

# AWS ETL PIPELINE ARCHITECTURE

## VISUAL EXAMPLES

BY ABDUL RAYHAAN

### 1)AWS S3 WILL LOOK LIKE:

YOU WILL CREATE TWO FOLDERS FOR LOG DATA AND CSV DATA

**my-careplus-data** info

Objects Properties Permissions Metrics Management Access Points

**Objects (2)** [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">support-logs/</a>	Folder	-	-	-
<input type="checkbox"/>	<a href="#">support-tickets/</a>	Folder	-	-	-

THEN, IN EACH FOLDER YOU WILL CREATE TWO MORE FOLDERS NAMED RAW AND PROCESSED

**support-logs/** [Copy S3 URI](#)

Objects Properties

**Objects (2)** [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">processed/</a>	Folder	-	-	-
<input type="checkbox"/>	<a href="#">raw/</a>	Folder	-	-	-

WHEN YOU DONE WITH INGESTION OF DATA, YOUR RAW FOLDERS WILL HAVE .LOG AND .CSV FILES AS SHOWN BELOW

**raw/** [Copy S3 URI](#)

Objects Properties

**Objects (7)** [Copy S3 URI](#) [Copy URL](#) [Download](#) [Open](#) [Delete](#) [Actions](#) [Create folder](#) [Upload](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">support_logs_2025-07-01.log</a>	log	August 31, 2025, 19:14:56 (UTC+05:30)	32.6 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-02.log</a>	log	September 2, 2025, 13:25:12 (UTC+05:30)	34.4 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-03.log</a>	log	September 2, 2025, 13:28:19 (UTC+05:30)	26.7 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-04.log</a>	log	September 22, 2025, 13:29:14 (UTC+05:30)	26.8 KB	Standard

**raw/** Copy S3 URI

**Objects** | Properties

**Objects (3)** Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">support_tickets_2025-07-01.csv</a>	csv	August 31, 2025, 16:05:23 (UTC+05:30)	3.0 KB	Standard
<input type="checkbox"/>	<a href="#">support_tickets_2025-07-02.csv</a>	csv	September 19, 2025, 19:26:58 (UTC+05:30)	2.8 KB	Standard

WHEN YOU ARE DONE WITH USING LAMBDA FUNCTION (automate\_support\_logs , automate\_support\_tickets) AND GLUE SERVICE , YOUR DATA WILL BE TRANSFORMED AND LOADED TO .processed folder AND WILL LOOK LIKE

**processed/** Copy S3 URI

**Objects** | Properties

**Objects (7)** Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">support_logs_2025-07-01.parquet</a>	parquet	September 2, 2025, 13:04:14 (UTC+05:30)	11.4 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-02.parquet</a>	parquet	September 2, 2025, 13:25:23 (UTC+05:30)	11.5 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-03.parquet</a>	parquet	September 2, 2025, 13:28:22 (UTC+05:30)	10.8 KB	Standard
<input type="checkbox"/>	<a href="#">support_logs_2025-07-04.parquet</a>	parquet	September 22, 2025, 13:29:25 (UTC+05:30)	10.8 KB	Standard

**processed/** Copy S3 URI

**Objects** | Properties

**Objects (3)** Copy S3 URI Copy URL Download Open Delete Actions Create folder Upload

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Find objects by prefix

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	<a href="#">run-1756806209275-part-block-0-r-00000-snappy.parquet</a>	parquet	September 2, 2025, 15:13:31 (UTC+05:30)	2.4 KB	Standard
<input type="checkbox"/>	<a href="#">run-1758290286104-part-block-0-r-00000-snappy.parquet</a>	parquet	September 19, 2025, 19:28:08 (UTC+05:30)	2.5 KB	Standard
<input type="checkbox"/>	<a href="#">run-1758546972835-part-block-0-r-00000-snappy.parquet</a>	parquet	September 22, 2025, 18:46:15 (UTC+05:30)	2.3 KB	Standard

