Page No.
Date
Sec. 1. 45
in the second
· · · · · · · · · · · · · · · · · · ·
data analytics on it
· · · · · · · · · · · · · · · · · · ·
The William Commence
1.1 Subj. 5.7 . 6
, , , , , , , , , , , , , , , , , , , ,
OL - like query lang
OL -like query lang IDFS. It was developed
interface to overvino
interface for querying I in Hadoup. Hive is
stem and enables
stem and enables known as Hivesgl
MIOWIT US THEOGE

Devanshu Surana PC-23, Panel C

BDT 1, Batch 2

1032210755

BDT Lab Assignment 7

Problem Statement:

Create Hive Database and perform data analytics on it

Objectéves:

1. To learn Hive concept

2. To perform data analytics on it

Theory:

Explain:

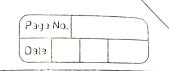
Introduction to Hive:

Hive is a data ware housing and 301-like query lang tool that is built on top of the HDFS. It was developed to provide a high-level, user-friendly interface for querying and analyzing large datasets stored in Hadoup. Hive is part of the Apache Hadoup ecosystem and enables users to write 301-like queries known as Hivesol on top of Hadoup data.

Key features of Hive:

1. SQL-Like query language: Hive provides Hive QL, making it easy for sol users to work with big data.

2. Scalability: It can handle massive datasets and scales horizontally.



- 3. Schema on Read: Hive interprets schema when data is read, allowing for flexibility with data structure.
- 4. Integration with Hadoop: Hive seamlessly integrates with Hadoop components.
- 6. Custom UPFo: Users can create custom functions for complex data transformation.

Hive Commands:

- 1) 'Create Database! : Creates a new ob in Hive.
- 2) 'Use Database': Sets the current working db.
- 3) 'Create Table': Defines a new schema there for table.
- 4) 'Insert into Table!: Inserts data into Hive table.
- 5) 'Select': Performs data retrieval and querying
- 6) 'Alter Table' Modifies structure of an existing table.
- 7) 'Drop Table': Deletes a table and its data.
- 8) Describe Tablé: Provides metadata about a tables structur

Platform: 64-bit Open Source Linux Windows.

Conclusion: Hence I learned to create tive Database and performed data analytics on it.

FAQ's.

- Ans. Hive was initially developed by Facebook and later contributed to the Apache software foundation. It is now maintained as an open-source project by the Apache Hive community.
- Ans. Hive provides the functionality to load pre-created data table entities either from our local file system or from HDFS. The load data statement is

Page No.			
Date			

		Date
	used to load data into the h	nive table
	Syntax: LOAD DATA INPATH '< The +	table data location >'
	INTO TABLE table-name;	
3.	State the difference between	Hive and Nysgl.
Ans:	. Hive	MySQL.
1.	Hive stores data in HDFS. 1.	Mysgl stores data in
0	111	A CONTROLL OF CHANGE
<b>◎</b> ¾.	Hive uses Hive &L for querying. 2.	. Mysgl uses standard 88L.
3 ·	Hive follows schema-on-read. 3	. Mysgl tollows schema-
4.	2.001111	-on-write 4. Mysal for traditional
		databases.
5 .	Hive is highly scalable. 5 Hive is designed for OLAP 6	. Mysgl has limitations. . Mysgl is optimized for
6.	Hive is designed for OLAP 6	. MySQL is optimized for
	on large datasets, often slower	OLAP.
	for real-time transaction.	
		(2)/12)
		11101
		,