**TALENTO TECH**

**ANALISIS DE DATOS NIVEL INTEGRADOR G2**

**UNIDAD 2**

**ACTIVIDAD 1 LECCIÓN 1**

**PROVEEDORES DE SERVICIOS EN LA NUBE**

**ESTUDIANTE:** Andrés Fernando Ruales Muñoz

**PROFESOR:** Andrés Felipe Escallón Portilla

**POPAYÁN – CAUCA**

**2024**

**ACTIVIDAD 1**

**Tipo actividad: Proveedores de servicios en la nube (2 horas)**

Para esta actividad se realizará el análisis de los proveedores en la nube con los estudiantes. Utilizando 10 minutos, durante la sesión, los estudiantes deben averiguar qué empresas y servicios existen que provean servicios de:

* Almacenamiento de datos en la nube
* Procesamiento en la nube
* Creación de servidores virtualizados
* Creación de entornos de jupyter lab
* Creación de entornos de DataBricks

Una vez se conozcan los proveedores, se distribuyen los estudiantes en grupos y a cada uno de ellos se le asigna la tarea de investigar más de cada proveedor de servicios. Como actividad deberán determinar el costo de algunas de las herramientas más comunes como son:

1. Almacenamiento de datos (storage, blobs, entre otros)
2. Creación de máquinas virtuales y tipos de máquinas que se pueden crear
3. Herramientas de analítica de datos automatizadas
4. Herramientas de machine learning automatizadas (sagemaker, vertex, azure cognitive services, entre otros)

Para finalizar la sesión se debaten los resultados con los estudiantes para que cada grupo pueda ver las principales diferencias y ventajas de los proveedores encontrados en la nube. Además, los estudiantes deben concluir acerca de si es posible ejecutar los servicios en la nube; se debe preguntar si en todos los proveedores es posible correr cualquiera de estos entornos.

**SOLUCIÓN**

**Almacenamiento de datos en la nube:**

Amazon Web Services (AWS): Ofrece Amazon S3 (Simple Storage Service) para almacenamiento de objetos.

Microsoft Azure: Ofrece Azure Blob Storage para almacenamiento de objetos y Azure Data Lake Storage para almacenamiento de datos escalable y seguro.

Google Cloud Platform (GCP): Ofrece Google Cloud Storage para Almacenamiento de objetos.

**Procesamiento en la nube:**

Amazon Web Services (AWS): Ofrece Amazon EC2 (Elastic Compute Cloud) para instancias virtuales de servidores y Amazon EMR (Elastic MapReduce) para procesamiento de datos distribuido.

Microsoft Azure: Ofrece Azure Virtual Machines para máquinas virtuales y Azure HDInsight para procesamiento de big data.

Google Cloud Platform (GCP): Ofrece Google Compute Engine para instancias virtuales de servidores y Google Cloud DataProc para procesamiento de big data.

**Creación de servidores virtualizados:**

Amazon Web Services (AWS): Ofrece Amazon EC2 (Elastic Compute Cloud) para la creación de instancias virtuales de servidores.

Microsoft Azure: Ofrece Azure Virtual Machines para la creación de máquinas virtuales.

Google Cloud Platform (GCP): Ofrece Google Compute Engine para la creación de instancias virtuales de servidores.

**Creación de entornos de Jupyter Lab:**

Jupyter Project: La organización detrás de Jupyter Notebook, que también ofrece Jupyter Lab, una interfaz de usuario interactiva basada en la web para escribir y ejecutar código en varios lenguajes de programación.

Google Colab: Ofrece entornos de Jupyter Notebook basados en la nube para escribir y ejecutar código de Python y otros lenguajes de programación.

Microsoft Azure Notebooks: Ofrece entornos de Jupyter Notebook basados en la nube para escribir y ejecutar código de Python, R y F#.

**Creación de entornos de DataBricks:**

Databricks: Es una plataforma de análisis de datos basada en la nube que simplifica el procesamiento de big data y el análisis de datos. Ofrece entornos de colaboración para Apache Spark y herramientas integradas para el desarrollo de flujos de trabajo de análisis de datos.

