|  |
| --- |
| МИНОБРНАУКИ РОССИИ |
| Федеральное государственное бюджетное образовательное учреждение высшего образования **«МИРЭА − Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ)**

**Кафедра прикладной математики (ПМ)**

**ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ**

по дисциплине «Методы анализа данных»

**Практическая работа № 3**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИНБО-05-22, Новиков Денис Владиславович* | (подпись) | |
| Преподаватель | *Шкерин Алексей Леонидович* | (подпись) | |
| Отчет представлен | «11» сентября 2024 г. | |  | |

Москва 2024 г.

## Постановка задачи:

Запустить агент Flume чтобы собрать данные из разных источников и сохранить их в HDFS или локальной файловой системе.

## Ход работы:

### Получение данных с помощью Apache Flume

Создаем файл конфигурации на Рисунке 1.



Рисунок 1 — Файл конфигурации

В agent1 есть источник, который слушает порт 3333 для получения данных, канал, который буферизирует данные событий в памяти и слив, который логгирует данные событий в консоль. Запишем файл конфигурации на Рисунке 2.

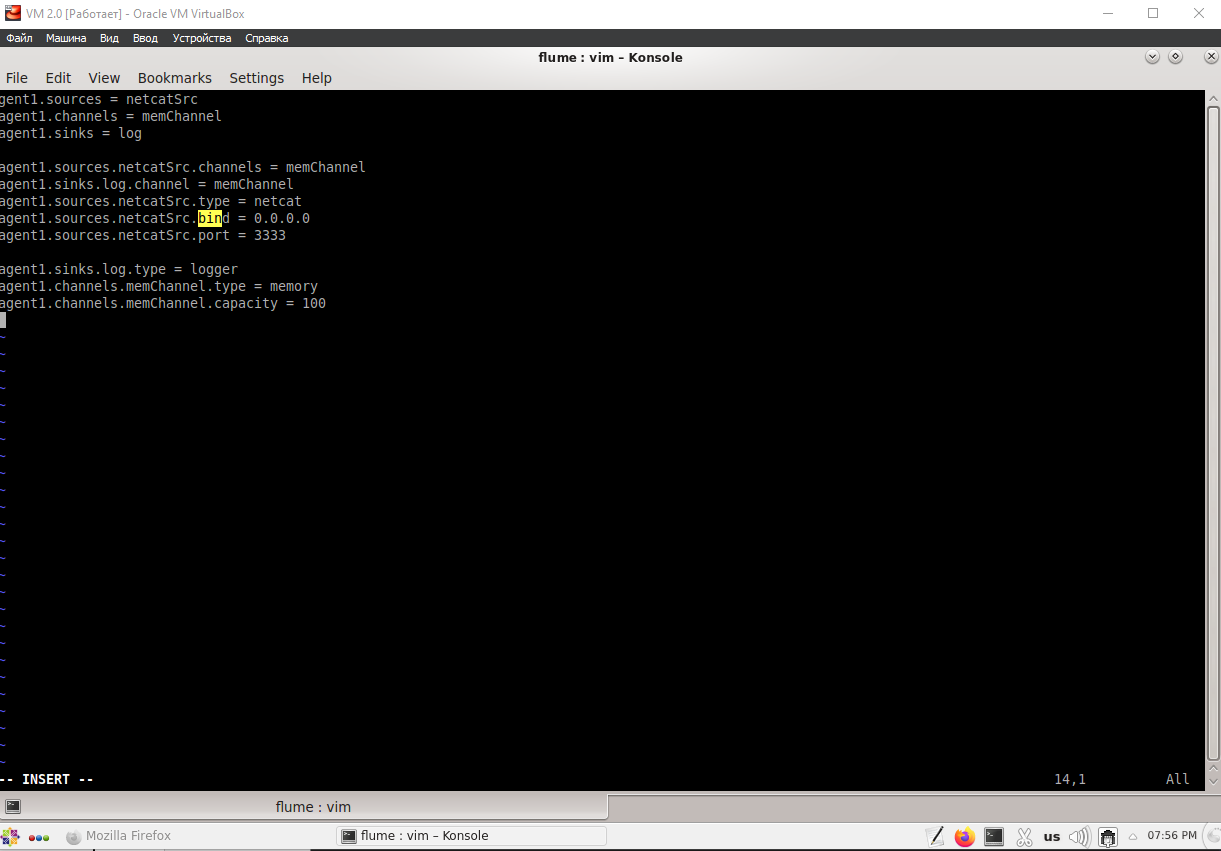


Рисунок 2 — Запись файла конфигурации

Запустим Flume agent1 на Рисунке 3.

