STEP ON IT!

Rapid Event-Based AWS Cost Control

With Step Functions

Damien Jones (he/him)

AWS Consultant @ Steamhaus

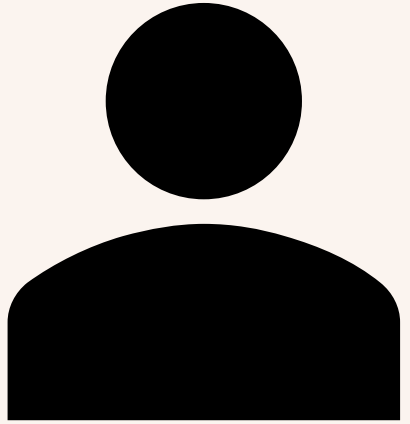
2025-09-18 AWS Cloud Day Dublin

aws

aws
aws

aws
aws
aws

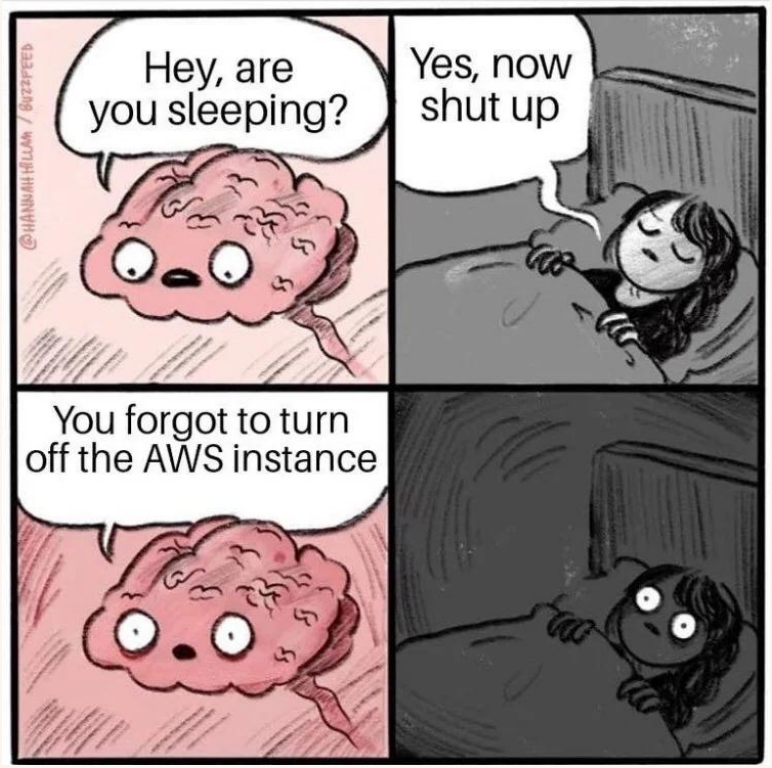
CLIENT AWS ACCOUNT



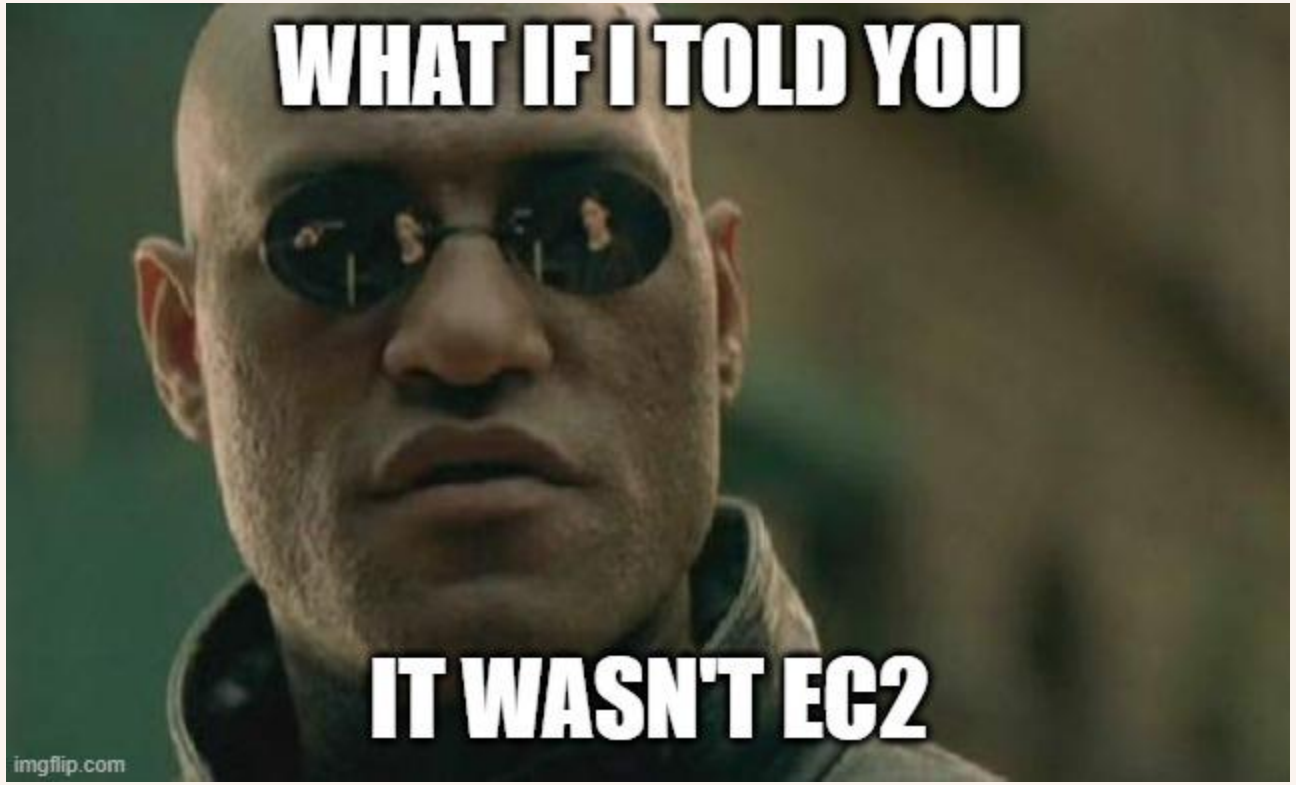
EMPLOYER AWS ACCOUNTS

MY AWS ACCOUNTS

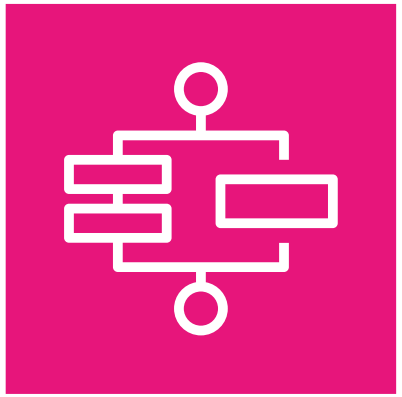
LOL ROFL LMAO ETC



NOPE



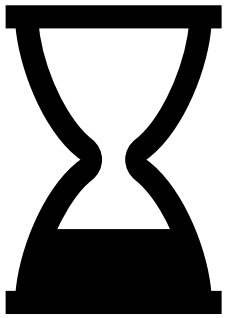
AWS STEP FUNCTIONS



STEP FUNCTION FEATURES

ORCHESTRATOR

INTEGRATOR



OBSERVER

BUT FIRST

Hi.

Steamhaus

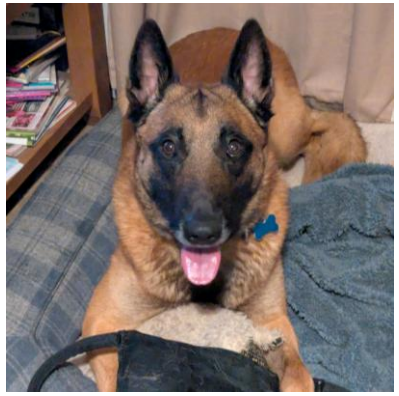
AWS Consultant

Steamhaus UK

Content Creator

amazonwebshark.com

aws certified



AWS Certified

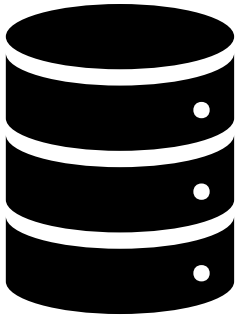
x5

Dog Dad

to this noodle

AWS User Group Leader

Lancashire UK



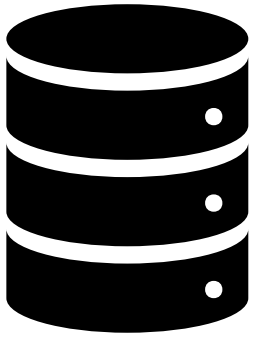
AWS Community Builder

Data Category

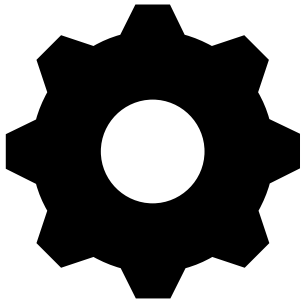
DATA INGESTION PIPELINE

SAVES MONEY ON...

