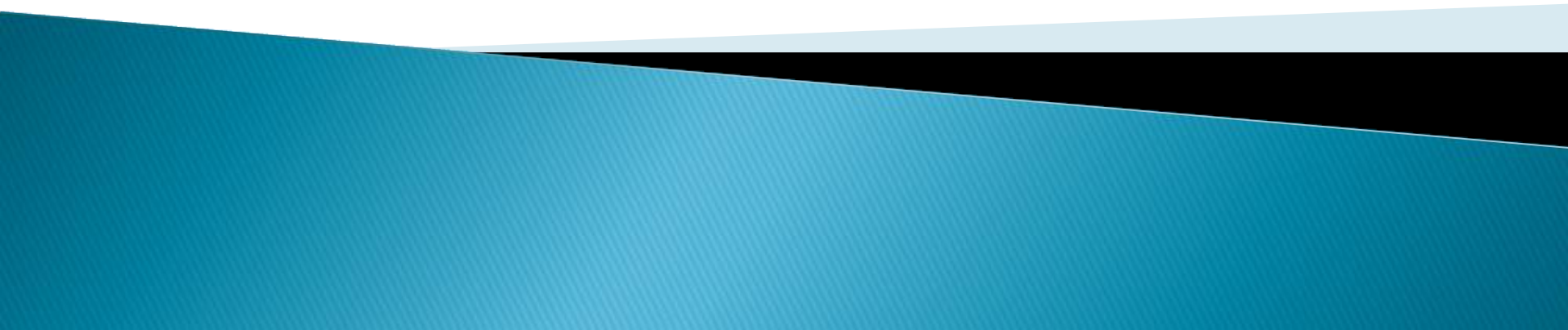



Capítulo 1

Introducción a las bases de datos



Temas

- 1.1.– Conceptos de bases de datos
 - 1.2.– Sistema de base de datos
 - 1.3.– Historia de la administración de datos
 - 1.4.– Aplicaciones de los sistemas de BD
 - 1.5.– Propósito de los sistemas de BD
 - 1.6.– Usuarios de la BD
 - 1.7.– Ciclo de vida de una BD
- 

1.1.– Conceptos de BD

► Dato (ANSI).

- Es una representación de hechos, conceptos o instrucciones en una forma normalizada conveniente para la comunicación, interpretación o procesamiento por humanos por medios automáticos.
- Cualquier representación de caracteres o cantidades analógicas para las cuales el significado es o puede ser asignado.
Generalmente, se ejecutan acciones sobre datos o ítems de datos para proporcionar alguna información acerca de una entidad.
- Hechos conocidos que pueden ser almacenados y que tienen un significado implícito.

Definiciones

- ▶ Base de datos.– Colección de datos relacionados. Una base de datos tiene la siguientes propiedades implícitas:
 - Representa algún aspecto del mundo real (miniuniverso)
 - Es una colección lógicamente coherente de datos con significado inherente
 - Está diseñada, construida y poblada con datos para un fin específico.

