

A Tour of Google Cloud Platform



Getting Started with Google's Infrastructure and Platform

+ColinSu

Developer Expert, Google Cloud Platform

Software Architect, Tagtoo



Google Cloud Platform



Slides are available on [slideshare.net
goo.gl/ARRyaq](http://slideshare.net/goo.gl/ARRyaq)



Outlines

- > Overview of Google Cloud Platform
- > Google App Engine - Platform-as-a-Service in Google Cloud
- > App Services in Google Cloud Platform
- > Google Compute Engine - Infrastructure-as-a-Service in Google Cloud
- > BigData Lab - What we did in GCDC 2013?





Overview

Google Cloud Platform



Google Cloud Platform



Why do we need Cloud?

Google Cloud Platform

> Why

- + too much data that your PC/servers couldn't store
- + too much computation your PC/servers couldn't deal with
- + your PC/servers are hard to scale

> Why not

- + need a website
- + it sounds cool



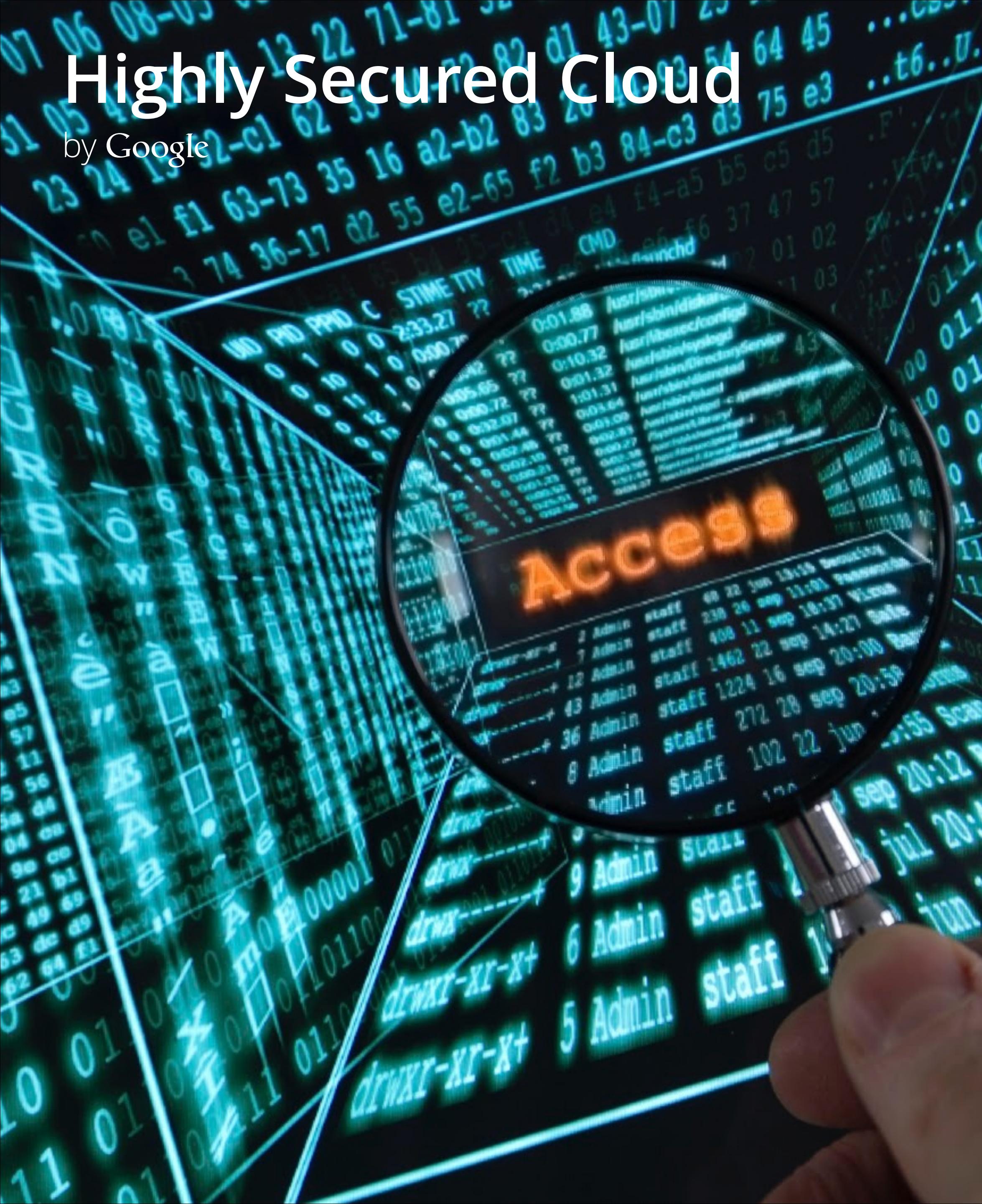


Why should we use Google Cloud Platform?

Google Cloud Platform

Highly Secured Cloud

by Google



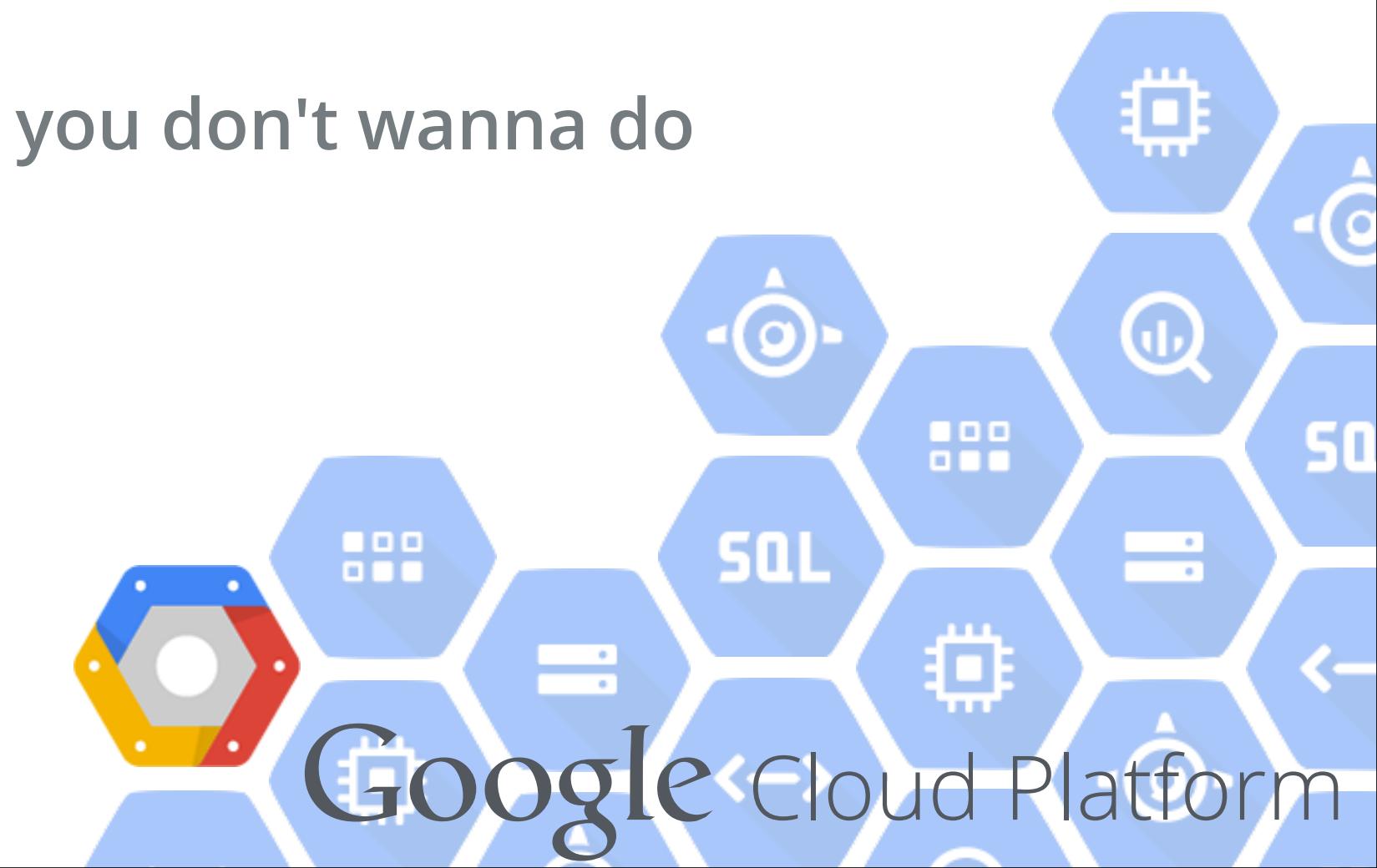
- > Access Control
 - public, private networks
 - block storage
- > Encryption
 - all block storages will be encrypted on the cloud, no worries about leaks
- > Encapsulation
 - all instances, virtual machines, networks or any resources will be encapsulated to prevent any other ones take over your precious stuffs



Powered by Google



- > You will be using Google's Infrastructure
 - Virtual Machines
 - Networks
 - Storage
- > And be placed in a safe place
 - Google's Data Center
- > And Google will do those for you
 - Scaling
 - Migrating
 - Maintenance
 - Take over anything you don't wanna do



So What is Google Cloud Platform

- > The best way how Google share their
 - + Cloud Infrastructure
 - + Cloud Knowledge
 - + Cloud Engineers
- > Your own data center, with minimum cost as possible as it could be



Google Cloud Platform

Google Cloud Family

Computing



Compute Engine



App Engine

Storage



Cloud SQL



Datastore



Cloud Storage

App Services



Cloud Endpoints



Big Query



Cloud Developer Console

< API Project



Overview

APIs & auth

Permissions

Settings

Support

Compute Engine

VM instances

Disks

Schemas

Images

Networks

Metadata

Load balancing

Zones

Operations

Quotas

Cloud Storage

Cloud SQL

BigQuery

Cloud Development

Networking

Create a new Instance

Name

lowercase, no spaces

Description

Optional

Tags

Optional

Metadata

Optional

comma separated

key

value

Location and resources

Zone

us-central1-a

Machine type

n1-standard-1

Boot source

New persistent disk from image

Note: Google-provided kernels are no longer supported in the Compute Engine v1 API. You must select an image that includes a valid OS kernel and boot loader. See the [Compute Engine v1 Migration Guide](#) to learn how to migrate your old images to Compute Engine v1.

