MEMSQL

PowerStream: Propelling Energy Innovation with Predictive Analytics

Nikita Shamgunov, CTO

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE

Topics

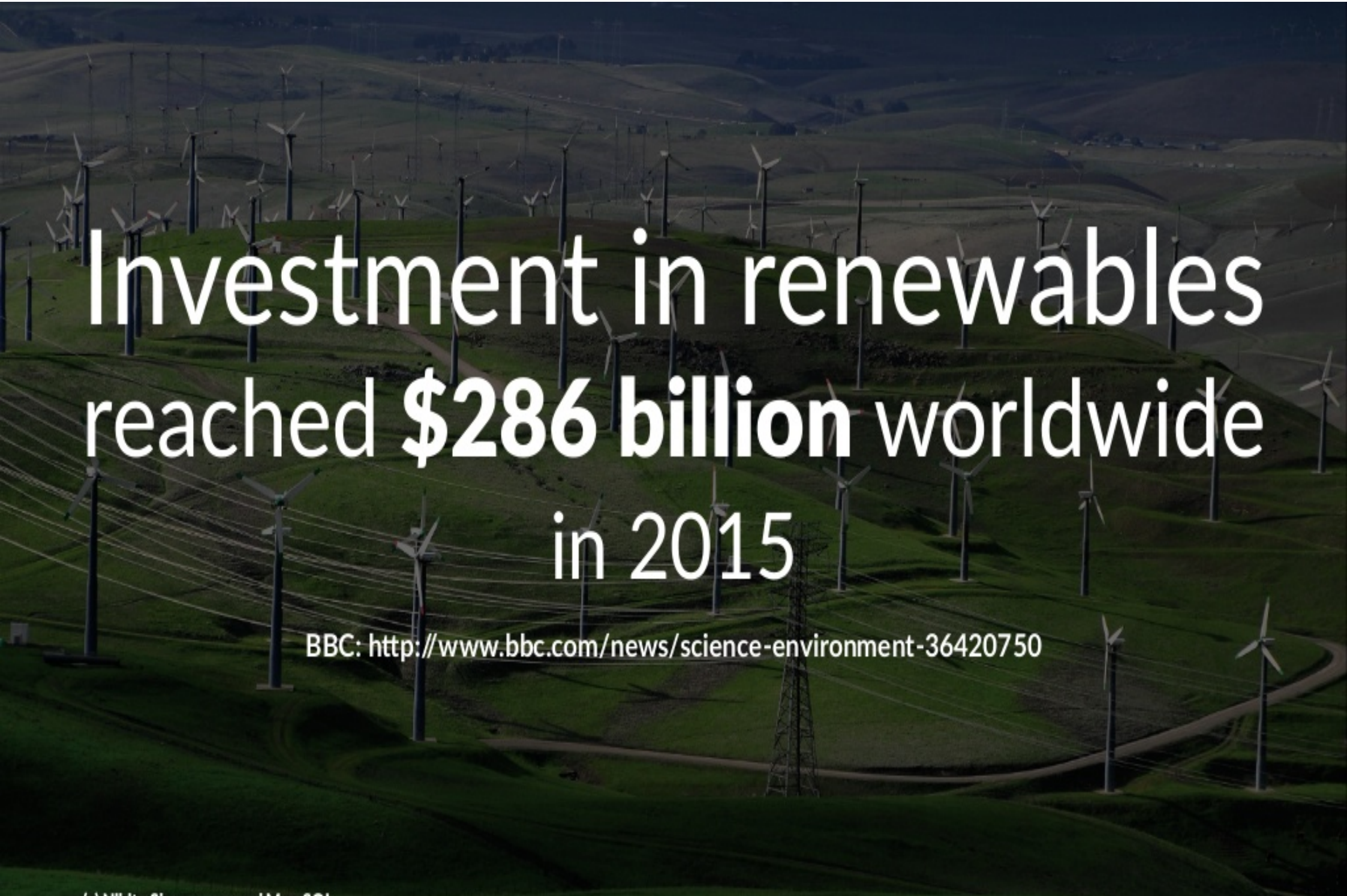
- Renewable Energy
- PowerStream
- Demo
- Q&A

