ASSESSMENT-17 Vinutha s

22/12/2023

**Apache Spark – Introduction**

Spark was introduced by Apache Software Foundation for speeding up the Hadoop computational computing software process.

As against a common belief, **Spark is not a modified version of Hadoop** and is not, really, dependent on Hadoop because it has its own cluster management. Hadoop is just one of the ways to implement Spark.

Spark uses Hadoop in two ways – one is **storage** and second is **processing**. Since Spark has its own cluster management computation, it uses Hadoop for storage purpose only.

1. Evolution of Apache Spark
2. Features of Apache Spark
3. Spark Built on Hadoop
4. Components of Spark

**Architecture of Apache Spark**

