AWS re:Invent

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

TRV301

Travel and hospitality data platform for democratizing enterprise data

Paul Armstrong (him/his)
Principal Solutions Architect EMEA
AWS

Bib Shukla (him/his)
Senior Customer Delivery Architect
AWS

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

Agenda

Data mesh

Data product

Domain specific lake house

Data mesh organization principles

Data mesh core concepts in AWS

Discussion

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

Common themes we hear from our customers

Sharing and searching data is difficult

Difficult to meet all requirements across differing business units

My team needs to own datasets, pipelines and repositories that are isolated from other teams

I wish to focus on innovating with data, not on maintaining and administering a data lake

My data science team should easily find the datasets they seek and have the ability to share them with others

If I share data, I've lost control

Why doesn't our organization treat data as a product?

There is a mismatch between executive leadership goals and business line deliverables and incentives

Need to create a model to support sharing from both producers and consumers of data

Current data architecture is complex and monolithic and slow to change

Our internal policies on what can be shared is unclear and there is lack of incentive to share

I just want to get access to the data I need

aws

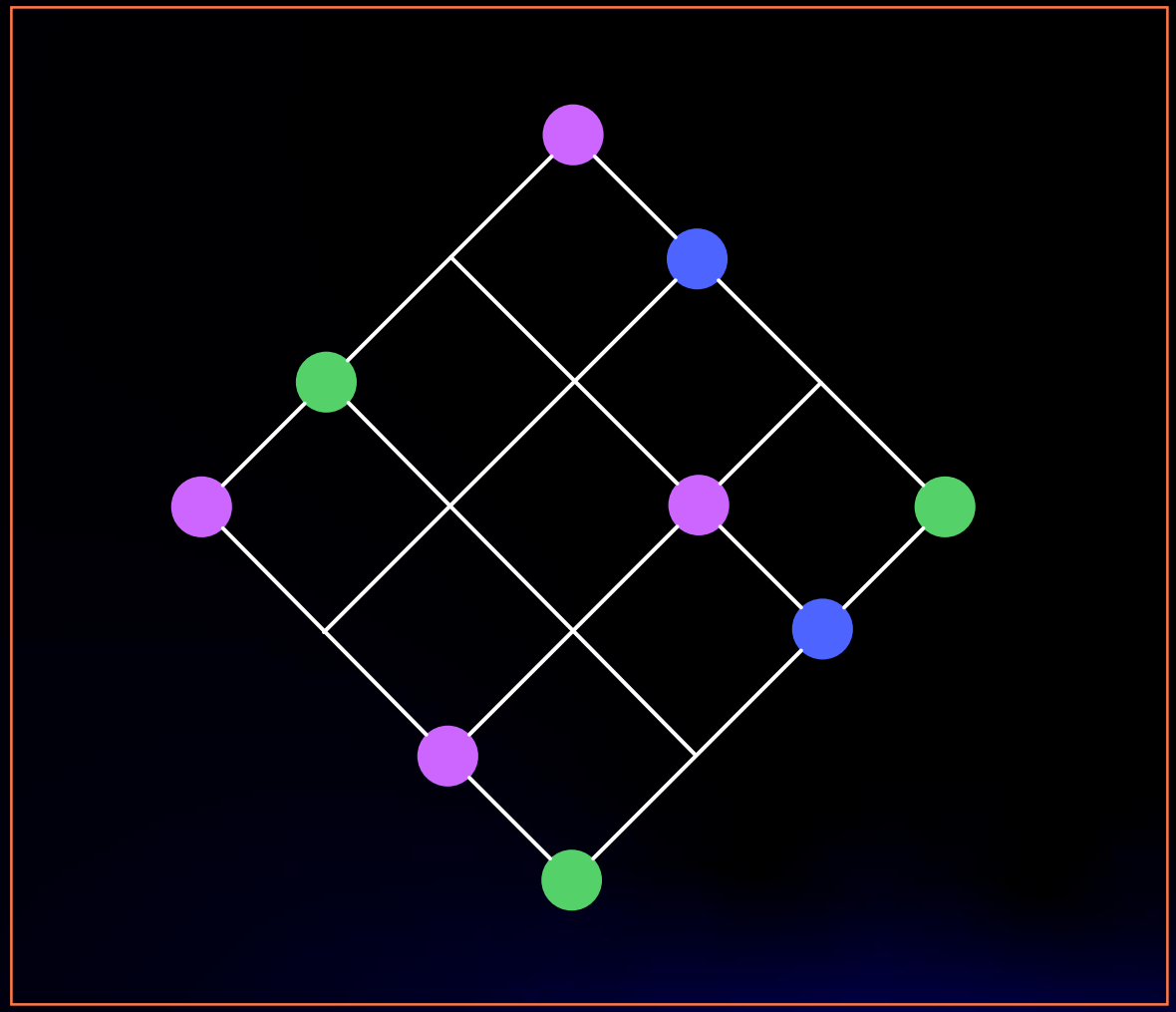
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is a data mesh?

