riskified

Riskified's Lessons Learned Migrating to CockroachDB for Greater Scale

Harel Safra, Data Platform Engineering, Team Lead
Yoav Shemesh, Senior Data Platform Engineer

RoachFest24

About us

Harel Safra
Data Platform Engineering
Team Lead,
Riskified

harel.safra@riskified.com

Yoav Shemesh
Senior Data
Platform Engineer,
Riskified

yoav.shemesh@riskified.com

Agenda

Riskified by the Numbers

750+

Global team, nearly 50% in **engineering & analytics**

185

Countries where we operate

50+

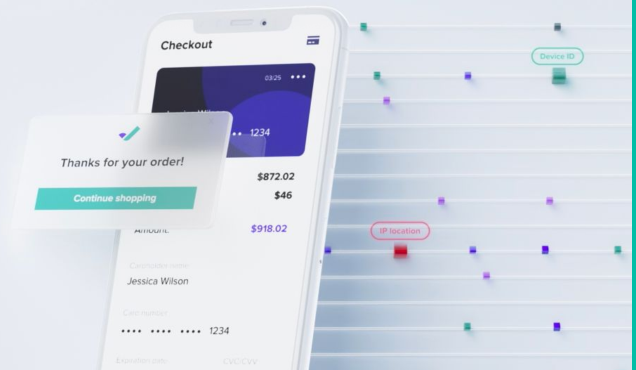
Publicly-traded companies among our clients

