Core Building Blocks

Howard Dierking
THINKER, BUILDER, LEADER

@howard_dierking blog.howarddierking.com

Topics

Compute

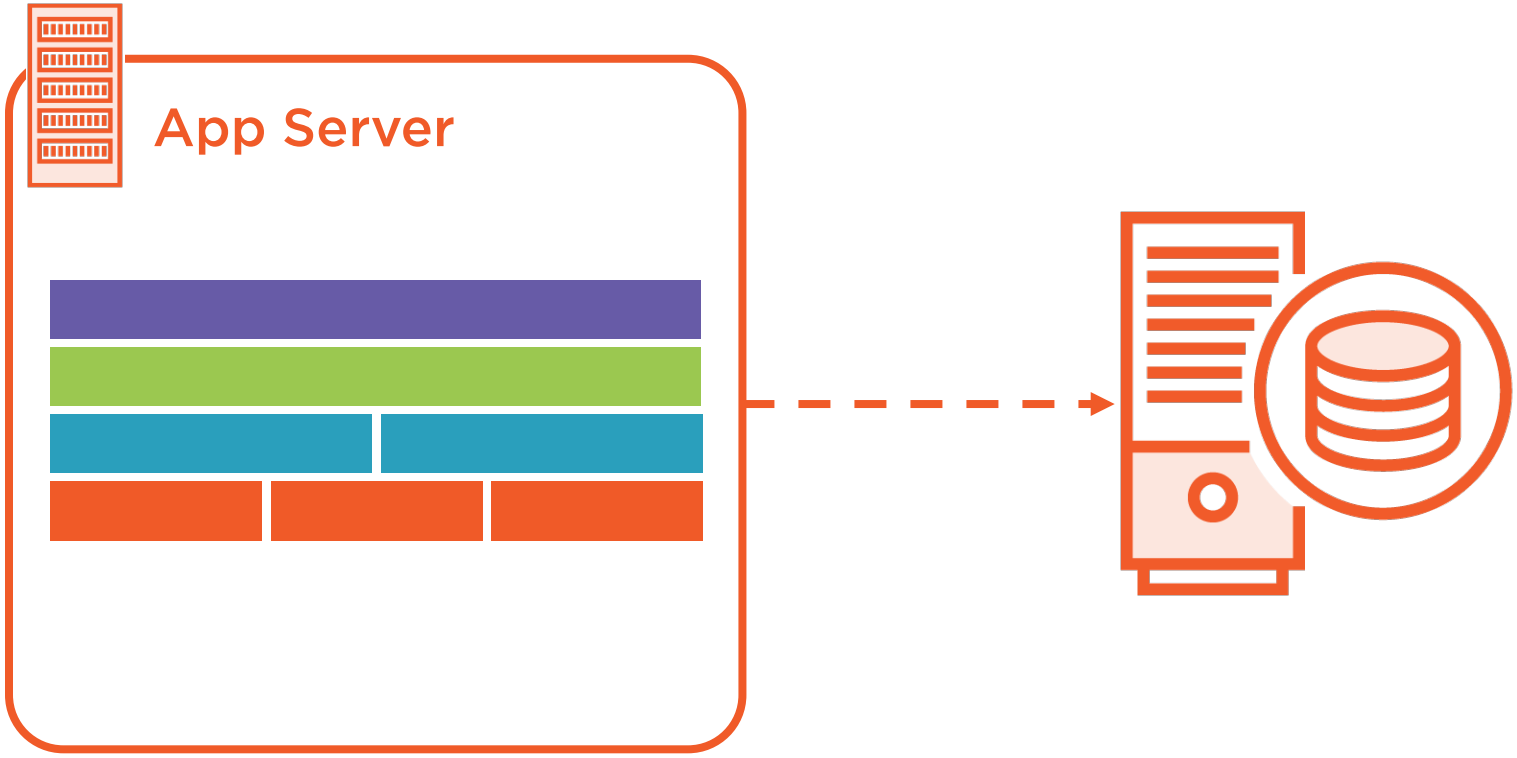
Storage

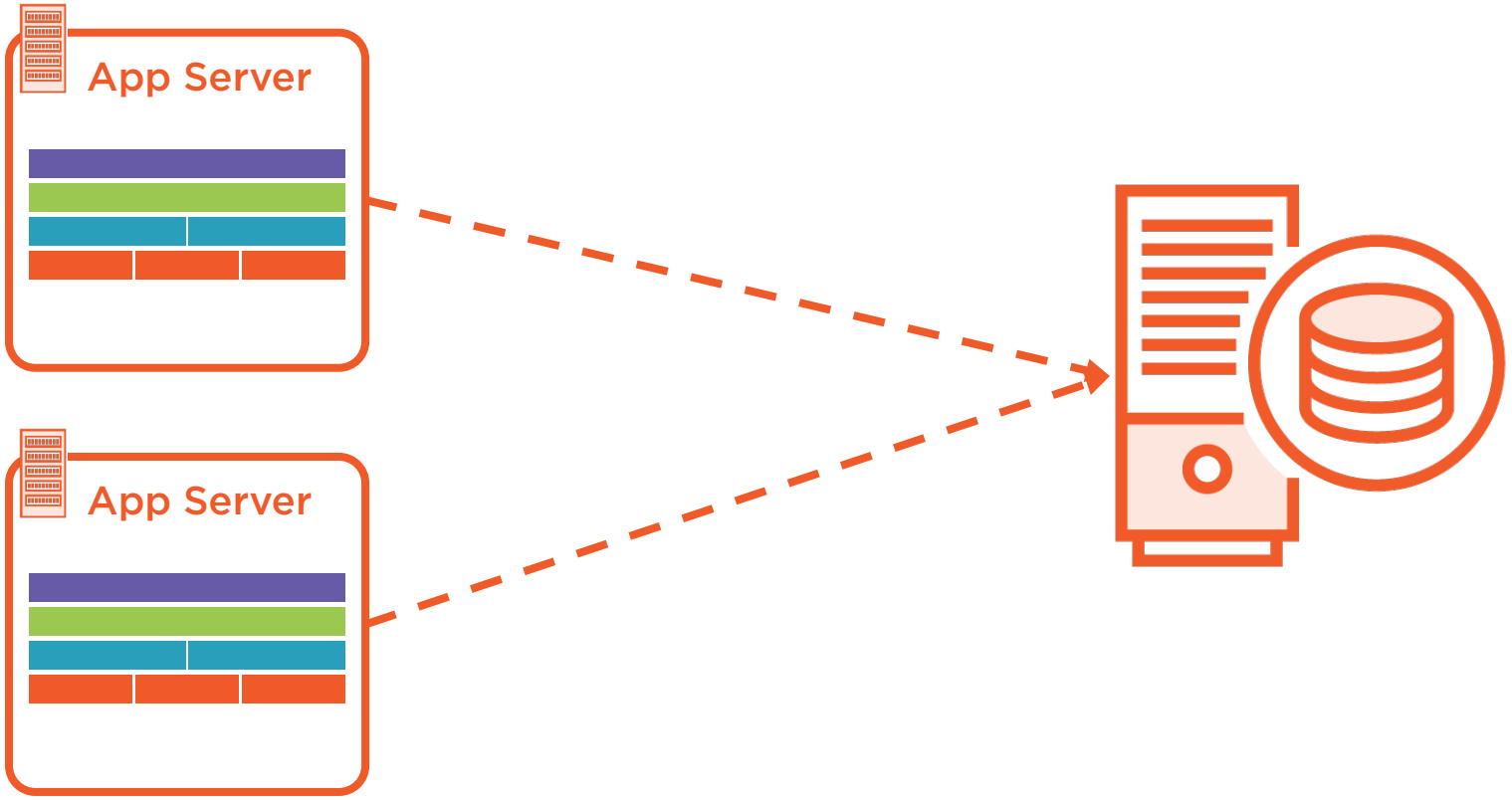
Networking

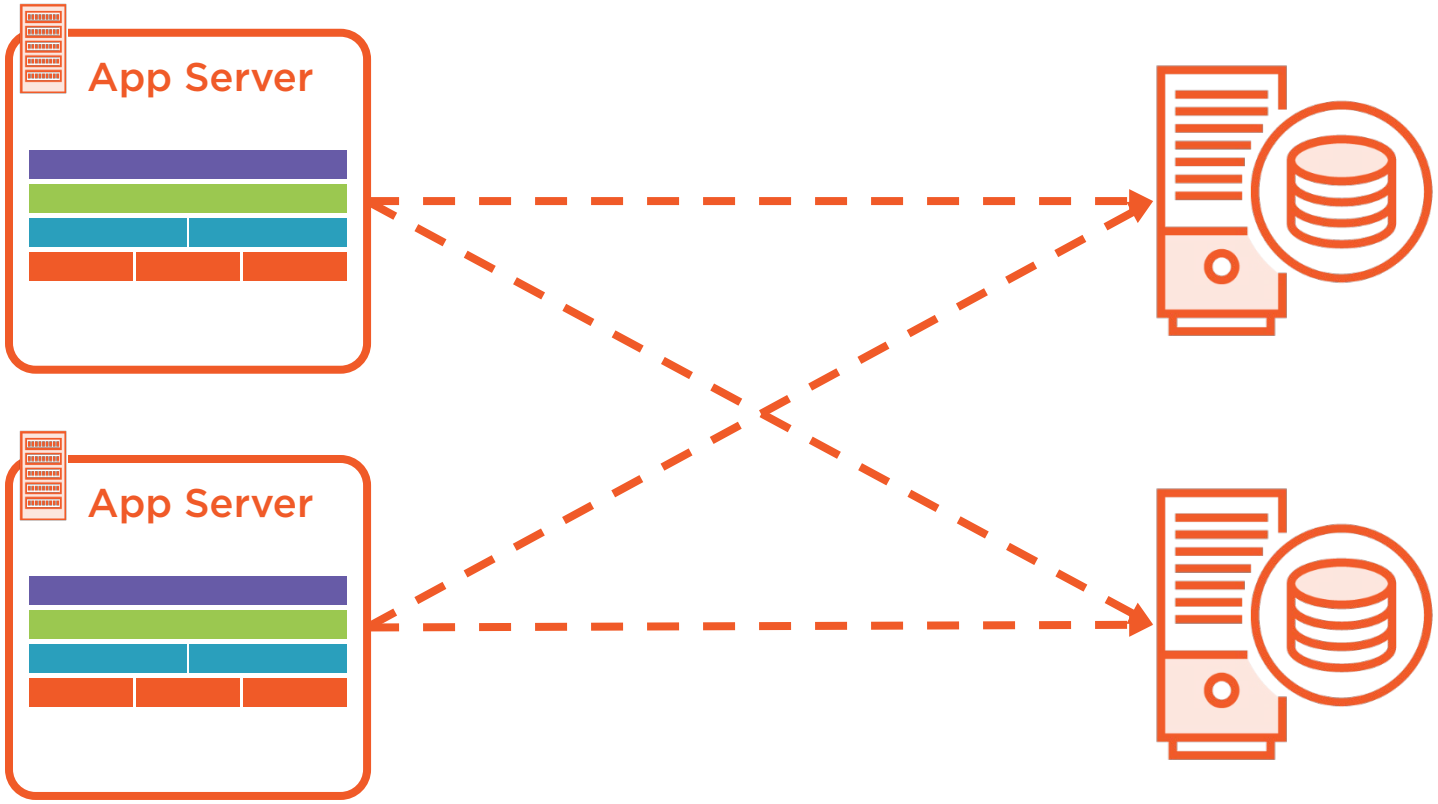
Compute

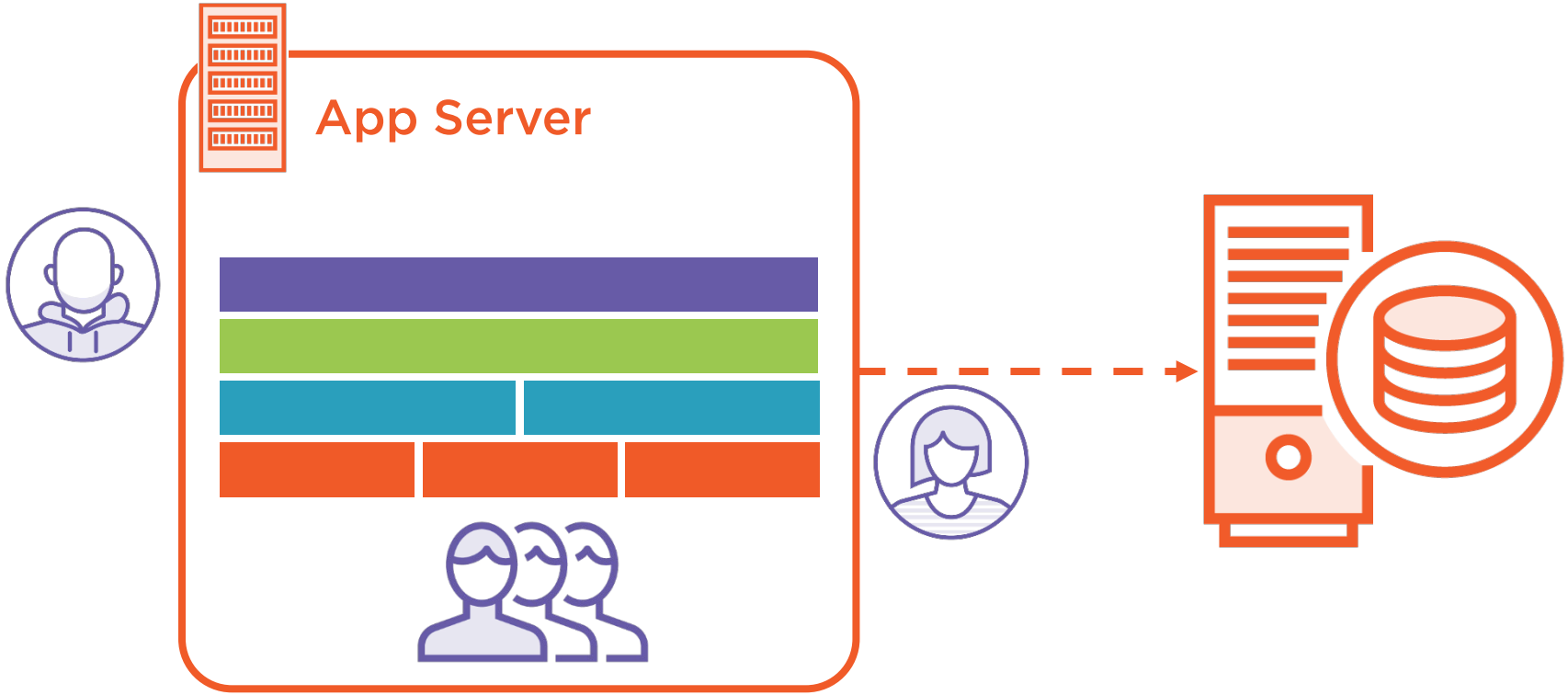

IAAS
PAAS

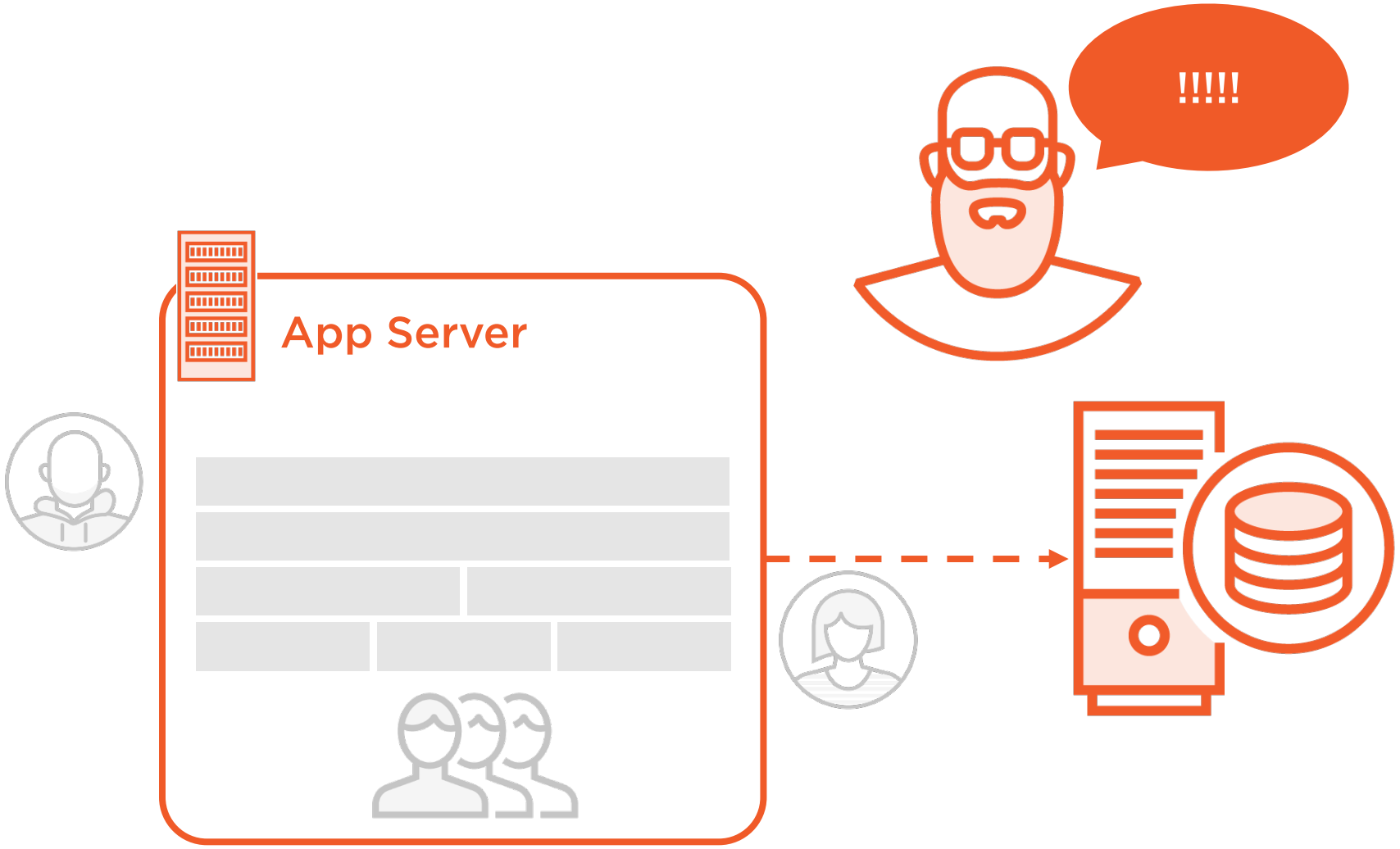
From IaaS to PaaS

