
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

PEX308

Reinvent your business with an AWS modern data strategy

Sudhir Gupta (he/him)

Principal Partner Solutions Architect, Analytics Specialist
AWS

Veena Vasudevan (she/her)

Senior Partner Solutions Architect, Analytics Specialist
AWS

Sanchit Jain (he/him)

Data Analytics and Cloud Practice Lead, AWS Hero and AWS Ambassador
Quantiphi Inc.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Voice of customers

AWS modern data strategy

Demo – Building a serverless modern data architecture

Customer case studies and reference architectures

Data mesh pattern

Call to action

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Voice of customers

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Decision-making in the past

In the past, decision-making revolved around the **enterprise data warehouse**



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data no longer fits

There is more data and more diversity of data than people think

Data growth

>10x
every 5 years

Data platforms needs

To live for
15+
years

To scale
1,000x



IDC, “Data Age 2025”

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Accessibility of data



Data scientists



Business users



Analysts



Applications



Machine learning



SQL analytics



Scientific



Real-time, streaming

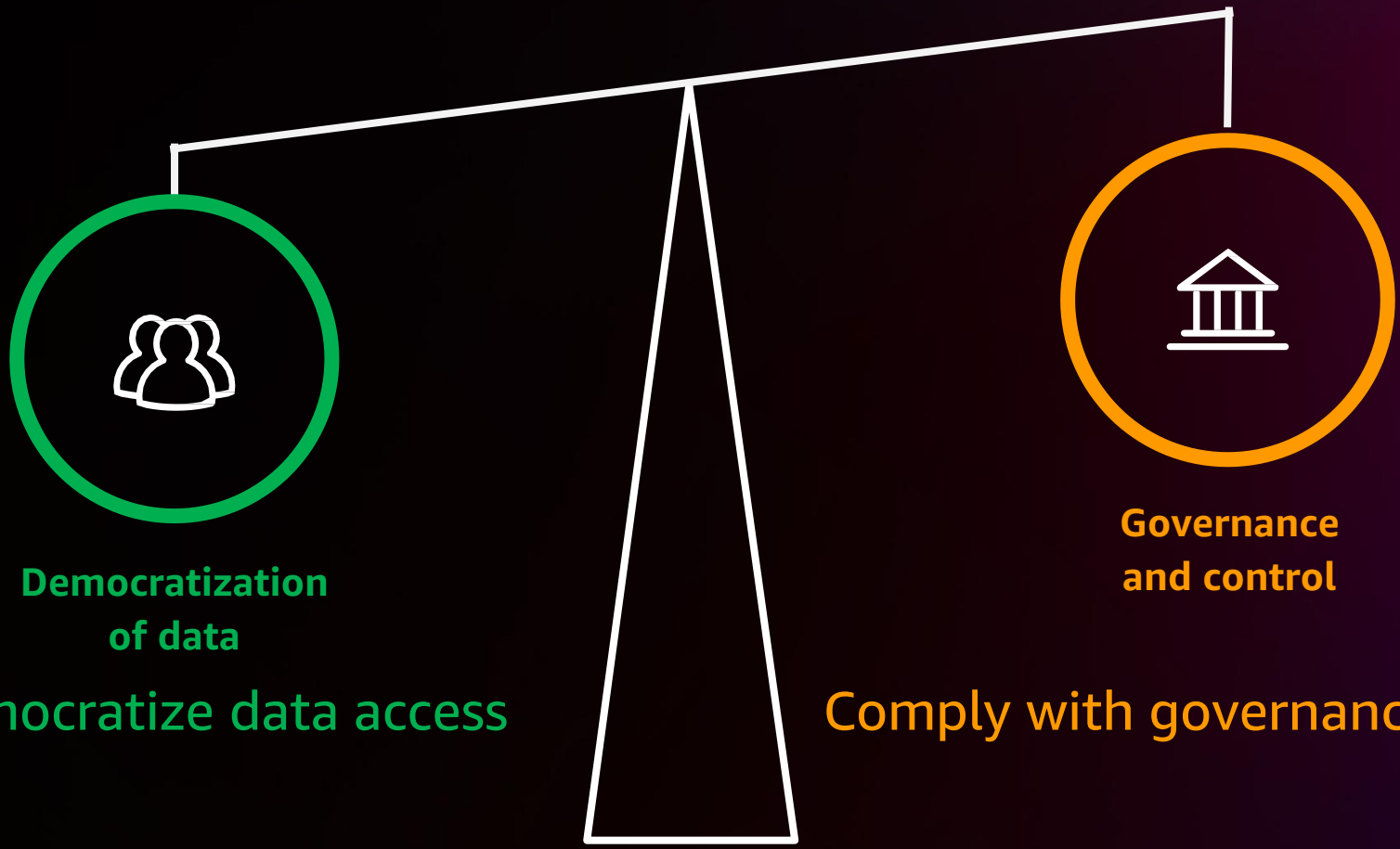
There are **more people** accessing data