Renewable Energy in the News

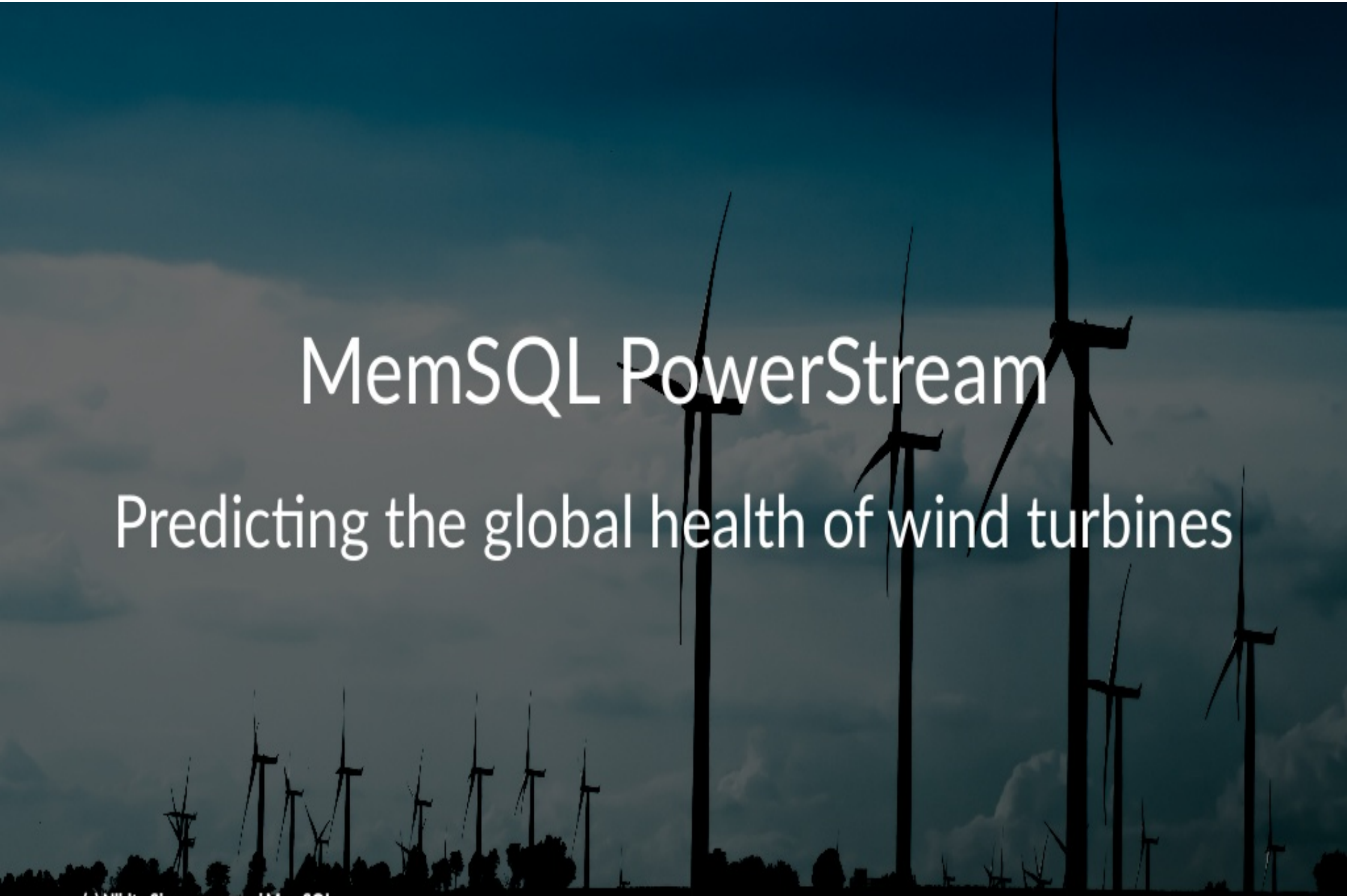
Germany Just Got Almost All of Its Power From Renewable Energy

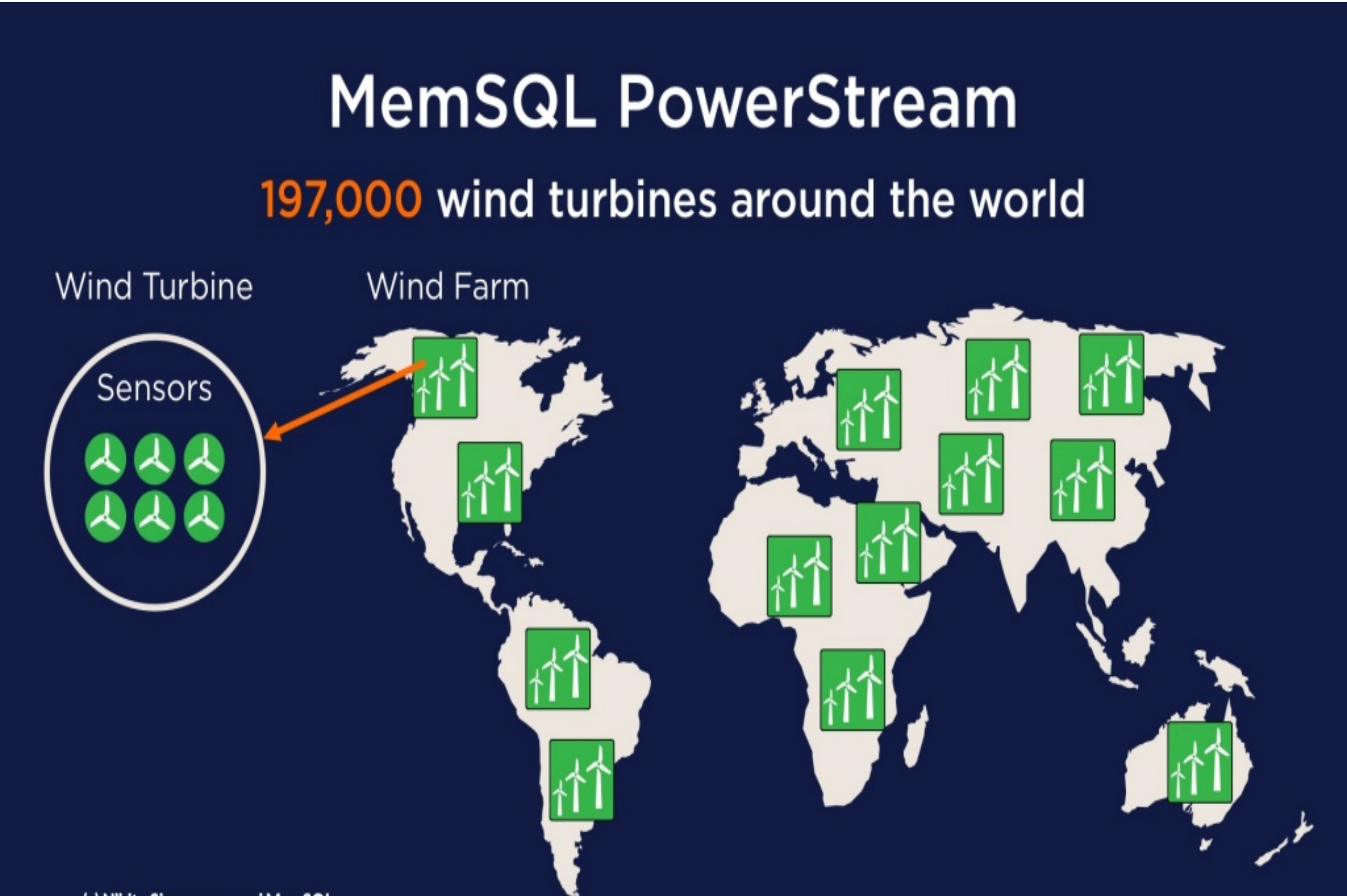
May 15, 2016

Bloomberg: http://www.bloomberg.com/news/articles/2016-05-16/germany-just-got-almost-all-of-its-power-from-renewable-energy