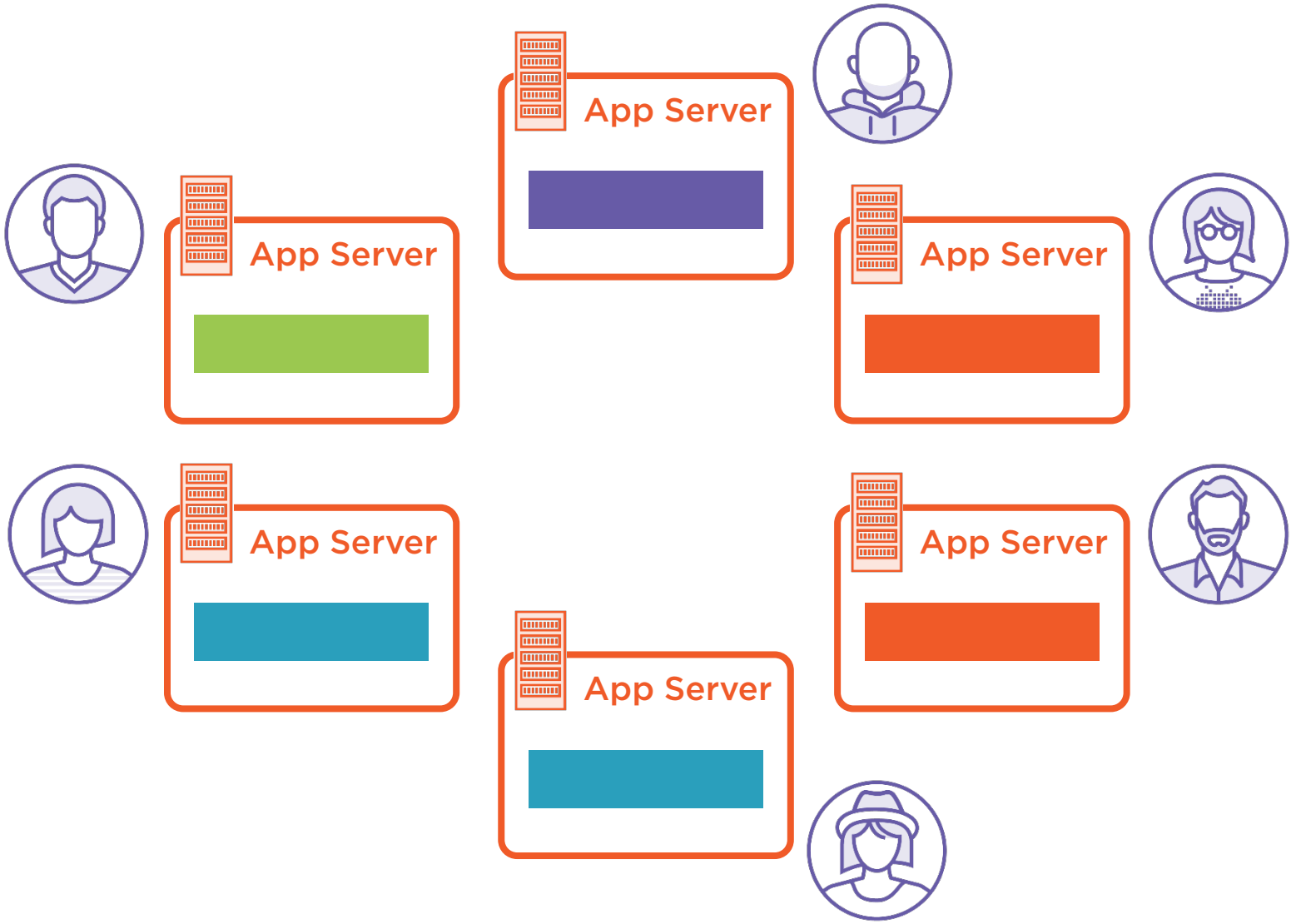


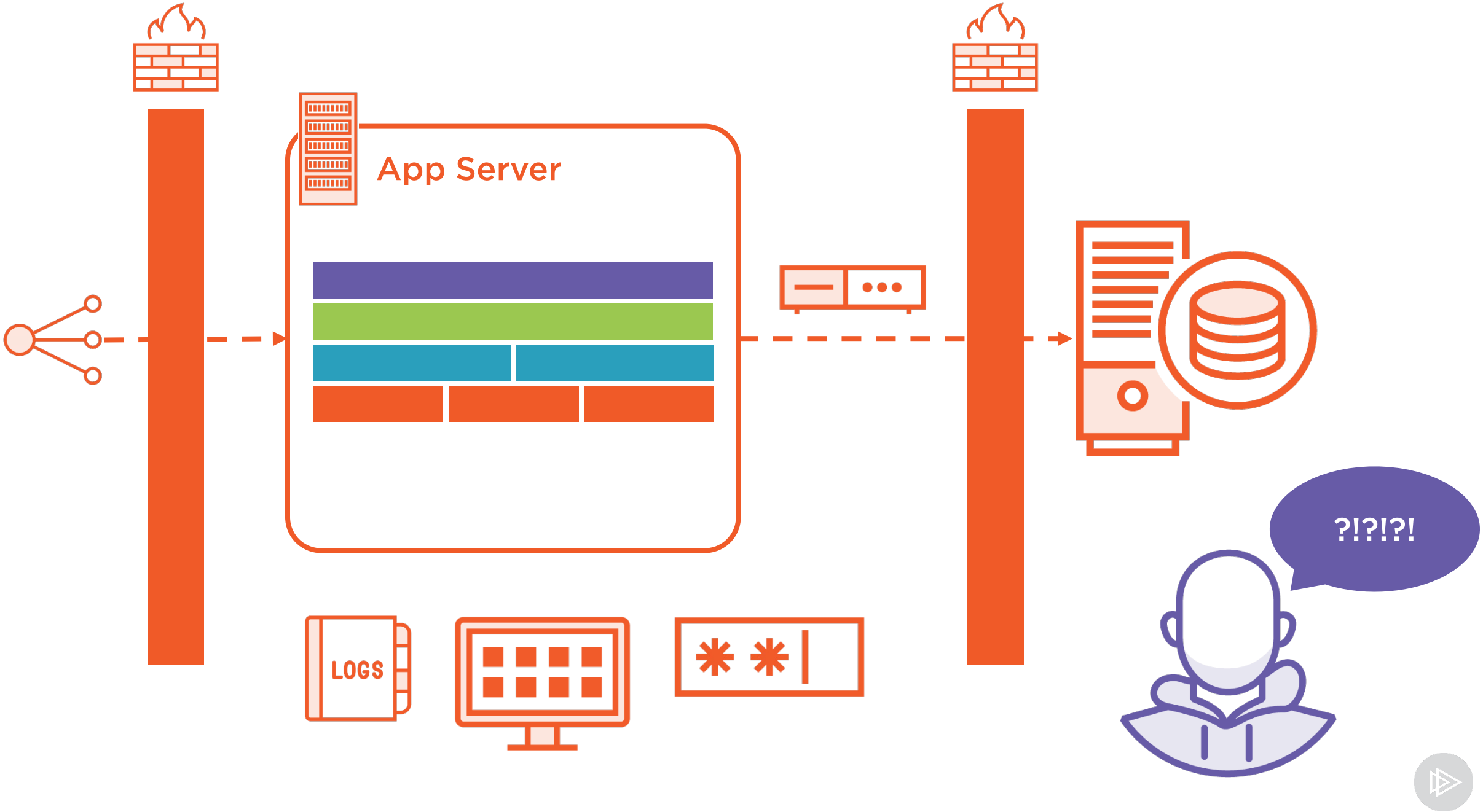

App Server
App Server


App Server
App Server


App Server


!!!!!
App Server


App Server
App Server
App Server
App Server
App Server
App Server


App Server
LOGS
?!?!?!

“Nanoservice is an antipattern where a service is too fine-grained. A nanoservice is a service whose overhead outweighs its utility.”

