

INF-239 ILI-239 Bases de Datos

Profesora
Cecilia Reyes C.
cecilia.reyes@usm.cl





Departamento de Informática
Universidad Técnica Federico Santa María



Unidad I

Conceptos y Propiedades de las Bases de Datos

Bases de Datos - Profesora Cecilia Reyes Covarrubias – Casa Central
Diapositivas realizadas con la colaboración Prof. J.Luis Martí – Campus San Joaquín



TEMARIO UNIDAD I

1.1 Definición de Bases de Datos

1.2 Enfoque Tradicional de Archivos versus Enfoque de Base de Datos

1.3 Tipos de Bases de Datos

1.4 Proceso de Diseño de Bases de Datos



1.1 DEFINICIÓN DE BASES DE DATOS



Departamento de Informática
Universidad Técnica Federico Santa María



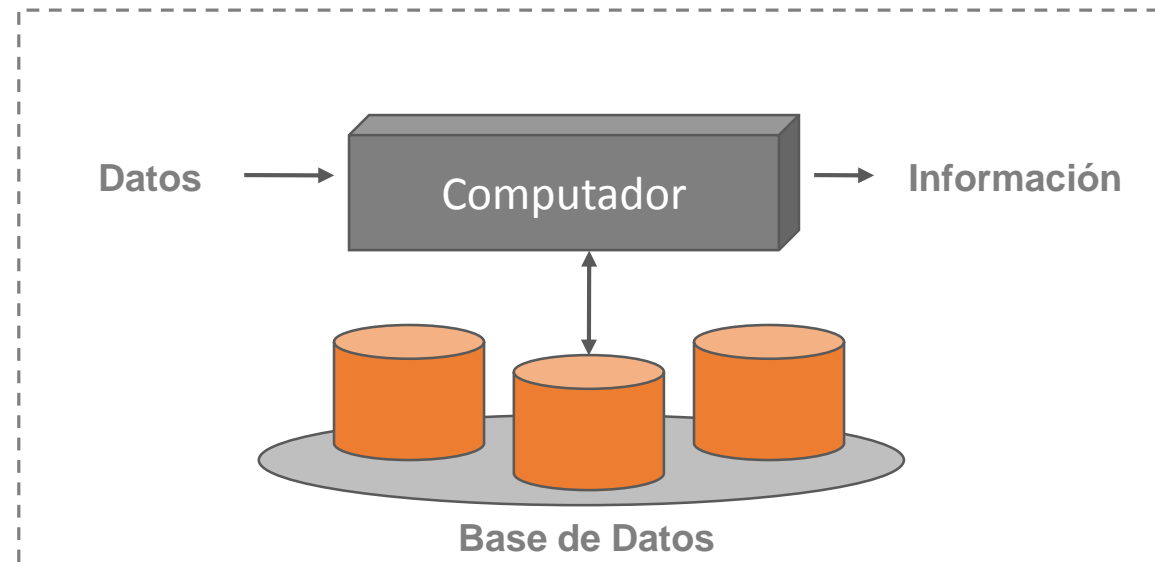
DATO COMO UN RECURSO

- En una organización existe una gran cantidad de recursos: humano, material (tecnológico), financiero y dato.
- El dato es considerado un recurso porque:
 - Tiene un valor intrínseco.
 - Aporta información valiosa al desempeño y a la toma de decisiones.
 - Su obtención, almacenamiento y control involucran gastos (inversión).



DEFINICIONES

- **Dato:** hecho relacionado con personas, objetos, lugares, eventos u otras entidades del mundo real. Pueden ser cualitativos (descriptivos) o cuantitativos, internos o externos, históricos o predictivos.
- **Información:** datos organizados o preparados (procesados y formateados) de una forma que sea adecuada para la toma de decisiones u otras actividades de la organización.
- **Base de Datos:** conjunto de archivos de datos relacionados entre sí donde se almacenan datos relevantes para la organización y que posteriormente serán recuperados para transformarlos en información .



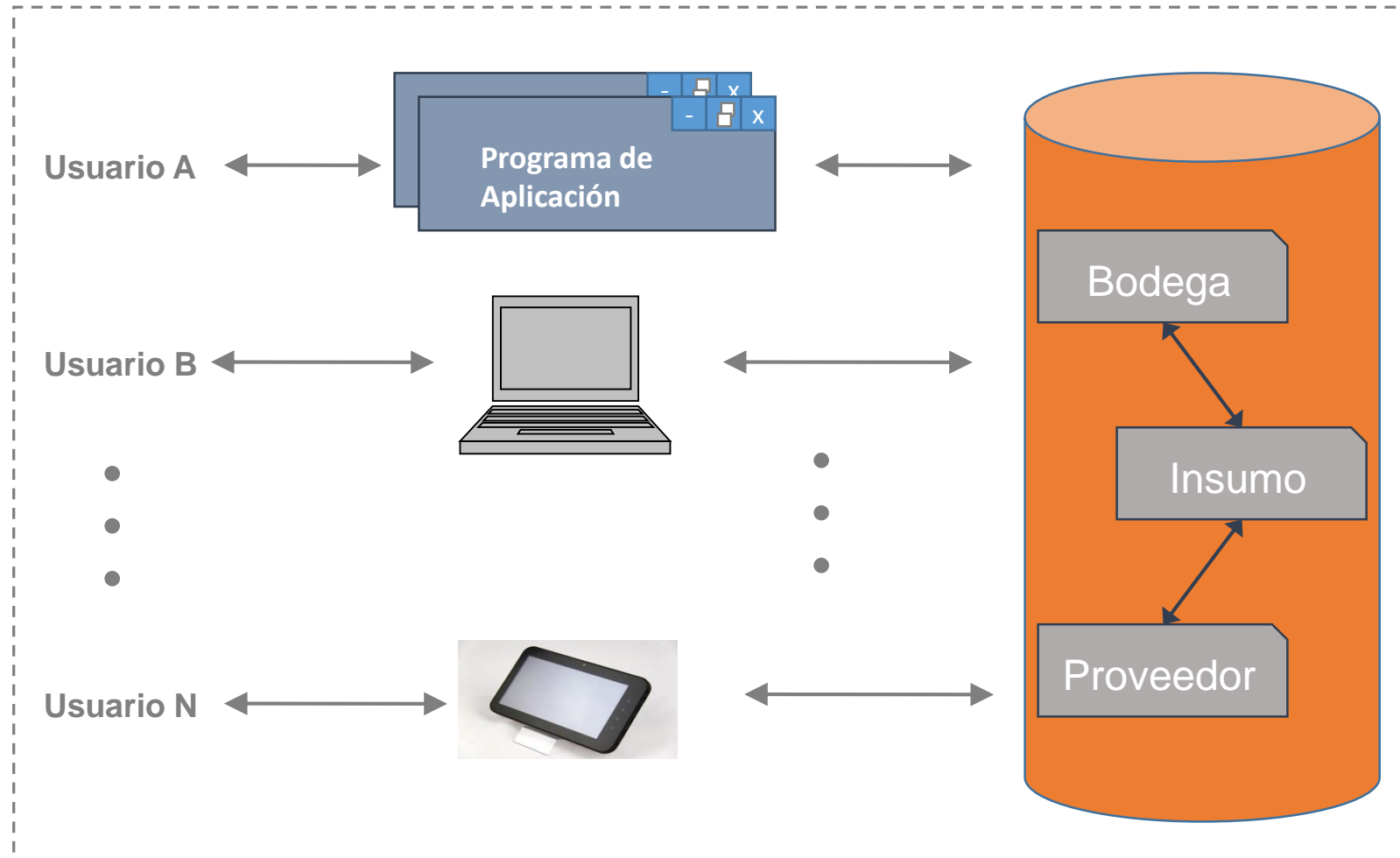
DEFINICIONES

- Los datos se pueden guardar, principalmente en archivos planos o bases de datos.
- Un **archivo plano** almacena datos que comparten una misma estructura y/o comportamiento similar. Por lo general, los datos de un archivo se refieren a un mismo **tipo de entidad** del mundo real. Por ejemplo, los Alumnos de la USM. Cada entidad representada por el archivo se guarda en un **registro** que se describe a través de sus **atributos**.

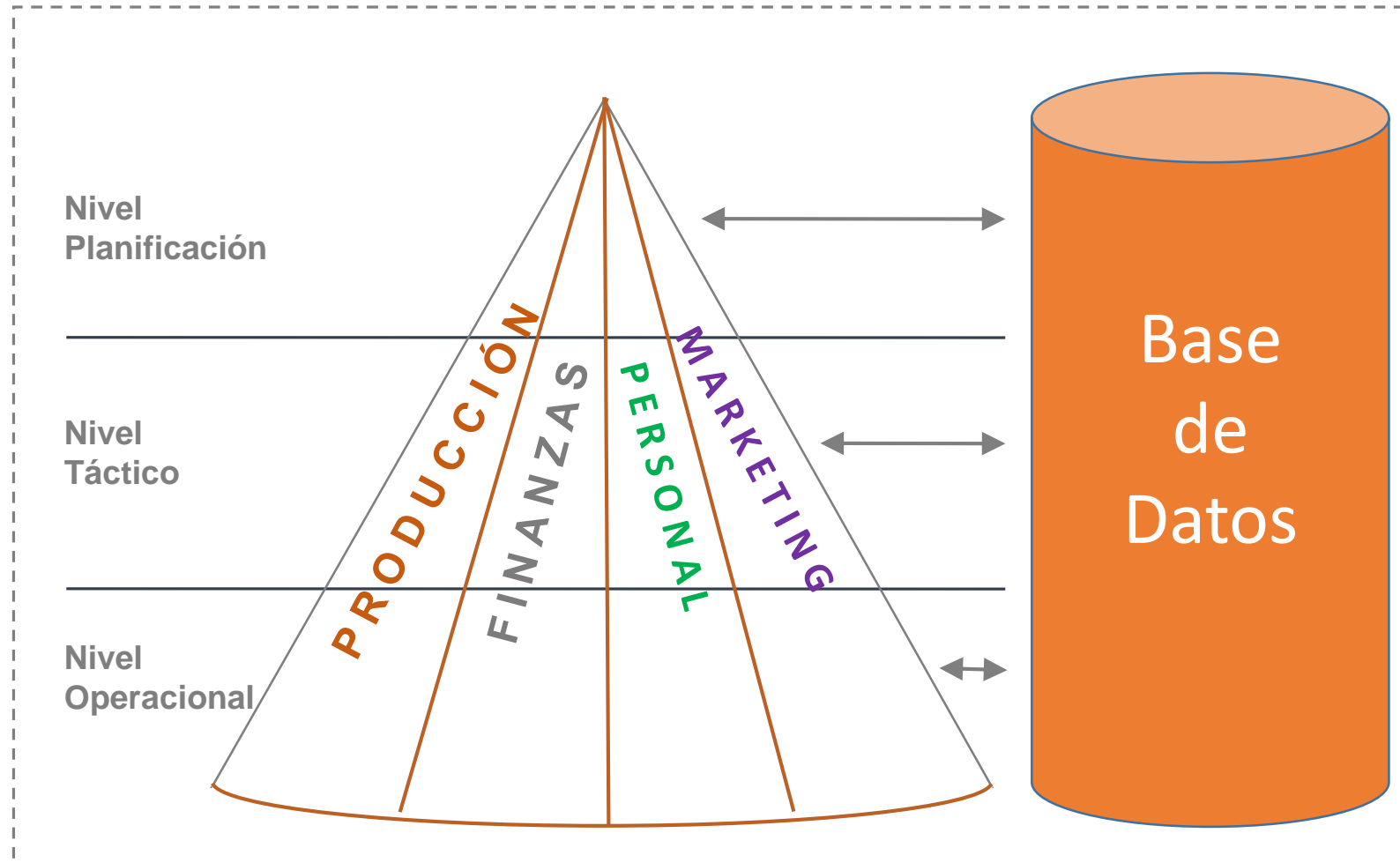
RUT	Nombre	Sexo	Región
18.345.678-9	Juan Pérez	M	5
18.223.344-5	María González	F	12
19.876.543-2	José Olivares	M	3
13.579.246-8	Rodrigo Martínez	M	5
12.121.212-2	Ana Castillo	F	8

- Una **base de datos** es un conjunto de archivos relacionados entre sí mediante alguna asociación lógica.

