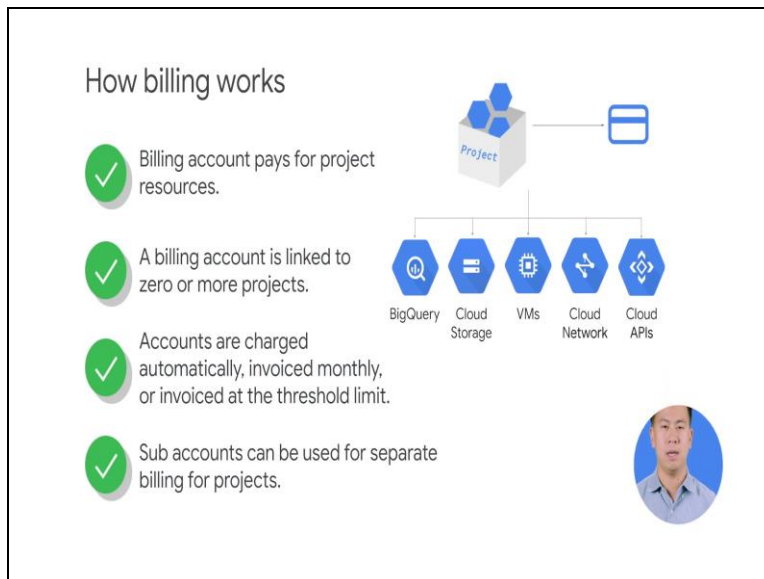


Google Cloud Computing Foundation Course
Jimmy Iran
SMB Growth Program Manager
Google Cloud

Lecture-9
Billing in GCP

(Refer Slide Time: 00:06)



In the next topic you will learn how billing works in GCP. Billing, it is no fun but a fact of life. Let us learn more about it billing and GCP is set up at the GCP project level when you define a GCP project you link a billing account to it. This billing account is where you will configure all your billing information including your payment option you can link your billing account to 0 or mobile projects. Projects that you do not link to any billing account can only use free GCP services.

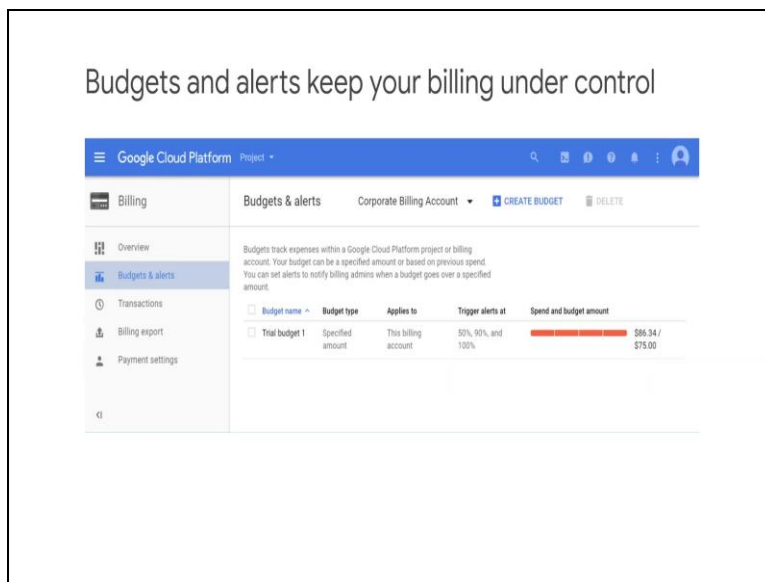
Your billing account can be charged automatically and invoiced every month or at every threshold limit. You can separate project Billings by sending up billing sub accounts. Some GCP customers who resell GCP services use sub accounts for each of their own clients.

(Refer Slide Time: 00:59)



You are probably thinking how can I make sure I do not accidentally run up a big GCP bill. GCP provides 4 tools to help. Budgets and alerts, billing export, reports and quotas. I will discuss each of these in more detail next.

