Enfoques de copia de seguridad y recuperación mediante AWS

Junio de 2016

Avisos

Este documento se suministra únicamente con fines informativos. Representa la oferta actual de productos y prácticas de AWS a partir de la fecha de publicación de este documento. Dichas prácticas y productos pueden modificarse sin previo aviso. Los clientes son responsables de realizar sus propias evaluaciones independientes de la información contenida en este documento y de cualquier uso de los productos o servicios de AWS, cada uno de los cuales se ofrece "como es", sin garantía de ningún tipo, ya sea explícita o implícita. Este documento no constituye ninguna garantía, representación, compromiso contractual, condición ni certeza por parte de AWS, sus filiales, proveedores o licenciantes. Las responsabilidades y obligaciones de AWS con sus clientes se rigen por los acuerdos de AWS y este documento no forma parte ni supone una modificación de ningún acuerdo entre AWS y sus clientes.

Contenido

Resumen

Este documento está dirigido a los arquitectos de soluciones empresariales, a los arquitectos de copias de seguridad y a los administradores de TI responsables de la protección de los datos en los entornos de TI de la compañía. Se abordan las cargas de trabajo de producción y las arquitecturas que se pueden implementar mediante AWS para aumentar o reemplazar una solución de copia de seguridad y recuperación. Mediante estos enfoques se pueden reducir los costos y obtener más escalabilidad y más durabilidad para cumplir con el objetivo de tiempo de recuperación (RTO, Recovery Time Objective), el objetivo de punto de recuperación (RPO, Recovery Point Objective) y los requisitos de cumplimiento.

Introducción

A medida que el crecimiento de los datos empresariales se acelera, la tarea de protegerlos se vuelve cada vez más desafiante. Son habituales las preguntas sobre la durabilidad y la escalabilidad de los métodos de copia de seguridad, incluida la siguiente: ¿Cómo me ayuda la nube en mis necesidades de copia de seguridad y archivado?

En este documento se abordan numerosas arquitecturas de copia de seguridad (aplicaciones nativas en la nube, entornos híbridos y locales) y de servicios de AWS asociados que se pueden utilizar para crear soluciones de protección de datos escalables y fiables.

¿Por qué debería usar AWS como una plataforma de protección de datos?

Amazon Web Services (AWS) es una plataforma de informática en la nube segura, de alto rendimiento, flexible, rentable y fácil de utilizar. AWS se encarga de las tareas pesadas no diferenciadas y proporciona herramientas y recursos que puede utilizar para crear soluciones de copia de seguridad y recuperación escalables.

Usar AWS como parte de su estrategia de protección de datos tiene muchas ventajas:

- **Durabilidad:** Amazon Simple Storage Service (Amazon S3) y Amazon Glacier están diseñados para ofrecer un 99,999999999% (11 nueves) de durabilidad para los objetos almacenados allí. Ambas plataformas son ubicaciones fiables para realizar copias de seguridad de datos.
- **Seguridad:** AWS proporciona una serie de opciones para el control de acceso y el cifrado de datos en tránsito y en reposo.
- **Infraestructura global:** los servicios de AWS están disponibles en todo el mundo, por lo que puede almacenar y hacer una copia de seguridad de los datos en la región donde se cumplan sus requisitos de cumplimiento.
- **Cumplimiento:** la infraestructura de AWS cuenta con certificaciones de cumplimiento de estándares como los Controles de las organizaciones de servicios (SOC, Service Organization Controls), la Declaración de normas para trabajos de atestación (SSAE, Statement on Standards for Attestation Engagements) 16, la Organización Internacional de Normalización (ISO, International Organization for Standardization) 27001, el estándar de seguridad de datos (DSS, Data Security Standard) del sector de las tarjetas de pago (PCI, Payment Card Industry), la Ley de Portabilidad y Responsabilidad de Seguros Médicos (HIPPA, Health Insurance Portability and Accountability Act), SEC[1] y el Programa federal de administración de riesgos y autorizaciones (FedRAMP, Federal Risk and Authorization Management Program) para que la solución de copia de seguridad se adecue fácilmente a sus requisitos de cumplimiento existentes.
- **Escalabilidad:** con AWS, no tiene que preocuparse por la capacidad. Puede aumentar o reducir su consumo según sus necesidades sin provocar gastos generales de administración.

[1] https://aws.amazon.com/about-aws/whats-new/2015/09/amazon-glacier-receives-third-party-compliance-assessment-for-sec-rule-17a-4f-from-cohasset-associates-inc/

- **TCO más bajo:** la escala de las operaciones de AWS recorta los costos de servicio y ayuda a reducir el costo total de propiedad (TCO, Total Cost of Ownership) del almacenamiento. AWS pasa los ahorros de costos a los clientes en forma de rebajas de precios.
- **Precio de pago por uso:** compre los servicios de AWS que necesite y solo durante el período que tiene pensado usarlos. Los precios de AWS no implican costos iniciales, multas por terminación ni contratos a largo plazo.

Servicios de almacenamiento de AWS para la protección de datos

Amazon S3 y Amazon Glacier son los servicios ideales para las copias de seguridad y el archivado. Son plataformas de almacenamiento duraderas y de bajo costo. Ambas ofrecen una capacidad ilimitada y no requieren de la administración de volúmenes o medios a medida que crecen los conjuntos de datos de copias de seguridad. El modelo de pago por uso y el bajo costo de GB por mes hacen que estos servicios sean la solución adecuada para la protección de datos.

Amazon S3

Amazon S3 proporciona un almacenamiento de objetos altamente seguro y escalable.

Amazon S3 le permite almacenar y recuperar cualquier cantidad de datos en cualquier momento y desde cualquier parte de la web. Amazon S3 almacena los datos como objetos dentro de recursos llamados *buckets*. AWS Storage Gateway y muchas soluciones de copia de seguridad de terceros pueden administrar los objetos de Amazon S3 por usted. Puede almacenar la cantidad de objetos que desee en un bucket y puede escribir, leer y eliminar objetos en este. Los objetos simples pueden tener un tamaño de hasta 5 TB.

Amazon S3 ofrece una gama de clases de almacenamiento diseñada para diferentes casos de uso, como los siguientes:

- **Amazon S3 Standard** para almacenamiento general de datos a los que se accede frecuentemente.

- **Amazon S3 Standard: acceso poco frecuente** para datos de duración prolongada a los que se accede con menos frecuencia.
- **Amazon Glacier** para un archivado a largo plazo.

Amazon S3 también ofrece políticas de ciclo de vida que puede configurar para administrar sus datos a través de este ciclo. Después de configurar una política, sus datos se migrarán a la clase de almacenamiento adecuada sin generar ningún cambio en su aplicación. Para obtener más información, consulte Clases de almacenamiento de S3.

Amazon Glacier

Amazon Glacier es un servicio de almacenamiento de costo extremadamente bajo y archivado en la nube que ofrece almacenamiento seguro y duradero para el archivado de datos y la copia de seguridad en línea. Para mantener bajos los costos, Amazon Glacier está optimizado para los datos a los que se accede con poca frecuencia y cuando son aceptables unos tiempos de recuperación de varias horas. Con Amazon Glacier, puede almacenar de manera fiable grandes o pequeñas cantidades de datos por tan solo 0,007 USD por gigabyte al mes, lo cual genera ahorros considerables en comparación con las soluciones locales. Amazon Glacier es adecuado para almacenar datos de copia de seguridad cuyos requisitos de conservación sean prolongados o indefinidos y para el archivado de datos a largo plazo. Para obtener más información, consulte Amazon Glacier.

