

Stream pipeline monitoring

Switching from ElasticSearch to ClickHouse



Ricky Thomas
Co-founder / CTO

Agenda

What is Streamkap

Our Data/Requirements

Why ElasticSearch didn't work for us

Solving with ClickHouse

What is Streamkap

Streamkap is a serverless SaaS platform for effortless streaming ETL



Real-time streaming of database data



Change data capture



Easy setup in minutes



In-stream transformations for filtering, routing, joins and more



Stream to ClickHouse and 14 other destinations

Sources

streamkap

Demo (Ricky)

Pipelines

Sources

Destinations

Logs

Chat

Docs & Updates >

Ricky Thomas >

Connect Source



MongoDB



DocumentDB



MySQL



PostgreSQL



SQL Server



kafka



elasticsearch



PostgreSQL

Name:

Hostname:

Port: 5432

Username:

Password:

Database:

Signal Table Schema:

Replication Slot Name: streamkap_pgoutput_slot

Publication Name: streamkap_pub

Advanced

Schema: 0 schemas + 

Destinations

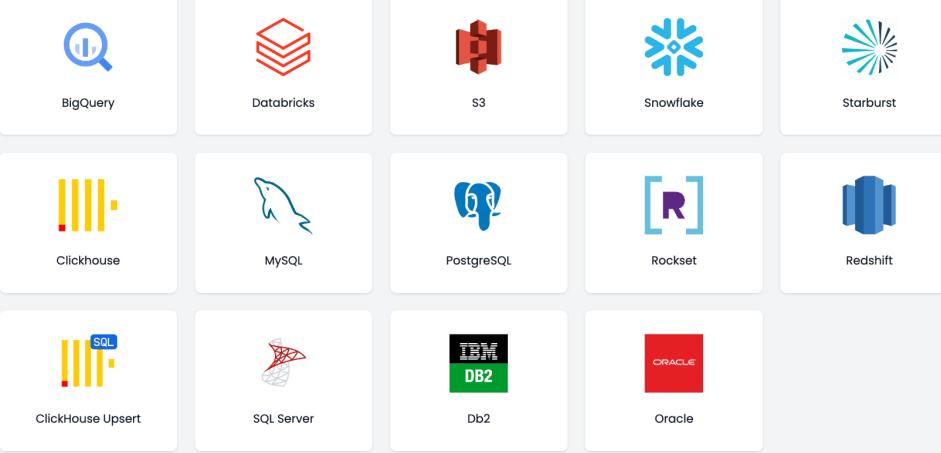
streamkap

Demo (Ricky) C
Pipelines
Sources
Destinations
Logs

Chat
Docs & Updates >

R Ricky Thomas >

Connect Destination



BigQuery Databricks S3 Snowflake Starburst
Clickhouse MySQL PostgreSQL Rockset Redshift
ClickHouse Upsert SQL Server Db2 Oracle

Clickhouse

Name:

Hostname:

Username:

Password:

Port:

Database:

SSL (?)

+ Save

Transforms

Running Jobs

Job Name	Start Time	Duration	End Time	Tasks	Status
transform_63d3dfedddcf4768d0004001.galleries.media_placements	2023-09-07 15:07:19	10d 4h 48m 52s	-	2 2	RUNNING
transform_63d3dfedddcf4768d0001002.libraries_filetree	2023-09-07 15:07:35	10d 4h 48m 36s	-	1 1	RUNNING
transform_63d3dfedddcf4768d0002002.libraries_filetree_folders	2023-09-07 15:08:00	10d 4h 48m 12s	-	1 1	RUNNING
transform_63d3dfedddcf4768d0007002.galleries.tags	2023-09-07 15:07:24	10d 4h 48m 47s	-	2 2	RUNNING
transform_63d3dfedddcf4768d0003002.libraries_filetree_folders_folders	2023-09-07 15:07:41	10d 4h 48m 31s	-	1 1	RUNNING
transform_63d3dfedddcf4768d0006001.galleries.tag_relationships	2023-09-07 15:07:46	10d 4h 48m 26s	-	2 2	RUNNING
transform_63d3dfedddcf4768d0007002.galleries.taghierarchy_transl_ation_aliases	2023-09-07 15:07:30	10d 4h 48m 42s	-	2 2	RUNNING
transform_63d3dfedddcf4768d0005001.galleries.tag_relationships_relations	2023-09-07 15:07:54	10d 4h 48m 17s	-	2 2	RUNNING
transform_63d3dfedddcf4768d0007002.galleries.taghierarchy_translations	2023-09-07 15:07:49	10d 4h 48m 22s	-	2 2	RUNNING