$105B+

Online volume (GMV) reviewed in 2022

99%+

Client retention for FY'2022



wayfair

PRADA

GoPro

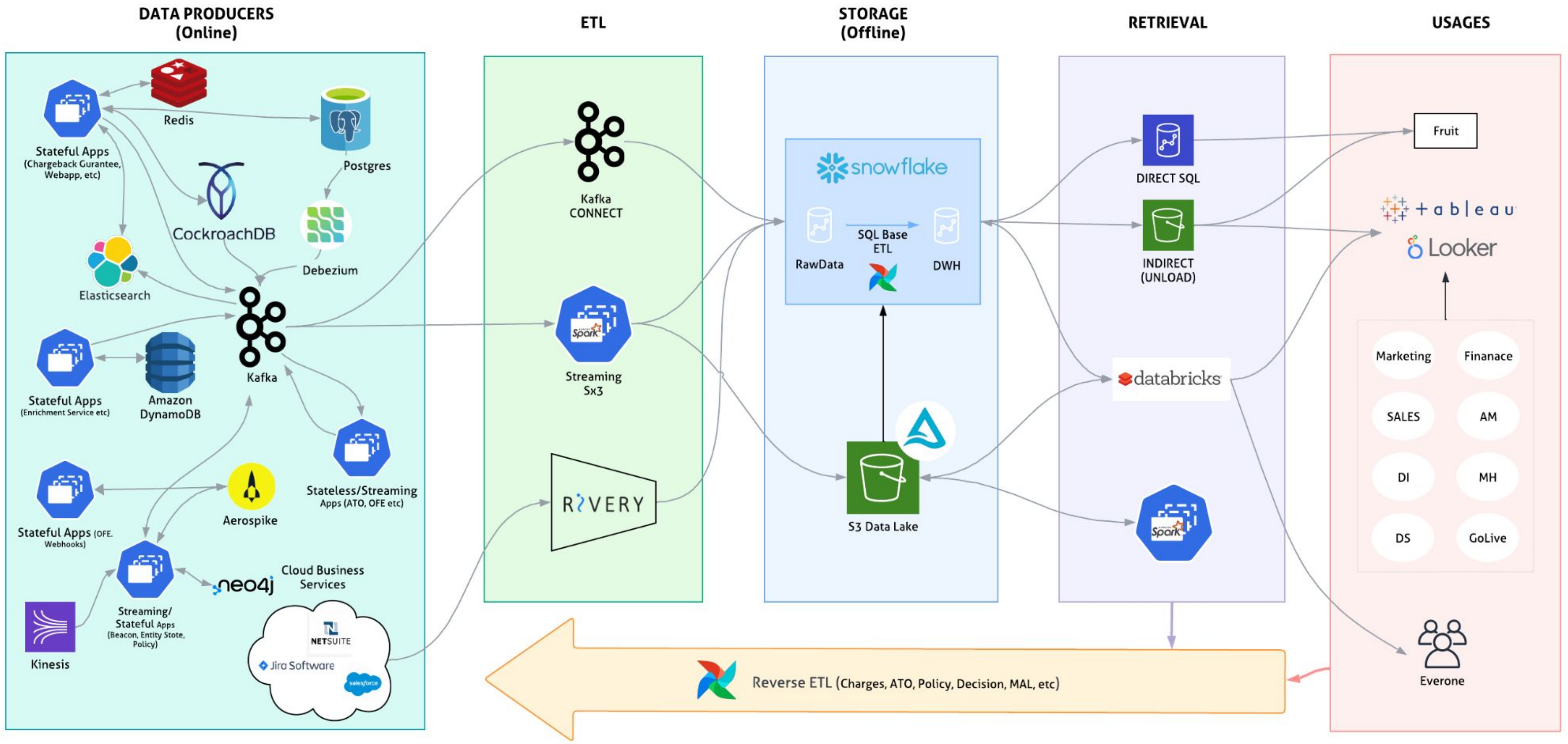
FINISH LINE

lastminute.com