AWS Storage Gateway

Mediante AWS Storage Gateway se conecta un dispositivo de software presente en sus instalaciones con el almacenamiento basado en la nube para ofrecer una integración completa y altamente segura entre su entorno de TI local y la infraestructura de almacenamiento de AWS. Para obtener más información, consulte AWS Storage Gateway.

Servicios de transferencia de AWS

Además de los gateways y los conectores de terceros, puede usar las opciones de AWS como AWS Direct Connect, AWS Snowball, AWS Storage Gateway y Amazon S3 Transfer Acceleration para transferir sus datos rápidamente. Para obtener más información, consulte Migración de datos a la nube.

Diseño de una solución de copia de seguridad y recuperación

Cuando desarrolla una estrategia integral para restaurar y realizar una copia de seguridad de los datos, primero debe identificar las situaciones de falla o desastre que pueden producirse y sus posibles repercusiones comerciales. En algunas industrias, debe tener en cuenta los requisitos reglamentarios en cuanto a la seguridad de los datos, la privacidad y la conservación de registros.

Debe implementar procesos de copia de seguridad con los que se ofrezca el nivel adecuado de detalle para cumplir con el RTO y el RPO de la compañía, lo que incluye lo siguiente:

- Recuperación en el nivel de archivos
- Recuperación en el nivel de volumen
- Recuperación en el nivel de la aplicación (por ejemplo, bases de datos)
- Recuperación en el nivel de imágenes

En las siguientes secciones, se describen los enfoques de copia de seguridad, recuperación y archivado según la organización de su infraestructura. La infraestructura de TI puede clasificarse a grandes rasgos como nativa en la nube, local e híbrida.

Infraestructura nativa en la nube

En este escenario se describe un entorno de carga de trabajo que existe por completo en AWS. Como se muestra en la siguiente figura, incluye servidores web, servidores de aplicación, servidores de monitorización, bases de datos y Active Directory.

Si ejecuta todos sus servicios desde AWS, puede aprovechar muchas funciones integradas para satisfacer sus necesidades de protección y recuperación de datos.

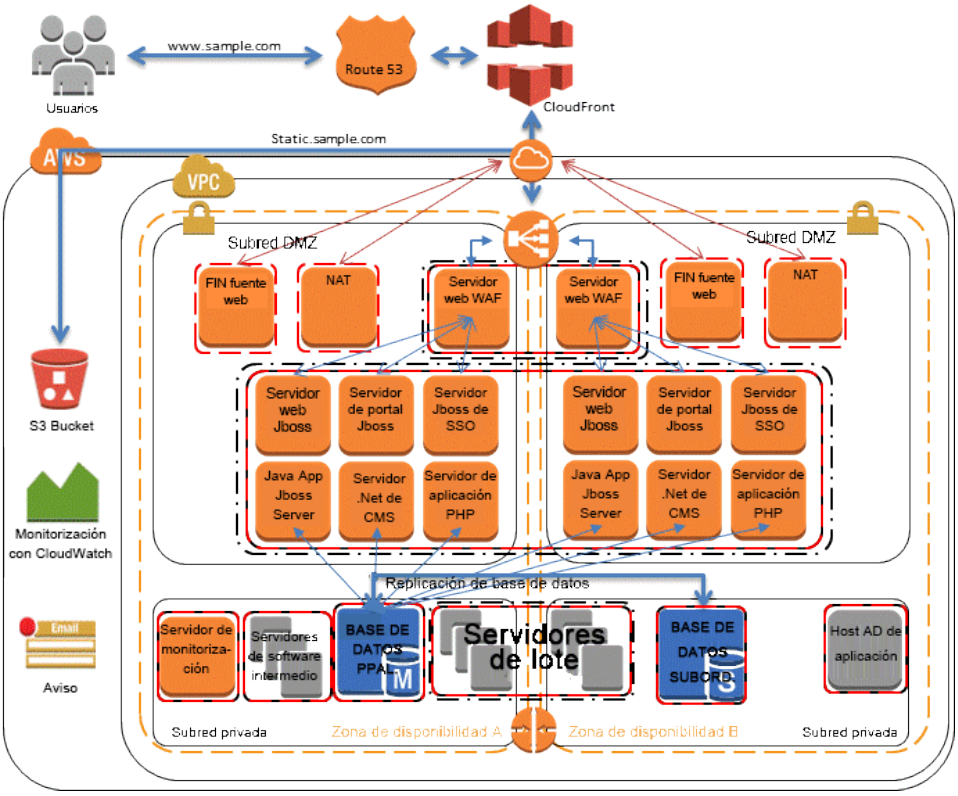

Figura 1: Escenario de infraestructura nativa de AWS

Protección basada en instantáneas de EBS

Cuando se ejecutan los servicios en Amazon Elastic Compute Cloud[2] (Amazon EC2), las instancias de computación pueden usar los volúmenes de Amazon Elastic Block Store (Amazon EBS) para almacenar datos primarios y acceder a estos. Puede usar este almacenamiento en bloque para los datos estructurados, como las bases de datos, o los datos no estructurados, como los archivos en un sistema de archivos en el volumen.

Amazon EBS también ofrece la posibilidad de crear instantáneas (copias de seguridad) de cualquier volumen de Amazon EBS. Realiza una copia del volumen y la coloca en Amazon S3, donde se almacena de manera redundante en varias zonas de disponibilidad. La primera instantánea es una copia completa del volumen, mientras que las instantáneas subsiguientes almacenan cambios graduales en el nivel del bloque únicamente.

[2] http://aws.amazon.com/ec2/

Esta es una manera rápida y confiable de restaurar todos los datos del volumen. Si solo necesita una restauración parcial, puede adjuntar el volumen a la instancia en ejecución con otro nombre de dispositivo, instalarlo y luego usar los comandos de copiado del sistema operativo para copiar los datos desde el volumen de copia de seguridad al volumen de producción.

Las instantáneas de Amazon EBS también pueden copiarse entre las regiones de AWS mediante el uso de la capacidad de copiado de instantáneas de EBS que se encuentra disponible en la consola o desde la línea de comandos, como se describe en la Guía del usuario de Amazon Elastic Cloud Compute[3]. Puede usar esta función para almacenar su copia de seguridad en otra región sin tener que administrar la tecnología de replicación subyacente.

Creación de instantáneas de EBS

Cuando crea una instantánea, protege sus datos directamente mediante un almacenamiento en disco duradero. Puede usar la Consola de administración de AWS, la interfaz de línea de comandos (CLI, Command Line Interface) o las interfaces de programación de aplicaciones (API, Application Programming Interface) para crear una instantánea de Amazon EBS.

En la consola de Amazon EC2, en la página **Volúmenes de Elastic Block Store**, seleccione **Crear instantánea** en el menú **Acciones**. En el cuadro de diálogo **Crear instantánea**, seleccione **Crear** para crear una instantánea que se almacenará en Amazon S3.

