



**ITSSAT**

## LISTA DE COTEJO PARA INVESTIGACION

<b>INTITUTO TECNOLOGICO SUPERIOR DE SAN ANDRES TUXTLA</b>		
CARRERA: INGENIERIA INFORMÁTICA		
DATOS GENERALES DEL PROCESO DE EVALUACION		
Nombre(s) del alumno(s): <b>MEZO BUSTAMANTE RICARDO</b>		Firma del alumno(s):
Producto: Investigación Unidad I	Nombre del Proyecto: Investigación: Objetivos de las base de datos distribuidas	Fecha: 23/Septiembre/2022
Asignatura: <b>TOPICOS DE BASE DE DATOS</b>	Grupo: 710 - A	Semestre: SEPTIMO
Nombre del Docente: MTI Lorenzo de Jesús Organista Oliveros		Firma del Docente:

<b>INSTRUCCIONES</b>				
Revisar las actividades que se solicitan y marque en los apartados "SI" cuando la evidencia se cumple; en caso contrario marque "NO". En la columna OBSERVACIONES indicaciones que puedan ayudar al alumno a saber cuáles son las condiciones no cumplidas, si fuese necesario.				
Valor del reactivo	Características a cumplir (Reactivo)	CUMPLE		OBSERVACIONES
		SI	NO	
1%	<b>Presentación.</b> El trabajo cumple con los requisitos de :	X		
1%	a. Buena presentación	X		
1%	b. No tiene faltas de ortografía	X		
1%	c. Mismo formato (letra arial 14, títulos con negritas)	X		
1%	d. Misma calidad de hoja e impresión	X		
1%	e. Maneja el lenguaje técnico apropiado	X		
2%	<b>Introducción y Objetivo.</b> La introducción y el objetivo dan una idea clara del contenido del trabajo, motivando al lector a continuar con su lectura y revisión.	X		
5%	<b>Sustento Teórico.</b> Presenta un panorama general del tema a desarrollar y lo sustenta con referencias bibliográficas y cita correctamente a los autores.	X		
2%	<b>Desarrollo.</b> Sigue una metodología y sustenta todos los pasos que se realizaron al aplicar los conocimientos obtenidos, es analítico y bien ordenado.	X		
2%	<b>Resultados.</b> Cumplió totalmente con el objetivo esperado, tiene aplicaciones concretas.	X		
2%	<b>Conclusiones.</b> Las conclusiones son claras y acordes con el objetivo esperado.	X		
2%	<b>Responsabilidad.</b> Entregó el reporte en la fecha y hora señalada.	X		
<b>20%</b>	<b>CALIFICACIÓN:</b>			<b>20%</b>



TECNOLÓGICO  
NACIONAL DE MÉXICO

**TECNOLÓGICO NACIONAL DE MÉXICO**



***Instituto Tecnológico Superior de San Andrés Tuxtla***

***Ingeniería Informática***

***Asignatura: Tópicos de Base de Datos***

***Semestre: Séptimo Semestre***

***Grupo: IINF710A***

***Docente: MTJ. Lorenzo de Jesús Organista Oliveros***

***Alumno: Ricardo Mezo Bustamante***

***-Presenta-***

***Objetivos de la Base de Datos Distribuidas***

***Período Escolar: Agosto-Diciembre 2022***

***San Andrés Tuxtla Ver, 14 de Septiembre de 2022***

## Introducción

Los datos son fundamentales para el funcionamiento de muchas de las aplicaciones y sitios web actuales. Los comentarios en un video viral, los puntajes cambiantes en un juego multijugador y los artículos que dejó en un carrito de compras en su tienda en línea favorita son fragmentos de información almacenados en algún lugar de una base de datos.

En términos generales, una base de datos es cualquier colección de información modelada lógicamente. Una base de datos no necesariamente tiene que estar almacenada en una computadora, y cosas como una pila de archivos de pacientes en un hospital, un conjunto de contactos o un archivador lleno de facturas antiguas, todos califican como ejemplos de bases de datos, los datos pueden ser hechos relacionados con cualquier objeto en consideración. Por ejemplo, tu nombre, edad, altura, peso, etc. son algunos datos relacionados contigo. Una fotografía, imagen, archivo, pdf, etc. también pueden considerarse datos.

En el contexto de los sitios web y las aplicaciones, cuando las personas se refieren a una "base de datos", a menudo se refieren a un programa informático que les permite interactuar con su base de datos. Estos programas, conocidos más formalmente como un sistema de gestión de bases de datos (DBMS), a menudo se instalan en un servidor privado virtual y se accede a ellos de forma remota. Redis, MariaDB y PostgreSQL son algunos ejemplos de DBMS de código abierto, pero hay muchos diferentes disponibles en la actualidad. Los diferentes DBMS generalmente tienen sus propias características únicas y conjuntos de herramientas asociados, pero generalmente se clasifican en una de dos categorías: bases de datos relacionales y no relacionales.

Analicemos un ejemplo de base de datos: un directorio telefónico en línea utiliza una base de datos para almacenar datos de personas, números de teléfono y otros detalles de contacto. Su proveedor de servicios de electricidad utiliza una base de datos para administrar la facturación, los problemas relacionados con el cliente, manejar los datos de fallas, etc. Consideremos también Facebook. Necesita almacenar, manipular y presentar datos relacionados con los miembros, sus amigos, las actividades de los miembros, los mensajes, los anuncios y mucho más. Podemos proporcionar un número incontable de ejemplos para el uso de bases de datos.

Una base de datos distribuida aparece para un usuario como una sola base de datos pero, de hecho, es un conjunto de bases de datos almacenadas en varias computadoras. Los datos en varias computadoras pueden ser accedidos y modificados simultáneamente usando una red. Cada servidor de base de datos en la base de datos distribuida está controlado por su DBMS local y cada uno coopera para mantener la consistencia de la base de datos global.

## **Base de Datos Distribuida**

Una base de datos distribuida es una base de datos que no está limitada a un sistema informático. Es como una base de datos que consta de dos o más archivos ubicados en diferentes computadoras o sitios, ya sea en la misma red o en una red completamente diferente. En lugar de almacenar todos los datos en una base de datos, los datos se dividen y almacenan en diferentes ubicaciones o sitios que no comparten ningún componente físico.

El sistema de base de datos distribuida es una base de datos que está integrada con una colección de bases de datos que están conectadas físicamente a una red informática. No se limita a un solo sistema, sino que se distribuye en diferentes redes informáticas o múltiples bases de datos. En el mundo moderno de hoy, los datos se administran para ayudar a las actividades comerciales que se han extendido a nivel mundial. Las bases de datos distribuidas se utilizan para el escalado horizontal y están diseñadas para cumplir con los requisitos de la carga de trabajo sin tener que realizar cambios en la aplicación de la base de datos o escalar verticalmente una sola máquina.

Las bases de datos distribuidas resuelven varios problemas, como disponibilidad, tolerancia a fallas, rendimiento, latencia, escalabilidad y muchos otros problemas que pueden surgir del uso de una sola máquina y una sola base de datos. Una base de datos distribuida representa varias bases de datos interconectadas distribuidas en varios sitios conectados por una red. Dado que todas las bases de datos están conectadas, aparecen como una sola base de datos para los usuarios.

Las bases de datos distribuidas utilizan múltiples nodos. Se escalan horizontalmente y desarrollan un sistema distribuido. Más nodos en el sistema proporcionan más poder de cómputo, ofrecen mayor disponibilidad y resuelven el problema del punto único de falla. Las diferentes partes de la base de datos distribuida se almacenan en varias ubicaciones físicas y los requisitos de procesamiento se distribuyen entre los procesadores en varios nodos de la base de datos. Un sistema centralizado de administración de bases de datos distribuidas (DDBMS) administra los datos distribuidos como si estuvieran almacenados en una ubicación física. DDBMS sincroniza todas las operaciones de datos entre las bases de datos y garantiza que las actualizaciones en una base de datos se reflejen automáticamente en las bases de datos de otros sitios.

