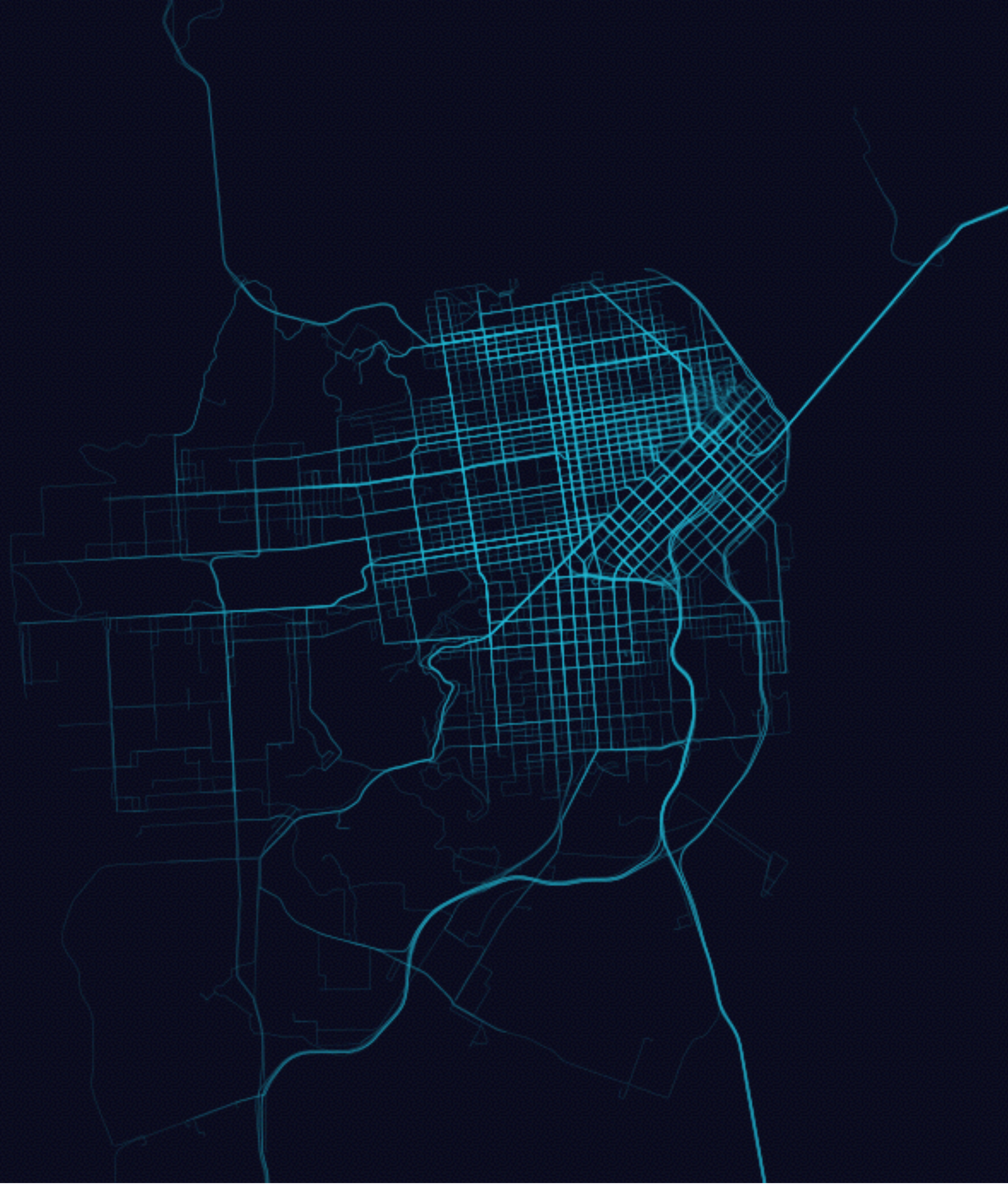
STREAM PROCESSING @ UBER

DANNY YUAN @ UBER



What is Uber

Transportation at your fingertips

05:00

Stream Data Allows Us To Feel The Pulse Of Cities

Marketplace Health

San Francisco

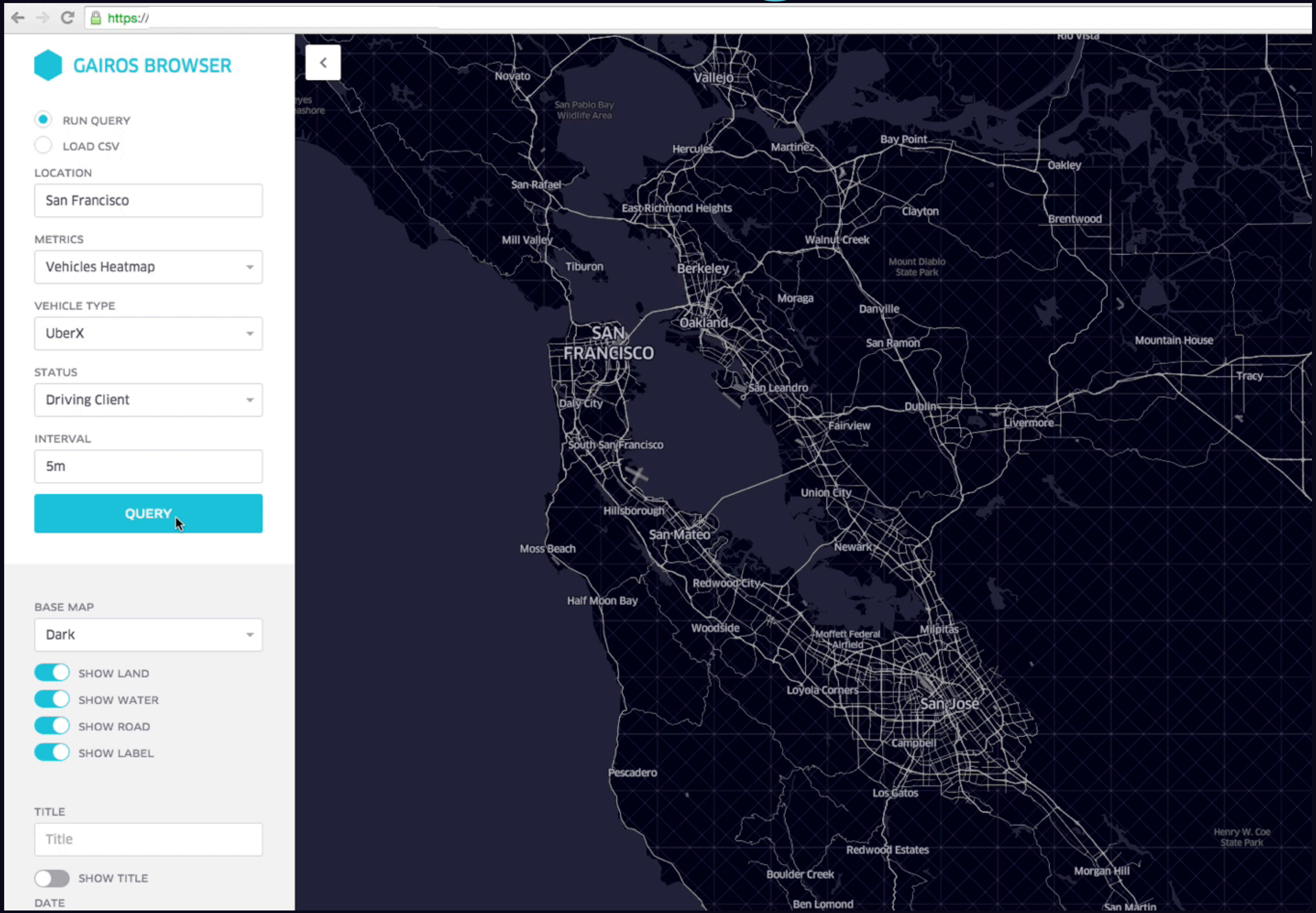
[redacted]

Trips Completed - Origin - [redacted]

San Francisco - All Vehicles

+
−
Golden Gate National Recreation Area
Alcatraz Island
Golden Gate Bridge
Text
Crissy Field
Ghirardelli Square
Palace of Fine Arts Theatre
The Walt Disney Family Museum
Francisco St
Lombard St
Filbert St
Green St
Vallejo St
Pacific Ave
Ferry Bldg Marketplace
Sacramento St
San Francisco Oakland Ba
The Fillmore
Ellis St
California St
San Francisco City Hall
Grove St
AT&T Park
Conservatory of Flowers
Haight St
32nd Ave
28th Ave
21st Ave
19th Ave
Golden Gate Park
California Academy of Sciences
Irving St
Judah St

What’s Going on Now



https://

GAIROS BROWSER

● RUN QUERY
○ LOAD CSV

LOCATION: San Francisco

METRICS: Vehicles Heatmap

VEHICLE TYPE: UberX

STATUS: Driving Client

INTERVAL: 5m

QUERY

BASE MAP: Dark

SHOW LAND

SHOW WATER

SHOW ROAD

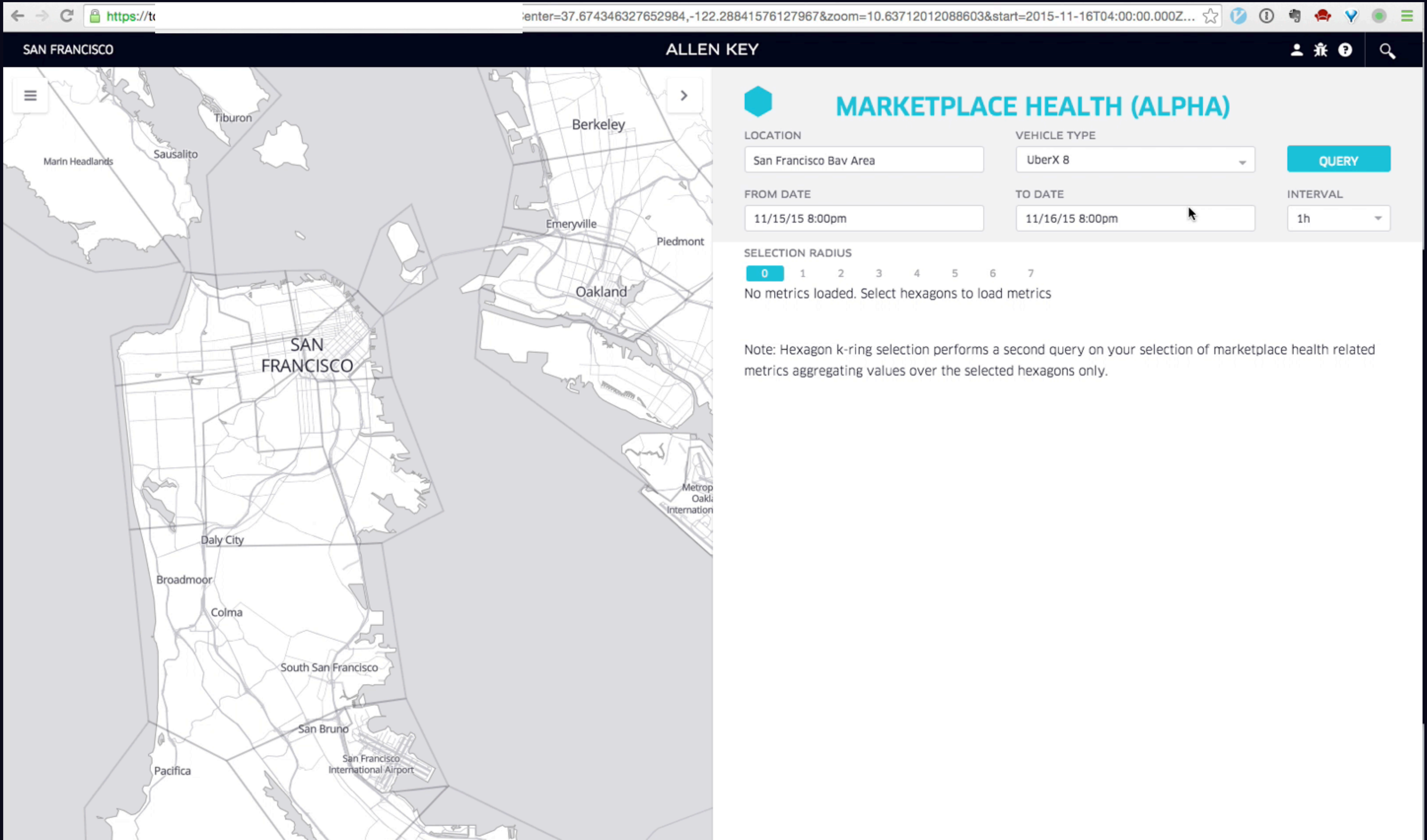
SHOW LABEL

TITLE: Title

SHOW TITLE

DATE

What’s Happened?



https://t enter=37.674346327652984,-122.28841576127967&zoom=10.63712012088603&start=2015-11-16T04:00:00.000Z...

