



GOOGLE CLOUD PLATFORM

# Best practices for optimizing your cloud costs

**Justin Lerma**

Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

## RELATED ARTICLE

Cloud cost optimization: principles for lasting success

READ ARTICLE ➔

## 1. Get to know billing and cost management tools

Due to the on-demand, variable nature of cloud, costs have a way of creeping up on you if you're not monitoring them closely. Once you understand your costs, you can start to put controls in place and optimize your spending. To help with this, Google Cloud provides a robust set of no-cost [billing and cost management tools](#) that can give you the visibility and insights you need to keep up with your cloud deployment.

At a high level, learn to look for things like *"which projects cost the most, and why?"* To start, [organize and structure your costs](#) in relation to your business needs. Then, drill down into the services using [Billing reports](#) to get an at-a-glance view of your costs. You should also learn how to attribute costs back to departments or teams using [labels](#) and build your own [custom dashboards](#) for more granular cost views. You can also use [quotas, budgets, and alerts](#) to closely monitor your current cost trends and forecast them over time, to reduce the risk of overspending.

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

your / aim is to get rid of resources that are no longer being used. Think about those

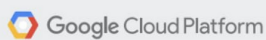


Always tread carefully when deleting a VM, though. Before deleting a resource, ask yourself, *“what potential impact will deleting this resource have and how can I recreate it, if necessary?”* [Deleting instances](#) gets rid of the underlying disk(s) and all of its data. One best practice is to take a [snapshot](#) of the instance before deleting it. Alternatively, you can choose to simply [stop](#) the VM, which terminates the instance, but keeps resources like disks or IP addresses until you detach or delete them.

For more info, read the [recommender documentation](#). And stay tuned as we add more usage-based recommenders to the portfolio.

**Schedule VMs to auto start and stop:** The benefit of a platform like [Compute Engine](#) is that you only pay for the compute resources that you use. Production systems tend to run 24/7; however, VMs in development, test or personal environments tend to only be used during business hours, and turning them off can save you a lot of money! For example, a VM that runs for 10 hours per day, Monday through Friday costs 75% less to run per month compared to leaving it running.

To get started, here’s a [serverless solution](#) that we developed to help you automate and manage automated VM shutdown at scale.

Production  
VMs

Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

...new recommendations for better advice and training. To help, our [lightweight recommendations](#) can...

7 instances could be resized to save you up to an estimated \$355 per month and increase performance. [Learn more](#)

<input type="checkbox"/>	Name	Zone	Machine type	Recommendation ^	In use by	Internal IP	External IP	Connect
<input type="checkbox"/>	instance-4	europa-west1-c	8 vCPUs, 30 GB	Save \$145 / mo		10.132.0.5	104.155.90.140	SSH
<input type="checkbox"/>	instance-10	europa-west1-c	4 vCPUs, 26 GB	Save \$101 / mo		10.132.0.11	104.155.111.113	SSH
<input type="checkbox"/>	instance-2	europa-west1-c	4 vCPUs, 15 GB	Save \$76 / mo		10.132.0.3	104.155.88.201	SSH
<input type="checkbox"/>	instance-7	europa-west1-c	1 vCPU, 3.75 GB	Save \$16 / mo		10.132.0.8	104.155.55.101	SSH
<input type="checkbox"/>	instance-9	europa-west1-c	1 vCPU, 3.75 GB	Save \$16 / mo		10.132.0.10	104.155.58.90	SSH
<input type="checkbox"/>	instance-11	europa-west1-c	1 vCPU, 0.6 GB	Increase perf.		10.132.0.2	104.155.108.6	SSH
<input type="checkbox"/>	instance-12	asia-east1-a	1 vCPU, 0.6 GB	Increase perf.		10.140.0.2	104.199.187.28	SSH
<input type="checkbox"/>	instance-3	europa-west1-c	1 vCPU, 0.6 GB			10.132.0.4	104.155.100.41	SSH

For organizations that use infrastructure as code to manage their environments, check out [this guide](#), which will show you how to deploy VM rightsizing recommendations at scale.

**Leverage preemptible VMs:** [Preemptible VMs](#) are highly affordable compute instances that live up to 24 hours and that are up to 80% cheaper than regular instances. Preemptible VMs are a great fit for fault tolerant workloads such as big data, genomics, media transcoding, financial modelling and simulation. You can also use a mix of regular and preemptible instances to finish compute-intensive workloads faster and cost-effectively, by setting up a specialized [managed instance group](#).

But why limit preemptible VMs to a Compute Engine environment? Did you know [GPUs](#),

 Find an article...

[Latest stories](#)
[Products](#)
[Topics](#)
[About](#)
[RSS Feed](#)

**Storage classes:** [Cloud Storage](#) offers a variety of [storage classes](#)—standard, nearline, coldline and archival, all with [varying costs](#) and their own best-fit [use cases](#). If you only use the standard class, it might be time to take a look at your workloads and reevaluate how frequently your data is being accessed. In our experience, many companies use standard class storage for archival purposes, and could reduce their spend by taking advantage of [nearline](#) or [coldline](#) class storage. And in some cases, if you are holding onto objects for cold-storage use cases like legal discovery, the new [archival class](#) storage might offer even more savings.

**Lifecycle policies:** Not only can you save money by using different storage classes, but you can make it happen automatically with object lifecycle management. By configuring a [lifecycle policy](#), you can programmatically set an object to adjust its storage class based on a [set of conditions](#)—or even delete it entirely if it's no longer needed. For example, imagine you and your team analyze data within the first month it's created; beyond that, you only need it for regulatory purposes. In that case, simply set a policy that adjusts your storage to coldline or archive after your object reaches 31 days.

