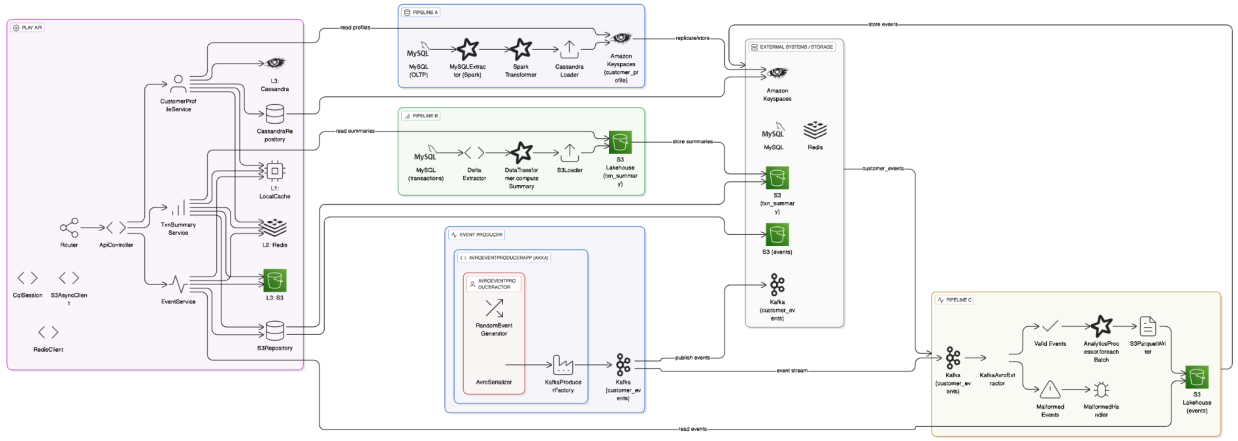
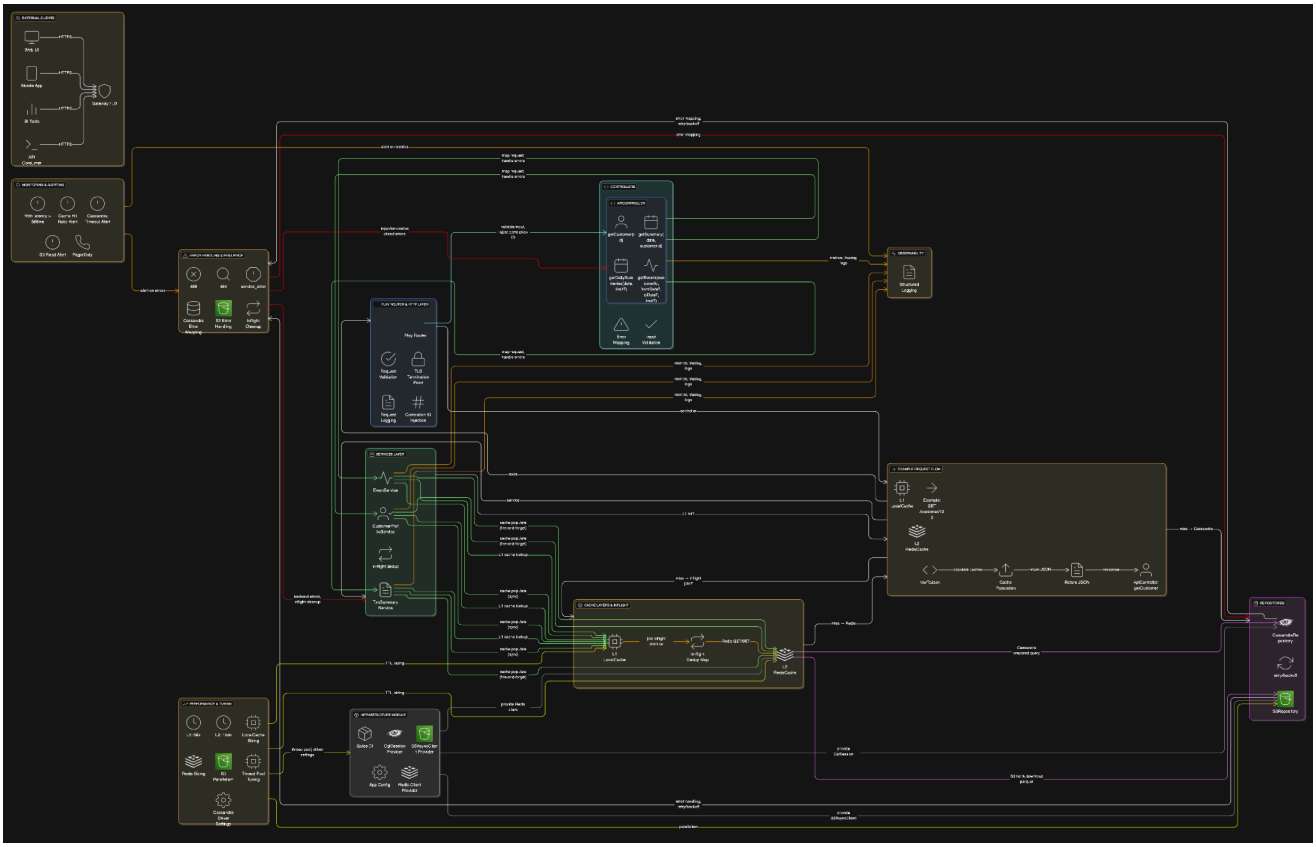
Customer Data Engineering

Owner: Sanjeev Kumar V
Date: Dec 12, 2025

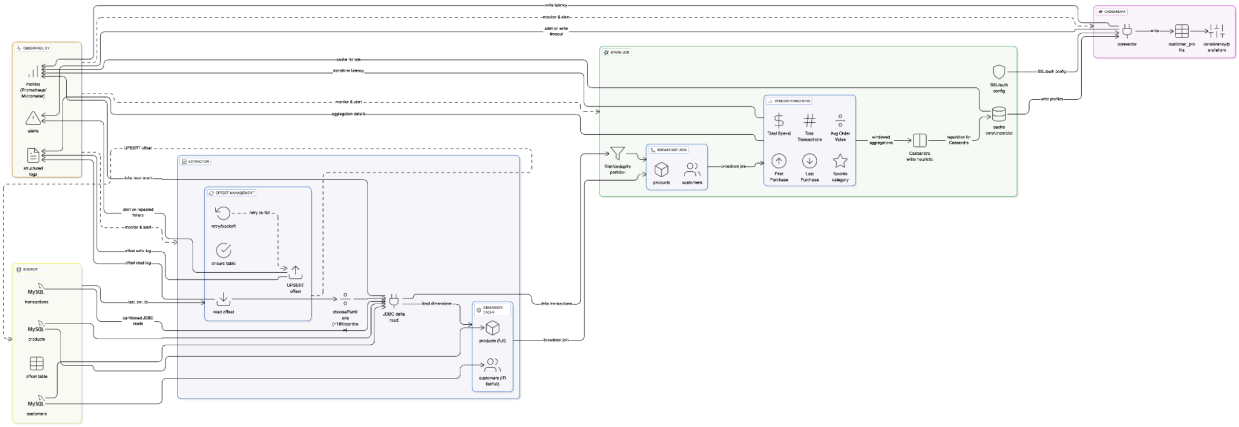
System Architecture:



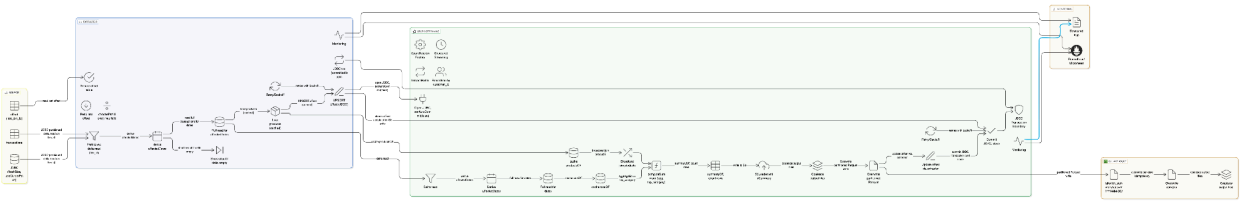
Play API Architecture



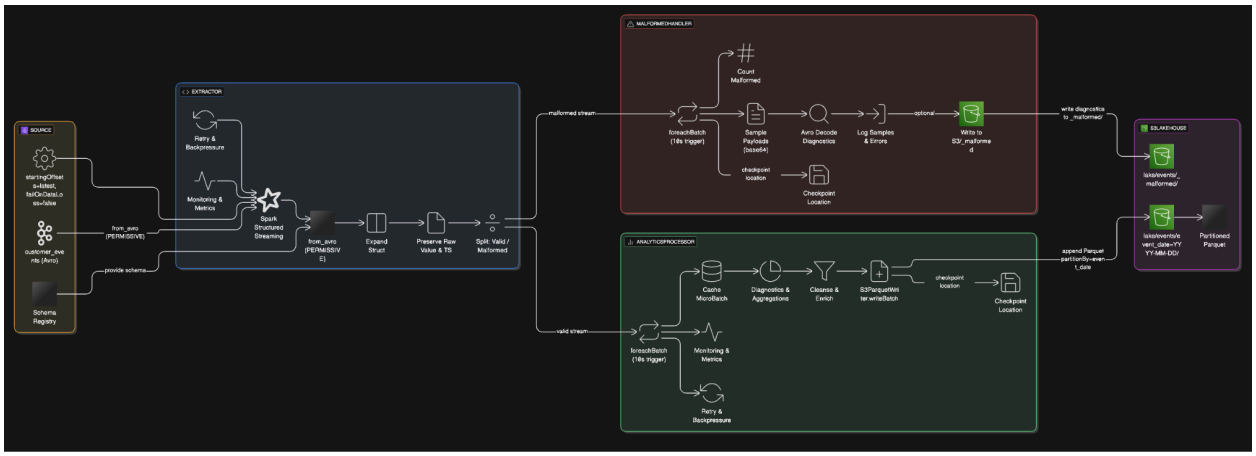
Pipeline A Architecture:



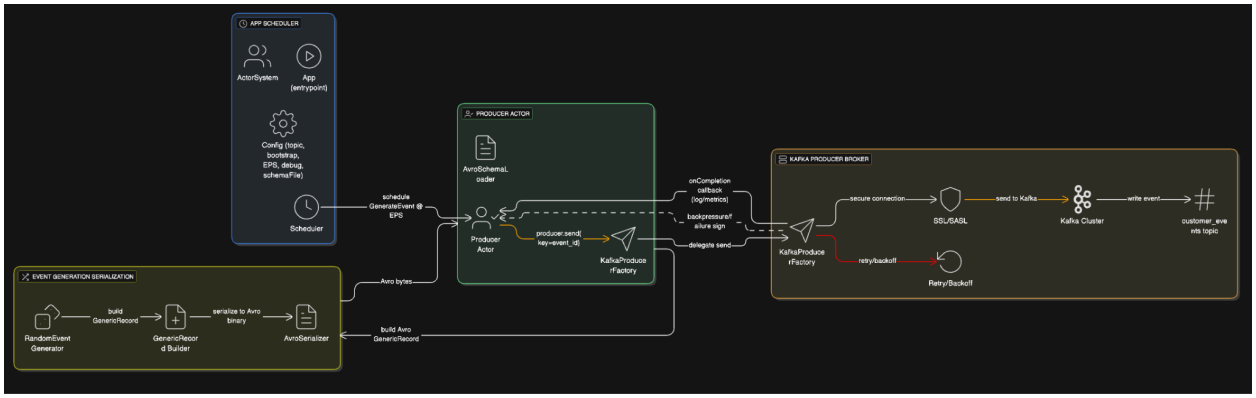
Pipeline B Architecture:



Pipeline C Architecture:



Producer Architecture:



DATA DICTIONARY

MySQL - Source System

Table: customers

This table stores all customer master data.

Field	Type	Description
customer_id	INT (PK)	Unique customer identifier
name	VARCHAR(100)	Customer full name
email	VARCHAR(150), UNIQUE	Customer email (unique across DB)
gender	ENUM('M','F','O')	Gender of the customer
signup_date	DATE	Date when customer account was created

Table: products

Stores details of all products purchased in transactions.

Field	Type	Description
product_id	INT (PK)	Unique product identifier
name	VARCHAR(150)	Product name

category	VARCHAR(50)	Product category (Electronics, Grocery, Sports…)
price	DECIMAL(10,2)	Unit price of the product

Table: transactions

This is the core **fact table** for Pipelines A & B (profiles + daily summaries).

