

Vasavi College of Engineering

Ibrahimbagh -31

Department of Computer Science and Engineering

A Report on 1 day Student Training Program on “**Big Data Analytics - Hadoop**” conducted on Sept 3rd, 2013

(Conducted Under TEQIP-II)

A Report on Student Training Program on “**Big Data Analytics- Hadoop**” was organized under TEQIP-II on Sept 3rd, 2013 for students audiences of 3/4 CSE in Seminar Hall, Ramanujan Block ,College campus.

The training program was conducted in two sessions by P. Kumar, Project Lead, Webcognize Softtech Pvt. Ltd.

The objective of the training program was to enlighten the students with concepts of

- Hadoop Distributed File System.
- Hadoop Echo System

Technical Session -1



Speaker Mr. P. Kumar has started off session 1 with Network File System followed by discussion on Hadoop Distributed File System with the greater emphasis on integration of HDFS with Hadoop MapReduce which overcomes the vulnerabilities faced by NFS.

Basically MapReduce is a programming model designed for processing large volumes of data in parallel by dividing the work into a set of independent tasks.

MapReduce programs are written in a particular style influenced by functional programming constructs, specifically idioms for processing lists of data.

MapReduce programs transform lists of input data elements into lists of output data elements. The first phase of a MapReduce program is called mapping. Reducing lets you aggregate values together. Reducing is often used to produce "summary" data, turning a large volume of data into a smaller summary of itself.

Technical session -2

Speaker Mr. P. Kumar has started off session 2 with Zookeeper Service which basically focuses on high availability of information. This high availability is achieved by making Zookeeper runs on set of several machines and allows storing of large amount of information in central location in order to provide high availability.



The distributed system components first contact ZooKeeper to get the definitive filename, and then fetch that file for the configurations. Before concluding the session Speakers have also shared the link to download the version of Hadoop with students.