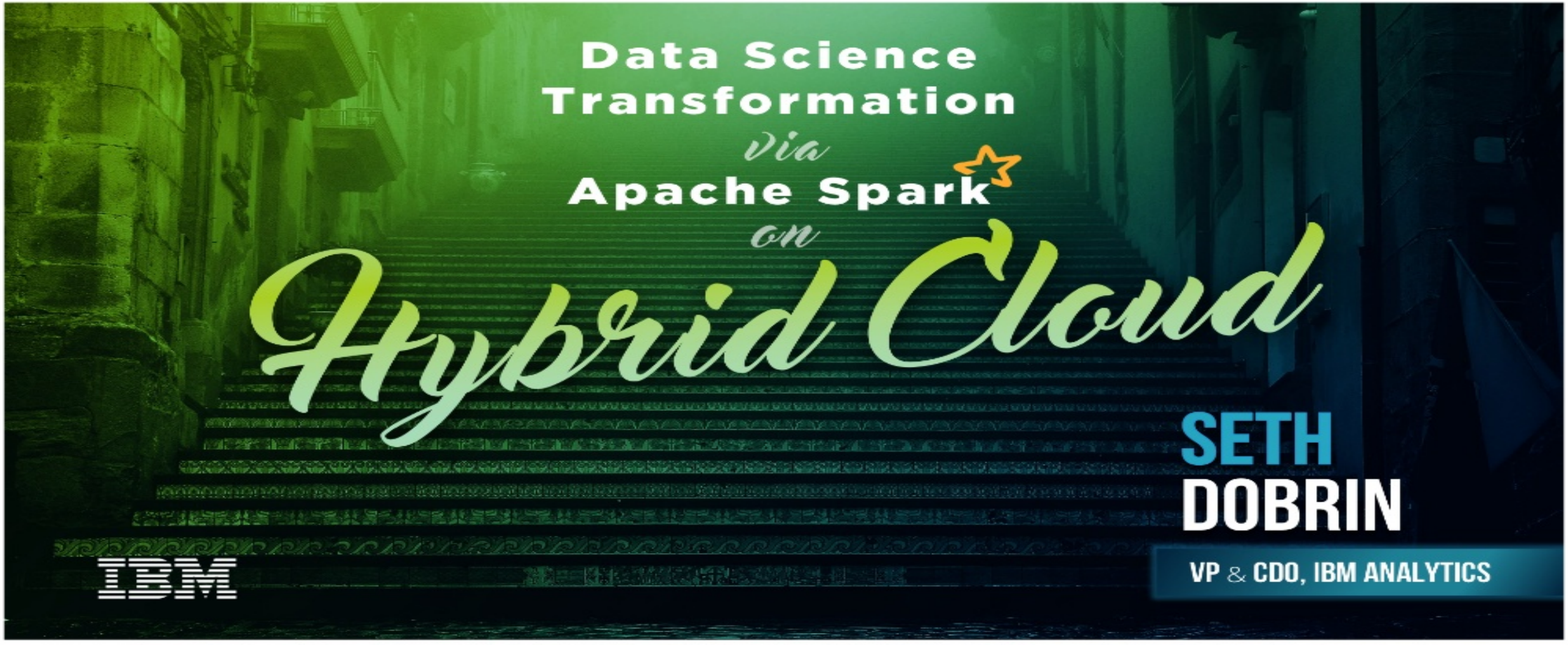
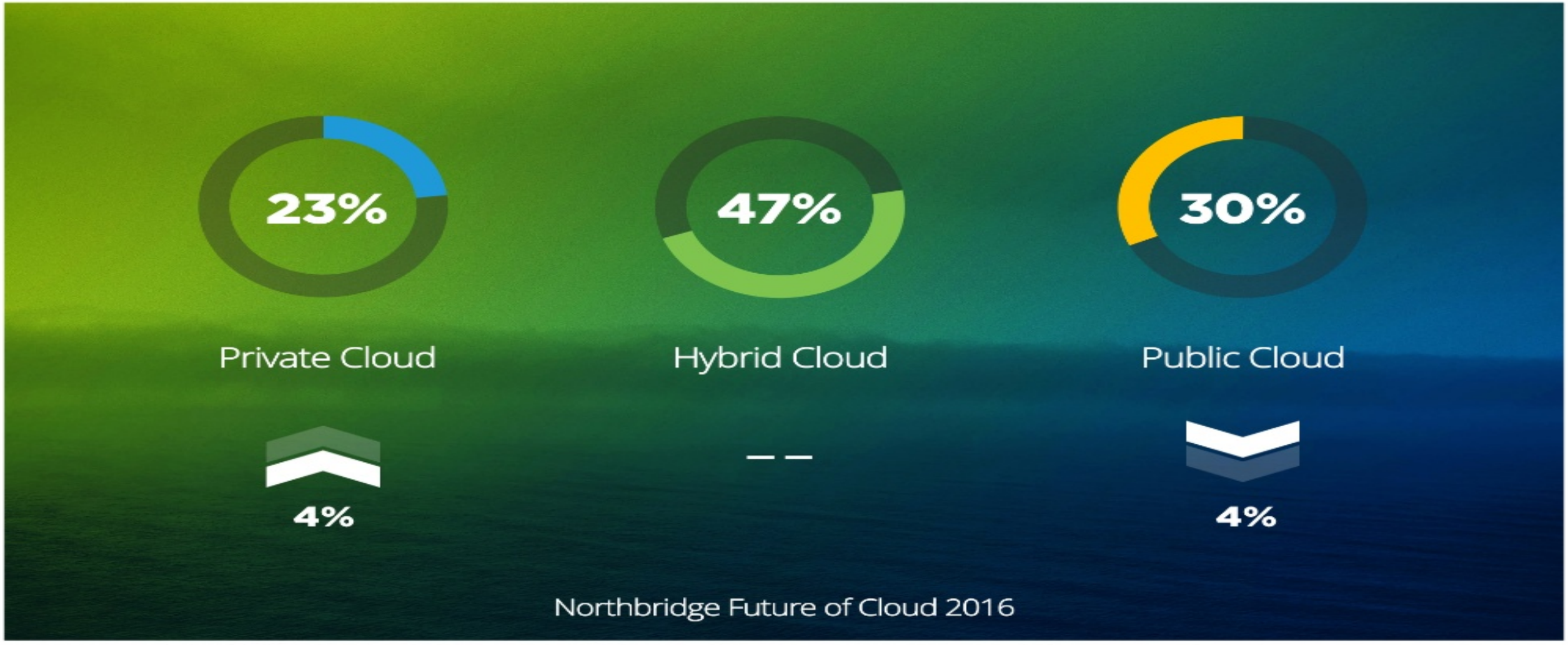
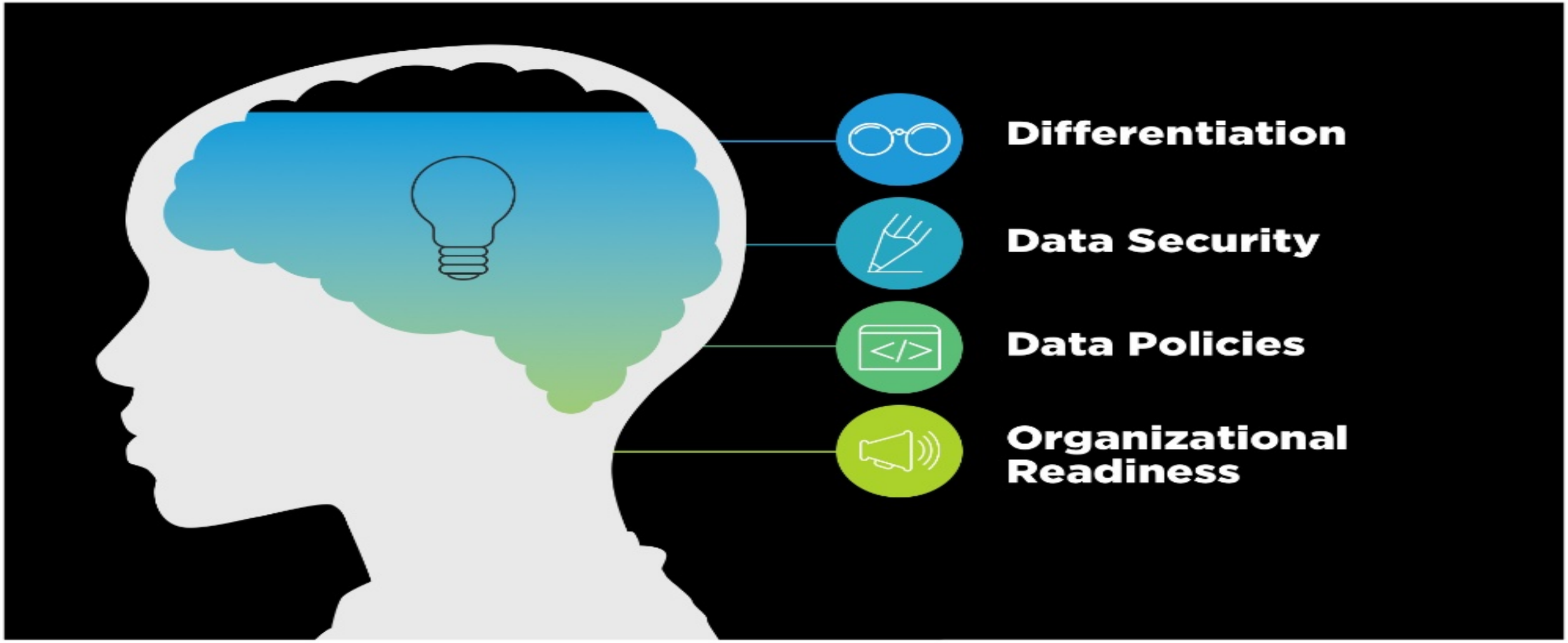
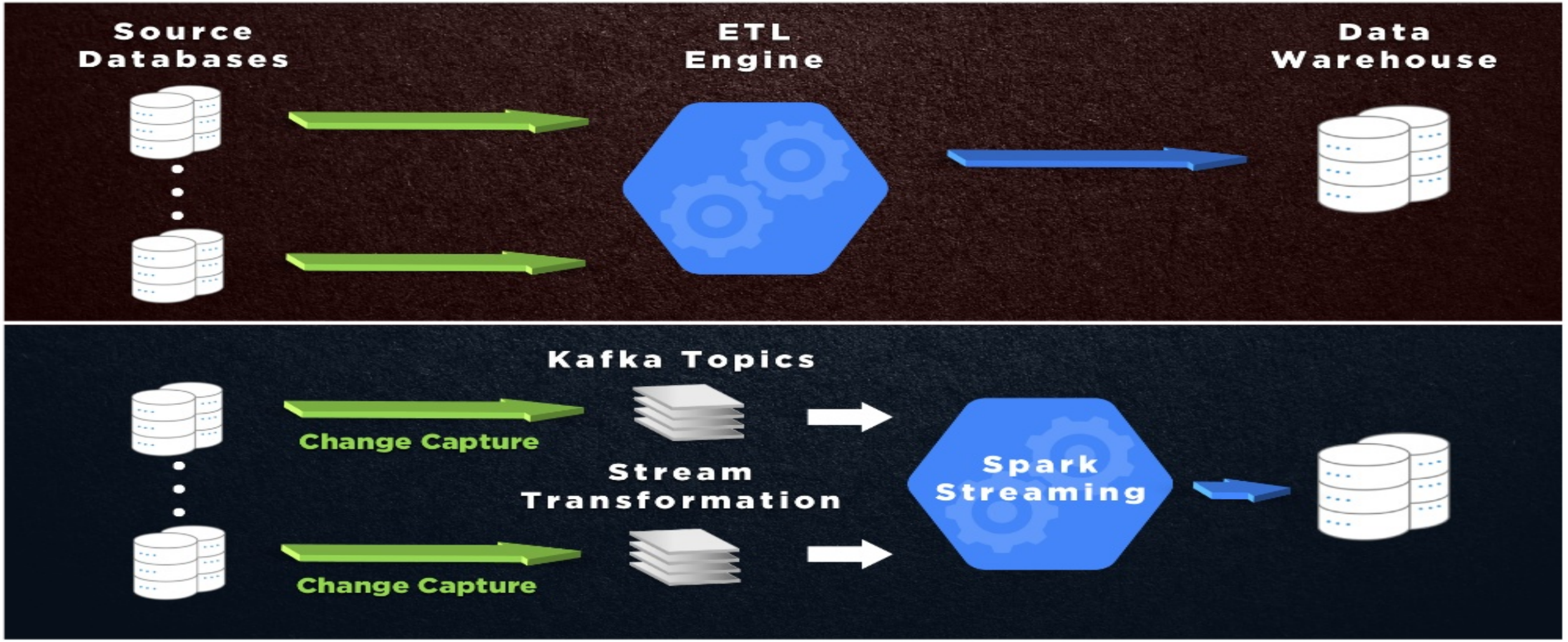

Data Science
Transformation
via
Apache Spark
on
Hybrid Cloud
SETH
DOBRIN
VP & CDO, IBM ANALYTICS
IBM


23%
47%
30%
Private Cloud
Hybrid Cloud
Public Cloud
4%
4%
Northbridge Future of Cloud 2016


Differentiation
Data Security
Data Policies
Organizational Readiness


Source Databases
ETL Engine
Data Warehouse
Kafka Topics
Change Capture
Stream Transformation
Change Capture
Spark Streaming

Spark + Kafka
Spark Streaming
Real-time Analytics
Apache Kafka
Spark ML, Analytics, Computation, Storage
Sensor Data
Apache Kafka

Sensor Data

Business Processes / Functions

Manufacturing

Smart Cities / Homes

Kafka

Bi-Directional

Redundancy

Sensor Integration

ETL Storage

Spark

High-Perf Hybrid Cloud

Real-Time Analytics

Semi-Supervised Learning

ML Algorithms

SparkR
HydraR
PySpark
SparkML
SystemML
Apache Spark
2016

Visit Us Today
booth #301
datascience.ibm.com
spark.tc/sethdobrin
SETH
DOBRIN
@sdobrin