
IBM DB2 OLAP SERVER



Procesamiento Analítico en Línea

IBARRA MARIA DE LOS ANGELES

2006

Índice

- ❑ Introducción
 - ❑ Data warehouse
 - ❑ Data mart
 - ❑ Sistemas OLTP
 - ❑ Sistemas OLAP
 - ❑ Operaciones analíticas básicas
 - ❑ Vista de datos
 - ❑ Modelo de datos
 - ❑ Utilidades de OLAP
 - ❑ Beneficios de OLAP
-

Índice

- ❑ Presentación del producto
 - ❑ Centro de depósitos
 - ❑ Definir área temática
 - ❑ Definir fuente de depósito
 - ❑ Definir destino de depósito
 - ❑ Definir movimiento y transformación de datos
 - ❑ Definir esquema estrella
 - ❑ Análisis modelo Multidimensional
 - ❑ Creación del modelos OLAP
 - ❑ Creación de un metaperfil
 - ❑ Creación de una Aplicación OLAP
-

Introducción

Estructura Multidimensional

En contraste con las bases de datos relacionales, las multidimensionales están compuestas por dimensiones que son atributos estructurales de un cubo, organizadas con jerarquías de categorías que describen los datos en tablas.

Por ejemplo se tienen las ventas de cada producto por región. Una compañía tiene tres productos (arandelas, tornillos, tuercas) que se venden en tres territorios (Este, Oeste, Central).

Introducción

Producto	Region #	Ventas
Arandelas	Este	50000
Arandelas	Oeste	60000
Arandelas	Central	100000
Tornillos	Este	40000
Tornillos	Oeste	70000
Tornillos	Central	80000
Tuercas	Este	90000
Tuercas	Oeste	120000
Tuercas	Central	30000

	Este	Oeste	Central
Arandelas	50000	60000	100000
Tornillos	40000	70000	80000
Tuercas	90000	120000	140000

Un camino para representar esta tabla en una forma mas óptima es a través de una matriz de dos dimensiones.

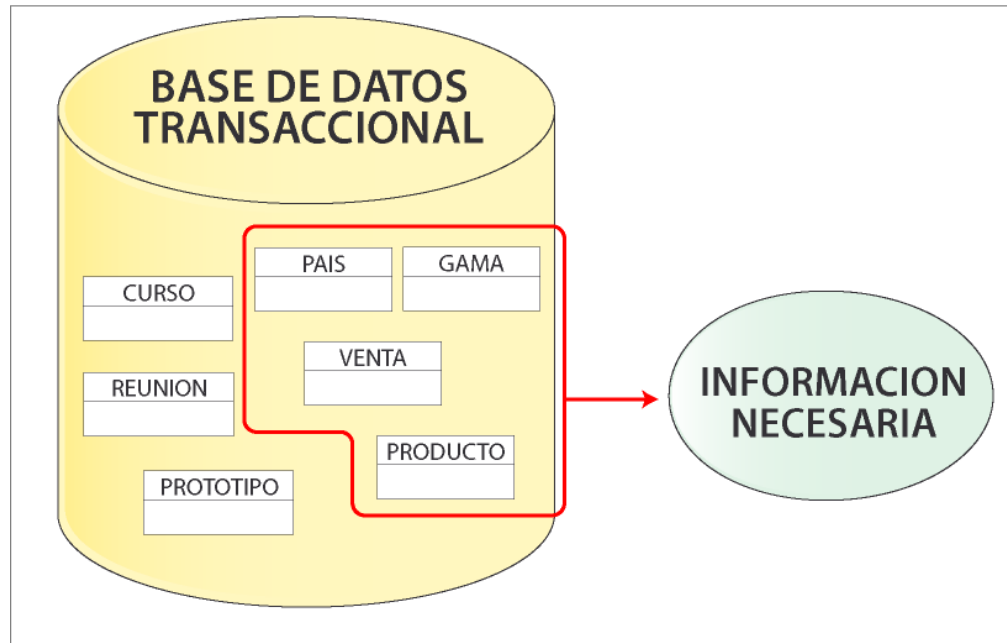
De esta forma se pueden realizar preguntas como ¿Cuáles fueron las ventas de arandelas en el Este?, ¿Cuáles fueron las ventas de Tornillos en el Oeste?.

Data Warehouse

Definición: Conjunto de datos orientados a temas integrado, no volátil de tiempo variante que se usa para el soporte del proceso de toma de decisiones gerenciales.

- ❑ **Orientado hacia la información relevante de la Organización**
Se diseña para consultar eficientemente información relativas a las actividades básicas de la organización como ser, comparas, ventas, etc.
-

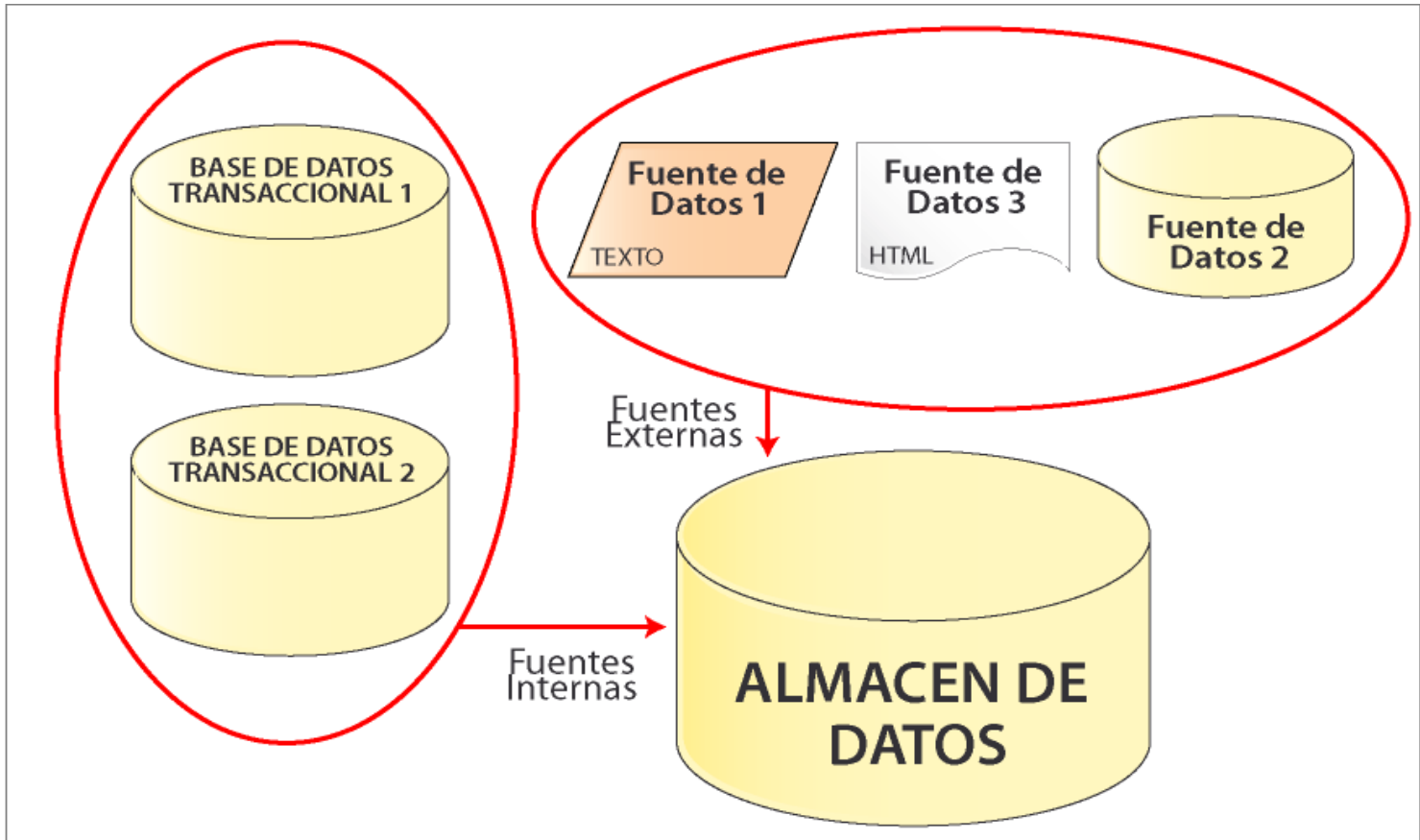
Data Warehouse



□ Datos Integrados

Integra datos recolectados de diferentes sistemas operacionales de la organización y o fuentes externas.

