

Universidad Central “Marta Abreu” de las Villas.

Facultad Matemática Física y Computación

Licenciatura en Ciencia de la Computación



TRABAJO DE DIPLOMA

*Mercado de Datos para el proceso de inventario del Sistema de
Gestión Contable VERSAT Sarasola*

Autora: Lisbet Hernández Cárdenas

Tutores: Dra. Luisa M. González González

Ing. Boris Pino Núñez

Santa Clara

2012-2013

DEDICATORIA

A mi mamá, por su apoyo incondicional y su confianza en mí.

A mi familia, por su confianza y por estar siempre pendientes de mí.

AGRADECIMIENTOS

A mi mamá, por estar siempre a mi lado apoyándome y ayudándome en todo lo posible.

A mi familia, por estar siempre pendientes y preocupados por mi.

A mis tutores, Dra. Luisa González González e Ing. Boris Pino Núñez por sus valiosas recomendaciones, por su esfuerzo y dedicación.

A las profesoras, MSc. Lisandra Díaz de La Paz y Lic. Beatriz Méndez Hernández, por darme su apoyo y brindarme su ayuda en todo lo que estuviera a su alcance.

A todos los profesores que hicieron posible mi formación profesional.

A los compañeros del grupo de desarrollo de la División DATAZUCAR Villa Clara por sus continuas colaboraciones, por su apoyo y por dedicar parte de su tiempo en brindarme y mostrarme la información necesaria para la realización del trabajo.

A todos, muchas gracias.

RESUMEN

El Sistema de Gestión Contable VERSAT Sarasola, es un software utilizado por todas las entidades presupuestadas del país para efectuar el control económico. Cuenta con varios módulos clientes y una base de datos centralizada diseñada sobre el Sistema de Gestión de Bases de Datos SQL Server2000. Entre los subsistemas que conforman al sistema integrado VERSAT Sarasola, el subsistema de inventario es uno de los más importantes debido al gran volumen de datos que maneja y a su complejidad. Este subsistema genera un gran número de información histórica, la cual debe tener un tratamiento especial, debido a su utilidad en el proceso de extracción de conocimiento que permite analizar la organización, prever su evolución y tomar decisiones estratégicas para el futuro. El hecho de cómo utilizar la información histórica de los distintos centros para que la misma deje de ser pasiva y ayude a la toma de decisiones de las empresas a las que prestan servicio es la principal problemática que se plantea.

En este trabajo se crea un Mercado de Datos con 11 dimensiones conformadas, 26 tablas de hechos, entre ellas 8 transaccionales, 2 consolidadas y 16 periódicas.

Para el desarrollo de la solución se utilizaron herramientas de la Suite de Pentaho, además del Sistema Gestor de Base de Datos PostgreSQL. Finalmente se hizo una comprobación del Mercado de Datos utilizando datos reales, tomados de la empresa ULAEX, SA.

ABSTRACT

Accounting Management System VERSAT Sarasola, is a software used by all the budgeted entities of the country to perform the economic control. It has several modules clients and a centralized database designed on the Database Management System SQL Server2000. Among the subsystems that make up the integrated system VERSAT Sarasola, the inventory subsystem is one of the most important due to the great volume of data that manages and to its complexity. This subsystem generates a great number of historical information, which should have a special treatment because of their usefulness in the knowledge extraction process to analyze the organization, consider developing and making strategic decisions for the future. The fact of how to use the historical information of the different centers so that the same one leaves to be passive and help to the decision-making of the companies they serve is the main problem that arises.

In this thesis a data mart is created with 11 conformed dimensions, 26 fact tables, including eight transactional, two consolidated and 16 snapshot periodic.

The Pentaho Suite tools were used to the development of the solution, besides the Database Management System PostgreSQL. Finally a validation of the data mart was made using real data, taken of the company ULAEX, INC.

TABLA DE CONTENIDOS

INTRODUCCIÓN	1
Capítulo 1. CASO DE ESTUDIO: PROCESO DE INVENTARIO DEL VERSAT	
SARASOLA	4
1.1. Fuente de datos: El sistema VERSAT Sarasola.	4
1.2. Subsistema de Inventario	6
1.3. Subsistema de Contabilidad	11
1.4. Subsistema de Planificación	14
1.5. Subsistema de Finanzas.....	17
1.6. Subsistema de Activos Fijos.....	19
1.7. Subsistema de Facturación.....	20
1.8. Conclusiones del capítulo	23
Capítulo 2. CARACTERÍSTICAS DE UN MODELO DE INVENTARIOS	25
2.1. Tipos de tablas de hechos.....	25
2.1.1. Transaccional.....	25
2.1.2. Snapshot Periódica	26
2.1.3. Snapshot Acumulativa	26
2.2. Caso de estudio inventario (Kimball and Ross, 2002).....	28
2.2.1. Snapshot periódica para el inventario	29
2.2.2. Transacciones en el inventario	32
2.2.3. Snapshot acumulativa para el inventario	33
2.3. Dimensiones conformadas	35
2.4. Rol jugado por las dimensiones.....	36
2.5. Tablas de hechos consolidadas	36
2.6. Conclusiones del capítulo	37
Capítulo 3. DISEÑO E IMPLEMENTACIÓN DEL MERCADO DE DATOS PARA EL	
PROCESO DE INVENTARIO DEL VERSAT SARASOLA.....	38
3.1. Pasos para el diseño lógico de la solución.	38
3.1.1. Paso 1. Seleccionar el proceso de negocio	38
3.1.2. Paso 2. Determinar el grano del proceso de negocio	38
3.1.3. Paso 3. Seleccionar las dimensiones	39

3.1.4.	Paso 4. Identificar los hechos	39
3.2.	Subprocesos identificados en el proceso de inventario.....	39
3.2.1.	Subproceso de compra	39
3.2.2.	Subproceso de venta.....	40
3.2.3.	Subprocesos de entrada a centros de costo y salida de centros de costo	40
3.2.4.	Subprocesos de entrada, salida y transferencia	41
3.2.5.	Subproceso de ajuste	41
3.2.6.	Subproceso de devolución	42
3.3.	Diseño del Mercado de Datos para el proceso de inventario.....	42
3.3.1.	Dimensiones identificadas	43
3.3.2.	Tablas de hechos transaccionales identificadas.....	44
3.3.3.	Tablas de hechos periódicas identificadas	46
3.3.4.	Tablas de hechos consolidadas identificadas	49
3.4.	Implementación del Mercado de Datos para el proceso de inventario	51
3.4.1.	Proceso de Extracción, Transformación y Carga (ETL).....	53
3.4.2.	Implementación de los cubos de datos.	58
3.5.	Reportes representativos	60
3.6.	Conclusiones del capítulo	61
	CONCLUSIONES	62
	RECOMENDACIONES.....	63
	BIBLIOGRAFÍA.....	64
	ANEXO I.....	66

LISTA DE FIGURAS

Figura 1.1. Subsistemas que integran el VERSAT Sarasola.....	5
Figura 2.1. Subconjunto de una cadena de valores de un vendedor al menudeo.....	28
Figura 2.2. Esquema snapshot periódico del inventario de una tienda.....	30
Figura 2.3. Inventario snapshot periódico mejorado, para apoyar el análisis de GMROI.....	32
Figura 2.4. Modelo transaccional de inventario en almacén.....	33
Figura 2.5. Inventario snapshot acumulativo en almacén.....	34
Figura 3.1. Esquema estrella del subproceso de compras a nivel de transacción.....	50
Figura 3.2. Esquema estrella del subproceso de compras a nivel de periódico.....	51
Figura 3.3. Diseño físico del Mercado de Datos.....	52
Figura 3.4. Proceso ETL para el subproceso de compras a nivel transaccional.....	55
Figura 3.5. Selección de los productos que no tienen fecha de vencimiento asociada.....	56
Figura 3.6. Selección de los productos que tienen fecha de vencimiento asociada.....	56
Figura 3.7. Proceso ETL para el subproceso de compras a nivel periódico.....	57
Figura 3.8. Proceso de carga del Mercado de Datos.....	57
Figura 3.9. Cubo de datos de la compras transaccional.....	59
Figura 3.10. Cubo de datos de la compras periódicas.....	59

LISTA DE TABLAS

Tabla 1.1. Relación entre tipo de operación y conceptos.....	8
Tabla 1.2. Matriz de procesos/dimensiones (Bus Matrix).....	23
Tabla 2.1. Comparación entre los tres tipos de tablas de hechos.....	27

INTRODUCCIÓN

El surgimiento de la Informática y su aplicación a los procesos contables y financieros, así como a otras actividades asociadas con la gestión de las entidades económicas, no rebasa los 30 años en el país y en particular en el Ministerio del Azúcar (MINAZ) está alrededor de los 28 años.

Una de las esferas donde más desarrollo se ha tenido en materia de sistemas computacionales, fueron los destinados a la información estadística a diferentes niveles.

Alrededor del año 2000 surge y se desarrolla el VERSAT Sarasola, el cual a finales del 2005 fue presentado para su certificación, lo que finalmente se logró con fecha febrero del 2006, para constituirse en el primer software cubano en lograr esta categoría.

El VERSAT Sarasola es un software de Gestión Contable, que es utilizado por todas las entidades presupuestadas del país para efectuar el control económico desde los centros de gestión Contables y las Tesorerías creadas en todos y cada uno de los municipios del país. Dicho sistema cuenta con varios módulos clientes: Inventario, Finanzas, Activos Fijos, Planificación, Facturación, Contabilidad, Nóminas y Costos y Procesos y una base de datos centralizada diseñada sobre el Sistema de Gestión de Bases de Datos SQL Server2000. La explotación y el estudio de este sistema han dado paso a analizar con detenimiento el funcionamiento y desarrollo de algunos subsistemas, principalmente el de inventario. Este subsistema genera un gran número de información histórica, la cual debe tener un tratamiento especial, debido a su utilidad en el proceso de extracción de conocimiento que permite analizar la organización, prever su evolución y tomar decisiones estratégicas para el futuro, y sin embargo no es tratada por el sistema operacional.

De esta forma surge el siguiente **planteamiento del problema**: cómo utilizar la información histórica de los distintos centros para que la misma deje de ser pasiva y ayude a la toma de decisiones de las empresas a las que prestan servicio.

En este sentido se propone como **objetivo general**: construir un Mercado de Datos para facilitar la toma de decisiones del proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola usando un enfoque multidimensional y validarlo con al menos una empresa.

Para lograr este objetivo se plantean los siguientes **objetivos específicos**:

1. Describir el caso de estudio para el proceso de inventario de VERSAT Sarasola.

2. Diseñar el Mercado de Datos.
3. Modelar los procesos de Extracción, Transformación y Carga de los datos.
4. Implementar el Mercado de Datos con herramientas de la suite de Pentaho.
5. Validar el Mercado de Datos con datos reales de al menos una empresa.

Se plantean las siguientes **preguntas de investigación**:

1. ¿Cuáles son los aspectos en el diseño de la aplicación que tienen un mayor interés técnico?
2. ¿Cuáles son las posibles herramientas de software libre disponibles para la realización de un Mercado de Datos?
3. ¿Qué metodología usar para el diseño del Mercado de Datos?
4. ¿Qué decisiones de diseño es necesario tomar para las particularidades de la aplicación?
5. ¿Cuáles son las herramientas a utilizar para el proceso de negocio?

Después de realizar la revisión del marco teórico y en respuesta a las preguntas de investigación, se plantea la siguiente **hipótesis de investigación**:

- La creación de un Mercado de Datos ofrece posibilidades de procesamiento de información histórica y la generación de reportes útiles para ayudar a la toma de decisiones.

La tesis, ha sido organizada en tres capítulos. A continuación se describe esta organización:

- **Capítulo 1: Caso de estudio del proceso de inventario del VERSAT Sarasola.** En este capítulo se describe el sistema operacional que constituye la fuente de datos para el Mercado de Datos que se modela en este trabajo. Además se hace una breve caracterización de otros subsistemas que tienen relación con los inventarios para poder definir las dimensiones conformadas del Mercado de Datos correspondiente a este proceso.
- **Capítulo 2: Características de un modelo de inventarios.** En este capítulo se realiza una descripción de un modelo genérico de inventario como un referente clásico para el Mercado de Datos que se muestra en la tesis, así como algunos conceptos que o bien son usados para resolver el caso de estudio o fueron necesarios para tomar algunas decisiones de diseño.

- **Capítulo 3: Diseño e implementación del Mercado de Datos para el proceso de inventario del VERSAT Sarasola.** En este capítulo se abordan aspectos concernientes a la descripción de la solución, específicamente, a la descripción de las fuentes a integrar, definición de las áreas de análisis, la arquitectura, el diseño, la interacción entre los componentes y la implementación del Mercado de Datos, el cubo de datos y el modelo multidimensional propuesto.

Capítulo 1. CASO DE ESTUDIO: PROCESO DE INVENTARIO DEL VERSAT SARASOLA

En este capítulo se describe el sistema operacional que constituye la fuente de datos para el Mercado de Datos que se modela en este trabajo. El sistema operacional denominado VERSAT Sarasola se presenta como un caso de estudio para el manejo de inventarios de todas las categorías en un almacén, así como también para el control de productos que son asignados a trabajadores de una empresa para su uso. En este capítulo también se hace una breve caracterización de otros subsistemas que tienen relación con los inventarios para poder definir las dimensiones conformadas del Mercado de Datos correspondiente a este proceso.

1.1. Fuente de datos: El sistema VERSAT Sarasola.

En julio de 1999, el Ministerio del Azúcar decidió elaborar un sistema económico integrado por especialistas del propio organismo. En esa fecha es que nace o se fortalece el grupo de trabajo que ya venía realizando las interrelaciones y completamiento de los sistemas que conforman el denominado Paquete Económico de Tránsito, pero en este caso se centra toda la actividad en Santa Clara, que es donde radicaba la dirección del mismo y una buena parte de los productores.

Alrededor del año 2000 surge y se desarrolla el sistema VERSAT Sarasola, certificado en febrero del 2006, para constituirse en el primer software cubano en lograr esta categoría.

El VERSAT Sarasola fue categorizado como el software de gestión más completo existente en el país hasta este momento, por tanto también se está utilizando como soporte tecnológico para los primeros pasos que se dan con vistas a la elaboración del primer ERP (Enterprise Resource Planning) cubano.

Un aspecto relevante lo constituye el hecho de que este software, cuyo objetivo básico original fue que se elaborara sólo para la agroindustria azucarera cubana, se ha generalizado a otros sectores de la economía, debido en gran parte al reconocimiento que ha tenido por parte de los organismos rectores de esta actividad en el país. En la realización del mismo, se ha tenido en cuenta, la interrelación con las aplicaciones instaladas en los niveles superiores, con vistas a garantizar que el flujo automatizado de la información se mantenga, pues este

producto está destinado al uso empresarial y este debe suministrar los elementos solicitados por ellos.

El VERSAT Sarasola es un sistema que cuenta con varios módulos clientes (Inventario, Finanzas, Activos Fijos, Planificación, Facturación, Contabilidad, Nóminas y Costos y Procesos) como se muestra en la Figura 1.1 y una base de datos centralizada diseñada sobre el Sistema de Gestión de Bases de Datos SQL Server2000 (Autores, 2005, DATAZUCAR, 2002, Autores).

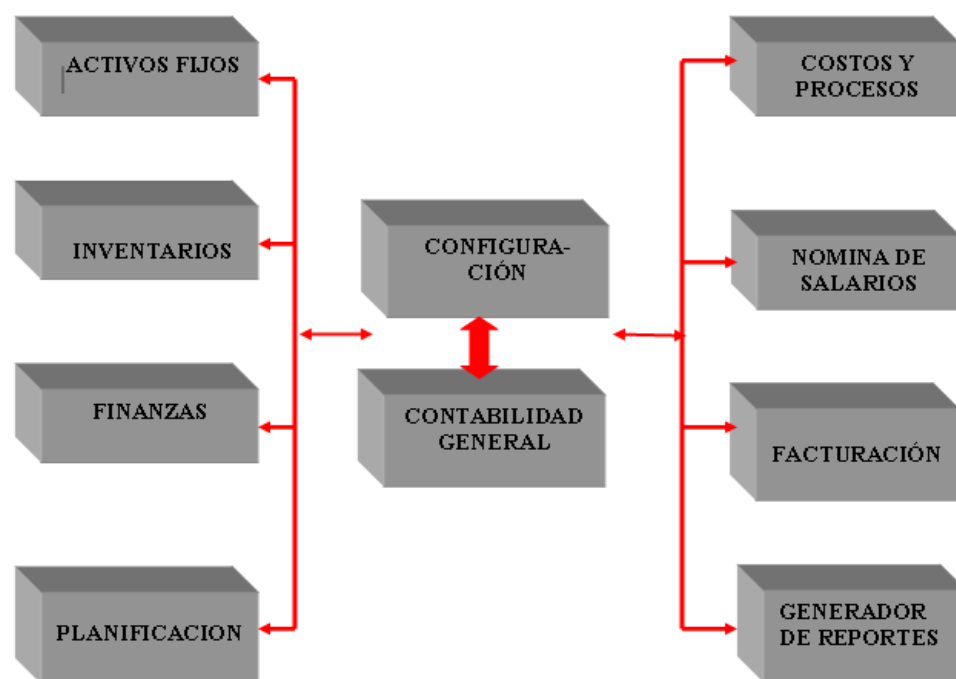


Figura 1.1 Subsistemas que integran el VERSAT Sarasola

Entre los subsistemas que conforman al sistema integrado VERSAT Sarasola, el subsistema de inventario es uno de los más importantes, debido al gran volumen de datos que maneja y a su complejidad. De esta forma, el estudio y explotación del mismo han marcado un gran interés en su funcionamiento y desarrollo, pues este subsistema genera un gran número de información histórica, la cual debe tener un tratamiento especial, debido a su utilidad en el proceso de extracción de conocimiento que permite analizar la organización, prever su evolución y tomar decisiones estratégicas para el futuro.

1.2. Subsistema de Inventario

El subsistema de control de los inventarios, como parte del sistema económico, tiene como finalidad la de procesar toda la documentación primaria inherente a la gestión de almacenes focalizando su acción en el control de los recursos materiales a través de lo establecido en el Manual de Normas de Control Interno, Sección II, Capítulo 2: Subsistema de Inventarios. Además tiene incorporado el control de los útiles y herramientas en uso por cada uno de sus custodios o paños.

El inventario se lleva a nivel de unidad contable, las cuales en el VERSAT son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa y constituye el elemento organizativo de mayor nivel e implica a los demás subsistemas. Las unidades contables, al igual que otros elementos como las entidades, los productos, conceptos, centros de costo y custodios, tienen un código, un nombre y un estado que indica si están activos o no. Además, el sistema mantiene un identificador artificial para cada elemento que compone dicho subsistema con el propósito de facilitar los acoples.

Dentro de cada unidad contable el inventario se realiza de manera excluyente a nivel de almacén o de custodio. El inventario se lleva a cabo en un almacén cuando se realizan movimientos de productos en el mismo, por todos los conceptos que se decida utilizar. Además de los atributos comunes descritos anteriormente para algunos elementos, también interesa conocer dónde está localizado el almacén, es decir su dirección, así como el jefe del mismo.

El subsistema de inventarios, tiene también como elemento organizativo al mismo nivel que los almacenes, la gestión de custodios. Los custodios son trabajadores que poseen la responsabilidad material de un producto que generalmente son productos que usan en su labor, por ejemplo herramientas.

Existe una estrecha e importante relación entre los almacenes y las unidades contables, pues cada almacén esta asociado solamente a una unidad contable, sin embargo, una unidad contable puede tener asociado muchos almacenes.

Como se mencionó anteriormente, en los almacenes se realizan movimientos de productos, por lo que se necesita conocer la existencia real de los productos en cada momento después de realizar cualquier movimiento. Los productos en el almacén tienen dos formas de tratamiento: unitario y por lotes. Los productos unitarios representan cada producto físico, a nivel atómico, ya sea a nivel de almacén, para la compra o la venta.

Los lotes no representan una entidad física a efectos del inventario, sino que es una forma particular de tratar a aquellos productos que tienen una misma entrada al almacén con una misma fecha de vencimiento. El manejo de los movimientos en el almacén por lotes, para aquellos productos que lo requieran, facilita conocer la fecha de vencimiento de los productos y llevar diferentes controles sobre su caducidad.

Los productos tienen una unidad de medida, acorde al Sistema Internacional de unidades (SI)(Cubana, 1988), (Kg., lts, sacos, unidad, latas, etc.) que permiten cuantificar el volumen de la operación que se realiza. Las unidades de medidas a su vez tienen su tratamiento especial, en que además de los atributos comunes a todo elemento del subsistema, tienen también un nombre abreviado, por ejemplo Kg.

En el proceso de inventario, la categoría de los productos define el destino que los mismos pueden tener en el almacén: insumo, mercancía para la venta, producción terminada, consignación. Cada categoría tiene un identificador y un nombre.

En el proceso de inventario cuando se realizan operaciones de compras y ventas, las empresas, unidades básicas o empresariales o cualquier otro centro, denominadas de forma genérica como entidades; pueden asumir el rol de proveedor o cliente según las necesidades específicas de cada una en el momento de la negociación. Una entidad puede participar en una compra, una venta o prestar determinado servicio. Los servicios se prestan solamente para una operación de compra y puede ser manipulación, transportación de la mercancía u otro tipo; estos servicios tienen asociado un código y una descripción.

Cada movimiento de productos en almacenes o custodios genera un tipo de operación específico. En este proceso de inventario, existen nueve tipos de operaciones, las cuales, a su

vez, pueden representar subprocesos dentro del propio proceso de inventario. Estas operaciones son:

- compras
- ventas
- devolución
- transferencia
- ajuste
- entrada
- salida
- entrada a centro de costo
- salida de centro de costo

Con el objetivo de garantizar el control de las operaciones en almacenes y custodios, se hace necesario asociar a cada una de estas operaciones diferentes conceptos; estos conceptos son definidos por el usuario. La Tabla 1.1 muestra ejemplos de cada tipo de operación y sus posibles conceptos asociados.

Tipo de operación	Conceptos asociados
Compras	Compra de insumo Compra a entidades
Ventas	Ventas a entidades Otras ventas categ. insumo
Devolución	Devolución de clientes Devolución de compra
Transferencia	Salida por transferencia Entrada por tranf.comedor
Ajuste	Ajuste por aumento Cancelando venta de insumo
Entrada	Sobranje de insumos Entradas a Consignación, Garantía y Reserva
Salida	Salida de útiles para uso Pérdidas y faltantes de insumos
Entrada a centro de costo	Devolución al comedor Entrada por devolución
Salida de centro de costo	Salida de Consignación al Costo Salidas de insumo

Tabla 1.1 Relación entre tipo de operación y conceptos asociados

Es importante tener en cuenta que las operaciones entrada a centro de costo y salida de centro de costo, indican entradas y salidas de productos para gastos, respectivamente; es decir estos productos entran y salen en determinados centros de costo.

Los centros de costo son una división interna, departamental de una entidad (empresa), pudiendo ser de servicios. En una entidad podrá haber tantos centros de costo ya sean productivos como de servicios se necesiten o requieran.

El proceso de inventario a nivel de gestión de custodio es similar al de los almacenes, por tanto, se debe tener en cuenta que los custodios pertenecen a una unidad contable y que a los mismos se le asignan determinados productos, con su unidad de medida; y además, dichos productos son adquiridos por diferentes conceptos. Otro elemento que interesa en el proceso de inventario a nivel de gestión de custodio, son los centros de costo, pues también existen productos que generan gastos y por tanto, entran y salen en determinados centros de costo. En el caso de los custodios también se deben tener en cuenta los tipos de operaciones existentes, destacándose que en este caso en particular para los custodios sólo son consideradas siete posibles operaciones:

- devolución
- transferencia
- ajuste
- entrada
- salida
- entrada a centro de costo
- salida de centro de costo

Como se mencionó anteriormente, el proceso de inventario a nivel de gestión de custodio es similar al de los almacenes, con la diferencia que no existe la categoría, es decir, no interesa saber si los productos son para insumo, mercancía para la venta, producción terminada o consignación. Además, como no se realizan operaciones de compras y ventas, tampoco existe una relación con las entidades y los servicios, ya que estas solo están involucradas con operaciones de compras y ventas.

La existencia de un producto, en general, expresa la cantidad del producto y su importe o valor, que pueden aumentar o disminuir según el tipo de operación.

Los movimientos de productos en el proceso de inventario, tanto para almacenes como para custodios, usualmente se realizan a diario, pero el control de estos movimientos se realiza en unidades de tiempo más extensas. Luego interesa conocer el resumen de los movimientos realizados en los siguientes intervalos de tiempo:

- quincena
- período
- trimestre
- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamadas períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario, por ejemplo:

- Definir ejercicio
 - o Inicio ejercicio: 5/03/2010
 - o Fin ejercicio: 5/03/2011
- Definir períodos dentro del ejercicio
 - o Inicio período: 5/03/2010 Fin período: 5/04/2010
 - o Inicio período: 6/04/2010 Fin período: 5/05/2010
 - o Inicio período: 6/05/2010 Fin período: 5/06/2010
 - o Inicio período: 6/06/2010 Fin período: 5/07/2010
 - o Inicio período: 6/07/2010 Fin período: 5/08/2010
 - o Inicio período: 6/08/2010 Fin período: 5/09/2010
 - o Inicio período: 6/09/2010 Fin período: 5/10/2010
 - o Inicio período: 6/10/2010 Fin período: 5/11/2010

- o Inicio período: 6/11/2010 Fin período: 5/12/2010
- o Inicio período: 6/12/2010 Fin período: 5/01/2011
- o Inicio período: 6/01/2011 Fin período: 5/02/2011
- o Inicio período: 6/02/2011 Fin período: 5/03/2011

Notar que en un almacén de una unidad contable, se realizan los movimientos de determinados productos con su unidad de medida correspondiente por todos los conceptos que se decida utilizar, desde o para una entidad respectivamente (según el concepto), en una unidad de tiempo determinada y para un fin determinado (categoría).

De igual forma un custodio de una unidad contable, tiene en su poder determinados productos con su unidad de medida, los cuales son adquiridos a través de un determinado concepto y en una unidad de tiempo específica.

Luego, interesa conocer, la cantidad e importe de productos que se mueven tanto en los almacenes como en los custodios, asociado a cada tipo de operación o subproceso. En el caso de aquellos productos que se mueven en el almacén por lotes mediante las operaciones de compras, ventas, transferencias, entradas o salidas, no sólo interesan cuántos productos se movieron, sino también en cuántos lotes; y a su vez, la cantidad de productos vencidos y no vencidos. Además, es necesario saber la cantidad de dinero invertido (importe) en servicios prestados cuando se realiza una compra en un almacén. De forma general, también se desea conocer la existencia real de un determinado producto con su unidad de medida de un almacén en una unidad contable, para una determinada categoría y en una unidad de tiempo dada, teniendo en cuenta la cantidad de productos vencidos y no vencidos; así como la existencia real de un determinado producto con su unidad de medida que posee un custodio en una unidad contable, para una unidad de tiempo determinada.

1.3. Subsistema de Contabilidad

En el sistema económico integrado VERSAT Sarasola, la concepción del registro de los hechos económicos está en el procesamiento de los documentos primarios que le dan origen, para ello se han elaborado un grupo de subsistemas, donde se procesan estos documentos.

