AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

DAT212

Simplify your Apache Cassandra workloads with Amazon Keyspaces

Michael Raney
Senior NoSQL Specialist Solutions Architect
Amazon Keyspaces

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Service Overview
- Service Architecture
- Migration

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Purpose Built Databases

	Relational	Key-value	Document	In-memory	Graph	Time-series	Ledger	Wide Column
	Referential integrity, ACID transactions, schema-on-write	High throughput, Low latency reads and writes, endless scale	Store documents and quickly access querying on any attribute	Query by key with microsecond latency	Quickly and easily create and navigate relationships between data	Collect, store, and process data sequenced by time	Complete, immutable, and verifiable history of all changes to application data	Scalable, highly available, and managed Apache Cassandra-compatible service
AWS Service(s)	Amazon Aurora, Amazon RDS	Amazon DynamoDB	Amazon DocumentDB	Amazon ElastiCache	Amazon Neptune	Amazon Timestream	Amazon QLDB	Amazon Keyspaces (for Apache Cassandra)
Common Use Cases	Lift and shift, ERP, CRM, finance	Real-time bidding, shopping cart, social, product catalog, customer preferences	Content management, personalization, mobile	Leaderboards, real-time analytics, caching	Fraud detection, social networking, recommendation engine	IoT applications, event tracking	Systems of record, supply chain, health care, registrations, financial	Build low-latency applications, leverage open source, migrate Cassandra to the cloud

© 2021, Amazon Web Services, Inc. or its Affiliates.

Amazon Keyspaces (for Apache Cassandra)

A scalable, highly available, and fully managed Apache Cassandra–compatible database service

Cassandra Query Language (CQL)



Works with your existing application code, drivers, and developer tools.

Fully managed



No nodes or servers to manage. Serverless offering.

Performance at scale



Consistent single digit millisecond performance at any scale

Highly available and secure



Built-in end-to-end encryption, available continuous backup and restore, and 99.99% availability SLA.

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

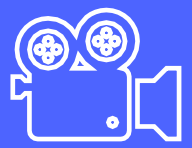
Common business applications that use Amazon Keyspaces



Social networking



Retail



Media & entertainment



Banking & finance



Gaming



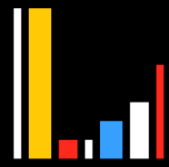
Network & IT operations

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Featured customers

Pricewaterhouse Coopers

Afterpay Limited



長橋 LONGBRIDGE





HYDRA X

© 2021, Amazon Web Services, Inc. or its Affiliates.

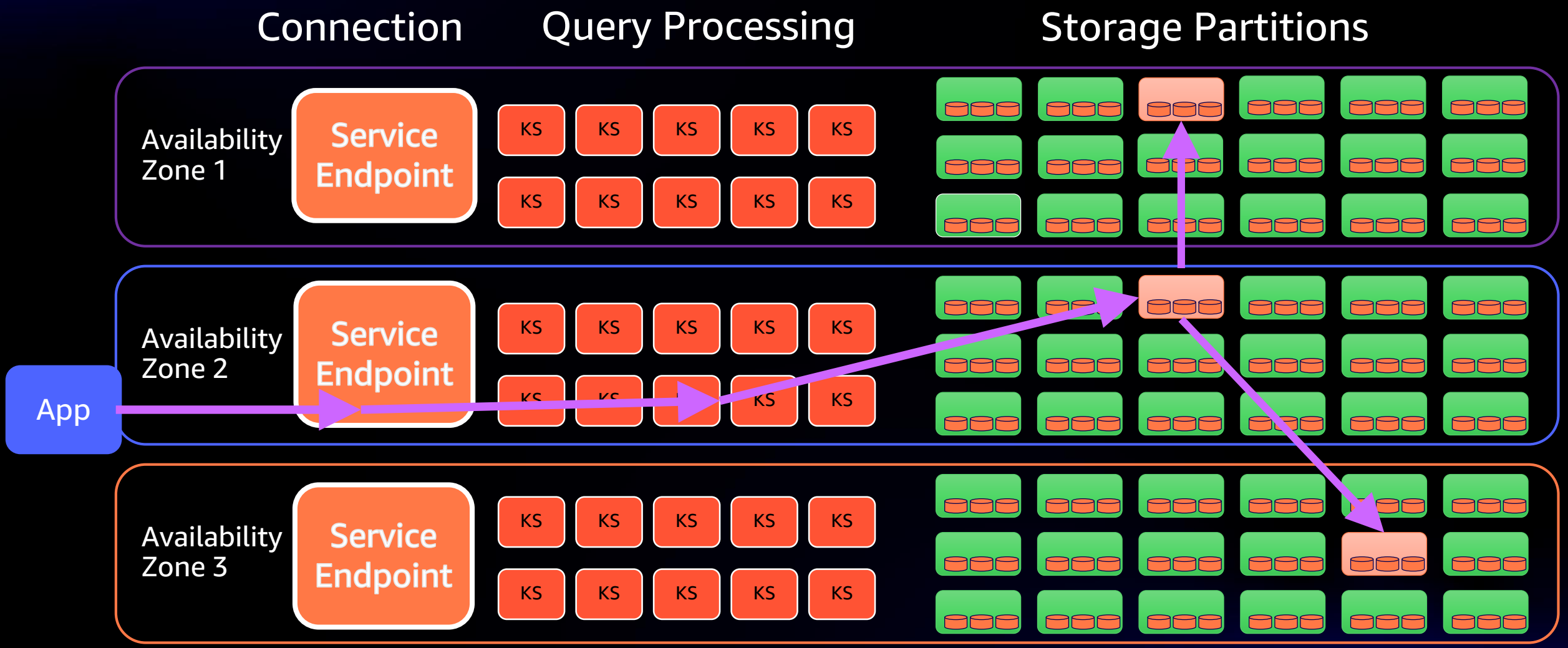
aws

Architecture

- Service architecture
- Point-in-time recovery (PITR)
- Time to Live (TTL)