Image

Choose an image

Additional disks

Optional

Choose a disk

- > Manage all API services on Google Cloud

(e.g. Translation API, Prediction API, Maps API...)

- > Compose equivalent commands for:
Command-line tools (Google Cloud SDK)
RESTful API

- > Dashboard for monitoring all resources on Google Cloud Platform



Google cloud SDK

```
(littleq@littleq-macbook-pro:s000) [731] o) gcloud components list
The following are the components available through the Google Cloud
SDK. You may choose to install one or more of the pre-configured
packages (which contain everything you need to get started), and/or
any of the individual components below.

|-----+-----+-----+-----+-----+-----+-----+
| Status | Name          | Description | Version | Last Update | Size | S
|-----+-----+-----+-----+-----+-----+-----+
| Not Installed | Cloud SDK for Go Developers | Go | v0.10.0 | 2015-07-06 | 1.1M | S
| Not Installed | Cloud SDK for Java Developers | Java | v0.10.0 | 2015-07-06 | 1.1M | S
| Installed     | Cloud SDK Core Command Line Tools | Core | v0.10.0 | 2015-07-06 | 1.1M | S
| Installed     | Cloud SDK for Python and PHP Developers | Python | v0.10.0 | 2015-07-06 | 1.1M | S
|-----+-----+-----+-----+-----+-----+-----+
|-----+-----+-----+-----+-----+-----+-----+
|-----+-----+-----+-----+-----+-----+-----+
| Status | Name          | Description | Version | Last Update | Size | S
|-----+-----+-----+-----+-----+-----+-----+
| Not Installed | App Engine SDK for Go (Mac OS X, x86_64) | Go | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | App Engine Launcher Application for Mac | Java | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | App Engine SDK for Java | Java | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | App Engine SDK for Python and PHP | Python | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | Big Query Command Line Tool | SQL | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | Cloud SDK Core Libraries | Core | v0.10.0 | 2015-07-06 | 1.1M | S
| Installed     | Cloud SQL Admin Command Line Interface | SQL | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | Cloud Storage Command Line Tool | Storage | v1.9.1 | 2015-07-06 | 1.1M | S
| Installed     | Compute Engine Command Line Tool | Compute | v1.9.1 | 2015-07-06 | 1.1M | S
|-----+-----+-----+-----+-----+-----+-----+
```

- > Install/uninstall/upgrade all command-line tools related to Google Cloud Platform
 - > Notification for new release of any Cloud SDK component
 - > Automatization





Storage

Storage Services in Google Cloud Platform



Google Cloud Platform

Cloud SQL

- > It's MySQL, but managed by Google
- > Relational Data Storage on Google Cloud
- > Use Cases
 - + LAMP Applications
 - + Google App Engine



Cloud Datastore

- > Non-relational database (NoSQL)
- > Schema-less data
- > Use Cases
 - + Highly scalable application



Cloud Storage

- > Protected
Your data is protected at multiple physical locations
- > Strong, configurable security
OAuth or simple access control on your data
- > Multiple usages
 - + **Serve static objects directly**
 - + **Use with other Google Cloud products (Bridge)**





App Services

Power Up Your Applications in Google Cloud Platform



Google Cloud Platform

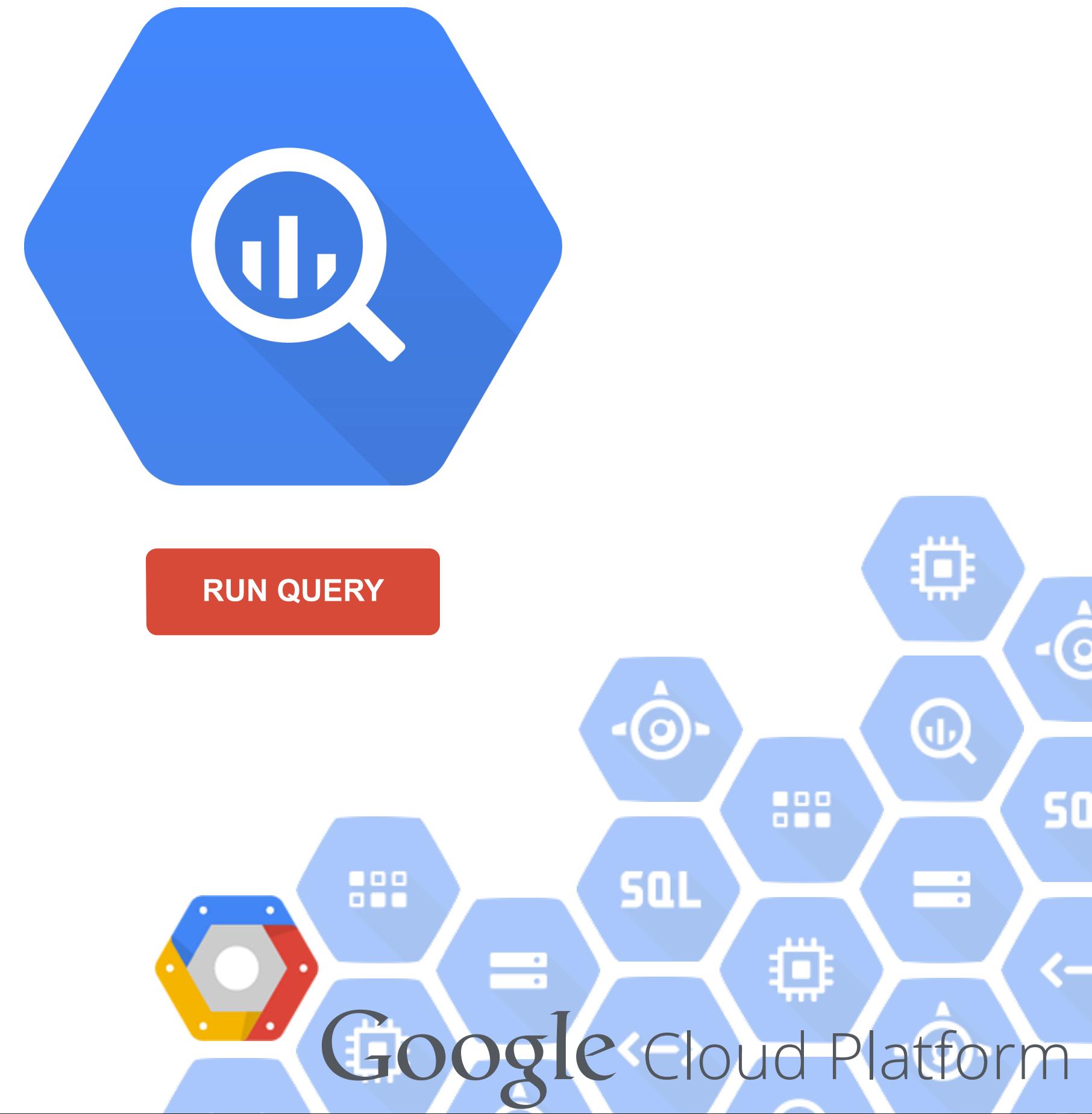
App Services in Google Cloud Platform

- > Data Analysis Tool
 - + BigQuery
 - + Google Prediction API
- > Cloud Endpoints
- > Google Cloud DNS



Big Query

- > Analyze terabytes of data, just a click of a button
- > Super-fast, SQL-like queries
- > Convenient import/export mechanism



BigQuery Browser Tool

Compose Query

```
SELECT timestamp, title, COUNT(*) as cnt
FROM publicdata:samples.wikipedia
WHERE LOWER(title) CONTAINS 'speed' AND wp_namespace = 0
GROUP BY title, timestamp ORDER BY cnt DESC LIMIT 20;
```

RUN QUERY Query complete (4.1s elapsed, 11.5 GB processed)

Query Results

Row	timestamp	title	cnt
1	1196276720	New Hampshire Motor Speedway	2
2	1187028345	Speedway World Team Cup	2
3	1043861144	Speed of gravity	2

Download as CSV Save as Table

- > Previewing of data
- > Statistics of tables
- > History/Cached Result
- > Save query result as another BigQuery table



Popular Languages on Github

BIGQUERY

```
SELECT repository.language, COUNT(repository.language) as num
FROM [publicdata:samples.github_nested]
GROUP BY repository.language
ORDER BY num DESC
LIMIT 10
```

1.6s elapsed, 12.8 MB processed



More Ways To Use BigQuery

- > Command-line Tool
 - a full-featured command-line tool is included in Google Cloud SDK, called `bq`
- > RESTful API
 - a set of APIs is ready for helping you to control all components and data on your BigQuery
- > BigQuery Connector for Excel
 - Microsoft Excel? No problem, we have an extension for your excel to query over your excel data
- > Third-party Tools
- > Make your own

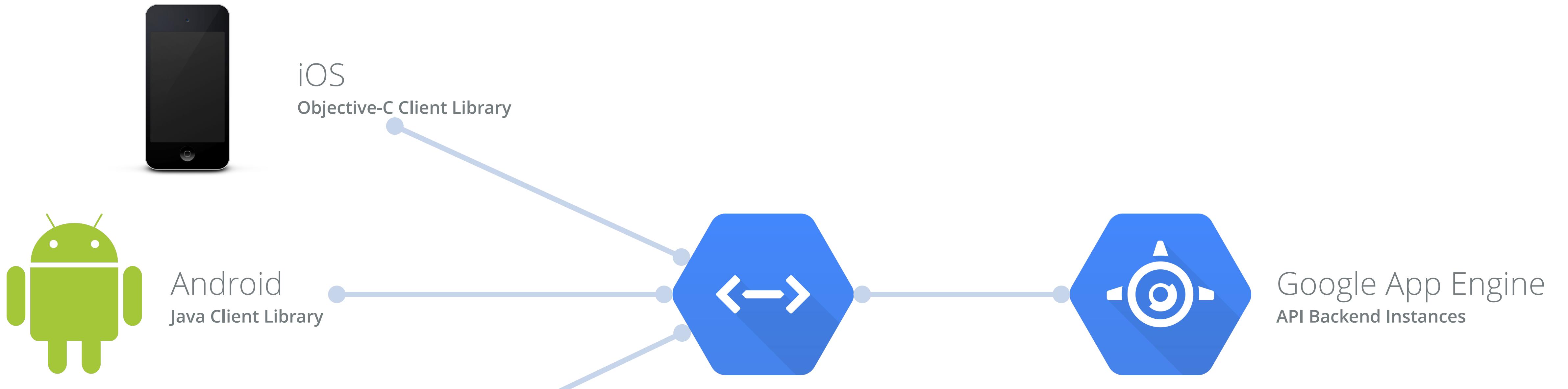


Cloud Endpoints

- > generate **APIs** and **client libraries** from an App Engine application
- > make it easier to share web backend for mobile applications



Cloud Endpoints Architecture



Use Cases



- > Website
 - solid integration, define APIs in Google App Engine application and generate JavaScript client library with Endpoints, no more handmade AJAX
- > API Server
 - define APIs with Endpoints API, then it will be a RESTful API service immediately
- > Mobile Applications
 - Backend-as-a-Service**
 - define reusable APIs with GAE various services, then generate client libraries for iOS, Android and web browser to share the resources you have



Prediction API

- > Machine Learning
 - + Categorical
 - + Regression
- > Pattern-matching
- > Simple API Interface



What can you do with Prediction API

- > Recommendation System
Predict what will be liked by your users
- > Filtering spams
Categorizes messages as spam or non-spam
- > Semantic Analysis
Know how your users feel, given your comments



Prediction Model Gallery

- > Language Identifier
- > Tag Categorizer
- > Sentiment Predictor



Try Prediction

Select a model:

Sentiment Predictor

Analyzes the sentiment of a short English-language text snippet.

