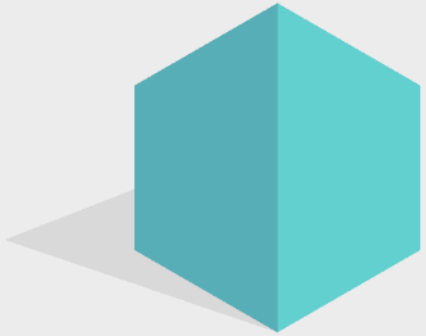


www.clonera.net Levent Karakaş

Real-time Streaming of Zabbix Metrics to Big Data Platforms

Zabbix Summit 2020

CLONERA

Storing Metrics in RDBMS

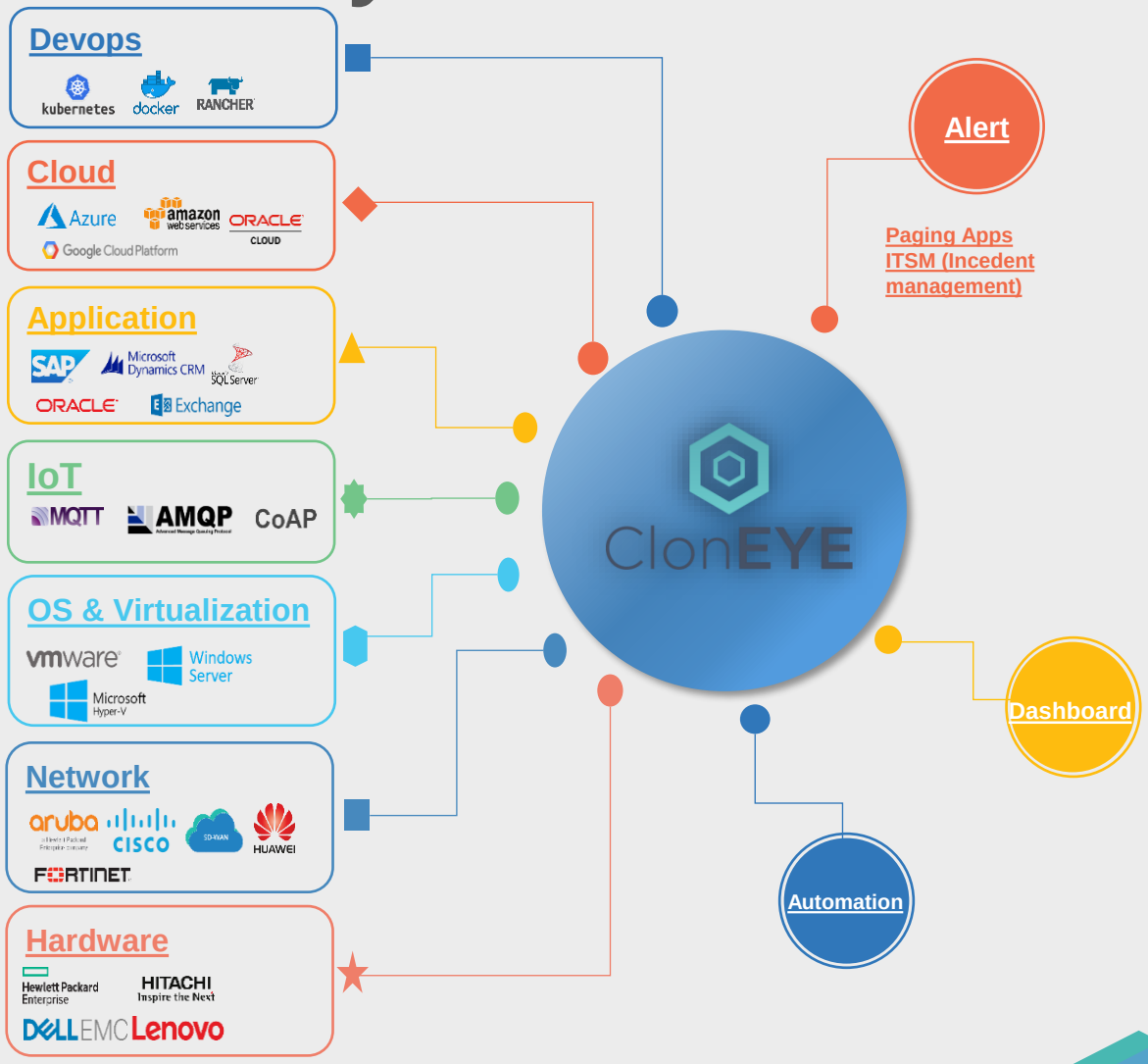
- Relational databases are not the ideal solution for storing metrics at scale
- Timescaledb support in Zabbix since version 4.2 is a huge relief (10x performance improvement)
- What if there are terabytes of metric data to manage?

CLONERA

Who needs to manage terabytes of metric data?

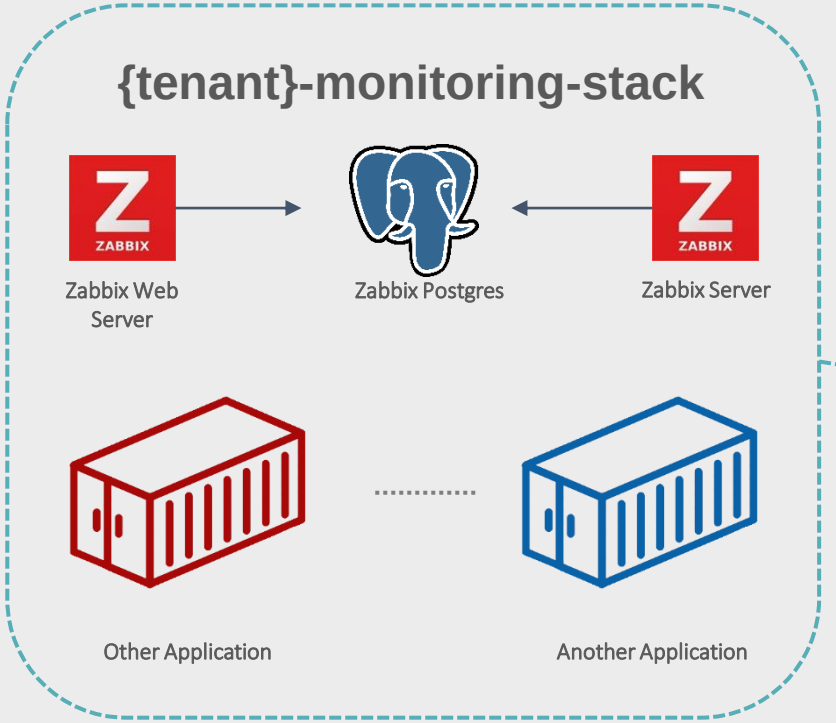
ClonEYE

- Best of breed monitoring tools, all integrated
- Zabbix is the main pillar
- Multi tenant
- Single-Sign-On (SSO)
- Everything runs on containers
- CI/CD Orchestration