DEFINICION TÉCNICA DE BD



DEFINICION ORGANIZACIONAL DE BD



1.2 ENFOQUE TRADICIONAL DE ARCHIVOS V/S BD



Departamento de Informática
Universidad Técnica Federico Santa María

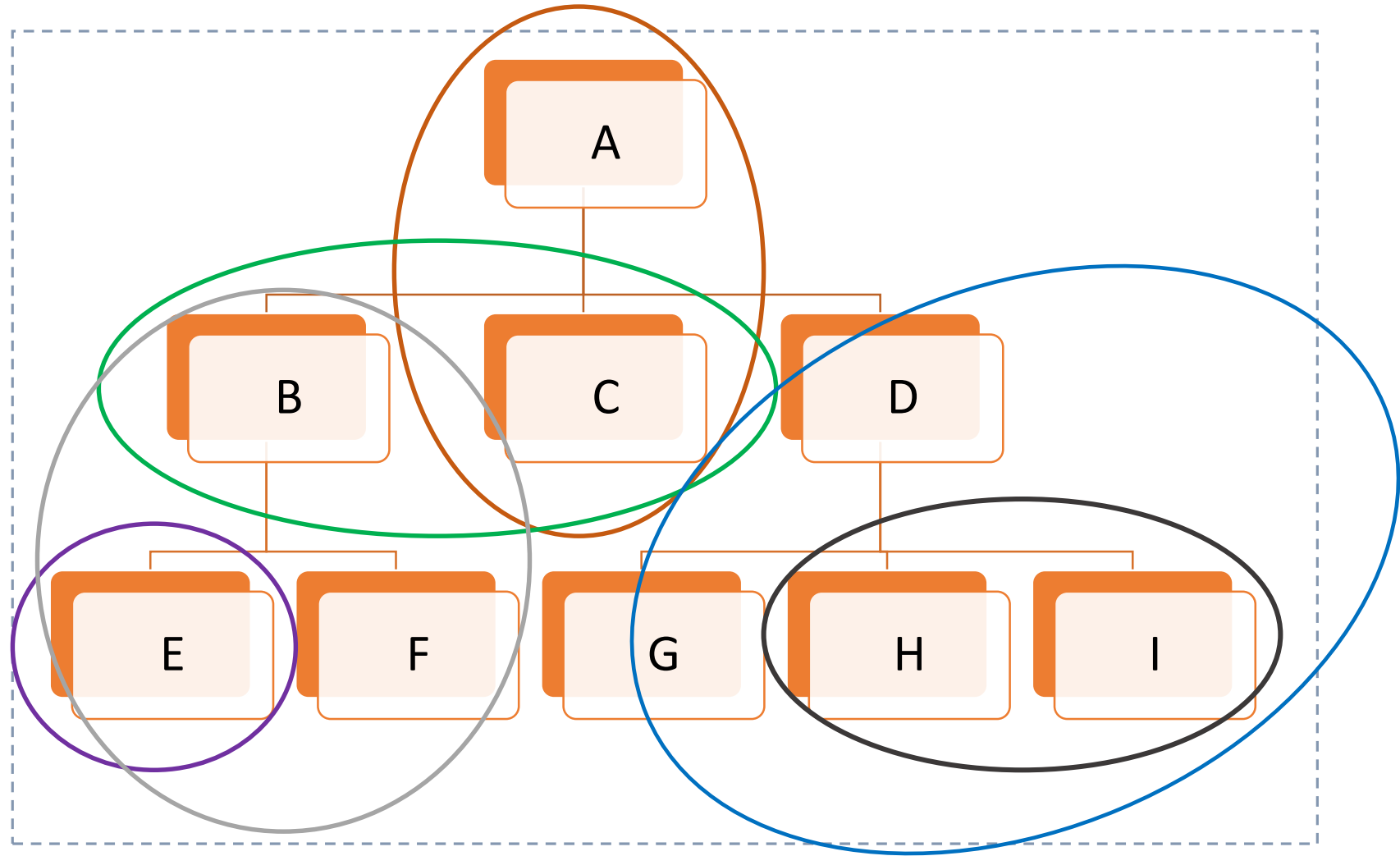


ENFOQUE DE ARCHIVOS

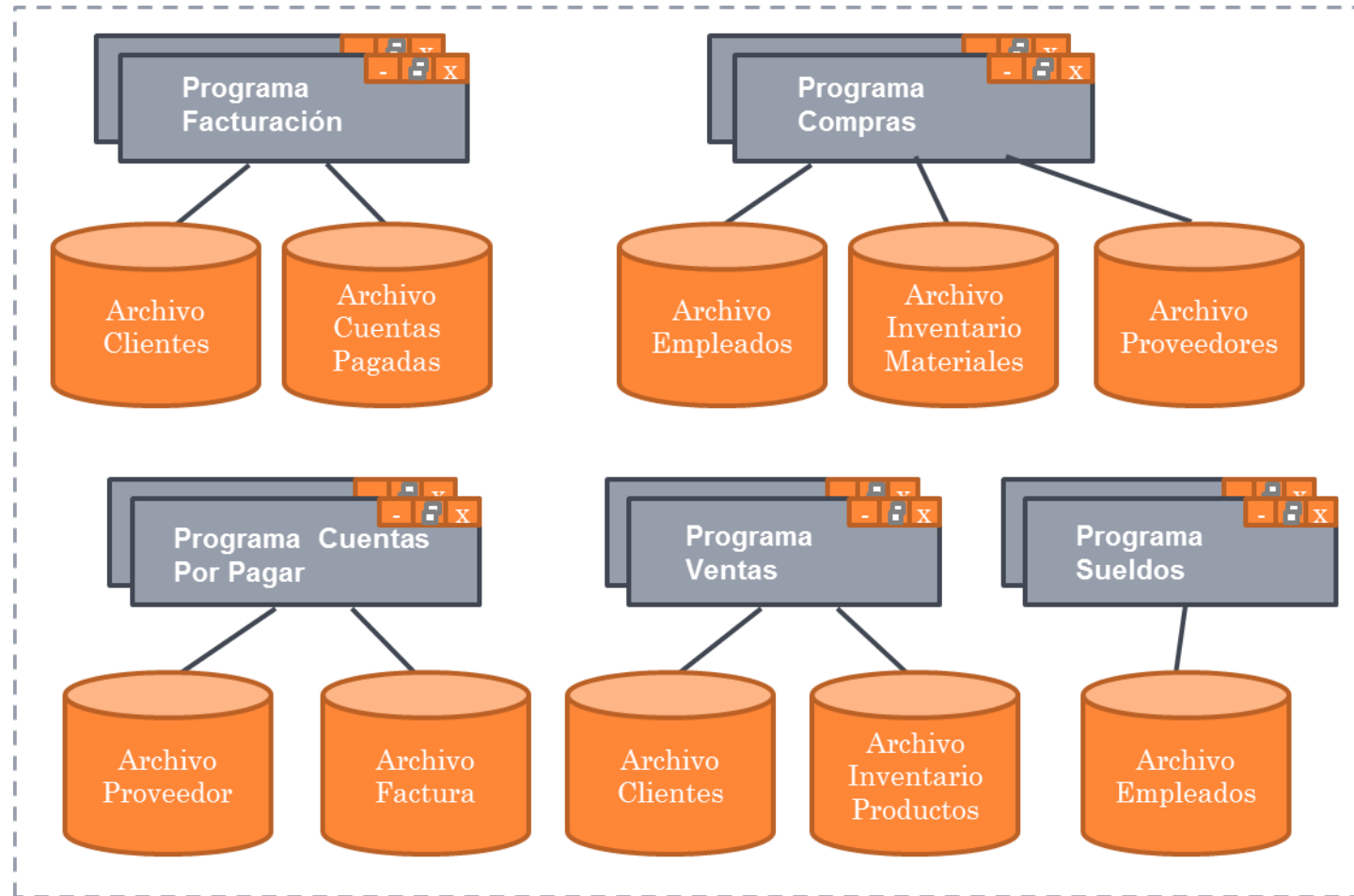
- Enfoque “del pasado” usado para el procesamiento de datos, también conocido como Enfoque por Agregación.
- Antiguamente, las organizaciones desarrollaban sus sistemas de información en forma aislada, sin existir una fuerte comunicación entre sus unidades.



ENFOQUE DE ARCHIVOS



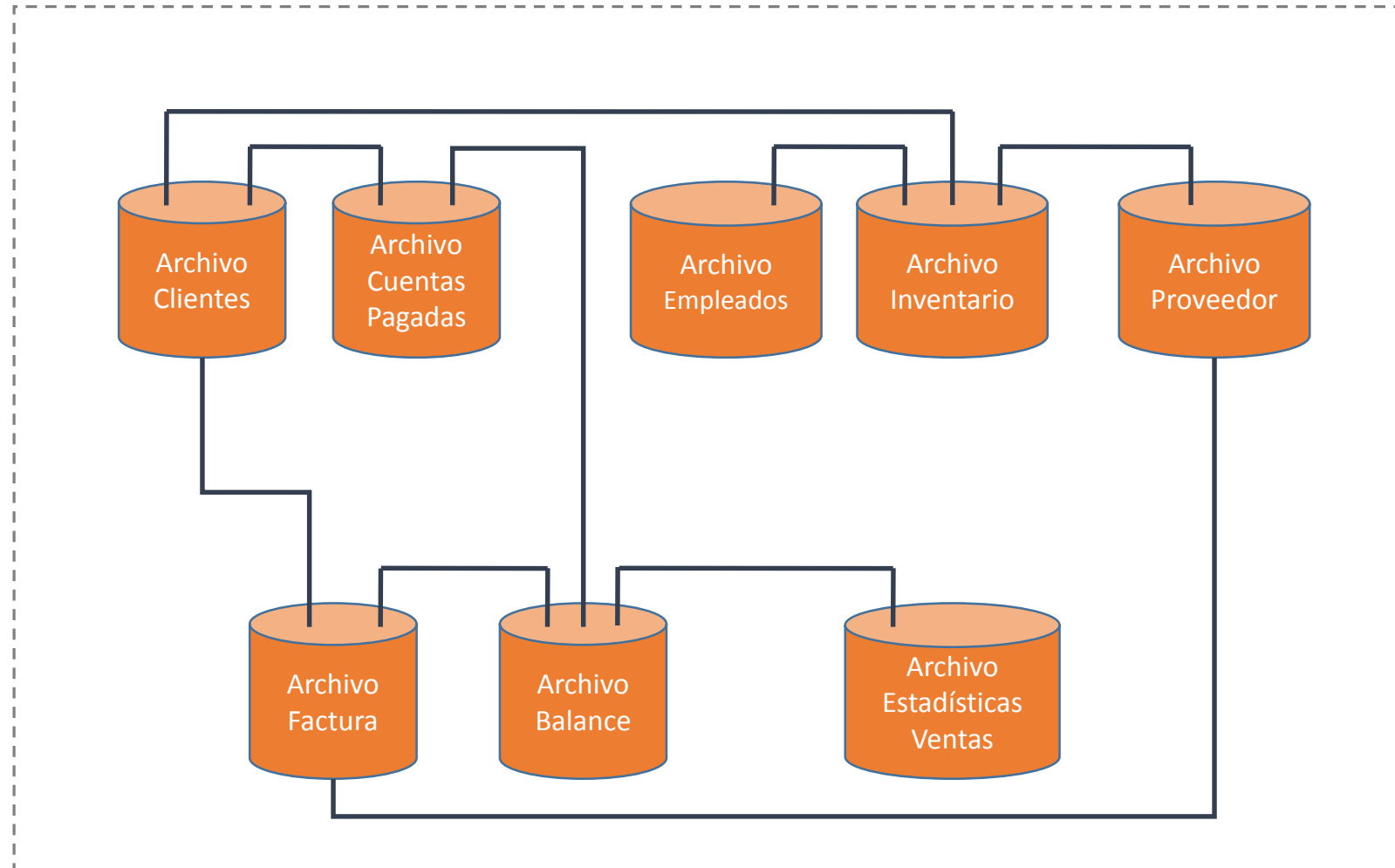
ENFOQUE DE ARCHIVOS



ENFOQUE DE ARCHIVOS - DESVENTAJAS

- Redundancia no controlada (subutilización del espacio en disco)
- Dependencia de los datos de los programas de aplicación
- Pobre estandarización
- Inconsistencia de los datos
- Problemas con el cliente
- Baja productividad del desarrollador

