CHANDIGARH UNIVERSITY

Discover. Learn. Empower.

Apex Institute of Technology

Department of Computer Science & Engineering

Bachelor of Engineering (Computer Science & Engineering)

Big Data Analytics and IoT– (CSD-432)

Prepared By: Ms. Bhanu Priyanka.Valluri (E8243)

DISCOVER . LEARN . EMPOWER

Big Data Analytics in IoT

Course Objective

The objective of this course is :

CO Number	Title	Level
CO1	To Understand the fundamental of big data analytics and computing in IoT domain.	Understand
CO2	To Acquire knowledge on data analytic tools and techniques.	Understand
CO3	To learn the practical implementation of big data analytics and solve the real time problem.	Understand

Will be covered in this lecture

Big Data Analytics in IoT

Course Outcome:

Upon successful completion of this course, students will be able to:

CO Number	Title	Level
CO1	Understand the big data analytics concepts with respective to IoT along with their challenges.	Understand
CO2	Explain the concepts of the development of smart systems.	Understand
CO3	Implement the use of big data tools to process IoT data in various fields of communication by find a solution.	Understand

Will be covered in this lecture

Welcome to the session of

Unit-1: ***Introduction to Big Data analytics in IoT Domain***

Chapter-2: A Role od BDA in IoT

...Activity...

The figure explain the importance or role of BDA in the developing technologies.

Consider and 2 from the adjacent fig which are relate to the field of production.

Agenda

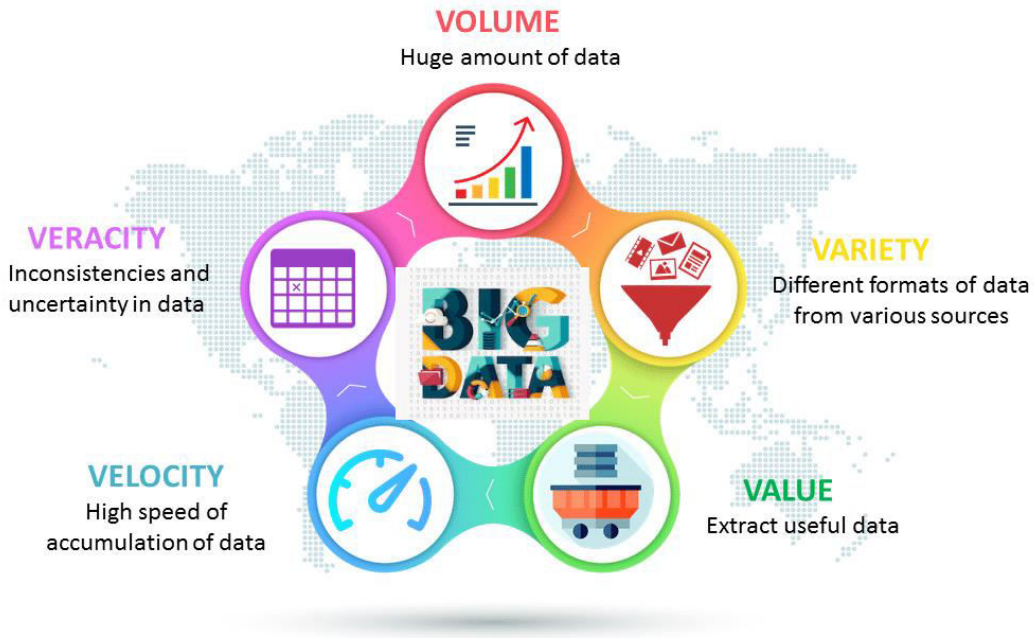
Chapter:2__Lecture: 4

- Introduction to BDA, IoT
- Hadoop ecosystems,
- Modules of Hadoop
- Big data process

Link: https://www.orangemantra.com/blog/java-plays-evolutionary-role-big-data-iot/

Introduction to BDA, IoT and Hadoop ecosystems

- The Big Data has been divided as per the five basic ingredients:
- volume, variety, velocity, veracity and value.