Data mesh is a pattern for defining how organizations can organize around one or more data domains with a focus on delivering enterprise data products

Key principals

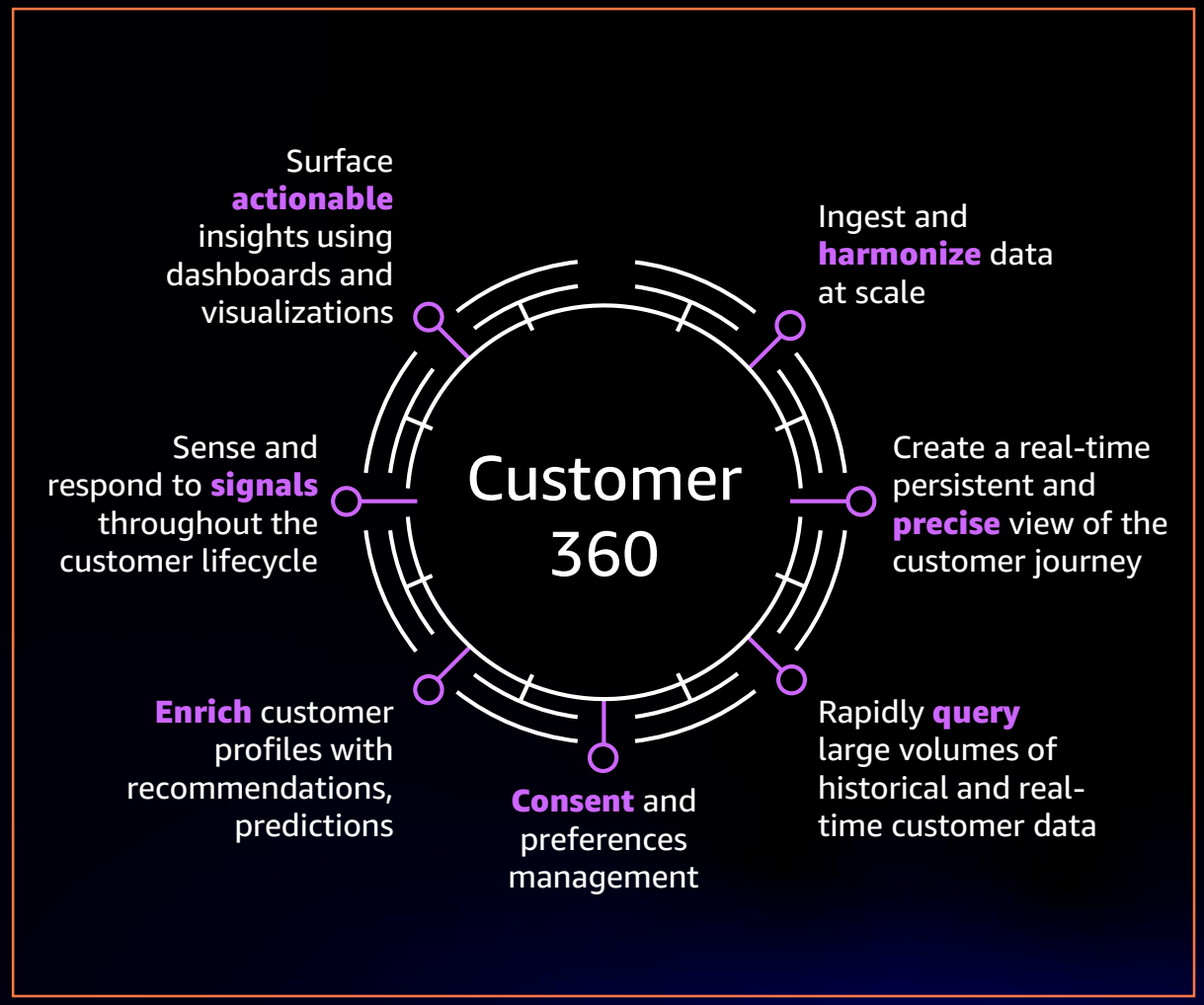
- Federated computational governance
- Self-serve data platform
- Data as a product
- Domain ownership



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

What is a data product?

