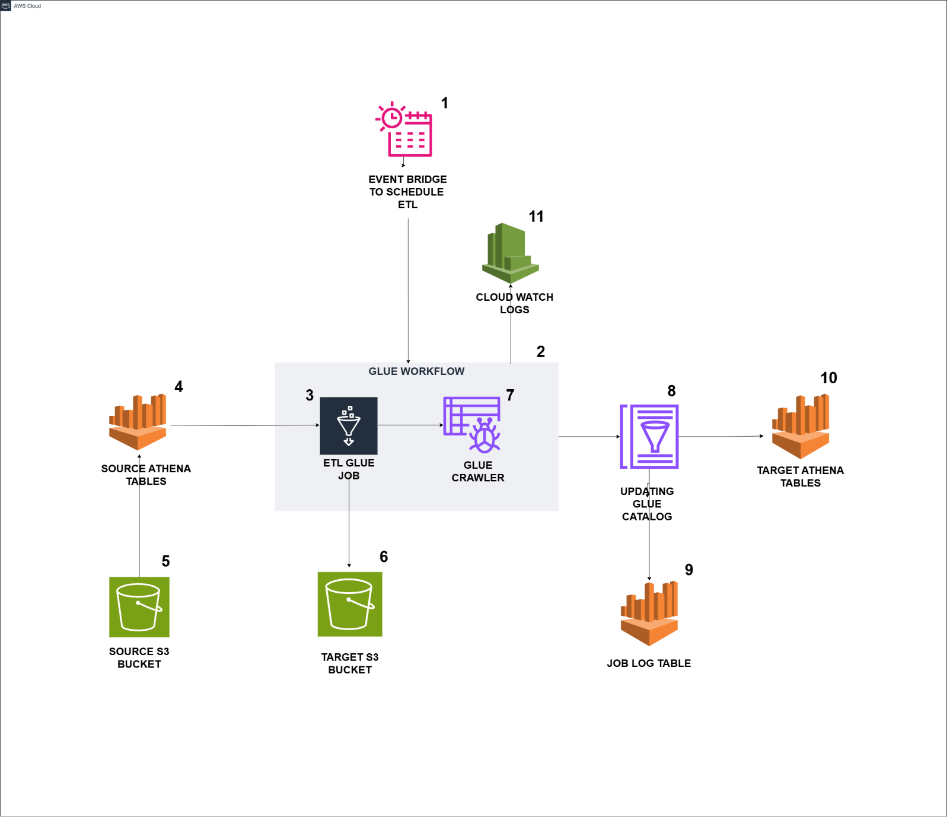
**CIAM Project Help Document**

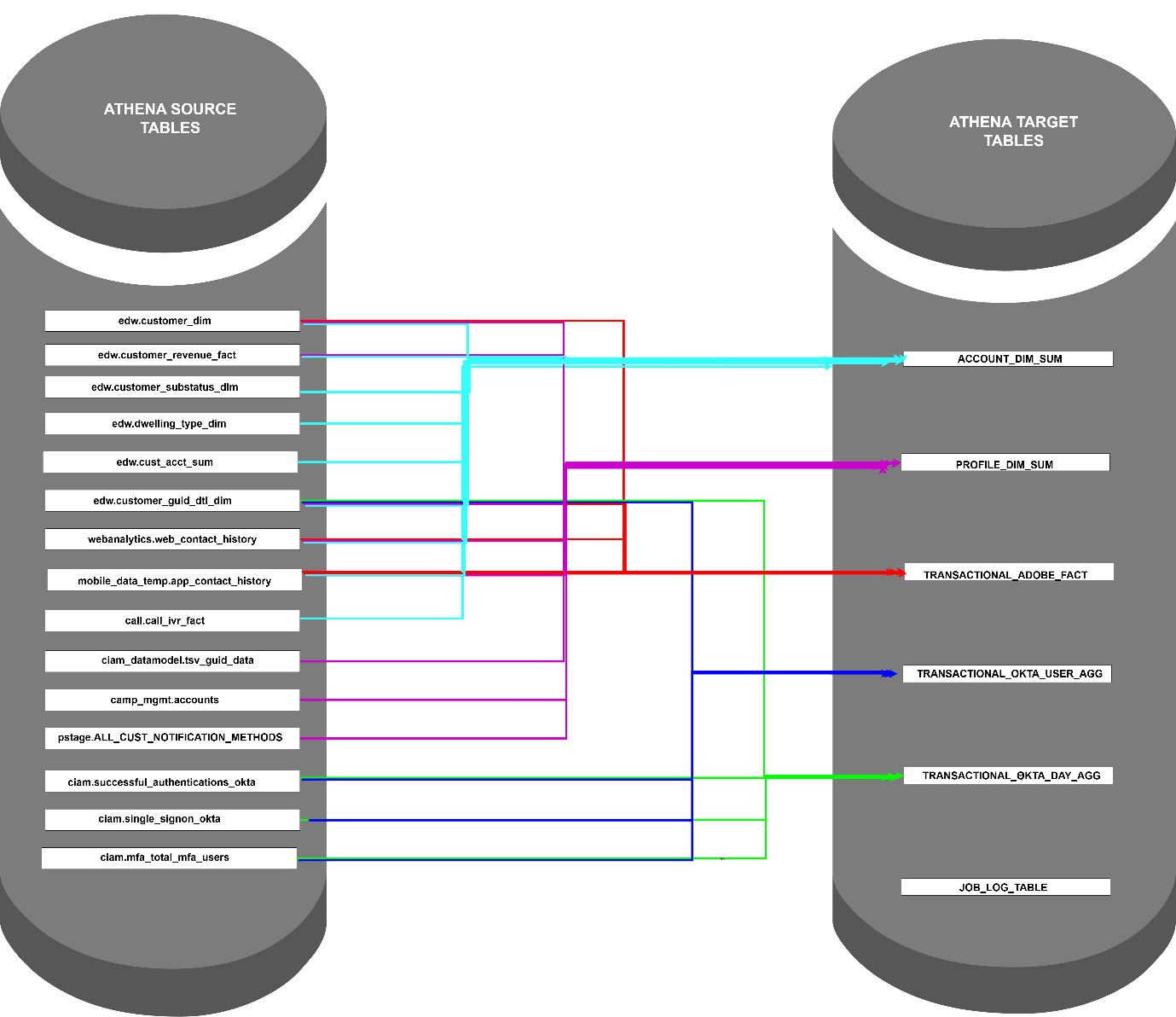
**ARICHITECTURE DIAGRAM FOR GLUE:**



**DESCRIPTION:**

1. EventBridge triggers the Glue workflow based on the scheduled time.
2. The Glue workflow is triggered by EventBridge and initiates the run.
3. A Glue job starts running within the Glue workflow using an On-Demand trigger.
4. The Glue job executes Athena queries to read data from the source tables.
5. The Glue job fetches data from the source S3 path based on the queries executed in Athena.
6. The data from the source tables is loaded into the target S3 path as Parquet files, organized into partitioned folders.
7. Upon successful completion of the Glue job, the Glue crawler is triggered and starts running.
8. The job log table is created, containing runtime information and table counts.
9. The target tables are available in Athena, where the data can be queried.
10. CloudWatch stores logs for both the Glue jobs and the Glue crawler.

**DATA FLOW DIAGRAM:**



**DESCRIPTION:**

**Source Databases:**

* edw
* webanalytics
* mobile\_data\_temp
* call
* ciam
* camp\_mgmt
* pstage
* ciam\_datamodel

**Source Tables :**

* edw.customer\_dim
* edw.customer\_revenue\_fact
