

IBM DB2 OLAP SERVER



PROCESAMIENTO ANALÍTICO EN LINEA

IBARRA MARIA DE LOS ANGELES

CORRINTES 2005

INDICE

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

INDICE

- ❖ 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

Base de Datos: conjunto de datos estructurados, fiables y homogéneos organizados independientemente en máquina, accesibles en tiempo real, compatible por usuarios concurrentes que tienen necesidades de información diferentes y no predecibles en el tiempo.

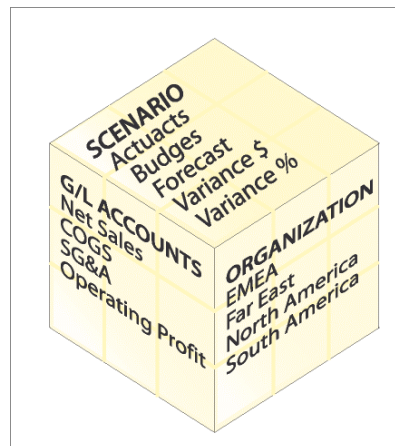
Tipos de Base de Datos:

- ❖ Jerárquica.
- ❖ Red .
- ❖ Relacional.
- ❖ Multidimensionales.

INTRODUCCIÓN

Estructura Multidimensional: esta estructura es una variación de la estructura relacional que utiliza estructuras multidimensionales para organizar datos y expresar las relaciones entre ellos.

Visualización de las Estructuras Multidimensionales



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.

Creación del DW: este se crea al extraer datos de una o más BD de aplicaciones operacionales.

Los datos extraídos se transforman para eliminar inconsistencias y resumir si es necesario, y luego son cargadas en el DW.

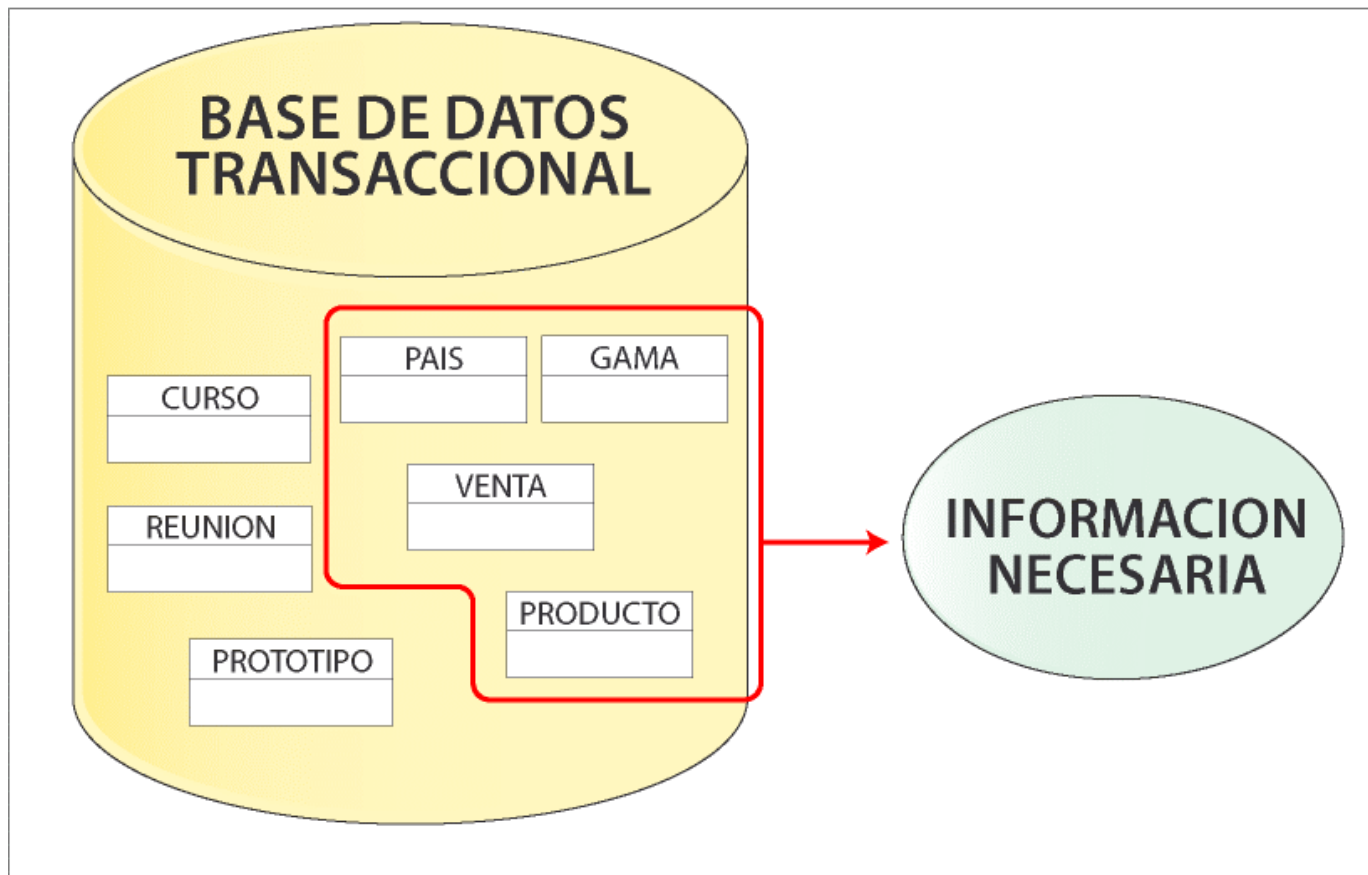
DATA WAREHOUSE

Características del DW

❖ *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.

Y no para exportar los procesos que se realizan en la organización como ser gestión de pedidos, facturación, etc.