STORAGE



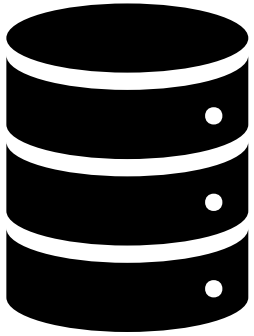
COMPUTE



THIRD PARTY APIS

DATA SOURCE

DATABASE

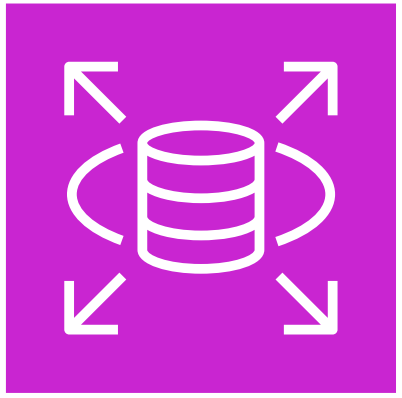


API

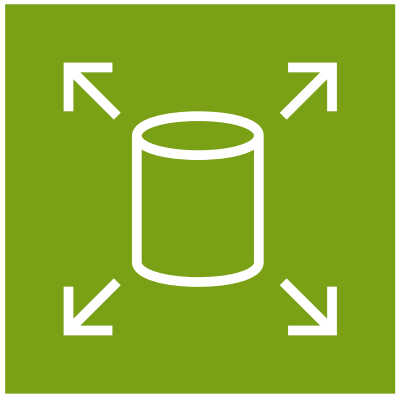
DATA TARGET

S3 BUCKET

RDS
DATABASE

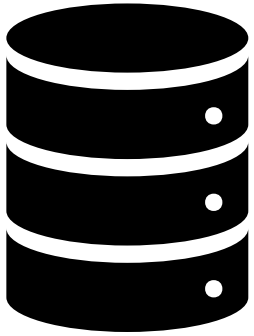


EBS VOLUME



INGEST DATA FROM SOURCE TO TARGET

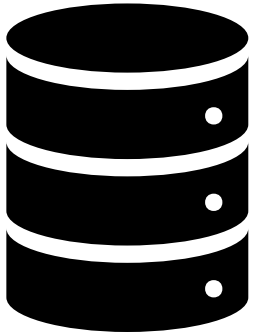
SOURCE



TARGET

DATA INGESTOR

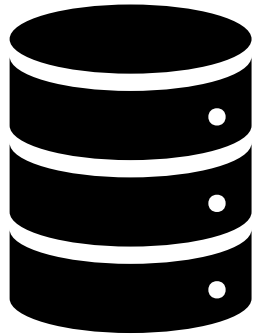
SOURCE



DATA INGESTOR

TARGET

AWS LAMBDA FUNCTION



API / DATABASE

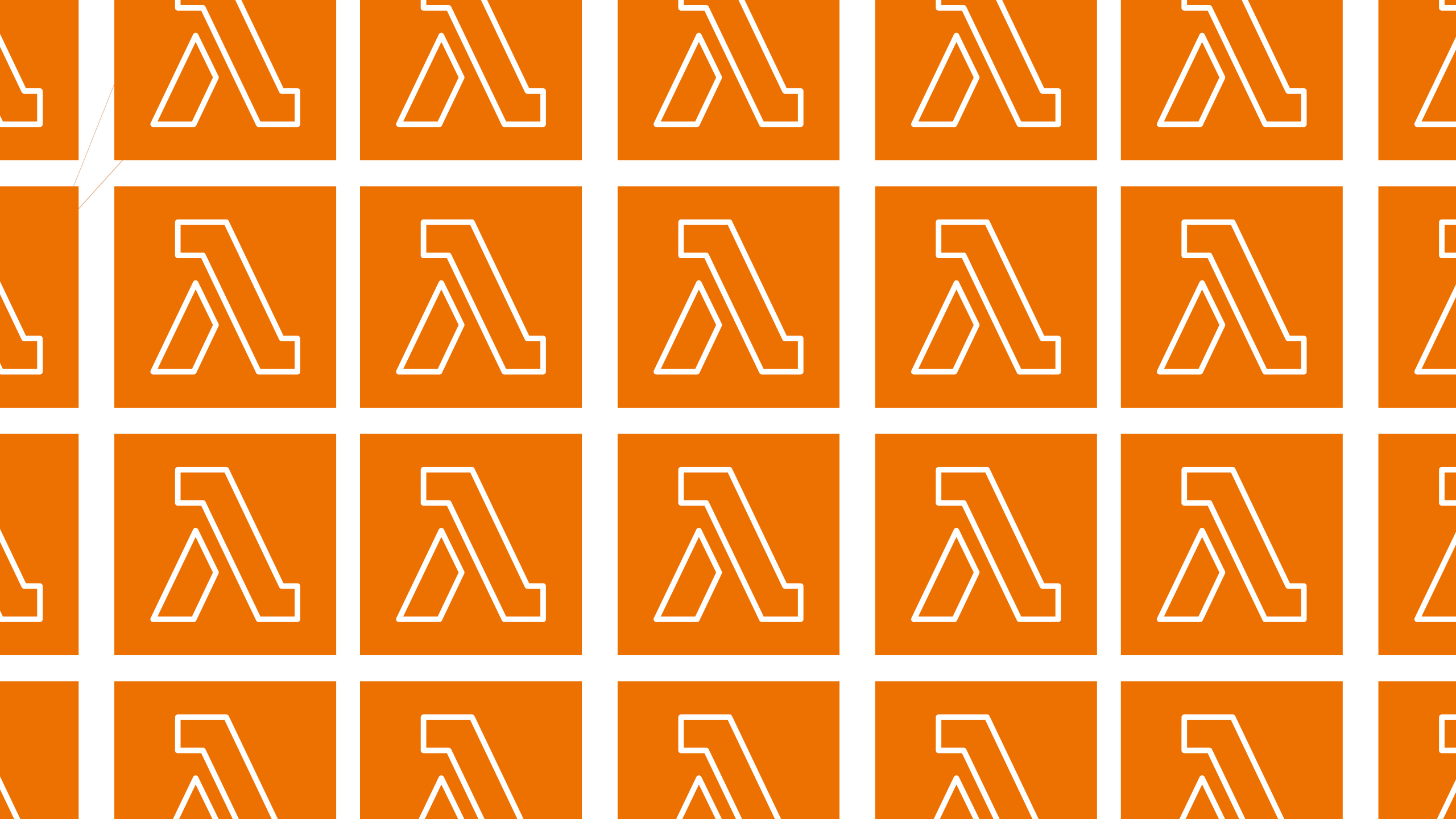
AMAZON S3 BUCKET

AWS LAMBDA

FUNCTION

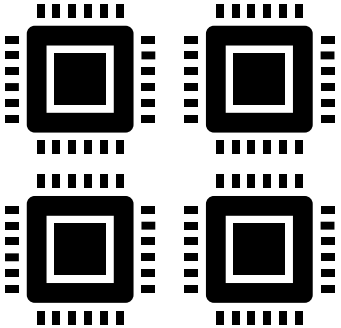
AWS LAMBDA

FUNCTION

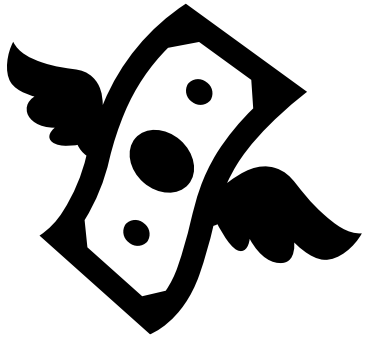


REWRITES SAME DATA
DUPLICATES IF VERSIONNED
RATE LIMITING & COST INCREASES AT SOURCE

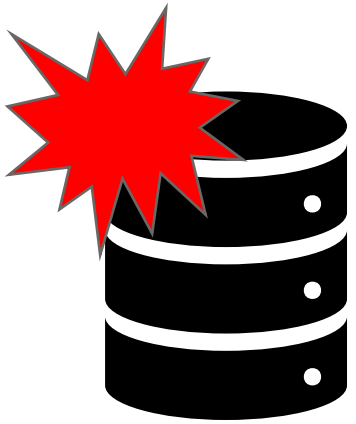
PROCESS DUPLICATION



COMPUTE & STORAGE COSTS



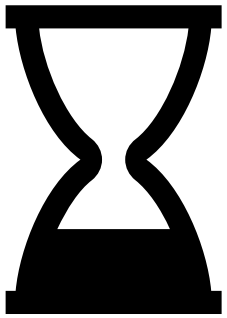
DATA SOURCE ISSUES



DATA SOURCE ISSUES

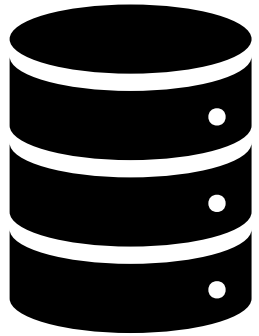
HIGHER ACCESS COSTS

RESTRICATED ACCESS



NO ACCESS

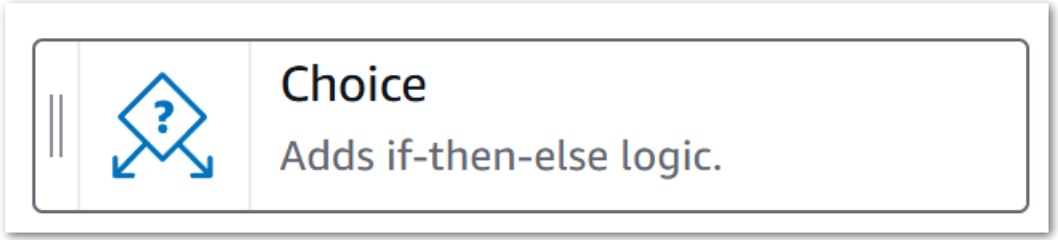
AWS LAMBDA FUNCTION

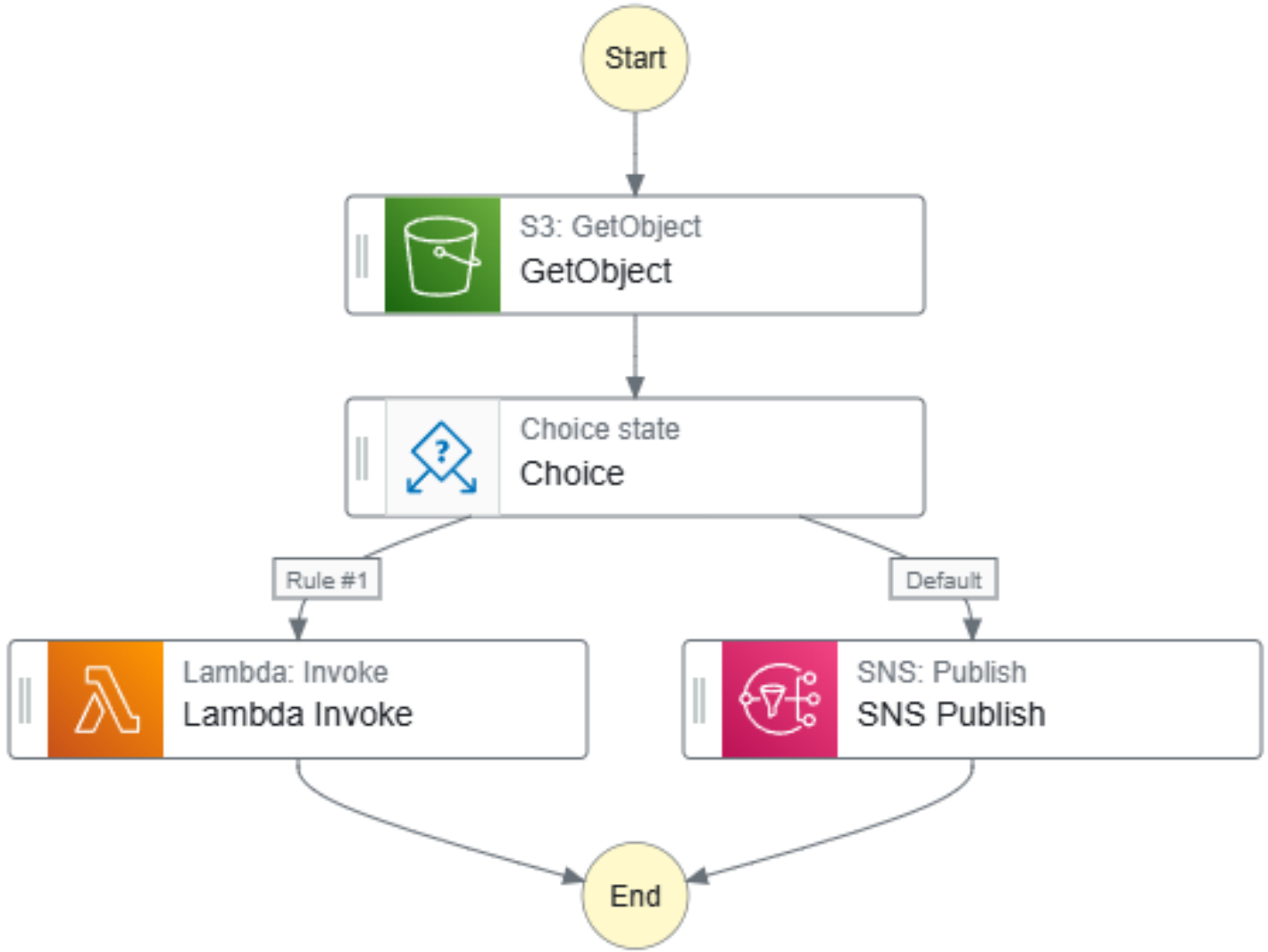


API / DATABASE

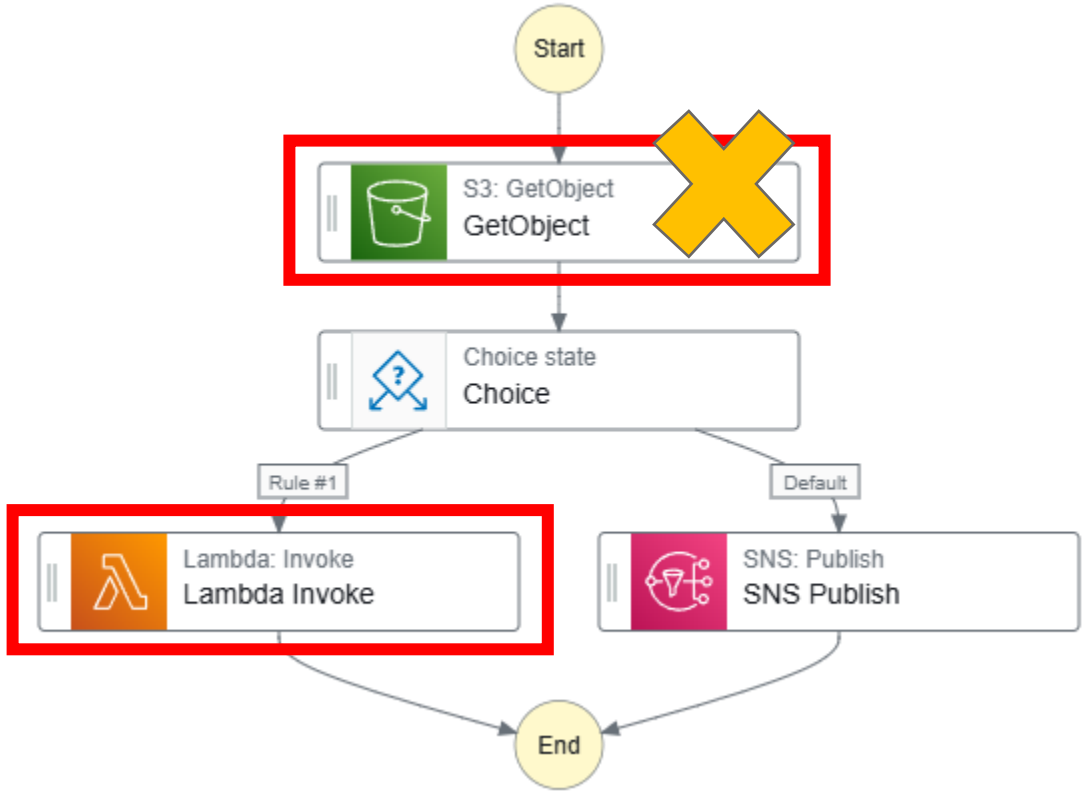
AMAZON S3 BUCKET

STEP FUNCTION CHOICE STATE

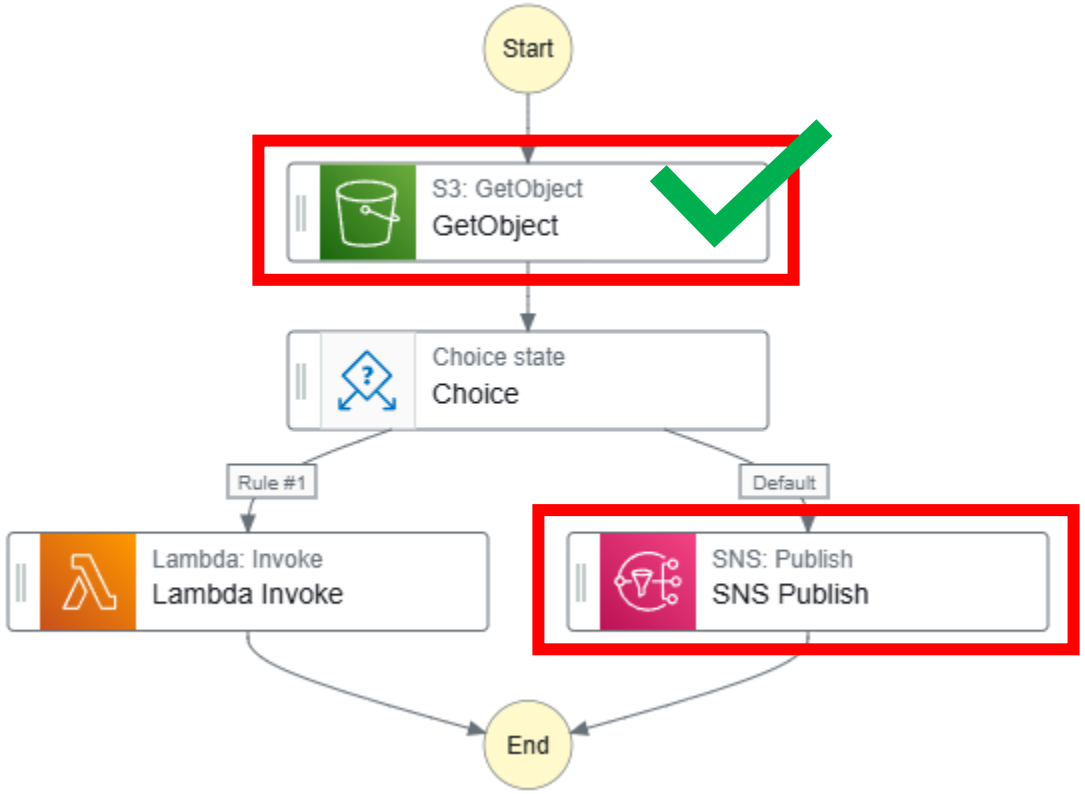



Start
S3: GetObject
GetObject
Choice state
Choice
Rule #1
Default
Lambda: Invoke
Lambda Invoke
SNS: Publish
SNS Publish
End

NO S3 DATA FOUND = CALL API

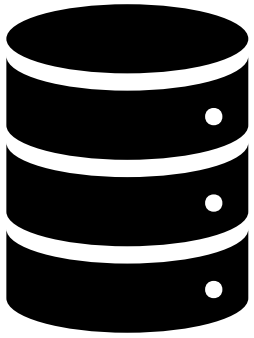


S3 DATA FOUND = NOTIFY USER

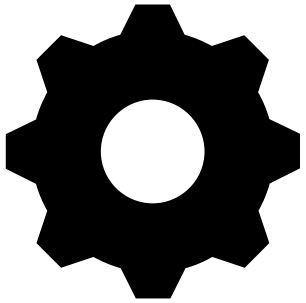


SAVES MONEY ON...

