



COMMUNITY DAY

AOTEAROA



COMMUNITY DAY



A path to better performance: Leveraging AWS services for DB indexing & archiving

Ali Haydar | Johnny Shao | 06/09/2023



Fergus

Job software for tradies



COMMUNITY DAY

AOTEAROA

Ali Haydar

Platform Lead



[/in/ahaydar/](#)



[@Alee-haydar](#)

Johnny Shao

Platform Engineer



[/in/johnnysea/](#)



Fergus

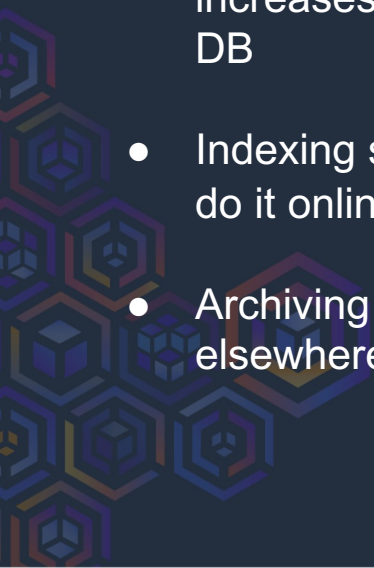


COMMUNITY DAY

AOTEAROA

Overview

- Maintain performance as data volume increases - let's index and archive our DB
- Indexing should be simple - can we do it online?
- Archiving - move that old data elsewhere
- Approach Selection
- Challenges
- Implementation





Customer Driven

We operate with acute customer focus to make their lives better

Dashboard loading took ~ 30s for some customers

Started doing some analysis

- Identifying pages that can be improved
- Queries originating from these page & targeted tables

RDS Performance Insights + Sentry



Indexing

- Indexing accelerates data retrieval, optimizing query performance and enabling swift access to specific information

Archiving

- Archiving simplifies data management, keeping active databases fast and efficient while safeguarding important historical records

Indexing





COMMUNITY DAY

AOTEAROA

Indexing - Considering Options



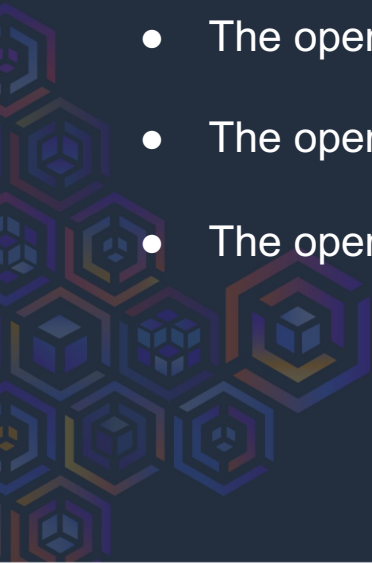


COMMUNITY DAY

AOTEAROA

Indexing Option 1 - Online MySQL Indexing

- MySQL allows online DDL changes after V5.6
- The operation would still consume limited DB resources
- The operation would cause a replica lag on the read replicas
- The operation cannot be paused or rollback



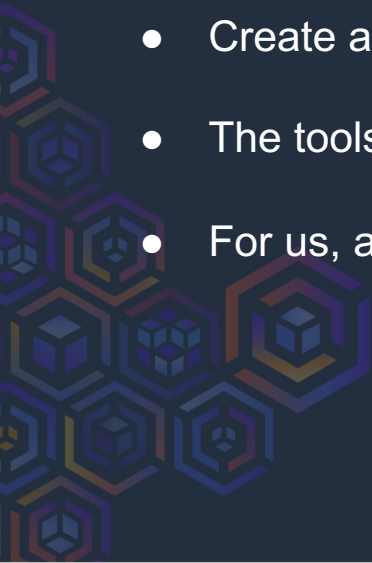


COMMUNITY DAY

AOTEAROA

Indexing Option 2 - External Tool

- Explored 2 tools: Percona (pt-online-schema-change) or Github (Ghost)
- Create a copy of the table, add the index, then copy the data over
- The tools either don't allow or have limitations with Foreign keys
- For us, adding indices was an infrequent operation

