

LECCIÓN 4.2: INCORPORACIÓN DE ALMACENAMIENTO A UNA INSTANCIA DE AMAZON EC2



INFORMACIÓN GENERAL DEL ALMACENAMIENTO DE AMAZON EC2

Volumen raíz

Este volumen siempre contiene el sistema operativo invitado



Almacén de instancias



Amazon EBS
(Solo con respaldo de SSD)

Volúmenes de datos

Para los datos a los que se accede mediante una sola instancia



Almacén de instancias



Amazon EBS

Para los datos a los que se accede desde varias instancias



Amazon Elastic File System (Amazon EFS) [Linux]



Amazon FSx for Windows File Server

Una instancia EC2 siempre tendrá un **volumen raíz** y, opcionalmente, puede tener uno más **volúmenes de datos**.

Las instancias EC2 tienen cuatro opciones de almacenamiento principales:

Almacén de instancias

Amazon EBS

Amazon Elastic File System (Amazon EFS)

Amazon FSx for Windows File Server

Las cuatro opciones se pueden utilizar para almacenar un volumen de datos. Sin embargo, solo se puede utilizar un almacén de instancias o un volumen de EBS respaldado por SSD para almacenar un volumen raíz.

Además, un almacén de instancias o un volumen de EBS debe utilizarse en una sola instancia a la vez. En el caso de un volumen de almacén de instancias, solo puede utilizarlo la instancia a la que se agrega el volumen.

Si desea compartir un volumen de datos en varias instancias al mismo tiempo, puede utilizar Amazon EFS o Amazon FSx for Windows File Server. Amazon EFS proporciona un sistema de archivos de uso compartido para las instancias de Linux, pero Amazon FSx proporciona un sistema de archivos de uso compartido para las instancias de Microsoft Windows.

En esta sección, analizará cada una de estas opciones.

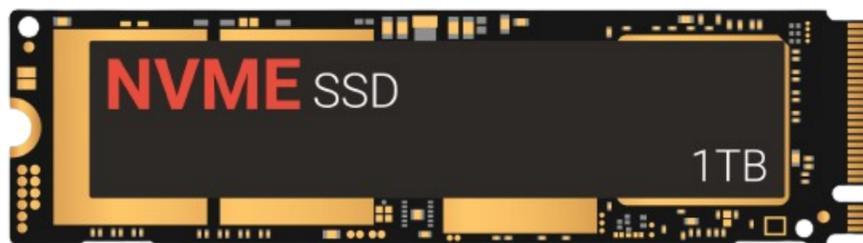
ALMACÉN DE INSTANCIAS

- Un almacén de instancias proporciona **almacenamiento no persistente** a una instancia:
 - Los datos se almacenan en el *mismo servidor físico* donde se ejecuta la instancia
- Características:
 - Almacenamiento temporal a nivel de bloque
 - Utiliza HDD o SSD
 - **Los datos del almacén de instancias se perderán cuando la instancia se *detenga o termine*.**
- Casos de uso de ejemplo:
 - Buffers
 - Caché
 - Datos de pruebas



Un volumen del almacén de instancias ofrece almacenamiento a nivel de bloque temporal para la instancia. Este almacenamiento se encuentra en discos que están conectados físicamente al equipo que aloja la instancia en ejecución. Un almacén de instancias es una buena opción para el almacenamiento temporal de información que cambia con frecuencia, como búferes, memorias caché, datos de pruebas y demás contenido temporal.

Un almacén de instancias consta de uno o varios volúmenes que se muestran como dispositivos de bloques. Los volúmenes residen en unidades de disco duro (HDD) o de estado sólido (SSD). Las SSD pueden ser SSD serie ATA o SSD de memoria rápida no volátil (NVMe). Los almacenes de instancias que utilizan NVMe tienen un rendimiento de E/S superior. Un almacén de instancias es exclusivo de una instancia concreta, pero el subsistema del disco se comparte entre las instancias de un equipo host.



El tipo de instancia determina los tamaños disponibles para un almacén de instancias y el tipo de hardware que se utiliza para los volúmenes de almacenes de instancias. La mayoría de los tipos de instancias admiten los almacenes de instancias, pero no todos.