ENFOQUE DE BASES DE DATOS

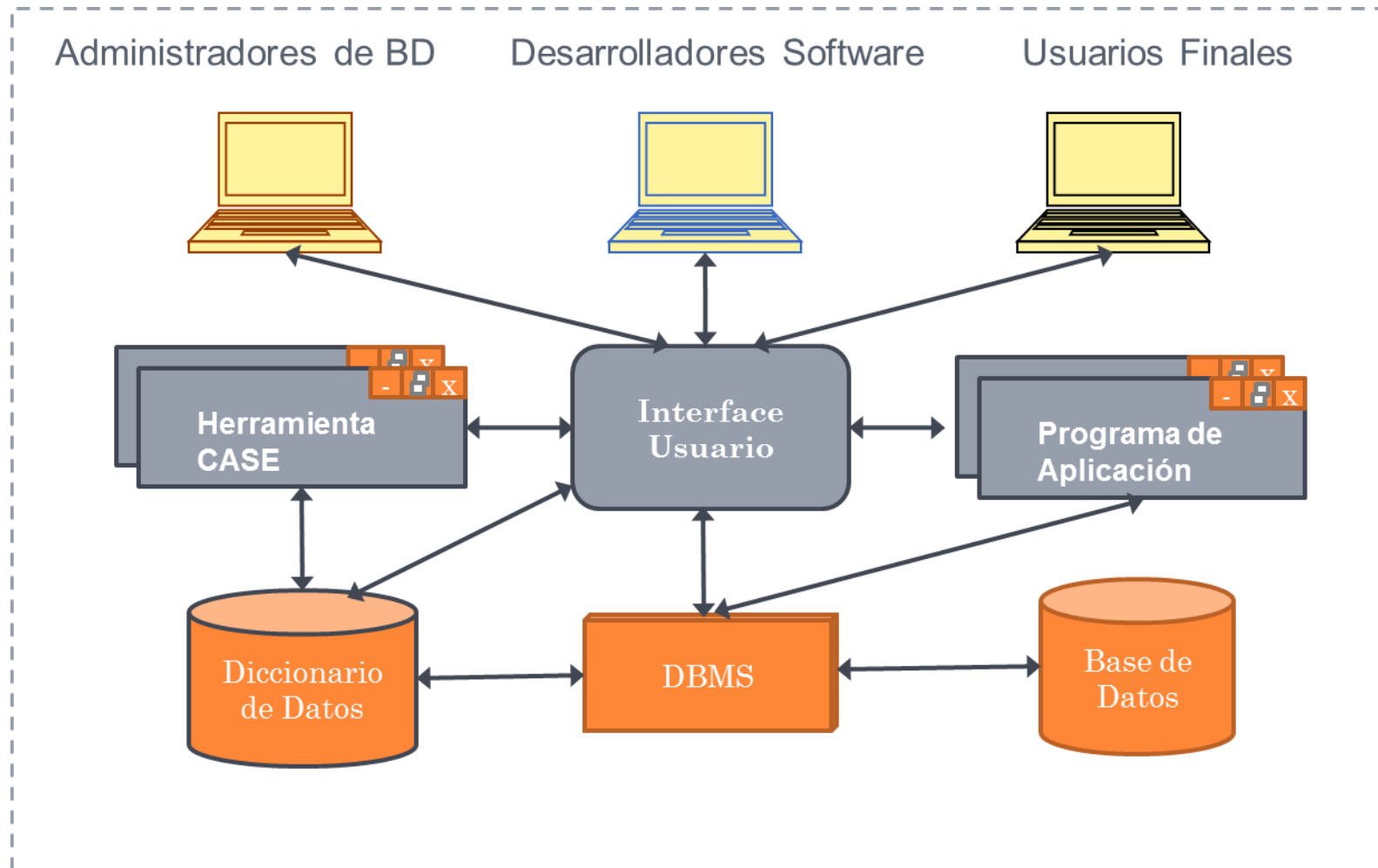


ENFOQUE DE BASES DE DATOS

COMPARACION CON DESVENTAJAS ARCHIVOS

- Redundancia no controlada (subutilización del espacio en disco)
 - Dependencia de los datos de los programas de aplicación
 - Pobre estandarización
 - Inconsistencia de los datos
 - Problemas con el cliente
 - Baja productividad del desarrollador
-
- Visión centralizada, compartida y única de los datos.
 - Minimización de la redundancia.
 - Independencia de los datos de los programas de aplicación.
 - Estandarización.
 - Integración y seguridad de datos, generan una mayor consistencia de ellos.
 - Facilidades para el diseño y desarrollo de aplicaciones, mejoran la relación con el cliente y la productividad del desarrollador.

COMPONENTES ENFOQUE DE BASES DE DATOS



COMPONENTES ENFOQUE DE BASES DE DATOS

USUARIOS

- Personas con requerimientos de información que realizan operaciones de mantención (ingreso, modificación, eliminación) y consulta a la base de datos, como así también especialistas en el área que se preocupan de su administración.
- Incluyen: usuario final, desarrollador de aplicaciones, diseñador de bases de datos, administrador de bases de datos (DBA), administrador de datos (arquitecto de datos).



COMPONENTES ENFOQUE DE BASES DE DATOS

DBMS (Data Base Management System)

- Sistema Administrador de Bases de Datos (SABD), software que permite crear y mantener una o más bases de datos en forma centralizada o distribuida.
- Conocido también como motor o servidor de datos.
- Sus principales funciones:
 - Definición de Datos (DDL): create, alter, drop...
 - Manipulación de Datos (DML): insert, update, delete...
 - Control de Datos (DCL): grant, revoke...
- RDBMS: Relational Data Base Management System



PostgreSQL



COMPONENTES ENFOQUE DE BASES DE DATOS

BASE DE DATOS (Data Base)

- Conjunto de datos operacionales, almacenados en el computador y accedidos por distintas aplicaciones.
- Lugar físico donde están almacenados los datos (base de datos física) o más específicamente los valores de dichos datos.

DICCIONARIO DE DATOS (Data Dictionary)

- Base de datos que guarda una descripción de los datos (o metadatos), como su tipo, largo, propietario, tamaño de los registros, etc.
- También se le conoce como base de datos lógica, schema, catalogo.



COMPONENTES ENFOQUE DE BASES DE DATOS

INTERFACE USUARIO

- Forma en que el SABD permite la interacción con la base de datos.
- Pueden ser líneas de comando, lenguajes, menus, pantallas, sistemas Web, sistemas de reconocimiento de la voz,

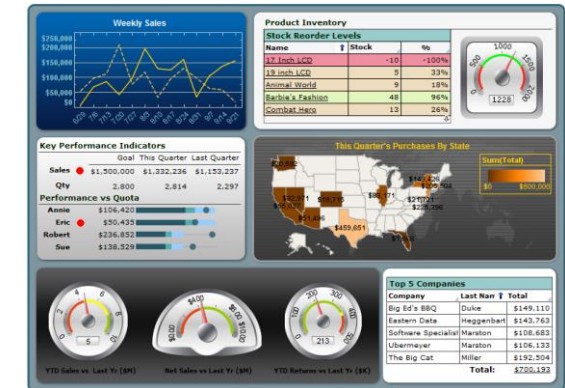
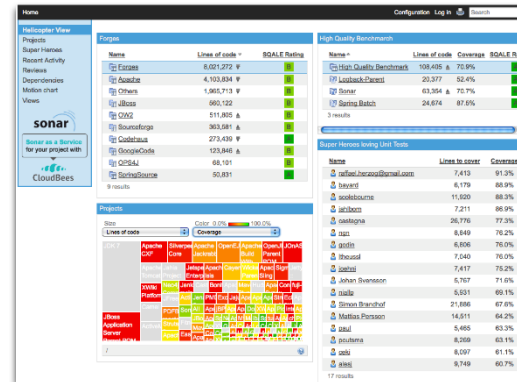
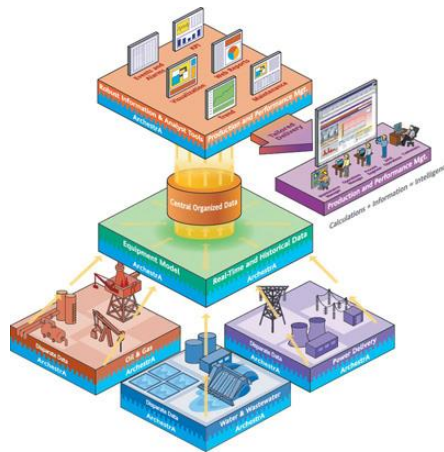
HERRAMIENTA CASE (Computer Aided Software Engineering)

- Software que ayuda al desarrollador de aplicaciones en todas las etapas del ciclo de vida de un software.
- En el caso de las BD ayudan a generar el modelo de datos, incluso algunas generan código SQL.
- Ejemplos de CASE usadas en BD: Erwin, EasyCASE, Oracle Designer, DBDesigner, MySQL Workbench

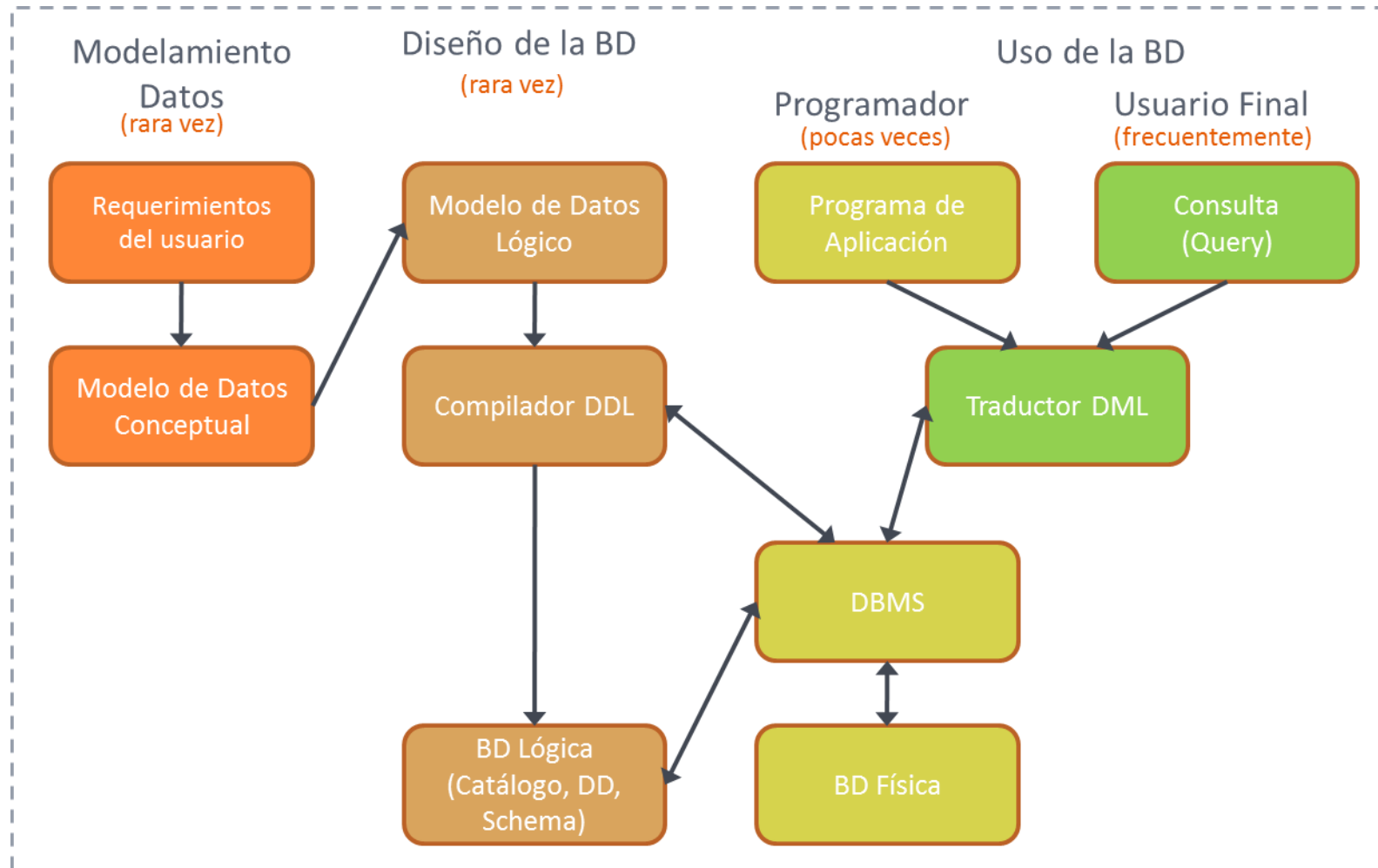
COMPONENTES ENFOQUE DE BASES DE DATOS

PROGRAMA DE APLICACIÓN

- Programas computacionales escritos por los desarrolladores principalmente para:
 - Poblar inicialmente la BD (importancia de la migración de datos desde distintas fuentes).
 - Mantener en el tiempo la BD.
 - Generar información a los usuarios a través de reportes, informes, gráficos, dashboard, etc.