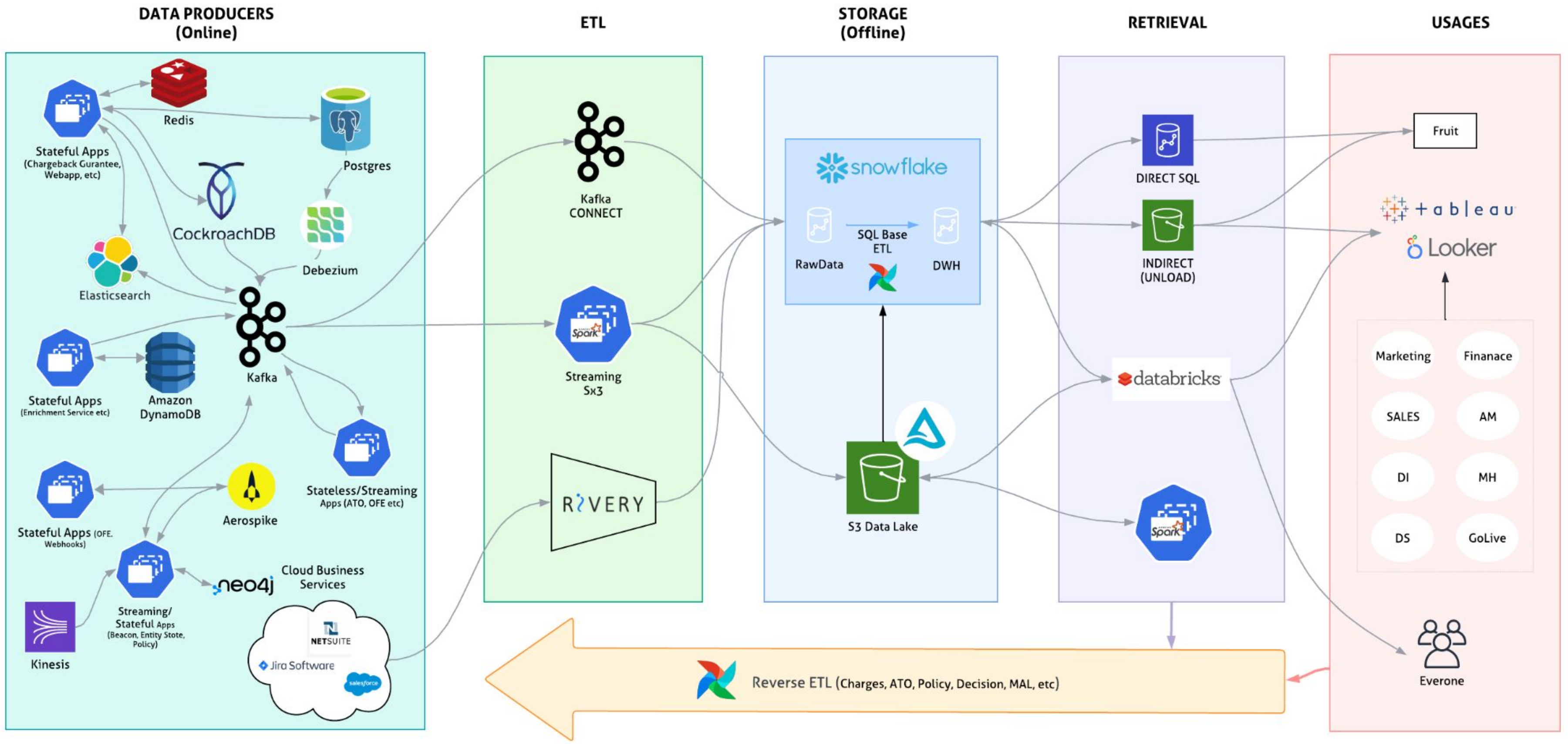
REVOLVE

CANADA GOOSE

Riskified's Data Platform

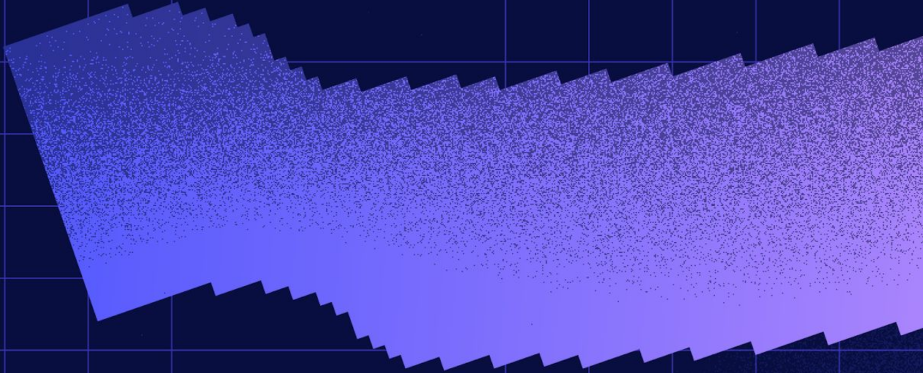


Riskified's Data Platform



SO, WHY DID WE LEAVE POSTGRES?