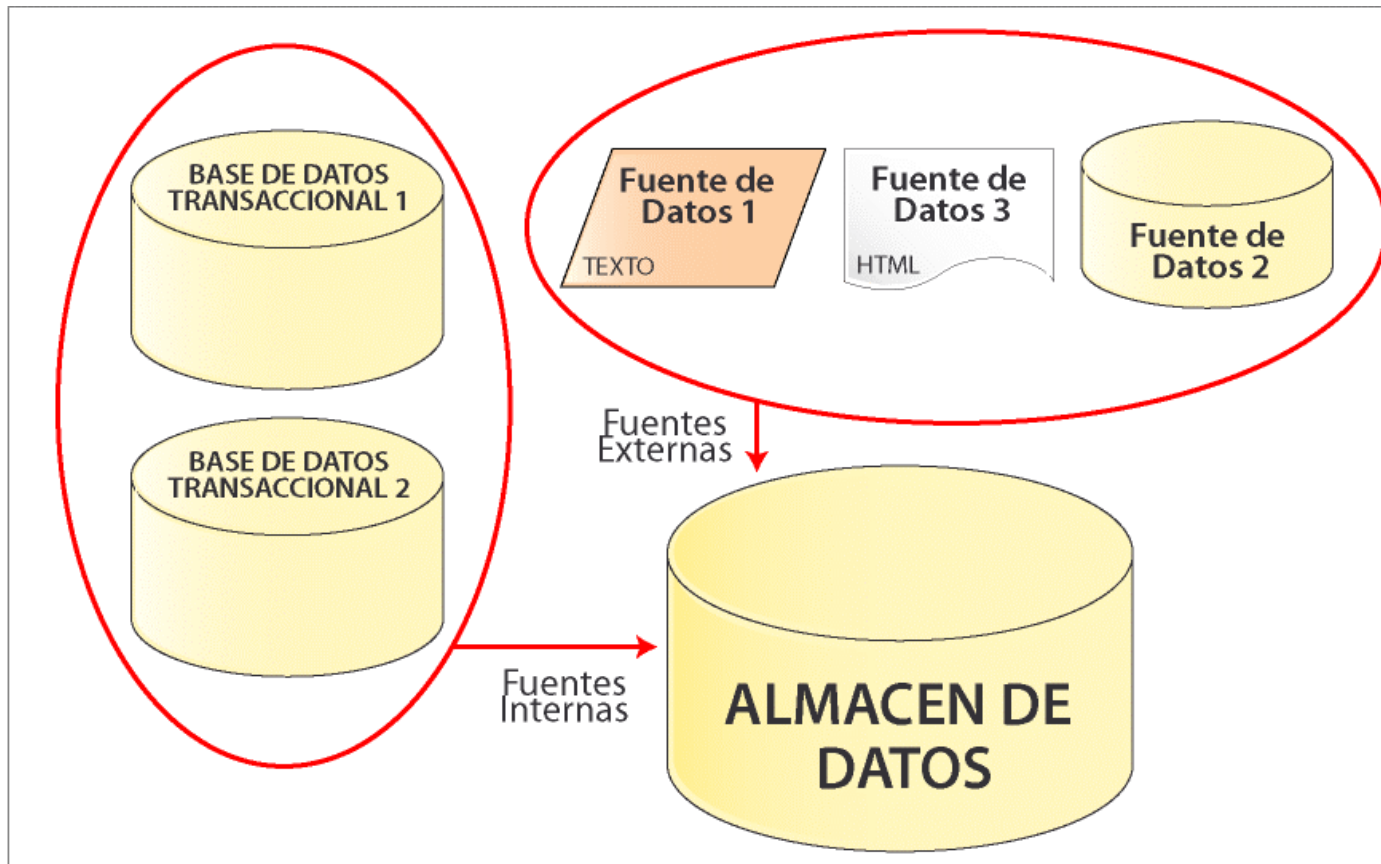
DATA WAREHOUSE



DATA WAREHOUSE

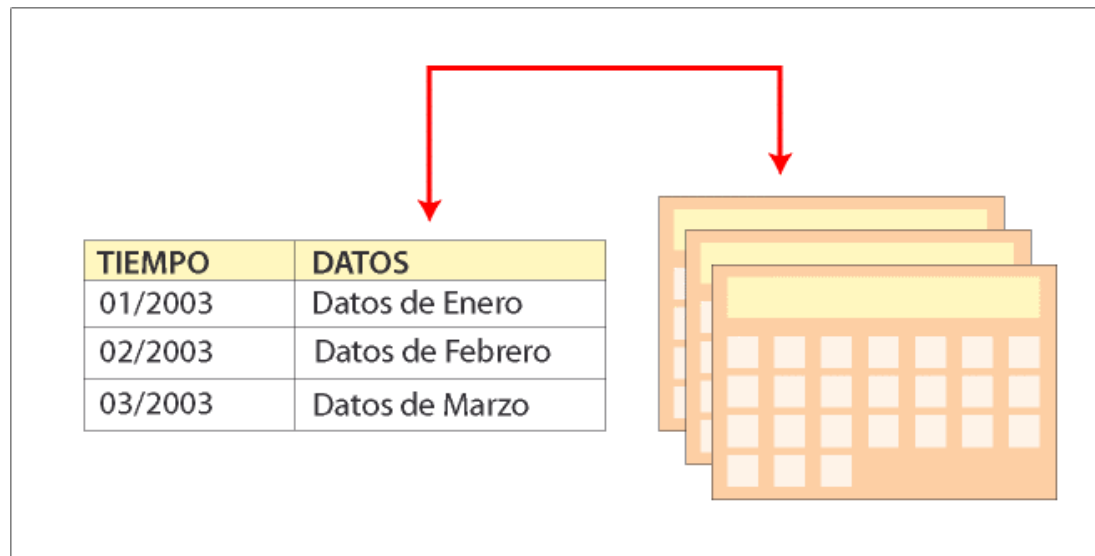
- ❖ *Datos Integrados*: integra datos recolectados de diferentes sistemas operacionales de la organización y o fuentes externas.

DATA WAREHOUSE



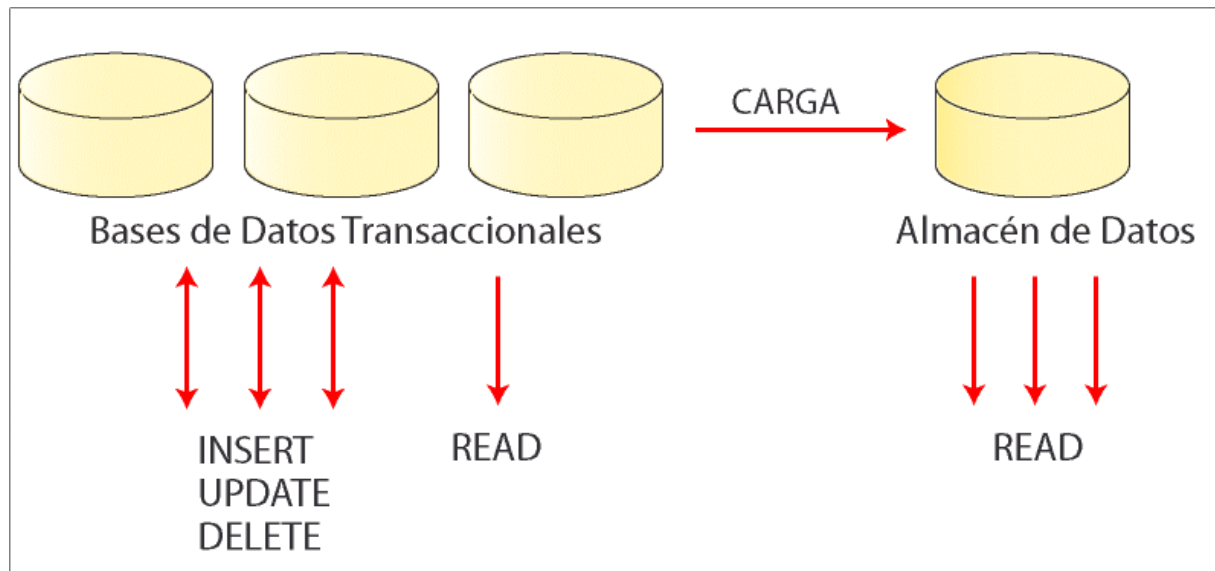
DATA WAREHOUSE

❖ *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 MART

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.

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.

Estos sistemas definen el comportamiento operacional en un entorno operacional de gestión.

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 modelos 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
	Toys	\$1,90
Q2	Electronics	\$8,90
	Toys	\$0,75

VISTA DE DATOS

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.

