**Universidad Nacional de La Matanza**

**Departamento de Ingeniería e Investigaciones Tecnológicas**

**Bases de Datos Aplicada**

**Grupo: 2**

**TP Integrador – Entrega N°2**

**Fecha de Presentación: 05/09/2023**

**Integrantes:**

|  |  |  |  |
| --- | --- | --- | --- |
| **DNI** | **Apellido** | **Nombre** | **Email** |
| 43308587 | Clivio | Lucas Ariel | lucas\_clivio@hotmail.com |
| 43991136 | Castillo | Alexis Ezequiel | alexis.fis@hotmail.com |
| 43920122 | Hoz | Lucas Nahuel | hozlucas28@gmail.com |
| 43458499 | Sosa | Gonzalo | gonzalososa2001@outlook.com |

[Primer informe – BBDD On-premise 3](#_Toc150523818)

[Introducción 3](#_Toc150523819)

[Requerimientos técnicos 3](#_Toc150523820)

[Software de base 3](#_Toc150523821)

[Hardware de base 3](#_Toc150523822)

[Costo de licenciamiento 4](#_Toc150523823)

[Segundo informe – BBDD Nube 6](#_Toc150523824)

[Introducción 6](#_Toc150523825)

[Conceptos importantes 6](#_Toc150523826)

[Operating Expense (OPEX) 6](#_Toc150523827)

[Capital Expense (CAPEX) 6](#_Toc150523828)

[Total Cost of Ownership (TCO) 6](#_Toc150523829)

[Análisis de los costos – Azure, AWS y Google Cloud 6](#_Toc150523830)

[Microsoft Azure 6](#_Toc150523831)

[Amazon RDS for MySQL 10](#_Toc150523832)

[Google Cloud - Cloud SQL for MySQL 10](#_Toc150523833)

[Conclusiones 11](#_Toc150523834)

[Bibliografía 11](#_Toc150523835)

[Tercer informe – Documentación del DBMS 12](#_Toc150523836)

[¿Qué es un DBMS? 12](#_Toc150523837)

[Documentación – Configuraciones del SQL Server 12](#_Toc150523838)

# Primer informe – BBDD On-premise

## Introducción

En el presente informe se analizan los aspectos necesarios para la implementación de una base de datos, utilizando MySQL[[1]](#footnote-2), junto con las herramientas necesarias para manipular la misma, como un DBMS[[2]](#footnote-3). Dicho informe está orientado a la firma CURE S.A. Análisis basado en la documentación otorgada por el sitio oficial de MySQL (<http://www.mysql.com>).

## Requerimientos técnicos

### Software de base

Recomendamos usar una distribución de Linux, tal como puede ser CentOS, Ubuntu, o RedHat. En este caso puntual, Ubuntu nos parece una buena opción ya que provee un entorno confiable para la puesta en producción de la base de datos, además su costo de licenciamiento no es tan elevado como RedHat y el soporte viene incluido, por lo cual es la opción adecuada.

### Hardware de base

Según el informe otorgado, CURE S.A prevé que el almacenamiento dispuesto para la base de datos ascenderá a 4 GB en los primeros dos años. Además, el SO[[3]](#footnote-4) ocupará aproximadamente 20 GB por lo que se deberá de disponer de un equipo con 128 GB ≈ 256 GB, preferentemente de la familia NVMe que otorga una alta tasa de lectura/escritura, impactando positivamente en la performance al acceder y/o escribir o sobrescribir datos.

Por otra parte, estimando una concurrencia de 50 usuarios en la etapa inicial del sistema, se recomienda implementar 32 GB de RAM ECC, que podrán hacer frente a la demanda de los 50 usuarios base, dejando una tolerancia suficiente para el añadido de más usuarios.

Respecto a la CPU, se debería optar por un procesador de 8 núcleos ≈ 16 núcleos, entre las alternativas podría considerarse sistemas embebidos EPYC o Intel Xeon. Ya que las exigencias del motor no requieren de un procesamiento intensivo asumiendo una carga de 50 usuarios. Es por eso que recomendamos un procesador Intel Xeon E5-2680 V4 que posee 14 núcleos, siendo una opción viable.