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

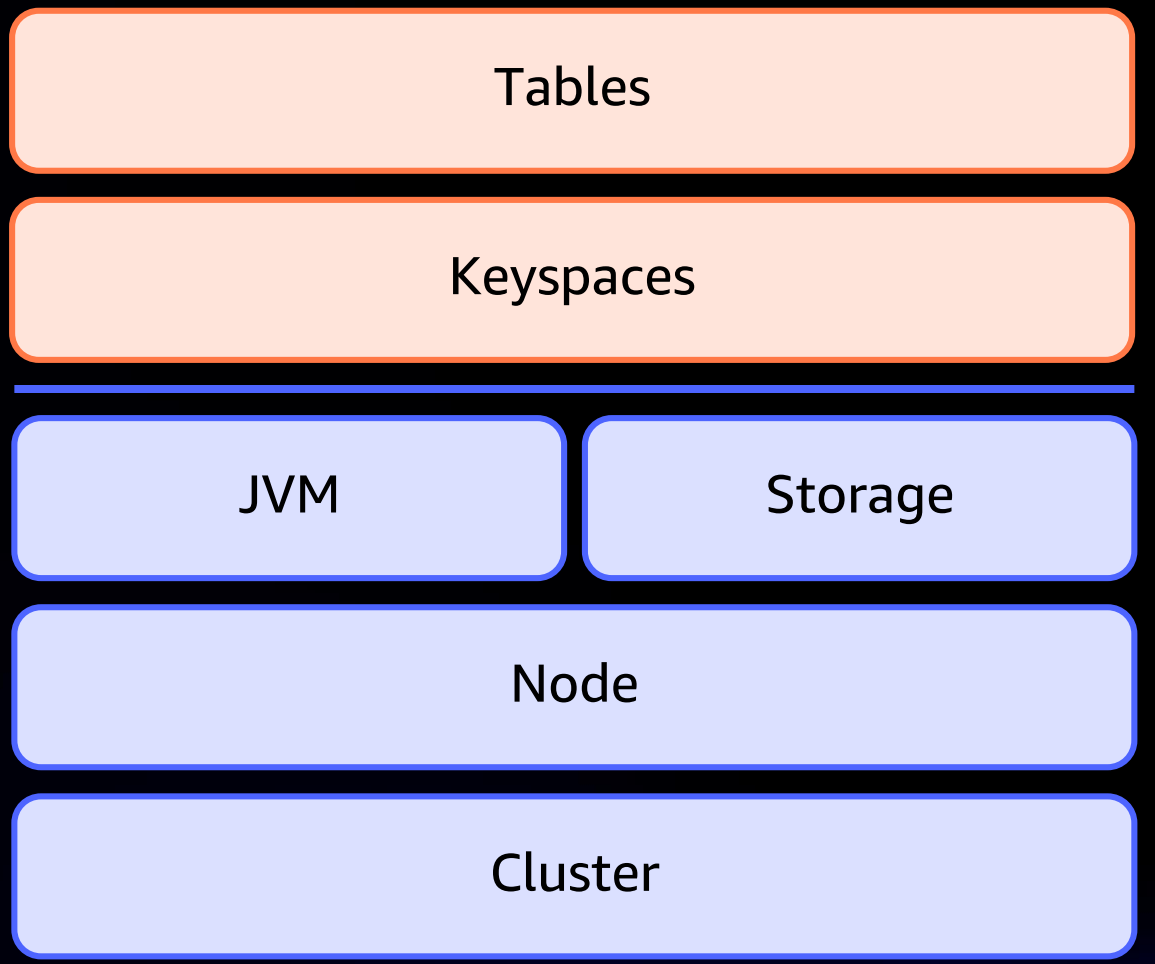
Highly available, massively scalable architecture



