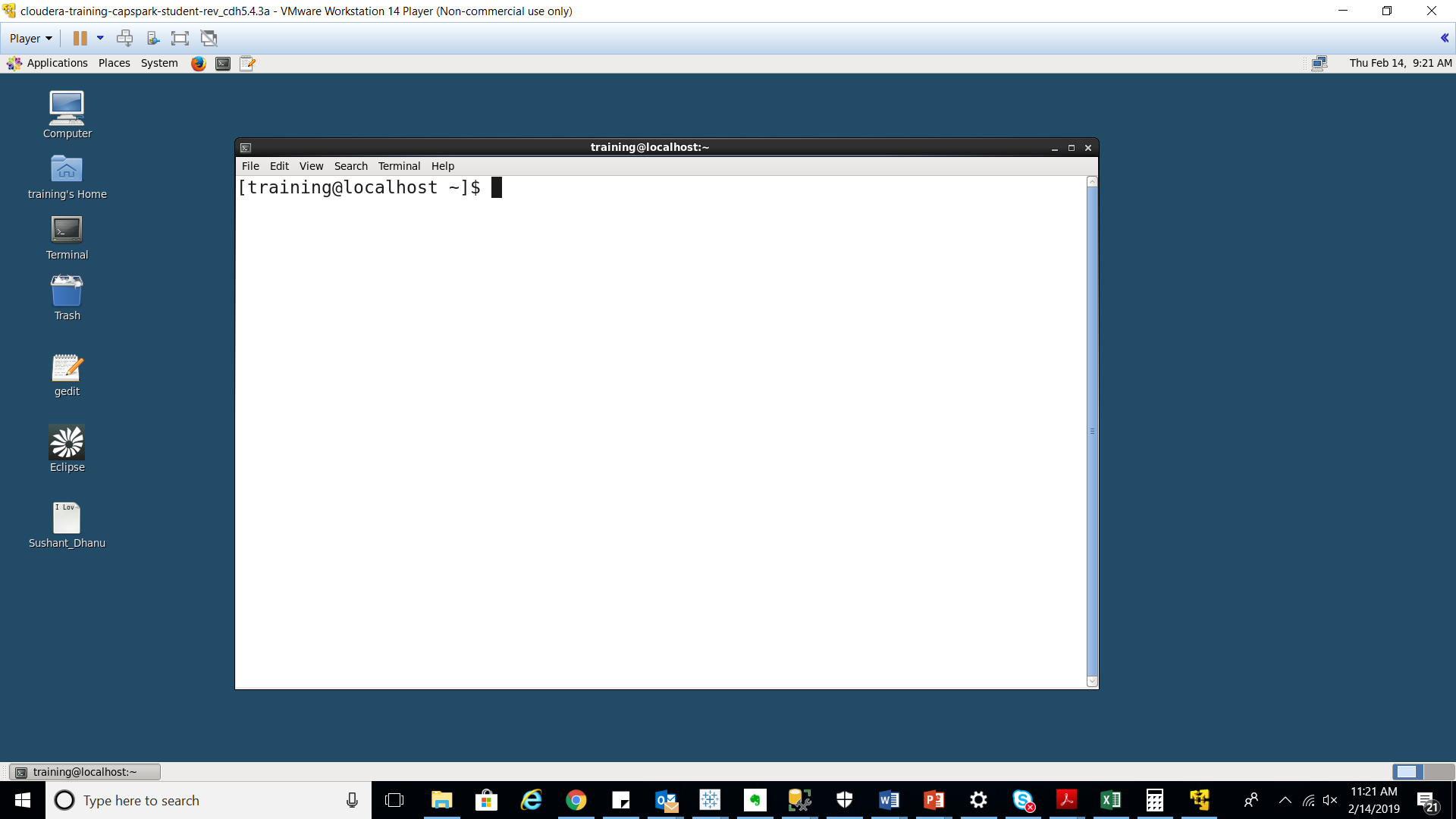
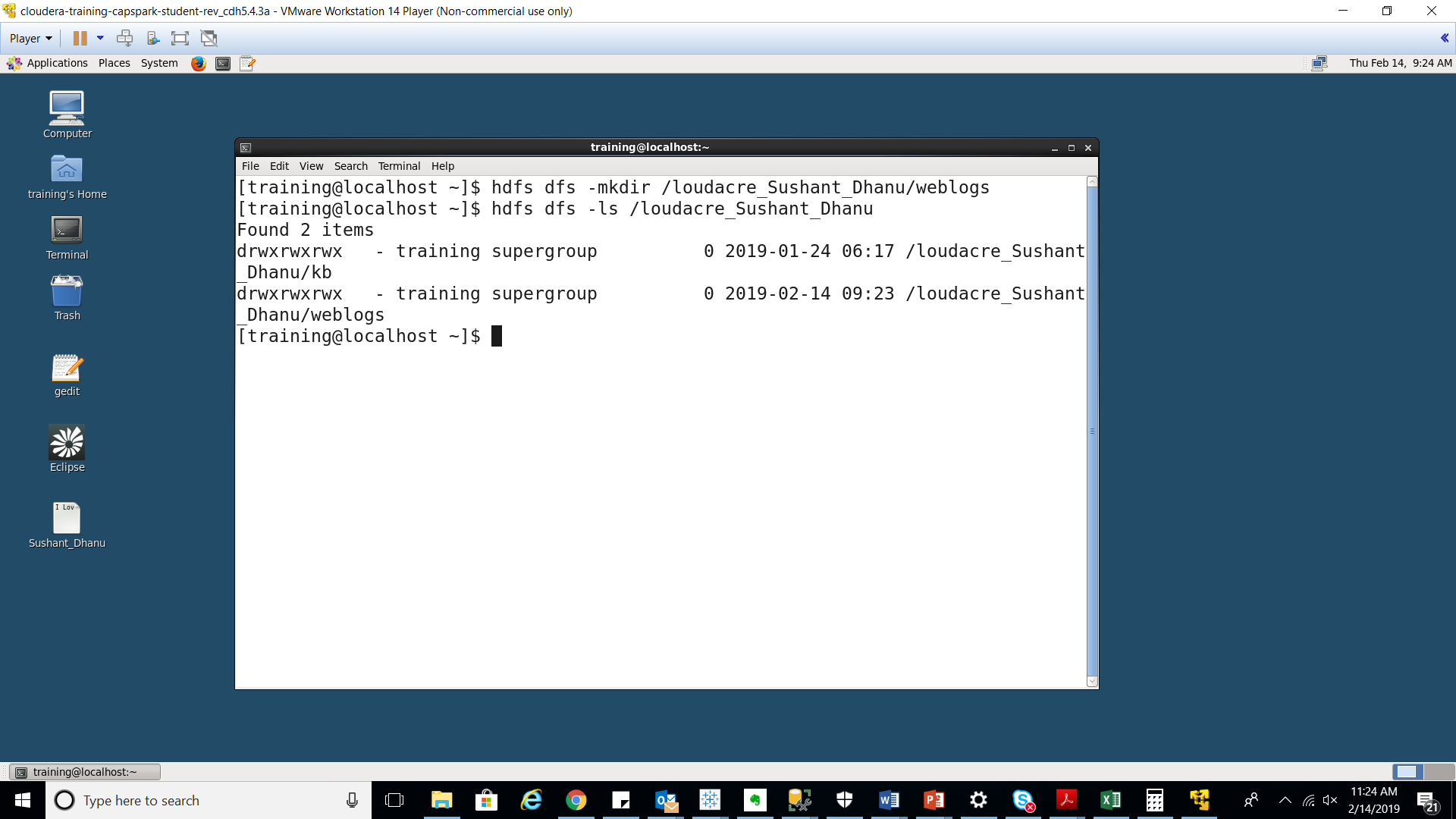
**Flume**

