

Big Data Practice
Environment

CONFIGURING SINGLE NODE HADOOP CLUSTER INSTALLATION SETUP



Steps To Build

Get your Cloud

Sign up for any of the Cloud Services

01

Understand Compute

2 VCPUs with 8GB of RAM
22GB of RAM

02

Get ubuntu on Compute

Setup passwordless login inside Linux

03

04

Setting up Python

Setting up Pyspark and Jupyterlab and ensure connectivity

05

Download installers

Hadoop, Hive, Spark3 and Postgres Docker

06

Initiate the installation

Requires good grasp of Linux commands

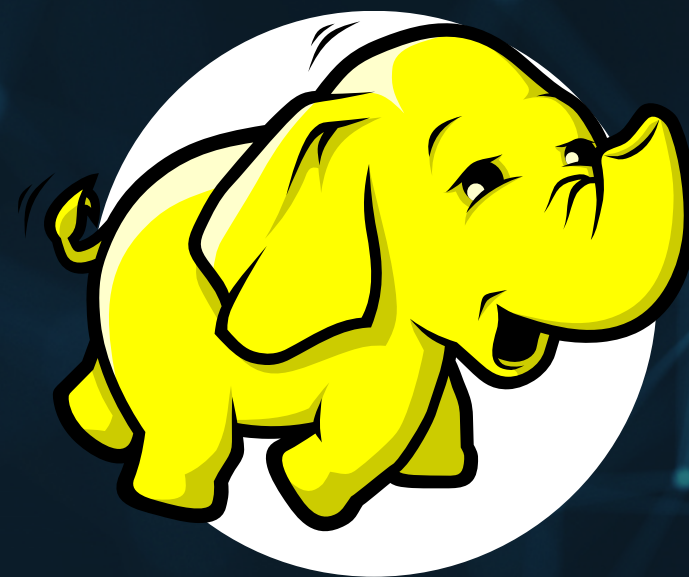
A decorative graphic on the left side of the slide. It features a stack of dominoes, some with white dots and others with red dots, arranged in a cascading fashion. Below the dominoes, there are white circuit-like lines on a dark blue background. The background of the entire slide is dark blue with a network of glowing blue lines and dots.

Steps & Actions

- **Download Hadoop, Hive and Spark executables**
- **Start HDFS installation and Format**
- **Complete Hive installation**
- **Complete Spark Installation with Pyspark3**
- **Configure JupyterLab**

The Ecosystem

Its all written by humans for humans...



File System

HADOOP



Query Engine

Spark



Programming

Python



THANKS FOR WATCHING

PRACTICE

PRACTICE

 **LIKE**

 **SHARE**

 **SUBSCRIBE**

PRACTICE

PRACTICE