http://arnon.me/2014/03/services-microservices-nanoservices/

Control

Agility

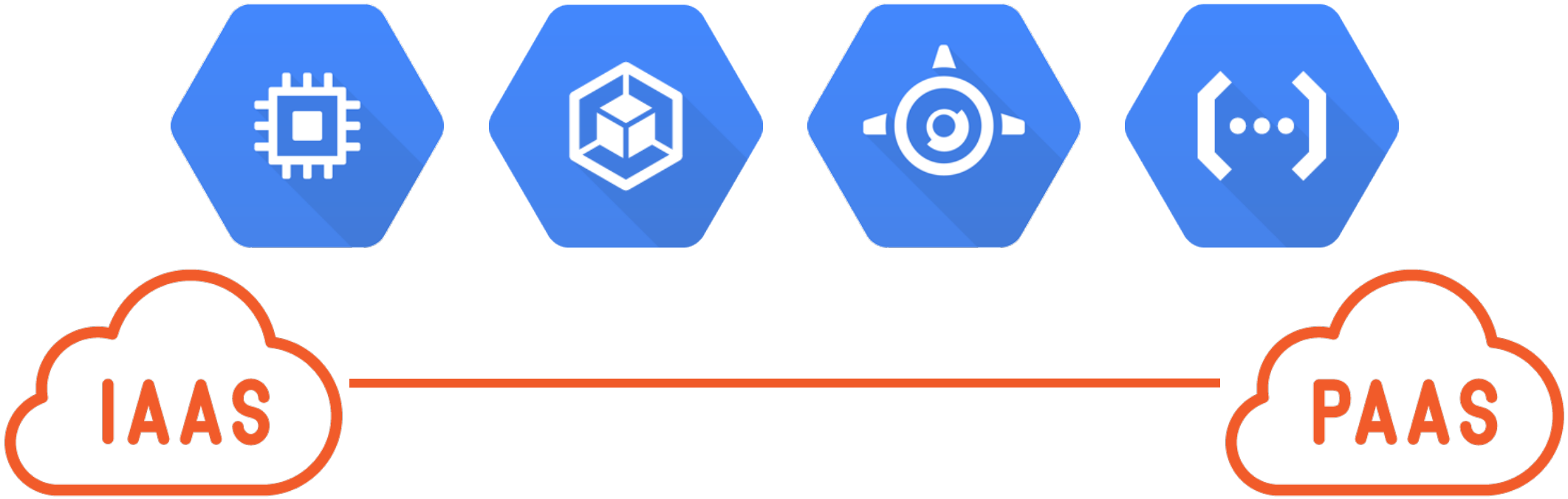
Compute Services

Compute Engine

Kubernetes Engine

App Engine

Cloud Functions


IAAS
PAAS

Serverless computing

Limited language support

Integrated lifecycle management

Ideal for reactive architecture

Cloud Functions

Managed code runtime

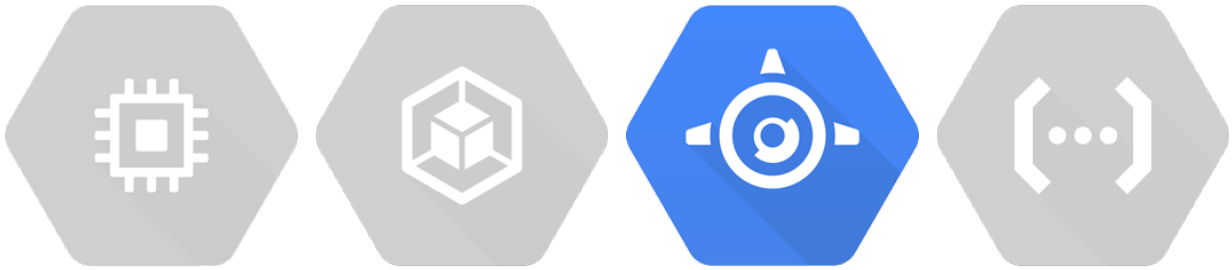
Integrated lifecycle management

Application stack

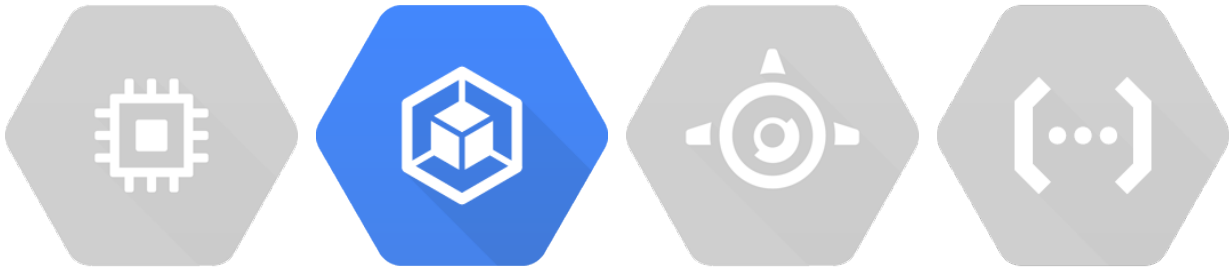
Multiple language support

App Engine Flexible

App Engine



Kubernetes Engine

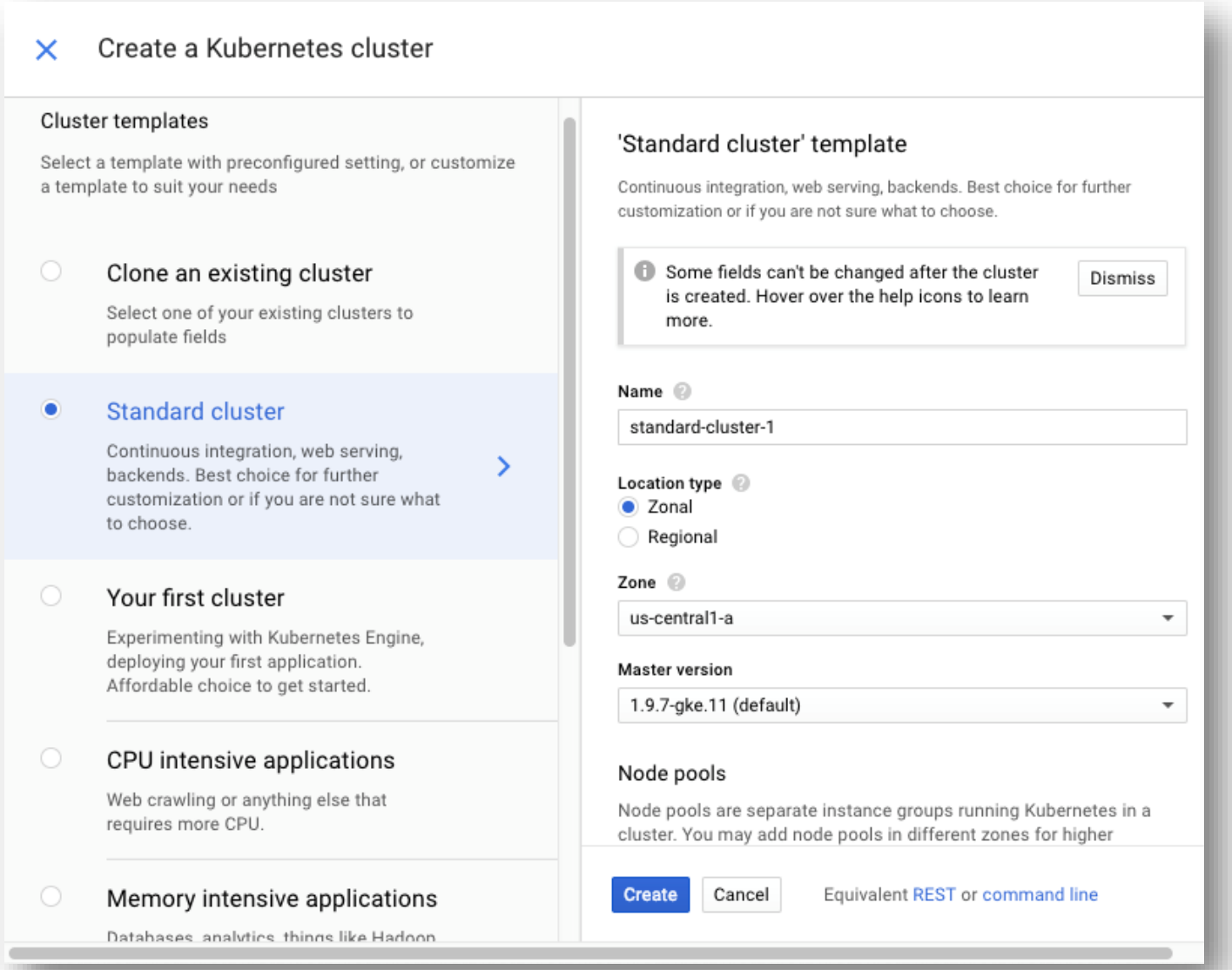


