

Série de Certificação GCP: 4.4 Gerenciando soluções de dados



Prashanta Paudel

12 de novembro de 2018 · 25 minutos de leitura

Este blog discute o gerenciamento de soluções de dados, incluindo muitos produtos e soluções GCP.

Como já discutimos coisas básicas em blogs anteriores, vamos nos aprofundar diretamente no conteúdo do curso.

=====

===

Execução de consultas para recuperar dados de instâncias de dados (por exemplo, Cloud SQL, BigQuery, Cloud Spanner, Cloud Datastore, Cloud Bigtable, Cloud Dataproc)

Consultas ajudam a exibir os dados no formato da instância.

Tradicionalmente, usamos consultas no banco de dados para mostrar os dados que satisfazem os campos que mencionamos na consulta.

vamos para as diferentes consultas da solução de dados.

SQL CLOUD

Depois que o banco de dados estiver ativo e em execução, você terá duas opções para consultar dados

1. do console
2. do Cloud Shell
3. Usando o Cloud SDK ou aplicativos de terceiros.

vamos ver o console e Shell como estes abrange todos os tipos de método

Do console

Só podemos realizar esses tipos de ações no Cloud SQL.

- Crie um banco de dados
- Exportar / importar um banco de dados
- duplicar / excluir um banco de dados

A consulta é feita a partir do Cloud Shell, VM ou outras ferramentas.

Do Cloud Shell

Primeiro, conecte-se à instância do shell

```
$ gcloud sql connect mydata --usuário = root --quiet
Lista de permissões do seu IP para conexão de entrada por 5
minutos ... pronto.
Conectando-se ao banco de dados com usuário SQL [root].
Digite a senha:
Bem-vindo ao monitor MariaDB. Comandos terminam com; ou \ g.
Seu ID de conexão do MySQL é 263
Versão do servidor: 5.7.14-google-log (Google)

Copyright (c) 2000, 2017, Oracle, MariaDB Corporation Ab e
outros.

Digite 'help;' ou '\ h' para ajuda. Digite '\ c' para limpar
a instrução de entrada atual.

MySQL [(nenhum)]>
```

Então podemos fazer várias operações no banco de dados nos comandos do MySQL.

Visualizando informações de resumo da instância

Você pode visualizar informações resumidas sobre suas instâncias do Cloud SQL

```
Instâncias de gcloud sql descrevem [INSTANCE_NAME]
```

Crie a instância

```
Instâncias de gcloud sql criam [INSTANCE_NAME] --tier =  
[MACHINE_TYPE] --region = [REGION]  
  
Instâncias de $ gcloud sql criam pop1011 --tier = db-n1-  
standard-1 - região = us-central1  
Criando a instância do Cloud SQL ... concluída.  
Criado  
[https://www.googleapis.com/sql/v1beta4/projects/fourpointfo  
ur-222308/instances/pop1011].  
NOME DATABASE_VERSION NÍVEL DA LOCALIZAÇÃO PRIMARY_ADDRESS  
STATUS  
PRIVATE_ADDRESS pop1011 MYSQL_5_7 us-central1-a db-n1-  
padrão-1 35.225.244.102 - RUNNABLE
```

Não inclua informações confidenciais ou de identificação pessoal no nome da sua instância; é visível externamente.

Você não precisa incluir o ID do projeto no nome da instância. Isso é feito automaticamente quando apropriado (por exemplo, nos arquivos de log).

MACHINE_TYPE é um dos valores da etapa anterior que começa com `db-` .

Por exemplo, o comando a seguir cria uma instância de segunda geração chamada `instance1` com o tipo de máquina `db-n1-standard-2` na região de Londres:

```
Instâncias de gcloud sql criam instance1 --tier = db-n1-  
standard-2 --region = europe-west2
```

Instância de clone

```
clone de instâncias de gcloud sql [SOURCE_INSTANCE_NAME]  
[TARGET_INSTANCE_NAME]
```

```
$ gcloud sql instancia clone pop1011 mypop
Clonando a instância do Cloud SQL ... pronto.
Criado
[https://www.googleapis.com/sql/v1beta4/projects/fourpointfo
ur-222308/instances/mypop].
NOME DATABASE_VERSION NÍVEL DA LOCALIZAÇÃO PRIMARY_ADDRESS
PRIVATE_ADDRESS STATUS
mypop MYSQL_5_7 us-central1-a db-n1-standard-1
35.232.220.175 - RUNNABLE
```

Excluindo Instância

```
instâncias de gcloud sql excluem [INSTANCE_NAME]

$ gcloud sql instances delete mypop
Todos os dados da instância serão perdidos quando a
instância for excluída.

Você quer continuar (S / n)? y

Excluindo a instância do Cloud SQL ... concluída.
Excluído
[https://www.googleapis.com/sql/v1beta4/projects/fourpointfo
ur-222308/instances/mypop].
```

QUERIES no MySQL

```
$ gcloud sql connect mydata --usuário = root --quiet
Lista de permissões do seu IP para conexão de entrada por 5
minutos ... pronto.
Conectando-se ao banco de dados com usuário SQL [root].
Digite a senha:
Bem-vindo ao monitor MariaDB. Comandos terminam com; ou \ g.
Seu ID de conexão do MySQL é 263
Versão do servidor: 5.7.14-google-log (Google)

Copyright (c) 2000, 2017, Oracle, MariaDB Corporation Ab e
outros.

Digite 'help;' ou '\ h' para ajuda. Digite '\ c' para limpar
a instrução de entrada atual.

MySQL [(none)]>
OR da VM
#mysql
```

```

MOSTRAR BASES DE DADOS - Mostrar
todas as bases de dados neste servidor
MySQL [(nenhuma)]> MOSTRAR BASES DE DADOS
->;
+ ----- +
| Banco de dados |
+ ----- +
| information_schema |
| classicmodels |
| mysql |
| performance_schema |
| sys |
+ ----- +
5 linhas no conjunto (0,04 seg.)

USE databaseName - Define o banco
de dados padrão (atual)

MySQL [(nenhum)]> use classicmodels
Leitura de informações da tabela para conclusão de nomes de
tabela e coluna
Você pode desativar esse recurso para obter uma
inicialização mais rápida com -A

Database alterado

- Database-Level
DROP DATABASE databaseName - Exclua o banco
de dados (irrecuperável!)
DROP DATABASE SE EXISTE databaseName - Exclua se
existir
CREATE DATABASE databaseName - Crie um novo
banco de dados
CREATE DATABASE SE NÃO EXISTE databaseName - Crie somente se
não existir não existe

SELECT DATABASE () - Mostra o
banco de dados padrão

MySQL [classicmodels]> selecione banco de dados ()
->;
+ ----- +
| banco de dados () |
+ ----- +
| classicmodels |
+ ----- +
1 linha no set (0,04 seg.)

SHOW CREATE DATABASE databaseName - Mostra a
instrução CREATE DATABASE

- Tabela-Level
DROP TABLE [SE EXISTE] tableName , ...
CREATE TABLE [SE NÃO EXISTE] tableName (
    columnName columnType columnAttribute , ...
    PRIMARY KEY ( columnName ),
    FOREIGN KEY ( columnNmae ) REFERENCES tableName (