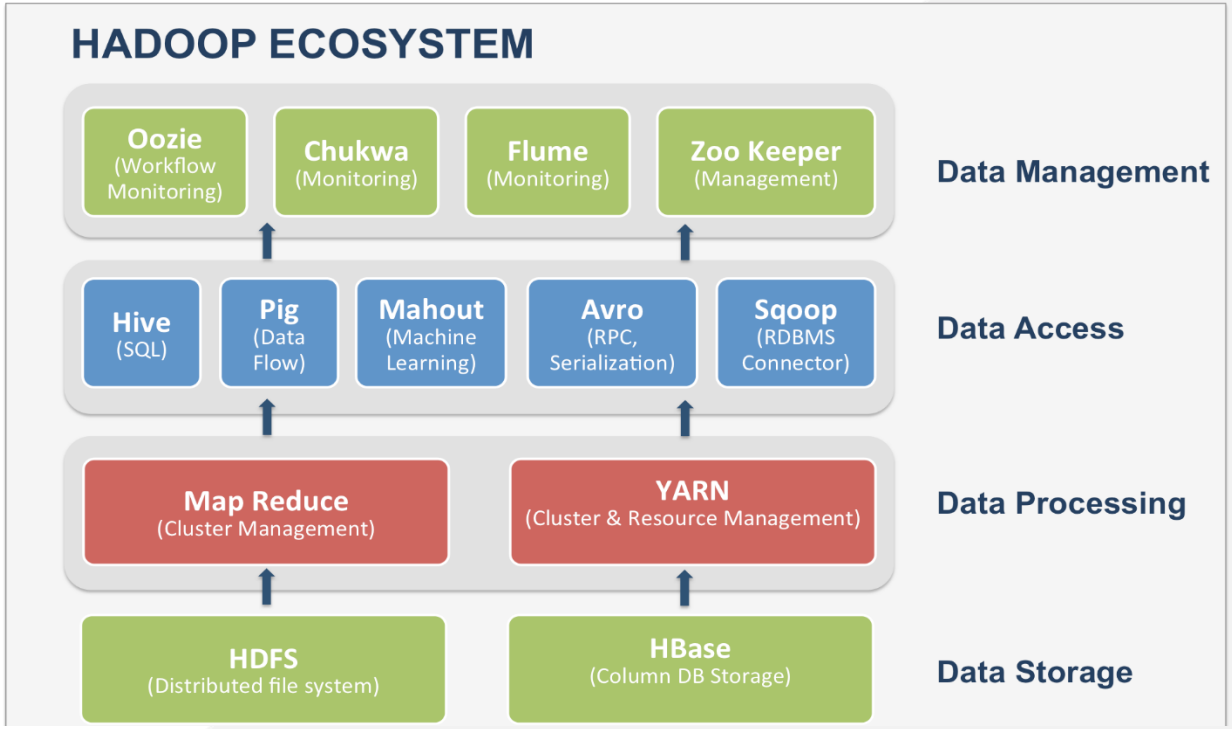


Link: https://www.edureka.co/blog/big-data-characteristics/

Introduction to BDA, IoT and Hadoop ecosystems

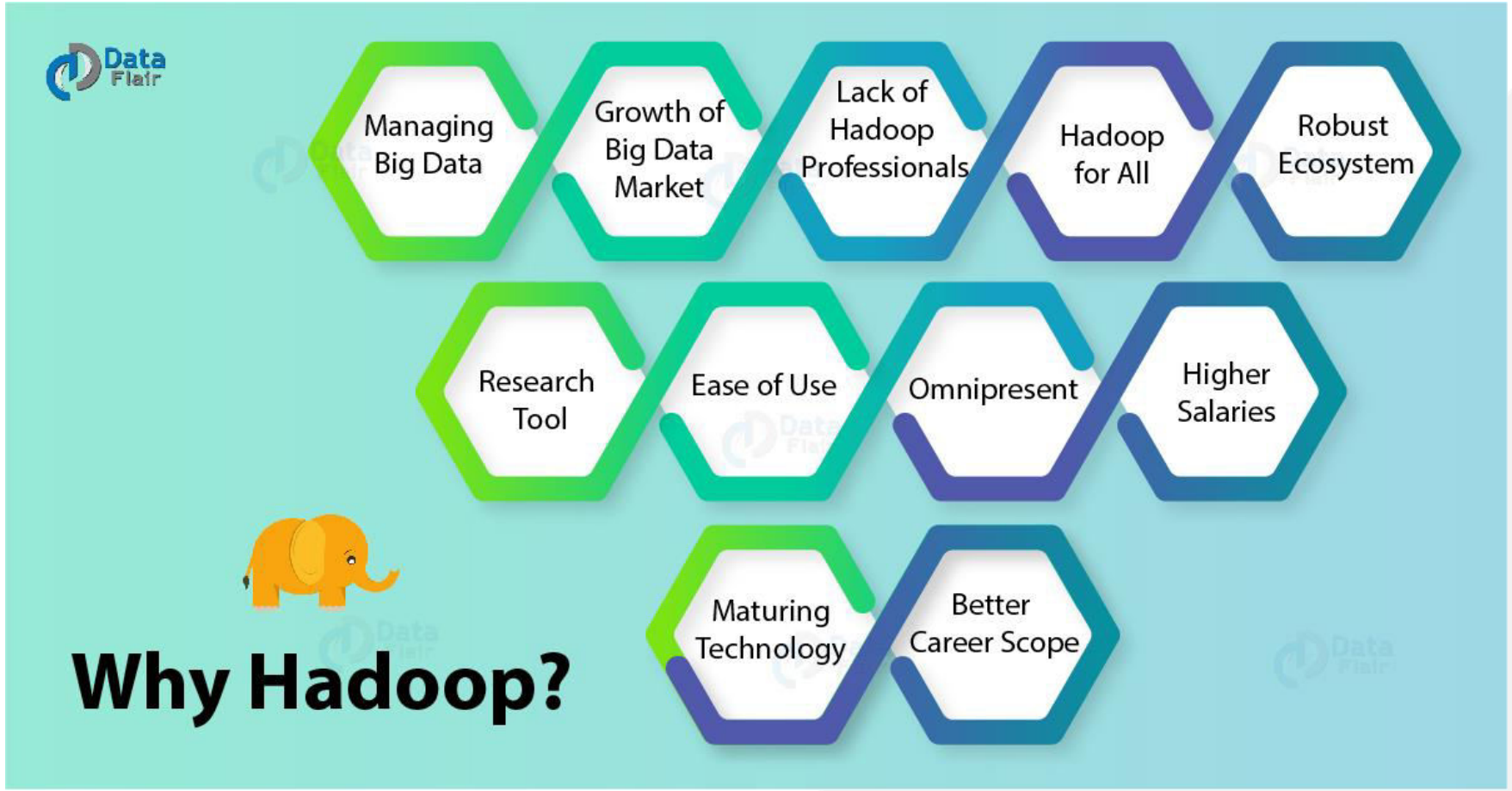
- The Hadoop is capable of taking care of several forms of unstructured as well as structured data, thus providing users better flexibility for gathering, processing as well as examining data collected than the data warehouses or relational databases afford.

Basically, Hadoop is made up of modules aimed for carrying out specific tasks.



Link: https://savvycomsoftware.com/what-you-need-to-know-about-hadoop-and-its-ecosystem/

What is Hadoop ?