Una base de datos distribuida es básicamente una base de datos que no se limita a un sistema, sino que se distribuye en diferentes sitios, es decir, en varias computadoras o en una red de computadoras. Un sistema de base de datos distribuida se encuentra en varios sitios que no comparten componentes físicos. Esto puede ser necesario cuando una base de datos particular necesita ser accedida por varios usuarios globalmente. Debe administrarse de tal manera que para los usuarios parezca una sola base de datos.

## Características De La Base De Datos Distribuida

Algunas características generales de las bases de datos distribuidas son:

- **Independencia de la ubicación** : los datos se almacenan físicamente en múltiples sitios y son administrados por un DDBMS independiente.
- **Procesamiento** de consultas distribuidas: las bases de datos distribuidas responden consultas en un entorno distribuido que administra datos en varios sitios. Las consultas de alto nivel se transforman en un plan de ejecución de consultas para una gestión más sencilla.
- **Gestión de transacciones distribuidas** : proporciona una base de datos distribuida consistente a través de protocolos de compromiso, técnicas de control de concurrencia distribuida y métodos de recuperación distribuidos en caso de muchas transacciones y fallas.
- **Integración perfecta** : las bases de datos de una colección suelen representar una sola base de datos lógica y están interconectadas.
- **Vinculación de red** : todas las bases de datos de una colección están vinculadas por una red y se comunican entre sí.
- **Procesamiento de transacciones** : las bases de datos distribuidas incorporan el procesamiento de transacciones, que es un programa que incluye una colección de una o más operaciones de base de datos. El procesamiento de transacciones es un proceso atómico que se ejecuta por completo o no se ejecuta en absoluto.

### Tipos de bases de datos distribuidas

Hay dos tipos de bases de datos distribuidas:

- **homogéneo**
- **heterogéneo**

#### Homogéneo

Una base de datos distribuida homogénea es una red de bases de **datos idénticas** almacenadas en varios sitios. Los sitios tienen el mismo sistema operativo, DDBMS y estructura de datos, lo que los hace fácilmente manejables. Las bases de datos homogéneas permiten a los usuarios acceder a los datos de cada una de las bases de datos sin problemas.

#### Heterogéneo

Una base de datos distribuida heterogénea utiliza **diferentes** esquemas, sistemas operativos, DDBMS y diferentes modelos de datos.

En el caso de una base de datos distribuida heterogénea, un sitio en particular puede ignorar por completo a otros sitios, lo que provoca una cooperación limitada en el procesamiento de las solicitudes de los usuarios. La limitación es por qué se requieren traducciones para establecer la comunicación entre sitios.

## **Almacenamiento De Bases De Datos Distribuidas**

El almacenamiento de bases de datos distribuidas se administra de dos maneras:

- **Replicación**
- **Fragmentación**

### **Replicación**

En la replicación de bases de datos, los sistemas almacenan copias de datos en diferentes sitios. Si una base de datos completa está disponible en varios sitios, es una base de datos completamente redundante. La ventaja de la replicación de la base de datos es que aumenta la disponibilidad de los datos en diferentes sitios y permite que se procesen las solicitudes de consulta en paralelo.

Sin embargo, la replicación de la base de datos significa que los datos requieren actualizaciones constantes y sincronización con otros sitios para mantener una copia exacta de la base de datos. Cualquier cambio realizado en un sitio debe registrarse en otros sitios, de lo contrario se producirán inconsistencias.

Las actualizaciones constantes provocan una gran sobrecarga del servidor y complican el control de la concurrencia, ya que se deben verificar muchas consultas simultáneas en todos los sitios disponibles.

### **Fragmentación**

Cuando se trata de la fragmentación del almacenamiento de bases de datos distribuidas, las relaciones se fragmentan, lo que significa que se **dividen en partes más pequeñas**. Cada uno de los fragmentos se almacena en un sitio diferente, donde se requiere.

El requisito previo para la fragmentación es asegurarse de que los fragmentos puedan reconstruirse posteriormente en la relación original sin perder datos.

La ventaja de la fragmentación es que no hay copias de datos, lo que evita la inconsistencia de datos.

Hay dos tipos de fragmentación:

- **Fragmentación horizontal:** el esquema de relación se fragmenta en grupos de filas y cada grupo (tupla) se asigna a un fragmento.

- **Fragmentación vertical:** el esquema de relación se fragmenta en esquemas más pequeños y cada fragmento contiene una clave candidata común para garantizar una unión sin pérdidas.

## **Ventajas Y Desventajas De La Base De Datos Distribuida**

A continuación, se presentan algunas ventajas y desventajas clave de las bases de datos distribuidas:

Las ventajas y desventajas se explican en detalle en las siguientes secciones.

### **Ventajas**

- **Desarrollo Modular** . El desarrollo modular de una base de datos distribuida implica que un sistema se puede expandir a nuevas ubicaciones o unidades agregando nuevos servidores y datos a la configuración existente y conectándolos al sistema distribuido sin interrupción. Este tipo de expansión no provoca interrupciones en el funcionamiento de las bases de datos distribuidas.
- **Confiabilidad** \_ Las bases de datos distribuidas ofrecen una mayor confiabilidad en contraste con las bases de datos centralizadas. En caso de falla de la base de datos en una base de datos centralizada, el sistema se detiene por completo. En una base de datos distribuida, el sistema funciona incluso cuando ocurren fallas y solo ofrece un rendimiento reducido hasta que se resuelve el problema.
- **Menor costo de comunicación** . El almacenamiento local de datos reduce los costos de comunicación para la manipulación de datos en bases de datos distribuidas. El almacenamiento local de datos no es posible en bases de datos centralizadas.
- **Mejor Respuesta** . La distribución eficiente de datos en un sistema de base de datos distribuida proporciona una respuesta más rápida cuando las solicitudes de los usuarios se cumplen localmente. En las bases de datos centralizadas, las solicitudes de los usuarios pasan por la máquina central, que procesa todas las solicitudes. El resultado es un aumento en el tiempo de respuesta, especialmente con muchas consultas.

### **Desventajas**

- **Software costoso** . Garantizar la transparencia y la coordinación de los datos en múltiples sitios a menudo requiere el uso de software costoso en un sistema de base de datos distribuido.
- **Sobrecarga grande** . Muchas operaciones en varios sitios requieren numerosos cálculos y una sincronización constante cuando se utiliza la

replicación de la base de datos, lo que provoca una gran sobrecarga de procesamiento.

- **Integridad de datos** . Un posible problema al usar la replicación de bases de datos es la integridad de los datos, que se ve comprometida al actualizar los datos en varios sitios.
- **Distribución de datos inadecuada** . La capacidad de respuesta a las solicitudes de los usuarios depende en gran medida de la distribución adecuada de los datos. Eso significa que la capacidad de respuesta se puede reducir si los datos no se distribuyen correctamente en varios sitios.

### **Ejemplos De Bases De Datos Distribuidas**

Aunque hay muchas bases de datos distribuidas para elegir, algunos ejemplos de bases de datos distribuidas incluyen Apache Ignite, Apache Cassandra , Apache HBase , Couchbase Server, Amazon SimpleDB , Clusterpoint y FoundationDB.

Apache Ignite se especializa en almacenar y computar grandes volúmenes de datos en grupos de nodos. En 2014, Ignite fue abierto por GridGain Systems y luego fue aceptado en el programa Apache Incubator. La base de datos de Apache Ignite utiliza RAM como nivel de almacenamiento y procesamiento predeterminado.

Apache Cassandra ofrece soporte para clústeres que abarcan varias ubicaciones y presenta su propio lenguaje de consulta, Cassandra Query Language (CQL). Además, las estrategias de replicación de Cassandra son configurables.

Apache HBase se ejecuta sobre el sistema de archivos distribuidos de Hadoop y proporciona una forma tolerante a fallas para almacenar grandes cantidades de datos dispersos. También presenta compresión, operación en memoria y filtros Bloom por columna. HBase no pretende reemplazar la base de datos SQL, aunque Apache Phoenix proporciona una capa SQL para HBase.

