Predict with Prometheus

Using prometheus metrics for online data analysis - OLAP

João Sousa - 24 March 2022

“To predict the future we have to learn from the past”

- Unknown Source

Collect metrics from applications in real-time

- Increase service offering
- Adjust the price on demand
- Create analytical dashboard
- Forecast based on previous events

But how can this data be ingested and ready to be used?

OLAP

Online Analytical Processing

What is OLAP?

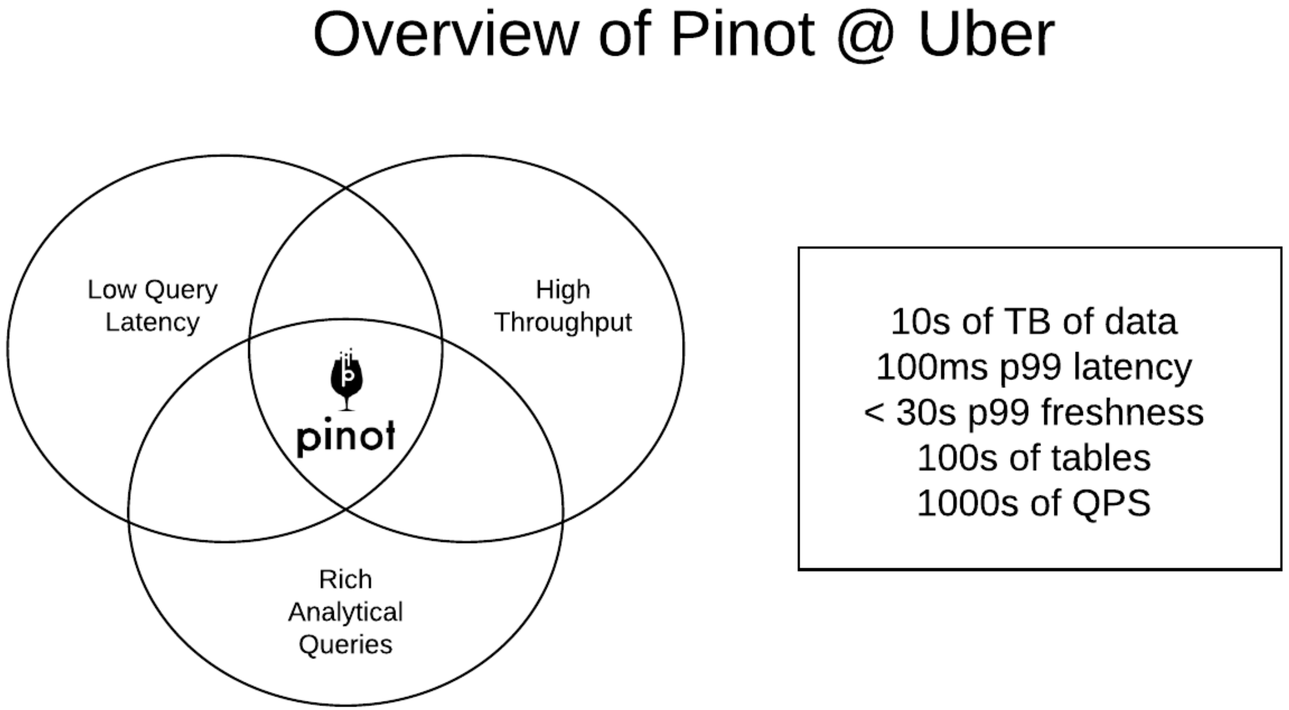
Online Analytical Processing

- Real-time (stream) data ingestion
- Freshness data can be queried
- Query large data sets
- Requires only read operations
- SQL statements can be limited