SAN FRANCISCO

ALLEN KEY

MARKETPLACE HEALTH (ALPHA)

LOCATION: San Francisco Bav Area

VEHICLE TYPE: UberX 8

QUERY

FROM DATE: 11/15/15 8:00pm

TO DATE: 11/16/15 8:00pm

INTERVAL: 1h

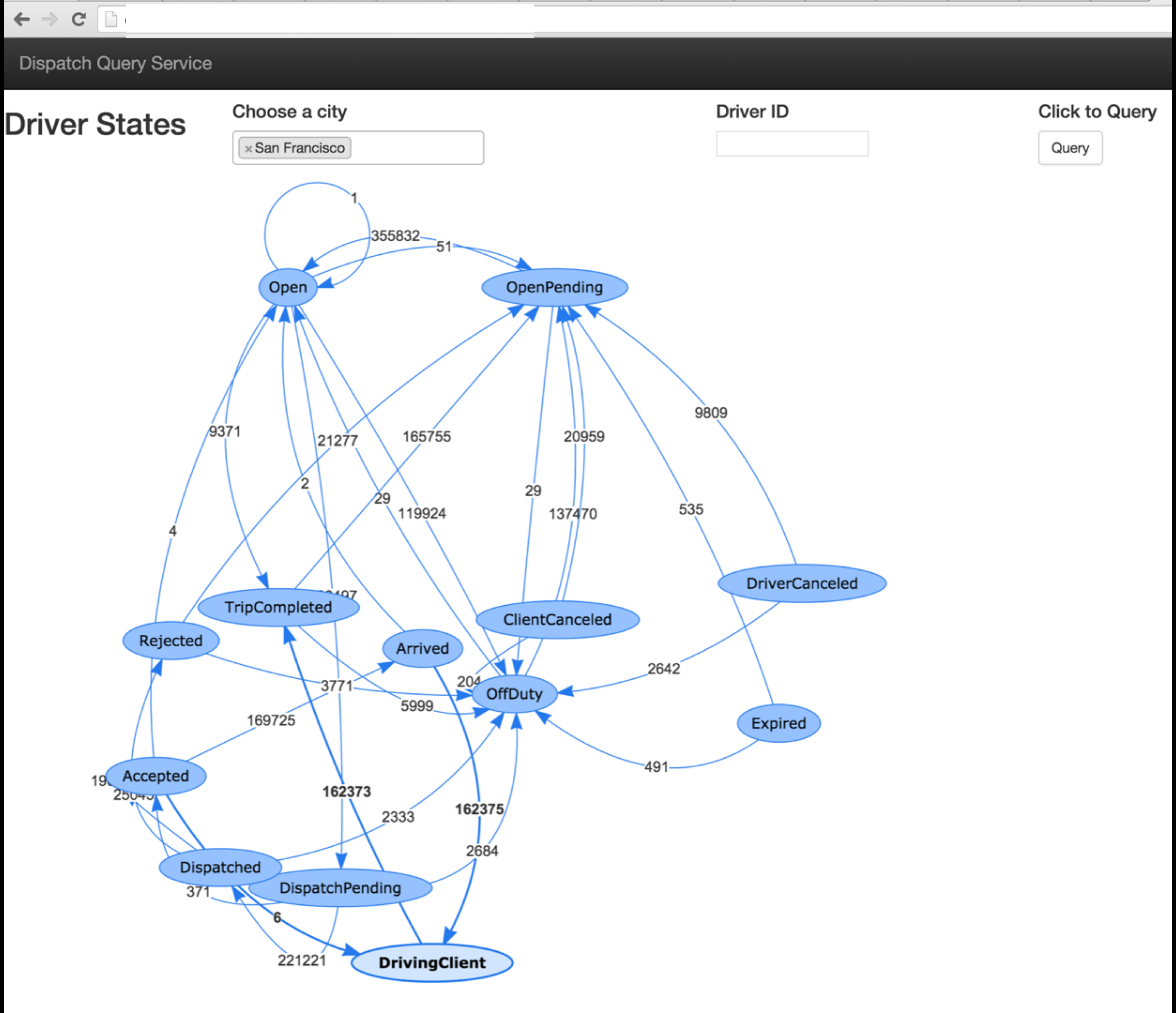
SELECTION RADIUS: [✓ 0] 1 2 3 4 5 6 7

No metrics loaded. Select hexagons to load metrics

Note: Hexagon k-ring selection performs a second query on your selection of marketplace health related metrics aggregating values over the selected hexagons only.

Status Tracking

OffDuty
OpenPending
Open
DispatchPending
Dispatched
Accepted
Arrived
DrivingClient
TripCompleted
OpenPending_from_TripCompleted
OffDuty_from_TripCompleted
Dispatched_from_Arrived
Open_from_Arrived
ClientCanceled
OpenPending_from_ClientCanceled
OffDuty_from_ClientCanceled
OffDuty_from_DriverCanceled
Dispatched_from_Accepted
DriverCanceled
Expired
Rejected
OffDuty_from_OpenPending
OpenPending_from_DriverCanceled
OpenPending_from_Rejected
Open_from_Dispatched
OffDuty_from_Open
OpenPending_from_Expired
Open_from_Accepted
OffDuty_from_DispatchPending
OffDuty_from_Dispatched
OffDuty_from_Expired
OffDuty_from_Rejected



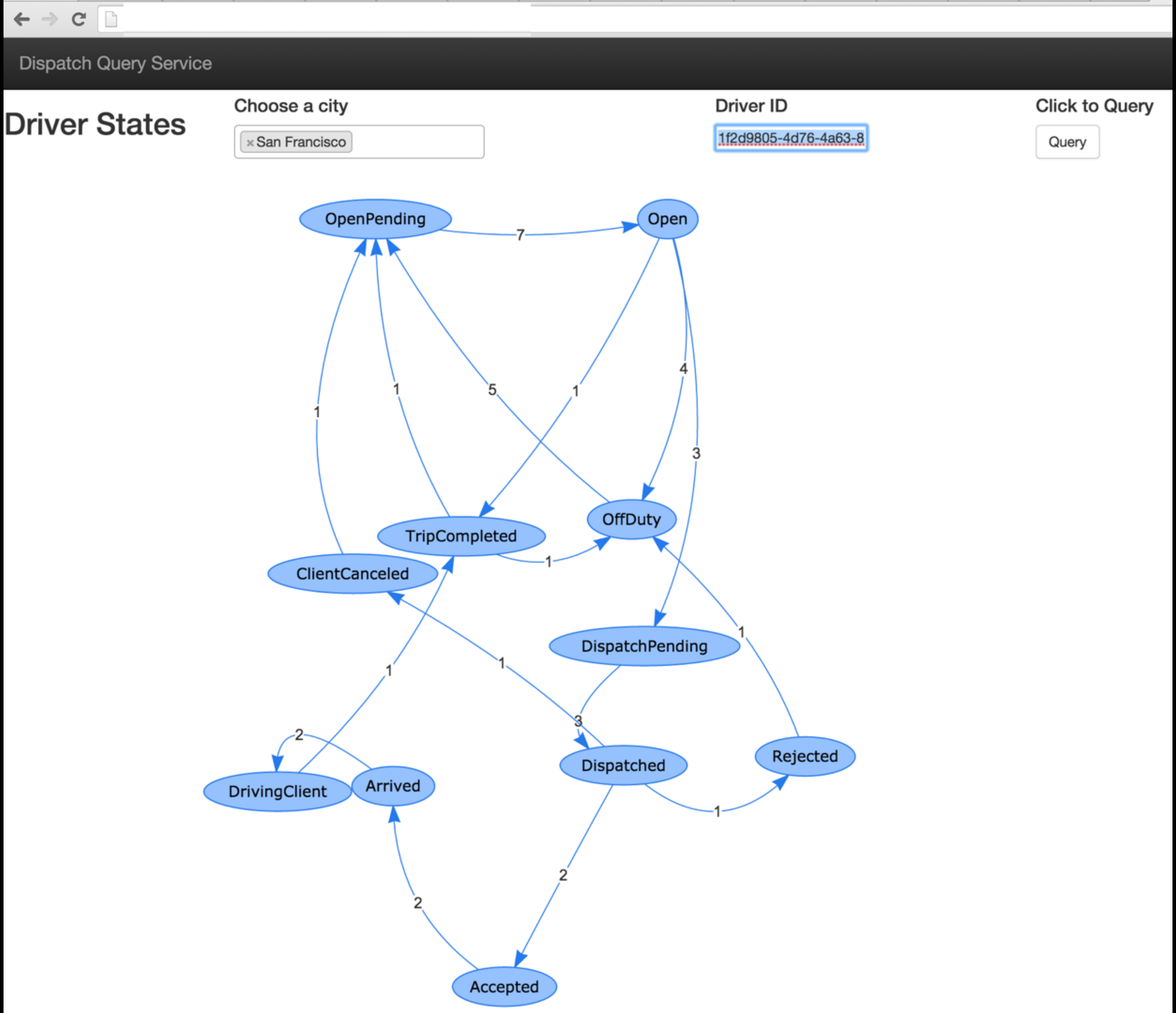
Dispatch Query Service

Driver States

Choose a city: × San Francisco

Driver ID:

Click to Query: Query



Dispatch Query Service

Driver States

Choose a city

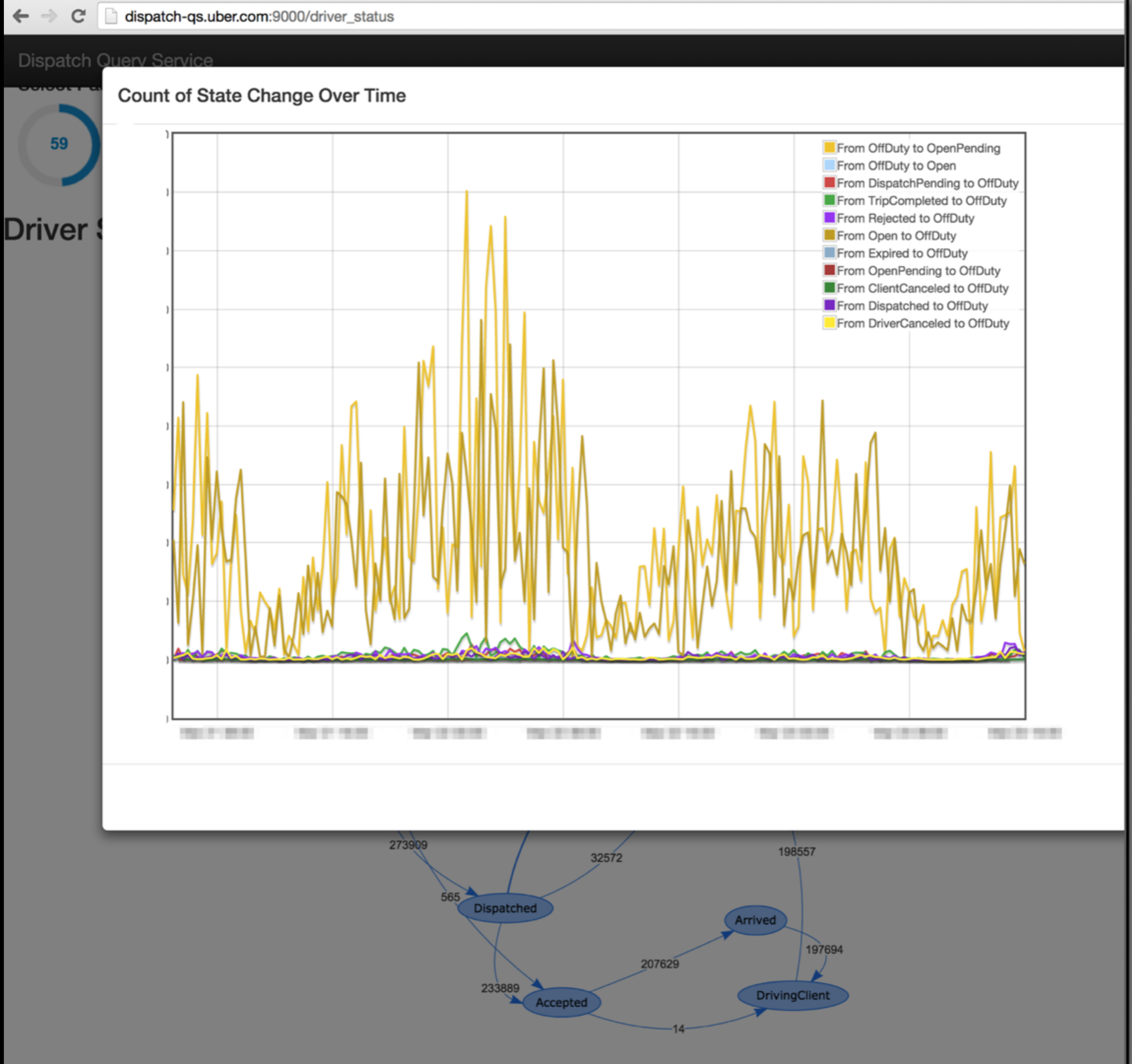
× San Francisco

Driver ID

1f2d9805-4d76-4a63-8

Click to Query

Query



dispatch-qs.uber.com:9000/driver_status

Dispatch Query Service

Count of State Change Over Time

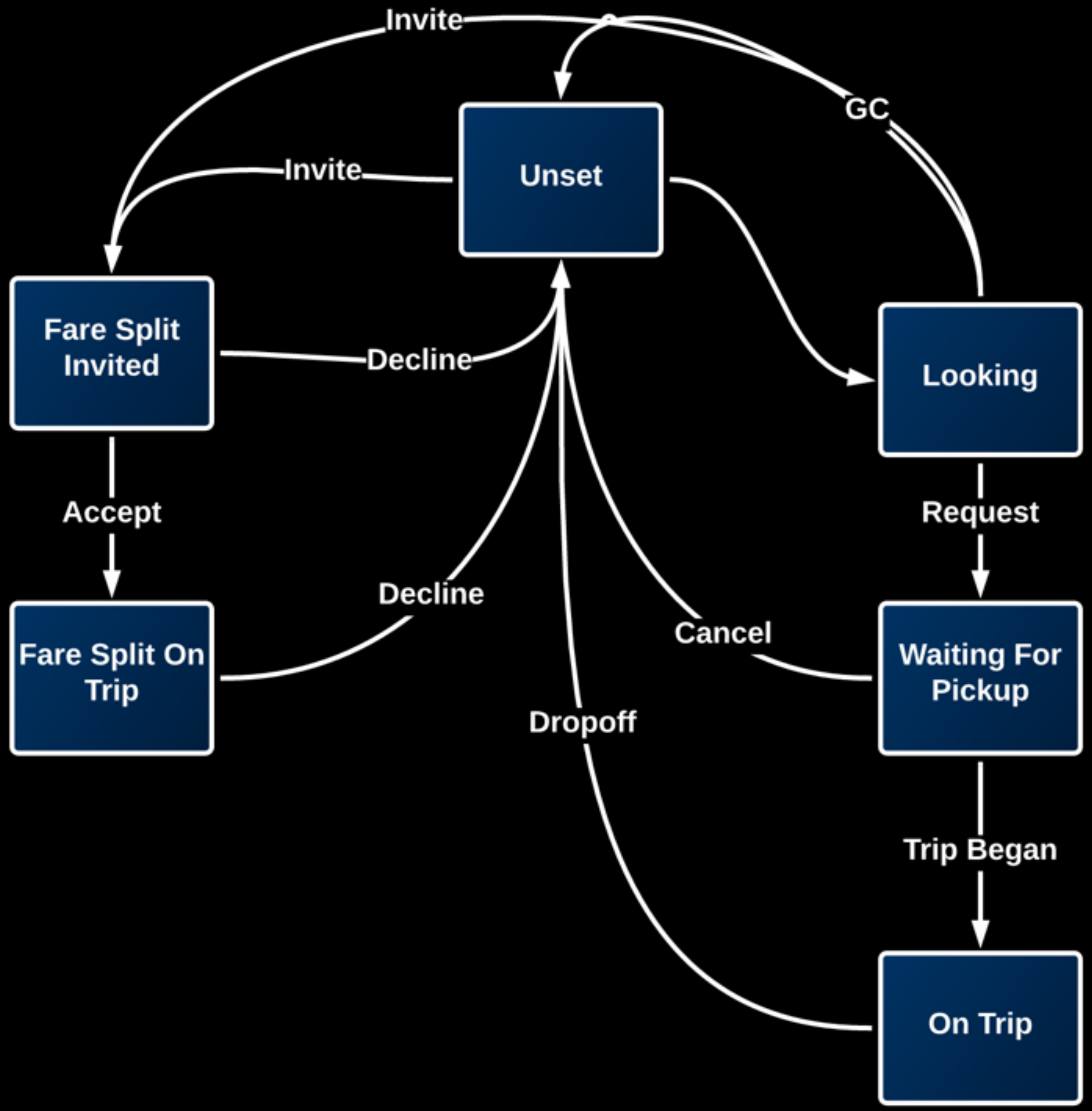
59
Driver S
From OffDuty to OpenPending
From OffDuty to Open
From DispatchPending to OffDuty
From TripCompleted to OffDuty
From Rejected to OffDuty
From Open to OffDuty
From Expired to OffDuty
From OpenPending to OffDuty
From ClientCanceled to OffDuty
From Dispatched to OffDuty
From DriverCanceled to OffDuty
273909
32572
198557
565
Dispatched
Arrived
197694
207629
233889
Accepted
DrivingClient
14

[redacted]

A Little Background

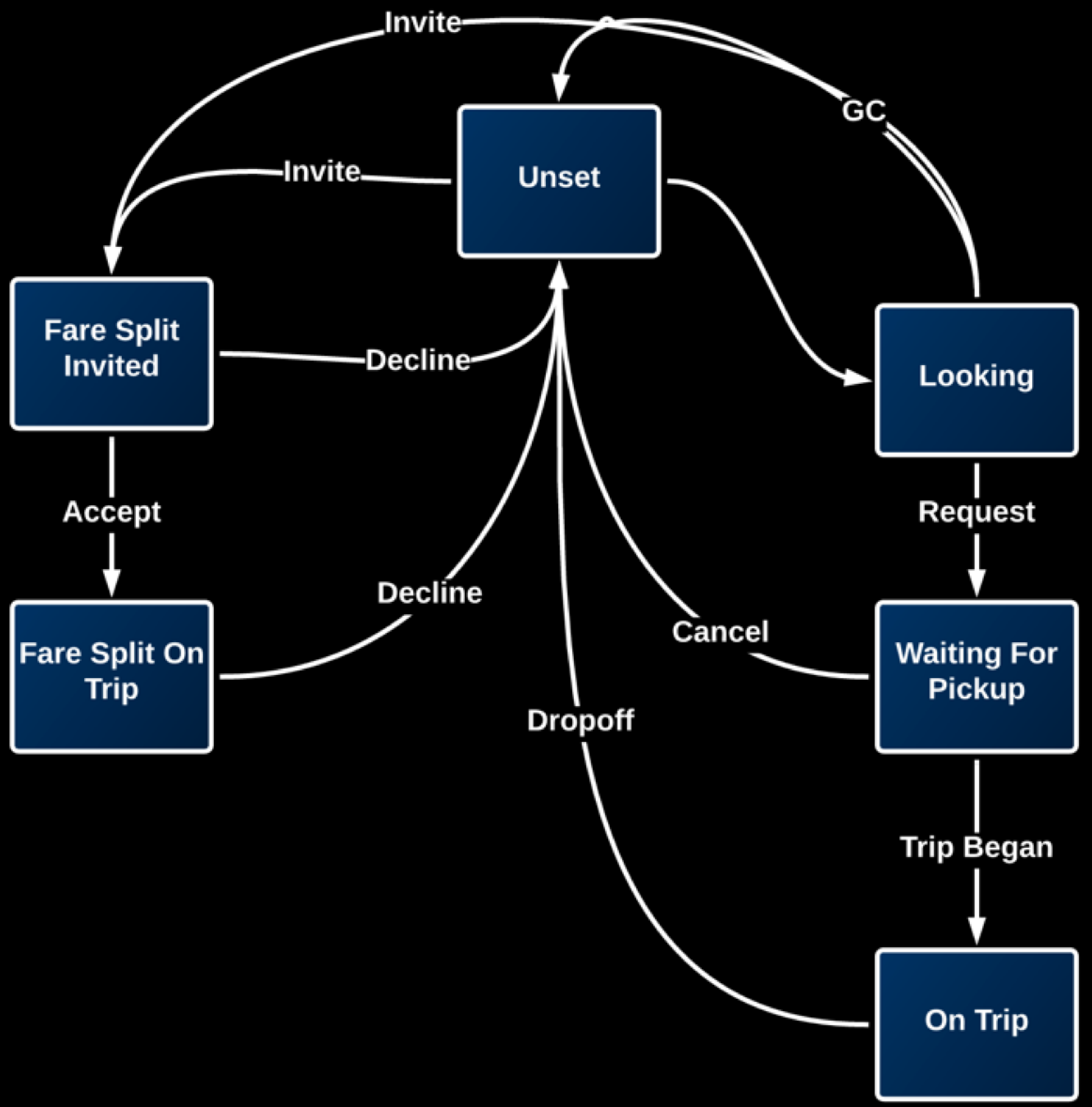
Uber’s Platform Is a Distributed State Machine

Rider States

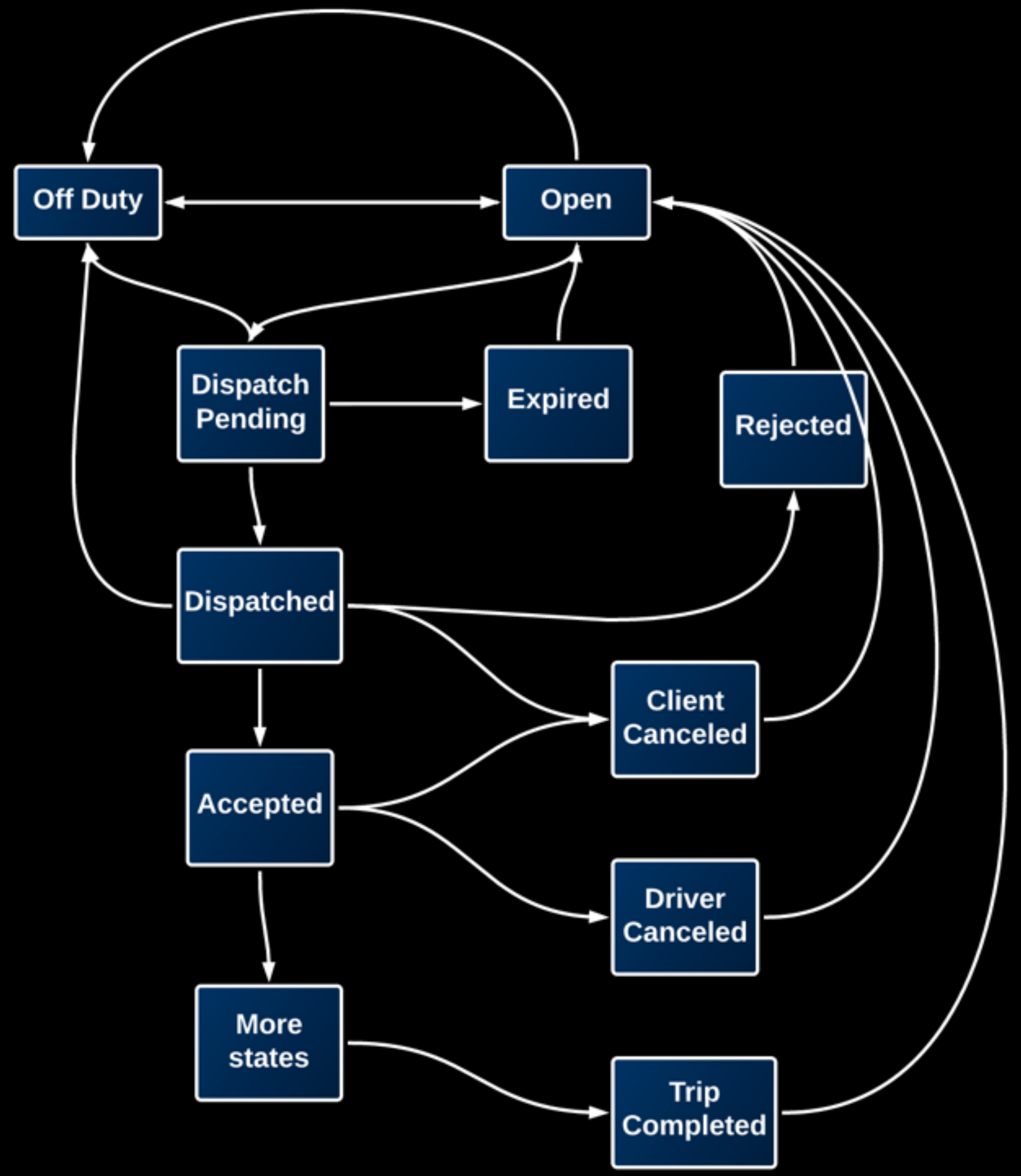


Uber’s Platform Is a Distributed State Machine

Rider States



Driver States



Applications can’t do everything

Instead, Applications Emit Events

Events Should Be Available In Seconds

Events Should Rarely Get Lost

Events Should Be Cheap And Scalable

Where are the challenges?

