An introduction to the Google Cloud Platform

a.k.a. Jamie's first public talk as a Googler

Jamie Kinney | Google Cloud Solutions Architect
@jamiekinney | jamiekinney@google.com

Google Cloud

Jamie Kinney

- Solutions Architect, Google Cloud Platform
- “Noogler” with 10+ years of cloud computing experience and a passion for Scientific Computing
- Formerly product manager for AWS Batch and HPC

Rob Simpson

- Product Manager, Android Search
- Creator of the first .Astronomy
- Formerly of Zooniverse at the University of Oxford

Google Cloud

Why are Googlers attending .Astronomy X?

Listen & Learn

Educate

Help

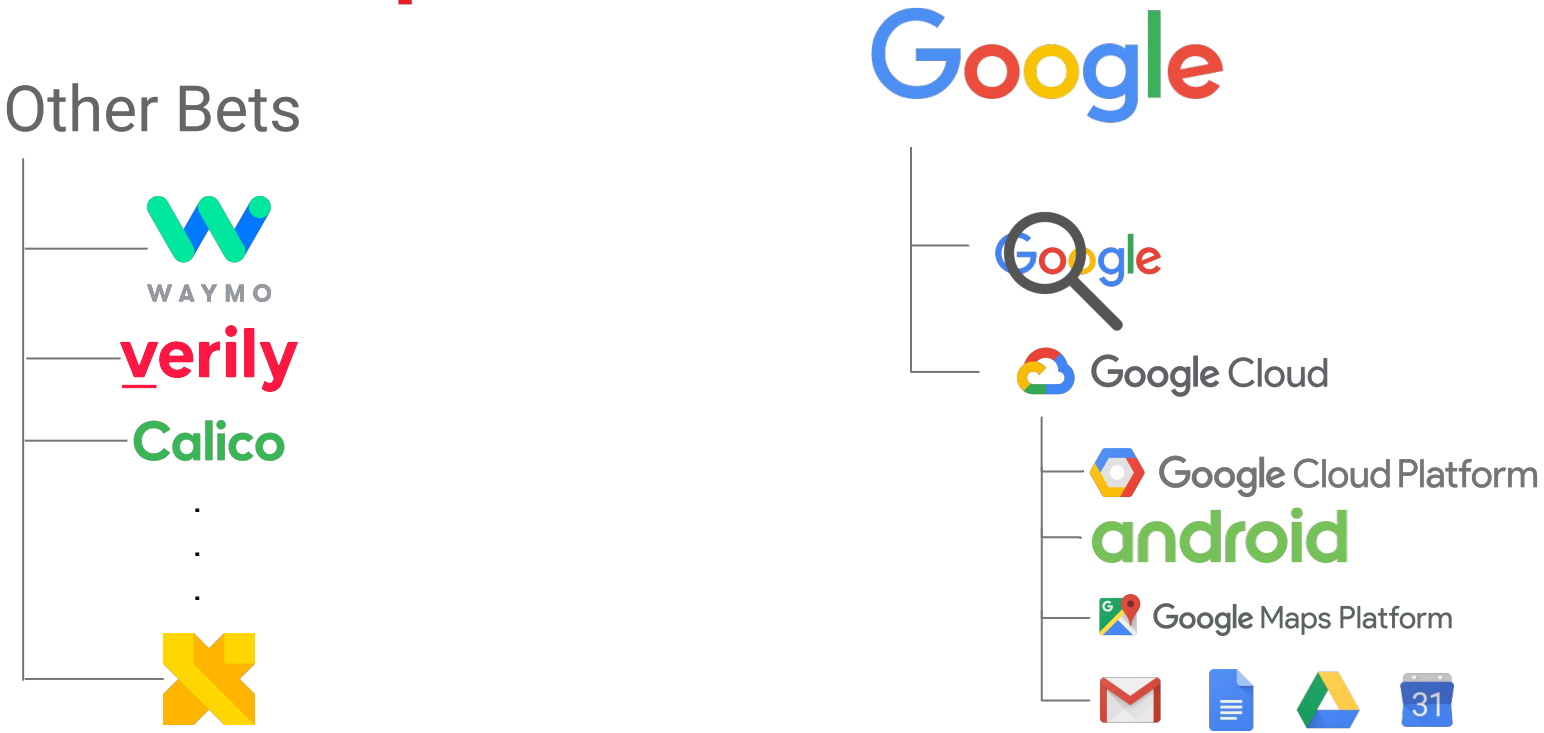
Google Cloud

Agenda

- Alphabet, Google, & GCP
- How are researchers using Google?
- Overivew of Relevant Services
- Things to try at .Astronomy X
- Questions?

Alphabet, Google, & GCP

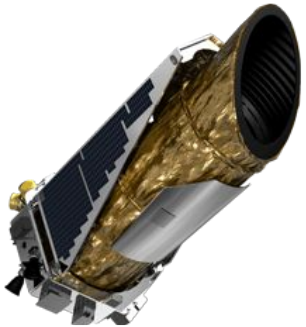
Alphabet

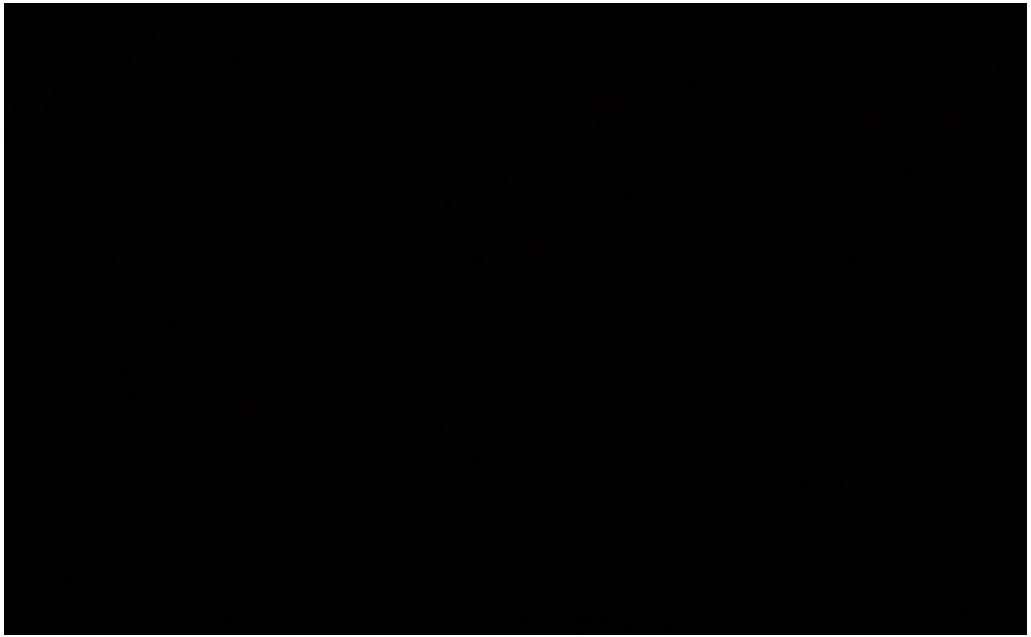

Other Bets
WAYMO
verily
Calico
Google
Google
Google Cloud
Google Cloud Platform
android
Google Maps Platform
31

Examples of Scientific Research on GCP

Astronet: A Neural Network for Identify Exoplanets in Light Curves

TensorFlow code for building neural network classification models which operate on Kepler `.fits` files to identify exoplanets in light curves





Soure code: https://github.com/tensorflow/models/tree/master/research/astronet

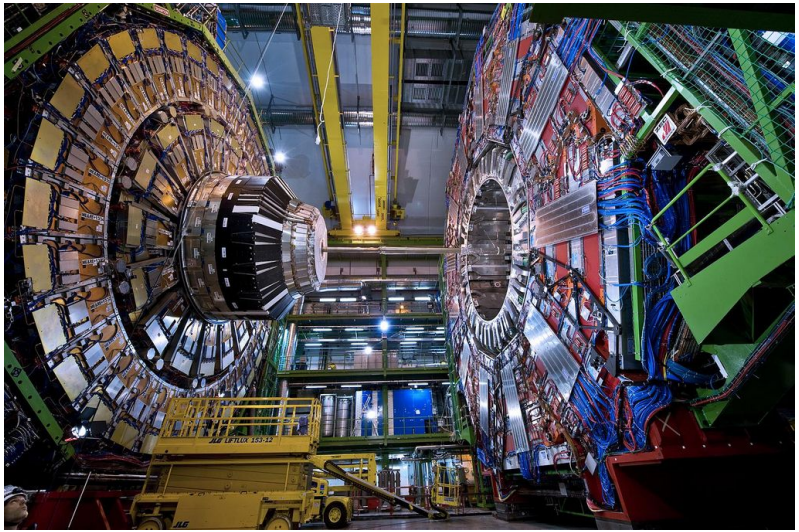
Research paper: http://adsabs.harvard.edu/abs/2018AJ....155...94S

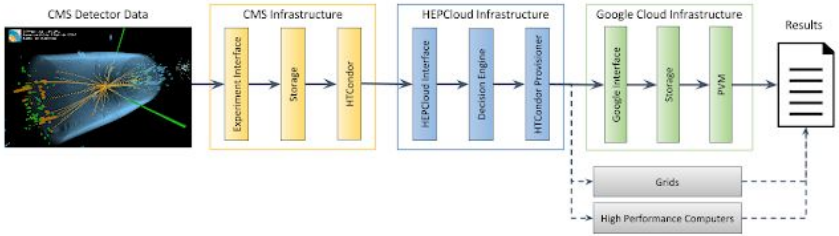
NASA Press release: https://www.nasa.gov/press-release/artificial-intelligence-nasa-data-used-to-discover-eighth-planet-circling-distant-star

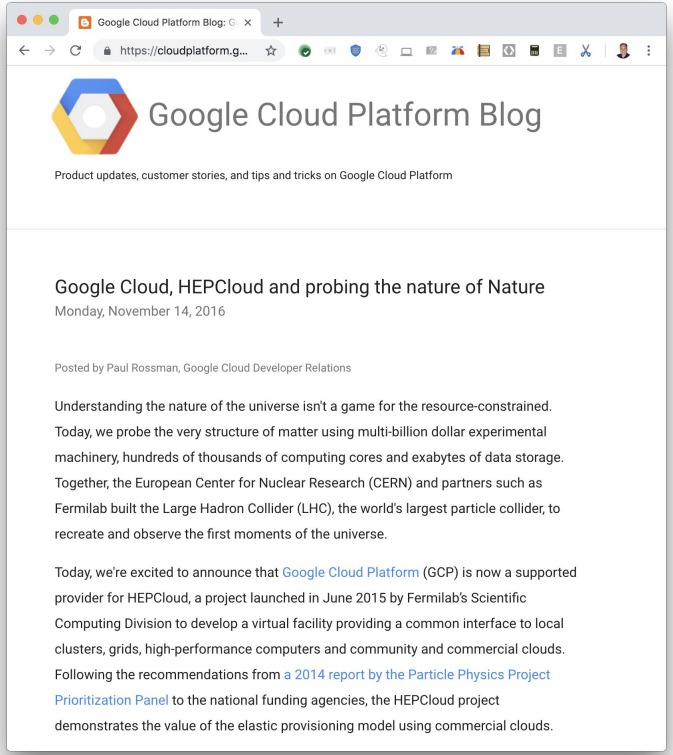
Google

Chris Shallue - Google
Andrew Vanderburg - UT Austin & Harvard-Smithsonian Center for Astrophysics

Google Cloud, HEPCloud and probing the nature of Nature







Google Cloud Platform Blog

Product updates, customer stories, and tips and tricks on Google Cloud Platform

Google Cloud, HEPCloud and probing the nature of Nature

Monday, November 14, 2016

Posted by Paul Rossman, Google Cloud Developer Relations

Understanding the nature of the universe isn't a game for the resource-constrained. Today, we probe the very structure of matter using multi-billion dollar experimental machinery, hundreds of thousands of computing cores and exabytes of data storage. Together, the European Center for Nuclear Research (CERN) and partners such as Fermilab built the Large Hadron Collider (LHC), the world's largest particle collider, to recreate and observe the first moments of the universe.

Today, we're excited to announce that Google Cloud Platform (GCP) is now a supported provider for HEPCloud, a project launched in June 2015 by Fermilab's Scientific Computing Division to develop a virtual facility providing a common interface to local clusters, grids, high-performance computers and community and commercial clouds. Following the recommendations from a 2014 report by the Particle Physics Project Prioritization Panel to the national funding agencies, the HEPCloud project demonstrates the value of the elastic provisioning model using commercial clouds.

