# aws INNOVATE
## DATA EDITION

**23 August, 2022**

aws

# Unlocking analytics capability by migrating your legacy on-premises data warehouse to Amazon Redshift

Randy Chng

Specialist Solutions Architect
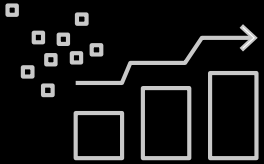Analytics Acceleration Lab
Amazon Web Services

Sean Beath

Specialist Solutions Architect
Analytics Acceleration Lab
Amazon Web Services

aws

# Trends driving data warehouse modernisation
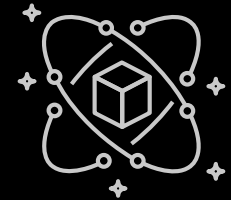
# Big data trends in the real world

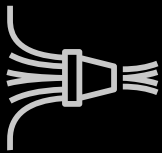New data sources,
sensors drive data growth

Empower personas – use
best-fit analytical services

Growth in open
data formats

Streaming data – analyze
without loading it first

Single source of truth –
unify analytics

Unified security,
compliance

**The complexity and growth of data brings its own set of unique challenges**

*What we're hearing from customers:*

"Analytics requirements are coming in from all corners. We don't have enough specialized skills."

"How do I derive value from growing, diverse data, living in data silos?"

"Performance has to be consistently fast. Costs are hitting the roof."

# Analyze all your data

# Amazon Redshift for scaling your data needs

## Analyze all your data

Break through data silos and gain real time and predictive insights on all your data in a few clicks, with no data movement or data transformation

## Price-Performance at scale

Gain up to 3x better price performance than any other cloud data warehouse, with performance innovation out of the box at no additional cost to you.
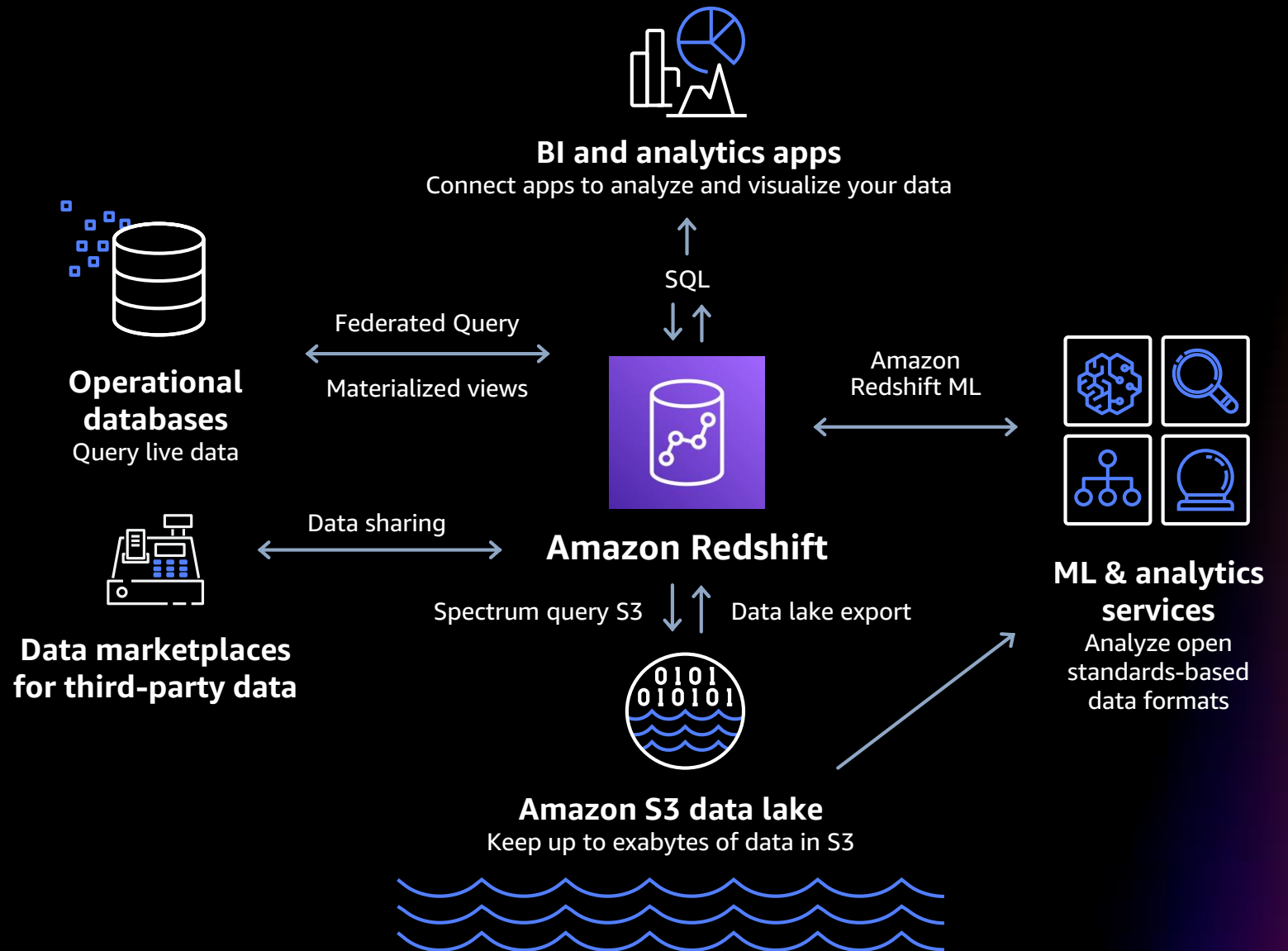
## Easy analytics for everyone