STORAGE



COMPUTE

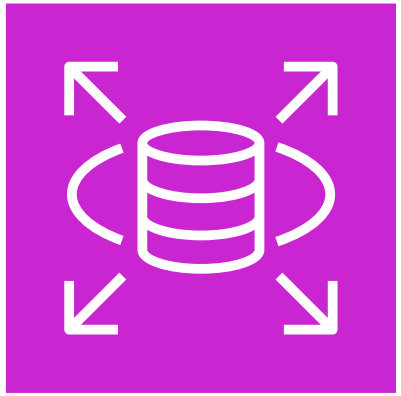


THIRD PARTY APIS

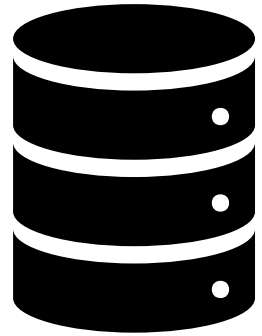
ALSO WORKS WITH...

DYNAMODB

RDS



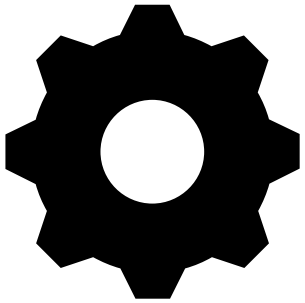
ONPREM / THIRD PARTY DATABASES



JOB CHAIN

SAVES MONEY ON...