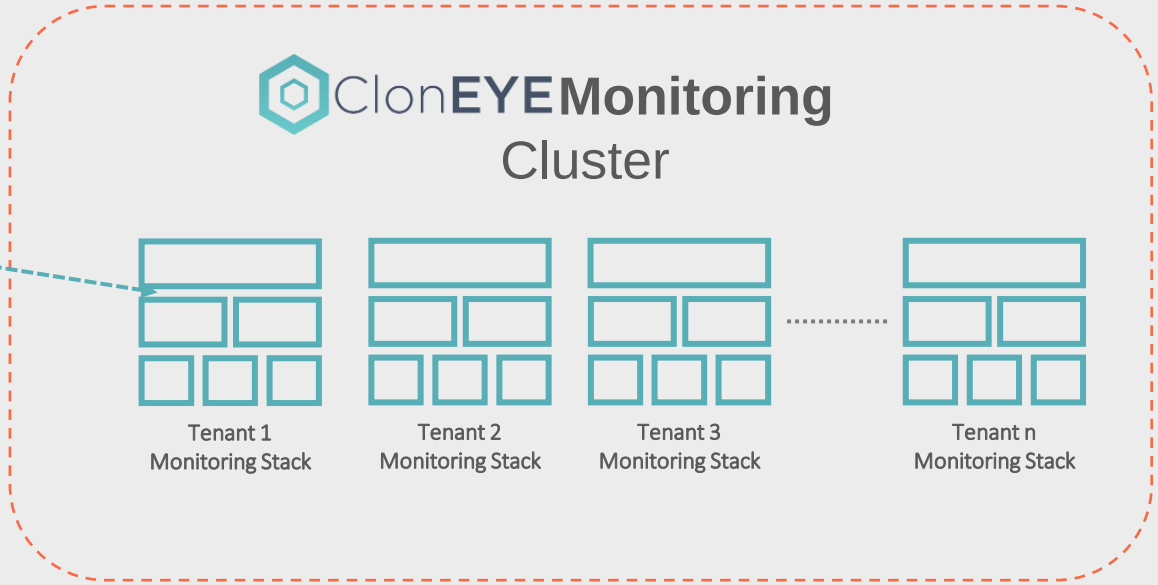


CLONERA

ClonEYE Plarform



Many containers in each stack



One stack for each tenant

CLONERA

How to store lots of metrics?

- Cost effective
- Large Datasets
- Replication
- Fault Tolerance
- High Availability
- Scalability
- High throughput

- Cloud Native
- Stream Native
- Batch Ingestion
- Time optimized partitioning
- Horizontal Scalibility
- Flexible Schema
- SQL Support

CLONERA

How to import data?

- Batch Ingestion
  - Bulk data import
  - Supports many file formats and file sources
  - Task based
- Streaming
  - Real-time
  - Druid service ingests directly from streams

CLONERA

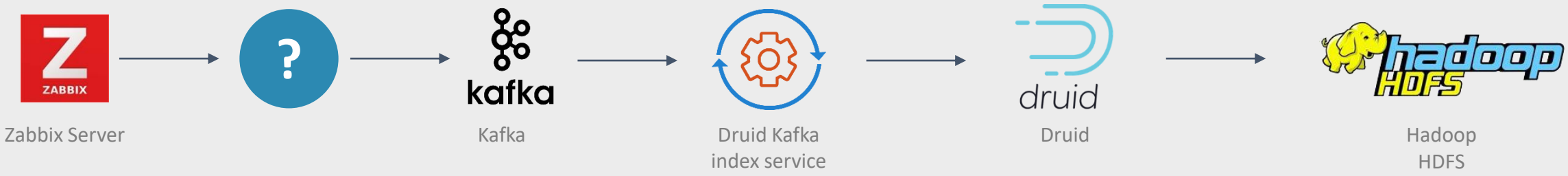
Zabbix Real-Time Export

- Zabbix supports exporting to file (which can be batch digested)
- File format is JSON (which is supported by Druid ingestion)
- A new file is created when the file reaches to a size (They should be moved, imported, deleted?)

https://www.zabbix.com/documentation/current/manual/appendix/install/real_time_export

CLONERA

Kafka for real-time streaming

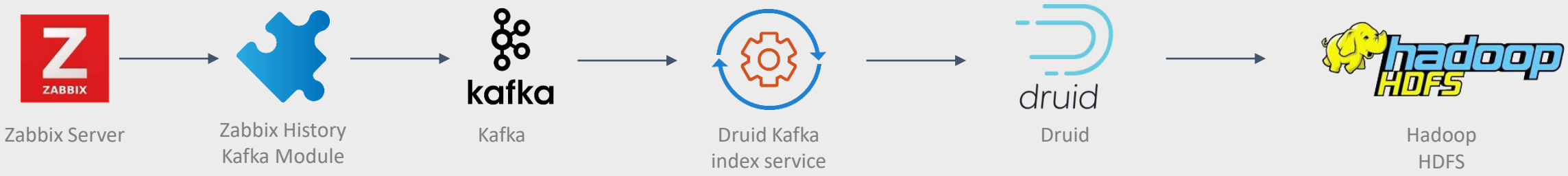


Zabbix Server | Kafka | Druid Kafka index service | Druid | Hadoop HDFS

- If we can stream zabbix events to kafka, Druid Kafka index service can ingest
- No files to manage!