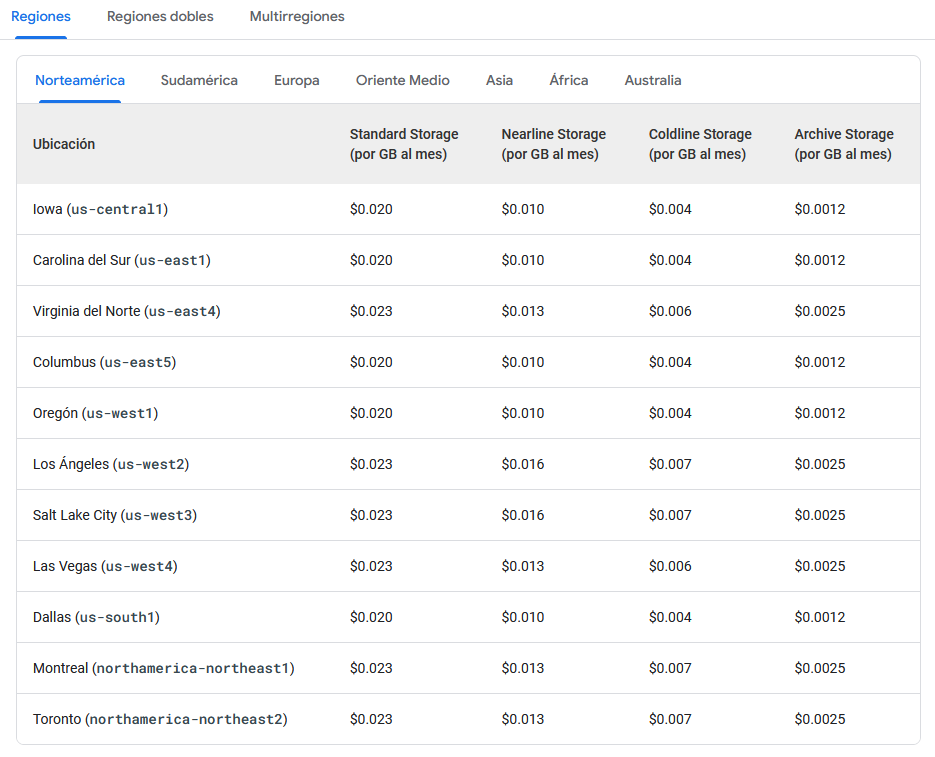
**1.** **Google Cloud Storage**

Los precios de Cloud Storage están relacionados con los siguientes componentes:

**Almacenamiento de datos:** cantidad de datos almacenados en tus buckets. Las tarifas de almacenamiento varían según la clase de almacenamiento de los datos y la ubicación de los buckets.

**Procesamiento de datos:** El procesamiento que realiza Cloud Storage, que incluye los cargos por operaciones, las tarifas de recuperación aplicables y la replicación entre regiones.