		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

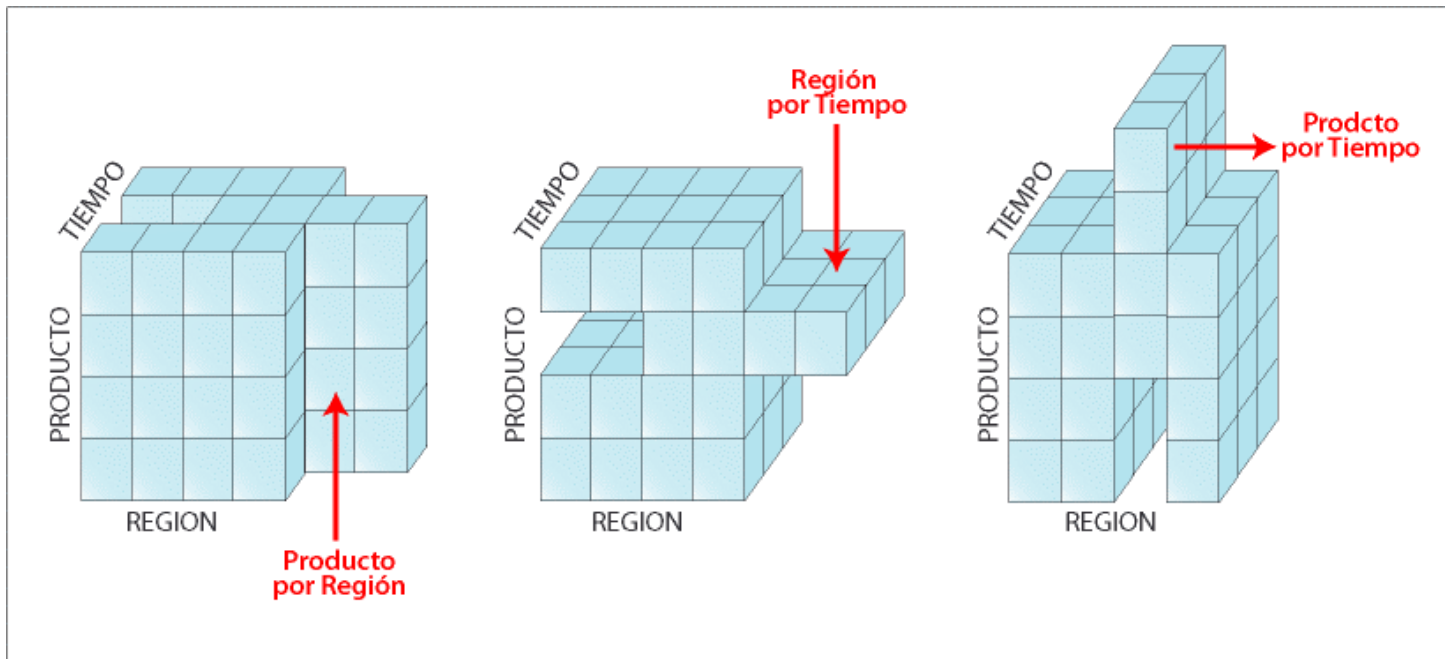
VISTA DE DATOS

Esto mismo puede ser visualizado mediante un cubo.

	MAYO	ABRIL	JUNIO
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
	Argentina	Brasil	Chile

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.
- ❖ La data puede ser compartida con aplicaciones SQL.
- ❖ Data 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 DE OLAP

- ❖ 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 DE OLAP

- ❖ Es de fácil uso y acceso flexible para los usuarios.
- ❖ Los datos están organizados en diferentes dimensiones los 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

Como primer medida se deben realizar los siguientes pasos:

1. 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

2. ANALISIS MODELO MULTIDIMENSIONAL

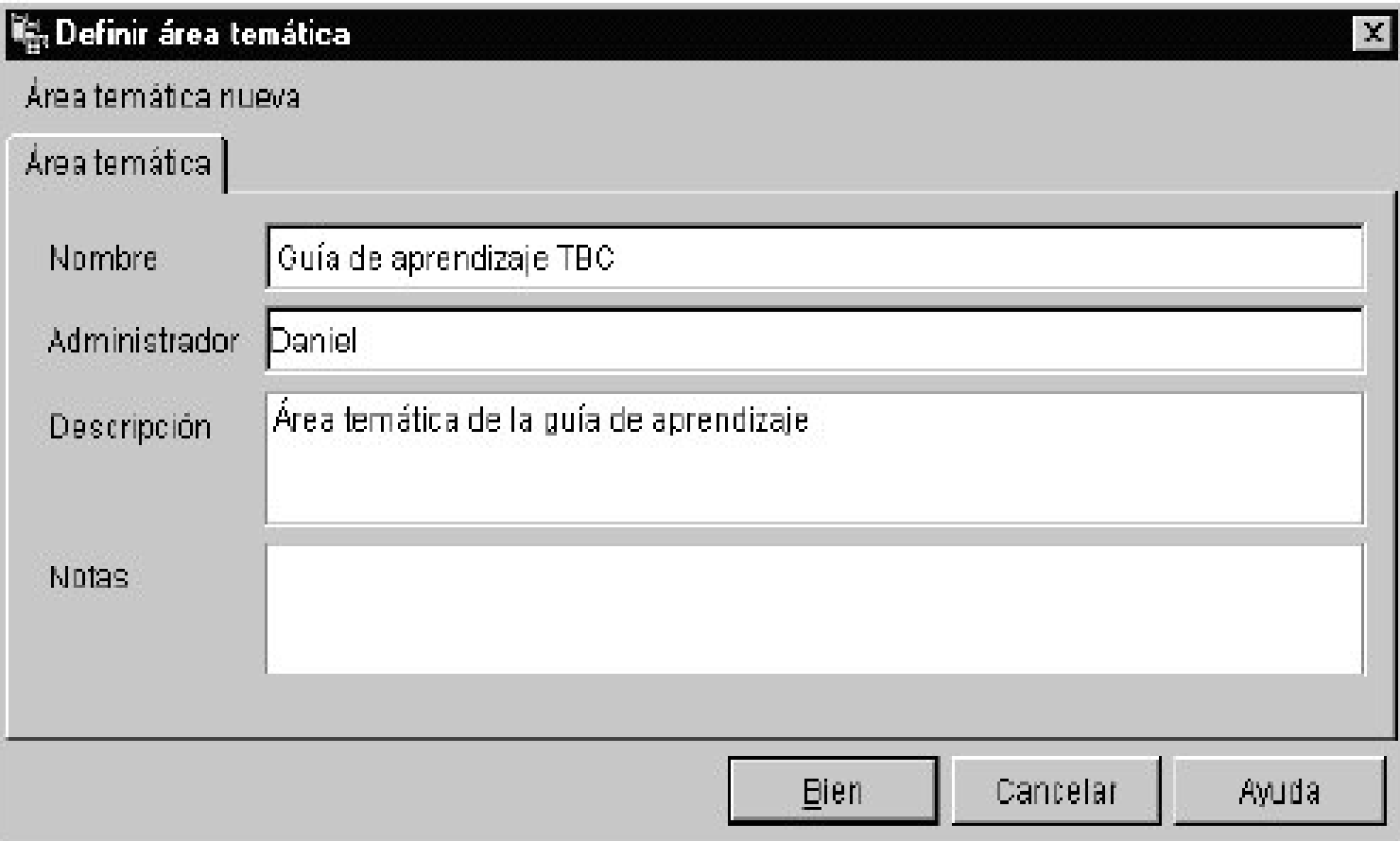
- ❖ Creación del modelo OLAP
 - ❖ Selección de la tabla de hechos.
 - ❖ Creación de dimensiones.
 - ❖ Unión de las tablas de mediciones.
 - ❖ Creación de Herencia.
 - ❖ Guardar modelo OLAP.
- ❖ Creación un metaperfil
 - ❖ Iniciar asistente.
 - ❖ Seleccionar un modelo olap en el que basar el metaperfil.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

DEFINICIÓN DE AREA TEMATICA

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



The image shows a software dialog box titled "Definir área temática" (Define thematic area). The dialog box has a standard Windows-style title bar with a close button (X) in the top right corner. Below the title bar, the text "Área temática nueva" (New thematic area) is displayed. The main content area is divided into several sections:

- Área temática:** A label followed by a vertical line, indicating the current section.
- Nombre:** A text input field containing the text "Guía de aprendizaje TBC".
- Administrador:** A text input field containing the text "Daniel".
- Descripción:** A text input field containing the text "Área temática de la guía de aprendizaje".
- Notas:** A large, empty text area for additional notes.