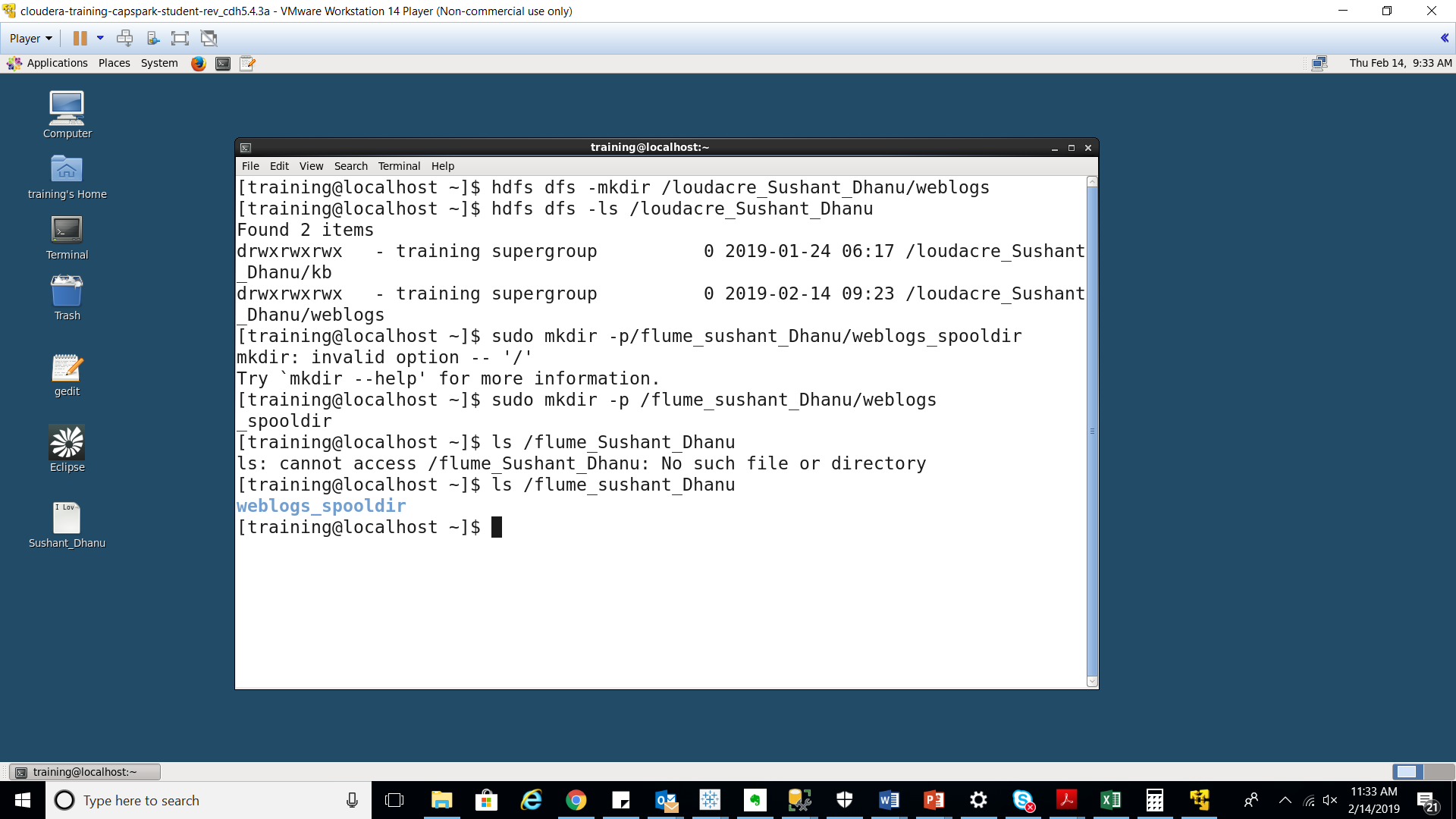
Start the Cloudera VM and then launch the “terminal” application.



