

AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

ANT310-R

Build a data mesh with AWS Lake Formation and AWS Glue

Dzenan Softic

Sr. Solutions Architect, Startups
AWS

David Greenshtein

Sr. Specialist Solutions Architect, Analytics
AWS



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

Agenda

Intro
Presentation
~15min

Self-Paced
Building
~1h 45min

Common data lake challenges we hear from our customers

"Finding the data I need is too difficult"

"I just want to get access to the data I need"

"Difficult to meet all requirements across differing business units"

"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"

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

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

"Current data architecture is complex and monolithic and slow to change"

"If I share data, I've lost control"

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

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

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



Centralized solution challenges

Central data team becomes the **bottleneck**

Fail to **scale** data consumers

Unable to **discover and consume** the data

No **central data governance**

Lack of data **auditability**



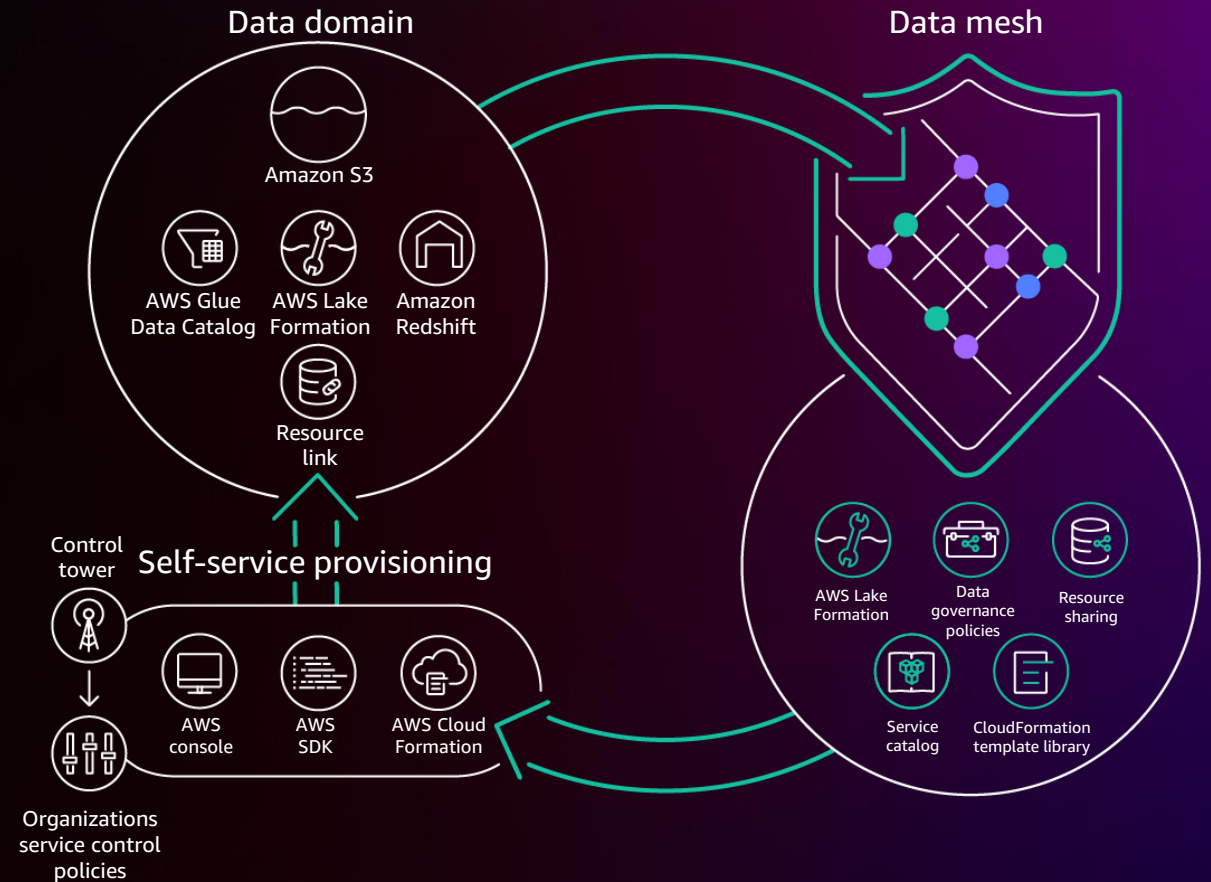
Why data mesh?