Couchbase Server es un paquete de software NoSQL que es ideal para aplicaciones interactivas que sirven a múltiples usuarios simultáneos al crear, almacenar, recuperar, agregar, manipular y presentar datos. Para respaldar estas numerosas necesidades de aplicaciones, Couchbase Server proporciona un valor de clave escalable y acceso a documentos JSON .

Amazon SimpleDB se utiliza como servicio web con Amazon Elastic Compute Cloud y Amazon S3 . Amazon SimpleDB permite a los desarrolladores solicitar y almacenar datos con una gestión de base de datos y una responsabilidad administrativa mínimas.

Clusterpoint elimina la complejidad, los problemas de escalabilidad y las limitaciones de rendimiento de las arquitecturas de bases de datos relacionales. Los datos se administran en formato XML o JSON mediante API abiertas. Debido a que

Clusterpoint es una base de datos de documentos sin esquemas, elimina los problemas de escalabilidad y rendimiento que enfrentan la mayoría de las arquitecturas de bases de datos relacionales.

FoundationDB es una base de datos multimodelo diseñada en torno a una base de datos central que expone un almacén de valores clave ordenado con cada transacción. Estas transacciones admiten propiedades ACID y son capaces de leer y escribir claves almacenadas en cualquier máquina dentro del clúster. Las características adicionales aparecen en capas alrededor de este núcleo.

### **Objetivos De La Base De Datos Distribuida**

El concepto de sistema de gestión de bases de datos distribuidas es almacenar los datos en diferentes ubicaciones. Hay varios objetivos para distribuir los datos.

**Fiabilidad:** El sistema de administración de bases de datos distribuidas es más confiable si alguno de los sistemas informáticos conectados no funciona, entonces otras computadoras pueden completar la tarea sin demora.

**Disponibilidad:** Si un servidor de computadora no funciona y dejó de funcionar en cualquier momento, los otros servidores de computadora pueden realizar la tarea según lo solicitado.

**Actuación:** Se puede acceder a los datos y la información desde sistemas de gestión de bases de datos distribuidas desde diferentes ubicaciones. Es muy fácil de manejar y mantener.

**Catalogación.** A medida que los datos en un sistema distribuido se distribuyen entre ubicaciones, se vuelve imperativo contar con una catalogación de lo que se almacena en qué ubicación junto con detalles sobre su autonomía y confidencialidad. Los detalles de la base de datos con respecto a la distribución de datos, la fragmentación de datos y la replicación deben expandirse en el catálogo de la base de datos, que es dinámico y está en constante expansión.

**Recuperabilidad de datos:** Una de las principales ventajas de DDMS es que se extiende a través de una red de computadoras con componentes independientes y, por lo tanto, la recuperación de datos en un sitio en particular se convierte en una función importante del sistema distribuido.

**Procesamiento de consultas distribuidas:** En un entorno donde los datos se encuentran en varios sitios en un sistema distribuido, el procesamiento de consultas distribuidas se utiliza para procesar las consultas.

**Gestión de transacciones de datos:** En un sistema distribuido, las transacciones tienen lugar en diferentes sistemas en varias ubicaciones físicas, por lo tanto, para completar una transacción, el Administrador de transacciones de datos se comunica con todos los administradores de transacciones locales. Por lo tanto, la gestión de transacciones de datos juega un papel muy importante.

## Conclusión

En las últimas décadas, las bases de datos distribuidas han recorrido un largo camino. Sin embargo, todavía tienen algunos desafíos clave que vale la pena mencionar. Las bases de datos distribuidas ofrecen transparencia de ubicación de datos con autonomía local. Esto significa que, aunque es posible que las aplicaciones no sepan dónde residen exactamente los datos, cada sitio tiene la capacidad de controlar los datos locales, administrar la seguridad, realizar un seguimiento de las transacciones y recuperarse cuando ocurren fallas en el sitio local. La autonomía está disponible incluso si las conexiones a otros sitios han fallado. Esto permite una mayor flexibilidad donde los datos específicos almacenados en sitios particulares pueden necesitar más seguridad y controles de cumplimiento frente a otros datos que no lo pueden hacer.

La extensibilidad de los sistemas de administración de bases de datos (DBMS) se ha investigado durante mucho tiempo en la academia y la industria, y se considera que es una gran historia de éxito para la investigación de bases de datos. Muchos de los resultados obtenidos ahora están disponibles en productos. Todavía no está claro cuál sería la mejor solución para un escenario de aplicación específico y si todos estos enfoques competitivos seguirán coexistiendo. Los criterios decisivos que afectarán el resultado de esta competencia son el soporte para el conjunto correcto de componentes, el soporte adecuado para el desarrollo de aplicaciones, el rendimiento, soporte adecuado para el desarrollo de componentes y gestión de bibliotecas de componentes. La tecnología de bases de datos ha recorrido un largo camino para enfrentar los nuevos desafíos que se encuentran en numerosas aplicaciones actuales.

La base de datos distribuida, como su nombre lo indica, es el proceso de distribuir una base de datos en varias o en una red de varias computadoras o distribuidas para un sistema de administración eficiente. Generalmente, la red de computadoras utilizadas para los sistemas distribuidos podría estar ubicada en la misma ubicación física o pueden estar ubicadas globalmente en varias partes del mundo conectadas a través de un único sistema de gestión. Estas bases de datos distribuidas podrían alojarse en un servidor de red o en intranets y extranets corporativas o, más recientemente y más popularmente, en computadoras independientes descentralizadas alojadas en Internet.

Los datos son el elemento vital del negocio, razón por la cual se necesita una base de datos en el centro de todo. Sin embargo, no todas las bases de datos pueden satisfacer las crecientes necesidades de datos de las empresas actuales. En particular, necesita un sistema de base de datos distribuido que le permita innovar y transformarse sin esfuerzo.

## Referencias Bibliográficas

[1]"Distributed Database System - GeeksforGeeks", GeeksforGeeks, 2022. [Online]. Available: <https://www.geeksforgeeks.org/distributed-database-system/>. [Accessed: 19- Sep- 2022].

[2]P. home, c. support and </form>, "What Is a Distributed Database? {Features, Benefits & Drawbacks}", Knowledge Base by phoenixNAP, 2022. [Online]. Available: <https://phoenixnap.com/kb/distributed-database>. [Accessed: 19- Sep- 2022].

[3]"Distributed Databases Tutorial", Tutorialride.com, 2022. [Online]. Available: <https://www.tutorialride.com/distributed-databases/distributed-databases-tutorial.htm>. [Accessed: 19- Sep- 2022].

[4]"Distributed Database (Goals, Types, Advantages and Disadvantages) - DatabaseTown", DatabaseTown, 2022. [Online]. Available: <https://databasetown.com/distributed-database-goals-types-advantages-and-disadvantages/>. [Accessed: 19- Sep- 2022].

[5]"Distributed Databases", Docs.oracle.com, 2022. [Online]. Available: [https://docs.oracle.com/cd/A57673\\_01/DOC/server/doc/SCN73/ch21.htm](https://docs.oracle.com/cd/A57673_01/DOC/server/doc/SCN73/ch21.htm). [Accessed: 19- Sep- 2022].



GUIA DE OBSERVACIÓN PARA RESOLUCIÓN DE EJERCICIOS PRACTICOS

NOMBRE DE LA ASIGNATURA: **TOPICOS DE BASE DE DATOS**

NOMBRE DE LA UNIDAD: **SISTEMAS DE BASE DE DATOS DISTRIBUIDAS**

ALUMNO: **MEZO BUSTAMANTE RICARDO**

**INSTRUCCIONES**

Revisar los documentos o actividades que se solicitan y marque en los apartados “SI” cuando la evidencia a evaluar se cumple; en caso contrario marque “NO”. En la columna “OBSERVACIONES” ocúpela cuando tenga que hacer comentarios referentes a lo observado.