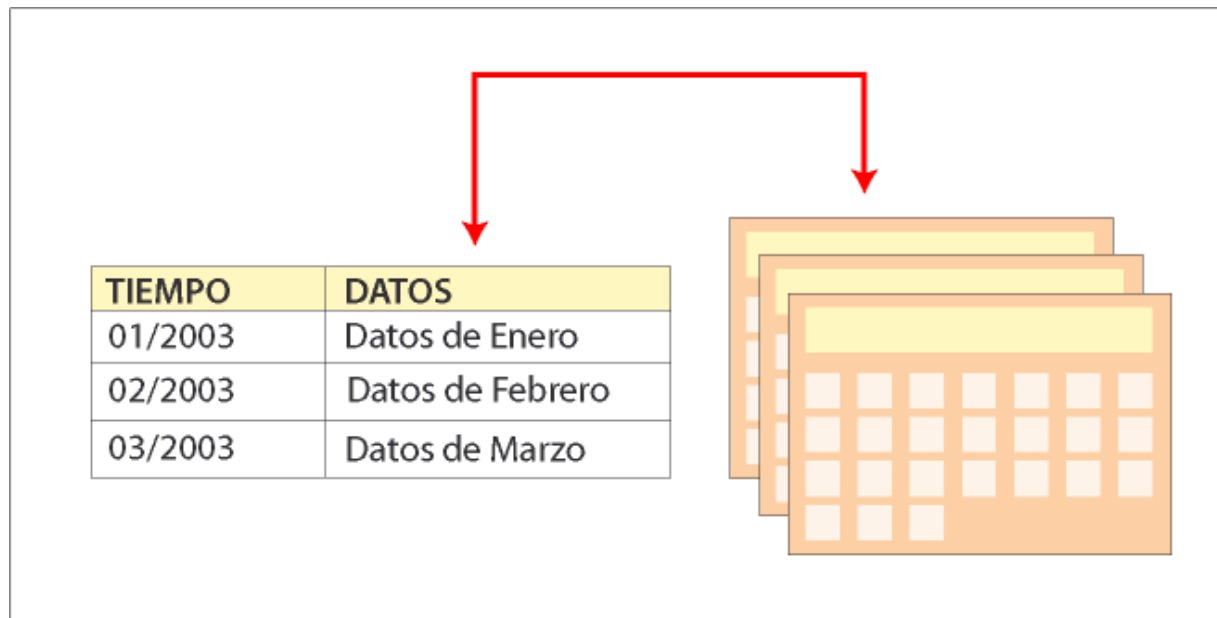
Data Warehouse



Data Warehouse

Variable en el Tiempo

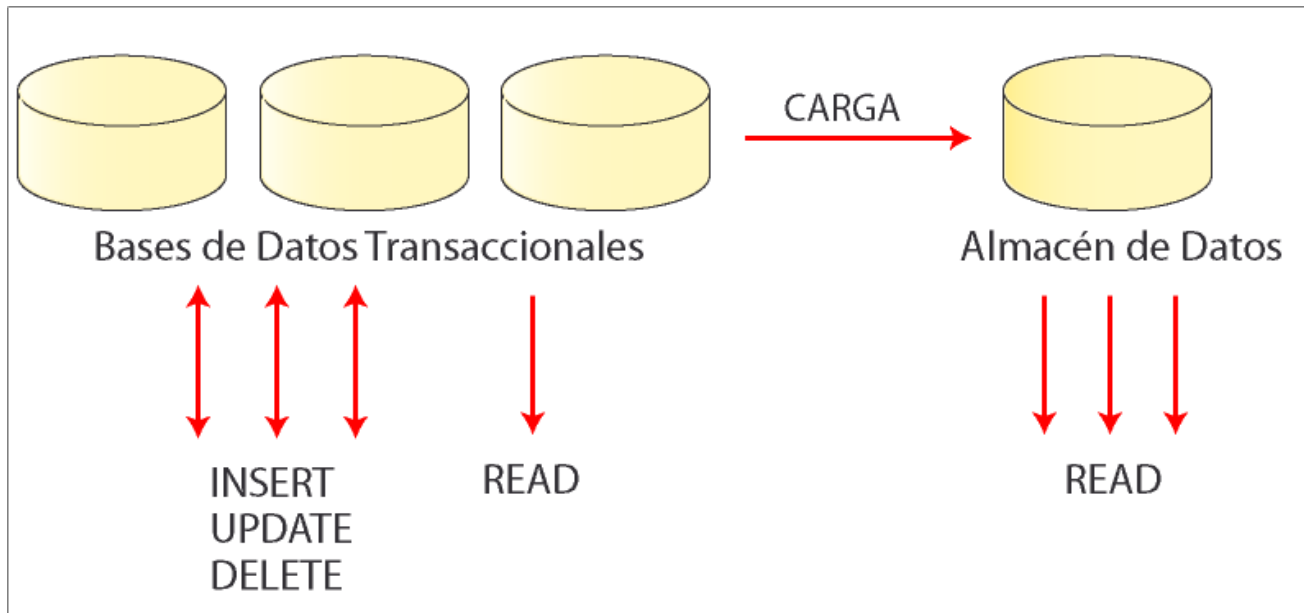
Los datos son relativos a un periodo de tiempo y estos deben ser integrados periódicamente, los mismos son almacenados como fotos que se corresponden a un periodo de tiempo.



Data Warehouse

- ❑ **No Volátil**

Los datos que son almacenados no sufren ninguna actualización solo son incrementados. El periodo cubierto para un DW va de 2 a 10 años.



Data Marts

Problema: al crecer el D.W. El rendimiento de las consultas decae y el modelo centralizado deja de ser óptimo.

Solución: crear almacenes de datos especializados por área, que reciben los datos desde el almacén centralizado, estos almacenes se conocen como Data Marts.

Ventajas: al usar menos usuarios se lo puede optimizar para recuperar más rápidamente los datos que necesitan.

Existen dos tipos de Data Marts:

- Dependientes: los datos son extraídos del D.W.
 - Independientes: los datos son extraídos de los sistemas operacionales.
-

Sistemas OLTP

Definición: procesan las transacciones en tiempo real de un negocio.

Desventaja: proporciona capacidades limitadas para la toma de decisiones.

Características	OLTP	OLAP
Tamaño BBDD	Gigabytes	Gigabytes a Terabytes
Origen de Datos	Interno	Interno Externo
Actualización	Actual	Histórico
Consultas	Predecible	Ad Hoc
Actividad	Operacional	Analítica

Sistemas OLAP

Definición: los datos son clasificados en diferentes dimensiones y pueden ser vistas unas con otras en diferentes combinaciones para obtener diferentes análisis de los datos que contienen.

- ❑ En este modelo los datos son vistos como cubos los cuales consisten en categorías descriptivas (dimensiones) y valores cuantitativos (medidas).
 - ❑ El modelo multidimensional de datos simplifica a los usuarios formular consultas complejas, arreglar datos en un reporte, cambiar de datos resumidos a datos detallados, etc.
-

Operaciones Analíticas Básicas

- ❑ **Consolidación:** este comprende el conjunto de datos. Esto puede involucrar acumulaciones simples o agrupaciones complejas que incluyen datos interrelacionados.
 - ❑ **Drill-Down:** OLAP puede moverse en la dirección contraria y presentar automáticamente datos detallados que abarcan datos consolidados.
 - ❑ **Slicing and Dicing:** se refiere a la capacidad de visualizar a la BD desde diferentes puntos de vista.
-

Operaciones Analíticas Básicas

VENTAS			
	Productos	Store 1	Store 2
Q1	Electronics	\$5,20	\$5,60
	Toys	\$1,90	\$1,40
	Clothing	\$2,30	\$2,60
	Cosmetics	\$1,10	\$1,10
Q2	Electronics	\$8,90	\$7,20
	Toys	\$0,75	\$0,40
	Clothing	\$4,60	\$4,60
	Cosmetics	\$1,50	\$0,50

SLICE y DICE

VENTAS		
	Productos	Store 1
Q1	Electronics	\$5,20
Q1	Toys	\$1,90
Q2	Electronics	\$8,90
Q2	Toys	\$0,75

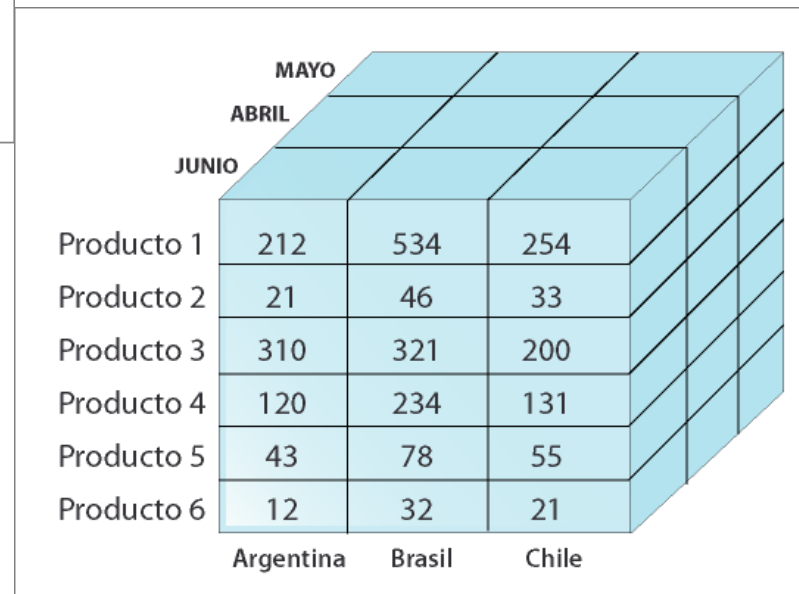
Vista de Datos

La vista de datos como cubos es una extensión de la manera normal en que los usuarios de negocios interactúan con los datos.

Por Ejemplo: la mayoría de los usuarios desearía ver como se desarrollan las ventas a lo largo del tiempo . Para ello se necesitaría ver varias planillas de cálculo.

