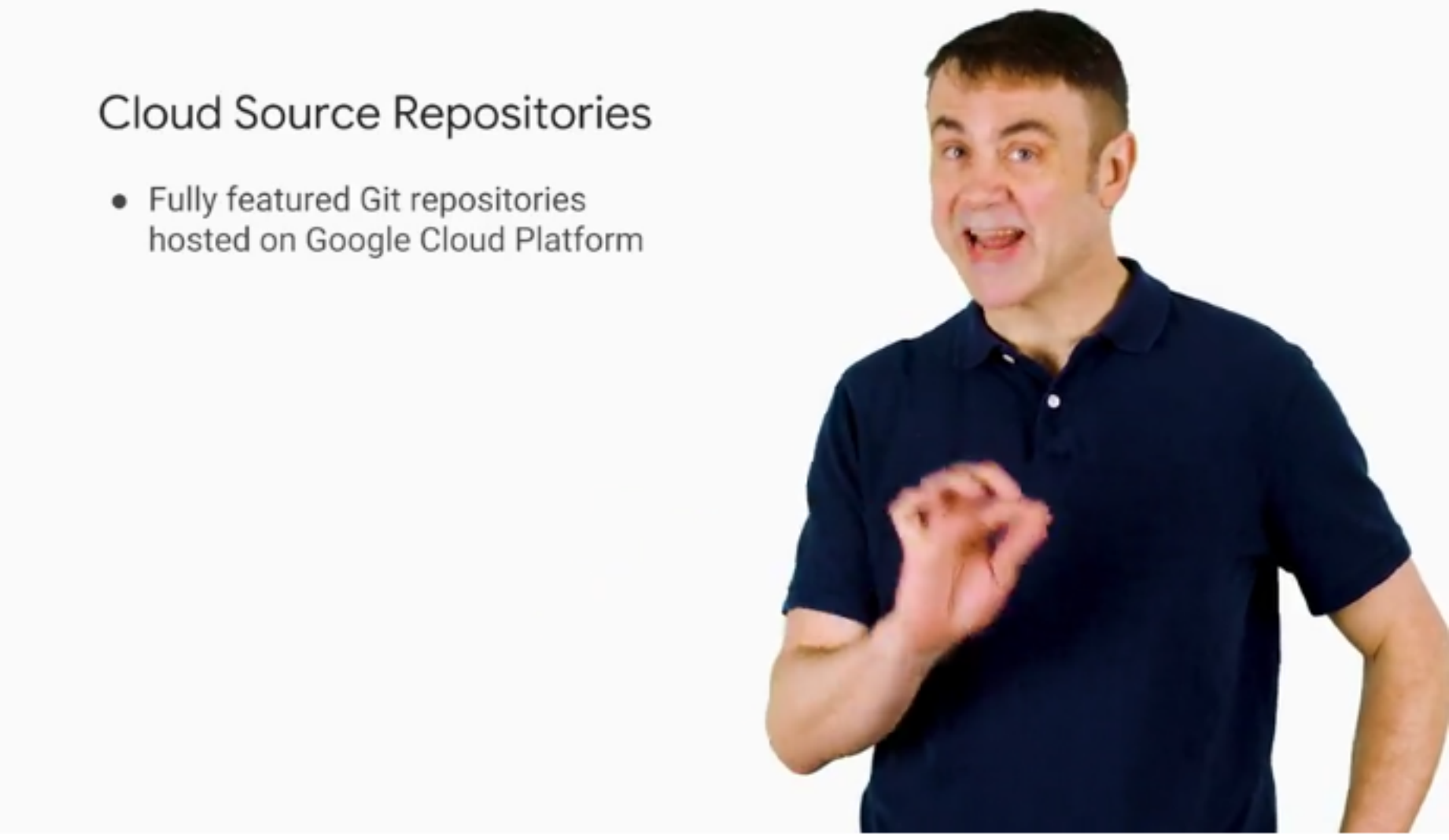
Development in the cloud

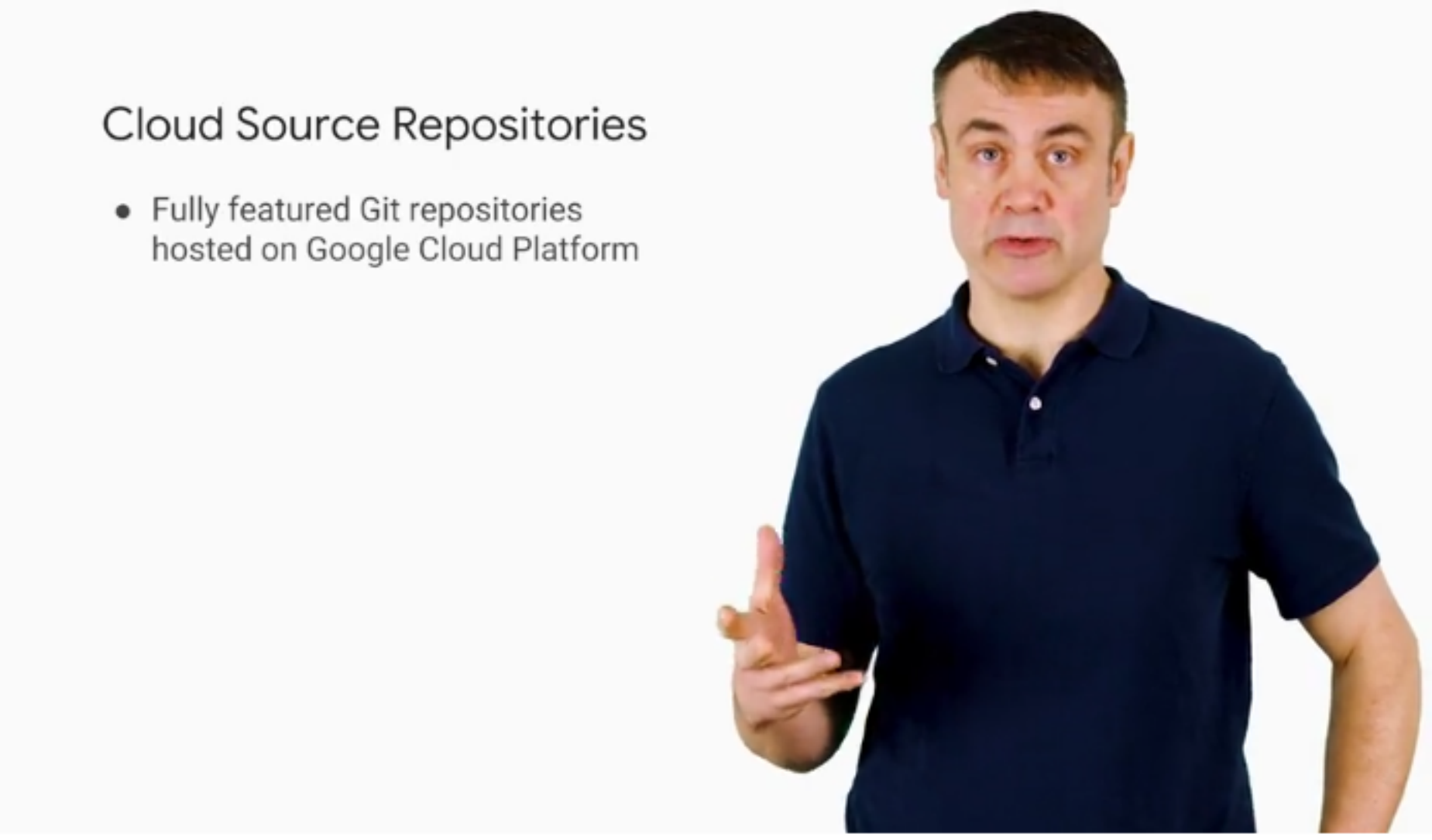


Lots of GCP customers use Git to store and manage their source code trees. That means, running their own Git instances or using a hosted Git provider. Running your own is great because you have total control. Using a hosted Git provider is great because it's less work.

What if there were a third way? Maybe a way to keep code private to a GCP project and use IAM permissions to protect it, but not have to maintain the Git instance yourself. That's what Cloud Source Repositories is.

It provides Git version control to sup-
port your team's development of any

Development in the cloud



