



training and
certification

**AWS Academy Cloud Architecting (ES)
Module 02 Student Guide
Versión 2.0.2**

200-ACACAD-20-ES-SG

© 2021 Amazon Web Services, Inc. o sus empresas afiliadas.
Todos los derechos reservados.

Este contenido no puede reproducirse ni redistribuirse, total ni parcialmente,
sin el permiso previo por escrito de Amazon Web Services, Inc. Queda prohibida
la copia, el préstamo o la venta de carácter comercial.

Para correcciones o comentarios relacionados con el curso,
contacte con nosotros en:
<https://support.aws.amazon.com/#/contacts/aws-training>.

Todas las marcas comerciales pertenecen a sus propietarios.

Contenido

Módulo 2: Presentación de la arquitectura en la nube

4

AWS Academy Cloud Architecting

Módulo 2: Presentación de la arquitectura en la nube



Bienvenido al Módulo 2: Presentación de la arquitectura en la nube.

Información general sobre el módulo



Secciones

1. ¿Qué es la arquitectura en la nube?
2. Marco de Buena Arquitectura de Amazon Web Services (AWS)
3. Prácticas recomendadas para crear soluciones en AWS
4. Infraestructura global de AWS



Revisión de conocimientos

Este módulo incluye las siguientes secciones:

1. ¿Qué es la arquitectura en la nube?
2. Marco de Buena Arquitectura de Amazon Web Services (AWS)
3. Prácticas recomendadas para crear soluciones en AWS
4. Infraestructura global de AWS

Al final de este módulo, se le solicitará que complete una evaluación de conocimientos que probará su nivel de comprensión de los conceptos clave analizados en este módulo.

Objetivos del módulo

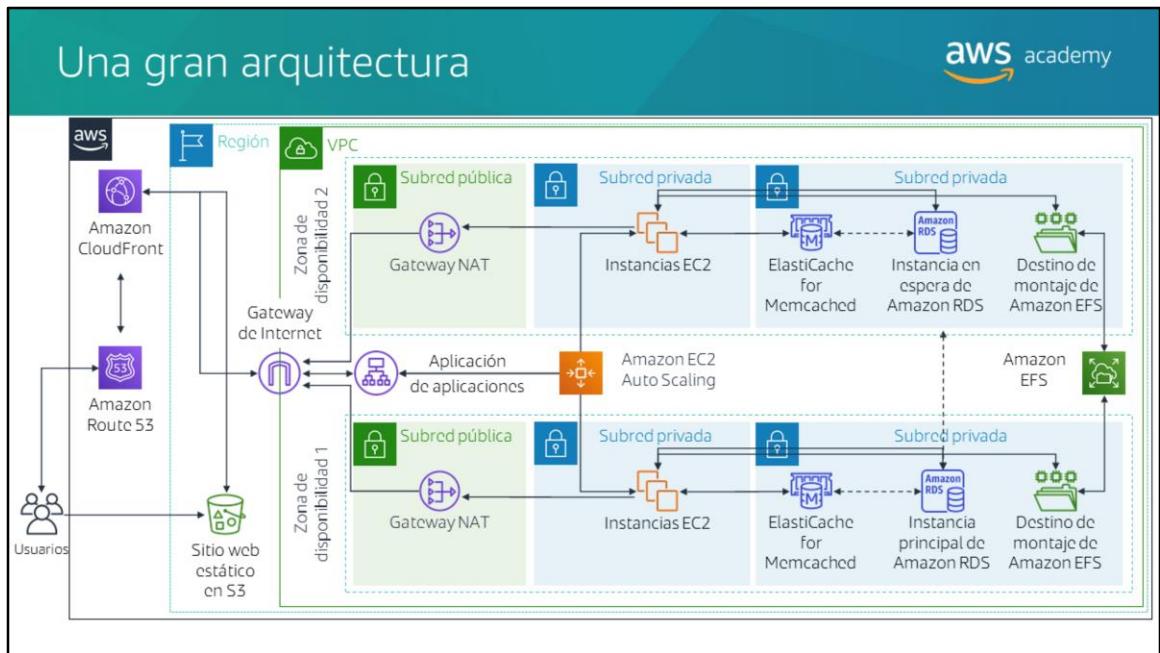


Una vez finalizado este módulo, debería ser capaz de lo siguiente:

- Definir la arquitectura en la nube
- Describir cómo diseñar y evaluar arquitecturas a través de AWS Well-Architected Framework
- Explicar las prácticas recomendadas para crear soluciones en AWS
- Describir cómo tomar decisiones fundamentadas acerca de dónde colocar los recursos de AWS

Una vez finalizado este módulo, debería ser capaz de lo siguiente:

- Definir la arquitectura en la nube
- Describir cómo diseñar y evaluar arquitecturas a través de AWS Well-Architected Framework
- Explicar las prácticas recomendadas para crear soluciones en AWS
- Describir cómo tomar decisiones fundamentadas acerca de dónde colocar los recursos de AWS



Al final de este curso, habrá aprendido sobre todos los componentes en este diagrama de arquitectura. También debería poder desarrollar sus propias arquitecturas de soluciones que sean tan grandes y robustas como la de este ejemplo. Verá que este diagrama se repite al comienzo de la mayoría de los módulos del curso. A medida que se presenten en el curso, se revelarán nuevos componentes del diagrama.

Módulo 2: Presentación de la arquitectura en la nube

Sección 1: ¿Qué es la arquitectura en la nube?



Introducción de la Sección 1: ¿Qué es la arquitectura en la nube?

Necesidad de arquitectura



Alrededor del año 2000, Amazon tenía dificultades para lograr que su nuevo sitio web de compras tuviera una alta disponibilidad y escalabilidad.

Para comprender qué es la arquitectura en la nube y por qué es importante, primero debe considerar un ejemplo de cómo es el desarrollo de software en su ausencia.

Alrededor del año 2000, Amazon intentaba crear un servicio de comercio electrónico que permitiera a los vendedores externos crear sus propios sitios de compras en línea sobre el motor de comercio electrónico de Amazon. La empresa tenía dificultades para lograr que su nuevo sitio web de compras tuviera una alta disponibilidad y escalabilidad.

Orígenes de AWS



- En ese momento, según Andy Jassy, CEO de AWS, las herramientas de comercio electrónico de Amazon eran “un completo desorden”.
 - Las aplicaciones y las arquitecturas se desarrollaban **sin una planificación adecuada**.
 - Era **difícil separar los servicios** entre sí.
- **Solución:** Amazon creó un conjunto de API bien documentadas que, a su vez, se convirtió en el estándar para el desarrollo de servicios.

En una [entrevista de TechCrunch sobre el origen de AWS](#), el director ejecutivo (CEO) de AWS, Andy Jassy, dijo que, al principio, las herramientas de comercio electrónico de Amazon eran un “verdadero desorden”. Las aplicaciones y las arquitecturas se desarrollaban sin una planificación adecuada. Jassy también dijo que era “un gran desafío separar los distintos servicios para hacer una plataforma de desarrollo centralizada”.

La solución a este problema fue crear un conjunto de interfaces de programación de aplicaciones (API) bien documentadas para organizar el entorno de desarrollo.

