



IoT を担うCassandraとSparkの利便性

INTHEFOREST





自己紹介

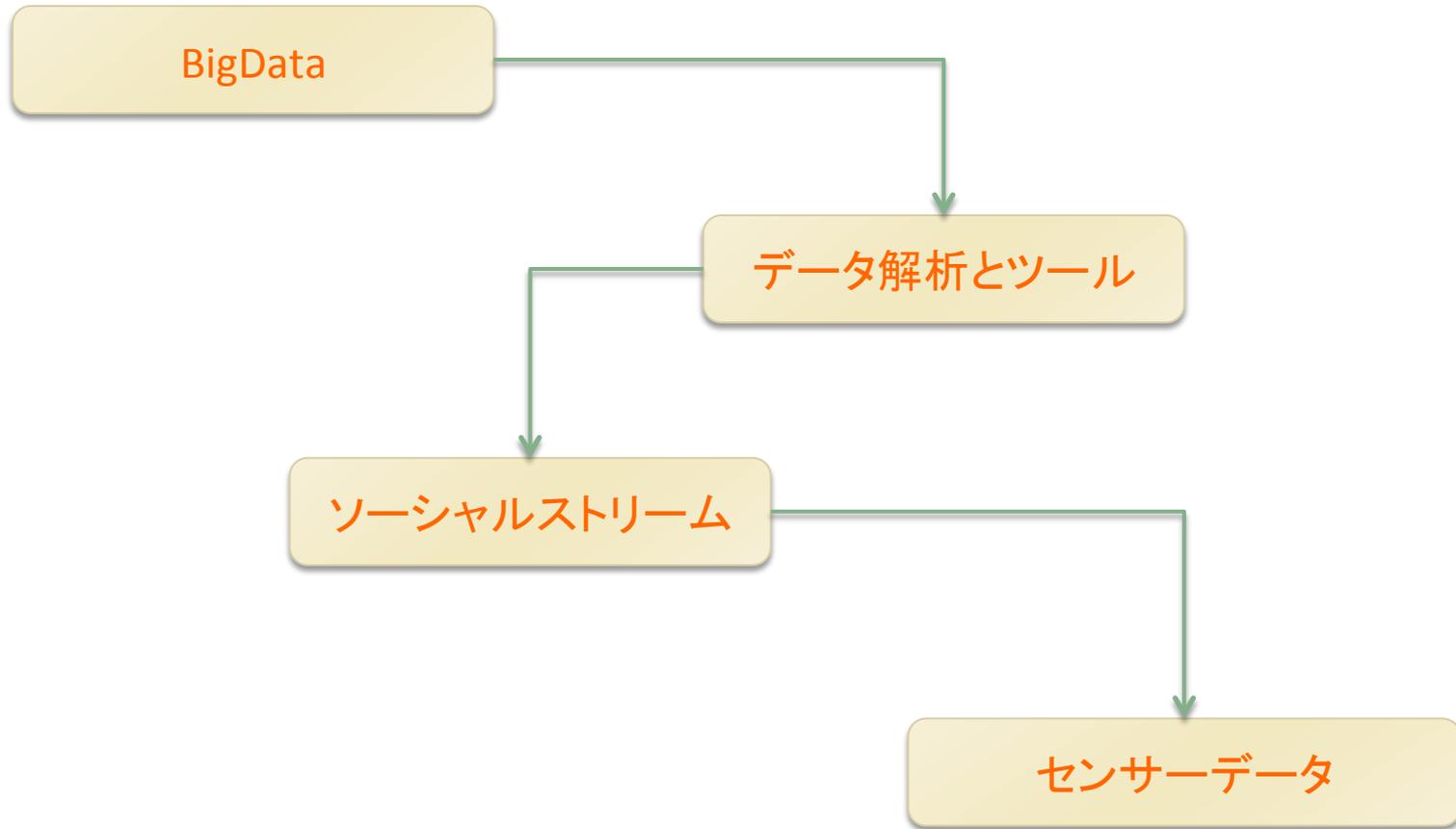
富田 和孝 (@railute)

株式会社INTHEFOREST 代表取締役社長
Cassandra商用サポート、Cassandraコンサルティング他

Python歴7年、最近はJavaが中心。
PHP、Ruby、JavaScript、Perl、Cは業務経験有り。

職種：本職はDB・インフラ系エンジニア
以前、某レストランサーチのDBA
高負荷・大容量・大規模のOracleRACとPostgreSQLと
MySQLに苦しめられ続けた経験あり。
最近のトレンドはかつての専攻「認知意味論」をベースにした
NLP

Agenda



What's bigdata?