Investment in renewables
reached $286 billion worldwide
in 2015
BBC: http://www.bbc.com/news/science-environment-36420750

Introducing PowerStream


MemSQL PowerStream
Predicting the global health of wind turbines


MemSQL PowerStream
197,000 wind turbines around the world
Wind Turbine
Wind Farm
Sensors

1 to 2 million data points per second

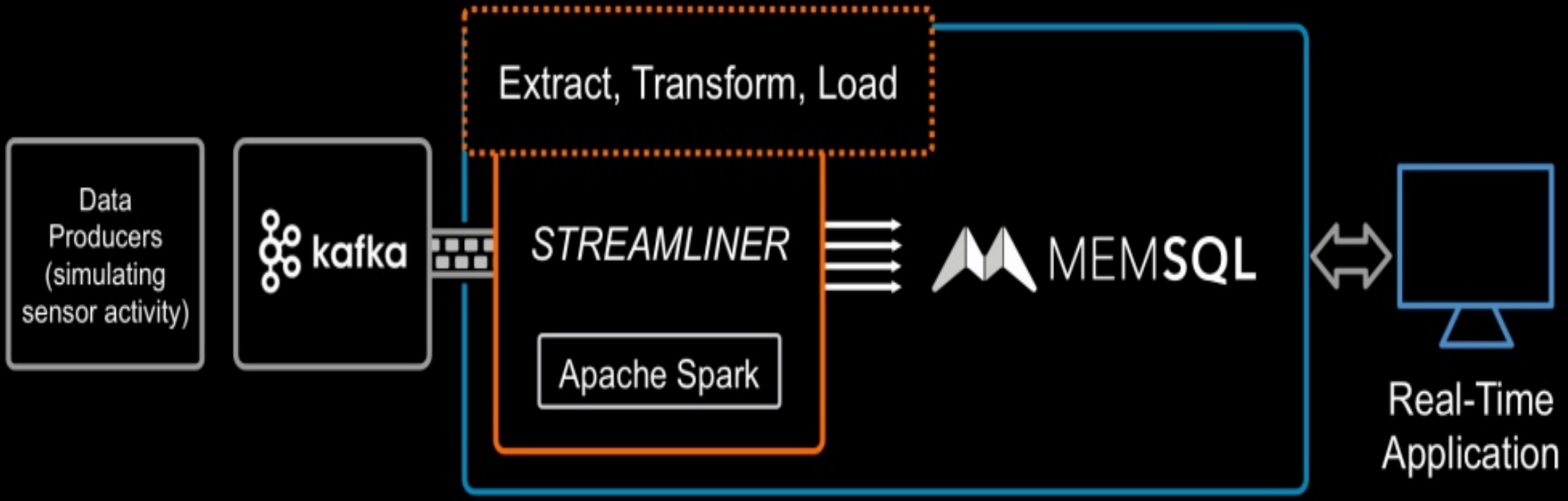
with MemSQL *Streamliner*

MemSQL PowerStream

Internet-of-Things simulation depicting health of wind turbines globally.

7 machines - AWS C4-2X large instances, at $0.311 per hour per machine, annual cost ~ $19,000.

Streamliner Architecture for PowerStream



MEMSQL | PowerStream

Visualizing data from **2 million sensors** and **197,000 wind turbines** from around the world. All data generated in real time by MemSQL.

Writes / Second	1.1M
Reads / Second	198K

Windfarm Sensor Breakdown

1K 0

Alerts

- Windfarm in Germany/Nordrhein-Westfalen/Steinfurt
- Windfarm in Germany/Sachsen-Anhalt
- Windfarm in Germany/Niedersachsen/Salzgitter
- Windfarm in Germany/Schleswig-Holstein/Wiemersdorf
- Windfarm in Germany/Niedersachsen
- Windfarm in Germany/Sachsen-Anhalt/Aschersleben
- Windfarm in Germany/Sachsen
- Windfarm in Germany/Schleswig-Holstein
- Windfarm in Germany/Nordrhein-Westfalen/Rüthen
- Windfarm in Germany/Baden-Württemberg

MEMSQL | PowerStream

Visualizing data from **2 million sensors** and **197,000 wind turbines** from around the world. All data generated in real time by MemSQL.

