



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES  
CUAUTITLÁN**

**REDES DE COMPUTADORAS**

**“ CREACIÓN DE UNA BASE DE DATOS EN UNA RED LOCAL  
PARA LA RECUPERACIÓN DE DATOS SÍSMICOS ”**

**TRABAJO DE SEMINARIO**

QUE PARA OBTENER EL TÍTULO DE:

LICENCIADO EN INFORMÁTICA

P R E S E N T A:

**ARTURO PORTILLO SANABRIA**

ASESOR: ING. MIGUEL ALVAREZ PASAYE

**TESIS CON  
FALLA DE ORIGEN**



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLÁN  
UNIDAD DE LA ADMINISTRACIÓN ESCOLAR  
DEPARTAMENTO DE EXÁMENES PROFESIONALES

U. N. A. M.  
FACULTAD DE ESTUDIOS  
SUPERIORES CUAUTITLÁN



DEPARTAMENTO DE  
EXÁMENES PROFESIONALES

DR. JAIME KELLER TORRES  
DIRECTOR DE LA FES-CUAUTITLÁN  
PRESENTE.

AT'N: ING. RAFAEL RODRIGUEZ CEBALLOS  
Jefe del Departamento de Exámenes  
Profesionales de la FES-C.

Con base en el art. 51 del Reglamento de Exámenes Profesionales de la FES-Cuautitlán, nos permitimos comunicar a usted que revisamos el Trabajo de Seminario:

"Redes de Copradoras: Creación de una Base de Datos en una Red Local para la recuperación de Datos Simicos"

que presenta al pasante: Arturo Portillo Sanabria.

con número de cuenta: 8738156-6 para obtener el Título de:  
Licenciado en Informática

Considerando que dicho trabajo reúne los requisitos necesarios para ser discutido en el EXAMEN PROFESIONAL correspondiente, otorgamos nuestro VISTO BUENO.

ATENTAMENTE.

"POR MI RAZA HABLARA EL ESPIRITU"

Cuautitlán Izcalli, Edo. de México, a 15 de Octubre de 19 97

MODULO:  
Modulo II  
Modulo III  
Modulo IV

PROFESOR:  
Ing. Miguel Alvarez Penayo  
Ing. Juan Manuel Hdz. Duarte  
Ing. Francisco I. Chávez G.

FIRMA:

---

## **AGRADECIMIENTOS**

### **A mis padres.**

Con un sincero agradecimiento por su cariño, amor, comprensión y todos los esfuerzos que realizaron, sin los cuales no hubiese llegado al agrado de desarrollo que a la fecha he logrado.

Deseo que hagan conciencia de que este resultado obtenido, no sólo es para mí, sino también para ustedes, ya que esto es el logro de una de las principales metas que para conmigo se fijaron.

### **Hermanos.**

Con una muestra de agradecimiento, por todo el apoyo moral que me brindaron para que terminara mi Carrera Profesional.

### **Amigos.**

Doy gracias a todas aquellas personas que colaboraron de una u otra forma en la realización de mi objetivo, en especial al Ing. Miguel Alvarez Pasaye, a mis compañeros de trabajo Cecilia, Paty, Miguel, David y Personas con una estimación especial Ana, Carlos, Erick, Jesús, Rogelio.

---

---

## INDICE

### INTRODUCCIÓN

#### CAPITULO I

1	Almacenamiento De Datos	2
1.1	Elección De Almacenamiento	2
1.2	Almacenamiento Secundario Para Acceso Directo	3
1.2.1	Almacenamiento En Disco Para Acceso Directo	3
1.2.2	Tipos De Discos Magnéticos	3
1.2.3	Almacenamiento De Datos En Discos Magnéticos	4
1.2.4	Ventajas Y Limitaciones De Los Discos Magnéticos	5
1.3	Acceso Directo Con Discos De Ram Y Burbuja Magnética	6
1.3.1	Almacenamiento De Discos En Ram	6
1.3.2	Almacenamiento En Burbuja Magnética	6
1.4	Acceso Directo A Disco Opticos	7
1.5	Almacenamiento Secundario Para Acceso Secuencial	8
1.5.1	Almacenamiento De Disco Para Acceso Secuencial	8
1.5.2	Almacenamiento En Cinta Magnética	8
1.5.3	Almacenamiento De Datos En Cinta Magnética	9
1.5.4	Acceso A Los Datos En Cinta Magnética	10
1.5.5	Ventajas Y Limitaciones De La Cinta Magnética	11

#### CAPITULO 2

2	Importancia De La Base De Datos	13
2.1	Enfoque De Las Bases De Datos	14
2.1.1	Conceptos Básicos	14
2.1.2	Conceptos Asociados A Bases De Datos	16
2.1.3	Necesidad De La Bases De Datos	17
2.1.4	Ventajas De Una Base De Datos	18
2.2	Sistema Manejador De Bases De Datos ( DBMS )	18
2.2.1	Características De Los DBMS	19
2.2.2	Elementos De DBMS	20
2.3	Bases De Datos Jerárquicas	21
2.3.2	Ventajas Del Modelo Jerárquico	23
2.3.3	Limitaciones De Un Modelo Jerárquico	23
2.4	Bases De Datos En Red	24
2.4.1	Ventajas De Un Modelo De Red	25
2.4.2	Limitaciones De Un Modelo De Red	25

---

---

2.4.2	Limitaciones De Un Modelo De Red	25
2.5	Bases De Datos Relacionales	26
2.5.1	Conceptos Fundamentales	27
2.5.2	Las 12 Reglas De COOD	29
2.5.3	Ventajas De Un Modelo Relacional	34
2.5.4	Limitaciones De Un Modelo Relacional	35

### **CAPITULO 3**

3.1	El Desarrollo De Sistemas Informáticos	36
3.1.1	Metodología Para El Desarrollo De Sistemas Informáticos	40
3.1.2	Importancia De La Verificación Del Requerimiento	45
3.2	Evaluación De Proyectos	45
3.2.1	Importancia De Los Proyectos	46
3.2.2	Metodología De La Evaluación De Proyectos	47

### **CAPITULO 4**

4.1	Verificación De Los Requerimientos	53
4.1.1	Definición Conceptual Del Proyecto	53
4.2	Planeación Y Control Del Proyecto	54
4.2.1	Identificación Del Grupo De Trabajo	56
4.3	Análisis Del Proyecto	60
4.4	Diseño Del Proyecto	69
4.5	Pruebas	71
4.6	Modelo Lógico De La Base De Datos	72

<b>CONCLUSIONES</b>	<b>83</b>
<b>REFERENCIAS BIBLIOGRAFICAS</b>	<b>84</b>

---

---

## **INTRODUCCION**

En todas las esferas de la vida y en todas las áreas de la industria y el comercio, las bases de datos amplían en gran medida las posibilidades de acción abiertas al hombre. Algunos de los aspectos sobresalientes del desarrollo industrial de nuestros días se deben al extraordinario crecimiento de la disponibilidad de información.

En la actualidad la información juega un papel importante dentro de cualquier organización donde los datos son recursos administrables y que con su almacenamiento y empleo adecuado pueden proporcionar una importante herramienta en la toma de decisiones de una determinada organización. Por tanto, mientras mayor sea la cantidad de datos a que tiene acceso a una base de datos, aumentará el potencial de información que proporcione.

Sin embargo el hecho, de que en la actualidad se tenga una gran disponibilidad de información en el momento requerido, no sólo se debe a la innovación tecnológica, sino también al desarrollo de nuevos métodos para el almacenamiento de datos que diariamente se requieren para administrar negocios, bancos, universidades, organismos gubernamentales y otras actividades humanas.

El medio ambiente de base de datos, con un software de apoyo y las técnicas de manejo que le acompañan, proporcionan el marco para tratar los datos como un recurso estandarizado, administrable y compatible. Compartir archivos de datos entre distintas aplicaciones y emplear descripciones de datos y proposiciones de acceso consistentes en los programas en el medio ambiente de base de datos, ayudan a tener una visión cuidadosa y estandarizada de todo el proceso de desarrollo de un sistema de información.

---

# **CAPITULO 1**

---

---

## 1. ALMACENAMIENTO DE DATOS

Los métodos para desarrollar un sistema de información están evolucionando de varias formas con respecto al uso de nuevas herramientas y técnicas, como lo hemos venido observando a lo largo de estos años, involucrando también a los usuarios finales en el proceso de desarrollo.

El centro de muchas actividades de los desarrolladores está en la forma de como se almacenan los datos en los sistemas para su recuperación cuando sea necesario. Aquí hay algo que permanece constante. Las computadoras, ya sea grandes o pequeñas, y los programas que se utilizan (lenguajes de tercera, cuarta o quinta generación), siguen almacenando los datos en dos únicas maneras: secuencialmente, con un registro detrás de otro, o aleatoriamente en lugares específicos de almacenamiento. Estos métodos son una característica arquitectónica fundamental de los dispositivos de almacenamiento de computadoras.

En cierto momento, los datos deben traducirse a una estructura de almacenamiento, ya sea secuencial o aleatoriamente. Ahí es donde se llega al punto importante. Para elegir un medio de almacenamiento de debe tener en consideración tres aspectos importantes:

1. **Velocidad de recuperación.** El tiempo de acceso de un dispositivo de almacenamiento es el tiempo que se requiere para localizar y recuperar los datos almacenados, en respuesta a una instrucción del sistema. Es preferible un tiempo de acceso corto.
2. **Capacidad de almacenamiento.** Debe tomarse en cuenta la capacidad de un dispositivo para almacenar la cantidad de datos que se requiere actualmente y en el futuro.
3. **Costo por bit de capacidad.** La reducción de este costo es un objetivo obvio.

### 1.1 ELECCION DE ALMACENAMIENTO

Si existiera un componente de almacenamiento que fuera el más rápido, el de mayor capacidad y de menor costo, no habría que tener una jerarquía de almacenamiento. La forma en que se organizan los datos determina el método elegido. Si un archivo se puede organizar de manera secuencial y si los registros sólo requieren actualizaciones periódicas, es muy probable que la opción con el menor costo sea el uso de un medio de almacenamiento secundario por acceso secuencial. Por otro lado, si se requiere de un acceso rápido a cualquier

---

registro de los archivos , será preciso instalar un dispositivos de almacenamiento por acceso rápido, por su rápida respuesta.

Para elegir un dispositivo de almacenamiento por acceso rápido se tiene que buscar un equilibrio entre rendimiento y costo. En algunas entidades es posible que la velocidad de procesamiento sea de primordial importancia y el costo sea secundario.<sup>2</sup>

## **1.2 ALMACENAMIENTO SECUNDARIO PARA ACCESO DIRECTO**

### **1.2.1 ALMACENAMIENTO EN DISCO PARA ACCESO DIRECTO**

Los discos magnéticos son el medio más popular de almacenamiento secundario para acceso directo. Un disco magnético es un plato metálico o plástico en donde los datos se escriben o se leen de los discos ( o diskettes) sólo cuando están girando dentro del drive, (unidad de disco). Al igual que un cinta de audio, los registros de los datos organizados en una cinta magnética deben leerse según la secuencia de almacenamiento. Como las melodías, los datos grabados en un disco magnético también pueden leerse en secuencia si se organizaron en una forma que permita este tipo de recuperaciones. Las cabezas de lectura/escritura, son electromagnetos pequeños capaces de leer, grabar o borrar las marcas polarizadas están montadas en un brazo de un dispositivo de almacenamiento en disco, pueden moverse rápidamente y directamente a cualquier punto del disco para almacenar y recuperar datos. Los datos grabados en un disco pueden leerse muchas veces. Permanecerán allí por tiempo indefinido a hasta que se borre y vuelva a utilizar la superficie de disco.

### **1.2.2 TIPOS DE DISCO MAGNETICOS**

Todos los discos magnéticos son platos redondos recubiertos de una material de grabación magnetizable, pero hasta allí llega el parecido. Existen varios tamaños de disco. Pueden ser portátiles o estar montados de manera permanente en sus dispositivos de almacenamiento (llamadas unidades de disco), pueden estar hechos de material rígido o plástico flexible. Algunas de las posibles opciones de discos:

- Discos de material grande ( 14 pulgadas) montadas permanentemente en envases sellados, libres de contaminación.

---

<sup>2</sup> James A. Seen " Análisis y Diseño de Sistemas "

- 
- Otros discos de metal de ( 14 pulgadas ) empacados en cartuchos removibles o paquetes de discos.
  - Discos rígidos más pequeños ( 9,8 y 5 1/4 pulgadas) montadas permanentemente en dispositivos de discos.
  - Discos flexibles portátiles de 8 y 5 1/4 pulgadas. empacados individualmente en fundas protectoras.
  - Discos flexibles y rígidos todavía compactos con un diámetro de menos de cuatro pulgadas. los diskettes más populares son los de 3 1/2 .

### **1.2.3 ALMACENAMIENTO DE DATOS EN DISCOS MAGNETICOS.**

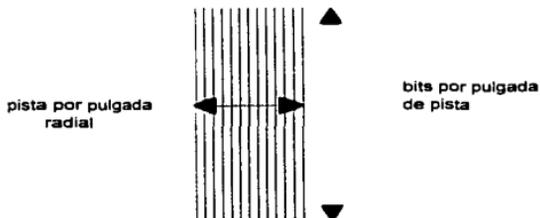
En un disco fonográfico, la música se almacena en un surco continuo que sigue un trayectoria espiral hacia el centro del disco. Pero en un disco magnético no hay surcos. En vez de ello, los datos se almacenan en todos los discos, en varios círculos concéntricos invisibles llamados pistas. Estas pistas, como los anillos de un árbol, comienzan de un borde exterior del disco y continúan hacia centro sin tocarse jamás. Todas las pistas están enumeradas.

Un motor hace girar el disco rápidamente. Los datos se graban en las pistas de la superficie del disco en movimiento y se leen de esa superficie, por medio de una o más cabezas de lectura/escritura. Si se utiliza un disco flexible, la cabeza entra en contacto con el disco. Si se emplean discos duros, la cabeza vuela sobre un cojín de aire unas cuantas micras por encima de la superficie.

Cuantas más superficies tenga un sistema de disco, mayor será su capacidad de almacenamiento. Los primeros sistemas de discos flexibles están diseñados de tal manera que sólo se grababan los datos en una superficie de disco. El acceso a la superficie de estos discos de un sólo lado se logra a través de una ranura o ventana de uno de los lados de la funda. Pero los sistemas posteriores se hicieron para utilizar discos flexibles de dos lados que tienen ranuras por ambas superficies. Naturalmente, la capacidad de almacenamiento de un disco de dos caras es casi el doble de la de los discos de un solo lado.

La capacidad de almacenamiento de un sistema de disco depende también de los bits por pulgada de pista y el número de pistas por pulgada radial, es decir. La capacidad de almacenamiento depende del número de bits que pueden almacenarse en una pulgada de pista y del número de pistas que se pueden grabarse en una pulgada en el sentido radial del disco. Los

primeros discos flexibles tenían un sólo lado y eran de densidad sencilla ( SD, single density), mas tarde aparecieron de un lado de doble densidad ( DAD, double density ), capaces de almacenar más bits por pulgada de pista. Ahora son comunes los discos de dos lados y doble densidad, pero a últimas fechas se han duplicado el número de pistas por pulgada para producir discos de cúaduple densidad ( QD, quad-density)



**Velocidad de acceso** . El tiempo de acceso de los datos almacenados en un disco está determinado básicamente por :

1. **El tiempo de posicionamiento** : es el tiempo que se requiere para colocar una cabeza sobre la pista apropiada.
2. **El tiempo de búsqueda**: es el tiempo que se requiere para girar el disco de manera que el dato deseado quede bajo la cabeza.

## 1.2.4 VENTAJAS Y LIMITACIONES DE LOS DISCOS MAGNETICOS

A continuación se hará un resumen de algunas de las ventajas de los discos:

- Los registros en los discos pueden almacenarse en forma secuencial y procesarse con los registros en cinta magnética o. pueden almacenarse para procesarlos en forma directa.

---

- Pueden tener acceso a cualquier registro de disco en línea y pueden actualizarse en unos cuantos milisegundos. No es necesario clasificar las transacciones.

- Puede utilizarse una sola transacción de entrada de datos para actualizar los registros de disco en línea de varios archivos relacionados entre sí.

Pero los discos también tienen limitaciones:

- Cuando se actualiza un archivo en una cinta magnética, la cinta muestra anterior no suele sufrir alteraciones y sigue disponible durante cierto tiempo, en el caso que se presenten fallas en los sistemas. En cambio, cuando se actualizan los registros de un archivo en disco, es posible que se borren los registros anteriores cuando los registros nuevos se graban en el disco.

- Ha habido gente que elude la seguridad y logra tener acceso a los archivos importantes de disco desde terminales remotas, estos archivos se han manipulado y hasta destruido. Es más fácil mantener la seguridad de los archivos en las cintas magnéticas.

### **1.3 ACCESO DIRECTO CON DISCO DE RAM Y BURBUJA MAGNETICA.**

#### **1.3.1 ALMACENAMIENTO EN DISCO DE RAM**

Este método de almacenamiento de RAM ( Random Access Memory, Memoria de Acceso Aleatorio) se utiliza en algunas computadoras personales y consiste en emplear un bloque de pastillas semiconductoras RAM como " disco" simulado para evitar operaciones mecánicas de los discos giratorios. Un disco RAM no es en realidad un disco, ni tiene partes móviles, sino es un conjunto de pastillas RAM configurado de tal manera que la computadora cree que en realidad es un disco.

La principal ventaja de los discos RAM es la velocidad, par agilizar sus sesiones de programación. los programadores pueden copiar sus programas de discos flexibles o de algún otro medio permanente al disco RAM y evitar en su totalidad los retrasos causados por llamadas frecuentes al disco.

#### **1.3.2 ALMACENAMIENTO EN BURBUJAS MAGNETICAS**

Los dispositivos de almacenamiento de burbuja magnética son también pastillas semiconductoras que no tienen partes móviles y son, por tanto, muy confiables. las pastillas de

---

burbujas conservan los datos almacenados cuando se deja de suministrar energía. La terminales portátiles utilizan dispositivos de burbuja para almacenar datos, hasta que puedan transmitirse a sistemas más grandes, también son utilizadas en talleres, robots y computadoras militares.

## **1.4 ACCESO DIRECTO A DISCO OPTICOS**

Las técnicas de almacenamiento se pueden basar en el empleo de la luz, en lugar de hacerlo en los principios de magnetismo, los discos ópticos utilizan un recubrimiento delgado de metal u otro material depositado sobre el disco, se queman o estampan diminutas perforaciones que solamente se pueden ver con el microscopio. Estas perforaciones pueden hacerse en un disco de 12 pulgadas a fin de grabar las imágenes y sonido, también se pueden hacer discos compactos para grabar música de alta calidad. En ambos casos, los patrones de las perforaciones representan los datos digitales mediante los cuales se codifican imágenes y sonidos. Cuando se colocan estos dispositivos reproductores, se utiliza un rayo láser para leer los patrones de perforaciones y convertirlos en las señales audiovisuales que se requieren para producir las imágenes y el sonido de la televisión o los sonidos de los discos compactos.

La tecnología de los discos compactos de audio se adaptó para desarrollar discos de almacenamiento óptico CD-ROM (disco compacto, memoria sólo lectura) y las unidades correspondientes que se pueden usar con computadoras personales. La densidad de almacenamiento de disco óptico es enorme, el costo de almacenamiento es muy bajo y el tiempo de acceso es relativamente breve. En un sólo disco CD-ROM caben cerca de 550 megabytes (un mega equivale a 1.048 millones de bytes, por lo que 550 megabytes son más de 570 millones de caracteres alfanuméricos).

Los CD-ROM son sólo una forma de almacenamiento óptico. Además de estos dispositivos pregrabados sólo lectura, existen dispositivos de almacenamiento óptico con base a los rayos láser en los cuales se puede escribir. Con los sistemas de escritura única y lectura múltiple (WORM, write-once, read-many), los usuarios graban sus propios datos en discos ópticos en blanco. La permanencia de los discos WORM es deseable en aplicaciones de archivo, que actualmente dependen de microfilm, papel o medios magnéticos. Los sistemas WORM se pueden usar en instalaciones de cómputo de todos los tamaños y también se pueden emplear cuando es preciso actualizar los datos. Dada la enorme capacidad de almacenamiento, los usuarios pueden seguir almacenando y recuperando información más reciente hasta llenar en su totalidad el disco.

---

## **1.5 ALMACENAMIENTO SECUNDARIO PARA ACCESO SECUENCIAL**

Los registros de un archivo secuencial a menudo se graban uno tras otro en el orden ascendente o descendente determinado por la llave de registro. Para tener acceso a los registros, la computadora comienza con el primero de la secuencia. Este registro se lee y puede ser que se procese o no. A continuación se lee el segundo registro, seguido del tercero, y así sucesivamente. Los archivos secuenciales se mantuvieron durante años en cajas de tarjetas perforadas, pero los medios de almacenamiento que se utilizan ahora para el acceso secuencial son los discos y cintas magnéticas.

### **1.5.1 ALMACENAMIENTO EN DISCO PARA ACCESO SECUENCIAL.**

Un ejemplo claro de este tipo de acceso puede ser grabar la música en un disco fonográfico en donde se sigue una secuencia determinada, pudiéndose más tarde reproducir en la misma secuencia si eso es lo que se desea. De manera similar, los datos grabados en un disco magnético pueden almacenarse y después recuperarse en un orden previamente determinado. Para apoyar una aplicación de procedimiento secuencial en un disco no se está utilizando realmente la capacidad de acceso directo de la unidad de disco y puede recuperarse de manera directa el primer registro de la secuencia, pero los demás de leerán y procesarán en orden como si estuvieran almacenados en una cinta magnética.