Phrase:

hello world

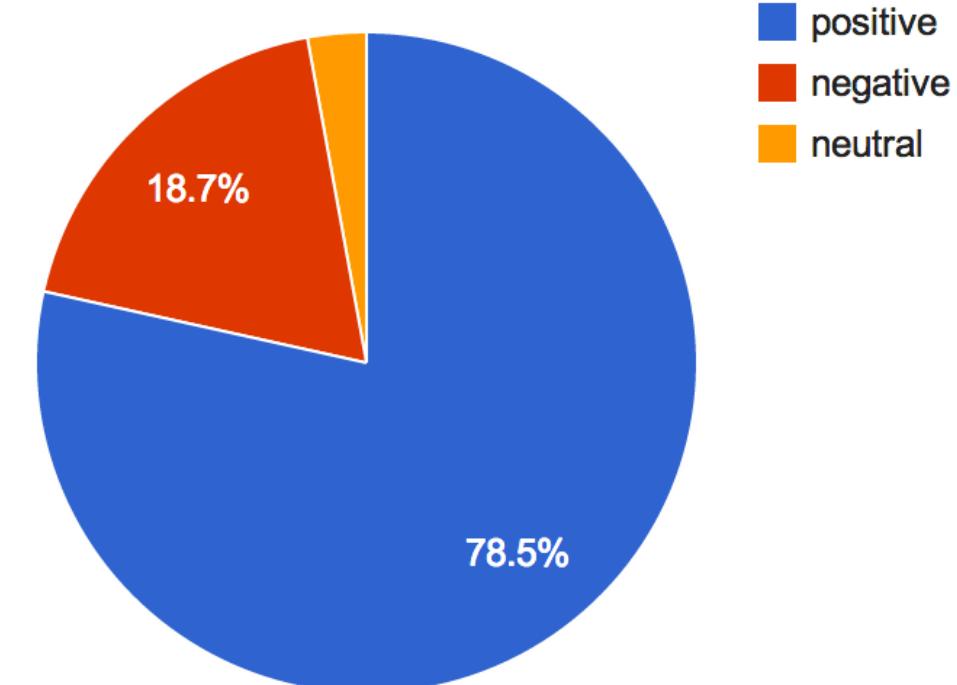
Enter a text comment for sentiment analysis

Predict

Results:

positive

[Switch chart type](#)





Google App Engine

Platform-as-a-Service in Google Cloud Platform



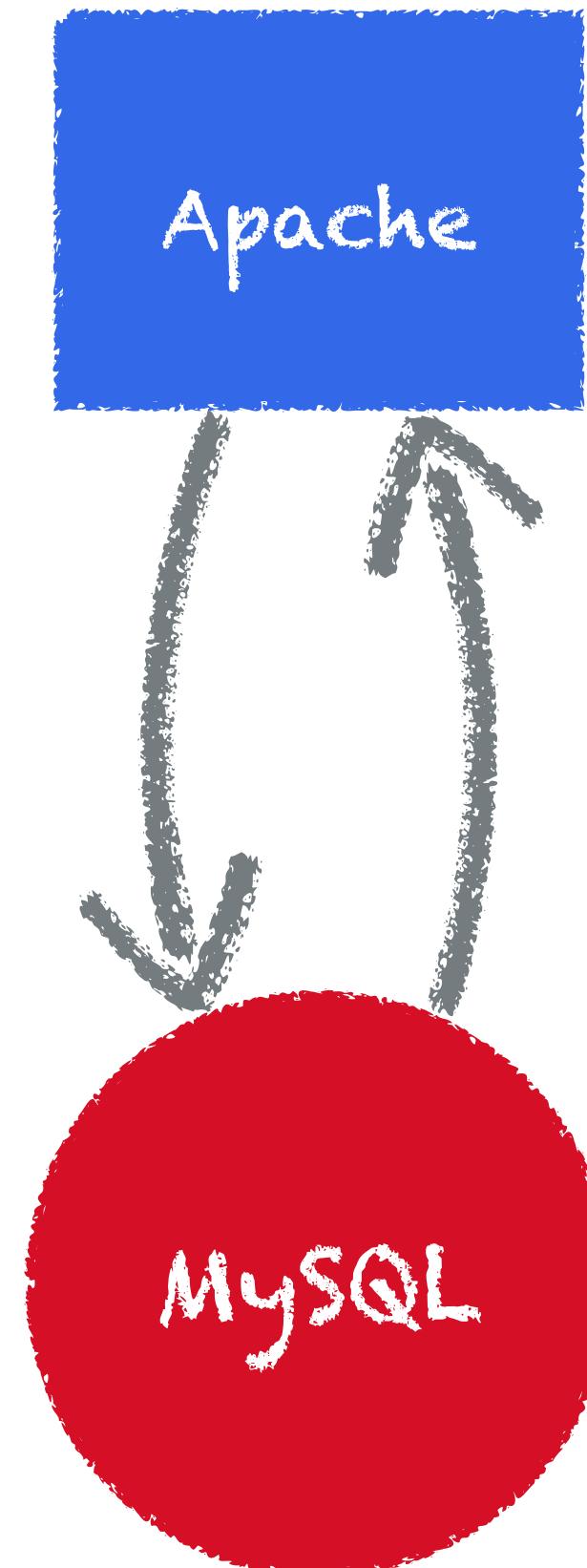
Google Cloud Platform

How did you build a full-functional web service?

a scalable, high-performance, fault-tolerant service

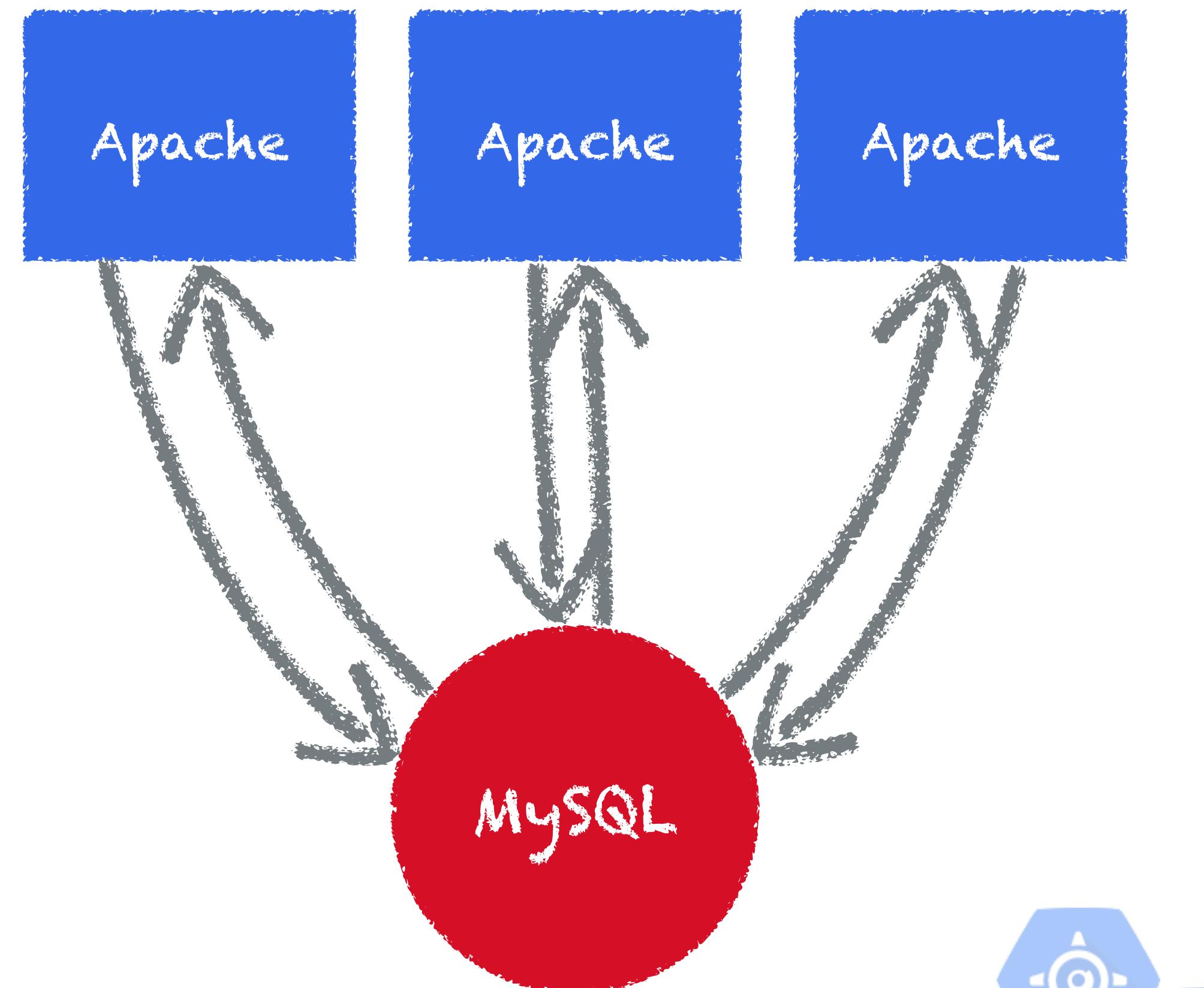
Traditional Way: LAMP

- > LAMP
 - + Linux
 - + Apache
 - + MySQL
 - + Programming Languages
- > Failed: not scalable



