

Armazenamento na nuvem

Fundamentos do GCP: infraestrutura principal



Primeiros passos com o Cloud Storage e o Cloud SQL



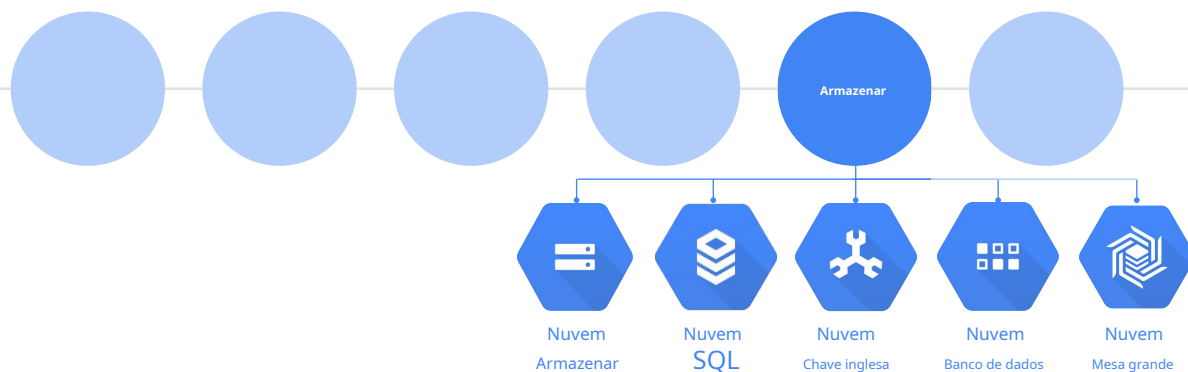
Google Cloud

Última modificação 2018-08-12

© 2017 Google Inc. Todos os direitos reservados. Google e o logotipo da Google são marcas registradas da Google Inc. Todos os outros nomes de empresas e produtos podem ser marcas registradas das respectivas empresas de seus países respectivos.

© 2017 Google Inc. Todos os direitos reservados. Google e o logotipo da Google são marcas registradas da Google Inc. Todos os outros nomes de empresas e produtos podem ser marcas registradas das respectivas empresas de seus países respectivos.

Google Cloud Platform



O Google Cloud Platform tem muitas opções de armazenamento que atendem a quase todos os casos de uso do cliente. Neste módulo, voltamos nossa atenção para as principais opções de armazenamento: Google Cloud Storage, Google Cloud SQL, Google Cloud Spanner, Cloud Datastore e Google Cloud Bigtable.

Agenda

Armazenamento na nuvem

Cloud Bigtable

Cloud SQL e Cloud Spanner

Armazenamento de dados em nuvem

Comparando opções de armazenamento

Teste e laboratório

O Cloud Storage é um armazenamento binário de objetos grandes

- Alto desempenho, escala de internet
- Administração simples
 - Não requer gerenciamento de capacidade
- Criptografia de dados em repouso
- Criptografia de dados em trânsito por padrão de Google para endpoint
- Os serviços de importação online e offline são acessível



 Google Cloud

O Google Cloud Storage oferece aos desenvolvedores e organizações de TI armazenamento de objetos durável e altamente disponível. Não avalia taxa mínima; você paga apenas pelo que usar. O provisionamento prévio de capacidade não é necessário.

O que é armazenamento de objetos? Não é o mesmo que armazenamento de arquivos, no qual você gerencia seus dados como uma hierarquia de pastas. Não é o mesmo que armazenamento em bloco, no qual seu sistema operacional gerencia seus dados como pedaços de disco. Em vez disso, armazenamento de objetos significa isso: você diz ao seu armazenamento: "Aqui, mantenha esta sequência arbitrária de bytes", e o armazenamento permite que você resolva isso com uma chave exclusiva. No Google Cloud Storage e em outros sistemas, essas chaves exclusivas estão na forma de URLs, o que significa que o armazenamento de objetos interage bem com tecnologias da Web.

O Google Cloud Storage sempre criptografa seus dados no servidor, antes de gravá-los no disco, sem custo adicional. Os dados que transitam entre o dispositivo de um cliente e o Google são criptografados por padrão usando HTTPS/TLS (Transport Layer Security). Na verdade, o Google foi o primeiro grande provedor de nuvem a habilitar HTTPS/TLS por padrão.

O Google Cloud Storage não é um sistema de arquivos, embora possa ser acessado como um por meio de ferramentas de terceiros, como o Cloud Storage FUSE. Os objetos de armazenamento oferecidos pelo Google Cloud Storage são "imutáveis", o que significa que você não os edita no local, mas cria uma nova versão. Armazenamento em nuvem do Google

O uso principal é sempre que o armazenamento de objetos grandes binários é necessário: conteúdo online, backup e arquivamento, armazenamento de resultados intermediários no processamento de fluxos de trabalho e muito mais.