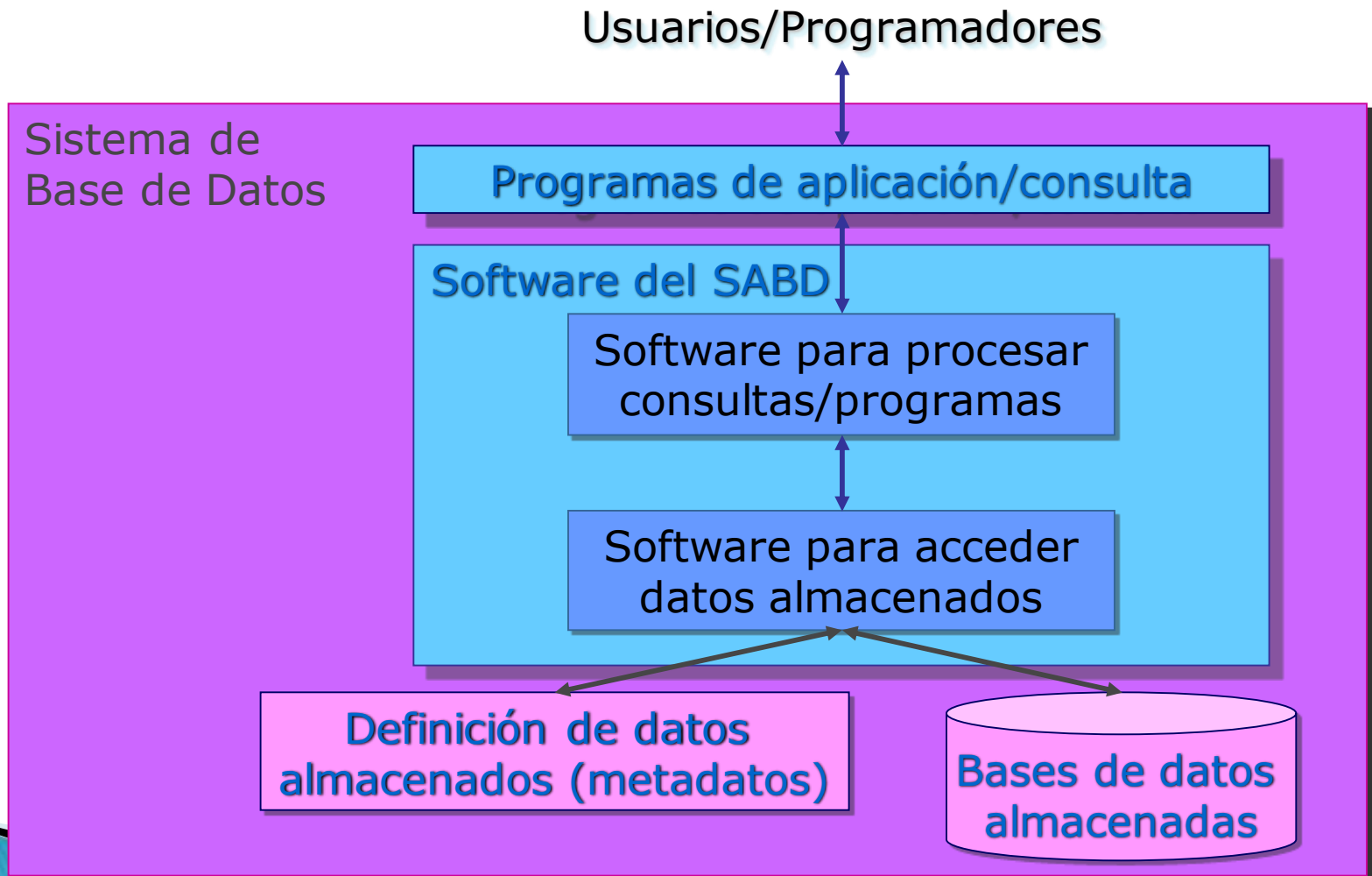


1.2 Sistema de BD

- ▶ Un Sistema Administrador de Bases de Datos (SABD) es una colección de programas de software que habilita a los usuarios a crear y mantener una base de datos. Es un sistema de software de uso general que facilita el proceso de definir, construir y manipular bases de datos para varias aplicaciones.
- ▶ Un Sistema de base de datos es el software que manipula a una base de datos.