```

```

columnNmae )
)
SHOW TABLES - Mostra todas as tabelas no
banco de dados padrão
DESCRIBE | DESC tableName - Descreve os detalhes de uma
tabela
ALTER TABLE tableName... - Modificar uma tabela, por
exemplo, ADD COLUMN e DROP COLUMN
ALTER TABLE tableName ADD ColumnDefinition
ALTER TABLE tableName GOTA columnName
ALTER TABLE tableName ADD FOREIGN KEY ( columnNmae )
REFERÊNCIAS tableName ( columnNmae )
ALTER TABLE tableName GOTA FOREIGN KEY constraintName
SHOW CREATE TABLE tableName - Mostra a instrução
CREATE TABLE para este tableName

- Nível de Linha
INSERT INTO tableName
VALUES ( column1Value , column2Value, ...)
- Inserir em todas as colunas
INSERT INTO tableName
VALUES ( coluna1Valor , coluna2Valor , ...), ... -
Inserir várias linhas
INSERT INTO tableName ( column1Name , ..., columnNName )
VALUES ( column1Value ,. . ., columnNValue ) - Inserir nas
colunas selecionadas
DELETE FROM tableName WHERE critérios
UPDATE tableName SET columnName = expr , ... WHERE criteria
SELECT * | column1NameAS alias1 , ..., columnNName como
aliasN
FROM tableName
WHERE critérios
GROUP BY columnName
ORDER BY columnName ASC | DESC, ...
HAVING groupConstraints Contagem
LIMIT | contagem de offset MySQL [classicmodels]> select
* de clientes onde postalcode = 41101; + ----- +
----- + --- ----- + -----
----- + ----- + ----- + -----
---- + ----- + ----- + ----- + ----- + ----
----- + - ----- +

| customerNumber | customerName | contactLastName |
contactFirstName | telefone | addressLine1 | addressLine2 |
cidade | estado | postalCode | país | salesRepEmployeeNumber
| creditLimit |
+ ----- + ----- + ----- + -----
----- + ----- + ----- + -----
----- + ----- + ----- + -----
-- + ----- + ----- + - ----- +
| 484 | Iberia Gift Imports, Corp. | Roel | José Pedro |
(95) 555 82 82 | C / Romero, 33 | NULL | Sevilha | NULL |
41101 | Espanha | 1702 | 65700,00 |
+ ----- + ----- + ----- + -----
----- + ----- + ----- + -----
----- + ----- + ----- + -----
-- + ----- + ----- + - ----- +
1 linha no set, 20 avisos (0.04 seg)

```

- Outros
SHOW WARNINGS; - Mostrar os avisos da declaração anterior

=====

=====

BigQuery

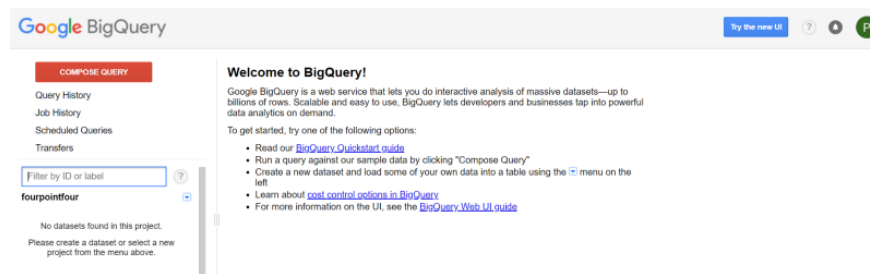
Bigquery é a plataforma onde podemos ***importar enormes bancos de dados em terrabytes e executar uma consulta sobre eles.***

É basicamente uma plataforma de consulta em execução, em vez de armazenar ou editar bancos de dados.

Podemos executar uma consulta no banco de dados público diretamente a partir da plataforma, importar o banco de dados para o BigQuery e, em seguida, executar a consulta.

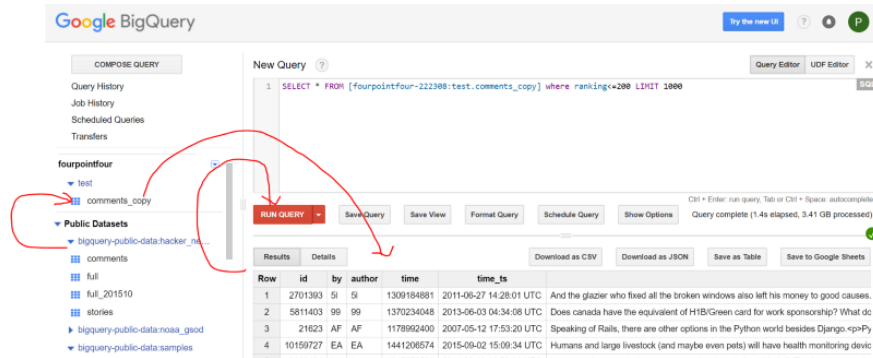
A grande diferença estará no tamanho do banco de dados que precisa ser armazenado no armazenamento em nuvem.

Primeiro, precisamos configurar um banco de dados.



Painel do BigQuery

Três etapas para executar uma consulta importando dados



três etapas para consulta de dados públicos

O principal objetivo é criar um job de consulta e recuperar o resultado por meio da API ou exportar o resultado para o Data Studio para análise e representação adicionais.

Na Shell

```
prashantapaudel_7 @ cloudshell: ~ (fourpointfour-222308) $
bq shell
```

```
Bem-vindo ao BigQuery! (Digite ajuda para mais
informações.)
Fourpointfour-222308>
```

```
fourpointfour-222308> SELECT * FROM [bigquery-public-
data.samples.shakespeare]
```

```
Esperando por bqjob_r517d95cd678a9617_000000167078ee625_1
... (1s) Estado atual: DONE
```

```
+-----+-----+-----+-----+
+
| palavra | word_count | corpus | corpus_date |
+-----+-----+-----+-----+
+
| LVII | 1 | sonetos | 0 |
| augurs | 1 | sonetos | 0 |
```

```
| lamento | 1 | sonetos | 0 |
| whereupon | 1 | sonetos | 0 |
| limpa | 1 | sonetos | 0 |
| porção | 1 | sonetos | 0 |
| juramentos | 1 | sonetos | 0 |
| estrangeiro | 1 | sonetos | 0 |
| aviso | 1 | sonetos | 0 |
| Ervas Daninhas | 1 | sonetos | 0 |
| ducaosa | 1 | sonetos | 0 |
| De abril | 1 | sonetos | 0 |
| season'd | 1 | sonetos | 0 |
| excede | 1 | sonetos | 0 |
| miss'd | 1 | sonetos | 0 |
+-----+-----+-----+-----+
```



```
+
fourpointfour-222308>
```

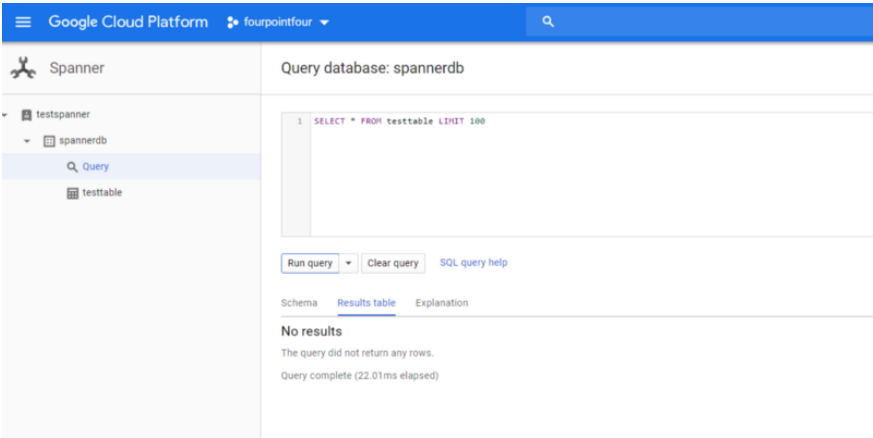
=====

===

Cloud Spanner

A pesquisa de chave de nuvem é semelhante ao BigQuery e ao SQL, pois todos seguem o formato SQL.