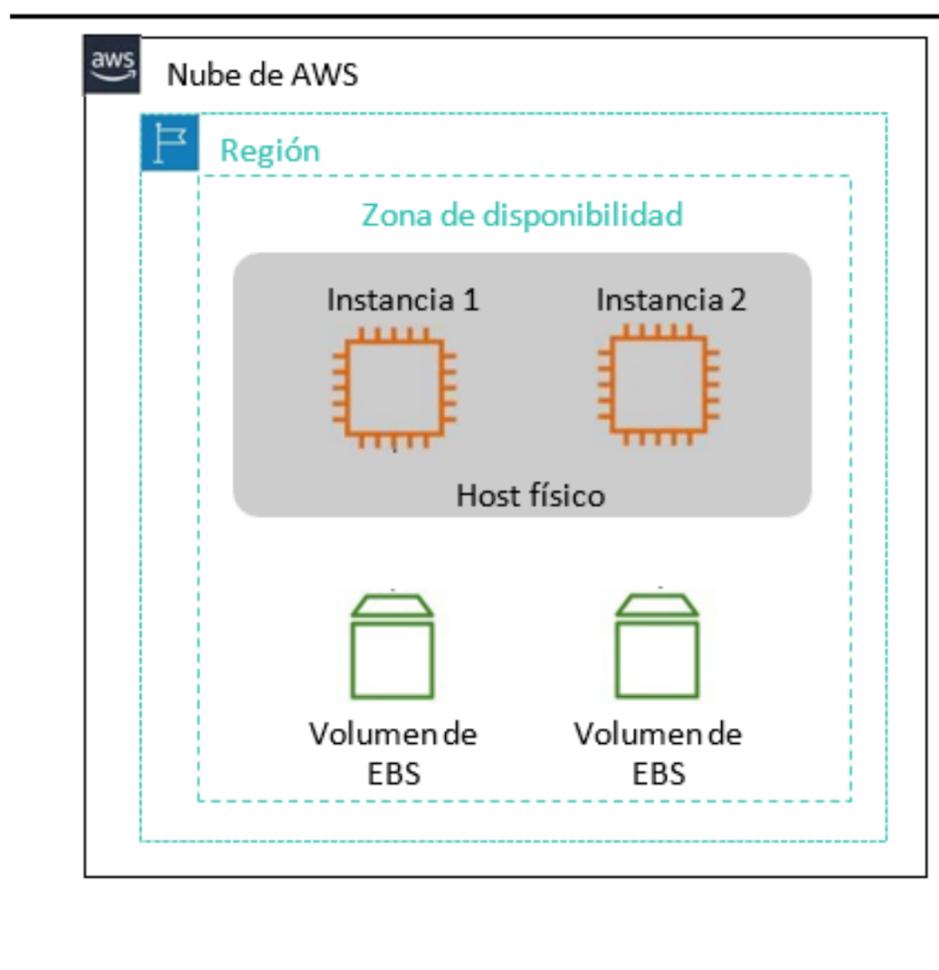
Los datos de un almacén de instancias se conservan cuando se reinicia la instancia, pero no cuando se termina. Las instancias con respaldo en el almacén de instancias no se pueden detener. Solo se pueden reiniciar o terminar. Si una instancia se reinicia, los datos del almacén de instancias persisten. Para obtener más información, consulte el tema Almacén de instancias de Amazon EC2 en la documentación en línea de Amazon EC2.

AMAZON EBS

Los volúmenes de Amazon EBS ofrecen almacenamiento persistente conectado a la red a una instancia EC2.

Características

- Con almacenamiento persistente a nivel de bloque
- Se puede asociar a cualquier instancia en la misma zona de disponibilidad
- Utiliza HDD o SSD
- Se puede cifrar
- Admite instantáneas que se almacenan de forma persistente en S3
- Los datos se almacenan de forma persistente independientemente de la vida de la instancia



Amazon EBS proporciona volúmenes de almacenamiento a nivel de bloque, similares a los de un disco duro externo, para utilizarlos con las instancias de Amazon EC2. Los volúmenes de Amazon EBS tienen una alta disponibilidad y fiabilidad, y se pueden asociar a cualquier instancia en ejecución en la misma zona de disponibilidad a través de la red.

Los volúmenes de Amazon EBS se conservan independientemente de la vida útil de la instancia y se pueden cifrar. Además, puede realizar una copia de seguridad de los datos de un volumen de EBS en Amazon S3 si toma una instantánea en un momento específico.