IMPLEMENTACIÓN ENFOQUE DE BASES DE DATOS



BD EN EL DESARROLLO DE SI

- **Sistema de Información** (SI): conjunto de aplicaciones (software), datos, recursos materiales (equipos) y personas (usuarios) que interactúan para procesar datos y convertirlos en información relevante para una organización.
- Etapas principales en el desarrollo de un SI:
 - Análisis
 - Diseño
 - Construcción (codificación o programación)
 - Implementación
 - Mantenimiento
- Relación con las BD:
 - Análisis: énfasis en el manejo integrado de los datos y en la generación de una estructura lógica de los datos, que se adapte a los **requerimientos de información** de los usuarios.
 - Diseño: conversión de la estructura lógica en especificaciones para archivos y programas que puedan ser implementadas por el DBMS disponible.

BD EN EL DESARROLLO DE SI

		Nivel Estratégico	Nivel Táctico	Nivel Operacional
	Decisión que Apoya	Planificación a largo plazo	Control gerencial	Control operacional
	Tipo de Decisión	No estructurada	Semi-estructurada	Estructurada
	Modelo más usado	Predictivo	Descriptivo	Normativo
Características de la información	<ul style="list-style-type: none"> ➤ Fuente ➤ Exactitud ➤ Amplitud ➤ Frecuencia ➤ Rango de Tiempo ➤ Uso 	Medio Ambiente Razonable Resumida A Solicitud Años Predicción	Registros Internos Buena Detallada Periódica Años Control	Operación Interna Exacta Muy Detallada Tiempo Real Meses Acción Diaria

TIPOS DE SI

OLTP

OnLine Transaction Processing

- SI Operacional o TPS (Transaction Processing Systems)
- SI Administrativos o MIS (Management Information Systems)

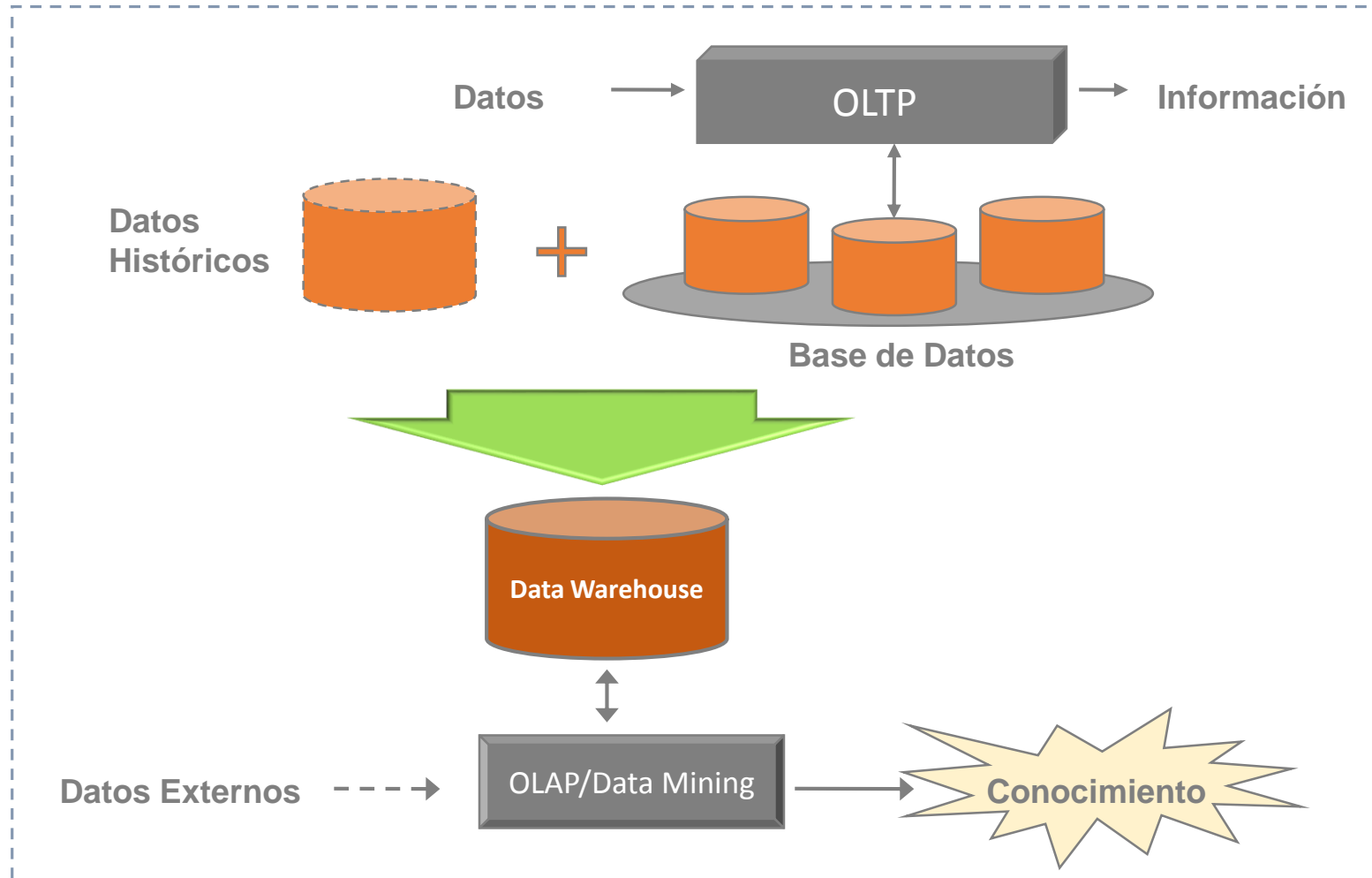
OLAP

OnLine Analytic Processing

- Sistemas de Apoyo a la toma de decisiones o DSS (Decision Support Systems)



TIPOS DE SI: OLTP v/s OLAP



1.3 TIPOS DE BD



Departamento de Informática
Universidad Técnica Federico Santa María



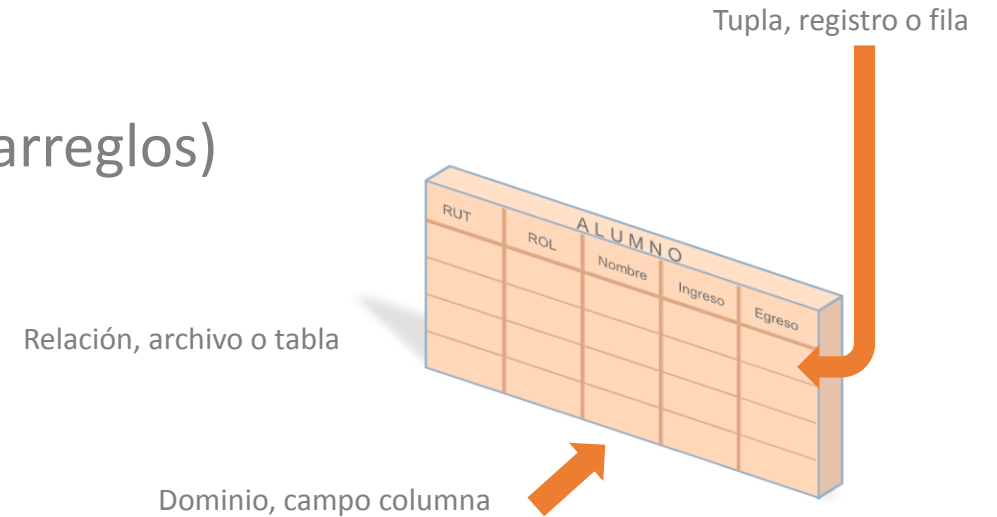
TIPOS DE BD

- Según estructura de datos usada
- Según nivel organizacional que apoyan
- Según tipo de dato almacenado
- Según ubicación de la copia principal de los datos
- Según número de procesadores que participan en el procesamiento de consulta
- Según número de sitios que participan en el almacenamiento de datos

TIPOS DE BD

SEGÚN ESTRUCTURA DE DATOS

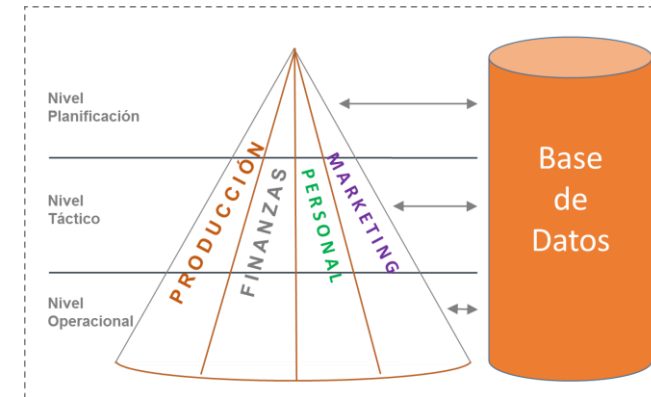
- Jerárquica (árbol)
- Reticular (grafo o red)
- Relacional (relation o tabla bidimensional)
- Orientada al objeto (clases de objetos)
- Multidimensional (cubo, hipercubo, conjunto arreglos)



TIPOS DE BD

SEGÚN NIVEL ORGANIZACIONAL QUE APOYAN

- BD Operacional (Transaccional - OLTP)
- BD de Gestión (Data Warehouse, Data Mart - OLAP)
- BD Estratégica (Data Warehouse – OLAP, Data Mining)



TIPOS DE BD

SEGÚN TIPO DE DATO ALMACENADO

- Estructurado y preciso (relacional)
- Agregado (multidimensional)
- Semiestructurado (espaciales, XML, textuales)
- No estructurado (web)

TIPOS DE BD

SEGÚN TIPO DE DATO ALMACENADO

RUT	NOMBRE	ESTADO CIVIL	OCUPACION	SEXO
1111-1	Felipe Camiroaga	Soltero	Animador	M
2222-2	Alexis Sánchez	Soltero	Futbolista	M
3333-3	Marlen Olivari	Divorciada	Show-woman	F
4444-4	Mauricio Isla	Casado	Futbolista	M

Dato estructurado, preciso

AÑO	2008 ▼
-----	--------

Dato agregado

Ventas Promedio	SEXO > ESTADO CIVIL									
	F					M				
PRODUCTOS	Soltero	Casado	Divorciado	Viudo	Total	Soltero	Casado	Divorciado	Viudo	Total
Notebook	50	20	100	30	200	100	50	500	100	750
MP4	10	5	30	5	50	20	30	100	50	200
iPhone	500	200	300	100	1100	300	200	200	300	1000
iPod	100	50	100	10	260	100	100	200	300	700
Pendrive	2000	50	1000	100	3150	1000	1000	1000	200	3200
Gran Total	2660	325	1530	245	4760	1520	1380	2000	950	5850

