



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO  
CENTRO UNIVERSITARIO UAEM-TEXCOCO

RODRÍGUEZ LEÓN KARLA MELISSA

TITULO:

ANÁLISIS DE LAS BASES DE DATOS DISTRIBUIDAS COMO UNA  
ALTERNATIVA PARA LOS SISTEMAS ERP EN LAS EMPRESAS

DIRECTOR(A) DE LA TESINA:

M. EN C. NIÑO MEMBRILLO YEDID ERANDINI

FECHA DE PRESENTACIÓN:

AGOSTO DEL 2013

## **AGRADECIMIENTOS**

A dios por darme a una familia que me ha motivado y apoyado a seguir adelante con los estudios a nivel licenciatura y por brindarme siempre la posibilidad de superarme.

Así mismo le agradezco a mi directora y a mis revisores de tesina por haberme proporcionado ayuda y observaciones para mejorar este trabajo de investigación

## **Índice**

Lista de cuadros y Figuras	1
Antecedentes	9
Introducción	11
Objetivos	
Objetivo general	12
Objetivos particulares	12
Materiales y métodos	13
<b>Capítulo I Sistemas ERP</b>	
1.1 Conceptos básicos	
1.1.1 Proceso de negocio	15
1.1.2 Procesos de negocio en las empresas	16
1.1.3 Sistema ERP	18
1.2 Importancia de los sistemas ERP en las compañías	25
1.3 Elementos de sistemas ERP	26
1.4 Funcionamiento de los sistemas ERP en las organizaciones	28
1.5 Gestión de procesos de negocios y reingeniería	30
1.6 Arquitectura Cliente/servidor	32
1.7 Principales aplicaciones del ERP	35
1.8 Proveedores de los ERP	37

1.9 Ventajas y Desventajas	42
----------------------------	----

## **Capítulo II** Bases de datos relacionales

2.1 Conceptos básicos	46
-----------------------	----

2.2 Normalización	53
-------------------	----

2.3 Sql	61
---------	----

2.4 Ventajas y Desventajas	66
----------------------------	----

### 2.5 Bases de datos centralizadas

2.5.1 Conceptos básicos	68
-------------------------	----

2.5.2 Uso de estas bases de datos en las Organizaciones	69
---	----

2.5.3 Sistemas de administración y bases de datos	70
---	----

2.5.4 Estructura	71
------------------	----

2.5.5 Almacenamiento	73
----------------------	----

2.5.6 Consultas	76
-----------------	----

2.5.7 Disponibilidad	81
----------------------	----

### 2.6 Bases de datos distribuidas

2.6.1 Conceptos básicos	82
-------------------------	----

2.6.2 Uso de estas bases de datos en las Organizaciones	83
---	----

2.6.3 Sistemas de administración de bases de datos	87
--	----

2.6.4 Estructura	89
------------------	----

2.6.5 Almacenamiento	90
----------------------	----

2.6.6 Consultas	91
-----------------	----

2.6.7 Disponibilidad	98
----------------------	----

**Capítulo III** Análisis y comparación, las bases de datos centralizadas contra las bases de datos distribuidas en los sistemas ERP.

3.1 Consultas	106
3.2 Disponibilidad	142
3.3 Ventajas y desventajas	144
Resultados, Conclusiones y recomendaciones	146
Glosario	155
Referencias bibliográficas	160



## **Lista de Cuadros y Figuras**

Figura 1.- Arquitectura Básica De Un Sistema ERP	20
Figura 2. Elementos y/o módulos básicos de un ERP	27
Figura 3.- Como funciona un sistema empresarial	28
Figura 4. Los ERP integración horizontal de los procesos operativos de la empresa	30
Figura 5. Modelo cliente/servidor	33
Figura 6. Gráfica de Senior Blog	36
Figura 7. Gráfica de Software Guru	39
Figura 8. El mercado del ERP	40
Figura 9; Magic Quadrant for Single-Instance ERP for Product-Centric Midmarket Companies	40
Figura 10.- Los tres niveles de abstracción de datos	47
Figura 11; Empleado	50
Figura 12. Pasos para el diseño de una base de datos.	51
Figura 13; Empleado	55
Figura 14; Empleado.	55
Figura 15; Empleado.	55
Figura 16; Domicilio	56
Figura 17; Empleado	57

Figura 18; Domicilio	57
Figura 19; Empleado	58
Figura 20; Domicilio	58
Figura 21; Departamento	58
Figura 22; Rel_Emp_Dep	58
Figura 23; Empleado	59
Figura 24; Empleado.	60
Figura 25; Localidad.	60
Figura 26; modelo relacional.	61
Figura 27; Base de datos centralizada monousuario.	71
Figura 28; Base de datos centralizada multiusuario.	72
Figura 29; Arquitectura de 3 niveles.	72
Figura 30; Sectores, clusters y pistas	74
Figura 31; cilindros	74
Figura 32; Árbol B+.	76
Figura 33; Consulta 1.	77
Figura 34; Consulta 2.	78

Figura 35; Consulta 3.	78
Figura 36; Consulta 4.	79
Figura 37; Consulta 5.	80
Figura 38; Consulta 6.	80
Figura 39; Consulta 7.	81
Figura 40; Sucursal	84
Figura 41; RSucuCuenta.	84
Figura 42; Cuenta.	84
Figura 43; RCuentaCliente.	84
Figura 44; Cliente.	84
Figura 45; Fragmento horizontal de Monterrey	85
Figura 46; Fragmento horizontal del Estado de México	85
Figura 47; Fragmento vertical de Monterrey	85
Figura 48; Fragmento vertical del Estado de México	85
Figura 49; Fragmento HV1 de Monterrey (S1)	86
Figura 50; Fragmento HV1 del Estado de México (S2)	86
Figura 51; Fragmento HV2 de Monterrey	86

Figura 52; Fragmento HV2 del Estado de México	86
Figura 53; Fragmento VH1 de Monterrey	87
Figura 54; Fragmento VH2 de Monterrey	87
Figura 55; Fragmento VH del Estado de México	87
Figura 56; Fragmento VH del Estado de México	87
Figura 57; Base de datos distribuida	89
Figura 58; Arquitectura para bases de datos distribuidas.	90
Figura 59; Replica de datos.	91
Figura 60; RSucuCuenta de Monterrey (RC1)	95
Figura 61; RSucuCuenta del Estado de México (RC2)	95
Figura 62; Cuenta de Monterrey (CUENTA1)	95
Figura 63; Cuenta del Estado de México (CUENTA 2)	95
Figura 64; Cuenta del cliente de Monterrey (CC1).	96
Figura 65; Cuenta del cliente del Estado de México (CC2)	96
Figura 66; Clientes de Monterrey (CLIE1).	96
Figura 67; Clientes del estado de México (CLIE2).	96
Figura 68. Zope Architecture Diagram	104

Figura 69; Descripción de atributos	109
Figura 70; Base de datos centralizada de personas y organizaciones	110
Figura 71; Base de datos distribuida de personas y organizaciones de México	111
Figura 72; Base de datos distribuida de personas y organizaciones de Monterrey	112
Figura 73; Dat_Per	113
Figura 74; Per_Rol	113
Figura 75; Rol_Per.	113
Figura 76; Per_Hab	114
Figura 77; Hab_Per	114
Figura 78; Per_Pro_Int.	114
Figura 79; Pro_Int_Per	114
Figura 80; Per_Nac	114
Figura 81; Nac_Per.	114
Figura 82; Per_Org	114
Figura 83; Dat_Org	115
Figura 84; Org_Rol.	115
Figura 85; Rol_Org.	115

Figura 86; Org_ Hab.	115
Figura 87; Hab_Org.	115
Figura 88; Cod_Act_Org	115
Figura 89; Dat_Per de México (R1Mex)	116
Figura 90; Dat_Per de Monterrey (R1Mon)	117
Figura 91; Per_Rol de México (R2Mex).	117
Figura 92; Per_Rol de Monterrey (R2Mon).	117
Figura 93; Rol_Per de México (R3Mex).	117
Figura 94; Rol_Per de Monterrey (R3Mon).	117
Figura 95; Per_Hab de México (R4Mex).	118
Figura 96; Per_Hab de Monterrey (R4Mon).	118
Figura 97; Hab_Per de México (R5Mex).	118
Figura 98; Hab_Per de Monterrey (R5Mon).	118
Figura 99; Per_Pro_Int de México (R6Mex).	118
Figura 100; Per_Pro_Int de Monterrey (R6Mon).	118
Figura 101; Pro_Int_Per de México (R7Mex).	118
Figura 102; Pro_Int_Per de Monterrey (R7Mon)	118

Figura 103; Per_Org de México (R8Mex).	118
Figura 104; Per_Org de Monterrey (R8Mon).	118
Figura 105; Per_Nac de México (R9Mex).	119
Figura 106; Per_Nac de Monterrey (R9Mon).	119
Figura 107; Nac_Per de México (R10Mex).	119
Figura 108; Nac_Per de Monterrey (R10Mon).	119
Figura 109; Dat_Org 1 de Mexico (R11Mex1)	119
Figura110; Dat_Org 2 de Mexico (R11Mex2)	120
Figura 111; Dat__Org 1 de Monterrey (R11Mon1)	120
Figura 112; Dat_Org 2 de Monterrey (R11Mon2)	120
Figura 113; Org_Rol de México (R12Mex).	120
Figura 114; Org_Rol de Monterrey (R12Mon).	120
Figura 115; Rol_Org de México (R13Mex).	121
Figura 116; Rol_Org de Monterrey (R13Mon).	121
Figura 117;Org_Hab de México (R14Mex).	121
Figura 118;Org_Hab de Monterrey (R14Mon).	121
Figura 119; Hab_Org de México (R15Mex).	121

Figura 120; Hab_Org de Monterrey (R15Mon).	121
Figura 121; Cod_Act_Org de México(R16Mex).	121
Figura 122; Cod_Act_Org de Monterrey (R16Mon).	121
Figura 123;Resultado de la consulta 1 de la base de datos centralizada	123
Figura 124; Resultado de la consulta 2 de la base de datos centralizada	124
Figura 125; Resultado de la consulta 3 de la base de datos centralizada	126
Figura 126; Resultado de la consulta 4 de la base de datos centralizada	128
Figura 127; Resultado de la consulta 5 de la base de datos centralizada	130
Figura 128; Resultado de la consulta 6 de la base de datos centralizada	133
Figura 129; Resultado de la sonsulta 7 de la base de datos centralizada	140
Figura 130; Cuadro comparativo	142
Figura 131; Cuadro comparativo de consultas	145

## **Antecedentes**

Se realizó un estudio para la implantación de un ERP en la Escola Técnica Superior d'Enginyeria (etse) por Romo (2008) “para una empresa de transportes en el 2008, esta investigación esta titulada como estudio para la implantación de un ERP en una empresa de transporte” surge porque una pyme de transporte requiere implantar un ERP.

Se habla sobre el ERP que de forma general se define como un software que está compuesto por módulos (básicos, opcionales, verticales) que integran las actividades que tiene una empresa, este sistema está automatizado y usa una base de datos compartida.

También se habla sobre los posibles módulos como es el caso de contabilidad y finanzas, así como de las características del ERP como: Integral, modular, adaptable, flexible, una base de datos centralizada, entre otras.

Para posteriormente explicar las características; ventajas; desventajas; recomendaciones para ver si una empresa necesita un ERP; se estudian los fabricantes de éstos; metodología para un ERP; se estudia la mejor selección; lo que conlleva la implantación, costos y se estudian tres sistemas ERP (My M Soft ERP, Easy Tech, Ekon Distributor), y finalmente se estudia la empresa para percibir el sistema ERP que necesita.

El ERP que mejor se adapta es My M Soft ERP pues es la que mejor se adapta a las necesidades de la pyme.

Del mismo modo se tomó en cuenta la investigación que lleva por título Estudio comparativo de paquetes ERP de Cuenca y Boza (2006) para estudiar los ERP ofertados de manera que a la empresa la auxilia a la selección de un ERP por lo que se hicieron comparaciones de paquetes de

ERP como SAP, entre otros; dando recomendaciones para el tamaño de empresa y el ámbito, así como exponer las plataformas y las bases de datos que usan.

“Por otro lado en la investigación ventajas y retos en el uso de bases de datos distribuidas del 2002, se estudian las ventajas y desventajas de las bases de datos distribuidas ya que muchas empresas se diversifican geográficamente. Se habla de ventajas en cuanto a la economía, interconexión, rendimiento, confiabilidad, disponibilidad, procesamiento, entre otros; llegando a la conclusión de que las bases de datos distribuidas permiten incorporar recursos de cómputo poco a poco pero que existen áreas que están en investigación” (López, 2002, p.3).

Atho (2006). “En el documento bases de datos distribuidas habla sobre el concepto, las características, arquitectura, procesamiento de consultas y diferencias entre una base de datos centralizada y una distribuida, en las conclusiones se dice que las bases de datos distribuidas encajan bien en las empresas y su uso ha incrementado, así mismo se dice que tanto las bases de datos centralizadas como las distribuidas se basan en hechos comunes y cercanos, entre otros.”

“En el trabajo de estudio para la implementación de un ERP en una empresa de aduanas” de la Escola Técnica Superior d’Enginyeria (etse) realizado en el 2009 se le asesora a una empresa aduanera para encontrar el ERP que mejor se adapta a sus necesidades y lo pueda usar, viendo la viabilidad del proyecto, información de la empresa, las soluciones ERP y la comparación de éstos, sugiriendo finalmente Microsoft Dynamics Navision para la empresa aduanera” (Díaz, 2009, p.1-70).

Como se aprecia en los antecedentes los sistemas ERP son cada vez más usados por las organizaciones para la eficiencia de las mismas pues permiten la comunicación de toda la organización en un mismo sistema. De igual forma se habla sobre bases de datos distribuidas, las cuales tienen beneficios para su implementación en las organizaciones. Es por eso que estos antecedentes fungen de punto de partida para la realización de la tesina.

## **Introducción**

Las empresas actualmente trabajan con sistemas de información para agilizar sus procesos, tales como los sistemas de información gerencial, de apoyo a las decisiones, los sistemas ERP, entre otros.

En este caso la investigación se enfoca en los sistemas de planificación de recursos empresariales (ERP), que son sistemas que integran y manejan los procesos de negocios (manera de coordinar y organizar actividades para elaborar un producto) que hay en las áreas funcionales de una organización; estos sistemas les proporcionan a las empresas mejor comunicación entre sus departamentos y mayor fiabilidad en la información por lo cual son de gran importancia para las organizaciones, ya que además esto hace que la productividad de la organización aumente significativamente.

La presente investigación se llevó a cabo con el fin de mostrar el uso de las bases de datos en los sistemas ERP.

Abordando información de los sistemas ERP tal como: Definición, funcionamiento, características, ventajas, desventajas, y posibles problemas que pudiera presentar. Aunado a esto se analizarán las bases de datos centralizadas contra las bases de datos distribuidas para identificar los problemas específicamente en cuanto a las consultas y la disponibilidad en el uso

de las bases de datos centralizadas con la finalidad de proponer el uso de las bases de datos distribuidas como una alternativa en los sistemas ERP, tomando como base el sistema ERP5.

## **Objetivos**

### **Objetivo general**

Analizar las bases de datos distribuidas, para identificar los elementos y detallar sus características, mejorías y deterioros del uso de estas bases de datos como alternativa en el desarrollo de un sistema ERP.

### **Objetivos particulares**

- Realizar una investigación bibliográfica de las bases de datos centralizadas y las bases de datos distribuidas en los sistemas ERP que sirva para la realización del marco teórico de la tesina.
- Comparar las bases de datos distribuidas y las bases de datos centralizadas considerando las consultas y la disponibilidad para seleccionar la opción más factible en los sistemas ERP.
- Generar recomendaciones para el uso de las bases de datos distribuidas como alternativa en el funcionamiento de los sistemas ERP.

## **Materiales y métodos**

Se compiló información en diversas fuentes bibliográficas, posteriormente se analizó, evaluó, comparó y luego se dieron a conocer los resultados de la viabilidad de la idea, para conseguir esto, la metodología se fragmentó en fases en las que se expresaron de una forma un poco más detallada lo que se llevó a cabo. A continuación se exponen las fases:

### Fase 1 Recopilación y Exploración

En esta fase se hizo una recolección de información sobre temas importantes para el desarrollo de la tesina como los sistemas ERP, bases de datos centralizadas y bases de datos distribuidas para posteriormente hacer un estudio a fondo con ayuda de libros, artículos y otras fuentes de información. La información obtenida fue la base para el desarrollo del marco teórico que apoyó en la realización de la tesina.

### Fase 2 Confrontación

Etapas en las que se evaluaron las consultas y la disponibilidad en las bases de datos centralizadas como de las bases de datos distribuidas para expresar sus ventajas y desventajas con la finalidad de realizar una comparación de las mismas en los ERP.

### Fase 3 Resultados y sugerencias

Con base a la evaluación se explicaron los resultados de la investigación y se propuso una nueva alternativa dado los siguientes puntos:

- Evaluación del uso de bases de datos distribuidas en los sistemas ERP
- Propuesta o sugerencia de una nueva opción para el funcionamiento de los sistemas ERP

Se dieron las conclusiones y las recomendaciones.

Para la obtención de la información de este documento que sustente la investigación de las bases de datos distribuidas como alternativa en los sistemas ERP se utilizaron libros de sistemas de información y libros de bases de datos de la Universidad Autónoma del Estado de México, Centro Universitario UAEM-Texcoco; así como libros, tesis, artículos, revistas en línea y sitios web con fin de poder complementar de manera correcta la investigación.

## **Capítulo 1 sistemas ERP.**

### **1.1 Conceptos básicos**

#### **1.1.1 Procesos de negocio**

Todas las organizaciones cuentan con procesos de negocio para llevar a cabo sus actividades diarias y mantener el orden y control de las mismas; para entender mejor los procesos de negocio es necesario primero definir un proceso.

Un proceso es un “Conjunto de tareas relacionadas entre sí en forma lógica y con cuyo desempeño se busca obtener un resultado definido” (Stair, Reynolds, 2000, p.6). El autor Hitpass (2012) lo define como “Una concatenación lógica de actividades que cumplen un determinado fin, a través del tiempo y lugar, impulsados por eventos.”

Para Alarcón (s.f) es un “Conjunto de actividades que, con un input recibido, es capaz de crear un producto de valor para el cliente.

Dichas actividades están relacionadas unas con otras e integradas de manera coherente y terminan con la entrega del producto o servicio al cliente. Necesitan expresamente la entrega de un producto o servicio.

Por tanto, los departamentos funcionales no constituyen un proceso, aunque realizan una misma tarea que alimenta a varios procesos. No es por tanto, un análisis de tareas”, cabe resaltar que un input suele traducirse como una entrada u aportación.

Debido a lo anterior se puede decir que un proceso es un conjunto de actividades o pasos a seguir para llegar a un fin. Por ejemplo, en el área de recursos humanos se lleva a cabo el proceso de negocio reclutamiento de personal, para ello es necesario realizar diversas actividades tales como

sacar la convocatoria de empleo (indicando el perfil que se requiere), solicitar el curriculum, validar perfil y curriculum, realizar pruebas psicométricas (para evaluar la personalidad y la aptitud) y hacer la entrevista que se puede realizar en dos partes, es decir por el encargado de la selección (psicólogo) y por el responsable de la empresa que contrata; estos pasos forman parte del proceso de negocio antes mencionado.

Un proceso de negocios para Laudon y Laudon (2008) es:

la manera en que se organiza, coordina y orienta el trabajo para producir un producto o servicio valioso. Los procesos de negocios son flujos de trabajo de materiales, información y conocimiento, es decir, conjunto de actividades. Los procesos de negocios también se refieren a las maneras particulares en que las empresas coordinan el trabajo, la información y el conocimiento, y a las formas en que la administración elige coordinar el trabajo. Todas las empresas se pueden considerar como conjunto de procesos de negocios. Algunos de estos procesos son parte de procesos más grandes.

Así que un proceso de negocios es un conjunto de actividades que se relacionan con el fin de conseguir un resultado en el negocio. Si una empresa tiene bien definidos sus procesos y se ejecutan correctamente, la productividad de la misma incrementará, lo que probablemente le brinde una ventaja competitiva.

### **1.1.2 Procesos de negocios en las empresas**

Los procesos de negocios son importantes en las organizaciones, como ya se mencionó tienden a incrementar la productividad de la misma; por tanto, mejoran su rendimiento y su eficiencia.

En las organizaciones en muchas ocasiones se piensa que el área de producción, de ventas u otras son como tales procesos de negocios. Sin embargo, no es así, ya que son áreas (o departamentos) funcionales de una empresa; es decir, X organización cuenta con un área de ventas y mercadotecnia, de finanzas, producción, entre otras, pero cada área tiene asociadas diversas actividades o funciones para cumplir con un objetivo que es el proceso de negocios. Por ejemplo, el departamento de ventas y marketing lleva a cabo la identificación de clientes, dar a conocer el producto al consumidor y vender el producto.

Tomando como base el ejemplo de la empresa X, supóngase que ha disminuido sus ingresos en los últimos tres meses, y tras hacer diversos análisis para conocer el motivo se llega a la conclusión de que el problema se está presentando en el área de ventas y marketing porque no se vende cierto producto, entonces se detecta que se debe a que no se identificó de manera correcta a los clientes, para llevar a cabo este proceso de negocios se tuvo que pensar hacia quien iba dirigido el producto (por ejemplo, si el producto fuesen unas papas de dieta, el cliente principal serían aquellas mujeres que no tienen un peso ideal); qué es lo que el cliente busca (si ese cliente desea no subir de peso o bajar de peso); cuáles son los problemas que tiene (porqué el cliente necesita unas papas de dieta, si es por salud o autoestima) y qué necesita para satisfacer estos problemas (si en verdad necesita comida de dieta como son las papas); con ello se percibe que si un proceso de negocio no se encuentra realizado correctamente, como en este caso que la identificación de clientes no fue correcta (porque en un inicio se pensó que los clientes eran personas diabéticas y no aquellas mujeres que no tienen peso ideal), puede afectar a toda la organización. Por ello, es importante que las empresas tengan bien definidos sus procesos de negocios.

### 1.1.3 Sistemas ERP.

Existen diversas definiciones de los sistemas de planificación de recursos empresariales (ERP), a continuación se presentan algunas de ellas:

“ERP son los sistemas de información que tratan de automatizar, de forma horizontal, el conjunto de procesos operacionales<sup>1</sup> o reflejo de la actividad de la organización (finanzas, compras, ventas, almacén, etc.)” (Giner, 2004, p. 114). Cabe resaltar que el autor al momento de hablar de que un ERP automatiza de forma horizontal los procesos operacionales hace referencia a que los procesos se manejan a la par, es decir, se puede tener el proceso de ventas y el de compras, y éstos se encuentran relacionados, por lo que trabajan en conjunto.

Muñis (2004) define ERP como “un sistema de planificación de los recursos y de gestión de la información que, de una forma estructurada, satisface la demanda de necesidades de la gestión empresarial. Se trata de un programa de software integrado que permite a las empresas evaluar, controlar y gestionar más fácilmente su negocio en todos los ámbitos”.

de Pablos, Lopez-Hermoso, Martín-Romo y Medina (2004) mencionan que “Un ERP Enterprise Resource Planning es un sistema integrado de gestión que cubre de forma total diferentes áreas funcionales de una empresa. Es un sistema de información estructurado que permite la gestión integral de las diversas áreas empresariales.”

Laudon y Laudon (2008) dicen que

---

<sup>1</sup> Son los que se relacionan con las operaciones (cantidad de actividades o tareas que se necesitan para realizar un determinado producto) que se llevan a cabo en cada una de las áreas o departamento de la empresa. Entre ellos se encuentran aquellos que generan valor en el cliente, es decir, asociados a la cadena del valor (identifica las principales actividades que crean un valor para los clientes y las actividades de apoyo relacionadas; permite identificar los costos en que incurre una organización, entre otros) del producto o servicio.

Los sistemas empresariales integran los procesos de negocios<sup>2</sup> clave de toda una empresa en un sólo sistema de software que permita un flujo transparente de la información a través de la organización. Estos sistemas se enfocan principalmente en procesos internos, aunque podrían abarcar transacciones con clientes y proveedores.

los sistemas empresariales, también conocidos como sistemas de planeación de recursos empresariales (ERP), los cuales se basan en un conjunto de módulos<sup>3</sup> de software integrados y en una base de datos<sup>4</sup> central común. La base de datos recopila información procedente de muchas divisiones y departamentos diferentes de una empresa, y de procesos de negocios clave relacionados con la manufactura y la producción, las finanzas y la contabilidad, las ventas y el marketing, así como los recursos humanos, y la pone a disposición de las aplicaciones que apoyan a casi todas las actividades de negocios internas de una organización.

La Figura 1 muestra los componentes de un ERP, los cuales se encuentran comunicados a una base de datos central, por lo tanto, se puede compartir información entre los distintos componentes (front office, back office, cadena de suministro), en el front office se encuentran los distintos módulos (procesos de negocios) que tienen cierta relación con el cliente (consumidor). Por el contrario, en el back office se aprecia que no se cuenta con una comunicación directa con el consumidor, y en el esquema también se maneja la cadena de suministros que tiene que ver con el material que se necesita para la elaboración de un producto o servicio hasta la distribución del mismo.

---

<sup>2</sup> Manera especial de coordinar y organizar las actividades laborales. Información y conocimiento para elaborar un producto o servicio.

<sup>3</sup> Son un software que agrupa subprogramas en un área, estos módulos permiten integrar las distintas operaciones en las diferentes áreas.

<sup>4</sup> Conjunto de datos almacenados sin redundancias innecesarias en un soporte informático y accesible simultáneamente por distintos usuarios y aplicaciones.

Un sistema ERP es una solución informática integral que está formada por unidades interdependientes denominadas Módulos: Los primeros y fundamentales son los denominados Módulos Básicos, de adquisición obligatoria, y alrededor de los cuales se agregan los otros módulos opcionales, que no se adquieren obligatoriamente y se agregan para incorporar nuevas funciones al sistema ERP. También existen los llamados Módulos verticales y corresponden a módulos opcionales diseñados específicamente para resolver las funciones y procesos del negocio de un sector económico específico. (Benvenuto, 2006, Vol 4, p.38)

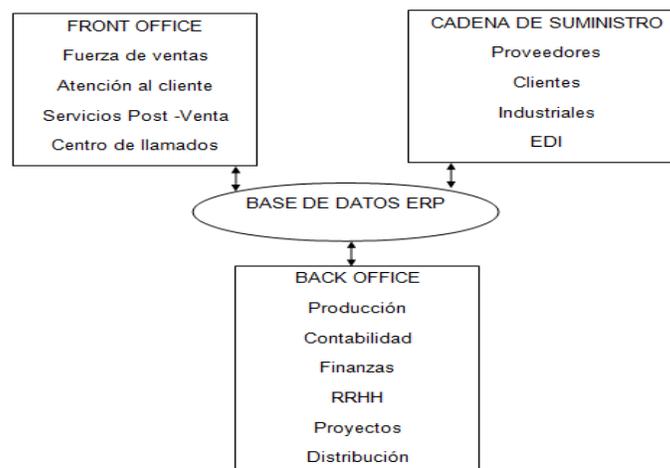


Figura 1.- Arquitectura Básica De Un Sistema ERP; Benvenuto,(2006).

Existen características primordiales las cuales distinguen a un ERP de cualquier otro software empresarial y son las siguientes:

- **Integrales:** Ya que permiten controlar los diferentes procesos de la compañía (todos los departamentos de una empresa se relacionan entre sí). Por ejemplo, en una empresa, si un cliente realiza un pedido, implica que se cree una venta, lo que lleva al proceso de producción, control de inventarios, planificación de distribución de cierto producto, entre otros. Si la organización no usara un sistema empresarial, tendría que tener varios

programas que intervengan en estos procesos, sin embargo con un ERP, el encargado de ventas sólo tendría que capturar el pedido y el sistema lleva a cabo todo lo demás; con ello se conseguiría que la información no sea manejada por varias personas, así que de cierta forma se contaría con información protegida.

- **Adaptables:** Los ERP se pueden adaptar a cualquier empresa, independientemente del sector al que se dediquen y de las peculiaridades de sus procesos de negocios. Esto se consigue a través de la parametrización<sup>5</sup>(configuración) de los procesos. Por ejemplo, en una organización, en el control de inventarios probablemente se requiera manejar la partición de lotes, sin embargo, otra organización puede no requerirla. Los ERP más adelantados permiten el desarrollo rápido de nuevos procesos.
- **Modularidad:** Cuentan con distintos módulos (partes de aplicación) que se pueden ir implementando continuamente, entonces los ERP están formados por una cantidad de módulos determinados, los cuales son independientes unos de otros, pero al mismo tiempo se mantienen comunicados. Algunos módulos básicos de estos sistemas son: producción, ventas, mercadotecnia, contabilidad financiera, contabilidad de gestión, gestión del proyecto, gestión del flujo de trabajo, logística y recursos humanos.
- **Independencia:** Cada módulo es independiente en su funcionamiento.
- **Estandarización:** Los módulos son compatibles pues se desarrollan en el mismo entorno de programación.

Hay otros autores que precisan que en los ERP existen diferencias, pero presentan ciertas características comunes, entre las que se pueden señalar:

---

<sup>5</sup> La parametrización es el valor añadido fundamental que se debe hacer con cualquier ERP para adaptarlo a las necesidades concretas de cada empresa.

- **Sentido global:** Buscan satisfacer las necesidades de la empresa en cuanto a su información.
- **La naturaleza estándar de la solución con la disposición de un entorno de desarrollo propio:** Ayuda a que la empresa realice las adaptaciones que necesita para llevar a cabo su actividad.
- **Base de datos centralizada:** Múltiples usuarios pueden acceder a ella y evita problemas (redundancia, e integridad).
- **Integridad de datos:** Ya que los datos son ingresados una vez en el sistema, se debe tratar de que sean consistentes; es decir, que no se tengan distintas versiones en diferentes lugares del sistema, además deben ser completos y comunes para reducir los errores.
- **Reingeniería de Proceso:** En ocasiones no se necesita llevar a cabo, pero algunas empresas que van a implantar un ERP se ven en la necesidad de cambiar algunos de sus procesos para que puedan estar alineados con los del sistema.
- **Es un todo:** Si bien un sistema empresarial puede contar con menús modulares configurables para cada usuario, es un solo programa que tiene acceso a una base de datos centralizada. **Automatizar los procesos de operaciones:** La empresa los controla y obtiene información para integrar y acumular la información en un datawarehouse<sup>6</sup>; es decir, que el datawarehouse puede obtener información de otros sistemas, tales como los sistemas ERP que trabajan con otras bases de datos (ya que al almacén de datos también se le conoce como base de datos corporativa) para integrar información que ayude al análisis de los mismos para la toma de decisiones.

---

<sup>6</sup> O Almacén de datos es la colección de datos orientados a un tema, integrados, no volátiles, e historizados, organizados para dar soporte al proceso de ayuda en la toma de decisiones.

- **Permiten incorporar las nuevas Tecnologías de la información y la comunicación (TIC):** Como por ejemplo: transferencias electrónicas (de dinero, entre otros), sistemas electrónicos de intercambio de datos (correo electrónico, entre otros), ejecución de procesos a través de internet (como ventas).
- **Arquitectura Cliente / Servidor:** Los sistemas ERP se encuentran basados en esta arquitectura, la cual consta de un computador central que se llama servidor, cuya actividad es atender de manera simultánea a los usuarios a los que se les conoce como clientes.
- **Elevado número de funcionalidades:** Los sistemas empresariales cuentan con numerosas funcionalidades, esto permite que en las organizaciones se abarquen la mayoría de los procesos de negocio.
- **Grado de abstracción:** Puede manejar cualquier tipo de circunstancias que se puedan presentar en la organización, además de que suele tolerar distintos grupos empresariales sin conexión entre ellos.
- **Orientación a los procesos de negocio:** Centrándose en el diseño de los sistemas ERP, las funcionalidades se encuentran organizadas con la ayuda de un modelo de referencia o con la descripción de las funcionalidades con base en la razón del negocio haciendo uso de cualquier herramienta sobre modelación de procesos de negocio.<sup>7</sup>

---

<sup>7</sup> Herramientas que ayudan al modelado de procesos de negocios(BPM). El BPM trata de representar los procesos de negocio de una empresa con objeto de que puedan ser analizados y mejorados; es decir se trata de una técnica para alinear los desarrollos con las metas y objetivos de las empresas.

- **Universalidad:** Es un software de tipo *World Class*,<sup>8</sup> por lo que un ERP se puede usar en cualquier empresa. Sin embargo, los proveedores hacen énfasis en que hay ERP para industrias específicas.