<b>Valor del reactivo</b>	<b>Características a cumplir (Reactivo)</b>	<b>CUMPLE</b>		<b>OBSERVACIONES</b>
		<b>Si</b>	<b>NO</b>	
<b>8%</b>	¿Identifico el problema planteado?	X		
<b>4%</b>	¿Identifico los datos de entrada del problema?	X		
<b>4%</b>	¿Identifico los datos de salida del problema?	X		
<b>8%</b>	¿Generó la solución del problema en forma clara y comprensible (orden)?	X		
<b>12%</b>	¿Elaboró el programa respetando la sintaxis del lenguaje de programación (orden)?	X		
<b>4%</b>	Comprobó los resultados esperados a través de pruebas de escritorio?	X		
<b>40%</b>	<b>CALIFICACIÓN:</b>			<b>40%</b>



TECNOLÓGICO  
NACIONAL DE MÉXICO

**TECNOLÓGICO NACIONAL DE MÉXICO**



**Instituto Tecnológico Superior de San Andrés Tuxtla**

**Ingeniería Informática**

**Asignatura: Tópicos de base de datos**

**Semestre: Séptimo Semestre**

**Grupo: IINF710A**

**Docente: MTJ. Lorenzo de Jesus Organista Oliveros**

**Alumno:**

- Ricardo Mezo Bustamante
- Roberto Carlos Delgado Ferman
- Oziel Alberto Morales Pot
- Ramon de Jesús Muñoz Chigüil
- Mario Ángel López Sánchez

**-Presenta-**

**Configuración de Base de Datos Distribuidas**

**Período Escolar: Agosto-Diciembre 2022**

**San Andrés Tuxtla Ver, 26 de Septiembre de 2022**

## Introducción

Las bases de datos son la piedra angular de cualquier aplicación de software. Necesitará una o más bases de datos para desarrollar casi todo tipo de aplicaciones de software: Web, Enterprise, Embedded Systems, Real-Time Systems, AI, ML, HPC, Blockchain, IoT y muchas otras aplicaciones.

Con el auge de los Microservicios, la Nube, las Aplicaciones Distribuidas, el Escalamiento Global, los Datos Semi-Estructurados, los Big Data, los Datos Rápidos, los Datos de Baja Latencia: a las bases de datos SQL tradicionales ahora se unen varios NoSQL . NewSQL y bases de datos en la nube .

Actualmente el uso de las tecnologías es vital en el ámbito empresarial y el laboral esto debido al uso de las cantidades masivas de información entre esas, información de la empresa, cliente, empleados, etc. Las empresas utilizan los datos almacenados en bases de datos para tomar decisiones comerciales informadas.

Para el manejo de esta información se utilizan las bases de datos ya que estas ayudan a la administración, difusión, edición y consulta de cierta información en el momento que uno o varios ya sean clientes o administradores necesiten en ese momento. Una base de datos es información que se configura para un fácil acceso, administración y actualización. Las bases de datos informáticas suelen almacenar conjuntos de registros de datos o archivos que contienen información, como transacciones de ventas, datos de clientes, información financiera y de productos.

Las bases de datos se utilizan para almacenar, mantener y acceder a cualquier tipo de datos. Recogen información sobre personas, lugares o cosas. Esa información se reúne en un solo lugar para que pueda ser observada y analizada. Las bases de datos se pueden considerar como una colección organizada de información.

Una base de datos está conformada por tablas, consultas, informes, formularios y la información en ellas, hay diferentes tipos de bases de datos, pero nos centramos en la base de datos distribuida ya que esta está conformada de varios equipos que se denominan nodos, y estos se comunican entre si mediante una red de comunicaciones.

Este tipo de base de datos se caracteriza por su disponibilidad a la hora de consultar la información además de que permite que cada equipo puede trabajar de manera autónoma y de manera local por lo que la información no tiene ningún grado de deterioro. Las bases de datos distribuidas tienen un grado de escalabilidad ya que un gran rendimiento y tiempo de respuesta ya que para un cliente es importante que la información este en tiempo y forma.

## Fragmentación

La fragmentación es una característica de servidor de bases de datos que le permite controlar dónde se almacenan los datos a nivel de tabla. La fragmentación le permite definir grupos de filas o claves de índice dentro de una tabla, según algún algoritmo o esquema. Puede almacenar cada grupo o fragmento (también conocido como partición) en un espacio de base de datos independiente asociado con un disco físico específico. Se utilizan sentencias SQL para crear los fragmentos y asignarlos a los espacios de base de datos.

Para acceder a una porción de datos en un SABDD, un programa busca y recupera el dato que necesita. Dicho dato no se mueve por todo el sistema, sino que reside en una localidad. Los procesos que se ejecutan utilizan operaciones algebraicas para determinar su ubicación y recuperarlo.

Existen tres tipos de fragmentación la horizontal, la vertical y la mixta

### 1.-Fragmentación Horizontal

Al utilizar una fragmentación horizontal esta es una división lineal de la tabla es decir una tabla pasa ser dos o más tablas.

Ejemplo: Una tabla de alumnos completa.

DNI	Escuela	Nombre	Nota ingreso	Beca
87633483	EUI	Concha Queta	5.6	No
99855743	EUI	Josechu Letón	7.2	Si
33887293	EUIT	Oscar Romato	6.1	Si
05399075	EUI	Bill Gates	5.0	No
44343234	EUIT	Pepe Pótamo	8.0	No
44543324	EUI	Maite Clado	7.5	Si
66553234	EUIT	Ernesto Mate	6.6	No

Al utilizar la fragmentación horizontal:

DNI	Escuela	Nombre	Nota ingreso	Beca
87633483	EUI	Concha Queta	5.6	No
99855743	EUI	Josechu Letón	7.2	Si
05399075	EUI	Bill Gates	5.0	No
44543324	EUI	Maite Clado	7.5	Si

Esta divide la tabla en secciones horizontales en este caso partiendo la tabla completa en dos tablas:

DNI	Escuela	Nombre	Nota ingreso	Beca
33887293	EUIT	Oscar Romato	6.1	Si
44343234	EUIT	Pepe Pótamo	8.0	No
66553234	EUIT	Ernesto Mate	6.6	No

## 2.-Fragmentación Vertical.

Al utilizar una fragmentación Vertical estamos dividiendo una tabla hacia abajo junto con los datos en ella. Ejemplo:

Tenemos una tabla con datos los datos de la table son escuela, situación y número de alumnos:

Escuela	Situación	Número alumnos
EUI	Campus sur	3000
EUIT	Campus sur	2800
TOPOGRAFIA	Campus sur	800
ETSIT	Ciudad Universitaria	2500
FI	Campus Montegancedo	2100

Al utilizar una fragmentación vertical divido los datos en tablas individuales. Ahora tenemos una tabla con la escuela y situación:

Escuela	Situación
EUI	Campus sur
EUIT	Campus sur
TOPOGRAFIA	Campus sur
ETSIT	Ciudad Universitaria
FI	Campus Montegancedo

Y una con escuela y número de alumnos:

Escuela	Número alumnos
EUI	3000
EUIT	2800
TOPOGRAFIA	800
ETSIT	2500
FI	2100

### 3.-Fragmentación Mixta.

Como el mismo nombre indica es una combinación de las dos anteriores vistas. Un ejemplo a partir de una tabla fragmentada horizontalmente:

DNI	Escuela	Nombre	Nota ingreso	Beca
87633483	EUI	Concha Queta	5.6	No
99855743	EUI	Josechu Letón	7.2	Si
05399075	EUI	Bill Gates	5.0	No
44543324	EUI	Maite Clado	7.5	Si

Al fragmentarla de forma mixta se quita la tabla nota:

DNI	Escuela	Nombre	Beca
87633483	EUI	Concha Queta	No
99855743	EUI	Josechu Letón	Si
05399075	EUI	Bill Gates	No
44543324	EUI	Maite Clado	Si

Y después:

DNI	Escuela	Nombre	Nota ingreso
87633483	EUI	Concha Queta	5.6
99855743	EUI	Josechu Letón	7.2
05399075	EUI	Bill Gates	5.0
44543324	EUI	Maite Clado	7.5

## Desarrollo de la Practica

Una de las sociedades médicas más importantes del país, con oficina central en Madrid, dispone de tres grandes centros médicos en esta misma ciudad. En cada uno de ellos se atienden determinadas especialidades médicas por el personal contratado por dicha sociedad, de manera que una especialidad pueda darse en un centro o varios y todos tienen al menos una especialidad.