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

No servers to manage

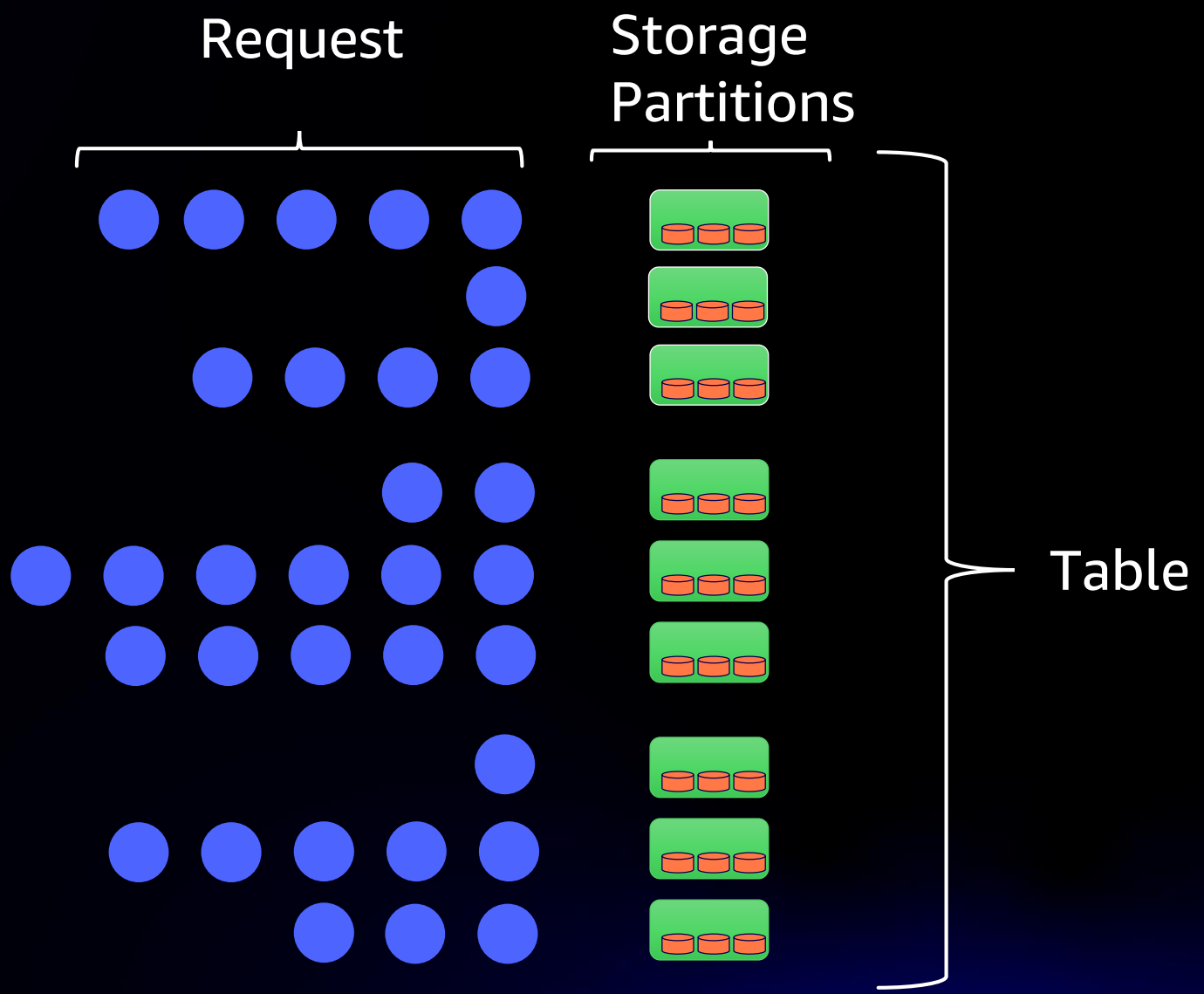
- Serverless – no managing nodes or low-level cluster settings
- No managing compaction strategies or tombstones
- No tuning Java virtual machine settings or garbage collection
- Keyspaces and tables are fundamental building blocks



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Table is made of many storage partitions

- Table made of many storage partitions
- Each partition can perform 3,000 RCU / 1,000 WCU
- Increase partitions to increase throughput and storage

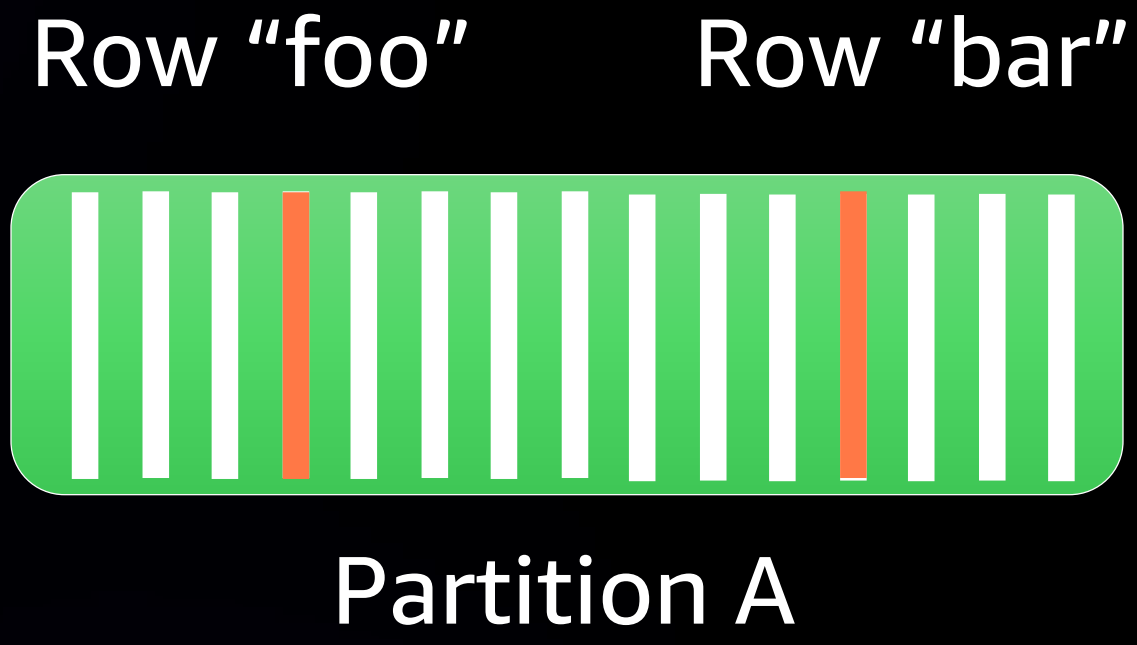


aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

High-traffic isolation

- Sustained traffic on a storage partition



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

High-traffic isolation

- Sustained traffic on a storage partition
- Split partition to increase throughput
- Split partition to increase storage



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways!

- Scalable architecture for throughput and storage
- Rebalances workload to reduce skewed access
- Support extremely wide Cassandra partitions
- Table is the basic abstraction

