

# Vitess

## Stateless Storage in the Cloud

---

Sugu Sougoumarane

Co-creator, CTO @ PlanetScale



@VITESSIO

# What is Vitess?

---

Cloud  
Native  
Database

---

Massively  
Scalable

---

HA

---

Based on  
MySQL

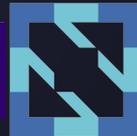




*Graduation!*

**Vitess**

**CLASS OF 2019**



VITESS.IO

# Vitess Stats

Started  
**2010**

Marquee  
Adopters



**180**

Contributors



**9,000**

Stars

**17,000**

 Commits

**1100**

Slack Members

**1200**

Forks 



# Key Adopters



“No database storage system other than Vitess truly fit all of Slack’s needs.”

- Michael Demmer, Principal Engineer



**“You have to keep on working on it,  
but Vitess does provide you  
essentially with near unlimited scale.”**

**- Jon Tirsen, Engineering Manager**

 Square

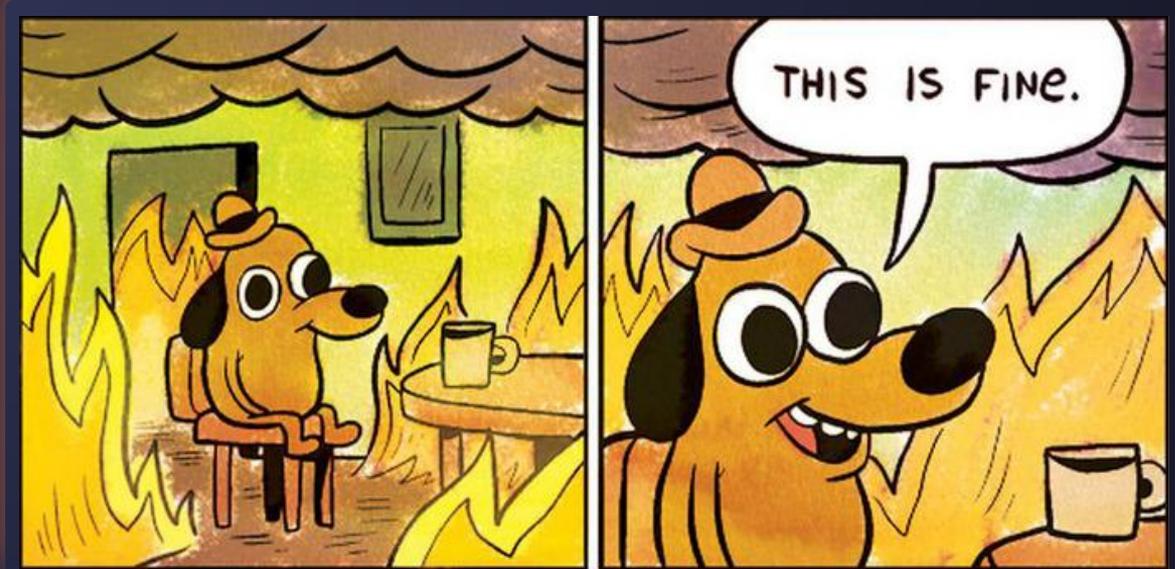


“Vitess gave us the horizontal scalability without any restriction on the DB clients.”

- Chengcheng Hu, Senior Engineering Manager



# YouTube in 2010



2013

## YouTube

- On-prem → Borg
- Remain Open Source
- Adapters for proprietary APIs

Storage Engine running as Stateless Application

---

# Stateless Storage



# 2015



## Google Cloud Platform Blog

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

### Scaling MySQL in the cloud with Vitess and Kubernetes

Friday, March 20, 2015

Your new website is growing exponentially. After a few rounds of high fives, you start scaling to meet this unexpected demand. While you can always add more front-end



kubernetes release date

All News Images Videos Shopping More Settings Tools

About 6,630,000 results (0.49 seconds)

Kubernetes / Initial release dates

July 21, 2015



kubernetes

# Kubernetes Workloads



Oldest, since  
2016



Hundreds of  
keyspaces



30,000+  
Tablets



Migrate from cloud to  
cloud MySQL 8.0



# Earlier in 2019



VITESS.IO



**Kelsey Hightower**   
@kelseyhightower

I'm always going to recommend people exercise extreme caution when running stateful workloads on Kubernetes. Most people who are asking "can I run stateful workloads on Kubernetes" don't have much experience with Kubernetes and often times the workload they are asking about.

