Modelos y Persistencia de Datos

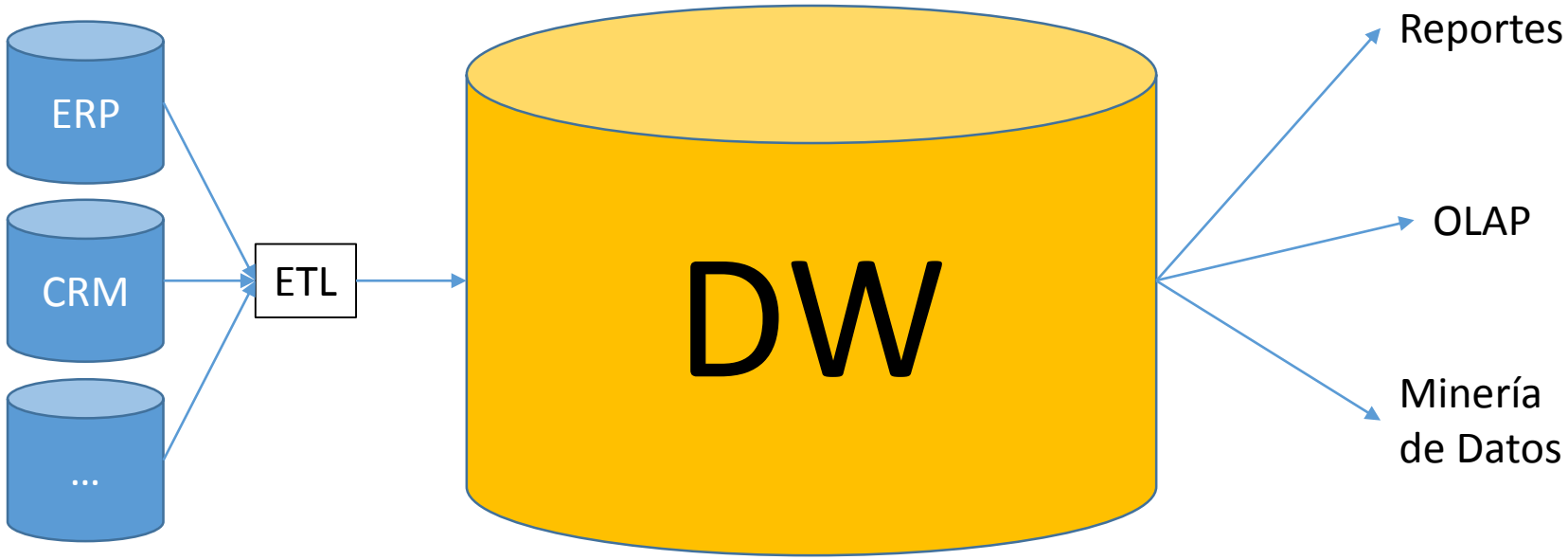
Inteligencia de Negocios
(Business Intelligence)

Alejandro Sierra

Sistemas de Apoyo a Decisiones (DSS)

- Características (Sprague 1980)
  1. Apunta a problemas de toma de decisiones poco especificados.
  2. Combinar técnicas y modelos analíticos con funciones de acceso y recuperación de datos.
  3. Fácil de usar e interactivo para personas no técnicas.
  4. Flexibilidad y capacidad de adaptarse a cambios en el ambiente de toma de decisiones.

