**Universidad Nacional de La Matanza**

**Departamento de Ingeniería e Investigaciones Tecnológicas**

**Bases de Datos Aplicada**

**Grupo: 9**

**TP Integrador – Entrega nro: 4**

**Fecha de Presentación: 15/11/2023**

**Integrantes:**

|  |  |  |  |
| --- | --- | --- | --- |
| **DNI** | **Apellido** | **Nombre** | **Email** |
| 43034043 | Agasi | Alejo | aagasi@alumno.unlam.edu.ar |
| 41894276 | Giasone | Santiago | sgiasone@alumno.unlam.edu.ar |
| 40784696 | Violi | Stefania | svioli@alumno.unlam.edu.ar |

**Contenido**

[**Unidad 1** 2](#_Toc150807279)

[**Introducción** 2](#_Toc150807280)

[**Desarrollo** 2](#_Toc150807281)

[**Conclusiones** 4](#_Toc150807282)

[**Bibliografía** 4](#_Toc150807283)

[**Glosario** 4](#_Toc150807284)

[**Unidad 2** 5](#_Toc150807285)

[**Introducción** 5](#_Toc150807286)

[**Desarrollo** 5](#_Toc150807287)

[**Conclusiones** 7](#_Toc150807288)

[**Bibliografía** 8](#_Toc150807289)

[**Glosario** 8](#_Toc150807290)

# **Unidad 1**

## **Introducción**

Durante el análisis del sistema a desarrollar para la firma “CURE S.A.”, se concluye que será necesario contar con una base de datos relacional.

El motor de base de datos elegido es Elasticsearch, del cual se detallará:

• Requerimientos técnicos   
 o Software de base   
 o Hardware recomendado (mínimo requerido y recomendado)   
 ▪ CPU, Memoria y Almacenamiento

• Costo y modo de licenciamiento   
 o La primera estimación es que el sistema se accederá desde 50 puestos  
 o Se estima que la base de datos acumulará 4 GB en los primeros dos años.  
 o No se cuenta con personal capacitado en la organización que pueda desempeñarse como dba.   
 o Costos en USD (BNA) o pesos a la cotización del día.

## **Desarrollo**

Elasticsearch es un motor de búsqueda y análisis de datos distribuido y de código abierto desarrollado por Elastic. Está diseñado para almacenar, buscar y analizar grandes volúmenes de datos de manera rápida y eficiente. Su principal enfoque es permitir búsquedas rápidas y relevantes en grandes conjuntos de datos no estructurados o semiestructurados.

Elasticsearch brinda velocidad, escalabilidad y flexibilidad. Sin embargo, su implementación y mantenimiento pueden requerir conocimientos técnicos y experiencia en operaciones distribuidas. Si se configura y administra correctamente, puede ser una herramienta valiosa para abordar una amplia gama de casos de uso, desde búsqueda de texto completo hasta análisis de registros y más.

**1) Requerimientos Técnicos**

**a) Software de Base:** Elasticsearch es el componente central, pero también se necesitarán otras herramientas en el stack ELK (Elasticsearch, Logstash, Kibana) para un sistema completo de análisis y visualización de datos. Estos son los componentes principales:

· **Logstash**

· **Beats**

· **Elasticsearch**

· **Kibana**

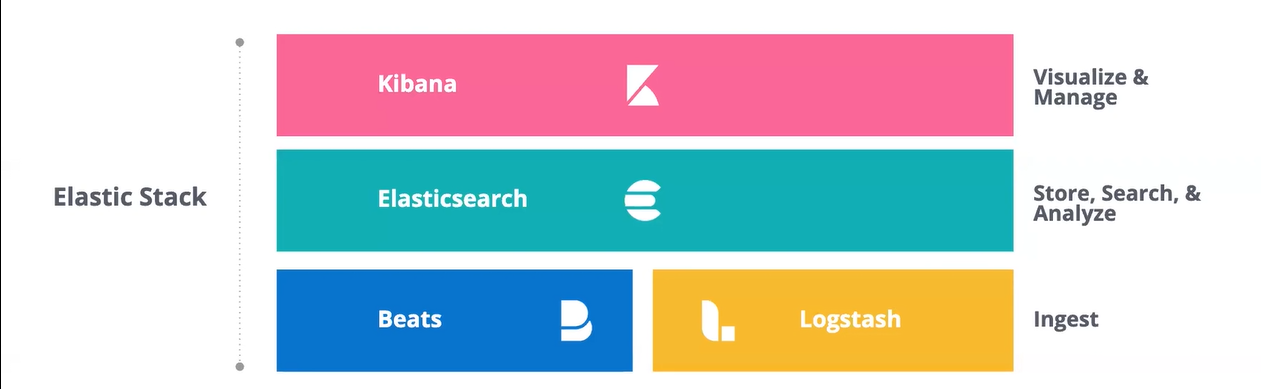
La ingesta de datos hacia Elasticsearch se realiza de dos formas (Beats y Logstash):

