



Servicios de almacenamiento y bases de datos

En este módulo, analizaremos los servicios de almacenamiento y bases de datos de Google Cloud. Todas las aplicaciones necesitan almacenar información, ya sean datos comerciales, contenido multimedia para transmitir o datos de sensores de dispositivos.

Desde una perspectiva centrada en la aplicación, se usa tecnología para almacenar y recuperar los datos. Elegir entre una base de datos o un almacén de objetos es menos importante que la capacidad del servicio para satisfacer los requisitos de la aplicación de almacenar y recuperar los datos con eficiencia, dadas sus características.

Google ofrece numerosos servicios de almacenamiento de datos. En este módulo, analizaremos Cloud Storage, Filestore, Cloud SQL, Cloud Spanner, Cloud Firestore y Cloud Bigtable. Primero, veamos una descripción general de estos diferentes servicios.

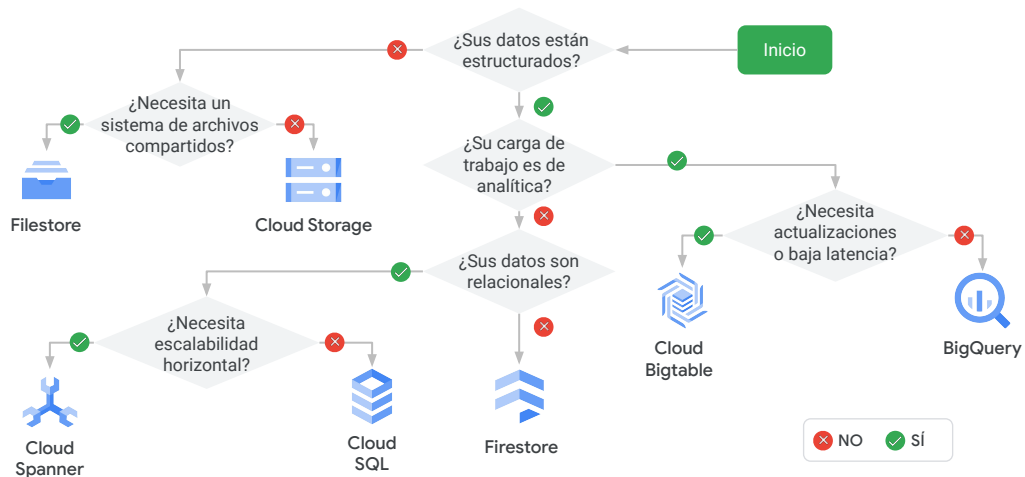
Servicios de almacenamiento y bases de datos



En esta tabla, se muestran los servicios de almacenamiento y bases de datos, y se destaca el tipo de servicio de almacenamiento (de objetos, de archivos, relacional, no relacional y almacén de datos), para qué sirve cada uno y su uso previsto.

BigQuery también aparece en el lado derecho. Mencionamos este servicio porque tiene características de almacenamiento y procesamiento de datos. Usted *puede* almacenar datos en BigQuery, pero está destinado al análisis de macrodatos y las consultas interactivas, por lo que hablaremos de este servicio más adelante en el curso.

Tabla de decisiones de almacenamiento y base de datos



Si las tablas no son de su preferencia, también agregamos este árbol de decisión para que pueda identificar la solución que se adapta mejor a su aplicación.

Analicemos esto juntos:

- En primer lugar, pregúntese lo siguiente: ¿Sus datos están estructurados? Si no es el caso, pregúntese entonces si necesita un sistema de archivos compartidos. De ser así, elija Firestore.
- Si no, elija Cloud Storage.
- Si sus datos están estructurados, ¿su carga de trabajo se enfoca en la analítica? Si es así, recomendamos que elija Cloud Bigtable o BigQuery según sus necesidades de latencia y actualizaciones.
- De lo contrario, compruebe si sus datos son relacionales. Si no son relacionales, elija Firestore.
- Si son relacionales, puede que Cloud SQL o Cloud Spanner sean buenas opciones, dependiendo de si necesita escalabilidad horizontal.

Según su aplicación, es posible que use uno o varios de estos servicios para realizar el trabajo. Si desea obtener más información para elegir entre estos diferentes servicios, consulte la sección de vínculos de este video

<https://cloud.google.com/storage-options/>,
<https://cloud.google.com/products/databases/>

Alcance

Itinerario sobre infraestructura

- Diferenciadores de los servicios
- Cuándo considerar el uso de cada servicio
- Cómo configurar un servicio y conectarse a él

Itinerario sobre ingeniería de datos

- Cómo usar un sistema de bases de datos
- Diseño, organización, estructura, esquema y uso para una aplicación
- Detalles sobre cómo un servicio almacena y recupera datos estructurados

Antes de estudiar cada uno de los servicios de almacenamiento de datos, definamos el alcance de este módulo.

El objetivo de este módulo es explicar qué servicios están disponibles y cuándo puede considerar usarlos desde la perspectiva de la infraestructura. Nuestro objetivo es que pueda configurar un servicio y conectarse a él sin tener conocimientos detallados sobre cómo usar un sistema de bases de datos.

Le recomendamos los cursos de ingeniería de datos de Google Cloud si busca más información sobre el diseño, las organizaciones, las estructuras, los esquemas y los detalles sobre cómo los datos se optimizan, entregan y almacenan adecuadamente en esos diferentes servicios.

Temario

Cloud Storage y Filestore

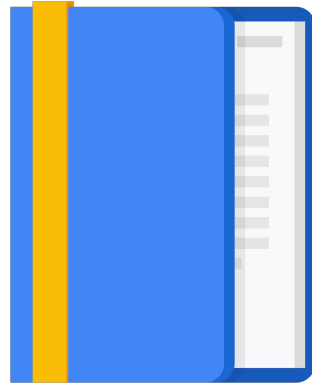
Cloud SQL

Cloud Spanner

Firestore

Cloud Bigtable

Memorystore



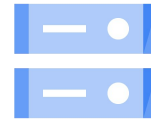
Analicemos el temario. Este módulo abarca todos los servicios que hemos mencionado hasta ahora. Para familiarizarse con estos servicios, los utilizará en dos labs.

También brindaremos una descripción general de Memorystore, el servicio de Redis completamente administrado que ofrece Google Cloud.

Comencemos con el análisis de Cloud Storage y Filestore.

Cloud Storage es un servicio de almacenamiento de objetos

- Contenido de sitios web
- Almacenamiento de datos para archivado y recuperación ante desastres
- Distribución de grandes objetos de datos a los usuarios mediante descarga directa

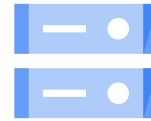


Cloud
Storage

Cloud Storage es el servicio de almacenamiento de objetos de Google Cloud. Permite almacenar y recuperar cualquier cantidad de datos en todo el mundo y en cualquier momento. Puede usar Cloud Storage en varias situaciones, como la entrega de contenido de un sitio web, el almacenamiento de datos con fines de archivado y recuperación ante desastres, o la distribución de grandes objetos de datos a los usuarios mediante descarga directa.

Características clave de Cloud Storage

- Se escala a exabytes
- El tiempo que tarda en enviar el primer byte es de milisegundos
- Disponibilidad sumamente alta en todas las clases de almacenamiento
- Se usa una sola API para todas las clases de almacenamiento



Cloud
Storage

Cloud Storage tiene algunas características clave:

- Se escala a exabytes de datos.
- El tiempo que tarda en enviar el primer byte es de milisegundos.
- Tiene una disponibilidad sumamente alta en todas las clases de almacenamiento.
- Se usa una sola API para todas las clases de almacenamiento.

Se cree que Cloud Storage consta de archivos almacenados en un sistema de archivos, pero, en realidad, no es así. Es una colección de buckets en los que se pueden guardar objetos. Puede crear directorios, por decirlo así, aunque, en realidad, un directorio solo es otro objeto que hace referencia a diferentes objetos del bucket. No podrá indexar fácilmente todos estos archivos como si fuera un sistema de archivos. Solo tiene una URL específica para acceder a los objetos.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.999999999% | | | |

Cloud Storage tiene cuatro clases de almacenamiento: Standard, Nearline, Coldline y Archive. Cada una de ellas ofrece 3 tipos de ubicación:

- Una multirregión es un área geográfica grande, como Estados Unidos, que contiene dos o más lugares geográficos.
- Una región doble es un par específico de regiones, como Finlandia y Países Bajos.
- Una región es un lugar geográfico específico, como Londres.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.99999999% | | | |

Los objetos almacenados en una región doble o una multirregión tienen redundancia geográfica. Ahora, repasemos cada una de las clases de almacenamiento:

Standard Storage es mejor para los datos a los que se accede con frecuencia (datos “activos”) o los que se almacenan por períodos breves. Es la clase más costosa, pero no tiene duración mínima del almacenamiento ni costo de recuperación.

Cuando se usa en una región, Standard Storage es adecuado para almacenar datos en la misma ubicación que los clústeres de Google Kubernetes Engine o las instancias de Compute Engine en las que se usan los datos. La ubicación conjunta de los recursos maximiza el rendimiento de los cálculos que consumen muchos datos y puede reducir los cargos de red.

Cuando se usa en una región doble, aún obtienes un rendimiento optimizado cuando accedes a los productos de Google Cloud que se encuentran en una de las regiones asociadas, pero también obtienes la disponibilidad mejorada que proviene del almacenamiento de datos en ubicaciones separadas a nivel geográfico.

Cuando se usa en una multirregión, Standard Storage es apropiado para almacenar datos a los que se accede en todo el mundo, como la entrega de contenido de sitios web y videos en streaming, la ejecución de cargas de trabajo interactivas o la entrega

de datos que respaldan aplicaciones para dispositivos móviles y de videojuegos.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.99999999% | | | |

Nearline Storage es un servicio de almacenamiento duradero, asequible y de alta calidad para almacenar datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos. Nearline Storage es mejor que Standard Storage en situaciones en las que una disponibilidad ligeramente menor, una duración mínima de almacenamiento de 30 días y los costos de acceso a datos son desventajas aceptables a cambio de menores costos de almacenamiento en reposo.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.99999999% | | | |

Coldline Storage es un servicio de almacenamiento de muy bajo costo y muy duradero para almacenar datos a los que se accede con poca frecuencia. Coldline Storage es mejor que Standard Storage o Nearline Storage en situaciones en las que una disponibilidad ligeramente menor, una duración mínima de almacenamiento de 90 días y mayores costos de acceso a datos son desventajas aceptables a cambio de menores costos de almacenamiento en reposo.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.99999999% | | | |

Archive Storage es el servicio de almacenamiento más duradero y de menor costo para el archivado de datos, la copia de seguridad en línea y la recuperación ante desastres. A diferencia de los servicios de almacenamiento “más fríos” que ofrecen otros proveedores de servicios en la nube, con este servicio los datos están disponibles en milisegundos, no en horas o días. En comparación con otras clases de almacenamiento de Cloud Storage, Archive Storage no tiene un ANS de disponibilidad, aunque la disponibilidad típica es similar a Nearline Storage y Coldline Storage. Archive Storage también tiene costos más altos para el acceso a los datos y las operaciones, y una duración mínima de almacenamiento de 365 días. Archive Storage es la mejor opción para los datos a los que planea acceder menos de una vez al año.

Descripción general de clases de almacenamiento

| | Standard | Nearline | Coldline | Archive |
|------------------------------------|--|---|---|--|
| Caso de uso | Datos “activos” o almacenados solo durante períodos breves, como cálculos de uso intensivo | Datos de acceso poco frecuente, como copias de seguridad, contenido multimedia de cola larga y archivado de datos | Datos de acceso poco frecuente que se leen o modifican una vez por trimestre, como máximo | Archivado de datos, copias de seguridad en línea y recuperación ante desastres |
| Duración mínima del almacenamiento | Ninguna | 30 días | 90 días | 365 días |
| Costo de recuperación | Ninguno | \$0.01 por GB | \$0.02 por GB | \$0.05 por GB |
| ANS de disponibilidad | 99.95% (multirregión/ región doble) 99.90% (región) | 99.90% (multirregión/región doble) 99.00% (región) | | Ninguno |
| Durabilidad | 99.99999999% | | | |

Enfoquémonos en la durabilidad y la disponibilidad. Todas estas clases de almacenamiento tienen 11 nueves de durabilidad, pero ¿qué significa eso? ¿Significa que puede acceder a sus archivos en cualquier momento? No. Significa que no perderá datos. Es posible que no pueda acceder a sus datos, lo que es parecido a ir al banco y ver que su dinero está ahí, tiene 11 nueves de durabilidad. Pero, cuando el banco está cerrado, no podemos acceder a él. Esa es la diferencia de disponibilidad entre las clases de almacenamiento y el tipo de ubicación.

Descripción general de Cloud Storage

Buckets

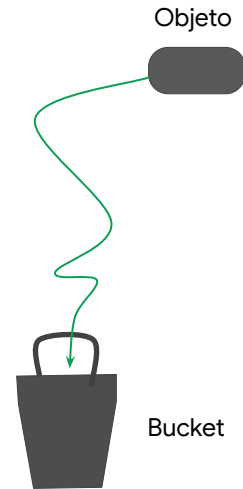
- Requisitos de nombres
- No se pueden anidar

Objetos

- Heredan la clase de almacenamiento de bucket cuando se crean
- Sin tamaño mínimo; almacenamiento ilimitado

Acceso

- Comando `gsutil`
- API (RESTful) de JSON o API de XML



Aquí, Cloud Storage se divide en varios elementos diferentes.

