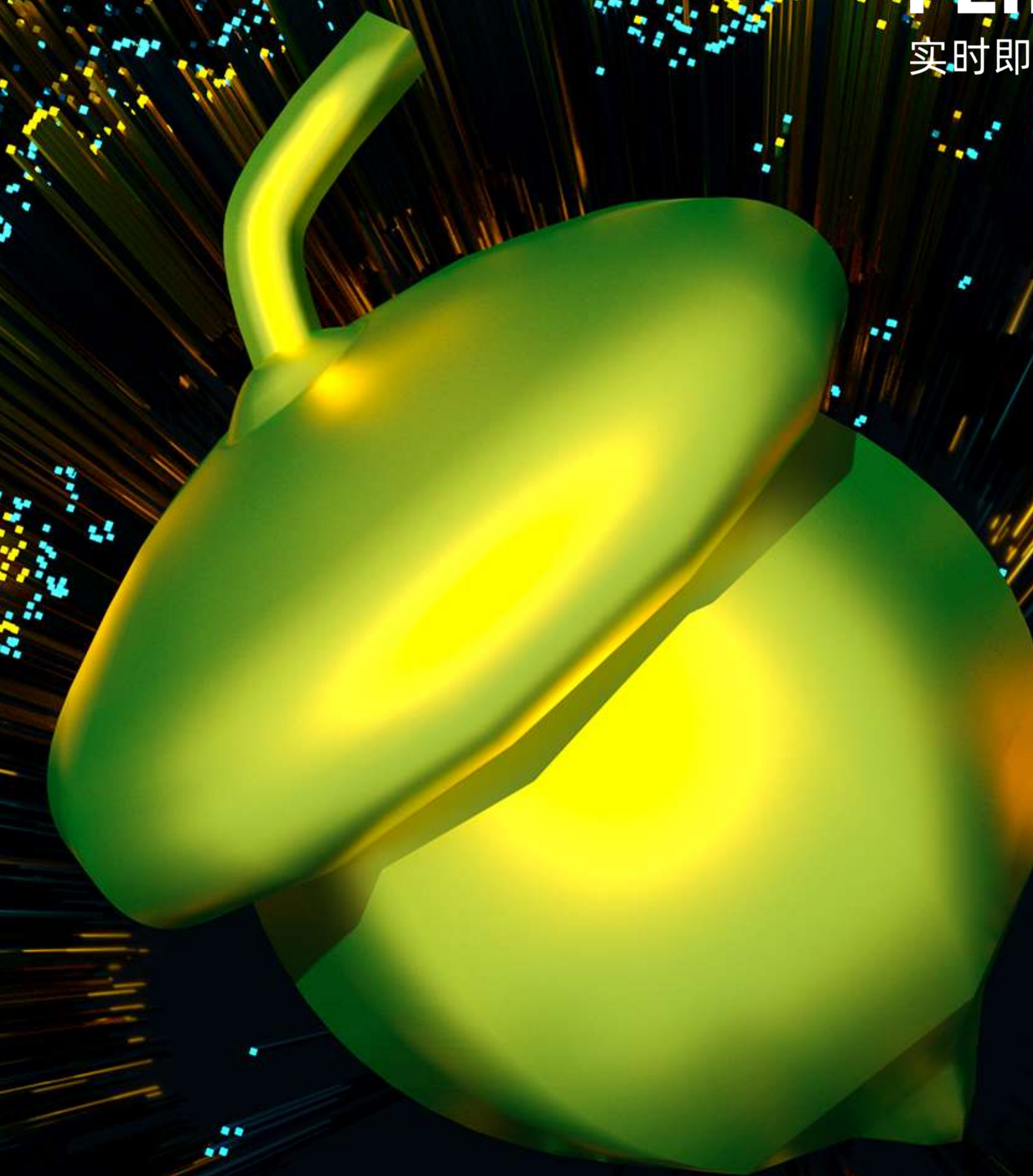


Storage Reimagined for a Streaming World

| **Srikanth Satya** VP Engineering, Dell Technologies

| **滕昱** 戴尔科技集团 软件开发总监



北京·国家会议中心
2019.11.28-11.30

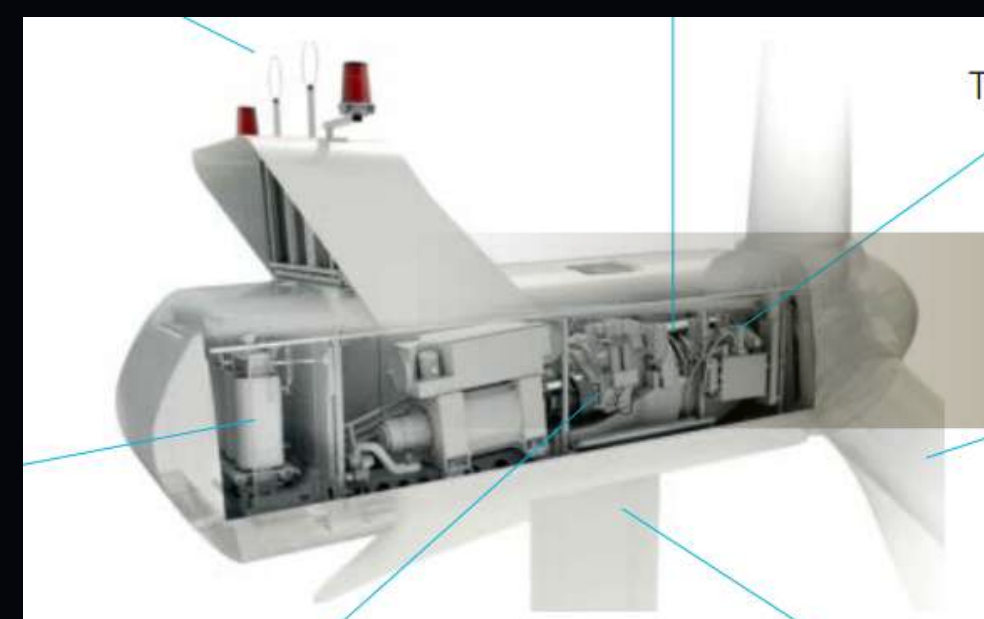
Why is Streaming Attractive?

- Shrink to zero the time it takes to derive business value from data
- Consolidate infrastructure with unified streaming + batch platform
- Key streaming capabilities
 - Treat data as **continuous and infinite** rather than finite and static
 - Deliver **accurate results** processing data continuously even with late arriving or out of order data
 - Deliver consistently fast results by **dynamically scaling** data ingest, storage, and processing capacity in coordination with the volume of data arriving
 - Process data at **point of ingest**