Mapa general de BI



Fuentes:
- (OLTP)
- Maestras

Bodega de Datos

Consultas
a la
Bodega

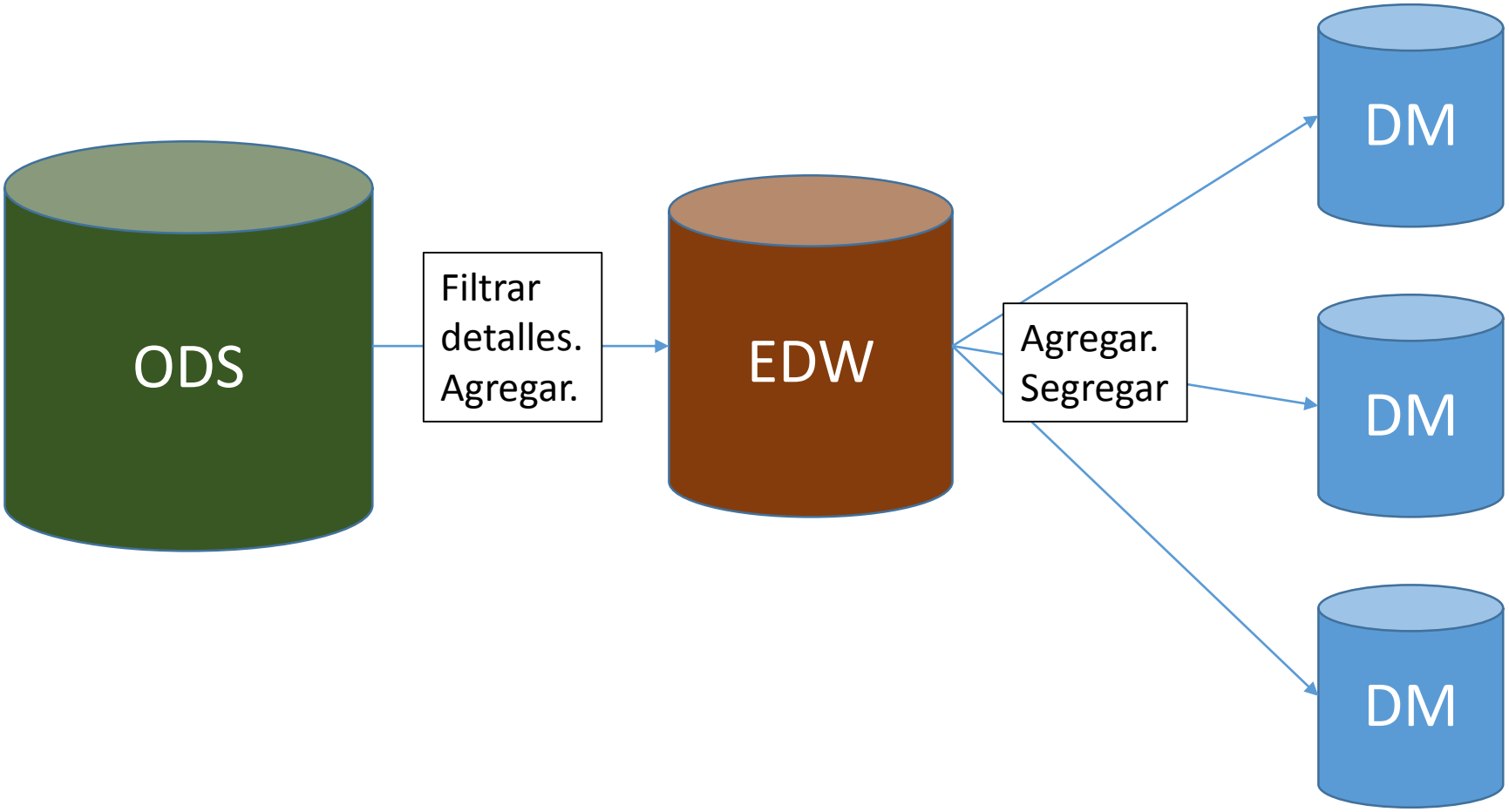
Bodega de Datos (DW)

- Repositorio Central
- Datos provienen de fuentes internas y externas
- Limpios y transformados
- Formateados para consulta rápida.
- Tipos de bodegas (no excluyentes)
  - Almacén de datos Operacionales (ODS)
  - Bodega de Datos Empresarial (EDW)
  - Data Mart (DM)

Tipos de DW

Tipo	Detalle	Alcance	Historial	Uso
ODS	Máximo nivel de detalle	Empresa	Corto	Táctico, análisis de la operación en el día a día
EDW	Detalle y agregaciones	Empresa	Amplio	Táctico y Estratégico
DM	Agregaciones. Poco detalle	Tema específico	Amplio	Datos con alto valor para un grupo de personas, unidad de negocio, …

Flujo de datos dentro de DW



Diseño Multidimensional

- Hecho
  - Es **lo que se analiza**
  - Normalmente solo contiene medidas (números)
    - Y relaciones con las dimensiones
  - Gran cantidad de filas
  - *Ejemplo: Cantidad de productos vendidos*
- Dimensión
  - Contiene **criterios de análisis**
  - Jerarquías
  - Sólo se conecta con otras dimensiones a través de los hechos
  - *Ejemplos: Cuándo, Tienda, Cliente, Producto*

