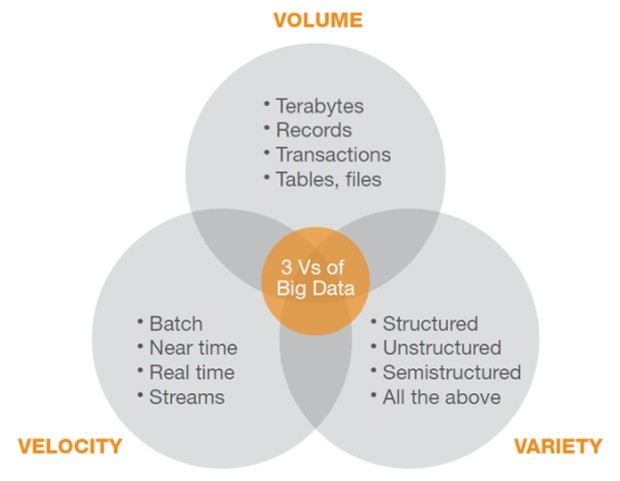
# Characteristics Of 'Big Data:

***Volume*** – The name 'Big Data' itself is related to a size which is enormous. Size of data plays very crucial role in determining value out of data. Also, whether a particular data can actually be considered as a Big Data or not is dependent upon volume of data. Hence, 'Volume' is one characteristic which needs to be considered while dealing with 'Big Data'.

***Variety*** – The next aspect of 'Big Data' is its variety. Variety refers to heterogeneous sources and the nature of data, both structured and unstructured. During earlier days, spreadsheets and databases were the only sources of data considered by most of the applications. Now days, data in the form of emails, photos, videos, monitoring devices, PDFs, audio, etc. is also being considered in the analysis applications. This variety of unstructured data poses certain issues for storage, mining and analyzing data.

***Velocity*** – The term 'velocity' refers to the speed of generation of data. How fast the data is generated and processed to meet the demands, determines real potential in the data. Big Data Velocity deals with the speed at which data flows in from sources like business processes, application logs, networks and social media sites, sensors, mobile devices, etc. The flow of data is massive and continuous.



# Explain the possible solutions to handle Big data:

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

***Scale up:***

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

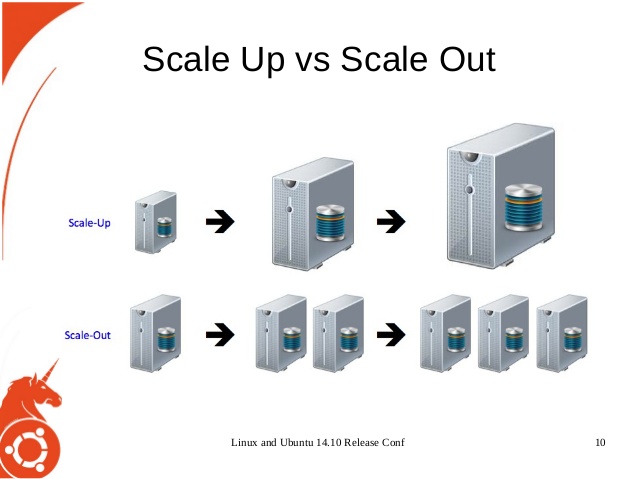
2. But it is costly, complex and time consuming process.

3. It is a monolithic system.

***Scale out:***

1. Use multiple commodity machines and distribute the load among these machines.

2. Quickly implements its solution among its machines.

3. It is not monolithic system instead increases its number of machines.  
  
  


# Explain the differences between scaling up and scaling out:

***Scale up:***

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

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

3. It is a monolithic system.

***Scale out: (Horizontal scaling)***

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

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

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

***Scale out: (vertical scaling)***

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

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

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