A continuación, describiremos los componentes de hardware mínimos y recomendados para dar una noción de las necesidades físicas de la base de datos:

|  |  |  |
| --- | --- | --- |
|  | Mínimos | Recomendados |
| Usuarios | 10 ≈ 50 | 50 ≈ 100 |
| Procesador | 4 ≈ 8 núcleos CPU | 8 ≈ 16 núcleos CPU |
| RAM | 32 ≈ 64 GB ECC | 128 ≈ 256 GB ECC |
| Almacenamiento (SO) | 20 GB | 20 GB |
| Almacenamiento (DB) | 128 GB | 256 GB |

## Costo de licenciamiento

Luego de un exhaustivo análisis del precio, alcance y necesidades de la base de datos de CURE S.A se ha concluido que la mejor versión de MySQL, ofrecida por Oracle (corporación dueña de MySQL), es: MySQL Standard Edition Subscription (On-Premises, 1-4 socket server).

Consideramos que la versión Standard de MySQL es la más adecuada, dado su costo, características y soporte incluido. Teniendo un costo significativamente inferior a la versión Enterprise, la cual decidimos no incluir al sistema dados sus casos de uso particulares y demasiado críticos para un desarrollo adecuado.

Dicha versión es ofertada por $2.140 USD a la fecha del 27/08/2023.

El sistema operativo que detallamos anteriormente en los requerimientos técnicos, basándonos en la opción recomendada: Ubuntu, oferta una suscripción profesional que incluye soporte para el mismo sistema operativo y, además, nos asegura obtener actualizaciones de seguridad para varios paquetes, dentro de los cuales se encuentra MySQL. Dicha suscripción profesional a Ubuntu oscila los $750 USD (anuales).

A continuación detallaremos los costos del hardware que recomendamos dados los requisitos del sistema:

|  |  |  |
| --- | --- | --- |
|  |  | Precio |
| Placa madre | MBD-X10SRL-F-O | $700 USD |
| Procesador | Intel Xeon E5 – 2680 V4 | $600 USD |
| RAM | Kingston Premier 32GB 2666MT/s DDR4 ECC (16GB x2) | $100 USD ($50 c/u) |
| Almacenamiento | HP EX920 M.2 256GB | $70 USD |

Según el informe de CURE S.A., no poseen personal capacitado para cumplir el rol de un DBA[[4]](#footnote-5) , por lo que se requiere un presupuesto adecuado para financiar la contratación de uno. Se estima que el salario de un DBA ronda los $800 USD ≈ $900 USD.

# Segundo informe – BBDD Nube

## Introducción

En el presente informe abordaremos los costos de implementar la base de datos de la organización CURE S.A utilizando servicios que proveen bases de datos en la Nube[[5]](#footnote-6). Analizaremos los costos de los proveedores: Google Cloud, Amazon Web Services y Microsoft Azure. Además, informaremos el salario promedio de un DBA, que será el encargado de administrar dicha base de datos.

El informe tiene como objetivo conocer la inversión inicial y mensual de la base de datos en la nube, considerando los proveedores mencionados. A continuación, para familiarizarnos con algunos conceptos importantes relacionados a los costos, definiremos OPEX, CAPEX y TCO.

## Conceptos importantes

### Operating Expense (OPEX)

Costos diarios generados por el negocio, pagos por internet, servicios de soporte técnico, entre otros. OPEX abarca todos los costos referidos a las operaciones regulares del negocio.

### Capital Expense (CAPEX)

Costos derivados de grandes adquisiciones fundamentales para el funcionamiento del negocio, tales como servidores On-premise[[6]](#footnote-7) o mismamente el desembolso inicial para un Data Center[[7]](#footnote-8). Involucrando todo lo referido a mejorar el hardware de los equipos, las instalaciones y la infraestructura.

### Total Cost of Ownership (TCO)

Combinación de OPEX y CAPEX refiriéndose al tiempo de vida de un producto o servicio. En otras palabras son los costos directos e indirectos asociados al producto o servicio.

Al combinar OpEx y CapEx se ven reflejados de forma más precisa todos los costos necesarios para mantener en pie al negocio.

## Análisis de los costos – Azure, AWS y Google Cloud

### Microsoft Azure

Dados los requisitos de hardware para la base de datos de CURE S.A, empleando MySQL, Microsoft Azure nos ofrece el servicio Azure Database for MySQL (PaaS).