Link: https://data-flair.training/blogs/why-hadoop/

Introduction to BDA, IoT and Hadoop ecosystems

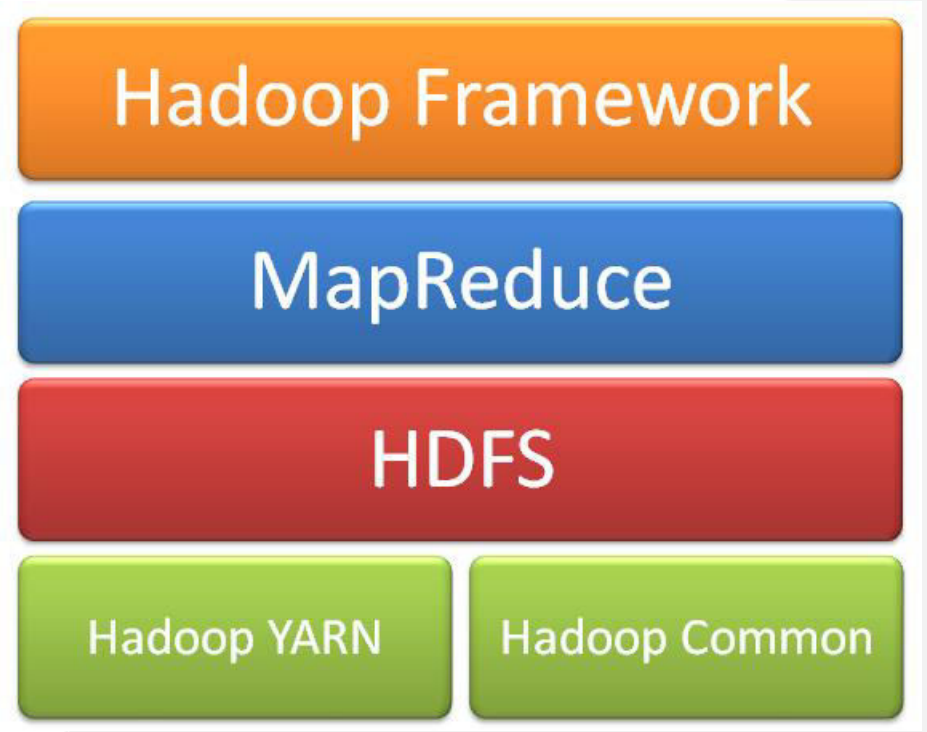
Following are the four modules associated with Hadoop system:

1. **Distributed File System Module**: The dispersed file organization permits data to be contained in a conveniently operable form through huge count of coupled stashed appliances.
2. **MapReduce Module**: This module has been named as per the two fundamental operations it carries out such as: reading data from the database and placing that in suitable format for the analysis purpose as well as carrying out analytical operations.

Introduction to BDA, IoT and Hadoop ecosystems

3. Hadoop Common: It affords the tool desired for the computer of the intended user.

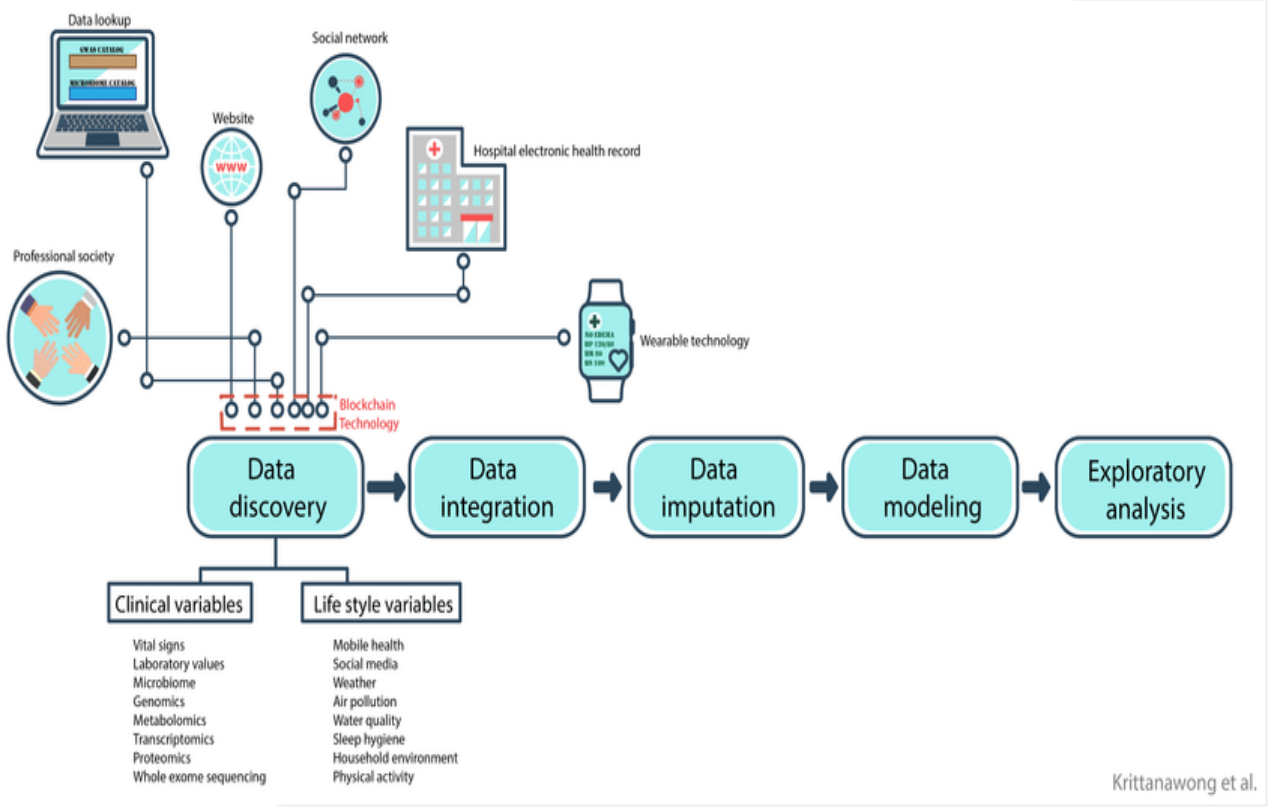
4. Hadoop Yarn Module: It is the last segment which manages the system resources preserving the data as well as executing the analysis.