Como se muestra cada autor tiene una manera de conceptualizar a los ERP, algunos son breves, otros mencionan sus componentes y otros explican las características que diferencian a los sistemas ERP de los demás sistemas. Pero de todas las definiciones se puede rescatar que un sistema ERP se trata de un software (en este caso sistema) integrado para la gestión de la empresa; se compone de módulos que se encuentran comunicados entre las diferentes áreas funcionales de una empresa. Las características que se pueden retomar, ya que al parecer son esenciales, es que los módulos son independientes, pero a la vez todos se integran, y los pueden ir agregando conforme la organización lo requiera; otra característica es que hacen uso de una base de datos central y los módulos se mantienen comunicados mediante la misma plataforma.

Los objetivos principales que buscan los sistemas ERP son:

- Brindar apoyo a los clientes de la organización en tiempo y manejo de información.
- Optimizar los procesos de la organización.
- Acceso a información confiable, precisa y pertinente (integridad de datos).
- Compartir información entre los módulos (departamentos) de la empresa.
- Eliminar datos y operaciones que se repiten.
- Reducir tiempos y costos de los procesos (mediante la reingeniería) lo que conlleva al aumento de la productividad de la organización.

---

<sup>8</sup> Es usado como sinónimo de excelencia, capacidad de cambio, mejora continua, productos y servicios de gran calidad.

## **1.2 Importancia de los sistemas ERP en las compañías**

Lazo (2010) dice que los “ERP son relevantes ya que proporcionan a los usuarios de este sistema una forma universal de entrar, ver y usar la información que se encuentra en estos sistemas, además de que las barreras entre los departamentos desaparecen.”

Los sistemas empresariales mejoran la comunicación de la información en la empresa, logrando con ello vincular y organizar sus operaciones cotidianas. Por ejemplo, en la transacción de un pedido, el almacén le informa a la fábrica que surta los productos que se han agotado, al mismo tiempo en el área de contabilidad se le hace una petición de que le mande la factura al cliente; por lo tanto, las personas encargadas del proceso del pedido se encuentran al tanto de lo que ocurre. Así mismo los sistemas empresariales le brindan a las organizaciones la facilidad de atender rápidamente las solicitudes de sus clientes, producir y almacenar la información necesaria en inventario.

También aumentan la rentabilidad de la empresa al minimizar costos, incrementan la satisfacción del cliente, mejoran la toma de decisiones administrativas (gracias al aporte de información valiosa) y ayudan a los gerentes a analizar la rentabilidad general de los productos al brindarles información de toda la empresa.

Los sistemas empresariales contienen herramientas analíticas que usan los datos que fueron capturados en el sistema, con ello evalúan el desempeño de la empresa.

### **1.3 Elementos de un ERP**

Como ya se mencionó; estos sistemas están compuestos por módulos (de acorde a cada área) con la finalidad de trabajar de manera independiente y a la vez trabajar en conjunto, algunos ejemplos son:

a) Planeación de los requerimientos de material (MRP).

Surgen ante la necesidad de integrar los artículos a fabricar en un inventario; así que son sistemas que integran las actividades de producción y compra, por lo tanto, tiene que ver con inventarios.

La lógica del MRP es simple, aunque su complejidad está en la cantidad de artículos a administrar y los niveles de explosión de materiales con que se cuenta. El MRP trabaja sobre la base de dos parámetros básicos del control de producción: tiempos y cantidades. El sistema debe de ser capaz de calcular las cantidades a fabricar de productos terminados, de los componentes necesarios y de las materias primas a comprar que permitan satisfacer la demanda independiente. Además, al hacer ésto debe considerar cuándo deben iniciar los procesos para cada artículo con el fin de entregar la cantidad completa en la fecha comprometida (Alviso, 2005, p. 40).

b) Sistemas de Gestión de la configuración del software (SCM).

Es un conjunto de operaciones y logística que busca la entrega de un producto a su cliente, contiene actividades como la obtención de materiales para la transformación del producto, hasta su colocación en el mercado.

c) Sistemas de administración de la relación con los clientes (CRM)

Tiene que ver con la parte estratégica que busca una empresa centrándose en la información de los clientes, con ello se busca ofrecer soluciones a éstos adecuándose a sus necesidades.

d) Sistemas de administración de recursos humanos (HRM)

Auxiliándose de estos sistemas se lleva a cabo la administración de los recursos humanos en actividades como pago de remuneraciones, contratación, ascensos, entre otros. Lo que ayuda a la organización a conocer información relevante acerca de sus empleados.

e) Sistemas de administración de recursos financieros (FRM)

Se encargan de integrar operaciones contables tales como ventas, activos, entre otros, para determinar los recursos que se necesitan en la organización para la realización de las proyecciones.

En la figura 2 podemos ver que un ERP está conformado por módulos básicos (mencionados anteriormente) que una empresa requiere para agilizar sus procesos mejorando su productividad.



Figura 2. Elementos y/o módulos básicos de un ERP

## 1.4 Funcionamiento de los sistemas ERP en las organizaciones

Como ya se mencionó los sistemas empresariales contienen un conjunto de módulos de software integrados y una base de datos central que permite compartir los datos entre muchos procesos de negocios y áreas funcionales diferentes de toda la empresa. Para entender mejor el funcionamiento de un ERP se explicará el mismo con apoyo de la figura 3.

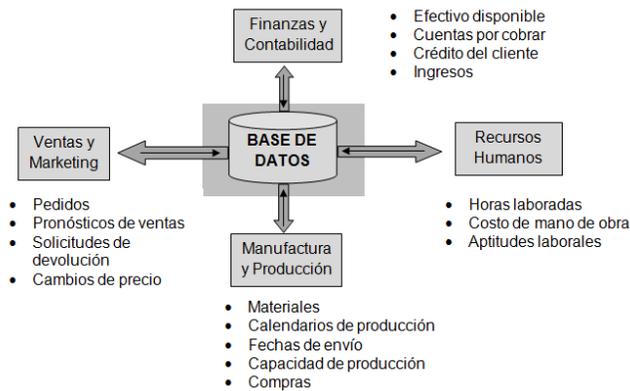


Figura 3.- Como funciona un sistema empresarial; Laudon y Laudon, (2008).

En la figura 3 se puede ver que existen distintos departamentos (módulos hablando del sistema ERP): el de finanzas y contabilidad, recursos humanos, manufactura y producción, ventas y marketing, cada departamento por si solo cuenta con procesos que puede llevar a cabo, y también puede realizar otras actividades con ayuda de otros departamentos, por ejemplo, centrándose en las siguientes áreas:

- Finanzas y contabilidad
- Manufactura y producción

El departamento de manufactura y producción puede saber con qué capacidad de producción cuenta de acuerdo a los ingresos que tiene la organización, los cuales se encuentran en el

departamento de finanzas y contabilidad, por lo que el ERP debe tener acceso a esta información para poder manipularla.

Los ERP se pueden ir complementando poco a poco, como ya se había indicado anteriormente, se puede empezar por un área funcional, después otra área y así consecutivamente, hasta que la empresa haya cubierto las necesidades y procesos de negocio que tiene con base en las áreas, es decir, si una organización con cinco departamentos ya cuenta con módulos en las áreas de finanzas, producción, administración y mercadotecnia, sólo requiere el de recursos humanos, así que decide implementar el módulo correspondiente, con lo que se estarán satisfaciendo las necesidades globales de la organización.

Otro claro ejemplo de cómo funciona un sistema ERP es basándose en departamentos como el de compras (adquisiciones que realiza la empresa); el de producción (encargado del desarrollo del producto); el de finanzas (cuestión económica que pueda tener una empresa basándose en balances generales, estados de resultados, entre otros); el de almacén (tiene que ver con la parte de inventarios); el de logística (información que tiene que ver con la distribución del producto); el de ventas (tiene que ver con información referente a los productos que se han vendido) y el de stock (que va de la mano con el área de almacenes pues también tiene que ver con inventarios) . Cada departamento cuenta con una base de datos, sin embargo, éstas se encuentran comunicadas entre sí, es decir, representa una base de datos a la cual todos los departamentos tienen acceso para visualizar, agregar, actualizar y eliminar datos, logrando que la información se vuelva confiable (véase la figura 4).

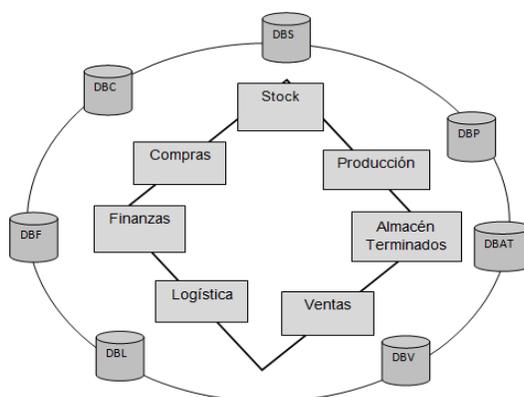


Figura 4. Los ERP integración horizontal de los procesos operativos de la empresa; Giner, (2004).

### 1.5 Gestión de procesos de negocio y reingeniería

La gestión de procesos y la reingeniería de procesos son conceptos diferentes aunque muchas veces se les asocia, ya que ambos conceptos ayudan a una empresa a mejorar su productividad, enfocándose en sus procesos de negocios, que como ya se mencionó, son primordiales para las organizaciones.

La gestión de procesos de negocios (BPM- *Business Process Management*), también conocida como administración de procesos de negocios, busca mejorar la eficiencia de las empresas, además de que les proporciona ventajas competitivas, ya que al estar al tanto de sus procesos de negocios la empresa se puede adaptar más fácilmente a los cambios que se presentan en el mercado.

El libro Club BPM (2010) considera al “*Business Process Management*, como Un conjunto de herramientas, tecnologías, técnicas, métodos y disciplinas de gestión para la identificación, modelización, análisis, ejecución, control y mejora de los procesos de negocio. Las mejoras incluyen tanto cambios de mejora continua como cambios radicales.

Otra definición es una “Disciplina de Gestión por Procesos de Negocio y de Mejora Continua apoyada fuertemente por las Tecnologías de la información”(Hitpass, 2012, p.19)

Por lo tanto, el BPM es una técnica o método que permite manejar los procesos (manuales y automatizados) de una empresa para lograr resultados con base a los objetivos alineados a la estrategia del negocio y se apoya de las tecnologías de información, aunque cabe aclarar que para algunos autores, como el libro Club BPM 2011, el BPM es sólo la disciplina del manejo de procesos de negocios (sin tecnologías de información), para otros, el libro de Club BPM 2010 y el libro de BPM por hitpass, es la gestión de los procesos de negocios con ayuda de las tecnologías de información.

En cuanto a la Reingeniería de procesos de negocios (BPR o RPN) implica un gran cambio en estos procesos para que se vea reflejado en el rendimiento de la organización.

Para Alarcón “Una **comprensión** fundamental y profunda de los **procesos** de cara al **valor añadido** que tienen para los **clientes**, para conseguir un rediseño **en profundidad** de los **procesos** e **implantar** un cambio **esencial** de los mismos para alcanzar mejoras **espectaculares** en medidas críticas del rendimiento (costes, calidad, servicio, productividad, rapidez,) modificando al mismo tiempo el **propósito del trabajo** y los **fundamentos** del negocio, de manera que permita establecer si es preciso unas nuevas **estrategias corporativas.**”

Por lo tanto, el BRP suele aplicarse una vez y lleva consigo mismo un cambio enorme en los procesos, ya que es el rediseño de los mismos y al personal le puede costar trabajo adecuarse a los nuevos procesos (disminución de productividad), es por eso que suele aplicarse a procesos de negocios estratégicos (1 o 2 por decir); en cambio, el BPM permite la mejora continua de uno o varios procesos de inicio a fin (modelado, administración, análisis, entre otros) con la finalidad

de conocer perfectamente cada proceso de negocio, lo que incrementará la eficiencia en la organización.

Los sistemas ERP no se enfocan directamente en los procesos de negocio, es decir, el ERP integra las áreas funcionales que cuentan con procesos de negocio de una organización, pero no maneja ni estudia (de inicio a fin) de manera directa los procesos de negocio de una empresa; en comparación con el BPR y el BPM (que se enfocan de lleno a los procesos de negocio), que si usan la gestión de procesos de negocios o la reingeniería, dependiendo de lo que la empresa necesite, y además se hace uso de un sistema ERP, por lo tanto la productividad de la misma va a aumentar no sólo porque los departamentos van a estar comunicados, sino también porque los procesos de negocios van a estar de acorde a lo que la organización requiere.

Como ya se comentó, en la implementación de los sistemas ERP a veces se tiene que hacer una reingeniería en los procesos para que estén acordes a la organización, es decir, si hay un proceso que no viene funcionando, se hace un rediseño del mismo y se adecua con el ERP, por lo que posiblemente se presente un cambio importante, en el mejor de los casos la organización puede mejorar su rendimiento y lograr la eficiencia en sus procesos y en el peor caso puede caer un poco su rendimiento, esto solo si no se hace una correcta implementación del sistema.

## **1.6 Arquitectura Cliente/Servidor**

La arquitectura Cliente/Servidor se utiliza en el área de redes y consta de dos elementos: un cliente y servidor, que son computadoras (ver figura 5), las cuales están en constante comunicación, ya que el cliente realiza las solicitudes necesarias para realizar una actividad y el servidor le proporciona una respuesta al cliente. Como ya se había mencionado, los sistemas ERP trabajan con esta arquitectura por lo que es importante conocer cómo funciona.



Figura 5. Modelo cliente/servidor

**El cliente** es el elemento encargado de hacerle una petición al servidor y espera a que éste le responda; se le conoce con el término *front-end* (llamados también procesos ligeros).

Sus funciones son:

- Gestionar la interfaz de usuario (manejar las pantallas).
- Interactuar, es decir, mantenerse en comunicación con el usuario (mantener y procesar todo el diálogo con el usuario).
- Procesar la lógica de aplicación y realizar validaciones a nivel local (entrada de datos y validación).
- Recoger resultados del *back-end* (servidor).
- Ocultar el servidor a la red.
- Recuperación de errores

**El servidor** es el elemento que se encarga de atender a los distintos clientes que le hacen peticiones sobre algún recurso que él administre. Se le conoce con el término *back-end* (llamados también procesos pesados) y sus funciones son:

- Aceptar los requerimientos de bases de datos que hacen los clientes.
- Procesar requerimientos de bases de datos.
- Procesar la lógica de la aplicación.
- Ejecutar validaciones sobre bases de datos, almacenamiento de datos y organización de los mismos.

## El modelo cliente/servidor para Tanenbaum (2003)

Es aplicable cuando el cliente y el servidor están en el mismo edificio (por ejemplo, cuando pertenecen a la misma compañía), pero también cuando están bastante retirados. Por ejemplo cuando una persona en casa accede a una página Web, se emplea el mismo modelo, en el que el servidor remoto de Web es el servidor y la computadora personal del usuario es el cliente. En la mayoría de los casos, un servidor puede manejar una gran cantidad de clientes.

Si vemos el modelo cliente-servidor en detalle, nos daremos cuenta de que hay dos procesos involucrados, uno en la máquina cliente y otro en la máquina servidor. La comunicación toma la siguiente forma: el proceso cliente envía una solicitud a través de la red al proceso servidor y espera una respuesta. Cuando el proceso servidor recibe la solicitud, realiza el trabajo que se le pide o busca los datos solicitados y devuelve una respuesta.

## Ventajas del esquema Cliente/Servidor

- Para el uso de sistemas Cliente/Servidor existen plataformas de hardware cada vez más baratas. Además, se puede utilizar hardware y software de varios fabricantes, lo cual contribuye considerablemente a la reducción de costos y favorece la flexibilidad en la implantación y actualización de soluciones.
- Facilita la integración entre sistemas diferentes y comparte información ( las máquinas ya existentes puedan ser utilizadas pero presentando interfaces más amigables al usuario)
- Al favorecer el uso de interfaces gráficas interactivas, los sistemas construidos bajo este esquema tienen mayor interacción con el usuario; presenta la ventaja, con respecto a uno

centralizado, de que no es siempre necesario transmitir información gráfica por la red pues ésta puede residir en el cliente.

- Es más rápido el mantenimiento y el desarrollo de aplicaciones.
- La estructura inherentemente modular facilita la integración de nuevas tecnologías y el crecimiento de la infraestructura computacional.
- Proporciona a los diferentes departamentos de una organización soluciones locales al mismo tiempo que permite la integración de la información de manera global.

#### Desventajas del esquema Cliente/Servidor

- Se cuenta con muy escasas herramientas para la administración y ajuste del desempeño sistemas.
- Es importante que los clientes y los servidores utilicen el mismo mecanismo (deben tener mecanismos generales que existan en diferentes plataformas).
- Hay que tener estrategias para el manejo de errores y para mantener la consistencia de los datos (se deben hacer verificaciones en el cliente y en el servidor).
- Pueden presentarse problemas en el desempeño por congestión en la red y dificultad de tráfico de datos.

### **1.7 Principales aplicaciones del ERP**

En el estudio del cuadrante mágico de Gartner (véase figura 8) publicado en 2012 (ERP para medianas empresas basadas en productos) se tienen como líderes a SAP con su solución *Business All-in-one* y a Microsoft *Dynamics* con su solución *Dynamics AX*, así que tomando como base estas soluciones es que se hablará de los sectores o industrias en los que se suelen aplicar los sistemas ERP. Ambas brindan soluciones para el sector minorista que se refiere a las unidades económicas que pueden comprar y vender un bien de consumo final (producto) o bien

dedicarse a la compra o venta. Así mismo, ofrecen soluciones para el sector público, que son organizaciones que realizan ciertas funciones para el estado.

Por otro lado, en la página Web el Senior Blog, quien hizo uso de información de evaluandoerp.com, habla de que 108 empresas que realizaron evaluaciones y comparaciones en el segundo trimestre del 2010 y se obtuvo la gráfica o figura 6 en la que se muestran las industrias o mercados que suelen hacer mayor uso de sistemas ERP, es decir, todas aquellas empresas que utilizan sistemas empresariales.

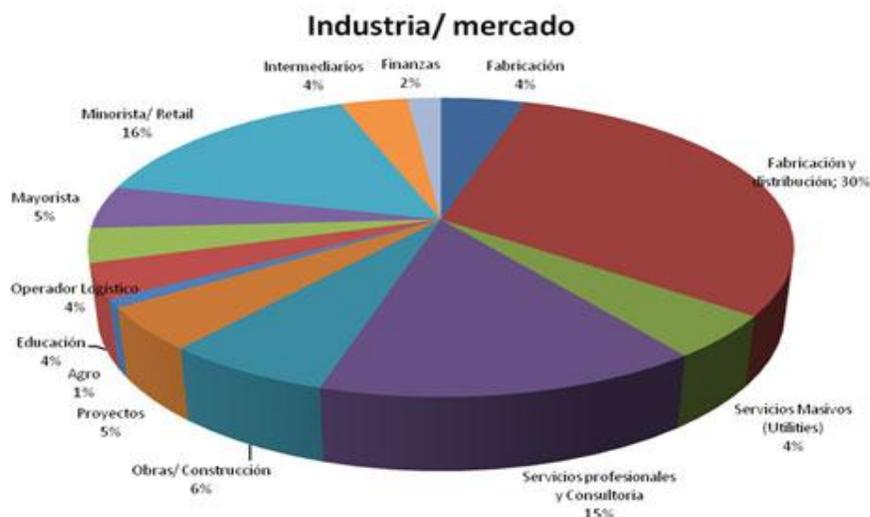


Figura 6. Gráfica de Senior Blog<sup>9</sup>. Recuperado

de<sup>10</sup> <http://blog.guiasenor.com/archives/2010/07/demanda-erp-2do-trimestre-2010.html>

Como se aprecia, de las 108 empresas que participaron en la evaluación de productos de ERP el sector que sobresale es el de fabricación y distribución. Por tal motivo se hablará brevemente de ellos.

---

9

Esta Imagen no tenía nombre de la página de donde se sacó, para efectos de esta tesina se le colocó el nombre de Gráfica de Senior Blog

La fabricación (manufactura): Es el proceso en el que determinados recursos (materias primas, mano de obra, energía, capital) se transforman para obtener bienes y servicios. Existen proveedores que ofrecen soluciones para este sector como: Microsoft Dynamics AX, Infor Global Solutions, entre otros.

La distribución: Se refiere a que un comerciante distribuye un producto al comerciante. Los proveedores que ofrecen soluciones para este sector son Microsoft Dinamycs AX, Infor Global Solutions, entre otros.

Así mismo, como se muestra en la figura 6, un ERP es implementado en diversos sectores, por lo que no es exclusivo de un sector, ya que se pueden adaptar a las organizaciones para que éstas puedan incrementar su eficiencia.

### **1.8 Proveedores de los ERP**

Existen diversos proveedores de sistemas ERP aunque los más mencionados en las distintas fuentes revisadas son SAP, Oracle, Microsoft, PeopleSoft, JD Edwards, SSA Global y Baan.

Romo (2008) comenta “A nivel mundial tenemos tres grandes fabricantes: las veteranas SAP y Oracle y la “novata” Microsoft. Destacar también otros fabricantes tales como Baan y Sage pero con una menor presencia en el mercado. El líder mundial de este tipo de software es SAP, uno de los inventores del concepto. A continuación tenemos en un segundo puesto a Oracle y un tercero a Microsoft.”

Rodríguez, Pineda y Sánchez (s.f) dicen “En el mercado existe una gran cantidad de sistemas erp comerciales, entre los más conocidos se encuentran: SAP, PeopleSoft, Oracle Financials, JD Edwards, Baan y Solomon.”

Cabe resaltar que Oracle compró Peoplesoft quien anteriormente había adquirido a la empresa JD Edwards (también encargada de desarrollo de ERP), con ello facilitando la entrada a las PYMES y abarcando un gran mercado.

Así mismo Infor Global Solutions (empresa de aplicaciones y servicios empresariales) adquirió SSA Global quien anteriormente compro Bann Co (también una empresa encargada del desarrollo de ERP ).

El mercado de los ERP ha sufrido cambios, así que los proveedores que permanecen en el mercado son SAP, Oracle, Microsoft e Infor Global Solutions, por lo que se hablará de manera breve de ellos.

- **SAP.** Fundada En Alemania en 1972 es considerada la mayor empresa proveedora de ERP mundialmente, destacándose en procesos de producción y gestión, logística y recursos humanos. Además, atiende a las empresas de todos los tamaños y fue pionera en los sistemas empresariales.

Su acrónimo es sistema de análisis y desarrollo de programas, aunque cabe mencionar que posteriormente se cambió a sistemas, aplicaciones y productos en procesamiento de datos (SAP AG).

- **Oracle.** Dedicado a los ERP desde 1987, sus clientes son organizaciones relacionadas con la producción y consumo de producto, así que es un competidor directo del SAP.

El nombre de *Oracle* se debe a que sus creadores trabajaron años antes (de 1987) para la agencia central de inteligencia (CIA) en un proyecto de base de datos en el que la clave era *Oracle*, lo que dió pie a que la corporación se llamara así.

- **Microsoft.** Inició operaciones en Latinoamérica y el Caribe en 1986. Unió sus productos de ERP bajo la división de soluciones de administración empresarial y cuenta con distintas versiones de *Dynamics*.
- **Infor Global Solutions.** Empresa privada fundada en Estados Unidos en 2002, inició bajo el nombre de *Agilisys* y posteriormente fue cambiado a *Infor Global Solutions*.

En la revista Software Guru se habla de que anteriormente se pensaba que los ERP sólo eran para empresas grandes, aunque era erróneo ya que las Pymes también cuentan con acceso a estos sistemas, pues en México existen cerca de 600,000 Pymes y se utilizan para disminuir costos y para que los procesos sean eficientes. El mercado de los ERP está compuesto por distintos proveedores como se muestra en la figura 7.

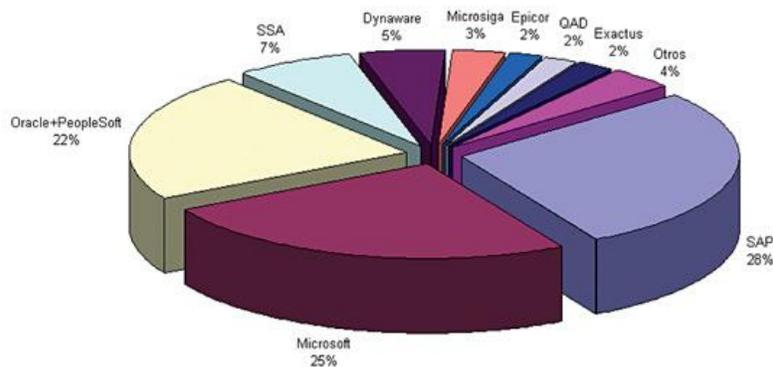


Figura 7. Gráfica de Software Guru.<sup>11</sup> Recuperado de <http://sg.com.mx/content/view/455>

Como se aprecia SAP predomina en el mercado siguiéndole Microsoft, Oracle y otros proveedores con menos reconocimiento.

En la página EvaluandoERP.com se habla de que según Gartner está de acuerdo con Himmelberger (Director de marketing de SAP Business) en que SAP contaba con el 26% de

<sup>11</sup> Esta imagen no tenía nombre de la página de donde se sacó, para efectos de esta tesis se le colocó el nombre de Gráfica de Software Guru

mercado en ERP en 2009 y Oracle con el 12% aunque Microsoft e *Infor* continúan en el mercado (véase Figura 8).

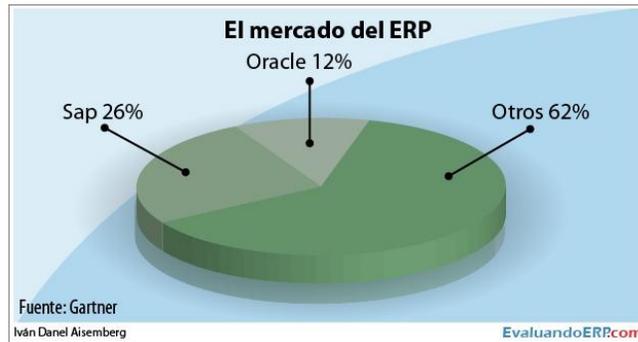


Figura 8. El mercado del ERP, EvaluandoERP.com; Iván (2012).

En un estudio que realizó Gartner (cuadrante mágico de Gartner) en junio del 2012 respecto a los ERP como instancia para medianas empresas centradas en producto, cataloga como a líderes a SAP (Business All-in-One) y a Microsoft Dynamics AX.

Figure 1. Magic Quadrant for Single-Instance ERP for Product-Centric Midmarket Companies

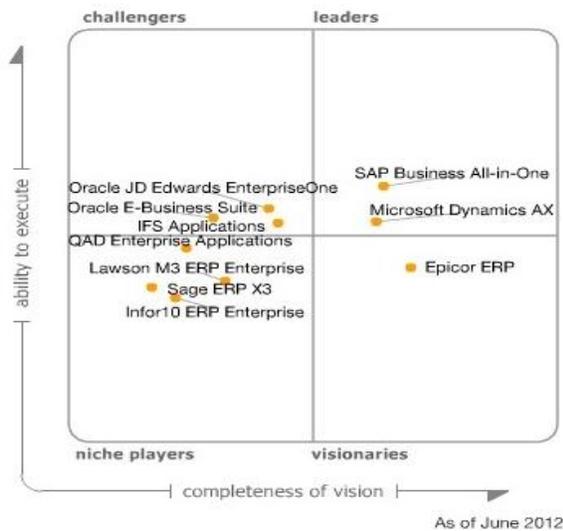


Figura 9. Magic Quadrant for Single-Instance ERP for Product-Centric Midmarket Companies; Gartner, (2012).

En la figura 9 se encuentran los líderes (*leaders*) que son aquellos que se encuentran a la cabeza de los ERP, es decir, cuentan con gran potencial y están preparados para hacer frente a los cambios. Por otro lado, están los aspirantes (*challengers*) que cuentan con sistemas ERP bastantes buenos, pero que no están acondicionados para soportar grandes cambios. Del mismo modo, se encuentran los visionarios (*visionaries*) que se pueden anticipar a las necesidades futuras (tienen visión), pero cuenta con incapacidad para ejecutar y tal vez al mercado le falta impulso. También se encuentran los jugadores de nicho (*niche players*), en los que los ERP están dirigidos a un mercado en especial, pero suelen carecer de experiencia, aunque pueden ser ideales para el mercado al que van dirigidos.

De los proveedores antes mencionados se ha llegado a la conclusión de que los proveedores más reconocidos son: El SAP, siguiéndole Microsoft y en tercer instancia Oracle; aunque a continuación se hablará específicamente de Business All-in-one (SAP) y Microsoft Dinamycs AX, dado que se encuentran como líderes en el cuadro mágico de Gartner y es la fuente consultada más actual.

- Business All-in-one, es un ERP que provee SAP que incluye planificación de los recursos empresariales, gestión de las relaciones con los clientes, *supplier relationship management*, *business intelligence*, *sap best practices*, plataforma tecnológica sap netweaver, con esto asegura que le ayuda a la empresa que adquiera Business- All-in-one a mejorar la gestión financiera, mantener la excelencia operativa, mejorar la agilidad y la productividad, unificar y simplificar; además, va dirigido para sectores como: aeronáutica, automotor, sector bancario, productos químicos, productos de consumo, defensa y seguridad, entre otros.

- Microsoft dynamics AX, se trata de un ERP que está dirigido a cinco sectores en específico como:
  - Fabricación: Permite ejecutar y mejorar sus procedimientos (fabricación contra pedidos, fabricación contra stock, informes, entre otros).
  - Distribución: Conecta empleados, operaciones y distribuidores en una solución, simplifica procesos en la cadena de suministro y se puede utilizar la inteligencia de negocio de autoservicio
  - Sector minorista: Para empresas medianas y grandes con capacidades para puntos de venta, gestión de tiendas, cadena de suministro, comercialización, entre otros.
  - Servicios: Permite la atención a clientes de todo el mundo, mensajería integrada, VoIP<sup>12</sup>, Plantillas de proyecto.
  - Sector público: Gestiona fianzas, compras, recursos humanos, relaciones con los clientes, fondos, proyectos y subvenciones y prestación de servicios.

Como se puede ver, existen diversos proveedores de ERP, cada uno de ellos con la finalidad de lograr satisfacer las necesidades de las organizaciones y cada uno de ellos ofrece distintas soluciones, sólo es cuestión de que la empresa que desee implementar un ERP estudie bien cada una de las alternativas para ver cuál es la más adecuada.

## **1.9 Ventajas**

La ventaja más importante de los ERP es la administración de la información en tiempo real, lo que ayuda a que la empresa tenga una gran interacción con la logística de información y

---

<sup>12</sup> Conocida también como la voz sobre IP o voz IP es una tecnología que permite la transmisión de voz a través de las redes IP (Internet).

productos, es decir, en las áreas cambia constantemente la información, así que se tiene una buena comunicación.

Otro beneficio es que la información no se duplica y hay menos errores en la información (cuando es manipulada), por ser un sistema integral. Además, su implementación adecuada involucra el aumento de la productividad en la organización, así como el mejor aprovechamiento del tiempo (ahorros de tiempo), es decir, si antes se ocupaba tiempo para realizar actividades como llevar un informe de un departamento a otro, ahora ese tiempo es empleado en otras actividades; de igual manera, se facilitan los procesos que se realizan de forma manual, por lo que también se percibe el crecimiento en la competitividad de la empresa.

Del mismo modo representa otros beneficios tales como:

- Constituye las áreas de una organización de manera que ésta tiene más control sobre su operación, formando la cooperación y coordinación entre los departamentos.
- Forma una solución de manera integrada para algunas de las actividades de la organización, lo que conlleva a la actualización continua e inmediata de los datos en los diferentes lugares donde se encuentre la organización, logrando con ello mejorar el proceso de la toma de decisiones.
- Se realiza una base de datos centralizada en la que se registran, procesan, revisan y controlan las funciones que son ejecutadas en la organización independientemente del lugar en donde se encuentren, de manera que al momento de acceder a información de una base de datos central e integrada favorece el proceso de toma de decisiones.
- En una organización un ERP refuerza todo el software en un solo sistema, a diferencia de otros sistemas que tienen distintos tipos de software integrados en ella.

Como se puede apreciar, existen diversas ventajas que las organizaciones pueden adquirir haciendo el uso debido de los sistemas ERP incrementando su productividad y optimizando sus procesos.

## **Desventajas**

A continuación se explican los factores por los cuales los ERP no serían viables para las empresas.

- **Inversión:** Es uno de los problemas primordiales que puede enfrentar una organización, pues además del costo del sistema se tienen otros costos como la capacitación, implementación, soporte y configuración, entre otros.
- **Tiempo y complejidad de implementación:** Es complicado implementar un ERP, ya que puede pasar mucho tiempo para la implementación de éste, lo que puede afectar la eficiencia de las actividades que se llevan a cabo en la empresa.
- **Personal:** Automatiza tareas que son ejecutadas por personas, si éstas no se encuentran capacitadas y no cuentan con habilidades para usar los sistemas ERP, la empresa se verá afectada.
- **Datos como saldos en inventarios, cuentas por pagar, y costos de productos son parámetros que el sistema calcula cada vez que éstos se requieren, y sólo se guarda el dato de cierre del período. Esto dificulta modificar los indicadores ante cambios en el entorno del negocio.**
- **Complejidad para integrar la información externa:** Es difícil integrar información externa en los ERP, dado que las organizaciones cuentan con sistemas independientes de distintos proveedores que cuentan con estructuras no compatibles con el ERP.