- Broad, cohesive collections of related data aligned to business use cases and goals
- Federated governance of data created, provided, stored, transformed in, or consumed from the data product
- Diverse data types sourced from distributed internal and external sources



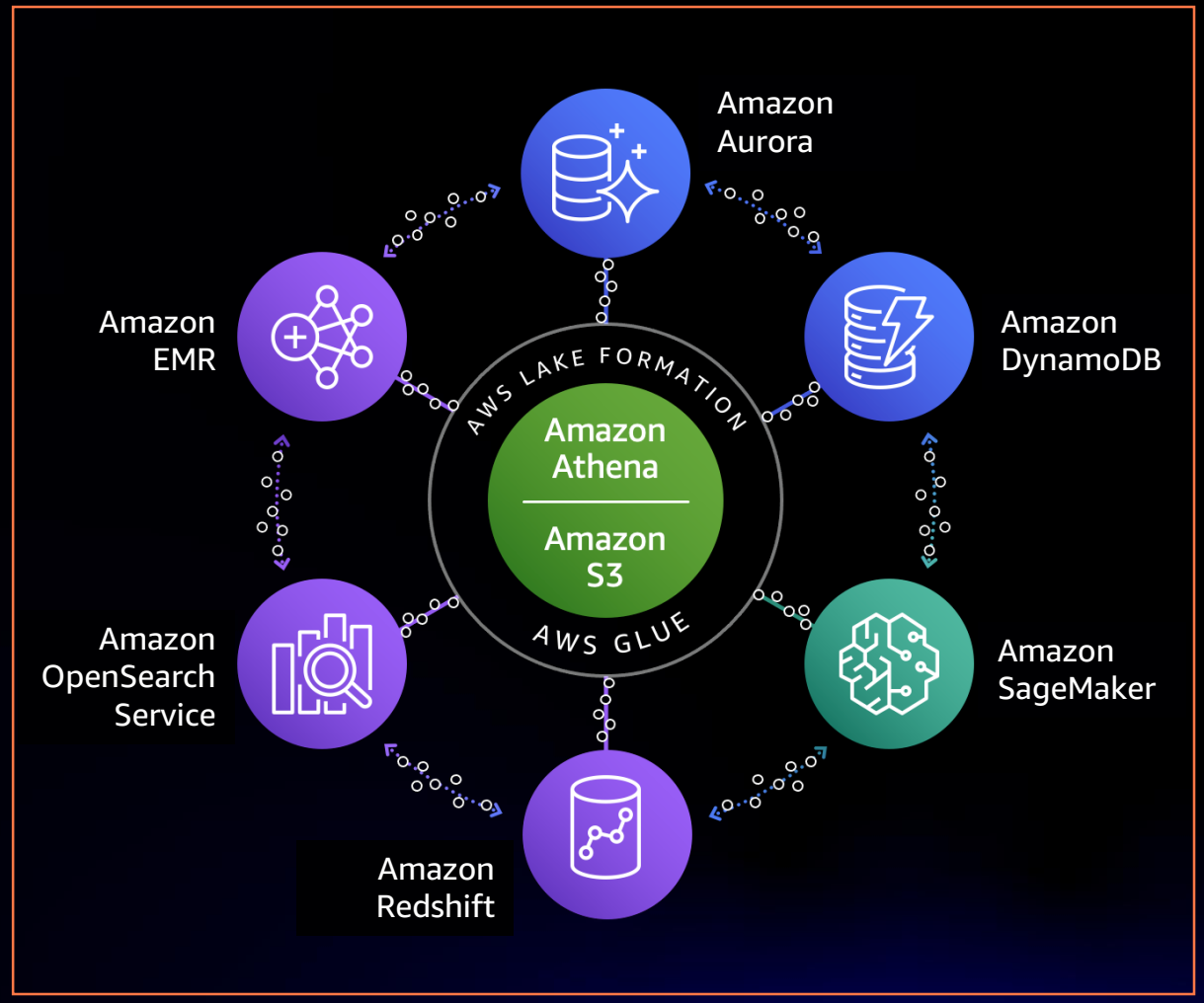
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is a lake house?