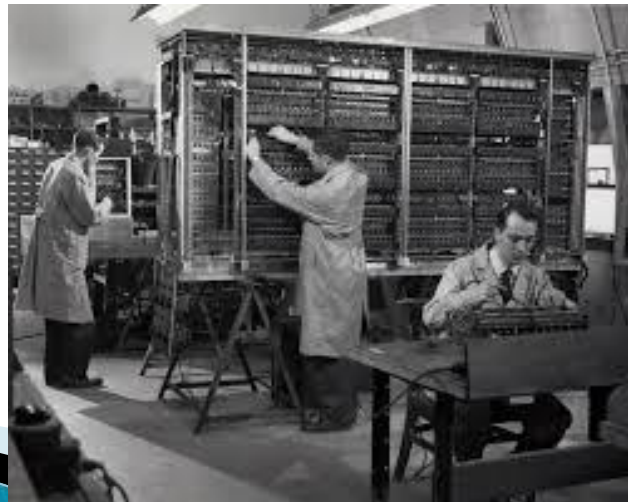


Sistema de bases de datos

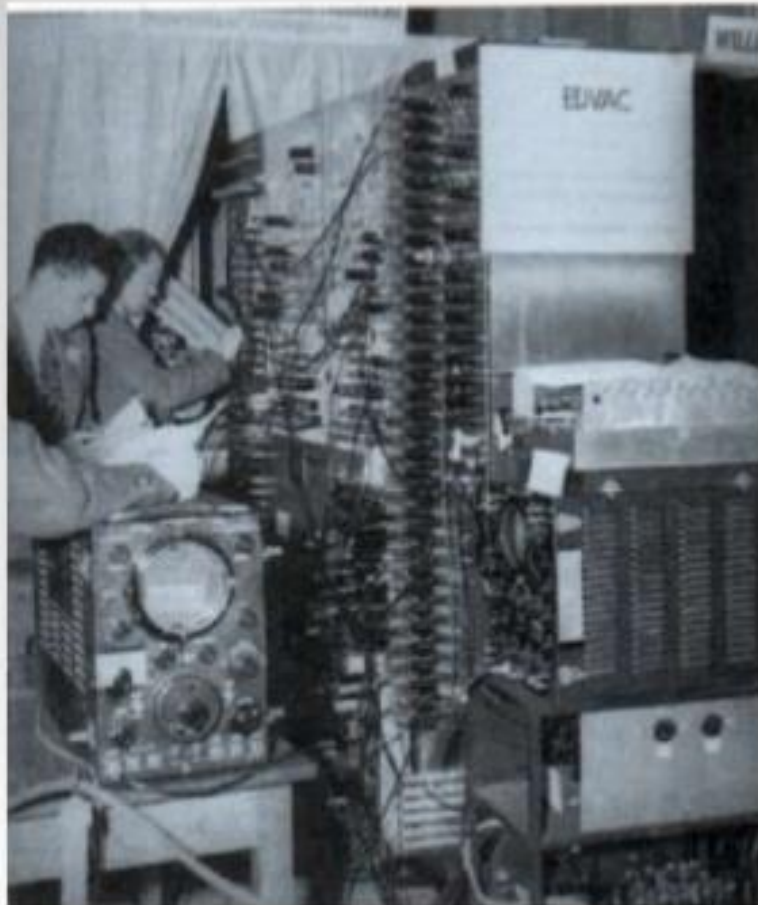


1.3.– Historia de la administración de datos

- ▶ Los datos no eran almacenados en discos
- ▶ El programador definía las estructuras lógicas y las estructuras físicas de los datos (estructuras de almacenamiento, métodos de acceso, modos de I/O, etc.)
- ▶ Un conjunto de datos por programa. Alta redundancia de datos.



LAS PRIMERAS COMPUTADORAS



1949, EDVAC (Electronic Discrete Variable Automatic Computer), por John Von Neumann, quien se unió al equipo de trabajo de Eckert y Mauchly, usaba 4000 bulbos de vacío, ésta máquina contaba ya con memoria por lo que la idea fundamental de Neumann fue permitir que en la memoria coexistieran datos e instrucciones y que entonces la computadora pueda ser programada en un lenguaje.

Problemas

- ▶ No había persistencia.
 - Todos los datos son transitorios y desaparecen cuando el programa termina
- ▶ La memoria de acceso aleatorio (RAM) era cara y limitada
 - Todos los datos no cabían en la memoria disponible
- ▶ Baja productividad del programador
 - El programador tenía que hacer un gran trabajo tedioso

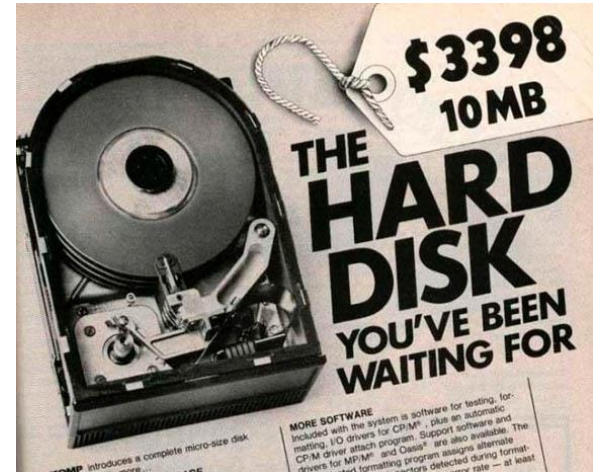


Tarjetas perforadas

- Las tarjetas perforadas están hechas de papel manila, la cual contiene información que una computadora podría leer con un lector de tarjetas perforadas, es decir, la información puede ingresar a la máquina. Estas están hechas con orificios que representan números y letras.

