

# Quick Introduction On Big Data and Hadoop

Tejas Saravana

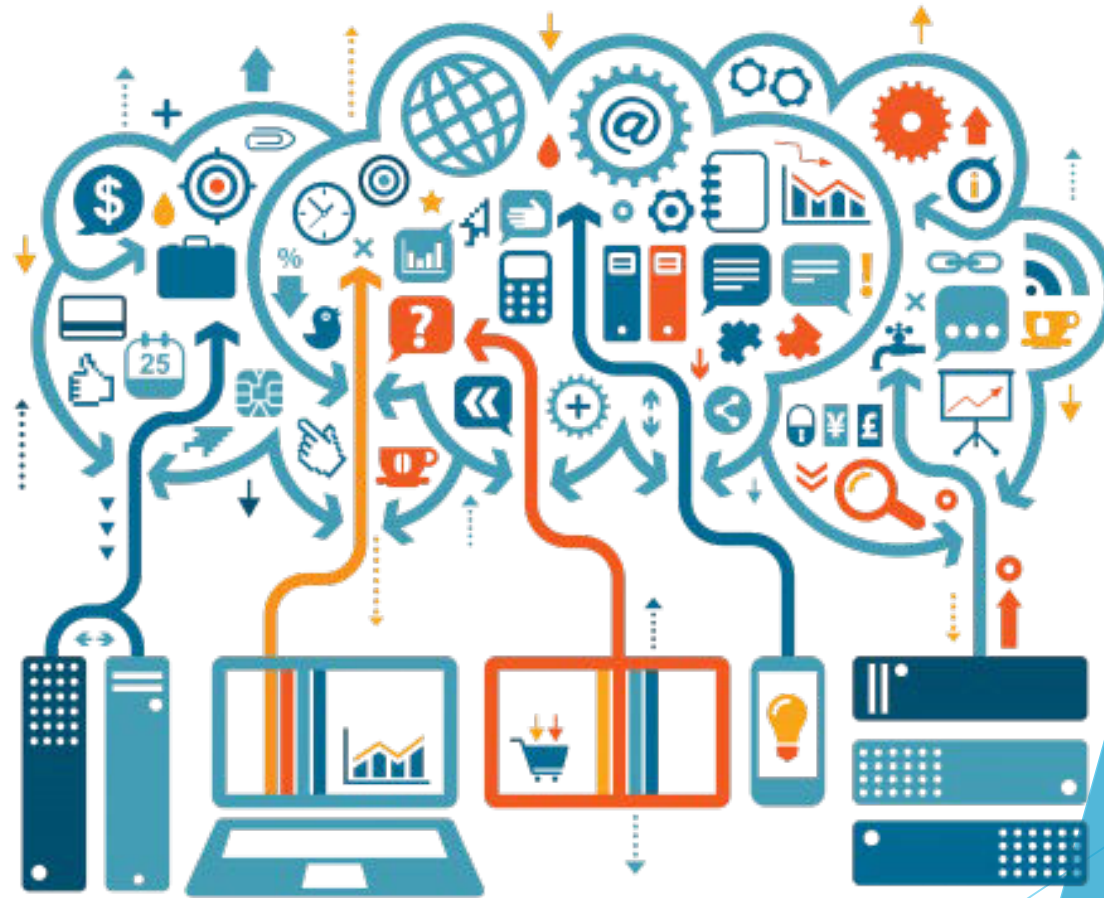
# Big Data

Just a popular term used to describe the exponential growth and availability of data

- ❖ Structured
- ❖ Semi Structured
- ❖ Un Structured

## Challenge

- ❖ Enormous volumes of data
- ❖ Many sources and types of data
- ❖ Pace at which data flows
- ❖ Abnormality in data



Big Data

# Big Data



Social media data

Click stream data



Connected Devices



Sensor data



# What It Can Tell You

---

Application Logs	User activity, fraud detection, application performance
Business Process Logs	Customer activity across channels, purchases, accounts
Call Detail Records	Billing, revenue assurance, customer assurance, marketing intelligence
POS Data	Buying pattern of customers

# Hadoop

---

Open source software platform managed by the Apache Software Foundation has proven to be very helpful in storing and managing vast amounts of data cheaply and efficiently.

Basically, it's a way of storing enormous data sets across distributed clusters of servers and then running "distributed" analysis applications in each cluster.

# Hadoop core Components and Daemons

## Two Components of Hadoop Master - Slave

1 HDFS(Storage)  
Name Node(Master)  
DataNode(Slave)

2 Yarn(Processing)  
ResourceManager(Master)  
NodeManager(Slave)

# Hadoop Cluster with Daemons



Master

Name Node  
Resource Manager

1 TB Hard Disk, 64GB RAM

5 Servers in Cluster

1 Master  
4 Slaves



Slave



Slave



Slave



Slave

Data Node  
Node Manager

6 TB Hard Disk  
16 GB RAM

# Hadoop Eco System

---

## Distributed Programming:

- Apache MapReduce
- Apache Pig
- Apache Hive
- Apache Spark



# Hadoop - The Swiss Knife Of 21<sup>st</sup> Century



Hadoop can be integrated with multiple analytic tools to get the best out of it, like M-Learning, R , Python, Spark, MongoDB etc.

**Thank You**

The right side of the slide features a decorative graphic composed of several overlapping, semi-transparent blue triangles and polygons. The colors range from a light sky blue to a deep navy blue, creating a layered, geometric effect.