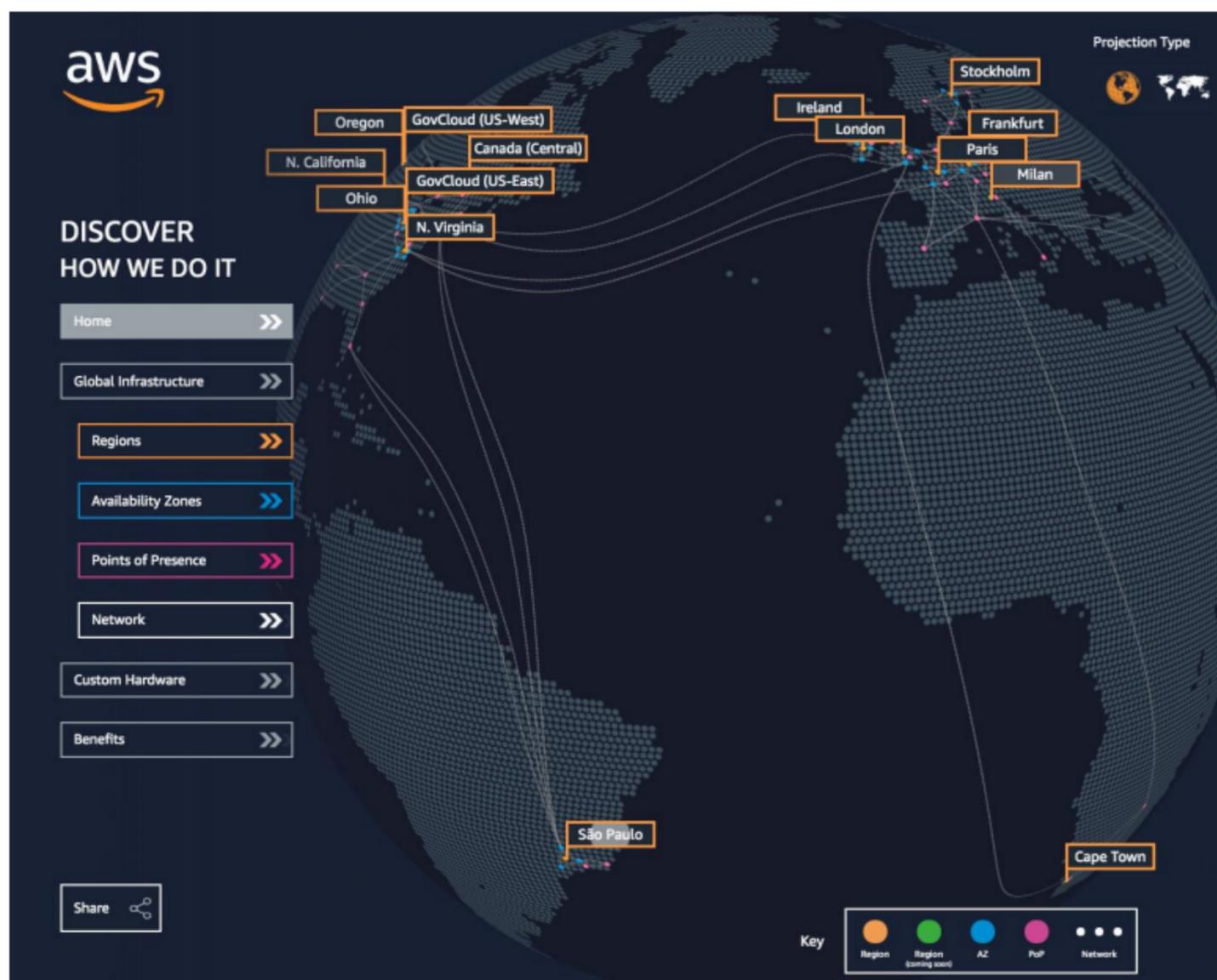


Lección 4

Donde colocar los servicios de AWS



Regiones de AWS

- Una región AWS es una zona geográfica
- Cada región de AWS consiste en dos o más zonas de disponibilidad
- La comunicación entre regiones utiliza la infraestructura de red troncal de AWS
- Usted habilita y controla la replicación de datos entre regiones

La infraestructura de la nube de AWS se basa en regiones. AWS cuenta con 33 regiones en todo el mundo. Una región de AWS es una ubicación geográfica física con dos 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 con varios proveedores de servicios de Internet (ISP, Internet Service Provider). Las regiones también están conectadas a una red troncal global privada, que proporciona un menor costo y una mayor uniformidad en la latencia de red entre regiones en comparación con la Internet pública.

