

'A TIME TO SPARK' **Torben Hoffmann** Architect @ basho.com @lehoff







Click 'engage' to rate sessions and ask questions

## **Basho Snapshot**

Distributed Systems Software for Big Data, IoT and Hybrid Cloud applications



#### 2011 Creators of Riak

- Riak KV: Highly distributed NoSQL database
- Riak S2: Large Object Storage

#### 2015 New Products

- Basho Data Platform: Integrated NoSQL databases, caching, in-memory analytics, and search
- Time Series: Riak optimised for time series data

100+ employees

#### **Global Offices**

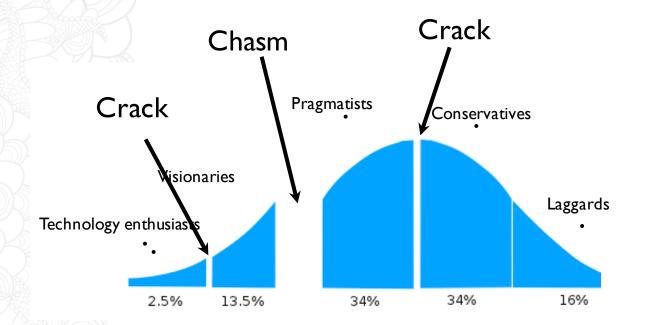
Seattle (HQ), Washington DC, London, Tokyo



## AN IoT RICH DATA DAY



## **Cracks and a Chasm**





### MEETING THE NEEDS OF THE ENTERPRISE

#### PRIORITIZED NEEDS

High Availability - Critical Data

High Scale - Heavy Reads & Writes

Geo Locality - Multiple Data Centers

Operational Simplicity – Resources

Don't Scale as Clusters

Data Accuracy – Write Conflict Options



#### RIAK KV USE CASES

User Data
Session Data
Profile Data
Real-time Data
Log Data

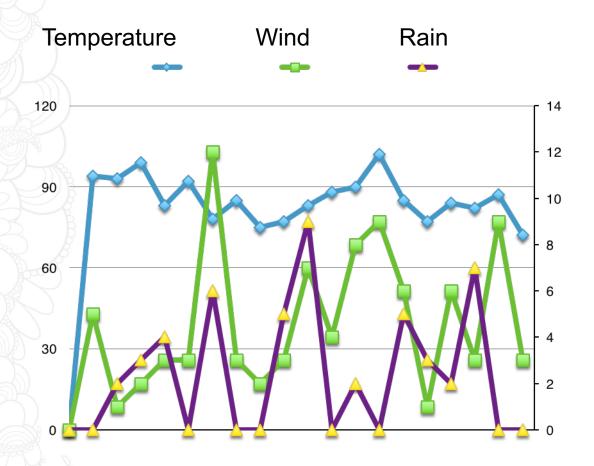


## TIME SERIES USE CASES

IoT/Devices
Financial/Economic
Scientific Observations
Log Data

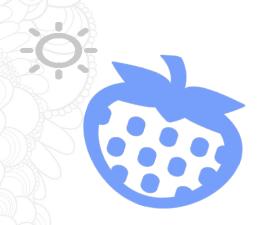




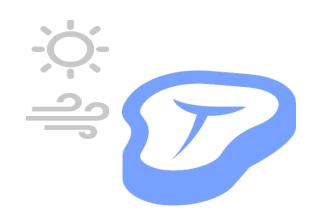




## WEATHER FORECAST PREDICTS Walmart > SALES



Ideal BERRY weather turns out to be low wind with temperatures below 80 degrees.



People are more likely to eat STEAK when it's warm out with higher winds but no rain, but not if it gets too hot.



## THE CONNECTED HOME

Smart Thermostat Market is expected to grow from \$146.9 million in revenue in 2014 to \$2.3 billion by 2023.

-Navigant

- Store historical date for analytics
- Store User Schedules
- Sessions Storage for Connected Users









## POS FOR OVER 20K SMALL BUSINESSES

Cloud-based management tools and data analytics usually only accessible to larger companies.

Millions of transactions each day.

- Quick Service
- Retail
- Restaurant & Bar









# 20 TERABYTES OF DATA PER DAY BILLIONS OF MOBILE DEVICES

- 10 BILLION data transactions a day
   150,000 a second Apple
- Forecasting 2.8 BILLION locations around the world
- Generates 4GB OF DATA every second



We're focusing on helping people make better decisions with the weather.