Field	Type	Description
txn_id	BIGINT (PK)	Unique transaction ID
customer_id	INT (FK)	FK → customers.customer_id
product_id	INT (FK)	FK → products.product_id
qty	INT	Number of items purchased
amount	DECIMAL(10,2)	Total amount of transaction (qty × price)
txn_timestamp	TIMESTAMP	Purchase timestamp

Offset Tracking Tables (Used by Pipelines A & B)

Table: profile_offsets

Used by **Pipeline A (MySQL → Cassandra)** to ensure incremental processing.

Field	Type	Description
source	VARCHAR(128)	Identifier for the pipeline source (e.g., "transactions")
last_txn_id	BIGINT	Last processed txn_id for profile updates

Table: txn_summary_offsets

Used by **Pipeline B (MySQL → S3 Daily Summaries)** for incremental loads.

Field	Type	Description

source	VARCHAR(128)	Identifier for the pipeline source
last_txn_id	BIGINT	Last processed txn_id for summary generation

---

Kafka Event Schema (Pipeline C Input)

(Event Producer → Kafka → Pipeline C → S3)

Field	Type	Description
event_id	STRING	Unique event UUID
customer_id	INT	Customer who performed the event
event_type	STRING	Behaviour type: WISHLIST, LIKE, CART_ADD.
product_id	INT NULL	Optional product reference
event_timestamp	STRING	ISO timestamp assigned at event creation

---

Cassandra Table (Pipeline A Output)

Keyspaces Table: customer_profile

Stores aggregated customer analytics.

Column	Type	Description
customer_id	INT (PK)	Unique ID
name	TEXT	Customer name
email	TEXT	Email
gender	TEXT	Gender
total_spend	DECIMAL	SUM(amount)
total_transactions	INT	COUNT(*)
avg_order_value	DECIMAL	total_spend / total_transactions
first_purchase	TIMESTAMP	MIN(txn_timestamp)

last_purchase	TIMESTAMP	MAX(txn_timestamp)
favorite_category	TEXT	Category with highest frequency

S3 Daily Summary Schema (Pipeline B Output)

Partition folder: lake/txn_summary/date=YYYY-MM-DD/*.parquet

Column	Type	Description
date	STRING	Partition key
customer_id	INT	Customer ID
total_amount	DOUBLE	Daily spend
total_items	INT	Total quantity
distinct_products	INT	Count of unique products
top_category	STRING	Most purchased category

2.6 S3 Events Schema (Pipeline C Output)

Partition folder: lake/events/event_date=YYYY-MM-DD/*.parquet

Column	Type	Description
event_id	STRING	Event identifier
customer_id	INT	Customer
event_type	STRING	Behaviour
product_id	INT	Nullable
event_timestamp	BINARY (INT96)	Original event timestamp
ingestion_timestamp	BINARY (INT96)	Ingestion time recorded by Spark

Pipeline-quality checks report

Pipeline A – MySQL → Cassandra (Customer Profile ETL)

1. Input Data Quality Checks

Check Type	Description
Primary Key Integrity	Ensures `customer_id` is not null and unique.
Email Format Validation	Rejects invalid or malformed email addresses.
Gender Enum Validation	Only accepts `M`, `F`, `O`.
Date Validation	Ensures `signup_date` is a valid ISO date.

---

2. Schema & Transformation Quality Checks

Check	Description
Column completeness	Ensures all required fields exist before transformation.
Type consistency	Converts MySQL types → Cassandra compatible types (e.g., VARCHAR → VARCHAR, DATE → DATE).
Null handling	Replaces unexpected nulls with defaults/logging warnings.

---

3. Offset / Incremental Load Checks

Check	Description
Offset table validation (`profile_offsets`)	Reads last processed `txn_id` before each run.
Monotonicity check	Ensures new `txn_id` > stored offset.
Duplicate prevention	Avoids re-processing already consumed rows.

---

4. Write Quality Checks (Cassandra)

Check	Description
Consistency level validation	Ensures write acknowledgements success.
Timeout handling	Retries on Cassandra `ReadTimeout`, `WriteTimeout`, `Unavailable` errors.
Row-level validation after write	Optional: read-back check (configurable).

Pipeline B – MySQL → S3 Lakehouse (Daily Transaction Summary)

1. Input Data Quality Checks

Check Type	Description
Primary key validation	Ensures `txn_id` exists and is unique.
Foreign key integrity	Valid `customer_id` and `product_id` reference.
Amount and Quantity check	Must be positive numeric values.
Timestamp validity	Ensures `txn_timestamp` is a valid UTC timestamp.

---

2. Aggregation Quality Checks

Check	Description
Group-by consistency	Ensures summaries are calculated correctly per customer per day.
Total amounts validation	No negative totals; amounts match sum(qty * price).

Distinct product count check	Ensures distinct count uses correct product_id.

---

3. Offset Checks (`txn_summary_offsets`)

Check	Description
Last processed txn_id validation	Ensures incremental summary updates.
No missing transactions	Detects gaps in txn_id sequence.

---

4. Write Quality Checks (S3 Parquet)

Check	Description
Partition validation	Must write to: `lake/txn_summary/date=YYYY-MM-DD/`
Schema consistency	Enforces Parquet schema: `{customer_id, total_amount, total_items, distinct_products, top_category}`
Atomic batch writes	Uses temporary files + rename for safe writes.
File-level validation	Validates Parquet file is readable after write.

Pipeline C – Kafka Avro → S3 Events Lake (Valid + Malformed Streams)

1. Kafka Input Quality Checks

Check	Description
Avro schema validation	Deserializes using schema registry with `PERMISSIVE` mode.
Required field checks	Ensures `event_id` & `event_type` must exist.
Event timestamp validation	Parses event timestamp or falls back to Kafka timestamp.

Malformed event detection	Missing keys → automatically routed to `malformed` stream.

---

2. Transformation Quality Checks

Check	Description
Product ID correction	Invalid product_id (<0) → null.
Timestamp normalization	Converts string timestamp → Spark timestamp.
Partition column generation	Adds `event_date` for partitioning.
Lineage metadata	Adds `ingestion_timestamp`.

---

3. Malformed Events Checks

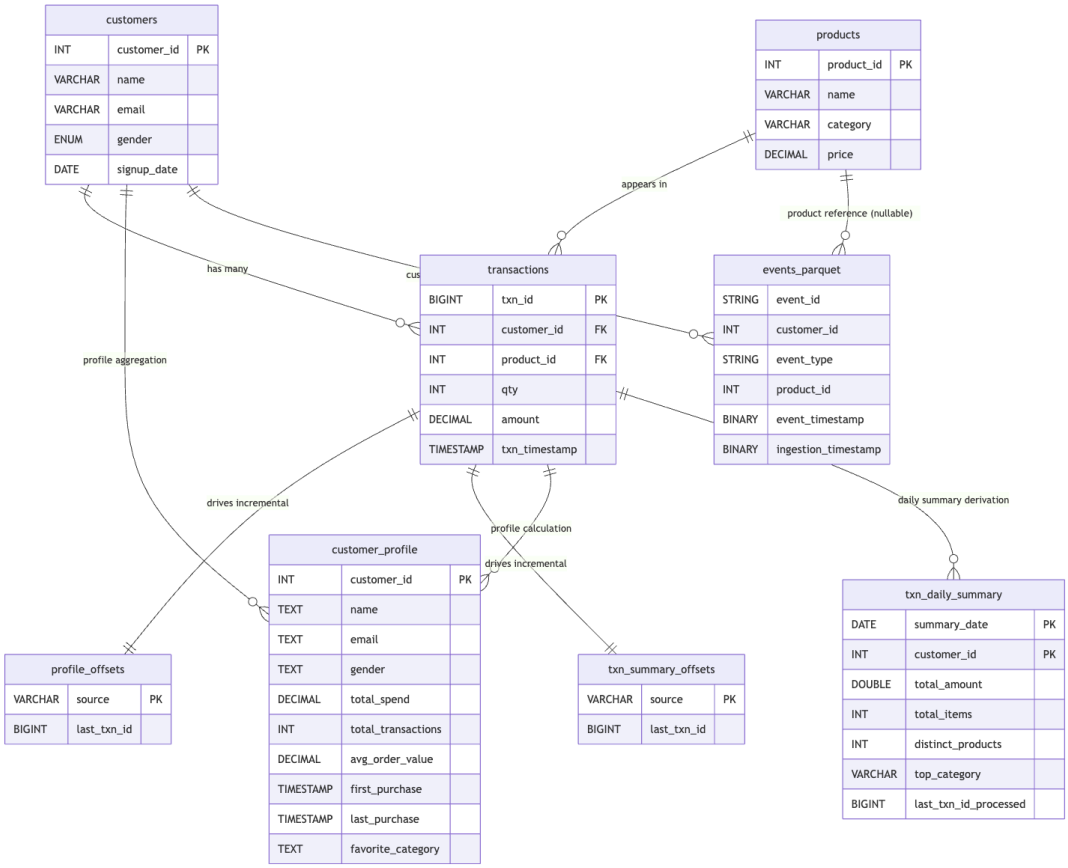
Check	Description
Sample logging	Logs up to 5 Base64 samples.
Avro decode attempt	Logs decoding failures for analysis.
Count mismatch detection	Alerts if malformed ratio exceeds threshold.

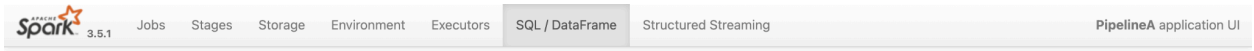
---

4. Write Quality Checks (S3 Parquet)

Check	Description
Partitioning validation	Writes to: `lake/events/event_date=YYYY-MM-DD/`
Schema consistency	Ensures event schema is stable across writes.
Exactly-once semantics	Achieved via streaming checkpoints.

Fail-safe writes	Writes via `foreachBatch` with exception handling.





SQL / DataFrame

Completed Queries: 7

Completed Queries (7)

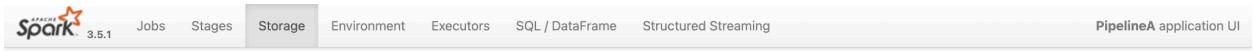
Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go

ID	Description	Submitted	Duration	Job IDs	Sub Execution IDs
16	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:18:00	3 s		
15	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:17:00	4 s		
14	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:16:00	3 s		
8	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:15:00	27 s	[33][34][35][36][37][38][39][40][41][42][43][44][45][46]	[9][10][11][12][13] +details
7	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:14:00	3 s		
6	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:13:11	3 s		
0	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e4... +details	2025/12/12 16:12:30	41 s	[7][8][9][10][11][12][13][14][15][16][17][18][19][20][21][22]	[1][2][3][4][5] +details

Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go



Storage

RDDs

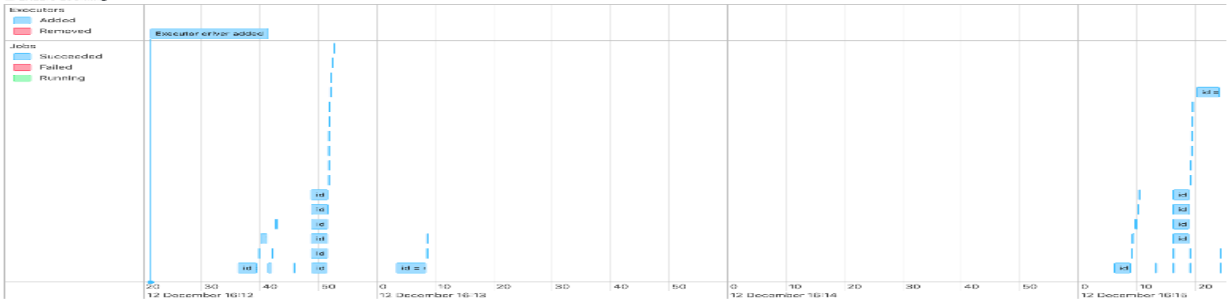
ID	RDD Name	Storage Level	Cached Partitions	Fraction Cached	Size in Memory	Size on Disk
44	*(1) Scan JDBCRelation(products) [numPartitions=1] [product_id#0,category#2] PushedFilters: [], ReadSchema: struct<product_id:int,category:string>	Disk Memory Deserialized 1x Replicated	1	100.00%	2.0 KiB	0.0 B

Spark Jobs (?)

