad General*. Barcelona : Mentensó, 1932. pág. 2.
- [2]. **AYAVIRI, G. D.** *Contabilidad básica y Documentos mercantiles*. Argentina : "N-DAG". págs. 1,2,3.
- [3]. **Maldonado, R.** *Estudio de la Contabilidad General*. 2006.
- [4]. **Tovar, J.C.T.** "Introducción a la contabilidad I."
- [5]. **Suárez Sánchez, Dayli y Silveira Palacio, Yudisleysis.** *Propuesta de procedimiento para la presentación de replicación de datos del Centro de Tecnologías y Gestión de Datos*. Facultad 8, Universidad de la Ciencias Informática. Ciudad de la Habana : s.n., 2010. págs. 7,8, Trabajo de diploma.
- [6]. **Valmaseda Ortiz, Marcos.** *Programación "Master-Master con MySQL 5.0 en Debian Etch"*. Universidad de la Ciencias Informáticas. La Habana : s.n., 2006, Revista de Software Libre de la UCI, Vol. 2.
- [7]. **Pupo Polanco, Rosell y González Pérez, Yenier.** *Implementación del componente réplica de base de datos para Akademos v2.0*. Universidad de las Ciencias Informáticas. Universidad de las Ciencias Informáticas. La Habana : s.n., 2009. trabajo de pregrado.
- [8]. *Propuesta de una solución de replicación para el gestor de bases de datos PostgreSQL en la Oficina Nacional de Recursos Minerales*. Facultad 9, Universidad de las Ciencias Informática. La Habana : s.n., 2010. trabajo de pregrado.
- [9]. **Rodríguez García, Kenny y Baños Fernández, Eddy Ernesto.** *Sistema de réplica para la Intranet Corporativa y el Sitio en Internet de PDVSA*. Universidad de las Ciencias Informática. La Habana : s.n., 2009. trabajo de pregrado.
- [10]. **PENTAHO.** *Pentaho Data Integration.Spoon 2.3.1*. Pentaho. 2008. Manual de Usuario.
- [11]. **Espinoza, Humberto.** *Una Alternativa de DBMS Open Source*. PostgreSQL. 2007, Consultores Informáticos Integrales, págs. 5,6,7,8,9.
- [12]. **Peña Pérez, Anaís y Miranda Estrada, Anay.** *Desarrollo de la funcionalidad para la programación de captura y envío de datos en el Replicador Reko*. Facultad 6, Universidad de las Ciencias Informáticas. La Habana : s.n., 2010. págs. 12,13, trabajo de pregrado.
- [13]. **Martín Escofet, Carme.** UOC. [En línea] [Citado el: 7 de 12 de 2011.] [http://ocw.uoc.edu/computer-science-technology-and-multimedia/bases-de-datos/bases-de-datos/P06\\_M2109\\_02149.pdf](http://ocw.uoc.edu/computer-science-technology-and-multimedia/bases-de-datos/bases-de-datos/P06_M2109_02149.pdf). P06/M2109/02149..
- [14]. [Guia-ubuntu.org](http://guia-ubuntu.org). [En línea] PgAdmin\_III. [Citado el: 7 de 12 de 2011.] [http://guia-ubuntu.org/index.php?title=PgAdmin\\_III](http://guia-ubuntu.org/index.php?title=PgAdmin_III).

- [15]. **Arévalo M., Mauricio.** Scribd. [En línea] 2011. [Citado el: 8 de 12 de 2011.] <http://es.scribd.com/doc/63764883/64/PgAdmin>.
- [16]. **Sierra, María.** *Trabajando con Visual Paradigm for UML.* [ed.] Univ. Cantabria Fac de Ciencias. 2008, INGENIERÍA DEL SOFTWARE I Práctica 1.
- [17]. **Fernández Fernández, Karina y Navarro Fonseca, Carlos.** *Análisis, diseño e implementación del mercado de datos para los indicadores sobre el Plan Turquino para la Oficina Nacional de Estadísticas.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2010.
- [18]. Gravatar. [En línea] 2011. [Citado el: 8 de 12 de 2011.] <http://www.gravatar.biz/index.php/herramientas-bi/pentaho/caracteristicas-pentaho/>.
- [19]. *Descripción de las funcionalidades del ERP.* Informática, Universidad de las Ciencias, La Habana : s.n., 2010, Centro para la Informatización de Gestión de Entidades, págs. 7,17,18,19,20.
- [20]. **Pérez Alfonso, Damián y Mena Rodríguez, Luis Edgardo.** *Módulo de Reko para la replicación entre bases de datos con estructuras diferentes.* Facultad 15, Universidad de las Ciencias Informáticas. La Habana : s.n., 2010. págs. 10,11,13,14,15,16, trabajo de pregrado.
- [21]. **Cabrera González, Lic. Miguel P.** *SISTEMA ECONÓMICO INTEGRADO.* Villa Clara : CASA DEL SOFTWARE, 2009, XV FORUM DE CIENCIA Y TECNICA, págs. 4,6.
- [22]. *Sistema Integral de Gestión Económica Administrativa.* La Habana : s.n., 2008, Unidad Empresarial de Base División de Sistemas Gerenciales, págs. 3,4,13.
- [23]. **Rodríguez Sánchez, Ing. Tamara , Fernández González, Ing. Mairelys y Cabrera Casas, Ing. Eliecer .** *La gestión empresarial de las entidades cubanas. Cedrux a la vuelta de la esquina. No. 2,* La Habana : s.n., abril- junio de 2011, Vol. 15. [24]. **Collado, Albert.** uniriioja. [En línea] 6 de 8 de 2009. [Citado el: 3 de 5 de 2011.] <http://dialnet.uniriioja.es/servlet/articulo?codigo=2241561>.
- [25]. SAP. [En línea] [Citado el: 4 de 6 de 2012.] <http://www.sap.com/mexico/solutions/business-suite/erp/sistemas-erp.ep>.
- [26]. **Guanche Cañizares, M.** *Desarrollo de un DataMart para la obtención de las razones financieras de los subsistemas de Cobros y Pagos, Caja y Banco del proyecto ERP-Cuba.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2011.
- [27]. **Valdés Yero, Livan.** *Desarrollo de un Data Mart para la obtención de los indicadores financieros de los subsistemas de Contabilidad y Costos y Procesos del proyecto ERP.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2011.

