AWS re:Invent

ADM401

Analyzing Amazon DSP data in an AWS data lake

David Beck

Programmatic Consultant

Amazon Advertising

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws

Agenda

Introduction and overview

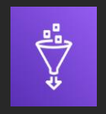
Demo



Amazon Simple Storage Service (Amazon S3)



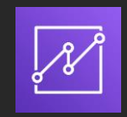
AWS Lambda



AWS Glue



Amazon Athena



Amazon QuickSight

Introduction and overview

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Introduction & overview

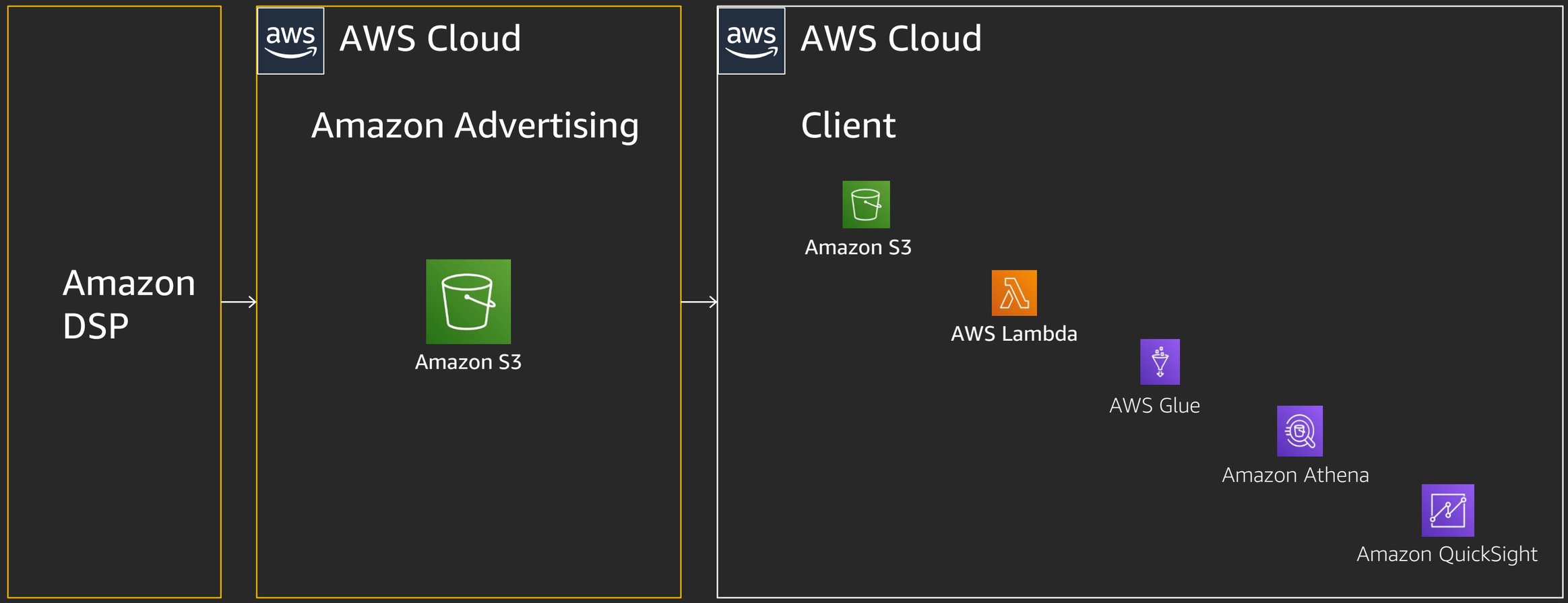


