# Série de Certificação GCP: 3.4 Implantação e implementação de soluções de dados



### Prashanta Paudel

31 de outubro de 2018 · 22 minutos de leitura

Você já deve saber que o Google é uma empresa de dados, e não uma empresa de TI, porque a maior parte de seu produto depende da análise de dados de seus usuários.

Por exemplo, a Pesquisa do Google, o mecanismo de pesquisa número um do mundo, depende das estatísticas dos usuários para exibir resultados de pesquisa relevantes e adicioná-los aos usuários. Então, em suma, ele usa dados de usuários para mostrar adicionar ao mesmo usuário. Tudo o que o Google faz é processar os dados.

Se olharmos para as **soluções de dados da** nuvem **do Google** , podemos dividi-las em

### Bancos de dados

- Cloud SQL
- Cloud Datastore
- Cloud Spanner
- · Cloud Bigtable

Existem outras soluções também, mas não exigimos isso agora.

### Análise de dados

- BigQuery
- Cloud Pub / Sub
- Cloud Dataproc

Existem muito mais soluções, mas não precisamos delas agora.

# Bancos de dados

Como já estudamos antes, existem basicamente dois tipos de bancos de dados

- 1. SQL
- Cloud SQL
- · Cloud Spanner

### 2. NoSQL

- · Cloud Datastore
- Cloud Bigtable

# **Cloud SQL**



Cloud SQL

O Cloud SQL é um serviço de banco de dados totalmente gerenciado que facilita a configuração, manutenção, gerenciamento e administração de seu banco de dados relacional. Suporta Cloud SQL

- 1. PostgreSQL e
- 2. MySQL

Ele oferece alto desempenho, escalabilidade e conveniência para os usuários criarem e implementarem um banco de dados para seus aplicativos. Um banco de dados é um dado estruturado que será hospedado na nuvem do Google, mas pode ser acessado de qualquer lugar.



### Focus on Your Application

Let Google manage your database, so you can focus on your applications. Cloud SQL is perfect for Wordpress sites, e-commerce applications, CRM tools, geospatial applications, and any other application that is compatible with MySQL or PostgreSQL.

### Simple & Fully Managed

Cloud SQL is easy to use. It doesn't require any software installation. It automates all your backups, replication, patches, and updates - while ensuring greater than 99.95% availability, anywhere in the world.





### Performance & Scalability

Cloud SQL delivers high performance and scalability with up to 10TB of storage capacity, 40,000 IOPS, and 416GB of RAM per instance.

### Reliability & Security

Easily configure replication and backups to protect your data. Go further by enabling automatic failover to make your database highly available (HA). Your data is automatically encrypted and Cloud SQL is SSAE 16, ISO 27001, PCI DSS v3.0, and HIPAA compliant.

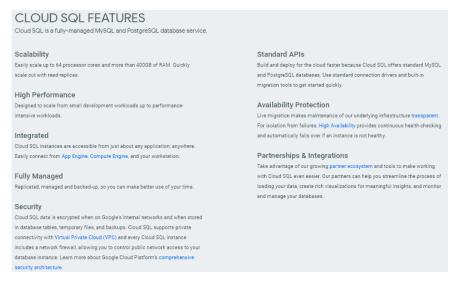




### Discounts Without Lock-in

Cloud SQL offers per-second billing, automatic sustained use discounts, and instance sizes to fit any budget. Database instances are easy to stop and start. There is no up-front commitment, and with sustained use discounts, you'll automatically get discounted prices for databases that run continuously.

Referências: https://cloud.google.com/sql/



Referências: https://cloud.google.com/sql/

O preço do Cloud SQL depende das regiões e zonas do GCP

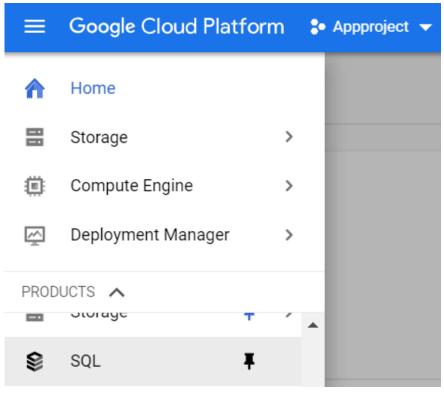
### Casos de uso



Exemplo de projeto 1: conexão MySQL mais simples

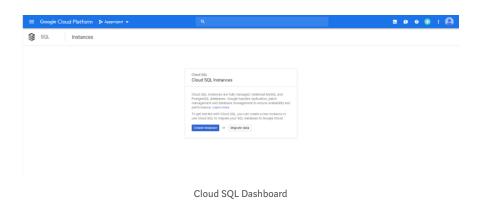
A forma mais simples de conexão SQL é criar uma instância do SQL na nuvem e conectá-la a partir do usuário e do IP autorizados para executar algumas tarefas de banco de dados.

Primeiro de tudo, crie um banco de dados no Cloud SQL.

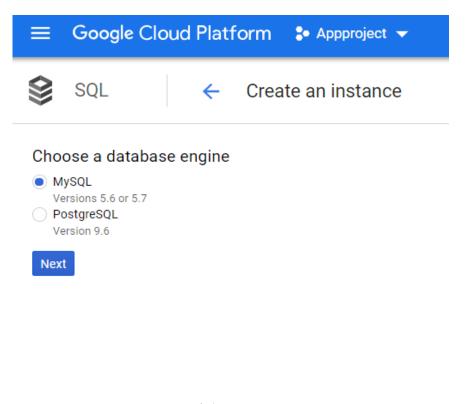


selecione SQL

# Então você verá um painel



Clique em Criar instância



Selecione o MySQL

### Agora crie o banco de dados



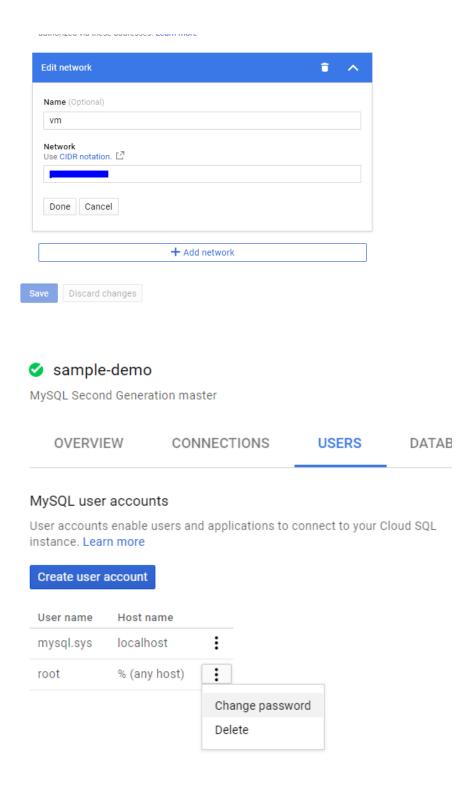
criando um banco de dados

Depois de criar um banco de dados, você deve instalar o Mysql Workbench na máquina cliente para conectar-se ao servidor MySQL. Para isso, faça o download na máquina Linux ou em qualquer outro, se desejar.

```
$ sudo apt-get install mysql-workbench
```

Isso instalará o ambiente de trabalho em sua máquina Linux.

Agora autentique a conexão, permitindo o IP do Linux na nuvem SQL e adicione a senha de root.



Agora vá para a máquina Linux que tem o MySQL workbench e se conecta ao CLoud SQL Server que acabamos de criar.

```
: ~ $ mysql --host = 35.193.199.18 --user = root - senha
```

Bem vindo ao monitor do MariaDB. Comandos terminam com; ou \ g.Seu ID de conexão do MySQL é 57Server version: 5.7.14-google-log (Google) Copyright © 2000, 2017, Oracle, MariaDB Corporation e outros. Digite 'help;' ou '\ h' para ajuda. Digite '\ c' para limpar a instrução de entrada atual.

Agora você está conectado ao Cloud SQL pela CLI.

Viva!