At the bottom of the dialog box, there are three buttons: "Elen" (likely "Elegir" or "Aceptar"), "Cancelar", and "Ayuda".

DEFINICIÓN DE FUENTE DE DEPÓSITOS

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.

Puede utilizar más de una fuente en un proceso.

DEFINICIÓN DE FUENTE DE DEPÓSITOS

The image shows a Windows-style dialog box titled "Definir fuente de depósito" (Define data source). The dialog has a tabbed interface with the following tabs: "Fuente de depósito", "Sitios de agente", "Base de datos", "Tablas y vistas", and "Seguridad". The "Fuente de depósito" tab is selected. The dialog contains the following fields:

- Nombre:** Fuente relacional de la guía de aprendizaje
- Tipo de fuente de depósito:** DB2 UDB para Windows NT
- Administrador:** Daniel
- Descripción:** Datos relacionales para la empresa TBC
- Notas:** (Empty text area)

At the bottom of the dialog are three buttons: "Ejecutar", "Cancelar", and "Ayuda".

DEFINICIÓN DE DESTINO DE DEPÓSITOS

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 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

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 PROCESOS

Nombre	Modifi...	Descr...
crear dimension producto	9 de abri...	
Crear dimension de mercado de la gua ...	9 de abri...	Proceso ...
Crear dimensión de entorno	9 de abri...	
Crear tabla de hechos de la guia de apr...	9 de abri...	
Creardimensión de tiempo	9 de abri...	

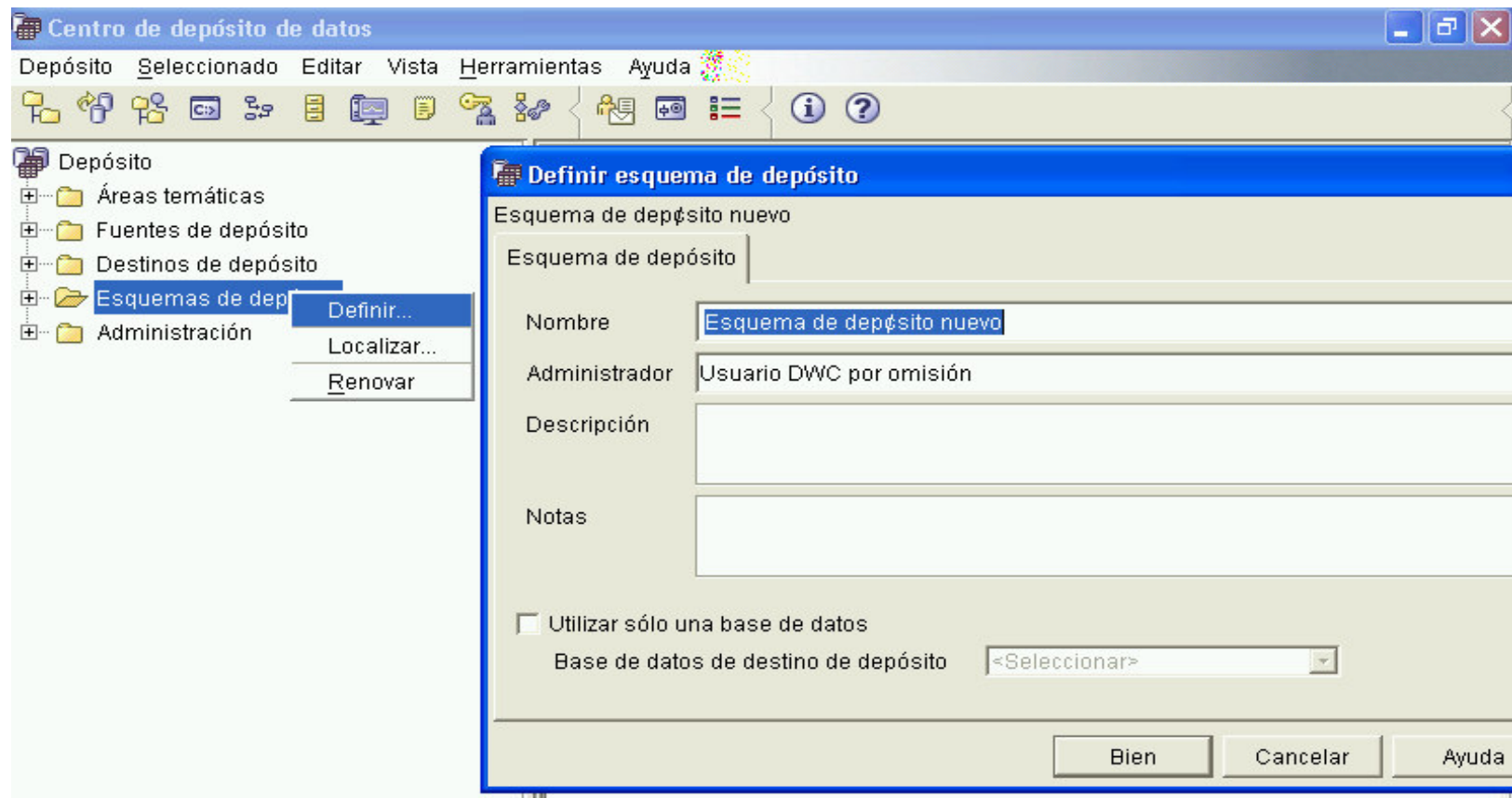
DEFINICIÓN ESQUEMA ESTRELLA

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

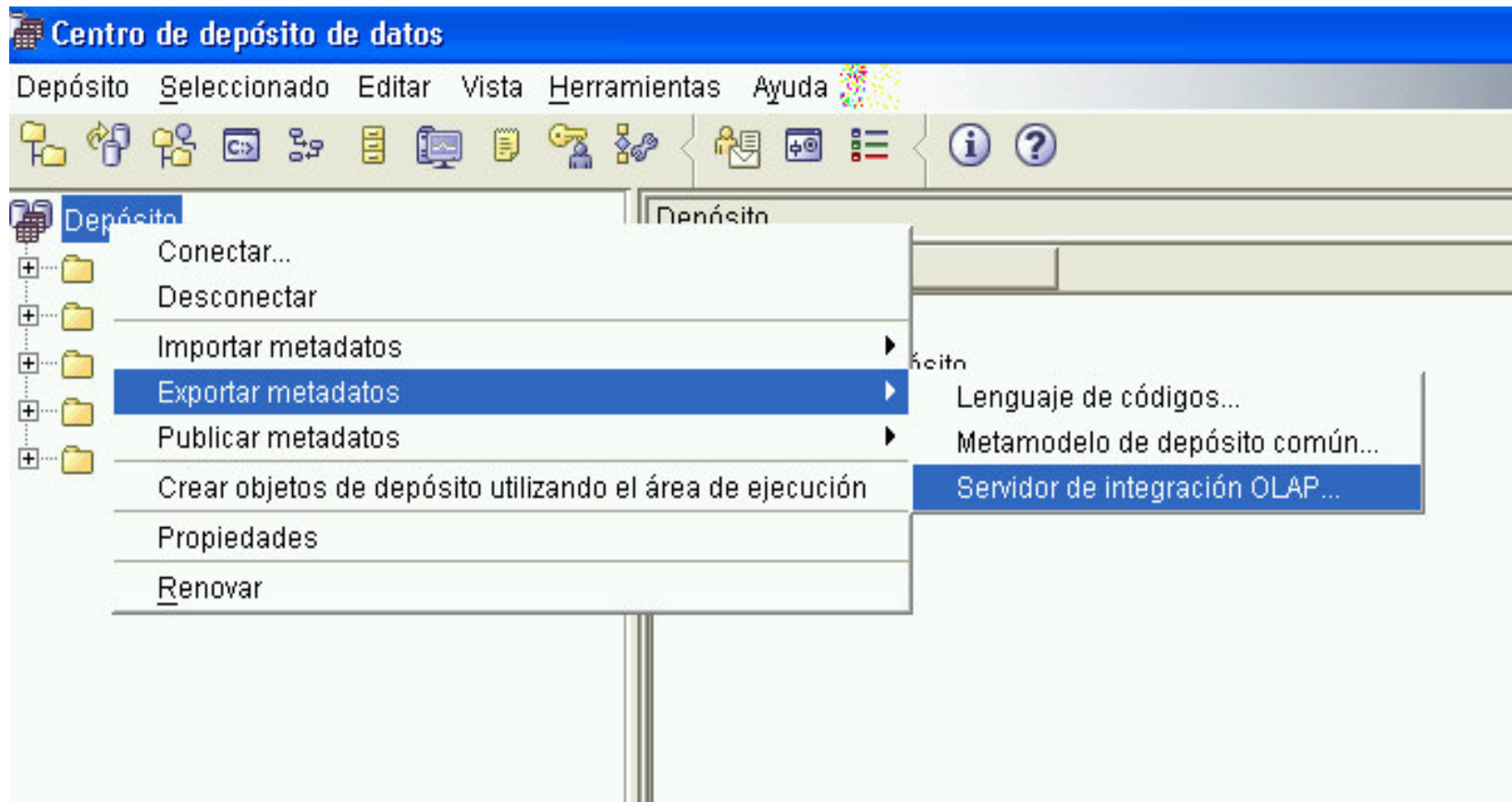
Se puede utilizar este esquema para consultas en la base de datos de depósito.