Los problemas continuaron

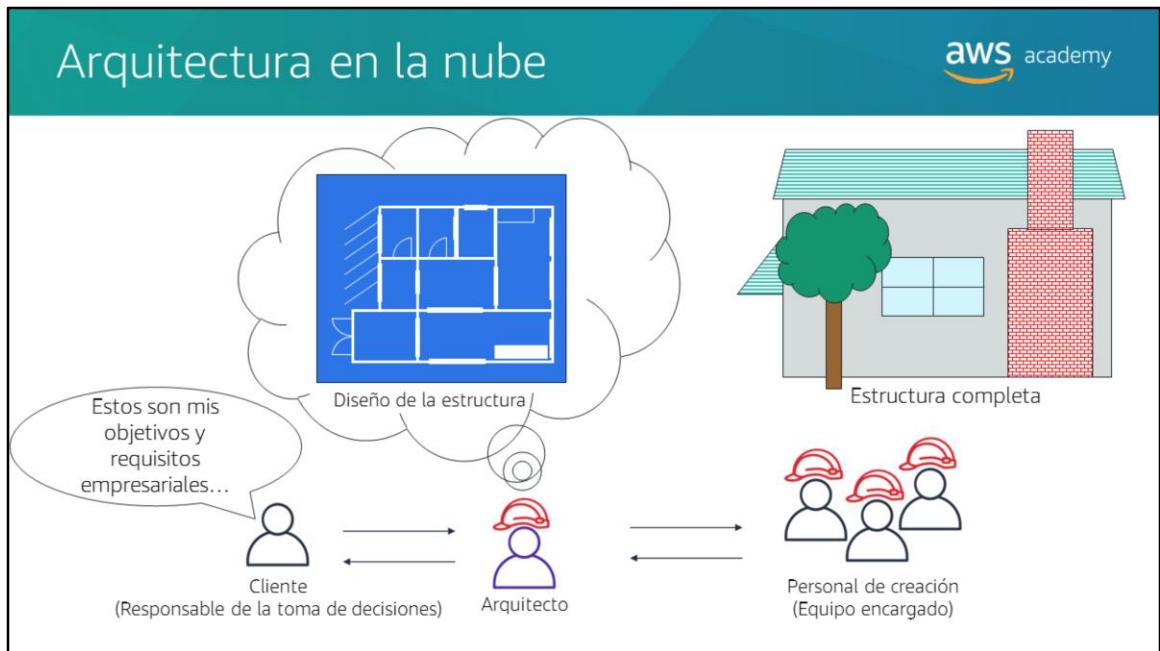


- Amazon todavía tenía dificultades para crear aplicaciones de forma rápida.
 - Los componentes de bases de datos, informática y almacenamiento tardaban **3 meses** en crearse.
 - Cada equipo creaba sus propios recursos, **sin planificar la escalabilidad ni la reutilización**.
- **Solución:** Amazon creó servicios internos para obtener arquitecturas de alta disponibilidad, escalabilidad y fiabilidad sobre su infraestructura. En 2006, comenzó a vender estos servicios como AWS.

Sin embargo, Amazon todavía tenía dificultades para crear aplicaciones rápidamente a medida que la empresa crecía y se contrataba a más ingenieros de software:

- Se tardó **3 meses** en crear los componentes de base de datos, informática y almacenamiento para un proyecto cuya duración total se estimaba que sería de tan solo **3 meses**.
- Cada equipo construyó sus propios recursos sin planificar la escalabilidad o la reutilización.

La solución fue crear servicios internos para diseñar arquitecturas de alta disponibilidad, escalabilidad y fiabilidad sobre la infraestructura de Amazon. En 2006, comenzó a vender estos servicios como AWS.



Entonces, ¿qué es la arquitectura en la nube? La arquitectura en la nube es la práctica de aplicar características de la nube a una solución que utiliza servicios y características en la nube para satisfacer las necesidades técnicas y los casos de uso empresariales de una organización. Una solución es similar a un plano para un edificio.

Los sistemas de software requieren que los arquitectos administren su tamaño y complejidad.

Los arquitectos en la nube se encargan de las siguientes tareas:

- Interactuar con los responsables de la toma de decisiones para identificar los objetivos empresariales y las capacidades que deben mejorarse.
- Garantizar la alineación entre los productos tecnológicos finales de una solución y los objetivos empresariales.
- Trabajar con los equipos de entrega que implementan la solución para garantizar que las características tecnológicas sean adecuadas.

Tener sistemas bien diseñados aumenta la probabilidad de que los productos tecnológicos ayuden a cumplir los objetivos empresariales.

Aprendizajes clave de la Sección 1



The slide features a large, ornate key lying on a teal-colored wooden surface. A small white rectangular card is tucked under the head of the key, with the word "Takeaway" written on it in a black, sans-serif font.

aws academy

- La arquitectura en la nube es la práctica de aplicar características de la nube a una solución que utiliza servicios y características en la nube para satisfacer las necesidades técnicas y los casos de uso empresariales de una organización
- Puede utilizar los servicios de AWS para crear arquitecturas de alta disponibilidad, escalabilidad y fiabilidad.

Estos son algunos de los aprendizajes clave de esta sección del módulo:

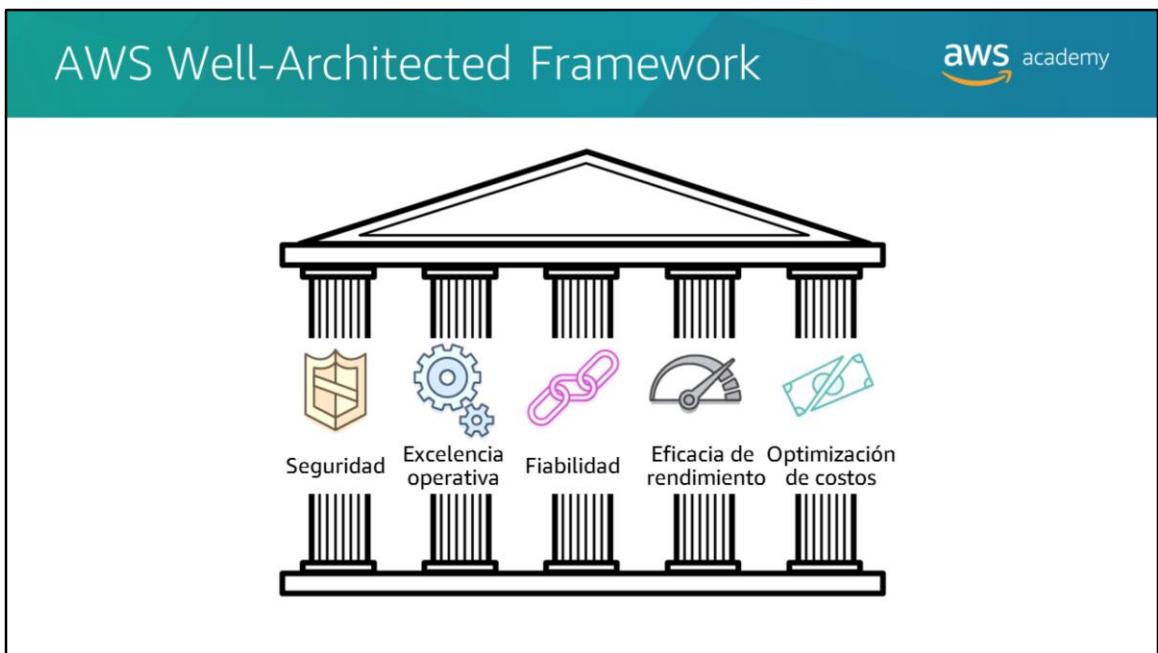
- La arquitectura en la nube es la práctica de aplicar características de la nube a una solución que utiliza servicios y características en la nube para satisfacer las necesidades técnicas y los casos de uso empresariales de una organización
- Puede utilizar los servicios de AWS para crear arquitecturas de alta disponibilidad, escalabilidad y fiabilidad.