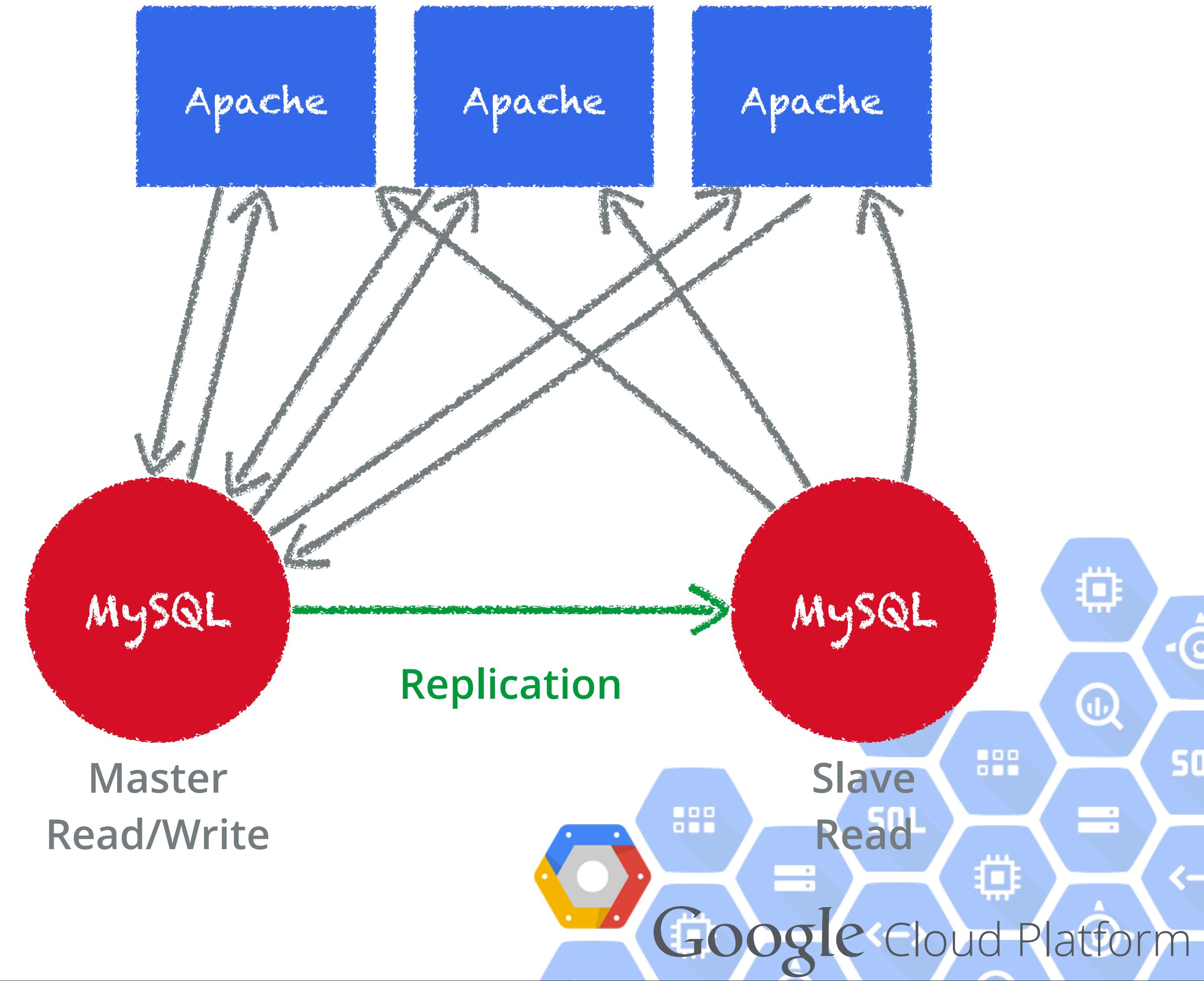
You Need Load Balancing

- > Power up Apache2 army!
- > Failed: database is alone, too busy



Now Scale Database

- > Replication (Master/Slave)
- > Failed: Master may die



Life is good, don't waste it



Platform-as-a-Service

- > You should care
 - + Application code
 - + Automated Scaling
- > You should not care
 - + Server management (networking, cores, memory, disks...)
 - + Bootstrapping
 - + Deployment



How PaaS Work

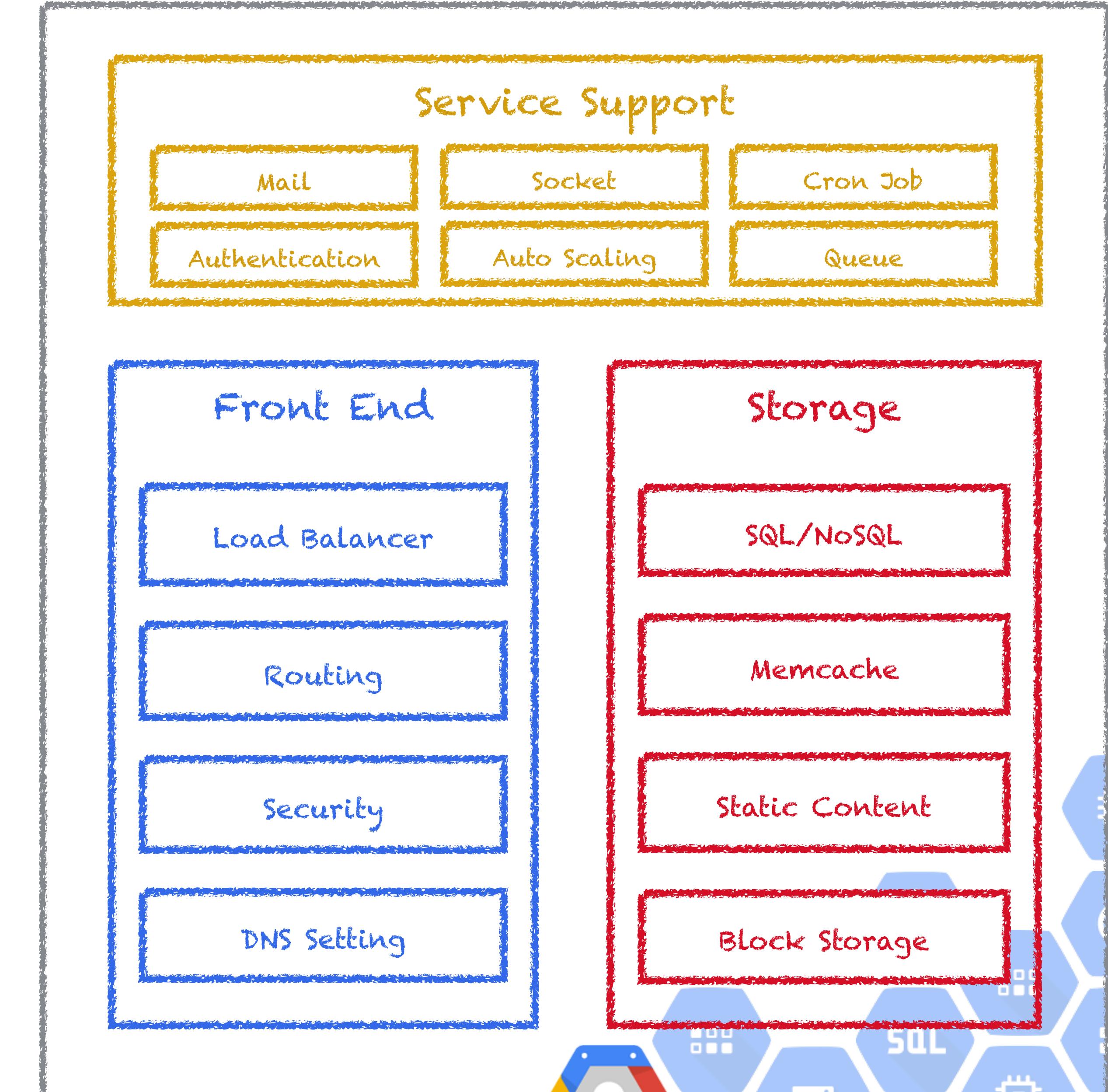
- > Application code gets **executed** (Runtime)
- > Static Content gets **served** (CDN)
- > Data gets **stored** (Database)
- > Server gets **secured** (Sandbox)
- > Service gets scaled (and automatically)



Your Responsibility

Application
Code

Google's Responsibility



Google App Engine

- > Sandboxed containers with various runtimes
- > Easy to build
All you need to do is preparing your application code
- > Easy to run
Deploy with a single command, and it works
- > Easy to scale
scale on GAE is automated and easy to configure



Languages and Runtimes

> Java

Java Servlets interface

Support for standard interfaces to App Engine scalable services such as JDO, JPA, JavaMail and JCache

> Python

Python 2.7 and full support for any pure Python libraries, tools and frameworks

Built-in Compiled C-extension libraries are good to go

> PHP

Currently in "**Preview**" stage

But enough for your "Wordpress"