Descripción general de Cloud Storage

Buckets

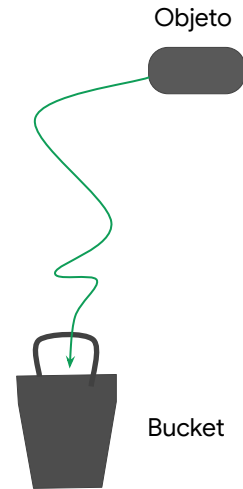
- Requisitos de nombres
- No se pueden anidar

Objetos

- Heredan la clase de almacenamiento de bucket cuando se crean
- Sin tamaño mínimo; almacenamiento ilimitado

Acceso

- Comando `gsutil`
- API (RESTful) de JSON o API de XML



En primer lugar, hay buckets que deben tener un nombre único a nivel global y no se pueden anidar.

Descripción general de Cloud Storage

Buckets

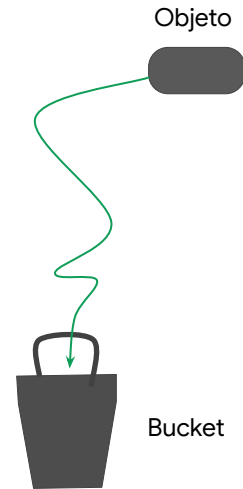
- Requisitos de nombres
- No se pueden anidar

Objetos

- Heredan la clase de almacenamiento de bucket cuando se crean
- Sin tamaño mínimo; almacenamiento ilimitado

Acceso

- Comando `gsutil`
- API (RESTful) de JSON o API de XML



Los datos que guarda en esos buckets son objetos que heredan la clase de almacenamiento del bucket, los que pueden ser archivos de texto, documentos, videos, etc. No hay un tamaño mínimo para esos objetos y puede escalarlo tanto como desee, siempre que la cuota lo permita.

Descripción general de Cloud Storage

Buckets

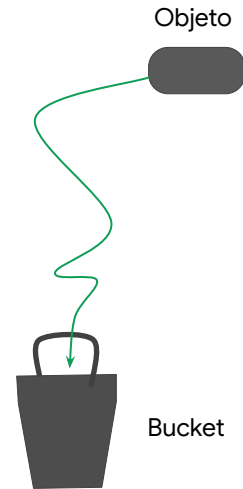
- Requisitos de nombres
- No se pueden anidar

Objetos

- Heredan la clase de almacenamiento de bucket cuando se crean
- Sin tamaño mínimo; almacenamiento ilimitado

Acceso

- Comando `gsutil`
- API (RESTful) de JSON o API de XML



Para acceder a los datos, puede usar el comando `gsutil` o las API de JSON o XML.

Cómo cambiar las clases de almacenamiento predeterminadas

- Se aplica la clase predeterminada a los objetos nuevos.
- Los buckets regionales nunca se pueden convertir en multirregionales o de región doble.
- Los buckets multirregionales nunca se pueden convertir en regionales.
- Los objetos no se pueden trasladar de un bucket a otro.
- Se pueden administrar las clases de los objetos mediante la Administración del ciclo de vida de los objetos.

Cuando sube un objeto a un bucket, el objeto se asigna a la clase de almacenamiento del bucket, a menos que especifique una. Puede cambiar la clase de almacenamiento predeterminada de un bucket, pero no puede cambiar el tipo de ubicación de regional a multirregional o de región doble, o viceversa.

También puede cambiar la clase de almacenamiento de un objeto que ya existe en el bucket sin mover el objeto a otro bucket ni cambiar la URL del objeto. Establecer una clase de almacenamiento por objeto es útil, por ejemplo, si tiene objetos que desea conservar en el bucket, pero a los que no espera acceder con frecuencia. En este caso, puede minimizar los costos cambiando la clase de almacenamiento de esos objetos específicos a Nearline Storage, Coldline Storage o Archive Storage.

Para facilitar la administración de las clases de objetos de los buckets, Cloud Storage ofrece la Administración del ciclo de vida de los objetos. Hablaremos sobre ese tema más adelante.

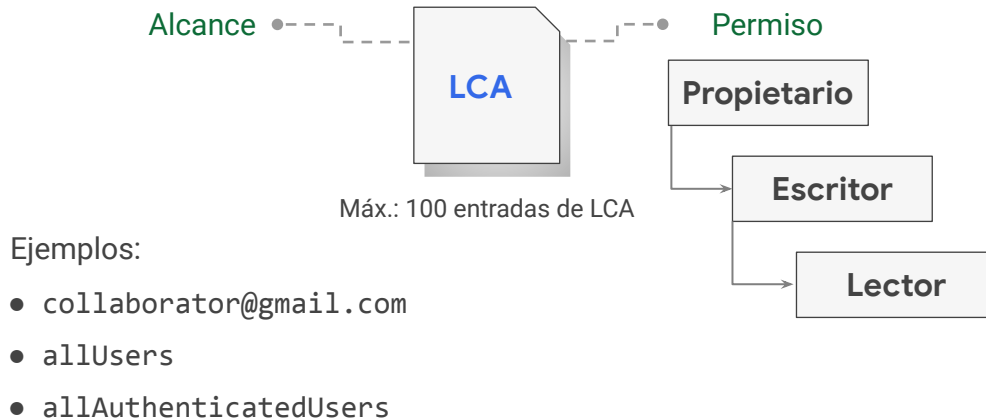
Control de acceso



Observemos el control de acceso a los objetos y buckets que forman parte de un proyecto.

- Podemos usar la IAM en el proyecto para controlar qué usuarios o cuentas de servicio pueden ver el bucket, mostrar una lista de los objetos almacenados dentro, ver los nombres de esos objetos o crear buckets nuevos. En la mayoría de los casos, la IAM es suficiente y los roles se heredan del proyecto al bucket y al objeto.
- Las listas de control de acceso (LCA) ofrecen control más detallado.
- Para un control más detallado aún, las URLs firmadas ofrecen una clave criptográfica que brinda acceso por tiempo limitado a un bucket o un objeto.
- Por último, un documento de políticas firmado permite definir mejor el control determinando el tipo de archivo que puede subir una persona con una URL firmada. Analicemos con mayor detalle las LCA y las URLs firmadas.

Listas de control de acceso (LCA)



Una LCA es un mecanismo que puede usar para definir quiénes tienen acceso a sus buckets y objetos, así como el nivel de acceso que tienen. La cantidad máxima de entradas de LCA que se pueden crear para un bucket o un objeto es 100.

Cada LCA consta de una o más entradas que constan de dos partes:

- Un alcance, que define quién puede realizar las acciones especificadas (por ejemplo, un usuario específico o un grupo de usuarios)
- Un permiso, que define qué acciones se pueden realizar (por ejemplo, leer o escribir)

El identificador allUsers que aparece en esta diapositiva representa a cualquier persona de Internet, con o sin una Cuenta de Google. En cambio, el identificador allAuthenticatedUsers representa a cualquier persona autenticada con una Cuenta de Google.

Para obtener más información sobre las LCA, consulte la sección de vínculos de este video [<https://cloud.google.com/storage/docs/access-control/lists>]

URLs firmadas

- Son como una clave de “estacionamiento” que otorga acceso a buckets y objetos por medio de tickets:
 - Cada ticket es una URL con firma criptográfica
 - Con límite de tiempo
 - Operaciones que se especifican en el ticket: HTTP GET, PUT y DELETE (no POST)
 - Cualquier usuario que tenga una URL puede invocar las operaciones permitidas
- Ejemplo en el que se usa la clave de cuenta privada y gsutil:

```
gsutil signurl -d 10m path/to/privatekey.p12  
gs://bucket/object
```

En el caso de algunas aplicaciones, es más fácil y eficiente otorgar tokens de acceso con límite de tiempo que cualquier usuario puede utilizar, en vez de recurrir a la autenticación basada en cuentas para controlar el acceso a los recursos. Un ejemplo de esto es cuando no desea exigir que los usuarios tengan Cuentas de Google.

Las URLs firmadas le permiten hacerlo en Cloud Storage. Usted crea una URL que otorga acceso de lectura o escritura a un recurso específico de Cloud Storage y especifica cuándo vence el acceso. Esa URL se firma con una clave privada asociada a una cuenta de servicio. Cuando se recibe una solicitud, Cloud Storage puede verificar que la URL que otorga acceso se emitió en nombre de una principal de seguridad confiable (en este caso, la cuenta de servicio) y delega su confianza en esa cuenta al titular de la URL.

Después de entregar la URL firmada, se pierde el control sobre ella, por lo que es recomendable que el acceso a la URL firmada venza después de un plazo razonable.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

Cloud Storage también incluye numerosas funciones. Por ahora, las analizaremos a grandes rasgos porque pronto revisaremos algunas de ellas con más detalle.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

Anteriormente en el curso, hablamos un poco sobre las Claves de encriptación proporcionadas por el cliente cuando se conectan discos persistentes a máquinas virtuales. Esta función permite enviar sus propias claves de encriptación en vez de las claves administradas por Google, lo cual también está disponible para Cloud Storage.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

Cloud Storage también ofrece Administración del ciclo de vida de los objetos, que permite borrar o archivar objetos automáticamente.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

Otra función es el control de versiones de objetos, que permite mantener varias versiones de objetos en los buckets. Debe tener en cuenta que se le cobra por las versiones como si fueran varios archivos.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

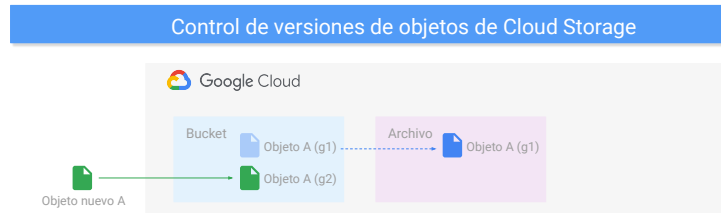
Además, Cloud Storage ofrece sincronización de directorios para que pueda sincronizar un directorio de VM con un bucket.

Características de Cloud Storage

- Claves de encriptación proporcionadas por el cliente (CSEK)
 - Permite usar su propia clave en vez de las claves administradas por Google.
- Administración del ciclo de vida de los objetos
 - Borra o archiva objetos automáticamente.
- Control de versiones de objetos
 - Mantiene varias versiones de objetos.
- Sincronización de directorios
 - Sincroniza un directorio de VM con un bucket.
- Notificaciones sobre cambios realizados en los objetos con Pub/Sub

Se pueden configurar notificaciones sobre cambios realizados en los objetos para Cloud Storage con Pub/Sub. Lo analizaremos más adelante.

El Control de versiones de objetos admite recuperar objetos que se borraron o reemplazaron



- Los objetos son inmutables.

- Control de versiones de objetos:
 - Permite mantener un historial de modificaciones de objetos.
 - Puede mostrar una lista de las versiones archivadas de un objeto, restablecer un objeto a un estado anterior o borrar una versión.

En Cloud Storage, los objetos son inmutables, por lo que un objeto subido no puede cambiar durante su ciclo de vida de almacenamiento. Para admitir la recuperación de objetos que se borran o reemplazan, Cloud Storage ofrece la función de Control de versiones de objetos.

El Control de versiones de objetos admite recuperar objetos que se borraron o reemplazaron



- Los objetos son inmutables.
- Control de versiones de objetos:
 - Permite mantener un historial de modificaciones de objetos.
 - Puede mostrar una lista de las versiones archivadas de un objeto, restablecer un objeto a un estado anterior o borrar una versión.

El Control de versiones de objetos se puede habilitar en un bucket, tras lo cual Cloud Storage crea una versión archivada de un objeto cada vez que se reemplaza o borra la versión publicada del objeto. La versión archivada conserva el nombre del objeto, pero se identifica de forma única con un número de generación, como se ilustra en esta diapositiva con g1.

Cuando está habilitado el Control de versiones de objetos, usted puede mostrar una lista de las versiones archivadas de un objeto, restablecer la versión publicada de un objeto a un estado anterior o borrar permanentemente una versión archivada, según sea necesario. Puede activar o desactivar el control de versiones de un bucket en cualquier momento. Si lo desactiva, las versiones de los objetos existentes quedarán en su lugar y el bucket dejará de acumular nuevas versiones de objetos archivadas.

Para obtener más información sobre el Control de versiones de objetos, consulte la sección de vínculos de este video:

<https://cloud.google.com/storage/docs/object-versioning>

Las políticas de Administración del ciclo de vida de los objetos especifican las acciones que se pueden realizar en objetos que cumplen con reglas determinadas

- Ejemplos:
 - Cambiar los objetos con más de un año de antigüedad a una clase de almacenamiento inferior
 - Borrar los objetos creados antes de una fecha específica
 - Conservar solo las 3 versiones más recientes de un objeto
- La inspección de objetos ocurre en lotes asíncronos.
- Los cambios pueden tardar 24 horas en aplicarse.

Cloud Storage ofrece la Administración del ciclo de vida de los objetos a fin de admitir casos de uso comunes, como configurar un tiempo de actividad para los objetos, archivar versiones anteriores de objetos o cambiar objetos a una clase de almacenamiento inferior para facilitar la administración de costos.

Puede asignar una configuración de administración del ciclo de vida a un bucket. La configuración es un conjunto de reglas que se aplican a todos los objetos en el bucket. Cuando un objeto cumple con los criterios de una de las reglas, Cloud Storage realiza de forma automática una acción específica en él.

Las políticas de Administración del ciclo de vida de los objetos especifican las acciones que se pueden realizar en objetos que cumplen con reglas determinadas

- Ejemplos:

- Cambiar los objetos con más de un año de antigüedad a una clase de almacenamiento inferior
- Borrar los objetos creados antes de una fecha específica
- Conservar solo las 3 versiones más recientes de un objeto
- La inspección de objetos ocurre en lotes asíncronos.
- Los cambios pueden tardar 24 horas en aplicarse.