amazonadvertising

David Beck
Programmatic Consultant
Amazon Advertising

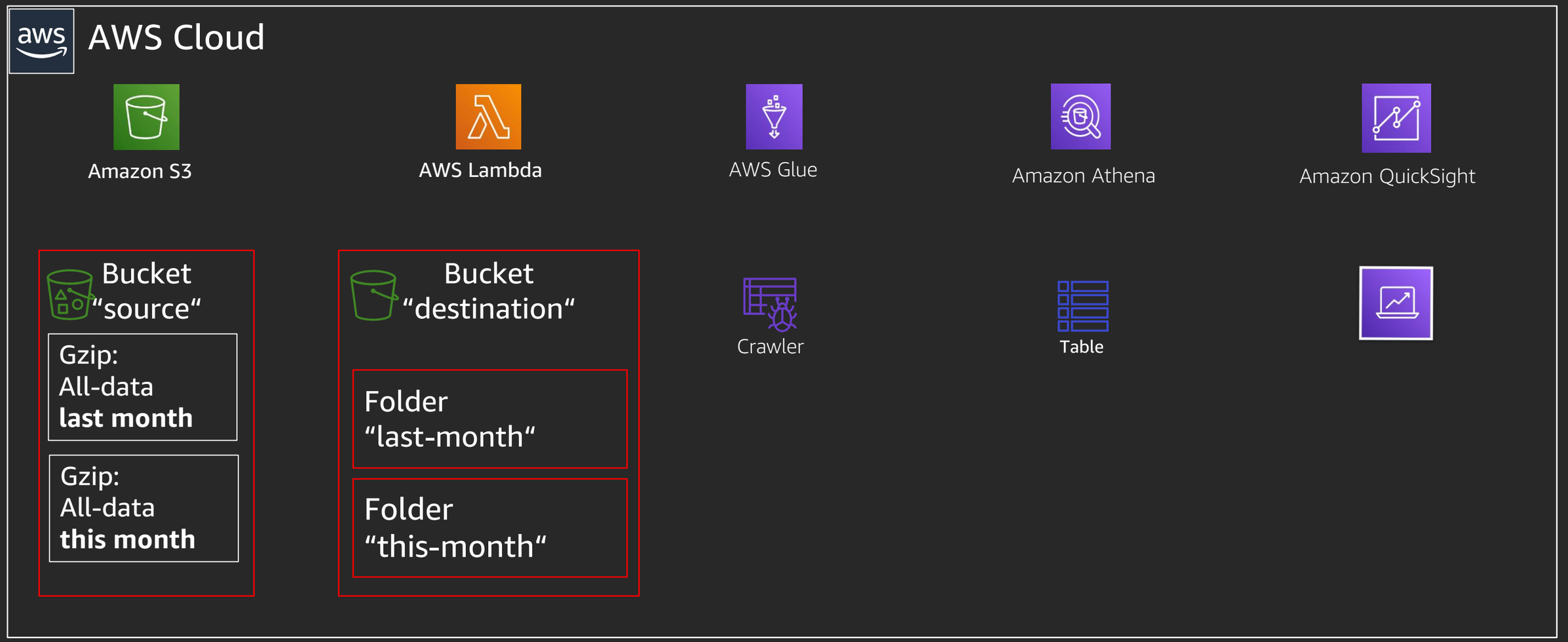
Introduction & overview

Client scenario



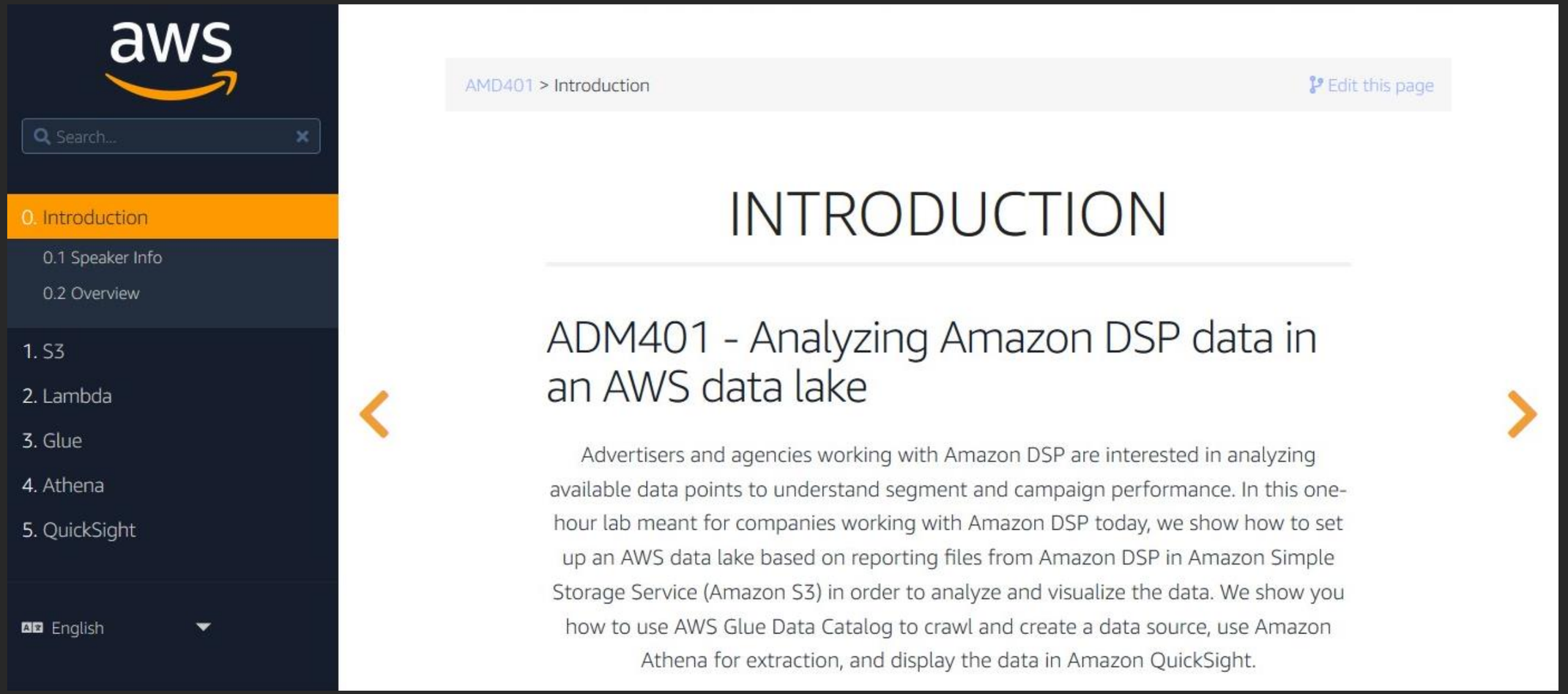
Introduction & overview

Demo scenario



Introduction & overview

Training material



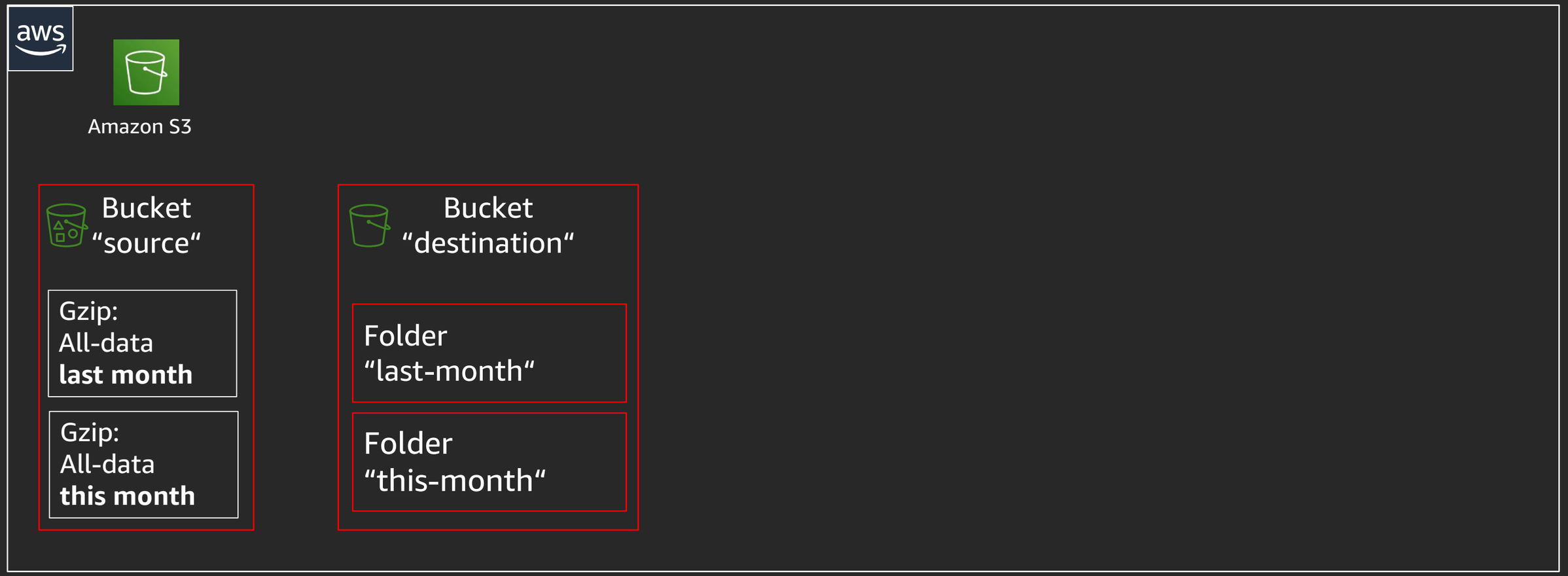
Demo

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Demo: Amazon S3

Create initial folder structure in Amazon S3

Please use US east (N. Virginia)" (us-east-1)



Demo: Lambda

Set up lambda function

Run function ad-hoc in order to copy and delete data



Demo: Lambda

```
[…]

# 1. identify the reports keys

for obj in source_bucket.objects.filter(Prefix=source_prefix):

        if "all-data_all-adv_this-month-daily" in obj.key:

                this_month_report_key = obj.key

                this_month_report_name = this_month_report_key.rsplit('/', 1)[1]

        elif "all-data_all-adv_last-month-daily" in obj.key:

                last_month_report_key = obj.key

                last_month_report_name = last_month_report_key.rsplit('/', 1)[1]

[…]
```