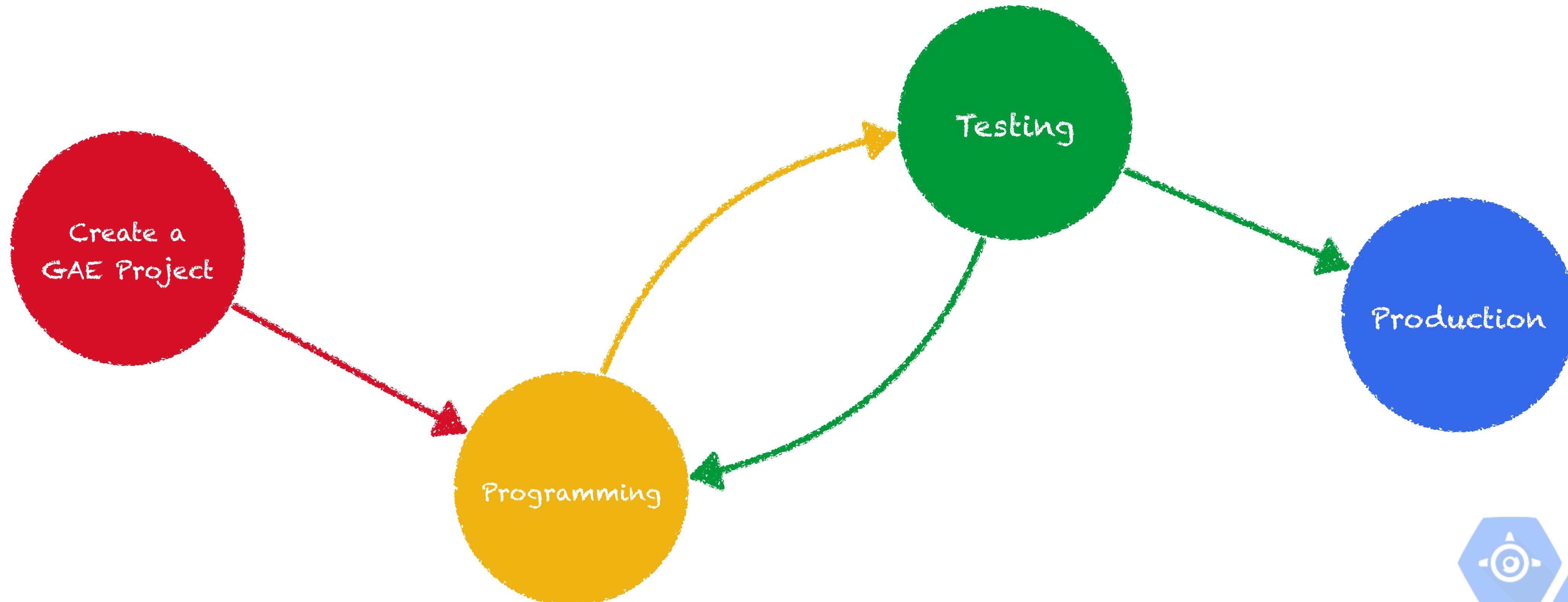
> Go

Currently in "**Experimental**" stage

automated build service included, no need to rebuild when code changed
and interface similar to the standard Go http package



Workflow of Google App Engine



Data on Google App Engine

- > Datastore
 - schema-less, scalable object data storage
 - rich data modeling API
 - SQL-like query language, GQL (Google Query Language)
- > Cloud Storage
 - strong, flexible, distributed storage service for serving or storing static files
- > Search
 - Google-like search on structured data, such as full text, numbers, dates and geographic locations
- > Memcache
 - a distributed, in-memory data cache to greatly improve your applications
- > Logs
 - programmatic access to logging system
 - a full functional control panel in Cloud Console, better than a gzip file
- > Migration/Backup Tools



Communication

- > Channel
 - Create a persistent connection between your application and Google servers
 - Send messages to JavaScript clients in **real-time**
- > Mail
 - send email messages **on behalf of admin or Google account users**
 - receive mails at various custom email addresses
- > URL Fetch
 - Efficiently issue **HTTP or HTTPS requests** on your web application
- > Outbound Socket
 - Socket support without requiring any special App Engine libs or any special App Engine (`import socket` in Python)
- > XMPP
 - Enable your application to send and receive chat messages to/from any **XMPP-compatible messaging service** (e.g. Facebook Chat, previous Google talks...)



Process Management

- > Task Queue
 - allow your application to synchronize user requests, and organize them to be executed later
- > Scheduled Task (Cron Job)
 - configure regular tasks at scheduled times or regular intervals



Computation

- > Modules

Create instances for exempting from request deadlines and request more memory and CPI resources for computing

- > MapReduce

optimized adaptation of the MapReduce computing model for efficient distributing computation on large data sets

- > Images API

Manipulate, combine and enhance images

Convert images between formats

Query metadata of images (height/width, colors)





Related Projects & Libraries

Extend Your App Engine to Work on Large-scale Computation



Google Cloud Platform

Pipeline

- > Connects together complex, time-consuming workflows
- > Asynchronize tasks
- > Built-in pipelines or implement your own pipelines



Create a Pipeline

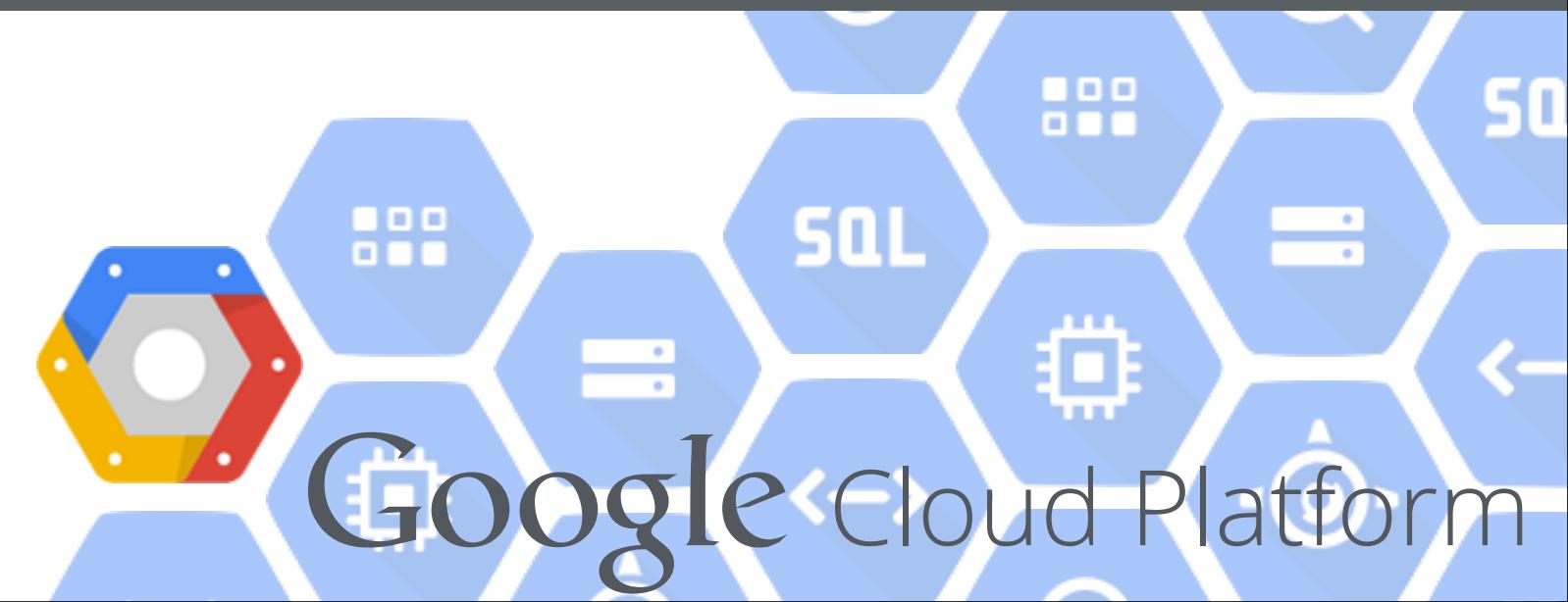
PYTHON

```
from pipeline import common

class CountReport(pipeline.Pipeline):

    def run(self, email_address, entity_kind, property_name, *value_list):
        split_counts = yield SplitCount(entity_kind, property_name, *value_list) # pipeline to gain
count result
        yield common.Log.info('SplitCount result = %s', split_counts)

        with pipeline.After(split_counts):
            with pipeline.InOrder():
                yield common.Delay(seconds=1)
                yield common.Log.info('Done waiting')
                yield EmailCountReport(email_address, split_counts) # another sending mail pipeline
```



Pipeline Console

main.CountReport - ID #92e411de07dc11e09e72414b2760c3c1
http://localhost:8080/_ah/pipeline/status?root=92e411de07dc11e09e72414b

Auto-pilot

main.CountReport
Run
Started: 3 seconds ago

pipeline.common.Delay
Run
Started: 1 seconds ago

main.SplitCount
Finalizing
Complete: 2 seconds ago

- main.LongCount
Finalizing
Complete: 2 seconds ago
- main.LongCount
Finalizing
Complete: 2 seconds ago
- pipeline.common.Append
Finalizing
Complete: 2 seconds ago
- pipeline.common.Log
Finalizing
Complete: 2 seconds ago
- pipeline.common.Log
Finalizing
Complete: 1 seconds ago

* main.UselessPipeline
Retry
Whoops, I need to retry
Will run: 13 seconds from now
[Console Home](#)

main.EmailCountReport
Waiting

pipeline.common.Log
Waiting

pipeline.common.Log
Waiting

Example console

This is an example console for a pipeline

This is where you would put all kinds of graphs and other useful things, like this:



main.CountReport - ID #db9713c207dc11e0b95c414b2760c3c1
http://localhost:8080/_ah/pipeline/status?root=db9713c207dc11e0b95c414b

Auto-pilot

main.CountReport
Done
Run time: 17.053 seconds
Complete: 21 seconds ago

main.SplitCount
Done
Run time: 2.594 seconds
Complete: 35 seconds ago

pipeline.common.Log
Done
Run time: 1.383 seconds
Complete: 36 seconds ago

pipeline.common.Delay
Done
Run time: 2.294 seconds
Complete: 34 seconds ago

pipeline.common.Log
Done
Run time: 1.331 seconds
Complete: 35 seconds ago

pipeline.common.Log
Done
Run time: 1.257 seconds
Complete: 33 seconds ago

main.EmailCountReport
Done
Run time: 1.208 seconds
Complete: 33 seconds ago

main.UselessPipeline
Done
Pretending to fail, will retry shortly.
Run time: 1.191 seconds
Complete: 21 seconds ago
[Console Home](#)

pipeline.common.Log
Done
Run time: 1.531 seconds
Complete: 20 seconds ago

Children

- main.SplitCount
- pipeline.common.Log
- pipeline.common.Delay
- pipeline.common.Log
- main.EmailCountReport
- main.UselessPipeline
- pipeline.common.Log



MapReduce Library

https://developers.google.com/appengine/docs/python/dataproCESSing/mapreduce_library

- > Programming model for **processing large data sets** in a **parallel** and **distributed** algorithms on a cluster
- > differ from map/reduce, one of functional programming conception, but has the same idea, "**divide and conquer**"
- > Proposed by **Google**
- > Hadoop-free



MapReduce in Functional Programming

- > map()/reduce() in Python
- > map(**func(elem)**, list) -> **list**
- > reduce(**func(elem1, elem2)**, list) -> **elem**