Valda (2011) dice:

Aunque los sistemas **ERP** puedan generar un incremento de productividad, para muchas empresas es casi imposible pagar el costo de las licencias, implementación y sobre todo del mantenimiento del mismo, pues son dinámicos, de nada sirve tener el mismo sistema en una empresa que crece y cambia día a día.

Además del costo, el tiempo que sugiere la implementación es un problema para las empresas, este problema empieza por la rigidez que tienen los ERP, es difícil que una empresa en particular desarrolle su propio sistema, los ERP que son sistemas genéricos, tienen que ser adaptados a las empresas desde su estructura principal.

El manejo del **ERP** tiene sus desventajas, se necesita instruir a los trabajadores de cada módulo que se vaya a asignar, la especialización de los trabajadores genera un costo y tiempo que tiene que emplear la persona para hacer un cambio en su estructura operativa, lamentablemente la resistencia al cambio presenta un problema muy grande en este punto.

## Capítulo II Bases de datos relacionales

### 2.1 Conceptos básicos

Para comprender mejor las bases de datos relacionales se empezará por definir una base de datos.

Para Coronel, Morris y Rob (2011) es “una estructura computarizada compartida e integrada que guarda un conjunto de:

- Datos del usuario final, es decir, datos sin elaborar que son de interés para el usuario final.
- **Metadatos**, o datos acerca de datos, por medio de los cuales los datos del usuario final son integrados y manejados.”

Es un “*Fondo común de información almacenada en una computadora para que cualquier persona o programa autorizado pueda acceder a ella, independientemente de su procedencia y del uso que haga.*” (Pons, Marín, Medina, Acid y Vila, 2005, p.6)

Por lo tanto una base de datos es un conjunto de datos relacionados que se encuentran recopilados (guardados) en un medio de almacenamiento para la organización de los mismos; los elementos que interactúan con una base de datos son: hardware, software, datos y usuarios.

En el ámbito de las bases de datos existen términos de suma importancia en cuanto a la abstracción de datos (Ver figura 10).

- Nivel interno o físico: La manera en que los datos están almacenados físicamente, éste es el nivel más bajo.
- Nivel conceptual o lógico: Tiene que ver con el administrador de base de datos.
- Nivel de vistas: La forma en que los usuarios ven los datos.

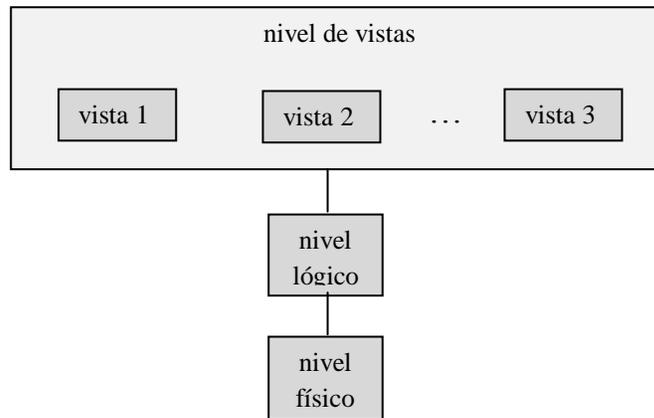


Figura 10.- Los tres niveles de abstracción de datos; Silberschatz, Korth y Sudarshan,(2007).

Los sistemas de bases de datos cuentan con esquemas según el nivel de abstracción que se maneje, es decir, el nivel de vistas tiene un esquema de vistas, el nivel lógico posee un esquema lógico y el nivel físico contiene un esquema físico; estos esquemas son la forma en que la base de datos representa el diseño conceptual, lógico y físico.

Así mismo hay conceptos básicos en cuanto a los errores o deficiencias que suele haber en una base de datos:

- Redundancia: Solo se necesita de una copia correcta de los nombres y números de teléfono del agente. Hacer que aparezcan en más de un lugar produce redundancia de datos. Existe redundancia de datos cuando los mismos datos se guardan innecesariamente en diferentes lugares (coronel, Morris y Rob). Cabe aclarar que al hablar de números y teléfono del agente los autores hacen referencia al número y teléfono de una persona que funge como agente.

La redundancia se refiere a la duplicidad de los mismos datos en diferentes partes de la base de datos.

- Inconsistencia: Cuando se almacenan datos de versiones diferentes que se refieren al mismo dato en distintas relaciones (tablas), es decir, los datos no coinciden en distintos lugares (relaciones) aunque hagan referencia a lo mismo.
- Falta de integridad: Que los datos sean consistentes, es decir, precisos y verificables.

Los términos que se deben conocer en bases de datos son:

- Sistema de administración de base de datos (DBMS o SGBD): Es el software que opera la base de datos y permite la interacción con el usuario.
- Lenguaje de definición de datos (LDD): Lo proporciona el DBMS y permite definir los datos.
- Lenguaje de manipulación de datos (LDM): Esta función debe ser proporcionada por el DBMS y permite hacer peticiones para agregar, actualizar, eliminar y borrar datos de una base de datos.
- Diccionario de datos: Tiene información de la base de datos, es decir datos de datos(metadatos). Supóngase que en una base de datos se tiene la tabla (relación, con base de datos relacional o entidad en el modelo entidad-relación) con el nombre empleado, esta tabla se encuentra compuesta por los atributos *id\_empleado*, *nom\_empleado*, *apellido\_p*, *apellido\_m*, *colonia*, *cod\_postal*, *dirección*, *id\_depto* y *nom\_depto*;  
 el *id\_empleado*, (nombre del atributo) tiene un contenido o descripción, es de tipo int (tipo), tiene un formato y además es llave primaria. Se puede hacer lo mismo con los demás atributos; al conjunto de toda la descripción de los datos se le llama diccionario de datos.

En el modelado de datos especificando el modelo de datos relacional facilita el conocimiento de los metadatos, pues se conocen las relaciones y los atributos, lo que ayuda a definir correctamente el diccionario de datos.

- Modelo de datos: Herramienta que permite describir los datos y relaciones, por lo tanto se describe el diseño de una base de datos en los 3 niveles de abstracción. Existen diversos modelos de datos como:
  - Modelo relacional: Está basado en registros, pues usa tablas para representar los datos y las relaciones, cada tabla tiene registros definidos.
  - Modelo entidad-relación: Toma la percepción del mundo real trabajando con entidades (cosa u objeto) y relaciones.

También se encuentran otros modelos de datos como: modelo de datos orientado a objetos, semiestructurados, de red y jerárquico; lo cuales no son expuestos dado que no son esenciales para la realización de la tesina.

- Consulta: Es una petición que se hace a la base de datos esperando una respuesta

Una vez explicado los conceptos fundamentales se proporcionará una definición de una base de datos relacional:

“Un esquema de base de datos relacional  $S$  es un conjunto de esquemas de relación  $S=\{R_1, R_2, \dots, R_m\}$  y de restricciones de integridad<sup>13</sup> RI.”(Elmasri y Navathe, 2007, p.132).

Silberchatz, Korth y Sudarshan (2007) dicen que las bases de datos relacionales “se basan en el modelo relacional y usan un conjunto de tablas para representar tanto los datos como las

---

<sup>13</sup> Coherencia entre los datos.

relaciones entre ellos. También incluyen un LMD y un LDD. La mayor parte de los sistemas de bases de datos relacionales comerciales emplean el lenguaje SQL,”

Para comprender mejor las bases de datos relacionales es esencial conocer algunos conceptos importantes como:

- Relación: Se le considera así a una tabla que consta de filas (tuplas) y columnas (atributos).
- Atributo: Se refiere a la característica de una relación.
- Tupla: se le llama así a una fila.
- Dominio: Conjunto válido de valores asociados a un atributo.

Retomando el ejemplo que se utilizó en el diccionario de datos se utiliza la figura 11 para ilustrar los conceptos de bases de datos relacionales.

<u>id empleado</u>	<u>nom empleado</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>	<u>direccion</u>	<u>id depto</u>	<u>nom depto</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240	huexotitla no.12	1	finanzas
2	rocio	sosa	ortiz	tepetlaoxtoc	56245	arandas no.2	2	mercadotecnia
3	samuel	lopez	vidal	san miguel	56293	benito juarez	3	administracion
4	mireya	estrada	gonzalez	san joaquin	56242	nopaltitla no.20	4	recursos humanos
5	melissa	campos	sosa	san miguel	56239	el cerrito no.3, limon s/n	5	produccion

Figura.- 11; Empleado.

Relación: Empleado

Atributo: Son los que se encuentran subrayados y con un fondo claro, es decir, *id\_empleado*, *nom\_empleado*, *apellido\_p*, *apellido\_m*, *colonia*, *cod\_postal*, *dirección*, *id\_depto* y *nom\_depto*.

Tupla: La fila que se encuentra con un tono más fuerte.

Dominio: Los valores que están entre llaves son el dominio de *id\_empleado*.

Para hacer el diseño lógico de una base de datos relacional se utiliza el modelo relacional, en ocasiones a partir del modelo entidad-relación (ER) se puede diseñar el modelo relacional; sin embargo, es necesario convertir el diagrama ER a modelo relacional mediante una serie de pasos. Una vez que se tiene el modelo relacional es necesario utilizar la técnica llamada normalización para eliminar redundancia, mantener la integridad y proporcionar información más eficiente.

Como anteriormente se habla de un diseño lógico, cabe aclarar que éste pertenece a una de las fases que se realizan para hacer el diseño de una base de datos, así que si se requiere hacer un diseño de base de datos se llevan a cabo ciertas fases.

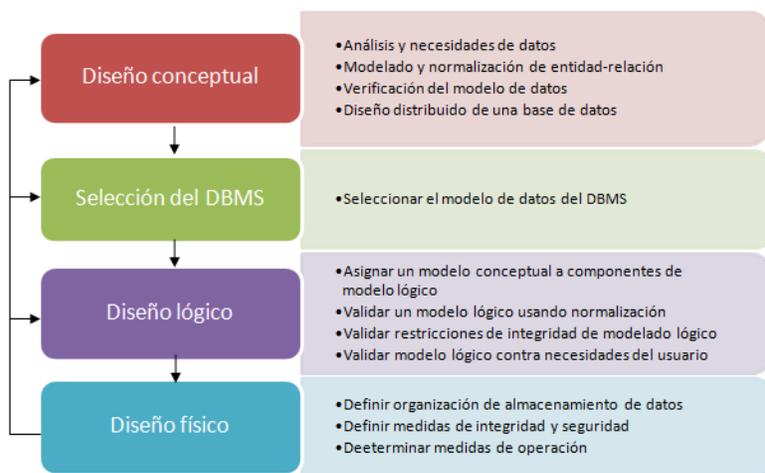


Figura 12. Pasos para el diseño de una base de datos.

En la figura 12 se pueden apreciar 4 pasos (fases) para realizar el diseño de la base de datos que son:

- **Diseño conceptual:** se hace el diseño de la base de datos independientemente del software y detalles físicos, para ello se usa el modelo (diagrama) entidad relación; para poder hacer este modelo se deben conocer 4 puntos importantes.

- Análisis y necesidades de datos: Se debe conocer las necesidades de los usuarios y la estructura u organización de la información.
- Modelado y normalización de entidad relación: Es importante estandarizar documentación y se crea el modelo entidad relación a partir de reglas de negocio.
- Verificación del modelo de datos: Es corroborar que el modelo de datos puede soportar los procesos que se requieren hacer.
- Diseño distribuido de una base de datos: En caso de que se vaya a hacer uso de una base de datos distribuida, es en este punto donde se debe estudiar la división de la base de datos y el lugar de almacenamiento de cada fragmento.
- Selección del DBMS
  - Se deben estudiar las características del DBMS como sus ventajas, desventajas, costos, funciones que ofrece y hardware que necesita.
- Diseño lógico: Se hace una estructura de la base de datos independientemente de los detalles físicos, por lo que se hace uso del modelo relacional, para realizar esta fase hace uso de 4 puntos que son:
  - Asignar el modelo conceptual a componentes de modelo lógico: Se convierte el modelo entidad relación (realizado en el diseño físico) al modelo relacional
  - Validar el modelo lógico mediante normalización: El modelo relacional debe estar normalizado por lo menos en la tercer forma normal, aunque cabe aclarar que la normalización se puede presentar en las diferentes etapas del diseño.
  - Validar restricciones de integridad de modelado lógico: Las restricciones se definen y son aceptadas.

- Validar modelo lógico contra necesidades del usuario: Se acepta lo que se definió en el modelo lógico.
- Diseño físico: Se establece el almacenamiento de datos, así como las características del acceso a los datos, por lo tanto se ejecutan 3 actividades que son:
  - Definir organización de almacenamiento de datos: Se debe conocer el volumen de datos y los patrones de uso para determinar la ubicación y organización física de cada relación (tabla)
  - Definir medidas de integridad y seguridad: A través del DBMS el administrador puede designar a los usuarios que tienen acceso a determinadas relaciones de la base de datos.
  - Determinar medidas de operación: Pues la operación de una base de datos se puede ver afectada por características físicas de almacenamiento (tiempo de búsqueda, tamaño del sector, número de discos, número de cabezas de lectura/escritura)

## **2.2 Normalización**

En la elaboración de bases de datos es primordial que se haga la normalización para eliminar redundancias e inconsistencia de datos.

La normalización “es un proceso para evaluar y corregir estructuras de tablas a fin de minimizar redundancias de datos, con lo cual se reduce la probabilidad de anomalías de datos.” (Coronel, Morris y Rob, 2011, p.175)

“Es una serie de reglas que involucra análisis y transformación de las estructuras de datos en relaciones que exhiben propiedades únicas de consistencia, mínima redundancia y máxima estabilidad.” (Osorio, 2008, p.36).

En la normalización se suele hacer uso de tres formas normales (3FN) Por lo que se estudiará cada una de ellas:

1) La primera forma normal para los autores Coronel, Morris y Rob (2011) “describe el formato tabular en el que:

- Todos los atributos llave están definidos.
- No hay grupos repetidores en la tabla. En otras palabras, cada intersección de renglón/ columna contiene un y sólo un valor, no un conjunto de ellos.
- Todos los atributos son dependientes de la llave primaria.”

El siguiente autor dice es “Una relación R se encuentra en primera forma normal (1FN) si y sólo si todos los dominios simples subyacentes contienen solo valores atómicos. Es decir que el cruce de una fila con una columna sólo tiene un dato, o sea, que no existen grupos repetitivos.”(Nevado, s.f, p.94)

Por lo tanto, cada atributo tiene un valor y cada tupla tiene una llave que la identifique, para ilustrar la 1FN (primer forma normal), se retomará la figura 11 nuevamente ahora figura 13 (que se encuentra sin normalizar). Como se aprecia en esta figura se tiene un valor asociado en la quinta tupla a un atributo (*direccion*), que no es atómico.

<u>id empleado</u>	<u>nom empleado</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>	<u>direccion</u>	<u>id depto</u>	<u>nom depto</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240	huexotitla no.12	1	finanzas
2	rocio	sosa	ortiz	tepetlaoxtoc	56245	arandas no.2	2	mercadotecnia
3	samuel	lopez	vidal	san miguel	56293	benito juarez	3	administracion
4	mireya	estrada	gonzalez	san joaquin	56242	nopaltitla no.20	4	recursos humanos
5	melissa	campos	sosa	san miguel	56239	el cerrito no.3, limon s/n	5	produccion

Figura 13; Empleado

Dado que se tienen dos valores: la del cerrito no.3 y limón s/n, el atributo (*ver figura 14*) *dirección* se quita y se pone con la clave primaria aparte.

<u>id empleado</u>	<u>nom empleado</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>	<del><u>direccion</u></del>	<u>id depto</u>	<u>nom depto</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240	<del>huexotitla no.12</del>	1	finanzas
2	rocio	sosa	ortiz	tepetlaoxtoc	56245	<del>arandas no.2</del>	2	mercadotecnia
3	samuel	lopez	vidal	san miguel	56293	<del>benito juarez</del>	3	administracion
4	mireya	estrada	gonzalez	san joaquin	56242	<del>nopaltitla no.20</del>	4	recursos humanos
5	melissa	campos	sosa	san miguel	56239	<del>el cerrito no.3, limon s/n</del>	5	produccion

Figura 14; Empleado.

Como el atributo *dirección* se pone en otra relación junto con el *id\_empleado* (que es la llave primaria) la primera forma normal quedaría como se ilustra en las tablas de las figuras 15 y 16.

<u>id empleado</u>	<u>nom empleado</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>	<u>id depto</u>	<u>nom depto</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240	1	finanzas
2	rocio	sosa	ortiz	tepetlaoxtoc	56245	2	mercadotecnia
3	samuel	lopez	vidal	san miguel	56293	3	administracion
4	mireya	estrada	gonzalez	san joaquin	56242	4	recursos humanos
5	melissa	campos	sosa	san miguel	56239	5	produccion

Figura 15; Empleado.

<u>id empleado</u>	<u>direccion</u>
1	huexotitla no.12
2	arandas no.2
3	benito juarez
4	nopalitla no.20
5	el cerrito
5	no.3, limon s/n

Figura 16; Domicilio

2) Los autores Coronel, Morris y Rob (2011) definen que una relación está en segunda forma normal cuando:

- “Está en **1F**.

*y también*

- No incluye dependencias parciales; esto es, ningún atributo es dependiente de sólo una parte de la llave primaria

Nótese que todavía es posible que una tabla en 2FN exhiba dependencia transitiva; esto es, la llave primaria puede apoyarse en uno o más atributos no primos para determinar funcionalmente otros atributos no primos, como está indicado por una dependencia funcional entre los atributos no primos.”

“Una relación R se encuentra en segunda forma normal si y sólo si está en primera forma normal y todos los atributos no clave, dependen funcionalmente de forma completa de la llave primaria.”(Nevado, s.f, p.95)

Nótese que se usan las figuras 15 y 16 (que ya están en 1FN anteriormente explicado) para llevar a cabo la 2FN.

Se observa que existe dependencia parcial (los atributos no son totalmente dependientes de la llave primaria) en la relación Empleado (Figura 15, mas no en la relación Domicilio, por lo que esta última se conserva igual), pues el atributo *nom\_depto* no depende de la llave primaria que es *id\_empleado*, sin embargo, depende del atributo *id\_depto*, por ello se quitan de la relación Empleado y se ponen en otra relación

<u>id empleado</u>	<u>nom emplea do</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>	<u>id depto</u>	<u>nom depto</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240	1	finanzas
2	rocio	sosa	ortiz	tepetlaoxtoc	56245	2	mercadotecnia
3	samuel	lopez	vidal	san miguel	56293	3	administracion
4	mireya	estrada	gonzalez	san joaquin	56242	4	recursos humanos
5	melissa	campos	sosa	san miguel	56239	5	produccion

Figura 17; Empleado

<u>id empleado</u>	<u>direccion</u>
1	huexotitla no.12
2	arandas no.2
3	benito juarez
4	nopaltitla no.20
5	el cerrito
5	no.3, limon s/n

Figura 18; Domicilio

Como se separan los atributos *id\_depto* y *nom\_depto* en otra relación la 2FN queda de la siguiente manera:

id empleado	nom emplea do	apellido p	apellido m	colonia	cod postal
1	carlos	rodriguez	rodriguez	la purificacion	56240
2	rocio	sosa	ortiz	tepetlaoxtoc	56245
3	samuel	lopez	vidal	san miguel	56293
4	mireya	estrada	gonzalez	san joaquin	56242
5	melissa	campos	sosa	san miguel	56239

Figura 19; Empleado

id empleado	direccion
1	huexotitla no.12
2	arandas no.2
3	benito juarez
4	nopaltitla no.20
5	el cerrito
5	no.3, limon s/n

Figura 20; Domicilio

id depto	nom depto
1	finanzas
2	mercadotecnia
3	administracion
4	recursos humanos
5	produccion

Figura 21; Departamento

id empleado	id depto
1	1
2	2
3	3
4	4
5	5

Figura 22; Rel\_EmpDep

Una vez explicada la 2FN se platicará de la 3FN.

3) “Una tabla está en **tercera forma normal (3NF)** cuando:

- Está en 2NF

y además

- No contiene dependencias transitivas.” (Coronel, Morris y Rob, 2011, p.187)

“Se dice que una relación está en forma normal si y sólo si está en segunda forma normal y todos los atributos no clave dependen de manera no transitiva de la clave primaria.” (Nevado, s.f, p.97)

Por lo que se está en 3FN cuando se tiene la 2FN y no hay dependencias transitivas, cabe aclarar que una dependencia transitiva es cuando los atributos que no son llave primaria no dependen de otros atributos que no son llave primaria, es decir, si los atributos dependen unívocamente de la llave primaria.

A continuación se ilustra la 3FN a partir de las figuras de las tablas 19, 20, 21 y 22 (ya que están en 2FN).

Dado que la única relación que contiene dependencias transitivas (atributo no primo dependiente de un atributo no primo) es la de EMPLEADO, pues el atributo *colonia* depende de *cod\_postal*, *colonia* se quita junto con el atributo del cual depende (Ver Figura 23).

<u>id empleado</u>	<u>nom empleado</u>	<u>apellido p</u>	<u>apellido m</u>	<u>colonia</u>	<u>cod postal</u>
1	carlos	rodriguez	rodriguez	la purificacion	56240
2	rocio	sosa	ortiz	tepetlaoxtoc	56245
3	samuel	lopez	vidal	san miguel	56293
4	mireya	estrada	gonzalez	san joaquin	56242
5	melissa	campos	sosa	san miguel	56239

Figura 23; Empleado

Ya que se quitó el atributo *colonia* (junto con *cod\_postal*) salen 2 relaciones, las cuales se muestran en las tablas de las figuras 24 y 25.

id empleado	nom empleado	apellido p	apellido m	cod postal
1	carlos	rodriguez	rodriguez	56240
2	rocio	sosa	ortiz	56245
3	samuel	lopez	vidal	56293
4	mireya	estrada	gonzalez	56242
5	melissa	campos	sosa	56239

Figura 24; Empleado.

cod postal	colonia
56240	la purificacion
56245	tepetlaoxtoc
56293	san miguel
56242	san joaquin
56239	san miguel

Figura 25; Localidad.

Por lo tanto, las demás relaciones como no tienen dependencias transitivas quedan como estaban, ello se muestra en las tablas de las figuras 20, 21 y 22.

id empleado	direccion
1	huexotitla no.12
2	arandas no.2
3	benito juarez
4	nopalitla no.20
5	el cerrito
5	no.3, limon s/n

id depto	nom depto
1	finanzas
2	mercadotecnia
3	administracion
4	recursos humanos
5	produccion

id empleado	id depto
1	1
2	2
3	3
4	4
5	5

Figura 20; Domicilio

Figura 21; Departamento

Figura 22; Rel\_EmpDep

Una vez que se tiene la normalización se hace el modelo relacional, el cual quedaría de la siguiente forma:

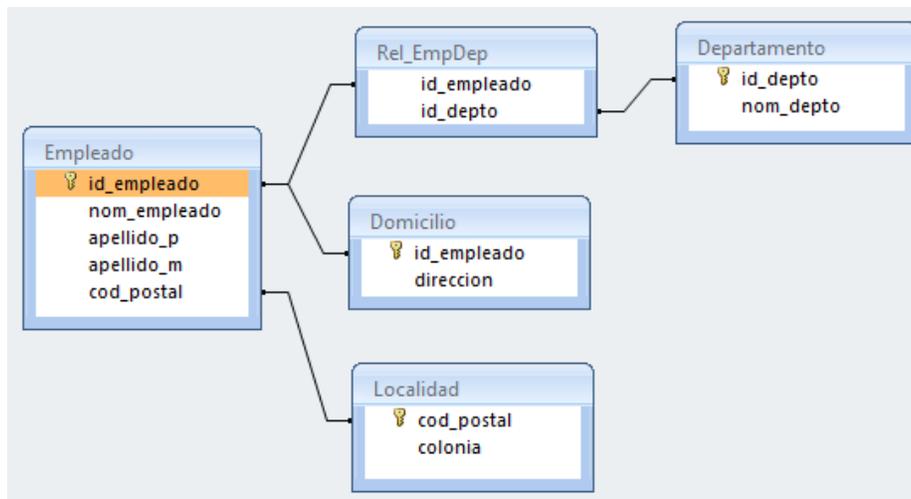


Figura 26; modelo relacional.

Ya que se tiene el modelo relacional (modelo lógico) que se muestra en la figura 26 se hace el diccionario de datos para ver el diseño físico, con ello todas las personas que se están encargando del diseño de la base de datos usaran la misma relación (tabla) con los mismos nombres y características de los atributos, con lo cual se reducirán los errores.

## 2.3 SQL

SQL significa lenguaje de consulta estructurado y fue realizado originalmente por la *International Business Machines* (IBM), en español: máquinas de negocios internacionales, se trata de un lenguaje de consulta estructurado utilizado en las bases de datos relacionales para la manipulación y la administración de los datos.

Elmasri y Navathe (2007) dicen “SQL es un lenguaje de bases de datos global: cuenta con sentencias para definir datos, consultas y actualizaciones. Por tanto, se comporta como DDL y como DML.”

Como se comentó anteriormente el DDL permite la definición de datos. A continuación se explican algunos comandos que permiten la definición.

- Crear objetos sobre bases de datos: Como tablas, índices y vistas; la sintaxis descrita por Coronel, Morris y Rob es: CREATE TABLE, CREATE INDEX Y CREATE VIEW, a continuación se muestran algunos ejemplos con base a la relación EMPLEADO que se ha estado manejando en la investigación.

- CREATE TABLE: Permite crear una tabla.

Ejemplo:

```
CREATE TABLE Empleado(  
    id_empleado integer,  
    nom_empleado char(15),  
    apellido_p char(15),  
    apellido_m char(15),  
    cod_postal varchar(25));
```

- CREATE INDEX: Permite crear un índice para una tabla

Ejemplo:

```
CREATE INDEX nombre_empind on Empleado (nom_empleado);
```

- CREATE VIEW: Permite crear una vista.

Así mismo anteriormente se comentó que el DML permite la manipulación de datos (se muestran ejemplos con base a la relación Empleado), esto lo hace a través de algunas sentencias que permiten:

- Insertar datos: Se refiere a insertar una tupla en una relación y su sintaxis básica según

Coronel, Morris y Rob es:

```
INSERT INTO nombre_de_tabla  
VALUES (valor_1, valor_2, ... , valor_n);
```

Ejemplo:

```
INSERT INTO Empleado  
VALUES (1, 'carlos', rodriguez ', 'rodriguez', 56240);
```

- Actualizar datos: Con este comando se modifican los datos en una relación, y según

Coronel, Morris y Rob se escribe:

```
UPDATE nombre_de_tabla  
SET nombre_de_columna=expresión [, nombre_de_columna = expresión]  
[WHERE lista_de_condiciones];
```

Ejemplo:

```
UPDATE Empleado  
SET nom_cliente= "ana", apellido_p= "gonzalez"  
WHERE nom_cliente= "carlos", apellido_p= "rodríguez";
```

- Guardar datos: Los datos se guardan cuando el usuario cierra la base de datos o cuando usa el comando COMMIT, es decir , guarda los cambios que se realizan en el contenido de una relación (tabla), su sintaxis la escriben Coronel, Morris y Rob como:

```
COMMIT [WORK ]
```

Para el ejemplo de Cliente:

COMMIT; la palabra WORK puede ser escrita o no, como en la tabla Empleado se realizó un cambio (con el comando UPDATE) ahora el valor asociado al atributo nombre será *ana* y no *carlos*, dado que se guardaron los cambios.

- **ROLLBACK:** Usando este comando se restablecen los valores de los datos, es decir, si se hacen cambios en una relación y se desea regresar a como estaba en un principio se hace uso de este comando, aunque cabe aclarar que funciona si previamente no se ha usado COMMIT. Su sintaxis es solo ROLLBACK;

En el ejemplo de Empleado como ya se usó COMMIT no se pueden restablecer los valores (*carlos* en lugar de *ana* y *rodriguez* en lugar de *gonzales*), con ROLLBACK, así que en los atributos *nom\_empleado* y *apellido\_p* se tienen asociados los valores *ana* y *gonzales*.

- **Eliminar datos.** A través del comando DELETE se borran las tuplas seleccionadas, de acuerdo con Coronel, Morris y Rob se escribe:

```
DELETE FROM nombre_de_tabla  
  
[WHERE lista_de_condiciones];
```

Ejemplo:

```
DELETE FROM Empleado  
  
Where id_empleado=2;
```

Para efectos de ésta tesina se profundizará en las consultas, que son una petición (o pregunta) que se le puede hacer a 1 o más relaciones y se le hace a un SGBD esperando una respuesta para brindar información que le sea de interés al usuario, algunas sentencias básicas que se usan para el uso de consultas son:

- **SELECT:** selecciona los atributos que se piden en la clausula WHERE.
- **Símbolo \*:** Hace una lista de todos los atributos.
- **FROM:** Se usa para hacer referencia a X tabla o tablas.
- **WHERE:** se emplea para especificar los criterios de lo que se desea recuperar.

Aunque también una consulta puede hacer uso de los siguientes comandos:

- **ORDER BY:** Permite mostrar los atributos de forma ascendente o descendente.

Ejemplo:

```
SELECT nom_empleado, cod_postal
FROM Empleado
ORDER BY nom_empleado;
```

- **GROUP BY:** Agrupa los atributos que se desean mostrar, debe ir acompañado de una función agregada como COUNT, MIN, MAX, AVEG Y SUM.

Ejemplo:

```
SELECT id_empleado, COUNT(*) as nom_empleado
FROM Empleado
Where cod_postal=56239
GROUP BY cod_postal;
```

Como la sentencia GROUP BY Hace uso de una función agregada, se explicará cada una de ellas con ejemplos de la tabla de la figura 24:

- **COUNT:** Proporciona la cantidad de elementos del atributo seleccionado.

Ejemplo:

```
SELECT COUNT(nom_empleado) FROM Empleado;
```

- **MIN:** Brinda el valor asociado más pequeño de un atributo.

Ejemplo:

```
SELECT MIN(id_empleado) FROM Empleado;
```

- **MAX:** Da como resultado el valor más grande asociado a cierto atributo.

Ejemplo:

```
SELECT MAX(id_empleado) FROM Empleado;
```

- **SUM:** Realiza la suma de los valores asociados de acuerdo al atributo indicado.

Ejemplo:

```
SELECT SUM(id_empleado) as total FROM Empleado;
```

- **AVG:** Calcula el promedio de los valores asociados a un atributo seleccionado.

Ejemplo:

```
SELECT AVG(id_empleado) FROM Empleado;
```

Como se puede ver SQL es un lenguaje de consultas estructurado que es de gran ayuda; ya que facilita la creación de objetos y la manipulación de los mismos, lo que lleva a que se puedan obtener consultas de forma sencilla que son de interés para el usuario final; si se requiere conocer más específicamente las consultas acudir al apartado 2.5.6.

## **2.4 Ventajas y Desventajas**

Como ya se mencionó el modelo de datos más utilizado es el modelo relacional, el cual se utiliza en las bases de datos relacionales. A continuación se presentan ventajas:

- Es uno de los modelos más usados en el mercado, pues se dirige hacia la parte lógica y no física que puede ser muy compleja, esto le brinda al diseñador la oportunidad de desligarse de los detalles físicos.
- Además se hace uso del lenguaje estructurado SQL que es un lenguaje declarativo, el cual permite decir que hacer pero no como hacerlo, lo que le facilita al diseñador su trabajo.
- Así mismo elimina la redundancia, lo que probablemente lleve a que se generen menos errores en la base de datos y, por lo tanto, halla mayor eficiencia en los mismos.
- Se hace uso del modelo relacional, el cual es un modelo de fácil entendimiento para el diseñador dado que se usan relaciones (tabla), atributos y tuplas; si el modelo relacional se realiza de manera correcta posiblemente disminuya errores.

#### Desventajas:

- Si no se hace un buen modelo relacional posiblemente la base de datos no funcione bien, lo que puede traer costos extras.
- No permite ver el comportamiento de las relaciones.
- Las personas que no conocen sobre bases de datos pueden hacer consultas con facilidad en un ambiente gráfico, pero sin pensar en un diseño de la base de datos, esto lleva a que si la base de datos crece (y no se hizo un buen diseño), se haga más lento el sistema.

## 2.5 Bases de datos centralizadas

### 2.5.1 Conceptos básicos

Las bases de datos centralizadas entran dentro de la clasificación de una base de datos según su ubicación, para comprender porque se dice que entra dentro de ésta clasificación se definirá una base de datos centralizada.

Elmasri y Navathe (2008) comentan que “Un DBMS es centralizado si los datos están almacenados en un solo computador. Un DBMS centralizado puede soportar varios usuarios, pero el DBMS y la base de datos residen en su totalidad en un solo computador.”

Una base de datos centralizada es “una base de datos que soporte datos ubicados en un solo sitio” (Coronel, Morris y Rob, 2011, p.9).