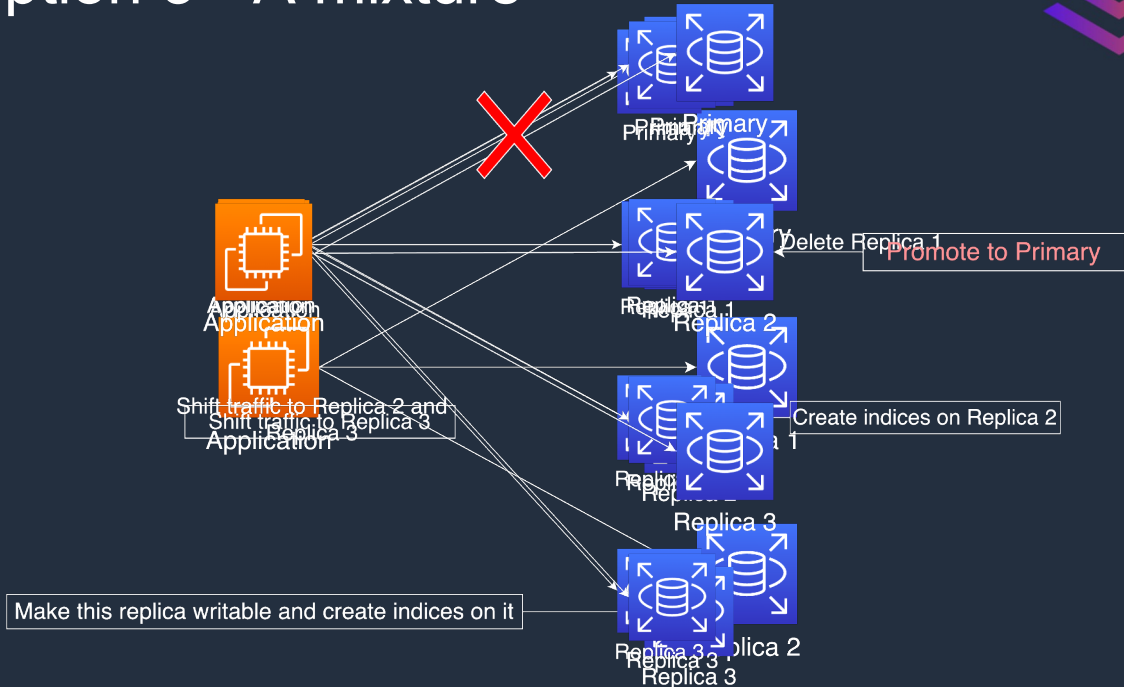




COMMUNITY DAY

AOTEAROA

Indexing Option 3 - A mixture

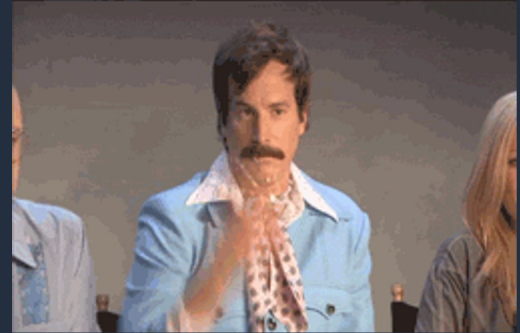
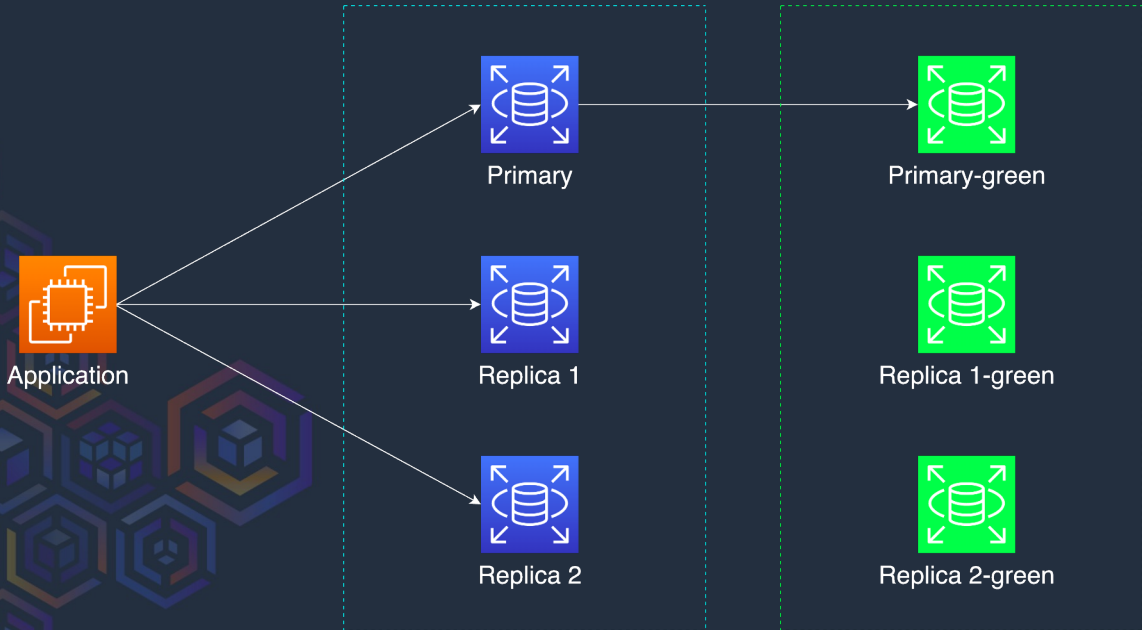




COMMUNITY DAY

AOTEAROA

Indexing - WAIT - RDS BLUE/GREEN Deployment



Blue / Green Deployment limitations

- The `event_scheduler` parameter must be disabled on the green env
- Blue/green deployments aren't supported for the following features:
 - Amazon RDS Proxy
 - Cascading read replicas
 - Cross-Region read replicas
 - AWS CloudFormation
 - Multi-AZ DB cluster deployments
- Cannot change encryption config, upgrade engine version on the blue env, and both env need to be in the same AWS Account

Archiving



Archiving