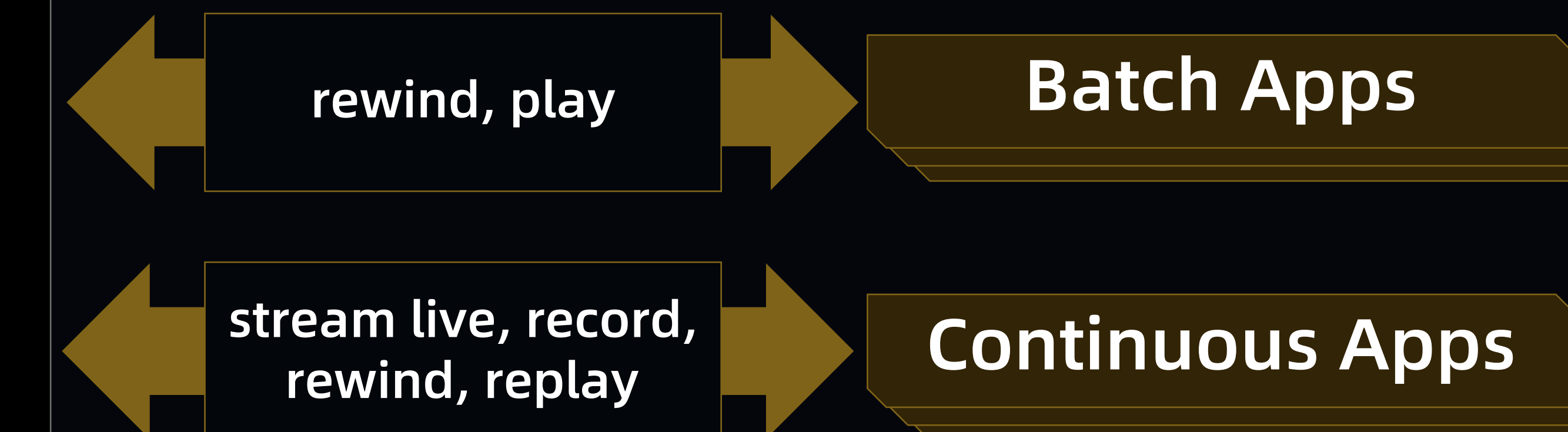
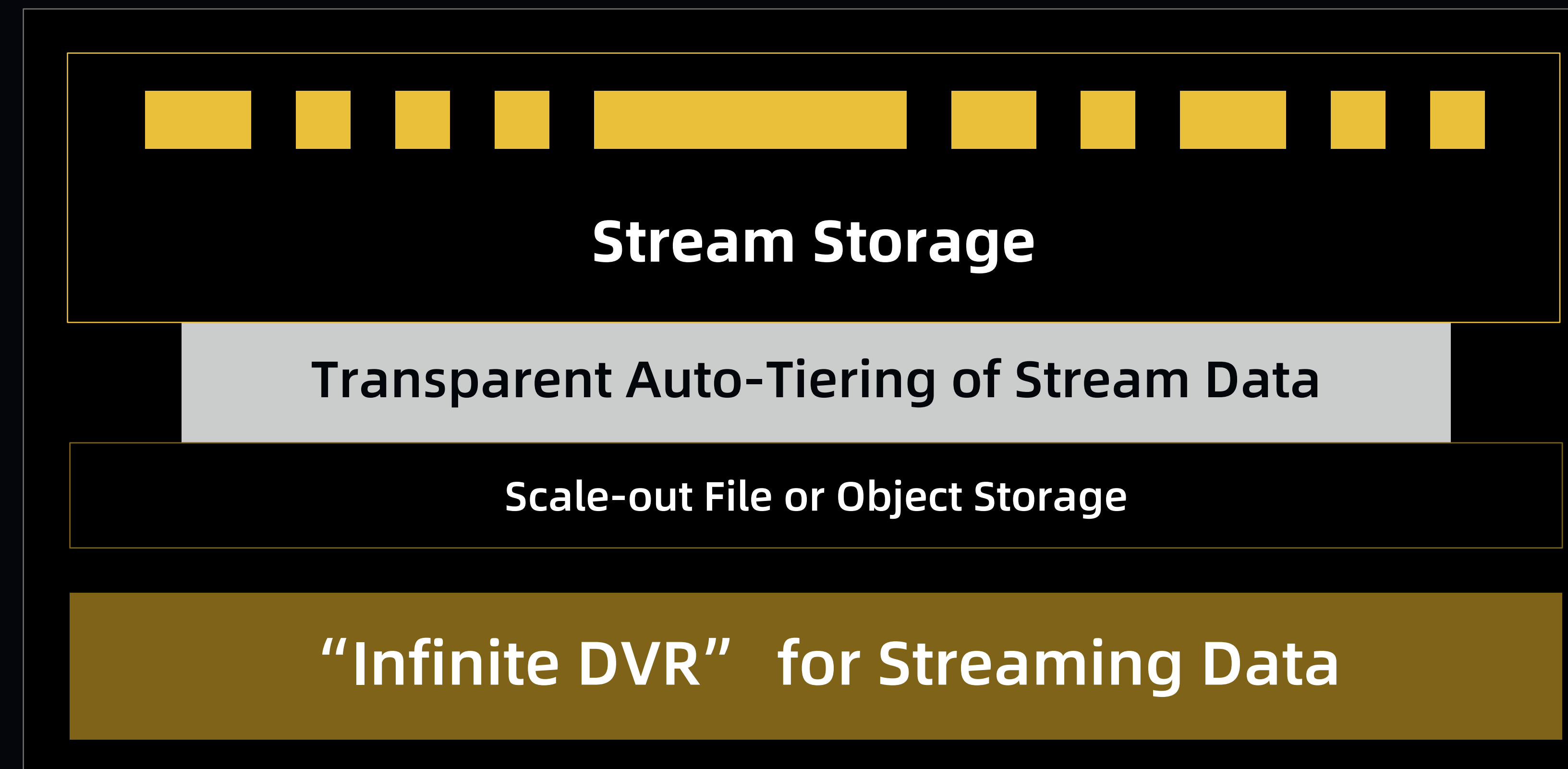
Keys To Enabling Adoption

- **Lowering the Complexity Barrier to Entry**
 - Easily consumable programming models, e.g. Flink SQL
 - Rich set of pre-packaged streaming data middleware: Ingest and pub/sub, Stream processing / SQL, Search, Time Series
 - Simplified storage infrastructure for infinite and continuous data
- **Providing a platform with robust production capabilities**
 - Stability + Security + Serviceability

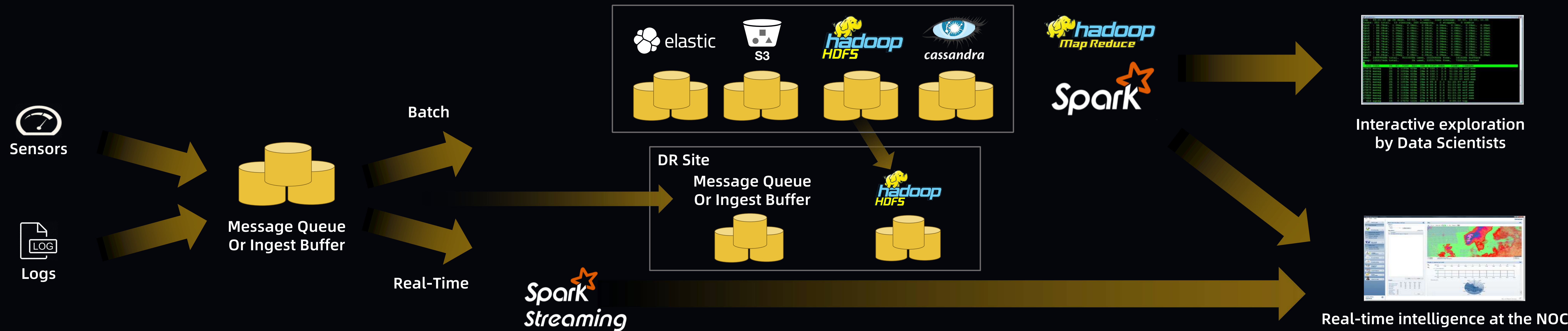
Simplified Storage Infrastructure: The DVR Analogy