Big Data Process

The Big Data process involves the following major phases

1. Data Extraction
2. Data Loading
3. Preprocessing
4. Data processing
5. Data Analysis
6. Data Transformation and Data Visualization



Link: https://www.researchgate.net/figure/Big-Data-Process-Flow-for-Cardiovascular-Medicine_fig1_327895555

Big Data Process

1. **Data Extraction**: Data extraction indicates the process or act of fetching data from data sources for subsequent storage or processing. The implication into the transitional organization is hence pursued by data metamorphosis as well as the possible inclusion of meta data before exporting to any other stage in the work flow.
2. **Data Loading**: The data loading represents the process that involves taking the transformed data and loading it where the users can access it.
3. **Preprocessing**: Data preprocessing is fundamentally a data mining approach which transforms raw data into more recognizable form. Practical data is mostly conflicting, deficient or lack in many features. Real world data may contain many errors. Therefore, for resolving these issues, data preprocessing is a proven method.

Big Data Process

4. Data processing: Data processing represents basically the collection and manipulation of data items for generating meaningful information. It may be considered as a sub set of information processing.

5. Data Analysis: Data analysis is basically a process of inspecting, transforming and modeling data with the aim to discover the useful information and assisting the decision making. Data analysis involves several strategies while compassing various techniques in science, business as well as social science domains.

6. Data Transformation and Data Visualization: Data transformation represents the process of converting data from one format or structure into another format or structure.

Summary

In todays session we tried understanding the below concept :

- ☞BDA + IoT
- ☞Hadoop ecosystems
- ☞What is Hadoop?
- ☞The different modules associate with Hadoop
- ☞Big data process and the sequential steps involved .

In the next session we will discuss about challenges, characteristics, big data ultimatums, types of data source

References:

- ✓ https://www.simplilearn.com/how-facebook-is-using-big-data-article?source=CTAexp
- ✓ https://www.icas.com/ca-today-news/10-companies-using-big-data
- ✓ https://www.bernardmarr.com/default.asp?contentID=1076
- ✓ Bryant, R.E., Katz, R.H., Lazowska, E.D.: Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society
- ✓ Sathi, A.: Implementation section (book 1). In: Big Data Analytics: Disruptive Technologies for Changing the Game, 1st ed. MC Press Online (2012)

Assessment Pattern

S.No.	Item	Number/semester	Marks
1	MSTs	2	20 per each
2	Quiz	2 per unit	4 per each quiz
3	Time bound surprise test	3 (one per unit)	12 per each test
4	Assignments	3 (one per unit)	10 per each Assignment
5	Engagement task (non gradable)	One per each topic	depends
6	Attendance + Engagement score	Above 90%	2
Internal (division as mentioned above points 1-6)			**40**
External			**60**
Total			**100**

THANK YOU

For queries
Email: bhanu.priyanka.e8243@cumail.in