It provides Git version control to support your team's development of any application or service, including those that run on App Engine, Compute Engine, and Kubernetes Engine.

With Cloud Source Repositories, you can have any number of private Git repositories, which allows you to organize the code associated with your cloud project in whatever way works best for you.

Cloud Source Repositories also contains a source viewer so that you can browse and view repository files from within the GCP console. Many applications contain event-driven

Cloud Functions[Beta]

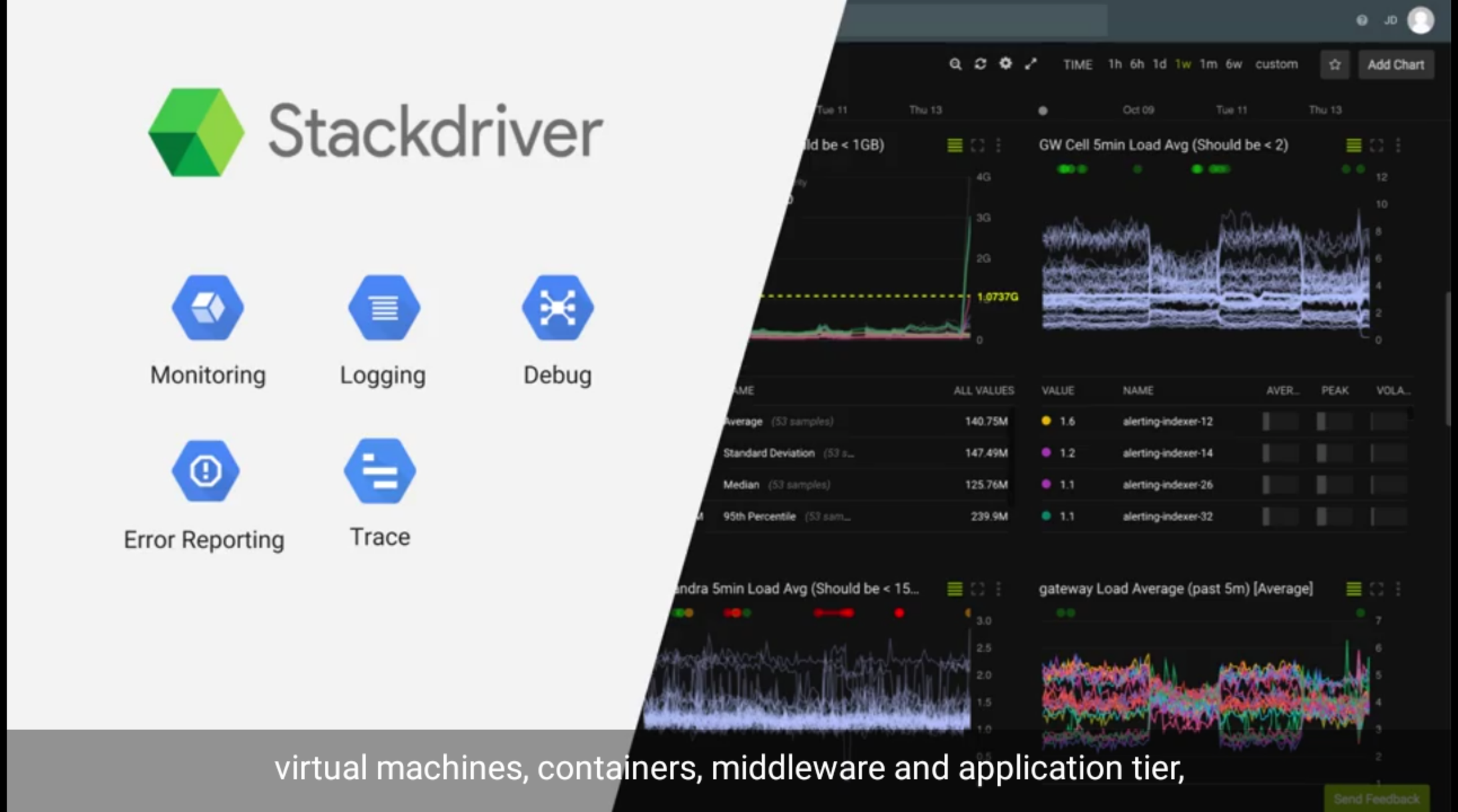
- Create single-purpose functions that respond to events without a server or runtime
- Written in Javascript; execute in managed Node.js environment on Google Cloud Platform