Demo: Lambda

```
[…]
# 2. copy the reports to the destination bucket
this_month_report = { 'Bucket' : source_bucket_name, 'Key': this_month_report_key}
        this_month_dest_key = dest_prefix + 'this-month/' + this_month_report_name
        dest_bucket.copy(this_month_report, this_month_dest_key)

last_month_report = { 'Bucket' : source_bucket_name, 'Key': last_month_report_key}
        last_month_dest_key = dest_prefix + 'last-month/' + last_month_report_name
        dest_bucket.copy(last_month_report, last_month_dest_key)
[…]
```

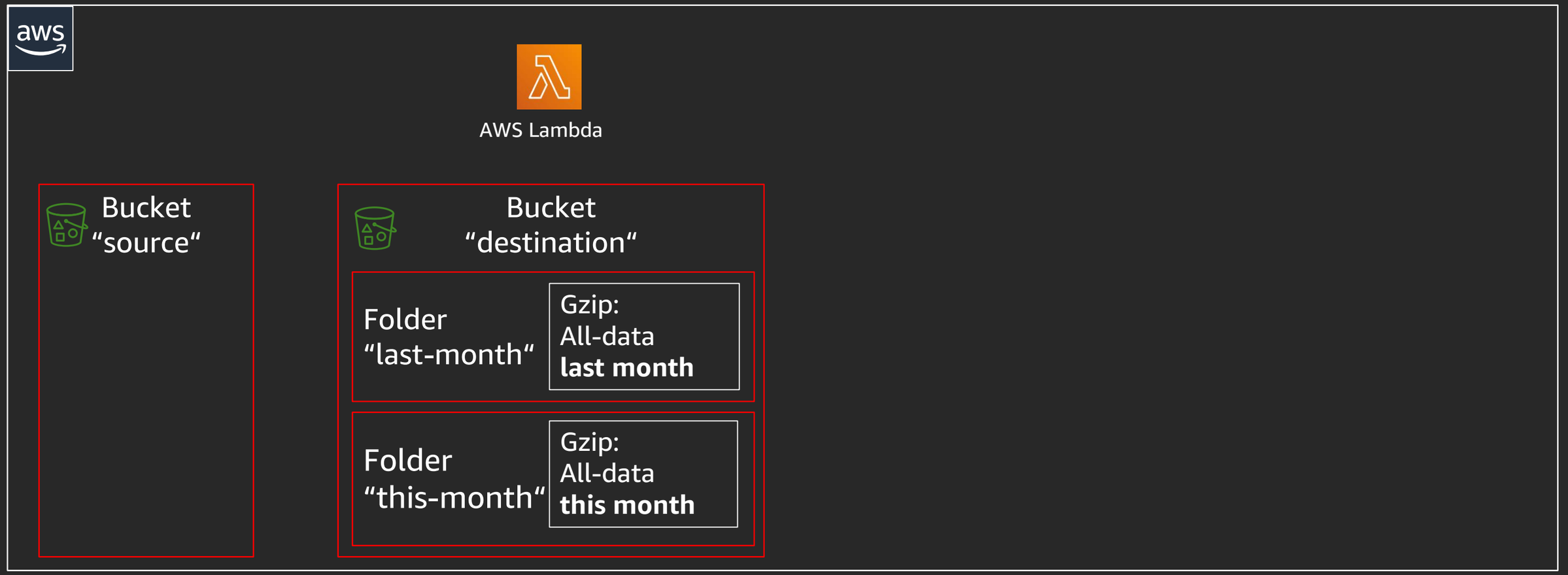
Demo: Lambda

```
[…]
# 3. delete the reports from the source bucket if they were copied successfully
for obj in dest_bucket.objects.filter(Prefix=dest_prefix):
        if this_month_report_name in obj.key:
                client.delete_object(Bucket=source_bucket_name, Key=this_month_report_key)
        elif last_month_report_name in obj.key:
                 client.delete_object(Bucket=source_bucket_name, Key=last_month_report_key)
[…]
```