### **1.5.2 ALMACENAMIENTO EN CINTA MAGNETICA.**

Una de las aplicaciones más frecuentes es emplear la cinta magnética para respaldar el contenido de las unidades de disco duro. Además, es muchas veces el medio elegido para almacenar archivos grandes que se leen y procesan de manera secuencial. En las bibliotecas de cinta magnética de los grandes sistemas de cómputo se mantienen miles de carretes de datos almacenados. Aunque las ventas de unidades de almacenamiento en disco magnético rígido y en disco óptico están creciendo rápidamente, la cinta magnética sigue siendo un competidor importante para fines de almacenamiento.

---

### 1.5.3 ALMACENAMIENTO DE DATOS EN CINTA MAGNETICA.

En esta forma de almacenamiento los datos se almacenan en forma de pequeñas marcas en el material magnetizable que cubre una cara de la cinta de plástico. La superficie recubierta de la cinta se divide en columnas verticales (o cuadros) e hileras horizontales (llamadas canales o pistas). En muchos sistemas de cinta actuales se emplea un código BCD ( Binary Coded Decimal, decimal codificado en binario) de ocho bits con un formato de cinta magnética de nueve pistas.

Hoy día es común utilizar nueve pistas para grabar datos en carretes de cinta estándar de media pulgada de ancho. Sin embargo, hace poco se idearon sistemas de cinta más modernos que graban 18 pistas de datos en la misma superficie. La cinta de estos sistemas se almacenan en cartuchos de 10 X 13 cm.

El código para cinta de nueve pistas se utiliza en muchas computadoras modernas. En este ejemplo se usó un formato EBCDIC ( Extended Binary-Coded Decimal Interchange Code, código extendido de intercambio decimal codificado en binario ) de ocho bits.

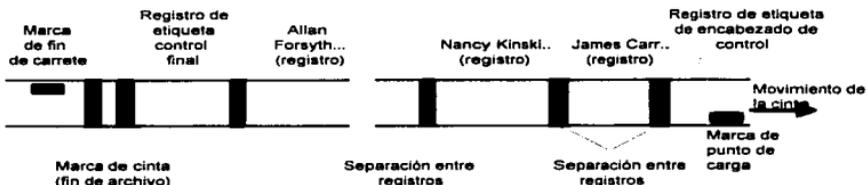
Una cinta de nueve pistas está recubierta con material magnetizable a base de óxido de hierro, pero los cartuchos nuevos utilizan un recubrimiento de dióxido de cromo. Pueden almacenarse hasta 6250 caracteres en una sola pulgada de cinta de nueve pistas, pero en ese mismo espacio pueden caber hasta 38000 caracteres cuando se utiliza un sistema de cartucho de 18 pistas.

Dado de cualquier cinta magnética es un medio continuo, cómo es posible identificar registros individuales en una cinta. La respuesta es que puedan separarse los registros mediante un espacio en blanco llamado separación entre registro, esta separación se crea automáticamente al grabar los datos en la cinta. Cuando se leen los datos desde un registro de un cinta en movimiento al procesador, la cinta se detiene al llegar a una separación, y permanece inmóvil hasta que se termina de procesar el registro y a continuación se mueve otra vez para pasar el siguiente registro a la computadora. El procedimiento se repite hasta que se procesa todo el archivo. La figura 1.1 muestra cómo podría representarse en cinta un archivo de empleados. Los registros pueden ser de longitud variable, y Si la cinta contiene un gran número de registros muy cortos, cada uno separado por medio de una separación entre registros, es posible que más de la mitad de la cinta esté en blanco y que se interrumpa constantemente el movimiento de la cinta. Para evitar esta situación ineficiente, es posible unir varios registros cortos para formar un **bloque de cinta** (véase la fig. 1.2).

## 1.5.4 ACCESO A LOS DATOS EN CINTA MAGNETICA.

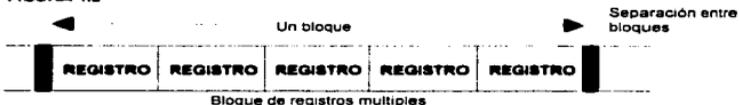
Para tener acceso a los datos almacenados en un cinta magnética fuera de línea, el operador de la computadora debe, en primer término, localizarla y montarla en una unidad de cinta ). A continuación pasa por debajo de las cabezas de lectura/escritura de la unidad en donde estas pueden leer los datos de una cinta ó grabar en ella.

FIGURA 1.1



Los primeros pies de la cinta se dejan en blanco para permitir que se enrollada en un equipo. Una marca reflejante conocida como punto de carga marca reflexiva, indica al mecanismo el comienzo de la cinta utilizable mientras que una marca similar de fin de carrete indica el fin de la cinta utilizable. Estas marcas se colocan en los bordes opuestos de la cinta para que la máquina pueda identificarlas. Entre la marca del punto de carga el primer registro de datos se encuentra una etiqueta de encabezado de control, que identifica el contenido de la cinta, el tipo de formato en que se encuentra grabado y proporciona el número de programa que se debe utilizar cuando se procesa la cinta, así como otros datos de control que ayudan a evitar que se borre por accidente una cinta importante. Después del último registro de un archivo se encuentra una etiqueta final de control que puede contener la cinta del número de registros en el archivo. Puede compararse el número de registros procesados y el número grabado en esta etiqueta para verificar que no faltó alguno. Es posible que un registro especial de un solo carácter marque el final de un archivo. Este carácter especial se llama marca de cinta.

FIGURA 1.2



Es común combinar varios registros cortos para formar un bloque el cual el procesador lee como una sola unidad para ahorrar espacio en la cinta y acelerar la entrada de datos. El programa de instrucciones separa los registros que forman un bloque para procesarlos.

## 1.5.5 VENTAJAS Y LIMITACIONES DE LA CINTA MAGNETICA

El almacenamiento en cinta magnética puede tener las siguientes ventajas:

- **Alta densidad de datos.** Un carrete común de cinta de 10 1/2 pulgadas de diámetro, tiene una longitud de 2400 pies y puede grabar 800,1600,6250, caracteres en cada pulgada, dependiendo de la unidad de cinta que utilice por tanto, si se graba 6250 caracteres en cada pulgada de cinta y si la cinta tiene una longitud de 28 800 pulgadas (2400 pies por 12 pulgadas), entonces la capacidad máxima de la cinta será de 180 millones de caracteres. Además, como ya se vio, los cartuchos modernos de 18 pistas pueden lograr una densidad de datos seis veces mayor.
- **Bajo costo y facilidad de manejo.** Un carrete de cinta 10 1/2 pulgadas cuesta menos de 20 dólares. Un paquete de discos removibles con capacidad de almacenamiento similar puede costar 25 veces más. Y dado que el carrete es compacto y pesa menos de un kilogramo y medio, ocupa muy poco espacio de almacenamiento y es fácil de manejar. Sin embargo, el uso de cinta magnética para almacenamiento secundario tiene algunas limitaciones, entre las que pueden mencionarse:

---

• **Falta de acceso directo a los registros.** La cinta magnética es un medio de almacenamiento secuencial que se utiliza para el procesamiento por lotes. Es necesario leer y procesar toda la cinta para actualizar los registros de un archivo organizados en forma secuencial. Si se requiere acceso frecuente a los registros del archivo en forma rápida y aleatoria, el archivo no deberá almacenarse en cinta magnética. Se requería demasiado tiempo de operador para montar y desmontar cintas, y desperdiciaría mucho tiempo de máquina en la lectura de registros que no necesitan.

• **Problemas ambientales.** Las partículas de polvo y la falta de control de los niveles de temperatura o humedad pueden causar errores en lectura de cintas y sus envases deben etiquetarse y controlarse cuidadosamente para no borrar por error un archivo importante.<sup>2</sup>

---

# **CAPITULO 2**

---

---

## **2 IMPORTANCIA DE LA BASE DE DATOS**

A lo largo de la historia de la informática el técnico ha tratado de reflejar el mundo real, con sus objetivos, propiedades, características y relaciones entre ellos, por medio de diferentes estructuras, más o menos sofisticadas, que hacían imposible su tratamiento..

A lo largo de esta historia las estructuras de datos, al igual que otros elementos del mundo informático como los lenguajes de programación , han ido acercándose al lenguaje natural o del usuario y esa será la tendencia general en los próximos años.

Al principio, los datos eran meros parámetros dados al programa para su ejecución y obtención de los resultados, estaban incluidos en el propio código fuente de los programas, por lo que si cambian los datos había que modificar el programa y recompilarlo.

Posteriormente, en lo que podríamos llamar época de los archivos, los datos salen de los programas y se agrupan en registros que se almacenan en dispositivos generalmente magnéticos ( cintas o discos ) para su tratamiento. El acceso a estos registros podía ser:

### **ACCESO SECUENCIAL:**

La lectura de los registros se realiza de forma consecutiva desde el inicio del archivo, por lo que para obtener un registro determinado de un fichero hay que leer todos los registros que le preceden. Es el acceso adecuado para almacenamiento en un soporte sin necesidad de acceso directo (como la cinta magnética ) y para ficheros usados para tratamiento masivo de registros.

### **ACCESO DIRECTO:**

En este caso el dispositivo sobre el que se almacenan los registros deberá tener la posibilidad de acceso o posicionamiento directo en una determinada posición o dirección dentro del fichero.

Para la localización de un registro es necesario asignarle un campo o clave que identifique el registro. A partir de esta clave hay dos modos de calcular la dirección donde se encuentra el registro:

a) Mediante la aplicación de un algoritmo sobre la clave que de una dirección dentro del fichero. Si al aplicar el algoritmo sobre las claves de registros diferentes obtuviésemos la misma dirección , dichos registros se considerarán como sinónimos. El almacenamiento de sinónimos se

---

realiza en una zona final del fichero llamada zona de sobrecarga en la que la localización de los registros es secuencial.

b) Mediante la utilización de diferentes niveles de índices que ofrecen la posibilidad de localizar la situación del registro a través del seguimiento de los mismos.

Si los ficheros permiten un acceso directo mediante índices y también un acceso secuencial ordenado a través del campo clave, tendremos un fichero llamado secuencial indexado.

## **2.1 ENFOQUE BASE DE DATOS**

La informática fué evolucionando y las necesidades de los usuarios fueron cada vez mayores. A los tratamiento masivos de archivos en los centros de cálculo sucedieron aplicaciones explotadas en tiempo real desde terminales conectadas al ordenador central a través de diferentes medios de comunicación.

Los tiempos de respuesta exigidos y el acceso concurrente de varios usuarios a un mismo fichero para su actualización o consulta, hicieron que las estructuras de ficheros resultaran insuficientes.

Cuestiones como la integridad de los archivos, sistemas de seguridad ante caídas de tensión, averías del ordenador ,etc. Son de gran preocupación y la imposibilidad de que todo programador se ocupe de estos temas en cada programa hace que aparezcan los primeros sistemas de gestión de bases de datos ( SGBD o DBMS, del inglés Data Base Management System ).En este caso, el sistema se ocupa de controlar los accesos a la base de datos, de gestionar la posible concurrencia de varios usuarios a la misma, de realizar las correspondientes operaciones para la seguridad e integridad de los datos y , en resumen, una serie de trabajos que resultan transparentes al programador y que resuelven las necesidades planteadas por los usuarios.

### **2.1.1 CONCEPTOS BASICOS**

#### **DATOS**

- Los datos son una representación del mundo real. Únicamente identificarán la porción que ejemplifique un hecho particular de interés para una o más aplicaciones.

---

## **INFORMACION**

- Es el resultado obtenido de la transformación de los datos; para una persona dada, con un fin determinado, o para satisfacer los requerimientos específicos. Este punto de vista es muy relativo, visto que, lo que son " datos procesados " para una persona, lo son en crudo para otra. Los mismos datos son objeto de diferentes clases de procesamiento en distintas circunstancias.

## **ENTIDAD**

- Una entidad es una persona, un lugar, una cosa, un evento o un concepto acerca del cual se desea registrar información. Ejemplos de entidades son clientes, proveedores, inventarios, empleados, etc.

## **ATRIBUTOS**

- Son las características básicas que identifican a la entidad ( campo de datos ).

## **VALORES DE LOS DATOS**

- El valor de los datos es la información o los datos mismos contenidos en cada campo. Estos valores pueden ser cuantitativos, cualitativos o descriptivos, dependiendo de la manera como los campos de datos ( atributos ) describan a una entidad.

## **REGISTROS**

- Un registro es una colección de datos relacionados. Así, un archivo es un conjunto de registros similares, esto es, del mismo tipo.

## **ARCHIVO DE DATOS**

- Los registros de datos forman un archivo de datos. Así, un archivo es un conjunto de registros similares, esto es, del mismo tipo.

---

- La base de datos puede definirse como una colección de datos interrelacionados, los cuales se encuentran almacenados en conjunto sin incluir datos redundantes o innecesarios; su finalidad es la de servir a una aplicación más, de la mejor manera posible; los datos se almacenan de modo que resulten independientes de los programas que los usan; se emplean métodos bien determinados para incluir datos nuevos y para modificar o extraer los datos almacenados.

La idea básica en la implementación de un base de datos es que los mismos datos sean aprovechados para tantas aplicaciones como sea posible, permitiendo no sólo la lectura de datos almacenados, sino la continua modificación de los que son necesarios para el control de las aplicaciones. Las bases de datos pueden ser diseñadas en procesamiento de lotes, en tiempo real o en línea. En ocasiones muchos datos encuentran simultáneamente guardados en varios archivos con distintas finalidades y también con diferentes fechas de actualización; en la base de datos se pretende eliminar esta redundancia; sin embargo en algunas bases de datos se admite cierta redundancia ( controlada o mínima ) con el objeto de reducir los tiempos de acceso o simplificar los métodos de direccionamiento.

## **2.1.2 CONCEPTOS ASOCIADOS A BASES DE DATOS**

### **REDUNDANCIA**

Se habla de redundancia cuando la información se repite en diferentes archivos de la misma base de datos y esto conduce a que existan datos duplicados. Por lo tanto, surge la necesidad de reducir el grado de redundancia con otras características deseables de la base, de modo que es preferible establecer esta como controlada o mínima en lugar de no redundancia como criterio de diseño.

### **INCONSISTENCIA**

Este problema surge como resultado de lo anterior, significa que al actualizar los datos almacenados en diversos archivos de la misma base de datos de una manera duplicada existe el riesgo de no realizarlo en todos los archivos y en este momento se genera la información inconsistente.<sup>3</sup>

---

<sup>3</sup> Mariela B. Vázquez " Introducción a las Bases de Datos "

---

## **INTEGRIDAD**

Son las medidas que aseguran el correcto mantenimiento los datos de la base de datos y que permiten superar las contingencias sin daño para estos. Además de protegerlos contra posibles problemas de software, deben incluirse también procedimientos de verificación que aseguren que los valores de los datos se ajustan a ciertas reglas preestablecidas. Hay diferentes maneras de asegurar la integridad de los datos:

**Validación de los datos:** el contenido de cada elemento de entrada debe coincidir con el tipo de dato descrito previamente.

**Validación del valor de los datos:** el contenido de un campo de entrada puede confirmarse para cierto rango de valores.

## **SEGURIDAD**

Para evitar que una persona no autorizada tenga acceso al sistema, es común el uso de claves de acceso. Existen varios aspectos de seguridad relacionados específicamente con los datos, como son impedir que vean los datos aquellas personas no autorizadas y evitar que modifiquen o consulten información que no deba alterarse.

## **CONCURRENCIA**

Se dice que hay concurrencia cuando varios elementos u ocurrencias intentan acceder al mismo tiempo a un proceso. Por ejemplo, en un sistema de usuarios múltiples como un sistema de reservaciones de línea aérea, cuando dos vendedores intentan actualizar en forma simultánea el mismo registro, interfieren entre sí, de manera que una de esas dos actualizaciones puede no tomarse en cuenta.

### **2.1.3 NECESIDAD DE LA BASES DE DATOS**

Considerando el manejo de archivos para un grupo de aplicaciones, sin el uso de una base de datos, cada programa de aplicación es considerado como un caso separado y distinto; en consecuencia, no hay coordinación entre los programas de aplicación del mismo grupo. Por lo general dos programas de aplicación necesitan recurrir a un mismo archivo y no solamente el

---

archivo puede tener multiples usuarios, sino que además cada usuario puede tener diferentes necesidades con respecto a dicho archivo.

Debido a los problemas que se enfrentan los analistas y diseñadores de sistemas, surge la necesidad de hacer uso de métodos que faciliten el diseño de bases de datos, a fin de ayudar en la conceptualización física y lógica de los datos, el establecimiento de las relaciones entre las mismas bases de datos, así como a determinar sus atributos; optimizando de esta manera los recursos con que se cuenta ( espacio en disco ) y evitando la duplicación de datos ( redundancia e inconsistencia ).

## **2.1.4 VENTAJAS DE UNA BASE DE DATOS**

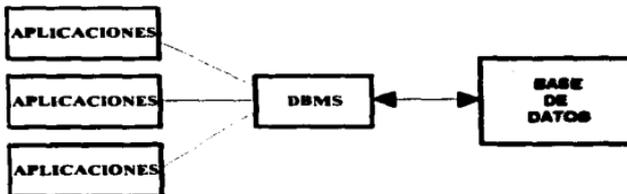
- Disminuye la redundancia.
- Evita la inconsistencia.
- Conserva la integridad de los datos.
- Permite compartir datos entre diferentes aplicaciones.
- Proporcionan independencia entre los datos y los programas.
- Facilita el desarrollo de aplicaciones.
- Reducción en el mantenimiento de programas.
- La uniformidad se encuentra centralizada.

## **2.2 SISTEMA MANEJADOR DE BASES DE DATOS ( DBMS )**

Se necesita un sistema que integre los archivos en una base de datos y que pueda proporcionar diferentes orientaciones a usuarios diferentes. El software, hardware y los procedimientos para manejar la base de datos conforman un sistema de manejo de base de datos ( DBMS, Data Base Management System ). Un sistema de manejo de la base de datos hace posible acceder datos integrados que cruzan límites operacionales, funcionales u organizativos dentro de una organización. En terminos generales, alguno de los objetivos

---

principales de DBMS son: eliminar la redundancia, mantener la consistencia de los datos, resolver los problemas de concurrencia y regular el acceso a los datos.



## 2.2.1 CARACTERISTICAS DE LOS DBMS

La administración de las bases de datos fusiona la necesidad que tiene el usuario por técnicas más complejas para la manipulación de los datos con la creciente tecnología de las computadoras. Un DBMS hace más fácil la actualización y modificaciones de las bases de datos existentes, reduce el modo substancial, el espacio necesario para almacenar y procesarlos, su uso mejora significativamente el manejo de la base de datos en comparación con los métodos más comunes de manipulación de archivos. Además:

- Proporciona independencia entre los programas de aplicación final de los usuarios y la estructura de almacenamiento física de los datos.
- Programas de utilidad para facilitar la creación, mantenimiento y reestructuración de las bases de datos.
- Facilita la reorganización de los datos
- Atiende de manera efectiva las diferentes necesidades de la organización
- Suministra información consistente y oportuna tiempo para la toma de decisiones.

- 
- **Genera seguridad de los datos y restringe el acceso a ellos.**
  - **Posibilidad de recuperar las operaciones en forma manual con esfuerzo mínimo.**

## **2.2.2 ELEMENTOS DE DMBS**

Son tres los principales elementos de un sistema de administración de base de datos, estos son:

### **LENGUAJE DE DEFINICIÓN DE DATOS ( DDL )**

El término lenguaje de definición de datos es una herramienta para el DBMS, la cual ayuda a especificar la forma como los datos pueden almacenarse; sin embargo, existen dos maneras de almacenamiento. La primera es una vista lógica de datos como la forma en que el diseñador percibe la organización, y la segunda es la vista física, refleja la manera en que realmente están almacenados los datos.

### **DICCIONARIO DE DATOS ( DD )**

En un medio de base de datos, uno de los principales objetivos se refiere a que muchos usuarios compartan datos comunes y que estos sean correctos. Para realizar los objetivos de tener datos correctos, redundancia mínima y control del uso de los datos, es necesario poseer un mecanismo de control, para ello se puede pensar en un diccionario de datos. Un diccionario de datos es un depósito que contiene información acerca de los datos de la base de datos.

El diccionario de datos almacena información sobre los datos relativos al origen de éstos, descripción, relación con otros datos, uso, responsabilidad y formato. Uno de los objetivos básico del diccionario de datos es permitir el manejo y la documentación de los datos.

### **LENGUAJE MANIPULADOR DE DATOS ( DML )**

---

El término de manipulación de datos se refiere en forma amplia al mecanismo empleado para recuperar datos almacenados en una base de datos, así como a la inserción de nuevos datos, la supresión y modificación de los que se encuentran almacenados.

El lenguaje de manipulación de datos es aquél que capacita a los usuarios a acceder o manipular los datos según estén organizados por el modelo adecuado. Existen dos maneras básicas de hacerlo; una de ella consiste en hacer que un programa de aplicación emita una instrucción al DBMS para que encuentre ciertos datos en la base y los devuelva al programa, a estas instrucciones se les denomina proposiciones integradas. La otra forma de tener acceso a ellos es cuando la persona que los busca desde una terminal, emita un comando especial de manera directa al DBMS para que los encuentre y devuelva a la pantalla, a esto se le conoce como operación de consulta.