En cada subsistema se procesan, registran y contabilizan todos los hechos económicos afines a las características y contenido para lo cual fue concebido, transfiriéndose sus resultados a la contabilidad general para asentarlos en los registros contables y con ello obtener los estados financieros y otras informaciones necesarias para la administración empresarial.

Una de las características fundamentales de este subsistema en su funcionamiento, es la de posibilitar la obtención de resultados en varios intervalos de tiempo, lo que le permite al usuario tener una información permanente y actualizada en el tiempo. De igual forma que se describió en el subsistema de inventario, los intervalos de tiempo que interesan son los siguientes:

- quincena
- período
- trimestre
- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamadas períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario.

Este subsistema es el rector de todo el sistema, ya que el contenido del resto de los subsistemas, está subordinado a las necesidades y requerimientos del registro contable, que se resumen en la contabilidad general, pues toda la información que se ofrece, está implícita íntegramente en las diferentes cuentas contables. Las cuentas contables son el instrumento que permite identificar, clasificar y registrar una determinada operación en dependencia de su naturaleza. Cada cuenta tiene asociada una clave, una descripción y un identificador. Además es necesario saber si la cuenta está activa o no, así como la naturaleza de la misma, pues indica si es una cuenta deudora o acreedora.

Por la interrelación que existe entre la contabilidad general y la de costos referida al registro de los gastos por centros de costo asociados a las cuentas que controla el subsistema de contabilidad general y a la necesidad de tener acceso desde éste a opciones que existen en el de costos y procesos, para agilizar el trabajo y obtener informaciones sin tener que acceder al mismo, se incluye la actividad costos y procesos dentro de este subsistema.

Al existir un clasificador de centros de costo independiente al clasificador de cuentas de la contabilidad general, es necesario identificar de la relación de cuentas, cuáles de ellas son de gastos y asociarlas con los centros de costo correspondientes para poder registrar los hechos económicos que se relacionan con gastos.

Los centros de costo son una división interna, departamental de una entidad, pudiendo ser de servicios. En una entidad podrá haber tantos centros de costo productivos como de servicios se necesiten o requieran. Los centros de costo, así como el resto de los elementos que componen dicho subsistema, dígame, las unidades contables, los subelementos de gasto, los elementos de gasto y las partidas, tienen un código, un nombre y un estado que indica si están activos o no. Además, al igual que en el subsistema de inventario, el sistema mantiene un identificador artificial para cada elemento que compone dicho subsistema con el propósito de facilitar los acoples.

Al seleccionar las cuentas de gastos, es necesario definir en cuales se van a registrar los gastos por subelementos, ya que debido a las características del cálculo y registro de los costos de producción, se utilizan subcuentas en las que no se registran gastos por subelementos.

Un aspecto importante e imprescindible en la contabilidad de costos, es que cuando se realiza el análisis de una cuenta por subelemento, también se le asocia a dicha cuenta el registro de los gastos por elementos de gasto y la conformación del costo por partidas.

En este subsistema el análisis contable se realiza a nivel de unidad contable, las cuales en el VERSAT son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa y constituye el elemento organizativo de mayor nivel e implica a los demás subsistemas.

Por lo tanto, como se mencionó anteriormente, todos los hechos económicos de cada uno de los subsistemas se procesan, registran y contabilizan, de acuerdo a las características y contenido para lo cual fueron concebidos. La resultante de este proceso termina transfiriéndose en forma de comprobantes de operaciones al subsistema de contabilidad general. Los comprobantes de operaciones, son documentos que registran los débitos y créditos que origina cada operación en la cuenta contable y que se expresa en la cuenta a través de un saldo.

Luego en el proceso de contabilidad general interesa conocer, el débito y el crédito que origina cada operación en la cuenta, así como el saldo de la misma; teniendo en cuenta que el débito es la suma de todas las cantidades situadas en el debe de una cuenta, mientras que el crédito es la suma de todas las cantidades del haber. De igual forma, podemos definir el saldo de una cuenta como la obtenida al comparar su débito con su crédito, es decir, la suma entre el débito y el crédito. Por lo tanto, estas medidas se analizan para una o más unidades contables, en un intervalo de tiempo dado, donde una cuenta determinada se puede analizar de forma general, o por centros de costo. Teniendo en cuenta que en este último caso también pueden interesar los subelementos de gasto y a su vez, los elementos de gasto y las partidas asociadas.

1.4. Subsistema de Planificación

El proceso de planificación se ha concebido con una orientación a la esfera presupuestada, sin dejar de tener posibilidades de generalización en el sector empresarial, predominando, de esta forma, la terminología utilizada en el sector presupuestado. Además de la planificación tradicional que está establecida por el país, en este subsistema se han implementado los mecanismos particulares que se aplican por el Ministerio de Finanzas y Precios para la elaboración del presupuesto del estado.

En el sistema económico integrado VERSAT Sarasola, la concepción de la planificación está relacionada con los principios contables – financieros en que se fundamenta el sistema, por tanto su procesamiento siempre tendrá un carácter totalmente contable. De esta forma podemos concluir que todo el procesamiento de planificación se traduce a una expresión contable, donde toda la información que se ofrece, está implícita íntegramente en las diferentes cuentas contables. Las cuentas contables son el instrumento que permite identificar, clasificar

y registrar una determinada operación en dependencia de su naturaleza. Cada cuenta tiene asociada una clave, una descripción y un identificador. Además es necesario saber si la cuenta está activa o no, así como la naturaleza de la misma, pues esta indica si es una cuenta deudora o acreedora.

Al existir un clasificador de centros de costo independiente al clasificador de cuentas de la contabilidad general, es necesario identificar de la relación de cuentas, cuáles de ellas son de gastos y asociarlas con los centros de costo correspondientes para poder registrar los hechos económicos que se relacionan con gastos.

Los centros de costo son una división interna, departamental de una entidad, pudiendo ser de servicios. En una entidad podrá haber tantos centros de costo productivos como de servicios se necesiten o requieran. Los centros de costo, así como el resto de los elementos que componen dicho subsistema, dígame, las unidades contables, los subelementos de gasto, los elementos de gasto, las partidas y los indicadores, tienen un código, un nombre y un estado que indica si están activos o no. Además, al igual que en los subsistemas descritos anteriormente, el sistema mantiene un identificador artificial para cada elemento que compone dicho subsistema con el propósito de facilitar los acoples.

Al seleccionar las cuentas de gastos, es necesario definir en cuales se van a registrar los gastos por subelementos, ya que debido a las características del cálculo y registro de los costos de producción, se utilizan subcuentas en las que no se registran gastos por subelementos.

Un aspecto importante e imprescindible en la contabilidad de costos, es que cuando se realiza el análisis de una cuenta por subelemento, también se le asocia a dicha cuenta el registro de los gastos por elementos de gasto y la conformación del costo por partidas.

Otro importante elemento es que este subsistema de planificación, incluye la presupuestación de los diferentes tipos de ingresos, gastos y costos de ventas, así como de los indicadores físicos o de valor que se requieran en cada entidad.

Un indicador es una información que no se expresa contablemente y que tiene una utilidad de análisis, por ejemplo los niveles de actividad (ventas, matrícula, promedio de trabajadores) y

los resultados productivos (producción bruta, valor agregado, producción mercantil, utilidad o pérdida), estos últimos se obtienen de procesar elementos contables junto a elementos estadísticos.

En este subsistema, al igual que en los otros que componen al VERSAT Sarasola, el análisis se realiza a nivel de unidad contable, las cuales en el VERSAT son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa y constituye el elemento organizativo de mayor nivel e implica a los demás subsistemas.

Una de las características fundamentales de este subsistema en su funcionamiento, es la de posibilitar la obtención de resultados en varios intervalos de tiempo, lo que le permite al usuario tener una información permanente y actualizada en el tiempo. De igual forma que se describió en el subsistema de inventario, los intervalos de tiempo que interesan son los siguientes:

- quincena
- período
- trimestre
- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamados períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario.

Por lo tanto, en el subsistema de planificación interesa conocer el plan de una cuenta determinada, la cual se puede analizar de forma general o por centros de costo; teniendo en cuenta que en este último caso también pueden interesar los subelementos de gasto y a su vez, los elementos de gasto y las partidas asociadas. Además, se desea conocer el total planificado de un indicador solicitado, así como el real del indicador, el acumulado real y el plan

acumulado. Todas estas medidas se analizan para una o más unidades contables, y en una unidad de tiempo determinada.

1.5. Subsistema de Finanzas

La actividad de finanzas dentro del VERSAT Sarasola es sin dudas la más abarcadora y donde la configuración de los diferentes conceptos, hacen que la misma sea muy transparente a los usuarios en cuanto al registro contable de sus operaciones y la actualización en tiempo real de todas los subsistemas que con la misma se interrelacionan. Controla las operaciones que se realizan sobre las diferentes cuentas bancarias definidas, permitiendo:

- Controlar los talonarios de los diferentes instrumentos de pagos por cada cuenta bancaria, incluyendo además las letras de cambio
- Procesar los estados de cuentas a partir de la información que se recibe en el mismo y las operaciones registradas con anterioridad en la entidad. Realizar conciliaciones de cada una de las cuentas bancarias
- Procesar los cobros y pagos automáticos recibidos.
- Emitir reportes referidos a documentos en tránsito, registro de cheques emitidos, registro de disponibilidad, etc.

Finanzas constituye uno de los módulos básicos del paquete económico-financiero VERSAT. El trabajo de este módulo se basa fundamentalmente en el procesamiento de los documentos primarios, teniendo en cuenta los diferentes conceptos de cobros y pagos definidos por el usuario, de forma tal que permiten modelar la actividad económico-financiera de su entidad. Una entidad es una empresa, una unidad básica o empresarial o cualquier otro centro, en la cual se desarrollan y controlan las operaciones que se realizan sobre las diferentes cuentas bancarias definidas y asociadas a cada entidad.

En este subsistema, al igual que en los otros que componen al VERSAT Sarasola, el análisis se realiza a nivel de unidad contable, las cuales en el VERSAT son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa y constituye el elemento organizativo de mayor nivel e implica a los demás subsistemas.

Las entidades y las unidades contables tienen asociado un código, un nombre y un estado que indica si están activos o no. Además, el sistema mantiene un identificador artificial para cada elemento que compone dicho subsistema con el propósito de facilitar los acoples.

Para garantizar una mejor distribución de las funciones financieras y una mejor aplicación de los principios del control interno de la contabilidad, se decidió agrupar dichas funciones en cinco actividades bien definidas: Banco, Caja, Cobros y Pagos, Créditos Bancarios y Otras operaciones financieras.

Este subsistema se interrelaciona además con otros del VERSAT. Para el de Contabilidad tributa sus operaciones contables en forma de comprobantes y del de Inventarios, Activos Fijos, Facturación y Nóminas recibe las obligaciones generadas por las operaciones efectuadas en ellos.

Además, una de las características fundamentales de este subsistema en su funcionamiento, es la de posibilitar la obtención de resultados en varios intervalos de tiempo, lo que le permite al usuario tener una información permanente y actualizada en el tiempo. De igual forma que se describió en los subsistemas anteriores, los intervalos de tiempo que interesan son los siguientes:

- quincena
- período
- trimestre
- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamados períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario.

1.6. Subsistema de Activos Fijos

El subsistema de activos fijos del VERSAT Sarasola, permite el control de todos los activos fijos, tanto los tangibles como los intangibles, estén activos, en almacén u otros estados. Además de su valor y reposición posee una considerable cantidad de atributos asociados a cada medio, que garantizan su depreciación sistemática, realización de inventarios físicos y otra apreciable cantidad de opciones o reportes.

Los activos fijos se definen como los bienes que una empresa utiliza de manera continua en el curso normal de sus operaciones; representan al conjunto de servicios que se recibirán en el futuro a lo largo de la vida útil de un bien adquirido.

De acuerdo a las características o funcionalidad similar, los activos pertenecen a distintos grupos, los cuales son imprescindibles y permiten un mejor control y manipulación de los mismos. Los grupos tienen asociado un identificador y un código. También es importante conocer la descripción y si el mismo amortiza o no.

En el módulo de Activos Fijos, como en los demás subsistemas que integran el paquete VERSAT Sarasola, se trabaja por unidades contables, las cuales pueden ser una o varias. En el VERSAT, las unidades contables son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa y constituye el elemento organizativo de mayor nivel e implica a los demás subsistemas. Al igual que los grupos, las unidades contables tienen un código, un nombre y un estado que indica si la misma esta activa o no; además de un identificador artificial generado por el sistema con el propósito de facilitar los acoples.

Una de las características fundamentales de este subsistema en su funcionamiento, es la de posibilitar la obtención de resultados en varios intervalos de tiempo, lo que le permite al usuario tener una información permanente y actualizada en el tiempo. De igual forma que se describió en los subsistemas anteriores, los intervalos de tiempo que interesan son los siguientes:

- quincena
- período
- trimestre

- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamados períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario.

En este subsistema es importante conocer las compras acumuladas hasta un intervalo de tiempo determinado, así como el total de compras.

1.7. Subsistema de Facturación

El VERSAT Sarasola está integrado por varios subsistemas y uno de ellos es el de Facturación. El mismo permite realizar facturas ya sea por: servicios, productos de insumo, mercancías para venta, producción terminada, activos fijos. Por tanto este subsistema puede operar independiente o vinculado a los subsistemas de Inventario, Activos Fijos y el de Finanzas, según la configuración que se establezca en la entidad.

El hecho de que una entidad ponga en explotación las bondades que le brinda el subsistema en cuestión, depende de sus propias características y de la necesidad que esta tenga de agilizar su información, sin pasar por alto los requerimientos técnicos para su instalación.

El vínculo entre Facturación y Contabilidad es mínimo, ya que son los demás subsistemas los encargados de contabilizar las operaciones que se generen desde Facturación. El único comprobante que podría generarse estaría asociado al ingreso o gasto/costo/faltante, ocasionado por la venta de insumo por debajo de su costo teniendo como contrapartida la cuenta transitoria asociada al movimiento de inventario.

Para declarar los tipos de facturas, en este subsistema se han tenido en consideración las orientaciones contenidas en la Instrucción 15-2006 del MFP sobre Factura Comercial y se le

incorporaron los elementos necesarios para el uso de multimonedas, este aspecto ha sido esencial por la demanda de los usuarios, en particular de las empresas mixtas en las cuales se factura en monedas extranjeras. Además respetando lo dispuesto en el documento mencionado, en la factura quedará identificado si es una venta: mayorista de bienes y servicios, interna donde medien operaciones de cobros y pagos, entregas de productos en consignación, ventas de AFT u operaciones por devolución.

Según los aspectos detallados anteriormente las facturas se han clasificado en cuatro grupos:

- Facturación por Servicios
- Facturación de Inventarios
- Facturación de Activos Fijos
- Facturación de Bienes y Servicios

Como se puede evidenciar el subsistema de facturación tiene una estrecha relación con el subsistema de inventario y es capaz de facturar todo tipo o categoría de venta que se realice por una entidad. Por tanto, en el subsistema de facturación intervienen elementos comunes al subsistema de inventario.

Como elemento organizativo de mayor nivel se encuentran las unidades contables, las cuales en el VERSAT son una unidad independiente desde el punto de vista contable, con su propio registro contable. La unidad contable generalmente pertenece a una empresa e implica a los demás subsistemas. Las unidades contables, al igual que otros elementos como las entidades, los productos y los conceptos, tienen un código, un nombre y un estado que indica si están activos o no. Además, el sistema mantiene un identificador artificial para cada elemento que compone dicho subsistema con el propósito de facilitar los acoples.

Dentro de cada unidad contable se encuentran los almacenes, en los cuales se realizan los movimientos de los productos por todos los conceptos que se decida utilizar. Además de los atributos comunes descritos anteriormente para algunos elementos, también interesa conocer dónde está localizado el almacén, es decir su dirección, así como el jefe del mismo.

Los productos tienen una unidad de medida, acorde al Sistema Internacional de unidades (SI) (Cubana, 1988), (Kg., lts, sacos, unidad, latas, etc.) que permiten cuantificar el volumen de la operación que se realiza. Las unidades de medidas a su vez tienen su tratamiento especial, en que además de los atributos comunes a todo elemento del subsistema, tienen también un nombre abreviado, por ejemplo Kg

Otros elementos que se deben tener en cuenta en el subsistema de facturación, dado que el mismo permite realizar facturas ya sea por servicios, productos de insumo, mercancías para venta y producción terminada, son los conceptos y la categoría. Los conceptos nombran e indican los diferentes movimientos que se realizan de un determinado producto y son fijados por el usuario. La categoría determina el destino de los productos y tiene asociado un identificador y un nombre.

En este subsistema es importante destacar la facturación de servicios, donde se debe tener en cuenta que los servicios son prestados por una determinada entidad y solo ocurre cuando estamos en presencia de una operación de compra; estos servicios tienen asociado un código y una descripción. Por otra parte cuando hablamos de entidad, nos referimos a una empresa, una unidad básica o empresarial o cualquier otro centro, la cual funge indistintamente de acuerdo a la operación que se realice.

Una de las características fundamentales de este subsistema en su funcionamiento, es la de posibilitar la obtención de resultados en varios intervalos de tiempo, lo que le permite al usuario tener una información permanente y actualizada en el tiempo. De igual forma que se describió en los subsistemas anteriores, los intervalos de tiempo que interesan son los siguientes:

- quincena
- período
- trimestre
- cuatrimestre
- semestre
- ejercicio contable

Se debe tener en cuenta que el ejercicio contable está compuesto por unidades de tiempo más pequeñas llamados períodos, por ejemplo, un ejercicio contable puede ser el año calendario y cada uno de los meses es un período. Además, tanto el ejercicio contable como el período, tienen un inicio y un fin, por tanto un ejercicio y los períodos del mismo pueden ser definidos por el usuario.

En general el subsistema de facturación tiene incorporada la prefactura como una modalidad muy utilizada en nuestro país y posee una cantidad apreciable de reportes asociados a la actividad comercial.

Después de haber descrito los subsistemas más importantes del VERSAT Sarasola pudimos apreciar la estrecha relación que existe entre los mismos, como muestra la Tabla 1.2, lo cual permite definir las dimensiones conformadas del Mercado de Datos correspondiente al proceso de inventarios.

Dimensiones Procesos	Fecha	Unidad Contable	Almacén	Custodio	Entidad	Categoría	Concepto	Producto	Unidad de medida	Centro de costo	Servicio
Inventario	X	X	X	X	X	X	X	X	X	X	X
Contabilidad	X	X								X	
Plan	X	X								X	
Finanzas	X	X			X						
Activos Fijos	X	X									
Facturación	X	X	X		X	X	X	X	X		X

Tabla 1.2 Matriz de procesos/dimensiones (Bus Matrix)

1.8. Conclusiones del capítulo

En este capítulo se describieron seis importantes subsistemas del software de Gestión Contable VERSAT Sarasola; entre ellos el subsistema de Inventario, pues este constituye nuestro caso

de estudio. Después de analizar dichos subsistemas, se observan las siguientes particularidades:

1. La relación existente entre el subsistema de inventario y los subsistemas de contabilidad, planificación, finanzas, activos fijos y facturación, dan paso a tratar los siguientes elementos del sistema como dimensiones conformadas:
 - a. Unidad contable
 - b. Fecha
 - c. Almacén
 - d. Entidad
 - e. Categoría
 - f. Concepto
 - g. Producto
 - h. Unidad de medida
 - i. Centros de costo
 - j. Servicios
2. El manejo de productos en el almacén de forma unitaria y por lotes, por su utilidad en la fecha de caducidad de los productos que lo requieren.
3. La estrecha relación existente entre los almacenes y las unidades contables, destacando que un almacén pertenece a una unidad contable, mientras que a una unidad contable pertenecen muchos almacenes
4. El análisis de inventario relacionado con los almacenes es independiente al análisis de inventario relacionado con los custodios, dando paso a dos procesos diferentes: el proceso de inventario en almacenes y el proceso de inventario en custodios.
5. La existencia de diferentes tipos de operaciones, tanto para el proceso de inventario de almacenes, como para el proceso de inventario de custodios, dando lugar a que las mismas sean consideradas subprocesos individuales dentro de los procesos de inventario de almacenes y custodios respectivamente.

Capítulo 2. CARACTERÍSTICAS DE UN MODELO DE INVENTARIOS

Se consideran ya conocidos los conceptos básicos sobre Almacenes de Datos.

El capítulo comienza con una descripción de los tipos de tablas de hechos, dado que son útiles para comprender el caso de estudio que se muestra a continuación sobre un modelo genérico de inventario como un referente clásico para el Mercado de Datos que se muestra en la tesis.

Finalmente se complementa con algunos conceptos que o bien son usados para resolver el caso de estudio o fueron necesarios para tomar algunas decisiones de diseño.

2.1. Tipos de tablas de hechos

Las tablas de hechos, siempre pueden reducirse a sólo tres tipos fundamentales: transaccional, periódica y acumulativa. Las tablas de hechos transaccionales rastrean cada transacción en un punto discreto en el tiempo, es decir cuando el evento de la transacción ocurrió. La tabla de hechos periódica captura los resultados acumulativos sobre intervalos de tiempos específicos. Generalmente, el diseño de la tabla de hechos acumulativa incluye algunos campos fecha para capturar las fechas cuando el artículo en cuestión pasa a través de cada uno de los procesos comerciales o hitos en la cadena de valor.

Se recomienda que en cada situación de diseño, se identifique el uso de cada uno de estos tres tipos de tablas. A continuación se describen estos tres tipos de tablas de hechos.

2.1.1. Transaccional

Los negocios funcionales fundamentalmente a nivel de transacción individual. El grano de la transacción representa una medida instantánea en un punto específico en el espacio y el tiempo; por tanto, las tablas de hechos transaccionales representan un evento que ocurrió en un momento instantáneo del tiempo (Kimball et al., 1998).

En general, las tablas de hechos transaccionales son un tipo de tabla de hecho en la cual, la granularidad es una fila para el más bajo nivel de detalle capturado por una transacción; es decir, existe una fila en la tabla, sólo si un evento de la transacción ocurrió. Así, una tabla de hecho transaccional puede estar casi vacía, o puede contener billones de archivos.

Por otra parte, las medidas numéricas en la tabla de hechos transaccional deben referirse al momento del evento de la medida, no a un espacio de tiempo o a algún otro tiempo. Además,

estas tablas de hechos transaccionales son más grandes y detalladas que los otros tipos de tablas de hechos (Kimball and Ross, 2002).

2.1.2. Snapshot Periódica

Las tablas de hechos periódicas, son necesarias para ver los resultados acumulativos del negocio en intervalos de tiempo regulares y predecibles, es decir representan un espacio de tiempo regularmente repetido. Este estilo de tabla de hecho está bien preparado para tratar procesos de larga duración como las cuentas bancarias y otras formas de reporte financiero. Al contrario de la tabla de hecho transaccional, dónde se carga una fila por cada ocurrencia de un evento, con la periódica se saca una foto (snapshot) de la actividad al final de un día, semana, o mes; luego otra foto al final del próximo período, y así sucesivamente. La tabla de hecho periódica es a menudo el único lugar para recuperar fácilmente una vista regular y predecible de las principales métricas resultantes del negocio. Este tipo de tabla es típicamente más compleja que las transaccionales (Kimball and Ross, 2002).

El diseño de la tabla de hecho periódica está estrechamente relacionado al diseño de la tabla de hecho transaccional que la acompaña. Las tablas de hechos, por lo general, comparten muchas tablas dimensionales, aunque las periódicas normalmente tienen un conjunto menor de dimensiones. Recíprocamente, a menudo en una tabla de hechos periódica hay más hechos que los encontrados en una tabla de hecho transaccional (Ponniiah, 2001).

En general, las tablas de hechos periódicas son un tipo de tabla de hecho que representa los resultados comerciales al final de cada período de tiempo regular y predecible. Los períodos diarios y mensuales son los más comunes. Además, las tablas de hechos periódicas son más probables a ser modificadas exitosamente durante su vida agregando más hechos al grano básico de la tabla.

2.1.3. Snapshot Acumulativa

El tercer tipo de tabla de hecho es la acumulativa. El grano de una tabla de hecho acumulativa es la historia completa de una entidad desde su creación al momento presente. Este tipo de tablas de hechos casi siempre tienen múltiples fechas, representando los eventos más predecibles o fases que tienen lugar durante el curso de una vida.

Las tablas de hechos snapshot acumulativas capturan los resultados de los eventos importantes en una serie relacionada de procesos. Las filas de la tabla de hechos son cargadas cuando el primer evento o hito ocurre. A diferencia de los otros tipos de tablas de hechos, la tabla de hecho acumulativa se vuelve a visitar y se actualizan las filas existentes para mostrar los resultados actuales o acumulados de cada evento (Ross, 2003).

Este tipo de tablas de hechos no son apropiadas para los procesos continuos de larga duración y además tienen varias características inusuales. La diferencia más obvia es el gran número de llaves extranjeras de fechas del calendario. Todas las tablas de hechos acumulativas tienen un conjunto de fechas que llevan a cabo la trama normal para la tabla.

Por tanto, se puede decir que las tablas de hechos acumulativas son un tipo de tabla de hecho con múltiples fechas que representan los hitos de un proceso relativamente efímero, denominado tubería. Además se usan para describir procesos que tienen un principio y un fin definido, como el procesamiento de una orden.

Notar que estas tres variaciones de tabla de hecho no son totalmente disímiles, pues ellas comparten dimensiones conformadas que son las principales para construir tablas de hechos separadas que pueden usarse juntas con etiquetas y filtros comunes y consistentes. Mientras las dimensiones son compartidas, la administración y ritmo de las tres tablas de hecho son bastante diferentes. A continuación la Tabla 2.1 muestra de forma resumida una comparación de estos tres tipos de tablas de hechos.

Características	Transaccional	Periódica	Acumulativa
Período de tiempo que representa	Un momento en el tiempo	Intervalos regulares y predecibles	Espacio de tiempo indeterminado, típicamente efímero
Granularidad	Una fila por cada transacción de un evento	Una fila por período	Una fila por vida
Cargas de las tablas de hechos	Insertar	Insertar	Insertar y actualizar
Actualizaciones de las filas de hecho	No se vuelve a visitar	No se vuelve a visitar	Se visita en cualquier actividad
Dimensión fecha	Fecha de transacción	Fecha al final del período	Múltiples fechas para hitos normales
Hechos	Actividad de transacción	Resultados en un intervalo de tiempo predefinido	Resultados durante el tiempo de vida

Tabla 2.1. Comparación entre los tres tipos de tablas de hechos (Kimball and Ross, 2002)

2.2. Caso de estudio inventario (Kimball and Ross, 2002)

La mayoría de las organizaciones tiene una cadena de valor subyacente que consiste en sus principales procesos comerciales. La cadena de valor identifica el flujo natural, lógico de las actividades primarias de una organización. Por ejemplo, en el caso de un vendedor al menudeo (retailer), la compañía puede emitir una orden de compra a un fabricante de un producto. Los productos se envían al almacén del minorista dónde se incluyen en el inventario. Se entregan a una tienda individual dónde de nuevo los productos se asientan en su inventario hasta que un consumidor hace una compra. A continuación se ilustra este subconjunto de la cadena de valor de un minorista o vendedor al detalle en la Figura 2.1.

