

## **Capítulo 2: Base de datos**

## 2.1 Modelado de la base de datos

### 2.1.1 Especificación de OpenGIS

Las principales compañías de desarrollo de aplicaciones GIS en el mundo se encargaron de crear una especificación estándar para el diseño de bases de datos espaciales, entre estas compañías encontramos a Environmental Systems Research Institute, IBM Corporation, Informix Software Inc, MapInfo Corporation y Oracle Corporation. El documento de especificación corresponde por su nombre en inglés a “OpenGIS Simple Features Specification for SQL, Revision 1.1”, liberado el 5 de mayo de 1999. A partir de entonces se han creado *schemas* de bases de datos que respetan tal especificación [Ryden,1999].

Existen tres representaciones principales con las cuales es posible modelar la base de datos según la especificación en OpenGIS [ArcSDEonline,2004].

- 1.- *Con geometrías normalizadas*: Almacena el *bounding box* del GML y normaliza todas las coordenadas inmersas en el documento.
- 2.- *Almacenamiento de geometrías en binario*: Almacena tal cual las coordenadas pero en binario.
- 3.- *Utilizar en SQL extendido para geometrías*: Almacena las coordenadas en un tipo de datos geométrico extendido para SQL espacial.

En la siguiente tabla se muestra los esquemas a utilizar para almacenar geometrías según el tipo de base de datos utilizada:

Tabla 3.1 (SGBD's y geometrías espaciales) [ArcSDEonline,2004]

| SGDB      | Almacenamiento geométrico                     | Tipo de la columna |
|-----------|---|--------------------|
| DB2 (IBM) | SQL extendido ( <i>Spatial Extender</i> )     | Tipo Geometría     |
| Oracle    | Esquema binario                               | Long Raw           |
|           | Esquema binario                               | Blob               |
|           | SQL extendido ( <i>Oracle Spatial</i> )       | Tipo Geometría     |
|           | Esquema normalizado ( <i>Oracle Spatial</i> ) | Integer            |
| MySQL     | SQL extendido                                 | Tipo Geometría     |
|           | Esquema binario                               | Blob               |
| Postgres  | SQL extendido                                 | Tipo Geometría     |
| Informix  | SQL extendido ( <i>Spatial DataBlade</i> )    | Tipo Geometría     |
| MS SQL    | Esquema binario                               | Image              |

MS SQL 2000 no maneja SQL extendido para la administración de datos geométricos. Por lo que es necesario implementar completamente el esquema geométrico en binario.

### 2.1.2 Metadatos geométricos

El modelado de la información para a la especificación de OpenGIS se basa en la estructura de dos tablas que contendrán todas las relaciones geométricas necesarias para almacenarla.

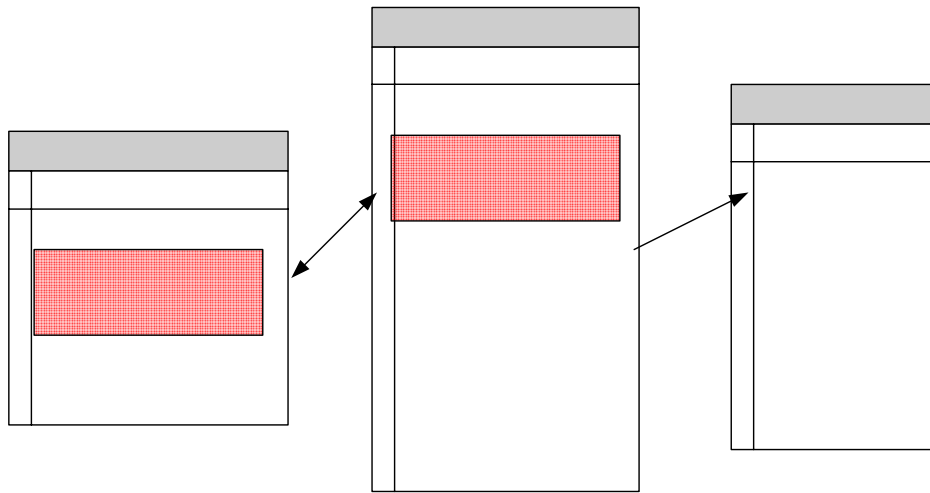


Imagen 2.1 [Tablas de los metadatos geométricos]

La tabla de SPATIAL\_REF\_SYS contiene la descripción del sistema espacial en el que se está trabajando la base de datos geográfica, esto implica además información del autor y las unidades en que se están describiendo los datos.

