



Modelos de Datos



El modelo de datos se define como una colección de herramientas conceptuales para describir los datos, sus relaciones, la semántica asociada a ellos y sus restricciones de consistencia.

Un modelo de datos, representa la organización conceptual o lógica de los datos que soportan las operaciones que se ejecuten en un sistema determinado.

Cada uno de los procesos de un sistema requiere de un conjunto de datos que soporten las transacciones a efectuar, de tal modo que un modelo de datos debe responder a: ¿Cómo se organizan los datos, de forma tal que su uso sea eficiente?.

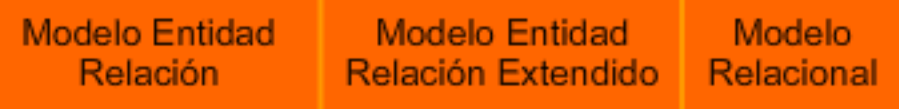
Un modelo establece una interpretación del mundo real al que se asocia, a través de una representación gráfica que relaciona los elementos que lo componen. Para responder a la acelerada transformación de la tecnología los modelos de datos han evolucionado y una clasificación que presenta los diferentes tipos de modelos de datos la constituyen los siguientes tres grupos:



Modelamiento de Datos



Modelos de Datos





Modelos de Datos

Modelos lógicos
basados en objetos

Modelos lógicos
basados en registros

Modelos físicos
de datos

Modelos lógicos basados en objetos

Estos modelos se usan para describir datos en el nivel conceptual y de visión. Su principal característica es la capacidad de estructuración ya que es bastante flexible y permiten que las restricciones de datos se realicen explícitamente. Algunos de los modelos más conocidos son:

- Modelo entidad-relación
- Orientado a objetos
- Modelo funcional de datos

ejempl 



Modelos de Datos

Modelos lógicos
basados en objetos

Modelos lógicos
basados en registros

Modelos físicos
de datos

Modelos lógicos basados en objetos

Estos modelos se usan para describir datos en el nivel conceptual y de visión. Su principal característica es la capacidad de estructuración ya que es bastante flexible y permiten que las restricciones de datos se realicen explícitamente. Algunos de los modelos más conocidos son:

- Modelo entidad-relación
- Orientado a objetos
- Modelo funcional de datos

ejempl 



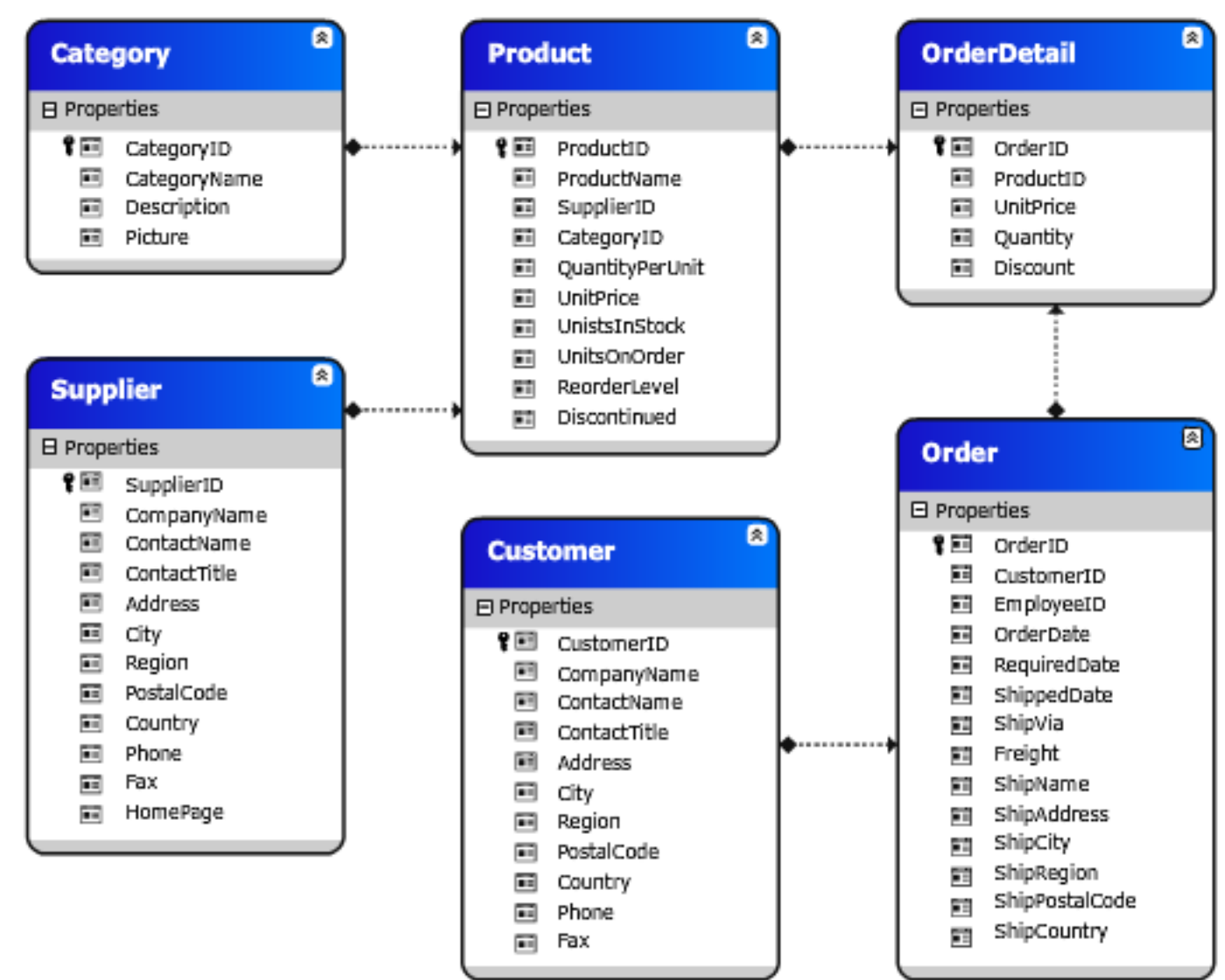
Modelos

M

Estos
visi
que
rea

- M
- O
- M

Modelos físicos
datos





Modelos de Datos

Modelos lógicos
basados en objetos

Modelos lógicos
basados en registros

Modelos físicos
de datos

Modelos lógicos basados en registros

Estos modelos se usan para describir datos en los modelos conceptual y físico. Permiten especificar la estructura lógica global de la Base de Datos y proporcionan una descripción a nivel más alto de la implementación.

Los tres modelos de datos más aceptados son los modelos: relacional, de red y jerárquico. El modelo relacional es el más utilizado. Siendo el modelo relacional el más utilizado.

- Modelo relacional
- Modelo de red
- Modelo jerárquico

ejempl 



Modelos

Modelos físicos
datos

M

Est
físic
y p

Los
red
mo

- Mo
- Mo
- Mo

CLIENTE

NOMBRE	CEDULA	CUENTA	CIUDAD
CANO	7.205.310	C-101	CALI
PEREZ	1.352.851	C-121	PASTO
TORO	9.874.115	C-203	BOGOTA
LOPEZ	9.705.004	C-302	BUGA
SERNA	2.454.498	C-109	TADO
VEGA	4.111.119	C-230	LIMA
CANO	7.205.310	C-309	CALI
PEREZ	1.352.851	C-209	PASTO