Writes / Second 2M

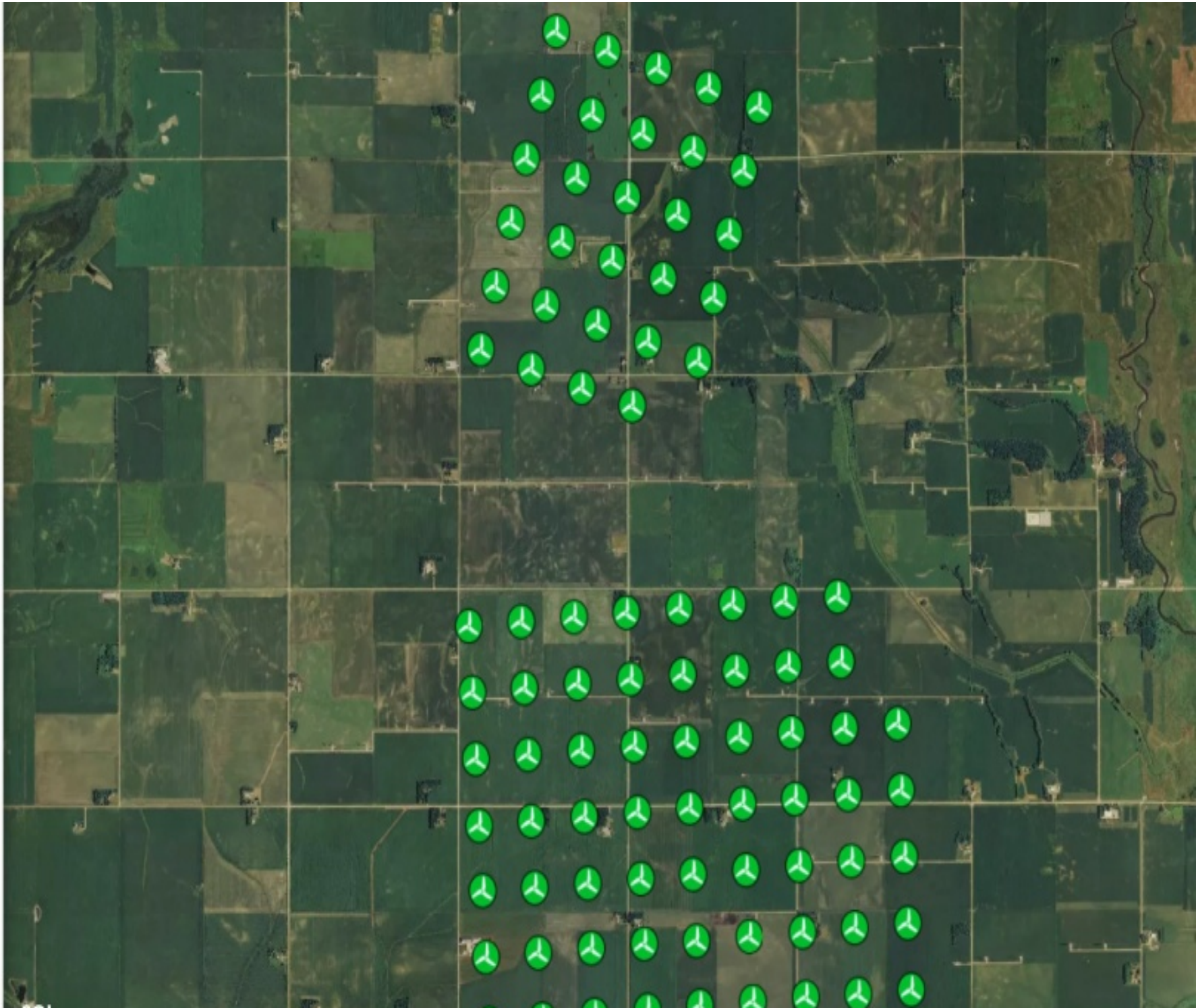
Reads / Second 2K

Turbine Sensor Breakdown

90 0

Alerts

- Windfarm in Germany/Nordrhein-Westfalen/Steinfurt
- Windfarm in Germany/Sachsen-Anhalt
- Windfarm in Germany/Niedersachsen/Salzgitter
- Windfarm in Germany/Schleswig-Holstein/Wiemersdorf
- Windfarm in Germany/Niedersachsen
- Windfarm in Germany/Sachsen-Anhalt/Aschersleben
- Windfarm in Germany/Sachsen
- Windfarm in Germany/Schleswig-Holstein
- Windfarm in Germany/Nordrhein-Westfalen/Rüthen
- Windfarm in Germany/Baden-Württemberg