[3] http://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ebs-copy-snapshot.html

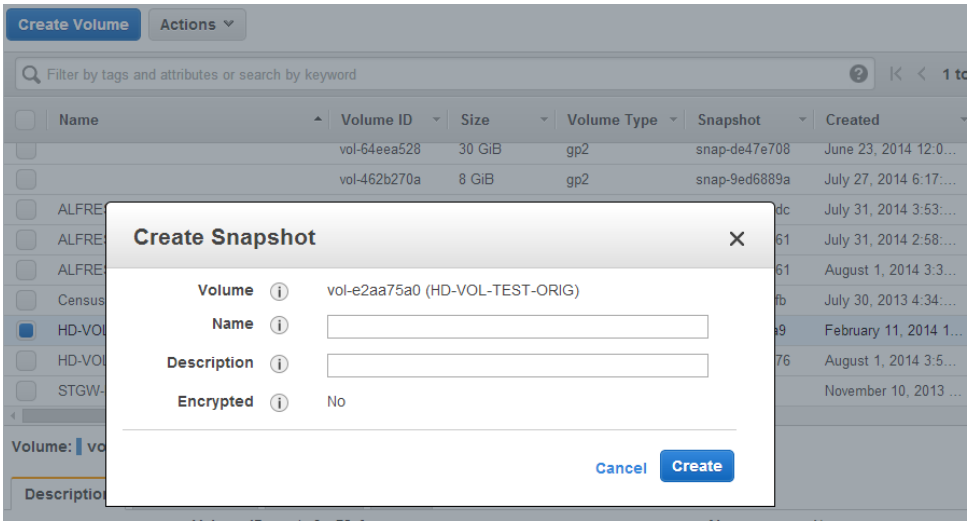

Figura 2: Uso de la consola EC2 para crear una instantánea

Para usar el comando de CLI para crear una instantánea, ejecute el siguiente comando:

```
➢  aws ec2 create-snapshot
```

Puede programar y ejecutar los comandos de `aws ec2 create-snapshot` regularmente para realizar una copia de seguridad de los datos de EBS. Los precios económicos de Amazon S3 le posibilitan conservar muchas generaciones de datos. Además, debido a que las instantáneas se basan en bloques, solo consume espacio para los datos que cambiaron luego de la creación de la primera instantánea.

Restablecimiento desde una instantánea de EBS

Para restaurar datos desde una instantánea, puede usar la Consola de administración de AWS, la CLI o las API para crear un volumen a partir de una instantánea existente.

Por ejemplo, siga estos pasos para restaurar un volumen a una copia de seguridad de un momento dado anteriormente:

1. Use el siguiente comando para crear un volumen a partir de la instantánea de copia de seguridad:

```
➢ aws ec2 create-volume –region us-west-1b –snapshot-
  id mysnapshot-id
```

2. En la instancia de Amazon EC2, desmonte el volumen existente.

   En Linux, use `Desmontar`. En Windows, use el Administrador lógico de volumen (LVM, Logical Volume Manager).

3. Use el siguiente comando para quitar el volumen existente de la instancia:

```
➢ aws ec2 detach-volume –volume-id oldvolume-id –
  instance-id myec2instance-id
```

4. Use el siguiente comando para incluir el volumen creado desde la instantánea:

```
➢ aws ec2 attach-volume –volume-id newvolume-id
  –instance-id myec2instance-id –device /dev/sdf
```

5. Vuelva a montar el volumen en la instancia en ejecución.

Creación de copias de seguridad coherentes o dinámicas

Cuando realiza una copia de seguridad, lo mejor es que el sistema se encuentre en un estado en el que no esté realizando ninguna E/S. Lo ideal es que la máquina no esté aceptando tráfico, pero esto es cada vez menos frecuente porque las operaciones de TI de todos los días a toda hora se convirtieron en la norma.

Por este motivo, debe pausar el sistema de archivos o la base de datos para realizar una copia de seguridad limpia. La forma de hacerlo depende de su base de datos o sistema de archivos.

El proceso para una base de datos es el siguiente:

- Si es posible, coloque la base de datos en el modo de copia de seguridad dinámica.
- Ejecute los comandos de la instantánea de Amazon EBS.
- Quite la base de datos del modo de copia de seguridad dinámica o, si usa una réplica de lectura, finalice la instancia de esta.

El proceso para un sistema de archivos es similar, pero depende de las capacidades del sistema operativo o del sistema de archivos. Por ejemplo, XFS es un sistema de archivos que puede vaciar sus datos para realizar una copia de seguridad coherente. Para obtener más información, consulte xfs_freeze[4].

Si su sistema de archivos no tiene la capacidad de congelarse, debe desinstalarlo, ejecutar el comando de la instantánea y volver a instalar el sistema de archivos. De otro modo, puede facilitar este proceso al usar un administrador lógico de volumen que tenga la capacidad de congelamiento de E/S.

Debido a que el proceso de instantáneas continúa en el fondo y la creación de la instantánea es rápida de ejecutar y captura un momento dado, solo debe desinstalar los volúmenes a los que está realizando una copia de seguridad durante algunos segundos. Debido a que la ventana de copia de seguridad es lo más pequeña posible, el tiempo de interrupción es predecible y se puede programar.

[4] https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/xfsfreeze.html

Realización de copias de seguridad de varios volúmenes

En algunos casos, puede distribuir los datos por varios volúmenes de Amazon EBS al usar un administrador lógico de volumen para aumentar el posible rendimiento. Cuando utiliza un administrador lógico de volumen (por ejemplo, mdadm o LVM), es importante realizar una copia de seguridad desde la capa del administrador de volumen en vez de a partir de los volúmenes de EBS subyacentes. Con esto se garantiza que los metadatos y los volúmenes del subcomponente sean coherentes.

Existen varias formas de lograrlo. Por ejemplo, puede usar el script que creó alestic.com[5]. Se deben vaciar al disco los búferes de memoria, se debe detener la E/S al disco del sistema de archivos y se debe iniciar una instantánea simultáneamente para todos los volúmenes que conforman el conjunto de matriz redundante de discos independientes (RAID, Redundant Array of Independent Disks). Luego de que se inicia la instantánea para los volúmenes (generalmente, en un segundo o dos), el sistema de archivos puede continuar con sus operaciones. Se deben etiquetar las instantáneas para que pueda administrarlas colectivamente durante una restauración.

También puede realizar estas copias de seguridad desde el administrador lógico de volumen o desde el nivel del sistema de archivos. En estos casos, el uso de un agente tradicional de copias de seguridad permite la copia de seguridad de los datos a través de la red. En Internet o en AWS Marketplace se encuentran disponibles numerosas soluciones de copia de seguridad basadas en agentes[6]. Recuerde que en los software de copias de seguridad basados en agentes se espera que el nombre del servidor y la dirección IP sean coherentes. Como resultado, usar estas herramientas con las instancias implementadas en una nube privada virtual (VPC, Virtual Private Cloud) de Amazon[7] es la mejor manera de garantizar confiabilidad.

---

[5] https://github.com/alestic/ec2-consistent-snapshot

[6] https://aws.amazon.com/marketplace/

[7] http://aws.amazon.com/vpc/

Un enfoque alternativo es crear una réplica de los volúmenes del sistema primario que existen en un solo volumen grande. Con esto se simplifica el proceso de copia de seguridad, ya que solo se debe hacer una copia de seguridad de un volumen grande y esta no se lleva a cabo en el sistema principal. Sin embargo, primero debe determinar si un solo volumen puede rendir lo suficiente durante la copia de seguridad y si el tamaño máximo del volumen es adecuado para la aplicación.

