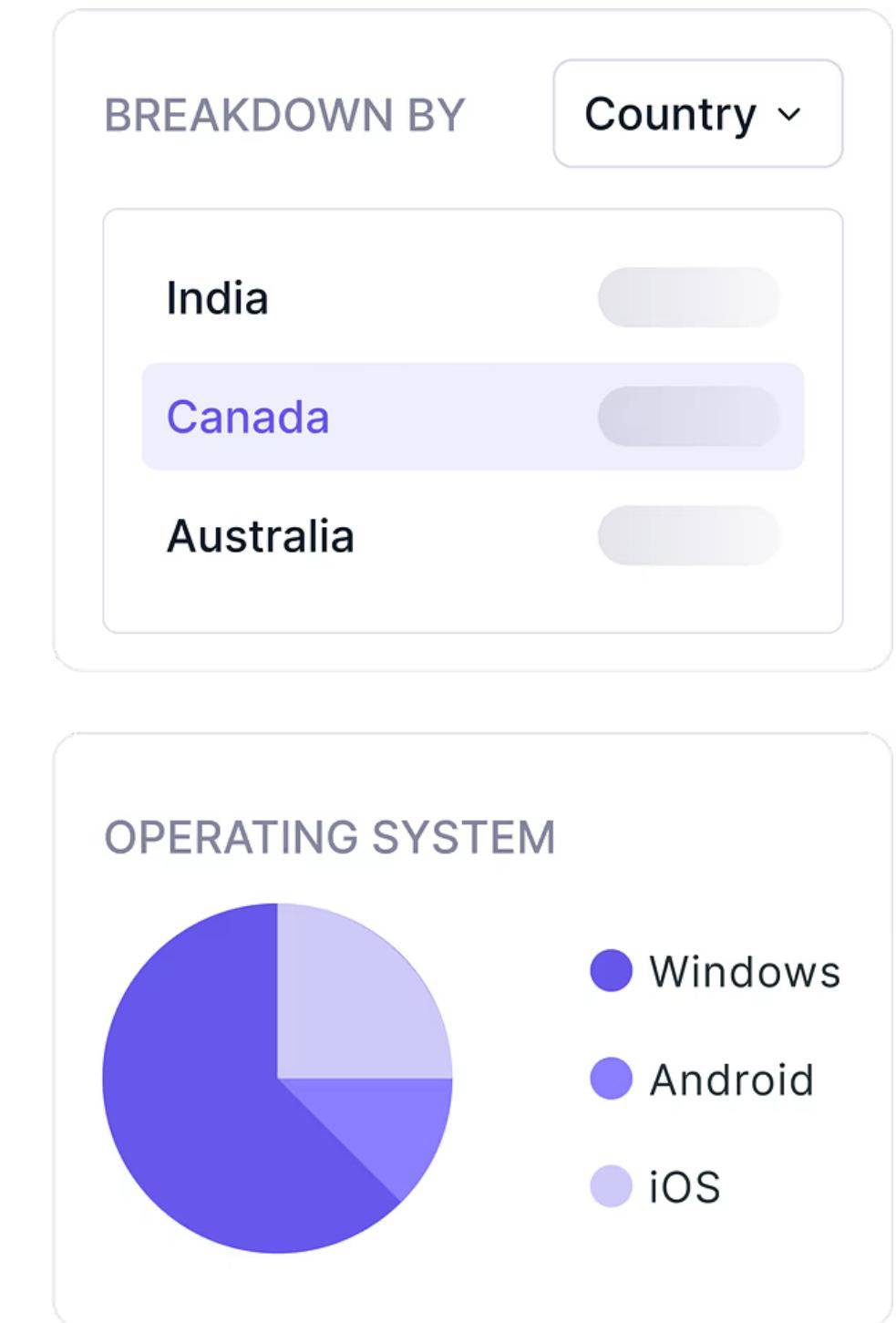
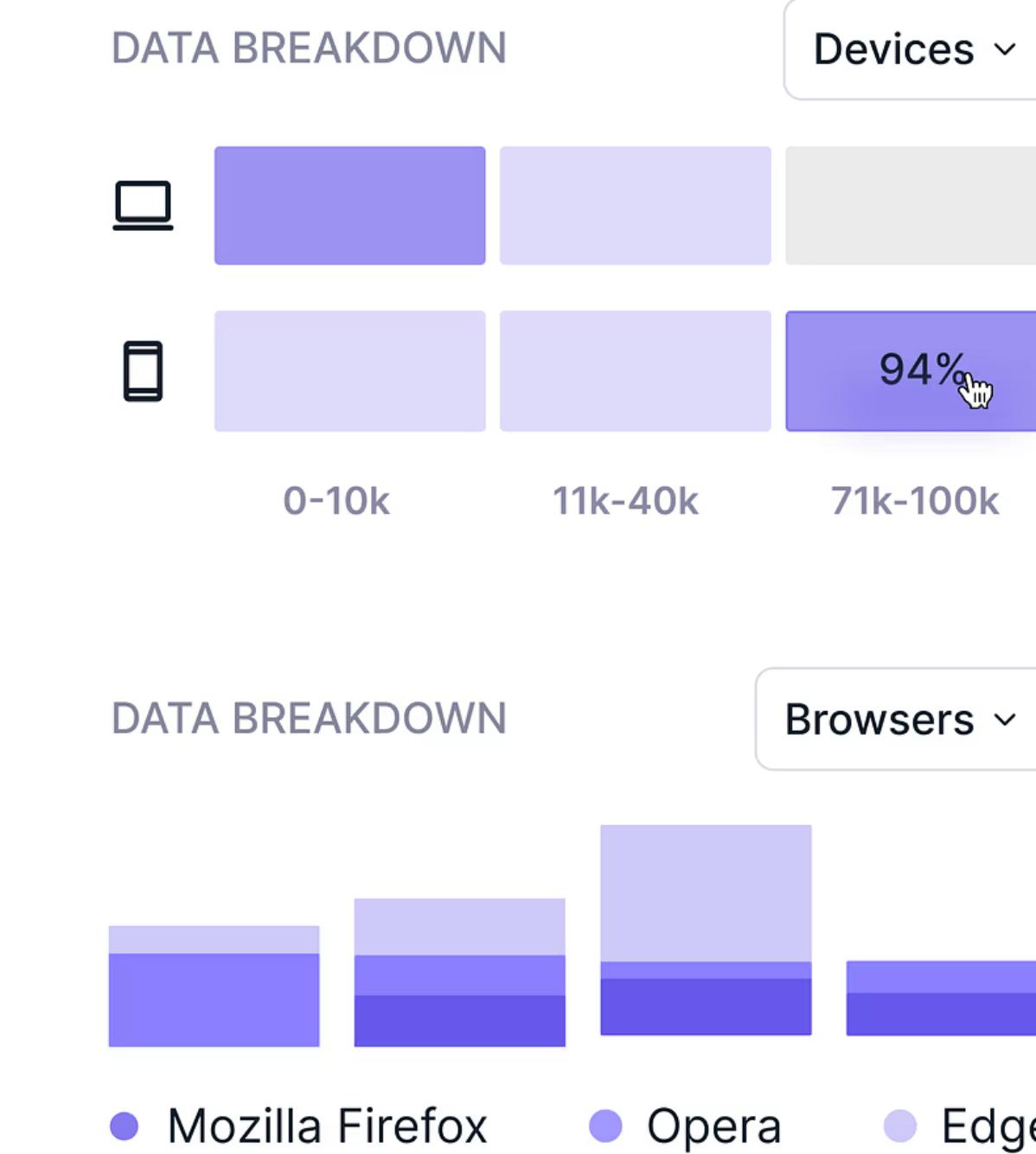