User: rac
Total Uptime: 3.9 min
Scheduling Mode: FIFO
Completed Jobs: 50

▾ Event Timeline

☐ Enable zooming



▾ Completed Jobs (50)

Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go

Job Id (Job Group) ▾	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
49 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:24	6 ms	1/1 (1 skipped)	1/1 (2 skipped)
48 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:24	10 ms	1/1	2/2
47 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:20	4 s	1/1 (5 skipped)	128/128 (5 skipped)
46 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	0.2 s	1/1 (5 skipped)	128/128 (5 skipped)
45 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	56 ms	1/1 (4 skipped)	1/1 (4 skipped)
44 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	78 ms	1/1 (1 skipped)	1/1 (1 skipped)
43 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	91 ms	1/1 (1 skipped)	1/1 (1 skipped)
42 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	0.1 s	1/1	1/1
41 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	0.1 s	1/1	1/1
40 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:19	63 ms	1/1	1/1
39 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:18	15 ms	1/1	1/1
38 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	11 ms	1/1	1/1
37 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	11 ms	1/1	1/1
36 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	3 s	1/1	1/1
35 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	3 s	1/1	1/1
34 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	3 s	1/1	1/1
33 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:16	3 s	1/1	1/1
32 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:13	0.2 s	1/1 (1 skipped)	200/200 (2 skipped)
31 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:10	0.3 s	1/1 (1 skipped)	200/200 (2 skipped)
30 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:10	0.1 s	1/1 (2 skipped)	1/1 (202 skipped)
29 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:09	0.5 s	1/1 (1 skipped)	200/200 (2 skipped)
28 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:09	0.5 s	1/1 (1 skipped)	200/200 (2 skipped)
27 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:09	0.1 s	1/1	2/2
26 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:15:06	3 s	1/1	2/2
25 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:13:29	13 ms	1/1 (1 skipped)	1/1 (2 skipped)
24 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:13:29	28 ms	1/1	2/2
23 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:13:23	5 s	1/1 (5 skipped)	128/128 (5 skipped)
22 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:52	0.3 s	1/1 (5 skipped)	128/128 (5 skipped)
21 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:52	0.3 s	1/1 (4 skipped)	1/1 (4 skipped)
20 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:52	97 ms	1/1 (1 skipped)	1/1 (1 skipped)
19 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	0.1 s	1/1 (1 skipped)	1/1 (1 skipped)
18 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	0.2 s	1/1	1/1
17 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	0.2 s	1/1	1/1
16 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	16 ms	1/1	1/1
15 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	36 ms	1/1	1/1
14 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	29 ms	1/1	1/1
13 (a12a1c35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:51	47 ms	1/1	1/1
12 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
11 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
10 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
9 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
8 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
7 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:48	3 s	1/1	1/1
6 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:46	0.3 s	1/1 (1 skipped)	200/200 (2 skipped)
5 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:42	0.4 s	1/1 (1 skipped)	200/200 (2 skipped)
4 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:42	54 ms	1/1 (2 skipped)	1/1 (202 skipped)
3 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:41	0.7 s	1/1 (1 skipped)	200/200 (2 skipped)
2 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:40	1 s	1/1 (1 skipped)	200/200 (2 skipped)
1 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:39	0.3 s	1/1	2/2
0 (a12a1d35-d35f-4a96-aa42-96e40537d120)	id = 4b78e16e-d9a3-495d-98a4-45bbdb113ba8 runId = a12a1d35-d35f-4a96-aa42-96e... start at ProfileConsolidation.Job.scala:267	2025/12/12 16:12:36	3 s	1/1	2/2

Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go

Stages for All Jobs

Completed Stages: 50
Skipped Stages: 46

Completed Stages (50)

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

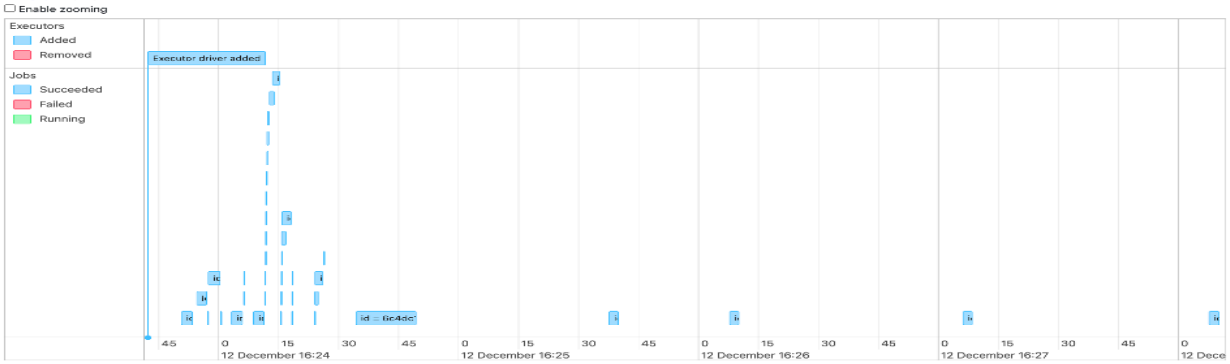
Stage Id	Description	Submitted	Duration	Tasks: Succeeded/Total	Input	Output	Shuffle Read	Shuffle Write
95	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:24	3 ms	1/1			118.0 B	
93	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:24	8 ms	2/2	6.7 KiB			118.0 B
92	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:20	4 s	128/128	167.7 KiB			
88	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	0.1 s	128/128			26.7 KiB	
83	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	51 ms	1/1			83.2 KiB	24.7 KiB
78	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	75 ms	1/1			83.6 KiB	26.9 KiB
73	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	85 ms	1/1			98.4 KiB	36.3 KiB
71	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	0.1 s	1/1	64.6 KiB			83.8 KiB
70	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	0.1 s	1/1	64.6 KiB			98.4 KiB
69	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:19	50 ms	1/1	10.3 KiB			
68	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:18	10 ms	1/1	10.3 KiB			
67	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	8 ms	1/1	2.0 KiB			
66	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	10 ms	1/1	2.0 KiB			
65	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	3 s	1/1	10.3 KiB			
64	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	3 s	1/1	64.6 KiB			
63	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	3 s	1/1				
62	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:16	3 s	1/1				
61	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:10	0.7 s	200/200	32.9 KiB			
59	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:10	0.8 s	200/200	32.9 KiB			
57	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:10	0.1 s	1/1			11.3 KiB	
54	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:09	0.6 s	200/200	32.9 KiB			11.3 KiB
52	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:09	0.5 s	200/200			9.5 KiB	
50	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:09	0.1 s	2/2	6.7 KiB			9.5 KiB
49	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:15:06	3 s	2/2				
48	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:13:08	7 ms	1/1			118.0 B	
46	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:13:08	19 ms	2/2	6.5 KiB			118.0 B
45	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:13:03	5 s	128/128	185.7 KiB			
39	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:52	0.2 s	128/128			24.7 KiB	
33	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:52	0.3 s	1/1			81.2 KiB	24.7 KiB
28	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:52	89 ms	1/1			96.7 KiB	54.9 KiB
26	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	0.1 s	1/1			61.8 KiB	26.3 KiB
24	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	0.7 s	1/1	62.6 KiB			96.7 KiB
23	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	0.2 s	1/1	62.6 KiB			61.8 KiB
22	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	10 ms	1/1	10.2 KiB			
21	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	30 ms	1/1	10.2 KiB			
20	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	22 ms	1/1	2.0 KiB			
19	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:51	41 ms	1/1	2.0 KiB			
18	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1	2.0 KiB			
17	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1	10.2 KiB			
16	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1	62.6 KiB			
15	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1				
14	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1				
13	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:48	3 s	1/1				
12	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:46	0.2 s	200/200	32.6 KiB			
10	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:42	0.4 s	200/200	32.6 KiB			
8	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:42	51 ms	1/1			11.3 KiB	
5	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:41	0.7 s	200/200	32.6 KiB			11.3 KiB
3	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:40	1 s	200/200			9.5 KiB	
1	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:39	0.8 s	2/2	6.5 KiB			9.3 KiB
0	id = 4b78e16c-d9a3-496d-98a4-45bbdb113ba8 runId = a12a1d35-d36f-4a96-aa42-86e4... start at ProfileConsolidation.Job.scala:257 +details	2025/12/12 16:12:36	3 s	2/2				

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Spark Jobs (?)

User: rac
Total Uptime: 4.9 min
Scheduling Mode: FIFO
Completed Jobs: 39

▾ Event Timeline
☐ Enable zooming



▾ Completed Jobs (39)

Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go

Job Id (Job Group)	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
38 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:28:07	3 s	1/1	1/1
37 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:27:06	3 s	1/1	1/1
36 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:26:07	3 s	1/1	1/1
35 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:25:37	2 s	1/1	1/1
34 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:34	15 s	1/1 (2 skipped)	1/1 (201 skipped)
33 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:26	28 ms	1/1 (2 skipped)	1/1 (201 skipped)
32 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:23	2 s	1/2	200/201
31 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:23	1 s	1/1 (1 skipped)	200/200 (1 skipped)
30 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:23	22 ms	1/1	1/1
29 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:18	5 ms	1/1 (3 skipped)	1/1 (202 skipped)
28 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:18	33 ms	1/1 (2 skipped)	1/1 (201 skipped)
27 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:18	34 ms	1/1 (2 skipped)	1/1 (201 skipped)
26 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	2 s	1/2	200/201
25 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	1 s	1/1 (1 skipped)	200/200 (1 skipped)
24 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	15 ms	1/1	1/1
23 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	10 ms	1/1 (3 skipped)	1/1 (202 skipped)
22 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	41 ms	1/1 (2 skipped)	1/1 (201 skipped)
21 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:15	60 ms	1/1 (2 skipped)	1/1 (201 skipped)
20 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:13	2 s	1/2	200/201
19 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:12	1 s	1/1 (1 skipped)	200/200 (1 skipped)
18 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:12	0.5 s	1/1	1/1
17 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	0.7 s	1/1 (1 skipped)	200/200 (1 skipped)
16 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	0.5 s	1/1 (1 skipped)	200/200 (1 skipped)
15 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	0.1 s	1/1	1/1
14 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	9 ms	1/1 (1 skipped)	1/1 (1 skipped)
13 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	13 ms	1/1	1/1
12 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	11 ms	1/1 (1 skipped)	1/1 (1 skipped)
11 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	19 ms	1/1	1/1
10 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	12 ms	1/1 (1 skipped)	1/1 (1 skipped)
9 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:11	15 ms	1/1	1/1
8 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:08	3 s	1/1	1/1
7 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:06	16 ms	1/1 (1 skipped)	1/1 (1 skipped)
6 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:06	44 ms	1/1	1/1
5 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:03	3 s	1/1	1/1
4 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:24:00	32 ms	1/1 (1 skipped)	1/1 (1 skipped)
3 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:23:57	3 s	1/1	1/1
2 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:23:57	82 ms	1/1 (1 skipped)	1/1 (1 skipped)
1 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:23:54	3 s	1/1	1/1
0 (450b8ce8-43a3-4712-b445-3d0bdfdbaa06)	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d... start at PipelineBJob.scala:187	2025/12/12 16:23:50	3 s	1/1	1/1

Page: 1

1 Pages. Jump to 1 . Show 100 items in a page. Go

Stages for All Jobs

Completed Stages: 40
Skipped Stages: 27

Completed Stages (40)

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Stage Id	Description	Submitted	Duration	Tasks: Succeeded/Total	Input	Output	Shuffle Read	Shuffle Write
66	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:29:06	3 s	1/1				
65	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:28:07	3 s	1/1				
64	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:27:06	3 s	1/1				
63	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:26:07	3 s	1/1				
62	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:25:37	2 s	1/1				
61	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:34	15 s	1/1		19.0 KiB	111.2 KiB	
58	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:26	26 ms	1/1			97.8 KiB	
55	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:23	1 s	200/200	203.0 KiB			97.8 KiB
54	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:23	1 s	200/200	203.0 KiB			111.2 KiB
52	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:23	17 ms	1/1	13.5 KiB			
51	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:18	2 ms	1/1			59.0 B	
47	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:18	30 ms	1/1			78.7 KiB	59.0 B
44	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:18	31 ms	1/1			78.7 KiB	
41	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	1 s	200/200	203.0 KiB			78.7 KiB
40	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	1 s	200/200	203.0 KiB			78.7 KiB
38	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	11 ms	1/1	13.5 KiB			
37	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	7 ms	1/1			59.0 B	
33	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	38 ms	1/1			78.7 KiB	59.0 B
30	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:15	56 ms	1/1			78.7 KiB	
27	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:13	1 s	200/200	203.0 KiB			78.7 KiB
26	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:12	1 s	200/200	203.0 KiB			78.7 KiB
25	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:12	30 ms	1/1	13.5 KiB			
24	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	0.2 s	200/200	203.0 KiB			
23	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	0.5 s	200/200			58.5 KiB	
21	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	0.1 s	1/1	35.7 KiB			58.5 KiB
20	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	4 ms	1/1			59.0 B	
18	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	9 ms	1/1	13.5 KiB			59.0 B
17	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	6 ms	1/1			59.0 B	
15	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	12 ms	1/1	35.7 KiB			59.0 B
14	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	6 ms	1/1			59.0 B	
12	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:11	9 ms	1/1	13.5 KiB			59.0 B
11	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:08	3 s	1/1				
10	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:06	10 ms	1/1			59.0 B	
8	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:06	33 ms	1/1	35.7 KiB			59.0 B
7	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:03	3 s	1/1				
6	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:24:00	20 ms	1/1			59.0 B	
4	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:23:57	3 s	1/1				59.0 B
3	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:23:57	60 ms	1/1			177.0 B	
1	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:23:54	3 s	1/1				177.0 B
0	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0... start at PipelineBJob.scala:187 +details	2025/12/12 16:23:50	3 s	1/1				