Issues encountered



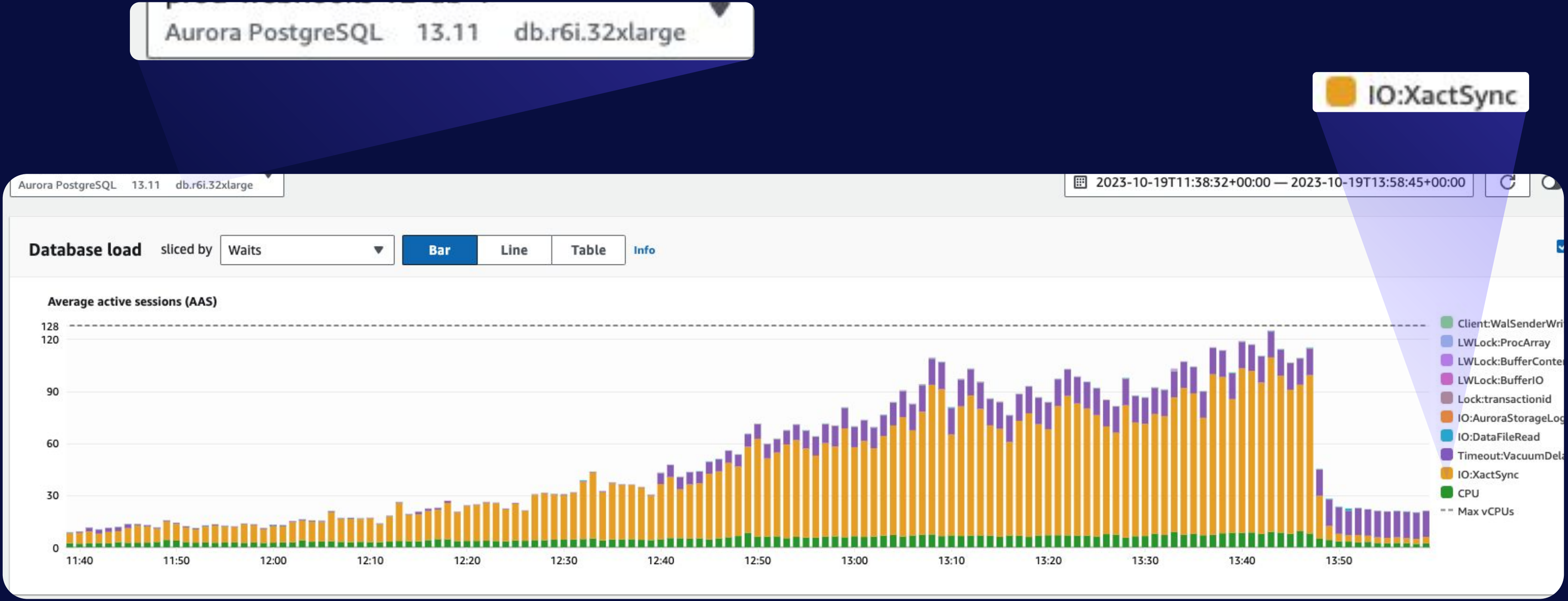
Aurora postgres

- Single writer server
- Transaction limit
- Multiple clusters

Scaling limitation

Aurora PostgreSQL 13.11 db.r6i.32xlarge

IO:XactSync

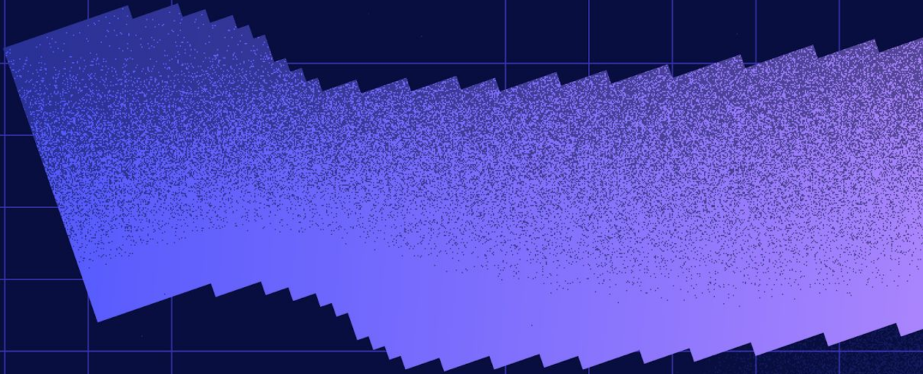

Aurora PostgreSQL 13.11 db.r6i.32xlarge

2023-10-19T11:38:32+00:00 — 2023-10-19T13:58:45+00:00

Database load sliced by Waits

CHOOSING RIGHT

POC summary



Test criteria

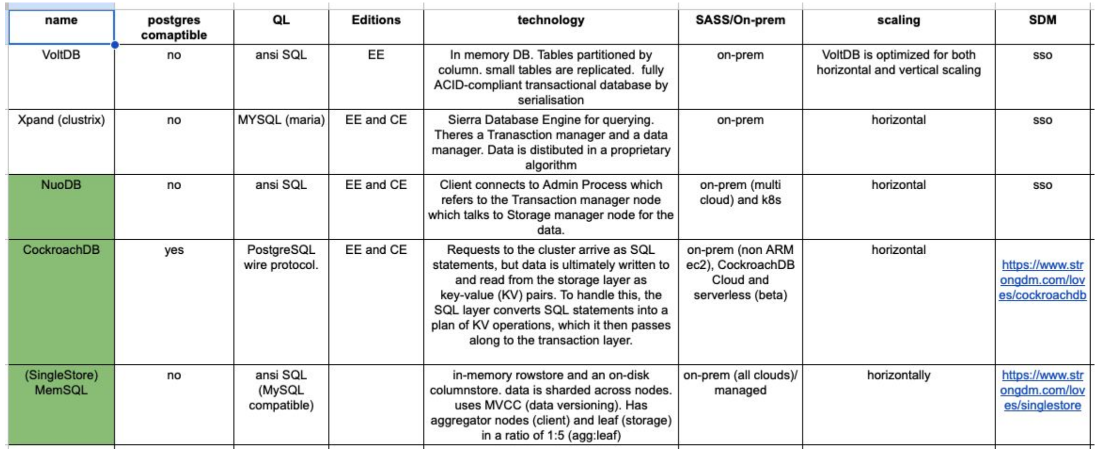
- Postgres compatible to avoid Application changes
- Scale, Resilience & Operations
  - Online upgrades, failures, backups …
- Installation & data loading
- Security
- Performance

POC: Candidates

- VoltDB
- Xpand (clustrix)
- NuoDB
- CockroachDB
- SingleStore (MemSQL)
- YugabyteDB