Focus on getting from data to insights in seconds, while we take care of data warehouse infrastructure and performance optimization for you

## Protect your data with confidence

Spend less time building custom solutions to monitor and protect your data, with built in security supported by AWS industry-leading compliance frameworks

# Break through data silos and analyse data across AWS services

**BI and analytics apps**
Connect apps to analyze and visualize your data

SQL

Federated Query

**Operational databases**
Query live data

Materialized views

Amazon Redshift ML

**Amazon Redshift**

Data sharing

**Data marketplaces for third-party data**

Spectrum query S3

Data lake export

**ML & analytics services**
Analyze open standards-based data formats

**Amazon S3 data lake**
Keep up to exabytes of data in S3

# 10 years of reinvention by listening to customers

**2012** → *Terabytes to petabytes to exabytes* → **2022**

## Customer challenges

Need a cost-effective way to run fast analytic queries on large volumes of data

Increasingly using data lakes and needing to query open-format data on Amazon S3 in place

Scale storage independently of compute

Elastically scale without user involvement

Analyze/query operational data in external data stores

Need to query data across clusters without data movement

Need analysts outside of the data science team to do ML (build, train, and apply models)

Simplified query interface for analysts

High-performance query acceleration without user involvement

Variety of third-party datasets and marketplace of data

Abstracting away complexity of capacity management

Real-time analytics in the warehouse

Concurrent scaling for writes

Granular authorization controls

Connections to more identity providers

## Amazon Redshift

Amazon Redshift launched as the first MPP cloud data warehouse at a price performance unattainable in on-premises MPP DWs