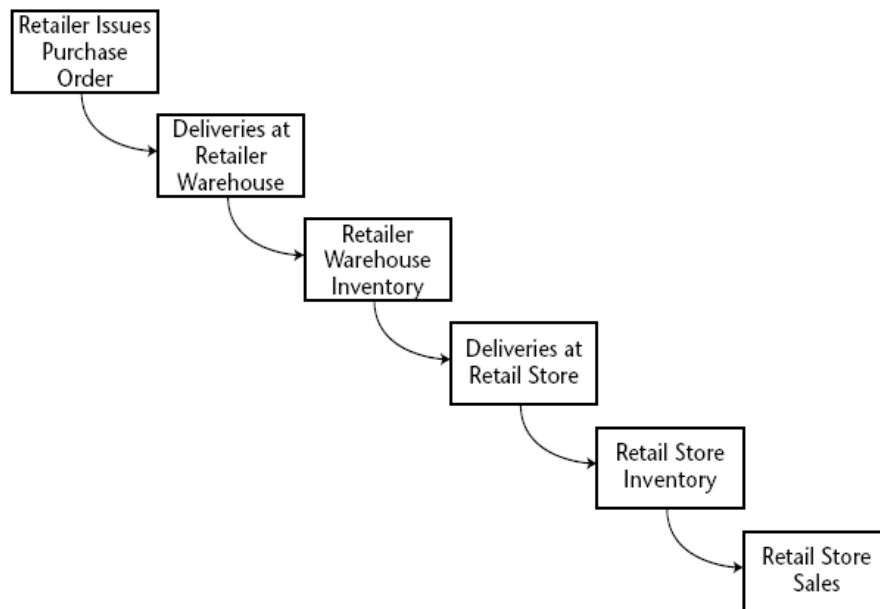


Figura 2.1. Subconjunto de una cadena de valores de un vendedor al menudeo.

Dado que cada proceso comercial produce métricas únicas, en intervalos de tiempo únicos con granularidad y dimensionalidad única, cada proceso típicamente genera una o más tablas de hecho.

Un problema de inventario se puede modelar con tres tipos de tablas de hechos. El primer tipo de tabla de hechos es la snapshot periódica. Todos los días (o para algún otro intervalo de tiempo regular), se miden los niveles de inventario de cada producto y se colocan como filas separadas en una tabla de hechos. Estas filas periódicas aparecen con el tiempo como una serie de capas de datos en el modelo dimensional. El segundo tipo de tabla de hechos es la transaccional, mediante este tipo de tabla de hechos se registra cada transacción que tiene un

impacto en los niveles del inventario en la medida que los productos se mueven a través del almacén. Finalmente, el tercer tipo de tabla de hechos es la snapshot acumulativa, con este tipo de tabla de hechos se construye una fila en la tabla de hechos para cada producto recibido y se actualiza la fila hasta que el producto deja el almacén. Cada uno de los tres modelos de inventario muestra una historia diferente. En algunas aplicaciones de inventario, dos o incluso los tres modelos pueden ser apropiados simultáneamente.

2.2.1. Snapshot periódica para el inventario

Un buen manejo de los niveles de inventario en las tiendas puede tener un gran impacto en la rentabilidad de una empresa. Al asegurar que el producto correcto está en la tienda correcta en el momento correcto se minimiza la no existencia o el agotamiento de alguna mercancía (el producto no está disponible en el estante a ser vendido) y reduce costos derivados por problemas con el inventario. Un minorista necesita tener la posibilidad de analizar la cantidad disponible diaria en los niveles de inventario por producto y tienda.

A continuación se muestra el diseño de este caso de estudio con una metodología que sigue cuatro pasos:

1. Selección del proceso de negocio. El proceso comercial que interesa analizar es el inventario de una tienda que vende al menudeo.
2. Granularidad. Interesa el inventario diario por producto en cada tienda individual, lo que se asume que es el nivel atómico de detalle proporcionado por el sistema de inventario operacional.
3. Las dimensiones resultan inmediatamente de esta declaración del grano: fecha, producto y tienda. Con la información que se tiene hasta el momento no es posible prever otras dimensiones descriptivas adicionales a esta granularidad. El inventario típicamente no está asociado con una dimensión de promoción. Aunque una promoción puede activarse mientras los productos están asentándose en el inventario, la promoción normalmente no es asociada con el producto hasta que realmente se venda. Después de que la promoción ha acabado, los productos todavía pueden permanecer en el inventario. Típicamente, una dimensión para la promoción se asocia con el movimiento del producto, como cuando el producto se pide, se recibe, o se vende.

4. La vista más simple de inventario involucra sólo un hecho: la cantidad disponible. Esto lleva a un diseño dimensional excepcionalmente sencillo, como se muestra en la Figura 2.2.

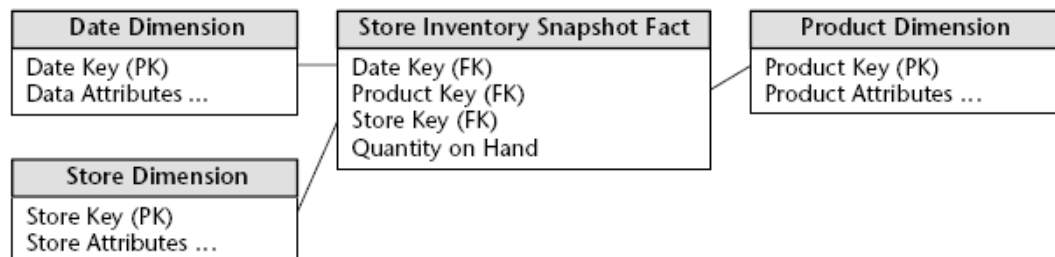


Figura 2.2. Esquema snapshot periódico del inventario de una tienda

Un esquema tan simple como este puede ser muy útil, del mismo puede derivarse numerosas visiones si los niveles del inventario son frecuentemente moderados para muchos productos en muchas situaciones de almacenamiento.

La tabla de hecho periódica enfrenta un desafío serio que por ejemplo una tabla de hechos transaccional de ventas no enfrenta. Una tabla de hechos de ventas es bastante esparcida, pues un por ciento de los productos se vende cada día realmente. Si un producto no se vende en una tienda en un día dado, entonces no habrá ninguna fila en la tabla de hechos para esa combinación de llaves. El inventario, por otro lado, genera tablas periódicas densas. Dado que el minorista se esfuerza por evitar situaciones de no existencia o agotamiento de alguna mercancía, dónde el producto no está disponible para la venta, de manera que hay una fila en la tabla de hecho por cada producto en cada tienda todos los días virtualmente.

En el esquema periódico de inventario, la cantidad disponible puede resumirse por productos o tiendas y puede producirse un total válido. Los niveles de inventario, sin embargo, no son aditivos por las fechas porque ellos representan instantáneas de un nivel o equilibrio en un momento dado en el tiempo. No es posible decir si el inventario de ayer es el mismo o diferente del inventario de hoy solamente mirando los niveles de inventario. Porque los niveles de inventario son aditivos por algunas dimensiones pero no por todas, en este caso nos referimos a ellos como hechos semiaditivos.

La vista simplista de inventario que se desarrolla en nuestra tabla de hecho periódica permite ver una serie de tiempo de niveles de inventario. Para análisis más complejos del inventario, la cantidad disponible no es suficiente. La cantidad disponible necesita ser usada junto con

hechos adicionales para medir la velocidad de movimiento del inventario y desarrollar otras métricas interesantes como el número de giros, el número de suministro diario y el retorno del margen bruto en el inventario (GMROI).

Si se agregara la cantidad vendida a cada fila de la tabla hecho de inventario, se podría calcular el número de giros y el suministro diario. Para el inventario periódico diario, el número de giros medido cada día es calculado como la cantidad vendida dividido por la cantidad disponible. Para un espacio de tiempo extendido, como un año, el número de giros es la cantidad total vendida dividida por la cantidad media diaria disponible. El número de suministros diarios es un cálculo similar. Para un período de tiempo, el número de suministro diario es la cantidad final disponible dividido por la cantidad media vendida.

Además de la cantidad vendida, se puede proporcionar también el valor extendido del costo del inventario, así como el último precio de venta. La diferencia entre estos dos valores es la ganancia bruta. El margen bruto es igual a la ganancia bruta dividida por el último precio de venta.

Finalmente, se puede multiplicar el número de giros por el margen bruto y conseguir el GMROI, como se expresa en la fórmula siguiente:

$$\text{GMROI} = \frac{\text{la cantidad total vendida} * (\text{el último precio de venta} - \text{el valor de costo})}{\text{la cantidad media disponible diaria} * \text{el último precio de venta}}$$

Aunque esta fórmula parece complicada, la idea que sugiere es simple, al multiplicar el margen bruto (gross margin) por el número de giros (turns), se crea una medida de efectividad de la inversión.

GMROI es una norma métrica usada por analistas para evaluar la calidad de la inversión que una compañía hace en su inventario.

Si se quiere ser más ambicioso que con el diseño inicial de la Figura 2.2, entonces se debe incluir la cantidad vendida (quantity sold), el valor de costo (value at cost), y el último precio de venta (value at the latest selling price) en la tabla de hechos snapshot periódica, como se ilustra en la Figura 2.3. Claro, si algunas de éstas métricas existen a diferente granularidad en tablas de hecho separadas, se necesitaría una aplicación para recuperar todos los componentes del cómputo de GMROI al mismo nivel.

Notar que la cantidad disponible es semiaditiva pero que las otras medidas en esta tabla de hechos periódica avanzada son totalmente aditivas por las tres dimensiones. La cantidad

venta se resume al grano particular de la tabla de hecho que es diario en este caso. El valor de GMROI no se guarda en la tabla de hechos, porque no es un hecho aditivo.



Figura 2.3. Inventario snapshot periódico mejorado, para apoyar el análisis de GMROI.

El esquema snapshot periódico es el más común para un inventario. No obstante, se mencionarán brevemente dos alternativas para complementar el diseño. En lugar de describir a estos modelos en el contexto del inventario en la tienda de menudeo, retomamos la cadena de valor para discutir el inventario desde el almacén.

2.2.2. Transacciones en el inventario

Una segunda forma de modelar el proceso de negocio de inventario es registrar cada transacción que afecta el inventario. Las transacciones del inventario en el almacén podrían incluir lo siguiente:

- Recibir el producto
- Poner el producto en el área de inspección
- Liberar el producto del área de inspección
- Devolver el producto al vendedor debido a fallas durante la inspección
- Colocar el producto en su canasta o contenedor.
- Autorizar el producto para la venta
- Escoger el producto de su canasta o contenedor
- Empaquetar el producto para el embarque
- Enviar el producto al cliente
- Recibir el producto del cliente
- Retornar el producto al inventario por devolución del cliente
- Remover el producto del inventario

Cada transacción del inventario identifica la fecha, el producto, el almacén, el vendedor, el tipo de transacción, y en la mayoría de los casos, un único valor que representa el impacto de la cantidad inventariada causada por la transacción. Asumiendo que la granularidad de nuestra tabla de hecho es una fila por transacción del inventario, el esquema resultante se ilustra en la Figura 2.4.

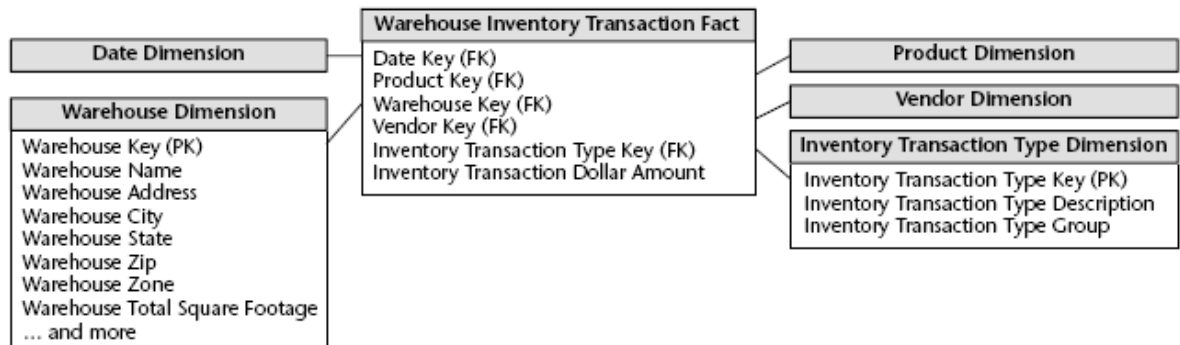


Figura 2.4. Modelo transaccional de inventario en almacén.

Aunque la tabla de hecho a nivel transaccional es de nuevo muy simple, contiene la información disponible más detallada sobre el inventario porque refleja las manipulaciones de inventario. La tabla de hecho a nivel transaccional es útil para medir la frecuencia y cronometrar los tipos de transacción específicos. Por ejemplo, solo una tabla de hecho de granularidad transaccional puede responder las siguientes interrogantes:

- ¿En un mismo día cuántas veces hemos colocado y escogido un determinado producto de la misma canasta o contenedor en momentos diferentes?
- ¿Cuántos embarques o envíos separados hemos recibido de un vendedor dado y cuándo lo hemos aceptado?

Aun así, es impráctico usar esta tabla como la única base para analizar el desempeño del inventario. Aunque es teóricamente posible reconstruir la posición exacta del inventario en cualquier momento en el tiempo dándole seguimiento a todas las transacciones ocurridas desde una posición conocida del inventario; es demasiado embarazoso e impráctico para responder preguntas sobre el almacén de datos que incluyan fechas o productos.

2.2.3. Snapshot acumulativa para el inventario

El modelo del inventario que se explora finalmente de manera breve es el acumulativo. En este modelo se coloca una fila en la tabla de hecho para cada envío de un producto particular al

almacén. En una única fila de la tabla de hecho se registra desde el envío del producto hasta que el mismo deja el almacén. El modelo acumulativo sólo es posible si se pueden distinguir fiablemente productos entregados en un envío de aquellos entregados en otro. Este enfoque también es apropiado si se puede dar seguimiento a los envíos con un alto nivel de detalle como el número de serie del producto o el número del lote.

Asumamos que el inventario pasa por una serie de eventos bien definidos o hitos como los movimientos a través del almacén, tales como, recibir, inspeccionar, colocar en canasta o similar, autorizar la venta, escoger, embalar, y enviar. La filosofía que sigue la tabla de hechos acumulativa es proporcionar un estado actual desde el embarque de un producto y como pasa por todas esas fases o hitos hasta que deja el almacén. Cada fila de la tabla de hecho se actualiza hasta que el producto deje el almacén. Como se ilustra en la Figura 2.5, la tabla de hechos del inventario acumulativo con sus múltiples fechas y hechos es bastante diferente de los esquemas transaccionales y periódicos.

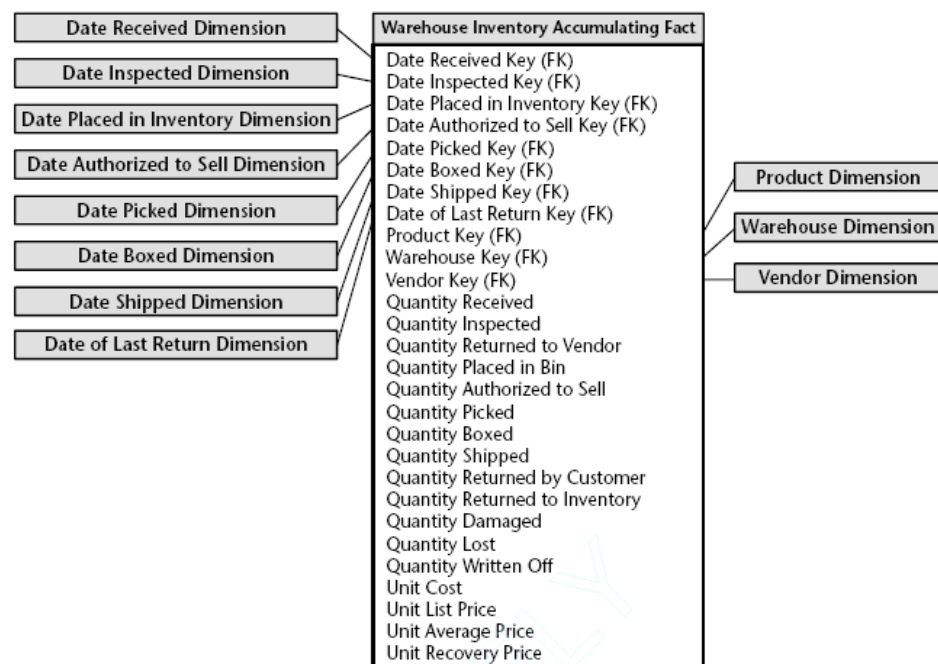


Figura 2.5. Inventario snapshot acumulativo en almacén

La tabla de hechos acumulativa es el tercer tipo de tabla de hechos; estas son interesantes debido a las múltiples llaves extranjeras de fechas estimadas de la lista de llaves y también porque la misma tabla de hechos se vuelve a visitar y modificar una y otra vez.

2.3. Dimensiones conformadas

Las dimensiones conformadas son subconjuntos idénticos o estrictos de una dimensión; granular y detalladas. Por otra parte, las dimensiones conformadas tienen llaves consistentes; además, los nombres de columnas que indican atributos, las definiciones de los atributos y los valores de los atributos, también son consistentes. Por tanto, se considera que las tablas de dimensiones no son conformadas si los atributos se etiquetan o contienen valores diferentes (Kimball and Ross, 2002).

De igual forma, se puede plantear que una dimensión conformada es una dimensión que significa lo mismo para cada posible tabla de hecho con la que se puede acoplar. Generalmente esto significa que una dimensión conformada es idénticamente la misma dimensión en cada Mercado de Datos (Kimball et al., 1998).

Las dimensiones conformadas son enormemente importantes en los Almacenes de Datos, por esta razón, se debe tener en cuenta que sin un uso estricto de las mismas, el Almacén de Datos no puede funcionar como un todo integrado. Además, si una determinada dimensión es común a varios mercados de datos y no es usada como una dimensión conformada, entonces estos mercados de datos separados, simplemente no pueden usarse juntos, o peor aun, en el caso de que se intente usarlos juntos se producirán resultados erróneos. De esta forma, las dimensiones conformadas hacen posible:

- Una única tabla de dimensión se puede usar contra múltiples tablas de hechos en un mismo Almacén de Datos.
- Las interfaces de usuario y el contenido de los datos son consistentes donde quiera que se use.
- Hay una interpretación consistente de atributos.

La creación de dimensiones conformadas es tanto una decisión política como técnica, el uso de las mismas debe apoyarse a los niveles ejecutivos más altos. Por otra parte, es responsabilidad del equipo diseñador establecer, publicar y mantener las dimensiones conformadas. También se debe tener en cuenta que la tarea de identificar y diseñar estas dimensiones puede llevar mucho tiempo y se deben diseñar al nivel de detalle más bajo que sea posible; además deben tener una clave no significativa para poder permitir cambios en el futuro.

2.4. Rol jugado por las dimensiones

Un rol en un Almacén de Datos es cuando una sola dimensión aparece varias veces en la misma tabla de hechos. Esto puede pasar de varias maneras, por ejemplo, en ciertos tipos de tablas de hechos, el tiempo puede aparecer repetidamente.

Una sola dimensión puede jugar varios roles simultáneos cuando usa una tabla de hechos. La dimensión subyacente puede existir como una sola tabla física, pero cada uno de los roles debe presentarse en una vista separadamente etiquetada.

Un rol-jugado ocurre cuando una sola dimensión física aparece varias veces en una tabla de hechos, cada uno representado como una tabla lógica separada con nombres de columnas únicos a través de las vistas (Kimball et al., 1998).

2.5. Tablas de hechos consolidadas

Nos referimos a tablas de hechos que combinan métricas a una granularidad común, como consolidadas o tablas de hechos de segundo nivel (o Mercados de Datos consolidados). Mientras las tablas de hechos consolidadas pueden ser muy útiles por lo que se refiere a actuación y utilidad, ellas representan a menudo una dimensionalidad compuesta, porque consolidan hechos al menor conjunto del común denominador de dimensiones (Kimball and Ross, 2002).

Un riesgo potencial asociado con las tablas de hechos consolidadas, es que a veces equipos del proyecto basan sus planes solamente en la granularidad de la tabla de hecho consolidada, mientras no reúnen requisitos del usuario que exigen la habilidad de buscar en los datos más granulares. Estos esquemas también se pueden encontrar en serios problemas si los equipos del proyecto intentan forzar una correspondencia uno a uno para consolidar los datos con diferente granularidad o dimensionalidad.

Cuando se combinan hechos de múltiples procesos de negocio en una tabla de hechos consolidada, ellos deben estar al mismo nivel de granularidad y dimensionalidad. Óptimamente, los hechos separados naturalmente están a un grano común. Por otra parte, estamos obligados a eliminar o agregar algunas dimensiones para apoyar la correspondencia uno a uno o mantenerlos en tablas de hechos separadas.

En general podemos decir que las tablas de hechos consolidadas combinan los resultados de procesos de negocios separados cuando los mismos se analizan juntos frecuentemente.

2.6. Conclusiones del capítulo

En este capítulo se realizó una descripción del caso de estudio de inventario propuesto por (Kimball and Ross, 2002), el cual dio paso al estudio de importantes conceptos que caracterizan las particularidades del mercado de datos que se muestra en la tesis. A partir del análisis y el estudio realizado de lo expuesto en este capítulo, se concluyó lo siguiente:

1. Las dimensiones conformadas son necesarias en el diseño del Mercado de Datos, debido a la estrecha relación que existe entre subsistemas.
2. El uso de tablas de hechos periódicas y transaccionales simultáneamente, es apropiado para modelar los subprocesos originados por los tipos de operaciones presentados en el subsistema de inventario.
3. Las tablas de hechos consolidadas son útiles en este Mercado de Datos, pues permiten realizar análisis globales de los distintos subprocesos modelados en el mismo.
4. Definir los roles jugado por las dimensiones, como por ejemplo la dimensión entidad, permite esclarecer que función realizan en una determinada operación.

Capítulo 3. DISEÑO E IMPLEMENTACIÓN DEL MERCADO DE DATOS PARA EL PROCESO DE INVENTARIO DEL VERSAT SARASOLA

En este capítulo se describe el diseño del caso de estudio. Además se muestra la descripción de los procesos de Extracción, Transformación y Carga con la herramienta Pentaho Data Integration; así como el diseño del cubo de datos y algunos reportes que permiten validar el mercado, tomando datos reales de la empresa ULAEX, SA.

Para ejemplificar y mostrar los esquemas y diseños mencionados anteriormente se seleccionó el subproceso de compras, siendo este el más grande y abarcador del Mercado de Datos.

3.1. Pasos para el diseño lógico de la solución.

El diseño de un Mercado de Datos es el sostén de la solución a las necesidades planteadas por el cliente. Después de analizar el caso de estudio propuesto en el Capítulo 1 de la tesis, estamos en condiciones de mostrar el diseño del Mercado de Datos para el proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola. La metodología que se utiliza para el diseño de este Mercado de Datos, contiene cuatro pasos propuesta por (Kimball et al., 1998, Kimball and Ross, 2002).

3.1.1. Paso 1. Seleccionar el proceso de negocio

Un proceso es una actividad de la organización soportada por un sistema OLTP del cual se puede extraer información con el propósito de construir el Mercado de Datos.

El proceso comercial que interesa analizar en este caso es el proceso de inventario del VERSAT Sarasola, a nivel de almacenes y de gestión de custodios.

3.1.2. Paso 2. Determinar el grano del proceso de negocio

La declaración del grano especifica exactamente qué representa una fila individual en la tabla de hechos. El grano es el nivel de detalle al que se desea almacenar información sobre la actividad a modelar. Define el nivel atómico de datos en el Almacén de Datos y determina las dimensiones básicas del esquema. Finalmente, el grano es la respuesta a la pregunta: ¿Cómo usted describe una fila en la tabla de hechos?

Luego, interesa el inventario diario por producto con su unidad de medida, en cada almacén o custodio de una unidad contable, para un concepto determinado; lo que se asume que es el nivel atómico de detalle proporcionado por el sistema de inventario operacional.

3.1.3. Paso 3. Seleccionar las dimensiones

Las dimensiones resultan de la pregunta: ¿Cómo las personas del negocio describen los datos que resultan del proceso comercial? Las dimensiones caracterizan la actividad al nivel de detalle (grano) que se ha elegido. De cada dimensión se debe decidir los atributos o propiedades relevantes para el análisis de la actividad.

Por tanto, en este caso, las dimensiones resultan inmediatamente de la declaración del grano que se planteó anteriormente: fecha, producto, unidad de medida, almacén, custodio, unidad contable, concepto. Es importante destacar que con la información que se tiene hasta el momento no es posible prever otras dimensiones descriptivas adicionales a esta granularidad.

3.1.4. Paso 4. Identificar los hechos

Los hechos constituyen información sobre la actividad que se desea almacenar en cada tupla de la tabla de hechos y que será el objeto del análisis. Los hechos se determinan contestando la pregunta: ¿Qué estamos midiendo? Todos los hechos candidatos en un diseño deben estar vinculados con el nivel de granularidad propuesto previamente.

La vista más simple de inventario involucra sólo tres hechos: la cantidad por concepto, el importe por concepto y la existencia.

3.2. Subprocesos identificados en el proceso de inventario

En el proceso de inventario del VERSAT Sarasola, cada movimiento de productos en almacenes o custodios genera un tipo de operación específico. En este caso existen nueve tipos de operaciones, las cuales a su vez, constituyen subprocessos dentro del propio proceso de inventario. Durante el análisis y descripción de estos subprocessos, se debe tener en cuenta que los productos en el almacén tienen dos formas de tratamiento: unitario y por lotes.

Cada uno de estos subprocessos genera un conjunto de hechos y dimensiones descriptivas adicionales que refinan el diseño del Mercado de Datos.

3.2.1. Subproceso de compra

En el proceso de inventario, las compras sólo interesan a nivel de almacenes, es decir, a nivel de gestión de custodios no se realizan operaciones de compras. Cuando estamos en presencia de una compra, las empresas, unidades básicas o empresariales o cualquier otro centro, denominado de forma genérica como entidades; pueden asumir el rol de proveedor. Hay que

destacar que en este tipo de operación, las entidades también pueden prestar determinado servicio; por ejemplo, manipulación, transportación de la mercancía u otro tipo.

Luego, en el subproceso de compra interesa: las compras diarias por producto con su unidad de medida y la fecha de vencimiento asociada al mismo, en cada almacén de una unidad contable, por categoría, por entidad (proveedora o de prestación de servicio), por servicio y para un concepto determinado. De estas compras interesa medir la cantidad que se compró y el importe, así como el importe en servicios y la existencia real después de realizada dicha operación. En el caso de aquellos productos que se mueven en el almacén por lotes, no sólo interesa cuántos productos se compraron, sino también en cuántos lotes y a su vez, la cantidad de productos vencidos y no vencidos.

3.2.2. Subproceso de venta

Las ventas, al igual que las compras, solo se realizan a nivel de almacenes. En este caso también participan las entidades, pero solo como clientes. Por tanto, en el subproceso de venta interesa: las ventas diarias por producto con su unidad de medida y la fecha de vencimiento asociada al mismo, en cada almacén de una unidad contable, por categoría, por entidad y para un concepto determinado. Además, interesa medir la cantidad vendida, el importe, así como la existencia real después de cada venta realizada. En el caso de aquellos productos que se mueven en el almacén por lotes, no solo interesa cuántos productos se vendieron, sino también en cuántos lotes y a su vez, la cantidad de productos vencidos y no vencidos.