12:10 AM · Mar 24, 2019 · [Twitter Web Client](#)

283 Retweets 914 Likes



**Kelsey Hightower**   
@kelseyhightower · Mar 24

Replying to [@kelseyhightower](#)

Some people believe that rubbing Kubernetes on a stateful workload turns it into a fully managed database offering rivaling RDS. This is false. Maybe with enough effort, and additional components, and an SRE team, you can build RDS on top of Kubernetes.

 18

 149

 494



**Sugu Sougoumarane**   
@ssougou · Mar 24

Replying to [@kelseyhightower](#)

[@vitessio](#) fills this gap. [@HubSpot](#) and [@JD\\_Corporate](#) have been running stateful workloads on it for over two years now. But it will be nice if Kubernetes put more thought into this problem area :).

 2

 1

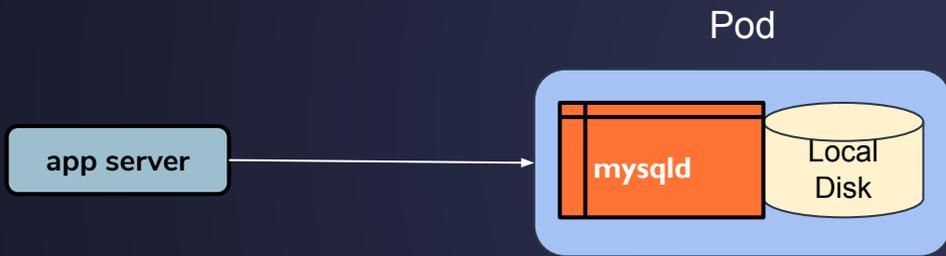
 10





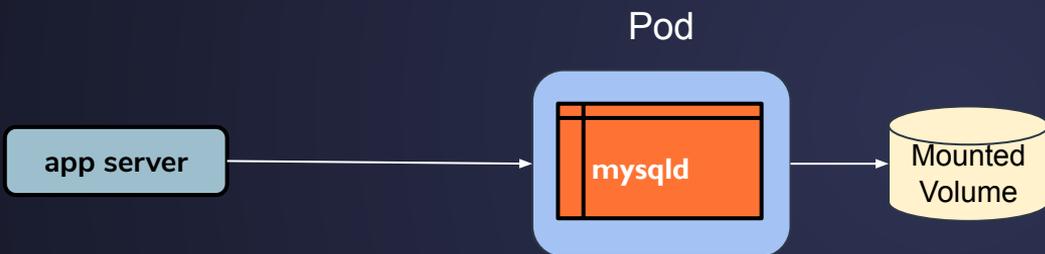
One does not  
simply...  
**move** MySQL to  
Kubernetes...

# Scenario 1: Local Disk



- Data loss on Pod reschedule

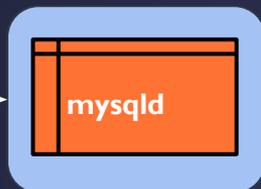
# Scenario 2: Mounted Volume



- Unavailability on Pod reschedule
- Performance impact

# Scenario 3: Failover

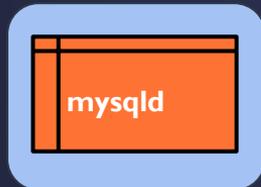
Pod 0: Master



app server

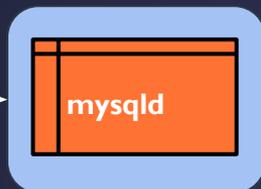
- Cannot designate master
- Failover on Pod reschedule
- Who informs the app?

Pod 1: Replica

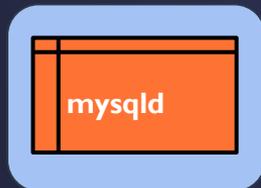


# Scenario 4: Discovery

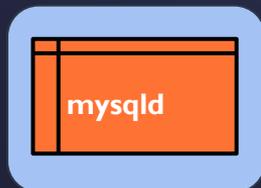
Pod 0: Master



Pod 1: Replica



Pod 2: New Replica



app server



- How do you initialize the replica?
- Who informs the app?

# Cultural **issues**

Life cycle

Data  
Size

IPs reused

Topo  
overload

# Later in 2019



**Kelsey Hightower** ✓  
@kelseyhightower

Replying to @kelseyhightower @chrislovecnm and 2 others

I challenge anyone to believe that Kubernetes alone will deliver RDS like database management out of the box. If people believe that then I'm not sure they understand how RDS works.