Estos son algunos ejemplos de casos de uso:

- Primero, cambiar a Coldline Storage la clase de almacenamiento de los objetos con más de un año de antigüedad.
- Segundo, borrar los objetos creados antes de una fecha específica (por ejemplo, el 1 de enero de 2017).
- Y tercero, conservar solo las 3 versiones más recientes de cada objeto de un bucket que tenga habilitado el control de versiones.

Las políticas de Administración del ciclo de vida de los objetos especifican las acciones que se pueden realizar en objetos que cumplen con reglas determinadas

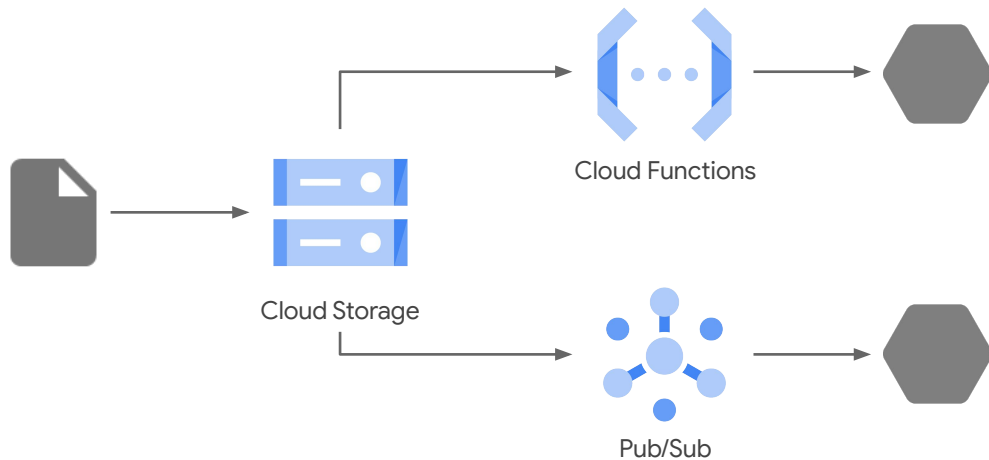
- Ejemplos:
 - Cambiar los objetos con más de un año de antigüedad a una clase de almacenamiento inferior
 - Borrar los objetos creados antes de una fecha específica
 - Conservar solo las 3 versiones más recientes de un objeto
- La inspección de objetos ocurre en lotes asíncronos.
- Los cambios pueden tardar 24 horas en aplicarse.

La inspección de objetos ocurre en lotes asíncronos, así que es posible que las reglas no se apliquen inmediatamente. Además, las actualizaciones de la configuración del ciclo de vida pueden tardar hasta 24 horas en surtir efecto. Esto significa que, cuando cambia la configuración del ciclo de vida, puede que la Administración del ciclo de vida de los objetos aún realice acciones basadas en la configuración anterior durante un período de hasta 24 horas. Téngalo en cuenta.

Para obtener más información sobre la Administración del ciclo de vida de los objetos, consulte la sección de vínculos de este video:

<https://cloud.google.com/storage/docs/lifecycle>

Notificaciones de Pub/Sub para Cloud Storage



Puede configurar las notificaciones de cambios realizados en los objetos para Cloud Storage con Pub/Sub. Las notificaciones de Pub/Sub envían información a este servicio sobre los cambios que se realizan en los objetos de los buckets y, luego, esta se agrega como mensajes al tema de Pub/Sub que usted elija. Por ejemplo, puede hacer un seguimiento de los objetos que se crean y borran en el bucket. Cada notificación contiene información en la que se describe el evento que la activó y el objeto que cambió. Puede enviar notificaciones a cualquier tema de Pub/Sub de cualquier proyecto en el que tenga permisos suficientes. Una vez que el tema de Pub/Sub recibe una notificación, el mensaje resultante se puede enviar a cualquier cantidad de suscriptores del tema.

Si desea obtener más información para conectar sus buckets de Cloud Storage con un tema de Pub/Sub, consulte la documentación de requisitos previos en la sección de vínculos de este video.

[\[https://cloud.google.com/storage/docs/reporting-changes#prereqs\]](https://cloud.google.com/storage/docs/reporting-changes#prereqs)

Recomendamos usar las notificaciones de Pub/Sub para realizar un seguimiento de los cambios que se realizan en los objetos de sus buckets de Cloud Storage, ya que son más rápidas, flexibles, rentables y fáciles de configurar. Pub/Sub es el servicio de mensajería en tiempo real distribuido que ofrece Google. Es uno de los temas del itinerario sobre desarrollo de aplicaciones.

Servicios de importación de datos

- **Transfer Appliance:** Almacene sus datos en un bastidor, captúrelos y, luego, envíelos a Google Cloud.
- **Servicio de transferencia de almacenamiento:** Importe datos en línea (desde otro bucket, un bucket de S3 o una fuente web).
- **Importación de medios sin conexión:** Un proveedor de terceros sube los datos desde soportes físicos de comunicación.



Cloud Console permite subir archivos individuales a los buckets, pero ¿qué pasaría si tuviera que subir terabytes o, incluso, petabytes de datos? Hay tres servicios diseñados para ello: Transfer Appliance, el Servicio de transferencia de almacenamiento y la Importación de medios sin conexión.

Servicios de importación de datos

- **Transfer Appliance:** Almacene sus datos en un bastidor, captúrelos y, luego, envíelos a Google Cloud.



- **Servicio de transferencia de almacenamiento:** Importe datos en línea (desde otro bucket, un bucket de S3 o una fuente web).
- **Importación de medios sin conexión:** Un proveedor de terceros sube los datos desde soportes físicos de comunicación.



Transfer Appliance es un dispositivo de hardware que puedes usar para migrar grandes volúmenes de datos (desde cientos de terabytes hasta 1 petabyte) a Google Cloud sin interrumpir las operaciones empresariales. Las imágenes de esta diapositiva son ejemplos de dispositivos para transferir datos.

Servicios de importación de datos

- **Transfer Appliance:** Almacene sus datos en un bastidor, captúrelos y, luego, envíelos a Google Cloud.



- **Servicio de transferencia de almacenamiento:** Importe datos en línea (desde otro bucket, un bucket de S3 o una fuente web).

- **Importación de medios sin conexión:**
Un proveedor de terceros sube los datos desde soportes físicos de comunicación.



El Servicio de transferencia de almacenamiento permite realizar importaciones de alto rendimiento de datos en línea. La fuente de datos puede ser otro bucket de Cloud Storage, un bucket de Amazon S3 o una ubicación HTTP/HTTPS.

Servicios de importación de datos

- **Transfer Appliance:** Almacene sus datos en un bastidor, captúrelos y, luego, envíelos a Google Cloud.
- **Servicio de transferencia de almacenamiento:** Importe datos en línea (desde otro bucket, un bucket de S3 o una fuente web).
- **Importación de medios sin conexión:** Un proveedor de terceros sube los datos desde soportes físicos de comunicación.



Por último, la Importación de medios sin conexión es un servicio de terceros en el que los soportes físicos de comunicación (arrays de almacenamiento, discos duros, cintas y unidades de memoria flash USB) se envían a un proveedor que sube los datos.

Para obtener más información sobre estos tres servicios, consulte la sección de vínculos de este video:

<https://cloud.google.com/transfer-appliance/>

<https://cloud.google.com/storage-transfer/docs/>

<https://cloud.google.com/storage/docs/offline-media-import-export>

Cloud Storage proporciona coherencia global sólida

- Read-after-write
- Lectura y, luego, actualización de metadatos
- Lectura y, luego, borrado
- Lista de depósitos
- Lista de objetos



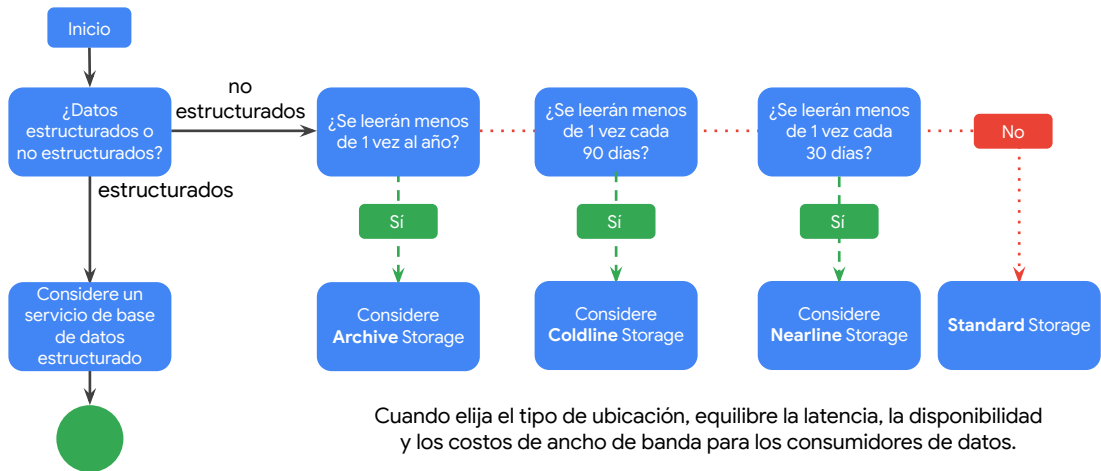
Cuando sube un objeto a Cloud Storage y recibe una respuesta exitosa, este está disponible de inmediato para descargarlo y realizar operaciones de metadatos en él desde cualquier ubicación en la que Google preste servicios. Esto se aplica cuando crea un objeto nuevo o reemplaza uno existente. Debido a que las cargas tienen coherencia sólida, nunca recibirá una respuesta 404 Not Found ni datos inactivos para una operación de read-after-write o read-after-metadata-update.

La coherencia sólida global se extiende también a las operaciones de eliminación en objetos. Si la solicitud de eliminación es exitosa, el intento inmediato de descarga de un objeto o sus metadatos arrojará un código de estado 404 Not Found, debido a que el objeto deja de existir después de esta operación.

La lista de buckets tiene coherencia sólida. Por ejemplo, si crea un bucket y, luego, realiza una operación para mostrar una lista de buckets, el bucket nuevo aparecerá en ella.

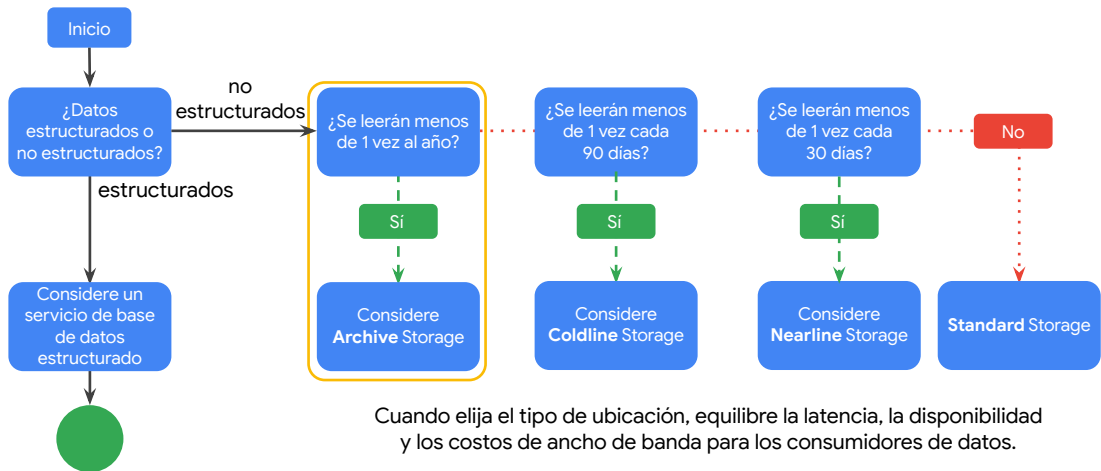
La lista de objetos también tiene coherencia sólida. Por ejemplo, si sube un objeto a un bucket y, luego, realiza una operación para mostrar una lista de objetos, el objeto nuevo aparecerá en ella.

Cómo elegir una clase de almacenamiento



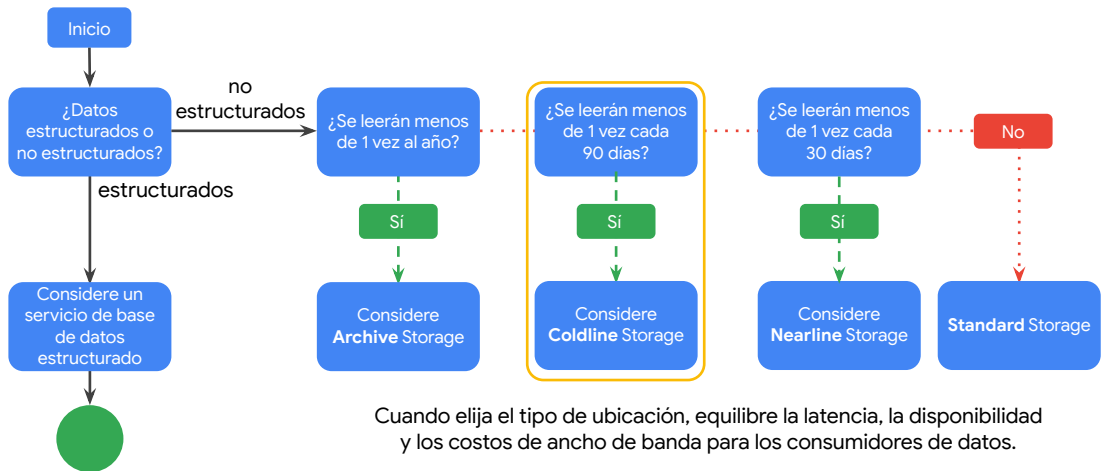
Exploremos el árbol de decisión para encontrar la clase de almacenamiento adecuada en Cloud Storage.

