



Elija soluciones de
almacenamiento y datos

Priyanka Vergadia
Developer Advocate, Google Cloud

En este módulo, analizaremos las soluciones de almacenamiento y datos de Google Cloud, además de cómo seleccionar la más apropiada para cumplir con sus requisitos comerciales y técnicos.

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

Google Cloud ofrece un conjunto amplio de opciones de almacenamiento diferentes orientadas a distintos tipos y tamaños de datos, ciclos de vida y patrones de acceso a datos.

Analizaremos cómo almacenar datos binarios con Cloud Storage, datos relacionales con Cloud SQL o Spanner, y datos NoSQL o no estructurados con Firestore y Bigtable.

Además, consideraremos almacenar en caché para acceder a los datos con más rapidez a través de Memorystore y, por último, agregar datos para informes y consultas con BigQuery como almacén de datos.

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

Google Cloud ofrece un conjunto amplio de opciones de almacenamiento diferentes orientadas a distintos tipos y tamaños de datos, ciclos de vida y patrones de acceso a datos.

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

Analizaremos cómo almacenar datos binarios con Cloud Storage,

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

datos relacionales con Cloud SQL o Spanner,

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

y datos NoSQL o no estructurados con Firestore y Bigtable.

Objetivos de aprendizaje

- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

Además, consideraremos almacenar en caché para acceder a los datos con más rapidez a través de Memorystore y,

Objetivos de aprendizaje

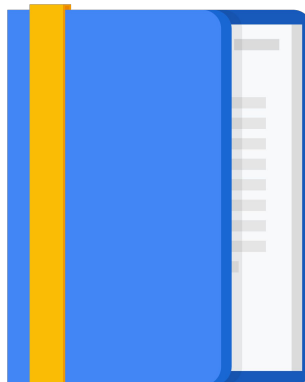
- Elegir el servicio de almacenamiento de datos de Google Cloud adecuado según el caso de uso, la durabilidad, la disponibilidad, la escalabilidad y el costo
- Almacenar datos binarios con Cloud Storage
- Almacenar datos relacionales con Cloud SQL y Spanner
- Almacenar datos NoSQL con Firestore y Cloud Bigtable
- Almacenar datos en caché para un acceso rápido con Memorystore
- Agregar datos para informes y consultas con BigQuery como almacén de datos

por último, agregar datos para informes y consultas con BigQuery como almacén de datos.

Temario








Características clave del
almacenamiento

Elija soluciones de
almacenamiento y datos de
Google Cloud



Comencemos analizando las características clave del almacenamiento.

Google Cloud: cartera administrada de almacenamiento y bases de datos

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore

Google Cloud tiene una amplia variedad de opciones administradas de almacenamiento y bases de datos en su cartera. Conocer las características de cada una y poder seleccionar una solución adecuada es tan importante como tener un arquitecto durante el proceso de diseño. A nivel general, los servicios abarcan desde almacenamiento de objetos relacionales NoSQL y almacén de datos hasta almacenamiento en memoria. Estos servicios son completamente administrados, escalables y están respaldados por ANS líderes de la industria.

Elegir la solución de almacenamiento es la correcta en función de sus requisitos conlleva un equilibrio entre una serie de características como tipos de datos, escala, durabilidad, disponibilidad y requisitos de ubicación, entre otras. En este módulo, analizaremos formas en las que puede tomar la mejor decisión con base en sus requisitos.

Los diferentes servicios de almacenamiento de datos tienen distintos ANS de disponibilidad

Opción de almacenamiento	Porcentaje de ANS de disponibilidad
Cloud Storage (bucket multirregional)	>=99.95
Cloud Storage (bucket regional)	99.9
Cloud Storage (Coldline)	99.0
Spanner (multirregión)	99.999
Spanner (región única)	99.99
Firestore (multirregión)	99.999
Firestore (región única)	99.99

Los diferentes servicios de almacenamiento de datos tienen distintos ANS de disponibilidad. A menudo, el ANS de disponibilidad de un servicio depende de la configuración de este. Por ejemplo, como se muestra en la diapositiva, la disponibilidad de Cloud Storage varía según si se crean buckets regionales, multirregionales o Coldline.

Lo mismo se observa en Cloud Spanner y Firestore, cuyos parámetros de configuración multirregionales ofrecen disponibilidad más alta que los de región única. Es aquí que los requisitos son extremadamente importantes, ya que ayudarán a tomar decisiones fundamentadas en cuanto al almacenamiento.

Los ANS de disponibilidad suelen definirse por mes. El término “porcentaje de tiempo de actividad mensual” corresponde al total de minutos en un mes, menos los minutos de tiempo de inactividad que afectó a todos los períodos de inactividad en un mes, divididos por la cantidad total de minutos en un mes.

Para conocer las cifras actualizadas de ANS, consulte la documentación.

La durabilidad representa las posibilidades de perder datos

Evitar la pérdida de datos es una responsabilidad compartida.

Opción de almacenamiento	Google Cloud proporciona	Qué debe hacer
Cloud Storage	Durabilidad de un 99.999999999% Control de versiones (opcional)	Active el control de versiones
Discos	Instantáneas	Programe trabajos de instantáneas
Cloud SQL	Copias de seguridad a pedido Copias de seguridad automáticas Recuperación de un momento determinado Servidor de conmutación por error (opcional)	Crear copias de seguridad en cualquier momento Ejecutar copias de seguridad de bases de datos SQL
Spanner	Replicación automática	Ejecutar trabajos de exportación
Firestore	Replicación automática	Ejecutar trabajos de exportación

La durabilidad de los datos representa las probabilidades de perderlos. Según la solución de almacenamiento, la durabilidad es una responsabilidad compartida. La responsabilidad de Google Cloud es garantizar que los datos sean duraderos en caso de que falle el hardware. Usted es responsable de crear copias de seguridad de sus datos.

Por ejemplo, Cloud Storage le proporciona un 99,999999999% de durabilidad y, además, incluye control de versiones. Sin embargo, es su responsabilidad determinar cuándo usar esta función. Se recomienda que active el control de versiones y que tenga versiones antiguas archivadas como parte de una política de administración de la vida útil de los objetos.

En el caso de otros servicios de almacenamiento, por lo general, lograr la durabilidad implica obtener copias de seguridad de datos. Para los discos, significa crear instantáneas, por lo que se deben programar este tipo de trabajos. En Cloud SQL, puede crear una copia de seguridad en cualquier momento (a pedido). Esto podría ser útil si piensa realizar una operación riesgosa en su base de datos o si necesita una copia de seguridad y no desea esperar al período asignado. Google Cloud también proporciona copias de seguridad automáticas, recuperación de un momento determinado y, de forma opcional, un servidor de conmutación por error. Puede crear copias de seguridad a pedido para cualquier instancia, sin importar si tienen habilitadas las copias de seguridad automáticas. Para mejorar la durabilidad, las también se deben ejecutar copias de seguridad de bases de datos SQL.

Spanner y Firestore ofrecen replicación automática, y usted debe ejecutar trabajos de exportación en que los datos se exporten a Cloud Storage.

La cantidad de datos y de operaciones de lectura y escritura es importante cuando se selecciona un servicio de almacenamiento de datos

Algunos servicios escalan horizontalmente agregando nodos.

- Bigtable
- Spanner

Otros servicios escalan verticalmente agrandando las máquinas.

- Cloud SQL
- Memorystore

Y otros escalan automáticamente sin límites.

