Using Spark and Riak for IoT apps

Patterns and Anti-patterns

Pavel Hardak
Basho Technologies

SPARK SUMMIT EAST 2017

IOT & INDUSTRY VERTICALS

Manufacturing

35% of manufacturers already use smart sensors, 10% plan to implement them within a year, and 8% plan to implement them within three years, according to PwC.

Oil, gas, and mining

We estimate 5.4 million IoT devices will be used on oil extraction sites by 2020. The devices will primarily be internet-connected sensors used to provide environmental metrics about extraction sites.

Transportation

Connected cars are a top IoT device. We estimate there will be over 220 million connected cars on the road by 2020.

Insurance

74% of insurance executives said they believe the IoT will disrupt insurance within the next five years, and 74% plan to invest in developing and implementing IoT strategies by 2016, according to an SMA Research survey.

Defense

We estimate spending on drones will reach $8.7 billion in 2020. In addition, 126,000 military robots will be shipped in 2020, according to Frost & Sullivan.

Connected Home

By 2030, we expect the majority of home devices shipped will be connected to the internet due to initiatives from device makers to connect everything they produce.

Agriculture

We estimate 75 million IoT devices will be shipped for agricultural uses in 2020, at a 20% CAGR. These devices are primary sensors placed in soil to track acidity levels, temperature, and variables that help farmers increase crop yields.

Food Services

We estimate 310 million IoT devices will be used by food services companies by 2020. The majority of these devices will be digital signs connected throughout grocery stores and fast-food companies.

Infrastructure

We estimate municipalities worldwide will increase their spending on IoT systems at a 30% CAGR, from $36 billion in 2014 to $133 billion in 2019. This investment will generate$421 billion in economic value for cities worldwide in 2019.

Utilities

Energy companies throughout the world are trying to meet the rising demand in energy. To do this, they will be installing nearly 1 billion smart meters by 2020.

Retail

Beacons, paired with mobile apps, are being used in stores to monitor customer behavior and push advertisements to customers. In the US, we estimate $44.4 billion will be generated from beacon-triggered messages.

Hospitality

31% of hotels use next-generation door locks, 33% have room control devices, 16% have connected TVs, and 15% use beacons throughout the hotel, according to Hospitality Technology's 2015 Lodging Technology survey.

Logistics

Tracking sensors placed on parcels and shipping containers will help reduce costs associated with lost or damaged goods. In addition, robots, such as the Amazon Kiva robot, help reduce labor costs in warehouses.

Healthcare

We estimate 646 million IoT devices will be used for healthcare by 2020. Connected healthcare devices can collect data, automate processes, and more. But these devices can also be hacked, thereby posing a threat to the patients who rely on them.

Banks

There are nearly 3 million ATMs installed globally in 2015, according the World Bank. Some teller-assist ATMs provide a live-stream video of a teller for added customer support.

Smart Buildings

43% of building managers in the US believe the IoT will affect how they run their building within the next two to three years, according to a survey from Daintree Networks.