既存のツールでは扱いきれないデータ

既存ツールも進化している

但しパラダイムシフトも起きている

立場によって変わる言葉

既存ツールの進化を超えるデータ増量

IoTというパラダイムシフト



IoTとM2M

Internet of Things

一意に識別可能な「もの」がインターネット/クラウドに接続され、情報交換することにより相互に制御する仕組みである

Machine to Machine

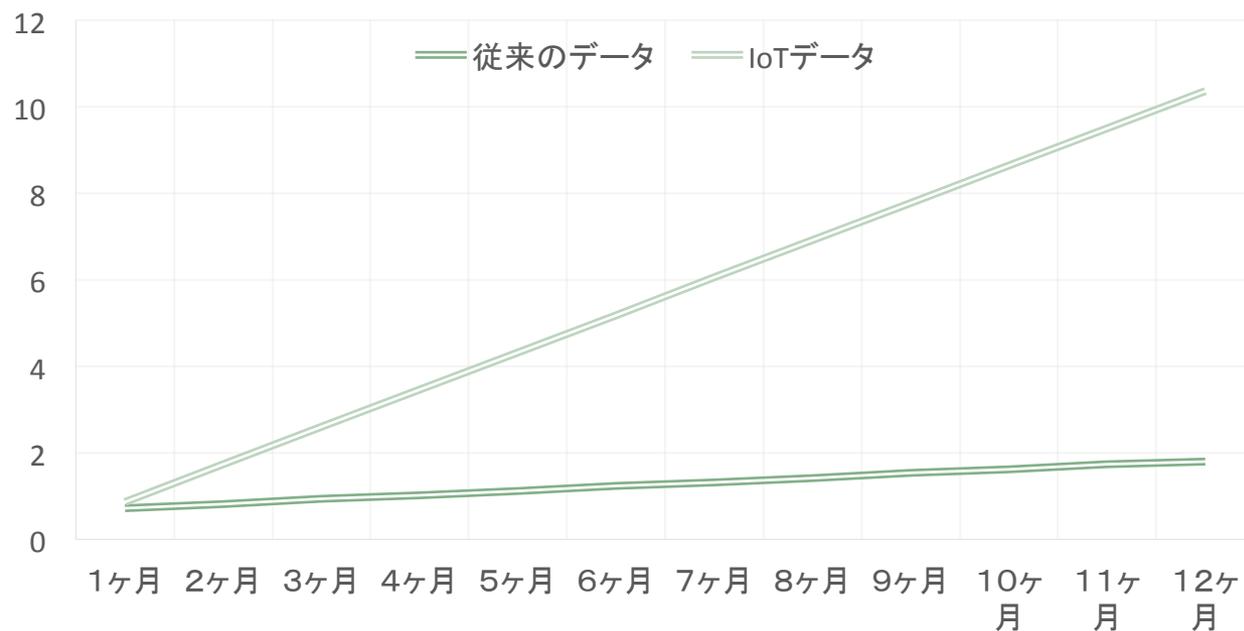
マシンツーマシン(Machine-to-Machine)とは、コンピュータネットワークに繋がれた機械同士が人間を介在せずに相互に情報交換し、自動的に最適な制御が行われるシステムを指す。