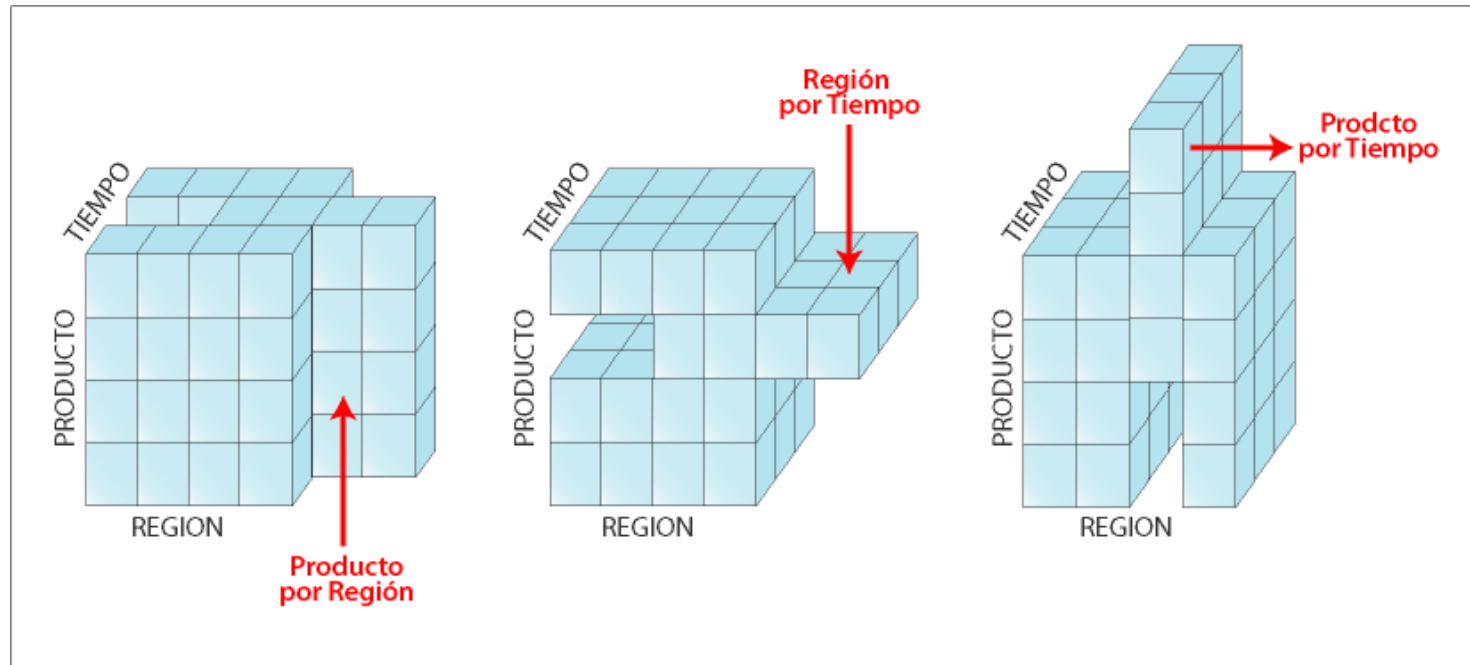
Vista de Datos

ABRIL				
MAYO				
JUNIO				
REGION				
	Argentina	Brasil	Chile	
PRODUCTOS	Producto 1	212	534	254
	Producto 2	21	46	33
	Producto 3	310	321	200
	Producto 4	120	234	131
	Producto 5	43	78	55
	Producto 6	12	32	21



Vista de Datos

Debido a su representación pueden ser tomadas rebanadas de datos de las mismas, para responder diversas preguntas.



Modelos de Datos

- ❑ **MOLAP:** en estos sistemas los datos se encuentran almacenados en una estructura de datos multidimensional (OLAP Multidimensional).

Ventajas:

- Mayor performance en el procesamiento de queries.
- Poco tiempo de cálculo realizado en el momento.
- Puede escribir sobre la base de datos.
- Posibilita hacer cálculos más complicados.

Desventajas:

- Tamaño limitado para la arquitectura del cubo.
- No puede acceder a datos que no están en el cubo.
- No puede explotar el paralelismo en de las Bases de Datos.

Modelos de Datos

- ❑ **ROLAP:** son sistemas en los cuales los datos se encuentran almacenados en una base de datos relacional (OLAP Relacional).

Ventajas:

- Uso total de la seguridad e integridad de la base de datos .
- Escalable para grandes volúmenes.
- Los datos puede ser compartida con aplicaciones SQL.
- Datos y estructura más dinámicos.

Desventajas:

- Queris más lentos.
- Construcción cara.
- Los cálculos están limitados a las funciones de las bases de datos.

Modelos de Datos

- ❑ **HOLAP:** estos sistemas mantienen los registros detallados en la base de datos relacional, mientras que los datos resumidos o agregados se almacenan en una base de datos Multidimensional separada.

Son conocidos como Híbridos debido a que contiene las características de los sistemas anteriores.

Comparaciones

Con MOLAP: los datos son precalculados y luego son almacenados en cubos de datos multidimensionales.

Con ROLAP: los datos son accedidos directamente desde el D.W. u otra fuente de datos relacional y no son almacenados por separado.

Con HOLAP: Mantiene los volúmenes de datos más grande en la base de datos relacional y las agregaciones en una base de datos MOLAP separada.

Utilidades

- ❑ Tienen acceso a grandes cantidades de datos.
 - ❑ Analizan las relaciones entre muchos tipos de elementos empresariales.
 - ❑ Involucran datos agregados.
 - ❑ Comparan datos agregados a través de periodos jerárquicos.
 - ❑ Presentan los datos en diferentes perspectivas.
 - ❑ Involucran cálculos complejos entre elementos de datos.
 - ❑ Pueden responder con rapidez a consultas de usuarios.
-

Beneficios

- ❑ Es de fácil uso y acceso flexible para los usuarios.
 - ❑ Los datos están organizados en diferentes dimensiones lo que permite un mejor análisis.
 - ❑ Ahorro generado por productividad de personas altamente profesionales.
 - ❑ Permite encontrar la historia en los datos.
 - ❑ Genera ciertas ventajas competitivas.
-

Presentación del Producto

Vamos a realizar los siguientes pasos:

CREAR UN DEPOSITO DE DATOS

- ❑ Definición de área temática.
 - ❑ Definición de fuente de depósito.
 - ❑ Definición de destino de depósito.
 - ❑ Definición de movimiento y transformación de datos.
 - ❑ Definición de proceso.
 - ❑ Definición de esquema estrella.
 - ❑ Exportación del esquema estrella.
-

Presentación del Producto

ANALISIS MODELO MULTIDIMENSIONAL

- ❑ Creación del modelo OLAP.
 - ❑ Selección de la tabla de hechos.
 - ❑ Creación de dimensiones.
 - ❑ Unión de las tablas de mediciones.
 - ❑ Creación de Herencia.
 - ❑ Guardar modelo OLAP.
 - ❑ Creación un metaperfil.
 - ❑ Iniciar asistente.
 - ❑ Seleccionar un modelo olap en el que basar el metaperfil.
 - ❑ Creación de una aplicación OLAP.
-

Definición de un Área Temática

En primer lugar se debe crear una BD para que contenga los datos para el depósito. Luego, esta debe ser registrada con el ODBC.

Se utiliza el Centro de depósito de datos para definir un área temática.

Un *área temática* identifica y agrupa procesos relativos a un área lógica de la empresa.

Definición de un Área Temática



The image shows a software dialog box titled "Definir área temática" (Define thematic area). The dialog box has a blue title bar with a close button (X) in the top right corner. The main content area is light beige and contains the following fields:

- Área temática:** A label with a vertical line next to it, indicating the current field.
- Nombre:** A text input field containing the text "CURSO".
- Administrador:** A text input field containing the text "marian".
- Descripción:** A text input field containing the text "Creacion de un AREA TEMATICA".
- Notas:** A larger, empty text input field.

At the bottom of the dialog box, there are three buttons: "Bien" (OK), "Cancelar" (Cancel), and "Ayuda" (Help).

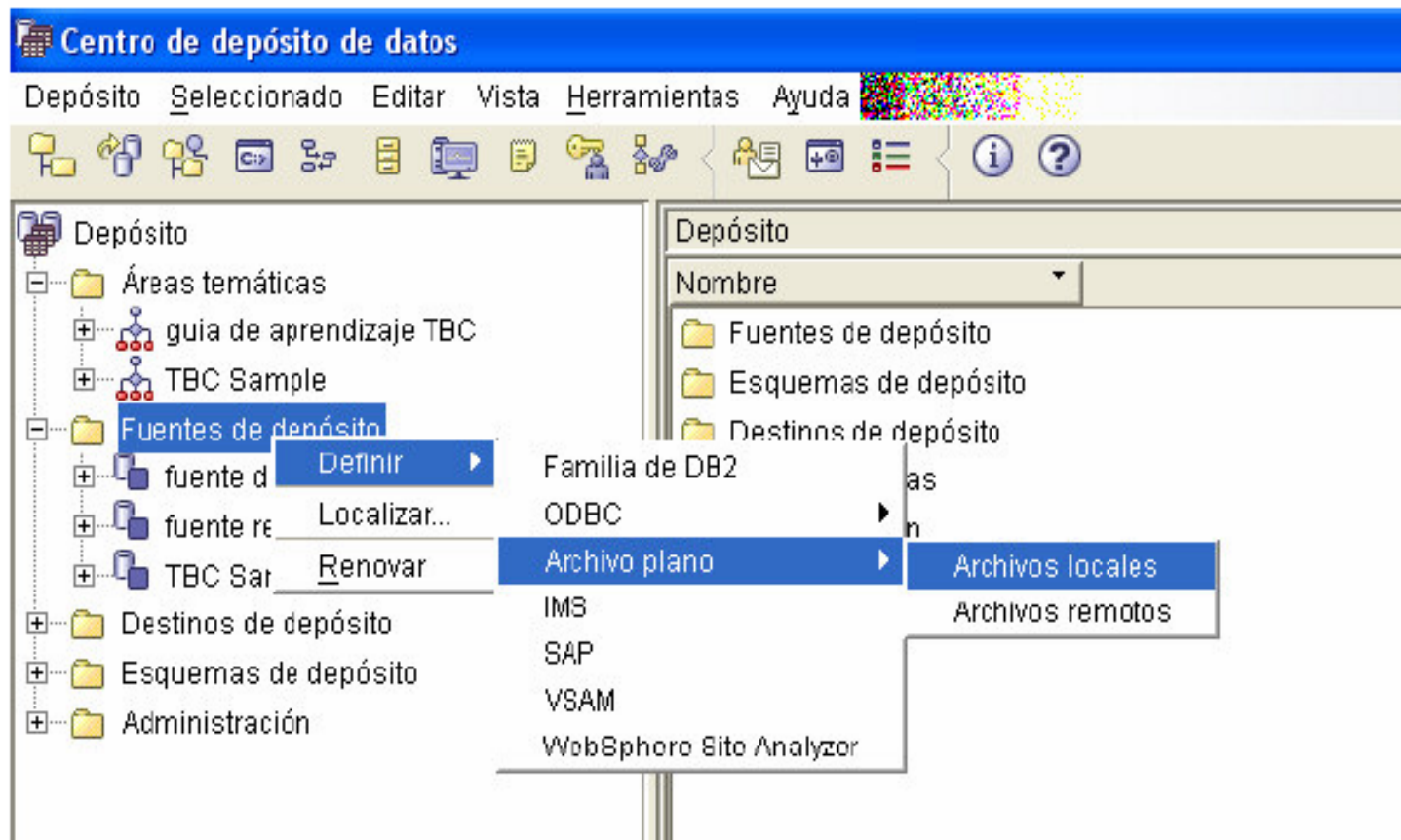
Definición de Fuente de Depósito

Las *fuentes de depósito* son las definiciones lógicas de las tablas y archivos que proporcionarán datos a la tabla de mediciones.

El centro de depósito de datos utiliza las especificaciones de las fuentes de depósito para acceder a los datos y seleccionarlos.