Lets take a look…

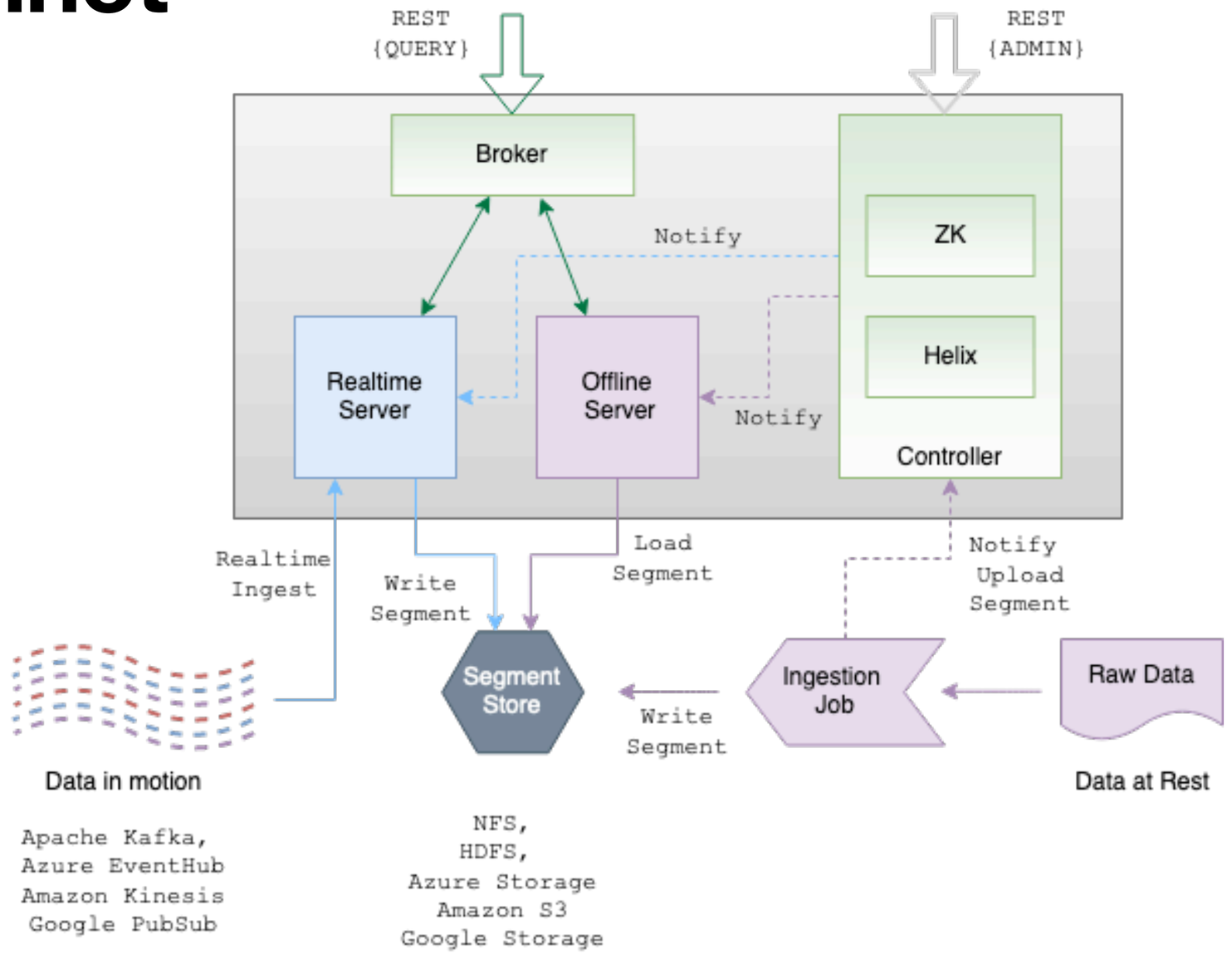
Uber Show case

- Dashboards
- Analytical applications
- Near-real-time exploration



Apache Pinot

Architecture



What a great story!!

