



BIG DATA
DEVELOPMENT

ACADGILD

Session 4: Hadoop Java API's

Practice Paper - 3



PRACTICE PAPER

Session 4: Practice Paper 3

Table of Contents

1. Introduction

2. Problem Statement

3. Output

1. Introduction

In this assignment you need to answer the scenario based questions

2. Problem Statement

1.If 7TB is the available disk space per node (9 disks with 1 TB, 2 disk for operating system etc. were excluded.). Assuming initial data size is 600 TB. How will you estimate the number of data nodes (n)?

2.Imagine that you are uploading a file of 500MB into HDFS.100MB of data is successfully uploaded into HDFS and another client wants to read the uploaded data while the upload is still in progress. What will happen in such a scenario, will the 100 MB of data that is uploaded will it be displayed?

3. Output

N/A