We are creating a directory in HDFS within the loudacre\_Sushat\_Dhanu directory that exists. This directory will hold all the data files that we will ingest using Flume.

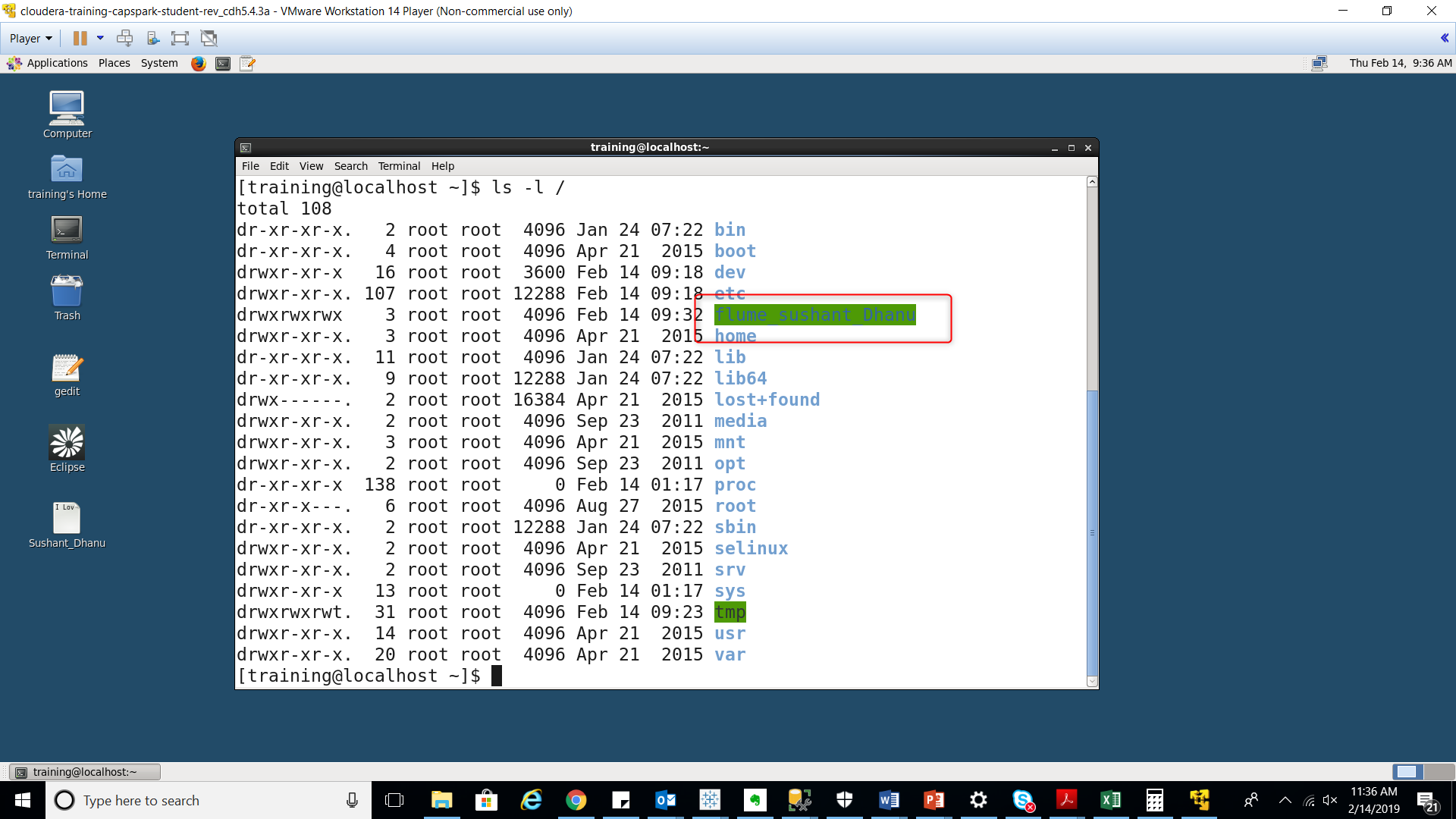


We are creating a directory on the local system to hold the flume logs that we want to ingest. This directory will serve as the “source” for our flume agent.