Módulo 2: Presentación de la arquitectura en la nube

Sección 2: AWS Well-Architected Framework



Introducción de la Sección 2: AWS Well-Architected Framework



AWS Well-Architected Framework es una guía diseñada para ayudarlo a crear la infraestructura más segura, resiliente, eficiente y de alto rendimiento posible para sus aplicaciones y cargas de trabajo en la nube. Ofrece un enfoque uniforme para evaluar arquitecturas en la nube e instrucciones para ayudarlo a implementar diseños. Documenta una serie de preguntas básicas y prácticas recomendadas que le permiten comprender si una arquitectura determinada sigue las prácticas recomendadas de la nube. AWS desarrolló este marco después de revisar miles de arquitecturas de clientes en AWS.

AWS Well-Architected Framework comprende cinco pilares: excelencia operativa, seguridad, fiabilidad, eficacia del rendimiento y optimización de costos.



El pilar de seguridad trata sobre la capacidad de proteger la información, los sistemas y los activos, mientras que se ofrece valor de negocio mediante evaluaciones de riesgo y estrategias de mitigación.

Su arquitectura presentará una presencia de seguridad mucho más sólida si implementa una gran base de identidades, habilita la trazabilidad, aplica seguridad en todas las capas, automatiza las prácticas recomendadas de seguridad y protege los datos en tránsito y en reposo.

Para obtener más información acerca de las prácticas recomendadas de seguridad, consulte el [documento técnico del Pilar de seguridad](#).

Pilar de excelencia operativa



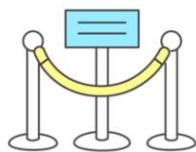
- Capacidad de ejecutar y monitorear sistemas
- Mejora continua de los procesos y los procedimientos de soporte



Implementado



Actualizado



Ejecutado

El pilar de excelencia operativa aborda la capacidad de ejecutar sistemas y obtener información sobre sus operaciones para ofrecer valor de negocio. También aborda la capacidad de mejorar continuamente los procesos y procedimientos de soporte.

Cuando diseña una carga de trabajo para operaciones, debe tener en cuenta cómo se implementará, actualizará y operará. Implemente prácticas de ingeniería que se alineen con las reducciones de defectos y las correcciones rápidas y seguras.

Habilite la observación con registros, instrumentación y métricas técnicas y empresariales para que pueda recibir información sobre lo que está sucediendo dentro de su arquitectura.

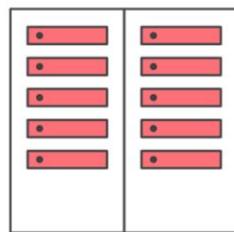
En AWS, puede ver toda su carga de trabajo (aplicaciones, infraestructura, políticas, gobernanza y operaciones) como código. Todo se puede definir y actualizar mediante el uso de código. Esto significa que puede aplicar la misma disciplina de ingeniería que utiliza para el código de las aplicaciones a cada elemento de su pila.

Para obtener más información acerca de las prácticas recomendadas de excelencia operativa, consulte el [documento técnico del Pilar de excelencia operativa](#).

Pilar de fiabilidad



- Recuperarse rápidamente de interrupciones en la infraestructura o el servicio
- Adquirir de manera dinámica recursos informáticos para satisfacer la demanda
- Mitigar interrupciones como las siguientes:
 - Errores de configuración
 - Problemas transitorios de red



En el pilar de fiabilidad, se aborda la capacidad de un sistema para recuperarse de interrupciones de infraestructura o servicio, y adquirir de manera dinámica recursos informáticos para satisfacer la demanda. También aborda la capacidad de un sistema para mitigar interrupciones como configuraciones erróneas o problemas transitorios de red.

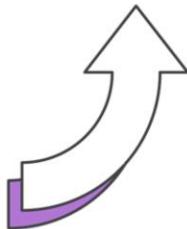
Puede ser difícil garantizar la fiabilidad en un entorno tradicional. Los problemas surgen de puntos únicos de error, falta de automatización y falta de elasticidad. Si aplica las prácticas recomendadas descritas en el pilar de fiabilidad, puede evitar muchos de estos problemas. Esto ayudará tanto a usted y a sus clientes a tener una arquitectura correctamente diseñada con respecto a la alta disponibilidad, la tolerancia a errores y la redundancia general.

Para obtener más información acerca de las prácticas recomendadas de fiabilidad, consulte el [documento técnico del Pilar de fiabilidad](#).

Pilar de eficacia del rendimiento



- Elegir recursos eficientes y mantener su eficiencia a medida que la demanda cambia
- Democratizar las tecnologías avanzadas
- Emplear compatibilidad mecánica



Cuando considera el rendimiento, debe maximizarlo mediante el uso eficiente de los recursos informáticos. También debe mantener esa eficiencia a medida que cambie la demanda.

Además, es importante democratizar las tecnologías avanzadas. En situaciones en las que resulta difícil implementar la tecnología por su cuenta, considere la posibilidad de utilizar un proveedor. Cuando el proveedor implementa la tecnología por usted, se hace cargo de la complejidad y el conocimiento, lo que libera a su equipo para que se concentre en hacer un trabajo de mayor valor agregado.

La *simpatía mecánica* se refiere a cuando se utiliza una herramienta o sistema con una comprensión de cómo funciona mejor. Utilice el enfoque tecnológico que se ajuste mejor a lo que intenta conseguir. Por ejemplo, tenga en cuenta los patrones de acceso a los datos cuando seleccione un enfoque de base de datos o de almacenamiento.

Para obtener más información acerca de las prácticas recomendadas de rendimiento, consulte el [documento técnico sobre el Pilar de eficacia del rendimiento](#).

Pilar de optimización de costos



- Medir la eficiencia
- Eliminar los gastos innecesarios
- Considerar el uso de servicios administrados



La optimización de costos es un requisito continuo de cualquier buen diseño de arquitectura. El proceso es iterativo y debe perfeccionarse y mejorarse durante toda la vida útil de su producción. Comprender cuán eficiente es su arquitectura actual en relación con sus objetivos puede eliminar gastos innecesarios. Considere el uso de servicios administrados, ya que estos operan a escala de nube y pueden ofrecer un costo menor por transacción o servicio.

Para obtener más información al respecto, consulte el [documento técnico del Pilar de optimización de costos](#).

AWS Well-Architected Tool



- Ayuda a revisar el estado de las cargas de trabajo y las compara con las prácticas recomendadas sobre arquitectura de AWS más recientes
- Brinda acceso a los conocimientos y a las prácticas recomendadas que utilizan los arquitectos de AWS, siempre que necesite estos recursos
- Ofrece un plan de acción con instrucciones paso a paso sobre cómo crear mejores cargas de trabajo para la nube
- Ofrece un proceso uniforme para revisar y medir las arquitecturas en la nube

Si desea ayuda para crear una solución bien diseñada, puede utilizar la herramienta AWS Well-Architected Tool. AWS Well-Architected Tool es una herramienta de autoservicio que le proporciona acceso bajo demanda a las prácticas recomendadas de AWS actuales. Estas prácticas recomendadas pueden ayudarle a crear una infraestructura de aplicaciones segura, de alto rendimiento, resiliente y eficiente en AWS.