Por lo tanto, una base de datos centralizada se encuentra ubicada físicamente en un solo sitio, aunque varios usuarios pueden tener acceso a ella, pero no por esto es una base de datos distribuida.

A continuación se explican conceptos más avanzados en el área de bases de datos:

- Índice: Es una estructura (organización) que permite acceder a las tuplas de una relación rápidamente.
- Integridad referencial: Se dice que se tiene integridad referencial cuando la llave foránea de una relación tiene el mismo valor o un valor nulo que el de una tupla que se encuentra en otra relación.
- Transacción: Consiste en realizar una o varias operaciones o procesos, ya sea que sólo se lean o se modifiquen (escriban).

- Atomicidad: Se hace una operación completa o no se hace, es decir, si se hace una transacción y por circunstancias no se completa, se aborta y regresa al inicio.
- Concurrencia: Cuando distintas actividades se ejecutan a la misma vez, en este caso transacciones.
- Disponibilidad: Significa la probabilidad de que el sistema se encuentre disponible en un tiempo.

### **2.5.2 Uso de estas bases de datos en las organizaciones**

Hoy en día las organizaciones requieren cada vez más, información ordenada, oportuna y veraz, lo cual se consigue utilizando una base de datos.

En general casi todas las empresas manejan bases de datos centralizadas para satisfacer sus necesidades, esto sin importar el sector al que estén dirigidas o el tamaño de la empresa; pues anteriormente se mencionó que una base de datos centralizada puede estar ubicada físicamente en una computadora y en la misma acceder a la base de datos o de diversas computadoras acceder a la base de datos. Por lo que bien le funciona a una farmacia que se encuentre en un sólo establecimiento hasta una tienda comercial que posea varias sucursales.

Las ventajas de la base de datos centralizada (información exacta, visualizada en el momento preciso y con la menor probabilidad de error) van a ser las mismas para cualquier caso, ya que para ambos casos se podrá conocer información importante como datos personales de los empleados, la cantidad de productos con la que se cuenta, la cantidad de productos que más se venden, descripción del producto (como fecha de caducidad), entre otras.

Además la base de datos ayudará a satisfacer las necesidades de sus clientes, lo que llevará a que la empresa aumente su productividad.

Así mismo es mucho más fácil implementar una base de datos centralizada en menos tiempo (dado que el análisis y el diseño es más sencillo) y con un costo menor (por ser menos compleja que otras bases de datos).

### **2.5.3 Sistemas de administración y bases de datos**

El sistema de administración de bases de datos es imprescindible para poder crear y usar una base de datos, como ya se mencionó es el intermediario entre el usuario final y la base de datos, además de que permite que los datos de la base de datos sean compartidos por los usuarios y los programas.

También cuenta con funciones como:

- Administración de un diccionario de datos: El DBMS contiene los atributos y características en el diccionario de datos, que es proveído por el usuario.
- Administración de almacenamiento de datos: El SGBD manipula el almacenamiento de los datos físicamente.
- Transformación y presentación de datos: El DBMS se encarga de hacer la transformación de datos lógicos al formato de datos físicos.
- Administración de seguridad: El sistema de administración de base de datos admite establecer autenticidad para los usuarios.
- Lenguaje de acceso a base de datos: El SGBD cuenta con un lenguaje de consulta estructurado que permite la definición (LDD) y manipulación de datos (LMD).

Coronel, Morris y Rob (2011) dicen:

el sistema de base de datos está formado por datos relacionados lógicamente en un solo depósito lógico. (La etiqueta “lógico” refleja el hecho de que, aun cuando el depósito de

datos parece ser una sola unidad para el usuario final, su contenido puede en realidad estar físicamente distribuido entre múltiples equipos de almacenamiento de datos o lugares). Como el depósito de datos de la base de datos es una sola unidad lógica, la base de datos representa un cambio importante en la forma en que los datos del usuario final son guardados, se tiene acceso a ellos y son administrados.

Por tanto, un SGBD se ve como una sola unidad sin importar el tipo de base de datos (centralizada o descentralizada) que esté auxiliando; en una base de datos centralizada el SGBD se encuentra instalado y trabaja en una sola computadora sin importar que ésta se encuentre comunicada a una red.

#### 2.5.4 Estructura

Como ya se ha comentado una base de datos centralizada se encuentra en un sólo lugar físico, por lo que su estructura es simple; se pueden tener dos alternativas:

- Una sola computadora con su base de datos centralizada (monousuario) que se puede ver en la figura 27.
- Una computadora (mainframe) con la base de datos centralizada, a la cual pueden acceder múltiples usuarios (multiusuario), se ilustra en la figura 28.

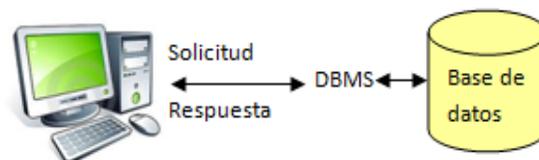


Figura 27; Base de datos centralizada monousuario.

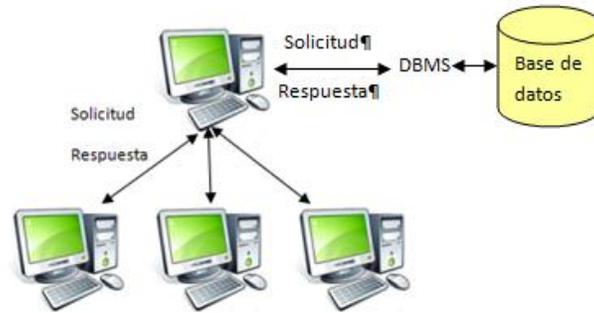


Figura 28; Base de datos centralizada multiusuario.

Cabe aclarar que el DBMS cuenta con una estructura llamada ANSI/SPARC (ver figura 29) o arquitectura de tres niveles (para una base de datos centralizada) según el grado de abstracción de los datos. Anteriormente en los conceptos básicos de las bases de datos relacionales se habló del grado de abstracción de datos en una base de datos, sin embargo, a continuación se explica más a detalle cada nivel.

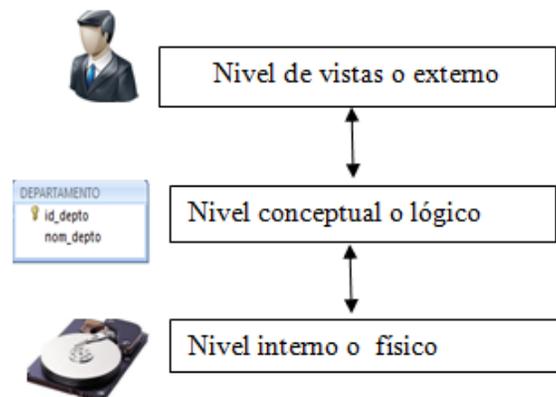


Figura 29; Arquitectura de 3 niveles.

- Nivel externo: Son las vistas que un usuario (final) percibe, para representar las vistas externas, es decir, vistas de usuario, el diseñador usa los diagramas entidad relación, (esto para identificar los datos que se requieren en cada vista y facilitar el desarrollo del

sistema) y la base de datos lo percibe como esquema externo. Es el nivel más alto de abstracción, por lo tanto, es independiente del software y del hardware.

- Nivel conceptual: El administrador de base de datos describe toda la base de datos mediante relaciones y atributos (modelo relacional), lo que facilita la elaboración de una base de datos; se dice que es dependiente del software pero independiente del hardware y que existe una independencia lógica de datos
- Nivel físico: Tiene que ver con guardar la base de datos en un medio de almacenamiento físico generalmente secundario (disco duro). En ocasiones el diseñador puede requerir trabajar en este nivel para aumentar la eficiencia de la base de datos mediante métodos de acceso a los datos. Es dependiente del software y del hardware, además posee independencia física de datos, pues se puede cambiar el esquema interno sin tener que cambiar el esquema conceptual y externo.

### **2.5.5 Almacenamiento**

Como ya se mencionó el SGBD es el medio que permite la comunicación entre el usuario y la base de datos; por lo tanto, a través de éste se pueden hacer actividades como insertar, cambiar, actualizar, eliminar y guardar datos de una base de datos, sin embargo, estas actividades son transparentes para el usuario final. Aunque cabe decir que un diseñador y administrador de base de datos trabaja con el DBMS a un nivel lógico (por lo general no se preocupa del nivel físico), es importante (en menor medida) conocer las técnicas de almacenamiento en disco (nivel físico), ya que con ello puede incrementar el rendimiento de la base de datos.

Una base de datos puede ser almacenada en medios de almacenamiento distintos clasificados en:

- Medio primario: Son medios de almacenamiento que operan en la unidad central de procesamiento (CPU) como la memoria cache (acelera la ejecución de los programas) y la memoria principal (o DRAM, RAM dinámica, que posee los programas y datos a utilizar), además son volátiles (se pierde la información cuando no hay luz).
- Medio secundario: Los cuales son no volátiles (no pierde su contenido cuando se pierde la luz) por lo que permiten que los datos se guarden permanentemente; los medios secundarios se dividen en almacenamiento magnético (discos flexibles, discos duros y cintas magnéticas) y almacenamiento óptico (CD).

Cabe mencionar que por lo general una base de datos se guarda en un medio de almacenamiento secundario magnético (disco duro) ya que las bases de datos suelen ser grandes y no cabrían en la memoria principal, además de que ésta es volátil; por tanto se hablará del almacenamiento de los datos en la base de datos en un disco duro (almacenamiento físico o nivel físico).

Es preciso decir que un disco duro está fabricado de material magnético en forma de platos (discos), así que cada plato está compuesto por pistas o *tracks* (se puede decir que son varios anillos unidos de diferentes tamaños) cada pista está dividida en sectores conocidos también como bloques o páginas (normalmente de 512 bytes), una agrupación de varios sectores se conoce como *cluster* y se le llama cilindro a las pistas que tienen el mismo tamaño en los platos (ver figuras 30 y 31)

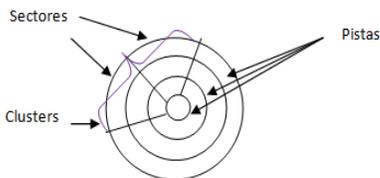


Figura 30; Sectores, clusters y pistas

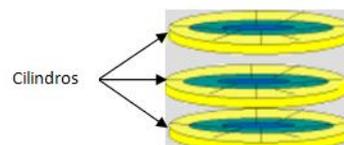


Figura 31; cilindros

Para almacenar los datos de una base de datos en el disco, los datos se organizan como archivos (secuencia de registros<sup>14</sup>), por lo tanto los datos se almacenan en forma de registros y éstos se almacenan en bloques del disco duro, posteriormente se hace una copia de el o los bloques que se necesiten en el búfer y de ahí pasa a la memoria principal, no obstante se puede tardar al hacer una búsqueda sobre algún dato en específico, por ello se usan técnicas para almacenar los registros de un archivo en el disco (métodos de organización de los datos) como la técnica del doble búfer, registros desordenados, registros ordenados, registros dispersos y los árboles B; ya que “Los árboles B+ surgen como una variante de los árboles B, para hacer más eficiente su uso ante las necesidades de implementación de un manejador de bases de datos”(Quiroga y Martínez, s.f, p.173) se explicarán a continuación.

Elmasri y Navathe (2007) dicen “Un **árbol** está formado por **nodos**. Cada nodo del árbol, excepto el nodo especial denominado **raíz**, tiene un nodo **padre** y varios (ninguno o más) nodos **hijo**. El nodo raíz no tiene padre. Un nodo que no tiene ningún hijo se llama nodo **hoja**; un nodo que no es hoja es un nodo **interno**. El **nivel** de un nodo siempre es uno más que el nivel de su padre, siendo cero el nivel del nodo raíz.”

Para entender el funcionamiento de un árbol B+, es necesario conocer los índices multiniveles, ya que esta estructura de datos trabaja con ellos.

El índice multinivel usa un archivo de índice (contiene registros de índices o entradas de índices y punteros a los bloques del archivo de datos) a este primer archivo de índice se le llama primer nivel (o base de un índice multinivel), el archivo de índice esta ordenado y tiene un valor distinto por cada clave (campo) de búsqueda (valor único que se requiere), entonces se crea un índice

---

<sup>14</sup> Normalmente describen entidades y sus atributos. Por ejemplo, un registro EMPLEADO representa una entidad empleado, y el valor de cada campo del registro especifica algún atributo de ese empleado como es *Nombre*, *FechaNac*, *Sueldo* o *supervisor*.

principal para el primer nivel, a este índice se le llama segundo nivel de índice multinivel, como el segundo nivel es un índice principal se usan anclas de bloque<sup>15</sup>, o sea que el segundo nivel tiene una entrada por cada bloque del primer nivel, si se requiere usar otro nivel se tiene que crear otro archivo de índices y así sucesivamente.

La característica principal de un árbol B+ es que los punteros a datos se almacenan solamente en los nodos hoja y éstos están enlazados, con ello se tiene un acceso ordenado a los registros mediante el campo de búsqueda.

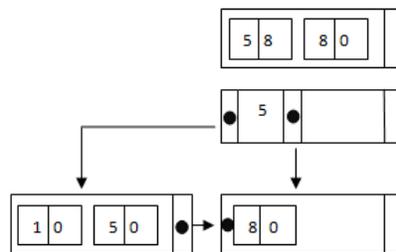


Figura 32; Árbol B+.

En la figura 32 se muestra un claro ejemplo de un árbol B+, como se aprecia el primer nodo hoja está enlazado al segundo nodo hoja.

### 2.5.6 Consultas

Anteriormente se explicó que una consulta es hacer una solicitud mediante el SGBD a una o varias relaciones con el fin de obtener información relevante para el usuario.

Las bases de datos en el mercado que se usan con mayor frecuencia son las relacionales (RDBMS), las cuales trabajan con el lenguaje SQL (anteriormente explicado) para la realización de consultas.

<sup>15</sup> El primer registro de cada bloque del fichero de datos es el registro ancla, o simplemente ancla de bloque.

Las consultas se pueden obtener de una relación (unaria), de dos relaciones (binarias) o de tres o más relaciones (enarias) esto con el fin de satisfacer las necesidades de los usuarios al requerir datos de diferentes relaciones lo que brinda información más completa.

A continuación se proporcionan algunos ejemplos de consultas unarias utilizando la tabla de la figura 24.

id_employado	nom_employado	apellido_p	apellido_m	cod_postal
1	carlos	rodriguez	rodriguez	56240
2	rocio	sosa	ortiz	56245
3	samuel	lopez	vidal	56293
4	mireya	estrada	gonzalez	56242
5	melissa	campos	sosa	56239

Figura 24; Tabla Empleado.

- 1) Se desea conocer los datos (nombre y apellidos) de todos los empleados

```
SELECT * FROM Empleado;
```

Se obtendrá:

id_employado	nom_employado	apellido_p	apellido_m	cod_postal
1	carlos	rodriguez	rodriguez	56240
2	rocio	sosa	ortiz	56245
3	samuel	lopez	vidal	56239
4	mireya	estrada	gonzales	56242
5	melissa	campos	sosa	56239

Figura 33; Consulta 1.

- 2) Se requiere conocer los empleados cuyo apellido paterno sea *lopez*

```
SELECT nom_empleado, apellido_p

FROM Empleado

WHERE apellido_p='lopez';
```

El resultado será:

nom_empleado	apellido_p
samuel	lopez

Figura 34; Consulta 2.

- 3) Se necesita saber los datos (id, nombre y apellidos de los empleados que tengan por id 1 o 3 o 5).

```
SELECT id_empleado, nom_empleado, apellido_p, apellido_m

FROM Empleado

Where id_empleado=1 OR id_empleado=3 OR id_empleado=5;
```

Lo que mostrará:

id_empleado	nom_empleado	apellido_p	apellido_m
1	carlos	rodriguez	rodriguez
3	samuel	lopez	vidal
5	melissa	campos	sosa

Figura 35; Consulta 3.

Ahora se muestran algunas consultas binarias:

4) Se solicita saber que empleados(apellidos y nombres) viven en la colonia de la purificación

```
SELECT Empleado.apellido_p, Empleado.apellido_m,  
Empleado.nom_empleado, Localidad.colonia  
  
FROM Empleado, Localidad  
  
WHERE empleado.cod_postal=localidad.cod_postal and colonia='la  
purificacion';
```

Aparecerá lo siguiente:

apellido_p	apellido_m	nom_empleado	colonia
rodriguez	rodriguez	carlos	la purificacion

Figura 36; Consulta 4.

5) Se requiere conocer los datos (apellidos y nombres) de los empleados que viven en la dirección de Benito juarez s/n o el cerrito no.3

```
SELECT Empleado.apellido_p, Empleado.apellido_m,  
Empleado.nom_empleado, Domicilio.direccion  
  
FROM Empleado, Domicilio  
  
WHERE empleado.id_empleado=domicilio.id_empleado and  
(direccion='benito juarez s/n' or direccion='el cerrito no.3');
```

Dando como resultado:

apellido_p	apellido_m	nom_Empleado	direccion
lopez	vidal	samuel	benito juarez s/n
campos	sosa	melissa	el cerrito no.3

Figura 37; Consulta 5.

A continuación se dan ejemplos de consultas terciarias:

- 6) Se desea conocer los datos (colonia y dirección) de los empleados carlos y rocio (también mostrar los nombres de los empleados)

```
SELECT localidad.colonia, domicilio.direccion,
Empleado.nom_Empleado
FROM Localidad, Domicilio, Empleado
WHERE empleado.cod_postal=localidad.cod_postal and
empleado.id_Empleado=domicilio.id_Empleado and
(nom_Empleado='carlos' or nom_Empleado='rocio');
```

Lo que mostrará:

colonia	direccion	nom_Empleado
la purificacion	huexotitla no.12	carlos
tepetlaoxtoc	arandas no.2	rocio

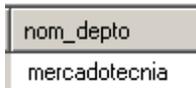
Figura 38; Consulta 6.

- 7) Se desea saber en qué departamento trabaja el empleado rocío

```
SELECT departamento.nom_depto
FROM Departamento, Rel_EmpDep, Empleado
```

```
WHERE empleado.id_empleado=rel_empdep.id_empleado and  
rel_empdep.id_depto=departamento.id_depto and  
nom_empleado='rocio';
```

Aparecerá lo siguiente:



nom_depto
mercadotecnia

Figura 39; Consulta 7.

Los comandos más utilizados son: SELECT, FROM, WHERE, UPDATE, INSERT, DELETE, entre otros.

Para conocer la descripción de estas cláusulas y cómo funcionan (ejemplos) revisar el apartado de SQL que se encuentra anteriormente.

### **2.5.7 Disponibilidad**

El administrador de base de datos es el encargado de asegurarse que una base de datos tenga una buena disponibilidad, es decir, que la base de datos permanezca activa para que los usuarios puedan acceder a ella en el momento que se requiera.

“Para garantizar la disponibilidad de datos, todo el sistema (no sólo el componente de datos debe ser protegido de degradación o interrupción de servicio causadas por cualquier fuente (interna o externa).” (Coronel, Morris y Rob, 2011, p.629).

Una base de datos centralizada ofrece un nivel bajo de disponibilidad dado que la base de datos se encuentra en un sólo sitio central, es decir, si falla la base de datos por alguna razón (falla del disco duro, caída del sistema, error en software de aplicación) ningún usuario que tenga los

privilegios de acceder a la base de datos (autenticación) podrá tener acceso a ésta para la realización de sus tareas, además en caso de un desastre natural se puede perder la base de datos, y por tanto su disponibilidad, por ello es importante realizar respaldos de la base de datos y que éstos sean recientes.

En caso de que se haga uso de una base de datos centralizada a la cual accedan varios clientes se tiene la probabilidad de que haya fallo en la red de comunicación, aunque el DBMS debe tener la capacidad para la recuperación de la misma, el administrador de base de datos debe coordinarse con especialistas de redes para minimizar el margen de error en la red.

Así mismo la disponibilidad de la base de datos se puede ver afectada si se realiza mantenimiento de la misma, es por eso, que el administrador de base de datos debe buscar herramientas que le permitan realizar el mantenimiento mientras la base de datos se encuentra disponible o buscar herramientas que utilicen poco tiempo para realizar el mantenimiento.

## **2.6 Bases de datos distribuidas**

### **2.6.1 Conceptos básicos**

Las bases de datos distribuidas son cada vez más utilizadas a pesar de que son más complejas, a diferencia de las centralizadas, éstas se encuentran físicamente distribuidas en varios sitios.

Coronel, Morris y Rob definen una base de datos distribuida como “Una base de datos que soporte datos distribuidos en varios lugares diferentes.”

“Un DBMS distribuido (DDBMS) puede tener la base de datos y el software DBMS distribuidos por muchos sitios, conectados por una red de computadoras”. (Elmasri y Navathe, 2008, p 44).

A continuación se explican algunos conceptos importantes para entender el funcionamiento de una base de datos distribuida.

- DDBMS: Así se le llama al sistema que permite el manejo de la base de datos distribuida, además de contar con las funciones que brinda un DBMS debe brindar otras funciones como control de concurrencia, transparencia de distribución, entre otros.
- Fragmentos de bases de datos: Son todos los fragmentos que componen la base de datos.
- Fragmentación de datos: Es la descomposición de una relación (tabla) en diversas partes llamadas fragmentos.
- Fragmentación horizontal: Una relación se divide en fragmentos por tuplas.
- Fragmentación vertical: Una relación se divide en fragmentos por atributos.
- Fragmentación mixta, híbrida o combinada: Una relación puede dividirse primero en forma horizontal (por tuplas) y luego en vertical (por atributos), llamada fragmentación HV; o una relación se puede dividir en fragmentos verticales y posteriormente en fragmentos horizontales, llamada VH .
- Procesamiento distribuido: En diversos lugares se procesa la base de datos.

### **2.6.2 Uso de estas bases de datos en las Organizaciones**

Las empresas que requieren compartir información entre diferentes lugares geográficos y necesitan que esto se haga de manera eficiente hacen uso de las bases de datos distribuidas.

Anteriormente se mencionó que una base de datos distribuida está compuesta por fragmentos, los cuales pueden ser verticales (por atributos), horizontales (por tuplas) o mixtos; a continuación se muestra un ejemplo de cada fragmentación tomando como base a un banco X que tiene sucursales en el estado de México, en Monterrey y en otros lugares, es decir los datos están

distribuidos en fragmentos de base de datos en México, en Monterrey y en los demás lugares (sitios), para este caso se usarán solo los sitios de México y Monterrey, cada uno cuenta con las relaciones Sucursal (con los atributos *id\_sucursal*, *nom\_sucursal*, *ciudad* y *dirección*), Cuenta (con los atributos *num\_cuenta* y *saldo*), Cliente (con sus atributos *id\_cliente*, *nom\_cliente*, *apellido\_p*, *apellido\_m*, *dirección* y *telefono*) RSucuCuenta( con los atributos *id\_sucursal* y *num\_cuenta*) y RCuentaCliente (con sus atributos *id\_cliente*, *num\_cuenta*)

id_sucursal	nom_sucursal	ciudad	dirección
1	trinidad	estado de mexico	texcoco
2	san jose	monterrey	nezahualcoyotl
3	centro	monterrey	quezada
4	norte	monterrey	benito juarez
5	hidalgo	estado de mexico	ecatepec
6	trinidad	monterrey	lopez portillo
7	sur	estado de mexico	coacalco

Figura 40; Sucursal

id_sucursal	num_cuenta
1	8242
1	1520
1	2030
2	15178
2	16182
3	30543
4	40506
5	56012
6	60248
7	71147

Figura 41; RSucuCuenta.

num_cuenta	saldo
8242	200
1520	30000
2030	4000
15178	20000
16182	100
30543	300
40506	78500
56012	6432
60248	2450
71147	5780

Figura 42; Cuenta.

id_cliente	num_cuenta
1	8242
2	1520
3	2030
4	15178
5	16182
6	30543
7	40506
8	56012
9	60248
10	71147

Figura 43; RCuentaCliente.

id_cliente	nom_cliente	apellido_p	apellido_m	dirección	telefono
1	carlos	gonzalez	lopez	la conchita	92806
2	rocio	loza	diaz	trinidad	95504
3	ana	rodriguez	rodriguez	venustiano	94809
4	pablo	vidal	perez	los sauces	94528
5	itzel	lopez	lopez	colibris	96833
6	juan	garcia	ramirez	jacaranda	97858
7	tania	fernandez	miranda	san martin	95611
8	liliana	romero	prado	9 de julio	92805
9	eduardo	mendoza	garcia	venustiano	96444
10	carmen	quezada	martinez	carranza	95701

Figura 44; Cliente.

Para ilustrar la fragmentación horizontal suponga que en el caso de que se requiera conocer la información sobre las sucursales que tiene Monterrey y las sucursales que tiene el Estado de México (criterio de selección) se hace uso de una fragmentación horizontal tomando como base

la figura 40, pues se requiere conocer la información (tuplas) perteneciente a la ciudad de Monterrey y al Estado de México, dando como resultado dos fragmentos el perteneciente a Monterrey y el de él Estado de México que se pueden ver en las figuras 45 y 46.

id_sucursal	nom_sucursal	ciudad	dirección
2	san jose	monterrey	nezahualcoyotl
3	centro	monterrey	quezada
4	norte	monterrey	benito juarez
6	trinidad	monterrey	lopez

id_sucursal	nom_sucursal	ciudad	dirección
1	trinidad	estado de mexico	texcoco
5	hidalgo	estado de mexico	ecatepec
7	sur	estado de mexico	coacalco

Figura 45; Fragmento horizontal de Monterrey      Figura 46; Fragmento horizontal del Estado de México

Para mostrar la fragmentación vertical pudiese ser que en Monterrey sólo se necesite conocer el nombre del cliente y sus apellidos; pero en el Estado de México se pudiese requerir sólo la dirección del cliente y el teléfono, por lo que se haría uso de una fragmentación vertical (por atributos de la figura 44) dando como resultado dos fragmentos, uno que va a estar compuesto por el *id\_cliente*, *nom\_cliente*, *apellido\_p* y *apellido\_m* (ver figura 47); y el otro formado por *id\_cliente*, *dirección* y *teléfono* que se muestra en la figura 48.

id_cliente	nom_cliente	apellido_p	apellido_m
1	carlos	gonzalez	lopez
2	rocio	loza	díaz
3	ana	rodriguez	rodriguez
4	pablo	vidal	perez
5	itzel	lopez	lopez
6	juan	garcia	ramirez
7	tania	fernandez	miranda
8	liliana	romero	prado
9	eduardo	mendoza	garcia
10	carmen	quezada	martinez

id_cliente	dirección	telefono
1	la conchita	92806
2	trinidad	95504
3	venustiano	94809
4	los sauces	94528
5	colibris	96833
6	jacaranda	97858
7	san martin	95611
8	9 de julio	92805
9	venustiano	96444
10	carranza	95701

Figura 47; Fragmento vertical de Monterrey      Figura 48; Fragmento vertical del Estado de México

En el caso de que se solicite saber la información de las sucursales (*nombre*) respectivas de Monterrey y del Estado de México tomando la figura 40 se usa una fragmentación mixta (horizontal-vertical), pues se requiere conocer la ciudad de las sucursales y el nombre de cada una de ellas. (Ver figuras 49 y 50).

id_sucursal	nom_sucursal	ciudad
2	san jose	monterrey
3	centro	monterrey
4	norte	monterrey
6	trinidad	monterrey

id_sucursal	nom_sucursal	ciudad
1	trinidad	estado de mexico
5	hidalgo	estado de mexico
7	sur	estado de mexico

Figura 49; Fragmento HV1 de Monterrey (S1)

Figura 50; Fragmento HV1 del Estado de México (S2)

id_sucursal	dirección
2	nezahualcoyotl
3	quezada
4	benito juarez
6	lopez portillo

id_sucursal	dirección
1	texcoco
5	ecatepec
7	coacalco

Figura 51; Fragmento HV2 de Monterrey

Figura 52; Fragmento HV2 del Estado de México

Así mismo se puede presentar una fragmentación mixta (vertical-horizontal) si se desea saber sólo el nombre de las sucursales que están en Monterrey, ya que se aplicó esta fragmentación da como resultado dos fragmentos de base de datos (ver figuras 53 y 54), y si también se requiere conocer el nombre de las sucursales que están en el Estado de México da como resultado las figuras 55 y 56.

id_sucursal	nom_sucursal	ciudad
2	san jose	monterrey
3	centro	monterrey
4	norte	monterrey
6	trinidad	monterrey

Figura 53; Fragmento VH1 de Monterrey

id_sucursal	dirección
2	nezahualcoyotl
3	quezada
4	benito juarez
6	lopez portillo

Figura 54; Fragmento VH2 de Monterrey

id_sucursal	nom_sucursal	ciudad
1	trinidad	estado de mexico
5	hidalgo	estado de mexico
7	sur	estado de mexico

Figura 55; Fragmento VH del Estado de México

id_sucursal	dirección
1	texcoco
5	ecatepec
7	coacalco

Figura 56; Fragmento VH del Estado de México

Cabe mencionar que cada fragmento puede estar ubicado en diferentes sitios, nodos o emplazamientos, lo que aumenta la fiabilidad (menor grado de error) y la disponibilidad de que el sistema se encuentre funcionando, del mismo modo incrementa la confiabilidad, dado que se tiene mejor control de los datos almacenados localmente, de la misma manera aumenta el rendimiento de la base de datos, ya que hay una distribución de trabajo.

### 2.6.3 Sistemas de administración de bases de datos

En bases de datos distribuidas, los sistemas de administración de base de datos (DBMS) se llaman sistemas de administración de base de datos distribuidas (DDBMS).

“Un DDBMS gobierna el almacenamiento y procesamiento de datos lógicamente relacionados, mediante sistemas interconectados computarizados en los que los datos y funciones de procesamiento están distribuidos en varios sitios.” (Coronel, Morris y Rob, 2011, p.485)

Prácticamente un DDBMS tiene las mismas funciones que un DBMS del cual se habló anteriormente en las bases de datos centralizadas, es decir, debe contar con las funciones de un DBMS, sin embargo es imprescindible que cuente con otras funciones para poder soportar una base de datos distribuida tales como:

- Seguimiento de datos: Debe ser capaz de controlar la distribución, fragmentación y replicación de los datos.
- Procesamiento de consultas distribuidas: Es la posibilidad de acceder a sitios, así como realizar consultas en los sitios a través de una red de comunicación.
- Administración de transacciones distribuidas: Diseñar consultas y transacciones con la capacidad de acceder a los datos desde diferentes sitios.
- Administración de datos replicados: distinguir a que copias de los datos se puede acceder, manteniendo la consistencia.
- Recuperación de una base de datos distribuida: Debe contar con la capacidad de recuperarse de posibles fallos, es decir, si se interrumpe la comunicación debe contar con la capacidad de poder recuperarla.
- Capacidad de mapeo: Para poder localizar los datos en los fragmentos locales o remotos.
- Seguridad: Privacidad en las bases de datos (locales o remotas).
- Control de concurrencia: Administración del acceso simultáneo.

Así mismo un DDBMS tiene los siguientes componentes:

- Estaciones de trabajos y equipos remotos.
- Componentes de la red (computadoras, sistemas operativos, medios de comunicación).
- Procesador de transacciones o administrador de transacciones.

- Procesador de datos o administrador de datos.

Como se puede apreciar los DDBMS toman un papel importante en las bases de datos distribuidas, para que éstas puedan realizar su trabajo de manera eficiente ayudando a controlar a través de sus componentes las funciones de una base de datos distribuida.

#### 2.6.4 Estructura

Como se mencionó anteriormente una base de datos distribuida está compuesta por fragmentos de bases de datos, los cuales se encuentran ubicados en sitios diferentes (por lo que el diseño de la base de datos es más complicado).



Figura 57; Base de datos distribuida

En la figura 57 se aprecia que se tienen tres sitios, España, México y Brasil; cada sitio cuenta con fragmentos de base de datos, los cuales pueden ser horizontales, verticales y/o mixtos. Para que los sitios tengan comunicación entre sí están conectados a través de una red de comunicaciones

Como la estructura de la base de datos distribuida como tal es diferente a una base de datos centralizada, de alguna manera también cambia la estructura del DBMS al DDBMS.

Tomando como base la arquitectura de 3 niveles (ANSI/SPARC) para el DDBMS quedaría la siguiente arquitectura:

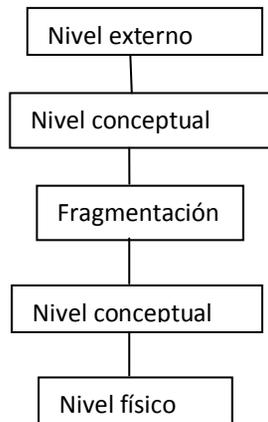


Figura 58; Arquitectura para bases de datos distribuidas.

En la figura 58 se muestran los niveles que se podrían adaptar a la arquitectura de 3 niveles, a continuación una breve descripción de cada nivel.

Nivel externo global: Son las vistas que se le muestran al usuario y pueden ser  $n$  vistas.