Definición de Fuente de Depósito

Se puede utilizar más de una fuente en un proceso.



Definición de Fuente de Depósito

Definir fuente de depósito

Fuente de depósito nuevo

Fuente de depósito | Sitios de agente | Base de datos | Tablas y vistas | Seguridad

Nombre: Fuente Relacional de la Guia de Aprendizaje

Tipo de fuente de depósito: Familia de DB2

Administrador: Marian

Descripción: Datos relacionales para la empresa TBC

Notas:

Bien Cancelar Ayuda

Definición de Destino de Depósito

Los *Destinos de depósito* identifican la base de datos y las tablas que el Centro de depósito de datos debe utilizar para el depósito de datos.

Normalmente, las tablas de destino que se definen en el destino de depósito también se utilizan como tablas de mediciones y de hechos del esquema en estrella.

Sin embargo, el destino de depósito puede incluir también tablas de destino intermedias que se utilizan para la transformación de datos.

Definición de Destino de Depósito

Definir destino de depósito

Destino de depósito nuevo

Destino de depósito | Sitios de agente | Base de datos | Tablas | Seguridad

Nombre: Destino de ejemplo

Tipo de destino de depósito: Familia de DB2

Administrador: marian

Descripción:

Notas:

Bien Cancelar Ayuda

Definición de Movimiento y Transformación de Datos

Se debe definir el modo en que el Centro de depósito de datos debe mover y transformar datos para darles un formato adecuado para el depósito de los mismos.

En primer lugar, se define un proceso que contenga una serie de pasos del proceso de transformación y movimiento de datos.

Definición de Movimiento y Transformación de Datos

Definir proceso

guía de aprendizaje TBC - Proceso nuevo

Proceso | Seguridad

Nombre: Crear Dimensión de Mercado

Administrador: marian

Descripción:

Notas:

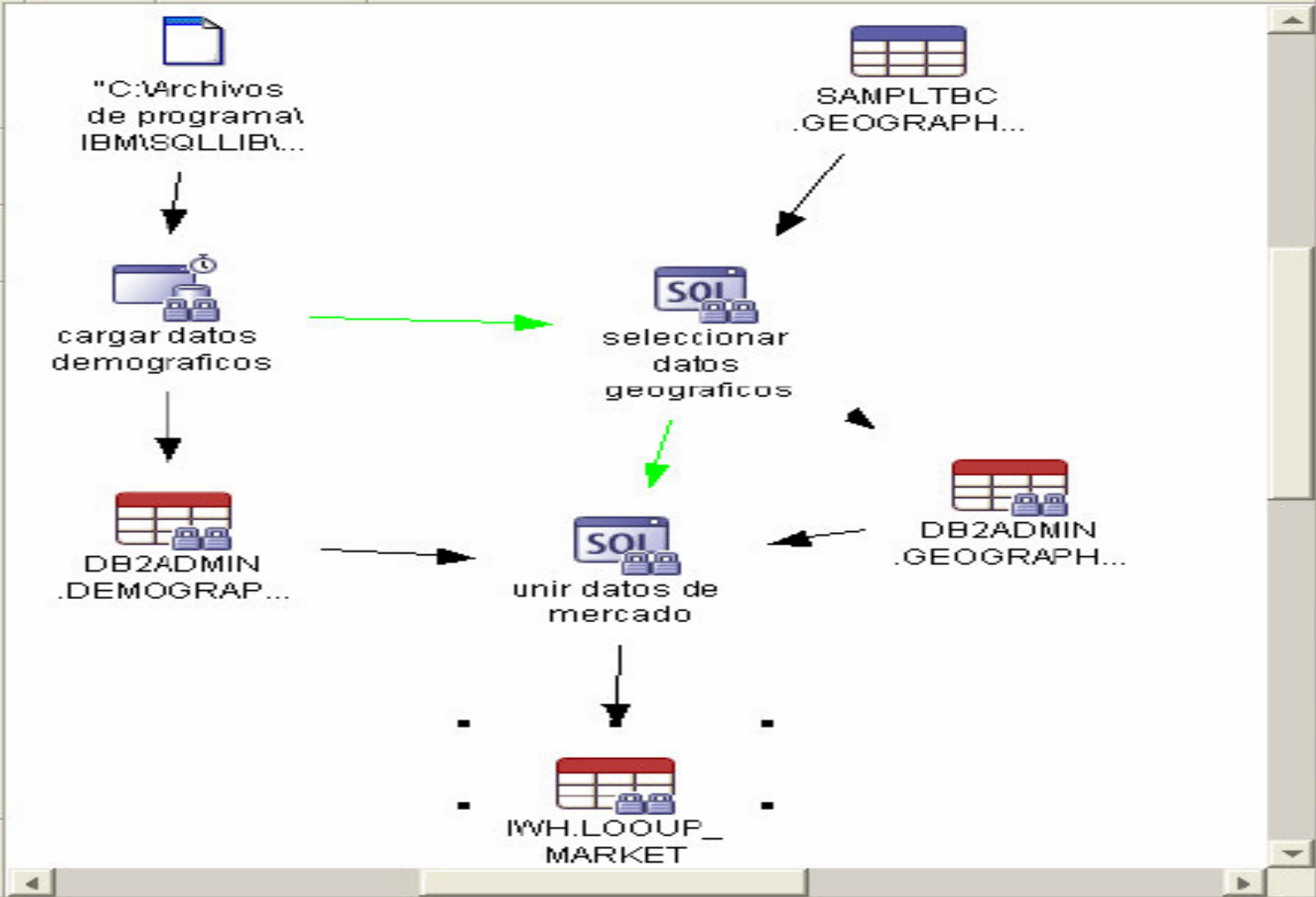
Bien Cancelar Ayuda

Definición de Movimiento y Transformación de Datos

Como primer medida se debe especificar las tablas fuente que deben transformarse para el depósito.

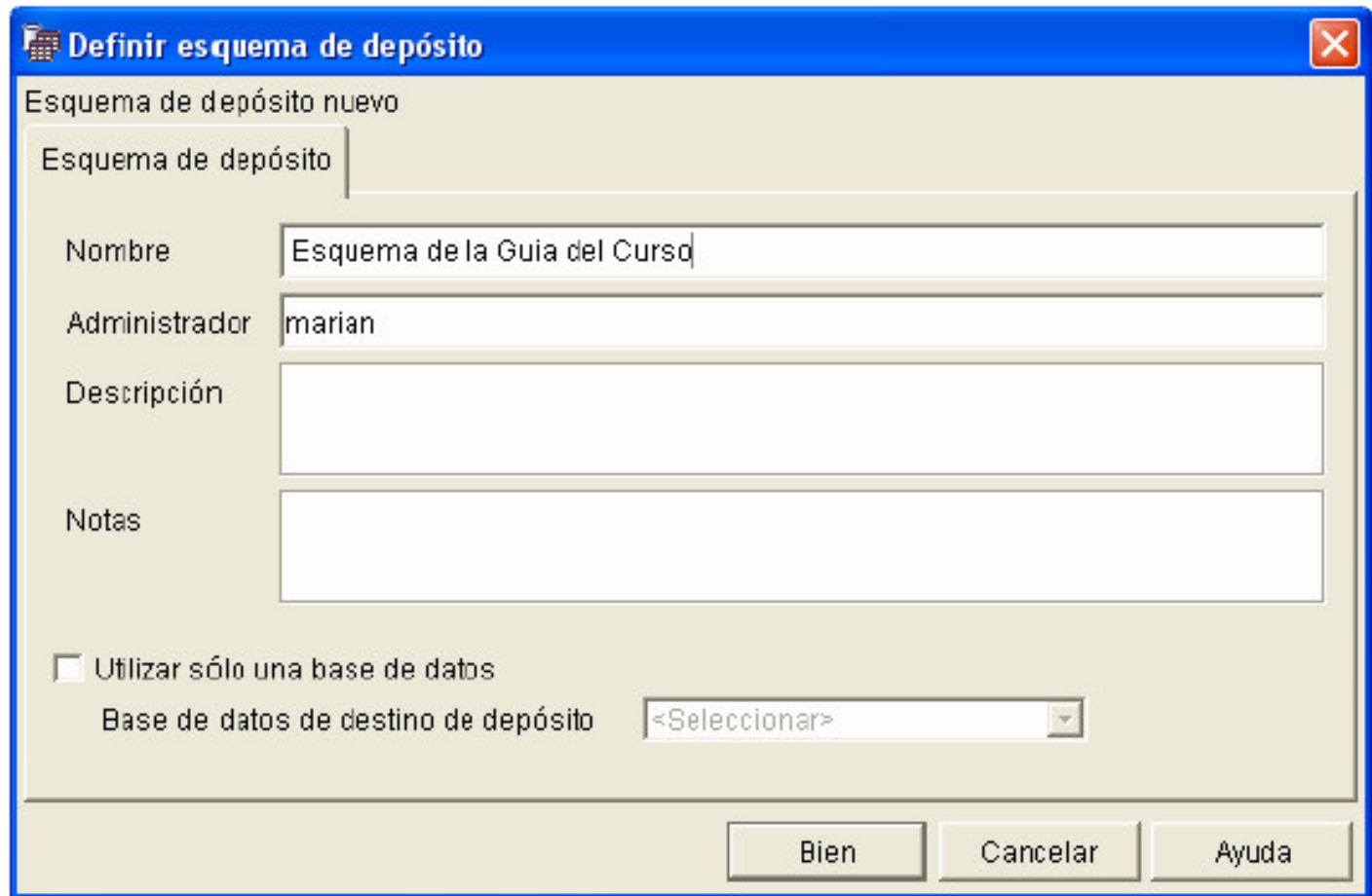
Finalmente se debe definir los pasos de transformación de datos que utilizan dos métodos de transformación diferentes:

- ❑ Cargar datos en la base de datos de depósito con un programa.
 - ❑ Seleccionar datos fuente y unir tablas con sentencias de SQL.
-



Definición de un Esquema en Estrella

Se puede crear un esquema en estrella desde la tablas de depósito.

