

Prototipo de mercado de datos para la División de Control y Fiscalización de la Dirección General de Aduanas

Ricardo Chinchilla Arley ¹
José Luis Vargas Murillo ²

La información administrada por el Servicio Nacional de Aduanas (SNA) en todo el país es almacenada actualmente en el Sistema de Información Aduanera (SIA).

Palabras claves

Bases de datos, sistemas de información, almacenes de datos

Resumen

El presente trabajo es un resumen del proyecto de graduación de ambos autores. Propone la aplicación de un mercado de datos para estructurar una división de la Dirección General de Aduanas, con el fin de realizar investigación fiscales y aduaneras. Presenta un prototipo basado en un esquema de bases de datos, con una tabla de hechos y 11 tablas dimensionales. Los datos son extraídos de la base de datos del Sistema de Información aduanero; son transformados e ingresados al prototipo, donde podrían ser consultados utilizando una interfase tipo OLAP. Se implementa también el acceso por Internet.

Introducción

La información administrada por el Servicio Nacional de Aduanas (SNA) en todo el país es almacenada actualmente en el Sistema de Información Aduanera (SIA), el cual contiene los datos de los ingresos fiscales por concepto de impuestos de importación, que constituyen aproximadamente el 50% del presupuesto nacional. El SIA consiste en un conjunto de aplicaciones de base de datos sobre una plataforma Oracle”, que procesa y registra las transacciones aduaneras que recibe.

A pesar de que el SIA mantiene los registros de las operaciones de los usuarios desde 1993, la información que alberga es controlada por aplicaciones desarrolladas para el registro, seguimiento y control de transacciones, que no están diseñadas para permitir el

-
- 1 Ricardo Chinchilla Arley, Master en Computacion ITCR, funcionario Centro Centroamericano de Población, Director Biblioteca Sede del Atlántico, Universidad de Costa Rica. (richar@ccp.ucr.ac.cr).
 - 2 José Luis Vargas Murillo, Master en Computacion ITCR, funcionario División de Control y Fiscalización de la Dirección General de Aduanas. (vargasmjl@hacienda.go.cr).

acceso rápido y preciso requerido en los procesos de análisis y toma de decisiones de la administración aduanera.

Este sistema tiene 10 años de estar en producción. La versión sobre la que está desarrollado no recibe soporte técnico del fabricante, debido a su antigüedad; realizar modificaciones estructurales profundas supondrían un costo muy elevado.

Tal circunstancia dificulta también la adopción de estrategias y el establecimiento de directrices de control que permitan definir con menor incertidumbre las empresas, procesos o mercancías que se habrán de seleccionar como sujetos de estudio para propósitos fiscales y aduaneros.

Es necesario contar con una tecnología que permita hacer comparaciones, encontrar patrones de comportamiento, hallar tendencias, establecer relaciones que ayuden a la detección o prevención de delitos aduaneros o el señalamiento de incumplimientos de las leyes y normas del comercio nacional e internacional. Es aquí donde el desarrollo de un Depósito de Datos (DD), definido por González (1998a: 11) como un “[...] repositorio único, completo y consistente de datos obtenidos de diferentes fuentes [...] que están disponibles al usuario de forma tal que pueda ser comprendido y usado en un contexto de empresa”, proporcionaría al SNA una herramienta capaz de suministrar información transformada, filtrada, resumida, agrupada y lista para satisfacer los requerimientos de los usuarios.

El proceso de análisis o toma de decisiones basado en la información extraída del SIA es una tarea cotidiana de la División de Control y Fiscalización (DCF), la cual está encargada de definir la pauta en la orientación y planeamiento de las

funciones fiscalizadoras del SNA en general. Entre las funciones de la DCF se encuentra velar por la definición y aplicación oportuna y precisa de criterios de control que son los que, a fin de cuentas, determinarán hacia quién, cómo y cuándo deben dirigirse los esfuerzos fiscalizadores.

Los criterios de control emanados de la DCF se basan en varias fuentes de información, entre las cuales están los informes estadísticos producidos por la División de Estadística, Registro y Divulgación de la DGA, generados utilizando la información almacenada en el SIA. Debido a que este sistema no es apropiado para el trabajo de análisis, los datos que se utilizan para satisfacer las solicitudes de información provienen de archivos de texto (planos) extraídos del SIA, generados por funcionarios en las aduanas o desde la misma División de Estadística.

Estos archivos planos son producidos por rutinas con instrucciones SQL (lenguaje estructurado de peticiones) que extraen los valores de las variables seleccionadas y registradas en un período determinado. Estos valores son descargados en tablas tipo DBF (por sus siglas en inglés de *Data Base File*), a partir de las cuales se extraen otras tablas con información específica, según los requerimientos de la DCF y otras oficinas o dependencias. Ninguna de las tablas tiene relación entre sí.

Es necesario contar con un sistema que permita la integración de la información obtenida del SIA y provea las herramientas necesarias para ser utilizadas en el análisis y la evaluación de la información extraída. Esa información contribuirá a disminuir la incertidumbre en el control de los riesgos aduaneros, tales como son la defraudación fiscal, el contrabando y el incumplimiento de la normativa aduanera.

La herramienta que se utilizó para la construcción del prototipo fue Visual FoxPro 6.0, por considerar que reúne las características necesarias para su desarrollo.

Corey (1997: 33) define el Mercado de Datos (MD) como un “(...) *subconjunto de un Depósito de Datos que se extrae para satisfacer las necesidades de una clase particular de usuario*”. Es una tecnología que resolvería la problemática descrita, debido a las facilidades de almacenamiento y procesamiento analítico de que dispone.

Un MD proporciona un esquema de almacenamiento que permite realizar consultas complejas, por medio de la interacción e interrelación de múltiples entidades, manteniendo solo un repositorio con la información agregada y ordenada, según las necesidades de los usuarios. Además, provee de una interfase de consulta que permite al usuario la posibilidad de análisis de la información para la toma de decisiones. Esto con el propósito de atacar los delitos e infracciones aduaneras, así como procurar el ejercicio del control fiscal.