CLONERA

Zabbix History Kafka Loadable Module



Zabbix Server → Zabbix History Kafka Module → Kafka → Druid Kafka index service → Druid → Hadoop HDFS

- Found a module in github
- Tested and realized that it is not working
- Tried fixing

- Tried writing a brand new one
- After lots of debugging, realized it is not possible to keep librdkafka state open and reuse for events
- Gave up

CLONERA

Use a REST to kafka middleware

- Kafka REST by Confluent

  https://github.com/confluentinc/kafka-rest

- Kafka Pixy by Mailgun

  https://github.com/mailgun/kafka-pixy

CLONERA

Zabbix History Webhook Module

- Formats and writes Zabbix history to a webhook in zabbix real-time export protocol in JSON (one additional tag: item key)
- Full support for float, integer, string, text and log
- Content type can be specified (default: application/json)
- SSL verification errors can be ignored

CLONERA

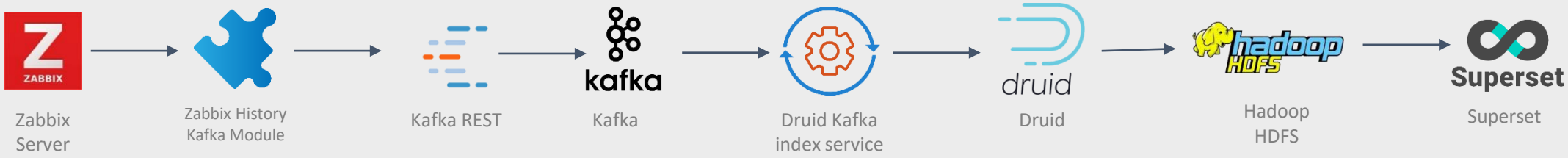
Zabbix History Webhook Module

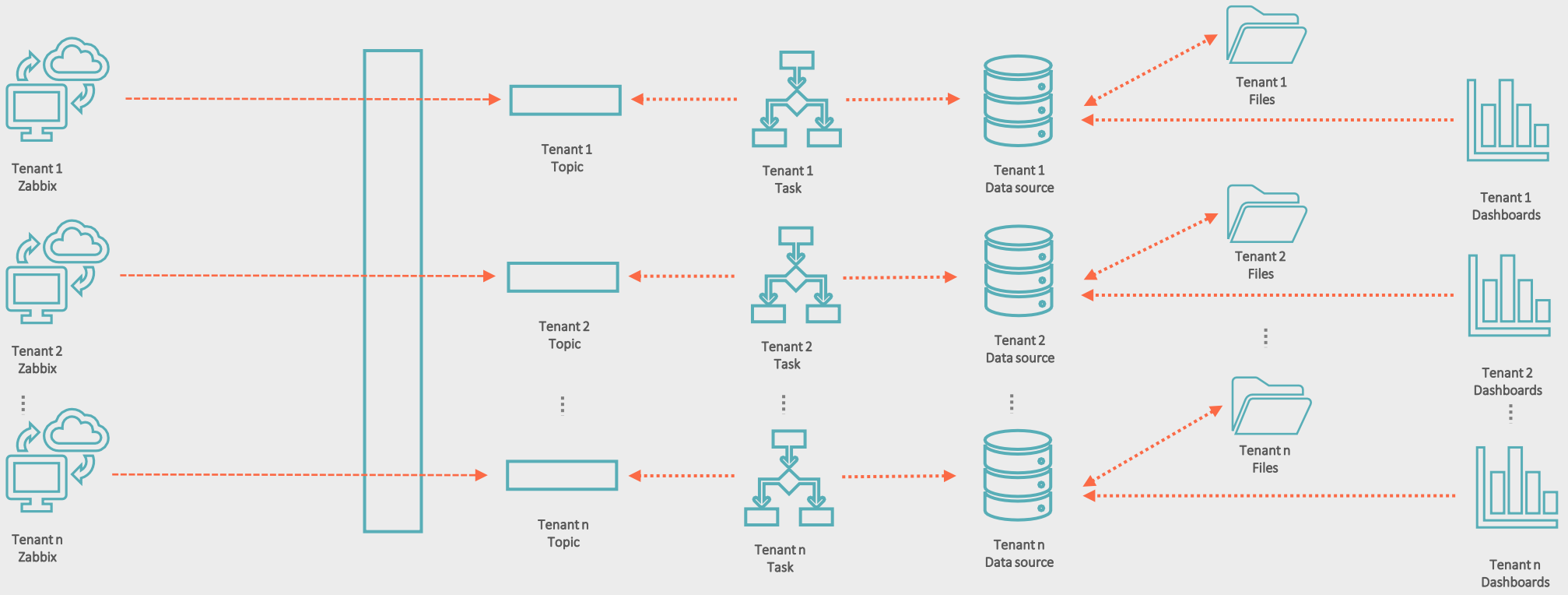
- Ability to call webhook per measurement or multiple measurements (bulk mode)
- Possibility to use custom tags while exporting in bulk mode (needed for kafka rest)
- Environment variables supported (for ease of use with containers)
- Can be used to export to any system that accepts web push

https://github.com/clonera/zabbix-history-webhook

CLONERA

Final Topology





CLONERA

Configuration #1: Zabbix Server

Environment variables to Zabbix Server Container

ZBX_LOADMODULE = "history_webhook.so"

ZBX_WEBHOOK_URL = "http://${kafka-rest}:38082/topics/${tenant}-zabbix"

ZBX_WEBHOOK_CONTENT_TYPE="application/vnd.kafka.json.v2+json"

ZBX_WEBHOOK_ENABLE_TEXT = "1"

ZBX_WEBHOOK_ENABLE_STRING = "1"

(float and integer is enabled by default we also enable text and string here)

https://github.com/clonera/zabbix-history-webhook

CLONERA

Configuration #2: Kafka Rest

Environment variables to Kafka Rest Container

KAFKA_REST_ZOOKEEPER_CONNECT = "${zookeper}:2181"

KAFKA_REST_HOST_NAME = "${kafka-host}"

KAFKA_REST_LISTENERS = "http://0.0.0.0:38082"

https://docs.confluent.io/current/kafka-rest/quickstart.html

CLONERA

Configuration #3: Kafka

Environment variables to Kafka Container

KAFKA_BROKER_ID = 3

KAFKA_ZOOKEEPER_CONNECT = "${zookeper}:2181"

KAFKA_ADVERTISED_HOST_NAME = "${hostname}"

KAFKA_LISTENERS = "PLAINTEXT://0.0.0.0:9092"

KAFKA_LOG_RETENTION_HOURS = 48

(9092 port is published on the host in this configuration)

CLONERA

Configuration #4: Druid HDFS

Download druid-hdfs-storage into extension folder and change configuration file:

vi conf/druid/cluster/_common/common.runtime.properties

druid.extensions.loadList=[**"druid-hdfs-storage"**, "druid-kafka-indexing-service", "druid-datasketches", "druid-influx-extensions", "postgresql-metadata-storage"]

druid.storage.type=hdfs

druid.storage.storageDirectory=hdfs://hadoop.local:8020/apps/druid/warehouse

https://druid.apache.org/docs/latest/development/extensions-core/hdfs.html

CLONERA

Configuration #5: Druid Kafka indexing

Download druid-kafka-indexing-service into extensions folder and change configuration file:

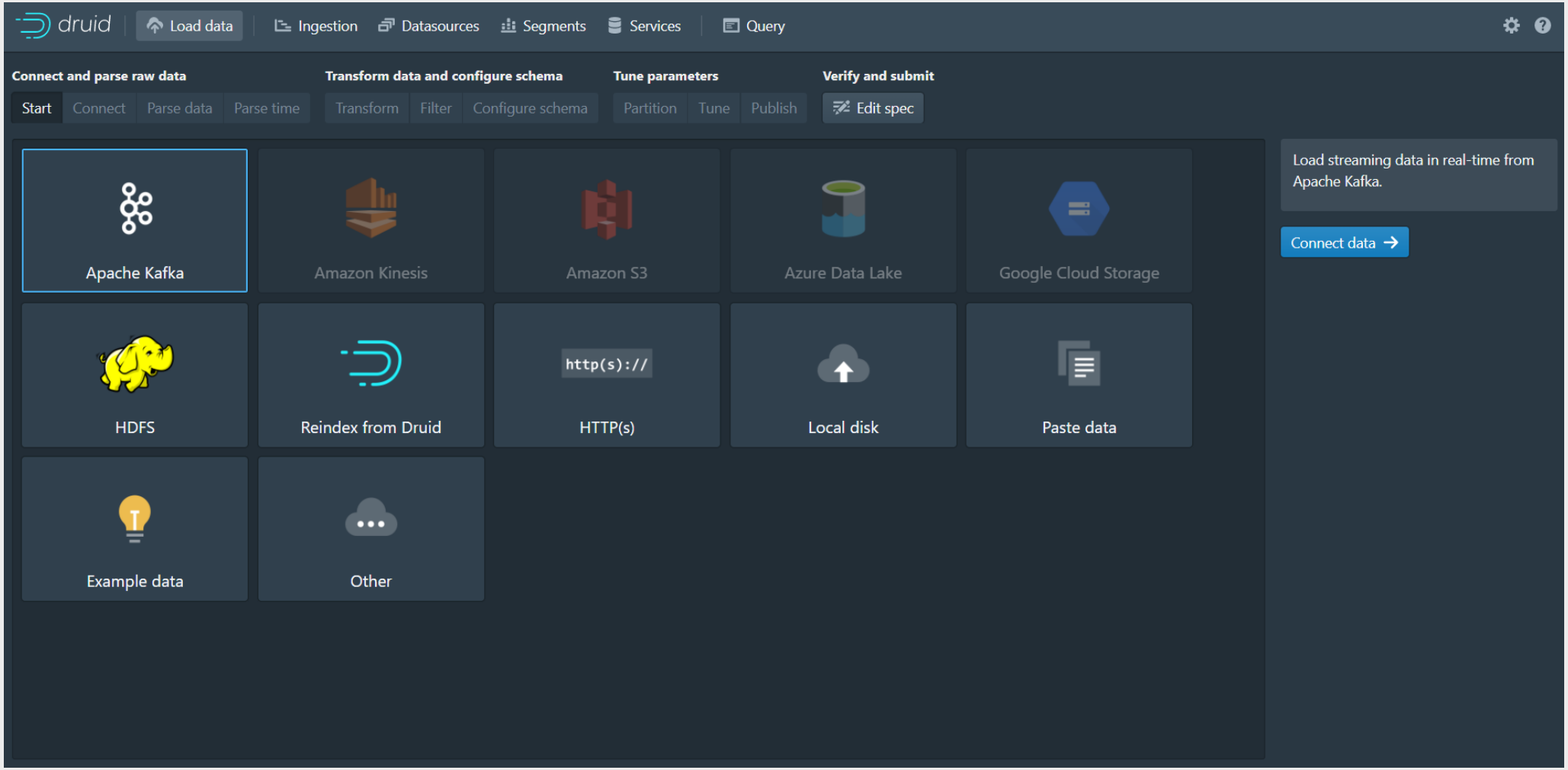
vi conf/druid/cluster/_common/common.runtime.properties

druid.extensions.loadList=["druid-hdfs-storage", **"druid-kafka-indexing-service"**, "druid-datasketches", "druid-influx-extensions", "postgresql-metadata-storage"]

https://druid.apache.org/docs/latest/development/extensions-core/kafka-ingestion.html

