



Hadoop Installation

HADOOP big data processing

Modes

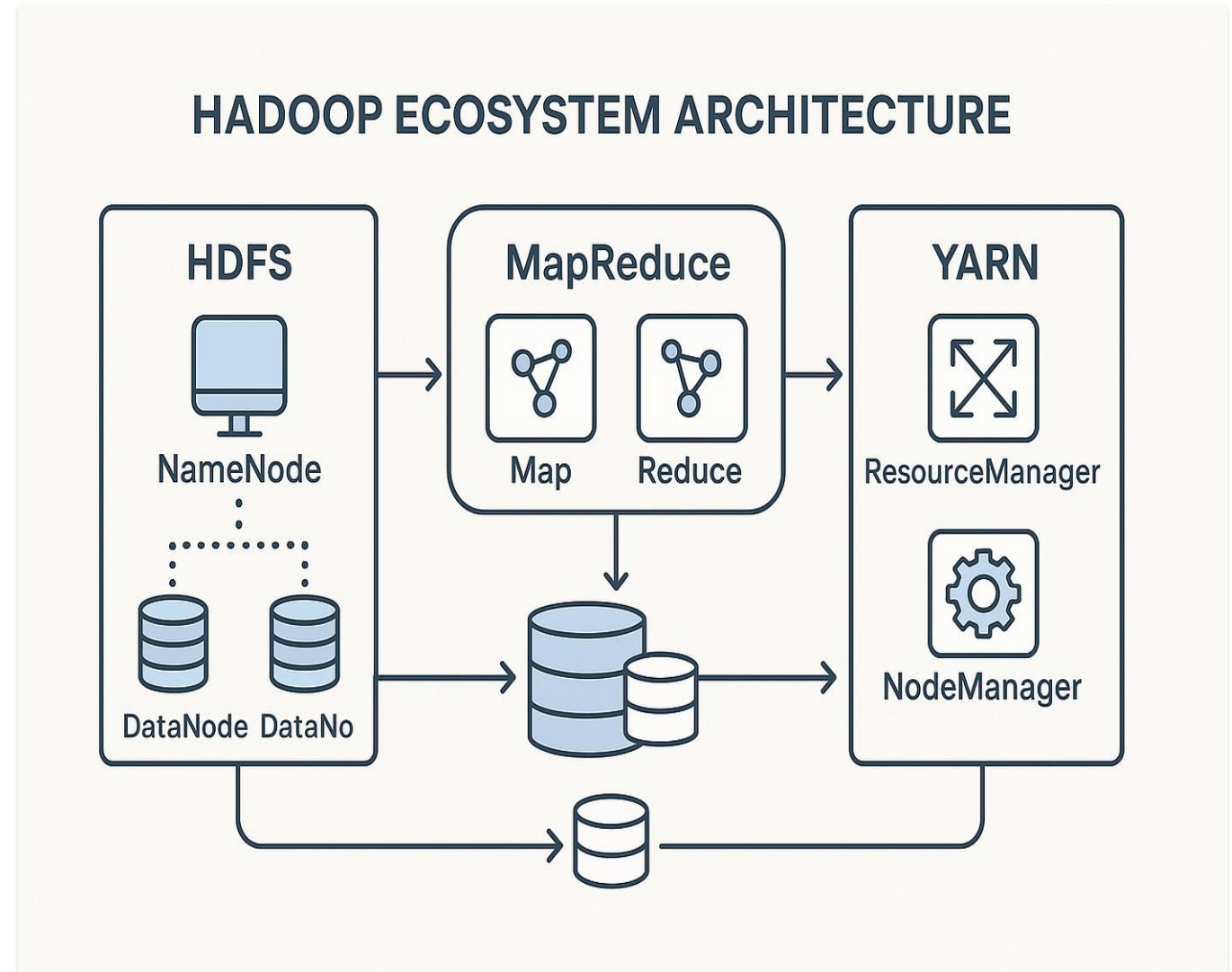
Understanding Different Deployment Strategies



What is Hadoop?