Container orchestration

Hosted Kubernetes

Multiple cluster configuration options

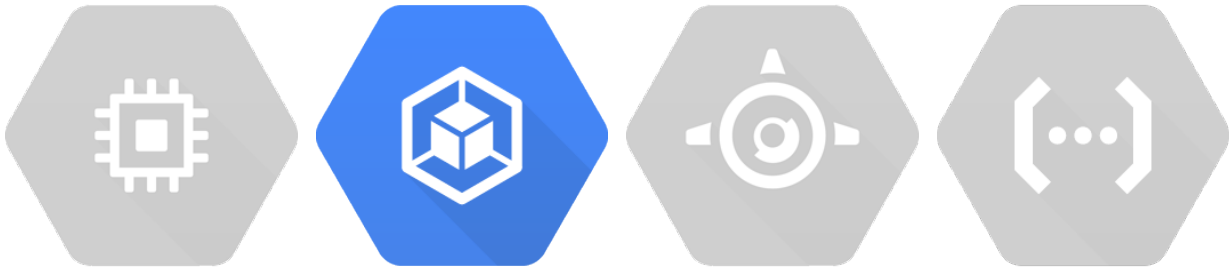
Cluster Templates



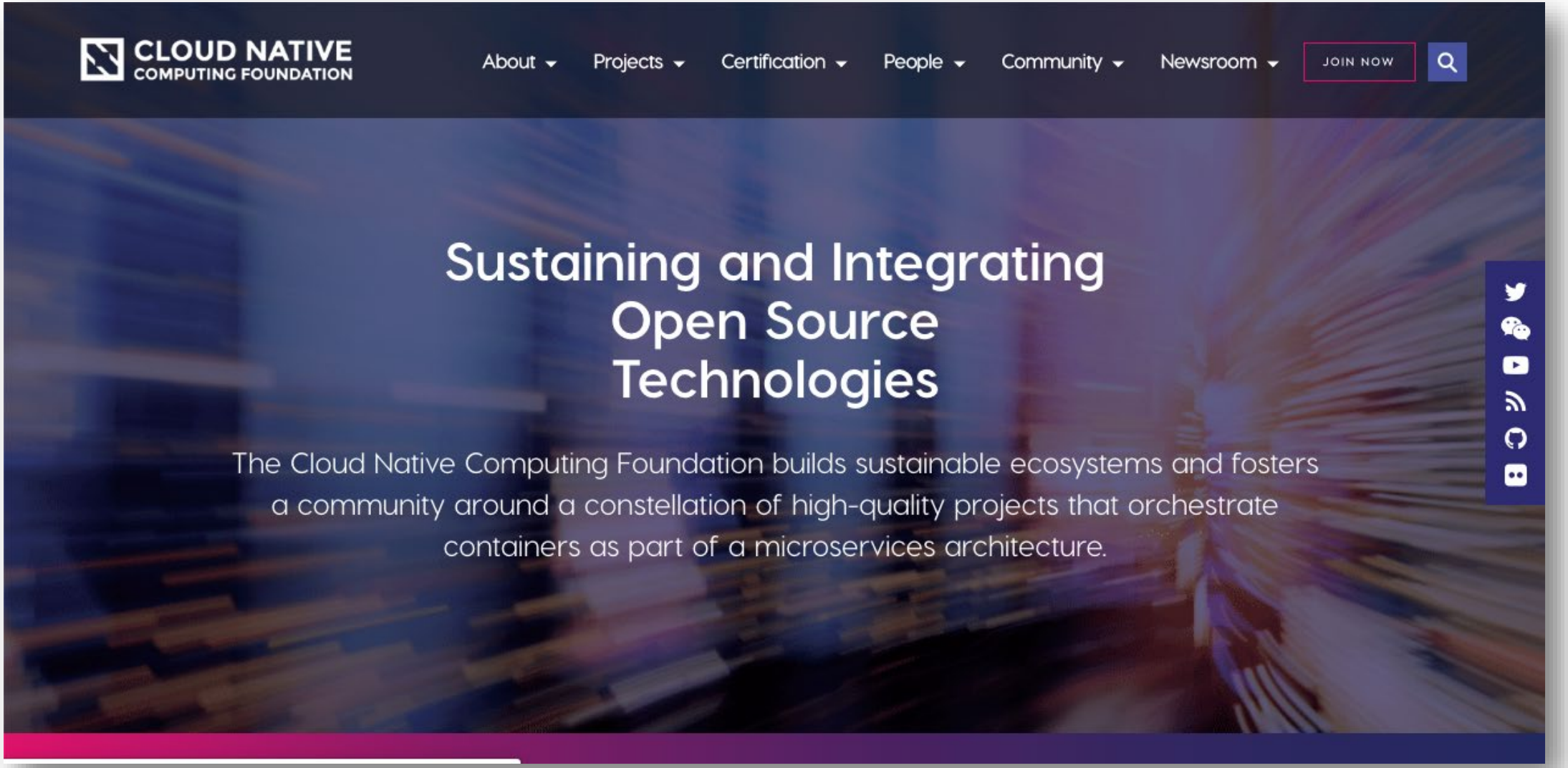
Hosted container registry service

Requires some infrastructure knowledge

Kubernetes Engine



https://cncf.io



Highly configurable virtual machines

Predefined and custom machine types

Cost optimization options

Sole tenant option

Instance groups and autoscaling

Compute Engine

Storage

Cloud SQL

Cloud Storage

Persistent Disk

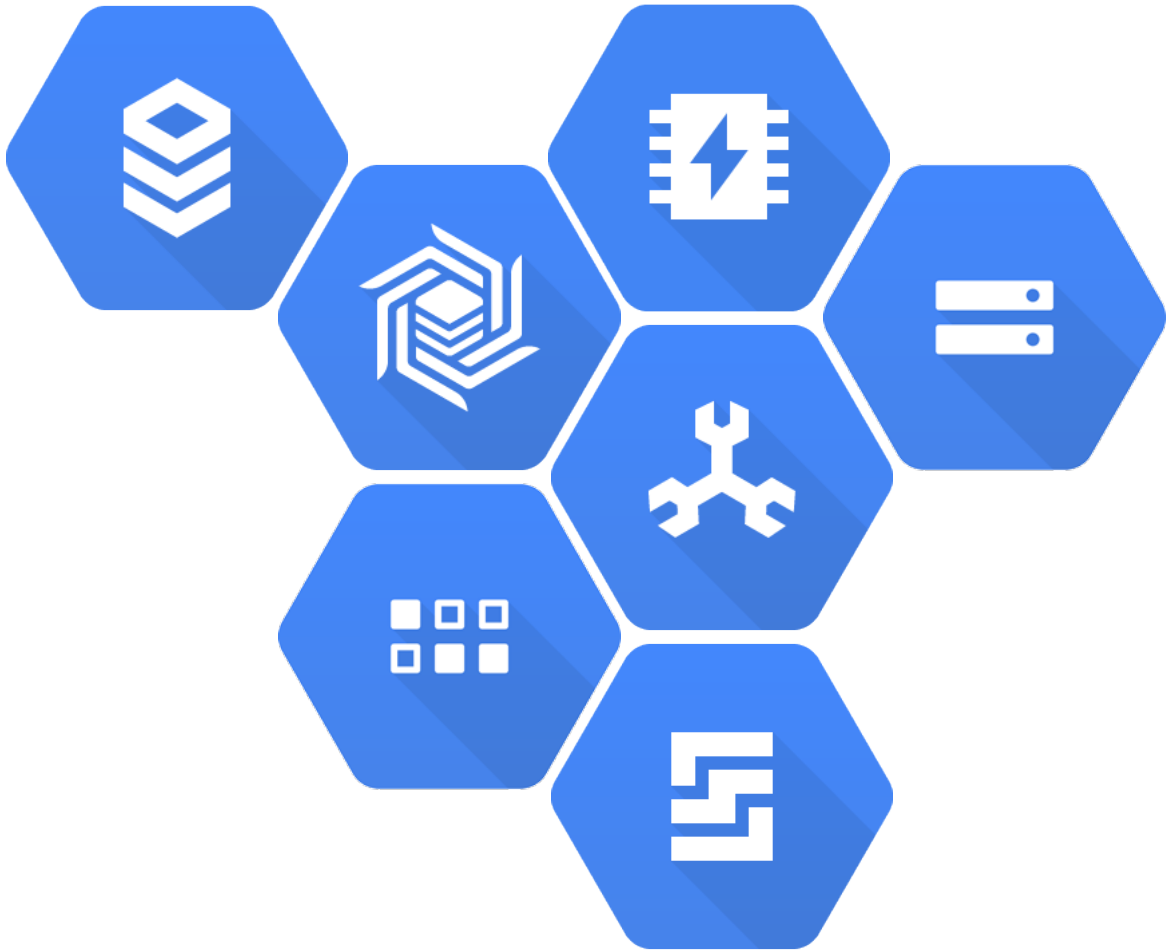
Cloud Filestore

Cloud Bigtable

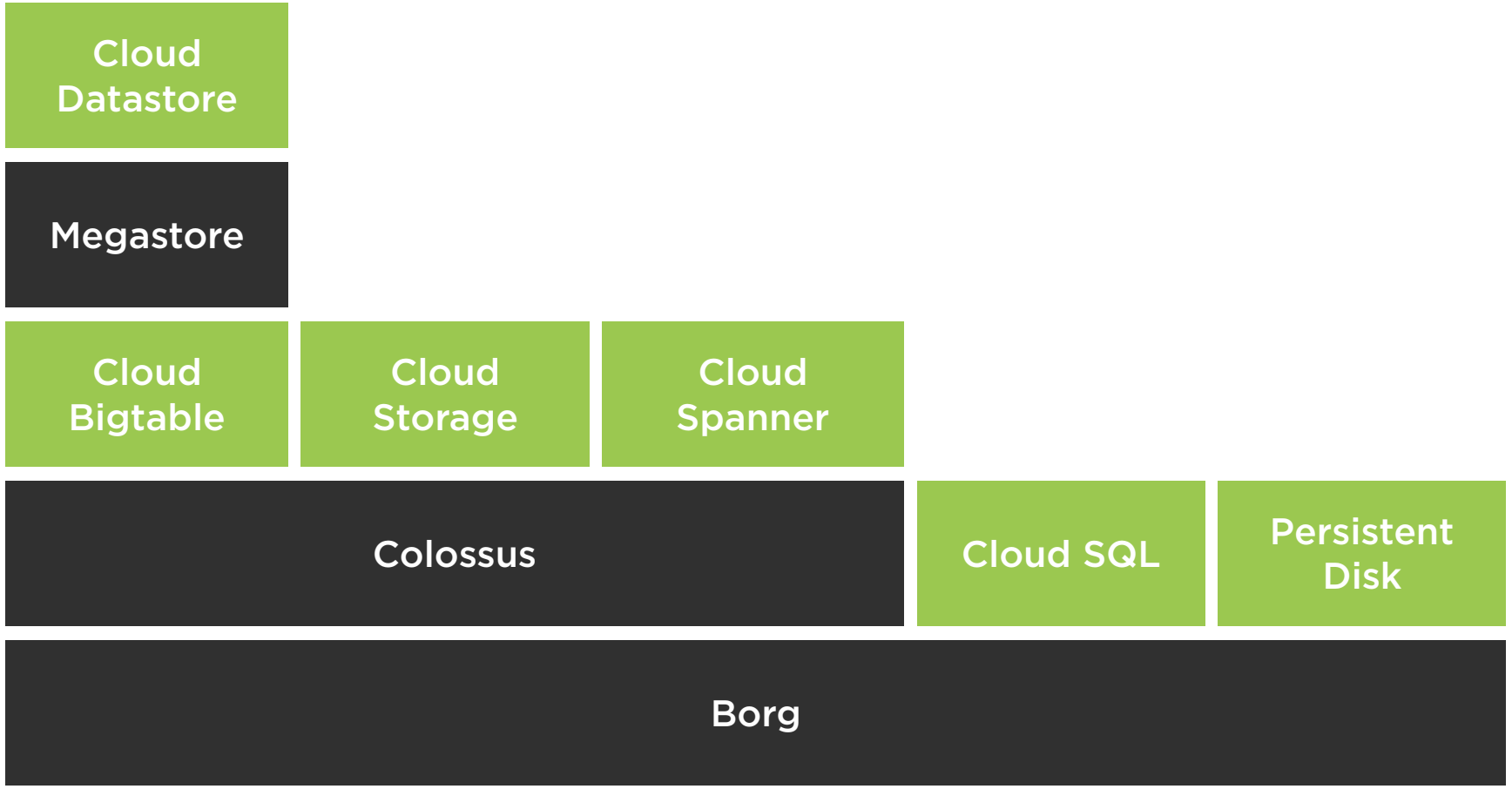