Alcances y limitaciones

El prototipo se limita a los datos suministrados en los informes estadísticos solicitados por la DCF a la División de Estadística Registro y Divulgación. Dichos datos consisten en aquellos valores contenidos en los campos del SIA correspondientes a las declaraciones aduaneras de importación definitiva, que la DCF ha considerado como necesarios para la atención de la demanda de información. Para esta aplicación se utilizó una base inicial de 250 mil registros.

La herramienta que se utilizó para la construcción del prototipo fue Visual FoxPro 6.0, por considerar que reúne las características necesarias para su desarrollo.

La base de datos fue implementada bajo el esquema estrella y sin tablas de resumen, por lo que el cálculo de las consultas es efectuado sobre la tabla de

hechos general. Por otro lado, el prototipo proporciona una interfase de consulta que pretende ser una aproximación a un OLAP (On-Line Analytical Processing). Esta interfase permite consultas multidimensionales, la reducción dimensional y visualizar la información agrupada desde distintas perspectivas, además de ofrecer salidas a distintos formatos de archivo.

Prototipo de mercado de datos

Definición de dimensiones

Las dimensiones utilizadas en esta aplicación son un subconjunto de las variables o campos que conforman parte de la información transaccional de los importadores que, en sociedad con las agencias de aduana, diariamente es transmitida a las aduanas del país.

Las dimensiones descriptivas que se proponen para el Mercado de Datos son:

- Aduana
- Agencia
- Importador
- Mercancía
- Lugar de descarga
- País de origen
- País de procedencia
- Modalidad
- Trámite
- Semáforo
- Tiempo

Las medidas asociadas a las dimensiones indicadas son:

- Número de bultos
- Valor Cif (costo, seguro y flete)
- Derechos arancelarios a la importación.
- Impuesto selectivo de consumo.
- Impuesto de la ley 6946.
- Impuesto de ventas.
- Impuesto a las mercancías con destino al depósito libre de Golfito.
- Impuesto a las mercancías del tratado libre de comercio con México.

- Total de impuestos pagados.
- Total de impuestos exonerados.

Metodología de desarrollo

La primera tarea que se llevó a cabo fue determinar las necesidades de información de los usuarios, aplicando entrevistas e investigando en los archivos históricos del Departamento de Estadística todo lo relacionado con las solicitudes de información. Esto dio como resultado el conjunto de dimensiones descriptivas y cuantitativas.

Posteriormente, se hizo el análisis de los datos dentro de cada dimensión considerada, modelando con ellos un primera visión del prototipo. La meta era buscar el acomodo de los requerimientos de los usuarios al concepto que se iba desarrollando. En esta etapa se utilizaron datos de prueba reales, con los cuales los usuarios exploraban si la interfase que se ofrecía satisfacía sus necesidades. Se continuó con ese tipo de evaluación hasta considerar que la información proporcionada por las dimensiones que formaban el modelo, así como las medidas que las delimitaban, cumplían aceptablemente sus pretensiones.

El concepto metodológico general que se ha empleado en esta propuesta ha sido desarrollar una solución de facilitación de información para análisis y toma de decisiones, a partir de sectores específicos, como es el caso de la DCF. Posteriormente, se pueden considerar nuevas etapas con la inclusión incremental de otras áreas de interés (nuevos mercados de datos), como por

ejemplo el de valoración de mercancías, la verificación arancelaria o los de control de procesos aduaneros.

Para el diseño lógico del MD se utilizó la propuesta de modelaje de hipercubos planteado por Sonia Cortez Araniva (Cortez 1999). Las razones por la cuales se adopta este modelo son, en primer término, su simplicidad y fácil comprensión, y debido a que los otros modelos no se han orientado, según dicha autora, a sugerir un modelo multidimensional gráfico y sin orientación a una implementación específica.

Diseño lógico

Modelo gráfico

En el Cuadro 1 muestra los elementos a modelar desde una perspectiva general:

Se implementa entonces un MD para las importaciones definitivas, tomando en cuenta las mercancías importadas, según los capítulos, partidas arancelarias, y el nivel de apertura nacional. Se considera, además, el período en que se realizan éstas. Dicho período puede ser en años fiscales o en años calendario; estos

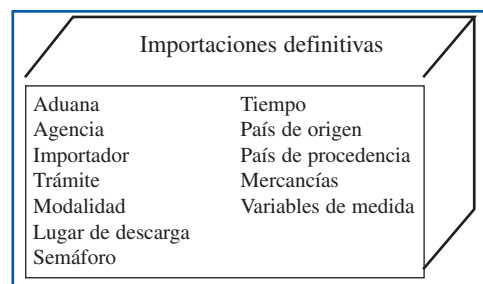


Figura 1
Modelo general del MD

Cuadro 1 Descripción de elementos a modelar				
Importaciones definitivas	Mercancías Período	Capítulos Años calendario Años fiscales	Partidas Trimestres	Apertura nacional Meses

últimos incluyen trimestres y meses. La descripción anterior define las jerarquías de las dimensiones *mercancía* y *tiempo* del prototipo.

El modelo general del MD, sería:

Por su parte, el modelo dimensional para el sistema propuesto sería el que se muestra en la Fig. 2.

Aquí se presenta cada una de las dimensiones con sus respectivos atributos, todas ellas unidas a la tabla *Importación_defi*. Por su parte, el modelo que representa las diferentes jerarquías se muestra en la Fig. 3.

En este modelo se despliegan las jerarquías existentes en el MD. Se puede apreciar el despliegue

correspondiente de las dimensiones *Tiempo* y *Mercancías*, además de los atributos de clasificación de la tabla de hechos.