Many Dimensions

Dozens of fields per event

Granular Data

San Francisco

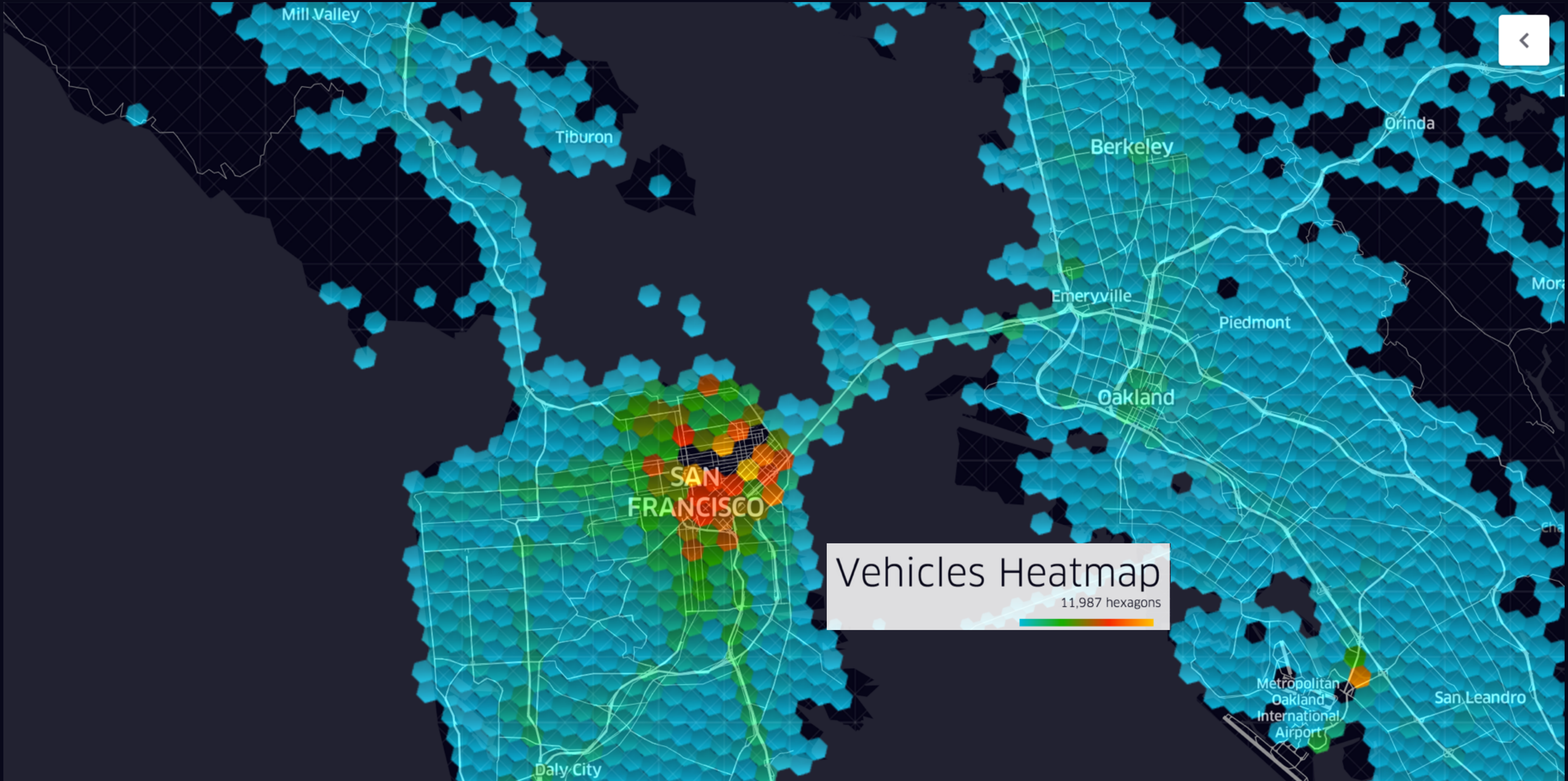
[redacted]

Trips Completed - Origin - [redacted]

San Francisco - All Vehicles

+
−
Golden Gate National Recreation Area
Alcatraz Island
Golden Gate Bridge
Crissy Field
The Walt Disney Family Museum
Text
Ghirardelli Square
Palace of Fine Arts Theatre
Francisco St
Lombard St
Filbert St
Green St
Vallejo St
Pacific Ave
Sacramento St
Ferry Bldg Marketplace
The Fillmore
Ellis St
San Francisco City Hall
Grove St
Haight St
AT&T Park
California St
Conservatory of Flowers
Golden Gate Park
California Academy of Sciences
Irving St
Judah St
San Francisco Oakland Bay

