Vitess: online schema migration automation

Shlomi Noach

PerconaLive EU 2020

About me

Engineer at PlanetScale

Author of **orchestrator**, **gh-ost**, **freno** and others

Blog at openark.org

github.com/shlomi-noach

@ShlomiNoach

PlanetScale

- Founded in February 2018
- Venture backed: a16z, SignalFire
- ~40 employees distributed
- Authors Vitess
- Offers **PlanetScaleDB** for cloud offering and **PlanetScaleDB BYOK**: Bring Your Own Kubernetes

Vitess

A database clustering system for horizontal scaling of MySQL

- CNCF graduated project
- Maintained by the community
- PlanetScale is a major contributor

Vitess architecture

Consider a simple replication topology

Vitess architecture

vttablet

- A daemon/sidecar
- Controls the **mysqld** process
- Interacts with the **mysqld** server
- *Typically* on same host as **mysqld**

Vitess architecture

Vitess architecture

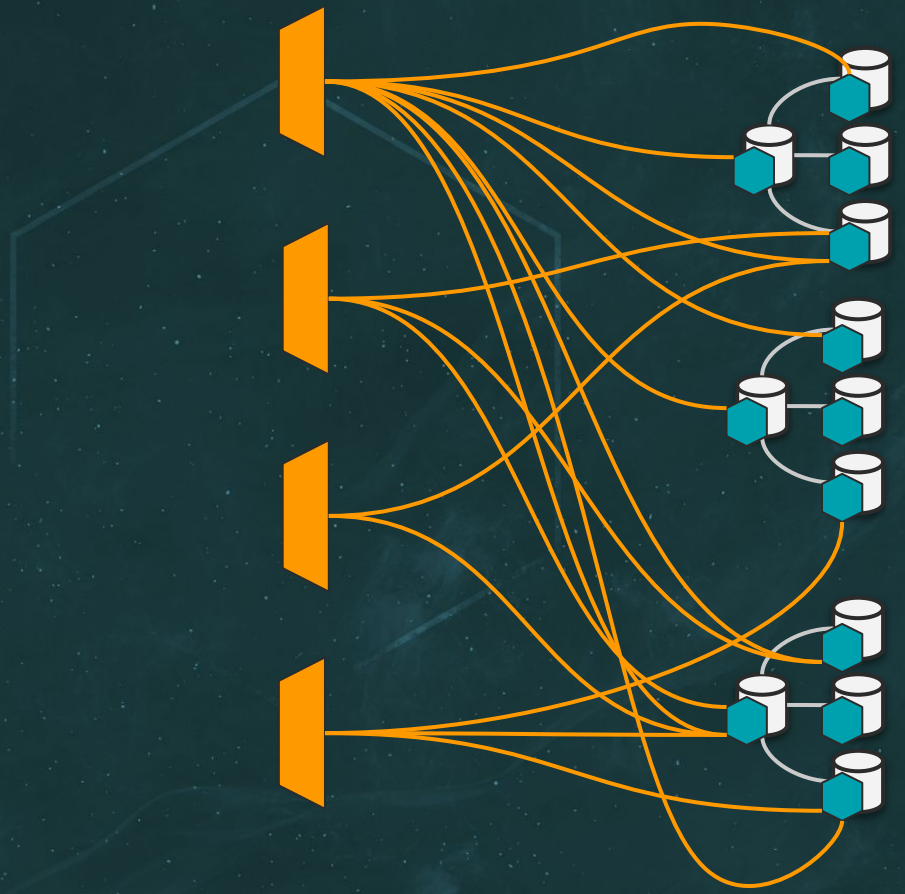
vtgate

- A smart proxy
- Speaks **mysql** protocol
- Impersonates the database
- Relays queries to tablets