Nivel conceptual global: Se hace una descripción lógica de la base de datos distribuida como si fuera una centralizada.

Fragmentación: Es la manera en que las relaciones se van a dividir lógicamente.

Nivel conceptual local: Se detalla la base de datos local.

Nivel físico local: Almacenamiento físico de las bases de datos locales.

### 2.6.5 Almacenamiento

A diferencia de las bases de datos centralizadas, el almacenamiento de una base de datos distribuida se realiza en diferentes sitios (nodos o emplazamientos), por lo que se tiene que decidir cómo dividir la base de datos en fragmentos, ver los fragmentos que se van a replicar y la ubicación de los fragmentos y las réplicas.

Como ya se mencionó anteriormente los fragmentos de la base de datos pueden ser horizontales, verticales o mixtos; así mismo puede haber copias de estos fragmentos en diversos nodos, es decir réplica de datos (ver figura 59).

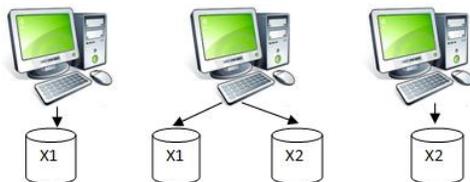


Figura 59; Replica de datos.

Para decidir la ubicación de los fragmentos y las replicas se usa la asignación de datos<sup>16</sup>, cabe aclarar que se puede usar sólo la fragmentación o sólo la réplica de datos o una combinación de las dos de acuerdo a las necesidades del usuario; para poder llevar a cabo la asignación de datos es necesario tomar en cuenta el tamaño de la base de datos, la frecuencia de uso de ciertos fragmentos y los costos para implementar la base de datos.

### 2.6.6 Consultas

En un DDBMS es más difícil realizar consultas, ya que la base de datos se encuentra dividida en fragmentos y además éstos pueden estar replicados, lo que puede causar problemas para decidir a qué fragmento acceder cuando se requiera llevar a cabo un proceso.

En caso de que un DDBMS no cuente con la transparencia de distribución, el usuario deberá realizar las consultas directamente con los fragmentos adecuados, es decir, especificando el fragmento o los fragmentos que se necesitan para realizar la consulta; en caso contrario, que el DDBMS soporte la transparencia de distribución o ubicación (que sería lo ideal), el usuario podrá realizar una consulta como si fuera un DBMS (por lo que el usuario no se da cuenta de que está trabajando en una base de datos distribuida).

<sup>16</sup> Describe el proceso de decidir dónde ubicar los datos.

Elmasri y Navathe (2007) comentan “Para las consultas, un módulo de descomposición de consulta debe partir, o descomponer, esas consultas en subconsultas que puedan ejecutarse individualmente en los distintos sitios. Por otro lado, es preciso generar un procedimiento para combinar los resultados de estas subconsultas y formar así el resultado final de la consulta global”

Como se mencionó anteriormente las consultas en un DDBMS con transparencia de distribución se hacen de la misma forma que en una base de datos centralizada, para ver ejemplos de estas consultas regresar al subtema 2.5.6 consultas.

Dado que una consulta en un DDBMS sin transparencia de ubicación es más compleja a continuación se muestran algunos ejemplos.

- 1) Se requiere saber los datos (*id\_sucursal*, *nom\_sucursal*, *ciudad*, y *direccion*) de las sucursales de Monterrey y del estado de México, basándose en las figuras 45 y 46 (fragmentación horizontal), esto suponiendo que:

R1= *id\_sucursal*, *nom\_sucursal*, *ciudad*, *direccion*.

R2= *id\_sucursal*, *nom\_sucursal*, *ciudad*, *direccion*.

Consulta:

```
SELECT id_sucursal, nom_sucursal, ciudad, direccion FROM R1
```

```
UNION
```

```
SELECT id_sucursal, nom_sucursal, ciudad, direccion FROM R2;
```

2) Se desea conocer toda la información de los clientes (*id\_cliente*, *nom\_cliente*, *apellido\_p*, *apellido\_m*, *dirección* y *teléfono*) tomando las figuras 47 y 48 (que son fragmentos verticales), para ello:

R1=*id\_cliente*, *nom\_cliente*, *apellido\_p*, *apellido\_m*.

R2 =*id\_cliente*, *dirección*, *teléfono*.

Consulta:

```
SELECT id_cliente, nom_cliente, apellido_p, apellido_m,
dirección, teléfono

FROM R1, R2

WHERE R1.id_cliente=R2.id_cliente;
```

3) Se necesitan todos los datos de todas las sucursales (Monterrey y del Estado de México), esto retomando las figuras 49, 50, 51 y 52 (fragmentación mixta HV)

Subconsulta:

```
R1=SELECT id_sucursal, nom_sucursal, ciudad FROM R1

UNION

SELECT id_sucursal, nom_sucursal, ciudad FROM R2;
```

Subconsulta:

```
R2=SELECT id_sucursal, direccion FROM R3
```

UNION

```
SELECT id_sucursal, direccion FROM R4
```

**Subconsulta:**

```
SELECT id_sucursal, nom_sucursal, ciudad, direccion
```

```
FROM R1, R2
```

```
WHERE R1.id_sucursal= R2.id_sucursal;
```

4) Al igual que en la consulta 4 se necesitan todos los datos de todas las sucursales

(Monterrey y del Estado de México), pero ahora retomando las figuras 53, 54, 55 y 56 (fragmentación mixta VH).

**Subconsulta:**

```
SELECT id_sucursal, nom_sucursal, ciudad, direccion
```

```
FROM R1, R2
```

```
WHERE R1.id_sucursal= R2.id_sucursal;
```

**Subconsulta:**

```
R1=SELECT id_sucursal, nom_sucursal, ciudad FROM R1
```

UNION

```
SELECT id_sucursal, nom_sucursal, ciudad FROM R2;
```

Subconsulta:

```
R2=SELECT id_sucursal, direccion FROM R3
```

UNION

```
SELECT id_sucursal, direccion FROM R4
```

- 5) Basándose en el ejemplo que fue dado en el subtema 2.6.2 supongase que se necesita recuperar el nombre de las sucursales que están en Monterrey y en el Estado de México (Figuras 49 y 50), así como las cuentas y saldos de cada sucursal y todos los datos de sus clientes respectivos esto intuyendo que la relación de sucursal está fragmentada en 2 particiones horizontales.

num_cuenta	saldo
15178	20000
16182	100
30543	300
40506	78500
60248	2450

num_cuenta	saldo
8242	200
1520	30000
2030	4000
56012	6432
71147	5780

Figura 60; RSucuCuenta de Monterrey (RC1)    Figura 61; RSucuCuenta del Estado de México (RC2)

id_sucursal	num_cuenta
2	15178
2	16182
3	30543
4	40506
6	60248

id_sucursal	num_cuenta
1	8242
1	1520
1	2030
5	56012
7	71147

Figura 62; Cuenta de Monterrey (CUENTA1)

Figura 63; Cuenta del Estado de México (CUENTA 2)

id_cliente	num_cuenta
4	15178
5	16182
6	30543
7	40506
9	60248

id_cliente	num_cuenta
1	8242
2	1520
3	2030
8	56012
10	71147

Figura 64; Cuenta del cliente de Monterrey (CC1). Figura 65; Cuenta del cliente del Estado de México (CC2)

id_cliente	nom_cliente	apellido_p	apellido_m	dirección	telefono
4	pablo	vidal	perez	los sauces	94528
5	itzel	lopez	lopez	colibris	96833
6	juan	garcia	ramirez	jacaranda	97858
7	tania	fernandez	miranda	san martin	95611
9	eduardo	mendoza	garcia	venustiano	96444

Figura 66; Clientes de Monterrey (CLIE1).

id_cliente	nom_cliente	apellido_p	apellido_m	dirección	telefono
1	carlos	gonzalez	lopez	la conchita	92806
2	rocio	loza	diaz	trinidad	95504
3	ana	rodriguez	rodriguez	venustiano	94809
8	liliana	romero	prado	9 de julio	92805
10	carmen	quezada	martinez	carranza	95701

Figura 67; Clientes del estado de México (CLIE2).

Subconsulta:

```
R1= SELECT nom_sucursal FROM S1
```

UNION

```
SELECT nom_sucursal FROM S2;
```

Subconsulta:

```
R2= SELECT id_sucursal, num_cuenta FROM RC1
```

UNION

```
SELECT id_sucursal, num_cuenta FROM RC2;
```

**Subconsulta:**

```
R3= SELECT num_cuenta, saldo FROM CUENTA1
```

UNION

```
SELECT num_cuenta, saldo FROM CUENTA2;
```

**Subconsulta:**

```
R4= SELECT id_cliente, num_cuenta FROM CC1
```

UNION

```
SELECT id_cliente, num_cuenta FROM CC2;
```

**Subconsulta:**

```
R5= SELECT id_cliente, nom_cliente, apellido_p, apellido_m,  
direccion, telefono FROM CLIE1
```

UNION

```
SELECT id_cliente, nom_cliente, apellido_p, apellido_m,  
direccion, telefono FROM CLIE2;
```

Una vez que se realizaron las subconsultas lo que sigue es unir las para brindar la información solicitada.

```

SELECT nom_sucursal, num_cuenta, saldo, nom_cliente, apellido_p,
apellido_m, direccion, telefono

FROM R1, R2, R3, R4, R5

WHERE R1.id_sucursal=R2.id_sucursal and
R2.num_cuenta=R3.num_cuenta and R3.num_cuenta= R4.num_cuenta and
R4.id_cliente=R5.id_cliente;

```

### **2.6.7 Disponibilidad**

La disponibilidad es una de las grandes ventajas que ofrecen las bases de datos distribuidas debido a que si falla un sitio los demás sitios continúan funcionando, esto se atribuye a que el DDBMS debe contar con lo que se llama independencia de falla<sup>17</sup>.

Así mismo la disponibilidad se da porque además de que la base de datos está distribuida en diferentes sitios (en fragmentos) también estos fragmentos pueden estar replicados, esto es que puede haber copias de los fragmentos de bases de datos en diferentes lugares, por lo que si se requiere acceder a información de algún sitio pero éste no se encuentra funcionando, se puede acceder a dicha información en otro sitio.

Silberchatz, Korth y Sudarshan (2007) dicen “Uno de los objetivos del empleo de bases de datos distribuidas es disfrutar de una disponibilidad elevada; es decir, la base de datos debe funcionar casi todo el tiempo. En concreto, dado que los fallos son más probables en los sistemas distribuidos de gran tamaño, las bases de datos distribuidas deben seguir funcionando aunque

---

<sup>17</sup> El sistema no es afectado por fallas en nodos. El sistema está en operación continua incluso en el caso de una falla de nodo o una expresión en la red.

sufran diferentes tipos de fallos. La posibilidad de continuar funcionando incluso durante los fallos se denomina robustez.”

Aunque el DDBMS debe ser robusto se mencionarán los posibles problemas que podrían surgir

- Mensaje erróneo, el mensaje puede no estar bien.
- Perdida de mensaje, es decir si del sitio 1 se manda un mensaje al sitio 2 y por x circunstancia no llegase a su destino se tendría que volver a enviar el mensaje.
- Incertidumbre del mensaje, o sea que se manda un mensaje del sitio 1 al sitio 2 pero no se conoce si el mensaje llego a su destino, puede ser que el mensaje no se haya recibido o que el mensaje se haya recibido pero no hubo una respuesta de confirmación.
- Fallo en la red, es decir, puede no haber comunicación entre los sitios.
- Fallo de una transacción, por lo que se presenta una falla del sitio.

Se recomienda que para solucionar estos problemas independientemente del DBMS el administrador de base de datos debe elegir dos o más servidores con fragmentos distintos y que más de un fragmento actualice un mismo elemento de datos replicado, pues con ello se reducirán los problemas mencionados en este apartado.

### **Capítulo III Análisis y comparación, las bases de datos centralizadas contra las bases de datos distribuidas en los sistemas ERP.**

Para realizar este capítulo se utilizará el ERP5 (desarrollado por Nexedi), dado que es un sistema empresarial *opensource* (de código abierto), es decir, es un software cuyo código puede ser modificado libremente, así como distribuido por los usuarios, trabaja bajo la licencia propia de zope ZPL, la cual es compatible con la licencia GPL, dado que este sistema empresarial es de código abierto, se puede encontrar mayor documentación que en un sistema empresarial propietario, es por ello que se eligió este ERP.

La documentación que se puede encontrar en ERP5.com es documentación para el usuario, a través de *one student one ERP* (OSOE), que es un proyecto colaborativo en el que se puede contribuir, es decir, en este proyecto se encuentran alumnos que tienen conocimientos sobre ERP5 y proporcionan información al usuario, además se cuenta con documentación en forma de FAQ (se refiere a las preguntas y respuestas que se hacen frecuentemente), *HowTo* (procedimiento corto para poder realizar un proceso en específico), directrices (documentos que orientan al principio, es decir, son documentos que dan una introducción o guía para usar el código, documentación, entre otros de ERP5) y con código abierto.

Así mismo ERP5 es un software gratuito que se basa en Zope que es un servidor de aplicaciones web<sup>18</sup> de código abierto, el cual trabaja con el lenguaje de programación llamado Python<sup>19</sup>; cabe hacer notar que ERP5.com ofrece información para aprender Python y utilizar zope.

---

<sup>18</sup> Es decir son servidores web, pero que tienen capacidad de almacenar y gestionar aplicaciones web. Entendiendo que una aplicación web es un servicio al que los usuarios acceden a través de la web.

<sup>19</sup> Python es un lenguaje de scripting independiente de plataforma y orientado a objetos, preparado para realizar cualquier tipo de programa, desde aplicaciones Windows a servidores de red o incluso, páginas web. Es un lenguaje interpretado, lo que significa que no se necesita compilar el código fuente

ERP5 es multiplataforma, es decir, se puede instalar en Windows y Linux por medio de la máquina virtual VMware Player, que es la instalación que se utilizará para esta tesina dado que solo se probará y no se requiere cambiar el código de ERP, también se puede realizar la instalación a través de SlapOS<sup>20</sup> (para las personas que desean contribuir al núcleo de ERP5), cabe resaltar que anteriormente ERP5 no utilizaba SlapOS sino TioLive<sup>21</sup>, por lo que los usuarios ya no se pueden suscribir a TioLive aunque podía hacer uso de él hasta finales del 2012, por lo que tendrán que emigrar a SlapOS.

La página web erp5.org dice que ERP5 abarca la contabilidad, la gestión de relaciones con clientes, comercio, gestión de almacenes, transporte, facturación, gestión de recursos humanos, diseño de producto, producción y gestión de proyectos.

En ERP5.com dice que el Modelo Unificado de Empresas ERP5 se utiliza para la contabilidad, gestión de almacenes, producción y gestión de recursos humanos en ERP5. Todas las características desarrolladas en un módulo están disponibles en los otros módulos debido al vocabulario unificado que es independiente de los procesos de negocio específicos.

El modelo de ERP5 es llamado así porque se basa en 5 puntos:

- Recursos: Detalla los recursos como un producto.

---

para poder ejecutarlo, lo que ofrece ventajas como la rapidez de desarrollo e inconvenientes como una menor velocidad.

<sup>20</sup> (Simple Language for Accounting and Provisioning Operating System) es una plataforma distribuida open source para automatizar el desarrollo, contabilidad y facturación de servicios cloud computing. De este modo, cualquiera puede vender servicios cloud (SaaS, PaaS o IaaS) sobre su infraestructura hardware de un modo sencillo y rápido.

Una nube SlapOS está compuesta de un maestro SlapOS y múltiples nodos. El maestro contiene un web service basado en REST que almacena la información de los nodos y asigna las tareas a ejecutar

<sup>21</sup> Es una plataforma de subcontratación total del sistema de información ([Total Information Outsourcing](#)). Se basa en tecnologías y [software de código abierto](#). TioLive ofrece a las empresas un conjunto de soluciones de aplicación disponibles en línea sobre Internet ([software como servicio](#)).

- Nodos: Recibe y manda recursos.
- Movimiento: Describe un movimiento sobre un recurso entre nodos.
- Path (rutas de acceso): Es el camino que un nodo tiene para acceder a un recurso.
- Items (tema): Es en elemento físico de un recurso.

En la página ERP5.com también dice “ERP5 es un completo ERP incluyendo la gestión de relaciones con clientes (CRM), gestión de la producción (MRP), gestión de la cadena de suministro (SCM), la gestión del diseño de producto (PDM), contabilidad, recursos humanos y el comercio electrónico.”

Una vez instalado ERP5 se confirma que cuenta con los siguientes módulos:

Base (Base): En este módulo se tiene información relativa a los empleados, a las organizaciones y se puede dar de alta los usuarios que tienen acceso a ERP5, así como sus cuentas.

CRM (CRM): En dicho módulo se tiene información que concierne tanto al cliente, como a los proveedores del ente social a través de campañas o tickets (que es como un informe donde se pueden ver todos los eventos relacionados a una actividad) y eventos (un evento es una acción, por ejemplo mandar un correo electrónico, una reunión, entre otros).

Comercio (*Trade*): Aquí se pueden ver información referente a inventarios, órdenes de compra, listas de empaque de compras, condiciones comerciales de compra, órdenes de venta, listas de empaque de ventas y condiciones comerciales de venta.

Gestión de datos de productos (*Product Data Management-PDM*): En éste se cuenta con información acerca del producto o servicio que se compra, se vende o se tiene en la organización.

Contabilidad (*Accounting*): Aquí se encuentra información de las transacciones de pagos, las transacciones de contabilidad, entre otros.

(DMS): En tal módulo se pueden crear diferentes documentos de texto, pdf, hoja de cálculo, y pagina web, entre otros.

Como se aprecia ERP5 es un buen sistema empresarial, cabe hacer notar que lo han utilizado tanto organizaciones de producción como un hospital y un banco, claro está haciéndole algunas modificaciones o agregándole los módulos requeridos.

Como se mencionó anteriormente ERP5 se basa en Zope, cuyo servidor de aplicaciones web trabaja con *Zope Object Data Base (ZODB)*, es decir, cuando se trabaja con zope se trabaja con objetos que se almacenan en ZODB, que es una base de datos de objetos transaccional, la cual se encarga de administrar los objetos (los cuales son persistentes, es decir, se mantienen guardados), los mantiene en memoria, los escribe y los lee. Además, ZODB proporciona atomicidad, consistencia, aislamiento, y durabilidad (propiedades ACID), lo que permite que ZODB sea eficiente.

Un punto importante de mencionar es que ZODB actúa como *front end* para una base de datos relacional específicamente MySQL/MariaDB (que actúa como *back end*) para ERP5. Para ilustrar esto ver la figura 68 en la que además de ZODB y MySQL se muestran otros componentes de ZOPE, los cuales no serán explicados dado que no son relevantes para la realización de esta tesina.

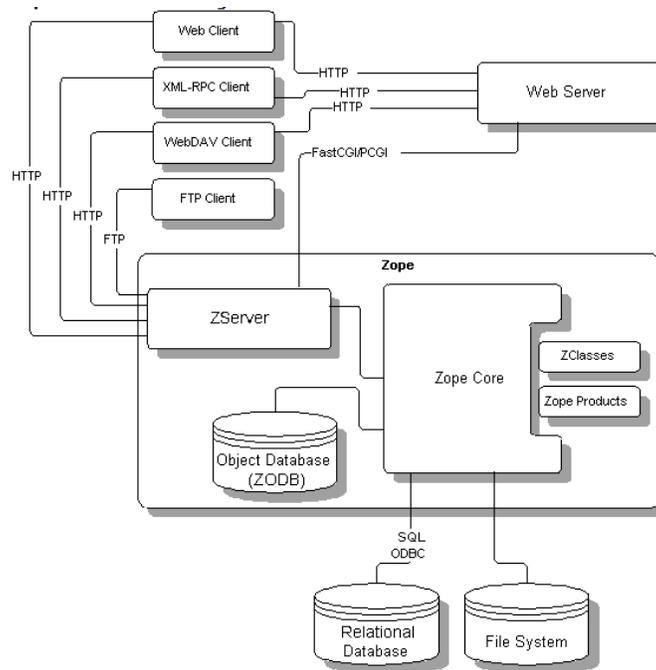


Figura 68. Zope Architecture Diagram\* Zope Components ;

<http://old.zope.org/WhatIsZope/ZopeArchitecture>

Cabe aclarar que ERP5 puede hacer uso de implementaciones del almacenamiento ZODB como *Zope Enterprise Object* (ZEO) que consta de un nodo servidor y múltiples nodos cliente o *Nexedi Enterprise Objects* (NEO) que a diferencia de ZEO trabaja con un clúster, ambas implementaciones tienen la finalidad de mejorar la escalabilidad de ERP5 según los requerimientos del usuario, es decir, tener la capacidad de adaptarse a circunstancias cambiantes sin disminuir la calidad. Con ello se pretende que ZODB incremente su eficiencia y por tanto que se pueda comunicar de forma más ágil con MySQL

Aunque ERP5 está basado en Zope, no usa el catálogo *Z catalog* para realizar búsqueda e indexación de documentos, ya que éste trabaja con la base de datos orientada a objetos de Zope, sin embargo, ERP5 hace uso del ZSQL Portal Catalog, cuyo catálogo es una herramienta que permite el acceso a la base de datos relacional, este ZSQL portal catalog contiene internamente

el adaptador de base de datos, en este caso el adaptador para MySQL es ZMySQLDA, cuyo componente es como tener una capa de abstracción con las características propias de la base de datos, también contiene los métodos ZSQL, los cuales permiten realizar una petición SQL y poder consultar la base de datos, de igual forma contiene *MySQL\_database\_connection* que es para hacer la conexión con la base de datos MySQL; claro está que todos estos componentes son transparentes para el usuario final, por lo tanto, el usuario final no sabe que está trabajando con todos ellos.

Otro punto relevante que es necesario mencionar es que ERP5 no usa un motor de persistencia<sup>22</sup> como tal, ya que un motor de persistencia puede usar algunas características propias de la orientación a objetos como el polimorfismo, pero en Zope no se maneja el polimorfismo (se refiere a acceder a múltiples funciones distintas mediante la misma interfaz). En cambio, en el caso de ERP5 los objetos son serializados, es decir, los objetos se transforman en una cadena de bytes para poder ser enviados por una red hacia el *backend* (MySQL/MariaDB según sea el caso), esta cadena de bytes se mantiene en memoria cuando los objetos son requeridos para leerse, pero en caso de que el objeto necesite ser modificado o escrito se guardan en el disco duro, es decir, en la base de datos, por lo tanto, la cadena de bytes se manda al backend y en éste la cadena de bytes se convierte en registros (actúa como un motor de persistencia) para poder ser almacenados en la base de datos. Cabe mencionar que la serialización o *marshelling* corresponde al proceso de modificar o escribir un objeto en la base de datos, y al proceso de leer el objeto se le llama deserialización, de igual forma, es relevante mencionar que los lenguajes de

---

<sup>22</sup> En el mundo de la programación este traductor no es más que un componente de software (concretamente, una capa de programación), al que se le dan los nombres de "capa de persistencia", "capa de datos", "correspondencia O/R" ("OR mapping") o "motor de persistencia".

El motor de persistencia traduce entre los dos formatos de datos: de registros a objetos y de objetos a registros.

programación como java, php y python permiten hacer la serialización a través de métodos propios de cada uno de ellos, en el caso de python se realiza comúnmente mediante el módulo *pickle*.

Por lo tanto, ERP5 hace uso de ZODB y MySQL, por lo que usa una base de datos centralizada, A continuación se explica por qué una base de datos distribuida sería una buena alternativa para ERP5.

### **3.1 Consultas**

Como ya se mencionó anteriormente las consultas son una petición que el usuario realiza al DBMS y son de suma importancia, dado que aportan información requerida por el usuario, así mismo en las bases de datos centralizadas las consultas son relativamente sencillas, ya que se accede a una única base de datos, es decir, la información se encuentra en un solo punto. Así que basta con conocer el nombre de la base de datos, de sus relaciones y la sintaxis de una consulta (ver apartado 2.5.6) para poder realizarla. En cambio, una consulta en una base de datos distribuida requiere mayor complejidad, pues además de conocer el nombre de la base de datos, de sus relaciones y la sintaxis de una consulta, se debe estar al tanto de los fragmentos de base de datos (fragmentación horizontal, vertical o mixta) y si es que éstos se encuentran replicados y en que emplazamientos.

Se utilizará el modulo base de ERP5, para mostrar el funcionamiento de éste con una base de datos centralizada y una distribuida, el módulo base consta principalmente de dos apartados el de personas y organización, como se aprecia cada uno tiene sus respectivas relaciones; a continuación se muestra la base de datos centralizada y la distribuida (nótese que la descripción de los atributos se encuentra en la figura 69 descripción de atributos y los nombres de las

relaciones se encuentran abreviados, para conocer el significado de cada relación ir al glosario que se encuentra al final de la investigación).

Atributo	Descripción
Id_per	Clave que identifica unívocamente a la persona
Titulo	Si la persona es señor, señora, señorita, entre otros.
Primer_nombre	Primer nombre de la persona
Segundo_nombre	Segundo nombre de la persona si es que lo tiene
Apellido	Apellido de la persona
Apellido-de_soltera	Apellido que ocupaba la mujer antes de casarse
Lugar_de_nacimiento	Zona donde nació la persona
Fecha_de_nacimiento	Fecha en la que nació la persona
Sexo	Genero de la persona
Foto	Imagen que muestra como es la persona
Direccion_P	Datos del lugar donde reside la persona
Cod_postal_P	Clave de la persona que identifica las zonas en un país
Ciudad_P	Ciudad donde vive la persona
Pais	País en el que se encuentra la persona
Email_P	Correo electrónico de la persona
Tel_P	Teléfono de la persona
Cel_P	Celular de la persona
Fax_P	Fax de la persona
Estado_civil	Situación en la que se encuentra la persona
Cabeza_de_familia	Si la persona es jefe de familia
Num_de_ninos	Cantidad de niños que tiene la persona
Grado	Si la persona es agente, gerente, entre otros
Num_de_socios	Cantidad de socios que tiene la persona
Sufijo	Nivel de puesto

Funcion_P	Puesto de la persona
Actividad_P	Las actividades que realiza la persona en la empresa
Num_de_empleado	Clave que tiene el empleado dentro de la empresa
Convenio_colectivo	Convenio donde las personas conocen sus derechos u obligaciones
Coeficiente_de_salario	Las variaciones que hay en los salarios para la actualización de los mismos
Nivel_salarial	Si la persona es gerente su nivel salarial probablemente será 2, es el tipo de salario que tiene la persona.
Num_de_identificacion_de_seguridad_social_P	Clave de la persona con la que está afiliada a un instituto de seguridad social
Descipcion_P	Detalles que se deseen poner sobre la persona
Id_rol_P	Clave que identifica al rol de la persona
Rol_P	Papel(es) que desempeña la persona
Id_hab_P	Clave que identifica de manera única las habilidades
Habilidades_P	Los talentos de la persona
Id_pro_int_P	Clave que identifica el producto de interés
Producto_de_Interes_P	Producto en que la persona está interesada
Id_Nac_Per	Clave que identifica la nacionalidad
Nacionalidad	Nacionalidad de la persona
Id_org	Clave que identifica de manera única a la organización
Nombre_oficial	Nombre de la organización con el que se le conoce ante una autoridad
Nombre_usual	Nombre que se le suele dar a la organización
Sitio	Si la empresa es la central (principal) o una sucursal (secundaria)
Grupo	Si es una empresa que pertenece a otra
Fecha_de_creacion	Fecha en la que la organización fue creada
Num_de_identificacion_de_seguridad_social_0	Clave de la empresa con la que está afiliada a una institución de salud.
Logotipo	Imagen que identifica a la empresa
Funcion_O	Cargo de la organización

Actividad_O	Actividades de la organización
Direccion_O	Datos del lugar donde se encuentra la empresa
Cod_postal_O	Clave de la organización que identifica las zonas en un país
Ciudad_O	Ciudad en la que está la empresa
Region	País en el que está localizada la empresa
Email_O	Correo electrónico de la empresa
Tel_O	Teléfono de la empresa
Fax_O	Fax de la empresa
Capital_registrado	Capital con el que cuenta la empresa oficialmente
Forma_juridica	Constitución de la empresa legalmente
Codigo_de_registro_empresarial	Es un instrumento de información de las empresas industriales, comerciales y de servicios que operan en el país, y sirve para identificar la oferta y demanda de productos y servicios de las empresas registradas
Codigo_de_incorporacion_geografica	Identifica una zona geográfica de forma exclusiva como puede ser el código de provincia, de municipio, de distrito, de sección censal o el código postal
Codigo_de_IVA	Valor del impuesto al valor agregado
Codigo_EAN_13	Son 13 dígitos que identifican a un producto (compuesto de dígitos que identifican al país, la empresa que lo fabrica, el núm. del producto y un dígito de control).
Descripcion_O	Detalles que se deseen agregar sobre la organización
Id_rol_O	Clave que identifica de manera unívoca el rol de la organización
Rol_O	Papel(es) que funge la organización
Id_hab_O	Clave que identifica unívocamente las habilidades con las que cuenta la empresa
Habilidades_O	Destrezas con las que cuenta la organización
Codigo_de_la_actividad	Clave de acuerdo al giro de la actividad de la empresa (a lo que se dedica la empresa).

Figura 69; Descripción de atributos

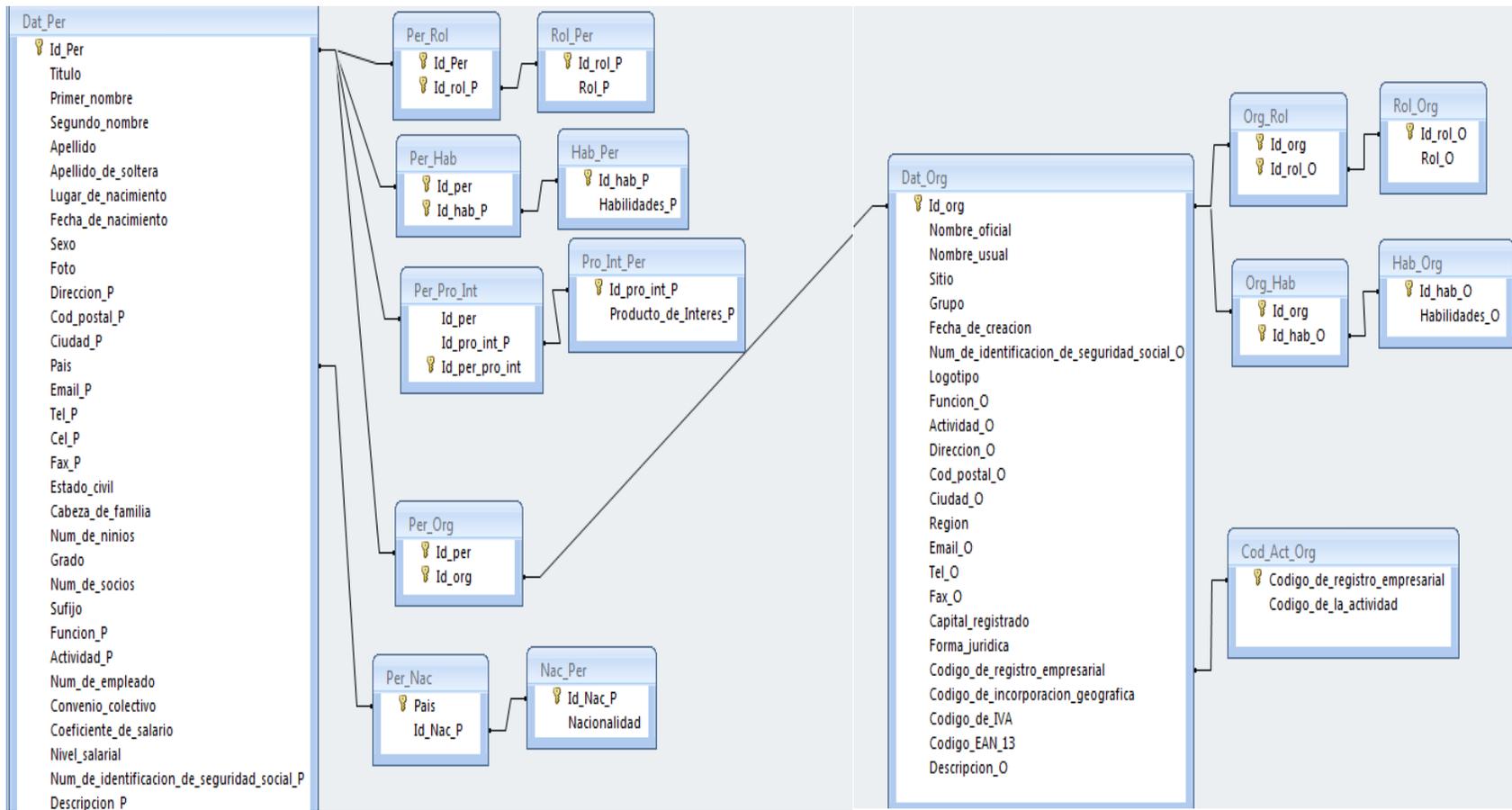


Figura 70; Base de datos centralizada de personas y organizaciones.