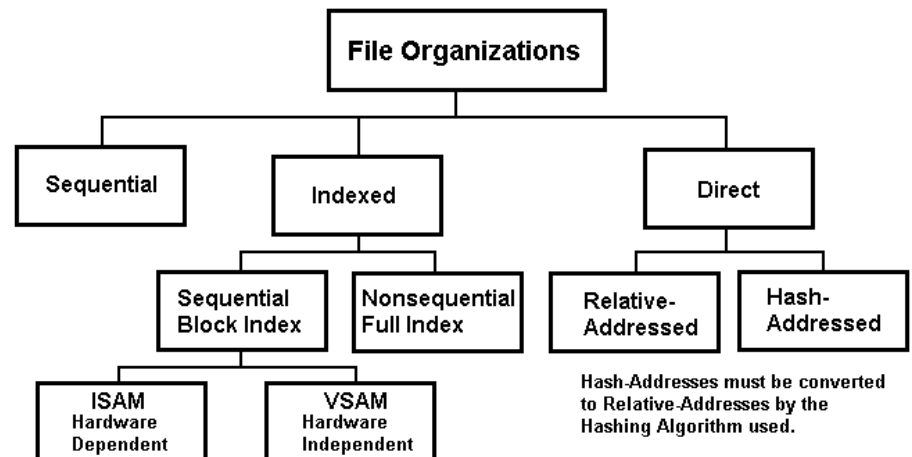
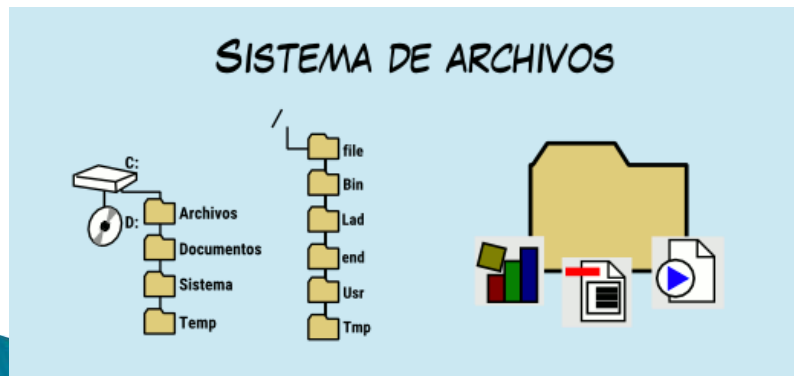


Siguiente



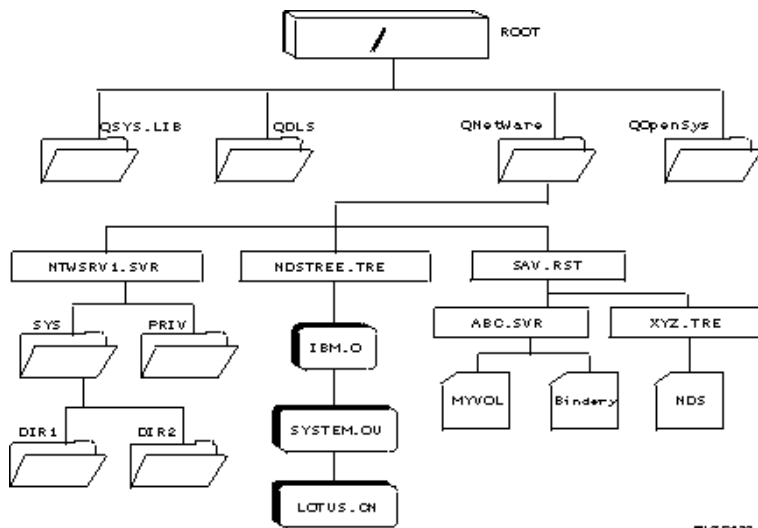
Procesamiento de archivos

- ▶ Los datos eran almacenados en archivos con interfaces entre programas y archivos
- ▶ Se desarrollaron varios métodos de acceso (ej. secuencial, indexado, aleatorio)
- ▶ Un archivo correspondía a uno o varios programas



Funciones del Sistema de Archivos

- ▶ Mapeo entre los archivos lógicos y los archivos físicos
 - Archivos lógicos: un archivo visto por usuarios y programas.
 - Los datos pueden ser vistos como una colección de bytes o como una colección de registros (colección de bytes con una estructura particular)
 - Los programas manipulan los archivos lógicos
 - Archivos físicos: un archivo como existe actualmente en el dispositivo de almacenamiento
 - Los datos son usualmente vistos como una colección de bytes localizados en una dirección física en el dispositivo
 - El sistema operativo manipula los archivos físicos.