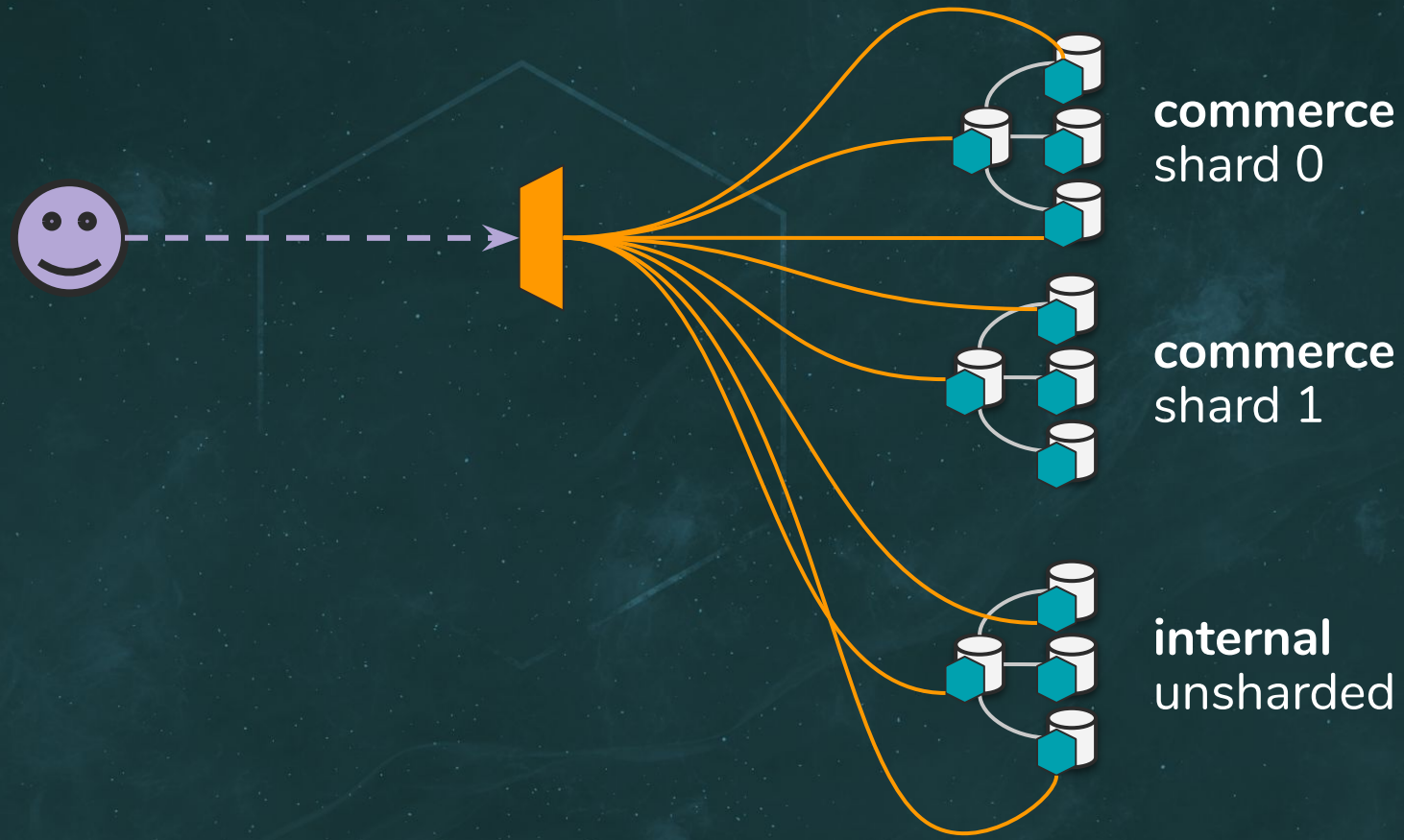
Vitess architecture

vtgate

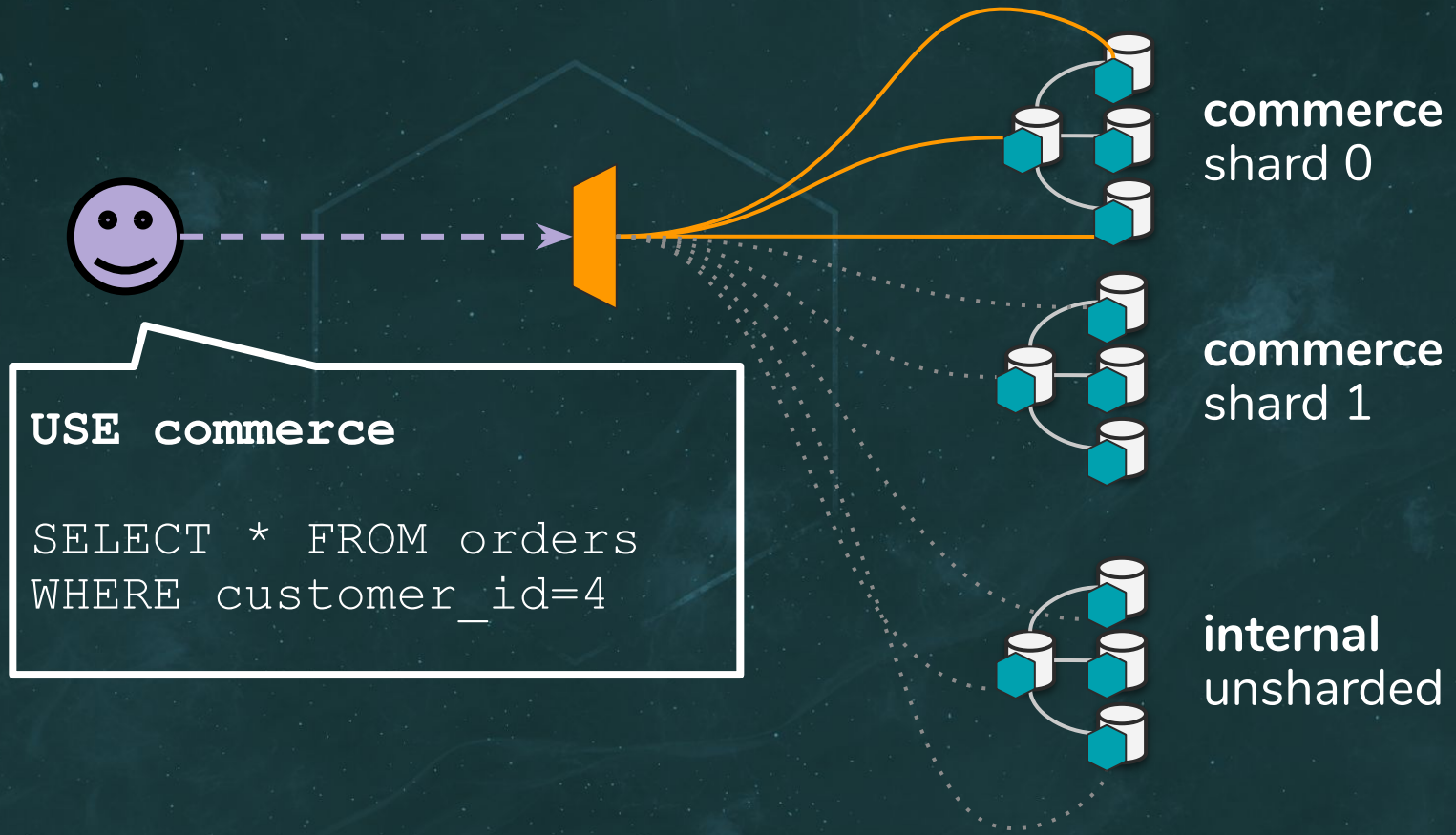
- Stateless
- Scale out



