



Universidad Nacional de La Matanza
Departamento de Ingeniería e Investigaciones Tecnológicas

Bases de Datos Aplicada

Grupo: 2

TP Integrador – Entrega nro: 3

Fecha de Presentación: [10/10/2023]

Integrantes:

DNI	Apellido	Nombre	Email
43316372	Castellani	Santiago	scastellani@alumno.unlam.edu.ar
43863195	Colantonio	Bruno	bcolantonio@alumno.unlam.edu.ar
44183677	Coulleri	Flavio	fcoulleri@alumno.unlam.edu.ar
41893248	Kowalski	Franco	fkowalski@alumno.unlam.edu.ar

Contenido

Unidad 1	2
Introducción	2
Desarrollo	2
Conclusiones	2
Bibliografía	2
Glosario	2
Unidad 2	3
Introducción	3
Desarrollo	3
Conclusiones	3
Bibliografía	3
Glosario	3
Unidad 3	4
Introducción	4
Desarrollo	4
Conclusiones	4
Bibliografía	4
Glosario	4
Unidad 4	5
Introducción	5
Desarrollo	5
Conclusiones	5
Bibliografía	5
Unidad 5	6
Introducción	6
Desarrollo	6
Conclusiones	6
Bibliografía	6
Glosario	6
Unidad 6	7
Introducción	7
Desarrollo	7
Conclusiones	7

Bibliografía	7
Glosario	7

Unidad 1

Introducción

Investigación acerca de los requerimientos técnicos, costo y características de un sistema de bases de datos relacionales implementando MySQL para la firma “CURE S.A.”.

Desarrollo

Requerimientos técnicos.

- Softwares de base: el motor MySQL cuenta con soporte en los sistemas operativos: Linux x86 64-bit, Mac OS, Windows x86 64-bit.

Requerimientos de hardware:

- Mínimo: 2 CPU Cores, 2 GB RAM, subsistema de I/O de disco aplicable a una base de datos de escritura intensiva.
- Recomendado: 4 CPU Cores o más, 8 GB RAM o más, RAID10 o RAID 0+1 disk setup.

Costo y modo de licenciamiento.

- La firma no cuenta con personal capacitado en la organización que pueda desempeñarse como Database Administrator, por lo que debemos considerar contratar uno. El sueldo promedio es de \$347.371/mes (pesos argentinos) o \$991,96 (dólares estadounidenses al cambio oficial).

MySQL posee 3 versiones on-premise, es decir, el software y la gestión de este es responsabilidad de la organización, y los servidores se establecen localmente.

Estas versiones son: Standard Edition, Enterprise Edition y Cluster Carrier Grade Edition, aunque nos enfocaremos en las primeras 2 ediciones, ya que la última ofrece características que superan ampliamente las necesidades de CURE.SA, debido a que utiliza clusters.

La Standard Edition es la versión más económica, \$2148 dólares estadounidenses/año.

La Enterprise Edition \$5370 dólares estadounidenses/año, brinda muchas características que la hacen más favorable. Entre ellas:

- Encriptación automática de los archivos físicos de la base de datos, con descriptación a la de leerlos. Esto es beneficioso ya que al tratarse de una base de datos de un hospital, se manejan datos sensibles. Además, brinda herramientas de criptografía, como generación de claves, firma digital, entre otras características para cumplir con los requisitos regulatorios de seguridad informática.
- Soporte Técnico 24/7 de ingenieros de soporte en base de datos que nos guían en la creación, despliegue y manejo de la base. Esto nos ayuda a compensar la falta de un dba.
- Backups incrementales y parciales, con recuperación 'point-in-time', permitiendo hacer una recuperación en un momento específico de la línea de tiempo. Además compresión de los backups.
- Tiene módulos de sistema de autenticación para integrar con PAM o Windows Active Directory, esto ahorra tiempo ya que no debemos gestionar cada usuario y permite mejorar la seguridad de la infraestructura extendiendo las políticas de usuario que aplican los sistemas de gestión.
- Herramientas de producción.
- Incluye una aplicación web llamada "MySQL Enterprise Monitor" que permite monitorear todos los aspectos de la base de datos desde un entorno protegido y seguro.
- Modelado.
- Alta disponibilidad de los datos y una buena tolerancia a fallos.
- Firewall integrado con protección a tiempo real para ataques específicos a base de datos, como SQL Injection. Esto es beneficioso debido a que la seguridad es un punto muy importante por la manipulación de información sensible.