3.2.3. Subprocesos de entrada a centros de costo y salida de centros de costo

Es importante tener en cuenta que las operaciones entrada a centros de costo (CC) y salida de CC, indican entradas y salidas de productos para gastos, respectivamente, en determinados CC. También es importante destacar que estas dos operaciones están presentes a nivel de almacén y de gestión de custodio.

Luego, a nivel de almacén interesa: las entradas o salidas diarias, según el tipo de operación, por producto con su unidad de medida y la fecha de vencimiento asociada al mismo, en cada almacén de una unidad contable, por categoría, por centro de costo y para un concepto determinado. Además, interesa medir la cantidad que entró o salió, el importe, y la existencia. En el caso de aquellos productos que se mueven en el almacén por lotes, no solo interesa

cuántos productos entraron o salieron en total, sino también la cantidad de productos vencidos y no vencidos.

A nivel de gestión de custodio interesa: las entradas o salidas diarias, según el tipo de operación, por producto con su unidad de medida, por custodio de una unidad contable, por centro de costo y para un concepto determinado. Además, interesa medir la cantidad que entró o salió, el importe, y la existencia.

3.2.4. Subprocesos de entrada, salida y transferencia

Similar a los subprocesos vistos en el punto anterior, estos tres tipos de operaciones están presentes a nivel de almacén y de gestión de custodio. Las operaciones de entradas y salidas, indican aquellos productos que se mueven en el almacén mediante donaciones, bajas, sobrante, pérdidas, entre otras.

Luego, para estos tres subprocesos a nivel de almacén, interesa: los movimientos diarios, según el tipo de operación, por producto con su unidad de medida y la fecha de vencimiento asociada al mismo, en cada almacén de una unidad contable, por categoría y para un concepto determinado. Además, interesa medir la cantidad movida, el importe, y la existencia. En el caso de aquellos productos que se mueven en el almacén por lotes, no sólo interesa cuántos productos se vendieron, sino también en cuántos lotes y a su vez, la cantidad de productos vencidos y no vencidos.

A nivel de gestión de custodio interesa: los movimientos diarios, según el tipo de operación, por producto con su unidad de medida, por custodio de una unidad contable y para un concepto determinado. Además, interesa medir la cantidad que se movió, el importe, y la existencia.

3.2.5. Subproceso de ajuste

Los ajustes son operaciones internas que realizan los responsables del inventario para facilitar el desarrollo y desempeño de este proceso. De igual forma, se realiza a nivel de almacén y de gestión de custodio.

A nivel de almacén interesa: el ajuste diario por producto con su unidad de medida, en cada almacén de una unidad contable, por categoría y para un concepto determinado. Además, interesa medir la cantidad y el importe.

Luego, a nivel de gestión de custodio interesa: el ajuste diario por producto con su unidad de medida, por custodio de una unidad contable, para un concepto determinado. De igual forma interesa medir la cantidad, el importe y la existencia.

3.2.6. Subproceso de devolución

El subproceso de devolución, al igual que los subprocesos vistos anteriormente, también se analiza a nivel de gestión de custodio y a nivel de almacén. En este subproceso interesa conocer aquellas entidades que están involucradas cuando se realiza alguna devolución.

Por tanto, en el subproceso de devolución a nivel de almacén interesa: las devoluciones diarias por producto con su unidad de medida y su fecha de vencimiento asociada, en cada almacén de una unidad contable, por categoría, por entidad y para un concepto de terminado. Además, interesa medir la cantidad devuelta, el importe, y la existencia. En el caso de aquellos productos que se mueven en el almacén por lotes, no solo interesa cuántos productos se devolvieron en total, sino también la cantidad de productos vencidos y no vencidos.

A nivel de gestión de custodio interesa: las devoluciones diarias por producto con su unidad de medida, por custodio de una unidad contable, para un concepto determinado. De igual forma interesa medir la cantidad, el importe y la existencia.

3.3. Diseño del Mercado de Datos para el proceso de inventario

El proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola que se describe en la tesis, presenta ciertas características y particularidades que lo distinguen del caso de estudio de inventario discutido por Kimball (Kimball and Ross, 2002). En las secciones anteriores pudimos identificar las principales dimensiones, hechos y granularidad del proceso de inventario general mediante la metodología que sigue cuatro pasos. Además, con la descripción de los subprocesos, se previeron otras dimensiones y hechos necesarios en el diseño del Mercado de Datos.

De igual forma nos pudimos percatar de la necesidad de usar distintos tipos de tablas de hechos con el objetivo de facilitar y viabilizar el diseño y el análisis posterior del Mercado de Datos.

A continuación se describen las dimensiones y los distintos tipos de tablas de hechos identificados.

3.3.1. Dimensiones identificadas

En el diseño del Mercado de Datos del proceso de inventario se identificaron 11 dimensiones, las cuales se describen a continuación:

1. dim_unidadcontable

Esta dimensión describe las propiedades del universo de valores de unidades contables existentes.

2. dim_almacen

Esta dimensión describe las propiedades del universo de valores de almacenes existentes.

3. dim_custodio

Esta dimensión describe las propiedades del universo de valores de custodios existentes.

4. dim_producto

Esta dimensión describe las propiedades del universo de valores de productos existentes.

5. dim_unidadmedida

Esta dimensión describe las propiedades del universo de valores que permite cuantificar el volumen de la operación que se realiza.

6. dim_categoria

Esta dimensión describe las propiedades del universo de valores que permite categorizar a los productos según el destino que los mismos puedan tener en el almacén.

7. dim_concepto

Esta dimensión describe las propiedades del universo de valores de conceptos existentes.

8. dim_entidad

Esta dimensión describe las propiedades del universo de valores de entidades existentes.

9. dim_centrocosto

Esta dimensión describe las propiedades del universo de valores de centros de costos existentes.

10. dim_servicio

Esta dimensión describe las propiedades del universo de valores de servicios existentes.

11. dim_fecha

Esta es la dimensión más común en el diseño de almacenes de datos debido a que define una línea de tiempo para enmarcar la información almacenada.

3.3.2. Tablas de hechos transaccionales identificadas

Teniendo en cuenta las características que presentan las tablas de hechos transaccionales y las formas de tratamiento de los productos (unitario y por lote) en el almacén, se decidió incorporar este tipo de tablas de hechos en el Mercado de Datos. En total hay ocho tablas de hechos transaccionales, las cuales están relacionadas a los subprocesos que tienen en cuenta el tratamiento de los productos por lotes.

Seguidamente se describen estas tablas de hechos transaccionales

1. hch_compras_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las compras de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_servicio, dim_unidadmedida, dim_categoria, dim_almacen, dim_producto, dim_entidad, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe, importe por servicio y existencia.

2. hch_ventas_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las ventas de productos realizadas. Se relaciona con las dimensiones: dim_entidad, dim_fecha, dim_concepto, dim_unidadcontable, dim_categoria, dim_almacen, dim_producto, dim_unidadmedida. Las medidas que almacena son: cantidad, importe y existencia.

3. hch_transferencia_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las transferencias de productos realizadas, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_unidadmedida, dim_concepto, dim_fecha, dim_unidadcontable, dim_categoria, dim_almacen, dim_producto. Las medidas que almacena son: cantidad, importe y existencia.

4. hch_entrada_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las entradas de productos realizadas, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_unidadcontable, dim_unidadmedida, dim_concepto, dim_categoria, dim_almacen, dim_producto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

5. hch_salida_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las salidas de productos realizadas, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_unidadmedida, dim_fecha, dim_unidadcontable, dim_categoria, dim_almacen, dim_producto, dim_concepto. Las medidas que almacena son: cantidad, importe y existencia.

6. hch_devolucion_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las devoluciones de productos realizadas, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_fecha, dim_unidadcontable, dim_unidadmedida, dim_categoria, dim_almacen, dim_entidad, dim_producto, dim_concepto. Las medidas que almacena son: cantidad, importe y existencia.

7. hch_entrada_cc_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las entradas de productos en los centros de costos, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_unidadmedidad, dim_unidadcontable, dim_centrocosto, dim_categoria, dim_almacen, dim_producto, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

8. hch_salida_cc_transaction

Esta tabla de hechos recoge toda la información diaria y detallada relacionada con las salidas de productos en los centros de costos, pero sólo a nivel de almacén. Se relaciona con las dimensiones: dim_unidadmedida, dim_unidadcontable, dim_centrocosto, dim_categoria, dim_almacen, dim_producto, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

3.3.3. Tablas de hechos periódicas identificadas

Uno de los principales objetivos de este Mercado de Datos, es obtener información de los distintos subprocesos al final de un determinado período (día, quincena, mes, etc.). Para ello se modelaron 16 tablas de hechos periódicas, las cuales están relacionadas con cada subproceso identificado anteriormente, tanto a nivel de almacén, como a nivel de gestión de custodio.

A continuación se describen las nueve tablas de hechos periódicas relacionadas con los subprocesos a nivel de almacén.

1. hch_compras_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las compras de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que se almacenan son: cantidad de lotes, cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

2. hch_ventas_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las ventas de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de lotes, cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

3. hch_transferencia_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las transferencias de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de lotes, cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

4. hch_entrada_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las entradas de productos realizadas. Se relaciona con las dimensiones:

dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de lotes, cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

5. hch_salida_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las salidas de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de lotes, cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

6. hch_devolucion_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las devoluciones de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

7. hch_entrada_cc_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las entradas de productos en los centros de costos. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

8. hch_salida_cc_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con las entradas de productos en los centros de costos. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad de

productos vencidos, cantidad de productos no vencidos, cantidad de productos total y el importe total.

9. hch_ajuste_periodic

Esta tabla de hechos recoge de una forma más compactada y resumida la información relacionada con los ajustes realizados. Se relaciona con las dimensiones: dim_fecha, dim_unidadcontable, dim_producto, dim_unidadmedida, dim_categoria, dim_almacen, dim_concepto. Las medidas que almacena son: cantidad que se ajusto y el importe.

A continuación se describen las siete tablas de hechos periódicas relacionadas con los subprocesos a nivel de gestión de custodio

1. hch_transferencia_custodio

Esta tabla de hechos recoge toda la información relacionada con las transferencias de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

2. hch_entrada_custodio

Esta tabla de hechos recoge toda la información relacionada con las entradas de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

3. hch_salida_custodio

Esta tabla de hechos recoge toda la información relacionada con las salidas de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

4. hch_devolucion_custodio

Esta tabla de hechos recoge toda la información relacionada con las devoluciones de productos realizadas. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

5. hch_entrada_cc_custodio

Esta tabla de hechos recoge toda la información relacionada con las entradas de productos en los centros de costos. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

6. hch_salida_cc_custodio

Esta tabla de hechos recoge toda la información relacionada con las salidas de productos en los centros de costos. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

7. hch_ajuste_custodio

Esta tabla de hechos recoge toda la información relacionada con los ajustes realizados. Se relaciona con las dimensiones: dim_unidadcontable, dim_producto, dim_unidadmedida, dim_custodio, dim_concepto, dim_fecha. Las medidas que almacena son: cantidad, importe y existencia.

3.3.4. Tablas de hechos consolidadas identificadas

Con el fin de combinar y agrupar los resultados de los distintos subprocesos vistos anteriormente, se modelaron dos tablas de hechos consolidadas: una a nivel de almacén y otra a nivel de gestión de custodio. El principal objetivo de estas dos tablas de hechos es registrar las existencias reales en un período determinado.

1. hch_existencia_almacen

Esta tabla de hechos está relacionada con las dimensiones: dim_fecha, dim_almacen, dim_unidadcontable, dim_categoria, dim_producto, dim_unidadmedida. La única medida que almacena es la existencia real al final de un determinado período.

2. hch_existencia_custodio

Esta tabla de hechos está relacionada con las dimensiones: dim_fecha, dim_custodio, dim_unidadcontable, dim_producto, dim_unidadmedida. La única medida que almacena es la existencia real al final de un determinado período.

La Figura 3.1 muestra el esquema estrella correspondiente al subproceso de compras a nivel de transacción, es decir cada entrada en la tabla de hechos representa la compra de un producto

que tiene tratamiento por lotes. Se seleccionó este subproceso en particular, de manera coherente con el resto de los epígrafes.

Este esquema describe detalladamente la relación existente entre la tabla de hechos y las dimensiones involucradas en este subproceso, así como los atributos de las dimensiones y las medidas de la tabla de hechos. Es importante destacar que en este esquema se evidencia como modelar y representar los roles que puede jugar una determinada dimensión, en este caso la dimensión entidad, que como se ha explicado anteriormente en este subproceso de compras funge como cliente y como entidad que presta servicios; y de igual forma la dimensión fecha.

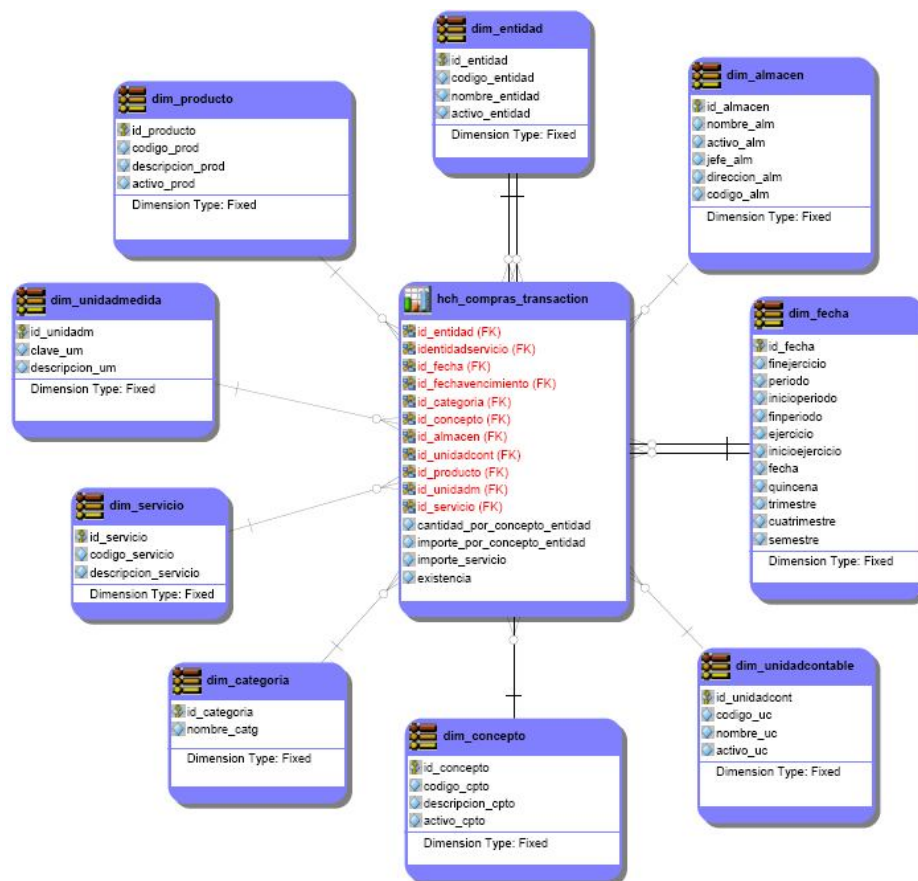


Figura 3.1. Esquema estrella del subproceso de compras a nivel de transacción.

Con el fin de mostrar las diferencias entre los tipos de de tablas de hechos transaccionales y periódicas, la Figura 3.2 muestra el diseño del esquema estrella relacionado con este mismo subproceso de compras, pero a un nivel más resumido y compactado; es decir la tabla de hechos que se muestra es del tipo snapshot periódica, ya descrita en el epígrafe 3.3. De igual

forma se describe la relación existente entre la tabla de hechos y las dimensiones involucradas en este subproceso, así como los atributos de las dimensiones y las medidas de la tabla de hechos.

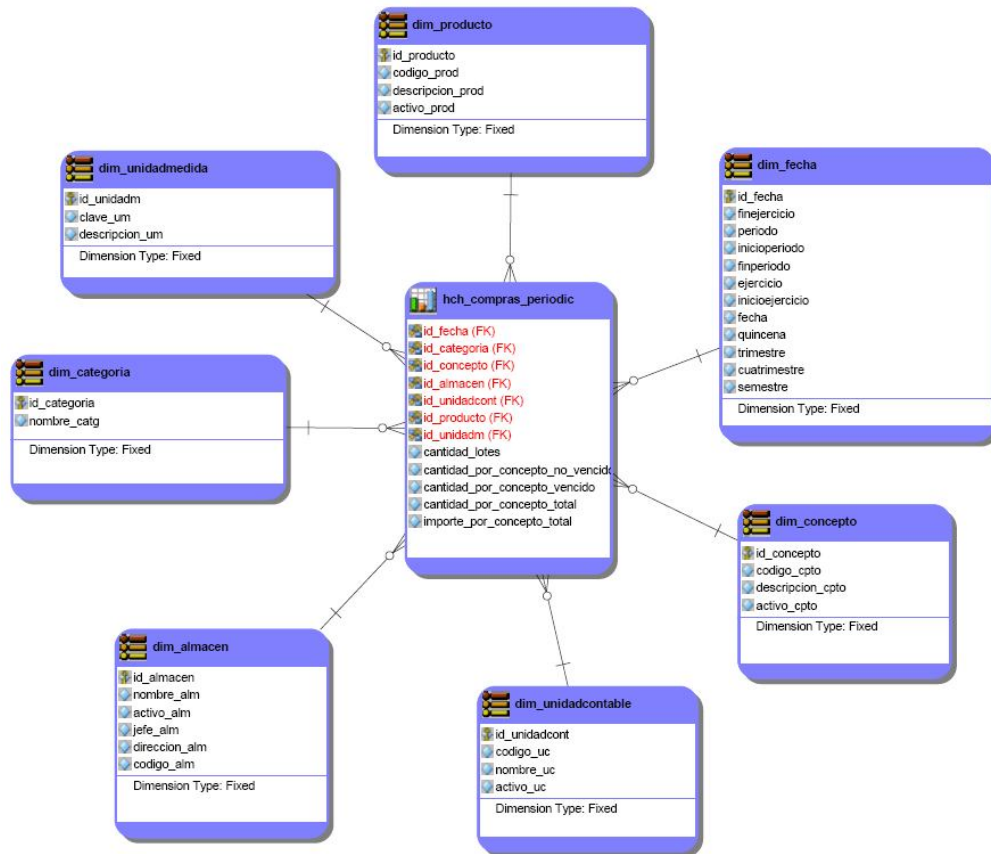


Figura 3.2. Esquema estrella del subproceso de compras a nivel periódico.

El diseño del resto de los subprocesos es similar a los esquemas de estrella vistos anteriormente, ver Anexo I.

3.4. Implementación del Mercado de Datos para el proceso de inventario

Posterior al diseño multidimensional del Mercado de Datos, se procede a la implementación física del mismo. En el modelo dimensional, tanto el diseño lógico como el físico poseen una gran semejanza. El modelo físico se diferenciará del lógico en términos de los detalles específicos del gestor incluyendo tipos de datos, relación entre tablas, etc.

El punto de partida del modelo físico es el modelo lógico, este debe reflejar al modelo lógico en la mayor medida posible, sin embargo algunos cambios en las estructuras de las tablas y las

columnas son necesarios para ajustarla a las características propias del Sistema Gestor de Base de Datos Relacionales (SGBDR) donde se almacena la información relativa al Mercado de Datos; en este caso el Sistema Gestor de Base de Datos a utilizar es el PostgreSQL en su versión 9.0.

Para desarrollar el modelo físico generalmente es necesario la utilización de herramientas que automaticen el proceso, en este caso el ERStudio en su versión 8.0. Entre las funcionalidades que brinda esta herramienta se encuentra la sincronización bidireccional de los diseños lógico y físico, la construcción automática de base de datos, la reingeniería inversa de base de datos, entre otras.

Una vez elaborado el modelo lógico, se genera el modelo físico, dando paso de esta forma, a la generación de la base de datos que almacena la información correspondiente al Mercado de Datos para la toma de decisiones en el Sistema Gestor de Base de Datos PostgreSQL, como se muestra en la Figura 3.3.

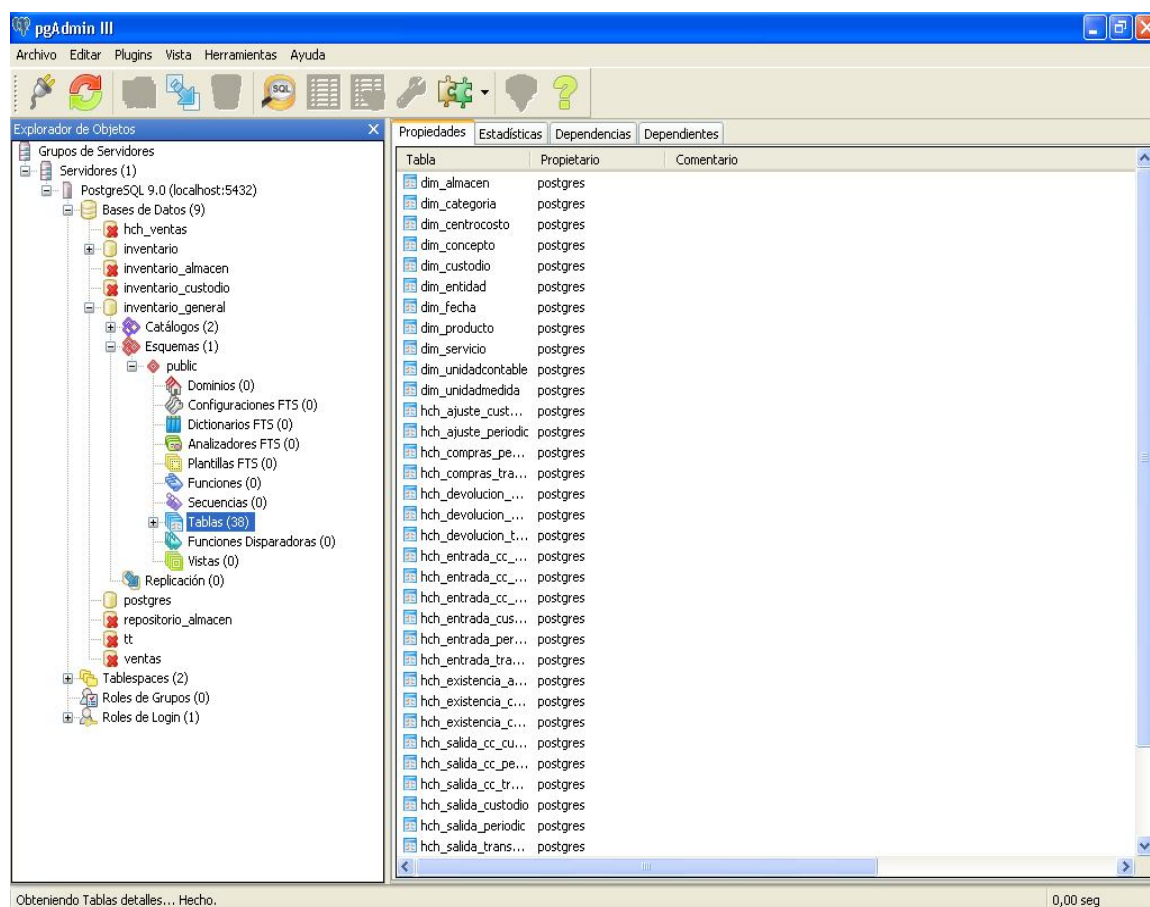


Figura 3.3. Diseño físico del Mercado de Datos.

3.4.1. Proceso de Extracción, Transformación y Carga (ETL)

Para construir un Mercado de Datos los diseñadores deben tener una amplia perspectiva del uso que se espera del mercado para ello es necesario seleccionar un esquema adecuado que refleje el uso previsto.

Muchas de las cuestiones que rodean a los sistemas de apoyo para la toma de decisiones, se refieren en primer lugar a las tareas de obtener y preparar los datos. Para poder explotar la información existente en los diferentes sistemas empresariales, hay que extraerla de ellos, transformarla y llevarla al Mercado de Datos mediante un conjunto de procesos y herramientas conocidas como ETL (Extract, Transform and Load). De forma muy resumida, la extracción consiste en extraer los datos de la fuente de datos original. La transformación se refiere al hecho de limpiar la información incoherente o no válida, adaptar los tipos de datos, los valores de los campos, etc. Por último la carga se refiere al hecho de guardar los datos en el formato adecuado dentro del Mercado de Datos. Cada una de estas operaciones involucra sus propias consideraciones especiales. Para una descripción más detallada de estos procesos, ver 34 subsistemas de Kimball (Casters et al., 2010).

Para este caso de estudio los datos deben ser extraídos, transformados, limpiados y cargados desde la base de datos centralizada diseñada sobre el Sistema de Gestión de Bases de Datos SQL Server 2000 hacia el Mercado de Datos ubicado en el Sistema de Gestión de Bases de Datos PostgreSQL.

Para los procesos de extracción, transformación y carga se utilizan las ventajas que brinda la herramienta Pentaho Data Integration (PDI), la cual es una herramienta ETL que nos permite extraer la información de una base de datos OLTP, transformar la información a través de un modelo dimensional y cargar los resultados de la transformación en una base de datos destino, para que luego pueda ser consultada y analizada a través de herramientas para desarrollar reportes especializados. PDI está formado por un conjunto de herramientas, cada una con un propósito específico.

- Spoon; es la herramienta gráfica que nos permite el diseño de las transformaciones y trabajos. Incluye opciones para previsualizar y comprobar los elementos desarrollados. Es la principal herramienta de trabajo de PDI y con la que se construyen y validan los procesos ETL.

- Pan: es la herramienta que nos permite la ejecución de las transformaciones diseñadas en Spoon, ya sea desde un fichero o desde el repositorio. Esta herramienta nos permite desde la línea de comandos preparar la ejecución mediante scripts.
- Kitchen: similar a Pan, pero para ejecutar los trabajos o jobs.
- Carte: es un pequeño servidor Web que permite la ejecución remota de transformaciones y jobs.

En la Figura 3.4 se muestra un ejemplo de la carga de la tabla de hechos, hch_compras_transaction, relacionada con el subproceso de compras a nivel transaccional. En este caso la información almacenada en la base de datos fuente se ha sido procesada en dos grupos diferentes; el primero incluye aquellos productos que en el sistema original no tienen una fecha de vencimiento asociada y el segundo grupo incluye a los productos que tienen una fecha de vencimiento asociada. Esta división se justifica porque en la tabla destino relacionada con este subproceso, cada entrada representa la compra de un producto que tiene tratamiento por lotes, como se expuso anteriormente, y por tanto el procesamiento de aquellos productos que tienen una fecha de vencimiento asociada en el sistema original, es diferente al tratamiento de aquellos productos que no la tienen.

Después de extraer esta información de la base de datos fuente, se efectúa un proceso de transformación y limpieza, ajustándola a la información del cliente. Posteriormente se realiza un proceso de búsqueda de las llaves subrogadas relacionadas con el hecho, para luego realizar la carga de los datos en la tabla destino.

Los puntos señalados en la Figura 3.4 indican lo siguiente:

1. En este paso se realiza la consulta a la base de datos fuentes para seleccionar la relación de aquellos productos que no tienen una fecha de vencimiento asociada, como se muestra en la Figura 3.5.
2. En este paso se realiza la consulta a la base de datos fuentes para seleccionar la relación de aquellos productos que tienen una fecha de vencimiento asociada, como se muestra en la Figura 3.6.