Vitess architecture



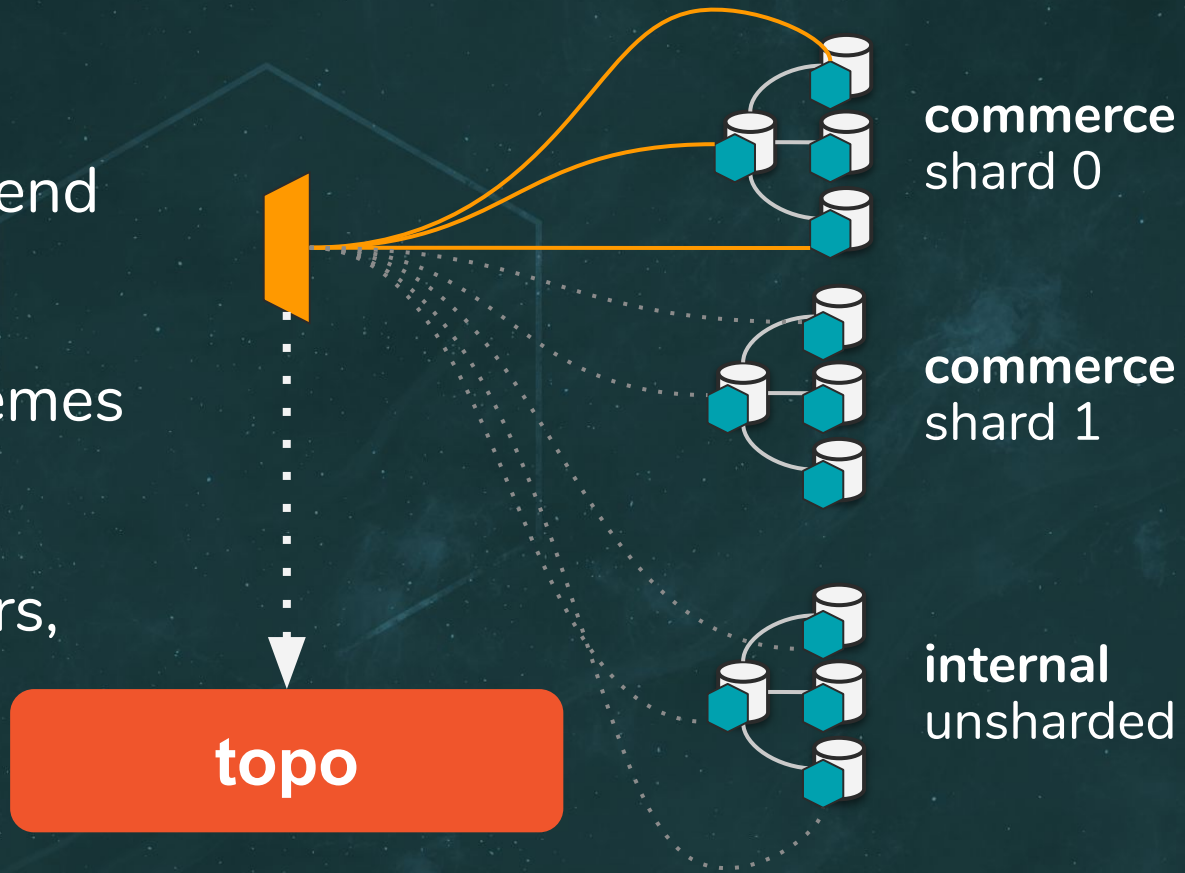
Vitess architecture



Vitess architecture

topo

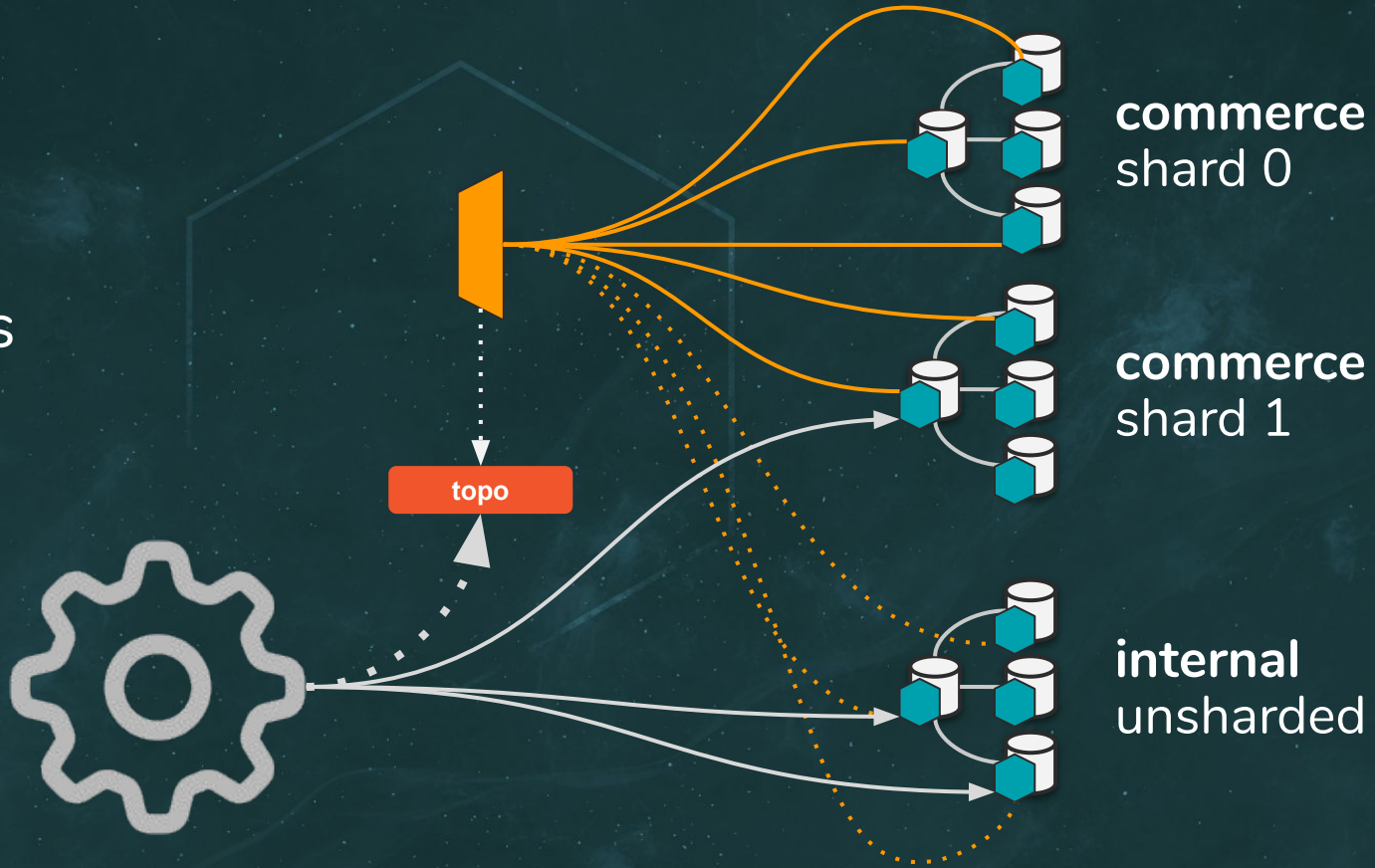
- Vitess state backend
- etcd/zk/consul
- Info:
  - Sharding schemes
  - Schemas
  - Clusters
  - Tablets, servers,
  - Operations
  - ...



