# Big Data Technologies

## Agenda

- Distributed systems
- Hadoop

## Hadoop Installation

**Local Installation**

- Installation steps
    - step 1: Download and extract Hadoop.
    - step 2: In .bashrc, set JAVA_HOME, HADOOP_HOME and PATH (for HADOOP_HOME/bin & HADOOP_HOME/sbin).
- Can access only local file system.
- terminal> hadoop fs -ls /
- terminal> hadoop fs -ls /home/
- terminal> hadoop fs -ls /home/nilesh
- terminal> hadoop fs -help
- terminal> hadoop fs -cat /home/nilesh/sep22/dbda/bigdata/day02/day02.md

**Single Node Cluster Installation**

- Installation steps: As given in PDF
- terminal> hdfs namenode -format
    - ~/bigdata/hd-data/nn/current --> fsimage_0000000000000000000 (metadata)
- terminal> start-dfs.sh
- terminal> jps
- terminal> start-yarn.sh
- terminal> jps
- terminal> stop-yarn.sh

- terminal> jps
- browser: http://localhost:9870
  - Summary
  - Data Nodes
  - Utilities -> Browse File systems
- terminal> hadoop fs -mkdir /user
- terminal> hadoop fs -mkdir /user/nilesh
- terminal> cd /tmp
- terminal> cat > welcome.txt
  - Create a new file on local machine
- terminal> hadoop fs -put welcome.txt /user/nilesh
  - Upload that file on HDFS
- terminal> rm welcome.txt
  - Delete that file from local machine.
- terminal> hadoop fs -ls /user/nilesh
- terminal> hadoop fs -cat /user/nilesh/welcome.txt
- terminal> hadoop fs -cat /user/nilesh/welcome.txt
- terminal> hadoop fs -head /user/nilesh/welcome.txt
- terminal> hadoop fs -get /user/nilesh/welcome.txt
- terminal> cat welcome.txt
- terminal> stop-dfs.sh