

RETO INTRO BIG DATA ENTREGABLE PARCIAL

Guía para el estudiante

Cada estudiante debe hacer entrega de:

1. Enlace a su sitio web: Este enlace debe redirigir al sitio web que cada estudiante ha desarrollado para este entregable con las secciones de contenido que se indican en este documento

https://grupo07-bigdata.github.io/practica_big_data/

2. Documento con capturas de pantalla: Un documento PDF que incluya capturas de pantalla del sitio web. Este documento servirá como prueba inmutable de la publicación en caso de que haya cambios posteriores.

ACERCA DE NOSOTROS

ASIR Maria Teresa Joven Sanagustin

DAW Ludmila Da Silva Marques

DAM Javier Cerejido Cortés

CONTACTO

t.jov@hotmail.com

ludmiladsilva1@gmail.com

lvajcc@hotmail.com

PROYECTO RETO BIG DATA

ENTREGABLE PARCIAL

Sección #1

Sección #2

Sección #3

Sección #4

BIG DATA

Haciendo click en los distintos botones de la página de inicio anterior, se abren los documentos con la solución a lo que se pide en cada una de las secciones siguientes:

Sección #1 Conociendo plataformas cloud

Plataformas Cloud

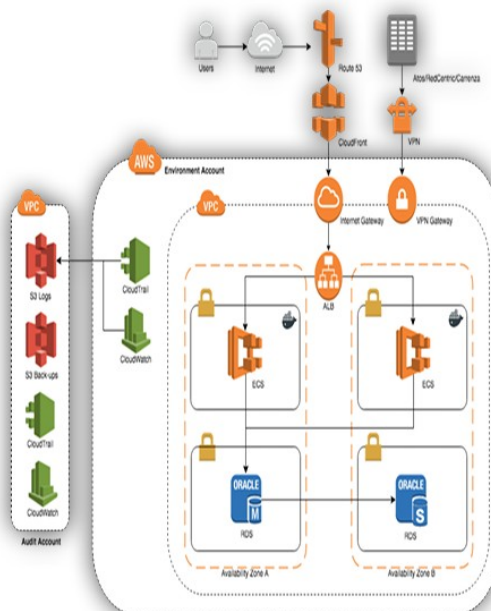
	AWS	Azure	Google Cloud	IBM Cloud
Precio	Pago por uso, modelo flexible. Descuentos por uso reservado a largo plazo.	Pago por uso y por capas (Pay-as-you-go). Opciones de ahorro con planes a largo plazo.	Precios similares, descuentos por compromiso de uso a largo plazo.	Modelos flexibles según el uso y tamaño del negocio. Opciones a medida.
Nivel de implantación	Más utilizado globalmente, líder en el mercado de cloud.	Muy popular, especialmente en empresas que ya usan productos Microsoft.	Tercero en el mercado, con foco en IA y Big Data.	Popular en empresas que requieren entornos híbridos y legacy systems.
Limitaciones	Complejidad en la gestión para usuarios novatos. Costos pueden aumentar rápidamente sin buena gestión.	Integrado fuertemente con productos Microsoft, lo que puede limitar en entornos no-Microsoft.	Menor variedad de servicios comparado con AWS. Ecosistema más pequeño.	Limitada oferta en algunos servicios avanzados de IA o machine learning.
Ventajas frente a otras	Amplia gama de servicios y presencia global en más regiones que cualquier otra.	Integración nativa con herramientas como Office 365 y Active Directory.	Fuerte en inteligencia artificial, aprendizaje automático y análisis de datos.	Excelente para empresas con sistemas heredados y modelos híbridos.
Facilidad de aprendizaje	Complejo al principio, pero bien documentado.	Fácil para usuarios de entornos Microsoft, buena curva de aprendizaje para otros.	Muy enfocado en desarrolladores, fácil si se tiene experiencia en Google.	Fuerte curva de aprendizaje pero con buenos recursos para desarrolladores.
Otra característica	Alta escalabilidad, ideal para empresas de todos los tamaños.	Seguridad integrada y compliance con diversas regulaciones globales.	Foco en sostenibilidad, con iniciativas de reducción de huella de carbono.	Fuerte enfoque en seguridad y entornos empresariales críticos.