Amazon Redshift Spectrum

Amazon Redshift RA3 nodes

Concurrency scaling

Federated query

Data sharing

Amazon Redshift ML

Amazon Redshift Query Editor V2

AQUA

Data sharing with AWS Data Exchange

Amazon Redshift Serverless

Amazon Kinesis streaming ingestion

Concurrency writes

Role-based access control, identity management updates

# Use cases with Amazon Redshift

# Visualize data from all data sources

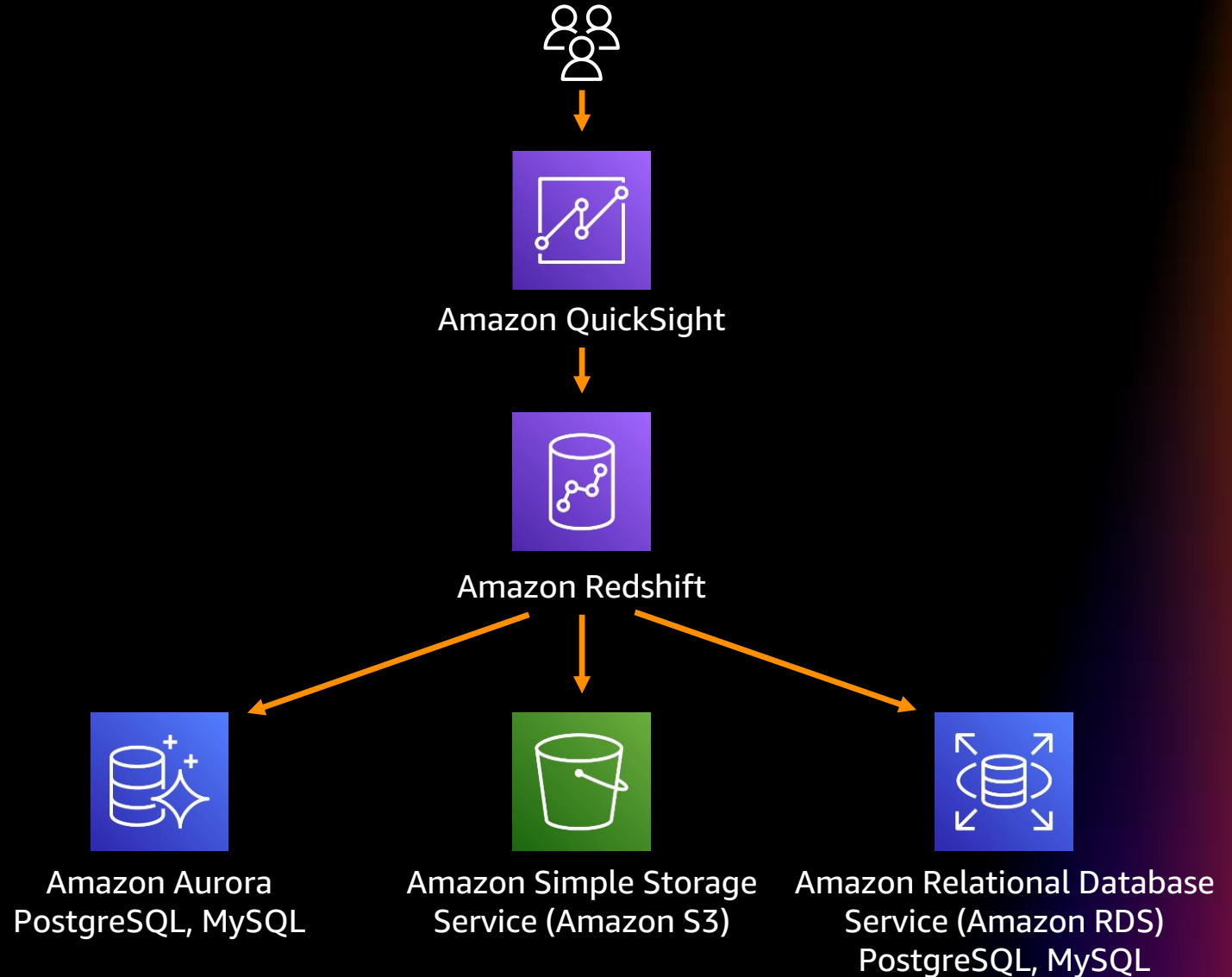Integrate operational data with data warehouse and data lake for real-time analytics

Analytics on operational data without data movement and ETL delays

Query and join data from one or more Amazon RDS and Aurora databases

Amazon RDS for MySQL and Amazon Aurora MySQL support

Amazon QuickSight

Amazon Redshift

Amazon Aurora
PostgreSQL, MySQL

Amazon Simple Storage
Service (Amazon S3)