- Cloud Storage
- BigQuery
- Firestore

Cuando se selecciona un servicio de almacenamiento de datos, es importante conocer la cantidad de datos y de operaciones de lectura y escritura. Algunos servicios escalan horizontalmente agregando nodos, como Bigtable y Spanner, lo que contrasta con Cloud SQL y Memorystore, que escalan máquinas de forma vertical. Otros servicios escalan automáticamente sin límites; por ejemplo, Cloud Storage, BigQuery y Firestore.

¿Necesita coherencia sólida?

Las bases de datos con coherencia sólida actualizan todas las copias de datos dentro de una transacción.

Este tipo de coherencia garantiza que todos obtengan la copia más reciente de los datos sobre operaciones de lectura.

- Storage
- Cloud SQL
- Spanner
- Firestore
- Cloud Bigtable (sin replicación de instancias)

Las bases de datos de coherencia eventual actualizan una copia de datos y el resto de forma asíncrona.

Pueden manejar un gran volumen de operaciones de escritura.

- Cloud Bigtable (comportamiento predeterminado)
- Réplicas de Memorystore

La coherencia sólida es otra característica importante que considerar cuando diseñe soluciones de datos. Una base de datos con coherencia sólida actualizará todas las copias de datos dentro de una transacción y garantizará que todos los usuarios obtengan la copia más reciente de los datos encargados sobre lecturas. Los servicios de Google Cloud que proporcionan coherencia sólida incluyen Cloud Storage, Cloud SQL, Spanner y Firestore. Si una instancia no usa la replicación, Cloud Bigtable proporciona coherencia sólida, ya que todas las operaciones de lectura y escritura se envían al mismo clúster.

Las bases de datos de coherencia eventual suelen tener varias copias de los mismos datos para ofrecer rendimiento y escalabilidad. Estas admiten manejar grandes volúmenes de operaciones de escritura. Su operación consiste en actualizar una copia de datos de forma síncrona y todas las copias de forma asíncrona, lo que significa que no se garantiza que todos los lectores puedan ver los mismos valores de un determinado punto en el tiempo. Con el paso del tiempo, los datos se harán más coherentes, pero no será un proceso inmediato. Bigtable y Memorystore son ejemplos de servicios de datos de Google Cloud que tienen coherencia eventual.

La replicación en Cloud Bigtable tiene coherencia eventual de forma predeterminada. Esto significa que, cuando escriba un cambio en un clúster, con el tiempo podrá leerlo en otros clústeres de la instancia, pero solo después de que el cambio se replique entre los clústeres.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Cuando diseñe una solución de almacenamiento de datos, es importante calcular el costo total por GB para determinar las consecuencias financieras de su decisión.

- Bigtable y Spanner se diseñaron para conjuntos de datos grandes y no son muy rentables para conjuntos de datos pequeños.
- Firestore es más económico por GB almacenado, pero se debe considerar el costo de las operaciones de lectura y escritura.
- Cloud Storage no es tan costoso, pero solo es adecuado para ciertos tipos de datos.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros y, además, genera costos por cada consulta.

Como puede ver, no es fácil elegir la solución de almacenamiento correcta. Debe basarse en el tipo y el tamaño de los datos, y los patrones de lectura o escritura.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Cuando diseñe una solución de almacenamiento de datos, es importante calcular el costo total por GB para determinar las consecuencias financieras de su decisión.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Bigtable y Spanner se diseñaron para conjuntos de datos grandes y no son muy rentables para conjuntos de datos pequeños.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Firestore es más económico por GB almacenado, pero se debe considerar el costo de las operaciones de lectura y escritura.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Cloud Storage no es tan costoso, pero solo es adecuado para ciertos tipos de datos.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros y, además, genera costos por cada consulta.

Calcule el costo total por GB cuando elija un servicio de almacenamiento

- Bigtable y Spanner serían demasiado costosos para almacenar cantidades pequeñas de datos.
- Firestore es más económico por GB, pero también debe pagar por las operaciones de lectura y escritura.
- Cloud Storage es relativamente económico, pero no puede ejecutar bases de datos en almacenamiento.
- El almacenamiento de BigQuery es relativamente económico, pero no proporciona acceso rápido a los registros, y debe pagar por ejecutar consultas.

Debe elegir las soluciones de almacenamiento correctas para cada servicio según sus requisitos.

Como puede ver, no es fácil elegir la solución de almacenamiento correcta. Debe basarse en el tipo y el tamaño de los datos, y los patrones de lectura o escritura.

Actividad 6: Defina las características del almacenamiento

Consulte el Cuaderno de ejercicios de Design and Process.

- Determine las características del almacenamiento para cada uno de sus servicios de casos de éxito.



Ahora definirá las características del almacenamiento para cada uno de sus servicios de casos de éxito.

En una arquitectura de microservicios, no hay una gran base de datos para almacenar todos los datos de toda la aplicación. Cada servicio debe mantener sus propios datos. Las características del almacenamiento incluyen si los datos son estructurados o no estructurados, y relacionales o NoSQL.

Pregúntese si necesita coherencia sólida o si la eventual resulta suficiente. También piense en cuántos datos tiene y si necesita operaciones de lectura y escritura, o solo de lectura.

Actividad 6: Defina las características del almacenamiento

Consulte el Cuaderno de ejercicios de Design and Process.

- Determine las características del almacenamiento para cada uno de sus servicios de casos de éxito.



Ahora definirá las características del almacenamiento para cada uno de sus servicios de casos de éxito.

Actividad 6: Defina las características del almacenamiento

Consulte el Cuaderno de ejercicios de Design and Process.

- Determine las características del almacenamiento para cada uno de sus servicios de casos de éxito.



En una arquitectura de microservicios, no hay una gran base de datos para almacenar todos los datos de toda la aplicación. Cada servicio debe mantener sus propios datos. Las características del almacenamiento incluyen si los datos son estructurados o no estructurados, y relacionales o NoSQL.

Pregúntese si necesita coherencia sólida o si la eventual resulta suficiente. También piense en cuántos datos tiene y si necesita operaciones de lectura y escritura, o solo de lectura.

Servicio	Estructurado o no estructurado	SQL o NoSQL	Coherencia sólida o eventual	Cantidad de datos (MB, GB, TB, PB o EB)	Solo lectura o lectura y escritura
<i>Cuenta</i>	<i>Estructurado</i>	<i>SQL</i>	<i>Sólida</i>	<i>GB</i>	<i>Lectura/escritura</i>

Esta es una tabla de ejemplo de un servicio de cuentas que especifica todos los requisitos comerciales y técnicos que acabo de mencionar.

Consulte la actividad 6 en su cuaderno de ejercicios de diseño a fin de completar una tabla similar para su servicio.

Revise la actividad 6: Defina las características del almacenamiento

- Determine las características del almacenamiento para cada uno de sus servicios de casos de éxito.



En esta actividad se le solicitó definir las características del almacenamiento para el caso de éxito que está diseñando. Estas características lo ayudarán a elegir los servicios de almacenamiento de Google Cloud más adecuados para su aplicación.

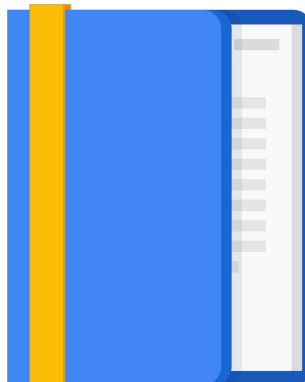
Servicio	Estructurado o no estructurado	SQL o NoSQL	Coherencia sólida o eventual	Cantidad de datos (MB, GB, TB, PB o EB)	Solo lectura o lectura y escritura
Inventario	Estructurado	NoSQL	Sólida	GB	Lectura/escritura
Cargas de inventario	No estructurado	N/A	N/A	GB	Solo lectura
Pedidos	Estructurado	SQL	Sólida	TB	Lectura/escritura
Analítica	Estructurado	SQL	Eventual	TB	Solo lectura

Este es un ejemplo de TurisClic, nuestro portal de viajes en línea. Nos enfocamos en el inventario, las cargas de inventario, los pedidos y los servicios de estadísticas. Como puede ver, cada uno de estos servicios tiene diferentes requisitos que pueden dar como resultado elegir diferentes servicios de Google Cloud.

