



# **- Diseño Físico -**

**Luis D. García**

**Abril de 2006**



## \* Diseño Físico \*

## Introducción

### AGENDA

- ▶ **Introducción**
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

El diseño de una base de datos está compuesto por tres etapas, el **Diseño Conceptual**, en el cual se descubren la semántica de los datos, definiendo las entidades, atributos y relaciones, el **Diseño Lógico**, en el que se especifica qué se guarda en la Base de Datos y el **Diseño Físico**, donde se especifica cómo se guardan los datos.

Para ello, el diseñador debe conocer muy bien toda la funcionalidad del SGBD concreto que se vaya a utilizar y también el sistema informático sobre el que éste va a trabajar. El diseño físico no es una etapa aislada, ya que algunas decisiones que se tomen durante su desarrollo pueden provocar una reestructuración del esquema lógico.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ **Metodología**
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Metodología

El diseño físico se divide de cuatro fases, cada una de ellas compuesta por una serie de pasos:

1. Traducir el esquema lógico global para el SGBD específico:

- Diseñar las relaciones base para el SGBD.
- Diseñar las reglas para el SGBD.

2. Diseñar la representación física:

- Analizar las transacciones.
- Escoger las organizaciones de ficheros.
- Escoger los índices secundarios.
- Estimar la necesidad de espacio en disco.



## \* Diseño Físico \*

## Metodología

### AGENDA

- ▶ Introducción
- ▶ **Metodología**
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

3. Diseñar los mecanismos de seguridad.

- Diseñar las vistas de los usuarios.
- Diseñar las reglas de acceso.

4. Monitorizar y afinar el sistema.

El objetivo de la etapa de Diseño Físico es producir una descripción de la implementación de la base de datos en memoria secundaria, incluyendo a su vez las estructuras de almacenamiento y los métodos de acceso que se utilizarán para conseguir un acceso eficiente a los datos.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ **Traducción del Esquema Lógico**
  - Diseñar las Relaciones Base
  - Diseñar las Reglas
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Traducción del Esquema Lógico

La primera fase del diseño lógico consiste en traducir el esquema lógico global en un esquema que se pueda implementar en el SGBD escogido. Para ello, es necesario conocer toda la funcionalidad que éste ofrece. El diseñador deberá saber:

- Si el sistema soporta la definición de claves primarias y foráneas.
- Si el sistema soporta la definición de datos requeridos (es decir, si se pueden definir atributos como no nulos).
- Si el sistema soporta la definición de dominios.
- Si el sistema soporta la definición de reglas.
- ¿Cómo se crean las relaciones base en el sistema?



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ **Traducción del Esquema Lógico**
  - **Diseñar las Relaciones Base**
  - Diseñar las Reglas
  
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Traducción del Esquema Lógico

### Diseñar las Relaciones Base para el SGBD

Las relaciones base se definen mediante el lenguaje de definición de datos del SGBD. Para ello, se utiliza el esquema lógico global y el diccionario de datos. El esquema lógico consta de un conjunto de relaciones junto a sus atributos, clave primaria y claves foráneas, si las tiene, así como sus reglas de integridad.

En el diccionario de datos por otro lado, se describen cada uno de los atributos definiendo a la vez, los tipo de datos, longitud y restricciones de dominio, el valor por defecto (opcional) y si admite o no valores nulos.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ **Traducción del Esquema Lógico**
  - **Diseñar las Relaciones Base**
  - Diseñar las Reglas
  
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Traducción del Esquema Lógico

### Diseñar las Relaciones Base para el SGBD

Ejemplo:

```
CREATE DOMAIN string AS VARCHAR(15);  
CREATE DOMAIN tipo AS VARCHAR(1)  
CHECK (VALUE IN ('O','A','E','P') );
```

```
CREATE TABLE empleado (  
    ci          INT4      NOT NULL,  
    nombre     STRING   NOT NULL,  
    tipo       TIPO     NOT NULL DEFAULT 'O',  
    sueldo     INT4      NOT NULL,  
    ci_jefe    INT4      NOT NULL,  
  
    PRIMARY KEY (ci),  
    FOREIGN KEY (ci_jefe) REFERENCES propietario  
        ON DELETE cascade );
```



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ **Traducción del Esquema Lógico**
  - Diseñar las Relaciones Base
  - **Diseñar las Reglas**
  
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Traducción del Esquema Lógico

### Diseñar las Reglas para el SGBD

Las actualizaciones que se realizan sobre las relaciones de la Base de Datos de una organización deben observar ciertas restricciones propias de la misma. Algunos SGBD proporcionan mecanismos que permiten definir estas restricciones y vigilan que no se violen.