Cómo elegir una clase de almacenamiento



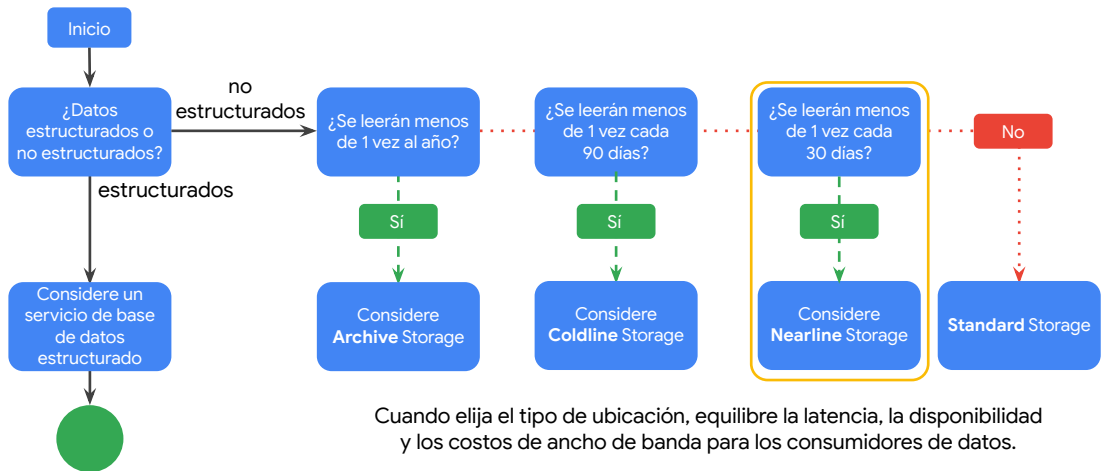
Si leerá sus datos menos de una vez al año, considere usar **Archive Storage**.

Cómo elegir una clase de almacenamiento



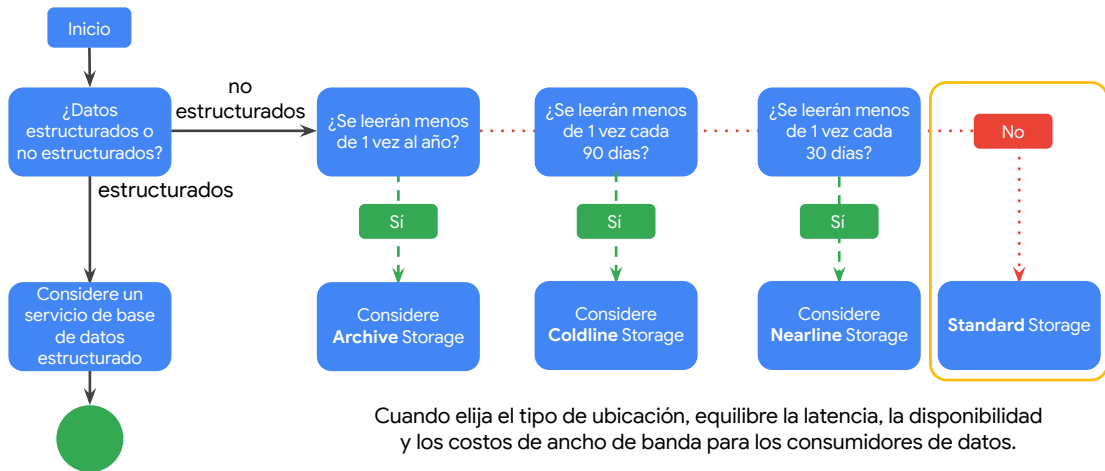
Si leerá sus datos menos de una vez cada 90 días, considere usar **Coldline Storage**.

Cómo elegir una clase de almacenamiento



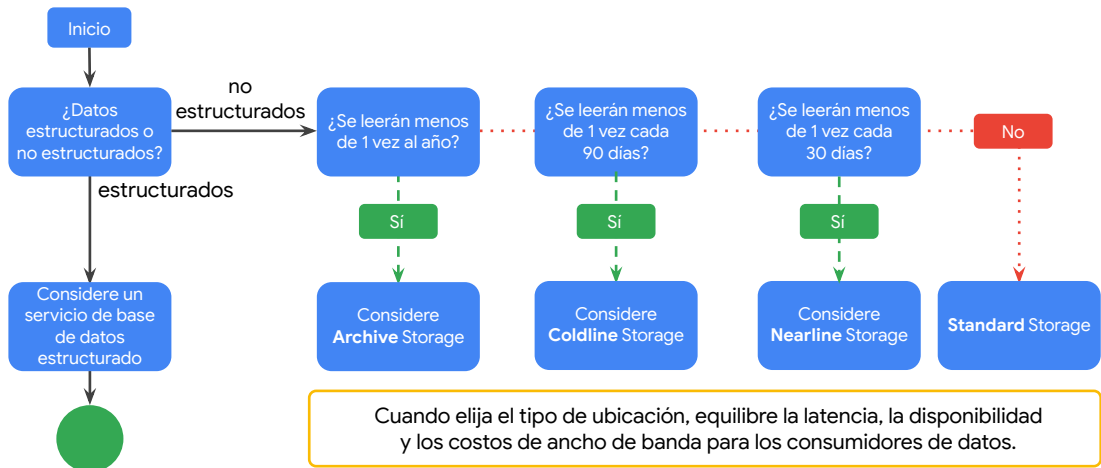
Si leerá sus datos menos de una vez cada 30 días, considere usar **Nearline Storage**.

Cómo elegir una clase de almacenamiento



Si hará lecturas y escrituras con mayor frecuencia, considere usar **Standard Storage**.

Cómo elegir una clase de almacenamiento



También debe tener en cuenta el tipo de ubicación:

- Use una **región** a fin de optimizar la latencia y el ancho de banda de red para los consumidores de datos (como las canalizaciones de analítica) que se agrupan en la misma región.
- Use una **región doble** cuando desee obtener ventajas de rendimiento similares a las de las regiones, pero también quiera aprovechar la disponibilidad que brinda la redundancia geográfica.
- Use una **multirregión** cuando desee entregar contenido a los consumidores de datos que están fuera de la red de Google y distribuidos en grandes áreas geográficas o cuando quiera aprovechar la disponibilidad mayor que brinda la redundancia geográfica.

Filestore es un servicio de almacenamiento de archivos administrado para aplicaciones



Filestore es un servicio de almacenamiento de archivos administrado destinado a aplicaciones que requieren una interfaz de sistema de archivos y un sistema de archivos compartido para datos.

Filestore es un servicio de almacenamiento de archivos administrado para aplicaciones

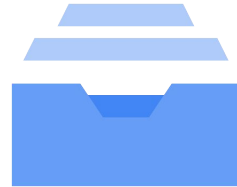
- Almacenamiento conectado a la red (NAS) completamente administrado para instancias de Compute Engine y GKE



Filestore brinda a los usuarios una experiencia simple y nativa para combinar el almacenamiento conectado a la red (NAS) administrado con sus instancias de Compute Engine y Google Kubernetes Engine.

Filestore es un servicio de almacenamiento de archivos administrado para aplicaciones

- Almacenamiento conectado a la red (NAS) completamente administrado para instancias de Compute Engine y GKE
- Rendimiento predecible



La posibilidad de ajustar la capacidad y el rendimiento de Filestore en forma independiente genera un rendimiento previsible más ágil para las cargas de trabajo basadas en archivos.

Filestore es un servicio de almacenamiento de archivos administrado para aplicaciones

- Almacenamiento conectado a la red (NAS) completamente administrado para instancias de Compute Engine y GKE
- Rendimiento predecible
- Compatibilidad completa con NFSv3



Filestore ofrece compatibilidad nativa con aplicaciones empresariales existentes y admite cualquier cliente compatible con NFSv3.

Filestore es un servicio de almacenamiento de archivos administrado para aplicaciones

- Almacenamiento conectado a la red (NAS) completamente administrado para instancias de Compute Engine y GKE
- Rendimiento predecible
- Compatibilidad completa con NFSv3
- Aumenta la escala de cargas de trabajo de alto rendimiento a cientos de TB



Las aplicaciones obtienen el beneficio de ciertas funciones (como el rendimiento del escalamiento horizontal, cientos de TB de capacidad y el bloqueo de archivos) sin necesidad de instalar ni mantener complementos especializados ni software del cliente.

Filestore tiene muchos casos de uso



Filestore tiene muchos casos de uso.

Filestore tiene muchos casos de uso

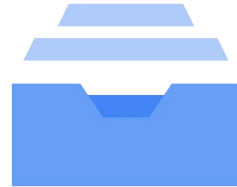
- Migración de aplicaciones



Con este servicio, puede acelerar la migración de aplicaciones empresariales. Muchas aplicaciones locales requieren una interfaz de sistema de archivos para los datos. A medida que estas aplicaciones continúan su migración a la nube, Filestore puede admitir una amplia variedad de aplicaciones empresariales que requieren un sistema de archivos compartidos.

Filestore tiene muchos casos de uso

- Migración de aplicaciones
- Renderización de contenido multimedia



Para renderizar contenido multimedia, puede activar fácilmente los archivos compartidos de Filestore en instancias de Compute Engine, lo que les permite a los artistas de efectos visuales colaborar en el mismo archivo compartido. Dado que los flujos de trabajo de renderización se suelen ejecutar en distintas flotas (“granjas de renderización”) de máquinas de procesamiento, las cuales activan un sistema de archivos compartidos, Filestore y Compute Engine pueden adaptarse para satisfacer las necesidades de renderización de los trabajos.

Filestore tiene muchos casos de uso

- Migración de aplicaciones
- Renderización de contenido multimedia
- Automatización de diseño electrónico (EDA)



La automatización de diseño electrónico (EDA) se basa en la administración de datos, requiere la capacidad de dividir cargas de trabajo en lotes en miles de núcleos y tiene grandes necesidades de memoria. Filestore ofrece la capacidad y el escalamiento necesarios para satisfacer las necesidades de los clientes de la industria manufacturera mediante una EDA de gran consumo y, también, garantiza que se pueda acceder a los archivos a nivel universal.

Filestore tiene muchos casos de uso

- Migración de aplicaciones
- Renderización de contenido multimedia
- Automatización de diseño electrónico (EDA)
- Análisis de datos



Las cargas de trabajo de análisis de datos incluyen modelos financieros de procesamiento complejo o análisis de datos del entorno. Estas cargas de trabajo son sensibles a la latencia. Filestore ofrece latencia baja para operaciones de archivos y, a medida que cambian las necesidades de rendimiento o capacidad, usted puede aumentar o disminuir el tamaño de las instancias fácilmente según lo necesite. Ya que Filestore es una capa de almacenamiento intercambiable y persistente, permite acceder inmediatamente a los datos para una analítica inteligente de alto rendimiento sin perder tiempo valioso cargando y descargando datos en las unidades de los clientes.

Filestore tiene muchos casos de uso

- Migración de aplicaciones
- Renderización de contenido multimedia
- Automatización de diseño electrónico (EDA)
- Análisis de datos
- Procesamiento de datos genómicos



La secuencia de datos genómicos requiere una enorme cantidad de datos sin procesar, unos miles de millones de datos por persona. Este tipo de análisis requiere velocidad, escalabilidad y seguridad. Filestore satisface las necesidades de las instituciones de investigación y las empresas que realizan investigaciones científicas, a la vez que ofrece precios predecibles para el rendimiento.

Filestore tiene muchos casos de uso

- Migración de aplicaciones
- Renderización de contenido multimedia
- Automatización de diseño electrónico (EDA)
- Análisis de datos
- Procesamiento de datos genómicos
- Administración de contenido web



Los desarrolladores web y los principales proveedores de hosting confían en Filestore para administrar y entregar contenido web, incluidas ciertas funciones como el hosting de WordPress.

Lab

Cloud Storage

Recordemos algunos de los conceptos de Cloud Storage que acabamos de analizar para aplicarlos en un lab.

En este lab, creará buckets y llevará a cabo muchas de las funciones avanzadas que se encuentran disponibles en Cloud Storage. Establecerá listas de control de acceso para limitar quiénes pueden acceder a sus datos y qué acciones se permiten. Usará la capacidad de proveer y administrar sus propias claves de encriptación para obtener seguridad adicional. Habilitará el control de versiones de objetos para hacer un seguimiento de los cambios que se realizaron en los datos y, además, configurará la administración de ciclos de vida con el fin de que los objetos se archiven o borren automáticamente después de un período determinado. Por último, usará la función de sincronización de directorios que mencionamos y compartirá sus buckets entre proyectos con la IAM.

Revisión del lab

Cloud Storage

En este lab, aprendió a crear y trabajar con buckets y objetos, y aplicó las siguientes funciones de Cloud Storage:

- Claves de encriptación proporcionadas por el cliente
- Listas de control de acceso
- Administración del ciclo de vida
- Control de versiones de objetos
- Sincronización de directorios
- Uso compartido de recursos entre proyectos con la IAM

Ahora que conoce muchas de las funciones avanzadas de Cloud Storage, recomendamos que las use en varias aplicaciones que antes no había considerado. Una forma común, rápida y fácil para comenzar a utilizar GCP es usar Cloud Storage como un servicio de copias de seguridad.

Puede continuar con un recorrido por el lab, pero recuerde que la interfaz de usuario de GCP puede cambiar, por lo que su entorno podría ser un poco diferente.