Temario

Características clave del
almacenamiento

Elija soluciones de
almacenamiento y datos de
Google Cloud



Ahora que documentó las características de los datos de sus servicios, hablemos sobre cómo elegir soluciones de almacenamiento y datos de Google Cloud.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Como se muestra en esta tabla, la cartera de almacenamiento y bases de datos de Google Cloud abarca datos relacionales, NoSQL, de objetos, almacén de datos y almacenes en memoria. Analicemos cada servicio de izquierda a derecha.

Cloud SQL es un almacén de datos de esquema fijo con un límite de almacenamiento de 30 terabytes. Se ofrece con MySQL, PostgreSQL y SQL Server. Estos servicios son adecuados para las aplicaciones web como comercio electrónico o CMS.

Cloud Spanner también es una base de datos relacional y de esquema fijo, pero escala de forma infinita y puede ser regional o multirregional. Los casos de uso de ejemplo incluyen bases de datos relacionales escalables de más de 30 GB con alta disponibilidad y también accesibilidad global, como administración de cadenas de suministro y fabricación.

Los almacenes de datos NoSQL de Google Cloud no tienen esquema.

Firestore es un almacén de datos de documentos completamente administrado con un tamaño máximo de documento de 1 MB. Es útil para datos jerárquicos; por ejemplo, un estado de juego de perfiles de usuario.

Cloud Bigtable también es un almacén de datos NoSQL que escala de forma infinita. Es recomendable para eventos de lectura y escritura pesados, y casos de uso como servicios financieros, Internet de las cosas y transmisiones de publicidad digital.

Para el almacenamiento de objetos, Google Cloud ofrece Cloud Storage. Cloud Storage no tiene esquema y está completamente administrado con escala infinita. Almacena datos de objetos binarios y, por lo tanto, es recomendable para almacenar imágenes, entregar multimedia y crear copias de seguridad.

BigQuery proporciona almacenamiento de datos. Este almacenamiento usa un esquema fijo y admite análisis de SQL completamente administrados de los datos almacenados. Es una excelente opción para realizar análisis y crear paneles de inteligencia empresarial.

En el caso del almacenamiento en memoria, Memorystore proporciona una base de datos administrada sin esquema de Redis. Esta es ideal para almacenar datos en caché de aplicaciones web y dispositivos móviles, y también para proporcionar acceso rápido al estado de arquitecturas de microservicios.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Como se muestra en esta tabla, la cartera de almacenamiento y bases de datos de Google Cloud abarca datos relacionales, NoSQL, de objetos, almacén de datos y almacenes en memoria. Analicemos cada servicio de izquierda a derecha.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Cloud SQL es un almacén de datos de esquema fijo con un límite de almacenamiento de 30 terabytes. Se ofrece con MySQL, PostgreSQL y SQL Server. Estos servicios son adecuados para las aplicaciones web como comercio electrónico o CMS.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Cloud Spanner también es una base de datos relacional y de esquema fijo, pero escala de forma infinita y puede ser regional o multirregional. Los casos de uso de ejemplo incluyen bases de datos relacionales escalables de más de 30 GB con alta disponibilidad y también accesibilidad global, como administración de cadenas de suministro y fabricación.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Los almacenes de datos NoSQL de Google Cloud no tienen esquema.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Firestore es un almacén de datos de documentos completamente administrado con un tamaño máximo de documento de 1 MB. Es útil para datos jerárquicos; por ejemplo, un estado de juego de perfiles de usuario.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTTP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Cloud Bigtable también es un almacén de datos NoSQL que escala de forma infinita. Es recomendable para eventos de lectura y escritura pesados, y casos de uso como servicios financieros, Internet de las cosas y transmisiones de publicidad digital.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTTP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

Para el almacenamiento de objetos, Google Cloud ofrece Cloud Storage. Cloud Storage no tiene esquema y está completamente administrado con escala infinita. Almacena datos de objetos binarios y, por lo tanto, es recomendable para almacenar imágenes, entregar multimedia y crear copias de seguridad.

Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

BigQuery proporciona almacenamiento de datos. Este almacenamiento usa un esquema fijo y admite análisis de SQL completamente administrados de los datos almacenados. Es una excelente opción para realizar análisis y crear paneles de inteligencia empresarial.

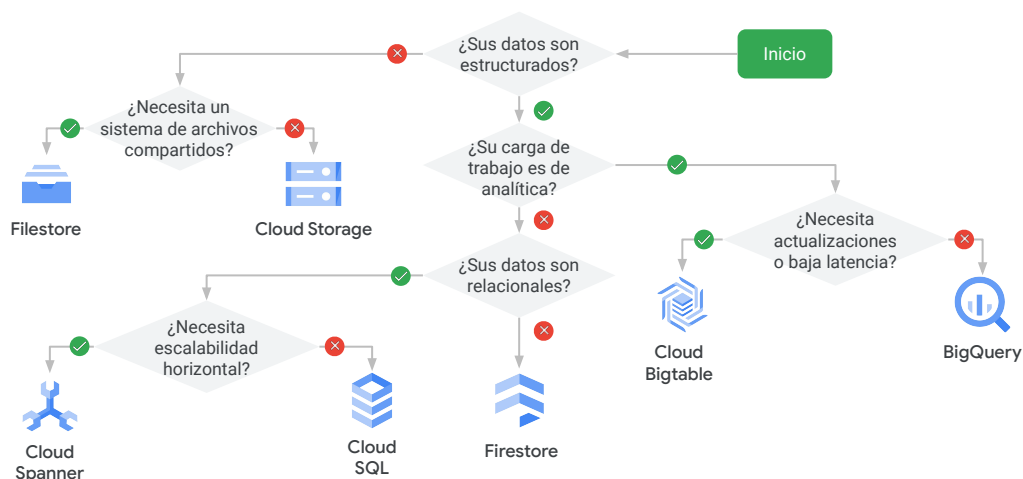
Cartera de almacenamiento y bases de datos de Google Cloud

Relacional		NoSQL		Objeto	Almacén	En la memoria
 Cloud SQL	 Cloud Spanner	 Firestore	 Cloud Bigtable	 Cloud Storage	 BigQuery	 Memorystore
Ideal para Frameworks web	Ideal para RDBMS + escala, alta disponibilidad, HTAP	Ideal para Jerarquías, dispositivos móviles y la Web	Ideal para Eventos y operaciones de lectura y escritura intensivas	Ideal para Datos de objetos binarios	Ideal para Almacenamiento de datos empresariales	Ideal para Almacenar datos en caché de aplicaciones web y para dispositivos móviles
Por ejemplo: CMS y comercio electrónico	Por ejemplo: Metadatos de usuario, tecnología publicitaria, financiera y de marketing	Por ejemplo: Perfiles de usuario, estado de juegos	Por ejemplo: AdTech, IoT y finanzas	Por ejemplo: Imágenes, entrega de contenido multimedia, copias de seguridad	Por ejemplo: Estadísticas y paneles	Por ejemplo: Estado de juegos y sesiones de usuarios
Escala hasta 30 TB MySQL PostgreSQL SQL Server	Escala de forma infinita Regional o multirregional	Completamente administrado Base de datos de documentos	Escala de forma infinita NoSQL de columna amplia	Completamente administrado Escalable de forma infinita	Completamente administrado Análisis de SQL	Bases de datos de Redis administradas
Esquema fijo	Esquema fijo	Sin esquemas	Sin esquemas	Sin esquemas	Esquema fijo	Sin esquemas

Google Cloud

En el caso del almacenamiento en memoria, Memorystore proporciona una base de datos administrada sin esquema de Redis. Esta es ideal para almacenar datos en caché de aplicaciones web y dispositivos móviles, y también para proporcionar acceso rápido al estado de arquitecturas de microservicios.