COMPUTE



INCIDENT MANAGEMENT

BUSINESS CONTINUITY

SITUATION

AMAZON ECS

TASK

OVERNIGHT PROCESS CHAIN

A

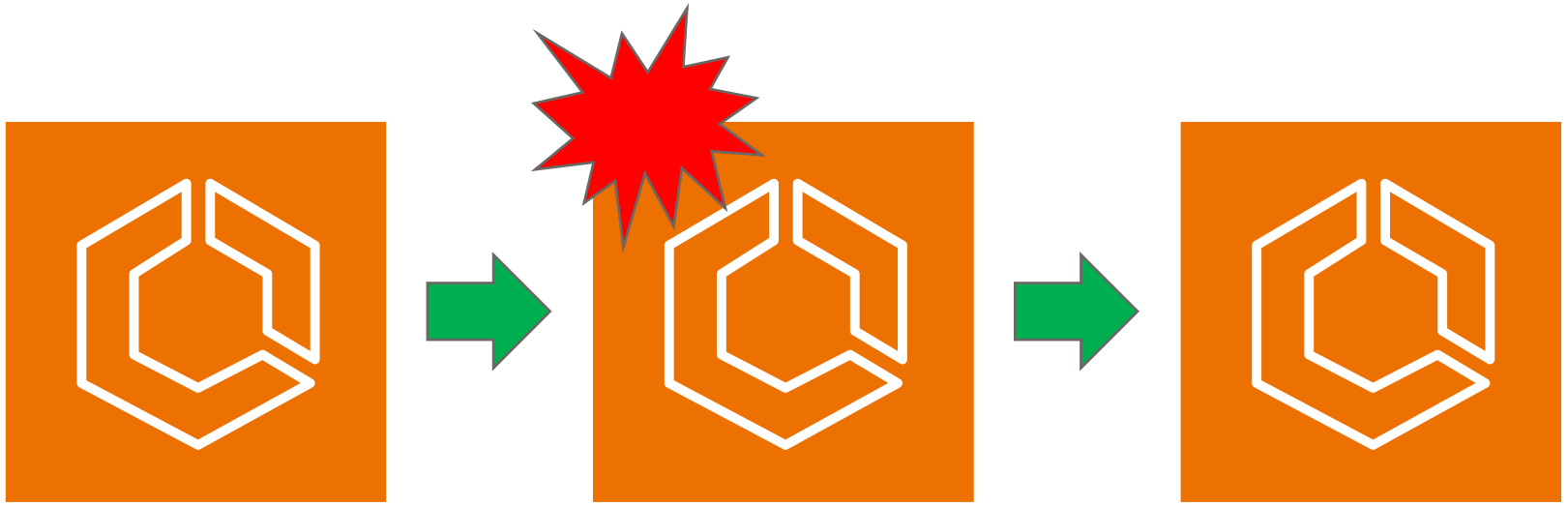
B

C

SEQUENTIAL CALLS

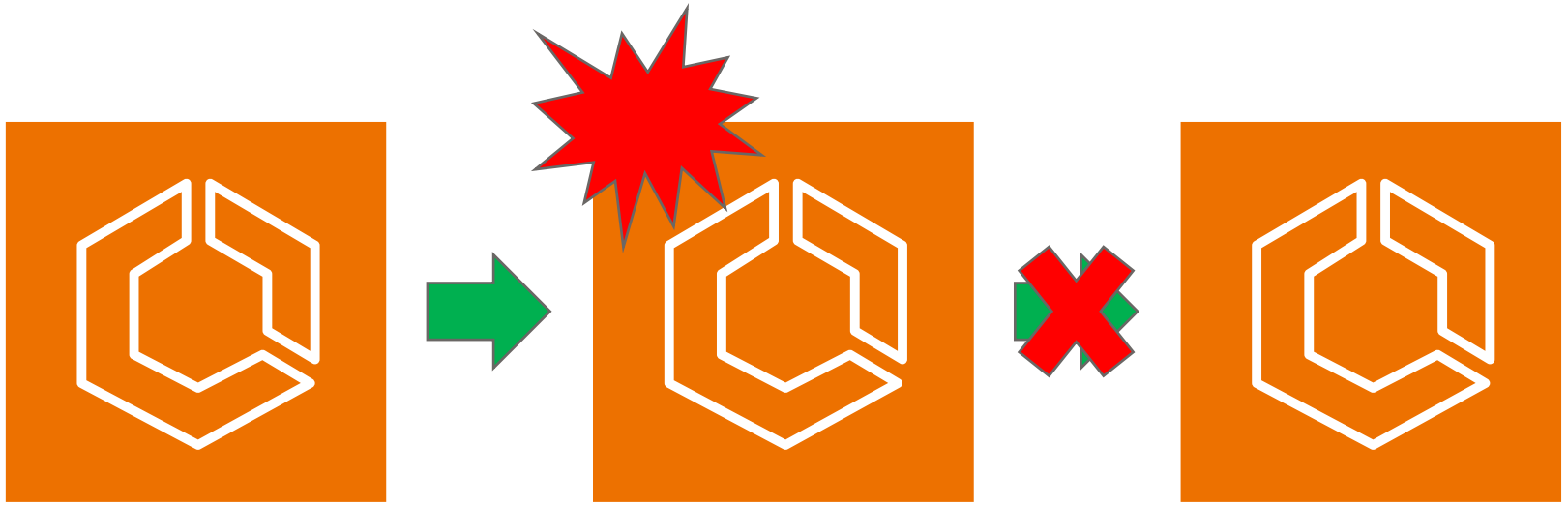
A B C

UH OH



A B C

DOUBLE UH OH



A B C

FAILURE ON B

ON CALL TRIGGERED

TASK C NEVER STARTS

DOWNSTREAM IMPACT

COST IMPACT

ON CALL PAYMENTS

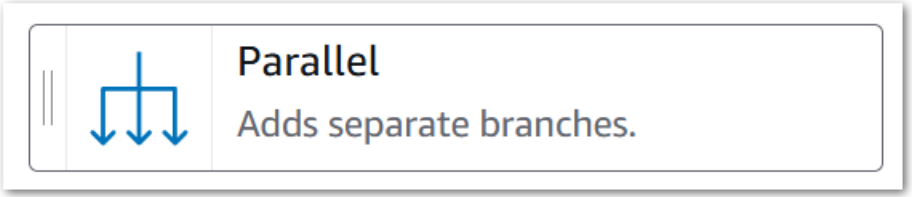
BUSINESS DISRUPTION

CUSTOMER DETRIMENT

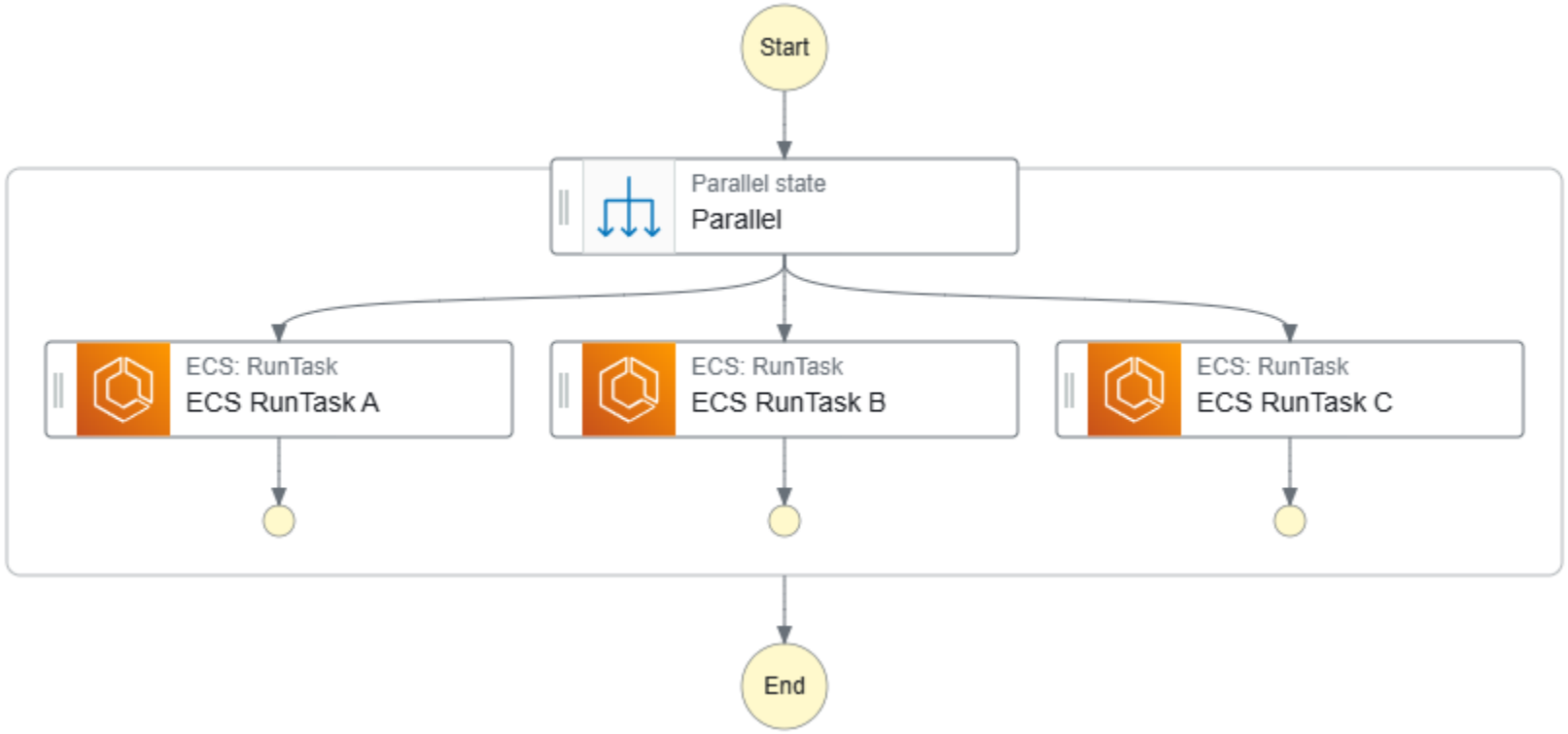
SEQUENTIAL CALLS

A B C

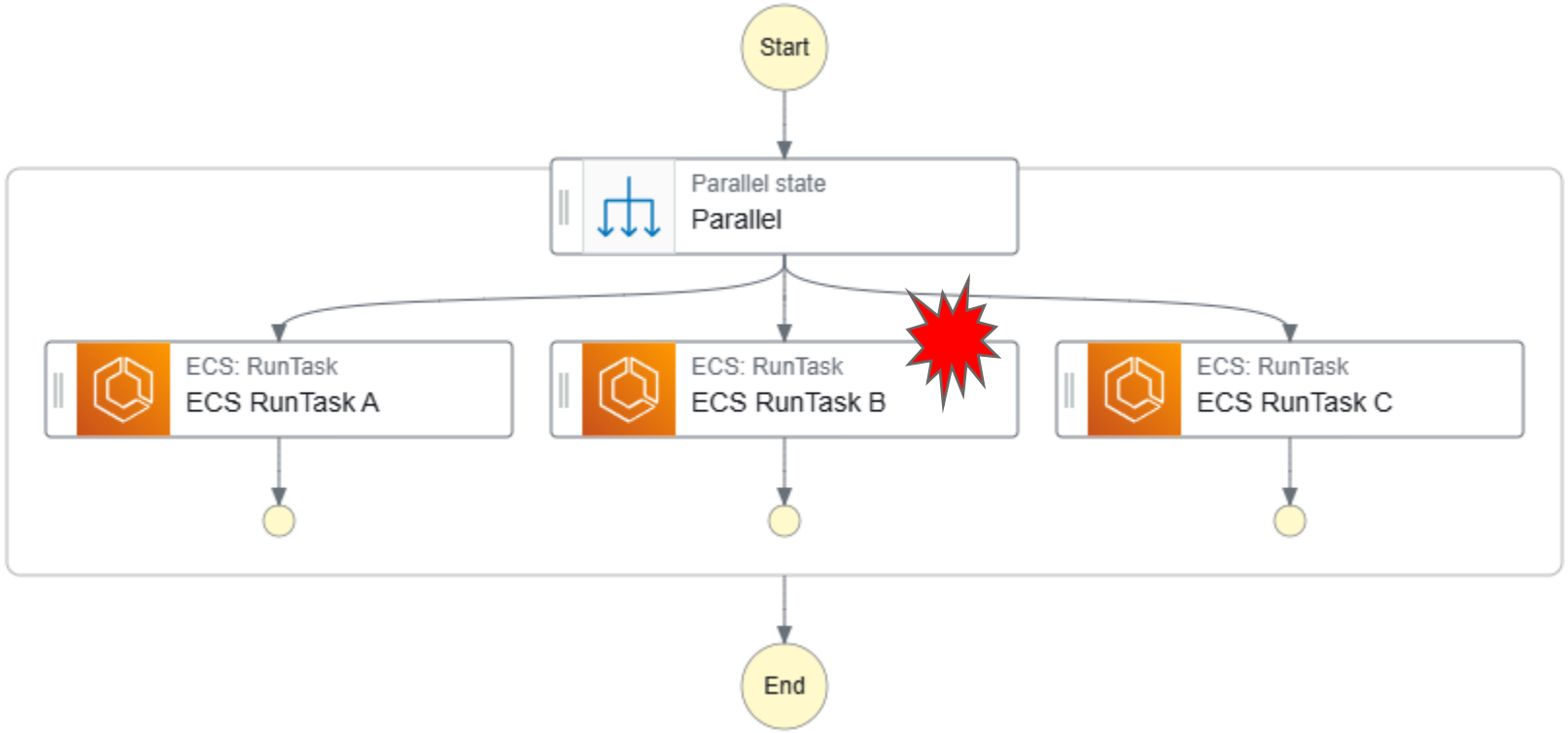
STEP FUNCTION PARALLEL STATE



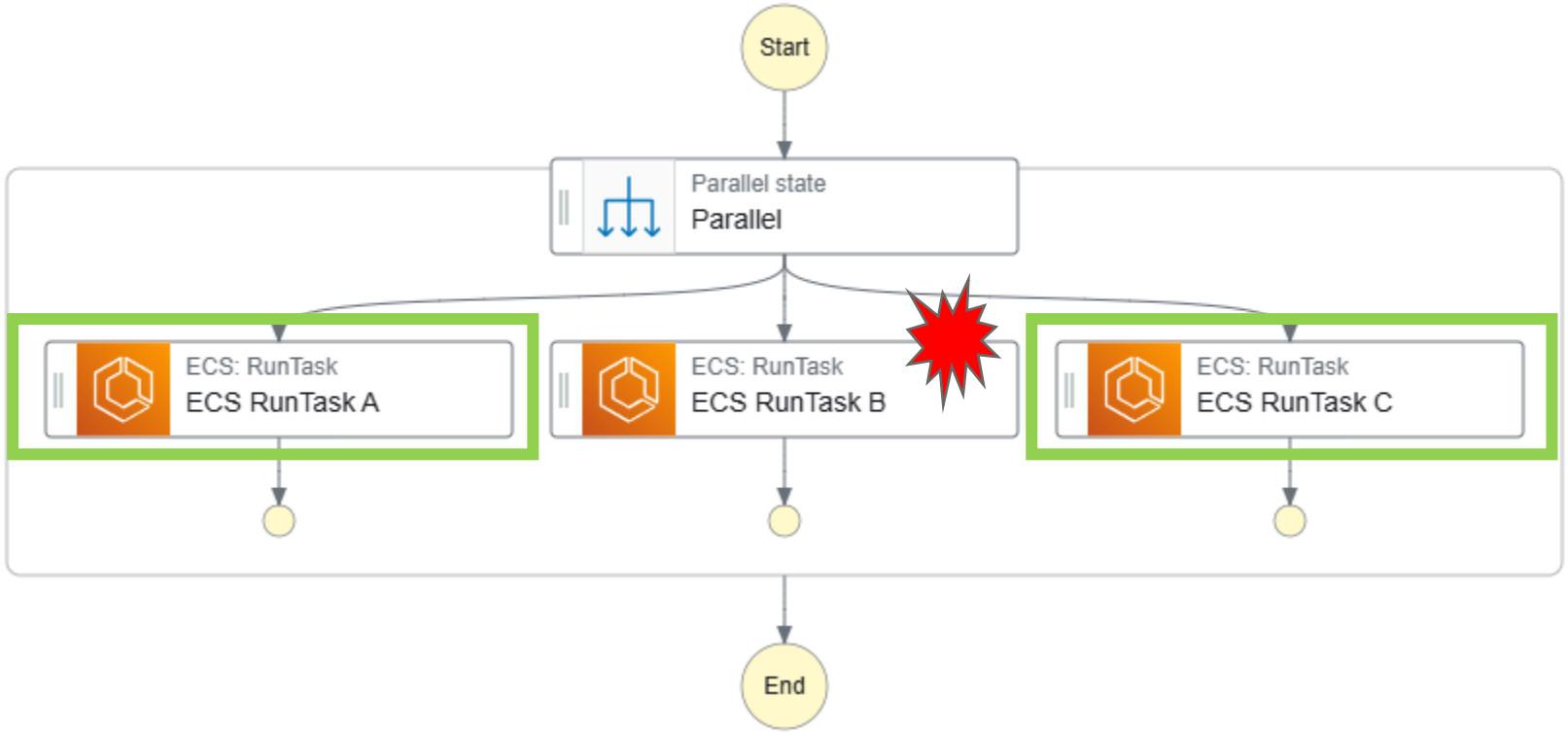
STEP FUNCTION PARALLEL WORKFLOW



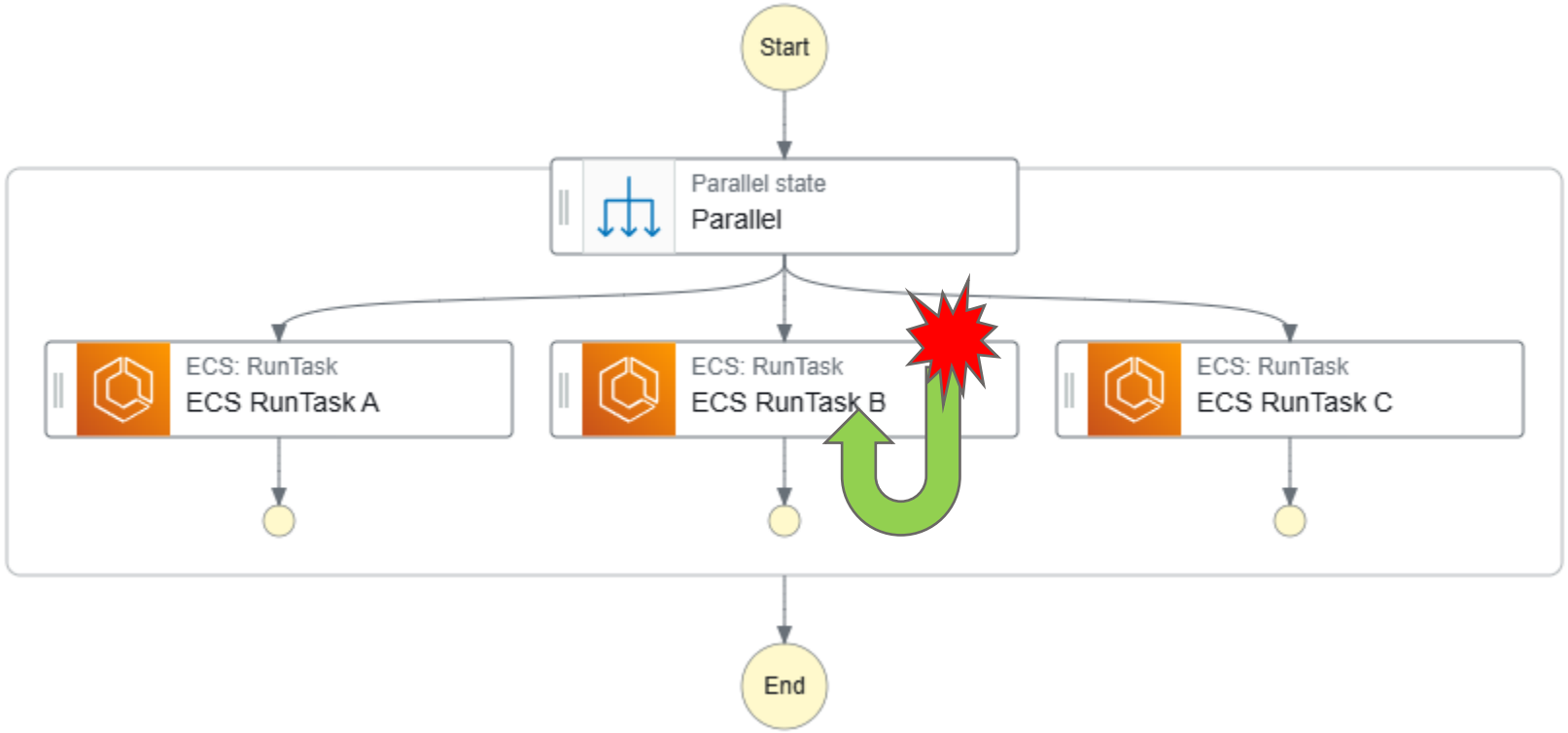
ISSUE ON B


Start
Parallel state
Parallel
ECS: RunTask
ECS RunTask A
ECS: RunTask
ECS RunTask B
ECS: RunTask
ECS RunTask C
End

ISSUE ON B

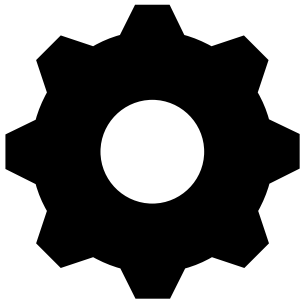

Start
Parallel state
Parallel
ECS: RunTask
ECS RunTask A
ECS: RunTask
ECS RunTask B
ECS: RunTask
ECS RunTask C
End

TRY-CATCH LOGIC



SAVES MONEY ON...