Alguns dos comandos são

```
MySQL [(none)]> show databases
  Database
 information schema
 mysql
 performance schema |
 rows in set (0.04 sec)
MySQL [(none)]> create database test
    -> ;
Query OK, 1 row affected (0.04 sec)
MySQL [(none)]> create table people;
ERROR 1046 (3D000): No database selected
MySQL [(none)]> CREATE TABLE Persons (
           PersonID int,
           LastName varchar(255),
           FirstName varchar(255),
           Address varchar (255),
```

Comandos SQL

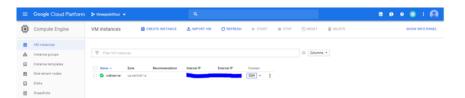
\_\_\_\_

# Eample 2: Mecanismo de computação do DB Localhost + LAMP (No cloud SQL)

### Website do livro de visitas

Primeiro de tudo, vamos criar um servidor web, por favor, siga os passos nos blogs anteriores para configurar um servidor web





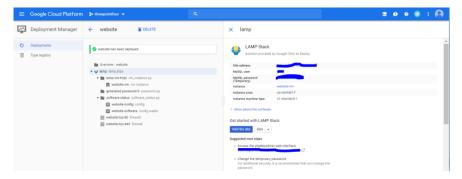
Novo servidor da web criado

# Aqui, criei um servidor web com o sistema operacional Debian e instalei

- Apache2
- MySQL Workbench
- PHP
- phpmyadmin

Em suma, é chamado de servidor LAMP.

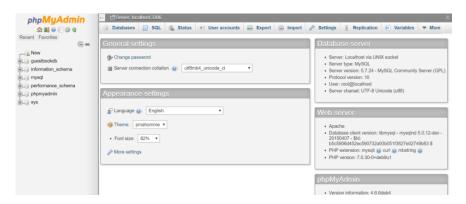
Então, em vez de instalar cada pacote, um por um, implantei um servidor LAMP do Marketplace na nuvem do Google.



Servidor LAMP

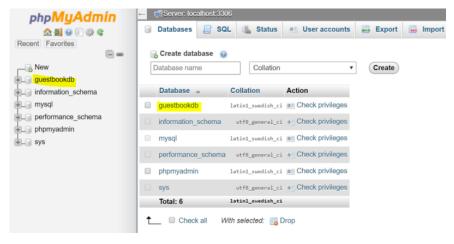
Agora que temos toda a infraestrutura, começaremos a construir o banco de dados primeiro.

### Ir para <a href="http://website/phpmyadmin">http://website/phpmyadmin</a>



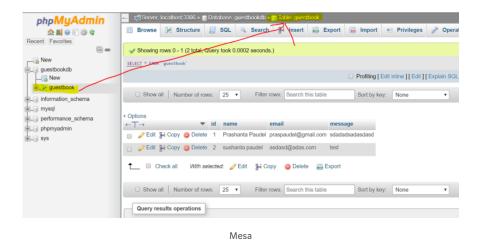
página phpmyadmin

### Agora crie um banco de dados e uma tabela



Crie um banco de dados

### Agora crie mesa



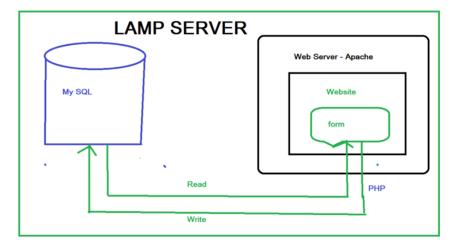
Agora vamos para a web part

```
root@website-vm:/var/www/html# ls
addcomment.php guestbook.php index.html
```

servidor web

Precisamos de três arquivos para serem hospedados no servidor da Web

- 1. index.html para a primeira página do site.
- 2. guestbook.php mostra toda a lista de convidados assinados.
- 3. addcomment.php para adicionar dados do formulário front-end da web ao banco de dados.



Servidor da lâmpada

### index.html

```
chead>
<script type="text/javascript">// <![CDATA[
function Validate()

var x=document.forms["guest"]["email"].value;
var y=document.forms["guest"]["name"].value;
if(y=null || y==")

lett(Fliesse enter your Name! ");
return false;

if(x=null || x==")

lett(Fliesse enter your email address!');
return false;

var atpos=x.indexOf(3");
var atpos=x.indexOf(3");
var dotpos=x.lastIndexOf(");
if (stpos(1 || dotpos=xaspos=2 ||="" dotpos+2="">=x.length)

lett(Not s valid s=nail address");
return false;

sise
return false;

clast (Not s valid s=nail address");
return false;

//atpos=2s// /]]>

<
```

### guestbook.php

## addcomment.php

```
Php
host="localhost"; //Add your SQL Server host here
user="root"; //SQL Username
pass="parific"; //SQL Password
dbname="questbockdb"; //SQL Database Name
con=mysqli_connect($host,$user,$pass,$dbname);
f (mysqli_connect_errno($con))

cho "<hi>>Failed to connect to MySQL: " . mysqli_connect_error() ."</hi>";

name=$ POST('name');
email=$ POST('massil');
message=$ POST('message');
sql='INSERI INTO guestbock(name,email,message) VALUES('$name','$email','$message')";
f (!mysqli_query($con,$sql))
ie('Error: ' . mysqli_error($con));
lse
cho "Values Stored in our Database!";
ysqli_close($con);
```

### Primeira página

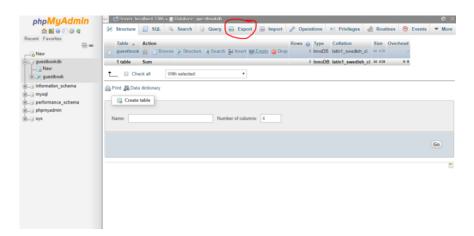
Display ALL			
Name:	Email:	Message:	Sign this in the Book
=====	======	primeira página	=====
====			

# Exemplo 3: Cloud SQL e Compute Engine (baseado em PHP)

# Usando a página da web para inserir e recuperar dados no Cloud SQL

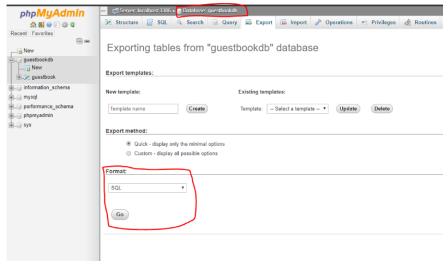
Agora vamos mover o banco de dados localhost para a nuvem usando o PHPMyAdmin

goto phpmyadmin



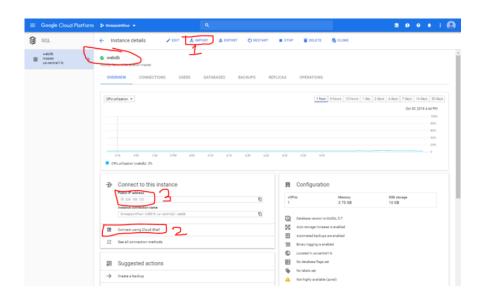
Exportar DB

Selecione o banco de dados e clique em exportar, no formato selecione SQL e salve o arquivo em sua área de trabalho

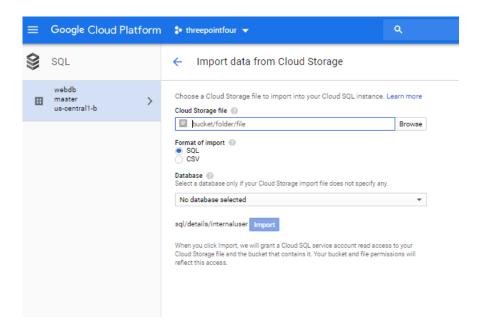


exportar db

### Agora, configure o Cloud SQL no GCP



Depois de iniciar o servidor Cloud SQL, a primeira coisa a fazer é importar o banco de dados que exportamos do phpmyadmin. Dessa forma, obteremos dados de antes.



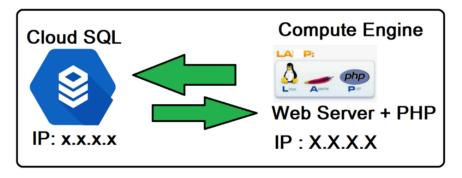
Uma coisa a lembrar é que não podemos importar diretamente, mas temos que fazer o upload para um intervalo no Cloud Storage e, em seguida, selecionar a partir daí.