**Deduplication:** Another common source of waste in storage environments is duplicate data. Of course, there are times when it's necessary. For instance, you may want to duplicate a dataset across multiple geographic regions so that local teams can access it quickly. However, in our experience working with customers, a lot of duplicate data is the result of lax version control, and the resulting duplicates can be cumbersome and expensive to manage.



Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)

there are several features you can use to keep your data safe without wasting space unnecessarily.

## 4. Tune your data warehouse

Organizations of all sizes look to BigQuery for a modern approach to data analytics. However, some configurations are more expensive than others. Let's do a quick check of your BigQuery environment and set up some guardrails to help you keep costs down.

**Enforce controls:** The last thing you need is a long query to run forever and rack up costs. To limit query costs, use the [maximum bytes billed setting](#). Going above the limit will cause the query to fail, but you also won't get charged for it, as shown below.

Maximum bytes billed <sup>?</sup>

Resource management

Job priority <sup>?</sup>

☒ Interactive

☐ Batch

Cache preference <sup>?</sup>

☒ Use cached results

Additional settings

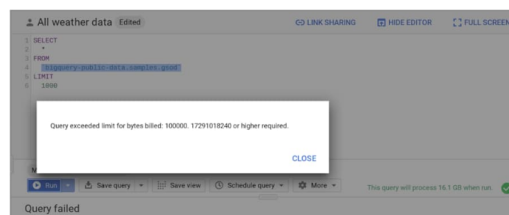
SQL dialect <sup>?</sup>


☒ Standard

☐ Legacy

Processing location <sup>?</sup>

Auto-select ▾



 Find an article...

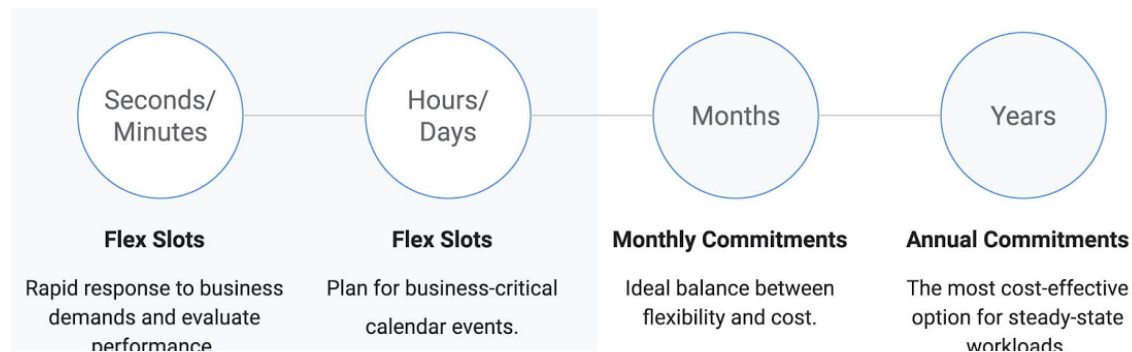
[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)



BigQuery costs, the first thing to do is check your bill and see if you are being charged for streaming inserts. And if you are, ask yourself, *“Do I need data to be immediately available (seconds instead of hours) in BigQuery?”* and *“Am I using this data for any real-time use case once the data is available in BigQuery?”* If the answer to either of these questions is no, then we recommend you to switch to batch loading data, which is free.

**Use Flex Slots:** By default, BigQuery charges you variable [on-demand pricing](#) based on bytes processed by your queries. If you are a high-volume customer with stable workloads, you may find it more cost effective to switch from on-demand to [flat rate pricing](#), which gives you an ability to process unlimited bytes for a fixed predictable cost.

Given rapidly changing business requirements, we recently introduced [Flex Slots](#), a new way to purchase BigQuery slots for duration as short as 60 seconds, on top of monthly and annual flat-rate commitments. With this combination of [on-demand and flat-rate pricing](#), you can respond quickly and cost-effectively to changing demand for analytics.



Find an article...

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)



and intercontinental [network egress cost](#). Checkout [this brief video](#) for an overview of Network Intelligence Center and Network Topology.

**Network Service Tiers:** Google Cloud lets you choose between two network service tiers: [premium and standard](#). For excellent performance around the globe, you can choose premium tier, which continues to be our tier of choice. Standard tier offers a lower performance, but may be a suitable alternative for some cost-sensitive workloads.

**Cloud Logging:** You may not know it, but you do have control over network traffic visibility by filtering out logs that you no longer need. Check out some [common examples](#) of logs that you can safely exclude. The same applies to [Data Access audit logs](#), which can be quite large and incur additional costs. For example you probably don't need to log them for development projects. For VPC Flow Logs and Cloud Load Balancing, you can also enable [sampling](#), which can dramatically reduce the volume of log traffic being written into the [database](#). You can set this from 1.0 (100% log entries are kept) to 0.0 (0%, no logs are kept). For troubleshooting or custom use cases, you can always choose to collect telemetry for a particular VPC network or subnet or drill down further to monitor a specific VM Instance or virtual interface.

## Want more?

Whether you're an early-stage startup or a large enterprise with a global footprint, everyone wants to be smart with their money right now. Following the tips in this blog

[Latest stories](#)[Products](#)[Topics](#)[About](#)[RSS Feed](#)





Get started

Blog

Menu ▾

Follow Us



Google

Privacy

Terms

About Google

Google Cloud Products

Language ▾



Help