Diseño Multidimensional

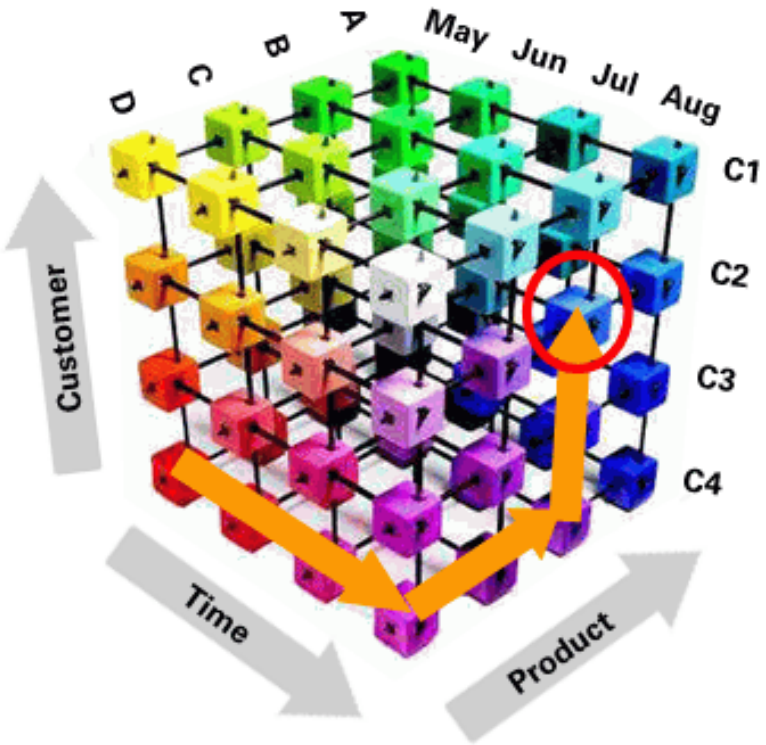
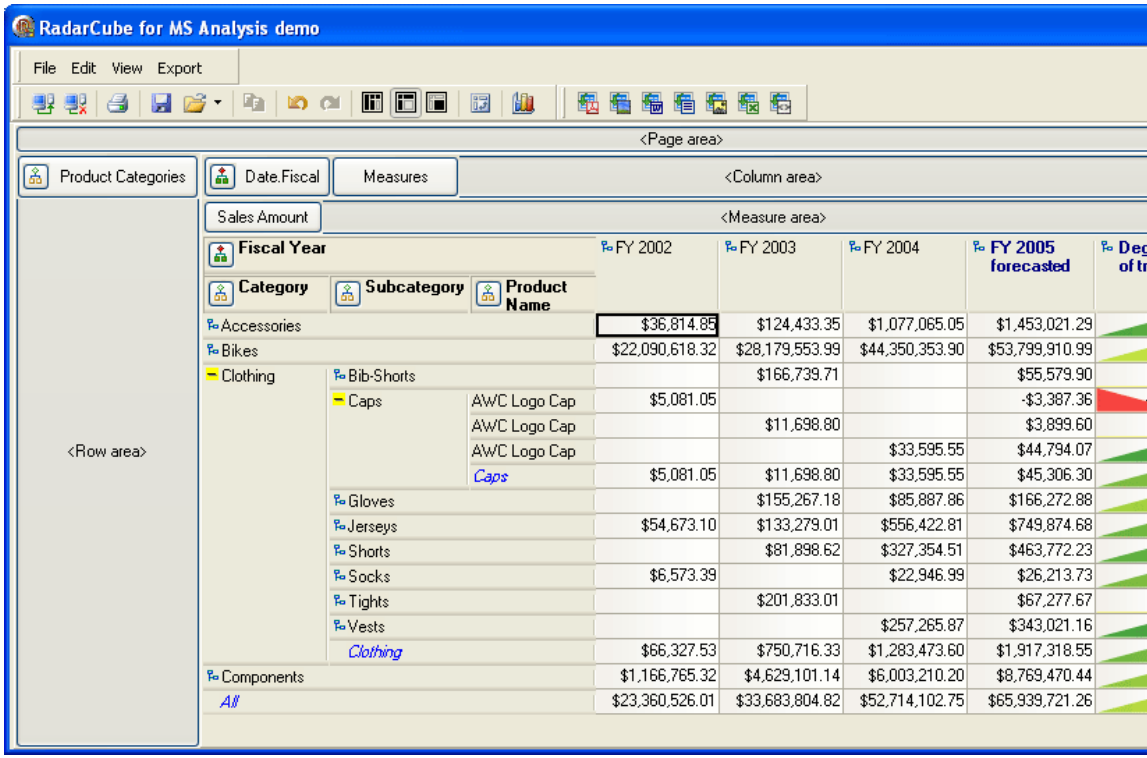


Imagen tomada de http://beagleresearch.com/which-tools-do-you-use/

Jerarquías en las dimensiones

- Permiten hacer análisis con diferentes niveles de detalle
- Ej.
  - Tiempo
    - Año,
    - Trimestre,
    - Mes,
    - Semana,
    - Día
  - Producto
    - Línea,
    - Modelo,
    - Producto