Em cloud spanner, temos que criar um banco de dados relacional ou criar um pipeline onde os bancos de dados são criados por meio da instância do Dataflow. então podemos consultar os dados como semelhantes aos anteriores



Cloud Spanner

=====

===

Cloud Bigtable

O Cloud Bigtable é um sistema de gerenciamento de banco de dados NoSQL. Ele armazena dados em grande formato tabular.

Instance ID
ID is permanent
mybigtable

Instance type ⓘ
☒ **Production (recommended)**
 Minimum of 3 nodes. High availability. Cannot downgrade later.
☐ **Development**
 Low-cost instance for development and testing. Does not provide high availability or replication. Can upgrade to Production later.

Storage type ⓘ
 Choice is permanent. Applies to all clusters. Affects cost.
☒ **SSD**
 Lower latency and higher read QPS. Typically used for real-time serving use cases, such as ad serving and mobile app recommendations.
☐ **HDD**
 Higher latency for random reads. Good performance on scans and typically used for batch analytics, such as machine learning or data mining.

Clusters
 mybigtable-c1 (us-central1-f -- 3 nodes) *Not saved* ✎
[+ Add replicated cluster](#)

Adding a cluster automatically replicates data between all clusters.
[Replication guidance](#)

[Create](#) [Cancel](#)

Monthly cost estimate
 \$1,593.50 per month (1,000 GB data, 3 nodes)
 Hourly rate effectively \$2.18
 Try another storage size (per cluster)
 1000 GB

Item	Estimated cost
1 cluster	\$1,423.50/month
1000 GB SSD	\$170.00/month
Total	\$1,593.50

Node charges are for provisioned resources, regardless of node usage. The same node charges apply even if your instance is inactive. [Learn more](#)

Comece criando uma instância com um único cluster. Use o `bigtable instances create` comando para criar uma instância:

```
gcloud bigtable instances create INSTANCE_ID \
  --cluster=CLUSTER_ID \
  --cluster-zone=CLUSTER_ZONE \
  --display-name=DISPLAY_NAME \
  [--cluster-num-nodes=CLUSTER_NUM_NODES] \
  [--cluster-storage-type=CLUSTER_STORAGE_TYPE] \
  [--instance-type=INSTANCE_TYPE]
```

\$ gcloud Instâncias bigtable criam abcxyz --cluster = hell12345 --cluster-zone = us-central1-f --display-name = yesplease
 Criando instância bigtable abcxyz ... pronto.

Forneça os seguintes valores:

- `INSTANCE_ID` : O identificador permanente da instância.
- `CLUSTER_ID` : O identificador permanente para o cluster.
- `CLUSTER_ZONE` : A zona onde o cluster é executado.
- Se você planeja usar a replicação, verifique se o Cloud Bigtable está disponível em pelo menos uma outra zona em sua região preferida. Veja a lista de zonas.

- `DISPLAY_NAME` : Um nome legível que identifica a instância no console do GCP.

O comando aceita os seguintes sinalizadores opcionais:

- `--cluster-num-nodes=CLUSTER_NUM_NODES` : O número de nós no cluster. Os clusters em uma instância de produção devem ter três ou mais nós. O valor padrão é `3` . Se você não tem certeza de quantos nós precisa, use o padrão. Você pode adicionar mais nós depois. [Saiba mais](#) .
- Não use este sinalizador para instâncias de desenvolvimento.
- `--cluster-storage-type=CLUSTER_STORAGE_TYPE` : O tipo de armazenamento a ser usado para o cluster. Cada cluster em uma instância deve usar o mesmo tipo de armazenamento. Aceita os valores `SSD` e `HDD` . O valor padrão é `SSD` .
- Na maioria dos casos, o valor padrão é o melhor. Essa escolha é permanente. Saber mais.
- `--instance-type=INSTANCE_TYPE` : O tipo de instância a criar. Aceita um dos seguintes valores:
- `PRODUCTION` (padrão): uma instância completa de alta disponibilidade. Essa escolha é permanente. Saber mais.
- `DEVELOPMENT` : Uma instância de baixo custo para desenvolvimento e teste, com desempenho limitado e sem SLA. Você pode atualizar para uma instância de produção posteriormente. Saber mais.

=====

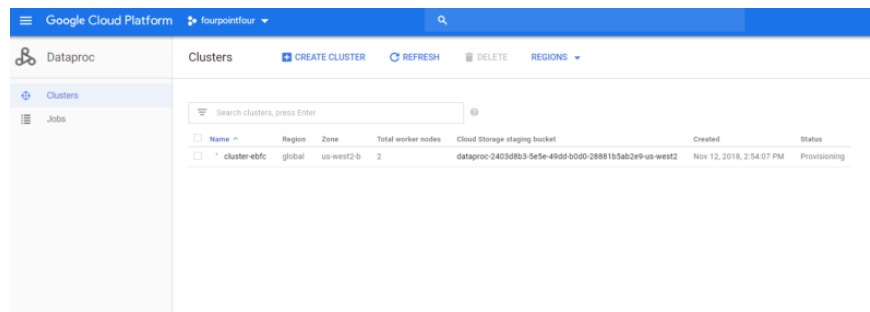
===

Cloud Dataproc

Para usar o Dataproc, você precisa ativar a API primeiro.

O Google Cloud Dataproc permite provisionar clusters do Apache Hadoop e se conecta a armazenamentos de dados analíticos subjacentes

Criar um Cluster



cluster de dataproc

Depois de criar o cluster, agora adicione um trabalho ao cluster.

=====

==

Estimando o custo da consulta do Bigquery

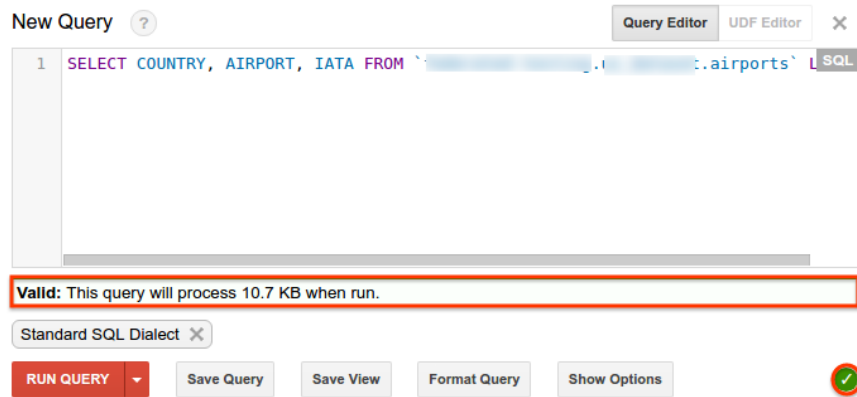
As consultas sob demanda são cobradas com base no número de bytes lidos. Para preços de consulta on-demand atuais, consulte a página de preços.