Temario

Cloud Storage y Filestore

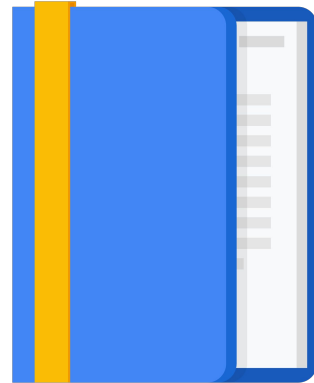
Cloud SQL

Cloud Spanner

Firestore

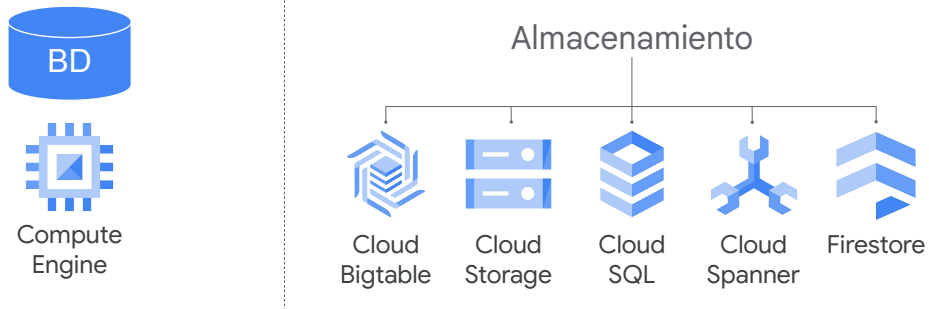
Cloud Bigtable

Memorystore



Analicemos los servicios estructurados o de base de datos relacional. El primero es Cloud SQL.

Cree su propia solución de base de datos o use un servicio administrado



¿Por qué debería usar un servicio de Google Cloud para SQL cuando puede instalar una imagen de la aplicación de SQL Server en una VM con Compute Engine?

La verdadera pregunta es si debería crear su propia solución de base de datos o usar un servicio administrado. Utilizar un servicio administrado tiene sus beneficios, así que analizaremos por qué debería usar Cloud SQL como servicio administrado en Google Cloud.

Cloud SQL es un servicio de base de datos completamente administrado (MySQL, PostgreSQL o Microsoft SQL Server)

- Los parches y las actualizaciones se aplican automáticamente.
- Usted administra los usuarios de MySQL.
- Cloud SQL admite numerosos clientes:
 - `gcloud sql`
 - Secuencias de comandos de App Engine y Google Workspace
 - Aplicaciones y herramientas
 - SQL Workbench y Toad
 - Aplicaciones externas mediante controladores estándar de MySQL



Cloud SQL

Cloud SQL es un servicio completamente administrado de bases de datos para MySQL, PostgreSQL o Microsoft SQL Server.

Cloud SQL es un servicio de base de datos completamente administrado (MySQL, PostgreSQL o Microsoft SQL Server)

- Los parches y las actualizaciones se aplican automáticamente.
- Usted administra los usuarios de MySQL.
- Cloud SQL admite numerosos clientes:
 - `gcloud sql`
 - Secuencias de comandos de App Engine y Google Workspace
 - Aplicaciones y herramientas
 - SQL Workbench y Toad
 - Aplicaciones externas mediante controladores estándar de MySQL



Esto significa que los parches y las actualizaciones se aplican automáticamente, pero, de todos modos, debe administrar los usuarios de MySQL con las herramientas de autenticación nativas que se incluyen en estas bases de datos.

Cloud SQL es un servicio de base de datos completamente administrado (MySQL, PostgreSQL o Microsoft SQL Server)

- Los parches y las actualizaciones se aplican automáticamente.
- Usted administra los usuarios de MySQL.
- **Cloud SQL admite numerosos clientes:**
 - `gcloud sql`
 - Secuencias de comandos de App Engine y Google Workspace
 - Aplicaciones y herramientas
 - SQL Workbench y Toad
 - Aplicaciones externas mediante controladores estándar de MySQL



Cloud SQL admite muchos clientes, como secuencias de comandos de Cloud Shell, App Engine y Google Workspace. También es compatible con otras aplicaciones y herramientas que ya conoce, como SQL Workbench, Toad y otras apps externas que usan controladores de MySQL estándar.

Instancia de Cloud SQL

Rendimiento:

- 30 TB de almacenamiento
- 40,000 IOPS
- 416 GB de RAM
- Escalamiento horizontal con réplicas de lectura

Opciones:

- MySQL 5.6, 5.7 (versión predeterminada) u 8.0
- PostgreSQL 9.6, 10, 11 o 12 (versión predeterminada)
- Microsoft SQL Server 2017



Cloud SQL ofrece un alto rendimiento y escalabilidad con un máximo de 30 TB de capacidad de almacenamiento, 40,000 IOPS y 416 GB de RAM por instancia. Puede aumentar la escala fácilmente hasta 64 núcleos de procesador y escalar horizontalmente con réplicas de lectura.

Instancia de Cloud SQL

Rendimiento:

- 30 TB de almacenamiento
- 40,000 IOPS
- 416 GB de RAM
- Escalamiento horizontal con réplicas de lectura

Opciones:

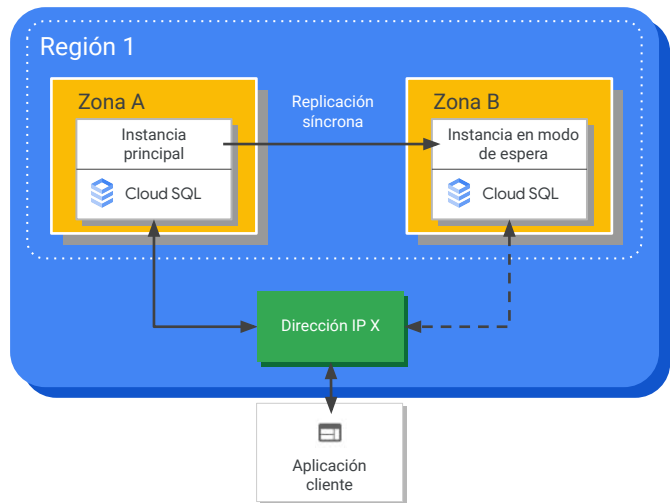
- MySQL 5.6, 5.7 (versión predeterminada) u 8.0
- PostgreSQL 9.6, 10, 11 o 12 (versión predeterminada)
- Microsoft SQL Server 2017



Actualmente, puede usar Cloud SQL con MySQL 5.6, 5.7 u 8.0, PostgreSQL 9.6, 10, 11 o 12, o cualquiera de las ediciones Web, Express, Standard o Enterprise de SQL Server 2017 disponibles en este momento.

Servicios de Cloud SQL

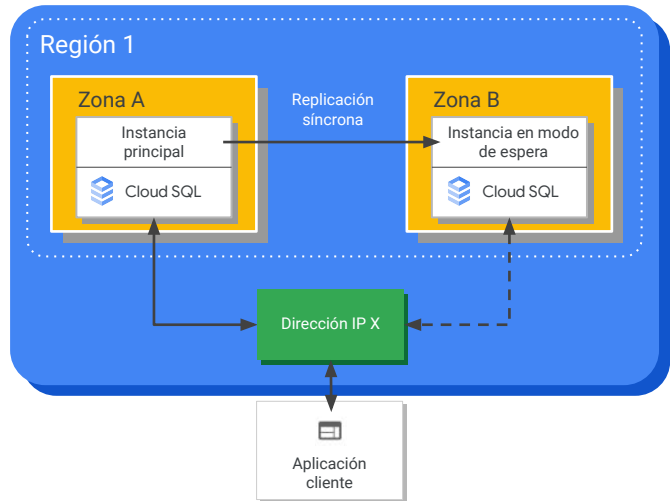
- Configuración de alta disponibilidad
- Servicio de copia de seguridad
- Importación y exportación
- Escalamiento
 - Vertical: Capacidad de máquina
 - Horizontal: Réplicas de lectura



Examinemos otros servicios que brinda Cloud SQL:

Servicios de Cloud SQL

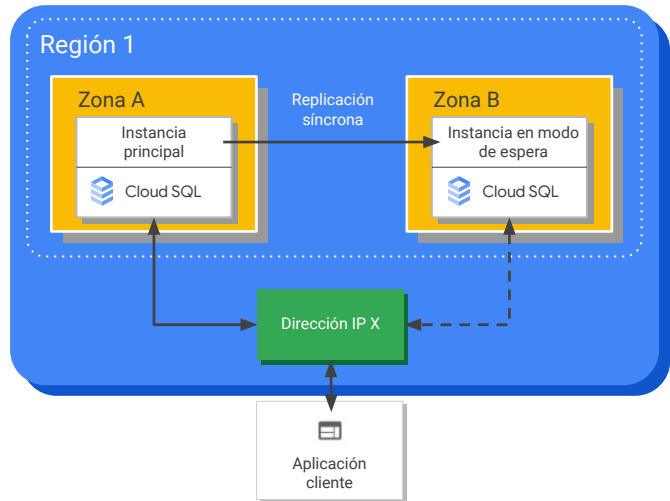
- Configuración de alta disponibilidad
- Servicio de copia de seguridad
- Importación y exportación
- Escalamiento
 - Vertical: Capacidad de máquina
 - Horizontal: Réplicas de lectura



En la configuración de alta disponibilidad, en una instancia regional, la configuración se compone de una *instancia principal* y otra *en espera*. A través de la [replicación síncrona](#) a los discos persistentes de cada zona, todas las operaciones de escritura realizadas en la instancia principal se replican en los discos de ambas zonas antes de que una transacción se informe como confirmada. Si ocurre una falla en una instancia o zona, el disco persistente se conecta a la instancia en espera y se convierte en la nueva instancia principal. Luego, los usuarios se redirigen a la nueva instancia principal. Este proceso se denomina *conmutación por error*.

Servicios de Cloud SQL

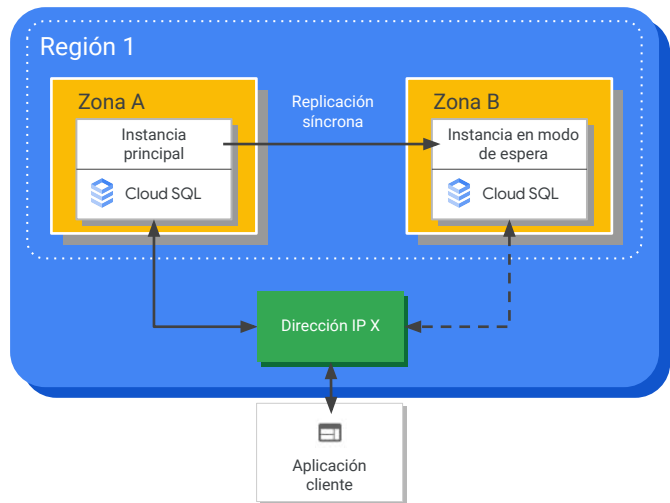
- Configuración de alta disponibilidad
- Servicio de copia de seguridad
- Importación y exportación
- Escalamiento
 - Vertical: Capacidad de máquina
 - Horizontal: Réplicas de lectura



Cloud SQL también ofrece copias de seguridad automáticas y a pedido con recuperación de un momento determinado.

Servicios de Cloud SQL

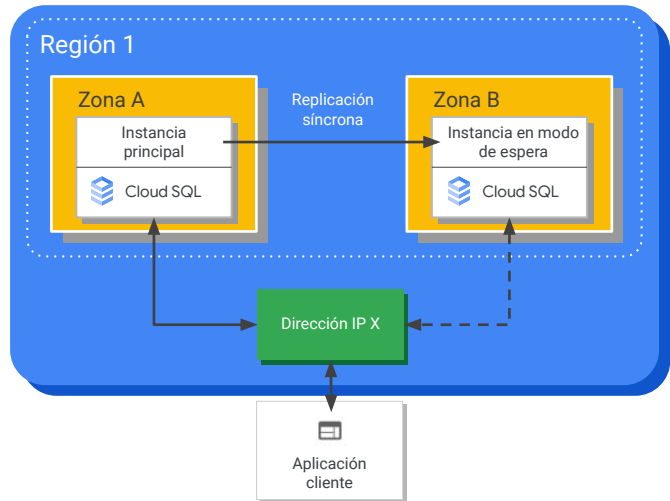
- Configuración de alta disponibilidad
- Servicio de copia de seguridad
- **Importación y exportación**
- Escalamiento
 - Vertical: Capacidad de máquina
 - Horizontal: Réplicas de lectura



Puede importar y exportar bases de datos con mysqldump o importar y exportar archivos CSV.

Servicios de Cloud SQL

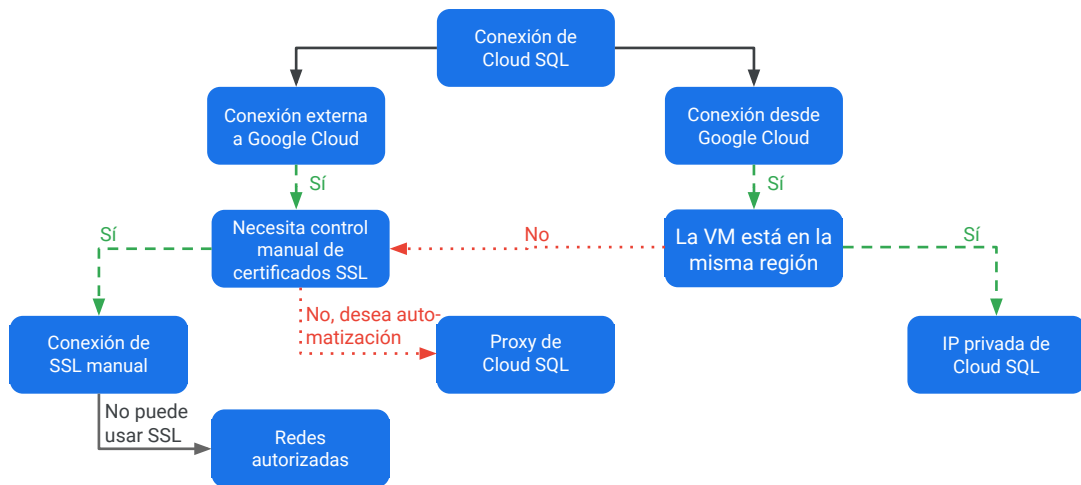
- Configuración de alta disponibilidad
- Servicio de copia de seguridad
- Importación y exportación
- Escalamiento
 - Vertical: Capacidad de máquina
 - Horizontal: Réplicas de lectura



Cloud SQL también puede escalar verticalmente (se necesita reiniciar las máquinas) o de manera horizontal con réplicas de lectura.