#### GEOMETRY\_COLUMNS

La tabla de LAYERS es utilizada solamente por geometrías normalizadas y se utiliza para guardar información específica de la capa (como el *bounding box*, unidades y tipos de almacenamiento), permitiendo así que se puedan normalizar todas las coordenadas geométricas encontradas en el documento.

SRID  
 F\_TABLE\_CATALOG  
 F\_TABLE\_SCHEMA  
 F\_TABLE\_NAME  
 F\_GEOMETRY\_COLUMN  
 STORAGE\_TYPE  
 COORD\_DIMENSION  
 MAX\_PPR

LAYER\_  
 DESCR  
 TABLE\_  
 OWNER  
 SPATIA  
 SRID  
 STORA  
 EGLAFC  
 MINX  
 MINY  
 MAXX  
 MAXY  
 CDATE  
 OPTIMA  
 STATS

GEOMETRY\_COLUMNS almacena las referencias a las tablas de atributos y geométricas, la información sobre el tipo de almacenamiento y el tipo de geometrías de una capa en particular, además de contener una llave foránea sobre algún sistema espacial de referencia.

El esquema FEATURES es utilizado generalmente para complementar la información que se tiene de la capa, esto implica que se almacena el *boundig box*, el tipo de almacenamiento y ciertas banderas. La tabla FEATURES generalmente es de uso de ESRI con ArcSDE y se utiliza para normalizar la información geométrica.

La relación entre estas tablas se encuentra expresada con el SRID que lo liga a un sistema espacial de referencia y con el nombre de la tabla que contiene los atributos.

### 2.1.3 Implementación del esquema geométrico binario

Como se mencionó con anterioridad, el esquema geométrico binario es otra forma de manejar información geométrica. Este tipo de manejo tiene la cualidad de que las coordenadas son almacenadas en un campo binario, es decir, en un Blob, binary o Image según sea el caso.

Las tablas utilizadas para administrar la introducción de nuevas capas en un esquema geométrico binario corresponden a SPATIAL\_REF\_SYS y a GEOMETRY\_COLUMNS [Ryden,1999].

La implementación utilizada hace uso de los principales campos del esquema, pero puede ser expandido cuando se requiera. A continuación se presenta el esquema que fue utilizado para el diseño de la base de datos. Si se verifica contra la especificación de OpenGIS [Ryden,1999] se aprecia que no se encuentra implementada en su totalidad, pero algunos campos fueron desechados por su poca relevancia en el proyecto, además de que pocos documentos incluyen tal información, a pesar de ello la estructura puede ser modificada sin ninguna consecuencia cuando se deseé.

Tabla 2.2 Esquema implementado

```
Create table SPATIAL_REF_SYS(  
  SRID int not null,  
  AUTH_NAME varchar(256) not null,  
  AUTH_SRID int not null,  
  SRTEXT varchar (2048) not null,  
  primary key (SRID )  
);  
Create table GEOMETRY_COLUMNS(  
  F_TABLE_NAME varchar(200),  
  F_GEOMETRY_COLUMN varchar(200),  
  G_TABLE_NAME varchar(256),  
  STORAGE_TYPE int,  
  GEOMETRY_TYPE int,  
  COORD_DIMENSION int,  
  MAX_PPR int,  
  SRID int  
);
```

A continuación se detallará cada una de las tablas representadas en el esquema implementado en la tabla 2.2 del documento.

Tabla: SPATIAL\_REF\_SYS:

SRID: Se refiere al valor único de identificación en el sistema de referencia espacial en la base de datos. Además es definido como llave primaria.

AUTH\_NAME: Especifica el nombre del autor.

AUTH\_SRID: La llave especificada en el AUTH\_NAME de la tabla de sistema espacial de referencia.

SRTEXT: La representación textual más conocida del sistema espacial de referencia.

GEOMETRY\_COLUMNS:

F\_TABLE\_NAME: Contiene el nombre de la tabla con los atributos de las geometrías de la capa.

F\_GEOMETRY\_COLUMN: Contiene el nombre de la columna que relaciona la información geométrica y de atributos.

G\_TABLE\_NAME: Es el nombre de la tabla con la información geométrica.

STORAGE\_TYPE: El tipo de almacenamiento.

0.- Es un almacenaje normalizado según la especificación.

1.- El almacenaje es en geometrías binarias.

2.- OGIS Implementación de la extensión de SQL.