CUENTA

CUENTA	CUENTA
C-101	50.000
C-121	120.000
C-203	30.000
C-302	90.000
C-209	42.000
C-109	100.500
C-230	50.000



Modelo Relacional

Modelo de Red

Modelo Jerárquico





Modelos de Datos

Modelos lógicos
basados en objetos

Modelos lógicos
basados en registros

Modelos físicos
de datos

Modelos físicos de datos

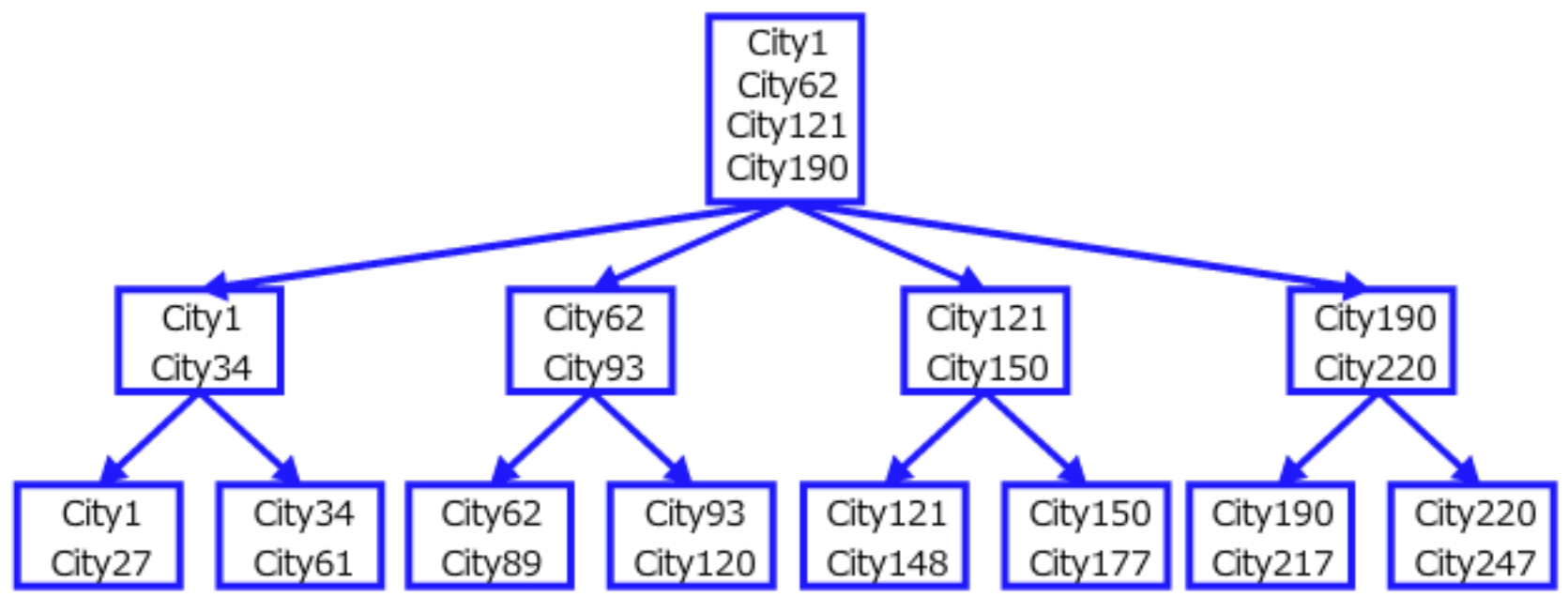
Estos modelos se usan para describir datos en el nivel más bajo. Hay muy pocos modelos físicos de datos en uso, siendo los más conocidos el modelo unificador y el de memoria de elementos. En estos casos las estructuras se implementan dentro del propio manejador de datos, por ejemplo: los Árboles binarios y las tablas Hash.

ejempl+



Modelo

físicos





Modelo Entidad Relación



El Modelo Entidad Relación (MER) es una metodología de diseño de Bases de Datos que consiste en representar a nivel conceptual los datos que soportan el funcionamiento de un sistema. El MER fue introducido originalmente por Peter Chen en 1976 y aunque ha sufrido variaciones en cuanto a los diagramas utilizados para representar sus elementos, su operación y utilidad siguen vigentes.