Após a importação, verifique o banco de dados, selecione o banco de dados e veja as tabelas e os dados.

Agora temos que editar o arquivo addcomment.php e guestbook.php. Basicamente, temos que redirecionar o endereço do servidor de banco de dados para o Cloud SQL e isso é 4 linhas nas linhas de conexão do banco de dados php

```
image: imag
```

Então, agora a arquitetura do nosso sistema é



Cloud SQL + Compute

# CARREGANDO DADOS PARA O SQL Cloud

- 1. Importar / Exportar: no exemplo
- 2. Carregar dados do Cloud Storage: baldes usados na transferência.

# **Cloud Datastore**



Cloud Datastore

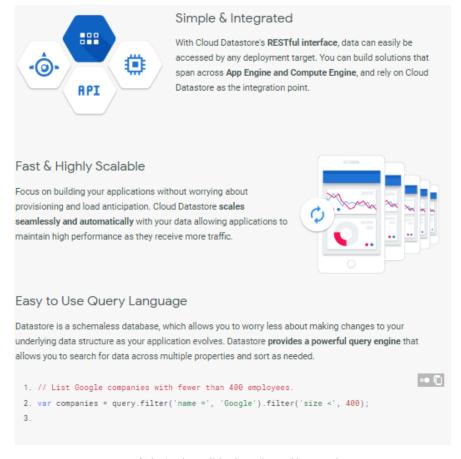
O Cloud Datastore é um banco de dados NoSQL altamente escalável para seus aplicativos da Web e móveis.

O Google Cloud Datastore fornece um banco de dados orientado a documentos elástico e altamente disponível como um serviço. Ele é

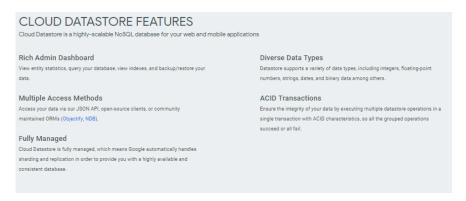
totalmente gerenciado, portanto, você não precisa implantar, atualizar, configurar ou gerenciar sua solução de banco de dados. O Cloud Datastore vem com um rico painel de administração, um poderoso mecanismo de consulta e vários métodos para acessar o banco de dados, o que o torna ideal para cargas de trabalho móveis e da Web.

# Banco de dados NoSQL altamente escalável

O Cloud Datastore é um banco de dados NoSQL altamente escalável para seus aplicativos. O Cloud Datastore **lida automaticamente com fragmentação e replicação**, fornecendo um banco de dados altamente disponível e durável que é dimensionado automaticamente para lidar com a carga de seus aplicativos. O Cloud Datastore fornece uma infinidade de recursos, como **transações ACID**, **consultas semelhantes a SQL**, **índices e muito mais**.



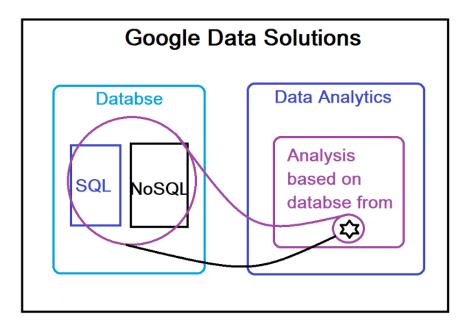
Referências: https://cloud.google.com/datastore/



referências: https://cloud.google.com/datastore/

Depois de ler a introdução, uma coisa que vem a cada mente é como usá-la. Uma vez que é plataforma de armazenamento de banco de dados para bancos de dados NoSQL, o aplicativo armazenará bancos de dados para ele não arquivos ou pastas.

Deixe-me esclarecer a diferença aqui por



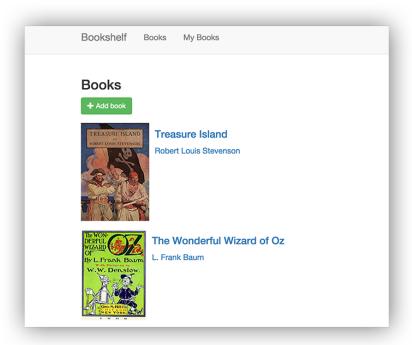
soluções

Então, vamos fazer um exemplo.

# **App Bookshelf PHP**

O aplicativo Bookshelf é um exemplo de aplicativo da web escrito em PHP que mostra como usar vários produtos do Google Cloud Platform, incluindo:

- Ambiente flexível do Google App Engine
- Google Cloud SQL
- Google Cloud Datastore
- Google Cloud Storage
- Google Cloud Pub / Sub



O aplicativo de amostra Bookshelf armazena uma coleção de títulos de livros. Qualquer pessoa que tenha acesso ao aplicativo pode adicionar livros à lista. O aplicativo de amostra oferece esses recursos:

- Os usuários podem visualizar a lista de livros, adicionar livros à lista e remover livros da lista.
- Os usuários podem editar detalhes do livro.
- Os usuários podem fazer upload de imagens de capa para livros.
- Os usuários podem fazer login com suas contas do Google e visualizar os livros que eles adicionaram à lista.

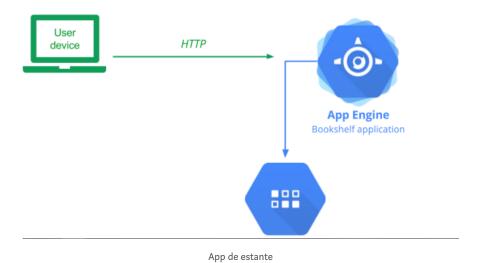
. . .

# **Objetivos**

- Clone ou faça o download do aplicativo de amostra.
- Construa o aplicativo e execute-o em sua máquina local.
- Implante o aplicativo no App Engine.
- Percorra o código de exemplo.
- Saiba como o aplicativo armazena dados estruturados.
- Saiba como o aplicativo armazena dados binários no Google Cloud Storage.
- Saiba como o aplicativo autentica usuários.
- Saiba como o aplicativo cria registros de eventos visíveis no console do Google Cloud Platform.

Este tutorial explora detalhadamente o aplicativo Bookshelf e discute como cada recurso do aplicativo é implementado usando tecnologias e serviços familiares fornecidos pelo Cloud Platform.

Crie um aplicativo de estante que seja executado no mecanismo do aplicativo, mas salve o banco de dados no Cloud Datastore.

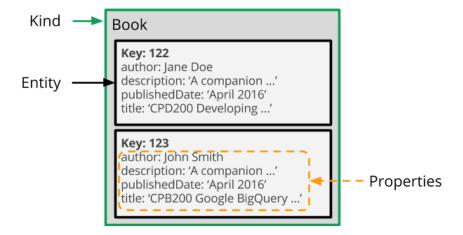


Um tipo no Cloud Datastore é aproximadamente equivalente a uma tabela em um banco de dados relacional. Os tipos são usados para categorizar entidades para consultas.

Os objetos de dados no Google Cloud Datastore são conhecidos como entidades. Uma entidade possui uma ou mais propriedades nomeadas (que são como campos em um banco de dados relacional). Cada propriedade pode ter um ou mais valores. As entidades são identificadas por uma chave exclusiva e são semelhantes a linhas em um banco de dados relacional. Entidades do mesmo tipo não precisam ter as mesmas propriedades. Essas características exclusivas implicam uma maneira diferente de projetar e gerenciar dados para aproveitar a capacidade de dimensionar automaticamente.

O Cloud Datastore também organiza dados em grupos de entidades. Um grupo de entidades consiste em uma entidade raiz e todos os seus descendentes. Os aplicativos geralmente usam grupos de entidades para organizar dados altamente relacionados para obter <u>uma consistência forte</u> e reforçar a <u>transacionalidade</u>. Por exemplo, um aplicativo pode usar um grupo de entidades para armazenar dados sobre um produto ou um perfil de usuário.

O modelo de dados do Bookshelf para o Cloud Datastore é ilustrado pelo diagrama a seguir. Observe que o modelo de dados simples usado no Bookshelf não faz uso de grupos de entidades.