A importação/exportação de mídia off-line é uma solução de terceiros que permite carregar dados no Google Cloud Storage enviando sua mídia física, como unidades de disco rígido (HDDs), fitas e unidades flash USB, para um provedor de serviços terceirizado que carrega dados em seu nome. A importação/exportação de mídia offline é útil se você estiver limitado a uma conexão de internet lenta, não confiável ou cara.

A importação offline está disponível por meio de provedores terceirizados: <https://cloud.google.com/storage/docs/offline-media-import-export> O serviço de transferência do Cloud Storage permite importar grandes quantidades de dados on-line para o Google Cloud Storage de forma rápida e econômica. Para usar o serviço de transferência do Cloud Storage, configure uma transferência de uma fonte de dados para um coletor de dados. As fontes de dados podem ser um bucket do Amazon Simple Storage Service (Amazon S3), um local HTTP/HTTPS ou outro bucket do Google Cloud Storage. Os coletores de dados são sempre um bucket do Google Cloud Storage.

Exemplos de uso do serviço de transferência do Cloud Storage incluem:

- Fazer backup de dados em um bucket do Google Cloud Storage de outros provedores de armazenamento.
- Mover dados de um bucket do Standard Storage para um bucket do Nearline Storage para reduzir seus custos de armazenamento.

Seus arquivos do Cloud Storage são organizados em buckets

Atributos do bucket	Conteúdo do bucket
Nome globalmente exclusivo	Arquivos (em um namespace plano)
Classe de armazenamento	
Local (região ou multirregião)	
Políticas do IAM ou listas de controle de acesso	Listas de controle de acesso
Configuração de versão de objeto	
Regras de gerenciamento do ciclo de vida do objeto	

Seus arquivos do Cloud Storage são organizados em buckets. Ao criar um bucket: você atribui a ele um nome globalmente exclusivo; você especifica uma localização geográfica onde o bucket e seu conteúdo são armazenados; e você escolhe uma classe de armazenamento padrão. Escolha um local que minimize a latência para seus usuários. Por exemplo, se a maioria de seus usuários estiver na Europa, você provavelmente desejará escolher um local europeu: uma região do GCP na Europa ou a multirregião da UE.

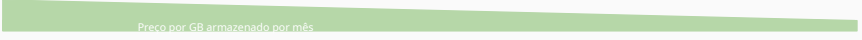

Há várias maneiras de controlar o acesso dos usuários aos seus objetos e buckets. Para a maioria dos propósitos, o Cloud IAM é suficiente. Os papéis são herdados do projeto para o bucket para o objeto. Se precisar de um controle mais preciso, você pode criar listas de controle de acesso (“ACLs”) que oferecem controle mais preciso, as ACLs definem quem tem acesso aos seus buckets e objetos, bem como qual nível de acesso eles têm. Cada ACL consiste em duas informações: Um escopo, que define quem pode executar as ações especificadas (por exemplo, um usuário específico ou grupo de usuários). E uma permissão, que define quais ações podem ser executadas (por exemplo, ler ou escrever).

Lembre-se de que os objetos do Cloud Storage são imutáveis. Você pode ativar o controle de versão de objetos em seus buckets, se desejar. Se você fizer isso, o Cloud Storage mantém um

histórico de modificações, ou seja, substituições ou exclusões, de todos os objetos no bucket. Você pode listar as versões arquivadas de um objeto, restaurar um objeto para um estado mais antigo ou excluir permanentemente uma versão, conforme necessário. Se você não ativar o controle de versão do objeto, o novo sempre substitui o antigo.

O Cloud Storage também oferece políticas de gerenciamento do ciclo de vida. Por exemplo, você pode dizer ao Cloud Storage para excluir objetos com mais de 365 dias ou para excluir objetos criados antes de 1º de janeiro de 2013; ou para manter apenas as 3 versões mais recentes de cada objeto em um bucket que tenha o controle de versão ativado.

Como escolher entre as classes do Cloud Storage

	Multirregional	Regional	Perto da linha	Linha Fria
Destinado a dados que são...	Mais frequente acessado	Acessado com frequência dentro de uma região	Acessou menos de uma vez por mês	Acessou menos de uma vez por ano
SLA de disponibilidade	99,95%	99,90%	99,00%	99,00%
APIs de acesso	APIs consistentes			
Tempo de acesso	Acesso de milissegundos			
<u>Preço de armazenamento</u>				
<u>Preço de recuperação</u>				
Casos de uso	Armazenamento de conteúdo e Entrega	Análise na região, transcodificação	Conteúdo de cauda longa, backups	Arquivamento, recuperação de desastres

O Cloud Storage permite escolher entre quatro tipos diferentes de classes de armazenamento: Regional, Multirregional, Nearline e Coldline. Multirregional e Regional são armazenamento de objetos de alto desempenho, enquanto Nearline e Coldline são armazenamento de backup e arquivamento. Todas as classes de armazenamento são acessadas de maneira análoga usando a API Cloud Storage e todas oferecem tempos de acesso de milissegundos.