IoTというパラダイムシフト



データ増加率



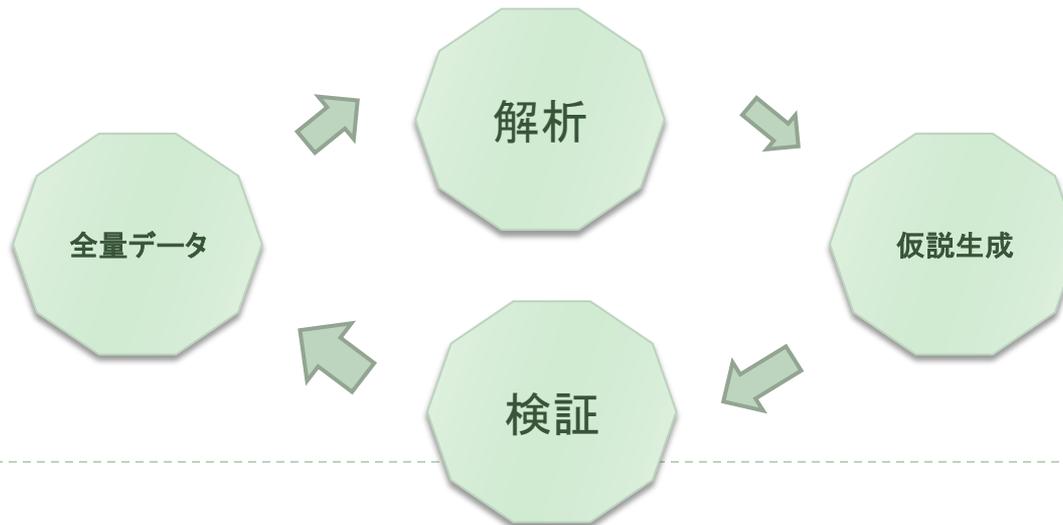
IoTというパラダイムシフト



従来のデータ解析



最近のデータ解析



データ解析とツール



リニアにスケールするCassandra

Wednesday, November 2, 2011

Benchmarking Cassandra Scalability on AWS - Over a million writes per second

by Adrian Cockcroft and Denis Sheahan

Netflix has been rolling out the Apache Cassandra NoSQL data store for production use over the last six months. As part of our benchmarking we recently decided to run a test designed to validate our tooling and automation scalability as well as the performance characteristics of Cassandra. Adrian presented these results at the [High Performance Transaction Systems](#) workshop last week.

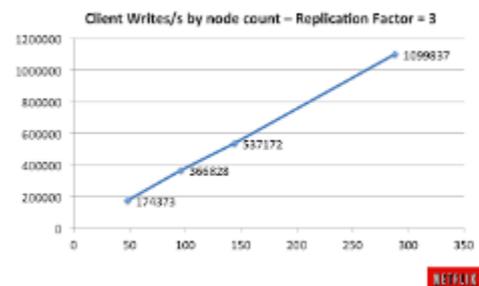
We picked a write oriented benchmark using the standard Cassandra "stress" tool that is part of the product, and Denis ran and analyzed the tests on Amazon EC2 instances. Writes stress a data store all the way to the disks, while read benchmarks may only exercise the in-memory cache. The benchmark results should be reproducible by anyone, but the Netflix cloud platform automation for AWS makes it quick and easy to do this kind of test.

The automated tooling that Netflix has developed lets us quickly deploy large scale Cassandra clusters, in this case a few clicks on a web page and about an hour to go from nothing to a very large Cassandra cluster consisting of 288 medium sized instances, with 96 instances in each of three EC2 availability zones in the US-East region. Using an additional 60 instances as clients running the stress program we ran a workload of 1.1 million client writes per second. Data was automatically replicated across all three zones making a total of 3.3 million writes per second across the cluster. The entire test was able to complete within two hours with a total cost of a few hundred dollars, and these EC2 instances were only in existence for the duration of the test. There was no setup time, no discussions with IT operations about datacenter space and no more cost once the test was over.

To measure scalability, the same test was run with 48, 96, 144 and 288 instances, with 10, 20, 30 and 60 clients respectively. The load on each instance was very similar in all cases, and the throughput scaled

50ノードから300ノードへのスケール ングベンチマークテスト

Scale-Up Linearity



Per-Instance Activity

The next step is to look at the average activity level on each instance for each of these tests to look for bottlenecks. A summary is tabulated below.