Enfoques para la copia de seguridad de bases de datos

AWS tiene muchas opciones para las bases de datos. Puede ejecutar su propia base de datos en una instancia de EC2 o usar una de las opciones de la base de datos de servicio administrada que proporciona Amazon Relational Database Service [8](Amazon RDS). Si ejecuta su propia base de datos en una instancia de EC2, puede hacer una copia de seguridad de los datos a los archivos mediante el uso de herramientas nativas (por ejemplo, MySQL[9], Oracle[10], MSSQL[11], PostgreSQL[12]) o crear una instancia de los volúmenes que contienen datos mediante el uso de uno de los métodos descritos en "Protección basada en instantáneas de EBS".

Uso de copias de seguridad de réplicas de bases de datos

Para las bases de datos integradas a los conjuntos de RAID de los volúmenes de Amazon EBS, puede quitar la carga de las copias de seguridad en la base de datos principal al crear una réplica de lectura de esta. Esta es una copia actualizada de la base de datos que se ejecuta en otra instancia de Amazon EC2. Se puede crear la réplica de la instancia de la base de datos con varios discos similares a la fuente, o bien, se pueden agrupar los datos en un solo volumen de EBS. Luego, puede usar uno de los procedimientos descritos en "Protección basada en instantáneas de EBS" para realizar una instantánea de los volúmenes de EBS. Este enfoque se usa a menudo para bases de datos grandes que deben funcionar todo el día, todos los días. Cuando ese es el caso, la ventana de copia de seguridad es demasiado extensa y la base de datos de producción no se puede desactivar durante períodos tan largos.

---

[8] https://aws.amazon.com/rds/

[9] http://dev.mysql.com/doc/refman/5.7/en/backup-and-recovery.html

[10] http://docs.oracle.com/cd/E11882_01/backup.112/e10642/rcmbckba.htm#BRADV8003

[11] http://msdn.microsoft.com/en-us/library/ms187510.aspx

[12] http://www.postgresql.org/docs/9.3/static/backup.html

Uso de Amazon RDS para copias de seguridad

En Amazon RDS se incluyen funciones para realizar copias de seguridad de bases de datos automáticamente. Amazon RDS crea una instantánea del volumen de almacenamiento de su instancia de base de datos, a través de lo cual se realiza una copia de seguridad de toda esta instancia, no solo de las bases de datos individuales.

Amazon RDS ofrece dos métodos diferentes para la realización de copias de seguridad y restauración de las instancias de base de datos:

- **Las copias de seguridad automatizadas** permiten la recuperación de su instancia de base de datos. Las copias de seguridad automatizadas se activan de forma predeterminada cuando crea una nueva instancia de base de datos. Amazon RDS realiza una copia de seguridad diaria y completa de sus datos durante una ventana que define cuando crea una instancia de base de datos. Puede configurar un período de conservación de hasta 35 días para la copia de seguridad automatizada. Amazon RDS usa estas copias de seguridad de datos periódicas junto con sus registros de transacciones para permitirle restaurar su instancia de base de datos a cualquier segundo durante su período de conservación, hasta `LatestRestorableTime` (generalmente, los últimos cinco minutos). Para encontrar el último momento de restauración de las instancias de sus bases de datos, puede usar la llamada a API `DescribeDBInstances` o consultar la pestaña **Descripción** de la base de datos en la Consola de administración de AWS.

  Cuando inicia una recuperación a un momento dado, se aplican los registros de transacciones a la copia de seguridad diaria más adecuada para restaurar su instancia de base de datos al momento que solicitó.

- **Las instantáneas de bases de datos** son copias de seguridad que ejecuta el usuario y que le permiten realizar una copia de seguridad de su base de datos a un estado conocido con la frecuencia que desee y, luego, regresar a ese estado en cualquier momento. Puede usar la Consola de administración de AWS o la llamada a API `CreateDBSnapshot` para crear instantáneas de bases de datos. Estas instantáneas tienen una capacidad de conservación ilimitada. Se conservan hasta que usa la consola o la llamada a API `DeleteDBSnapshot` para eliminarlas explícitamente.

Cuando restaure una base de datos a un momento dado o desde una instantánea de base de datos, se creará una instancia de base de datos nueva con un nuevo punto de conexión. De esta manera, puede crear varias instancias de base de datos desde una instantánea de base de datos específica o desde un momento dado.

Puede usar la Consola de administración de AWS o una llamada `DeleteDBInstance` para eliminar la instancia de base de datos anterior.

Uso de AMI para realizar copias de seguridad de las instancias de EC2

AWS almacena imágenes del sistema en las llamadas imágenes de máquina de Amazon (AMI, Amazon Machine Images). Estas imágenes constan de una plantilla para el volumen raíz que se necesita para lanzar una instancia. Puede usar la Consola de administración de AWS o el comando de CLI `aws ec2 create-image` para realizar una copia de seguridad del volumen de raíz como una AMI.

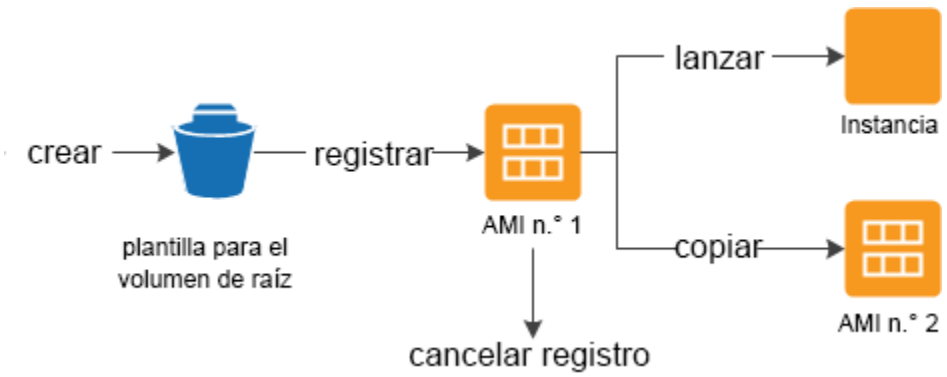

Figura 3: Uso de una AMI para lanzar y realizar una copia de seguridad de una instancia

Cuando registra una AMI, se almacena en su cuenta mediante el uso de instantáneas de Amazon EBS. Estas instantáneas se alojan en Amazon S3 y duran mucho tiempo.