- Log related tables that are frequently accessed - but customers only request new data
- How about we archive old data?



COMMUNITY DAY

AOTEAROA

Archiving - Considering Options



Archiving Option 1 - A new table

Move data into a new table in DB

- 2 ways:
 - Create a script to copy data in batches
 - Use a tool such as Percona

Challenge:

- Possible impact on CPU / DB resources





COMMUNITY DAY

AOTEAROA

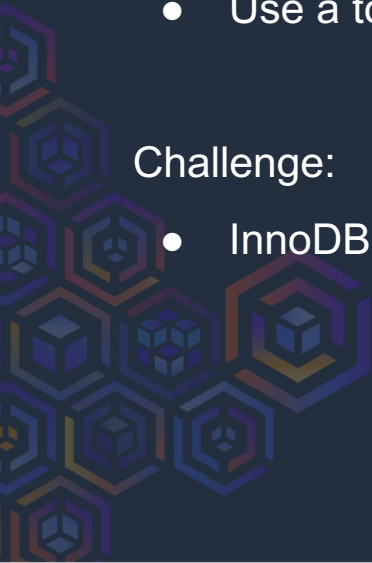
Archiving Option 2 - Partitioning

MySQL table partitioning

- Use a tool such as Percona pt-online-schema-change

Challenge:

- InnoDB tables with foreign key constraints cannot be partitioned





COMMUNITY DAY

AOTEAROA

Archiving

Option 3?