Apache Hadoop Framework


- Open-source distributed computing framework
- **HDFS** - Distributed file system
- **MapReduce** - Processing model
- **YARN** - Resource management



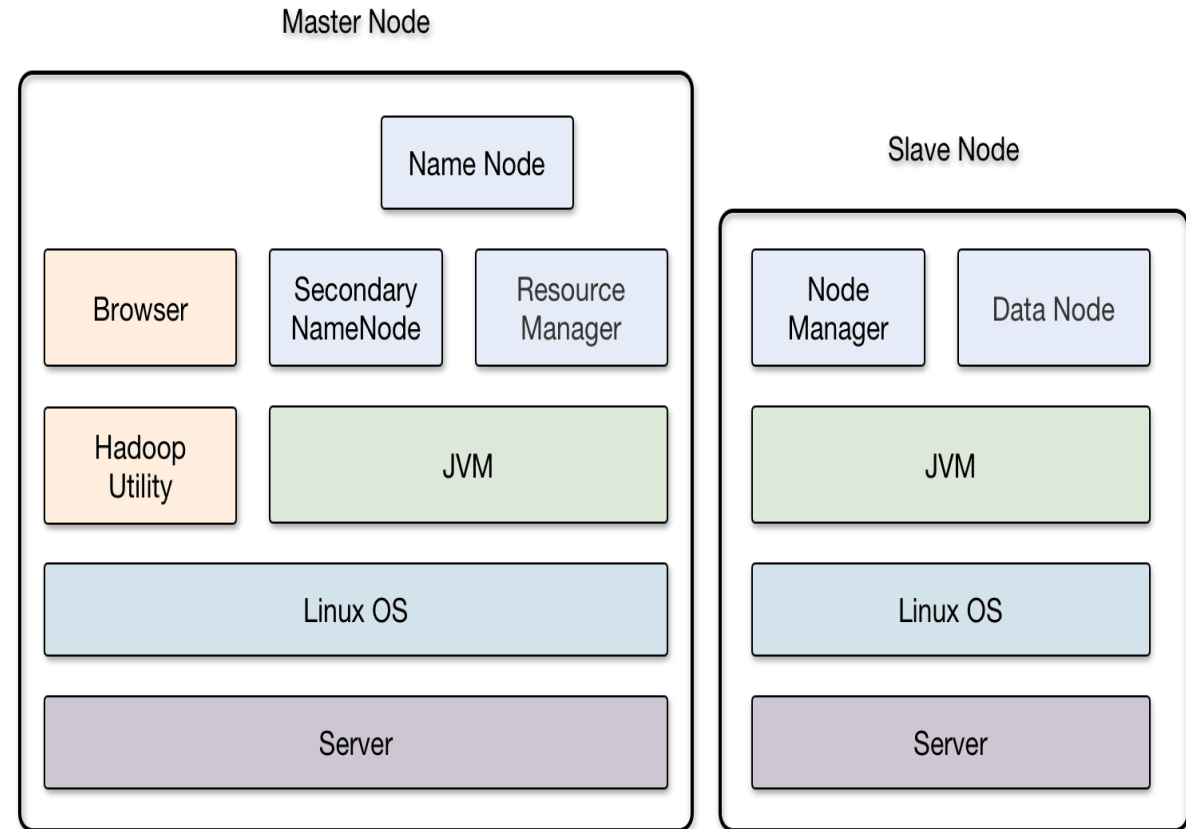
Installation Modes Overview

Three Main Modes

 **Standalone**
Single Machine