Las regiones de AWS que se establecieron antes del 20 de marzo de 2019 están habilitadas de forma predeterminada. Las regiones que se establecieron 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 sobre AWS en China, consulte la página de 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 cumplimiento específicos.

Para obtener 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 corresponde, sobre el estado donde reside cada región. Usted es responsable de seleccionar la región donde debe almacenar los datos, en función de los requisitos de cumplimiento y latencia de red.

Los servicios y productos de AWS se encuentran disponibles por región, de modo que tal vez no vea todas las regiones disponibles para un determinado servicio. Para ver una lista de servicios de AWS ofrecidos por región, consulta la Tabla de regiones.

Para obtener más información sobre la infraestructura global de AWS, consulte el sitio web de infraestructura global. Para ver un mapa interactivo actual de la infraestructura global de AWS, consulte el Mapa interactivo de la infraestructura global de AWS.

Zonas de disponibilidad de AWS

- Cada zona de disponibilidad:
 - Se encuentra conformada por uno o más centros de datos
 - Está diseñada para el aislamiento de errores
 - Está interconectada con otras zonas de disponibilidad en una región por medio de enlaces privados de alta velocidad

- Para algunos servicios, puede elegir la zona de disponibilidad
- Para lograr resistencia, AWS recomienda la replicación entre 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 para constituir una zona de error independiente. Eso significa que las zonas de disponibilidad se encuentran separadas físicamente dentro de una región metropolitana típica. También se encuentran en llanuras con menor riesgo de inundación (la clasificación específica de las zonas de inundación varía según la región). Además de tener un suministro eléctrico ininterrumpido discreto e instalaciones de generación de energía de respaldo en el lugar, la alimentación proviene de redes eléctricas diferentes de empresas independientes de servicios públicos para reducir aún más los puntos únicos de error.

Las zonas de disponibilidad están conectadas de forma redundante a varios proveedores de tránsito de nivel 1. Una zona de disponibilidad es el nivel más granular de especificación que se puede hacer para algunos servicios, como Amazon EC2.



Usted es responsable de seleccionar las zonas de disponibilidad donde se encontrarán los sistemas. Los sistemas pueden abarcar varias zonas de disponibilidad. Debe diseñar los 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 les permite mantenerse resistentes en la mayoría de las situaciones de error, incluidos los desastres naturales o errores de sistema.

Zonas locales de AWS

- Le permite ejecutar partes de las aplicaciones sensibles a la latencia más cercana de los usuarios finales y los recursos en una ubicación geográfica específica
- Son una extensión de una región de AWS en la que puede usar servicios de AWS en proximidad geográfica de los usuarios finales
- Le permiten ubicar servicios de cómputo, almacenamiento, base de datos y otros servicios seleccionados de AWS más cerca de la población, industrias y los centros de TI donde no hay una región en la actualidad
- Las administra y reciben soporte de AWS
- La zona local de AWS de Los Ángeles (LA) está disponible por invitación

Las Zonas locales de AWS son un tipo nuevo de infraestructura de implementación que ubica cómputo, almacenamiento, bases de datos y otros servicios seleccionados de AWS cerca de grandes poblaciones, industrias y centros de TI donde no hay una región de AWS en la actualidad. Con las Zonas locales de AWS, puede ejecutar partes de las aplicaciones sensibles a la latencia más cerca de los usuarios finales y los recursos en una ubicación geográfica específica. Puede usar las Zonas locales de AWS para ofrecer latencia de milisegundos de un solo dígito para casos prácticos como la creación de contenido para medios y entretenimiento, juegos en tiempo real, simulaciones de yacimientos, Electronic Design Automation y machine learning.