COMPUTE



INCIDENT MANAGEMENT

BUSINESS CONTINUITY

STOPPING ABANDONED JOBS

MY AWS BILL

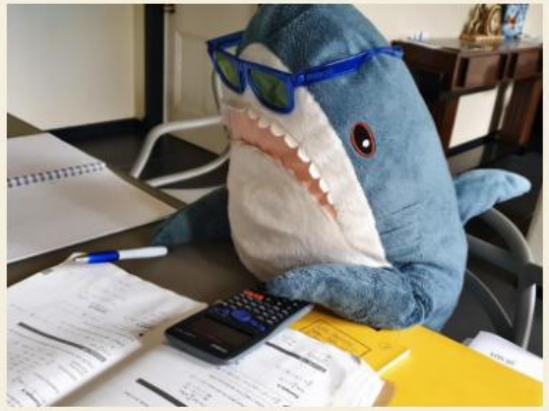
AWS GLUE

JOB

THE CAUSE

Gold Layer PySpark ETL With AWS Glue Studio

November 15, 2024



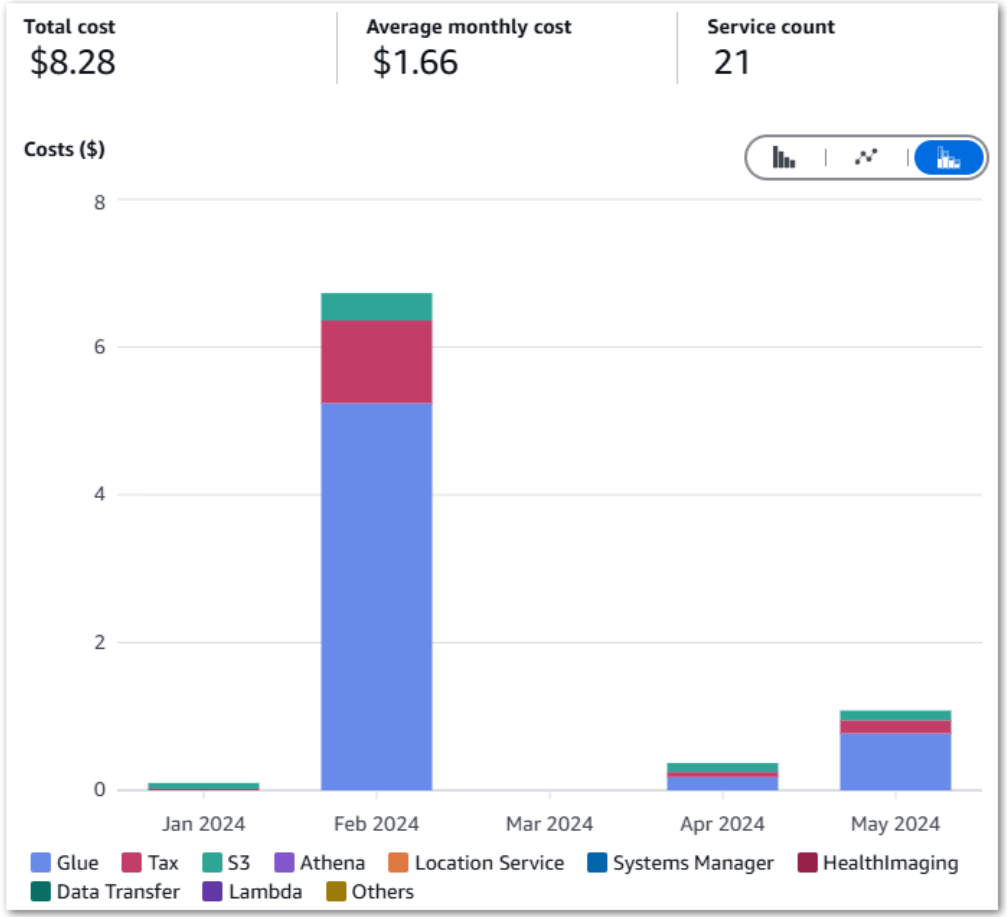
Bespoke Veracity Checks With AWS Glue Data Quality

September 2, 2024

Silver Layer Python ETL With The AWS Glue ETL Job Script Editor

August 12, 2024

COSTS: 2024 H1



COSTS: 2024 JUNE

COST & USAGE: API OPERATION

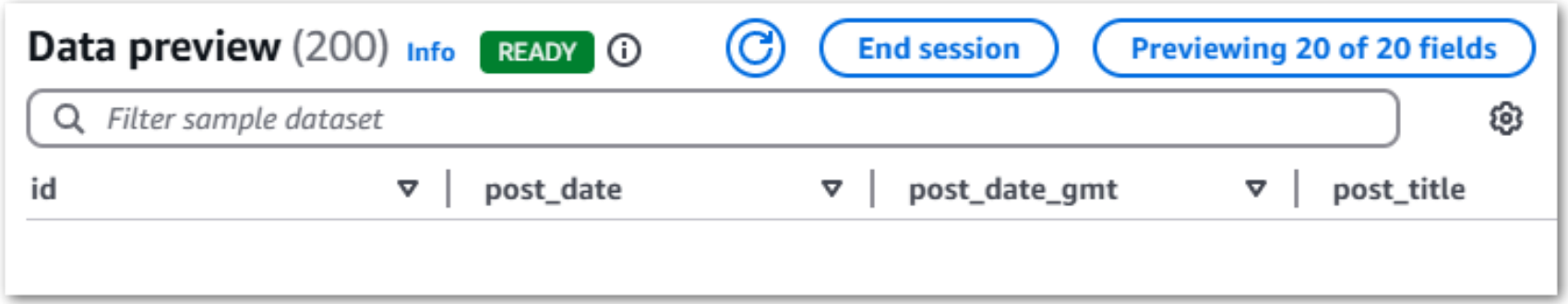
COST & USAGE: API OPERATION – TAX

COST: API OPERATION – GLUE JOB RUN

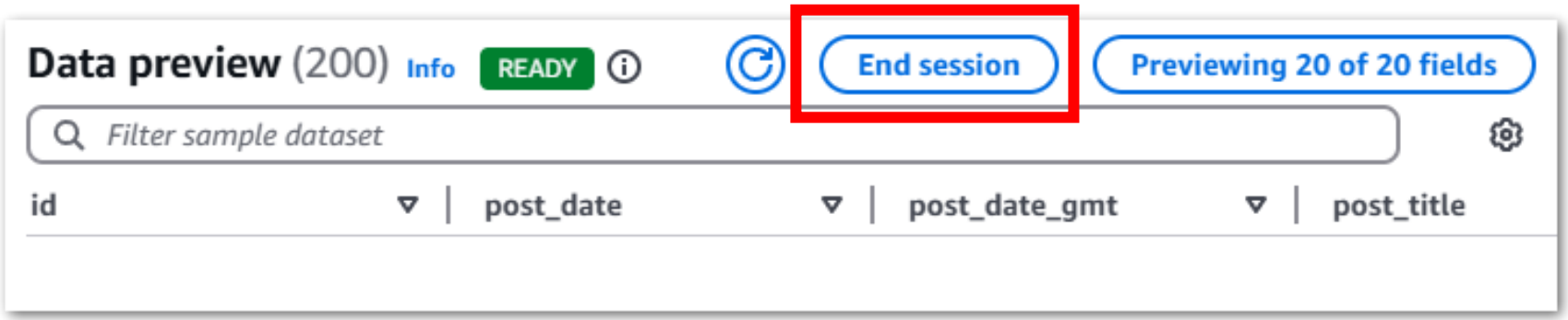
COST: API OPERATION – GLUE INTERACTIVE SESSION

COST DRIVER #1: OPEN SESSIONS

COST DRIVER: ACTIVE SESSION



COST DRIVER: ACTIVE SESSION



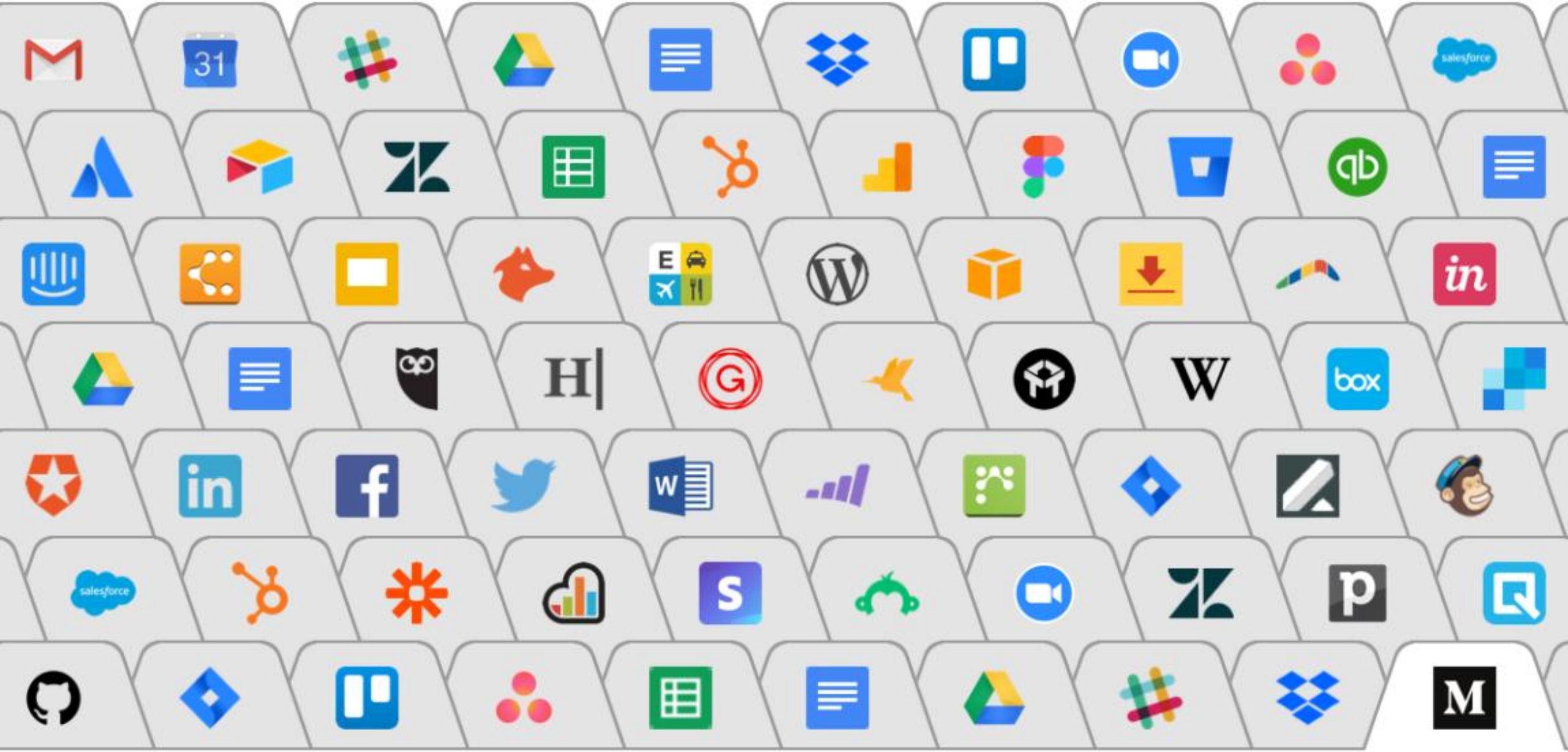
COST DRIVER #2: MULTIPLE SESSIONS

COST DRIVER: MULTIPLE SESSIONS

ONE TAB

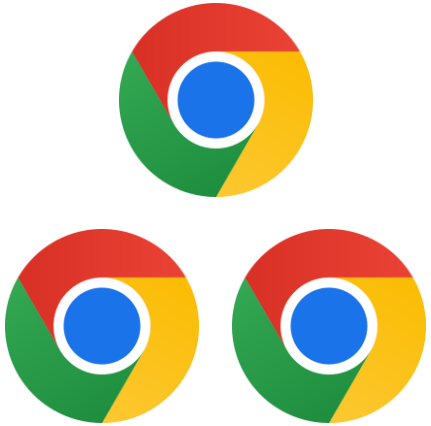
ONE SESSION

ONE BILLABLE RESOURCE

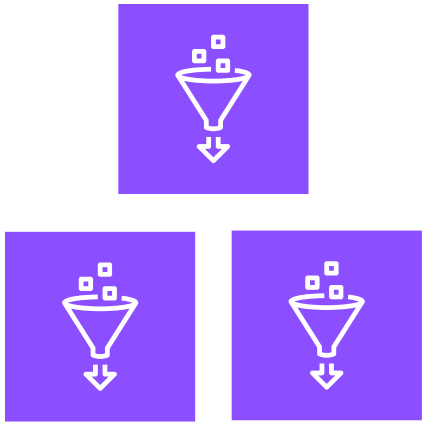

31
salesforce
in
box
salesforce

COST DRIVER: MULTIPLE SESSIONS

MANY TABS

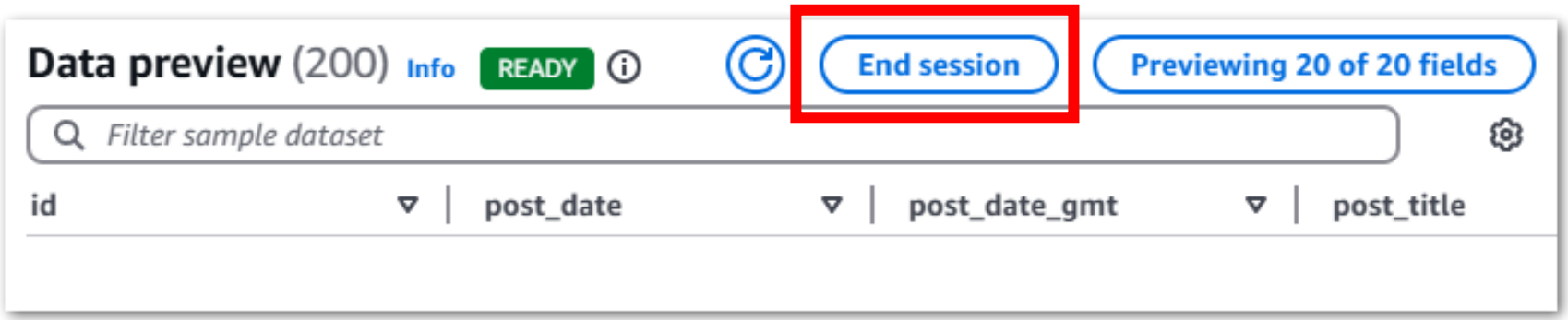


MANY SESSIONS

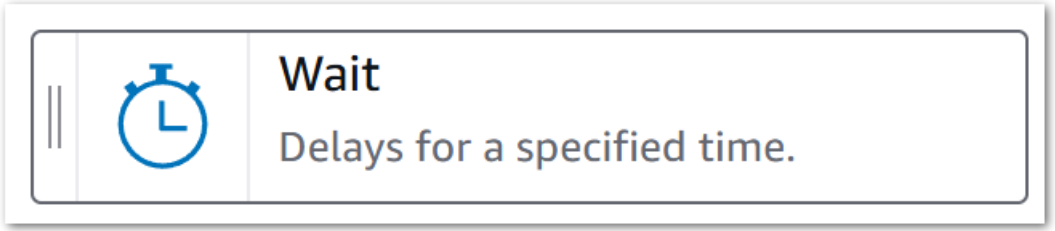


MANY BILLABLE RESOURCES!