Demo: Lambda

Set up lambda function

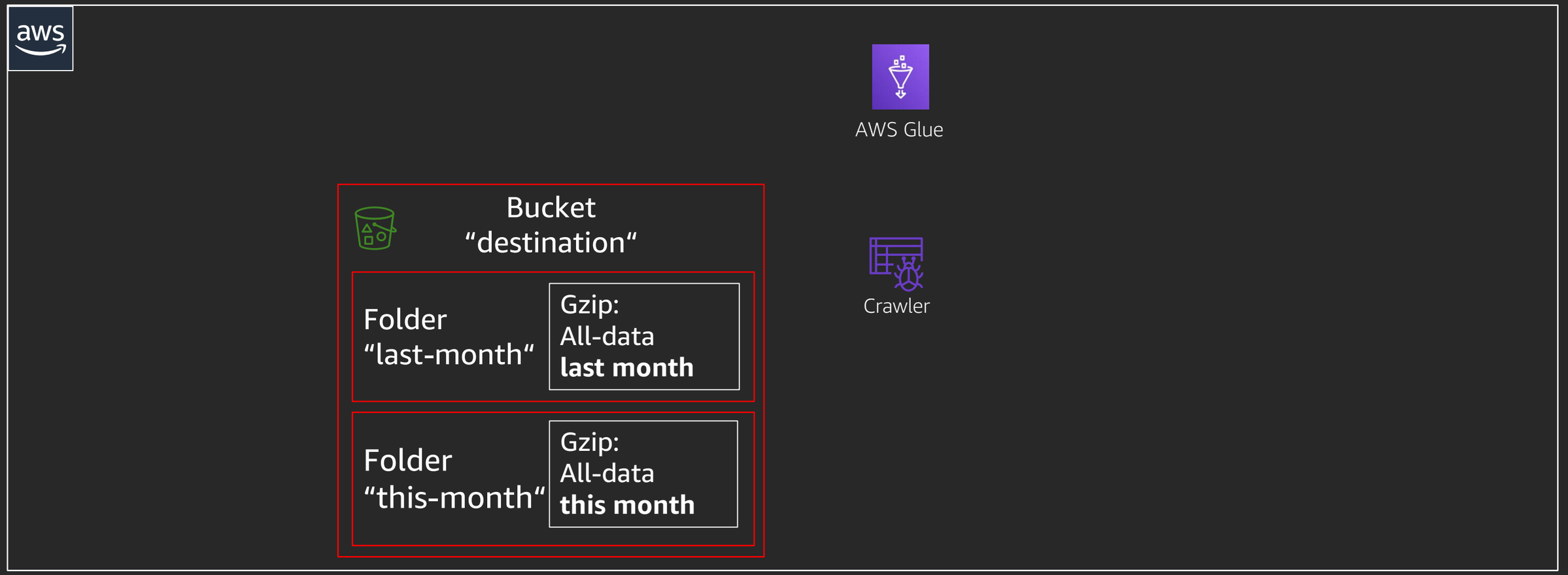
Run function ad-hoc in order to copy and delete data



Demo: AWS Glue

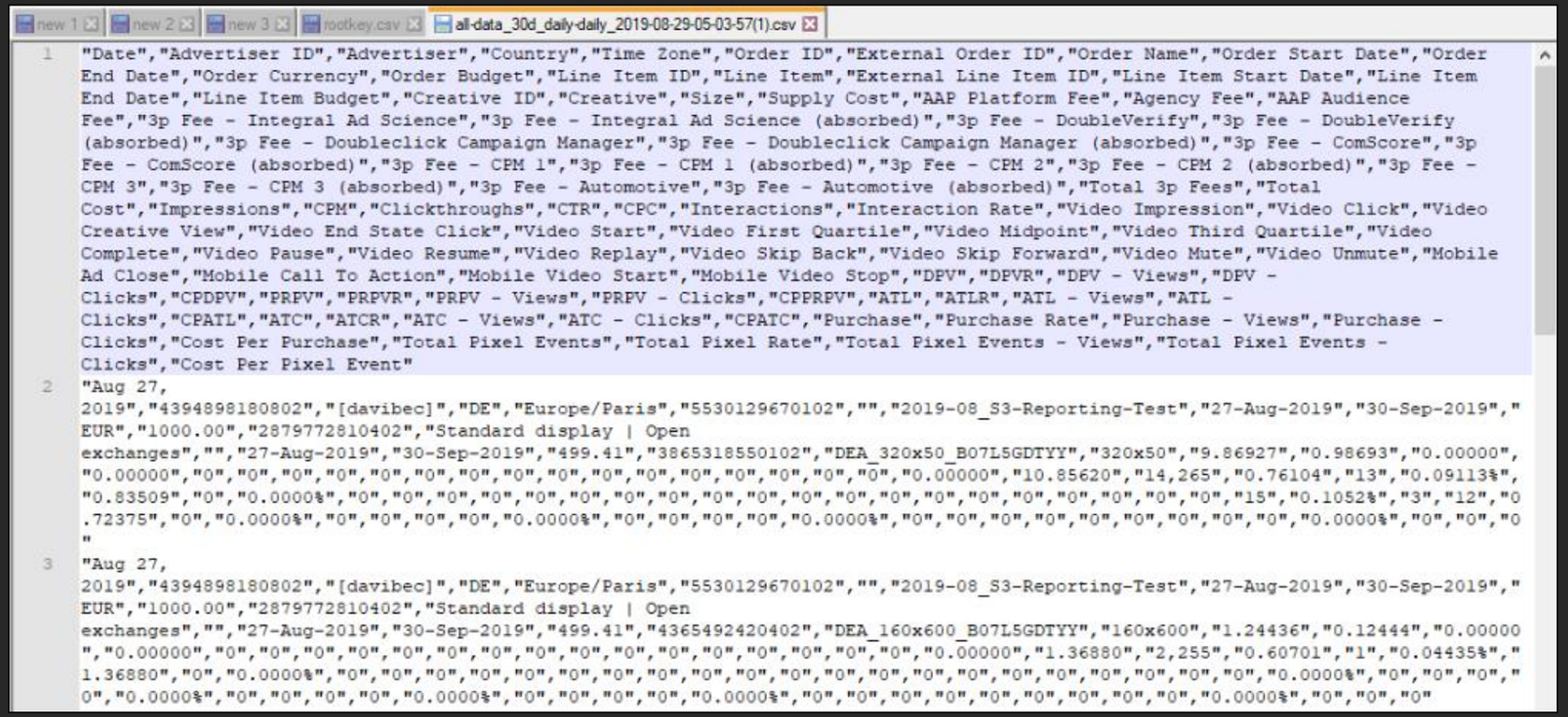
Create table via crawler with classifier