Descripción del modelo

La descripción del modelo se detalla a continuación:

Descripción del modelo

Hipercubo: *Importaciones definitivas*

Definición de dimensiones

Importación_defi: Variables de medida en la importación que consideran el valor de las mercancías, los impuestos pagados, las exoneraciones y la cantidad de bultos

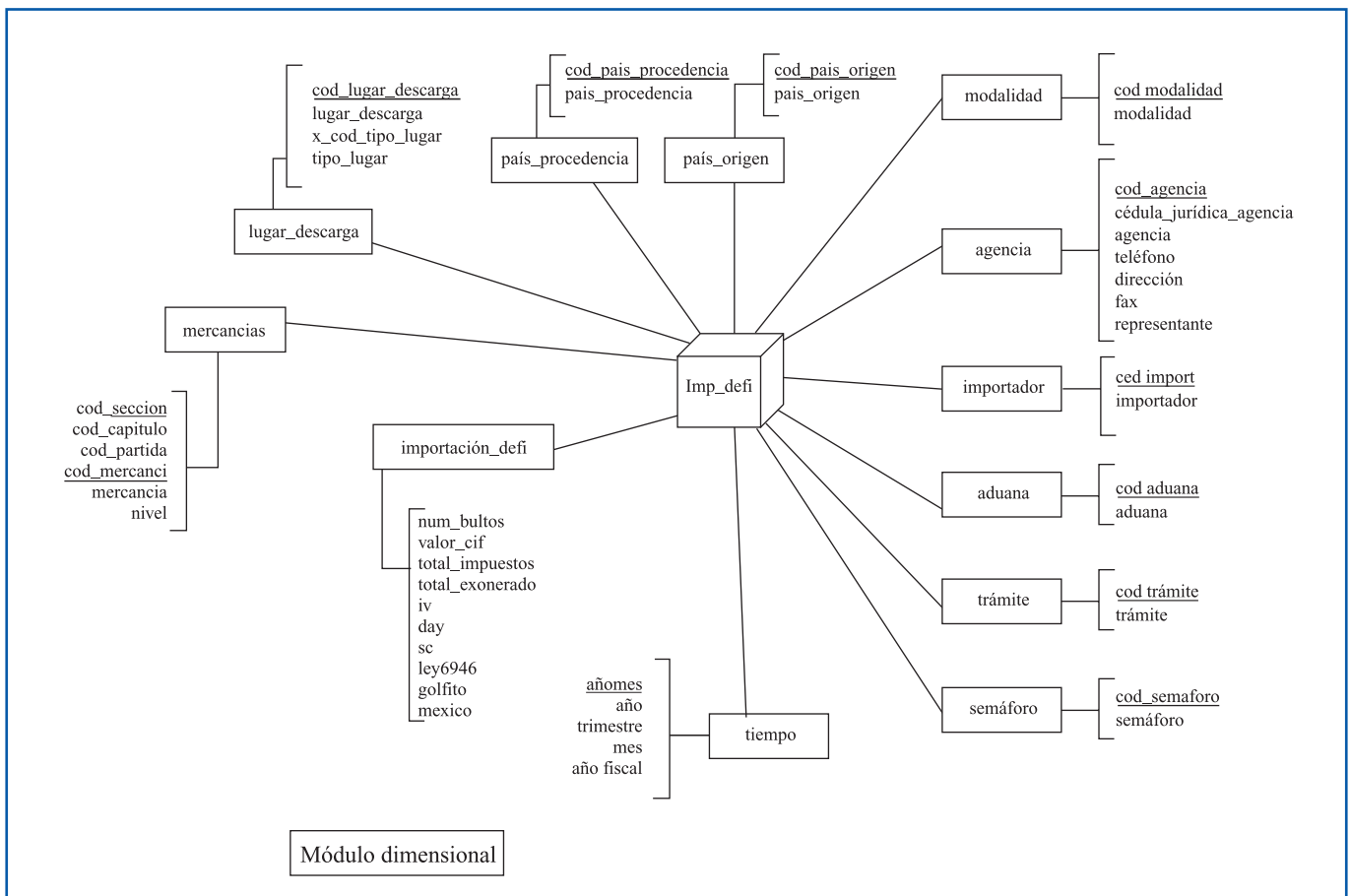


Figura 2
Esquema del modelo dimensional del prototipo

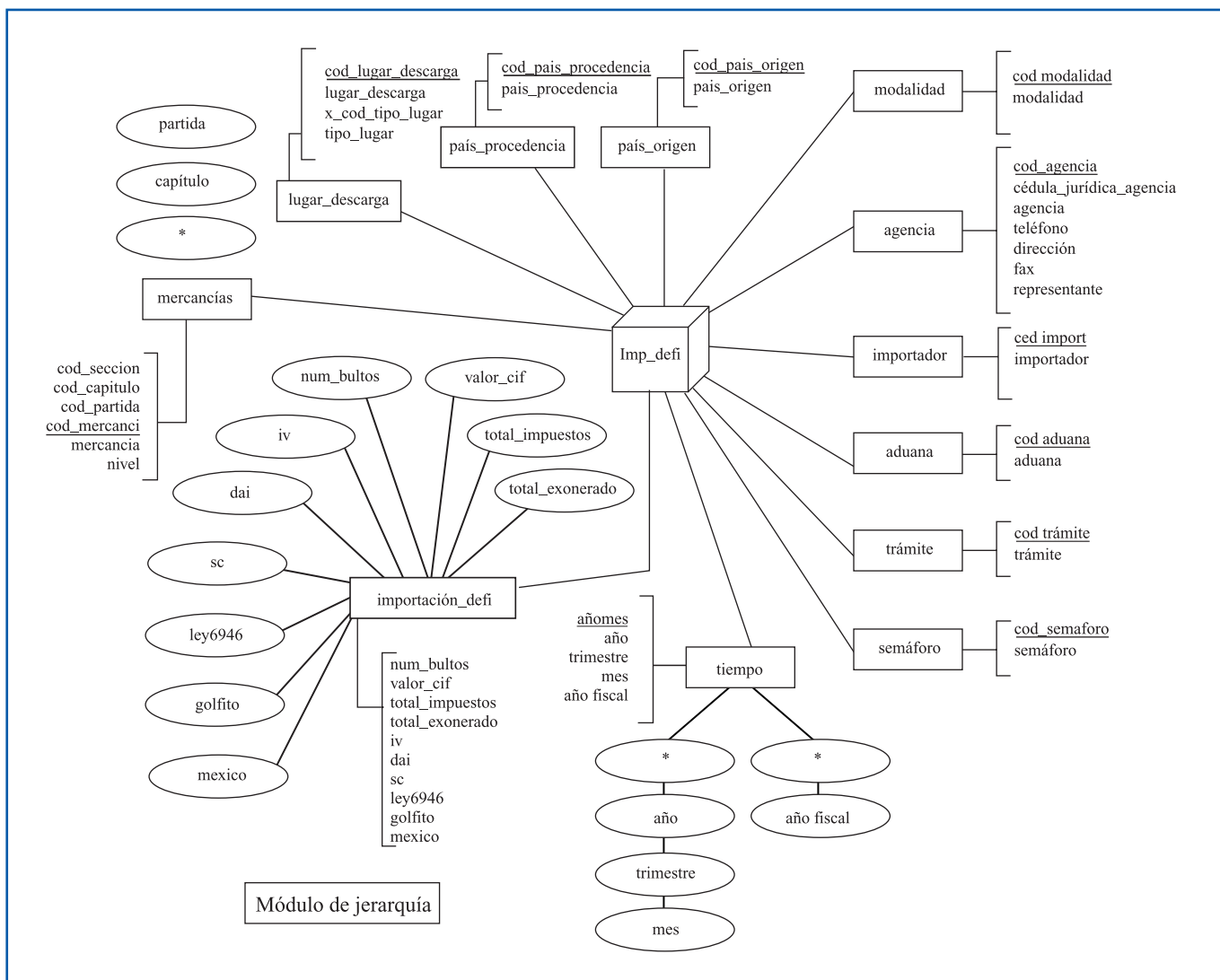


Figura 3
Esquema del modelo de jerarquías del prototipo

Modalidad: Categoría sobre el tipo de mercancía y destinación

Agencia: Nombre y descripción de la agencia aduanal encargada de la importación.

Importador: Nombre y descripción del importador de mercancías.

Aduana: Nombre y descripción de la aduana por donde se realiza el trámite de desalmacenaje.

Trámite: Tipos de trámite para el desalmacenaje.

Lugar_descarga: Lugar físico donde se encuentra la mercancía a desalmacenar

País_origen: País de donde es originaria la mercancía importada

País_procedencia: País del cual procede la mercancía importada

Semáforo: Tipo de revisión al cual se somete la mercancía

Mercancías: Descripción de las mercancías importadas

Tiempo: Período de tiempo en el que se consideran las variables

Definición de dominio

[Cód_modalidad]: número entero de dos dígitos.

[Modalidad]: nombre de todos los tipos de operación.

[Cód_agencia]: número entero de tres dígitos.

[Céd_jurídica_agencia]: número entero de 11 dígitos.

[Agencia]: nombre de todas las agencias del país.

[Céd_importador]: número entero de 12 dígitos.

[Importador]: nombre de todos los importadores del país.

[Cód_aduana]: número entero de dos dígitos.

[Aduana]: nombre de todas las aduanas del país.

[Cód_tramite]: código alfabético de un carácter.

[Trámite]: nombre de los tipos de trámite aplicados a la mercancía

[Cód_lugar_descarga]: código alfanumérico de tres caracteres.

[Lugar_descarga]: nombre de los tipos de lugar de descarga de mercancías.

[Cód_país_origen]: número entero de 4 dígitos correspondiente al código ISO de países.

[País_origen]: nombre de todos los países de donde son originarias las mercancías.

[Cód_país_procedencia]: número entero de 4 dígitos correspondiente al código ISO de países.

[País_procedencia]: nombre de todos los países de donde proceden las mercancías.

[Cód_semaforo]: número entero de un dígito, el cual puede ser 0, 1 o 2.

[Semáforo]: nombre del tipo de revisión aplicado a la mercancía.