3:53 PM · Oct 8, 2019 · [Twitter Web App](#)

2 Likes



**Chris Love** @chrislovecnm · Oct 8

@kelseyhightower u "recommend people exercise extreme caution when running stateful workloads on K8s". Which I agree with. You need to be an expert at Kafka and k8s to run it k8s. I think you in the past recommended not doing it at all? Just doing some fact-checking for a preso.

1



**Kelsey Hightower** ✓  
@kelseyhightower

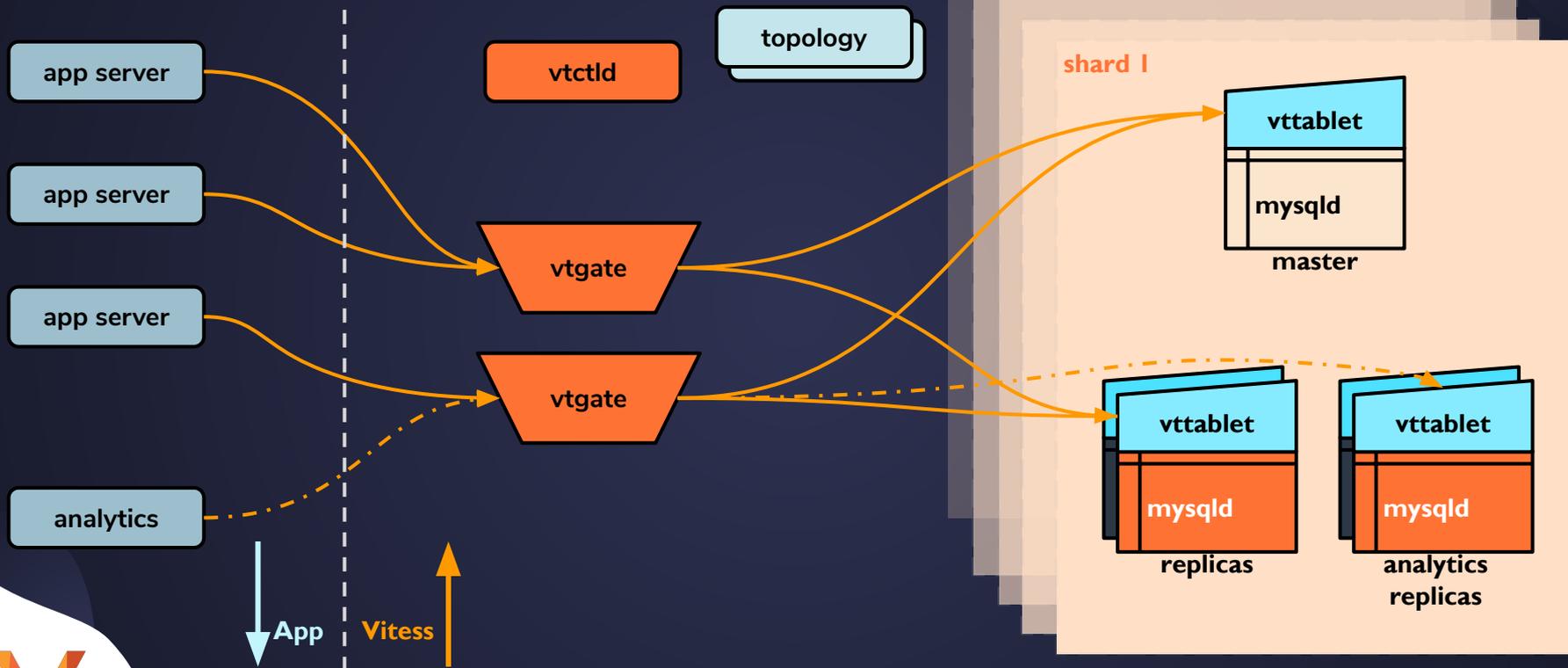
Replying to @chrislovecnm

I can't suggest people never do it, but I stand by Kubernetes is not enough. You'll need tooling and help from the stateful system. [@cockroachdb](#) is a good example of a database that meets Kubernetes part way.

For others such as mysql, then something like [@vitessio](#) helps.

3:42 PM · Oct 8, 2019 · [Twitter Web App](#)

# Architecture



We're hiring!  
[planetscale.com/careers](https://planetscale.com/careers)



VITESS.IO

# What's Next

Vitess.io

Do the  
tutorial

Bring up  
a cluster

 slack  
Join Vitess  
Slack

[console.planetscale.com](https://console.planetscale.com)

Don't leave your data behind.



VITESS.IO