IoT market - growth prediction

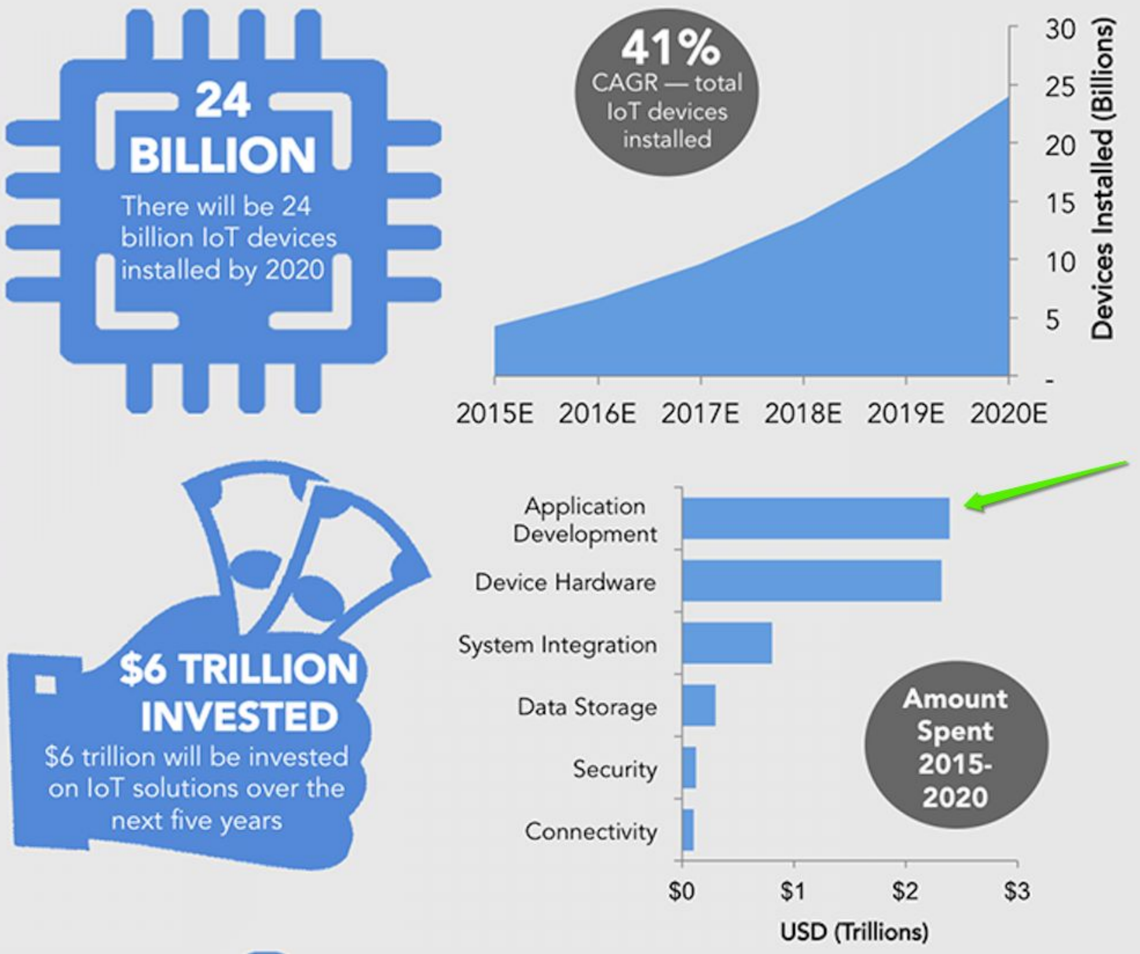
Number of connected "things"

- 2016 – about 6.4 B
  - 30% YoY growth, 5.5M activations per day
- 2020 – about 21 B

"By 2020 **more than half** of new **major business processes** and **systems** will incorporate some element of **Internet of Things**"