RV30070-2

	FAT16	FAT32	HFS+	ext3	NTFS 5.0	NTFS 6.0	ext4
Año de creación	1984	1996	1998	1999	2001	2006	2006
Empresa	Microsoft	Microsoft	Apple Computer	Stephen Tweedie	Microsoft	Microsoft	Varios
Sistema operativo inicial	MS DOS 3	Windows 95	Mac OS 8.1	Linux Kernel 2.4.15	Windows XP	Windows Vista	Linux Kernel 2.6.19
Tamaño máximo de nombre de fichero	8+3	8+3	255 caracteres UTF-16	255 bytes	255 caracteres	255 caracteres	256 bytes
Tamaño máximo de fichero	2/4 GB	4 GB	8 EB	2 TB	16 EB	16 EB	16 TB
Tamaño máximo de partición	2/4 GB	2/16 TB	8 EB	32 TB	16 EB	16 EB	1 EB

1.4.– Aplicaciones de los sistemas de BD

- **Bancos:** gestión de cuentas, información del cliente, transacciones, préstamos,...
- **Líneas áreas:** reservaciones, horarios,...usan bases de datos distribuidas geográficamente.
- **Universidades:** matrículas, expedientes, evaluaciones...
- **Ventas:** clientes, productos, inventarios,...
- **Transacciones de tarjetas de crédito:** cargos, desgloses, generación de reportes...
- **Telecomunicaciones:** registros de llamadas, generación de pagos...
- **Recursos Humanos:** registros de empleados, salarios, impuestos, descuentos, ingresos...



1.5.– Propósitos de una base de datos (1)

- ▶ Esquema integrado
 - los usuarios tienen una vista uniforme de los datos
 - ven las “cosas” sólo como relaciones (tablas) en el modelo relacional
- ▶ Integridad declarativa y cumplimiento de consistencia
 - $24000 \leq \text{Salario} \leq 50000$
 - No hay empleados que puedan tener un salario más alto que su jefe
 - Los usuarios especifican y el sistema obliga
- ▶ Vistas individualizadas
 - Hay restricciones a ciertas relaciones
 - Reorganización de relaciones para ciertas clases de usuarios

Propósitos de una base de datos (2)

▶ Acceso declarativos

◦ Lenguaje de consultas – SQL

- Encontrar los nombres de todos los ingenieros eléctricos

```
SELECT nombre, apellido  
FROM empleado  
WHERE titulo = 'Ing. Eléctrico'
```

- Encuentra los nombres de todos los empleados que han trabajado en un proyecto como administradores por más de 12 meses

```
SELECT emp.nombre, emp.apellido  
FROM empleado emp, asignacion asg  
WHERE resp = "administrador"  
AND duracion > 12  
AND emp.NoEmp = asg.IdEmp
```

▶ Ejecución determinada por el sistema

Propósitos de una base de datos (3)

► Transacciones


- Ejecutan peticiones de usuario como unidades atómicas
- Puede contener una consulta o múltiples consultas
- Proporciona
 - Transparencia de concurrencia
 - Múltiples usuarios pueden acceder la base de datos, pero cada uno la base de datos como sus datos personales
 - Control de concurrencia
 - Transparencia de fallo
 - Incluso cuando un fallo del sistema ocurra, la consistencia de la base de datos no es violada
 - Bitácora y recuperación

Propósitos de una base de datos (4)


► Propiedades de las transacciones

- Atomicidad
 - propiedad de todo o nada
- Consistencia
 - Cada transacción es correcta y no viola la consistencia de la base de datos
- Aislamiento
 - transacciones concurrentes no interfieren con las otras
- Durabilidad
 - Una vez que la transacción termina su trabajo (compromete), se garantiza que sus efectos sean reflejados en la base de datos sin importar que pueda ocurrir

