**HDFS:**

HDFS stands for Hadoop Distributed File System.

It is mainly used for storage purpose in hadoop system.

HDFS files are stored across the more number of clusters.

Each files are divided into blocks. The blocks are stored into more number of servers.

This system has replication of files which is used to achieve the scalable and fault tolerance properties.

Fault tolerance means if one system fails it will not affect the entire cluster. We can easily retrieve the data from replication blocks.

HDFS is very less expensive because of scalable and fault tolerance properties.

HDFS contains following features:

It is more suitable for very large files. For example mega bytes, giga bytes, tera bytes, peta bytes.

In HDFS, tha data is written only once but we can read it many times.

It contains commodity hardware system. This system is less expensive.

**Hadoop Cluster**:

Hadoop cluster is a commodity hardware type which is used to contain the large amount of unstructured data across the more number of servers.

Usually commodity hardware is less expensive which is achieved among the distribured environment.

It is mainly for analysing and storing the data.

It has one top node which is called name node. It contains all the information about the remaining nodes.

It contains another machine which is called job tracker. These nodes are called master nodes. The remaining nodes which are in the cluster is called data node and task tracker.

The replication of data is available which is used to achieve the fault tolerance and scalable.

**HDFS Blocks:**

In HDFS, more number of nodes are available in single cluster. Each node has more number of blocks which is used to store the data. Each block has 128 MB.

The blocks are fixed size. Each files are stored in each blocks. The size of the file is less than block size then the data is easily stored in block.

HDFS achieves fault tolerance and availability because of replication.

One data is stored into more than one blocks. So the data is retrieved easily.

The data in these blocks are stored in contiguous manner