### Ejemplo:

No es posible que un empleado tenga un sueldo menor que cero, por lo tanto, se debe definir entonces una restricción en la sentencia CREATE TABLE de la relación EMPLEADO.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ **Traducción del Esquema Lógico**
  - Diseñar las Relaciones Base
  - **Diseñar las Reglas**
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Traducción del Esquema Lógico

### Diseñar las Reglas para el SGBD

- Creación de la Restricción:

```
CREATE FUNCTION validar_sueldo() RETURNS trigger as $validar$  
BEGIN
```

```
IF NEW.sueldo < 0 THEN  
    RAISE EXCEPTION '% no puede tener un sueldo negativo',  
                    NEW.nombre;
```

```
END IF;  
RETURN NEW;  
END;
```

```
$validar$ LANGUAGE plpgsql;
```

```
CREATE TRIGGER validar BEFORE INSERT OR UPDATE  
ON empleado FOR EACH ROW EXECUTE PROCEDURE  
validar_sueldo();
```



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - Considerar Control de Redundancias
  - Estimar la necesidad de Espacio
  - Analizar las Transacciones
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Diseño de la Representación Física

Uno de los objetivos principales del diseño físico es almacenar los datos de modo eficiente. Para medir la eficiencia hay varios factores que se deben tener en cuenta:

- **Productividad de Transacciones:** es el número de transacciones que se quiere procesar en un intervalo de tiempo.

- **Tiempo de Respuesta:** es el tiempo que tarda en ejecutarse una transacción. Desde el punto de vista del usuario, este tiempo debería ser el mínimo posible.

- **Espacio en Disco:** es la cantidad de espacio en disco que hace falta para los ficheros de la base de datos. Normalmente, el diseñador querrá minimizar este espacio.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - Considerar Control de Redundancias
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Diseño de la Representación Física

En algunos casos todos estos factores no se pueden satisfacer a la vez. Por ejemplo, para conseguir un tiempo de respuesta mínimo, puede ser necesario aumentar la cantidad de datos almacenados, ocupando más espacio en disco. Por lo tanto, se deberán ir ajustando estos factores para conseguir un equilibrio razonable.

Para satisfacer de la manera más equitativa posible cada uno de estos factores nos basaremos en cinco actividades:

- Analizar las Transacciones
- Escoger la Organización de Ficheros
- Escoger los Índices Secundarios
- Considerar Control de Redundancias
- Estimar la necesidad de Espacio



## \* Diseño Físico \*

## Diseño de la Representación Física

### Analizar las Transacciones

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - **Analizar las Transacciones**
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - Considerar Control de Redundancias
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

Para realizar un buen diseño físico es necesario conocer las consultas y las transacciones que se van a ejecutar sobre la base de datos. Esto incluye tanto información cualitativa, como cuantitativa. Para cada transacción, hay que especificar:

- La frecuencia con que se va a ejecutar.
- Las relaciones y los atributos a los que accede la transacción, y el tipo de acceso: consulta, inserción, modificación o eliminación.
- Si es una consulta, los atributos involucrados en el *join* de dos o más relaciones.
- Las restricciones temporales impuestas sobre la transacción.



## \* Diseño Físico \*

## Diseño de la Representación Física

### Escoger la Organización de Ficheros

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - **Escoger la Organización de Ficheros**
    - Escoger los Índices Secundarios
    - Considerar Control de Redundancias
    - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

El objetivo de este paso es escoger la organización de ficheros óptima para cada relación. Por ejemplo, un **Fichero Desordenado** es una buena estructura cuando se va a cargar gran cantidad de datos en una relación en el momento de inicializarla, ya que la relación tiene pocas tuplas.

Por otra parte, los **Ficheros Dispersos** son apropiados cuando se accede a las tuplas a través de los valores exactos de alguno de sus campos (condición de igualdad en el WHERE).

Las organizaciones de ficheros a elegir en el diseño deben documentarse, justificando en cada caso la opción escogida.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - **Escoger los Índices Secundarios**
  - Considerar Control de Redundancias
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Diseño de la Representación Física

### Escoger los Índices Secundarios

Los *índices secundarios* permiten especificar caminos de acceso adicionales para las relaciones base. Por ejemplo, la relación EMPLEADO se puede haber almacenado en un fichero disperso a través del atributo '*ci*'. Si se accede a menudo a esta relación a través del atributo '*nombre*', se puede plantear la creación de un índice sobre dicho atributo para favorecer estos accesos.

A la hora de seleccionar los índices, se pueden seguir las siguientes indicaciones:

- Construir un índice sobre la clave primaria de cada relación base.
- No crear índices sobre relaciones pequeñas.
- Añadir un índice sobre los atributos que se utilizan para acceder con mucha frecuencia.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - **Escoger los Índices Secundarios**
  - Considerar Control de Redundancias
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

## Diseño de la Representación Física

### Escoger los Índices Secundarios

- Añadir un índice sobre las claves foráneas que se utilicen con frecuencia para hacer *joins*.
- Evitar los índices sobre atributos que se modifican a menudo.
- Evitar los índices sobre atributos poco selectivos (aquellos en los que la consulta selecciona una porción significativa de la relación).
- Evitar los índices sobre atributos formados por cadenas de caracteres largas.
- Los índices creados se deben documentar, explicando las razones de su elección.



## \* Diseño Físico \*

## Diseño de la Representación Física

### Considerar el Control de Redundancias

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - **Considerar el Control de Redundancias**
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

Existen casos en los cuales las relaciones ya normalizadas no proporcionan la máxima eficiencia para nuestra Base de Datos, por lo que es necesario volver atrás y “*desnormalizar*” estas relaciones.

