**TEMA 1 – BANCO DE DADOS – FP DAM**

**CONTENIDO 1**

**Fichero o archivo:**conjunto de información relacionada, tratada como un todo y organizada de forma estructurada. Es una secuencia de dígitos binarios que organiza información relacionada con un mismo aspecto.

Los ficheros están formados por **registros lógicos** que contienen datos relativos a un mismo elemento u objeto (por ejemplo, los datos de usuarios de una plataforma educativa). A su vez, los registros están divididos en campos que contienen cada una de las informaciones elementales que forman un registro (por ejemplo, el nombre del usuario o su dirección de correo electrónico).

Hemos de resaltar que los datos están almacenados de tal forma que se puedan añadir, suprimir, actualizar o consultar individualmente en cualquier momento.

Como los ficheros suelen ser muy voluminosos, solo se pueden llevar a la memoria principal partes de ellos para poder procesarlos. La cantidad de información que es transferida entre el soporte en el que se almacena el fichero, y la memoria principal del ordenador, en una sola operación de lectura/grabación, recibe el nombre de **registro físico**o **bloque***.*

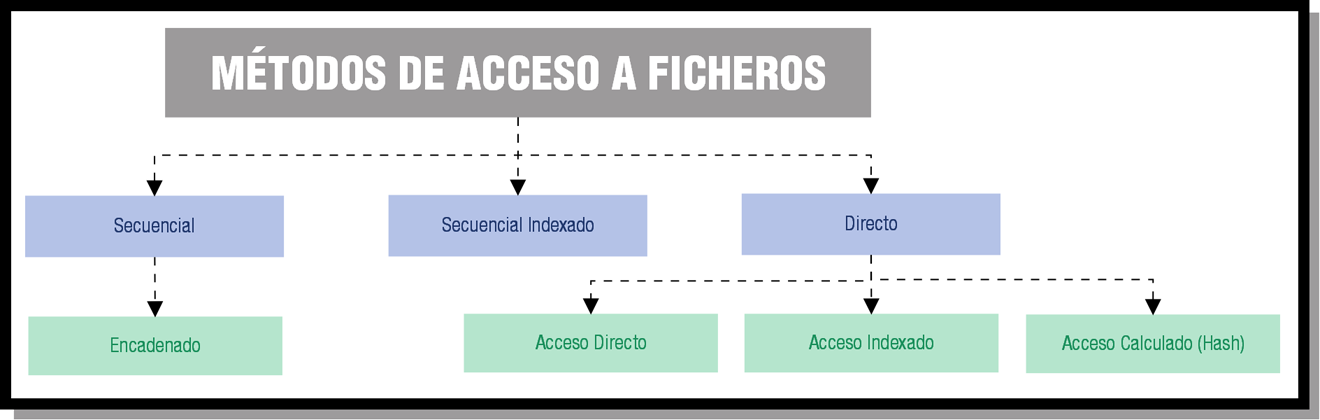
Normalmente en cada operación de lectura/grabación se transfieren varios registros del fichero, es decir un bloque suele contener varios registros lógicos. Al número de registros que entran en un bloque se le conoce con el nombre de **factor de blocaje**, y a esta operación de agrupar varios registros en un bloque se le llama **bloqueo de registros**.