Armazenamento Regional permite armazenar seus dados em uma região específica do GCP, us-central1, europe-west1 ou asia-east1. É mais barato que o armazenamento multirregional, mas oferece menos redundância.

Armazenamento multirregional custa um pouco mais, mas é geo-redundante. Isso significa que você escolhe uma localização geográfica ampla, como Estados Unidos, União Europeia ou Ásia, e o Cloud Storage armazena seus dados em pelo menos duas localizações geográficas separadas por pelo menos 160 quilômetros.

O armazenamento multirregional é apropriado para armazenar dados de armazenamento acessados com frequência: conteúdo de sites, cargas de trabalho interativas ou dados que fazem parte de aplicativos móveis e de jogos. Por outro lado, as pessoas usam o armazenamento regional para armazenar dados próximos às máquinas virtuais do Compute Engine ou aos clusters do Kubernetes Engine. Isso oferece melhor desempenho para cálculos com uso intensivo de dados.

Armazenamento Nearline é um serviço de armazenamento de baixo custo e altamente durável para armazenar dados acessados com pouca frequência. Essa classe de armazenamento é uma escolha melhor do que o armazenamento multirregional ou o armazenamento regional em cenários em que você planeja ler ou modificar seus dados em média uma vez por mês ou menos. Por exemplo, se você deseja adicionar arquivos continuamente ao Cloud Storage e planeja acessar esses arquivos uma vez por mês para análise, o Nearline Storage é uma ótima opção.

Armazenamento de linha fria é um serviço de armazenamento altamente durável e de custo muito baixo para arquivamento de dados, backup online e recuperação de desastres. O Coldline Storage é a melhor escolha para dados que você planeja acessar no máximo uma vez por ano, devido à disponibilidade um pouco menor, duração mínima de armazenamento de 90 dias, custos de acesso a dados e custos mais altos por operação. Por exemplo, se você deseja arquivar dados ou ter acesso no caso de um evento de recuperação de desastres.

A disponibilidade dessas classes de armazenamento varia, sendo multirregional com maior disponibilidade de 99,95%, seguida de regional com 99,9% e nearline e coldline com 99,0%.

Quanto ao preço, todas as classes de armazenamento incorrem em um custo por gigabyte de dados armazenados por mês, com multirregional tendo o preço de armazenamento mais alto e a linha fria o preço de armazenamento mais baixo. Taxas de saída e transferência de dados também podem ser aplicadas.

Além dessas cobranças, o armazenamento Nearline também incorre em uma taxa de acesso por gigabyte de dados lidos, e o armazenamento Coldline incorre em uma taxa mais alta por gigabyte de dados lidos.

Há várias maneiras de trazer dados para o Cloud Storage



Transferência online

Cópias autogerenciadas
usando linha de comando
ferramentas ou arrastar e soltar



Serviço de transferência de armazenamento

Transferências em lote programadas e gerenciadas



Dispositivo de transferência **Beta**

Aparelhos em rack para enviar seus dados
com segurança



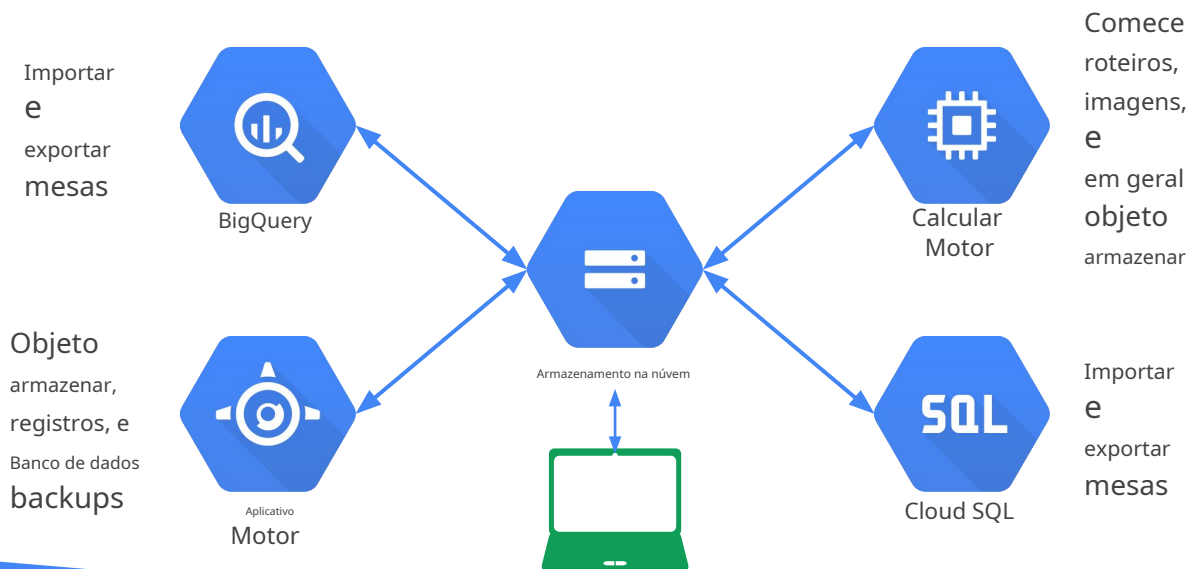
Independentemente da classe de armazenamento escolhida, há várias maneiras de trazer dados para o Cloud Storage.