Granular Data



Granular Data

Over 10,000 hexagons in the city

Granular Data

7 vehicle types

Granular Data

1440 minutes in a day

Granular Data

13 driver states

Granular Data

300 cities

Granular Data

1 day of data: 300 x 10,000 x 7 x 1440 x 13 = 393 billion possible combinations

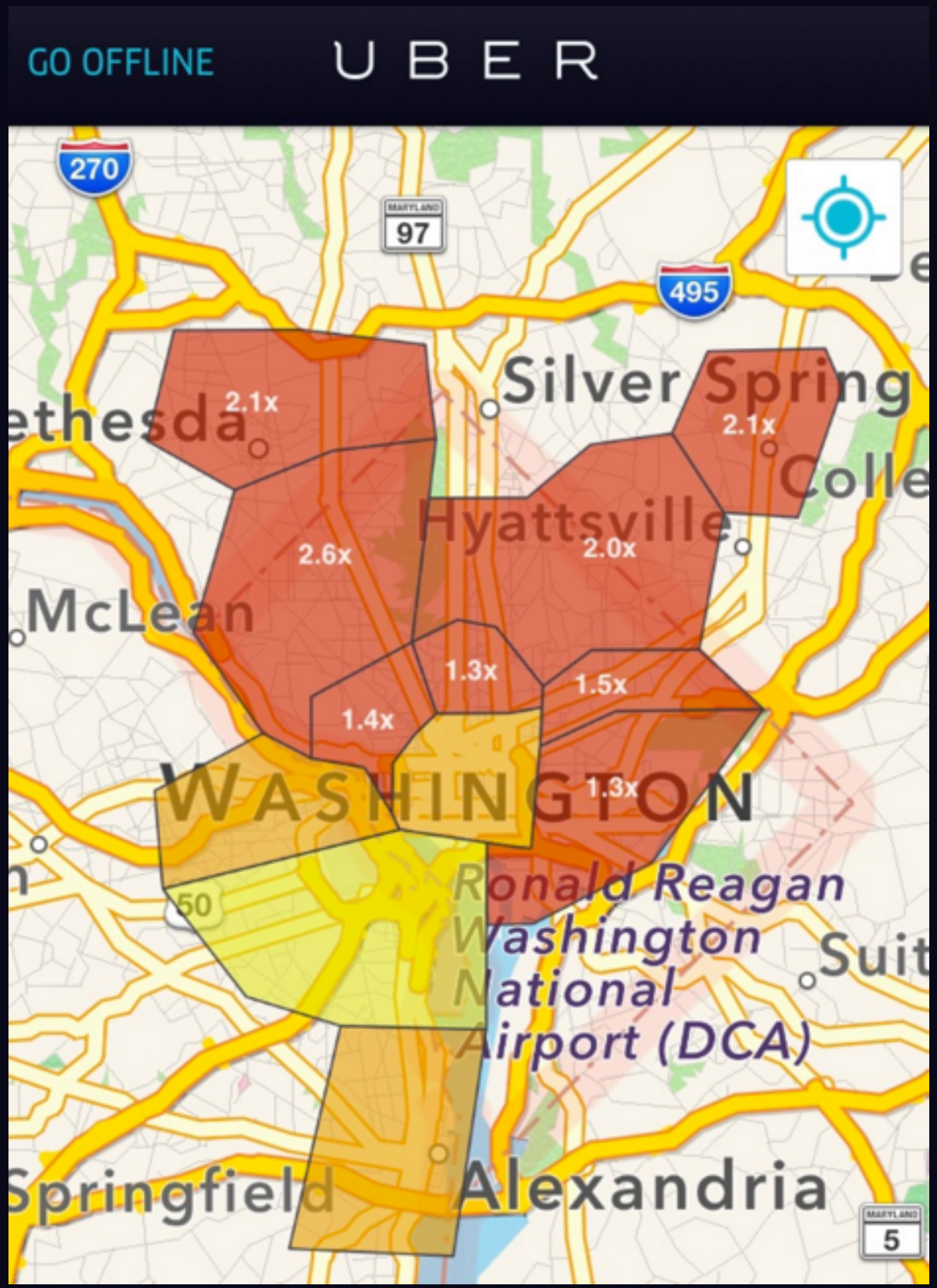
Unknown Query Patterns

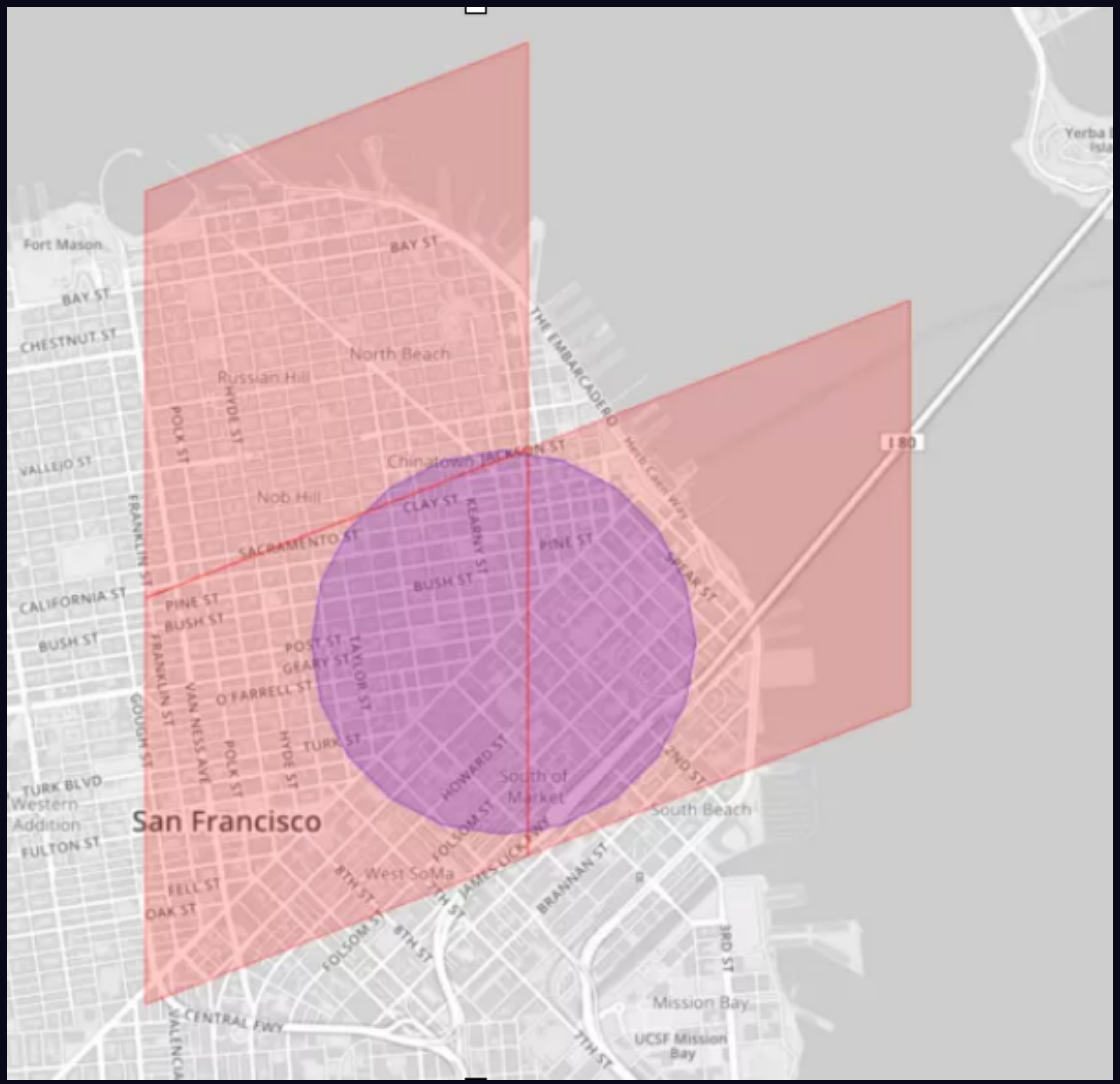
Any combination of dimensions

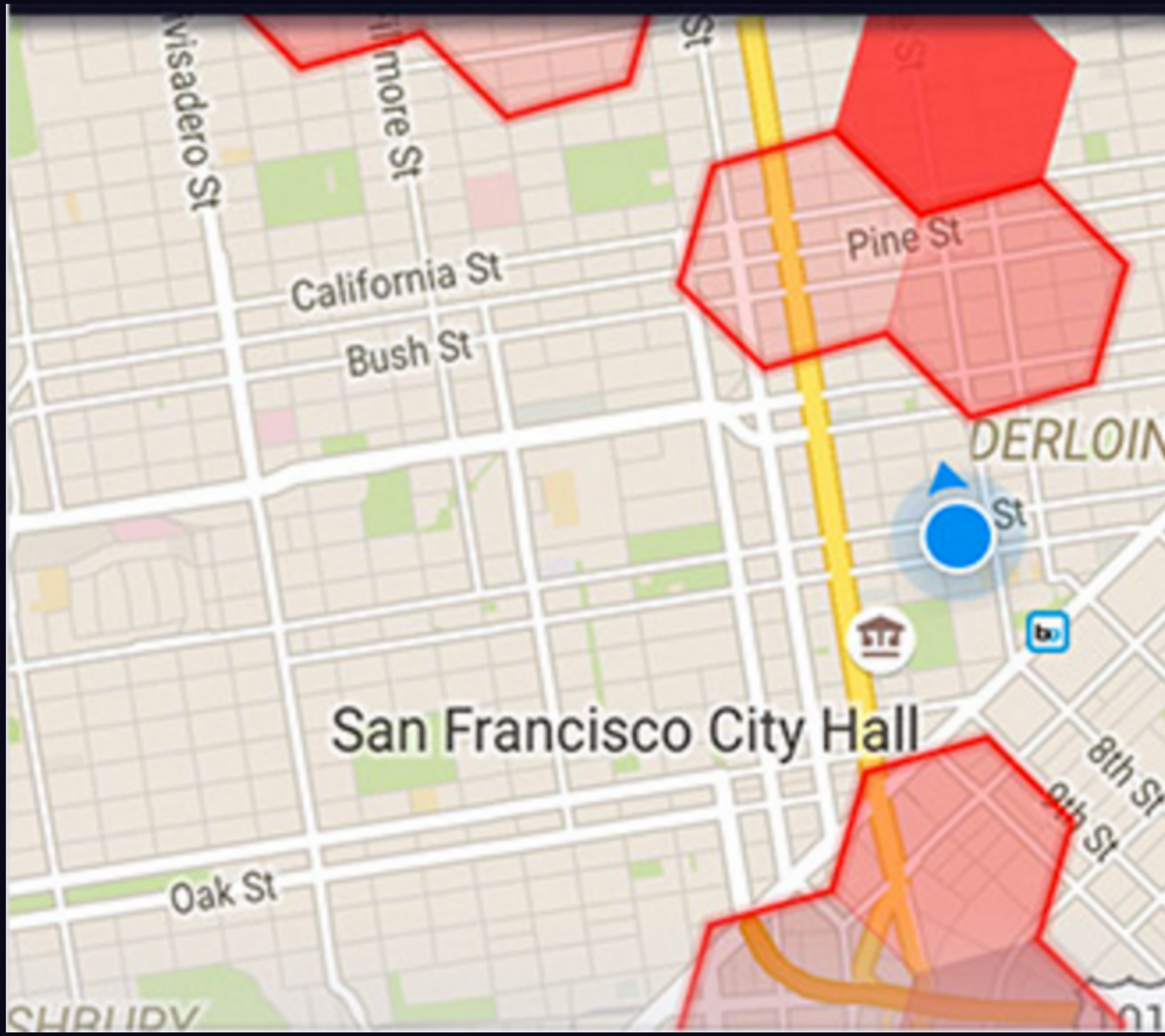
Variety of Aggregations

- Heatmap
- Top N
- Histogram
- count(), avg(), sum(), percent(), geo

Different Geo Aggregation







Large Data Volume

- Hundreds of thousands of events per second, or billions of events per day
- At least dozens of fields in each event

Tight Schedule

Key: Generalization

Data Type

- Dimensional Temporal Spatial Data

Dimension	Value
state	driver_arrived
vehicle type	uber X
timestamp	13244323342
lattitude	12.23
longitude	30.00

Data Query

- OLAP on single-table temporal-spatial data

```
SELECT <agg functions>, <dimensions>
FROM <data_source>
WHERE <boolean filter>
GROUP BY <dimensions>
HAVING <boolean filter>
ORDER BY <sorting criterial>
LIMIT <n>
DO <post aggregation>
```

Finding the Right Storage System

Minimum Requirements

- OLAP with geospatial and time series support
- Support large amount of data
- Sub-second response time
- Query of raw data

It can’t be a KV store

Challenges to KV Store

Pre-computing all keys is $O(2^n)$ for both space and time

It can’t be a relational database

Challenges to Relational DB

- Managing multiple indices is painful
- Scanning is not fast enough

A System That Supports

- Fast scan
- Arbitrary boolean queries
- Raw data
- Wide range of aggregations

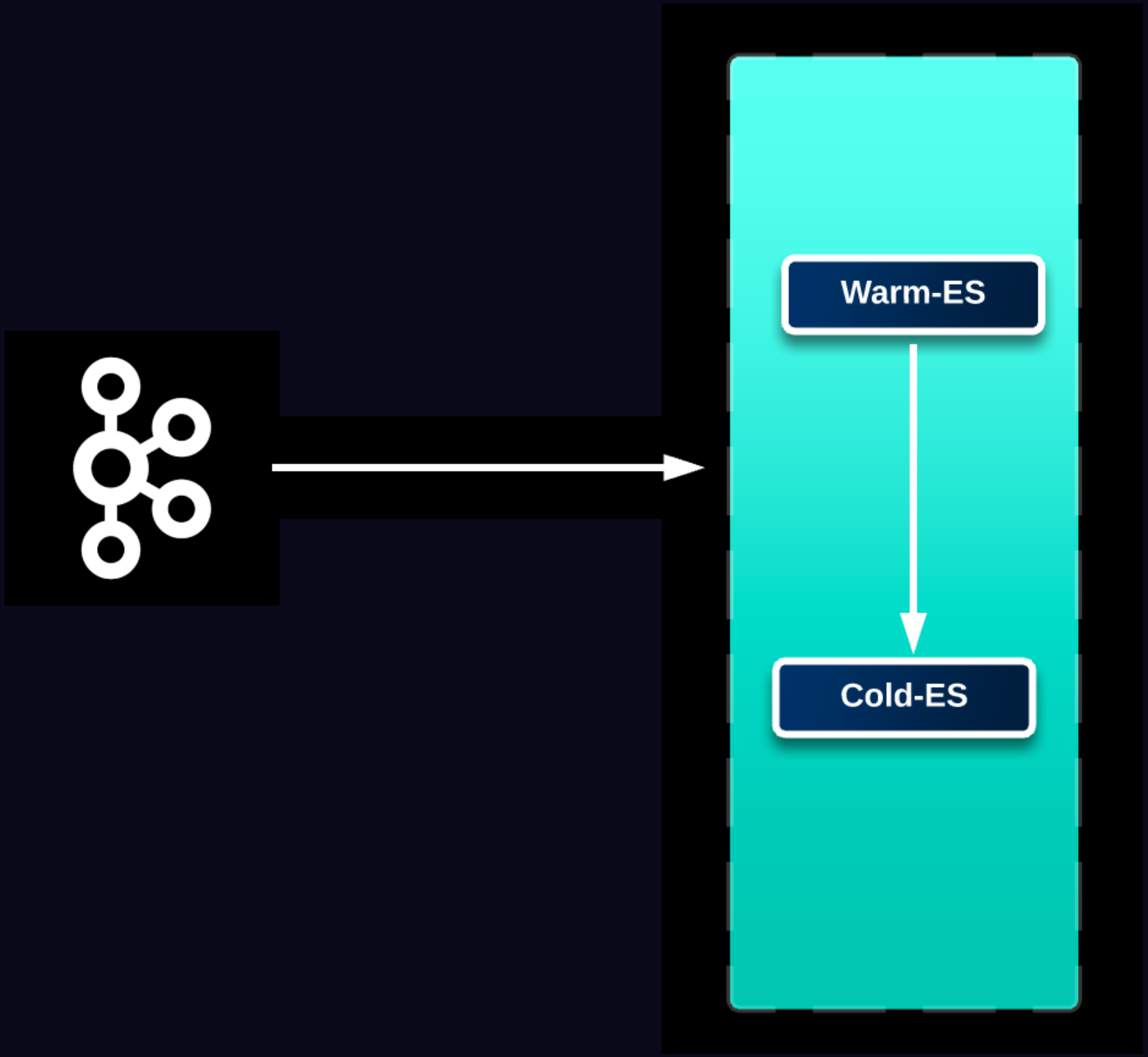
Elasticsearch

Highly Efficient Inverted-Index For Boolean Query

Built-in Distributed Query

Fast Scan with Flexible Aggregations

Storage

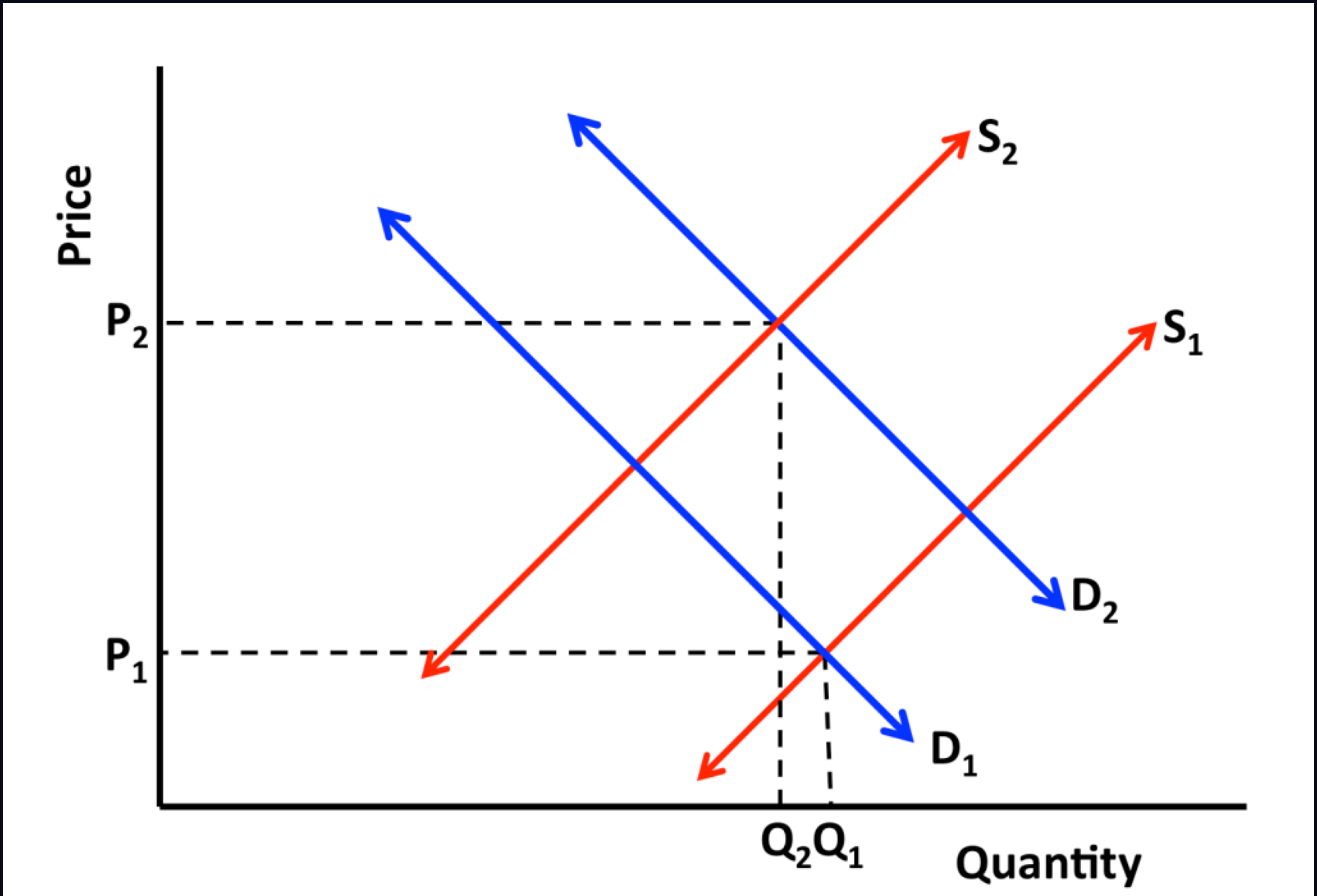


Are We Done?

