

ISC	Infraestructuras y Servicios Cloud
25/26	Casos de Estudio
GIIA	T1-T7

I. Preguntas de Razonamiento y Justificación (RA: Justificación de Conceptos)

Estas preguntas buscan la comprensión profunda de los conceptos clave y la capacidad de argumentación del estudiante.

#	Pregunta de Razonamiento y Justificación
1	Impacto de la Virtualización: Explique por qué la virtualización es el pilar tecnológico fundamental que hizo posible el modelo de negocio y técnico del Cloud Computing.
2	CapEx vs. OpEx: Describa la diferencia entre Gasto de Capital (CapEx) y Gasto Operativo (OpEx). Justifique por qué el Cloud Computing cambia la balanza de CapEx a OpEx para la mayoría de las empresas.
3	Seguridad en la Nube: Explique el concepto de Modelo de Responsabilidad Compartida en el contexto de la seguridad en la nube. Proporcione un ejemplo de una tarea de seguridad que sea responsabilidad del Cliente y otra del Proveedor (para un servicio IaaS).
4	Definición de Cloud: Mencione y explique brevemente al menos cuatro características esenciales del Cloud Computing según el NIST (por ejemplo, Autoservicio Bajo Demanda, Elasticidad Rápida, etc.).
5	Latencia y Nube: Para una aplicación crítica (ej. juegos online o trading de alta frecuencia) donde la latencia es mínima, ¿por qué la nube

#	Pregunta de Razonamiento y Justificación
	pública podría no ser la solución óptima? ¿Qué alternativa relacionada con la infraestructura de nube podría considerarse (ej. Cloud de Borde)?

SOLUCIONES

1. Impacto de la Virtualización

La **virtualización** es el pilar fundamental porque permite la **abstracción del hardware**. Gracias a ella, un proveedor de nube puede dividir un único servidor físico en múltiples recursos independientes (VMs), permitiendo la multi-tenencia, la escalabilidad inmediata y el modelo de pago por uso que define al Cloud Computing.

2. CapEx vs. OpEx

- **CapEx (Gasto de Capital):** Inversión inicial fuerte en activos físicos (servidores, racks) que se deprecian con el tiempo.
- **OpEx (Gasto Operativo):** Costos continuos por el uso de un servicio o producto.
- **Justificación:** El Cloud cambia la balanza a **OpEx** porque las empresas ya no compran hardware, sino que pagan mensualmente solo por los recursos que consumen, eliminando la barrera de entrada de la inversión inicial.

3. Modelo de Responsabilidad Compartida

Este modelo delimita las tareas de seguridad entre el proveedor y el cliente.

- **Responsabilidad del Proveedor (IaaS):** Seguridad "de" la nube (protección física de los centros de datos, mantenimiento del hardware y del hipervisor).
- **Responsabilidad del Cliente (IaaS):** Seguridad "en" la nube (actualización del Sistema Operativo instalado, configuración de firewalls/Security Groups y cifrado de sus datos).

4. Características Esenciales del NIST

1. **Autoservicio bajo demanda:** El usuario puede aprovisionar recursos automáticamente sin interacción humana con el proveedor.
2. **Elasticidad rápida:** Los recursos pueden aumentar o disminuir rápidamente para ajustarse a la demanda.
3. **Acceso ubicuo a la red:** Los servicios están disponibles a través de la red mediante mecanismos estándar (internet).
4. **Servicio medido:** El uso de recursos se monitorea y factura de forma transparente (pago por uso).

5. Latencia y Nube

En aplicaciones de trading o juegos online, la nube pública puede no ser óptima debido al retardo (latencia) que introduce la distancia física a los centros de datos regionales.

- **Alternativa: Cloud de Borde (Edge Computing)**, que acerca el procesamiento al usuario final para minimizar los tiempos de respuesta.

II. Preguntas de Conceptos Básicos y Definiciones (RA: Comprensión de Fundamentos)

Preguntas directas para verificar la comprensión de la terminología básica del Cloud Computing.