**Uso de red:** cantidad de datos que se leen o transfieren entre los buckets.



Interfaz de usuario gráfica, Texto, Correo electrónico

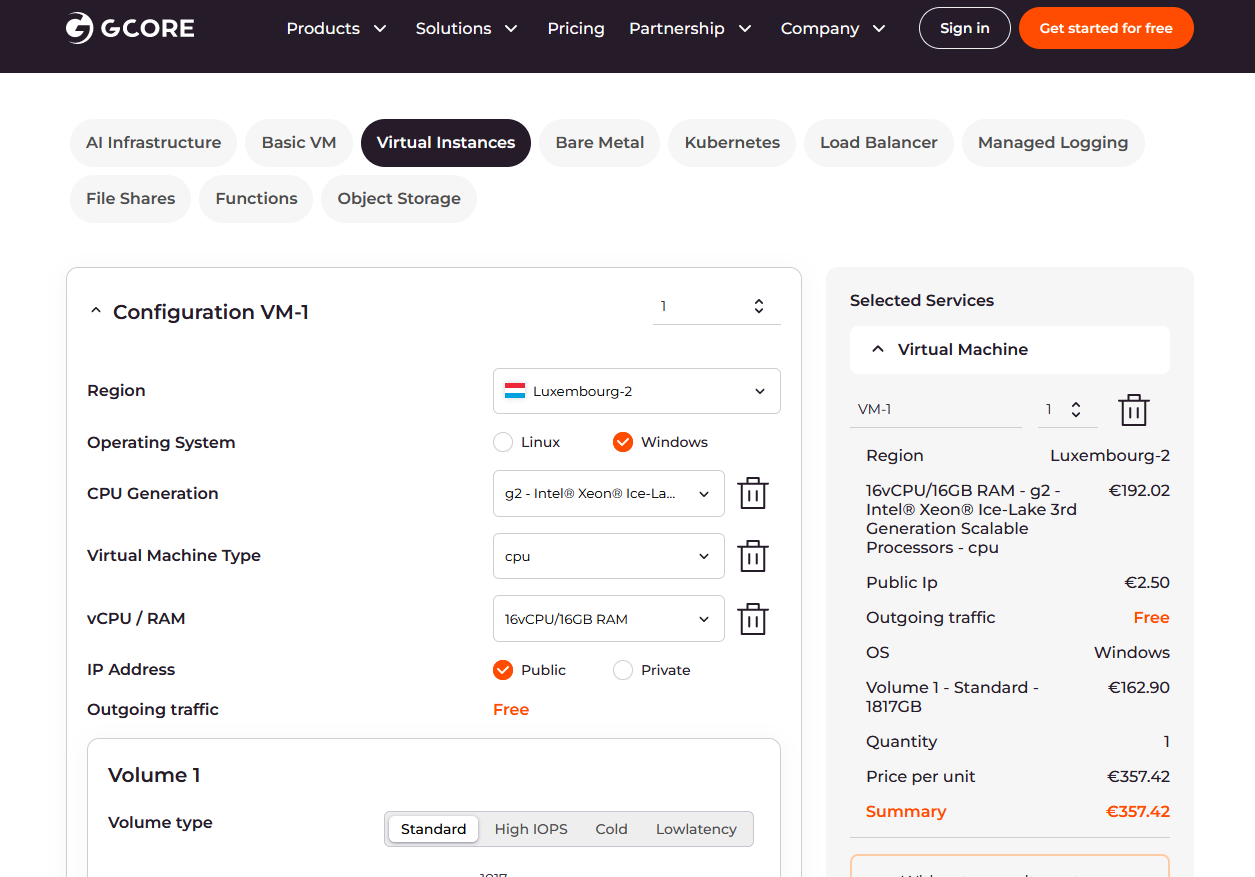
Descripción generada automáticamente

Tabla

Descripción generada automáticamente

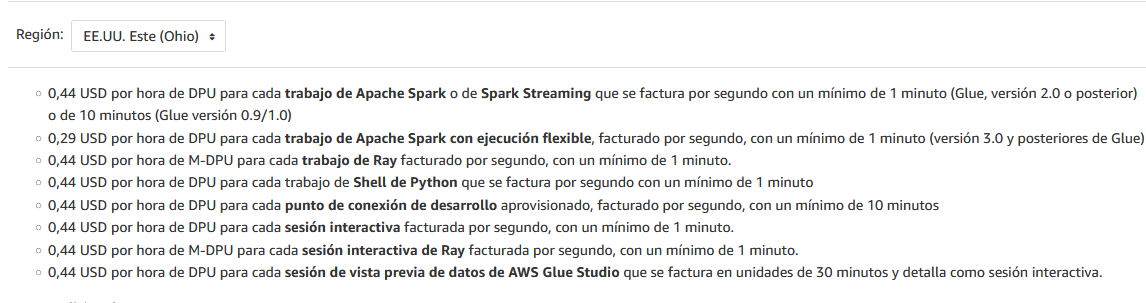
**2. GCORE**

Empresa que ofrece el servicio de maquinas virtuales en la nube dependiendo de las especificaciones de nuestra maquina varia el costo.