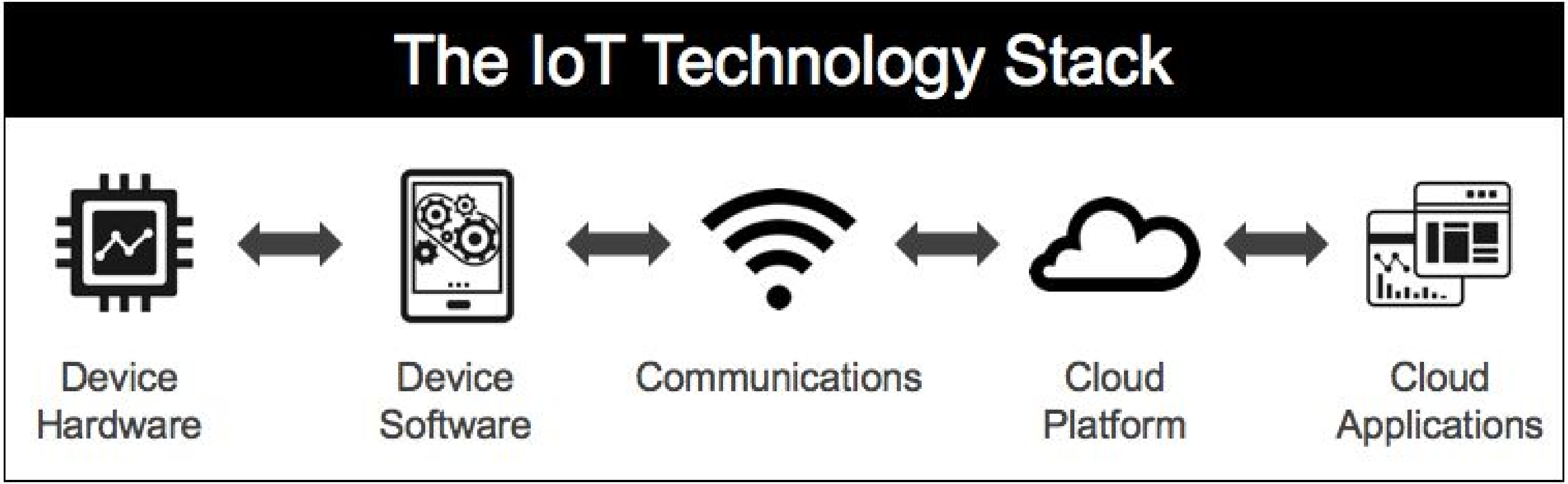
Gartner®

Sizing The Market



SPARK SUMMIT EAST 2017

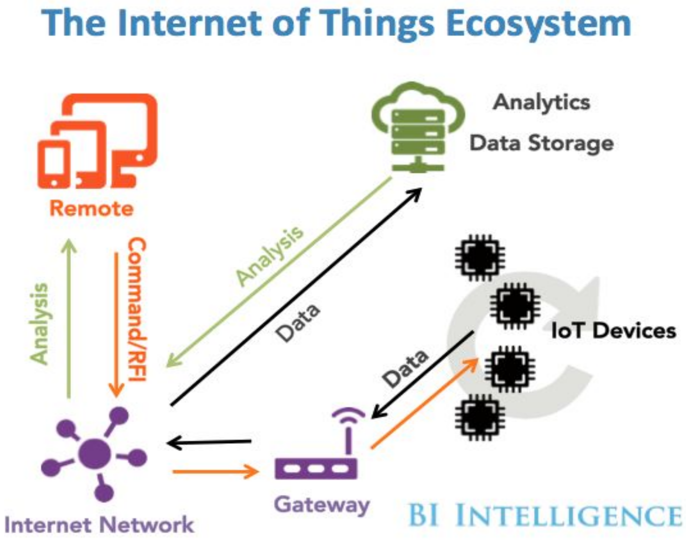
BI INTELLIGENCE
Intelligence.businessinsider.com



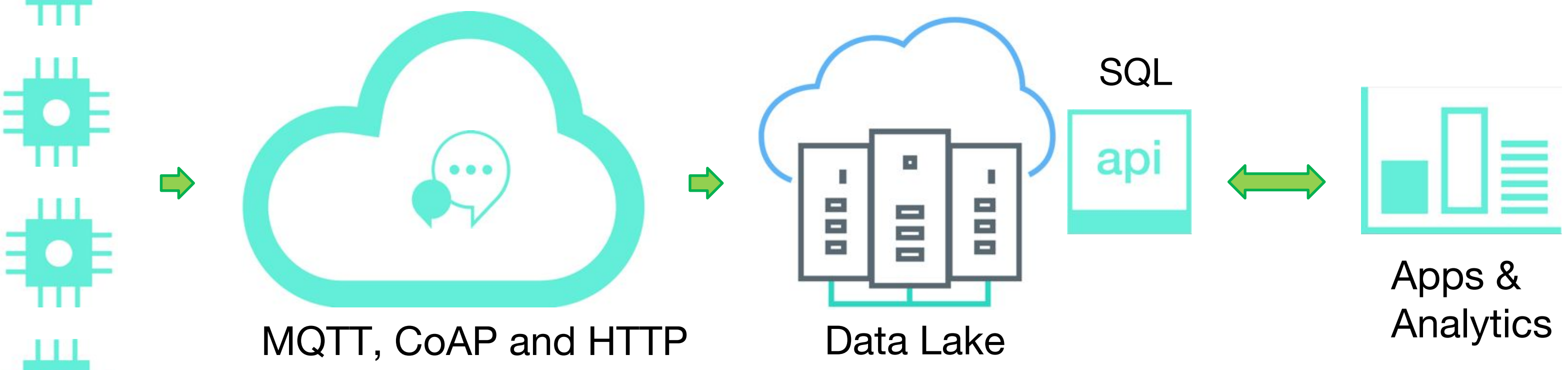
We want to be here!

IoT Project Plan

- Investigate those “things” and figure out
  - What protocols they support (CoAP, MQTT, HTTP, ...)
  - What data they generate (temperature, humidity, location, speed, ...)
- Collect this data in our data center
  - Implement protocols and parsing routines
  - Store into persistent storage (“Data Lake” architecture)
- Once stored in Data Lake
  - Analyze, summarize, “slice and dice”
  - Predict, make recommendations, discover insights
- Declare a victory (make profit, go for IPO, ...)



REFERENCE ARCHITECTURE (?)