- [28]. **Texidó Ricardo, Lisney.** *Modelado de los procesos Agregación y Consolidación del subsistema Contabilidad General.* La Habana : s.n., 2011. trabajo de pregrado.
- [29]. Real Academia Española. [En línea] 2010. [Citado el: 12 de 1 de 2011.] <http://buscon.rae.es/draeI/SrvltConsulta?LEMA=modelo>.
- [30]. *Conferencia 2: Modelo de Datos.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2011. Conferencia .
- [31]. *Conferencia 3: Modelo Relacional.* Universidad de las Ciencias Informáticas. La Habana : s.n., 2011.
- [32]. **Colombia, Universidad Nacional de.** [En línea] [Citado el: 12 de 1 de 2011.] <http://www.virtual.unal.edu.co/cursos/sedes/manizales/4060029/lecciones/cap4-1.html>.
- [33]. **Ralph Kimball, Margy Ross.** *The Data WareHouse Toolkit Second Edition.* [ed.] John Wiley and Sons, Inc. Canada : s.n., 2002.
- [34]. **Colomé, Yanersy Díaz.** Entorno Virtual de Aprendizaje. [En línea] [Citado el: 25 de 2 de 2012.] [http://eva.uci.cu/file.php/84/3.\\_Tema\\_2/Semana\\_9/Conferencia\\_6/Bibliografia\\_de\\_Consulta/Modelo\\_Multidimensional.\\_Sus\\_elementos.pdf](http://eva.uci.cu/file.php/84/3._Tema_2/Semana_9/Conferencia_6/Bibliografia_de_Consulta/Modelo_Multidimensional._Sus_elementos.pdf).
- [35]. **Llerena, Michel Díaz.** *DATA MART PARA LA GESTIÓN DEL CONOCIMIENTO.* 2007.
- [36]. **Colomé, Yanersy Díaz.** Entorno Virtual de Aprendizaje. [En línea] [Citado el: 25 de 2 de 2012.] [http://eva.uci.cu/file.php/84/3.\\_Tema\\_2/Semana\\_9/Conferencia\\_6/Bibliografia\\_de\\_Consulta/Introduccion\\_Modelo\\_Multidimensional.pdf](http://eva.uci.cu/file.php/84/3._Tema_2/Semana_9/Conferencia_6/Bibliografia_de_Consulta/Introduccion_Modelo_Multidimensional.pdf).
- [37]. —. Entorno Virtual de Aprendizaje. [En línea] [Citado el: 4 de 2 de 2012.] [http://eva.uci.cu/file.php/84/3.\\_Tema\\_2/Semana\\_9/Conferencia\\_8/Conf8\\_Modelo\\_Multidimensional.pdf](http://eva.uci.cu/file.php/84/3._Tema_2/Semana_9/Conferencia_8/Conf8_Modelo_Multidimensional.pdf).
- [38]. **Ramírez Hernández, Ing. Javier y Bermúdez Pérez, Ing. Henry Ernesto.** *SOLUCIÓN BASE PARA EL DESARROLLO DE SISTEMAS DE CONTROL LOGÍSTICO DE CEDRUX.* Universidad de las Ciencias Informáticas. La Habana : s.n., septiembre 2010.
- [39]. **del Toro Ríos, José Carlos y González Brito, Henry Raúl.** *Documento Visión proyecto ERP Cuba v2.0.* La Habana: s.n., 2008.
- [40]. **Rojas, T., Pérez, M. y A., Grimán.** *MODELO DE DECISIÓN PARA SOPORTAR LA SELECCIÓN DE HERRAMIENTAS CASE.* Caracas : Universidad Simón Bolívar, 2010, Departamento de Procesos y Sistemas – LISI.
- [41]. —. Entorno Virtual de Aprendizaje. [En línea] [http://eva.uci.cu/file.php/84/3.\\_Tema\\_2/Semana\\_9/Conferencia\\_6/Conf7\\_Modelo\\_Multidimensional.pdf](http://eva.uci.cu/file.php/84/3._Tema_2/Semana_9/Conferencia_6/Conf7_Modelo_Multidimensional.pdf).