GEOMETRY\_TYPE:

- 0.- Geometría.
- 1.- Punto.
- 2.- Curva.
- 3.- Polilíneas.
- 4.- Superficie.
- 5.- Polígono.
- 6.- Colección.
- 7.- Multipunto.
- 8.- Multicurva.
- 9.- Multipolilínea.
- 10.-Multisuperficie.
- 11.- Multipolígono.

COORD\_DIMENSION: Número de coordenadas utilizadas en el sistema. (Sólo aplicable al esquema de normalización).

MAX\_PPR: Máxima cantidad de puntos por fila. (Sólo aplicable al esquema de normalización).

SRID: ID correspondiente al sistema geométrico del cual forma parte.



Lo anterior permite que se puedan crear las tablas de cada capa dinámicamente especificando la información geográfica y sus características (lo que en inglés se conoce como *Features*). El esquema de características será titulado con el nombre de la capa y la tabla geométrica que tendrá el mismo nombre de la capa y un sufijo detallando su carácter geométrico ( *tabla\_geom* ).

El esquema de características contiene un *feature ID* o identificador de característica (fid) y funge como la llave primaria en la relación, inmediatamente seguido se encuentra la lista de atributos que son propios de la capa y finalmente el *geographic ID* o identificador geográfico que permite asociar las dos tablas *Feature* y *Geometry*. La cantidad de atributos de la tabla de características depende de la capa que se está representando.

La creación de la tabla geométrica no es dinámica como la tabla de características. El esquema utilizado para la tabla geométrica es el siguiente:

Tabla 2.3 Esquema de *FEATURE\_GEOM*

```
create table FEATURE_GEOM(  
  GID int not null,  
  XMIN int not null,  
  XMAX int not null,  
  YMIN int not null,  
  YMAX int not null,  
  primary key (GID)  
);
```

**GID:** Es el identificador de la entidad geométrica almacenada, además de ser la llave primaria de la tabla.

**XMIN:** Corresponde al valor mínimo en x del *bounding box* de la geometría almacenada en la tabla.

**XMAX:** Corresponde al valor máximo en x del *bounding box* de la geometría almacenada en la tabla.

**YMIN:** Corresponde al valor mínimo del *bounding box* de la geometría almacenada en la tabla.

**YMAX:** Corresponde al valor máximo del *bounding box* en la geometría almacenada en la tabla.

**WKB:** También conocido como *Well known binary*. Corresponde a la información geométrica almacenada en forma binaria.

La estructura propuesta para la información almacenada en el WKB es la propuesta dentro de la especificación de OpenGIS:

**Polilíneas:** Se incluye la información geométrica de la polilínea en binario, se encuentra entre paréntesis en la parte derecha y la característica geométrica como cabecera. La información geométrica incluye una secuencia de parejas de puntos: ordenada y abscisa, separada por comas y cada coordenada separada por comas.

Polyline(23,23 45,43 34,34)

Polígonos: Al igual que las polilíneas, la información geográfica de los polígonos se encuentra codificada a forma binaria, incluyendo al igual el nombre de la característica geométrica como cabecera y dentro de los paréntesis la información espacial. El polígono, tiene como cualidad, que está formado por un espacio geométrico cerrado pero la figura no es obligatoriamente convexa, es decir, puede incluir espacios geométricos dentro de la misma que no correspondan a la región del polígono. La región externa del polígono se representa primero dentro de la región geométrica y después se adicionan, si existen, los arcos internos. En particular las geometrías especificadas dentro de ambas regiones deben representar espacios cerrados. [Rigaux,2002]

Polygon( (0,0 10,0 10,10 0,10, 0,0) (0,0 0,5 5,5 0,5 0,0))

Es importante hacer notar que la representación se encuentra exclusivamente en un espacio bidimensional, en caso de que se deseen incluir coordenadas tridimensionales se puede asociar como un atributo de la geometría y para la presente tesis solamente se tomaron en cuenta estas dos dimensiones pero el modelado puede incluir cualquiera que respete la especificación de OpenGIS.

Las tablas anteriormente descritas contendrán toda la información necesaria para almacenar la información geométrica. El esquema propuesto; debido a que está implementado utilizando la especificación de OpenGIS es posible escalarlo al resto de

atributos y tablas que propone el documento. Los anteriores atributos y relaciones presentadas son las mínimas necesarias para poder modelar la información geográfica en el sistema administrador de base de datos. Adicionalmente, es posible escalar el modelo aquí presentado sin ningún impacto sobre el diseño, lo que permitiría guardar más información del documento geográfico.