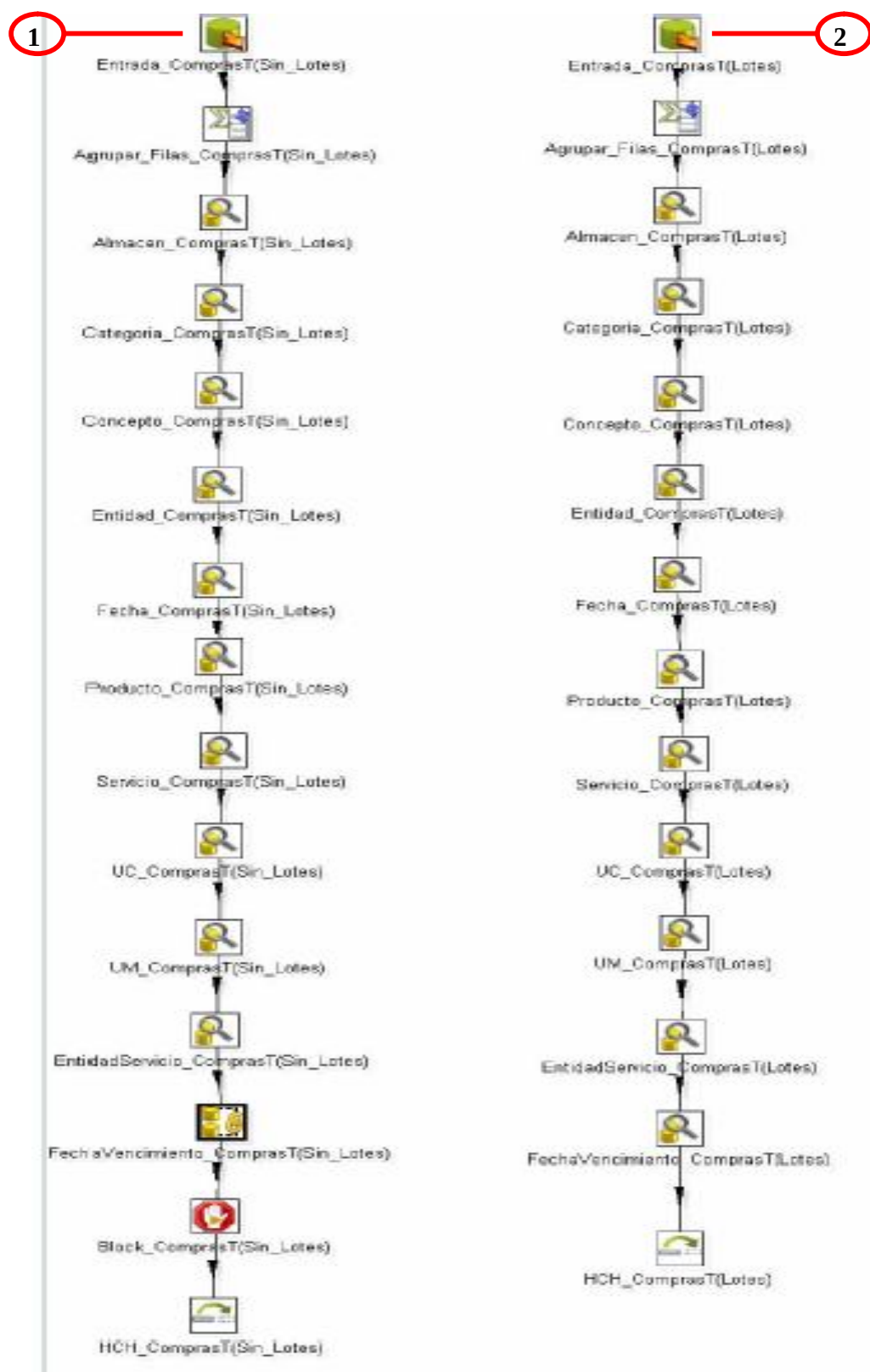


Figura 3.4. Proceso ETL para el subproceso de compras a nivel transaccional

```

SELECT dbo.inv_documento.fecha AS fecha_sqls, dbo.inv_documentoserv.identidad AS id_entidadServicio,
       dbo.inv_movimiento.idproducto, dbo.inv_documento.idconcepto, dbo.inv_documento.idunidad,
       dbo.inv_documentoentidad.identidad, dbo.inv_documentoserv.idservicio, dbo.inv_concepto.idcategoria,
       dbo.inv_documentoalm.idalmacen, dbo.inv_movimiento.idmedida, dbo.inv_movimiento.importe,
       dbo.inv_movimiento.cantidad, dbo.inv_documentoserv.importe AS importeServicio,
       dbo.inv_movimiento.existencia AS Existencia
FROM dbo.gen_producto INNER JOIN dbo.inv_documentoentidad INNER JOIN dbo.inv_documentoalm ON
       dbo.inv_documentoentidad.iddocumento = dbo.inv_documentoalm.iddocumento INNER JOIN dbo.inv_documento ON
       dbo.inv_documentoentidad.iddocumento = dbo.inv_documento.iddocumento AND
       dbo.inv_documentoalm.iddocumento = dbo.inv_documento.iddocumento INNER JOIN dbo.inv_concepto ON
       dbo.inv_documento.idconcepto = dbo.inv_concepto.idconcepto INNER JOIN dbo.inv_movimiento ON
       dbo.inv_documento.iddocumento = dbo.inv_movimiento.iddocumento ON
       dbo.gen_producto.idproducto = dbo.inv_movimiento.idproducto INNER JOIN dbo.inv_documentoserv ON
       dbo.inv_documento.iddocumento = dbo.inv_documentoserv.iddocumento INNER JOIN dbo.inv_ServAsoc ON
       dbo.inv_documentoserv.idservicio = dbo.inv_ServAsoc.idServicio INNER JOIN dbo.inv_regdocum ON
       dbo.inv_concepto.idregdocum = dbo.inv_regdocum.idregdocum
WHERE (dbo.inv_regdocum.descripcion LIKE 'Compra%') AND (dbo.inv_regdocum.Almacen = 1)
       AND NOT EXISTS (SELECT dbo.inv_documentoentidad.identidad, dbo.inv_documentoalm.idalmacen, dbo.inv_documento.idunidad,
       dbo.inv_documento.fecha AS fecha_sqls, dbo.inv_documento.idconcepto, dbo.inv_documento.idservicio,
       dbo.inv_documentoserv.identidad AS id_enterv, dbo.inv_lotefechavenc.fechavencimiento,
       dbo.inv_movimiento.idproducto, dbo.inv_concepto.idcategoria, dbo.inv_movimiento.idmedida,
       dbo.inv_movimiento.cantidad, dbo.inv_movimiento.existencia, dbo.inv_documentoserv.importe AS impServ,
       dbo.inv_movimiento.importe AS impEnt
       FROM dbo.inv_lote INNER JOIN dbo.inv_lotefechavenc ON
       dbo.inv_lote.idlote = dbo.inv_lotefechavenc.idlote INNER JOIN dbo.gen_producto INNER JOIN
       dbo.inv_documentoentidad INNER JOIN dbo.inv_documentoalm ON
       dbo.inv_documentoentidad.iddocumento = dbo.inv_documentoalm.iddocumento INNER JOIN
       dbo.inv_documento ON dbo.inv_documentoentidad.iddocumento = dbo.inv_documento.iddocumento AND
       dbo.inv_documentoalm.iddocumento = dbo.inv_documento.iddocumento INNER JOIN dbo.inv_concepto ON
       dbo.inv_documento.idconcepto = dbo.inv_concepto.idconcepto INNER JOIN dbo.inv_movimiento ON
       dbo.inv_documento.iddocumento = dbo.inv_movimiento.iddocumento ON
       dbo.gen_producto.idproducto = dbo.inv_movimiento.idproducto INNER JOIN dbo.inv_documentoserv ON
       dbo.inv_documento.iddocumento = dbo.inv_documentoserv.iddocumento INNER JOIN dbo.inv_ServAsoc ON
       dbo.inv_documentoserv.idservicio = dbo.inv_ServAsoc.idServicio ON
       dbo.inv_lote.idproducto = dbo.gen_producto.idproducto INNER JOIN dbo.inv_regdocum ON
       dbo.inv_concepto.idregdocum = dbo.inv_regdocum.idregdocum
WHERE (dbo.inv_regdocum.descripcion LIKE 'Compra%') AND (dbo.inv_regdocum.Almacen = 1))

```

Figura 3.5. Selección de los productos que no tienen fecha de vencimiento asociada

```

SELECT dbo.inv_documentoentidad.identidad, dbo.inv_documentoalm.idalmacen, dbo.inv_documento.fecha AS fecha_sqls,
       dbo.inv_documento.idunidad, dbo.inv_documento.idconcepto, dbo.inv_concepto.idcategoria, dbo.inv_movimiento.idmedida,
       dbo.inv_movimiento.idproducto, dbo.inv_documentoserv.identidad AS id_enterv, dbo.inv_documentoserv.idservicio,
       dbo.inv_lotefechavenc.fechavencimiento, dbo.inv_movimiento.cantidad, dbo.inv_movimiento.importe AS impEnt,
       dbo.inv_documentoserv.importe AS impServ, dbo.inv_movimiento.existencia
FROM dbo.inv_lote INNER JOIN dbo.inv_lotefechavenc ON dbo.inv_lote.idlote = dbo.inv_lotefechavenc.idlote INNER JOIN
       dbo.gen_producto INNER JOIN dbo.inv_documentoentidad INNER JOIN dbo.inv_documentoalm ON
       dbo.inv_documentoentidad.iddocumento = dbo.inv_documentoalm.iddocumento INNER JOIN dbo.inv_documento ON
       dbo.inv_documentoentidad.iddocumento = dbo.inv_documento.iddocumento AND
       dbo.inv_documentoalm.iddocumento = dbo.inv_documento.iddocumento INNER JOIN dbo.inv_concepto ON
       dbo.inv_documento.idconcepto = dbo.inv_concepto.idconcepto INNER JOIN dbo.inv_movimiento ON
       dbo.inv_documento.iddocumento = dbo.inv_movimiento.iddocumento ON
       dbo.gen_producto.idproducto = dbo.inv_movimiento.idproducto INNER JOIN dbo.inv_documentoserv ON
       dbo.inv_documento.iddocumento = dbo.inv_documentoserv.iddocumento INNER JOIN dbo.inv_ServAsoc ON
       dbo.inv_documentoserv.idservicio = dbo.inv_ServAsoc.idServicio ON
       dbo.inv_lote.idproducto = dbo.gen_producto.idproducto INNER JOIN dbo.inv_regdocum ON
       dbo.inv_concepto.idregdocum = dbo.inv_regdocum.idregdocum
WHERE (dbo.inv_regdocum.descripcion LIKE 'Compra%') AND (dbo.inv_regdocum.Almacen = 1)

```

Figura 3.6. Selección de los productos que tienen fecha de vencimiento asociada

Similar al epígrafe anterior, en la Figura 3.7 se muestra el proceso ETL para el subproceso de compras periódico. En este caso se selecciona la información que se encuentra en la tabla de hechos correspondiente a este subproceso pero a nivel de transacción, vista anteriormente, y se realizan las agregaciones necesarias para luego realizar la carga de los datos en la tabla destino de una forma más resumida y compactada.

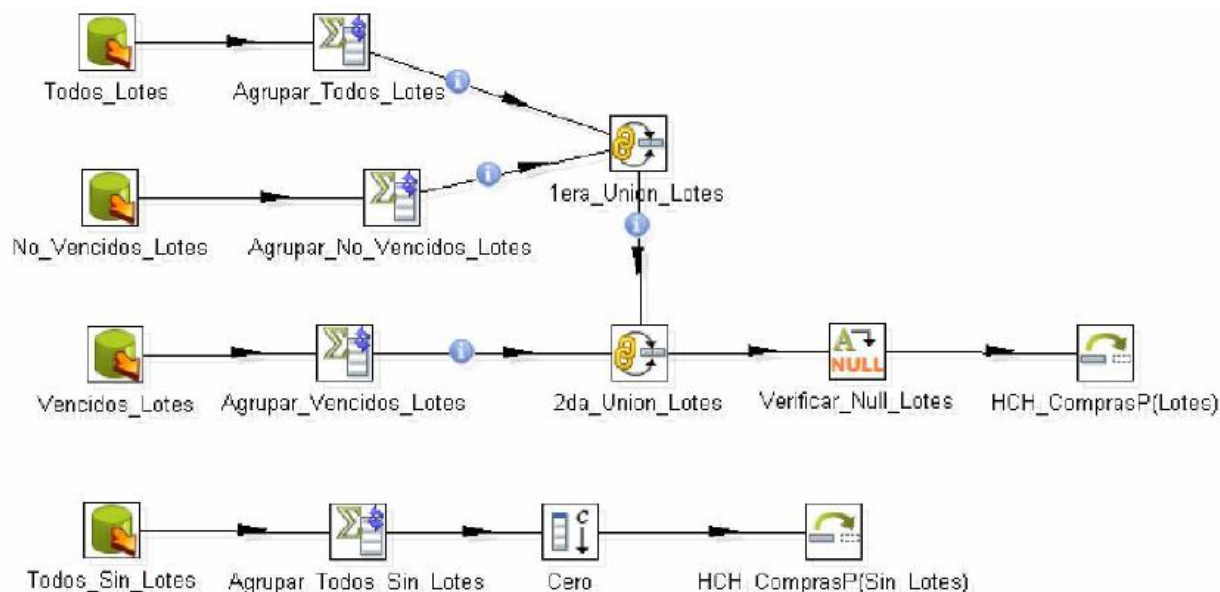


Figura 3.7. Proceso ETL para el subproceso de compras a nivel periódico

Los procesos de extracción, transformación y carga, del resto de los subprocesos involucrados en este Mercado de Datos, es similar a los descritos anteriormente, teniendo en cuenta las particularidades y detalles de cada uno.

Luego de haber realizado las transformaciones a los datos en cada subproceso, es necesario organizar el orden de las cargas de las tablas de hechos. El trabajo o *job* define el orden en que se van a ejecutar las transformaciones, el horario y la frecuencia de las cargas. La Figura 3.8 muestra el trabajo que resume el proceso general de carga del Mercado de Datos.

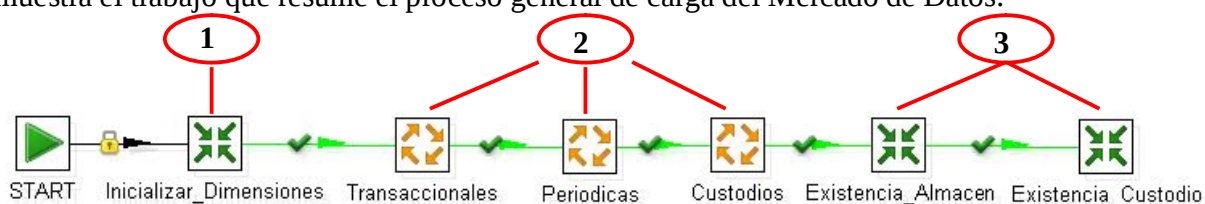


Figura 3.8. Proceso de carga del Mercado de Datos

1. En este paso se ejecuta una transformación para poblar las tablas correspondientes a las dimensiones.
2. En estos pasos se desarrollan tres trabajos, los cuales a su vez ejecutan un conjunto de transformaciones para poblar las tablas de hechos relacionadas con los subprocesos a nivel de almacenes (Transaccionales y Periódicas) y de gestión de custodio (Custodios). Destacar que fue necesario establecer un orden entre los trabajos

relacionados con los subprocesos a nivel de almacén, debido a la dependencia que existe entre las tablas de hechos transaccionales y periódicas.

3. En estos pasos se ejecutaron las transformaciones necesarias para poblar las tablas de hechos consolidadas relacionadas con los almacenes y custodios.

Notar que la transformación de los datos, se realiza de acuerdo a las necesidades de los usuarios; tales como cambios de formato, la cual asegura la unicidad y estandarización de los tipos de datos, por ejemplo la fecha, deberá ser tratada como un solo tipo de dato y siempre de la misma forma. La duplicidad de datos es otro tipo de transformación, donde se debe evitar que la información proveniente de una o varias fuentes de datos se multiplique a la hora de su unión o juntura. Otros tipos de transformaciones son los agregados y los campos integrados. La extracción, transformación y carga de los datos en este sistema en particular no va a ser de la envergadura que normalmente estos procesos conllevan debido a que la información histórica se encuentra con un alto grado de limpieza, estandarización y calidad. Sólo habría que centrar los esfuerzos en dos áreas específicas: en la extracción de los datos de la fuente y en la carga hacia el Mercado de Datos. Vale aclarar que no se descarta la posibilidad que aparezca algún tipo de transformación sencilla, principalmente de formato y conversión, que se requiera hacer pero no sería el grueso para la realización del proceso.

3.4.2. Implementación de los cubos de datos.

El éxito final del Mercado de Datos depende en gran medida de la presentación de la capa de visualización y de la forma en que se organiza la información estructural.

Uno de los elementos fundamentales donde se organiza la información son los cubos multidimensionales. Estos se realizan utilizando la herramienta Pentaho Schema Workbench que genera un fichero de configuración XML. En este fichero de esquema se pueden definir las dimensiones, los niveles de jerarquía de dimensiones y los hechos.

En este Mercado de Datos se modelaron 26 cubos multidimensionales, los cuales están relacionados con 11 dimensiones conformadas. La Figura 3.9 muestra el cubo de datos correspondiente al subproceso de compras a nivel de transacción, formado por 11 dimensiones y cuatro medidas, de las cuales las tres primeras son aditivas y la última es semiaditiva. De igual forma la Figura 3.10 muestra el cubo de datos correspondiente al subproceso de compras a nivel periódico formado por siete dimensiones y cinco medidas, las cuales son aditivas.

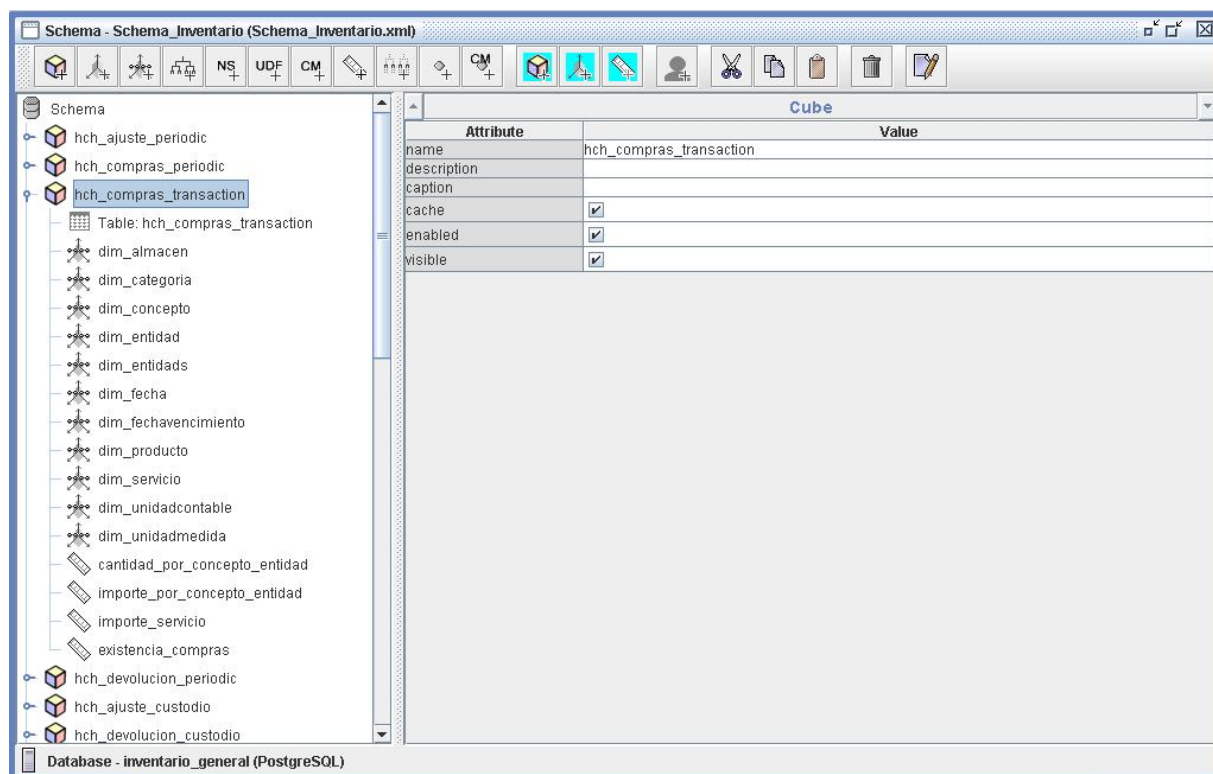


Figura 3.9. Cubo de datos de las compras transaccional

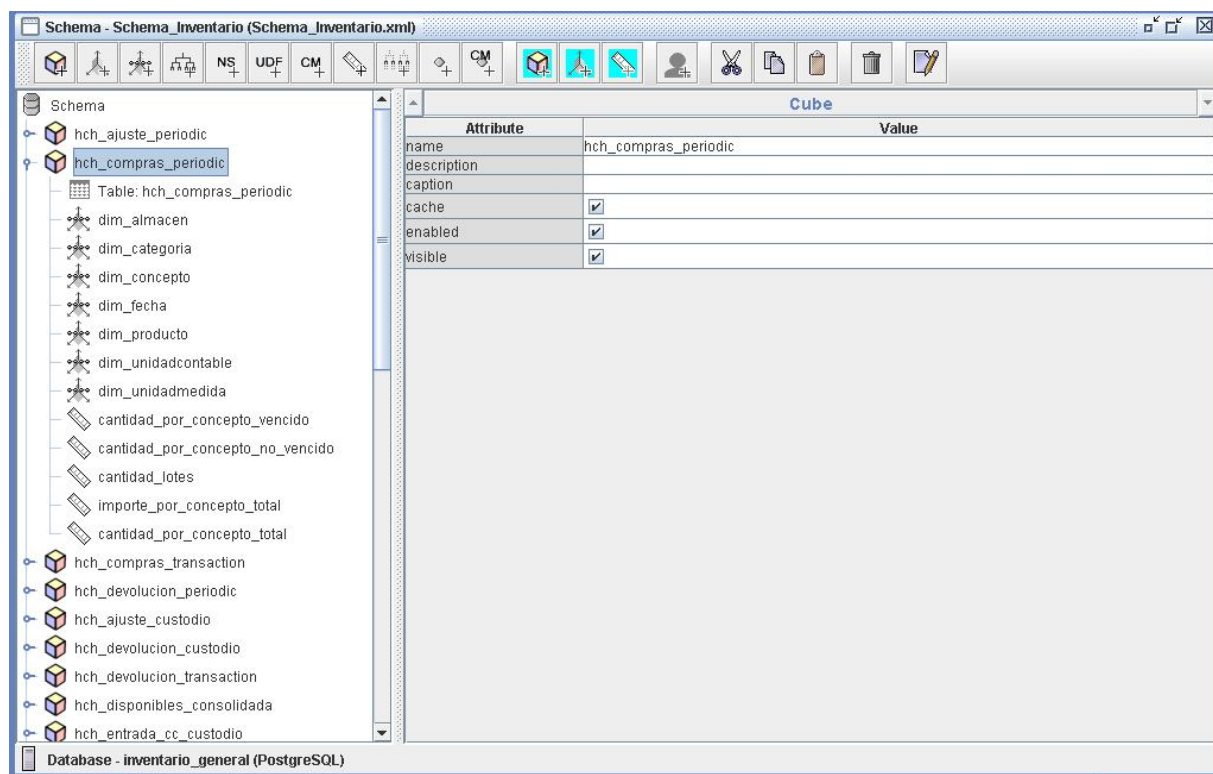


Figura 3.10. Cubo de datos de las compras periódicas

3.5. Reportes representativos

Los reportes representan la información que el cliente desea que se muestre como finalidad del producto. Los mismos fueron seleccionados luego de realizar un análisis del caso de estudio propuesto en el Capítulo 1, epígrafe 1.2 de la presente tesis, donde se recoge toda la información referente al subsistema de inventario del VERSAT Sarasola. A continuación se muestran algunos reportes representativos que fueron identificados.

- Reportes de las compras a nivel de transacción
 - o ¿Cómo se comportaron las compras de productos en cada almacén por entidades durante los años 2009 y 2010?
 - o Importe total en servicios de las compras realizadas durante un período determinado.
 - o ¿Qué entidades han prestado servicio y cuáles servicios en un período determinado.
- Reportes de las compras a nivel periódico
 - o Importe total de las compras realizadas durante un periodo determinado
 - o ¿Cuántos productos he comprado en un período determinado y en cuántos lotes?
 - o ¿Cuántos productos vencidos se han comprado por almacenes y por categorías?
- Reportes a nivel de gestión de custodio.
 - o Importe total de las transferencias realizadas durante un periodo determinado por custodios
 - o ¿Cuántos productos entraron con el 50% para gasto por custodio y por unidad contable?

De igual forma se pueden hacer reportes similares relacionados con los restantes subprocesos. Para ver los resultados de los reportes expuestos anteriormente se utilizaron las siguientes herramientas de la Suite de Pentaho:

1. Pentaho Report Designer: Editor con capacidad de personalización de informes útil a los desarrolladores. Incluye asistentes para facilitar la configuración de propiedades. Incluye un editor de consultas para facilitar la confección de los datos que se usan en los informes.

2. Pentaho Report Design Wizard: Herramienta de diseño de informes. Está destinada a usuarios con menos conocimientos técnicos.
3. Web ad-hoc reporting: Es similar a la herramienta anterior pero vía Web. Extiende la capacidad de los usuarios finales para la creación de informes a partir de plantillas pre configuradas y siguiendo un asistente de creación.

3.6. Conclusiones del capítulo

En este capítulo se describió el diseño y la implementación del Mercado de Datos para el proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola. Se detalló la arquitectura, el diseño lógico y físico, la estrategia de extracción, transformación y carga, así como la implementación de los cubos de datos. Todo esto posibilitó llegar a las siguientes conclusiones:

1. La metodología propuesta por R. Kimball, facilitó el proceso de determinar e identificar las principales dimensiones, hechos y granularidad del proceso de inventario.
2. El diseño del modelo dimensional del Mercado de Datos presenta:
 - a. 11 dimensiones conformadas
 - b. 26 tablas de hechos:
 - i. Ocho transaccionales
 - ii. Dos consolidadas
 - iii. 16 periódicas:
 - Nueve a nivel de almacén
 - Siete a nivel de gestión de custodio
3. Se definió la arquitectura de integración del sistema dando paso a la implementación del proceso de integración, lográndose la carga exitosa de los datos.
4. Como parte de la estructura OLAP del sistema se diseñaron 26 cubos dimensionales relacionados con 11 dimensiones, los cuales dan paso a la organización de la información estructuralmente.

CONCLUSIONES

Al finalizar el trabajo de diploma “Mercado de Datos para el proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola” se obtiene.

1. Una descripción del caso de estudio: proceso de inventario del Sistema de Gestión Contable VERSAT Sarasola.
2. Un resumen de la solución a un caso de estudio descrito por Kimball como un referente para tener en cuenta en la solución a un problema similar.
3. El diseño del Mercado de Datos mediante el uso de la herramienta ERStudio en su versión 8.0
4. La implementación de los procesos de ETL mediante el uso de la herramienta Pentaho Data Integration.
5. Se hizo una comprobación del Mercado de Datos utilizando datos reales tomados de la empresa ULAEX, SA.