Pipelines

 streamkap

 Demo (Ricky) ▼

 Pipelines

 Sources

 Destinations

 Logs

 Chat

 Docs & Updates

 Ricky Thomas ›

Pipelines

 Create pipeline

 Search by name...

 Status ▼

 Source ▼

 Destination ▼

Name	Status	Source	Destination	Latency	Topics	⋮
Mongo > Snowflake  	Paused	Paused Idle - No new records	Active	-	1	⋮
Vendors  	Active	Active Streaming	Active	0s 1ms	1	⋮
Customers  	Active	Active Streaming	Active	0s 1ms	1	⋮

Metrics - App

Source Details

Express Dev

Status Schema Auth Settings Logs

Active • Streaming

Latency 0s 163ms Topics 196 Events Received 125.47K Errors 0

Activity

Time Since Last Event: 0m 3s
Latency Behind Source: 0s 163ms

of times

Events

Total # of Snapshot Records Total # of Create Events Received Total # of Update Events Received Total # of Delete Events Received

Streaming Breakdown	Delete Events	Update Events	Create Events	Time Since Last Event	Last Streaming Event
0s 163ms	12,926	33,459	156,485	0m 3s	XID/169487626000 from server 2886747454

Topics

Volume Event Retries Event Errors Events Written Last Event Snapshot

200/page

Data - Logs

Sep 17 20:09:36.456	prod-kafka-connect-01	kafka-connect	[SF_KAFKA_CONNECTOR] Fetched offsetToken for channelName:EXPRESS_REPL.DEV_APPETIZE...
Sep 17 20:09:36.429	prod-kafka-connect-01	kafka-connect	[SF_KAFKA_CONNECTOR] Fetched offsetToken for channelName:EXPRESS_REPL.DEV_APPETIZE...
Sep 17 20:09:36.394	prod-kafka-connect-02	kafka-connect	Flushing 9 buffered records BigQueryDatabaseDialect
Sep 17 20:09:36.393	prod-kafka-connect-02	kafka-connect	Flushing 3 buffered records BigQueryDatabaseDialect
Sep 17 20:09:36.393	prod-kafka-connect-02	kafka-connect	Received 4 records. First record kafka coordinates:(sales_flat_order-0-2792445). W...
Sep 17 20:09:36.388	prod-kafka-connect-02	kafka-connect	Flushing 2 buffered records BigQueryDatabaseDialect
Sep 17 20:09:36.066	prod-kafka-connect-03	kafka-connect	Flushing 2 buffered records BigQueryDatabaseDialect
Sep 17 20:09:35.977	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Register blob request preparing for blob=[2023/9/17/19/9/s1597y_h0kQX3...
Sep 17 20:09:35.977	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Start registering blobs in client=KC_CLIENT_destination_645e14d8e03fc1...
Sep 17 20:09:35.977	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Finish uploading file=2023/9/17/19/9/s1597y_h0kQX3B1xz6cSY9myn69Y9shnd...
Sep 17 20:09:35.823	prod-kafka-connect-02	kafka-connect	Processing 0 records
Sep 17 20:09:35.656	prod-kafka-connect-01	kafka-connect	Exported 55002306 of 196228491 records for table 'parklu_live.zr_kol_post_content_...
Sep 17 20:09:35.648	prod-redpanda-node-04	agent	2023-09-17 19:09:35 UTC CORE WARN (pkg/logs/internal/launchers/file/launcher...
Sep 17 20:09:35.648	prod-redpanda-node-04	agent	2023-09-17 19:09:35 UTC CORE INFO (pkg/logs/internal/tailers/file/tailer_nix...
Sep 17 20:09:35.648	prod-redpanda-node-04	agent	2023-09-17 19:09:35 UTC CORE INFO (pkg/logs/internal/launchers/file/launcher...
Sep 17 20:09:35.258	prod-redpanda-node-04	agent	2023-09-17 19:09:35 UTC CORE INFO (pkg/logs/internal/launchers/file/launcher...
Sep 17 20:09:34.925	prod-redpanda-node-04	rpk	INFO 2023-09-17 19:09:34,925 [shard 1] kafka - connection_context.h:141 - Connecti...
Sep 17 20:09:34.885	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Start uploading file=2023/9/17/19/9/s1597y_h0kQX3B1xz6cSY9myn69Y9shndL...
Sep 17 20:09:34.884	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Finish building chunk in blob=2023/9/17/19/9/s1597y_h0kQX3B1xz6cSY9myn...
Sep 17 20:09:34.877	prod-kafka-connect-02	kafka-connect	Got brand-new compressor [.gz]
Sep 17 20:09:34.877	prod-kafka-connect-02	kafka-connect	[SF_INGEST] Finish building chunk in blob=2023/9/17/19/9/s1597y_h0kQX3B1xz6cSY9myn...
Sep 17 20:09:34.874	prod-kafka-connect-02	kafka-connect	Got brand-new compressor [.gz]
Sep 17 20:09:34.874	prod-kafka-connect-02	kafka-connect	[SF_INGEST] buildAndUpload task added for client=KC_CLIENT_destination_645e14d8e03...
Sep 17 20:09:34.288	prod-kafka-connect-02	kafka-connect	Flushing 4 buffered records BigQueryDatabaseDialect