#	Pregunta Básica	Respuesta Esperada (Breve)
6	IaaS: Defina Infraestructura como Servicio (IaaS) y cite un ejemplo de recurso proporcionado.	Provisión de recursos fundamentales de computación (VMs, redes, almacenamiento) a través de Internet. Ej: Máquinas Virtuales (VM).
7	PaaS: ¿Cuál es la principal diferencia entre PaaS y IaaS en términos de gestión del sistema operativo?	En PaaS, el proveedor de la nube gestiona el Sistema Operativo y el <i>runtime</i> . En IaaS, el cliente lo gestiona.
8	SaaS: Defina Software como Servicio (SaaS) y proporcione un ejemplo de uso diario.	Software alojado y gestionado centralmente por un proveedor, accesible vía web o app. Ej: Gmail, Office 365, Salesforce.
9	Nube Híbrida: ¿Qué combinación de modelos de despliegue define la Nube Híbrida ?	La combinación de una Nube Privada y una o más Nubes Públicas , que operan de forma conjunta.
10	Contenedores: ¿Cómo ayudan los contenedores (ej.	Empaquetan la aplicación con todas sus dependencias,

#	Pregunta Básica	Respuesta Esperada (Breve)
	Docker, Kubernetes) a la portabilidad de las aplicaciones en la nube?	asegurando que se ejecute de manera consistente en cualquier entorno (local, público, híbrido).

III. Casos de Estudio para la Toma de Decisiones (RA: Análisis y Decisión)

Estos casos requieren que el estudiante evalúe el escenario, compare opciones y recomiende la solución en la nube más adecuada (IaaS, PaaS, SaaS, Nube Pública, Privada, Híbrida).

#	Escenario del Caso de Estudio	Preguntas de Razonamiento y Justificación
1	Migración de un CRM Una PYME ha estado usando un software CRM (Customer Relationship Management) instalado en un servidor local. Los costos de mantenimiento y las fallas de disponibilidad son un problema creciente. El CEO considera migrarlo a la nube.	¿Qué modelo de servicio (IaaS, PaaS, SaaS) es el más recomendable para esta migración, considerando la necesidad de una gestión mínima? Justifique por qué la opción elegida es superior a las otras dos en este contexto.
2	Startup de Desarrollo de Apps Una startup necesita un entorno donde sus desarrolladores puedan construir, probar y desplegar rápidamente nuevas aplicaciones web sin preocuparse por la gestión de sistemas operativos o servidores. La previsión de crecimiento es alta, pero variable.	¿Qué modelo de servicio (IaaS, PaaS, SaaS) le proporcionaría la mayor agilidad y menor sobrecarga operativa? ¿Qué implicación tiene la elasticidad de la nube para su modelo de negocio?
3	Almacenamiento de Datos Sensibles Una institución financiera maneja una gran	¿Debería optar por una Nube Pública, Privada o Híbrida ? ¿Qué consideraciones de

#	Escenario del Caso de Estudio	Preguntas de Razonamiento y Justificación
	cantidad de datos de clientes altamente sensibles (regulados por normativas estrictas) y necesita una solución de almacenamiento con alta seguridad y soberanía del dato, pero también requiere acceso a herramientas de análisis de datos a gran escala.	cumplimiento normativo son cruciales para su justificación?
4	Picos de Tráfico de E-commerce Un comercio electrónico experimenta picos de tráfico extremos durante el "Black Friday" y la temporada navideña, pero el resto del año su tráfico es moderado. Mantener infraestructura propia para el pico es costoso e ineficiente.	¿Cómo resuelve el Cloud Computing este problema? Mencione el concepto clave. ¿Qué estrategia de infraestructura (por ejemplo, subir solo la web temporalmente) recomendaría para optimizar los costes?
5	Infraestructura de Laboratorio Virtual Una universidad necesita proporcionar a 500 estudiantes acceso a máquinas virtuales con software especializado (CAD, análisis de datos) solo durante 4 meses al año. Necesitan un aprovisionamiento y desaprovisionamiento rápidos.	¿Qué modelo de servicio (IaaS, PaaS, SaaS) y modelo de despliegue es el más adecuado para esta necesidad temporal y a gran escala? Justifique el impacto del modelo de pago por uso en el presupuesto de la universidad.
6	Disaster Recovery (DR) Una empresa con infraestructura local quiere implementar un plan de recuperación ante desastres sin invertir en un segundo centro de datos físico.	¿Cómo facilita la nube la implementación de una estrategia de Disaster Recovery ? ¿Qué modelo de despliegue (público/híbrido)