Pasaremos ahora a ver los diferentes tipos de sistemas de gestión de bases de datos, de ellos tres claramente diferenciados surgidos a lo largo del tiempo:

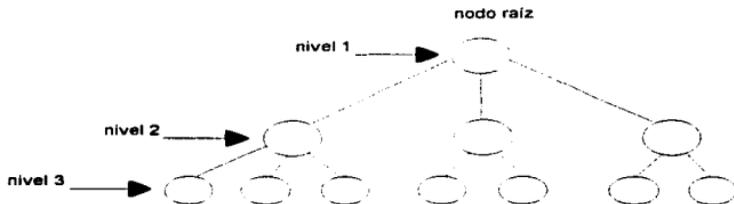
- BASES DE DATOS JERÁRQUICO.
- BASES DE DATOS EN RED.
- BASES DE DATOS RELACIONAL.

## 2.3 BASES DE DATOS JERARQUICAS

La estructura jerárquica de árbol se constituye con nodos y ramas. Un nodo es una colección de atributos de datos que describen a la entidad de ese nodo. El nodo más alto de una estructura jerárquica de árbol se reconoce como **raíz**. Los nodos dependientes se encuentran en los niveles más bajos en el árbol. El nivel de estos nodos depende de la distancia del nodo raíz ( véase la fig. 2.1 ). ( El árbol jerárquico se encuentra en un orden inverso si se le compara con un árbol real, el cual tiene su raíz en el fondo y sus ramas creciendo hacia arriba ).<sup>4</sup>

---

<sup>4</sup> Atre Shakuntala " Técnica de Bases de Datos "



**Fig 2.1**

Una estructura jerárquica de árbol está constituida por nodos o ramas.  
 Un nodo es una colección de atributos de datos que describen a la entidad en ese nodo.  
 El nodo más alto de una estructura jerárquica de árbol se llama el nodo "nodo raíz" ( tipo de entidad dominante ).  
 El nodo raíz está en el primer nivel. Los nodos dependientes ( tipos de entidad subordinada ) están en el 2 y 3 nivel

Un modelo jerárquico es un modelo que organiza a los datos en una estructura jerárquica de árbol. Cada vez que aparece el nodo raíz da principio a un registro lógico de base de datos, esto es una base de datos jerárquica está constituida por varios árboles. En el modelo jerárquico, los nodos que encuentran en el nivel 2 se conocen como hijos del nodo que se encuentran en el nivel 1, y el nodo del nivel 1 se conoce como padre de los nodos en el nivel 2. Los nodos del nivel 3, que corresponden a un nodo en el nivel 2 se conoce como padre, y así sucesivamente.

Una estructura jerárquica de árbol tiene que satisfacer las condiciones siguientes:

- 1) Un modelo jerárquico siempre empieza con el nodo raíz.
- 2) Cada nodo consta de uno o más atributos que describen a las entidades de ese nodo.
- 3) Los nodos dependientes pueden aparecer en dos niveles consecutivos. El nodo en el nivel precedente se convierte en el nodo padre de los nuevos nodos dependientes. Los nodos dependientes se pueden añadir tanto horizontal como verticalmente sin ningún límite.
- 4) Cada nodo que se presenta en el nivel 2 tiene que conectarse con uno y sólo un nodo que se presenta en el nivel 1. Cada nodo que se presenta en el nivel 3 tiene que conectarse a uno y sólo un nodo que se presenta en el nivel 2, y así sucesivamente. Debido a que como máximo puede haber sólo una conexión ( unión ) entre dos nodos cualquiera.

---

5) Un nodo padre puede tener uno a varios nodos hijos bajo su dependencia. Si no tiene ningún nodo bajo su dependencia no es nodo padre.

6) Cada nodo, excepto la raíz, tiene que accesarse a través de su nodo padre. El nodo representado en una verdadera jerarquía debe recuperarse sólo a través de su padre, ya que en éste radica el verdadero significado de la existencia de esos datos.

7) Un nodo aparece varias veces a cada nivel. Cada vez que aparece un nodo ( excepto cuando aparece un nodo raíz ) se tiene que conectar con un nodo padre, esto es, puede presentarse varias veces el nodo A. Cada vez que se presenta el nodo A da principio a un modelo lógico. Puede no aparecer una o muchas veces el nodo B por cada vez que se aparezca el nodo A, y así sucesivamente.

### **2.3.1 VENTAJAS DEL MODELO JERARQUICO**

- La ventaja mayor del modelo jerárquico es la existencia de sistemas de manejo de base de datos probados que se usan en el modelo jerárquico como estructura básica.
- La relativa simplicidad y facilidad de uso del modelo jerárquico y la familiaridad de los usuarios del procesamiento de datos con la jerarquía, son ventajas importantes.
- Existe una reducción de la dependencia de los datos. Cuando se usan los dos modelos jerárquicos.
- La predicción del funcionamiento se simplifica a través de relaciones predefinidas.

### **2.3.2 LIMITACIONES DE UN MODELO JERARQUICO**

- Las relaciones varios a varios no pueden implementarse de una manera eficiente. Esto puede traer como consecuencia redundancia en los datos almacenados. Sabemos que en nivel lógico, la redundancia no necesariamente es mala, por el contrario, promueve la simplicidad. Sin embargo, a nivel físico la redundancia es indeseable.
- Como resultado del estricto ordenamiento jerárquico, las operaciones conocidas como de inserción y de supresión se vuelven extremadamente complejas.

- 
- La eliminación de padres trae como consecuencia la eliminación de hijos. Como resultado de esto, los usuarios deben tener cuidado cuando tengan las necesidades de poner en funcionamiento la operación.
  - Los comandos jerárquicos tienden a ser de procedimiento, debido a lo estricto de la estructura.
  - La raíz es el tipo de nodo dominante. Cualquier nodo hijo es accesible solamente a través de su nodo padre.

## **2.4 BASE DE DATOS EN RED.**

Otro enfoque de los sistemas de manejo de bases de datos se fundamenta en estructuras denominadas redes. Durante la década de los 70's se desarrollaron varios DBMS orientados a redes bajo la dirección de la Conference on Data System Languages (CODASYL). El modelo de red interconecta las entidades de una empresa en una red, en donde utiliza bloques y flechas.

Un bloque representa una entidad o un tipo de registro. Cada tipo de registro está compuesto de cero, uno o más atributos ( también llamados campos de datos o campos ). Una flecha dirigida conecta dos o más tipos de registros y se utiliza para representar un conjunto de tipos de registros. El tipo de registro localizado en la cola de la flecha, funge como un tipo de registro propietario, y el registro localizado en la cabeza de la flecha funge como tipo de registro miembro. La flecha que va del propietario al miembro se le llama conjunto de tipos. Un conjunto de tipos muestra una relación lógica uno a varios, entre un propietario y un miembro.

La Estructura de un modelo de red debe de cumplir con las siguientes características:

- Un conjunto de tipos es un agrupamiento de registros relacionados que posee un nombre.
- Sólo existe una ocurrencia " propietario " en una ocurrencia del conjunto.
- Puede haber cero, una o varias ocurrencia " miembro " en una ocurrencia de conjunto.

- 
- Se dice que un conjunto de tipos está " vacío " cuando ninguna ocurrencia de registro miembro está asociada con una de registro propietario.
  - Una ocurrencia del conjunto existe cuando se almacena un registro.

### **2.4.1 VENTAJAS DE UN MODELO DE RED**

- La principal ventaja de un modelo de red consiste en que, como para el modelo jerárquico , existen sistemas de manejo de datos que utilizan, con éxito, el modelo de red como estructura básica.
- La relación varios a varios, que ocurre con frecuencia en la vida real, se puede implantar fácilmente.
- El modelo de red esta respaldado por el grupo de Tares de Bases de Datos ( DBTG ) de CODASYL ( Conferencia Sobre Lenguajes de Sistemas de Datos).

### **2.4.2 LIMITACIONES DE UN MODELO DE RED**

- La principal limitación del modelo de red es su complejidad. El programador debe estar familiarizado con las estructuras lógicas de la base de datos, ya que en ella tiene que navegar a través de diferentes ocurrencias de conjunto con la ayuda de las ocurrencias del tipo de registro del conector.
- Cuando la base de datos se reorganiza, es posible, a menos que se tenga mucho cuidado, que se pierda independencia de los datos. Esto ocurre cuando los conjuntos se eliminan. Asimismo, debido a que las observaciones proporcionadas a los programadores de aplicaciones son menos simples que los enfoques basados en el modelo jerárquico, la programación se vuelve más compleja. La Fig 2.2 muestra el esquema de una base de datos en red.

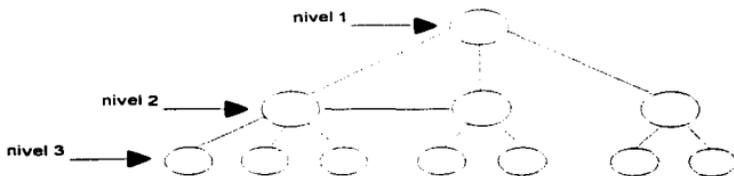


Fig. 2.2 Esquema de una base de datos en red

## 2.5 BASES DE DATOS RELACIONALES

Las bases de datos relacionales aparecieron a finales de los 70's y ofrecían varias ventajas sobre el modelo jerárquico. Los métodos para manipular los datos en las tablas son bien conocidos y razonablemente eficaces. El álgebra relacional proporciona una base matemática formal capaz de soportar la metodología de análisis de datos, con la posibilidad de unir dinámicamente los datos en las tablas virtuales durante la ejecución, las bases de datos relacionales facilitaron el desarrollo de aplicaciones con un nivel más alto de independencia de datos que sus predecesoras.

Sin embargo, el modelo relacional sufre al menos de un inconveniente notable.

Es difícil expresar la semántica de objetos complejos con un sólo modelo de tabla para el almacenamiento de datos. Aunque las bases de datos relacionales son adecuadas para contabilidad u otras aplicaciones típicas de procesamiento de transacciones, en las que los tipos de datos son simples y escasos en número, el modelo relacional ofrece ayuda limitada cuando los tipos se vuelven numerosos y complejos.

Al igual que el paso de jerárquica a relacional fue dado para reducir la atención hacia la estructura, también lo hizo así el paso de relacional a orientada a objetos. Las nuevas aplicaciones de los años 90 requieren soporte para estructuras de datos extensibles y complejas, acceso a datos complejos y alto rendimiento. Los conceptos tradicionales de bases de datos se está ampliando gradualmente para satisfacer las necesidades de los usuarios. La evolución de la tecnología de bases de datos durante los últimos años, representa este giro hacia la gestión de bases de datos más compleja.

---

A continuación se dará una explicación más detallada de las características y componentes de una base de datos relacional, ya que servirá de modelo para la creación de la base de datos del proyecto.

## 2.5.1 CONCEPTOS FUNDAMENTALES

El modelo relacional se distingue por la reducción de los datos estructurales almacenados en planas o **tablas** con filas y columnas. A estas tablas se les denomina **relaciones** y equivalen cada una de ellas a lo que en terminología tradicional se conocía como archivo.

Hay una correspondencia entre los términos utilizados al hablar archivos, tablas y relaciones. Así, si un archivo corresponde a una tabla, cada registro del archivo se corresponde con una fila y cada campo dentro del registro con una columna.

En terminología relacional cada fila de la tabla se conoce por **tupla** y cada columna por **atributo**.

La cardinalidad de una relación será el número de tuplas que la componen y grado el número de atributos o columnas de la misma.

Las relaciones son tablas con las siguientes propiedades:

- No hay tuplas iguales.
- El orden de las tuplas no es significativo.
- El orden de los atributos no es significativo.
- En una tupla determinada, cada atributo puede tomar sólo un valor ( no se admiten grupos repetitivos ).

## TABLAS

El principio de la organización de una base de datos es una tabla, es decir una disposición rectangular fila-columna de los valores de los datos. Cada tabla de una base de datos tiene un nombre de tabla único que identifica sus contenidos. En realidad, cada usuario puede definir sus propios nombres de tablas sin preocuparse acerca de los nombres elegidos por otros usuarios.

También es importante conocer que el conjunto de valores que una columna puede contener se denomina dominio de la columna.

---

Las columnas de una tabla tiene un orden de izquierda a derecha, que se define cuando la tabla se crea por primera vez. Una tabla simple tiene al menos una columna. El estándar SQL ( Structure Query Language ( lenguaje de Estructuras y Consultas ) ANSI-ISO ( American National Standards Institute, Instituto Nacional Estadounidense de Estándares) no especifica un número máximo de columnas de una tabla, pero casi todos los productores comerciales SQL imponen un limite que generalmente es de 255 columnas por tabla o mas. A diferencia de las columnas, las filas de una tabla no tienen orden particular.

De hecho si se utilizan dos consultas de base de datos consecutivos para visualizar el contenido de una tabla no hay garantía que las filas sean listadas dos veces en el mismo orden. Naturalmente se puede solicitar a SQL que odene las filas antes de visualizarlas, pero el orden de clasificación no tiene nada que ver con la disposición efectiva de las filas dentro de la tabla.

Una tabla puede tener cualquier número de filas. Una tabla de cero filas es perfectamente válido y se le denomina una tabla vacía ( por razones obvias). Un tabla vacía sigue teniendo una estructura, impuesta por sus columnas; simplemente no tiene datos. El estándar ANSI-ISO no limita el número de filas de una tabla, y muchos productos de SQL permiten que una tabla crezca hasta que se agote el espacio en disco disponible en el computador. Otros productos SQL imponen un limite máximo, pero este es siempre muy generoso dos millones de filas o mas en común.

## CLAVES PRIMARIAS

Puesto que las filas de una tabla relacional no están ordenadas, no se puede relacionar una fila específica por su posición en la tabla. No hay **primera fila** y **última fila**, o decimatercera fila de una tabla.

En una base de datos relacional bien diseñada cada tabla tiene una columna o combinación de columnas cuyos valores identifican unívocamente cada fila en la tabla. Esta columna o columnas se denominan clave primaria de la tabla.

La clave primaria tiene un valor único diferente para cada registro, de modo que no hay dos filas de una tabla con una clave primaria que sean duplicados exactos la una de otra. Una tabla en donde cada fila es diferente de todas se llama relación en términos matemáticos. El nombre de base de datos relacional proviene de estos terminos ya que las relaciones ( las tablas con filas distintas) son el corazón de una base de datos relacional.<sup>3</sup>

---

<sup>3</sup> Ollison L. Mark " Database Step By Step "

---

## RELACIONES

Una de las principales diferencias entre el modelo relacional y los demás datos primitivos es que los punteros explícitos, tales como las relaciones padre-hijo de una base de datos jerárquica, están prohibidas en las bases de datos relacionales. Obviamente estas relaciones siguen existiendo en una base de datos relacional. La relación padre/hijo entre una oficina de ventas y la gente que trabaja en ella no se pierde en el modelo relacional, tan solo no se presenta por un puntero explícito almacenado en la base de datos. En vez de ello, la relación está representada por valores de datos comunes almacenados en dos tablas de una base de datos. Todas las relaciones de una base de datos relacional están representadas de este modo.

## CLAVES FORÁNEAS

Una columna de una tabla cuyo valor coincide con la clave primaria de alguna otra tabla se denomina una clave foránea. Juntas una clave primaria y una clave foránea crean una relación padre-hijo entre las tablas que las contienen, del mismo modo que las relaciones padre-hijo de una base de datos jerárquica.

Lo mismo con una combinación de columnas puede servir como clave primaria de una tabla, una clave foránea puede ser también una combinación de columnas. De hecho, la clave foránea será siempre una clave compuesta ( multicolumna) cuando se referencia a una tabla con una clave primaria compuesta.

Obviamente el número de columnas y los tipos de datos de las columnas en la clave foránea y en la clave primaria deben ser idénticos unos de otros.

Las claves foráneas son parte fundamental del modelo relacional ya que se crean relaciones entre tablas en la base de datos.

### 2.5.2 LAS 12 REGLAS DE CODD.

En su artículo de 1985 en computerworld, Tedd Codd presentó doce reglas que una base de datos debe obedecer para que sea considerada verdaderamente relacional. Las reglas se derivan del trabajo teórico de Codd sobre el modelo relacional, y representan realmente más un objetivo ideal que una definición de una base de datos relacional.

---

Ningún DBMS relacional actualmente disponible satisface totalmente las doce reglas de Codd. De hecho, se esta convirtiendo en una práctica popular elaborar tarjetas de tanteo para productos DBMS comerciales, que muestran lo bien o mal que estos satisfacen cada una de las reglas. Desgraciadamente, las reglas son subjetivas.

- **La regla 1** es básicamente la definición informal de una base de datos relacional. Por que expresa toda la información de una base de datos relacional esta representada explícitamente a nivel lógico y exactamente mediante valores en tablas.

- **La regla 2** refuerza la importancia de las claves primarias para localizar datos en la base de datos. El nombre de la tabla localiza la tabla correcta, el nombre de la columna encuentra la columna correcta y el valor de la clave primaria encuentra la fila que contiene un datos individual de interés.

- **La regla 3** requiere soporte para la falta de datos mediante el uso de valores NULL. Esto quiere decir que los valores nulos distintos de la cadena de caracteres vacía o de una cadena de caracteres en blanco y distinto de cero de cualquier otro número se soporta en los DBMS completamente relacionales. para representar la falta de información inaplicable de un modo sistemático e independiente al tipo de dato.

- **La regla 4** requiere que la base de datos relacional sea autodescriptiva. En otras palabras, la base de datos debe contener ciertas tablas de sistema cuyas columnas describan la estructura de la propia base de datos.

- **La regla 5** ordena la utilización de un lenguaje de base de datos relacional, tal como SQL, aunque no se requiera específicamente SQL. El lenguaje debe de ser capaz de soportar todas las funciones básicas de un DBMS creación de un base de datos, recuperación y entrada de datos, implementación de la seguridad de la base de datos.

- **La regla 6** trata de las vistas, que son tablas virtuales utilizadas para dar a diferentes usuarios de una base de datos, diferentes vistas de su estructura. Es una de las reglas mas difíciles de implementar en la práctica, y ningún producto comercial la satisface totalmente hoy en día.

- **La regla 7** refuerza la naturaleza orientada a conjuntos de base de datos relacional, requiere que las filas sean tratadas como conjuntos en operaciones de inserción eliminación y actualización. La regla esta diseñada para prohibir implementaciones que sólo soportan la modificación o recorrido fila a fila de la base de datos.

- **La regla 8 y la regla 9** aíslan al usuario o al programa de aplicación de la implementación de bajo nivel de la base de datos. Especifican que las técnicas de acceso a almacenamiento especificadas y utilizadas por el DBMS, e incluso los cambios a la estructura de los tablas en la base de datos, no deberían afectar a la capacidad del usuario de trabajar con los datos.

---

• **La regla 10** dice que el lenguaje de base de datos debería soportar las restricciones de integridad que restringen los datos que pueden ser introducidos en la base de datos y las modificaciones que pueden ser efectuadas en esta. Esta es otra de las reglas que no soportan la mayoría de los productos comerciales DBMS.

• **La regla 11** dice que el lenguaje de base de datos debe ser capaz de manipular datos distribuidos localizados en otros sistemas informáticos; Es decir que un DBMS relacional tiene independencia de distribución.

• **La regla 12** impide otros caminos en la base de datos que pudieran sustituir su estructura, su estructura relacional y su integridad; más explícitamente podemos decir que si un sistema relacional tiene un lenguaje de bajo nivel ( un sólo registro cada vez ), ese bajo nivel no puede ser utilizado para sustituir o suprimir las reglas de integridad y las restricciones expresadas en el lenguaje relacional de nivel superior ( múltiples registros a la vez ) .

SQL esta basado en el modelo relacional de datos que organiza los datos en una base de datos como una colección de tablas.

- Cada tabla tiene un nombre que la identifica unívocamente.
- Cada tabla tiene una o más columnas nominadas, que estan dispuestas en un orden específico de izquierda a derecha.
- Cada tabla tiene cero o más filas, conteniendo cada una un único valor en cada columna. Las filas estan desordenadas.
- Todos los valores de una columna determinada tienen el mismo tipo de datos, y estos estan extraídos de un conjunto de valores legales llamado el dominio de la columna.

Las tablas están relacionadas unas con otras por los datos que contienen . El modelo de datos relacional utiliza claves primarias y claves foráneas para representar estas relaciones entre tablas:

- Una clave primaria es una columna o combinación de columnas dentro de una tabla cuyo (s) valor(s) identifica(n) unívocamente a cada fila de la tabla.
- Una tabla tiene una única clave primaria.
- Una clave foránea es una columna o combinación de columnas en una tabla cuyo(s) valor(es) es(son) un valor de clave primaria para alguna tabla.
- Una tabla puede contener mas de una clave foránea, enlazandola a una o más tablas.

- 
- Una combinación clave primaria/clave foránea crea una relación padre/hijo entre las tablas que la contienen.

## **NOMBRES DE TABLA**

Quando se especifica un nombre de tabla en una sentencia SQL, se presume que se esta refiriendo a una de las tablas propias ( es decir un tabla ya creada).

Con el permiso adecuado, también se puede referir a tablas propiedad de otros usuarios, utilizando un nombre de tabla cualificado. Un nombre de tabla cualificada especifica el nombre del propietario de la tabla junto con el nombre de la tabla, separados por un punto.

