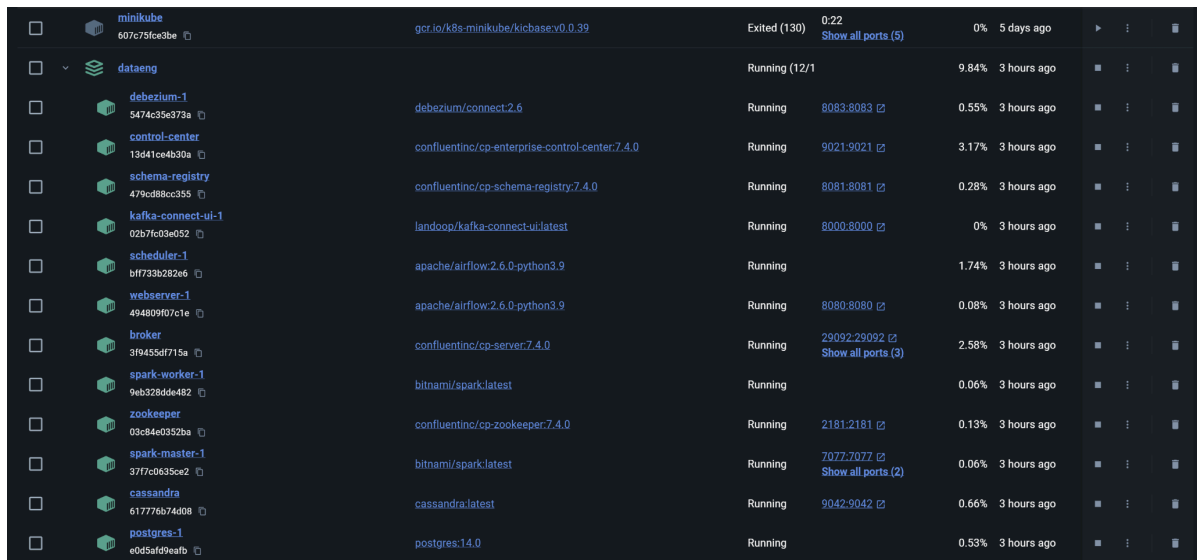


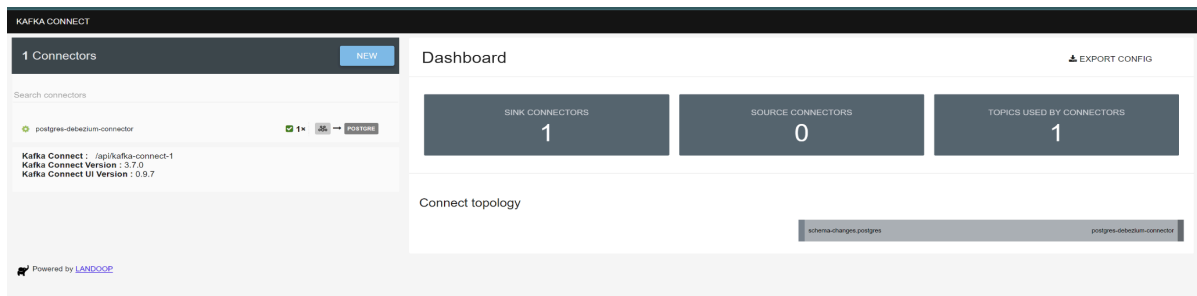
Change Data Capture POC using Debezium

- Performed Debezium POC - captured Create, Update, and Delete changes performed on a PostgreSQL Table through Debezium and displayed them in the Kafka Control Center.
- For the POC, I used:
 - Debezium - To capture the changes
 - PostgreSQL - The database where the changes were made
 - Kafka Connector UI - To view and manage the Debezium Connector
 - Zookeeper, Schema Registry, Broker - For Kafka
 - Control Center - For Kafka UI for viewing the captured changes



