Vitess

Jurisdiction Aware data locality

Deepthi Sigireddi
Jitendra Vaidya

Who are these people?

Deepthi Sigireddi
Vitess maintainer
Software Engineer
PlanetScale

Jitendra Vaidya
CEO, PlanetScale
Vitess contributor

Vitess

The Problem

- As governments pass data-locality laws, jurisdiction-aware database clusters are becoming important

- Supporting data-locality often requires re-architecting the applications

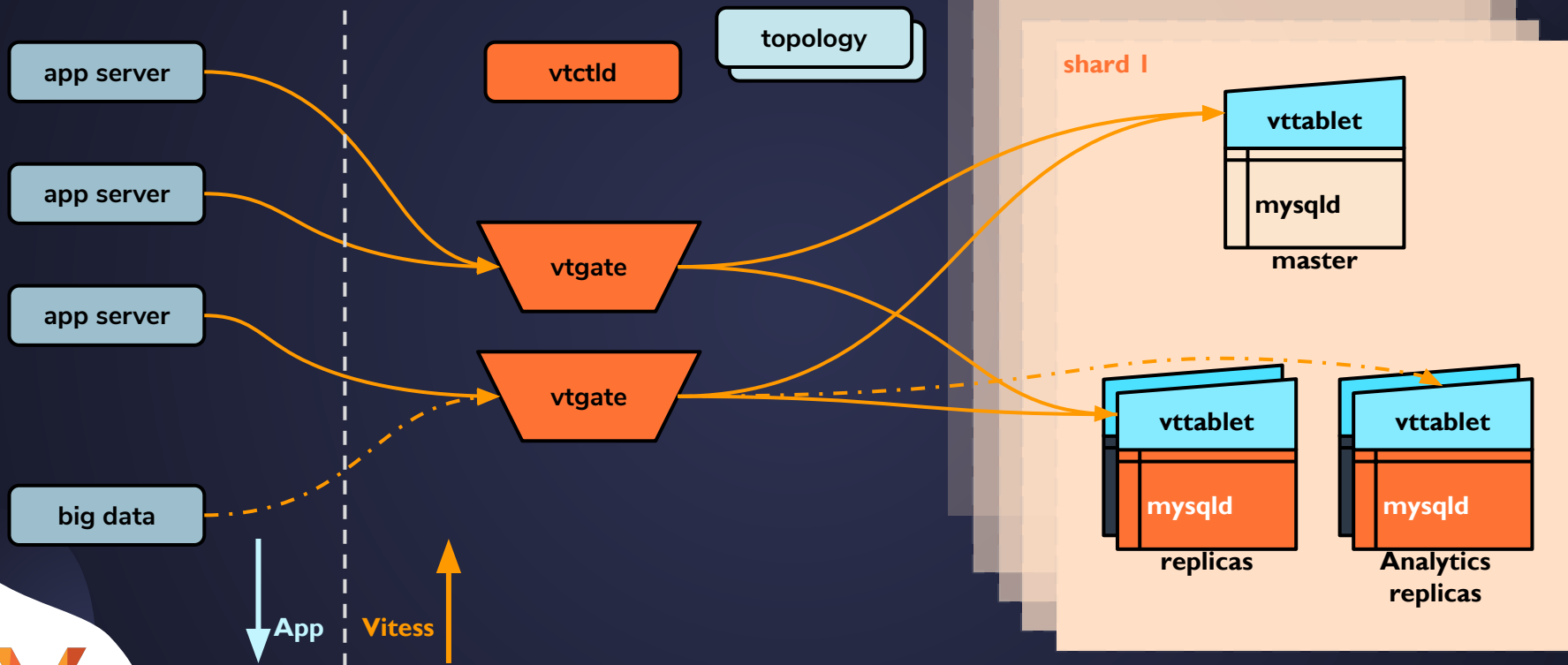- Migrating existing databases into databases resident in multiple locations is an operational challenge

Vitess

The Solution

Map jurisdiction to set of keyranges
This in turn maps a jurisdiction to a shard

Shard is located in a cell bounded by jurisdiction

Area corresponds to a data column in database

Vitess

Architecture

app server

app server

app server

big data

↓ App   ↑ Vitess

Vitess

vtctld

topology

vtgate

vtgate

shard n

shard l

vttablet

mysqld

master

vttablet

mysqld

replicas

vttablet

mysqld

Analytics
replicas

VITESS.IO

Concepts: *cell*

- A group of servers and network infrastructure
- A failure domain: isolated from failures in other cells.
- Examples:
  - a full data center
  - a subset of a data center, aka availability zone.
  - a Kubernetes cluster

Vitess

Concepts: *keyspace*

- logical database.
- If no sharding, *keyspace* => single MySQL database
- If using sharding, *keyspace* => multiple MySQL databases (all with identical schema)
- In either case, a keyspace appears as a single database from the standpoint of the application.