Plataforma finalista para realizar la migración

Como grupo, reconocemos a AWS como líder del mercado por su amplia gama de servicios y su infraestructura global, asegurando alta disponibilidad y rendimiento. La flexibilidad en precios y la variedad de herramientas, desde almacenamiento básico hasta inteligencia artificial, nos permiten satisfacer tanto necesidades actuales como futuras. Aunque su manejo puede ser complejo inicialmente, la extensa documentación y el soporte comunitario facilitan el aprendizaje y la adaptación, haciendo de AWS una opción estratégica para nuestro desarrollo tecnológico.

Sección #2 AWS

Comentario sobre diagrama



¿Están todos los servicios correctamente ubicados?

Creemos que sí. Los elementos clave como los balanceadores de carga (ALB), servicios en contenedores (ECS) y bases de datos (RDS) están distribuidos en diferentes zonas de disponibilidad dentro de una VPC. Además, hay conexiones a través de gateways y se utiliza CloudFront y Route 53 para la distribución de tráfico, lo cual está alineado con las mejores prácticas de AWS.

¿Por qué no se usa EC2 y se usa ECS? ¿Diferencias?

Consideramos que Amazon EC2 (Elastic Compute Cloud) es una solución más versátil para lanzar y gestionar servidores virtuales en la nube, brindando mayor control sobre el sistema operativo y la infraestructura. Por otro lado, Amazon ECS (Elastic Container Service) está más optimizado para ejecutar aplicaciones en contenedores, como Docker. Desde nuestra perspectiva, ECS simplifica la orquestación de contenedores, facilitando el despliegue, la escalabilidad y la gestión de aplicaciones en microservicios. Además, creemos que ECS reduce la sobrecarga operativa de gestionar instancias EC2 individuales, proporcionando un enfoque más automatizado y eficiente para ejecutar aplicaciones contenedorizadas.

¿Cuántas redes privadas hay?

En el diagrama se observan dos redes privadas virtuales (VPC). Una de ellas permite almacenar logs y copias de seguridad, además de monitorizar el rendimiento y el estado de los servicios. La otra VPC se divide en dos zonas de disponibilidad (Availability Zone A and B) que contienen servicios críticos como ECS y RDS.

¿En cuántos CPDs se ubica el servicio?

El servicio está distribuido en dos zonas de disponibilidad, lo que indica que se encuentra en al menos dos centros de procesamiento de datos diferentes. Esto garantiza alta disponibilidad y ofrece una mayor tolerancia a fallos."

Sección #3 – Primer servidor web en AWS

Procedimiento

Lanzamiento de una Instancia EC2:

Comenzamos el proceso seleccionando una Amazon Machine Image (AMI), que nos proporcionó una plantilla preconfigurada para nuestra instancia, adecuada a nuestras necesidades específicas de aplicación.

Elegimos un tipo de instancia basado en los recursos que requeríamos, como CPU y memoria, adecuados para soportar nuestra carga de trabajo.

Configuramos las redes y las opciones de seguridad, lo que incluyó establecer grupos de seguridad para controlar el acceso a la instancia.

Definimos scripts de datos de usuario que automatizaron la configuración inicial de la instancia, facilitando su despliegue y configuración operativa desde el primer arranque.

Configuración de Seguridad:

Tras el lanzamiento, ajustamos las reglas de seguridad para asegurarnos de que sólo los tráficos necesarios y seguros pudieran acceder a nuestra instancia. Esto fue vital para proteger los datos y los servicios que se ejecutaban en la instancia.

Modificamos las reglas de entrada del grupo de seguridad para permitir conexiones HTTP, permitiendo el acceso al servidor web que configuramos, asegurando que nuestro contenido web fuera accesible externamente de manera.

Monitoreo y Diagnóstico:

Utilizamos las herramientas de monitoreo proporcionadas por AWS para observar continuamente el desempeño y la salud de nuestra instancia. Estas herramientas nos ayudaron a verificar que todo funcionara como debía y a detectar cualquier problema operativo o de rendimiento tempranamente.

Revisamos los registros del sistema y utilizamos las capturas de pantalla de la consola de AWS para diagnósticos más profundos, lo cual fue crucial para mantener la estabilidad y la eficiencia operativa.

Protección de la Instancia:

Implementamos configuraciones para proteger nuestra instancia de detenciones o terminaciones accidentales. Activamos protecciones específicas para garantizar que nuestra instancia continuara operando sin interrupciones, salvaguardando así contra pérdidas de servicio o datos.