Per Node Activity

Per Node	48 Nodes	96 Nodes	144 Nodes	288 Nodes
Per Server Writes/s	10,900 w/s	11,460 w/s	11,900 w/s	11,456 w/s
Mean Server Latency	0.0117 ms	0.0134 ms	0.0148 ms	0.0139 ms
Mean CPU %Busy	74.4 %	75.4 %	72.5 %	81.5 %
Disk Read	5,600 KB/s	4,590 KB/s	4,060 KB/s	4,280 KB/s
Disk Write	12,800 KB/s	11,590 KB/s	10,380 KB/s	10,080 KB/s
Network Read	22,400 KB/s	23,610 KB/s	21,390 KB/s	23,640 KB/s
Network Write	18,600 KB/s	19,600 KB/s	17,810 KB/s	19,770 KB/s

Node specification - Xen Virtual Images, AWS US East, three zones
• Cassandra 0.8.0, CentOS, SunDMS
• AWS EC2 m1 Extra Large - Standard price \$ 0.68/hour



IoT向けデータベースとしてのCassandraの特徴

- 書込みに強い。
 - 書込み先が分散化されているので同時多数書込みに強い（秒間100万書込み等）
 - 結果整合性による柔軟な書込み精度を選択可能
- 解析ツールとの親和性
 - 多彩なドライバ(ODBC、JDBC、PHP、Ruby、Perl等)
 - Apache Hadoop、Apache Spark、Presto等の多彩な解析ツールを利用可能
- マルチベンダー
 - Windows、Linux、各種クラウド、JVMが稼働すれば使用する事が出来ます。Windowsでの採用実績もあります。

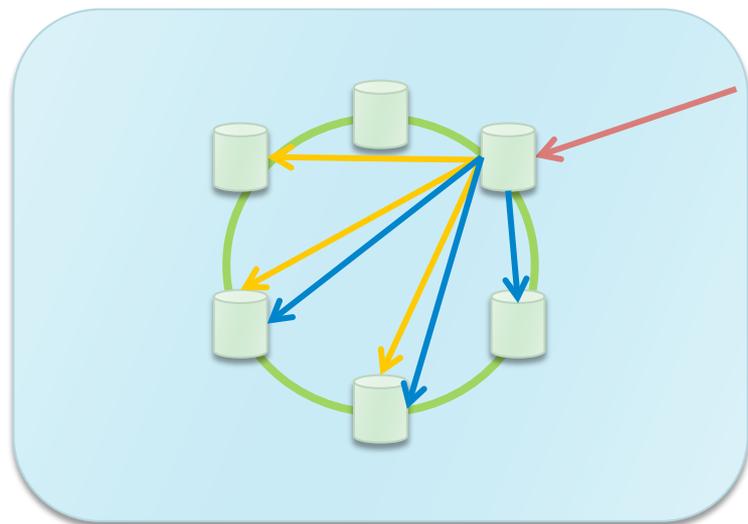


データ解析とツール



Cassandraは横断検索が苦手

RowKeyベースのConsistent Hashingの為、連続したKeyが同じノードに存在するとは限らないので連続した領域のデータ取得はCassandra単体では苦手

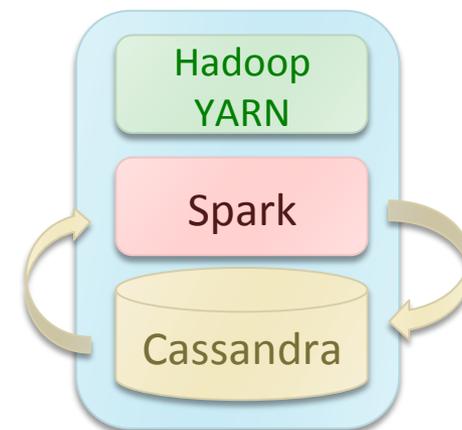
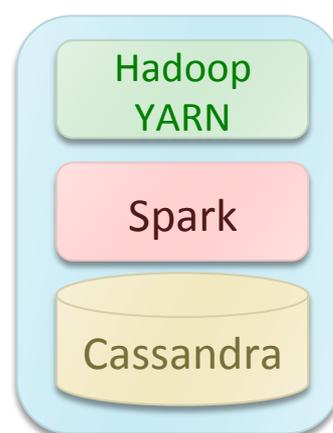
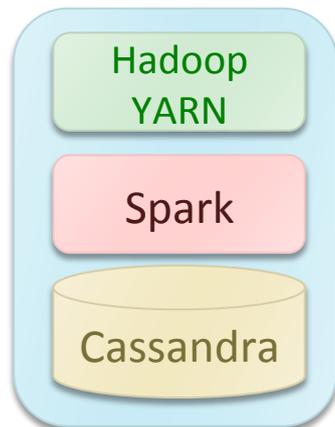
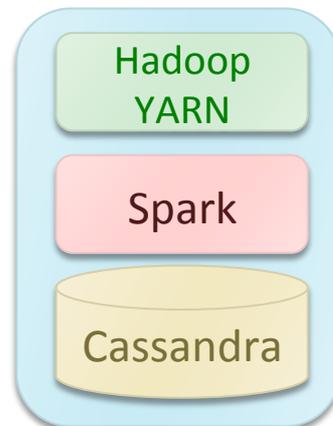
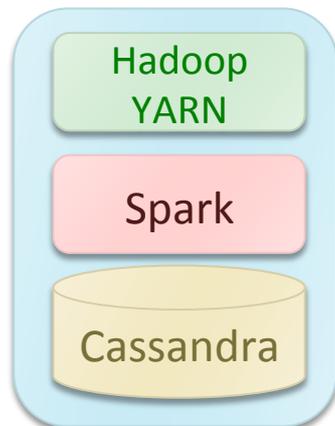
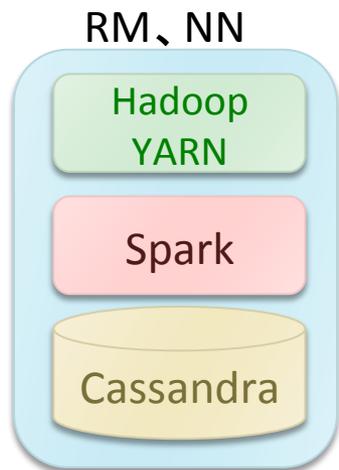