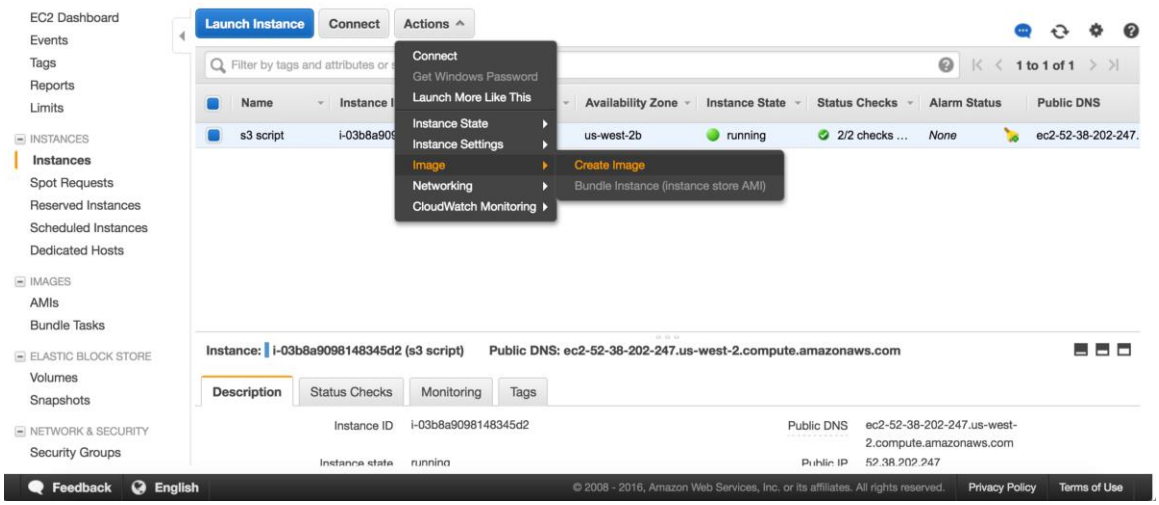

Figura 4: Uso de la consola EC2 para crear una imagen de máquina

Después de crear una AMI de su instancia de Amazon EC2, puede usarla para volver a crear la instancia o lanzar más copias de esta. También puede copiar las AMI de una región a otra en caso de una migración de la aplicación o una recuperación ante desastres.

Infraestructura local a AWS

En este escenario se describe un entorno de carga de trabajo sin componentes en la nube. Todos los recursos, incluidos los servidores web, los servidores de aplicación, los servidores de monitorización, las bases de datos, Active Directory, entre otros, se encuentran alojados en el centro de datos del cliente o a través de la coubicación.

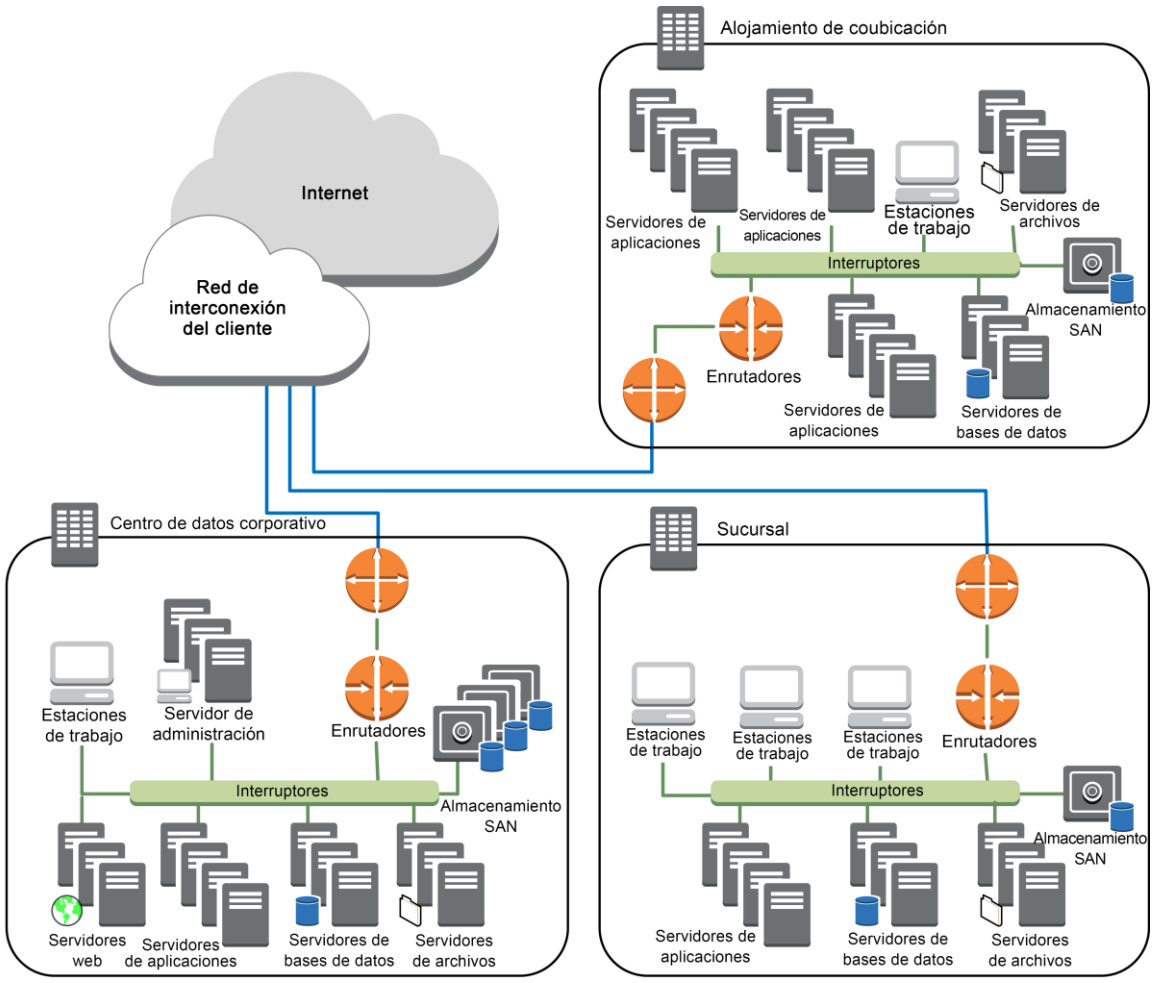

Figura 5: Entorno local

Mediante el uso de los servicios de almacenamiento de AWS, se puede concentrar en las tareas de copia de seguridad y archivado. No tiene que preocuparse por el escalado de almacenamiento ni la capacidad de la infraestructura para llevar a cabo la tarea de copia de seguridad.

Amazon S3 y Amazon Glacier se basan en la API de forma nativa y se encuentran disponibles en Internet. Esto permite a los proveedores de software de copia de seguridad integrar directamente sus aplicaciones con las soluciones de almacenamiento de AWS, como se muestra en la siguiente figura.

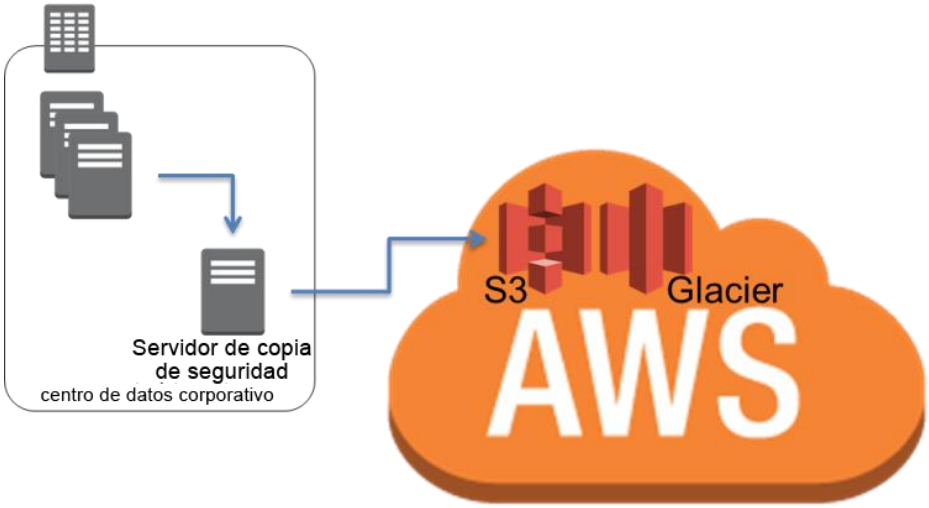

Figura 6: Conector de copia de seguridad a Amazon S3 o Amazon Glacier

En este escenario, el software de copia de seguridad y archivado interactúa directamente con AWS a través de las API. Debido a que el software de copia de seguridad depende de AWS, se realizará una copia de seguridad de los datos desde los servidores locales directamente a Amazon S3 o Amazon Glacier.