- [42]. **Dario, Ing. Bernabeu Ricardo.** *Metodología para la Construcción de un Data Warehouse.* Córdoba : s.n., 2010.
- [43]. Aula de Economía.com. [En línea] 2005. [Citado el: 11 de 10 de 2011.] <http://www.auladeeconomia.com/mercados2.htm>.
- [44]. **Moret Bonillo, Vicente, Mosqueira Rey, Eduardo y Hernández Pereira, Elena.** *VALIDACIÓN Y USABILIDAD DE SISTEMAS INFORMÁTICOS.* Departamento de Calidad. 2010. Curso de Doctorado.
- [45]. **Colomé, Yanersy Díaz.** Entorno Virtual de Aprendizaje. [En línea] [Citado el: 2012 de 2 de 4.] [http://eva.uci.cu/file.php/84/4.\\_Tema\\_3/Semana\\_15/Laboratorio\\_11/Materiales\\_para\\_el\\_Laboratorio/Validacion\\_de\\_Bases\\_de\\_Datos.pdf](http://eva.uci.cu/file.php/84/4._Tema_3/Semana_15/Laboratorio_11/Materiales_para_el_Laboratorio/Validacion_de_Bases_de_Datos.pdf).
- [46]. articuloz. [En línea] [Citado el: 6 de 1 de 2012.] <http://www.articuloz.com/software-articulos/herramienta-para-la-ejecucion-de-las-pruebas-886032.html>.

## ANEXOS

### Anexo 1: Mapa lógico de datos.

Mapa lógico de datos					
Nombre BD fuente: Cedrux			Nombre BD destino: Data Mart		
Nombre de la tabla	Nombre del campo	Tipo de dato	Nombre de la tabla	Nombre del campo	Tipo de dato
dat_estructurae	idestructurae	numeric(19,0)	dim_estructurae	idestructurae	numeric(19,0)
dat_estructurae	codigoestructura	varchar(50)	dim_estructurae	codigoestructura	character varying(50)
dat_estructurae	descripcion	varchar(255)	dim_estructurae	descripcion	character varying(255)
dat_estructurae	idestructuraepadre	numeric(19,0)	dim_estructurae	idestructuraepadre	numeric(19,0)
conf_consolidacion	idconsolidacion	numeric(19,0)	dim_configuracion_agregacion	idconsolidacion	numeric(19,0)
conf_consolidacion	nombre	character varying(50)	dim_configuracion_agregacion	nombre	character varying(50)



			n		
conf_consolidacion	idestructurae	numeric(19,0)	dim_configuracion_agregacion	idestructuraepadre	numeric(19,0)
conf_consolidacion	idestructuracomun	numeric(19,0)	dim_configuracion_agregacion	idestructurapadre	numeric(19,0)
dat_consolidacionentidades	idconsolidacion	numeric(19,0)	dim_configuracion_agregacion_hijos	idconsolidacion	numeric(19,0)
dat_consolidacionentidades	idestructurae	numeric(19,0)	dim_configuracion_agregacion_hijos	idestructuraehijo	numeric(19,0)

## Anexo # 2: Vista de las trazas de las pruebas de carga mediante el Jmeter.

**Configuración de la Conexión JDBC**

Nombre: Data Mart Prueba

Comentarios:

Nombre Variable Enlazado al Pool:

Nombre de Variable: postgres

**Configuración del Pool de Conexiones**

Número Máximo de Conexiones: 10

Timeout del Pool: 10000

Intervalo de Limpieza por Inactividad (ms): 60000

Auto Commit: True

**Validación de Conexión por Pool**

Keep-Alive: True

Edad máxima de las Conexiones (ms): 5000

Query de Validación: Select 1

**Configuración de la Conexión a Base de Datos**

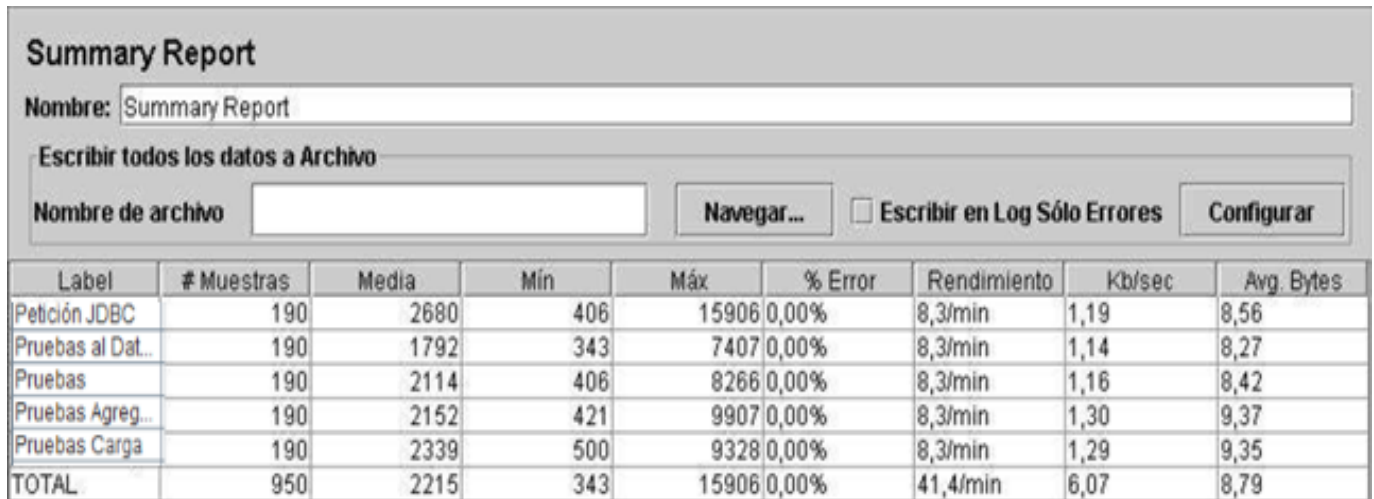
URL de la Base de Datos: jdbc:postgresql://localhost:otraa