データ解析は総当たりの逐次処理



データ解析とツール

Spark
システム構造



レンジ範囲内の
逐次処理



データ解析とツール



Sparkのライブラリ

- GraphX

- グラフとグラフ並列計算API

- MLlib

機械学習アルゴリズムAPI

Spark Streaming

データの逐次時系列処理

- Spark SQL

- SQLライクなQuery言語

データ解析とツール



▶ データ解析は軸が大事

- ▶ 何のための解析か、何を説明するためのレポートかをはっきりさせる。
- ▶ 業務分析をしっかりと行い、エビデンスとしてデータを使用する。



データ解析とツール



▶ HARUNIRE

データ解析ソリューション「HARUNIRE」は、データ分析を行うにあたって機械学習やスケーラブルなシステム設計など、ハードルの高い構築作業をデータ収集からデータ蓄積、データ解析、解析結果の取得までを目的に合わせてシームレスな環境として総合的にサポート。





サービス内容

- ▶ データ解析コンサルティング: データ解析を行う為に必要なデータ収集方法や、データストアの仕組み、解析軸の提案、表示方法の提案などマーケティングに必要なデータ解析を行う為のデータ解析設計をご提案します。
 - ▶ データ解析の為のデータストックの仕組みを構築します。
 - ▶ 機械学習など複雑で難しくなりがちな解析方法にあわせたアプリケーションを作成します。
-



業務例)



- ✓ ソーシャルストリーム
- ✓ センサーデータ



ソーシャルストリーム



SNSとAPI

Twitter

FaceBook





ソーシャルストリーム

例)