Los volúmenes de Amazon EBS residen en unidades de disco duro (HDD) o de estado sólido (SSD) (SSD serie ATA o SSD NVMe).

Debido a que están montados en la instancia, los volúmenes de EBS pueden proporcionar una latencia de acceso extremadamente baja. Por esta razón, los volúmenes de EBS se pueden utilizar para ejecutar una base de datos en una instancia EC2, por ejemplo.



	SSD de uso general (gp2)	SSD de IOPS provisionadas (io1)
Descripción	Equilibrio en el precio y el rendimiento para una gran variedad de cargas de trabajo	<ul style="list-style-type: none"> • Volumen SSD de mayor rendimiento • Bueno para cargas de trabajo de importancia crítica, de baja latencia o de alto rendimiento
Casos de uso	<ul style="list-style-type: none"> • Recomendado para la mayoría de las cargas de trabajo • Pueden ser un volumen de arranque 	<ul style="list-style-type: none"> • Aplicaciones de negocios fundamentales que requieren un rendimiento de IOPS sostenido • Cargas de trabajo de base de datos grande • Cargas de trabajo transaccionales • Puede ser un volumen de arranque

Amazon EBS proporciona tipos de volúmenes que difieren en las características de rendimiento y el precio, lo que le permite adaptar el rendimiento y el costo del almacenamiento a las necesidades de las aplicaciones. Se dividen en dos categorías: HDD y SSD.

Los volúmenes respaldados por SSD están optimizados para cargas de trabajo transaccionales que impliquen operaciones de lectura o escritura frecuentes de tamaño de E/S pequeño, en las que el atributo de rendimiento dominante es IOPS (Operaciones de entrada y salida por segundo).

Los dos tipos de volúmenes con respaldo de Amazon SSD son:

SSD de uso general (gp2): este es el tipo de volumen de EBS predeterminado para las instancias de Amazon EC2. Estos volúmenes son adecuados para una variedad de cargas de trabajo transaccionales, incluidos los entornos de desarrollo o prueba, las aplicaciones interactivas de baja latencia y los volúmenes de arranque.

SSD de IOPS provisionadas (io1): este tipo de volumen es la opción de almacenamiento de EBS de mayor rendimiento. Está diseñado para cargas de trabajo de aplicaciones y bases de datos esenciales de uso intensivo de E/S, así como cargas de trabajo de bases de datos y almacenamientos de datos de alto rendimiento.

Los volúmenes respaldados por SSD de Amazon EBS son ideales para cargas de trabajo de alto rendimiento y uso intensivo de IOPS que requieren una latencia extremadamente baja.

TIPOS DE VOLÚMENES CON RESPALDO DE:

	HDD de capacidad optimizada (st1)	HDD en frío (sc1)
Descripción	<ul style="list-style-type: none"> Tipo de volumen de bajo costo Diseñado para cargas de trabajo que requieren un rendimiento intensivo, a las que se accede con frecuencia 	<ul style="list-style-type: none"> Volumen HDD de menor costo Diseñado para cargas de trabajo a las que se accede con menos frecuencia
Casos de uso	<ul style="list-style-type: none"> Cargas de trabajo de <u>streaming</u> Big data Almacenes de datos Procesamiento de registros No puede ser un volumen de arranque 	<ul style="list-style-type: none"> Almacenamiento orientado al rendimiento para grandes volúmenes de datos a los que se accede con poca frecuencia Casos de uso donde el costo de almacenamiento más bajo es importante No puede ser un volumen de arranque

Los volúmenes respaldados por HDD están optimizados para cargas de trabajo de streaming grandes en las que la velocidad de procesamiento (medida en MiB/s) es una medición del rendimiento mejor que IOPS.

Hay dos tipos de volúmenes respaldados por HDD de Amazon:

HDD de rendimiento optimizado (st1): este tipo de volumen es ideal para cargas de trabajo de rendimiento intensivo a las que se accede con frecuencia y que poseen grandes conjuntos de datos y tamaños de E/S.

HDD en frío (sc1): este tipo de volumen proporciona el costo más bajo por GB de todos los tipos de volúmenes de EBS. Es una buena opción para las cargas de trabajo con grandes conjuntos de datos en frío a las que se accede con menos frecuencia.

Para obtener más información, consulte el tema Tipos de volumen de Amazon EBS en la guía del usuario de Amazon EC2.