#	Escenario del Caso de Estudio	Preguntas de Razonamiento y Justificación
		es típicamente el más rentable para este propósito?
7	Desarrollo vs. Producción Un equipo de IT necesita entornos idénticos para el desarrollo de software y para la producción real, pero quiere que el entorno de desarrollo sea más económico cuando no se está usando.	Describa cómo los principios de la nube (por ejemplo, Infraestructura como Código - IaC) pueden ayudar a mantener la paridad de entornos (dev/prod) y a la vez controlar los costes en el entorno de desarrollo.
8	Proveedor Único vs. Múltiple Una empresa está evaluando si usar un solo proveedor de nube (mono-cloud) o varios (multi-cloud) para sus diferentes aplicaciones. Les preocupa la dependencia de un solo proveedor y la portabilidad.	Enumere una ventaja y una desventaja del enfoque Multi-Cloud . ¿Qué implicación tiene el concepto de interoperabilidad ?
9	Refactorización de una Aplicación Legacy Una aplicación antigua monolítica tiene problemas de escalabilidad. El equipo de IT considera reescribirla usando microservicios y contenedores.	¿Por qué un entorno PaaS/Contenedores como Servicio es más adecuado para una arquitectura de microservicios que un IaaS tradicional? ¿Cómo se relaciona esto con el concepto de Nativo en la Nube ?
10	Monitoreo y Optimización de Costes Un equipo de Finanzas observa que su factura de la nube está creciendo sin control, aunque la carga de trabajo no ha aumentado significativamente.	¿Qué prácticas de FinOps o Gestión de Costes en la Nube deberían aplicar? Mencione al menos dos estrategias clave (ej. dimensionamiento, autoscaling, reservas).

1. Migración de un CRM

- **Modelo recomendado: SaaS** (Software como Servicio).
- **Justificación:** Es la opción ideal para una gestión mínima, ya que el proveedor se encarga de todo el mantenimiento, actualizaciones y disponibilidad. Supera a **IaaS** y **PaaS** porque elimina la necesidad de gestionar servidores, bases de datos o el sistema operativo, permitiendo a la PYME centrarse únicamente en el uso de la herramienta.

2. Startup de Desarrollo de Apps

- **Modelo de servicio: PaaS** (Plataforma como Servicio).
- **Agilidad:** Proporciona mayor agilidad al permitir que los desarrolladores desplieguen código sin gestionar la infraestructura subyacente.
- **Elasticidad:** Permite que la infraestructura crezca o se reduzca automáticamente según la demanda variable, asegurando que la startup solo pague por lo que necesita mientras garantiza el rendimiento para sus usuarios.

3. Almacenamiento de Datos Sensibles

- **Modelo de despliegue: Nube Híbrida.**
- **Justificación:** Permite mantener los datos altamente sensibles en una **Nube Privada** local para cumplir con normativas de soberanía y seguridad estricta, mientras se utiliza la **Nube Pública** para acceder a herramientas de análisis de datos a gran escala que serían muy costosas de replicar localmente.
- **Cumplimiento:** Es crucial garantizar que los datos no salgan de la jurisdicción legal requerida por la normativa financiera.

4. Picos de Tráfico de E-commerce

- **Concepto clave: Elasticidad Rápida.**
- **Resolución:** El Cloud permite ajustar los recursos hacia arriba o hacia abajo automáticamente según la demanda.
- **Estrategia recomendada:** Implementar una arquitectura de **Cloud Bursting** (Nube Híbrida), donde la infraestructura local maneja el tráfico moderado y se recurre a la nube pública solo para absorber los picos de eventos como el Black Friday, optimizando así los costes anuales.