POC: Tested

- YugabyteDB
- CockroachDB



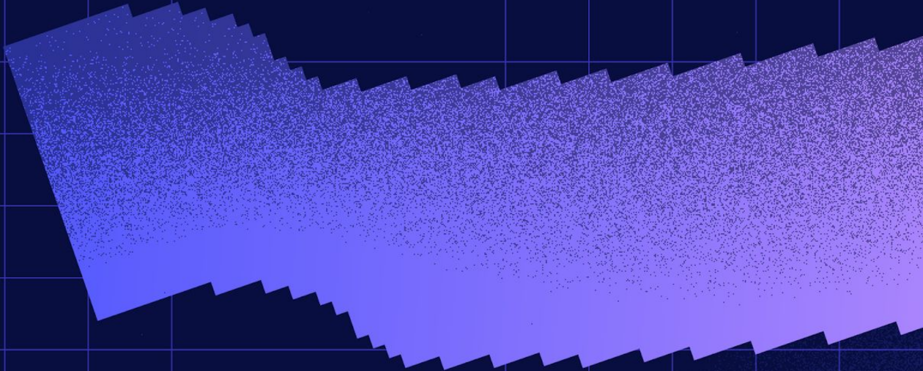
name	postgres comaptible	QL	Editions	technology	SASS/On-prem	scaling	SDM
VoltDB	no	ansi SQL	EE	In memory DB. Tables partitioned by column. small tables are replicated. fully ACID-compliant transactional database by serialisation	on-prem	VoltDB is optimized for both horizontal and vertical scaling	sso
Xpand (clustrix)	no	MYSQL (maria)	EE and CE	Sierra Database Engine for querying. Theres a Tranasction manager and a data manager. Data is distibuted in a proprietary algorithm	on-prem	horizontal	sso
NuoDB	no	ansi SQL	EE and CE	Client connects to Admin Process which refers to the Transaction manager node which talks to Storage manager node for the data.	on-prem (multi cloud) and k8s	horizontal	sso
CockroachDB	yes	PostgreSQL wire protocol.	EE and CE	Requests to the cluster arrive as SQL statements, but data is ultimately written to and read from the storage layer as key-value (KV) pairs. To handle this, the SQL layer converts SQL statements into a plan of KV operations, which it then passes along to the transaction layer.	on-prem (non ARM ec2), CockroachDB Cloud and serverless (beta)	horizontal	https://www.strongdm.com/loves/cockroachdb
(SingleStore) MemSQL	no	ansi SQL (MySQL compatible)		in-memory rowstore and an on-disk columnstore. data is sharded across nodes. uses MVCC (data versioning). Has aggregator nodes (client) and leaf (storage) in a ratio of 1:5 (agg:leaf)	on-prem (all clouds)/ managed	horizontally	https://www.strongdm.com/loves/singlestore

Performance metrics

- **Jmeter**
  - Postgres < 1ms
  - CockroachDB ±3ms
- **Main App sample process**
  - Postgres ±60ms
  - CockroachDB ±80ms

OUR SOLUTION

Implementation details



Our list of requirements

Any Tech

- Single server auto replacement
- Replace all nodes in the cluster
- Automatic deployments
- Change management
- Rolling restarts
- Monitoring