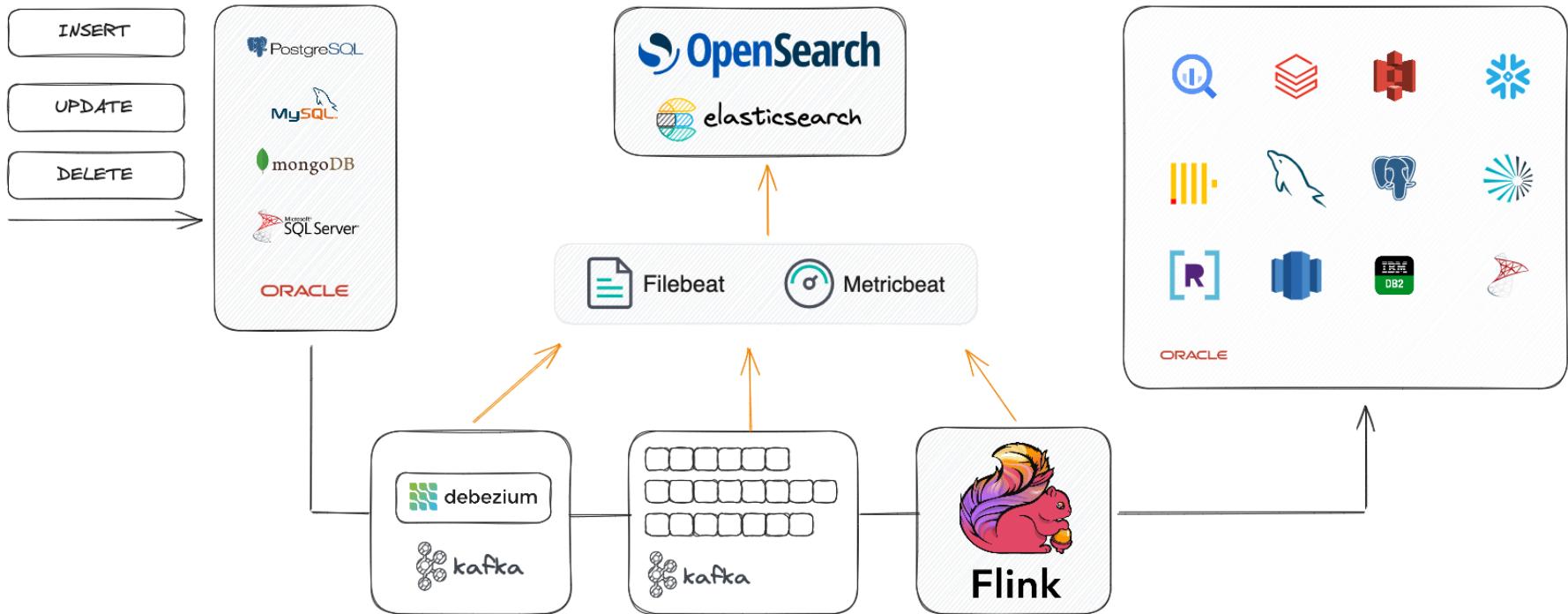
Data - Logs

	⌚ timestamp	ABC log
1	2023-09-19 08:11:24	[2023-09-19 08:11:24,806] INFO [task-thread-destination_6397538a676fa8f369d552]
2	2023-09-19 08:11:24	[2023-09-19 08:11:24,175] INFO [task-thread-destination_650194d7b5ae34ab6de2c5]
3	2023-09-19 08:11:24	[2023-09-19 08:11:24,125] INFO [task-thread-destination_650194d7b5ae34ab6de2c5]
4	2023-09-19 08:11:23	[2023-09-19 08:11:23,883] INFO [task-thread-destination_650198dbe2a77a14a03675]
5	2023-09-19 08:11:23	[2023-09-19 08:11:23,829] INFO [task-thread-destination_650198dbe2a77a14a03675]
6	2023-09-19 08:11:23	[2023-09-19 08:11:23,572] INFO [task-thread-destination_6397538a676fa8f369d552]
7	2023-09-19 08:11:23	[2023-09-19 08:11:23,045] INFO [ingest-register-thread] [destination_645e14d8e03fc]
8	2023-09-19 08:11:22	[2023-09-19 08:11:22,603] INFO [task-thread-destination_645927530133b40730fa44]
9	2023-09-19 08:11:22	[2023-09-19 08:11:22,663] INFO [task-thread-destination_6397538a676fa8f369d552]
10	2023-09-19 08:11:22	[2023-09-19 08:11:22,645] INFO [ingest-register-thread] [destination_645e14d8e03fc]
11	2023-09-19 08:11:22	[2023-09-19 08:11:22,645] INFO [ingest-register-thread] [destination_645e14d8e03fc]
12	2023-09-19 08:11:22	[2023-09-19 08:11:22,645] INFO [ingest-build-upload-thread-6] [destination_645e14c]
13	2023-09-19 08:11:22	[2023-09-19 08:11:22,630] INFO [task-thread-destination_6397538a676fa8f369d552]
14	2023-09-19 08:11:22	[2023-09-19 08:11:22,630] INFO [task-thread-destination_6397538a676fa8f369d552]