En la actualidad tienen una base de datos relacional centralizada, a la cual acceden todos los centros, con el esquema siguiente.

Realice el diseño de los esquemas de fragmentación y asignación de base de Datos Distribuida para conseguir la mayor autonomía local, tanto en los centros médicos como en la central, sabiendo que las operaciones realizadas en cada uno de ellos son:

- En la sede central (Cod\_Centro = 01) se elaboran la nominas del personal contratados por la Sociedad Medica y de aquí se envían a los Centros para que sean repartidas a los trabajadores.

- Cada Centro medico (Cod\_Centro = {02,03,04}) debe tener la relación de personal sanitario que trabaja en el mismo, así como el horario de la consulta y la especialidad que desarrolla.

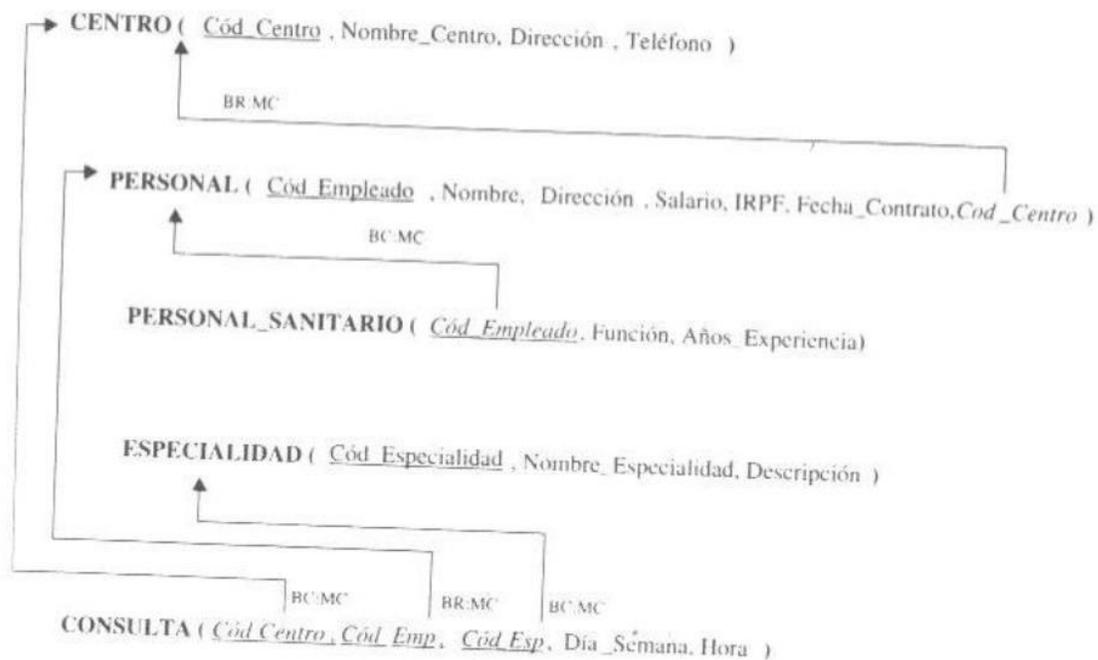


Figura 4.7. Esquema relacional correspondiente a una BD centralizada de Sociedades Médicas

### 1.- Verificación de Red (IP)

Se hizo uso del sistema operativo Windows 10, con el servidor MySQL server en su última versión, en esta parte se verificó en cada máquina la IP asignada mediante el comando "ipconfig", este es útil para obtener una visión rápida y concisa de los ajustes de configuración principales en la red TCP/IP y cubre todos los adaptadores e interfaces de red existentes (por ejemplo, LAN, WLAN, Bluetooth, RDSI y adaptadores VPN).

Maquina Ricardo:

```

    Símbolo del sistema

    Sufijo DNS específico para la conexión. . . : home.sercomm
    Vínculo: dirección IPv6 local. . . . . : fe80::4504:d31b:3e5c:555e%6
    Dirección IPv4. . . . . : 202.168.0.69
    Máscara de subred . . . . . : 255.255.255.0
    Puerta de enlace predeterminada . . . . . : 202.168.0.1
  
```

## Maquina Roberto:

```
C:\Windows\system32\cmd.exe

Adaptador de Ethernet Ethernet:

Sufijo DNS específico para la conexión. . . : home.sercomm
Descripción . . . . . : Realtek PCIe FE Family Controller
Dirección física. . . . . : 40-B0-34-13-97-B7
DHCP habilitado . . . . . : sí
Configuración automática habilitada . . . : sí
Vínculo: dirección IPv6 local. . . . . : fe80::8186:59fc:ee7d:8436%11(Preferido)
Dirección IPv4. . . . . : 202.168.0.67(Preferido)
Máscara de subred . . . . . : 255.255.255.0
Concesión obtenida. . . . . : domingo, 25 de septiembre de 2022 03:40:02 p. m.
La concesión expira . . . . . : lunes, 26 de septiembre de 2022 03:40:01 p. m.
Puerta de enlace predeterminada . . . . . : 202.168.0.1
Servidor DHCP . . . . . : 202.168.0.1
IAID DHCPv6 . . . . . : 188788788
DUID de cliente DHCPv6. . . . . : 00-01-00-01-2A-83-60-A3-40-B0-34-13-97-B7
Servidores DNS. . . . . : 202.168.0.1
NetBIOS sobre TCP/IP. . . . . : habilitado
```

## Maquina Oziel:

```
C:\WINDOWS\system32\CMD.exe

Sufijo DNS específico para la conexión. . . : home.sercomm
Vínculo: dirección IPv6 local. . . . . : fe80::bd7a:f108:f7b6:9021%19
Dirección IPv4. . . . . : 202.168.0.64
Máscara de subred . . . . . : 255.255.255.0
Puerta de enlace predeterminada . . . . . : 202.168.0.1
```

## Maquina Ramón de Jesús:

```
Adaptador de Ethernet Ethernet:

Sufijo DNS específico para la conexión. . . : home.sercomm
Vínculo: dirección IPv6 local. . . . . : fe80::ec3c:1e81:80ce:fec4%21
Dirección IPv4. . . . . : 202.168.0.65
Máscara de subred . . . . . : 255.255.255.0
Puerta de enlace predeterminada . . . . . : 202.168.0.1
```

## 2.- Comprobación de Red

Para comprobar que todos los equipos estaban en red se ocupó el comando “ping”, este se usa para determinar el estado de un host remoto. Al ejecutar el comando ping, el protocolo ICMP envía al host un determinado datagrama para solicitar una respuesta. El protocolo ICMP se ocupa de los errores en las redes TCP/IP.

```
Símbolo del sistema

C:\Users\RMSystems>ping 202.168.0.64

Haciendo ping a 202.168.0.64 con 32 bytes de datos:
Respuesta desde 202.168.0.64: bytes=32 tiempo=2ms TTL=128
Respuesta desde 202.168.0.64: bytes=32 tiempo=2ms TTL=128
Respuesta desde 202.168.0.64: bytes=32 tiempo=1ms TTL=128

Estadísticas de ping para 202.168.0.64:
    Paquetes: enviados = 4, recibidos = 4, perdidos = 0
            (0% perdidos),
    Tiempos aproximados de ida y vuelta en milisegundos:
        Mínimo = 1ms, Máximo = 2ms, Media = 1ms

C:\Users\RMSystems>ping 202.168.0.67

Haciendo ping a 202.168.0.67 con 32 bytes de datos:
Respuesta desde 202.168.0.67: bytes=32 tiempo=1ms TTL=128
Respuesta desde 202.168.0.67: bytes=32 tiempo=1ms TTL=128
Respuesta desde 202.168.0.67: bytes=32 tiempo=1ms TTL=128

Estadísticas de ping para 202.168.0.67:
    Paquetes: enviados = 4, recibidos = 4, perdidos = 0
            (0% perdidos),
    Tiempos aproximados de ida y vuelta en milisegundos:
        Mínimo = 1ms, Máximo = 1ms, Media = 1ms

C:\Users\RMSystems>ping 202.168.0.65

Haciendo ping a 202.168.0.65 con 32 bytes de datos:
Respuesta desde 202.168.0.65: bytes=32 tiempo=1ms TTL=128
Respuesta desde 202.168.0.65: bytes=32 tiempo=1ms TTL=128
Respuesta desde 202.168.0.65: bytes=32 tiempo=1ms TTL=128

Estadísticas de ping para 202.168.0.65:
    Paquetes: enviados = 4, recibidos = 4, perdidos = 0
            (0% perdidos),
    Tiempos aproximados de ida y vuelta en milisegundos:
        Mínimo = 1ms, Máximo = 1ms, Media = 1ms

C:\Users\RMSystems>
```





