

# HOW APACHE SPARK IS HELPING TAME THE WILD WEST OF WI-FI

Tomasz Magdanski

Director, Big Data and Analytics, iPass



# Who Are We?

iPass: the world's largest Wi-Fi network



**57M+**  
HOTSPOTS



**120+**  
COUNTRIES



**160+**  
NETWORK  
PROVIDERS

- Global operations, Silicon Valley headquarters
- On Nasdaq since 2003
- 40+ patents
- 800 of the Fortune 2000
- Launched iPass SmartConnect™ in Fall 2015

# Wi-Fi Is Unpredictable



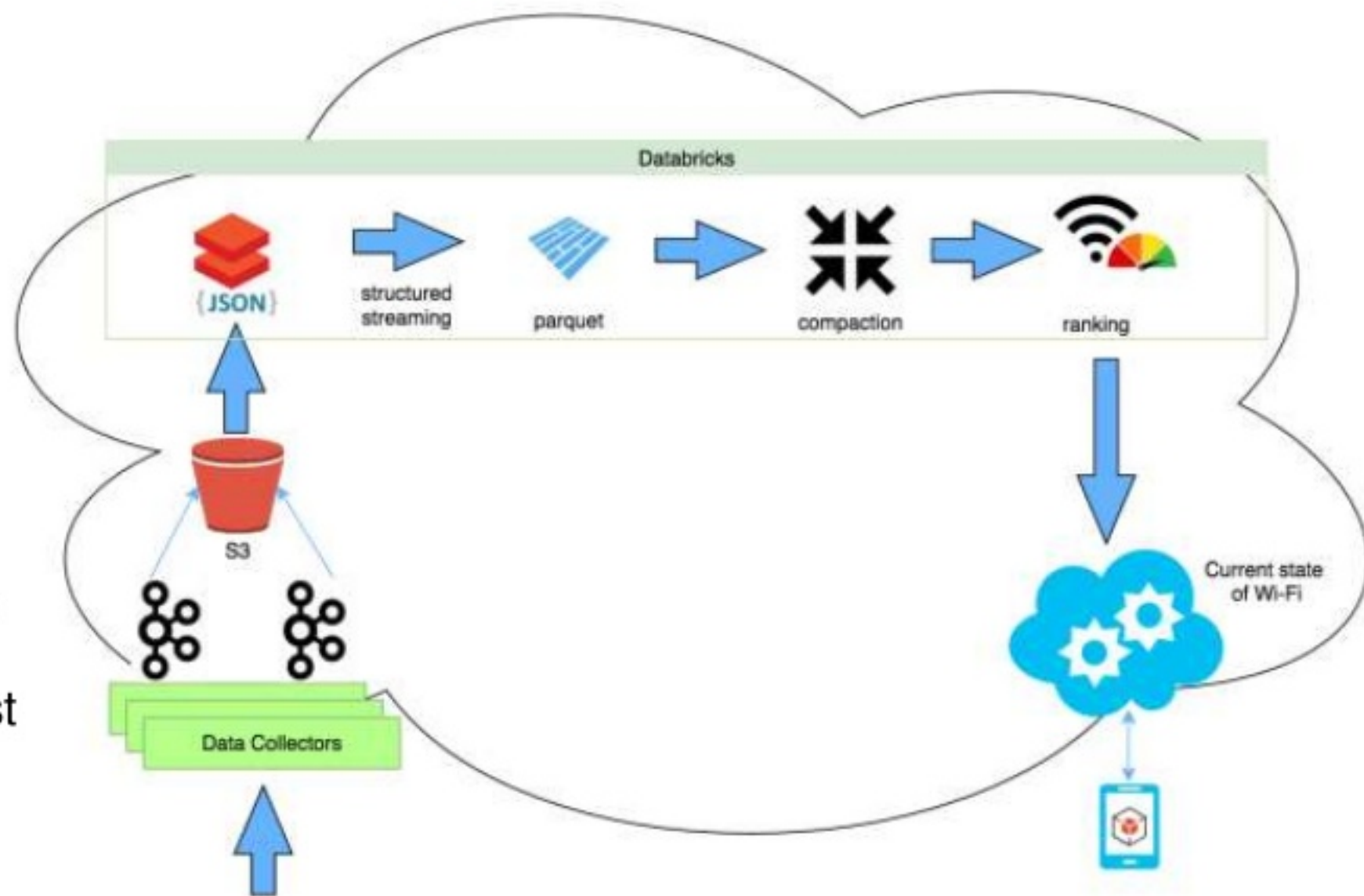
# So what's the solution?