Un nombre de tabla cualificado puede ser utilizado generalmente dentro de una sentencia SQL en cualquier lugar que pueda aparecer un nombre de tabla.

## **NOMBRES DE COLUMNA**

Quando se especifica un nombre de columna en una sentencia SQL, se puede determinar normalmente a que columna se refiere a partir del contexto. Sin embargo, si la sentencia afecta a dos columnas con el mismo nombre correspondientes a dos tablas diferentes, debe utilizarse un nombre de columna cualificado para identificar sin ambigüedad la columna designada. Un nombre de columna cualificada especifica tanto el nombre de la tabla que contiene la columna como el nombre de la columna, separados por un punto (.). Por ejemplo, la columna de nombre VENTAS en la tabla REPVENTAS tiene el nombre de columna cualificado.

**REPVENTAS . VENTAS**

Si la columna procede de una tabla propiedad de otro usuario, se utiliza un nombre de tabla cualificado en el nombre de una columna cualificado. Por ejemplo, la columna DIA\_MES en la tabla CUMPLEAÑOS propiedad del usuario SAM se especifica mediante el nombre de columna cualificado completo:

**SAM. CUMPLEAÑOS.DIA.MES**

Los nombres de columna cualificados pueden ser utilizados generalmente en una sentencia SQL en cualquier lugar en el que pueda aparecer un nombre de columna simple ( no cualificado ); las excepciones se indican en las descripciones de las sentencias SQL individuales.

---

## TIPOS DE DATOS

El estándar SQL ANSI/ISO especifica varios tipos de datos que pueden ser almacenados en una base de datos basada en SQL y manipulados por el lenguaje. Los tipos de datos especificados por el estándar son sólo un conjunto mínimo, pero casi todos los productos SQL comerciales, los soportan o tienen tipos de datos que son muy similares. Los tipos de datos ANSI/ISO consisten en lo siguiente:

- **Cadena de caracteres de longitud fija.** Las columnas que contienen este tipo de datos almacenan típicamente nombres de personas y empresas, direcciones, descripciones, etc.
- **Enteros.** Las columnas que contienen este tipo de datos almacenan típicamente cuentas, cantidades, edades, etc. Las columnas de valores enteros también se utilizan con frecuencia para obtener números identificadores, tales como números de cliente, de empleado y de pedido.
- **Números decimales.** Las columnas con este tipo de datos almacenan números que tienen parte fraccionaria y deben ser calculados exactamente, tales como porcentajes y tasas. También se utilizan frecuentemente para almacenar importes monetarios.
- **Números en coma flotante.** Las columnas con este tipo de datos se utilizan para almacenar números científicos que pueden ser calculados aproximadamente, tales como pesos y distancias. Los números en coma flotante pueden representar mayores rangos de valores que los números decimales, pero pueden producir errores de redondeo en los cálculos.

### Tipos de datos extendidos

La mayoría de los productos SQL comerciales ofrecen un conjunto mucho más extenso de tipos de datos que lo especificados en el estándar ANSI/ISO.

Algunos de los tipos de datos extendidos más populares o importantes son:

- **Cadena de caracteres de longitud variable.** casi todos los productos SQL soportan los datos VARCHAR, que permiten que una columna almacene cadenas de caracteres que varían de longitud de una fila a otra, hasta una cierta longitud máxima. El estándar ANSI/ISO solamente especifica cadenas de longitud fija, que son rellenas por la derecha con caracteres en blanco adicionales.
- **Importes monetarios.** Muchos productos SQL soportan un tipo MONEY o CURRENCY, que se almacena generalmente como número decimal o en coma flotante. El tener un tipo monetario distinto permite al DBMS formatear adecuadamente los importes monetarios cuando son visualizados.

- 
- Fechas y horas. Soportar valores para fecha y horas es también habitual en los productos SQL, aunque los detalles varían ampliamente de un producto a otro. Generalmente se soportan variadas combinaciones de fechas, horas, instantes, intervalos de tiempo y aritmética de fecha y hora.
  - Datos Boleanos. Algunos productos SQL (incluyendo dBASE IV) soportan los valores lógicos ( True o False ) como un tipo explícito.
  - Texto extenso. Varias bases de datos basadas en SQL soportan columnas que almacenan cadenas de textos extensas ( típicamente de hasta 32000 o 65000 caracteres, o entre 10 a 20 páginas escritas a un solo espacio). El DBMS restringe generalmente el uso de estas columnas en consultas y búsquedas interactivas.
  - Flujo de bytes no estructurados. Oracle y otros varios productos permiten almacenar y recuperar secuencias de bytes de longitud variable sin estructurar. Las columnas que contienen estos datos se utilizan para almacenar imágenes de video comprimidas, código ejecutable y otros tipos de datos sin estructurar.
  - Caracteres asiáticos. DB2 soporta cadenas de longitud fija y de longitud variable de caracteres de 16 bits utilizados para representar caracteres Kanji y otros caracteres asiáticos. Sin embargo, la búsqueda y ordenación de estos tipos GRAPHIC y VARGRAPHIC no está permitida.

## 2.5.3 VENTAJAS DE UN MODELO RELACIONAL

**SIMPLICIDAD**. El usuario final se representa con un modelo de datos. Sus solicitudes se formulan en términos de contenido de la información y no refleja la complejidad de los aspectos relacionados con el sistema. Un modelo relacional es lo que el usuario ve, y no necesariamente lo que se implementará físicamente.

**CONSULTAS NO PLANTEADAS**. Debido a que no hay una dependencia de posición entre las relaciones, las consultas no tienen que reflejar ninguna estructura preferida y por lo tanto pueden ser de tipo procedimientos no estándar de consulta.

**INDEPENDENCIA DE LOS DATOS**. Esto debe constituir uno de los principales objetivos de cualquier sistema de manejo de base de datos. El modelo relacional elimina detalles relativos a la estructura del almacenamiento y la estrategia de acceso de la interfase con el usuario. Este

---

modelo proporciona un grado de independencia de los datos relativos más alta que los dos modelos descritos anteriormente. Sin embargo, para poder usar esta propiedad del modelo relacional, el diseño de las relaciones debe ser exacto y completo.

**FUNDAMENTOS TEORICOS.** El modelo relacional está basado en la bien desarrollada teoría matemática de las relaciones. El riguroso método del diseño de la base de datos da a este modelo un fundamento sólido. Esta clase de fundamentación no existe para otros modelos de datos.

## **2.5.4 LIMITACIONES DEL MODELO RELACIONAL**

Aunque en la actualidad ya existen algunos sistemas de manejo de base de datos basados en los modelos relacionales disponibles comercialmente, su funcionamiento no ha sido comparado con el de un DBMS versión de claves múltiples. Como consecuencia, la principal cuestión por atender se refiere al funcionamiento. ¿ puede utilizarse un modelo relacional para un DBMS que pueda proporcionar un conjunto completo de capacidades operativas con la eficiencia requerida en gran escala ? Actualmente parece que esta cuestión puede ser atendida positivamente tomando en cuenta los adelantos tecnológicos en cuanto al aprovisionamiento de un mejor y más confiable software.

---

# **CAPITULO 3**

---

---

### **3.1 EL DESARROLLO DE SISTEMAS INFORMÁTICOS**

En la actualidad el sentido de desarrollo de los sistemas informáticos ha cambiado, debido a que ahora existen ya diversos paquetes estándar y se comercializan en las empresas, dichos paquetes atienden a las necesidades más comunes, como la automatización de oficinas (proveedores de palabras, hojas de cálculo, correos electrónicos, organizadores, etc. ) y en la labor cotidiana de los sistemas administrativos básicos ( contabilidad, nómina, crédito, cuentas por cobrar, cuentas por pagar, inventarios, etc. ) así como adelantos tecnológicos en cuanto al hardware y comunicación entre computadores.

Ahora, a los especialistas de sistemas y cómputo les compete una responsabilidad mayor: la de seleccionar, justificar, adecuar e implantar la mejor tecnología en hardware y software en la industria y cualquier entidad.

El cambio que se refiere al punto anterior es que ahora se desarrollan sistemas de mayor magnitud que los de hace cinco años, y son comercializados por empresas dedicadas a ese negocio. Dentro de las empresas también se desarrollan proyectos de gran magnitud (comunmente demanda una alta inversión ) y entonces la visión no es de construir sino de seleccionar, implantar y controlar nuevos sistemas.

Por otro lado, existe la dificultad de que los paquetes ya hechos, no se ajustan en todo a las necesidades de la empresa, ya sea por que no alcanzarán a cubrir la operación , o bien, empresas tienen vicios administrativos y no cambian al ritmo de la tecnología, por ejemplo:

- existen problemas con los usuarios a nivel operativo.
- no hay buenos sistemas de control y seguridad en las operaciones de los sistemas.
- comunmente se suguen manteniendo sistemas obsoletos dejando en operación métodos viejos en equipos de tecnología nueva.
- el personal de sistemas se orienta a corregir programas y no a liderear cambios en los sistemas de trabajo, y en los flujos de información.

Una solución consiste en realizar cambios y avances en lo administrativo para marchar al mismo ritmo de la tecnología; y estructurar las metodologías de trabajo y desarrollo e implantación de proyectos, ya que habría que hacer énfasis en las fases de verificación de la necesidad y su justificación, el análisis de lo que se solucionará y los mecanismos de

---

implementación; con la idea de incluir procedimientos adecuados a la nueva formula de trabajo.

Con respecto a los sistemas muy propios de cada empresa, de los cuales es difícil encontrar una aplicación generalizada, se estila realizar desarrollo internos para cubrir las necesidades especiales. En este caso, se debe utilizar en la empresa, una metodología para realizar el trabajo de desarrollo de sistemas, el trabajo de evaluación e implementación de proyectos, que permita controlar y hacer congruentes los esfuerzos de todo el personal del área de informática y con el personal usuario.

Actualmente dentro de las empresas, no existe desarrollo de sistemas grandes; por lo tanto, las tareas de análisis y programación detallada de sistemas ha disminuido, y a su vez han aumentado las tareas de administración y control de proyectos encaminados a implementar nuevas tecnología ( hardware y software, redes, etc.). Muchas veces el personal informático se dedica a hacer estudios de evaluación de paquetes de software, entrenamientos en paquetes de manufactura probada, concursos entre diferentes tipos de equipos, evaluando la calidad de servicios de comunicación, etc.; las labores de corrección o desarrollos propios, son ahora mucho menores. Por lo tanto, las metodologías también son cambiantes y ahora refuerzan más en la planeación y control de proyectos.

No existe una metodología mágica, que contenga la receta ideal para desarrollar cualquier proyecto, pero si existen bases ordenadas para el desarrollo de cualquier trabajo profesional referente a casi cualquier área de aplicación. En el caso de proyectos informáticos ambos tienen que ver con la tecnología y su aplicación para resolver problemas de la realidad, generalmente los relacionados a los flujos de información ( su producción distribución y ordenación ), que tienen que ser modelados para su simulación o bien para crear herramientas que ayuden a los trabajos de diversa índole.

Realmente, lo interesante de esto es que siempre hay una necesidad que satisfacer, un requerimiento que cubrir, de ahí siempre nace un nuevo proyecto o el mantenimiento ( adecuación, corrección o mejora ) a algunos ya existentes; por ello, la metodología tradicionalmente utilizada en proyectos informáticos guardan en su generalidad la tradicional estructura, la cual se muestra en la figura 3.1 .

Una necesidad nace de un requerimiento hecho por un usuario o un grupo de usuarios, hacia su departamento , éste es validado, aprobado, desarrollado, implantado con el usuario para

cubrir su necesidad y finalmente liberarlo, comprobándose a través del control, si la necesidad planteada fue satisfecha.

El proceso de desarrollo involucra desde la verificación real de las necesidades, la evaluación del proyecto y las fases de : planeación, análisis, diseño , construcción, pruebas, implantación, liberación y por último el mantenimiento, que nuevamente es receptor de una necesidad y se reciclan las funciones para satisfacerla.

Se presenta en la figura 3.2 el modelos propuesto para el ciclo de vida del desarrollo de proyectos que indica como se va desarrollando etapa por etapa. Se puede apreciar que cada fase de desarrollo deja un documento de una actividad realizada que sirve de base para la siguiente fase, y como punto de validación y control de actividades para el coordinador de proyectos, finalmente al liberarse empieza una etapa de mantenimiento en el cual se cierra el ciclo.

Figura 3.1. PROCESO DEL DESARROLLO DE SISTEMAS INFORMATICOS.

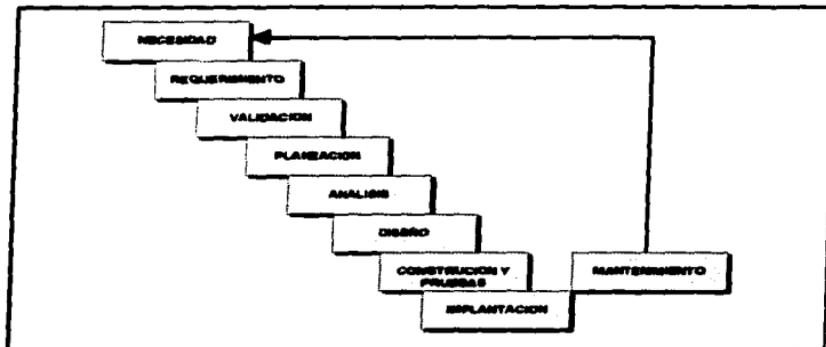
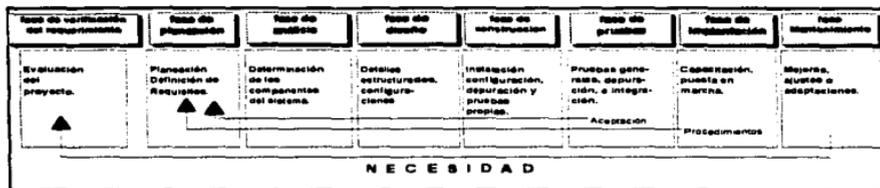


Figura.3.3 MODELO DE FASES PARA EL CICLO DE VIDA DE DESARROLLO DE SISTEMAS INFORMATICOS.



La participación del personal en el desarrollo de proyectos es muy variada, comunmente interviene mucha gente en diversas ramas profesionales ya sea interna o externa, según las necesidades. En específico un proyecto cualquiera que sea no podrá ser llevado a cabo únicamente por profesionales de una sola rama de estudios ni totalmente personal interno de una empresa. Ello debe resultar de los esfuerzos coordinados de especialistas en procesos geofísicos como del área de informática, administradores, ejecutivos y directores, etc.

Sin embargo, este esfuerzo combinado, debe de ser dirigido por un sólo individuo, capaz de guiar a los demás participantes, de anticipar los problemas rutinarios y programar los diversas fases de trabajo. Por lo tanto, en todo el proyecto es asignado a un sólo individuo, la responsabilidad global del control de proyecto, proporcionándole un trabajo de equipo interno, así como el apoyo de los directores para ejercer esa responsabilidad. Esta labor demanda que los coordinadores tengan conocimientos de informática, administración, finanzas, economía, etc. aunque no sea necesario un experto en esas ramas debe tener conocimientos suficientes para coordinar las actividades de todas ellas.<sup>6</sup>

\* Ballesteros N. Atal " Sistema de Diagnóstico "

---

Cada empresa en particular tiene una área de investigación y de mantenimiento en informática, pero para un proyecto grande y nuevo se utilizan servicios externos, pues de lo contrario se tendría que contratar personal especializado que sólo se necesitaría una sola vez. Los servicios externos se utilizarían para la realización de algunos trabajos o la contratación de algunos servicios de los cuales sean especialistas que sólo se requerirán una vez. Y por último, resulta importante la manera en que ese coordinador del proyecto administre el proyecto, es decir la metodología que se ocupará para desarrollarlo y concluirlo.

### **3.1.1 METODOLOGIA PARA EL DESARROLLO DE SISTEMAS INFORMATICOS**

Para la producción de sistemas de información son necesarios una serie de recursos más un metodo de desarrollo. Un principio fundamental en el desarrollo de sistemas, es la descomposición de un gran sistema en subsistemas de menor tamaño que sean manejables con interfaces claramente definidas. La filosofía "Divide y vencerás" suele utilizarse en disciplinas que incurren en el analisis de conceptos complejos, las unidades básicas son llamadas módulos, con interfaces de control y datos, la primera queda establecida por procedimientos de llamado entre módulos mientras que las interfaces de datos se manifiestan por medio del pase de parámetros entre módulos.

También es necesario incluir técnicas para la estimación de costos y confiabilidad de un productos, existe principios fundamentales en el analisis, diseño, construcción , pruebas, implantación y mantenimiento de productos informaticos, lo que origina metodologias para desarrollar en forma adecuada un producto. A continuacion de presentarán los pasos de la metodologia para desarrollo de los sistemas informáticos.

- a) Verificación de requerimientos.
- b) Planeación.
- c) Análisis.
- d) Diseño.
- e) Construcción.
- f) Pruebas.
- g) Implantación.
- h) Mantenimiento.

En proyectos que implican un alto costo y gran impacto dentro de la empresa, como es el caso de un proyecto de modernización , ya sea de compra de equipo, desarrollo de sistemas, etc.; es necesario dar fundamental énfasis en la verificación del requerimiento.

El método se describirá en forma general en los siguientes incisos, se presentarán sus subactividades, y el significado :

---

### **a) Verificación del requerimiento**

- **Definición conceptual del proyecto.**

Se refiere a desarrollar un enunciado definitivo del problema por resolver, incluir una descripción de la situación actual, restricciones del problema y de las metas que se lograrán. El enunciado debe hacerce ocupado terminología del usuario.

- **Evaluación del proyecto.**

se realiza una verificación de la existencia de la necesidad y se realizan estudios de capacidades, factibilidad técnica y justificación costo-beneficio del proyecto.

### **b) Planeación y control de proyecto.**

En esta fase se hacen los planes que describen los mecanismos que se ocuparán para lograr las metas y requisitos. Un logro es un hecho significativo en el ciclo de la vida del desarrollo del producto, ejemplo: Terminación del análisis de requisitos, documento de aceptación firmado por la dirección. Es necesario establecer cuántos logros son apropiados, cuándo ocurren, qué recursos necesitan, quién será el responsable de ejecutarlos qué se debe cumplir para que ocurra un logro. Las siguientes actividades componen la fase de planeación:

- **Investigación Preliminar.**

- **Estudio del requerimiento.**

- Antecedentes de la planeación
- Conocimiento de la solicitud formal.

- **Identificación del grupo de trabajo (internos y externos)**

- **Estrategia de solución.**

Se refiere al ¿Cómo?, es decir, como se pretende resolver el problema, y determinar las herramientas por fase, técnicas y notación a utilizar.

- **Desarrollo de premisas.**

**Planear actividades de administración de la configuración, control de calidad.  
Prever situaciones que afecten el desarrollo del proyecto.**

- **Proyecto del sistema.**

- **Determinar la estructura organizacional preliminar.**
- **Realinar el programa preliminar del desarrollo**
- **Realizar un estimado preliminar de recursos.**
- **Establecimiento de prioridades.**
- **Determinación de puntos de control y validación.**

- **Aceptación formal.**

- **Documentación**

---

### c) Análisis.

Se elabora estudio específico de los componentes del sistema y de la elección y participación de los proveedores, dejando claro que se necesita realizar para satisfacer los requerimientos, como obtenerlos y quién los realizará bajo las siguientes actividades:

- Estudio de necesidades.
- Definición del sistema.
- Definición de los componentes de sistema.
- Elección de proveedores.
- Auditoría (Comparar los resultados con los requisitos, validar puntos de control)
- Documentación.
- Aceptación formal de documentos de análisis.

### d) Diseño.

Se trabaja en los detalles de construcción y configuración del sistema, definiendo nombres, direcciones, estructura, en fin el diseño lógico y físico de la conectividad de los componentes, a través de las siguientes actividades:

- Definición de la estructura general del sistema.
- Definición de la estructura.
  - Diseño físico (características de conectividad, modelos y tipos de equipos)
  - Diseño lógico (definición de configuraciones, direcciones, determinación de servicio, etc.)
- Definición de procesos.
- Documentación.
- Auditorías.

### e) Construcción.

En esta etapa se lleva a cabo todo definitivo durante las etapas de análisis y diseño, es cuando se realiza la instalación directa y la configuración de cada módulo del proyecto. Con las siguientes actividades:

- Preparación de los lugares para instalar los equipos
- Preparación de cableado de terminales.
- Instalación y configuración.
- Ejecución pruebas propias.

- 
- Evaluar.
  - Integrar equipos.
  - Documentar
  - Auditorias

#### **f) Pruebas.**

- Preparación de datos de prueba.
  - Reales.
  - Volumen.
  - Elaboración de muestras.
- Realización de pruebas.
  - Individuales.
  - De volumen
  - Integrales
  - Respaldo y recuperación.
  - Especiales.
    - Rendimientos.
    - Tiempo de respuesta.
- Depuración
- Documentación.
- Auditorias.
- Aceptación del usuario. (Comparar con los requisitos del usuario definidos en la planeación ).

#### **g) Implantación.**

- Elaboración de documentos de referencia.
  - Manual de usuario.
  - Manual de instalación
  - Manual de sistemas.
  - Manual de procedimientos.
- Definición de seguridad
- Conversión.
- Capacitación a usuarios.
- Operación y evaluación.
- Auditorias.
- Liberación. (Aceptación formal).

---

## **h) Mantenimiento.**

- Solicitud o detección de Mantenimiento.
  - Adaptación.
  - Corrección.
  - Mejora.
- Análisis de la petición.
- Diseño.
- Construcción.
- Pruebas.
- Implantación.
- Documentación.
- Auditoria.