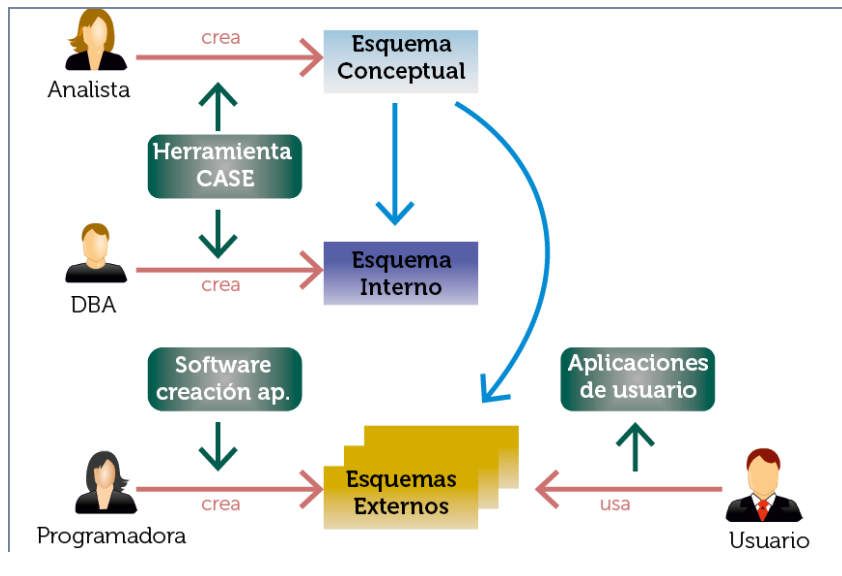
Ventajas de las bases de datos

- ▶ Los datos constituyen una ventaja organizacional \Rightarrow el control integrado
 - Reducción de redundancia
 - Eliminación de inconsistencia
 - Compartición
 - Estándares
 - Seguridad integrada
 - Integridad de datos
 - Mantenimiento simplificado
- 

Desventajas de las bases de datos

- ▶ Complejidad
 - ▶ Tamaño
 - ▶ Costos
 - SABD
 - Hardware
 - Actualizaciones
 - Migración
 - ▶ Prestaciones
 - ▶ Mayor impacto en los fallos
- 

1.6 Usuarios de la base de datos



► Usuario final

- usuario ingenuo o casual
- accede la base de datos a través de formas o por aplicaciones cliente (front-ends)
- alguien más sofisticado genera consultas ad hoc usando un DML

► Diseñador de la BD

- Lógico.– Identificar las entidades y atributos, así como las relaciones entre los datos y las restricciones que hay que aplicar.
- Físico.– Materializa el diseño lógico de la base de datos en un SADB.