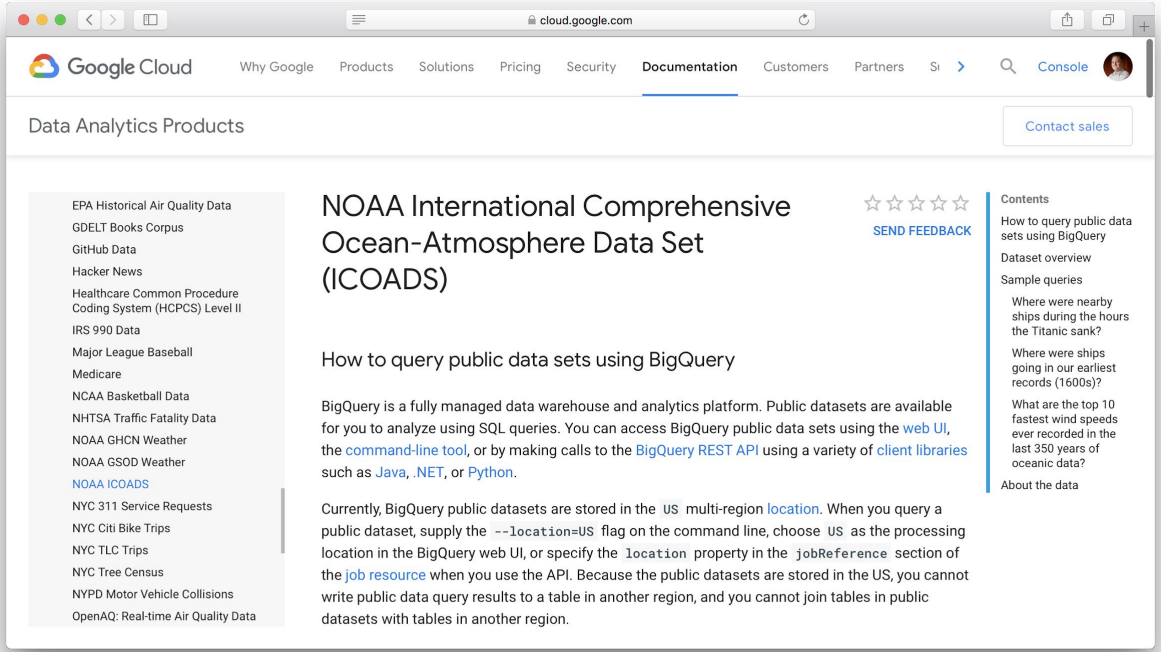
Google

https://cloudplatform.googleblog.com/2016/11/Google-Cloud-HEPCloud-and-probing-the-nature-of-Nature.html

NOAA Public Data Sets in BigQuery and Cloud Storage

“The great thing about what Google is doing with the public datasets program is they are combining various valuable datasets that are not readily available, nevermind all available in one analytics tool. Getting, for example, NOAA, USGS, EPA, and NSF data all in one place is incredibly valuable for data science and we are continually finding requests for this type of access. This is something that is not being done right now outside of the Google Cloud program.”

Dr. Ed Kearns, NOAA Chief Data Officer



http://www.noaa.gov/big-data-project
https://cloud.google.com/noaa-big-data/
https://cloud.google.com/public-datasets/

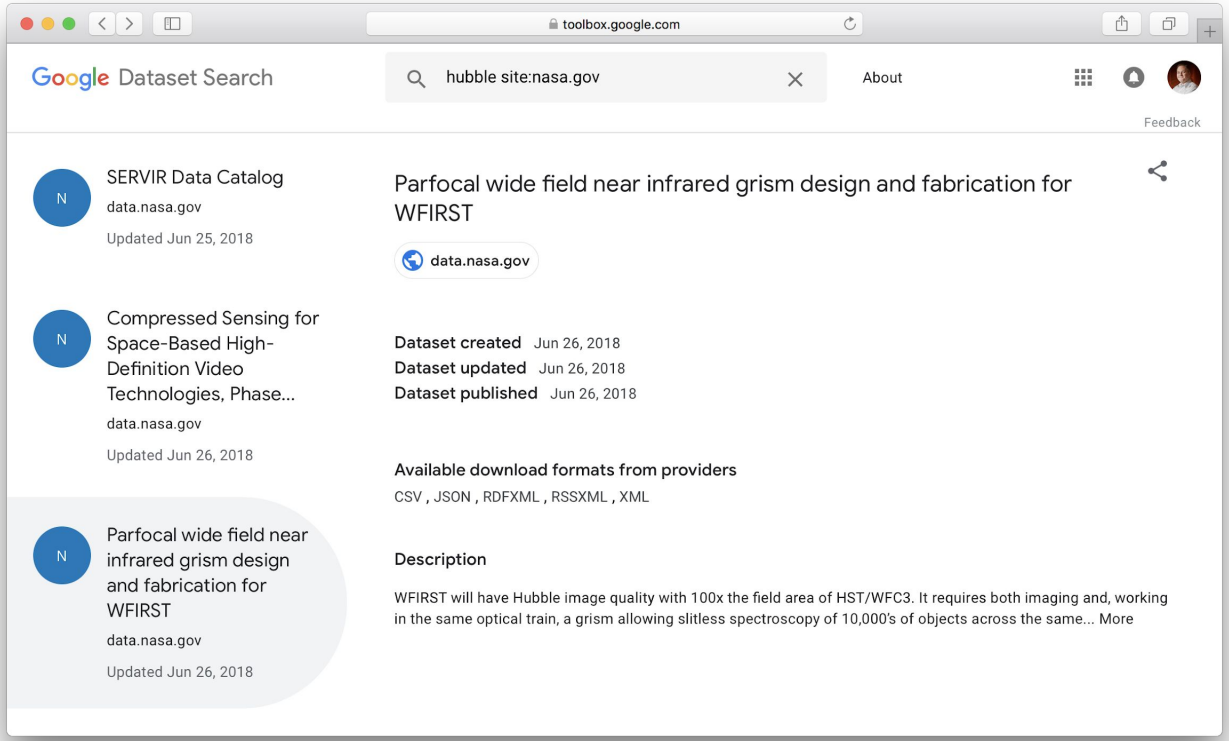
https://cloud.google.com/bigquery/sample-tables
https://www.youtube.com/watch?v=vTJWb4twnM8

Google Dataset Search

Find and discover public datasets with metadata and links to source systems.

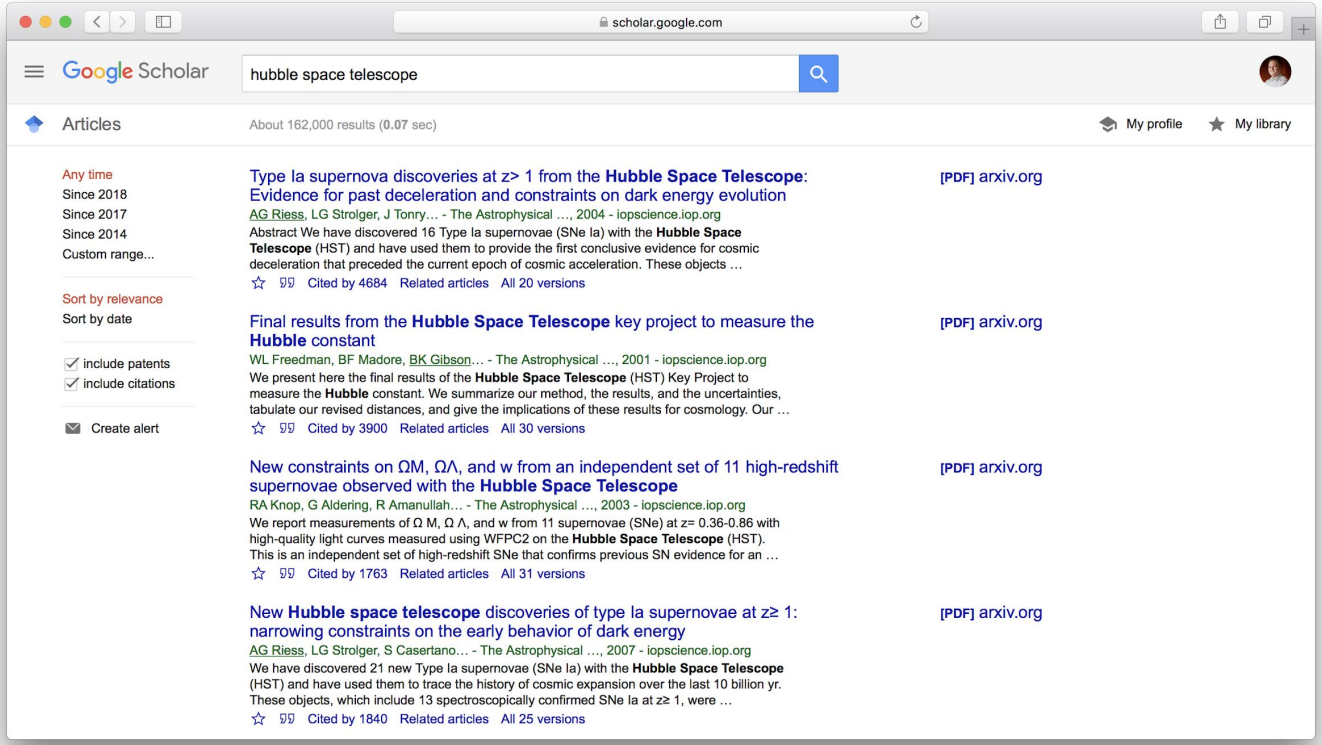
For example:
Hubble site:nasa.gov

CMIP5 site:noaa.gov



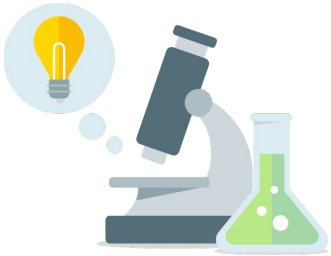
https://toolbox.google.com/datasetsearch

Google Scholar



https://scholar.google.com

Cloud Credits & Faculty Research Awards



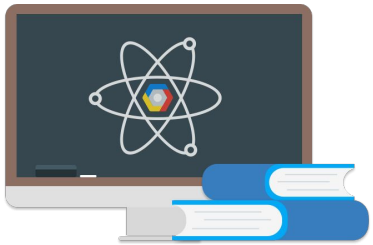
GCP Research Credits

- Free access to GCP for research workloads

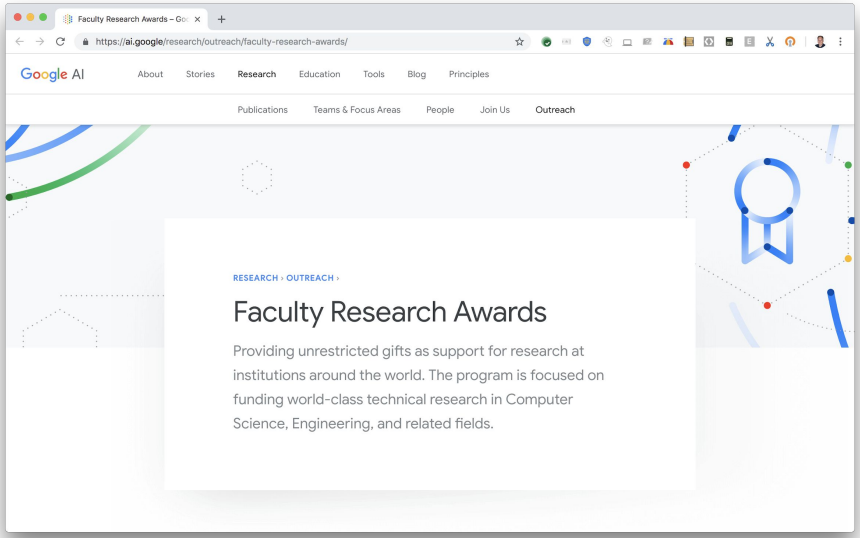
https://cloud.google.com/edu/?options=research-credits

Faculty Grants

- GCP credits for computer science faculty and students
- Curated resources designed to help CS faculty and students learn GCP



https://cloud.google.com/edu/?options=faculty-grants



- Unrestricted gifts as support for research at institutions around the world.
- Focused on funding world-class technical research in Computer Science, Engineering, and related fields.

https://ai.google/research/outreach/faculty-research-awards/

Google Credits for the NSF BIGDATA Program

Through the National Science Foundation's BIGDATA grants program, we're offering researchers $3M in Google Cloud Platform credits to use the same infrastructure, analytics and machine learning that we use to drive innovation at Google.

Has inspired innovative climate analysis and modeling at Columbia University and the launch of pangeo.io

GOOGLE CLOUD

Google Cloud supports $3M in grant credits for the NSF BIGDATA program

Bram Bout
Director, Google for Education

Karan Bhatia
Cloud Solutions Architect, Google Cloud

Published Feb 22, 2017

Google Cloud Platform (GCP) serves more than one billion end-users, and we continue to seek ways to give researchers access to these powerful tools. Through the National Science Foundation's BIGDATA grants program, we're offering researchers $3M in Google Cloud Platform credits to use the same infrastructure, analytics and machine learning that we use to drive innovation at Google.

https://www.blog.google/products/google-cloud/google-cloud-supports-3m-grant-credits-nsf-bigdata-program/
https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=504767