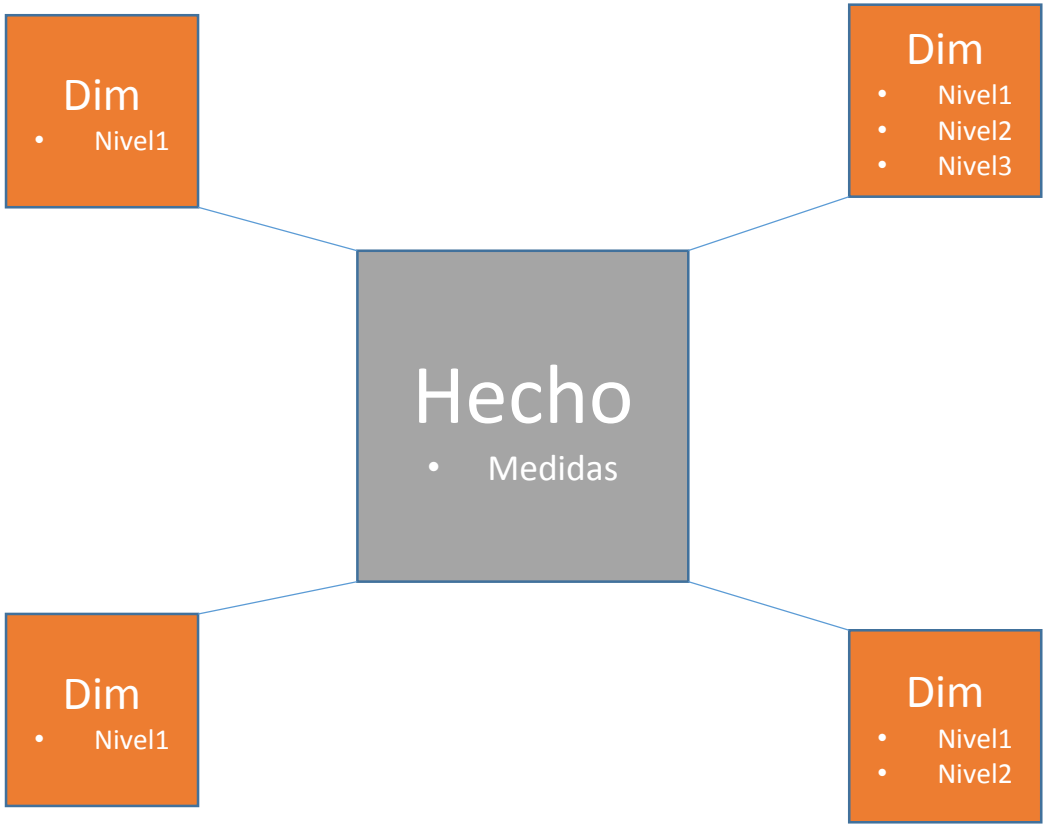


Diseño Físico

- Diseño físico de un modelo multidimensional
  - Esquema en Estrella
    - Dimensiones completamente denormalizadas
  - Esquema en Copo de Nieve
    - Dimensiones normalizadas
  - Hibrido
- Las relaciones (PK y FK) se establecen mediante llaves subrogadas
  - Llaves sin significado de negocio.
  - Autoincrement, Hash, secuencias,….

Esquema en Estrella

- Denormalizado



Jerarquías en Esquema Estrella (Dimensiones)

- Ejemplo de Jerarquía
  - Oficina, Ciudad, Departamento, Región
- Todos los niveles en una única tabla.
  - Alta redundancia
  - Muy pocos Joins para las consultas