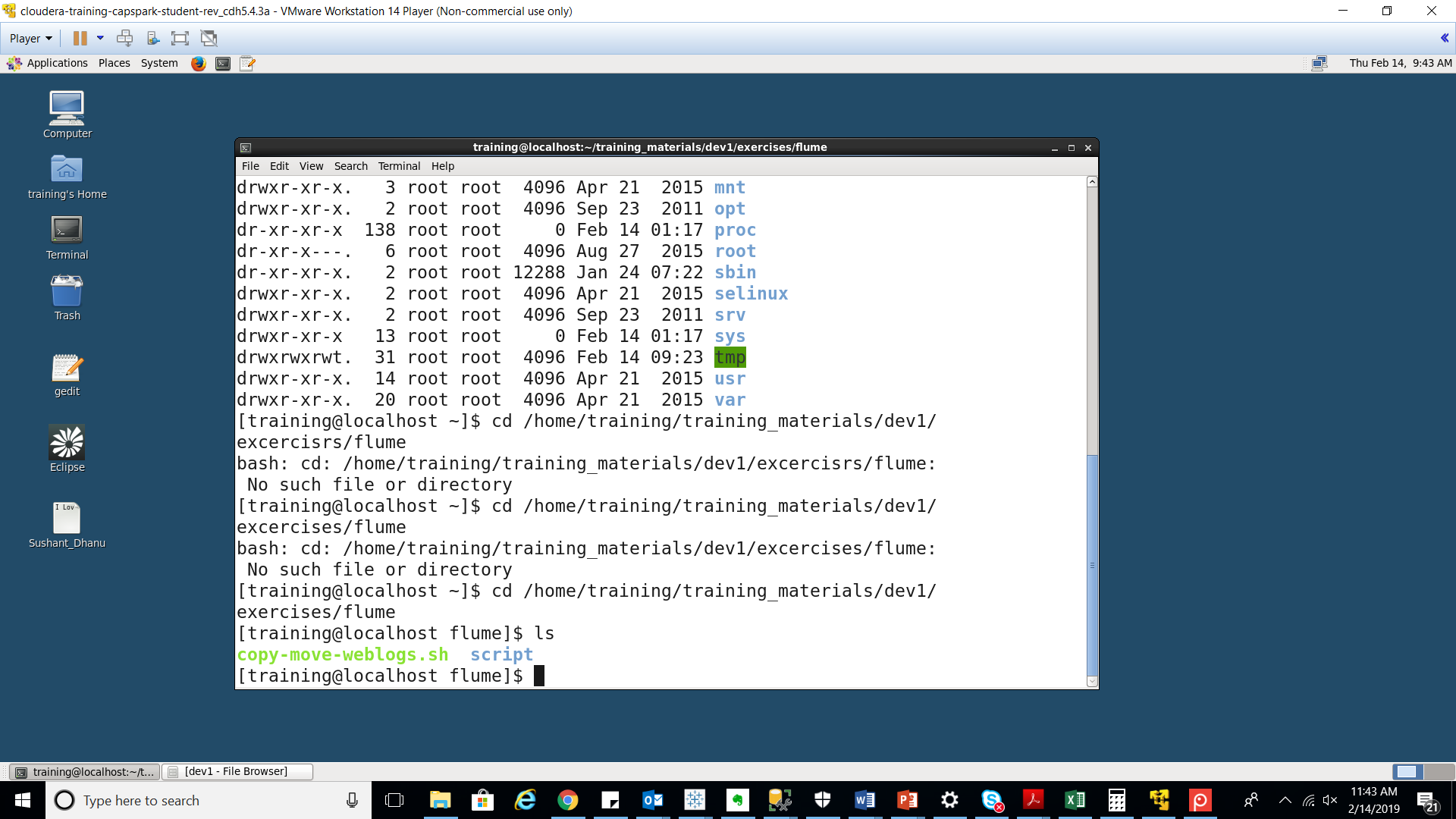


We are assigning all users permission to write to the flume\_Sushant\_Dhanu directory by using cmd “sudo chmod a+w -R /fume\_Sushant\_Dhanu”

And then looking up the permissions using “ls -l /”



We are changing directories using the “cd” command to the flume directory where we will create our flume configuration file.



Now we open a VI Editor file with flume\_conf\_Sushant\_Dhanu.conf name

Go with insert mode “ i ”

Refer sceenshot below

Configure the file using below mentioned code

Explanation:

Lines 1-3: We are aliasing the source, sink and channel for our flume agent.

Lines 5-7: We are configuring the properties for the channel

