# Assignment 2:

Q.1. What constitutes to Big Data?

a.volume

b.velocity

c.variety

d.All the above

ANSWER:The big data constitutes of all the three:Volume,variety and velocity.

2. Solution Identified to handle Big Data is?

a.Hadoop

b.storm

c.spark

d.All of these

ANSWER:Hadoop is the solution.Others are based on it.

3. Doug cutting added DFS &MapReduce support to:

a)Nutch

b)spark

c)GFS

d)facebook

ANSWER:He supported Nutch.

------------------------------------------------------------------------------------------------------------------------------------------

4.Which file system is followed by Hadoop?

a.DFS

b.GFS

c.HDFS

d.NTFS

ANSWER:Hadoop uses HDFS

------------------------------------------------------------------------------------------------------------------------------------------

5.How much data can be taken as a BigData in regards to volume?

a.Mega Bytes

b.Terabytes

c.Petabytes

d. b & c

ANSWER:B

6 .Hadoop can process data in Real time?

a.Yes

b.No

c.None of these

ANSWER:Yes.(I guess that is what Hadoop is meant for)

7 .Hadoop is developed using?

a. C

b. C++

c. Java

d. Ruby

ANSWER:Java

8 .Hadoop reduces cost of operations by?

a. using commodity Hardware

b. Using less number of nodes

c. Limited cluster size

d. None of these

ANSWER:d.None of these.It uses commodity hardware for Namenodes but for master node good h/w should be used.It cuts down cost because it reaches the data and then processes it.

9 .Search engine adopted in Hadoop is?

a.Google

b.yahoo

c.Solr

c.AOL

ANSWER:Solr

10. Stream processing can be achieved in Hadoop using?

a.Storm

b.Spark

c.Both of these

d.None of these

ANSWER:a.Storm

-------------------------------------------------------------------------------------------------------------------------------------------

11 What is the highest data unit we have at present?

a. petabyte

b. exabyte

c. terabyte

d. Brontobyte

ANSWER:Exabyte

12.Hadoop is well suitable for clusters where

a. Number of nodes keep changing

b. Number of nodes are constant

c. Most number of free nodes

d. None of these

ANSWER:b.Number of nodes keep changing

13 .what is batch processing?

a.processing of previously collected jobs in a single batch.

b.processing of previously collected jobs in different batches.

c.processing of jobs at that moment they are collcted

d.None of these

ANSWER:b.processing of previously collected jobs in different batches.

-----------------------------------------------------------------------------------------------------------------------------------

14 .what is stream processing?

a. Processing data as streams

b. Processing data in batches

c. Both of these

d. None of these

ANSWER:d.None of these

-------------------------------------------------------------------------------------------------------------------------------------

15 .Hadoop uses?

a. Batch processing

b. Near Real time processing

c. Real time processing

d. All the above

ANSWER:c.Real time processing.

Completed by

Shraddha bag