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

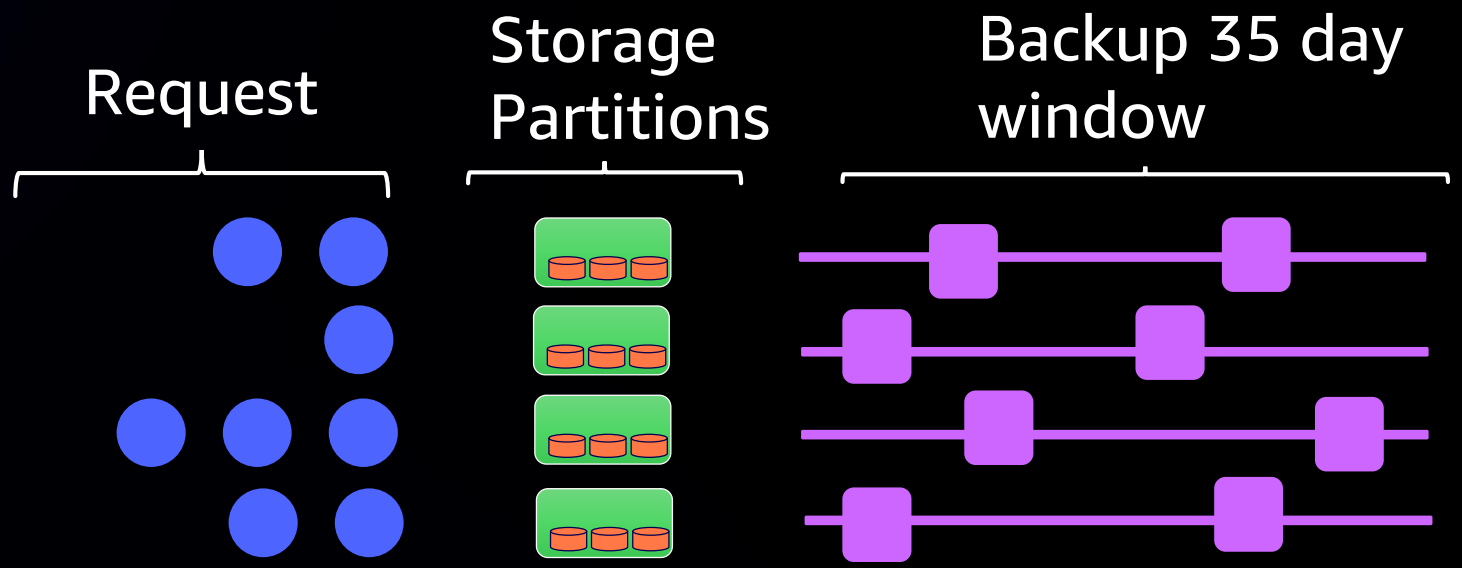
Modernizing Cassandra

- Point-in-time recovery
- Time to Live

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Point-in-time recovery (PITR)