Line 5: Defining the type of channel (memory)

Line 6: Defining the capacity of the channel (10,000). This means that our channel can store 10,000 events

Line 7: Defining the transaction capacity of the channel (10,000)

Lines 9-11: We are configuring the properties for the source

Line 9: Defining the type of source (spooldir)

Line 10: Defining the location of the source

Line 11: Attaching the channel to this source

Lines 13-19: We are configuring the properties for the sink

Line 13: Defining the type of sink (HDFS)

Line 14: Defining the location of the sink

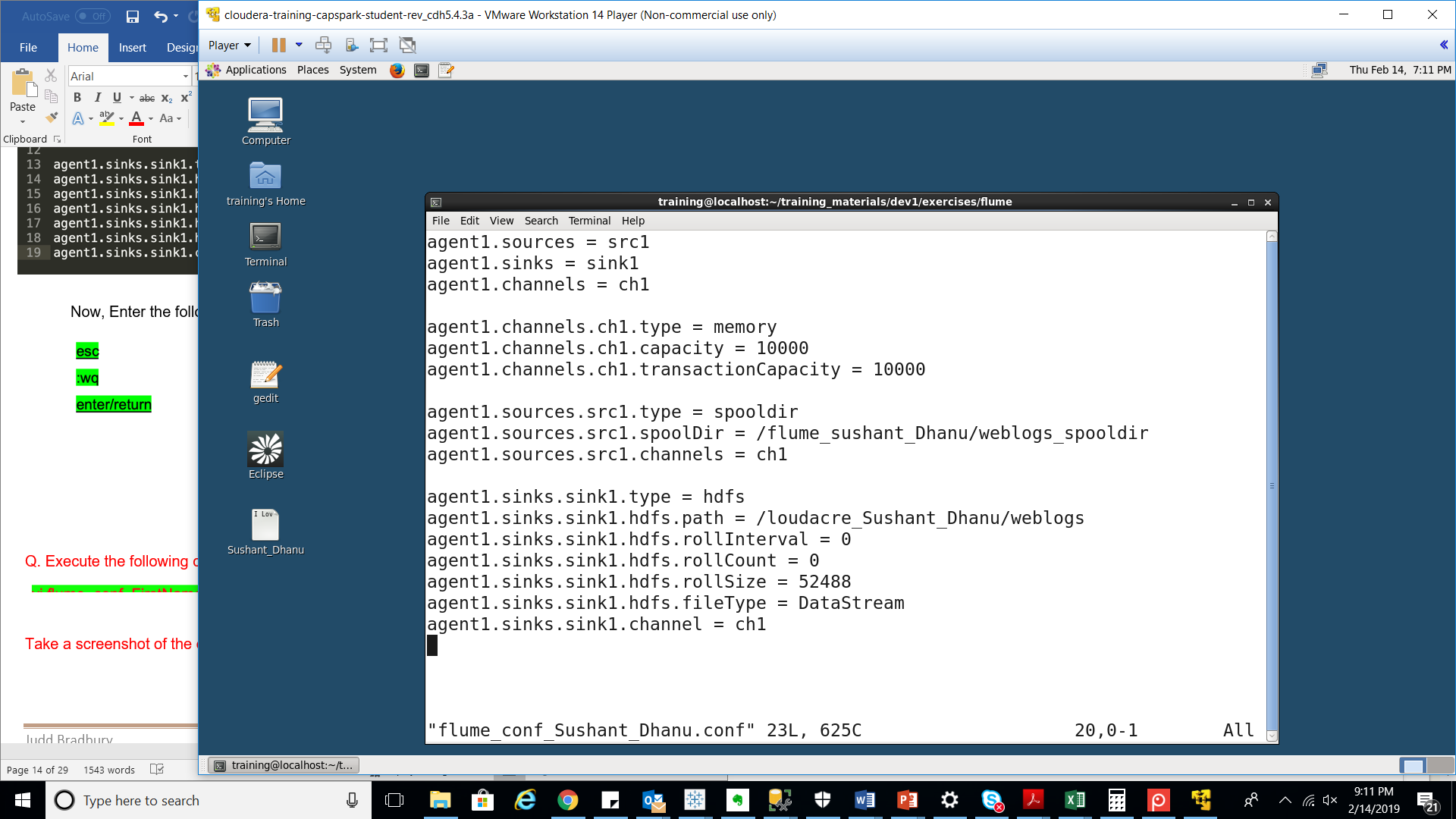
Line 15: Disabling time-based file rolling