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

DEFINICIÓN ESQUEMA ESTRELLA



EXPORTACIÓN DEL EXQUEMA



ANALISIS MODELO 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.

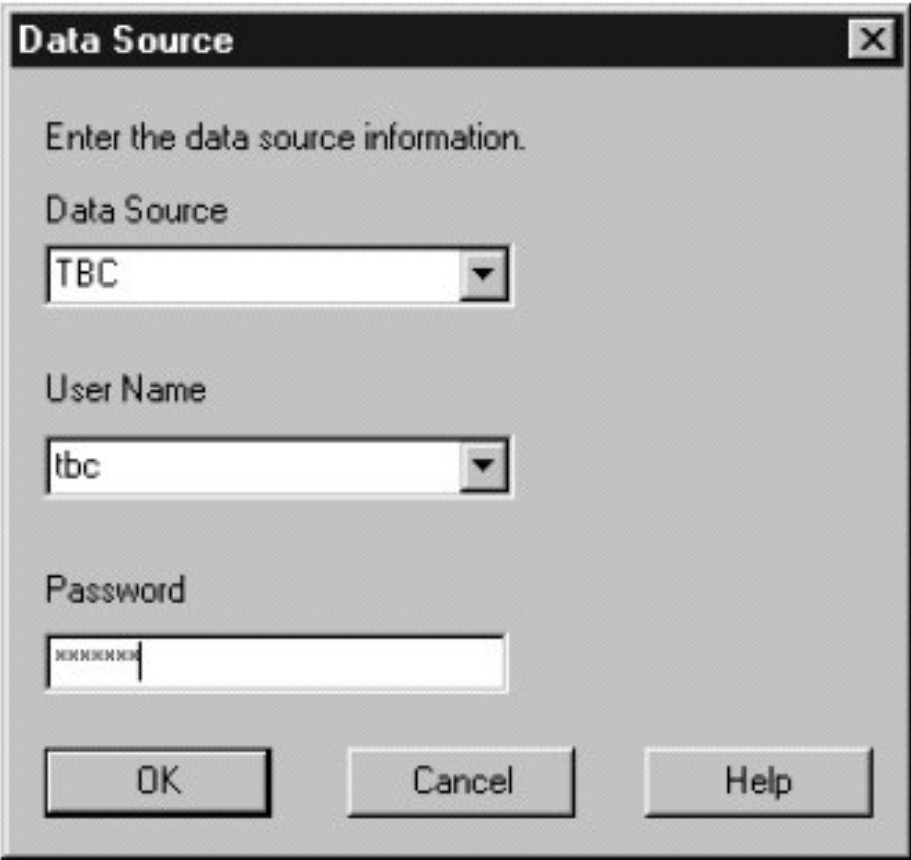
CREACIÓN DEL MODELO OLAP

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 bar reads 'IBM DB2 OLAP Integration Server' followed by 'Powered by Hyperion Essbase' and a logo. 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 note: 'When working with a metaoutline, specify Essbase login information.' It contains three input fields: 'Server', 'User Name', and 'Password'. At the bottom of the dialog are 'OK', 'Cancel', and 'Help' buttons.

CREACIÓN DEL 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.



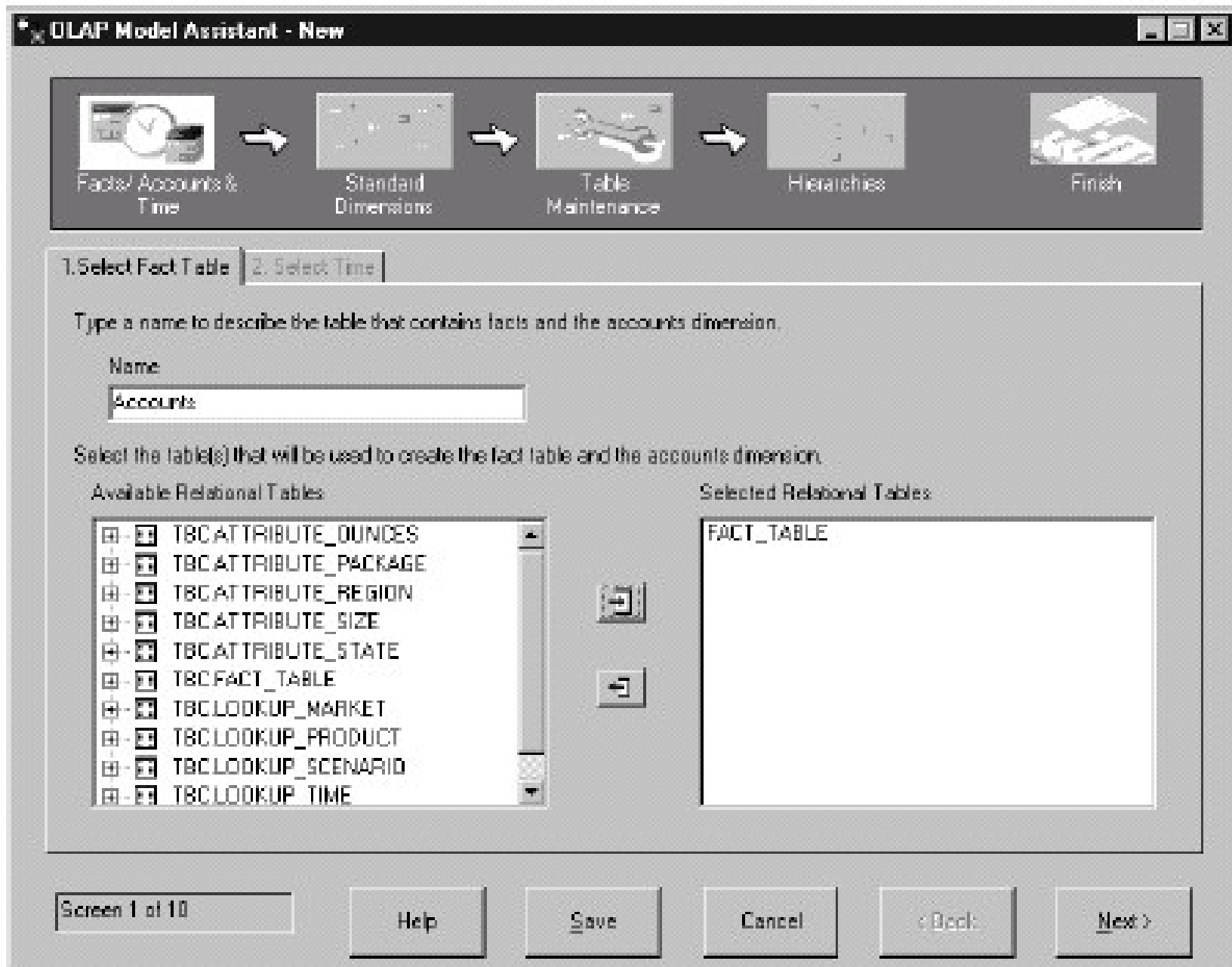
The image shows a 'Data Source' dialog box with the following fields and controls:

- Data Source:** A dropdown menu with 'TBC' selected.
- User Name:** A dropdown menu with 'tbc' selected.
- Password:** A text input field with masked characters (XXXXXXXXXX).
- Buttons:** 'OK', 'Cancel', and 'Help' buttons at the bottom.

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.



CREACIÓN DE DIMENSIONES

OLAP Model Assistant - New

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

1. Select Fact Table | 2. Select Time

Type a name to describe the time dimension.

Name
Time

Select how to create the time dimension.

Based on Fact Table

By Selecting Relational Tables

Available Relational Tables

- TBC.ATTRIBUTE_SIZE
- TBC.ATTRIBUTE_STATE
- TBC.FACT_TABLE
- TBC.LOOKUP_MARKET
- TBC.LOOKUP_PRODUCT
- TBC.LOOKUP_SCENARIO
- TBC.LOOKUP_TIME

Selected Relational Tables

LOOKUP_TIME

Screen 2 of 10

Help Save Cancel < Back Next >

UNION 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.

OLAP Model Assistant - New

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

TIME_ID
SEASON
QUARTER_ID
QUARTER
MONTH_ID
MONTH

Facts

TRANSDATE
TIME_ID
SCENARIO_ID
SALES
PRODUCT_KEY
PAYROLL
OPENING_INVENTORY
MISC
MARKETING
ENDING_INVENTORY
COGS
CITY_ID
ADDITIONS

Screen 5 of 10

Help Save Cancel < Back Next >

CREACIÓN DE HERENCIA

Las jerarquías organizan relaciones padre-hijo entre las columnas de una dimensión y se visualizan como estructura en árbol.


Por ejemplo, en la dimensión Time, puede definir el miembro Year como parte superior de la jerarquía. El miembro Quarter sería un hijo de Year y Month sería un hijo de Quarter.

El aspecto d la ventana es el siguiente



1. Define Hierarchies | 2. Preview Hierarchies

If necessary, add hierarchies to the dimensions and define the structure to use for consolidating data.