And in **different ways**

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

More regulatory pressure



Democratization of data

Democratize data access

Governance and control

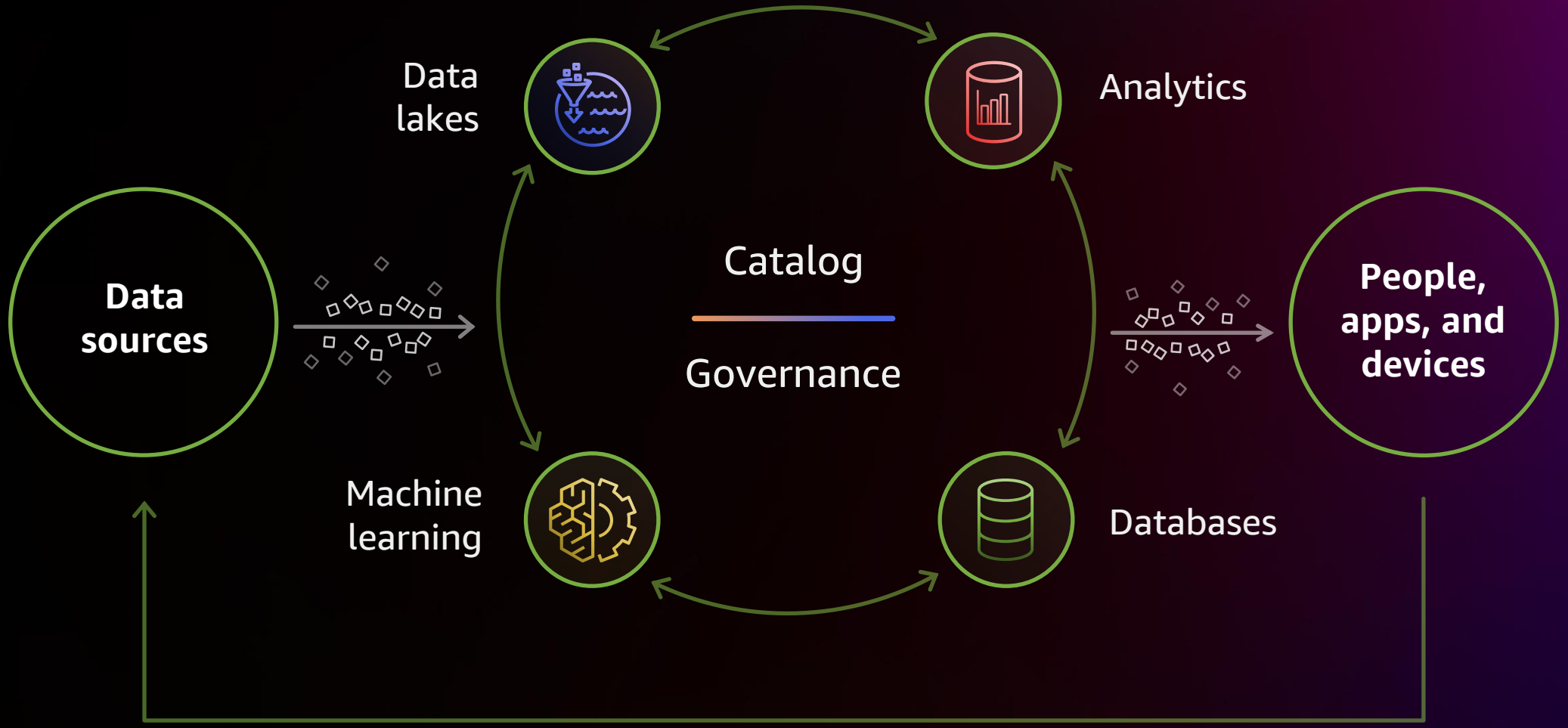
Comply with governance regulations

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data strategy
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data strategy



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS modern data strategy components



Scalable data lakes



Unified data access, security, and governance



Purpose-built for performance and cost



Serverless and easy to use



Built-in machine learning

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Scalable data lakes

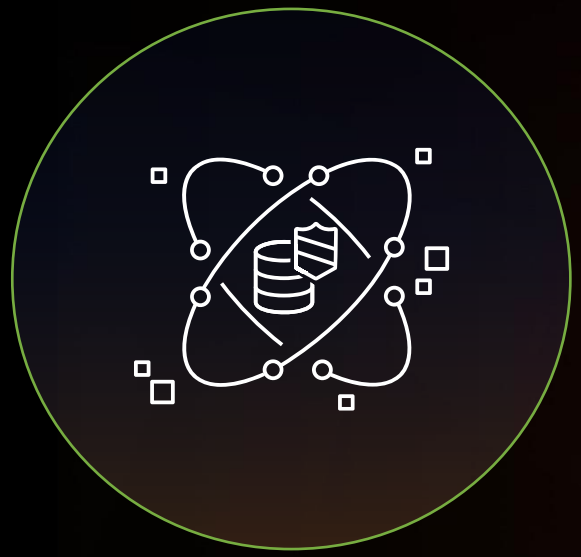
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon S3: Data lakes on AWS

- Store unlimited data in open file formats
- Unmatched durability, availability, and scalability
- Decouple storage from compute
- Choice of analytical and ML engines
- Pay as you go



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



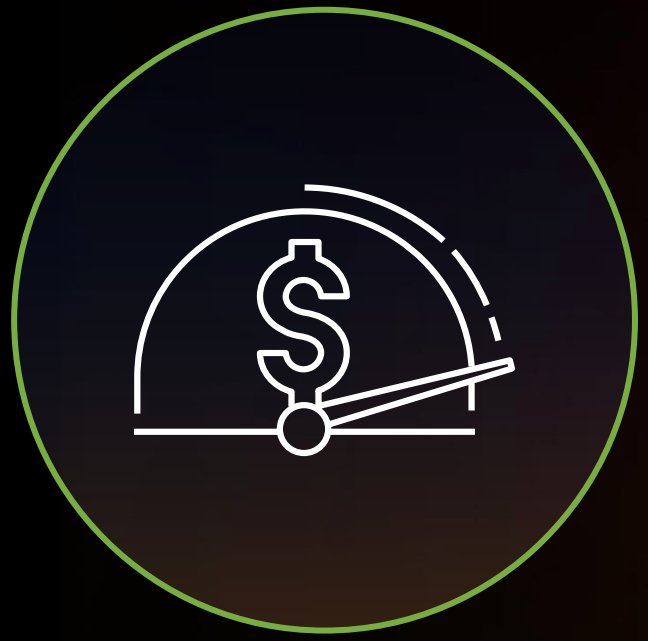
Unified data access, security, and governance

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Lake Formation: Unified data governance



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Purpose-built for performance and cost

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A family of purpose-built data services

Business intelligence and machine learning

- **Amazon QuickSight** — Visualizations
- **AWS Data Exchange** — Data exchange
- **Amazon SageMaker** — ML
- **Amazon Comprehend** — NLP
- **Amazon Transcribe** — Speech-to-text
- **Amazon Textract** — Extract text
- **Amazon Personalize** — Recommendations
- **Amazon Forecast** — Forecasts
- **Amazon Translate** — Translation
- **Amazon CodeGuru** — Code reviews
- **Amazon Kendra** — Enterprise search