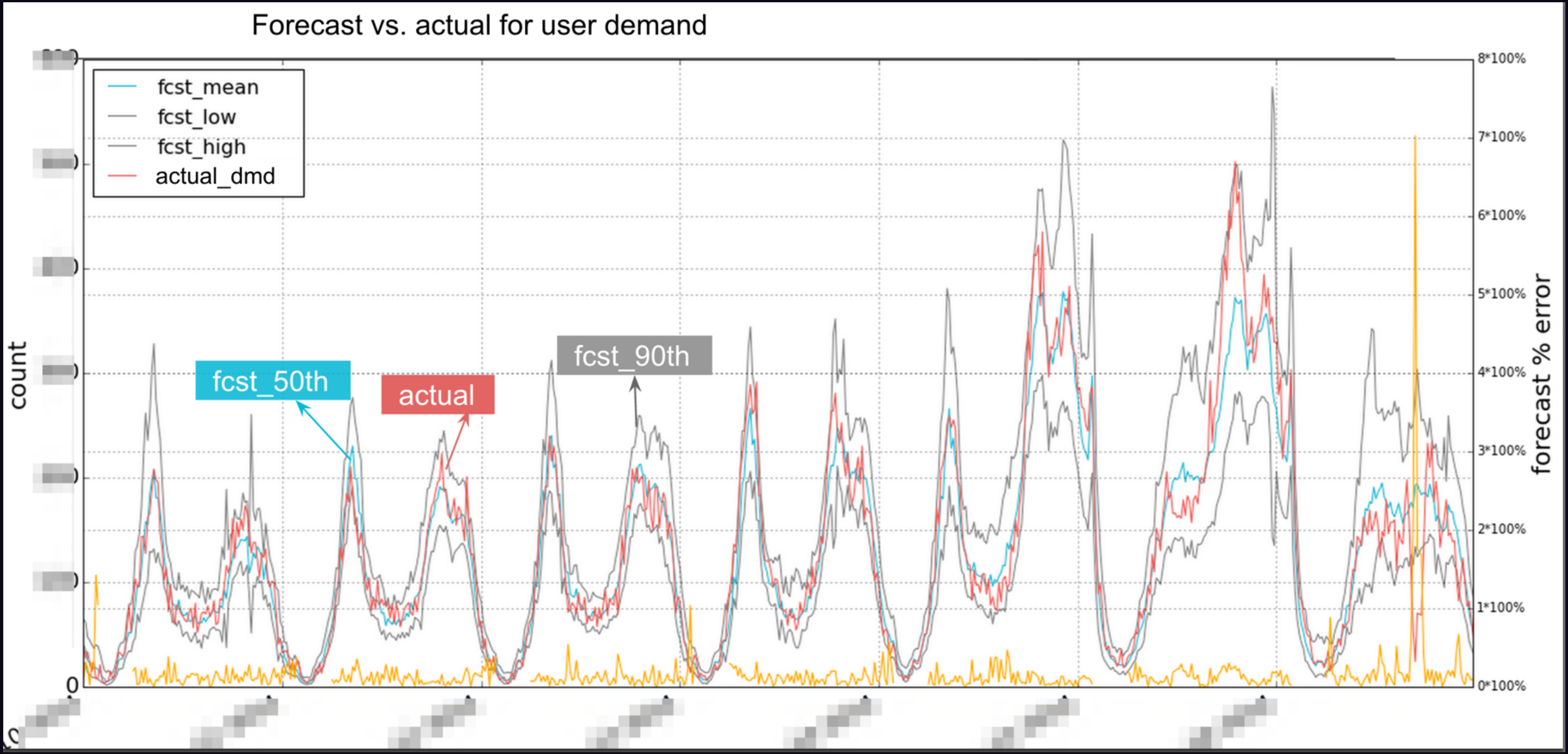
Transformation

e.g. (Lat, Long) -> (zipcode, hexagon)

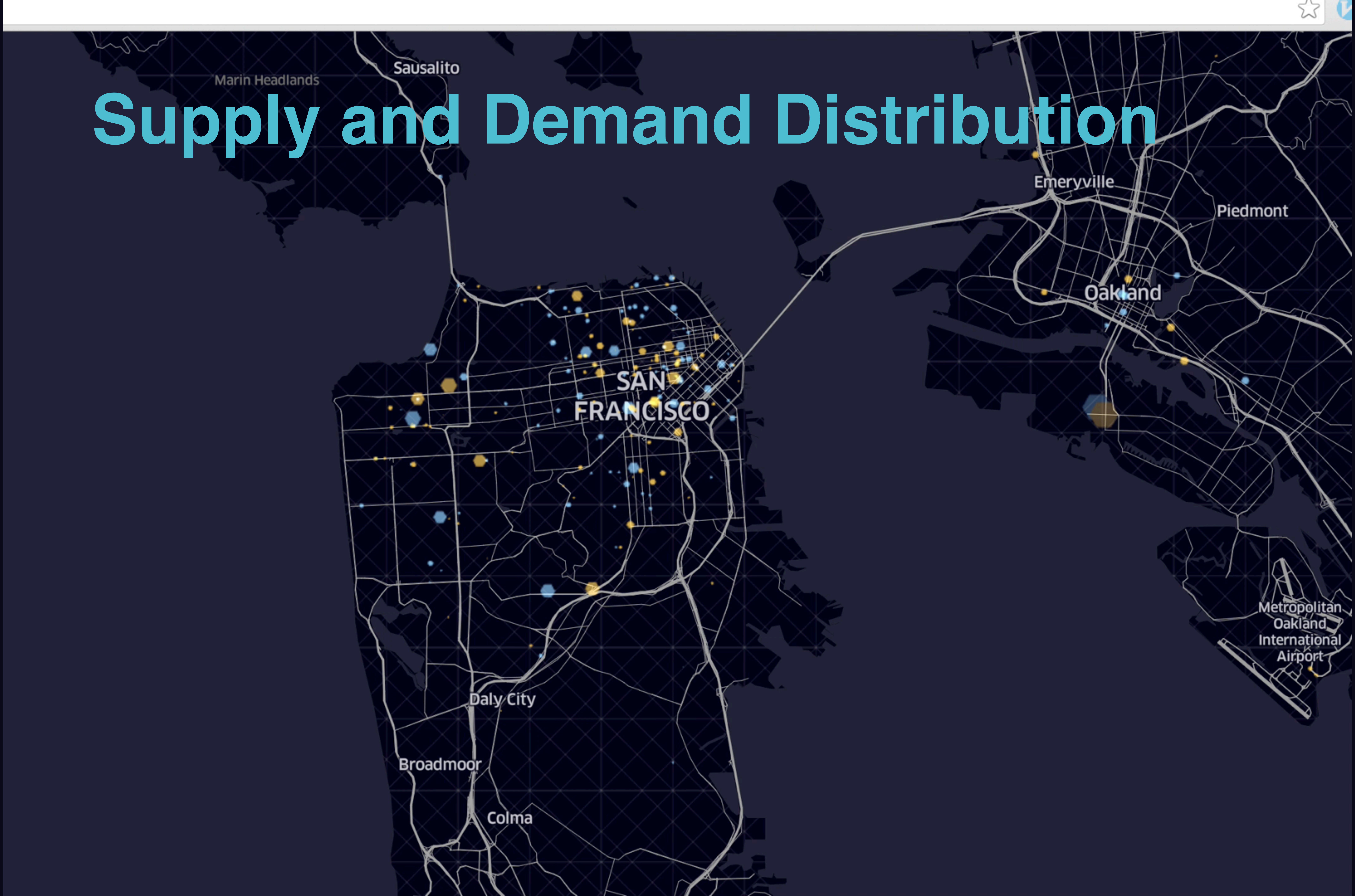
Dynamic Pricing



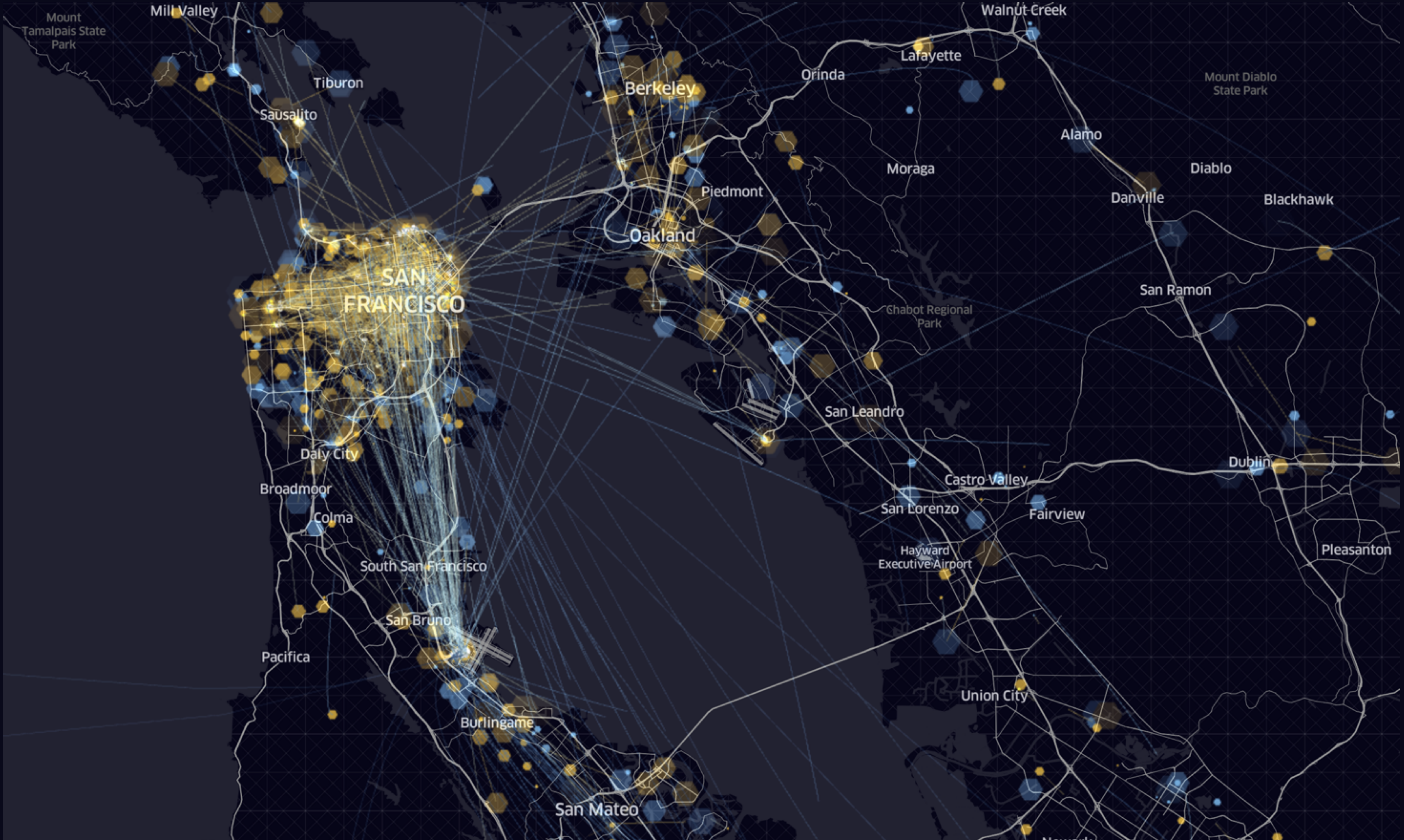
Trend Prediction



Supply and Demand Distribution



Technically Speaking: Clustering & Pr(D, S, E)