[Cód_capitulo]: número entero correspondiente a los primeros dos dígitos del arancel de mercancías.

[Cód_partida]: número entero correspondiente a los primeros cuatro dígitos del arancel de mercancías.

[Cód_mercancia]: número entero correspondiente a los 10 dígitos del Sistema Arancelario Centroamericano.

[Mercancia]: nombre de las mercancías.

[Añomes]: número entero de 6 dígitos.

[Cód_año]: número entero de dos dígitos.

[Año]: número entero de cuatro dígitos.

[Cód_trimestre]: nombre de número ordinal correspondiente a los trimestres del año.

[Trimestre]: número y nombre del trimestre.

[Cód_mes]: nombre de número ordinal correspondiente a los meses del año.

[Mes]: número y nombre de los meses del año.

[Cód_añofiscal]: número entero de dos dígitos.

[Año fiscal]: número entero de cuatro dígitos.

Definición de jerarquías

Dimensión Mercancías:

J ₁ :	P ₁ : * → capítulo P ₂ : capítulo → partida P ₃ : partida → mercancía	Pliega todos los capítulos Pliega todas las partidas por capítulo Pliega todas las mercancía por partida
------------------	--	--

Dimensión Tiempo

J ₁ :	P ₁ : * → años P ₂ : año → trimestre P ₃ : trimestre → mes	Pliega todos los años calendario Pliega todos los trimestres por año Pliega todos los meses por trimestre
------------------	---	---

J ₂	P ₁ : * → Años fiscal	Pliega todos los años fiscales
----------------	----------------------------------	--------------------------------

J₁ \longleftrightarrow J₂

Las jerarquías J₁ y J₂ son mutuamente excluyentes

Cada uno de los atributos requiere un dominio, el cual se coloca junto a ellos encerrado entre corchetes.

Se han definido jerarquías en las dimensiones de Mercancías y Tiempo. En el caso de la dimensión Tiempo las jerarquías año calendario y año fiscal son mutuamente excluyentes; por tanto se pueden hacer consultas sobre una jerarquía determinada, pero no sobre ambas al mismo tiempo.

Cada una de las dimensiones propuestas está estructurada por los atributos descriptivos asociados, uno de los cuales necesariamente corresponde al que identifica unívocamente a cada registro, o sea su llave primaria.

Por su parte la tabla de hechos (importacion_defi) que contiene las variables de medida, se enlaza con las dimensiones a través de las llaves foráneas provenientes de cada una de ellas. La unión de las llaves foráneas declaradas en la tabla de hechos conforma su llave primaria, completando así el esquema de estrella.