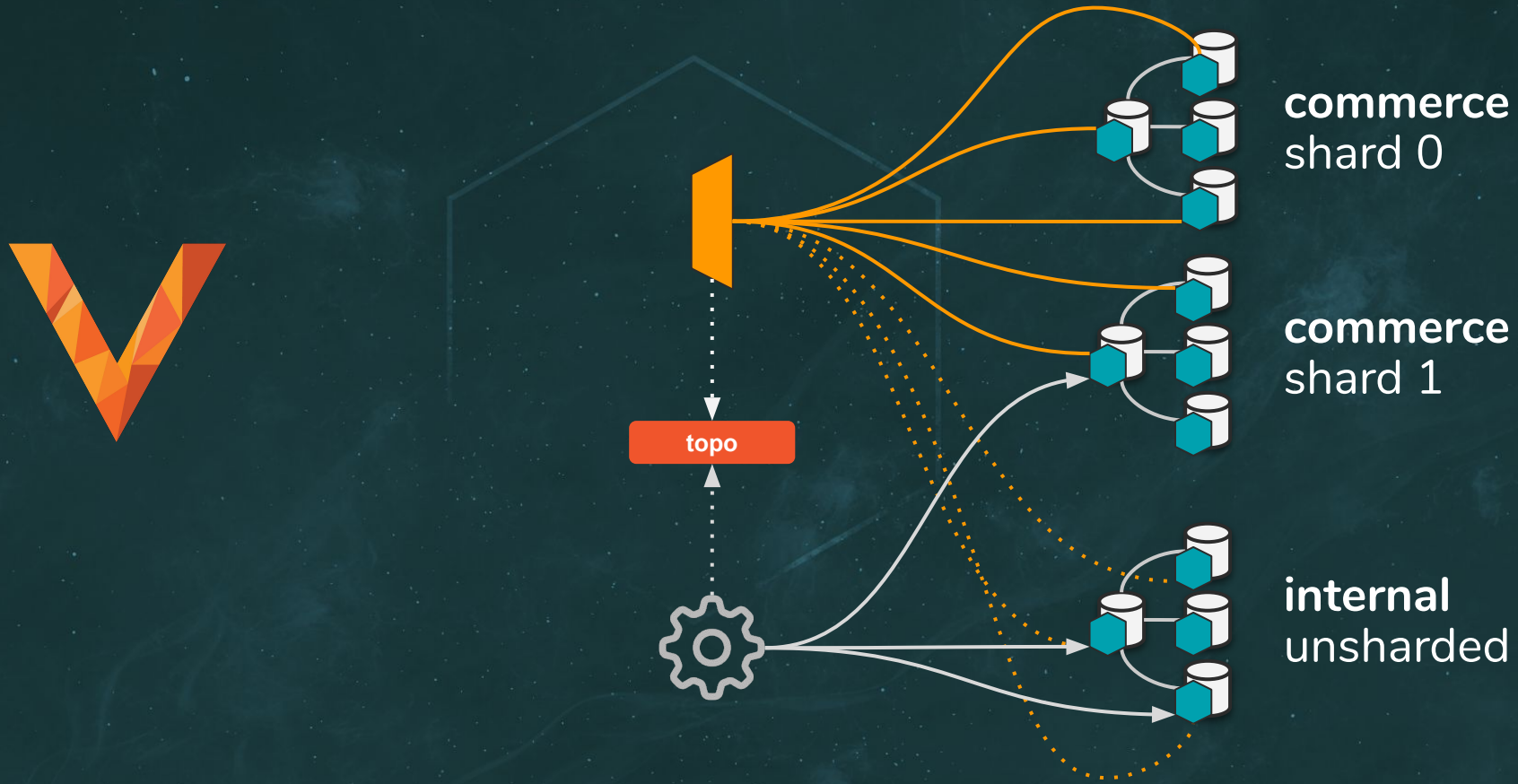
Vitess architecture

vtctld

- Daemon
- API server
- Operations



Vitess architecture



Mystery card

A hidden truth

Issues with schema migrations

- **ALTER TABLE** is blocking

Issues with schema migrations

- ALTER TABLE is blocking
- **gh-ost**, **pt-online-schema-change**, **fb-osc** as alternatives

  Installing, formalizing, command line options

Issues with schema migrations

- ALTER TABLE is blocking
- gh-ost, pt-online-schema-change, fb-osc as alternatives
- **Discovery**

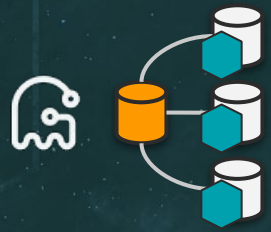
internal
unsharded

commerce
shard 0

commerce
shard 1

Issues with schema migrations

- ALTER TABLE is blocking
- gh-ost, pt-online-schema-change, fb-osc as alternatives
- Discovery
- **Running**

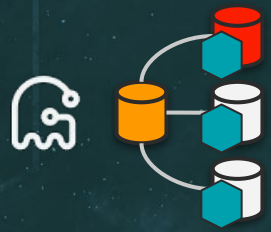


commerce
shard 0

commerce
shard 1

Issues with schema migrations

- ALTER TABLE is blocking
- gh-ost, pt-online-schema-change, fb-osc as alternatives
- Discovery
- Running
- **Throttling**

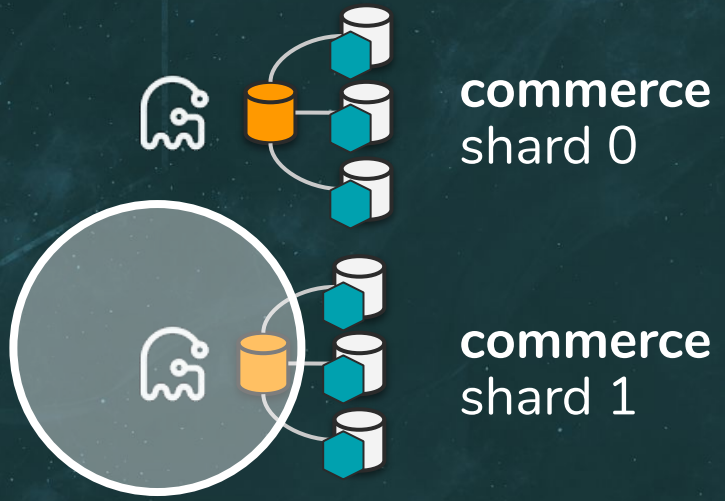


commerce
shard 0

commerce
shard 1

Issues with schema migrations

- ALTER TABLE is blocking
- gh-ost, pt-online-schema-change, fb-osc as alternatives
- Discovery
- Running
- Throttling
- **Visibility**



Issues with schema migrations

- ALTER TABLE is blocking
- gh-ost, pt-online-schema-change, fb-osc as alternatives
- Discovery
- Running
- Throttling
- Visibility
- **Control**

commerce
shard 0

commerce
shard 1

Issues with schema migrations

Vitess' architecture has the knowledge and the capability to address the above.

Vitess Online DDL

```
ALTER WITH 'gh-ost' TABLE
orders ADD COLUMN ts TIMESTAMP
```

ZeroDependenciesNoReally™

Vitess Online DDL

```
ALTER WITH 'pt-osc' TABLE
orders ADD COLUMN ts TIMESTAMP
```