Data - Metrics

	⌚ updated_at	ABC metric_name	.labels	123 value
▼ 185	2023-09-19 08:08:55	kafka_consumer_fetch_manager_records_lag	▼ {availability_zone=us-west-2a,clientId=connector-consumer-de,availability_zone=us-west-2a,clientId=connector-consumer-destination_6397538a676fa8f369d,cluster_id=streamkap-prod,instance=prod-kafka-connect-02,job=kafka-connect,partition=0,private_dns_name=ip-10-0-3-142.us-west-2.compute.internal,prometheus=prometheus/prometheus-kube-prometheus-prometheus/prometheus_replica=prometheus-prometheus-kube-prometheus-region=us-west-2,topic=source_63975020676fa8f369d5526c_brandalley_live_inver}	0
185.1				
185.2				
185.3				
185.4				
185.5				
185.6				
185.7				
185.8				
185.9				
185.10				
185.11				
▼ 186	2023-09-19 08:08:55	debezium_metrics_lastevent	▼ {availability_zone=us-west-2a,cluster_id=streamkap-prod,content_type=application/json,context=streaming,instance=prod-kafka-connect-02,job=kafka-connect,lastevent=position: {ord: 2}, key: {"id": {"\$oid": "650957040fd72"}, name=source_63d3dfedddcf4768d13ac00b, task=0, plugin=mongodb,private_dns_name=ip-10-0-3-142.us-west-2.compute.internal,prometheus=prometheus/prometheus-kube-prometheus-prometheus/prometheus_replica=prometheus-prometheus-kube-prometheus-region=us-west-2}	1
186.1				
186.2				
186.3				
186.4				
186.5				
186.6				
186.7				
186.8				
186.9				
186.10				
186.11				
186.12				