Conclusiones

Se recomienda que la firma "CURE S.A." implemente la versión Enterprise Edition de MySQL, debido a que los beneficios que esta brinda serán de gran relevancia debido a la sensibilidad de los datos que se gestionan, junto con el equipo de soporte que incluye, el cual será de gran relevancia a la hora de desarrollar el sistema y mantenerlo.

Bibliografía

<https://shop.oracle.com/apex/product?p1=MySQL>

Glosario

[completar en caso de ser necesario]

Unidad 2

Introducción

En este informe desarrollaremos el servicio en la nube de bases de datos relacionales. Las principales opciones que se presentan son Microsoft Azure, Amazon Web Services, Google Cloud. Analizaremos las ventajas y desventajas de cada servicio.

Desarrollo

Primero analizaremos **Amazon Web Services**, esta plataforma ofrece múltiples opciones configurables para obtener una base de datos acorde a nuestras necesidades. Para nuestras necesidades en esta firma "CURE S.A", optamos contratar el servicio "Amazon Aurora MySQL - Compatible", en este, especificamos las características de 1 nodo (máquina virtual), almacenamiento de 4gb, almacenamientos de copias de seguridad de 4 gb, familia de instancias "general purpose", el tipo de instancia db.t3.medium la cual incluye procesamiento virtual de 2 VCores y ram de 4gb, OnDemand. La estimación de este servicio es de **\$161 dólares americanos por mes**.

Las versiones disponibles son Standard y optimizado para Entrada/Salida. La versión Optimizada para E/S cobra un valor fijo por acceso al servidor de la nube, por lo cual no implica un costo acumulado si se realizan consultas intensamente, aunque el costo por la capacidad de almacenamiento es mayor a la versión estándar. Por otra parte, la versión estándar sí contempla un costo por acceso al servidor (Entrada/Salida) que, según la página de Amazon, implica un 25% del costo total del servicio brindado. Por lo tanto, si se van a realizar pocas operaciones de Entrada/Salida conviene la versión standard, de otra forma, la versión optimizada es superior, debido a que no solo no se facturan dichos accesos sino que la capacidad de almacenamiento para el sistema solicitado es considerablemente baja. Para almacenar 4GB (primeros dos años) conviene utilizar la versión optimizada, y se ahorrarían aproximadamente **\$70 dólares americanos por mes**.

Export date: **6/9/2023**

Language: **English**

Estimate title: **My Estimate**

Estimate URL: <https://calculator.aws/#/estimate?id=d45802b3fe9c4ef75660d44d501ed5b3f962749b>

Estimate summary

Upfront cost	Monthly cost	Total 12 months cost
0.00 USD	161.06 USD	1,932.72 USD
		Includes upfront cost

Detailed Estimate

Name	Group	Region	Upfront cost	Monthly cost
Amazon Aurora MySQL-Compatible	No group applied	South America (Sao Paulo)	0.00 USD	161.06 USD
Status: - Description: Config summary: Aurora MySQL Cluster Configuration Option (Aurora I/O-Optimized), Change records per statement (0.38), Nodes (1), Instance Type (db.t3.medium), Utilization (100 %Utilized/Month), Instance Family (General purpose), Pricing Model (OnDemand), Storage amount (4 GB), Additional backup storage (4 GB), Total Size of Backup Processed for Export (GB) (4 per month)				

Por otra parte, **Google Cloud** ofrece una configuración más sencilla y rápida. Entre ellas incluye 1 instancia, edición Enterprise de MySQL, localización del servidor en Brasil, tipo de instancia (db-standard-2) instancia de vCPUs 2 y ram de 7.5gb, almacenamiento de 10 gb(es el mínimo), backup 4gb. Su valor estimado es de un total de **\$112 dólares americanos**.