· **Logstash:** Utilizado para enriquecer y transformar los datos, me permite conectarme con sistemas de información a través de APIs

· **Beats:** Utiliza agentes rápidos instalados en servidores, principalmente para enviar información tipo log y métricas

· **Elasticsearch:** Es el motor de búsquedas analítico.

· **Kibana:** Para la visualización y análisis de datos.



**b) Hardware Recomendado:**

· **CPU:** Mínimo 2 núcleos, recomendado 4 núcleos o más.

· **Memoria RAM:** Mínimo 8 GB, recomendado 16 GB o más.

· **Almacenamiento:** Depende del volumen de datos. Se recomienda SSD para un mejor rendimiento.

**2) Costo y Modo de Licenciamiento**

Es conveniente comenzar con la prueba gratuita de 14 días ofrecida. Cuando ésta caduque, la mejor opción es la suscripción anual estándar, puesto que ofrece las características básicas necesarias y también soporte (cubriendo así la situación planteada anteriormente), además de que nos permite obtener un descuento. El precio de este no está detallado ya que dicha modalidad implica contactarse con Elastic, pero para tener una noción la suscripción mensual estándar cuesta 95 USD. Incluso también se puede obtener una suscripción de varios años para ahorrar aún más. Las suscripciones anuales nuevas se facturan con el [modelo de consumo prepago](https://www.elastic.co/guide/en/cloud/current/ec-billing-models.html#ec-prepaid-consumption) y su uso se mide en unidades de consumo de Elastic. En este modelo, compras unidades de consumo de Elastic (ECU) por anticipado, y todo uso posterior se debita del saldo de ECU.

## **Conclusiones**

Consideramos que Elasticsearch no se adecúa de la mejor manera a las necesidades del cliente, puesto que se trata de un motor de Base de Datos No Relacional y está orientado para el análisis de grandes volúmenes de datos. Dado el modelo de negocio del cliente y sus proyecciones de escalabilidad para los primeros dos años, creemos conveniente utilizar una Base de Datos Relacional como podría ser SQL Server, MySQL, PostgreSQL, etc.

## **Bibliografía**

[Elasticsearch Platform: Encuentra respuestas en tiempo real a escala | Elastic](https://www.elastic.co/es/)

[Precios de Elasticsearch oficial: Elastic Cloud, Elasticsearch gestionado](https://www.elastic.co/es/pricing)

[Preguntas frecuentes sobre precios de Elastic](https://www.elastic.co/es/pricing/faq)

[Elasticsearch: Primeros pasos | Elastic Videos](https://www.elastic.co/es/webinars/getting-started-elasticsearch?elektra=en-elasticsearch-page)

[Hardware prerequisites | Elastic Cloud Enterprise Reference [3.6]](https://www.elastic.co/guide/en/cloud-enterprise/current/ece-hardware-prereq.html)

## **Glosario**

Ingesta de Datos: Es el proceso de importar grandes archivos de datos de múltiples fuentes a un único sistema de almacenamiento basado en la nube, desde el que se puede acceder a los mismos y analizarlos. Como los datos tienen diferentes formas y proceden de centenares de fuentes, se limpian y transforman en un formato único utilizando un proceso de extraer/transformar/cargar (ETL).

# **Unidad 2**

## **Introducción**

Luego de presentar la alternativa al cliente, el mismo planteó la posibilidad de alojar la base de datos en la nube.

Por lo tanto, se repetirá el análisis de la parte 1 para un servicio alojado en la nube, considerando los mismos lineamientos de capacidad y cálculo que fueron estimados.