Tabla de decisiones de almacenamiento y base de datos



Google Cloud

Resumamos los servicios de este módulo con una tabla de decisiones.

En primer lugar, pregúntese lo siguiente: ¿Sus datos son estructurados? y ¿necesitará acceder a ellos con un formato de datos estructurados?

Si la respuesta es no en ambos casos, pregúntese si necesita un sistema de archivos compartidos. Si es así, elija Filestore. De lo contrario, elija Cloud Storage.

Si sus datos son estructurados y necesita acceder a ellos de esta forma, pregúntese lo siguiente: ¿mi carga de trabajo se enfoca en la analítica? Si es así, le recomendamos elegir Cloud Bigtable o BigQuery, según sus necesidades de latencia y actualización. De lo contrario, verifique si sus datos son relacionales.

Si no lo son, elija Firestore. Si son relacionales, le recomendamos elegir Cloud SQL o Cloud Spanner, según su necesidad de 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 distintos servicios, consulte los vínculos sobre opciones de almacenamiento y bases de datos en los materiales del curso.

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

<https://cloud.google.com/products/databases/>

Tabla de decisiones de almacenamiento y base de datos

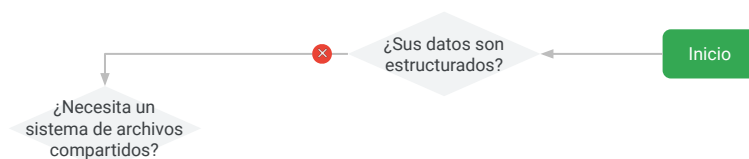
Resumamos los servicios de este módulo con una tabla de decisiones.

Tabla de decisiones de almacenamiento y base de datos



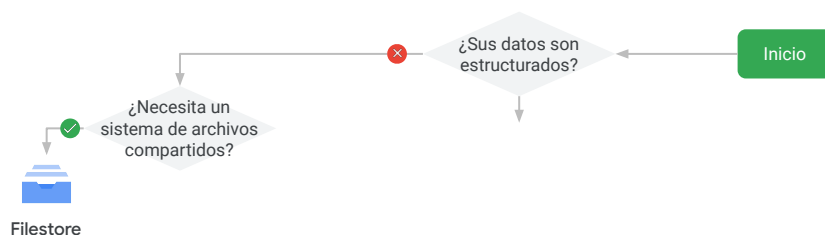
En primer lugar, pregúntese lo siguiente: ¿Sus datos son estructurados? y ¿necesitará acceder a ellos con un formato de datos estructurados?

Tabla de decisiones de almacenamiento y base de datos



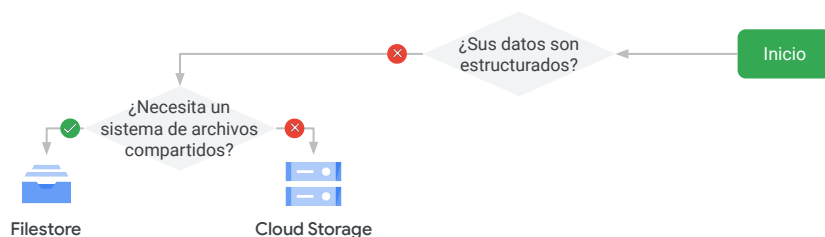
Si la respuesta es no en ambos casos, pregúntese si necesita un sistema de archivos compartidos.

Tabla de decisiones de almacenamiento y base de datos



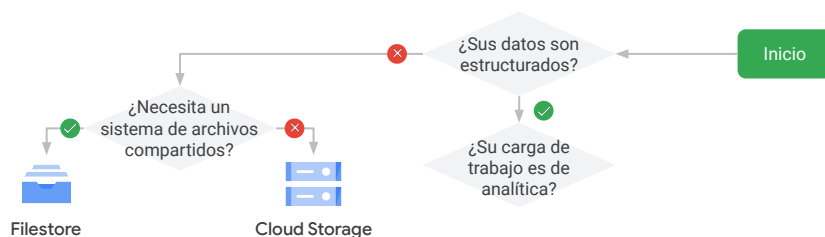
Si es así, elija Filestore.

Tabla de decisiones de almacenamiento y base de datos



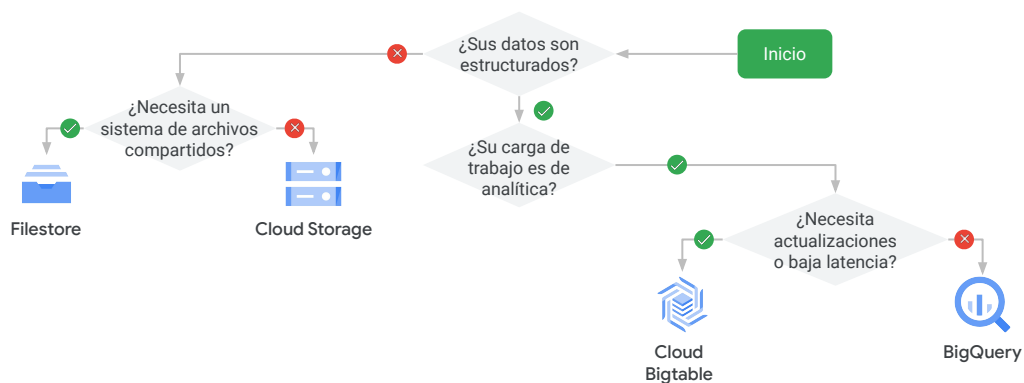
De lo contrario, elija Cloud Storage.

Tabla de decisiones de almacenamiento y base de datos



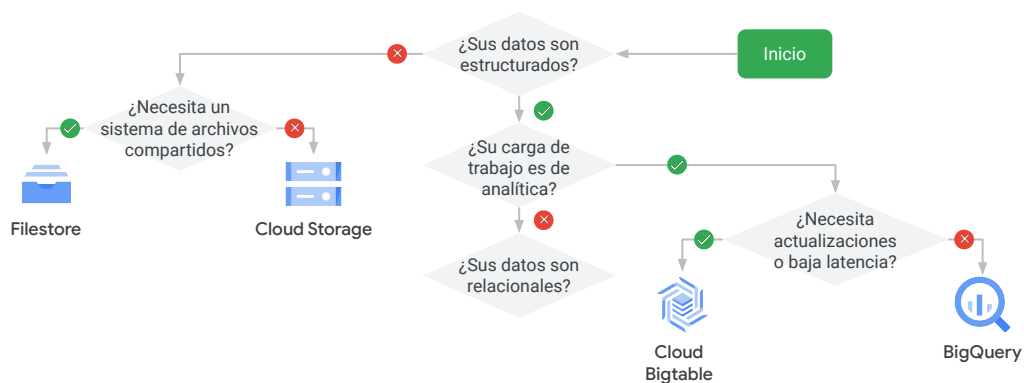
Si sus datos son estructurados y necesita acceder a ellos de esta forma, pregúntese lo siguiente: ¿mi carga de trabajo se enfoca en la analítica?

