



Architecture Diagrams

# Scale Geospatial Data Lakes on AWS



---

# Scale Geospatial Data Lakes on AWS: Architecture Diagrams

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

---

# Table of Contents

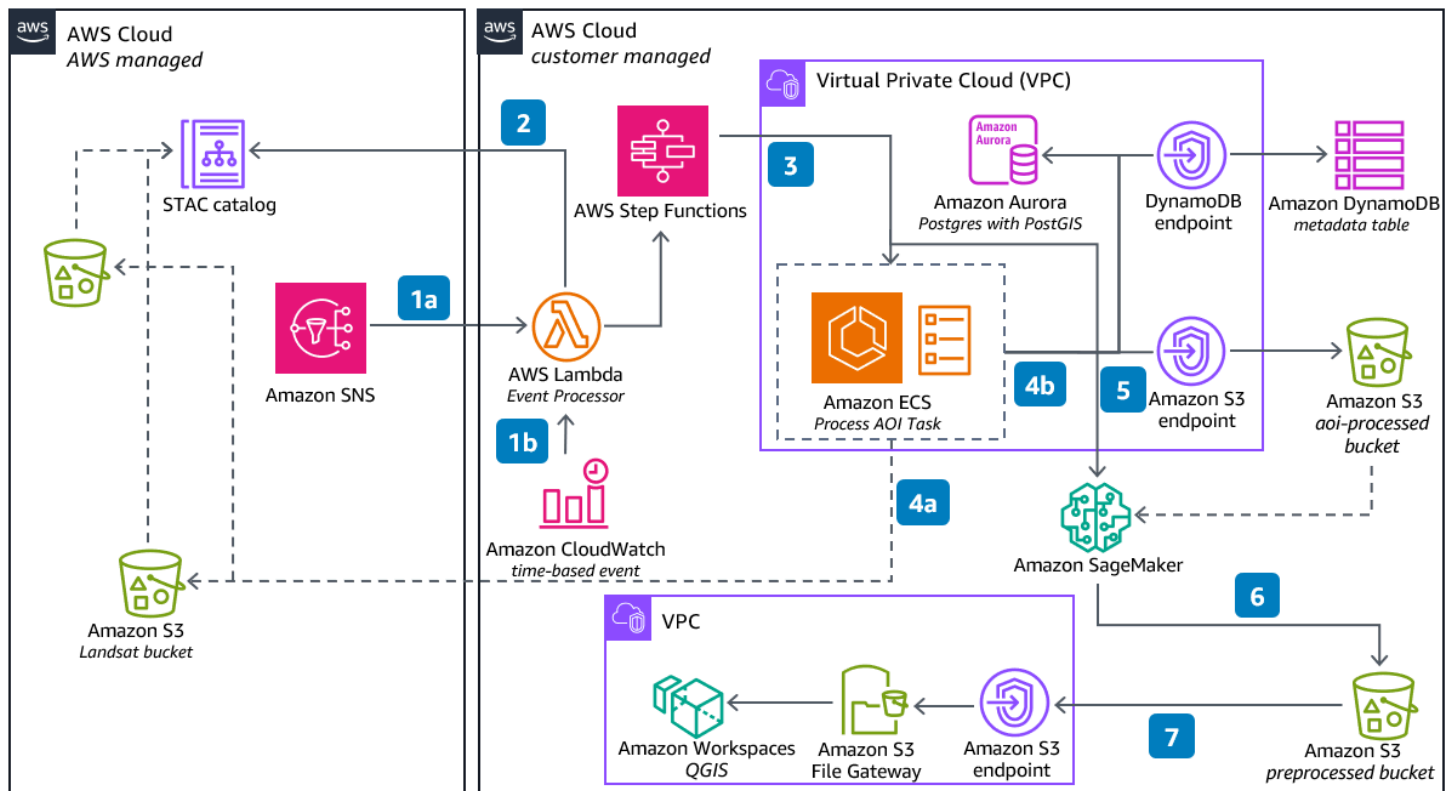
<b>Home</b> .....	<b>1</b>
Scale Geospatial Data Lakes on AWS Diagram .....	1
Download editable diagram .....	2
Create a free AWS account .....	2
Further reading .....	2
Contributors .....	3
Diagram history .....	3

# Scale Geospatial Data Lakes on AWS

Publication date: August 14, 2023 ([Diagram history](#))

Repositories of geospatial data are becoming [increasingly important](#) in many organizations – for use in everything from logistics and insurance to supply chain optimization. This reference architecture shows you how to build scalable geospatial data repositories on AWS. It simplifies the design of geospatial data pipelines, allowing accelerated access to raw data by integrating AWS-managed datasets from the [Registry of Open Data on AWS](#), eliminating the need to store it on your data lake. It integrates with a variety of dissemination mechanisms and supports diverse processing demands.

## Scale Geospatial Data Lakes on AWS Diagram



1. Initiate a data ingestion pipeline based on new scene detection. Subscribe to **Amazon Simple Notification Service** (Amazon SNS) topics for managed datasets with appropriate filters.

A time-based event can be configured using **Amazon CloudWatch** rules to begin ingestion during specific time windows.

2. **AWS Lambda** queries the [SpatioTemporal Asset Catalogs \(STAC\) API](#) for a respective dataset to get product details. It initiates the data processing pipeline through **AWS Step Functions**.
3. **Step Functions** orchestrates the processing tasks. Parallel or sequential processing can be configured based on task characteristics and requirements.
4. **Amazon Elastic Container Service** (Amazon ECS) runs containerized tasks to:
  - Download the products from datasets hosted on the Registry of Open Data on AWS.
  - Process (crop and geomosiac) the tiles to area of interest and store them in **Amazon Simple Storage Service** (Amazon S3) aoi-processed bucket. Build metadata and store them in **Amazon DynamoDB**. Store vector data in **Amazon Aurora** Postgres with PostGIS extensions.
5. **Step Functions** initiates the next processing task with **Amazon SageMaker AI**.
6. **SageMaker AI**-hosted ML models perform cloud removal and band math.
7. **Amazon WorkSpaces** hosted GIS workbenches can be used for visualization. The data is stored in an **Amazon S3** preprocessed bucket.

## Download editable diagram

To customize this reference architecture diagram based on your business needs, [download the ZIP file](#) which contains an editable PowerPoint.

## Create a free AWS account

[Sign up now](#)

Sign up for an AWS account. New accounts include 12 months of [AWS Free Tier](#) access, including the use of Amazon EC2, Amazon S3, and Amazon DynamoDB.

## Further reading

For additional information, refer to

- [AWS Architecture Icons](#)
- [AWS Architecture Center](#)
- [AWS Well-Architected](#)
- [Video, AWS Summit, Brussels 2022](#)

- [Registry of Open Data on AWS](#)
- [SpatioTemporal Asset Catalogs \(STAC\) API](#)

## Contributors

Contributors to this reference architecture diagram include:

- Ajit Rajdeosingh, Senior Solutions Architect, Amazon Web Services

## Diagram history

To be notified about updates to this reference architecture diagram, subscribe to the RSS feed.

Change	Description	Date
<a href="#">Initial publication</a>	Reference architecture diagram first published.	August 14, 2023

### Note

To subscribe to RSS updates, you must have an RSS plugin enabled for the browser you are using.