Redimensionamiento de Recursos:

Según fue necesario, ajustamos los recursos asignados a nuestra instancia. Esto incluyó cambiar el tipo de instancia a una más potente cuando las demandas de procesamiento o memoria aumentaron y expandir el almacenamiento disponible para acomodar un volumen mayor de datos.

Estos ajustes nos permitieron escalar de manera efectiva y mantener un rendimiento óptimo a medida que evolucionaron nuestras necesidades y condiciones operativas.

Consola con un servidor web con la protección contra terminación habilitada

i-07a6a74c43bd4d2e0 (Web Server)	
Required	 <code>arn:aws:ec2:us-east-1:409744288719:instance/i-07a6a74c43bd4d2e0</code>
▼ Detalles de la instancia Información	
Plataforma  Amazon Linux	ID de AMI  <code>ami-06b21ccaff8cd686</code>
Detalles de la plataforma  Linux/UNIX	Nombre de AMI  <code>al2023-ami-2023.6.20241010.0-kernel-6.1-x86_64</code>
Detener la protección desactivado	Hora de lanzamiento  <code>Wed Oct 23 2024 12:09:24 GMT+0200 (hora de verano de Europa central)</code> (24 minutes)
Recuperación automática de instancias Predeterminada	Ciclo de vida normal
Índice de lanzamiento de AMI 0	Par de claves asignado en el lanzamiento  <code>vockey</code>
Especificación de crédito standard	ID de kernel  <code>kernel-6.1-x86_64</code>
	Monitoreo desactivado
	Proceso de terminación Habilitado
	Ubicación de AMI  <code>amazon/al2023-ami-2023.6.20241010.0-kernel-6.1-x86_64</code>
	Comportamiento de detención de hibernación desactivado
	Motivo de transición de estado -
	Mensaje de transición de estado

A screenshot of a web browser window. The address bar at the top shows a back arrow, a refresh icon, a lock icon, the text 'Não segura', and the IP address '54.162.70.31'. The main content area of the browser displays the text 'Hello world' in a large, bold, black serif font.

The screenshot shows the AWS CloudWatch console for an EC2 instance. The 'Monitoreo' (Monitoring) tab is active. The dashboard displays eight metrics:

- Utilización de la CPU (%):** A line graph showing CPU usage fluctuating between approximately 1.6% and 3.2%.
- Entrada de red (bytes):** A line graph showing network input in bytes, fluctuating between approximately 180 and 360.
- Salida de red (bytes):** A line graph showing network output in bytes, fluctuating between approximately 180 and 360.
- Paquetes de entrada de red (recuento):** A line graph showing network input packets, fluctuating between approximately 2 and 4.
- Paquetes de salida de red (recuento):** A line graph showing network output packets, fluctuating between approximately 2 and 4.
- Uso de créditos de CPU (recuento):** A line graph showing CPU credit usage, starting at 0.276 and decreasing to 0.
- Saldo de créditos de CPU (recuento):** A line graph showing CPU credit balance, starting at 30.7 and decreasing to 15.4.

Grupos de seguridad (1/5) Información

Acciones

Exportar los grupos de seguridad a CSV

Crear grupo de seguridad

	Name	ID de grupo de seguridad	Nombre del grupo de seguridad	ID de la VPC	Descripción	Propietario
<input type="checkbox"/>	-	sg-0562d4dd1391c226f	default	vpc-07926095e384e46d5	default VPC security group	409744288719
<input type="checkbox"/>	-	sg-01d73df44ca145ba9	default	vpc-0db280a9501551c23	default VPC security group	409744288719
<input type="checkbox"/>	-	sg-00281b2c013683a62	default	vpc-0a1fbca70781d495f	default VPC security group	409744288719
<input type="checkbox"/>	-	sg-0459b07e6b06392109	Ec2SecurityGroup	vpc-0a1fbca70781d495f	VPC Security Group	409744288719
<input checked="" type="checkbox"/>	-	sg-056d642111cf79dc4	Web Server security group	vpc-0db280a9501551c23	Security group for my web server	409744288719

sg-056d642111cf79dc4 - Web Server security group

Detalles

Reglas de entrada

Reglas de salida

Etiquetas

Reglas de entrada (1)

Administrar etiquetas

Editar reglas de entrada

	Name	ID de la regla del gr...	Versión de IP	Tipo	Protocolo	Intervalo de puertos	Origen	Descripción
<input type="checkbox"/>	-	sgr-007a5fc3d235a1bf58	IPv4	HTTP	TCP	80	0.0.0.0/0	-

i-07a6a74c43bd4d2e0 (Web Server)

