

Stream Data Ingestion using SQS, Eventbridge, Lambda and S3

Objective

Build a simulated data pipeline for booking data that integrates various AWS services, demonstrating real-time data processing, filtering, and storage.

Step 1: Create 3 IAM roles- one for Lambda to access SQS, one for Eventbridge to read from SQS and trigger Lambda and lastly for lambda to write the filtered records to S3

Step 2: Create an Amazon SQS Queue with DLQ

- Create an SQS Standard Queue named AirbnbBookingQueue. Setup a Dead Letter Queue (DLQ): Create another SQS queue named AirbnbBookingDLQ. Configure the AirbnbBookingQueue to send messages to AirbnbBookingDLQ after 3 unsuccessful delivery attempts.

Amazon SQS > Queues

① SQS fair queues is now available for standard queues.
Automatically manage noisy neighbors in your queues with fair queues, a new feature that limits noisy neighbor impact across all message groups. Add a group identifier to your messages, and SQS re-orders messages to ensure no single tenant impact the time in queue for any other tenants.

Queues (2)

Search queues by prefix

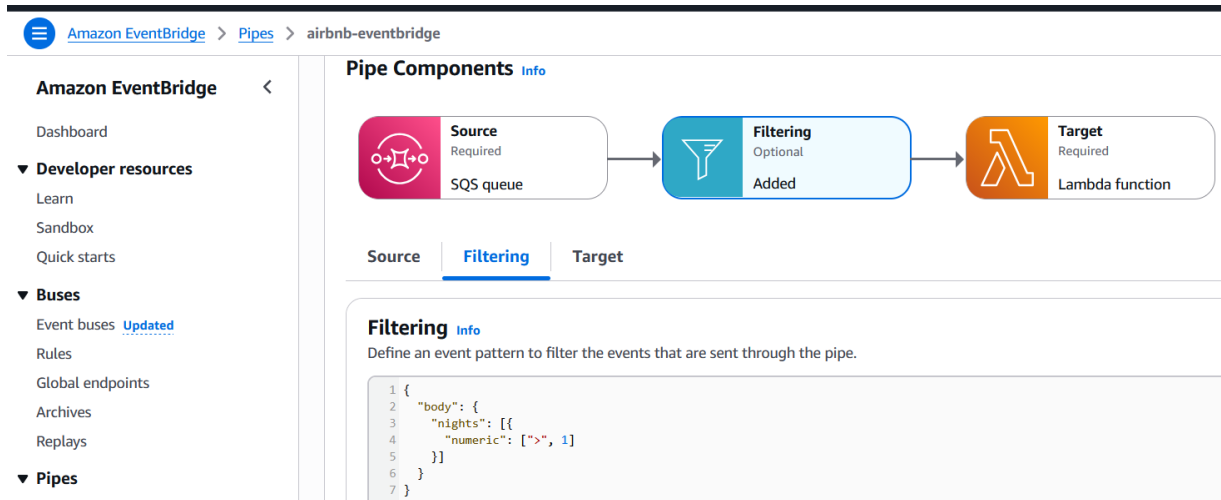
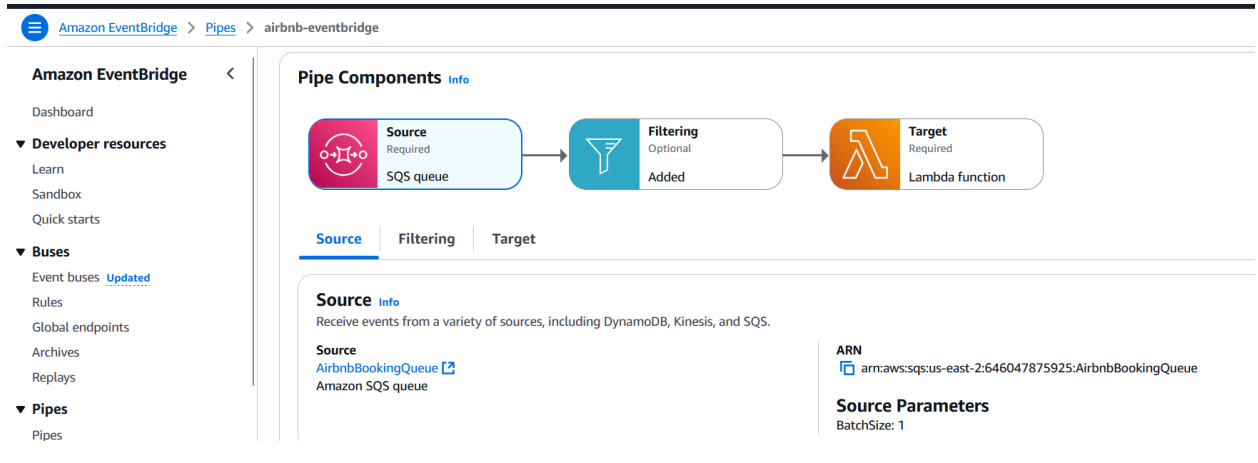
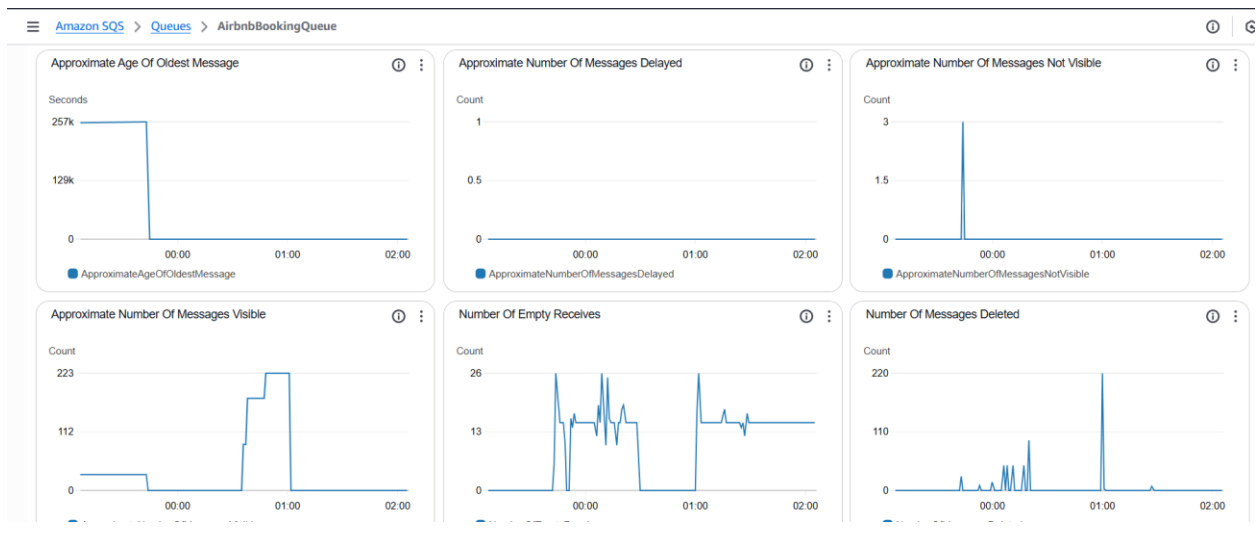
	Name	Type	Created	Messages available	Messages in flight	Encryption	Content-
	AirbnbBookingDLQ	Standard	2025-08-12T17:54-04:00	3	0	Amazon SQS key (SSE-SQS)	-
	AirbnbBookingQueue	Standard	2025-08-12T17:54-04:00	0	0	Amazon SQS key (SSE-SQS)	-