With a lake house architecture on AWS, customers store domain-specific data in a data lake and use a ring of purpose-built data services around the lake, allowing them to make decisions with speed and agility, at a scale and price-performance unmatched in the market

The lake house enables customers to:

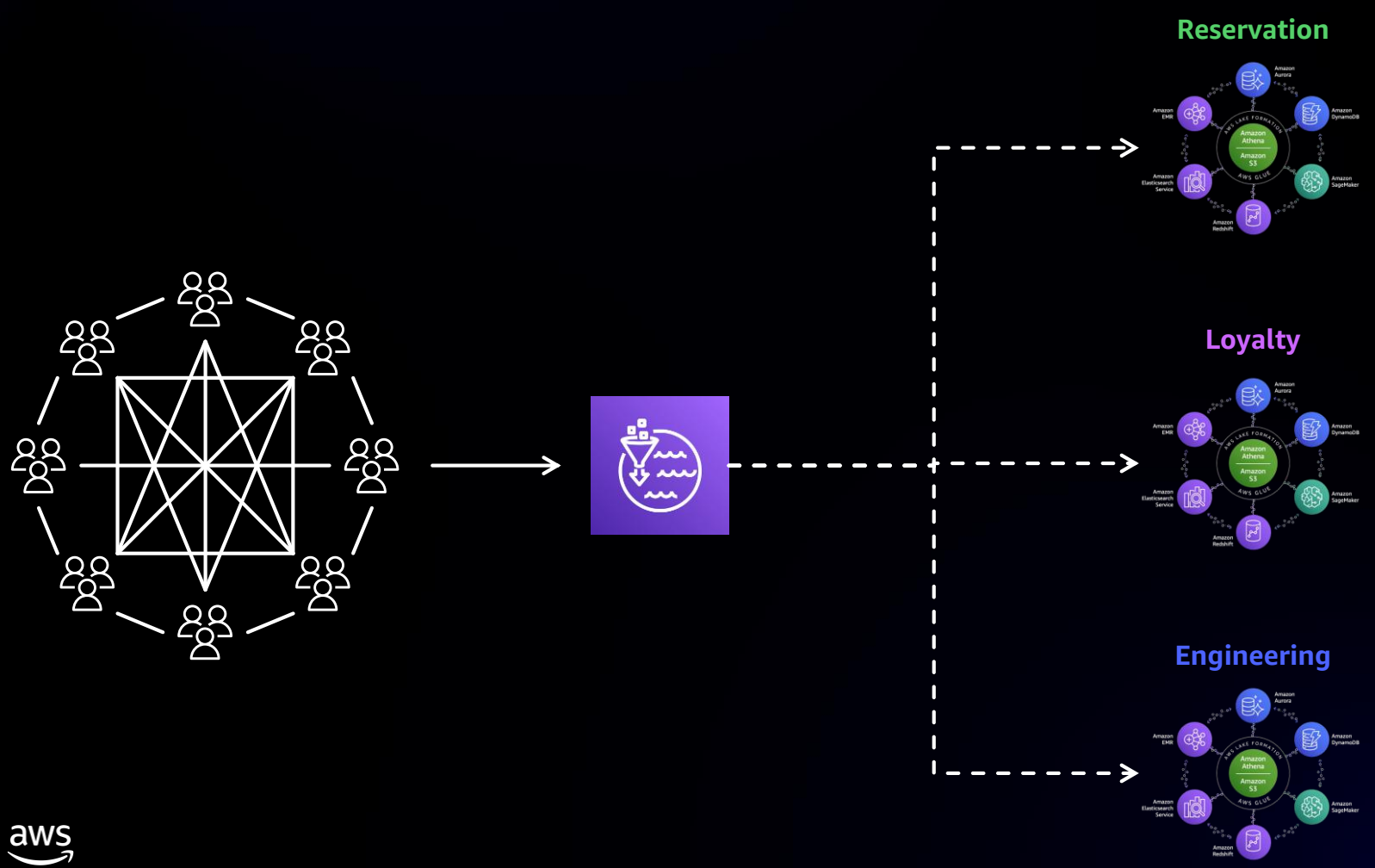
- Deliver a collection of domain-specific data products
- Rapidly build scalable data lakes
- Use a broad and deep collection of purpose-built data services
- Ensure compliance via a unified way to secure, monitor, catalog, and manage access to your data
- Scale your systems at a low cost without compromising performance



