# A Successful migration from ElasticSearch to ClickHouse

CONTENTSQUARE

# Agenda

What are we doing at Contentsquare

What have we done with ClickHouse

Our challenges during this migration

CONTENTSQUARE

# Who are we ?

Christophe Kalenzaga
Data Engineer

Vianney Foucault
Platform Engineer

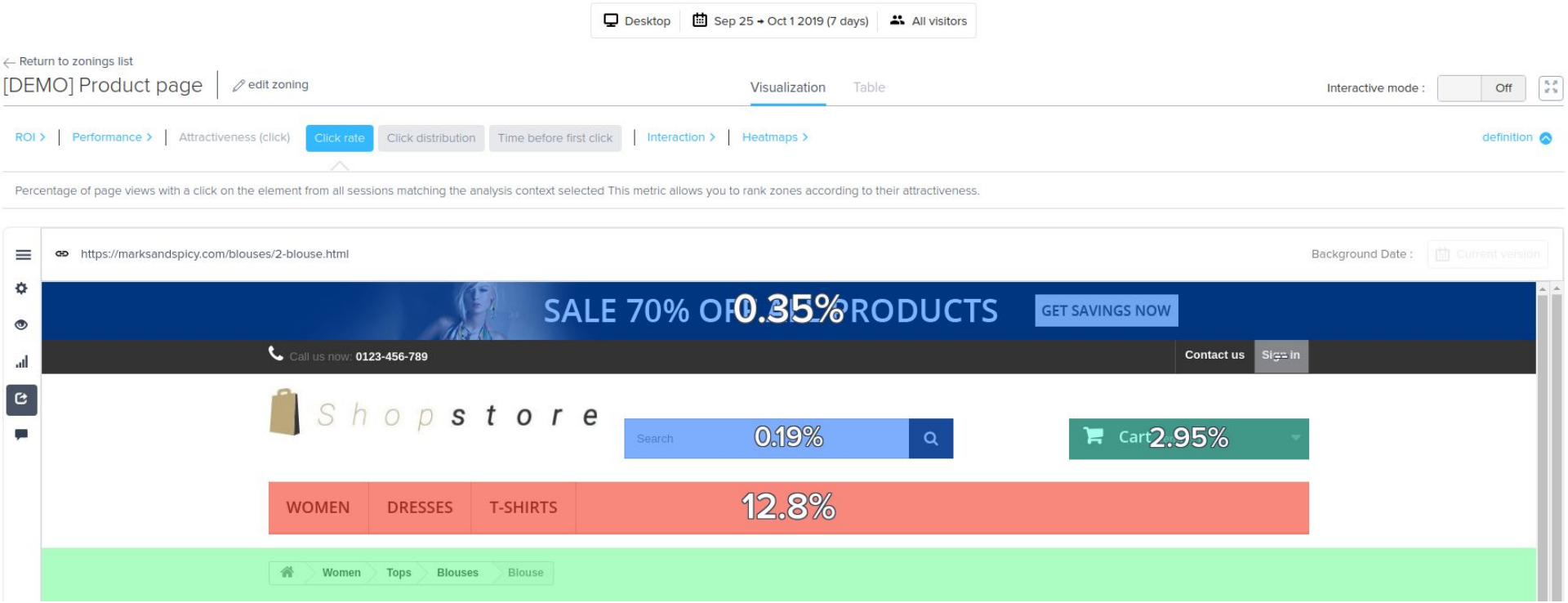
CONTENTSQUARE

# ContentSquare in a few words

- Data analytics Company
  - Analysis of websites and mobile applications
  - 1.3 TBytes collected per day
  - 13 months retention (~500 TBytes)

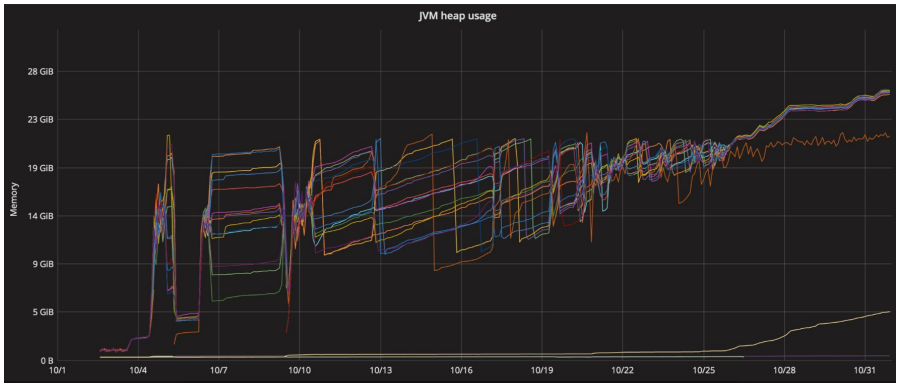
- 600+ clients Worldwide
  - multiple time zones
  - 60k queries per day

# One of the many challenges at ContentSquare

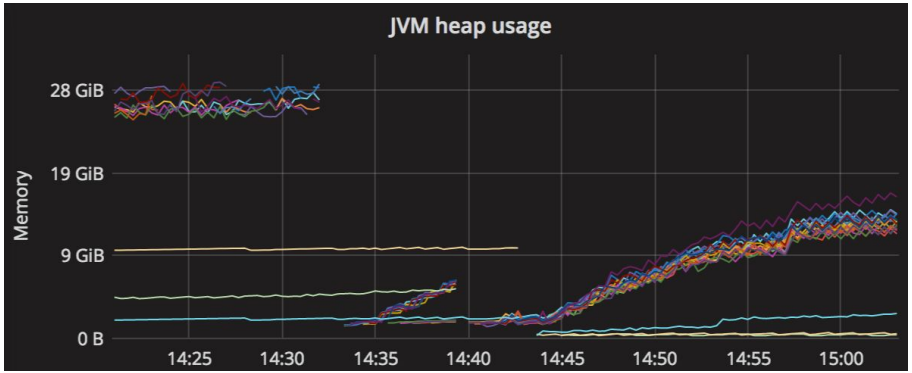


# The road from ElasticSearch to ClickHouse

# Guess what this is




An ElasticSearch Cluster that's going to crash




The crash of an ElasticSearch Cluster in production

# Reasons to leave ElasticSearch

- Stability
- Cost
- Scalability
- Analytics features

# Why ClickHouse

- Open Source
- Fast
- Free
- Created by people in the same industry
- No cloud provider lock-in
- community