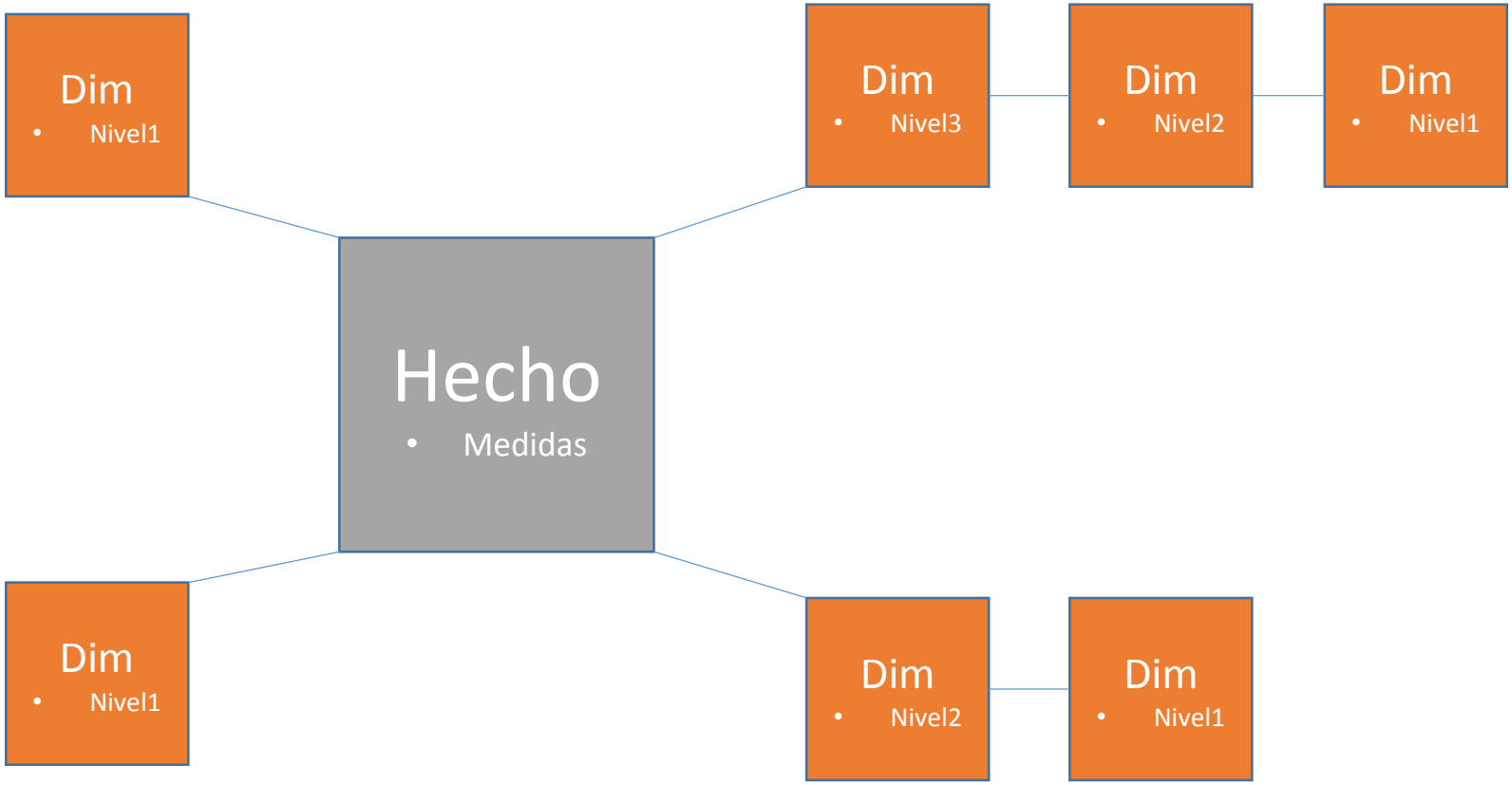
Ejemplo Dimensión de Tiempo denormalizada

Fecha ID	Fecha	Dia Semana	Número Mes	Mes	Año
12654	09/09/15	Martes	9	Sept.	2015
12655	10/09/15	Miércoles	9	Sept.	2015
…					
12684	06/10/15	Martes	10	Oct	2015

Copo de Nieve

- Se mantienen normalizadas las dimensiones
  - Cada nivel de la jerarquía está en una tabla diferente
- La tabla de hechos se relaciona <u>con el nivel más detallado de la jerarquía</u>
- Más fácil de actualizar
- Menos redundancia
- <u>Más Joins</u> para una consulta agregada