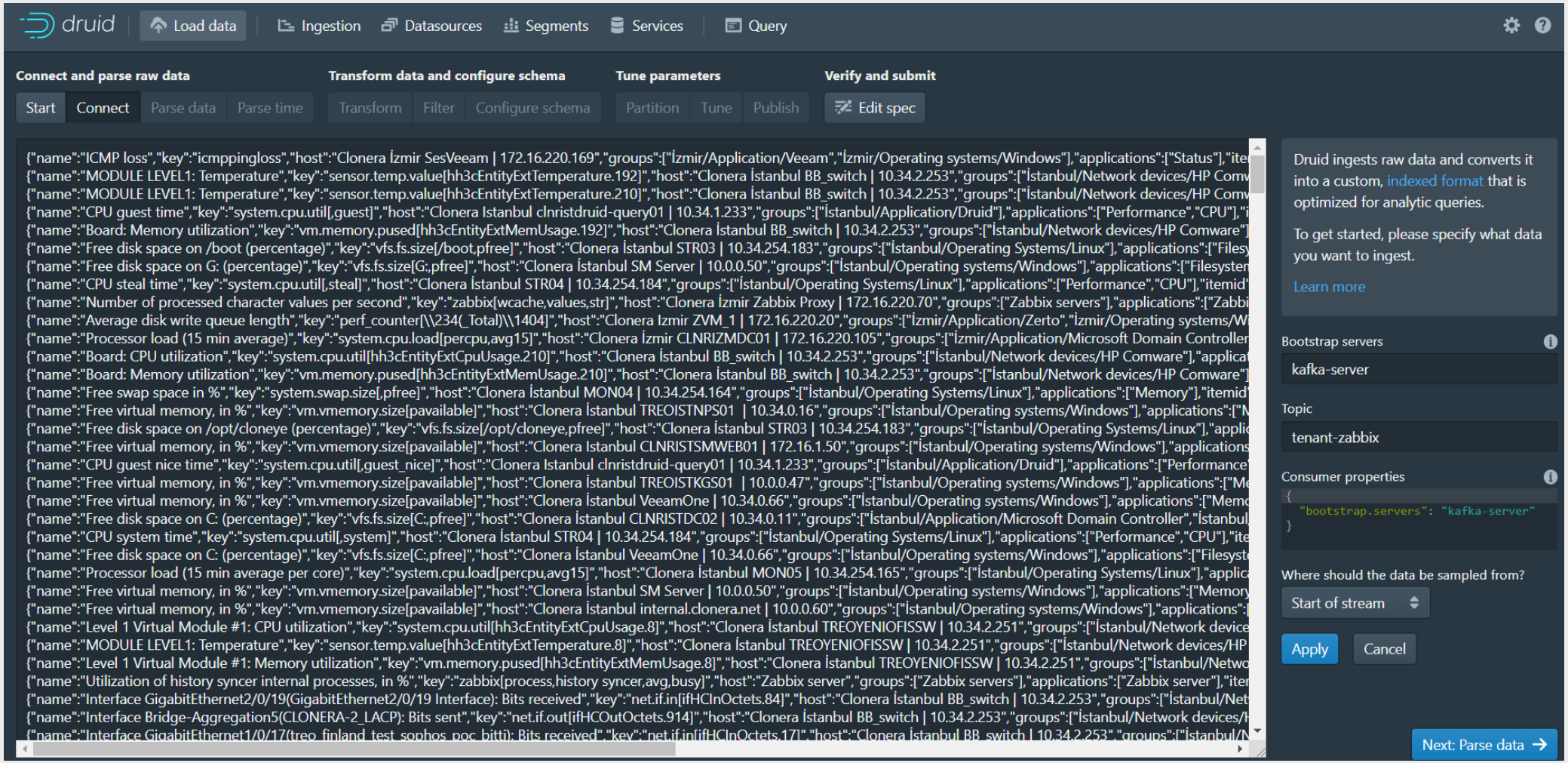
CLONERA

Configuration #6: Druid ingestion



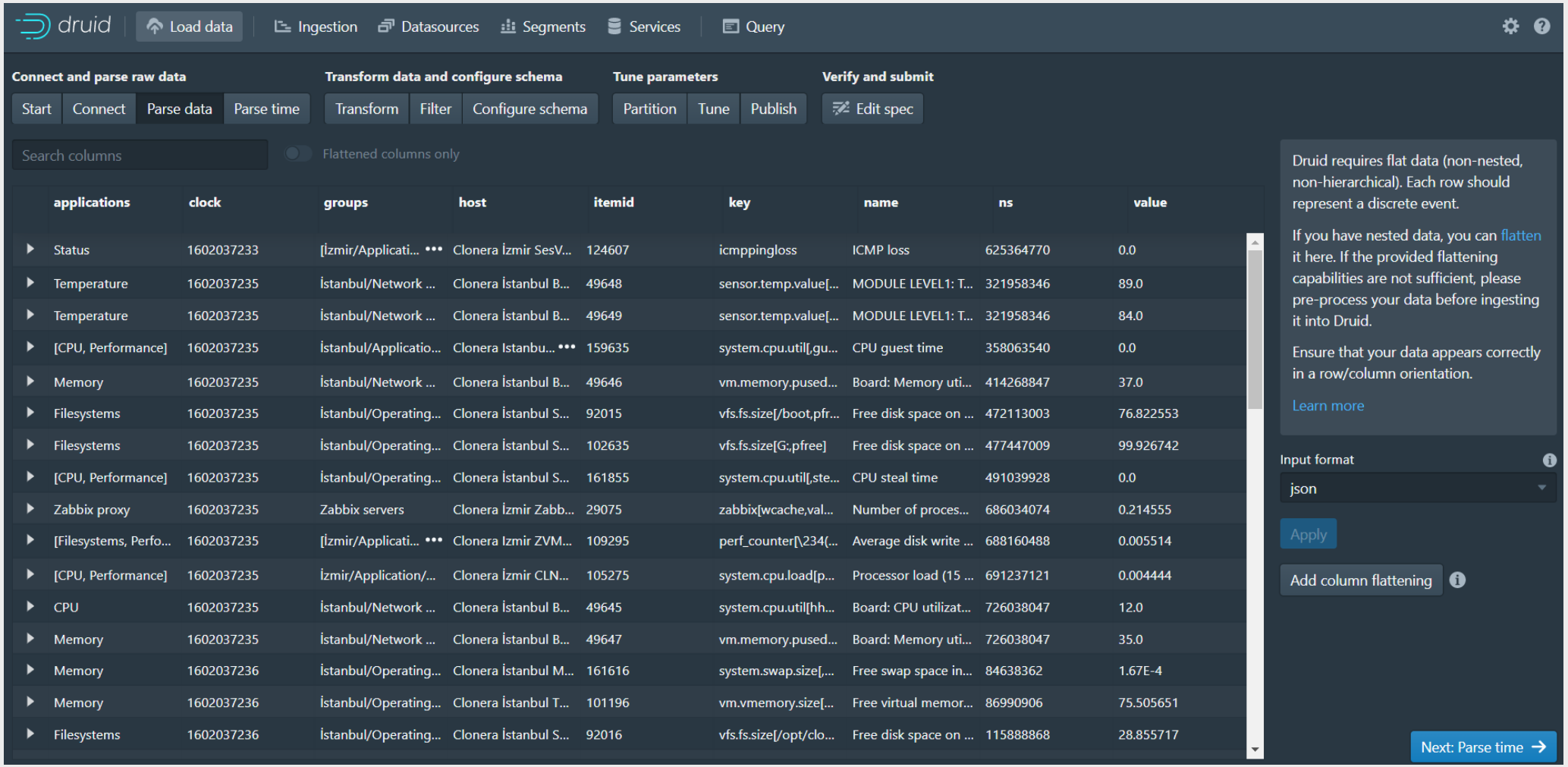
CLONERA

Configuration #6: Druid ingestion



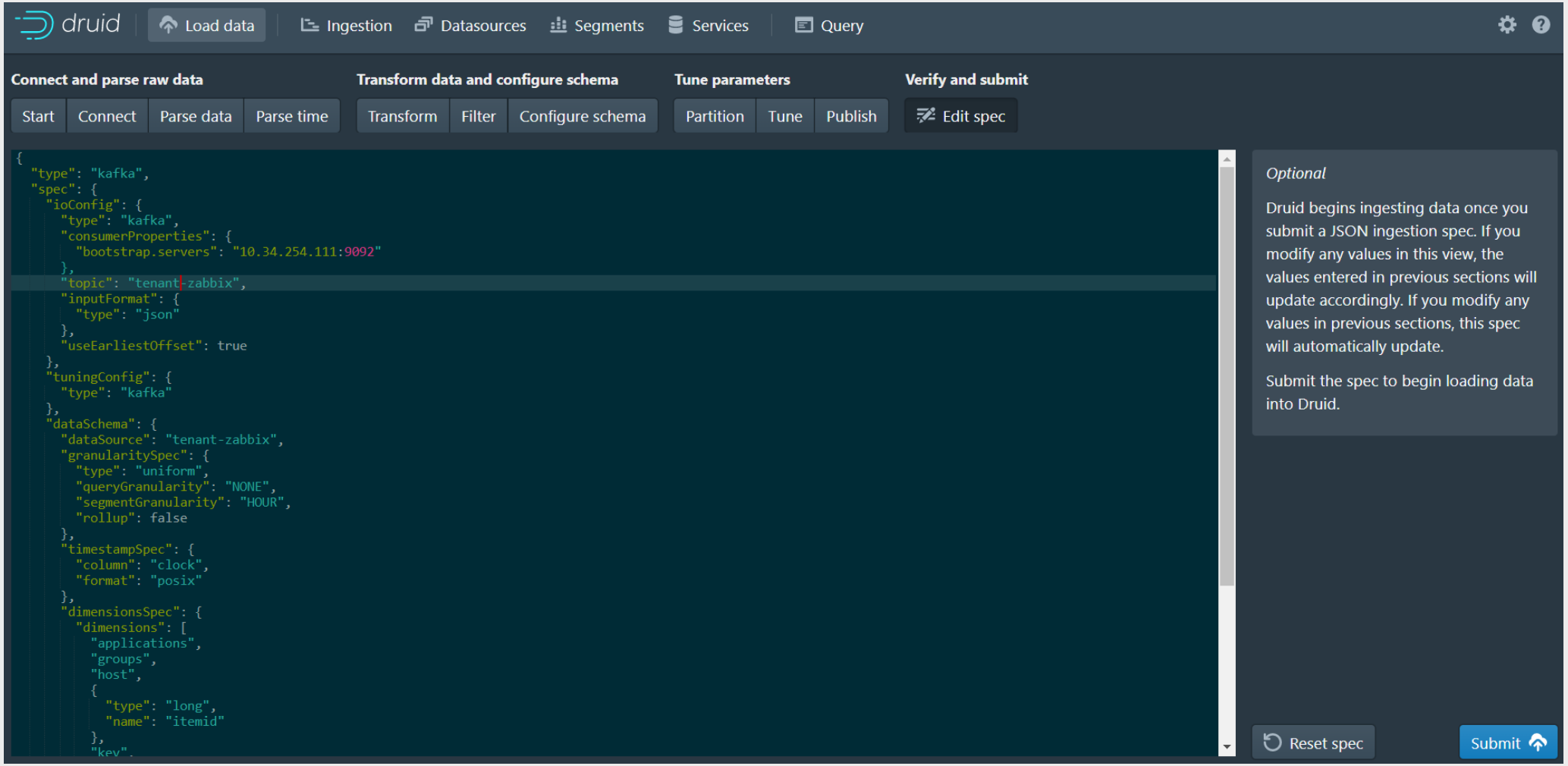