Not so fast, my friend.

What is wrong with “Data Lake” for IoT ?

SPARK SUMMIT EAST 2017

SPARK SUMMIT EAST 2017

What is ~~different~~ special about IoT?
It is about the “things”... and more.

SPARK SUMMIT EAST 2017

fitbit

Application Class / Type	No. of Connections 2025	Requirement
CCTV, mHealth, Electronic billboards, automotive infotainment	**2 Bn** Wired, WiFi, Cellular	>10Mb/s Not cost sensitive, fixed power
Telematics, smart home, M2M backhaul	**8 Bn** Wired, WiFi, Cellular, Satellite	**<1Mb/s Low cost (<$15 UE), fixed power or regularly rechargeable battery**
Sensors, meters, wearables, 'thing' tracking, assisted living, logistics	**20 Bn** WiFi, Zigbee, Bluetooth, PLC, sub-GHz License-Exempt (e.g. M-Bus, etc.)	**<10Kb/s Very low-cost (<$5 UE), Ultra Low power (>10yrs battery life)**

IoT Networks and Protocols

2G
3G

4G

Lte

Z WAVE

ZigBee

SIGFOX

Bluetooth

LoRa

Wi Fi

IoT Devices & IoT Network Protocols

- Wireless technologies
- Limited range
- Limited bandwidth
- Shared transmission media
- Mesh or Ad-hoc Topology
- Possible signals interference

- Low cost hardware components
- Low power radio transmitters
- Very small antennas
- “Custom-made” firmware
- Constrained Application Protocol (CoAP)
- “Best Effort” QoS (“shoot and forget”)

IoT is “Big Data” - by definition.
Actually, lots and lots of Big Data.

IoT Data Categories

	Category	Description
Metadata & Profiles	Devices	Device info (model, SN, firmware, sensors, ..), configuration, owner, …
	Users	Personal info, preferences, billing info, registered devices, …
Time Series	Ingested ("Raw")	Measurements, statuses and events from devices.
	Aggregated ("Derived")	Calculated data - from devices & profiles • Rollups – aggregate metrics from low resolution to higher ones (min - hour – day) using min, max, avg, ... • Aggregations – aggregate measurements, configuration and profiles (model, region, …) over time ranges

Five “V”s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries: user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*

Five “V”s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries, user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*
Variety	Sensors data (time series), users and devices profiles, also time series “derived” data (e.g. rollups, aggregations).

Five "V"s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries, user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*
Variety	Sensors data (time series), users and devices profiles, also time series "derived" data (e.g. rollups, aggregations).
Volume	Starts small, grows quickly, keeps coming 24x7x365 (nights, weekends and holidays). Spikes up on new model launches or successful marketing campaign. Can slow down, but will keep growing. *Efficient data retention policy is critical to prevent overflows.*

Five "V"s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries, user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*
Variety	Sensors data (time series), users and devices profiles, also time series "derived" data (e.g. rollups, aggregations).
Volume	Starts small, grows quickly, keeps coming 24x7x365 (nights, weekends and holidays). Spikes up on new model launches or successful marketing campaign. Can slow down, but will keep growing. *Efficient data retention policy is critical to prevent overflows.*
Veracity	Generally trustworthy, but beware of "low cost" sensors with low accuracy. Sent over not-so-reliable transport - expect that some data will be corrupted or arrive late or might be lost. *(Hopefully the devices were not hijacked or impersonated by hackers)*

Five "V"s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries, user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*
Variety	Sensors data (time series), users and devices profiles, also time series "derived" data (e.g. rollups, aggregations).
Volume	Starts small, grows quickly, keeps coming 24x7x365 (nights, weekends and holidays). Spikes up on new model launches or successful marketing campaign. Can slow down, but will keep growing. *Efficient data retention policy is critical to prevent overflows.*
Veracity	Generally trustworthy, but beware of "low cost" sensors with low accuracy. Sent over not-so-reliable transport - expect that some data will be corrupted or arrive late or might be lost. *(Hopefully the devices were not hijacked or impersonated by hackers)*
Value	Profiles and summaries are much more valuable than raw data samples. The value of "raw" time series quickly goes down after it was processed and clock advanced. Aggregated ("derived") data are more valuable than raw data. *Exceptions: financial transactions, life support, nuclear plants, oil rigs, …*

