

BigData

Ravi Nagvekar

What is Big Data?

What makes data, “Big” Data?

“Big Data” is data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it...

Who's Generating Big Data



Social media and networks

(all of us are generating data)



Scientific instruments

(Collecting all sorts of data)



Mobile devices

(Tracking all objects all the time)

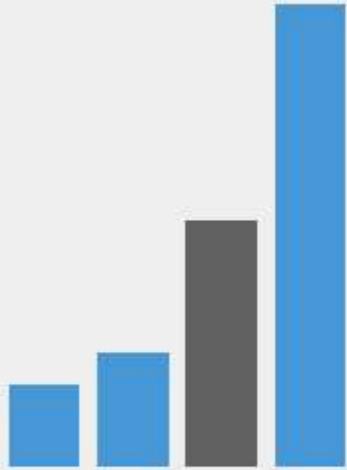


Sensor technology and networks

(measuring all kinds of data)

- The progress and innovation is no longer hindered by the ability to collect data
- But, by the ability to manage, analyze, summarize, visualize, and discover knowledge from the collected data in a timely manner and in a scalable fashion

3 V's of Bigdata



Volume



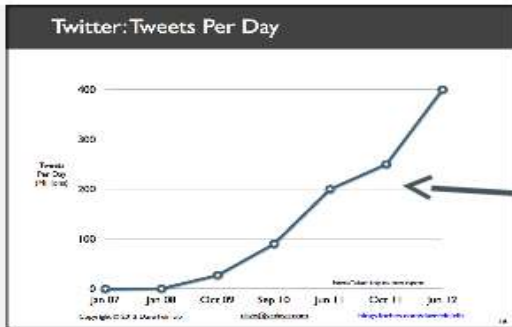
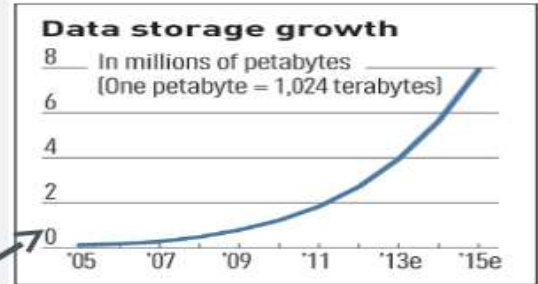
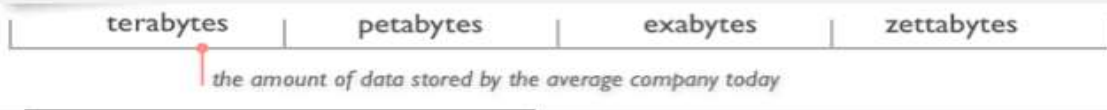
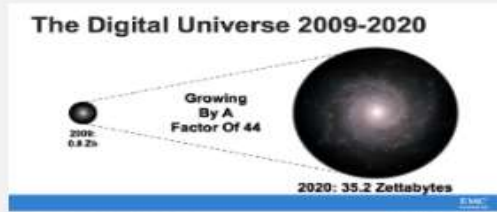
Velocity



Variety

Characteristics of Big Data: 1-Scale (Volume)

- **Data Volume**
 - 44x increase from 2009 to 2020
 - From 0.8 zettabytes to 35zb
- Data volume is increasing exponentially



*Exponential increase in
collected/generated data*

Characteristics of Big Data: 2-Speed (Velocity)

- Data is begin generated fast and need to be processed fast
- Online Data Analytics
- Late decisions □ missing opportunities
- **Examples**
 - **E-Promotions:** Based on your current location, your purchase history, what you like □ send promotions right now for store next to you
 - **Healthcare monitoring:** sensors monitoring your activities and body □ any abnormal measurements require immediate reaction