TIPOS DE BD

SEGÚN TIPO DE DATO ALMACENADO

```
<SalesOrder SONumber="12345">
  <Customer CustNumber="543">
    <CustName>ABC Industries</CustName>
    <Street>123 Main St.</Street>
    <City>Chicago</City>
    <State>IL</State>
    <PostCode>60609</PostCode>
  </Customer>
  <OrderDate>981215</OrderDate>
  <Item ItemNumber="1">
    <Part PartNumber="123">
      <Description>Stainless steel </Description>
      <Price>9.95</Price>
    </Part>
    <Quantity>10</Quantity>
  </Item> ...
</SalesOrder>
```

Dato semiestructurado



Dato no estructurado

TIPOS DE BD

SEGÚN UBICACIÓN COPIA PRINCIPAL DE LOS DATOS

- Basada en memoria principal (*in-memory database*) (1 nivel)
- Basada en el disco (2 niveles)
- Basada en almacenamiento terciario (3 niveles)

Oracle Database In-Memory

Powering the Real-Time Enterprise



Figure 1. Oracle's unique dual-format architecture

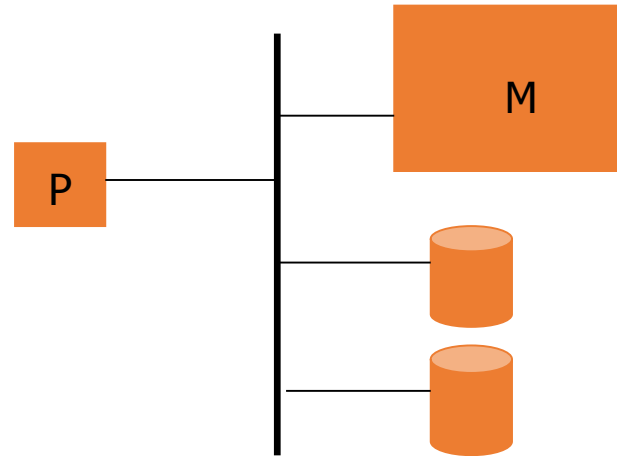
TIPOS DE BD

SEGÚN NUMERO DE PROCESADORES

- Serial (secuencial)
- Paralela:
 - Memoria Compartida (MC)
 - Nada Compartido (NC)
 - Disco Compartido (DC)
 - Arquitectura Híbrida (AC: Algo Compartido)

TIPOS DE BD

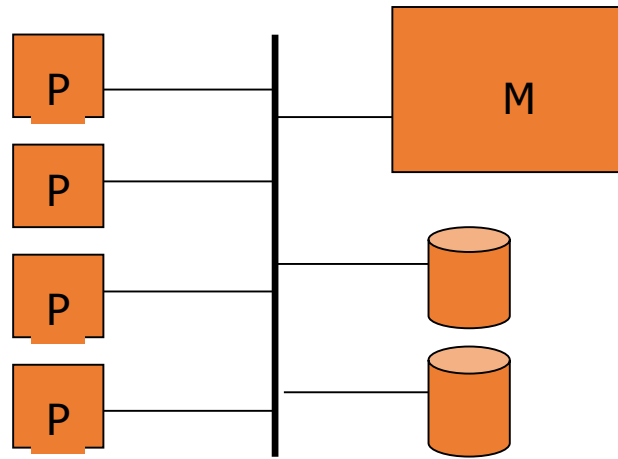
SEGÚN NUMERO DE PROCESADORES – BD SERIAL



Un procesador (P), una Memoria Principal (M) y discos de Memoria Secundaria para almacenar la BD

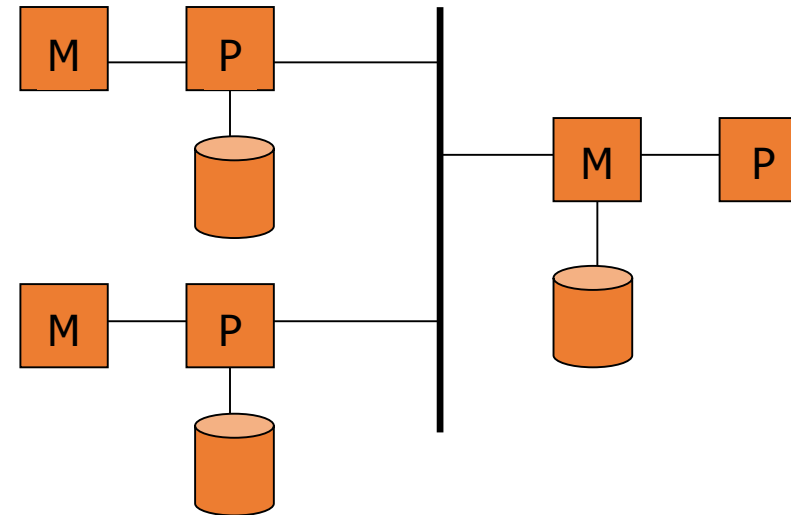
TIPOS DE BD

SEGÚN NUMERO DE PROCESADORES – BD PARALELA



Memoria Compartida

Varios procesadores (P) compartiendo Memoria Principal (M) y Discos (BD).

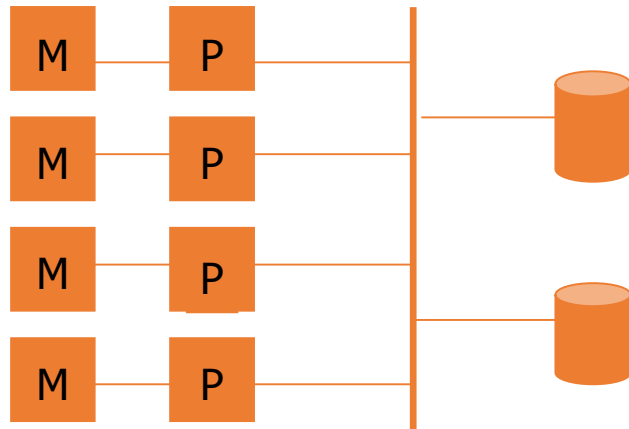


Nada Compartido

Varios procesadores (P) sin compartir Memoria Principal ni Discos (BD)

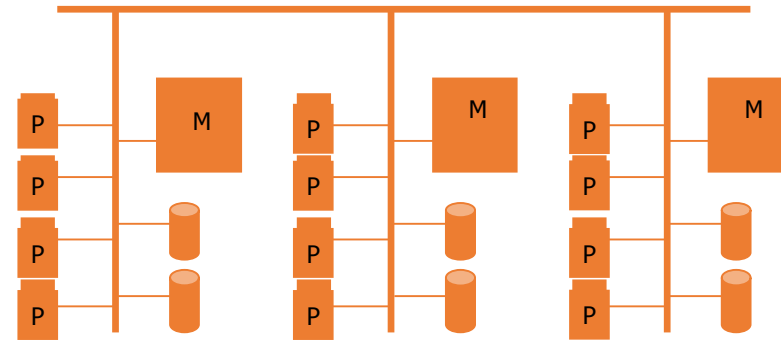
TIPOS DE BD

SEGÚN NUMERO DE PROCESADORES – BD PARALELA



Disco Compartido

Varios procesadores (P)
compartiendo Discos (BD), pero
tienen una memoria privada.



Arquitecturas Híbridas

Ejemplo de Clusters

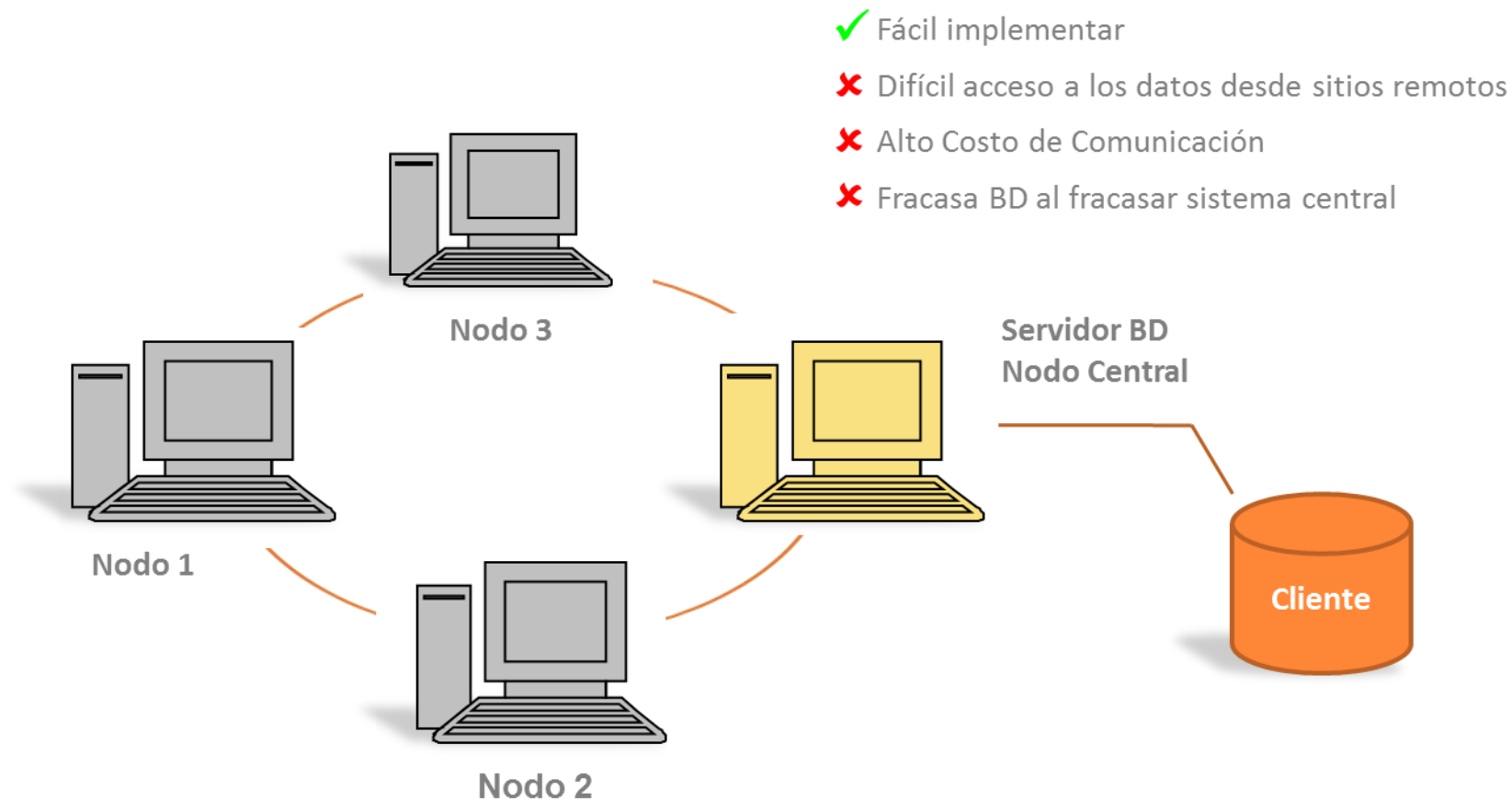
TIPOS DE BD

SEGÚN NUMERO DE SITIOS

- Centralizada
- Distribuida
- Otras:
 - Web
 - SMBD (Sistemas Múltiples BD o BD Federadas)
 - Móvil

TIPOS DE BD

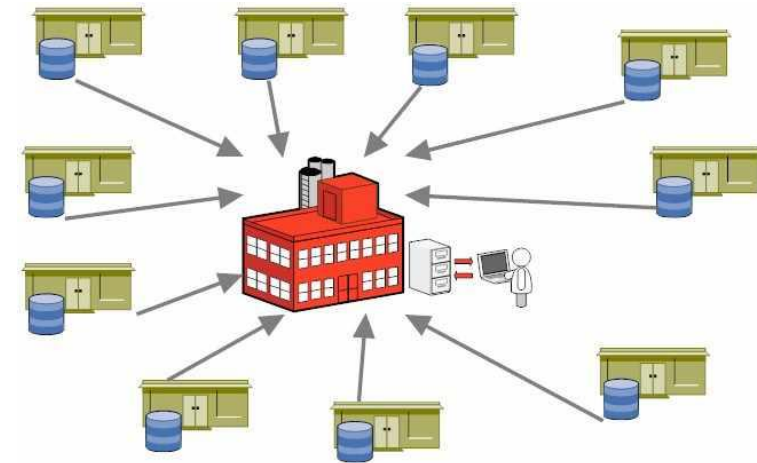
SEGÚN NUMERO DE SITIOS – BD CENTRALIZADA