Se realizará el cálculo de costo para:   
 • Amazon Web Services   
 • Microsoft Azure   
 • Google Cloud

Además, el cliente comenta que RRHH está analizando la posibilidad de contratar un DBA, por lo que se investigará el sueldo promedio de un DBA en el DBMS asignado (o el que se considera más apropiado), especificando el seniority del perfil que buscan.

Por otra parte, se incluyen conclusiones respecto a aspectos de implementación para el cliente, planteando los correspondientes argumentos que respalden dicha decisión.

Por último, en el glosario se incluyen términos importantes como CAPEX, OPEX y TCO.

## **Desarrollo**

Teniendo en cuenta las conclusiones de la Unidad 1, realizaremos las estimaciones de costos con el motor de Bases de Datos MySQL.

-Para la nube de **Google Cloud:**

*Requerimientos Mínimos*

**Número de instancias**: 1  
 **vCPU**: 2 núcleos  
 **RAM:** 7.5 GB  
 **Almacenamiento:** 10 GB HDD (es el mínimo disponible)  
 **Tamaño de Backup:** 4GB  
 **Tiempo promedio DB ejecutando:** 24 horas x 7 días a la semana **Ubicación:** Santiago de Chile **Costo Total Estimado:** 139.75 dólares al mes

Si se cambia el almacenamiento por 10 GB SSD, el costo estimado es de: 140.87 dólares al mes.

*Requerimientos Recomendados*

**Número de instancias**: 1  
 **vCPU**: 4 núcleos  
 **RAM:** 15 GB  
 **Almacenamiento:** 10 GB HDD (es el mínimo disponible)  
 **Tamaño de Backup:** 4GB  
 **Tiempo promedio DB ejecutando:** 24 horas x 7 días a la semana **Ubicación:** Santiago de Chile **Costo Total Estimado:** 277.79 dólares al mes

Si se cambia el almacenamiento por 10 GB SSD, el costo estimado es de: 278.91 dólares al mes.

-Para la nube de **Microsoft Azure:**

**Número de instancias**: 1  
 **vCPU**: 2 núcleos  
 **RAM:** no especifica  
 **Almacenamiento:** 5 GB HDD  
 **Tamaño de Backup:** 4GB  
 **Tiempo promedio DB ejecutando:** 24 horas x 7 días a la semana **Ubicación:** Sur de Brasil  
 **Costo Total Estimado:** 105.94 dólares al mes

-Para la nube de **Amazon Web Services:**

*Requerimientos Mínimos*

**Número de instancias**: 1  
 **vCPU**: 2 núcleos  
 **RAM:** 7.5 GB   
 **Almacenamiento:** 5 GB HDD  
 **Tamaño de Backup:** 4GB  
 **Tiempo promedio DB ejecutando:** 24 horas x 7 días a la semana **Ubicación:** San Pablo (Brasil)  
 **Costo Total Estimado:** 172.88 dólares al mes

*Requerimientos Recomendados*

**Número de instancias**: 1  
 **vCPU**: 2 núcleos  
 **RAM:** 16 GB   
 **Almacenamiento:** 5 GB HDD  
 **Tamaño de Backup:** 4GB  
 **Tiempo promedio DB ejecutando:** 24 horas x 7 días a la semana **Ubicación:** San Pablo (Brasil)  
 **Costo Total Estimado:** 325,45 dólares al mes

**Sueldos DBA**

Respecto al requerimiento del cliente de contratar un DBA, se concluye que lo más adecuado para la situación planteada es un DBA Semi Senior, siendo el sueldo promedio del mismo en MySQL:

* U$S 8500-12000/mes (Estados Unidos).
* € 1600-2500/mes (España).
* MXN$ 27500/mes (México).
* ARS$ 450.000/mes (Argentina)

## **Conclusiones**

Teniendo en cuenta todo el desarrollo anterior, recomendamos contratar un DBA nivel Semi Senior, ya que el mismo se encuentra capacitado para administrar una Base de Datos del estilo y los requerimientos que el cliente plantea. Un DBA SSr puede entre otras cosas:

* Manejar tareas complejas de configuración, optimización y seguridad.
* Es capaz de realizar tareas de copia de seguridad y recuperación avanzadas.
* Puede colaborar en proyectos más grandes y ayudar a los DBAs junior en su desarrollo (en caso de que la Base de Datos crezca y se requieran más DBAs).
* Tiene varios años de experiencia en la administración de bases de datos.