AWS Well-Architected Tool lo ayuda a revisar el estado de las cargas de trabajo y las compara con las prácticas recomendadas sobre arquitectura de AWS. Le permite acceder a los conocimientos y a las prácticas recomendadas que utilizan los arquitectos de AWS, siempre que lo necesite.

Esta herramienta está disponible en la consola de administración de AWS. Usted define la carga de trabajo y responde a una serie de preguntas en las áreas de excelencia operativa, seguridad, fiabilidad, eficacia del rendimiento y optimización de costos. Luego, AWS Well-Architected Tool ofrece un plan de acción con instrucciones paso a paso sobre cómo mejorar la carga de trabajo para la nube.

AWS Well-Architected Tool ofrece un proceso coherente para revisar y medir las arquitecturas en la nube. Puede utilizar los resultados que proporciona la herramienta para identificar los próximos pasos a fin de mejorar, impulsar las decisiones sobre la arquitectura e integrar las consideraciones sobre la arquitectura a su proceso de control corporativo.

Para obtener más información acerca de AWS Well-Architected Tool, consulte el [sitio web de AWS Well-Architected Tool](#).



- AWS Well-Architected Framework ofrece un **enfoque uniforme** para evaluar las arquitecturas en la nube y las instrucciones, de forma que pueda implementar diseños
- AWS Well-Architected Framework comprende **cinco pilares**
- Cada pilar documenta **una serie de preguntas básicas** que le permiten comprender si una arquitectura determinada sigue las prácticas recomendadas de la nube
- **AWS Well-Architected Tool** permite revisar el estado de las cargas de trabajo y compararlas con las prácticas recomendadas más recientes de AWS en el ámbito de la arquitectura

Estos son algunos de los aprendizajes clave de esta sección del módulo:

- AWS Well-Architected Framework ofrece un enfoque uniforme para evaluar arquitecturas en la nube e instrucciones para ayudarlo a implementar diseños.
- AWS Well-Architected Framework comprende cinco pilares.
- Cada pilar documenta una serie de preguntas básicas que le permiten comprender si una arquitectura determinada se adecua a las prácticas recomendadas de la nube.
- AWS Well-Architected Tool le ayuda a revisar el estado de las cargas de trabajo y las compara con las prácticas recomendadas sobre arquitectura de AWS.

Módulo 2: Presentación de la arquitectura en la nube

Sección 3: Prácticas recomendadas para desarrollar soluciones en AWS



Presentación de la Sección 3: Prácticas recomendadas para desarrollar soluciones en AWS

Compensaciones de diseño



- Evalúe las compensaciones para que pueda seleccionar un enfoque óptimo
- Algunos ejemplos de compensaciones incluyen lo siguiente:
 - Intercambiar la uniformidad, la durabilidad y el espacio por el tiempo y la latencia para ofrecer un rendimiento mayor
 - Priorizar la velocidad de comercialización de las nuevas características sobre el costo
- Basar las decisiones de diseño en datos empíricos



A medida que diseñe una solución, piense cuidadosamente sobre las compensaciones para que pueda elegir un enfoque óptimo. Por ejemplo, puede intercambiar consistencia, durabilidad y espacio por tiempo y latencia para ofrecer un mayor rendimiento. O bien, puede priorizar la velocidad de comercialización por encima del costo.

Las compensaciones pueden aumentar el costo y la complejidad de su arquitectura, por lo que sus decisiones de diseño deben basarse en datos empíricos. Por ejemplo, es posible que necesite realizar pruebas de carga para asegurarse de que se obtenga un beneficio mensurable en el rendimiento. O bien, tal vez deba realizar evaluaciones con un punto de referencia para lograr la carga de trabajo más rentable a lo largo del tiempo. Cuando evalúa las mejoras relacionadas con el rendimiento, también querrá considerar cómo las elecciones de diseño de arquitectura afectarán a los clientes y a las eficiencias de las cargas de trabajo.

En esta sección, aprenderá sobre las prácticas recomendadas para diseñar soluciones en AWS. También aprenderá sobre prácticas no recomendadas (o diseños de soluciones malas) que debe evitar.

1. Habilitar la escalabilidad (1 de 2)



Asegúrese de que la arquitectura puede gestionar los cambios demandados.



Cuando ejecuta sus cargas de trabajo en la nube de AWS, puede escalar su infraestructura de forma rápida y proactiva. Asegúrese de implementar escalabilidad en cada capa de su infraestructura.

Para comprender la importancia de la escalabilidad, considere esta práctica no recomendada, donde el escalado se realiza de forma reactiva y manual.

En este caso, cuando los servidores de aplicaciones alcanzan la capacidad máxima, se impide que los usuarios accedan a la aplicación. Entonces, los administradores lanzan manualmente una o más instancias nuevas para administrar la carga. Desafortunadamente, una instancia tarda algunos minutos en estar disponible para su uso después de su lanzamiento. Esto aumenta el tiempo en que los usuarios no pueden acceder a la aplicación.

1. Habilitar la escalabilidad (2 de 2)



Asegúrese de que la arquitectura puede gestionar los cambios demandados.



Si habilita la escalabilidad, puede mejorar su diseño para anticipar la necesidad de más capacidad y entregarla antes de que sea demasiado tarde.

Por ejemplo, puede utilizar una solución de monitoreo como Amazon CloudWatch para detectar si la carga total de su flota de servidores ha alcanzado un límite específico. Puede definir este límite para *Permanecer por encima del 60 % de utilización de la CPU por más de 5 minutos*, o cualquier caso relacionado con el uso de los recursos. Con CloudWatch, también puede diseñar métricas personalizadas basadas en aplicaciones específicas que puedan activar el escalado de recursos que sea necesario.

Cuando se dispara la alarma, Auto Scaling de Amazon EC2 lanza una nueva instancia de forma inmediata. Esa instancia estará lista antes de que se alcance la capacidad, lo que proporciona una experiencia perfecta para los usuarios.

Idealmente, también debería diseñar su sistema para que *escale de forma descendente* cuando la demanda desciende, de modo que no ejecute (y pague) instancias que ya no necesite.

2. Automatizar entornos



Automatece el aprovisionamiento, la terminación y la configuración de los recursos cuando sea posible.

Práctica no recomendada



Práctica recomendada



AWS ofrece herramientas de monitoreo y automatización integradas para prácticamente todas las capas de su infraestructura. Aproveche estas herramientas para garantizar que su infraestructura responda rápidamente a los cambios.

Puede utilizar herramientas como CloudWatch y Amazon EC2 Auto Scaling para detectar recursos que no están en buen estado y automatizar el lanzamiento de recursos de reemplazo. También puede recibir notificaciones cuando cambien las asignaciones de recursos.

3. Tratar los recursos como desechables



Aproveche la naturaleza de aprovisionamiento dinámico de la informática en la nube.

Práctica no recomendada

- Con el tiempo, los servidores diferentes acaban teniendo configuraciones distintas
- Los recursos se ejecutan cuando no son necesarios
- Las direcciones IP no modificables impiden la flexibilidad
- Probar nuevas actualizaciones en el hardware que se está utilizando puede ser difícil y poco práctico.

Práctica recomendada

- Automatizar la implementación de nuevos recursos con configuraciones idénticas
- Terminar recursos que no están en uso
- Cambiar a las nuevas direcciones IP automáticamente
- Probar las actualizaciones en los nuevos recursos y, a continuación, sustituir los antiguos por los actualizados

La práctica recomendada de tratar a los recursos como desechables se refiere a la idea de pensar en su infraestructura como software en lugar de hardware.