existing applications without having to worry about scaling.

Deployment Manager

- Provides repeatable deployments
- Create a .yaml template describing your environment and use Deployment Manager to create resources

Then, you give the template to Deployment Manager, which figures out and



virtual machines, containers, middleware and application tier,

Stackdriver offers capabilities in six areas

Monitoring

Platform, system, and application metrics

Uptime/health checks

Dashboards and alerts

Error Reporting

Error notifications

Error dashboard

Logging

Platform, system, and application logs

Log search, view, filter, and export

Log-based metrics

Debugger

Debug applications

Trace

Latency reporting and sampling

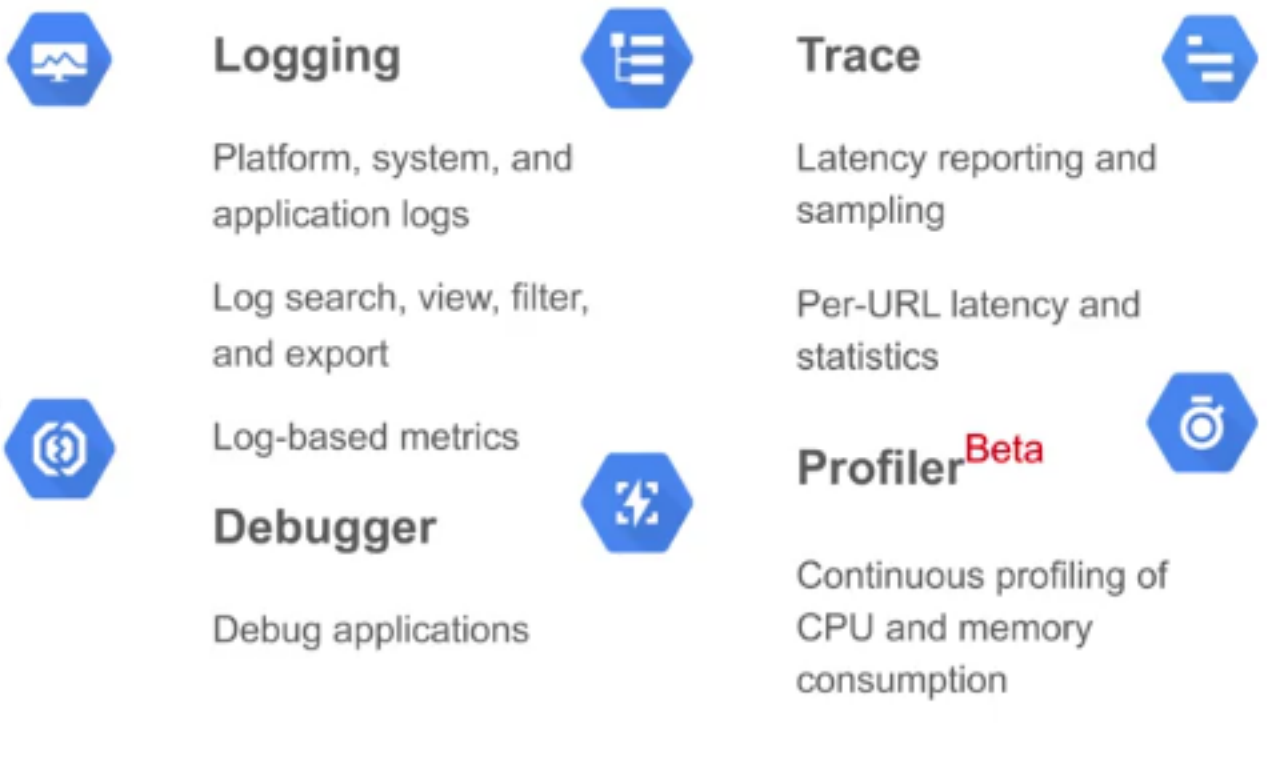
Per-URL latency and statistics

Profiler Beta

Continuous profiling of CPU and memory consumption

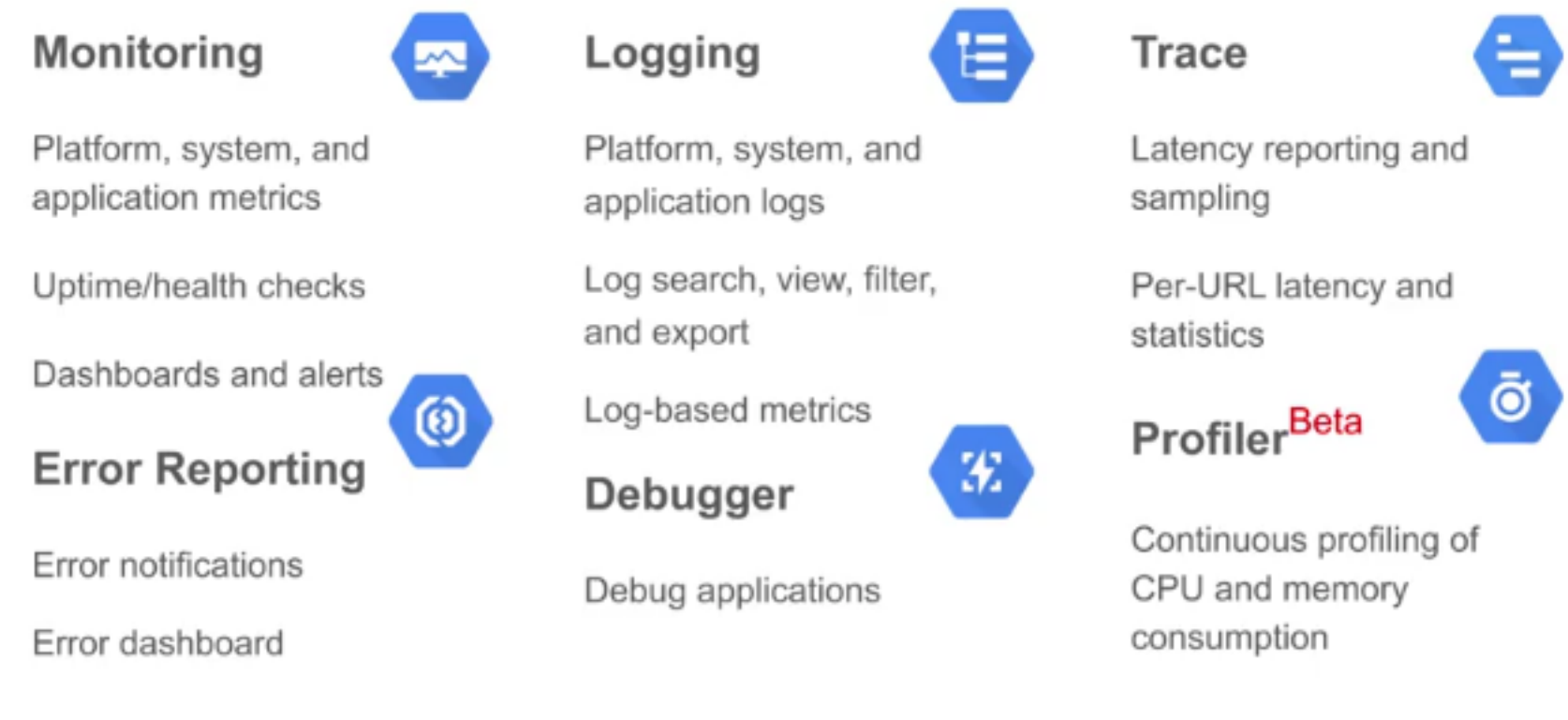
Logging, Trace, Error Reporting and Debugging.

Monitoring:...umentation



Stackdriver Logging lets you view logs from your applications and filter and search on them. Logging also lets you define metrics, based on log contents that are incorporated into dashboards and alerts. You can also export logs to BigQuery, Cloud Storage and Cloud PubSub.

Stackdriver Error Reporting tracks and groups the errors in your cloud applications. And it notifies you when new errors are detected. With Stackdriver Trace, you can sample the latency of app engine applications and report Per-URL statistics. How about



A painful way to debug an existing application is to go back into it and add lots of logging statements. Stack-driver Debugger offers a different way. It connects your applications production data to your source code.

So you can inspect the state of your application at any code location in production. That means you can view the application stage without adding logging statements. Stackdriver Debugger works best when your application source code is available, such as in Cloud Source repositories.

5/7

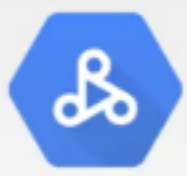
Step 3

Use opencv to process the video.

```
import cv2
import numpy as np`
url = 'Your IP here/video'
cap = cv2.VideoCapture(url)
while(True):
    ret, frame = cap.read()
    if frame is not None:
        cv2.imshow('frame',frame)
    q = cv2.waitKey(1)
    if q == ord("q"):
        break
cv2.destroyAllWindows()
```

