Scala Cassandra Tableau

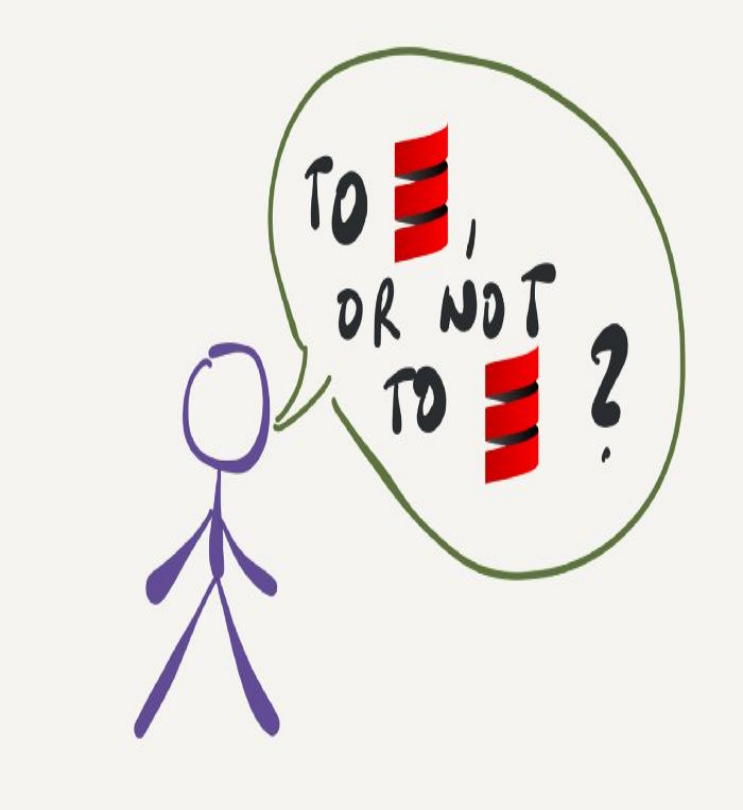
Sheng Huang

Scala

- Designed started in 2001, Released in 2004
  By Martin Odersky
- Scalable Language
  - Object Oriented and Functional Programming
- Runs on JVM and is statically typed

BIG CONCEPT

- Machine Learning concepts
- Similar to Java but concise
- Spark - Scala
- Versatile



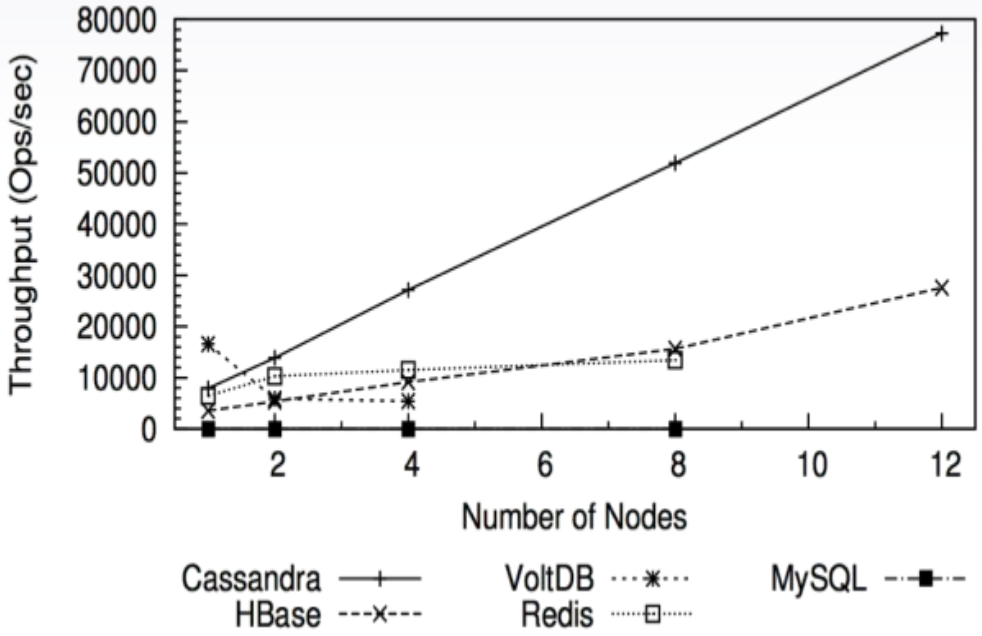
1

Cassandra Database

Overview of Cassandra

- Released in July 2008
- Avinash Lakshman and Prashant Malik
- Fault Tolerant
- Decentralized
- Scalable

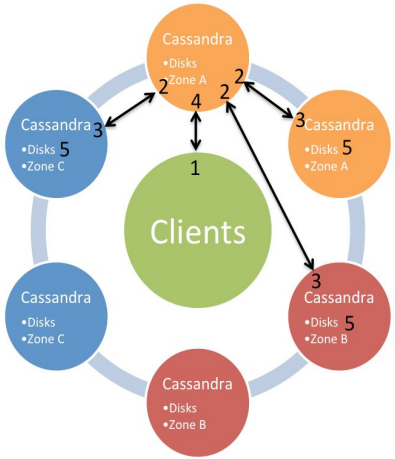
Architecture



Cassandra Write Data Flows

Single Region, Multiple Availability Zone

1. Client Writes to any Cassandra Node
2. Coordinator Node replicates to nodes and Zones
3. Nodes return ack to coordinator
4. Coordinator returns ack to client
5. Data written to internal commit log disk



If a node goes offline, hinted handoff completes the write when the node comes back up.