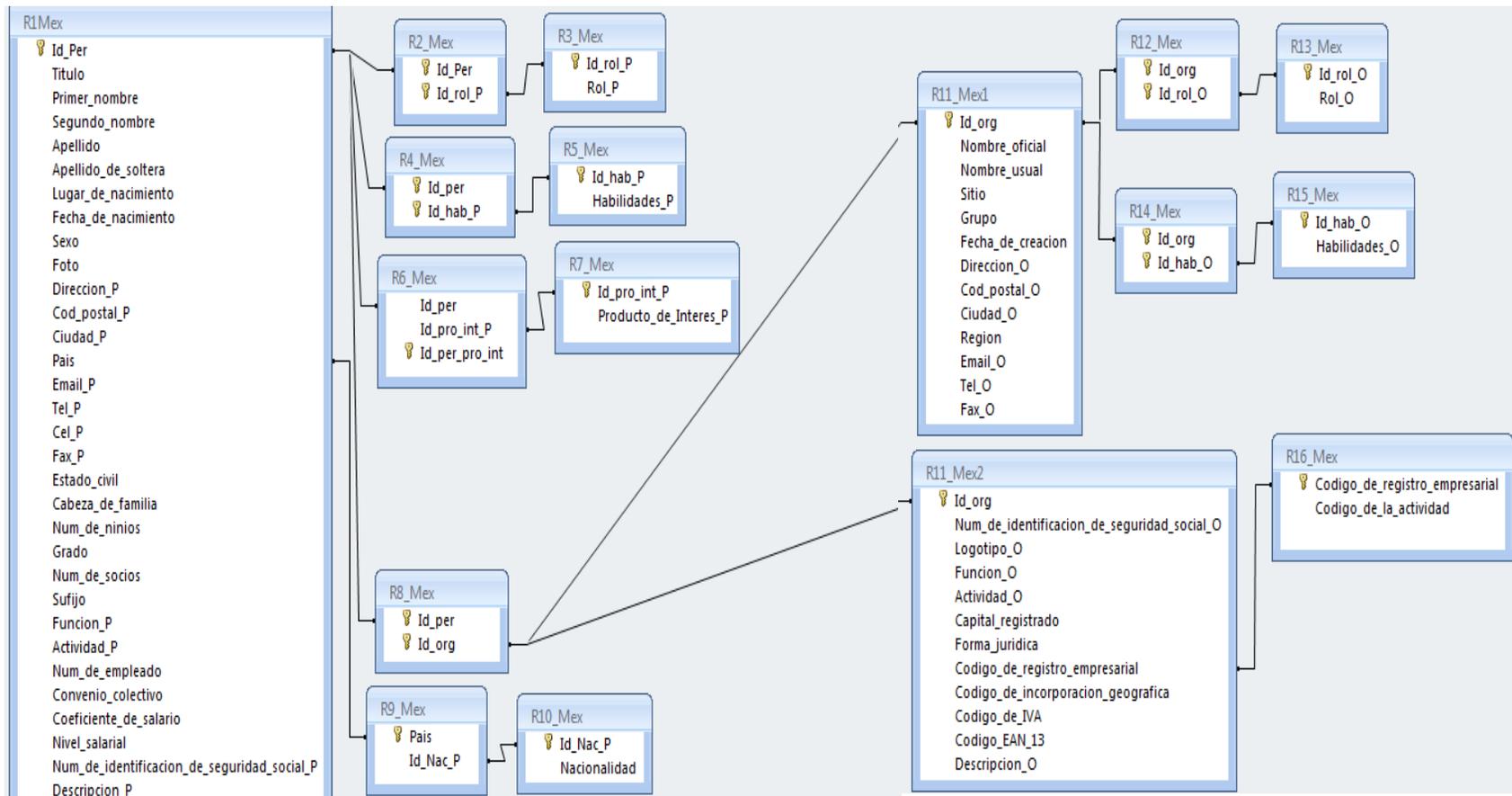


Figura 71; Base de datos distribuida de personas y organizaciones de México

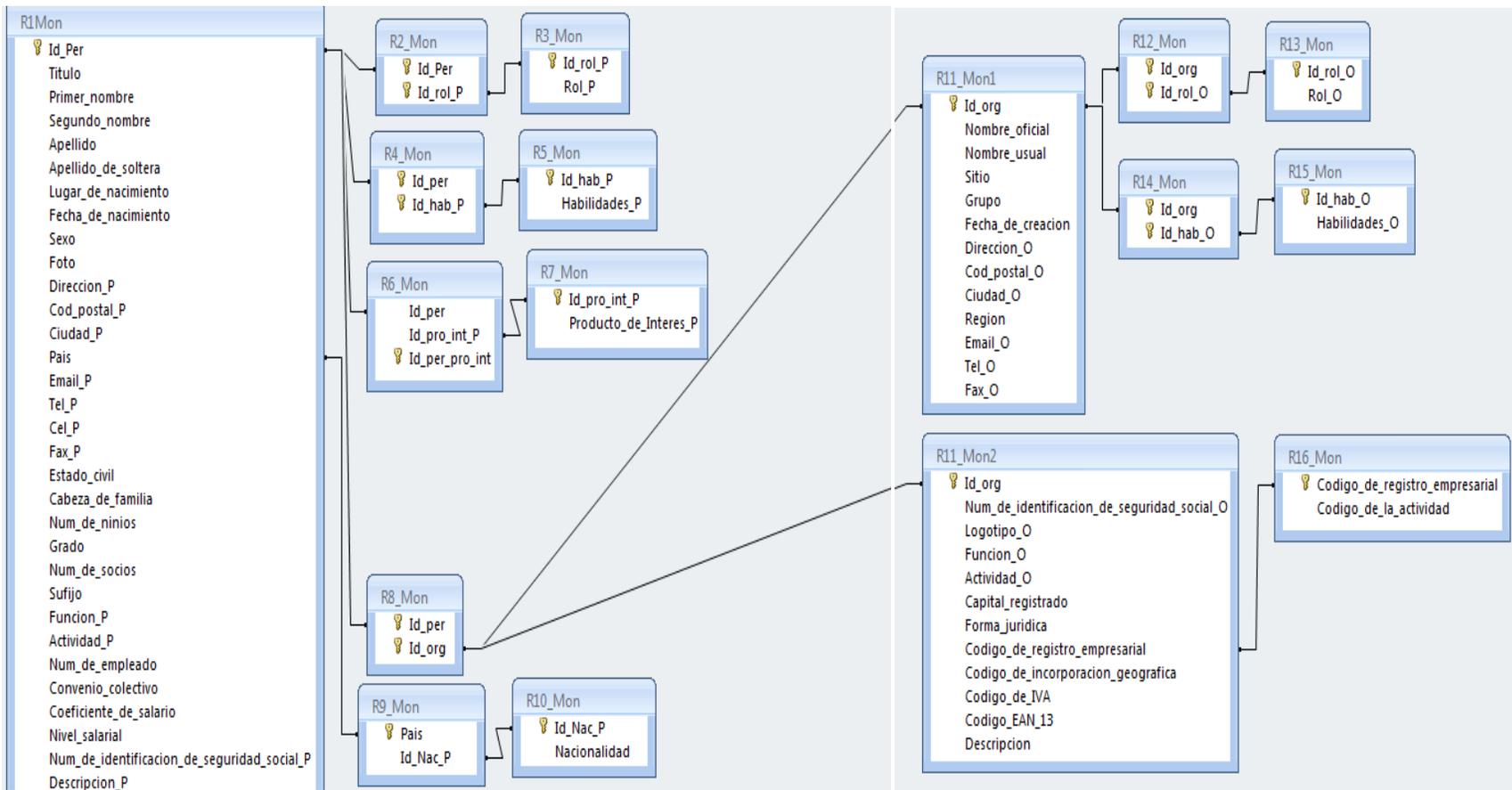


Figura 72; Base de datos distribuida de personas y organizaciones de Monterrey

Base de datos Centralizada

- Apartado de personas

Id_per	Titulo	Primer_nombre	Segundo_nombre	Apellido	Apellido_de_soltera	Lugar_de_nacimiento	Fecha_de_nacimiento	Sexo	Foto	Direccion_P	Cod_postal_P	Ciudad_P	Pais	Email_P	Tel_P	Cel_P
1	Señorita	Rocio	del Carmen	Cardenas		Coatlinchan	01/05/1980	Femenino		La purificacion num 12	56240	Mexico	Mexico	<a href="mailto:rocio@hotmail.com">rocio@hotmail.com</a>	224466	40248
2	Señor	Carlos		Sosa		Los reyes	15/20/1878	Masculino		Tepetlaotoc num54	56230	Mexico	Mexico	<a href="mailto:carlos@yahoo.com">carlos@yahoo.com</a>	40246	80646
3	Señora	Ana		Bracho	Sosa	Texcoco	30/03/1982	Femenino		Jarritos num12	56220	Monterrey	Mexico	<a href="mailto:ana@hotmail.com">ana@hotmail.com</a>	552030	402030
4	Señora	Ana	Paola	Rodriguez	Lopez	Tepexpan	20/08/1975	Femenino		La condesa s/n	56221	Monterrey	Mexico	<a href="mailto:paola@gmail.com">paola@gmail.com</a>	558899	402068

Fax_P	Estado_civil	Cabeza_de_familia	Num_de_ninos	Grado	Num_de_socios	Sufijo	Funcion_P	Actividad_P	Num_de_empleado_P	Convenio_colectivo	Coficiente_de_salario	Nivel_salarial	Num_de_identificacion_de_seguridad_social_P	Descripcion_P
	Cohabitando		0	Gerente	0		Gerente de contabilidad y finanzas	Intermediacion financiera excepto seguros/seguros y planes de pensiones	1	SYNTEC		Nivel 2	152040	Empleado del area de contabilidad
156	Soltero	√	0	Agente	0		Agente de la compañía	Administracion publica y defensa				Nivel1		Es una empleado de la empresa y a la vez un cliente de la misma
	Casado		2	Accionista	1		Agente de ventas	Inmobiliario, alquileres y actividades de negocios	2	SYNTEC		Nivel1	113333	Persona que trabaja en la empresa y ademas trabaja en otras empresas como medios de comunicacion, entre otros
112	Casado		1	Agente	0		Ejecutivo de la empresa	Transporte, almacenamiento y comunicaciones	3	SYNTEC		Nivel 3	203080	Empleado de la empresa que no trabaja en ninguna otra

Figura 73; Dat\_Per

Id_per	Id_rol_P
1	1
2	2
2	3
3	1
3	4
3	5
3	6
4	1

Figura 74; Per\_Rol.

Id_rol_P	Rol_P
1	Interno
2	Proveedor
3	Cliente
4	Administracion
5	Medios de comunicación
6	Perspectiva de ventas

Figura 75; Rol\_Per.

Id_per	Id_hab_P
1	1
2	2
3	2
3	3
3	4
4	5

Figura 76; Per\_Hab.

Id_hab_P	Habilidades_P
1	Contabilidad
2	Administración
3	Ventas
4	Mercadeo
5	Tecnologías de la información

Figura 77; Hab\_Per

Id_per_pro_int	Id_per	Id_pro_int_P
1	1	0
2	2	1
3	2	2
4	3	3
5	4	4

Figura 78; Per\_Pro\_Int.

Id_pro_int_P	Producto_de_interes_P
0	
1	Equipos industriales
2	Electrodomesticos
3	Productos publicados
4	Componentes y suministros de equipos

Figura 79; Pro\_Int\_Per.

Pais	Id_Nac_P
Mexico	1

Figura 80; Per\_Nac.

Id_Nac_P	Nacionalidad
1	Mexicano

Figura 81; Nac\_Per.

Id_per	Id_org
1	1
1	2
2	2
3	3
4	3

Figura 82; Per\_Org

- Apartado de organizaciones

Id_org	Nombre_oficial	Nombre_usual	Sitio	Grupo	Fecha_de_creacion	Num_de_identificacion_de_pseudonimo	Localidad	Funcion_O	Actividad_O	Direccion_O	Cod_postal_O	Ciudad_O	Region	Email_O	Tel_O	Fax_O	Capital_requerida	Forma_juridica	Codigo_de_registro_empresarial	Codigo_de_inscripcion_geografica	Codigo_de_IVA	Codigo EAM-13	Descripcion_O
1	Usom texcoco	Usom texcoco	Principal	Usom	22/01/1975	1357		Ventar	Manufactura textil	Loyar num 15	56229	Mexico	Mexico	usom@hotmail.com	24545	514	800,000	Sociedad de responsabilidad limitada	567		25	5.67349E+11	Realiza ventar de manufactura textil
2	Chapingo	Chapingo		Chapingo	19/07/1950	990		Producción de fabricación	Muebler	Santiago num 54	55440	Mexico	Mexico	shapingo@amajlsam	33112	620	200,000	Consumidor de cooperacion	606				
3	Upem san andres	Upem san andres		Upem	22/12/1965	750		Mercadotecnia	Publicacion e impresión de reportes grabador	Bonita juarez	55470	Manterrey	Mexico	usom@yahoo.com	77001	880		Propietaria unica	505				

Figura 83; Dat\_Org

Id_org	Id_rol_O
1	1
2	2
3	2
3	3

Figura 84; Org\_Rol.

Id_rol_O	Rol_O
1	Interno
2	Administracion
3	Medios de comunicación

Figura 85; Rol\_Org.

Id_org	Id_hab_O
1	1
2	2
3	3
3	4

Figura 86; Org\_Hab.

Id_hab_O	Habilidades_O
1	Ventas
2	Administración de proyecto
3	Administración por producción
4	Mercadotecnia

Figura 87; Hab\_Org.

Codigo_de_registro_empresarial	Codigo_de_la_actividad
567	412
606	213
505	111

Figura 88; Cod\_Act\_Org.

Base de datos distribuida

- Apartado de personas

Supóngase que la empresa “Z” usa ERP5 en dos sitios diferentes, es decir, en Monterrey y en México, en cada uno de ellos se requiere conocer los datos respectivos (*Titulo, Primer\_Nombre, Segundo\_nombre, Apellido, Apellido\_de\_soltera, Lugar\_de\_nacimiento, Fecha\_de\_nacimiento, Sexo, Foto, Direccion\_P, Cod\_Postal\_P, Ciudad\_P, Pais, Email\_P, Tel\_P, Cel\_P, Fax\_P, Estado\_civil, Cabeza\_de\_familia, Num\_de\_ninos, Grado, Num\_de\_socios, Sufijo, Funcion\_P, Actividad\_P, Num\_de\_empleado\_P, Convenio\_colectivo, Coeficiente\_de\_salario, Nivel\_salarial, Num\_de\_identificacion\_de\_seguridad\_social\_P, Descripcion\_P*) de las personas involucradas a la organización de cada zona, por tanto, se aplica una fragmentación horizontal a la Figura 73, de la cual salen los fragmentos que se muestran a continuación.

Id_per	Titulo	Primer_nombre	Segundo_nombre	Apellido	Apellido_de_soltera	Lugar_de_nacimiento	Fecha_de_nacimiento	Sexo	Foto	Direccion_P	Cod_postal_P	Ciudad_P	Pais	Email_P	Tel_P	Cel_P
1	Señorita	Rocio	del Carmen	Cardenas		Coatlinchan	01/05/1980	Femenino		La purificacion num 12	56240	Mexico	Mexico	<a href="mailto:rocio@hotmail.com">rocio@hotmail.com</a>	224466	40248
2	Señor	Carlos		Sosa		Los reyes	15/20/1878	Masculino		Tepetlaotoc num54	56230	Mexico	Mexico	<a href="mailto:carlos@yahoo.com">carlos@yahoo.com</a>	40246	80646

Fax_P	Estado_civil	Cabeza_de_familia	Num_de_ninos	Grado	Num_de_socios	Sufijo	Funcion_P	Actividad_P	Num_de_empleado_P	Convenio_colectivo	Coeficiente_de_salario	Nivel_salarial	Num_de_identificacion_de_seguridad_social_P	Descripcion_P
	Cohabitando		0	Gerente	0		Gerente de contabilidad y finanzas	Intermediacion financiera excepto seguros/seguros y planes de pensiones	1	SYNTEC		Nivel 2	152040	Empleado del area de contabilidad
156	Soltero	√	0	Agente	0		Agente de la compañía	Administracion publica y defensa				Nivel1		Es una empleado de la empresa y a la vez un cliente de la misma

Figura 89; Dat\_Per de México (R1Mex)

Id_per	Titulo	Primer_nombre	Segundo_nombre	Apellido	Apellido_de_soltera	Lugar_de_nacimiento	Fecha_de_nacimiento	Sexo	Foto	Direccion_P	Cod_postal_P	Ciudad_P	Pais	Email_P	Tel_P	Cel_P
3	Señora	Ana		Bracho	Sosa	Texcoco	30/03/1982	Femenino		Jarritos num12	56220	Monterrey	Mexico	<a href="mailto:ana@hotmail.com">ana@hotmail.com</a>	552030	402030
4	Señora	Ana	Paola	Rodriguez	Lopez	Tepexpan	20/08/1975	Femenino		La condessa s/n	56221	Monterrey	Mexico	<a href="mailto:paola@gmail.com">paola@gmail.com</a>	558899	402068

Fax_P	Estado_civil	Cabeza_de_familia	Num_de_ninos	Grado	Num_de_socios	Sufijo	Funcion_P	Actividad_P	Num_de_empleado_P	Convenio_colectivo	Coficiente_de_salario	Nivel_salarial	Num_de_identificacion_de_seguridad_social_P	Descripcion_P
	Casado		2	Accionista	1		Agente de ventas	Inmobiliario, alquileres y actividades de negocios	2	SYNTEC		Nivel1	113333	Persona que trabaja en la empresa y ademas trabaja en otras empresas como medios de comunicacion, entre otros
112	Casado		1	Agente	0		Ejecutivo de la empresa	Transporte, almacenamiento y comunicaciones	3	SYNTEC		Nivel 3	203080	Empleado de la empresa que no trabaja en ninguna otra

Figura 90; Dat\_Per de Monterrey (R1Mon)

Además, en cada sitio se requiere conocer el rol (*Rol\_P*), las habilidades (*Habilidades\_P*), el producto de interés (*Producto\_de\_interes\_P*), la organización en la que laboran (*Nombre\_oficial*), y la nacionalidad (*Nacionalidad*) del personal asociado a la empresa que usa ERP5.

Id_per	Id_rol_P
1	1
2	2
2	3

Figura 91; Per\_Rol de México (R2Mex).

Id_per	Id_rol_P
3	1
3	4
3	5
3	6
4	1

Figura 92; Per\_Rol de Monterrey (R2Mon).

Id_rol_P	Rol_P
1	Interno
2	Proveedor
3	Ciente
4	Administracion
5	Medios de comunicaci3n
6	Perspectiva de ventas

Figura 93; Rol\_Per de México (R3Mex).

Id_rol_P	Rol_P
1	Interno
2	Proveedor
3	Ciente
4	Administracion
5	Medios de comunicaci3n
6	Perspectiva de ventas

Figura 94; Replica Rol\_Per de México en Monterrey (R3Mon)

Id_per	Id_hab_P
1	1
2	2

Figura 95; Per\_Hab de México (R4Mex).

Id_per	Id_hab_P
3	2
3	3
3	4
4	5

Figura 96; Per\_Hab de Monterrey (R4Mon).

Id_hab_P	Habilidades_P
1	Contabilidad
2	Administración
3	Ventas
4	Mercadeo
5	Tecnologías de la informacion

Figura 97; Hab\_Per de México (R5Mex).

Id_hab_P	Habilidades_P
1	Contabilidad
2	Administración
3	Ventas
4	Mercadeo
5	Tecnologías de la informacion

Figura 98; Replica Hab\_Per de México en Monterrey (R5Mon).

Id_per pro_int	Id_per	Id_pro_int _P
1	1	0
2	2	1
3	2	2

Figura 99; Per\_Pro\_Int de México (R6Mex).

Id_per pro_int	Id_per	Id_pro_int _P
4	3	3
5	4	4

Figura 100; Per\_Pro\_Int de Monterrey (R6Mon).

Id_pro_int _P	Producto_de_interes_P
0	
1	Equipos industriales
2	Electrodomesticos
3	Productos publicados
4	Componentes y suministros de equipos

Figura 101; Pro\_Int\_Per de México (R7Mex).

Id_pro_int _P	Producto_de_interes_P
0	
1	Equipos industriales
2	Electrodomesticos
3	Productos publicados
4	Componentes y suministros de equipos

Figura 102; Replica de Pro\_Int\_Per de México en Monterrey (R7Mon)

Id_per	Id_org
1	1
1	2
2	2

Figura 103; Per\_Org de México (R8Mex).

Id_per	Id_org
3	3
4	3

Figura 104; Per\_Org de Monterrey (R8Mon).

Pais	Id_Nac_P
Mexico	1

Figura 105; Per\_Nac de México (R9Mex).

Pais	Id_Nac_P
Mexico	1

Figura 106; Replica de Per\_Nac de México en Monterrey (R9Mon).

Id_Nac_P	Nacionalidad
1	Mexicano

Figura 107; Nac\_Per de México (R10Mex).

Id_Nac_P	Nacionalidad
1	Mexicano

Figura 108; Replica de Nac\_Per de México en Monterrey (R10Mon).

- Apartado de Organizaciones

Ahora en México por lo general se consulta el *Nombre\_oficial*, *Nombre usual*, *Sitio*, *Grupo*, *Fecha\_de\_creacion*, *Direccion\_O*, *Cod\_postal\_O*, *Ciudad\_O*, *Region*, *Email\_O*, *Tel\_O* y el *Fax* de una Organización, y en Monterrey se suele recurrir mas al *Num\_de\_identificacion\_de\_seguridad\_social\_O*, *Logotipo*, *Funcion\_O*, *Actividad\_O*, *Capital\_registrado*, *Forma\_juridica*, *Codigo\_de\_registro\_empresarial*, *Codigo\_de\_incorporacion\_geografica*, *Codigo\_de\_IVA*, *Codigo\_EAN\_13* y *Descripcion\_O*; de igual forma, en cada sitio (México y Monterrey) se requiere saber toda la información perteneciente a cada lugar (de la relación Datos\_O), es decir se aplica una fragmentación mixta VH) a la figura 83, dando como resultado los siguientes fragmentos.

Id_org	Nombre_oficial	Nombre_usual	Sitio	Grupo	Fecha_de_creacion	Direccion_O	Cod_postal_O	Ciudad_O	Region	Email_O	Tel_O	Fax_O
1	Uaem texcoco	Uaem texcoco	Principal	Uaem	22/01/1975	Leyes num 15	56229	Mexico	Mexico	<a href="mailto:uaem@hotmail.com">uaem@hotmail.com</a>	24545	514
2	Chapingo	Chapingo	Chapingo	Chapingo	19/07/1950	Santiago num 54	55440	Mexico	Mexico	<a href="mailto:chapingo@email.com">chapingo@email.com</a>	33112	620

Figura 109; Dat\_Org 1 de Mexico (R11Mex1)

Id_org	Num_de_identificacion_de_seguridad_social_O	Logotipo	Funcion_O	Actividad_O	Capital_registrado	Forma_juridica	Codigo_de_registro_empresarial	Codigo_de_incorporacion_geografica	Codigo_de_IVA	Codigo_EAN_13	Descripcion_O
1	1357		Ventas	Manufactura textil	800,000	Sociedad de responsabilidad limitada	567		25	567348901234	Realiza ventas de manufactura textil
2	990		Producción de fabricación	Muebles	200,000	Consumidores de cooperacion	606				

Figura110; Dat\_Org 2 de Mexico (R11Mex2)

Id_org	Nombre_oficial	Nombre_usual	Sitio	Grupo	Fecha_de_creacion	Direccion_O	Cod_postal_O	Ciudad_O	Region	Email_O	Tel_O	Fax_O
3	Upem san andres	Upem san Andri		Upem	22/12/1965	Benito juarez	55470	Monterrey	Mexico	<a href="mailto:upem@yahoo.com">upem@yahoo.com</a>	77001	880

Figura 111; Dat\_Org 1 de Monterrey (R11Mon1)

Id_org	Num_de_identificacion_de_seguridad_social_O	Logotipo	Funcion_O	Actividad_O	Capital_registrado	Forma_juridica	Codigo_de_registro_empresarial	Codigo_de_incorporacion_geografica	Codigo_de_IVA	Codigo_EAN_13	Descripcion_O
3	750		Mercadotecnia	Publicaciones e impresión de soportes grabados		Propietario unico	505				

Figura 112; Dat\_Org 2 de Monterrey (R11Mon2)

De igual forma, en México y Monterrey se necesitan conocer los roles (*Rol\_O*), las habilidades (*Habilidades\_O*), y el código de la actividad (*Codigo\_de\_la\_actividad*) de las empresas que se encuentran en dichos lugares, así que se implementa una fragmentacion horizontal para las figuras 84, 85, 86, 87 y 88(con el mismo criterio), dado que en éstas no se puede aplicar una fragmentacion vertical, por lo que quedan los siguientes fragmentos.

Id_org	Id_rol_O
1	1
2	2

Figura 113; Org\_Rol de México (R12Mex).

Id_org	Id_rol_O
3	2
3	3

Figura 114; Org\_Rol de Monterrey (R12Mon).

Id_rol_O	Rol_O
1	Interno
2	Administracion
3	Medios de comunicaci3n

Figura 115; Rol\_Org de M3xico (R13Mex).

Id_rol_O	Rol_O
1	Interno
2	Administracion
3	Medios de comunicaci3n

Figura 116; Replica de Rol\_Org de M3xico en Monterrey (R13Mon).

Id_org	Id_hab_O
1	1
2	2

Figura 117;Org\_Hab de M3xico (R14Mex).

Id_org	Id_hab_O
3	3
3	4

Figura 118;Org\_Hab de Monterrey (R14Mon).

Id_hab_O	Habilidades_O
1	Ventas
2	Administraci3n de proyecto
3	Administraci3n por producci3n
4	Mercadotecnia

Figura 119; Hab\_Org de M3xico (R15Mex).

Id_hab_O	Habilidades_O
1	Ventas
2	Administraci3n de proyecto
3	Administraci3n por producci3n
4	Mercadotecnia

Figura 120; Replica de Hab\_Org de M3xico en Monterrey (R15Mon).

Codigo_de_registro_empresarial	Codigo_de_la_actividad
567	412
606	213

Figura 121; Cod\_Act\_Org de M3xico (R16Mex).

Codigo_de_registro_empresarial	Codigo_de_la_actividad
505	111

Figura 122; Cod\_Act\_Org de Monterrey (R16Mon).

Ahora se realizar3n las mismas consultas tanto en la base de datos centralizada como en la distribuida para ver la diferencia entre ambas.

1) En Monterrey se necesitan los datos de las personas como el *Id\_per*, *Titulo*, *Primer\_nombre*, *Segundo\_nombre*, *Apellido*, *Apellido\_de\_soltera*, *Lugar\_de\_nacimiento*, *Fecha de nacimiento*, *Sexo*, *Foto*, *Direccion\_P*, *Cod\_postal\_P*, *Ciudad\_P*, *Pais*, *Email\_P*, *Tel\_P*, *Cel\_P*, *Fax\_P*, *Estado\_civil*, *Cabeza\_de\_familia*, *Num\_de\_ninos*, *Grado*, *Num\_de\_socios*, *Sufijo*, *Funcion\_P*, *Actividad\_P*, *Num\_de\_empleado\_P*, *Convenio\_colectivo*, *Coficiente\_de\_salario*, *Nivel\_salarial*, *Num\_de\_identificacion\_de\_seguridad\_social\_P* y la *Descripcion\_P*.

Base de datos centralizada (tomando la figura 73 del apartado 3.1)

```
SELECT * FROM Dat_Per where Ciudad_P='monterrey';
```

Base de datos distribuida(tomando la figura 90 del apartado 3.1)

```
SELECT * FROM R1Mon;
```

A continuación se utiliza el modelo de costo simple descrito por Ozsu y Valduriez para demostrar el beneficio de las bases de datos distribuidas contra las centralizadas, cabe mencionar que este modelo se usará para todas las consultas.en el cual se asume que el acceso a una tupla *tupacc* es 1 unidad y la transferencia a una tupla, es decir, *tuptrans* es de 10 unidades,

Base de datos centralizada

Acceso a Dat_Per	4*tupacc= 4
	———
El costo total es	4

Base de datos distribuida

Acceso a R1Mon  $2 * \text{tupacc} = 2$   
 El costo total es  $2$

Como se aprecia en la consulta de la base de datos centralizada el acceso es mayor que en la distribuida, además de que se tiene que especificar la ciudad (Monterrey), ello implica que la búsqueda para poder mostrar los registros que cumplan con dicho criterio se realice en todas las tuplas de la relación, haciendo mas lenta la respuesta; ésto es algo que no sucede en la base de datos distribuida pues unicamente se recurre al fragmento de Monterrey y se muestra el resultado, consiguiendo así optimizar el tiempo de respuesta, es decir, aumentar el rendimiento de la base de datos.

Id_per	Título	Primer_nombre	Segundo_nombre	Apellido	Lugar_de_nacimiento	Fecha_de_nacimiento	Sexo	Direccion_P	Cod_postal_P	Ciudad_P
3	señora	ana		bracho	texcoco	30-03-1982	femenino	jaritos num 12	56220	monterrey
4	señora	ana	paola	rodriguez	tepexpan	20-08-1975	femenino	la condesa s/n	56221	monterrey

Pais	Email_P	Tel_P	Cel_P	Fax_P	Estado_civil	Cabeza...	Num_de_ninios	Grado	Num_de_socios	Sufijo	Funcion_P
mexico	ana@hotmail.com	552030	402030	0	casado		2	accionista	1		agente de ventas
mexico	paola@gmail.com	558899	4020698	112	casado		1	agente	0		ejecutivo de la empresa

Actividad_P	Num_de_empleado	Convenio_colectivo	Coficiente_de_salario	Nivel_salarial	Numero_de_identificacion_de_seguridad_social	Descripcion_P
inmobiliario, alquileres y actividades de negocio	2	SYNTEC		nivel1	113333	persona que trabaja e
transporte, almacenamiento y comunicaciones	3	SYNTEC		nivel3	203080	empleado de la empre

Figura 123; Resultado de la consulta 1 de la base de datos centralizada

- En México se requiere conocer el *Título*, *Primer\_nombre*, *Segundo\_nombre*, *Apellido*, *Apellido\_de\_soltera*, *Sexo*, *Direccion\_P*, *Cod\_postal\_P*, *Email\_P* y *Tel\_P* (relación *Dat\_Per*); también se necesita conocer el *Rol\_P*, las *Habilidades\_P*, y la organización (*Nombre\_oficial*) en la que trabajan las personas

Base de datos centralizada (Véase figuras 73, 74, 75, 76, 77, 82 y 83 del apartado 3.1)

```
SELECT Titulo, Primer_nombre, Segundo_nombre, Apellido,
Apellido_de_soltera, Sexo, Direccion_P, Cod_postal_P, Email_P, Tel_P,
Rol_P, Habilidades_P, Nombre_oficial
FROM Dat_Per, Per_Rol, Rol_Per, Per_Hab, Hab_Per, Per_Org, Dat_Org
WHERE ((Dat_Per.Id_per=Per_Rol.Id_per and
Per_Rol.Id_rol_P=Rol_Per.Id_rol_P) and (Dat_Per.Id_per=Per_Hab.Id_per and
Per_Hab.Id_hab_P=Hab_Per.Id_hab_P)and (Dat_Per.Id_per=Per_Org.Id_per and
Per_Org.Id_org=Dat_Org.Id_org) and Ciudad_P='Mexico');
```

Titulo	Primer_nombre	Segundo_nombre	Apellido	Apellido_de_soltera	Sexo	Direccion_P	Cod_postal_P	Email_P	Tel_P	Rol_P	Habilidades_P	Nombre_oficial
Señorita	Rocio	del carmen	Cardenas		femenino	La purificacion num	56240	rocio@hotmail.com	224466	Interno	Contabilidad	Usam texcoco
Señorita	Rocio	del carmen	Cardenas		femenino	La purificacion num	56240	rocio@hotmail.com	224466	Interno	Contabilidad	Chapingo
Señor	Carlos		Sosa		masculino	Tepetlaotoc num 54	56230	carlos@yahoo.com	40246	Proveedor	Administracion	Chapingo
Señor	Carlos		Sosa		masculino	Tepetlaotoc num 54	56230	carlos@yahoo.com	40246	Cliente	Administracion	Chapingo

Figura 124; Resultado de la consulta 2 de la base de datos centralizada

Base de datos distribuida (tomando las figuras 89, 91, 93, 95, 97, 103, 109 del apartado 3.1)

```
SELECT Titulo, Primer_nombre, Segundo_nombre, Apellido,
Apellido_de_soltera Sexo, Direccion_P, Cod_postal_P, Email_P, Tel_P,
Rol_P, Habilidades_P, Nombre_oficial
FROM R1Mex, R2Mex, R3Mex, R4Mex, R5Mex, R8Mex, R11Mex1
WHERE ((R1Mex.Id_per=R2Mex.Id_per and R2Mex.Id_rol_P=R3Mex.Id_rol_P) and
(R1Mex.Id_per=R4Mex.Id_per and R4Mex.Id_hab_P=R5Mex.Id_hab_P) and
(R1Mex.Id_per=R8Mex.Id_per and R8Mex.Id_org=R11Mex1.Id_org));
```

A continuación se presenta el modelo de costo para estas consultas.