Muitos clientes simplesmente usam `gsutil`, que é o comando do Cloud Storage do SDK do Cloud. Você também pode mover dados arrastando e soltando no Console do GCP, se usar o navegador Google Chrome. Mas e se você tiver que fazer upload de terabytes ou mesmo petabytes de dados? O Google Cloud Platform oferece o Storage Transfer Service online e o Transfer Appliance offline para ajudar.

O serviço de transferência de armazenamento permite agendar e gerenciar transferências em lote para o Cloud Storage de outro provedor de nuvem, de uma região diferente do Cloud Storage ou de um endpoint HTTP(S).

O Transfer Appliance é um servidor de armazenamento em rack e de alta capacidade que você aluga do Google Cloud. Basta conectá-lo à sua rede, carregá-lo com dados e enviá-lo para uma instalação de upload onde os dados são enviados para o Cloud Storage. O serviço permite que você transfira com segurança até um petabyte de dados em um único dispositivo. A partir desta gravação, ainda é beta e não está disponível em todos os lugares, portanto, verifique o site para obter detalhes.

O Cloud Storage funciona com outros serviços do GCP



Google Cloud

Existem outras maneiras de colocar seus dados no Cloud Storage, pois essa opção de armazenamento é totalmente integrada a muitos produtos e serviços do Google Cloud Platform.

Por exemplo, você pode importar e exportar tabelas de e para o BigQuery, bem como o Cloud SQL.

Você também pode armazenar registros do App Engine, backups do Cloud Datastore e objetos usados por aplicativos do App Engine, como imagens. O Cloud Storage também pode armazenar scripts de inicialização de instâncias, imagens do Compute Engine e objetos usados por aplicativos do Compute Engine.

Resumindo, o Cloud Storage geralmente é o ponto de ingestão de dados que estão sendo movidos para a nuvem e frequentemente é o local de armazenamento de dados de longo prazo.

Agenda

Armazenamento na nuvem

Cloud Bigtable

Cloud SQL e Cloud Spanner

Armazenamento de dados em nuvem

Comparando opções de armazenamento

Integrações com outros serviços

Teste e laboratório

O Cloud Bigtable é NoSQL gerenciado

- NoSQL totalmente gerenciado, banco de dados de colunas amplas
serviço para aplicativos de terabyte
- Integrado
 - Acessado usando a API HBase
 - Compatibilidade nativa com big data,
ecossistemas Hadoop



 Google Cloud

O Cloud Bigtable é o serviço de banco de dados de big data NoSQL do Google. É o mesmo banco de dados que alimenta muitos serviços principais do Google, incluindo Pesquisa, Analytics, Maps e Gmail.

O que significa NoSQL? Aqui está uma comparação informal com bancos de dados relacionais tradicionais criados para oferecer suporte a consultas SQL. Um banco de dados relacional oferece tabelas nas quais cada linha tem o mesmo conjunto de colunas, e o mecanismo de banco de dados impõe essa regra e outras regras que você especifica para cada tabela: o “esquema de banco de dados”. Um esquema rigorosamente aplicado e que não muda com frequência ajuda muitos aplicativos a manter a integridade dos dados. Mas alguns aplicativos exigem uma abordagem muito mais flexível. Para esses aplicativos, nem todas as linhas precisam ter as mesmas colunas e, na verdade, o banco de dados pode ser projetado para tirar proveito disso, preenchendo as linhas de forma esparsa. Isso é parte do que torna um banco de dados NoSQL o que é.

O Cloud Bigtable é oferecido como um serviço totalmente gerenciado, o que significa que você gasta seu tempo desenvolvendo aplicativos valiosos em vez de configurar e ajustar seu banco de dados para desempenho e escalabilidade. Além disso, a própria equipe de operações do Bigtable do Google monitora o serviço para garantir que os problemas sejam resolvidos rapidamente.

O Cloud Bigtable é ideal para aplicativos que precisam de taxa de transferência muito alta e

escalabilidade para dados de chave/valor não estruturados, em que cada valor normalmente não é maior que 10 MB. O Cloud Bigtable também se destaca como mecanismo de armazenamento para operações MapReduce em lote, processamento/análise de fluxo e aplicativos de aprendizado de máquina.