La desnormalización de una relación puede ser una opción viable cuando las prestaciones que se obtienen no son las deseadas y la relación se actualiza con poca frecuencia, pero se consulta muy a menudo. Las redundancias que se pueden incluir al desnormalizar son de varios tipos:

- Se pueden introducir datos derivados (calculados a partir de otros datos).
- Se pueden duplicar atributos.
- Se pueden hacer **joins** de relaciones.



## \* Diseño Físico \*

## Diseño de la Representación Física

### Considerar el Control de Redundancias

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - **Considerar el Control de Redundancias**
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

No se pueden establecer de manera concreta una serie de reglas que determinen cuándo desnormalizar relaciones y cuando no, pero hay algunas situaciones muy comunes en donde puede considerarse esta posibilidad:

- Combinar Relaciones (1:1): cuando hay relaciones involucradas en relaciones de uno a uno, se accede a ellas de manera conjunta con frecuencia y casi no se les accede separadamente, por lo que se pueden combinar en una sola relación.

- Duplicar atributos no clave en Relaciones (1:N) para reducir los **joins**: para evitar operaciones de **join**, se pueden incluir atributos de la relación padre en la relación hijo de las relaciones de uno a muchos.



## \* Diseño Físico \*

## Diseño de la Representación Física

### Considerar el Control de Redundancias

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - **Considerar el Control de Redundancias**
  - Estimar la necesidad de Espacio
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

- Tablas de Referencia (*lookup*): estas tablas son listas de valores, cada uno de los cuales tiene un código. Por ejemplo, digamos que existe una tabla de referencia para los *tipos* de EMPLEADO, con las descripciones de estos tipos y un código asociado.

Ahora, si se coloca en la relación EMPLEADO una clave foránea a esta tabla para indicar el tipo de empleado es mucho más fácil validar los datos, además de que se ahorra espacio escribiendo sólo el código y no la descripción para cada empleado.

- Duplicar Claves Foráneas en Relaciones (1:N) para reducir los *joins*: para evitar operaciones de *join*, se pueden incluir claves foráneas de una relación en otra relación con la que se relaciona.



## \* Diseño Físico \*

## Diseño de la Representación Física

### Estimar la necesidad de Espacio

#### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ **Diseño de la Representación Física**
  - Analizar las Transacciones
  - Escoger la Organización de Ficheros
  - Escoger los Índices Secundarios
  - Considerar Control de Redundancias
  - **Estimar la necesidad de Espacio**
- ▶ Diseño de los Mecanismos de Seg.
- ▶ Monitorizar y afinar el Sistema

En caso de que se tenga que adquirir nuevo equipamiento informático, el diseñador debe estimar el espacio necesario en disco para la base de datos.

Esta estimación depende del SGBD que se vaya a utilizar y del hardware. En general, se debe estimar el número de tuplas de cada relación y su tamaño. También se debe estimar el factor de crecimiento de cada relación.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ **Diseño de los Mecanismos de Seg.**
- ▶ Monitorizar y afinar el Sistema

## Diseño de los Mecanismos de Seguridad

Durante el Diseño Lógico se habrán especificado los requerimientos en cuanto a seguridad que en el Diseño Físico se deben implementar. Para llevar a cabo esta implementación, el diseñador debe conocer las posibilidades que ofrece el SGBD que se vaya a utilizar.

- Diseñar las Vistas de los Usuarios: el objetivo de este paso es diseñar las vistas correspondientes a los esquemas lógicos locales.

Las vistas, además de preservar la seguridad, mejoran la independencia de datos, reducen la complejidad y permiten que los usuarios vean los datos en el formato deseado.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ **Diseño de los Mecanismos de Seg.**
- ▶ Monitorizar y afinar el Sistema

## Diseño de los Mecanismos de Seguridad

- Diseñar las Reglas de Acceso: el administrador debe asignar a cada usuario un identificador, que tendrá un contraseña asociada por motivos de seguridad.

Para cada usuario o grupo de usuarios se otorgarán permisos para realizar determinadas acciones sobre determinados objetos de la base de datos. Por ejemplo, los usuarios de un determinado grupo pueden tener permiso para consultar los datos de una relación base concreta y no tener permiso para actualizarlos.



## \* Diseño Físico \*

### AGENDA

- ▶ Introducción
- ▶ Metodología
- ▶ Traducción del Esquema Lógico
- ▶ Diseño de la Representación Física
- ▶ Diseño de los Mecanismos de Seg.
- ▶ **Monitorizar y afinar el Sistema**

## Monitorizar y Afinar el Sistema

Una vez implementado el esquema físico de la base de datos, el mismo debe ser puesto en marcha para observar sus prestaciones. Si éstas no son las deseadas, el esquema deberá cambiar para intentar satisfacerlas.

Una vez afinado el esquema, no permanecerá estático, ya que tendrá que ir cambiando conforme lo requieran los nuevos requisitos de los usuarios.

Los SGBD proporcionan herramientas para monitorizar el sistema mientras que el mismo se encuentra en funcionamiento.

**FIN**