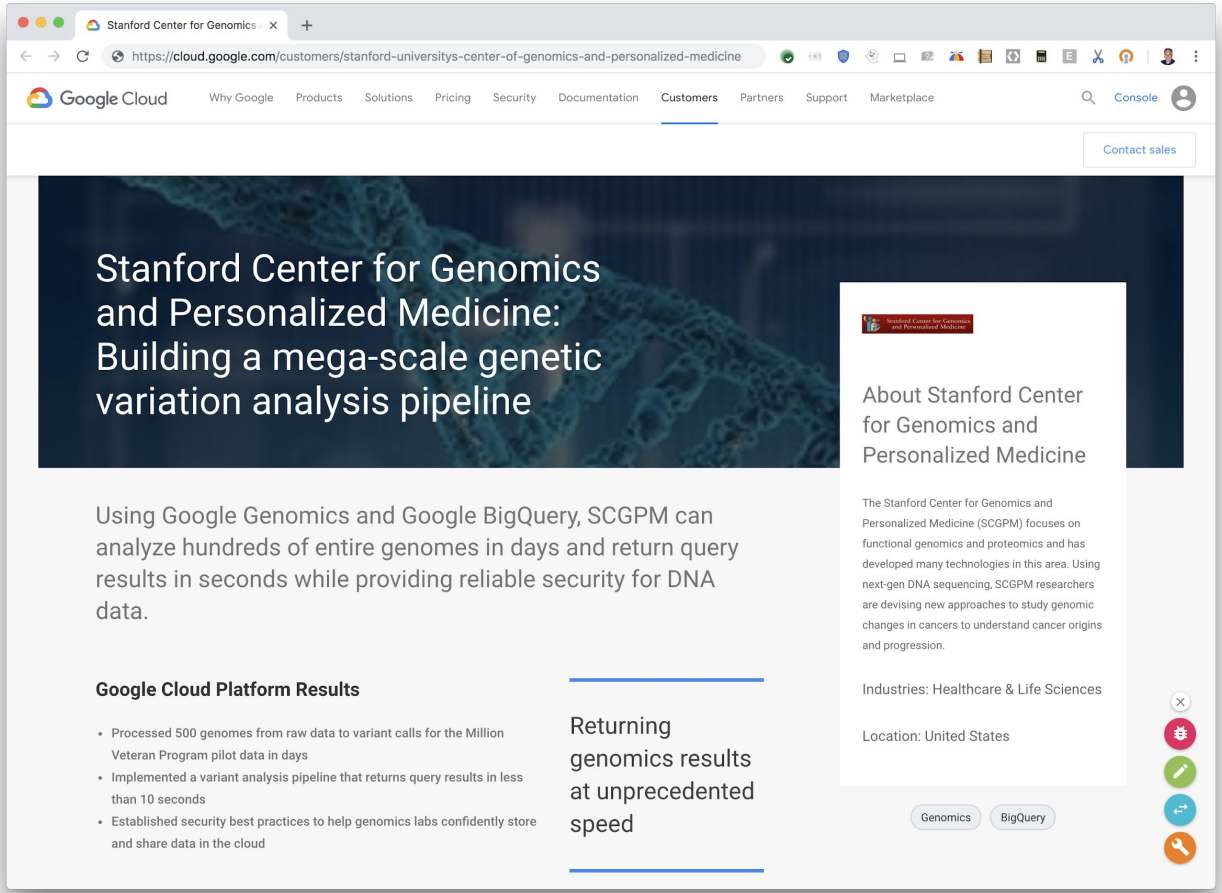
Northeastern University uses GCP to model Zika's Spread

WHAT RESEARCHERS ARE SAYING

"We have the flexibility to scale up to several thousand independent virtual instances in parallel, so we can generate a full analysis for a single epidemic scenario—which may consist of up to 250,000 independent simulations—in less than a day."

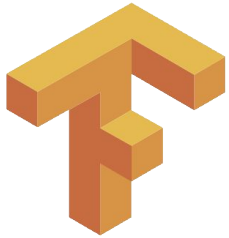
Matteo Chinazzi, Associate Research Scientist, Northeastern University

https://edu.google.com/latest-news/stories/neu-gcp/



Google

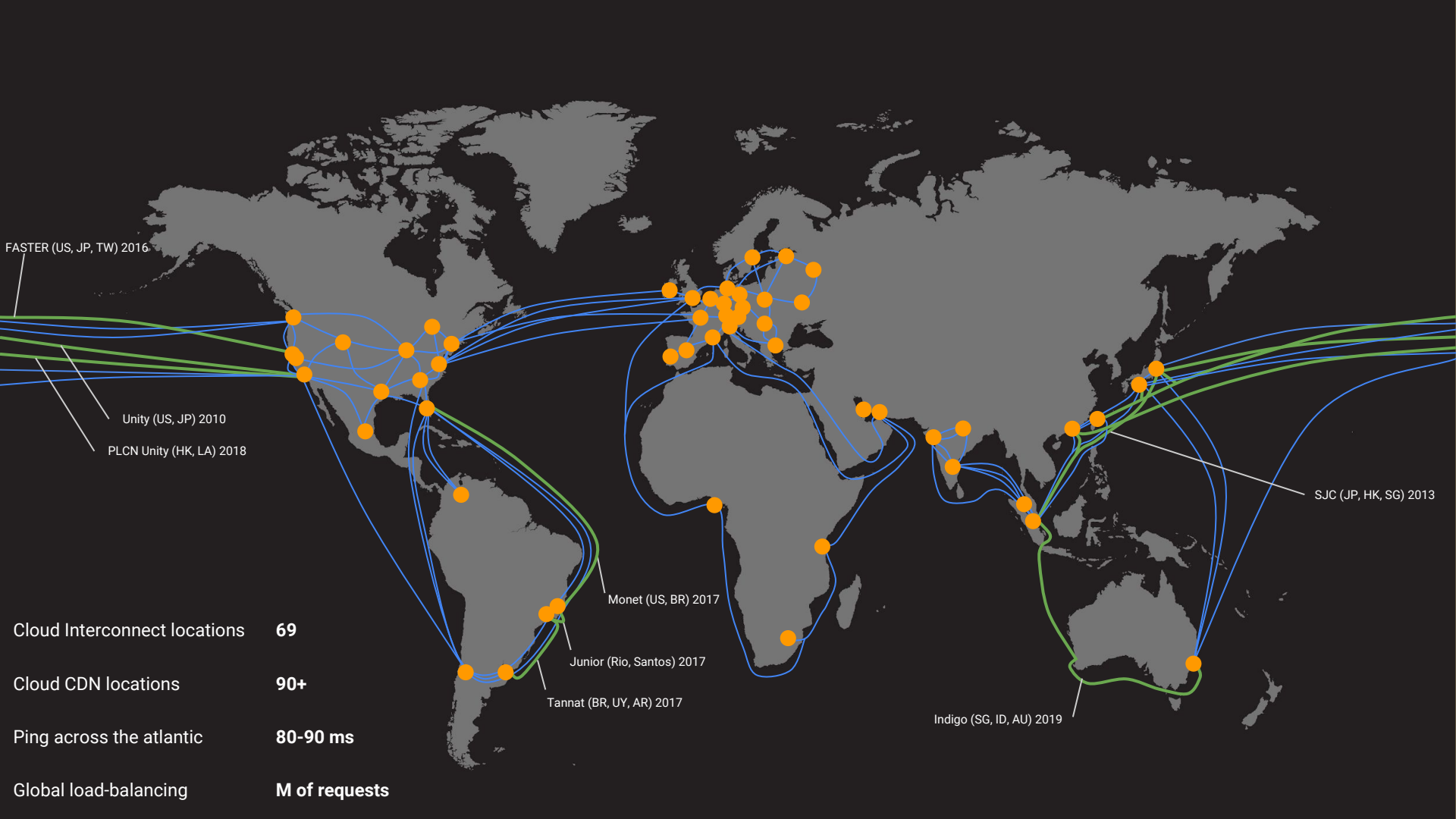
https://cloud.google.com/customers/stanford-universitys-center-of-genomics-and-personalized-medicine



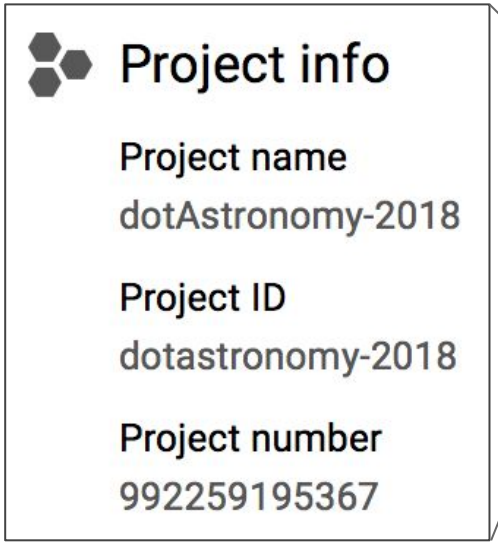
Overview of Relevant Google Cloud Services

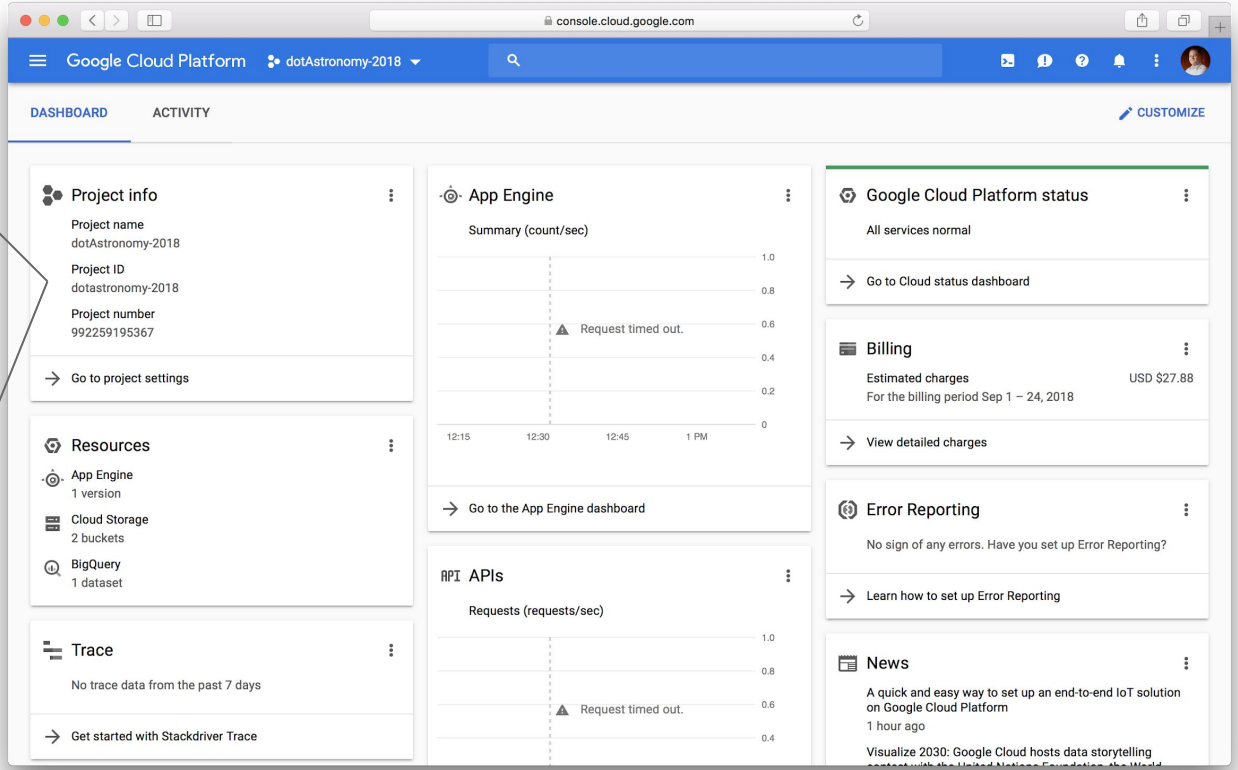

Oregon
3
Los Angeles
3
Iowa
4
3
3
N Virginia
S Carolina
3
Montreal
Belgium
London
3
3
3
Netherlands
Frankfurt
3
Finland
3
Zürich
3
Mumbai
3
Osaka
3
Tokyo
HK
3
3
Taiwan
3
Singapore
3
São Paulo
Sydney
3



Cloud Interconnect locations	**69**
Cloud CDN locations	**90+**
Ping across the atlantic	**80-90 ms**
Global load-balancing	**M of requests**

Organizations vs. Projects





GCP ↔ On-Prem Connectivity options

Public internet/VPN

- Simple
- Resilient
- Use Google's existing edge network

Dedicated interconnect

- Lower GCP egress cost
- Dedicated bandwidth and SLA at up to 100 Gbps
- Private space RFC-1918 addressing
- 70+ locations in 33 countries

Direct peering

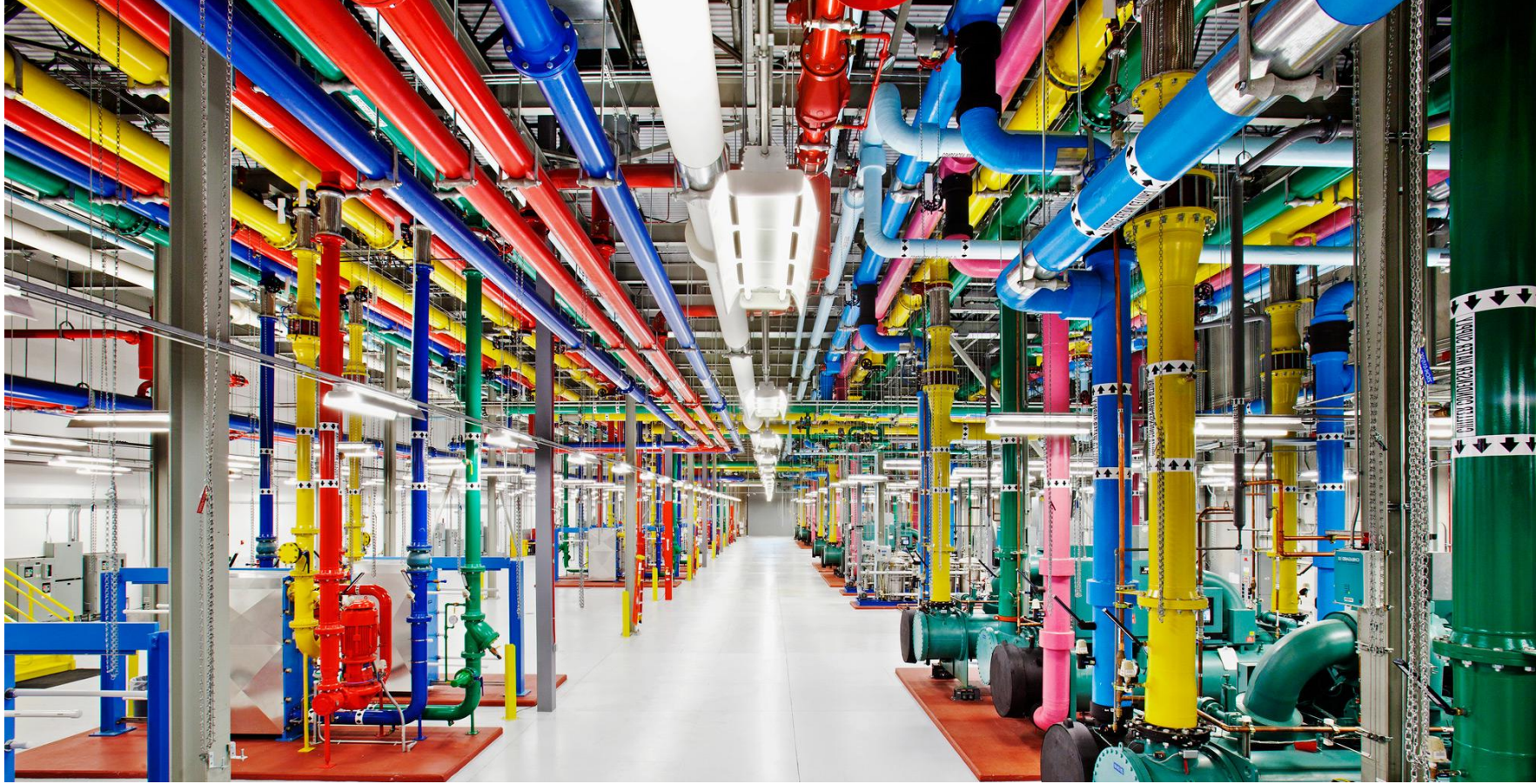
- Utilizes existing BGP route selection and internet routing
- Greater control of peering facilities
- Requirements for peering