minikube	607c73fce3be	gcr.io/k8s-minikube/kicbase:v0.0.39	Exited (130)	0/22 Show all ports (5)	0%	5 days ago	
dataeng			Running (12/1)		9.84%	3 hours ago	
debezium-1	5474c39e379a	debezium/connect:2.6	Running	8083:8083	0.55%	3 hours ago	
control-center	13d41ce4b30a	confluentinc/cp-enterprise-control-center:7.4.0	Running	9021:9021	3.17%	3 hours ago	
schema-registry	479cd88cc355	confluentinc/cp-schema-registry:7.4.0	Running	8081:8081	0.28%	3 hours ago	
kafka-connect-ui-1	02b7fc03e052	landoop/kafka-connect-ui:latest	Running	8000:8000	0%	3 hours ago	
scheduler-1	b1f733b282e6	apache/airflow:2.6.0-python3.9	Running		1.74%	3 hours ago	
webserver-1	494809f07c1e	apache/airflow:2.6.0-python3.9	Running	8080:8080	0.08%	3 hours ago	
broker	3f9455df715a	confluentinc/cp-server:7.4.0	Running	29092:29092 Show all ports (3)	2.58%	3 hours ago	
spark-worker-1	9eb328dde482	bitnami/spark:latest	Running		0.06%	3 hours ago	
zookeeper	03c94e0352ba	confluentinc/cp-zookeeper:7.4.0	Running	2181:2181	0.13%	3 hours ago	
spark-master-1	3777c0635ce2	bitnami/spark:latest	Running	7077:7077 Show all ports (2)	0.06%	3 hours ago	
cassandra	61777b674d08	cassandra:latest	Running	9042:9042	0.66%	3 hours ago	
postgres-1	e0d5af09eafb	postgres:14.0	Running		0.53%	3 hours ago	

- Created a Connector



KAFKA CONNECT

1 Connectors [NEW](#)

Search connectors

postgres-debezium-connector 1x 3x PORTS

Kafka Connect : /api/kafka-connect-1
Kafka Connect Version : 3.7.0
Kafka Connect UI Version : 0.9.7

Powered by [LANDOOPE](#)

Dashboard [EXPORT CONFIG](#)

SINK CONNECTORS 1

SOURCE CONNECTORS 0

TOPICS USED BY CONNECTORS 1

Connect topology


schema-changes.postgres → postgres-debezium-connector

- Created a Kafka Topic to display the messages of the captured Database changes


postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 0 

Consumers

Bytes out/sec 0 

+ Produce a new message to this topic

Message fields

No new messages

The message browser shows messages that have arrived since this page was opened.