CockroachDB

- Backups
- CDC

Single server auto replacement

+

Replace all nodes in the cluster

Automatic deployments

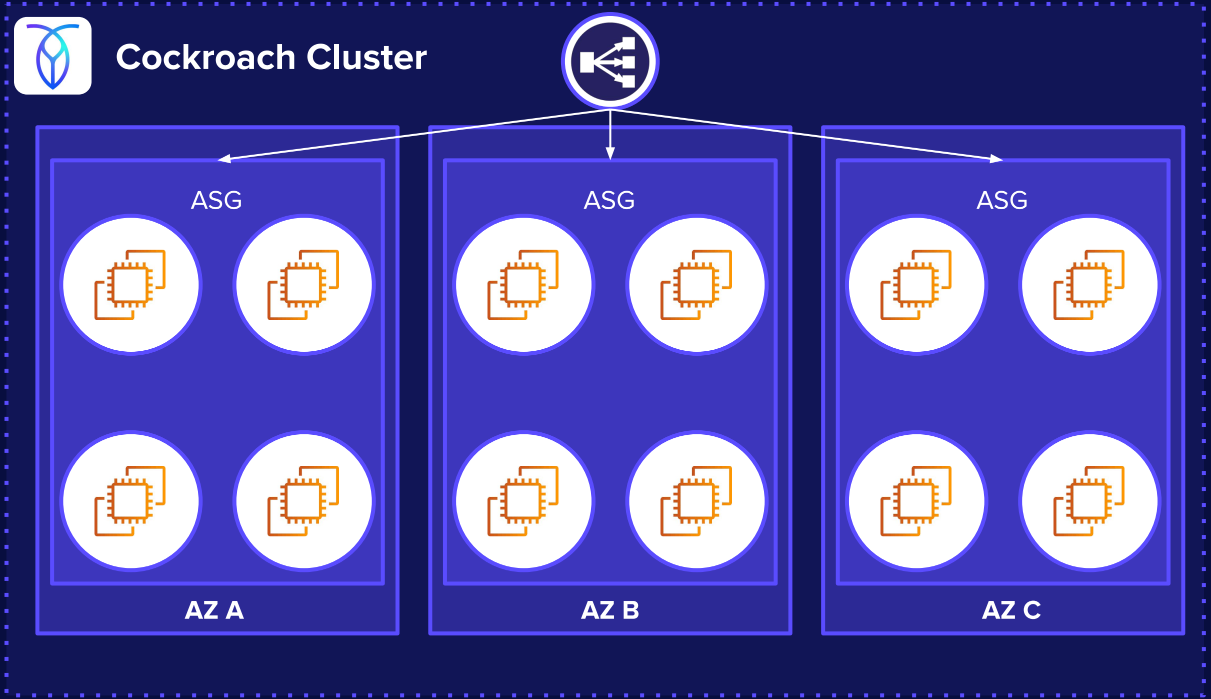
+

Change management

Rolling restarts

Monitoring & Backup

Easy management using ASG per AZ



Single server auto replacement

+

Replace all nodes in the cluster

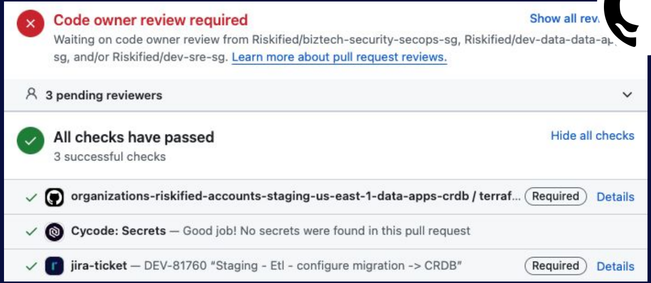
Automatic deployments

+

Change management

Rolling restarts

Monitoring & Backup



Code owner review required — Show all rev

Waiting on code owner review from Riskified/biztech-security-secops-sg, Riskified/dev-data-data-ap sg, and/or Riskified/dev-sre-sg. Learn more about pull request reviews.

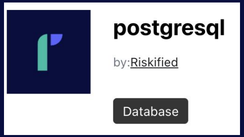
3 pending reviewers

All checks have passed — Hide all checks

3 successful checks

- ✓ organizations-riskified-accounts-staging-us-east-1-data-apps-crdb / terraf... Required Details
- ✓ Cycode: Secrets — Good job! No secrets were found in this pull request
- ✓ jira-ticket — DEV-81760 "Staging - Etl - configure migration -> CRDB" Required Details

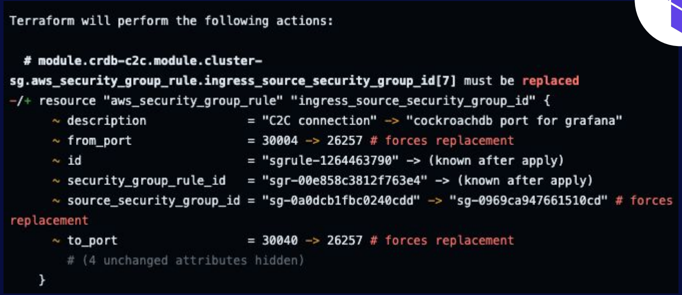
GitHub Actions for trigger plan and apply on PR and merge



postgresql

by:Riskified

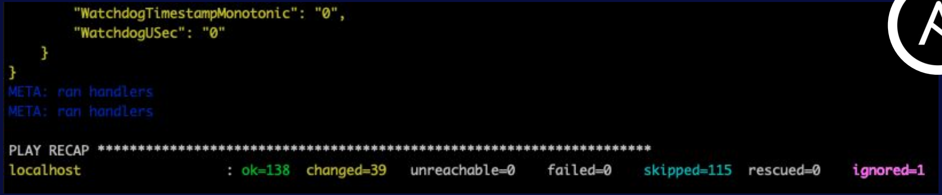
Database



```
Terraform will perform the following actions:

  # module.crdb-c2c.module.cluster-
sg.aws_security_group_rule.ingress_source_security_group_id[7] must be replaced
-/+ resource "aws_security_group_rule" "ingress_source_security_group_id" {
      ~ description              = "C2C connection" -> "cockroachdb port for grafana"
      ~ from_port                = 30004 -> 26257 # forces replacement
      ~ id                       = "sgrule-1264463790" -> (known after apply)
      ~ security_group_rule_id   = "sgr-00e858c3812f763e4" -> (known after apply)
      ~ source_security_group_id = "sg-0a0dcb1fbc0240cdd" -> "sg-0969ca947661510cd" # forces
replacement
      ~ to_port                  = 30040 -> 26257 # forces replacement
        # (4 unchanged attributes hidden)
    }
```

Terraform for managing AWS components



```
            "WatchdogTimestampMonotonic": "0",
            "WatchdogUSec": "0"
        }
}
META: ran handlers
META: ran handlers

PLAY RECAP ***************************************************************************
localhost                  : ok=138  changed=39   unreachable=0    failed=0    skipped=115   rescued=0    ignored=1
```