… with Thanos

Prometheus

What is it? The Basics…

- OpenSource Project
- Pull metrics from exporters
- Metrics stored in TSDB
- Alert Manager
- Support PromQL query language
- Widely used in K8s solutions

What it isn’t

- Design for long-term storage
- Advance analytical queries
- Aggregation of huge historical data
- Slice and dice large data sets

```
select u.namespace,  u.day, u.sum/h.sum
from
(
    select namespace,  date_format(_timestamp_, '%d') as "day", sum(value) as "sum", max(_timestamp_) as maxtime
    from kube_resourcequota
    where prometheus_replica='prometheus-cmy-prometheus-operator-prometheus-0'
        and resource = 'requests.storage'
        and type = 'used'
    group by namespace,  date_format(_timestamp_, '%d')
) as "u"
join
(
    select namespace, date_format(_timestamp_, '%d') as "day" , sum(value) as "sum"
    from kube_resourcequota
    where prometheus_replica='prometheus-cmy-prometheus-operator-prometheus-0'
        and resource = 'requests.storage'
        and type = 'hard'
    group by namespace, date_format(_timestamp_, '%d')
) as "h"
on u.day=h.day and u.namespace = h.namespace
order by u.namespace, maxtime
```

Per namespace evolution of storage used/requested ratio over last week

Prometheus + Apache Pinot

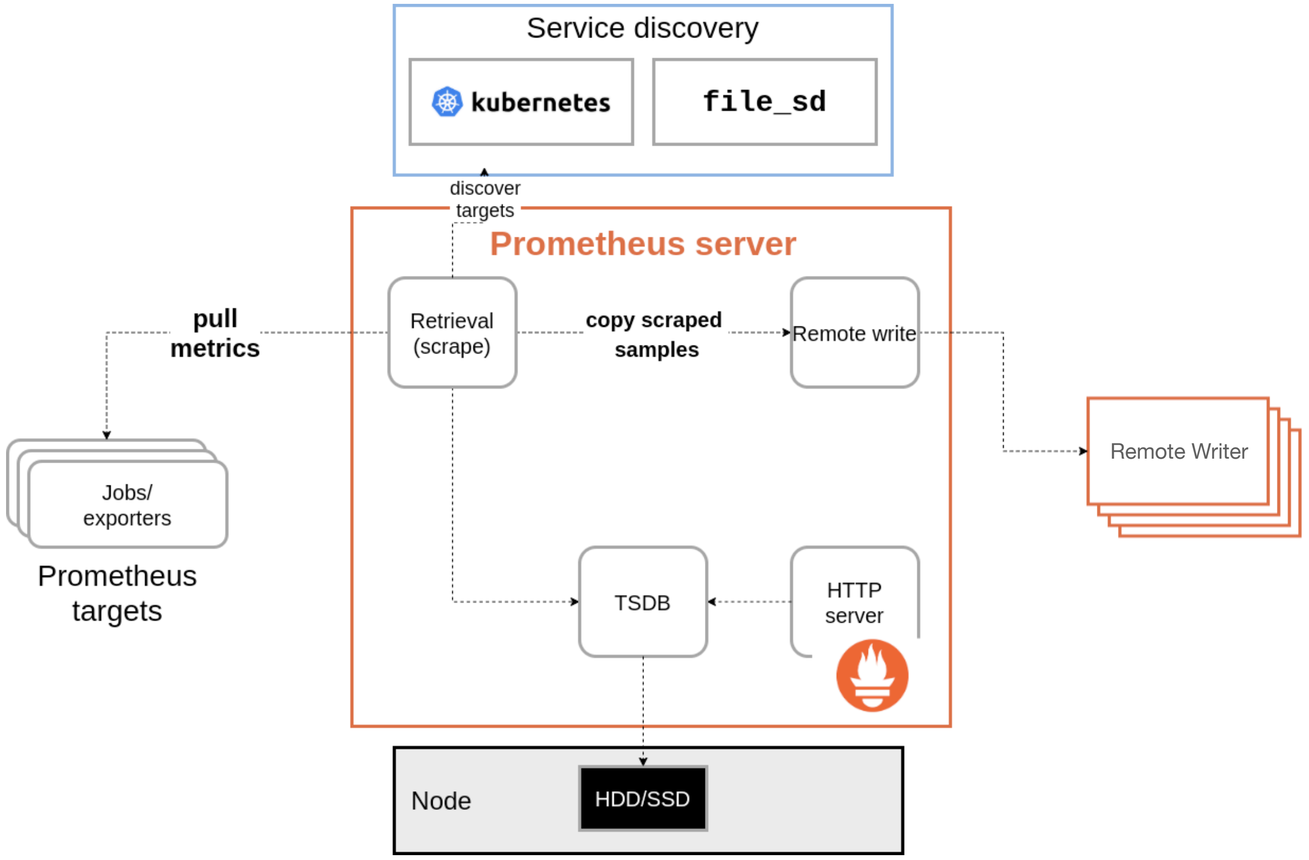
How to Integrate?