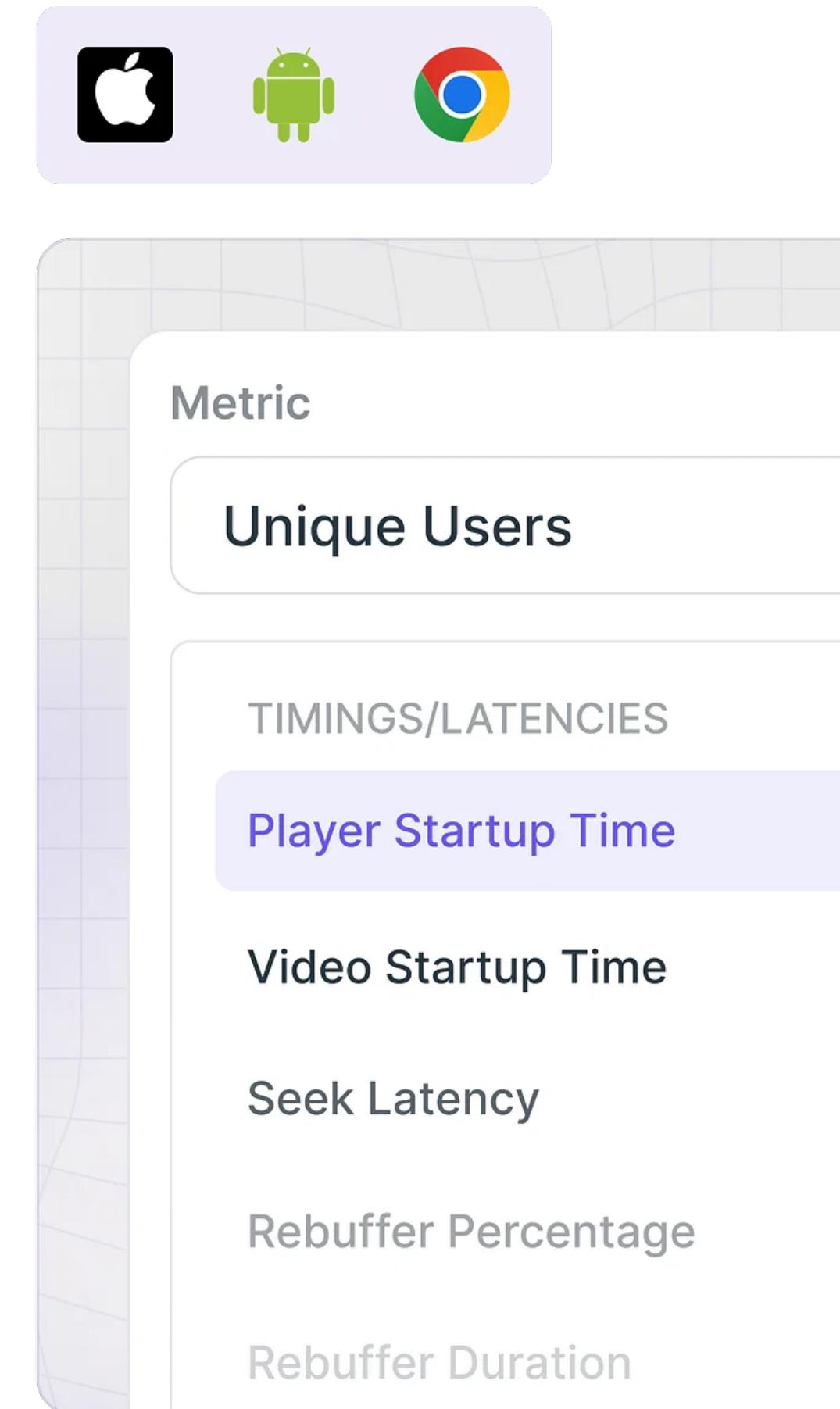