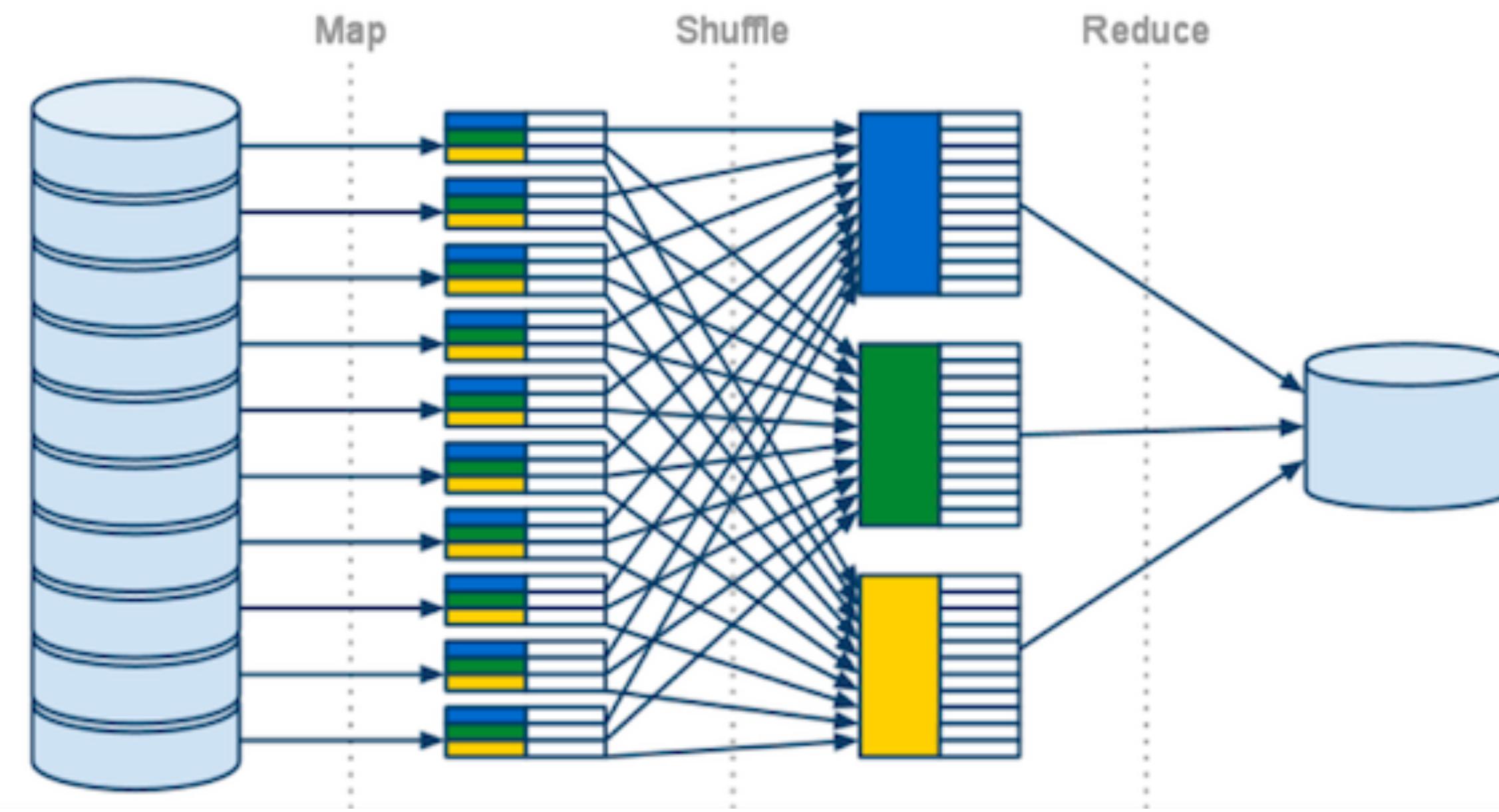
PYTHON

```
>>> map(lambda x: x*x, [1,2,3,4])  
[1, 4, 9, 16]  
  
>>> reduce(lambda x,y: x+y, [1,2,3,4])  
10
```



MapReduce: Stages

- > Map
- > Shuffle
- > Reduce



Configure a MapReduce Pipeline

PYTHON

```
class WordCountPipeline(base_handler.PipelineBase):

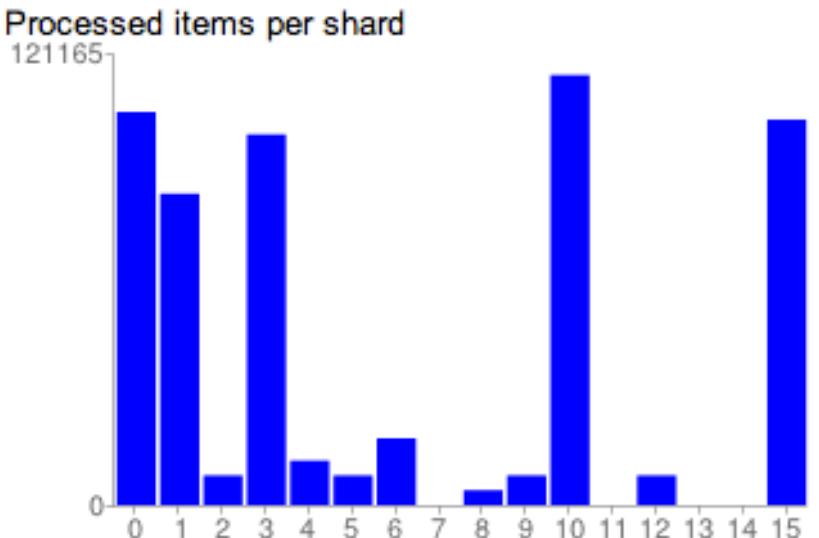
    def run(self, filekey, blobkey):
        output = yield mapreduce_pipeline.MapreducePipeline(
            "word_count", # name of mapreduce job
            "main.word_count_map", # mapper function
            "main.word_count_reduce", # reducer function
            "mapreduce.input_readers.BlobstoreZipInputReader", # input reader
            "mapreduce.output_writers.FileOutputWriter", # output writer
            mapper_params={ # parameters to supply to the input reader
                "input_reader": {
                    "blob_key": blobkey,
                },
            },
            reducer_params={ # parameters to supply to the output writer
                "output_writer": {
                    "mime_type": "text/plain",
                    "output_sharding": "input",
                    "filesystem": "blobstore",
                },
            },
            shards=16) # number of shards
        yield StoreOutput("WordCount", filekey, output)
```



MapReduce Monitoring

test-map

Job #1578697716674A90B9934



Overview

- Running
- Elapsed time: 21 days, 04:56:17
- Start time: 2/20/2014 6:26:57 PM
- input_reader: {"entity_kind":"ad.models.ProductAd","batch_size":50}

Counters

- mapper-calls: 602925 (0.33/sec avg.)
- mapper-waittime-ms: 75273461 (41.08/sec avg.)

[« Back to Overview](#) | [Abort Job](#)

Mapper status

Shard	Status	Description	Last work item	Time elapsed
0	success	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'sinyi_house:product:52747V', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'treemall:product:JcRK9Zis1u5s8Vikh8sI2LwacotjEDJ%2FNKPZohJADvyMpYcYnOZ5lZJYuJ4DRi9=', _app=u's~tagtooadex2'))	Key('ProductAd', 'treemall:product:JcRK9Zis1u5s8Vikh8sI2LwacotjEDJ%2FMgCp2blgOdmost3HiMBQTSyqQIWmc0Q=')	05:41:36
1	success	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'yahoo:product:4702351', _app=u's~tagtooadex2') to None)	Key('ProductAd', 'zakka:product:ZA144701')	00:53:56
2	running	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'yahoo:product:3249432', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'yahoo:product:4130356', _app=u's~tagtooadex2'))	Key('ProductAd', 'yahoo:product:3323350')	01:27:43
3	success	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'yahoo:product:4130356', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'yahoo:product:4465394', _app=u's~tagtooadex2'))	Key('ProductAd', 'yahoo:product:4465393')	12:35:19
4	running	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'howshopping:product:27500', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'sinyi_house:product:52747V', _app=u's~tagtooadex2'))	Key('ProductAd', 'howshopping:product:7147')	01:16:40
5	running	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'treemall:product:JcRK9Zis1u5s8Vikh8sI8WIF9L1TrmTljK4wAleylg%3D=', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'treemall:product:JcRK9Zis1u5s8Vikh8sIwagMhELJShwSO95iYIPVHU%3D=', _app=u's~tagtooadex2'))	Key('ProductAd', 'treemall:product:JcRK9Zis1u5s8Vikh8sI8WIF9L1TrmTciuHt8aNmk%3D')	00:28:24
6	running	KeyRangesIterator for Single KeyRange ASC[datastore_types.Key.from_path(u'ProductAd', u'herhuiv:product:311495797', _app=u's~tagtooadex2') to datastore_types.Key.from_path(u'ProductAd', u'herhuiv:product:9958842A')]	Key('ProductAd', 'herhuiv:product:9958842A')	01:16:40