Treats existing data platforms as **independent** domains

Improves **data governance**

Establishes a **central** mechanism for **data discovery**

Provides **self-service data sharing** features



Data mesh – Four core principles



Data owner

Data domain ownership



Data steward

Federated computational governance



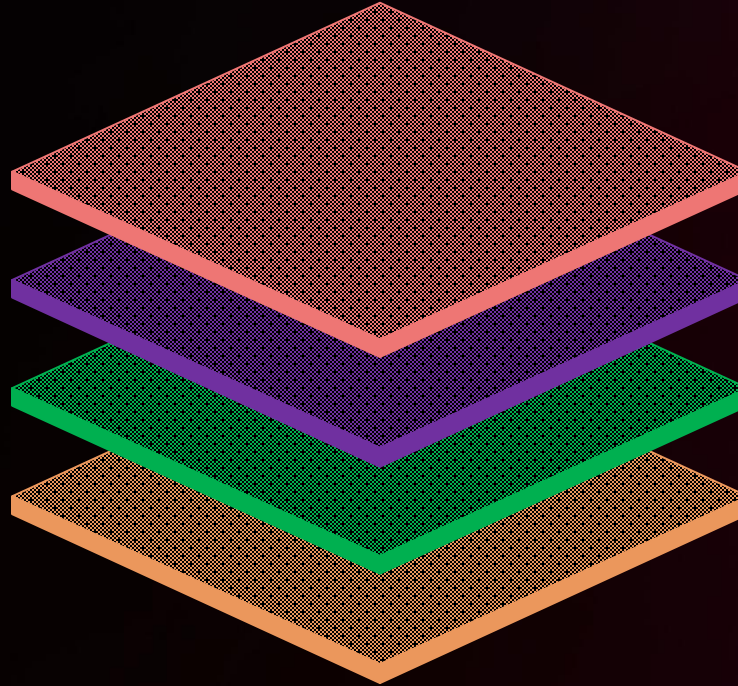
Data engineer

Data as a product



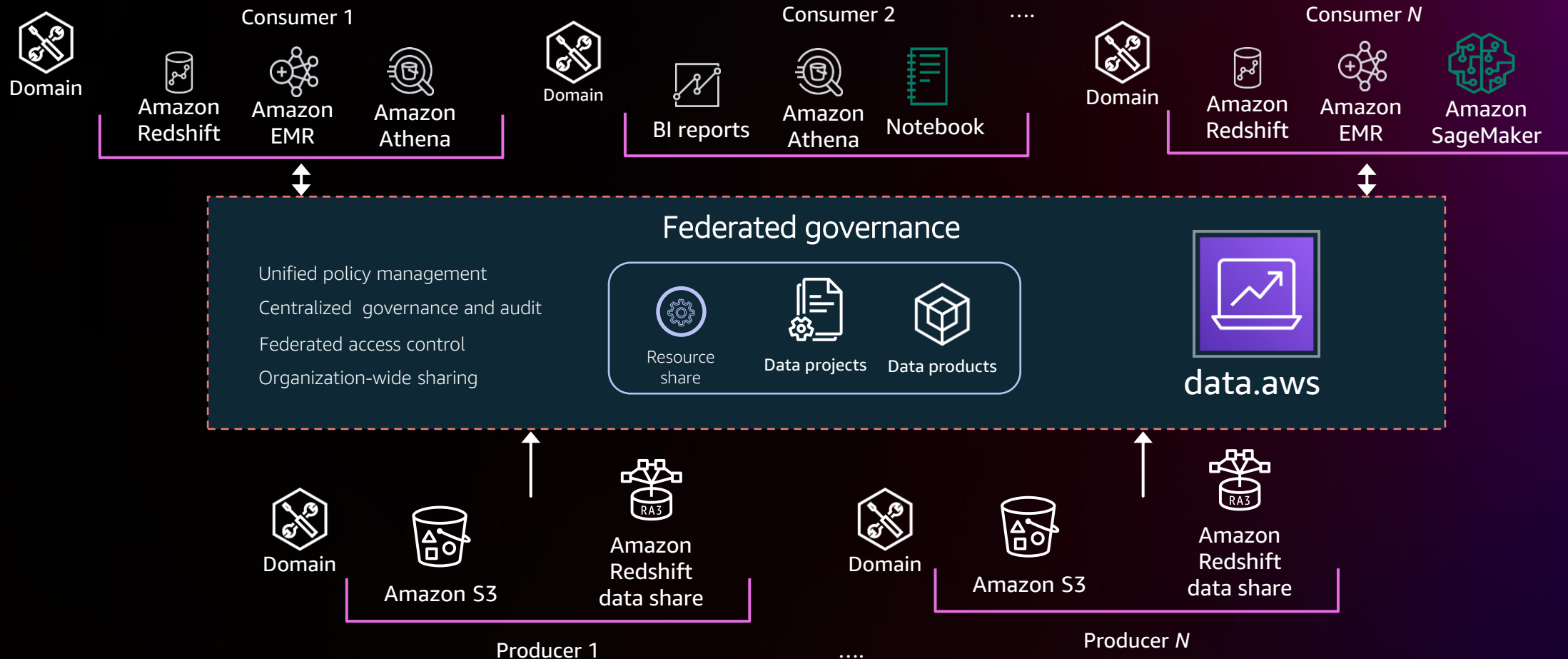
Data consumer

Self-serve sharing



Data mesh architecture

Decentralized, light-weight, federated governance across domain-oriented data systems to drive governed sharing