Para estimar custos antes de executar uma consulta usando o:

- Validador de consulta na interface do usuário da web
- `--dry_run` sinalizar no CLI
- `dryRun` parâmetro ao enviar um trabalho de consulta usando a API
- Calculadora de preços do Google Cloud Platform

Estimando custos de consulta

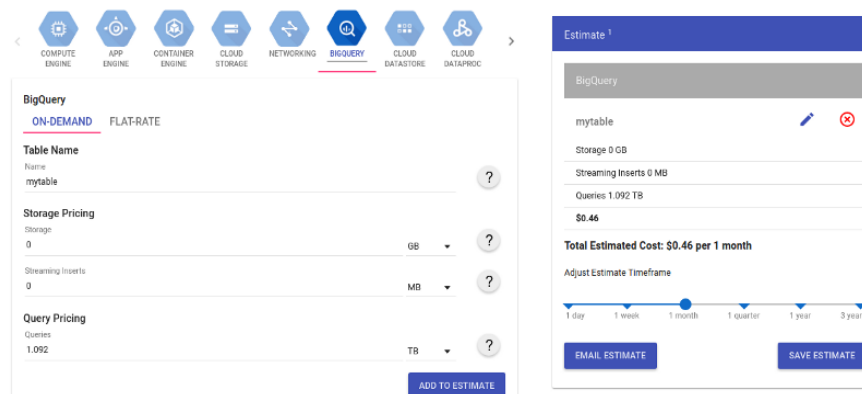
Quando você insere uma consulta na interface do usuário da web, o validador de consulta verifica a sintaxe da consulta e fornece uma estimativa do número de bytes lidos. Você pode usar essa estimativa para calcular o custo da consulta na calculadora de preços.



Se sua consulta processar uma pequena quantidade de dados, talvez seja necessário converter os bytes processados de KB para MB. MB é a menor medida usada pela calculadora de preços.

Estimando os custos de consulta usando a Calculadora de preços do Google Cloud Platform

Para estimar os custos de consulta sob demanda na Calculadora de preços do Google Cloud Platform, insira o número de bytes processados pela consulta como MB, GB, TB ou PB. Se sua consulta processa menos de 1 TB, a estimativa é de US \$ 0, pois o BigQuery oferece gratuitamente 1 TB de processamento de consultas sob demanda por mês.



Para estimar o custo de uma consulta usando a calculadora de preços:

1. Abra a Calculadora de preços do Google Cloud Platform,

2. Clique em BigQuery.
3. Clique na guia **On-Demand**.
4. Para **Nome da Tabela**, digite o nome da tabela. Por exemplo, `airports`.
5. Para **Preços de armazenamento**, insira `0` no campo **Armazenamento**.
6. Para **Query Pricing**, insira os bytes estimados lidos em sua execução a seco ou no validador de consulta. Se o valor for menor que 1 MB, você deverá convertê-lo em MB para a calculadora de preços. Usando 10918 bytes como exemplo, o valor é aproximadamente 0,01091 MB.

BigQuery

ON-DEMAND FLAT-RATE

Table Name

Name

airports

Storage Pricing

Storage

0 GB

Streaming Inserts

MB

Query Pricing

Queries



0.010918 MB

ADD TO ESTIMATE

1. Clique em **Adicionar à estimativa**.
2. A estimativa aparece à direita. Observe que você pode salvar ou enviar a estimativa por e-mail.

Estimate ¹

BigQuery

airports  

Storage 0 GB

Streaming Inserts 0 MB

Queries 0.000 TB

\$0.00

Total Estimated Cost: \$0.00 per 1 month

Adjust Estimate Timeframe

1 day

1 week

1 month

1 quarter

1 year

3 years

EMAIL ESTIMATE

SAVE ESTIMATE

Nesse caso, o número de bytes lidos pela consulta está abaixo do processamento de 1 TB sob demanda fornecido pelo nível gratuito. Como resultado, o custo estimado é de US \$ 0.

Incluindo preços de taxa fixa na calculadora de preços

Se você tiver preços fixos aplicados à sua conta de faturamento, poderá clicar na guia **Taxa fixa**, escolher seu plano de taxa fixa e adicionar seus custos de armazenamento à estimativa.

The screenshot shows the Google Cloud Pricing Calculator interface. On the left, a sidebar lists various Google Cloud services: Compute Engine, App Engine, Container Engine, Cloud Storage, Networking, BigQuery (selected), Cloud Datastore, and Cloud Dataproc. The main panel is titled 'BigQuery' and shows the 'Flat-rate' pricing plan. Under 'Flat-rate plan', 'FlatRate-40' is selected. The 'Plan info' section shows 'Number of slots: 2000' and 'Query Throughput: 1'. The 'Storage Pricing' section shows 'Storage' as a variable. On the right, a summary box titled 'Estimate 1' shows 'BigQuery Flat-rate' with a 'flatRate-40' plan, 'Slots 2000', and 'Storage 20,480 GB'. The 'Total Estimated Cost' is '\$40,409.60 per 1 month'. A timeline slider is set to '1 month'. At the bottom, there are buttons for 'EMAIL ESTIMATE' and 'SAVE ESTIMATE'.

Para mais informações, consulte Preços fixos.

=====

=====

Fazendo backup e restaurando instâncias de dados (por exemplo, Cloud SQL, Cloud Datastore, Cloud Dataproc)

Cloud SQL

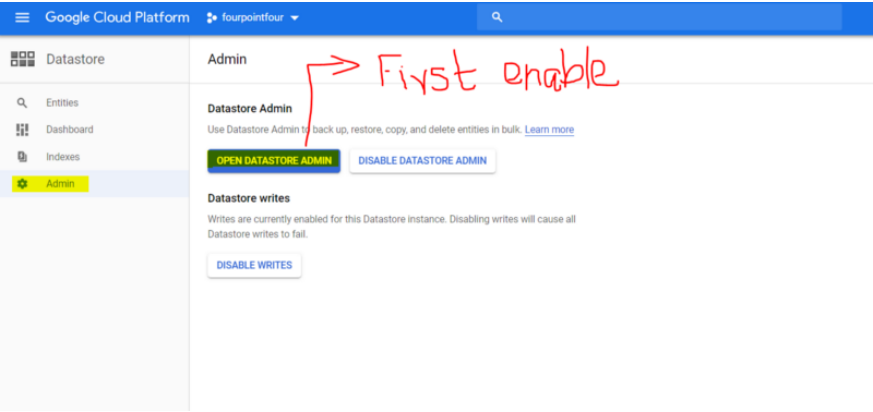
você pode fazer backup do banco de dados SQL em nuvem no Bucket.

The screenshot shows the Google Cloud Platform interface for exporting data from Cloud SQL to Cloud Storage. The top navigation bar shows 'Google Cloud Platform' and the user 'fourpointfour'. The left sidebar shows the 'SQL' service and the instance 'mydata-master-us-central1-f'. The main panel is titled 'Export data to Cloud Storage'. It prompts the user to 'Choose a Cloud Storage location and format for your Cloud SQL export'. Under 'Cloud Storage export location', there is a text input field with 'bucket/folder/file' and a 'Browse' button. Under 'Format', the 'SQL' option is selected, which 'Exports a plain text file with a sequence of SQL commands, like the output of mysqldump.' The 'CSV' option is also available, which 'Exports a plain text file with one line per row and comma-separated fields. Requires SQL SELECT query.' There is a 'Show advanced options' link and an 'Export' button. A note at the bottom states: 'When you click Export, we will grant a Cloud SQL service account write access to your bucket. Your bucket permissions will reflect this access.'

cópia de segurança

e Restaurar usando a importação do mesmo local ou gerenciando manualmente o banco de dados bac = kup.