Previous Architecture

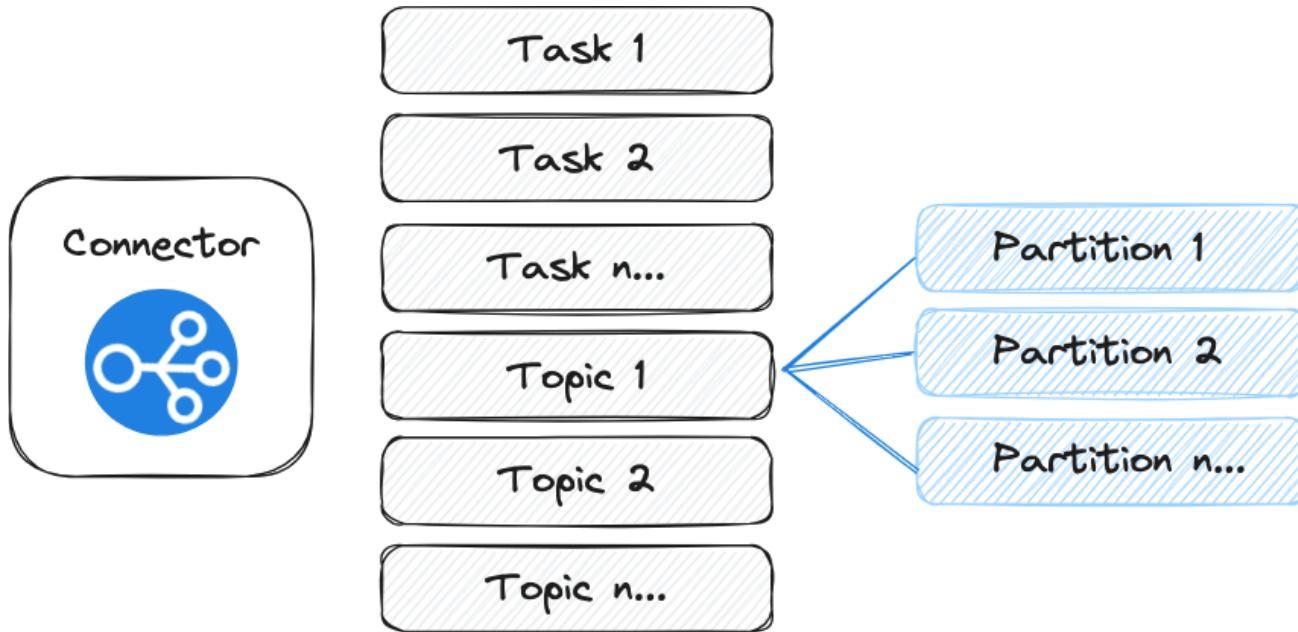


Granularity

Metric = recordsLag

Reporting by = client-id,tasks, topics, partitions at various intervals

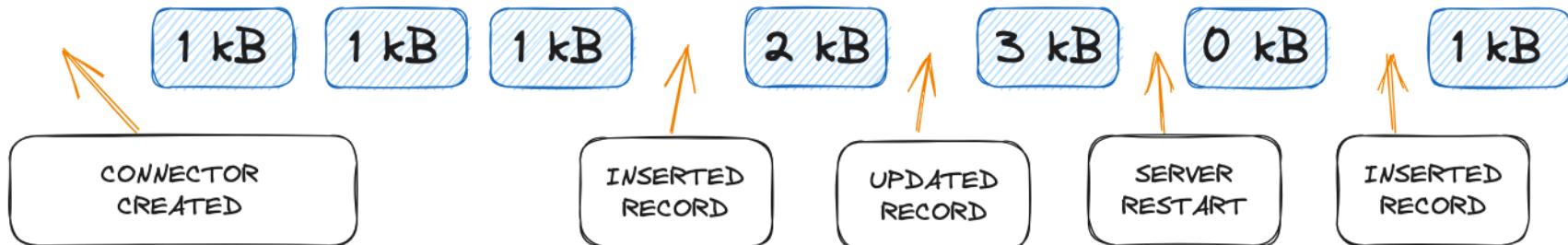
Lag for a connector with 5 tasks, 100 topics, 10 partitions for topic, we must rollup $5 \times 100 \times 10 = 5,000$ separate metrics



State at Time Reporting

Example: Metric = bytesConsumedTotal

JMX metrics are state at time and do not survive restarts or provide changes over time



Key Metrics needed in real-time

Customer Facing

- Status (Connectors, Streaming, Snapshots)
- Volume (GB)
- Events (Read, Written, Last Event, Lag)
- Job Status
- Logs - Errors, Linking to entity
- Latency

Internal

- All above at Connector/Topic/Partition level
- Producer and Consumer Health
- Throughput (Snapshots, Streaming, Jobs, Network)
- Partitions (Quantity, Distribution)
- Cluster Health
- Kubernetes
- System Metrics - i.e. Grafana/Datadog
- Storage (Hot, Tiered Storage)
- Limits

So what happened?

