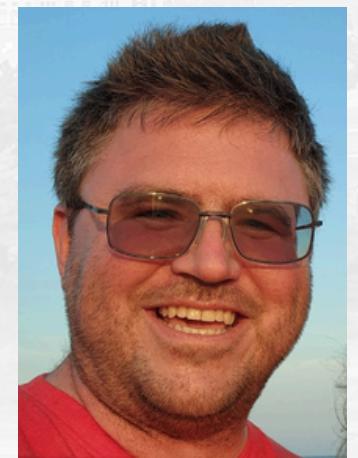




DEMO - SHIP DATA

July, 2024

www.singlestore.com



Tyler Denton,
Enterprise Architect



Goals

- Quickly Ingest Data from Parquet MMSI Files
- Represent that data visually
- Identify MMSI based on Polygon Search
- Aggregate Data Points into Routes
- Demonstrate data streaming through Kafka



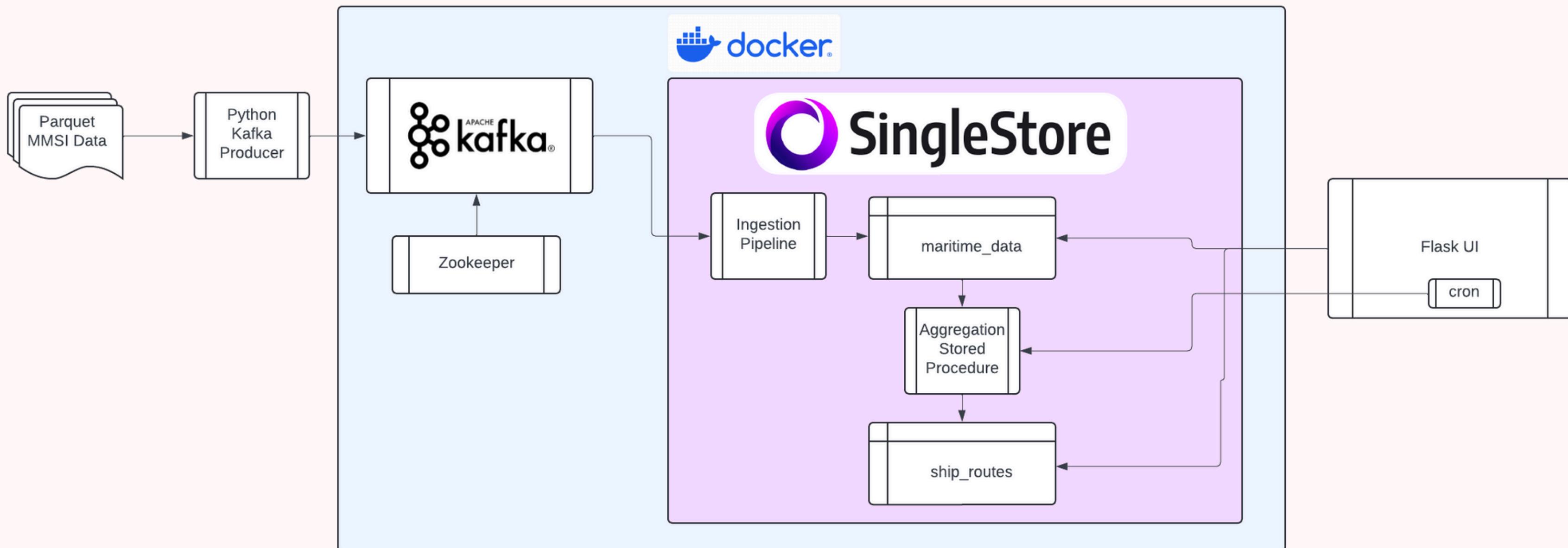
Current State

- We have many Historical Parquet files
- Parquet files contain geopoints. Very hard to query against and results in a very large dataset
- Have data streaming from ships in addition to parquet data
- No idea which ships tracks are accurate



SingleStore is dedicated to helping businesses adapt more quickly, embrace diverse data and accelerate innovations.

Solution



Definition of Success

Fast Data Ingestion

By specifying shard keys we can really speed up ingestion.

Kafka Integration

Pipelines allow for very simple integration with kafka that consumes different schemas. We are using JSON as our schema.



Easy Queries

SingleStore is ANSI SQL compliant and MySQL wire compatible. Making writing queries and integrating existing tooling very simple.

Spatial Data Support

SingleStore supports a subset of WKT and Spatial data query functions to make spatial data access very simple and fast. You do not have to build specialized inverted indexes goeindexes to be able to quickly search across geodata.

Learnings and Limitations

- Spatial Data Functions are very limited and do not follow the ST_* Standard
- No python or Javascript UDFs, although there are WASM based UDAF
- No Scheduled triggers, materialized views, partial aggregation indexes
- Shard Keys and Sort Keys can make a big difference in query performance
- GeoData has to be stored in a Rowstore
- Kafka Integration is very easy
- Queries can be very fast even for large result sets

Links

SingleStore

Kaggle Dataset

<https://github.com/tdenton8772/singlestore-project>

<https://docs.singlestore.com/cloud/reference/sql-reference/geospatial-functions/>

<https://docs.singlestore.com/cloud/reference/information-schema-reference/pipelines/pipelines/>

<https://docs.singlestore.com/cloud/reference/sql-reference/procedural-sql-reference/create-procedure/>



THANK YOU

FOR YOUR ATTENTION