# GUMLET

Analysing 2B+ video and image  
requests per day using ClickHouse

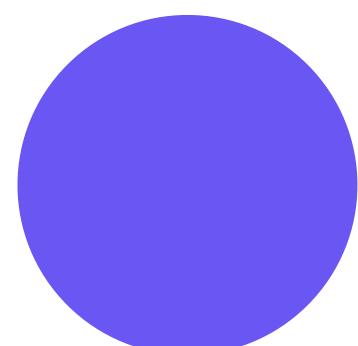




# Analytics is lifeline for video

- Unique users
- Playback time
- Views
- Re-buffer metrics
- Latency metrics
- Bandwidth stats
- Image bandwidth
- Requests
- CDN hit ratio
- Error statistics
- Popular images and videos
- Transformation statistics

## Metrics Tracked

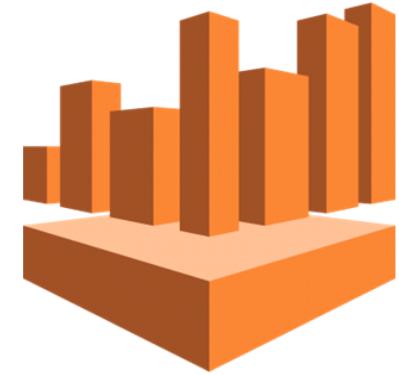


# Previous Architecture

- Real-time ingestion via BigQuery
- BigQuery on-demand queries for all analytics

## Challenges

- High ingestion / analytics cost
- Limited to 100 concurrent queries
- Storage cost was logical bytes and not physical



Amazon Athena



amazon  
REDSHIFT



## Choosing New System

- Scalable storage - tens of TBs data to start with
- No price per query - we can end up with millions of queries
- Exceptional documentation
- No maintenance (ideally)