Filtrar dispositivos de bloques

<input checked="" type="checkbox"/>	ID de volumen	Nombre del di...	Tamaño del volu...	Estado de la cone...	Hora de conexión	Cifrado	ID de clave de KMS	Eliminar cuando ter...
<input checked="" type="checkbox"/>	vol-06abaccb04fe16723	/dev/xvda	8	Asociado	2024/10/23 12:09 GMT+2	No	-	Sí

Monitorización del volumen (1)

i-07a6a74c43bd4d2e0 (Web Server)								
<input type="text" value="Filtrar dispositivos de bloques"/>								
<input checked="" type="checkbox"/>	ID de volumen	Nombre del di...	Tamaño del volu...	Estado de la cone...	Hora de conexión	Cifrado	ID de clave de KMS	Eliminar cuando term
<input checked="" type="checkbox"/>	vol-06abaccb04fe16723	/dev/xvda	10	Asociado	2024/10/23 12:09 GMT+2	No	–	Sí

Sección #4 – Balanceo y autoscaling

Infraestructura VPC creada

Balanceadores de carga (1/1)

Elastic Load Balancing escala automáticamente la capacidad del equilibrador de carga en respuesta a los cambios en el tráfico entrante.

Filtrar equilibradores de carga

<input checked="" type="checkbox"/>	Nombre	Nombre de DNS	Estado	ID de VPC	Zonas de disponibilidad	Tipo	Fecha creada
<input checked="" type="checkbox"/>	LabELB	LabELB-2067764169.us-e-	Aprovisionado	vpc-089ee05c65a095...	2 Zonas de disponibilidad	application	24 de octubre de 2024, 09:56 (UTC+02:00)

Equilibrador de carga: LabELB

Detalles |
 Agentes de escucha y reglas |
 Mapeo de red |
 Mapa de recursos - nuevo |
 Seguridad |
 Monitorización |
 Integraciones |
 Atributos |
 Etiquetas

Mapeo de red

Los destinos en las zonas y subredes enumeradas están disponibles para el tráfico procedente del equilibrador de carga con las direcciones IP que se indican.

VPC
vpc-089ee05c65a095373

CIDR de VPC IPv4: 10.0.0.0/16

IPv6 :-

Tipo de dirección IP del equilibrador de carga



IPv4

Mapeos

La selección de dos o más zonas de disponibilidad y las subredes correspondientes aumenta la tolerancia a errores de las aplicaciones.

Zona	Subred	Dirección IPv4 privada	Dirección IPv6
us-east-1a (us-east-1a-z2)	subnet-eafcf18b1aa8911bc0	Asignado por AWS	No aplicable
us-east-1b (us-east-1a-z4)	subnet-dc547eae951d4dc150	Asignado por AWS	No aplicable