El desarrollo de un sistema bajo una metodología , con apoyo directo de la Dirección de la empresa, y la administración, da vida a un sistema por el grado de prevención y estudio de quien lo desarrolla y de las participaciones completas del usuario, así como por el ritmo de cambios en la operación general de la actividades que le dieron origen.

Para el desarrollo de sistemas, las necesidades y limitaciones del usuario deben de estar determinadas y explícitamente establecidas, y apoyadas por estudios que justifiquen el desarrollo del sistema de información. El producto debe diseñarse considerando tanto a los que los desarrollan como a los usuarios, y a quienes les den mantenimiento. La configuración e instalación debe realizarse con cuidado y probarse profundamente, además de que debe prepararse la documentación de apoyo, así como los principios de la operación. En la documentación no se debe olvidar: El manual del usuario, las instrucciones de instalación, las guías de entrenamiento y los manuales de documentación técnica.

Las tareas de mantenimiento incluyen solicitudes de análisis de cambio, diseño y modificación del sistema, pruebas completas de la versión modificada, actualización de la documentación para mostrar dichos cambios, así como la distribución de la nueva versión en los lugares apropiados.

Las metas principales en el desarrollo de los sistemas, son mejorar la calidad de los sistemas, aumentar la productividad en la empresa a través del apoyo a la toma de decisiones.

---

### **3.1.2 IMPORTANCIA DE LA VERIFICACIÓN DEL REQUERIMIENTO.**

Dentro de las grandes empresas, cuando se realizan proyectos costosos es importante realizar la verificación de la necesidad que los generó, definición del sistema y una evaluación concreta de su tamaño, factibilidad y costo-beneficio.

El origen de todo es la necesidad, y por tanto, la descripción de la problemática especificando los elementos principales del problema a resolver, los resultados que se esperan obtener, la delimitación del problema al realizar una solicitud formal con ciertos compromisos por ambas partes (usuario-desarrolladores).

En esta fase se trata de verificar la existencia de la necesidad potencial insatisfecha y hacer un estudio de viabilidad, desde el punto de vista operativo, demostrar que tecnológicamente es posible producirlo, una vez que se verificó que no existe impedimento alguno para obtener todos los recursos necesarios para la producción y finalmente demostrar que es económicamente rentable realizarlo.

Los objetivos del proyecto están en función de las intenciones de quienes los promueven, y se pueden agregar, cuales son limitaciones que se imponen, la localización, el monto máximo de la inversión y otros elementos. Dado que el proyecto a desarrollar es de mucha importancia en la empresa y se nota que requerirá una inversión alta, se profundizará en la parte de evaluación de proyectos.

### **3.2 EVALUACIÓN DE PROYECTOS**

Para realizar la evaluación de un proyecto normalmente se reúnen grupos interdisciplinario sobre diferentes áreas de estudio, cada uno de los especialistas desarrolla la parte que le corresponde. El resultado de esta interacción es un estudio completo acerca de la viabilidad técnica, económica y de mercado que sirve de base para decidir la realización de alguna inversión.

Actualmente hace falta que el que toma la decisión tenga un cierto integrador para que comprenda la información presentada y sea capaz de comprenderla, criticarla y elegir la mejor decisión y al mismo tiempo pueda evaluar sus impactos en los diferentes ambientes dentro y fuera de la empresa.

---

Un proyecto es la búsqueda de una solución inteligente a un problema tendiente a resolverse, por lo que se convierte en una necesidad humana. En esta forma, puede haber diferentes ideas, inversiones de diversos montos, tecnología y metodología con diversos enfoques, pero todas ellas dedicadas a resolver las necesidades del ser humano en ciertas facetas, como pueden ser: educación, alimentación, salud, ambiente, cultura, etc. El proyecto de inversión cualquiera que sea, tiene por objeto conocer su rentabilidad económica y social, de tal manera que asegure resolver una necesidad humana en forma eficiente segura y rentable.

### **3.2.1 IMPORTANCIA DE LOS PROYECTOS**

Siempre que exista una necesidad humana de un bien o servicio, habrá una necesidad de invertir, pues hacerlo es la única forma de producir un artículo, o invertir en un proyecto que finalmente se traducirá en dinero. En la actualidad, una inversión fuerte debe tener una base que la justifique, dicha base es precisamente un proyecto bien estructurado y evaluado que indique la pauta que debe seguirse y precise los indicadores adecuados para medir el impacto de su beneficio una vez desarrollado e implementado, y a la vez deje claros los beneficios con respecto al costo de inversión para la empresa, de ahí la importancia de elaborar proyectos.

Para decidir sobre el desarrollo de proyectos, es necesario que éste sea sometido a un análisis multidisciplinario, una decisión de este tipo no puede ser tomada por una sola persona con un enfoque limitado, o ser analizado sólo desde un punto de vista. Aunque no se puede hablar sobre una metodología rígida que guíe en la toma de decisiones sobre un proyecto, fundamentalmente debido a la gran diversidad de proyectos y a sus diferentes aplicaciones, si es posible afirmar que una decisión siempre debe de estar basada en el análisis de un sin número de antecedentes con la aplicación de una metodología lógica que abarque la consideración de todos los factores que participan y afectan al proyecto.

Sin embargo, El hecho de realizar un análisis aun cuando se considere lo más completo posible, no implica que, la inversión está exento de riesgo. También se dice que los estimados en los estudios costo-beneficio, sobre todo financieros no necesariamente se darán de la manera analizada. depende de como se maneje el proyecto y las consideraciones reales de su implantación , pero permite predecir algunos contratiempos y estimados globales de montos de inversión.

El proyecto de evaluación de proyectos y toma de decisiones no proporciona resultados iguales en un tiempo que en otro, si en un proyecto de inversión privada se diera a evaluar a dos grupos multidisciplinarios distintos, es seguro que los resultados no serían iguales. Esto se debe a que conforme avanza el estudio, las alternativas de selección son multiples en el tamaño, la localización, el tipo de tecnología que se emplee, la organización, enfoque, prioridad, etc.

---

Particularmente la realidad económica, política, social y cultural de la entidad donde se piense invertir, marcará los criterios que se seguirán para realizar la evaluación adecuada, independientemente de la metodología empleada.

### **3.2.2 METODOLOGIA DE LA EVALUACION DE PROYECTOS**

Aunque existen metodologías recomendadas en la evaluación de proyectos, sólo se pueden tomar las consideraciones que se aplique dependiendo del proyecto. La metodología presentada debe emplearse en un buen porcentaje, ya que se conoce que se ha aplicado en casos de estudio de factibilidad en el área de informática. Con respecto al proyecto de conversión de formato y la creación de base de datos puede ser aplicable a variaciones.

Se distinguen tres niveles de profundidad en los estudios de evaluación de proyectos: Al más simple se la llama "perfil", "gran visión" o "identificación de la idea", la cual se elabora partir de la información existente, el juicio común y la opinión que da la experiencia. En términos monetarios sólo presenta cálculos globales de las inversiones, los costos de los ingresos, sin entrar a investigaciones de terreno.<sup>7</sup>

El siguiente nivel se denomina "estudio de la prefactibilidad" o "anteproyecto". Este estudio profundiza la investigación en fuentes secundarias y primarias en la investigación de mercado, detalla la tecnología que se empleará, determinará los costos totales y la rentabilidad económica del proyecto, y es la base en que se apoyan los inversionistas para la toma de decisiones.

El nivel más profundo y final es conocido como "proyecto definitivo" contiene básicamente toda la información del anteproyecto y las pautas para su desarrollo y finalmente la evaluación de los indicadores después de su preparación para evaluar que tan eficiente es y cuáles serán las mejoras que requiere. A continuación se presentará en forma de lista los pasos en la evaluación de los proyectos:

- a) Análisis del entorno
- b) Detección de necesidades
- c) Análisis de oportunidades para satisfacer las necesidades.
- d) Definición conceptual del proyecto.
- e) Estudio del proyecto.
- f) Evaluación del proyecto.
- g) Decisión del proyecto.
- h) Realización del proyecto.

---

<sup>7</sup> Baca Urbina "Evaluación de Proyectos"

---

En concreto, se tomarán algunos los pasos de la metodología antes presentada; es decir, el método para la verificación de requerimiento se realiza a través de la definición conceptual del proyecto, el estudio del proyecto, el estudio técnico y el análisis económico, que nos dará una idea clara de que se va a lograr y si es necesario y factible que se haga. A continuación se desglosan sólo los pasos de la metodología de evaluación de proyectos que se ocuparán la verificación de las necesidades de desarrollar el proyecto.

**a) Definición conceptual del proyecto.**

Aquí se incluye una descripción inicial, retomando el motor o motivo que originó el proyecto, un análisis del entorno y finalmente se presentarán los requerimientos y partes que conformarán el sistema para satisfacer las necesidades planteadas, es así como se llega a la definición conceptual de lo que se hará en el desarrollo.

**b) Estudio del proyecto**

En punto se refiere a realizar un estudio para ratificar la existencia de la necesidad insatisfecha, cuantificar el servicio, la manera en que éste será proporcionado y cuales son las necesidades potenciales en un futuro pronto. Durante esta etapa el servicio es visto a través de la demanda de los usuarios del sistema, la oferta y comercialización del servicio, para cuantificar las necesidad potencial a ser cubierta. En los siguientes puntos se presentan las características de los pasos en el estudio del proyecto:

**i) Análisis de la oferta del servicio**

El propósito de este análisis es determinar o medir las cantidades y las condiciones en que se esta ofreciendo el servicio de las necesidades y comparar con la demanda para determinar si está bien la capacidad que se pretende instalar para cubrir los servicios demandados es correcta.

**ii) Análisis de la demanda del servicio.**

El principal propósito que se persigue con el análisis de la demanda, es determinar o medir cuáles son las fuerzas que afectan los requerimientos del mercado con respecto a un bien o un servicio, así como determinar la posibilidad de participación del producto en la satisfacción de dicha demanda.

Existe varios tipos de demanda, en relación con su oportunidad que se pueden clasificar como sigue:

- **Demanda insatisfecha.** Aquella en la que lo producido u ofrecido no alcanza a cubrir los requerimientos del mercado.

---

• Demanda satisfecha. Aquella en la que se le ofrece al mercado exactamente lo que se requiere. Se pueden conocer dos tipos de demanda satisfecha: La saturada que no puede soportar mayor cantidad del bien o servicio, pues se está usando plenamente y la no saturada que se encuentra aparentemente satisfecha pero que aún puede crecer.

• Determinación de la demanda potencial insatisfecha. Se refiere a la cantidad de bienes y servicios que es probable que el mercado consuma en los años futuros; sobre el cual se ha determinado que ningún productor actual podrá satisfacer si prevalecen las condiciones en las cuales se hizo el cálculo. Se obtiene una simple diferencia, del balance de la oferta y demanda y con los datos proyectados se puede calcular la probable demanda potencial o insatisfecha en el futuro.

Cuando se habla por ejemplo de la constitución de un maquinaria, la demanda potencial insatisfecha son los pedidos no surtidos o el servicio no prestado por el equipo debido a la capacidad insuficiente; a futuro, la demanda potencial insatisfecha se calcula considerando que el nivel actual del servicio se mantiene constante, es decir, no se sustituye el equipo y la demanda del servicio crece, de forma que al paso del tiempo, se dejan de sentir más pedidos o se deja de prestar servicio por parte del equipo.

### iii) **Análisis de la comercialización.**

Es el aspecto de la mercadotecnia más descuidado, regularmente siempre se mide la factibilidad del proyecto, pero no la manera en la que será organizada la atención del servicio con los usuarios. de esto también depende el éxito de éste.

En la realidad, cuando el equipo ya está instalado surgen todos los problemas que la comercialización representa, se puede tener el mejor equipo, pero si no se sabe operar o si no se proporciona el procedimiento o el medio para hacer llegar el servicio adecuado finalmente se degrada todo el conjunto.

La comercialización no es la simple transferencia de productos hasta las manos del consumidor; esta actividad debe conferirle al producto los beneficios de tiempo y lugar; es decir, una buena comercialización es aquella que coloca el producto en el sitio y momento adecuados, para dar al consumidor la satisfacción que él espera.

Regularmente después de implementar un nuevo servicio, no se cuenta con la capacidad del personal para proporcionar dicho servicio, por lo tanto es buscar apoyos adicionales mientras se gana el conocimiento y experiencia para estos casos, es pensar en asesoría de instalación e implementación del servicio por despachos que proporcionen ciertas ventajas, como asumir los riesgos sobre el mantenimientos, funcionalidad y respaldos con otros equipos e ingeniería en casos de fallas.

---

Los medios de distribución son la ruta que toma un producto o servicio para pasar del productor al consumidor, en ese caso se tratará del procedimiento por el cual se determina la relación del usuario con el departamento que ofrece el servicio y la manera de canalizar los servicios necesarios, su utilización y canalización de las nuevas necesidades.

También será necesario realizar estudios de capacidad, tanto de diseño, como del sistema para reducir errores y aprovechar la capacidad total del sistema para alcanzar eficiencia en el servicio:

- Capacidad de Diseño. Reducido por mezcla de productos y condiciones del mercado a largo plazo , altas especificaciones de calidad y el balance inadecuado entre equipo y mano de obra en su operación.
- Capacidad del Sistema. Reducida por efectos a corto plazo como la demanda actual, desempeño de los directivos ( mala programación estrategias, y control eficiente, etc), ineficiencia de los trabajadores ( falta de aptitudes y bajo nivel de esfuerzo), ineficiencia de las máquinas ( paros, mantenimiento, reemplazo, etc), y finalmente la producción real que es la que el consumidor ve directamente.

La comparación entre oferta y demanda y la comercialización da como resultado, una visión del tamaño e impacto del proyecto. El tamaño del proyecto está determinado por varios factores a analizar como :

- El tamaño de la demanda
- El nivel de suministros e insumos
- La tecnología empleada y los equipos
- El personal y la organización necesaria para desarrollarlo e implementación.

### **c) Estudio Técnico**

Durante este estudio se verifica la posibilidad técnica ( factibilidad ) de realizar el proyecto, analizar y determinar las capacidades de los equipos y de los medios de comunicación , así como las instalaciones y la organización necesaria para realizar el proyecto, todo esto es sólo a nivel compatibilidad y factibilidad. Comúnmente se logra a partir del análisis de las siguientes etapas:

- La determinación del tamaño de la planta, que en nuestro caso de traduciría en la preparación de las instalaciones físicas para la conexión de los equipos.
- Determinación de la localización óptima de los equipos.
- Ingeniería del proyecto y análisis administrativo ( definición de la factibilidad de adquisición , compatibilidad y conexión de todos los elementos a ser incluidos ).

---

En el caso de la instalación de equipo resulta un tanto difícil analizarlo y determinarlo, pues depende de una investigación muy minuciosa y concursos anticipados de proveedores. Por lo anterior el estudio realmente se limitará a checar compatibilidades entre equipos, instalaciones necesarias y capacidades de los sistemas y enlaces sin entrar en cosas detalladas que se llevarán a cabo en la etapa del análisis.

El estudio mide la facilidad del proyecto a corto y mediano plazo, es muy importante destacar como objetivo en esta parte un documento de factibilidad técnica que dé la tranquilidad de que existe la posibilidad tecnológica que soporte el sistema.

#### d) Estudio económico.

La parte de análisis económico pretende determinar cuál es el monto de los recursos económicos necesarios para la realización del proyecto, cuál será el costo total de operación, así como los beneficios directos que se tendrán con cada servicio, Lo anterior se llevara a cabo en los siguientes pasos:

- Investigación de servicios y costos estimados según capacidades determinadas.
- Determinación de la inversión inicial
- Determinación del monto de Gastos de operación mensual.
- Determinación de beneficios directos que se tendrán para hacer un balance.
- Determinación de las premisas para lograr el punto de equilibrio, es decir, de que manera se comercializará (utilizará e implantará), para lograr los costos más bajos con el beneficio más alto.
- Análisis del riesgo sobre la inversión.

Hablando sobre el análisis económico, es decir que la palabra "costo" nadie a logrado definirla con exactitud, debido a si amplia aplicación, pero se puede decir que el costo es un desembolso en efectivo o en especie hecho en el pasado, en el presente, en el futuro o en forma virtual. Por ejemplo, los costos pasados, que no tienen efecto para propósitos de evaluación económica se les llama "inversión", en un estado de resultados proforma o proyectado en un evaluación, se utilizarán los costos futuros, y el llamado "costo de oportunidad" sería un buen ejemplo de costo virtual, así como también lo es el desembolso hecho.

También es importante señalar que la evaluación de proyectos es una técnica de planeación y la forma de tratar el aspecto contable no es tan rigurosa, lo cual demuestra cuando por simplicidad, las cifras se redondean al millar más cercano. Esto es así pues no hay que olvidar que se está tratando de predecir lo que sucederá en el futuro, y sería absurdo decir, por ejemplo, que los costos de producción para el tercer año de funcionamiento del proyecto serán de una cantidad exacta. No hay forma de precedir con tanta exactitud el futuro.

---

Por lo anterior, debe quedar claro y aceptado que el redondeo de las cifras a miles no afecta en absoluto la evaluación económica y no se está violando ningún principio contable puesto que aquí no se trata de controlar las cifras del proyecto, pues sería tanto como querer controlar con esta rigurosidad el futuro, lo cual no es posible.

En este momento surge el problema sobre el método de análisis que se empleará para comprobar la rentabilidad económica del proyecto. Se sabe que el dinero disminuye su valor real con el paso del tiempo a una tasa aproximadamente igual al nivel de inflación vigente. Esto implica que el método de análisis empleado deberá tomar en cuenta en este cambio de valor real del dinero a través del tiempo.

Cuando se realiza la sustitución de un equipo o maquinaria, la evaluación económica difiere ligeramente de los métodos comúnmente presentados para la evaluación, en el reemplazo de equipo. En estos casos se presentan dos situaciones claramente definidas, que a su vez, obligan a definir un método específico de evaluación económica. La primera situación surge cuando la maquinaria a sustituir sólo es parte de un proceso productivo y no produce ingresos por sí misma, es decir, contribuye a la elaboración de un producto y es muy difícil cuantificar con precisión con cuanto contribuye el trabajo de esa máquina, al costo real del producto.

Si el proyecto produce directamente un artículo terminado o servicio que al venderse produciera ingreso, aunque la misma empresa produzca una gran variedad de artículos, es posible aislar la evaluación económica de esa maquinaria por el método de análisis incremental, el cual permite introducir al análisis toda la serie de datos reales que se pueden originar, como son aumento de productividad, disminución de costos, depreciación, impuestos, etc.

Se llama análisis incremental por que cuantifica aumentos de inversión a los cuales debe corresponder aumentos de ingresos, es decir, se tiene un equipo trabajando normalmente y éste produce determinado ingreso, la inversión actual es cero, puesto que el equipo se compró hace tiempo. Como se pretende reemplazar dicho equipo se produce un incremento de inversión por la compra de equipo nuevo, a este incremento de inversión debe corresponder un aumento proporcional de ingresos, de no ser así la inversión tendría que rechazarse.

La evaluación en este caso se hará por el método, y dará como resultado el estudio costo-beneficio del desarrollo del proyecto, comparando costos contra beneficios en su productividad, disminución de otros rubros de costos, impuestos, trabajos externos etc.

En concreto, en la etapa de verificación del requerimiento se tendrán como resultado: Un estudio de validación de la necesidad existente, un estudio de factibilidad técnica del desarrollo del proyecto que satisfaga la necesidad y un estudio de costo-beneficio que justifique y proyecte la inversión necesaria para llevar a cabo el proyecto.

---

# **CAPITULO 4**

---

---

## **4.1 VERIFICACION DE LOS REQUERIMIENTOS**

### **DEFINICION CONCEPTUAL DEL PROYECTO**

La exploración sísmica es un estudio que se hace en una determinada zona geográfica del país con el propósito de localizar hidrocarburos. Una exploración sísmica se puede realizar en tres formas diferentes.

Exploración marina, exploración terrestre y exploración en vibrador. Cada exploración utiliza equipos especiales y formatos para su grabación, por ejemplo ( la exploración terrestre utiliza dinamita , se hace una perforación en el suelo de 20 a 25 mts. de profundidad, se carga el pozo con dinamita, y se detona el pozo, el registro un geofono lo envía a un cajetín y este a su vez lo envía a la cinta magnética para su grabación ).

A lo largo de estos años se han realizado diversas exploraciones sísmicas, en la que toda la información se ha registrado en las cintas magnéticas con diversos tipos de formatos. Se tiene aproximadamente 200 mil cintas por compactar y cambiar el tipo de formato ( SEG-Y ), estas 200 mil cintas están almacenadas en bodegas sin las condiciones apropiadas para su conservación, tienen una duración aproximada de 20 años en condiciones propias, pero si no se tiene una instalación para su cuidado las cintas se reducen a una vida de 10 aproximadamente.

Además el volumen físico que ocupan estas cintas es bastante considerable, las bodegas que se tienen ya no son suficientes , por lo que se tiene que construir nuevas instalaciones para su almacenamiento y conservación. Esto ocasiona gastos que no contemplados para la entidad. De esta problemática es como surge el proyecto de transformación contenida en las cintas magnéticas a cartuchos 3490E con un formato de salida estándar ( SEG-Y).