Tabla de decisiones de almacenamiento y base de datos



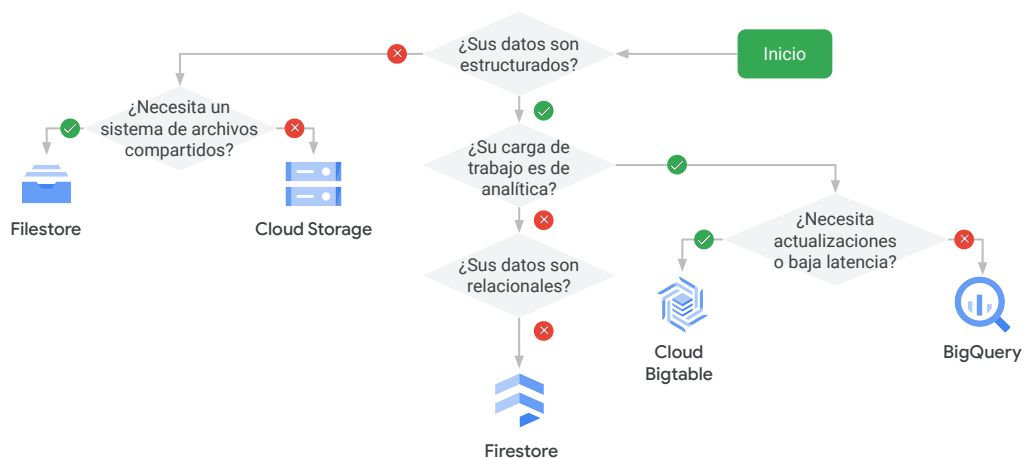
Si es así, le recomendamos elegir Cloud Bigtable o BigQuery, según sus necesidades de latencia y actualización.

Tabla de decisiones de almacenamiento y base de datos



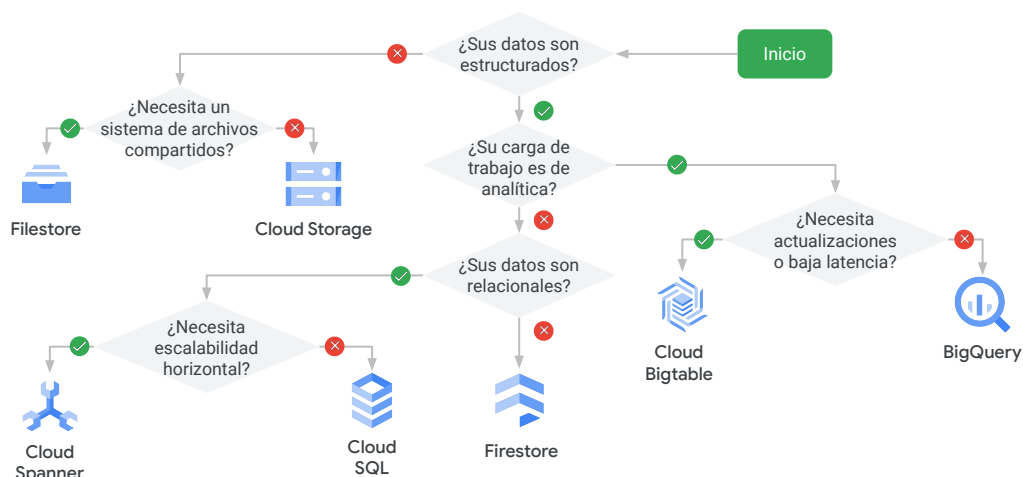
De lo contrario, verifique si sus datos son relacionales.

Tabla de decisiones de almacenamiento y base de datos



Si no lo son, elija Firestore.

Tabla de decisiones de almacenamiento y base de datos



Google Cloud

Si son relacionales, le recomendamos elegir Cloud SQL o Cloud Spanner, según su necesidad de 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 distintos servicios, consulte los vínculos sobre opciones de almacenamiento y bases de datos en los materiales del curso.

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

<https://cloud.google.com/products/databases/>

Transferir datos a Google Cloud puede ser un desafío

	1 Mbps	10 Mbps	100 Mbps	1 Gbps	10 Gbps	100 Gbps
1 GB	3 h	18 min	2 min	11 s	1 s	0.1 s
10 GB	30 h	3 h	18 min	2 min	11 s	1 s
100 GB	12 días	30 h	3 h	18 min	2 min	11 s
1 TB	124 días	12 días	30 h	3 h	18 min	2 min
10 TB	3 años	124 días	12 días	30 h	3 h	18 min
100 TB	34 años	3 años	124 días	12 días	30 h	3 h
1 PB	340 años	34 años	3 años	124 días	12 días	30 h
10 PB	3,404 años	340 años	34 años	3 años	124 días	12 días
100 PB	34,048 años	3,404 años	340 años	34 años	3 años	124 días

Google Cloud

También le recomendamos considerar cómo transferir datos en Google Cloud. Debe tener presentes varios factores, como el costo, el tiempo, las opciones de transferencia en línea y sin conexión, y la seguridad.

Si bien transferir a Cloud Storage es gratuito, habrá costos relacionados con el almacenamiento de datos y quizá incluso costos de dispositivos, si se usa un dispositivo de transferencia, o costos de salida si se realizan transferencias desde otro proveedor de servicios en la nube.

Si tiene grandes conjuntos de datos, el tiempo requerido para transferirlos en una red podría resultar poco realista. Incluso si es realista, los efectos en la infraestructura de su organización podrían ser perjudiciales mientras se realiza la transferencia.

En esta tabla, se muestra el desafío que involucra trasladar grandes conjuntos de datos. Por ejemplo, si tiene 1 TB de datos que debe transferir con una conexión de 100 Mbps, la transferencia tardará aproximadamente 12 días. Las celdas codificadas por color destacan cronogramas poco realistas que requieren soluciones alternativas.

Veamos las opciones de transferencia de datos en línea y sin conexión.

Para cargas de datos más pequeñas o programadas, use el Servicio de transferencia de Cloud Storage

Importe datos en línea a Cloud Storage

- Amazon S3
- Ubicación HTTP/HTTPS
- Transferencia de datos entre buckets de Cloud Storage

Google Cloud

Si tiene cargas de datos más pequeñas o programadas, use el Servicio de transferencia de Cloud Storage, que le permite hacer lo siguiente:

- Mover datos al bucket de Cloud Storage o crear copias de seguridad de datos en él desde otros proveedores de almacenamiento en la nube, como Amazon S3, su almacenamiento local o cualquier ubicación HTTP/HTTPS
- Mover datos de un bucket de Cloud Storage a otro, de modo que estén disponibles para diferentes grupos de usuarios o aplicaciones
- Mover los datos periódicamente como parte de una canalización de procesamiento de datos o de un flujo de trabajo analítico

Para cargas de datos más pequeñas o programadas, use el Servicio de transferencia de Cloud Storage

Importe datos en línea a Cloud Storage

- Amazon S3
- Ubicación HTTP/HTTPS
- Transferencia de datos entre buckets de Cloud Storage

Trabajos programados

- Únicos o recurrentes, la importación se realiza a una hora programada del día
- Opciones para borrar objetos después de la transferencia o que no se encuentran en la fuente
- Filtro por nombre del archivo o fecha de creación

Google Cloud

El Servicio de transferencia de almacenamiento proporciona opciones que facilitan la transferencia y sincronización de datos. Por ejemplo, puede hacer lo siguiente:

- Programar operaciones de transferencia únicas o recurrentes
- Borrar objetos existentes en el bucket de destino si no tienen un objeto que coincida en la fuente
- Borrar objetos de la fuente de datos después de transferirlos
- Programar la sincronización periódica de una fuente de datos con un receptor de datos mediante filtros avanzados basados en fechas de creación de archivos, filtros por nombre de archivo y las horas del día en que prefiere importar datos

Use el Servicio de transferencia de almacenamiento para subir datos locales a gran escala desde su centro de datos

- Instale agentes locales en sus servidores
- El agente se ejecuta en un contenedor de Docker
- Configure una conexión a Google Cloud
- Requiere un mínimo de 300 Mbps de ancho de banda
- Escala a miles de millones de archivos y cientos de TB
- Seguro
- Retiro automático
- Con registro
- Fácil de supervisar a través de Cloud Console

Google Cloud

Use el Servicio de transferencia de almacenamiento para subir datos locales a gran escala desde su centro de datos

El Servicio de transferencia de almacenamiento para datos locales permite transferir datos en línea a gran escala del almacenamiento local a Cloud Storage. En este servicio, se integran la validación de datos, la encriptación, los reintentos de error y la tolerancia a errores. El software local se instala en sus servidores, es decir, el agente está incluido como un contenedor de Docker, y se configura una conexión a Google Cloud. Los directorios que se transferirán a Cloud Storage se seleccionan en Cloud Console.

Una vez que comienza la transferencia de datos, el servicio la paralelizará en muchos agentes que admiten el escalamiento a miles de millones de archivos y cientos de TB. A través de Cloud Console, un usuario puede ver los registros de transferencia detallados, así como la creación, la administración y la supervisión de los trabajos de transferencia.

A fin de usar el Servicio de transferencia de almacenamiento para datos locales, se requiere una fuente que cumpla con los requisitos de POSIX y una conexión de red de al menos 300 Mbps. También se necesita un servidor de Linux compatible con Docker que pueda acceder a los datos que se transferirán, con los puertos 80 y 443 abiertos para las conexiones salientes.

El caso de uso es para transferencias de datos locales cuyo tamaño sea superior a

1 TB.

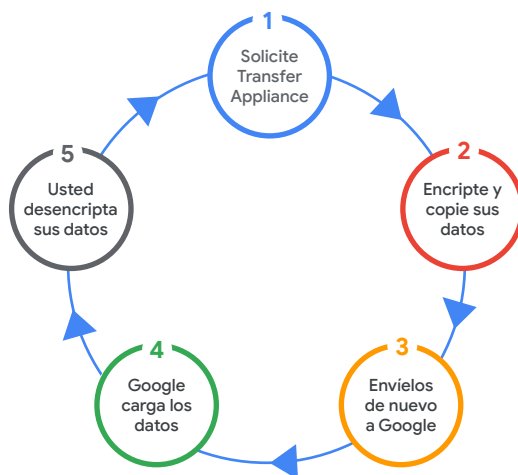
Use Transfer Appliance para grandes cantidades de datos

Dispositivo en bastidores de hasta 1 PB enviado a Google

Use Transfer Appliance si tarda demasiado tiempo en subir sus datos.

Garantice lo siguiente:

- Que controla la clave de encriptación
- Que Google borre el contenido del dispositivo de manera segura después de usarlo



Google Cloud

Para grandes cantidades de datos locales que tardan demasiado en subirse, use Transfer Appliance.

Transfer Appliance es un servidor de almacenamiento en bastidores, seguro y de alta capacidad que puede instalar en su centro de datos. Almacene datos en él y envíelos a una ubicación de transferencia para subir la información a Google. Los datos están protegidos, usted controla la clave de encriptación, y Google borra el contenido del dispositivo después de que se completa la transferencia.

El proceso para usar un dispositivo de transferencia es que usted solicita un dispositivo y este se envía en un caso de manipulación evidente. Luego, los datos se transfieren al dispositivo. Este se envía de nuevo a Google, se cargan los datos en Cloud Storage y se le notifica que está disponible. Google usa sellos de manipulación evidente en casos de envío desde y hacia el sitio de transferencia de datos. Los datos se encriptan de acuerdo con el estándar AES256 cuando se capturan. Cuando se completa la transferencia, se borra el contenido del dispositivo de acuerdo con los estándares NIST-800-88. Usted deberá descriptar los datos cuando desee usarlos.

También hay un servicio de transferencia para BigQuery

Tipo de fuente

Elige la fuente de datos de la siguiente lista.

Fuente*

Google Cloud Storage

Esta es la configuración de Google Cloud Storage. [Más información](#)

Nombre de la conexión

Nombre visible*

Migración de BigQuery

Opciones de programación

☐ Comenzar ahora ☒ Comenzar a la hora definida

Se repite*

Diario

Fecha de inicio y hora de ejecución*

30/11/19, 2:00 a.m.

Configuración de destino

Selecciona el destino de los datos transferidos.

Conjunto de datos de destino*

sales_dataset

Detalles de la fuente de datos

Tabla de destino*

daily_sales

URI de Cloud Storage*

☒ doug-rehnstrom-public/*.*.csv

☒ Borrar archivos de la fuente de datos

Formato del archivo*

CSV

Opciones de transferencia

Todos los formatos

Cantidad de errores permitidos

0

JSON, CSV

☐ Ignorar valores desconocidos

CSV

Delimitador de campos

,

Filas del encabezado que se omitirán

0

Google Cloud

También existe un servicio de transferencia para BigQuery. El Servicio de transferencia de datos de BigQuery automatiza mover datos desde aplicaciones de SaaS hacia BigQuery de manera programada y administrada. El Servicio de transferencia de datos es compatible inicialmente con fuentes de aplicaciones de Google, como Google Ads, Campaign Manager, Google Ad Manager y YouTube. También hay conectores de datos que facilitan la transferencia de datos desde Teradata, Amazon Redshift y Amazon S3 hasta BigQuery.

En las capturas de pantalla de la diapositiva, se muestra que se selecciona un tipo de fuente para la transferencia, se configura una programación y se selecciona un destino de datos. Para la transferencia, también se configuran los formatos de datos.

Actividad 7: Elija servicios de almacenamiento y datos de Google Cloud

Consulte el Cuaderno de ejercicios de Design and Process.

- Elija los servicios de almacenamiento para cada uno de sus servicios de casos de éxito.



Ahora seleccionará productos de almacenamiento para cada uno de sus servicios de casos de éxito.

Use las características de almacenamiento y sus conocimientos de los diversos servicios de Google Cloud que abarcamos en este módulo para elegir qué servicios de almacenamiento serían los más adecuados para cada uno de sus microservicios.

Actividad 7: Elija servicios de almacenamiento y datos de Google Cloud

Consulte el Cuaderno de ejercicios de Design and Process.

- Elija los servicios de almacenamiento para cada uno de sus servicios de casos de éxito.



Ahora seleccionará productos de almacenamiento para cada uno de sus servicios de casos de éxito.








Actividad 7: Elija servicios de almacenamiento y datos de Google Cloud

Consulte el Cuaderno de ejercicios de Design and Process.

- Elija los servicios de almacenamiento para cada uno de sus servicios de casos de éxito.



Use las características de almacenamiento y sus conocimientos de los diversos servicios de Google Cloud que abarcamos en este módulo para elegir qué servicios de almacenamiento serían los más adecuados para cada uno de sus microservicios.

Servicio							
	Persistent Disk	Cloud Storage	Cloud SQL	Firestore	Cloud Bigtable	Cloud Spanner	BigQuery
Servicio de cuentas			X				

Esta es una tabla de ejemplo de un servicio de cuentas que debe usar Cloud SQL, ya que necesitamos una base de datos relacional que tenga coherencia sólida con GB de datos y capacidades de lectura y escritura.