Google Compute Engine

Infrastructure-as-a-Service in Google Cloud Platform



Google Cloud Platform

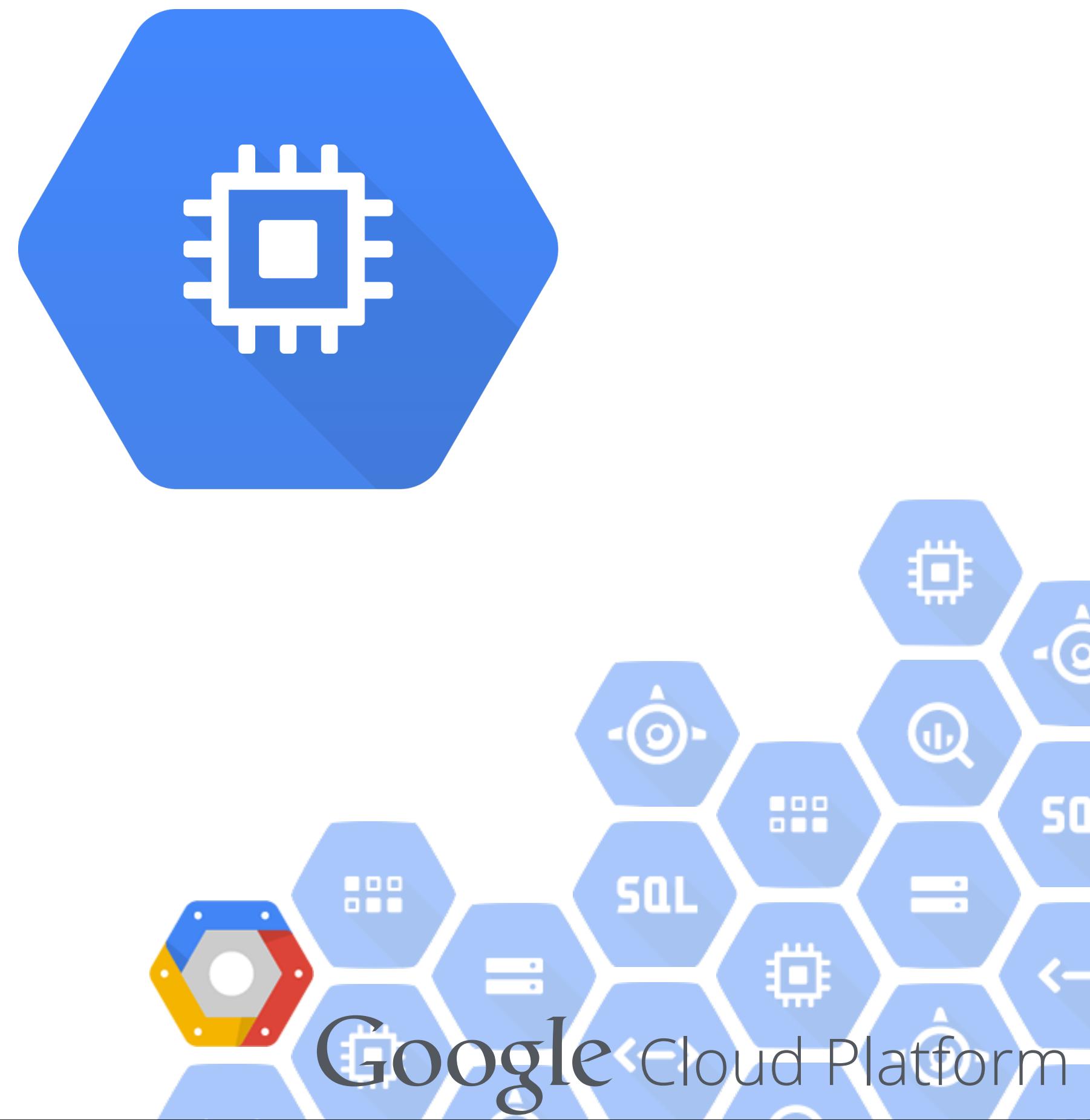
Infrastructure-as-a-Service

- > Google has resources
 - + CPU Cores
 - + Memory
 - + Networking
 - + Persistency (Disks, Snapshot, Cloud Storage...)
 - + Well-trained engineering monkeys
- > You have business and be busy



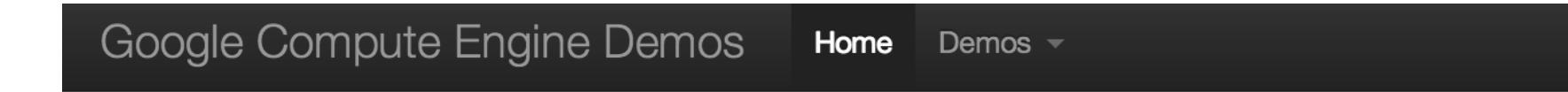
Google Compute Engine

- > High-performance virtual machines
from micro-VM to large instance
- > Powered by Google's global network
you could build a large cluster with strong and consistent bandwidth, provided by
Google
- > Load Balancing
spread incoming traffic across instances
- > Fast Bullet Reloading
quick deployment of large VMs
command-line interface
web-based console
- > Highly secured
All data written to disk in Compute Engine will be encrypted by high-class
encryption algorithm



Instances on GCE

- > KVM-based Virtual machines
- > Fast booting time
routinely takes less than 30 secs
- > Various OS support
- > Various machine types



Welcome to Quick Start!

This demo shows just how quick it is to start hundreds of instances.

[Edit](#) your Compute Engine project information.

0000 Timer: 00:00:31



<http://gce-demos.appspot.com>



Operating Systems



• • • • •



Operating Systems

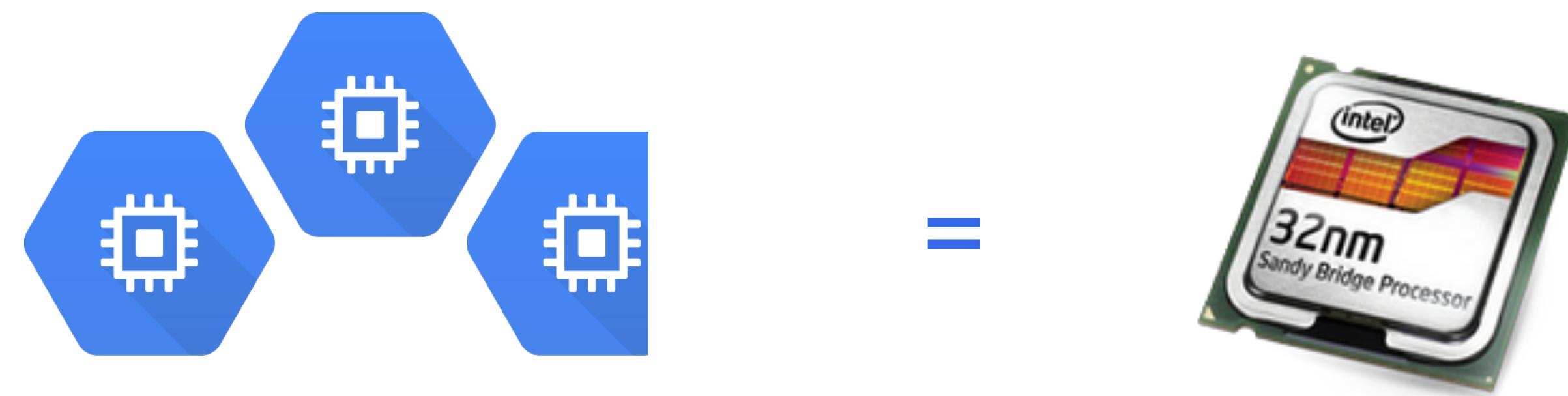


Premier Only



Google Compute Engine Units (GCEUs)

- > a unit of CPU capacity used to describe the compute power of instance types
- > 2.75 GCEUs = 1 minimum power of 1 logical core on Sandy Bridge platform



Machine Types (Shared-Core)

f1-micro

Shared Core

0.6 GB Memory

⋮

g1-small

1 Core

1.7 GB Memory

1.38 GCEUs



Machine Types (Standard)

n1-standard-n

Starts from 1 Core

Start from 3.75 GB Memory

n	Virtual CPUs	Memory	GCEUs
1	1	3.75 GB	2.75
2	2	7.50 GB	5.50
4	4	15 GB	11
8	8	30 GB	30
16	16	60 GB	60



Machine Types (High Memory)

n1-highmem-n

Starts from 2 Core

Start from 13 GB Memory

⋮

n	Virtual CPUs	Memory	GCEUs
2	2	13 GB	5.50
4	4	26 GB	11
8	8	52 GB	22
16	16	104 GB	44



Machine Types (High CPU)

n1-highcpu-n

Starts from 2 Core

Start from 1.8 GB Memory

⋮

n	Virtual CPUs	Memory	GCEUs
2	2	1.8 GB	5.50
4	4	3.6 GB	11
8	8	7.2 GB	22
16	16	14.4 GB	44



Regions & Zones



Google Cloud Platform

Live Migration

- > Transparent Maintenance
- > Auto restart instances shutdown by system events
- > During transparent maintenance, you could set GCE to handle your instance in two ways:
 - + Live migrate
affect performance in some degree
but remain your instances online (no downtime)
 - + Terminate and reboot



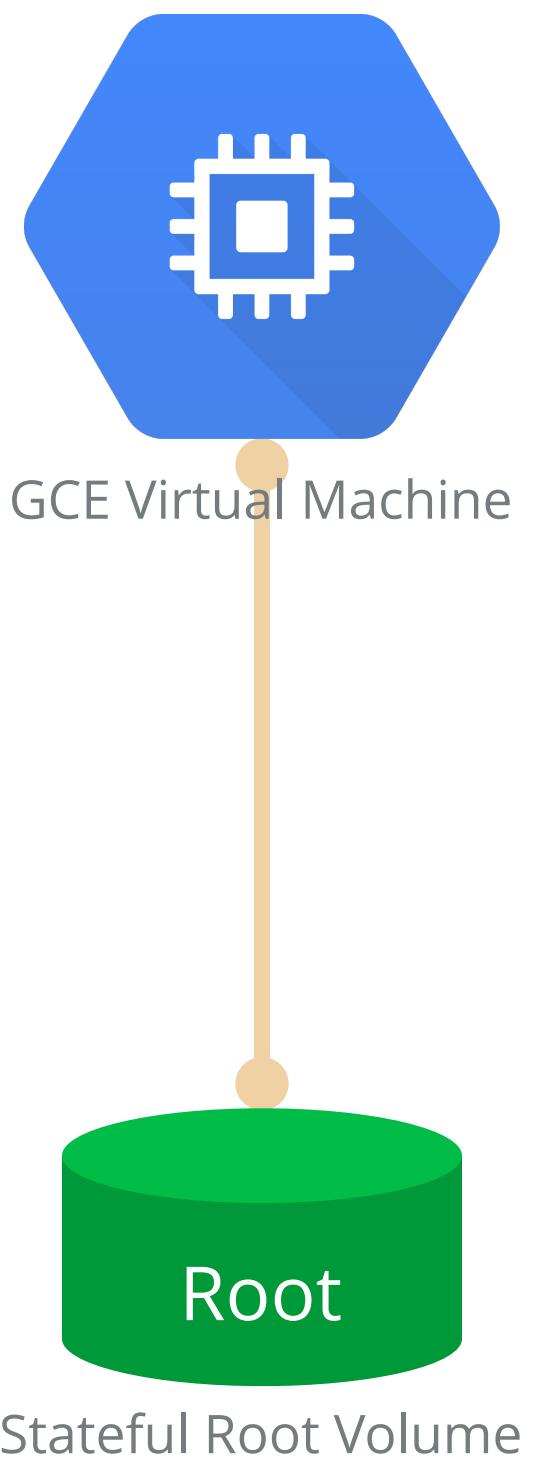
Persistent Disk

- > Virtual SCSI device
- > Block Storage
- > Persistent until deleted
- > Hot-plug to GCE instances (attach/deattach)