Plus many more

Analytics

- **Amazon Redshift** — Data warehousing
- **Amazon EMR** — Hadoop + Spark
- **Amazon Athena** — Interactive analytics
- **Amazon OpenSearch Service** — Operational analytics
- **Amazon Kinesis Data Analytics** — Real time

Databases

- **Amazon Aurora** — MySQL, PostgreSQL
- **Amazon RDS** — MySQL, PostgreSQL, MariaDB, Oracle, SQL Server, **Amazon RDS on VMware**
- **Amazon DynamoDB** — Key value, Document
- **Amazon DocumentDB** — Document
- **Amazon ElastiCache** — Redis, Memcached
- **Amazon Keyspaces (for Apache Cassandra)** — Wide column
- **Amazon Neptune** — Graph
- **Amazon QLDB** — Ledger database
- **Amazon Timestream** — Time series

Blockchain

- **Amazon Managed Blockchain**
- **AWS Blockchain Templates**

Data lake

- **Amazon S3/ Amazon S3 Glacier**
- **AWS Lake Formation** — Data lakes
- **AWS Glue** — ETL and Data Catalog

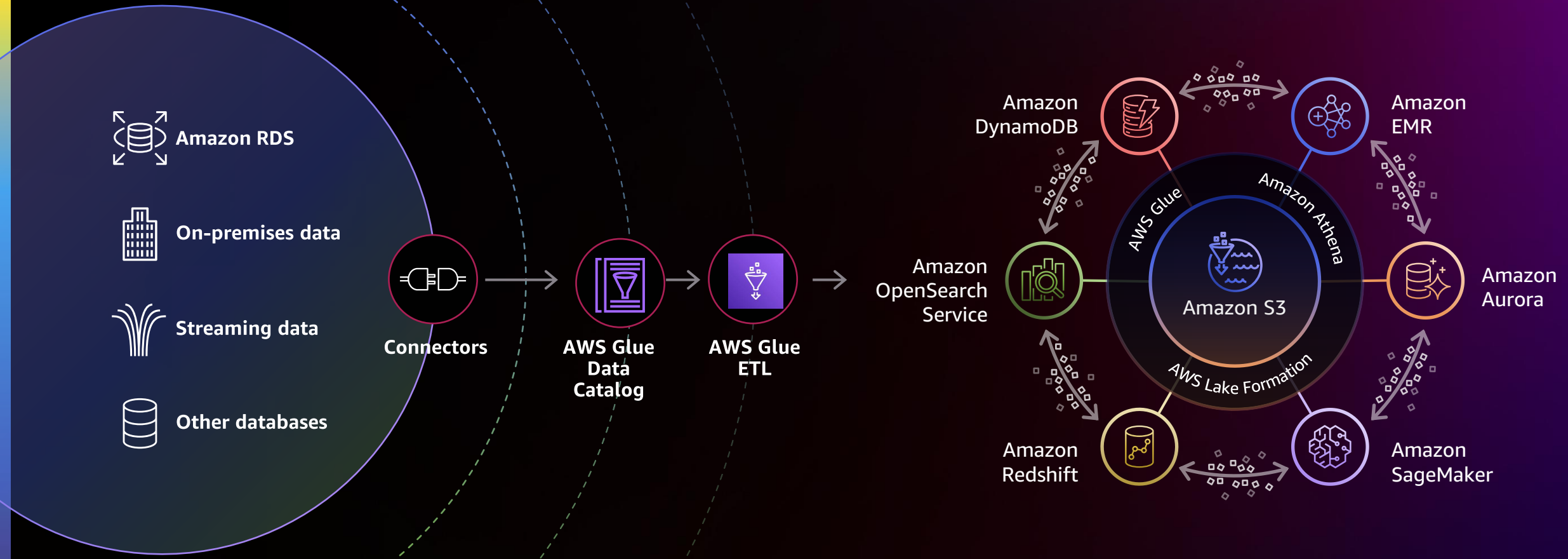
Data movement

AWS DMS | AWS Snowball | AWS Snowmobile | Amazon Kinesis Data Firehose | Amazon Kinesis Data Streams | Amazon MSK

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Seamless data integration



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Integrate third-party data products

Better technology

Multiple delivery methods for third-party data



Secure and compliant with encrypted data and AWS Identity and Access Management (IAM) integration



Natively integrate with AWS and non-AWS services





Easier administration



One place to exchange data publicly or privately



Migrate existing subscriptions at no additional cost



Simplified contracts and consolidated, secure billing

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Out-of-the-box benefits

PERFORMANCE, SCALE, AND COST



AWS Glue

AWS Glue – <1 min. startup time



Amazon S3

Amazon S3 Intelligent-Tiering saves up to 70% on storage costs



Amazon EMR runtimes

- Queries run 3.1x faster compared to standard Apache Spark at 40% of the cost
- Queries run up to 2.6x faster than standard Presto



Amazon OpenSearch Service

- Store 9x more log data at a lower cost
- UltraWarm storage tier reduces costs by up to 90%



Amazon Redshift

- Amazon Redshift Serverless pricing reduced by 25% in GA
- Up to 5x better price performance compared to other cloud data warehouses

aws Compute integration

With AWS Graviton2 instances, customers save 25.7% for typical workloads

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Serverless and easy to use

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Benefits of serverless

1. Think big
2. Start small
3. Go fast



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Serverless options for data analytics in the cloud



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Low-code/no-code environments