Five “V”s	IoT data
Velocity	Torrent of small writes (sensors). Reads – millions of low-latency queries, user and device profiles, range queries for TS data (slices). Stream of updates (profiles) - *beware of conflicts.*
Variety	Sensors data (time series), users and devices profiles, also time series “derived” data (e.g. rollups, aggregations).
Volume	Starts small, grows quickly, keeps coming 24x7x365 (nights, weekends and holidays). Spikes up on new model launches or successful marketing campaign. Can slow down, but will keep growing. *Efficient data retention policy is critical to prevent overflows.*
Veracity	Generally trustworthy, but beware of “low cost” sensors with low accuracy. Sent over not-so-reliable transport - expect that some data will be corrupted or arrive late or might be lost. *(Hopefully the devices were not hijacked or impersonated by hackers)*
Value	Profiles and summaries are much more valuable than raw data samples. The value of “raw” time series quickly goes down after it was processed and clock advanced. Aggregated (”derived”) data are more valuable than raw data. *Exceptions: financial transactions, life support, nuclear plants, oil rigs, …*
Complexity	Poly-structured using simple schemas and simple relations (usually implicit). Some data is treated as unstructured (”opaque”) for speed or flexibility. *Note: expect schema or structure changes without preliminary notice.*


Gartner.
Through 2018, 75% of Internet of Things projects will take up to twice as long as planned.
gartner.com/events
Source: Gartner
© 2016 Gartner, Inc. and/or its affiliates. All rights reserved.

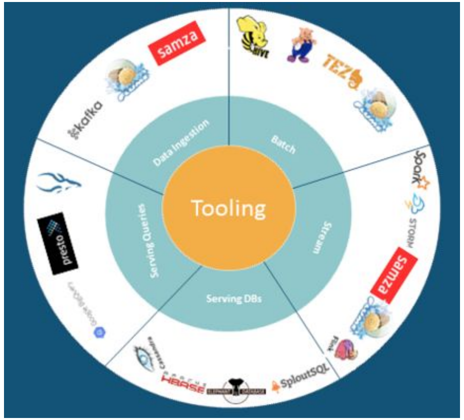
How are we going to solve it ?

IoT Data & Processing

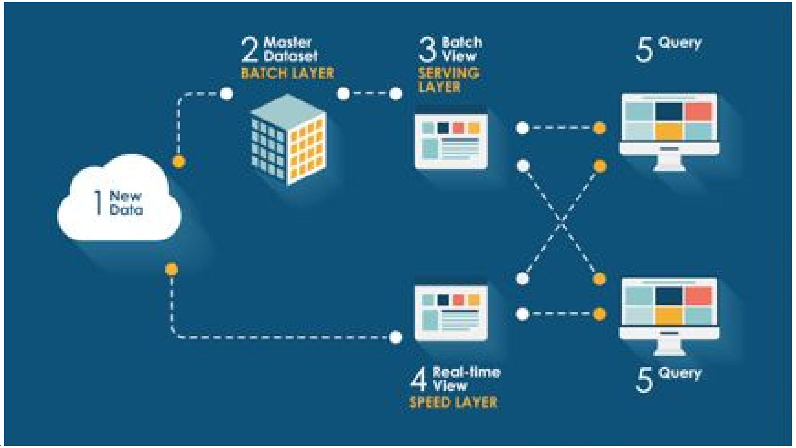
- Data
  - Huge amounts of data records - arriving 24x7x365
  - Some data records will arrive out-of-order, be late (minutes or hours) or lost
  - Expect “unexpected” - e.g. errors, nulls, schema or type changes, drops
- Processing
  - Preprocessing - validation and cleansing
  - Translation (format, type, version, ...) and enrichment
  - Aggregations - min, max, avg, sum, top or bottom N, percentile, ...
  - Grouping - device vendor and model, location, service, subscription type, ...
  - Rollups - from 10 sec raw samples to 1 min, 1 hour, 1 day, 1 week, 1 month, ..
  - Alarms (e.g. threshold crossing), anomaly detection (using ML)
  - Predefined reports (e.g. daily, weekly, ...)
  - Ad-hoc reports or exploratory queries
  - Insights, predictions, ...