TIPOS DE BD

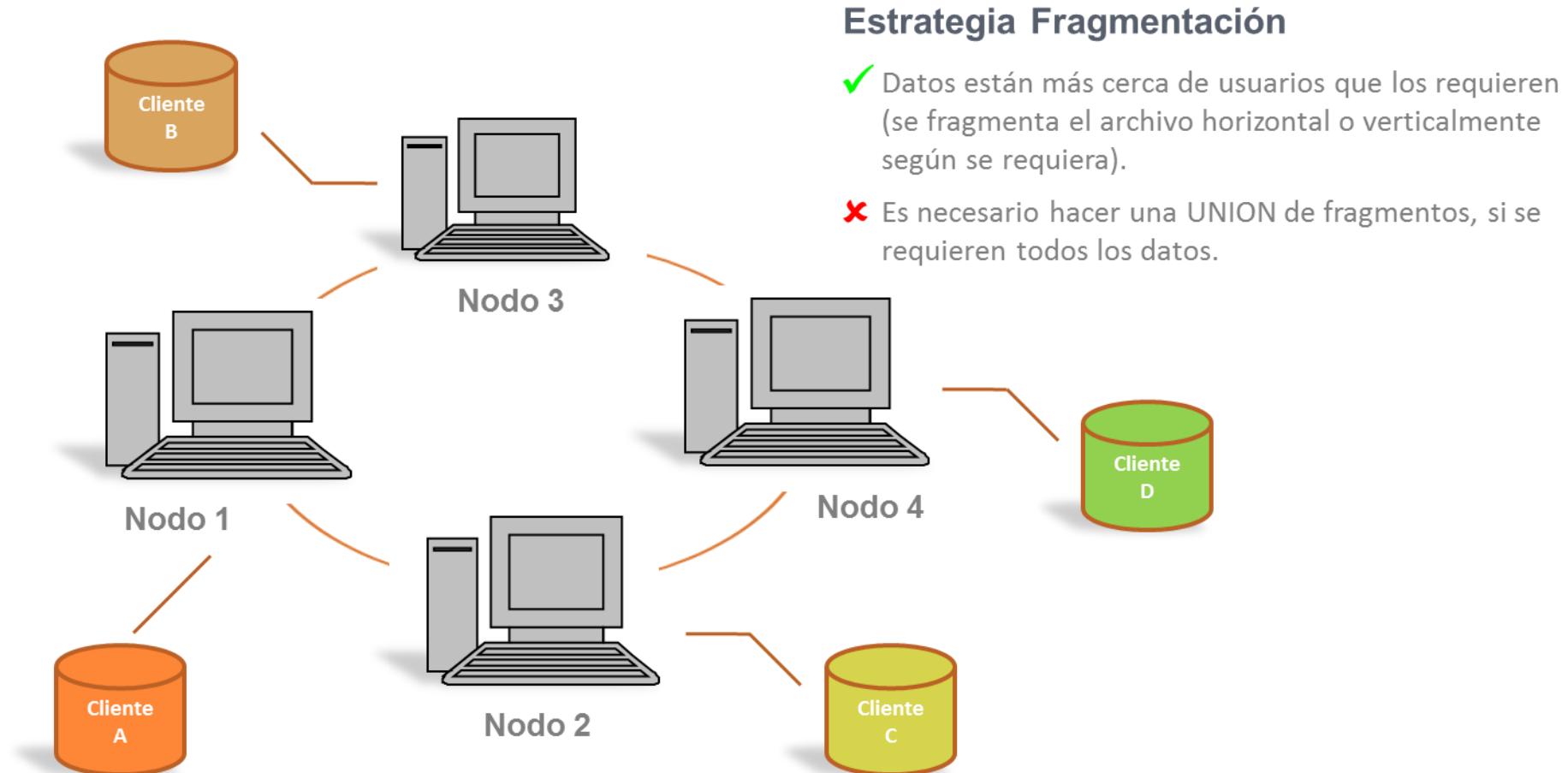
SEGÚN NUMERO DE SITIOS – BD DISTRIBUIDA

- BD Distribuida: base de datos lógica que es repartida físicamente entre computadores que están en distintos lugares pero conectados por una red.
- Estrategias de distribución de datos:
 - Fragmentación o Particionamiento
 - Horizontal
 - Vertical
 - Replicación
 - Híbrida



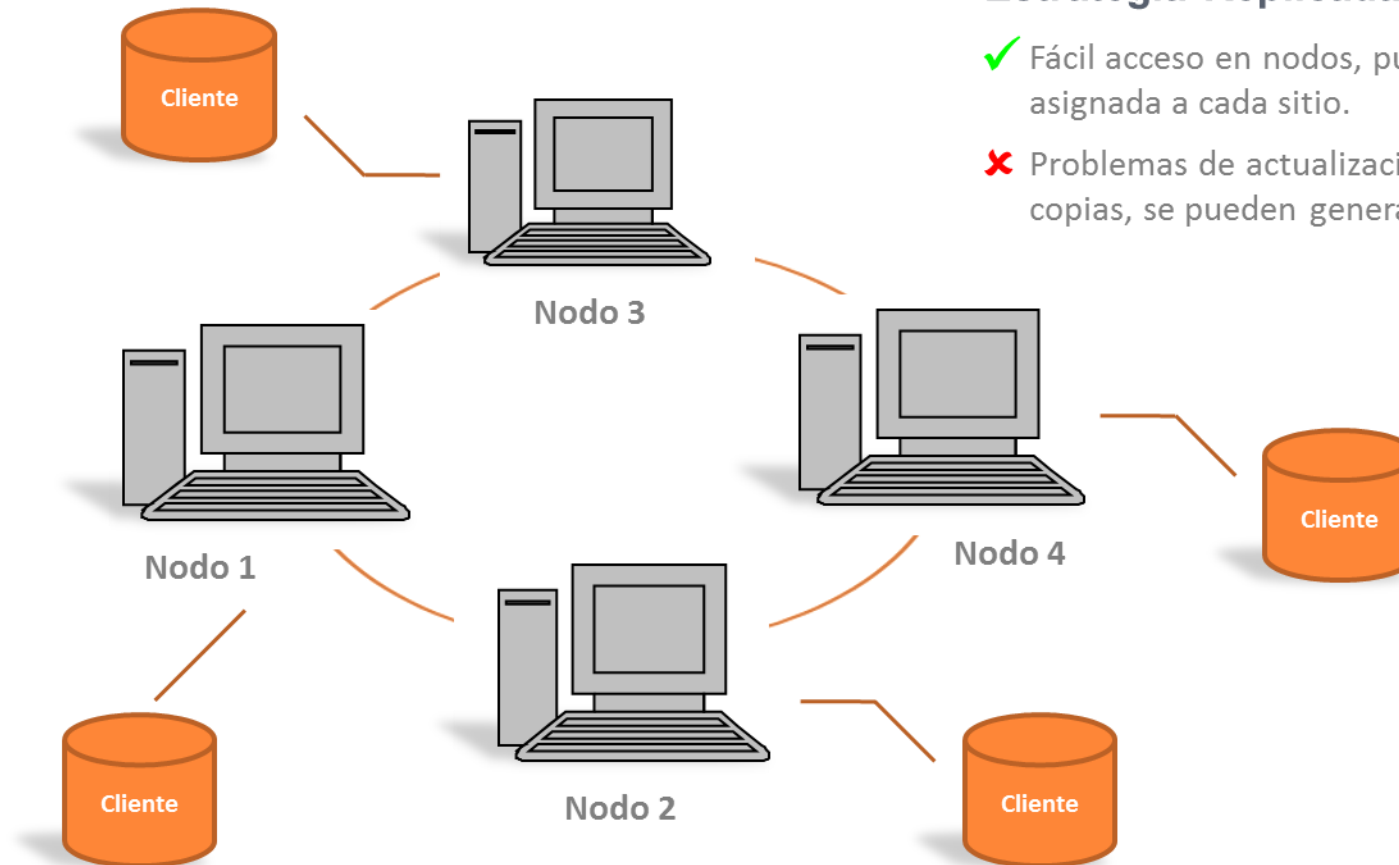
TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD DISTRIBUIDA



TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD DISTRIBUIDA



Estrategia Replicada

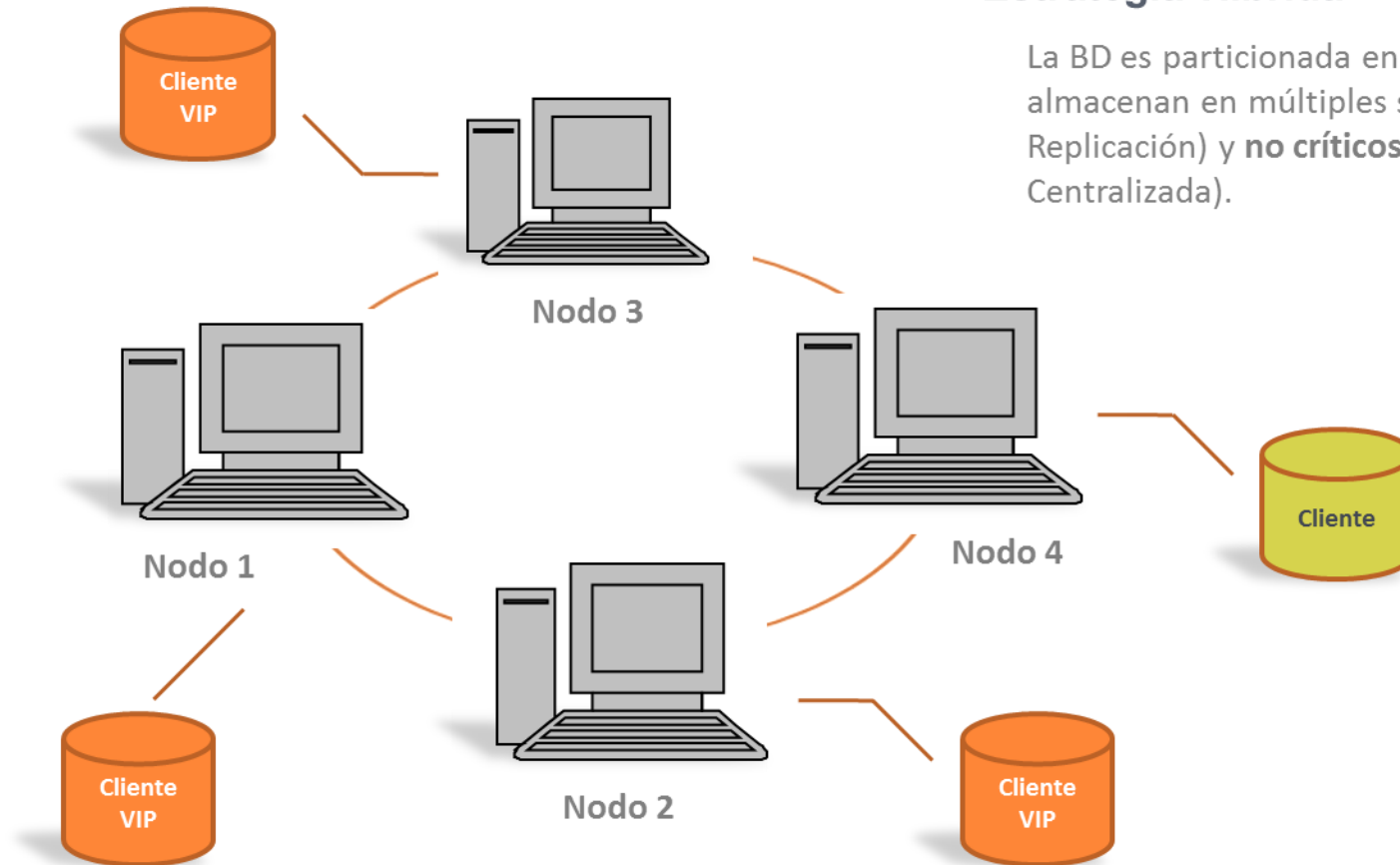
- ✓ Fácil acceso en nodos, pues una copia del archivo es asignada a cada sitio.
- ✗ Problemas de actualización al existir múltiples copias, se pueden generar inconsistencias.

TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD DISTRIBUIDA

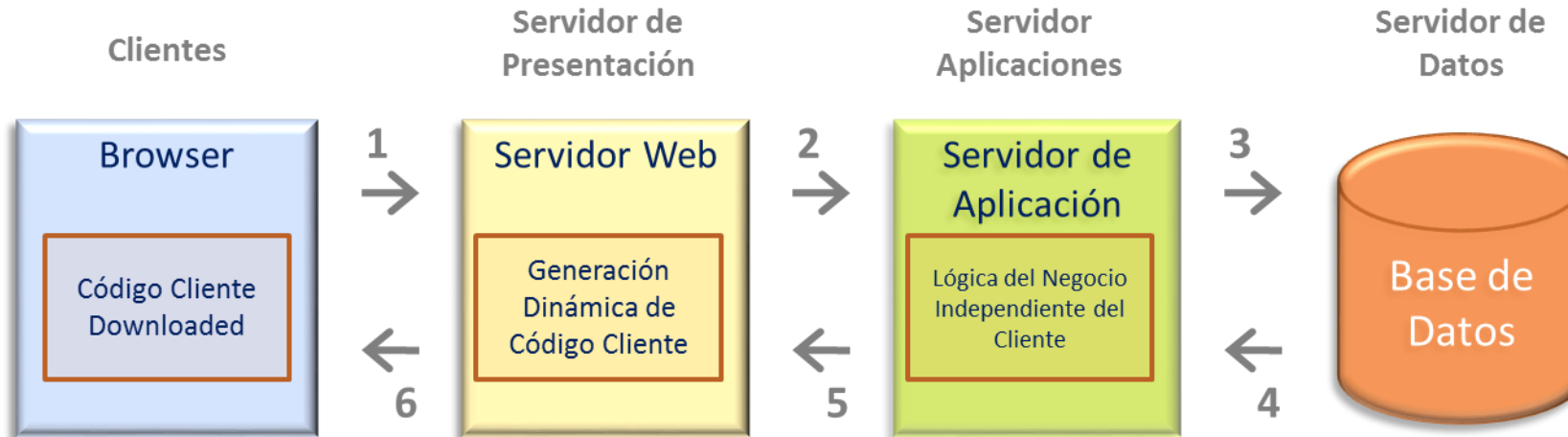
Estrategia Híbrida

La BD es particionada en fragmentos **críticos** (se almacenan en múltiples sitios: Fragmentación y Replicación) y **no críticos** (se almacenan en un sitio: Centralizada).



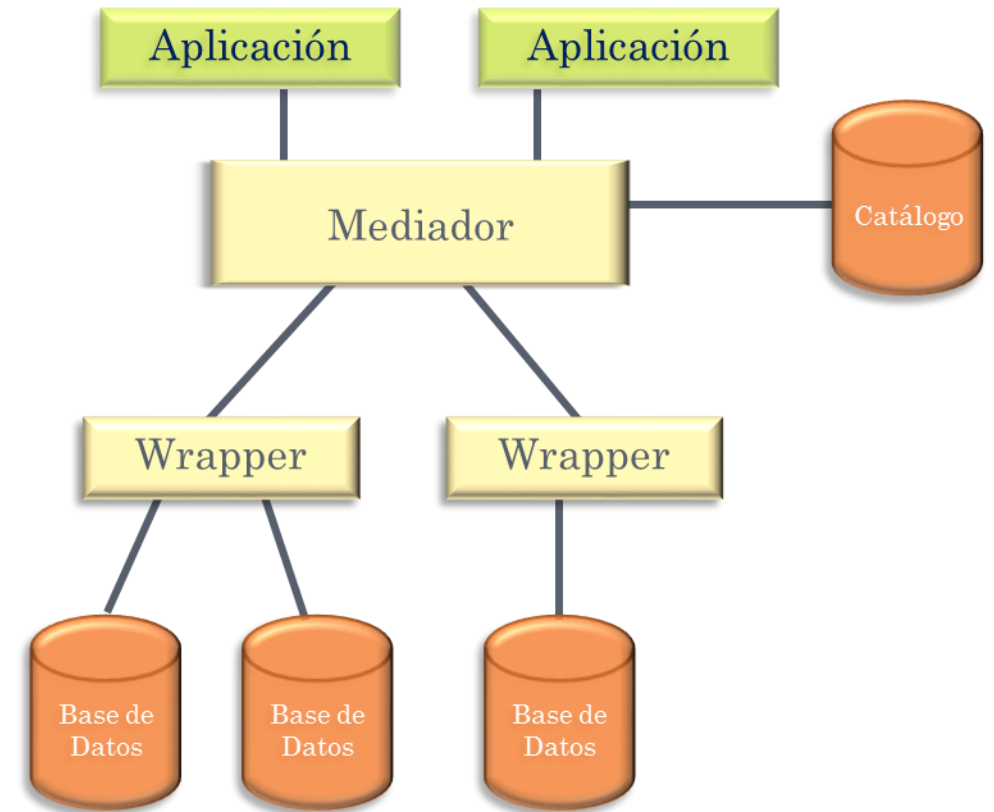
TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD EN LA WEB



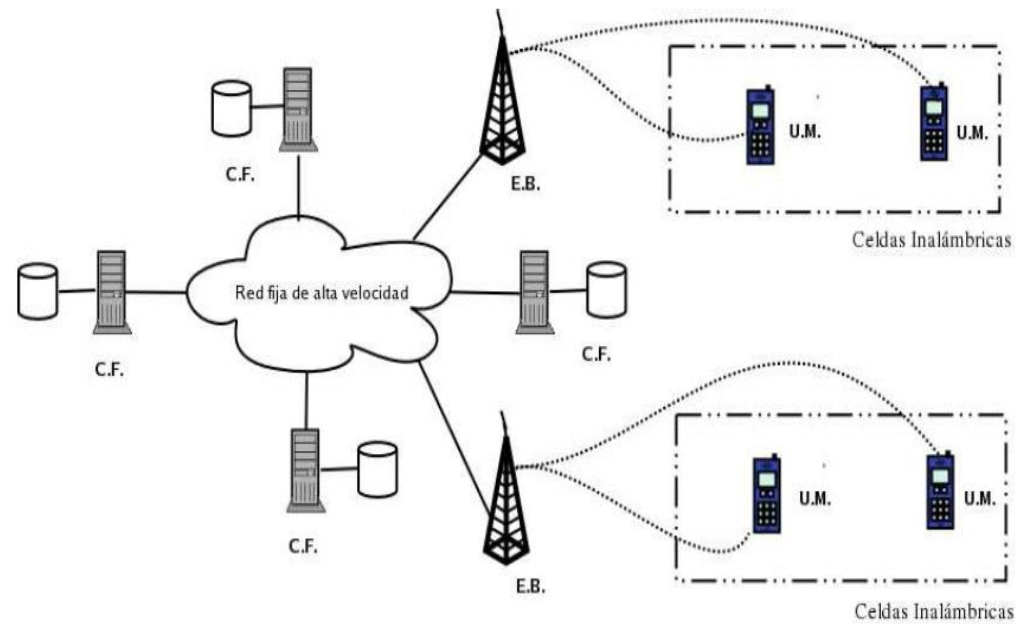
TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD FEDERADA



TIPOS DE BD

SEGÚN NUMERO DE SITIOS – BD MOVIL



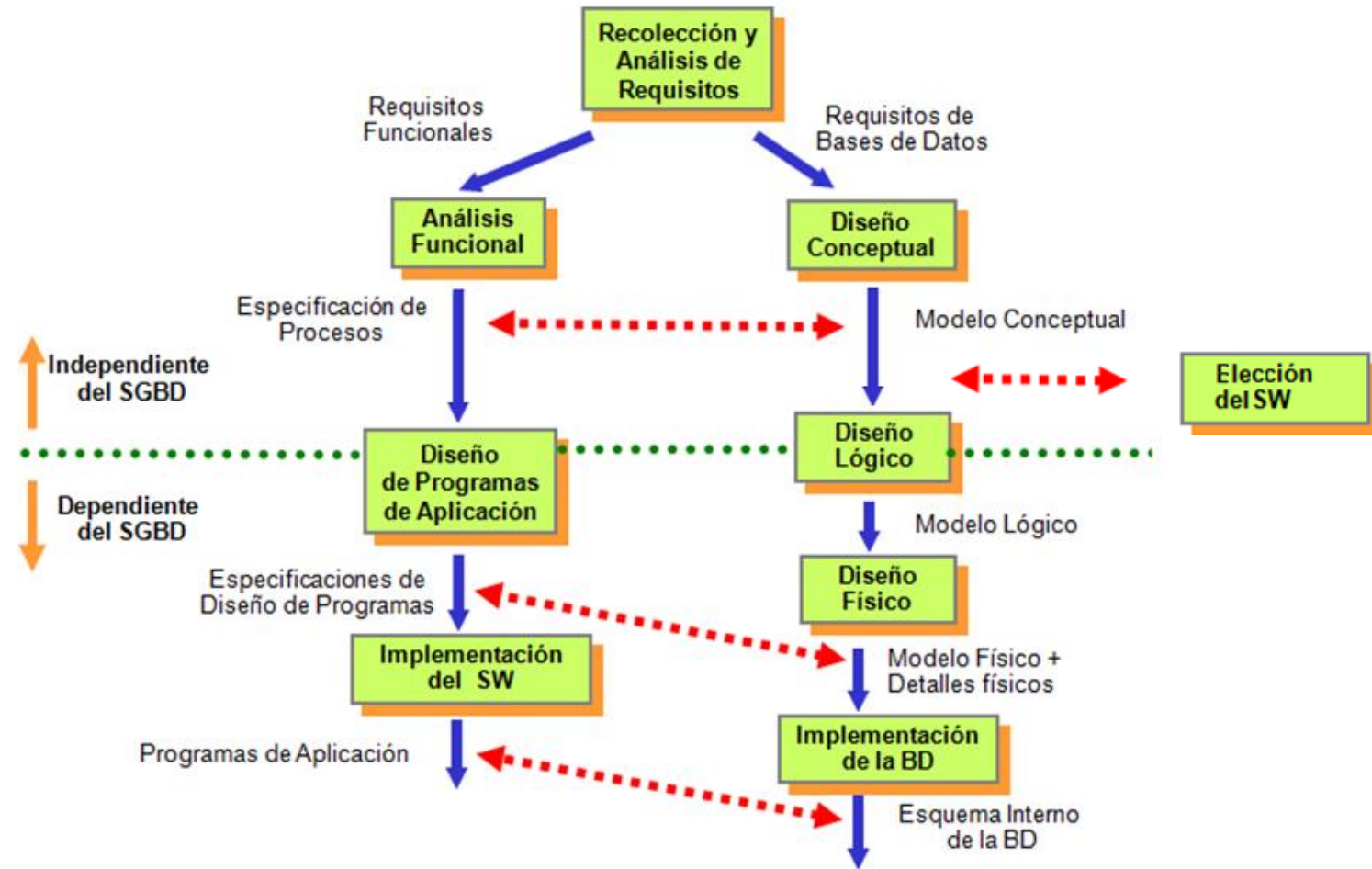
1.4 PROCESO DE DISEÑO DE BD



Departamento de Informática
Universidad Técnica Federico Santa María



ETAPAS PARA DISEÑO DE BD



ETAPAS PARA DISEÑO DE BD

ETAPA 1: RECOLECCION Y ANALISIS DE REQUISITOS

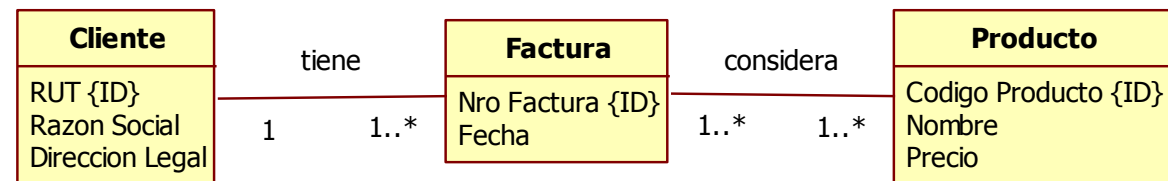
- **Objetivo:** identificar las necesidades de información de los usuarios (requisitos, requerimientos o vistas).