5. Infraestructura de Laboratorio Virtual

- **Modelo de servicio: IaaS** (proporcionando VMs).
- **Modelo de despliegue: Nube Pública.**
- **Impacto presupuestario:** El modelo de **pago por uso** es fundamental, ya que la universidad solo paga por los recursos consumidos durante los 4 meses de clase, evitando la inversión en hardware que quedaría ocioso el resto del año.

6. Disaster Recovery (DR)

- **Facilitación:** La nube permite replicar datos y aplicaciones en regiones geográficamente distantes sin el coste de construir un centro de datos físico secundario.
- **Modelo rentable:** La **Nube Pública** o un enfoque **Híbrido** es lo más rentable, ya que se puede mantener una infraestructura mínima "en espera" y escalarla solo en caso de un desastre real.

7. Desarrollo vs. Producción

- **Infraestructura como Código (IaC):** Permite definir los entornos mediante archivos de configuración, garantizando que el entorno de desarrollo sea una copia exacta del de producción (paridad).
- **Control de costes:** Con IaC se pueden programar scripts para "apagar" o destruir el entorno de desarrollo automáticamente cuando no se esté usando (noches o fines de semana) y recrearlo en minutos cuando sea necesario.

8. Proveedor Único vs. Múltiple

- **Ventaja Multi-Cloud:** Evita la dependencia de un solo proveedor (*vendor lock-in*) y aumenta la resiliencia.
- **Desventaja Multi-Cloud:** Aumenta la complejidad operativa y los costes de transferencia de datos entre nubes.
- **Interoperabilidad:** Es la capacidad de que las aplicaciones y datos funcionen correctamente a través de diferentes plataformas de nube sin necesidad de cambios drásticos.

9. Refactorización de una Aplicación Legacy

- **Justificación de PaaS/Contenedores:** Estos servicios (como EKS o ECS) gestionan automáticamente la orquestación, el escalado y el despliegue de microservicios, algo que en IaaS sería complejo de configurar manualmente.

- **Nativo en la Nube:** Se refiere a aplicaciones diseñadas desde cero para aprovechar las capacidades de escalabilidad, resiliencia y gestión automática de la nube.

10. Monitoreo y Optimización de Costes

- **Estrategias FinOps:**
 1. **Dimensionamiento (Rightsizing):** Ajustar el tamaño de las instancias a la carga real de trabajo para no pagar por recursos que no se usan.
 2. **Reservas e Instancias Spot:** Usar instancias reservadas para cargas constantes o instancias **Spot** para procesos que toleren interrupciones, reduciendo drásticamente la factura.
-

20 Casos de Estudio y Ejercicios de ISC

I. Casos de Decisión Arquitectónica (Diseño detallado y Componentes) (1-10)

Estos ejercicios simulan escenarios del mundo real y requieren que los estudiantes tomen decisiones técnicas específicas (VMs, Contenedores, *Serverless*, Redes, Almacenamiento, etc.) y justifiquen su elección.

1. Caso: API de Inferencia de Modelo Ligero

Una aplicación necesita una API REST para clasificar texto. El tráfico es muy bajo la mayor parte del tiempo, pero tiene picos impredecibles. El modelo de NLP es ligero (< 250 MB).

- **Decisión y Justificación (RA):** ¿Qué **modelo de cómputo** (VM/Contenedor/Serverless FaaS) es el más costo-eficiente, escalable y simple de gestionar? Nombre el servicio de AWS más adecuado (por ejemplo, AWS Lambda).

- **Pregunta de Razonamiento:** Explique la principal ventaja de este modelo de cómputo elegido frente a una Máquina Virtual (EC2) si la carga de trabajo está inactiva el 90% del tiempo.

SOLUCIÓN

- **Modelo de cómputo: Serverless (FaaS).**
- **Servicio AWS: AWS Lambda.**
- **Ventaja:** Si la carga está inactiva el 90% del tiempo, el coste es **cero**, mientras que una VM (EC2) cobraría por cada hora que esté encendida, aunque no procese peticiones.