Amazon EMR Studio



AWS Glue DataBrew



AWS Glue Studio



Amazon Redshift query editor v2



AWS Data Exchange

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Built-in machine learning

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS brings ML closer to data



Databases



Data warehouses and data lakes



Business intelligence tools

Amazon Aurora ML



Amazon Neptune ML



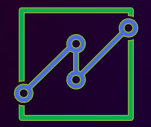
Amazon Redshift ML



Amazon Athena ML



Amazon QuickSight ML



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

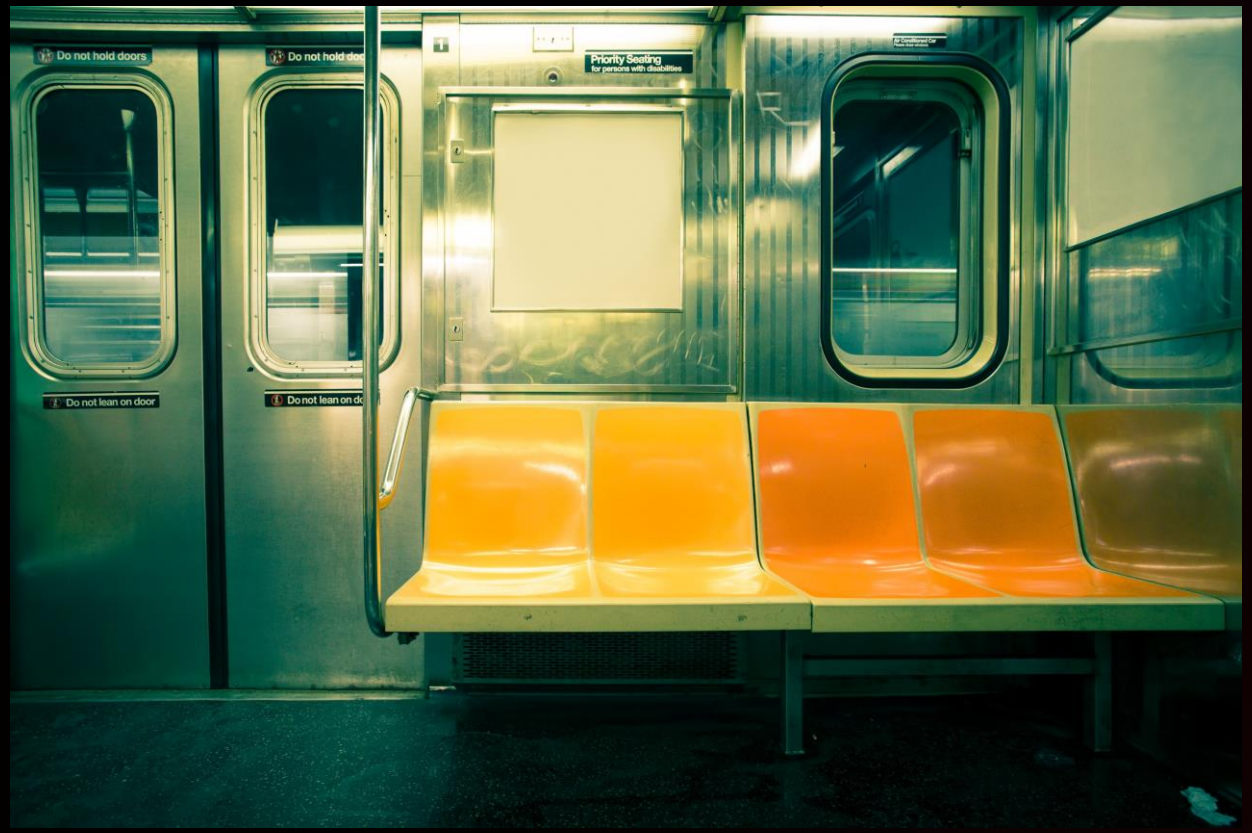
Demo
Building a serverless modern data architecture

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo: Use case

TRANSFORM, ANALYZE, AND VISUALIZE NEW YORK CITY TRANSIT DATA IN REAL TIME

Metropolitan Transportation Authority



Sign up for MTA API developer resources:
https://api.mta.info/#/landing

Train Routes	MTA API Endpoints
ACE	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-ace
BDFM	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-bdfm
G	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-g
JZ	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-jz
NQRW	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-nqrw
L	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-l
123456	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs
7	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-7
SIR	https://api-endpoint.mta.info/Dataservice/mtagtfsfeeds/nyct%2Fgtfs-si

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo: Architecture

TRANSFORM, ANALYZE, AND VISUALIZE NEW YORK CITY TRANSIT DATA USING FULLY SERVERLESS SERVICES

Kinesis producers running on AWS Lambda functions publish NYC Subway raw data feeds from MTA APIs to Kinesis data streams



AWS Lambda
Routes – A, C, E



AWS Lambda
Routes – B, D, F, M



AWS Lambda
Routes – J, Z



AWS Lambda
Routes – 1, 2, 3, 4, 5, 6

Kinesis data streams receive NYC Subway data



Amazon Kinesis Data Streams

Spark streaming application consumes raw feeds from Kinesis, sanitizes and transforms data, and stores it in Amazon S3



Amazon EMR Serverless

Transformed data files from Spark streaming application are deposited



Amazon S3

Run interactive queries and perform exploratory analysis on transformed data residing in Amazon S3



Amazon Redshift Serverless



Amazon Athena

Create live dashboards and visualize data from the Amazon Redshift table



Amazon QuickSight

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

GitHub link: https://github.com/vasveena/reinvent_demo



Case studies and reference architectures

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Case study: Connected vehicle platform

quantiphi
Solving What Matters

ONE OF THE WORLD'S LEADING LEASING COMPANIES, WITH 1.7 MILLION VEHICLES IN OVER 30 COUNTRIES

Challenge

- Rudimentary system
- No real-time analytics
- Lack of dynamic scalability
- Difficulty in troubleshooting

Solution

- Scalable and secure platform
- Enhanced syncing capabilities
- Serverless ingestion
- Single-click deployment