Streaming
Data Sources



Today's Pipeline

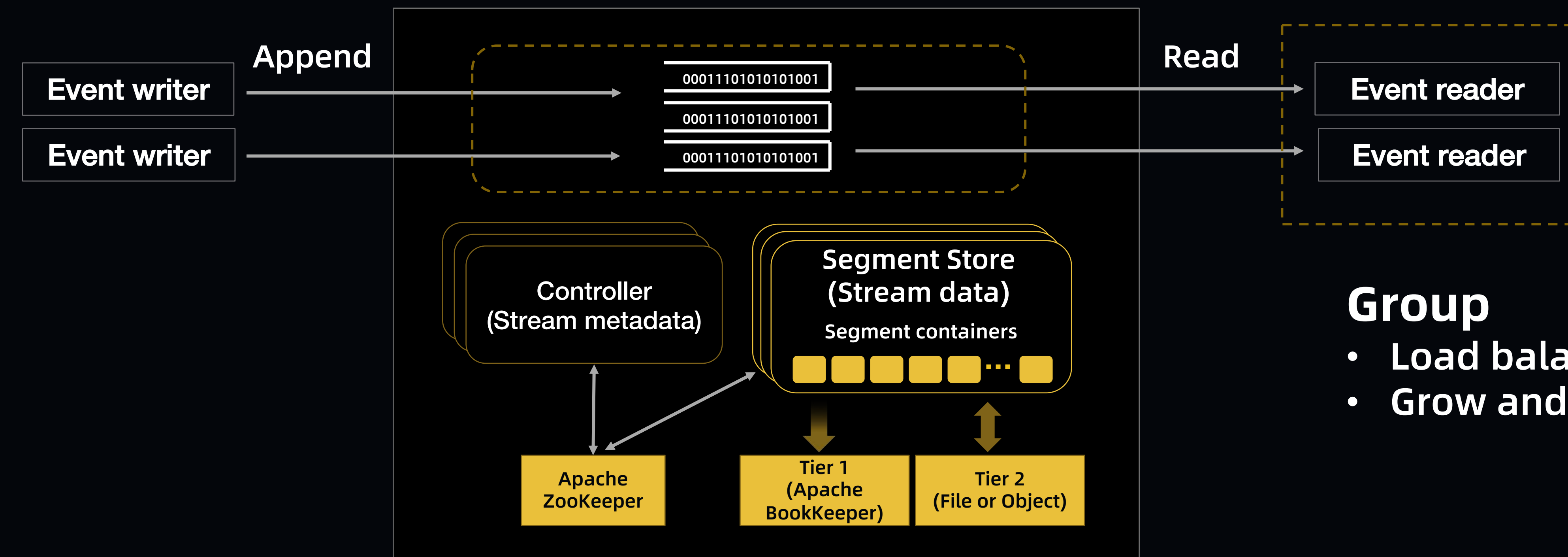


Pravega and Streams

Ingest stream data

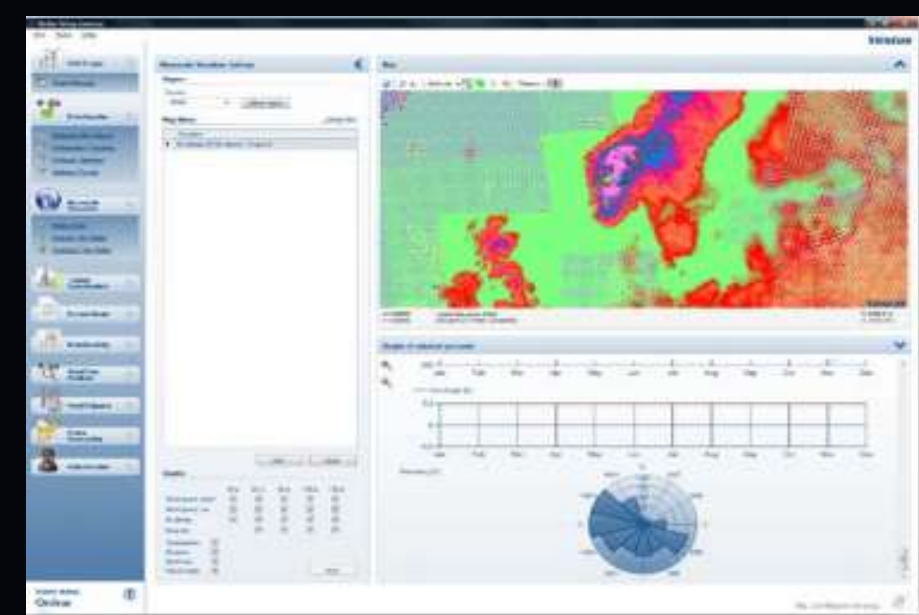
Pravega

Process stream data

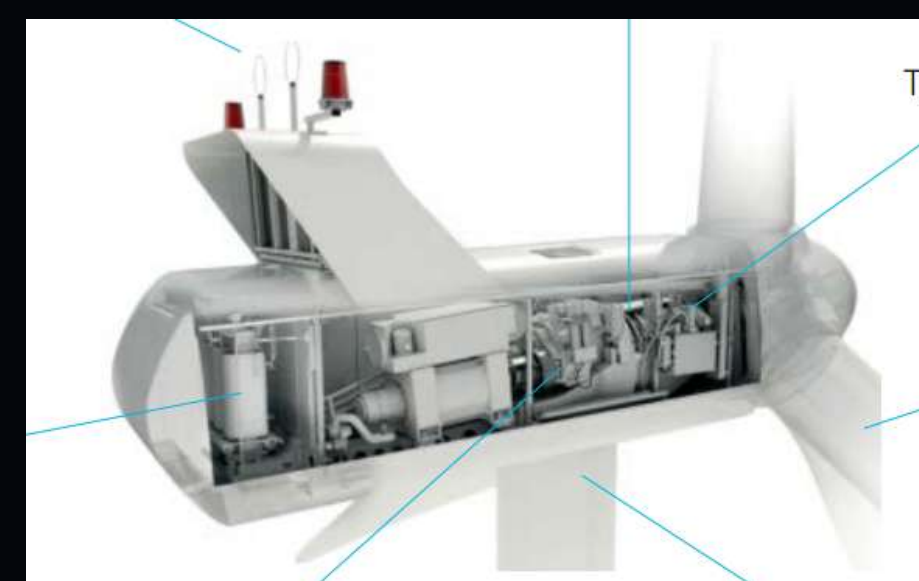


Operational Scenario: Today

How to test & deploy new version of analytics business logic?



Streaming
Application
Business Logic



Recent Data: Kafka Logs

... x=5 z=6 y=2 x=4 a=7 y=5 ...
older newer

ETL

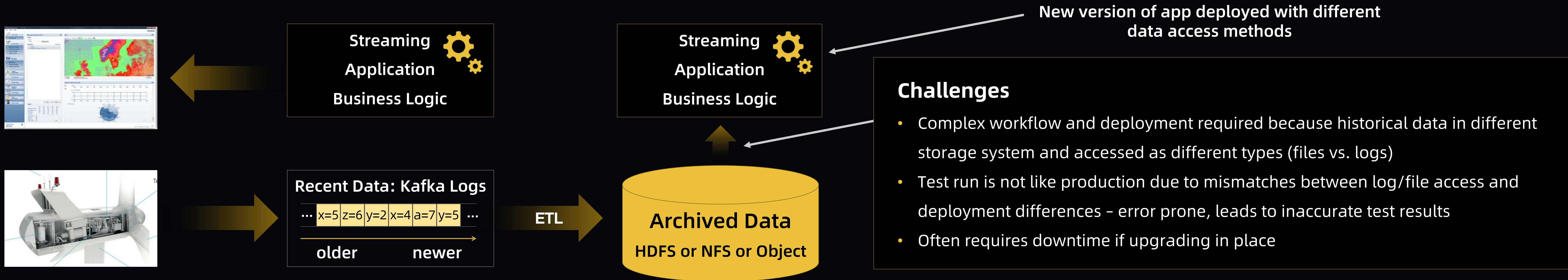
Archived Data
HDFS or NFS or Object

Requirements

- Run new business logic against historical data sets
- Validate correct results for problematic scenarios
- Ensure no regression
- Deploy new business logic in production
- Ensure new version is functioning properly before switching users/apps
- Revert to prior if something fails

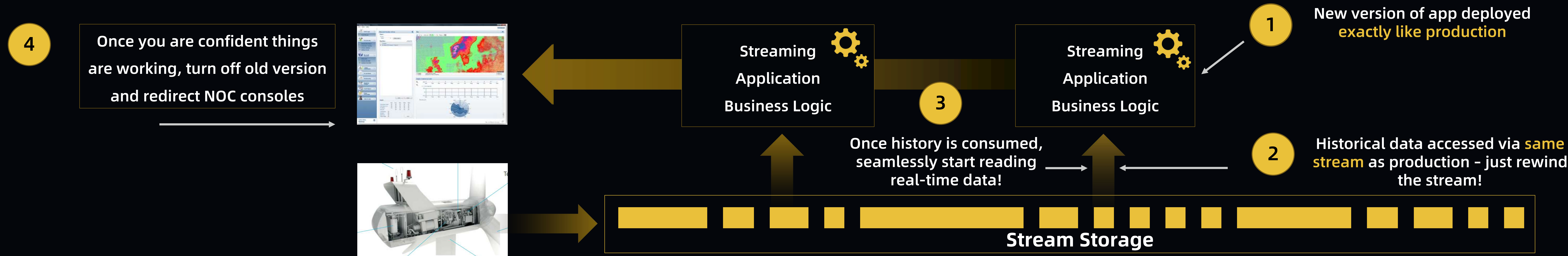
Operational Scenario: Today

How to test & deploy new version of analytics business logic?