Amazon Relational Database
Service (Amazon RDS)
PostgreSQL, MySQL

# Stream and analyse data in near-real time

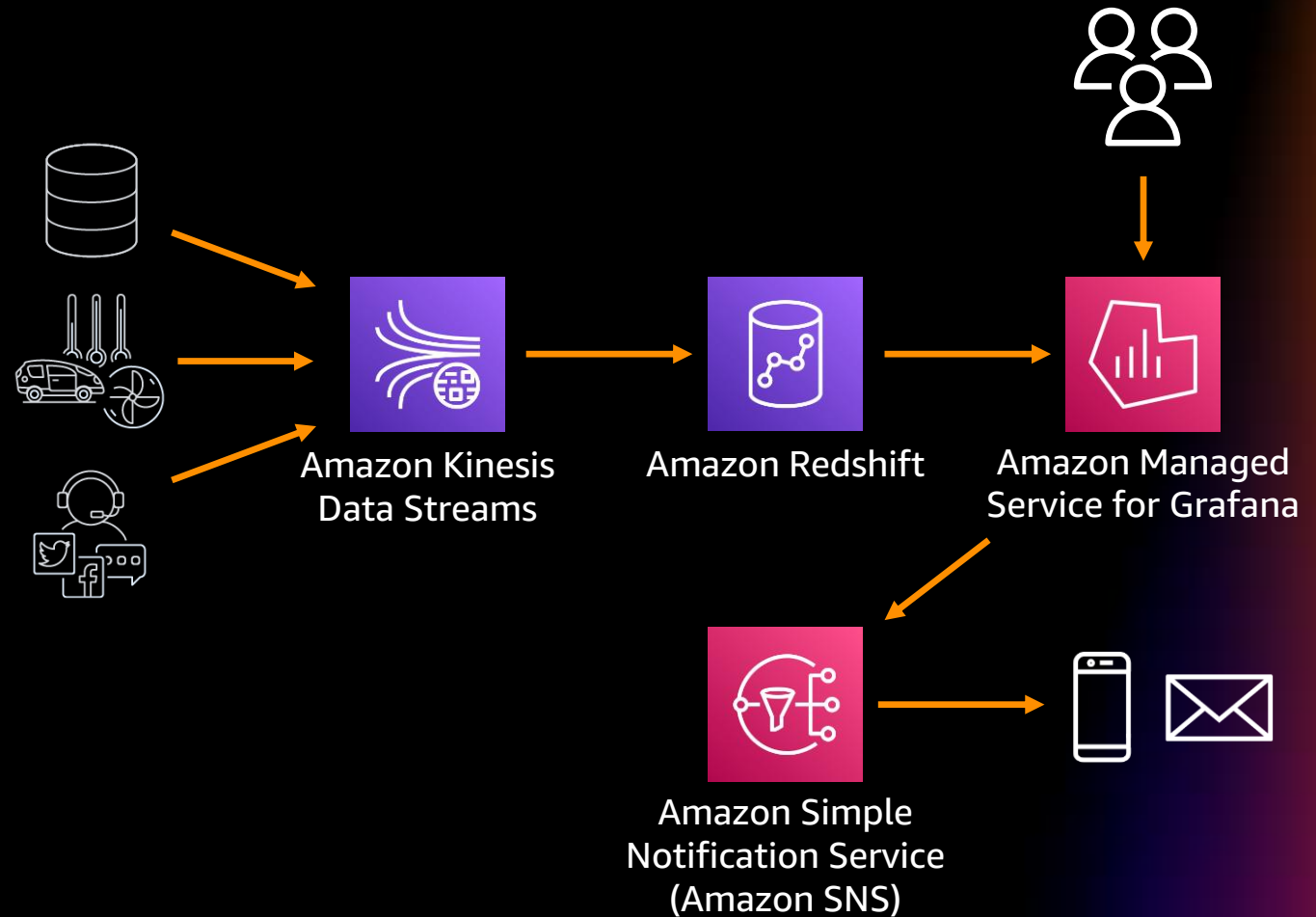Ingest hundreds of MB/s into Redshift with single-digit second latency with Amazon Kinesis Data Streams

Entirely configurable with SQL

Use Amazon Managed Service for Grafana to visualise metrics and send alerts in near-real time

Connect to an existing Amazon Kinesis Data Stream, Amazon Managed Streaming for Apache Kafka, or Kafka stream