- Table data backed up continuously
- No impact to performance

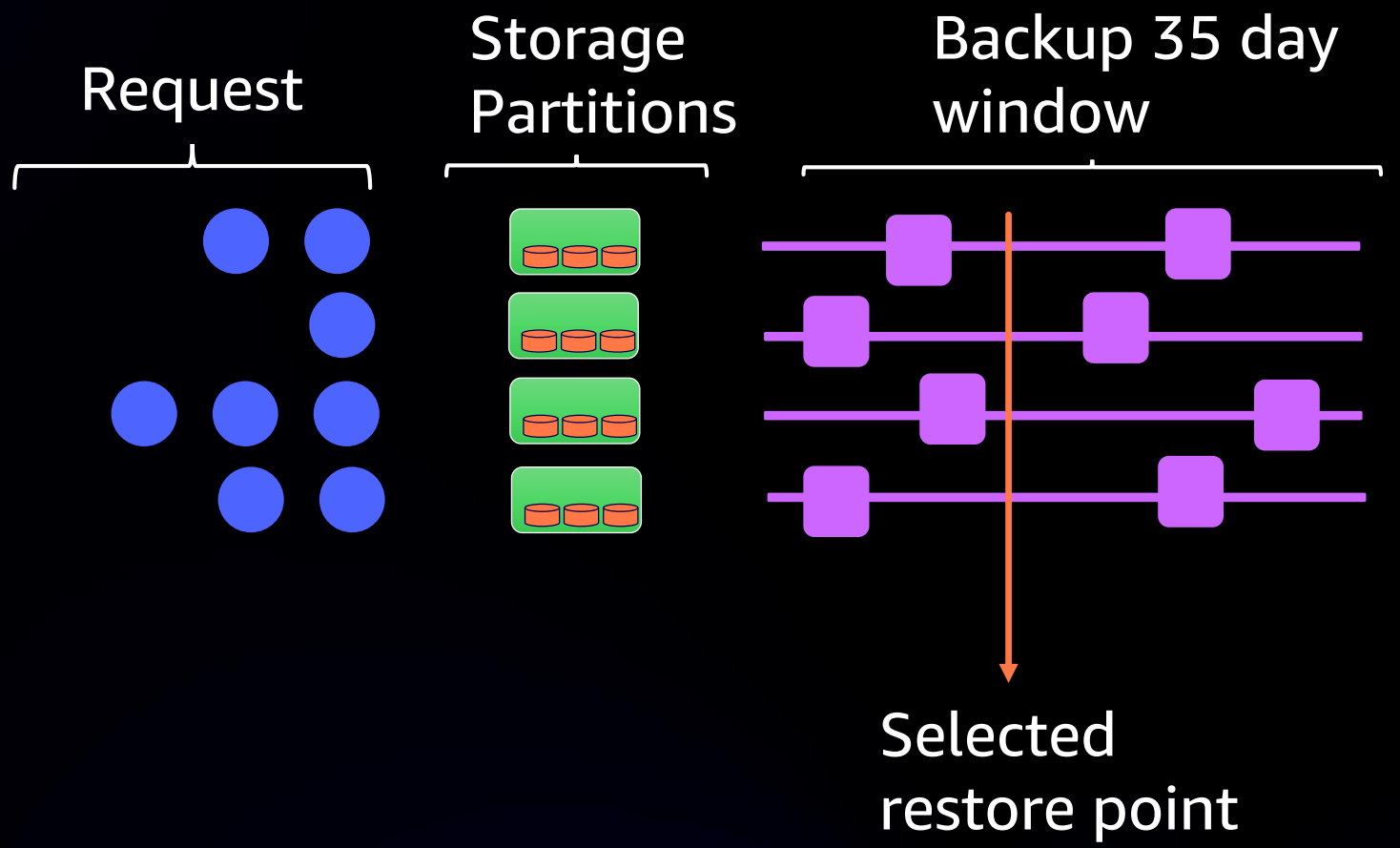


aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Point-in-time recovery

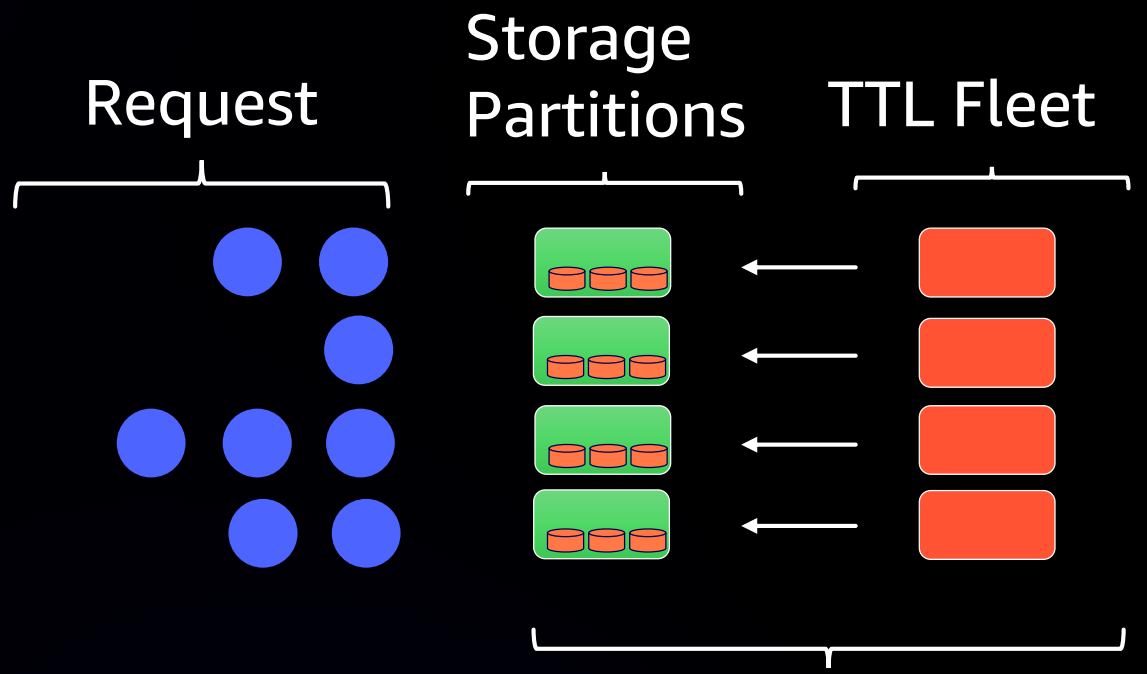
- Table data backed up continuously
- No impact to performance
- Table data can be restored to any second in the previous 35 days
- Fully consistent restore



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Time to Live (TTL)

- Insert new cell with Time to Live
- A query will not return expired results
- Records removed during background process
- Will not impact performance



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways!

- Enabling PITR and TTL will not impact performance
- No managing compaction or tombstones
- Fully managed with per table granularity

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

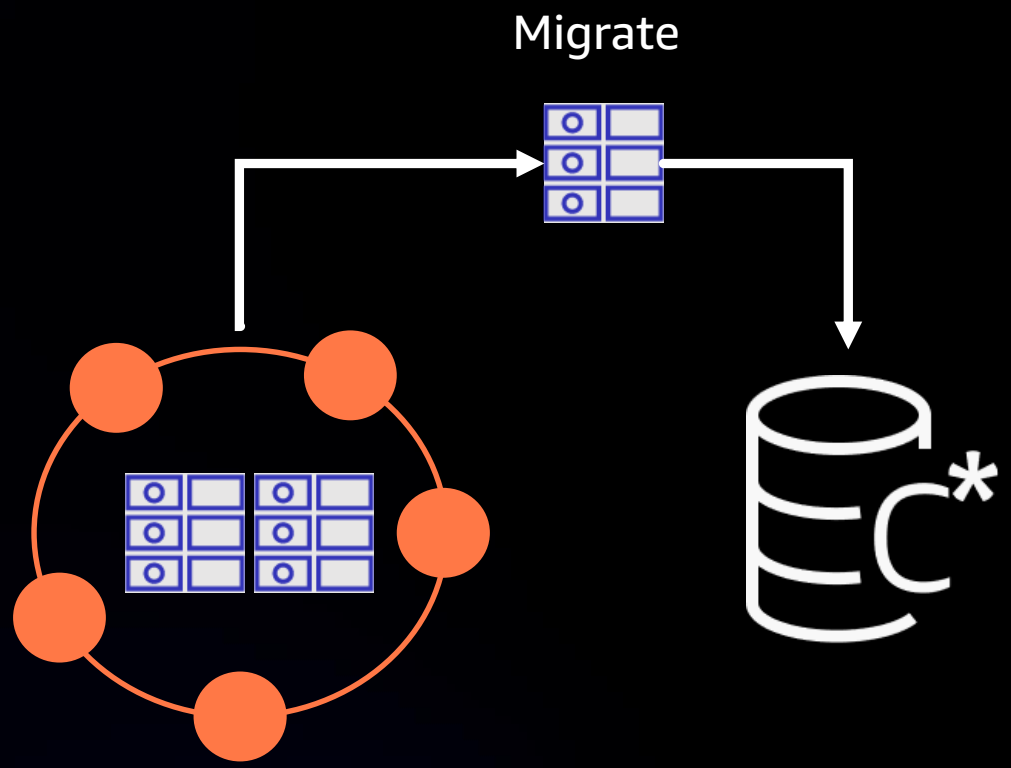
Migration from Apache Cassandra

- Table-based migration
- Dual writes

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Table-based migration

- Cassandra Cluster binds together many independent tables
- Keyspaces basic deployment model is a table
- Plan phased approach to migration
- Choose good first candidates