CLONERA

Configuration #6: Druid ingestion



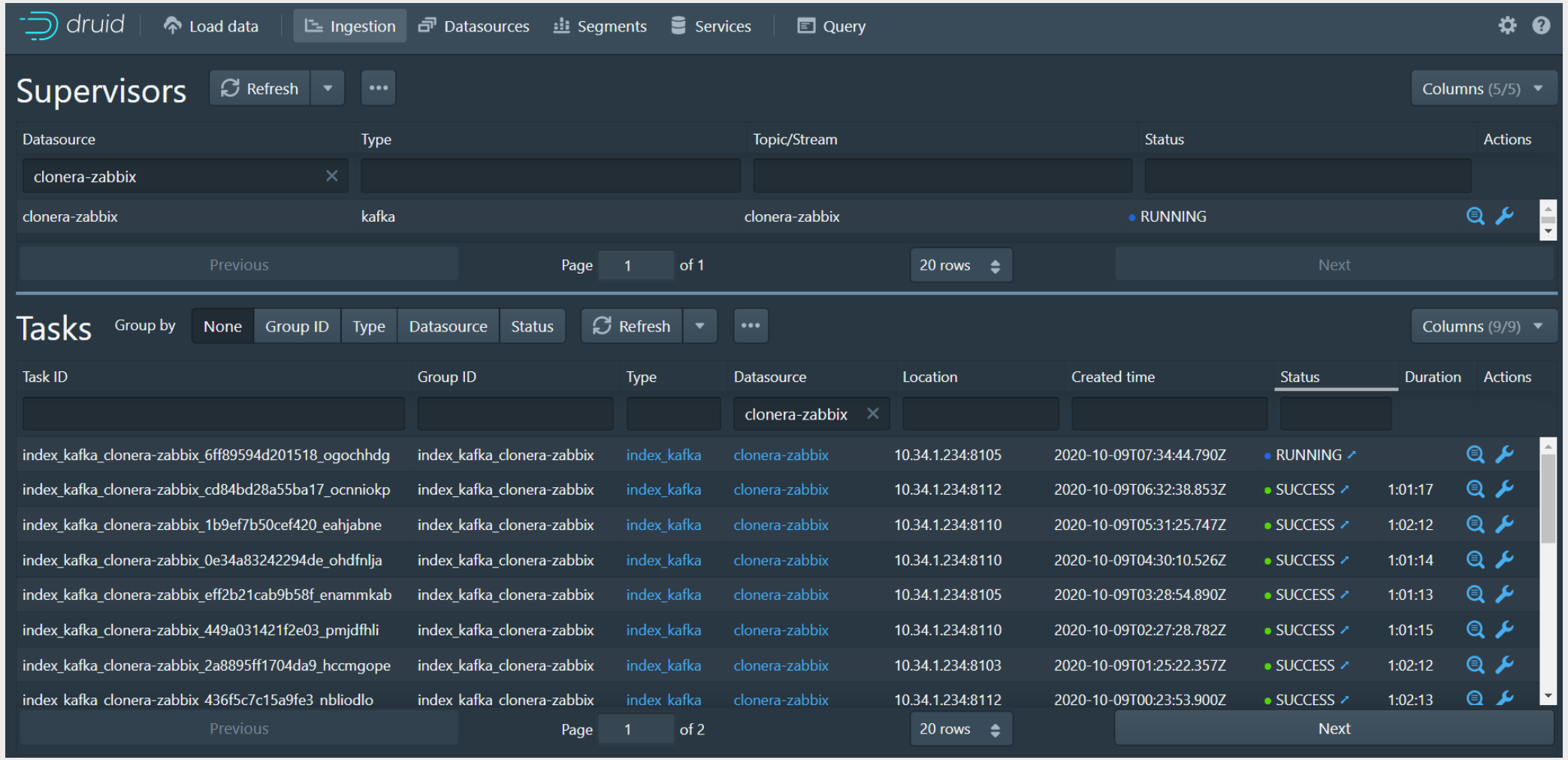
CLONERA

Configuration #6: Druid ingestion



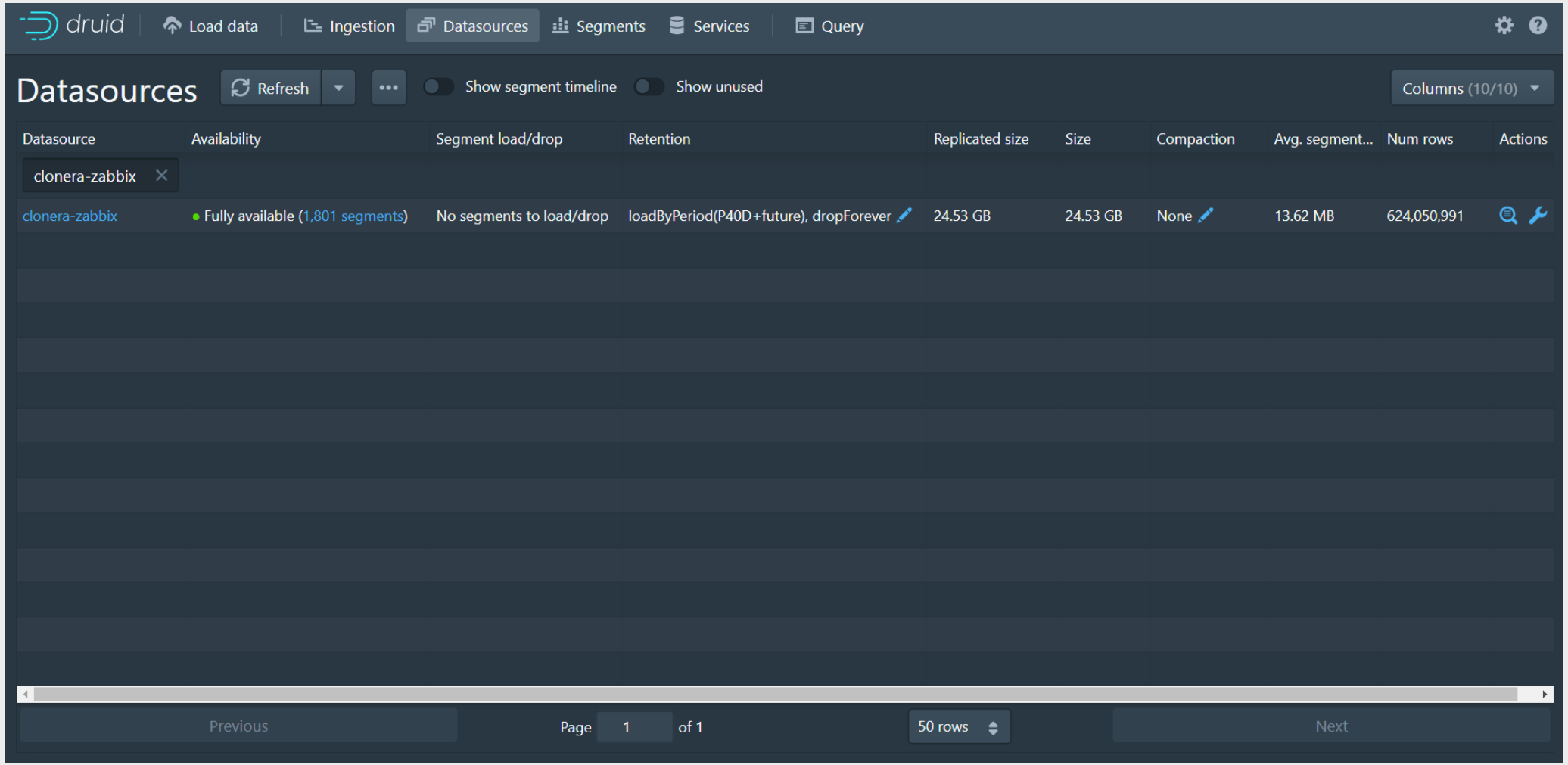
CLONERA

Configuration #6: Druid ingestion