# Antes de você começar

- Crie um novo projeto do GCP e, em seguida, crie um aplicativo do App Engine e ative o faturamento nesse projeto.
- Ative as APIs Cloud Datastore, Cloud Pub / Sub, Cloud Storage JSON, Stackdriver Logging, Google+ e Google Cloud SQL.

Verifique se o seu projeto padrão está correto:

```
lista de configuração do gcloud
```

Se o ID do projeto listado na saída não for o projeto que você pretendia usar para este tutorial, configure o projeto inserindo este comando:

```
gcloud config set project [YOUR_PROJECT_ID]
```

Onde [YOUR\_PROJECT\_ID] está o ID do projeto que você criou ou escolheu para usar neste tutorial.

Clone o repositório de amostra:

```
git clone https://github.com/GoogleCloudPlatform/getting-
started-php.git
```

Alternativamente, você pode <u>baixar o exemplo</u> como um arquivo zip e extraí-lo.

Agora, aqui vem os principais passos:

Etapa 1: repositório Clone do github

```
prashantagcppaudel @ cloudshell: ~ (pfour-221109) clone do $
git https://github.com/GoogleCloudPlatform/getting-started-
php.git
  Clonando em 'getting-started-php' ...
remoto: Enumerando objetos: 7, concluído .
remoto: Contando objetos: 100% (7/7), pronto.
remote: Compactando objetos: 100% (7/7), pronto.
remoto: Total 2143 (delta 0), reutilizado 2 (delta 0),
reutilizado na embalagem 2136
Objetos recebidos: 100% (2143/2143), 1,04 MiB | 0 bytes / s,
pronto.
Resolvendo deltas: 100% (1419/1419), pronto.
```

### Etapa 2: Ir para pasta de início e pasta de 2 dados estruturados

```
$ cd getting-started-php /
$ cd 2-dados estruturados /
```

Etapa 3: copie settings.yaml.dist para settings.yaml

```
$ cp config / settings.yml.dist config / settings.yml
```

Etapa 4: Novamente, vá para a pasta de configuração

```
$ cd config /
```

Etapa 5: Editar configurações .yml da seguinte maneira

```
GNU nano 2.7.4
                                                               File:
# Google credentials and configuration
google_client_id: YOUR_CLIENT_ID
google client secret: YOUR CLIENT SECRET
                        pfour-221109
google_project_id:
# options are "mysql", "postgres", "mongodb", or "datastore"
bookshelf_backend: datastore
# configure CloudSQL for MySQL or PostgreSQL backend
cloudsql_connection_name: INSTANCE_CONNECTION_NAME
cloudsql_database_name: YOUR_DATABASE_NAME
cloudsql_user: YOUR_CLOUDSQL_USER
cloudsql password: YOUR CLOUDSQL PASSWORD
cloudsql_port: 3306
# configure Mongo backend
mongo_url: YOUR_MONGODB_URL
mongo database: YOUR MONGO DATABASE
mongo_collection: YOUR_MONGO_COLLECTION
```

Etapa 6: Instalar o compositor

```
$ composer install
 Carregando repositórios do compositor com informações do
Instalando dependências (incluindo require-dev) do arquivo
de bloqueio
Operações do pacote: 52 instalações, 0 atualizações, 0
remoções
  - Instalando o firebase / php-jwt (v4.0.0): Fazendo
download (100%)
  - Instalando o rize / uri-template (0.3.2): Baixando
(100\%)
  - Instalando o psr / http-message (1.0.1): Baixando (100%)
  - Instalando o psr / log (1.0.2): Baixando (100% )
  - Instalando monolog / monolog (1.23.0): Baixando (100%)
  - Instalando guzzlehttp / psr7 (1.4.2): Baixando (100%)
  - Instalando guzzlehttp / promises (v1.3.1): Baixando
(100\%)
  - Instalando o guzzlehttp / guzzle (6.3.0): Baixando
(100\%)
  - Instalando o psr / cache (1.0.1): Baixando (100%)
  - Instalando o google / auth (v0.11.1): Baixando (100%)
  - Instalando google / cloud (v0.21.1): Baixando (100%)
  - Instalando symfony / routing (v3.0.9): Baixando (100%)
  - Instalando symfony / polyfill-mbstring (v1.7.0):
Baixando (100%)
  - Instalando symfony / http-foundation (v3.0.9): Baixando
(100\%)
  - Instalando o symfony / event-dispatcher (v3.0.9):
Baixando (100%)
  - Instalando o symfony / debug (v3.4.4): Baixando (100%)
  - Instalando o symfony / http-kernel (v3.0.9): Baixando
(100\%)
  - Instalação de espinha / pimple (v1.1.1): Download (100%)
  - Instalação de silex / silex (v1.3.6): Download (100%)
monolog / monolog sugere a instalação de php-amqplib / php-
amgplib (Permitir o envio de mensagens de log para um
servidor AMQP usando php-amqplib) monolog
/ monolog sugere instalar php-console / php-console
(Permitir envio de mensagens de log para o Google Chrome)
monolog / monolog sugere instalar rollbar / rollbar
(Permitir envio de mensagens de log para Rollbar)
monolog / monolog sugere instalar ruflin / elastica
(Permitir envio de mensagens de log para um servidor Elastic
Search)
monolog / monolog sugere instalar sentinela / sentinela
(Permitir envio de mensagens de log para um servidor Sentry)
google / cloud sugere instalar google / gax (Necessário para
suportar gRPC)
O google / cloud sugere a instalação do google / proto-
client-php (necessário para suporte ao gRPC)
google / cloud sugere a instalação de james-heinrich /
getid3 (o cliente do Google Cloud Speech determina a taxa de
amostragem e a codificação das entradas de áudio)
symfony / routing sugere a instalação doutrina / anotações
(Para usar o carregador de anotações)
symfony / routing sugere instalar o symfony / config (Para
usar o roteador all-in-one ou qualquer carregador)
symfony / routing sugere instalar o symfony / dependency-
injection (Para carregar rotas de um serviço)
```

```
symfony / routing sugere instalar o symfony / expression-
language (Para usar a correspondência de expressões)
symfony / event-dispatcher sugere instalar symfony /
dependency-injection 0
symfony / http-kernel sugere a instalação do symfony /
class-loader
symfony / http-kernel sugere a instalação do symfony /
config
symfony / http-kernel sugere a instalação do symfony /
dependency-injection
symfony / http-kernel sugere a instalação do symfony /
symfony / http-kernel sugere a instalação do symfony / var-
dumper o
behat / mink sugere a instalação do behat / mink-selenium2-
driver (driver lento, mas compatível com JS para qualquer
aplicativo (requer Selenium2)) 0
behat / mink sugere a instalação do driver behat / mink-
zombie (driver headless rápido e habilitado para JS para
qualquer aplicativo (requer node.js ))
symfony / console sugere a instalação do symfony / lock
sebastian / global-state sugere a instalação de ext-uopz (*)
phpunit / phpunit-mock-objects sugere a instalação do ext-
soap (*)
phpunit / php-code-coverage sugere a instalação do ext-
xdebug (> = 2.2.1)
phpunit / phpunit sugere instalar o phpunit / php-invoker (~
1.1)
Gerando arquivos de autoload
```