Você pode usar o Cloud Bigtable para armazenar e consultar todos os seguintes tipos de dados:

- Dados de marketing, como históricos de compras e preferências do cliente
- Dados financeiros, como históricos de transações, preços de ações e taxas de câmbio
- Dados da Internet das Coisas, como relatórios de uso de medidores de energia e eletrodomésticos
- Dados de séries temporais, como uso de CPU e memória ao longo do tempo para vários servidores

O Cloud Bigtable é oferecido por meio da mesma API de código aberto que o HBase, o banco de dados nativo do Hadoop. Isso permite a portabilidade de aplicativos entre o HBase e o Bigtable.

Por que escolher o Cloud Bigtable?

- Armazenamento replicado
- Criptografia de dados em trânsito e em repouso
- ACLs baseadas em papéis
- Impulsiona os principais aplicativos, como o Google Analytics e Gmail



 Google Cloud

Os clientes costumam escolher o Bigtable se os dados forem: *Grande*

- Grandes quantidades (> 1 TB) de dados semiestruturados ou estruturados *Rápido*
- Os dados são de alta taxa de transferência ou mudam rapidamente
- *NoSQL*
- Transações, semântica relacional forte não é necessária

E principalmente se for:

Série temporal

- Os dados são séries temporais ou têm ordenação semântica natural

Big data

- Você executa processamento assíncrono em lote ou em tempo real nos dados

Aprendizado de máquina

- Você executa algoritmos de aprendizado de máquina nos dados

O Bigtable foi projetado para lidar com cargas de trabalho massivas com baixa latência consistente e alta taxa de transferência, por isso é uma ótima opção para aplicativos operacionais e analíticos, incluindo IoT, análise de usuários e análise de dados financeiros.

Para mais informações sobre o Cloud Bigtable, consulte

<https://www.google.com/url?q=https://cloudplatform.googleblog.com/2015/05/introducing-Google-Cloud-Bigtable.html>

Padrões de acesso do Bigtable



Como o Cloud Bigtable faz parte do ecossistema do GCP, ele pode interagir com outros serviços do GCP e clientes de terceiros.

Do ponto de vista da API do aplicativo, os dados podem ser lidos e gravados no Cloud Bigtable por meio de uma camada de serviço de dados, como VMs gerenciadas, o servidor HBase REST ou um servidor Java usando o cliente HBase. Normalmente, isso servirá para fornecer dados a aplicativos, painéis e serviços de dados.

Os dados também podem ser transmitidos por meio de uma variedade de estruturas populares de processamento de fluxo, como Cloud Dataflow Streaming, Spark Streaming e Storm.

Se o streaming não for uma opção, os dados também poderão ser lidos e gravados no Cloud Bigtable por meio de processos em lote, como Hadoop MapReduce, Dataflow ou Spark. Muitas vezes, os dados resumidos ou recém-calculados são gravados no Cloud Bigtable ou em um banco de dados downstream.

Agenda

Armazenamento na nuvem

Cloud Bigtable

Cloud SQL e Cloud Spanner

Armazenamento de dados em nuvem

Comparando opções de armazenamento

Teste e laboratório

O Cloud SQL é um RDBMS gerenciado

- Oferece bancos de dados MySQL e PostgreSQL como um serviço
- Replicação automática
- Backups gerenciados
- Escala vertical (leitura e gravação)
- Escala horizontal (leitura)
- Segurança do Google



 Google Cloud

O Cloud SQL é um serviço fácil de usar que oferece bancos de dados relacionais totalmente gerenciados. O Cloud SQL permite que você entregue ao Google as tarefas comuns, mas necessárias e muitas vezes demoradas, como aplicar patches e atualizações, gerenciar backups e configurar replicações, para que você possa se concentrar na criação de ótimos aplicativos.

Cada instância do Cloud SQL inclui um firewall de rede, permitindo que você controle o acesso de rede à sua instância de banco de dados concedendo acesso.

O Cloud SQL é fácil de usar: não requer instalação ou manutenção de software.

Escale facilmente até 64 núcleos de processador e mais de 100 GB de RAM. Escala horizontalmente rapidamente com réplicas de leitura.

Replicação automática

O Google Cloud SQL é compatível com os seguintes cenários de réplica de leitura:

- Instâncias do Cloud SQL replicando de uma instância mestre do Cloud SQL. As réplicas são outras instâncias no mesmo projeto e local da instância mestre. Este recurso está em Beta.
- Instâncias do Cloud SQL replicando de uma instância mestre externa. A instância mestre é externa ao Google Cloud SQL. Por exemplo, ele

- pode estar fora da rede do Google ou em uma instância do Google Compute Engine. Este recurso está em Beta.
 - Instâncias externas do MySQL replicando de uma instância mestre do Cloud SQL
- As réplicas externas estão em ambientes de hospedagem, fora do Cloud SQL.