TIPOS DE FICHEROS

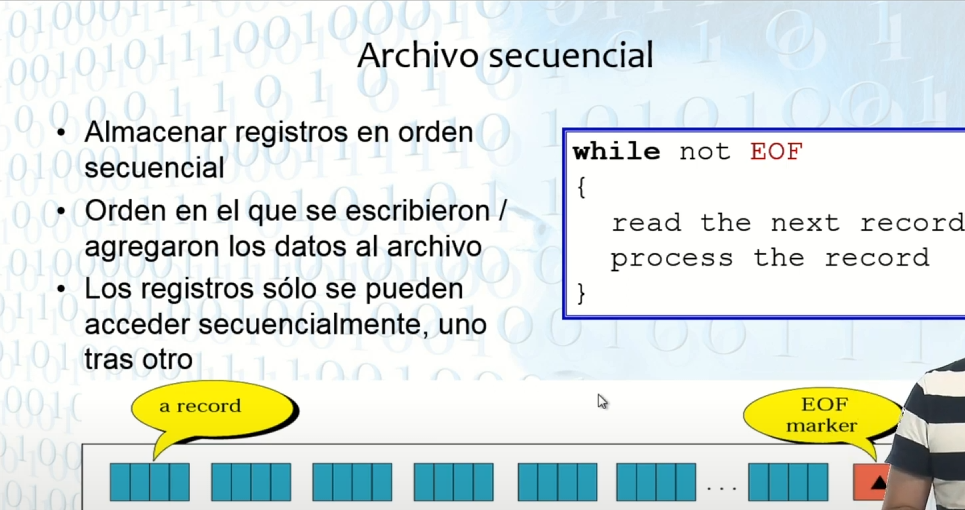
1. **Ficheros permanentes:**contienen información relevante para una aplicación. Es decir, los datos necesarios para el funcionamiento de ésta. Tienen un periodo de permanencia en el sistema amplio. Estos se subdividen en:
   * **Ficheros maestros:** contienen el estado actual de los datos que pueden modificarse desde la aplicación. Es la parte central de la aplicación, su núcleo. Podría ser un archivo con los datos de los usuarios de una plataforma educativa.
   * **Ficheros constantes:** son aquellos que incluyen datos fijos para la aplicación. No suelen ser modificados y se accede a ellos para realización de consultas. Podría ser un archivo con códigos postales.
   * **Ficheros históricos:**contienen datos que fueron considerados como actuales en un periodo o situación anterior. Se utilizan para la reconstrucción de situaciones. Podría ser un archivo con los usuarios que han sido dados de baja en la plataforma educativa.
2. **Ficheros temporales:**Se utilizan para almacenar información útil para una parte de la aplicación, no para toda ella. Son generados a partir de datos de ficheros permanentes. Tienen un corto periodo de existencia. Estos se subdividen en:
   * **Ficheros intermedios:**almacenan resultados de una aplicación que serán utilizados por otra.
   * **Ficheros de maniobras:** almacenan datos de una aplicación que no pueden ser mantenidos en memoria principal por falta de espacio.
   * **Ficheros de resultados:** almacenan datos que van a ser transferidos a un dispositivo de salida.