- The database is initially empty (Primary Key is 'id'):

```
dataeng-postgres-1
postgres:14.0
e0d5afd9eafb

Logs Inspect Bind mounts Exec Files Stats

airflow=# SELECT * FROM created_users;
 id | first_name | last_name | email | username
----+-----+-----+-----+-----
(0 rows)

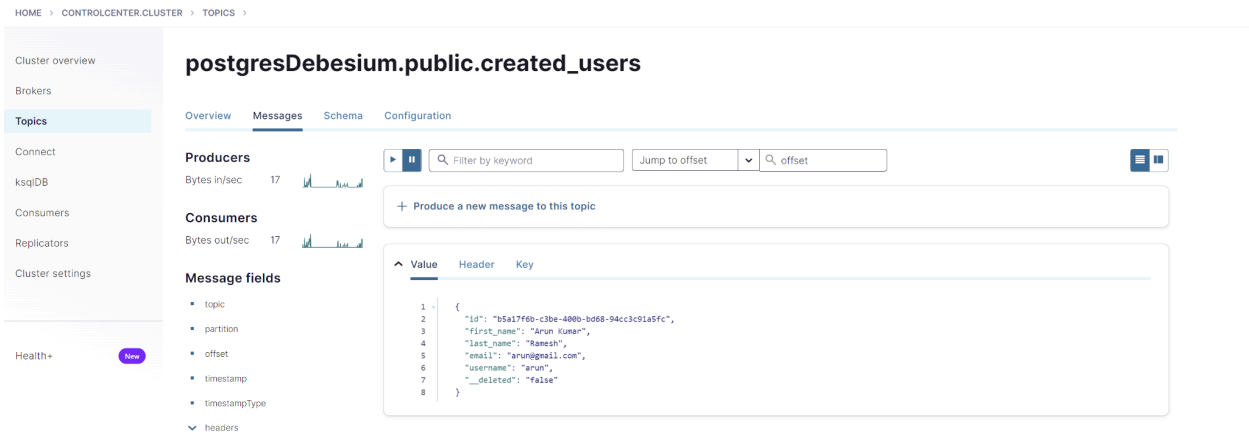
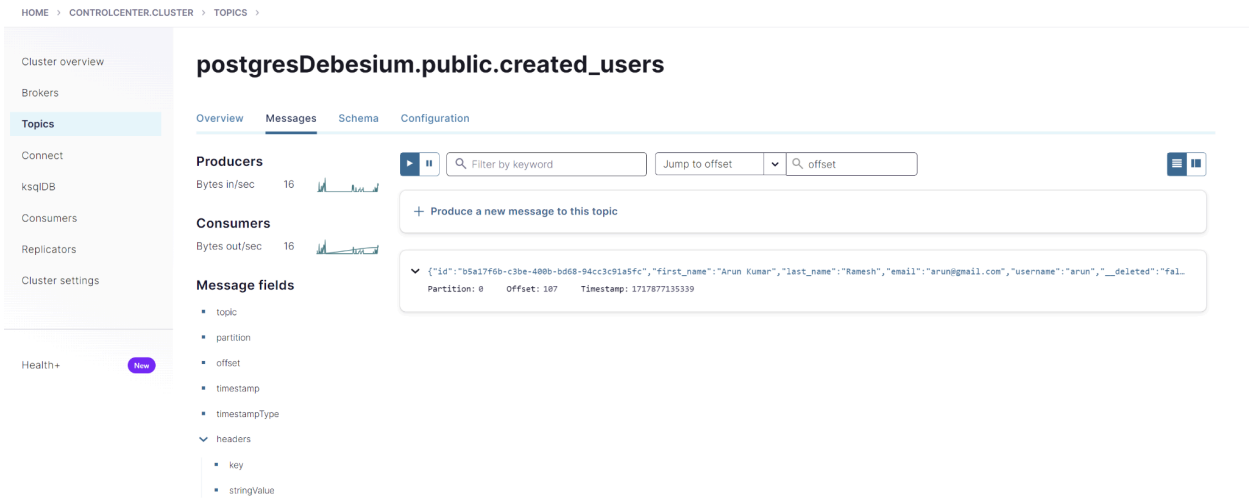
airflow=#
```

- Performing Create Operation:

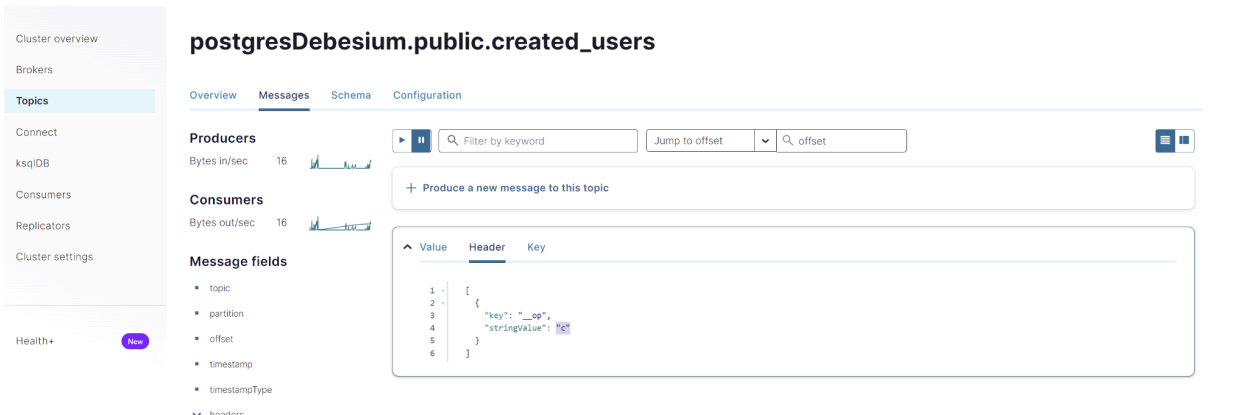
```
airflow=# INSERT INTO created_users (id, first_name, last_name, email, username)
VALUES (
    gen_random_uuid(),
    'Arun Kumar',
    'Ramesh',
    'arun@gmail.com',
    'arun'
);
INSERT 0 1
airflow=# SELECT * FROM created_users;
          id          | first_name | last_name | email      | username
-----+-----+-----+-----+-----
b5a17f6b-c3be-400b-bd68-94cc3c91a5fc | Arun Kumar | Ramesh    | arun@gmail.com | arun
(1 row)

airflow=#
```