Etapa 7: implantar o aplicativo no App Engine

```
$ gcloud app deploy
3acd409e5aea: Camada já existe
363b7c2debba:
Push ed59ef025e1e: Layer já existe
7420484a2283:
Push 5bdaf3588472: Layer já existe
03fce3219a27: Layer já existe
56f56dd1c2db: Layer já existe
2f4254df2a74: Layer já existe
3bfe3bd8f9e5:
Push 89d0200fdea0: Layer já existe
f7fb3b0f4fc6: A camada já existe
d464d4f089a3: A camada já existe
62f57a66e54a: A camada já existe
1c1171b5f438: A camada já existe
b00f214424ac: A camada já existe
96e208368397: A camada já existe
8c468b8e19fa: A camada já existe
b056439591da: A camada já existe
84ff92691f90: A camada já existe
mais recente: digest: sha256:
401c366752609e012efa055233becd40dce65dcb6ad70a23e820a2370fd5
```

```
f50f tamanho: 4911
FEITO

Serviço de atualização [ padrão] (isso pode levar alguns minutos) ... pronto.

Definir divisão de tráfego para o serviço [padrão] ... pronto.

Serviço implantado [padrão] para [https://pfour-221109.appspot.com]

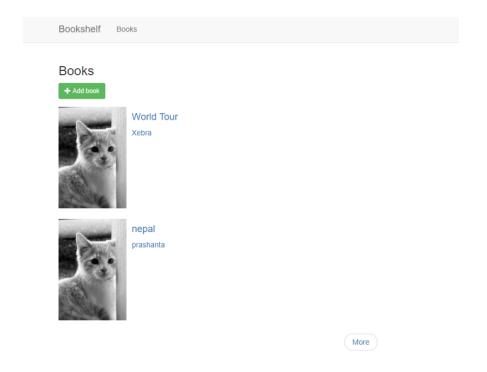
Você pode transmitir logs a partir da linha de comando executando:

$ gcloud app logs tail -s default

Para visualizar seu aplicativo no navegador da web, execute:

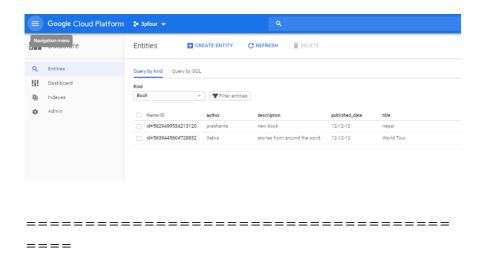
$ gcloud app browse
```

Passo 8: Agora verifique o site em <a href="https://pfour-221109.appspot.com">https://pfour-221109.appspot.com</a> .



## TERMINADO!

As informações são armazenadas no armazenamento de dados como segue



# **Google BigQuery**



Big Query

O BigQuery é um data warehouse em nuvem rápido, altamente escalável, econômico e totalmente gerenciado para análise, com aprendizado de máquina integrado.

O BigQuery é o data warehouse corporativo, altamente escalável e sem servidor do Google, projetado para tornar todos os analistas de dados produtivos com um desempenho de preço inigualável. Como não há infraestrutura para gerenciar, você pode se concentrar na análise de dados para encontrar insights significativos usando SQL familiar sem a necessidade de um administrador de banco de dados.

Analise todos os seus dados criando um data warehouse lógico sobre o armazenamento colunar gerenciado, bem como dados de armazenamento de objetos e planilhas. Construa e operacionalize

soluções de aprendizado de máquina com SQL simples. Compartilhe insights de forma fácil e segura em sua organização e além, como conjuntos de dados, consultas, planilhas e relatórios.

O BigQuery permite que as organizações capturem e analisem dados em tempo real usando sua poderosa capacidade de processamento de fluxo contínuo, para que suas percepções sejam sempre atualizadas e haja até 1 TB de dados analisados por mês e 10 GB de dados armazenados.

This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

Learn More about Medium's DNT policy

This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

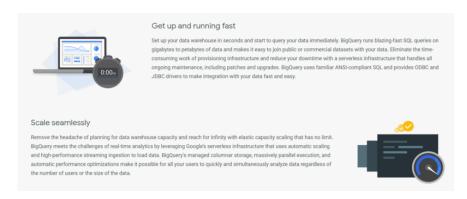
Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

Learn More about Medium's DNT policy

This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

## Learn More about Medium's DNT policy



Referência: https://cloud.google.com/bigquery/



Referência: https://cloud.google.com/bigquery/

### **BIGQUERY FEATURES**

A fast, highly scalable, cost-effective, and fully managed enterprise data warehouse for analytics.

#### Serverless

Serverless data warehousing gives you the resources you need, when you need them. With BigQuery, you can focus on your data and analysis, rather than operating and sizing computing resources.

#### Real-time Analytics

BigQuery's high-speed streaming insertion API provides a powerful foundation for real-time analytics. BigQuery allows you to analyze what's happening now by making your latest business data immediately available for analysis.

#### Automatic High Availability

Free data and compute replication in multiple locations means your data is available for query even in the case of extreme failure modes. BigQuery transparently and automatically provides durable, replicated storage and high availability with no extra charge and no additional setup.

#### Standard SO

BigQuery supports a standard SQL dialect which is ANSI:2011 compliant, reducing the need for code rewrite and allowing you to take advantage of advanced SQL features. BigQuery provides free ODBC and JDBC drivers to ensure your current applications can interact with BigQuery's powerful engine.

### Petabyte Scale

BigQuery is fast and easy to use on data of any size. With BigQuery, you'll get great performance on your data, while knowing you can scale seamlessly to store and analyze petabytes more without having to buy more capacity.

#### Flexible Pricing Models

BigQuery enables you to choose the pricing model that best suits you. Ondemand pricing lets you pay only for the storage and compute that you use. Flat-rate pricing enables high-volume users or enterprises to choose a stable monthly cost for analysis. For more information see BigQuery pricing.

#### Data Encryption and Security

You have full control over who has access to the data stored in BigQuery, BigQuery makes it easy to maintain strong security with fine-grained identity and access management with Cloud Identity and Access Management, and your data is always encrypted at rest and in transit.

#### Data Locality

You have the option to store your BigQuery data in US, Japan, and European locations while continuing to benefit from a fully managed service. BigQuery gives you the option of geographic data control, without the headaches of setting up and managing clusters and other computing resources in-region.

Referência: https://cloud.google.com/bigquery/

#### Standard SQL

BigQuery supports a standard SQL dialect which is ANSI:2011 compliant, reducing the need for code rewrite and allowing you to take advantage of advanced SQL features. BigQuery provides free ODBC and JDBC drivers to ensure your current applications can interact with BigQuery's powerful engine.

#### Federated Query and Logical Data Warehousing

BigQuery breaks down data silos so you can analyze all your data assets from one place. Through powerful federated query, BigQuery can process data in object storage (Cloud Storage), transactional databases (Cloud Bigtable), or spreadsheets in Google Drive — all without duplicating data. One tool lets you query all your data sources.

### Storage and Compute Separation

BigQuery provides you with fine-grained control of cost and access. With BigQuery's separated storage and compute, you pay only for the resources you use. You have the option to choose the storage and processing solutions that make sense for your business and control access for each.

### Automatic Backup and Easy Restore

BigQuery automatically replicates data and keeps a seven-day history of changes, reducing worries about unexpected data changes. This allows you to easily restore and compare data from different times.

#### Data Locality

You have the option to store your BigQuery data in US, Japan, and European locations while continuing to benefit from a fully managed service. BigQuery gives you the option of geographic data control, without the headeches of setting up and managing clusters and other computing resources in-region.

#### Foundation for AI

BigQuery provides a flexible, powerful foundation for machine learning and artificial intelligence. Besides bringing ML to your data with BigQuery ML, integrations with Cloud ML Engine and TensorFlow enable you to train powerful models on structured data. Moreover, BigQuery's ability to transform and analyze data helps you get your data in shape for machine learning.

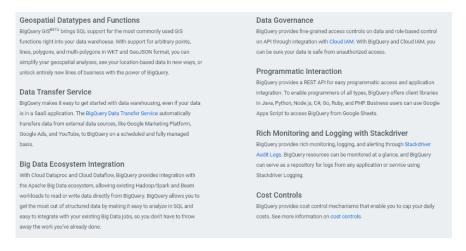
### Foundation for BI

BigQuery forms the data warehousing backbone for modern BI solutions, and enables seamless data integration, transformation, analysis, visualization, and reporting with tools from Google and our technology partners.

### Flexible Data Ingestion

Load your data from Cloud Storage or stream it into BigQuery at thousands of rows per second to enable real-time analysis of your data. Use familiar data integration tools like Informatica, Talend, and others out of the box.