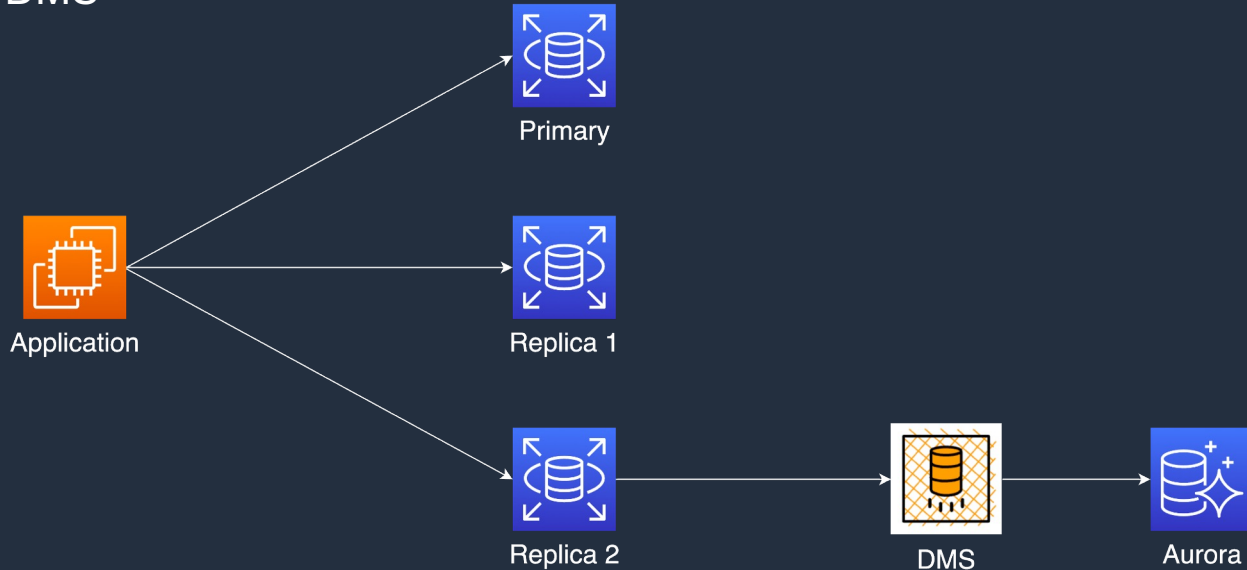


COMMUNITY DAY

AOTEAROA

Archiving Option 3 - A new storage

- Aurora + DMS



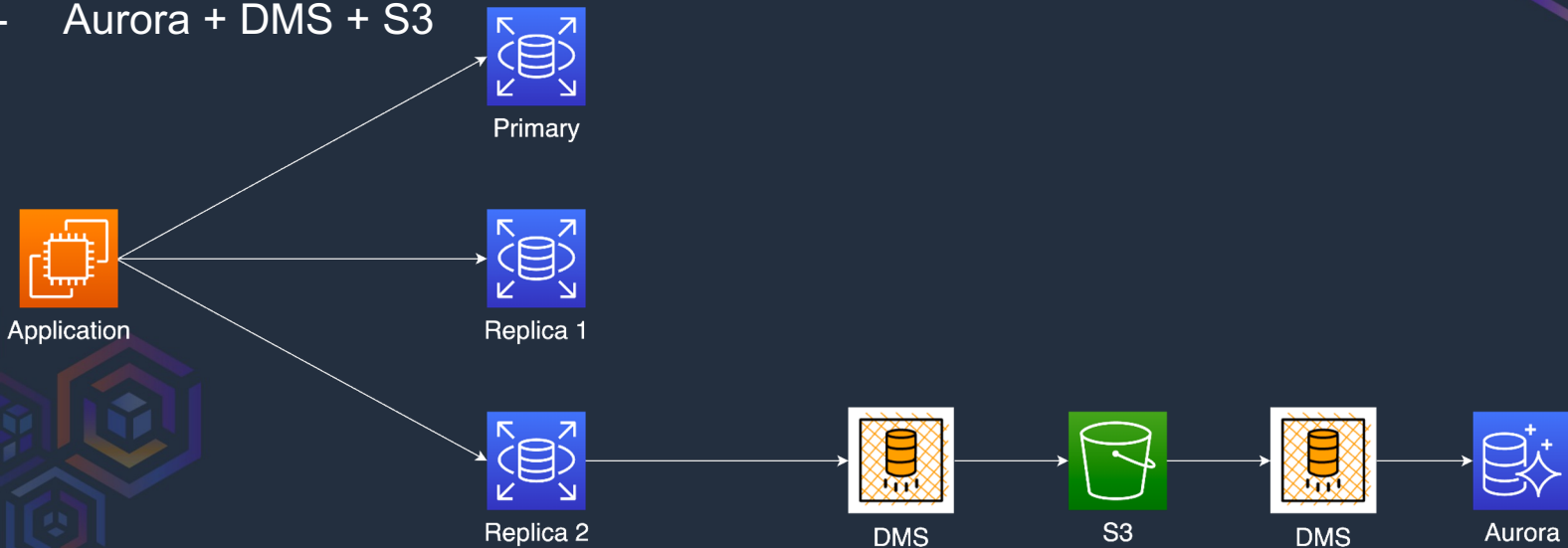


COMMUNITY DAY

AOTEAROA

Archiving Option 3 - A new storage

- Aurora + DMS + S3

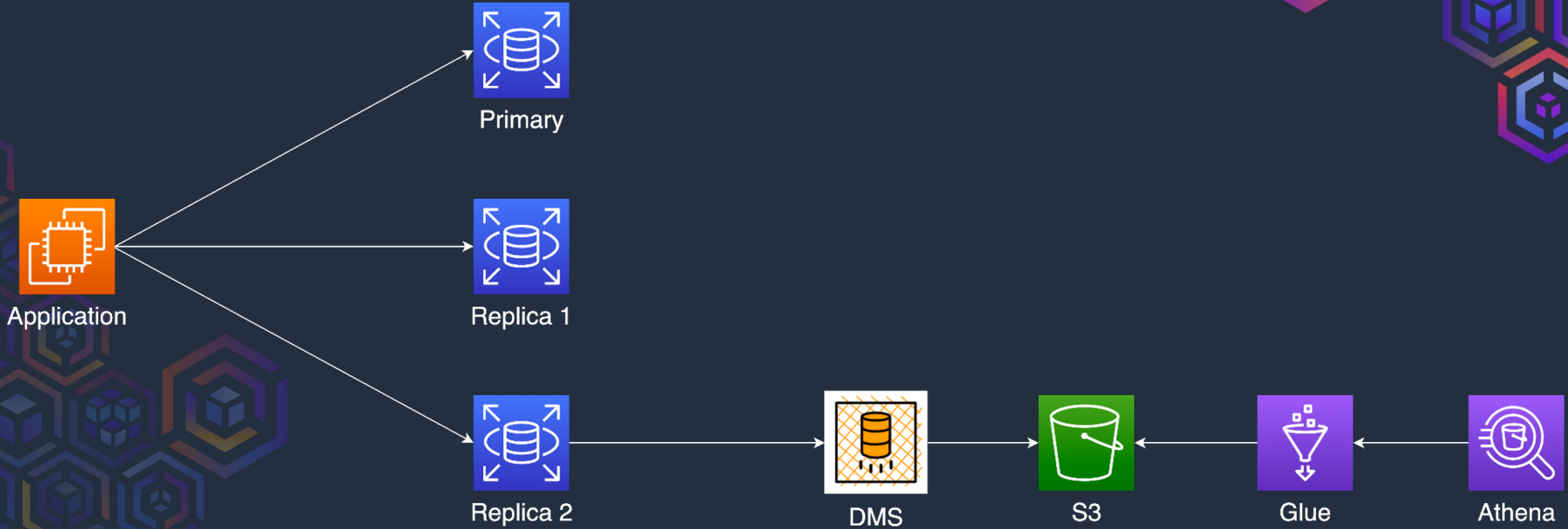




COMMUNITY DAY

AOTEAROA

Migrate data to S3 using DMS

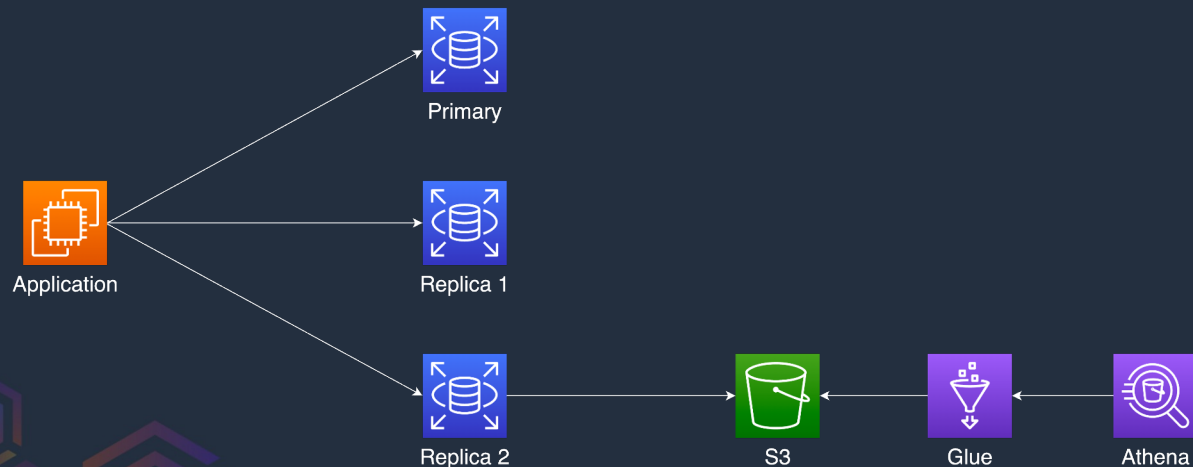




COMMUNITY DAY

AOTEAROA

Exporting RDS DB snapshot data to Amazon S3



Challenge:

- Define a table from multiple snapshots
- Filter out duplicate data when populating the table

DMS Approach Challenges

- DMS requires monitoring and alarming
- Changing the source data store means we would need to restart the DMS tasks
 - A restart would overwrite the existing data
 - A new target data store should be created
- Hard to delete/modify items in the archived data

Summary

- Indexing
 - RDS Blue/Green deployment (~2 mins of downtime)
 - For frequent operations, it might be worth it to invest into an external tool (e.g. Percona, Ghost) - these tools have their own limitations
- Archiving
 - DMS to replicate data into S3, then delete the archived data from the source database
 - Challenges with changing data source



COMMUNITY DAY

AOTEAROA

Success - Improved customer-facing dashboard by 152%.

Before

After

30.11s



3.2s

6.71s

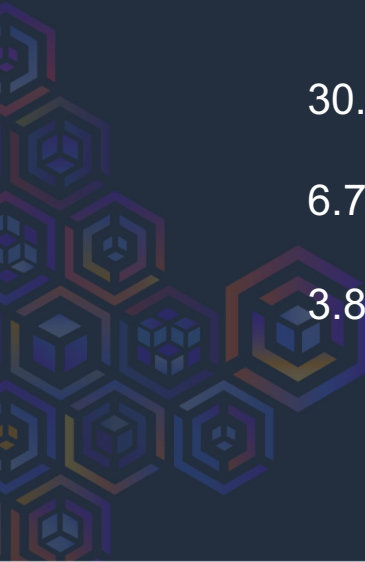
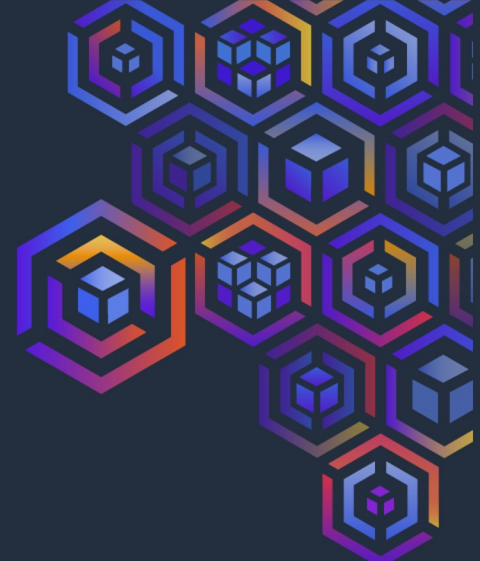