Line 16: Disabling event-based file rolling

Line 17: Enables size-based file rolling

Line 18: Outputs raw text files instead of sequence file format files

Line 19: Attaching the channel to the sink



Now execute the following command

Flume-ng agent –conf /etc/flume-ng/conf \

--conf-file /home/training/training\_materials/dev1/execises/flume \

/flume\_conf\_Sushant\_Dhanu.conf \

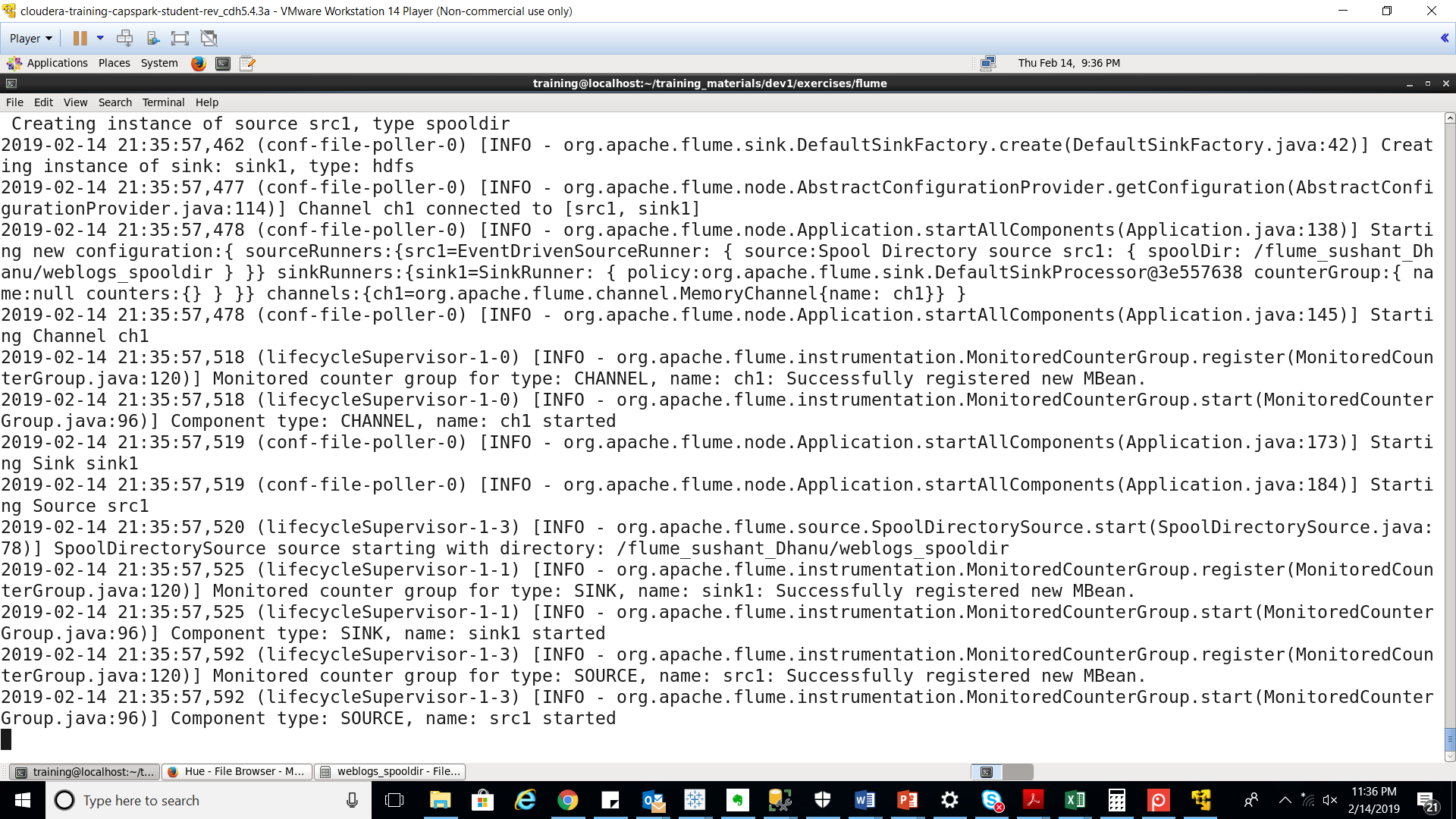
--name agent1 \

-Dflume.root.logger=INFO,console