Carrier Peering

- Offers benefits of peering when requirements cannot be met
- Service provider partners can provide SLA

Google Cloud

Google Cloud



Google Cloud

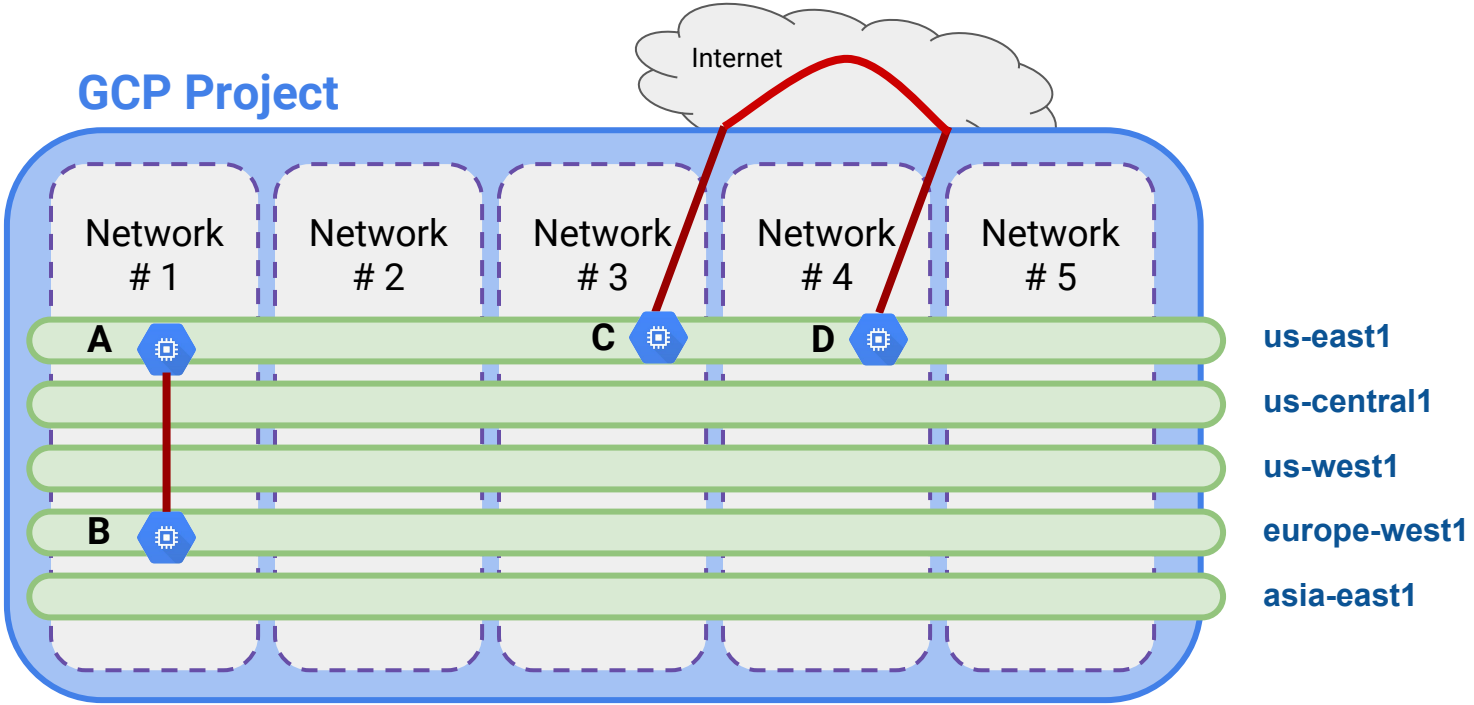
Google Cloud

A **GCP project** is a global and locked-down resource
Communication between resources in different projects require Shared VPC

A project can contain many **GCE networks**, which are global
Machines in different zones/regions can live on the same network

Load balancing is just a forwarding rule
A single anycast IP to all backend instances in any region

Network egress is charged when data is crossing regions
Irrespective if they are connected through a GCP network or not

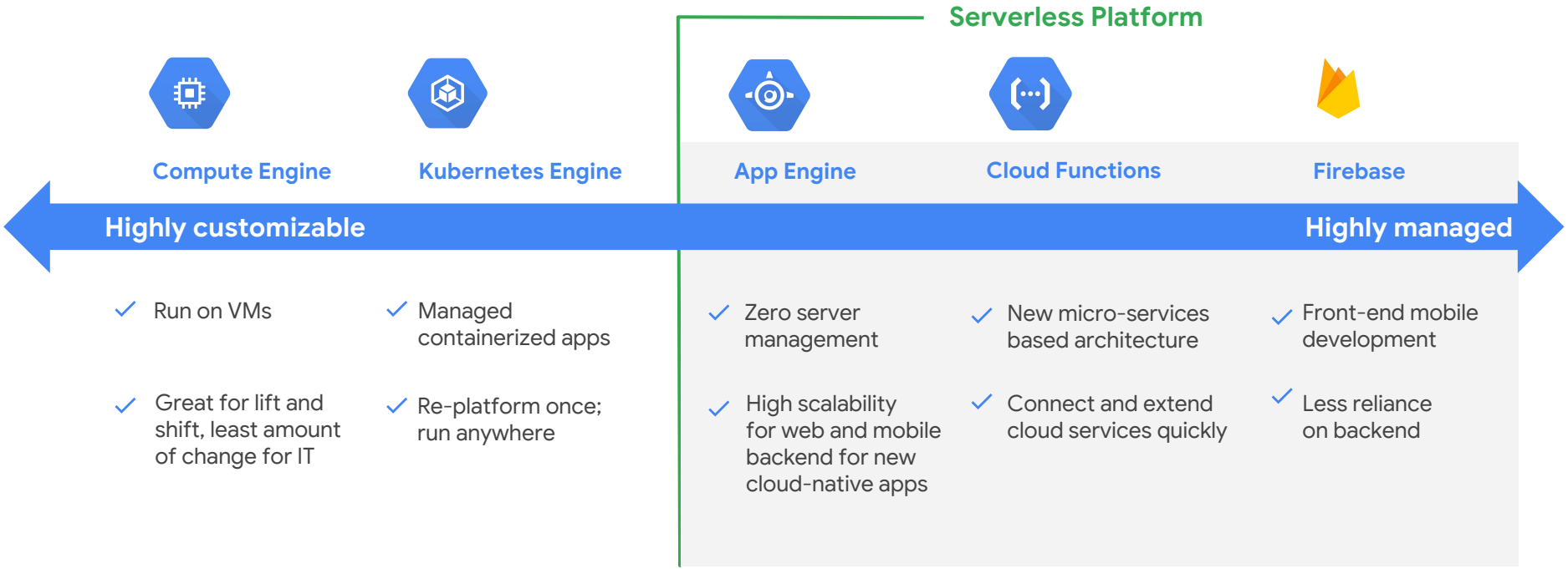


1. A and B can communicate over Internal IPs *even though they are in different regions*
2. C and D must communicate over External IPs *even though they are in the same region*

The compute continuum

Hosting infrastructure, containers, apps, and code

Compute continuum



Compute Engine

Full control: VMs for Linux and Windows Server

Machine type

Customize to select cores, memory and GPUs.

Basic view

Cores

96 vCPU 1 - 96

Memory

460 GB 86.4 - 624

Extend memory

CPU platform

Intel Skylake or later

GPUs

Choosing a machine type

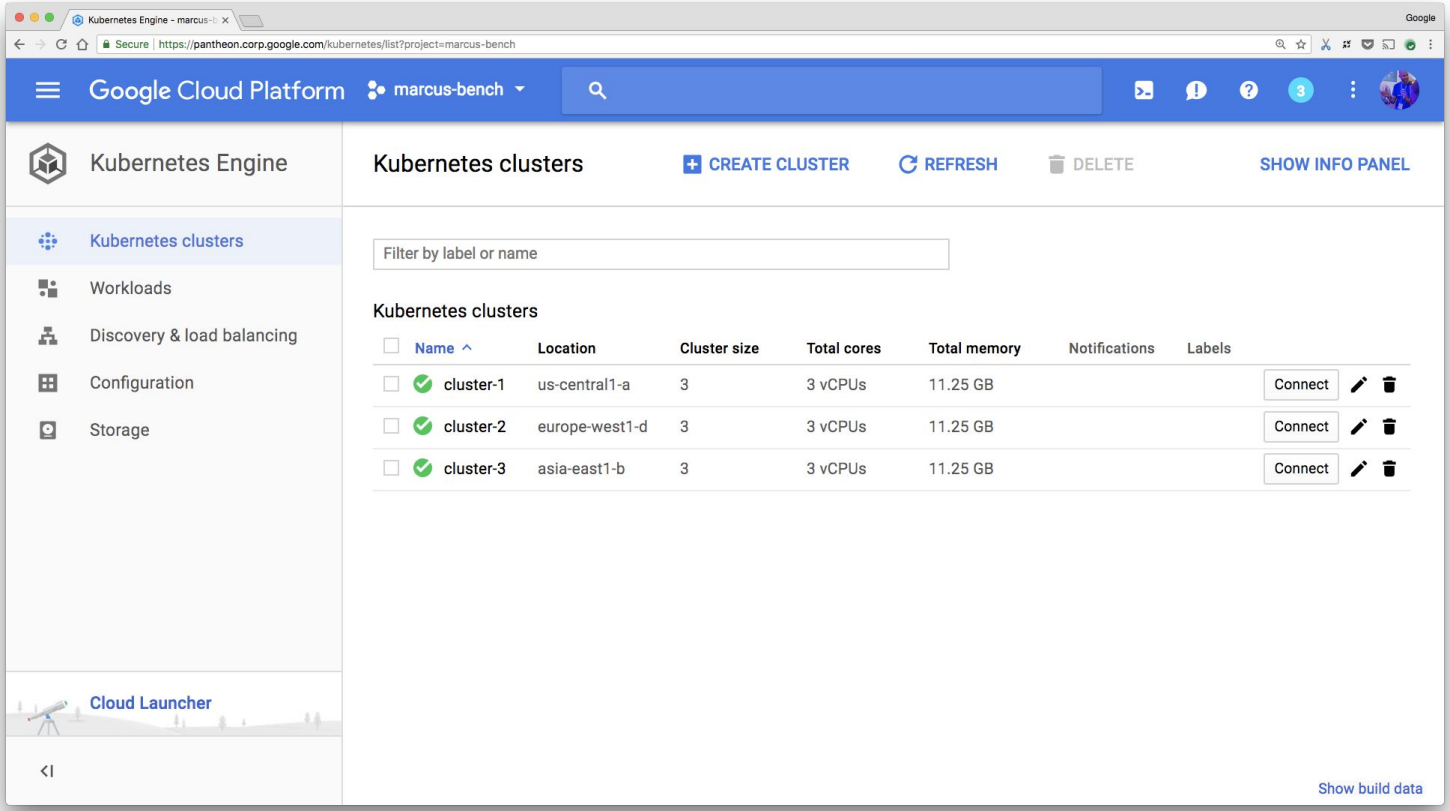
How to create your first GCE VM image

```
# Start with a small ubuntu instance
gcloud compute instances create --image-project ubuntu-os-cloud --image ubuntu-1804-bionic-v20180911
--machine-type n1-highmem-8 --boot-disk-size 250GB --boot-disk-type pd-ssd --zone us-west1-b
ubuntu-18-04

# Delete the VM, but keep the disk
gcloud compute instances delete ubuntu-18-04 --keep-disks all

# Create a VM image from the disk
gcloud compute images create my-ubuntu-18-04 --source-disk ubuntu-18-04

# Launch a new VM using the image
gcloud compute instances create --image my-ubuntu-18-04 --machine-type n1-highmem-2 my-new-ubuntu-vm

# Now that you have verified that the image works, you can delete the disk from your source image
gcloud compute disks delete ubuntu-18-04
```

Kubernetes Engine

Open source Kubernetes-based container orchestration

- Managed by Google
  - Runs a stable, vetted and modern K8s release
    - Optionally use time-limited ‘alpha clusters’
  - Master node managed and upgraded by Google
  - Optional auto-upgrades for worker nodes
  - Zonal and regional clusters
- Nodes run Container-optimised OS
  - Backed by Chromium team
  - Security hardened
  - In-place upgrades and downgrades