INSTANCIAS OPTIMIZADAS PARA AMAZON EBS

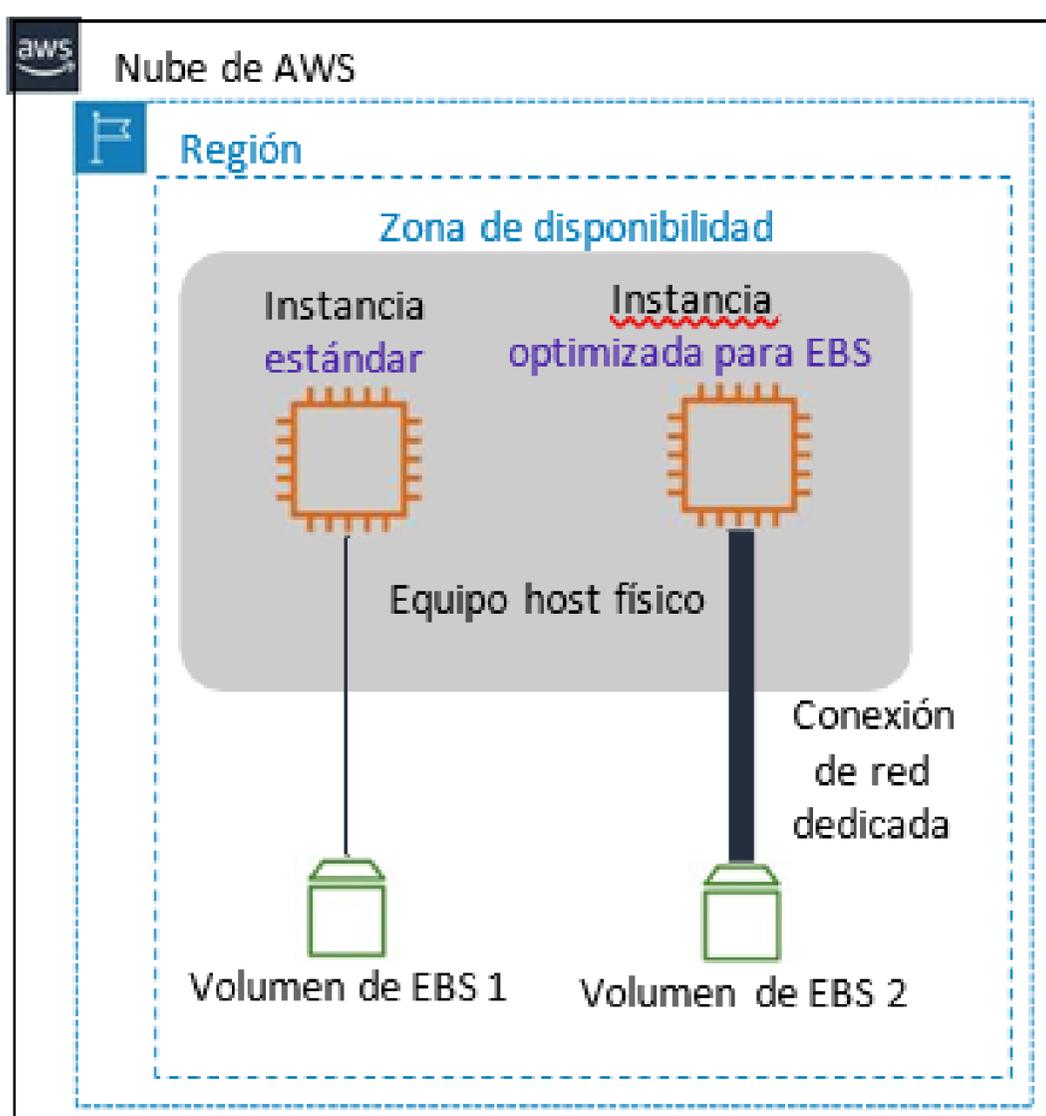
Determinados tipos de instancias EC2 se pueden optimizar para EBS

Beneficios:

- Ofrece una conexión de red dedicada para los volúmenes de EBS asociados
- Aumenta el rendimiento de E/S
- Si se utiliza un tipo de instancia basado en el sistema Nitro de Amazon EC2, se obtiene un mayor rendimiento

Uso:

- Para los tipos de instancias optimizadas para EBS, la optimización está habilitada de forma predeterminada.
- Para otros tipos de instancias que admiten la optimización, se la debe habilitar de forma manual.



