



MODULE III PART 2  
HDFS

# At the end of the session you will be able

- Explain Hadoop Architecture.
- HDFS



# Basic Terminologies



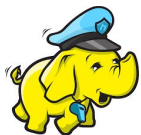
- Node
  - A computer.
  - A non-enterprise commodity hardware.
- Rack
  - A collection of 30-40 nodes.
  - Connected to the same network switch.
  - Physically stored close together.
- Hadoop Cluster
  - A collection of racks.

# HADOOP ARCHITECTURE



- Master-Slave technology
- Master Node
  - Assign task to slave nodes
  - Manage resources
  - Stores metadata
- Slave Nodes
  - Do the actual computing
  - Stores data

# HADOOP ARCHITECTURE



## HADOOP ARCHITECTURE

APPLICATION LAYER



Other  
Applications

RESOURCE  
MANAGEMENT  
LAYER



STORAGE LAYER



# HADOOP ARCHITECTURE



- Consists of three layers
  - HDFS
    - Storage Layer
  - YARN
    - Resource Management Layer
  - MapReduce
    - Application Layer

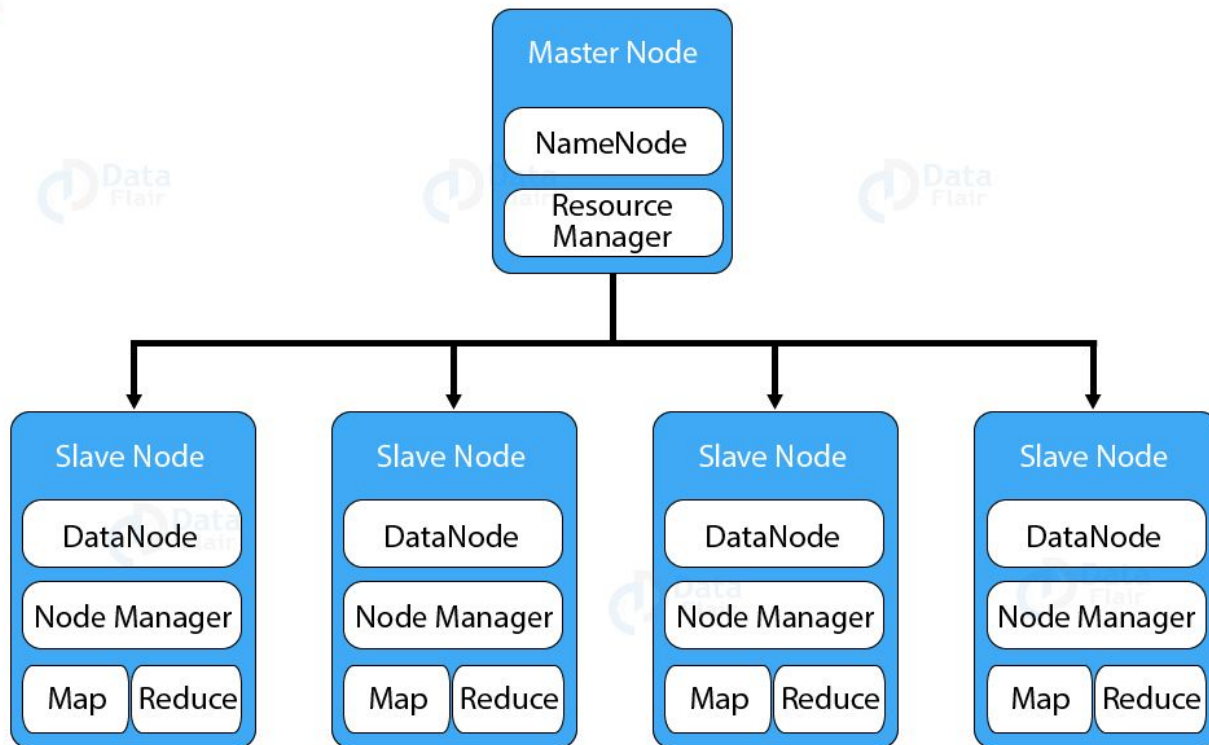


# HDFS

Hadoop Distributed File System

- Provides data storage for Hadoop
- Splits data into smaller blocks
- Stores blocks in a distributed manner.
- 2 daemons
  - NameNode
    - For master node
  - DataNode
    - For slave nodes

# HDFS Master-Slave Architecture







# HDFS

a.NameNode and DataNode

- NameNode
  - A software that runs on commodity hardware.
  - Regulates clients' file access.
  - Executes file system operations such as renaming, closing, and opening files and directories.
  - Manages file system namespace.



# HDFS

a. NameNode and DataNode

- DataNode
  - Process client's request.
  - Performs read/write operations.
  - Stores actual business data.
  - Perform operations such as block creation, deletion, and replication according to the instructions of the namenode



# HDFS

## Goals

- Fault detection and recovery.
- Huge datasets.
- Hardware at Data.

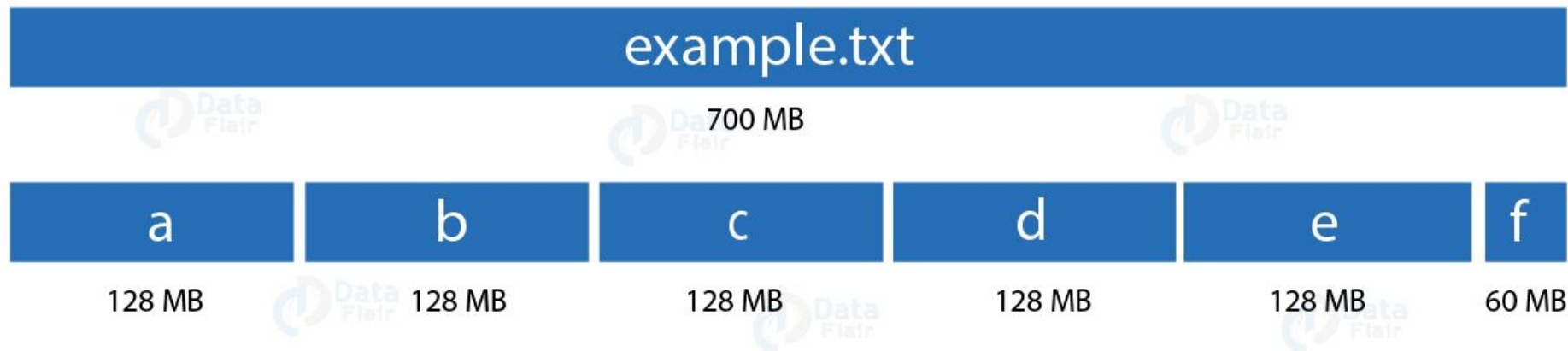


# HDFS

## Blocks

- To store user data.
- Smallest storage unit.
- Files are divided into one or more segments.
- Block is the smallest contiguous storage allocated to a file.
- Default size is 128MB.
- Can be increased as per requirement.

# HDFS Block



A rectangular blackboard with a dark, textured surface is framed by a thick, rustic wooden border. The words "Thank" and "You" are written in a white, serif typeface, centered on the board. "Thank" is on the top line and "You" is on the bottom line.

Thank  
You