Cloud Datastore



enable admin

então você pode ver o painel de administração para fazer backup e restaurar

Datastore Admin of fourpointfour-222308

The backup and restore functionality is being phased out, and will no longer be available as of Feb 28th, 2019. Please use the Cloud Datastore exports service for [Exporting and Importing](#).

Entities				
Entity Kind	# Entities	Avg. Size/Entity	Entities Size	Total Size
name	Stats not available	Stats not available	Stats not available	Stats not available

Backup EntitiesDelete Entities

Backups

Import Backup Information

=====

Revendo o status do trabalho no BigQuery

Gerenciando trabalhos do BigQuery

Depois de enviar um trabalho do BigQuery, você pode visualizar dados do trabalho, listar trabalhos, cancelar um trabalho ou reexecutar um

trabalho.

Quando um trabalho é enviado, ele pode estar em um dos três estados:

- `PENDING` - agendado
- `RUNNING`
- `DONE` - relatado como `SUCCESS` ou `FAILURE` (se o trabalho foi concluído com erros)

Visualizar dados do trabalho

Você pode visualizar dados e metadados do trabalho usando a interface do usuário da web, a CLI e a API. Esses dados incluem detalhes como o tipo de trabalho, o estado do trabalho e o usuário que executou o trabalho.

Permissões necessárias

Para obter dados e metadados do trabalho, você deve ter `bigquery.jobs.get` permissões. A seguinte função predefinida do IAM no nível do projeto inclui `bigquery.jobs.get` permissões:

- `bigquery.admin`

Se você conceder uma conta à `bigquery.admin` função, o usuário poderá visualizar todos os dados do trabalho no projeto, independentemente de quem enviou o trabalho.

As funções a seguir recebem `bigquery.jobs.get` permissões para tarefas criadas por você. Esses usuários podem ver somente os dados do trabalho para trabalhos enviados:

- `bigquery.user`
- `bigquery.jobUser`

Para mais informações sobre funções e permissões do IAM no BigQuery, consulte o controle de acesso.

Visualizar informações sobre trabalhos

Para ver informações sobre um trabalho:

1. No painel de navegação, clique em **Histórico de Trabalhos**.
2. Na seção **Trabalhos Recentes**, clique no trabalho para visualizar os detalhes.

Listando jobs em um projeto

Seu projeto mantém seu histórico profissional para todos os trabalhos criados nos últimos seis meses. Para solicitar a exclusão automática de tarefas com mais de 50 dias, entre em contato com o suporte.

Você pode visualizar seu histórico de tarefas do BigQuery por meio do Console do Google Cloud Platform, da CLI ou da API. Este histórico inclui trabalhos que estão no `RUNNING` estado e trabalhos que são `DONE` (indicados pelo relatório do estado como `SUCCESS` ou `FAILURE`).

Permissões necessárias

Para listar trabalhos, você deve ter `bigquery.jobs.list` permissões. As seguintes funções do IAM predefinidas no nível do projeto incluem `bigquery.jobs.list` permissões:

- `bigquery.user`
- `bigquery.admin`

Quando você recebe `bigquery.jobs.list` permissões, pode listar todas as tarefas em um projeto, mas os detalhes e os metadados são redigidos para trabalhos enviados por outros usuários. `bigquery.jobs.list` As permissões permitem que você veja todos os detalhes dos trabalhos criados por você.

Para listar todos os trabalhos, incluindo detalhes de trabalhos criados por outros usuários, você deve ter `bigquery.jobs.listAll` permissões. Apenas o `bigquery.admin` papel tem `bigquery.jobs.listAll` permissões.

A função a seguir é concedida `bigquery.jobs.list` apenas para tarefas criadas por você. Esses usuários só podem listar os trabalhos enviados por eles:

- `bigquery.jobUser`

Para mais informações sobre funções e permissões do IAM no BigQuery, consulte o controle de acesso.

Listando trabalhos

Quando você lista trabalhos em um projeto, não é necessário fornecer um local. Atualmente, os trabalhos são listados para todos os locais.

Para listar trabalhos em um projeto:

1. No painel de navegação, clique em **Histórico de Trabalhos**.
2. Na seção **Recentes**, seus trabalhos são listados por tempo de criação com os trabalhos mais recentes no topo. A lista inclui trabalhos apenas para o usuário atual. Para ver todos os trabalhos, use a ferramenta de linha de comando ou a API.

Cancelando trabalhos

Você pode cancelar um trabalho `RUNNING` ou `PENDING` na interface do usuário da web, na CLI ou na API. No entanto, nem todos os tipos de trabalho podem ser cancelados. Se o trabalho não puder ser cancelado, um erro será retornado.

Mesmo que o trabalho possa ser cancelado, o sucesso não é garantido. O trabalho pode ter sido concluído no momento em que a solicitação de cancelamento é enviada ou o trabalho pode estar em um estágio em que não pode ser cancelado.

Nota: Você ainda pode incorrer em custos depois de cancelar um trabalho, dependendo do estágio em que foi cancelado. Para mais informações, consulte Preços.

Permissões necessárias

Para cancelar um trabalho, você deve ter `bigquery.jobs.update` permissões. A seguinte função predefinida do IAM no nível do projeto inclui `bigquery.jobs.update` permissões:

- `bigquery.admin`

Se você conceder uma conta à `bigquery.admin` função, o usuário poderá cancelar qualquer trabalho qualificado, independentemente de quem a enviou.

As funções a seguir podem cancelar tarefas criadas por você. Esses usuários podem cancelar apenas os trabalhos enviados:

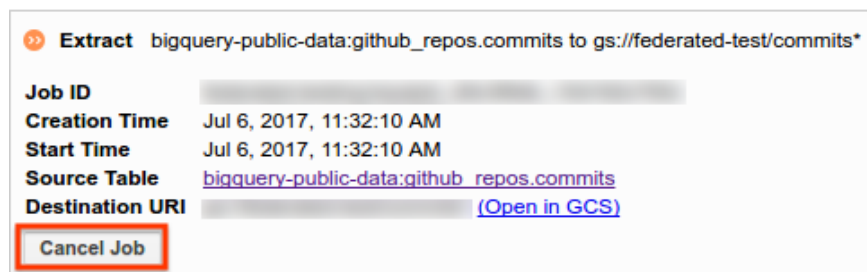
- `bigquery.user`
- `bigquery.jobUser`

Para mais informações sobre funções e permissões do IAM no BigQuery, consulte o controle de acesso.

Cancelando um trabalho

WEB UI

1. No painel de navegação, clique em **Histórico de Trabalhos**.
2. **Nota:** Você não pode cancelar um trabalho de consulta usando a interface do usuário da web.
3. Na seção **Trabalhos recentes**, clique no trabalho que você está cancelando. Os trabalhos mais recentes aparecem no topo da lista.
4. Nos detalhes do trabalho, clique em **Cancelar trabalho**.



Repetindo um trabalho

Não é possível reexecutar um trabalho usando o mesmo ID de trabalho. Em vez disso, você cria um novo trabalho com a mesma configuração. Quando você envia o novo trabalho na interface do usuário da web ou

no CLI, um novo ID de trabalho é atribuído. Quando você envia o trabalho usando a API ou as bibliotecas do cliente, deve gerar um novo ID de trabalho.

