

Big Data Overview

April 15

| Copyright © 2015 Tata Consultancy Services Limited

- What is Big Data ?
- Characteristics of Big Data
- Who's Generating Big Data ?
- Challenges in Handling Big Data
- How Big Data is Handled?
- Applications of Big Data
- Use cases for Big Data

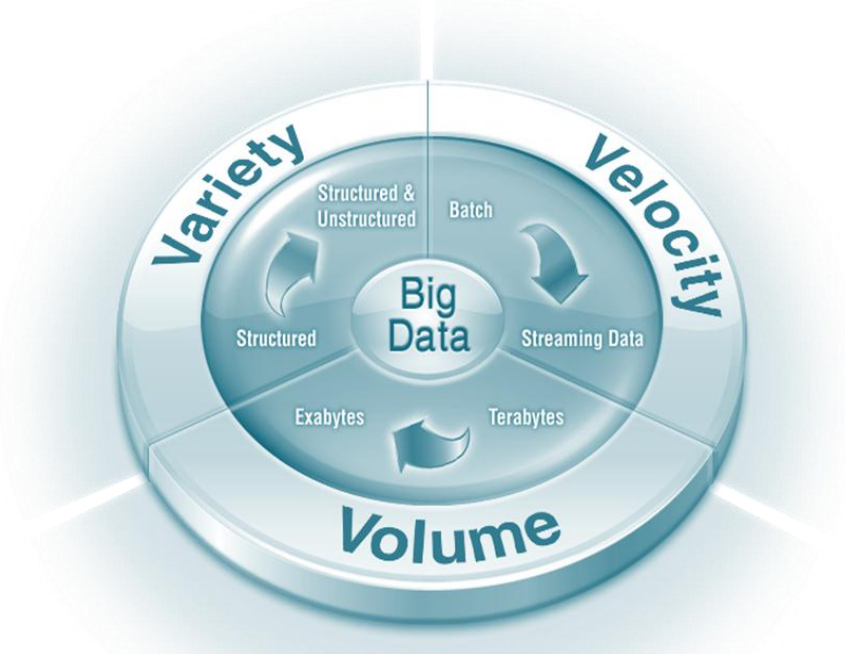
What is Big Data ?

Big data is a general term used to describe the voluminous amount of unstructured and semi-structured data a company creates.

The size of big data is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time.