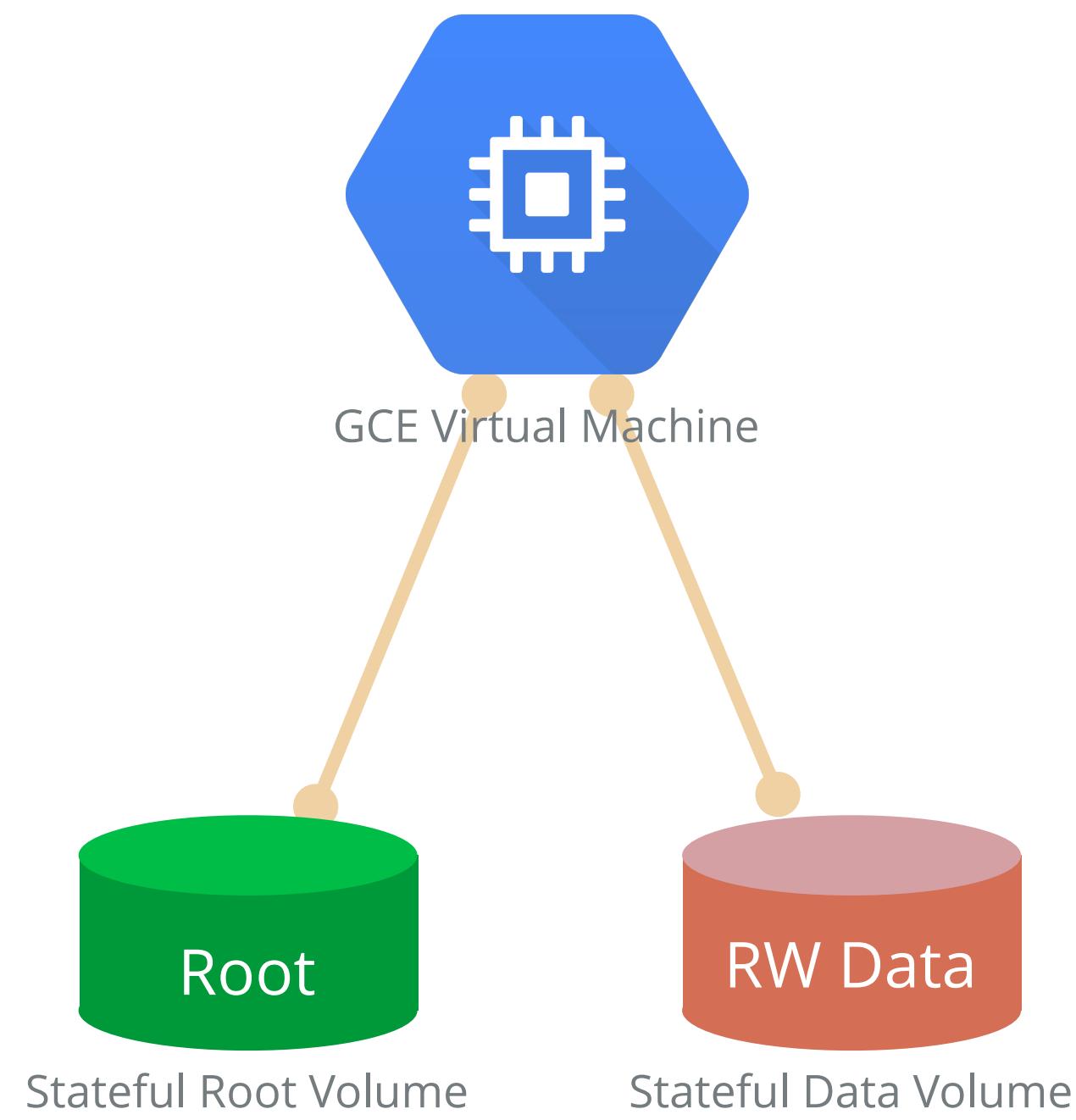
Persistent Disk Mode

- > Primary Disk: OS boot volume



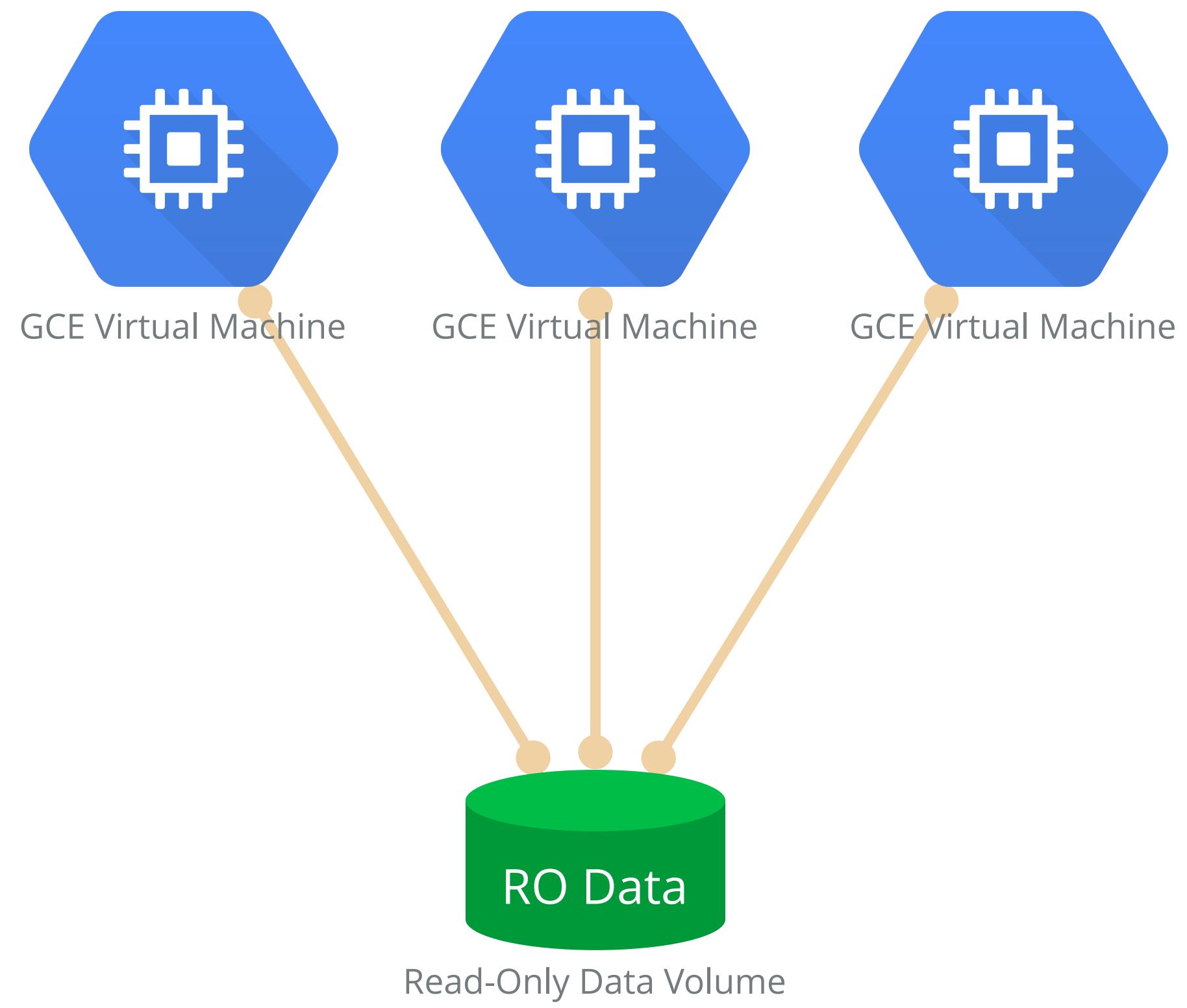
Persistent Disk Mode

- > Additional Disk: Read/Write Mode for user managed data volume



Persistent Disk Mode

- > Distribution Disk: Instant distribution of static content



Load Balancing

- > Target Pools
- > Health Checking
- > Forwarding Rules
iptables for target pools



Persistent IP Addresses

- > CGE reserved IP for instance, won't change with the reboot of VMs
- > You can promote ephemeral IP to persistent IP
- > no DNS changing anymore



Integrated Networking

- > Networking is first-class object on GCE which means you could apply/unapply it anytime easily
- > pre-defined networks before the first instance started



Multi-Region Resources

- > Those resources are global resources
 - + Images (OS Images)
 - + Snapshots
 - + Network
 - + Firewalls
 - + Routes
- > And they're also first-class object in GCE



Services Host GCE

- > Scalr
- > Right Scale



BigData Lab

One of the finalists in Google Cloud Developer Challenge 2013



Google Cloud Platform

Google Cloud Developer Challenge

- > <http://www.google.com/events/gcdc2013/>
- > developers in 6 regions envolved
- > goals
 - + effective use of Google App Engine
 - + originality of concept
 - + integration and creative use of Google Products



BigData Lab

- > Provide a simple web interface to perform 4 big data operations:
 - + Storing (Data Source)
 - + MapReduce
 - + Prediction (Machine Learning)
 - + Visualization



What Are We Using?

- > Google App Engine & Google Compute Engine
- > Cloud Endpoints
- > Google Cloud Storage
- > MapReduce Module for Google App Engine
- > Pipeline Module for Google App Engine
- > Prediction API
- > BigQuery



Demo

BigData Lab Data MapReduce Prediction Visualization Sign out Signed in Colin Su

MapReduce Lab

Function Editor **Mapper Editor** Reducer Editor Job Manager

Mapper Name WordCount Mapper

```
1 def main(line):
2     for word in line.split():
3         yield (word, 1)
```

Character Count Mapper
Synced **Delete**

WordCount Mapper
Synced **Delete**

Function Counter Mapper
Synced **Delete**

Save All Code

Create a new mapper

Save **Validate**



EOF

Thanks For Listening
+GDTaipei



Google Cloud Platform

Questions?



Google Cloud Platform