ZeroDependenciesNoReally™

Vitess Online DDL

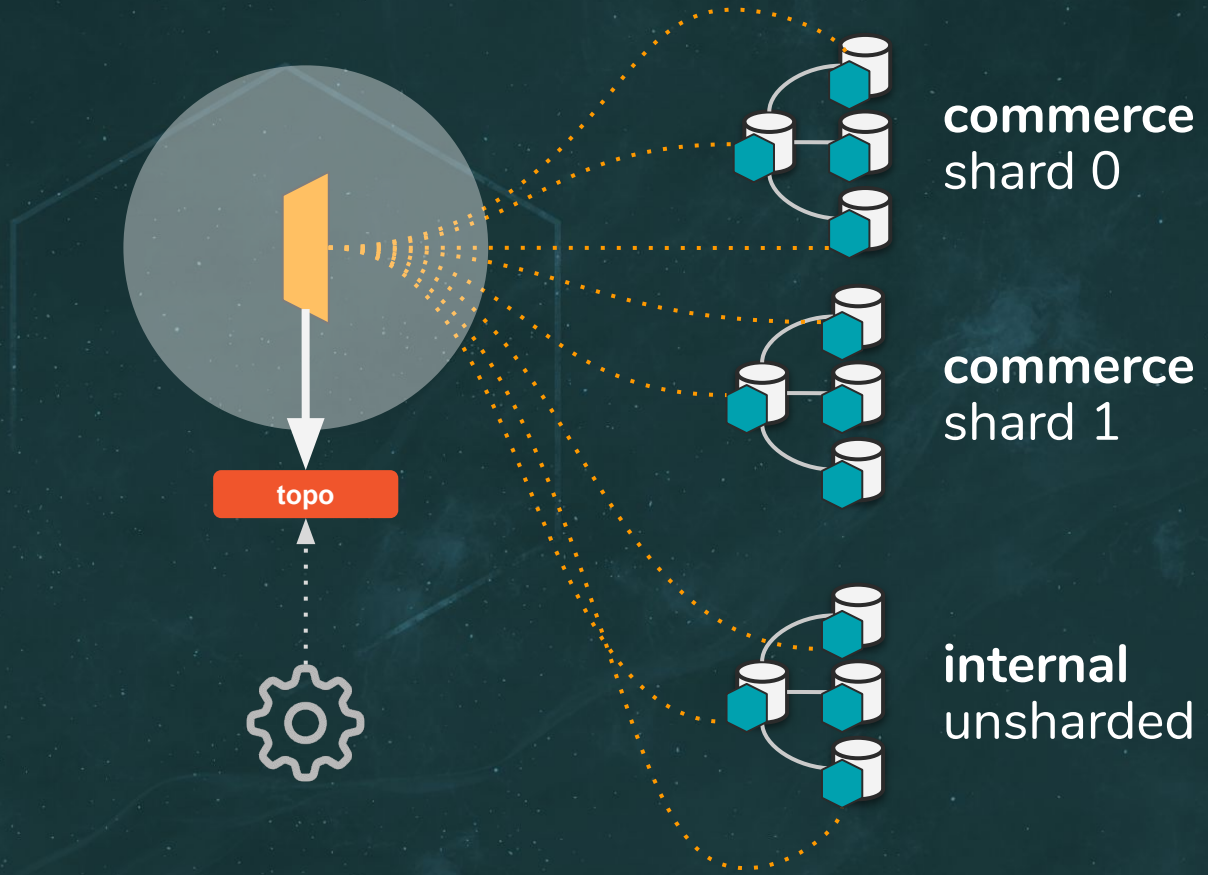
DEMO

What just happened?