Step 3: Create Producer Lambda Function

- Lambda Function - Producer: Create a Lambda function named ProduceAirbnbBookingData. This function will generate Airbnb booking data and publish it to AirbnbBookingQueue.

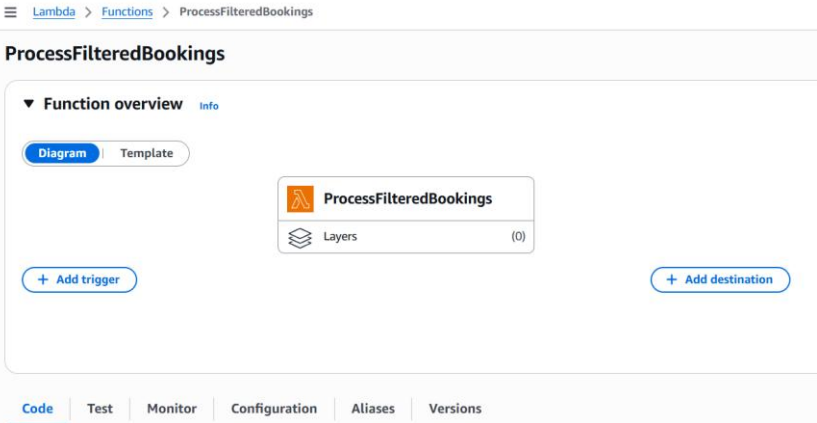
Step 4: Setup EventBridge Pipe

- EventBridge Pipe: Create an EventBridge Pipe to consume messages from AirbnbBookingQueue. Filter messages where the booking duration is more than 1 day.



Step 5: Create Destination Lambda Function

- **Lambda Function - Consumer:** Create a Lambda function named `ProcessFilteredBookings`. This function will be triggered by the EventBridge Pipe and will write the filtered records to an S3 bucket.



- S3 Bucket: Ensure the S3 bucket is created beforehand to store the booking records. The bucket can be named airbnb-booking-records, The data gets stored in hive partition format