# Timeline of our migration

**Mar 2018** — Start our brand new analytics solution for mobile application on ClickHouse

**Jun 2018** — the mobile analytics solution is released

**Aug 2018** — Start working an a new web analytics solution on ClickHouse


**Jun 2019** — The first web client is migrated to our new ClickHouse platform

**Sep 2019** — The last web client is migrated to our new ClickHouse platform

# Gains from switching from ElasticSearch to ClickHouse

- CH is 11 times cheaper !
- Queries are 4 times faster on average
- Queries are 10 times faster for the 99 percentile of latencies
- Rock solid stability
- ClickHouse stores  6 times more data

# First Challenge
# How to reduce the overhead of the insertion?

# First Challenge: how to scale insertion without scaling ClickHouse?

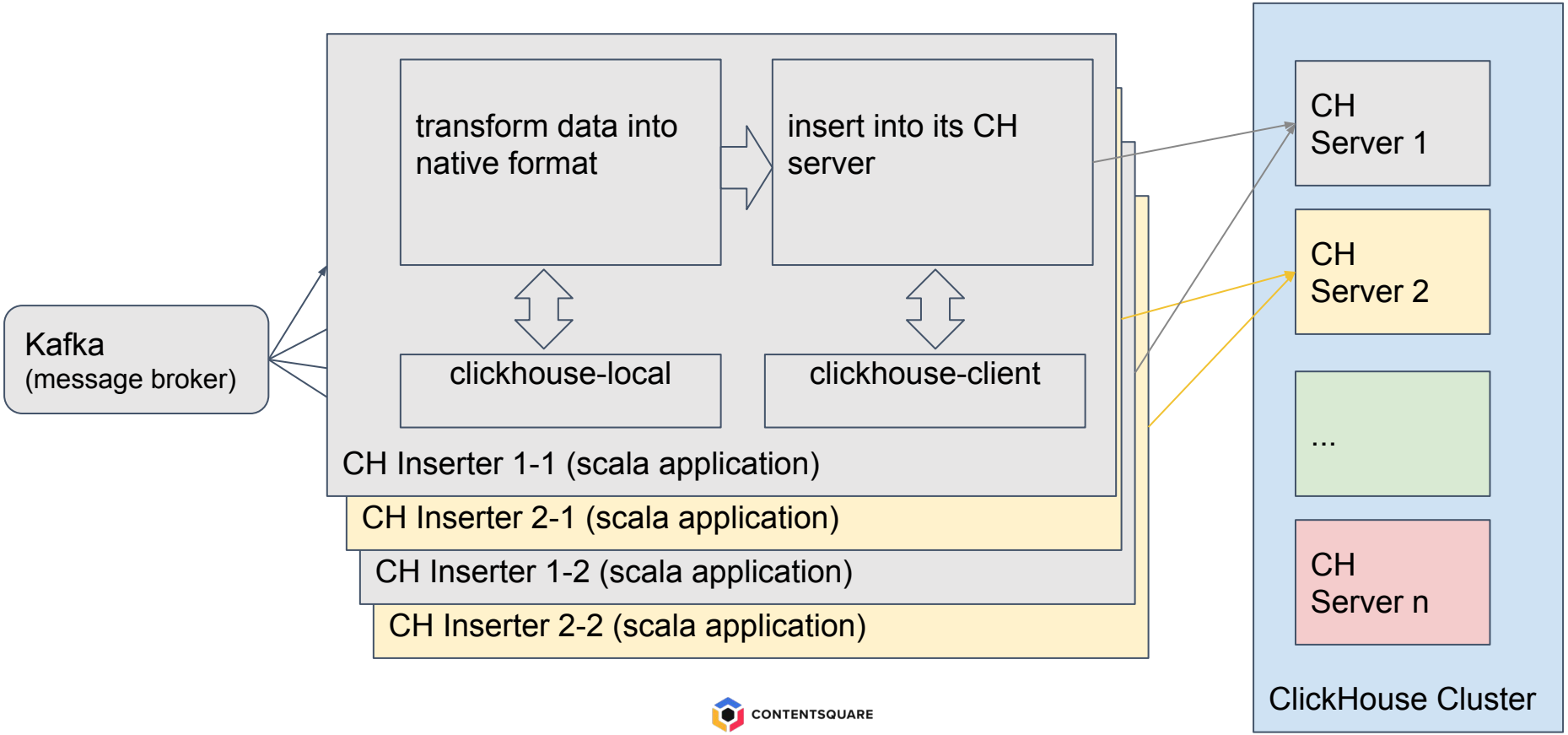
- Massive insertions consume CPU and memory

Solution: insert in clickhouse using the native format

- Insertions in distributed tables consume network and disks

Solution: insert data in clickhouse directly in the right server

# First Challenge: how to scale insertion without scaling ClickHouse?



# Second Challenge
# How to design queries?

CONTENTSQUARE

# Second Challenge: how to design optimized queries?

## Things to know when designing queries

- CH evaluates all conditions of a query

```
SELECT count(1) FROM my_table
WHERE cond_1 AND cond_2 AND cond_3
=> cond_3 will be evaluated even if cond_1 is false
```