New Use Cases —> New Requirements

Pre-aggregation

Joining Multiple Streams

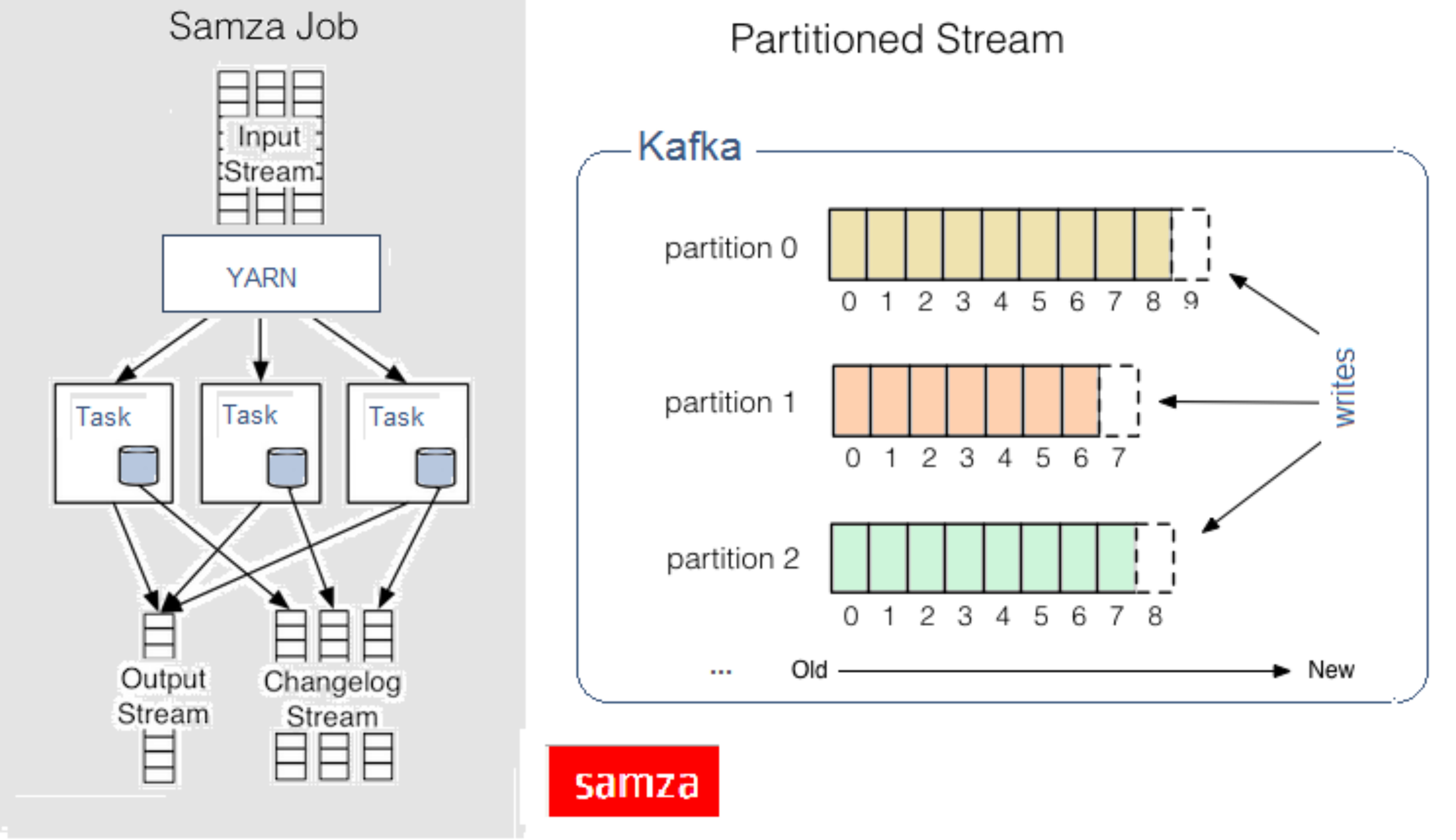
Sessionization

Multi-Staged Processing

State Management

Apache Samza

Why Apache Samza?



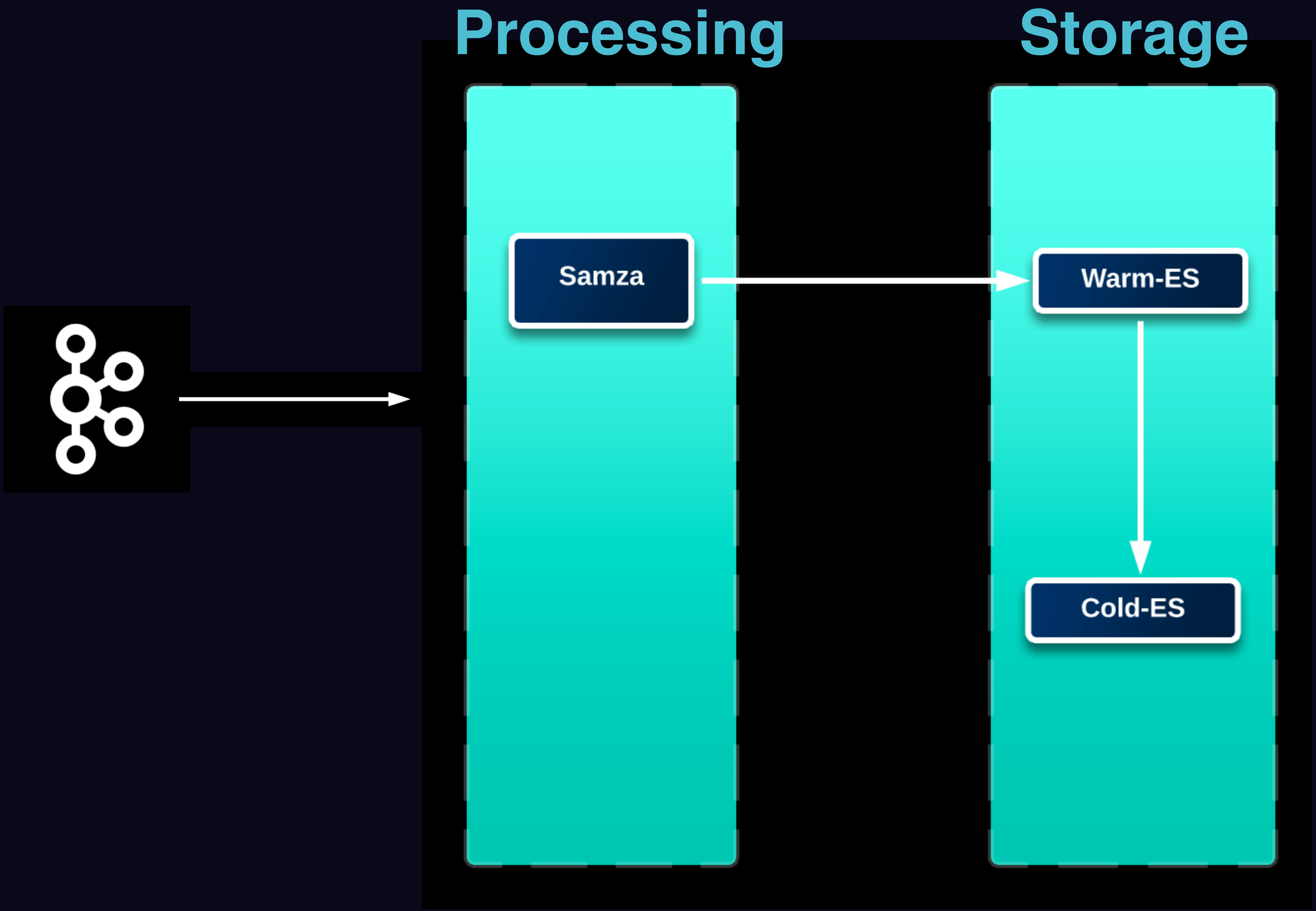
DAG on Kafka

Excellent Integration with Kafka

Excellent Integration with Kafka

Built-in Checkpointing

Built-in State Management



Processing

Storage

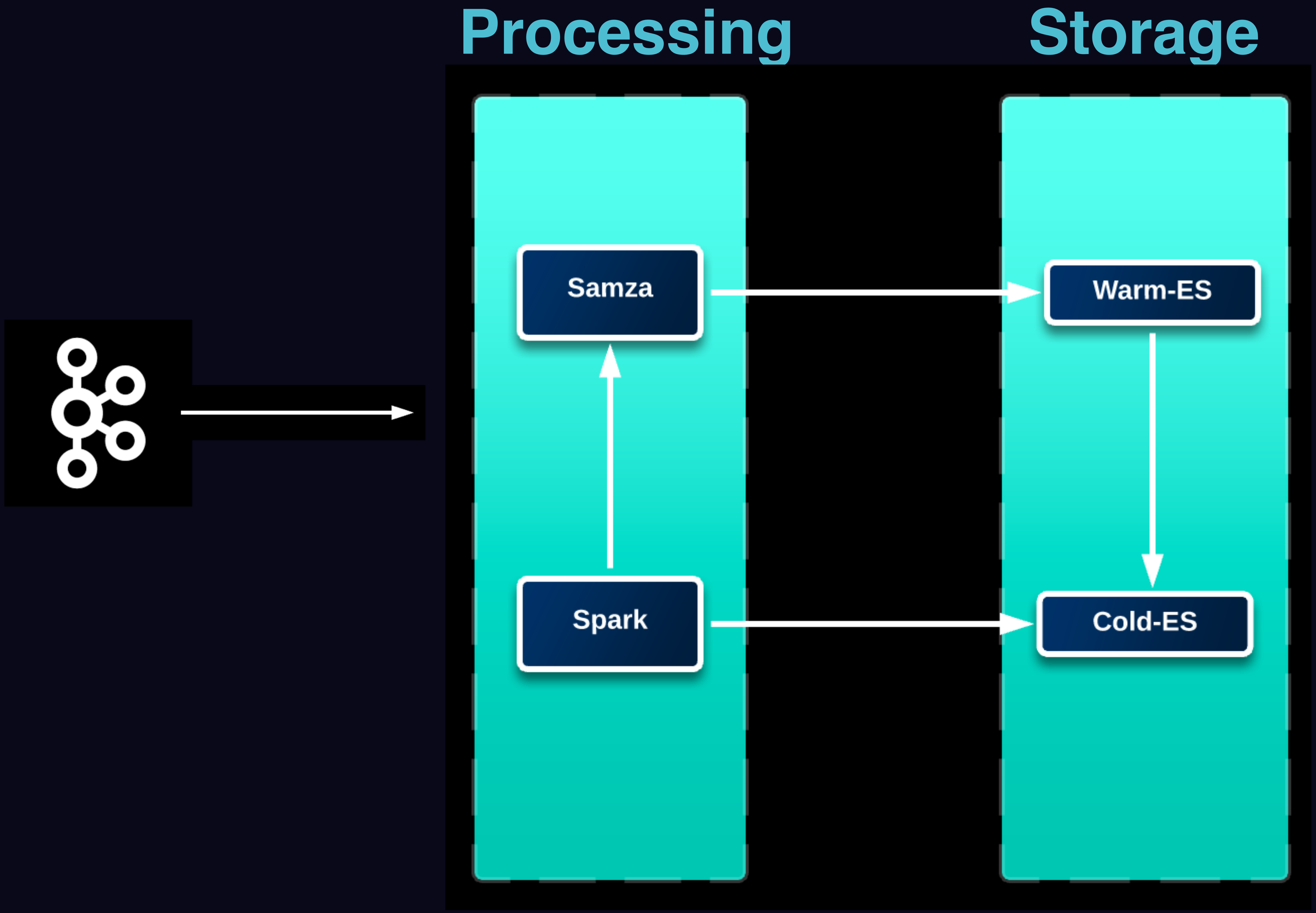
Samza

Warm-ES

Cold-ES

What If Storage Is Down?

What If Processing Takes Long?


Processing
Storage
Samza
Warm-ES
Spark
Cold-ES

Are We Done?