 **Pseudo-Distributed**
One JVM, All Daemons

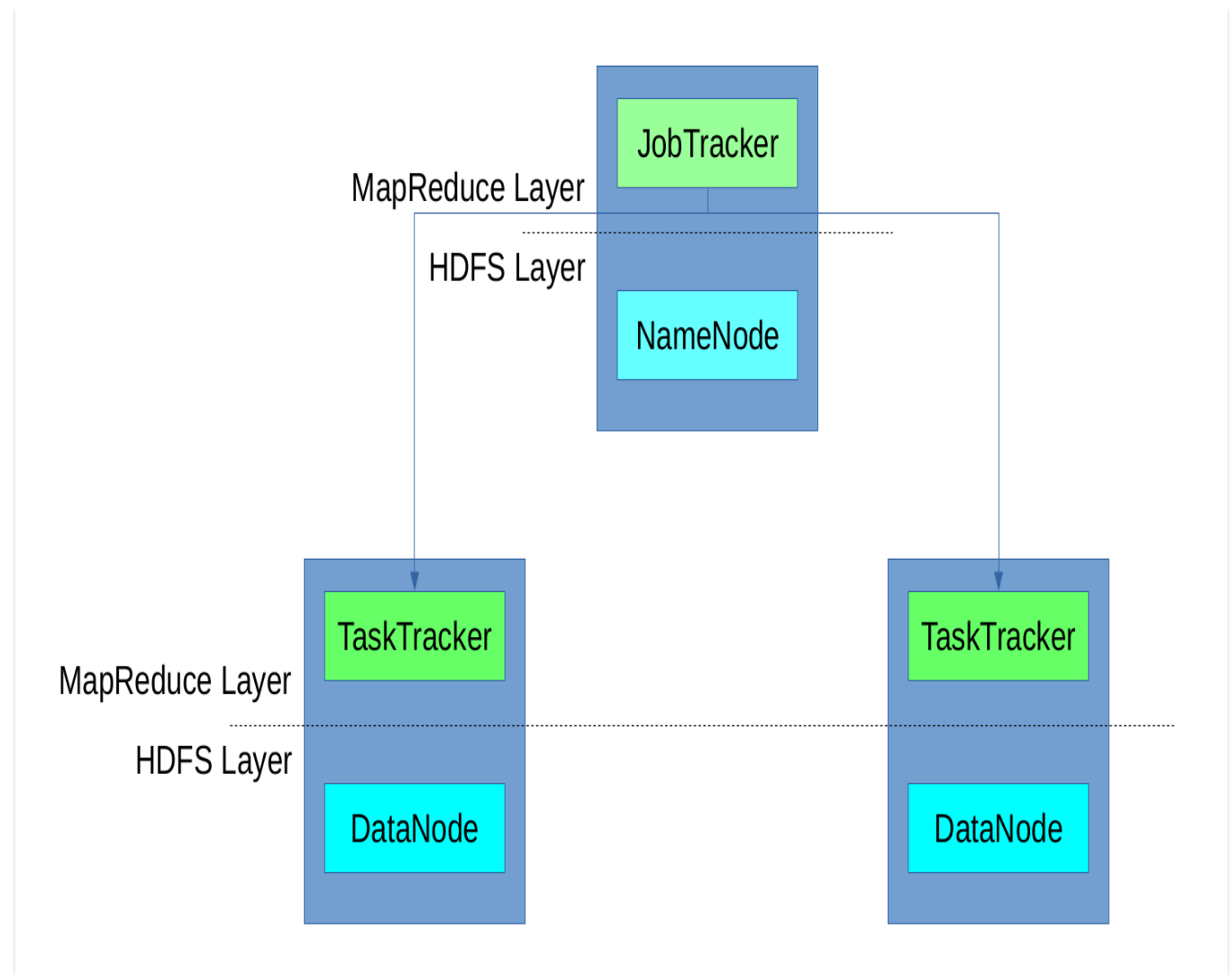
 **Fully-Distributed**
Multi-Node Cluster



Standalone Mode

Key Features

- Single machine deployment
- Non-distributed environment
- **Development & Testing**
- Minimal configuration
- Quick setup process