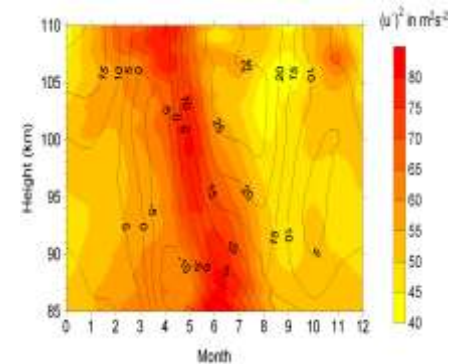
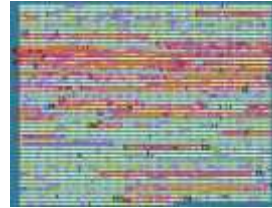
Characteristics of Big Data: 3-Complexity (Variety)

Various formats, types, and structures

Text, numerical, images, audio, video,
sequences, time series, social media data,
multi-dim arrays, etc...

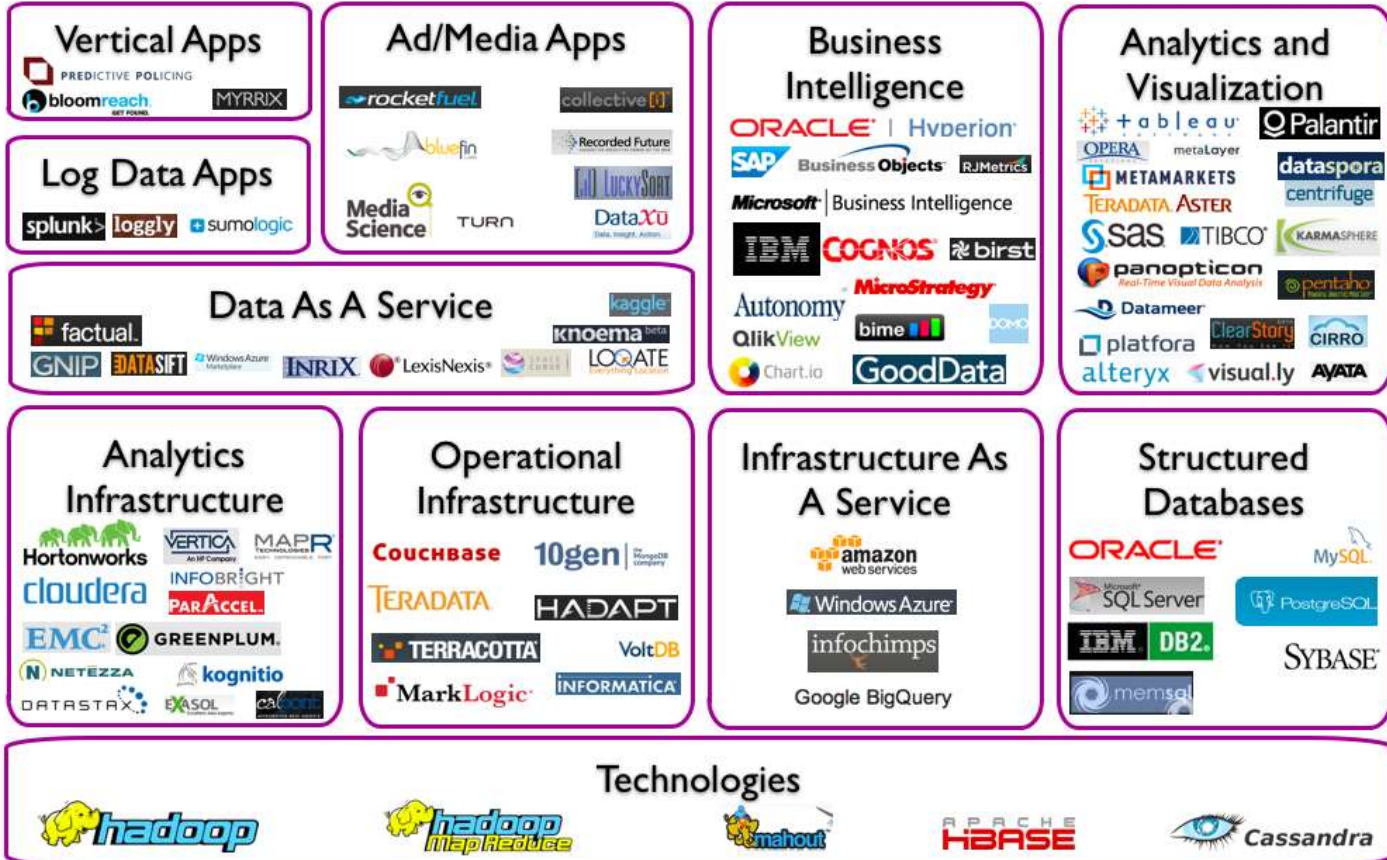