Determinados tipos de instancias EC2 se pueden optimizar, de manera que haya un acceso de E/S mayor a un volumen de EBS. Estas instancias se denominan instancias optimizadas para Amazon EBS.

Una instancia optimizada para EBS tiene una conexión de red dedicada entre la instancia misma y un volumen de EBS. Esta optimización suministra el mejor rendimiento para acceder a sus volúmenes de EBS al reducir la contención entre la E/S de Amazon EBS y el tráfico de red restante de la instancia. Específicamente, proporciona un ancho de banda dedicado para Amazon EBS con opciones de entre 425 y 14 000 Mbps (en función del tipo de instancia que se utilice). Este ancho de banda dedicado proporciona consistencia en el rendimiento. Por ejemplo, para un volumen de SSD de uso general (gp2) que está asociado a una instancia optimizada para EBS, puede representar una mejora del 10 % con respecto a su rendimiento de referencia.





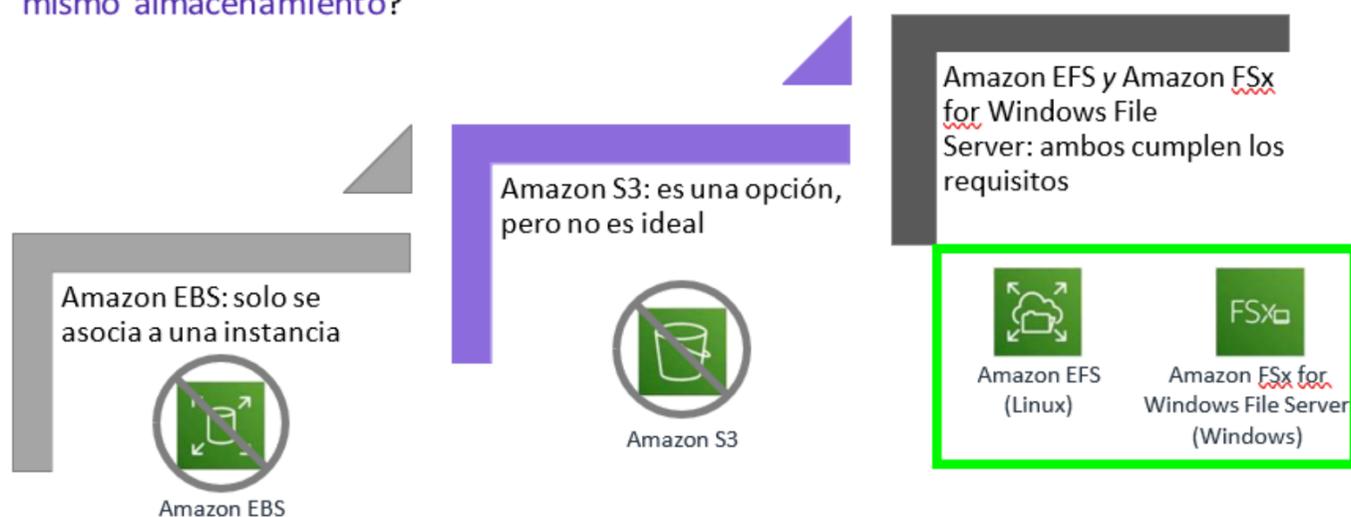
Si un tipo de instancia EC2 admite la optimización de EBS, estará habilitada de forma automática cuando la categoría del tipo de instancia sea optimizada para EBS. De lo contrario, debe habilitar la optimización de forma manual cuando lance la instancia mediante la configuración del atributo optimizada para Amazon EBS.

Además, si se crea un tipo de instancia optimizada para EBS en el sistema Nitro de AWS, podrá aprovechar otros aumentos de rendimiento. Este sistema es un conjunto de componentes de hardware y software creados por AWS que suministran un alto nivel de rendimiento, disponibilidad y seguridad, además de capacidades sin sistema operativo, para eliminar la sobrecarga de visualización.

Para obtener más información, consulte el tema Instancias optimizadas para Amazon EBS en la guía del usuario de Amazon EC2.

SISTEMAS DE ARCHIVOS COMPARTIDOS PARA INSTANCIAS EC2

¿Qué sucede si existen **varias instancias** que necesitan usar el **mismo almacenamiento**?