Copo de Nieve



Copo de Nieve
Ejemplo dimensión de tiempo

Fecha	Dia Semana
08/09/2015	Martes
09/09/2015	Miercoles
…	

Mes	Nombre
09/2015	Sept.
10/2015	Oct.
…	

Año
2015
2016

Hibrido

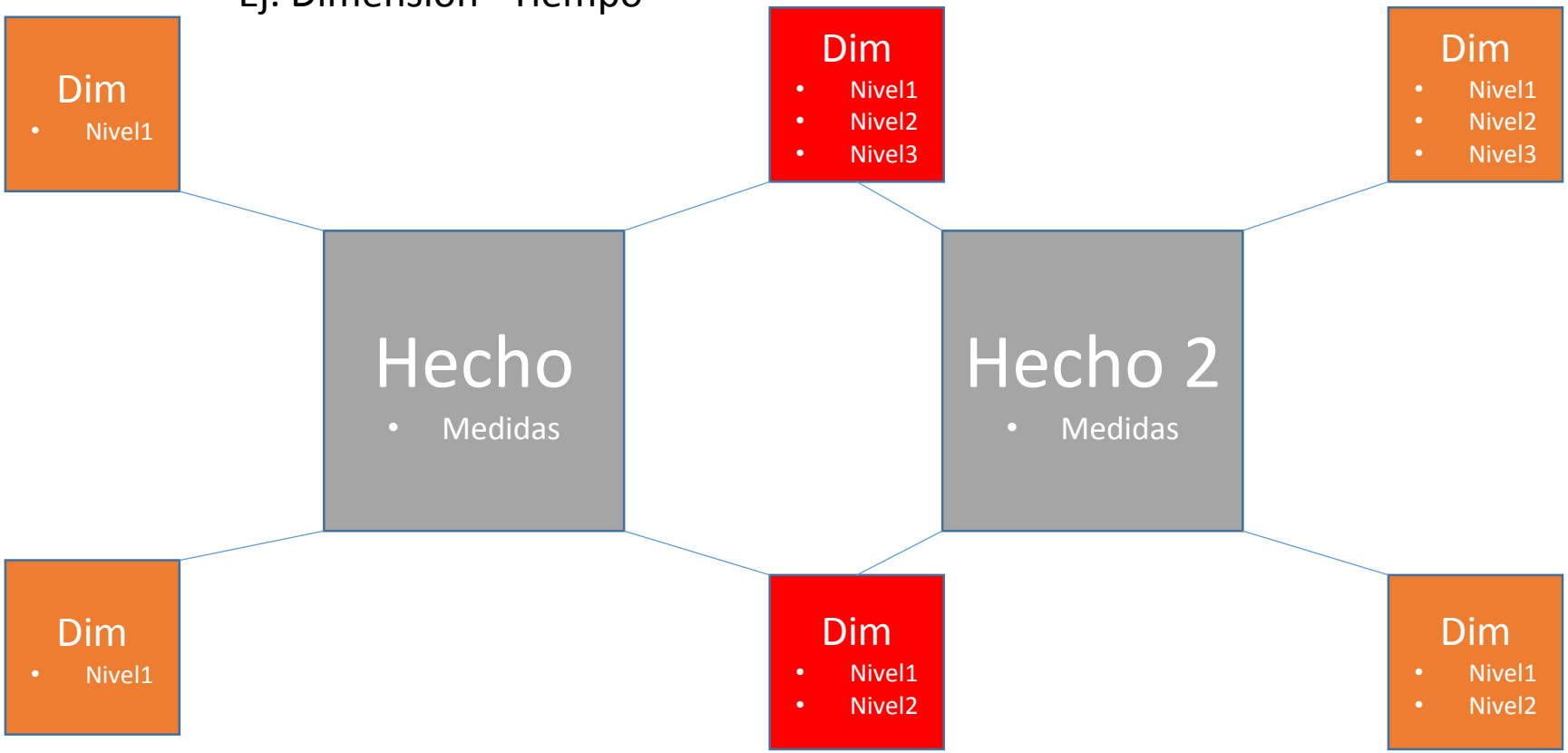
- Dimensiones que menos cambian denormalizadas
- Dimensiones que más cambian y relaciones muchos-muchos normalizadas.

Ejemplo (Tablero)

- Diseñar Estrella
- Hecho
  - Ventas
    - Cantidad, Valor
- Dimensiones
  - Tiempo
    - Día, Mes, Año
  - Producto
    - Producto, Línea
  - Cliente
  - Tienda
    - Tienda, Ciudad