Static data vs. streaming data

A single application can be
generating/collecting many types of data

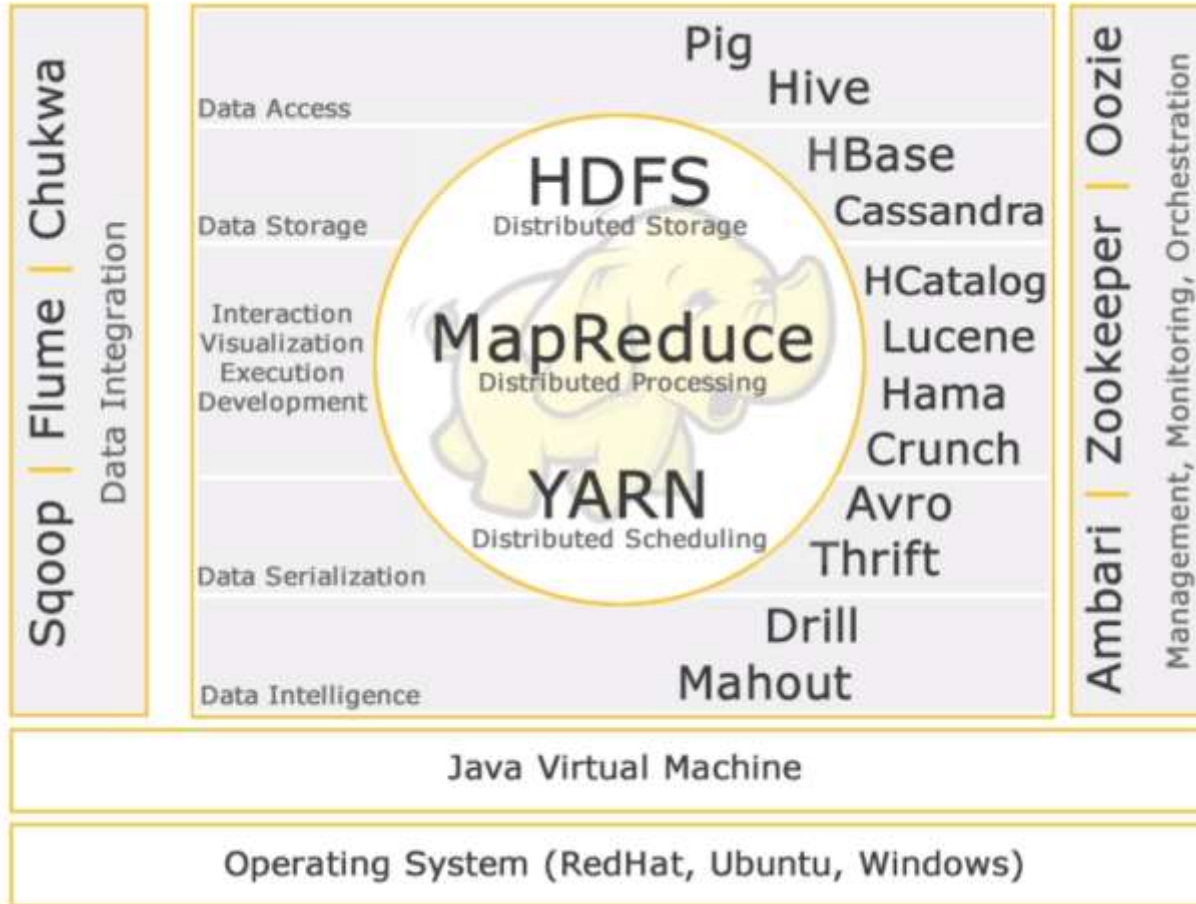


What Technology Do We Have
For Big Data ??

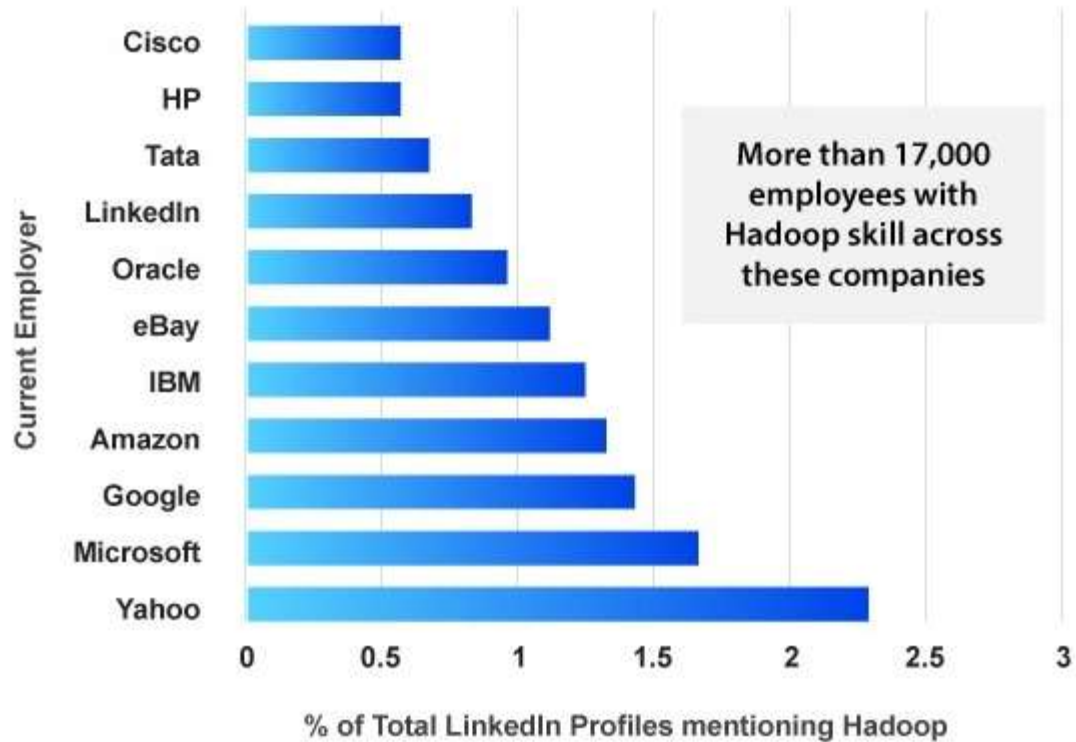
Big Data Landscape



Hadoop Ecosystem



Largest Employer of Hadoop Skills



References

<http://www.slideshare.net/raghav125/big-data-58325582>

<http://bigdatabloggin.blogspot.com/>

<http://www.slideshare.net/sbgoyal/s-b-goyal>

<http://www.slideshare.net/AnkitaSharma104/big-data-ankita1>

<http://www.slideshare.net/dhan1989/a-big-data-concept>

<http://www.forbes.com/forbes/welcome/>

http://www.slideshare.net/Hadoop_Summit/what-is-the-point-of-hadoop