Cloud Spanner

Cloud Datastore (Firestore)

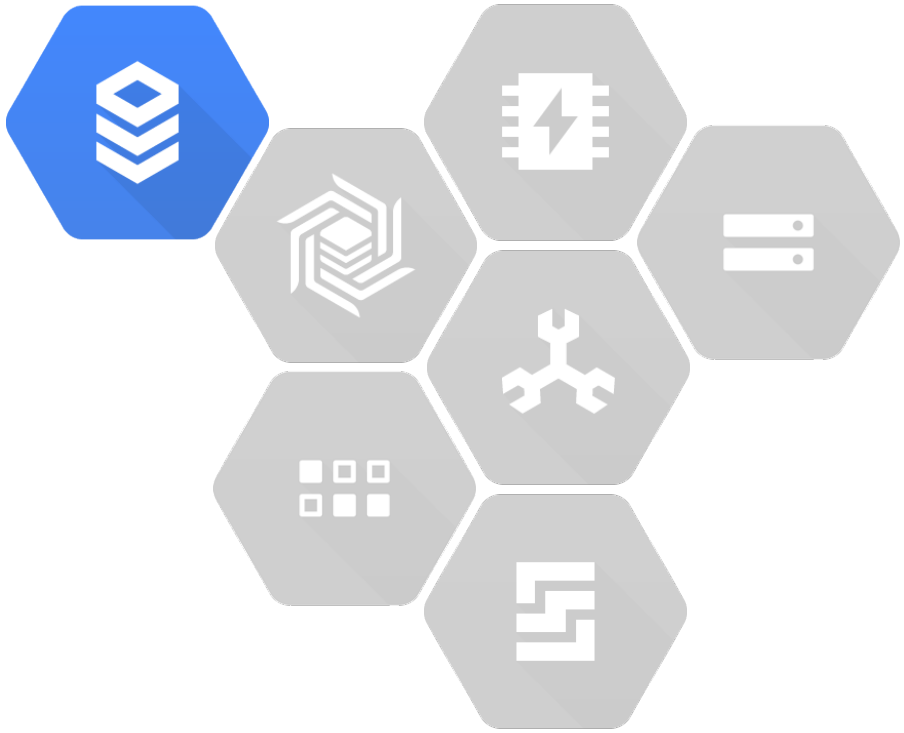
Cloud Memorystore



Technology Layering



Cloud SQL



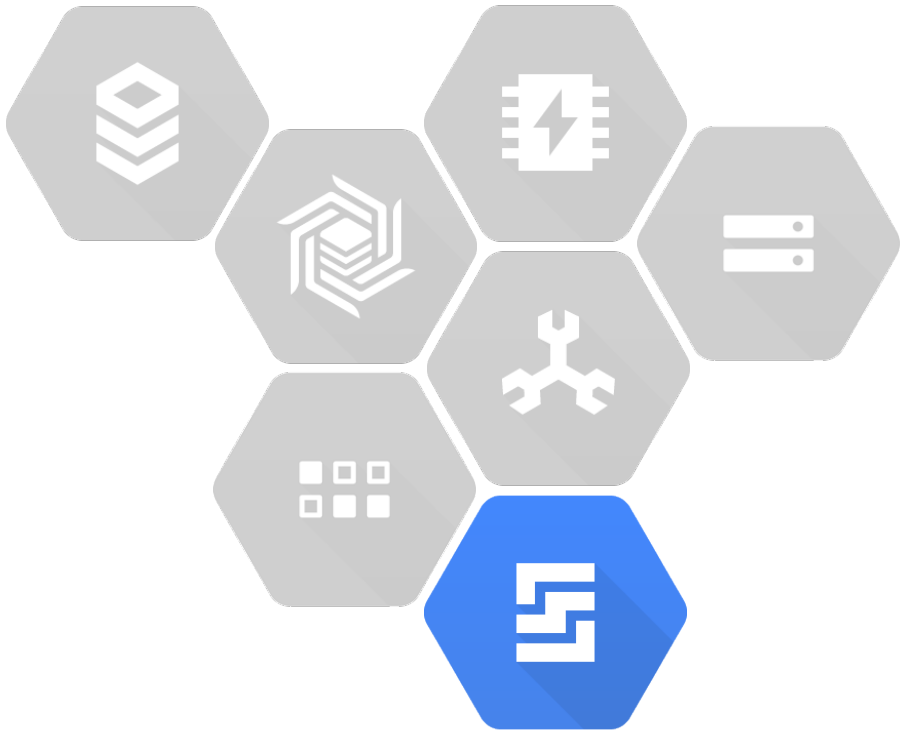
Managed Postgres and MySQL

Scalable for most OLTP workloads

Benefits from Compute Engine cost optimizations

Runtime and management configuration

Persistent Disks



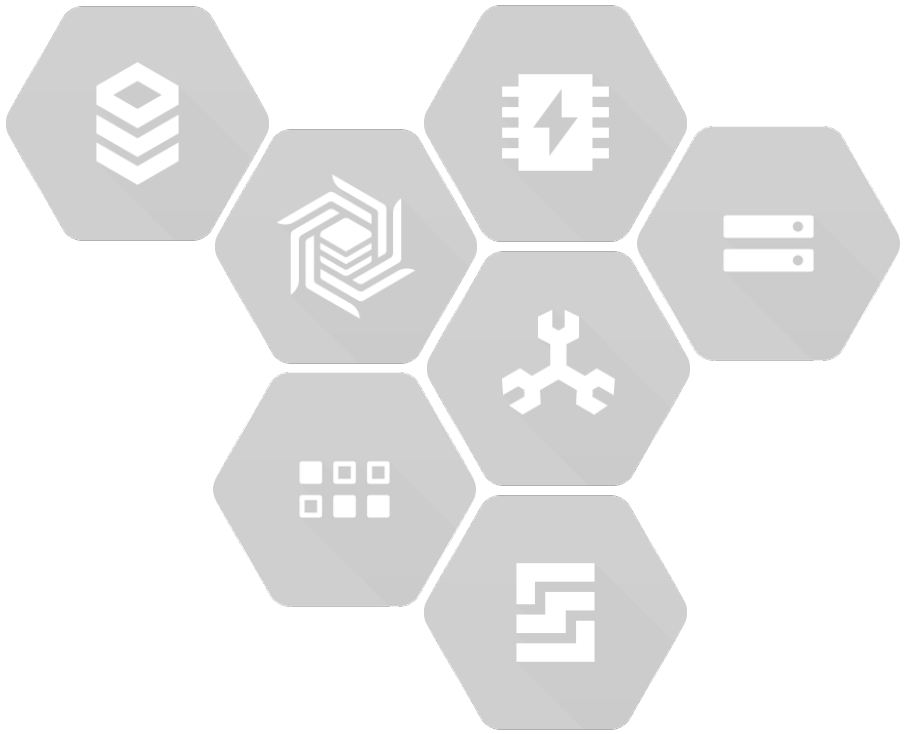
Block storage for use with compute VMs

Independent of VMs

Highly configurable

Managed and optimized infrastructure

Cloud Filestore

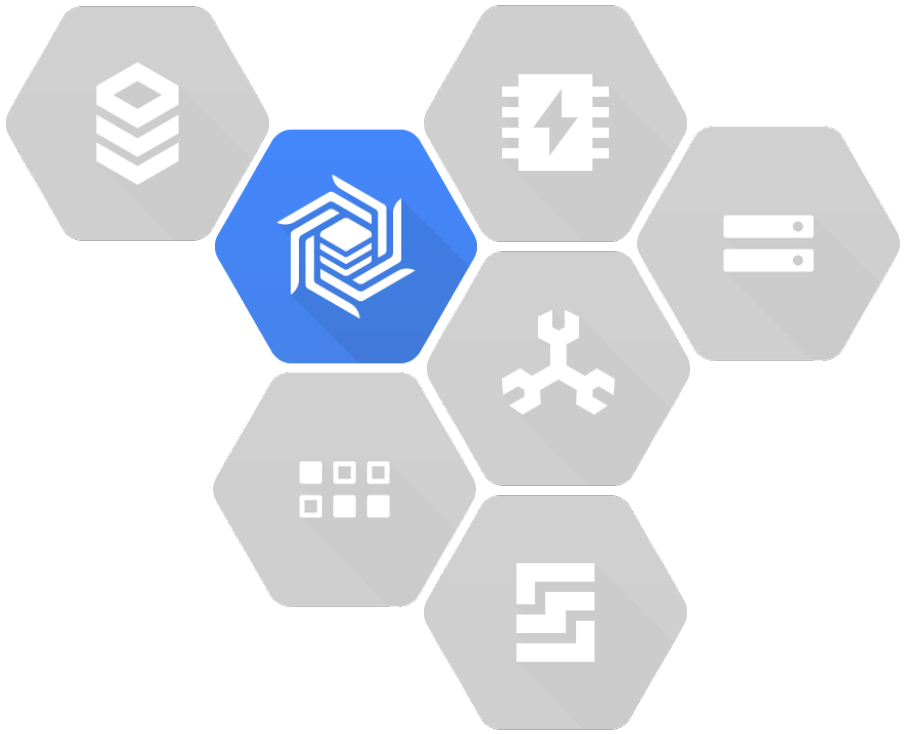


Managed Network Attached Storage (NAS)