Cada Zona local de AWS es una extensión de una región de AWS. Puede ejecutar aplicaciones sensibles a la latencia en una Zona local de AWS mediante servicios de AWS como Amazon EC2, Amazon Virtual Private Cloud (Amazon VPC), Amazon Elastic Block Store (Amazon EBS), Amazon FSx y Elastic Load Balancing en proximidad geográfica de los usuarios finales. Las Zonas locales de AWS ofrecen una conexión segura y de gran ancho de banda entre las cargas de trabajo locales y aquellas que se ejecutan en la región de AWS.

Las Zonas locales de AWS le permiten conectarse sin problemas con sus demás cargas de trabajo que se ejecutan en AWS, además de conectarse a la gama completa de servicios dentro de la región a través de las mismas API y conjuntos de herramientas.

Las Zonas locales de AWS las administra y reciben soporte de AWS, y le ofrecen toda la elasticidad, la escalabilidad y los beneficios de seguridad de la nube. Con las Zonas locales de AWS, puede crear e implementar aplicaciones sensibles a la latencia más cerca de los usuarios finales mediante un conjunto uniforme de servicios de AWS.

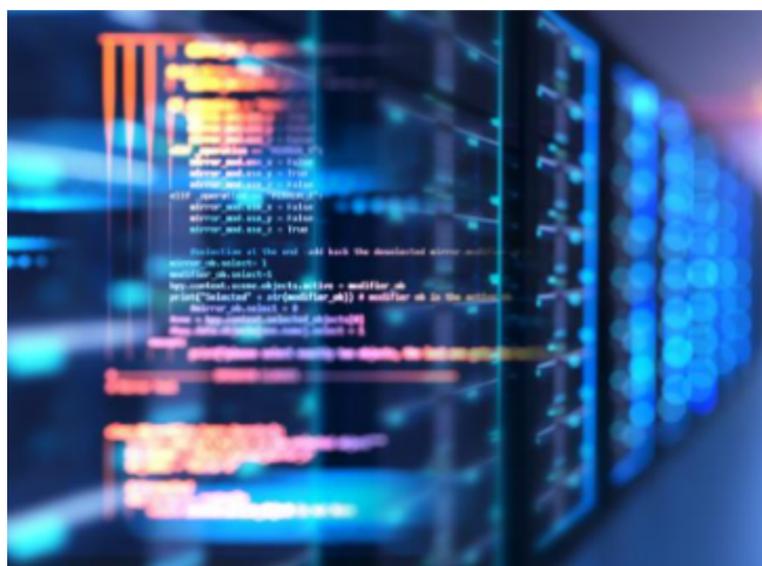
También puede escalar o reducir verticalmente y pagar solo por los recursos que utilice.

La Zona local de AWS de Los Ángeles está disponible de forma generalizada por invitación en la actualidad, y puede esperar que se agreguen más Zonas locales en el futuro.

Para obtener más información sobre las Zonas locales de AWS, consulte la página de [Zonas locales de AWS](#).

Centros de datos de AWS

- En los centros de datos se almacenan y se procesan los datos
- Un centro de datos suele alojar decenas de miles de servidores
- Todos los centros de datos se encuentran en línea y a disposición de los clientes
- Equipos de red personalizados de AWS:
 - Provenientes de varios ODM
 - Tienen una pila de protocolo de red personalizada