← Google Cloud B...Data Platform

Google Cloud's big data services are fully managed and scalable



Cloud Dataproc

Managed Hadoop MapReduce, Spark, Pig, and Hive service

Cloud Dataflow

Stream and batch processing; unified and simplified pipelines

BigQuery

Analytics database; stream data at 100,000 rows per second

Cloud Pub/Sub

Scalable and flexible enterprise messaging

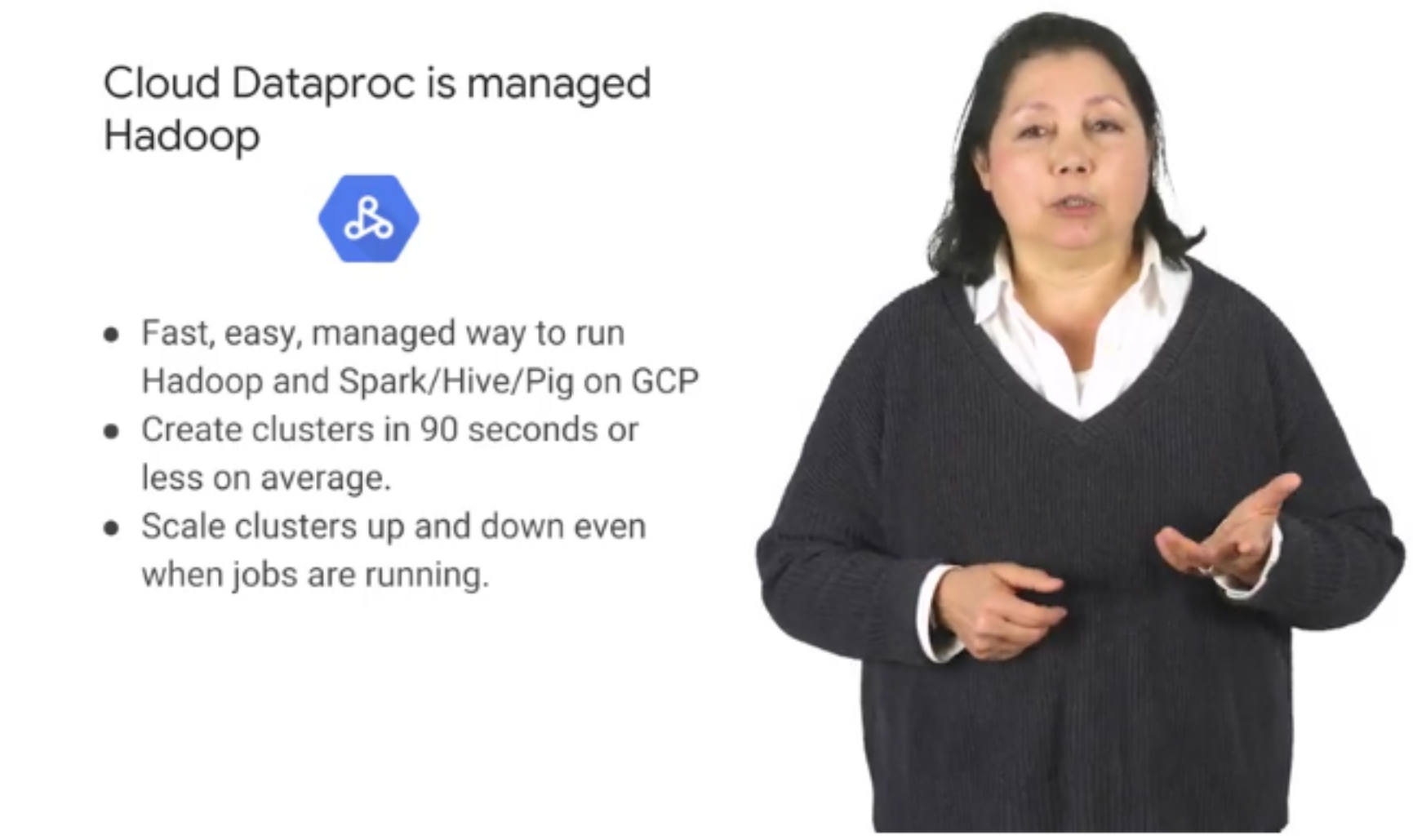
Cloud Datalab

Interactive data exploration

Serverless means you don't have to worry about provisioning Compute Instances to run your jobs. The services are fully managed, and you pay only for the resources you consume. The platform is integrated, so that GCP data services work together to help you create custom solutions.

Apache Hadoop is an open source framework for big data. It is based on the MapReduce programming model which Google invented and published.

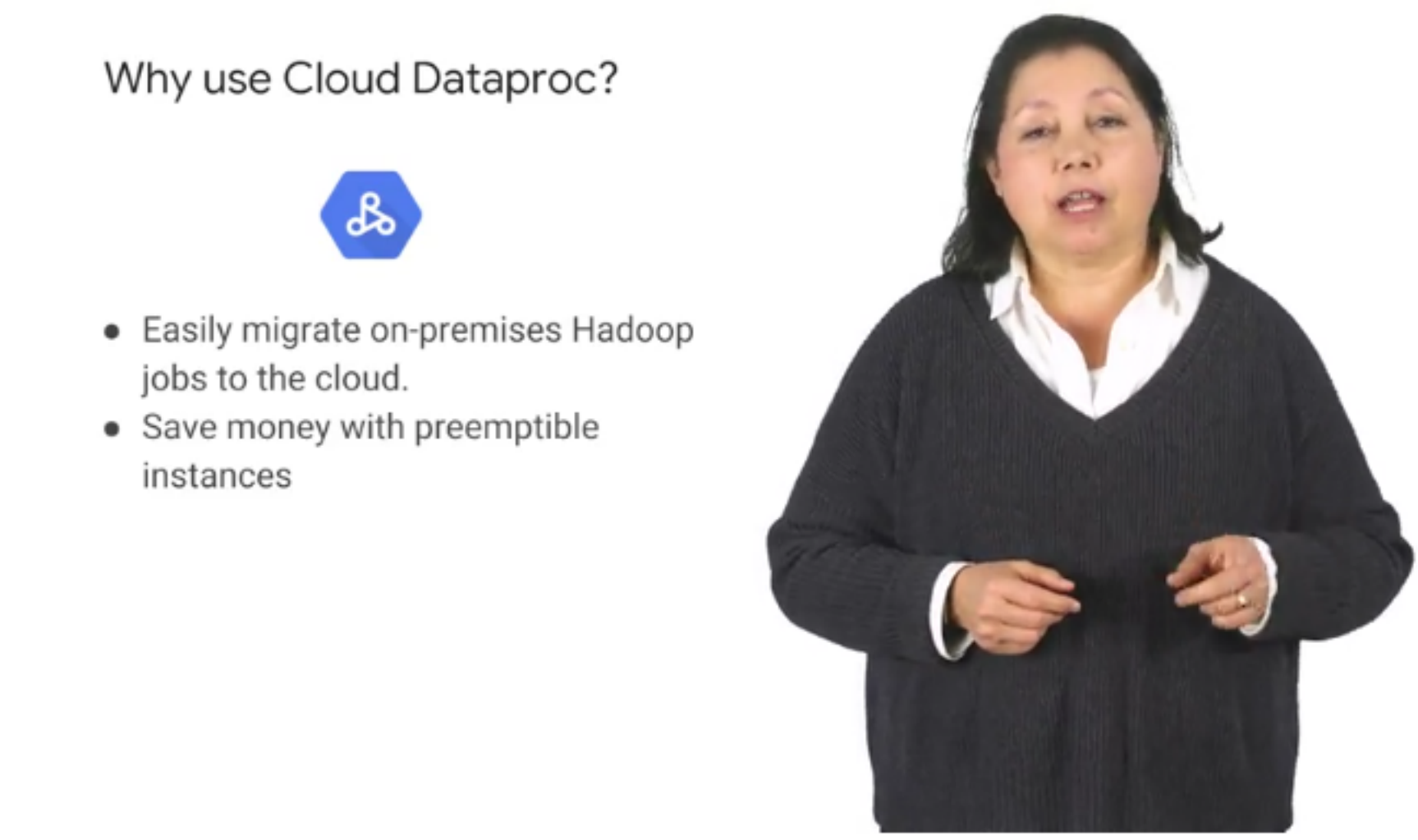
The MapReduce model is, at its simplest, means that one function, traditionally called the "Map function," runs in parallel with a massive dataset