Con el hardware, es fácil comprar componentes más específicos de los que necesita para que esté preparado para picos de uso. Eso es costoso e inflexible: es más difícil de actualizar debido a los costos irrecuperables.

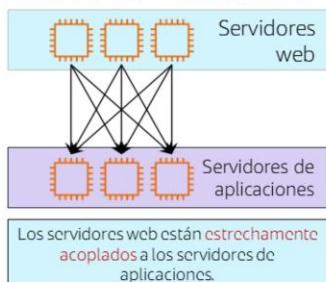
En su lugar, cuando trata sus recursos como desechables, migrar entre instancias u otros recursos discretos es bastante sencillo. Puede responder rápidamente a los cambios en las necesidades de capacidad, actualizar aplicaciones y administrar el software subyacente.

4. Utilizar componentes de bajo acoplamiento

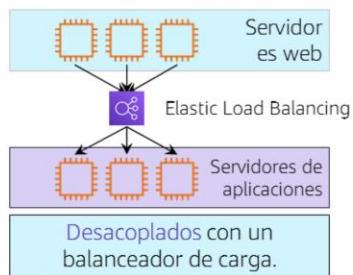


Diseñe las arquitecturas con componentes independientes.

Práctica no recomendada



Práctica recomendada



Las infraestructuras tradicionales tienen cadenas de servidores estrechamente integrados, cada uno con una función concreta. El problema es que cuando uno de esos componentes o capas cae, la interrupción del sistema puede ser fatal. También impide el escalado. Si agrega o quita servidores en una capa, también debe conectar todos los servidores de cada capa de conexión.

El ejemplo de la izquierda muestra un conjunto de servidores web y servidores de aplicaciones con un acoplamiento ajustado. Si un servidor de aplicaciones cae, se producirá un error porque los servidores web intentan conectarse a él y no pueden hacerlo.

El bajo acoplamiento le permite utilizar soluciones administradas como intermediarias entre las capas del sistema. Con este diseño, el intermediario maneja automáticamente tanto los fallos como el escalado de componentes o capas.

El ejemplo de la derecha muestra un balanceador de carga (en este caso, una instancia de Elastic Load Balancing) que dirige las solicitudes entre los servidores de aplicaciones y los servidores web. Si falla uno de los servidores de aplicaciones, el balanceador de carga dirigirá todo el tráfico automáticamente a los dos servidores en buen estado.

Las dos soluciones principales de desacoplamiento de componentes son los **balanceadores de carga** y las **colas de mensajes**.

5. Diseñar servicios, en lugar de servidores



Utilice la amplitud de los servicios de AWS. No limite su infraestructura a servidores.

Práctica no recomendada

- Las aplicaciones sencillas se ejecutan en servidores persistentes
- Las aplicaciones se comunican directamente las unas con las otras
- Los recursos web estáticos se almacenan localmente en instancias
- Los servidores de backend gestionan la autenticación y el almacenamiento de estado de los usuarios

Práctica recomendada

- Cuando sea apropiado, considere el uso de contenedores o una solución sin servidor
- Las colas de mensajes administran la comunicación entre aplicaciones.
- Los activos web estáticos se almacenan externamente, como en Amazon Simple Storage Service (Amazon S3)
- Los servicios administrados de AWS manejan la autenticación y el almacenamiento de estado de los usuarios

La siguiente práctica recomendada es diseñar servicios, no servidores. Aunque Amazon Elastic Compute Cloud (Amazon EC2) ofrece una gran flexibilidad para diseñar y configurar su solución, no siempre debe ser la primera (o única) solución que utilice para cada necesidad. En algunos casos, los contenedores o una solución sin servidor podrían ser más adecuados. Por lo tanto, es importante considerar cuáles son sus necesidades y cuál es la solución adecuada.

Con las soluciones sin servidor y los servicios administrados de AWS, no es necesario aprovisionar, configurar y administrar toda una instancia de Amazon EC2.

Las soluciones administradas que tienen un perfil más bajo y tienen más rendimiento pueden reemplazar las soluciones basadas en servidor a un costo menor. Algunos ejemplos son AWS Lambda, Amazon Simple Queue Service (Amazon SQS), Amazon DynamoDB, Elastic Load Balancing, Amazon Simple Email Service (Amazon SES) y Amazon Cognito.

6. Elegir la solución de base de datos correcta



Adecue la tecnología a la carga de trabajo, no al revés.

Tenga en cuenta lo siguiente:

- Las necesidades de lectura y escritura
- Los requisitos de almacenamiento totales
- El tamaño habitual de los objetos y cómo se accede a ellos
- Los requisitos de durabilidad
- Los requisitos de latencia
- La cantidad máxima de usuarios conectados al mismo tiempo
- La naturaleza de las consultas
- La intensidad necesaria de los controles de integridad

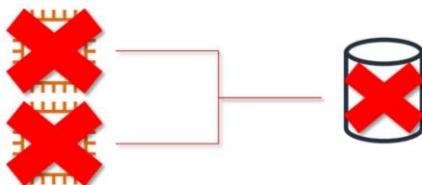
Es importante que elija la solución de base de datos adecuada. En los centros de datos y entornos en las instalaciones tradicionales, las limitaciones del hardware y las licencias disponibles pueden restringir su elección de una solución de almacenamiento de datos. AWS recomienda que elija un almacén de datos en función de sus necesidades para su entorno de aplicaciones.

7. Evitar los puntos únicos de error (1 de 2)



*Suponga que todo falla.
Luego, desarrolle los diseños
a partir de los errores.*

Siempre que sea posible, utilice la redundancia para evitar que un único punto derribe un sistema completo.



Práctica no recomendada

Servidores de aplicaciones

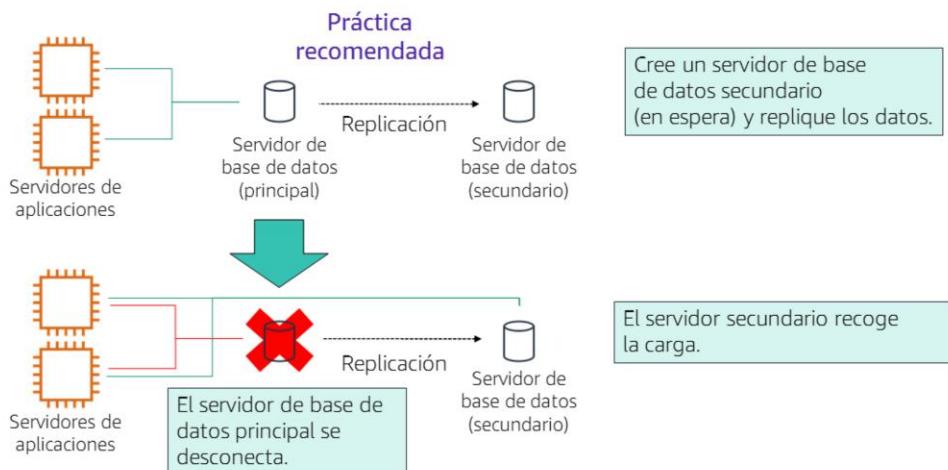
Servidor de base de datos

Elimine de las arquitecturas los puntos únicos de error siempre que sea posible. Esto no significa que siempre deba duplicar cada componente. Dependiendo de sus acuerdos de nivel de servicio (SLA) de tiempo de inactividad, puede utilizar soluciones automatizadas que solo lanzan componentes cuando es necesario. También puede utilizar un servicio administrado, en el que AWS reemplaza automáticamente el hardware subyacente que funciona mal.