Este servicio es una instancia administrada de MySQL, donde es posible configurar una alta disponibilidad, backups automáticos, escalado elástico, grados de seguridad, entre otras configuraciones avanzadas.

Al configurar los requerimientos del sistema, hemos decidido alojar el servidor de la base de datos en la región sur de Brasil, buscando minimizar la latencia[[8]](#footnote-9). Con respecto al tipo de servidor, decidimos descartar la opción: Single Server, debido a que esta orientada a aplicaciones ya existentes en la nube. Por lo que optamos por la opción: Flexible Server, que permite migrar las cargas de trabajo de producción al servicio de Azure.

Para cumplir con los requisitos de hardware nos decantamos por una máquina virtual con la siguiente configuración:

|  |  |
| --- | --- |
| Núcleos | 8 (d8as-4v) |
| RAM | 32 GiB ≈ 35 GB |
| Almacenamiento (base de datos) | 120 GiB ≈ 128 GB |

Dicha configuración de la máquina virtual oscila los USD $0,9240 (3/9/2023) por hora.

Al momento de elegir uno de los soportes ofrecidos, optamos por el estándar con un costo mensual de $100 USD. Con respecto a los backups, los mismos se realizan en el servidor, sin costo alguno. En cuanto a la disponibilidad del servicio, tenemos la opción de alta disponibilidad (HA), que garantiza que el servidor no se caiga al mantener réplicas en diferentes zonas. Sin embargo, debido a su costo significativo, hemos preparado dos presupuestos que varían en este aspecto: uno sin alta disponibilidad y otro con alta disponibilidad.

* Costo mensual (incluyendo HA): $961,75 USD
* Costo mensual (excluyendo HA): $530,87 USD

Cabe recalcar que los precios corresponden a la fecha 3/9/2023. Además, estos poseen un descuento del 40% debido a la reserva de compra por todo un año, en el caso de reservar por 3 años el descuento seria del 60%.

A continuación, se encuentran las imágenes de los presupuestos con y sin HA:





### Amazon RDS for MySQL

Dentro de las opciones de bases de datos ofrecidas por AWS, la opción Amazon RDS for MySQL es la opción más factible para el motor de base de datos que buscamos incorporar, ya que es el servicio DBaaS[[9]](#footnote-10) de Amazon para MySQL.

Este servicio ofrece facilidades sobre la base de datos, como elasticidad, supervisión y seguridad, por lo que el equipo de desarrollo puede centrarse en otras tareas productivas. El presupuesto necesario para contratar dicho servicio consta de $718 USD mensuales (3/9/2023). Incluyendo las siguientes prestaciones:

|  |  |
| --- | --- |
| Núcleos | 8 (db.m3.2xlarge) |
| Almacenamiento (backups) | 40 GB |
| Almacenamiento (snapshots) | 40 GB |
| Almacenamiento (base de datos) | 128 GB SSD (GP2) |

También se pueden considerar otras opciones ofrecidas por Amazon, como:

* **Aurora:** servicio similar a Amazon RDS for MySQL respecto a la administración y mantenimiento de las base de datos, aunque con una capa extra de seguridad que nos facilita su uso, tal y como podría ser el manejo del pool de conexiones, aunque para la concurrencia de usuarios que se presenta, no justificaría su precio.
* **Instancia EC2:** servicio que nos permitiría tener control sobre el SO subyacente, aunque tendríamos que destinar un esfuerzo significativo de despliegue y CI/CD[[10]](#footnote-11) para poner en marcha la base de datos. Si bien la inversión podría parecer menor, la propuesta OnDemand[[11]](#footnote-12) generaría costos adicionales cuando el sistema se encuentre altamente demandado.

