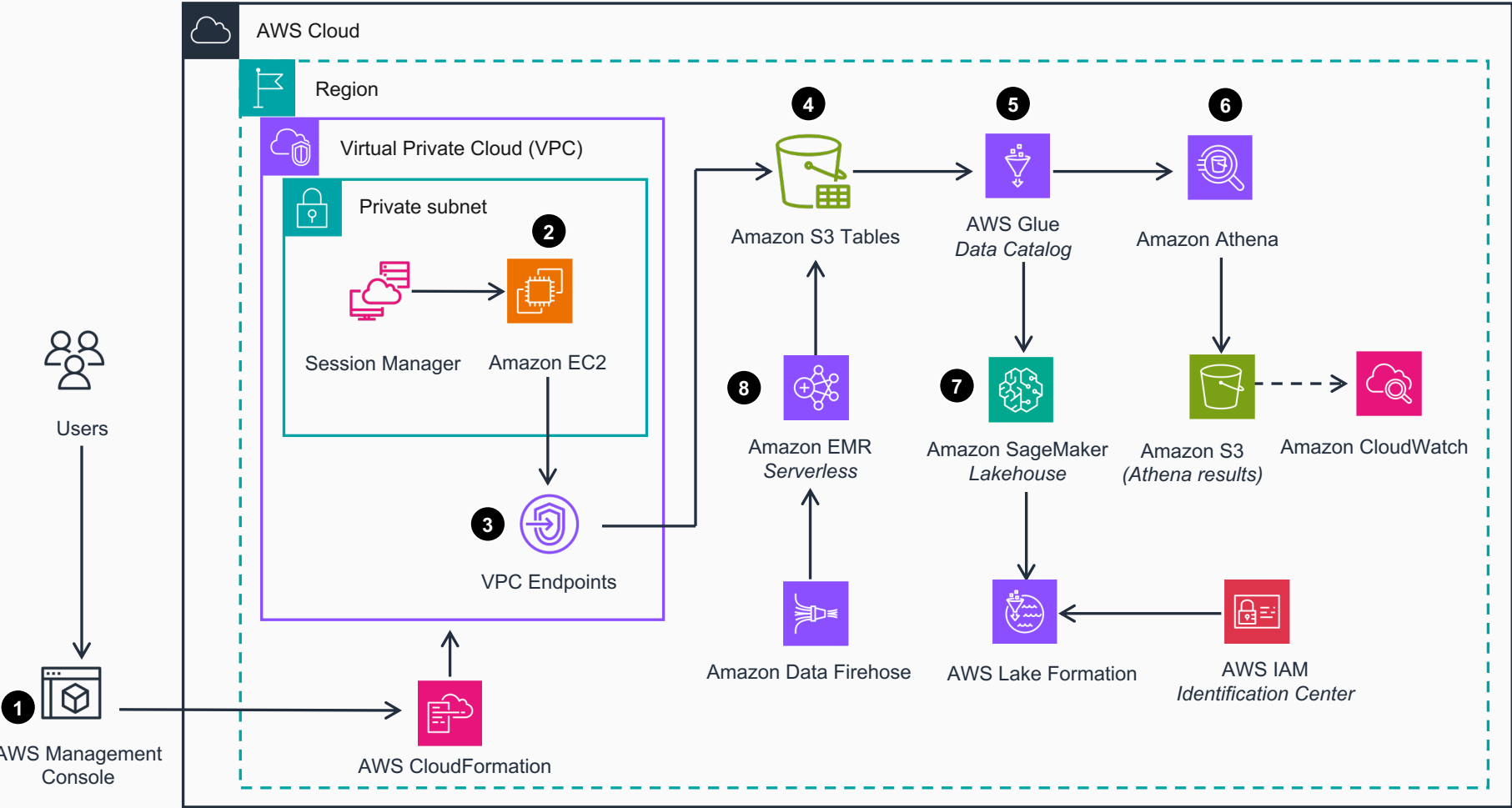


Guidance for Deploy a PoC for Amazon S3 Tables

This architecture diagram illustrates how to build and operate a Proof of Concept (PoC) for Amazon S3 Tables on AWS. It shows the key components and their transactions.



- 1 AWS Management Console**
The user logs into the AWS Management Console and deploys the proof-of-concept (PoC) template provided in this Guidance, using AWS CloudFormation to create a new stack that deploys an Amazon Virtual Private Cloud (Amazon VPC).
- 2 Amazon Elastic Compute Cloud (Amazon EC2)**
Private subnet instance accessed through SSM Session Manager. No internet gateway or NAT Gateway - all traffic routes through VPC endpoints.
- 3 VPC Endpoints**
S3 (Gateway), S3 Tables, SSM, SSM Messages, EC2 Messages, Glue, and Athena endpoints provide private connectivity to AWS services.
- 4 Amazon S3 Tables**
Table Bucket stores data in Apache Iceberg V2/V3 format. Namespace PoC data contains sensor_readings and maintenance_test tables. Automatic maintenance handles compaction, snapshot expiry, orphan file removal, and intelligent-tiering, with Amazon CloudWatch providing S3 Tables maintenance metrics.
- 5 AWS Glue Data Catalog**
Catalog integration from S3 Tables registers metadata for query engines.
- 6 Amazon Athena**
Workgroup queries S3 Tables through the Glue Data Catalog. Results stored in a dedicated S3 bucket for Athena Results & Spark Staging.
- 7 Amazon SageMaker Lakehouse & AWS Lake Formation**
Governance layer (optional) for fine-grained access control on S3 Tables data. AWS IAM roles and policies grant SSM access and Athena, Glue, and Lake Formation permissions.
- 8 Amazon EMR Serverless & Amazon Data Firehose**
Optional paths: EMR Serverless (Scenario 4b) for Spark-based processing; Data Firehose (Scenario 5) for streaming ingestion into S3 Tables.