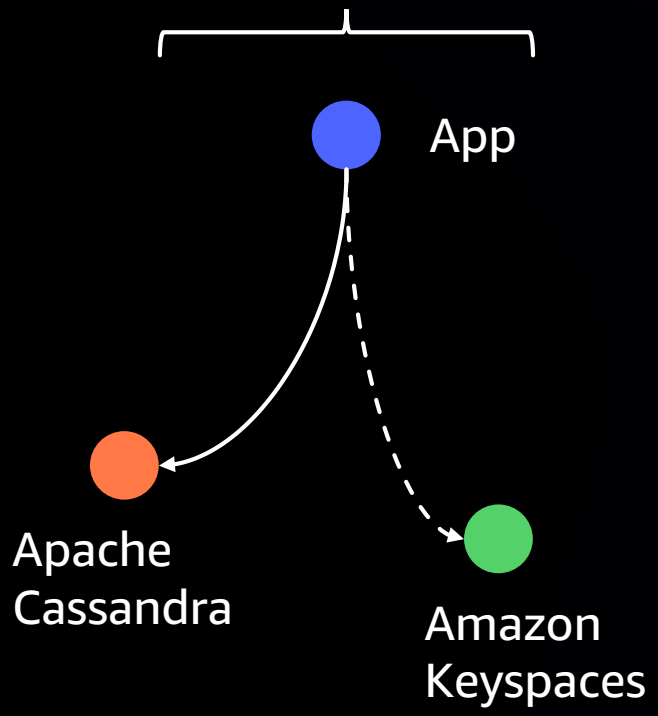


aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual writes migration

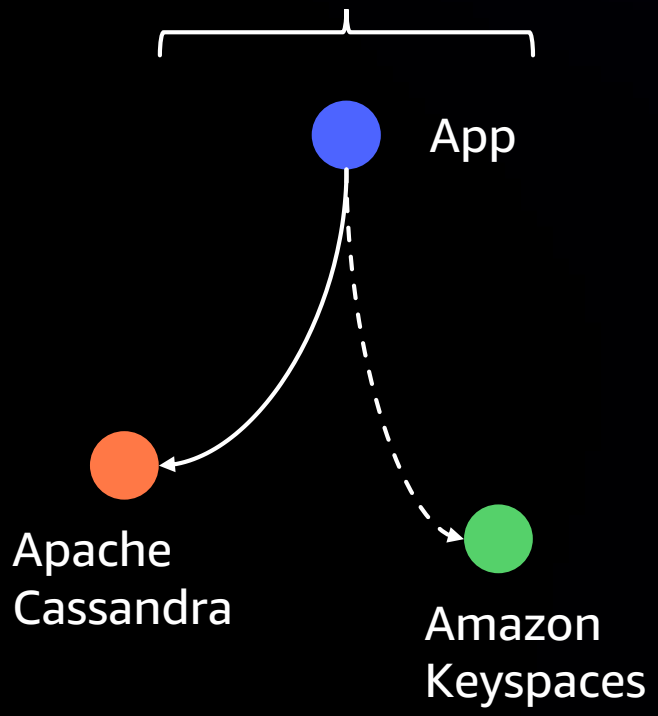
Cassandra as Primary and Keyspaces as secondary



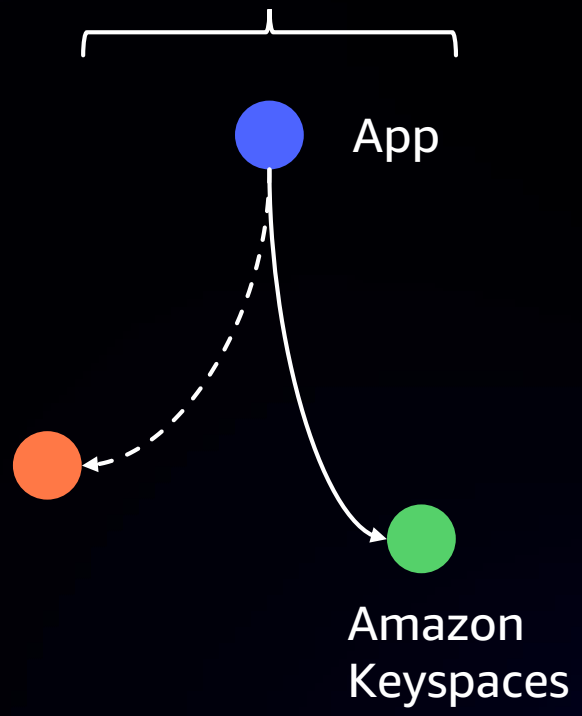
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual writes migration