Related session

ANT336

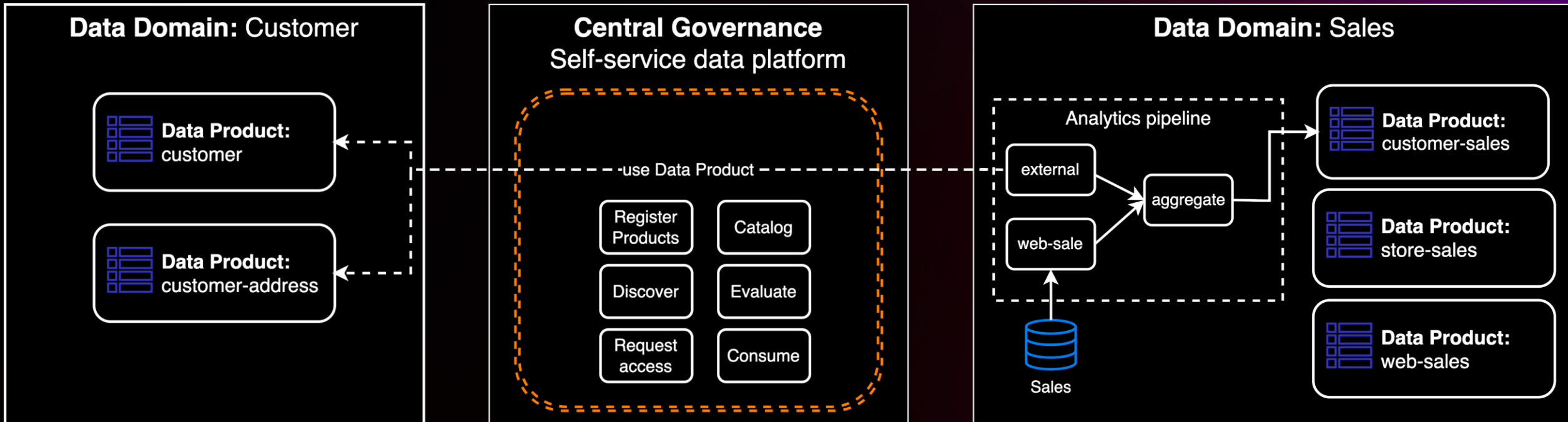
Building data mesh architectures on AWS



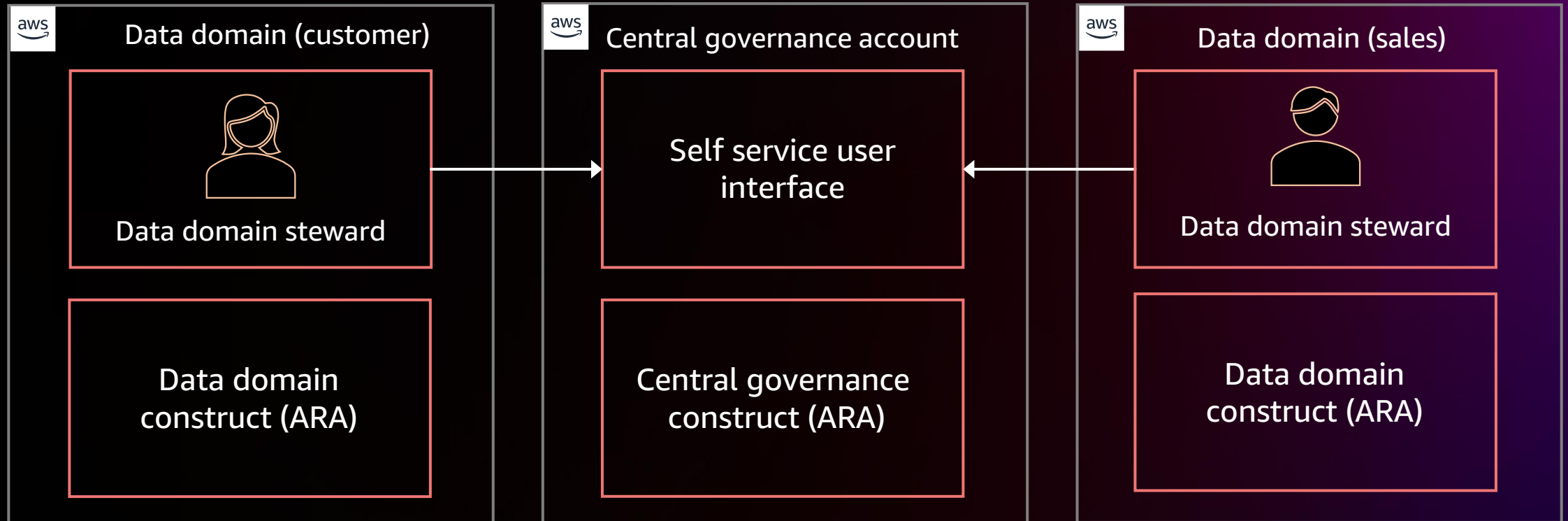
Today's workshop



What we will build today



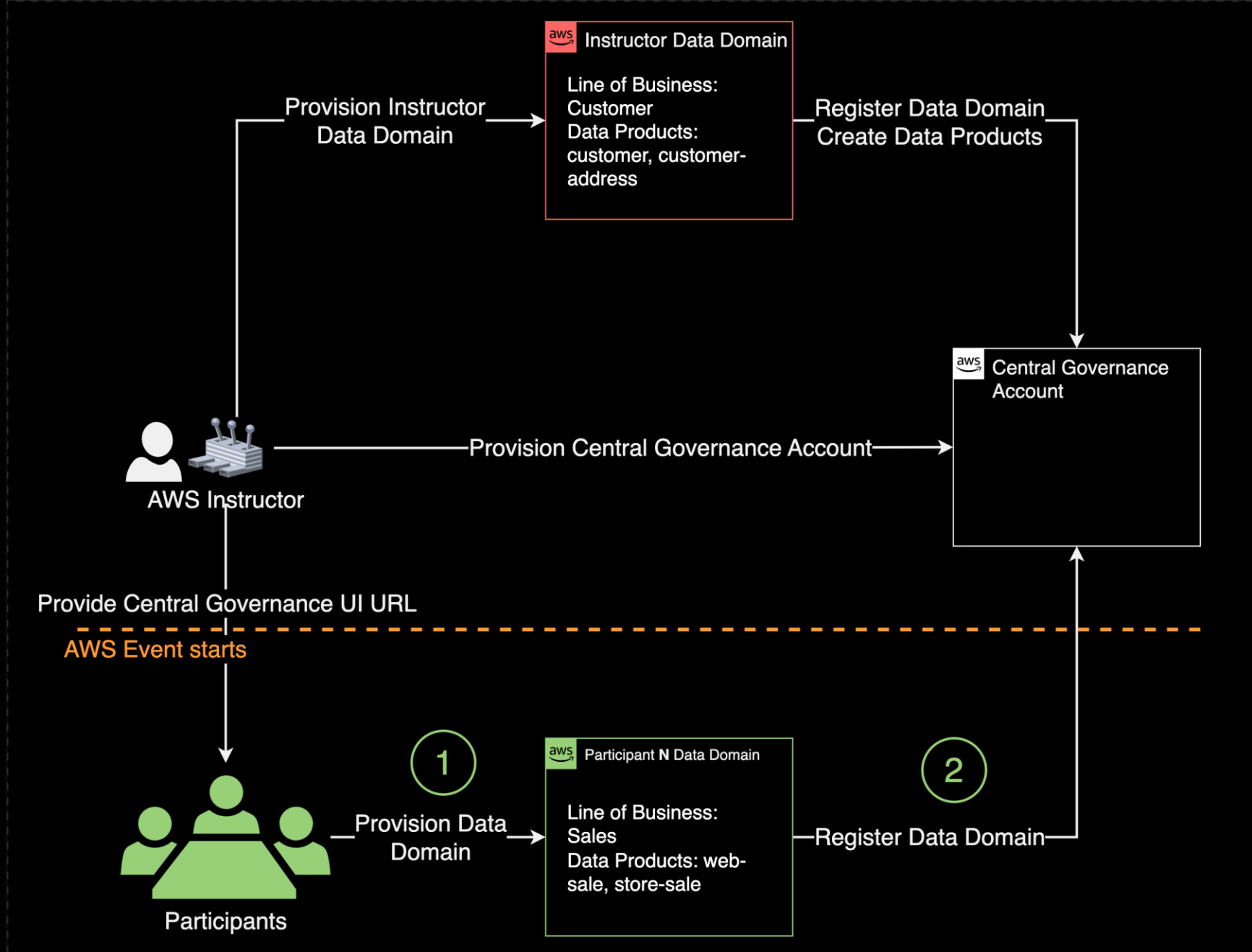
Main building blocks



<https://github.com/aws-samples/aws-analytics-reference-architecture>

Workshop setup





Workshop access

Join link: <https://s12d.com/ANT310>

Getting started with this workshop

- As a participant, you will have access to an AWS account with any optional pre-provisioned infrastructure and IAM policies needed to complete this workshop
- The AWS account will only be available for the duration of this workshop; you will lose access to the account thereafter
- The optional pre-provisioned infrastructure will be deployed to a specific Region; check your workshop content to determine whether other Regions will be used
- Be sure to review the terms and conditions of the event; do not upload any personal or confidential information in the account

Step 1: Sign-in on your laptop