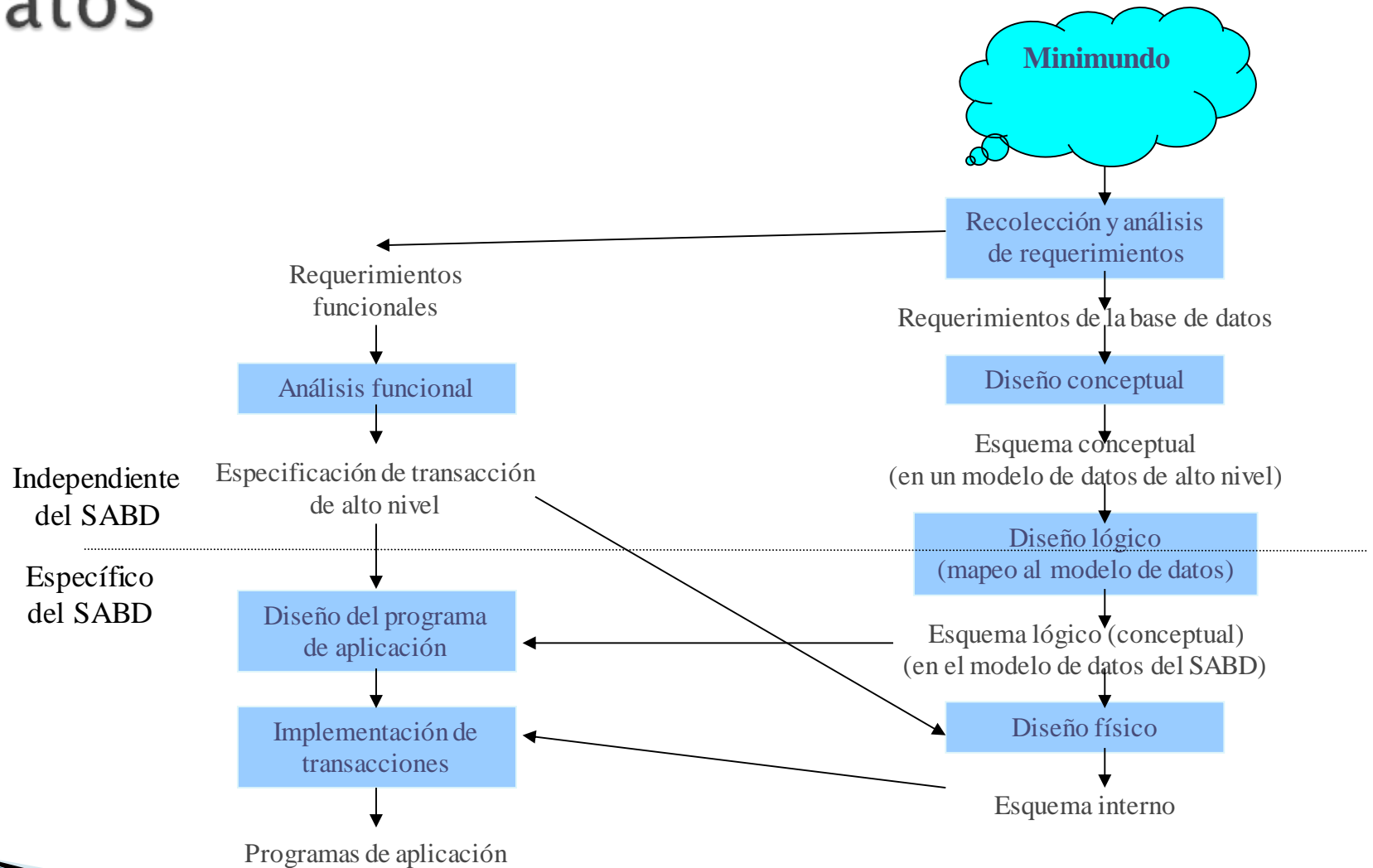
► Desarrollador/programador de aplicaciones

- diseña e implementa aplicaciones para acceder a base de datos (algunos pueden ser usados por usuarios finales)


► Administrador de la base de datos

- define las aplicaciones y las vistas de usuario
- monitorea y afina el rendimiento del SADB (define/modifica el esquema interno)
- carga y reformatea la base de datos
- es responsable de la seguridad y factibilidad

22 1.7 Ciclo de vida de una base de datos



Ciclo de vida de una base de datos(1)

- 1.- *Análisis de requerimientos*: se diseña la aplicación de base de datos para comprender que datos van a ser almacenados en la base de datos, que aplicaciones deben de construirse para accederlos y que operaciones serán las más frecuentes por parte de los usuarios.
 - 2.- *Diseño conceptual de la base de datos*: la información entregada por el análisis es usada para desarrollar una descripción de alto nivel de los datos para ser almacenados en la base de datos, así como las restricciones que aplicarán.
 - 3.- *Diseño lógico de la base de datos*: se debe elegir el SABD para implementar el diseño de la base de datos y convertir el diseño conceptual en el esquema (conceptual) de la base de datos con el modelo de datos elegido.
- 

Ciclo de vida de una base de datos(2)

- 4.- *Refinación del esquema*: el cuarto paso en el diseño de una base de datos es analizar la colección de relaciones en nuestro esquema de base de datos relacionales para identificar problemas potenciales y refinarlo (mediante la normalización).
- 5.- *Diseño físico de la base de datos*: aquí se deben considerar las cargas de trabajo típicas esperadas que debe soportar nuestra base de datos y realizar los refinamientos para satisfacer los criterios de rendimiento requeridos.
- 6.- *Diseño de seguridad*: se identifican los diferentes grupos de usuarios y los roles de estos, así como las partes de la base de datos a las que tendrán acceso.