Si su software de copia de seguridad existente no es compatible de forma nativa con la nube de AWS, puede usar los productos AWS Storage Gateway. AWS Storage Gateway[13] es un dispositivo virtual que proporciona una integración completa y segura entre su centro de datos y la infraestructura de almacenamiento de AWS. Este servicio le permite almacenar datos de forma segura en la nube de AWS para obtener un almacenamiento escalable y rentable. Storage Gateway admite protocolos de almacenamiento estándar del sector que funcionan con sus aplicaciones existentes y, al mismo tiempo, almacena de forma segura todos sus datos cifrados en Amazon S3 o Amazon Glacier.

[13] http://aws.amazon.com/storagegateway/

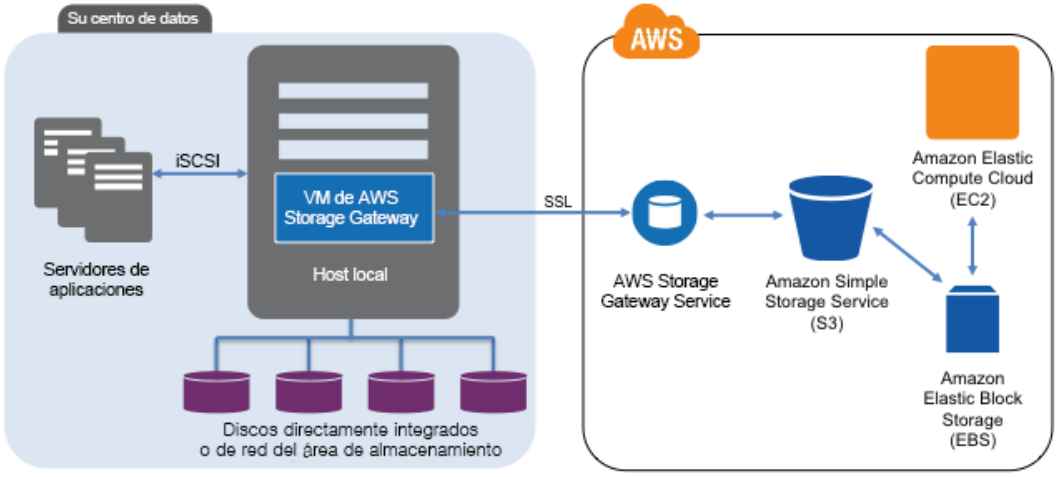

Figura 7: Conexión de almacenamiento local a AWS

AWS Storage Gateway es compatible con las siguientes configuraciones:

- **Puertas de enlace de volumen:** las puertas de enlace de volumen proporcionan volúmenes de almacenamiento cuyas copias de seguridad se encuentran en la nube y que puede instalar como dispositivos de interfaces estándares de equipos pequeños de Internet (iSCSI, Internet Small Computer System Interface) desde sus servidores de aplicaciones locales. La puerta de enlace es compatible con las siguientes configuraciones de volumen:
  - **Volúmenes en caché en la puerta de enlace:** puede almacenar sus datos principales en Amazon S3 y conservar localmente los datos a los que accede con frecuencia. Estos volúmenes proporcionan considerables ahorros de costos en almacenamiento principal, reducen al mínimo la necesidad de escalar el almacenamiento a nivel local y mantienen un acceso de baja latencia a sus datos de acceso frecuente.
  - **Volúmenes almacenados en la puerta de enlace:** en caso de necesitar acceso de baja latencia a todo su conjunto de datos, puede configurar su puerta de enlace de datos a nivel local para almacenar sus datos principales localmente y realizar, de forma asíncrona, copias de seguridad de las instantáneas de un momento dado de estos datos en Amazon S3. Los volúmenes almacenados en la puerta de enlace proporcionan copias de seguridad duraderas y económicas fuera del sitio que puede recuperar localmente o desde Amazon EC2.

- **Biblioteca de cintas virtuales (VTL) de puerta de enlace:** con la biblioteca de cintas virtuales (VTL, Virtual Tape Library) de puerta de enlace puede obtener una colección ilimitada de cintas virtuales. Cada cinta virtual se puede almacenar en una biblioteca de cintas virtuales respaldada por Amazon S3 o una estantería de cintas virtuales respaldada por Amazon Glacier. La biblioteca de cintas virtuales expone una interfaz iSCSI estándar del sector que proporciona a su aplicación de copia de seguridad acceso en línea a las cintas virtuales. Cuando ya no necesite acceder inmediatamente o con frecuencia a los datos incluidos en una cinta virtual, puede usar su aplicación de copia de seguridad para mover estos datos de la biblioteca de cintas virtuales a la estantería de cintas virtuales con el fin de reducir aun más los costos de almacenamiento.

Estas puertas de enlace actúan como dispositivos de instalación automática que proporcionan dispositivos iSCSI estándar, que se pueden integrar en su plataforma de copia de seguridad o archivado. Puede usar los dispositivos de disco iSCSI como grupos de almacenamiento para su software de copia de seguridad o VTL de puerta de enlace para descargar la copia de seguridad o archivado basado en cinta directamente en Amazon S3 o Amazon Glacier.

Mediante este método, su copia de seguridad y archivos quedan automáticamente fuera del sitio (por cuestiones de cumplimiento) y se almacenan en dispositivos duraderos, lo cual elimina la complejidad y los riesgos de seguridad de la administración de cintas fuera del sitio.

Entornos híbridos

Las dos implementaciones de infraestructura analizadas hasta el momento, nativa de la nube o local, se pueden combinar en un escenario híbrido donde el entorno de las cargas de trabajo tiene componentes de infraestructura locales y en AWS. Los recursos, incluidos los servidores web, los servidores de aplicación, los servidores de monitorización, las bases de datos, Active Directory, entre otros, se encuentran alojados en el centro de datos del cliente o en AWS. Las aplicaciones que se ejecutan en la nube de AWS se conectan a las aplicaciones que se ejecutan localmente.

Esto se está convirtiendo en un escenario común para las cargas de trabajo de la compañía. Muchas compañías tienen sus propios centros de datos y usan AWS para aumentar su capacidad. Estos centros de datos del cliente a menudo están conectados a la red de AWS a través de enlaces de red de alta capacidad. Por ejemplo, con AWS Direct Connect[14], puede establecer una conexión privada dedicada entre sus instalaciones y AWS. Esto proporciona un ancho de banda y una latencia coherente para cargar datos a la nube a fin de proteger los datos y obtener un rendimiento y una latencia coherentes para las cargas de trabajo híbridas.