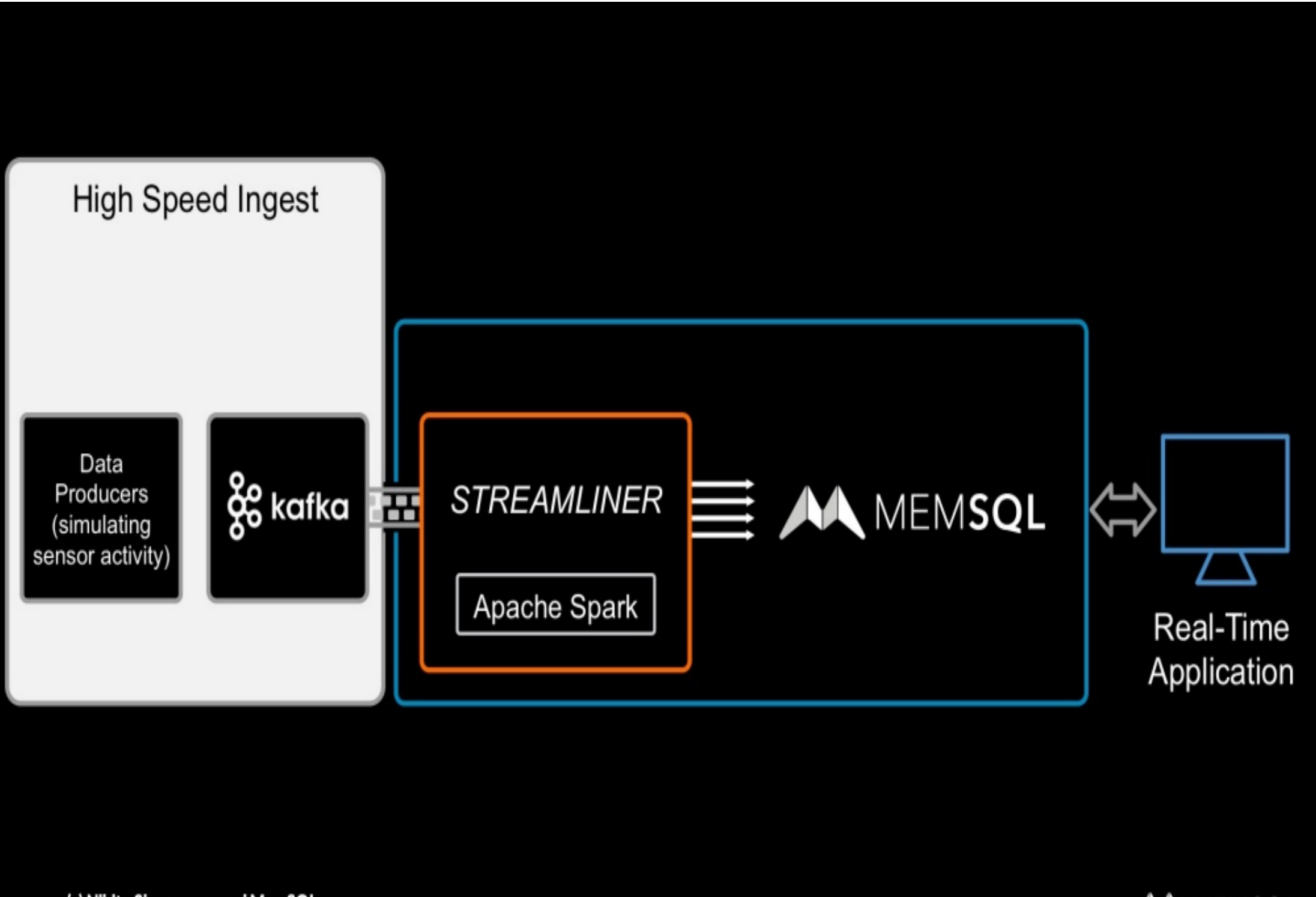


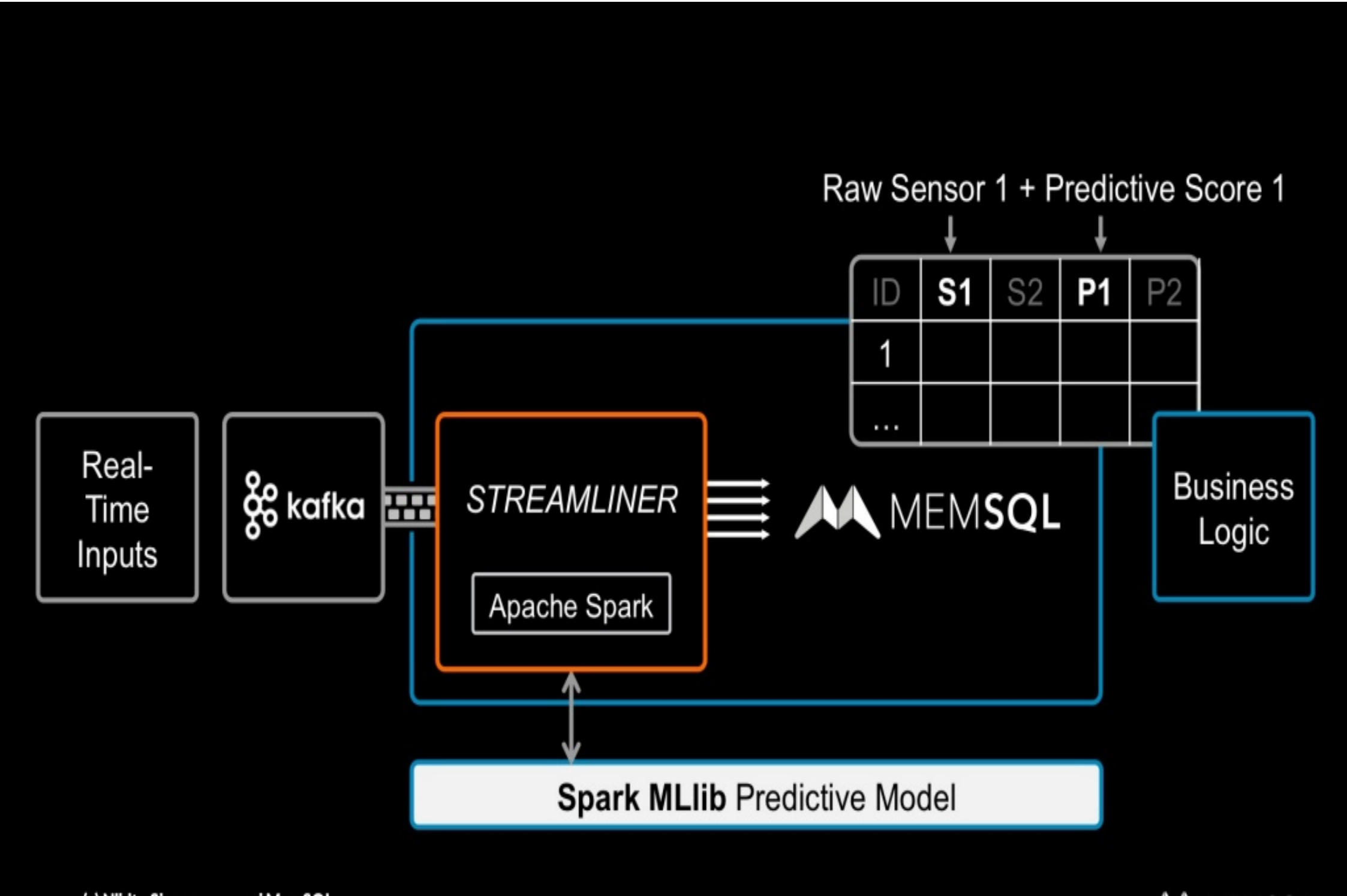
Demo Sequence

- Part I
  - High Speed Ingest
  - Predictive Analytics
  - Business Intelligence
- Part II
  - Spark SQL Pushdowns