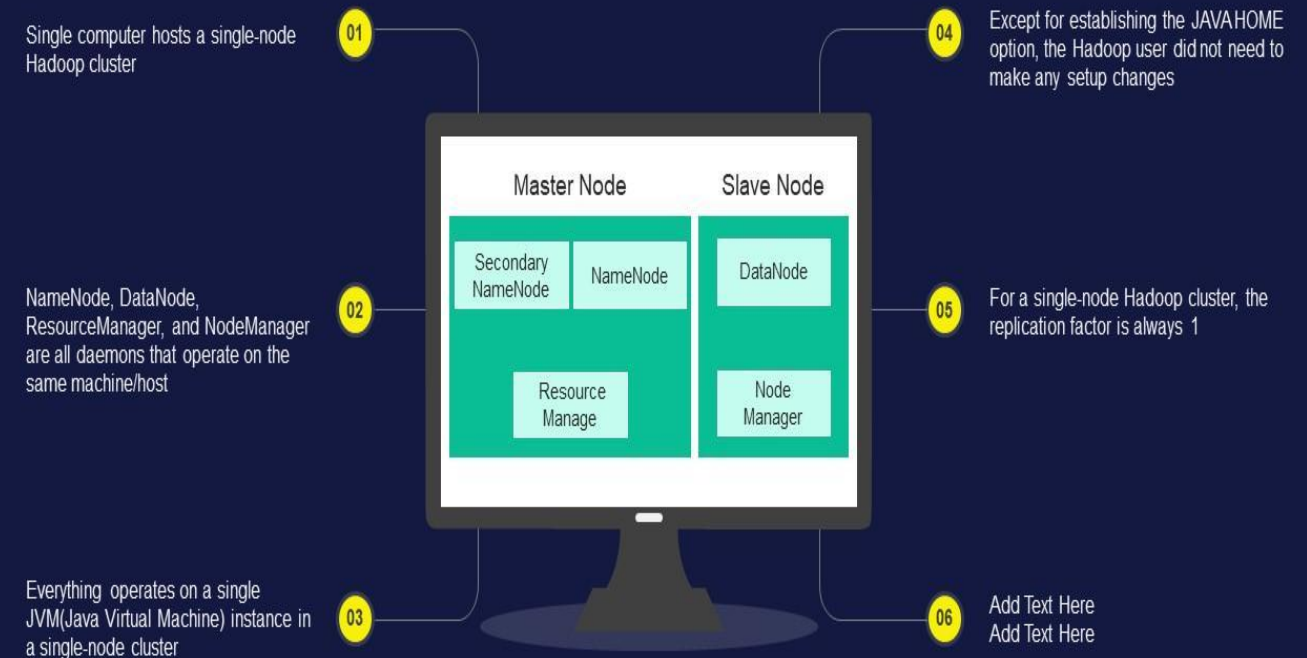
Pseudo-Distributed Mode

Key Features

- Single machine deployment
- Multiple daemon processes
- All processes on **one JVM**
- Simulates cluster environment
- Suitable for **testing applications**

Single Node Hadoop Cluster or Pseudo Distributed Mode

This slide depicts the single-node Hadoop cluster and its various components such as master node, slave node, secondary Name Node, resource manager, and Data Node.



This slide is 100% editable. Adapt it to your needs & capture your audience's attention.

Fully-Distributed Mode

Key Features

- Multiple nodes deployment
- Master-Slave Architecture**
- Production environment ready
- High availability**
- Scalable data processing

Multi Node Hadoop Cluster or Fully Distributed Mode

This slide depicts the multi-node Hadoop cluster, and its structure, including its deployment, architecture style, nodes that run on master and slave nodes.

01

Implemented on multiple computers

05

DataNodes and NodeManagers are daemons that operate on slave nodes which are low-cost commodity equipment

02

Multi-node Hadoop cluster's daemons are all up and running on separate machines/hosts

06

In a multi-node Hadoop cluster, slave computers can be located anywhere, regardless of where the master server is physically located

03

Architecture of a multi-node Hadoop cluster is master-slave

07

Add text here

04

On the master nodes, which are high-end computer machines, the daemons Namenode and ResourceManager run