Representación de operaciones

Drill Down - Roll Up

Es posible, de acuerdo con la estructura de jerarquías diseñada en este modelo, ejecutar las operaciones de Drill Down y Roll Up sobre las dimensiones Tiempo y Mercancías conforme a la siguiente figura:

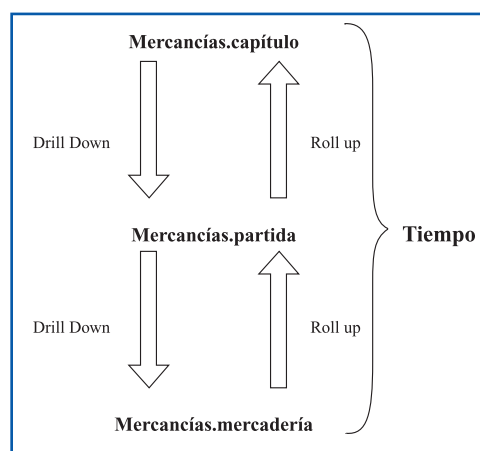


Figura 4
Drill Down y Roll Up
en la dimensión mercancía

Mientras se mantiene cualquier jerarquía de la dimensión Tiempo, se puede profundizar o plegar desde la jerarquía de Capítulo hasta la jerarquía de Mercancía y viceversa.

Así mismo, se puede realizar la misma operación para la dimensión Tiempo a través de la combinación de cualquiera de las demás dimensiones consideradas en el modelo:

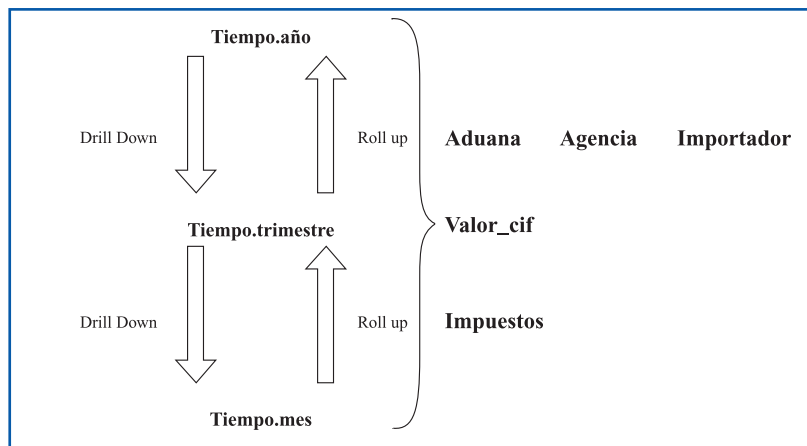


Figura 5
Drill Down y Roll Up en la dimensión Tiempo

En este caso se mantiene cualquier combinación de, por ejemplo, Aduana – Agencia – Importador, y los valores de medida Valor_cif e Impuestos, mientras se profundiza o pliega desde la jerarquía Año hasta la jerarquía Mes de la dimensión Tiempo, y viceversa.

En este caso concreto, el mínimo nivel de granularidad de la jerarquía es la de mes. Esta decisión se basa en tres razones: 1. por cuestiones académicas de demostración de la funcionalidad del prototipo, 2. por cuestiones prácticas de ahorro de espacio físico y velocidad de procesamiento, y 3. se determinó que este último nivel satisface las necesidades de información de los usuarios potenciales del prototipo. Sin embargo, la incorporación del nivel de jerarquía día, o inclusive, hora, no

representaría ningún problema técnico ni funcional, y podría ser implementado sin ninguna dificultad.

Diseño físico

Base de datos

Con base en el modelo lógico representado gráficamente con los símbolos de Cortez (1999), se desarrolla el modelo físico.

Para la implementación de la base de datos se han creado tablas para cada una de las dimensiones descriptivas del modelo conceptual, las cuales contendrán los atributos de clasificación respectivos y la llave primaria. La Fig. 6 muestra el mapeo del modelo utilizando el esquema de implementación tipo estrella, donde todas las dimensiones se relacionan con la tabla de hechos con una cardinalidad de uno a muchos.

La tabla de hechos contiene las variables de medida. La llave primaria de la tabla de hechos está compuesta por el conjunto total de llaves foráneas, las cuales, a su vez, son llaves primarias de sus respectivas dimensiones $PK = (FK1 + FK2 + \dots + FK_n)$.

Con respecto a la integridad referencial, se debe recordar que se trabaja con una base de datos histórica, razón por la cual las dimensiones son raramente modificadas o actualizadas, ejerciendo un estricto control sobre los datos desde los procesos de carga. La integridad referencial es administrada por los dominios de los atributos previamente definidos. De esta forma, no es posible insertar o actualizar un dato si el valor propuesto para un atributo no satisface las reglas de su dominio.

Para todas las dimensiones, la tabla secundaria común es Importación_defi. En la regla de actualización se posibilita actualizar solamente la tabla secundaria, ya que los registros de las tablas primarias no son modificados. En la