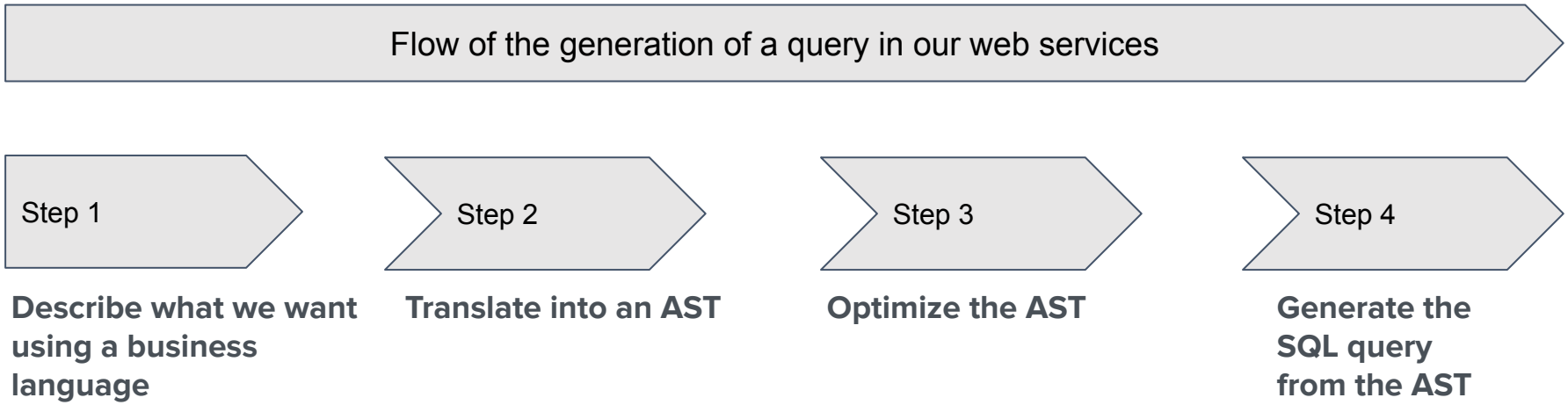
- CH evaluates duplicates conditions only once

```
SELECT  countIf(cond_1 AND cond_2),
        countIf(cond_1 AND cond_3)
FROM my_table
=> cond_1 will be evaluated only once
```

- CH doesn't behave well with many subqueries or joins

- ...

# Second Challenge: how to design optimized queries?

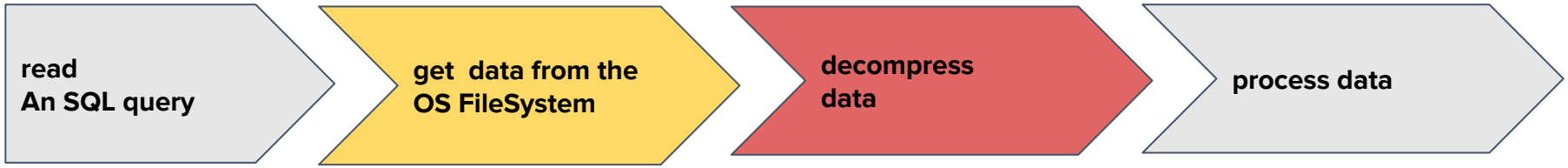


# Third Challenge
# How to reduce storage costs?

CONTENTSQUARE

# Third Challenge: how to reduce storage costs?

- Understand how ClickHouse manages data during a query




- Understand the concept of COLD and HOT queries on ClickHouse

- Understand the pros and cons of each compression codecs
  - Generic: ZSTD, LZ4, None
  - Specialized: Dictionary Encoding, Gorilla, T64, Delta

# Third Challenge: how to reduce storage costs?

- Understand the pattern of your queries
  - Are all columns evenly used?
  - Are people doing most queries on the same dataset?

- Benchmark on realistic workloads
  - Benchmark on worst case scenarios
    - CH 10 times faster on fast disks than on slow disks
  - Benchmark with queries from production
    - CH 1.2 times faster on fast disks than on slow disks

# Fourth Challenge
# How to avoid functional regression?

CONTENTSQUARE

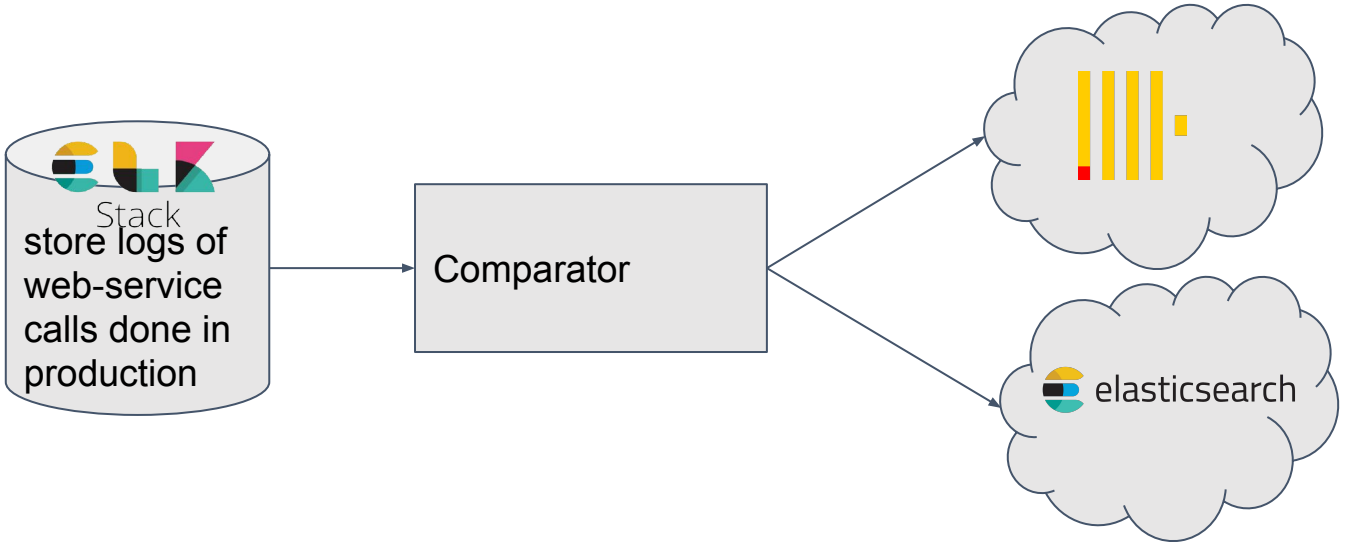
# Forth Challenge: how to avoid functional regression?