Impact

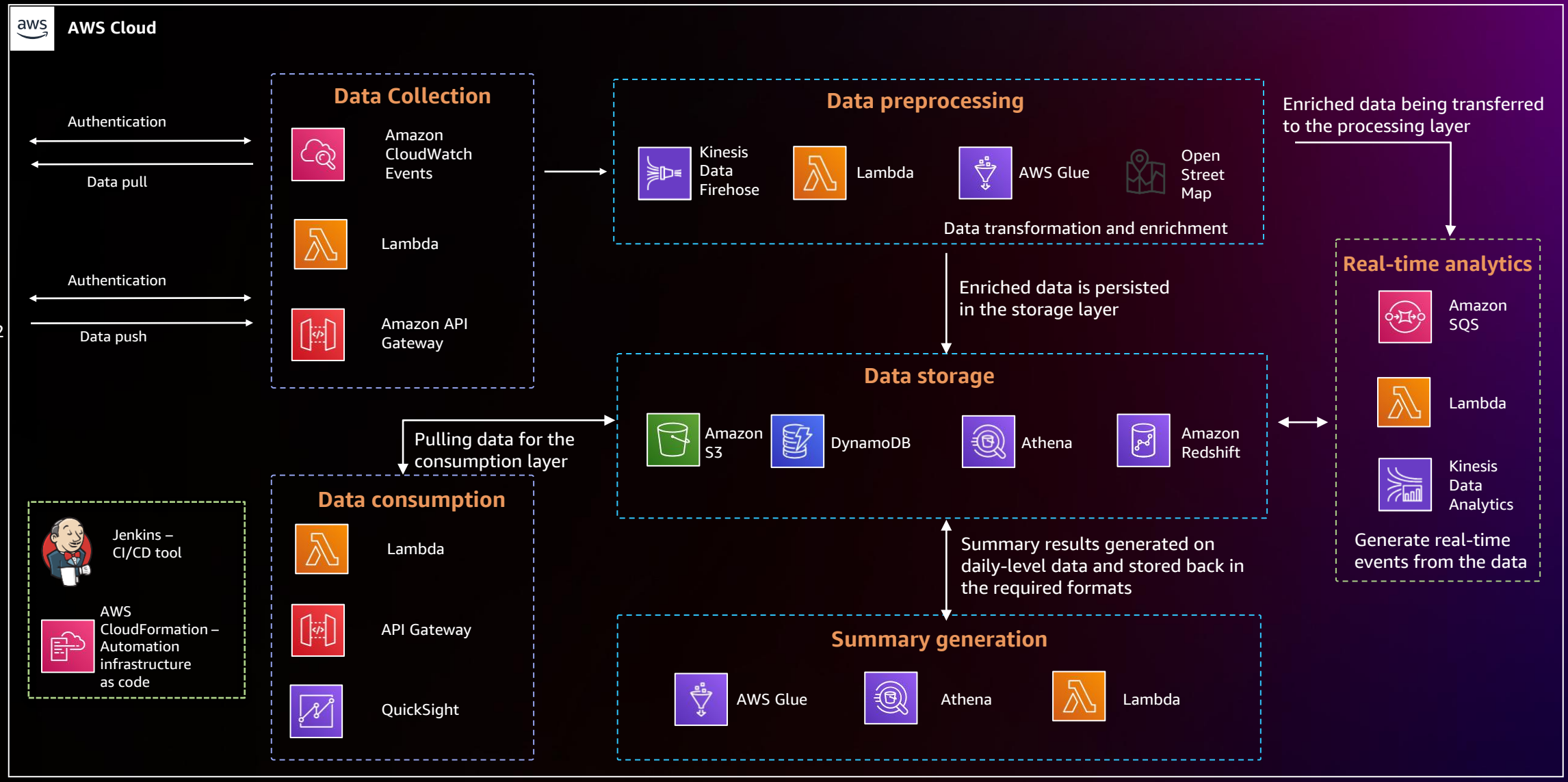
- **50%** Reduction in infrastructure cost
- **70%** Reduction in deployment time
- Facilitated real-time analytics
- Generated deeper insights for workflows
- Flexible and modular platform

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture: Connected vehicle platform

quantiphi
Solving What Matters



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Case study: Smart manufacturing

quantiphi
Solving What Matters

THE WORLD'S LEADING SUPPLIER OF TWO-PIECE CAN AND END-MAKING MACHINERY

Challenge

- Reactive decision-making
- Lack of data pipeline
- Lack of analytics and BI workflow
- Poor process efficiency

Solution

- Dynamic data ingestion pipelines
- Automated ETL pipelines
- A robust and scalable lake house
- A self-service BI platform