Captured Insertion message -



‘c’ stands for ‘Create’ (As we did an ‘Insert’ operation)



HOME > CONTROLCENTER.CLUSTER > TOPICS >

Cluster overview
Brokers
Topics
Connect
ksqldb
Consumers
Replicators
Cluster settings

Health+ New

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers
Bytes in/sec 17

Consumers
Bytes out/sec 17

Message fields

- topic
- partition
- offset
- timestamp
- timestampType
- headers

Filter by keyword
Jump to offset
offset

+ Produce a new message to this topic

Value	Header	Key
{ "id": "b5a17f6b-c3be-400b-bd68-94cc3c91a5fc" }		

Notice the “id” of the message in Kafka and the “id” of the inserted record match.

- Inserting two more records and displaying the captured messages -

```
airflow=# INSERT INTO created_users (id, first_name, last_name, email, username)
VALUES (
    gen_random_uuid(),
    'Alice',
    'Brown',
    'Alice@gmail.com',
    'alicebrown'
);
INSERT 0 1
airflow=# INSERT INTO created_users (id, first_name, last_name, email, username)
VALUES (
    gen_random_uuid(),
    'Jane',
    'Smith',
    'Jane@gmail.com',
    'janesmith'
);
INSERT 0 1
airflow=# SELECT * FROM created_users;
          id          | first_name | last_name | email          | username
-----+-----+-----+-----+-----
 b5a17f6b-c3be-400b-bd68-94cc3c91a5fc | Arun Kumar | Ramesh    | arun@gmail.com | arun
 4aa08d39-eb73-4017-b0fd-88627bd7f67d | Alice      | Brown     | Alice@gmail.com | alicebrown
 c1844796-d4d4-418d-b641-5bcb319a4584 | Jane       | Smith     | Jane@gmail.com  | janesmith
(3 rows)

airflow=#
```

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+

New

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 0

Consumers

Bytes out/sec 0

Message fields

- topic
- partition
- offset
- timestamp
- timestampType
- headers
 - key
 - stringValue
- key



Filter by keyword

Jump to offset

offset



+ Produce a new message to this topic

{ "id": "c1844796-d464-418d-b641-5bcb319a4584", "first_name": "Jane", "last_name": "Smith", "email": "Jane@gmail.com", "username": "janesmith", "__deleted": "false", "Partition": 0, "Offset": 109, "Timestamp": 1717877273640 }

{ "id": "4aa88d39-eb73-4817-b0fd-88627bd7f67d", "first_name": "Alice", "last_name": "Brown", "email": "Alice@gmail.com", "username": "alicebrown", "__deleted": "false", "Partition": 0, "Offset": 108, "Timestamp": 1717877268113 }

{ "id": "b5a17f6b-c3be-480b-bd68-94cc3c91a5fc", "first_name": "Arun Kumar", "last_name": "Ramesh", "email": "arun@gmail.com", "username": "arun", "__deleted": "false", "Partition": 0, "Offset": 107, "Timestamp": 1717877135339 }