Permissões necessárias

Para executar um trabalho, você deve ter

`bigquery.jobs.create` permissões. As seguintes funções do IAM predefinidas no nível do projeto incluem `bigquery.jobs.create` permissões:

- `bigquery.user`
- `bigquery.jobUser`
- `bigquery.admin`

Para mais informações sobre funções e permissões do IAM no BigQuery, consulte o controle de acesso.

Executando novamente um trabalho

Para repetir um trabalho de consulta:

1. No painel de navegação, clique em **Histórico de consultas**.
2. Na seção **Consultas**, à direita da consulta, clique em **Abrir Consulta**.
3. Clique em **Executar consulta**.

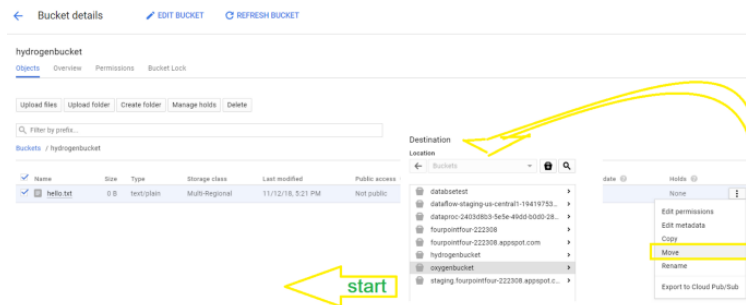
Para repetir um trabalho de carga:

1. No painel de navegação, clique em **Histórico de Trabalhos**.
2. Na seção **Trabalhos recentes**, clique no trabalho que você deseja repetir. Os trabalhos mais recentes aparecem no topo da lista.
3. Nos detalhes do trabalho, clique em **Repetir Carregar Trabalho**.
4. **Nota:** Você não pode repetir um trabalho de exportação ou um trabalho de cópia usando a interface do usuário da web.

=====

=====

Movendo objetos entre intervalos do Cloud Storage



objetos em movimento em baldes

No Cloud Shell

```
#gsutil mv -p gs://hydrogenbucket/hello.txt gs://oxygenbucket
```

Mudando Classes de Armazenamento de Objetos

Esta página descreve como alterar a classe de armazenamento de objetos em um bloco, sobrescrevendo o objeto. Para aprender como alterar as classes de armazenamento de objeto sem sobrescrever um objeto, consulte o recurso Gerenciamento de ciclo de vida do objeto. Para saber mais sobre as classes de armazenamento por objeto, consulte Classes de armazenamento por objeto.

Use o `-s` sinalizador em um `rewrite` comando. Por exemplo:

```
gsutil reescreva -s [STORAGE_CLASS] gs:// [PATH_TO_OBJECT]
```

Onde `[STORAGE_CLASS]` está a nova classe de armazenamento para o seu objeto e `[PATH_TO_OBJECT]` é o nome do objeto cuja classe você está

alterando.

Nota: Este guia envolve a substituição de dados. Os objetos Nearline Storage ou Coldline Storage têm encargos de exclusão antecipada se forem sobregravados a menos de 30 ou 90 dias do tempo de criação, respectivamente. Além disso, observe que, se você tiver ativado o controle de versão de objeto para o seu intervalo, o objeto original permanecerá no seu intervalo até que seja explicitamente excluído.

Alterando a Classe de Armazenamento Padrão de um Balde

Esta página mostra como alterar a *classe de armazenamento padrão* para seus intervalos. Quando você faz o upload de um objeto para o intervalo, se você não especificar uma classe de armazenamento para o objeto, o objeto é atribuído à classe de armazenamento padrão do intervalo. Para uma visão geral dos intervalos, leia os Termos-chave. Para aprender mais sobre classes de armazenamento, consulte Classes de armazenamento.

Para alterar a classe de armazenamento padrão de um intervalo existente:

Use o `gsutil defstorageclass set` comando com a classe de armazenamento desejada, substituindo `[VALUES_IN_BRACKETS]` pelos valores apropriados:

```
gsutil defstorageclass set [STORAGE_CLASS] gs: //  
[BUCKET_NAME]
```

A resposta parece com o seguinte exemplo:

```
Definir a classe de armazenamento padrão como "  
[STORAGE_CLASS]" para o intervalo gs: // [BUCKET_NAME]
```

```
=====  
===
```


Gerenciamento de ciclo de vida de objetos

Para oferecer suporte a casos de uso comuns, como definir um tempo de vida (TTL) para objetos, arquivar versões antigas de objetos ou "rebaixar" classes de armazenamento de objetos para ajudar a gerenciar custos, o Cloud Storage oferece o recurso Object Lifecycle Management. Esta página descreve o recurso, bem como as opções disponíveis ao usá-lo. Para aprender como ativar o Object Lifecycle Management e para exemplos de políticas de ciclo de vida, consulte [Gerenciando](#) ciclos de vida .

Introdução

Você pode atribuir uma configuração de gerenciamento de ciclo de vida a um intervalo. A configuração contém um conjunto de regras que se aplicam a objetos atuais e futuros no bloco. Quando um objeto atende aos critérios de uma das regras, o Cloud Storage executa automaticamente uma ação especificada no objeto. Aqui estão alguns exemplos de casos de uso:

- Faça o downgrade da classe de armazenamento de objetos com mais de 365 dias para o Coldline Storage.
- Excluir objetos criados antes de 1º de janeiro de 2013.
- Mantenha apenas as três versões mais recentes de cada objeto em um intervalo com a versão ativada.

Configuração do ciclo de vida

Cada configuração de gerenciamento do ciclo de vida contém um conjunto de regras. Ao definir uma regra, você pode especificar qualquer conjunto de condições para qualquer ação. Se você especificar várias condições em uma regra, um objeto deverá corresponder a *todas* as condições para a ação a ser executada. Se você especificar várias regras que contenham a mesma ação, a ação será executada quando um objeto corresponder à (s) condição (ões) em *qualquer uma* das regras. Cada regra deve conter apenas uma ação.

Se um único objeto estiver sujeito a várias ações, o Cloud Storage executará apenas uma das ações, e o objeto será reavaliado antes que ações adicionais sejam tomadas. Uma ação Excluir tem precedência

sobre uma ação `SetStorageClass`. Se várias ações `SetStorageClass` forem especificadas, a ação de alternar para a classe de armazenamento com o menor preço de armazenamento em repouso é escolhida.

Portanto, por exemplo, se você tiver uma regra que exclua um objeto e outra regra que altere a classe de armazenamento do objeto, mas ambas as regras usem exatamente a mesma condição, a ação de exclusão sempre ocorrerá quando a condição for atendida. Se você tiver uma regra que altera a classe do objeto para `Nearline Storage` e outra regra que altera a classe do objeto para `Coldline Storage`, mas ambas as regras usam exatamente a mesma condição, a classe do objeto sempre muda para `Coldline Storage` quando a condição é atendida.