Impact

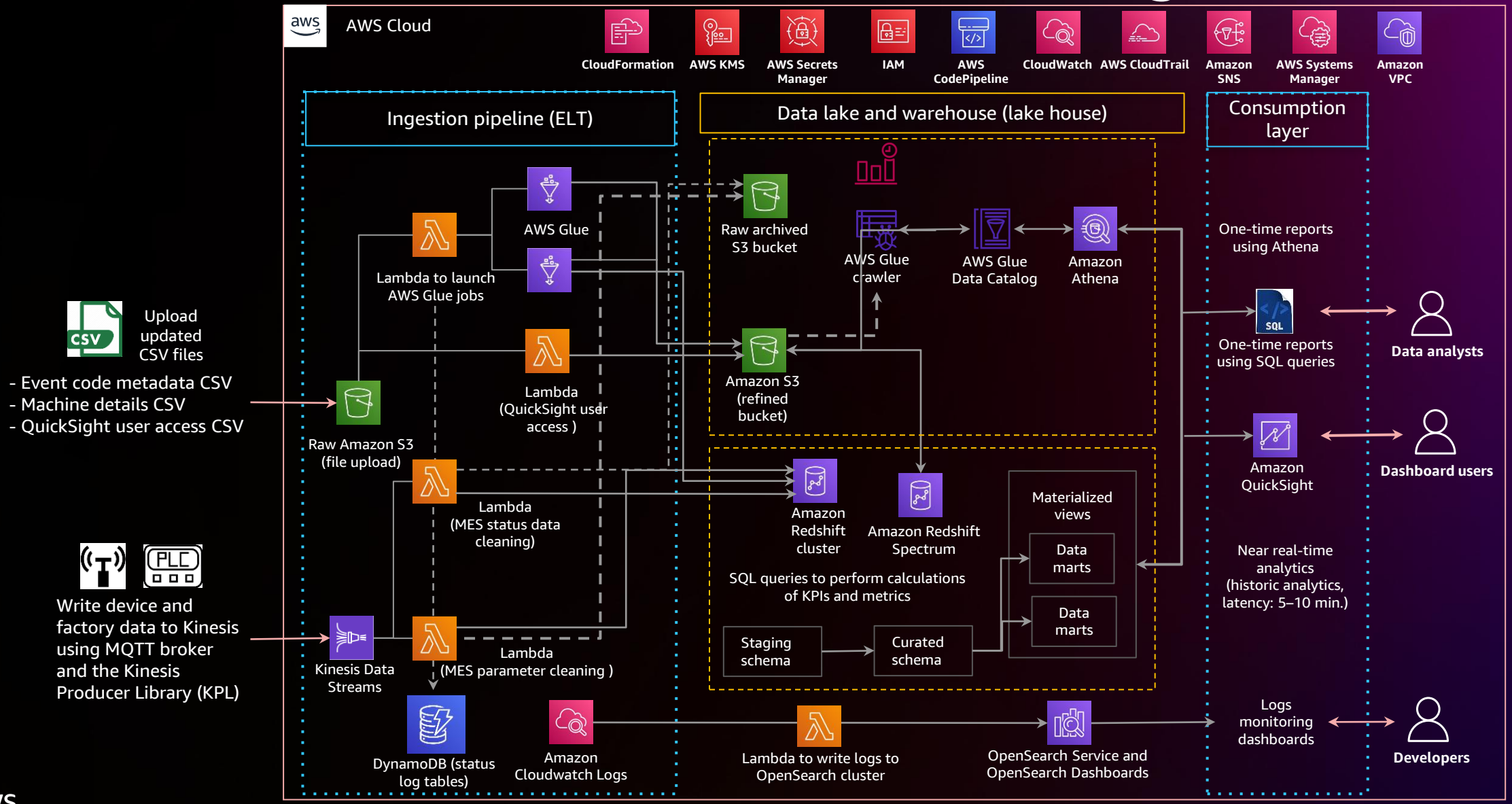
- **80%** Improvement in business efficiency
- Consolidated monitoring and reporting
- Self-service BI capabilities for insights
- Better inventory visibility

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Architecture – Smart manufacturing

quantiphi
Solving What Matters



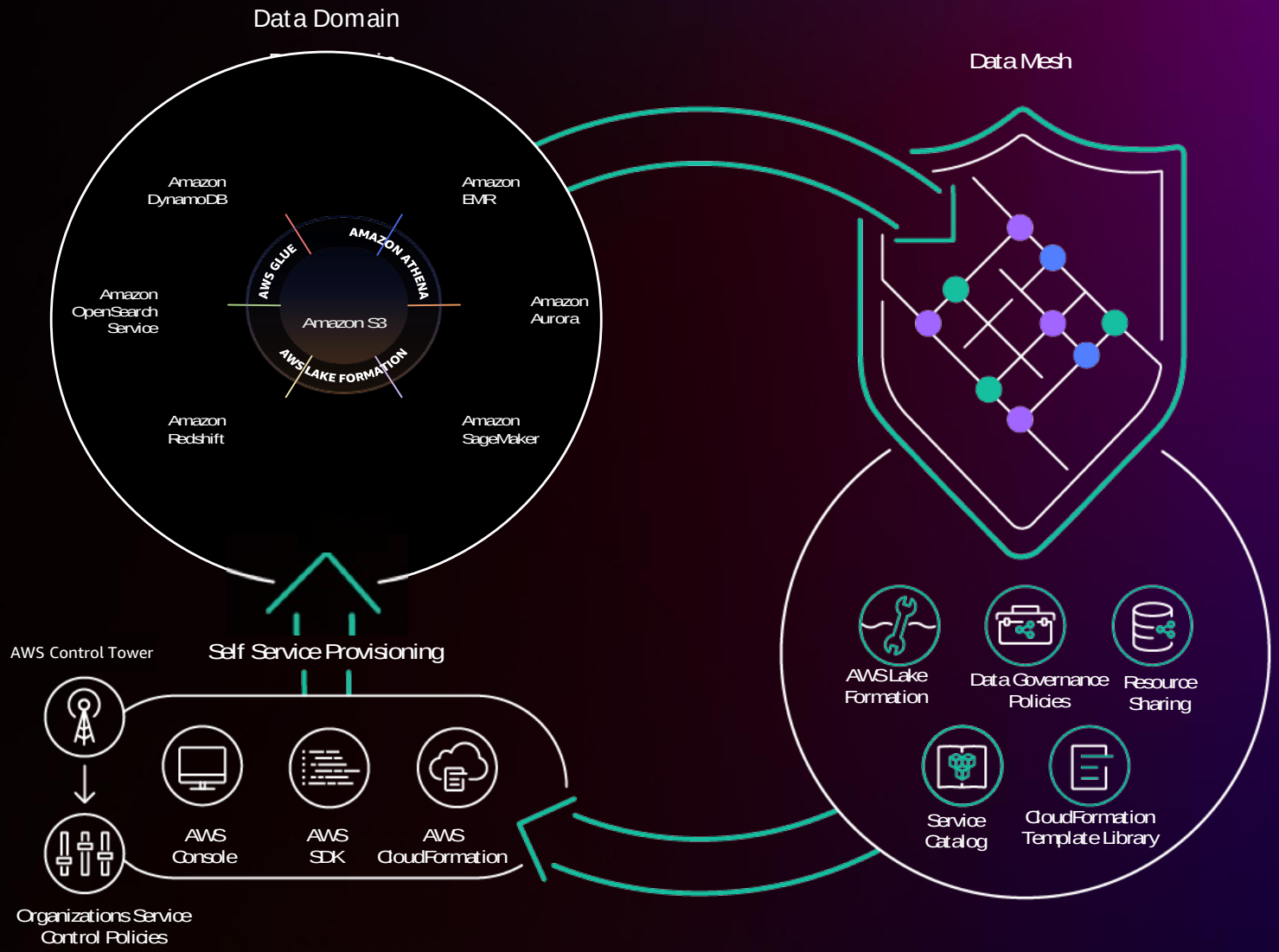
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh pattern with an AWS modern data architecture