▶ Twitter

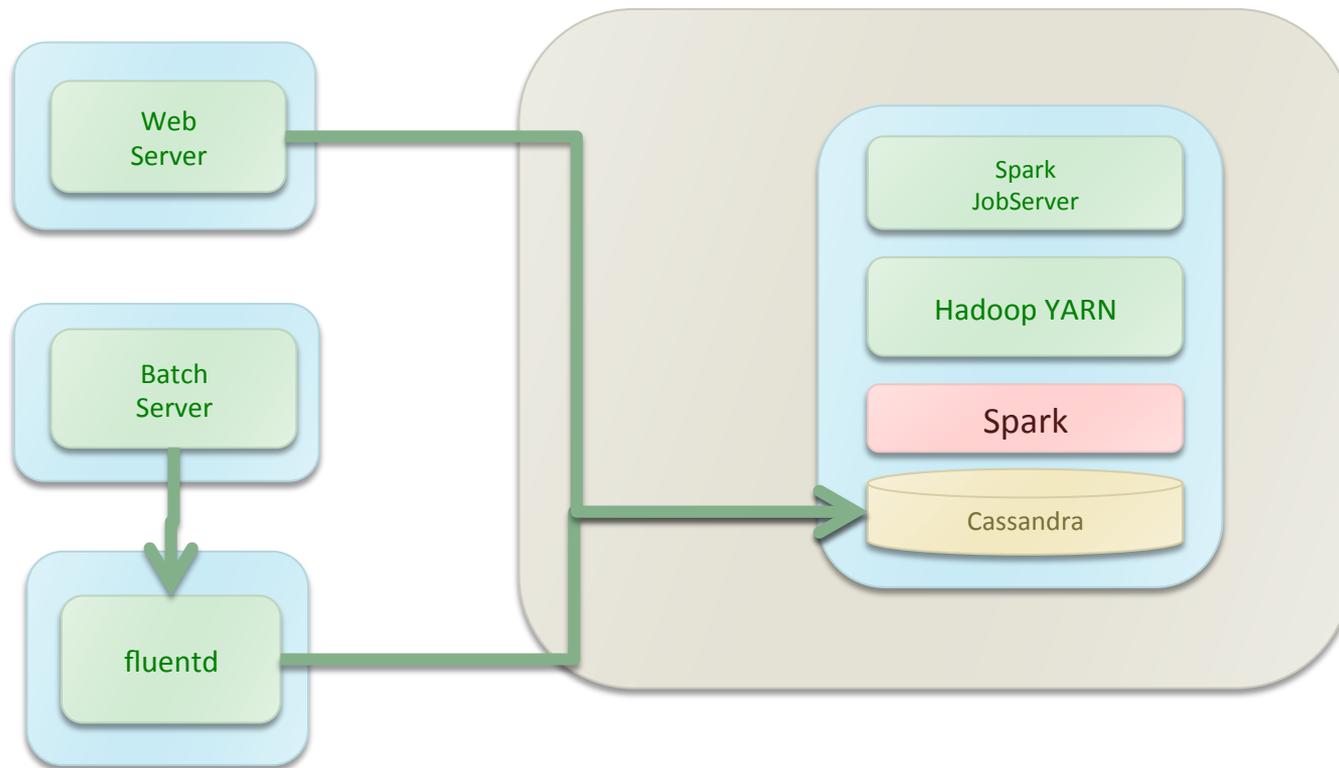
- ▶ APIより取得(二ヶ月間で3000万件超30,493,408)
 - ▶ 特定キーワードによるデータ収集
 - ▶ Reply除去
 - ▶ ワードカウントを利用した傾向分析
 - ▶ k-meansを使用した15クラスレベルのクラスタリング
 - ▶ 特定のワードを含むデータにラベル付けした教師データを元にした Naive Bayesによるクラシフィケーション
 - ▶ ラベル毎のワードの強さをを用いた傾向分析
-