Dicho esto, si le preocupa la escalabilidad horizontal, recomendamos que considere usar Cloud Spanner, que analizaremos más adelante en este módulo.

Conexión a una instancia de Cloud SQL

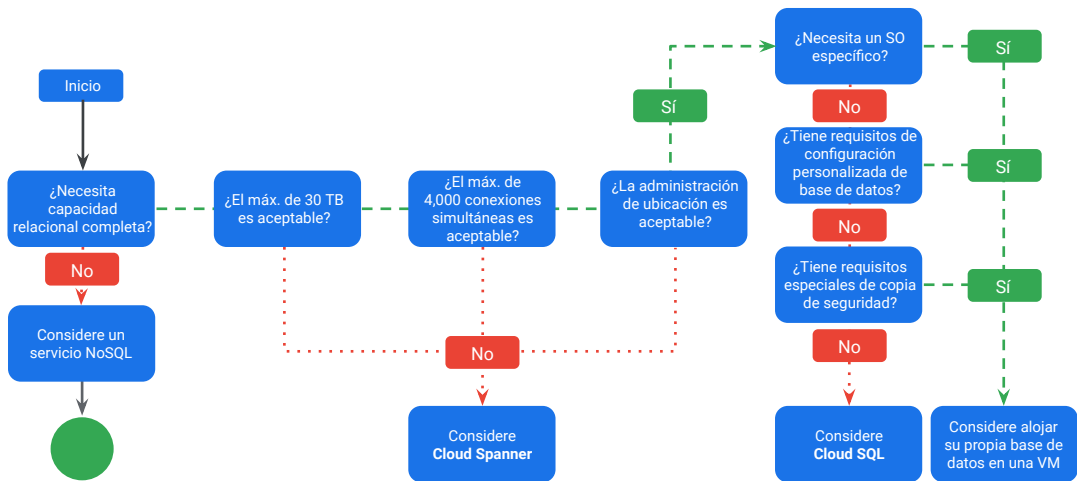


Elegir un tipo de conexión a su instancia de Cloud SQL influirá en la seguridad, el rendimiento y la capacidad de automatización. Si desea conectar una aplicación alojada en el mismo proyecto de Google Cloud que su instancia de Cloud SQL y ubicada en la misma región, elegir la conexión IP privada brindará la conexión con mejor rendimiento y seguridad mediante conectividad privada. En otras palabras, el tráfico nunca está expuesto a la Internet pública. Tenga en cuenta que conectarse a una dirección IP privada de Cloud SQL desde VMs ubicadas en la misma región solo es una recomendación basada en el rendimiento y no un requisito.

Si la aplicación está alojada en otra región o proyecto, o si desea conectarse a su instancia de Cloud SQL desde el exterior de Google Cloud, tiene 3 opciones. En este caso, recomendamos que use el proxy de Cloud SQL, que controla la autenticación, encriptación y rotación de claves por usted. Si necesita tener control manual de la conexión SSL, puede generar y rotar periódicamente los certificados por su cuenta. De lo contrario, puede usar una conexión no encriptada por medio de la autorización de una dirección IP específica para conectarse a su servidor SQL por una dirección IP externa.

Explorará estas opciones en un próximo lab, pero, si desea obtener más información sobre la IP privada, consulte la sección de vínculos de este video [\[https://cloud.google.com/sql/docs/mysql/private-ip\]](https://cloud.google.com/sql/docs/mysql/private-ip).

Elija Cloud SQL



A modo de resumen, exploremos este árbol de decisión que le permitirá encontrar el servicio de almacenamiento de datos adecuado con capacidad relacional completa.

Si necesita más de 30 TB de espacio de almacenamiento o más de 4,000 conexiones simultáneas a su base de datos, o si desea que su diseño de aplicación sea responsivo para escalamiento, disponibilidad y administración de ubicación cuando escale verticalmente a nivel global, considere usar Cloud Spanner, que explicaremos más adelante en este módulo.

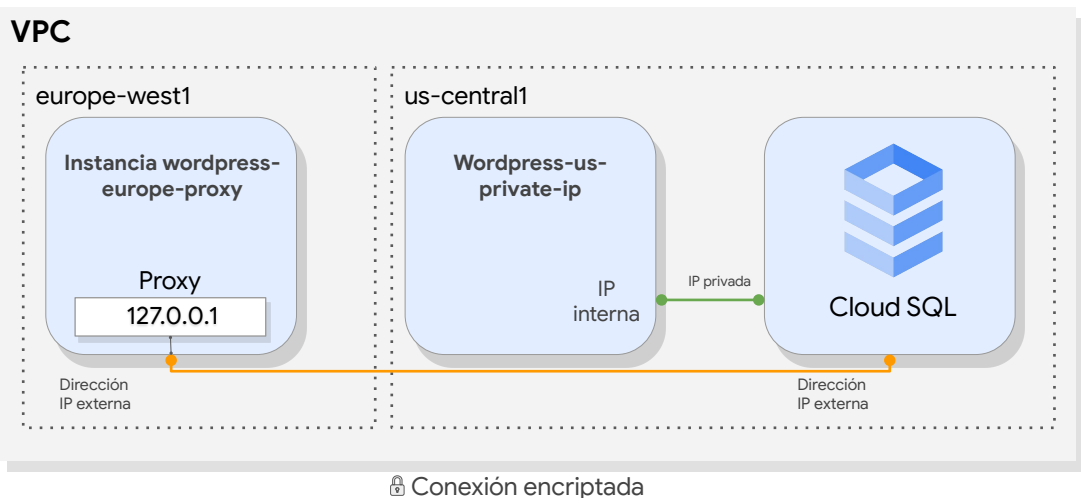
Si no le preocupan estas restricciones, pregúntese si necesita un SO específico o si tiene requisitos de configuración personalizada de la base de datos o requisitos especiales de copia de seguridad. De ser así, considere alojar su propia base de datos en una VM por medio de Compute Engine. De lo contrario, le recomendamos que use Cloud SQL como un servicio completamente administrado para sus bases de datos relacionales.

Si se convenció de que usar Cloud SQL como un servicio administrado es mejor que utilizar o reimplementar su solución de MySQL existente, consulte la sección de vínculos, en la que encontrará una solución para migrar de MySQL a Cloud SQL [<https://cloud.google.com/solutions/migrating-mysql-to-cloudsql-concept>]

Lab

Cómo implementar Cloud SQL

Recordemos algunos de los conceptos de Cloud SQL que discutimos para aplicarlos en un lab.



En este lab, configurará un servidor de Cloud SQL y aprenderá a conectarle una aplicación con un proxy a través de una conexión externa. Además, configurará una conexión a través de un vínculo de IP privada que ofrece beneficios de rendimiento y seguridad. En este lab, la aplicación elegida para hacer la demostración es WordPress. No obstante, la información y las prácticas recomendadas sirven para cualquier aplicación que requiera SQL Server.

Al final de este lab, tendrá 2 instancias de trabajo de un frontend de WordPress conectadas a través de 2 tipos de conexión diferentes a su backend de la instancia de SQL, como se muestra en este diagrama.

Revisión del lab

Cómo implementar Cloud SQL

En este lab, creó una base de datos de Cloud SQL y la configuró para usar una conexión externa a través de un proxy seguro y una dirección IP privada, que es más segura y eficiente.

Recuerde que solo puede conectarse a través de una IP privada si la aplicación y el servidor de Cloud SQL se encuentran en la misma región y son parte de la misma red de VPC. Si su aplicación está alojada en otra región, VPC o incluso proyecto, use un proxy para proteger su conexión a través de la conexión externa.

Puede continuar con un recorrido por el lab, pero recuerde que la interfaz de usuario de GCP puede cambiar, por lo que su entorno podría ser un poco diferente.

Temario

Cloud Storage y Filestore

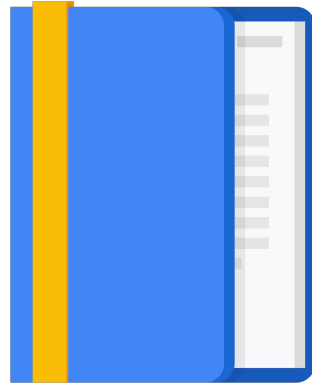
Cloud SQL

Cloud Spanner

Firestore

Cloud Bigtable

Memorystore



Si Cloud SQL no se adapta a sus requisitos porque necesita escalabilidad horizontal, considere usar Cloud Spanner.

Cloud Spanner combina los beneficios de la estructura de las bases de datos relacionales con el escalamiento horizontal no relacional

- Escala a petabytes
- Coherencia sólida
- Alta disponibilidad
- Se usa para aplicaciones financieras y de inventario
- Tiempo de actividad mensual
 - Multi-Regional: 99.999%
 - Regional: 99.99%



Cloud Spanner es un servicio creado específicamente para la nube y que combina los beneficios de la estructura de las bases de datos relacionales con el escalamiento horizontal no relacional.

Este servicio puede entregar petabytes de capacidad y brinda coherencia transaccional a escala global, esquemas, SQL y replicación síncrona automática para ofrecer alta disponibilidad. Los casos de uso incluyen aplicaciones financieras y de inventario que suelen entregarse mediante tecnología de bases de datos relacionales.

Como se muestra en esta diapositiva, tendrá diferentes ANS de tiempo de actividad mensual dependiendo de si creó una instancia regional o multirregional. Sin embargo, para ver los porcentajes actualizados, siempre debe consultar la documentación, que encontrará en la sección de vínculos de este video:

[\[https://cloud.google.com/spanner/sla\]](https://cloud.google.com/spanner/sla)

Características

| | Cloud Spanner | | BD relacional | | BD no relacional | |
|----------------|---------------|------------|---------------|-----------------------|------------------|--------------|
| Esquema | ✓ | Sí | ✓ | Sí | ✗ | No |
| SQL | ✓ | Sí | ✓ | Sí | ✗ | No |
| Coherencia | ✓ | Sólida | ✓ | Sólida | ✗ | Eventual |
| Disponibilidad | ✓ | Alta | ✗ | Conmutación por error | ✓ | Alta |
| Escalabilidad | ✓ | Horizontal | ✗ | Vertical | ✓ | Horizontal |
| Replicación | ✓ | Automática | ↻ | Configurable | ↻ | Configurable |

Comparemos Cloud Spanner con las bases de datos relacionales y no relacionales. Como una base de datos relacional, Cloud Spanner tiene esquema, SQL y coherencia sólida. Además, como una base de datos no relacional, Cloud Spanner ofrece alta disponibilidad, escalabilidad horizontal y replicación configurable.

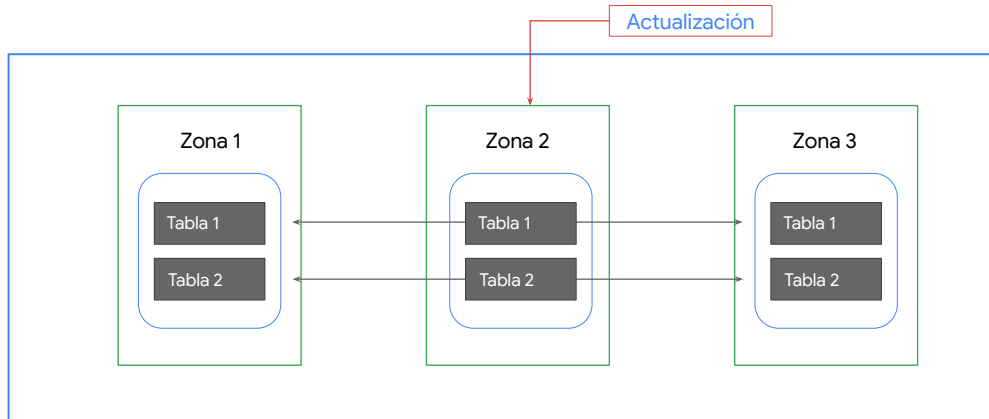
Como mencionamos, Cloud Spanner ofrece lo mejor de las bases de datos relacionales y no relacionales. Estas funciones permiten casos de uso fundamentales, como la creación de sistemas coherentes para transacciones y administración de inventario en los sectores de servicios financieros y venta minorista. Para comprender mejor cómo funciona todo esto, estudiemos la arquitectura de Cloud Spanner.

Arquitectura de Cloud Spanner



Una instancia de Cloud Spanner replica datos en N zonas de la nube, que se pueden encontrar en una o varias regiones. Puede ubicar la base de datos en la región que usted elija. Esta arquitectura brinda alta disponibilidad y ubicación global.