### Base de datos centralizada

Acceso a Dat_Per (4 tuplas) y Per_Rol (8 tuplas)	32*tupacc= 32
Acceso a Per_Rol (8 tuplas) y Rol_Per (6 tuplas)	48*tupacc=48
Acceso a Dat_Per (4 tuplas)y Per_Hab (6 tuplas)	24*tupacc=24
Acceso a Per_Hab (6 tuplas) y Hab_Per (5 tuplas)	30*tupacc=30
Acceso a Dat_Per (4 tuplas) y Per_Org (5 tuplas)	20*tupacc=20
Acceso a Per_Org (5 tuplas) y Dat_Org (3 tuplas)	15*tupacc=15
	<hr/>
El costo total es	169

### Base de datos distribuida

Acceso a R1Mex (2 tuplas) y R2Mex (3 tuplas)	6*tupacc= 6
Acceso a R2Mex (3 tuplas) y R3Mex (6 tuplas)	18*tupacc=18
Acceso a R1Mex (2 tuplas)y R4Mex (2 tuplas)	4*tupacc= 4
Acceso a R4Mex (2 tuplas) y R5Mex (5 tuplas)	10*tupacc=10
Acceso a R1Mex (2 tuplas) y R8Mex (3 tuplas)	6*tupacc= 6
Acceso a R8Mex (3 tuplas) y R11Mex1 (2 tuplas)	6*tupacc= 6
	<hr/>
El costo total es	50

Como se puede ver el costo total de la base de datos centralizada es mayor que la distribuida, así mismo se puede notar que en el criterio de selección se necesitan los datos donde la ciudad sea

México, pero se solicitan datos no solamente de una relación, sino de varias, por lo que en una base de datos centralizada al hacer una reunión, también llamada combinación, unión o junta interna, en ocasiones llamada por Coronel Morris y Rob como “antigua ”(realizada en la cláusula where), consume mas recursos(hablando de memoria y tiempo) que en una distribuida, ya que además de que se necesita acceder a varias tablas también se tiene que acceder a todos los registros para realizar la búsqueda de dicho criterio.

3) En México se necesita saber el *Nombre\_oficial*, *Nombre\_usual*, *Sitio*, *Grupo*, *Fecha\_de\_creacion*, *Direccion\_O* , *Cod\_postal\_O*, *Ciudad\_O*, *Region*, *Email\_O*, *Tel\_O* y el *Fax\_O* de las empresas pertenecientes a esta ciudad.

Base de datos centralizada (Ver figura 83 del apartado 3.1)

```
Select Nombre_oficial, Nombre_usual, Sitio, Grupo, Fecha_de_creacion,
Direccion_O, Cod_postal_O, Ciudad_O, Region, Email_O, Tel_O, Fax_O

From Dat_Org

Where Ciudad_O='Mexico';
```

Nombre_oficial	Nombre_usual	Sitio	Grupo	Fecha_de_creacion	Direccion_O	Cod_postal_O	Ciudad_O	Region	Email_O	Tel_O	Fax_O
Uaem texcoco	Uaem texcoco	Principal	Uaem	22-01-1975	Leyes num 15	56229	Mexico	Mexico	uaem@hotmail.com	24545	514
Chapingo	Chapingo	NULL	Chapingo	19-07-1950	Santiago num 54	55440	Mexico	Mexico	chapingo@gmail.com	33112	620

Figura 125; Resultado de la consulta 3 de la base de datos centralizada

Base de datos distribuida (Vease figura 109 del apartado 3.1)

```
Select Nombre_oficial, Nombre_usual, Sitio, Grupo, Fecha_de_creacion,
Direccion_O, Cod_postal_O, Ciudad_O, Region, Email_O, Tel_O, Fax_O

From R11Mex1;
```

Modelo de costo para estas consultas

Base de datos centralizada

Acceso a Dat_Org (3 tuplas)	3*tupacc= 3
	——
El costo total es	3

Base de datos distribuida

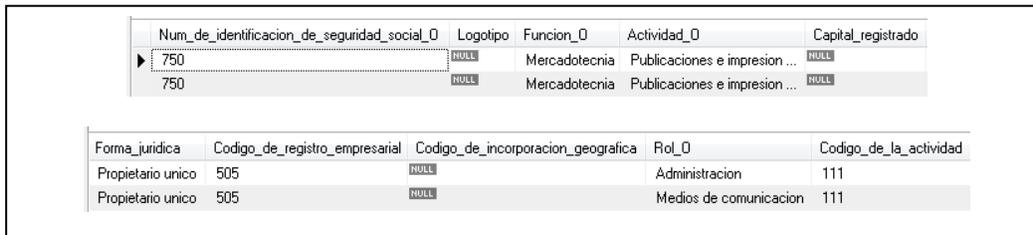
Acceso a R11Mex1 (2 tuplas)	2*tupacc= 2
	——
El costo total es	2

Se puede apreciar que el costo total de realizar la consulta es más grande en la centralizada que en la distribuida, de igual forma es visible que la consulta es casi la misma, salvo que en la base de datos centralizada se tiene que usar un criterio de búsqueda y en la distribuida no porque el fragmento sólo contiene datos de México, además si se observan las figuras se puede notar que la relación Dat\_Org es grande a comparación de la relación Dat\_Org de México (R11Mex1), esto significa que la consulta es más ágil en la distribuida que en la centralizada, ya que es más eficiente acceder a una relación con pocos atributos y tuplas que a una relación con muchos atributos y tuplas.

- 4) En Monterrey se necesita conocer el *Num\_de\_identificacion\_de\_seguridad\_social\_O*, *Logotipo*, la *Funcion\_O*, *Actividad\_O*, *Capital\_registrado*, *Forma\_Juridica*, el *Codigo\_de\_registro\_empresarial*, el *Codigo\_de\_incorporacion\_geografica*, el *Rol\_O*, y el *Codigo\_de\_la\_actividad*, de las organizaciones que se encuentran en Monterrey.

Base de datos centralizada (Ver las figuras 83, 84, 85, 88 del apartado 3.1)

```
SELECT Dat_Org.Num_de_identificacion_de_seguridad_social_O,  
Dat_Org.Logotipo, Dat_Org.Funcion_O, Dat_Org.Actividad_O,  
Dat_Org.Capital_registrado, Dat_Org.Forma_juridica,  
Dat_Org.Codigo_de_registro_empresarial,  
Dat_Org.Codigo_de_incorporacion_geografica, Rol_Org.Rol_O,  
Cod_Act_Org.Codigo_de_la_actividad  
  
FROM Dat_Org, Org_Rol, Rol_Org,Cod_Act_Org  
  
WHERE (Dat_Org.Id_org=Org_Rol.Id_org and  
Org_Rol.Id_rol_O=Rol_Org.Id_rol_O) and  
(Dat_org.Codigo_de_registro_empresarial=Cod_Act_Org.Codigo_de_registro_emp  
resarial) and Ciudad_O='Monterrey';
```



The screenshot displays two tables of data. The top table has columns: Num\_de\_identificacion\_de\_seguridad\_social\_O, Logotipo, Funcion\_O, Actividad\_O, and Capital\_registrado. The bottom table has columns: Forma\_juridica, Codigo\_de\_registro\_empresarial, Codigo\_de\_incorporacion\_geografica, Rol\_O, and Codigo\_de\_la\_actividad.

Num_de_identificacion_de_seguridad_social_O	Logotipo	Funcion_O	Actividad_O	Capital_registrado
750	NULL	Mercadotecnia	Publicaciones e impresio ...	NULL
750	NULL	Mercadotecnia	Publicaciones e impresio ...	NULL

Forma_juridica	Codigo_de_registro_empresarial	Codigo_de_incorporacion_geografica	Rol_O	Codigo_de_la_actividad
Propietario unico	505	NULL	Administracion	111
Propietario unico	505	NULL	Medios de comunicacion	111

Figura 126; Resultado de la consulta 4 de la base de datos centralizada

Base de datos distribuida (Vease figuras 112, 114, 116, 122 del apartado 3.1)

```
SELECT Num_de_identificacion_de_seguridad_social_O, Logotipo, Funcion_O,  
Actividad_O, Capital_registrado, Forma_juridica,  
Codigo_de_registro_empresarial, Codigo_de_incorporacion_geografica, Rol_O,  
Codigo_de_la_actividad  
  
FROM R11Mon2, R12Mon,R13Mon, R16Mon;
```

```

WHERE (R11Mon2.Id_Org=R12Mon.Id_org and
R12Mon.Id_rol_O=R13Mon.Id_rol_O) and
(R11Mon2.Codigo_de_registro_empresarial=R16_Mon.Codigo_de_registro_empresa
rial);

```

A continuación se muestra el modelo de costos para dichas consultas

#### Base de datos centralizada

Acceso a Dat_Org (3 tuplas) y Org_Rol (4 tuplas)	12*tupacc= 12
Acceso a Org_Rol (4 tuplas) y Rol_Org (3 tuplas)	12*tupacc= 12
Acceso a Dat_Org (3 tuplas) y Cod_Act_Org (3 tuplas)	9*tupacc= 9
	—————
El costo total es	33

#### Base de datos distribuida

Acceso a R11Mon2 (1 tuplas) y R12Mon (2 tuplas)	2*tupacc= 2
Acceso a R12Mon (2 tuplas) y R13Mon (3 tuplas)	6*tupacc= 6
Acceso a R11Mon2 (1 tuplas)y R16Mon (1 tuplas)	1*tupacc= 1
	—————
El costo total es	9

Se puede percibir que es mejor la consulta en la base de datos distribuida, ya que es menor el costo total que en la centralizada y se accede a los fragmentos que se encuentran en Monterrey, por lo que no es necesario buscar los datos exclusivos de Monterrey como en la base de datos centralizada, como ya se mencionó esto es bueno porque se utilizan menos recursos y la consulta se realiza de forma más rápida.

5) En México se requiere conocer los datos (*Id\_per*, *Titulo*, *Primer\_nombre*, *Segundo\_nombre*, *Apellido*, *Apellido\_de\_soltera*, *Lugar\_de\_nacimiento*, *Fecha\_de\_nacimiento*, *Sexo*, *Foto*, *Dirección\_P*, *Cod\_postal\_P*, *Ciudad\_P*, *Pais*, *Email\_P*, *Tel\_P*, *Cel\_P*, *Fax\_P*, *Estado\_civil*, *Cabeza\_de\_familia*, *Num\_de\_ninios*, *Grado*, *Num\_de\_socios*, *Sufijo*, *Funcion\_P*, *Actividad\_P*, *Num\_de\_empleado\_P*, *Convenio\_colectivo*, *Coficiente\_de\_salario*, *Nivel\_salarial*, *Num\_de\_identificacion\_de\_seguridad\_social\_P* y *Descripcion\_P*) de las personas que trabajen en Monterrey y en México.

Base de datos centralizada (tomando la figura 73 del apartado 3.1)

SELECT \* FROM Dat\_Per;

Id_per	Titulo	Primer_nombre	Segundo_nombre	Apellido	Lugar_de_nacimiento	Fecha_de_nacimiento	Sexo	Foto	Direccion_P	Cod_postal_P
1	señorita	rocio	del carmen	cardenas	coatlincan	01-05-1980	femenino	BLOB	La purificacion num	56240
2	señor	carlos		sosa	los reyes	15-20-1878	masculino	BLOB	tepetlaoxtoc num 54	56230
3	señora	ana		bracho	texcoco	30-03-1982	femenino	BLOB	jarritos num 12	56220
4	señora	ana	paola	rodriguez	tepepaxan	20-08-1975	femenino	BLOB	la condesa s/n	56221

Ciudad_P	Pais	Email_P	Tel_P	Cel_P	Fax_P	Estado_civil	Cabeza...	Num_de_ninios	Grado	Num_de_socios	Sufijo
mexico	mexico	rocio@hotmail.com	224466	40248	0	cohabitando		0	gerente	0	
mexico	mexico	rcarlos@yahoo.com	40246	80646	156	soltero	si	0	agente	0	
monterrey	mexico	ana@hotmail.com	552030	402030	0	casado		2	accionista	1	
monterrey	mexico	paola@gmail.com	558899	4020698	112	casado		1	agente	0	

Funcion_P	Actividad_P	Num_de_empleado	Convenio_colectivo	Coficiente_de_salario	Nivel_salarial	Numero_de_identificacion_de_seguridad_social	Descripcion_P
gerente de contabilidad y finanzas	intermediacion financiera excepto seguros/seg	1	SYNTEC		nivel2	152040	empleado del area de conta
agente de la compañía	administracion publica y defensa	0			nivel1		es un empleado de la empre
agente de ventas	inmobiliario, alquileres y actividades de negocio	2	SYNTEC		nivel1	113333	persona que trabaja en la er
ejecutivo de la empresa	transporte, almacenamiento y comunicaciones	3	SYNTEC		nivel3	203080	empleado de la empresa qu

Figura 127; Resultado de la consulta 5 de la base de datos centralizada

Base de datos distribuida (tomando las figuras 89 y 90 del apartado 3.1)

R1Mex= *Id\_per*, *Titulo*, *Primer\_nombre*, *Segundo\_nombre*, *Apellido*, *Apellido\_de\_soltera*, *Lugar\_de\_nacimiento*, *Fecha de nacimiento*, *Sexo*, *Foto*, *Dirección\_P*, *Cod\_postal\_P*, *Ciudad\_P*, *Pais*, *Email\_P*, *Tel\_P*, *Cel\_P*, *Fax\_P*, *Estado\_civil*, *Cabeza\_de\_familia*,

*Num\_de\_ninios, Grado, Num\_de\_socios, Sufijo, Funcion\_P, Actividad\_P, Num\_de\_empleado\_P, Convenio\_colectivo, Coeficiente\_de\_salario, Nivel\_salarial, Num\_de\_identificacion\_de\_seguridad\_social\_P y Descripcion\_P*

R1Mon= *Id\_per, Titulo, Primer\_nombre, Segundo\_nombre, Apellido, Apellido\_de\_soltera Lugar\_de\_nacimiento, Fecha\_de\_nacimiento, Sexo, Foto, Dirección\_P, Cod\_postal\_P, Ciudad\_P, Pais, Email\_P, Tel\_P, Cel\_P, Fax\_P, Estado\_civil, Cabeza\_de\_familia, Num\_de\_ninios, Grado, Num\_de\_socios, Sufijo, Funcion\_P, Actividad\_P, Num\_de\_empleado\_P, Convenio\_colectivo, Coeficiente\_de\_salario, Nivel\_salarial, Num\_de\_identificacion\_de\_seguridad\_socia\_Pl y Descripcion\_P*

R1= SELECT \* FROM R1Mex

UNION

SELECT \* FROM R1Mon

Modelo de costos para la consulta 5 de la base de datos centralizada y distribuida

Base de datos centralizada

Acceso a Dat_Per (4 tuplas)	4*tupacc= 4
-----------------------------	-------------

El costo total es	4
-------------------	---

Base de datos distribuida

Transferencia de R1Mon (2 tuplas) a México	2*tuptrans= 20
--	----------------

El costo total es	20
-------------------	----

En este caso la consulta es más viable en la centralizada que en la distribuida dado que el costo es más elevado, además la consulta es más ostentosa para la base de datos distribuida que para la centralizada, pero sólo cuando se hacen estas consultas globales es cuando se pueden utilizar más recursos, aunque dichas consultas no se suelen realizar con frecuencia, ya que las que se realizan con frecuencia son las locales, es decir, son consultas que se realizan a los fragmentos de base de datos que se encuentran en esos sitios.

- 6) Se solicitan los datos *Titulo, Primer\_nombre, Segundo\_nombre, Apellido, Apellido\_de\_soltera, Direccion\_P, Cod\_postal\_P, Tel\_P, Funcion\_P, Actividad\_P, Rol\_P, Habilidades\_P, Producto\_de\_Interes\_P* de las relaciones *Dat\_Per, Rol\_Per, Hab\_Per* y *Pro\_Int\_Per* respectivamente, así como el *Nombre\_oficial* de la organización en la que trabajan, la *Funcion\_O* y el *Rol\_O* de las relaciones *Dat\_Org* y *Rol\_Org* de todas las personas que sean mujeres que se encuentran en México y en Monterrey.

Base de datos centralizada (Figuras 73, 74, 75, 76, 77, 78, 79, 82, 83, 84 y 85 del apartado 3.1 )

```
SELECT Titulo, Primer_nombre, Segundo_nombre, Apellido,
Apellido_de_soltera, Direccion_P, Cod_postal_P, Tel_P, Funcion_P,
Actividad_P, Rol_P, Habilidades_P, Producto_de_Interes_P, Nombre_oficial,
Funcion_O, Rol_O

FROM Dat_Per, Per_Rol, Rol_Per, Per_Hab, Hab_Per, Per_Pro_Int,
Pro_Int_Per, Per_Org, Dat_Org, Org_Rol, Rol_Org

WHERE (Dat_Per.Id_per=Per_Rol.Id_per and
Per_Rol.Id_rol_P=Rol_Per.Id_rol_P) and (Dat_Per.Id_per=Per_Hab.Id_per and
Per_Hab.Id_hab_P=Hab_Per.Id_hab_P) and (Dat_Per.Id_per=Per_Pro_Int.Id_per
and Per_Pro_Int.Id_pro_int_P=Pro_Int_Per.Id_pro_int_P) and
```

(Dat\_Per.Id\_per=Per\_Org.Id\_per and Per\_Org.Id\_org=Dat\_Org.Id\_org) and  
 (Dat\_Org.Id\_org=Org\_Rol.Id\_org and Org\_Rol.Id\_rol\_O=Rol\_Org.Id\_rol\_O) and  
 Sexo='Femenino';

Titulo	Primer_nombre	Segundo_nombre	Apellido	Apellido_de_soltera	Direccion_P	Cod_postal_P	Tel_P	Funcion_P
Señorita	Rocio	del carmen	Cardenas		la purificacion ...	56240	224466	Gerente de contabilid...
Señorita	Rocio	del carmen	Cardenas		la purificacion ...	56240	224466	Gerente de contabilid...
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana		Bracho	Sosa	Jarritos num 12	56220	552030	Agente de ventas
Señora	Ana	Paola	Rodriguez	Lopez	La condesa s/n	56221	558899	Ejecutivo de la empre..
Señora	Ana	Paola	Rodriguez	Lopez	La condesa s/n	56221	558899	Ejecutivo de la empre..

Actividad_P	Rol_P	Habilidades_P	Producto_de_Inter...	Nombre_oficial	Funcion_O	Rol_O
Intermediacion financi...	Interno	Contabilidad	NULL	Uaem texcoco	Ventas	Interno
Intermediacion financi...	Interno	Contabilidad	NULL	Chapingo	Produccion d...	Administracion
Inmobiliario, alquileres ...	Interno	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Administracion	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Medios de comunica...	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Perspectiva de ventas	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Interno	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Administracion	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Medios de comunica...	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Perspectiva de ventas	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Interno	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Administracion	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Medios de comunica...	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Perspectiva de ventas	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Administracion
Inmobiliario, alquileres ...	Interno	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Administracion	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Medios de comunica...	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Perspectiva de ventas	Administracion	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Interno	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Administracion	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Medios de comunica...	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Perspectiva de ventas	Ventas	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Interno	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Administracion	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Medios de comunica...	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Inmobiliario, alquileres ...	Perspectiva de ventas	Mercadeo	Productos publicados	Upem san andres	Mercadotecnia	Medios de comunicacion
Transporte, almacena...	Interno	Tecnologias ...	Componentes y su...	Upem san andres	Mercadotecnia	Administracion
Transporte, almacena...	Interno	Tecnologias ...	Componentes y su...	Upem san andres	Mercadotecnia	Medios de comunicacio

Figura 128; Resultado de la consulta 6 de la base de datos centralizada

Base de datos distribuida (Tomando las figuras 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101 y 102 del apartado 3.1)

#### **Subconsulta**

```
R1= SELECT Id_per, Titulo, Primer_nombre, Segundo_nombre, Apellido,  
Apellido_de_soltera, Direccion_P, Cod_postal_P, Tel_P, Funcion_P,  
Actividad_P, FROM R1Mex
```

UNION

```
SELECT Id_per per, Titulo, Primer_nombre, Segundo_nombre, Apellido,  
Apellido_de_soltera, Direccion_P, Cod_postal_P, Ciudad_P, Funcion_P,  
Actividad_P, ROM R1Mon;
```

#### **Subconsulta**

```
R2=SELECT Id_per, Id_rol_P FROM R2Mex
```

UNION

```
SELECT Id_per, Id_rol_P FROM R2Mon;
```

#### **Subconsulta**

```
R4=SELECT Id_per, Id_hab_P FROM R4Mex
```

UNION

```
SELECT Id_per, Id_hab_P FROM R4Mon;
```

### Subconsulta

```
R6=SELECT Id_per, Id_pro_int_P FROM R6Mex
```

```
UNION
```

```
SELECT Id_per, Id_pro_int_P FROM R6Mon;
```

### Subconsulta

```
R7=SELECT Id_Per, Id_Org FROM R8Mex
```

```
UNION
```

```
SELECT Id_Per, Id_Org FROM R8Mon
```

### Subconsulta

```
R8=SELECT * FROM R11Mex1, R11Mex2 WHERE R11Mex1.Id_org =R11Mex2.Id_org
```

### Subconsulta

```
R9=SELECT * FROM R11Mon1, R11Mon2 WHERE R11Mon1.Id_org =R11Mon2.Id_org
```

### Subconsulta

```
R10= SELECT * FROM R9 U R8 WHERE R9.Id_org=R8.id_org
```

### Subconsulta

```
R11= SELECT Id_org, Id_rol_O FROM R12Mex
```

```
UNION
```

```
SELECT Id_org, , Id_rol_O FROM R12Mon
```

## Subconsulta

```
SELECT Titulo, Primer_nombre, Segundo_nombre, Apellido,  
Apellido_de_soltera, Direccion_P, Cod_postal_P, Tel_P, Funcion_P,  
Actividad_P, Rol_P, Habilidades_P, Producto_de_interes_P, Nombre_oficial,  
Funcion_O, Rol_O  
  
FROM R1, R2, R4, R6, R7, R8, R9, R10, R11, R3Mon, R5Mon, R7Mon, R13Mon  
  
WHERE ((R1.Id_per=R2.Id_per and R2.Id_rol_P=R3Mon.Id_rol_P) and  
  
(R1.Id_per=R4.Id_per and R4.Id_hab_P= R5Mon.Id_hab_P) and  
  
(R1.Id_per=R6.Id_per and R6.Id_pro_int_P=R7Mon.Id_pro_int_P) and  
  
(R1.Id_per=R7.Id_per and R7.Id_org=R10.Id_org) and (R10.Id_org=R11.Id_org  
and R11.Id_rol_O=R13Mon.Id_rol_O) and Sexo='Femenino');
```

## Modelo de costos para dichas consultas

### Base de datos centralizada

Acceso a Dat_Per (4 tuplas) y Per_Rol (8 tuplas)	32*tupacc= 32
Acceso a Per_Rol (8 tuplas) y Rol_Per (6 tuplas)	48*tupacc= 48
Acceso a Dat_Per (4 tuplas) y Per_Hab (6 tuplas)	24*tupacc= 24
Acceso a Per_Hab (6 tuplas) y Hab_Per (5 tuplas)	30*tupacc= 30
Acceso a Dat_Per (4 tuplas) y Per_Pro_Int (5 tuplas)	20*tupacc= 20
Acceso a Per_Pro_Int (5 tuplas) y Pro_Int_Per (5 tuplas)	25*tupacc= 25
Acceso a Dat_Per (4 tuplas) y Per_Org (5 tuplas)	20*tupacc= 20

Acceso a Per_Org (5 tuplas) y Dat_Org (3 tuplas)	15*tupacc= 15
Acceso a Dat_Org (3 tuplas) y Org_Rol (4 tuplas)	12*tupacc= 12
Acceso a Org_Rol (4 tuplas) y Rol_Org (3 tuplas)	12*tupacc= 12
	<hr/>
El costo total es	238
Base de datos distribuida	
Transferir R1Mex a Monterrey	2*tuptrans= 20
Producir R1=R1Mex U R1Mon	2*tupacc= 2
Transferir R2Mex a Monterrey	3*tuptrans= 30
Producir R2=R2Mex U R2Mon	3*tupacc= 3
Transferir R4Mex a Monterrey	2*tuptrans= 20
Producir R4=R4Mex U R4Mon	2*tupacc= 2
Transferir R6Mex a Monterrey	3*truptrans= 30
Producir R6=R6Mex U R6Mon	3*tupacc= 3
Transferir R8Mex (3 tuplas) a Monterrey	3*tuptrans= 30
Producir R7=R8Mex U R8Mon	3*tupacc= 3
Reunión R1 (4 tuplas) con R2 (8 tuplas)	32
Reunión R2 (8 tuplas) con R3Mon (6 tuplas)	48

Reunión R1 (4 tuplas) con R4 (6 tuplas)	24
Reunión R4 (6 tuplas) con R5Mon (5 tuplas)	30
Reunión R1 (4 tuplas) con R6 (5 tuplas)	20
Reunión R6 (5 tuplas) con R7Mon (4 tuplas)	20
Reunión R1 con R7	20
Producir R8= R11Mon1 (1 tupla) reunión con R11Mon2 (1 tupla)	1
Producir R9=R11Mex1 (2 tuplas) reunión con R11Mex2 (2 tupla)	4
Transferir R9 (2 tuplas) a Monterrey	2*tuptrans=20
Producir R10=R9 U R8	2*tupacc= 2
Reunión R7 (5 tuplas) con R10 (3 tuplas)	15
Transferir R12Mex (2 tuplas) a Monterrey	2*tuptrans=20
Producir R11=R12Mex U R12Mon	2*tupacc= 2
Reunión R10 (3 tuplas) con R11(4 tuplas)	12
Reunión R11(4 tuplas) con R13Mon (3 tuplas)	12
	<hr/>
El costo total es	425

Como se muestra la consulta de la base de datos distribuida realiza más operaciones que la consulta de la centralizada, dado que se trata de una consulta global, pero como ya se mencionó

esta clase de consultas no suelen hacerse muy a menudo en el sitio, por lo tanto sólo afecta el rendimiento de la base de datos cuando se hacen consultas globales.

- 7) Se desea conocer todos los datos (de las relaciones Dat\_Org1 de México, Dat\_Org2 de México, Rol\_Org de México, Hab\_Org de México, Cod\_Act\_Org de México) de las organizaciones que hay en México y que se muestren en orden de acuerdo al nombre usual.

Base de datos centralizada (de las figuras 83 a la 88 del apartado 3.1)

```
SELECT  Dat_Org.Id_org, Dat_Org.Nombre_oficial, Nombre_usual, Sitio,
Grupo, Fecha_de_creacion, Direccion_O, Cod_postal_O, Ciudad_O, Region,
Email_O, Tel_O, Fax_O, Num_de_identificacion_de_seguridad_social_O,
Logotipo, Funcion_O, Actividad_O, Capital_registrado, Forma_juridica,
Dat_Org.Codigo_de_registro_empresarial,
Codigo_de_incorporacion_geografica, Codigo_de_IVA, Codigo_EAN_13,
Descripcion_O, Rol_O, Habilidades_O, Codigo_de_la_actividad

FROM  Dat_Org, Org_Rol, Rol_Org, Org_Hab, Hab_Org, Cod_Act_Org

WHERE (Dat_Org.Id_org=Org_Rol.Id_org and
Org_Rol.Id_rol_O=Rol_Org.Id_rol_O) and (Dat_Org.Id_Org=Org_Hab.Id_org and
Org_Hab.Id_hab_O=Hab_Org.Id_hab_O)
and(Dat_Org.Codigo_de_registro_empresarial=Cod_Act_Org.Codigo_de_registro_
empresarial) and Ciudad_O='Mexico' order by Nombre_usual;
```

Id_org	Nombre_oficial	Nombre_usual	Sitio	Grupo	Fecha_de_creacion	Direccion_O	Cod_postal_O	Ciudad_O	Region
2	Chapingo	Chapingo	NULL	Chapingo	19-07-1950	Santiago num 54	55440	Mexico	Mexico
1	Uaem texcoco	Uaem texcoco	Principal	Uaem	22-01-1975	Leyes num 15	56229	Mexico	Mexico

Capital_registrado	Forma_juridica	Codigo_de_registro_empresarial	Codigo_de_incorporacion_geografica	Codigo_de_IVA
200,000	Consumidores de cooperacion	606	NULL	NULL
800,000	Sociedad de responsabilidad limitada	567	NULL	25

Email_O	Tel_O	Fax_O	Num_de_identificacion_de_seguridad_social_O	Logotipo	Funcion_O	Actividad_O
chapingo@gmail.com	33112	620	990	NULL	Produccion de fabricacion	Muebles
uaem@hotmail.com	24545	514	1357	NULL	Ventas	Manufactura textil

Codigo_EAN_13	Descripcion_O	Rol_O	Habilidades_O	Codigo_de_la_actividad
NULL	NULL	Administracion	Administracion de proyecto	213
567348901234	Realiza ventas de manufactura textil	Interno	Ventas	412

Figura 129; Consulta 7 de la base de datos centralizada

Base de datos distribuida (Figuras 109, 110, 113, 115, 117, 119 y 121 del apartado 3.1)

```

SELECT R11Mex1.Id_org, R11Mex1.Nombre_oficial, Nombre_usual, Sitio, Grupo,
Fecha_de_creacion, Direccion_O, Cod_postal_O, Ciudad_O, Region, Email_O,
Tel_O, Fax_O, Num_de_identificacion_de_seguridad_social_O, Logotipo,
Funcion_O, Actividad_O, Capital_registrado, Forma_juridica,
R11Mex2.Codigo_de_registro_empresarial,
Codigo_de_incorporacion_geografica, Codigo_de_IVA, Codigo_EAN_13,
Descripcion_O, Rol_O, Habilidades_O, Codigo_de_la_actividad
FROM R11Mex1, R11Mex2, R12Mex, R13Mex, R14Mex, R15Mex, R16Mex
WHERE (R11Mex1.Id_org= R11Mex2.Id_org and
(R11Mex1.Id_org=R12Mex.Id_Org and R12Mex.Id_rol_O= R13Mex.Id_rol_O) and
(R11Mex1.Id_org=R14Mex.Id_org and R14Mex.Id_hab_O=R15Mex.Id_hab_O) and
(R11Mex2.Codigo_de_registro_empresarial=R16Mex.Codigo_de_registro_empresarial)order by Nombre_usual);

```

A continuación se muestra el modelo de costos para estas consultas

Base de datos centralizada

Reunión de Dat_Org (3 tuplas) y Org_Rol (4 tuplas)	12
Reunión de Org_Rol (4 tuplas) y Rol_Org (3 tuplas)	12
Reunión de Dat_Org (3 tuplas) y Org_Hab (4 tuplas)	12
Reunión de Org_Hab (4 tuplas) y Hab_Org (4 tuplas)	16
Reunión de Dat_Org (3 tuplas) y Cod_Act_Org(3 tuplas)	9
	<hr/>
El costo total es	61

Base de datos distribuida

Reunión de R11Mex1 (2 tuplas) y R11Mex2 (2 tuplas)	4
Reunión de R11Mex1 (2 tuplas) y R12Mex (2 tuplas)	4
Reunión de R12Mex (2 tuplas) y R13Mex (3 tuplas)	6
Reunión de R11Mex1 (2 tuplas) y R14Mex (2 tuplas)	4
Reunión de R14Mex (2 tuplas) y R15Mex (4 tuplas)	8
Reunión de R11Mex2 (2 tuplas) y R16Mex (2 tuplas)	4
	<hr/>
El costo total es	30

Se puede notar que el costo de la consulta distribuida es mejor que el de la centralizada, como ya se mencionó el hecho de hacer una consulta local en una distribuida consume menos recursos, cabe hacer notar que es más rápido acceder a muchas relaciones con pocos registros que a pocas

tablas con muchos registros, aunque lo ideal sería acceder a una base de datos que tenga un equilibrio en ambas cosas.

A continuación se muestra el cuadro comparativo de las consultas de la base de datos centralizada versus distribuida

Numero de Consulta	Costo total	
	Centralizada	Distribuida
1	4	2
2	169	50
3	3	2
4	33	9
5	4	20
6	238	425
7	61	30

Figura 130; Cuadro comparativo de consultas.