La base del MER está en identificar los elementos o entes importantes del sistema, los datos que componen cada uno de ellos y la interacción entre dichos elementos.

Los componentes básicos de un MER son: Entidades, Atributos y Relaciones.

- ENTIDAD:**
Se denomina entidad a todo ente (conceptual o físico) del cual se desea mantener información.
- ATRIBUTO:**
Conjunto de elementos de información que caracterizan a una entidad, identificándola, calificándola, cuantificándola, o declarando su estado.
- RELACIONES:**
Identifica la interacción que existe entre dos entidades.

ejempl+



Modelamiento de Datos



Modelo Entidad Relación



1 Identificar las Entidades

Inicialmente se requiere detectar los elementos u objetos de los cuales se requiere mantener información, a continuación identificar las operaciones que se ejecutan sobre estos elementos con lo cual es posible que aparezcan más entidades, pero solo deben quedar las que estén directamente relacionadas con el proceso a modelar. Los nombres de las entidades se deben escribir en mayúscula y singular.



Modelo Entidad Relación

Pasos para la
construcción
de un MER

Representación
Gráfica del Modelo

Matriz
Entidad – Entidad

Representación Gráfica del Modelo

Se han construido diferentes notaciones a través de los años, pero la esencia del modelo es la misma, a continuación se presentan algunas de estas para su comparación y análisis:

- **Notación de Peter Chen**
- **Notación CASE**

ejemplo 



Modelo Entidad Relación



Notación de Peter Chen

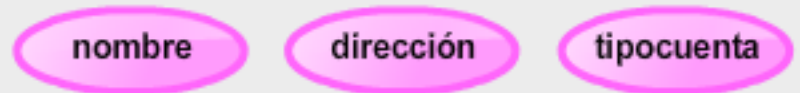
Notación de Case Method

Notación de Peter Chen:

Entidades:



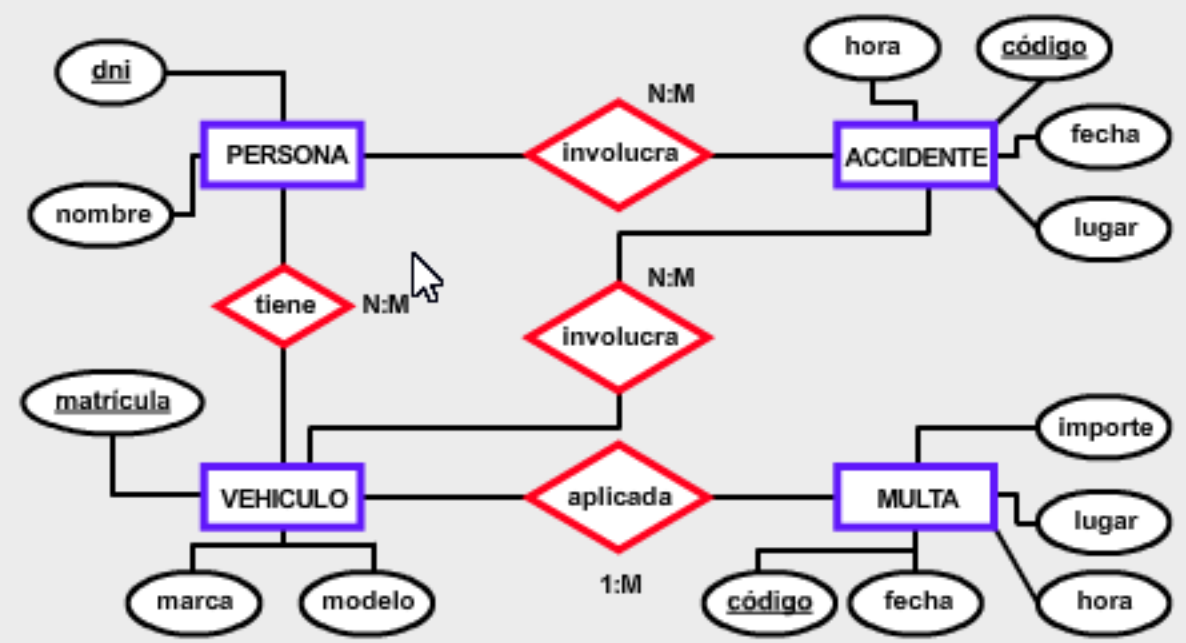
Atributos:



Relaciones:



Diagrama completo:



Modelo Entidad Relación con Notación de Peter Chen



Modelo Entidad Relación

Pasos para la construcción de un MER

Representación Gráfica del Modelo

Matriz Entidad – Entidad

Matriz Entidad - Entidad

Para validar el MER es posible utilizar una matriz Entidad x Entidad, donde se observan las entidades y sus relaciones. Para construirla se hace una matriz con las entidades en filas y columnas, en cada intersección se escribe el nombre de la relación en singular, indicando además la cardinalidad y la obligatoriedad usando pares ordenados (el primer término indica opcional -0- u obligatorio -1- y el segundo indica cardinalidad uno -1- ó varios -n-). En la casilla donde se crucen dos entidades sin relación alguna se coloca una "X".

La siguiente representa la matriz entidad-entidad de un modelo propuesto para controlar la realización de eventos en una institución.

	EVENTO	TIPO	PARTICIPANTE	INSTITUCION
EVENTO	X	Pertenece (1,1)	Se inscriben (1,n)	Los realiza (1,n)
TIPO	Pertenecen (0,n)	X	X	X
PARTICIPANTES	Se inscribe (1,n)	X	X	X
INSTITUCION	Realiza (0,n)	X	X	X





Modelo Relacional



Es un modelo formal propuesto por E.W. Codd en 1970, basado en los conceptos matemáticos de Relación y Teoría de Conjuntos. Este modelo esta soportado por la mayoría de los productos de bases de datos en el mercado actual y se representa por tablas en forma bidimensional.

Características

Transformación

ejempl





Modelo Relacional

Normalización

Reglas de Integridad
(Constraints)

Almacenamiento

Características

- Los renglones, líneas o filas, poseen datos acerca de una entidad.
- Las columnas contienen datos acerca de los atributos de la entidad.
- Cada posición en la tabla almacena un valor simple, atómico, de un atributo.
- Todos los valores almacenados en una columna son del mismo tipo, es decir, están definidos sobre el mismo dominio.
- Cada columna o atributo debe poseer un nombre único.
- El orden de las columnas no es relevante.
- El orden de las filas no es relevante.
- No pueden existir dos filas idénticas en la tabla.



Modelo Relacional

Normalización

Reglas de Integridad
(Constraints)