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

Data mesh is a collection of domain-specific lake houses

ENTERPRISE VIEW ACROSS DISTRIBUTED DATA STORE



- Decentralized data domains
- Enterprise data products
- Federated governance
- Self-service sharing

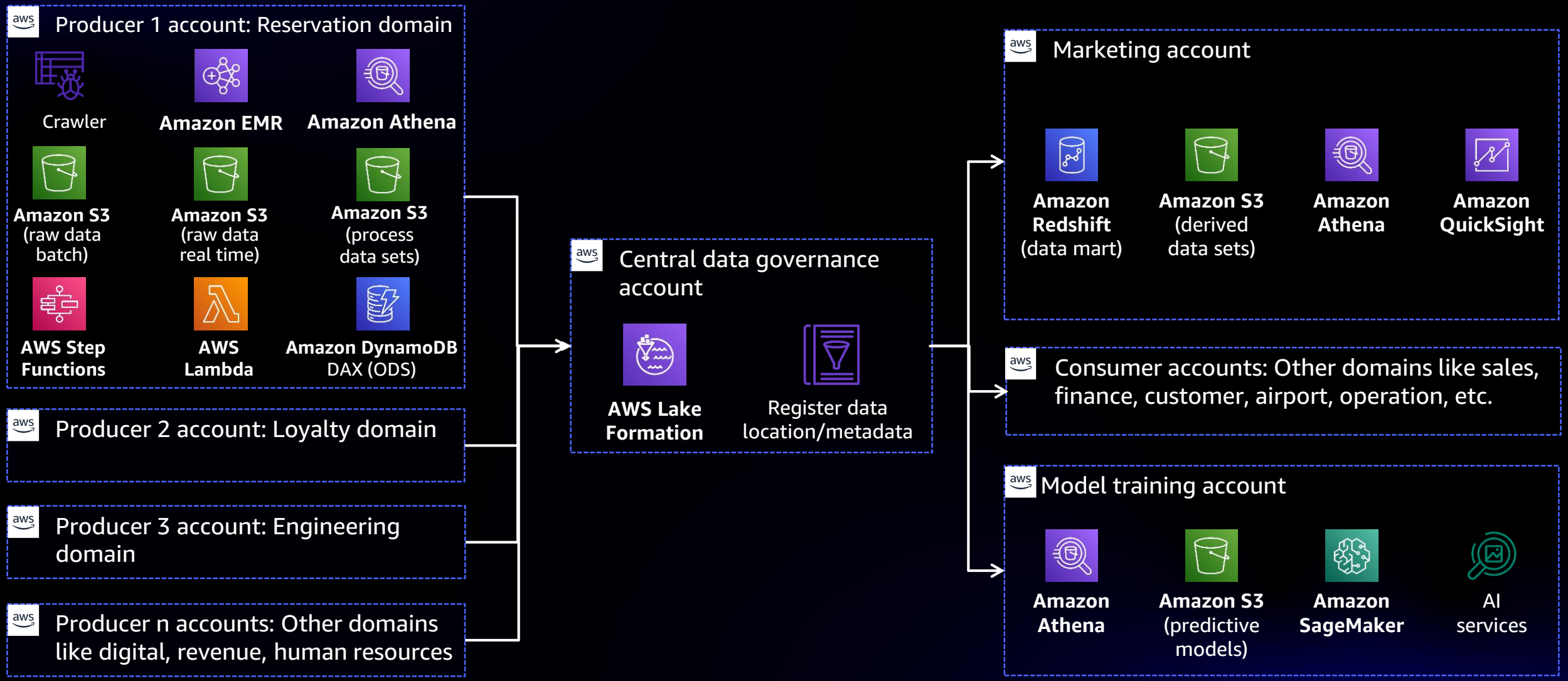
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh organizational principles

1. Allow data producers to create the data model and contract for the data they own
2. Provide efficient means to safely find, publish, and securely exchange data
3. Create a single model for identity and data classification
4. Build logging, monitoring, and auditing into this environment
5. Mandate this system's use for data sharing using resource share
6. Set organizational goals on adoption and create an inspection mechanism
7. Publicize success of data sharing between teams to create an entitlement policy