### 3.2 Disponibilidad

Como se vio en el capítulo anterior, la disponibilidad es la posibilidad de que la base de datos se encuentre activa, dicha característica de las bases de datos es imprescindible para el buen funcionamiento de la misma.

La disponibilidad en una base de datos centralizada es baja, dado que la información se encuentra en un sólo lugar y aunque se debe tener una copia de seguridad se pierde la

disponibilidad en caso de un desastre natural o de alguna otra circunstancia como falla del disco duro, caída del sistema, error en software de aplicación o mantenimiento (vistas en el punto 2.5.7) cuando lo ideal sería que la base de datos siempre esté disponible para los usuarios.

En cambio, en una base de datos distribuida la probabilidad de que la disponibilidad sea nula es menos propensa, ya que los fragmentos de bases de datos se encuentran en diferentes lugares y además los fragmentos pueden estar replicados en distintos sitios; así que si falla un sitio no hay de que preocuparse pues los usuarios pueden acceder a otros sitios sin saberlo, que contienen la misma información que el sitio que ha fallado, por tanto, la base de datos será disponible.

Además las bases de datos distribuidas tienen robustez (funcionan aunque haya un fallo) y el DBMS cuenta con independencia de falla.

Como se puede apreciar la disponibilidad es mucho mejor en las bases de datos distribuidas que en las centralizadas aún tomando medidas para prevenir el fallo de una base de datos, por tanto, sería una buena opción para ERP5 pues aunque posee una buena disponibilidad, ésta aumentaría todavía más.

### 3.3 Ventajas y Desventajas

Las ventajas y desventajas que se muestran a continuación son sobre las consultas y la disponibilidad en base de datos centralizadas y distribuidas, ya que son los puntos que se evaluarán para la realización de la tesina.

	CONSULTAS		DISPONIBILIDAD	
	VENTAJAS	DESVENTAJAS	VENTAJAS	DESVENTAJAS
<b>CENTRALIZADAS</b>	Las consultas son más sencillas	Las consultas se hacen a un sólo sitio	Se puede hacer una copia de seguridad que esté actualizada	Se pierde la disponibilidad en caso de fallar el disco duro
		No se pueden hacer consultas locales y globales	Para el mantenimiento se pueden buscar herramientas que permitan que la base de datos esté disponible	Se anula la disponibilidad si cae el sistema
		La combinación utiliza mayor tiempo dado que busca en toda la base de datos		No hay disponibilidad si se produce un error de aplicación
				No existe disponibilidad si falla la red de comunicación
				En caso de un desastre natural se elimina la disponibilidad hasta implementar nuevamente la base de datos
		Se pueden hacer consultas locales (se accede a un sólo sitio) y globales (se accede a varios sitios)	En caso de que el DBMS no tenga transparencia de distribución las consultas son complejas pues se deben conocer los fragmentos de base de datos	En caso de que falle el disco duro de un sitio la disponibilidad se mantiene

<b>DISTRIBUIDAS</b>	Si el DBMS tiene transparencia de distribución las consultas se hacen de la misma forma que en una base de datos centralizada		Si cae el sistema de un sitio la disponibilidad se conserva	
	La combinación requiere menos tiempo pues sólo se efectúa en fragmentos de base de datos		Si se produce un error en el software de aplicación de un sitio persiste la disponibilidad	
			Si falla la red de comunicación de un sitio la disponibilidad no se elimina	
			En caso de un desastre natural en un sitio específico la disponibilidad sigue (pues hay fragmentos y réplicas de bases de datos en diferentes sitios)	
			El DBMS cuenta con independencia de falla	
			Tiene robustez	
			Se puede hacer una copia de seguridad que esté actualizada de cada sitio	
			Para el mantenimiento se pueden buscar herramientas que permitan que la base de datos esté disponible	

Figura 131; Cuadro comparativo

## Resultados, conclusiones y recomendaciones

### Resultados

En la presente investigación se realizó una comparación entre las consultas y la disponibilidad de las bases de datos centralizadas contra las bases de datos distribuidas, con el fin de dar a conocer la viabilidad de una base de datos distribuida en un ERP; para llevar a cabo este trabajo de investigación, se tomaron dos apartados (el de personas y organización) del módulo base de ERP5, dichos apartados fueron realizados a través del modelo relacional en una base de datos centralizada para posteriormente poder hacer el modelo relacional de la base de datos distribuida.

Para hacer la base de datos distribuida se requería hacer una fragmentación horizontal en el apartado de personas, ya que se necesitaba conocer todos los datos de las relaciones Dat\_Per, Per\_Rol, Rol\_Per, Per\_Hab, Hab\_Per, Per\_Pro\_Int, Pro\_Int\_Per, Per\_Org, Per\_Nac y Nac\_Per de acuerdo a la zona donde se encuentran, es decir, suponiendo que en México se necesitaban todos los datos de las relaciones mencionadas anteriormente de las personas de este lugar y en Monterrey los datos de las relaciones mencionadas anteriormente de las personas de dicha ciudad. Para el apartado de organizaciones se realizó una fragmentación mixta (vertical horizontal) en la relación Dat\_Org, ya que en México sólo se necesitaba recurrir a ciertos atributos (*Nombre\_oficial, Nombre usual, Sitio, Grupo, Fecha\_de\_creacion, Direccion\_O, Cod\_postal\_O, Ciudad\_O, Region, Email\_O, Tel\_O y Fax\_O*) y en Monterrey se requerían otros atributos (*Num\_de\_identificacion\_de\_seguridad\_social\_O, Logotipo, Funcion\_O, Actividad\_O. Capital\_registrado, Forma\_juridica,Codigo\_de\_registro\_empresarialCodigo\_de\_incorporacion\_geografica, Codigo\_de\_IVA, Codigo\_EAN\_13 y Descripcion\_O*), y al igual que en el apartado de personas se necesitaba conocer los datos de acuerdo a la ciudad; en

el caso de las otras relaciones (Dat\_Org, Org\_Rol, Org\_Hab, Hab\_Org y Cod\_Act\_Org) sólo se aplicó una fragmentación horizontal debido a que se necesitaban los datos respectivos de acuerdo al lugar y no se requería ninguna fragmentación vertical (la cual no se puede aplicar a las otras relaciones, dado que no habría forma), Véase apartado 3.1 Consultas.

Una vez realizado lo anterior se hicieron las mismas consultas para la base de datos centralizada y la distribuida basándose en la sintaxis de Mysql (para la centralizada), ya que la base de datos centralizada se realizó en dicho gestor de base de datos para que no hubiera errores de sintaxis; la base de datos distribuida no fue necesario aplicarla al gestor de base de datos (Mysql), ya que se baso en la misma sintaxis que se utilizó en la base de datos centralizada.

Así mismo para comprobar las ventajas de las consultas en las bases de datos distribuidas respecto a la centralizada se utilizó un modelo de costo en el que se toman en cuenta el acceso (tupacc) a las relaciones, así como la transferencia (tuptrans) de las mismas mostrando claramente un desempeño más óptimo en la distribuida que en la centralizada, el cual se puede apreciar mejor en la figura 130 del apartado 3.1 consultas.

Por lo tanto, después de realizar dichas comparaciones en las consultas y de estudiar la disponibilidad en ambas bases de datos se llegaron a los siguientes resultados:

- Las consultas son más complicadas en una base de datos distribuida que en una centralizada, ya que para realizarlas se tiene que conocer el nombre de todos los fragmentos y si es que éstos se encuentran replicados y en que sitio. No obstante, se puede utilizar un SGBD que cumpla con la mayor cantidad de características (independencia de sitio local, transparencia de fragmentación, transparencia de replicación, entre otras) para soportar una base de datos distribuida (aunque en el

mercado es difícil encontrar un DBMS que cumpla con tales características), en cuyo caso las consultas se realizan de la misma forma que en una base de datos centralizada.

- Al realizar consultas locales de una o más relaciones en una base de datos distribuida se consumen menos recursos, pues se accede a relaciones de menor magnitud y ya no se tiene que andar haciendo la búsqueda en todos los registros para ver cuales cumplen con el criterio de selección, en cambio, en una centralizada la clausula WHERE (consulta global) es más amplia, ya que se tiene que poner el criterio de búsqueda, además de que la búsqueda de los registros que cumplan con dicho criterio es más tardada, dado que se tiene que buscar en una relación con mayor cantidad de registros, por lo tanto, se utilizan más recursos en memoria.
- En las bases de datos distribuidas las consultas globales requieren mayor rendimiento de tiempo y memoria, aunque este tipo de consultas se realizan con menos constancia que las locales, por lo que no afectaría significativamente el uso de la base de datos, en cambio, tanto las consultas locales como las globales, que son términos de las bases de datos distribuidas (las cuales pueden aumentar o disminuir su rendimiento), en las bases de datos centralizadas serían el mismo tipo de consultas (consultas globales), ya que no existen las consultas locales.
- En las bases de datos distribuidas es mucho más rápido realizar la combinación entre relaciones porque éstas son más chicas que en una base de datos centralizada, donde las relaciones son más grandes.
- Las bases de datos distribuidas aumentan significativamente la disponibilidad de la misma, dado que la base de datos no se encuentra físicamente en un sólo lugar, como es el caso de las bases de datos centralizadas, así que si se presenta algún problema de falla

de disco, desastre natural, entre otros, en algún determinado sitio y en otros sitios se requiere realizar consultas locales a la base de datos, la disponibilidad se mantiene, además de que si se presenta alguna falla de las antes mencionadas en un sitio, pero se tienen réplicas de éste en otros sitios, se puede acceder a éste para realizar las consultas, y por tanto la disponibilidad se mantiene.

## **Conclusiones**

Los sistemas de planificación de recursos empresariales (ERP) son de gran ayuda para las empresas, dado que les ayudan a llevar un buen control en todas las áreas funcionales de la organización de forma integral, es decir, si hay un cambio en el área de ventas, automáticamente se verá reflejado en el área de contabilidad o administración, suponiendo que en alguna de estas áreas se maneje el inventario.

Los ERP poseen ciertas características, por ejemplo, están compuestos de módulos de acuerdo a los procesos de negocios de las organizaciones, trabajan de forma integral, pero a la vez cada módulo es independiente, trabajan con una base de datos centralizadas, entre otras características.

Como se mencionó, los ERP hacen uso de una base de datos centralizada, es decir, se encuentran en un sólo sitio, es importante tomar en cuenta la base de datos (tipo) en un sistema porque es donde residen todos los datos con los que cuenta la organización, es por ello que se propuso la opción de utilizar una base de datos distribuida en los ERP, por tanto, se realizó una investigación bibliográfica sobre las bases de datos centralizadas y las distribuidas en un sistema empresarial, cumpliendo así el primer objetivo del trabajo de investigación.

Con base al tema análisis de las bases de datos distribuidas como una alternativa para los sistemas ERP en las empresas se puede concluir que las bases de datos distribuidas aumentan el rendimiento en cuanto a tiempo y recursos, ya que los fragmentos de base de datos se encuentran en varios sitios o emplazamientos, por tanto, hay mayor probabilidad de que un ERP que trabaje con una base de datos distribuida aumente su eficiencia respecto a una base de datos centralizada, es por ello que para la realización de este trabajo de investigación se escogió ERP5.

En la presente investigación se estudiaron esencialmente dos puntos en cuanto a los dos tipos de base de datos según el sitio (centralizadas y distribuidas), las consultas y la disponibilidad.

Cabe destacar que cada base de datos tiene sus ventajas y desventajas, por ejemplo, en cuanto a las consultas en una base de datos centralizada, son mucho más sencillas, pero el rendimiento es más tardado, ya que se accede a relaciones con muchos registros, en cambio, en una base de datos distribuida se pueden hacer consultas locales y globales; las locales tienen la ventaja de que utilizan más recursos, dado que se accede generalmente al o los fragmentos más cercanos de donde se solicita la información, aunque se pueden hacer consultas globales, las cuales no suelen hacerse constantemente, que disminuyen el rendimiento de la base de datos.

Por otro lado, en las bases de datos centralizadas la disponibilidad (el lapso de tiempo en que una base de datos se encuentra activa) es más vulnerable porque se puede presentar una falla en el disco duro, por la realización del mantenimiento, por un error en el software, entre otras fallas, aunque claro está que se pueden tener respaldos, pero se pierde tiempo en poner en marcha nuevamente la base de datos. Cosa que en una base de datos distribuida es difícil que suceda, dado que los fragmentos de la base de datos se encuentran en diferentes sitios, cada fragmento puede tener replicas en otros sitios y además se cuenta con respaldo de ellos.

Como se puede apreciar las ventajas de utilizar una base de datos distribuidas son varias respecto a una centralizada en cuanto a las consultas y la disponibilidad en ERP5 y, por tanto, sería una buena alternativa para dicho sistema. Con esto se demuestra que se realizó una comparación entre las bases de datos distribuidas y las centralizadas considerando las consultas y la disponibilidad, mostrando que una base de datos distribuida es una buena alternativa para ERP5, logrando así el segundo y tercer objetivo de esta tesina.

Cabe mencionar que este sistema empresarial también fue elegido dado que es un sistema de código abierto y se ha implementado en empresas de producción como en organismos de salud, además de que se pueden modificar o añadir módulos según las necesidades de la empresa, ERP5 se basa en Zope, que es un servidor de aplicaciones web, y trabaja con una base de datos orientada a objetos llamada *Zope Object Data Base (ZODB)* para la distribución de procesos, pero utiliza como *backend* una base de datos relacional llamada Mysql o María DB según sea el caso, por tanto, usa una base de datos centralizada.

## **Recomendaciones**

Con base en la presente investigación se recomienda utilizar una base de datos distribuida en los ERP, específicamente en ERP5, ya que aunque las consultas son más complejas también es cierto que aumenta la eficiencia de la base de datos, pues al realizar una consulta local se accede a una menor cantidad de tuplas optimizando el tiempo de búsqueda, es decir, es más eficiente acceder a relaciones pequeñas y con pocas tuplas que a comparación de una base de datos centralizada en la cual se tienen relaciones grandes (por ende con más tuplas), y que al momento de realizar una búsqueda además de que se va a tardar más tiempo, va a consumir más recursos como se explicó en el apartado 3.1 Consultas de esta tesina; en este apartado también se puede

notar que aunque las consultas globales tardan mayor tiempo y consumen más recursos, son consultas que probablemente pocas veces se realizan, pues en el sitio de Monterrey por lo general se requiere acceder a datos propios de dicho lugar, muy esporádicamente se necesitará realizar consultas globales para conocer información sobre otros emplazamientos.

Así mismo la implementación de una base de datos distribuida en ERP5 es recomendable, ya que al realizar una combinación entre las relaciones consume menos tiempo y memoria que una base de datos centralizada, debido a que se tiene una menor cantidad de tuplas que en una base de datos centralizada (Ver apartado 3.1 Consultas).

De igual forma una base de datos distribuida tiene robustez, lo que significa que una base de datos de este tipo en ERP5 incrementaría la disponibilidad de la misma y, por tanto, disminuiría la posibilidad de que el sistema falle, es decir, si se presenta alguna falla del disco, caída del sistema, falla de red, desastre natural, o algún otro error en determinado sitio, la base de datos (los fragmentos de la base de datos) de los otros sitios sigue disponible para su sistema local, ya que en esos otros sitios se podrá hacer uso de sus fragmentos de base de datos y, por tanto, realizar consultas locales a excepción de las globales, en caso de que no se cuente con réplicas en otros sitios del emplazamiento que falló, esto es una ventaja significativa sobre una base de datos centralizada, ya que si falla la base de datos no podrá estar disponible para nadie, porque como ya se mencionó anteriormente sólo se encuentra en un sitio, así que sólo estará disponible hasta que se restaure la base de datos, como se muestra en el apartado 3.2 Disponibilidad.

Del mismo modo se recomienda hacer uso de ERP5, ya que es un sistema empresarial de código abierto, cuenta con documentación Faq, How To y directrices, además de que se puede adaptar a las necesidades de la empresa, sin importar si es una empresa que vende productos, un banco o

hasta una institución de salud; al mismo tiempo es recomendable, ya que se puede instalar tanto en Windows como en Linux, aunque si se requiere instalar en Macintosh se necesita poner en contacto con Nexedi que es la empresa que lo respalda.

Para usar ERP5 se debe tener en cuenta que la empresa debe contar con personal capacitado en base de datos y programación sin necesidad de conocer específicamente el lenguaje python, ya que en [erp5.com](http://erp5.com) se ofrecen tutoriales sobre el mismo y sobre zope, además de que la implementación del mismo puede ser complicada, dado que se puede necesitar información importante que se ofrece en un curso en línea por OSOE que tiene por costo inicial 500 euros y un costo por clase de 200 euros.

Además es importante notar que ERP5 trabaja en línea, por lo que no será de gran ayuda para empresas que tengan un servicio de internet no muy viable.

Es necesario tomar en cuenta que en la página de [erp5.com](http://erp5.com) puede haber información no actualizada, dado que se hicieron cambios recientemente a ERP5, por lo que es aconsejable ponerse en contacto con Nexedi si se tiene alguna duda.

Al instalar ERP5 tal cual, sin hacer ningún cambio al sistema, cuenta con seis módulos, uno donde se muestra información personal relacionada con los empleados, con el cliente, con los proveedores, entre otros; otro módulo de gestión de relación con el cliente, otro módulo de comercio, otro módulo de gestión de datos del producto, otro módulo de contabilidad, y un último módulo que presenta documentos, por lo que es recomendable para empresas dedicadas a la producción y ese tipo de cosas.

En caso de utilizar ERP5, es recomendable asimilar las necesidades de la empresa para estudiar la posibilidad de utilizar una implementación de almacenamiento de ZODB, ya sea ZEO o NEO,

dado que aumentan el rendimiento del sistema empresarial, aunque esto no significa que es obligatorio utilizar alguna implementación de ZODB.

## **GLOSARIO**

**Cod\_Act\_Org:** Proporciona el Código de la actividad que tiene la organización.

**Cod\_Act\_Org de México (R21Mex):** Proporciona el código de la actividad de las empresas de México.

**Cod\_Act\_Org de Monterrey (R21Mon):** Da a conocer el código de la actividad de las empresas de Monterrey.

**Dat\_Org:** Relación que contiene los datos generales de el ente social.

**Dat\_Org 1 de México (R14Mex1):** Relación que muestra la primer parte de datos generales de las empresas que están en México.

**Dat\_Org 2 de México (R14Mex2):** Proporciona la segunda parte de los datos generales de las organizaciones que se encuentran en México.

**Dat\_Org 1 de Monterrey (R14Mon1):** Nombre de la relación que contiene la primer parte de los datos generales de las empresas que están en Monterrey.

**Dat\_Org 1 de Monterrey (R14Mon2):** Brinda la segunda parte de los datos generales de las organizaciones que se encuentran en Monterrey.

**Dat\_Per:** Nombre de la relación que contiene datos personales y generales de una persona.

**Dat\_Per de México (R1Mex):** Relación que contiene los datos generales de una persona en México.

**Dat\_Per de Monterrey (R1Mon):** Nombre de la relación que tiene los datos generales de una persona en Monterrey.

**Hab\_Org:** Nombre de la relación que contiene las habilidades con las que cuenta la empresa.

**Hab\_Org de México (R19Mex):** Relación que tiene las habilidades de las empresas que se encuentran en México.

**Hab\_Org de Monterrey (R19Mon):** Proporciona las habilidades con las que cuentan las organizaciones que están en Monterrey.

**Hab\_Per:** Permite conocer las habilidades de la persona.

**Hab\_Per de México (R5Mex):** Brinda las habilidades de la persona en México.

**Hab\_Per de Monterrey (R5Mon):** Muestra las habilidades de las personas en Monterrey.

**Nac\_Per:** Relación que brinda la nacionalidad de la persona.

**Nac\_Per de México (R13Mex):** Contiene la nacionalidad de las personas que se encuentran en México.

**Nac\_Per de Monterrey (R13Mon):** Muestra la nacionalidad de las personas que se encuentran en Monterrey.

**Org\_Hab:** Brinda la relación entre Dat\_Org y Hab\_Org.

**Org\_Hab de México (R18Mex):** Muestra la relación que existe entre Dat\_Org 1 de México (R14Mex1) y Hab\_Org de México (R19Mex).

**Org\_Hab de Monterrey (R18Mon):** Brinda la relación que existe entre Dat\_Org 1 de Monterrey (R15Mon1) y Hab\_Org de Monterrey (R19Mon).

**Org\_Per:** Proporciona el nombre oficial de la Organización a la que la persona pertenece.

**Org\_Per de México (R9Mex):** Relación que contiene el nombre oficial de las organizaciones que están en México.

**Org\_Per de Monterrey (R9Mon):** Proporciona el nombre oficial de las empresas que se encuentran en Monterrey.

**Org\_Rol:** Relación que proporciona la interacción con Dat\_Org y Rol\_Org.

**Org\_Rol de México (R16Mex):** Vincula las relaciones Dat\_Org 1 de México (R14Mex1) con Rol\_Org de México (R17Mex).

**Org\_Rol de Monterrey (R16Mon):** Relaciona Dat\_Org 1 de Monterrey (R14Mon1) con Rol\_Org de Monterrey (R17Mon).

**Per\_Hab:** Nombre de la relación que contiene las claves de la relación Dat\_Per y Hab\_Per.

**Per\_Hab de México (R4Mex):** Es la relación intermedia entre Dat\_Per de México (R1Mex) y Hab\_Per de México (R5Mex).

**Per\_Hab de Monterrey (R5Mon):** Proporciona la relación entre Dat\_Per de Monterrey (R1Mon) y Hab\_Per de Monterrey (R5Mon).

**Per\_Nac:** Nombre de la relación intermedia entre Dat\_Per y Nac\_Per.

**Per\_Nac de México (R12Mex):** Relación intermedia entre Dat\_Per de México (R1Mex) y Nac\_Per de México (R13Mex), cuya función es relacionar ambas tablas.

**Per\_Nac de Monterrey (R12Mon):** Nombre de la relación intermedia que vincula Dat\_Per de Monterrey (R1Mon) con Nac\_Per de Monterrey (R13Mon).

**Per\_Org:** Permite la relación entre Dat\_Per y Org\_per.

**Per\_Org de México (R8Mex):** Muestra la relación entre Dat\_Per de México (R1Mex) y Org\_Per de México (R9Mex).

**Per\_Org de Monterrey (R8Mon):** Brinda la integridad referencial entre Dat\_Per de Monterrey (R1Mon) y Org\_Per de Monterrey (R9Mon).

**Per\_Pro\_Int:** Relación llave entre Dat\_per y Pro\_Int\_Per.

**Per\_Pro\_Int de México (R6Mex):** Relación intermedia entre Dat\_Per de México (R1Mex) y Pro\_Int\_Per de México (R7Mex).

**Per\_Pro\_Int de Monterrey (R6Mon):** Proporciona la integridad referencial entre Dat\_Per de Monterrey (R1Mon) y Pro\_Int\_Per de Monterrey (R7Mon).

**Per\_Rol:** Relación intermedia entre Dat\_Per y Rol\_Per que permite relacionar ambas relaciones.

**Per\_Rol de México (R2Mex):** Proporciona la integridad referencial entre Dat\_Per de México (R1Mex) y Rol\_Per de México (R3Mex).

**Per\_Rol de Monterrey (R2Mon):** Relaciona Dat\_Per de Monterrey (R1Mon) y Rol\_Per de Monterrey (R3Mon).

**Pro\_Int\_Per:** Nombre de la relación que tiene el producto en el que la persona está interesada.

**Pro\_Int\_Per de México (R7Mex):** Producto(s) en el que puede estar interesada la(s) persona(s) de México.

**Pro\_Int\_Per de Monterrey (R7Mon):** Nombre de la relación que lleva el producto (s) en el que la(s) persona(s) pueden estar interesadas.

**Rol\_Org:** Permite mostrar los roles de la organización.

**Rol\_Org de México (R17Mex):** Nombre de la relación que contiene el rol de las empresas en México.

**Rol\_Org de Monterrey (R17Mon):** Proporciona el rol de las organizaciones que se encuentran en Monterrey.

**Rol\_Per:** Relación que contiene los roles de la persona.

**Rol\_Per de México (R3Mex):** Brinda el rol de las personas que están en México.

**Rol\_Per de Monterrey (R3Mon):** Muestra el rol de las personas que se encuentran en Monterrey.

## Referencias Bibliográficas

Alarcón, J. A. (s.f) *Reingeniería de procesos empresariales Teoría y práctica de la Reingeniería de la empresa a través de su estrategia, sus procesos y sus valores corporativos*. Recuperado de <http://books.google.com.mx/books?id=TeveToI-e5MC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Alviso, A. D.(2005) : *MRP II Evolución y desarrollo*. (Tesis de ingeniería). Recuperado de <http://dgsa.uaeh.edu.mx:8080/bibliotecadigital/bitstream/231104/334/1/MRP%20II%20evolucion%20y%20desarrollo.pdf>

Athó, F.(2006)Bases de datos distribuidas. Recuperado de <http://seccperu.org/.../fetch.php?...databases...bases de datos distribuidas...>

Benvenuto, Á.(2006). *Implementación de sistemas ERP, su impacto en la gestión de la empresa e integración con otras tic*. Capiv Review, Vol 4, 37-46. Recuperado de <http://www.capic.cl/capic/media/ART3Benvenuto.pdf>

Club BPM. (2010). *El libro del BPM 2010 Tecnologías, Conceptos, Enfoques Metodológicos y Estándares*. Recuperado de [http://books.google.com.mx/books?id=-RKPoU2h\\_DMC&printsec=frontcover&hl=es#v=onepage&q&f=false](http://books.google.com.mx/books?id=-RKPoU2h_DMC&printsec=frontcover&hl=es#v=onepage&q&f=false)

Club BPM.(2011). *El libro del BPM 2011 Tecnologías, Conceptos, Enfoques Metodológicos y Estándares*. Recuperado de <http://books.google.com.mx/books?id=8Hc5Q00RXEkC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Coronel, C., Morris S., Rob, P.(2011) *Bases de Datos Diseño, implementación y administración*. México: CENGAGE Learning.

de Pablos, C., Lopéz-Hermoso, J. J., Martín-Romo, S., Medina, S.(2004). *Informática y comunicaciones en la empresa*. Recuperado de <http://books.google.com.pe/books?id=U0MXWtqjxtsC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Díaz, M. A.(2009) *Estudio para la implementación de un erp en una empresa de aduanas*. Recuperado de [http://www.recercat.net/bitstream/handle/2072/43775/PFC\\_MiguelAngelDiazSilva.pdf?sequence=1](http://www.recercat.net/bitstream/handle/2072/43775/PFC_MiguelAngelDiazSilva.pdf?sequence=1)

Elmasri, R., Navathe, S. B.(2007) Fundamentos de sistemas de bases de datos. México: PEARSON EDUCACIÓN.

ERP5 (s.f) ERP5 Recuperado de <http://www.erp5.com/>

EvaluandoERP.com(2012). *¿SAP, ORACLE o MICROSOFT?* .Recuperado de <http://www.evaluandoerp.com/nota-2559-SAP-ORACLE-o-MICROSOFT.html>

Gartner. (2012). *Magic Quadrant for Single-Instance ERP for Product-Centric Midmarket Companies.* Recuperado de <http://www.computing.es/gestion/tendencias/1061277001401/gartner-cuadrante-magico-erp-unica.1.html>

Giner, F.(2004).*Los sistemas de información en la sociedad del conocimiento.* Recuperado de <http://books.google.com.mx/books?id=94sv48wCJAMC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Hitpass, B.(2012) *BPM Business Process Management Fundamentos y Conceptos de Implementación.* Recuperado de [http://books.google.com.mx/books?id=Dm4-MGAy5vMC&hl=es&source=gbs\\_similarbooks](http://books.google.com.mx/books?id=Dm4-MGAy5vMC&hl=es&source=gbs_similarbooks)

Laudon, C. K., y Laudon, P. J.(2008).Sistemas de información gerencial administración de la empresa digital. México: Pearson Educación

Lazo, S.(2010). Impacto del Enterprise Resource Planning (ERP) en las Empresas. 360, N°5. Recuperado de <http://cremc.ponce.inter.edu/360/revista360/tecnologia/Impacto%20del%20Enterprise%20Resource%20Planning%20ERP.pdf>

López, F. A.(2002) Ventajas y retos en el uso de bases de datos distribuidas. Recuperado de <http://www.utm.mx/temas/temas-docs/ensayo1t18.pdf>

Muñis, L.(2004). *ERP Guía práctica para la selección e implantación.* Recuperado de [http://books.google.es/books?id=etQc3\\_PXnQoC&printsec=frontcover&hl=es#v=onepage&q&f=false](http://books.google.es/books?id=etQc3_PXnQoC&printsec=frontcover&hl=es#v=onepage&q&f=false)

<http://blog.guiasenior.com/archives/2010/07/demanda-erp-2do-trimestre-2010.html>

Nevado, M., V. (s.f) Introducción a las bases de datos relacionales. Recuperado de

<http://books.google.com.mx/books?id=0lUpB1lNUdIC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Osorio, F., L.(2008) Bases de datos relacionales Teoría y práctica. Recuperado de <http://books.google.es/books?id=buM5rlZME-cC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Ozsu, T., Valduriez, P.(s.f)Principles of Distributed Database System. Canada: Springer.

Pons, O., Marin, N., Medina, J. N., Acid, S., Vila, M. A. (2005) Introducción a las bases de datos El modelo relacional. Recuperado de <http://books.google.com.mx/books?id=ijM1S2tvk0EC&printsec=frontcover&hl=es#v=onepage&q&f=false>

Quiroga, E., Martínez, R.(s.f) Estructuras de datos . Recuperado de <http://books.google.com.mx/books?id=0lLGBtZuq1wC&pg=PT181&lpg=PT181&dq=ARBOL+ES+B%2B+DE+QUIROGA&source=bl&ots=bq3MUuRM9P&sig=CkQUdhwX6dHID4YP6PX26VpxTTs&hl=es&sa=X&ei=oNz9Uav-MJDU9ATluIHwCQ&ved=0CDcQ6AEwAg#v=onepage&q=ARBOLES%20B%2B%20DE%20QUIROGA&f=false>

Rodríguez M, A. F., Pineda M, J. A., Sánchez O, R.(2002). *Sistemas de planificación de recursos empresariales: un caso real*. Recuperado de <http://www.iie.org.mx/bolDPATY02/apli.pdf>

Romo, J. M.(2008).*Estudio para la implantación de un erp en una empresa de transportes*. Recuperado de <http://www.recercat.net/bitstream/handle/2072/14082/PFC%20Jose%20Miguel%20Romo%20Arguello.pdf?sequence=1>

Senior Blog. (2010). *Demanda ERP | 2do. Trimestre 2010*. Recuperado de <http://blog.guiasenior.com/archives/2010/07/demanda-erp-2do-trimestre-2010.html>

Silberschatz, A., Korth, H. F., Sudarshan, S (2007) *Fundamentos de bases de datos*. Mexico: McGraw Hill

Software Guru.(s.f). *ERPs. Apostándole a las PyMes*. Recuperado de <http://sg.com.mx/content/view/455>

Stair, R. M., Reynolds, G. W.(2000). *Principios de Sistemas de información Enfoque administrativo*. Recuperado de

[http://books.google.com.mx/books?id=k\\_sKKIF0iCgC&printsec=frontcover&hl=es#v=onepage&q&f=false](http://books.google.com.mx/books?id=k_sKKIF0iCgC&printsec=frontcover&hl=es#v=onepage&q&f=false)

Tanenbahum, A. S.(2003). *Redes de computadoras*. México:Pearson Prentice Hall.

Valda, J. C.(2011). *Sistema ERP. Definición, funcionamiento, ventajas y desventajas*. Recuperado de <http://jcvalda.wordpress.com/2011/06/26/sistema-erp-definicion-funcionamiento-ventajas-y-desventajas/>

Texcoco, México a 11 de julio de 2013

M. EN C. JUAN MANUEL MUÑOZ ARAUJO  
SUDIRECTOR ACADEMICO DEL  
CENTRO UNIVERSITARIO UAEM  
TEXCOCO

PRESENTE:

M. en P.P Antonio Inoue Cervantes

AT'N RESPONSABLE DEL DEPARTAMENTO DE TITULACIÓN.

Con base en las revisiones efectuadas al trabajo escrito titulado "Análisis de las bases de datos distribuidas como una alternativa para los sistemas ERP en las empresas" que para obtener el título de Licenciado en Informática administrativa presenta el (la) sustentante Rodríguez León Karla Melissa, con número de cuenta 0823631 respectivamente, se concluye que cumple con los requisitos teórico metodológicos necesarios para su aprobación, pudiendo continuar con la etapa de digitalización del trabajo escrito.

ATENTAMENTE

Dra. en C. Lisbeth Rodríguez Mazahua,

Revisora

M. en C. José Sergio Ruíz Castilla

Revisor

M. en C. Yedid Erandini Niño Membrillo

Directora

C.c.p sustentante Karla Melissa Rodríguez León

C.c.p M en C. Yedid Erandini Niño Membrillo

C.c.p M en P.P Antonio Inoue Cervantes