## Tabla nomina personal:

centro - Table   consulta - Table   especialidad - Table   inf\_personal - Table   **nom\_personal - Table** x

Table Name:  Schema: **soc\_medicas**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
cod_empleado	VARCHAR(45)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
salario	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
IRPF	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
fecha_contrato	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						

## Tabla personal sanitario:

centro - Table   consulta - Table   especialidad - Table   inf\_personal - Table   nom\_personal - Table   **personal\_sanitario - Table** x

Table Name:  Schema: **soc\_medicas**

Charset/Collation:   Engine:

Comments:

Column Name	Datatype	PK	NN	UQ	B	UN	ZF	AI	G	Default/Expression
cod_empleado	VARCHAR(45)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
funcion	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						
años_experiencia	VARCHAR(45)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>						

## 6.- Carga de registros en las tablas:

Usando sentencias SQL se cargaron los datos en cada una de las tablas.

```

C:\Windows\System32\cmd.exe - mysql -u root -p
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.
MariaDB [(none)]> use soc_medicas;
Database changed
MariaDB [soc_medicas]> INSERT INTO `centro` (`cod_centro`, `nombre_centro`, `direccion`, `telefono`) VALUES('01', 'CENTRO', 'XALAPA', '5348329087'),('02', 'EL VALLE', 'MARTINEZ', '4398235671'),('03', 'ALEMAN', 'MISANTLA', '6254317832'),('04', 'MISANTLA', 'MISANTLA', '5438012865');
Query OK, 4 rows affected (0.058 sec)
Records: 4 Duplicates: 0 Warnings: 0
MariaDB [soc_medicas]> INSERT INTO `inf_personal` (`cod_empleado`, `nombre`, `direccion`, `cod_centro`) VALUES('CE1', 'Tomas Tuna', 'Xalapa', '01'),('CE2', 'Ana Alto', 'Martinez', '01'),('CE3', 'Pedro Peña', 'Misantla', '01'),('CE4', 'Bruno Brisa', 'Santa Rita', '01');
Query OK, 4 rows affected (0.079 sec)
Records: 4 Duplicates: 0 Warnings: 0
MariaDB [soc_medicas]> INSERT INTO `nom_personal` (`cod_empleado`, `salario`, `IRPF`, `fecha_contrato`) VALUES('CE1', '7500', '10.1%', '12/04/2000'),('CE10', '5900', '10.1%', '15/07/2012'),('CE11', '7000', '10.1%', '01/01/2012'),('CE12', '1500', '10.1%', '30/10/2001'),('CE13', '4550', '10.1%', '30-02-2012'),('CE14', '7500', '10.1%', '31/08/2005'),('CE15', '4000', '10.1%', '09/09/2007'),('CE16', '6000', '10.1%', '29/11/2008'),('CE2', '5000', '10.1%', '01/01/2011'),('CE3', '4900', '10.1%', '15/08/2011'),('CE4', '5000', '10.1%', '30/12/2010'),('CE5', '6000', '10.1%', '14/05/2009'),('CE6', '6500', '10.1%', '20/11/2009'),('CE7', '6900', '10.1%', '26/09/2000'),('CE8', '4900', '10.1%', '26/09/2000'),('CE9', '5500', '10.1%', '10/10/2010');
Query OK, 16 rows affected (0.096 sec)
Records: 16 Duplicates: 0 Warnings: 0
MariaDB [soc_medicas]> INSERT INTO `personal_sanitario` (`cod_empleado`, `funcion`, `años_experiencia`) VALUES('CE1', 'Enfermera', '12'),('CE2', 'Recepcionista', '5'),('CE3', 'Enfermero Auxiliar', '4'),('CE4', 'Medico Especialista', '12');
Query OK, 4 rows affected (0.145 sec)
Records: 4 Duplicates: 0 Warnings: 0
MariaDB [soc_medicas]> INSERT INTO `consulta` (`cod_centro`, `cod_empleado`, `cod_especialidad`, `dia_semana`, `hora`) VALUES('01', 'CE1', 'C01', 'lunes', '9:00'),('01', 'CE2', 'C01', 'martes', '10:00'),('01', 'CE3', 'C01', 'miércoles', '11:00'),('01', 'CE4', 'C01', 'jueves', '12:00');
Query OK, 4 rows affected (0.119 sec)
Records: 4 Duplicates: 0 Warnings: 0
MariaDB [soc_medicas]> INSERT INTO `especialidad` (`cod_especialidad`, `nombre_especialidad`, `descripcion`) VALUES('C01', 'Traumatología', 'Traumatología');
Query OK, 1 row affected (0.058 sec)
MariaDB [soc_medicas]> _
    
```

## 7.- Configuraciones en el Servidor

Se usaron instrucciones SQL para configurar mediante la terminal de MySQL server:

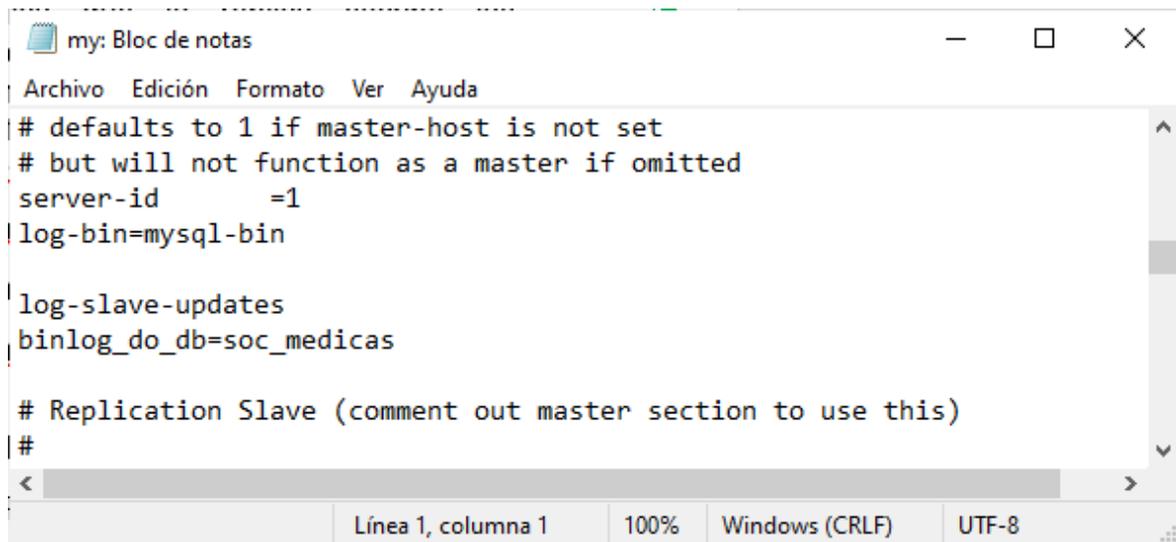
-Cambiar ID de servidor en cada maquina: Esto se hizo en el archivo de configuraciones de mysql llamado "my.ini", las instrucciones utilizadas son las siguientes:

"server-id=tu\_numero\_aqui": Este ID de servidor se utiliza para identificar servidores individuales dentro de la topología de replicación y debe ser un número entero positivo entre 1 y (232)-1. El valor predeterminado de server\_id de MySQL 8.0 es 1. Puede cambiar el valor de server\_id dinámicamente emitiendo una declaración como esta:

"log-bin=mysql-bin": Es el nombre base para la secuencia de registro. El servidor crea archivos de registro binarios en secuencia agregando un sufijo numérico al nombre base. Si no proporciona la opción --log-bin, MySQL usa binlog como el nombre base predeterminado para los archivos de registro binarios.