Add Hierarchy...  **Define Relationships**

Dimension Columns

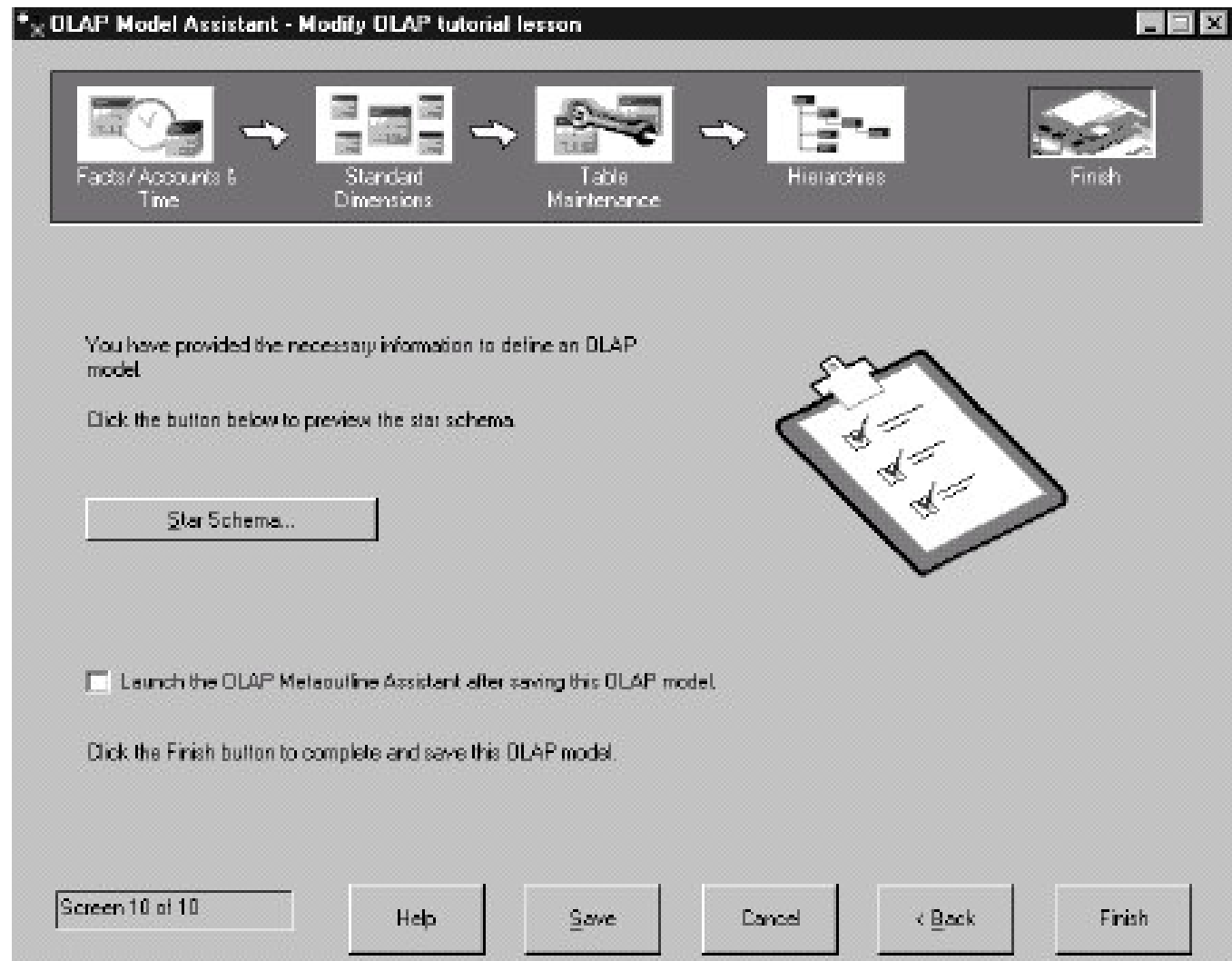
- CITY_ID
- POPULATION
- REGION_TYPE_ID
- SIZE_ID
- STATE
- STATE_TYPE_ID

Parent /Child Relationship

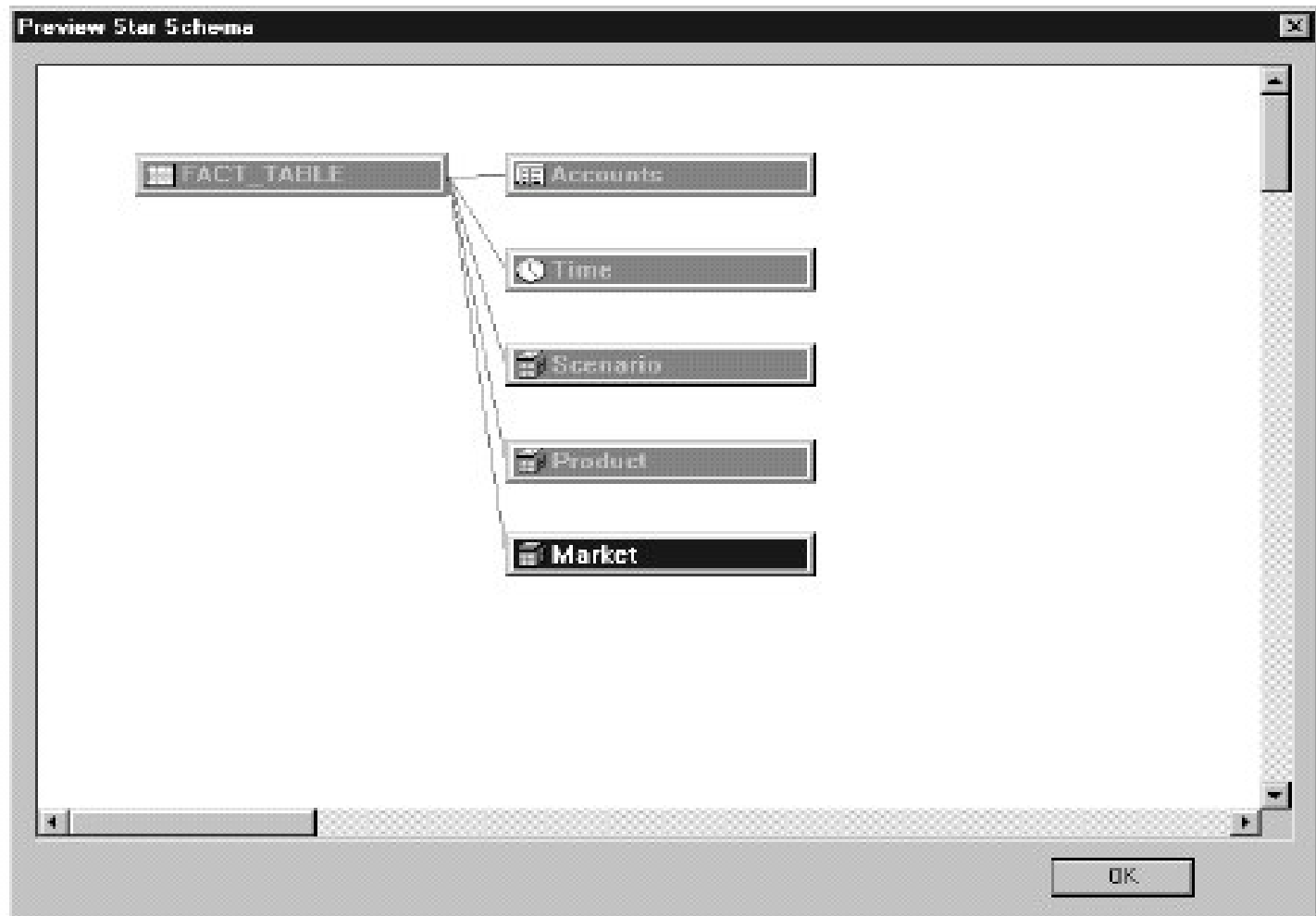
- Region-City
 - LOOKUP_MARKET.REGION
 - LOOKUP_MARKET.CITY

GUARDAR MODELO OLAP



GUARDAR 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.

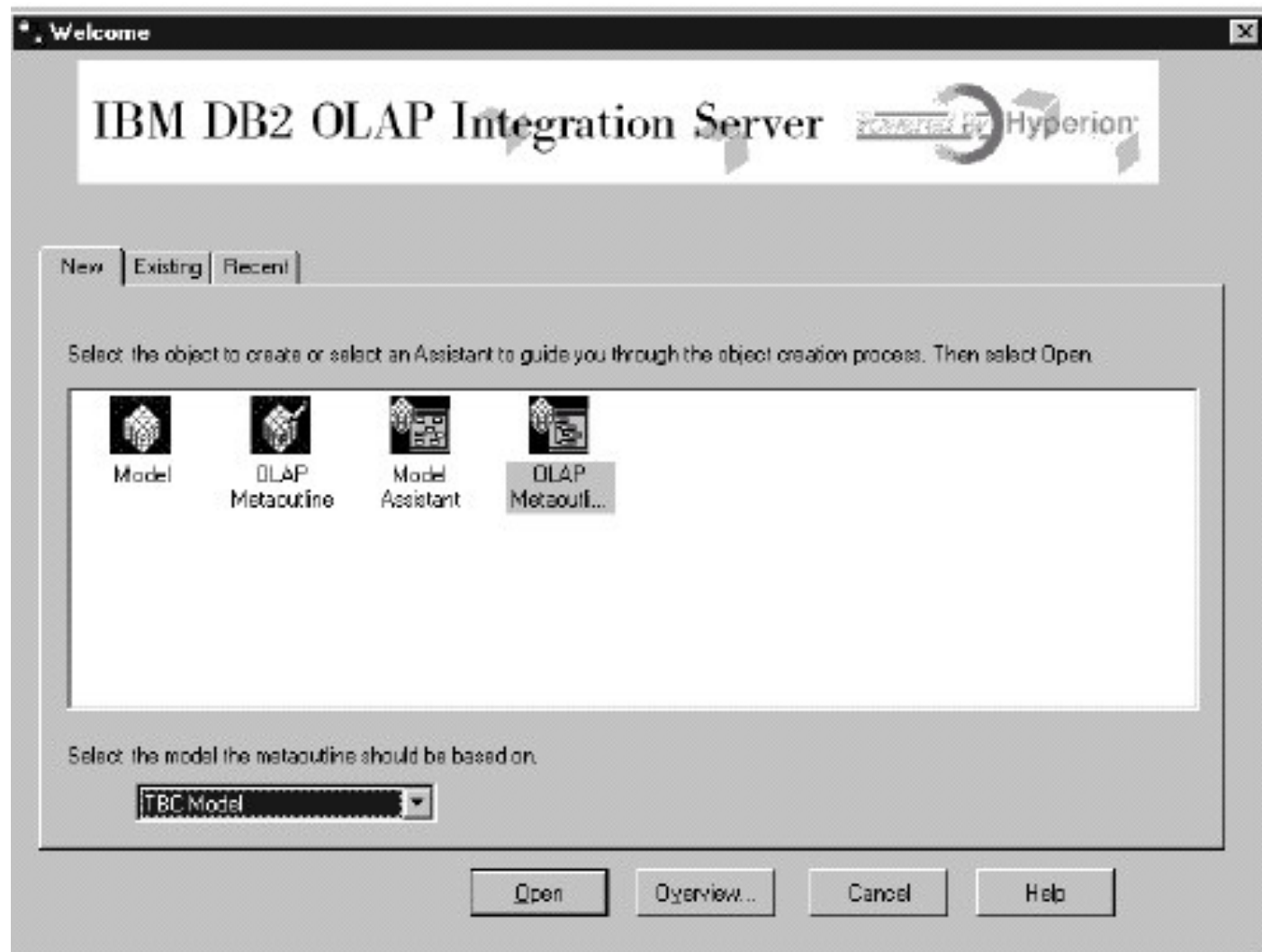
El primer paso al crear un metaperfil OLAP es decidir si utilizar la Interfaz del metaperfil OLAP, que ofrece todas las funciones, o el asistente de metaperfiles, que ofrece un planteamiento guiado más simple.