Backups gerenciados

O Cloud SQL cuida do armazenamento seguro dos dados de backup e facilita a restauração de um backup e a recuperação pontual para um estado específico de uma instância. O Cloud SQL retém até sete backups para cada instância, que estão incluídos no custo da sua instância.

Os dados do cliente do Cloud SQL são criptografados nas redes internas do Google e quando armazenados em tabelas de banco de dados, arquivos temporários e backups.

(As instâncias do MySQL estão disponíveis na primeira geração ou na segunda geração. O Google recomenda o uso de instâncias de segunda geração para a maioria dos casos de uso. As instâncias de primeira geração são recomendadas principalmente quando a compatibilidade com o MySQL 5.5 é necessária. Além disso, as instâncias de primeira geração podem ser econômicas para instâncias de banco de dados usadas com pouca frequência ou de teste/desenvolvimento, devido ao plano de cobrança por uso disponível e à política de ativação ON DEMAND disponível, que faz com que sua instância se desligue automaticamente após 15 minutos de inatividade.)

O Cloud SQL pode ser usado com outros serviços do GCP



O Cloud SQL pode ser usado com o App Engine usando drivers padrão.

Você pode configurar uma instância do Cloud SQL para seguir um aplicativo do App Engine.



As instâncias do Compute Engine podem ser autorizadas a acessar o Cloud SQL instâncias usando um endereço IP externo.

As instâncias do Cloud SQL podem ser configuradas com uma zona preferencial.



O Cloud SQL pode ser usado com aplicativos e clientes externos.

Ferramentas padrão podem ser usadas para administrar bancos de dados.

As réplicas de leitura externas podem ser configuradas.

Outro benefício das instâncias do Cloud SQL é que elas podem ser acessadas por outros serviços do GCP e até mesmo serviços externos. Você pode usar o Cloud SQL com o App Engine usando drivers padrão, como Connector/J para Java ou MySQLdb para Python.

Você pode autorizar instâncias do Compute Engine a acessar instâncias do Cloud SQL e configurar a instância do Cloud SQL para estar na mesma zona que sua máquina virtual.

O Cloud SQL também é compatível com outros aplicativos e ferramentas com os quais você pode estar acostumado, como SQL Workbench, Toad e outros aplicativos externos que usam drivers MySQL padrão.

O Cloud Spanner é um RDBMS escalonável horizontalmente

O Cloud Spanner é compatível com:

- Replicação automática
- Forte consistência global
- Instâncias gerenciadas com alta disponibilidade
- SQL (ANSI 2011 com extensões)



 Google Cloud

O Cloud Spanner oferece suporte a uma consistência forte, incluindo índices secundários altamente consistentes, SQL e instâncias gerenciadas com alta disponibilidade por meio de replicação de dados síncrona e integrada. Testado em batalha pelos próprios aplicativos e serviços de missão crítica do Google, Spanner impulsiona os negócios de US\$ 80 bilhões do Google.

O Cloud Spanner é especialmente adequado para aplicativos que exigem:

- Um SQL RDBMS, com junções e índices secundários
- Alta disponibilidade integrada
- Forte consistência global
- Tamanhos de banco de dados que excedem ~ 2 TB
- Muitos IOPS (dezenas de milhares de leituras/gravações por segundo ou mais)

Para uma visão geral técnica do Cloud Spanner, consulte <https://cloudplatform.googleblog.com/2017/02/inside-Cloud-Spanner-and-the-CAP-Theory.html>.

Agenda

Armazenamento na nuvem

Cloud Bigtable

Cloud SQL e Cloud Spanner

Armazenamento de dados em nuvem

Comparando opções de armazenamento

Teste e laboratório

O Cloud Datastore é um NoSQL horizontalmente escalonável

BD

- [NoSQL](#) projetado para back-ends de aplicativos
- Totalmente gerenciado
 - Usa uma arquitetura distribuída para gerenciar automaticamente o dimensionamento
- Redundância incorporada
- Apoia [ÁCIDO](#) transações



 Google Cloud

O Cloud Datastore é um banco de dados NoSQL altamente escalável para seus aplicativos. Assim como o Cloud Bigtable, não há necessidade de provisionar instâncias de banco de dados. O Cloud Datastore usa uma arquitetura distribuída para gerenciar o escalonamento automaticamente. Suas consultas são dimensionadas com o tamanho do conjunto de resultados, não com o tamanho do conjunto de dados.

O Cloud Datastore é executado em data centers do Google, que usam redundância para minimizar o impacto de pontos de falha. Seu aplicativo ainda poderá usar o Cloud Datastore quando o serviço receber um upgrade planejado.

O tamanho total dos bancos de dados do Cloud Datastore pode chegar a terabytes ou mais.