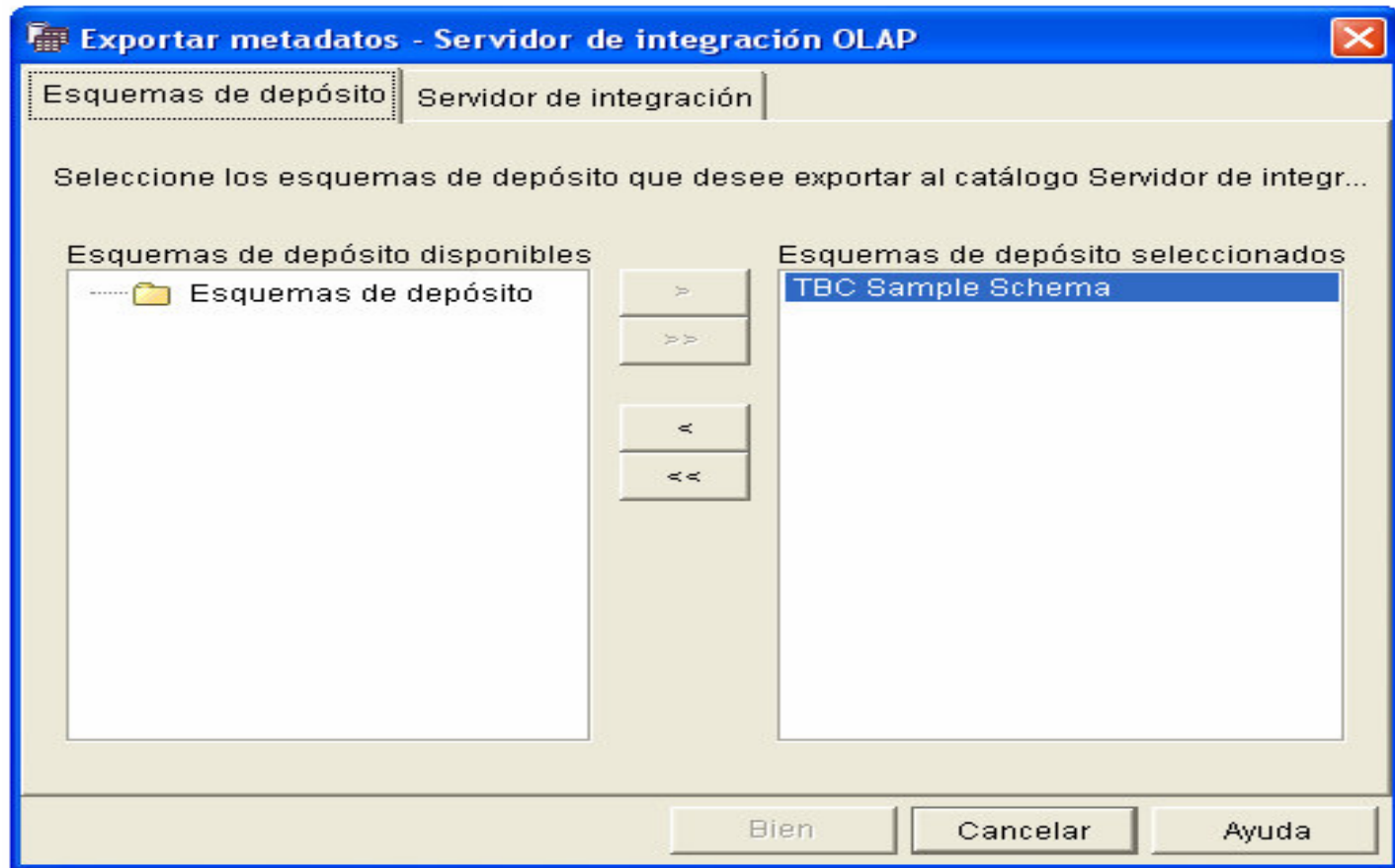


The image shows a Windows-style dialog box titled "Definir esquema de depósito" (Define deposit schema). The dialog has a blue title bar with a close button (X) in the top right corner. The main content area is light beige and contains the following elements:

- A tab labeled "Esquema de depósito" (Deposit schema).
- A "Nombre" (Name) text box containing "Esquema de la Guia del Curso" (Course Guide Schema).
- An "Administrador" (Administrator) text box containing "marian".
- A "Descripción" (Description) text box, currently empty.
- A "Notas" (Notes) text box, currently empty.
- An unchecked checkbox labeled "Utilizar sólo una base de datos" (Use only one database).
- A "Base de datos de destino de depósito" (Deposit destination database) label next to a dropdown menu showing "<Seleccionar>" (Select).
- At the bottom, three buttons: "Bien" (OK), "Cancelar" (Cancel), and "Ayuda" (Help).

Exportar Esquema Estrella

Luego se puede exportar el esquema en estrella a OLAP Integration Server para crear una base de datos OLAP.



Análisis Multidimensional

Dentro del Kit de iniciación de OLAP de DB2, la herramienta primaria para crear aplicaciones OLAP es el DB2 OLAP Integration Server, que ejecuta en el servidor Essbase multidimensional.

Con estas aplicaciones, los usuarios pueden analizar los datos de DB2 utilizando Lotus 1-2-3 o Microsoft Excel.

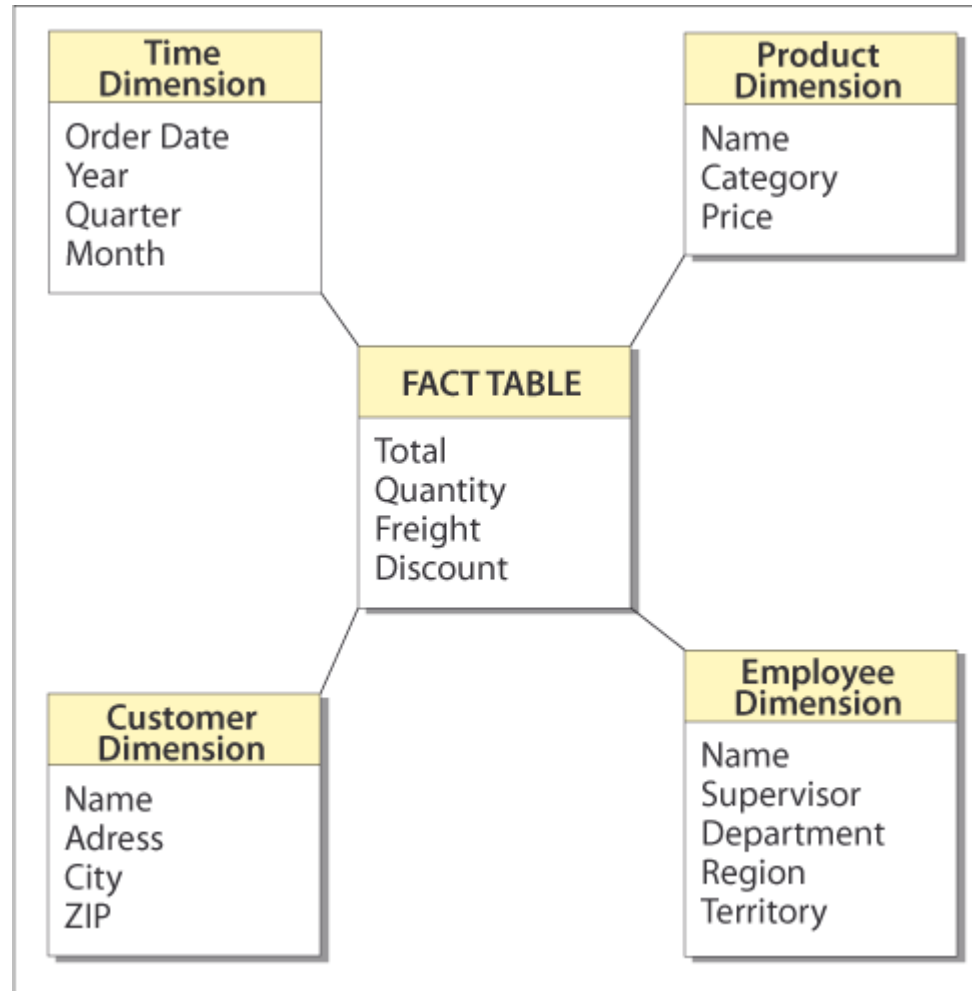
Creación de un Modelo OLAP

Es una estructura lógica que describe el plan comercial general. El modelo toma la forma de un *esquema en estrella*, que representa la relación entre sus componentes con una estructura en estrella.

En el centro del esquema en estrella se encuentra una *tabla de hechos*, que contiene los datos reales que desea analizar, como las cifras de ventas de productos.

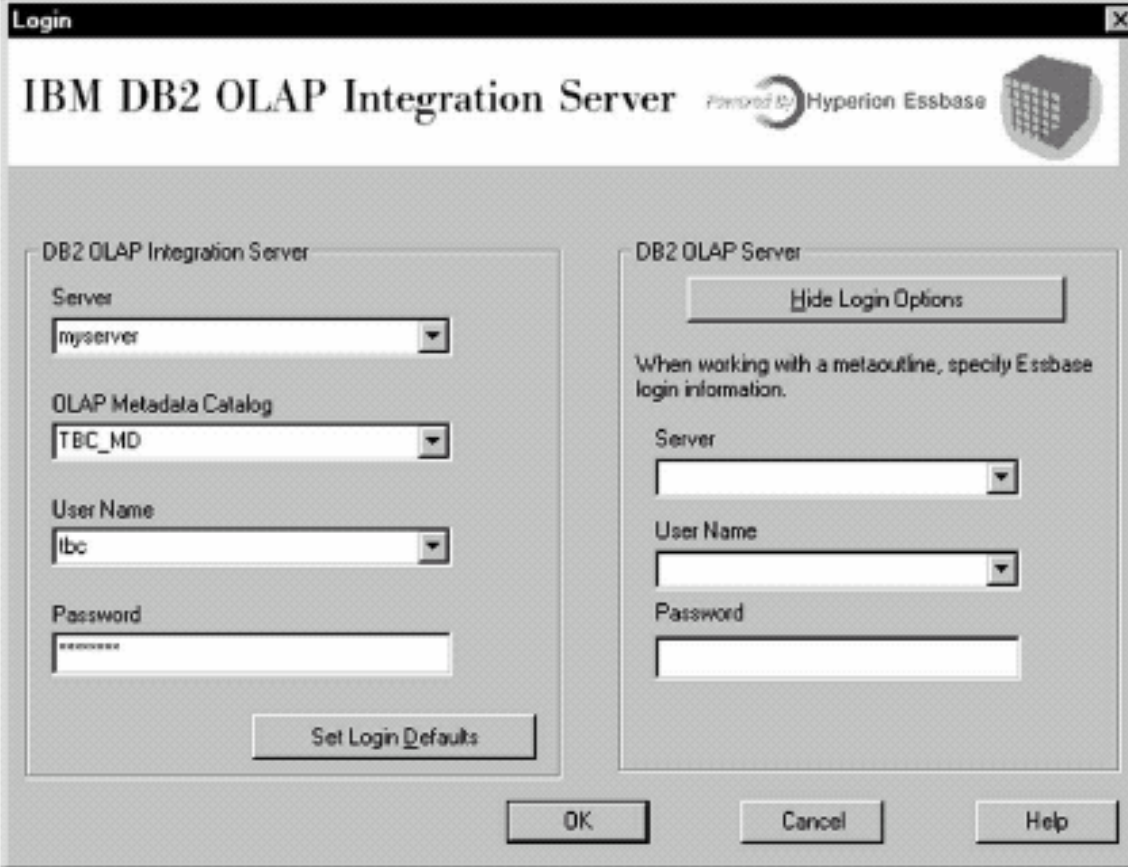
De la tabla de hechos salen las *tablas de mediciones* que contienen los datos que definen las dimensiones OLAP, como por ejemplo números de cuenta, meses, nombres de productos, etc.