Explanation: This command starts the flume agent.

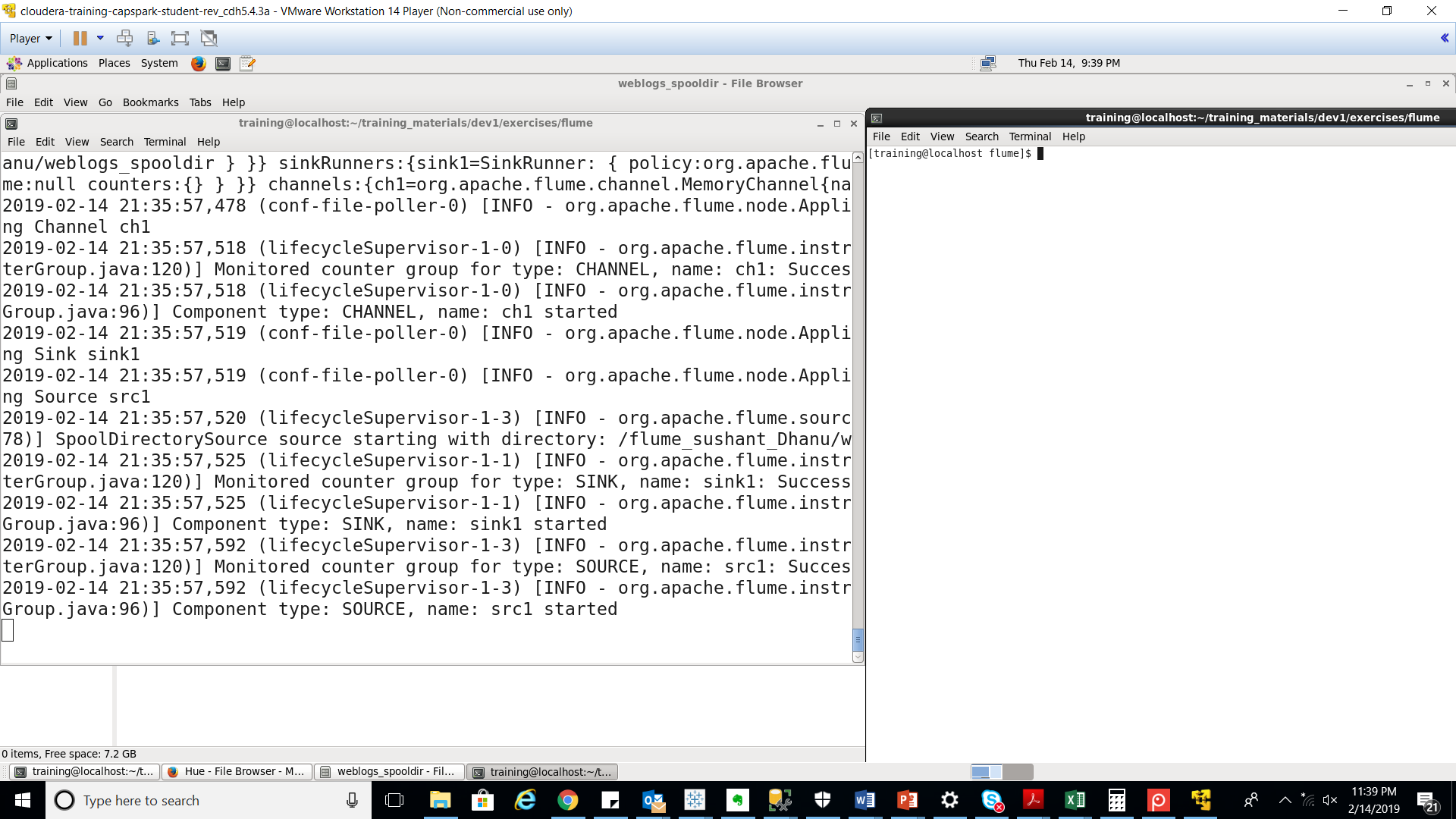
* The first line of the command specifies the path to additional configuration files that are needed to run the flume agent
* The second and third line of the command specifies the location of the flume configuration file that we created
* The fourth line of the command specifies the name of the flume agent
* The fifth line of the command tells flume to display the log messages in the terminal