MemSQL and Spark Real-time application

Demo part I

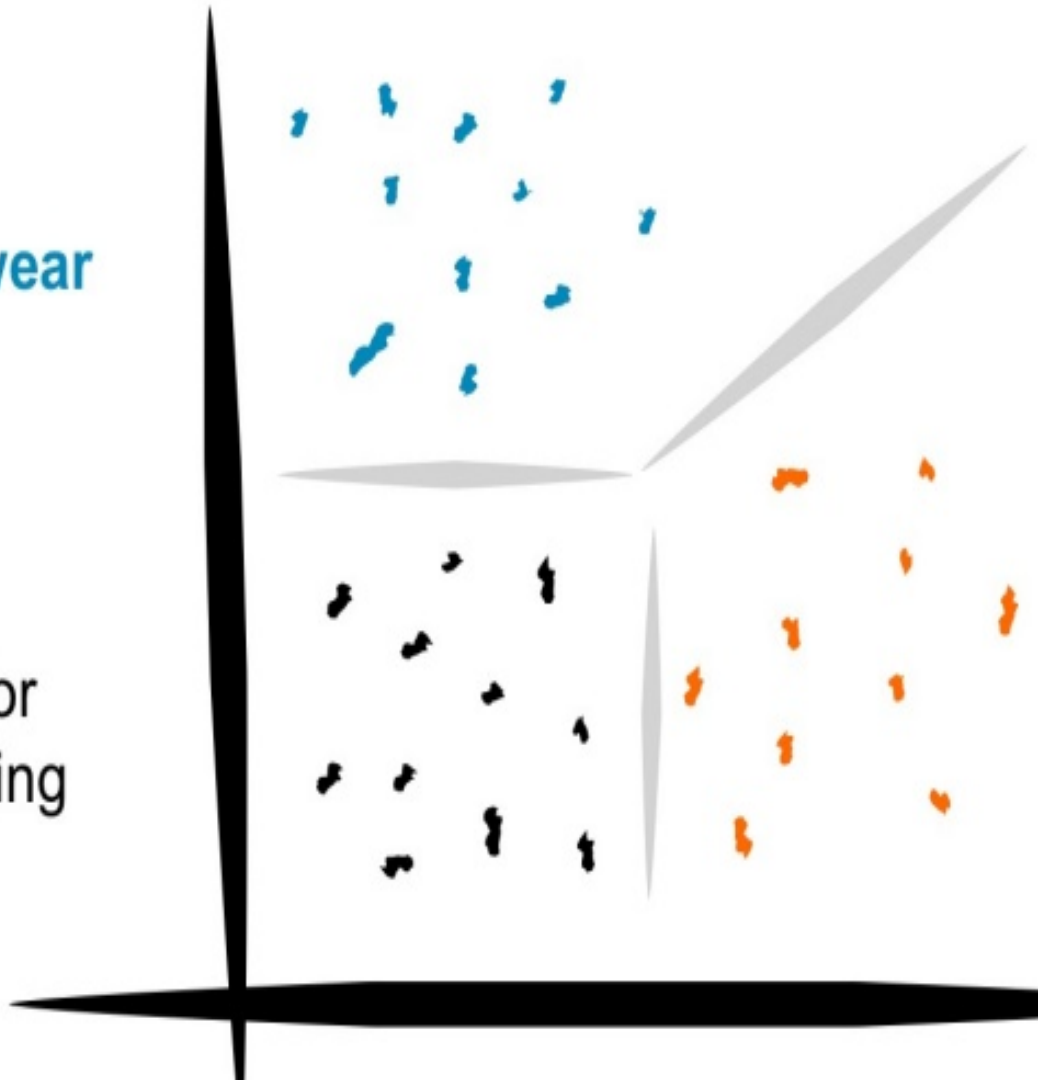

High Speed Ingest
Data Producers (simulating sensor activity)
kafka
STREAMLINER
Apache Spark
MEMSQL
Real-Time Application


Raw Sensor 1 + Predictive Score 1
ID
S1
S2
P1
P2
1
...
Real-Time Inputs
kafka
STREAMLINER
Apache Spark
MEMSQL
Business Logic
Spark MLlib Predictive Model