Creación de un Modelo OLAP



Creación de un Modelo OLAP

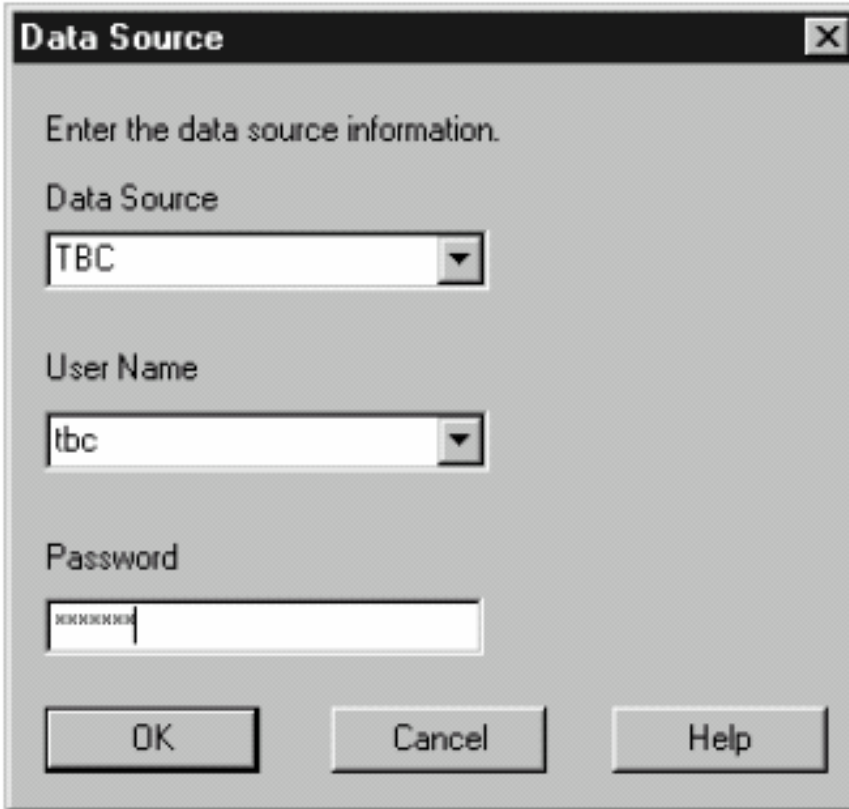
Una vez iniciada la interfaz del escritorio de Integration Server, se debe conectar a los metadatos del catálogo OLAP.



The screenshot shows a 'Login' dialog box for the 'IBM DB2 OLAP Integration Server'. The window title is 'Login'. The main title is 'IBM DB2 OLAP Integration Server' with a 'Powered by Hyperion Essbase' logo and a cube icon. The dialog is divided into two main sections: 'DB2 OLAP Integration Server' and 'DB2 OLAP Server'. The 'DB2 OLAP Integration Server' section contains four dropdown menus: 'Server' (set to 'myserver'), 'OLAP Metadata Catalog' (set to 'TBC_MD'), 'User Name' (set to 'lbc'), and 'Password' (masked with asterisks). Below these is a 'Set Login Defaults' button. The 'DB2 OLAP Server' section has a 'Hide Login Options' button and a text instruction: 'When working with a metaoutline, specify Essbase login information.' Below this are three input fields: 'Server', 'User Name', and 'Password'. At the bottom of the dialog are 'OK', 'Cancel', and 'Help' buttons.

Creación de un Modelo OLAP

Luego de esto se debe iniciar el asistente de modelos OLAP y conectar a la base de datos de DB2 que contiene la fuente de datos.



Data Source [X]

Enter the data source information.

Data Source
TBC

User Name
tbc

Password
XXXXXXXXXX

OK Cancel Help

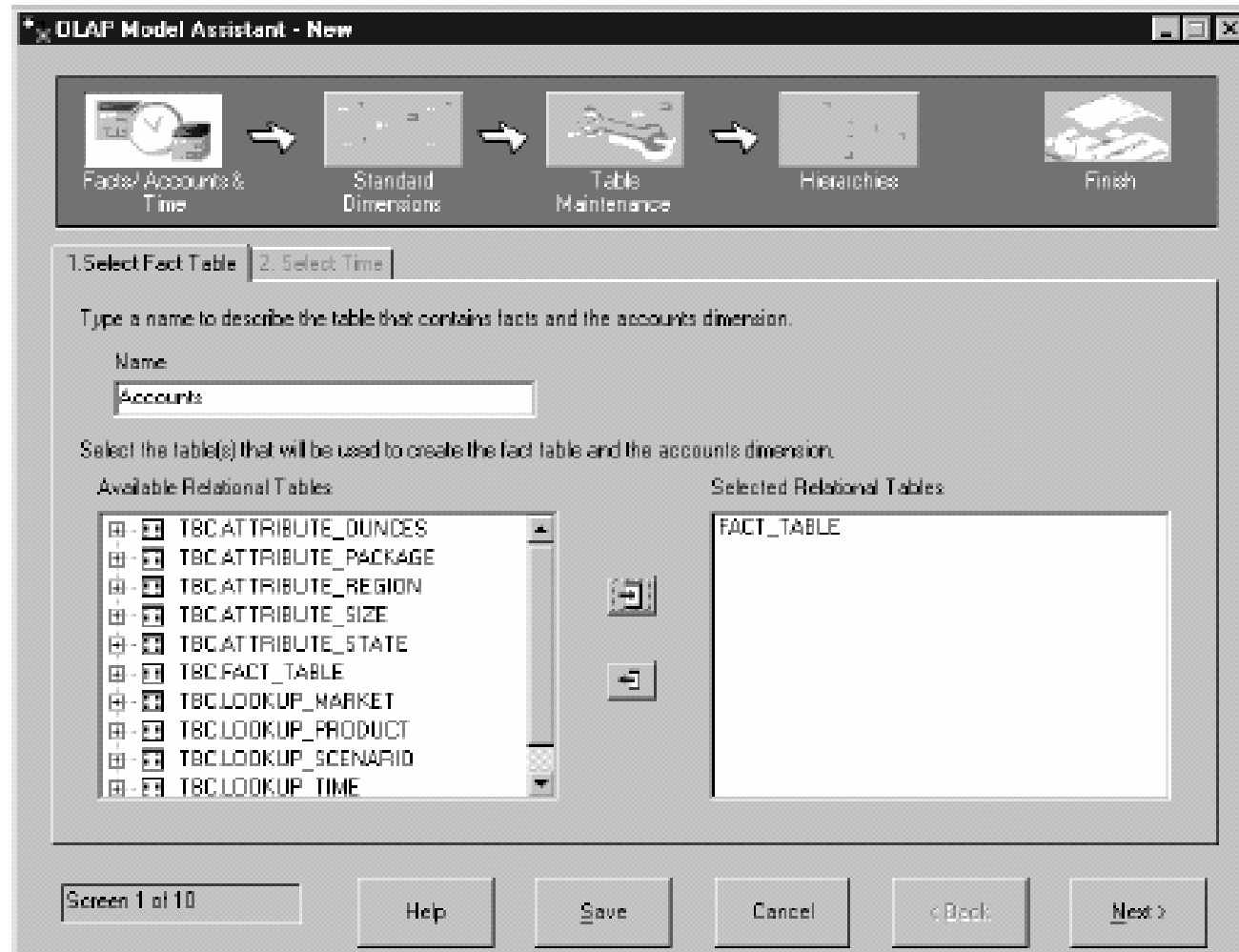
Selección de la Tabla de Hechos

Cada modelo necesita una tabla de hechos, que es el centro del esquema en estrella.