Kubernetes Engine - marcus-b
Google
Secure | https://pantheon.corp.google.com/kubernetes/list?project=marcus-bench
Google Cloud Platform
marcus-bench
3
Kubernetes Engine
Kubernetes clusters
CREATE CLUSTER
REFRESH
DELETE
SHOW INFO PANEL
Kubernetes clusters
Workloads
Discovery & load balancing
Configuration
Storage
Filter by label or name
Kubernetes clusters
Name
Location
Cluster size
Total cores
Total memory
Notifications
Labels
cluster-1 us-central1-a 3 3 vCPUs 11.25 GB Connect
cluster-2 europe-west1-d 3 3 vCPUs 11.25 GB Connect
cluster-3 asia-east1-b 3 3 vCPUs 11.25 GB Connect
Cloud Launcher
Show build data

App Engine

Deploy your code and we scale it for you

- PHP, Python 2.7/3.5, Java, Go, Node.JS, Ruby, & Custom Runtimes
- Applications run within containers which can be extracted and run elsewhere
- Build **services**, deploy **versions**, auto-scale **instances**
- Integrated debugging, tracing, logging and monitoring

Cloud Functions

Serverless platform for event-based microservices

- Isolated, secure execution context and lifecycle
- Written in Javascript, executed in Node.js runtime
  - Portability: write and test locally
  - More runtimes to follow
- Triggered by
  - Cloud Pub/Sub topics
  - Cloud Storage object notifications
  - HTTP endpoints
- Access all GCP services using standard GCP client libraries

Stackdriver

Monitoring

Platform, system, & application metrics
Uptime/health checks
Dashboards
Alerts

Logging

Platform, system, & application logs
Log search, view, filter and logs-based metrics

Debug

Production debug snapshots
Conditional snapshots
IDE integration
Live breakpoints

Trace

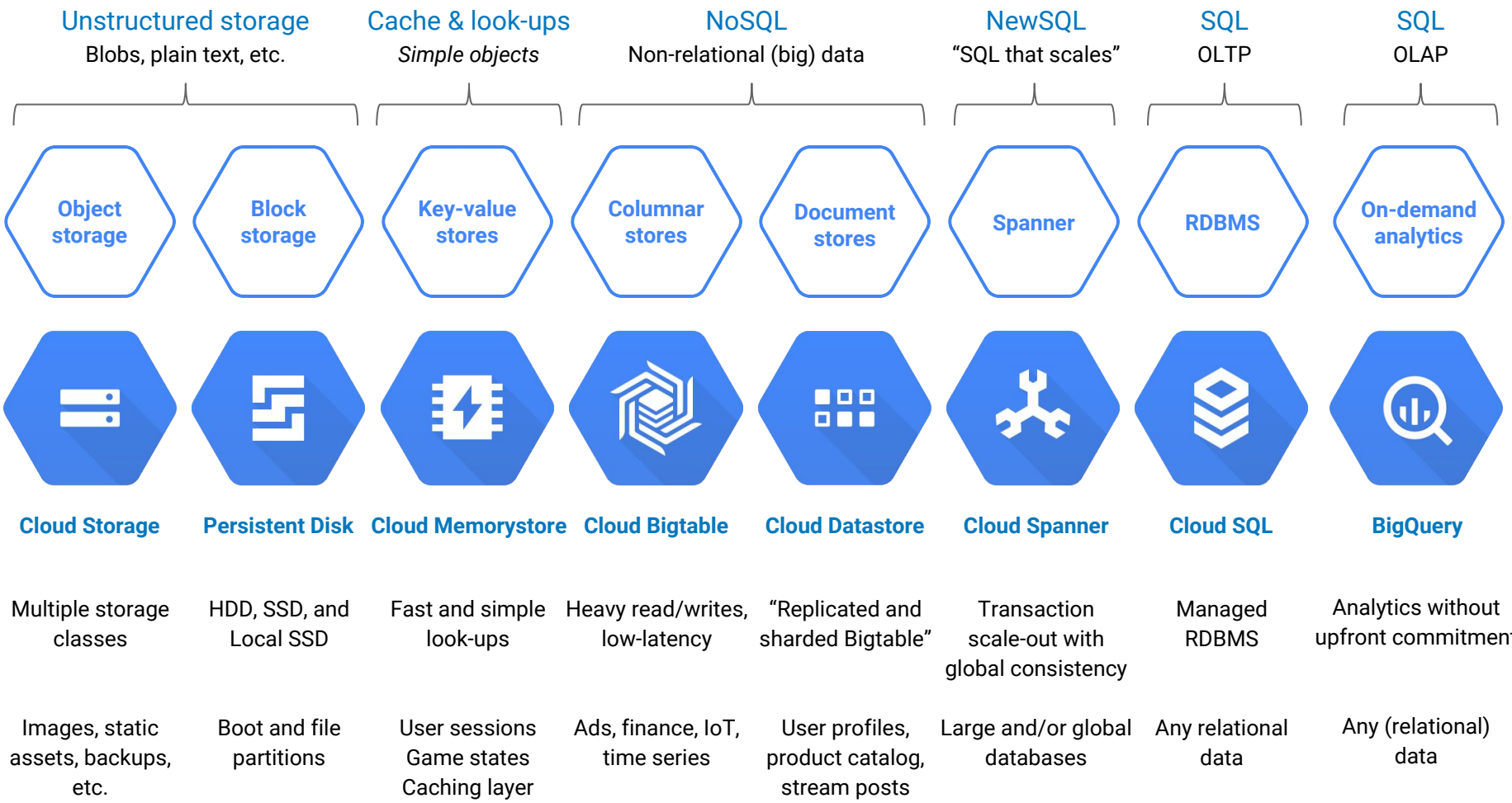
Invocation statistics
Latency Reporting
Per-URL latency sampling

Error Reporting

Error notifications
Error dashboard

The storage suite

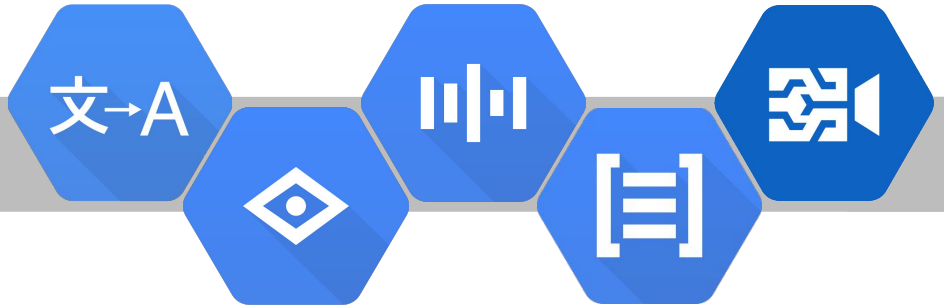
Structured, unstructured, and everything in between


Unstructured storage
Blobs, plain text, etc.
Cache & look-ups
Simple objects
NoSQL
Non-relational (big) data
NewSQL
"SQL that scales"
SQL
OLTP
SQL
OLAP
Object storage
Block storage
Key-value stores
Columnar stores
Document stores
Spanner
RDBMS
On-demand analytics
Cloud Storage
Persistent Disk
Cloud Memorystore
Cloud Bigtable
Cloud Datastore
Cloud Spanner
Cloud SQL
BigQuery
Multiple storage classes
HDD, SSD, and Local SSD
Fast and simple look-ups
Heavy read/writes, low-latency
"Replicated and sharded Bigtable"
Transaction scale-out with global consistency
Managed RDBMS
Analytics without upfront commitment
Images, static assets, backups, etc.
Boot and file partitions
User sessions
Game states
Caching layer
Ads, finance, IoT, time series
User profiles, product catalog, stream posts
Large and/or global databases
Any relational data
Any (relational) data

Cloud Launcher
Explore, deploy, and manage 3rd party solutions

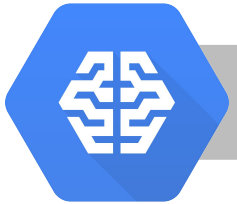
Big data and machine learning

Fully managed lifecycle of data and ML jobs



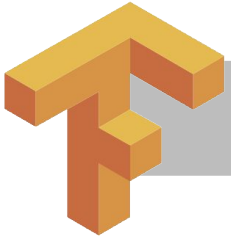
Google-trained ML APIs

Translate, Vision, Speech, Natural Language, and Video



Cloud ML Engine

Train and host your own Tensorflow model



Tensorflow

NVIDIA GPUs

Ideal for simulations, analysis, transcoding, and ML

Attach up to 8 GPUs per VM

Attached directly to the VM via PCIe x16
to achieve bare-metal performance.

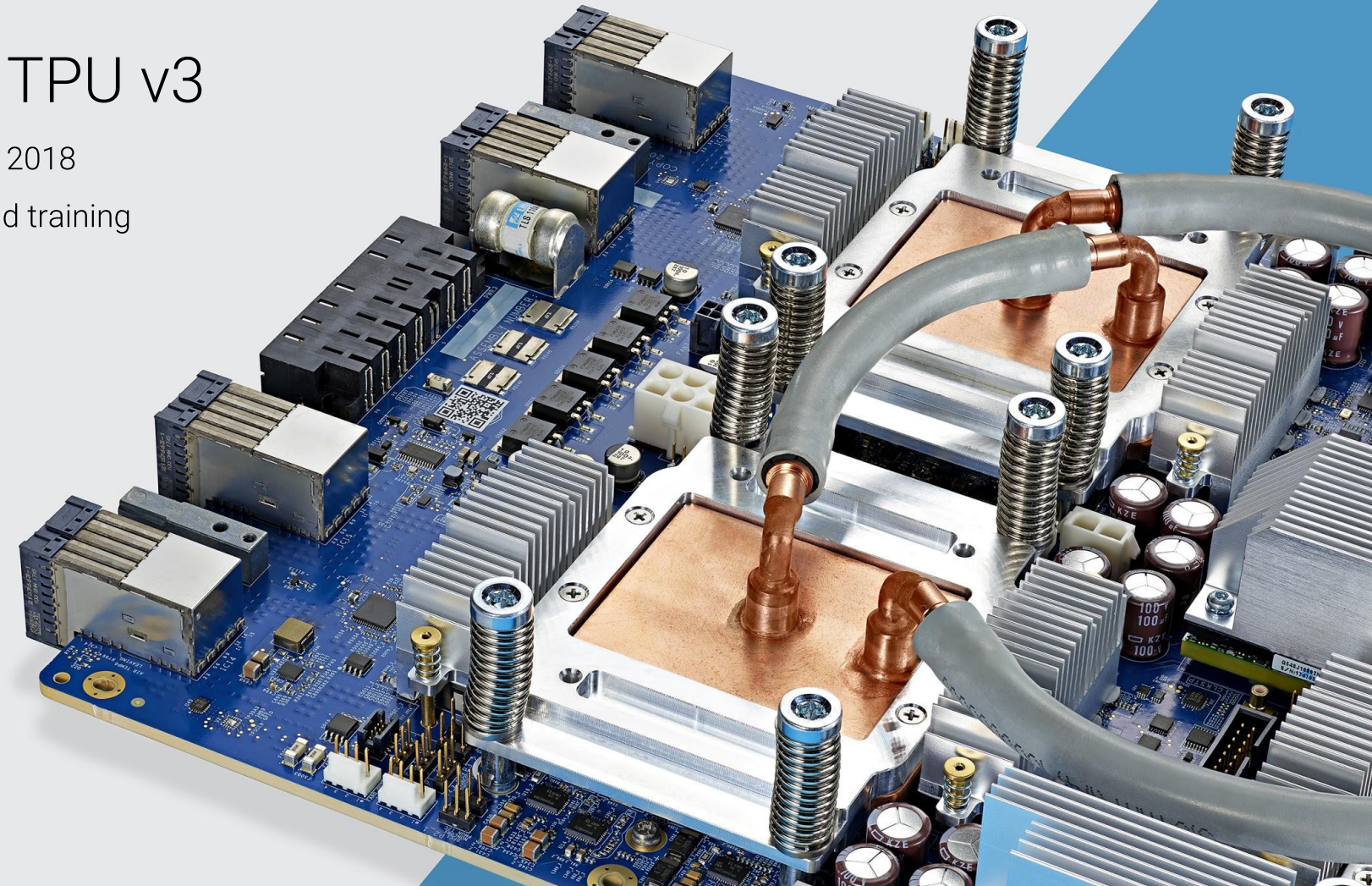
Per second billing
Preemptible support (~50% off)

NVIDIA Tesla K80
NVIDIA Tesla P100
NVIDIA Tesla V100

Cloud TPU v3

Launched in 2018

Inference and training

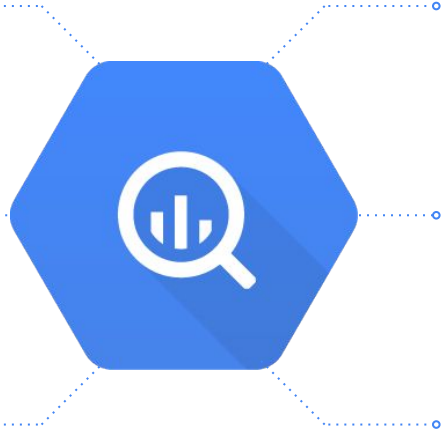


What is BigQuery?