Se decidió utilizar el cartucho 3490E por su capacidad de almacenamiento, además que se encuentra totalmente cubierto para su conservación a posibles daños ambientales. Se pretende cubrir las siguientes características:

- Reducir el espacio físico para su almacenamiento.
- Cambio de los diversos formatos de todas las cintas magnéticas a cartuchos 3490E con un formato secuencial estándar ( SEG-Y ).
- Recuperación de la información ya que las cintas se encuentran en malas condiciones, ya que necesitan tratamiento de especial de cintas.
- La seguridad del cartucho es buena ya que se encuentra totalmente sellado, su duración es aproximada de 30 años.

---

Los objetivos principales del proyecto:

- Convertir los formatos de grabación contenidas en cintas magnéticas a cartuchos 3490E con un formato secuencial de salida ( SEG-Y ).
- Crear de una base de datos que contendrá toda la información de los diversos prospectos realizados de las diferentes zona geográficas del país, así como el equipo utilizado para su exploración , y datos generales quea indentifican particularmente un prospecto.

## **4.2 PLANEACION Y CONTROL DEL PROYECTO**

Las principales actividades que cubre este proyecto son:

- Revisión, clasificación, limpieza y recuperación de cintas.
- Compactación de cintas a cartuchos 3490E
- Control de calidad de la información recuperada.
- Registro en la base de datos del manejo de las cintas
- Almacenamiento óptico de los reportes de observador.
- Informes parciales y finales.

**La revisión, clasificación, limpieza y recuperación de cintas y cartuchos se efectuará según el siguiente procedimiento.**

- a) Revisión del material recibido tanto de cintas magnéticas como de reportes de observador.
- b) Clasificación por región , área y prospecto.
- c) Inspección física del material, para conocer las condiciones en que se encuentra el material.
- d) En caso de que alguna cinta o cartucho que encuentre en malas condiciones para su compactación se procedera a realizar el tratamiento para despegar y recuperar la información grabada en ella.

---

**Realizar la compactación de las cintas magnéticas de acuerdo al siguiente procedimiento:**

- a) Grabación de la información en formato SEG-Y en el cartucho de salida tipo 3490E
- b) Duplicación del cartucho de salida.
- c) Etiquetado del cartucho de salida con un código de barra, la numeración será consecutiva para la correcta clasificación de la información.
- d) La información correspondiente a cada línea sismológica deberá ir grabada en cartuchos individuales, no debiendo tener cada cartucho, información de 2 o más líneas.

**Procedimiento para el control de calidad.**

- a) Verificación de los cartuchos para revisión de parámetros de grabación, tales como: tiempo de registro, intervalo de muestreo, canales utilizados ,primer y último registro de grabación.
- b) Historia de la transcripción indicando, fecha de compactación , errores de lectura recuperables y no recuperables, longitud de los registros, intervalo de muestreo, puntos de tiro y sus números de registro y demás información que se considere necesaria.
- c) Del cartucho grabado, generar una grafica de la traza más carcana a cada uno de los puntos observados, en escala de 8 trazas por centimetro por segundo.
- d) Generar a partir del cartucho grabado , una grafica de todas las trazas sismicas cada 100 registros en la misma escala. la escala puede ser modificada más o menos 50% de común acuerdo.
- e) Registrar en el programa de manejo de cintas, todos los cartuchos de salida, generados con sus datos de región, área, prospecto, línea y punto de tiro contenidos, número de cartucho, y demás información que se considere necesaria.

**Almacenamiento optico de reportes del observador**

Los reportes del observador serán digitalizados ópticamente ( escaneados) y las imagenes serán grabadas directamente de los discos ópticos de 5 1/4 pulgadas ( WORM ). El registro se llevará en un sistema teniendo como manejador de base de datos ORACLE , e interface grafica del usuario tipo windows , de tal forma que sea posible llevar el control preciso de todos los

---

datos grabados de cada disco. Deberá grabarse en ambos lados de los discos para permitir la grabación de 5000 a 8000 páginas de información por disco.

#### **Informes mensuales de la actividades**

Se entregarán dos tipos de informes los parciales y los finales.

##### **INFORME MENSUAL:**

- Número de cintas compactadas obtenidos área y prospecto
- Número de cintas que fuerón tratadas para su recuperación.

##### **INFORME FINAL:**

- Región, área y prospecto.
- Relación de las líneas transcritas y compactadas.
- Relación de cartuchos por cada línea.
- Relación de puntos de tiro de cada línea, en cada cartucho, con sus puntos iniciales y finales de su observación.
- Relación de los cartuchos pegados, dañados o con problemas de desprendimiento del material ferromagnético, que fueron recuperados para la recuperación de la información.

Se proporcionará la siguiente información al finalizar el trabajo de compactación:

1. Cartuchos tipo 3490E con información compactada en formato secuencial de grabación SEG-Y
2. Duplicado de los cartuchos con la información recuperada.
3. Disco ópticos de 5 1/4 pulgadas conteniendo los reportes de observador.
4. Listado de la impresora con la historia de la transcripción de información.
5. Graficas de control de calidad.
6. Entrega de los reportes finales.

#### **4.2.1 IDENTIFICACION DEL GRUPO DE TRABAJO**

Comúnmente el tipo de administración varía dependiendo de la etapa en que se encuentre el proyecto, ya que en la etapa de análisis y diseño se requiere ser más democrático, es decir las

---

metas se definen por concenso y el liderazgo depende de las tareas y las capacidades que tenga cada miembro.

Sin embargo en la etapa de construcción es necesario que el líder coordinador asigne tareas, asista a visiones y recorridos, detecte áreas de problemas, balancee cargas de trabajo y participe en actividades . También es necesario que el coordinador establezca los métodos para dirigir y controlar las actividades de los participantes del proyecto.

Al inicio del proyecto debe de estar clara la participación , basándose en la percepción que haga el coordinador del proyecto, esta se revisa por el grupo de trabajo y por la dirección y se difine en su caso, de manera que para todos queden claras sus responsabilidades y participación. También se diseñan objetivos de responsabilidades claros, concretos y explícitos, tangibles,y factibles, que finalmente puedan ser validados por el cordinador del proyecto.

Se hace notar que una vez que se dividió el trabajo y las responsabilidades, todos deben estar perfectamente enterados y de acuerdo con su función, de los resultados que específicamente entregarán todos al mismo tiempo, todo esto servirá de alimentación por los modulos participantes en el proyecto. Enseguida se determinará en forma genérica los equipos de trabajo necesarios para el desarrollo del proyecto, y durante el análisis se determinará quién participará específicamente en que actividad .

### **4.3 ANALISIS DEL PROYECTO**

#### **PERSONAL REQUERIDO:**

- **JEFE DE PROYECTO:** Es el encargado de coordinar las actividades de cada uno de los participantes del proyecto.
- **JEFE DE SOFTWARE:** Es el encargado de las configuraciones de los equipos y el mantenimiento de estos mismos.
- **ING DE SISTEMAS :** Al igual qu el Ing. de software se encargará de las configuraciones de los equipos y mantenimiento de los mismos.
- **ANALISTAS:** Estos se encargarán de la interpretación de los tipos de formatos y la realización de programas de dichos formatos.

- 
- **CINTOTECARIO:** Realizará la clasificación y revisión de los reportes del observador y de la cintas.
  - **OPERACION:** Serán las personas que manejarán los equipos y mandarán la ejecución de los programas hechos por los analistas.
  - **CAPTURISTAS:** Se encargarán de la captura de datos de topografía.

## **EQUIPOS REQUERIDOS:**

### **3 SISTEMAS PECC**

- Computadora 486/66 Mhz- memoria 16 Mb - sistema operativo UNIX
- Bus IDE - BUS EISA - BUS SCSI 2
- Controlador SCSI estandar - devices 0 a 7
- Controlador de cintas STC - desarrollado por PECC
- Controlador de cintas PERTEC - desarrollado por PECC
- Terminal operador ( ASCII )
- Impresora epson 570 - matriz de punto ( impresión en línea a cada registro )
- Unidad de cartuchos FUJITSU 2483K- 3490E
- Unidad de cintas STC 1968 con formatter - tridimensional 800/1600/6250 - 125 ips
- Unidad de cintas STC 1963 sin formatter - tridimensional 800/1600/6250 - 125 ips
- Unidad de cintas STC 9914 - M4 streamer -50/100 ips. - estandar PERTEC

### **UNIDADES DE CINTA ADICIONALES:**

- Unidad CYPHER 910 - 45 ips - 800/1600 bpi - estandar PERTEC
- Unidad CYPHER 910 - 60 ips - 800/1600 bpi - estandar PERTEC
- Unidad KENNEDY 9800 - 25 ips - 800-1600 bpi

### **3 SISTEMAS IBM RS/600 -250**

- Computadora mod. 7011- memoria 32 Mb - sistema operativo AIX (UNIX)
- Disco interno 2 Gb - Bus SCSI interno - 2 Bus SCSI externos
- Unidad EXABYTE 8500 - 5 Gb

- 
- Controlador VERSATEC 130 para plotter
  - Controlador de red ETHERNET
  - Terminal X NCD 15"
  - Terminal X NCD 19 "
  - Impresora HP laserjet IV SI
  - Plotter termico ATLANTEK 24 "
  - Unidad de cintas STC 9914 - M4 - estandar SCSI con Buffer ( streaming mode )
  - Unidad de cartuchos FUJITSU 2483K- 3490E - con cargador automático
  - Unidad de cartuchos FUJITSU 2483K- 3490E - con cargador automático

### **SISTEMA IBM RS/6000 - 250 SISTEMA BASE DE DATOS**

- Computadora mod. 7011 - memoria de 32 Mb - sistema operativo AIX ( UNIX )
- Disco internos 2 Gb - Bus SCSI interno - 2 Bus SCSI externos
- Controlador de red ETHERNET
- Terminal X NCD 19 "
- Terminal IBM ASCII
- Scanner FUJITSU 3096G - formatos carta y boble carta
- 2 controladores SCSI externos
- Unidades de discos ópticos 1.2 Gb - SCSI express MDI - 5 1/4
- Lector de código de barras
- Discos externos 2 Gb ( base de datos ORACLE )

### **SISTEMA COPIADO DE CARTUCHOS**

- Computadora 586 pentium 90 Mhz - Memoria 16 Mb - Sistema operativo DOS
- Controlador SCSI Fast Wide
- Terminal color y mouse
- Unidad de cartuchos FUJITSU 3490E 2483K con cargador automático
- Unidad de cartuchos FUJITSU 3490E 2483K con cargador automático

### **EQUIPO PARA TRATAMIENTO DE CINTAS**

- Horno BLUE M ELECTRIC - con rampas programables - capacidad 120 cintas
- Horno GRIEVE- con bomba de vacio - capacidad 15 cintas
- 3 limpiadores de cintas Mod. 3
- Limpiador de cintas KYBE 401 - con detector óptico

- 
- Limpiador de cartuchos GRAHAM DETECTOR 80 - con cargador automático

## **SOFTWARE UTILIZADO EN COMPACTACION:**

### **SISTEMA PECC**

- Sistema operativo LYNX OS ( UNIX )
- Para la demultiplexación, cambio de formato, edición . Paquete RAPPORT
- Utileria de copia, dump, listado, analisis. Paquete MSU

### **SISTEMA IBM**

- Sistema operativo AIX ( UNIX )
- Cambio de fomato, arreglo de secuencia de los punto de tiro, control de listados, jobs de plots, pegado de cartuchos, y todo tipo de trabajos especiales. Paquete " GEOVECTEUR ".
- Software X11-OSF/MOTIF para el manejo de pantallas gráficas.

### **SISTEMA PARA COPLADO DE CINTAS**

- Sistema operativo DOS MICROSOFT
- Copiado de cartucho 3490E a 3490E. Paquete intergrafico " SEMIR "

### **SISTEMA PC**

- Sistema operativo MSDOS MICROSOFT
- Edición de formularios de producción, reportes de compactación. Paquete EXCEL y WORD de MICROSOFT

### **SISTEMA BASE DE DATOS**

- Base de datos relacional ( ORACLE )
- Paquete GEOBASE, GEODOC, GEOVIEW., para el scaneado , display de documentos, manejo de discos ópticos, manejo de datos sísmicos en general.

---

## **RECEPCION Y PREPARACION DE LA INFORMACION A COMPACTAR**

- Se tendrá que clasificar la información tanto de las cintas como de los reportes observador de acuerdo al área y prospecto. Una vez hecha esta clasificación se tendrá que hacer una nueva clasificación, pero esta clasificación se tendrá que realizar ahora por línea de acuerdo al área y al prospecto en donde se ha realizado la observación sísmica.

- El procesamiento de las cintas como de los cartuchos físicamente dañados se tendrá que realizar a través de hornos y limpiadores, el promedio para tratar una cinta es de 24 a 36 horas a 55 grados C, con rampas de ascenso y descenso de temperatura de 6 horas, la limpieza se realiza de 1 a 3 ciclos completos.

- Entrega de los reportes del observador a los analistas. Estos se encargarán de realizar un Dump ( vaciado de la información contenida en la cinta en formato hexadecimal ) de la cinta para saber que tipo de información esta contenida en ella, una vez analizada e interpretado el tipo de formato de grabación que se utilizó durante la observación sísmica, se realizará un programa ( jobs ) para la compactación a cartucho 3490E y el cambio de formato secuencial SEG-Y

Antes de la realización de los programas para la interpretación de los datos sísmicos se dará una explicación de los diversos formatos que existen hasta el momento.

## **FORMATOS DE GRABACION PARA INFORMACION SISMICA**

En general, se pueden agrupar en dos tipos de formatos: formatos multiplexados y formatos demultiplexados (o secuenciales). Para cada uno de estos tipos existen formatos que podrían ser considerados como estándar y no-estándar.

### **FORMATOS MULTIPLEXADOS**

Los formatos multiplexados estándar son los formatos propuestos por la SEG (Society of Exploration Geophysicist) : SEG-A, SEG-B, SEG-C y SEG-D.

El formato EPR que es un formato utilizado para la transcripción de cintas analógicas a cintas digitales, es un formato no estándar.

En general los datos multiplexados consisten de tres bloques de información:

- **HEADER BLOCK** (Bloque del Encabezado) Donde se identifica el registro.
- **DATA BLOCK** (Bloque de Datos) : Contiene los Datos de los Canales.
- **TRAILER** (Fin de File) : Separa los Registros.

El Bloque del Encabezado contiene un número de identificación (No. de File), puede contener la información relativa a las ganancias de preamplificación (Early Gain) excepto para SEG-B y SEG-D. Este bloque puede ser separado (Gapped) o no separado (No Gapped) del Bloque de Datos. La longitud de este bloque es el "Header Length".

El Bloque de Datos consiste de las muestras de los datos para todos los canales de cada intervalo de muestreo. Los datos están arreglados en series de estructuras encabezadas por canales de sincronización donde se escribe el código de sincronía.

La identificación de un formato en particular parte en principio del reconocimiento de su código de sincronía. En la siguiente gráfica se presenta el tipo de código de sincronía para cada tipo de formato.

<b>FORMATO</b>	<b>CODIGO DE SINCRONIA</b>
<b>SEG-A</b>	FF FF FF
<b>SEG-B</b>	FD FD FF ; 01 01 01 ; 03 03 03
<b>SEG-C</b>	FF FF FF
<b>SEG-D</b>	01 01 00
<b>EPR</b>	FF FF FF

En ocasiones, la sola identificación del código de sincronía no basta para una adecuada descripción del formato de grabación (pues eventualmente puede encontrarse un código diferente al descrito en la tabla anterior); por lo que es necesario determinar el número de canales (auxiliares como de producción) en el equipo de registro. Para determinar tal cantidad, es necesario conocer cuántos "bytes por scan" se registran por cada muestra grabada en el sismógrafo. Para conocer esta cantidad en los formatos estándar sólo es necesario identificar el byte número 11 (byte 21 en SEG-D) en el "dump" de la cinta en cuestión, siempre que el formato de grabación sea estándar. Este valor aparece escrito en decimal aún cuando el "dump" sea hecho en hexadecimal.

En el caso de un formato no-estándar (EPR) los bytes por scan se determinan contando los bytes desde el inicio de un código de sincronía hasta el byte inmediatamente anterior al siguiente código.

Determinar los "bytes por scan" permite saber cuántos canales se utilizaron en el registro sísmico y de ellos cuantos corresponden a canales de producción. Generalmente, hay una correspondencia única entre cada formato y el número de "bytes por scan".

Asimismo, es importante tener una idea con qué tipo de palabra se hizo el registro sísmico: con palabra de 2 bytes = 16 bits (generalmente grabaciones de dos décadas atrás) o con palabra de 4 bytes = 32 bits (equipos de grabación más recientes); los cuales son los valores más comunes; aunque existen grabaciones en palabra de  $2 \frac{1}{2}$  bytes = 20 bits, 1 byte = 8 bits, etc. (estos últimos para los formatos SEG-D).

#### BYTES POR SCAN SEGUN FORMATO ESTANDAR Y NUMERO DE CANALES

FORMATO	24 CANALES	48 CANALES	96 CANALES	192 CANALES	240 CANALES
SEG-A	64	128	-	-	-
SEG-B	74	134	254	494	614
SEG-C	128	256	512	-	-
SEG-D*	78	138	258	498	618

\*Se muestran los valores para el SEG-D 15 que es el más común.

De la tabla anterior, se debe notar que se señalan los canales sísmicos o de producción y que para obtener el total de canales con que cuenta el sísmógrafo se debería de dividir el "número de bytes por scan" entre el valor de la palabra correspondiente a cada dato (palabra de 2 o 4 bytes). Esto para los formatos SEG-A, SEG-B y SEG-C. Para los formatos SEG-D es recomendable recurrir a la documentación disponible pues el formato de los datos varía ya que además de los canales de sincronía, de la palabra de tiempo y de los canales registrados con exponente; puede haber canales que se registran sin exponente de acuerdo al SEG-D en particular que se tenga.

---

Para el caso particular del formato EPR se podría esperar una similitud con el formato SEG-A. Esta similitud se daría en el "número de bytes por scan".

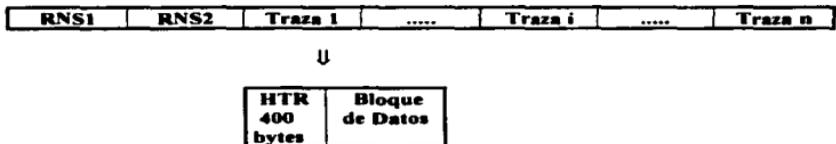
## FORMATOS SECUENCIALES

A diferencia de los formatos multiplexados, los formatos secuenciales tienen por característica que cada canal sísmico está registrado como un sólo bloque y uno tras otro formando cada registro sísmico, repitiéndose esta secuencia sucesivamente hasta el fin de la información.

De igual manera que en los formatos multiplexados hay dos tipos de formatos secuenciales: estándar y no estándar.

El formato secuencial estándar por antonomasia es el formato **SEG-Y** ( Society of Exploration Geophysicist ).

La estructura de este formato la muestra el siguiente esquema:



Donde **RNS1** es un bloque de 3200 bytes escrito en código EBCDIC y en el se escriben la información general de campo. **RNS2** es un bloque de 400 bytes escrito en código binario y en el se escribe la información general de proceso. El elemento **HTR** es el encabezado de cada una de las trazas y tiene una longitud de 240 bytes; en el se encuentra la información relativa a cada traza: No. secuencial de traza, No. de "file", No. de canal, No. de Punto de tiro, código de traza, No. de muestras, intervalo de muestreo, etc.

Cada uno de estos elementos está dividido en palabras de 2 y 4 bytes.

---

Hay dos tipos de programas que permiten la interpretación y recuperación de información sísmica esto se hace a través de los programas GEOVECTEUR Y RAPPORT.

En el sistema Geovecteur ( Programa para la interpretación y recuperación de datos sísmicos ), el bloque de 3200 bytes se codifica como HF1 (encabezado de cinta, elemento no sísmico 1). El bloque de 400 bytes es identificado como HF2 (encabezado de file, elemento no sísmico 2) mientras que el elemento de 240 bytes es HTR ( encabezado de cada una de las trazas ). Para el sistema RAPPORT ( Programa utilizado para el cambio de los distintos formatos ) la correspondiente codificación para estos elementos es: EBCHDR, BINHDR, TRCHDR respectivamente.

**EBCHDR** ( Es el encabezado de cintas con una longitud de 3200 bytes, que para el sistema Geovecteur corresponde a HF1 ).

**BINHDR** ( Es el encabezado de file con una longitud de 400 bytes , que para el sistema Geovecteur corresponde a HF2).

**TRCHDR** ( Es el encabezado de la traza con una longitud de 240 bytes, que para el sistema Geovecteur corresponde a HTR).

Los formatos secuenciales no estándar tienen la siguiente estructura general:

<b>RNS1</b>	<b>EOF</b>	<b>RNS2</b>	.....	<b>RNSn</b>	<b>EOF</b>	<b>Traza I</b>	.....	<b>Traza n</b>
-------------	------------	-------------	-------	-------------	------------	--------------------	-------	--------------------

Dentro de esta clasificación se encuentran los formatos PHOENIX "I", PHOENIX "F", WES-2 y WES4.

### **PHOENIX "I"**

Para PHOENIX "I" el elemento RNS1 es un elemento de 256 bytes seccionado en palabras de 16 bits. En este elemento se tiene una información relativa a la línea contenida en la cinta, es decir, se trataría de un Encabezado de línea. Después los elementos RNS2 a RNSn son generalmente 32 bloques de 80 bytes cada uno y contienen información histórica de proceso. Ocasionalmente, el tamaño de estos bloques y su cantidad podrían ser otros. Para PECC es necesario discriminar tales bloques, lo cual se realiza con la instrucción

Está instrucción salto los encabezados de 80 bytes.