ソーシャルストリーム



実際のシステム構成



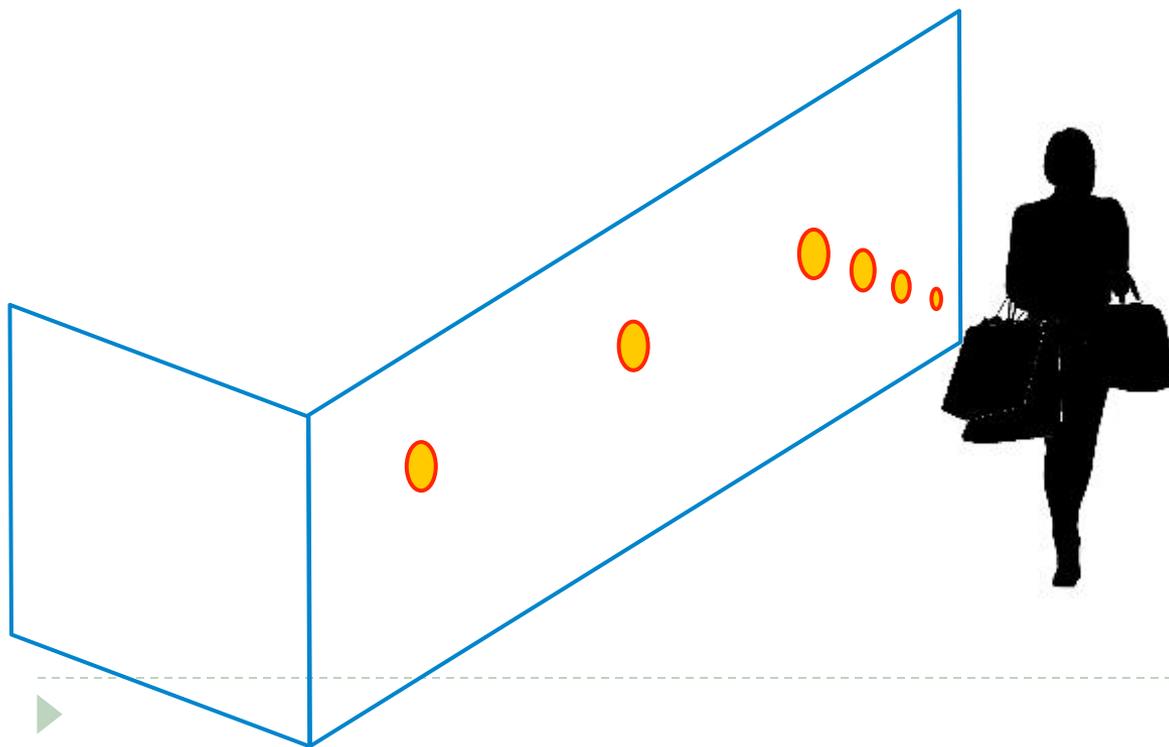
センサーデータ



例1)

iBeacon

BLE(Bluetooth low Energy)を使える端末(スマートフォンなど)を発信器として、その端末の持つエリアへの入出検知や距離計測を行える仕組み



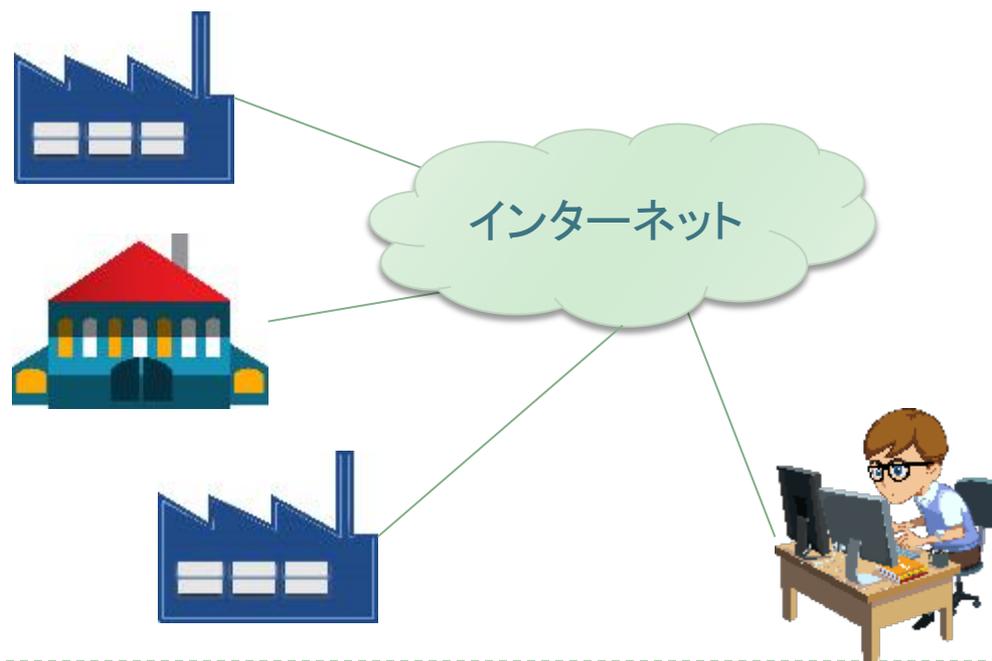
センサーデータ



例2)

Industrial Internet

産業用機器とITの融合に関するコンセプト。
高機能の機器、低コストのセンサー、インターネット、ビッグデータ収集・分析技術などを組み合わせ故障率検地など機器の効率的な運用を行う。





まとめ

- ▶ 解析には軸が必要
- ▶ Apache Spark 及び Apache Cassandraの組み合わせは環境構築のハードルが高い
- ▶ 解析用のQUERYはインタラクティブ
- ▶ 解析用のアプリケーションもインタラクティブに作成できることが重要