## 1300 MESSAGES A SECOND 80 MILLION PEOPLE

- 24 Hours a Day / 365 Days a Year
- Information is requested and amended more than 2.6 BILLION times a year
- Has enabled over 42 MILLION Summary Care Records to be created and stored
- Has transmitted over 1.3 BILLION prescription messages



## bounday

## 1.5 TRILLION RECORDS PER DAY

400% Year Over Year Increase

- Ability to Monitor entire IT environment from Single Portal
- Array of Real-Time Statistics and Insight
   Centralize & Correlate Events, Alerts, and Notifications





## WHAT ARE THE IMPLICATIONS?

Enterprises must choose: Modernize or Sink!



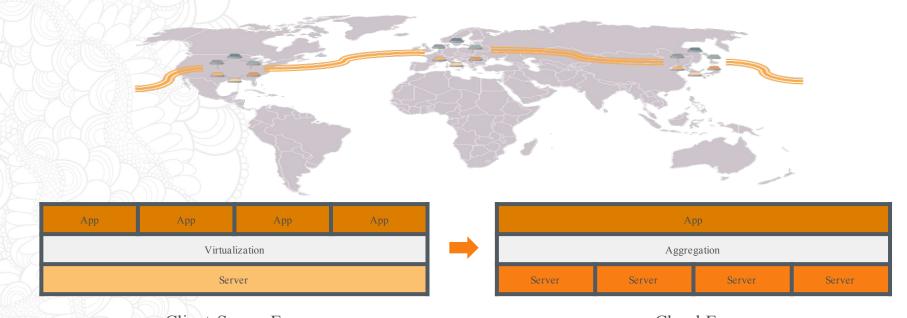




Making Time Work for You



## DISTRIBUTED WORKLOADS



Client-Server Era:

SMALL APPS

BIG SERVERS

ONE LOCATION

Cloud Era:
BIG APPS
SMALL SERVERS
MANY LOCATIONS



## WHY DISTRIBUTED SYSTEMS?

Everything works at small scale

What happens when something goes wrong

The customer experience matters



#### **DISTRIBUTED SYSTEMS**

Scale out, up and down predictably and linearly

Survive server, network or data center failures

Data locality enables data operations close to end-users



#### **DEVELOPERS**



#### **OPERATIONS**



#### **CUSTOMERS**



#### **SALES**



## DISTRIBUTED SYSTEMS AND NoSQL





















## Cleaning up







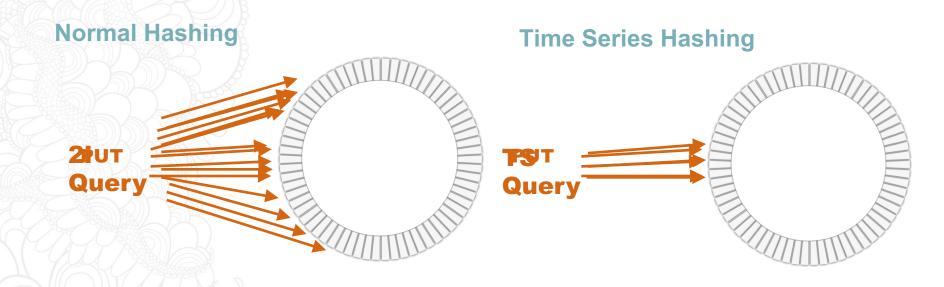






### BASICALLY ITS ALL ABOUT WHERE THE DATA IS

- Group some data together logically
- Write that data to the same vnode
- Run queries that go direct to the data





## NoSQL OPTIMIZED FOR IoT

Time Series Data is different

Higher write volumes of immutable data

**Data location matters** 

Data needs to be easy to retrieve using range queries

Eventually rolled up, compressed, with the details expired





select \*
from devices
where time >= 2015-08-06 1:00:00
and time <= 2015-08-06 01:10:00
and errorcode = 555123



All while still being highly available!