RECOMENDACIONES

- Extender el trabajo hacia los otros subsistemas del VERSAT Sarasola y de esta forma lograr la elaboración de un Almacén de Datos para este sistema de gestión contable, lo cual permite afianzar la posición líder de la aplicación en el país.

BIBLIOGRAFÍA

Finanzas al día [Online]. Available:

http://www.egrafip.cu/index.php?option=com_content&task=category§ionid=47&id=92&Itemid=46 [Accessed 25 de marzo del 2013].

AUTORES, C. D. Manual del Usuario y de Explotación VERSAT Sarasola.

AUTORES, C. D. 2005. Sistema Económico Integrado, VERSAT Sarasola. *VII Evento Nacional de la Mujer Economista y Contadora de Cuba*. Camagüey, Cuba.

BONIFATI, A., CATTANEO, F., CERI, S., FUGGETTA, A. & PARABOSCHI, S. 2001. Designing Data Marts for Data Warehouses. *ACM Transactions on Software Engineering and Methodology*, Vol. 10, No. 4, Pages 452–483.

BOUMAN, R. & DONGEN, J. V. 2009. *Pentaho Solutions: Business Intelligence and Data Warehousing with Pentaho and MySQL*. Indianapolis, Indiana, Wiley Publishing, Inc.

CASTERS, M., BOUMAN, R. & DONGEN, J. V. 2010. Pentaho® Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration. Indianapolis, Indiana: Wiley Publishing, Inc.

COLLIAT, G. 1996. OLAP, Relational and Multidimensional Database Systems. *SIGMOND Record*, 25.

CUBANA, N. (ed.) 1988. *Sistema Internacional de Unidades*: Editorial Pueblo y Educación.

CHAUDHURI, S. & DAYAL, U. An Overview of Data Warehousing and OLAP Technology.

CHINCHILLA ARLEY, R. 2011. Mercado de datos: conceptos y metodologías de desarrollo. *Tecnología en Marcha*, Vol. 24, N.º 3, Pag. 55-66.

DATAZUCAR, V. C. 2002. VERSAT Sarasola. *IV Feria Internacional de la Agroindustria Azucarera. International Sugar Agro-Industry Fair*. La Habana, Cuba.

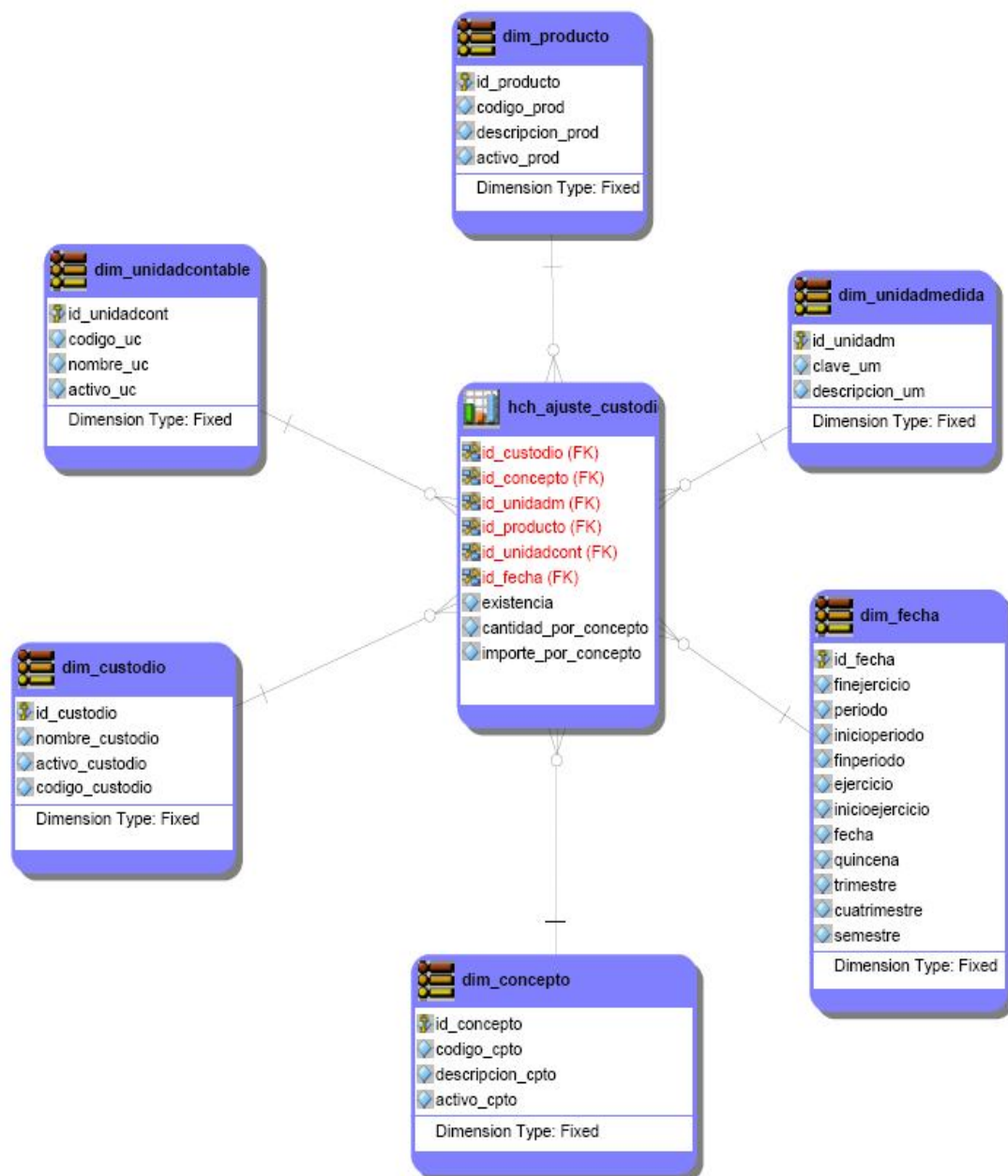
DOLK, D. R. 2000. Integrated model management in the data warehouse era. *European Journal of Operational Research*, 122, 199-218.

GARCÍA, A. B., BAS, A. O., SALORT, E. V. & GONZALEZ, L. C. 2004. Data Warehouse para la gestión por procesos en el sistema productivo. *Second World Conference on POM and 15th Annual POM Conference*. Cancun, Mexico.

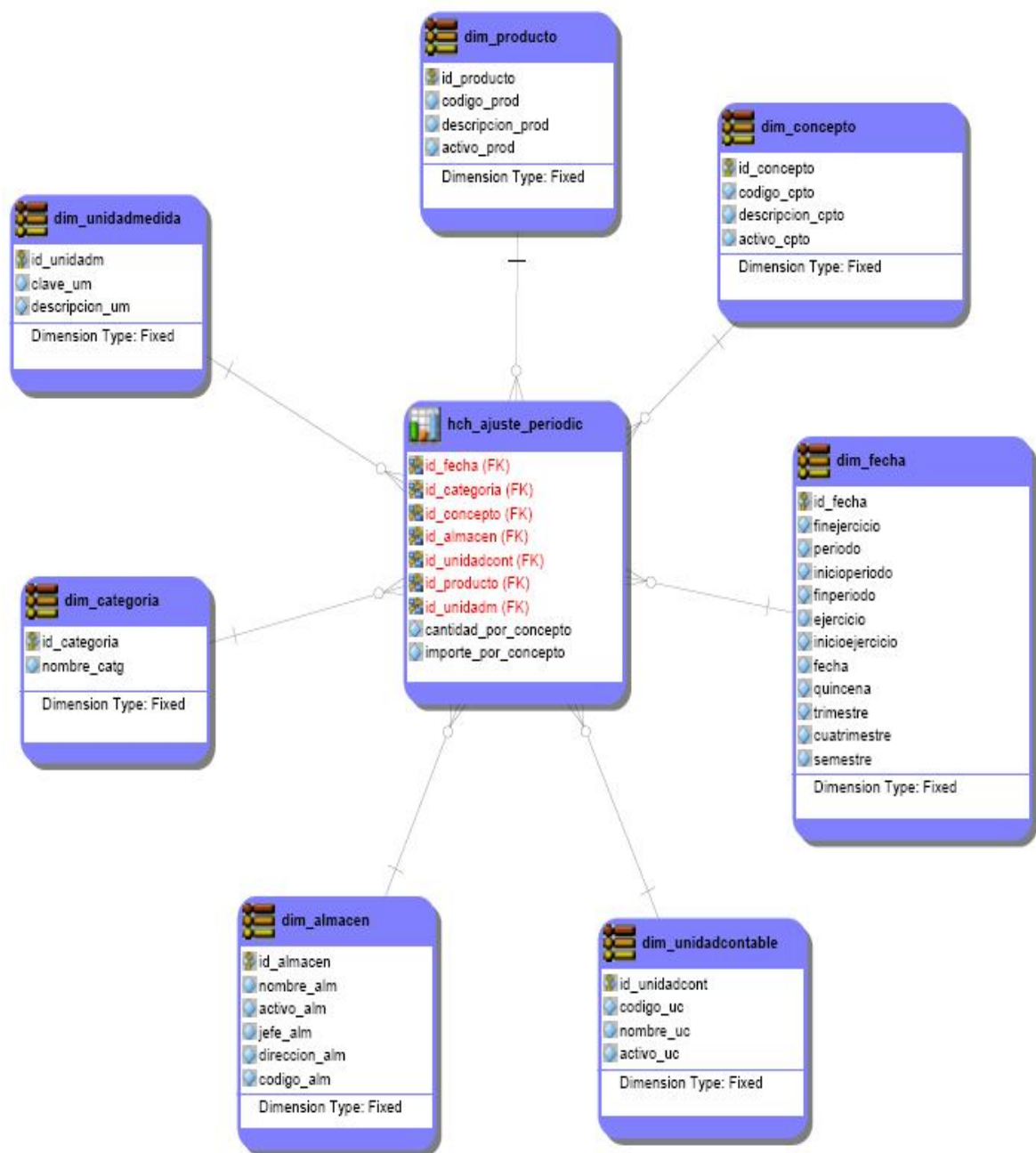
IMHOFF, C., GALEMMO, N. & GEIGER., J. G. 2003. Mastering Data Warehouse Desing: Relational and Dimensional Techniques. In: ELLIOTT, R. M. (ed.). Indianapolis, Indiana: Wiley Publishing, Inc.

- INMON, W. H. 2002. *Building The Data Warehouse*. 3rd ed, Canada, John Wiley & Sons, Inc.
- JARKE, M., LIST, T. & KÖLLER, J. 2000. The Challenge of Process Data Warehousing. *26th International Conference on Very Large Databases*. Cairo, Egypt.
- KIM, W. 2002. On Three Major Holes in Data Warehousing Today. *JOURNAL OF OBJECT TECHNOLOGY*, Vol. 1, no. 4, pages 39-47.
- KIMBALL, R. & CASERTA, J. 2004. *The data warehouse ETL toolkit : practical techniques for extracting, cleaning, conforming, and delivering data*. Canada, Wiley Publishing, Inc.
- KIMBALL, R. & ROSS, M. 2002. *The data warehouse toolkit : the complete guide to dimensional modeling*. 2nd ed, New York, John Wiley and Sons, Inc.
- KIMBALL, R., ROSS, M., THORNTHWAITE, W. & REEVES, L. 1998. *The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing and Deploying Data Warehouses*. New York, John Wiley & Sons, Inc.
- MURTAZA, A. H. 2000. A framework for Developing an Enterprise Data Warehousing Solution. *Enterprise System Integration*. Boca Raton, Florida.
- NADER, J. 2003. *Sistema de Apoyo Gerencial Universitario*.
- POE, V., KLAUER, P. & BROBST, S. 1998. *Building a Data Warehouse for Decision Support*. Prentice Hall.
- PONNIAH, P. 2001 *Data Warehousing Fundamentals: A Comprehensive Guide for IT Professionals.*, New York, John Wiley & Sons, Inc.
- POOLE, J., CHANG, D., TOLBERT, D. & MELLOR, D. 2003. *Common Warehouse Metamodel Developer's Guide*. Indianapolis, Indiana, Wiley Publishing, Inc.
- ROGERS, R. I. M. 2004. *CERTIFICACIÓN DE DEPOSITO LEGAL FACULTATIVO DE OBRAS PROTEGIDAS*. La Habana, Cuba patent application.
- ROSS, M. 2003. No Detail Too Small. Available:
<http://www.kimballgroup.com/category/business-intelligence-and-data-warehouse-articles/no-detail-too-small/> [Accessed 5 de abril del 2013].
- SIMITSIS, A., VASSILIADIS, P. & SELLIS, T. Extraction-Transformation-Loading Processes.
- WOLFF, C. G. 2002. Modelamiento multidimensional.

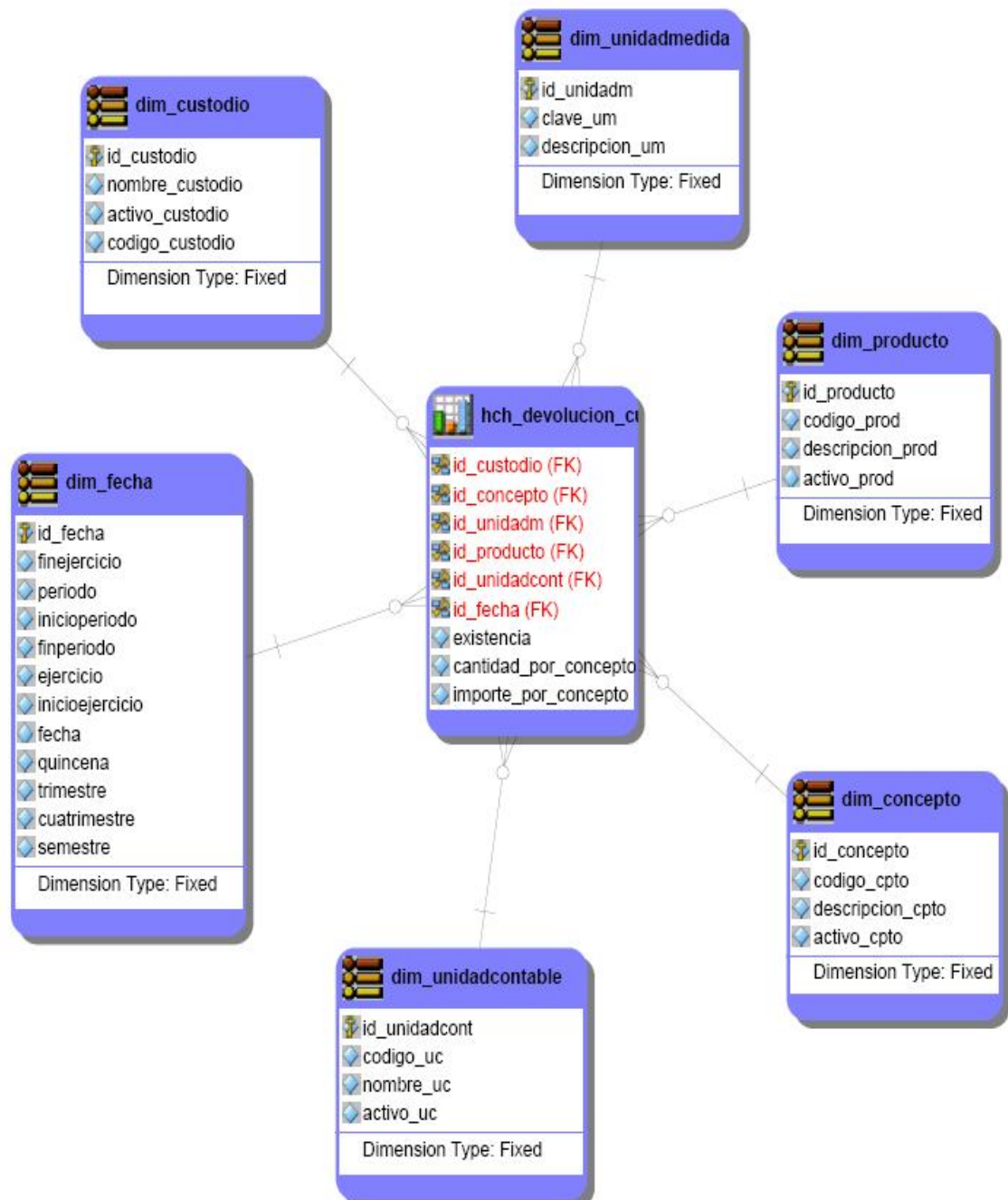
ANEXO I



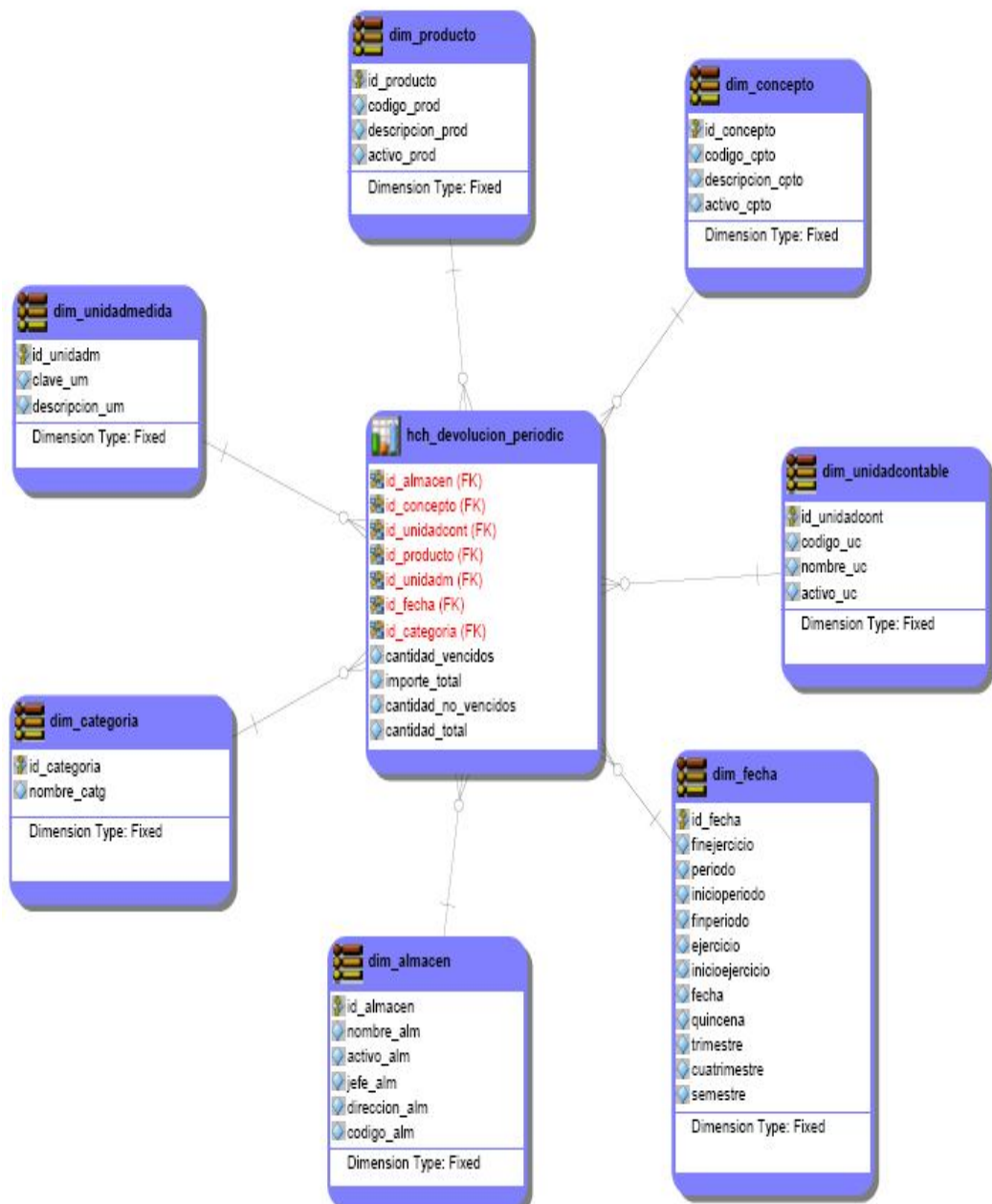
I.1 Esquema estrella del subproceso de ajuste a nivel de gestión de custodio



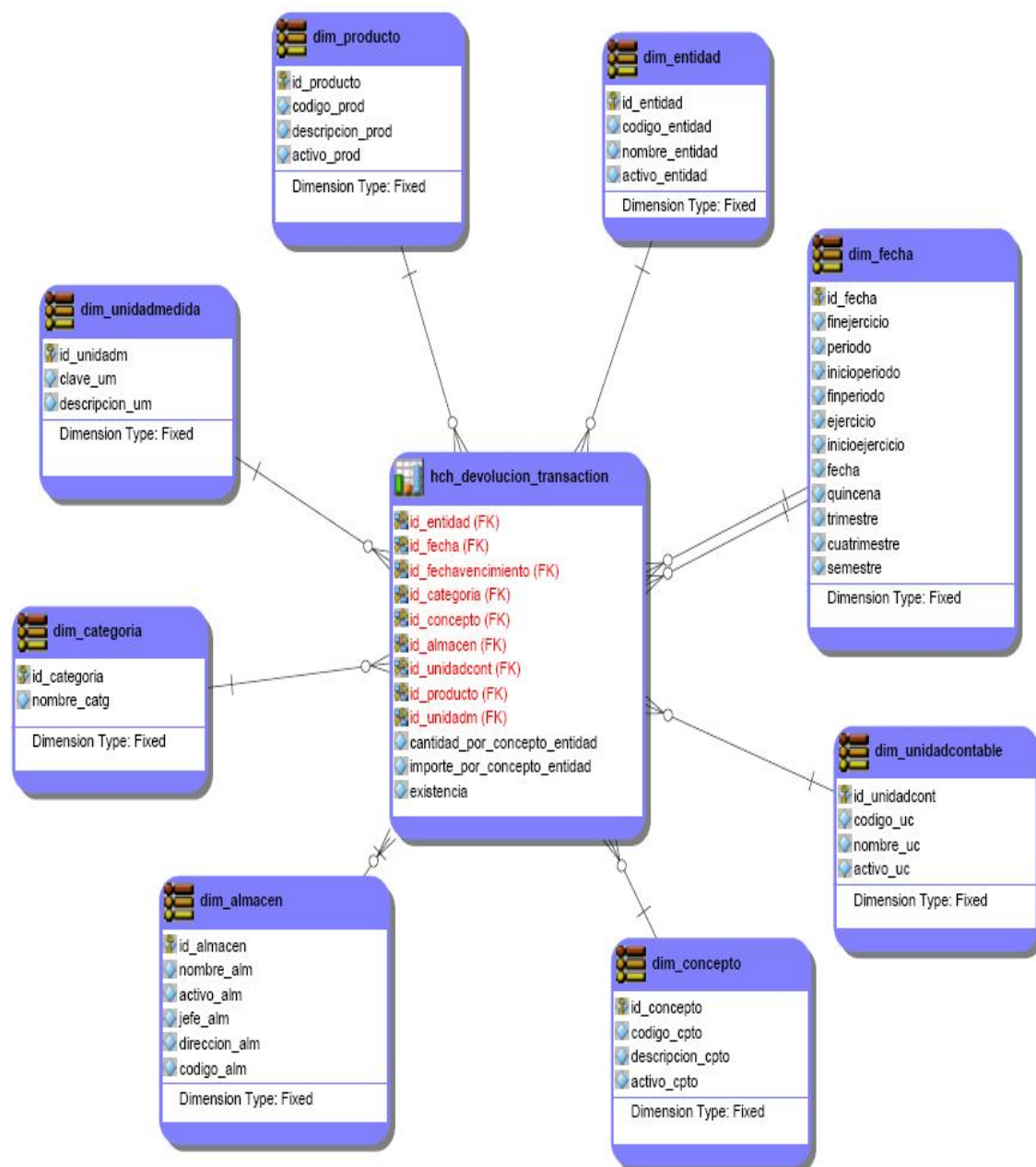
I.2 Esquema estrella del subproceso de ajuste a nivel periódico



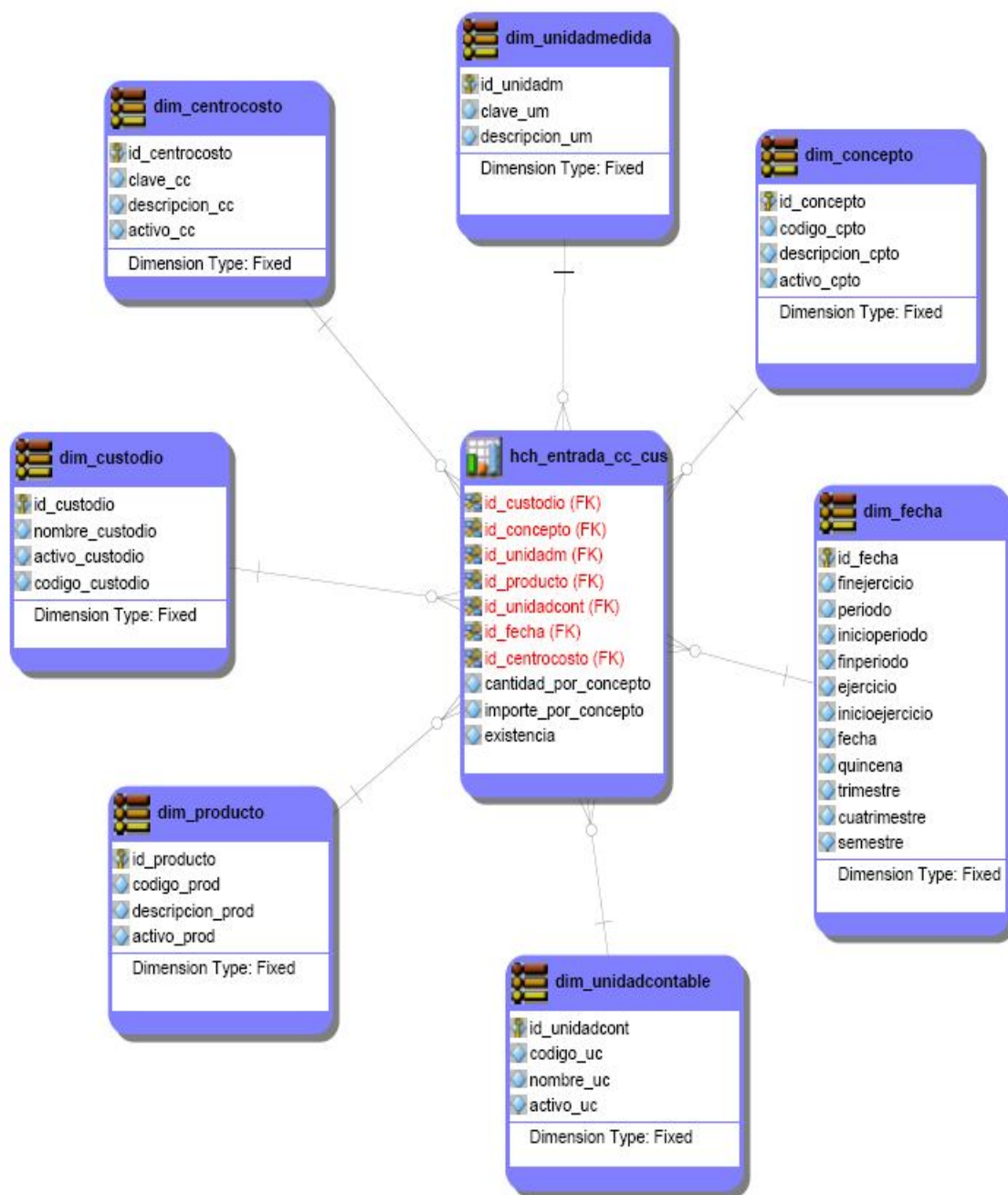
I.3 Esquema estrella del subproceso de devolución a nivel de gestión de custodio