Architectural Blueprints

- **Lambda Architecture** by *Nathan Marz* (ex-Twitter)
- **Kappa Architecture** by *Jay Kreps* (Confluent)
- **Zeta Architecture** by *Jim Scott* (MapR)
- ... and their variants



Zeta



Lambda

Kappa

Data Processing Framework for IoT

- Uses “Best of breed” OSS technologies
- Combines two paradigms
  - “Speed Layer” – pipeline for Stream Processing for “Data in Motion”
  - “Serving Layer” – analytics for “Data in Motion” and “Data at Rest”
- Every component is “Distributed by Design”
  - Collection Layer
  - Message Queue
  - Stream Processing
  - Data Storage (Database, Object System, Data Warehouse)
  - Query and Analytics Engines

Data store for IoT – "Wish list"

- **Ingested (Raw) Time Series**
  - Very high write throughput
  - Fast slice (time range) reads
- **Aggregated (Derived) Time Series**
  - Auto-distributed + slice locality
  - SQL-like queries
  - Aggregations
  - Bulk queries (analytics)
  - Secondary Indexes (Tags)
- **Efficient Storage**
  - Auto Data Retention (TTL)
  - Build-in anti entropy
  - Compression
  - Hot Backups
- **Profiles and Metadata**
  - Many concurrent reads with low latency
  - Reliable writes (ACID or conflict resolution)
  - Unstructured or partially structured
  - Secondary Indexes + Text Search
- **Scalability and Availability**
  - Distributed architecture, no SPoF
  - Linearly scalable - up and down
- **Operational simplicity**
  - Masterless architecture
  - Automatic rebalancing
  - Metrics, logs, events
  - Rolling upgrades

What DB type is a good fit for TS use cases?

Data Access Patterns

	Category	Description	R:W %
Metadata & Profiles	Devices & Users	Many low latency small reads - all over the dataset. Occasional updates – possibly by different “actors” (web, device, app), conflicts need to be prevented or resolved. Fewer creates and deletes.	90:10
Time Series			

Data Access Patterns

	Category	Description	R:W %
Metadata & Profiles	Devices & Users	Many low latency small reads - all over the dataset. Occasional updates – possibly by different "actors" (web, device, app), conflicts need to be prevented or resolved. Fewer creates and deletes.	90:10
Time Series	Ingested ("Raw")	Very high throughout of relatively small writes. Most reads are over recent time range "slice". Updates are rare (corrections). *This category is a biggest part of the IoT application dataset.*	10:90

Data Access Patterns

	Category	Description	R:W %
Metadata & Profiles	Devices & Users	Many low latency small reads - all over the dataset. Occasional updates – possibly by different "actors" (web, device, app), conflicts need to be prevented or resolved. Fewer creates and deletes.	90:10
Time Series	Ingested ("Raw")	Very high throughout of relatively small writes. Most reads are over recent time range "slice". Updates are rare (corrections). *This category is a biggest part of the IoT application dataset.*	10:90
	Aggregated ("Derived")	Mostly reads – users, platform services, reports. Writes are periodical on each time interval or from batch jobs.	80:20

Database Type For IoT or Time Series

Relational	Key Value	Document	Wide Column	Graph
MySQL	Riak KV	MongoDB	Cassandra	Neo4J
PostgreSQL	DynamoDB	CouchBase	HBase	Titan
Oracle	Voldemort	RethinkDB	Accumulo	Infinite Graph

None of existing DB types was designed to handle time series data

- Wide column DBs have high write throughput, but reads and updates are not their strength
- Key Value and Document DBs handle metadata well, but struggle with heavy writes and time-slicing reads
- Relational - good with metadata (unless number of updates is high), but a bad choice for TS data
- Graph DB – not a good choice for either time series or metadata, can be added later on

We need a new type of NoSQL database – Time Series

Database Type For IoT or Time Series

Relational	Key Value	Document	Wide Column	Graph
MySQL	Riak KV	MongoDB	Cassandra	Neo4J
PostgreSQL	DynamoDB	CouchBase	HBase	Titan
Oracle	Voldemort	RethinkDB	Accumulo	Infinite Graph

Time Series		
InfluxDB	Riak TS	Blueflood
KairosDB	Prometeus	Druid
OpenTSDB	Dalmatiner	Graphite

Iot Sensors Data – Hot to Cold

SENSORS DATA – HOT N’ COLD

Temp	Purpose	Description	Immutable?
Boiling Hot	App usage	Last known value(s) and/or for last N minutes, useful for immediate responses, very frequently accessed	No
Hot	Operational dataset	Last 24 hours to several days or weeks (rarely months), frequently accessed, dashboards and online analytics	Almost*
Warm	Historical data	Older data, less frequently accessed, used mostly for offline analytics and historical analysis	Yes
Cold	Archives	Used only in rare situations, kept in long term storage for regulatory or unpredicted purposes	Yes