- Major reasons for regressions

- Difficulties to spot regressions

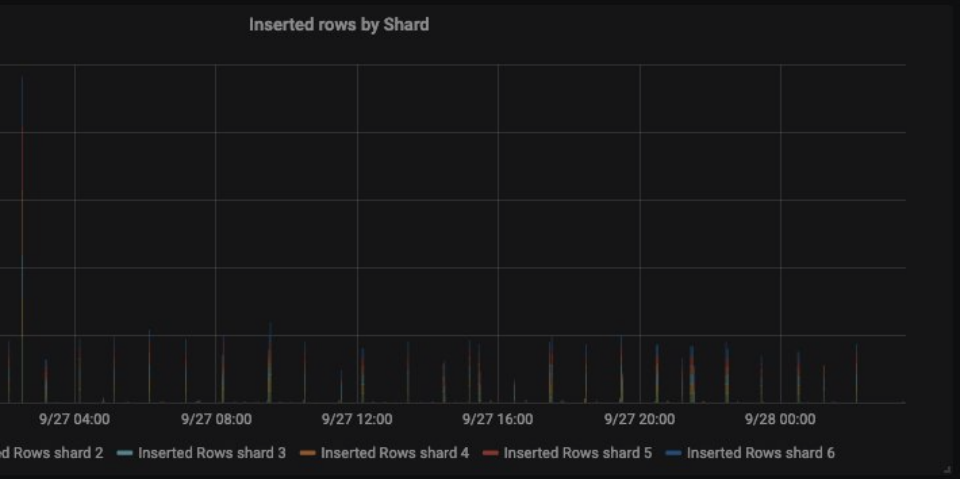
# Forth Challenge: how to avoid functional regression?

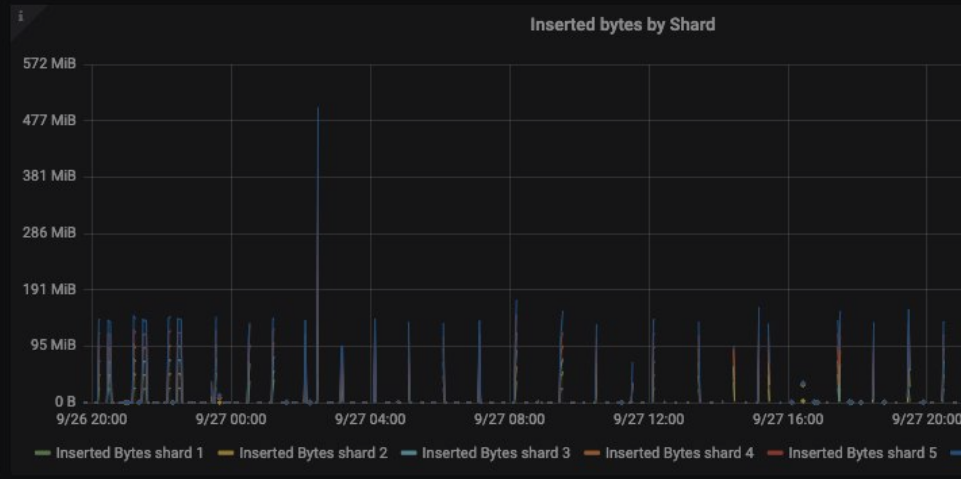
- **How we did it**

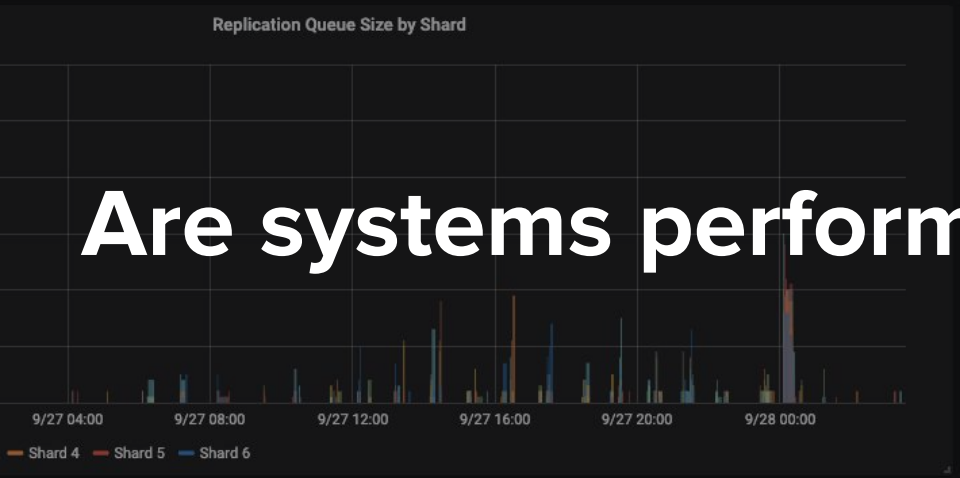



[Watch our detailed presentation on youtube]

# Fifth Challenge
# How to make ClickHouse production ready?





# Are systems performing well ?

# Observability

- **Get relevant metrics**

  Data inserted per shard
  Replication queues,
  Clickhouse Query time

- **Get relevant logs**

  Clickhouse logs, clickhouse err logs
  Application logs

- **Create relevant alerts**

  A replica is down
  A shard is down
  Clickhouse query time goes up...

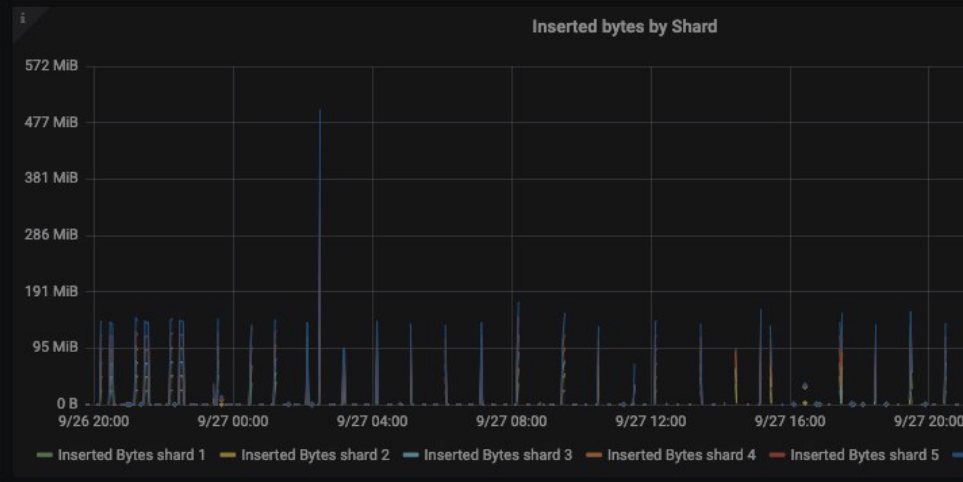
```
alert: clickhouse_replicas_unhealthy
expr: clickhouse:replicas:unhealthy
  != 0
for: 5m
labels:
  alertname: clickhouse_replicas_unhealthy
  severity: critical
  team: de
annotations:
  description: 'Presence of unhealthy replicas on
ClickHouse table {{ $labels.table}}
    on {{ $labels.Name }} : {{ $value }} != 0'
  title: 'ClickHouse unhealthy replicas on table {{
$labels.table }} on {{ $labels.Name
    }} : {{ $value }} != 0'
```