2. Caso: Entrenamiento de LLM con Presupuesto Limitado

Un equipo va a entrenar un Transformer muy grande que tarda 40 horas en un clúster de GPUs. El entrenamiento se puede reanudar desde *checkpoints* si es interrumpido.

- **Decisión y Justificación (RA):** ¿Qué familia de instancias EC2 (ej. Serie P o G) y qué modelo de precios (On-Demand, Reservadas o Spot) ofrece la mejor relación rendimiento-coste para esta carga de trabajo tolerante a fallos?
- **Pregunta de Razonamiento:** ¿Cómo mitiga la empresa el riesgo de interrupción inherente al modelo de precios elegido? (Pista: Mencionando el mecanismo de persistencia de datos).

SOLUCIÓN

- **Instancia:** Familia P o G (GPU).
- **Modelo de precios:** Instancias Spot.
- **Mitigación de riesgo:** Al ser tolerante a fallos, se usan **Checkpoints** guardados en almacenamiento persistente (como S3 o EBS) para reanudar el entrenamiento si AWS retira la instancia Spot.

3. Caso: Sistema de Ciberseguridad Híbrido

Una empresa quiere usar la potencia de la nube pública para un análisis masivo de logs, pero las máquinas que alojan los logs críticos están en su centro de datos privado por regulaciones estrictas.

- **Decisión y Justificación (RA):** ¿Qué **modelo de despliegue** (Público/Privado/Híbrido) es obligatorio? ¿Qué servicio de **conectividad híbrida** (VPN/Direct Connect) elegiría para la transferencia inicial de terabytes de logs, priorizando la seguridad y el ancho de banda consistente?
- **Pregunta de Razonamiento:** Explique la diferencia de coste y rendimiento clave entre la VPN Site-to-Site y la Conexión Dedicada (Direct Connect/ExpressRoute).

SOLUCIÓN

- **Despliegue: Híbrido.**
- **Conectividad: Direct Connect.**
- **Diferencia:** La VPN es más barata pero usa internet público (rendimiento variable); Direct Connect es una conexión física dedicada con ancho de banda consistente y mayor seguridad.

4. Caso: Data Lake para Imágenes Satelitales

Una agencia almacena petabytes de imágenes (datos no estructurados) que se usan poco después del mes 3, pero deben conservarse 10 años por motivos regulatorios.

- **Decisión y Justificación (RA):** ¿Qué **servicio de almacenamiento de objetos** debe ser el pilar del Data Lake (ej. Amazon S3)? Proponga una **política de ciclo de vida** (Life Cycle Policy) para optimizar el coste a largo plazo (ej. moviendo los datos a Glacier).
- **Pregunta de Razonamiento:** Explique la diferencia entre la clase S3 Standard y Glacier Deep Archive en términos de coste y tiempo de recuperación (latencia).

SOLUCIÓN

- **Servicio: Amazon S3.**
- **Ciclo de vida:** Mover datos a **S3 Glacier Deep Archive** después de 3 meses.
- **Diferencia:** S3 Standard tiene acceso instantáneo y mayor coste; Glacier Deep Archive es extremadamente barato pero la recuperación puede tardar horas (alta latencia)

5. Caso: Aplicación Web de Alto Tráfico Global

Un servicio de IA tiene usuarios en América y Europa. Necesita responder con la mínima latencia. El modelo está desplegado en clústeres idénticos en dos regiones diferentes.

- **Decisión y Justificación (RA):** ¿Qué **servicio DNS inteligente** (ej. Route 53) debe usarse para dirigir a los usuarios a la región con la mejor conexión? ¿Qué **tipo de enrutamiento** (ej. Latencia o Geográfico) es el más apropiado para minimizar el tiempo de respuesta percibido?
- **Pregunta de Razonamiento:** ¿Qué otro servicio de red global (CDN, ej. CloudFront) podría utilizarse junto con esta estrategia para reducir aún más la latencia de entrega de la interfaz web estática?