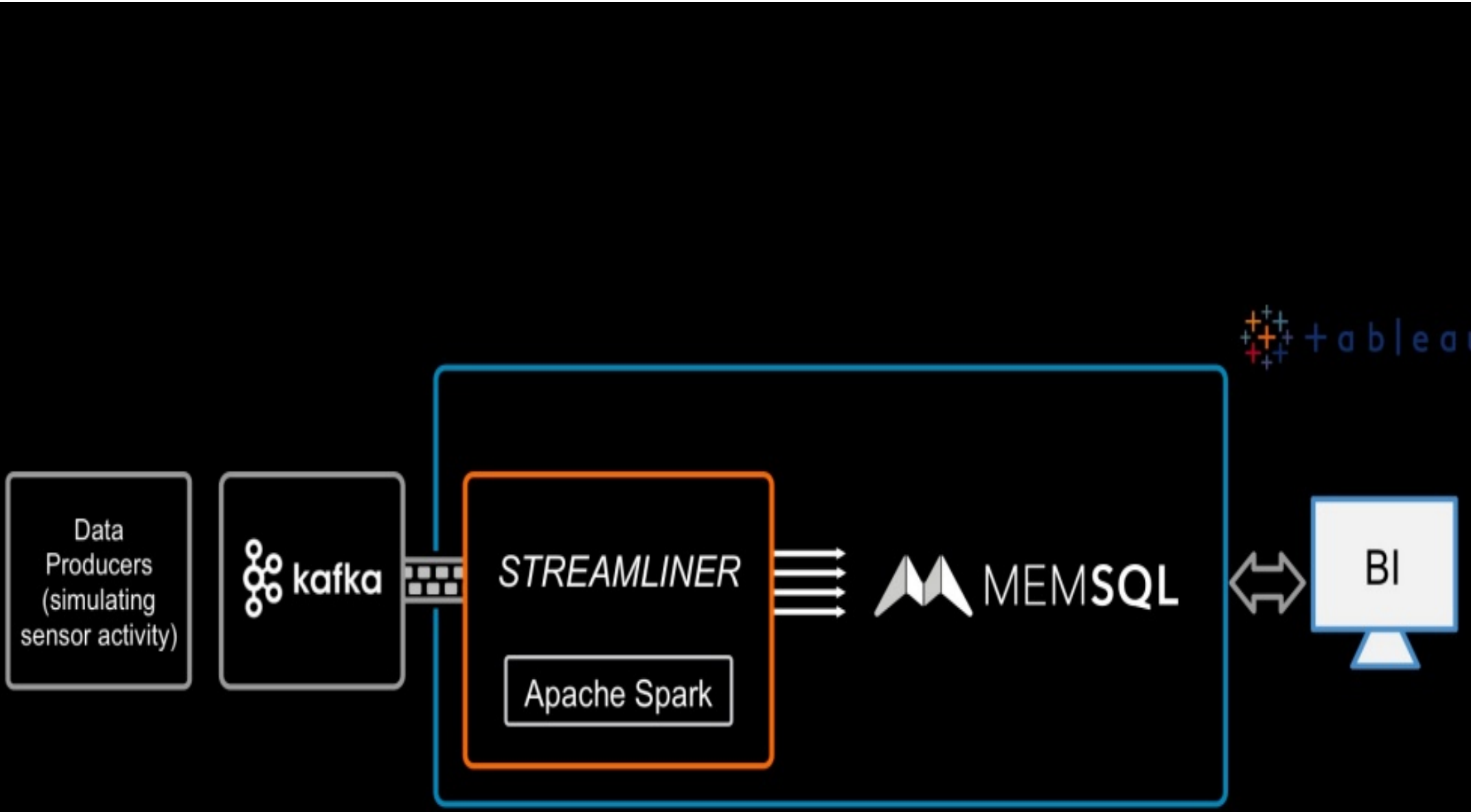
Classification

BLUE
Failure
Mechanical wear
Type 1

BLACK
training data for
turbine operating
normally



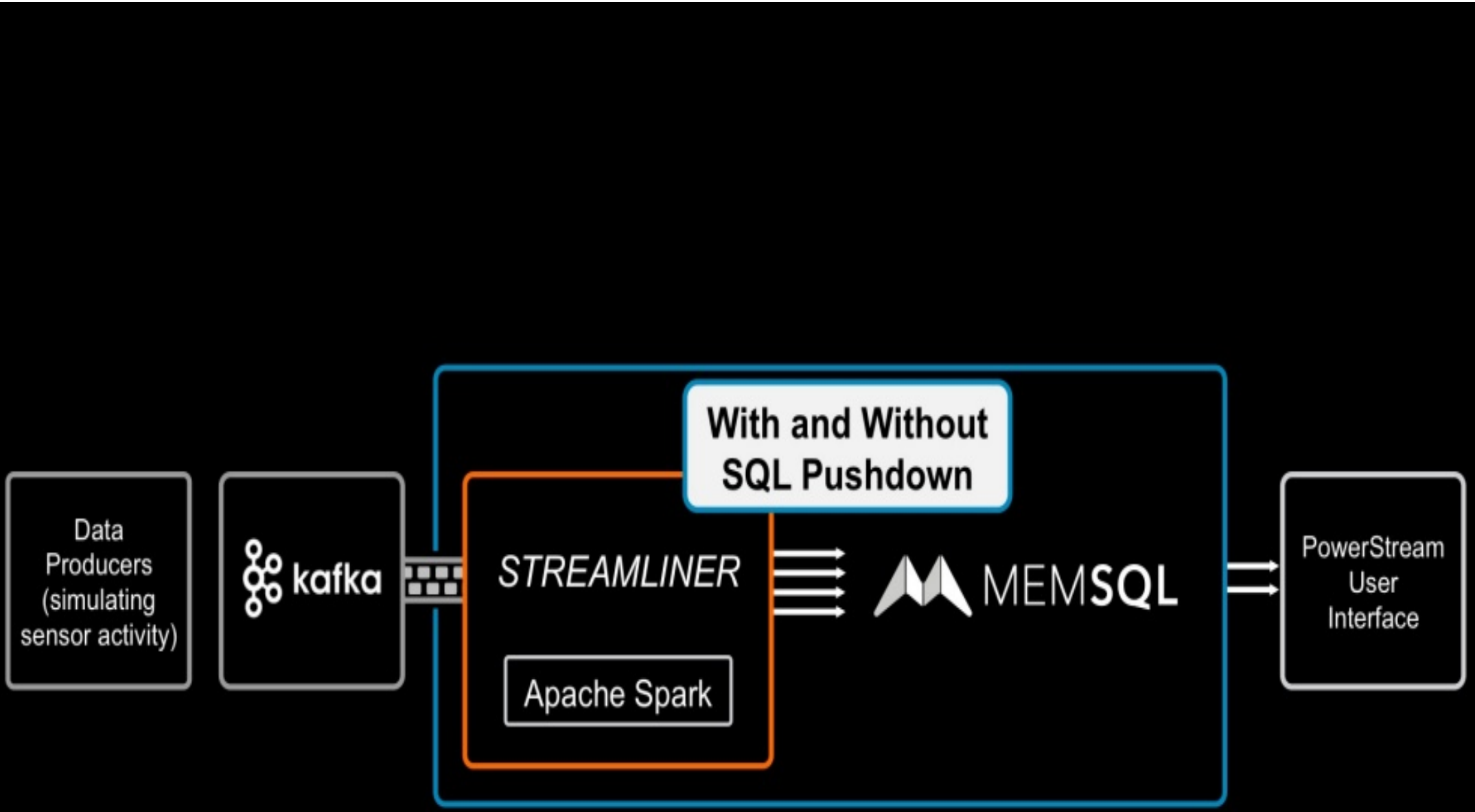
ORANGE
Failure
Mechanical wear
Type 2


Data Producers (simulating sensor activity)
kafka
STREAMLINER
Apache Spark
MEMSQL
BI
tableau

Demo Part I

MemSQL and Spark
Better together

Demo part II


Data Producers (simulating sensor activity)
kafka
STREAMLINER
Apache Spark
With and Without SQL Pushdown
MEMSQL
PowerStream User Interface