08

Add text here

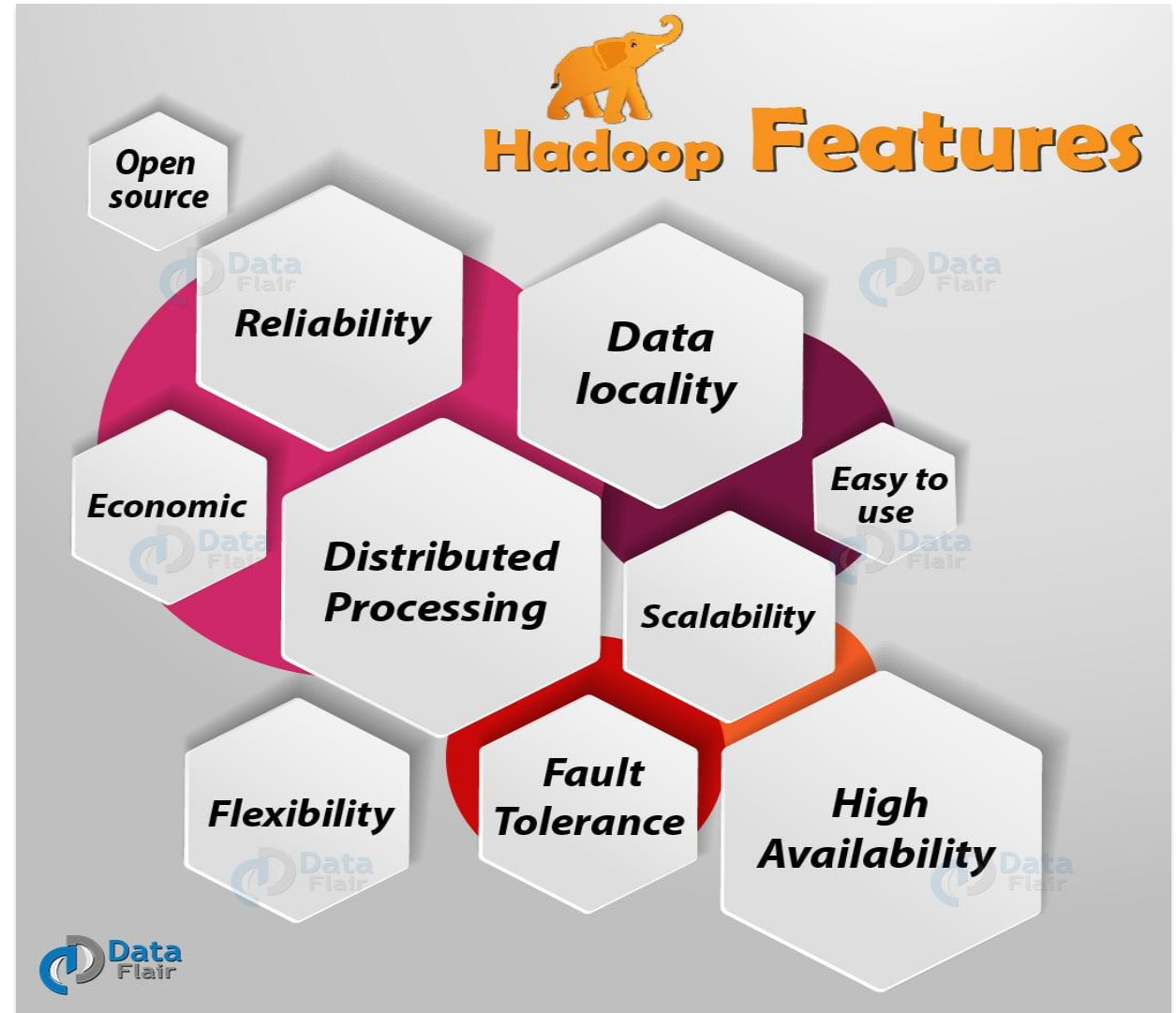
This slide is 100% editable. Adapt it to your needs & capture your audience's attention.

Comparison Table

Installation Modes Comparison

Mode	Nodes	Use Case	Scale
Standalone	Single	Dev/Test	None
Pseudo-Dist.	One JVM	Testing	Limited
Fully-Dist.	Multiple	Production	High

- Standalone: Single machine, minimal config
- Pseudo-Distributed: Simulated cluster on one JVM
- Fully-Distributed: Multi-node cluster, production ready