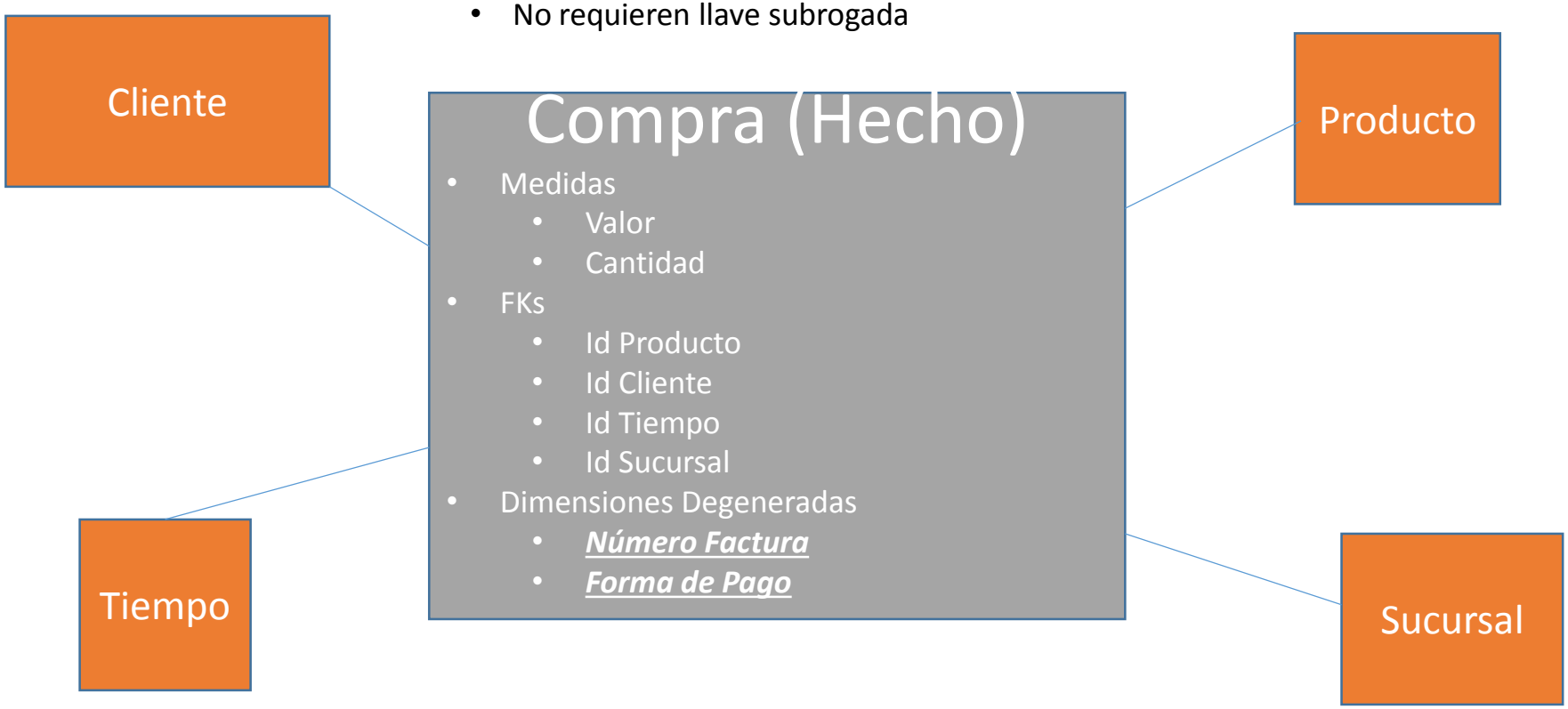
Dimensiones Conformadas

- Dimensiones relacionadas con varios Hechos
  - Ej. Dimensión “Tiempo”



Dimensiones Degeneradas

- Se incluyen en la tabla de hechos
- Muy pocos atributos (Id de negocio, Etiqueta)
- No cambian constantemente
- No requieren llave subrogada



Dimensiones que cambian en el tiempo

Opciones

- No mantener historial
  - Registro original
  - Sobreescribir
- Agregar filas
  - Se deben agregar campos para controlar versiones
  - Versiones infinitas
- Agregar columnas
  - Versiones limitadas
- Tabla con historial

Modelo Físico

- ROLAP
  - Relational OLAP
  - Utilizar BD Relacionales para guardar dimensiones y hechos
    - Esquema Estrella y Copo de Nieve
- MOLAP
  - Multidimensional OLAP
  - Propietario
  - Cálculos agregados preparados y almacenados
  - Alto desempeño
- HOLAP (Híbrido)
  - Detalles ROLAP
  - Agregados MOLAP