Online DDL flow

vtgate

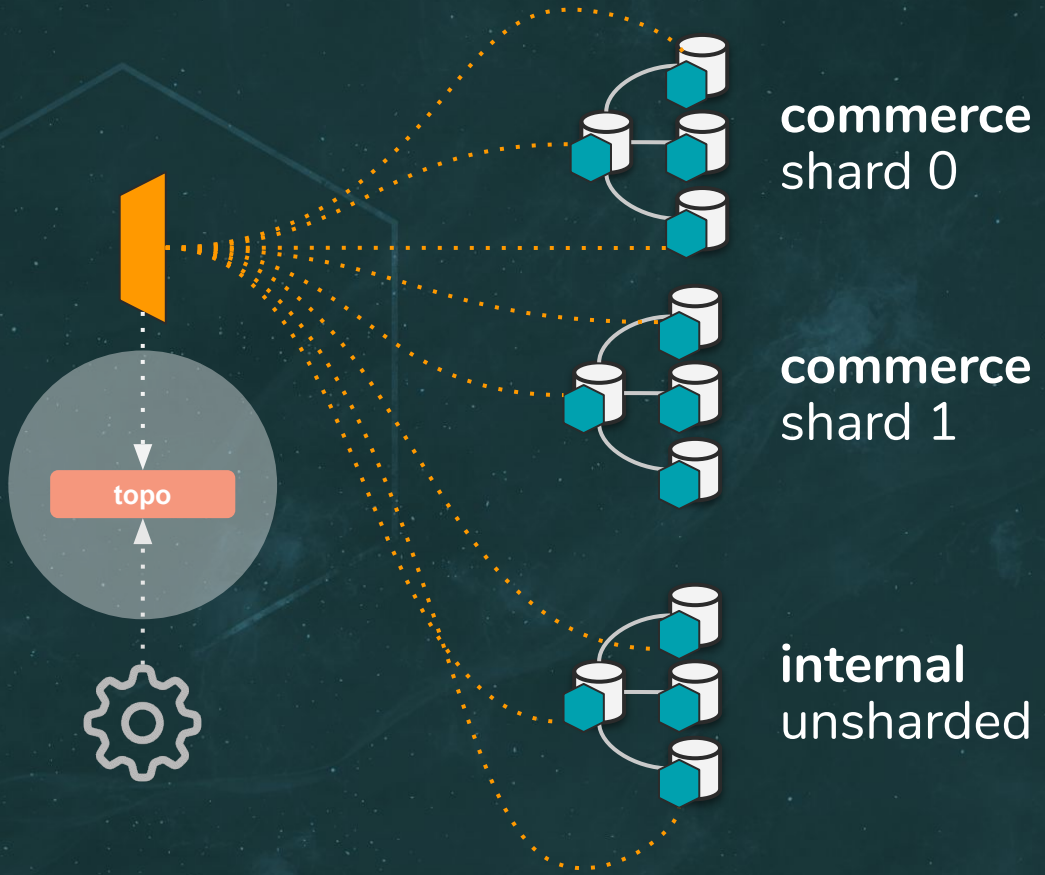
- Intercepts
- Parses
- Persists



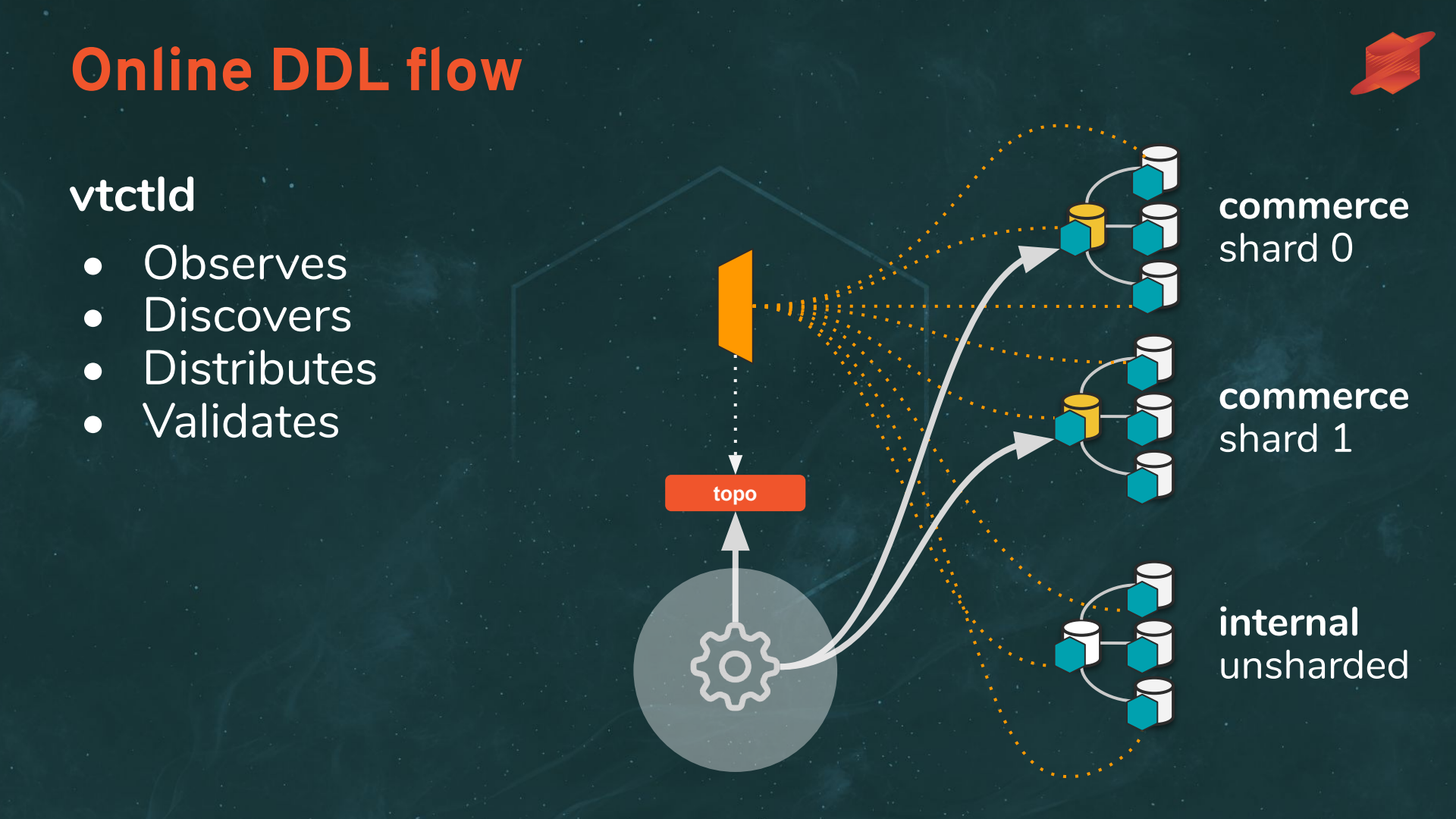
Online DDL flow

topo

- Stores



Online DDL flow



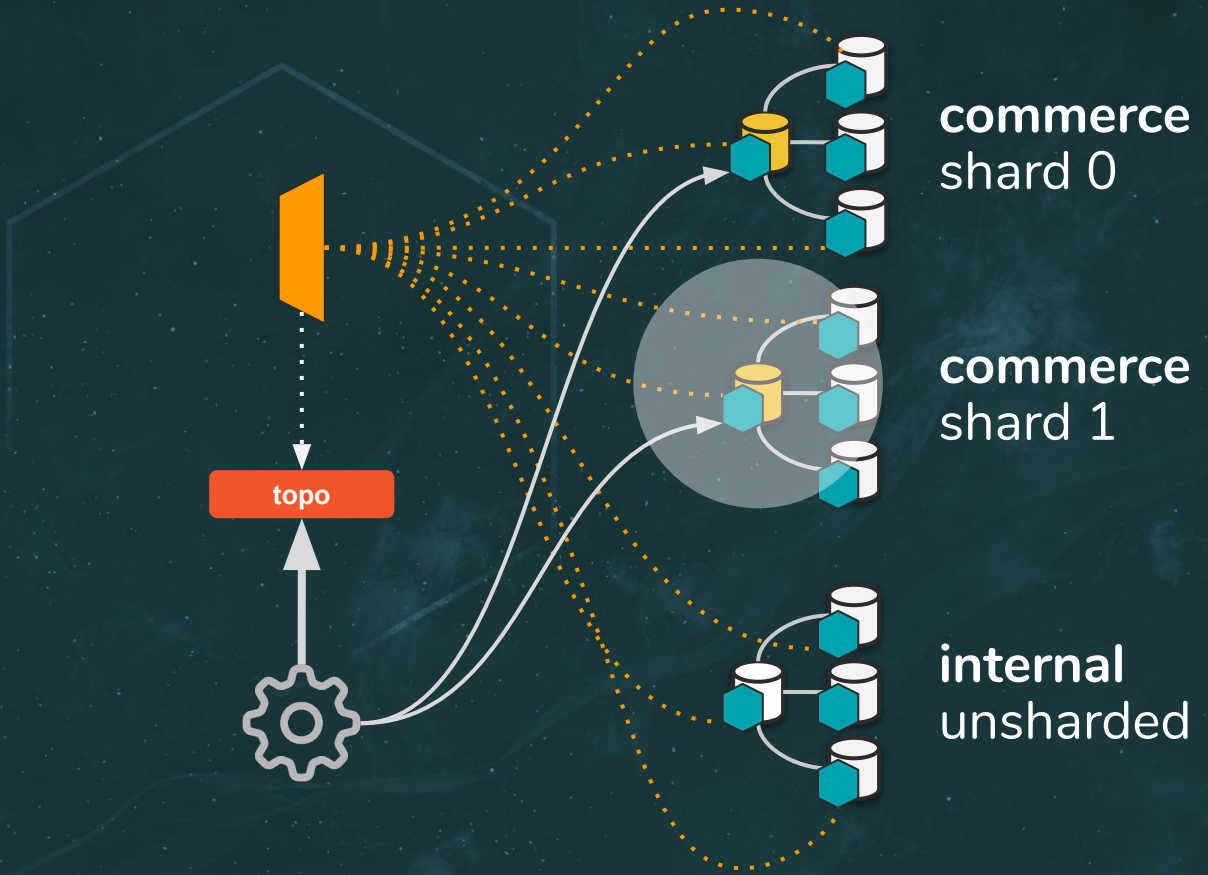
vtctld

- Observes
- Discovers
- Distributes
- Validates

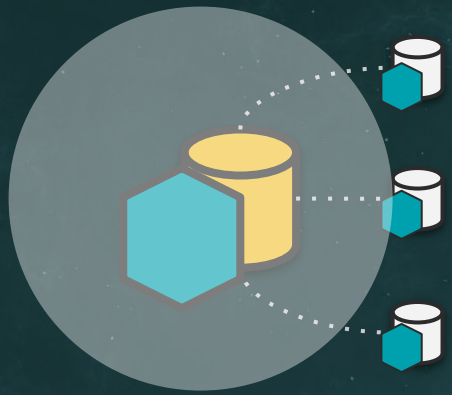
Online DDL flow

vttablet

- Has all the fun



Online DDL flow



vttablet on primary

- Schedules
- Prepares script(s)
- Runs **gh-ost** or **pt-osc**
- Tracks
- Throttles
- Cleans up
- Tracks state in local **_vt** meta schema

DEMO, continued

Vitess Online DDL

ZeroDependenciesNoReally™

Q: Which MySQL account and privileges do I need for the migration user? How do I tell Vitess the credentials?

A: Vitess creates a migration user with random password for each migration, on your behalf.
You do not need to do anything.

Q: How do I setup the gh-ost hooks?

A: Vitess creates **gh-ost** hooks directory and scripts for you.
Vitess creates a **pt-osc** plugin for you.
You do not need to do anything.

Q: Where do I install gh-ost?

A: On linux amd64 and with recent glibc, Vitess comes with gh-ost binary pre-compiled and embedded, and will extract the binary on your behalf.
You do not need to do anything.

Q: Where do I install gh-ost? (Cont.)

Otherwise use **-gh-ost-path**.

pt-online-schema-change is not bundled. Install it on all MySQL servers. Use **-pt-osc-path**

Q: How do I configure the throttling replicas?

