

# INTRODUCTION TO BIG DATA

ECAP456

Dr. Rajni Bhalla  
Associate Professor

# Learning Outcomes



After this lecture, you will be able to

- learn Pre-installation Setup

# Introduction

Before installing Hadoop into the Linux environment, we need to set up Linux using ssh (Secure Shell). Follow the steps given below for setting up the Linux environment.

1. Creating a User
2. SSH Setup and Key Generation
3. Installing Java
4. Downloading Hadoop
5. Hadoop Operation Modes

# Introduction

Before installing Hadoop into the Linux environment, we need to set up Linux using ssh (Secure Shell). Follow the steps given below for setting up the Linux environment.

6. Installing Hadoop in Standalone Mode
7. Installing Hadoop in Pseudo Distributed Mode
8. Verifying Hadoop Installation

# Creating a User

- Open the root using the command “su”.
- Create a user from the root account using the command “useradd username”.
- Now you can open an existing user account using the command “su username”.

# Commands

```
$ su
```

```
password:
```

```
# useradd hadoop
```

```
# passwd hadoop
```

```
New passwd:
```

```
Retype new passwd
```

# SSH Setup and Key Generation

- SSH setup is required to do different operations on a cluster such as
  - starting, stopping, distributed daemon shell operations.
- To authenticate different users of Hadoop, it is required to provide public/private key pair for a Hadoop user and share it with different users.

# Command

```
$ ssh-keygen -t rsa
```

```
$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

```
$ chmod 0600 ~/.ssh/authorized_keys
```

# Installing Java

Java is the main prerequisite for Hadoop. First of all, you should verify the existence of java in your system using the command “java -version”. The syntax of java version command is given below.

```
$ java -version
```

# Output

**java version "1.7.0\_71"**

**Java(TM) SE Runtime Environment (build 1.7.0\_71-b13)**

**Java HotSpot(TM) Client VM (build 25.0-b02, mixed mode)**

If java is not  
installed in your  
system, then  
follow the steps  
given below for  
installing java.

# Step 1

Download java (JDK <latest version> - X64.tar.gz)  
by visiting the following link [www.oracle.com](http://www.oracle.com)

# Step 2

- \$ cd Downloads/
- \$ ls
- jdk-7u71-linux-x64.gz
- \$ tar zxf jdk-7u71-linux-x64.gz
- \$ ls
- jdk1.7.0\_71 jdk-7u71-linux-x64.gz

# Step 3

- \$ su
- password:
- # mv jdk1.7.0\_71 /usr/local/
- # exit

# Step 4

- `export JAVA_HOME=/usr/local/jdk1.7.0_71`
- `export PATH=$PATH:$JAVA_HOME/bin`

Now apply all the changes into the current running system.

- `$ source ~/.bashrc`

# Step 5

Use the following commands to configure java alternatives

# Commands

- # alternatives --install /usr/bin/java java  
usr/local/java/bin/java 2
- # alternatives --install /usr/bin/javac javac  
usr/local/java/bin/javac 2
- # alternatives --install /usr/bin/jar jar  
usr/local/java/bin/jar 2

# Commands

- # alternatives --set java usr/local/java/bin/java
- # alternatives --set javac usr/local/java/bin/javac
- # alternatives --set jar usr/local/java/bin/jar



**That's all for now...**