**ASSIGNMENT 1.4**

1. **Characteristics of Big data:**

The characteristics of bigdata are described by 5 v’s namely

* Volume:The quantity of stored data.The size of the data which actually determines whether the it is a big one or small one. The vast amount of data generated every second .It is stored in Zetabytes or Brunobytes.
* Velocity:  The speed at which the data is generated and processed to meet the demands and challenges that lie in the path of growth and development without any time lagging in between.
* Variety: The type and nature of the data.The data may be structured(tables),semi-structured(set of logs) and unstructured(youtube).
* Veracity:The trustworthiness or messiness of the data in terms of accuracy. The quality of captured data can vary greatly, affecting accurate analysis which is overcame by analysis.
* Value:The outcome of processing of data is refered to as value.without value having bigadata is of no use.

1. **The possible solutions to handle Big data:**

* Scale Up: It refers to increasing the computing resources on a single node like disk capacity, RAM, data transfer speed.

But it is complex,costly and time consuming.

* **Scale Out**:It refers to distribution of resources in various systems by increasing the nodes like using many commodity hardwares.

It is economical and quick to implement.

1. **The differences between scaling up and scaling out:**

*The root of the difference is the existence (or lack) of a shared address space.*

Scaling up:

* It is known as concurrent programming,
* It is one of the most common methods for utilizing multi-core architecture in the context of a single application.
* It is done through multi-threading and in-process message passing.

Scaling out:

* It also known as distributed programming,
* Its does something similar by distributing jobs across machines over the network.
* There are different patterns associated with this model such as Master/Worker, Tuple Spaces, BlackBoard, and Map/Reduce.