Almacenamiento

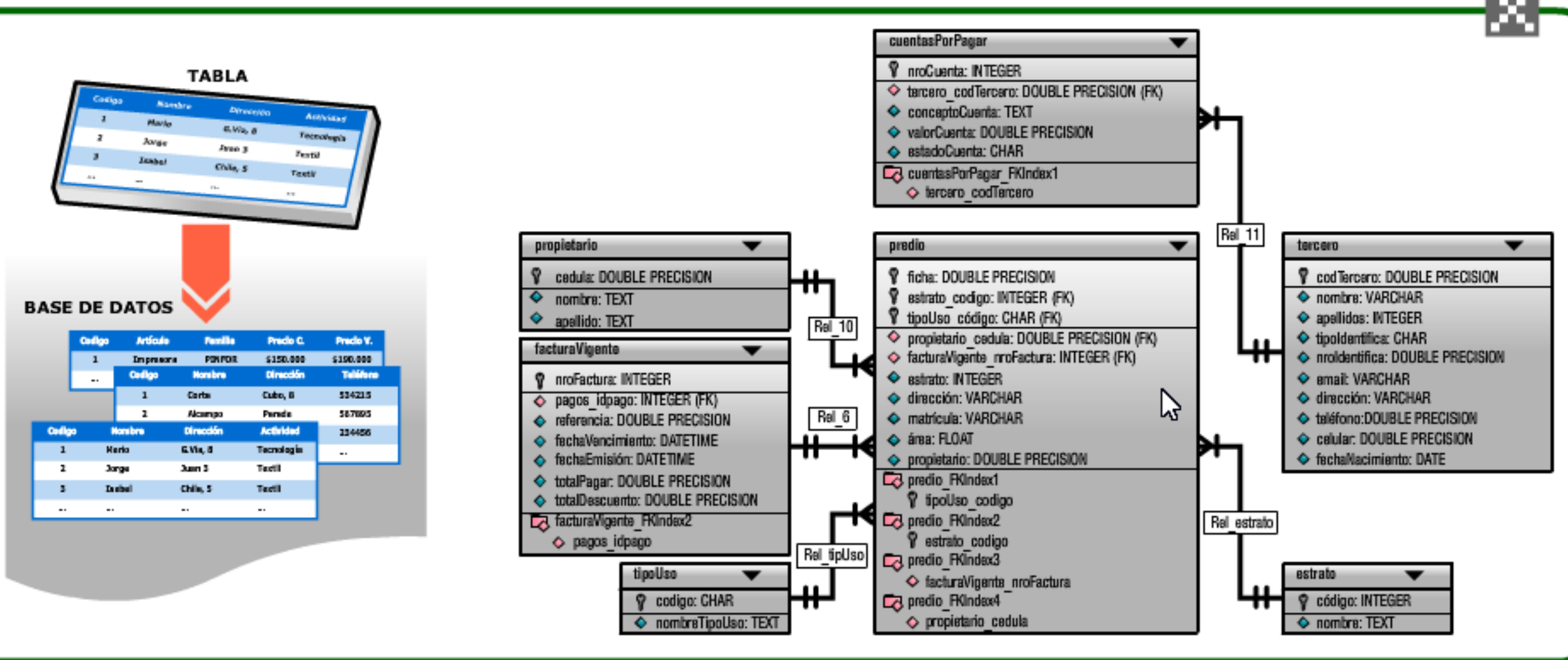
Transformación

Es posible transformar el Modelo Entidad Relación (MER) al Modelo Relacional, a partir de la utilización de una serie de reglas, algunas de estas son:

- Toda entidad se transforma en una tabla.
- Todo atributo se transforma columna o campo dentro de una tabla.
- El identificador único de la entidad se convierte en clave primaria de la tabla.
- Aplicar la preparación de los esquemas entidad-relación mediante la aplicación de las reglas que faciliten y garanticen la fiabilidad del proceso de transformación, las cuales se basan en la Normalización.
- Realizar la Transformación de relaciones, donde es necesario verificar la cardinalidad y la verificación de si la relación es obligatoria u opcional.



Modelo Relacional





Modelo Relacional



Normalización

La normalización es un proceso que permite garantizar un mínimo de redundancia y dependencias incoherentes sobre los datos, consiste en organizar los datos en tablas y establecer relaciones entre las tablas según reglas predefinidas.

La redundancia sobre los datos crea problemas de mantenimiento en la base de datos y ocasiona desperdicio de espacio en disco. Si hay que cambiar datos que existen en más de una tabla, se hace necesario hacer la actualización en todas sus ubicaciones. Además se generan problemas en las inserciones de nuevos datos, consultas y eliminaciones.

El proceso consiste en aplicar reglas de normalización sobre las tablas de una base de datos, cada regla se denomina "Forma Normal". Si una tabla cumple la primera regla, se dice que está en la "primera forma normal" y si cumple la regla N esta en forma normal N. Aunque son posibles otros niveles de normalización, la tercera forma normal se considera el máximo nivel necesario para la mayor parte de las aplicaciones.