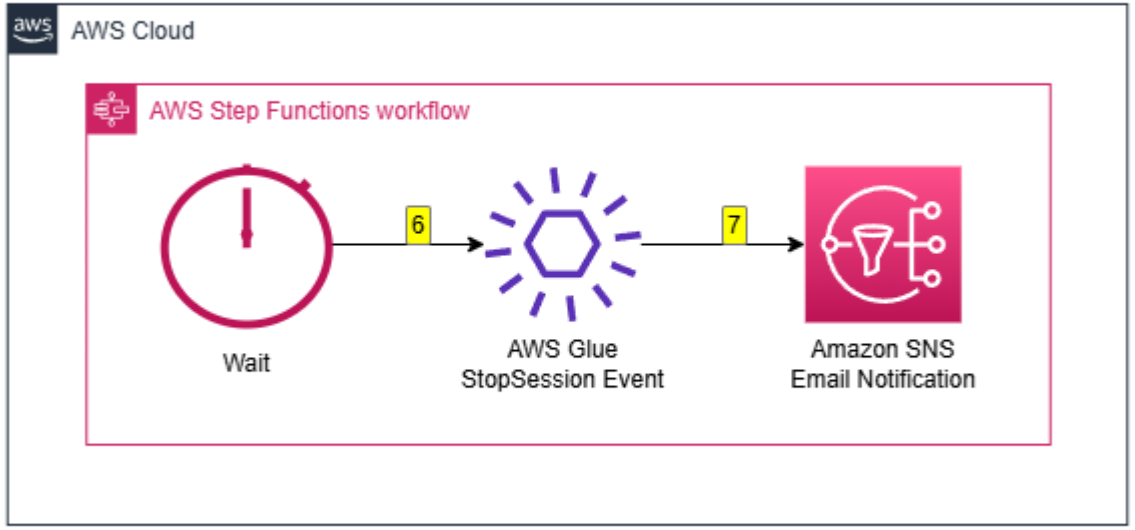
COST DRIVER: ACTIVE SESSION



STEP FUNCTION WAIT STATE

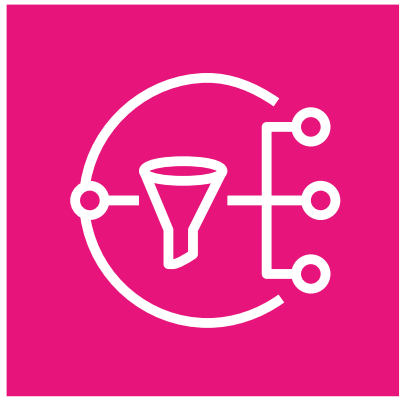


ARCHITECTURE DIAGRAM STEP F



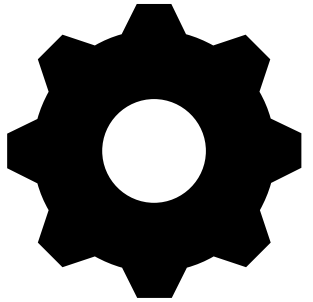
OTHER USES:
PERFORM OTHER ACTIONS

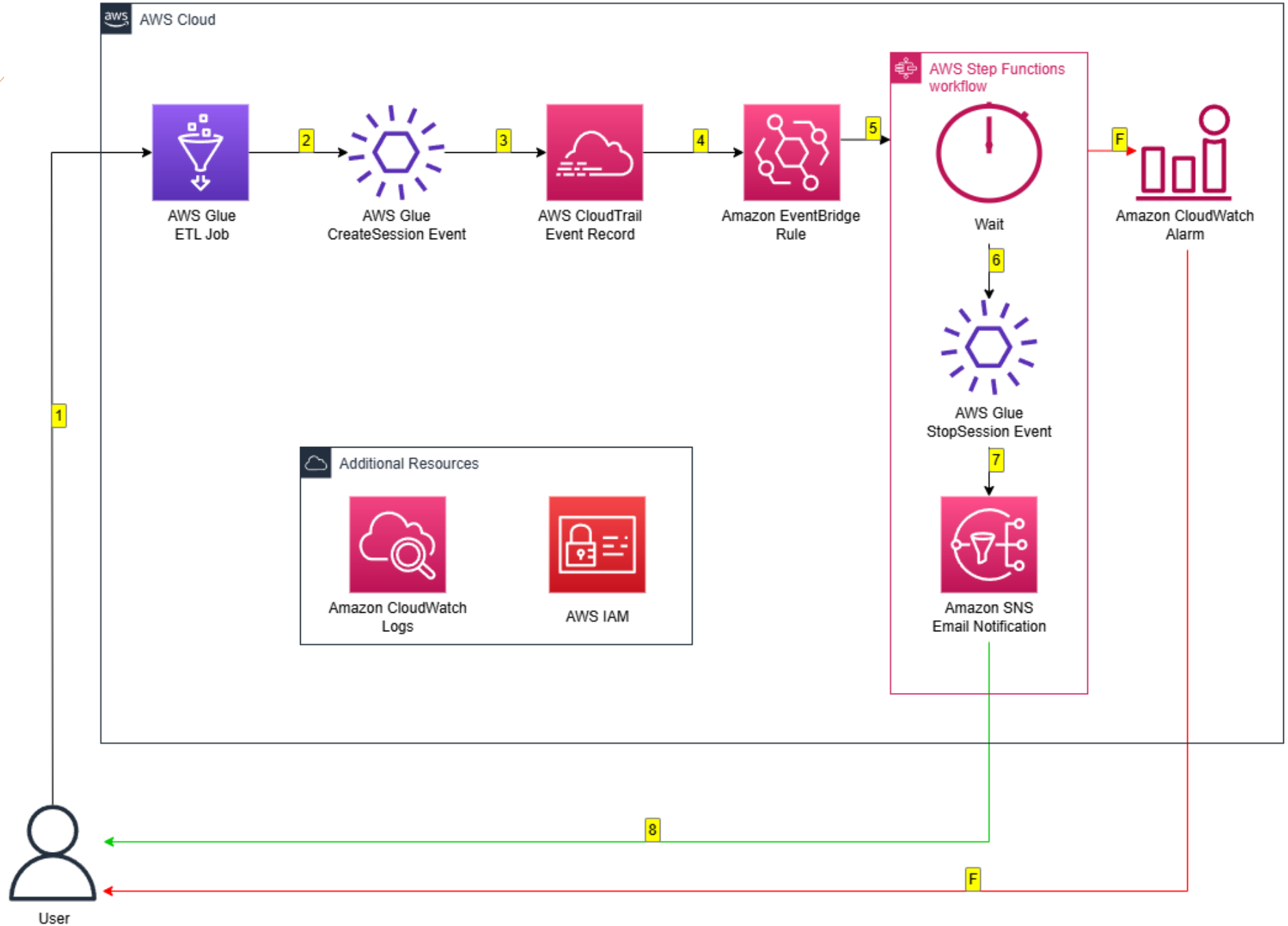
EMAIL & SLACK NOTIFICATIONS

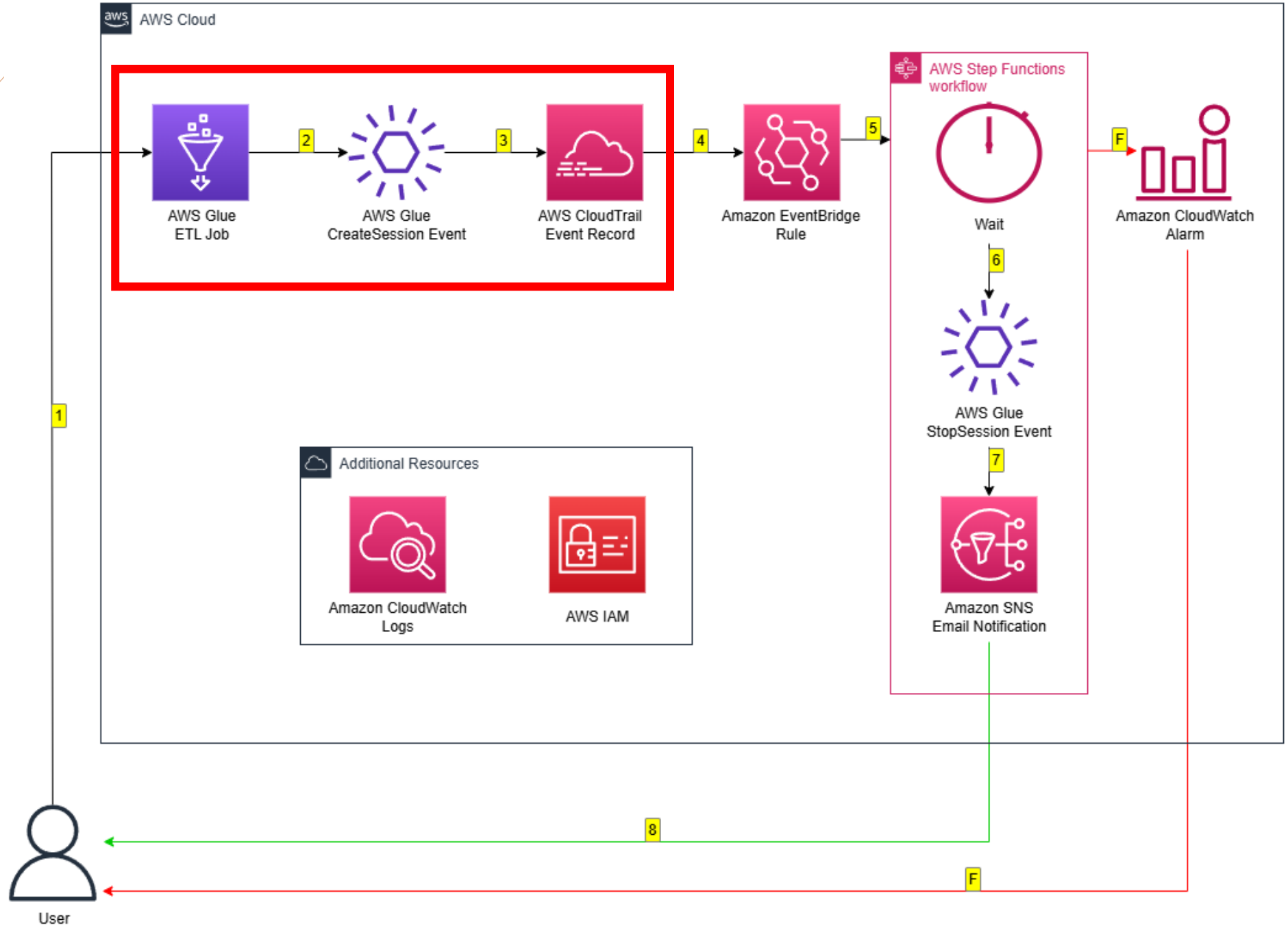


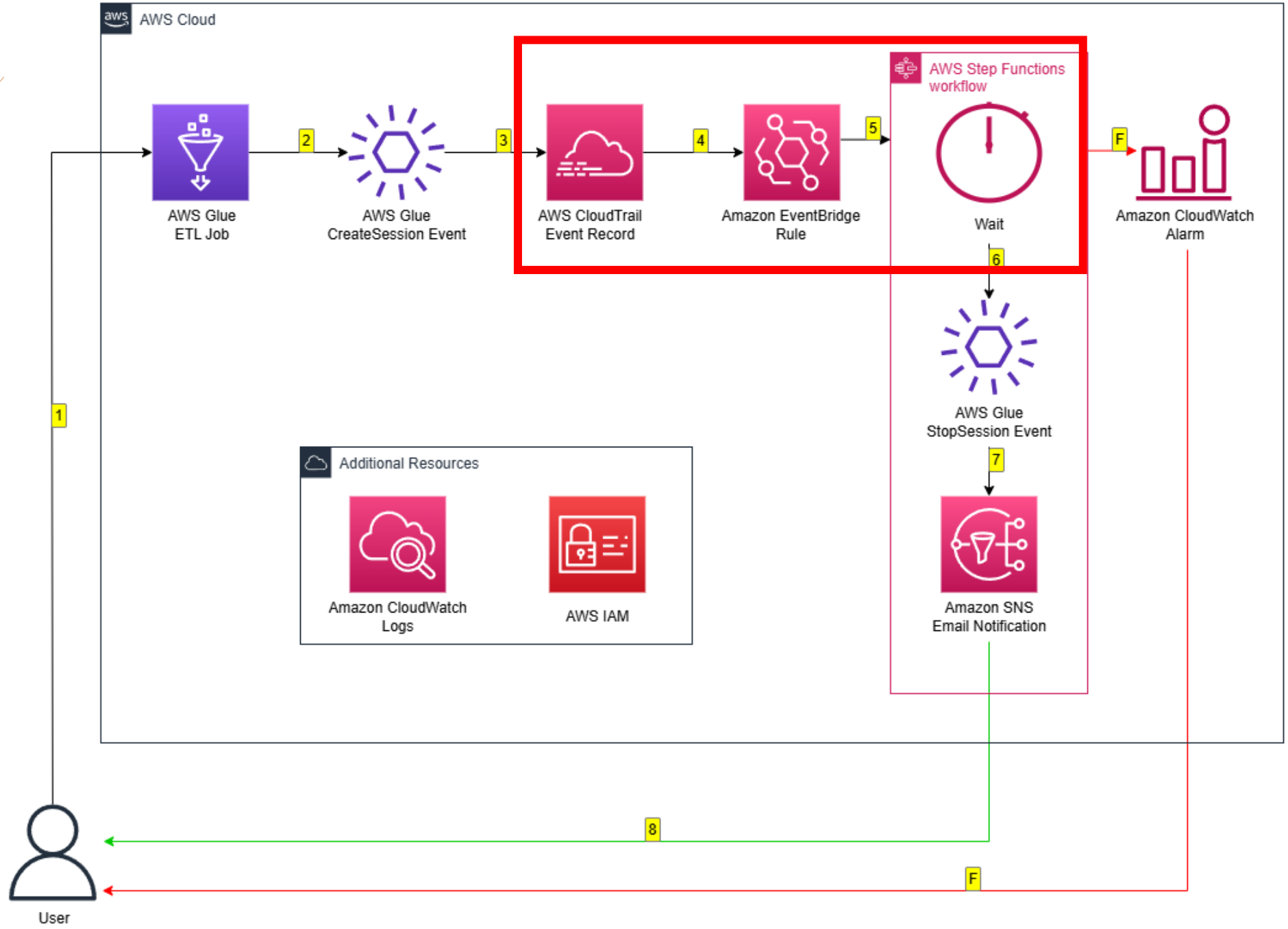
BESPOKE LOGGING

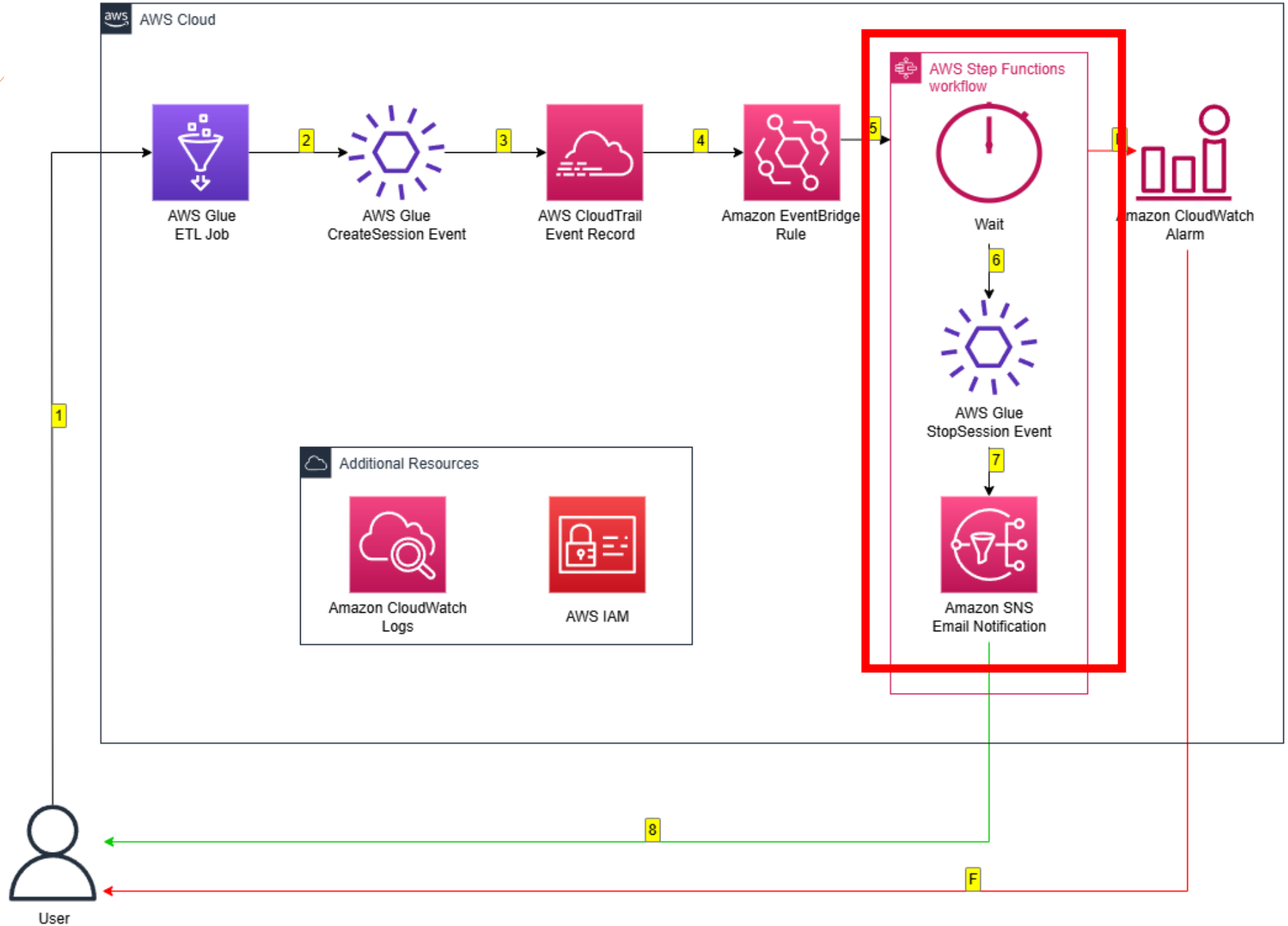
TRIGGER SEPARATE WORKFLOW



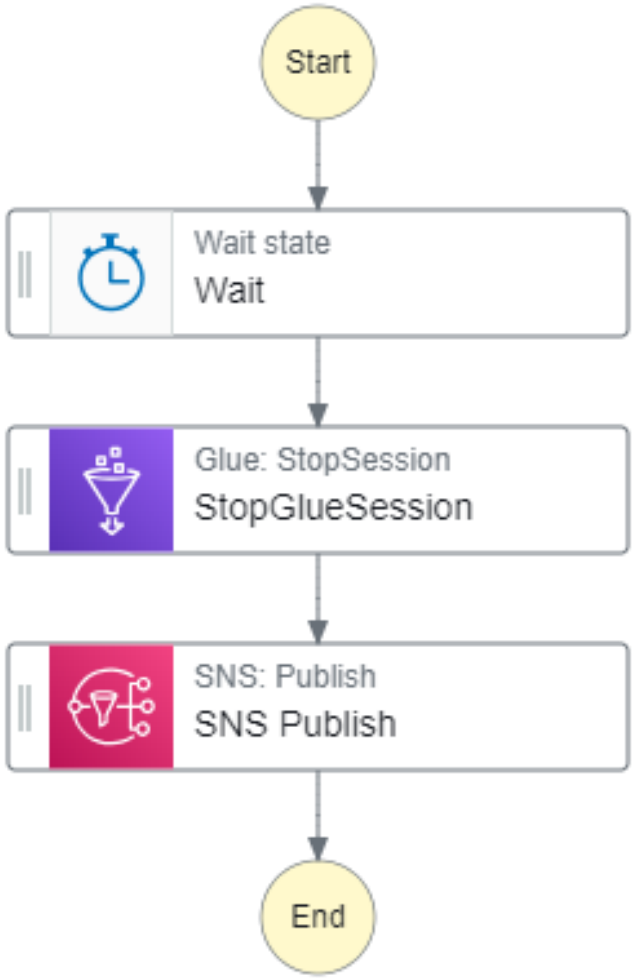

AWS Cloud
AWS Step Functions workflow
AWS Glue ETL Job
AWS Glue CreateSession Event
AWS CloudTrail Event Record
Amazon EventBridge Rule
Wait
Amazon CloudWatch Alarm
AWS Glue StopSession Event
Amazon SNS Email Notification
Additional Resources
Amazon CloudWatch Logs
AWS IAM
User
1
2
3
4
5
6
7
8
F
F


AWS Cloud
AWS Glue
ETL Job
2
AWS Glue
CreateSession Event
3
AWS CloudTrail
Event Record
4
Amazon EventBridge
Rule
5
AWS Step Functions
workflow
Wait
6
AWS Glue
StopSession Event
7
Amazon SNS
Email Notification
F
Amazon CloudWatch
Alarm
1
Additional Resources
Amazon CloudWatch
Logs
AWS IAM
8
F
User


AWS Cloud
AWS Step Functions workflow
AWS Glue ETL Job
AWS Glue CreateSession Event
AWS CloudTrail Event Record
Amazon EventBridge Rule
Wait
Amazon CloudWatch Alarm
AWS Glue StopSession Event
Amazon SNS Email Notification
Additional Resources
Amazon CloudWatch Logs
AWS IAM
User
1
2
3
4
5
6
7
8
F
F


AWS Cloud
AWS Step Functions workflow
AWS Glue ETL Job
AWS Glue CreateSession Event
AWS CloudTrail Event Record
Amazon EventBridge Rule
Wait
Amazon CloudWatch Alarm
AWS Glue StopSession Event
Amazon SNS Email Notification
Additional Resources
Amazon CloudWatch Logs
AWS IAM
User
1
2
3
4
5
6
7
8
F

STEP FUNCTIONS STUDIO



DEMO: GLUE

AWS Glue
Getting started
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Data Catalog tables
Data connections
Workflows (orchestration)
Zero-ETL integrations New
Data Catalog
Databases
Tables
Stream schema registries
Schemas
Connections
Crawlers
Classifiers
Catalog settings
Data Integration and ETL
Zero-ETL integrations New
ETL jobs
Visual ETL
Notebooks
Job run monitoring
Interactive Sessions
Data classification tools