STORAGE TIERS – FROM HOT TO COLD

RAM → Database (TSDB) → Object Storage → Archive

Temp	Purpose	Storage Products	Immutable?
Boiling Hot	App usage	Internal app cache, Redis or Memcached	No
Hot	Operational dataset	NoSQL Database (preferably Time Series DB) Riak TS, OpenTSDB, KairosDB, Cassandra, HBase	Almost*
Warm	Historical data	Object storage – HDFS (Hadoop), Ceph, Minio, Riak S2 or AWS S3	Yes
Cold	Archives	Various	Yes

Data Lake

STORAGE TIERS – REALITY CHECK

RAM → Database (TSDB) → Object Storage → Archive

Elastic Cache (Redis) → Database (Postgres, DynamoDB) → AWS S3 → Glacier

Temp	AWS Service	Storage price, GB per month
Boiling Hot	Elastic Cache (Redis)	$15-45
Hot	DynamoDB RDS (Postgres)	$ 0.25-0.35 (SSD) from $0.1 (Magnetic)
Warm	Simple Storage Service (S3)	$0.024 to $0.030
Cold	Glacier	$0.007

Data Lake

OSS technologies for scalable IoT apps

Component	Open Source Technologies
Load balancer	Ngnix, HA Proxy
Ingestion	Kafka, RabbitMQ, ZeroMQ, Flume
Stream Computing	Spark Streaming, Apache Flink, Kafka Streams, Samza
Time Series Store	InfluxDB, KairosDB, Riak, Cassandra, OpenTSDB
Profiles Store	CouchBase, Riak, MySQL, Postgres, MongoDB
Search	Solr, Elastic Search
Object Storage	HDFS (Hadoop), Minio, Riak S2, Ceph
Analytics Framework	Apache Spark (& MLlib), MapReduce, Hive
SQL Query Engine	Spark SQL, Presto, Impala, Drill
Cluster Manager	Mesosphere DC/OS or Mesos, Kubernetes, Docker Swarm

Checklist for IoT technology stack

- ❑ Is it vendor lock-in or <u>open source software</u>? Are there open APIs?
- ❑ Can it be deployed in cloud? At the edge? In a data center? Using hybrid approach?
- ❑ Can it be used it for free or low cost (no big upfront investment)?
- ❑ Are the components <u>pre-integrated</u> or can be easily integrated together?
- ❑ Can you develop your app on your laptop? How many "moving parts"?
- ❑ Can you <u>easily scale each component</u> in this architecture by 2x? 10x? **50x**?
- ❑ Is there a roadmap, actively worked on, which is aligned with <u>your vision</u>?
- ❑ Is there a <u>company</u> behind the technology to provide 24x7 support when needed?

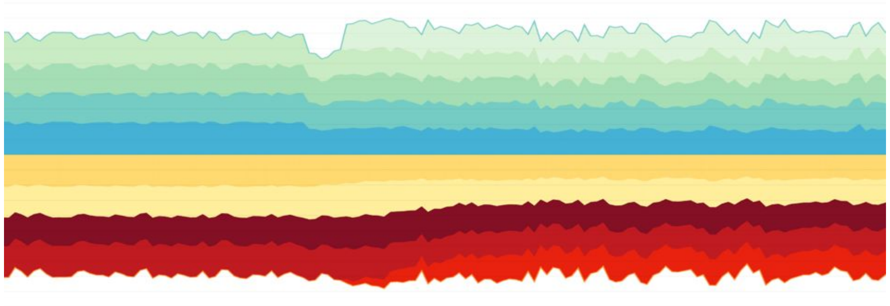
OSS technologies for IoT apps - the "**opinionated**" choice

Component	Open Source Technologies
Load balancer	HA Proxy
Ingestion	Apache Kafka
Stream Computing	**Structured Spark Streaming**
Time Series Store	**Riak (TS tables)**
Profiles Store	**Riak (KV buckets)**
Search	**Riak Search** (based on Solr)
Object Storage	**Riak S2**
Analytics Framework	**Apache Spark (& MLlib)**
SQL Query Engine	**Apache Spark SQL**
Cluster Manager	Mesosphere DC/OS or Kubernetes