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Skipped Stages (27)

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Spark 3.5.1 Jobs Stages Storage Environment Executors SQL / DataFrame Structured Streaming **Pipeline B - Incremental Daily T...** application UI

Storage

RDDs

ID	RDD Name	Storage Level	Cached Partitions	Fraction Cached	Size in Memory	Size on Disk
22	*(1) Project [txn_id#55L, customer_id#56, product_id#57, qty#58, amount#59, txn_timestamp#60, cast(txn_timestamp#60 as date) AS date#67] +- *(1) Scan JDBCRelation((SELECT txn_id, customer_id, product_id, qty, amount, txn_timestamp FROM transactions WHERE DATE(txn_timestamp) IN ('2025-12-11','2025-12-12','2025-12-10')) AS tx_full) [numPartitions=1] [txn_id#55L,customer_id#56,product_id#57,qty#58,amount#59,txn_timestamp#60] PushedFilters: [], ReadSchema: struct<txn_id:bigint,customer_id:int,product_id:int,qty:int,amount:decimal(10,2),txn_timestamp:ti...	Disk Memory Deserialized 1x Replicated	1	100.00%	35.7 KiB	0.0 B
34	*(1) Scan JDBCRelation(products) [numPartitions=1] [product_id#297,name#298,category#299,price#300] PushedFilters: [], ReadSchema: struct<product_id:int,name:string,category:string,price:decimal(10,2)>	Disk Memory Deserialized 1x Replicated	1	100.00%	13.5 KiB	0.0 B
66	AdaptiveSparkPlan isFinalPlan=false +- Exchange hashpartitioning(customer_id#56, 200), REPARTITION_BY_COL, [plan_id=203] +- Project [txn_id#55L, customer_id#56, product_id#57, qty#58, amount#59, txn_timestamp#60, cast(txn_timestamp#60 as date) AS date#434] +- Filter isnotnull(product_id#57) +- InMemoryTableScan [amount#59, customer_id#56, product_id#57, qty#58, txn_id#55L, txn_timestamp#60], [isnotnull(product_id#57)] +- InMemoryRelation [txn_id#55L, customer_id#56, product_id#57, qty#58, amount#59, txn_timestamp#60, date#67], StorageLevel(disk, memory, deserialized, 1 replicas) +- *(1) Project [txn_id#55L, customer_id#56, product_id#57, qty#58, amount#59, txn_timestamp#60, cast(txn_timestamp#60 as date) AS date#67] +- *(1) Scan JDBCRelation((SELECT txn_id, customer_id, product_id, qty, amount, txn_timestamp FROM transactions WHERE DATE(txn_timestamp) IN ('2025-12-11','2025-12-12','2025-12-10')) AS tx_full) [numPartitions=1] [txn...	Disk Memory Deserialized 1x Replicated	200	100.00%	203.0 KiB	0.0 B

Spark 3.5.1 Jobs Stages Storage Environment Executors SQL / DataFrame Structured Streaming **Pipeline B - Incremental Daily T...** application UI

SQL / DataFrame

Completed Queries: 7

Completed Queries (7)

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

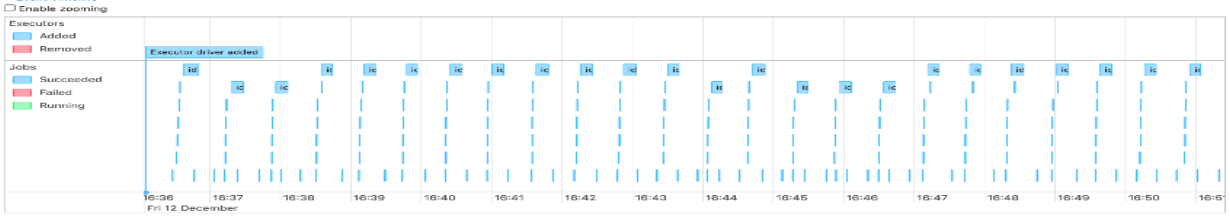
ID	Description	Submitted	Duration	Job IDs	Sub Execution IDs
21	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 6 +details	2025/12/12 16:30:00	10 s		[22] +details
19	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 5 +details	2025/12/12 16:29:00	9 s		[20] +details
17	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 4 +details	2025/12/12 16:28:00	11 s		[18] +details
15	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 3 +details	2025/12/12 16:27:00	9 s		[16] +details
13	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 2 +details	2025/12/12 16:26:00	10 s		[14] +details
11	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 1 +details	2025/12/12 16:25:31	9 s		[12] +details
0	id = 6c4dc180-b837-41fc-9890-d0e75ad52e55 runId = 450b8ce8-43a3-4712-b445-3d0bdfdbaa06 batch = 0 +details	2025/12/12 16:23:44	1.8 min		[1][2][3][4][5][6][7][8][9][10] +details

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Spark Jobs (?)

User: rac
Total Uptime: 15 min
Scheduling Mode: FIFO
Completed Jobs: 195

▾ Event Timeline
☐ Enable zooming



▾ Completed Jobs (195)

Page: 1 2 3 4 >

4 Pages. Jump to 1 . Show 50 items in a page. Go

Job Id (Job Group)	Description	Submitted	Duration	Stages: Succeeded/Total	Tasks (for all stages): Succeeded/Total
194 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:51:19	0.5 s	2/2	4/4
193 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:51:00	0.3 s	2/2	4/4
192 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:94	10 s	1/1	3/3
191 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:50	0.5 s	2/2	203/203
190 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:90	0.4 s	2/2	203/203
189 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:50	0.3 s	2/2	203/203
188 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:50	7 ms	1/1	1/1
187 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:49	0.5 s	2/2	4/4
186 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:50:42	0.4 s	2/2	4/4
185 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:50:25	0.6 s	2/2	4/4
184 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:15	11 s	1/1	3/3
183 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:12	0.2 s	2/2	203/203
182 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:12	0.3 s	2/2	203/203
181 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:12	0.2 s	2/2	203/203
180 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:12	7 ms	1/1	1/1
179 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:50:11	0.5 s	2/2	4/4
178 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:50:09	0.5 s	2/2	4/4
177 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:49:51	0.5 s	2/2	4/4
176 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:37	11 s	1/1	3/3
175 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:34	0.2 s	2/2	203/203
174 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:49:34	0.5 s	2/2	4/4
173 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:34	0.4 s	2/2	203/203
172 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:34	0.2 s	2/2	203/203
171 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:34	9 ms	1/1	1/1
170 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:33	0.4 s	2/2	4/4
169 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:49:19	0.5 s	2/2	4/4
168 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:49:01	0.5 s	2/2	4/4
167 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:49:01	11 s	1/1	3/3
166 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:57	0.4 s	2/2	203/203
165 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:57	0.4 s	2/2	203/203
164 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:56	0.3 s	2/2	203/203
163 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:56	10 ms	1/1	1/1
162 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:56	0.3 s	2/2	4/4
161 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:48:44	0.6 s	2/2	4/4
160 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:48:25	0.5 s	2/2	4/4
159 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:22	11 s	1/1	3/3
158 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:18	0.3 s	2/2	203/203
157 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:18	0.4 s	2/2	203/203
156 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:18	0.3 s	2/2	203/203
155 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:18	7 ms	1/1	1/1
154 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:48:17	0.5 s	2/2	4/4
153 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:48:07	0.3 s	2/2	4/4
152 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:47:49	0.4 s	2/2	4/4
151 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:46	10 s	1/1	3/3
150 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:43	0.2 s	2/2	203/203
149 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:43	0.4 s	2/2	203/203
148 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:42	0.3 s	2/2	203/203
147 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:42	7 ms	1/1	1/1
146 (e074dee9-ee82-484a-9f35-bc4782f3c2ac)	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4... start at PipelineApp.scala:127	2025/12/12 16:47:42	0.3 s	2/2	4/4
145 (f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0)	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-587... start at PipelineApp.scala:109	2025/12/12 16:47:31	0.6 s	2/2	4/4

Page: 1 2 3 4 >

4 Pages. Jump to 1 . Show 50 items in a page. Go

Spark 3.5.1 Jobs Stages Storage Environment Executors SQL / DataFrame Structured Streaming Pipeline C - Kafka Avro Events A... application UI

Storage

RDDs

ID	RDD Name	Storage Level	Cached Partitions	Fraction Cached	Size in Memory	Size on Disk
1853	*(1) Scan ExistingRDD[event_id#28,customer_id#29,event_type#30,product_id#40,event_timestamp#57,event_date#72,ingestion_timestamp#80]	Disk Memory Deserialized 1x Replicated	3	100.00%	25.1 KiB	0.0 B

Spark 3.5.1 Jobs Stages Storage Environment Executors SQL / DataFrame Structured Streaming Pipeline C - Kafka Avro Events A... application UI

SQL / DataFrame

Completed Queries: 84

Completed Queries (84)

Page: 1 2 > 2 Pages. Jump to 1 . Show 50 items in a page. Go

ID	Description	Submitted	Duration	Job IDs	Sub Execution IDs
301	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 93 +details	2025/12/12 16:53:08	0.5 s		[302] +details
299	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 92 +details	2025/12/12 16:52:51	0.6 s		[300] +details
292	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 51 +details	2025/12/12 16:52:41	20 s		[293][294][295][296][297][298] +details
290	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 91 +details	2025/12/12 16:52:32	0.6 s		[291] +details
288	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 90 +details	2025/12/12 16:52:12	0.4 s		[289] +details
281	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 50 +details	2025/12/12 16:52:04	21 s		[282][283][284][285][286][287] +details
279	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 89 +details	2025/12/12 16:51:54	0.3 s		[280] +details
277	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 88 +details	2025/12/12 16:51:35	0.6 s		[278] +details
270	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 49 +details	2025/12/12 16:51:26	20 s		[271][272][273][274][275][276] +details
268	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 87 +details	2025/12/12 16:51:19	0.5 s		[269] +details
266	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 86 +details	2025/12/12 16:51:00	0.3 s		[267] +details
259	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 48 +details	2025/12/12 16:50:49	19 s		[260][261][262][263][264][265] +details
257	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 85 +details	2025/12/12 16:50:42	0.5 s		[258] +details
255	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 84 +details	2025/12/12 16:50:25	0.6 s		[256] +details
248	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 47 +details	2025/12/12 16:50:11	19 s		[249][250][251][252][253][254] +details
246	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 83 +details	2025/12/12 16:50:09	0.5 s		[247] +details
244	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 82 +details	2025/12/12 16:49:51	0.6 s		[245] +details
240	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 81 +details	2025/12/12 16:49:34	0.6 s		[241] +details
235	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 46 +details	2025/12/12 16:49:33	20 s		[236][237][238][239][242][243] +details
233	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 80 +details	2025/12/12 16:49:18	0.6 s		[234] +details
231	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 79 +details	2025/12/12 16:49:01	0.6 s		[232] +details
224	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 45 +details	2025/12/12 16:48:56	20 s		[225][226][227][228][229][230] +details
222	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 78 +details	2025/12/12 16:48:43	1 s		[223] +details
220	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 77 +details	2025/12/12 16:48:25	0.5 s		[221] +details
213	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 44 +details	2025/12/12 16:48:17	21 s		[214][216][216][217][218][219] +details
211	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 76 +details	2025/12/12 16:48:07	0.3 s		[212] +details
209	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 75 +details	2025/12/12 16:47:49	0.4 s		[210] +details
202	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 43 +details	2025/12/12 16:47:42	19 s		[203][204][205][206][207][208] +details
200	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 74 +details	2025/12/12 16:47:31	0.6 s		[201] +details
198	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 73 +details	2025/12/12 16:47:13	0.4 s		[199] +details
191	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 42 +details	2025/12/12 16:47:06	19 s		[192][193][194][195][196][197] +details
189	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 72 +details	2025/12/12 16:46:55	0.5 s		[190] +details
187	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 71 +details	2025/12/12 16:46:38	0.3 s		[188] +details
180	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 41 +details	2025/12/12 16:46:28	20 s		[181][182][183][184][185][186] +details
178	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 70 +details	2025/12/12 16:46:19	0.3 s		[179] +details
176	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 69 +details	2025/12/12 16:46:02	0.5 s		[177] +details
169	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 40 +details	2025/12/12 16:45:51	20 s		[170][171][172][173][174][175] +details
167	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 68 +details	2025/12/12 16:45:44	0.3 s		[168] +details
165	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 67 +details	2025/12/12 16:45:25	0.6 s		[166] +details
158	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 39 +details	2025/12/12 16:45:15	19 s		[159][160][161][162][163][164] +details
156	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 66 +details	2025/12/12 16:45:06	0.6 s		[157] +details
154	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 65 +details	2025/12/12 16:44:49	0.3 s		[155] +details
147	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 38 +details	2025/12/12 16:44:37	20 s		[148][149][150][151][152][153] +details
145	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 64 +details	2025/12/12 16:44:30	0.6 s		[146] +details
143	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 63 +details	2025/12/12 16:44:13	0.6 s		[144] +details
136	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 37 +details	2025/12/12 16:44:03	18 s		[137][138][139][140][141][142] +details
134	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 62 +details	2025/12/12 16:43:54	0.6 s		[135] +details
132	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 61 +details	2025/12/12 16:43:37	0.3 s		[133] +details
125	id = 83aa934f-4d58-4e5c-a3d0-651fc72b3e58 runId = e074dee9-ee82-484a-9f35-bc4782f3c2ac batch = 36 +details	2025/12/12 16:43:25	19 s		[126][127][128][129][130][131] +details
123	id = d9e02216-a60c-4a3a-91ec-44ea71232773 runId = f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0 batch = 60 +details	2025/12/12 16:43:20	0.4 s		[124] +details