Clase del Driver JDBC: org.postgresql.Driver

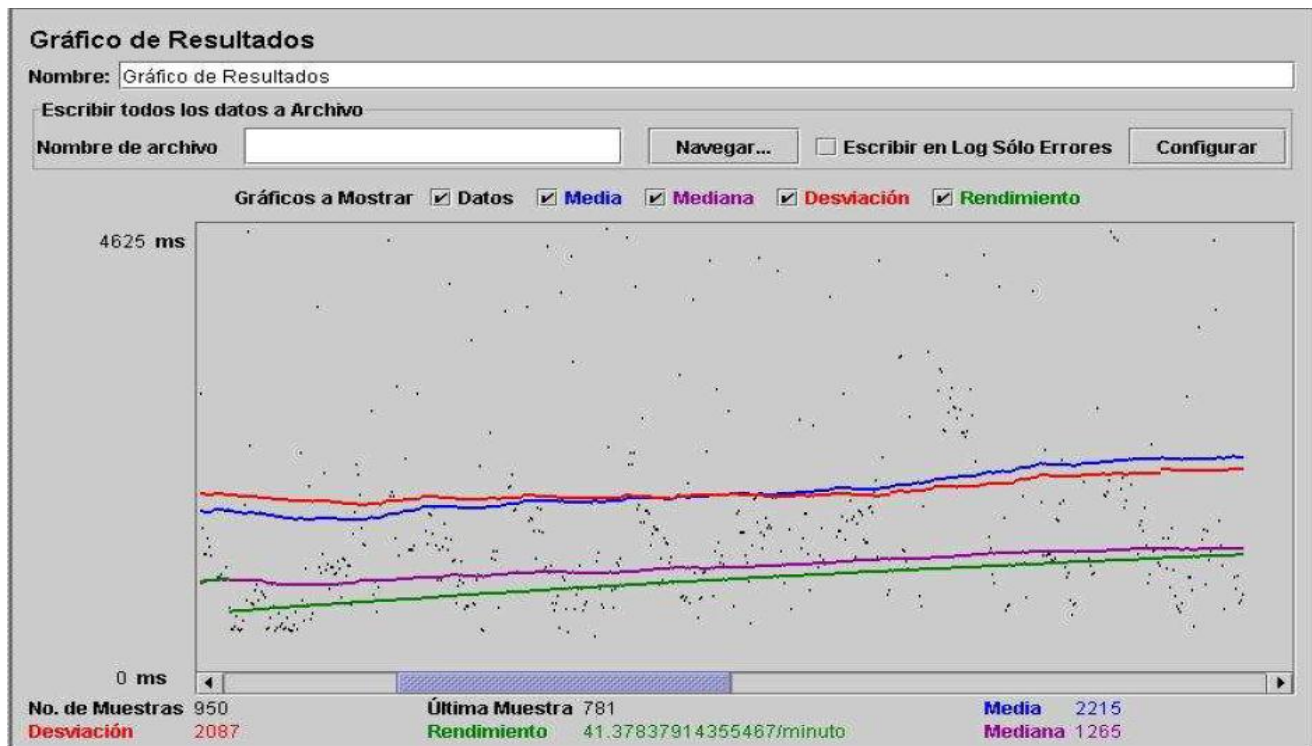
Nombre de Usuario: postgres

Password: postgres

### Anexo # 3: Vista del informe agregado mediante el Jmeter



### Anexo # 4: Vista del grafico de resultados mediante el Jmeter



### Anexo # 5: Volúmenes de dato de las tablas a utilizar.

- dim tiempo

id dim tiempo [PK] numeric(1)	quinquenio character varying	anno character varying	semestre character varying	trimestre character varying	mes character varying	nummes numeric(19,0)
99000000977	2006-2010	2010	Semestre2	Trimestre 1	Enero	1
99000000978	2006-2010	2010	Semestre2	Trimestre 1	Febrero	2
99000000979	2006-2010	2010	Semestre2	Trimestre 1	Marzo	3
99000000980	2006-2010	2010	Semestre2	Trimestre 2	Abril	4
99000000981	2006-2010	2010	Semestre2	Trimestre 2	Mayo	5
99000000982	2006-2010	2010	Semestre2	Trimestre 2	Junio	6
99000000983	2006-2010	2010	Semestre2	Trimestre 3	Julio	7
99000000984	2006-2010	2010	Semestre2	Trimestre 3	Agosto	8
99000000985	2006-2010	2010	Semestre2	Trimestre 3	Septiembre	9
99000000986	2006-2010	2010	Semestre2	Trimestre 4	Octubre	10
99000000987	2006-2010	2010	Semestre2	Trimestre 4	Noviembre	11
99000000988	2006-2010	2010	Semestre2	Trimestre 4	Diciembre	12
99000000989	2011-2015	2011	Semestre2	Trimestre 1	Enero	1
99000000990	2011-2015	2011	Semestre2	Trimestre 1	Febrero	2
99000000991	2011-2015	2011	Semestre2	Trimestre 1	Marzo	3
99000000992	2011-2015	2011	Semestre2	Trimestre 2	Abril	4
99000000993	2011-2015	2011	Semestre2	Trimestre 2	Mayo	5
99000000994	2011-2015	2011	Semestre2	Trimestre 2	Junio	6
99000000995	2011-2015	2011	Semestre2	Trimestre 3	Julio	7
99000000996	2011-2015	2011	Semestre2	Trimestre 3	Agosto	8
99000000997	2011-2015	2011	Semestre2	Trimestre 3	Septiembre	9

➤ dim cuenta

id cuenta [PK] numeric	grupo character varying(255)	contenido character varying(255)	concat character varying	descripcion character varying(255)	id cuenta padre numeric(19,0)
90000	Ingresos de perÃ-odos futuros	Contenido de prueba	012	Cuenta de prueba	0
90001	Activo Circulante	Efectivo en Caja	101	Efectivo en caja	0
90002	Activo Circulante	Efectivo en Caja	101/01	Efectivo en caja para pagos menores	90001
90003	Activo Circulante	Efectivo en Caja	101/02	Efectivo en caja por depositar en banco	90001
90004	Activo Circulante	Efectivo en Caja	101/03	Efectivo en caja para dietas en divisas	90001
90005	Activo Circulante	Efectivo en Caja	161	Anticipos a justificar	0
90006	Activo Circulante	Efectivo en Caja	161/01	Moneda nacional	90005
90007	Activo Circulante	Efectivo en Caja	161/02	Moneda CUC	90005
90008	Activo Circulante	Efectivo en Caja	109	Efectivo en banco	0
90009	Activo Circulante	Efectivo en Caja	109/01	Cheque en trÃ-nsito	90008
90010	Activo Circulante	Efectivo en Caja	109/02	Cta banc 01	90008
90011	Otros Activos	Faltantes de Bienes en InvestigaciÃ-n	333	Faltantes en investigaciÃ-n	0
90012	Otros Pasivos	Sobrantes en InvestigaciÃ-n	555	Sobrantes en investigaciÃ-n	0
90013	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	822	Gastos generales y de administraciÃ-n	0
90014	Grupo de cuentas nominales	CUENTAS NOMINALES ACREEDORAS	950	Otros ingresos	0
90015	Plazo de pago	AmortizaciÃ-n de Activos Fijos Tangibles	430	Reconocimiento anticipado	0
90016	Otros Activos	Faltantes de Bienes en InvestigaciÃ-n	334	Derecho de cobro	0
90017	Otros Activos	Faltantes de Bienes en InvestigaciÃ-n	349	Derecho fiscal	0
90018	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	822/01	Moneda nacional	90013
90019	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	822/02	Moneda CUC	90013
90020	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	801	Gasto prueba 1	0