### Google Cloud - Cloud SQL for MySQL

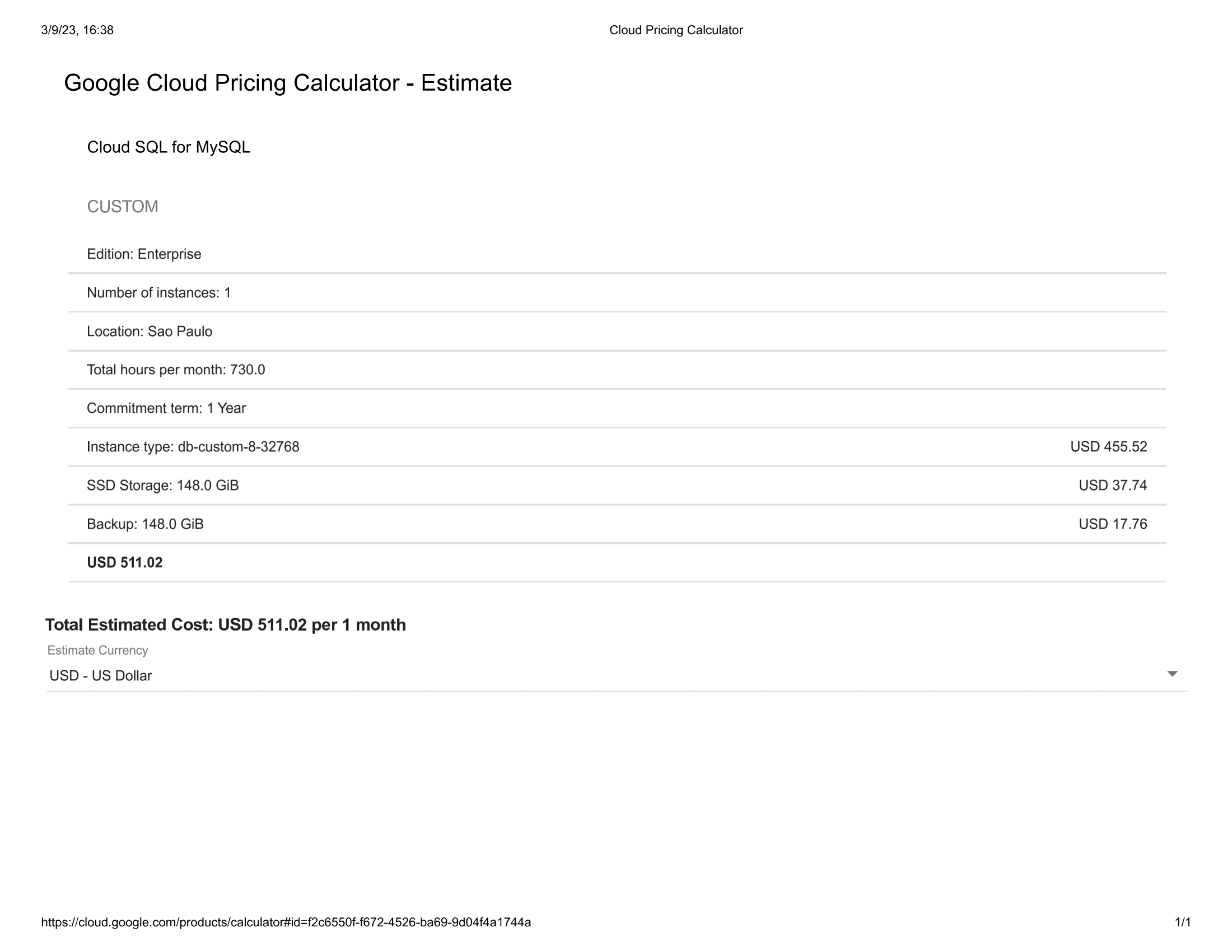
De los servicios de bases de datos en la Nube ofrecidos por Google Cloud, hemos tomado la decisión de seleccionar el servicio: Cloud SQL for MySQL, ya que permite utilizar el lenguaje MySQL que nos permite configurar múltiples aspectos del servicio, pudiendo obtener un sistema a la talla de las necesidades de CURE S.A.

Al ingresar los requisitos del sistema, el prepuesto para una base de datos con las siguientes características:

|  |  |
| --- | --- |
| RAM | 32 GB |
| Núcleos | 8 |
| Ubicación | Sao Paulo - Brasil |
| Almacenamiento (backups) | 148 GB |
| Almacenamiento (base de datos) | 148 GB SSD |

asumirá un costo mensual de $511,02 USD (3/9/2023), empleando la edición Enterprise.