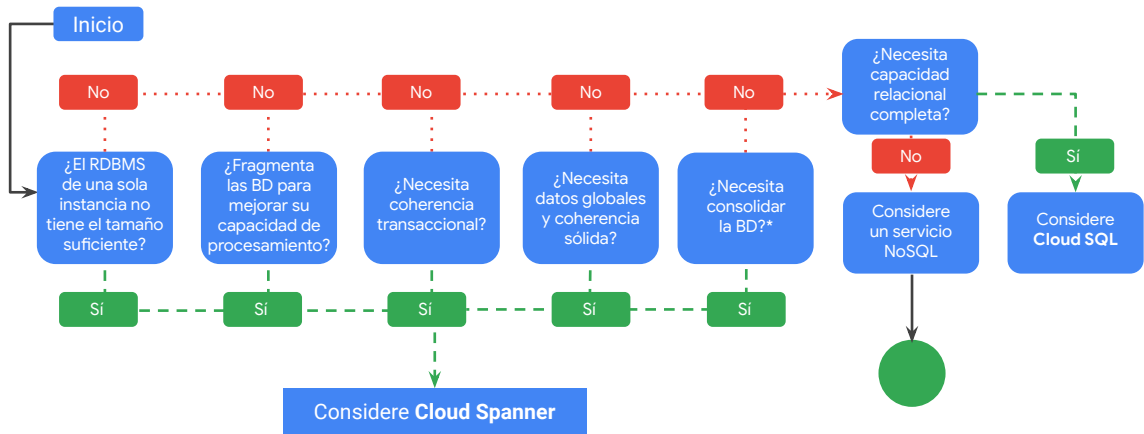
La replicación de datos se sincroniza entre zonas mediante la red de fibra óptica global de Google



La replicación de datos se sincronizará entre zonas mediante la red de fibra óptica global de Google. Se utilizan relojes atómicos para garantizar atomicidad en cualquier momento que actualice sus datos.

Eso es todo lo que explicaremos sobre Cloud Spanner. Como el enfoque de este módulo es comprender las circunstancias en las que usaría Cloud Spanner, analicemos el árbol de decisión.

Elija Cloud Spanner



Si las bases de datos relacionales no tienen el tamaño suficiente; está fragmentando las bases de datos para aumentar el rendimiento y la capacidad de procesamiento; necesita coherencia transaccional, datos globales y coherencia sólida, o solo desea consolidar su base de datos, considere usar Cloud Spanner.

Si no necesita nada de eso ni capacidades relacionales, considere usar un servicio NoSQL como Cloud Firestore, que explicaremos a continuación.

Si se convenció de que usar Cloud Spanner como un servicio administrado es mejor que utilizar o reimplementar su solución de MySQL existente, consulte la sección de vínculos, en la que encontrará una solución para migrar de MySQL a Cloud Spanner [<https://cloud.google.com/solutions/migrating-mysql-to-spanner>]

Temario

Cloud Storage y Filestore

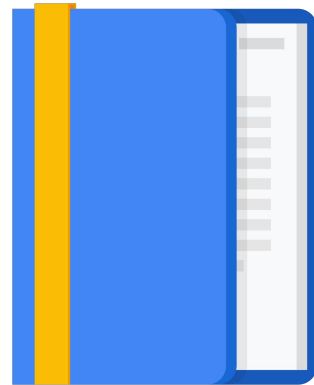
Cloud SQL

Cloud Spanner

Firestore

Cloud Bigtable

Memorystore



Si está buscando una base de datos NoSQL altamente escalable para sus aplicaciones, considere usar Firestore.

Firestore es una base de datos de documentos NoSQL

- Simplifica el almacenamiento, la sincronización y la consulta de datos
- Aplicaciones web, de IoT y para dispositivos móviles a escala global
- Soporte sin conexión y sincronización en vivo
- Funciones de seguridad
- Transacciones ACID
- Replicación multirregión
- Motor de consultas potente



Firestore

Firestore es una base de datos NoSQL de documentos nativa de la nube, rápida, sin servidores y completamente administrada que simplifica almacenar, sincronizar y consultar datos en apps web, de la IoT y para dispositivos móviles a escala global. Las bibliotecas cliente ofrecen sincronización en vivo y soporte sin conexión, y las funciones de seguridad y las integraciones en Firebase y Google Cloud agilizan la creación de apps sin servidores.

Firestore también admite transacciones ACID. Por ello, si falla cualquiera de las operaciones de la transacción y no se puede intentar, fallará toda la transacción.

Además, sus datos están protegidos y disponibles con la replicación multirregional automática y la coherencia sólida, incluso si ocurre algún desastre. Firestore incluso le permite ejecutar consultas sofisticadas en los datos NoSQL sin degradar el rendimiento. Esta función brinda una mayor flexibilidad en la forma en la que estructuras los datos.

Firestore es la próxima generación de Datastore

Modo Datastore (proyectos nuevos de servidor):

- Compatible con aplicaciones de Datastore
- Coherencia sólida
- Sin límites de grupo de entidades

Modo nativo (nuevas aplicaciones web y para dispositivos móviles):

- Capa de almacenamiento con coherencia sólida
- Modelo de datos de colección y documento
- Actualizaciones en tiempo real
- Bibliotecas cliente para dispositivos móviles y la Web

En realidad, Firestore es la próxima generación de Datastore. Puede operar en modo Datastore, por lo que es retrocompatible con esa plataforma. Crear una base de datos de Firestore en modo Datastore le permite acceder a una capa de almacenamiento mejorada de Firestore sin perder el comportamiento del sistema de Datastore.

Esto quita las siguientes limitaciones de Datastore:

- Las consultas ya no tienen coherencia eventual, sino sólida.
- Las transacciones ya no están limitadas a 25 grupos de entidades.
- Las operaciones de escritura en un grupo de entidad ya no están limitadas a 1 por segundo.

Estas son algunas de las funciones nuevas que incluye Firestore en modo nativo:

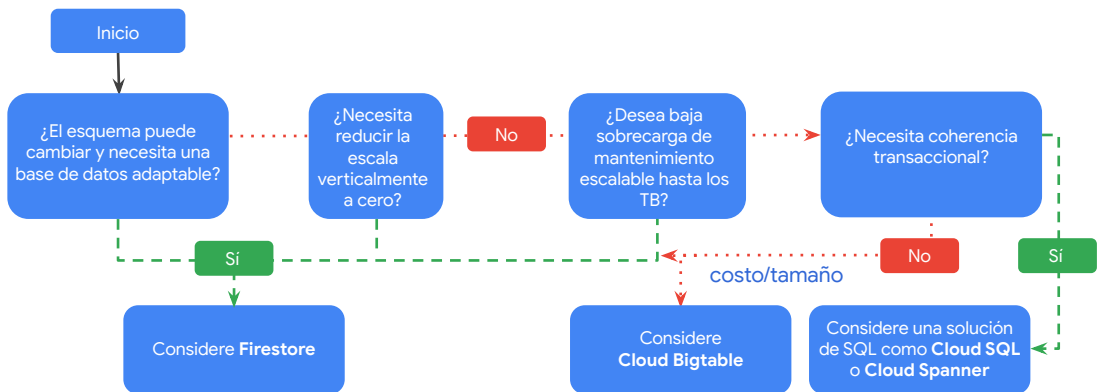
- Una nueva capa de almacenamiento de coherencia sólida
- Un modelo de recopilación y datos de documentos
- Actualizaciones en tiempo real
- Bibliotecas cliente para dispositivos móviles y la Web

Firestore es retrocompatible con versiones anteriores de Datastore, pero el nuevo modelo de datos, las actualizaciones en tiempo real y las funciones de bibliotecas cliente para dispositivos móviles y la Web no lo son. Para acceder a todas las nuevas funciones de Firestore, debe usar Firestore en modo nativo. La regla general es usar

Firestore en modo Datastore para los nuevos proyectos de servidor y el modo nativo con las nuevas apps web y para dispositivos móviles.

Dado que es la nueva generación de Datastore, Firestore es compatible con todas las bibliotecas cliente y API de Datastore. Los usuarios existentes de Datastore recibirán una actualización en vivo de Firestore automáticamente en una fecha próxima. Para obtener más información, consulte la sección de vínculos de este video: [<https://cloud.google.com/datastore/docs/firestore-or-datastore>, <https://cloud.google.com/datastore/docs/upgrade-to-firestore>]

Elija Firestore



A modo de resumen, exploremos este árbol de decisión que le permitirá determinar si Firestore es el servicio de almacenamiento adecuado para sus datos.

Si su esquema está sujeto a cambios y necesita una base de datos adaptable o reducir la escala a cero, o desea baja sobrecarga de mantenimiento hasta los terabytes, considere usar Firestore.

Además, si no requiere coherencia transaccional, recomendamos que considere Cloud Bigtable en función del costo o el tamaño.

A continuación, veremos Cloud Bigtable.

Temario

Cloud Storage y Filestore

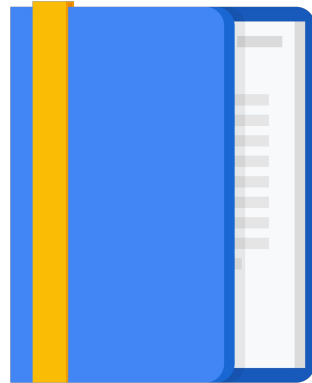
Cloud SQL

Cloud Spanner

Firestore

Cloud Bigtable

Memorystore



Si no requiere coherencia transaccional, recomendamos que considere Cloud Bigtable.

Cloud Bigtable es un servicio de base de datos de macrodatos NoSQL

- Escala de petabytes
- Latencia coherente inferior a 10 ms
- Escalabilidad sin interrupciones de la capacidad de procesamiento
- Aprende y se ajusta a patrones de acceso
- Ideal para AdTech, IoT y tecnología financiera
- Motor de almacenamiento para aplicaciones de AA
- Integración fácil en herramientas de macrodatos de código abierto



Cloud
Bigtable



APACHE
HBASE

Cloud Bigtable es un servicio de base de datos NoSQL completamente administrado a escala de petabytes con muy baja latencia. Ofrece un escalamiento sin interrupciones de la capacidad de procesamiento y es capaz de aprender para ajustarse a patrones de acceso específicos. De hecho, Cloud Bigtable es la misma base de datos que utilizan muchos de los servicios centrales de Google, como la Búsqueda, Analytics, Maps y Gmail.

Cloud Bigtable es una excelente opción para aplicaciones operativas y analíticas, como la IoT, las estadísticas de usuarios y los análisis de datos financieros, ya que admite alta capacidad de procesamiento de lectura y escritura con baja latencia. También es un motor de almacenamiento ideal para las aplicaciones de aprendizaje automático.

Cloud Bigtable se integra fácilmente a herramientas populares de macrodatos como Hadoop, Cloud Dataflow y Cloud Dataproc. Además, Cloud Bigtable es compatible con la API de HBase, un estándar de la industria de código abierto que ayuda a los equipos de desarrollo a dar sus primeros pasos. Cloud Dataflow y Cloud Dataproc se explican más adelante en el curso. Para obtener información sobre la API de HBase, consulte la sección de vínculos de este video: [<https://hbase.apache.org/>]

Modelo de almacenamiento de Cloud Bigtable

The diagram illustrates a Cloud Bigtable storage model. It features a table with a family of columns named "sigue a". The table has four columns: "gwashington", "jadams", "tjefferson", and "wmckinley". The rows represent different users: "gwashington", "jadams", "tjefferson", and "wmckinley". The table contains multiple versions of data, indicated by the number "1" in the cells. A green arrow points to the "tjefferson" column, labeled "versiones múltiples".

| Clave de fila | gwashington | jadams | tjefferson | wmckinley |
|---------------|-------------|--------|------------|-----------|
| gwashington | | 1 | | |
| jadams | 1 | | 1 | |
| tjefferson | 1 | 1 | | 1 |
| wmckinley | | | 1 | |

Cloud Bigtable almacena datos en tablas escalables de forma masiva, cada una de las cuales es un mapa clave-valor ordenado. La tabla consta de filas que, por lo general, describen una sola entidad y columnas que contienen valores individuales para cada fila. Cada fila se indexa según una clave de fila, y las columnas que se relacionan entre sí se suelen agrupar en una familia de columnas. Cada columna se identifica con una combinación de la familia de columnas y un calificador de columna, que es un nombre único en la familia.

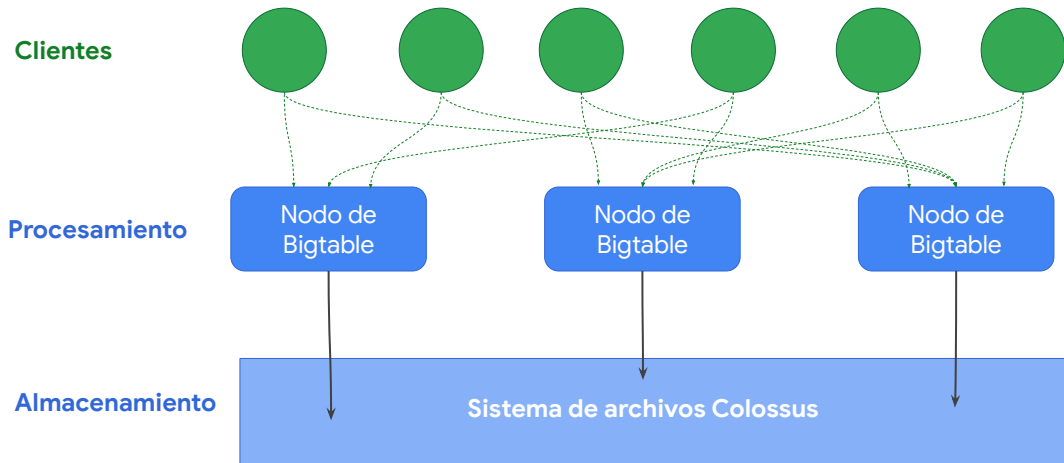
Cada intersección entre fila y columna puede contener varias celdas, o versiones, en distintas marcas de tiempo, lo que permite obtener un registro de cómo han cambiado los datos almacenados con el tiempo. Las tablas de Cloud Bigtable están dispersas. Si una celda no contiene ningún dato, no ocupará espacio.