* edw.customer\_substatus\_dim
* edw.dwelling\_type\_dim
* edw.cust\_acct\_sum
* edw.customer\_guid\_dtl\_dim
* webanalytics.web\_contact\_history
* mobile\_data\_temp.app\_contact\_history
* call.call\_ivr\_fact
* ciam\_datamodel.tsv\_guid\_data
* camp\_mgmt.accounts
* pstage.ALL\_CUST\_NOTIFICATION\_METHODS
* ciam.successful\_authentications\_okta
* ciam.single\_signon\_okta
* ciam.mfa\_total\_mfa\_users

**Target Database:** ciam\_datamodel

**Target Tables:**

* account\_dim\_sum
* profile\_dim\_sum
* transaction\_adobe\_fact
* transaction\_okta\_use\_agg
* transaction\_okta\_day\_agg
* job\_log\_table

**Database**: ciam\_datamodel

**Athena Tables -** For Digital Adoption Analysis

* digital\_adoption\_omni\_intent
* digital\_adoption\_omni\_intent1
* digital\_adoption\_omni\_intent\_detail
* digital\_adoption\_omni\_intent\_detail\_final\_1
* digital\_adoption\_omni\_intent\_final
* digital\_adoption\_omni\_intent\_final\_1
* digital\_adoption\_omni\_inter\_intent
* digital\_adoption\_omni\_inter\_intent1
* digital\_adoption\_omni\_inter\_intent\_detail
* digital\_adoption\_omni\_inter\_intent\_detail\_final\_1
* digital\_adoption\_omni\_inter\_intent\_final
* digital\_adoption\_omni\_inter\_intent\_final\_1