"log-slave-updates": en una réplica significa que la réplica escribe las actualizaciones que se reciben de un maestro en sus propios registros binarios. Entonces, una réplica puede convertirse en un maestro intermedio de otra réplica. Esta opción es un trampolín para varias configuraciones de replicación de nivel simple a complejo y nuevas tecnologías como Group Replication.

"binlog\_do\_db"=nombre\_de\_base\_de\_datos\_a\_replicar

A screenshot of a Notepad window titled "my: Bloc de notas". The window contains the following MySQL configuration text:

```
# defaults to 1 if master-host is not set
# but will not function as a master if omitted
server-id      =1
log-bin=mysql-bin

log-slave-updates
binlog_do_db=soc_medicas

# Replication Slave (comment out master section to use this)
#
```

The status bar at the bottom of the window shows "Línea 1, columna 1", "100%", "Windows (CRLF)", and "UTF-8".

## 8.- Creamos Un Usuario Remoto En Cada Maquina

Esto se realizó en la consola del servidor de mysql, en cada una de las maquinas con su respectivo usuario.

```
CREATE USER 'nombre_usuario'@'%' IDENTIFIED BY 'tu_contraseña';
```

```
GRANT SELECT, REPLICATION SLAVE, REPLICATION CLIENT ON *.* to  
nombre_usuario@'ip_remota' IDENTIFIED BY 'tu_contraseña';
```

#### 8.- Asignamos Permisos

Una vez aplicada las configuraciones anteriores se asignan los permisos necesarios para que pueda hacer la conexión.

```
FLUSH PRIVILEGES;
```

```
SHOW MASTER STATUS;
```

Posteriormente procedemos a detener el proceso del servidor para configurar las direcciones IP de cada usuario en cada máquina.

```
STOP SLAVE;
```

#### 9.- Configuramos Conexiones

El siguiente código se deberá aplicar a cada uno de los equipos, esto apuntando al equipo siguiente, ejemplo el equipo 2 con los datos de equipo 3, el 3 al 4, y el 4 al 1.

```
CHANGE MASTER TO
```

```
MASTER_HOST='ip_host_maestro',
```

```
MASTER_PORT=puerto_servidor,
```

```
MASTER_USER='nombre_usuario',
```

```
MASTER_PASSWORD='tu_contraseña',
```

```
MASTER_LOG_FILE='mysql-bin.000001',
```

```
MASTER_LOG_POS= 837;
```

Una vez cargada la configuración aplicamos al servidor e iniciamos servicio nuevamente.

```
START SLAVE;
```

# Aplicación Desarrollada en PHP

Se desarrollo una aplicación en lenguaje PHP para mostrar los registros:

Sociedades Médicas:



Menu

- Consultar Centros
- Consultas Vigentes
- Especialidades
- Personal en Turno
- Personal Sanitario

Consulta de Centros:



Codigo	Nombre	Direccion	Telefono
O1	CENTRO	XALAPA	5348329087
O2	EL VALLE	MARTINEZ	4390235671
O3	ALEMAN	MISANTLA	6254317832
O4	MISANTLA	MISANTLA	5438012865

Consulta de Vigencias:



Codigo de Centro	Codigo de Empleado	Codigo de Especialidad	Dia de Semana	Hora
O1	CE1	CO1	lunes	9:00
O3	CE10	CO2	viernes	13:00
O3	CE11	CO3	lunes	09:00
O3	CE12	CO3	martes	10:00
O4	CE13	CO3	miércoles	11:00
O4	CE14	CO3	jueves	12:00
O4	CE15	CO3	viernes	13:00
O4	CE16	CO4	lunes	09:00
O1	CE2	CO1	martes	10:00
O1	CE3	CO1	miércoles	11:00
O1	CE4	CO1	jueves	12:00
O2	CE5	traumatologia	viernes	13:00
O2	CE6	Pediatria	lunes	09:00
O2	CE7	Pediatria	martes	10:00
O2	CE8	Pediatria	miércoles	11:00
O3	CE9	CO2	jueves	12:00

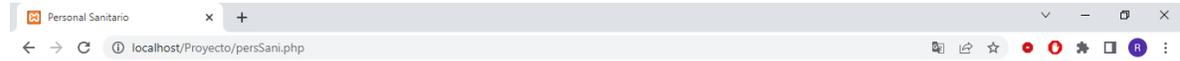
## Consulta de Especialidades:



### Especialidades

Codigo de Especialidad	Nombre de Especialidad	Descripcion
CO1	Traumatologia	Traumatologia
CO2	Pediatria	
CO3	Ginecologia	Ginecologia
CO4	Neurologia	Neurologia

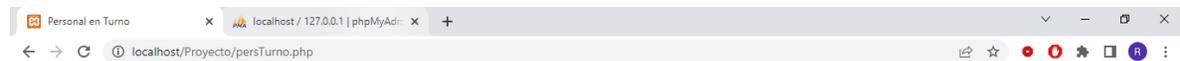
## Consulta de Personal Sanitario:



### Personal Sanitario

Codigo de Empleado	Funcion	Años de Experiencia
CE1	Enfermera	12
CE10	Enfermero Auxiliar	1
CE11	Farmacéutico	8
CE12	Doctor	11
CE2	Recepcionista	5
CE3	Enfermero Auxiliar	4
CE4	Medico Especialista	12
CE5	Medico Especialista	3
CE6	Técnico Administrativo	15
CE7	Director general	14
CE8	Auxiliar Administrativo	9
CE9	Medico Especialista	4

## Consulta de Personal en Turno:



### Personal en Turno

Codigo de Empleado	Nombre	Direccion	Salario	IRPF	Fecha de Contrato	Codigo de Centro
CE1	Tomasa Tuna	Xalapa	7500	10.1%	12/04/2000	O1
CE10	Sonia Salas	Los Idolos	5900	10.1%	15/07/2012	O3
CE11	José Juan	Misantla	7000	10.1%	01/01/2012	O3
CE12	Fito Fosa	El Porvenir	1500	10.1%	30/10/2001	O3
CE13	Olga Olmo	Xalapa	4550	10.1%	30-02-2012	O4
CE14	Heri Humo	Xalapa	7500	10.1%	31/08/2005	O4
CE15	Victor Vidal	Xalapa	4000	10.1%	09/09/2007	O4
CE16	Gina Gómez	Atzalan	6000	10.1%	29/11/2008	O4
CE2	Ana Alto	Martinez	5000	10.1%	01/01/2011	O1
CE3	Pedro Peña	Misantla	4900	10.1%	15/08/2011	O1
CE4	Bruno Brisa	Santa Rita	5000	10.1%	30/12/2010	O1
CE5	Caty Cortez	Tenochtitlan	6000	10.1%	14/05/2009	O2
CE6	Damián Diaz	Guerrero	6500	10.1%	20/11/2009	O2
CE7	Elsa Elvira	Misantla	6900	10.1%	26/09/2000	O2
CE8	Marcos Mina	Vega	4900	10.1%	26/09/2000	O2
CE9	Yair Yunes	Pozón	5500	10.1%	10/10/2010	O3

# Pruebas de Conexión

## Maquina Ricardo

The screenshot shows the phpMyAdmin interface for a MySQL server. The main content area displays the replication status, indicating that the server is configured as a replica. A table lists the replication variables and their values.

El servidor está configurado como réplica en un proceso de replicación. Deseas:

- Ver la tabla de estado de réplica

Variable	Valor
Slave_IO_State	Waiting for master to send event
Master_Host	202.168.0.67
Master_User	roberto
Master_Port	3306
Connect_Retry	60
Master_Log_File	mysql-bin.000003
Read_Master_Log_Pos	8189
Relay_Log_File	mysql-relay-bin.000002
Relay_Log_Pos	555
Relay_Master_Log_File	mysql-bin.000003
Slave_IO_Running	Yes
Slave_SQL_Running	Yes
Replicate_Do_DB	
Replicate_Ignore_DB	

## Maquina Roberto

The screenshot shows the phpMyAdmin interface for a MySQL server. The main content area displays the replication status, indicating that the server is configured as a replica. A table lists the replication variables and their values.