Cloud Dataproc is a fast, easy, managed way to run Hadoop, Spark, Hive, and Pig on Google Cloud Platform. All you have to do is request a Hadoop cluster. It will be built for you in 90 seconds or less, on top of Compute Engine virtual machines whose number and type you control.

If you need more or less processing power while your cluster is running, you can scale it up or down. You can use the default configuration for the Hadoop software in your cluster or you can customize it. And you can monitor your cluster using Stackdriver.

Running on-premises, Hadoop jobs re-

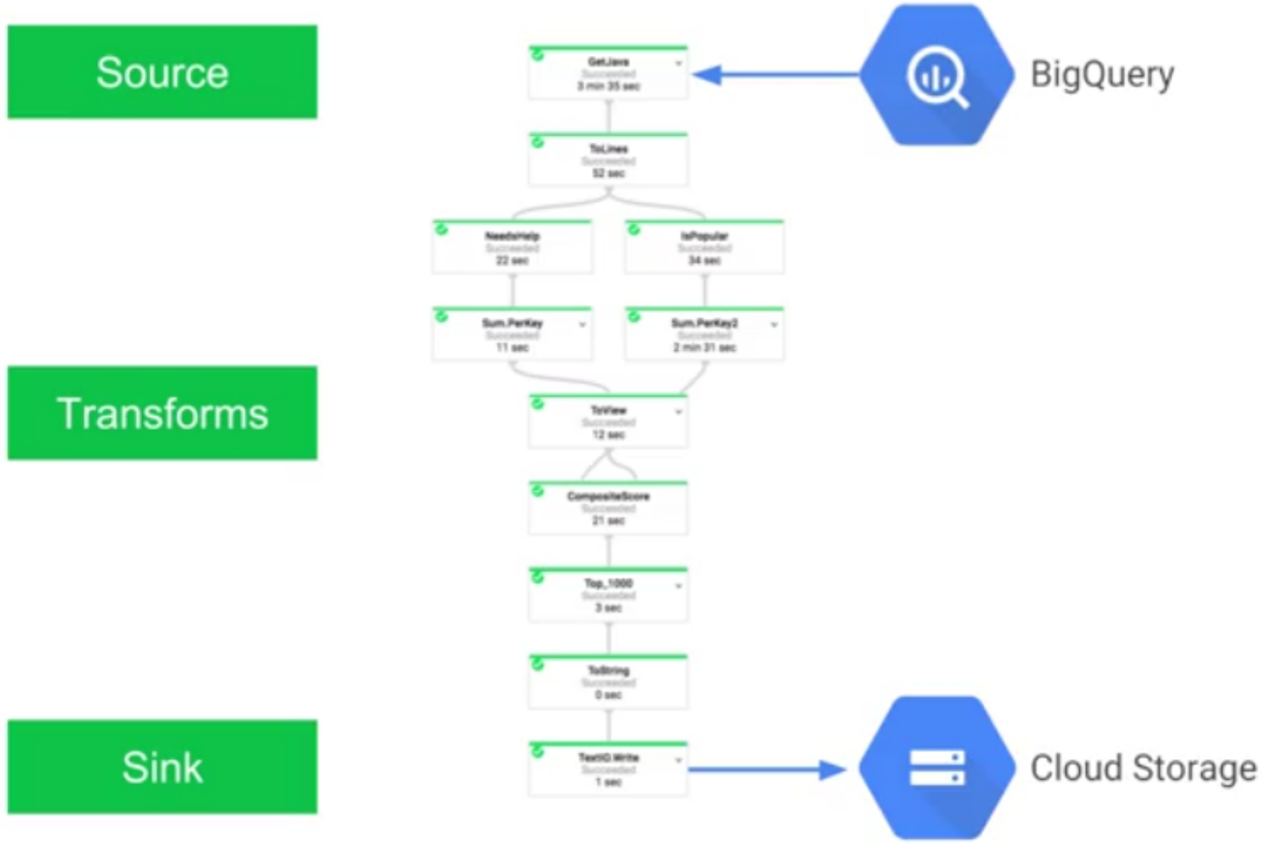


Running on-premises, Hadoop jobs requires a capital hardware investment. Running these jobs in Cloud Dataproc, allows you to only pay for hardware resources used during the life of the cluster you create. Although the rate for pricing is based on the hour, Cloud Dataproc is billed by the second.

Our Cloud Dataproc clusters are billed in one-second clock-time increments, subject to a one minute minimum billing. So, when you're done with your cluster, you can delete it, and billing stops. This is much more agile use of resources than on-premise hardware assets.