Shared block storage for compute VMs or containers

Low latency, low maintenance

Cloud Bigtable



High performance at scale

Managed infrastructure

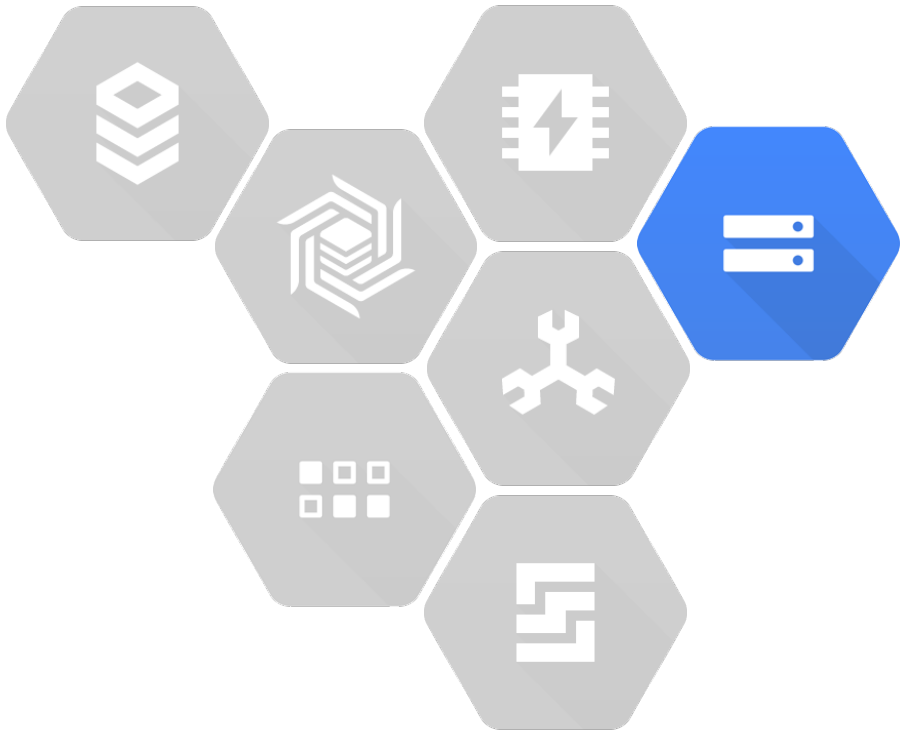
Library support for multiple languages

All entity information modeled as a row with a single index

No transactional guarantees beyond a row operation

Optimized for sparsely-populated rows

Cloud Storage



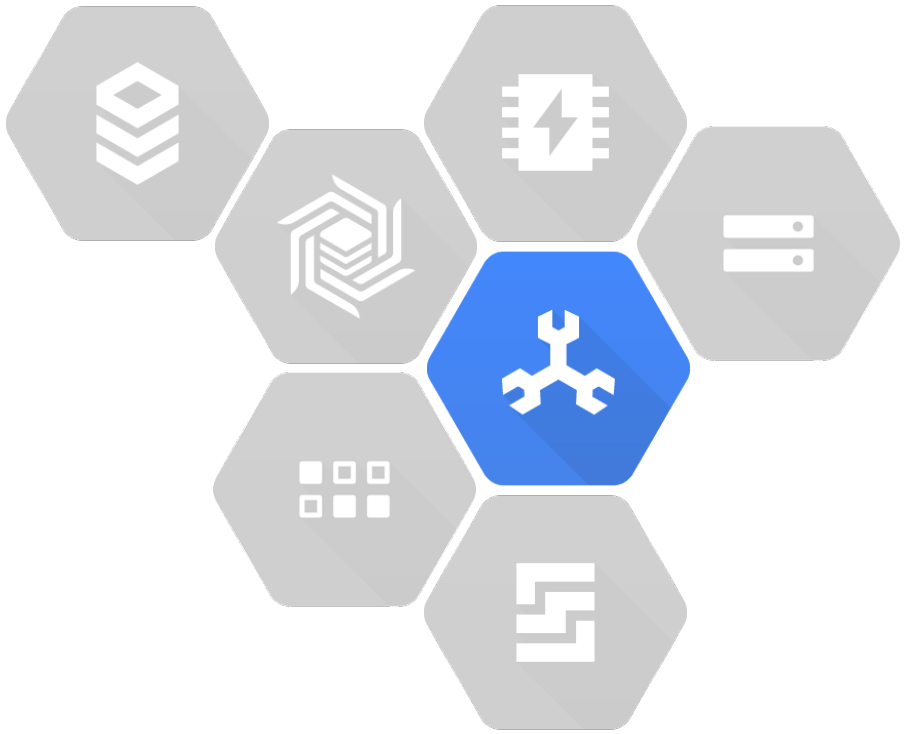
Global blob storage

Automatic edge caching

Different storage classes for different data types

Object lifecycle management

Cloud Spanner

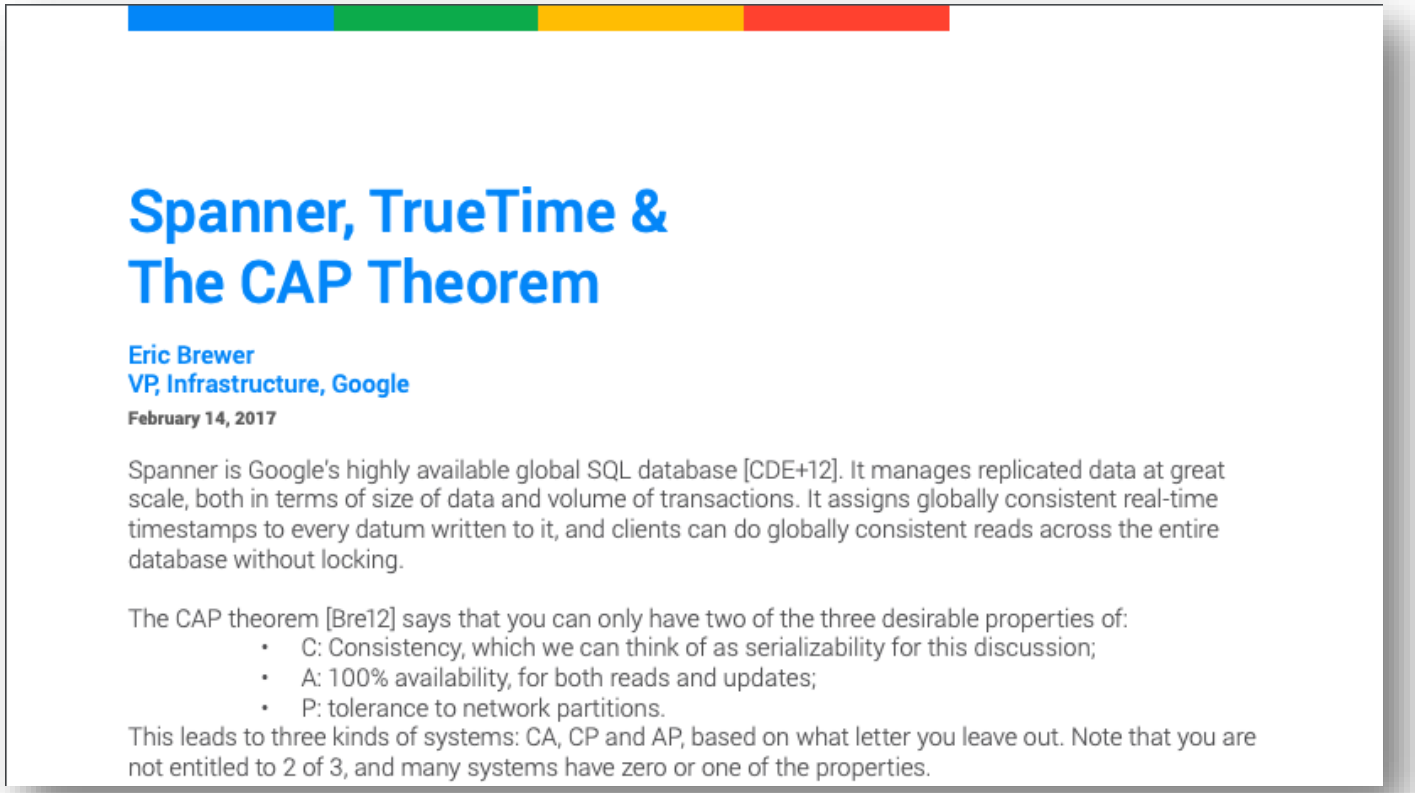


Horizontally-scalable, managed RDBMS

Distributed transaction support

Built for Google's own applications

https://goo.gl/wKy25P



Spanner, TrueTime & The CAP Theorem

Eric Brewer
VP, Infrastructure, Google

February 14, 2017

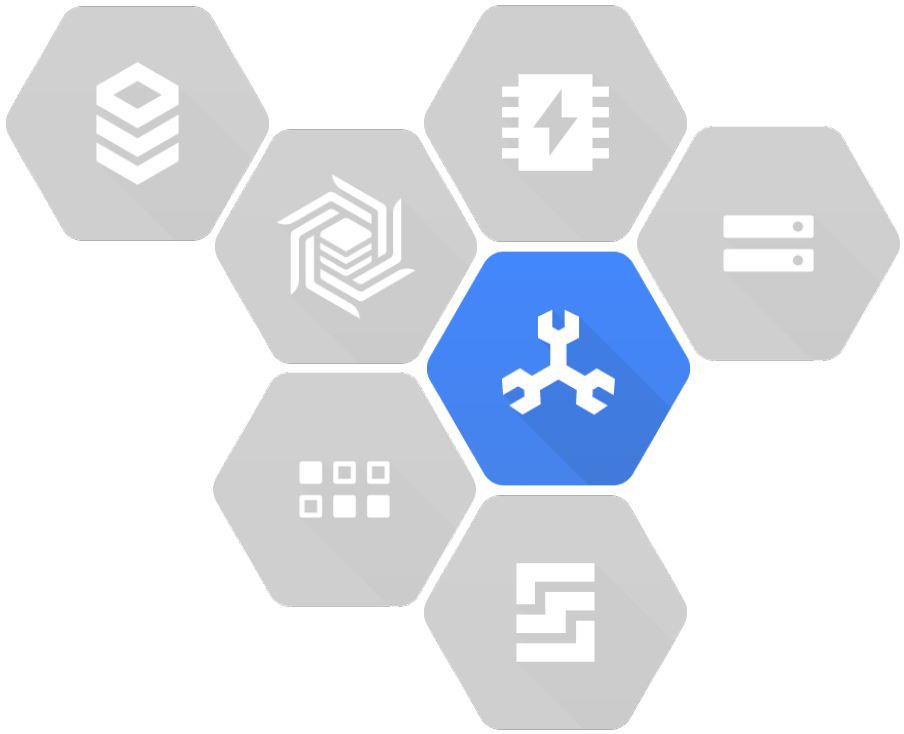
Spanner is Google's highly available global SQL database [CDE+12]. It manages replicated data at great scale, both in terms of size of data and volume of transactions. It assigns globally consistent real-time timestamps to every datum written to it, and clients can do globally consistent reads across the entire database without locking.