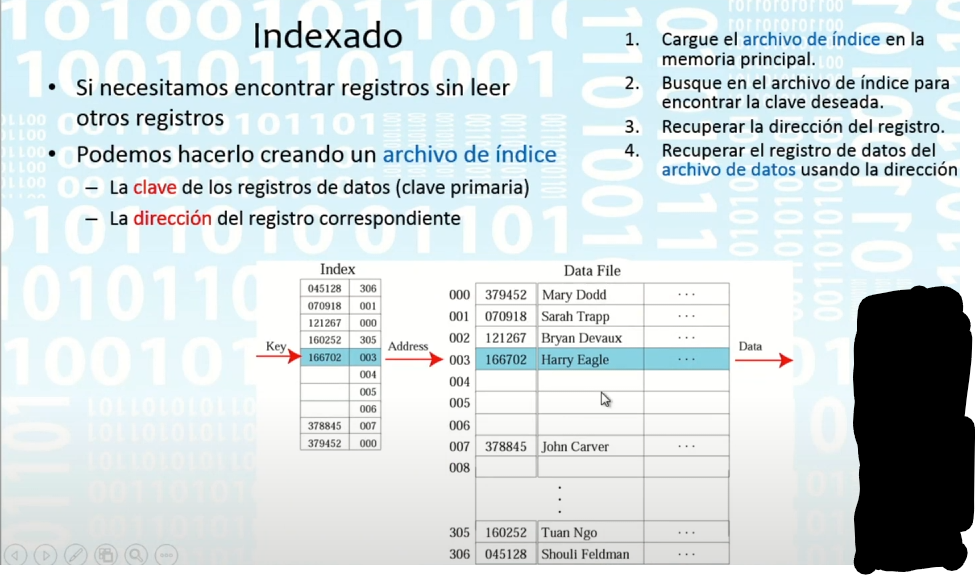
SUPORTE DE ACCESO DIRECTO

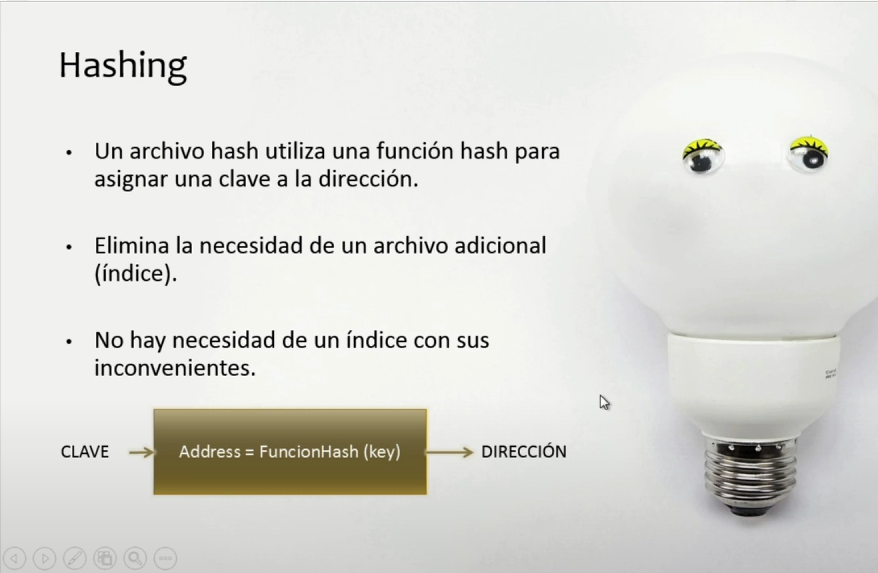
SUPORTE DE ACCESO SEQUENCIAL



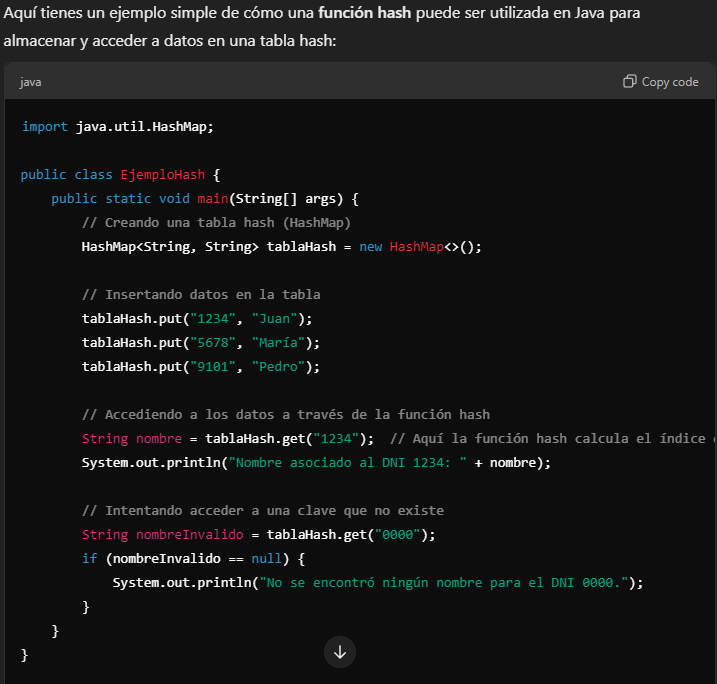
**Registro:** es básicamente una serie de campos que almacenan un dato relacionado a algo. EX.: lista de dni’s y correos electrónicos serian **campos** que tienen relación con el **registro personas**







Una **función hash** en la programación es un algoritmo que transforma una entrada (clave) en un número, generalmente un índice de un array o tabla. El valor generado por el hash se utiliza para localizar rápidamente los datos en estructuras como tablas hash. Se usa para **almacenamiento eficiente** y **búsqueda rápida**.



En función del uso que se vaya a dar al fichero, serán adecuados unos tipos u otros de organización. Mediante la utilización de **parámetros de referencia** , podremos determinar el uso de un fichero. Estos parámetros son:

1. **Capacidad o volumen:** es el espacio, en caracteres, que ocupa el fichero. La capacidad podrá calcularse multiplicando el número previsto de registros por la longitud media de cada registro.
2. **Actividad:** permite conocer la cantidad de consultas y modificaciones que se realizan en el fichero. Para poder especificar la actividad se deben tener en cuenta:
   * **Tasa de consulta o modificación:** que es el porcentaje de registros consultados o modificados en cada tratamiento del fichero, respecto al número total de registros contenidos en él.
   * **Frecuencia de consulta o modificación:** número de veces que se accede al fichero para hacer una consulta o modificación en un periodo de tiempo fijo.
3. **Volatilidad:**mide la cantidad de inserciones y borrados que se efectúan en un fichero. Para determinar la volatilidad es necesario conocer:
   * **Tasa de renovación:** es el tanto por ciento de registros renovados en cada tratamiento del fichero, respecto al número total de registros contenidos en él.
   * **Frecuencia de renovación:** es el número de veces que se accede al fichero para renovarlo en un periodo de tiempo fijo.
4. **Crecimiento:**es la variación de la capacidad del fichero y se mide con la tasa de crecimiento, que es el porcentaje de registros en que aumenta el fichero en cada tratamiento.

Los ficheros permiten organizar y memorizar conjuntos de datos del mismo tipo o naturaleza con una determinada estructura, siendo un medio para el almacenamiento de los datos o resultados de una aplicación específica. Pero si las aplicaciones, al ser diseñadas, deben depender directamente de sus ficheros o archivos, se pierde independencia y surgen serios inconvenientes: como información duplicada, incoherencia de datos, fallos de seguridad, etc.

Estos problemas debían ser solucionados, es cuando aparece el concepto de base de datos. Una base de datos permitirá reunir toda la información relacionada en un único sistema de almacenamiento, pudiendo cualquier aplicación utilizarla de manera independiente y ofreciendo una mejora en el tratamiento de la información, así como una evolución para el desarrollo de aplicaciones.

**Base de datos:** Es una colección de datos relacionados lógicamente entre sí, con una definición y descripción comunes y que están estructurados de una determinada manera. Es un conjunto estructurado de datos que representa entidades y sus interrelaciones, almacenados con la mínima redundancia y posibilitando el acceso a ellos eficientemente por parte de varias aplicaciones y usuarios.

La base de datos no sólo contiene los datos de la organización, también almacena una descripción de dichos datos. Esta descripción es lo que se denomina **metadatos**, se almacena en el **diccionario de datos o catálogo**y es lo que permite que exista **independencia de datos** lógica-física.

Una base de datos constará de los siguientes elementos:

* **Entidades:**objeto real o abstracto con características diferenciadoras de otros, del que se almacena información en la base de datos. Dicho de otra forma es algo acerca de lo cual se desea almacenar información. En una base de datos de una clínica veterinaria, posibles entidades podrían ser: ejemplar, doctor, consulta, etc.
* **Atributos:**son los datos que se almacenan de la entidad. Cualquier propiedad o característica de una entidad puede ser atributo. Continuando con nuestro ejemplo, podrían ser atributos: raza, color, nombre, número de identificación, etc.
* **Registros:** donde se almacena la información de cada entidad. Es un conjunto de atributos que contienen los datos que pertenecen a una misma repetición de entidad. En nuestro ejemplo, un registro podría ser: 2123056, Sultán, Podenco, Gris, 23/03/2009.
* **Campos:**donde se almacenan los atributos de cada registro. Teniendo en cuenta el ejemplo anterior, un campo podría ser el valor Podenco.

**Sistemas Gestores de Base de Datos (SGBD)**

Para poder tratar la informacion contenida en las bases de datos se utilizan los sistemas gestores de bases de datos o SGBD, también llamados DBMS (DataBase Management System), que ofrecen un conjunto de programas que permiten acceder y gestionar dichos datos.

El objetivo fundamental de los SGBD es proporcionar eficiencia y seguridad a la hora de recuperar o insertar información en las bases de datos. Estos sistemas están diseñados para la manipulación de grandes bloques de información.

**Sistema Gestor de Base de Datos:** Conjunto coordinado de programas, procedimientos, lenguajes, etc., que suministra, tanto a los usuarios no informáticos, como a los analistas, programadores, o al administrador, los medios necesarios para describir y manipular los datos contenidos en la base de datos, manteniendo su integridad, confidencialidad y seguridad.

**Nivel Externo**: Representa las vistas personalizadas de los usuarios. Está orientado a cómo los datos son vistos por diferentes personas o aplicaciones.

**Nivel Conceptual**: Proporciona una vista lógica centralizada de la base de datos completa. Define los datos y las relaciones entre ellos.