Referência: https://cloud.google.com/bigquery/



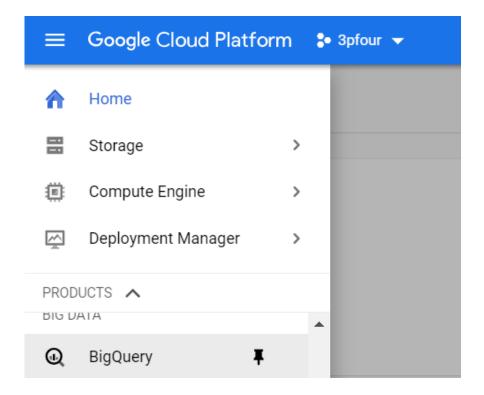
Referência: https://cloud.google.com/bigquery/

**Exemplo do BigQuery** 

# Executar uma consulta no conjunto de dados público

Antes de mais nada, você precisa ter sua conta do GCP e sua configuração de faturamento prontas.

Agora vá para o BigQuery em Big Data.



Você receberá uma interface do usuário antiga ou uma nova interface do usuário com base na abertura quando a nova interface do usuário estiver na versão beta.



BigQuery

Basicamente falando, esta ferramenta de interface do usuário e CLI é o mecanismo para processar todos os seus bancos de dados para fins analíticos.

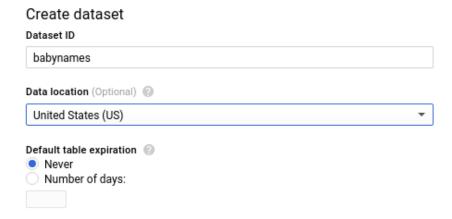
A tarefa mais básica é carregar os dados no armazenamento ou no próprio BigQuery e executar a consulta nele.

Primeiro, baixe os dados públicos do link <a href="http://www.ssa.gov/OACT/babynames/names.zip">http://www.ssa.gov/OACT/babynames/names.zip</a>

Descompacte este arquivo e verifique os dados. Selecione um ano e copie o arquivo txt para a área de trabalho

Novamente, vá para a interface do usuário do BigQuery Web e navegue até a seção **Recursos**, clique no nome do seu projeto.

- 1. No lado direito, no painel de detalhes, clique em **Criar conjunto** de dados .
- 2. Na página **Criar conjunto de dados** :
- Para ID do conjunto de dados, insira babynames.
- Para localização de dados , escolha Estados Unidos (EUA) .
   Atualmente, os conjuntos de dados públicos são armazenados na us região multi- localização . Para simplificar, você deve colocar seu conjunto de dados no mesmo local.

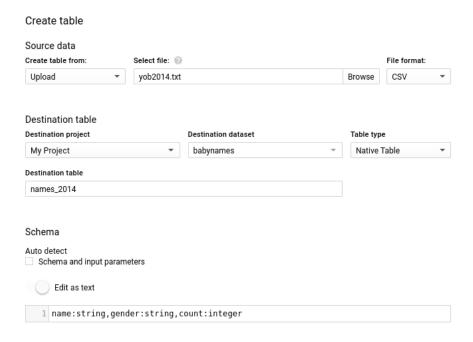


Deixe todas as outras configurações padrão no lugar e clique em **Criar conjunto de dados** .

# Carrega os dados em uma nova tabela

Em seguida, carregue os dados em uma nova tabela.

- No painel de navegação, na seção Recursos, clique no conjunto de dados babynames que você acabou de criar.
- 2. No lado direito, no painel de detalhes, clique em Criar tabela.
- 3. Use os valores padrão para todas as configurações, a menos que seja indicado de outra forma.
- 4. Na página Criar tabela:
- Para Fonte de Dados , clique em Tabela vazia e selecione
   Carregar .
- Para o arquivo Select, clique em Browse, navegue até o yob2014.txt arquivo e clique em Open.
- Para o formato de arquivo , clique em Avro e escolha CSV .
- Para a tabela de destino, insira names\_2014.
- Na seção **Esquema**, clique na opção **Editar como texto para** alternar e cole a seguinte definição de esquema na caixa.
- nome: string, gênero: string, contagem: integer



- 1. Clique em Criar Tabela.
- 2. Aguarde o BigQuery criar a tabela e carregar os dados. Enquanto o BigQuery carrega os dados, uma sequência (1 em execução) é exibida ao lado do histórico de tarefas no painel de navegação. A string desaparece depois que os dados são carregados.

# Visualizar a tabela

Depois que a sequência (1 em execução) desaparecer, você poderá acessar a tabela. Para visualizar as primeiras linhas dos dados:

- 1. Selecione **babynames > names\_2014** no painel de navegação.
- 2. No painel de detalhes, clique na guia **Visualizar**.



# Consulta a mesa

Agora que você carregou dados em uma tabela, pode executar consultas em relação a ela. O processo é idêntico ao exemplo anterior, exceto que, desta vez, você está consultando sua tabela em vez de uma tabela pública.

- 1. Se necessário, clique no botão **Compor nova consulta** . A menos que você tenha ocultado a janela de consulta anteriormente, ela ainda deve estar visível.
- 2. Copie e cole a consulta a seguir na área de texto da consulta. Esta consulta recupera os 5 principais nomes de bebês para homens dos EUA em 2014.
  - name, count
    FROM
    `babynames.names\_2014`
    WHERE
    gender = 'M'
    ORDER BY contagem DESC LIMITE 5

Clique em **Executar consulta** . Os resultados são exibidos abaixo da janela de consulta.

Query results		<b>≛</b> SA	VE AS ▼		M EXPLORE IN DATA STUDIO	
Query o	omplete (0	).945 sec ela	apsed, 621.	82 KB pro	cessed)	
Job information Results		JSON	JSON Execution deta		ls	
Row	name	count				
1	Noah	19263				
2	Liam	18440				
3	Mason	17177				
4	Jacob	16842				
5	William	16798				
			Result	ado da coi	nsulta	
	====					========
=						

# **Cloud Spanner**



Cloud Spanner

O Cloud Spanner é um sistema de banco de dados relacional totalmente gerenciado e de missão crítica . Tem propriedades especiais como

- 1. Alta disponibilidade
- 2. Consistência transacional
- 3. escala global

Com a semântica relacional tradicional, como esquemas, transação ACID, o SQL com chave de sincronização de replicação automática e síncrona é o único produto desse tipo no mercado.

A chave inglesa de nuvem é inicialmente escalonável horizontalmente. serviço de banco de dados relacional altamente consistente.

O Cloud Spanner é o único serviço de banco de dados de nível corporativo, globalmente distribuído e fortemente consistente criado para a nuvem, especificamente para combinar os benefícios de uma estrutura de banco de dados relacional com a escala horizontal nãorelacional. Essa combinação oferece transações de alto desempenho e consistência forte em linhas, regiões e continentes com um SLA de disponibilidade de 99,999%, líder do setor, sem tempo de inatividade planejado e segurança de nível corporativo. O Cloud Spanner revoluciona a administração e o gerenciamento de bancos de dados e torna o desenvolvimento de aplicativos mais eficiente.

No mundo sempre distribuído e globalmente distribuído, a eficiência de TI e de desenvolvedores, medida no tempo de inatividade e no

tempo de lançamento do aplicativo, é um dos recursos mais preciosos de uma organização. O desafio de gerenciar com eficiência os backends do banco de dados de aplicativos e, ao mesmo tempo, oferecer aos desenvolvedores as ferramentas necessárias para construir com eficiência era antes um desafio.

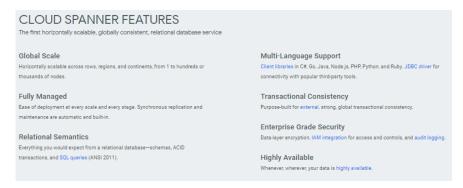


Referências: https://cloud.google.com/spanner/

#### Casos de Uso



Referências: https://cloud.google.com/spanner/



Referências: https://cloud.google.com/spanner/

\_\_\_\_\_

Exemplo: Chave Inglesa

A chave inglesa Cloud é usada apenas quando você precisa de melhor escalabilidade e maior disponibilidade.

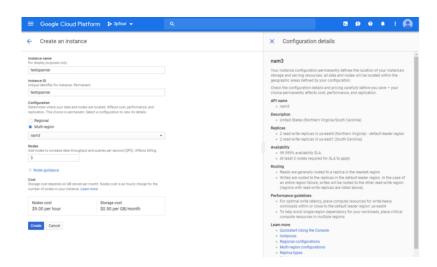
This embedded content is from a site that does not comply with the Do Not Track (DNT) setting now enabled on your browser.