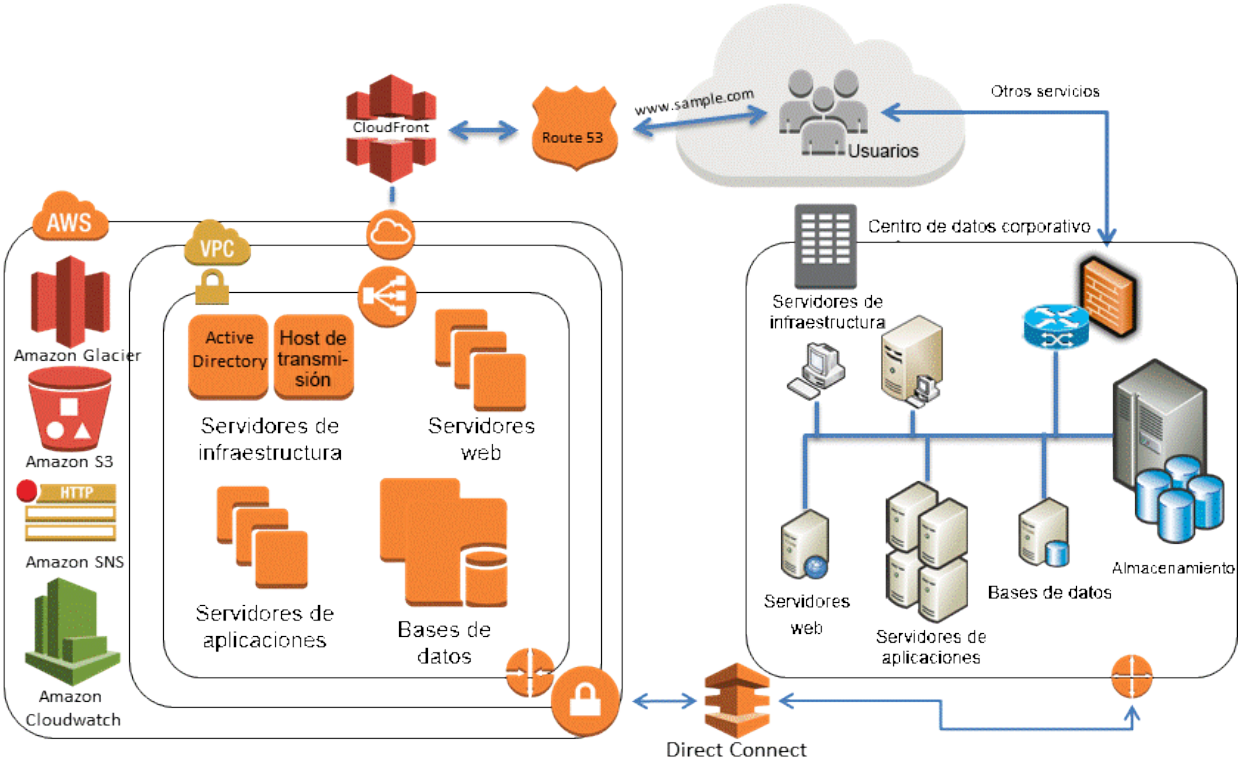

Figura 8: Escenario de infraestructura híbrida

Por lo general, las soluciones de protección de datos bien diseñadas usan una combinación de los métodos descritos en las soluciones nativas en la nube y locales.

[14] http://aws.amazon.com/directconnect/

Copia de seguridad de las aplicaciones basadas en AWS a su centro de datos

Si ya tiene una plataforma con la cual hace copias de seguridad de los datos para sus servidores locales, es fácil extenderla a los recursos de AWS mediante una conexión de red privada virtual (VPN, Virtual Private Network) o a través de AWS Direct Connect. Puede instalar el agente de copia de seguridad en las instancias de Amazon EC2 y respaldarlas conforme a sus políticas de protección de datos.

Migración de la administración de copias de seguridad a la nube por cuestiones de disponibilidad

De acuerdo con su arquitectura de copia de seguridad, puede tener un servidor de copia de seguridad principal y uno o más servidores de medios o almacenamiento ubicados localmente con los servicios que protege. En este caso, probablemente deba mover el servidor de copia de seguridad principal a una instancia de Amazon EC2 para protegerlo de los desastres locales y tener una infraestructura de copia de seguridad altamente disponible.

Para administrar los flujos de datos de copia de seguridad, probablemente deba crear uno o más servidores de medios en las instancias de Amazon EC2. Los servidores de medios cercanos a las instancias de Amazon EC2 le permitirán ahorrar dinero en transferencias por Internet y aumentar el rendimiento general de copia de seguridad y recuperación cuando realiza una copia de seguridad a S3 o Amazon Glacier.