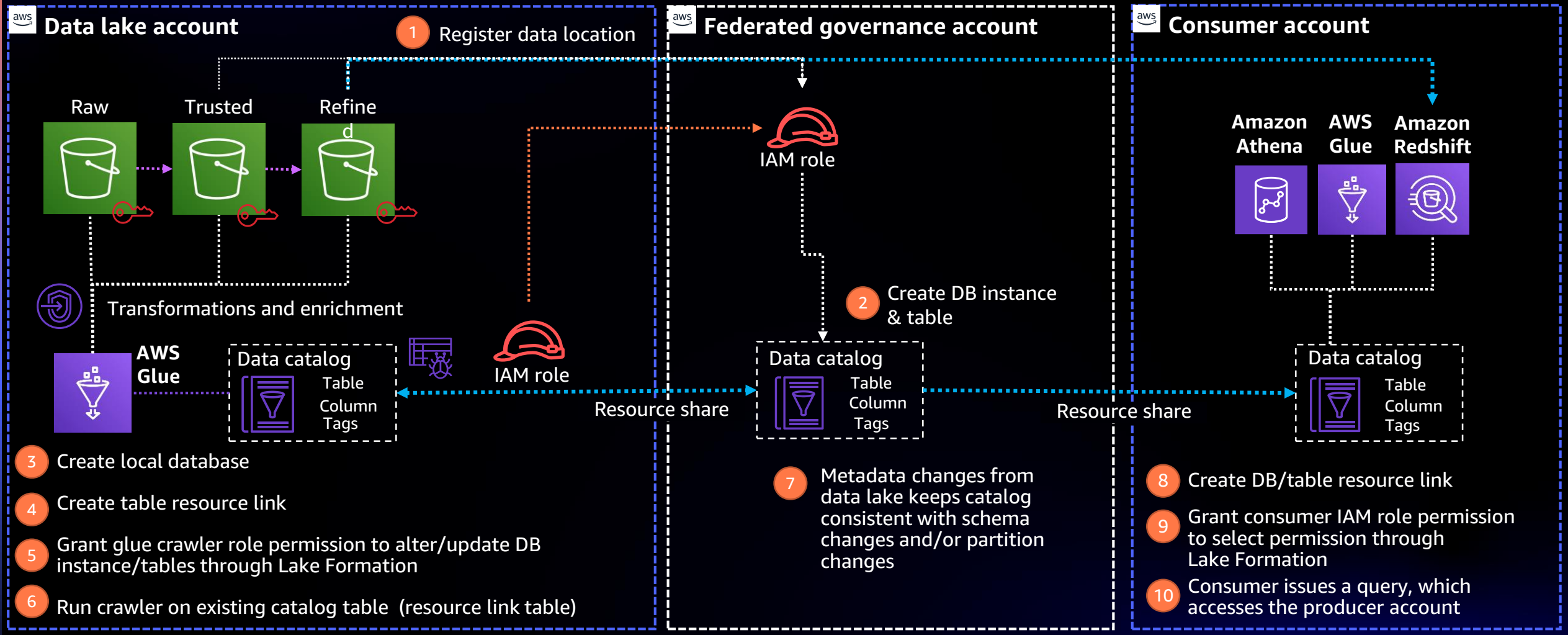
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data mesh: Core concepts in AWS



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

Data mesh build in AWS environment: End-to-end flow



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

Additional resources

THESE ARE ADDITIONAL RESOURCES FOR DATA MESH IN AWS

- Implementing travel and hospitality data mesh
  Build Data platform using domain owned design
  https://d1.awsstatic.com/architecture-diagrams/ArchitectureDiagrams/implementing-travel-and-hospitality-data-mesh-ra.pdf
- Data mesh using AWS Lake Formation and AWS Glue
  Implement a data mesh using AWS native services
  https://aws.amazon.com/blogs/big-data/design-a-data-mesh-architecture-using-aws-lake-formation-and-aws-glue/
- How JPMorgan Chase built a data mesh architecture
  How data mesh architecture supports data product strategy
  https://aws.amazon.com/blogs/big-data/how-jpmorgan-chase-built-a-data-mesh-architecture-to-drive-significant-value-to-enhance-their-enterprise-data-platform/

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

Thank you!

Paul Armstrong
armpaul@amazon.co.uk

Bib Shukla
bibshukl@amazon.co.uk

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