Page: 1 2 > 2 Pages. Jump to 1 . Show 50 items in a page. Go

Spark 3.5.1 | Jobs | Stages | Storage | Environment | Executors | SQL / DataFrame | Structured Streaming | Pipeline C - Kafka Avro Events A... application UI

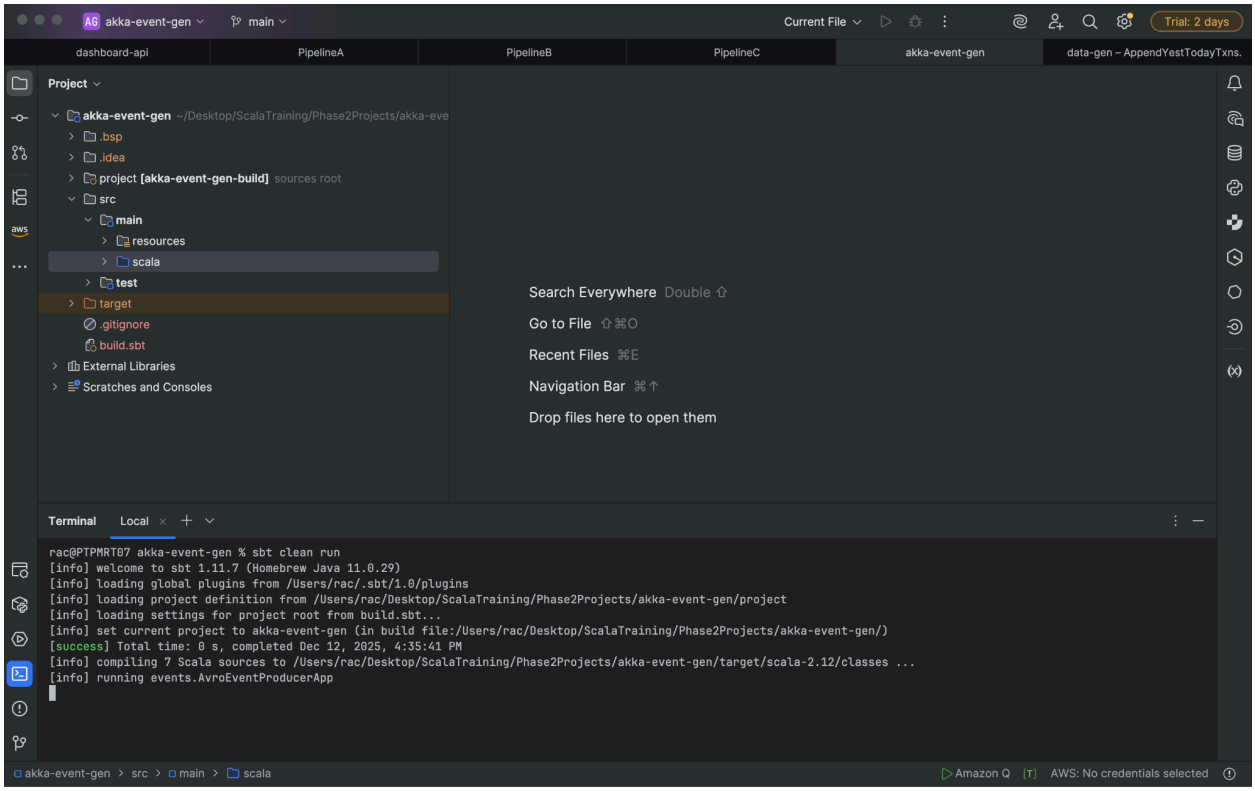
Streaming Query

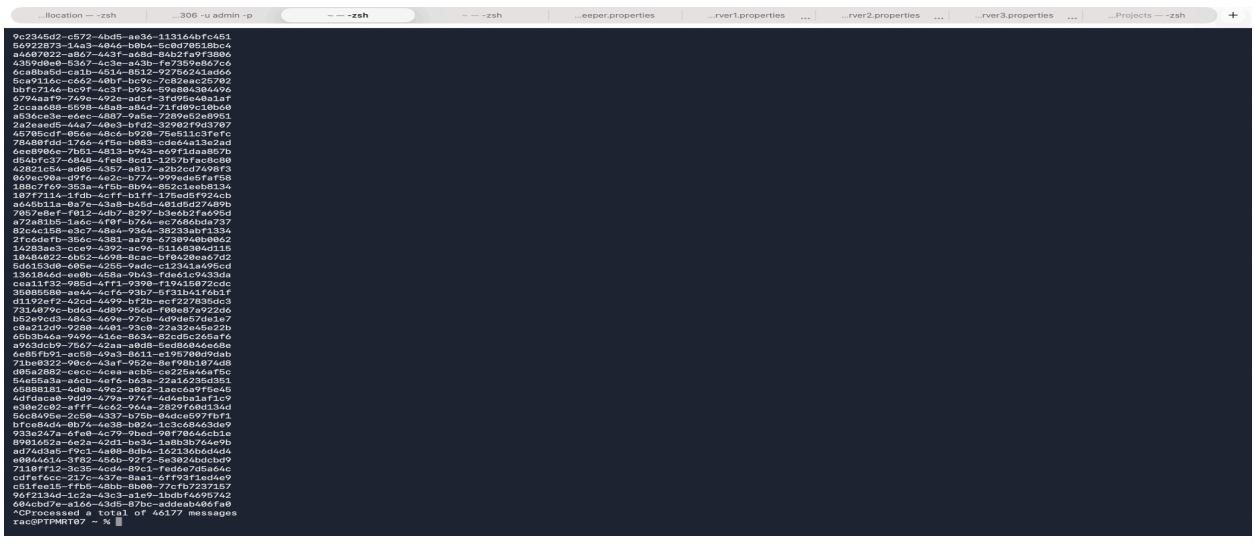
▾ Active Streaming Queries (2)

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go

Name	Status	ID	Run ID	Start Time ▾	Duration	Avg Input /sec	Avg Process /sec	Latest Batch
<no name>	RUNNING	83aa934f-4d58-4e5c-a3d0-651fc72b3e58	e074dee9-ee82-484a-9f35-bc4782f3c2ac	2025/12/12 16:36:18	18 minutes 1 second	9.62	9.91	52
<no name>	RUNNING	d9e02216-a60c-4a3a-91ec-44ea71232773	f5a4ecf2-d4c7-45e8-8c2e-5873c54c9cc0	2025/12/12 16:36:13	18 minutes 6 seconds	9.81	10.06	96

Page: 1 — 1 Pages. Jump to 1 . Show 100 items in a page. Go





Explanation of partitioning logic

Overview

All three pipelines use partitioning to balance parallelism, reduce network shuffle, and make downstream storage/querying efficient. Partitioning happens at three places:

(1) how JDBC reads from MySQL are partitioned,

(2) how Spark repartitions data for groupBy/joins and downstream writes, and

(3) how final data is partitioned in storage (Cassandra partition key or S3 hive-style date partitions).

Pipeline A — MySQL → Spark → Cassandra (Customer profiles)

Partitioning steps & heuristics

1. JDBC partitioning (extractor.readDeltaTransactions / readTransactionsForDates)
    - Uses `option("partitionColumn", "txn_id")` together with Spark `lowerBound`/`upperBound` and `numPartitions`.
    - `choosePartitions(minId, maxId, cfg)` computes numPartitions with the heuristic `~10K rows / partition`:

```
val approxPerPartition = 10000L
val p = Math.max(2, Math.min(cfg, (span / approxPerPartition + 1).toInt))
```

    - Purpose: parallelize JDBC reads across multiple connections/threads while avoiding too many tiny partitions.
2. Spark repartition for aggregation
    - Pre-aggregation: transactions are `.repartition($"customer_id")` before joins/aggregations. That colocates rows for each customer on the same partition and reduces shuffle during `groupBy(customer_id)`.
    - After aggregation, before writing to Cassandra, code computes `repartitions`:

```
val repartitions = Math.max(8, Math.min(writeParallelism, Math.max(1, affectedCountHint * 4)))
```

    - This caps/limits the number of output partitions to balance parallel Cassandra writes vs too many small SSTable writes.
3. Cassandra partitioning (storage layout)
    - Cassandra table `customer_profile` uses `customer_id` as the partition key (single row per customer). This enables very fast point lookups (API reads).
    - Write strategy: repartition by `customer_id` attempts to align Spark partitions to Cassandra partitioning so each Spark task drives writes for a subset of customers.

Pipeline B — MySQL → Spark Structured Streaming (ForeachBatch) → S3 (Daily summaries)

Partitioning steps & heuristics

1. JDBC partitioning for delta
    - Same `txn_id` partitioning heuristic via `choosePartitions` when reading delta ranges. This parallelizes `readDeltaTransactions`.
2. Derive affected dates
    - The pipeline reads only `affectedDates` (distinct `to_date(txn_timestamp)` from delta) and then reads the full transactions for those dates using `readTransactionsForDates(dates)`. This minimizes the scope of re-computation.
3. Spark repartition & aggregation
    - Preprocess transactions with `.repartition($"customer_id")` in the transformer to colocate per-customer rows for the daily summary aggregations.
4. S3 partitioning
    - Final writes to S3 use `partitionBy("date")` and `SaveMode.Overwrite` for idempotent writes of the affected date partition:

```
.partitionBy("date")
.parquet(lakeBase)
```

    - The `S3Loader.writeSummary` coalesces output files (`coalesce(coalesceNum)`) prior to write to control number/size of files per partition.

Pipeline C — Kafka Avro → Spark Structured Streaming → S3 (Events)

Partitioning steps & heuristics

1. Kafka partitioning (Producer behavior)
    - Your `AvroEventProducerActor` sets the Kafka message key to `eventId`:

```
val record = new ProducerRecord[String, Array[Byte]](topic, eventId,
bytes)
```

    - This means events are partitioned by `event_id` hashing. If you need ordering per customer, you'd want to key by `customer_id` instead.
2. Spark streaming partitioning & processing

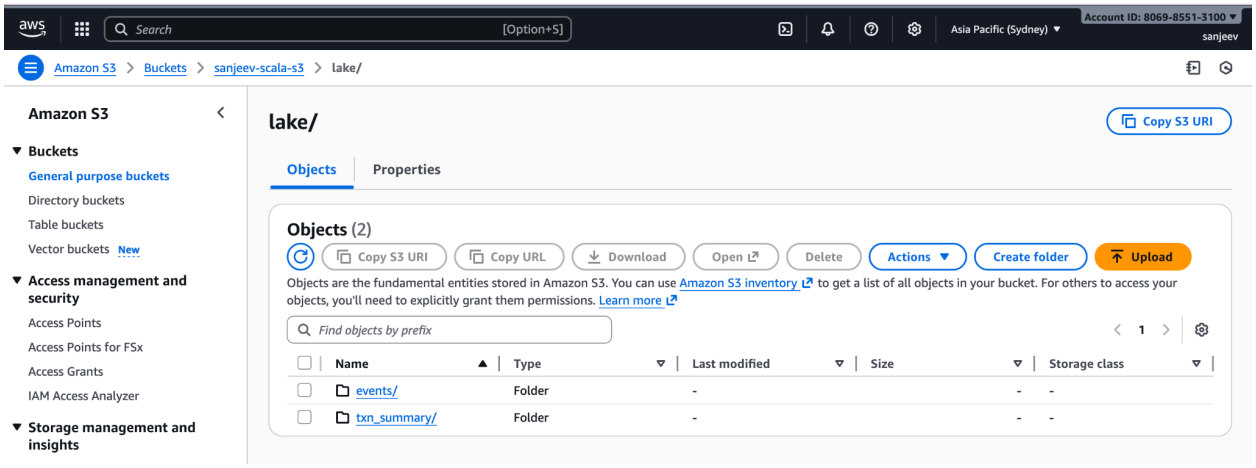
- Incoming Kafka stream is handled by Spark Structured Streaming. There is no explicit repartitioning in the extractor code beyond usual streaming micro-batch partitioning, but:
  - `from_avro` expands records.
  - `event_date = to_date(event_timestamp)` is added and used as partition column for final writes.

