Session 20: Oozie, Sqoop & Flume

Assignment 1

Big data and Hadoop development

Session 20: Oozie, Sqoop & Flume

Assignment

Table of Contents

Big data and Hadoop development

1. Introduction

In this assignment you need to fetch the tweets from Twitter using Flume and copy that to HDFS.

2. Objective

This assignment will help you to understand operations in Flume

3. Prerequisites

Acadgild's VM, or Linux operating system with Hadoop and Flume installed in it.

4. Associated Data Files

N/A

DATE SET DESCRIPTION

Link: https://acadgild.com/blog/streaming-twitter-data-using-flume/

5. Problem Statement

Link: https://acadgild.com/blog/streaming-twitter-data-using-flume/

You need to fetch the data from Twitter by following the steps given in the above blog and then send us the screenshot of the fetched data in HDFS.

6. Expected output

N/A

7. Approximate Time to Complete Task

120 min