The CAP theorem [Bre12] says that you can only have two of the three desirable properties of:

- C: Consistency, which we can think of as serializability for this discussion;
- A: 100% availability, for both reads and updates;
- P: tolerance to network partitions.

This leads to three kinds of systems: CA, CP and AP, based on what letter you leave out. Note that you are not entitled to 2 of 3, and many systems have zero or one of the properties.

Cloud Spanner

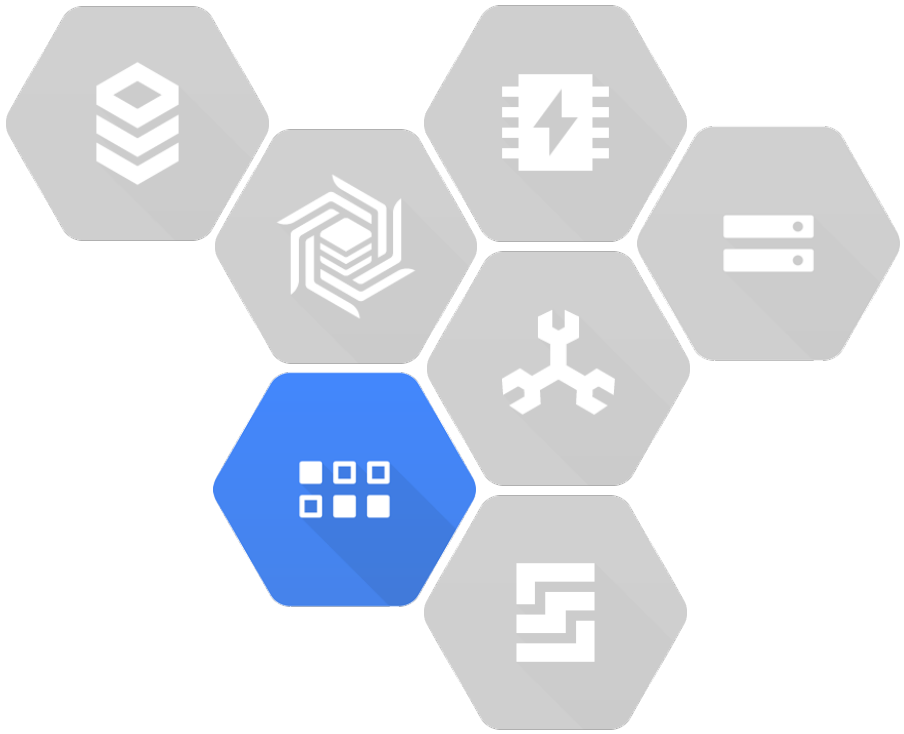


Horizontally-scalable, managed RDBMS

Distributed transaction support

Built for Google's own applications

Cloud Datastore



Used for managing structured data

Limited transaction support

Scales based on size of query results

Cloud Firestore

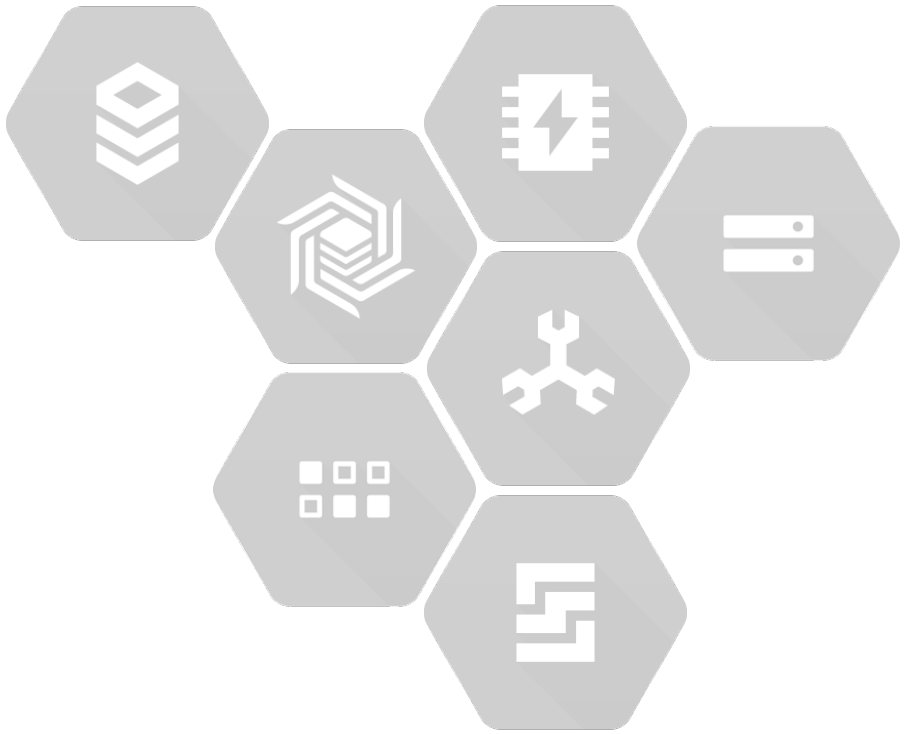
Brings together Cloud Datastore and the Firebase real-time database

Changes data model and storage model

Full transaction support

3 modes for transitioning from Cloud Datastore

Cloud Memorystore

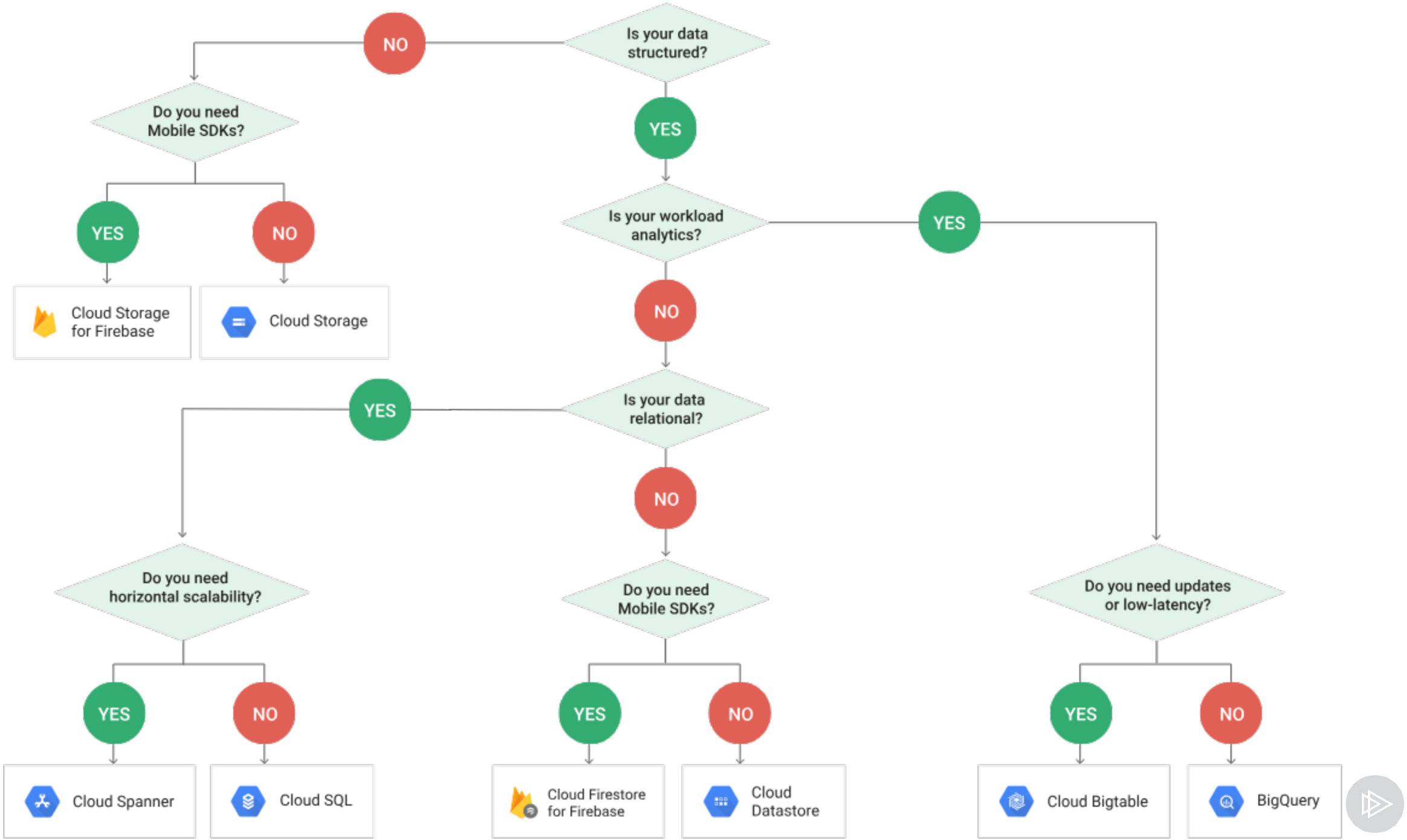


In-memory database

Ideal for caching

Supports Redis application protocol

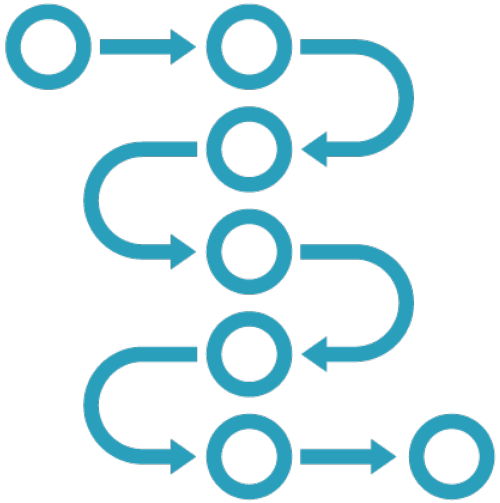
Managed infrastructure


Is your data structured?
NO
YES
Do you need Mobile SDKs?
YES
NO
Cloud Storage for Firebase
Cloud Storage
Is your workload analytics?
YES
NO
Is your data relational?
YES
NO
Do you need horizontal scalability?
YES
NO
Cloud Spanner
Cloud SQL
Do you need Mobile SDKs?
YES
NO
Cloud Firestore for Firebase
Cloud Datastore
Do you need updates or low-latency?
YES
NO
Cloud Bigtable
BigQuery