- **Pasos:**
 - Identificación de las áreas de aplicación y grupos de usuarios. Elección de los actores o participantes principales.
 - Análisis y estudio de la documentación existente en el sistema actual, incluye aplicaciones, manuales de políticas, reportes, diagramas organizacionales, etc.
 - Estudio del actual ambiente operativo y uso de la información. Incluye un análisis de los tipos de transacciones y sus frecuencias, y del flujo de información en el sistema.
 - Respuestas de cuestionarios obtenidas desde los potenciales usuarios, permiten identificar prioridades.
 - Formalización de requisitos.

ETAPAS PARA DISEÑO DE BD

ETAPA 2: DISEÑO CONCEPTUAL

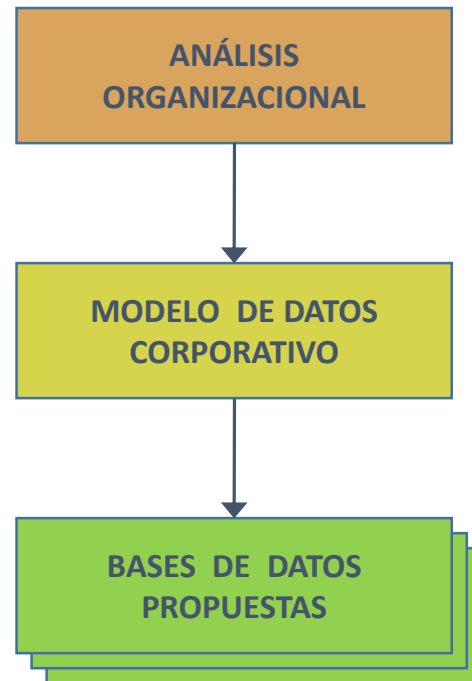
- Objetivo: construir un esquema conceptual que represente los datos necesarios para el sistema de información, que sea independiente del motor de datos a utilizar.
- El modelo conceptual sirve como:
 - Medio de comunicación entre usuarios y especialistas; por ende debe ser expresivo, simple, mínimo, formal, diagramático.
 - Mecanismo para validar entendimiento alcanzado del problema, por parte del especialista.
 - Descripción estable del contenido.



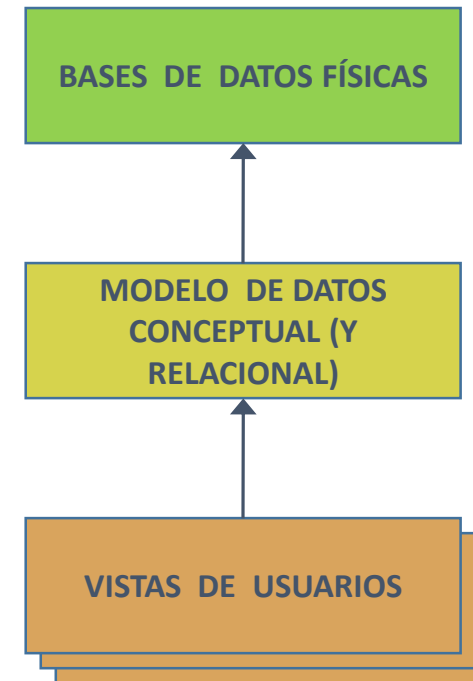
ETAPAS PARA DISEÑO DE BD

ETAPA 2: DISEÑO CONCEPTUAL

- Dos estrategias a seguir:



Top – Down



Bottom-Up

ETAPAS PARA DISEÑO DE BD

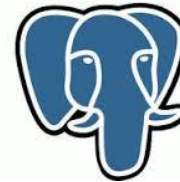
ETAPA 3: ELECCIÓN DEL SOFTWARE

- **Objetivo:** seleccionar aquel tipo de software (DBMS) que mejor se adecúe a las necesidades del sistema a construir.
- **Criterios a considerar:**
 - Costos: adquisición de hardware y software; operación y mantención; migración, ...
 - Requisitos del sistema: funcionales y no funcionales.
 - Estructuración de los datos

TERADATA

ORACLE®

PostgreSQL



ETAPAS PARA DISEÑO DE BD

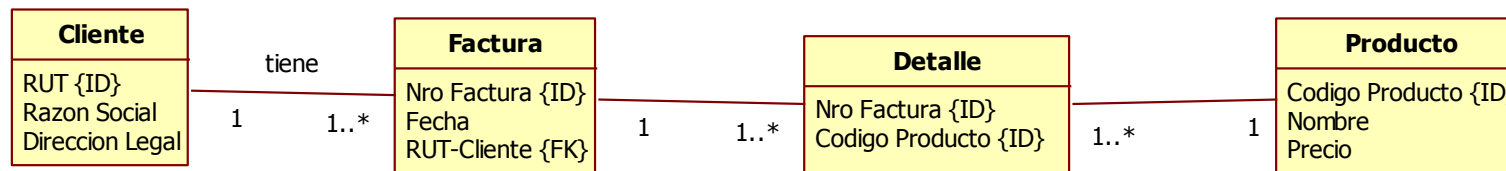
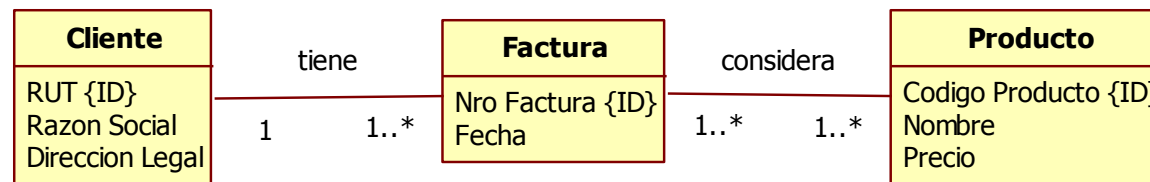
ETAPA 4: DISEÑO LOGICO

- **Objetivo:** generar un esquema basado en el modelo de datos soportado por el software escogido.
- **Pasos:**
 - Transformación independiente del sistema a un modelo relacional, orientado al objeto u otro.
 - Conversión de los esquemas a un software de bases de datos específico.

ETAPAS PARA DISEÑO DE BD

ETAPA 4: DISEÑO LOGICO

Modelo Conceptual



Modelo Relacional

ETAPAS PARA DISEÑO DE BD

ETAPA 5: DISEÑO FISICO

- **Objetivo:** escoger las estructuras de almacenamiento y métodos de acceso además de la ubicación de los archivos de bases de datos, para obtener un buen rendimiento de las distintas aplicaciones que interactúan con la base de datos.
- **Criterios a considerar:**
 - Transformación independiente del sistema a un modelo relacional, orientado al objeto u otro.
 - Tiempo de Respuesta: es el tiempo que transcurre desde el ingreso de la transacción hasta el recibo de su respuesta
 - Rendimiento del Sistema: número promedio de transacciones que pueden ser procesadas por minuto
 - Utilización del espacio en disco: cantidad de memoria ocupada por los archivos e índices

ETAPAS PARA DISEÑO DE BD

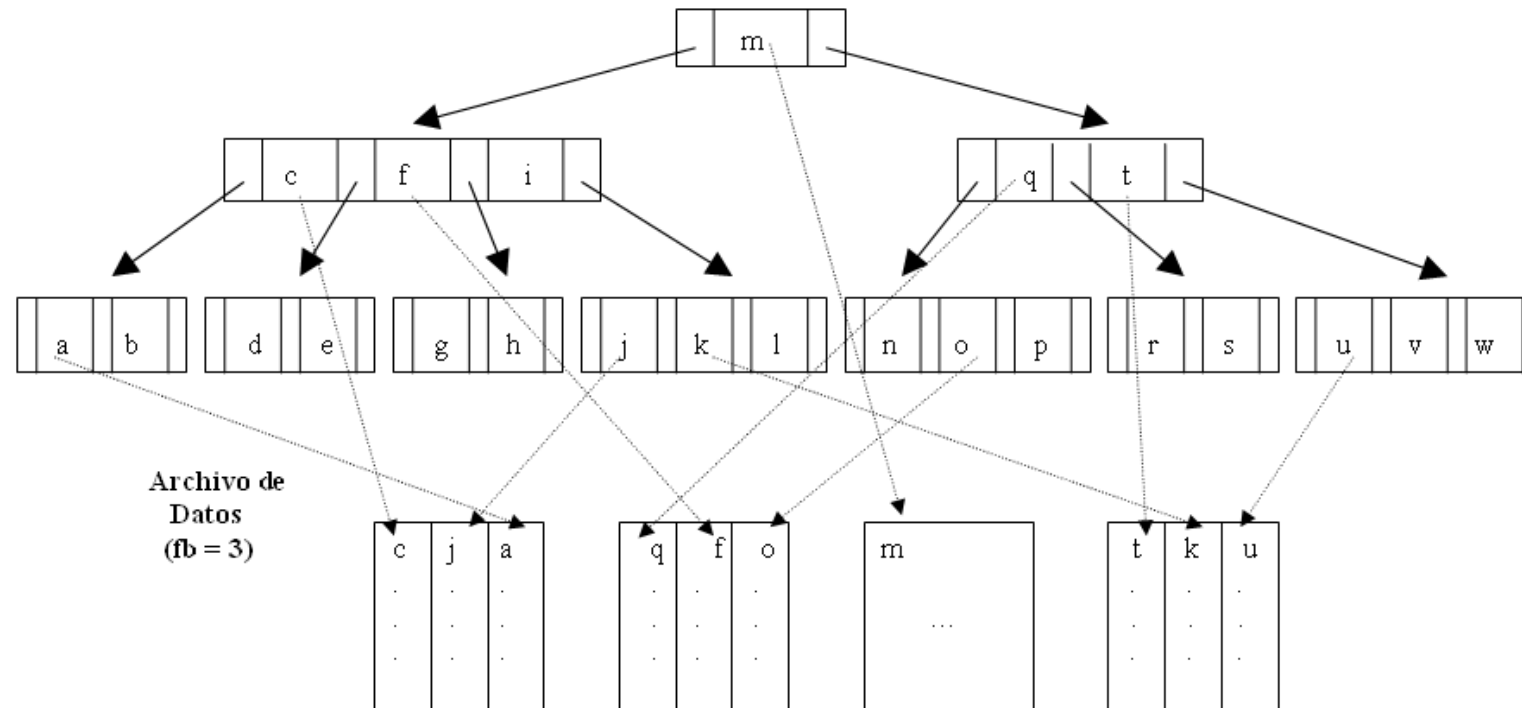
ETAPA 5: DISEÑO FISICO

■ Herramientas:

- Estructuras de almacenamiento:
 - Secuenciales: desordenados, ordenados
 - Directo: *hashing* estático, o con expansión dinámica
 - De tipo árbol: Arbol B.
- Índices
 - Dinámicos: hashing con expansión dinámica, de tipo Arbol B o B+
 - Bitmap

ETAPAS PARA DISEÑO DE BD

ETAPA 5: DISEÑO FISICO



ETAPAS PARA DISEÑO DE BD

ETAPA 6: IMPLEMENTACIÓN DE LA BASE DE DATOS

- **Objetivo:** codificación de sentencias para la definición y la manipulación de la base de datos, para crear los archivos y su poblamiento.

```
CREATE TABLE alumno
```

```
SELECT rut, nombre FROM alumno;
```

```
SELECT * FROM alumno WHERE carrera = 'INF';
```