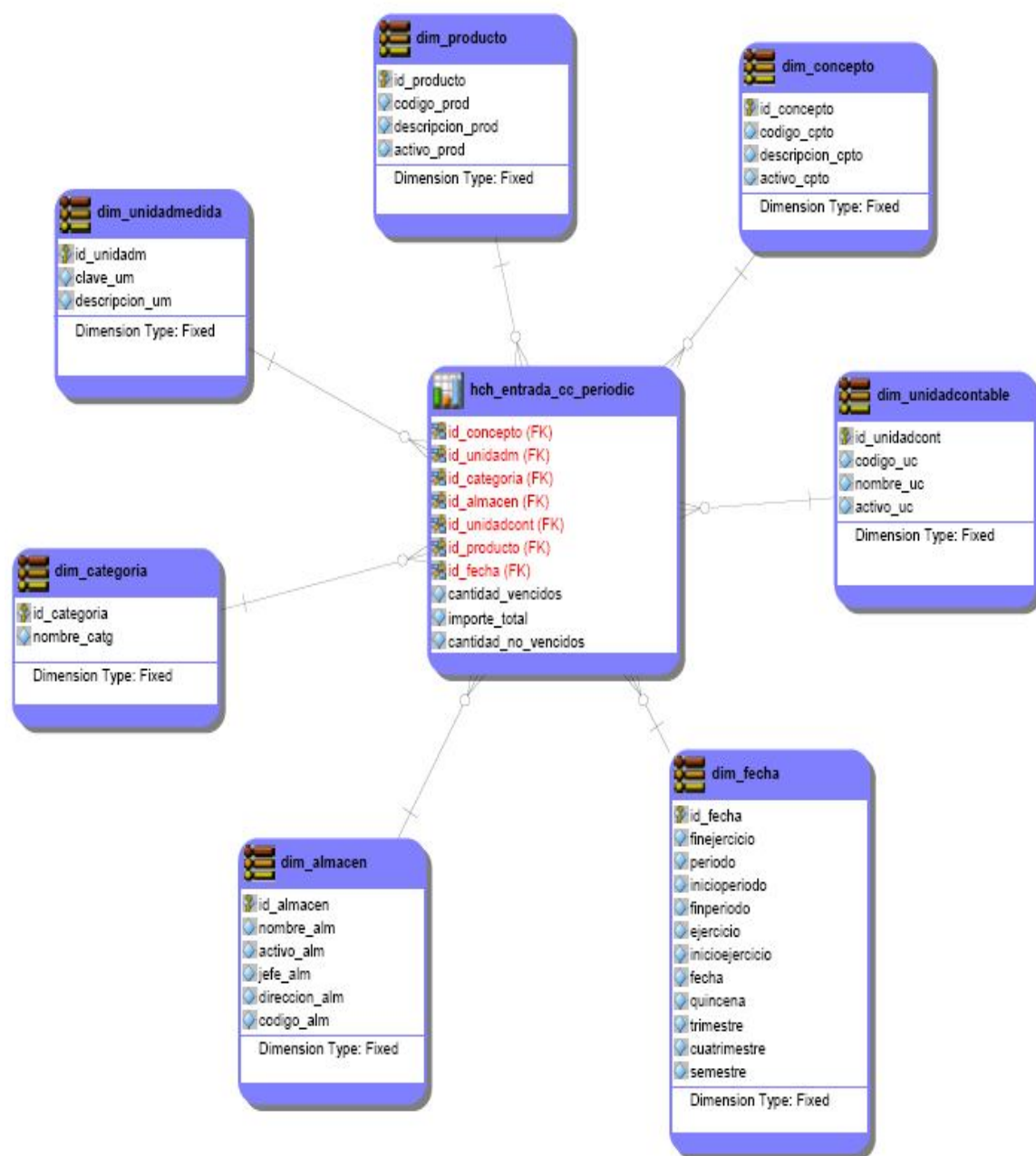
I.4 Esquema estrella del subproceso de devolución a nivel periódico



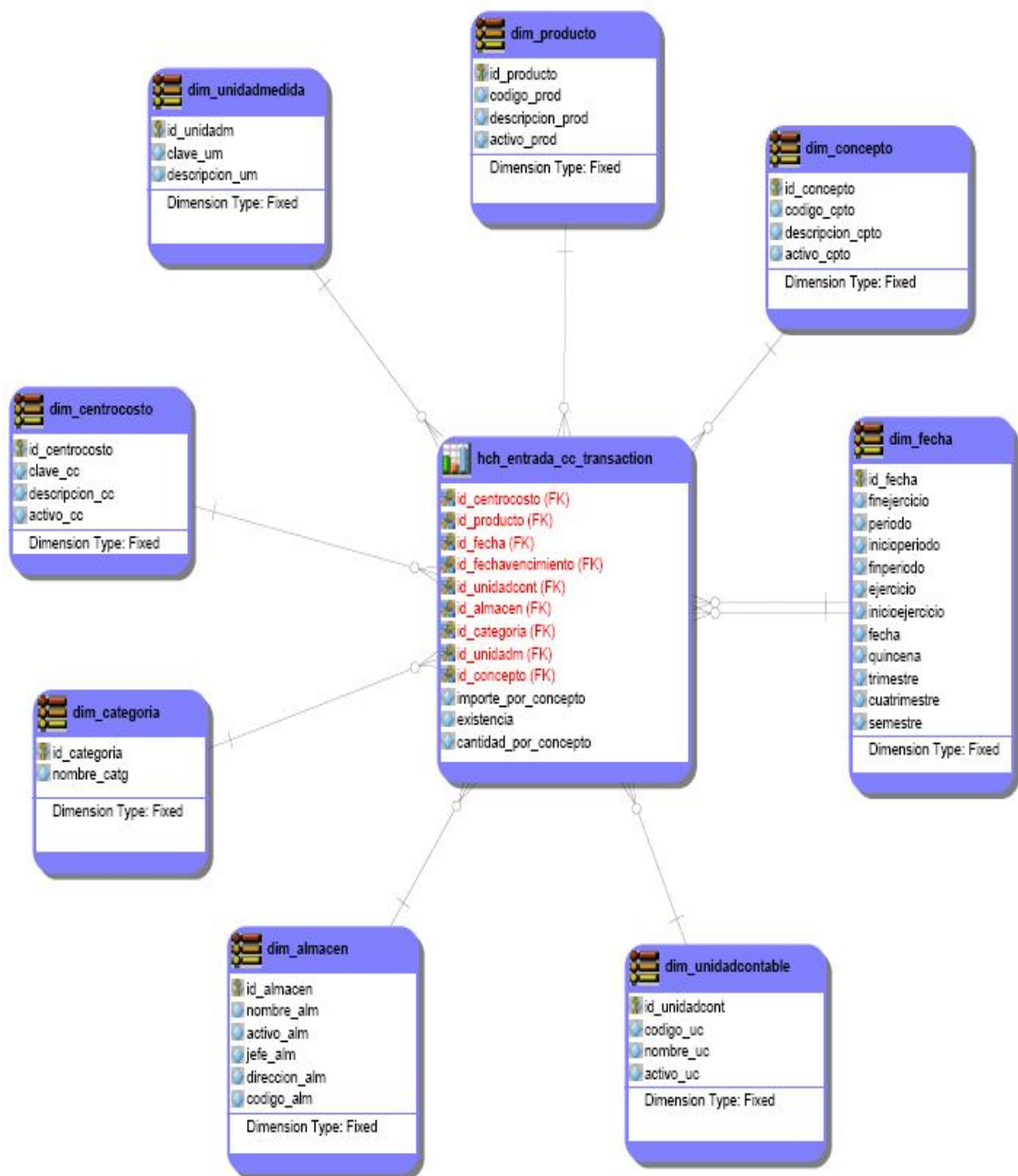
I.5 Esquema estrella del subproceso de devolución a nivel transaccional



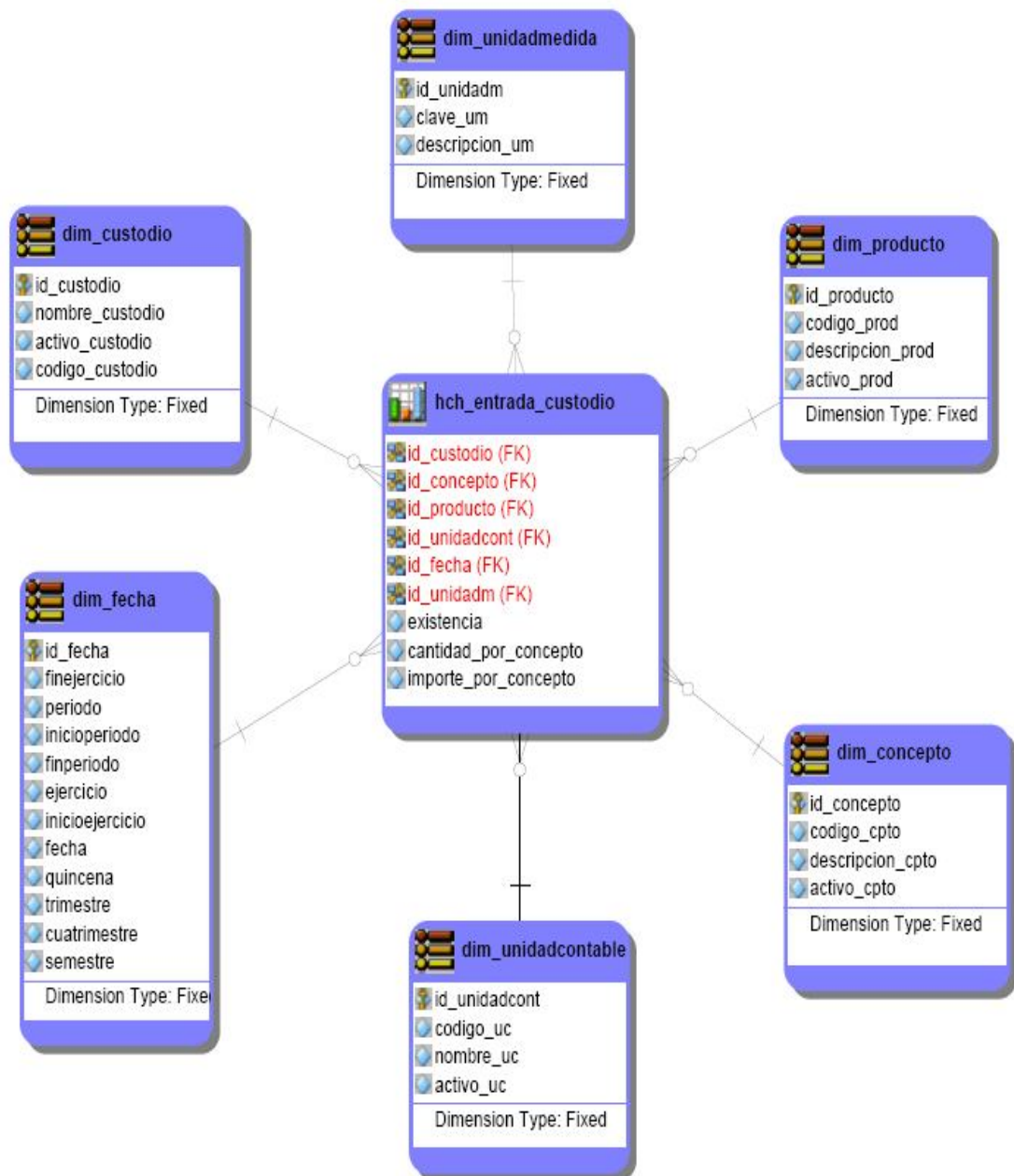
I.6 Esquema estrella del subproceso de entrada CC a nivel de gestión de custodio



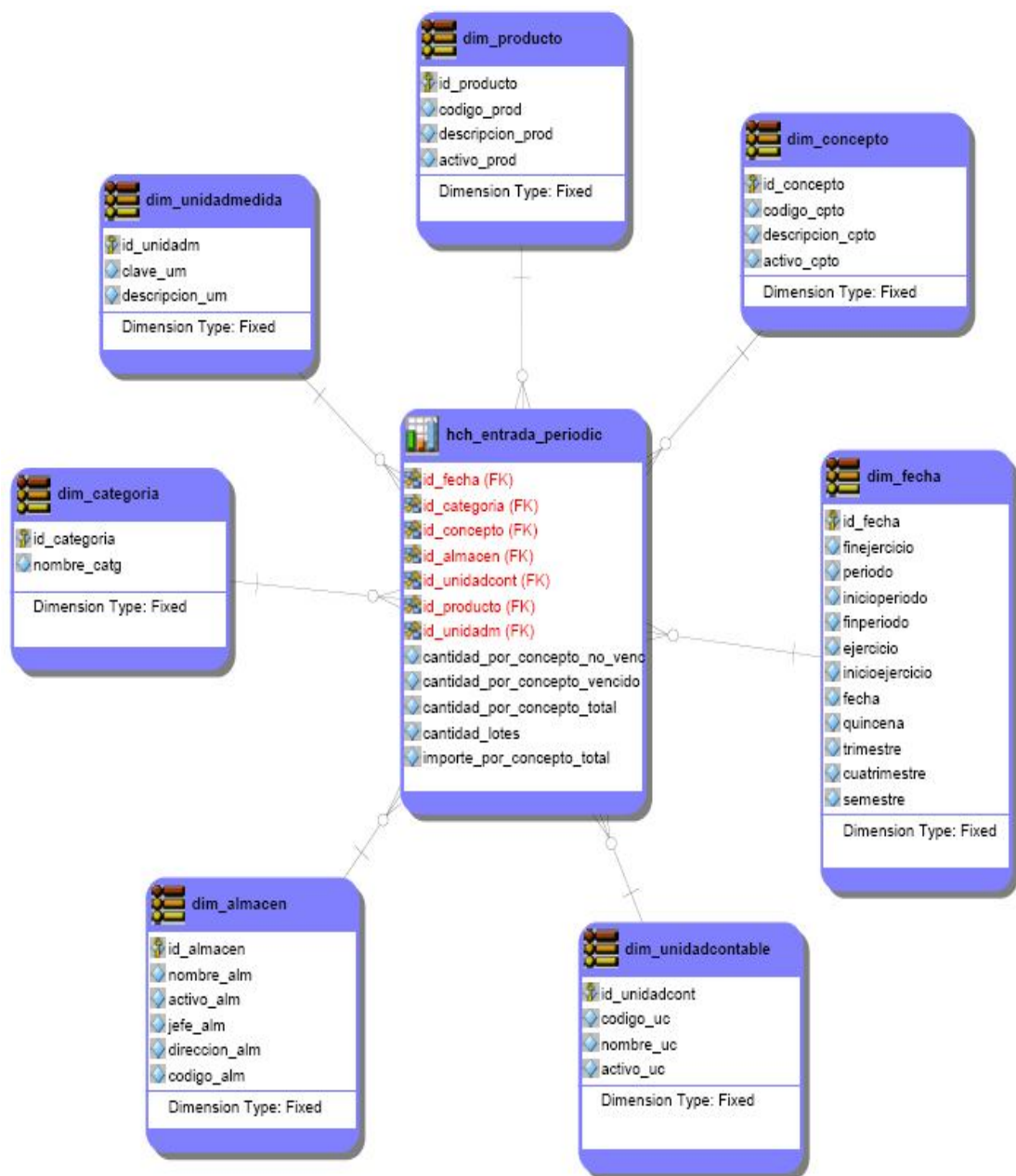
I.7 Esquema estrella del subproceso de entrada CC a nivel periódico



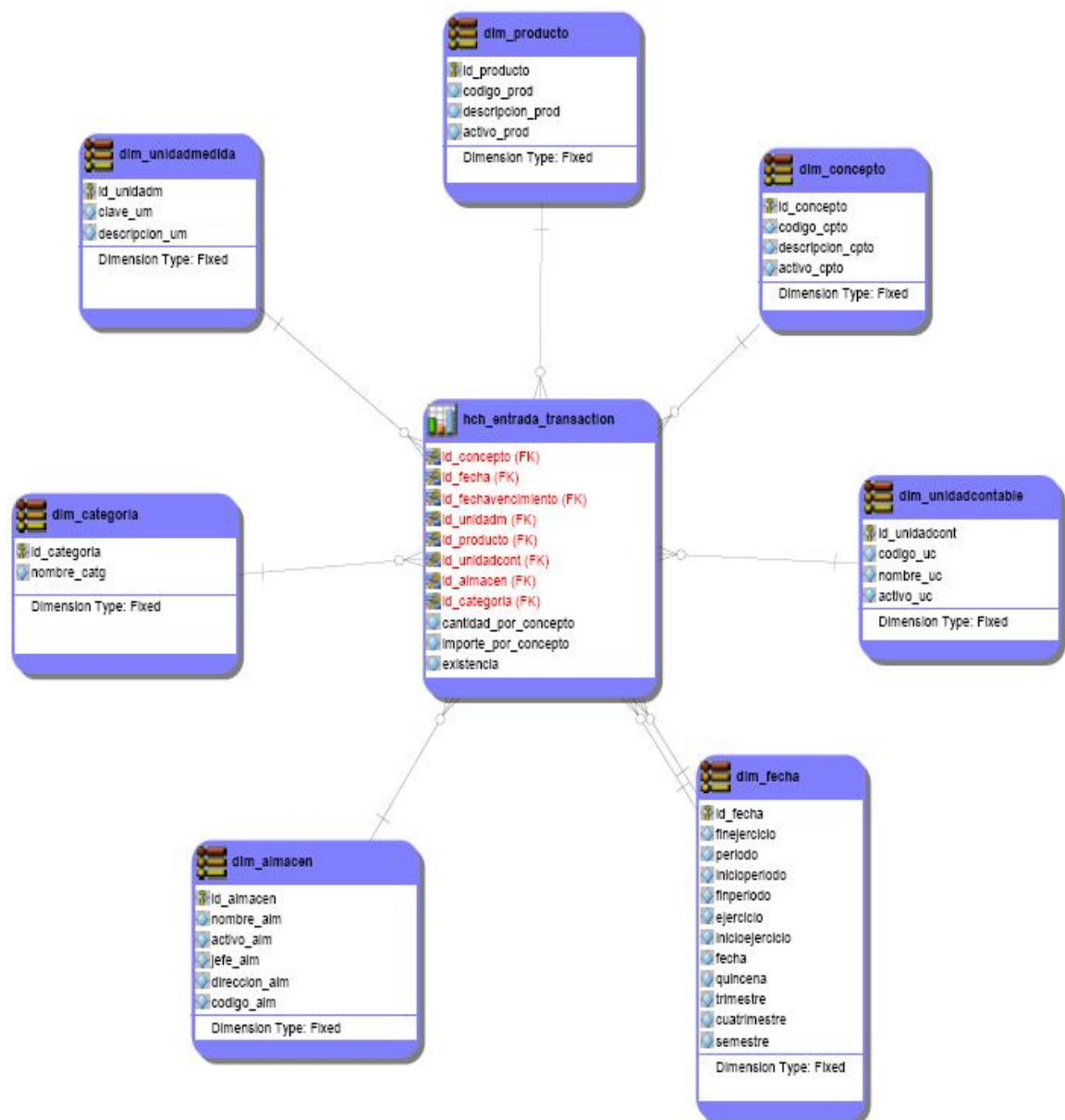
I.8 Esquema estrella del subproceso de entrada CC a nivel transaccional



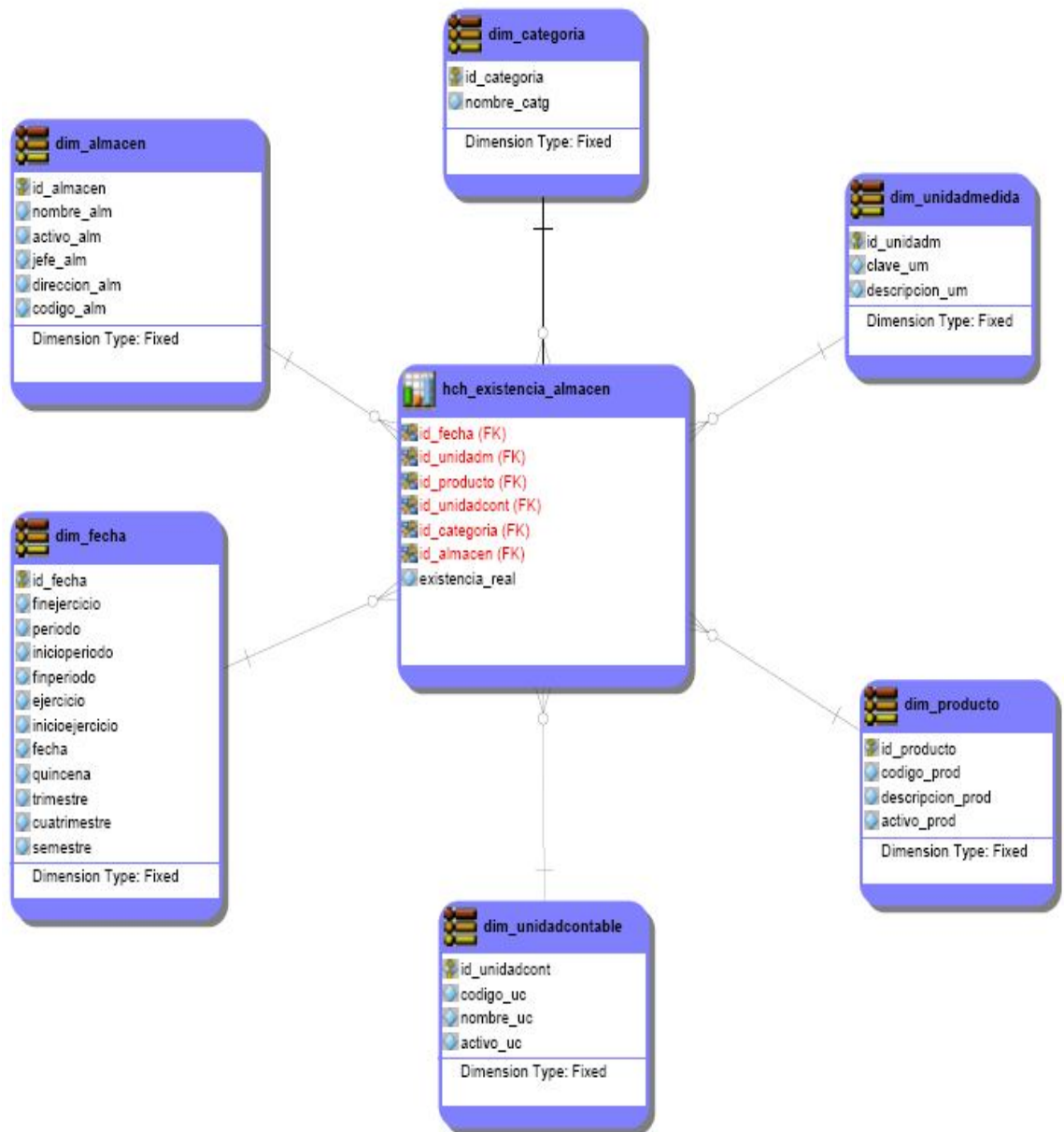
I.9 Esquema estrella del subproceso de entrada a nivel de gestión de custodio



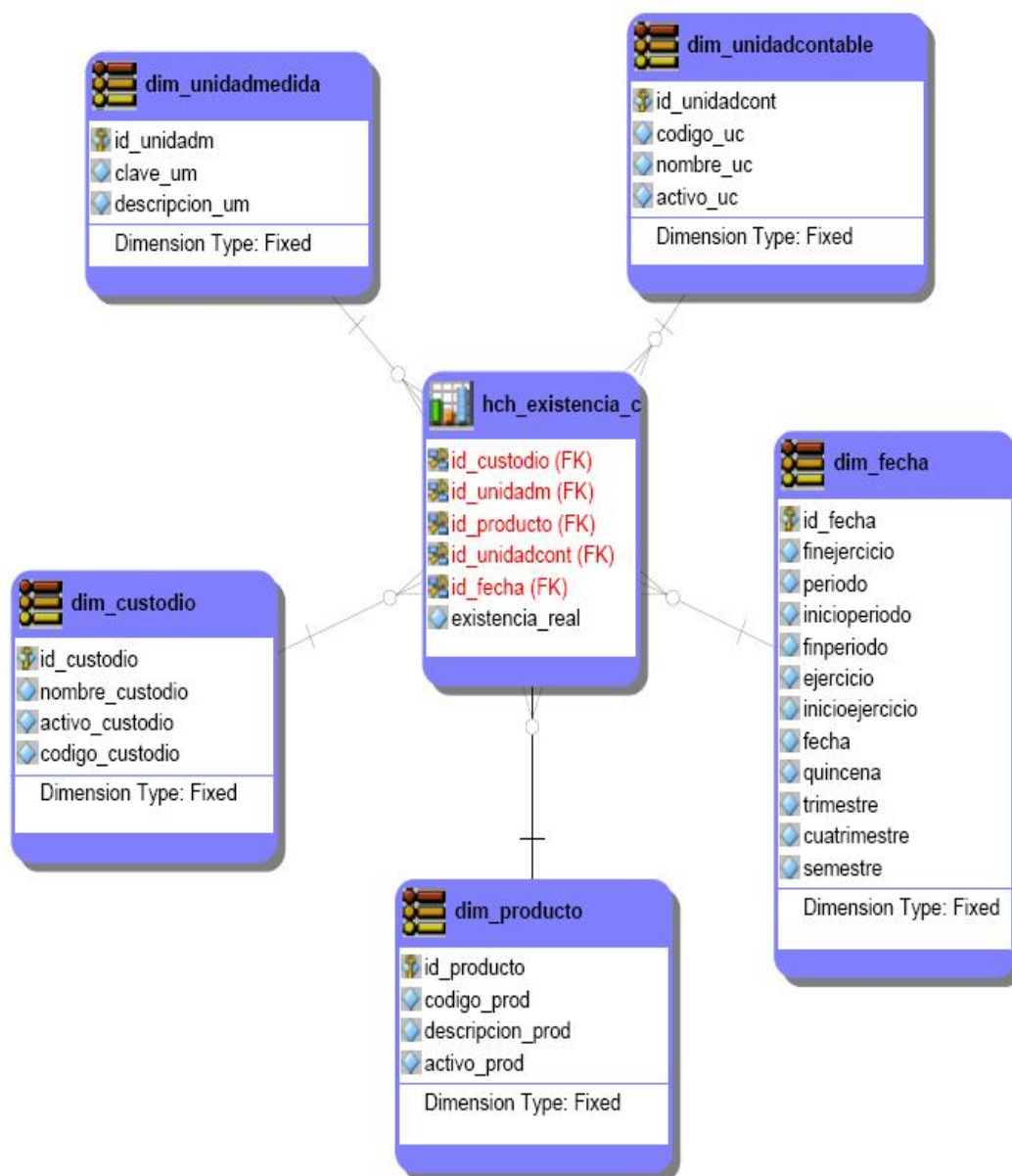
I.10 Esquema estrella del subproceso de entrada a nivel periódico