With no data loss even with a huge number of sources



## HOW DO WE SOLVE THIS





- Data Co-Location
- Composite Keys time and data family, series id
- Time quantization (quantum)

Simplified Data Modeling



DDL – Table and field definitions support structured and semi-structured data

Fast Queries and Analysis



- Range Queries (SQL like)
- LevelDB filtering
- Spark Connector



## DEVELOPER TEAM BENEFITS

Traditionally in NoSQL NoSQL designed for Time Series Developer's Responsibility Validate Process and and query check the data data Database's Validate Process Responsibility and and query check the data KV Store data KV Store



#### STORING STRUCTURED DATA

**Buckets used to Define Tables** 

#### Key format

Objects have a composite key
 (partition key and local key)

#### Tables

- Buckets can be defined as tables
- Tables can have a schema defined using DDL
- Columns in the table can be typed

#### Data Validation

Data is validated on input

San Francisco Water Data					
Site	ISO 8601 Timestamp	Temperature (C)	Specific Conductance (µS/cm)	Salinity (PSU)	Turbidity (FNU)
PIER_17	2015-09-18T00:00:00-08:00	17.8	48500	31.6	3
PIER_17	2015-09-18T00:15:00-08:00	17.8	48500	31.6	2.8
PIER_17	2015-09-18T00:30:00-08:00	17.7	48400	31.6	2.8
PIER_17	2015-09-18T00:45:00-08:00	17.8	48500	31.6	3.6
PIER_17	2015-09-18T01:00:00-08:00	17.7	48500	31.6	2
PIER_17	2015-09-18T01:15:00-08:00	17.3	48500	31.6	2.8
PIER_17	2015-09-18T01:30:00-08:00	17.4	48500	31.6	2.4
PIER_17	2015-09-18T01:45:00-08:00	17.3	48500	31.6	3.8
PIER_17	2015-09-18T02:00:00-08:00	17.4	48400	31.6	3
PIER_17	2015-09-18T02:15:00-08:00	17.4	48500	31.6	2.3
PIER_17	2015-09-18T02:30:00-08:00	17	48600	31.7	1.4
PIER_17	2015-09-18T02:45:00-08:00	17	48600	31.7	2.1
PIER_17	2015-09-18T03:00:00-08:00	16.9	48700	31.8	2.4
PIER_17	2015-09-18T03:15:00-08:00	16.8	48700	31.8	5.6

Form of the Mileton Bak



### **RANGE QUERIES**

Query Like SQL

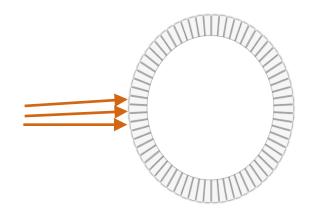
- Use Cases
  - Range queries
- Implementation Details
  - SQL like query language
  - Filtering rows based on column expressions
  - Filtering executed in backend
  - Specific columns are extracted
  - Simple select with WHERE clause
    - for numbers <,>=,<,<=,=,!=</li>
    - for other data types =, !=
    - AND, OR (nesting operators are supported)

select \*
from devices
where time >= 2015-08-06
1:00:00
and time <= 2015-08-06
01:10:00
and errorcode = 555123
and host = 'comaxb'

### **DATA LOCALITY**

Co-location is Key for Performance Gains

- Use Cases
  - Data co-located by time
  - Data co-located by geohashing
- Implementation details
  - Composite key defines how data is stored
  - Data is collocated on vnodes to speed up reads.
  - A quantum can be configured to prevent hotspots
  - This results in efficient range scans
  - Additional optimizations done for LevelDB scans





### **PEFORMANCE IMPROVEMENTS**

Fast Reads and Writes

- Improved write throughput
  - Backend improvements
  - Object marshalling and demarshalling process improvements
- Efficient range scans
  - Data locality
  - LevelDB scan optimizations
  - Query co-ordination



## **Open Issues aka WIP**

- Multi-Data Center / Cluster Replication
- Support for aggregation.
  - Out of the box count, sum, mean, min, max, stddev, top-n.
  - Can be done today by the developers on the client or by using the Spark connector.
- Only LevelDB backend is supported for the performance advantage
- Support for bucket-level deletion, bulk deletes
- Support for data expiry
- Support for SOLR / Riak Search
- Support for Spark Streaming
- Bulk / Buffered Uploads
- Query enhancements
- Additional performance improvements