Google Cloud Platform's enterprise data warehouse for analytics

Convenience of standard SQL

Fully managed and serverless



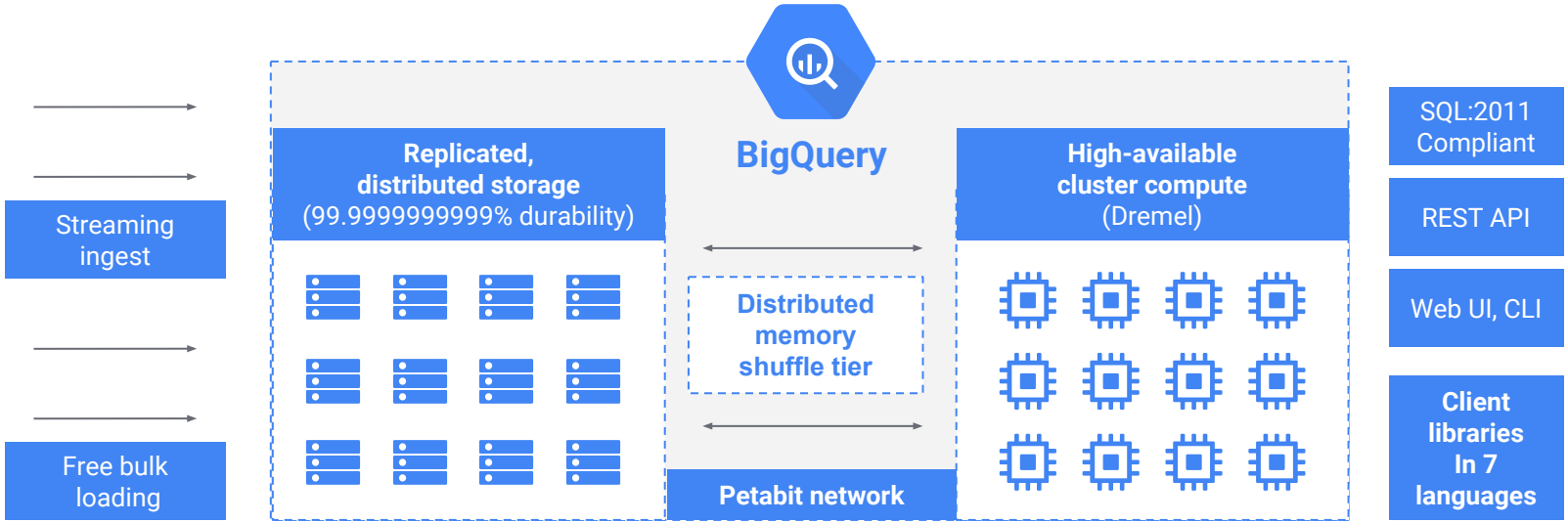
Petabyte-scale storage and queries

Encrypted, durable and highly available

Real-time analytics on streaming data

BigQuery: Architecture

Serverless, decoupled storage and compute for maximum flexibility



Google Cloud

Introducing BigQuery ML

Execute ML initiatives without moving data from BigQuery

Iterate on models in SQL in BigQuery to increase development speed

Automate model selection, and hypertuning

BigQuery ML empowers data analysts and data scientists

Google Cloud

BigQuery ML

Use SQL to...

Step 1:
Create and train your model

```
#standardSQL
CREATE MODEL `bqml_tutorial.sample_model`
OPTIONS(model_type='logistic_reg') AS
SELECT
  IF(totals.transactions IS NULL, 0, 1) AS label,
  IFNULL(device.operatingSystem, "") AS os,
  device.isMobile AS is_mobile,
  IFNULL(geoNetwork.country, "") AS country,
  IFNULL(totals.pageviews, 0) AS pageviews
FROM
  `bigquery-public-data.google_analytics_sample.ga_sessions_*`
WHERE
  _TABLE_SUFFIX BETWEEN '20160801' AND '20170630'
```

Step 2:
Evaluate your model

```
#standardSQL
SELECT
  *
FROM
  ML.EVALUATE(MODEL `bqml_tutorial.sample_model`, (
SELECT
  IF(totals.transactions IS NULL, 0, 1) AS label,
  IFNULL(device.operatingSystem, "") AS os,
  device.isMobile AS is_mobile,
  IFNULL(geoNetwork.country, "") AS country,
  IFNULL(totals.pageviews, 0) AS pageviews
FROM
  `bigquery-public-data.google_analytics_sample.ga_sessions_*`
WHERE
  _TABLE_SUFFIX BETWEEN '20170701' AND '20170801'))
```

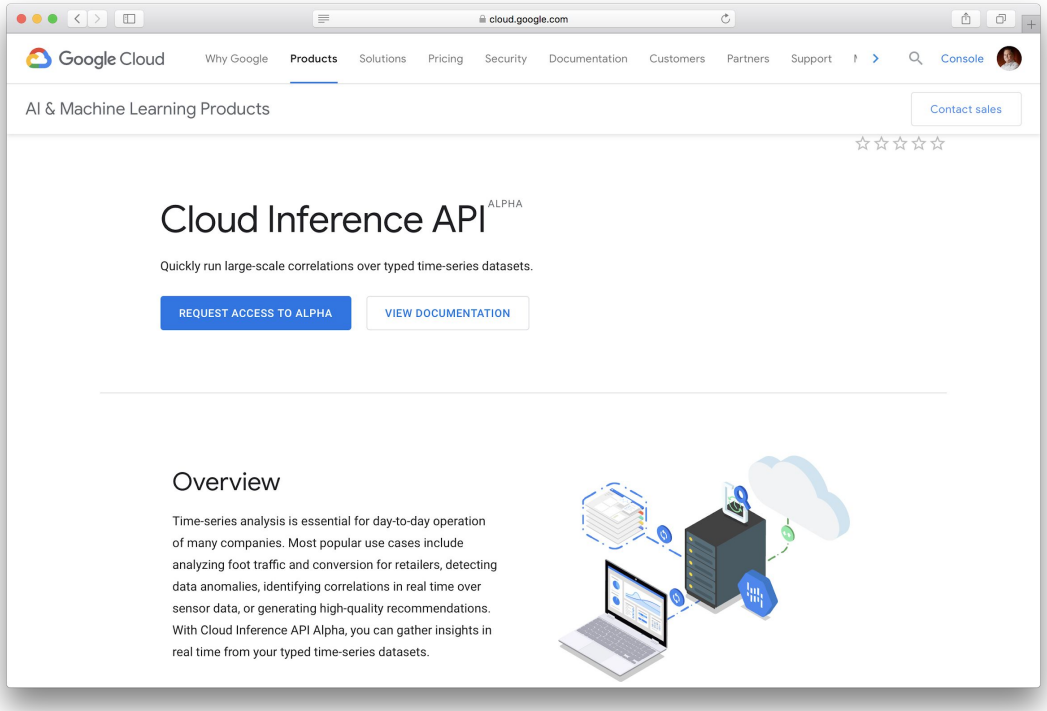
Step 3:
Make predictions

```
#standardSQL
SELECT
  country,
  SUM(predicted_label) as total_predicted_purchases
FROM
  ML.PREDICT(MODEL `bqml_tutorial.sample_model`, (
SELECT
  IFNULL(device.operatingSystem, "") AS os,
  device.isMobile AS is_mobile,
  IFNULL(totals.pageviews, 0) AS pageviews,
  IFNULL(geoNetwork.country, "") AS country
FROM
  `bigquery-public-data.google_analytics_sample.ga_sessions_*`
WHERE
  _TABLE_SUFFIX BETWEEN '20170701' AND '20170801'))
GROUP BY country
ORDER BY total_predicted_purchases DESC
LIMIT 10
```

Cloud Inference API

Recently launched in Alpha:

- Time-series analysis
- Detecting data anomalies
- Identifying correlations in real time over sensor data
- Generate high-quality recommendations