Remote Writer

Prometheus

Remote Writer

- Data from write-ahead log (WAL)
- Protocol Buffers for less payload
- Snappy compression



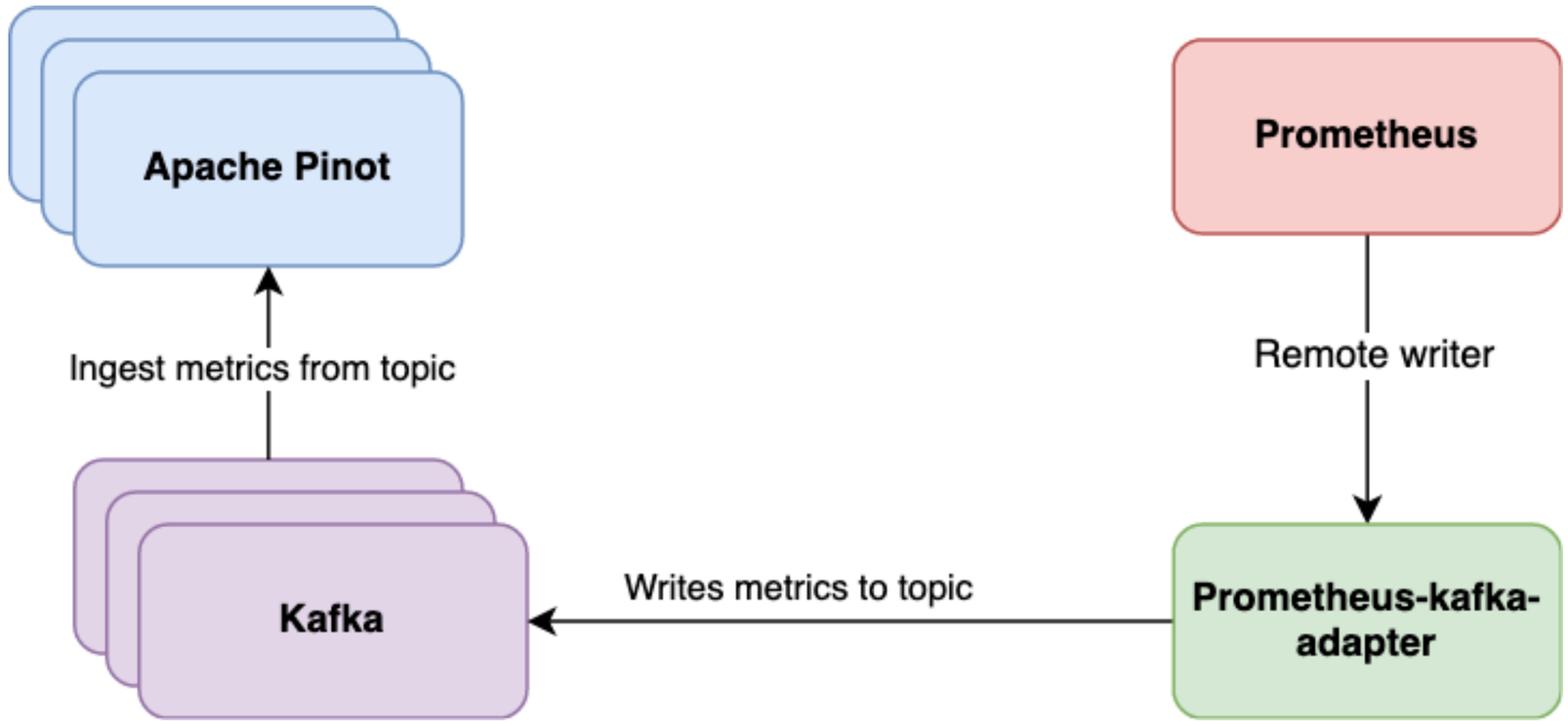
Prometheus with Apache Pinot

Prometheus Kafka Remote Writer - by Telefonica

- Receive metrics from prometheus
- Transform metrics in to JSON format
- Create Kafka topics - with rules capability
- Produce compressed data