Hadoop Installation Modes Mind Map

Installation Modes Hierarchy

Standalone

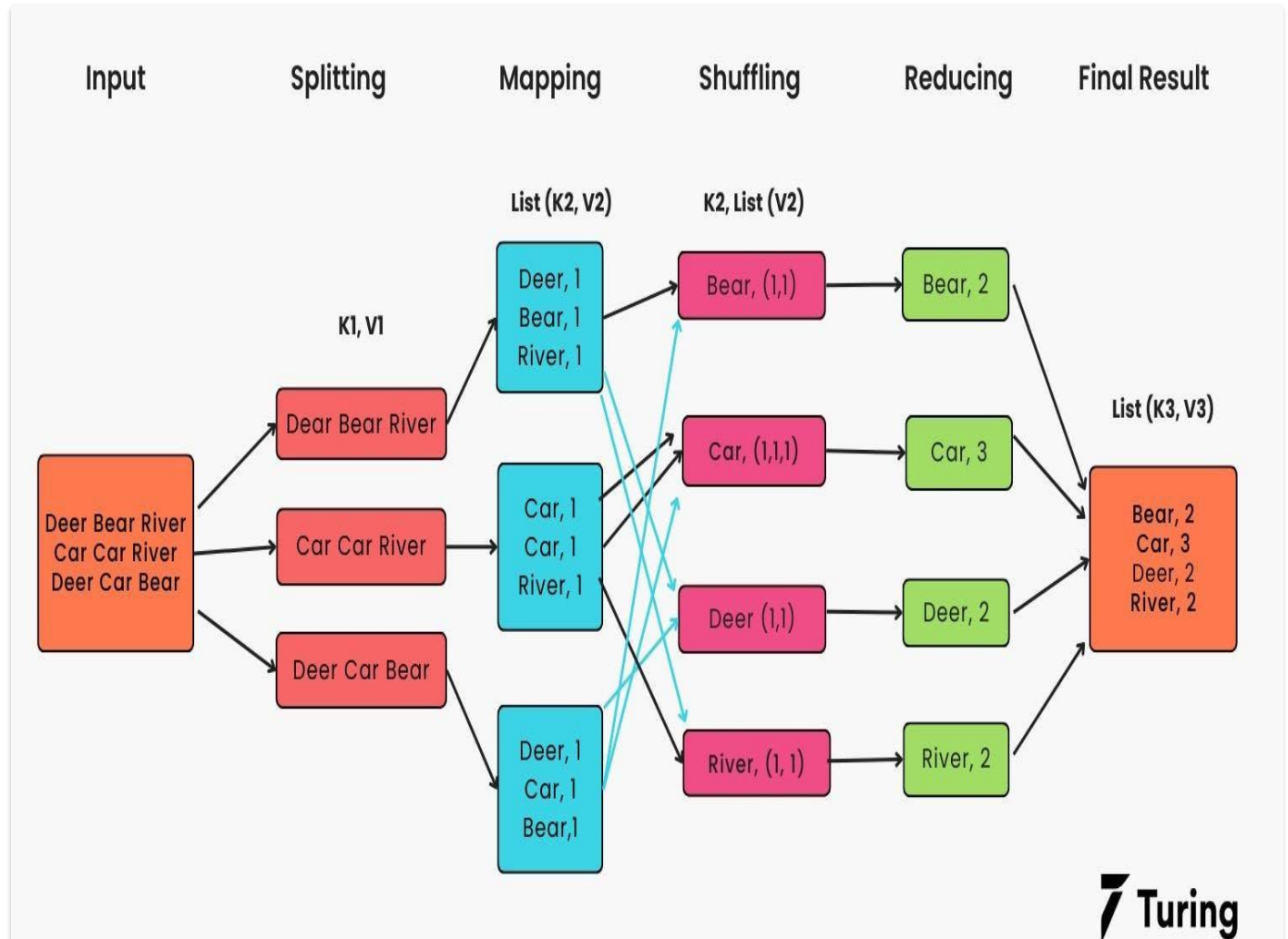
- Single machine
- Development & Testing
- Minimal configuration