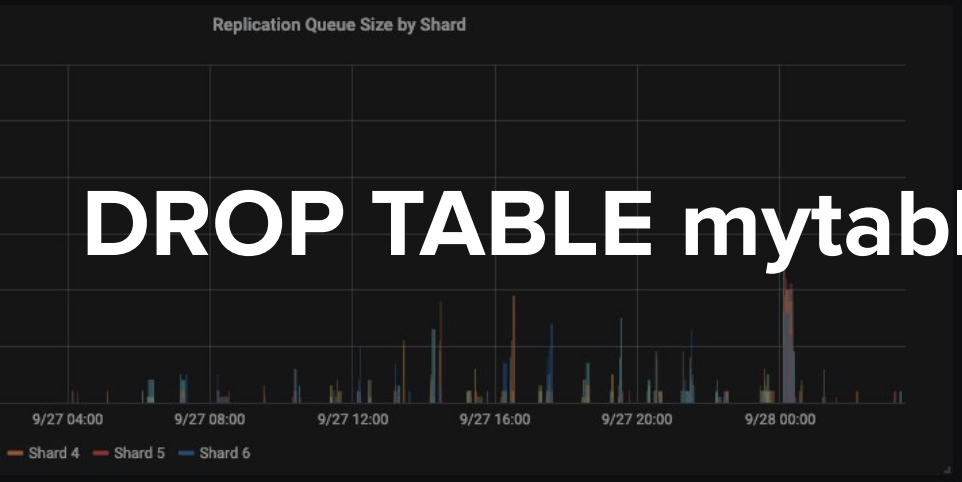
| 152.641 | 313.924 | 794.332 | 1,648.09 | 3,157.422 |
|---|---|---|---|---|
| 50th percentile of latency (ms) | 75th percentile of latency (ms) | 95th percentile of latency (ms) | 99th percentile of latency (ms) | 99.9th percentile of latency (ms) |