Cassandra as Primary and Keyspaces as secondary



Keyspaces as Primary and Cassandra as secondary

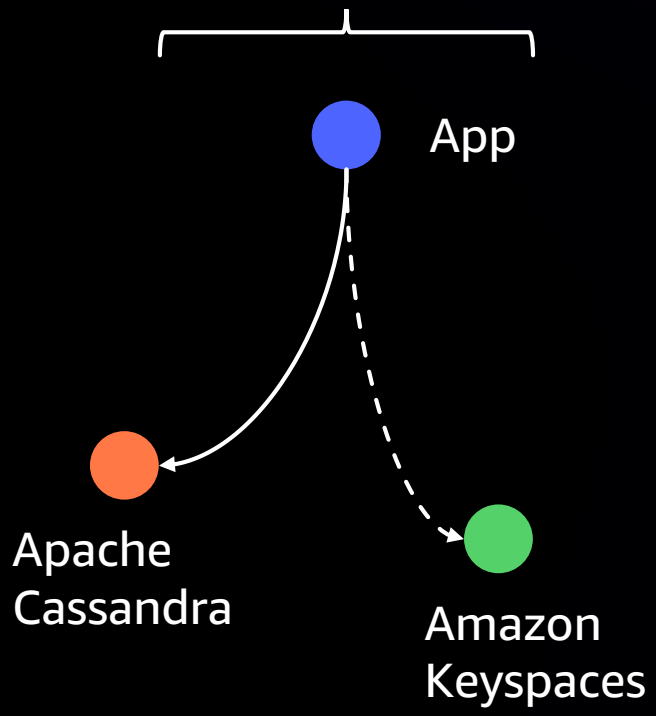


aws

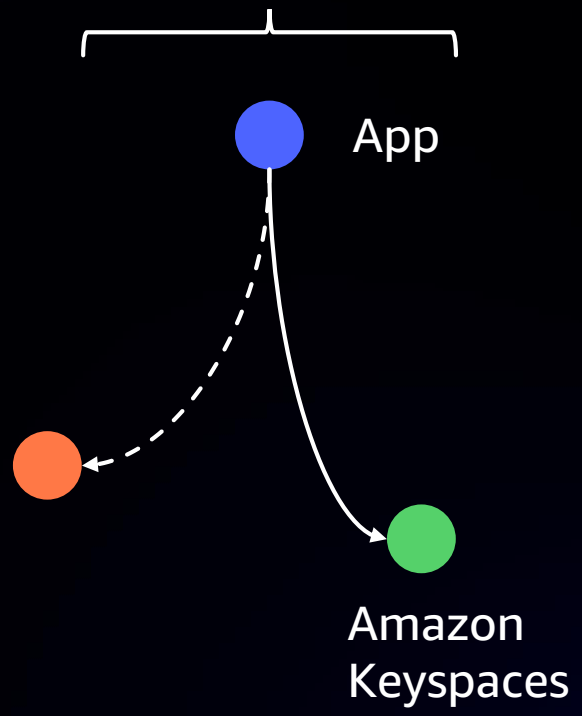
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual writes migration

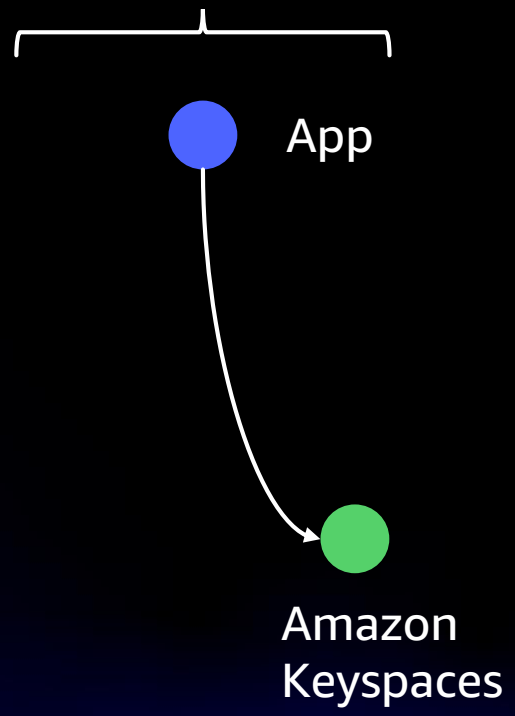
Cassandra as Primary and Keyspaces as secondary