Ações de ciclo de vida

As seguintes ações são suportadas por uma regra de ciclo de vida:

- **Apagar** : Apagar objetos ao vivo e / ou arquivados. (A *ao vivo* objeto é aquele que não é uma geração arquivado. Ver versão de objeto para detalhes). Esta ação pode ser aplicada a ambos os objetos versionados e não versionados. Em um intervalo com o controle de versão ativado, a exclusão de um objeto ativo arquiva o objeto, enquanto a exclusão de um objeto arquivado exclui o objeto permanentemente.
- **Aviso: depois** que um objeto é excluído, ele não pode ser recuperado. Tome cuidado ao configurar suas regras de ciclo de vida para que você não faça com que mais dados sejam excluídos do que você pretende. É recomendável que você teste suas regras de ciclo de vida em dados de desenvolvimento antes de aplicar a produção e observe os metadados de tempo de expiração para garantir que sua regra tenha o efeito desejado.
- **SetStorageClass** : **altera** a classe de armazenamento de objetos **ativos** e / ou arquivados. Essa ação pode ser aplicada a objetos com versão e sem versão. Esta ação suporta as seguintes transições de classe de armazenamento:
 - Classe de armazenamento original
 - Nova classe de armazenamento
 - Armazenamento multi-regional `StorageNearline`

- Armazenamento Coldline Armazenamento Regional Armazenamento de Rosto
- Armazenamento de linhas frias Armazenamento de linha de frente Armazenamento de linha
- **Nota:** A `SetStorageClass` ação incorre em encargos de operação de Classe A. Os encargos são determinados pela classe original do objeto. Por exemplo, a alteração de 1.000 objetos do Armazenamento Multi-regional para o Armazenamento do Coldline conta como 1.000 operações de Classe A e é faturada na taxa de operações de Classe A para o Armazenamento Multi-regional.

Condições do ciclo de vida

As seguintes condições são suportadas por uma regra de ciclo de vida:

- **Idade** : esta condição é satisfeita quando um objeto atinge a idade especificada (em dias). Ao especificar a `Age` condição, você está especificando um Time to Live (TTL) para objetos em um bucket com gerenciamento de ciclo de vida configurado. A hora em que a `Age` condição é considerada satisfeita é calculada adicionando o valor especificado ao tempo de criação do objeto. Por exemplo, se a hora de criação do objeto for 2017/01/10 10:00 UTC e a `Age` condição for 10 dias, a condição será satisfeita em e após 2017/01/20 10:00 UTC. Isso é verdade mesmo se o objeto for arquivado por meio da versão do objeto após sua criação.
- **CreatedBefore** : esta condição é satisfeita quando um objeto é criado antes da meia-noite da data especificada em UTC.
- **IsLive** : Se o valor for `true` , essa condição de ciclo de vida corresponderá apenas a objetos **ativos** ; se o valor for `false` , ele corresponderá apenas aos objetos arquivados. Para os propósitos dessa condição, objetos em buckets sem versão são considerados ativos.
- **MatchesStorageClass** : essa condição é satisfeita quando um objeto no bucket é armazenado como a classe de armazenamento especificada. Geralmente, se você pretende usar essa condição em objetos de armazenamento regional ou de armazenamento regional, deve incluir também `STANDARD` e

`DURABLE_REDUCED_AVAILABILITY` na condição de garantir que todos os objetos de classe de armazenamento semelhantes sejam cobertos.

- **Cuidado:** Não use `MULTI_REGIONAL` in `if` o balde está localizado em um local regional, e não use `REGIONAL` em `MatchesStorageClass` se o balde está localizado em um local multi-regional. Isso resulta em um erro quando você tenta definir uma configuração de gerenciamento de ciclo de vida.
- **NumberOfNewerVersions** : Relevante apenas para objetos com versão. Se o valor dessa condição for definido como N , um objeto satisfará a condição quando houver pelo menos N versões (incluindo a versão ao vivo) mais novas que ele. Para objetos ativos, o número de versões mais recentes é considerado 0. Para a versão arquivada mais recente, o número de versões mais recentes é 1 (ou 0 se não houver nenhum objeto ativo) e assim por diante.

Todas as condições são opcionais, mas pelo menos uma condição é necessária. Se você tentar definir uma configuração de ciclo de vida inválida, por exemplo, usando uma ação ou condição que não existe, receberá uma `400 Bad request` resposta de erro e qualquer configuração de ciclo de vida existente permanecerá em vigor.

Consulte Gerenciando Ciclos de Vida de Objeto para exemplos de uso de configurações de ciclo de vida. Para o formato geral de um arquivo de configuração de ciclo de vida, consulte a representação do recurso de bucket para JSON ou o formato de configuração do ciclo de vida para XML.

Comportamento do ciclo de vida do objeto

- O Cloud Storage inspeciona regularmente todos os objetos em um bucket para o qual o Object Lifecycle Management está configurado e executa todas as ações aplicáveis de acordo com as regras do bucket.
- **Cuidado:** o Cloud Storage realiza uma ação de forma assíncrona, de modo que pode haver um atraso entre quando as condições são atendidas e quando a ação é executada. Seus aplicativos não devem confiar em ações de ciclo de vida que ocorram dentro de

um determinado período de tempo após o cumprimento de uma condição de ciclo de vida.

- Por exemplo, se um objeto atender às condições de exclusão, o objeto pode não ser excluído imediatamente. Você continuará vendo o objeto até que a ação do ciclo de vida seja executada nele. Você não é cobrado pelo armazenamento do objeto, mas o acesso ao objeto incorre em qualquer operação aplicável e encargos de largura de banda, conforme descrito em [preços](#).
- Atualizações na configuração do seu ciclo de vida podem levar até 24 horas para entrar em vigor. Isso significa que, quando você altera a configuração do ciclo de vida, o Object Lifecycle Management ainda pode executar ações com base na configuração antiga por até 24 horas.
- Por exemplo, se você alterar uma `Age` condição de 10 dias para 20 dias, um objeto com 11 dias de idade poderá ser excluído pelo Object Lifecycle Management até 24 horas depois, devido aos critérios da configuração antiga.
- Uma `Delete` ação de ciclo de vida do objeto não terá efeito sobre um objeto enquanto o objeto tiver uma retenção de objeto colocada nele ou uma política de retenção que ainda não tenha sido atendida. Qualquer `Delete` ação que ocorra enquanto um objeto tiver uma restrição de política de retenção ou retenção ocorrerá depois que as restrições não se aplicarem ao objeto.
- Uma `SetStorageClass` ação do ciclo de vida do objeto não é afetada pela existência de retenções de objeto ou políticas de retenção.

Exclusão antecipada dos objetos Nearline Storage e Coldline Storage

O Object Lifecycle Management não reescreve um objeto ao alterar sua classe de armazenamento. Isso significa que quando um objeto é transferido para Nearline Storage ou Coldline Storage usando o `SetStorageClass` recurso, qualquer exclusão antecipada subsequente e cobranças associadas são baseadas no tempo de criação original do objeto, independentemente de quando a classe de armazenamento foi alterada.

Por exemplo, digamos que você faça upload de um objeto como Armazenamento regional e, 20 dias depois, a configuração do seu ciclo de vida altera a classe de armazenamento do objeto para Nearline Storage. Se você excluir o objeto imediatamente, haverá uma cobrança antecipada de 10 dias, já que o objeto existia por 20 dias. Se você excluir o objeto 10 dias após alterar sua classe de armazenamento para Nearline Storage, não haverá cobrança antecipada de exclusão, pois o objeto existiu por 30 dias.

Em comparação, digamos que você faça o upload de um objeto como Armazenamento Regional e, 20 dias depois, altere a classe de armazenamento usando uma reconfiguração (novamente para o Nearline Storage). Se você excluir o objeto imediatamente depois, incorreria em uma cobrança antecipada de 30 dias, já que o tempo de regravação se torna o novo horário de criação. Da mesma forma, se você esperasse 10 dias após a reconfiguração para excluir o objeto, incorreria em uma cobrança antecipada de 20 dias.