**GLUE JOB FUNCTIONALITIES:**

* When the job starts running, it checks if there is a JOB\_LOG\_TABLE in the target database.
* If the JOB\_LOG\_TABLE is not present or if it does not have an entry for 13\_MONTHS\_LOAD, the job will directly read the last 13 months of data from multiple source tables and load it into the target S3 path in Parquet file format. The data will be partitioned by a specific column for each target table.
* If the JOB\_LOG\_TABLE is present, the job will check the LOAD\_TYPE column. If the LOAD\_TYPE column contains an entry for 13\_MONTHS\_LOAD, the job will read only the latest current month’s data from the source tables and overwrite the corresponding partitioned folder in S3.
* Each time the job runs, it will create an entry in the JOB\_LOG\_TABLE that includes the table counts, start time, end time, and runtime details of the Glue job.
* After loading the data into the target S3 path, the Glue job will check if there is more than 14 months' worth of data. If so, it will delete the oldest month's data. Otherwise, it will skip the deletion.
* A Glue crawler has been created and configured with all the target S3 paths and the target database name.
* After the successful completion of the Glue job, the crawler will be triggered once to populate the tables in Athena.

**GLUE JOB RUNTIME** :

* For 13 months - 1 hr 6 mins
* For latest current months - 22 mins

**CIAM – CUSTOMER PROFILE AND ACTIVITY DASHBOARD:**

**CUSTOMER VIEW:**

* These reports provide a snapshot of customer growth, retention, and churn.
* It tracks the total customer base, new sign-ups, and registration completion to understand engagement levels.
* The churn analysis helps identify when customers are most likely to leave, while the tenure breakdown highlights customer loyalty over time.
* Product usage insights show which services are most popular among customers.

**PROFILE VIEW**

* These reports give an overview of user profiles and their verification status.
* It tracks total users, their account type, and profile completion to understand how engaged and verified they are.
* TSV (Two-Step Verification) adoption helps measure security awareness among users, showing how many have enabled additional verification methods.
* The breakdown of verification methods used (SMS, Email, Call) provides insights into preferred security options.

**CUSTOMER CONTACT VIEW:**

* These reports focuses on how customers are contacted and their communication preferences
* It tracks the percentage of users with verified email and phone numbers, as well as those who have opted out of communication.
* It also highlights potential issues with customer reachability, such as bounced emails and undelivered SMS messages.
* The single vs. multiple email analysis provides insights into how many accounts have multiple email addresses and single email address

**CUSTOMER ENGAGEMENT:**

* These reports highlights customer interactions across different channels over time. It tracks engagement through apps, communication platforms, and IVR systems, showing trends over the last 12 months and 30 days.
* It also provides insights into how recently customers have been contacted, helping identify those who are engaged frequently versus those who haven't been reached in a long time or ever

**ACTIVITY Tab:**

**REGISTRATION INSIGHTS:**

* These reports track customer registration attempts, highlighting success rates, It helps identify trends in registration performance
* Additionally, it analyzes new registrations based on customer tenure, showing when users sign up after joining. This helps in understanding user behavior and optimizing the registration process to improve conversions.

**AUTHENTICATION INSIGHTS:**

* These reports track how users authenticate across different channels and applications. It includes the total number of authentication attempts, the percentage of successful authentications
* This helps in understanding user behavior and identifying areas for improvement in the authentication process.

**CREDENTIAL RECOVERY OVERVIEW:**

* These reports provide insights into how users recover their credentials, including forgotten passwords, user IDs, and authentication IDs. It tracks the number of recovery attempts, the success rate
* it highlights the different recovery methods used, such as calls, emails, and text messages, and shows how often users attempt single versus multiple recovery attempts. This data helps in identifying trends and improving the user experience for credential recovery.

**TSV AUTHENTICATION OVERVIEW:**

* These reports track the verification and reset attempts for Two-Step Verification (TSV). It measures the number of attempts, the success rate, and the volume of successful verifications and resets.
* This data helps in understanding authentication trends and identifying areas for improvement in the verification and reset experience.

**PREFERENCE & NOTIFICATION OVERVIEW:**

* This report tracks user engagement with notification and preference settings. The **Preference Engagement Rate** measures how often users interact with and update their notification preferences.