Requests can choose to wait for one node, a quorum, or all nodes to ack the write

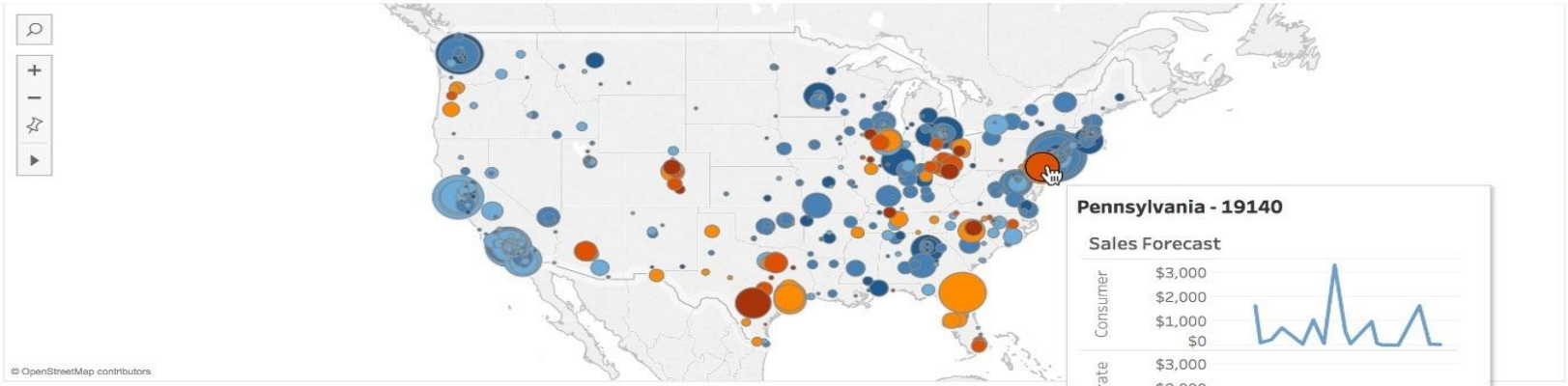
SSTable disk writes and compactions occur asynchronously

NETFLIX

tableau®

Executive Overview - Profitability (All)

Sales	Profit	Profit Ratio	Profit per Order	Sales per Customer	Avg. Discount	Quantity
$1,820,895	$231,963	12.7%	$59.04	$2,313.72	15.71%	29,793



Region
(All)

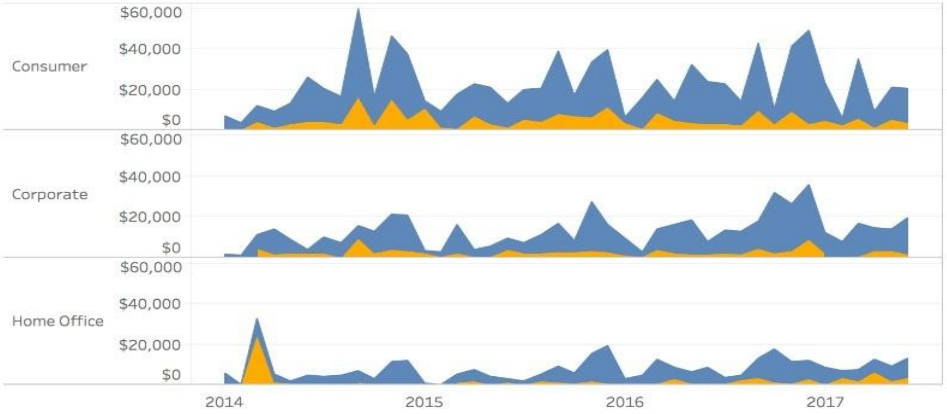
Order Date
1/3/2014 6/30/2017

Profitable
Unprofitable

Pennsylvania - 19140

Sales Forecast

Monthly Sales by Segment - States: **All**

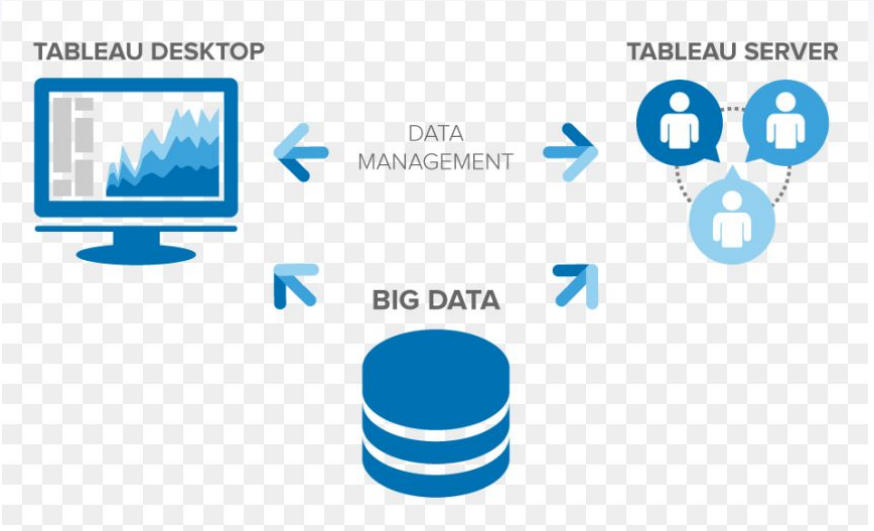


Monthly Sales b

Overview | Forecast

Tableau

- Intuitive User Experience
- Uses for everyone
- Ease of portability
  - Mobile-Friendliness
- High Performance
- Visually Pleasing
- Community



Example of Dashboard

https://public.tableau.com/profile/shenghao.huang#!/vizhome/Titanic_16093671703820/Dashboard1