Refer below screenshot as output of the command



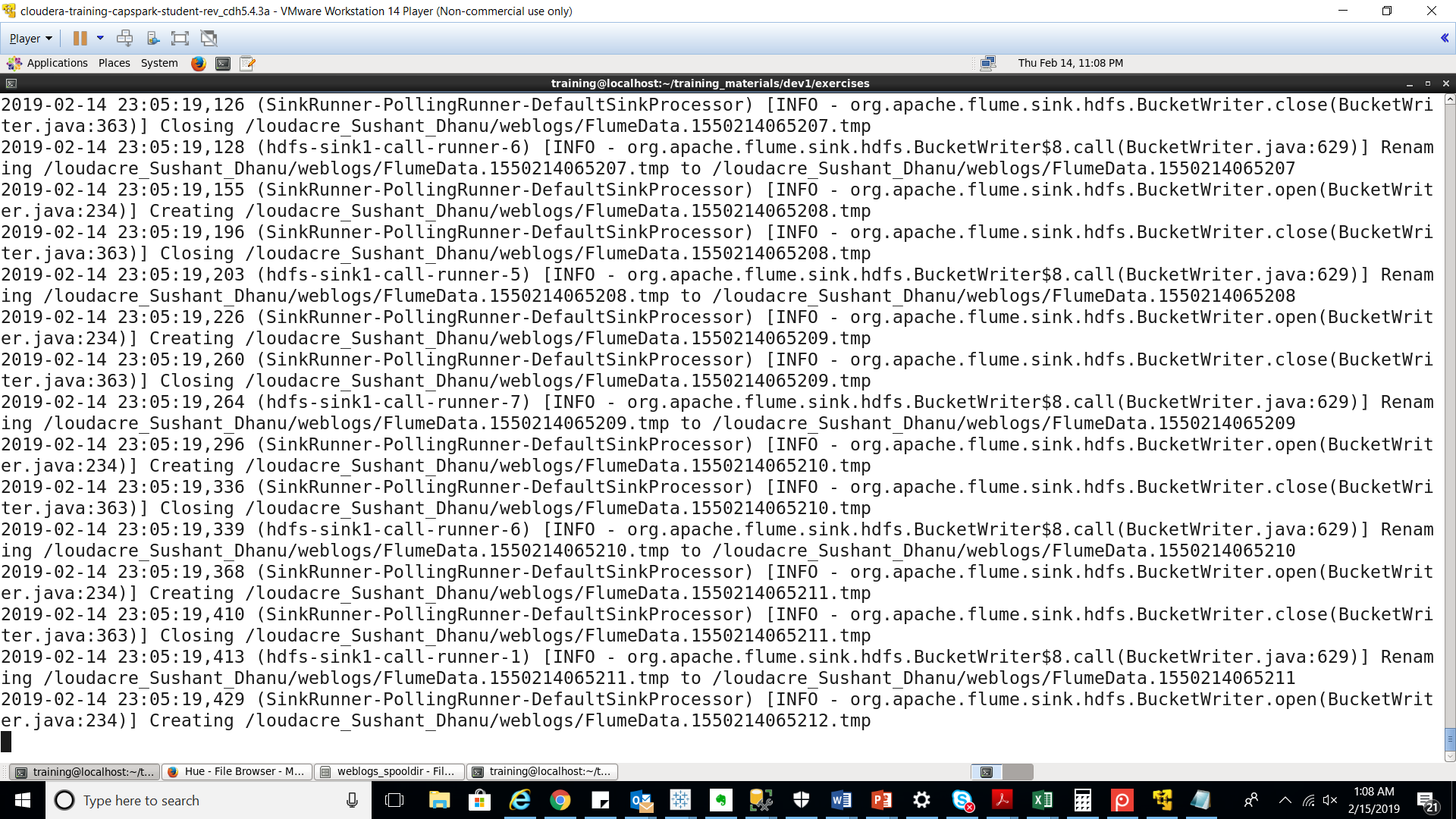
After executing the above command flume agent didn’t not start ingesting files

Now open another tab in the terminal as follows



Prior to this step, the flume agent had started, the source, sink, and channel had been configured but do not have any files to ingest. This step copies weblog files and puts them at our source location, soon as you execute this script, you will see output on the terminal window running your Flume agent.

Now if you see the following screenshot, ingestion of files has began



Once the Flume agent has finished ingesting files, enter the following command in the terminal to terminate Flume:

“Ctrl” +”c”

Using this command we are shutting down the flume agent.

.

Using the file browser in HUE, go to the sink location

(/loudacre\_Sushant\_Dhanu/weblogs)

the screenshot is of the log files in the weblogs directory