Este sencillo sistema muestra dos servidores de aplicación conectados a un único servidor de base de datos. El servidor de la base de datos representa un punto único de error y se debe evitar. Cuando se cae, los servidores de aplicaciones también caen.

Los servidores de aplicaciones deben seguir funcionando incluso si el hardware físico subyacente falla, se elimina o se reemplaza.

7. Evitar los puntos únicos de error (2 de 2)



Una forma común de evitar puntos únicos de error es crear un servidor de base de datos secundario (en espera) y replicar los datos. De esta manera, si el servidor de base de datos principal se desconecta, el servidor secundario puede encargarse de la carga.

En este ejemplo, cuando la base de datos principal se desconecta, los servidores de aplicaciones envían de forma automática sus solicitudes a la base de datos secundaria. Esto también ejemplifica la práctica recomendada n.º 3: considerar los recursos como desechables y diseñar las aplicaciones de modo que admitan cambios en el hardware.

8. Optimizar para costos



Aproveche la flexibilidad de AWS para aumentar su rentabilidad.

Tenga en cuenta lo siguiente:

- ¿Mis recursos son del tamaño y tipo adecuados para el trabajo?
- ¿Qué métricas debo monitorear?
- ¿Cómo me aseguro de desactivar los recursos que no están en uso?
- ¿Con qué frecuencia tendré que utilizar este recurso?
- ¿Puedo sustituir alguno de mis servidores por servicios administrados?

La informática en la nube le permite cambiar la inversión de capital por gastos variables. La *inversión de capital (CapEx)* son los fondos que utiliza una empresa para adquirir, actualizar y mantener los activos físicos, como los inmuebles, los edificios industriales o el equipo. Bajo este modelo, pagará por los servidores del centro de datos, estén activos o no.

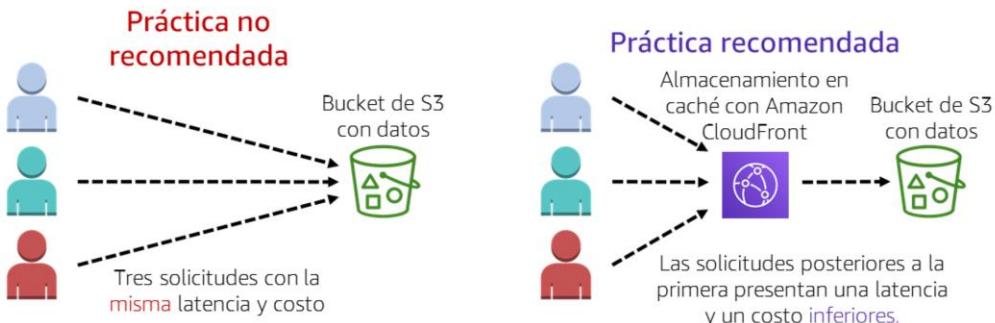
Por el contrario, los servicios de AWS utilizan un *modelo de costo variable*, lo que significa que solo pagará por los servicios individuales que necesita, mientras los utilice. Dentro de cada servicio, puede optimizar el costo. Muchos servicios ofrecen diferentes niveles de precios, modelos o configuraciones.

Tenga en cuenta que puede ser muy caro replicar una configuración de centro de datos en las instalaciones de servidores que se ejecutan las 24 horas del día, los 7 días de la semana en la nube. Por lo tanto, la mejor manera de construir su infraestructura, desde una perspectiva de costos, es solo aprovisionar los recursos que necesita y detener los servicios cuando no se están utilizando.

9. Utilizar el almacenamiento en caché



El almacenamiento en caché minimiza las operaciones de recuperación de datos redundantes, lo que mejora el rendimiento y el costo.



El *almacenamiento en caché* es una técnica para agilizar las solicitudes futuras y reducir el procesamiento de la red mediante el almacenamiento temporal de datos en una ubicación intermedia entre el solicitante y el almacenamiento permanente.

En el ejemplo de prácticas no recomendada, no se utiliza ningún servicio de almacenamiento en caché. Cuando alguien solicita un archivo de uno de los buckets de Amazon Simple Storage Service (Amazon S3), cada solicitud tarda la misma cantidad de tiempo en completarse y cada solicitud cuesta lo mismo.

En el ejemplo de práctica recomendada, la infraestructura utiliza Amazon CloudFront frente a Amazon S3 para proporcionar almacenamiento en caché. En este caso, la primera solicitud busca el archivo en Amazon CloudFront. Si no se encuentra, CloudFront solicita el archivo a Amazon S3. A continuación, CloudFront almacena una copia del archivo en una ubicación perimetral cercana al usuario y envía una copia al usuario que realizó la solicitud. Las solicitudes posteriores para el archivo se recuperan desde la ubicación perimetral (ahora más cercana) en CloudFront en lugar de Amazon S3.

Esto reduce la latencia y el costo porque, después de la primera solicitud, ya no pagará por que el archivo se transfiera fuera de Amazon S3.

10. Proteger toda la infraestructura



Implemente la seguridad en todas las capas de la infraestructura.

Tenga en cuenta lo siguiente:

- Aíslle partes de la infraestructura
- Cifre los datos en tránsito y en reposo
- Fuerce un control de acceso granular con el principio de privilegios mínimos
- Utilice la autenticación multifactor (MFA)
- Utilizar servicios administrados
- Registre el acceso a los recursos
- Automatice las implementaciones para mantener la consistencia de la seguridad

La seguridad no se trata sólo de atravesar el límite exterior de su infraestructura. También implica garantizar que sus entornos individuales y sus componentes estén protegidos unos de otros.

Por ejemplo, en Amazon EC2 puede configurar grupos de seguridad que le permiten especificar los puertos de las instancias que pueden enviar y recibir tráfico, así como dirigir ese tráfico. Los grupos de seguridad también pueden determinar de dónde puede venir o hacia dónde puede ir ese tráfico.

Puede utilizar los grupos de seguridad para reducir la probabilidad de que una amenaza para la seguridad de una instancia se extienda al resto de las instancias del entorno. Debe tomar precauciones similares con otros servicios. Las maneras concretas de implementar esta práctica recomendada se exponen a lo largo del curso.

Aprendizajes clave de la Sección 3



The image shows a large, antique-style key lying on a teal-colored wooden surface. Next to the key is a small, white rectangular card with the word "Takeaway" written on it in a black, cursive font.

aws academy

- A medida que diseña soluciones, evalúe las compensaciones y base sus decisiones en datos empíricos
- Para crear soluciones en AWS, siga estas prácticas recomendadas:
 - Habilitar la escalabilidad
 - Automatizar entornos
 - Tratar los recursos como desechables
 - Utilizar componentes de bajo acoplamiento
 - Diseñar servicios, en lugar de servidores
 - Elegir la solución de base de datos correcta
 - Evitar los puntos únicos de error
 - Optimización para costos
 - Utilizar el almacenamiento en caché
 - Proteger toda la infraestructura

Los aprendizajes clave de esta sección del módulo son los siguientes:

- A medida que diseña soluciones, evalúe las compensaciones y base sus decisiones en datos empíricos
- Para crear soluciones en AWS, siga estas prácticas recomendadas:
 - Habilitar la escalabilidad
 - Automatizar entornos
 - Tratar los recursos como desechables
 - Utilizar componentes de bajo acoplamiento
 - Diseñar servicios, en lugar de servidores
 - Elegir la solución de base de datos correcta
 - Evitar los puntos únicos de error
 - Optimización para costos
 - Utilizar el almacenamiento en caché
 - Proteger toda la infraestructura