Pseudo-Distributed

- Single machine, multiple daemons
- One JVM
- Simulated cluster

Fully-Distributed

- Multi-node cluster



Key Points Summary



Standalone

For learning & development



Pseudo-Distributed

For testing applications

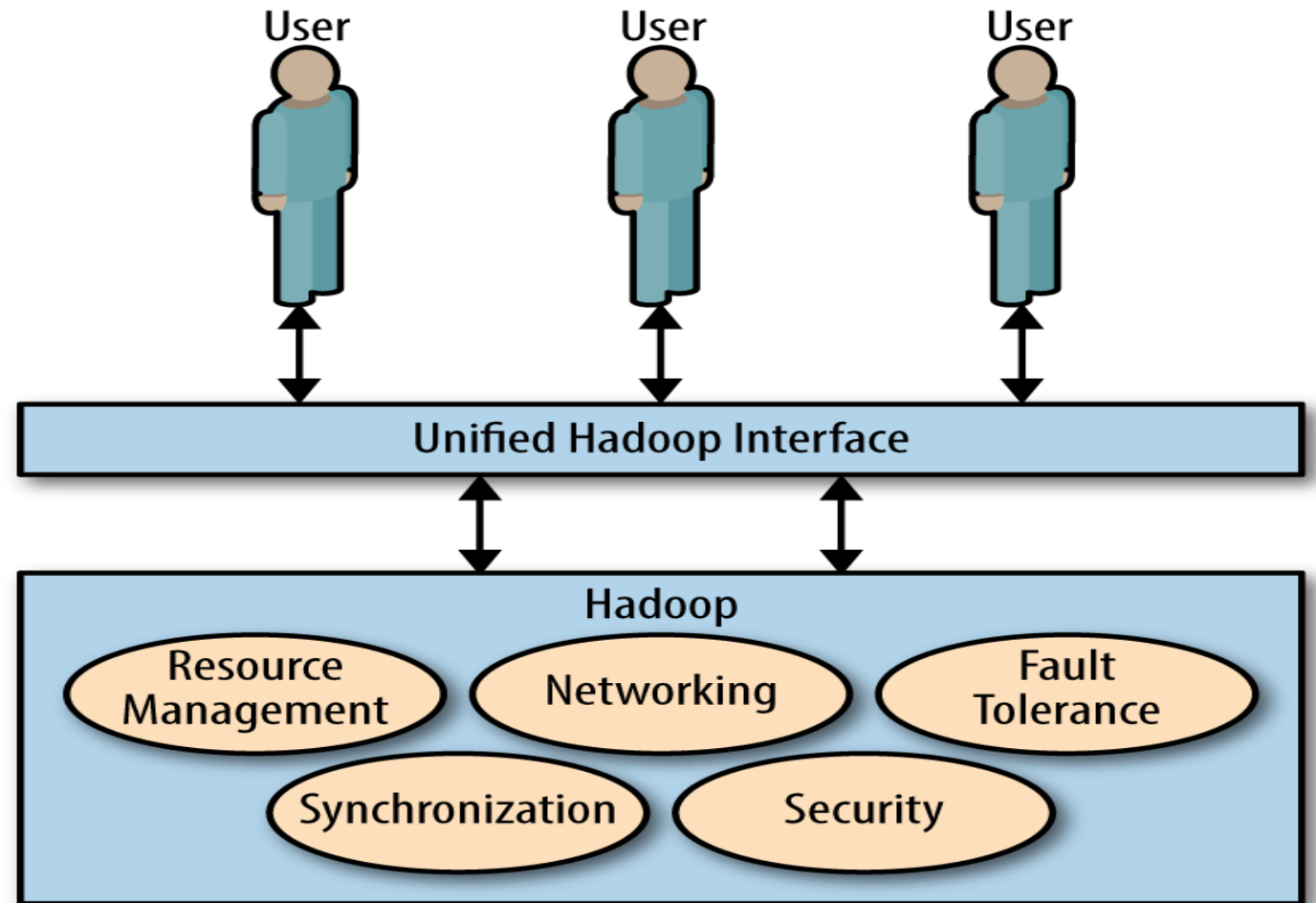


Fully-Distributed




For production environment

Understanding these modes enables proper deployment strategy

Abstraction Provided by Hadoop



Final Takeaways

-  **Choose Wisely**
Select installation mode based on requirements
-  **Unique Purpose**
Each mode serves different development needs
-  **Scale Accordingly**
Match deployment to project scale and complexity

Next Steps: Practice installation and configuration