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



CONEXIÓN CON LA BASE DE DATOS FUENTE

En la ventana Data Source, se deben escribir los siguientes valores:

Data Source

Nombre de la base de datos de DB2 donde se encuentran almacenados los datos comerciales.

User Name

ID de usuario que utilizará para acceder a DB2 UDB.

Password

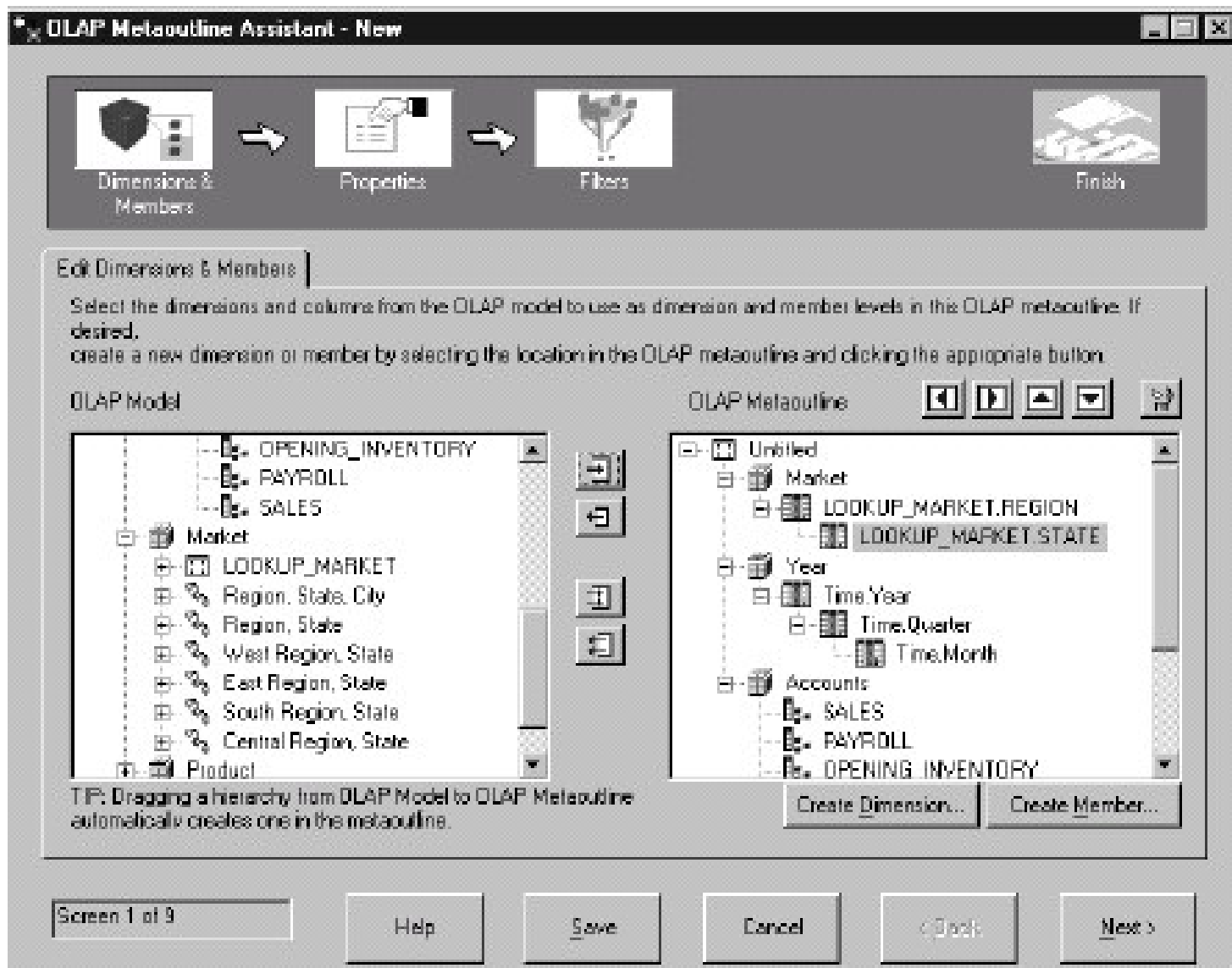
Contraseña para el ID de usuario que ha entrado en *User Name*.

SELECCIÓN DE LA DIMENSION Y MIEMBRO

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

El aspecto de la ventana es el siguiente:

SELECCIÓN DE LA DIMENSION Y MIEMBRO



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.

ESTABLECIMIENTO DE FILTROS

OLAP Metaoutline Assistant - New

Dimensional & Members → Properties → Filters → Finish

1. Name Filters | 2. Assign Member Filters | 3. Assign Measures Filters | 4. Review Filters

Type a name and an optional description for each OLAP metaoutline filter to be created in the OLAP metaoutline.

— OLAP Metaoutline Filter Detail —

Name

Description

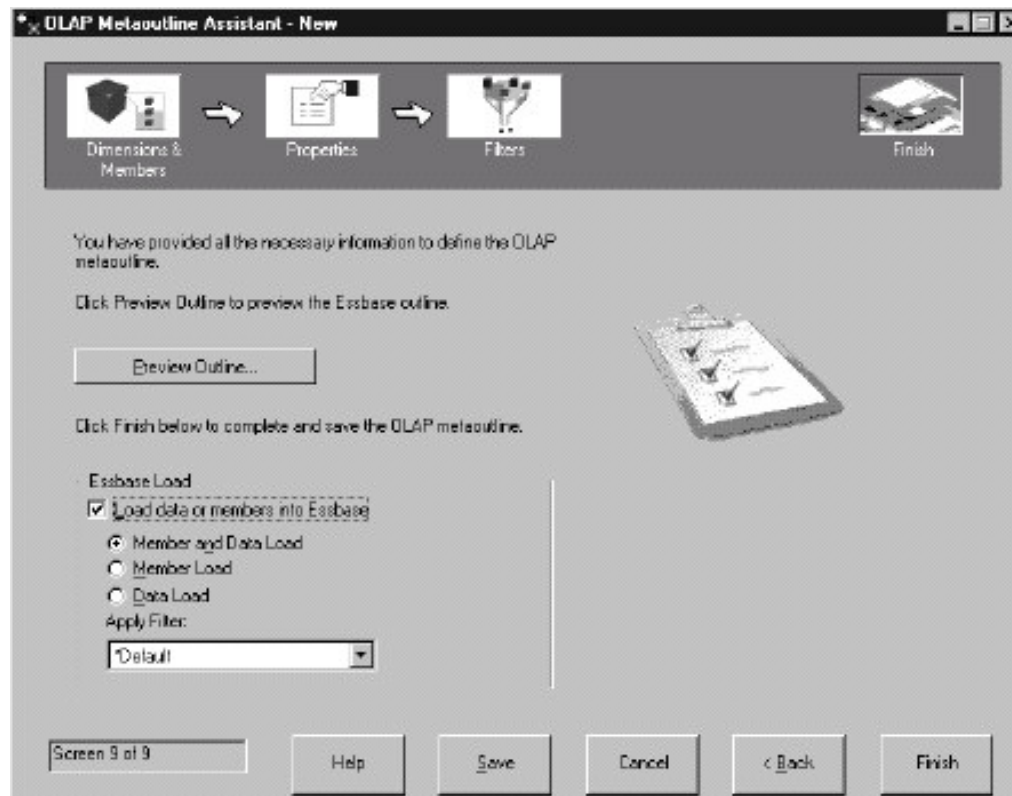
OLAP Metaoutline Filter List

Name	Description
Default	
SalesSG	

Screen 5 of 9

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.

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