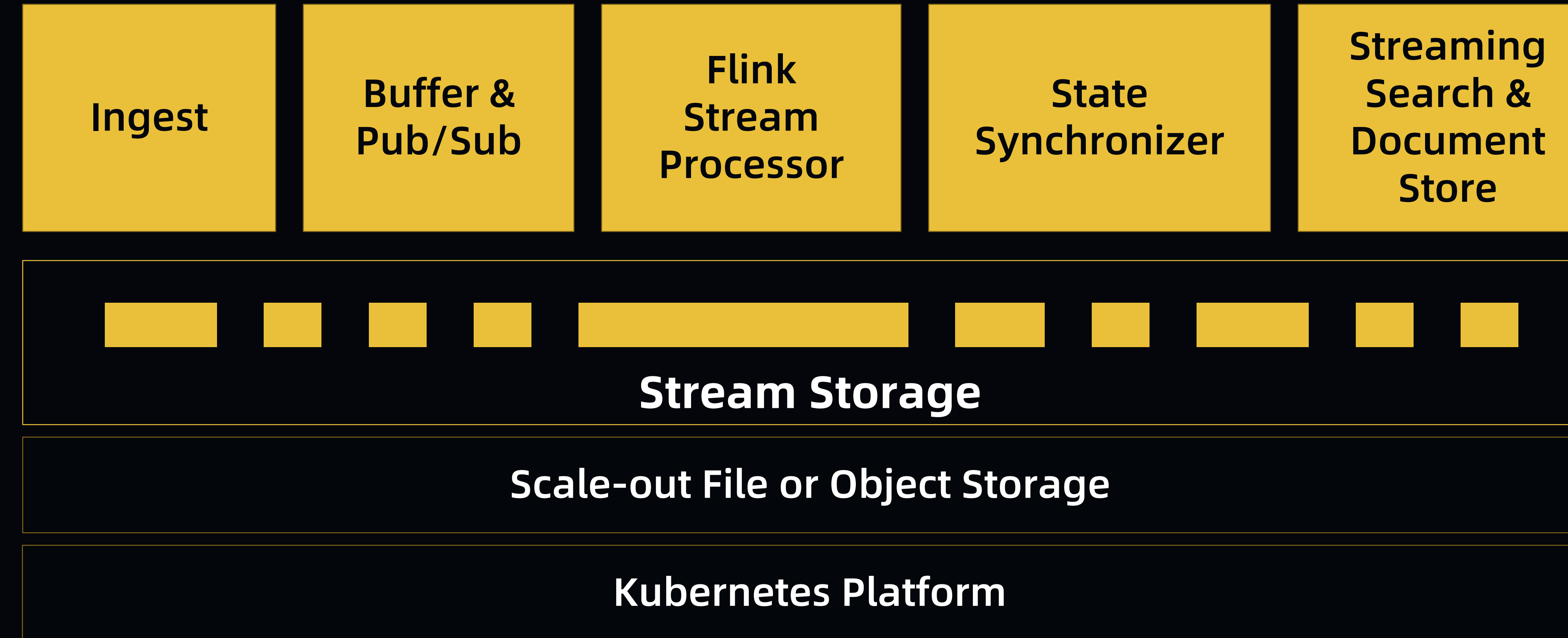
With Streaming Real-time/Batch Unified

How to test & deploy new version of analytics business logic?



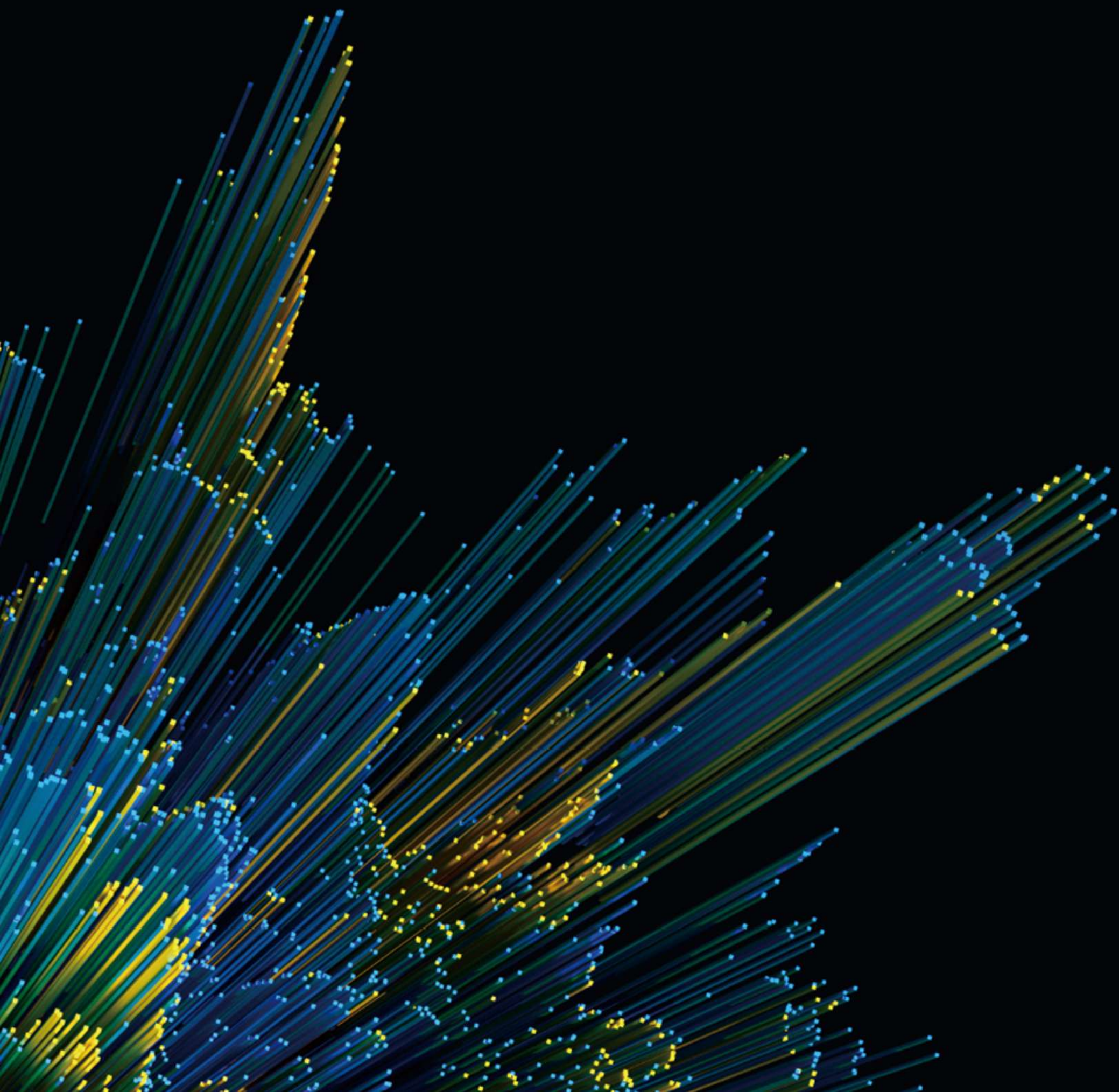
A Modern Streaming Data Platform

Enabling a new generation of
distributed data middleware
reimagined as streaming
infrastructure



- Tiering for “Infinite Streams”
- Dynamic Scaling
 - Automatically scale read/write parallelism based on load & SLO
 - No service interruptions
 - No manual reconfiguration of clients or service resources
- Smart Workload Distribution
 - No need to over-provision servers for peak load
- I/O Path Isolation
 - For tail writes, tail writes, catch-up reads
- Transactions For “Exactly Once”
- Storage Consolidation