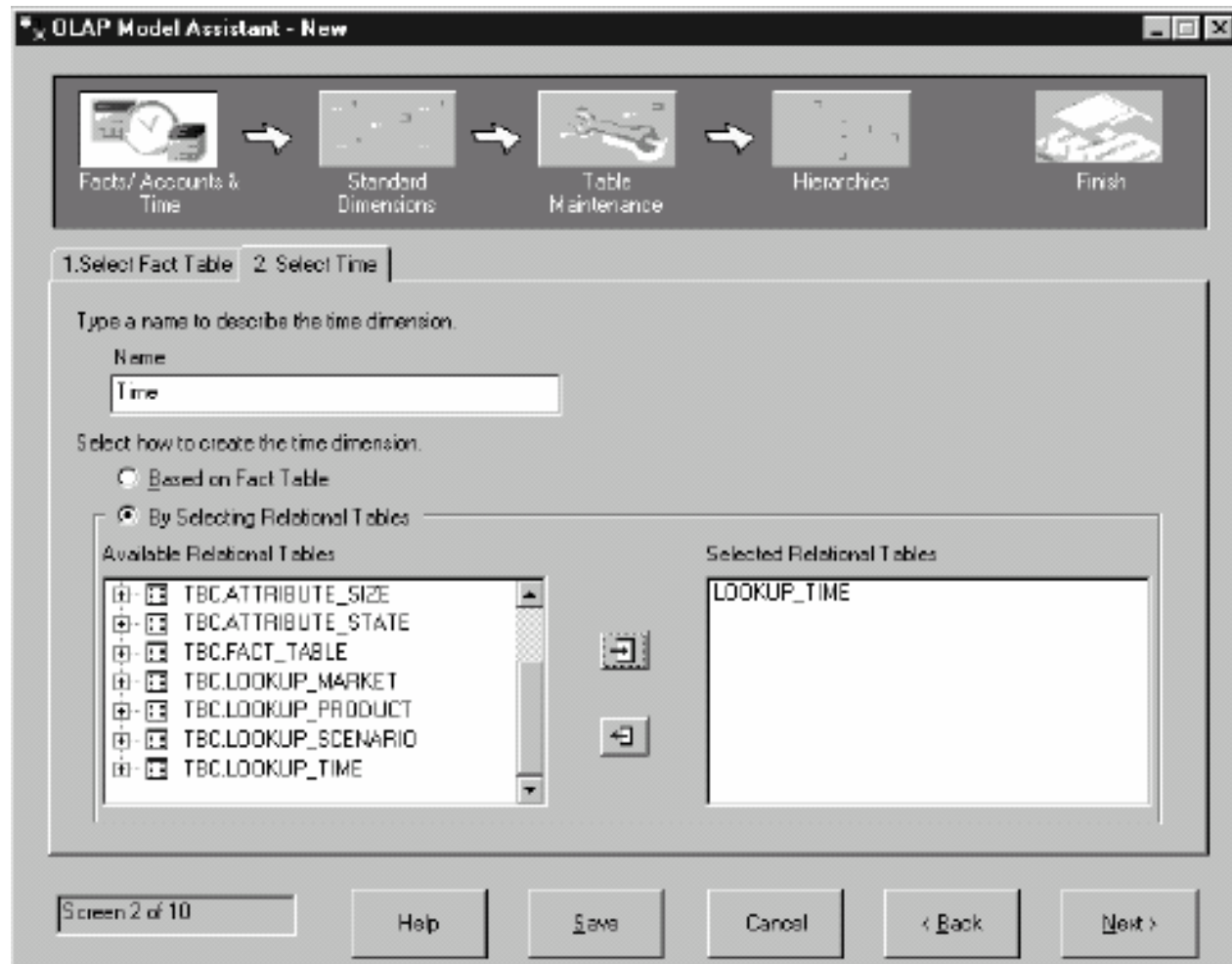
La dimensión Accounts contiene los datos numéricos que desea analizar y de los cuales desea hacer un seguimiento a lo largo del tiempo, como por ejemplo, ventas o inventario.

Estos datos también reciben el nombre de datos de *medidas* variables y permiten una contabilidad inteligente durante el proceso analítico en línea.

Selección de la Tabla de Hechos



Creación de las Dimensiones



Unión de la Tabla de Mediciones

El esquema en estrella representa las relaciones entre la tabla de hechos y las otras dimensiones del modelo. Se puede definir la estructura del esquema en estrella mediante uniones entre las tablas de mediciones y la tabla de hechos.

La parte izquierda de la página Fact Table Joins lista todas las dimensiones del modelo.

La parte derecha muestra las columnas que se unen entre las tablas de mediciones y la tabla de hechos, si existe una unión.

Unión de la Tabla de Mediciones

OLAP Model Assistant - New

Facts/Accounts & Time → Standard Dimensions → Table Maintenance → Hierarchies → Finish

1. Fact Table Joins | 2. Dimension Table Joins | 3. Edit Tables

If necessary, join each primary dimension table to the fact table. In the dimension list, select a dimension. In the join detail box, select a column in the dimension table and a column in the facts table.

Dimension	Joined
Accounts	X
Time	X
Scenario	X
Product	X
Market	X

Join Detail

Time	Facts
TIME_ID	TRANSDATE
SEASON	TIME_ID
QUARTER_ID	SCENARIO_ID
QUARTER	SALES
MONTH_ID	PRODUCT_KEY
MONTH	PAYROLL
	OPENING_INVENTORY
	MISC
	MARKETING
	ENDING_INVENTORY
	COGS
	CITY_ID
	ADDITIONS

Screen 5 of 10 | Help | Save | Cancel | < Back | Next >

Guardar el Modelo OLAP

OLAP Model Assistant - Modify OLAP tutorial lesson

Facts/Accounts & Time → Standard Dimensions → Table Maintenance → Hierarchies → Finish

You have provided the necessary information to define an OLAP model.

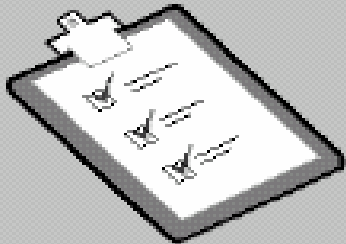
Click the button below to preview the star schema.

Star Schema...

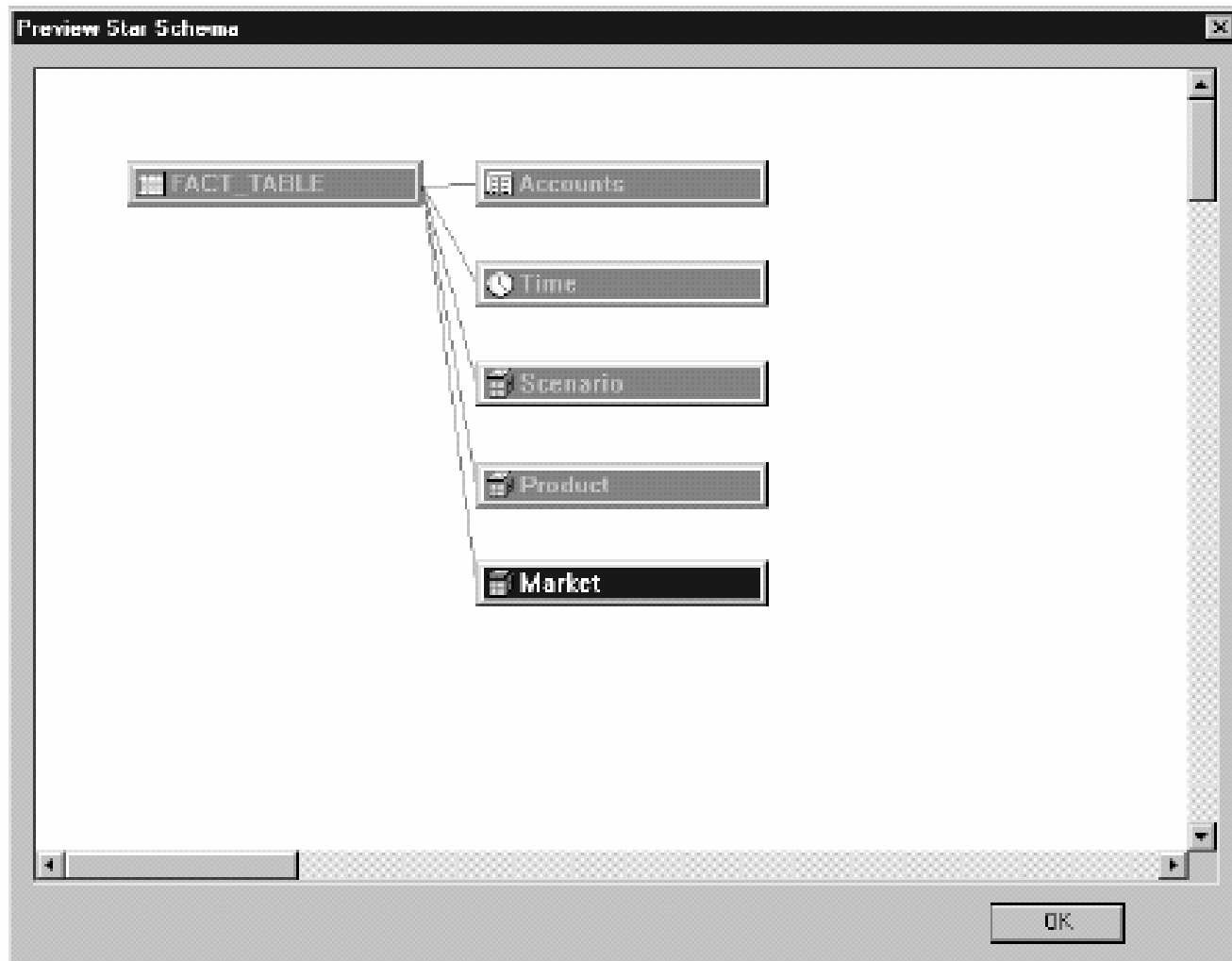
Launch the OLAP Metamodel Assistant after saving this OLAP model.

Click the Finish button to complete and save this OLAP model.

Screen 10 of 10 Help Save Cancel < Back Finish



Guardar el Modelo OLAP



Creación de un Metaperfil OLAP

Normalmente es un subconjunto del modelo que se utiliza para crear una aplicación OLAP.

La idea es crear uno o algunos modelos OLAP, a partir de los cuales puede crear muchos metaperfiles, cada uno de los cuales puede estudiar un aspecto específico de la empresa.

Los metaperfiles describen el modo en que el perfil de base de datos multidimensional aparece ante el usuario de OLAP.

Creación de un Metaperfil OLAP

Puede adaptar el ámbito de un metaperfil seleccionando las dimensiones que serán visibles para los usuarios de OLAP y estableciendo filtros que determinen los datos que se recuperan.