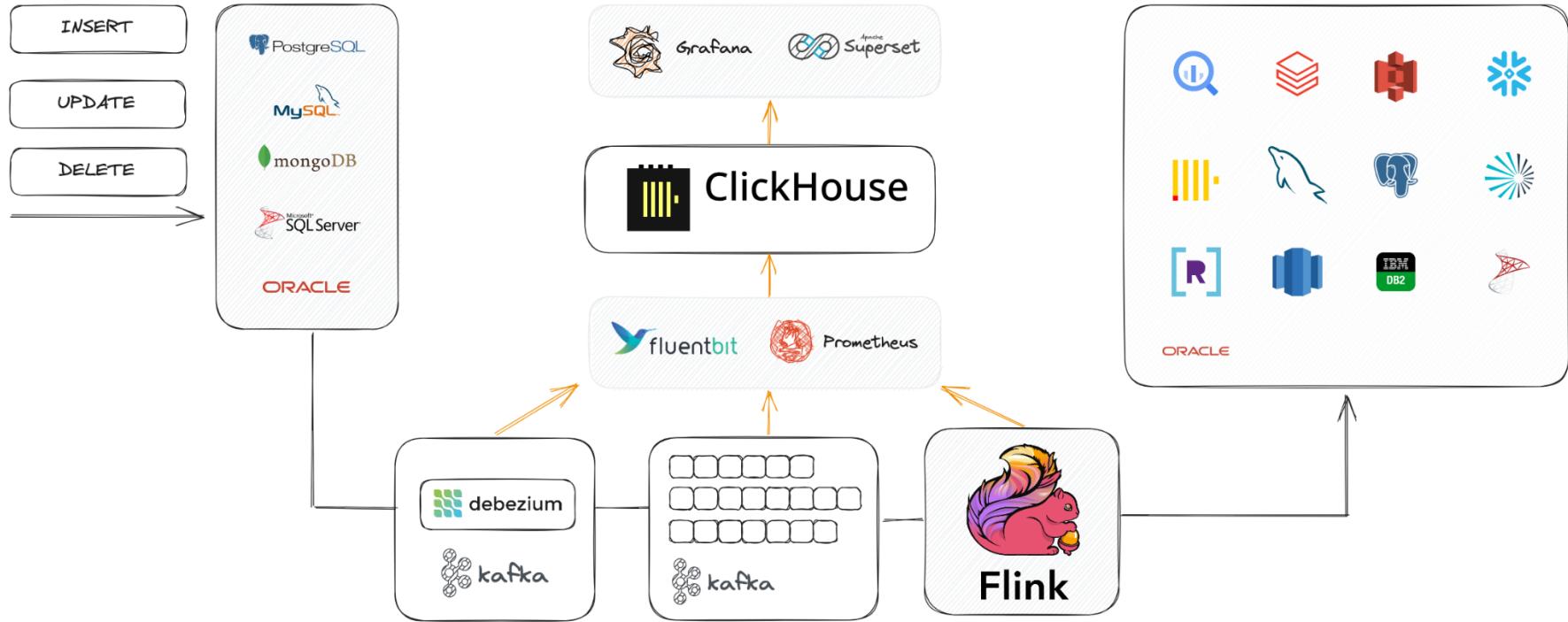
Despite (to name a few)

- Indices
- Elastic's roll up feature - Unable to handle massive indices of historical data
- Pagination
- Hot & warm caches
- Slowing down updates (bad for us)
- Sizing

We faced

- Slow queries
- Delayed results
- High CPU/contention
- Out of memory
- Modelling issues
 - Like all NoSQL databases, joins are expensive
 - Couldn't aggregate and group by categories efficiently (i.e. sum, group by topic)
 - Painful to work with without SQL
- High costs

New Architecture



How did we solve the earlier challenges

SQL allows us to execute a single query, and efficiently group at level we want to report (e.g partition, connector, client_id, task, topic, context)

```
select updated_at, metric_name, name, partition, connector, client_id, task, topic,
      groupArray(status) as status_array, groupArray(value) as values_array
   FROM default.metrics group by updated_at, metric_name, name, partition, connector,
client_id, task, topic
```

Arrays and array functions used to apply custom aggregations -
views store calculated changes in state (e.g. bytes) used for aggregating over time

```
case when metric_name in
('kafka_producer_topic_byte_total','debezium_metrics_totalnumberofeventsseen"...') then
    arrayCumSumNonNegative(x -> greatest(0, x), arrayDifference(values_array))
[length(values_array)]
```

Views combined with TTL to rollup historical data

```
SELECT toStartOfInterval(updated_at_to_minute, INTERVAL X minute) as updated_at_to_X_minute ...
```

Metric Flow



Rolled Up Historical Data TTL

Materialized View -
ReplicatedMergeTree

Delete expired Raw
and Differential records

Moving to ClickHouse Cloud

ClickHouse solved our metric issues but we were left with the management

- Kubernetes
- Monitoring/Alerting
- Availability
- Scaling
- Tuning
- Backups
- Multi-Region
- Labour

We moved to ClickHouse Cloud to solve this

Summary of moving to ClickHouse

- Speed - 98% faster
- Costs 90% lower
- Scalable
- No management
- No limits
- Support
- Happy customers

CDC to ClickHouse

 streamkap