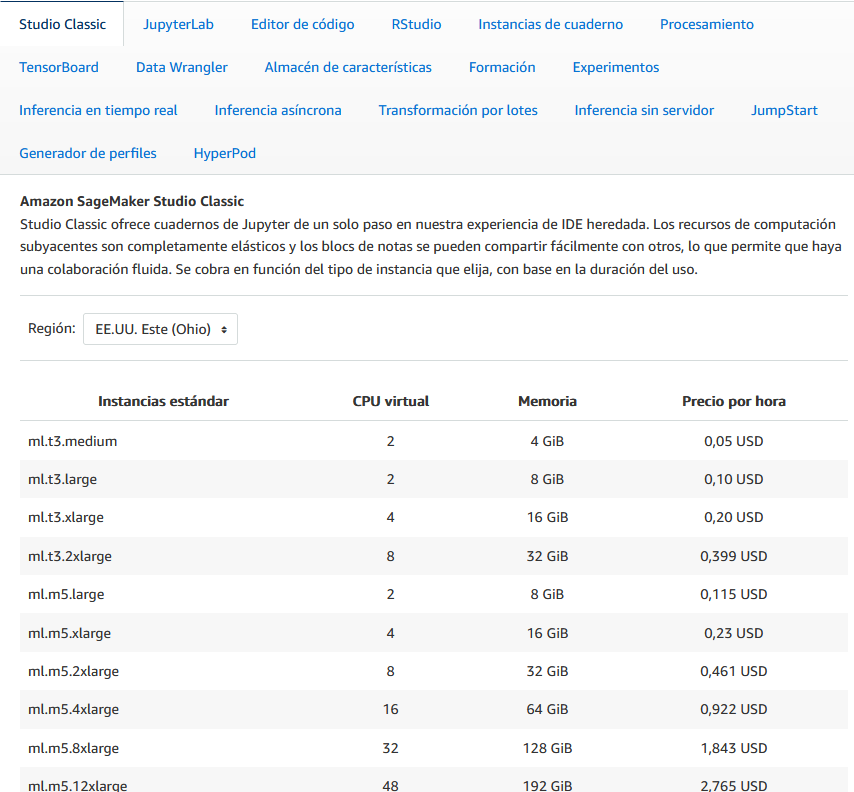
**3. AWS Glue**

La integración de datos sin servidor facilita la detección, preparación y combinación de datos para análisis, machine learning (ML) y desarrollo de aplicaciones.



**4. Amazon SageMaker**

Es un servicio completamente administrado que brinda a todos los científicos de datos y desarrolladores la capacidad de crear, entrenar e implementar modelos de machine learning de forma rápida.



**CONCLUSIÓN**

Es posible ejecutar los servicios en la nube mencionados en la mayoría de los proveedores de servicios en la nube, como Amazon Web Services (AWS), Microsoft Azure y Google Cloud Platform (GCP). Sin embargo, hay algunas consideraciones importantes a tener en cuenta:

Compatibilidad de servicios: Cada proveedor de servicios en la nube ofrece una variedad de servicios y productos, y la disponibilidad de ciertos entornos y herramientas puede variar entre proveedores. Es importante revisar la documentación de cada proveedor para confirmar la disponibilidad de los servicios específicos que necesitas.

Capacidades de la infraestructura: Algunos servicios en la nube pueden tener requisitos específicos de infraestructura o recursos subyacentes para ejecutar ciertos entornos. Por ejemplo, ejecutar entornos de DataBricks puede requerir una infraestructura que admita Apache Spark, mientras que la creación de servidores virtualizados puede requerir una capacidad de cómputo específica.

Costos y precios: Los costos asociados con la ejecución de servicios en la nube pueden variar entre proveedores y pueden depender de factores como el tipo y tamaño de los recursos utilizados, la región geográfica y el nivel de servicio. Es importante considerar los costos asociados con la ejecución de los servicios en la nube y comparar entre proveedores para determinar la opción más adecuada.

En resumen, si bien es posible ejecutar la mayoría de los servicios en la nube en varios proveedores, es importante investigar y evaluar las opciones disponibles para determinar el proveedor que mejor se adapte a tus necesidades específicas en términos de disponibilidad de servicios, compatibilidad de infraestructura y costos.