Please note, if you click through and view it anyway, you may be tracked by the website hosting the embed.

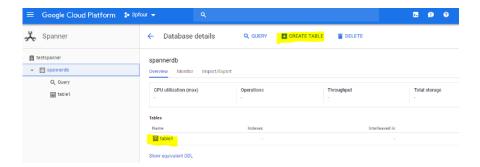
**Learn More about Medium's DNT policy** 

#### Criando uma Instância

Nos nós de computação do Cloud Spanner, o armazenamento é separado para permitir o dimensionamento de cada um independentemente. Além disso, seus nós e dados são replicados em várias zonas ou mesmo regiões para alta disponibilidade.

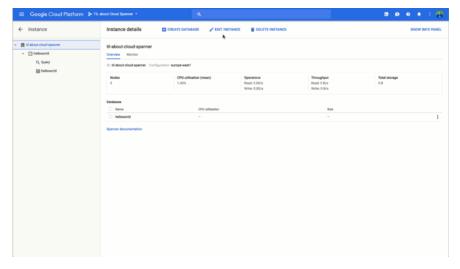


Além disso, criei tabela e coluna no banco de dados



#### Dimensionando uma Instância

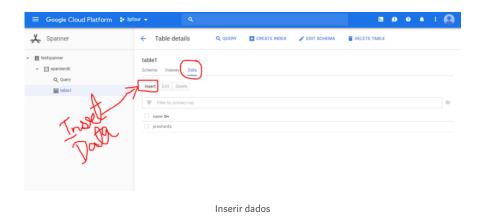
Para cada nó solicitado para sua instância, você obtém recursos de computação em todas as zonas e regiões, dependendo da configuração selecionada. Não há necessidade de planejamento e previsão difíceis sobre quantos recursos serão provisionados antecipadamente. Você pode apenas alterar o número de nós de instância sempre que precisar de mais ou menos.



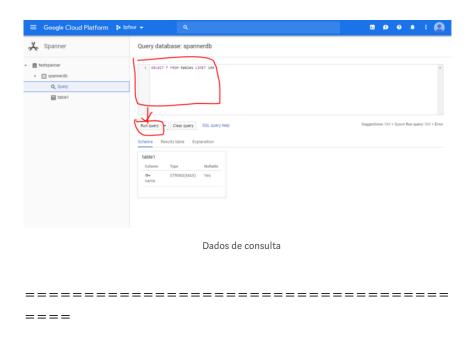
É fácil escalar sua instância do Cloud Spanner!

#### Leitura e Escrita de Dados

Leitura e escrita podem ser feitas a partir de máquinas clientes e através da API.



#### Dados de consulta



#### **CLUBE PUB / SUB**

Informe fluxos de eventos de qualquer lugar, em qualquer escala, para análise de fluxo simples, confiável e em tempo real

# Forneça dados de eventos onde você precisar

O Cloud Pub / Sub é uma base simples, confiável e escalável para análise de fluxo e sistemas de computação orientados a eventos. Como parte da solução de análise de fluxo do Google Cloud, o serviço ingere fluxos de eventos e os fornece para o Cloud Dataflow para processamento e o BigQuery para análise como uma solução de data warehousing. Confiar no serviço Cloud Pub / Sub para a entrega de dados de eventos permite que você se concentre em transformar seus negócios e sistemas de dados com aplicativos como:

- · Personalização em tempo real nos jogos
- Relatórios rápidos, segmentação e otimização em publicidade e mídia
- Dados do dispositivo de processamento para saúde, manufatura, petróleo e gás e logística
- Sindicando Fluxos de Dados Relacionados ao Mercado para Serviços Financeiros

# Construa aplicações multi-cloud e híbridas em arquitetura aberta

Distribuir dados em projetos e aplicativos executados em outras nuvens ou entre aplicativos na nuvem e locais. O Cloud Pub / Sub cabe facilmente em seu ambiente existente por meio de bibliotecas cliente eficientes para vários idiomas, APIs abertas de serviço REST / HTTP e gRPC e um conector Apache Kafka de software livre.

# Escala responsivamente e automaticamente

Escale para centenas de milhões de mensagens por segundo e pague apenas pelos recursos que você usa. Não há partições ou instâncias locais para gerenciar, reduzindo a sobrecarga operacional. Os dados são distribuídos de maneira automática e inteligente pelos datacenters por meio de nossa rede privada exclusiva de alta velocidade.

# Traga ferramentas de confiabilidade e segurança para aplicativos em tempo real

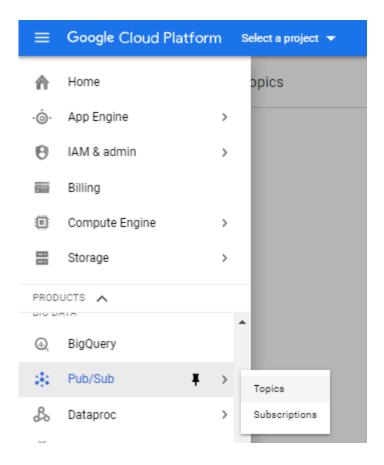
Use o Cloud Pub / Sub para simplificar sistemas distribuídos e escalonáveis. Todos os dados publicados são replicados de forma síncrona nas zonas de disponibilidade para garantir que as mensagens estejam disponíveis aos consumidores para processamento assim que estiverem prontas. Controles de acesso refinados permitem um compartilhamento de dados organizacional e organizacional sofisticado. E a criptografia de ponta a ponta adiciona segurança aos seus pipelines.

# At-least-once delivery Synchronous, cross-zone message replication and per-message receipt tracking ensures at-least-once delivery at any scale. Exactly-once processing Cloud Dataflow supports reliable, expressive, exactly-once processing of Cloud Pub/Sub streams. No provisioning, auto-everything Cloud Pub/Sub does not have shards or partitions. Just set your quota, publish and consume. Integrated Take advantage of integrations with multiple services, such as Cloud Storage and GMall update events and Cloud Functions for serverless event-driven computing.

Ref: https://cloud.google.com/pubsub/

Exemplo: Cloud Pub / Sub

Iniciando o pub / sub da nuvem



#### **Tutorial do Cloud Pub / Sub**

Este tutorial simples demonstra como escrever, implantar e acionar uma <u>função de nuvem em segundo plano</u> com um <u>disparador Cloud Pub / Sub</u>. Para saber mais sobre o Cloud Pub / Sub, consulte a <u>documentação</u> do Cloud Pub / Sub .

**Nota:** Você também pode usar <u>funções disparadas por HTTP</u> para ouvir <u>assinaturas push Pub / Sub</u>. Isso permite que uma única função do Cloud seja inscrita em vários tópicos.

# **Objetivos**

- Escreva e implemente uma <u>função de nuvem em segundo plano</u>.
- Acione a função publicando uma mensagem em um tópico do <u>Cloud Pub / Sub</u>.

# Preparando o aplicativo

- 1. Clone o repositório de aplicativos de amostra em sua máquina local:
- PYTHON (BETA)
- git clone <a href="https://github.com/GoogleCloudPlatform/python-docs-samples.gi">https://github.com/GoogleCloudPlatform/python-docs-samples.gi</a> t
- Como alternativa, você pode <u>baixar a amostra</u> como um arquivo zip e extraí-lo
- 2. Mude para o diretório que contém o código de amostra Cloud Functions para acessar o Cloud Pub / Sub:
- PYTHON (BETA)
- cd python-docs-samples / funções / helloworld /
- 1. Dê uma olhada no código de exemplo:
- PYTHON (BETA)
- 1. <u>funções / helloworld / main.py</u>
- 2. VISTA NO FEEDBACK DE GITHUB
- def hello\_pubsub (dados, contexto):

```
"" "Background Cloud Função a ser disparada pelo Pub / Sub.
```

Args:

data (dict): O dicionário com dados específicos para esse tipo de evento.

context (google.cloud.functions.Context): os metadados do evento Cloud Functions .