Streaming for Connected Cars

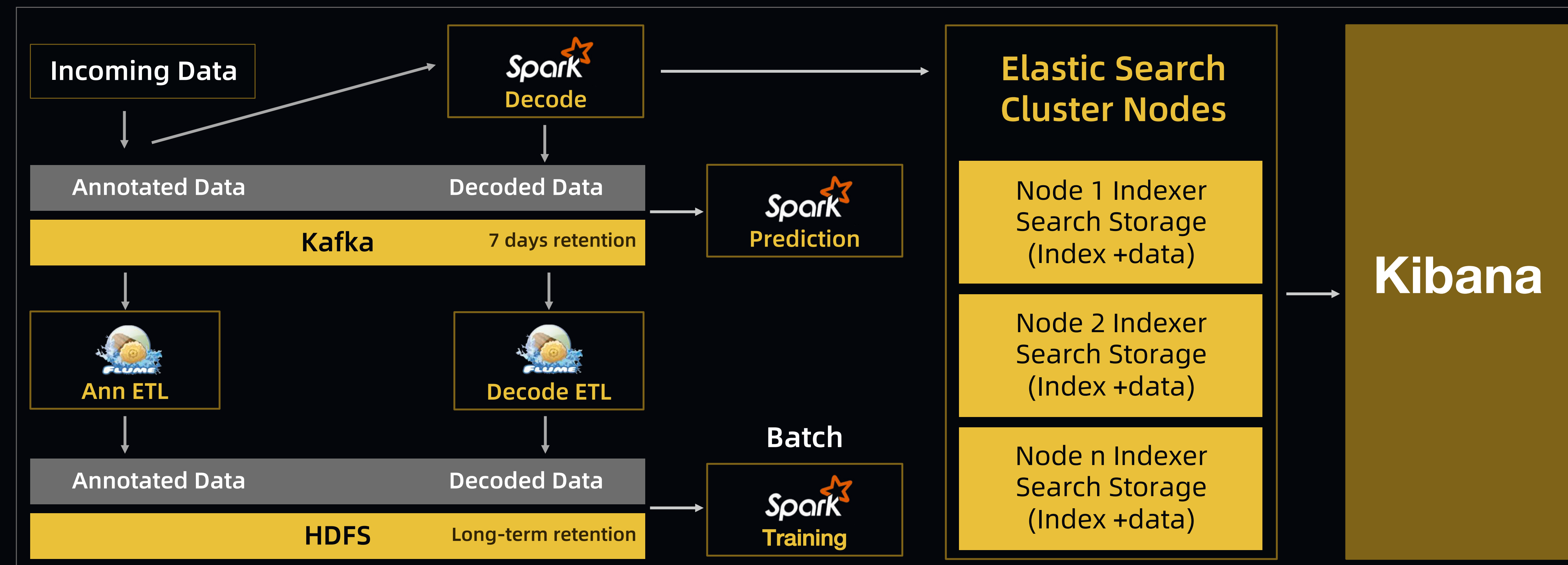


Utilizing Streaming for Connected Cars

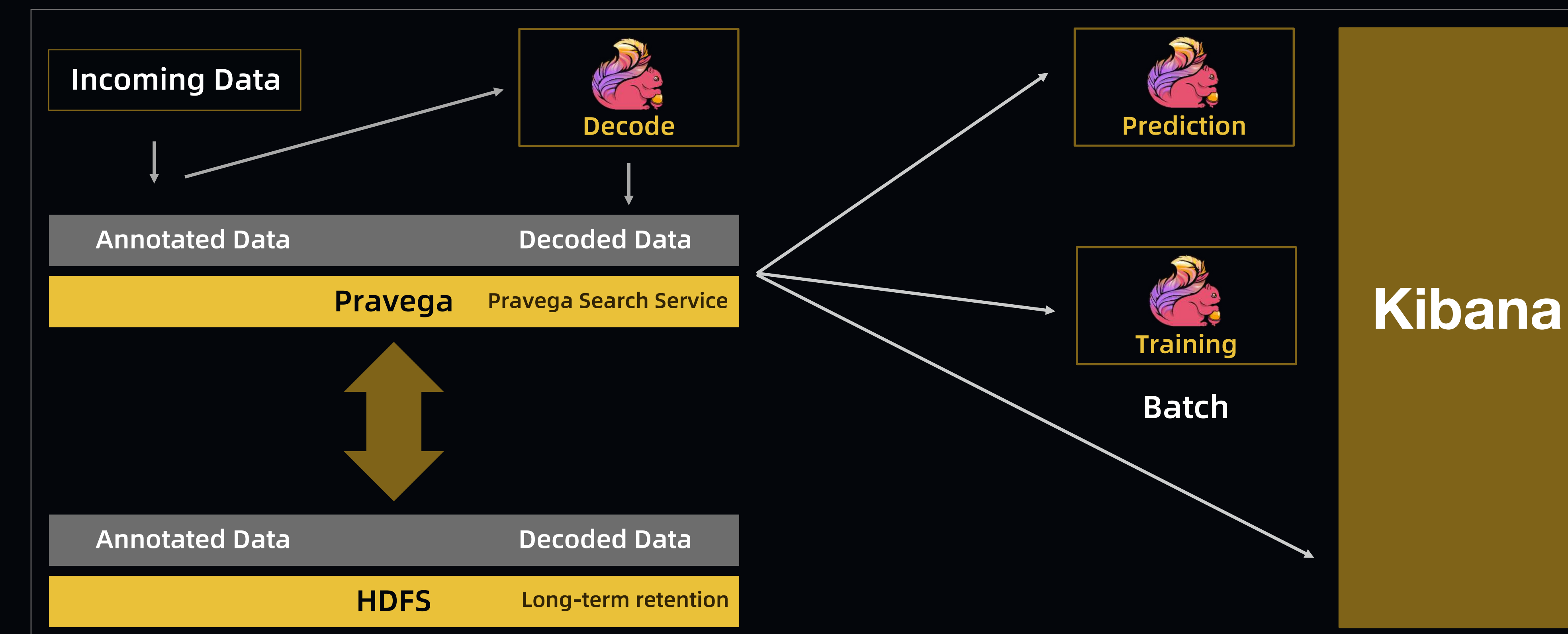
- Vehicular anomaly detection via real-time streaming
- Model training for prediction via batch processing
- Using generated model to predict anomalies via real-time inferencing
- Key success criteria
 - End-to-end event processing bandwidth (events/sec)
 - Model training execution time
 - Reduce storage resource consumption / costs