ELB creado

Inicio > Grupos de Auto Scaling > Grupo de Auto Escalado Lab			
<h2>Grupo de Auto Escalado Lab</h2>			
<div> Detalles Actividad Escalado automático Administración de instancias Monitoreo Actualización de instancias </div>			
<div> <div>Detalles del grupo</div> <div> <div> <div> <div>Nombre del grupo de Auto Scaling</div> <div>Grupo de Auto Escalado Lab</div> </div> <div> <div>Fecha de creación</div> <div>Thu Oct 24 2024 10:13:25 GMT-0200 (hora de verano de Europa central)</div> </div> </div> <div> <div>Capacidad deseada</div> <div>2</div> </div> <div> <div>Capacidad mínima</div> <div>2</div> </div> <div> <div>Capacidad máxima</div> <div>6</div> </div> </div> <div> <div> <div>Tipo de capacidad deseado</div> <div>Unidades (número de instancias)</div> </div> <div> <div>Estado</div> <div>-</div> </div> </div> <div> <div> <div>Nombre de recurso de Amazon (ARN)</div> <div>arn:aws:autoscaling:us-east-1:667288089457:autoScalingGroup:63a42d8d-555e-4c5f-8acc-06d6756ca15:autoScalingGroupName/Grupo de Auto Escalado Lab</div> </div> </div> <div> <div>Editar</div> </div> </div>			
<div> <div>Plantilla de lanzamiento</div> <div> <div> <div>Plantilla de lanzamiento</div> <div>  ti-02b97260a46f1a816-LabGenRng </div> </div> <div> <div>Versión</div> <div>Default</div> </div> <div> <div>Descripción</div> <div>-</div> </div> </div> <div> <div> <div>ID de AMI</div> <div>  ami-0ce90429c8be67104 </div> </div> <div> <div>Grupos de seguridad</div> <div>-</div> </div> <div> <div>Almacenamiento (volumen)</div> <div>-</div> </div> </div> <div> <div> <div>Tipo de instancia</div> <div>t2.micro</div> </div> <div> <div>ID de grupos de seguridad</div> <div>sg-04f930d6b7bced0ff</div> </div> <div> <div>Nombre del par de claves</div> <div>vokeyey</div> </div> </div> <div> <div> <div>Propietario</div> <div>arn:aws:sts::667288089457:assumed-role/vncLabo/user358197-Javier_Cerejido_Cortes</div> </div> <div> <div>Hora de creación</div> <div>Thu Oct 24 2024 10:04:51 GMT-0200 (hora de verano de Europa central)</div> </div> <div> <div>Solicitar instancias de spot</div> <div>No</div> </div> <div> <div>Editar</div> </div> </div> </div>			
<div> <div>Ver detalles en la consola de la plantilla de lanzamiento</div> </div>			
<div> <div>Red</div> <div> <div> <div>Zonas de disponibilidad</div> <div>us-east-1a, us-east-1b</div> </div> <div> <div>ID de subred</div> <div>subnet-0f94b18090c6b66, subnet-0c3910423cfa6e80b</div> </div> </div> </div> <div> <div>Editar</div> </div>			

Autoescaping creado

Balanceros de carga

Con el software se puede administrar cualquier tipo de aplicación de carga repartida en varios servidores.

Filtrar aplicaciones de carga

Número	Número de DNS	Estatus	ID de VPC	Zona de disponibilidad	Tipo	Fecha creada
1	cafe2d-207741018.us-east-1.elb.amazonaws.com	Activo	vpc-0f86c5d2e45c	2 Zona de disponibilidad	aplicación	24 de octubre de 2024, 18:35 UTC-05:00


X


Equilibrador de carga ALB


Detalles | Agentes de resolución y reglas | Páginas de red | Reglas de recursos - reglas | Seguridad | Monitorización | Integraciones | Archivos | Etiquetas


Detalles

Tipo de equilibrador de carga: Aplicación	Estatus: Activo	VPC: vpc-0f86c5d2e45c	Tipo de dirección IP del equilibrador de carga: IPv4
Ubicoma: Internaz Baling	Zona de disponibilidad: 2 Zona de disponibilidad us-east-1-elb-us-east-1	Zona de disponibilidad: 2 Zona de disponibilidad us-east-1-elb-us-east-1	Fecha creada: 24 de octubre de 2024, 18:35 UTC-05:00
DNS del equilibrador de carga: cafe2d-207741018.us-east-1.elb.amazonaws.com	Número de DNS AWS: cafe2d-207741018.us-east-1.elb.amazonaws.com (Región: us-east-1)		

 No security

 label-206775413-us-east-1-ib.amazonaws.com

 Load Test

 RDS

Meta-Data	Value
InstanceId	i-08578b9ee124c4dc
Availability Zone	us-east-1g

Current CPU Load: 4%

Alarmas en Cloudwatch

CloudWatch > Alarmas

Alarmas (2)

☐ Ocultar alarmas de Auto Scaling

Acciones ▾

Estado de la alarma: cualquier ▾

Tipo de alarma: cualquiera ▾

Estado de las acciones: cualquier ▾

< 1 > ⌵

<input type="checkbox"/>	Nombre ▾	Estado ▾	Última actualización del estado (Local) ▾	Condiciones	Acciones ▾
<input type="checkbox"/>	TargetTracking-Grupo de Auto Escalado Lab-AlarmHigh-885964a-90ab-4245-9136-9c0093a82ad	⚠ En modo alarma	2024-10-24 10:26:40	CPUUtilization > 60 para 3 puntos de datos dentro de 3 minutos	🟢 Acciones habilitadas
<input type="checkbox"/>	TargetTracking-Grupo de Auto Escalado Lab-AlarmLow-0b2680ef-b488-466f-ac8d-c1ba5cb4220c	🟢 CORRECTO	2024-10-24 10:24:16	CPUUtilization < 45 para 15 puntos de datos dentro de 15 minutos	🟢 Acciones habilitadas