Importar base64

• se 'dados' em dados:

```
nome = base64.b64decode (dados ['dados']). decode ('utf-8')
outro:
nome = 'Mundo'
print ('Olá, {}!'. format (nome) )
```

### Implantando a função

Para implantar a função com um disparador Cloud Pub / Sub, execute o seguinte comando no diretório em que o código de amostra está localizado:

PYTHON (BETA)

```
gcloud funções implantar hello_pubsub --runtime python37 --
trigger-resource SEU_TOPIC_NAME --trigger-event
google.pubsub.topic.publish
```

onde YOUR\_TOPIC\_NAME é o nome do tópico do Cloud Pub / Sub para o qual a função será inscrita.

**Nota:** A implementação da função cria o tópico Cloud Pub / Sub especificado para você. Você também pode especificar um tópico existente em seu projeto para assinar.

# Acionando a função

- 1. Publique uma mensagem no seu tópico do Cloud Pub / Sub. Neste exemplo, a mensagem é um nome que a função incluirá em uma saudação:
- os tópicos do geloud pubsub são publicados vour\_topic\_name -- message vour\_name
- 1. Substitua YOUR\_TOPIC\_NAME pelo nome do seu tópico do Cloud Pub / Sub e YOUR\_NAME por uma string arbitrária.
- 2. Verifique os logs para ter certeza de que as execuções foram concluídas:
- gcloud funciona logs de leitura limite 50
- Nota: Os registros podem demorar alguns instantes para aparecer.
   Se você não os ver imediatamente, verifique novamente em um minuto ou dois.

. . .

## Limpando

Para evitar cobranças à sua conta do Google Cloud Platform pelos recursos usados neste tutorial:

# **Excluindo o projeto**

A maneira mais fácil de eliminar o faturamento é excluir o projeto criado para o tutorial.

\_\_\_\_\_

# **Cloud Bigtable**

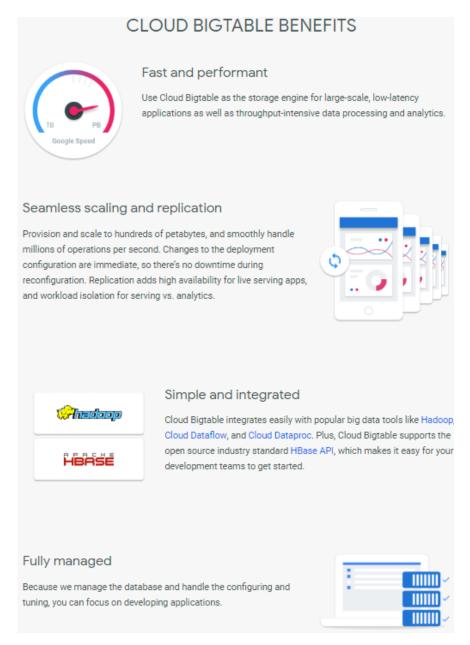


Cloud Bigtable

O Cloud Bigtable é um serviço de banco de dados NoSQL totalmente gerenciado que pode escalar até petabyte para cargas de trabalho analíticas e operacionais. Tem latência muito baixa tão pequena quanto 10ms. Tem opção de replicação em caso de falhas zonais.

A área mais utilizada para o Bigtable é o Machine Learning e o AI.

É facilmente integrável com as ferramentas atuais de big data, como Hadoop e HBase.



Referência: https://cloud.google.com/bigtable/

#### Casos de Uso

Cloud Bigtables são usados em tecnologia de anúncios, Fintech e IoT

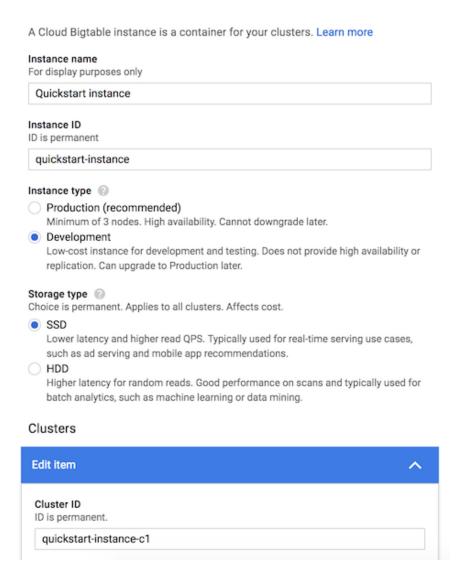
## Início rápido usando cbt

Esta página explica como usar o como usar o como para se conectar a uma instância do Cloud Bigtable, executar tarefas administrativas básicas e ler e gravar dados em uma tabela.

Se você estiver familiarizado com o HBase, convém seguir o início rápido usando o shell do HBase.

### Crie uma instância do Cloud Bigtable

- 1. Abra a página Criar Instância no Console do Google Cloud Platform.
- 2. ABRA A PÁGINA DE INSTÂNCIA DE CRIAR



- 1. Para o nome da instância, insira Quickstart instance.
- 2. Para ID da Instância, insira quickstart-instance.
- 3. Para o tipo de instância, selecione Desenvolvimento.

- 4. Para o tipo de armazenamento, selecione SSD.
- 5. Para **ID** do cluster, insira quickstart-instance-c1.
- 6. Para Região, selecione us-east1.
- 7. Para **Zone**, selecione **us-east1-c**.
- 8. Clique em **Criar** para criar a instância.

#### Conecte-se à sua instância

- 1. <u>Instale o Cloud SDK</u>, se ainda não o fez.
- 2. **Observação:** você também pode se conectar usando o Cloud Shell, fornecido com o Cloud SDK pré-instalado.
- 3. Abra uma janela de terminal, seja localmente ou com o Cloud Shell:
- 4. SHELL DE NUVEM ABERTA
- 5. Instale o cbt comando:
- gcloud components update
   gcloud components install cbt
- 1. Configure cbt para usar seu projeto e instância criando um .cbtrc arquivo, substituindo [PROJECT\_ID] pelo ID do projeto no qual você criou sua instância do Cloud Bigtable:
- echo project = [PROJECT\_ID] > ~/.cbtrc
   echo instance = quickstart-instance >> ~/.cbtrc

Agora você pode usar o cbt comando com sua instância!

#### Leia e escreva dados

O Cloud Bigtable armazena dados em *tabelas* , que contêm *linhas* . Cada linha é identificada por uma *chave de linha* .

Os dados em uma linha são organizados em *famílias* de colunas ou grupos de colunas. Um *qualificador de coluna* identifica uma única coluna dentro de uma família de colunas.

Uma *célula* é a interseção de uma linha e uma coluna. Cada célula pode conter várias *versões* de um valor.

- 1. Crie uma tabela nomeada my-table.
- cbt createtable my-table
- 1. Liste suas tabelas:
- cbt ls
- 1. O comando exibe uma saída semelhante à seguinte:
- my-table
- 1. Adicione uma família de colunas chamada cf1:
- cbt createfamily my-table cf1
- 1. Listar suas famílias de colunas:
- cbt ls my-table
- 1. O comando exibe uma saída semelhante à seguinte:
- cf1
- 1. Coloque o valor test-value na linha r1, usando a família de colunas cf1 e o qualificador de coluna c1
- cbt set my-table r1 cf1:c1=test-value
- 1. Use o cbt read comando para ler os dados que você adicionou à tabela:
- cbt read my-table
- 1. O shell exibe uma saída semelhante à seguinte:
- r1
  cf1:c1 @ 2016/10/31-15:05:38.840000
  "test-value"

- 1. Excluir a tabela my-table
- cbt deletetable my-table

. . .

#### Limpar

Para evitar cobranças à sua conta do Google Cloud Platform pelos recursos usados neste início rápido:

- 1. Abra a lista de instâncias do Cloud Bigtable no console do GCP.
- 2. ABRA A LISTA DE INSTÂNCIAS
- 3. Clique em instância de início rápido.
- 4. Clique em Excluir instância.

#### Delete instance and cluster?

Deleting this instance will permanently remove the instance, its cluster, and all the cluster's data. You cannot undo this later.

Confirm deletion by typing the instance ID below: quickstart-instance

quickstart-instance

CANCEL DELETE

- 1. Digite quickstart-instance e clique em **Excluir** para excluir a instância.
- 2. No seu terminal, exclua o .cbtrc arquivo:
- rm ~ / .cbtrc

-----

====

Desta forma, terminamos 3,4