Connected Car Solution: Architecture Comparisons

Kafka + Spark + Elastic Search Solution



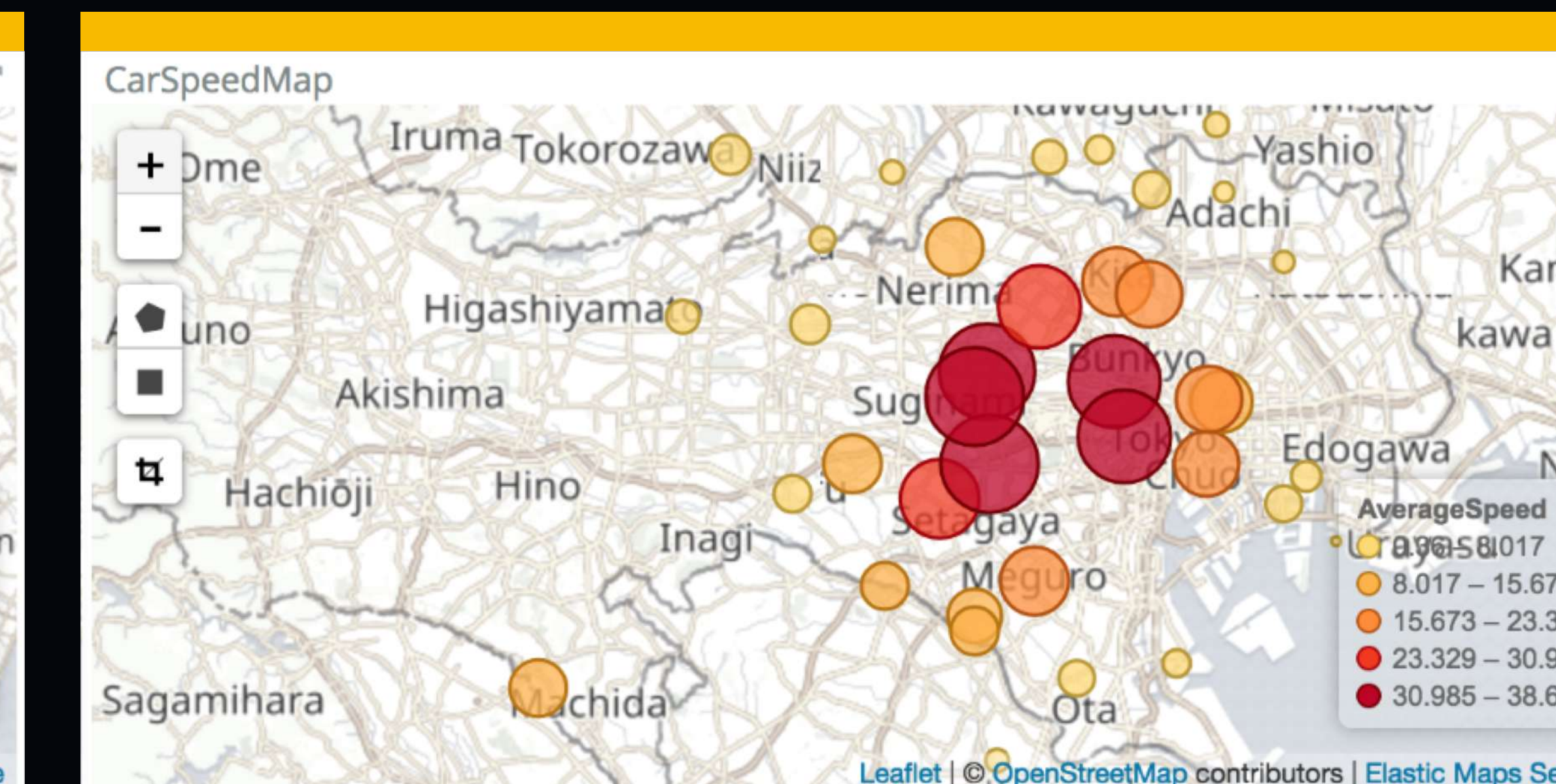
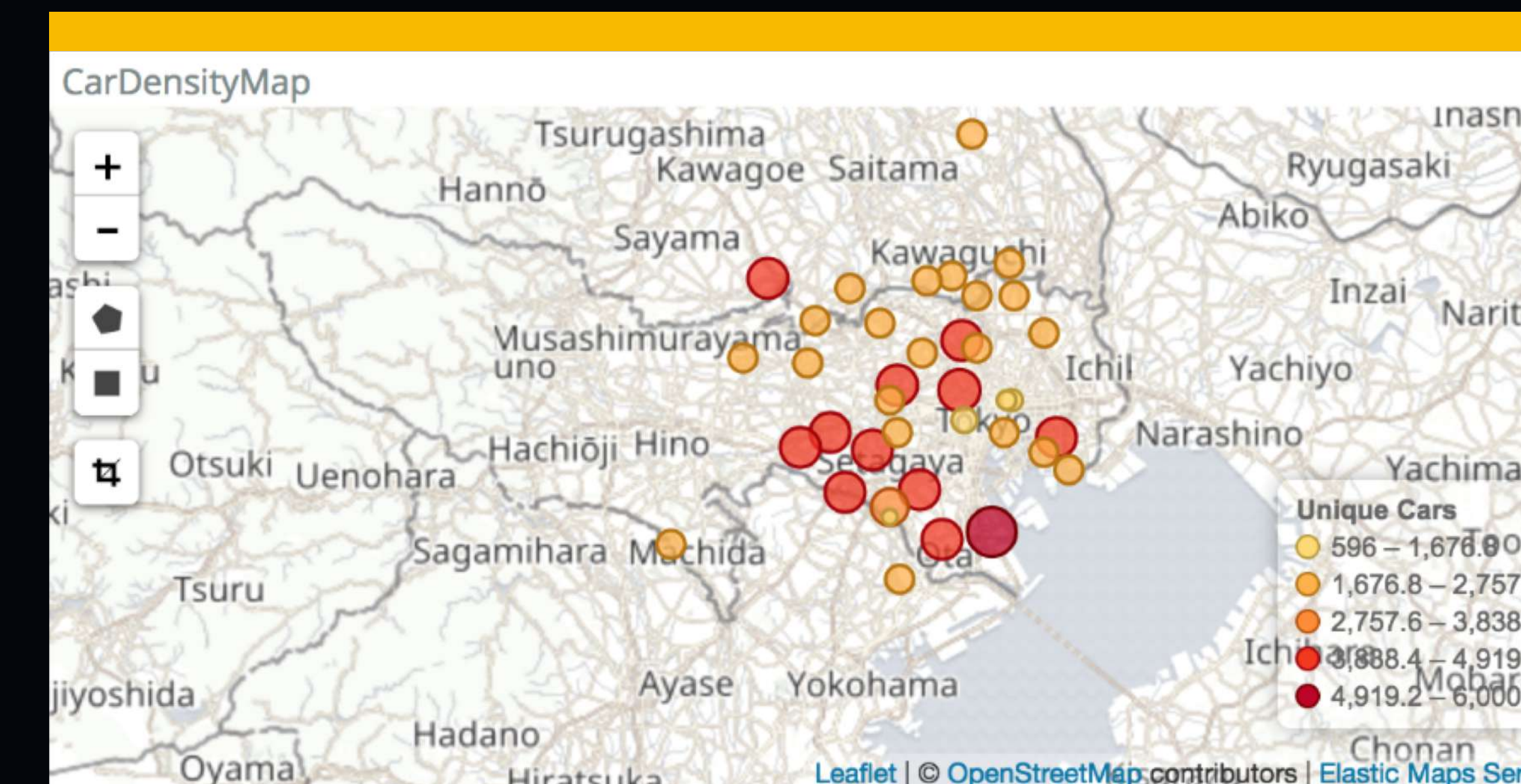
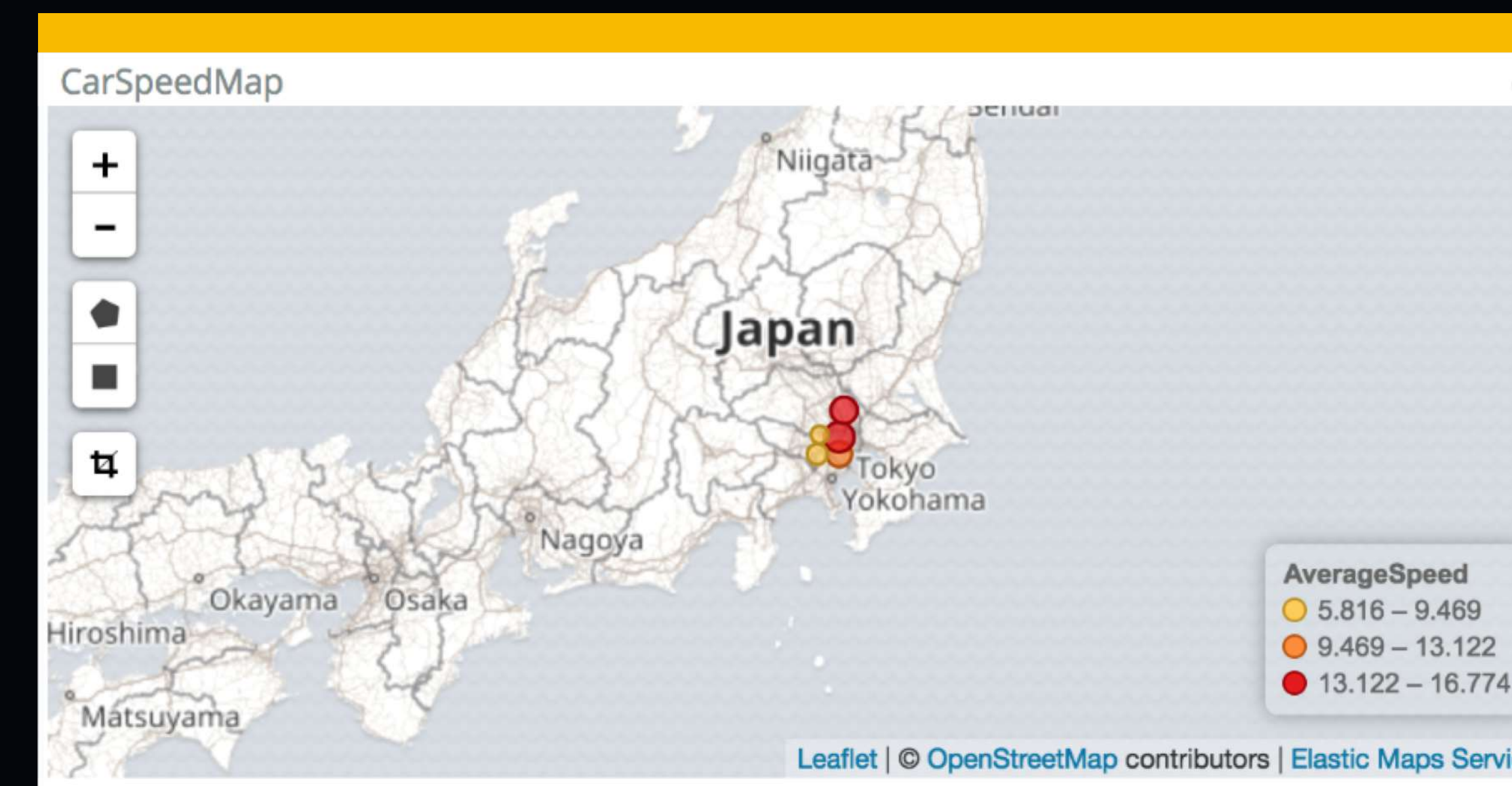
Pravega + Flink Solution