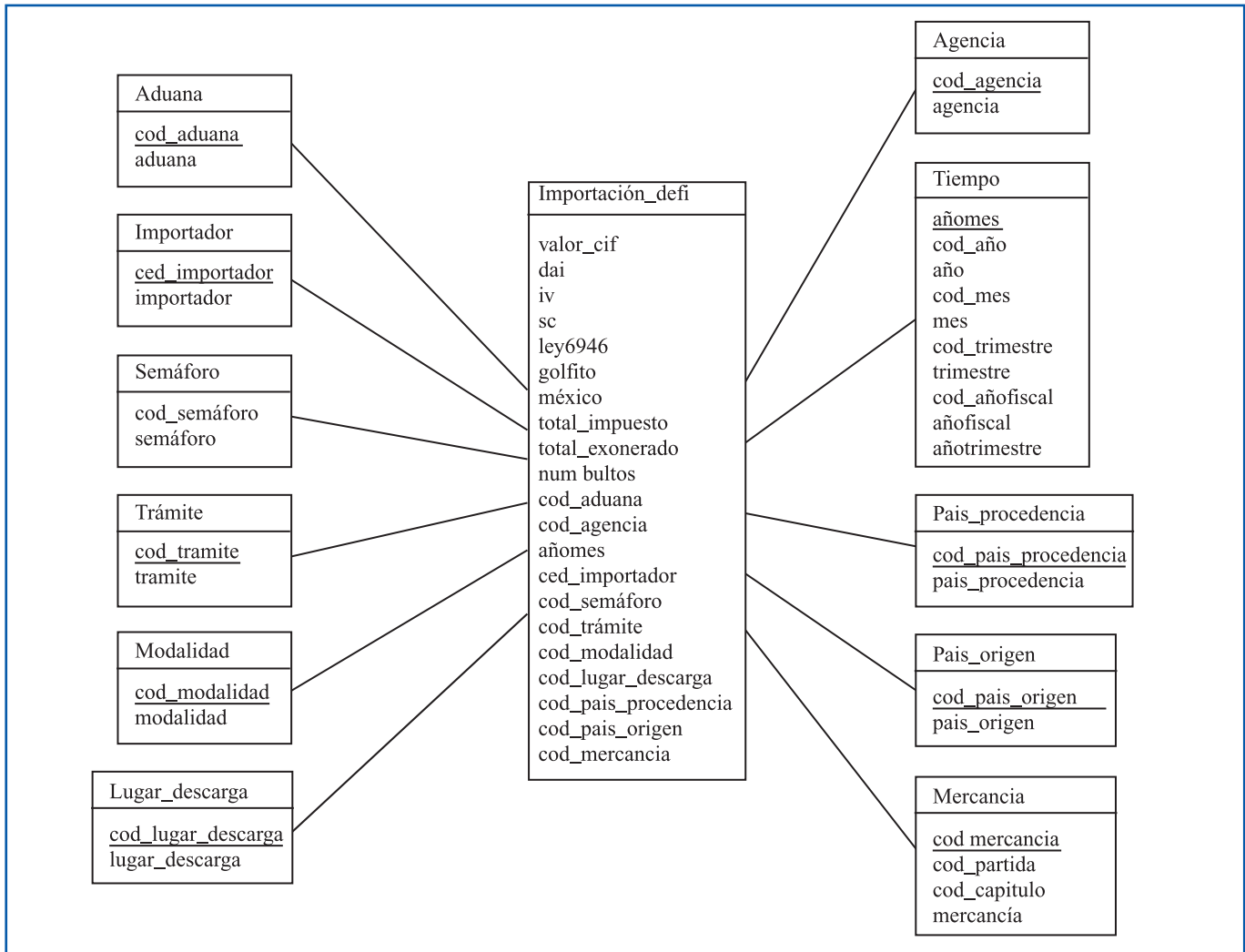


Figura 6
Modelo físico de la base de datos

regla de inserción, se permite insertar registros a la tabla secundaria aunque no existan registros coincidentes en las tablas primarias, esto debido a que los datos son previamente filtrados antes de su carga y se eliminan los registros sin coincidencias. En la regla de eliminación, es posible quitar registros de las tablas primarias aunque estén vinculados a la secundaria, esto porque, al ser una base de datos histórica, dichos registros nunca serán eliminados.

Por último, los datos a cargar han sido extraídos de un sistema transaccional que

aplica su propio esquema de control de integridad referencial en su operación normal. Esto debería garantizar que los datos que ingresan al sistema de carga del prototipo ya han sido depurados al nivel de integridad del sistema transaccional.

Extracción, transformación y carga de datos

El sistema de extracción del prototipo consta de tres subsistemas: el de transformación, el de agregación y el de carga. La Fig. 7 muestra el proceso con sus respectivos subsistemas. El proceso de extracción, que recolecta los datos

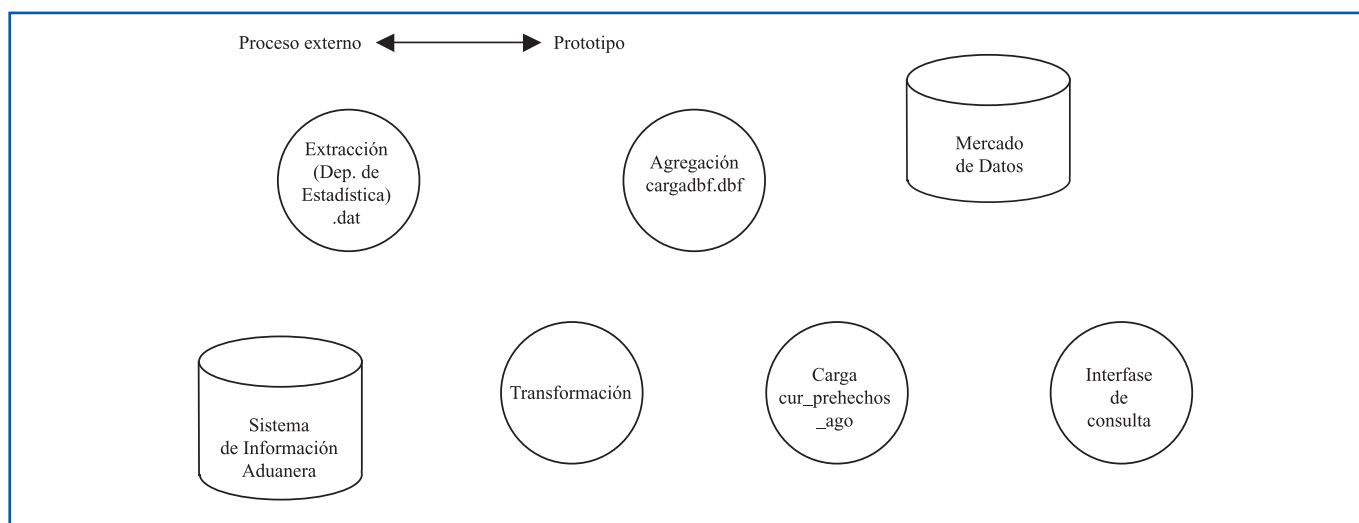


Figura 7
Proceso de transformación y carga

requeridos por el sistema alimentador, es realizado en el Departamento de Estadística.