You can also save money by telling

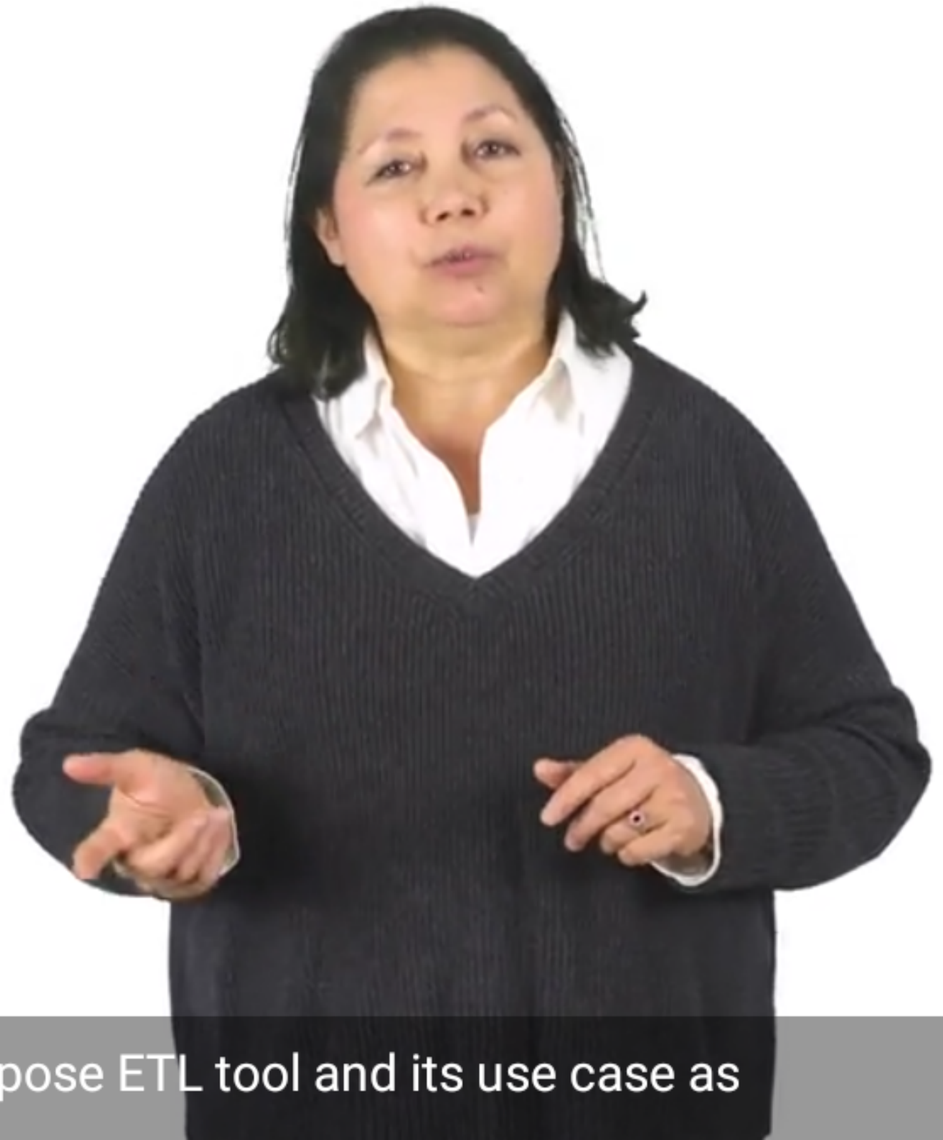
Dataflow pipelines flow data from a source through transforms



and writes its output to a cloud storage, the Sink.

Why use Cloud Dataflow?

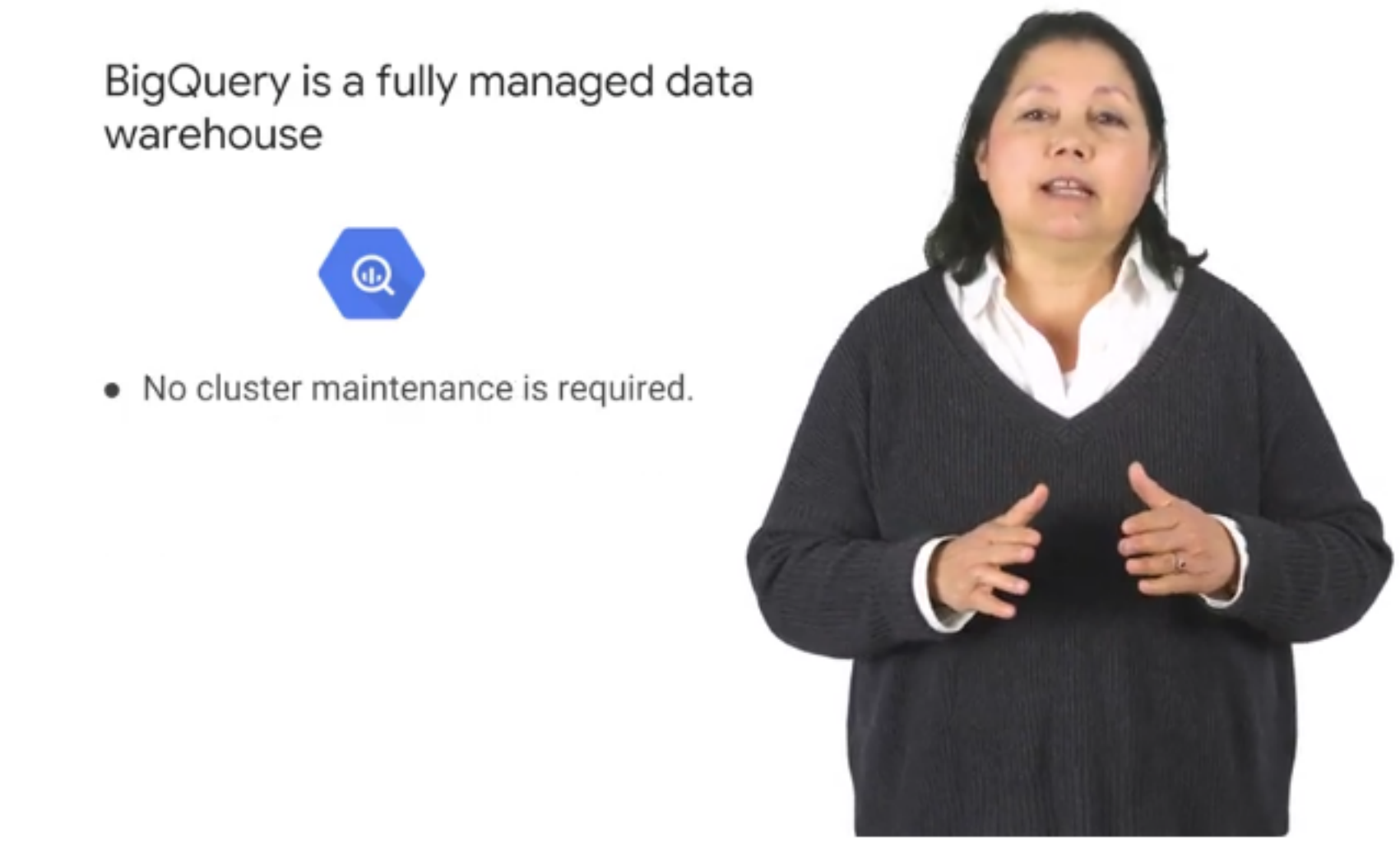
- ETL (extract/transform/load) pipelines to move, filter, enrich, shape data
- Data analysis: batch computation or continuous computation using streaming