A: Each **vttablet** runs a throttler service, which automatically and dynamically detects the throttling replicas. It evaluates replication lag autonomically.
Both **gh-ost** and **pt-osc** are configured to use it.
You do not need to do anything (but you can configure `-throttle_tablet_types="replica,rdonly"`)

Q: How do I drop pt-osc leftover triggers in case of failure?

A: Vitess will identify a failed migration, even if **vttablet** itself fails, and will drop the triggers on your behalf.
You do not need to do anything.

Q: How do I drop the leftover tables?

A: Each primary **vttablet** runs a table lifecycle process. It will collect the leftover tables (whether migration is successful or failed) and will safely and slowly (using the throttler mechanism) purge and drop them.
You do not need to do anything.

Mystery card reveal

A hidden truth,

In something not said

Vitess as an infrastructure framework

None of the above requires you to run traffic through Vitess.

You may keep running your production traffic as normal, if you wish, and only use Vitess for infrastructure.

Acknowledgements

- GitHub and GitHub team:
  - github.com/github/gh-ost
  - github.com/github/freno
  - Lessons learned
- Percona
  - percona.com/doc/percona-toolkit
- Community
- Vitess maintainers
- PlanetScale team

Resources

- Docs: vitess.io/docs
- Code: github.com/vitessio/vitess
- Slack: vitess.slack.com
- Demo:
  ```
  asciinema play https://git.io/JUhMa
  ```

Thank you!

Questions?

github.com/shlomi-noach
@ShlomiNoach