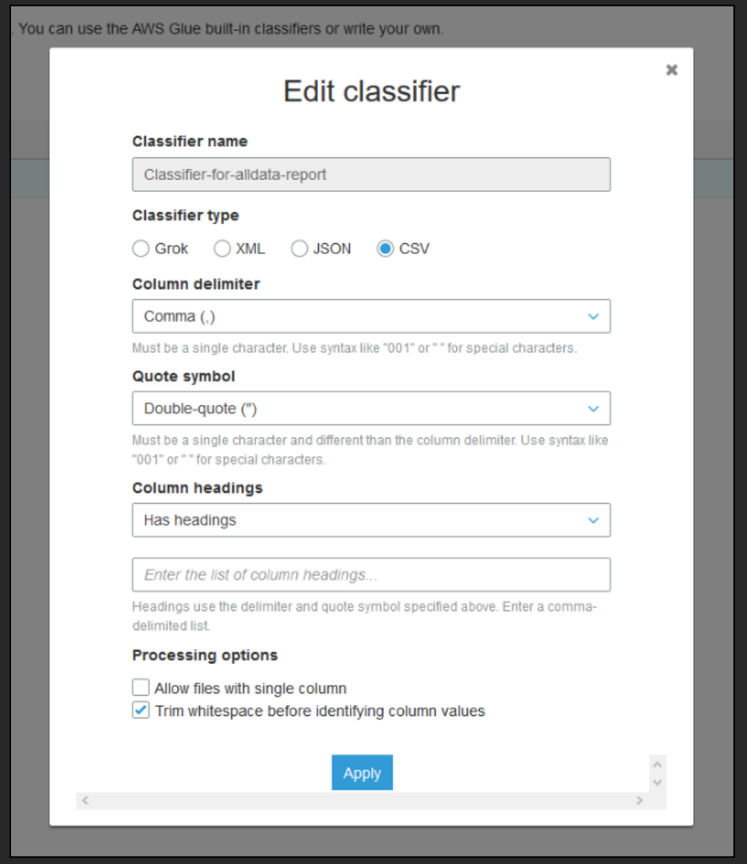
Run ad-hoc in order to create new data source