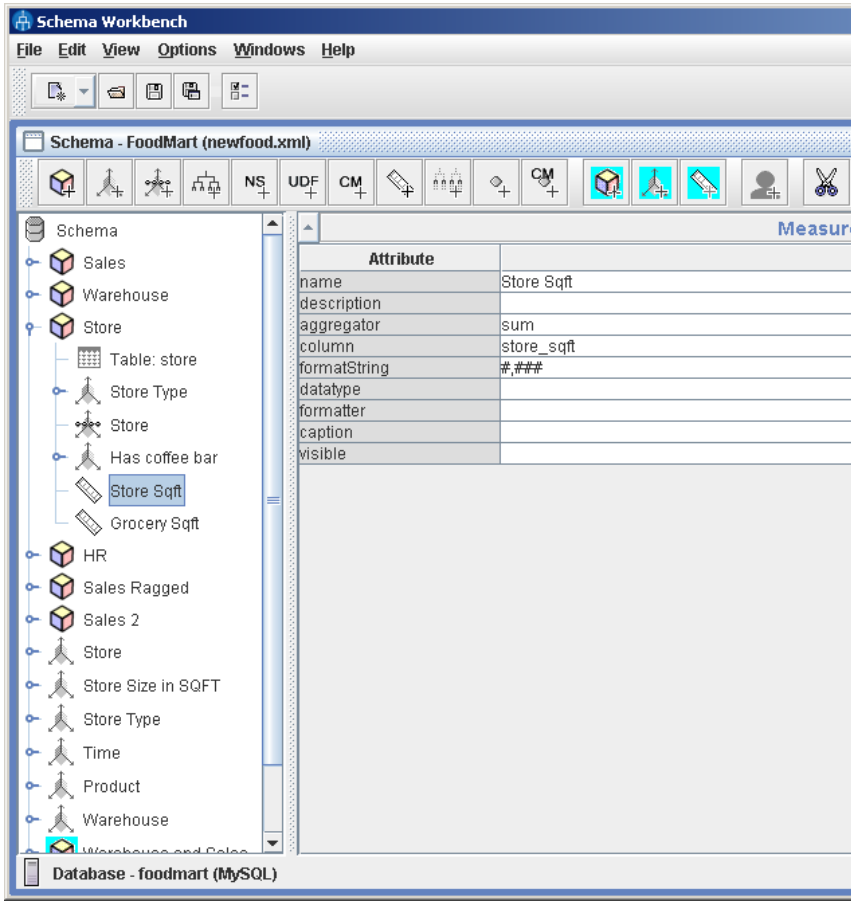
Análisis

- Reportes
  - Predefinidos, estáticos.
  - Cumplimiento de req. Gubernamentales
- OLAP
  - Interactivo
  - Drill Down – Bajar en una Jerarquía
    - Ej. Ventas por año -> Ventas por Semestre
  - Roll Up – Subir en una Jerarquía
  - Agregar/Remover dimensiones del análisis.
- Minería de Datos
  - Métodos estadísticos, aprendizaje, predicciones
  - Encontrar relaciones implícitas

Análisis OLAP

- Mapeo tablas-negocio
  - Tablas
  - Jerarquías
  - Nombres de Negocio
  - Significado de llaves foráneas

- El usuario final utiliza los **metadatos de negocio** que el motor OLAP convierte en **metadatos técnicos** para formular **SQL**



[Read Only] Oracle BI Administration Tool - SampleAppLite_BI0001

File Edit View Manage Tools Diagram Window Help

Presentation

- Sample Sales Lite
  - Time
  - More Time Objects
  - Products
  - More Product Objects
  - Offices
  - More Office Objects
  - Orders
  - Orders Dates
  - Base Facts
  - Calculated Facts
  - Time Series
- Sample Targets Lite
  - Time
  - More Time Objects
  - Products
  - Offices
  - More Office Objects
  - Base Facts

Business Model and Mapping

- SampleApp Lite
  - H0 Time
  - H1 Products
  - H2 Offices
  - H3 Orders
  - D0 Time
  - D1 Products
  - D2 Offices
  - D3 Orders (Facts Attributes)
  - F0 Revenue Base Measures
  - F1 Targets Base Measures

Physical

- Sample App Lite Data
  - SampleApp_Lite_Xml
  - D0 Time Dimension Tables ----------
  - D01 Time Day Grain
  - D02 Time Month Grain
  - D10 Product
  - D20 Offices
  - F0 Fact Tables ----------
  - F10 Revenue Facts (Order grain)
  - F11 Revenue Facts Aggregate
  - F20 Targets Facts
  - Samp_ Lite Source Tables ==========
  - SAMP_OFFICES_D
  - SAMP_PRODUCTS_D
  - SAMP_REVENUE_A
  - SAMP_REVENUE_F
  - SAMP_TARGETS_F
  - SAMP_TIME_DAY_D
  - SAMP_TIME_MTH_D
- System DB (Update me)

For Help, press F1

Inhibidores para BI

- Política / Poder
  - Los administradores se sienten intimidados y sienten que pierden poder.
  - BI no toma decisiones. Ayuda a tomarlas y sustentarlas
- Visión a corto plazo
  - BI da respuestas que posiblemente ya se conocían.
  - http://blog.consultorartesano.com/2006/11/no-te-fes-de-los-consultores.html
- Resistencia al cambio
  - Una estrategia de BI completa requiere la participación de muchas áreas

Taller Modelo Multidimensional

- Construir Modelo estrella Y copo de nieve
  - Por lo menos 2 dimensiones con jerarquías.
  - Telco
    - Llamadas
  - Salud
    - Autorizaciones, Atenciones
  - Banca
    - Transacciones de tc
  - Seguros
    - Siniestros
  - Otro …