SOLUCIÓN

- **Servicio DNS: Route 53.**
- **Enrutamiento: Por Latencia.**
- **Mejora adicional:** Usar **Amazon CloudFront (CDN)** para cachear contenido estático en ubicaciones de borde cerca de los usuarios.

6. Caso: Base de Datos de Metadatos para MLOps

Un equipo utiliza un *Feature Store* y necesita una base de datos con alta consistencia (**ACID**) y alta disponibilidad para almacenar el linaje y los metadatos de los modelos entrenados.

- **Decisión y Justificación (RA):** ¿Qué **modelo de BD** (SQL/Relacional o NoSQL) y **servicio gestionado (PaaS)** (ej. Amazon RDS/Aurora) es el más adecuado dada la necesidad de garantías ACID?
- **Pregunta de Razonamiento:** Describa la diferencia entre **Multi-AZ** y **Read Replicas** en RDS y explique cuál de los dos mecanismos prioriza la *disponibilidad* y la *tolerancia a fallos* frente a la *escalabilidad de lectura*.

SOLUCIÓN

- **Modelo: Relacional (SQL)** gestionado con **Amazon Aurora o RDS**.
- **Disponibilidad: Multi-AZ** garantiza alta disponibilidad y conmutación por error automática; las **Read Replicas** solo mejoran el rendimiento de lectura (escalabilidad).

7. Caso: Aislamiento en Red de 3 Capas

Diseñe el aislamiento de red de una arquitectura de 3 capas: **1. Balanceador** (público), **2. API de Inferencia** (privado, Flask) y **3. Base de Datos** (privada, RDS).

- **Decisión y Justificación (RA):** Explique por qué los componentes 2 y 3 deben estar en **subredes privadas** (Private Subnets). Demuestre cómo un **Security Group** (SG) para la BBDD asegura que solo la API pueda acceder, aplicando el Principio de Mínimo Privilegio.
- **Pregunta de Razonamiento:** Si las instancias API de la capa 2 necesitan descargar librerías de internet, ¿qué componente de red es necesario en la subred pública para que puedan iniciar ese tráfico saliente? (Pista: NAT Gateway).

SOLUCIÓN

- **Privacidad:** Las capas 2 (API) y 3 (DB) deben estar en subredes privadas para que no sean accesibles directamente desde internet, reduciendo la superficie de ataque.
- **Componente de salida:** Se necesita un **NAT Gateway** en la subred pública para que las instancias privadas descarguen librerías sin ser expuestas.

8. Caso: Caché para Resultados de Inferencia

Una API de inferencia tiene un tiempo de respuesta de 600ms, pero recibe muchas peticiones idénticas. El resultado es un JSON. Se necesita reducir la latencia a sub-milisegundos para respuestas repetidas.

- **Decisión y Justificación (RA):** ¿Qué **tipo de Base de Datos NoSQL** (Clave-Valor, Grafo, En Memoria) es la solución ideal para actuar como una caché de resultados? Mencione el servicio de AWS más adecuado (ej. ElastiCache) y justifique la elección en términos de **latencia**.
- **Pregunta de Razonamiento:** ¿Por qué la latencia de un sistema In-Memory (En Memoria) es superior a la de una BD NoSQL basada en disco, como DynamoDB?

SOLUCIÓN

- **Tipo de BD: En Memoria (NoSQL).**
- **Servicio: Amazon ElastiCache.**
- **Justificación:** La latencia es superior (microsegundos) porque los datos están en la RAM, evitando los tiempos de búsqueda física de los discos (incluso SSD) de bases como DynamoDB.

9. Caso: Orquestación de Microservicios Portables

Una aplicación de IA usa 10 microservicios en contenedores Docker y el equipo prioriza la **portabilidad** para evitar la dependencia de un solo proveedor de nube.

- **Decisión y Justificación (RA):** ¿Qué **tecnología de orquestación** es el estándar de facto en la industria (Kubernetes) y por qué es la base para una estrategia *Multi-Cloud*? Nombre el servicio equivalente gestionado en AWS (**EKS**) y Azure (**AKS**).
- **Pregunta de Razonamiento:** Explique la diferencia principal entre un servicio nativo de AWS como ECS y un servicio basado en Kubernetes como EKS en términos de **ecosistema** y **dependencia de proveedor**.