Metadados de tempo de expiração

Se uma `Delete` ação for especificada para um bucket com a `Age` condição (e sem `NumberOfNewerVersions` condição), alguns objetos poderão ser marcados com metadados de tempo de expiração. O tempo de expiração de um objeto indica a hora em que o objeto se torna (ou tornou-se) elegível para exclusão pelo Object Lifecycle Management. O prazo de expiração pode mudar à medida que a configuração do ciclo de vida ou a política de retenção do bucket muda.

Observe que a ausência de metadados de tempo de expiração não significa necessariamente que o objeto não será excluído, mas sim que não há informações suficientes disponíveis para determinar quando ou se ele será excluído. Por exemplo, se a hora de criação do objeto for 2013/01/10 10:00 UTC e a `Age` condição for definida como 10 dias, o tempo de expiração do objeto será 2013/01/20 10:00 UTC. No entanto, o tempo de expiração não estará disponível para o objeto se:

- A `NumberOfNewerVersions` condição também é especificada. Nesse caso, versões mais antigas do objeto ainda podem ser excluídas se novas versões forem adicionadas.
- A `CreatedBefore` condição também é especificada e definida como "2013-01-01" porque o objeto não satisfaz essa condição.

- O objeto está em espera porque o Cloud Storage não pode saber quando a retenção será removida.

Você não é cobrado pelo armazenamento após o tempo de expiração do objeto, mesmo que o objeto não seja excluído imediatamente. Você pode continuar a acessar o objeto antes que ele seja excluído e é responsável por outras cobranças (solicitação, largura de banda da rede). Se o tempo de expiração não estiver disponível para um objeto, o objeto será cobrado pelo armazenamento até o momento em que for excluído.

Você pode acessar o tempo de expiração de um objeto em seus metadados, se estiver disponível. A API REST retorna o tempo de expiração do objeto no `x-goog-expiration` cabeçalho de resposta.

Ao trabalhar com tempos de expiração, tenha em mente o seguinte:

- Se o bucket tiver uma política de retenção, o tempo de expiração será o mais tardar da condição de idade do Object Lifecycle Management e a hora em que o objeto atender ao período de retenção especificado pela política de retenção.
- Se houver vários tempos de expiração conflitantes aplicáveis a um objeto devido a diferentes regras de gerenciamento de ciclo de vida, o tempo de expiração aplicável mais antigo será usado.

Opções para rastrear ações do ciclo de vida

Para rastrear as ações de gerenciamento do ciclo de vida que o Cloud Storage executa, use uma das seguintes opções:

- Use os registros de acesso do Cloud Storage. Esse recurso registra a ação e quem executou a ação. Um valor `GCS Lifecycle Management` no `cs_user_agent` campo da entrada de log indica que a ação foi executada pelo Cloud Storage de acordo com uma configuração de ciclo de vida.
- Ative as notificações do Cloud Pub / Sub para o Cloud Storage para seu intervalo. Esse recurso envia notificações para um Cloud Pub / Subtopic de sua escolha quando ocorrem ações especificadas. Observe que esse recurso não registra quem executou as ações

Definir ciclo de vida do balde

Por padrão, os buckets não possuem o gerenciamento do ciclo de vida ativado. Para definir ou modificar a configuração do ciclo de vida de um bucket existente, você faz uma solicitação PUT com escopo para o bucket e usa o `lifecycle` parâmetro da string de consulta. Você deve incluir um documento XML no corpo da solicitação que contém a configuração do ciclo de vida. Observe que você não pode definir a configuração do ciclo de vida em um novo bloco que está criando.

Você deve ter `FULL_CONTROL` permissão para definir ou modificar a configuração do ciclo de vida de um intervalo existente. Além disso, você deve estar autenticado para usar o método PUT Bucket.

Parâmetros de string de consulta

ParameterDescriptionRequired `lifecycle` Você usa isso para alterar a configuração do ciclo de vida em um bloco existente. Você deve fornecer o documento de configuração do ciclo de vida no corpo da solicitação. Não

Solicitar Cabeçalhos

Veja [os cabeçalhos de pedidos comuns](#).

Solicitar elementos do corpo

Os elementos do corpo da solicitação a seguir são aplicáveis apenas se você usar o `lifecycle` parâmetro da cadeia de consulta para especificar a configuração do ciclo de vida para um bloco existente.

ElementDescription `LifecycleConfiguration` Define as políticas de gerenciamento de ciclo de vida para o intervalo, que contém 0 ou mais (até 100) regras. Use um elemento vazio (por exemplo, `<LifecycleConfiguration/>`) para desativar o gerenciamento do ciclo de vida do bucket. `Rule` Define uma regra de gerenciamento de ciclo de vida, que é feita de uma ação e a condição sob a qual a ação será executada. `Action` Define a ação a ser executada, que deve conter um e apenas um elemento de ação específico. Requeridos.

Delete Elemento de ação para excluir objetos no intervalo.

SetStorageClass Elemento de ação para alterar a classe de armazenamento de um objeto. **Condition** Define a condição sob a qual a ação será executada, que deve conter pelo menos um elemento de condição específico. Requeridos. **Age** Elemento de condição que corresponde a objetos com a idade especificada (em dias). **CreatedBefore** Elemento de condição que corresponde a objetos criados antes da meia-noite da data especificada em UTC. O valor é uma string de data ISO sem fuso horário, por exemplo, "2013-01-15".

IsLive Elemento de condição relevante apenas para objetos com versão. Corresponde objetos ativos se o valor for `true` ou objetos arquivados, se o valor for `false`. **MatchesStorageClass** Elemento de condição que corresponde a objetos da classe de armazenamento especificada. Essa condição pode ser adicionada várias vezes à mesma regra para abranger mais de uma classe de armazenamento.

NumberOfNewerVersions Elemento de condição relevante apenas para objetos com versão. Se o valor for N, a condição é satisfeita quando houver pelo menos N versões (incluindo a versão ao vivo) mais novas que esta versão do objeto.

Solicitar sintaxe

```
PUT /? Ciclo de vida HTTP / 1.1
Host: <bucket> .storage.googleapis.com
Data: <data e hora da solicitação>
Content-Length: <tamanho do corpo da solicitação>
Tipo de conteúdo: <tipo MIME do corpo>
Autorização: <cadeia de autenticação>
```

```
<xml_document_defining_lifecycle_configuration>
```

Cabeçalhos de resposta

A solicitação pode retornar vários cabeçalhos de resposta, dependendo dos cabeçalhos de solicitação usados.

Elementos do corpo de resposta

A resposta não inclui um documento XML no corpo da resposta.

Exemplo

A amostra a seguir ativa o gerenciamento do ciclo de vida de um bucket chamado example.com. Para mais exemplos, [veja a página Gerenciando Lifecycles](#).

Pedido

```
PUT /example.com?lifecycle HTTP / 1.1
Anfitrião: storage.googleapis.com
Content-Length: 220
Autorização: Bearer ya29.AHES6ZRVmB7fkLtd1XTmq6mo0S1wqZZi3-
Lh_s-6Uw7p8vtgSwg

<? xml version = "1.0"?>
<LifecycleConfiguration>
  <Regra>
    <Action>
      <Delete />
    </ Action>
    <Condição>
      <Age> 30 </ Age>
    </ Condition>
  </ Rule>
</ LifecycleConfiguration>
```

Resposta

```
HTTP / 1.1 200 OK
Data: Thu, 12 Mar 2012 03:38:42 GMT
Expires: Seg, 01 Jan 1990 00:00:00 GMT
Cache-Control: não-cache, no-store, deve-revalidar
Content-Length: 0 tipo de
conteúdo: text / html
```