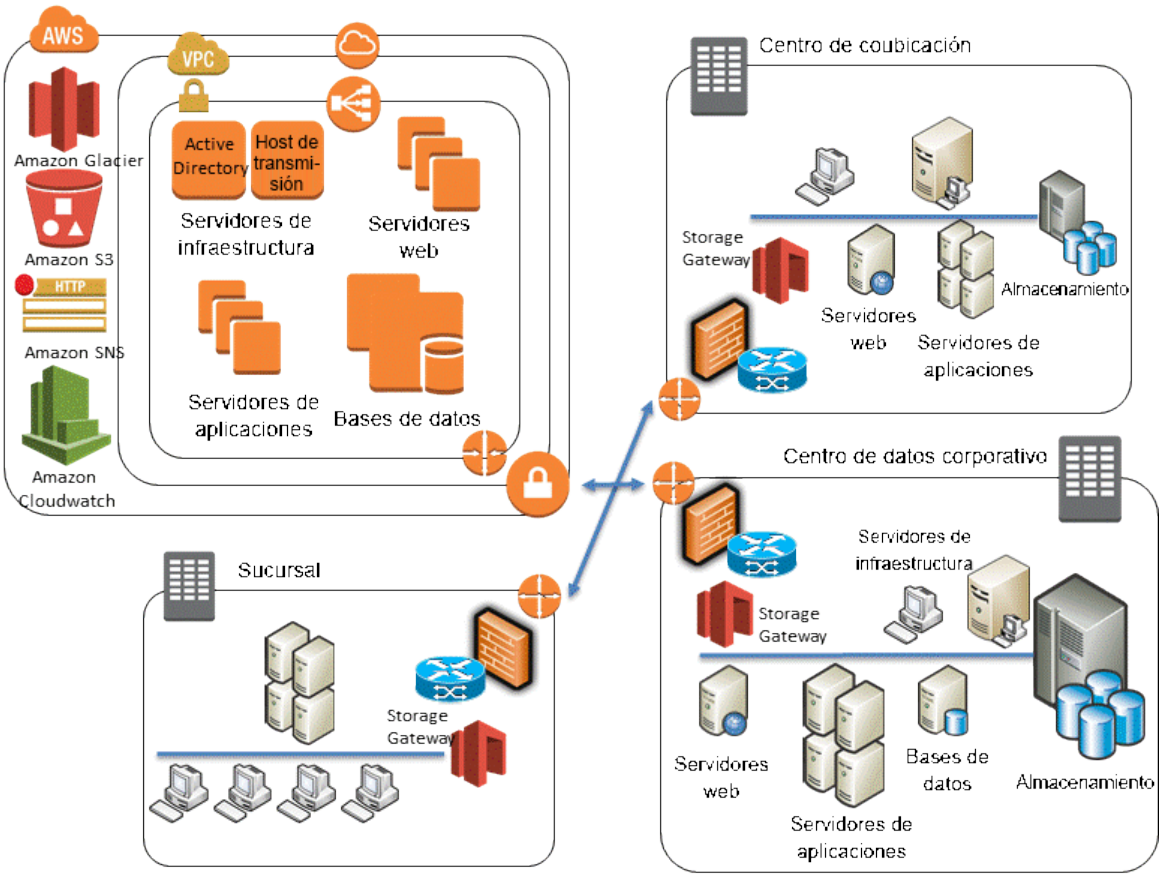

Figura 9: Uso de puertas de enlace en el escenario híbrido

Ejemplo de escenario híbrido

Suponga que está administrando un entorno donde está realizando una copia de seguridad de las instancias de Amazon EC2, los servidores autónomos, las máquinas virtuales y las bases de datos. Este entorno tiene 1 000 servidores y usted realiza una copia de seguridad del sistema operativo, los datos del archivo, las imágenes de la máquina virtual y las bases de datos. Hay 20 bases de datos (una combinación entre MySQL, Microsoft SQL Server y Oracle) a las que se les debe realizar una copia de seguridad.

Su software de copia de seguridad tiene agentes que realizan copias de seguridad de los sistemas operativos, las imágenes de la máquina virtual, los volúmenes de datos, las bases de datos de SQL Server y las de Oracle (mediante el uso de un administrador de recuperación [RMAN, Recovery Manager]). En el caso de las aplicaciones como MySQL para las cuales su software de copia de seguridad no tiene un agente, puede usar la utilidad del cliente mysqldump para crear un archivo de volcado de base de datos al disco donde los agentes de copia de seguridad estándares pueden proteger los datos.

Para proteger este entorno, su software de copia de seguridad de terceros probablemente tenga un servidor de catálogo global o servidor principal que controle las actividades de copia de seguridad, archivado y restauración, así como también servidores de varios medios que están conectados al almacenamiento basado en discos, las unidades de cintas abiertas lineales (LTO, Linear Tape-Open) y los servicios de almacenamiento de AWS.

La forma más simple de aumentar sus soluciones de copia de seguridad con los servicios de almacenamiento de AWS es aprovechar el respaldo de su proveedor de copia de seguridad de Amazon S3 o Amazon Glacier. Le sugerimos que se comunique con su proveedor para comprender sus opciones de integración y conectores. Para obtener una lista de proveedores de software de copia de seguridad que trabajen con AWS, consulte nuestro directorio de socios[15].

Si su software de copia de seguridad existente no admite de forma nativa el almacenamiento en la nube para la copia de seguridad o el archivado, puede usar un dispositivo de puerta de enlace de almacenamiento como un puente entre el software de copia de seguridad y Amazon S3 o Amazon Glacier.

Existen muchas soluciones de puerta de enlace de terceros. También puede utilizar los dispositivos virtuales de AWS Storage Gateway para acortar esta brecha porque utiliza técnicas genéricas como volúmenes basados en iSCSI y bibliotecas de cintas virtuales (VTL). Para esta configuración se necesita un hipervisor compatible (VMware o Microsoft Hyper-V) y almacenamiento local para alojar el dispositivo.

---

[15] http://www.aws-partner-directory.com/PartnerDirectory/PartnerSearch?type=ISV

Archivar datos con AWS

Cuando necesita conservar datos por cuestiones de cumplimiento o de la compañía, los archiva. A diferencia de las copias de seguridad, que generalmente se realizan para conservar una copia de los datos de producción durante un breve período por cuestiones de recuperación en caso de daños en los datos o pérdida de datos, mediante el archivado se conservan todas las copias de datos hasta que vence la política de conservación.

Una buena forma de archivar datos tiene las siguientes características:

- Duración de los datos para una integridad a largo plazo
- Seguridad de los datos
- Capacidad de fácil recuperación
- Bajo costo

El almacenamiento de datos inmutables puede ser otro requisito reglamentario o de cumplimiento.

Amazon Glacier proporciona la capacidad de archivar a bajo costo, el cifrado nativo de datos en reposo, un altísimo porcentaje de durabilidad y capacidad ilimitada.

Amazon S3 Standard: acceso poco frecuente es una buena elección para los casos en los que se requiere una recuperación rápida de los datos. Amazon Glacier es una buena elección para aquellos casos en los que se accede a los datos con poca frecuencia y cuando son aceptables unos tiempos de recuperación de varias horas.

En Amazon Glacier, los objetos se pueden dividir en niveles a través de reglas de ciclo de vida en S3 o la API de Amazon Glacier. La función de bloqueo de almacenes de Amazon Glacier le permite implementar y cumplir fácilmente los controles de cumplimiento en los almacenes individuales de Amazon Glacier con una política de bloqueo de almacenes. Puede especificar controles como el de grabación única, lectura múltiple (WORM, Write Once Read Many) en una política de bloqueo de almacenes y bloquear la política para que no se puedan hacer ediciones en un futuro. Para obtener más información, consulte Amazon Glacier.

Protección de los datos de copia de seguridad en AWS

La seguridad de los datos es una inquietud habitual. AWS toma la seguridad muy en serio. Esta es la base de cada servicio que lanzamos. Los servicios de almacenamiento como Amazon S3 proporcionan funcionalidades sólidas para el control de acceso y el cifrado de los datos en reposo y en tránsito. Todos los puntos de conexión de API de Amazon S3 y Amazon Glacier son compatibles con el cifrado SSL de datos en tránsito. Amazon Glacier cifra todos los datos en reposo de forma predeterminada. Con Amazon S3, los clientes pueden elegir un cifrado de servidor de los objetos en reposo al permitir a AWS administrar las claves de cifrado, proporcionar sus propias claves cuando se cargue un objeto o usar la integración de AWS Key Management Service (AWS KMS) [16]para las claves de cifrado. Además, los clientes siempre pueden cifrar sus datos antes de cargarlos a AWS. Para obtener más información, consulte Amazon Web Services: Información general sobre los procesos de seguridad.

Conclusión

Gartner reconoció a AWS como líder en los servicios de almacenamiento en la nube pública[17]. AWS tiene la capacidad de ayudar a las organizaciones a mover sus cargas de trabajo a plataformas basadas en la nube; la próxima generación de copias de seguridad. AWS proporciona soluciones económicas y escalables que ayudan a las organizaciones a lograr un equilibrio de sus requisitos de copia de seguridad y archivado. Estos servicios se integran bien con las tecnologías que usa actualmente.

---

[16] http://docs.aws.amazon.com/AmazonS3/latest/dev/UsingKMSEncryption.html

[17] http://www.gartner.com/technology/reprints.do?id=1-1WWKTQ3&ct=140709&st=sb

Colaboradores

Las siguientes personas participaron en la elaboración de este documento:

- Pawan Agnihotri, arquitecto de soluciones, Amazon Web Services
- Lee Kear, arquitecta de soluciones, Amazon Web Services
- Peter Levett, arquitecto de soluciones, Amazon Web Services

Revisiones del documento

Actualizado en mayo de 2016