Google Cloud Pricing Calculator - Estimate

Cloud SQL for MySQL

DB-STANDARD-2

Edition: Enterprise

Number of instances: 1

Location: Sao Paulo

Total hours per month: 730.0

Commitment term: 1 Year

Instance type: db-standard-2	USD 111.01
------------------------------	------------

HDD Storage: 10.0 GiB	USD 1.35
-----------------------	----------

Backup: 4.0 GiB	USD 0.48
-----------------	----------

USD 112.84

Total Estimated Cost: USD 112.84 per 1 month

Estimate Currency

USD - US Dollar

Por último, el servicio de **Microsoft Azure** el cual dispone de un tipo de instancia Gen 5 la cual cuenta con 2 VCore, servidor en la región sur de Brasil, almacenamiento de 5 gb, redundancia LRS, almacenamiento de copia de seguridad de 4 gb, soporte Professional Direct el cual brinda servicios de asesoramiento, api de soporte técnico, su desventaja es que el tiempo de respuesta inicial puede ser de hasta 1 hora. A su vez, se cuenta con soporte de nivel "Standard", con un costo de **\$100 dólares americanos**, orientado a pequeñas y medianas empresas. Si se desea, podría incluirse la versión "Profesional Direct" del soporte, la cual cuesta **\$1000 dólares americanos** y está orientada a negocios con mayor intensidad en la utilización de Azure. Utilizando el soporte Standard, el costo es de **\$194 dólares americanos**.

Microsoft Azure Estimate

Your Estimate

Service category	Service type	Region	Description	Estimated monthly cost
Databases	Azure Database for MySQL	Brazil South	Single Server Deployment, General Purpose Tier, 1 Gen 5 (2 vCore), 1 year reserved, 5 GB Storage with ZRS redundancy, 4 GB Additional Backup storage - LRS redundancy	\$94,36
Support	Support Licensing Program	Support	Microsoft Customer Agreement (MCA)	\$100,00
	Billing Account	Billing Profile		
	Total			\$194,36

Disclaimer

All prices shown are in United States – Dollar (\$) USD. This is a summary estimate, not a quote. For up to date pricing information please
This estimate was created at 9/6/2023 8:59:18 PM UTC.

Ante la posibilidad de contratar un **administrador de bases de datos**, el sueldo mensual promedio de un administrador de bases de datos en MySQL, en empresas similares dedicadas al rubro de salud es de alrededor de **\$1000 dólares americanos**. El nivel de seniority buscado es de semisenior, este nivel de seniority es necesario ya que será el responsable de gestionar los permisos de uso de la base de datos y por lo tanto un rol que requiere de una alta responsabilidad y experiencia.

Conclusiones

Tras la investigación de ventajas y desventajas de tres posibles opciones decidimos que **Azure Database for MySQL** es la opción que más se adapta a nuestras necesidades, ya que, tiene grandes prestaciones en los aspectos de rendimiento, almacenamiento y seguridad. Respecto a las otras opciones nos conviene debido a que se nos brinda soporte incluido en el paquete y no como un plan adicional como en el caso de AWS. Al mismo tiempo ofrece un sistema de redundancia tanto para el storage como el backup permitiendo almacenar la información.

Bibliografía

<https://calculator.aws/>

<https://azure.microsoft.com/en-us/pricing/calculator/>

<https://cloud.google.com/products/calculator>

Glosario

CAPEX: Se refiere a las inversiones en activos a largo plazo. Se trata de aquellas inversiones que crean beneficios.

OPEX: Se relaciona con los gastos operativos diarios. Son los gastos para mantener en funcionamiento el sistema.

TCO: Es un análisis que considera todos los costos asociados con la propiedad y operación de un activo durante su vida útil.

Unidad 3

Introducción

En esta entrega (entregas 3 y 4) se llevará a cabo la creación de la base de datos, objetos y restricciones del sistema de turnos médicos del hospital Cure SA. A su vez, se indicará un paso a paso de cómo instalar SQL Server y las configuraciones indicadas para este trabajo.

Desarrollo

Pasos para instalar SQL Server y configuración utilizada para este proyecto:

1. Ingresamos a la siguiente página:

<https://www.microsoft.com/es-es/sql-server/sql-server-downloads>.

2. Deslizamos para abajo y presionamos el botón "descargar ahora" de la versión desarrollador.

3. Ejecutamos el archivo que se nos descargó.

4. Seleccionamos el tipo de instalación personalizada.

5. Elegir ruta de descarga o dejar por default.

6. Presionar instalar.

7. Se nos abrirá una nueva ventana con el Installation Center, en el cual abrimos la pestaña de instalación y clickeamos sobre "New SQL Server standalone installation".