Armazenamento de dados do Google Cloud: benefícios

- Acesso sem esquema
 - Não há necessidade de pensar na estrutura de dados subjacente
- Ferramentas de desenvolvimento local
- Inclui uma cota diária gratuita
- Acesse de qualquer lugar através de um [Interface RESTful](#)



 Google Cloud

Recursos do Cloud Datastore:

- *Transações atômicas*
O Datastore pode executar um conjunto de operações em que todas são bem-sucedidas ou nenhuma ocorre.
- *Alta disponibilidade de leituras e gravações*
O Datastore é executado em data centers do Google, que usam redundância para minimizar o impacto de pontos de falha.
- *Escalabilidade massiva com alto desempenho*
O Datastore usa uma arquitetura distribuída para gerenciar o dimensionamento automaticamente. O Datastore usa uma combinação de índices e restrições de consulta para que suas consultas sejam dimensionadas com o tamanho do conjunto de resultados, não com o tamanho do conjunto de dados.
- *Armazenamento flexível e consulta de dados*
O armazenamento de dados mapeia naturalmente para linguagens de script e orientadas a objetos e é exposto a aplicativos por meio de vários clientes. Ele também fornece um SQL-like [Linguagem de consulta](#).
- *Equilíbrio de consistência forte e eventual*
O Datastore garante que as pesquisas de entidade e consultas de ancestrais sempre recebam dados fortemente consistentes. Todas as outras consultas são eventualmente consistentes. Os modelos de consistência permitem que seu aplicativo ofereça uma ótima experiência ao usuário enquanto lida com grandes quantidades de dados e usuários.



Criptografia em repouso

O Datastore criptografa automaticamente todos os dados antes de serem gravados no disco e descriptografa automaticamente os dados quando lidos por um usuário autorizado. Para mais informações, veja [Criptografia do lado do servidor](#).



Totalmente gerenciado sem tempo de inatividade planejado

O Google lida com a administração do serviço Datastore para que você possa se concentrar em seu aplicativo. Seu aplicativo ainda poderá usar o Datastore quando o serviço receber um upgrade planejado.

Agenda

Armazenamento na nuvem

Cloud Bigtable

Cloud SQL e Cloud Spanner

Armazenamento de dados em nuvem

[Comparando opções de armazenamento](#)

Integrações com outros serviços

Teste e laboratório

Comparando opções de armazenamento: detalhes técnicos

	Nuvem Banco de dados	Mesa grande	Nuvem Armazenar	Nuvem SQL	Nuvem Chave inglesa	BigQuery
Tipo	NoSQL documento	NoSQL coluna larga	Blobstore	Relacional SQL para OLTP	Relacional SQL para OLTP	Relacional SQL para OLAP
Transações	Sim	Fila única	Não	Sim	Sim	Não
Complexo consultas	Não	Não	Não	Sim	Sim	Sim
Capacidade	Terabytes+	Petabytes+	Petabytes+	Até ~ 10 TB	Petabytes	Petabytes+
O tamanho da unidade	1 MB/entidade	~10 MB/célula ~100 MB/linha	5 TB/objeto	Determinado por motor DB	10.240 MiB/10 MB/linha de linha	



Agora que abordamos as principais opções de armazenamento do GCP, vamos compará-las para ajudar você a escolher o serviço certo para seu aplicativo ou fluxo de trabalho.

Esta tabela se concentra nos diferenciais técnicos dos serviços de armazenamento. Cada linha é uma especificação técnica e cada coluna é um serviço. Deixe-me cobrir cada serviço da esquerda para a direita.

Considere usar o Cloud Datastore se precisar armazenar objetos estruturados ou se precisar de suporte para transações e consultas semelhantes a SQL. Esses serviços de armazenamento fornecem terabytes de capacidade com tamanho máximo de unidade de 1 MB por entidade.

Considere usar o Cloud Bigtable se precisar armazenar uma grande quantidade de objetos estruturados. O Cloud Bigtable não é compatível com consultas SQL nem com transações de várias linhas. Esse serviço de armazenamento fornece petabytes de capacidade com tamanho máximo de unidade de 10 MB por célula e 100 MB por linha.

Considere usar o Cloud Storage se precisar armazenar blobs imutáveis com mais de 10 MB, como imagens ou filmes grandes. Esse serviço de armazenamento fornece petabytes de capacidade com tamanho máximo de unidade de 5 TB por objeto.

Considere usar o Cloud SQL ou o Cloud Spanner se precisar de suporte completo ao SQL para um sistema de processamento de transações on-line. O Cloud SQL fornece até 10.230 GB, dependendo do tipo de máquina, enquanto o Cloud Spanner fornece petabytes. Se o Cloud SQL não atende às suas necessidades porque você precisa de escalabilidade horizontal, não apenas

por meio de réplicas de leitura, considere usar o Cloud Spanner.