90022	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	803	Gastos prueba 3	0
90023	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	804	Gasto prueba 4	0
90024	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	801/01	gasto MN	90020
90025	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	801/02	Gasto ME	90020
90026	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	802/01	Gasto CUC	90021
90027	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	802/02	Gasto CUP	90021
90028	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	805	Cuenta gastos costos 1	0
90029	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	807	Cuenta gastos costos 2	0
90030	Activo Circulante	Efectivo en Caja	183	Inventario	0
90031	Plazo de pago	Amortizaci3n de Activos Fijos Tangibles	401	Cuentas por pagar	0
90032	Grupo de cuentas nominales	CUENTAS NOMINALES ACREEDORAS	901	Ingreso por servicios	0
90033	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	801	Gasto de operaciones	0
90034	Otros Activos	Faltantes de Bienes en Investigaci3n	331	Cuenta por cobrar	0
90035	Ingresos de per3odos futuros	Contenido de prueba	013	Cuenta para productos	0
90036	Inventario	contenido de inventario	022	cuenta para apertura	0
90037	Grupo de cuentas nominales	CUENTAS NOMINALES ACREEDORAS	951	Ingresos	0
90038	Activo Circulante	Efectivo en Caja	101/04	Pagos menores	90001
90039	Activo Circulante	Efectivo en Caja	101/05	Por depositar	90001
90040	Activo Circulante	Efectivo en Caja	101/06	Divisa	90001
90041	Activo Circulante	Efectivo en Caja	101/67	Local Pagos menores	90001
90042	Activo Circulante	Efectivo en Caja	101/68	Local por de positar	90001
90043	Activo Circulante	Efectivo en Caja	101/69	Local divisa	90001
90044	Activo Circulante	Efectivo en Caja	101/70	Local dieta divisa	90001
90045	Activo Circulante	Efectivo en Caja	110	Cuenta Banco CUC Pruebas	0
90046	Activo Circulante	Efectivo en Caja	183	Inventario	0
90047	Plazo de pago	Amortizaci3n de Activos Fijos Tangibles	401	Cuenta por pagar	0
90048	Otros Activos	Faltantes de Bienes en Investigaci3n	335	Cuenta por cobrar	0
90049	Grupo de cuentas nominales	CUENTAS NOMINALES DEUDORASAA	801	Gastos por operaciones	0
90050	Grupo de cuentas nominales	CUENTAS NOMINALES ACREEDORAS	901	Ingreso por operaciones	0
90051	Otros Activos	Faltantes de Bienes en Investigaci3n	332	Transferencias activas	0
90052	Otros Pasivos	Sobrantes en Investigaci3n	556	Transferencias pasivas	0
90053	Grupo de cuentas nominales	CUENTAS NOMINALES ACREEDORAS	950	cuenta de efectivo	0
90054	Otros Activos	Faltantes de Bienes en Investigaci3n	335	Cuenta para nueva apertura	0
90055	Activo Circulante	Efectivo en Caja	120	Efectivo en caja para depositar	0
90056	Activo Circulante	Efectivo en Caja	121	Efectivo para pagos menores	0
90057	Activo Circulante	Efectivo en Caja	122	Efectivo en cuc para dietas	0
90059	Otros Pasivos	Sobrantes en Investigaci3n	570	Prueba	0
90061	Efectos y cuentas por cobrar me	Pasivos ficticios	976	srtfttt	0
90063	Ingresos de per3odos futuros	Contenido de prueba	980	susi	0
90064	Otros Activos	Faltantes de Bienes en Investigaci3n	332	Cuenta Bancaria	0

➤ dim\_cuenta\_closure

<b>idcuenta numeric(19,0)</b>	<b>idcuentapadr numeric(19,0)</b>	<b>distancia numeric(19,0)</b>
90063	0	1
90063	90063	0
90145	0	1
90145	90145	0
90206	0	2
90206	90200	1
90206	90206	0
90001	0	1
90001	90001	0
90158	0	1
90158	90158	0
90237	0	1
90237	90237	0
90127	0	1
90127	90127	0
90036	0	1
90036	90036	0
90267	0	2
90267	90267	0
90267	90266	1
90283	0	1
90283	90283	0
90098	0	2
90098	90098	0
90098	90096	1
90175	0	2
90175	90175	0
90175	90001	1
90047	0	1
90047	90047	0
90220	0	2
90220	90209	1
90220	90220	0
90021	0	1
90021	90021	0
90252	0	1
90252	90252	0
90144	0	1
90144	90144	0
90064	0	1
90064	90064	0

➤ dim\_estructura

<b>idestructura [PK] numeric</b>	<b>codigo character vai</b>	<b>denominacion character varying(255)</b>	<b>idestructurapadre numeric(19,0)</b>
900000000	123	UCI	0
900000001	01	MININT	0
900000003	1321	FACULTAD	900000000
900000006	4562	Almacen de Javier	900000000
900000019	567	esta	0
900000041	9889	klkl	0
900000044	333333	ETECSA	0
900000051	656899898	DivisiÃ³n Territorial Oeste	900000044
900000131	8559	Entidad de prueba	900000000
900000132	666	Carga Inicial COPA	0
900000134	0222121	Entidad ultima	900000000
900000136	4654	Estructuta lester	0
900000138	124	Johanny	900000000
900000141	435	d	900000000
900000142	345	dfg	900000000
900000144	1232	MININT 2	900000001
900000145	12359	Entidad Cierre	900000000
900000148	99	xxxxxxxxxxxxxxxxxxxxxx	0
900000155	998	EconomÃ­a	900000051
900000165	8	hhhh	900000138
900000166	136	MEP	0