Modelo Relacional

Normalización

Reglas de Integridad
(Constraints)

Almacenamiento

Reglas de Integridad (Constraints)

Aun cuando no se puede garantizar que los datos sean fidedignos por cuanto dependen en su gran mayoría de los usuarios, debemos apoyar la funcionalidad y cumplimiento de algunas reglas que puedan garantizar su Integridad.

En general, las condiciones que garantizan la integridad de los datos pueden ser de dos tipos:

1

Las restricciones de integridad de usuario

2

Las reglas de integridad de modelo



Modelo Relacional



Reglas de Integridad (Constraints)

Están asociados a las reglas del negocio, son condiciones específicas de una base de datos concreta que no son necesariamente relevantes en otra base de datos.

Están determinadas por la funcionalidad de los procesos sobre los que se va a manipular y por las validaciones mínimas que se esperan sobre los datos.



- 1 Las restricciones de integridad de usuario
- 2 Las reglas de integridad de modelo



Modelo Relacional



Reglas de Integridad (Constraints)



Son las condiciones generales de un modelo de datos. Estás son genéricas y en el caso del modelo relacional son:

- Unicidad de la clave Primaria: toda clave primaria que se defina no debe admitir valores repetidos.
- Integridad de entidad de la clave primaria: los atributos de la clave primaria no pueden tener valores nulos.

•Integridad Referencial: Definición de las políticas de inserción y eliminación para las filas cuyos que tengan campos referenciados.

•Integridad de Dominio: Los valores almacenados en un atributo deben ser del dominio declarado para dicho atributo (tipo de datos, rango de valores admitidos, opcional/obligatorio, cumplimiento de condiciones especiales).

1 Las restricciones de integridad de usuario

2 Las reglas de integridad de modelo

ejemplo



Modelo Relacional

Normalización

Reglas de Integridad
(Constraints)

Almacenamiento

Almacenamiento

Uno de los problemas comunes a los que se enfrenta el Administrador de la Base de datos es el relacionado al tema del espacio que debe asignar a las bases de datos en los discos duros para no tener dificultades en su almacenamiento, funcionamiento y rendimiento.

Complementario al diseño de la base de datos, es necesario realizar una estimación del tamaño que tendrá la base de datos cuando se encuentre en producción. Esta estimación puede ayudar a determinar la configuración de hardware que será necesaria para conseguir el rendimiento que se requiere asegurando el espacio en disco necesario para el almacenamiento de los datos y demás objetos asociados a las bases de datos.

Esta estimación también ayuda a determinar si el diseño requiere ajustes, ya que si el espacio requerido es muy grande, se podría aplicar mayor nivel de normalización, o en caso contrario si el tamaño calculado es inferior al esperado, se puede reducir la normalización para la optimización de tiempos de ejecución de las consultas.

Para comenzar se puede realizar una estimación del tamaño de cada tabla por separado y sumar los valores obtenidos. Se debe tener en cuenta si se tiene índices y el tipo de índices también como parte del tamaño de la tabla.