8. Se nos abrirá un wizard el cual iremos dando a siguiente hasta el paso "Feature Selection".

9. Seleccionamos las features a instalar: Todas dentro de la categoría 'Instance Features' y presionamos en siguiente.

10. Configuramos el rango de puertos para los servicios PolyBase (dejamos por defecto el rango 16450-16460).

11. Dentro de la configuración de TempDB seleccionamos el tamaño inicial de la misma (8MB) y el tamaño de auto crecimiento (64MB).

12. Dentro de la configuración de memoria, definimos el mínimo y el máximo (0 y 2TB respectivamente).

13. Avanzamos y presionamos el botón de instalar.

14. Fin

La resolución del trabajo práctico para esta entrega se encuentra en un script entregado junto con este archivo.

Conclusiones

A lo largo del desarrollo de la solución, nos encontramos con diversas problemáticas, tales como:

Para obtener insertar los datos a partir de los archivos .CSV otorgados, fue necesario utilizar SQL dinámico para la inserción de los datos sobre una tabla temporal. Se utilizó SQL dinámico ya que la ruta del archivo es un parámetro que recibe el procedimiento almacenado, por lo cual puede variar y la función implementada para hacer esta inserción (Bulk Insert) no admite variables. A su vez, fue necesario indicar el argumento CODEPAGE (valor 65.001) para que la inserción reconozca tildes y caracteres especiales. Esta codepage es la del sistema operativo que se utilizó para el trabajo práctico y posee información respecto a la codificación de los caracteres.

El archivo pacientes.csv posee la columna "Apellidos" que contiene tanto el nombre paterno como el materno, y "dirección" que contiene tanto el nombre de calle como su altura o número. Para manejar esta información, se debió procesar la cadena de texto separando y validando el contenido de la misma.

El archivo medicos.csv contiene las columnas "Nombre" y "Apellido" intercambiadas, por lo cual en nombre figura el apellido del médico y viceversa. En lugar de modificar el archivo de origen, se optó por hacer una inserción contemplando esta característica y tomando la columna "Nombre" del archivo como si fuera la columna "Apellido" en la tabla final. Además, la columna que contiene el apellido posee el prefijo "Dr.", por lo cual este contenido se debió filtrar.

Para todas las tablas involucradas en los archivos CSV se optó por implementar una clave primaria autoincrementada (identity) para facilitar la inserción y asociación con las tablas correspondientes.

El archivo Prestador.csv contiene 2 separadores ';' adicionales en cada fila. Para resolver esto, se creó una tercera columna auxiliar en la tabla temporal donde se inserta la información del archivo, para evitar que se tomen como caracteres propios de los registros de la tabla de Prestadores.

El archivo JSON con la información de los estudios posee un campo con el signo '\$' que dificultó la obtención de la información, debido a que este símbolo se usa para obtener los datos propios de un objeto en JSON. Además, algunos campos poseen espacios en el propio nombre. Para resolverlo, se indicó el nombre de estos campos entre comillas dobles, evitando que se tengan errores al realizar la lectura del archivo.

Bibliografía

[colocar links -nombre de documentos desde donde se obtuvo la informacion]

Glosario

[]

Unidad 4

Introducción

[describir brevemente el objetivo del TP]

Desarrollo

[describir/narrar la información recopilada/obtenida]

Conclusiones

[Narrar la/s conclusiones y decisiones del equipo]

Bibliografía

[colocar links -nombre de documentos desde donde se obtuvo la informacion]

Glosario

[]

Unidad 5

Introducción

[describir brevemente el objetivo del TP]

Desarrollo

[describir/narrar la información recopilada/obtenida]

Conclusiones

[Narrar la/s conclusiones y decisiones del equipo]

Bibliografía

[colocar links -nombre de documentos desde donde se obtuvo la informacion]

Glosario

[]

Unidad 6

Introducción

[describir brevemente el objetivo del TP]

Desarrollo

[describir/narrar la información recopilada/obtenida]

Conclusiones

[Narrar la/s conclusiones y decisiones del equipo]

Bibliografía

[colocar links -nombre de documentos desde donde se obtuvo la informacion]

Glosario

[]