➤ dim\_estructura\_closure

<b>idestructura numeric(19,0)</b>	<b>idestructurapadre numeric(19,0)</b>	<b>distancia numeric(19,0)</b>
900000166	0	1
900000166	900000166	0
900000131	0	2
900000131	900000131	0
900000131	900000000	1
900000167	0	2
900000167	900000166	1
900000167	900000167	0
900000132	0	1
900000132	900000132	0
900000000	0	1
900000000	900000000	0
900000041	0	1
900000041	900000041	0
900000145	0	2
900000145	900000000	1
900000145	900000145	0
900000148	0	1
900000148	900000148	0
900000019	0	1
900000019	900000019	0

➤ dim\_estructurae

<b>idestructurae [PK] numeric</b>	<b>codigoestruc character vai</b>	<b>descripcion character vai</b>	<b>idestructurae numeric(19,0)</b>
9000004	123	NUEVA ESTRU	0
9000011	123	DD	9000004
9000014	1	FACULTAD	0
9000016	456	ALMACEN DE	0
9000030	4562	PRUEBA	9000016
9000049	4562	CCCCCCCCCCC	9000030
9000064	123	Ñ&STO ES U	9000011
9000073	123	UCI	9000064
9000080	123	CENTRO PRUE	9000073
9000092	666	DPTO DE CON	0
9000093	333333	XXXXXXXXXXXX	0
9000094	0222121	ENTIDAD ULT	0
9000096	656899898	CENTRO RECT	0
9000097	8559	ENTIDAD DE	0
9000098	345	FACULTAD	9000073

➤ dim\_estructurae\_closure

<b>idestructurae numeric(19,0)</b>	<b>idestructuraepadre numeric(19,0)</b>	<b>distancia numeric(19,0)</b>
9000080	0	5
9000080	9000011	3
9000080	9000064	2
9000080	9000080	0
9000080	9000004	4
9000080	9000073	1
9000030	0	2
9000030	9000030	0
9000030	9000016	1
9000098	0	5
9000098	9000011	3
9000098	9000064	2
9000098	9000098	0
9000098	9000004	4
9000098	9000073	1
9000097	0	1
9000097	9000097	0
9000004	0	1
9000004	9000004	0
9000092	0	1
9000092	9000092	0

➤ dim\_configuracion\_agregacion

<b>idconsolidacion [PK] numeric(19,0)</b>	<b>nombre character vai</b>	<b>idestructuraepadre numeric(19,0)</b>	<b>idestructurapadre numeric(19,0)</b>
900000000000000	PRUEBA	9000000	900000000
900000000000001	PRUEB1	9000000	900000000
900000000000002	PRUEB2	9000000	900000000
900000000000003	PRUEB3	9000000	900000000
900000000000004	PRUEBA	9000000	900000000

➤ dim\_configuracion\_agregacion\_hijos

<b>idconsolidacion [PK] numeric(19,0)</b>	<b>idestructuraehijo [PK] numeric(19,0)</b>
9000000000000000	9000000
9000000000000000	9000001
9000000000000001	9000000
9000000000000001	9000001
9000000000000002	9000000
9000000000000002	9000001
9000000000000002	9000002
9000000000000003	9000000
9000000000000003	9000002
9000000000000003	9000006
9000000000000004	9000002
9000000000000004	9000006

➤ cub hecho contabilidad cuenta

<b>idhecho [PK] numeric</b>	<b>idestructura numeric(19,0)</b>	<b>idcuenta numeric(19,0)</b>	<b>idcentro numeric(19,0)</b>	<b>idelemento numeric(19,0)</b>	<b>iddimtiempo numeric(19,0)</b>	<b>creditomb numeric(19,2)</b>	<b>debitomb numeric(19,2)</b>	<b>idestructurae numeric(19,0)</b>	<b>creditoacumu numeric(19,2)</b>	<b>debitoacumu numeric(19,2)</b>
9900001336	900000000	90000	900000000006	90000000012	9900000977	160.00	54.00	9000004	67648.00	276.00
9900001337	900000000	90000	900000000006	90000000026	9900000977	110.00	110.00	9000004	67648.00	276.00
9900001338	900000000	90000	900000000028	90000000012	9900000977	700.00	100.00	9000004	67648.00	276.00
9900001339	900000000	90000	900000000028	90000000012	9900000977	66678.00	12.00	9000004	67648.00	276.00
9900001340	900000000	90002	900000000054	90000000018	9900000977	0.00	80.00	9000004	0.00	80.00
9900001341	900000000	90003			9900000978	400.00	0.00	9000004	1027.00	1027.00
9900001342	900000000	90004			9900000978	1000.00	1300.00	9000004	1540.00	1600.00
9900001343	900000000	90006			9900000978	190.00	190.00	9000004	642.00	330.00
9900001344	900000000	90006			9900000990	50.00	0.00	9000004	642.00	330.00
9900001345	900000000	90006			9900000980	150.00	0.00	9000004	642.00	330.00
9900001346	900000000	90006			9900000981	120.00	0.00	9000004	642.00	330.00
9900001347	900000000	90006			9900000982	120.00	140.00	9000004	642.00	330.00
9900001348	900000000	90007			9900000978	1100.00	0.00	9000004	2349.99	1952.67
9900001349	900000000	90007			9900000990	1249.99	416.67	9000004	2349.99	1952.67
9900001350	900000000	90007			9900000982	0.00	240.00	9000004	2349.99	1952.67
9900001351	900000000	90007			9900000984	0.00	1000.00	9000004	2349.99	1952.67
9900001352	900000000	90009			9900000977	35805.00	48878.00	9000004	60453.03	50518.00
9900001353	900000000	90009			9900000978	930.00	0.00	9000004	60453.03	50518.00
9900001354	900000000	90009			9900000990	526.67	0.00	9000004	60453.03	50518.00
9900001355	900000000	90009			9900000982	860.00	0.00	9000004	60453.03	50518.00
9900001356	900000000	90009			9900000984	22311.36	0.00	9000004	60453.03	50518.00