A continuación, se encuentra la imagen del presupuesto mencionado:



## Conclusiones

Luego de analizar los servicios de bases de datos SQL en la nube proporcionados por Microsoft Azure, AWS y Google Cloud, llegamos a la conclusión de que Google Cloud emerge como la opción más económica y adecuada para nuestras necesidades. Con su paquete de $511 USD (3/9/2023), encontramos una solución que cumple con nuestros requisitos sin la necesidad de optar por la alta disponibilidad, la cual resulta costosa y no es esencial actualmente para el sistema. En consecuencia, consideramos que el servicio Cloud SQL for MySQL de Google Cloud es la elección más idónea para las necesidades del sistema de CURE S.A.

## Bibliografía

Presupuesto de AWS (Amazon RDS for MySQL): <https://calculator.aws/#/estimate?id=c7cdfeb535a3c1de6f320357c0d11087d2cea096>Presupuesto de AWS (Instancia EC2): <https://calculator.aws/#/estimate?id=e5bc1546d2da20ae35f2deec906e4159dffcf36c>Presupuesto de Google Cloud (Cloud SQL for MySQL): <https://cloud.google.com/products/calculator#id=20218ad9-083f-4ded-8935-bf93cc871a40>

# Tercer informe – Documentación del DBMS

## ¿Qué es un DBMS?

Un DBMS (Sistema de Gestión de Bases de Datos) es un software que proporciona una interfaz entre la base de datos y los usuarios. Su función principal es gestionar eficientemente la creación, manipulación y administración de bases de datos. Además, el DBMS protege al sistema en caso de fallas de software/hardware, permitiendo el acceso multiusuario a las bases de datos que se encuentran bajo su resguardo.

Para CURE S.A se empleará el DBMS: “SQL Server”, en la siguiente sección se detallarán las configuraciones de este. Con el objetivo de conocer las capacidades y funcionalidades avanzadas que se han implementado al momento de instalarlo.

## Documentación – Configuraciones del SQL Server

En la presente sección se nombrarán las configuraciones empleadas al momento de instalar el DBMS “SQL Server”:

* Edición: Express.
* Servicios implementados:
  + SQL Server Replication.
* Servicios no implementados:
  + Data Quality Services.
  + Machine Learning Services.
  + Búsqueda de texto completo.
  + Servicio de consultas PolyBase.
* Se deshabilito la extensión de Azure para SQL Server.
* Se nombro a la instancia bajo el nombre “ExpressCURESA”.
* Se le asigno al servidor, inicialmente, una memoria máxima de 8 GB.
* Se configuro la base de datos temporal con un tamaño inicial de 8 MB, y con un máximo de 64 MB.
* Ubicación del DBMS: “C:\Program Files\Microsoft SQL Server”.
* Ubicación del directorio: “C:\Program Files\Microsoft SQL Server\MSSQL16.MSSQLSERVER\MSSQL\Data”.
* Ubicación de los registros (logs): “C:\Program Files\Microsoft SQL Server\MSSQL16.MSSQLSERVER\MSSQL\Data”.
* Ubicación de los respaldos (backups): “C:\Program Files\Microsoft SQL Server\MSSQL16.MSSQLSERVER\MSSQL\Backup”.

Además, se optó por instalar el SQL Server Management Studio (SSMS) en caso de que el sistema no pueda ser accedido por un cliente externo. Funcionando como una “puerta trasera” en caso de un bloqueo de clientes/conexiones remotas que comprometa la operabilidad del sistema de CURE S.A.

1. MySQL: sistema de gestión de bases de datos relacional desarrollado por Oracle Corporation. [↑](#footnote-ref-2)
2. DBMS: sistema de gestión de bases de datos, software que permite administrar una base de datos. [↑](#footnote-ref-3)
3. SO: sistema operativo. [↑](#footnote-ref-4)
4. DBA: administrador de base de datos responsable de los aspectos técnicos y tecnológicos de una base de datos. [↑](#footnote-ref-5)
5. Nube: modelo de almacenamiento de datos que está basado en redes de computadoras. [↑](#footnote-ref-6)
6. On-premise: software instalado en sistemas dentro de las instalaciones de la persona u organización. [↑](#footnote-ref-7)
7. Data Center: centro de datos. [↑](#footnote-ref-8)
8. Latencia: retardos temporales dentro de una red. [↑](#footnote-ref-9)
9. DBaaS: data base as a service. [↑](#footnote-ref-10)
10. CI/CD: métodos de integración continua. [↑](#footnote-ref-11)
11. OnDemand: bajo demanda. [↑](#footnote-ref-12)