# Gumlet Stack After ClickHouse

Ingestion process was updated  
to make it scalable



## Fixed Costs (almost)

Compute is now fixed cost. Only storage is variable

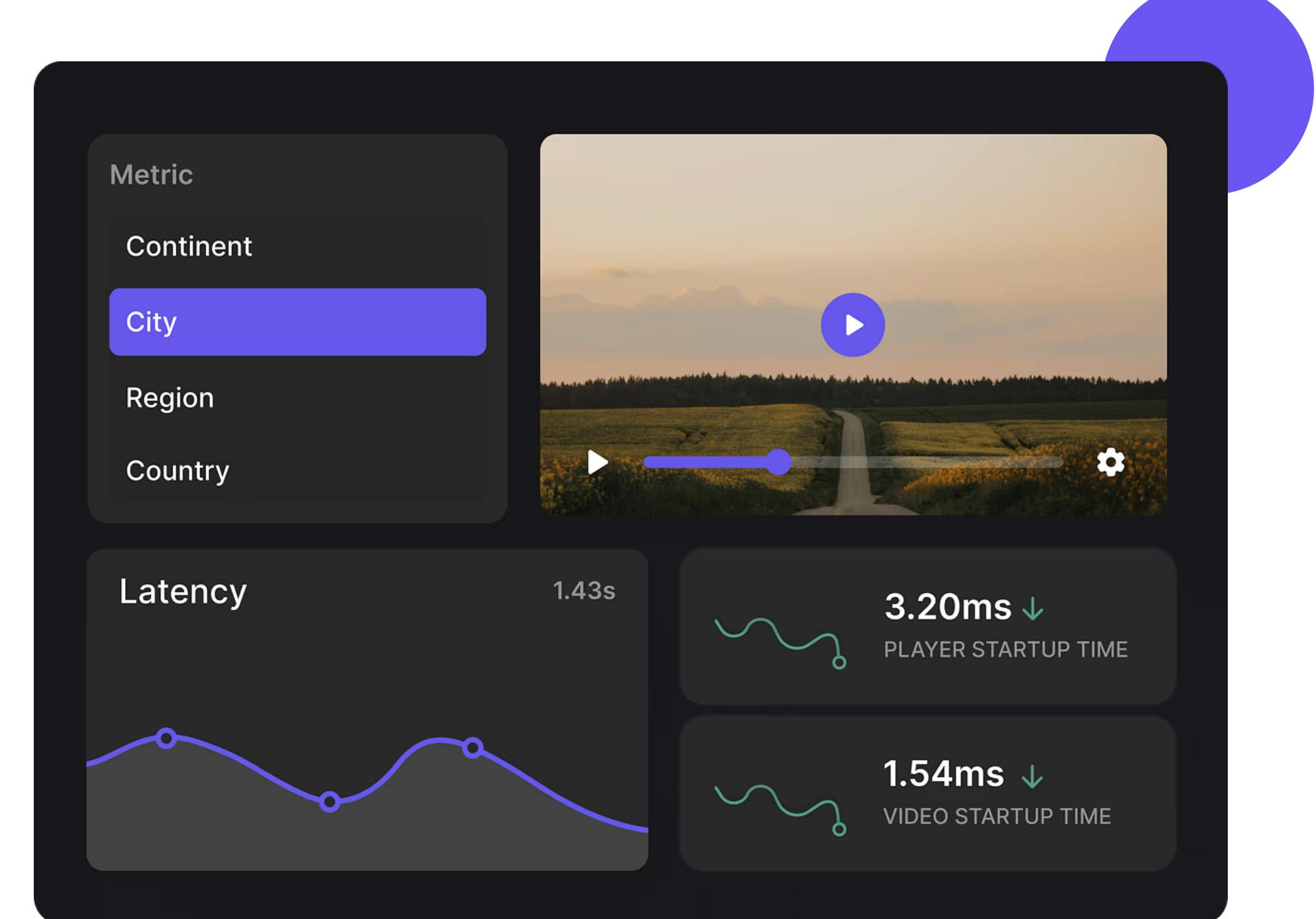
## Ingestion Speed from S3

CSV ingest from S3 went from 10-15 min an hour to < 2 min

## Fast Response Times

Our dashboards are now powered by ClickHouse

# ClickHouse Impact



## Changes needed

---

### BI Tool Change

BI tool was changes from Data Studio to Metabase

### Data Migration

We had to carefully export data as Parquet files and import in CH. No native migration is available

### ETL tool changes

Not all ETL tools support ClickHouse. If not supported, that needs to be changed.

## Scale Issues

---

### DateTime64 is not performant

In a large DB, DateTime must be used.

### Schema changes needed for Sampling

CH provides native sampling but it must be defined at time of table creation.

### LowCardinality and Delta

If data size is huge, LC and Delta can provide big savings in query and storage performance / cost.



## Caveats



**GUMLET**

[@adityapatadia](https://twitter.com/adityapatadia)   [@gumletha](https://twitter.com/gumletha)