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+

New

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 9

Consumers

Bytes out/sec 9

Message fields

- topic
- partition
- offset
- timestamp
- timestampType
- headers
 - key
 - stringValue
- key
 - id
- value
 - id
 - first_name
 - last_name
 - email



Filter by keyword

Jump to offset

offset



+ Produce a new message to this topic

Value Header Key

```
1 {
2   "id": "c1844796-d464-418d-b641-5bcb319a4584",
3   "first_name": "Jane",
4   "last_name": "Smith",
5   "email": "Jane@gmail.com",
6   "username": "janesmith",
7   "__deleted": "false"
8 }
```

Value Header Key

```
1 {
2   "id": "4aa88d39-eb73-4817-b0fd-88627bd7f67d",
3   "first_name": "Alice",
4   "last_name": "Brown",
5   "email": "Alice@gmail.com",
6   "username": "alicebrown",
7   "__deleted": "false"
8 }
```

{ "id": "b5a17f6b-c3be-480b-bd68-94cc3c91a5fc", "first_name": "Arun Kumar", "last_name": "Ramesh", "email": "arun@gmail.com", "username": "arun", "__deleted": "fa..." }

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+ New

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 9

Consumers

Bytes out/sec 9

Message fields

topic

partition

offset

timestamp

timestampType

headers

key

stringValue

key

id

value

id

first_name

Filter by keyword

Jump to offset

offset

Produce a new message to this topic

Value

Header

Key

1

2

3

4

5

6

{

"key": "__op",

"stringValue": "c"

}

Value

Header

Key

1

2

3

4

5

6

{

"key": "__op",

"stringValue": "c"

}

Value

Header

Key

● Performing Update Operation:

```
airflow=# UPDATE created_users
SET email = 'jane.smith@gmail.com'
WHERE username = 'janesmith';
UPDATE 1
airflow=# SELECT * FROM created_users;
```

id	first_name	last_name	email	username
b5a17f6b-c3be-400b-bd68-94cc3c91a5fc	Arun Kumar	Ramesh	arun@gmail.com	arun
4aa08d39-eb73-4017-b0fd-88627bd7f67d	Alice	Brown	Alice@gmail.com	alicebrown
c1844796-d4d4-418d-b641-5bcb319a4584	Jane	Smith	jane.smith@gmail.com	janesmith

(3 rows)

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+ New

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 0

Consumers

Bytes out/sec 0

Message fields

topic

partition

offset

timestamp

timestampType

headers

key

stringValue

key

id

value

id

first_name

last_name

email

Filter by keyword

Jump to offset

offset

Produce a new message to this topic

Value

Header

Key

1

2

3

4

5

6

7

8

{

"id": "c1844796-d4d4-418d-b641-5bcb319a4584",

"first_name": "Jane",

"last_name": "Smith",

"email": "Jane.smith@gmail.com",

"username": "janesmith",

"__deleted": "false"

}

Value

Header

Key

{

"id": "c1844796-d4d4-418d-b641-5bcb319a4584",

"first_name": "Jane",

"last_name": "Smith",

"email": "Jane.smith@gmail.com",

"username": "janesmith",

"__deleted": "false"

}

Value

Header

Key

{

"id": "4aa08d39-eb73-4017-b0fd-88627bd7f67d",

"first_name": "Alice",

"last_name": "Brown",

"email": "Alice@gmail.com",

"username": "alicebrown",

"__deleted": "false"

}

Value

Header

Key

{

"id": "b5a17f6b-c3be-400b-bd68-94cc3c91a5fc",

"first_name": "Arun Kumar",

"last_name": "Ramesh",

"email": "arun@gmail.com",

"username": "arun",

"__deleted": "false"

}

Value

Header

Key

The stringValue is “u” (stands for “Update”)

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers
Bytes in/sec 4

Consumers
Bytes out/sec 4

Message fields

- topic
- partition
- offset
- timestamp
- timestampType

headers

- key
- stringValue

key

- id

value

- id
- first_name
- last_name
- email
- username

Value Header Key

```
1 - {
2 -   {
3 -     "key": "__op",
4 -     "stringValue": "u"
5 -   }
6 - }
```