Sensitive data detection
Record Matching
Triggers
Workflows (orchestration)
Blueprints
Security configurations
Cost management New
Legacy pages
Connections (legacy)
Jobs (legacy)

AWS Glue Studio Info

Create job Info

Author in a visual interface focused on data flow.
Visual ETL

Author using an interactive code notebook.
Notebook

Author code with a script editor.
Script editor

Example jobs Info

Create example job

Your jobs (6) Info

Actions | Run job

Filter jobs by property

1

Job name	Type	Created by	Last modified	AWS Glue version
WordPress_Gold_statisticspagespostsjoin_botodelete	Glue ETL	Script	28/10/2024, 22:29:26	4.0
WordPress_Gold_statisticspagespostsjoin_mod	Glue ETL	Script	28/10/2024, 17:51:15	4.0
WordPress_Gold_statisticspagespostsjoin	Glue ETL	Visual	23/08/2024, 11:04:05	4.0
wordpressapi_silver	Python shell	Script	09/08/2024, 13:30:58	
bronze-statistics_pages_crawler	Glue ETL	Visual	12/03/2024, 15:30:57	4.0
bronze-statistics_pages	Glue ETL	Visual	12/03/2024, 13:09:19	4.0

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

DEMO: STEP FUNCTION

Step Functions
State machines
Activities
Developer resources
Online learning workshop
Local Development
Data flow simulator
Feature spotlight
Documentation
Join our feedback panel

GlueSession-WaitAndStop

Edit | Actions | Start execution

Details

Arn
arn:aws:states:eu-west-1:973122011240:stateMachine:GlueSession-WaitAndStop

IAM role ARN
arn:aws:iam::973122011240:role/service-role/StepFunctions-GlueSession-WaitAndStop-role-b8l7sy5xc

Type
Standard

Status
Active

Creation date
Jun 5, 2024, 19:31:49 (UTC+01:00)

X-Ray tracing
Disabled

Executions | Monitoring | Logging | Definition | Aliases | Versions | Tags

Executions (0/0)

View details | Stop execution | Redrive | Start execution

Filter executions by property or value | All | Last 15 months | Local timezone | 0 matches | 1

Name	Status	Start Time (local)	End Time (local)	Duration Info	Version	Alias

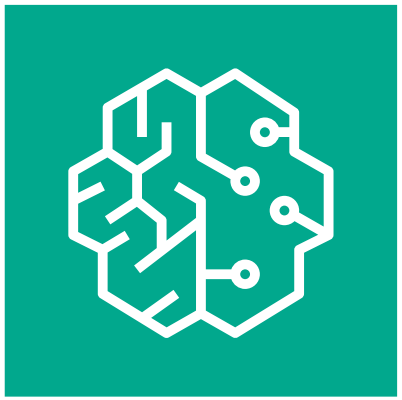
No executions
Start execution

© 2025, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

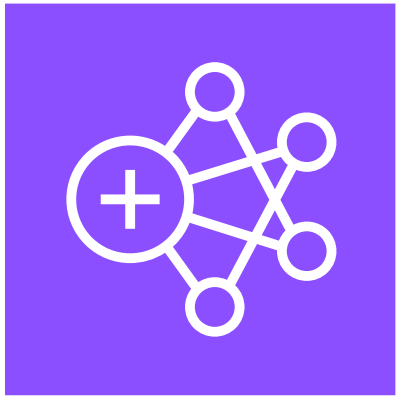
GLUE INTERACTIVE SESSIONS 2024 H2

ALSO WORKS WITH...