elasticsearch

Prometheus / Alerts Manager

pagerduty



# DROP TABLE mytable;


BACK UP YOURSELVES
DATA LOSS IS COMING

## Not so Built in features

Backup

Restore

Data Expiration

Clickhouse 19.14.3.3

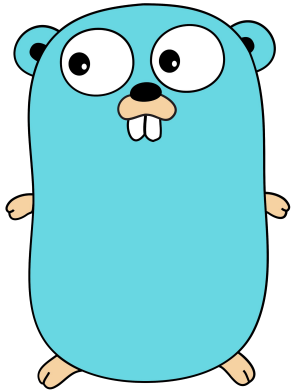
## Not so Orchestrated features

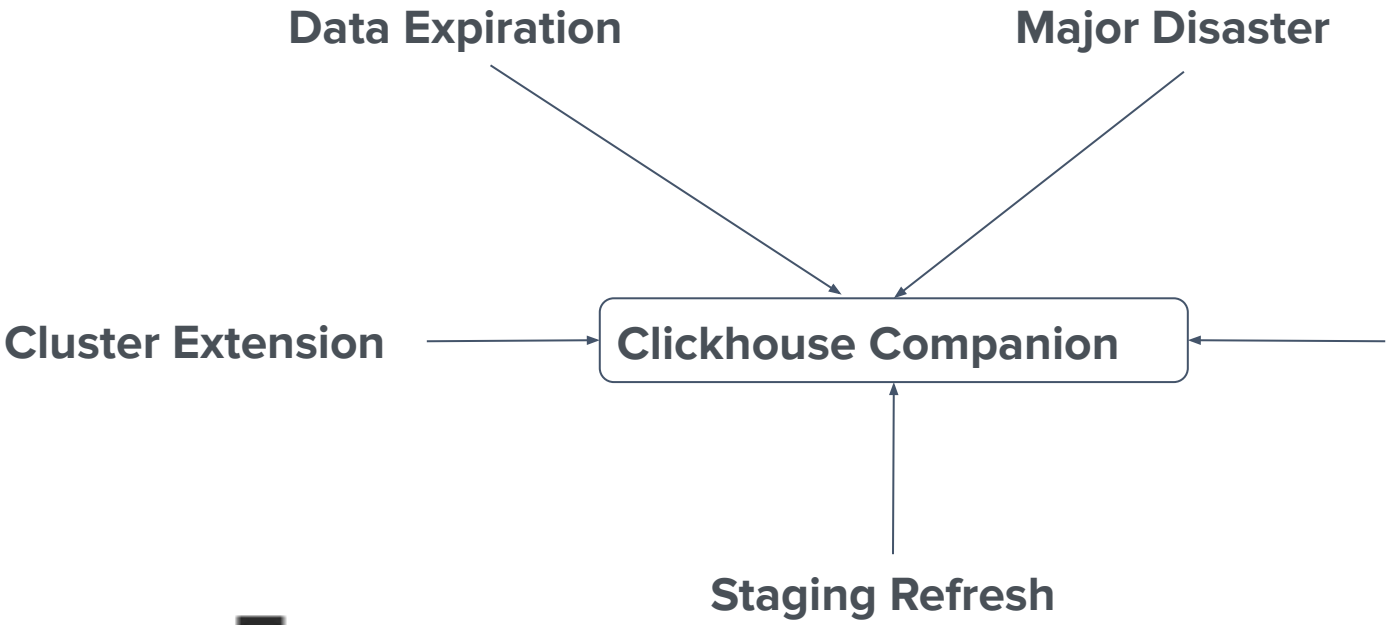
Clickhouse copier

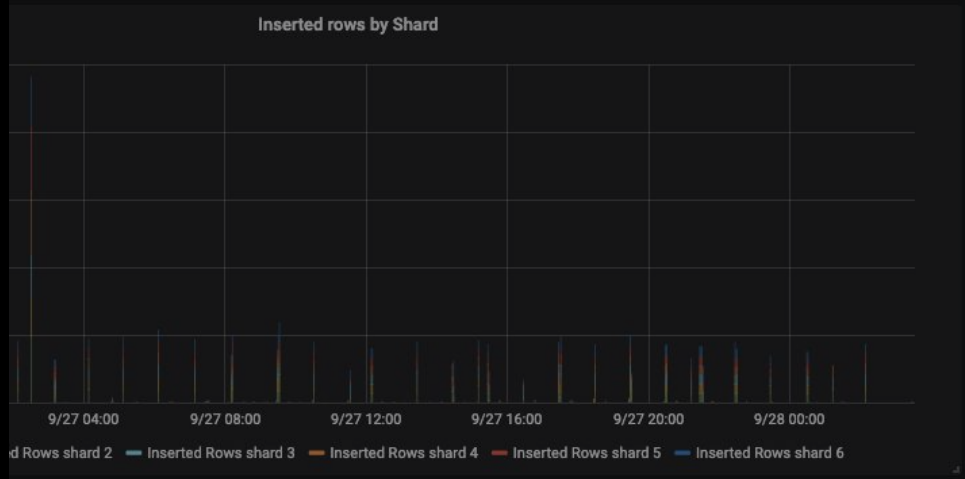
# Clickhouse Companion: Challenges

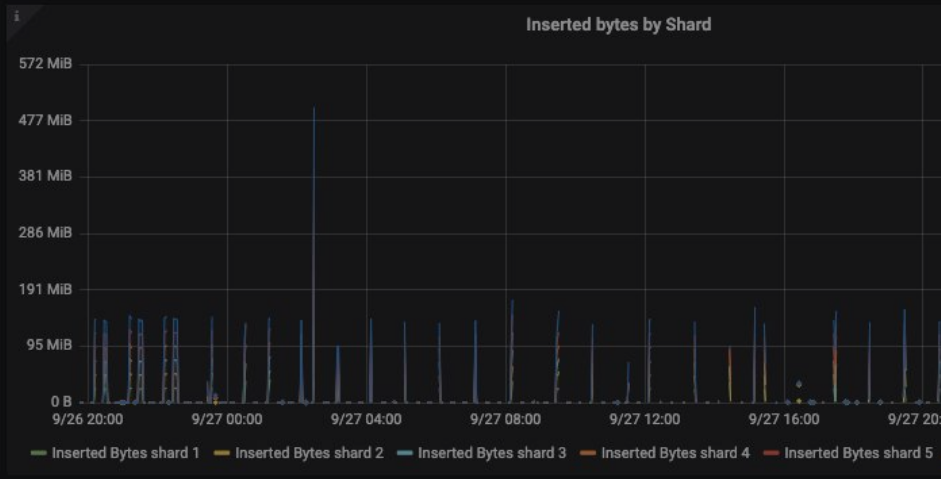
- Copy Partition data to an object storage
- Keep track of previous backups
- Update partition data when parts are merged
- Cleanup data and backup when data is expired
- Extra: admin/info tasks
- Orchestrate clickhouse copier