El subsistema de transformación convierte los datos en el formato de destino y elimina los campos que no son requeridos. El subsistema de agregación totaliza los datos hasta la jerarquía de Mes de la dimensión Tiempo. El subsistema de carga inserta los datos dentro del MD.

Un aspecto importante es el hecho de que, al provenir los datos de una sola fuente (SIA), no es necesario implementar opciones para solucionar los problemas de diferencias lógicas, de tipo o de tamaño en los formatos de los archivos. Por tal motivo, los datos se cargan sin conciliación.

La consulta de información en el prototipo

El centro neurálgico de la aplicación prototipo lo constituye la herramienta que permite extraer los datos agregados para un posterior análisis (Fig. 8). Para el diseño de una interfase de consulta capaz de realizar esta función, fue

necesario definir la secuencia lógica de las consultas, a fin de facilitar su operación.

Se decidió establecer, como punto de partida, la definición del periodo en que se desearía efectuar el análisis. Posteriormente, la selección de las dimensiones y sus instancias, así como el orden de aparición. Por ultimo, las unidades de medida y su ordenamiento.

Para efectuar estas operaciones, existe un procedimiento central, que reúne todos las distintas variables seleccionadas a lo largo de la construcción de la consulta.

Todas las variables seleccionadas o modificadas en la interfase, son manejadas por el código de un método de la aplicación que permite el cambio interactivo, reflejado directamente en el resultado de la consulta. Cualquier modificación en la interfase, envía las nuevas variables o valores al procedimiento llamado bldSQL, para la construcción interactiva y automática de la sentencia SQL que determina los resultados de la consulta final.

La instrucción SQL final que genera el procedimiento bldSQL es la siguiente:

```
lcSelect = "SELECT "+ lcCamposGrl + lcDimMedida+;
" FROM "+ lcFromGrl+;
lcWhereGrl+;
lcGroupBy+;
lcOrderBy
```

Cada uno de los elementos, con el prefijo lc, contiene cadenas de texto que la configuran, según sean las selecciones, filtros y restricciones definidas interactivamente por el usuario.

Acceso al mercado de datos por Internet

Con el desarrollo de un enlace por Internet al MD, se pretende facilitar el acceso de la información a los usuarios alejados en las aduanas. Además,

pretende servir de base para una futura expansión del prototipo, en un eventual proyecto de Data Warehouse a escala general.

Internet Database Connector (IDC)

Para el acceso a los datos contenidos en el MD, se utiliza un servidor con Microsoft NT Server como sistema operativo, Internet Information Server (IIS) como servidor web y la herramienta ODBC (*Open DataBase Connectivity*) Internet Database Connector (IDC) para el enlace a los datos.

Conceptualmente, el acceso a la bases de dato es realizado por el IIS, como se muestra en la Fig. 9. El visualizador del cliente envía las consultas al servidor Web a través del protocolo de transferencia de hipertexto (http). El servidor Web responde con un documento en formato HTML montado sobre una plantilla (htx). El acceso a la base de datos es proporcionado por el componente IDC, que es un programa *Internet Server Application Program Interface* o DLL/ISAPI (httpodbc.dll) que utiliza ODBC para establecer dicho acceso.

Construcción del enlace

Para la construcción del enlace del MD a Internet, se definieron las consultas que podrían requerir los altos ejecutivos del Ministerio de Hacienda, los gerentes o las jefaturas de departamento de las aduanas. Se toman estos usuarios clave como base para la definición debido a que, al encontrarse geográficamente distanciados de la DCF, no es posible el acceso directo y personal a la interfase de consulta principal.

Las consultas disponibles que se plantean son:

- Consultas generales
 - Aduana e importador
 - Aduana y agencia
 - Aduana y lugar de descarga



Figura 8
Interfase de consulta

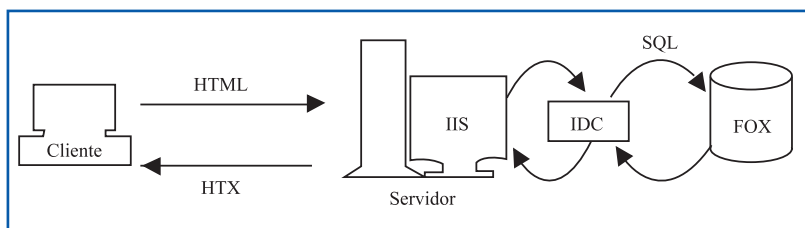


Figura 9
Acceso a bases de datos por IIS

- Consultas específicas por aduana y lugar de descarga
 - Aduana e importador
 - Aduana y agencia
 - Aduana y lugar de descarga por agencia
- Consulta por mercancías
 - General por aduana
 - Aduana e importador
 - Aduana y agencia
 - Aduana y lugar de descarga
 - Aduana y país de origen
- Consulta por tipo de revisión
- Mayores importadores mensuales

Una vez especificadas, se ejecuta el primer paso, en el cual se construye una tabla libre agregada sobre la cual se realizarán las consultas

El segundo paso es la configuración del origen de datos ODBC mediante el administrador de orígenes de datos ODBC de NT. Para ello se define un origen tipo sistema, utilizando Microsoft VisualFoxPro Drivers direccionado al directorio de la tabla libre.

Como tercer paso se encuentra la creación de los archivos IDC, los cuales especifican los parámetros que contienen información empleada para acceder una base de datos ODBC y el SQL para la ejecución de una consulta determinada. Para la interfase se construyó un archivo IDC para cada una de ellas, los cuales reciben los parámetros enviados por los formularios. Un archivo IDC tiene la siguiente estructura:

Datasource: <nombre del origen de datos>

Template: <nombre del archivo de plantilla htx donde se enviarán los datos>

SQLStatement:

+<estatutos SQL para el acceso a los datos>

El cuarto paso fue crear los formularios de captura de la consulta. Dichos formularios envían parámetros ingresados por el usuario en elementos *select* tipo *multiple*, y controlados por Java Script. Este último realiza dicho control por medio de funciones llamadas por el manejador de eventos *onChange*.