Amazon Kinesis
Data Streams

Amazon Redshift

Amazon Managed
Service for Grafana

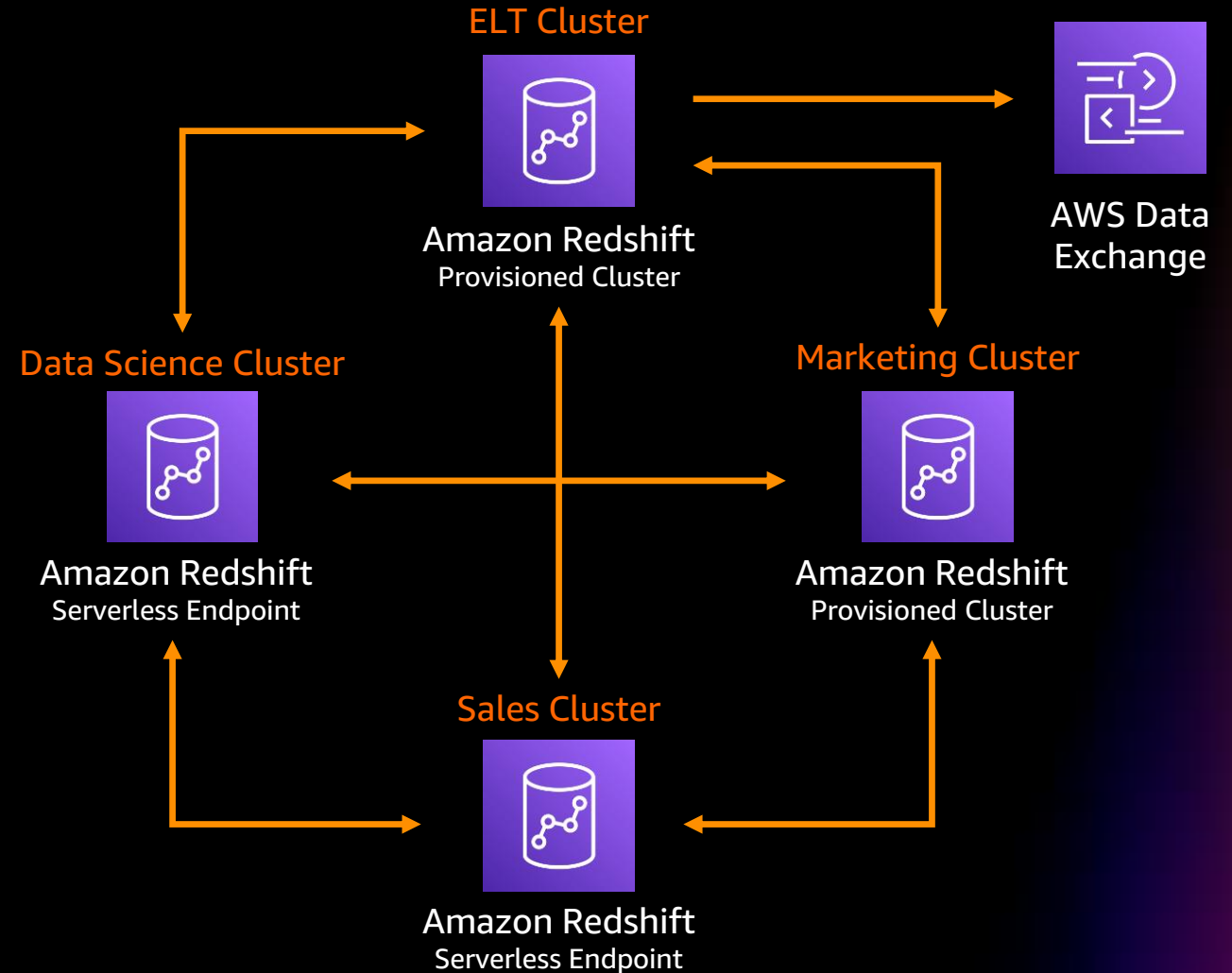Amazon Simple
Notification Service
(Amazon SNS)

# Combine data sharing with Amazon Redshift Serverless and AWS Data Exchange

Share data seamlessly between provisioned cluster to serverless end point and vice versa

Instant, live data access within and across AWS accounts and across AWS regions

Pay per second of runtime with no management required for Amazon Redshift Serverless

Package your Amazon Redshift data as data products, with pricing terms and conditions, with AWS Data Exchange

**ELT Cluster**

Amazon Redshift
Provisioned Cluster

AWS Data Exchange

**Data Science Cluster**

Amazon Redshift
Serverless Endpoint

**Marketing Cluster**

Amazon Redshift
Provisioned Cluster

**Sales Cluster**

Amazon Redshift
Serverless Endpoint

# Services to help you migrate

# Services to help you migrate



AWS Schema Conversion Tool
(AWS SCT)