Para saltar estos bloques en Geovecteur es suficiente con codificar PREB0.

En PHOENIX "I" las trazas contienen un encabezado (HTR) de 96 bytes dividido en palabras de 16 bits y el cual se localizan parámetros similares a los encontrados en el HTR del SEG-Y, en el entendido de que, por ser un formato diferente, tales parámetros tendrán una ubicación diferente en el encabezado.

#### **WES-2 y WES-4**

Los formatos WES-2 y WES-4 son formatos de cinta de Western Geophysical ( Cia. dedicada a la exploración sísmica estos formatos son propios de Western Geophysical "Wes" ).

En WES-2 los elementos RNS1 y RNS2 a RNSn no existen y la información comienza inmediatamente con los bloques correspondientes a las trazas. Aquí, cada bloque comienza con un encabezado de traza (HTR) con una longitud de encabezado de 34 bytes dividido en palabras de 16 bits o 2 bytes o media palabra que es lo mismo. De hecho, es conveniente buscar la palabra 1 pues en ella se localiza el identificador del formato (2) y también buscar la palabra 2 (que comienza a partir del byte 3) ya que en ella se encuentra codificada, en hexadecimal, la localización en palabra de 2 bytes del primer dato sísmico de ese bloque. Este valor es 11 equivalente a 17 en decimal. Una característica adicional es que en la palabra 7 siempre se tendrá el valor 4040.

En las palabras 8 a 10 se localiza el Identificador, que es un número compuesto de 6 caracteres correspondiendo los dos primeros al número de "file". Los cuatro siguientes caracteres el número de cinta al inicio de la línea y se incrementan según sea necesario (normalmente cada 100 "files"). De aquí se ve que la recuperación del número de file requiere de una serie de operaciones aritméticas factibles en Geovecteur.

Como se ve, aquí de igual manera que en PHOENIX "I" el encabezado de traza contiene información similar a la que se localiza en el HTR del SEG-Y.

---

En WES-4 existe una diferencia respecto a WES-2. Para WES-4 existe un encabezado identificador que contiene información relativa a los datos sísmicos. Algo similar a los bloques de 80 bytes en PHOENIX "I" sólo que la cantidad de bloques es variable y no se puede determinar un número característico. Este grupo de bloques de 80 bytes corresponden a los elementos RNS2 a RNSn.

Después de ellos siguen las trazas con su correspondiente encabezado de trazas donde es posible localizar: en la palabra 1 el identificador de formato: 4, en la palabra 2 el "Índice de inicio de datos sísmicos" y que generalmente es igual a 101 lo que determina, por otra parte, una longitud de encabezado de traza (HTR) igual a 100 bytes seccionado en palabras de 16 bits. En la palabra 3 se encuentra el número de "file" que, como en WES2, puede estar combinado con algún otro valor siendo necesario entonces recuperarlo a través de alguna operación aritmética.

A continuación se presentará un Dump tanto de formatos multiplexados como de los demultiplexados, el Dump es un vaciado en hexadecimal en donde se encuentran los valores de intervalo de muestreo, tiempo de grabación, número de file , número de canales , etc. Es importante el Dump para saber la localización de los valores a recuperar. Estos valores nos servirán para la realización de los programas ( jobs ) y poner en el jobs en que palabras localiza los datos que nos indican a recuperar.

## DUMP PARA FORMATOS MULTIPLEXADOS

Block1 ( file1 , block1) block size: 228

```
4364 0200 2905 9300 6523 2542 1500 0453 0698 1230 0000 6040 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600 2600
0000 0000
```

los números que se encuentran marcados indican los valores a recuperar por ejemplo:

el valor 4364 indica el número de file.

el valor 2542 indica dos valores, el 254 indica los bytes por scan y el valor 2 indica el intervalo de muestreo.

el valor 0698 indica el tiempo de grabación.  
el valor 288 indica el tamaño del bloque.

Block1 ( file1 , block2 ) block size: 780288

```
FFFF FFFF 0000 FFFC FFFC FFEF 111C 354C 5678 2356 4D8C 8978 E9B0 FF60 CEEE
1E0E EEA6 151E 134I ICBD 43TR 57UI FF60 CT91 01C1 6822 1500 3336 6578 8776
1E0E EEA6 151E 134I ICBD 43TR 57UI FF60 CT91 01C1 6822 1500 3336 6578 8776
1E0E EEA6 151E 134I ICBD 43TR 57UI FF60 CT91 01C1 6822 1500 3336 6578 8776
1E0E EEA6 151E 134I ICBD 43TR 57UI FF60 CT91 01C1 6822 1500 3336 6578 8776
```

el valor FFFF FFFF indica el código de sincronía.

## DUMP PARA FORMATOS DEMULTIPLEXADOS

Block1 ( File 1, block 1) block size: 3200

```
CEF1 F140 4040 4040 D751 D551 C1C3 C9D9 D540 D4C5 D675 4578 9000 8998 5421
E4C3 C3C9 4040 4040 89D1 H781 CL21 323D D213 DE31 DE11 2321 2322 3211 1231
E4C3 C3C9 4040 4040 89D1 H781 CL21 323D D213 DE31 DE11 2321 2322 3211 1231
CEF1 F140 4040 4040 D751 D551 C1C3 C9D9 D540 D4C5 D675 4578 9000 8998 5421
E4C3 C3C9 4040 4040 89D1 H781 CL21 323D D213 DE31 DE11 2321 2322 3211 1231
```

el primer bloque tiene un longitud de 3200 bytes en donde se encuentra información general contenida en la cinta ( indica la línea , área, prospecto, equipo utilizado, en que año se realizó la exploración, etc ).

Block1 ( File 1, block 2) block size: 400

```
0000 0001 0000 028A 0000 0001 0030 0005 07D0 07D0 0A00 0000 0001 0030 0001 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000
```

el segundo bloque tiene una longitud de 400 bytes, los números marcados indican los siguientes valores:

el número 028A indica la línea.

el número 0030 indica el número de canales utilizados en la grabación.  
 el número 07D0 indica el intervalo de muestreo.  
 el número 0A00 indica el número de muestras contenidas en la cinta, si convertimos este valor a valor decimal obtendremos 2560, si hacemos lo mismo con el valor 0A00 tendremos en decimal 2000 esto multiplicado por 2560 obtendremos 5120 que es el tiempo de grabación.

**Block1 ( File 1, block 3) block size: 10480**

```
0000 0001 0000 0001 0000 03E7 0000 0031 0000 0001 0000 0000 0000 0000 0009 0000
0000 0001 0000 0001 0000 03E7 0000 0031 0000 0001 0000 0000 0000 0000 0009 0000
0000 0000 0000 0000 0000 0000 0000 0001 0000 4120 0000 4140 0000 4160 2121 2221
0000 0000 0000 0000 0000 0000 0000 0001 0000 4120 0000 4140 0000 4160 2121 2221
```

el tercer bloque representa el encabezado de la traza en donde encontraremos secuencial de traza, el punto de tiro, trazas auxiliares, etc). Una vez explicado la función del Dump se procederá al diseño de programas de los distintos formatos de grabación.

### 4.3 DISEÑO DEL PROYECTO

En este punto se refiere a la creación de los programas para la interpretación de los formatos de las cintas y por otra parte las configuraciones de los diferentes sistemas utilizados por la compactación.

La preparación de los programas tanto para los sistemas **RAPPORT** como de los programas en **GEOVECTEUR** para la lectura de las cintas. En los dos sistemas ( **RAPPORT**, **GEOVECTEUR**) se pueden leer las cintas aunque en más recomendable utilizar el sistema **RAPPORT** por la facilidad y localización de las palabras a recuperar.

A continuación se presentarán algunos programas tanto en los sistemas **RAPPORT** como para el sistema **GEOVECTEUR**.

#### PROGRAMA PARA SISTEMA RAPPORT

*JOB	0001	5072	2	6144	
JOURNAL					
*DMX	SEGB	1600	134	48	HDRGAIN
HEADER	GAPPED	132	32		25
ERROR	0	0	ASKOPER		
RECORD	ALL				
	1	2	3	4	

*SEGY0	CART	00001	1750
TRCHDR	TAPETECH		
	SHOTSEQ	INT4	017
	SHOTTRC	INT4	237
<b>EBCHDR</b>			
C01	REGION		: MARIMA
C02	ZONA		: NORTE
C03	PROSPECTO		: F.E.S. C
C04	No DE LINEA		: 2
C05	BRIGADA		: NES-4
C06	EQUIPO		: SERCEL
C07	No DE CANALES		: 48
C08	TIEMPO DE GRA		: 6 SEG
C09	INTERVALO DE MUESTREO		: 2M
C10	FORMATO DE GRABACION		: SEG-B
<b>BINHDR</b>	<b>GENERATE</b>		
	0001	INT2	3255
	0001	INT2	3257
	0001	INT2	3227

\*END.

### PROGRAMA PARA SISTEMA GEOVECTEUR

* LIBRI	SI	01	(M0056-M0058),F1-F100,
LIBRI	FI	01	SEG-D
* LIBRI	BD	01	C00007, (RW),
LIBRI	F0	01	SEG-Y, FOR3, PRELE=1,3,5,39,40
			FWORD19,
			HF(2,2)= 449
			HF(2,4)=94
			HTR5=CGG22(1,32),
			HTR3=CGG2(1,32),
			RL6000,S12,
*BOUCL		1	
*FINBO			
*PROCS			X(B1)

---

Estos dos tipos de programas nos permiten la recuperación de los datos sísmicos, solamente se presentarán algunos ejemplos de dos de los formatos, pero son muchos programas elaborados para los diversos formatos, esta es una de las partes que conforma el proyecto. A continuación se presentarán las configuraciones para los distintos sistemas que conforman la compactación, solamente se presentarán gráficamente las configuraciones.

- Configuración del sistema PECC. ver ( fig. 4.1 )
- Configuración para el sistemas GEOVECTEUR. ver ( fig 4.2 )
- Configuración para el control de calidad en equipo IBM ( plotteado ). ver ( fig 4.3 )
- Configuración para el sistemas SEMIR. ver fig (4.4)

#### **4.5 PRUEBAS DE LA INFORMACION RECUPERADA**

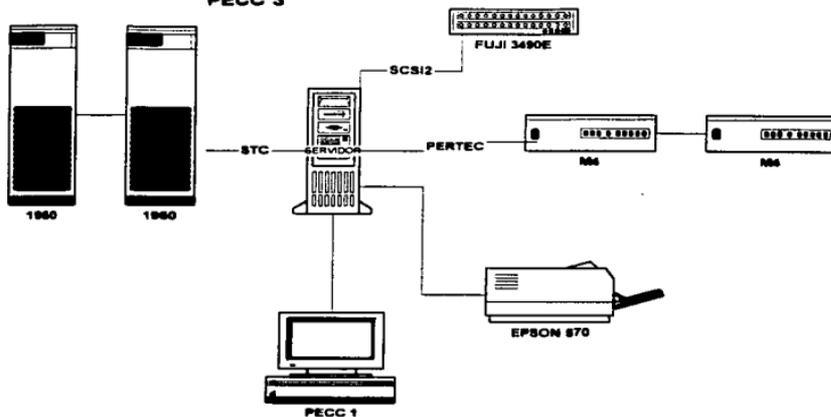
Se necesita comprobar que los programas diseñados realmente este recuperando la información deseada, para ello su comprobación será por medio de gráficas, en donde se revisará los datos indicados en la recuperación, los datos que se revisarán serán:

- Tiempo de grabación.
- Recuperación del número de file de toda la línea
- Recuperación del consecutivo de la traza.
- Intervalo de muestreo.
- Los canales de grabación.

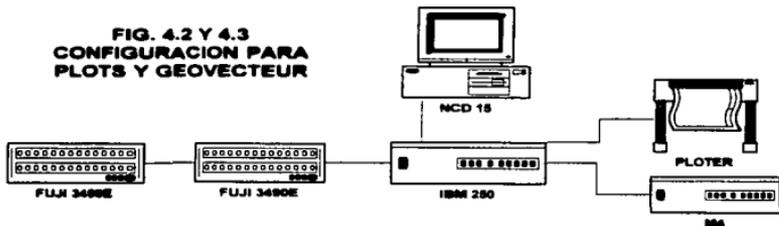
A continuación se muestra gráficamente los datos recuperados en una colección de traza y los puntos de tiro utilizados en la grabación de la línea del prospecto, ver ( Fig 4.5 y 4.6 ). Una vez aprobada la comprobación de la información , el paso siguiente será hacer un respaldo del cartucho original esto se logrará por medio del programa SEMIR. La única función del programa es copiar la información bit a bit del cartucho original a un cartucho 3490E para tener siempre un respaldo de la información, teniendo así una mayor seguridad de la información.

El cartucho original no se utilizará de ninguna manera, solamente el cartucho de respaldo se usará para cualquier tipo de proceso que se solicite. Una vez realizada la copia de respaldo el proceso de comprobación de la información esta completo, el paso siguiente será enviarlo a la base de datos para su captura y procedimientos que se tengan que hacer.

**FIG. 4.1 CONFIGURACION PECC 3**



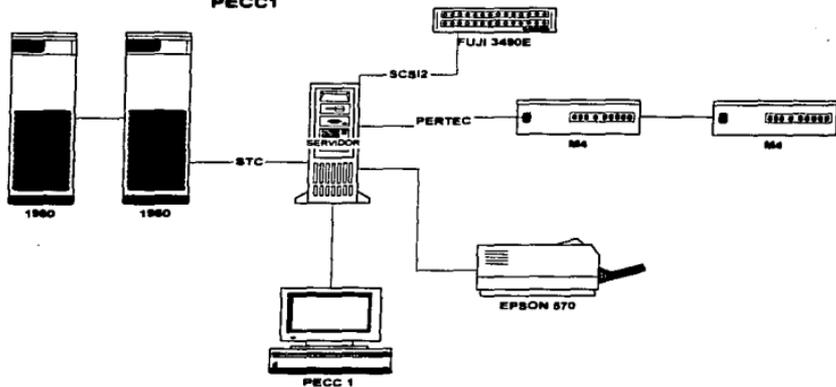
**FIG. 4.2 Y 4.3 CONFIGURACION PARA PLOTS Y GEOVECTEUR**



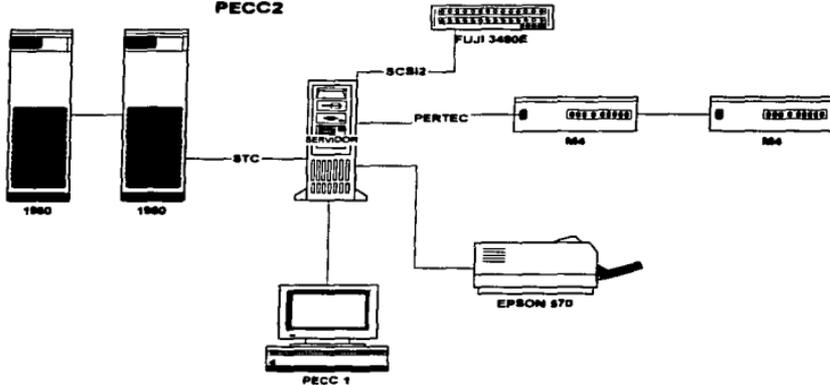
**FIG. 4.4 CONFIGURACION PARA COPIA DE CARTUCHO**



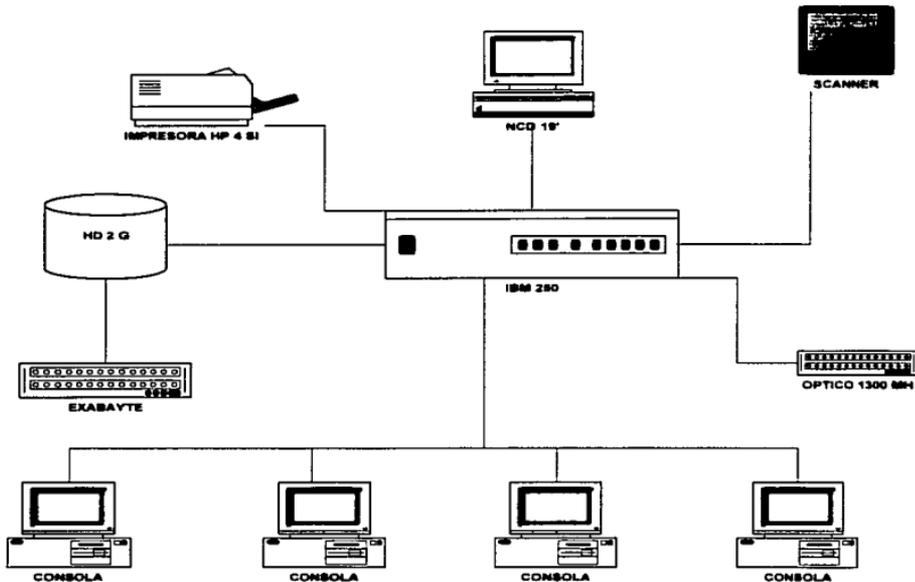
**FIG. 4.1 CONFIGURACION PECC1**



**FIG. 4.1 CONFIGURACION PECC2**



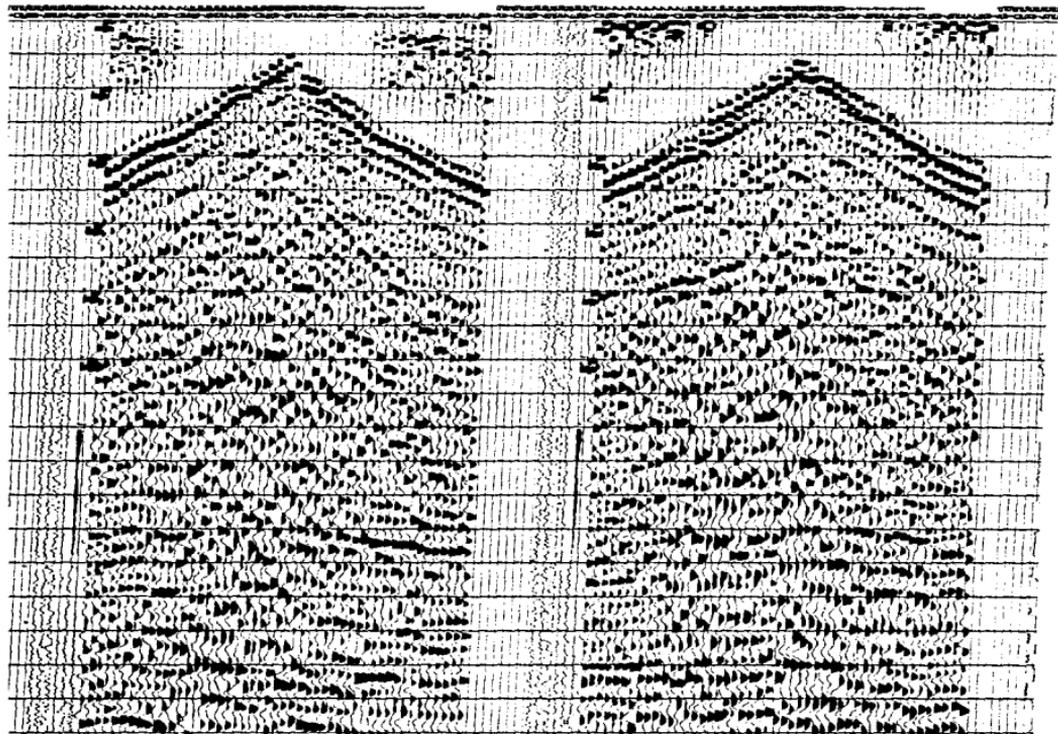
## CONFIGURACION HARDWARE BASE DE DATOS



# COLECCION DE TRAZA FIG. 4.5



PUNTO DE TIRO ( FIG. 4.6)



---

## 4.6 MODELO LOGICO DE LA BASE DE DATOS

Un sistema administrador de base de datos ( DBMS) utiliza un modelo de datos para definir la estructura fundamental de los mismos. Un modelo de datos expresa las entidades y sus relaciones y es la herramienta usada para representar la organización conceptual de los datos; siendo necesario para ello el uso de cierta terminología, que a continuación se describe:

### 1) ENTIDAD

Una entidad es una persona, un lugar, un evento o un objeto identificado en forma única del cual se registra información y que además cae dentro del alcance del sistema. Por lo que éste debe mantener, correlacionar y desplegar información. Las entidades son sustantivas y son representadas por medio de tablas. y pueden ser:

**tangibles.** Líneas, cartuchos, reportes del observador y topografía

**intangibles.** un nombre de actividad, la línea de un prospecto o un concepto abstracto.

### 2) TABLA

Es un arreglo de dos dimensiones compuesto por renglones y columnas. Una columna contiene un tipo de datos y un renglon incorpora similar tipo de información; ambos contienen datos pertenecientes a una entidad. Consideramos las entidades LINEAS, CARTUCHOS, REPORTES DEL OBSERVADOR y TOPOGRAFIA estas son las que se utilizarán en la base de datos.

A continuación se describen las cuatro tablas de la base de datos:

- **Tabla de Líneas:** Esta tabla contendrá todas las líneas de los prospectos así como la región , zona, área , y prospecto que permite identificar en que lugar se realizó la exploración sísmica.
- **Tabla de Cartuchos:** Esta tabla permite identificar que número consecutivo se le asigno a las líneas capturadas en la tabla de líneas por medio de código de barra único para cada cartucho de las líneas.
- **Tabla de Reportes del Observador.** Esta tabla se digitalizarán todos los documentos pertenecientes a cada línea de un prospecto, esto permitirá asociar las imágenes con la líneas capturadas en la tabla de líneas.