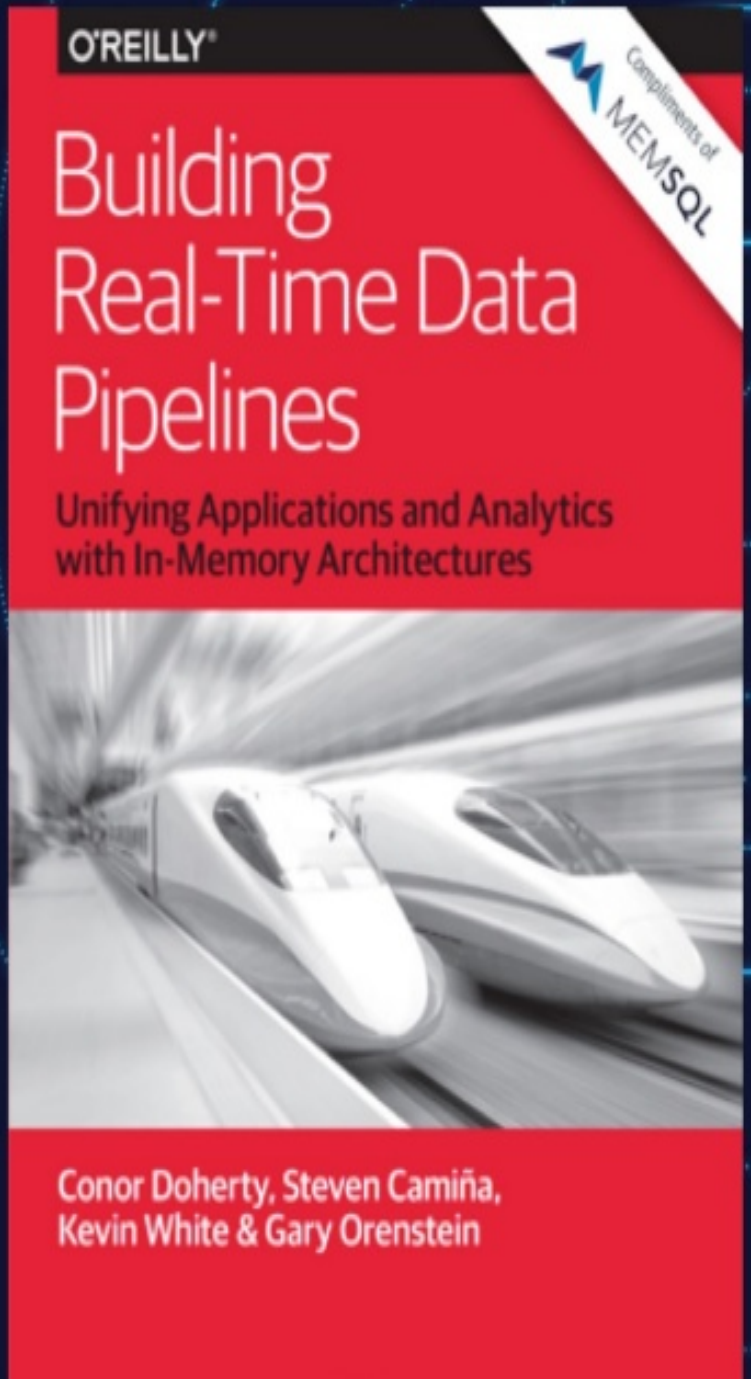
SQL Pushdown

- Delegate SQL query processing to a database
- Enhance speed and concurrency
- Complement other Spark capabilities
- Use Spark as a high level interface
- Command line example with and without pushdown

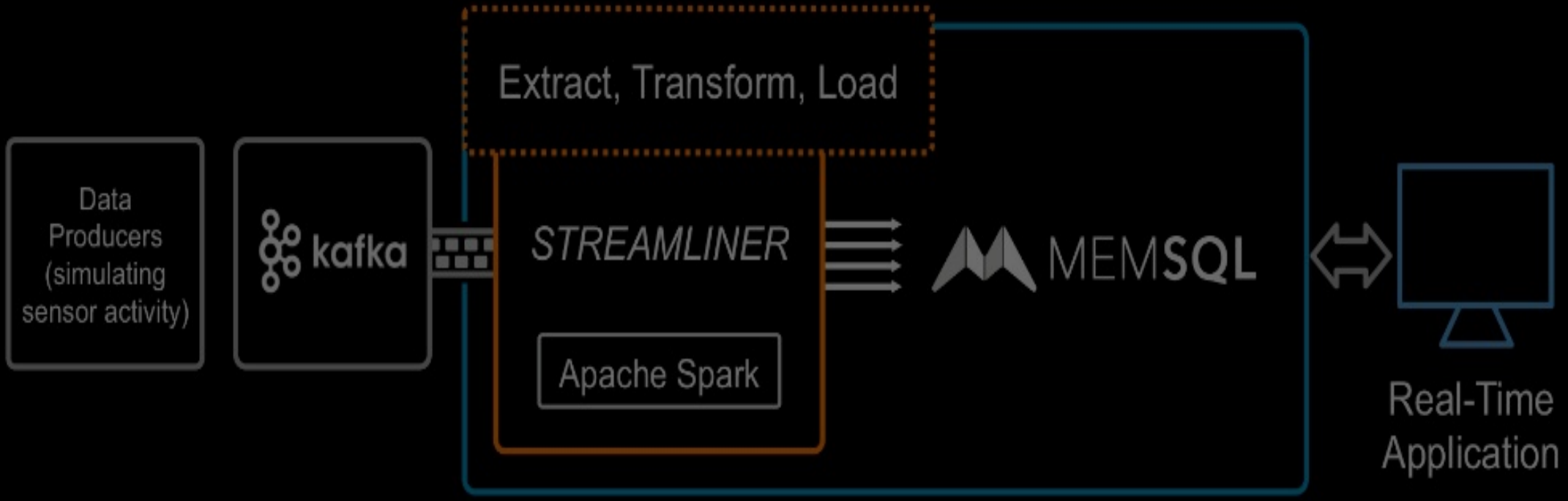
Demo Part II

From a Real-Time Dashboard to Predictive Applications

Get the blueprint



Questions and answers



MEMSQL

THANK YOU.

@nikitashamgunov

SPARK SUMMIT 2016
DATA SCIENCE AND ENGINEERING AT SCALE