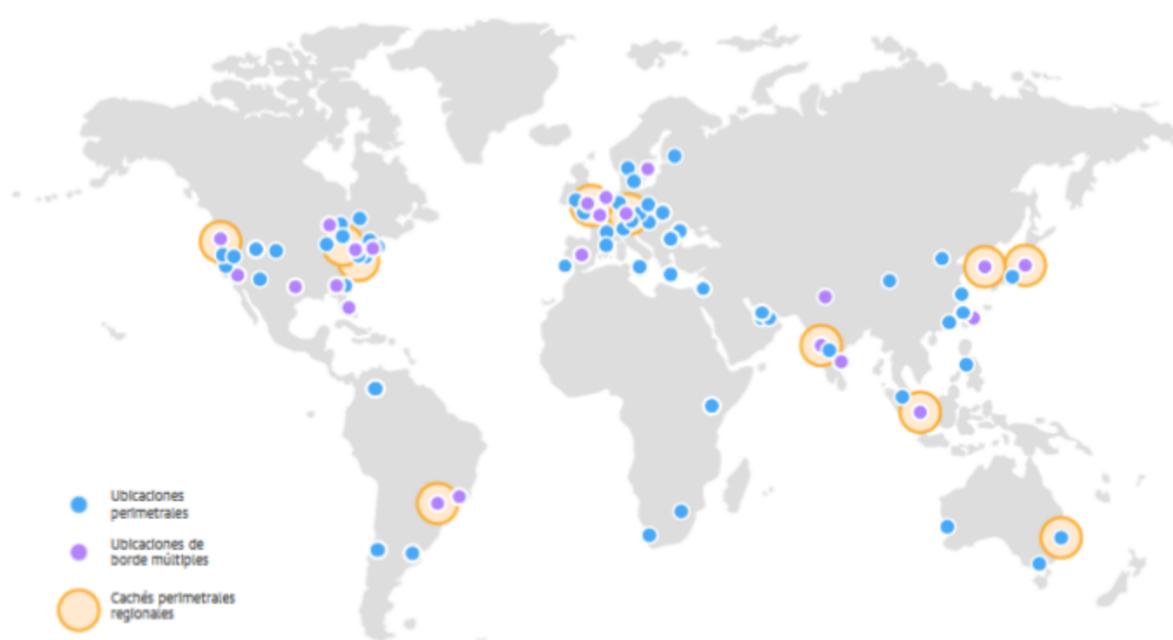


La base de la infraestructura de AWS son los centros de datos. Usted no especifica un centro de datos para la implementación de recursos. Sin embargo, un centro de datos es la ubicación donde residen los datos reales. Amazon opera centros de datos de vanguardia de alta disponibilidad. Aunque es poco frecuente, pueden ocurrir fallas 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. Si se produce un error, los procesos automatizados redirigen el tráfico de datos de los clientes fuera del área afectada. Las aplicaciones principales se implementan en una configuración N+1, de modo que, si se produce un error en un centro de datos, hay suficiente capacidad para poder equilibrar la carga del tráfico hacia los sitios restantes. AWS utiliza equipos de red personalizados provenientes de varios original device manufacturers (ODM, fabricantes de dispositivos originales). 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 sobre los centros de datos de AWS, consulte [Conozca cómo protegemos los centros de datos de AWS desde su diseño](#).

Puntos de presencia de AWS



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 perimetrales regionales.

Las ubicaciones perimetrales se encuentran en América del Norte, Europa, Asia, Australia, América del Sur, el Medio Oriente, África y China. Las ubicaciones perimetrales admiten servicios de AWS, como Amazon Route 53 y AmazonCloudFront.

Las cachés perimetrales regionales se utilizan con Amazon CloudFront. Se utilizan cuando tiene contenido al que no se accede con la frecuencia suficiente para que permanezca en una ubicación perimetral.

Las cachés perimetrales regionales absorben este contenido y proporcionan una alternativa a recuperar ese contenido del servidor de origen. Para obtener más información sobre la infraestructura de Amazon CloudFront, consulte la infraestructura de [Amazon CloudFront](#) [infraestructure](#).



Los puntos clave de esta sección del módulo incluyen:

- La infraestructura global de AWS consta de regiones y zonas de disponibilidad.
- La elección de una región suele basarse en los requisitos de cumplimiento o en reducir la latencia.
- Cada zona de disponibilidad se encuentra separada de forma física de otras zonas de disponibilidad y tiene alimentación, redes y conectividad redundantes.
- Las ubicaciones perimetrales y las cachés perimetrales regionales mejoran el rendimiento al almacenar en caché el contenido más cercano a los usuarios.

En resumen, en este módulo aprendió a hacer lo siguiente:

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



Ahora es el momento de completar la evaluación de conocimientos para esta unidad.