- Riak TS (Time Series) - highly scalable NoSQL database for IoT and Time Series

... and more

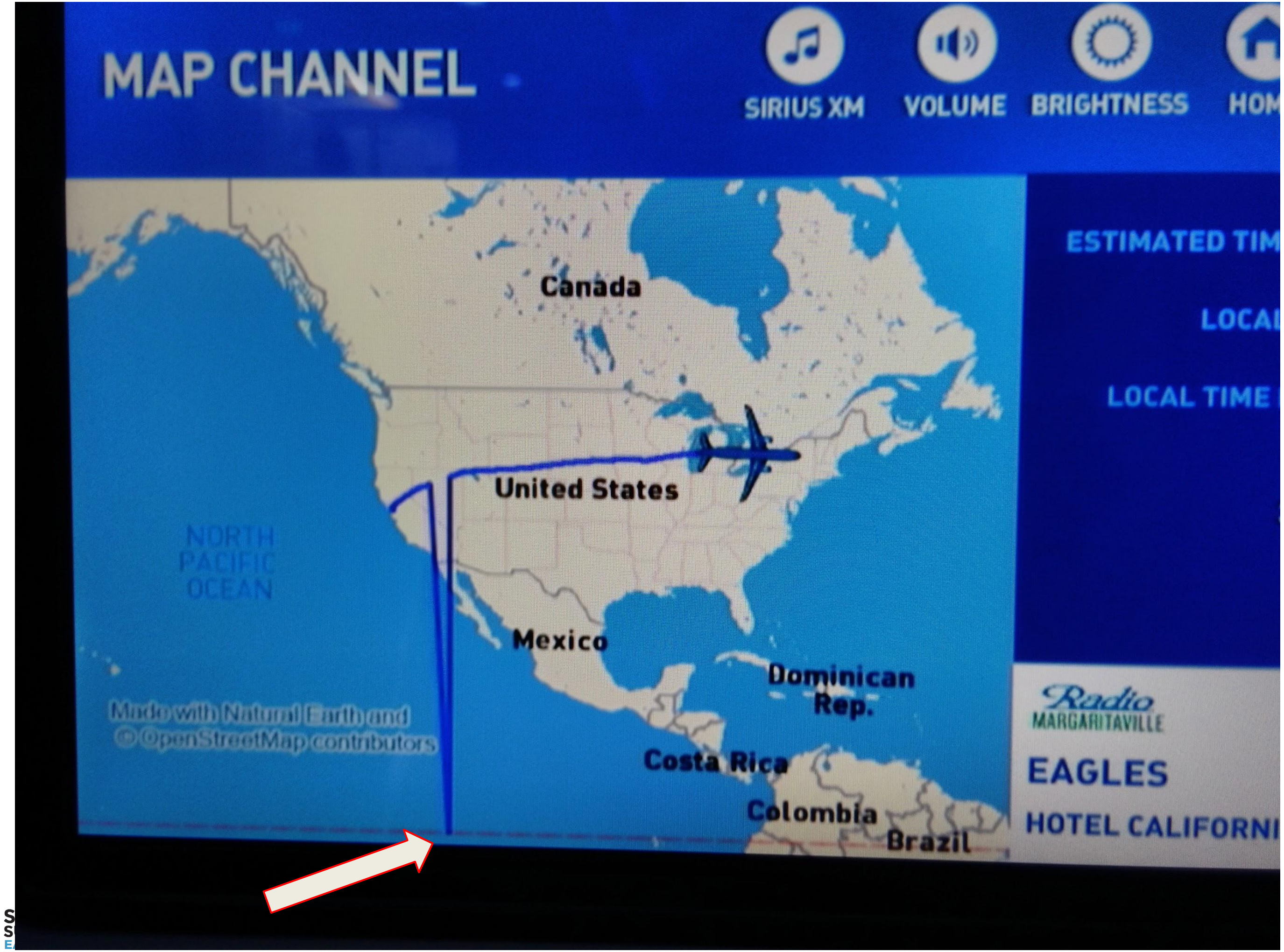
- Riak **Spark** Connector for **Apache Spark**
- Riak Integrations with Redis and Kafka
- Riak Mesos Framework (RMF) for DC/OS



Thank You!

Contact me at [pavel at basho dot com]

SPARK SUMMIT EAST 2017



A321

A Condor A321

Role	Narrow-body jet airliner
National origin	Multi-national[a]
Manufacturer	Airbus
First flight	11 March 1993
Introduction	1994 with Lufthansa
Status	In service
Primary users	American Airlines China Southern Airlines Turkish Airlines China Eastern Airlines
Produced	1992–present
Number built	1,447 as of 31 January 2017[1]
Unit cost	US$114.9 million[2]