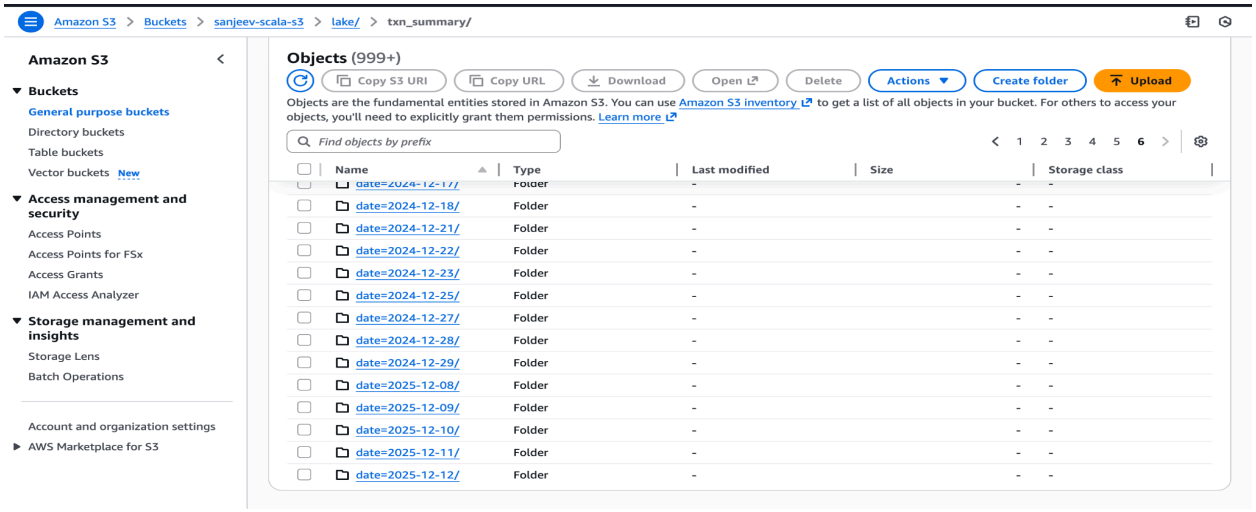
3. S3 partitioning
   - Final writes use `partitionBy("event_date")` and append mode for streaming:

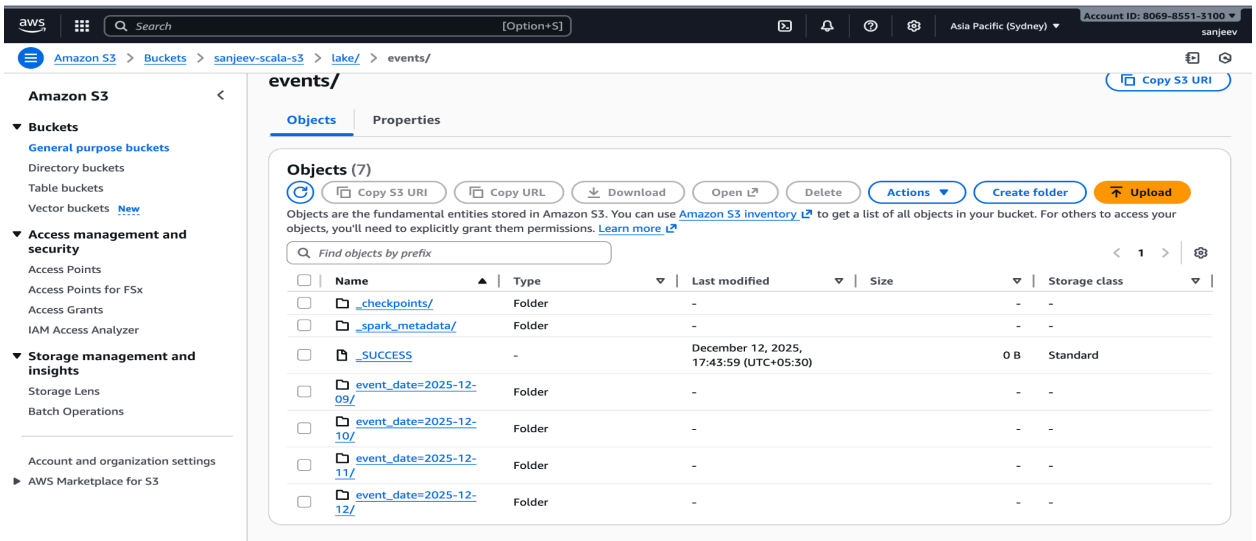
```
df.write.mode("append").partitionBy("event_date").parquet(basePath)
```

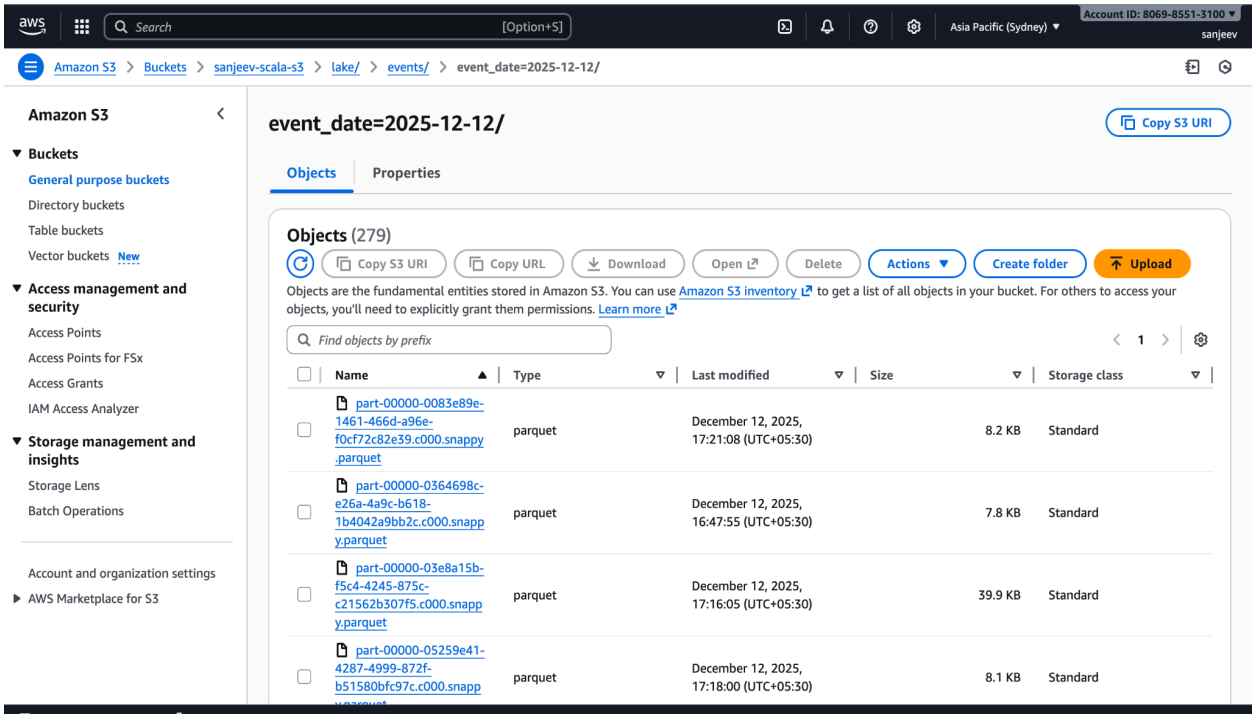
- Malformed records are handled by a separate `malformed` stream and can be written to a dedicated malformed path.

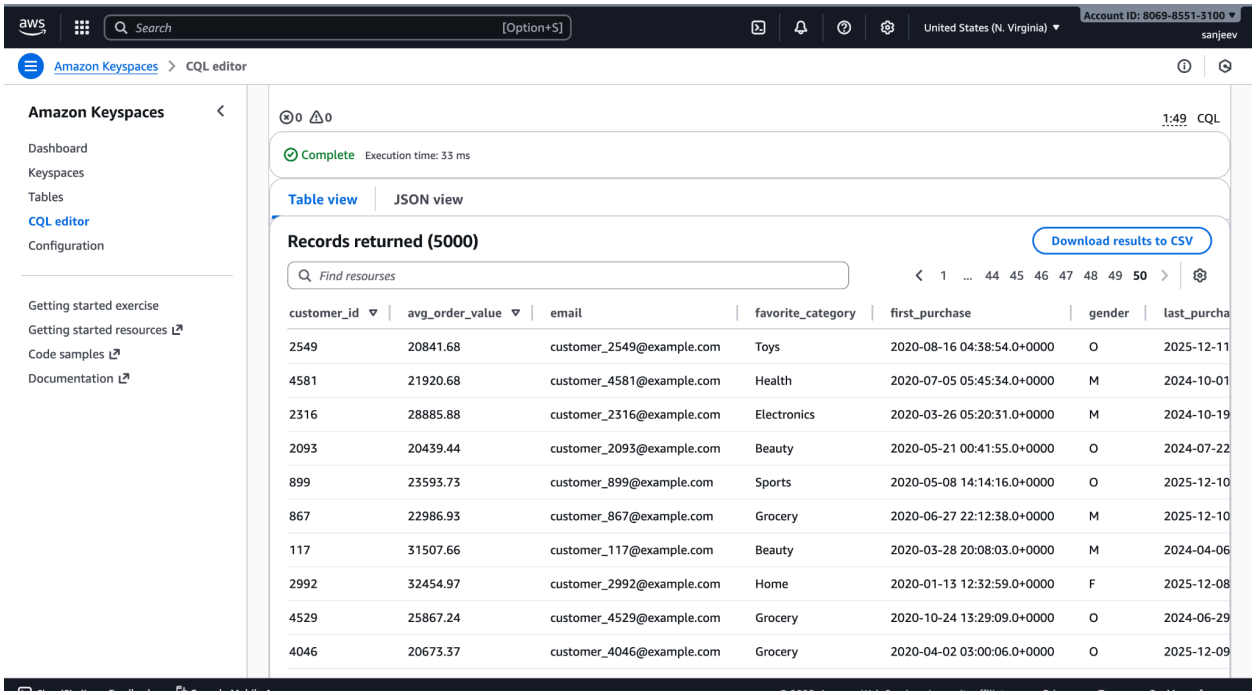
Screenshots:

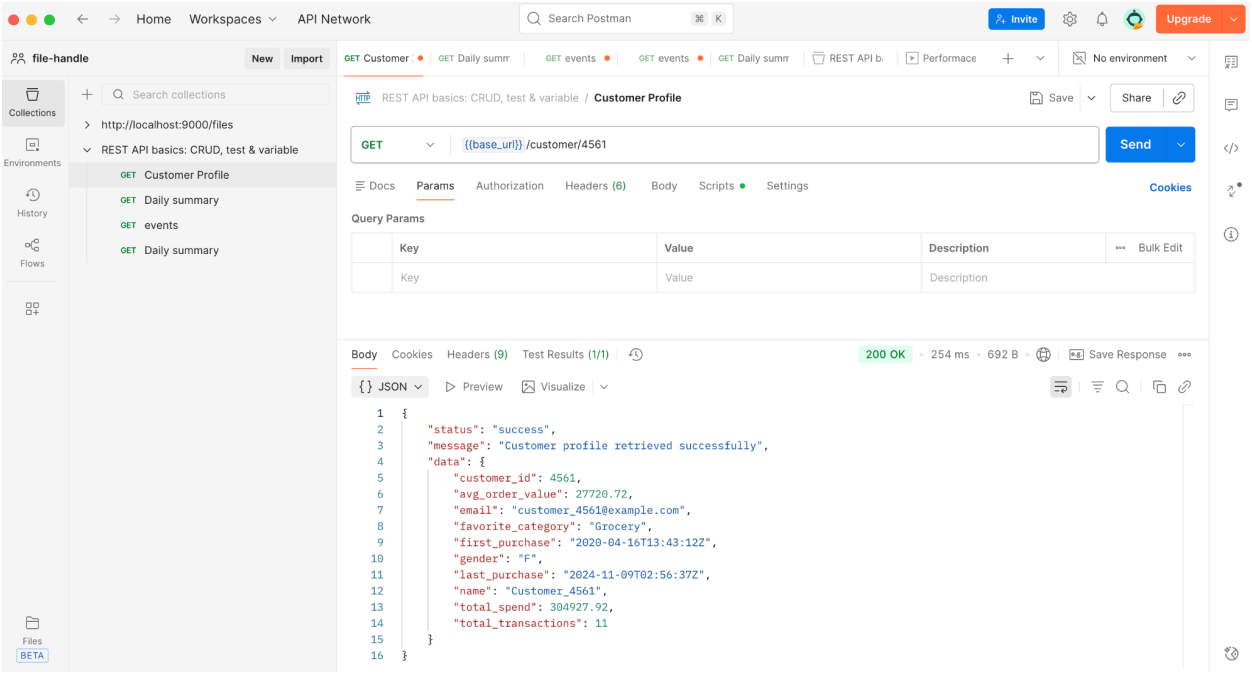


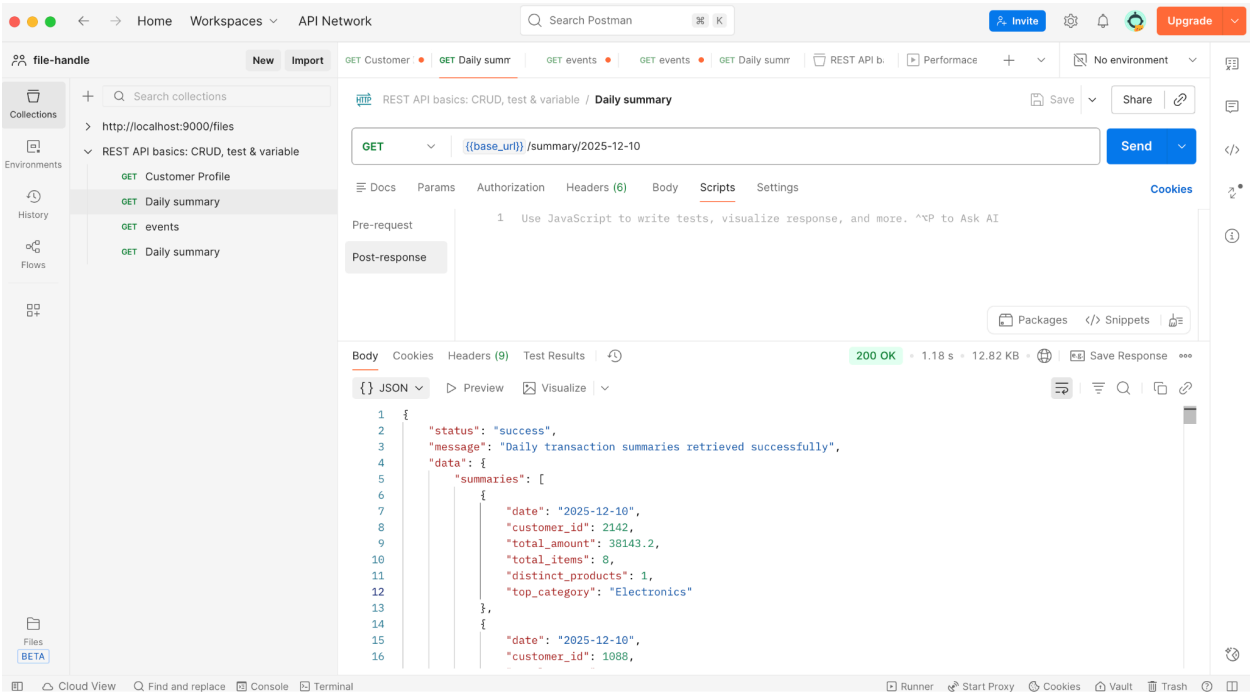


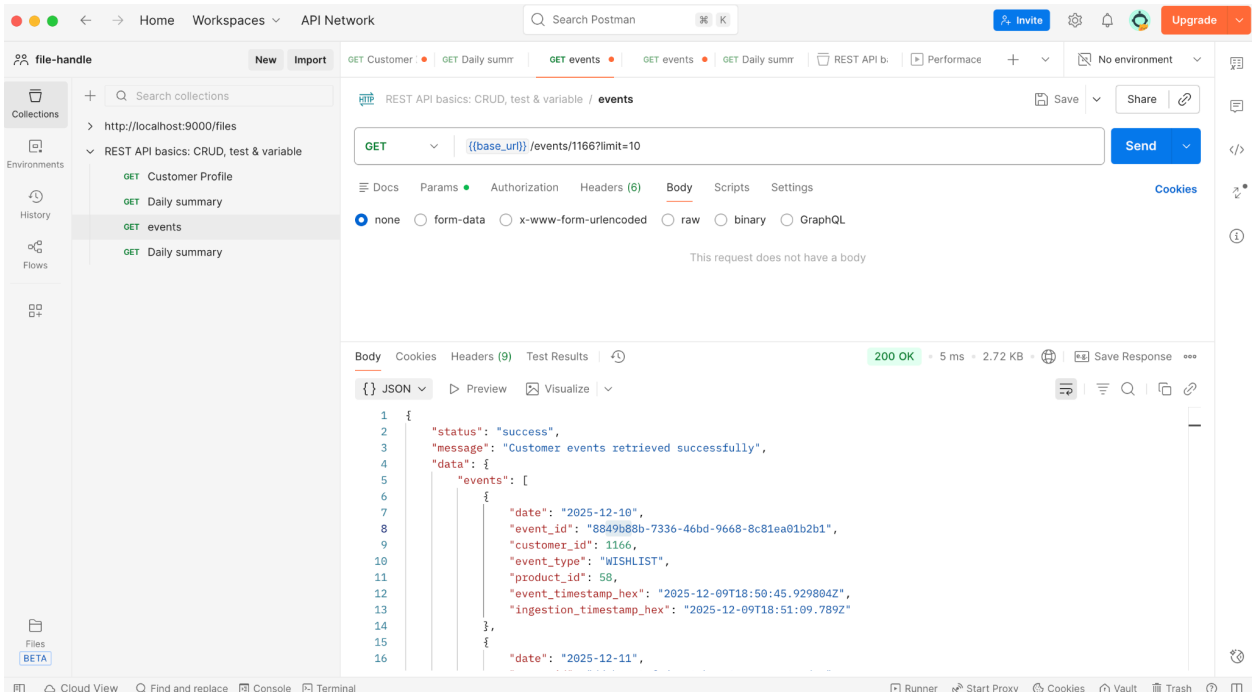


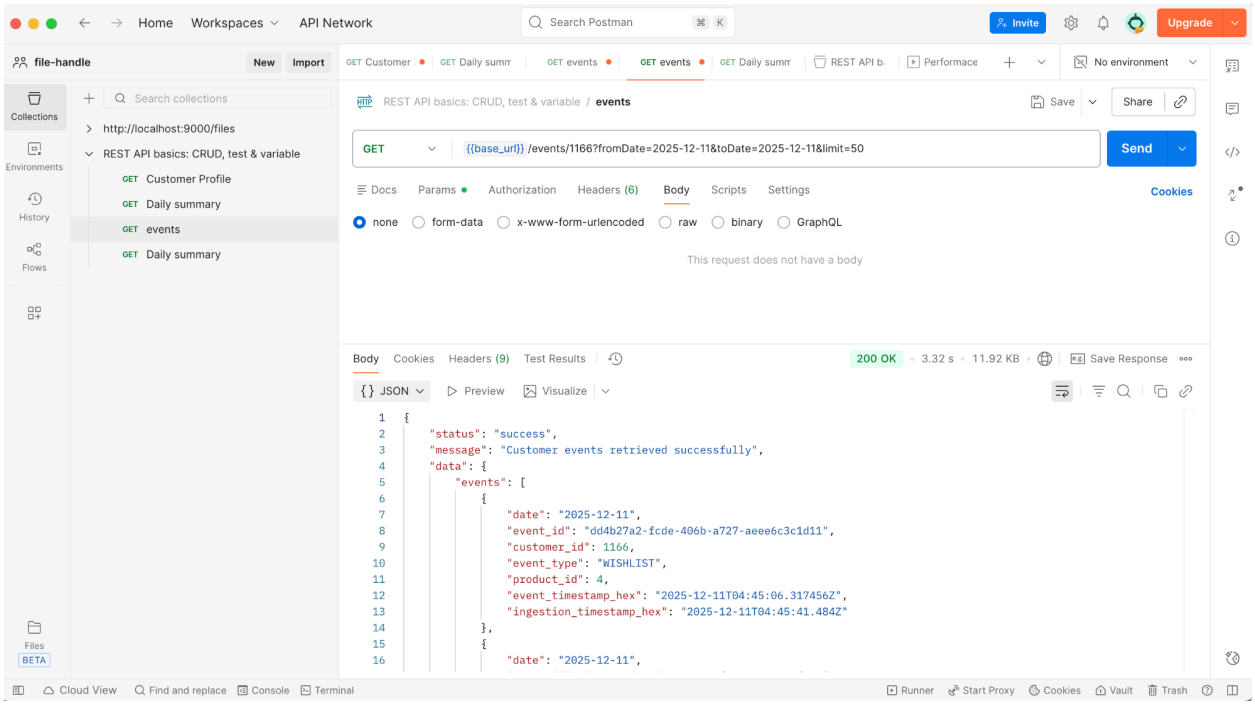

Search [Option+S]
Asia Pacific (Sydney)
Account ID: 8069-8551-3100
sanjeev
Amazon S3 > Buckets > sanjeev-scala-s3 > lake/ > events/ > event_date=2025-12-12/
Amazon S3
Buckets
General purpose buckets
Directory buckets
Table buckets
Vector buckets New
Access management and security
Access Points
Access Points for FSx
Access Grants
IAM Access Analyzer
Storage management and insights
Storage Lens
Batch Operations
Account and organization settings
AWS Marketplace for S3
event_date=2025-12-12/
Copy S3 URI
Objects
Properties
Objects (279)
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
Name
Type
Last modified
Size
Storage class
part-00000-0083e89e-1461-466d-a96e-f0cf72c82e39.c000.snappy.parquet
parquet
December 12, 2025, 17:21:08 (UTC+05:30)
8.2 KB
Standard
part-00000-0364698c-e26a-4a9c-b618-1b4042a9bb2c.c000.snappy.parquet
parquet
December 12, 2025, 16:47:55 (UTC+05:30)
7.8 KB
Standard
part-00000-03e8a15b-f5c4-4245-875c-c21562b307f5.c000.snappy.parquet
parquet
December 12, 2025, 17:16:05 (UTC+05:30)
39.9 KB
Standard
part-00000-05259e41-4287-4999-872f-b51580bfc97c.c000.snappy.parquet
parquet
December 12, 2025, 17:18:00 (UTC+05:30)
8.1 KB
Standard


Search [Option+S]
United States (N. Virginia)
Account ID: 8069-8551-3100
sanjeev
Amazon Keyspaces > CQL editor
Amazon Keyspaces
Dashboard
Keyspaces
Tables
CQL editor
Configuration
Getting started exercise
Getting started resources
Code samples
Documentation
0 0
1:49 CQL
Complete Execution time: 33 ms
Table view
JSON view
Records returned (5000)
Download results to CSV
Find resources
1 ... 44 45 46 47 48 49 50
customer_id | avg_order_value | email | favorite_category | first_purchase | gender | last_purcha
2549 | 20841.68 | customer_2549@example.com | Toys | 2020-08-16 04:38:54.0+0000 | O | 2025-12-11
4581 | 21920.68 | customer_4581@example.com | Health | 2020-07-05 05:45:34.0+0000 | M | 2024-10-01
2316 | 28885.88 | customer_2316@example.com | Electronics | 2020-03-26 05:20:31.0+0000 | M | 2024-10-19
2093 | 20439.44 | customer_2093@example.com | Beauty | 2020-05-21 00:41:55.0+0000 | O | 2024-07-22
899 | 23593.73 | customer_899@example.com | Sports | 2020-05-08 14:14:16.0+0000 | O | 2025-12-10
867 | 22986.93 | customer_867@example.com | Grocery | 2020-06-27 22:12:38.0+0000 | M | 2025-12-10
117 | 31507.66 | customer_117@example.com | Beauty | 2020-03-28 20:08:03.0+0000 | M | 2024-04-06
2992 | 32454.97 | customer_2992@example.com | Home | 2020-01-13 12:32:59.0+0000 | F | 2025-12-08
4529 | 25867.24 | customer_4529@example.com | Grocery | 2020-10-24 13:29:09.0+0000 | O | 2024-06-29
4046 | 20673.37 | customer_4046@example.com | Grocery | 2020-04-02 03:00:06.0+0000 | O | 2025-12-09


