Q1.Describe the characteristics of Big data in detail.

2

A1.

3

Volume:

4

Volume is a key contributor to the problem of why traditional relational database management systems

5

(RDBMS, data warehouses as we know them today) fail to handle Big Data.  Underlying that failure are more complex issues

6

of cost, reliability, long query times, and their inability to handle new sources of unstructured or semi-structured data like text.

7

​

8

Variety:

9

Variety describes different formats of data that do not lend themselves to storage in structured relational database systems.

10

These include a long list of data such as documents, emails, social media text messages, video, still images, audio, graphs, and the

11

output from all types of machine-generated data from sensors, devices, RFID tags, machine logs, cell phone GPS signals,

12

DNA analysis devices, and more.  This type of data is characterized as unstructured or semi-structured and has existed all along.

13

​

14

Velocity:

15

This is one of the major determinants of Big Data storage, retrieval, analysis, and deployment architecture that companies

16

must work through today. When you sign on to Amazon or Netflix and see recommended purchases or views just for you the same process

17

has taken place.  The architecture of capture, analysis, and deployment must support real-time turnaround (in this case fractions of a

18

second) and must do this consistently over thousands of new visitors each minute.  Real Time Big Data Analytics (RTBDA) is one of the main

19

frontiers of development in Big Data today.

20

​

21

Value:

22

It is equally true of both big and little data that if we are making the effort to store and analyze it then it must

23

be perceived to have value.

24

​

25

Veracity:

26

It tells how trustworthy the data is that leads to the right direction while processing it.

27

​

28

Q2.Explain the possible solutions to handle Big data.

29

A2. There are two possible solutions to handle the big data,these are scale up and scale out.

30

1. scale up: 1.It increase the configuration of single file system like disk capacity,RAM, speed and data transfer,etc.

31

2.but it is costly,complex and time consuming process.

32

3.it is a monolithic system.

33

scale out: 1.use multiple commodity machines and distribute the load among these machines.

34

35

3. quickly implements its solution among its machines.

36

4. it is not monolithic system instead increases its number of machines.

37

Q3. Explain the differences between scaling up and scaling out.

38

A3.Scale out:(Horizontal scaling)

39

1.Horizontal scaling means that you scale by adding more machines into your pool of resources.

40

2.It is often based on partitioning of the data i.e. each node contains only part of the data.

41

3.With horizontal-scaling it is often easier to scale dynamically by adding more machines into the existing pool.

42

Scale out:(vertical scaling)

43

1.Vertical scaling means that you scale by adding more power (CPU, RAM) to an existing machineIn a database world.

44

2.In vertical-scaling the data resides on a single node and scaling is done through multi-core.

45

3.It is often limited to the capacity of a single machine, scaling beyond that capacity often involves downtime and comes with an

46

upper limit.

47