Stream data analytics on Google Cloud Platform

Ingest

Ingest and distribute data reliably

Cloud Pub/Sub

Transform

Fast, correct computations quickly and simply

Cloud Dataflow

Analyze

Machine learning & data warehouse

BigQuery

Cloud Machine Learning

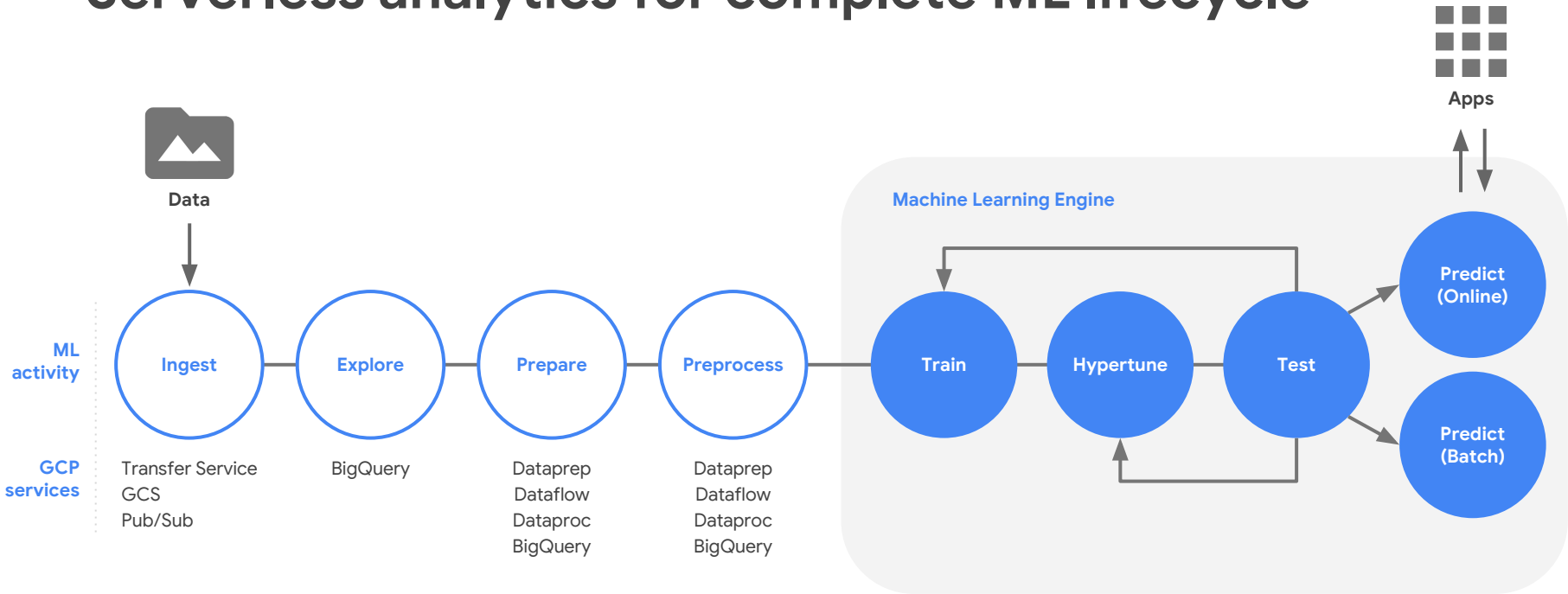
Cloud Translation API

Cloud Vision API

Cloud Natural Language API

Google Cloud

Serverless analytics for complete ML lifecycle



Google Cloud

Google Colaboratory

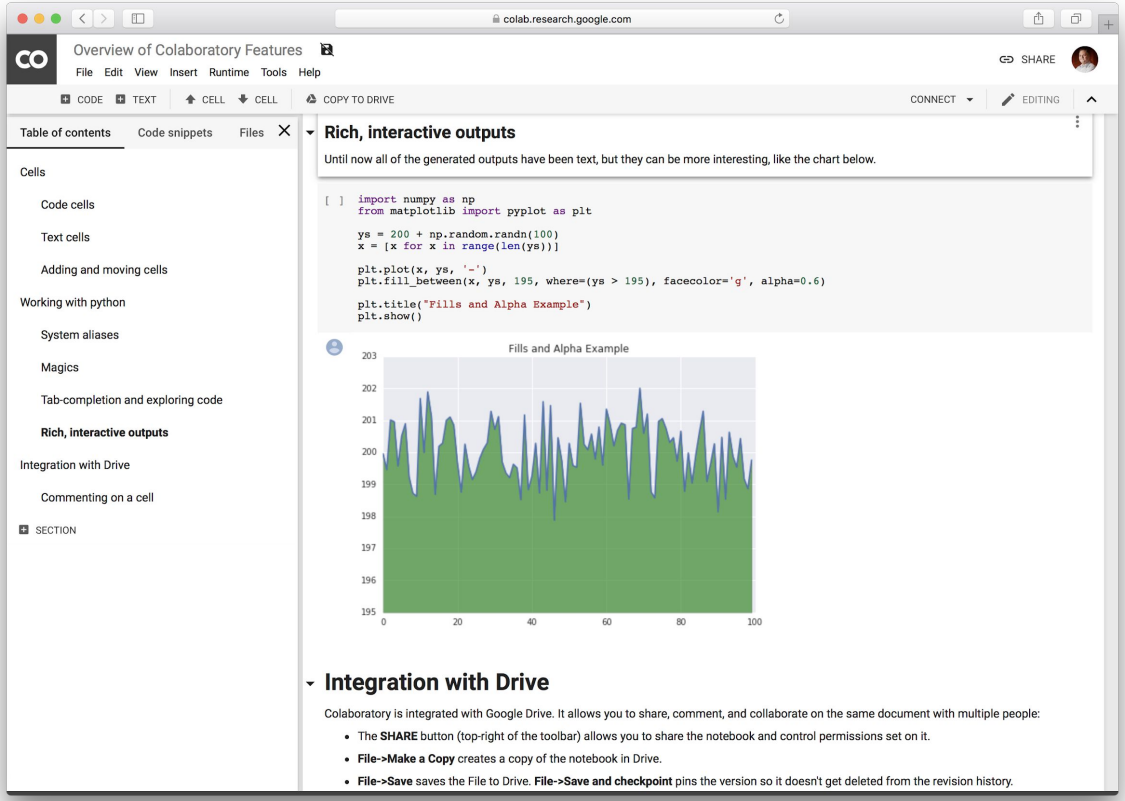
Jupyter notebook environment

Fully managed by Google

Free-to-use

Notebooks stored in Google drive

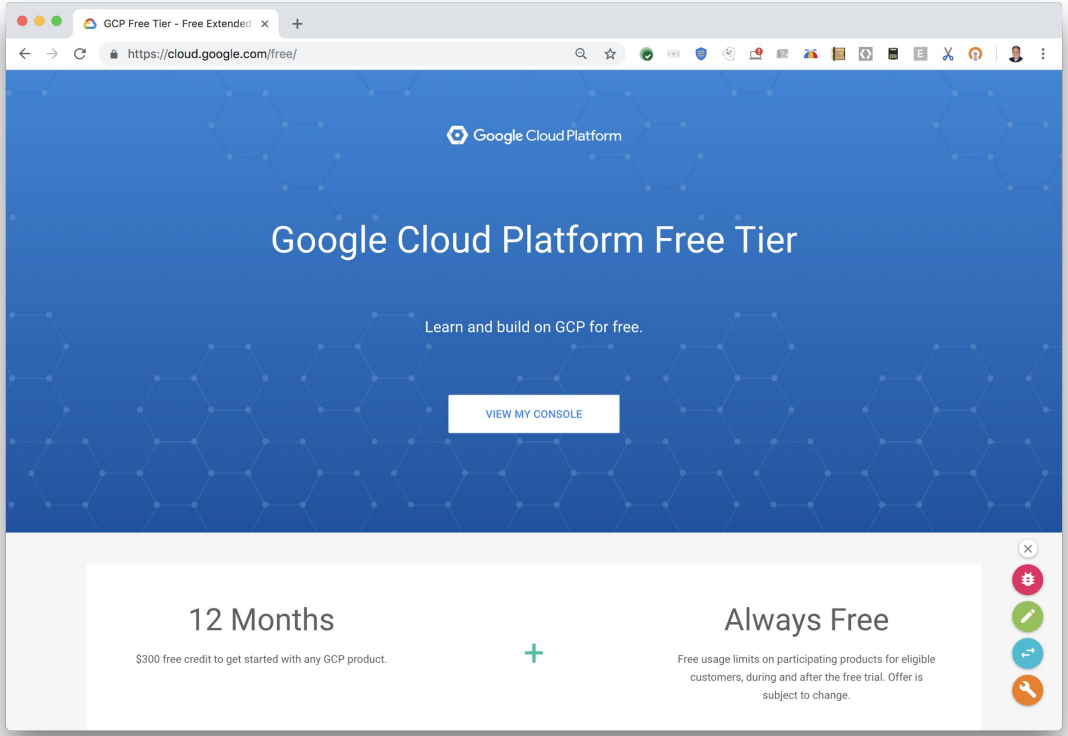
Also includes GPU support



https://colab.research.google.com

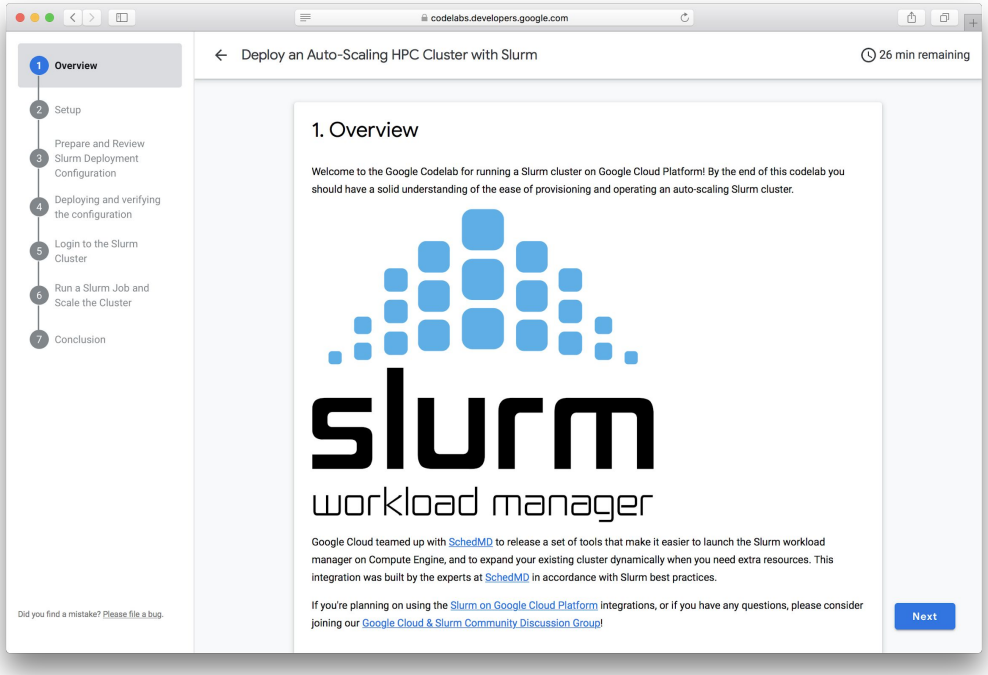
Things to try at .Astronomy X

Enjoy the GCP Free Tier



https://cloud.google.com/free/

Check out the GCP Codelabs



https://codelabs.developers.google.com/codelabs/hpc-slurm-on-gcp/

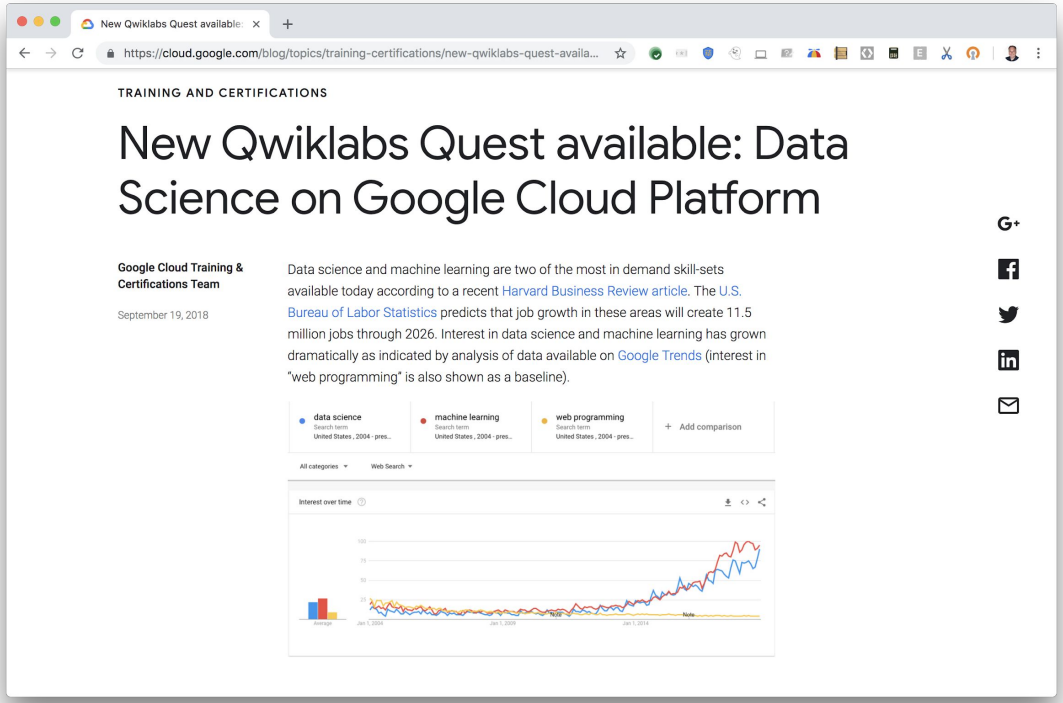
Codelabs are self-paced tutorials that you run in your own GCP project.

Take as long as you like.

Be sure to clean up all resources or simply delete the project when you are finished.

Many codelabs are available at codelabs.developers.google.com

Take advantage of free Qwiklabs



Qwiklabs are self-paced tutorials that run within a GCP account managed by Google.

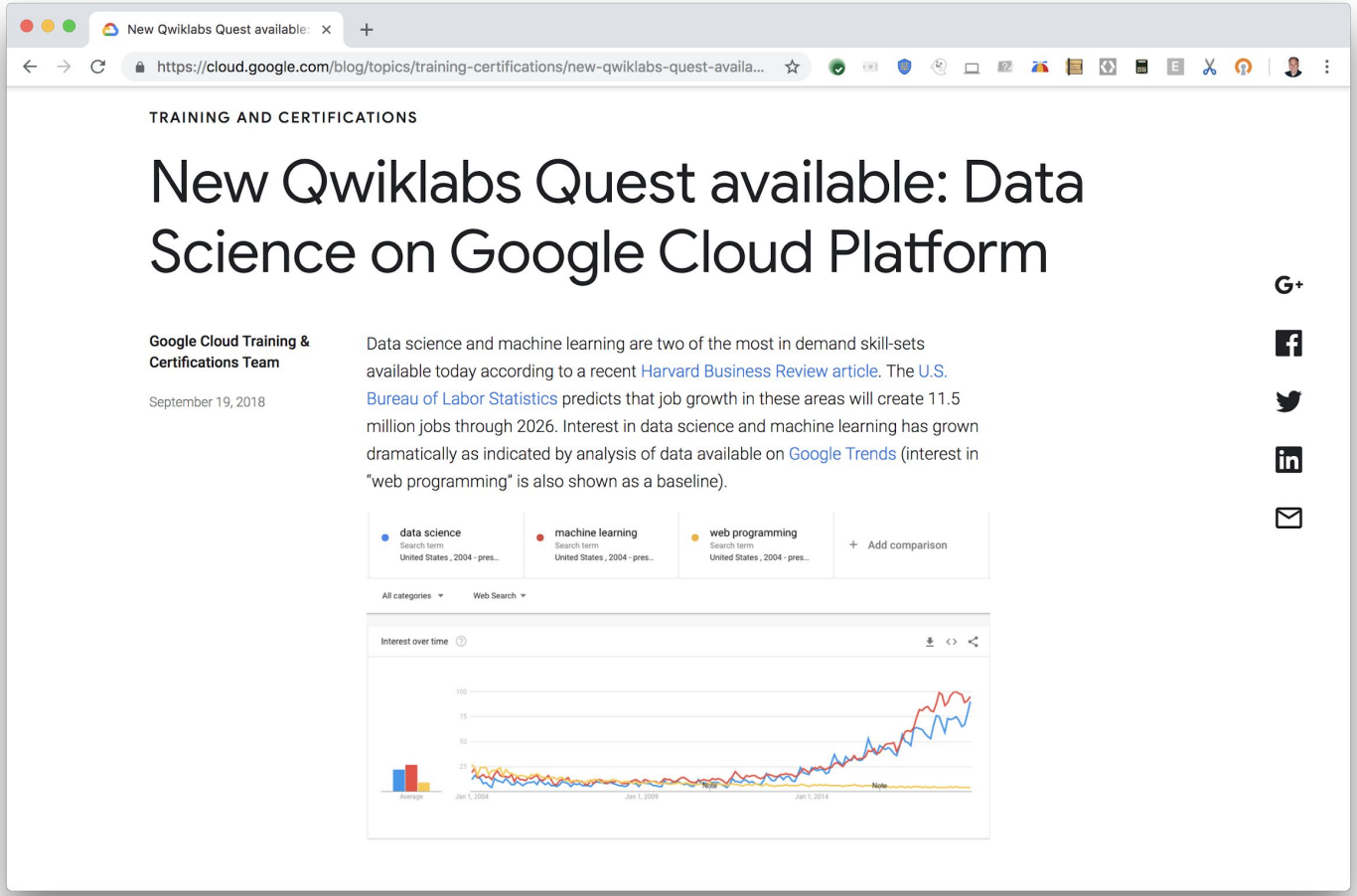
They typically take 30-60 minutes to complete.

Environments disappear after a few hours.