# Spark & Databricks

- 21B scans -> 500M records -> 100M hotspots
- Spark helped us make sense of the data
- We needed a solution that can automatically scale and handle real time analytics

# Spark: From Concept to Production

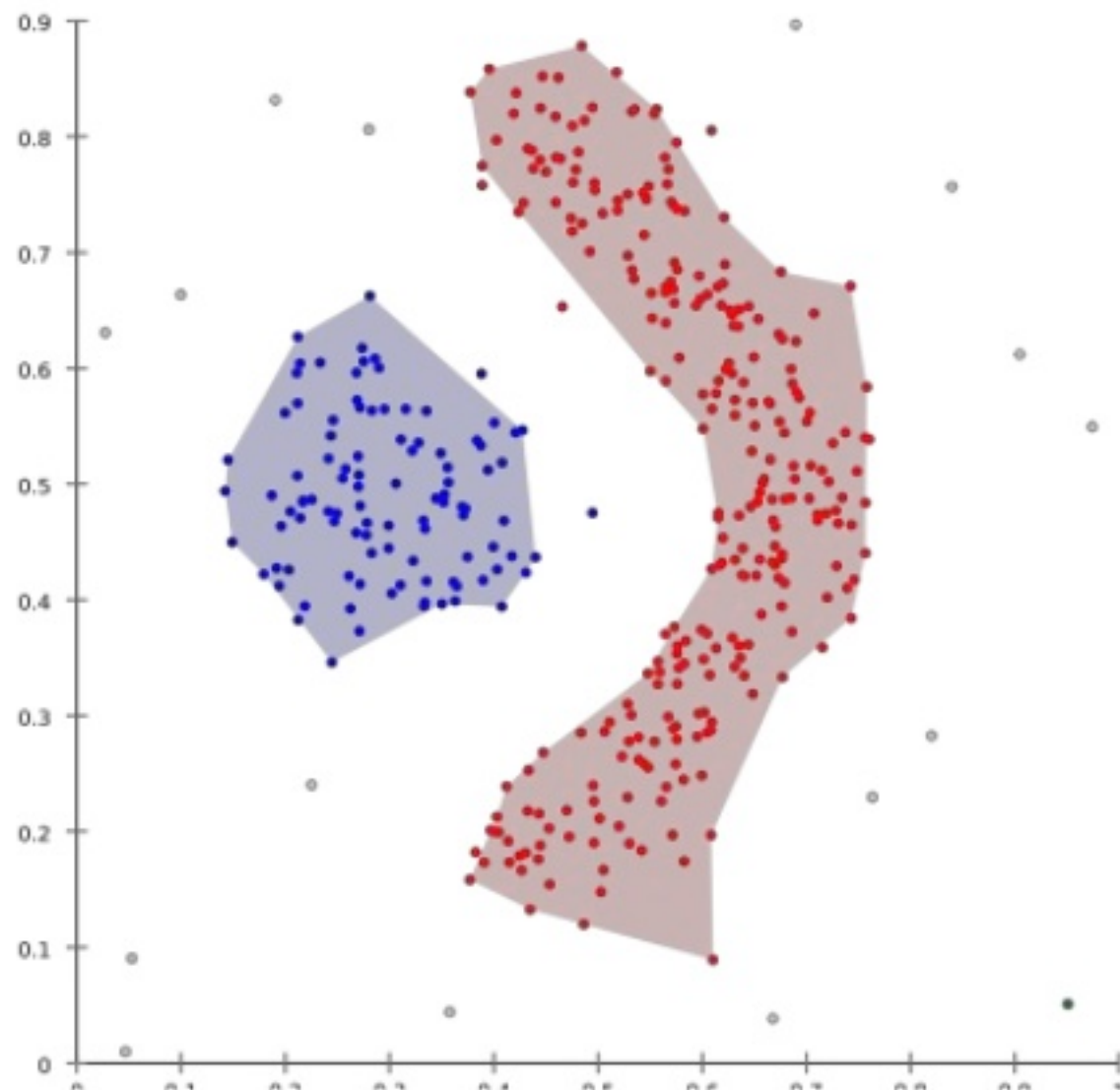
- Past:
  - in-house prototyping
  - Spark 1.3
  - RDD
- Present:
  - AWS and Databricks
  - Spark 2.0
  - Datasets
  - UDFs
  - Window aggregations
  - Full advantage of Tungsten and Catalyst





# Building Wi-Fi Network Characteristics

- Future:
  - Moving Hotspot
  - Changing SSID
  - Grouping and Graphframes - to find relationships
- Most of our code is written in Scala notebooks
- Ready to switch to structure streaming



# Conclusions

- Now we know
- Thanks to Databricks platform
  - Smaller team - big result
  - Focus on building scalable business logic, not infrastructure
- Small companies can successfully run big data projects without breaking the bank





# THANK YOU.

[tmagdanski@ipass.com](mailto:tmagdanski@ipass.com)

The logo for Spark Summit Europe 2016 features a stylized, circular graphic composed of many thin, radiating lines, resembling a sunburst or a stylized letter 'C'.

**SPARK SUMMIT**  
**EUROPE** 2016