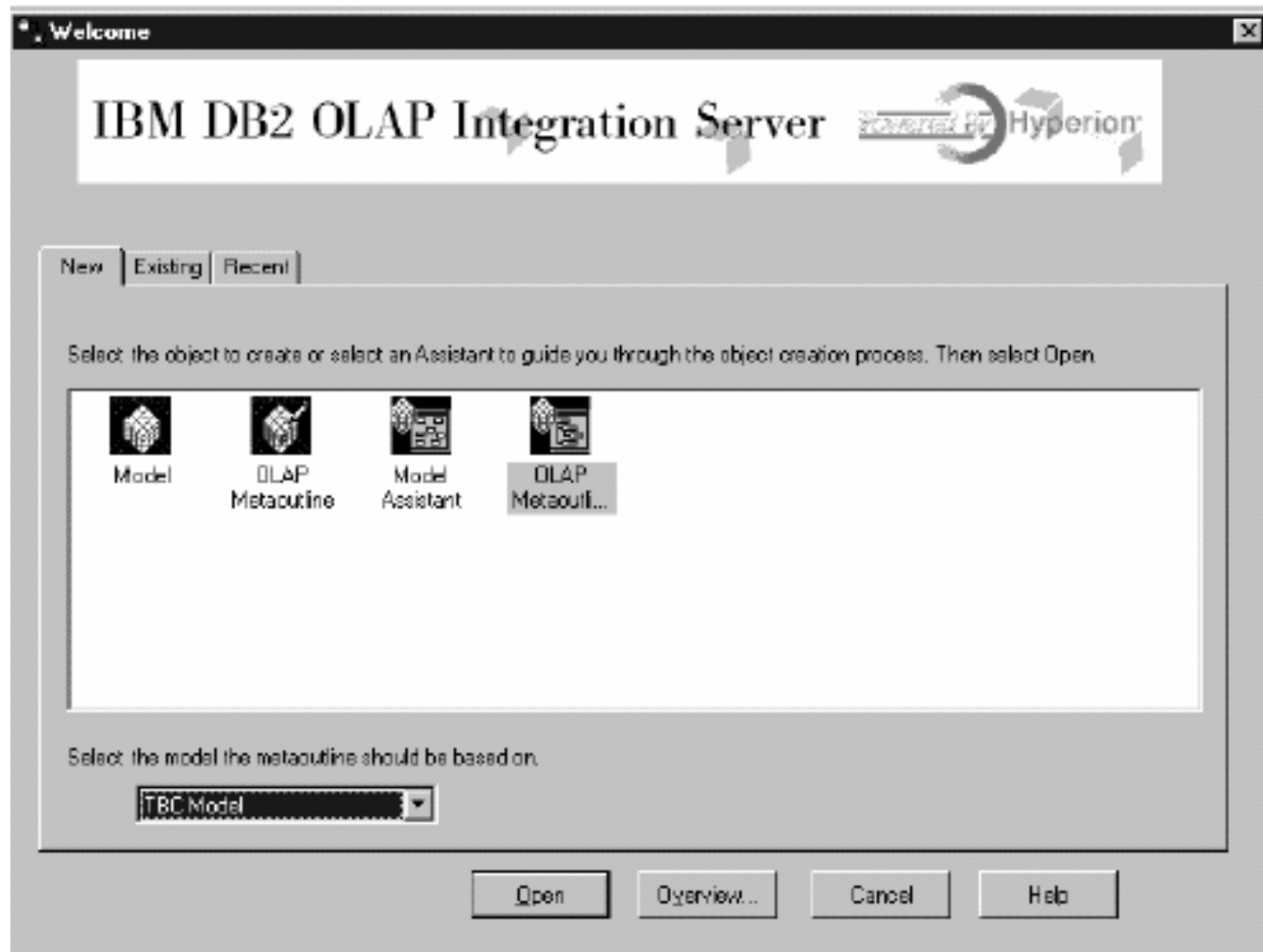
El primer paso al crear un metaperfil OLAP es decidir si utilizar la Interfaz del metaperfil OLAP, que ofrece todas las funciones, o el asistente de metaperfiles, que ofrece un planteamiento guiado más simple.

Creación de un Metaperfil OLAP

Los pasos para realizar un metaperfil con el asistente son los siguientes:

- ❑ Iniciar el asistente de metaperfiles OLAP.
 - ❑ Seleccionar un modelo OLAP en el que basar el metaperfil
 - ❑ Realizar una conexión a la base de datos.
 - ❑ Selección de dimensiones y miembros.
 - ❑ Definir propiedades.
-

Inicio del Asistente de Metaperfiles



Selección de la Dimensión y Miembros

Se deben seleccionar del modelo las dimensiones y miembros que desea utilizar en el metaperfil.

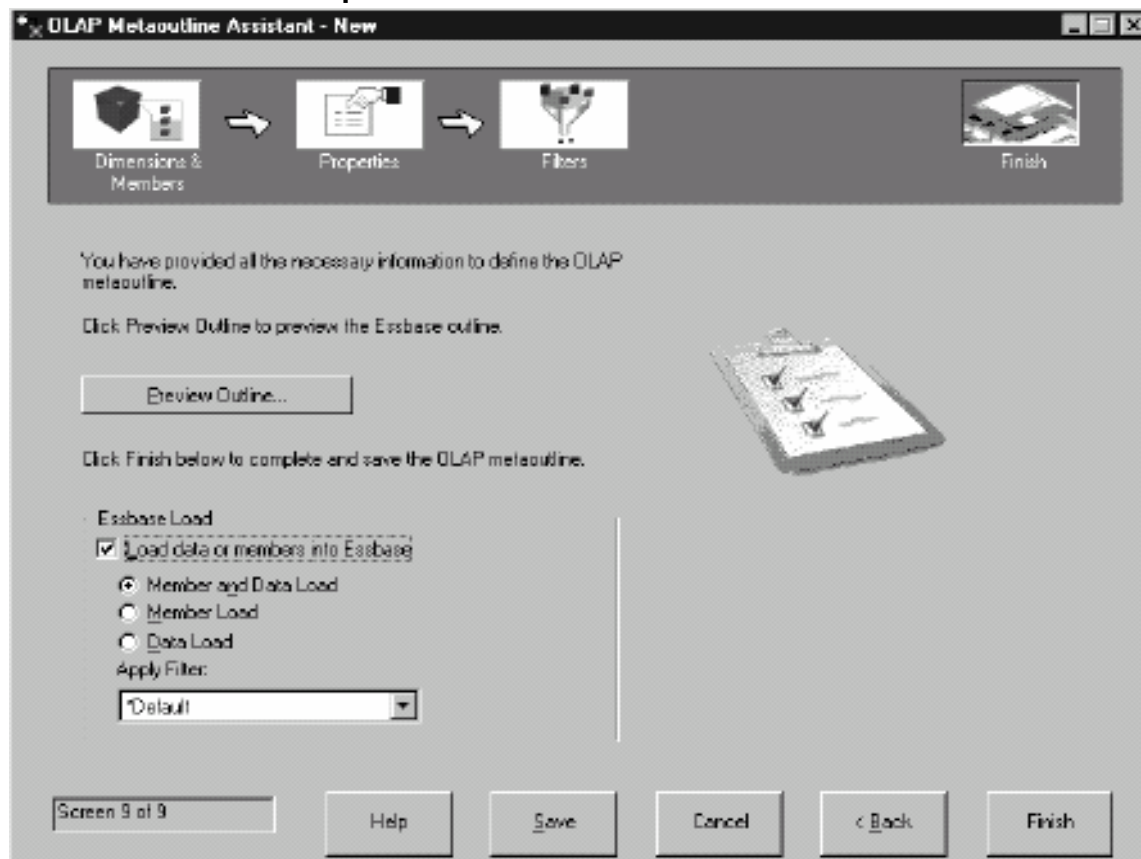
Establecimiento de Filtros

Al establecer los filtros estos limitan los miembros o datos que se cargan en una aplicación OLAP.

Por ejemplo, si el modelo OLAP contiene datos para todo el año, se puede establecer un filtro para cargar sólo los datos para el segundo trimestre.

Creación de una Aplicación OLAP

Una aplicación OLAP contiene datos estructurados por un *perfil Essbase*, o plantilla, basado en el metaperfil.



Creación de una Aplicación OLAP

Se le solicitará una serie de información.

- ❑ Un nombre.
 - ❑ El nombre de la aplicación OLAP que contendrá la base de datos en la que desea cargar los datos.
 - ❑ El nombre de la base de datos OLAP en la que desea cargar los datos.
 - ❑ Scripts de cálculo.
 - ❑ Cuándo deben cargarse los datos.
-

Creación de una Aplicación OLAP

Se iniciará el proceso Load and Calculate.

Una vez finalizado el proceso, se crea la aplicación OLAP y se puede analizar mediante la utilización de los programas de hoja de cálculo Microsoft Excel o Lotus 1-2-3.

Conclusión

Beneficios al adoptar el análisis OLAP

- ❑ Le permite calcular algunos indicadores financieros más importantes como ganancias por acción, retorno en valor líquido, flujo de inventario y días de ventas sobresalientes entre otros.
 - ❑ El análisis de flujo de caja le da la habilidad de ver qué segmentos del negocio están generando dinero y cuales no. Responde a la pregunta ¿Está el dinero en lugar indicado y en los niveles adecuados para un mejor retorno?.
-

Conclusión

- ❑ Análisis de venta de productos para determinar cuáles productos y segmentos están aportándole más a la compañía.
 - ❑ El análisis de rentabilidad de productos y clientes identifica los clientes y productos más rentables de una compañía; conocer sus mejores clientes y productos es crítico para la supervivencia de una compañía y para la identificación de nuevas oportunidades.
 - ❑ Análisis de mercadeo estratégico para organizar clientes desde los más valiosos a los menos valiosos (basado en el ingreso o margen, o en ambos) en un periodo de tiempo determinado.
-

Bibliografía

- ❑ Mc Graw Hill Alex Benson, Stephen J. Smith. Data Warehouse, Data mining and OLAP. USA 1997.
 - ❑ María José Remírez Quintana, José Hernandez Orallo. Extracción Automática de Conocimientos en Base de Datos e Ingeniería del Software. España 2003.
 - ❑ Jhon Wiley W.H. Inmon and Sons. Building the Data Warehouse. USA 1996.
 - ❑ IBM Press 2001. IBM DAB2 UDB Business Intelligence Tutorial.
-

FIN DE LA PRESENTACIÓN

Gracias por su Atención