Ansible for installing the software

Single server auto replacement

+

Replace all nodes in the cluster

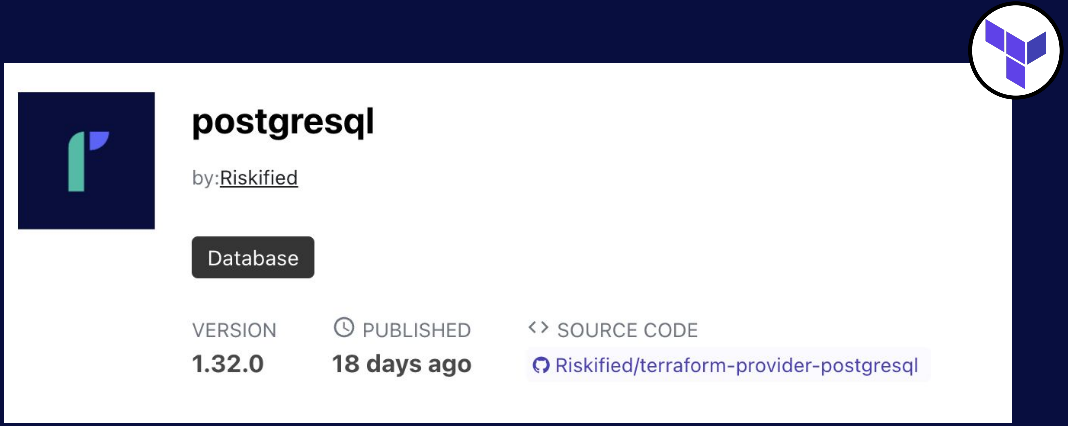
Automatic deployments

+

Change management

Rolling restarts

Monitoring & Backup



postgresql

by:Riskified

Database

VERSION
1.32.0

PUBLISHED
18 days ago

SOURCE CODE
Riskified/terraform-provider-postgresql

Feel free to use it :)

Single server auto replacement

+

Replace all nodes in the cluster

Automatic deployments

+

Change management

Rolling restarts

Monitoring & Backup

Ruby from Airflow

Single server auto replacement

+

Replace all nodes in the cluster

Automatic deployments

+

Change management

Rolling restarts

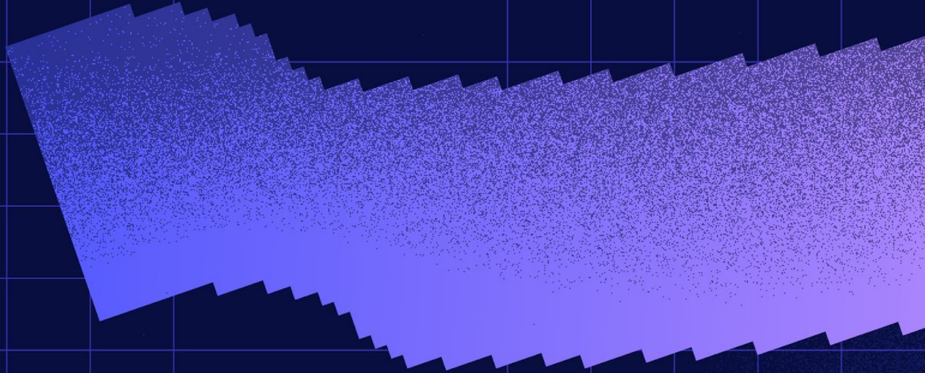
Monitoring & Backup

- Prometheus, Grafana & built in console
- Using the built-in exporters

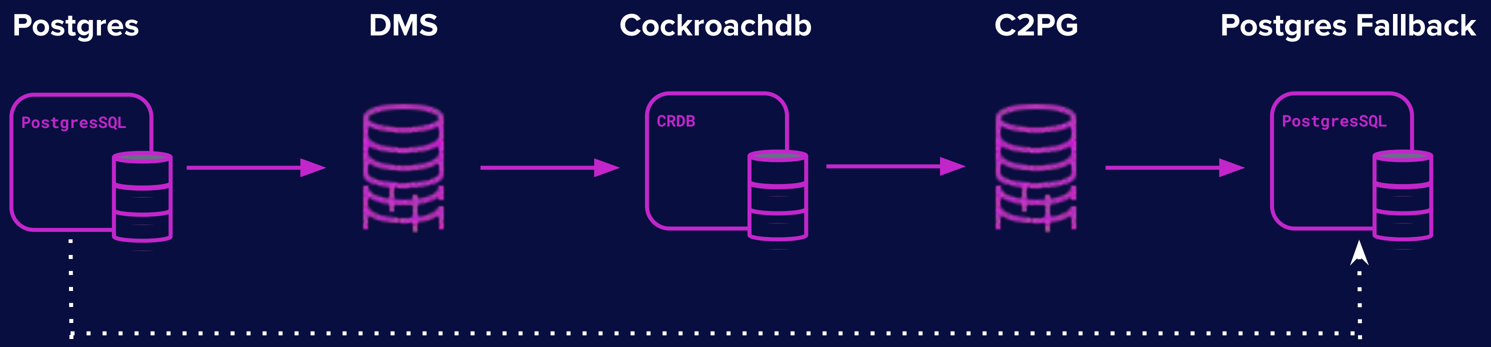
- Scheduled backups to S3
- Local & DR regions