AWS Database Migration Service
(AWS DMS)

Amazon S3 Transfer Acceleration

AWS Transfer Family

AWS DataSync

AWS Snow Family

AWS Direct Connect

AWS Storage Gateway
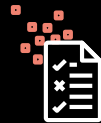
ISV connectors

Amazon Kinesis Data Firehose

# Data warehouse migration journey

## Analyse and plan

Discovery & planning

Migration business case

Migration expertise

Migration plan

## Workshop and pilot

### Workload analysis and pilot

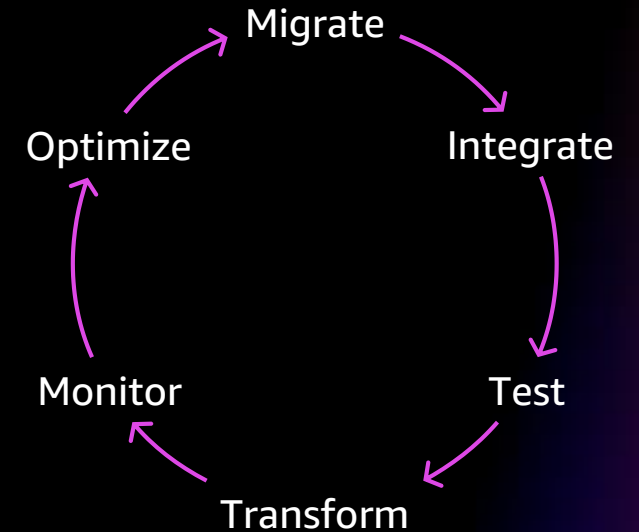Create new target

—

Modify/develop BI and apps to dual target

Modify/develop ETL to dual target

—

Skills/CoE

## Continuous workload migrations

### Migration

Migrate

Integrate

Test

Transform

Monitor

Optimize

AWS support throughout the journey

# Demo

aws

# Recap

- Trends driving data warehouse modernization

- Analyse all your data

- Use-cases with Amazon Redshift

- Services to help you migrate

- Demo

aws

# Experience our other Amazon Redshift sessions

| Session Name |
| --- |
| Unleash your organisation's data analytics potential with Amazon Redshift |
| Go beyond insights to predictive analytics with Amazon Redshift ML and Amazon SageMaker |
| Accelerate your time-to-insight with Amazon Redshift streaming |
| Build highly scalable and extensible SQL-based data pipeline using Managed Workflow for Apache Airflow and Amazon Redshift |
| Go beyond insights to predictive analytics with Amazon Redshift ML and Amazon SageMaker Canvas |

# Visit the AWS Data resource hub

A modern data strategy can help you manage, act on, and react to your data so you can make better decisions, respond faster, and uncover new opportunities. Dive deeper with these resources today.

- Harness data to reinvent your organization
- In unpredictable times, a data strategy is key
- Make data a strategic asset
- Rewiring your culture to be data-driven
- Put your data to work with a modern analytics approach
- … and more!

https://tinyurl.com/data-hub-aws

**Visit resource hub**

aws
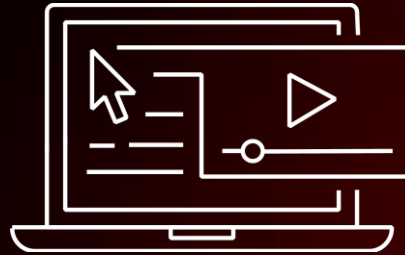
# AWS Training and Certification for Data and Analytics

**AWS Data & Analytics
FREE Training Resources**

Discover how to harness data, one of the world's most valuable resources, and innovate at scale.

https://bit.ly/3Ntlhy7

**AWS Data Analytics
Learning Plan**

This learning plan expose you to the fastest way to get answers from all your data to all your users. It can also help prepare you for the AWS Certified Data Analytics - Specialty certification exam.

https://bit.ly/3wBVjD1

**AWS Certified Data
Analytics - Specialty**

Earning AWS Certified Data Analytics – Specialty validates expertise in using AWS data lakes and analytics services.

https://go.aws/3lwF0RR

# Thank you!

aws