3.77s

3.82s



2.23s

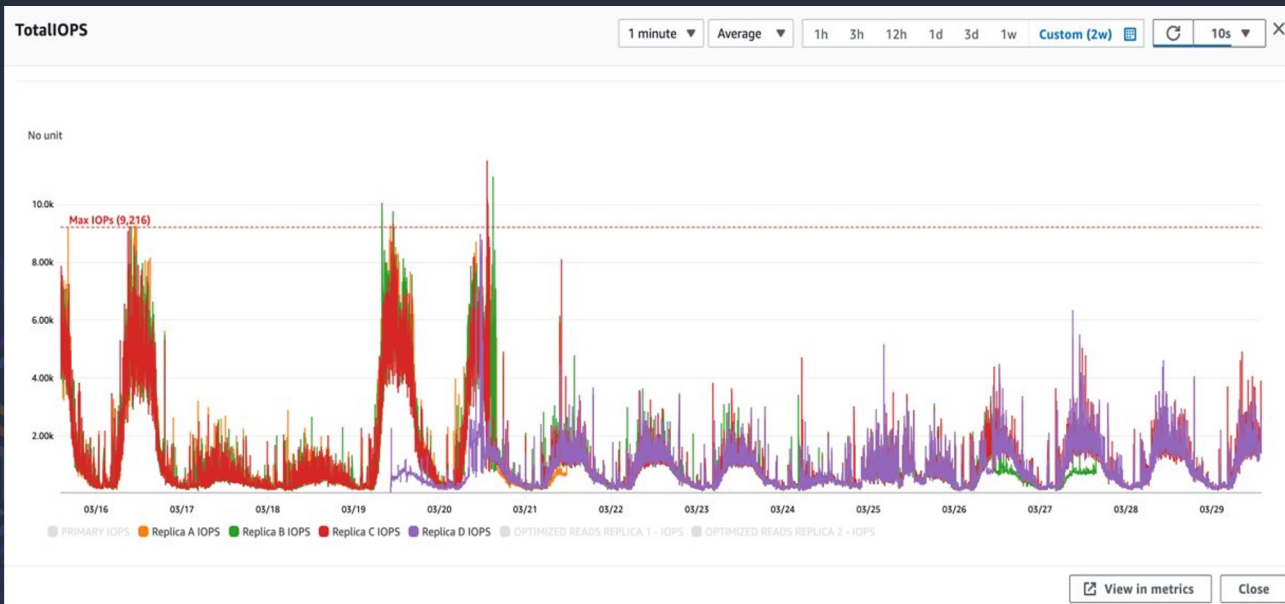




COMMUNITY DAY

AOTEAROA

Success



IOPS stands for Input/Output Operations Per Second

It measures how fast a computer can read or write data from/to a storage device.

Every storage device has a maximum IOPS it can handle.





COMMUNITY DAY

AOTEAROA

THANK
YOU





COMMUNITY DAY

AOTEAROA





COMMUNITY DAY

AOTEAROA

Fonts, sizes, colors, and layouts are all pre-built in this template.

Please do not use shadows or outlines on shape elements. Limit color use for chart graphics to grayscale plus one accent color.



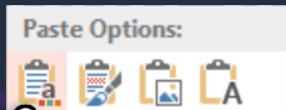


COMMUNITY DAY

AOTEAROA

Copy & Paste Content

When pasting content from another presentation please paste using “Destination Theme.”





COMMUNITY DAY

AOTEAROA

Re-sizing Assets

Always hold down shift key and drag from corner when scaling assets





COMMUNITY DAY

AOTEAROA

This is an illustrated title slide

PRESENTER | DATE





COMMUNITY DAY



This is a simple title slide

PRESENTER | DATE





COMMUNITY DAY

AOTEAROA

Title + Content





COMMUNITY DAY

AOTEAROA

Alt title + content 1





COMMUNITY DAY

AOTEAROA

Code Snippet





COMMUNITY DAY

AOTEAROA

Section Title





COMMUNITY DAY

AOTEAROA

Two columns





COMMUNITY DAY

AOTEAROA

Comparison





COMMUNITY DAY

AOTEAROA

Three column





COMMUNITY DAY

AOTEAROA

Four column with graphics





COMMUNITY DAY

AOTEAROA

Six section with graphics





COMMUNITY DAY

AOTEAROA