El almacén de instancias de Amazon EC2 y Amazon EBS ofrecen buenas opciones para almacenar volúmenes raíz y de datos para una única instancia. Sin embargo, ¿qué pasa si necesita compartir los datos en varias instancias de forma simultánea? ¿Cómo se maneja una aplicación que se ejecuta en varias instancias que deben utilizar el mismo sistema de archivos?

Anteriormente en esta sección, aprendió que un volumen de EBS debe asociarse solo a una instancia a la vez. Por lo tanto, no puede resolver el problema señalado. Amazon S3 es otra opción, pero es un sistema de almacenamiento de objetos, no un almacén de bloques. Por lo tanto, los cambios en Amazon S3 sobrescriben archivos enteros, no bloques de caracteres en los archivos. Para realizar cambios de alto rendimiento en archivos de varios tamaños, un sistema de archivos funciona mejor que un sistema de almacenamiento de objetos.

Este requisito necesita el rendimiento y la coherencia de lectura y escritura de un sistema de archivos de red. Amazon EFS y Amazon FSx for Windows File Server proporcionan un sistema de archivos de uso compartido para las instancias de Linux y de Windows. Podrá obtener información sobre estos dos servicios a continuación.

AMAZON EFS



Amazon EFS ofrece almacenamiento del sistema de archivos para cargas de trabajo **basadas en Linux**.

- Sistema de archivos elástico completamente administrado
- Incremento o reducción de escala automáticamente a medida que se agregan y eliminan los archivos
- Petabytes de capacidad
- Compatibilidad con protocolos del sistema de archivos de red (NFS)
 - Capacidad de montar el sistema de archivos en la instancia EC2
- Compatibilidad con todas las AMI basadas en Linux para Amazon EC2

Amazon EFS es un servicio completamente administrado que ofrece un almacenamiento de sistema de archivos para cargas de trabajo basadas en Linux. El acceso a los archivos se realiza a través de una interfaz de sistema de archivos (mediante API de E/S de archivos estándar del sistema operativo). Estos admiten semántica de acceso completo al sistema de archivos, como la consistencia sólida y el bloqueo de archivos. Amazon EFS usa la versión 4.x del protocolo Network File System (NFS). Se puede utilizar con cualquier AMI que admita este protocolo.



Varias instancias EC2 pueden acceder a un sistema de archivos de EFS al mismo tiempo. Por lo tanto, Amazon EFS puede proporcionar un origen de datos común para las cargas de trabajo y las aplicaciones que se ejecutan en más de una instancia de Amazon EC2. Además, los sistemas de archivos de EFS pueden escalarse automáticamente desde los gigabytes hasta los petabytes de datos sin necesidad de aprovisionar almacenamiento.

CASOS DE USO DE AMAZON

Casos de uso de Amazon



Cargas de trabajo y aplicaciones comunes:

- Directorios principales
- Sistema de archivos para aplicaciones empresariales
- Pruebas y desarrollo de aplicaciones
- Copias de seguridad de base de datos
- Administración de contenido y servicios web
- Flujos de trabajo multimedia
- Análisis de big data

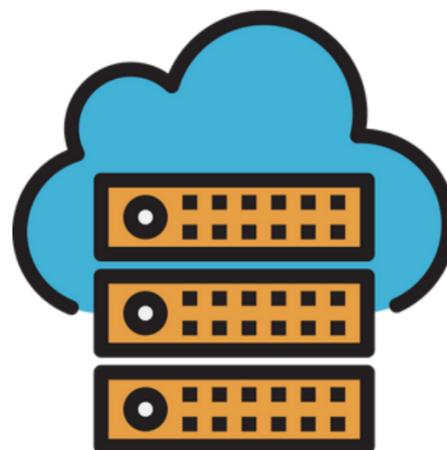
Ejemplo de comando para montar el sistema de archivos en cada SO invitado:

```
$ sudo mount -t nfs4 mount-target-DNS:/ ~/efs-mount-point
```

Entre los principales casos de uso de Amazon EFS se incluyen:

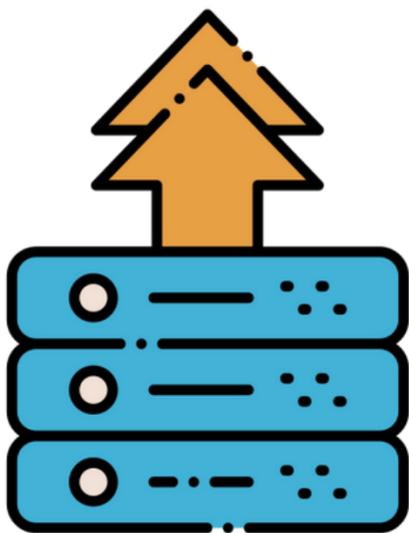
Directorios principales:

proporcionan almacenamiento a organizaciones que cuentan con muchos usuarios que deben obtener acceso a conjuntos de datos comunes y compartirlos.



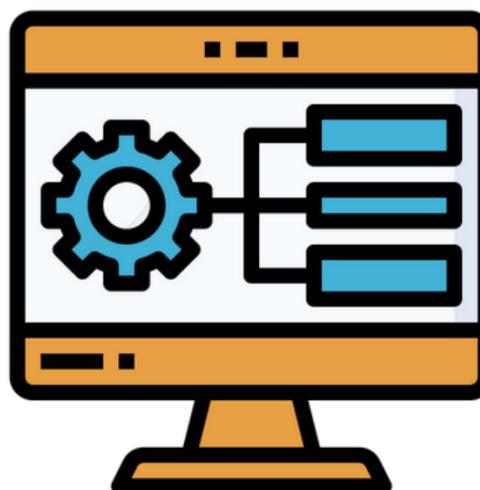
Sistema de archivos para aplicaciones empresariales:

proporcionan la escalabilidad, la elasticidad, la disponibilidad y la durabilidad necesarias para ser el almacenamiento de archivos para aplicaciones empresariales y para las aplicaciones ofrecidas como servicio.



Pruebas y desarrollo de aplicaciones:

proporcionan un repositorio de almacenamiento común que le permite compartir código y otros archivos de una forma segura y organizada.

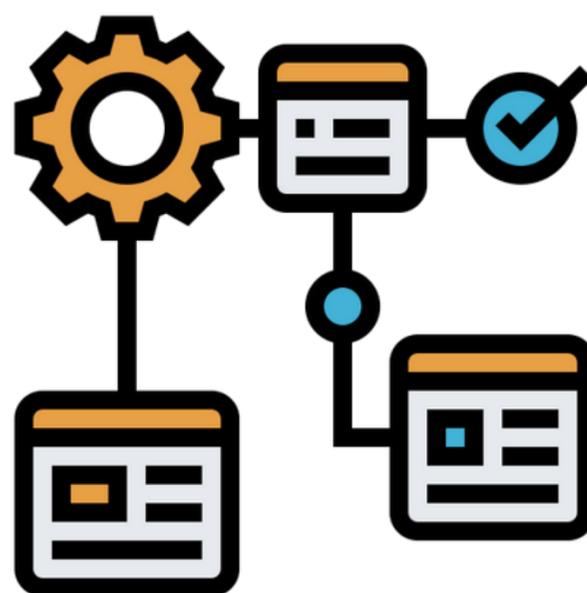


Entrega web y administración de contenido:

proporciona un sistema de archivos duradero y de alto rendimiento para sistemas de administración de contenido que almacenan y brindan información a diversas aplicaciones (como sitios web, publicaciones en línea y archivos).

Copias de seguridad de bases de datos:

pueden montarse con NFSv4 a partir de servidores de bases de datos.





Análisis de big data:

proporciona la escala y el rendimiento para aplicaciones de big data que requieren un rendimiento elevado para nodos de computación con coherencia de lectura tras escritura y operaciones de archivo de baja latencia.

Flujos de trabajo multimedia:

proporciona un modelo sólido de consistencia de datos, alto rendimiento y acceso a archivos compartidos que pueden reducir el tiempo que se tardan en realizar los trabajos de edición de video, producción en estudio, procesamiento de la transmisión, diseño de sonido y renderizado. Al mismo tiempo, consolida varios repositorios de archivos locales en una única ubicación para todos los usuarios.



Para obtener más información, consulte la guía del usuario de Amazon EFS.

AMAZON FSX FOR WINDOWS FILE SERVER

Proporciona almacenamiento de sistema de archivos compartido completamente administrado para las instancias de Microsoft Windows EC2.