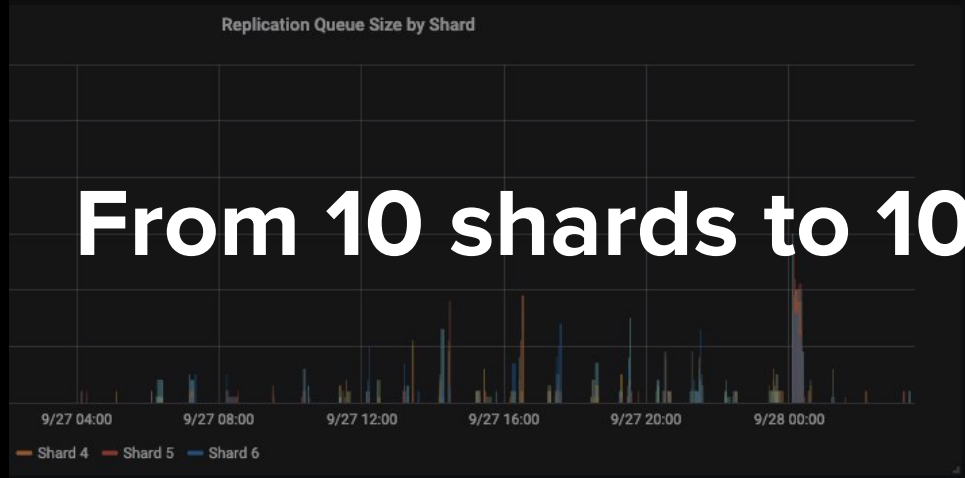
# Clickhouse Companion







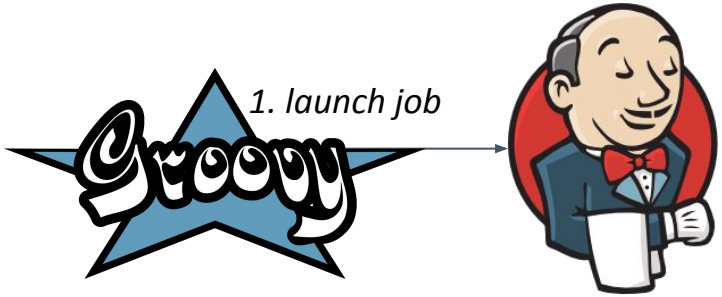





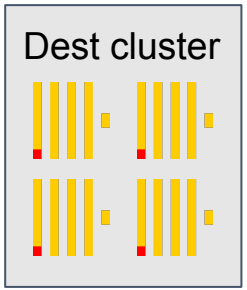
# From 10 shards to 100 shards


ONE DOES NOT SIMPLY
USE CLICKHOUSE COPIER
imgflip.com

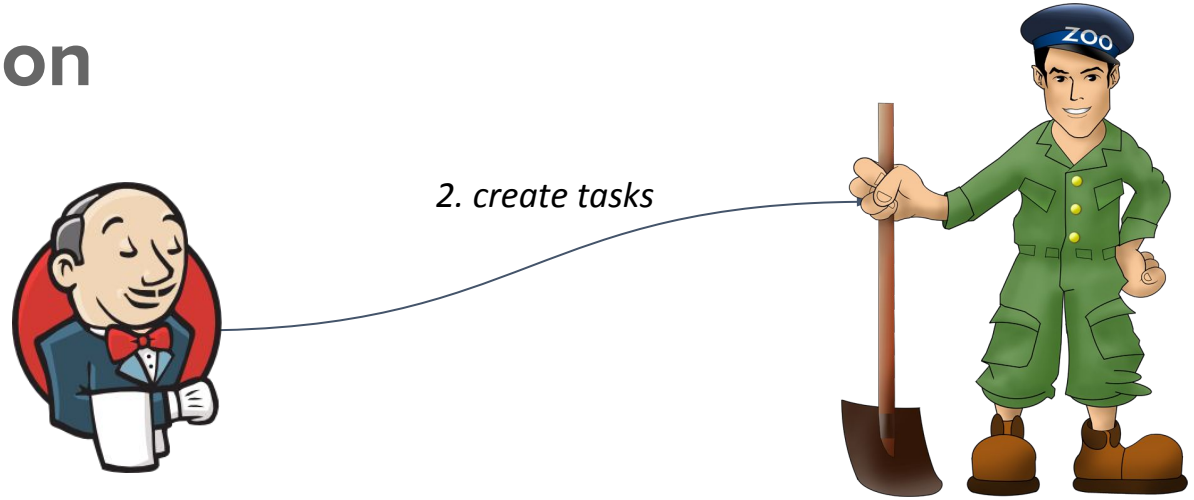
# Data migration



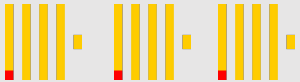


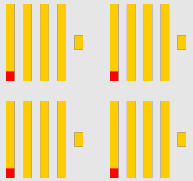
# Data migration



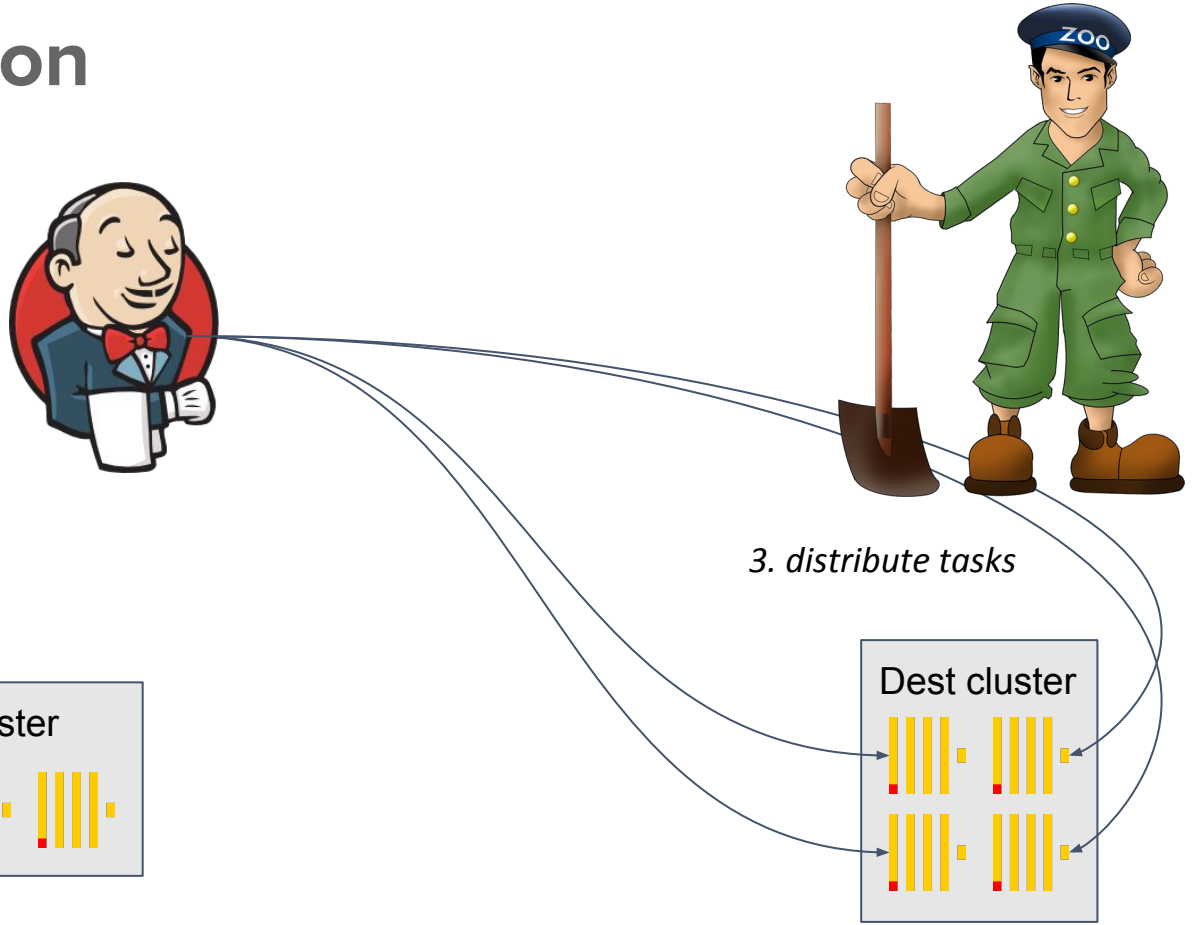

Source cluster



Dest cluster



# Data migration



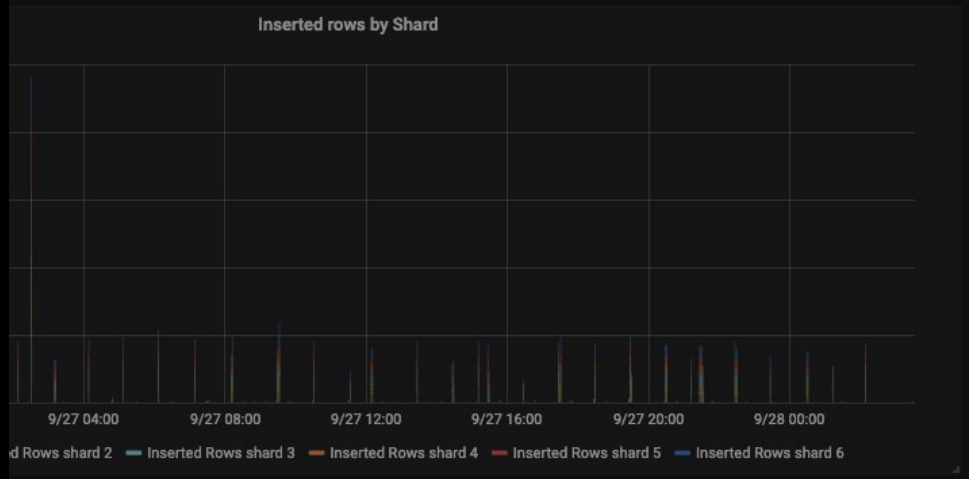
# Data migration

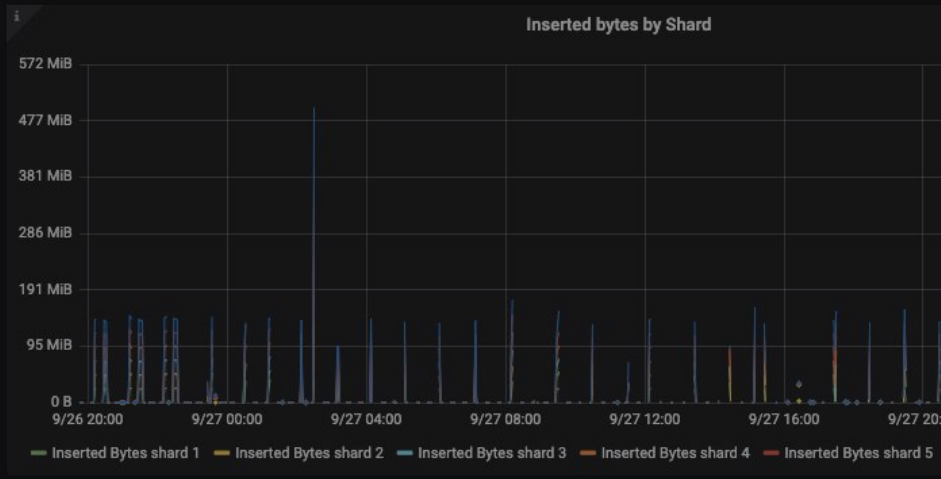

4. get tasks

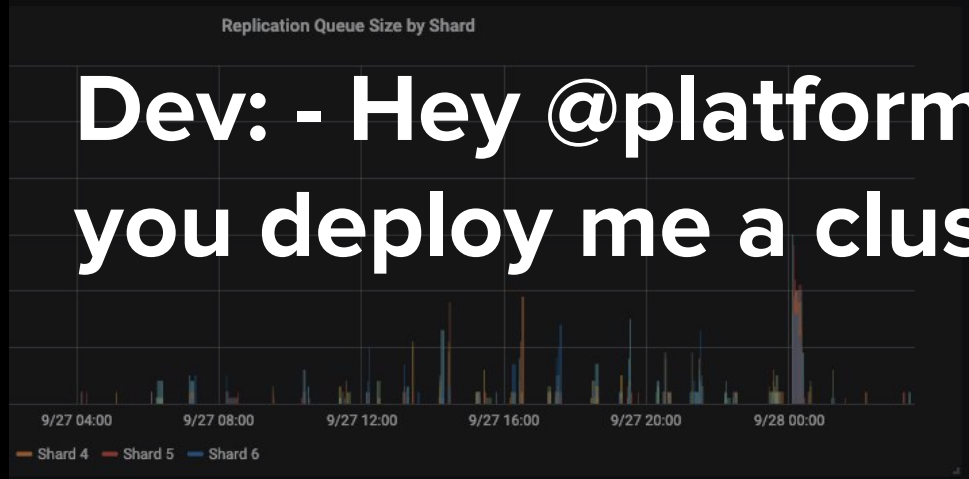
5. execute
clickhouse-copier tasks

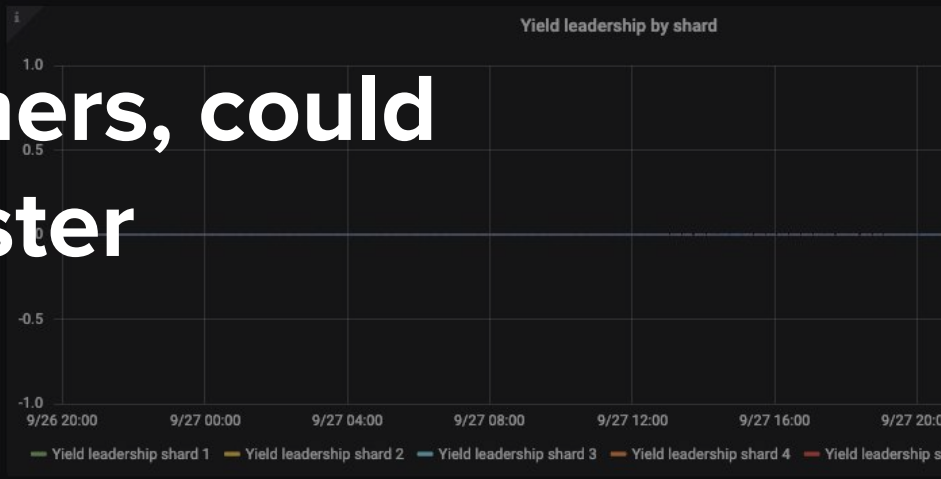
Source cluster

Dest cluster




# Dev: - Hey @platformers, could you deploy me a cluster


DEPLOYMENT ?
EASY PEASY LEMON SQUEEZY
imgflip.com

# Tools

Terraform + ANSIBLE

```
instances type          = "m5.xlarge"
instances count r1      = "3"
instances count r2      = "0"
key name                = "sshkey"
dns domain              = "internal"
zookeeper project       = "zookeeper"
environment             = "dev"
data volume size        = 1024
data volume type        = "st1"
nbs data disks          = 3
project name            = "clickhouse"
clickhouse backup       = "v2"
encrypt ebs             = true
aws_region              = "us-east-1"
```

# Takeaways

CONTENTSQUARE

Automate as soon as possible

Use common metrics to define custom alerts

Empower developers

Test your tools !

Bench, don’t guess

Take your time to understand CH before starting devs

# We’re hiring!

CONTENTSQUARE

## Q&A TIME !