Try Riak TS today basho.com/contact

Riak TS 1.0 generally available in mid-december





Write it like Riak Analyze it like Spark



#### **BASHO DATA PLATFORM**

Simplifies Big Data, IoT and Hybrid Cloud Applications

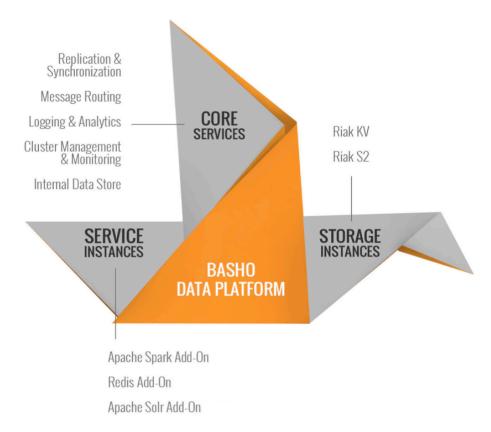
Simplify Complexity

High Availability and Fault Tolerance

Integrate Real-time Analytics with Apache Spark

Faster Application Performance with integrated Redis

Enriched search with Solr

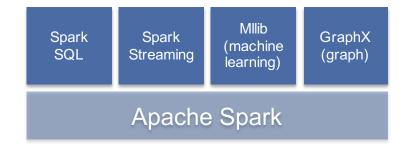






**Apache Spark™** is a fast and general engine for large-scale data processing.

- Open Sourced originally developer in AMPLab at UC Berkley
- Designed for running iterative algorithms & interactive analytics



#### **Speed**

Run programs up to 100x faster than Hadoop MapReduce in memory

or 10x faster on disk

#### **Ease of Use**

Over 80 high-level operators that make it easy to build parallel apps. And you can use it interactively from the Scala, Python and R shells.

#### Runs Everywhere

Including Mesos, standalone, or in the cloud and accesses diverse data sources



## APACHE SPARK ADD-ON

#### Write it like RIAK. Analyze it like Spark.

#### Cluster Management

Manage Spark clusters at scale using built-in leader election enabled by the Spark Connector for Riak.

#### Data Mover

Intelligently load data into Spark clusters to minimize network traffic and processing overhead.

#### Write-Back to Riak

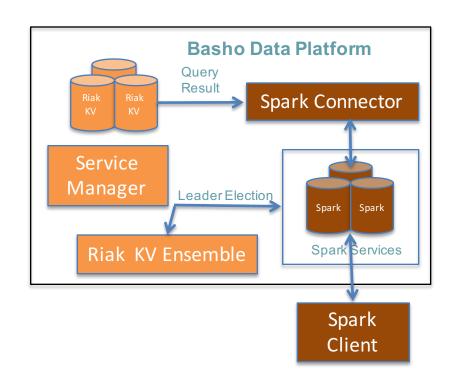
Store intermediate and final results back into Riak KV for further processing by Spark or other Big Data application components.

#### Perform at Scale

Apache Spark is architected for high performance, real-time analysis and persistence of Big Data.

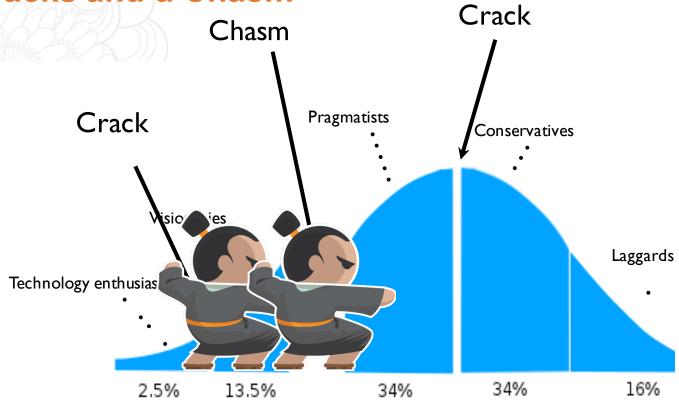
#### Operational Simplicity

Quickly deploy and configure Spark clusters with Riak KV. Auto-start failed Spark instances to reduce manual operations.





## **Cracks and a Chasm**









## VISIT OUR STAND!

Basho.com







## Remember to rate session

Thank you!