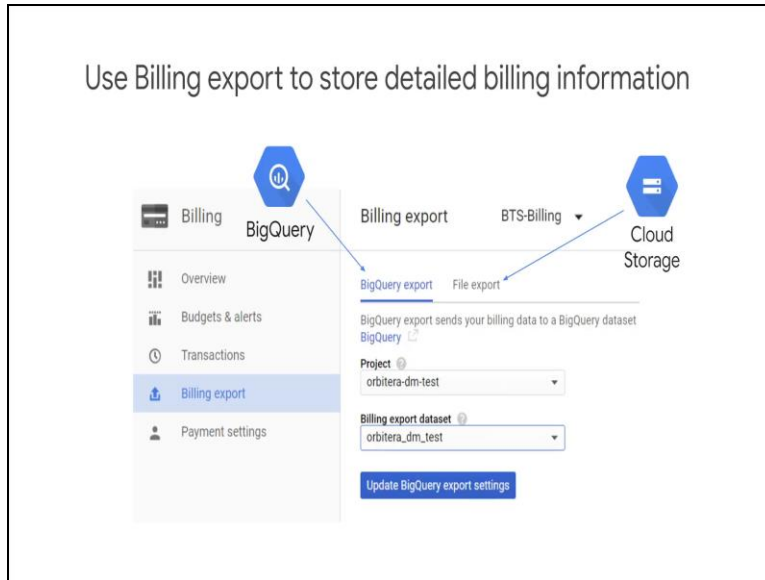
(Refer Slide Time: 01:19)



You can define budgets at the billing account level or at the project level. To be notified when costs approach your budget limit you can create an alert for example with a budget limit of \$20 and an alert set at 90% you receiving notification alert when your expenses reach \$18. You can also set up a web hook to be called in response to an alert. This Web hook can control

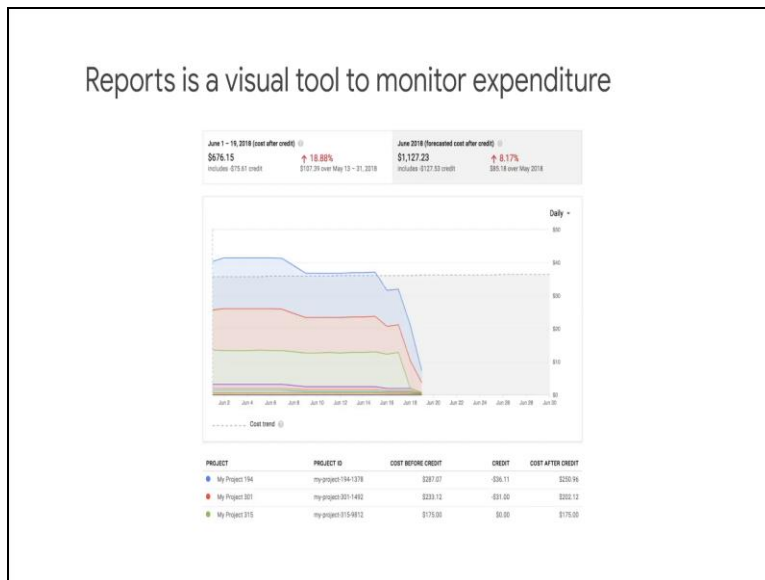
automation based on billing alerts. For example you could trigger a script to shutdown resources when a billing alert occurs.

(Refer Slide Time: 01:55)



Billing export allows you to store detailed billing information in places where it is easy to retrieve for external analysis, such as a big query data set or cloud storage bucket.

(Refer Slide Time: 02:06)



And reports is a visual tool and the console that allows you to monitor expenditure based on a project or services.

(Refer Slide Time: 02:15)



GCP also implements quotas which limit unforeseen extra billing charges. Quotas are designed to prevent the overconsumption of resources because of an error or a malicious attack, quotas apply at the level of the GCP project. There are two types of quotas rate quotas and allocation quotas. Rate quotas reset after a specific time for example by default the google kubernetes engine service implements a quota of 1,000 calls to its api from each GCB project every 100 seconds.

Allocation quotas governed a number of resources you can have in your projects. For example depending on your region the number of GPUs permitted varies by type and region. You can change quotas by requesting an increase from Google Cloud support. You can also use the console to request a quota change. GCP quotas also protect the community of GCP users by reducing the risk of unforeseen spikes in usage.