- **Tabla de Topografía:** En esta tabla contendrá información de la elevaciones y coordenadas de las líneas de los diferentes prospectos.

#### TABLA DE LINEAS

DOC_REFDOC	VARCHAR2 (8) /*identificador unico de fila
DOC_REFERENCE	VARCHAR2 (100) /* línea pto area zona region
DOC_DATECREATFIC	DATE /* fecha de creación de ficha
DOC_DATEMAJFIC	DATE /* fecha de ultima modificación
LST_RZAPLNN1	VARCHAR2 (8) /* REGION
LST_RZAPLNN2	VARCHAR2 (5) /* ZONA
LST_RZAPLNN3	VARCHAR2 (50) /* AREA
LST_RZAPLNN4	VARCHAR2 (50) /* PROSPECTO
LIGNE	VARCHAR2 (10) /* LINEA
ANNEE	NUMBER (4) /* AÑO INICIO DE PROSPECTO
LIGPTINI	NUMBER (9,1) /* PT INICIAL
LIGPTFIN	NUMBER (9,1) /* PT FINAL
DATE DEB	DATE /* INICIO PROSPECTO
DATE FIN	DATE /* FINAL PROSPECTO
NB CART	NUMBER (3) NUMERO DE CARTUCHOS

#### TABLA DE CARTUCHOS

DOC_REFDOC	VARCHAR2 (8) /* identificador unico de fila
DOC_REFERENCE	VARCHAR2 (50) /* mismo valor de no cartucho
DOC_REVIS	NUMBER (38) /* cuenta de cuantas modificaciones
DOC_DATECREATFIC	DATE /* fecha de creación de ficha
DOC_DATEMAJFIC	DATE /* fecha de ultima modificación
DOC_REFFERE	VARCHAR2 (1) /* línea pto area zona región
NO CARTUCHE	NUMBER (6) /* numero de cartuchos
CARPTINI	NUMBER (9,1) /* pt inicial
CARPTFIN	NUMBER (9,1) /* pt final
CONTRACTOR	VARCHAR2 (25) /* compañía
DATATERR	VARCHAR2 (6) /* CAMPO O MULTIPLEXADO
EQUIPE	VARCHAR2 (9) /* brigada
LONG_LIG	NUMBER (9,3) /* distancia en Km de línea
DISTENTR_PT	NUMBER (7) /* distancia entre pt
NBR_PT_TIR	NUMBER (6,1) /* numero total de pts
LST_DURENRN1	VARCHAR2 (9) /* longitud de grabción
NBR_CART_LIG	NUMBER (3) /* numero de cartuchos
REC_INIT	NUMBER (5) /* file inicial
REC_FINAL	NUMBER (5) /* file final

NB_TROUS	NUMBER (2) /* numero de pozos
LST_PASECHANN1	VARCHAR2 (3) /* intervalo de muestreo
DATE_NEB_ENR	DATE /* inicio de grabación
DATE_FIN_ENR	DATE /* fin de grabación
DATE_CRE	DATE /* grabación del cartucho
INTERFACE	NUMBER (7) /* distancia entre estacas
TX_COUV	NUMBER (5) /* apilado
LST_SOURCEN1	VARCHAR2 (25) /* energia (dinamita, vibrador, etc)
LST_INSTRENRN1	VARCHAR2 (10) /* equipo
LST_NBRCANN1	VARCHAR2 (3) /* canales
LST_GEOSOURCEN1	VARCHAR2 (10) /* arreglo de pozos
LST_GEORECEPTN1	VARCHAR2 (20) /* tendido
LST_ORILOCANN1	VARCHAR2 (4) /* lugar de cartuchos originales
LST_ORILOCANN2	VARCHAR2 (25) /* bodega de cartuchos originales
ORI_NOCANN	NUMBER (4) /* numero de caja de originales
LST_BKPLOCANN1	VARCHAR2 (4) /* lugar de cartuchos copias
LST_BKPLOCANN2	VARCHAR2 (25) /* bodega de cartuchos copias
BKP_NOCANN	NUMBER (4) /* numero de caja de copias

#### TABLA DE REPORTES DE OBSERVADOR

DOC_REFDOC	VARCHAR2 (8) /* identificador unico de fila
DOC_REFERENCE	VARCHAR2 (100) /*
DOC_LABEL	VARCHAR2 (15) /* nombre del disco optico
DOC_ASSOC	NUMBER (38) /* valor 2 cuando hay imagen asociada
DOC_ARCHIV	NUMBER (38) /* valor 1 cuando se grabo en WRK
DOC_DATECREATFIC	DATE /* fecha de creación de la ficha
DOC_DATEMAJFIC	DATE /* fecha de ultima modificación
DOC_REPPERE	VARCHAR2 (100) /* linea pto area zona región
DOC_SAVE	NUMBER (38) /* valor 1 cuando se grabo en SAV
NB_PAGES	NUMBER (6) /* numero de paginas
SP_INIT	NUMBER (9,1) /* pt inicial
SP_FINAL	NUMBER (9,1) /* pt final
OBS_COMMENT	VARCHAR2 (255) /* comentarios

#### TABLA DE ARCHIVOS DE TOPOGRAFIA

DOC_REFDOC	VARCHAR2 (8) /* identificador unico de fila
DOC_REFERENCE	VARCHAR2 (100) /*
DOC_LABEL	VARCHAR2 (15) /* nombre del disco optico
DOC_ASSOC	NUMBER (38) /* valor 1 cuando hay file asociado
DOC_ARCHIV	NUMBER (38) /* valor 1 cuando se grabo en WRK
DOC_DATECREATFIC	DATE /* fecha de creación de la ficha

DOC\_LABEL  
 DOC\_ASSOC  
 DOC\_ARCHIV  
 DOC\_DATECREATFIC  
 DOC\_DATEMAJFIC  
 DOC\_REFFERE  
 DOC\_CLASSE  
 DOC\_SAVE  
 NB\_PAGES  
 ST\_INIT  
 ST\_FINAL  
 TOPO\_COMMENT

VARCHAR2 (15) /\* nombre del disco optico  
 NUMBER (38) /\* valor 1 cuando hay file asociado  
 NUMBER (38) /\* valor 1 cuando se grabo en WRK  
 DATE /\* fecha de creaci3n de la ficha  
 DATE /\* fecha de ultima modificaci3n  
 VARCHAR2 (100) /\* linea pto area zona regi3n  
 VARCHAR2 (15) /\* siempre valor NORMAL  
 NUMBER (38) /\* valor 1 cuando se grabo en SAV  
 NUMBER (6) /\* numero de paginas  
 NUMBER (9,1) /\* pt inicial  
 NUMBER (9,1) /\* pt final  
 VARCHAR2 (255) /\* comentarios

**TABLA DE LINEAS**

Linea	Posic	Geo	Area
1551	Sur	Zs	Tabasco
2001	Marina	Zm	Golfo de M3xico
3050	Norte	Zn	Reynosa

RENGLONES

COLUMNA

**TABLA DE CARTUCHOS**

Cartucho	Linea	Posic	Area
100458	1551	900	3079
500780	2001	500	5045
300600	3050	700	1599

RENGLONES

COLUMNA

**TABLA DE REPORTES DE OBSERVADOR**

<b>Núm. de Páginas</b>	<b>Línea</b>	<b>Area</b>	<b>Proyecto</b>
25	1551	Tabasco	Frontera
30	2001	Golfo de México	Yum
15	3050	Reynosa	Coyame

RENGLONES

COLUMNA

**TABLA DE ARCHIVOS DE TOPOGRAFIA**

<b>Línea</b>	<b>Región</b>	<b>Area</b>	<b>Proyecto</b>
1551	Sur	Tabasco	Frontera
2001	Marina	Golfo de México	Yum
3050	Norte	Reynosa	Coyame

RENGLONES

COLUMNA

Cada renglón de la tabla incluye información perteneciente a una sólo línea y una columna contiene datos del mismo tipo; por ejemplo la columna Línea tiene únicamente el número de la línea de un prospecto, aunque pueden haber líneas repetidas en la tabla, pero que no pertenecen al mismo prospecto y más no datos de diferente tipo. observe que:

- Cada tabla debe tener un nombre particular de tabla
- Cada tabla posee un nombre de columna.
- El nombre de la columna debe ser único dentro de las tablas.
- El orden de columnas y renglones no es significativo.
- Las columnas necesitan ser atómicas ( no divisibles ).

### 3) ATRIBUTOS

Son las características propias de la entidad, la misma posee atributos básicos que la caracterizan. Los atributos se modelan como columnas de la entidad. La forma de diferenciar entidades es por medio de atributos, y cada una de ellas deben tener por lo menos un atributo diferente.

#### 4) TUPLA

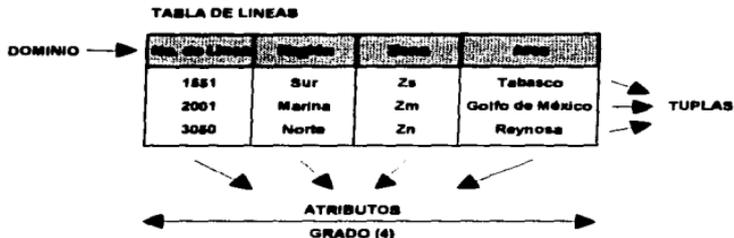
Colección ordenada de uno o más elementos de datos que forman un registro. Es el renglón  $n$  de las tablas.

#### 5) GRADO

Son los diferentes valores que puede tomar la tabla con respecto al dominio, es decir, la amplitud de la tabla en cuanto al número de atributos.

#### 6) DOMINIO

Es la colección de valores de los cuales uno o más atributos obtienen sus valores reales. De acuerdo con las diferentes tablas LINEAS, CARTUCHOS, REPORTE DE OBSERVADOR, y TOPOGRAFIA, la terminología anterior se ubica de la siguiente manera;



### LLAVES

Se denomina llave ( o clave) al atributo que permite significar de manera única a una entidad. Una llave es, en otras palabras, el campo a partir del cual se pueden inferir otros campos de una tabla; por lo que, cada tupla debe estar asociada con una llave que permita su identificación. En la tablas LINEAS, CARTUCHOS, REPORTE DEL OBSERVADOR y TOPOGRAFIA .

- En la tabla de Línea la llave que identifica únivocamente a un elemento de la tabla es el atributo **DOC\_REFDOC**.

- En la tabla de Cartucho la llave que identifica únivocamente a un elemento de la tabla es el atributo **DOC\_REFDOC**.

- En la tabla de Reporte del Observador la llave que identifica únivocamente a un elemento de la tabla es el atributo **DOC\_REFDOC**.

- En la tabla de Topografía la llave que identifica únivocamente a un elemento de la tabla es el atributo **DOC\_REFDOC**.

## LLAVES PRIMARIAS

La llave primaria ( primary key , PK) de una tabla identifica en forma única a cada renglón de la tabla.

En la tabla **LÍNEA**, la PK es el atributo **DOC\_REFDOC**, con el cual se relaciona a una línea en particular.

**TABLA DE LINEA**

DOC_REFDOC	LINE	PROJECT	AREA	ZONE	REGION	FIRST SP	LAST SP
k003245	1000	HIGUERON	TAMPICO	ZN	NORTE	2133.5	2766.5
k003246	1001	HIGUERON	TAMPICO	ZN	NORTE	2820.5	2823.5
k003247	1002	HIGUERON	TAMPICO	ZN	NORTE	2097.5	2001.5
k003248	1003	HIGUERON	TAMPICO	ZN	NORTE	2088.6	2964.5
k003249	1004	HIGUERON	TAMPICO	ZN	NORTE	2085.5	2739.5
k003250	1005	HIGUERON	TAMPICO	ZN	NORTE	2370.5	2656.5
k003251	1006	HIGUERON	TAMPICO	ZN	NORTE	2151.5	2718.5
k003252	1007	HIGUERON	TAMPICO	ZN	NORTE	2910.5	2148.5

PK

En la tabla **CARTUCHOS**, la PK es el atributo **DOC\_REFDOC**, con el cual se relaciona a un cartucho en particular.

### TABLA DE CARTUCHOS

DOC_REFDOC	DOC_REPPER	CARTRIDGE NO	FIRST SP	LAST SP
cd009353	1000    HIGUERON	113990	2133.5	2766.5
cd009354	1001    HIGUERON	113991	2820.5	2823.5
cd009356	1005    HIGUERON	113995	2370.5	2856.5
cd009359	1006    HIGUERON	113996	2151.5	2718.5
cd009380	1007    HIGUERON	113997	2910.5	2148.5
cd009361	1008    HIGUERON	113998	2142.5	2778.5
cd009362	1009    HIGUERON	113999	2754.5	2046.5
cd009355	1002    HIGUERON	113992	2097.5	2001.5
cd009356	1003    HIGUERON	113993	2068.5	2964.5
cd009357	1004    HIGUERON	113994	2085.5	2739.5

PK

En la tabla REPORTE DEL OBSERVADOR, la PK es el atributo DOC\_REFDOC, con el cual se relaciona a un documento de una línea en particular.

### TABLA DE REPORTE DEL OBSERVADOR

DOC_REFDOC	DOC_REPPER	DOC_LABEL	DOC_ASSOC	DOC_ARCHV	DOC_SAVE
ob003227	1-13    HIGUERON	WRKAMAR0001	2	1	1
ob003229	1-14    HIGUERON	WRKAMAR0001	2	1	1
ob003230	1-15    HIGUERON	WRKAMAR0001	2	1	1
ob003231	1-16    HIGUERON	WRKAMAR0001	2	1	1
ob003232	1-17    HIGUERON	WRKAMAR0001	2	1	1

PK

En la tabla TOPOGRAFIA, la PK es el atributo DOC\_REFDOC, con el cual se relaciona a un documento de una línea en particular.

ESTA YESIS NO DEBE  
CALIB DE LA BIBLIOTECA

---

## TABLA DE TOPOGRAFIA

DOC_REFDOC	DOC_REFFERE	DOC_LABEL	DOC_ASSOC	DOC_ARCHIV
ov000734	423    HIGUERON	WRKASUR0002	3	1
ov000735	424    HIGUERON	WRKASUR0002	3	1
ov000736	425    HIGUERON	WRKASUR0002	3	1
ov000737	426    HIGUERON	WRKASUR0002	3	1
ov000738	427    HIGUERON	WRKASUR0002	3	1
ov000739	428    HIGUERON	WRKASUR0002	3	1

← PK →

En una tabla, la llave primaria no puede contener valores nulos ( faltantes o espacios en blanco ) ni valores duplicados, y no se permite cambios sobre los valores de la llave primaria.

Generalmente los valores de la llave primaria son claves o valores que se inventan para poder identificar una tupla específica, los valores que se la asignan a la llave primaria son ajenos a la información de la tupla, estos valores deben cumplir con las características antes mencionadas de las llaves primarias.

## LLAVES FORANEAS

Una llave foránea ( foreign key FK ) es la llave primaria de una tabla, y al mismo tiempo forma parte de otra tabla únicamente como atributo. Retomando las tablas **CARTUCHOS** y **REPORTES DEL OBSERVADOR** , el atributo **DOC\_REFERENCE** se encuentra en ambas tablas; en la tabla **CARTUCHOS** forma parte de atributo y en la tabla **CARTUCHOS** forma parte de ella como un atributo más, es la **PK** de la misma.

El atributo **DOC\_REFERENCE** aparecen en las tablas **CARTUCHOS** y **REPORTES DEL OBSERVADOR** y ambos atributos son del mismo tipo. Generalmente las llaves foraneas ( **FK** ) indican asociaciones entre tablas, por lo que sus valores pueden ser nulos o repetidos.

## ASOCIACIONES

Una asociación es la unión o enlace entre dos o más entidades ( u otras asociaciones ), las cuales se encuentran dentro del enlace del sistema, y por ello el sistema debe mantener, correlaciones y desplegar información .

---

Generalmente las asociaciones requieren de al menos dos entidades ( es decir tablas primas ), estas deben estar dentro del enlace del sistema. Existen tres tipos de asociaciones:

Asociaciones UNO a UNO ( 1:1)

Asociaciones UNO a VARIOS ( 1: M )

Asociaciones VARIOS a VARIOS ( M:M)

Para esta base se utilizará asociaciones UNO a UNO por que no debe repetirse ningún valor con respecto a las líneas de un prospecto y UNO a VARIOS ya que una línea pueden contener más de un cartucho.

#### Asociaciones UNO a UNO ( 1:1)

Las ocurrencias de una entidad se pueden relacionar sólo a una ocurrencia de la otra entidad . Para el siguiente ejemplo las asociaciones se considerarán las Tablas de LINEA y CARTUCHO.

El proposito es proporcionar a cada línea un número consecutivo de cartucho , en la asociación uno a uno teniendo las tablas LINEA y CARTUCHO la asociación se establece de la siguiente manera:



Se interpretaría es la siguiente: Una línea puede tener asignado sólo un número de cartucho y uno de éstos sólo puede ser asignado a una línea; por lo que , su asociación es uno a uno ( a una línea un cartucho y a un cartucho una línea 1:1 ). A continuación se presentan las tablas de las entidades LINEAS y CARTUCHOS y sus respectivos atributos.

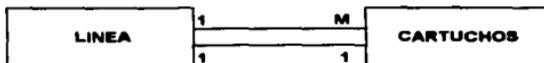
Se emplearán los siglas ND para indicar que no se admiten duplicados. En este ejemplo, es necesario indicarlo puesto que dos líneas no pueden tener el mismo número de cartucho asignado. Una característica importante es que las asociaciones 1:1 son simétricas; es decir, el campo DOC\_REFERENCE es FK en la tabla LINEAS y PK en la tabla CARTUCHOS .

Al modelar este tipo de asociaciones hay que hacerlo cuidadosamente, de manera que los valores nulos se minimicen o se eviten totalmente.

#### ASOCIACIONES UNO A MUCHOS ( 1: M)

---

Se dice que una relación entre entidades es uno a muchos, si las ocurrencias de una entidad están relacionadas con diversas ocurrencias de otra entidad. De acuerdo con las tablas de LINEA y CARTUCHOS, por ejemplo La tabla LINEA tiene n líneas y de igual manera tiene m cartuchos; para llevar el control del las líneas, si se considera las tablas LINEAS y CARTUCHOS, su asociación es la siguiente:



En este caso muchos cartuchos pueden ser asignado a una sola línea y de manera contraria, una sólo cartucho pueden ser asignado a una sola línea, por lo que, aquí la asociación que se crea es de uno a muchos ( 1:M).

El manejo de los datos de las líneas y cartuchos se lleva a cabo a través del paquete GEODOC. De este se utilizan dos módulos GEOBASE y GEOVIEW.

El primero permite la adquisición de la información y la ordena en una base de datos, permite la consulta de las mismas, permite la impresión de la información de las líneas y cartuchos así como también permite la asociación de las imágenes de los reportes del observador y los archivos que contienen las elevaciones y coordenadas con las líneas correspondientes. También permite la grabación de las imágenes y los archivos ya asociados a los discos ópticos.

GEOVIEW es la herramienta con la cual se manejan imágenes de los reportes del observador. Con éste se maneja el scanner, se visualizan las imágenes, se pueden borrar e imprimir documentos.

La base de datos en la cual se ordena la información de las líneas y de los cartuchos es ORACLE. La programación con la cual se especifica la forma como se almacenan los datos en la base de datos se lleva a cabo con la ayuda de SQL, es un lenguaje para acceso a la información almacenada en bases de datos relacionales.

---

## **CONCLUSIONES**

Con las ayuda de la base de datos se tendrá un mejor control y manejo de grandes volúmenes de la información, la clasificación de los distintos prospectos de acuerdo a la región, zona y área geográfica se asignaron por medio de un código de barra, con este código identificará exclusivamente a un sólo cartucho de un determinado prospecto. Lo cual evitará que un código se duplique en otra línea o en otro prospecto.

En cuanto al cambio de formato será de gran utilidad tener toda la información en un sólo formato de grabación , evitando con esto que los Geofísicos pierdan tiempo en analizar la información contenida en la cinta para poder interpretar el tipo de formato, con esta transformación de formato los analistas geofísicos harán los programas enfocados a el formato secuencial SEG-Y.

---

## **REFERENCIAS BIBLIOGRAFICAS**

1. **James A. Seen** " Análisis y Diseño de Sistemas de Información " Ed. Mc Graw Hill , Segunda Edición, Pág 620-625.
2. **Donald H. Sanders** " Informática Presente y Futuro" Ed. Mc Graw Hill, Tercera Edición, Pag. 257-282.
3. **Marisela B. Vázquez** " Introducción a las Bases de Datos " Guías y Textos de ómputo" Primera Edición, Pag. 9-13.
4. **Atre Shakuntala** " Técnicas de Bases de Datos " Ed. Trillas, Segunda Edición, Pag. 117-132.
5. **Gillenson L. Mark** " Database Step By Step " Segunda Edición, New York, 1994.
6. **Ballesteros N. Etal** " Sistema de Diagnóstico Orientado a la Calidad en el Desarrollo de Sistemas de Información" Primera Edición, UPIICSA.
7. **Baca U. Gabriel** " Evaluación de Proyectos " , Ed. Mc Graw Hill, Segunda Edición, México, 1994.