1

2

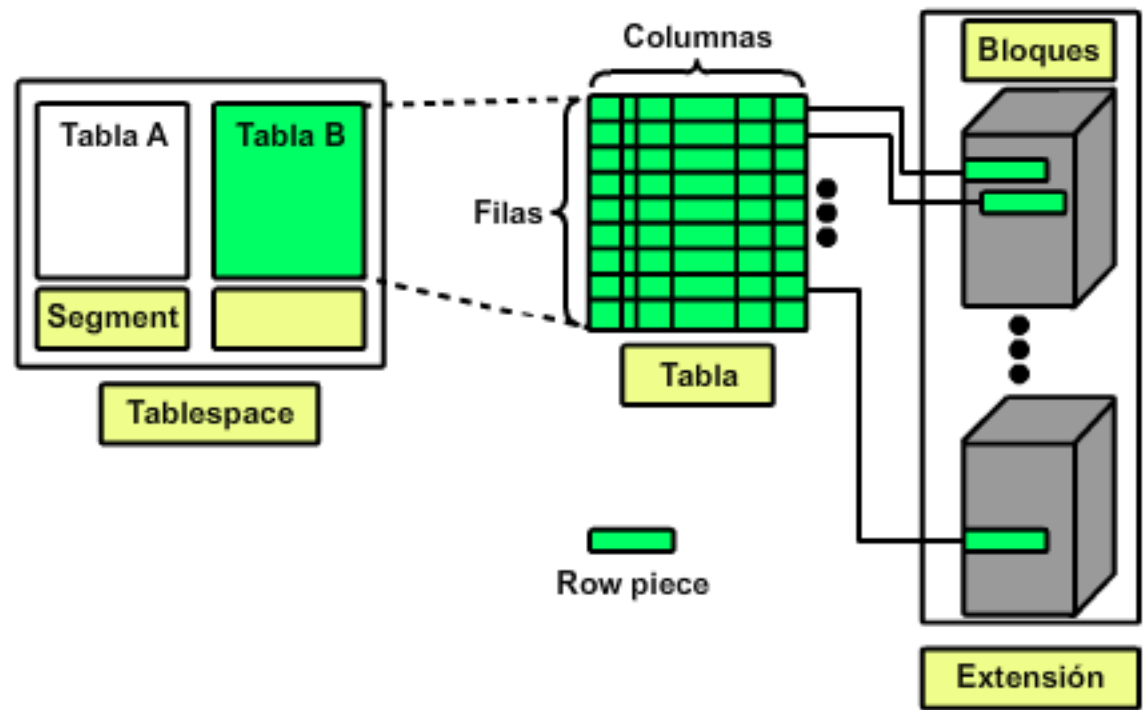
3



Modelo Relacional



Almacenamiento



Un cálculo base del espacio requerido lo entrega el modelado de las bases de datos al determinarse cuántos registros se esperan por cada tabla en un período determinado. Por ejemplo en un mes. Este valor se debe proyectar o multiplicar por el período de vida útil activa de la información por ejemplo 12 meses, 24, 36, etc.

El ejercicio se realiza para cada tabla u objeto de la base de datos. Al final se suman todos los datos y se obtiene una aproximación al tamaño de las bases de datos.

- 1
- 2
- 3



Modelo Relacional

Normalización

Reglas de Integridad
(Constraints)

Almacenamiento

Almacenamiento

Por cada tabla se debe tener en cuenta el tamaño de cada atributo o columna, el cual depende a su vez del tipo de información que contenga. Esta información se encuentra registrada en el **diccionario de datos**. Se debe verificar de acuerdo al Sistema Manejador de Base de Datos, cuántos bytes utiliza por cada tipo de dato. Por ejemplo, si el SMBD utiliza para un tipo Numeric 4 bytes y en la tabla hay tres campos Numeric, ya se tienen 12 bytes a ocupar.

Cuando el campo ha sido definido de tipo variable, se debe calcular, inicialmente, al máximo posible. Esto puede variar de acuerdo a la oportunidad del dato pero es mejor pecar por exceso que por defecto.

Es muy importante hacer la consulta de los tamaños que utiliza el Gestor de Base de Datos así como la forma en que almacena la información (por ejemplo campos Null, seguimiento interno de datos y cambios a los mismos, etc). La arquitectura de almacenamiento de cada SMBD influye también en la estimación de tamaños.

1

2

3

Diseño de Base de Datos

Una de las tareas más importantes para un Administrador de Bases de Datos es la del diseño de base de datos, que abarca desde el modelamiento de datos en el ámbito lógico conceptual, hasta la definición de las estructuras de almacenamiento, para lo cual se debe poseer total claridad y entero conocimiento del modelo de negocios, requerimientos de información y de los recursos e infraestructura con la que cuenta la organización tanto para el almacenamiento como para su procesamiento.