SOLUCIÓN

- **Tecnología: Kubernetes.**
- **Servicios: EKS (AWS) y AKS (Azure).**
- **Diferencia:** ECS es propietario de AWS (mayor dependencia/vendor lock-in); EKS usa el estándar Kubernetes, facilitando mover la carga a otros proveedores o entornos locales.

10. Caso: Procesamiento de Archivos por Eventos

Cada vez que un científico de datos sube un nuevo fichero CSV a un bucket S3, debe activarse automáticamente una función que lo preprocese y lance un trabajo de entrenamiento.

- **Decisión y Justificación (RA):** Describa el flujo de trabajo Cloud ideal.
¿Qué servicio de S3 se usa para iniciar el proceso (Event Notifications)?
¿Qué **modelo de cómputo** (VM/FaaS) es mejor para la tarea de preprocesamiento, dado que es corta y asíncrona?
- **Pregunta de Razonamiento:** ¿Qué concepto de AWS Lambda se utiliza para empaquetar grandes librerías como Pandas o NumPy para compartirlas entre múltiples funciones (Layers)?

SOLUCIÓN

- **Flujo:** S3 subida -> S3 Event Notifications -> AWS Lambda.
- **Modelo:** FaaS (Serverless) es mejor por ser una tarea corta, puntual y asíncrona.
- **Concepto:** Lambda Layers para empaquetar librerías como Pandas.

II. Preguntas de Justificación y Razonamiento (Conceptos y Modelos) (11-15)

Estas preguntas se centran en la comprensión de los modelos, principios y la sinergia entre Cloud e IA.

11. Modelos de Servicio (IaaS vs. PaaS vs. SaaS)

Dibuje la pirámide de los 3 modelos de servicio. Explique dónde traza la línea el **Modelo de Responsabilidad Compartida** para la gestión del **Sistema Operativo** en cada uno de ellos.

Gestión del Sistema Operativo (SO) según el Modelo de Responsabilidad Compartida:

- **IaaS (Infraestructura):** La línea de responsabilidad se traza **encima del hipervisor**. El proveedor es responsable del hardware físico, pero el **cliente** es responsable de instalar, parchear y mantener el Sistema Operativo.
- **PaaS (Plataforma):** La línea se traza **debajo del runtime**. El **proveedor** gestiona por completo el Sistema Operativo, el middleware y el entorno de ejecución. El cliente solo se preocupa por su aplicación y datos.
- **SaaS (Software):** La línea se traza en la **interfaz de usuario**. El **proveedor** gestiona todo el stack tecnológico, incluido el SO. El cliente solo es responsable de la configuración del software y sus propios datos

12. Ventaja de Docker para IA

El problema clásico es "Funciona en mi máquina". Explique cómo el concepto de **Contenedores Docker** resuelve los desafíos de **Consistencia de Entornos** y **Reproducibilidad de Experimentos** para un científico de datos. El uso de contenedores Docker es fundamental para el flujo de trabajo de un científico de datos por las siguientes razones:

- **Consistencia de Entornos:** Los contenedores empaquetan el código junto con sus dependencias, librerías (como TensorFlow o PyTorch) y configuraciones específicas. Esto garantiza que la aplicación se ejecute de forma idéntica en el portátil del científico, en el clúster de entrenamiento y en el servidor de producción.
- **Reproducibilidad de Experimentos:** Al definir el entorno en un archivo (Dockerfile), cualquier miembro del equipo puede recrear exactamente el mismo entorno de experimentación. Esto elimina el problema de "en mi máquina funciona" provocado por versiones de librerías incompatibles o variables de entorno mal configuradas.

13. Seguridad de IAM (Roles vs. Keys)

Un ingeniero de MLOps debe configurar una instancia EC2 para que acceda a un bucket S3. Explique por qué usar un **Rol IAM** es la mejor práctica de seguridad y por qué guardar las **Claves de Acceso** es inseguro.