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

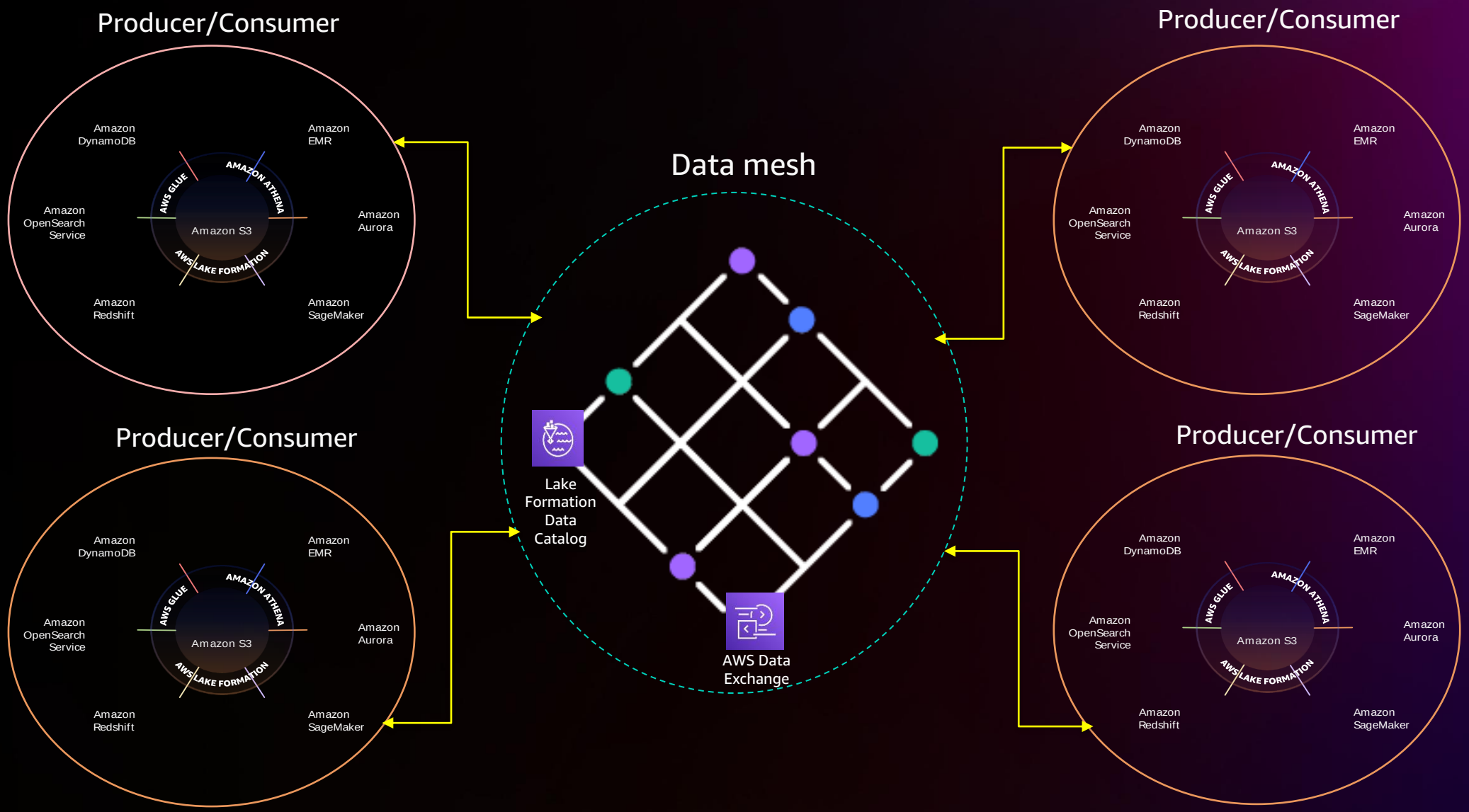
What is a data mesh pattern?

- Decentralized
- Domains ownership
- Data as a product
- Self-service data platform
- Federated governance



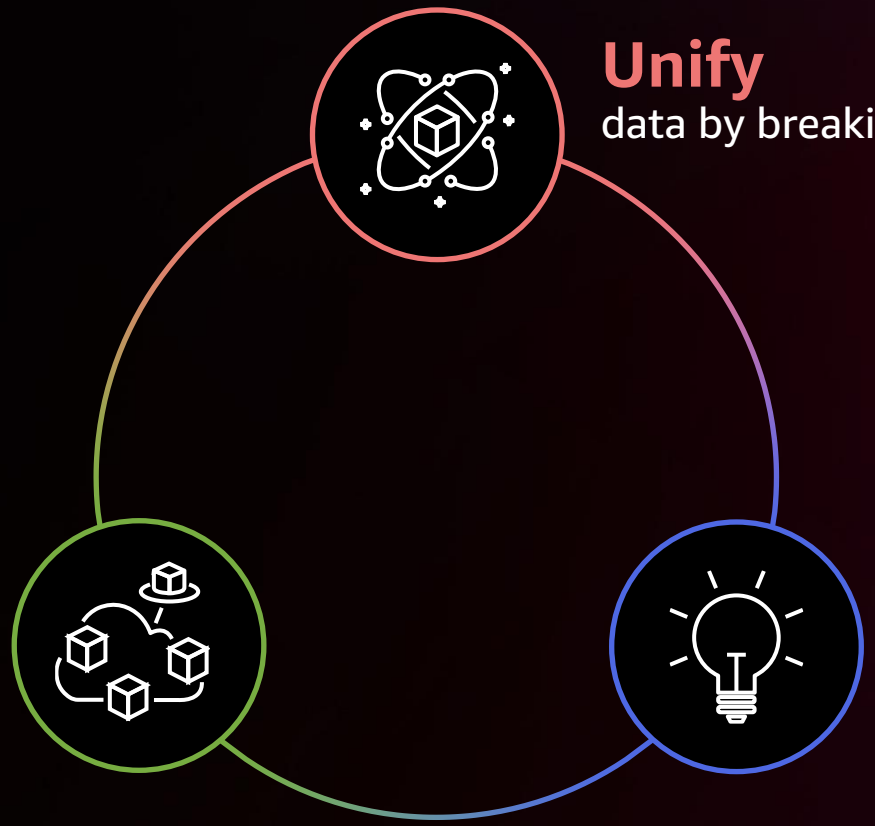
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh pattern on AWS