Slave_IO_State	Waiting for master to send event
Master_Host	202.168.0.64
Master_User	oriel
Master_Port	3306
Connect_Retry	60
Master_Log_File	mysql-bin.000003
Read_Master_Log_Pos	3433
Relay_Log_File	mysql-relay-bin.000009
Relay_Log_Pos	3356
Relay_Master_Log_File	mysql-bin.000003
Slave_IO_Running	Yes
Slave_SQL_Running	Yes
Replicate_Do_DB	
Replicate_Ignore_DB	
Replicate_Do_Table	
Replicate_Ignore_Table	
Replicate_Wild_Do_Table	

## Maquina Oziel:

The screenshot shows the phpMyAdmin interface for a MySQL server. The left sidebar displays a tree view of databases, including 'soc\_medicas' and 'test'. The main content area shows the 'Estado actual' (Current Status) tab for a replication setup. The table below lists various replication parameters and their values.

Slave_IO_State	Waiting for master to send event
Master_Host	202.168.0.65
Master_User	Jesus
Master_Port	3306
Connect_Retry	60
Master_Log_File	mysql-bin.000004
Read_Master_Log_Pos	664
Relay_Log_File	mysql-relay-bin.000002
Relay_Log_Pos	555
Relay_Master_Log_File	mysql-bin.000004
Slave_IO_Running	Yes
Slave_SQL_Running	Yes
Replicate_Do_DB	
Replicate_Ignore_DB	
Replicate_Do_Table	
Replicate_Ignore_Table	

## Maquina Ramón de Jesús

The screenshot shows the phpMyAdmin interface for a MySQL server. The left sidebar displays a tree view of databases, including 'soc\_medicas' and 'test'. The main content area shows the 'Estado actual' (Current Status) tab for a replication setup. The table below lists various replication parameters and their values.

Slave_IO_State	Waiting for master to send event
Master_Host	202.168.0.69
Master_User	ricardo
Master_Port	3306
Connect_Retry	60
Master_Log_File	mysql-bin.000005
Read_Master_Log_Pos	1019
Relay_Log_File	mysql-relay-bin.000002
Relay_Log_Pos	555
Relay_Master_Log_File	mysql-bin.000005
Slave_IO_Running	Yes
Slave_SQL_Running	Yes
Replicate_Do_DB	
Replicate_Ignore_DB	
Replicate_Do_Table	
Replicate_Ignore_Table	

## Conclusión

Una base de datos distribuida (DDB) es una colección integrada de bases de datos que se distribuye físicamente a través de sitios en una red informática. Un sistema de gestión de bases de datos distribuidas (DDBMS) es el sistema de software que gestiona una base de datos distribuida de forma que los aspectos de distribución sean transparentes para los usuarios. Para formar un sistema de base de datos distribuida (DDBS), los archivos deben estar estructurados, lógicamente interrelacionados y físicamente distribuidos en varios sitios. Además, debe haber una interfaz común para acceder a los datos distribuidos.

Se puede configurar una base de datos de valores clave distribuida para almacenar los mismos datos en múltiples nodos en todas las ubicaciones. Si un solo nodo falla, los datos aún están disponibles. No tiene que esperar a que se restablezca la base de datos. Una base de datos distribuida geográficamente mantiene nodos simultáneos en regiones geográficas para brindar resiliencia en caso de un corte de energía o comunicaciones regional. La capacidad de almacenar una sola base de datos en varias computadoras requiere un algoritmo para replicar datos que sea transparente para los usuarios.

Las ventajas de las bases de datos distribuidas incluyen muchos beneficios, pero los principales objetivos de los sistemas de bases de datos distribuidas son la escalabilidad, el rendimiento de las consultas (velocidad) y la disponibilidad.

El uso de las bases de datos es de vital importancia y entender su funcionamiento, así como saber usarlas puede hacer que el inicio de un pequeño proyecto pueda terminar en una gran y exitosa empresa, capaz de darle a sus clientes lo que buscan, un sistema confiable, eficaz, rápido, fácil de entender y usar.

Al decir bases de datos entendemos que nos referíamos a un conjunto de datos con sistemas que están integrados y basados en medios que contienen información. Y que nos va a facilitar la forma de trabajar con datos personales, que son evidentemente trabajados en las empresas.

Las bases de datos nos permiten llevar un buen control y una buena organización dentro de la empresa, así mismo mantener una buena economía con la recolección de datos.

Esto nos permitirá en un futuro que la empresa en la que trabajemos o en nuestra empresa crezca de manera óptima, evitando la pérdida de datos, empleando las mejores tecnologías, para la protección de la información, de nuestro cliente, nuestros usuarios y nuestros proveedores.

## Referencias Bibliográficas

Ceri S. y Pelagatti G. Bases de datos distribuidas: principios y sistemas. McGraw-Hill, 1984.

Bell D. y Grimson J. Sistemas de bases de datos distribuidas. Addison-Wesley, 1992.

Halevy AY, Ives ZG, Suciu D. y Tatarinov I. Mediación de esquemas para compartir datos semánticos a gran escala. VLDB J., 14(1):68–83 2005.

Pentaris F. y Ioannidis YE Optimización de consultas en redes distribuidas de sistemas de bases de datos autónomos. ACM Trans. Sistema de base de datos, 31(2):537–583, 2006.

Selinger PG Una descripción general de la arquitectura de R\*: un sistema de gestión de bases de datos distribuidas. En Proc. Quinto Taller de Berkeley sobre Gestión de Datos Distribuidos y Redes Informáticas, 1981, p. 187.

Título

## Evaluación - Unidad I

---

Instrucciones (opcional)

- Descargar el archivo y responder lo que se pide.
- Subir el archivo final resultante [EvaluaciónUI-Iniciales.PDF]

-Valor de la Evaluación 40%



Examen\_UI.doc

Word

Los alumnos pueden ver el archivo



Adjuntar



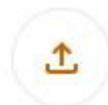
Drive



YouTube



Crear



Subir



Enlace



Nombre: Ricardo Mezo Bustamante

Calif: \_\_\_\_\_

**I. Subraya la respuesta correcta. (40%)**

1. Los datos pueden estar almacenados en la ubicación donde son usados más frecuentemente para que la mayoría de las operaciones sean locales y se reduzca el tráfico en la red.

- a) Independencia de Replicación      b) Independencia de Ubicación      c) **Independencia de Fragmentación**

2. El sistema sería vulnerable; es decir, si el sitio central falla, también fallará todo el sistema.

- a) Independencia de Falla      b) **No dependencia de un sitio central**      c) Independencia de Fragmentación

3. Soporte para un gran número de máquinas diferentes.

- a) Independencia de Replicación      b) **Independencia de HW**      c) Independencia de S.O

4. El punto básico es que en una consulta que involucra a varios sitios, habrá muchas formas posibles de mover los datos en el sistema para satisfacer la solicitud.

- a) Redundancia de datos      b) Independencia de Ubicación      c) **Procesamiento de consultas distribuidas**

5. Los usuarios no tienen que saber dónde están almacenados físicamente los datos.

- a) No dependencia de un sitio central      b) Independencia de almacenamiento      c) **Independencia de ubicación**

6. Pueden continuar operando cuando hay alguna falla en algún componente independiente.

- a) Independencia de falla      b) **Operación continúa**      c) Confiabilidad

7. Es necesario tener la posibilidad de soportar también una variedad de redes de comunicación distintas.

- a) Independencia de comunicación      b) **Independencia de red**      c) Independencia de S.O

8. Define Base de Datos Distribuidas:.

R: Es un conjunto de datos compartidos lógicamente interconectados y sus descripciones, que se distribuyen físicamente en alguna red informática.