VS

Pravega Search Example - Car Density and Speed

Time interval: 30m



Results Summary

Metric	Kafka / Spark	Pravega / Flink
Storage Capacity (GB)	148.99	45.84
Prediction Pipeline Rate: (events/sec)	78	676
Batch ML Training: Execution time (minutes)	13	6.5

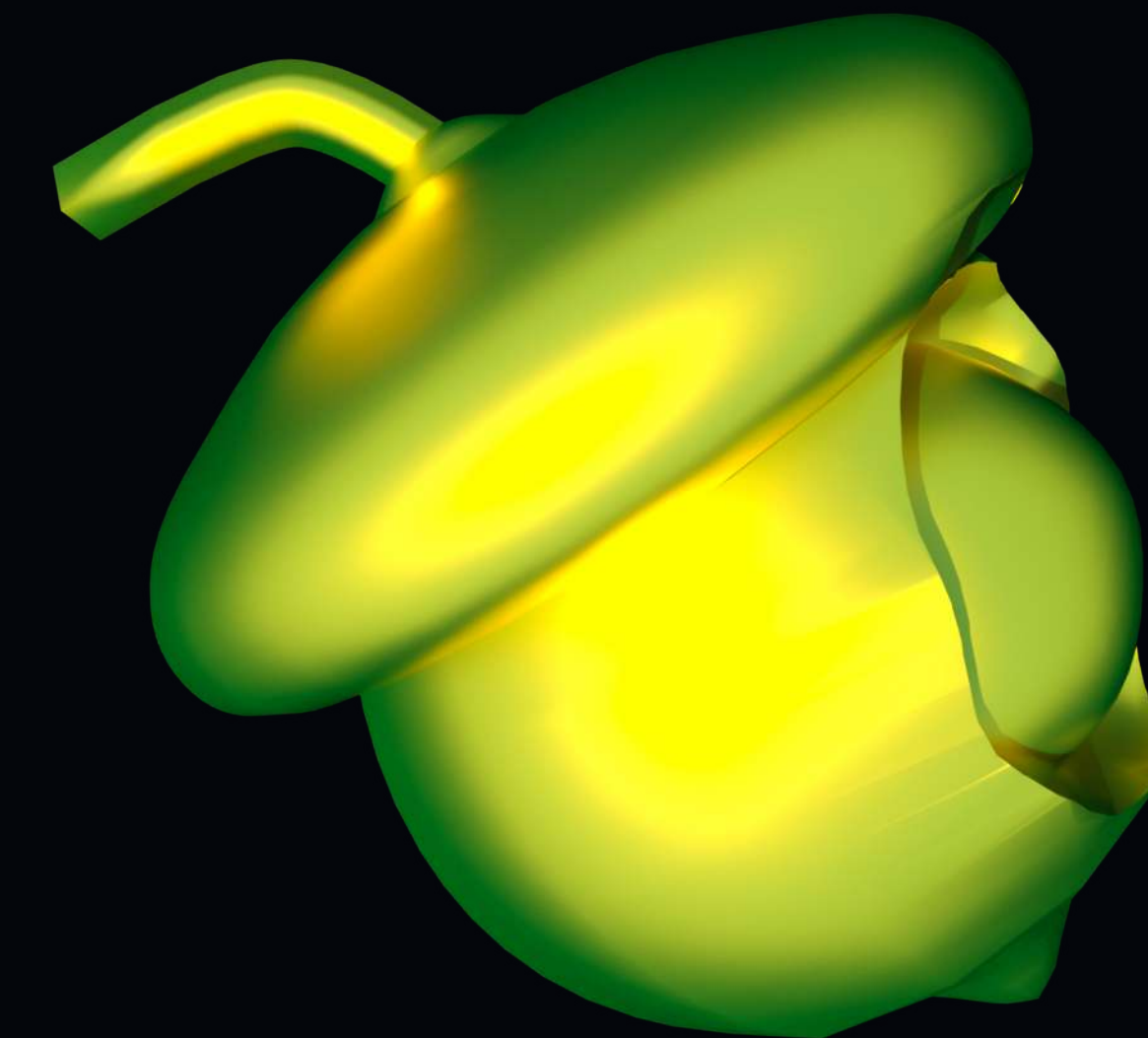
Pravega: Reliable Streaming Storage

- Pravega: an open source project with an open community
 - Stream Storage
 - Ingest Buffer with Pub/Sub built on top of streams
 - And Flink connectors!
- Project Website: <http://pravega.io>
- Source Repo: <http://github.com/pravega/pravega>
- Flink Connectors: <http://github.com/pravega/flink-connectors>



DELL Technologies

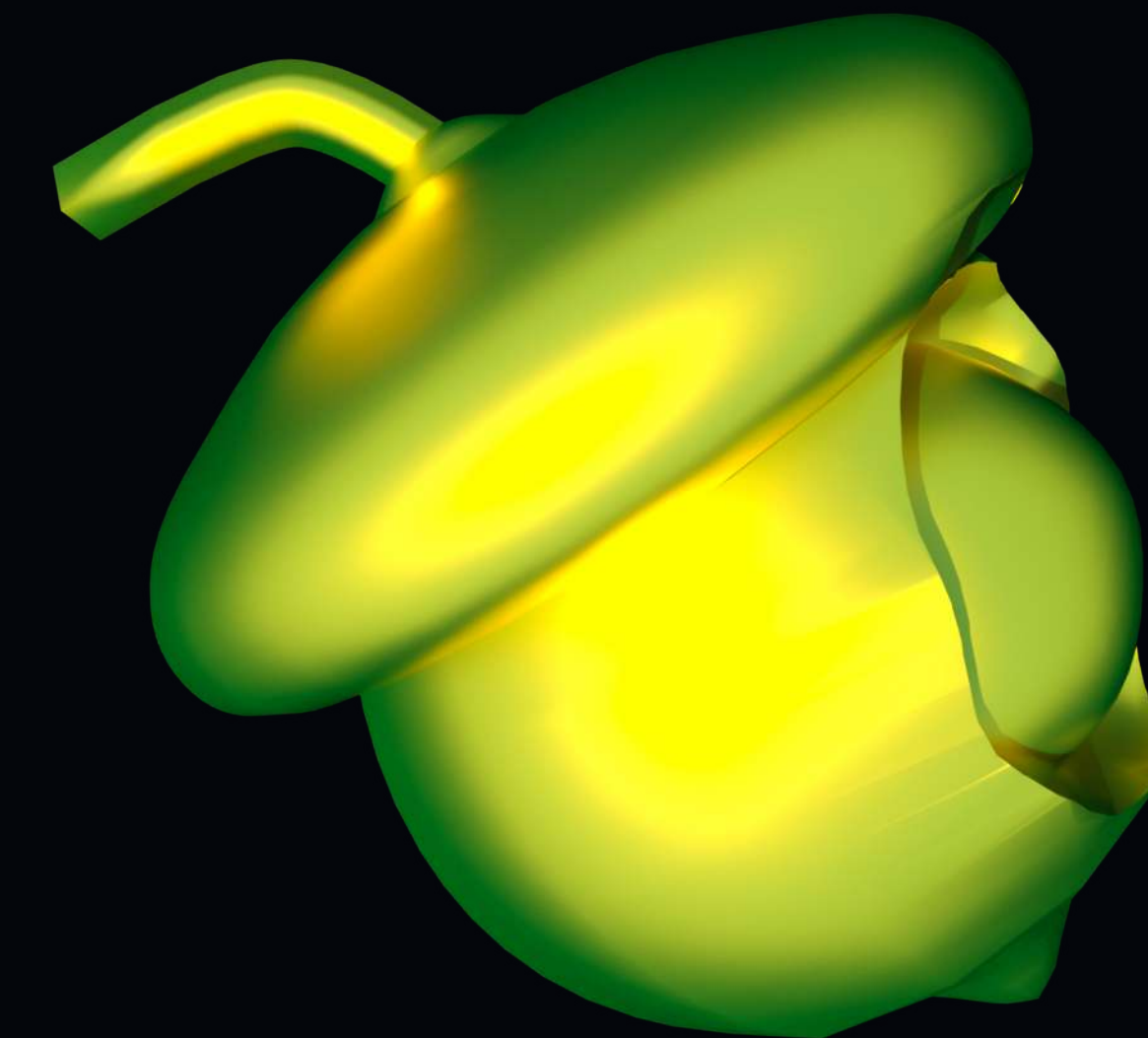
Flink Forward Asia



2019

Thanks!