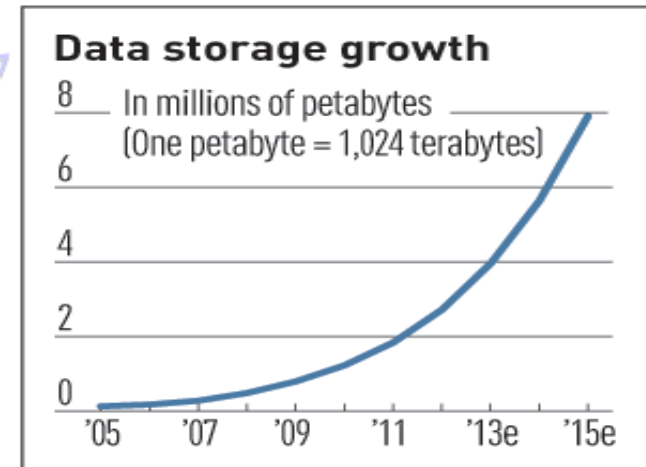
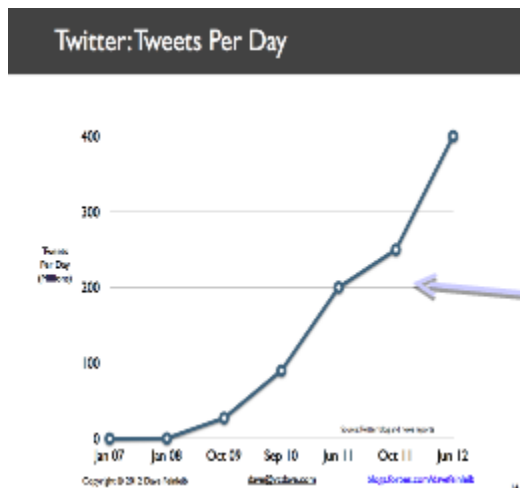
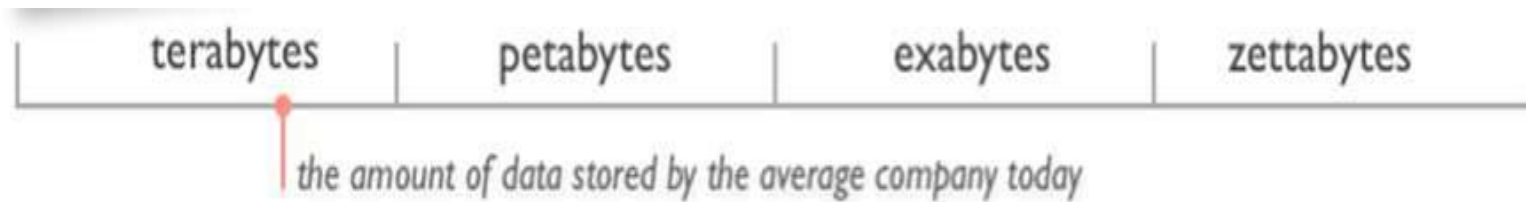
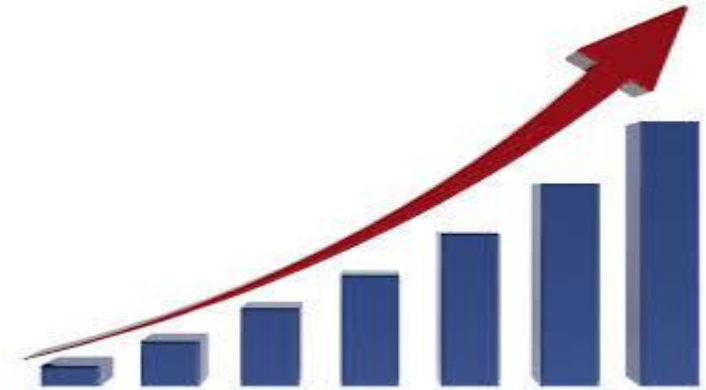
Big data spans three dimensions :

- Volume: Amount of data
- Velocity: Speed of data in/out
- Variety: Range of data types



Characteristics of Big Data - Volume

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



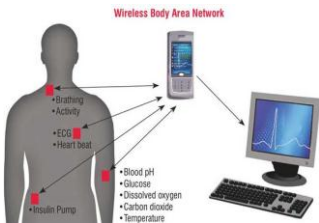
Characteristics of Big Data - Velocity

- Data is being generated fast and need to be processed fast
- Real time Data Analytics
- Late Decisions leads to Missing Opportunities



Examples

- ✓ **E-Promotions**: Based on your current location and Your purchase history and What you like, Retailers can send you the details about the promotions of their nearest store
- ✓ **Healthcare monitoring**: Sensors monitoring your activities and body, can alert any abnormal measurements that require immediate action



Characteristics of Big Data - 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 generate/collect many types of data



Who's Generating Big Data ?



Social media and Networks



Scientific instruments



Mobile devices



Sensor technology and networks

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

Structured



Relational
Databases



Spreadsheets

Semi Structured



XML



JSON

Un Structured



Social Media



Video

Challenges in Handling Big Data

- Difficulties – Capture, storage, search, sharing, analytics, visualizing data
- Data Storage – Physical storage, Acquisition, Space & Power costs
- Data Management – Skills, People, Time
- Data Processing (Information and Content management)



Traditional Way of Handling Data

Single High performing Machines



Cannot do everything...

