Vitess, k8s & sharding

Sugu Sougoumarane, Co-creator
CTO, PlanetScale
What is Vitess

- Sharding middleware for MySQL
- Massively scalable
- HA
- Cloud-native
Vitess stats

- Started 2010
- 8,000+ stars
- 16,000+ commits

Marquee adopters

- 100+ contributors
- 800+ Slack members
- 1000+ forks
Key adopters

- slack  25% Migrated to Vitess
- Square  Cash App fully runs on Vitess
- Pinterest  All of advertising campaign management
Top Charts

Free Apps

1. Cash App
   Send, Spend Money, Buy Bitcoin
   OPEN

2. YouTube: Watch, Listen, Stream
   Videos, Music and Live Streams
   OPEN

3. BURGER KING® App
   Coupons & Mobile Order
   GET

4. Instagram
   Photo & Video
   Download

Paid Apps

326 509 3.2K
The Community

- YouTube
- Slack
- Square
- JD.COM
- Pinterest
- HubSpot
- GitHub
- Nozzle
- Stitch Labs
- Flipkart
- BetterCloud
- Quiz of Kings
- AXON
- PixelFederation
Vitess on Kubernetes
Scaling MySQL in the cloud with Vitess and Kubernetes

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

About 6,680,000 results (0.49 seconds)

Kubernetes / Initial release dates

July 21, 2015
Production Workloads

- stitchlabs: Oldest, since 2016
- HubSpot: Hundreds of keyspaces
- JD.COM: 10,000+ Tablets
- Migrate from cloud to cloud
- MySQL 8.0
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 - 24 Mar 2019

283 Retweets 895 Likes

Some people believe that running 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.

Kelsey Hightower @kelseyhightower · Mar 24

Sugu Sougoumsrane @ssougou · Mar 24

Replying to @kelseyhightower

@vitessa 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 :).
ONE DOES NOT SIMPLY

MOVE MYSQL TO KUBERNETES
Starting with Vitess

- app server
- app server
- app server

Vitess

- Connection pooling
- Deadlines
- Hot row protection
- Row count limit
- Blacklisting

MainDB
Growing with Vitess

Replica routing
Load balancing
Master promotion with Orchestrator
Growing with Vitess

app server -> Vitess

Vitess -> DB2, Replica
Vitess -> DB1, Replica

Unified View
Growing with Vitess

Sharding agnostic
Growing with Vitess

app server

Vitess

Multi-zone Isolation

Zone 1
DB2
Replica
DB1:S1
Replica
DB1:S2
Replica
DB1:S3
Replica

Zone 2
Replica
Replica
Replica

Zone 3
Replica
Replica
Replica

app server

app server
Demo
Marketplace schema

- **product**
  - pid
  - description

- **customer**
  - cid
  - name

- **orders**
  - oid
  - cid
  - pid
  - mname

- **merchant**
  - name
VReplication

```
select pid, sum(price) as amount from orders
select * from orders where in_keyrange(mname, 'unicode_loose_md5', '-80')
select * from product
```
Use cases

- Materialized Views
- Real-time Rollups
- Resharding
- Backfilling of lookup vindexes
- Schema deployment
- Data migration
- Change notification
What’s next?

- Sign into [slack](https://slack.com) channel at [vitess.io](https://vitess.io)
- Try bringing up a cluster on your laptop or kubernetes
- [https://vitess.io/docs/tutorials/](https://vitess.io/docs/tutorials/)
- [https://github.com/vitessio/vitess](https://github.com/vitessio/vitess)
- Square Cash blog posts on sharding with Vitess
- ![Slack](https://slack.com) @sougou
- ![Twitter](https://twitter.com) @vitessio @planetscaledata @ssougou

Questions?