Não abordamos o BigQuery neste módulo, pois ele fica no limite entre o armazenamento de dados e o processamento de dados, mas você aprenderá mais sobre ele no módulo "Big Data and Machine Learning in the Cloud". O motivo usual para armazenar dados no BigQuery é usar seus recursos de análise de big data e consulta interativa. Você não gostaria de usar o BigQuery, por exemplo, como loja de apoio para um aplicativo on-line.

Comparando opções de armazenamento: casos de uso

	Nuvem Banco de dados	Nuvem Mesa grande	Nuvem Armazenar	Nuvem SQL	Nuvem Chave inglesa	BigQuery
Tipo	NoSQL documento	NoSQL coluna larga	Blobstore	Relacional SQL para OLTP	Relacional SQL para OLTP	Relacional SQL para OLAP
Melhor para	Semi-estruturado dados de aplicativos, durável dados de valor-chave	Dados "planos", Pesado ler escrever, eventos, dados analíticos	Estruturado e não estruturado binário ou objeto dados	Rede estruturas, existir formulários	Grande escala base de dados aplicativos (> ~2TB)	Interativo consulta, off-line análise
Casos de uso	Começando, App Engine formulários	AdTech, Financeiro e Dados de IoT	Imagens, grandes arquivos de mídia, backups	As credenciais do usuário, Pedidos de clientes	Sempre que alto E/S, global consistência é precisava	Dados armazenagem

Considerar os diferenciais técnicos dos diferentes serviços de armazenamento ajuda algumas pessoas a decidir qual serviço de armazenamento escolher, outras gostam de considerar casos de uso. Deixe-me passar por cada serviço mais uma vez.

O Cloud Datastore é melhor para dados de aplicativos semiestruturados usados em aplicativos do App Engine.

O Bigtable é melhor para dados analíticos com eventos pesados de leitura e gravação, como AdTech, dados financeiros ou IoT.

O Cloud Storage é melhor para dados binários ou de objetos estruturados e não estruturados, como imagens, arquivos de mídia grandes e backups.

O Cloud SQL é melhor para frameworks da Web e aplicativos existentes, como armazenamento de credenciais de usuários e pedidos de clientes.

O Cloud Spanner é melhor para aplicativos de banco de dados de grande escala com mais de 2 TB. Por exemplo, para casos de uso de comércio financeiro e comércio eletrônico.

Como mencionei no início do módulo, dependendo de sua aplicação, você pode usar um ou vários desses serviços para realizar o trabalho.

Questionário

Seu aplicativo transcodifica arquivos de vídeo grandes. Qual serviço de armazenamento você deve considerar primeiro?

Você transmite grandes quantidades de dados de dispositivos com sensores. Qual serviço de armazenamento você deve considerar primeiro?

Respostas do questionário

Seu aplicativo transcodifica arquivos de vídeo grandes. Qual serviço de armazenamento você deve considerar primeiro?

Armazenamento em nuvem do Google

Você transmite grandes quantidades de dados de dispositivos com sensores. Qual serviço de armazenamento você deve considerar primeiro?

Respostas do questionário

Seu aplicativo transcodifica arquivos de vídeo grandes. Qual serviço de armazenamento você deve considerar primeiro?

Armazenamento em nuvem do Google

Você transmite grandes quantidades de dados de dispositivos com sensores. Qual serviço de armazenamento você deve considerar primeiro?

Google Cloud Bigtable

Agenda

Armazenamento em nuvem do Google

Google Cloud Bigtable

Google Cloud SQL e Google Cloud Spanner

Armazenamento de dados do Google Cloud

Comparando opções de armazenamento

Integrações com serviços de armazenamento

Teste e laboratório

Instruções de laboratório

Neste laboratório, você criará um bucket do Google Cloud Storage e colocará uma imagem nele. Você também configurará um aplicativo em execução no Google Compute Engine para usar um banco de dados gerenciado pelo Google Cloud SQL e fazer referência à imagem no bucket do Cloud Storage.

- Crie um bucket do Cloud Storage e coloque uma imagem nele
- Crie uma instância do Cloud SQL e configure-a
- Conecte-se a uma instância do Cloud SQL de um servidor da Web
- Use uma imagem armazenada em um bucket do Cloud Storage em uma página da Web

Mais recursos

Visão geral do armazenamento em nuvem

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

Primeiros passos com o Google Cloud SQL <https://cloud.google.com/sql/docs/quickstart>

Cloud Bigtable

<https://cloud.google.com/stackdriver/docs/>

Cloud Spanner

<https://cloud.google.com/spanner/docs/>

Cloud Datastore

<https://cloud.google.com/datastore/docs/>



© 2017 Google Inc. All rights reserved. Google and the Google logo are trademarks of Google Inc. All other company and product names may be trademarks of the respective companies with which they are associated.

© 2017 Google Inc. All rights reserved. Google and the Google logo are trademarks of Google Inc. All other company and product names may be trademarks of the respective companies with which they are associated.