Consulte la actividad 7 en su cuaderno de ejercicios de diseño a fin de completar una tabla similar para los servicios de su aplicación.

Revise la actividad 7: Elija servicios de almacenamiento y datos de Google Cloud

- Elija los servicios de almacenamiento para cada uno de sus servicios de casos de éxito.



En esta actividad, se le solicitó seleccionar los servicios de Google Cloud Storage adecuados para su caso de éxito.

Servicio	 Persistent Disk	 Cloud Storage	 Cloud SQL	 Firestore	 Cloud Bigtable	 Cloud Spanner	 BigQuery
Inventario				X			
Cargas de inventario		X					
Pedidos			X				
Analítica							X

Este es un ejemplo de TurisClic, nuestro portal de viajes en línea:

- Para el servicio de inventario, utilizaremos Cloud Storage en las cargas sin procesar. Los proveedores subirán su inventario como datos en formato JSON almacenados en archivos de texto. Luego, el inventario se importará a una base de datos de Firestore.
- El servicio de pedidos almacenará sus datos en una base de datos relacional que se ejecuta en Cloud SQL.
- El servicio de estadísticas agregará datos de varias fuentes a un almacén de datos para el que usaremos BigQuery.

Repaso

Elija soluciones de
almacenamiento y datos

En este módulo, abordamos los diversos servicios de almacenamiento disponibles en Google Cloud. Estos incluyen Cloud Storage para datos binarios, Cloud SQL y Spanner para bases de datos relacionales, Cloud Bigtable y Firestore para bases de datos NoSQL, y BigQuery para almacenes de datos. También hablamos de diferentes características de almacenamiento y cómo puede usarlas para elegir dónde almacenar sus datos.

Cuestionario

Describe las diferencias entre la coherencia sólida y la eventual.

Describe las diferencias entre la coherencia sólida y la eventual.

Cuestionario

Describa las diferencias entre la coherencia sólida y la eventual.

En una base de datos distribuida, existen varias copias de los datos. Con la coherencia sólida, cuando una copia cambia, todas las copias cambian antes de estar disponibles para su lectura. Con la coherencia eventual, cuando se actualizan los datos, una copia cambia y las demás se actualizan de forma asíncrona.

La coherencia eventual permite realizar operaciones de lectura con más rapidez, pero es posible que los usuarios obtengan datos obsoletos sobre una lectura.

La respuesta anterior abarca los puntos principales. Las bases de datos no relacionales, también conocidas como bases de datos NoSQL, surgieron en los últimos años como una alternativa a las bases de datos relacionales. Para lograr la escalabilidad, estas bases de datos suelen mantener varias copias de los mismos datos distribuidos. Diseñar un sistema que aproveche las bases de datos no relacionales puede ser un desafío para los desarrolladores que están más acostumbrados al modelo relacional, ya que posiblemente no conozcan bien algunas de sus características y prácticas. La coherencia eventual es una de estas características.

La coherencia eventual es una garantía teórica de que, mientras no se realicen actualizaciones nuevas en una entidad, todas sus lecturas mostrarán el valor que se actualizó más recientemente. El sistema de nombres de dominio (DNS) de Internet es un buen ejemplo de un sistema con un modelo de coherencia eventual. Los servidores DNS no siempre reflejan los valores más recientes, sino los que se almacenan en caché y se replican en muchos directorios de Internet. Replicar los valores modificados en todos los clientes y servidores DNS lleva cierto tiempo.

En cambio, las bases de datos relacionales tradicionales se diseñaron en función del concepto de coherencia sólida, que también se denomina coherencia inmediata. Esto significa que todos los observadores de una entidad verán los mismos datos inmediatamente después de una actualización. Esta característica fue una premisa fundamental para muchos desarrolladores que utilizan bases de datos relacionales. Sin embargo, los desarrolladores deben hacer concesiones en cuanto a la

escalabilidad y el rendimiento de la aplicación para poder tener una coherencia sólida. En términos simples, es necesario bloquear los datos durante el período de actualización o el proceso de replicación para garantizar que ningún otro proceso actualice los mismos datos.

En una base de datos distribuida, existen varias copias de los datos. Con la coherencia sólida, cuando una copia cambia, todas las copias cambian antes de estar disponibles para su lectura. Con la coherencia eventual, cuando se actualizan los datos, una copia cambia y las demás se actualizan de forma asíncrona.

La coherencia eventual permite realizar operaciones de lectura con más rapidez, pero es posible que los usuarios obtengan datos obsoletos sobre una lectura.

Cuestionario

Suponga que está creando un sitio web para el que necesita almacenar preferencias de los usuarios, información de los productos y opiniones. No habrá una cantidad enorme de datos. ¿Cuál sería una solución sencilla, rentable y administrada?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Suponga que está creando un sitio web para el que necesita almacenar preferencias de los usuarios, información de los productos y opiniones. No habrá una cantidad enorme de datos. ¿Cuál sería una solución sencilla, rentable y administrada?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Cuestionario

Suponga que está creando un sitio web para el que necesita almacenar preferencias de los usuarios, información de los productos y opiniones. No habrá una cantidad enorme de datos. ¿Cuál sería una solución sencilla, rentable y administrada?

A. Firestore

B. Spanner

C. Cloud SQL

D. BigQuery

- A. Respuesta correcta. Firestore proporciona transacciones ACID con ajuste de escala automático y sincronización en vivo, y está integrado en Google Cloud y Firebase. También cuenta con un nivel gratuito.
- B. Respuesta incorrecta. Una función clave de Spanner es el escalamiento para datos relacionales, y el requisito indica que no habrá una enorme cantidad de datos. Usar Spanner probablemente sería más costoso que Cloud SQL y, en este caso, no necesitamos la escala que ofrece Spanner.
- C. Respuesta incorrecta. Si bien Cloud SQL con sus esquemas podría servir como solución, Firestore ofrece más flexibilidad como almacén de documentos y es una mejor opción para el tipo de datos con el que necesitamos trabajar en este caso (especialmente para las opiniones y preferencias de los usuarios). Además, la escala de Cloud SQL puede ser una limitación para cargas máximas, por ejemplo, en Black Friday, o si la base de clientes aumenta. Por último, el modelo de precios es más costoso que el de Firestore.
- D. Respuesta incorrecta. BigQuery es un almacén de datos que se usa para analizarlos, y su modelo y estructura de precios no cumplen con los requisitos de la aplicación.

La respuesta A es correcta. Firestore proporciona transacciones ACID con ajuste de escala automático y sincronización en vivo, y está integrado en Google Cloud y Firebase. También cuenta con un nivel gratuito.

La respuesta B no es correcta. Una función clave de Spanner es el escalamiento para datos relacionales, y el requisito indica que no habrá una enorme cantidad de datos. Usar Spanner probablemente sería más costoso que Cloud SQL y, en este caso, no necesitamos la escala que ofrece Spanner.

La respuesta C no es correcta. Si bien Cloud SQL con sus esquemas podría servir como solución, Firestore ofrece más flexibilidad como almacén de documentos y es una mejor opción para el tipo de datos con el que necesitamos trabajar en este caso (especialmente para las opiniones y preferencias de los usuarios). Además, la escala de Cloud SQL puede ser una limitación para cargas máximas, por ejemplo, en Black Friday, o si la base de clientes aumenta. Por último, el modelo de precios es más costoso que el de Firestore.

La respuesta D no es correcta. BigQuery es un almacén de datos que se usa para analizarlos, y su modelo y estructura de precios no cumplen con los requisitos de la aplicación.