**Nivel Interno**: Describe cómo los datos se almacenan físicamente en la base de datos.

**RESTRICIONES, DEFINICIONES Y MANIPULACION CON UN SGBD**

**LENGUAGES DE LA BASE DE DATOS. (DDL, DML, DCL < SQL)**

* **Función de descripción o definición:**Permite al diseñador de la base de datos crear las estructuras apropiadas para integrar adecuadamente los datos. Esta función es la que permite definir las tres estructuras de la base de datos: Estructura interna, Estructura conceptual y Estructura externa. (Estos conceptos se verán más adelante en el epígrafe sobre arquitectura del SGBD).

Esta función se realiza mediante el **lenguaje de descripción de datos**o **DDL**. Mediante ese lenguaje: se definen las estructuras de datos, se definen las relaciones entre los datos y se definen las reglas (restricciones) que han de cumplir los datos.

* **Función de manipulación:**permite a los usuarios de la base buscar, añadir, suprimir o modificar los datos de la misma, siempre de acuerdo con las especificaciones y las normas de seguridad dictadas por el administrador. Se llevará a cabo por medio de un **lenguaje de manipulación de datos** ( **DML**) que facilita los instrumentos necesarios para la realización de estas tareas.

También se encarga de definir **la vista externa** de todos los usuarios de la base de datos o vistas parciales que cada usuario tiene de los datos definidos con el DDL.

Por manipulación de datos entenderemos:

La recuperación de información almacenada en la base de datos, lo que se conoce como **consultas**.

La inserción de información nueva en la base de datos.

El borrado de información de la base de datos.

La modificación de información almacenada en la base de datos.

* **Función de control:** permite al administrador de la base de datos establecer mecanismos de protección de las diferentes visiones de los datos asociadas a cada usuario, proporcionando elementos de creación y modificación de dichos usuarios. Adicionalmente, incorpora sistemas para la creación de copias de seguridad, carga de ficheros, auditoría, protección de ataques, configuración del sistema, etc. El lenguaje que implementa esta función es el **lenguaje de control de datos**o **DCL.**

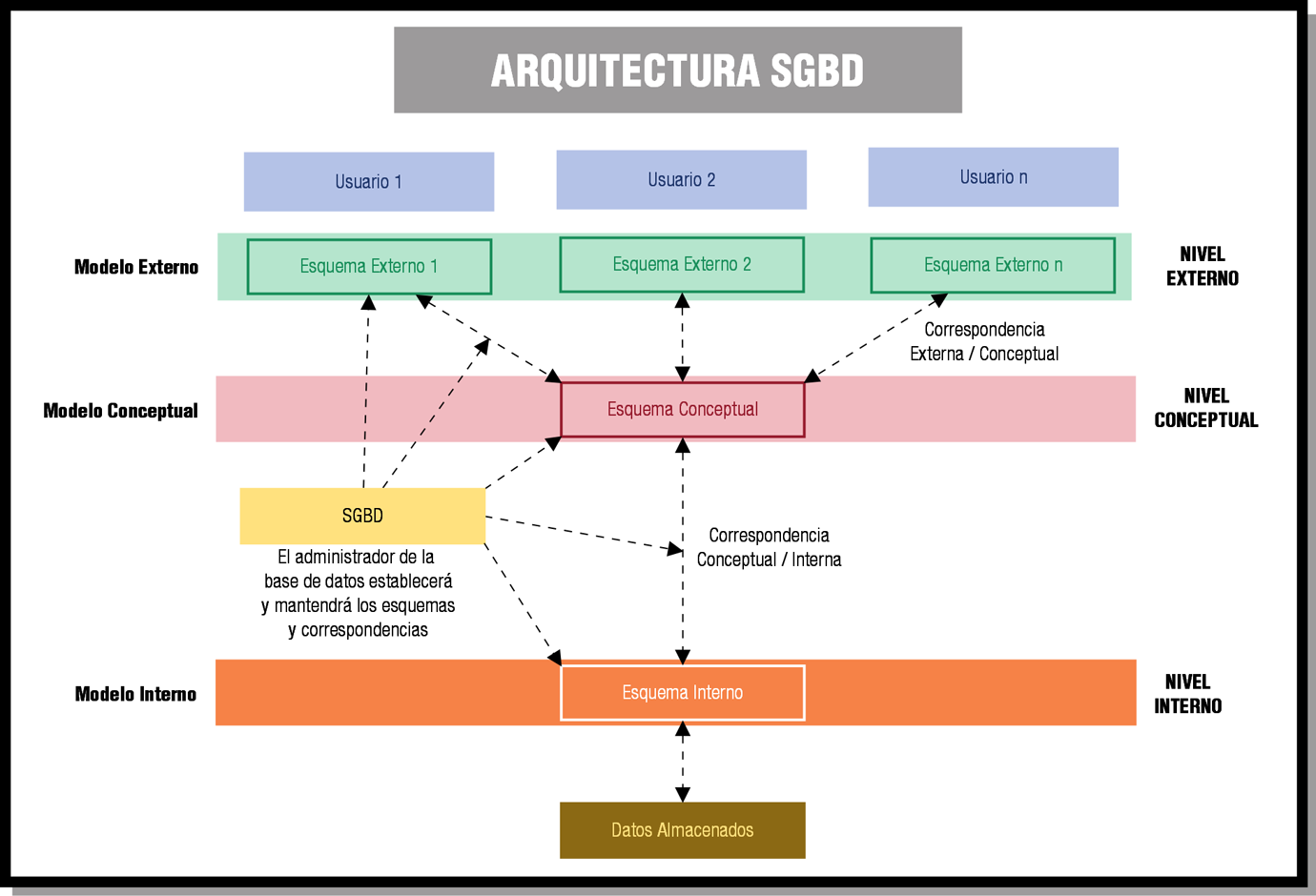
¿Y a través de qué lenguaje podremos desarrollar estas funciones sobre la base de datos? Lo haremos utilizando el **Lenguaje Estructurado de Consultas** ( **SQL**: Structured Query Language). Este lenguaje proporciona sentencias para realizar operaciones de DDL, DML y DCL.