As we've discussed, it's a general purpose ETL tool and its use case as

BigQuery

You're using Coursera offline



Suppose, instead of a dynamic pipeline, your data needs to run more in the way of exploring a vast sea of data. You want to do ad-hoc SQL queries on a massive data set. That's what BigQuery is for. It's Google's fully-managed, petabyte-scale, low-cost analytics data warehouse.

Because there's no infrastructure to manage, you can focus on analyzing data to find meaningful insights, use familiar SQL and take advantage of our pay-as-you-go model. It's easy to get data into BigQuery.

You can load it from cloud storage or cloud data store, or stream it into

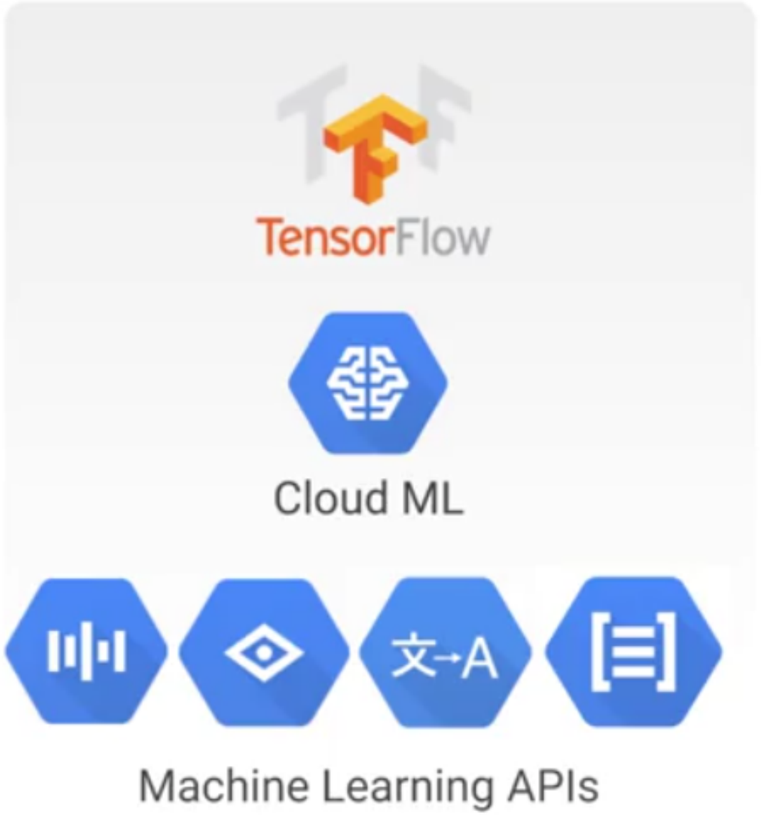
BigQuery

US and Asia locations are also available. Because BigQuery separates storage and computation, you pay for your data storage separately from queries. That means, you pay for queries only when they are actually running.

You have full control over who has access to the data stored in BigQuery, including sharing data sets with peo-ple in different projects. If you share data sets that won't impact your cost or performance. People you share with pay for their own queries, not you.

Long-term storage pricing is an

Cloud Machine Learning Platform



Open source tool to build and run neural network models

- Wide platform support: CPU or GPU; mobile, server, or cloud

Fully managed machine learning service

- Familiar notebook-based developer experience
- Optimized for Google infrastructure; integrates with BigQuery and Cloud Storage

Pre-trained machine learning models built by Google

- Speech: Stream results in real time, detects 80 languages
- Vision: Identify objects, landmarks, text, and content
- Translate: Language translation including detection
- Natural language: Structure, meaning of text

with pre-trained models and a platform to generate your own tailored models.

Why use the Cloud Machine Learning platform?

For structured data

- Classification and regression
- Recommendation
- Anomaly detection

For unstructured data

- Image and video analytics
- Text analytics

Generally, they fall into two categories, depending

Why use Cloud Pub/Sub?

- Building block for data ingestion in Dataflow, Internet of Things (IoT), Marketing Analytics
- Foundation for Dataflow streaming
- Push notifications for cloud-based applications
- Connect applications across Google Cloud Platform (push/pull between Compute Engine and App Engine)