Networking

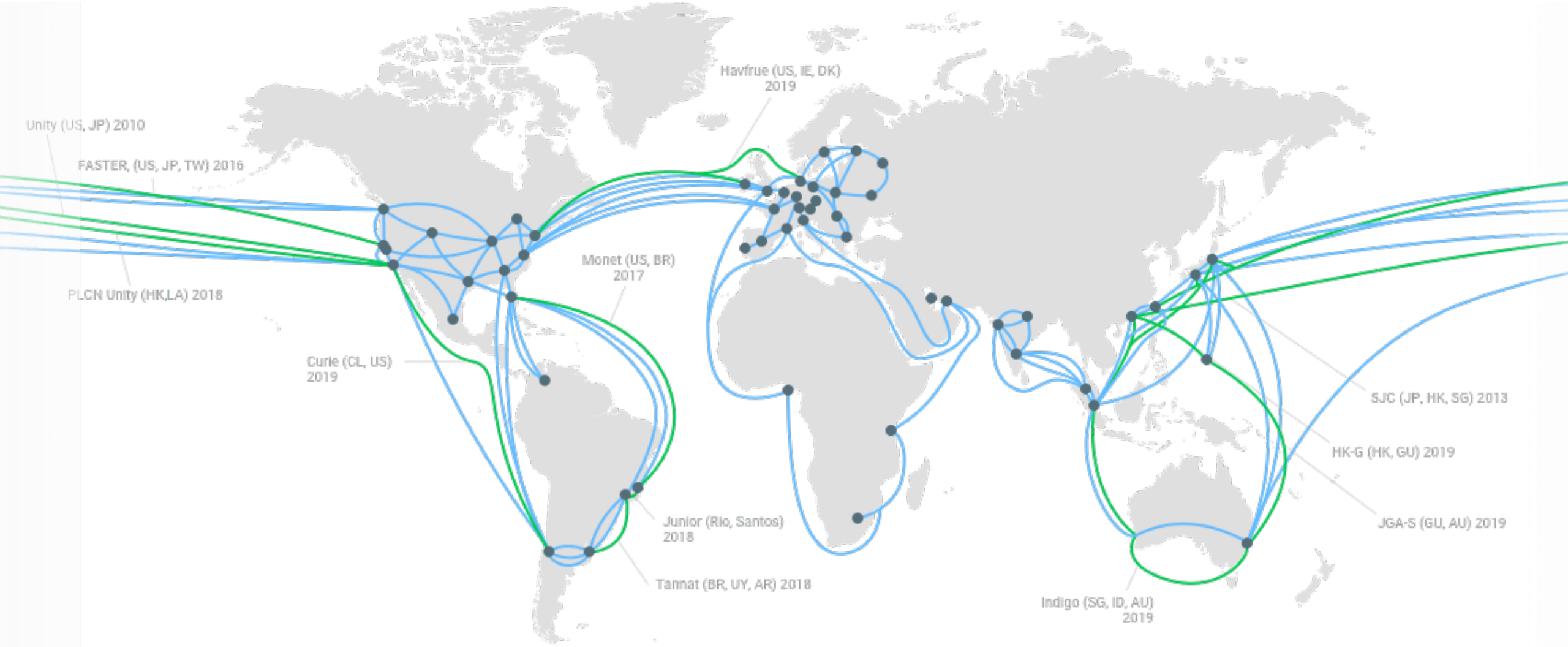
What Makes Google's Network Special?

Size and scale



Agility

Performance


Havfrue (US, IE, DK)
2019
Unity (US, JP) 2010
FASTER, (US, JP, TW) 2016
PLCN Unity (HK,LA) 2018
Monet (US, BR)
2017
Curie (CL, US)
2019
Junior (Rio, Santos)
2018
Tannat (BR, UY, AR) 2018
SJC (JP, HK, SG) 2013
HK-G (HK, GU) 2019
JGA-S (GU, AU) 2019
Indigo (SG, ID, AU)
2019

Google Cloud Platform Blog

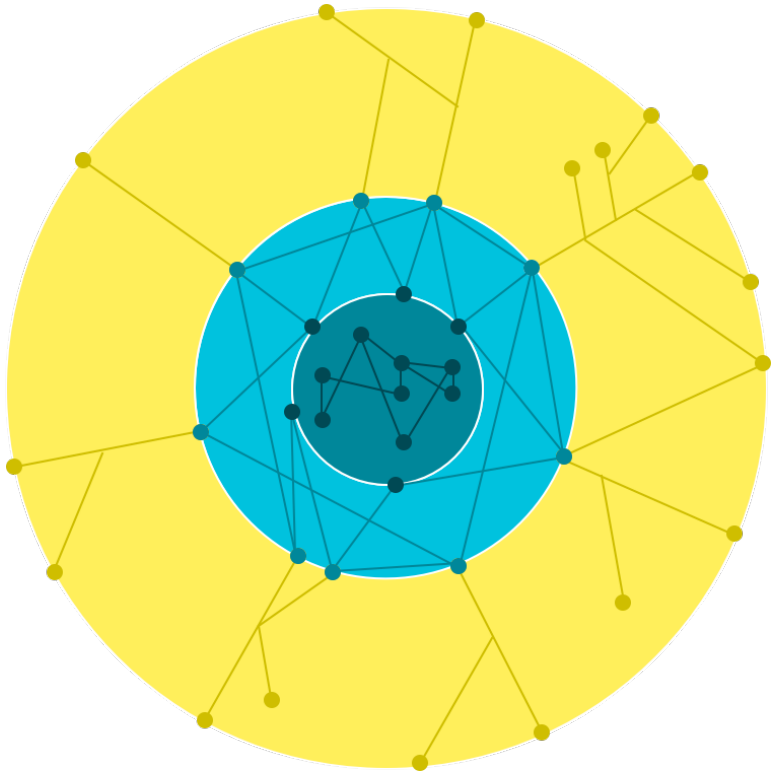
Product updates, customer stories, and tips and tricks on Google Cloud Platform

Our Los Angeles cloud region is open for business

Monday, July 16, 2018

By Kirill Tropin, Product Manager

Hey, LA — the day has arrived! The Los Angeles Google Cloud Platform region is officially open for business. You can now store data and build highly available, performant applications in Southern California.



Data center network

Edge Points of Presence (PoPs)

Edge caching nodes (GGC)

Software Defined Networking (SDN)

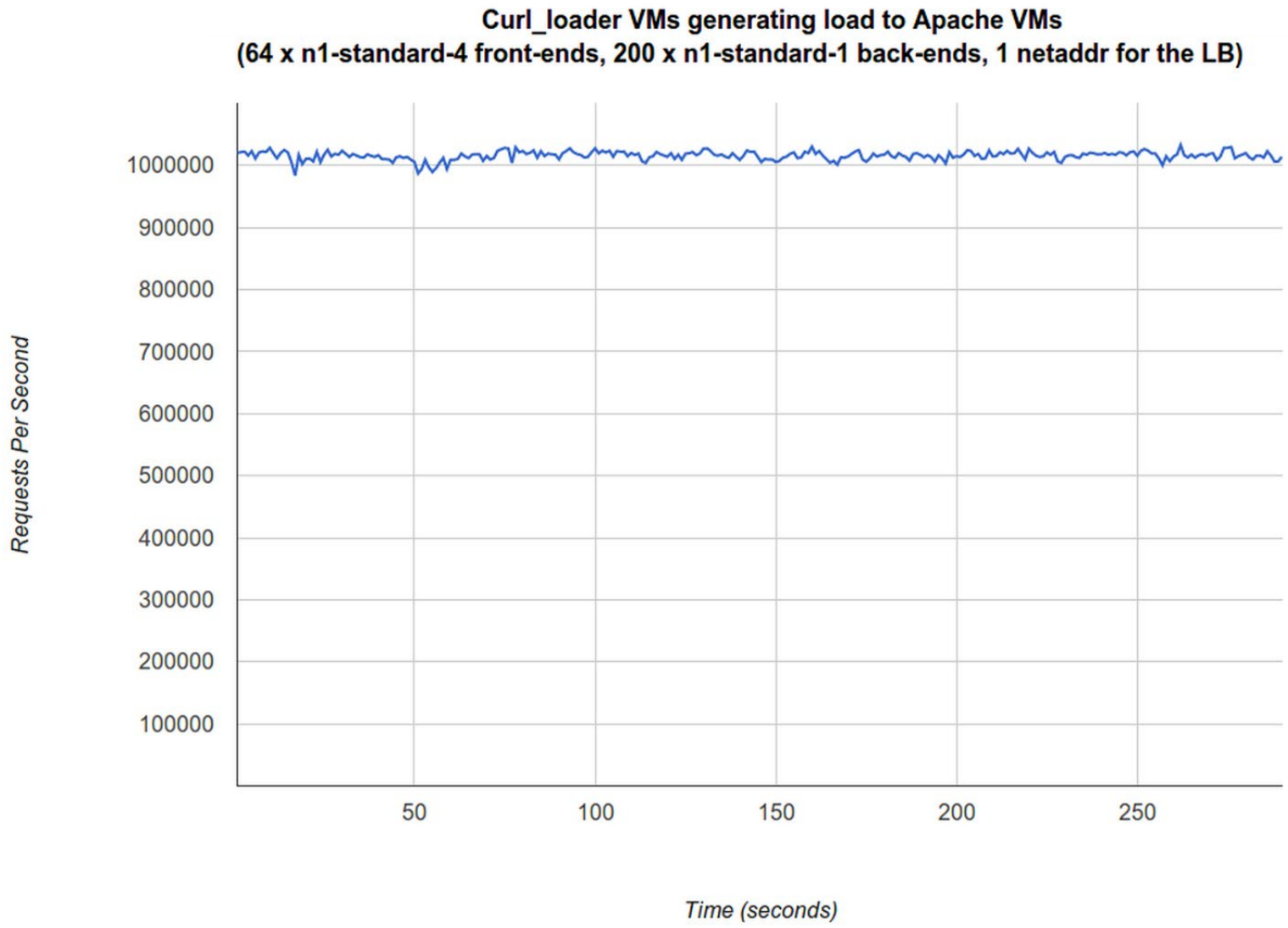
B4 (2013)

Andromeda (2014)

Jupiter (2015)

Espresso (2017)

Performance



Virtual Private Cloud

Cloud Load Balancing

Cloud Armor and Telemetry

Content Delivery Network

Cloud DNS

Cloud Interconnect

Network Service Tiers

Virtual Private Cloud (VPC)

Private network space

Metadata-driven approach to policy

Shared VPCs for large, federated systems

Cloud Load Balancing

Single, load-balanced IP address

Uses anycast IP addresses

Cloud Armor and Telemetry

Protection against DDoS attack

Applies policy on top of Cloud Load Balancer

Telemetry provides detailed inspection of all VPC ingress and egress

Content Delivery Network (CDN)

Extends caching beyond peering edge

Caches content on Google Global Cache nodes

Cloud DNS

Leverages existing Google DNS infrastructure

Flexible DNS configuration management

Cloud Interconnect

Connects existing network infrastructure to Google network

Includes both VPN and peering connections

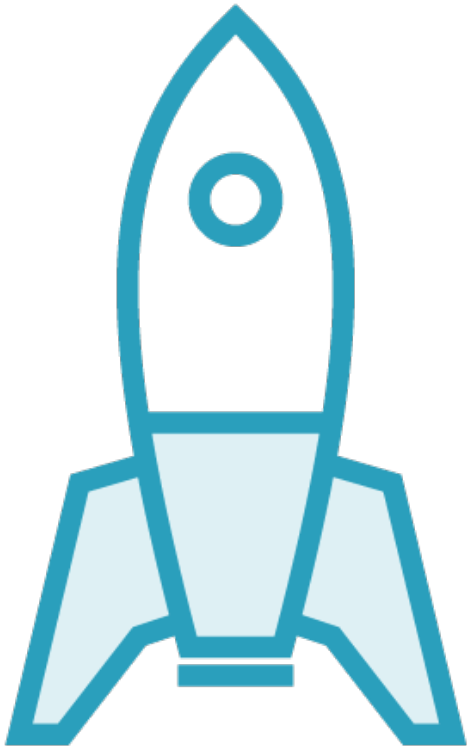
Supports direct and partner-mediated connections

Network Service Tiers

Base level and premium tiers

Difference based on how long traffic stays in Google's network

Enabling “Cloud 3.0”



Applications and functions, not VMs

Storage disaggregation, not disks

SLAs, not load balancing and scheduling

Intelligence, not data processing

Policy, not “middle boxes”

Summary

Core building blocks include compute, storage, and networking

Layering provides choice and agility

Common infrastructure across all of Google