Disadvantages:

- High Hardware Cost
- High Software Cost
- High Risk of Failure

Distribute and work - Hadoop Framework



Advantages:

- Commodity Hardware
- Free Software
- Reduced Risk of Failure
- 10 times processing power in 1/10th of cost

Applications of Big Data

A primary goal for looking at big data is to discover repeatable business patterns

Big data examples :

- Google processes about 24 petabytes of data per day
- The experiments in the Large Hadron Collider produce about 15 petabytes of data per year.
- The 2009 movie Avatar is reported to have taken over 1 petabyte of local storage at Weta Digital for the rendering of the 3D CGI effects



Applications of Big Data

Industry	Data Sources	Applicability
Supply chain, logistics and manufacturing	RFID Sensors, Handheld scanners, On-board GPS Vehicle and shipment tracking ...	<ul style="list-style-type: none"> Route optimization Cost savings Operational efficiency
Online services and Web Analytics	Click Stream Data, Web Logs...	<ul style="list-style-type: none"> Online Customer behaviour Website usage
Financial Services	Stock market and Banking transaction data	<ul style="list-style-type: none"> Maximize trading opportunities Identify potentially fraudulent cases
Energy and utilities	Smart instrumentation such as “smart grids” and electronic sensors attached to machinery, oil pipelines	<ul style="list-style-type: none"> Uncover and fix potential problems before they result in costly or even disastrous failures
Media and Telecommunications	Data from streaming media, smartphone, tablets, Call Detail Records	<ul style="list-style-type: none"> Gain knowledge on user behaviour Prevent customer churn Improve service.
Health care and Life sciences	Medical Records	<ul style="list-style-type: none"> Provide patient treatment options Analyze data for clinical studies
Retail and consumer products	Sales Transaction data	<ul style="list-style-type: none"> Unearth patterns in user behaviour Brand monitoring with social networking data
BPO	Customer call details	<ul style="list-style-type: none"> Identify major problems customer face Frequency of customers looking for help.

Use cases for Big Data

Research & Development

- Use customer insights to eliminate unnecessarily costly features and add features which has a higher value for the customer.
- Improve gross margins

After-Sales Support

- Obtain real-time input on emerging defects and adjust the production process immediately.
- R&D operations could use these data for redesign, new product development

Police departments

- Target crime hotspots and prevent crime waves

Public utilities

- Usage of data from sensors on water & sewer usage
- Detect leaks and reduce water consumption


Electric power utilities

- Smart meters to better manage resources and avoid blackouts

BD is being used to predict traffic flow in Rio de Janeiro, which is hosting both the FIFA World Cup and the Olympics (2014 and 2016)

- Understanding Big Data- by Chris Eaton, Dirk Deroos, Tom Deutsch, George Lapis, Paul Zikopoulos
- Pentaho- <http://www.pentaho.com/big-data/>
- <http://www.retailsolutions.com/company/overview.php>
- http://en.wikipedia.org/wiki/Big_data
- http://www.gits.waseda.ac.jp/GITS/workshop/2013/docs/ICT_Wang.ppt

Thank You



TATA

Promise what we deliver.
Deliver what we promise. That's
certainty

Critical situations. Ruthless competition. Unforgiving customers. Thankfully you can be absolutely sure of your IT solutions with Tata Consultancy Services (TCS). As one of the world's fastest growing technology and business solutions providers, TCS has built a reputation of delivery excellence based on world class IT solutions that are on time, within budget and consistently deliver superior quality. So it comes as no surprise that we pioneered the concept of the Global Network Delivery Model, Developed Innovation Labs and Solution Accelerators. Achieving a level of delivery excellence that provides greater value to our customers and is the industry benchmark. Enabling our clients to experience certainty.

TATA CONSULTANCY SERVICES
Experience certainty.
IT Services • Business Solutions • Outsourcing

To learn how your business can experience certainty, visit www.tcs.com