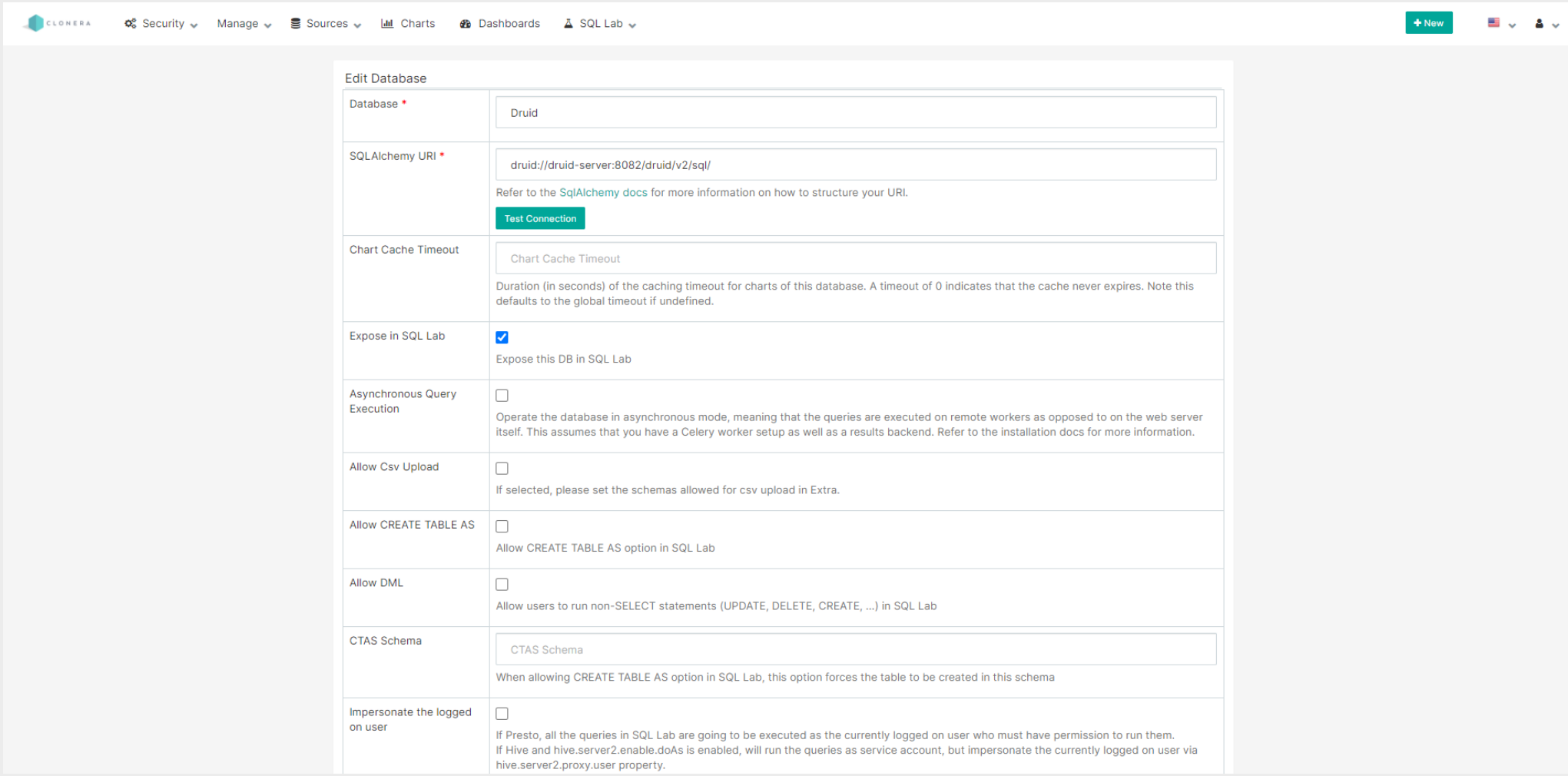
CLONERA

Configuration #6: Druid ingestion



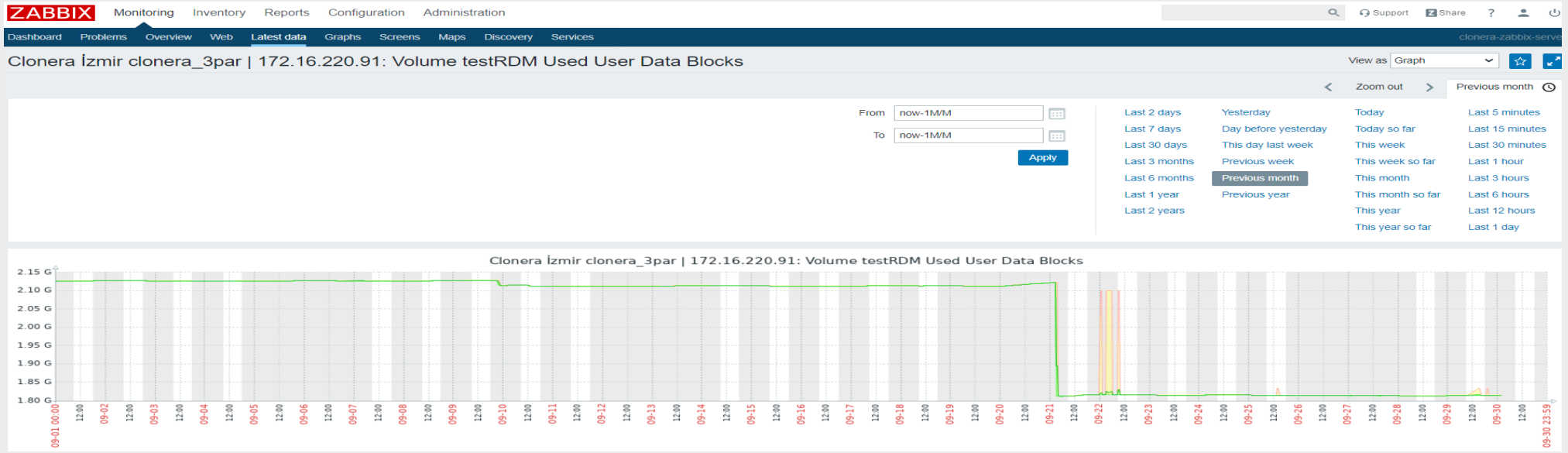
CLONERA

Configuration #7: Superset

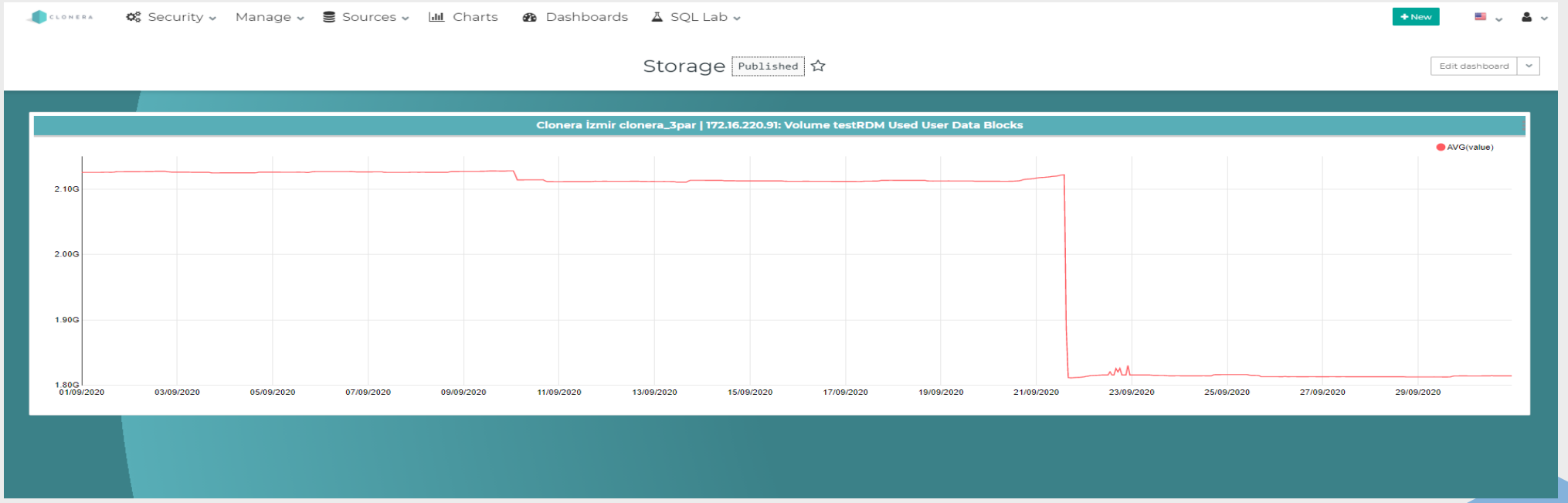


CLONERA

Zabbix



Superset



CLONERA

Development Team

DevOPS Engineer	Software Developer	BI Expert	Software Developer
Doğuş Peynirci	Burak Köseoğlu	Rahma Bayhatun	Nurdan Kolay

CLONERA

Thanks