SELECTION RADIUS

0 1 [✓ 2] 3 4 5 6 7

Presidio
PINE ST
SAN FRANCISCO
Golden Gate Park
18TH ST
GREAT HWY
Daly City

SELECTION RADIUS

0 1 [✓ 2] 3 4 5 6 7

Presidio
PINE ST
SAN FRANCISCO
n Ga e Par
18TH ST
GREAT HWY
Daly City

Post Processing

Results Transformation and Smoothing

Scale of Post Processing

10,000 hexagons in a city

Scale of Post Processing

331 neighboring hexagons to look at

Scale of Post Processing

331 x 10,000 = 3.1 Million Hexagons to Process for a Single Query

Scale of Post Processing

99%-ile Processing Time: 70ms

Post Processing

- Each processor is a pure function
- Processors can be composed by combinators

Post Processing

- Highly parallelized execution
- Pipelining

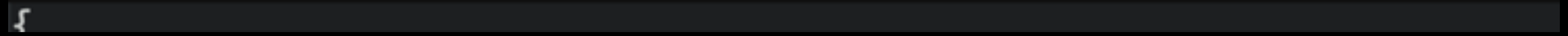
Post Processing

- Each processor is a pure function
- Processors can be composed by combinators
- Highly parallelized execution

Practical Considerations

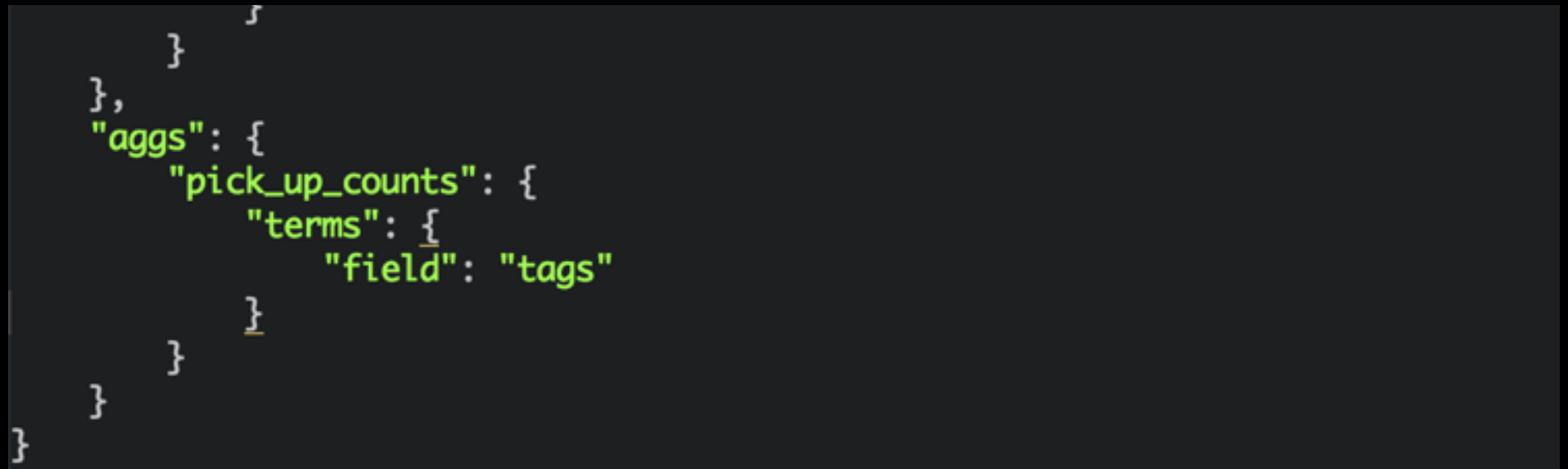
Data Discovery

Elasticsearch Query Can Be Complex



```
{
```

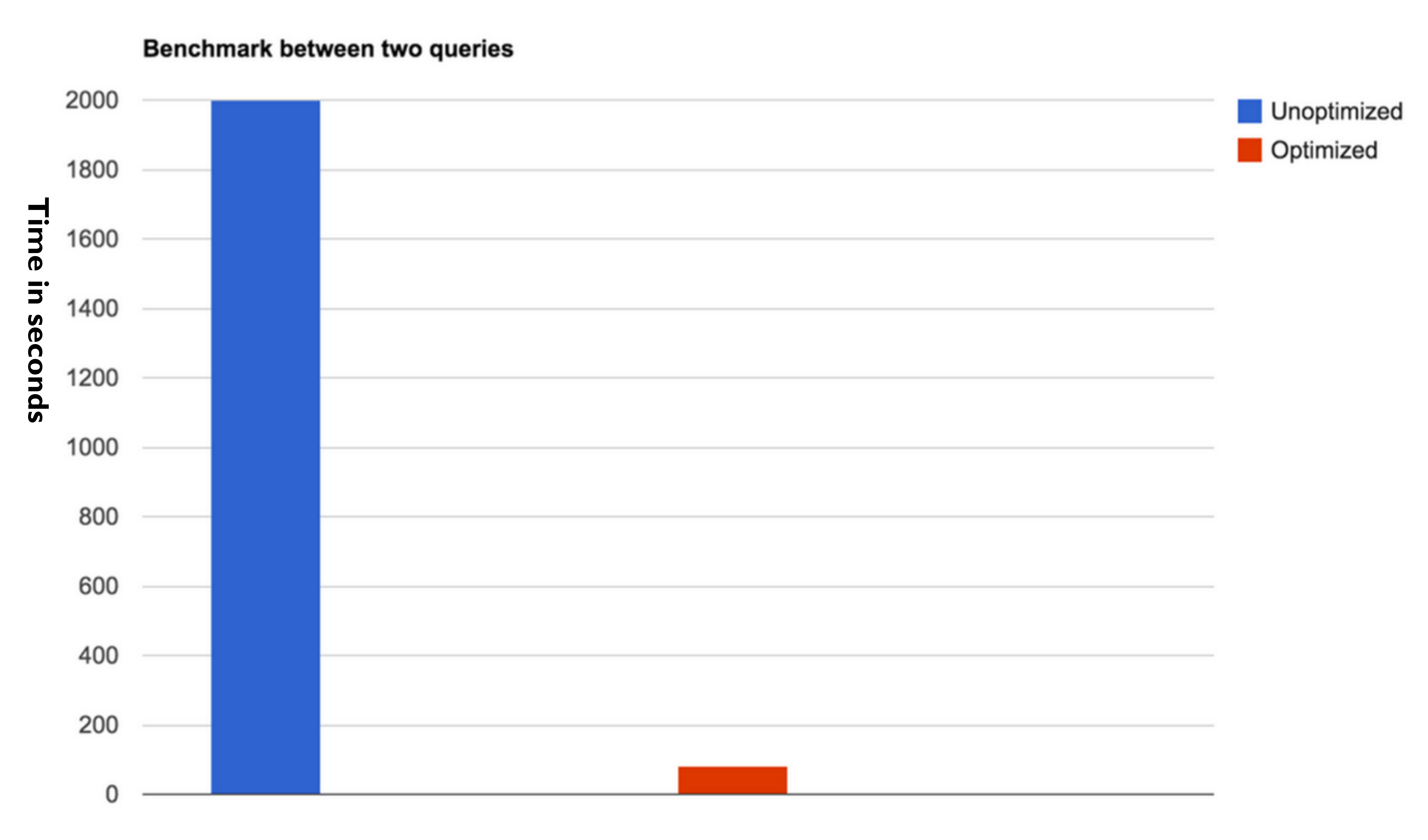
```
/driverAcceptanceRate?
geo_dist(10, [37, 22])&
time_range(2015-02-04,2015-03-06)&
aggregate(timeseries(7d))&
eq(msg.driverId,1)
```



```
        }
    },
    "aggs": {
        "pick_up_counts": {
            "terms": {
                "field": "tags"
            }
        }
    }
}
```

Elasticsearch Query Can Be Optimized

- Pipelining
- Validation
- Throttling



Benchmark between two queries
Time in seconds
2000
1800
1600
1400
1200
1000
800
600
400
200
0
🟦 Unoptimized
🟥 Optimized

Elasticsearch Can Be Replaced

Processing
Storage
Query
Samza
Spark
Warm-ES
Cold-ES
Query Layer
Applications

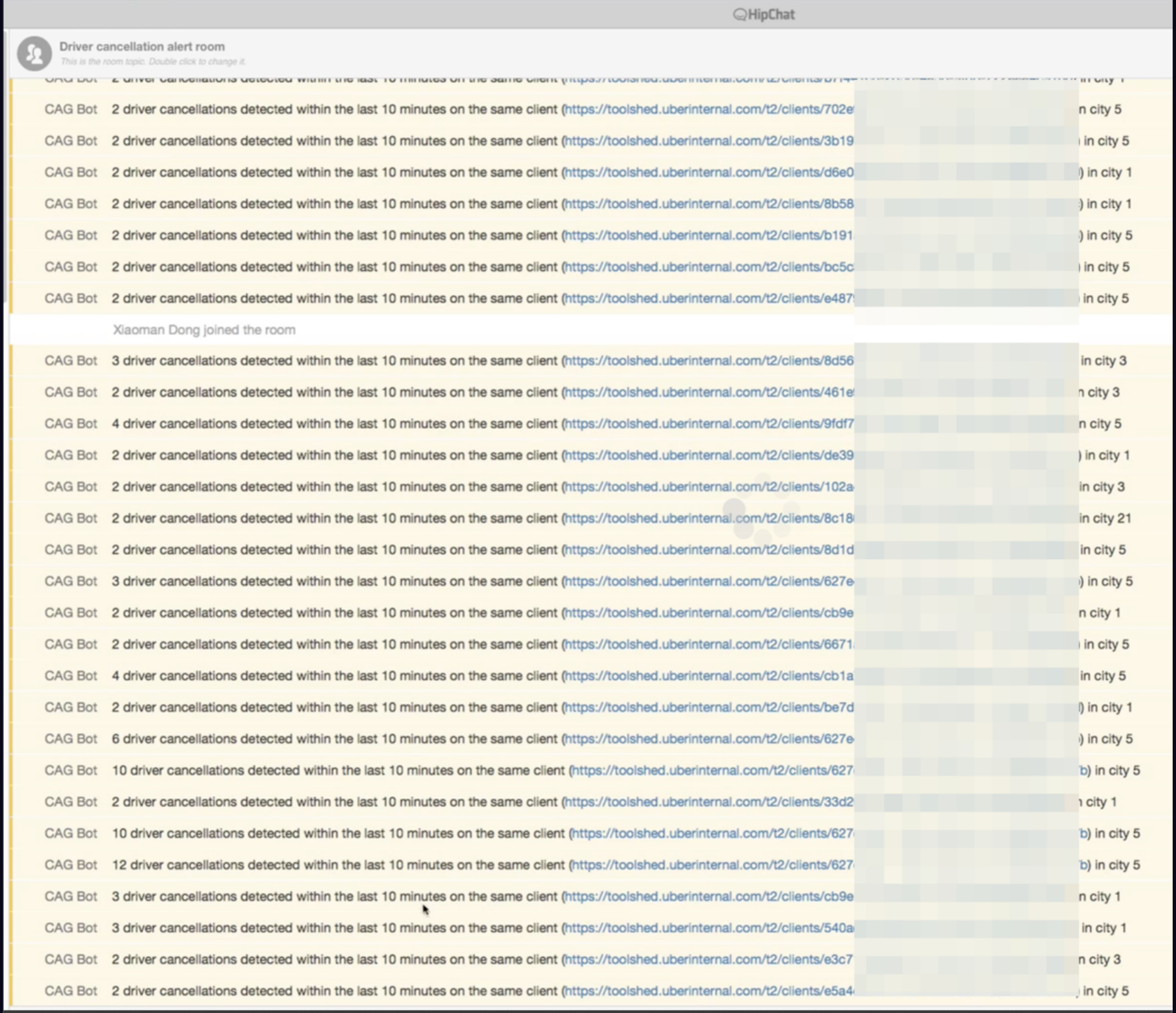
There’s one more thing

There are always patterns in streams

There is always need for quick exploration

How many drivers cancel a request 10 times in a row within a 5-minute window?

Which riders request a pickup from 100 miles apart within a half hour window?



HipChat

Driver cancellation alert room

This is the room topic. Double click to change it.

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/b71 [redacted] in city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/702e [redacted] n city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/3b19 [redacted] in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/d6e0 [redacted]) in city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/8b58 [redacted]) in city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/b191 [redacted]) in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/bc5c [redacted] in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/e487 [redacted] in city 5

Xiaoman Dong joined the room

CAG Bot 3 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/8d56 [redacted] in city 3

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/461e [redacted] n city 3

CAG Bot 4 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/9fdf7 [redacted] n city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/de39 [redacted]) in city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/102a [redacted] in city 3

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/8c18 [redacted] in city 21

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/8d1d [redacted] in city 5

CAG Bot 3 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/627e [redacted]) in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/cb9e [redacted] n city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/6671 [redacted] in city 5

CAG Bot 4 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/cb1a [redacted] in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/be7d [redacted]) in city 1

CAG Bot 6 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/627e [redacted]) in city 5

CAG Bot 10 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/627 [redacted] b) in city 5

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/33d2 [redacted] n city 1

CAG Bot 10 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/627 [redacted] b) in city 5

CAG Bot 12 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/627 [redacted] b) in city 5

CAG Bot 3 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/cb9e [redacted] n city 1

CAG Bot 3 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/540a [redacted] in city 1

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/e3c7 [redacted] n city 3

CAG Bot 2 driver cancellations detected within the last 10 minutes on the same client (https://toolshed.uberinternal.com/t2/clients/e5a4 [redacted] in city 5

Complex Event Processing

```
FROM driver_canceled#window.time(10 min)
SELECT clientUUID, count(clientUUID) as cancelCount
GROUP BY clientUUID HAVING cancelCount > 10
INSERT INTO hipchat(room);
```

Implementation Becomes Easy

Thank You!

UBER