Home
Workspaces
API Network
Search Postman
Invite
Upgrade
file-handle
New
Import
Collections
Environments
History
Flows
Files
BETA
Search collections
http://localhost:9000/files
REST API basics: CRUD, test & variable
GET Customer Profile
GET Daily summary
GET events
GET Daily summary
No environment
REST API basics: CRUD, test & variable / Customer Profile
Save
Share
GET
{{base_url}} /customer/4561
Send
Docs
Params
Authorization
Headers (6)
Body
Scripts
Settings
Cookies
Query Params
Key
Value
Description
Bulk Edit
Body
Cookies
Headers (9)
Test Results (1/1)
200 OK
254 ms
692 B
Save Response
JSON
Preview
Visualize
{
"status": "success",
"message": "Customer profile retrieved successfully",
"data": {
"customer_id": 4561,
"avg_order_value": 27720.72,
"email": "customer_4561@example.com",
"favorite_category": "Grocery",
"first_purchase": "2020-04-16T13:43:12Z",
"gender": "F",
"last_purchase": "2024-11-09T02:56:37Z",
"name": "Customer_4561",
"total_spend": 304927.92,
"total_transactions": 11
}
}


Home
Workspaces
API Network
Search Postman
Invite
Upgrade
file-handle
New
Import
Collections
Environments
History
Flows
Files
BETA
Search collections
http://localhost:9000/files
REST API basics: CRUD, test & variable
GET Customer Profile
GET Daily summary
GET events
GET Daily summary
No environment
REST API basics: CRUD, test & variable / Daily summary
Save
Share
GET
{{base_url}} /summary/2025-12-10
Send
Docs
Params
Authorization
Headers (6)
Body
Scripts
Settings
Cookies
Pre-request
Post-response
Use JavaScript to write tests, visualize response, and more. ^⌥P to Ask AI
Packages
Snippets
Body
Cookies
Headers (9)
Test Results
200 OK
1.18 s
12.82 KB
Save Response
JSON
Preview
Visualize
{
"status": "success",
"message": "Daily transaction summaries retrieved successfully",
"data": {
"summaries": [
{
"date": "2025-12-10",
"customer_id": 2142,
"total_amount": 38143.2,
"total_items": 8,
"distinct_products": 1,
"top_category": "Electronics"
},
{
"date": "2025-12-10",
"customer_id": 1088,
Cloud View
Find and replace
Console
Terminal
Runner
Start Proxy
Cookies
Vault
Trash


Home
Workspaces
API Network
Search Postman
Invite
Upgrade
file-handle
New
Import
Collections
Environments
History
Flows
Files
BETA
Search collections
http://localhost:9000/files
REST API basics: CRUD, test & variable
GET Customer Profile
GET Daily summary
GET events
GET Daily summary
GET events
GET Daily summ
REST API b
Performace
No environment
REST API basics: CRUD, test & variable / events
Save
Share
GET
{{base_url}} /events/1166?limit=10
Send
Docs
Params
Authorization
Headers (6)
Body
Scripts
Settings
Cookies
none
form-data
x-www-form-urlencoded
raw
binary
GraphQL
This request does not have a body
Body
Cookies
Headers (9)
Test Results
200 OK
5 ms
2.72 KB
Save Response
JSON
Preview
Visualize
{
"status": "success",
"message": "Customer events retrieved successfully",
"data": {
"events": [
{
"date": "2025-12-10",
"event_id": "8849b88b-7336-46bd-9668-8c81ea01b2b1",
"customer_id": 1166,
"event_type": "WISHLIST",
"product_id": 58,
"event_timestamp_hex": "2025-12-09T18:50:45.929804Z",
"ingestion_timestamp_hex": "2025-12-09T18:51:09.789Z"
},
{
"date": "2025-12-11",
Cloud View
Find and replace
Console
Terminal
Runner
Start Proxy
Cookies
Vault
Trash


Home
Workspaces
API Network
Search Postman
Invite
Upgrade
file-handle
New
Import
Collections
Environments
History
Flows
Files
BETA
Search collections
http://localhost:9000/files
REST API basics: CRUD, test & variable
GET Customer Profile
GET Daily summary
GET events
GET Daily summary
GET events
GET Daily summ
REST API b
Performace
No environment
REST API basics: CRUD, test & variable / events
Save
Share
GET
{{base_url}} /events/1166?fromDate=2025-12-11&toDate=2025-12-11&limit=50
Send
Docs
Params
Authorization
Headers (6)
Body
Scripts
Settings
Cookies
none
form-data
x-www-form-urlencoded
raw
binary
GraphQL
This request does not have a body
Body
Cookies
Headers (9)
Test Results
200 OK
3.32 s
11.92 KB
Save Response
JSON
Preview
Visualize
{
"status": "success",
"message": "Customer events retrieved successfully",
"data": {
"events": [
{
"date": "2025-12-11",
"event_id": "dd4b27a2-fcde-406b-a727-aeee6c3c1d11",
"customer_id": 1166,
"event_type": "WISHLIST",
"product_id": 4,
"event_timestamp_hex": "2025-12-11T04:45:06.317456Z",
"ingestion_timestamp_hex": "2025-12-11T04:45:41.484Z"
},
{
"date": "2025-12-11",
Cloud View
Find and replace
Console
Terminal
Runner
Start Proxy
Cookies
Vault
Trash

Performace Report of api calls - Dec 11, 2025 (#1)

Open in Postman

Postman collection: REST API basics: CRUD, test & variable

Report exported on: Dec 11, 2025, 12:29:19 (GMT+5:30)

Test setup

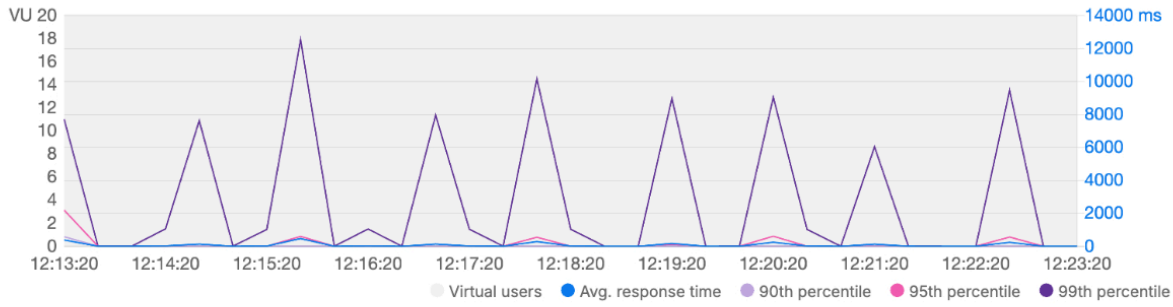
Virtual users	Start time	Load profile
20 VU	Dec 11, 12:13:20 (GMT+5:30)	Fixed
Duration	**End time**	**Environment**
10 minutes	Dec 11, 12:23:27 (GMT+5:30)	-

1. Summary

Total requests sent	Throughput	Average response time	Error rate
39,968	65.93 requests/second	50 ms	0.00 %

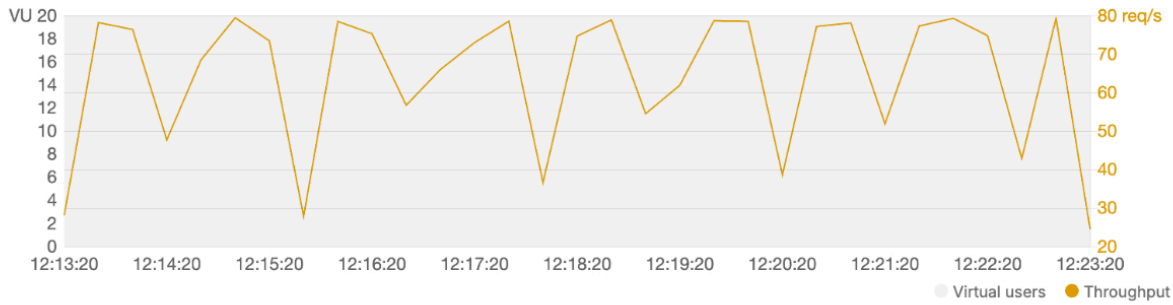
1.1 Response time

Response time trends during the test duration.



1.2 Throughput

Rate of requests sent per second during the test duration.



1.3 Requests with slowest response times

Top 5 slowest requests based on their average response times.

Request	Resp. time (Avg ms)	90th (ms)	95th (ms)	99th (ms)	Min (ms)	Max (ms)
GET events {{base_url}}/events/1166?limit=100	161	3	4	8,441	1	12,943
GET Daily summary {{base_url}}/summary/2025-12-10	23	4	5	1,045	1	3,040
GET Daily summary {{base_url}}/summary/2025-12-10/3156	13	2	3	565	1	1,063
GET Customer Profile {{base_url}}/customer/456	3	3	4	15	1	386

2. Metrics for each request

The requests are shown in the order they were sent by virtual users.

Request	Total requests	Requests/s	Min (ms)	Avg (ms)	90th (ms)	Max (ms)	Error %
GET Customer Profile {{base_url}}/customer/456	9,992	16.48	1	3	3	386	0
GET Daily summary {{base_url}}/summary/2025-12-10	9,992	16.48	1	23	4	3,040	0
GET events {{base_url}}/events/1166?limit=100	9,992	16.48	1	161	3	12,943	0
GET Daily summary {{base_url}}/summary/2025-12-10/3156	9,992	16.48	1	13	2	1,063	0

Sample JSON responses

1. Customer Profile API

GET /customer/:id

Sample Success Response


```
{
  "status": "success",
  "message": "Customer profile retrieved successfully",
  "data": {
    "customer_id": 456,
    "name": "Customer_456",
    "email": "customer_456@example.com",
    "gender": "M",
    "total_spend": 143949.5,
    "total_transactions": 6,
    "avg_order_value": 23991.58,
    "first_purchase": "2022-03-07T06:30:00Z",
    "last_purchase": "2024-05-18T18:17:40Z",
    "favorite_category": "Grocery"
  }
}
```


Sample Not-Found Response


```
{
  "status": "error",
  "message": "not-found",
  "data": {}
```

```
}
```

---

2. Daily Summary API

GET /summary/:date/:customerId

Sample Success Response


```
{
  "status": "success",
  "message": "Daily transaction summary retrieved successfully",
  "data": {
    "summary": {
      "date": "2025-12-10",
      "customer_id": 3156,
      "total_amount": 12895.28,
      "total_items": 2,
      "distinct_products": 1,
      "top_category": "Electronics"
    }
  }
}
```


Sample Not-Found Response


```
{
  "status": "error",
  "message": "not-found",
  "data": {}
}
```

Sample Invalid Date Response


```
{
  "status": "error",
  "message": "invalid_date_format",
  "data": {}
}
```


---

3. Events API

GET /events/:customerId?fromDate=&toDate=&limit=

Sample Success Response


```
{
  "status": "success",
  "message": "Customer events retrieved successfully",
  "data": {
    "events": [
      {
        "date": "2025-12-09",
        “event_id”: “dd4b27a2-fcde-406b-a727-aeee6c3c1d11”
        "customer_id": 2942,
        "event_type": "CART_ADD",
        "product_id": 491,
        "event_timestamp": "2025-12-09T18:18:50.850395Z",
        "ingestion_timestamp": "2025-12-09T18:18:59.149Z"
      }
    ]
  }
```

```
}
```

Sample Invalid fromDate Response

```
{
  "status": "error",
  "message": "invalid_fromDate_format",
  "data": {}
}
```

Sample Invalid toDate Response

```
{
  "status": "error",
  "message": "invalid_toDate_format",
  "data": {}
}
```

Sample Invalid Limit Response

```
{
  "status": "error",
  "message": "invalid_limit",
  "data": {}
}
```

Sample Customer Not Found Response

```
{
  "status": "error",
  "message": "not-found",
  "data": {}
}
```

---

4. Daily Summaries API

GET /summary/:date

Sample Success Response


```
{
  "status": "success",
  "message": "Daily transaction summaries retrieved successfully",
  "data": {
    "summaries": [
      {
        "date": "2025-12-10",
        "customer_id": 1088,
        "total_amount": 57714.9,
        "total_items": 6,
        "distinct_products": 1,
        "top_category": "Sports"
      },
      {
        "date": "2025-12-10",
        "customer_id": 3156,
        "total_amount": 12895.28,
        "total_items": 2,
        "distinct_products": 1,
        "top_category": "Electronics"
      }
    ]
  }
}
```

Sample Not-Found Response


```
{
  "status": "error",
  "message": "not-found",
  "data": {}
}
```


Sample Invalid Date Response


```
{
  "status": "error",
  "message": "invalid_date_format",
  "data": {}
}
```