MIGRATION PROCESS

Road to success



Migrations with AWS DMS and Replicator (belt & suspenders)



1. Create schemas, users
2. Full load + CDC with AWS Data Migration Service
3. Create a copy from source
4. CDC using C2PG (partitioned tables need custom script)

Migration tips & guidelines - Docs

- Use Cockroach cloud migration tool
- Replace IDENTITY columns with **unordered_unique_rowid** on busy tables
- Read the documentation!

Migration tips & guidelines

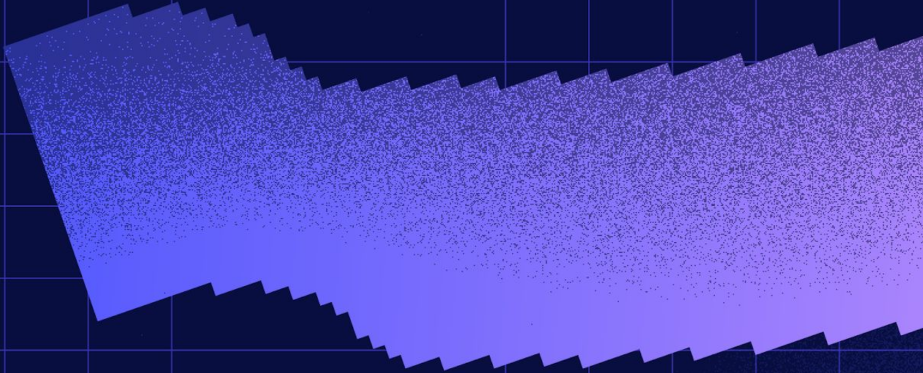
- Use port 5432 in addition to 26257
- Reuse existing DNS
- Read committed transaction isolation
- Wrap schema migrations in explicit transactions

Changefeeds

- Replication slots $\longrightarrow$ Changefeeds
- Message format is a bit different
- Don't need a Kafka Connect cluster
- We use Terraform automation
    - Using SQL migration provider
    - Creating changefeeds, Kafka and registry connections

LESSONS LEARNED
(SO YOU DON'T HAVE TO)

Clients & SDKs



Ruby

- Ruby on Rails works as is with the Postgres adaptor
- Use the SSL flag for production
- Rake - needs adapting
    - db: migrate
    - db: test:prepare - adapted due to db:migrate

Node & ORM

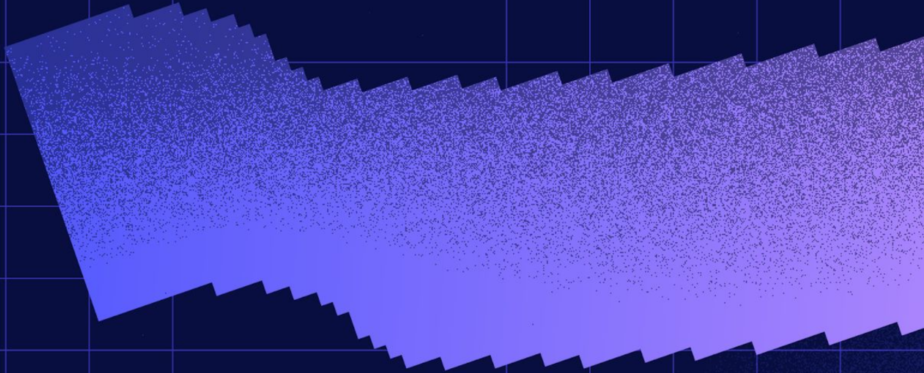
- Define SSL at the connection
  - We use global certificates
  - Postgres doesn't need it

Scala & JDBC

JUST WORKS!

TLDR;

Let’s wrap it up



Wrap Up

- Scalability issues were solved by CockroachDB
- AWS Auto Scaling Groups to manage the cluster
- Migration required minor application changes
- Replaced Debezium with Changefeeds

ANY QUESTIONS?

riskified tech;

Thank you for your time!

Harel Safra
harel.safra@riskified.com
www.linkedin.com/in/harelsafra

Yoav Shemesh
yoav.shemesh@riskified.com
www.linkedin.com/in/yoavshemesh