El ejemplo que se muestra aquí es una red social hipotética para presidentes de Estados Unidos, en la que cada presidente puede seguir publicaciones de otros. Destaquemos algunos aspectos:

- La tabla contiene una familia de columnas llamada "sigue a" que contiene varios calificadores de columnas.
- Los calificadores de columnas se usan como datos. En este diseño, se aprovecha la dispersión de las tablas de Cloud Bigtable y el hecho de que se pueden agregar nuevos calificadores de columnas según los cambios de sus datos.
- El nombre de usuario se usa como la clave de fila. Si suponemos que los

- nombres de usuario se extienden de manera uniforme en el alfabeto, el acceso a los datos será parejo en toda la tabla.

El procesamiento está separado del almacenamiento

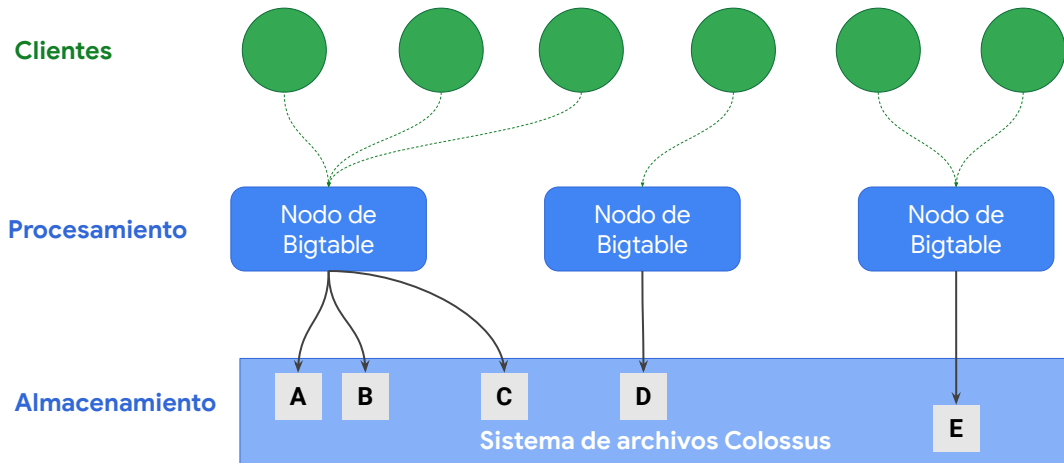


En este diagrama, se muestra una versión simplificada de la arquitectura general de Cloud Bigtable. Se ilustra que el procesamiento, realizado a través de un grupo de servidor y nodos de frontend, se controla de forma separada del almacenamiento.

Una tabla de Cloud Bigtable se fragmenta en bloques de filas contiguas, llamados tablets, que permiten balancear la carga de trabajo de las consultas. Si utilizó la API de HBase, las tablets son similares a las regiones de esa plataforma.

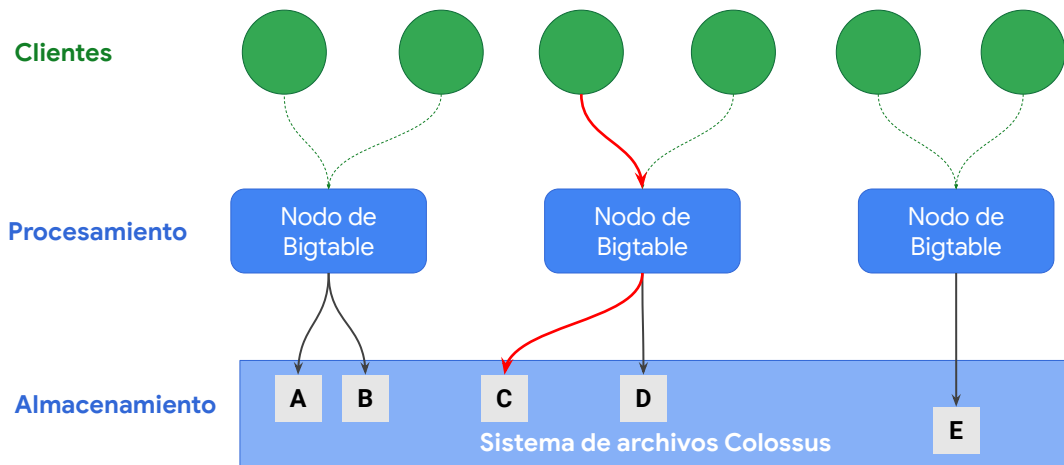
Las tablets se almacenan en Colossus, el sistema de archivos de Google, en formato SSTable. Un archivo SSTable proporciona un mapa inmutable, ordenado y persistente de claves a valores, en el que los valores y las claves son strings de bytes arbitrarios.

Aprende patrones de acceso



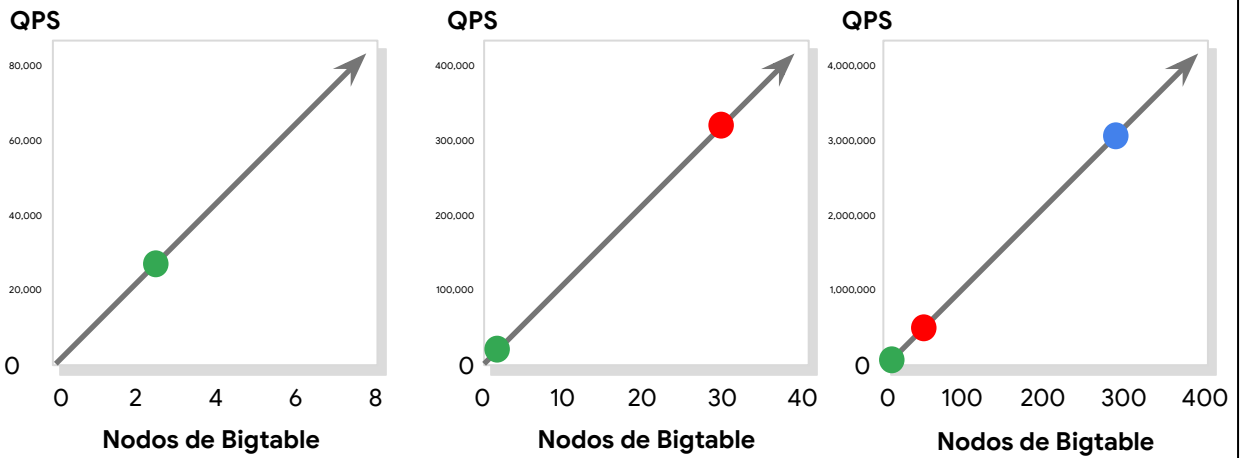
Como mencionamos antes, Cloud Bigtable aprende para ajustarse a patrones de acceso específicos. Si un nodo de Bigtable determinado accede frecuentemente a un subconjunto de datos...

Se rebalancea sin mover datos



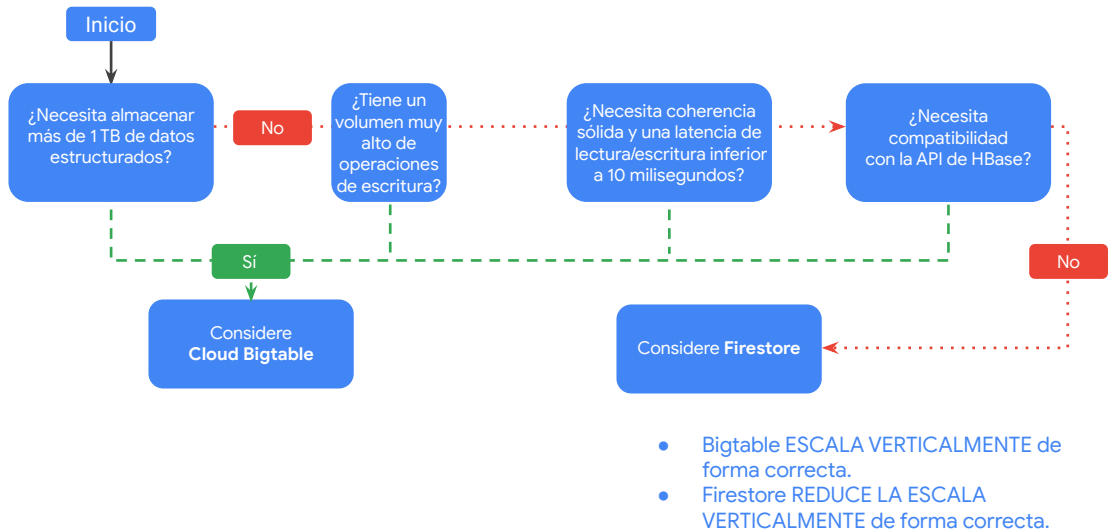
... Cloud Bigtable actualizará los índices para que otros nodos puedan distribuir esa carga de trabajo uniformemente, como se muestra aquí.

La capacidad de procesamiento se escala de forma lineal



Esa capacidad de procesamiento escala de forma lineal, así que, por cada nodo que agregue, verá una escala lineal de rendimiento de capacidad de procesamiento, hasta cientos de nodos.

Elija Cloud Bigtable



En resumen, si necesita almacenar más de 1 TB de datos estructurados, tienen un gran volumen de operaciones de escritura, necesita una latencia de lectura/escritura inferior a 10 milisegundos con coherencia sólida o necesita un servicio de almacenamiento que sea compatible con la API de HBase, considere usar Cloud Bigtable.

Si no necesita nada de eso y busca un servicio de almacenamiento que reduzca la escala verticalmente de forma correcta, considere usar Firestore.

Hablando de escalamiento, el clúster de Cloud Bigtable más pequeño que se puede crear tiene tres nodos y admite 30,000 operaciones por segundo. Recuerde que paga por los nodos mientras están operativos, sin importar si su aplicación los usa o no.

Temario

Cloud Storage y Filestore

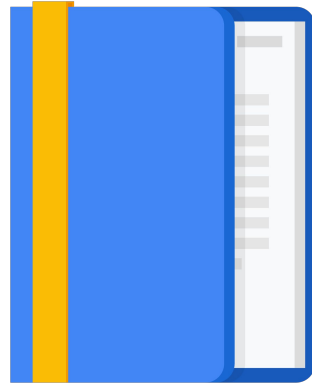
Cloud SQL

Cloud Spanner

Firestore

Cloud Bigtable

[Memorystore](#)



Veamos una descripción general rápida de Memorystore.

Memorystore es un servicio de Redis completamente administrado

- Servicio de almacenamiento de datos en memoria
- Enfóquese en crear apps excelentes
- Alta disponibilidad, conmutación por error, aplicación de parches y supervisión
- Latencia de menos de un milisegundo
- Instancias de hasta 300 GB
- Capacidad de procesamiento de 12 Gbps
- Migración lift-and-shift sencilla



Memorystore

Memorystore para Redis proporciona un servicio de almacenamiento de datos en la memoria completamente administrado, con base en la infraestructura escalable, segura y con alta disponibilidad que administra Google. Las aplicaciones que se ejecutan en Google Cloud pueden lograr un rendimiento extremo aprovechando el servicio de Redis altamente escalable, disponible y seguro sin la carga de administrar implementaciones complejas de Redis, lo que permite dedicar más tiempo a escribir código para que pueda enfocarse en crear aplicaciones excelentes.

En Memorystore, también se automatizan tareas complejas, como la habilitación de alta disponibilidad, la conmutación por error, la aplicación de parches y la supervisión. Las instancias de alta disponibilidad se replican en dos zonas y proporcionan un ANS de disponibilidad de un 99.9%.

Puede lograr con facilidad latencias inferiores a un milisegundo y obtener la capacidad de procesamiento que necesitan sus aplicaciones. Comience con el nivel de servicio y el tamaño más bajos y, luego, aumente la instancia sin esfuerzo y con un efecto mínimo en la disponibilidad de las aplicaciones. Memorystore admite instancias de hasta 300 GB y una capacidad de procesamiento de red de hasta 12 Gbps.

Como Memorystore para Redis es completamente compatible con el protocolo Redis, puede realizar la migración lift-and-shift de sus aplicaciones desde Redis de código abierto a Memorystore sin cambios de código gracias a la función de importación/exportación. No es necesario que aprenda a usar herramientas nuevas

porque todas las herramientas y bibliotecas cliente disponibles se pueden usar en este proceso.

Repaso

Servicios de almacenamiento y bases de datos



En este módulo, hablamos sobre diferentes servicios de bases de datos y almacenamiento que ofrece Google Cloud. Específicamente, aprendimos sobre Cloud Storage, un almacén de objetos completamente administrado; Cloud SQL, un servicio de base de datos MySQL y PostgreSQL completamente administrado; Cloud Spanner, un servicio de base de datos relacional con coherencia transaccional, escala global y alta disponibilidad; Firestore, una base de datos de documentos NoSQL completamente administrada; Cloud Bigtable, una base de datos de columnas anchas NoSQL completamente administrada, y Memorystore, un servicio de almacenamiento de datos en memoria completamente administrado para Redis.

Desde la perspectiva de la infraestructura, el objetivo era comprender qué servicios están disponibles y cómo se usan en diferentes circunstancias. Definir una estrategia de datos completa está fuera del alcance de este curso, pero Google ofrece cursos sobre ingeniería de datos y aprendizaje automático en Google Cloud en los que se aborda la estrategia de datos.