Cuestionario

Suponga que tiene una empresa global de servicios financieros con usuarios en todo el mundo. Necesita un servicio de bases de datos que pueda proporcionar baja latencia a nivel mundial con coherencia sólida. ¿Qué servicio elegiría?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Suponga que tiene una empresa global de servicios financieros con usuarios en todo el mundo. Necesita un servicio de bases de datos que pueda proporcionar baja latencia a nivel mundial con coherencia sólida. ¿Qué servicio elegiría?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Cuestionario

Suponga que tiene una empresa global de servicios financieros con usuarios en todo el mundo. Necesita un servicio de bases de datos que pueda proporcionar baja latencia a nivel mundial con coherencia sólida. ¿Qué servicio elegiría?

A. Firestore

B. Spanner

C. Cloud SQL

D. BigQuery

- A. Respuesta incorrecta. Firestore proporciona transacciones ACID con ajuste de escala automático, pero es probable que el modelo de datos de documentos no sea el apropiado para servicios financieros.
- B. Respuesta correcta. Una función clave de Spanner es la escala para datos relacionales con coherencia sólida. Además, está distribuido a nivel global a fin de proporcionar baja latencia. La alta disponibilidad y la replicación automática también son funciones robustas para los servicios financieros.
- C. Respuesta incorrecta. Cloud SQL no proporcionará baja latencia a nivel mundial porque es un servicio regional. Además, en el caso de los servicios financieros, la replicación y la disponibilidad deben administrarse, mientras que con Spanner estas opciones están integradas.
- D. Respuesta incorrecta. BigQuery es un almacén de datos usado para análisis de datos, y su modelo y estructura de precios no cumple con los requisitos de la aplicación.

La respuesta B es correcta. Una función clave de Spanner es la escala para datos relacionales con coherencia sólida. Además, está distribuido a nivel global a fin de proporcionar baja latencia. La alta disponibilidad y la replicación automática también son funciones robustas para los servicios financieros.

La respuesta A no es correcta. Firestore proporciona transacciones ACID con ajuste de escala automático, pero es probable que el modelo de datos de documentos no sea el apropiado para servicios financieros.

La respuesta C no es correcta. Cloud SQL no proporcionará baja latencia a nivel mundial porque es un servicio regional. Además, en el caso de los servicios financieros, la replicación y la disponibilidad deben administrarse, mientras que con Spanner estas opciones están integradas.

La respuesta D no es correcta. BigQuery es un almacén de datos usado para análisis de datos, y su modelo y estructura de precios no cumple con los requisitos de la aplicación.

Cuestionario

Suponga que desea analizar las tendencias de ventas. Para lograrlo, quiere combinar la información de su base de datos de Oracle local con los datos de Google Analytics y sus registros del servidor web. ¿Dónde podría almacenar los datos para consultarlos de manera fácil y rentable?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Suponga que desea analizar las tendencias de ventas. Para lograrlo, quiere combinar la información de su base de datos de Oracle local con los datos de Google Analytics y sus registros del servidor web. ¿Dónde podría almacenar los datos para consultarlos de manera fácil y rentable?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Cuestionario

Suponga que desea analizar las tendencias de ventas. Para lograrlo, quiere combinar la información de su base de datos de Oracle local con los datos de Google Analytics y sus registros del servidor web. ¿Dónde podría almacenar los datos para consultarlos de manera fácil y rentable?

- A. Firestore
- B. Spanner
- C. Cloud SQL
- D. BigQuery

Las respuestas A, B y C no son correctas. En la situación que se plantea, es necesario combinar datos de tres fuentes distintas, y al menos una de ellas es de transmisión. Firestore, Spanner y Cloud SQL no proporcionan funciones para transferir datos, por lo que implicaría más trabajo integrar y almacenar datos que en BigQuery.

La respuesta correcta es la D. BigQuery es un almacén de datos que se usa para analizarlos y, por lo tanto, está diseñado para este tipo de caso de uso. Ofrece la infraestructura necesaria para transferir datos desde diversas fuentes, lo que también es un requisito en la situación planteada. Otra ventaja es su modelo de costos, según el cual se paga por el almacenamiento y, luego, solo por las consultas que se ejecutan.

La respuesta correcta es la D. BigQuery es un almacén de datos que se usa para analizarlos y, por lo tanto, está diseñado para este tipo de caso de uso. Ofrece la infraestructura necesaria para transferir datos desde diversas fuentes, lo que también es un requisito en la situación planteada. Otra ventaja es su modelo de costos, según el cual se paga por el almacenamiento y, luego, solo por las consultas que se ejecutan.

Las respuestas A, B y C no son correctas. En la situación que se plantea, es

necesario combinar datos de tres fuentes distintas, y al menos una de ellas es de transmisión. Firestore, Spanner y Cloud SQL no proporcionan funciones para transferir datos, por lo que implicaría más trabajo integrar y almacenar datos que en BigQuery.

Cuestionario

Suponga que actualmente usa Firestore para almacenar información sobre productos, opiniones y sesiones de usuario, pero quiere acelerar el acceso a los datos de forma sencilla y rentable. ¿Cuál sería la mejor opción?

- A. Mover los datos a Spanner
- B. Mover los datos a BigQuery
- C. Mover los datos a Cloud Bigtable
- D. Almacenar los datos en caché con Memorystore

Suponga que actualmente usa Firestore para almacenar información sobre productos, opiniones y sesiones de usuario, pero quiere acelerar el acceso a los datos de forma sencilla y rentable. ¿Cuál sería la mejor opción?

- A. Mover los datos a Spanner
- B. Mover los datos a BigQuery
- C. Mover los datos a Cloud Bigtable
- D. Almacenar los datos en caché con Memorystore

Cuestionario

Suponga que actualmente usa Firestore para almacenar información sobre productos, opiniones y sesiones de usuario, pero quiere acelerar el acceso a los datos de forma sencilla y rentable. ¿Cuál sería la mejor opción?

- A. Mover los datos a Spanner
- B. Mover los datos a BigQuery
- C. Mover los datos a Cloud Bigtable
- D. Almacenar los datos en caché con Memorystore

- A. Respuesta incorrecta. Spanner es para datos relacionales a escala global. El modelo de datos y el de costos no son la mejor opción para este requisito.
- B. Respuesta incorrecta. BigQuery sirve para el análisis de datos, no para el almacenamiento.
- C. Respuesta incorrecta. El requisito básico aquí es la velocidad. Aunque Bigtable ofrece baja latencia, se diseñó para datos a gran escala, y ese no es un requisito en este caso.
- D. Respuesta correcta. Memorystore proporciona la mejor opción si se consideran el modelo de datos, el rendimiento, la escala, el costo y la disponibilidad.

La respuesta correcta es la D. Memorystore proporciona la mejor opción si se consideran el modelo de datos, el rendimiento, la escala, el costo y la disponibilidad.

La respuesta A no es correcta. Spanner es para datos relacionales a escala global. El modelo de datos y el de costos no son la mejor opción para este requisito.

La respuesta B no es correcta. BigQuery sirve para el análisis de datos, no para el almacenamiento.

La respuesta C no es correcta. El requisito básico aquí es la velocidad. Aunque Bigtable ofrece baja latencia, se diseñó para datos a gran escala, y ese no es un requisito en este caso.

Más recursos

Opciones de almacenamiento de Google Cloud

<https://cloud.google.com/products/storage/>

Opciones de datos de Google Cloud

<https://cloud.google.com/products/databases/>

Esta es una lista de recursos útiles para obtener más detalles sobre los temas analizados en esta sección.