✓ {"id":"c1844796-d4d4-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"Jane@gmail.com","username":"janesmith","__deleted":"false.."}
Partition: 0 Offset: 109 Timestamp: 1717877273648

✓ {"id":"4aa08d39-eb73-4017-b0fd-88627bd7f67d","first_name":"Alice","last_name":"Brown","email":"Alice@gmail.com","username":"alicebrown","__deleted":"false.."}
Partition: 0 Offset: 108 Timestamp: 1717877268113

✓ {"id":"b5a17f6b-c3be-400b-bd68-94cc3c91a5fc","first_name":"Arun Kumar","last_name":"Ramesh","email":"arun@gmail.com","username":"arun","__deleted":"false.."}
Partition: 0 Offset: 107 Timestamp: 1717877135339

Performing Delete Operation:

```
airflow=# DELETE FROM created_users WHERE username = 'janesmith';
DELETE 1
airflow=# SELECT * FROM created_users;
```

id	first_name	last_name	email	username
b5a17f6b-c3be-400b-bd68-94cc3c91a5fc	Arun Kumar	Ramesh	arun@gmail.com	arun
4aa08d39-eb73-4017-b0fd-88627bd7f67d	Alice	Brown	Alice@gmail.com	alicebrown

(2 rows)

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers
Bytes in/sec 0

Consumers
Bytes out/sec 0

Message fields

- topic
- partition
- offset
- timestamp
- timestampType

headers

- key
- stringValue

key

- id

value

- id
- first_name
- last_name
- email
- username

Value Header Key

```
1 - {
2 -   {
3 -     "id": "c1844796-d4d4-418d-b641-5bcb319a4584",
4 -     "first_name": null,
5 -     "last_name": null,
6 -     "email": null,
7 -     "username": null,
8 -     "__deleted": "true"
9 -   }
10 - }
```

✓ {"id":"c1844796-d4d4-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"Jane.smith@gmail.com","username":"janesmith","__deleted":"false.."}
Partition: 0 Offset: 110 Timestamp: 1717877382858

✓ {"id":"c1844796-d4d4-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"Jane@gmail.com","username":"janesmith","__deleted":"false.."}
Partition: 0 Offset: 109 Timestamp: 1717877273648

✓ {"id":"4aa08d39-eb73-4017-b0fd-88627bd7f67d","first_name":"Alice","last_name":"Brown","email":"Alice@gmail.com","username":"alicebrown","__deleted":"false.."}
Partition: 0 Offset: 108 Timestamp: 1717877268113

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+

postgresDebesium.public.created_users

OverviewMessagesSchemaConfiguration

Producers

Bytes in/sec4

Consumers

Bytes out/sec4

Message fields

topic

partition

offset

timestamp

timestampType

headers

key

stringValue

key

id

value

id

first_name

last_name

email

Filter by keyword

Jump to offset

offset

Produce a new message to this topic

ValueHeaderKey

1- {

2- {

3- "key": "__op",

4- "stringValue": "id"

5- }

6- }

▼ {"id":"c1844796-d4dd-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"jane.smith@gmail.com","username":"janesmith","__deleted":false,"Partition":0,"Offset":110,"Timestamp":1717877382859}

▼ {"id":"c1844796-d4dd-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"jane@gmail.com","username":"janesmith","__deleted":false,"Partition":0,"Offset":109,"Timestamp":1717877273640}

▼ {"id":"4aa98d39-eb73-4017-b0fd-88627bd7f67d","first_name":"Alice","last_name":"Brown","email":"Alice@gmail.com","username":"alicebrown","__deleted":false,"Partition":0,"Offset":108,"Timestamp":1717877260113}

- Deleting all the records (We should see two new messages as two records are deleted) -