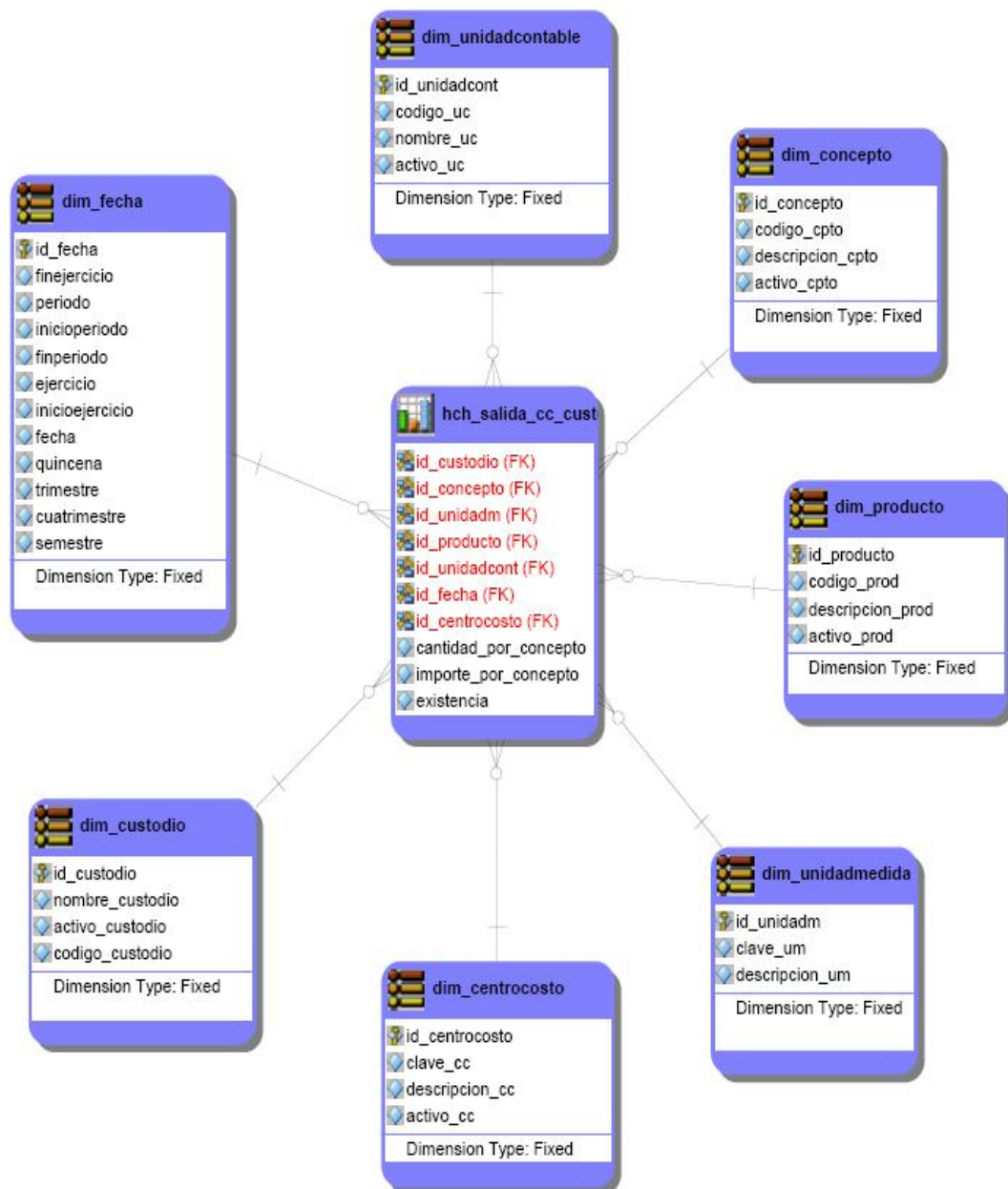
I.11 Esquema estrella del subproceso de entrada a nivel transaccional



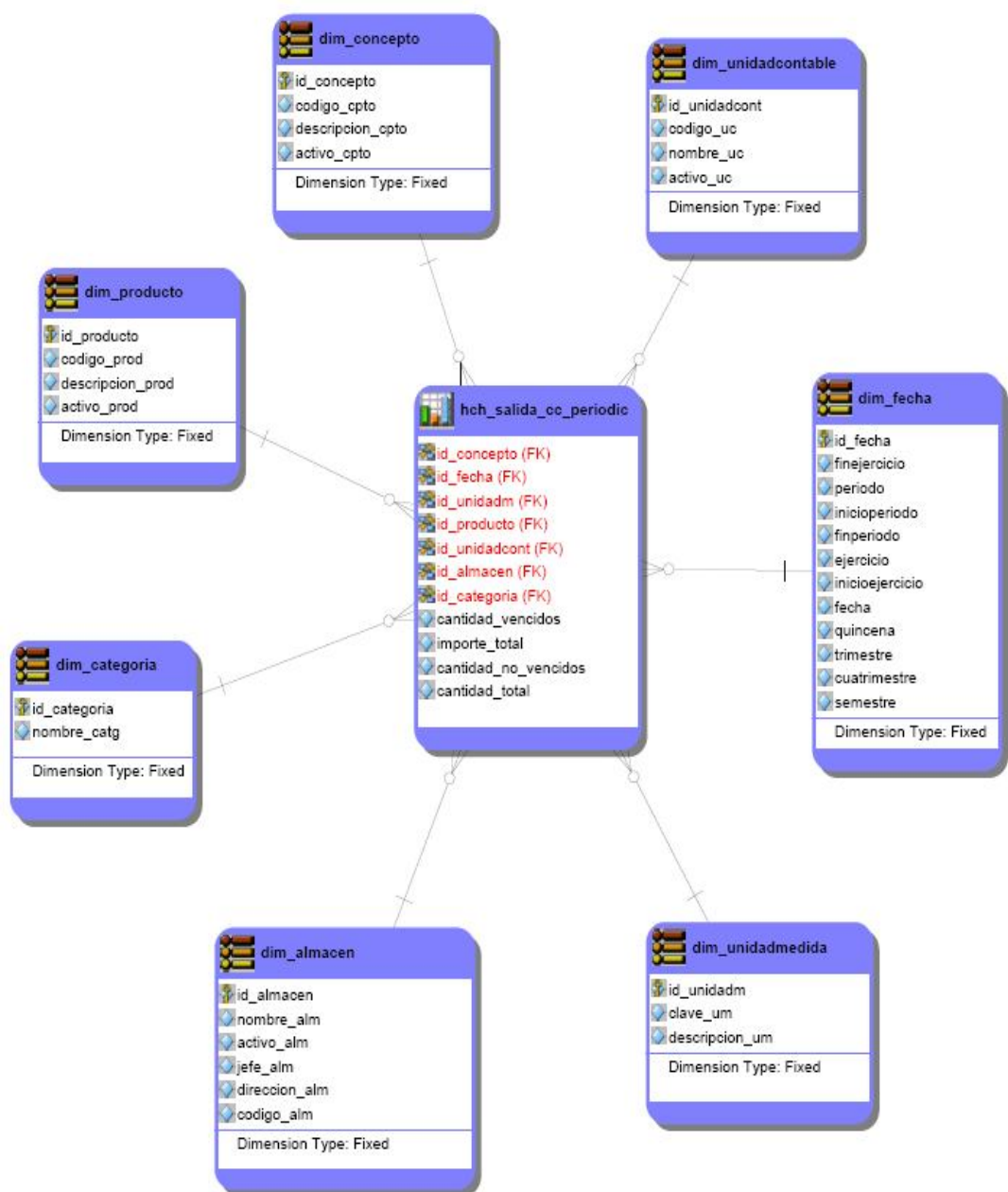
I.12 Esquema estrella existencia almacén



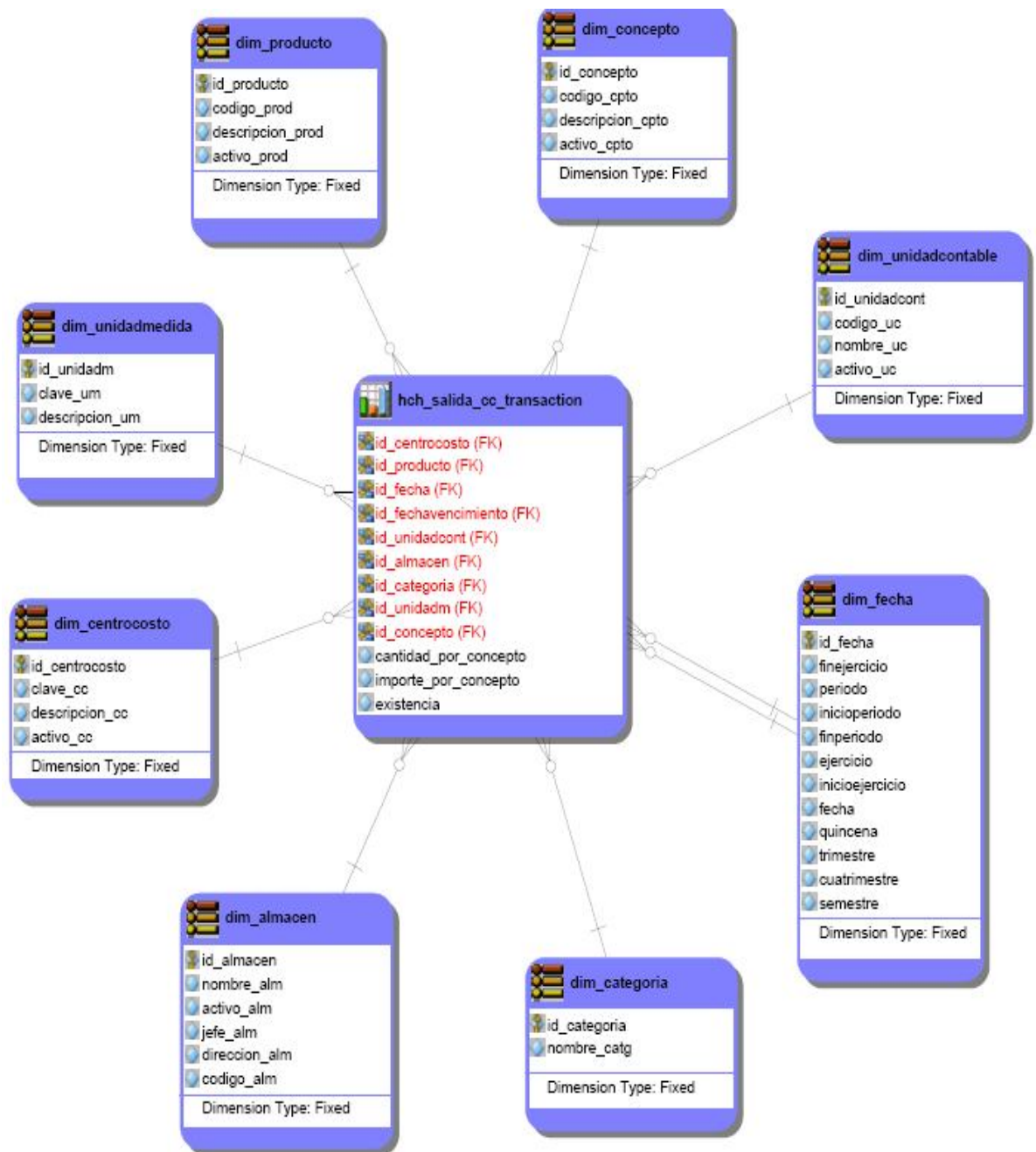
I.13 Esquema estrella existencia custodio



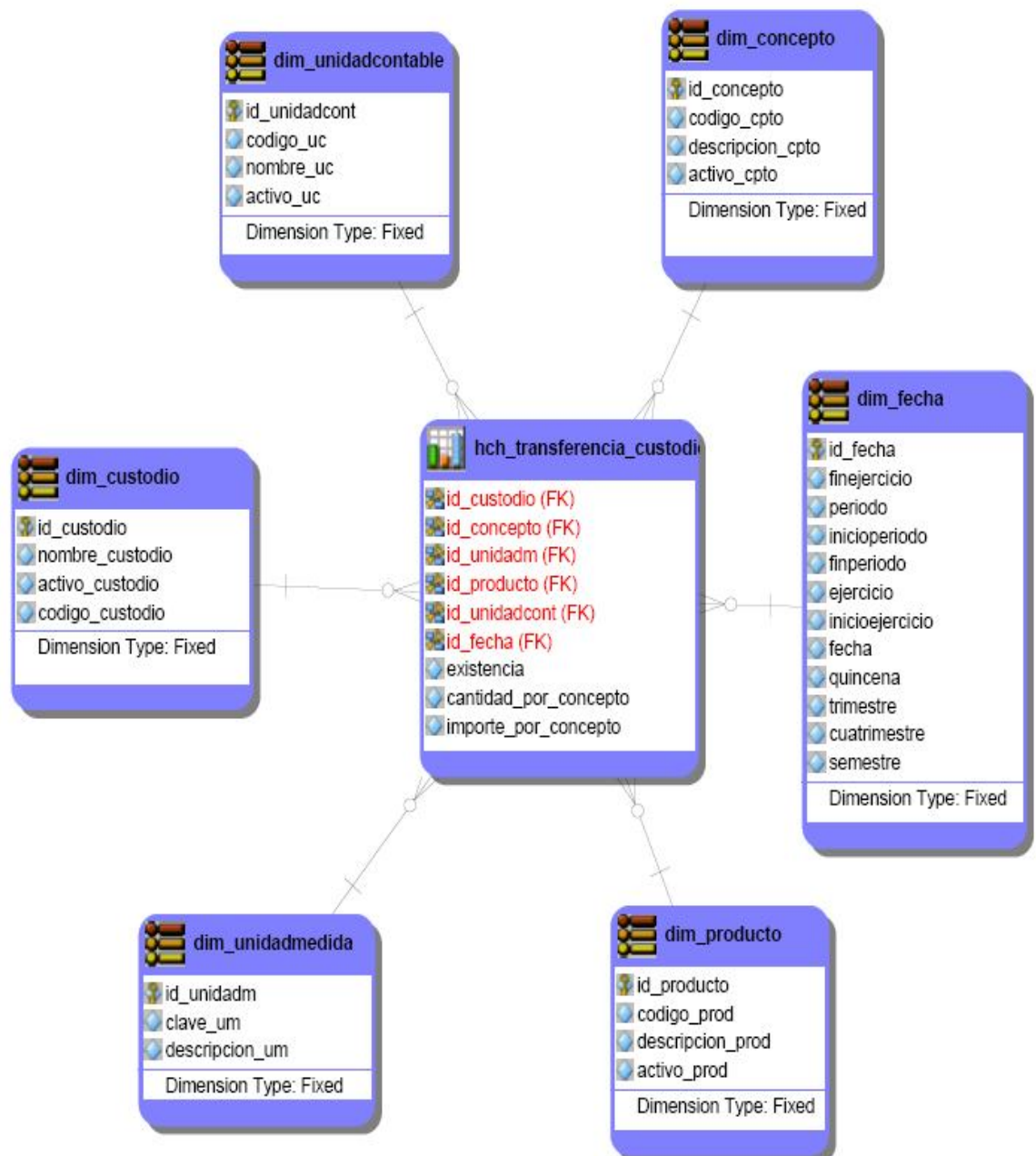
I.14 Esquema estrella del subproceso de salida CC a nivel de gestión de custodio



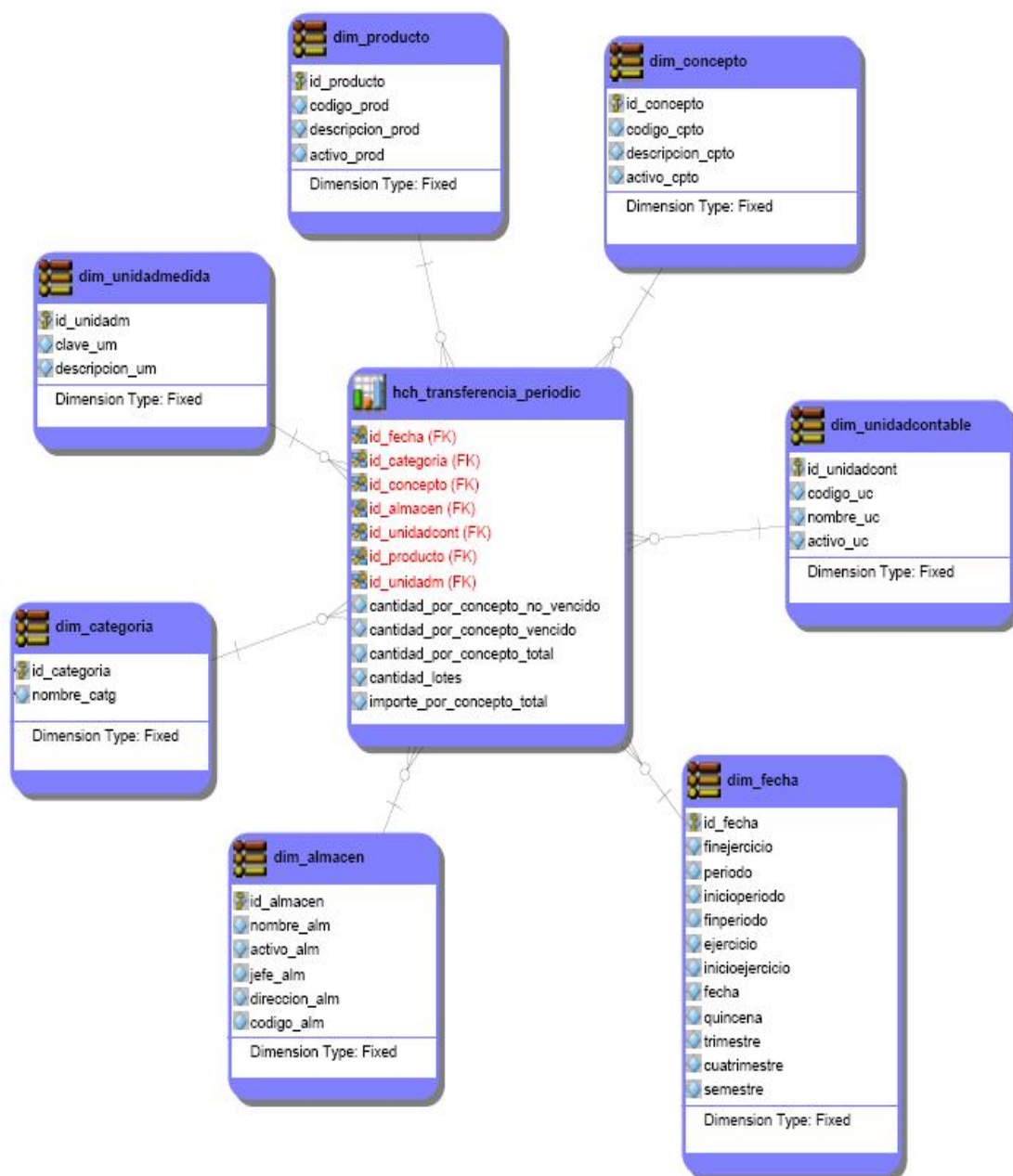
I.15 Esquema estrella del subproceso de salida CC a nivel periódico



I.16 Esquema estrella del subprocesso de salida CC a nivel transaccional



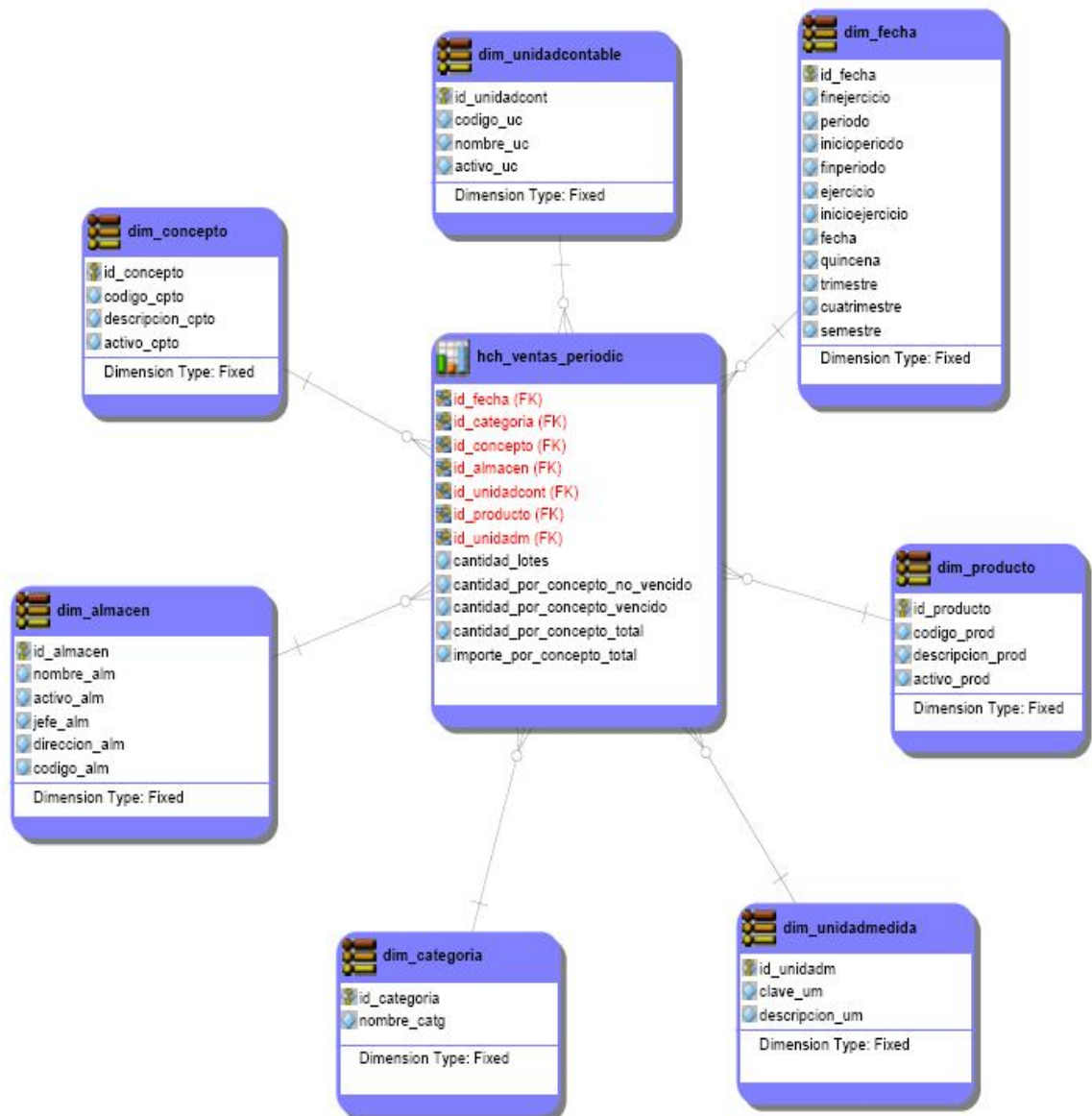
I.17 Esquema estrella del subproceso de transferencia a nivel de gestión de custodio



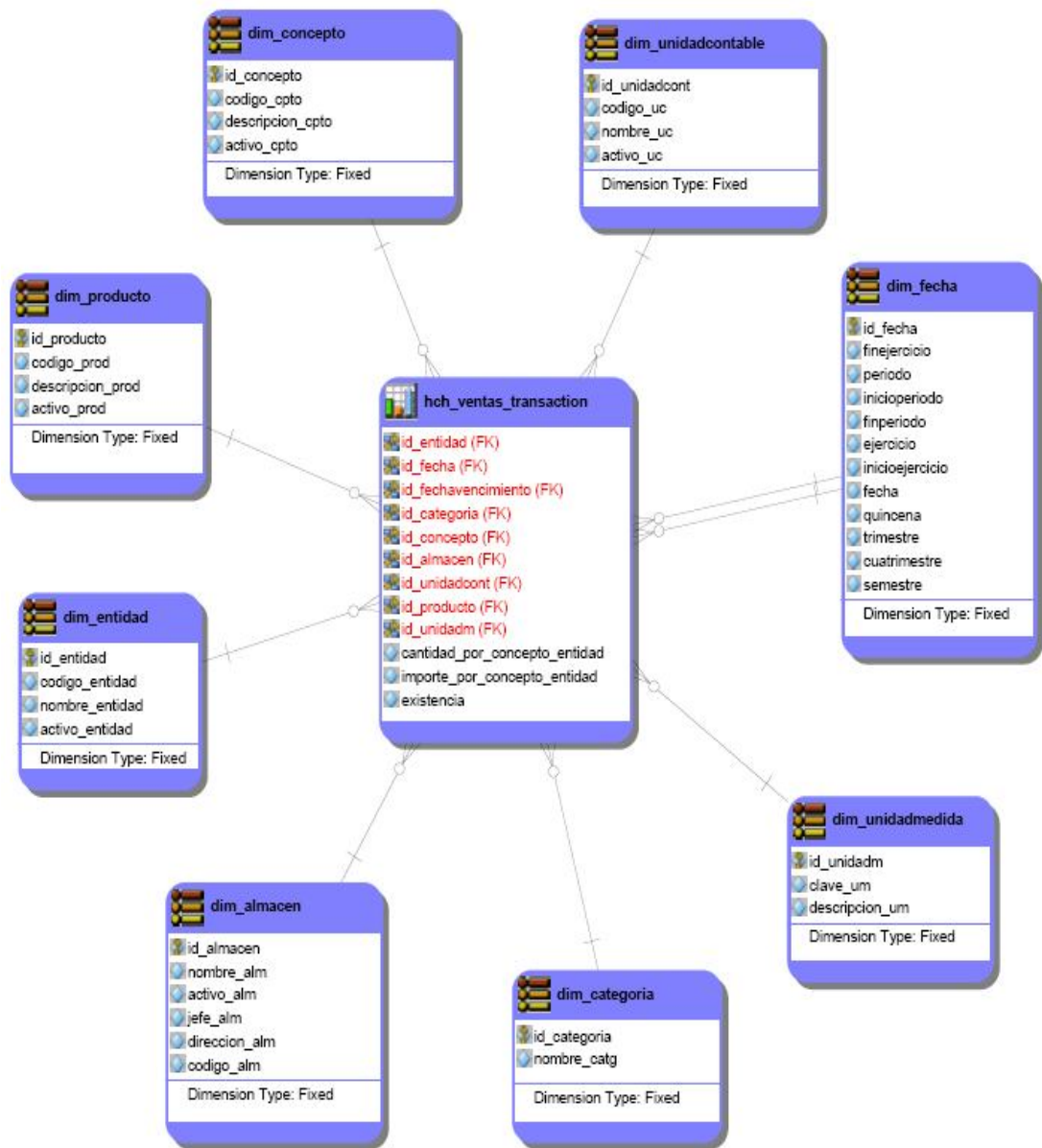
I.18 Esquema estrella del subproceso de transferencia a nivel periódico



I.19 Esquema estrella del subproceso de transferencia a nivel transaccional



I.20 Esquema estrella del subproceso de ventas a nivel periódico



I.21 Esquema estrella del subprocesso de ventas a nivel de transaccional