Keyspaces as Primary and Cassandra as secondary



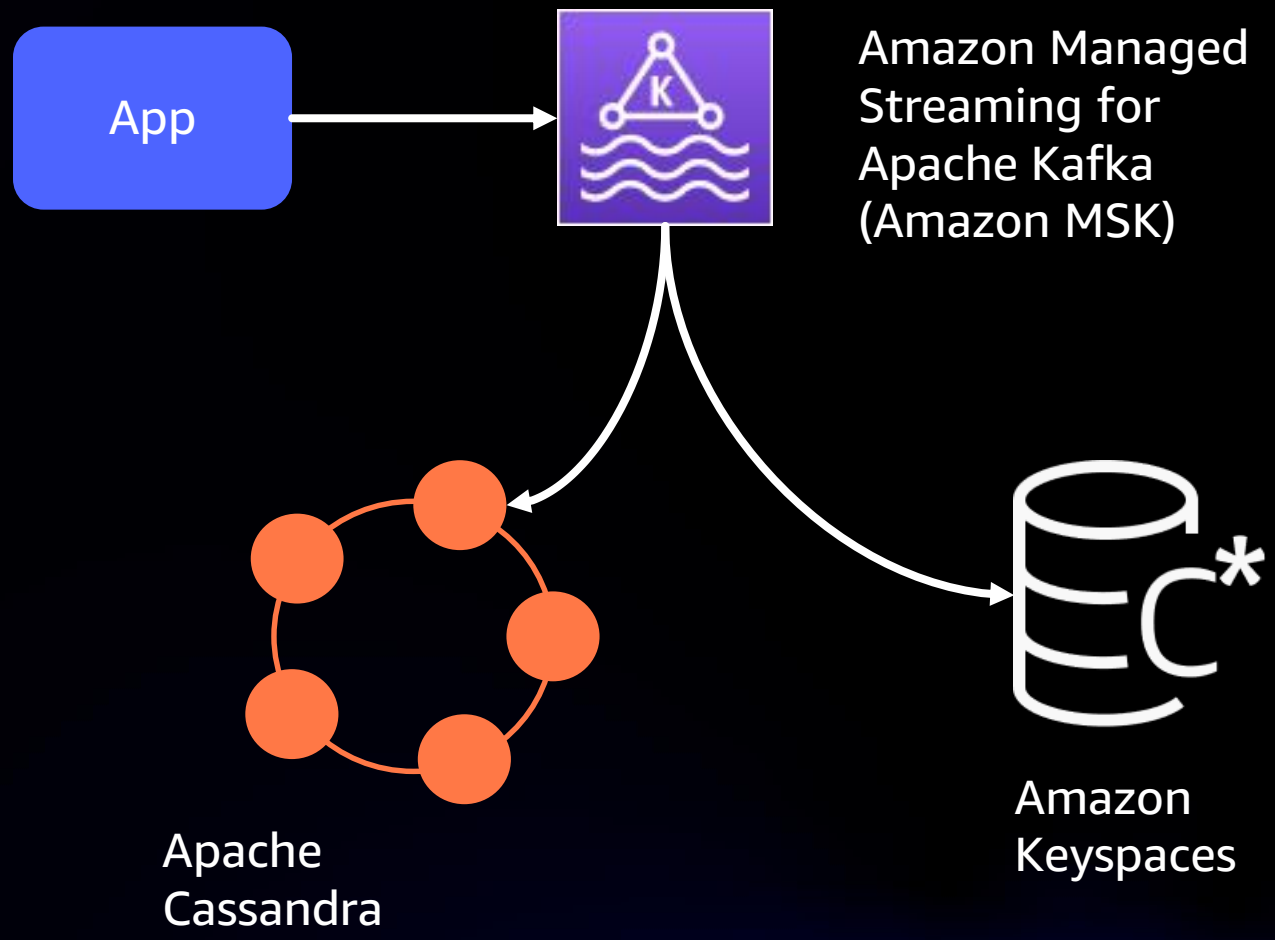
Keyspaces as Primary



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual writes migration from stream

- Write to Amazon Managed Streaming for Kafka
- Consumers for Amazon Keyspaces and Apache Cassandra
- Read from Amazon Keyspaces



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual writes migration Saga pattern



Insert transaction and details to transaction table

Table to insert / update

Duplicate insert / update

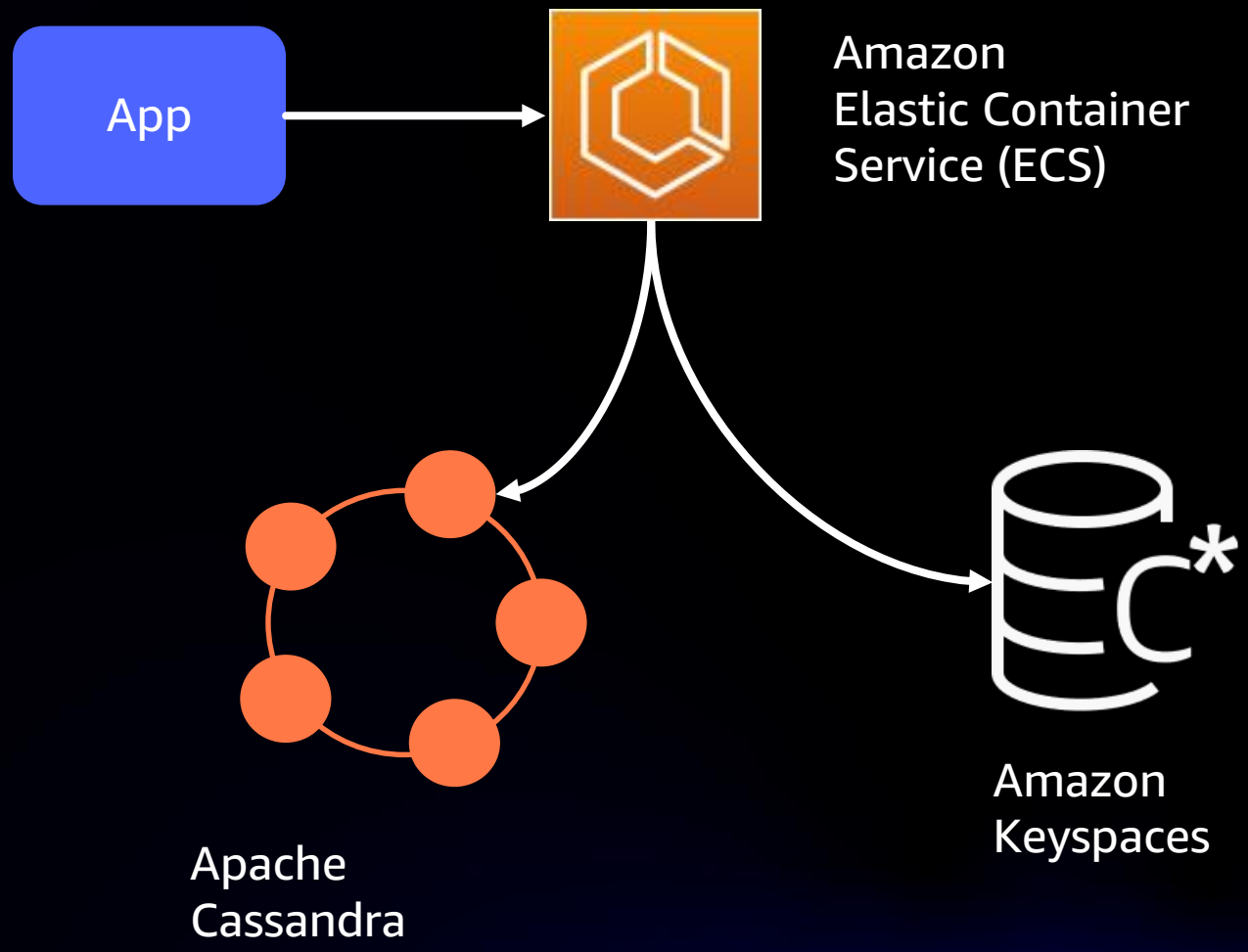
Remove/update transaction table

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Dual Writes migration from proxy

- Standup proxy application on Amazon ECS
- Perform dual writes to Apache Cassandra and Amazon Keyspaces
- Implement proxy as REST or gRPC service



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways!

- Break down clusters into tables and plan phased approach
- Choose the right migration strategy
- Work with AWS Solution Architects

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.