## 2)AWS LAMBDA FUNCTIONS:

YOUR AWS LAMBDA WILL LOOK LIKE THIS

☰ [Lambda](#) > Functions

**Functions (5)** Last fetched 0 seconds ago Actions Create function

🔍 Filter by attributes or search by keyword

<input type="checkbox"/>	Function name	Description	Package type	Runtime	Last modified
<input type="checkbox"/>	<a href="#">supports_log_etl</a>	-	Zip	Python 3.12	3 weeks ago
<input type="checkbox"/>	<a href="#">automate_support_logs_etl</a>	-	Zip	Python 3.12	3 weeks ago
<input type="checkbox"/>	<a href="#">incremental_automate_load_tickets</a>	-	Zip	Python 3.12	4 hours ago
<input type="checkbox"/>	<a href="#">automate_support_tickets_etl</a>	-	Zip	Python 3.13	3 days ago
<input type="checkbox"/>	<a href="#">incremental_automate_load</a>	-	Zip	Python 3.12	10 hours ago

FOR EACH LAMBDA FUNCTION CREATE AN **IAM ROLE** WITH REQUIRED PERMISSIONS AND ADD REQUIRED LAYERS , ADJUST YOUR CONFIGURATIONS AS REQUIRED AND MAKE SURE TO CREATE EVENT NOTIFICATIONS FOR EACH FUNCTIONS SO THAT THEY CAN BE TRIGGERED

### 3)EVENT NOTIFICATIONS:

aws 🔍 Search [Alt+S] Europe (Stockholm) Account ID: 9146-5494-9253 Abdul Rayhan

Amazon S3 > Buckets > my-careplus-data

**Amazon S3**

- General purpose buckets
- Directory buckets
- Table buckets
- Vector buckets
- Access Grants
- Access Points (General Purpose Buckets, FSx file systems)
- Access Points (Directory Buckets)
- Object Lambda Access Points
- Multi-Region Access Points
- Batch Operations
- IAM Access Analyzer for S3

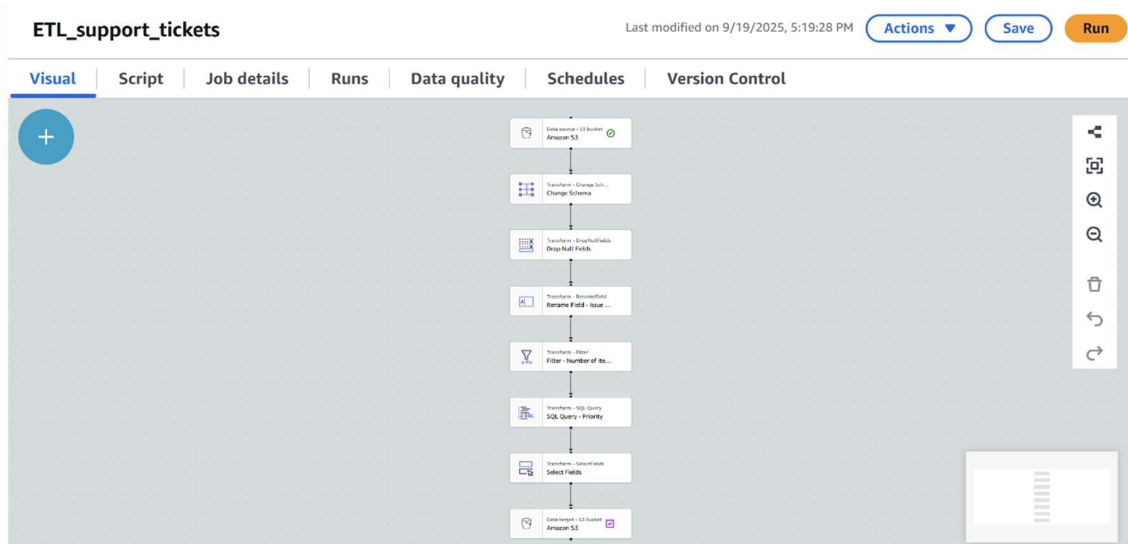
**Event notifications (4)** Edit Delete Create event notification

Send a notification when specific events occur in your bucket. [Learn more](#)

<input type="checkbox"/>	Name	Event types	Filters	Destination type	Destination
<input type="checkbox"/>	trigger_automate_etl	Put	support-logs/raw/.log	Lambda function	<a href="#">automate_suppo</a>
<input type="checkbox"/>	trigger_automatic_tickets_etl	Put	support-tickets/raw/.csv	Lambda function	<a href="#">automate_suppo</a>
<input type="checkbox"/>	trigger_increment_automate_tickets_load	Put	support-tickets/processed/.parquet	Lambda function	<a href="#">incremental_auti</a>
<input type="checkbox"/>	trigger_incremental_load	Put	support-logs/processed/.parquet	Lambda function	<a href="#">incremental_auti</a>