**COMPOSICIÓN DE UN SGBD**

1. **El diccionario de datos:** Meta dados/documentación interna/estructura
2. **El gestor de la base de datos.**Es la parte de software encargada de garantizar el correcto, seguro, íntegro y eficiente acceso y almacenamiento de los datos. RESPONSABLE POR MANEJAR, DE NORMAL ES UNA PERSONA ESPECIALIZADA EN BD
3. **Usuarios de la base de datos.**
   * + 1. **administrador de la base de datos o Database Administrator ( DBA)** Tiene el control centralizado de la base de datos y es el responsable de su buen funcionamiento. Es el encargado de autorizar el acceso a la base de datos, de coordinar y vigilar su utilización y de adquirir los recursos software y hardware que sean necesarios.
       2. **usuarios de la base de datos**, que serán diferentes usuarios de la BD con diferentes necesidades sobre los datos, así como diferentes accesos y privilegios. Podemos establecer la siguiente clasificación:

* Diseñadores.
* Operadores y personal de mantenimiento.
* Analistas y programadores de aplicaciones.
* Usuarios finales: ocasionales, simples, avanzados y autónomos.

1. **Herramientas de la base de datos.**Son un conjunto de aplicaciones que permiten a los administradores la gestión de la base de datos, de los usuarios y permisos, generadores de formularios, informes, interfaces gráficas, generadores de aplicaciones, etc.

ARQUITETURA

Para una base de datos, sólo existirá un único esquema interno, un único esquema conceptual y podrían existir varios esquemas externos definidos para uno o varios usuarios.

Gracias a esta arquitectura se consigue la **independencia de datos** a dos niveles:

* **Independencia lógica:**Podemos modificar el esquema conceptual sin alterar los esquemas externos ni los programas de aplicación.
* **Independencia física:**Podemos modificar el esquema interno sin necesidad de modificar el conceptual o el externo. Es decir, se puede cambiar el sistema de almacenamiento, reorganizar los ficheros, añadir nuevos, etc., sin que esto afecte al resto de esquemas.
* **Propósito General:**pueden ser utilizados para el tratamiento de cualquier tipo de base de datos y aplicación.
* **Propósito Específico:**Cuando el rendimiento es fundamental, se puede diseñar y construir un software de propósito especial para una aplicación específica, y este sistema no sirve para otras aplicaciones. Muchos sistemas de reservas de líneas aéreas son de propósito especial y pertenecen a la categoría de **sistemas de procesamiento de transacciones en línea**, que deben atender un gran número de transacciones concurrentes sin imponer excesivos retrasos.