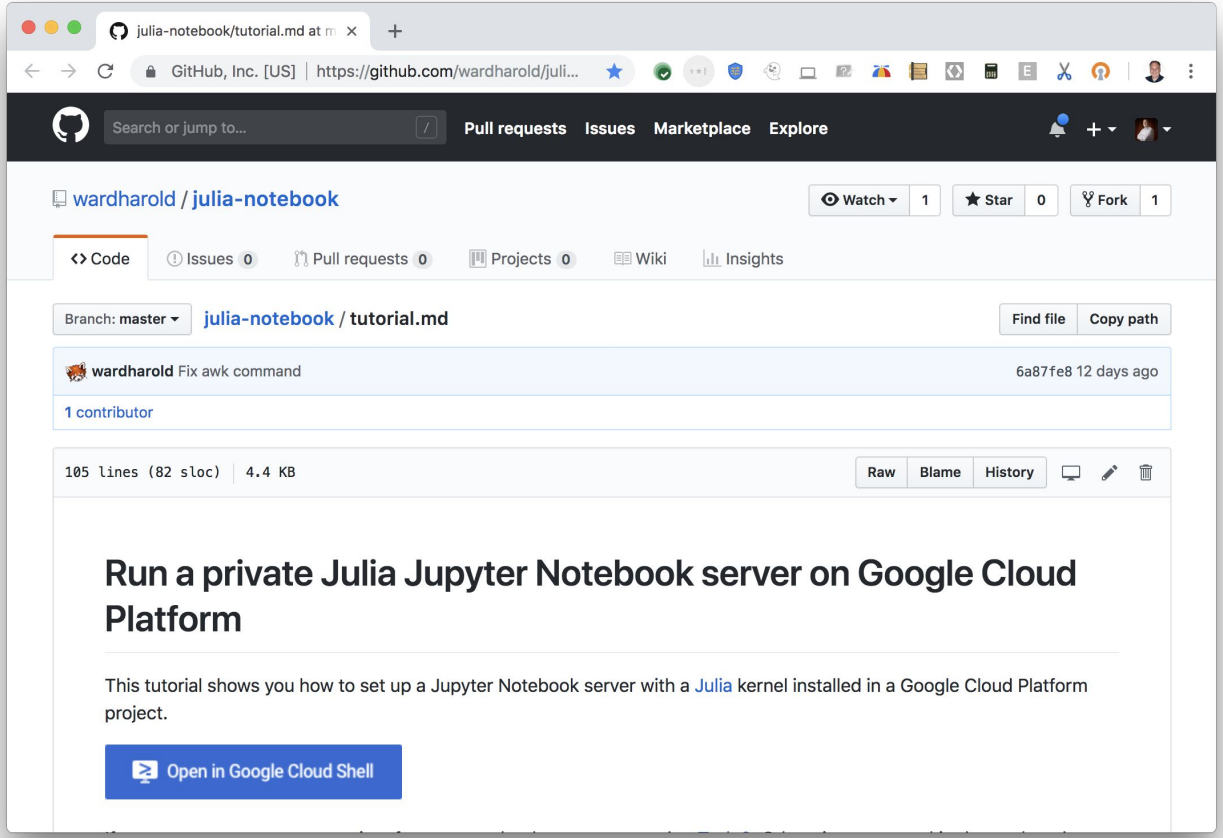
We typically charge by the lab, but some ‘quests’ or tracks are free. For example, the “Data Science on GCP” ‘quest’.

google.qwiklabs.com/home

https://cloud.google.com/blog/topics/training-certifications/new-qwiklabs-quest-available-data-science-google-cloud-platform

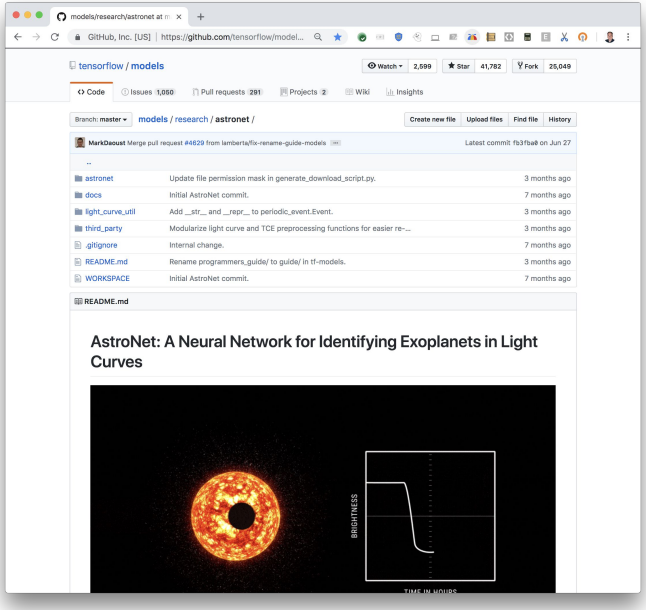


Google https://cloud.google.com/blog/topics/training-certifications/new-qwiklabs-quest-available-data-science-google-cloud-platform



Sandbox: https://github.com/wardharold/julia-notebook/blob/master/tutorial.md
Production: https://github.com/wardharold/community/blob/master/tutorials/julia-jupyter-notebook-server/index.md

Analyze Kepler Data Using Tensorflow

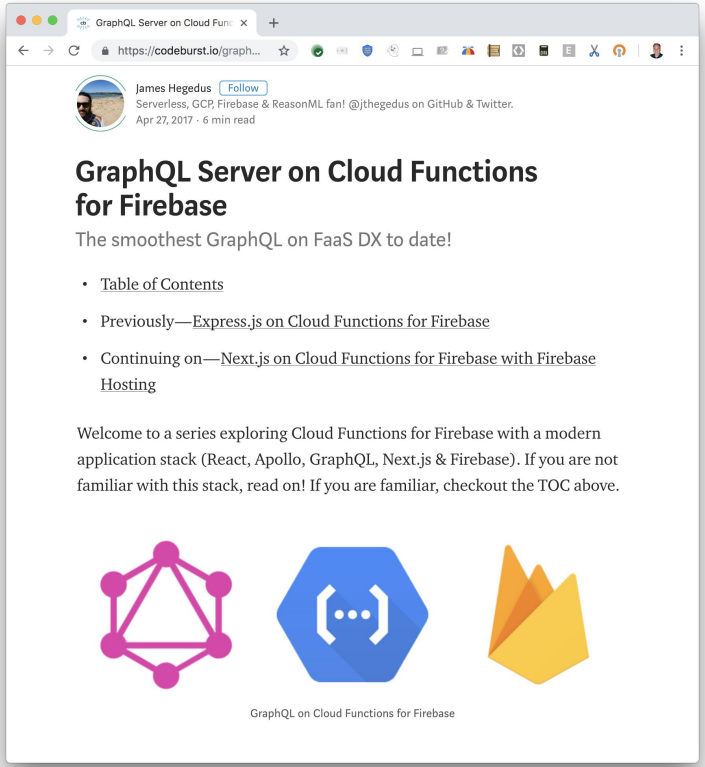


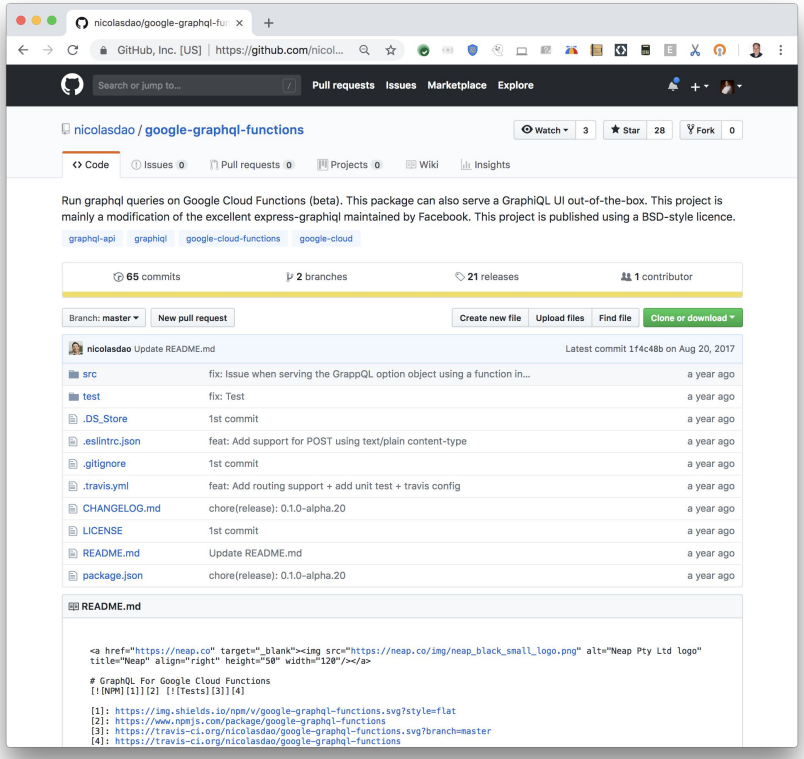
Note that this tutorial is designed to use GCE VMs with GPUs or TPUs. These aren't included in the GCP free tier. Also, before you can launch VMs with GPUs or TPUs, you will need to submit a quota request by visiting the GCP web console and searching for "quota."

In the meantime, let me know if you are interested in trying this out. If so, I can grant you access to one of my projects with access to GPUs and TPUs.

https://github.com/tensorflow/models/tree/master/research/astronet

“Serverless” GraphQL on Cloud Functions (for Arfon)





https://github.com/nicolasdao/google-graphql-functions
https://codeburst.io/graphql-server-on-cloud-functions-for-firebase-ae97441399c0

Google

Resources

Getting up to speed

Documentation: https://cloud.google.com/docs/
Product pages and docs

Google Support Center: http://enterprise.google.com/supportcenter
File support cases through the portal, via email, or by phone

GCP tags on Stack overflow: https://www.stackoverflow.com
Actively monitored by Eng/PM staff

Slack: http://googlecloud-community.slack.com/
Unofficial, but full of customers and googlers talking about GCP

Training: https://cloud.google.com/training/
Three categories of training, each with multiple webinars/classes

Certifications: https://cloud.google.com/certification/
Cloud Architect (coming soon) and Data Engineer (available now)

Coursera: https://www.coursera.org/learn/gcp-fundamentals
Highly recommended self-study on GCP Fundamentals

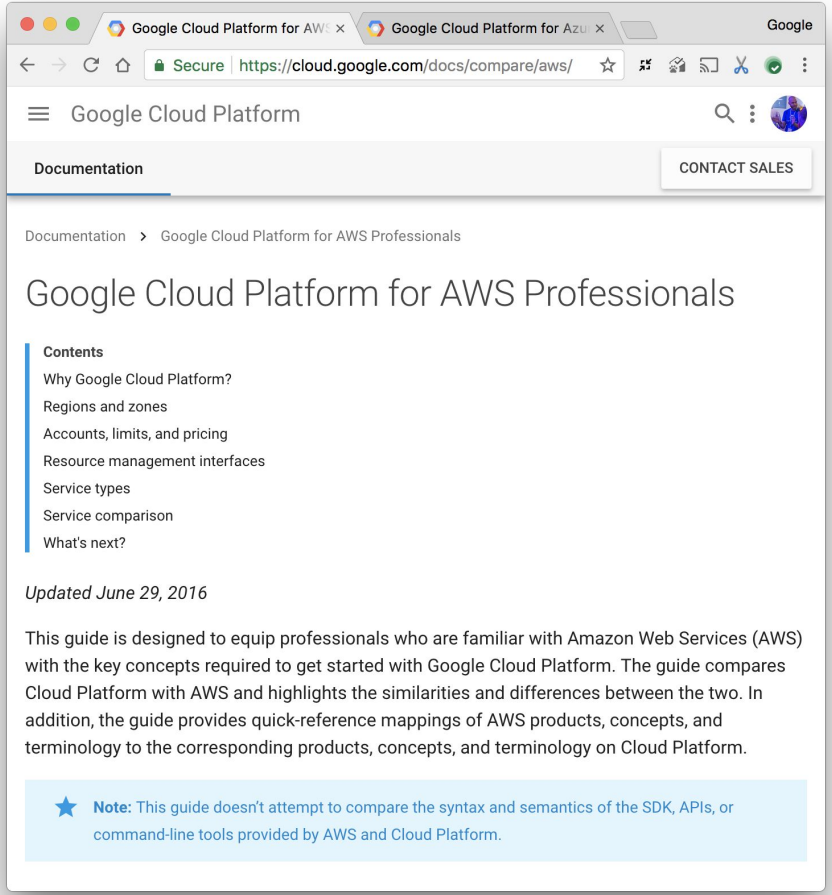
Code labs: https://codelabs.developers.google.com/
Covers all things Google, but also GCP and Kubernetes extensively

Google SRE book: https://landing.google.com/sre/book/index.html
This book, but online and free

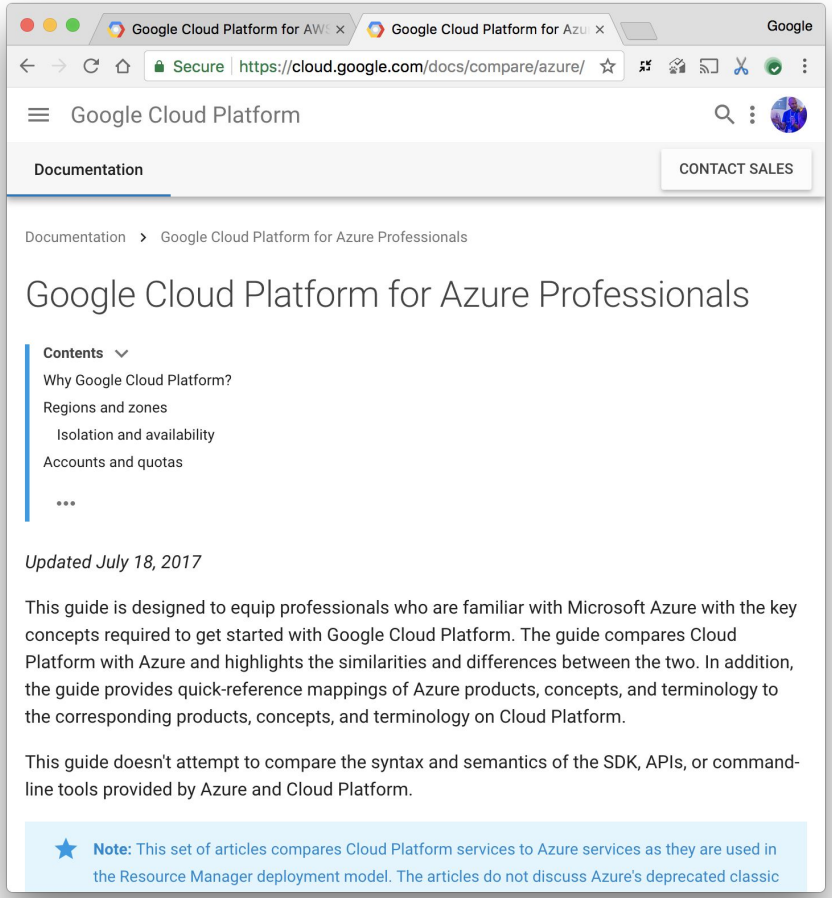
Research papers: https://research.google.com/
Lots of research, including several on container technology: 1, 2, and 3

Solution papers: https://cloud.google.com/docs/tutorials
Hidden gem. Lots of good architecture solution papers.

Google Cloud Next videos: https://goo.gl/c1Vs3h
All recorded content from Google Cloud Next 2017



https://cloud.google.com/docs/compare/aws/



https://cloud.google.com/docs/compare/azure/

Questions?