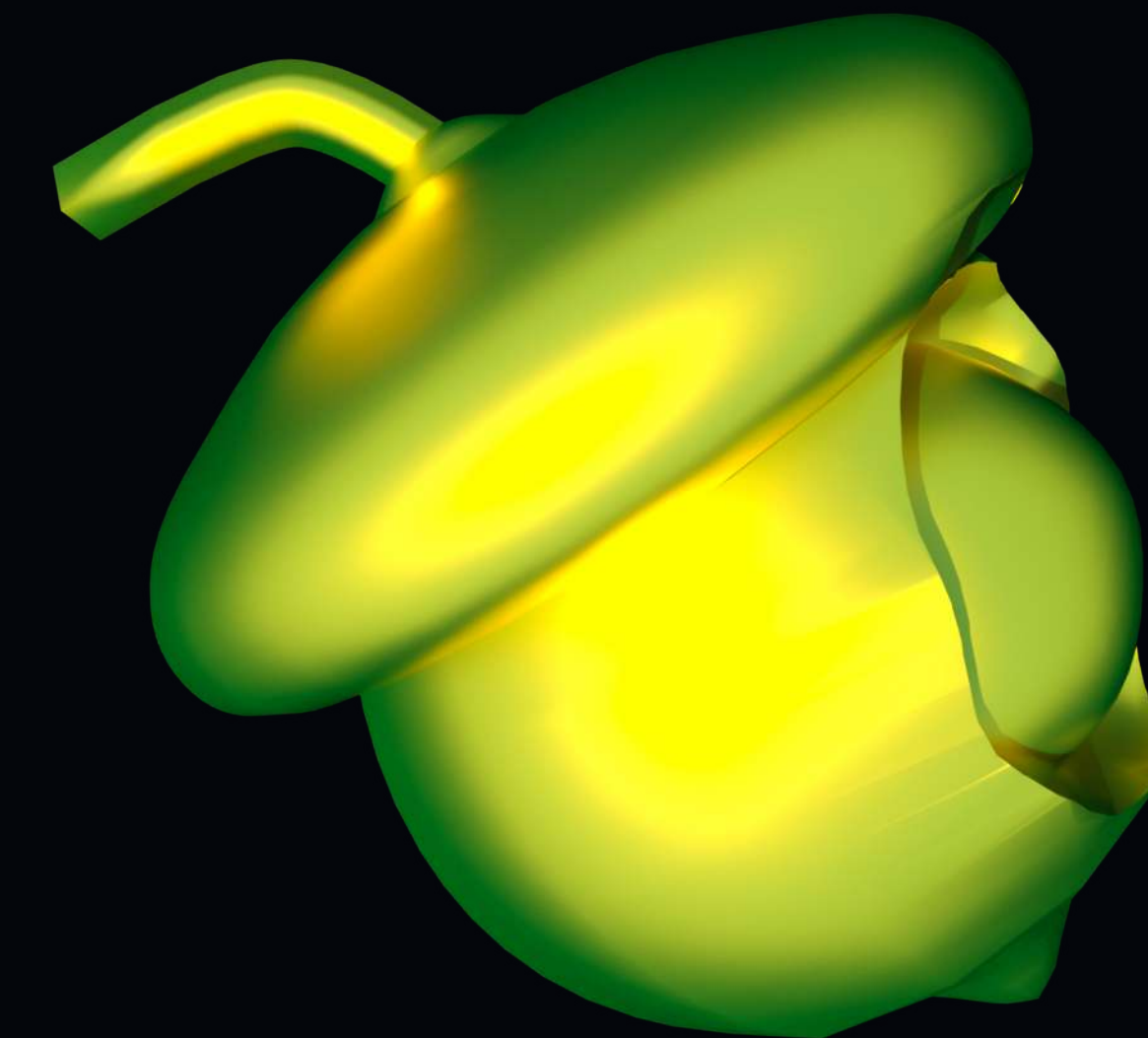
Flink Forward Asia



2019

Thanks!

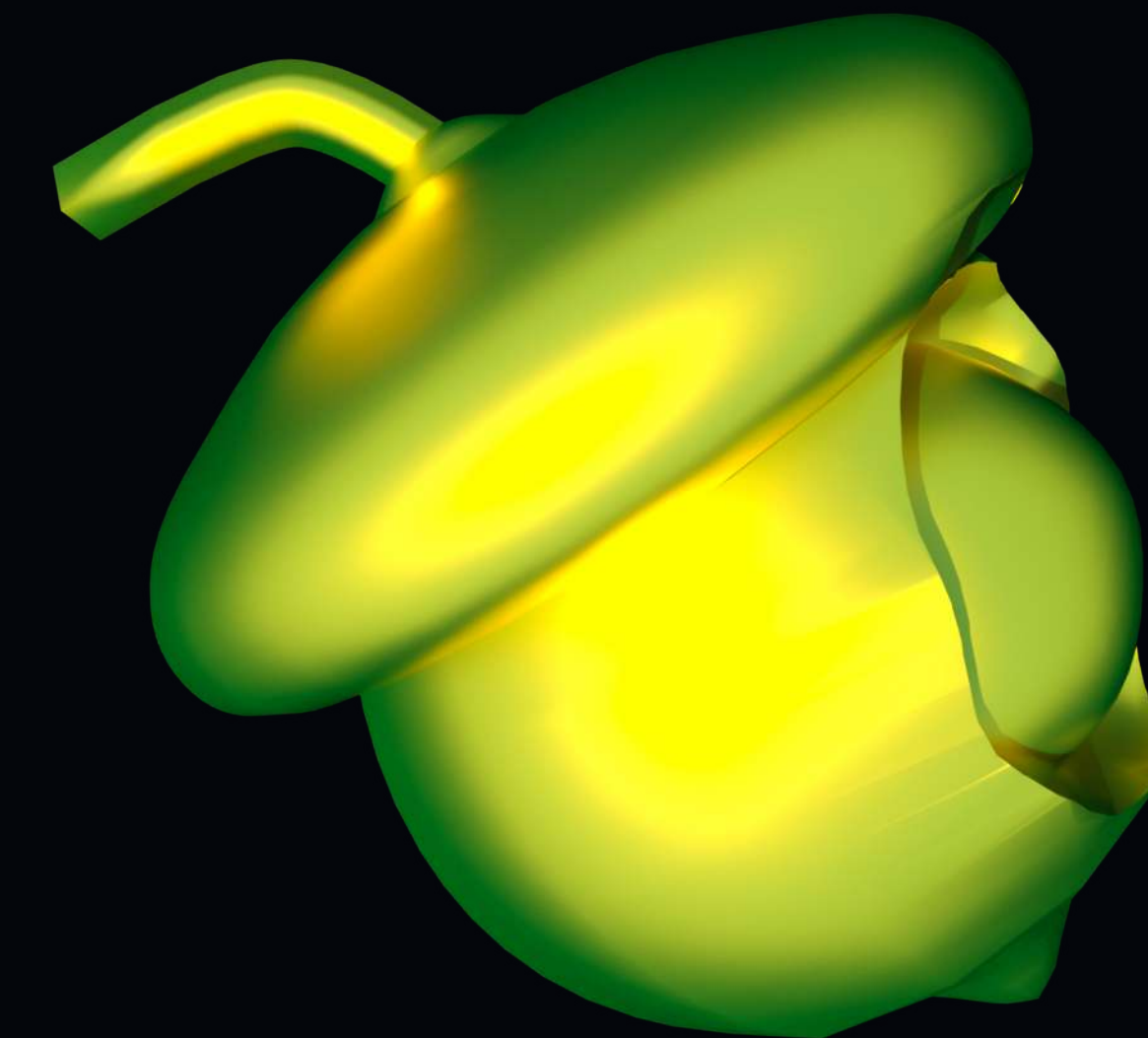
Flink Forward Asia



2019

Thanks!

Flink Forward Asia



2019

Thanks!