- Compatibilidad nativa con Microsoft Windows
- New technology file system (NTFS)
- Protocolo Native Server message Block (SMB) versión 2.0 a 3.1.1
- Espacios de nombres distributed File system (DFS) y DFS Replication
- Se integra con Microsoft Active directory y admite listas de control de acceso (ACL) de Windows
- Respaldo por almacenamiento SSD de alto rendimiento

Si necesita almacenamiento basado en sistemas de archivos para las instancias de Microsoft Windows EC2, Amazon FSx for Windows File Server proporciona servidores de archivos de Windows completamente administrados, respaldados por un sistema de archivos de Windows completamente nativo. Se basa en Microsoft Windows Server y tiene compatibilidad nativa con las características y los protocolos del sistema de archivos de Windows, entre los que se incluyen:

New Technology File System (NTFS), el sistema de archivos predeterminado de los servidores de Windows

Server Message Block (SMB), el protocolo estándar abierto para acceder al almacenamiento de archivos a través de una red

Distributed File System (DFS), un servicio que permite la organización de múltiples recursos compartidos de archivos SMB distribuidos en un sistema de archivos distribuido

Microsoft Active Directory, el servicio de directorio de Microsoft que administra y organiza los recursos empresariales (incluidos los usuarios, los grupos y los recursos de red)

Amazon FSx utiliza almacenamiento SSD para proporcionar rendimiento rápido, altos niveles de rendimiento e IOPS, y bajas latencias. También reduce la sobrecarga administrativa de la configuración y el aprovisionamiento de servidores de archivos y volúmenes de almacenamiento.



CASOS DE USO DE AMAZON FSX FOR



Amazon FSX Windows File Server admite una gran variedad de cargas de trabajo de Microsoft Windows.

- Directorios principales
- Cargas de trabajo de aplicaciones de migración "lift-and-shift"
- Flujos de trabajo multimedia y de entretenimiento
- Análisis de datos
- Administración de contenido y servicios web
- Entornos de desarrollo de software

Para obtener más información, consulte la guía del usuario de Amazon FSx for Windows File Server.

Los casos de uso de Amazon FSx for Windows File Server (FSx) incluyen las siguientes cargas de trabajo de Microsoft Windows:

Directorios principales

Crean un sistema de archivos al que pueden acceder cientos de miles de usuarios y establecen permisos a nivel de archivo o carpeta.

Flujos de trabajo de multimedia y de entretenimiento

Los flujos de trabajo de multimedia, como la transcodificación, el procesamiento y el streaming de archivos multimedia, a menudo dependen del almacenamiento compartido de archivos de Windows con un rendimiento constante. FSx for Windows File Server le proporciona un sistema de archivos compartido con el alto rendimiento y las bajas latencias que necesitan estas cargas de trabajo de Windows.

Servidores web y administración de contenido

Por lo general, varios servidores web o de contenido necesitan acceso a los mismos archivos. FSx for Windows File Server proporciona un sistema de archivos al que se puede acceder a través de miles de instancias simultáneamente.

Aplicaciones de migración “lift-and-shift”

Proporcionan recursos compartidos de archivos de Windows nativos completamente administrados con características, como la integración de Microsoft Active Directory y las copias de seguridad automáticas.

Análisis de datos

Debido a que proporciona sistemas de archivos escalables con alto rendimiento y bajas latencias, FSx for Windows File Server le permite ejecutar cargas de trabajo de análisis con uso intensivo de datos.

Entornos de desarrollo de software

FSx for Windows File Server le permite migrar al desarrollo en la nube sin cambios en los flujos de trabajo de desarrollo ya que proporciona almacenamiento de archivos compartido al que sus desarrolladores y sus entornos pueden acceder.

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

- Entre las opciones de almacenamiento para las instancias EC2 se incluyen el almacén de instancias, Amazon EBS, Amazon EFS y Amazon FSx for Windows File Server
- Para un volumen raíz, utilice el almacén de instancias o Amazon EBS con respaldo de SSD
- Para un volumen de datos que solo trabaja con una instancia, utilice el almacén de instancias o el almacenamiento de Amazon EBS
- Para un volumen de datos que funciona con varias instancias de Linux, utilice Amazon EFS
- Para un volumen de datos que funciona con varias instancias de Microsoft Windows, utilice Amazon FSx for Windows File Server