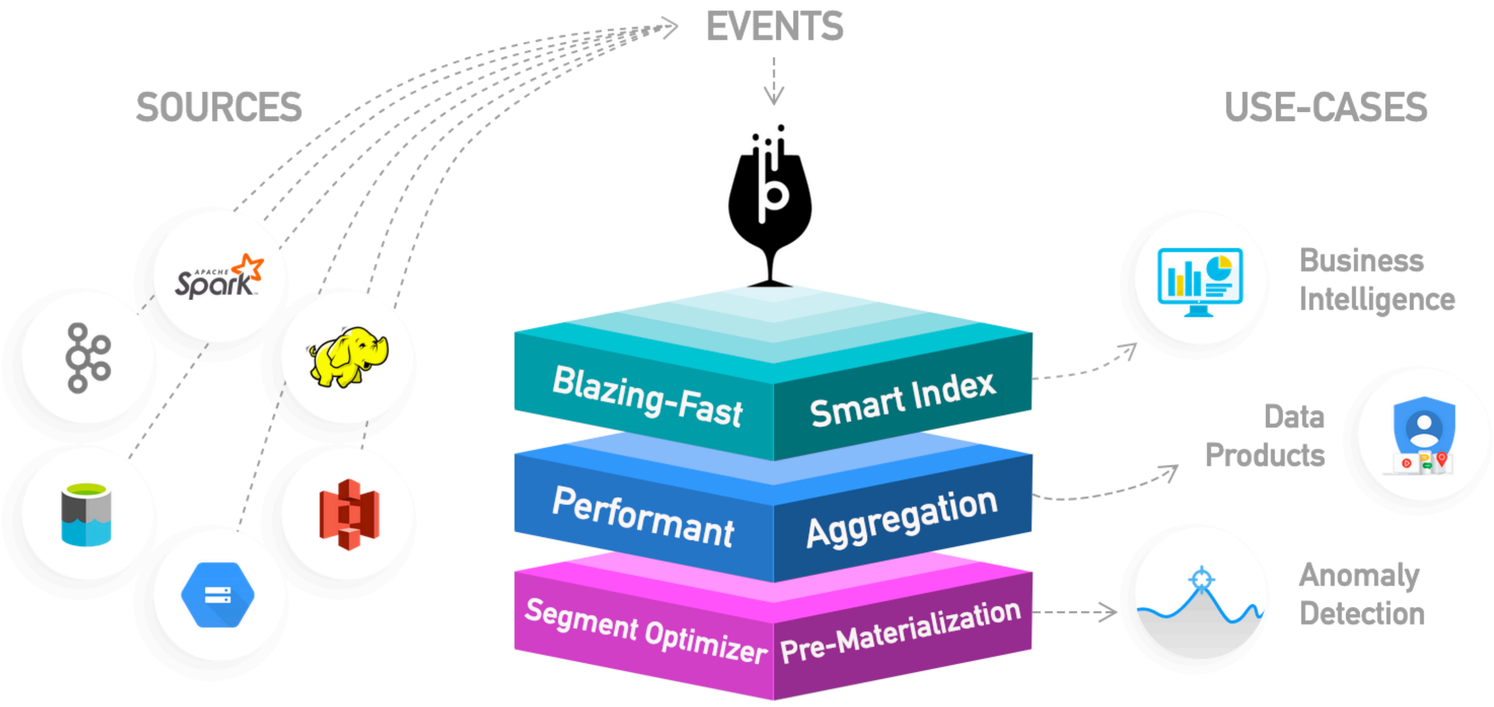
Prometheus with OLAP

Integrate with OLAP



Apache Pinot

And now…what can we do with it?



Hope you enjoyed!

Q&A