Para que una instancia EC2 acceda a S3, la mejor práctica es asignar un Rol IAM a la instancia.

- Por qué el Rol es mejor: Los roles utilizan credenciales temporales que AWS rota automáticamente. No es necesario configurar credenciales dentro del código ni de la instancia, eliminando el riesgo de robo de identidad si alguien accede al servidor.
- Por qué las Claves (Keys) son inseguras: Guardar *Access Keys* (claves permanentes) en archivos de configuración o código es peligroso. Si estas claves se filtran (por ejemplo, al subirlas a un repositorio de código), un atacante tendría acceso total y permanente a los recursos hasta que las claves sean revocadas manualmente.

14. Data Warehouse vs. Data Lake

Un analista necesita hacer un análisis exploratorio rápido sobre 2 años de datos brutos (JSONs) en S3. ¿Qué arquitectura está optimizada para esta tarea (OLAP)? ¿Cuál permite la filosofía *Schema-on-read*?

Ante la necesidad de analizar 2 años de datos brutos (JSONs) almacenados en S3:

- Arquitectura optimizada (OLAP): El Data Warehouse está tradicionalmente optimizado para tareas OLAP (procesamiento analítico en línea), pero requiere datos estructurados y procesados previamente.
- Filosofía *Schema-on-read*: El Data Lake es la arquitectura que permite esta filosofía. A diferencia del almacén tradicional, en un Data Lake puedes almacenar los datos en su formato original (bruto como JSON) y solo definir la estructura (esquema) en el momento en que decides leerlos para el análisis.

15. Sinergia Cloud e IA

Justifique la afirmación: "La Inteligencia Artificial moderna sería impracticable a escala sin el Cloud Computing". Mencione al menos dos pilares de hardware/servicio que la nube democratiza.

La Inteligencia Artificial moderna depende del Cloud Computing debido a la escala masiva de recursos necesarios para el entrenamiento de modelos complejos. Sin la nube, solo grandes corporaciones con centros de datos propios podrían innovar en IA.

Pilares que la nube democratiza:

1. **Computación Acelerada (GPUs/TPUs):** Permite alquilar por horas hardware extremadamente caro y especializado que es esencial para el Deep Learning.
2. **Almacenamiento Masivo y Escalable:** Proporciona la capacidad de almacenar petabytes de datos necesarios para alimentar los modelos de IA de forma económica y accesible.

III. Preguntas de Conceptos Básicos y Definiciones (Fundamentos) (16-20)

Preguntas rápidas para evaluar la comprensión de la terminología fundamental vista en las sesiones.

16. Principio NIST Clave

¿Cuál de los 5 principios del NIST (National Institute of Standards and Technology) define la capacidad del Cloud Computing para ajustar los recursos rápidamente hacia arriba o hacia abajo según la demanda?

Elasticidad Rápida.

17. Componentes de una AMI

Mencione dos de los componentes principales que se empaquetan dentro de una **AMI (Amazon Machine Image)**.

Sistema Operativo, software preinstalado y permisos de lanzamiento.

18. Almacenamiento de Bloque (EBS)

¿A qué recurso de cómputo se **adjunta** un volumen de **EBS** (Elastic Block Store)? ¿Está ligado a una **Región** o a una **Zona de Disponibilidad (AZ)**?

Se adjunta a una instancia EC2 y está ligado a una Zona de Disponibilidad (AZ) específica.

19. Seguridad VPC

¿Cuál es la principal diferencia funcional entre un **Security Group (SG)** y un **Network ACL (NACL)**?

Los **Security Groups** son a nivel de instancia (con estado/stateful); las **Network ACLs** son a nivel de subred (sin estado/stateless).

20. Servicios Serverless FaaS (AWS)

¿Cuál es el servicio de **Funciones como Servicio (FaaS)** de AWS? ¿Qué significa que tu código es invocado por un **Trigger**?

Lambda es el servicio FaaS. Un **Trigger** significa que el código se ejecuta automáticamente en respuesta a un evento externo (ej. subir un archivo).