### 4)GLUE SERVICE:

YOU CAN USE VISUAL ETL INSTEAD OF SCRIPT AND PERFORM REQUIRED TRANSFORMATIONS, AT THE END THE TRANSFORMATION SEQUENCE WILL LOOK LIKE



## 5)IAM ROLES:

**Identity and Access Management (IAM)** > Roles

Search IAM

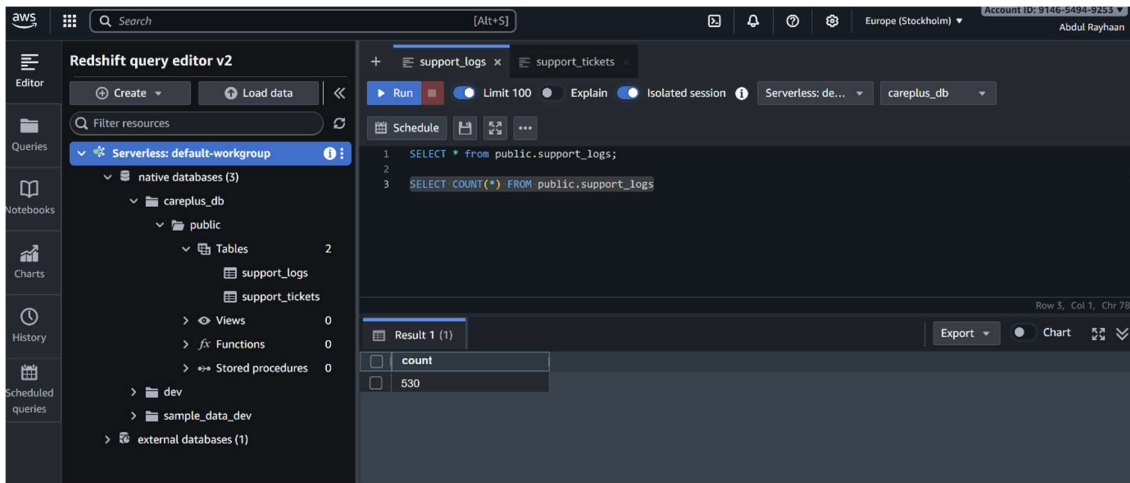
**Roles (12)** Info Delete Create role

An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.

Role name	Trusted entities	Last activity
<a href="#">AmazonRedshift-CommandsAccessRole-20250922T002127</a>	AWS Service: redshift-serverless, <a href="#">arn</a>	18 hours ago
<a href="#">automate_support_logs_etl-role-03cw3uex</a>	AWS Service: lambda	20 days ago
<a href="#">automate_support_tickets_etl-role-n9mkrcab</a>	AWS Service: lambda	3 days ago
<a href="#">AWSGlueServiceRole-support_logs</a>	AWS Service: glue	2 days ago
<a href="#">AWSGlueServiceRole-support_tickets</a>	AWS Service: glue	2 days ago
<a href="#">AWSServiceRoleForRedshift</a>	AWS Service: redshift (Service-Linker)	7 hours ago
<a href="#">AWSServiceRoleForSupport</a>	AWS Service: support (Service-Linker)	-
<a href="#">AWSServiceRoleForTrustedAdvisor</a>	AWS Service: trustedadvisor (Service-Linker)	-
<a href="#">glue_support_tickets</a>	AWS Service: glue	3 days ago
<a href="#">incremental_automate_load_tickets-role-54qffb4n</a>	AWS Service: lambda	1 hour ago
<a href="#">incremental_automate_load-role-nhbw0mkz</a>	AWS Service: lambda	-

## 6)AWS REDSHIFT:

IN AWS REDSHIFT, AFTER YOU HAVE CREATED DATABASE AND LOADED TRANSFORMED DATA TO REDSHIFT, YOU WILL SEE SOMETHING LIKE THIS



## 7) POWERBI VISUALS