Por otro lado, se puede presentar la duda de si alojar la Base de Datos en la nube o hacerlo On-Premise. Recomendamos utilizar una de las nubes anteriormente descritas, ya que las mismas poseen la flexibilidad y escalabilidad de reducir o incrementar los requerimientos de hardware para que se adapten a las necesidades del cliente. Si se optara por alojar la Base de Datos On-Premise, es probable que cierta tecnología (como las memorias RAM, la capacidad de Almacenamiento o incluso los modelos de procesadores) pueda quedar obsoleta al cabo de unos años, y se tenga que volver a comprar nuevos componentes de Hardware más actualizados o que se adecúen a las nuevas necesidades del cliente.

## **Bibliografía**

[Calculadora de precios de Google Cloud](https://cloud.google.com/products/calculator?hl=es)

[Pricing Calculator | Microsoft Azure](https://azure.microsoft.com/en-us/pricing/calculator/)

[AWS Pricing Calculator](https://calculator.aws/#/addService)

[openqube: Resultados de la encuesta de sueldos de Julio 2023](https://sueldos.openqube.io/encuesta-sueldos-2023.02/)

[Bolsa de trabajo de dba mysql Agosto 2023 - OCCMundial](https://www.occ.com.mx/empleos/de-dba-mysql/)

[Salario medio para Administrador Datos Mysql en México 2023](https://mx.talent.com/salary?job=administrador+datos+mysql)

[Salario medio para Remote Dba en Estados Unidos 2023](https://www.talent.com/es/salary?job=remote+dba)

[Sueldo de Administrador/a de Bases de Datos (2023) ¿Cuánto Cobra?](https://www.kiwiremoto.com/sueldo/administrador-de-bases-de-datos/)

[Sueldo: Database Administrator (DBA) en Buenos Aires, Argentina 2023 | Glassdoor](https://www.glassdoor.com.ar/Sueldos/buenos-aires-database-administrator-dba-sueldo-SRCH_IL.0,12_IM963_KO13,39.htm)

[Senior Database Administrator (DBA) Salary in 2023 | PayScale](https://www.payscale.com/research/US/Job=Senior_Database_Administrator_(DBA)/Salary/9ea6bd00/Mid-Career)

[Capex - Wikipedia, la enciclopedia libre](https://es.wikipedia.org/wiki/Capex)

[CAPEX - Qué es, definición y concepto | 2023 | Economipedia](https://economipedia.com/definiciones/capex.html)

[Fundamentos de OCI. CAPEX vs OPEX](https://www.youtube.com/watch?v=rkIjtvw21fI)

[Opex - Wikipedia, la enciclopedia libre](https://es.wikipedia.org/wiki/Opex)

[Coste total de propiedad - Wikipedia, la enciclopedia libre](https://es.wikipedia.org/wiki/Coste_total_de_propiedad)

[Comprendiendo el TCO (Total Cost of Ownership): origen, definición, cálculo y ventajas](https://www.manutan.es/blog/comprendiendo-tco-total-cost-of-ownership/#productBeginIndex:0&orderBy:7&)

## **Glosario**

**Conceptos**

**CAPEX:** Las inversiones en bienes de capital, gastos en capital o CAPEX son inversiones de capital que crean beneficios. Los CAPEX son utilizados por una compañía para adquirir o mejorar los activos fijos tales como equipamientos, propiedades o edificios industriales. En un servicio basado en la nube, no se tiene esta inversión.

**OPEX**: Es un costo permanente para el funcionamiento de un producto, negocio o sistema. Puede traducirse como gasto de funcionamiento, gastos operativos, o gastos operacionales. Alojar una Base de Datos en la nube implica tener un costo de funcionamiento u OPEX

Por ejemplo, la compra de una fotocopiadora implica gastos de capital, y el consumo anual de papel, tóner, energía y el coste de mantenimiento representa los gastos operativos. Para sistemas más grandes como las empresas, los gastos operativos también pueden incluir el costo de los trabajadores y los gastos de instalación como la renta y servicios (agua, luz, etc).

**TCO**: El costo total de propiedad es un método de cálculo diseñado para ayudar a los usuarios y a los gestores empresariales a determinar los costes directos e indirectos, así como los beneficios, relacionados con un producto o sistema. Básicamente, es un método de cálculo que determina el coste total de un producto o servicio a lo largo de su ciclo de vida.