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Modern data strategy journey



Unify
data by breaking down silos

Modernize
data infrastructure to a scalable, trusted, and secure cloud provider

Innovate
by inventing new experiences and reimagining existing processes

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

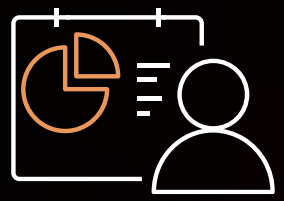
AWS Data & Analytics Partner trainings



Super classes



Self-service courses



Positioning & service roadmap



Train the trainer workshops

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Data & Analytics Partner programs

Analytics SI Acceleration Program (ASAP)

Use-case oriented, prescriptive and packaged offering to help SI partners

Get trained on migration methodologies, build assets, and migration content

Develop a strong go-to-market strategy and drive customer wins

AWS Service Delivery Program

https://aws.amazon.com/partners/programs/service-delivery/

AWS Data and Analytics Competency

https://aws.amazon.com/partners/competencies/

Validated AWS Partners with deep technical knowledge, experience, and proven customer success

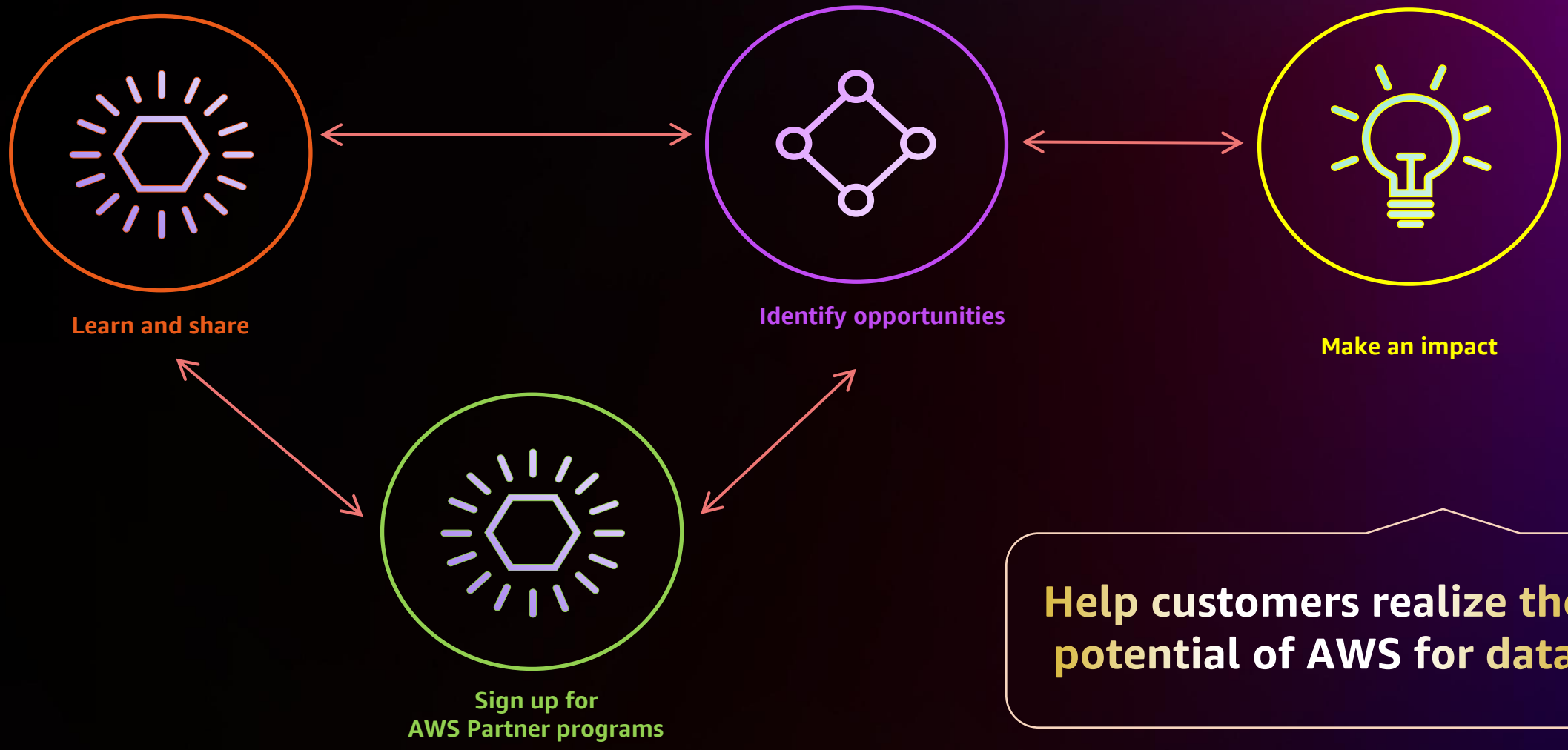
Benefits: Listing on AWS webpage, AWS Partner badge, exclusive Deeper Learning Series,

Co-sell with AWS, Marketing Development Fund, APN TV and APN Blog opportunities, and more

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Call to action



Help customers realize the potential of AWS for data

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Sudhir Gupta
LinkedIn: @sudhigaws

Veena Vasudevan
LinkedIn: @vasudevanveena

Sanchit Jain
LinkedIn: @sanchitdilipjain



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.