Módulo 2: Presentación de la arquitectura en la nube

Sección 4: Infraestructura global de AWS

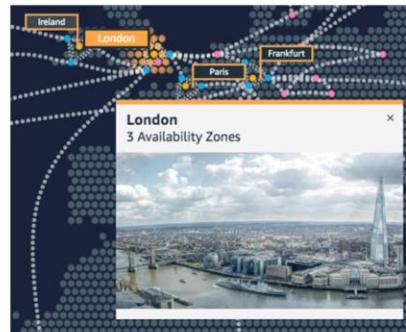


Presentación de la sección 4: Infraestructura global de AWS

Regiones de AWS



- Una **región de AWS** es una zona geográfica.
- Cada región de AWS consta de **dos o más zonas de disponibilidad**
- Para la comunicación entre regiones, se utiliza la infraestructura de **red troncal de AWS**
- Usted habilita y controla **la replicación de datos** entre regiones.



Ejemplo: región de Londres

La infraestructura de la nube de AWS se basa en regiones. AWS cuenta con 22 regiones en todo el mundo. Una *Región de AWS* es una ubicación geográfica física con una o más *Zonas de disponibilidad*. Las Zonas de disponibilidad, a su vez, constan de uno o más *centros de datos*.

Las regiones de AWS están conectadas a varios proveedores de Internet (ISP). Las regiones de AWS están conectadas a una red troncal global privada que proporciona menor costo y una latencia de red entre regiones más coherente en comparación con el Internet público.

Las regiones de AWS que se presentaron antes del 20 de marzo de 2019 están *habilitadas* de forma predeterminada. Las regiones que se presentaron después del 20 de marzo de 2019, como Asia Pacífico (Hong Kong) y Medio Oriente (Baréin), están *deshabilitadas* de forma predeterminada. Debe habilitar estas regiones para poder usarlas. Puede utilizar la consola de administración de AWS para habilitar o deshabilitar una región.

Algunas regiones tienen acceso restringido. Una cuenta de AWS (**China**) solo proporciona acceso a las regiones de Pekín y Ningxia. Para obtener más información acerca de AWS en China, consulte la [página AWS en China](#). La región aislada **AWS GovCloud (EE. UU.)** está diseñada para permitir a los organismos gubernamentales y a los clientes de EE. UU. transferir a la nube cargas de trabajo confidenciales y cumplir con sus requisitos normativos y de conformidad específicos.

Para lograr la tolerancia a errores y estabilidad, las regiones están aisladas unas de otras. Los recursos en una región no se replican automáticamente en otras regiones. Cuando almacena datos en una región específica, no se replica fuera de esa región. Si las necesidades del negocio así lo requieren, es su responsabilidad replicar los datos en las regiones. AWS ofrece información sobre el país y —si es aplicable— el estado donde reside cada región. Es responsable de elegir la región donde debería almacenar los datos, en función de sus requisitos de conformidad y latencia de red.

Los servicios y los productos de AWS se encuentran disponibles por región, por lo que es posible que no vea todas las regiones disponibles para un determinado servicio. Para obtener una lista de los servicios de AWS ofrecidos por región, consulte la [Tabla de regiones](#).

Para obtener más información acerca de la infraestructura de nube global de AWS, consulte el [sitio web Infraestructura global](#). Para obtener un mapa actual e interactivo de la infraestructura global de AWS, consulte el [mapa interactivo de infraestructura global de AWS](#).

Zonas de disponibilidad de AWS



- Cada zona de disponibilidad tiene estas características:
 - Está conformada por **uno o más** centros de datos.
 - Está diseñada para el **aislamiento de errores**
 - Está interconectada con otras zonas de disponibilidad de una región mediante enlaces **privados** de alta velocidad.
- Para ciertos servicios, puede elegir las zonas de disponibilidad
- Para lograr resiliencia, AWS recomienda la replicación entre las zonas de disponibilidad.



Cada región de AWS tiene dos o más ubicaciones aisladas, conocidas como *Zonas de disponibilidad*. Cada zona de disponibilidad abarca uno o más centros de datos, con algunas zonas de disponibilidad que tienen hasta seis centros de datos. Sin embargo, ningún centro de datos puede ser parte de dos zonas de disponibilidad.

Cada zona de disponibilidad está diseñada como una zona de error independiente. Esto significa que las zonas de disponibilidad están físicamente separadas dentro de una región metropolitana típica. También se ubican en llanuras de menor riesgo de inundación (la categorización específica de las zonas de inundación varía según la región). Además de tener sistemas de alimentación ininterrumpida discretos y generación de copia de seguridad en las instalaciones, la alimentación se suministra desde distintas redes de empresas de servicios públicos independientes para reducir todavía más los puntos de error únicos. Las zonas de disponibilidad están conectadas de forma redundante a varios proveedores de tránsito de nivel 1.

Una Zona de disponibilidad es el nivel de especificación más granular que puede realizar para determinados servicios, como Amazon EC2.

Usted es responsable de seleccionar las zonas de disponibilidad donde se encontrarán sus sistemas. Los sistemas pueden abarcar varias zonas de disponibilidad. Debe diseñar sus sistemas para que sobrevivan a un error temporal o prolongado de una zona de disponibilidad si ocurre un desastre. La distribución de las aplicaciones en varias zonas de disponibilidad permite que se mantengan resilientes en la mayoría de las situaciones de error, incluidos los desastres naturales o los errores de sistema.

AWS Local Zones



- Le permiten ejecutar partes de aplicaciones **sensibles a la latencia** más cerca de los usuarios finales y los recursos en una geografía específica
- Son una extensión de una región de AWS donde puede utilizar los servicios de AWS en una estrecha **proximidad geográfica a los usuarios finales**
- Le permiten colocar la informática, el almacenamiento, las bases de datos y otros servicios selectos de AWS más cerca de grandes centros de población, industria y TI donde no existe ninguna región en la actualidad
- Son administrados y respaldados por AWS
- **AWS Local Zone de Los Ángeles (LA)** está disponible por invitación

AWS Local Zones es un nuevo tipo de implementación de infraestructura de AWS que acercan los servicios de cómputo, almacenamiento, base de datos y otros servicios seleccionados a los centros poblacionales, industriales y de TI en donde no existe una región de AWS hoy en día. Con AWS Local Zones, puede ejecutar partes de aplicaciones sensibles a la latencia más cerca de los usuarios finales y los recursos en una geografía específica. Puede utilizar AWS Local Zones para ofrecer latencia de milisegundos de un solo dígito para casos de uso como creación de contenido multimedia y de entretenimiento, juegos en tiempo real, simulaciones de reservas, automatización de diseño electrónico y aprendizaje automático.

Cada ubicación de AWS Local Zone es una extensión de una región de AWS. Puede ejecutar aplicaciones sensibles a la latencia en una AWS Local Zone mediante servicios como Amazon EC2, Amazon Virtual Private Cloud (Amazon VPC), Amazon Elastic Block Store (Amazon EBS), Amazon FSx y Elastic Load Balancing en una estrecha proximidad geográfica con los usuarios finales. AWS Local Zones ofrecen una conexión segura y de gran ancho de banda entre las cargas de trabajo locales y las cargas de trabajo que se ejecutan en la región de AWS. Por lo tanto, AWS Local Zones permiten conectarse sin problemas a las demás cargas de trabajo que se ejecutan en AWS y conectarse a toda la gama de servicios dentro de la región a través de las mismas API y conjuntos de herramientas.