Vitess

VITESS.IO

Concepts: *Keyspace ID*

- keyspace is sharded by keyspace ID ranges.
- Each row is assigned a keyspace ID
- "Street Address" of a row
- Used to decide on which shard a given row lives
- ~~Keyspace ID: equivalent of a NoSQL sharding key~~
- Internal to Vitess, The application does not need to know anything about it.
- Not stored, computed

Vitess

Concepts: *Vindex*

- A way to compute Keyspace ID for any row in a table
- Vindex for a Table is defined by
  - Column name
  - Sharding function name
- KeySpaceIDForRow = ShardingFunction(ColumnValueForRow)
- For example, table name: customer, sharding column: id sharding function: hash
- For a row where id is 123, KeySpaceId = hash(123)

Vitess

Concepts: *Shard*

- Keyspace ID range (Begin, End)
- If Begin < KeyspaceID <= End, then row with KeyspaceID belongs in that shard.
- One master
- Multiple replicas
- Located in one or more *cells*

Vitess

Concepts: *Sharding Functions*

Or, add your own custom sharding function!

binary	Identity
binary_md5	md5 hash
hash	3DES null-key hash
numeric	Identity
numeric_static_map	A JSON file that maps input values to keyspace IDs
unicode_loose_md5	Case-insensitive (UCA level 1) md5 hash
reverse_bits	Bit Reversal

Vitess

Demo: *Goals & Setup*

- Vitess cluster that stores data for a given country in its region
- Client application need not be aware of where data is stored
- Four regions/Eight countries
  - **Americas -> USA, Canada**
  - **Europe -> France, Germany**
  - **East Asia -> China, Japan**
  - **South Asia -> India, Indonesia**

Vitess

Cluster Design

- Vitess "cell" for each region
  - **gcpuscentral1a -> Americas**
  - **gcpeuropewest3a -> Europe**
  - **gcpasiaeast1a -> East Asia**
  - **gcpasiasouth2a -> South Asia**
- "keyspace" with four shards
  - one shard resident in each cell

Vitess

Sharding scheme

Region_vindex

- Based on new "multi column" vindex
- Maps an (id, country) tuple to keyspace_id
- Looks up region_byte for country using a map

Source

Keyspace -> Cell

- Split keyspace into keyranges
  - -40 -> gcpuscentral1a (Americas)
    - {US: 0x1, Canada: 0x2}
  - 40-80 -> gcpeurope3a (Europe)
    - {France: 0x40, Germany: 0x41}
  - 80-c0 -> gcpeastasia1a (East Asia)
    - {China: 0x80, Japan: 0x81}
  - c0- -> gcpsouthasia2a (South Asia)
    - {India: 0xc0, Indonesia: 0xc1}

Country to Region Map

```
{
    "United States": 1,
    "Canada": 2,
    "France": 64,
    "Germany": 65,
    "China": 128,
    "Japan": 129,
    "India": 192,
    "Indonesia": 193
}
```

Vitess

Sample Data

name, national_id, country
Philip Roth,123-456-7890,United States
Gary Shteyngart,234-567-8910,United States
Margaret Atwood,345-678-9120,Canada
Alice Munro,456-789-1230,Canada
....
Arundhati Roy,567-891-2340,India
Shashi Tharoor,678-912-3450,India
Andrea Hirata, 607-891-2340, Indonesia
Ayu Utami, 708-912-3450, Indonesia

Vitess

VITESS.IO

Demo

https://console.cloud.google.com/kubernetes/list?organizationId=575365795705&project=vitess-kubecon&supportedpurview=project

Google Cloud Platform   vitess-kubecon

Kubernetes Engine

Kuberne...usters   **CREATE CLUSTER**   **DEPLOY**   SHOW INFO PANEL

- Clusters
- Workloads
- Services & Ingress
- Applications
- Configuration
- Storage

A Kubernetes cluster is a managed group of VM instances for running containerized applications. Learn more

Filter by label or name

	Name ^	Location	Cluster size	Total cores	Total memory	Notifications	Labels	
☐	✅ demo-asiaeast2a	asia-east2-a	3	12 vCPUs	45.00 GB			Connect
☐	✅ demo-asiasouth1a	asia-south1-a	3	12 vCPUs	45.00 GB			Connect
☐	✅ demo-europewest3a	europe-west3-a	3	12 vCPUs	45.00 GB	💡 Node upgrade available		Connect
☐	✅ demo-uscentral1a	us-central1-a	3	12 vCPUs	45.00 GB	💡 Node upgrade available		Connect

Marketplace

The best is yet to be

Start with monolithic database and reshard

Move a country from one region/cell to another

Break a region into smaller shards

Vitess

Credits

Dan K.
@MinisterOfEng

Anthony Yeh
@enisoc

Shaun M.
@shaun_alan

Vitess

More Vitess talks

Gone in 60 Minutes: Migrating 20 TB

from AKS to GKE in an Hour with Vitess

Derek Perkins, Nozzle

Thursday 2:25pm - 3:00pm Room 14AB

Vitess

Questions?

Deepthi Sigireddi

@ATechGirl

Jitendra Vaidya

@planetscaledata

@yaempiricist

Vitess

@VITESSIO   VITESS.IO