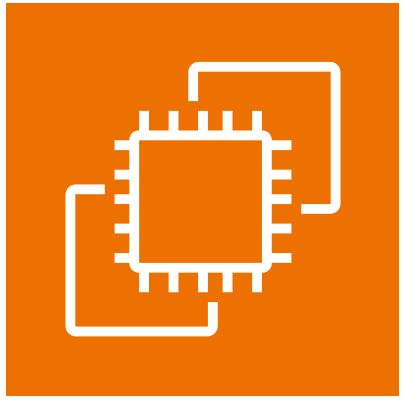
AMAZON SAGEMAKER



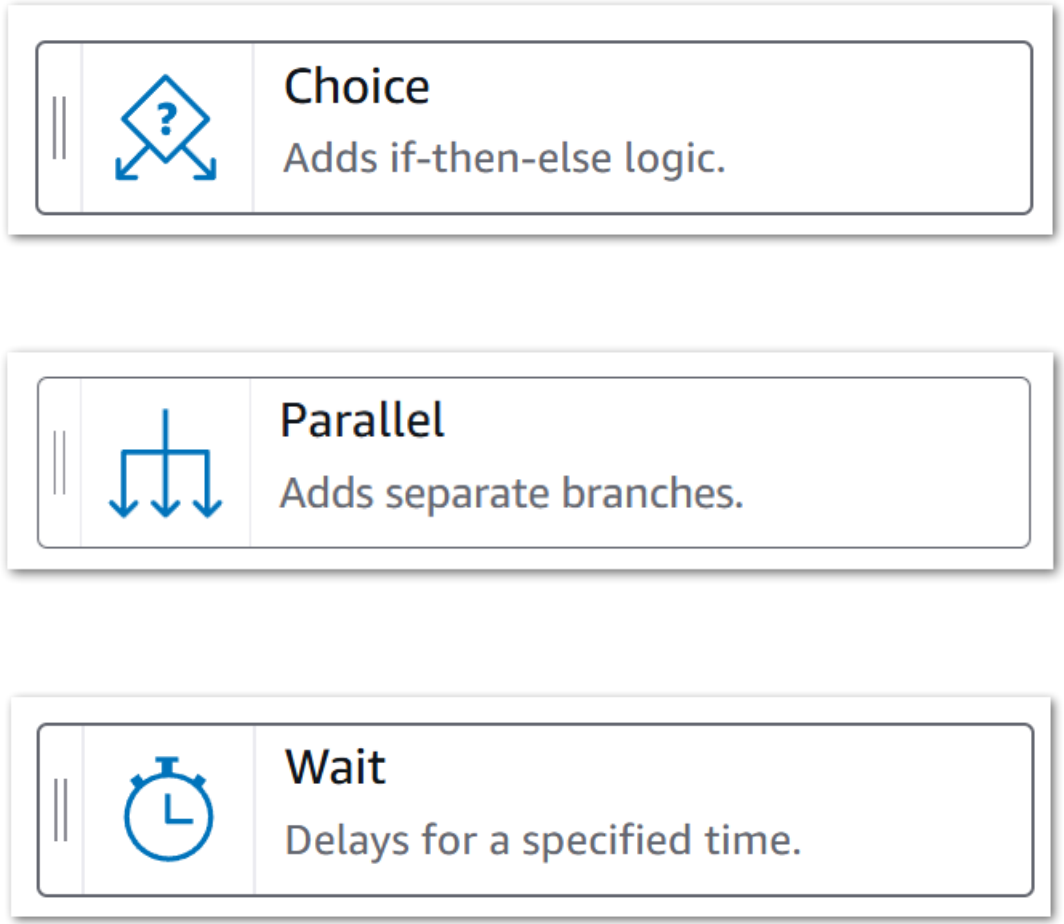
AMAZON EMR

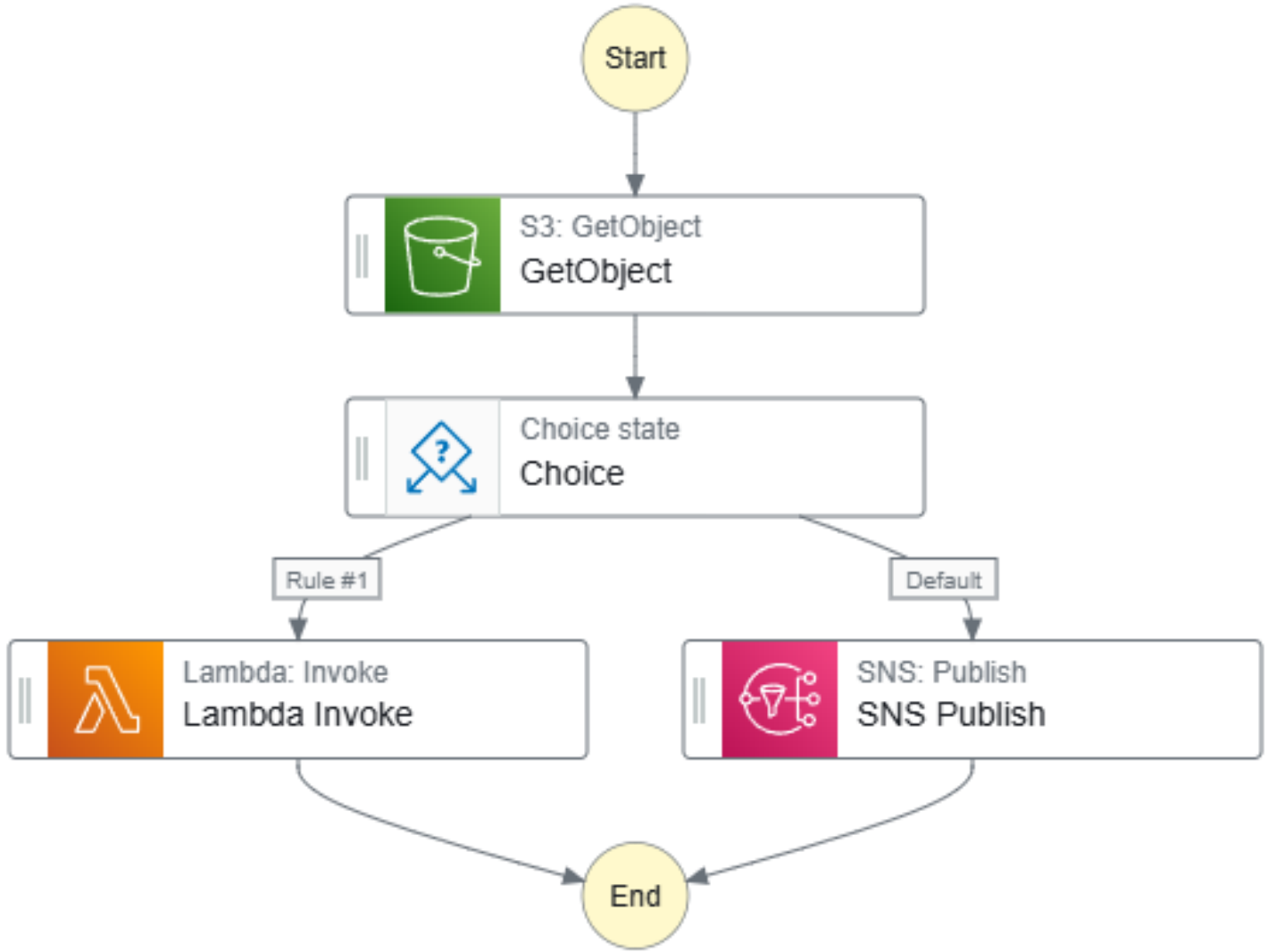


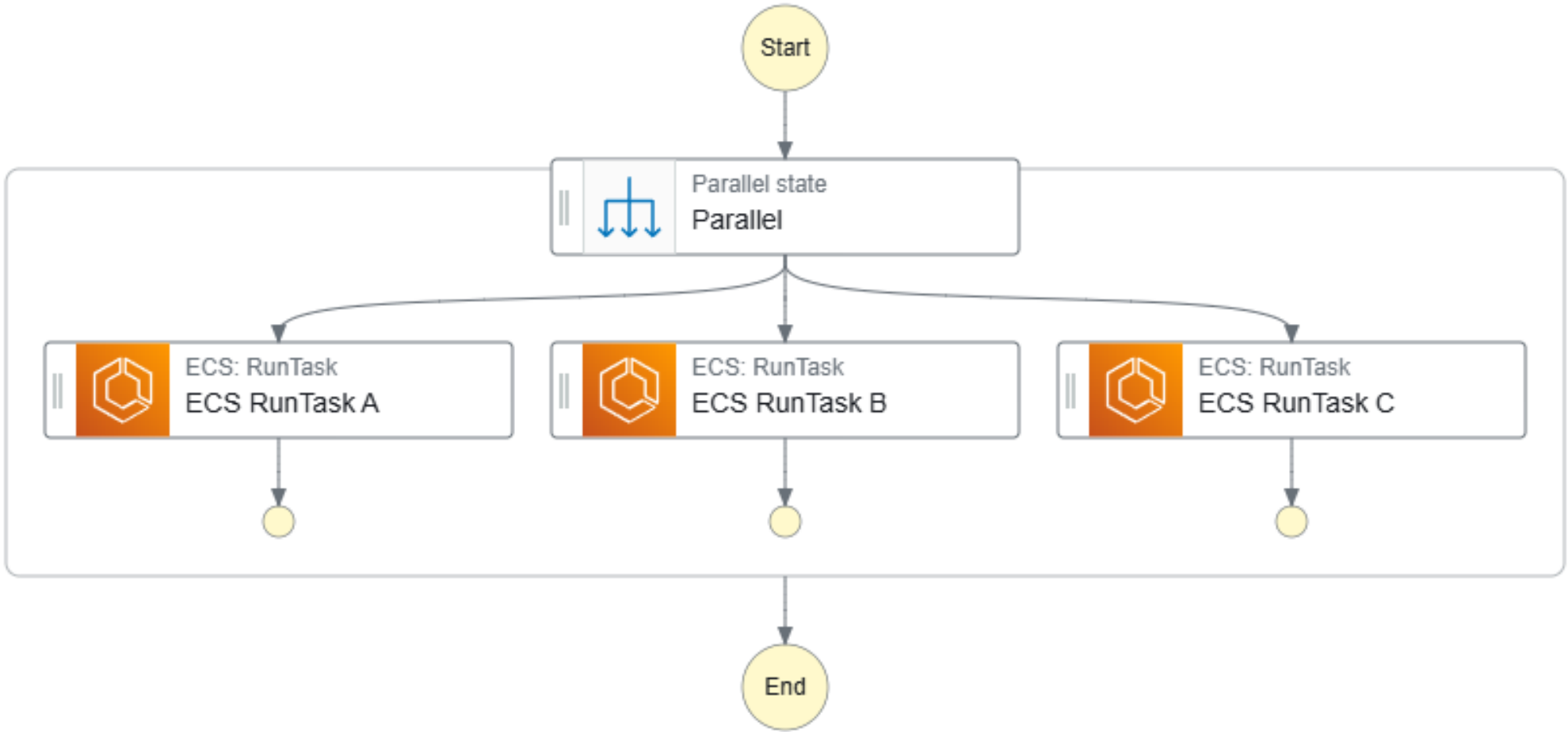
AMAZON EC2

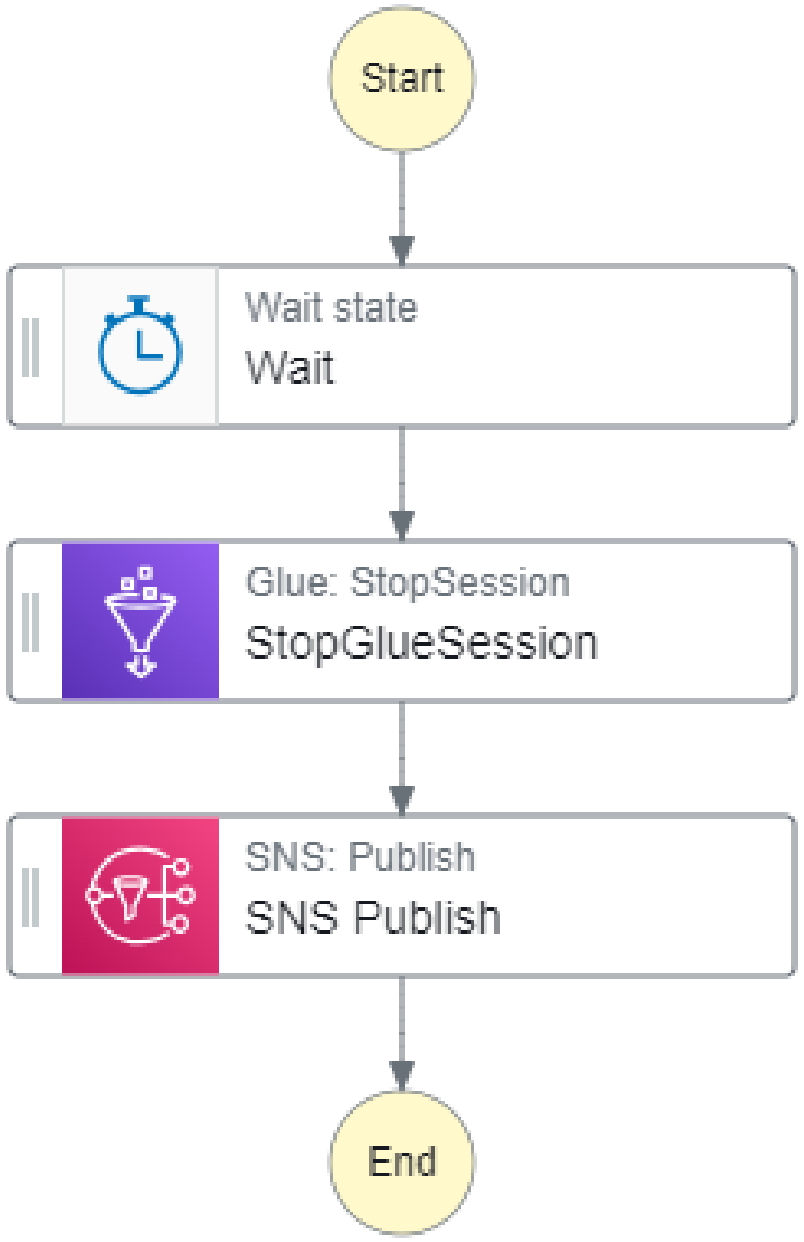


SUMMARY


Choice
Adds if-then-else logic.
Parallel
Adds separate branches.
Wait
Delays for a specified time.


Start
S3: GetObject
GetObject
Choice state
Choice
Rule #1
Default
Lambda: Invoke
Lambda Invoke
SNS: Publish
SNS Publish
End


Start
Parallel state
Parallel
ECS: RunTask
ECS RunTask A
ECS: RunTask
ECS RunTask B
ECS: RunTask
ECS RunTask C
End


Start
Wait state
Wait
Glue: StopSession
StopGlueSession
SNS: Publish
SNS Publish
End

THANK YOU

Me & Session Resources

Session Feedback