Sign in

Choose a preferred sign-in method

Email one-time password (OTP)

Enter your personal or corporate email to receive a one-time password

Login with Amazon

Login with your Amazon.com retail account

Amazon employee

Login with your Amazon Corporate account. Only for Amazon Employees.

Step 2: Review terms and join event

Terms and Conditions

Read and accept before joining the event

1. By using AWS Workshop Studio for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivate works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through Event Engine and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of AWS Workshop Studio will comply with these terms and all applicable laws, and your access to AWS Workshop Studio will immediately and automatically terminate if you do not comply with any of these terms or conditions.

☒ I agree with the Terms and Conditions

CancelPreviousJoin event

Step 3: Get started with the workshop

Workshop

Get started >

| | | | |
|-------------------------|------------------|---|-----------|
| Title | Complexity level | AWS services | Topics |
| Data Mesh Immersion Day | 300 | AWS Glue, AWS Lake Formation, AWS Step Functions | Analytics |

Description

Enable customers with hands-on experience to build a data mesh architecture on AWS step-by-step from scratch. **Instructor module: <https://gitlab.aws.dev/softicd/instructor-module-for-build-a-data-mesh-on-aws-workshop>**

ANT310

Have fun building data mesh on AWS!

Join link: <https://s12d.com/ANT310>

Workshop details: <https://s12d.com/ANT310-teams>



Resources



Workshop link:

<https://s12d.com/data-mesh-workshop>



Blog post:

<https://s12d.com/data-mesh-blog>

Got any feedback?

Evaluate this session using the **AWS Events mobile app**:
tap on More → My Evaluations → **ANT310**

or reach out to the speakers directly!

Dzenan Softic

✉ softicd@amazon.com

in linkedin.com/in/dzenansoftic

David Greenshtein

✉ dgreensh@amazon.com

in linkedin.com/in/davidgreenshtein

Thank you!



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