AWS Local Zones son administradas y admitidas por AWS, lo cual ofrece toda la elasticidad, escalabilidad y los beneficios de seguridad de la nube. Con AWS Local Zones, puede crear e implementar aplicaciones sensibles a la latencia más cerca de los usuarios finales mediante un conjunto coherente de servicios de AWS. También puede aumentar o disminuir los recursos, y pagar solo por los que utilice.

AWS Local Zone de Los Ángeles normalmente está disponible por invitación actualmente y se agregarán más zonas locales en el futuro.

Para obtener más información acerca de AWS Local Zones, consulte la [página AWS Local Zones](#).

Centros de datos de AWS



- En los centros de datos, se almacenan y se procesan los datos
- Un único centro de datos suele alojar cientos de miles de servidores
- Todos los centros de datos se encuentran en línea y a disposición de los clientes
- El equipo de red personalizado de AWS presenta las siguientes características:
 - Se obtiene de múltiples fabricantes de dispositivos originales (ODM)
 - Tiene una pila de protocolos de red personalizada



La base de la infraestructura de AWS son los centros de datos. No se especifica un centro de datos para la implementación de recursos. Sin embargo, el centro de datos es la ubicación donde residen los datos reales. Amazon opera centros de datos de vanguardia de alta disponibilidad. Aunque es poco frecuente, pueden ocurrir errores que afecten la disponibilidad de las instancias que están en la misma ubicación. Si aloja todas las instancias en una misma ubicación y se produce una falla en ella, ninguna de las instancias estará disponible.

Todos los centros de datos se encuentran en línea y a disposición de los clientes. En caso de fallo, los procesos automáticos redirigen el tráfico de datos de los clientes fuera de la zona afectada. Las aplicaciones principales se implementan con una configuración N+1 por lo que, si se produce un error en el centro de datos, se dispone de suficiente capacidad como para balancear la carga del tráfico a las ubicaciones restantes.

AWS utiliza equipos de red personalizados procedentes de varios fabricantes de dispositivos originales (ODM). Los ODM diseñan y fabrican productos en función de las especificaciones de una segunda empresa. Luego, la segunda empresa cambia la marca de los productos para la venta.

Para obtener más información acerca de los centros de datos de AWS, consulte [Conozca cómo protegemos los centros de datos de AWS desde su diseño](#).



Para entregar contenido a los usuarios finales con menor latencia, Amazon CloudFront utiliza una red global que incluye más de 200 puntos de presencia que se componen de ubicaciones perimetrales y cachés de borde regionales.

Las ubicaciones perimetrales se encuentran en América del Norte, Europa, Asia, Australia, América del Sur, Oriente Medio, África y China. Las ubicaciones perimetrales son compatibles con servicios de AWS, como Amazon Route 53 y Amazon CloudFront.

Las cachés de borde regionales se utilizan de forma predeterminada con Amazon CloudFront. Se utilizan cuando tiene contenido al que no se accede con la suficiente frecuencia como para permanecer en una ubicación perimetral. Las cachés de borde regionales absorben este contenido y proporcionan una alternativa para recuperarlo del servidor de origen.

Para obtener más información acerca de la infraestructura de Amazon CloudFront, consulte [la Infraestructura de Amazon CloudFront](#).

Aprendizajes clave de la sección 4



The image shows a large, antique-style key lying on a teal-colored wooden surface. A small, white rectangular tag is tied to the key's handle, with the word "Takeaway" written on it in a cursive font.

aws academy

- La infraestructura global de AWS se compone de **regiones, zonas de disponibilidad** y **ubicaciones perimetrales**
- La elección de una región se basa normalmente en los **requisitos de conformidad** o en la intención de **reducir la latencia**
- Cada **zona de disponibilidad** está separada físicamente de las demás y posee alimentación, redes y conectividad redundantes.
- **Las ubicaciones perimetrales** y **las cachés de borde regionales** mejoran el rendimiento almacenando en caché el contenido más cerca de los usuarios

Estos son algunos de los aprendizajes clave de esta sección del módulo:

- La infraestructura global de AWS se compone de regiones y zonas de disponibilidad
- La elección de una región se basa normalmente en los requisitos de conformidad o en la intención de reducir la latencia
- Cada zona de disponibilidad está separada físicamente de las demás y posee alimentación, redes y conectividad redundantes
- Las ubicaciones perimetrales y las cachés de borde regionales mejoran el rendimiento almacenando en caché el contenido más cerca de los usuarios

Módulo 2: Presentación de la arquitectura en la nube

Conclusión del módulo



Ahora es el momento de revisar el módulo y terminar con una evaluación de conocimientos.

Resumen del módulo



A modo de resumen, en este módulo, aprendió a hacer lo siguiente:

- Definir la arquitectura en la nube
- Describir cómo diseñar y evaluar arquitecturas a través de AWS Well-Architected Framework
- Explicar las prácticas recomendadas para crear soluciones en AWS
- Describir cómo tomar decisiones fundamentadas acerca de dónde colocar los recursos de AWS

A modo de resumen, en este módulo, aprendió a hacer lo siguiente:

- Definir la arquitectura en la nube
- Describir cómo diseñar y evaluar arquitecturas a través de AWS Well-Architected Framework
- Explicar las prácticas recomendadas para crear soluciones en AWS
- Describir cómo tomar decisiones fundamentadas acerca de dónde colocar los recursos de AWS

Completar la evaluación de conocimientos



Ha llegado el momento de completar la evaluación de conocimientos de este módulo.

Recursos adicionales



- [Página de Infraestructura global de AWS](#)
- [Mapa interactivo de la Infraestructura global de AWS](#)
- [Documento técnico de AWS Well-Architected Framework](#)
- [Documento técnico sobre el Pilar de seguridad](#)
- [Documento técnico sobre el Pilar de excelencia operativa](#)
- [Documento técnico sobre el Pilar de fiabilidad](#)
- [Documento técnico sobre el Pilar de la eficacia del rendimiento](#)
- [Documento técnico sobre el Pilar de optimización de costos](#)

Si desea obtener más información acerca de los temas que se trataron en este módulo, puede que le resulten útiles los siguientes recursos adicionales:

- [Página de Infraestructura global de AWS](#)
- [Mapa interactivo de la Infraestructura global de AWS](#)
- [Documento técnico de AWS Well-Architected Framework](#)
- [Documento técnico sobre el Pilar de seguridad](#)
- [Documento técnico sobre el Pilar de excelencia operativa](#)
- [Documento técnico sobre el Pilar de fiabilidad](#)
- [Documento técnico sobre el Pilar de la eficacia del rendimiento](#)
- [Documento técnico sobre el Pilar de optimización de costos](#)



Gracias

© 2020, Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados. Este contenido no puede reproducirse ni redistribuirse, total ni parcialmente, sin el permiso previo por escrito de Amazon Web Services, Inc. Queda prohibida la copia, el préstamo o la venta de carácter comercial. Envíenos sus correcciones o comentarios relacionados con el curso a: aws.course.feedback@amazon.com. Si tiene cualquier otra duda, contáctese con nosotros en: https://aws.amazon.com/contact-us/aws_training/. Todas las marcas comerciales pertenecen a sus propietarios.



Gracias por completar este módulo.