Demo: AWS Glue

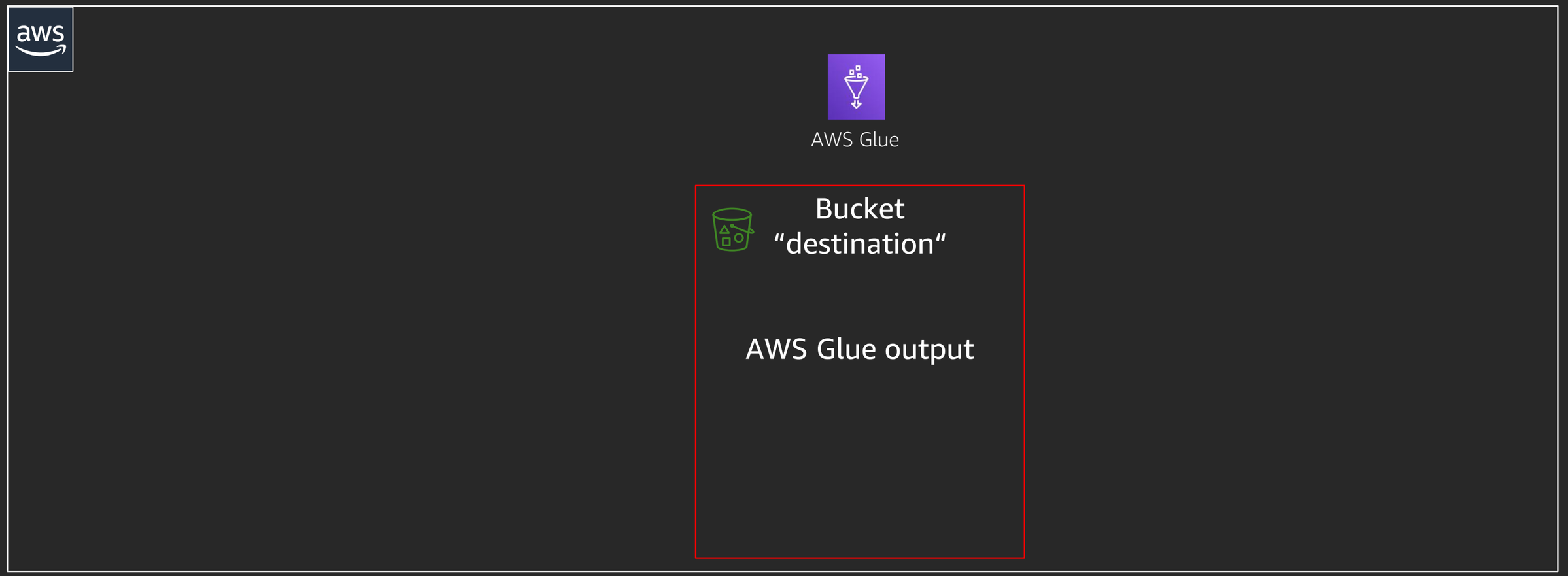
The file setup requires a dedicated classifier:





Demo: AWS Glue

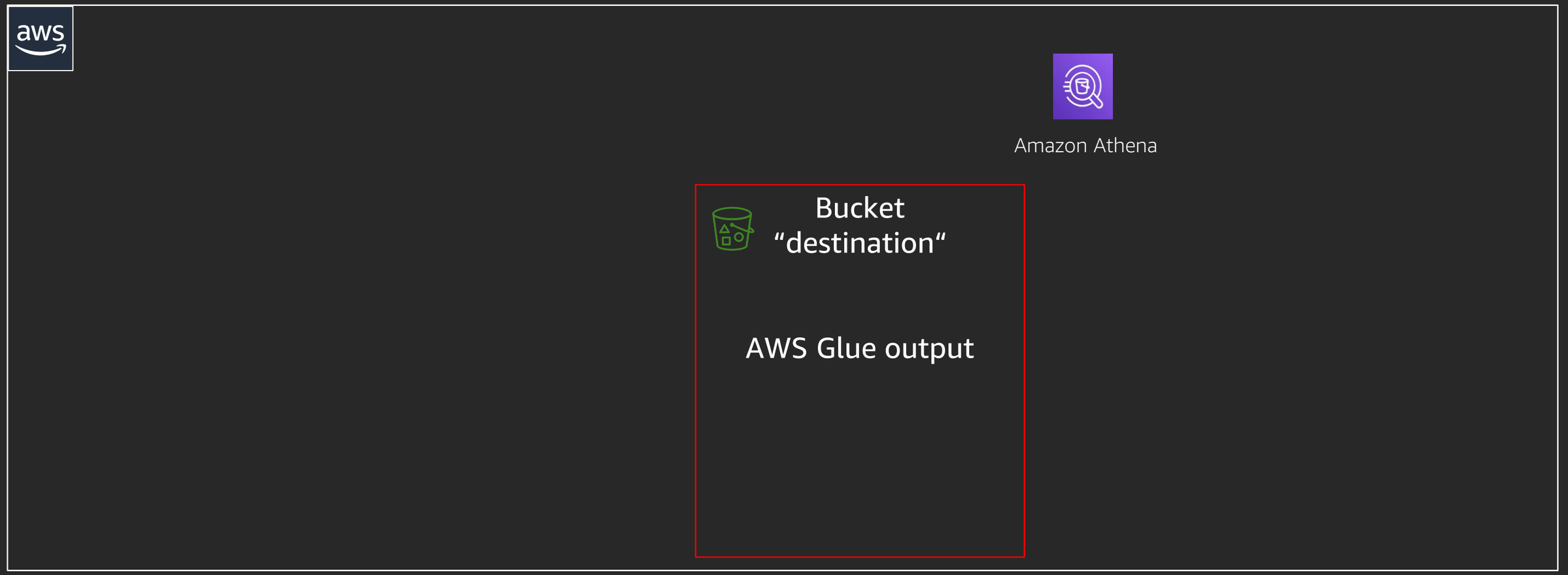
A new data source was created



Demo: Amazon Athena

Query relevant data via Amazon Athena

Create a “view” on your data



Demo: Amazon Athena

```
var pd SELECT

date,

'advertiser id' as advertiser_ID,

advertiser,

country,

'order id' as order_ID,

'order name' as order_name,

'linte item id' as line_item_ID,

'creative id' as creative_ID,

'supply cost' as supply_cost,

'aap platform fee' as aap_platform_fee,

'total cost' as total_cost,

impressions,

clickthroughs,

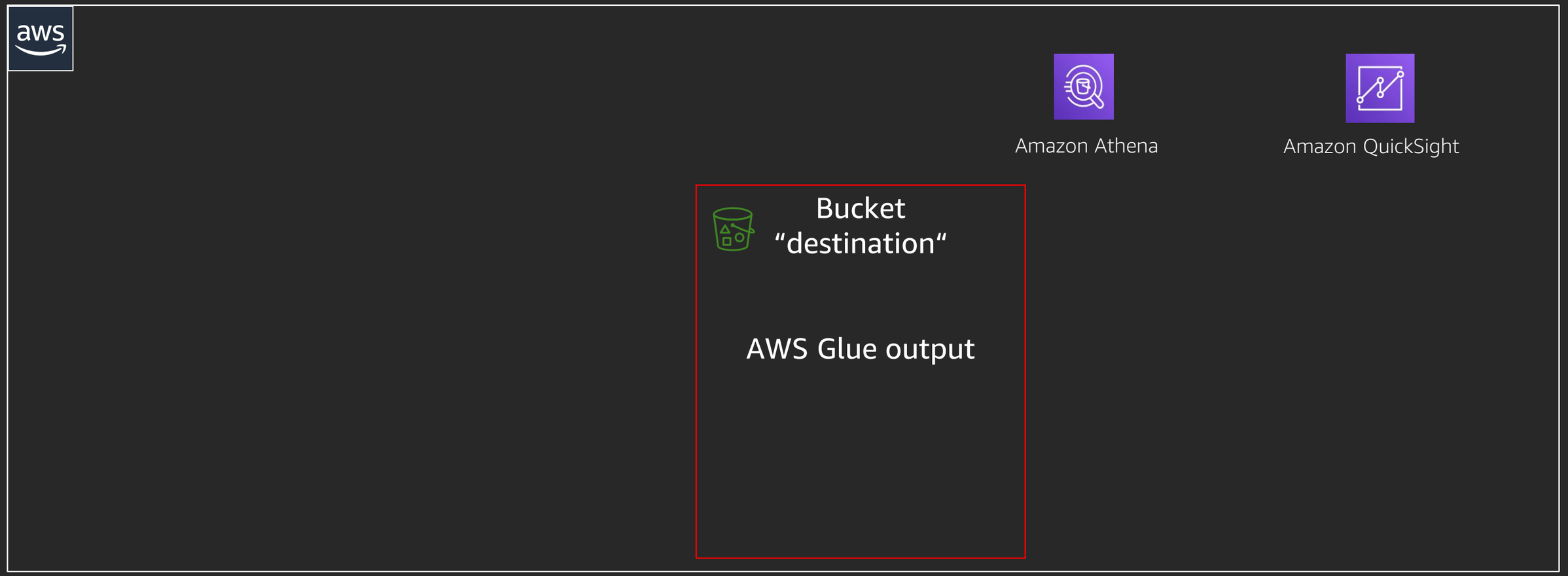
dpv


FROM "2019-08-29_all-
data"."davibec_destination ";
```

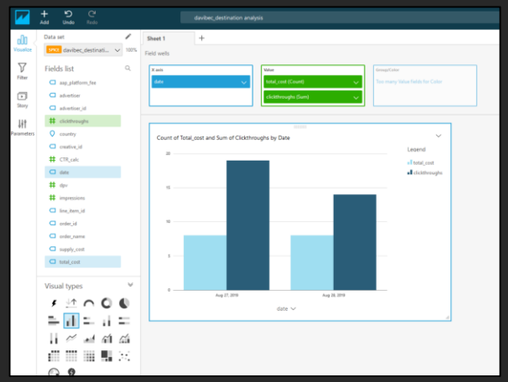
Demo: Amazon QuickSight

Use Amazon Athena as source for Amazon QuickSight

Load and visualize the data



Demo: Amazon QuickSight



Use Amazon Athena as source for Amazon QuickSight

Load and visualize the data

Amazon QuickSight

Thank you!

David Beck

davibec@amazon.com

AWS re:Invent

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.

aws



Please complete the session survey in the mobile app.

© 2019, Amazon Web Services, Inc. or its affiliates. All rights reserved.