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+ New

postgresDebesium.public.created_users

Overview

Messages

Schema

Configuration

Producers

Bytes in/sec 0

Consumers

Bytes out/sec 0

Message fields

topic

partition

offset

timestamp

timestampType

headers

key

stringValue

key

id

value

id

first_name

last_name

email

Filter by keyword

Jump to offset

offset

Produce a new message to this topic

Value

Header

Key

1

-

{

2

"id": "4aa08d39-eb73-4017-b0f6-8862bd7f67d",

3

"first_name": null,

4

"last_name": null,

5

"email": null,

6

"username": null,

7

"__deleted": "true"

8

}

Value

Header

Key

1

-

{

2

"id": "b5a17f6b-c3be-400b-bd68-94cc3c91a5fc",

3

"first_name": null,

4

"last_name": null,

5

"email": null,

6

"username": null,

7

"__deleted": "true"

8

}

Value

Header

Key

1

-

{

2

"id": "c1844796-d4d4-418d-b641-5bcb319a4584",

3

"first_name": null,

4

"last_name": null,

5

"email": null,

6

"username": null,

7

"__deleted": "true"

8

}

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+ New

postgresDebesium.public.created_users

Overview

Messages

Schema

Configuration

Producers

Bytes in/sec 7

Consumers

Bytes out/sec 7

Message fields

topic

partition

offset

timestamp

timestampType

headers

key

stringValue

key

id

value

id

first_name

last_name

email

username

__deleted

Filter by keyword

Jump to offset

offset

Produce a new message to this topic

Value

Header

Key

1

-

[

2

-

{

3

"key": "__op",

4

"stringValue": "00"

5

}

6

]

Value

Header

Key

1

-

[

2

-

{

3

"key": "__op",

4

"stringValue": "00"

5

}

6

]

Value

Header

Key

1

-

[

2

-

{

3

"key": "__op",

4

"stringValue": "00"

5

}

6

]

- Finally, displaying all the captured change messages (3 Creates + 1 Update + 3 Deletes)

Cluster overview

Brokers

Topics

Connect

ksqlDB

Consumers

Replicators

Cluster settings

Health+ View

postgresDebesium.public.created_users

Overview Messages Schema Configuration

Producers

Bytes in/sec 0

Consumers

Bytes out/sec 0

Message fields

- topic
- partition
- offset
- timestamp
- timestampType
- headers
 - key
 - stringValue
- key
 - id
- value
 - id
 - first_name
 - last_name
 - email
 - username
 - __deleted

Filter by keyword

Jump to offset

offset

+ Produce a new message to this topic

["id":"4aa08d39-a673-4017-b0fd-88627bd7f67d","first_name":null,"last_name":null,"email":null,"username":null,"__deleted":"true"]

Partition: 0 Offset: 113 Timestamp: 1717877627346

["id":"b5a17f6b-c3be-400b-bd68-94cc391a5fc","first_name":null,"last_name":null,"email":null,"username":null,"__deleted":"true"]

Partition: 0 Offset: 112 Timestamp: 1717877627346

["id":"c1844796-d6d4-418d-b641-5bcb319a4584","first_name":null,"last_name":null,"email":null,"username":null,"__deleted":"true"]

Partition: 0 Offset: 111 Timestamp: 1717877627346

["id":"c1844796-d6d4-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"jane.smith@gmail.com","username":"janesmith","__deleted":false]

Partition: 0 Offset: 110 Timestamp: 17178777382858

["id":"c1844796-d6d4-418d-b641-5bcb319a4584","first_name":"Jane","last_name":"Smith","email":"jane@gmail.com","username":"janesmith","__deleted":false]

Partition: 0 Offset: 109 Timestamp: 17178777273648

["id":"4aa08d39-a673-4017-b0fd-88627bd7f67d","first_name":"Alice","last_name":"Brown","email":"Alice@gmail.com","username":"alicebrown","__deleted":false]

Partition: 0 Offset: 108 Timestamp: 1717877208113

["id":"b5a17f6b-c3be-400b-bd68-94cc391a5fc","first_name":"Arun Kumar","last_name":"Ramesh","email":"arun@gmail.com","username":"arun","__deleted":false]

Partition: 0 Offset: 107 Timestamp: 1717877135339