Por último, para el despliegue de los datos se utilizan plantillas HTX para cada una de las consultas, las cuales contienen las etiquetas HTML `<%BeginDetail%>` y `<%EndDetail%>` que proporcionan una estructura de bucle que se ejecutará a través del cursor resultante. Cualquier HTML que esté entre estas dos etiquetas, se genera para cada uno de los registros y los nombres de los campos se pueden incrustar empleando la sintaxis `<%nombre campo%>`.

Conclusiones

La implementación de la solución contribuirá significativamente en la labor de investigación y en el ejercicio del control fiscal, en primer término en el rendimiento y tiempo de análisis y en segundo término en la calidad de los hallazgos derivados de su aplicación. Esto gracias a la interfase de consulta, la cual facilita el análisis multidimensional de los datos, mediante la construcción de tablas de resultados que almacenan la información de las distintas dimensiones

de forma agrupada y utilizando reducción dimensional.

Cabe destacar que la metodología definida por Cortez constituye un aporte para realizar el diseño conceptual y define un rumbo que facilita el desarrollo de la tecnología de DD. La incorporación de algunas consultas generales a través de Internet, por medio del acceso a una tabla agregada, especialmente construida para este fin, se ha enfocado en resolver problemas puntuales de información de aquellos usuarios que requieren acceder el Mercado de Datos de forma remota. Esta facilidad proporciona una mayor cobertura del servicio y acceso; sin embargo, puede ser ampliado incorporando mayor flexibilidad en los formularios y en los programas asociados que efectúan la extracción de información de la tabla agregada. A futuro, es recomendable una conexión directa al Mercado de Datos completo

Para una implementación final se deben incorporar tablas agregadas y registros sumariados para cada una de las combinaciones posibles. Esto, con el fin de permitir una mayor velocidad en el procesamiento de consultas. Además, es necesaria la incorporación de funciones graficadoras y de sumariación a la interfase de consulta, a fin de convertirla en un OLAP relacional y aumentar el potencial de análisis de los usuarios.

Una vez instalado el prototipo y evaluadas sus capacidades y posibilidades de apoyo a la administración, es recomendable su desarrollo futuro sobre tecnologías de base de datos multidimensionales más robustas.

Bibliografía

Cataldo, Joseph. "Care feeding data warehouse". *Database programming and design*. December, 1997. <http://www.dbpd.com/vault/9712toc.htm>.

Chinchilla Arley, Ricardo; Vargas Murillo, José Luis. *Prototipo de mercado de datos para la División de Control, y Fiscalización de la Dirección General de Aduanas*. Cartago. ITCR. 2002.

Corey, Michael; Abbey, Michael. *Oracle Data Warehousing: guía práctica para analizar, construir e implantar con éxito un sistema data warehouse*. Madrid. McGraw-Hill. 1997.

Cortés, Carlos; Meléndez, Jaime; Tobar, Mercedes. "Internet Database Connector". *Interfaz CGI para Servidores Web y Sistemas de Administración de Bases de Datos*. San Salvador: Universidad Centroamericana José Simeón Cañas. 1997. [http:// 168.243.1.4 /investigacion/ bdweb/reportes/ idc.html](http://168.243.1.4/investigacion/bdweb/reportes/idc.html).

Cortez Araniva, Sonia Elizabeth. *Desarrollo de un Modelo para el Diseño de Bases de Datos Multidimensionales*. Cartago. ITCR. 1999.

Cox Alvarado, Alexander. *Propuesta de un álgebra para modelar bases de datos multidimensionales*. Cartago. ITCR. 2000.

González Alvarado, Carlos. *Depósito de datos*. Cartago. ITCR. (1998a).

González Alvarado, Carlos. *Procesamiento analítico en línea*. Cartago. ITCR. 1998b.

Guderian, Dave; Leer, Doug; Molini, Steve. "Keepin the data warehouse in track" *Database programming and design*. January, 1998.

Harinarayan, V.; Rajaraman, A.; Ullman, J. "Implementing data cubes efficiently" SIGMOD'96 6/96, Montreal. Canada. 1996. <http://www-db.stanford.edu/pub/harinarayan/1995/cube.ps>

Herrán Gascón, Manuel de la; Castellar-Busó, Vicent. "Cómo diseñar grandes variables en bases de datos multidimensionales" *Revista digital universitaria*, jul, 2000. <http://www.revista.unam.mx/vol.1/art5/index.html>

Kimball, Ralph. "A dimensional modeling manifesto". *DBMS*. August, 1997.

Madsen, Mark. "Warehouse Design in the Agregate". *Database programming and design*. July, 1996. <http://www.dbpd.com/vault/julytoc.htm>

Oracle Corporation. *Design a database for OLAP*. March/April, 1996. <http://www.Oracle.com/oracle/96-mar/26meth.html>

Parsaye, Kamran. "OLAP and Data Mining:

Una vez instalado el prototipo y evaluadas sus capacidades y posibilidades de apoyo a la administración, es recomendable su desarrollo futuro sobre tecnologías de base de datos multidimensionales más robustas.

- Bridging the Gap". *Database programming design*. February, 1996. <http://www.dbpd.com/vault/parsfeb.htm>
- Parsaye, Kamran "Surveying Decision Support: New Realms of Analysis". *Database programming desing*. April, 1996.
- Pérez Martínez, Cayetano. *Replicación de bases de datos por Internet*. México: Instituto Politécnico Nacional, 1999. <http://148.204.20.3/PROYECTOS/TESIS/cayetano/DTeMae.htm>
- Poe, Vidette. *Building a data warehouse for decision support*. New Jersey: Prentice-Hall, 1998.
- Reese, Joseph. "Making the datawarehouse work overtime". *Database programming design*. September, 1997.
- Simon, Alan. "Beyond the warehouse". *Database programming desing*. december, 1996.
- Weldon, Jay; Weldon, Louise. "Managing multidimensional data". *Database programming and design*. august., 1995.
- Weldon, Jay; Weldon, Louise. "Choosing Tools for Multidimensional Data". *Database programming and design*. February, 1996.
- Wolff, Carmen. "Implementando un DataWarehouse". *Revista Ingeniería Informática: revista electrónica del DIICC*. Edición 5, año 3 [2000]. <http://www.inf.udec.cl/revista/edicion5/cwolff.htm>