9900001357	900000000	90010			9900000977	0.00	10000.00	9000004	0.00	121819.00
9900001358	900000000	90010			9900000978	0.00	650.00	9000004	0.00	121819.00
9900001359	900000000	90010			9900000990	0.00	1150.00	9000004	0.00	121819.00
9900001360	900000000	90010			9900000984	0.00	19.00	9000004	0.00	121819.00
9900001361	900000000	90011	90000000006	90000000006	9900000977	2.00	55.00	9000004	2.00	55.00
9900001362	900000000	90015			9900000978	15.00	0.00	9000004	107.00	4000085.00
9900001363	900000000	90015			9900000992	0.00	50.00	9000004	107.00	4000085.00
9900001364	900000000	90015			9900000998	25.00	0.00	9000004	107.00	4000085.00
9900001365	900000000	90016			9900000977	10002.00	2.00	9000004	110864.66	207.00
9900001366	900000000	90016			9900000978	7.00	0.00	9000004	110864.66	207.00
9900001367	900000000	90016			9900000979	500.00	0.00	9000004	110864.66	207.00
9900001368	900000000	90016			9900000998	16.66	0.00	9000004	110864.66	207.00
9900001369	900000000	90017			9900000977	501.00	0.00	9000004	2253.00	1741.00
9900001370	900000000	90018	90000000002	90000000010	9900000977	49770.00	150.00	9000004	50740.00	1275.00
9900001371	900000000	90018	90000000002	90000000010	9900000978	70.00	150.00	9000004	50740.00	1275.00
9900001372	900000000	90018	90000000002	90000000010	9900000981	40.00	100.00	9000004	50740.00	1275.00
9900001373	900000000	90018	90000000002	90000000010	9900000982	90.00	200.00	9000004	50740.00	1275.00
9900001374	900000000	90018	90000000002	90000000012	9900000978	70.00	20.00	9000004	50740.00	1275.00
9900001375	900000000	90018	90000000002	90000000012	9900000980	500.00	120.00	9000004	50740.00	1275.00
9900001376	900000000	90018	90000000018	90000000006	9900000977	100.00	500.00	9000004	50740.00	1275.00



## **GLOSARIO DE TÉRMINOS**

**Disparadores:** Procedimiento que se ejecuta cuando se cumple una condición establecida al realizar una operación de inserción (INSERT), actualización (UPDATE) o borrado (DELETE).

**DML:** Data Manipulation Language (Lenguaje de Manipulación de Datos) es un lenguaje proporcionado por el sistema de gestión de bases de datos que permite a los usuarios del mismo llevar a cabo las tareas de consulta y manipulación de los datos (INSERT, UPDATE, DELETE).

### **Estados financieros**

Constituyen el producto final de la Contabilidad; ellos describen la situación económico-financiera de una empresa.

### **Estado o cuenta de resultados.**

Muestra los saldos de las cuentas de ingresos y gastos generados en un periodo, así como los resultados de forma escalonada. Permite evaluar la situación económica de la empresa para un periodo de tiempo determinado lo que hace que se considere un estado dinámico.

### **Estado de situación o Balance general**

Muestra los saldos de las diferentes cuentas de activo, pasivo y patrimonio o capital, en una fecha determinada. Podría decirse que este estado es una fotografía de la situación financiera de la empresa. Esta información es preparada periódicamente

**ETL:** Extract, Transform and Load (Extraer, transformar y cargar) es el proceso que permite mover datos desde múltiples fuentes, reformatearlos y limpiarlos, y cargarlos en otra base de datos, DM, DW, o en otro sistema operacional.

**Inteligencia de negocios o BI** (del inglés Business Intelligence): Conjunto de estrategias y herramientas enfocadas a la administración y creación de conocimiento mediante el análisis de datos existentes en una organización o empresa.

**Log:** Registro oficial de eventos durante un rango de tiempo en particular. Es usado para registrar datos sobre quién, qué, cuándo, dónde y por qué un evento ocurre para un dispositivo en particular o aplicación.

**Mapeo:** Transformación que consiste en relacionar columnas del origen y del destino con nombres distintos o iguales.

**Transformaciones:** Conjunto de operaciones que se realizan para transformar los datos a replicar, o los metadatos asociados a estos.

**UCID:** Unidad de Compatibilización, Integración y Desarrollo de Productos Informáticos para la Defensa.