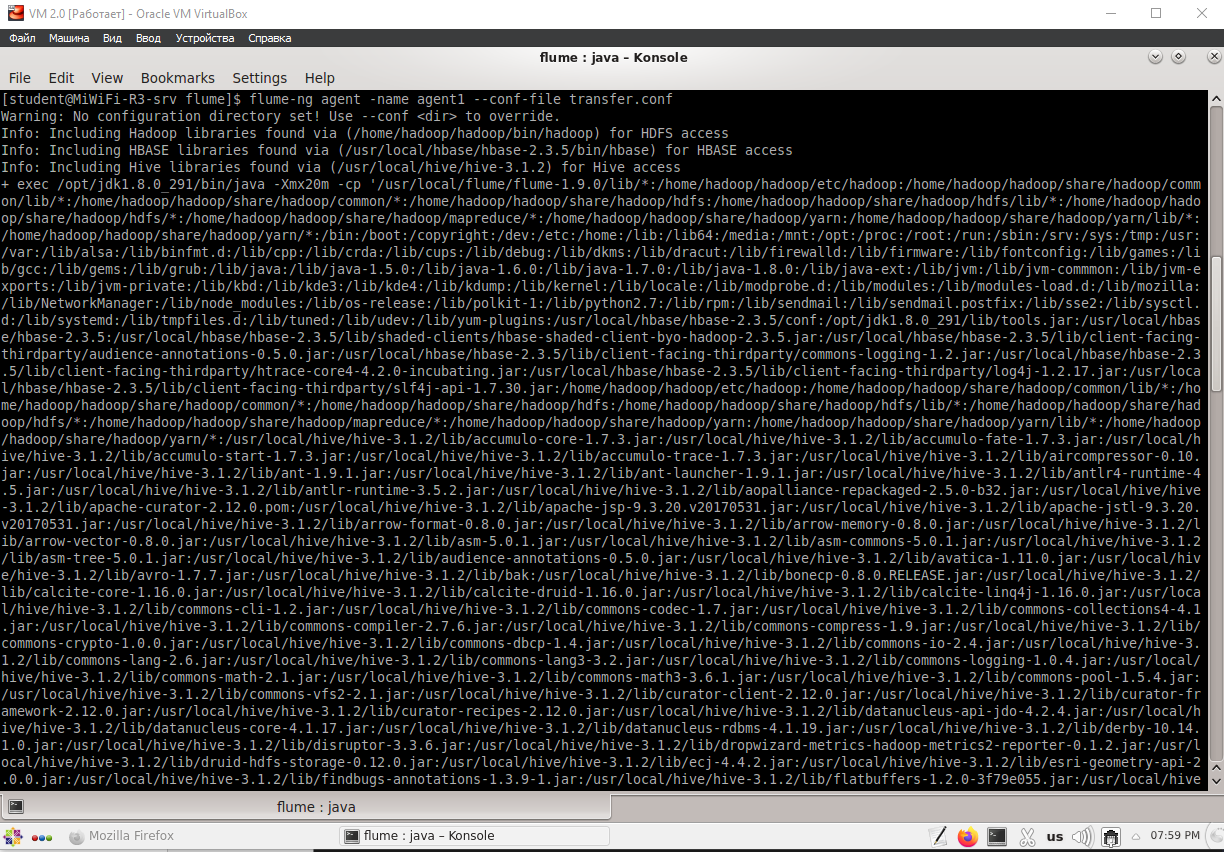


Рисунок 3 — Запуск Flume agent1

Откроем ещё одно окно терминала и выполним команду telnet на Рисунке 4.

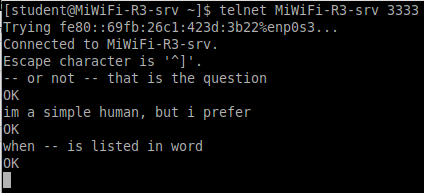


Рисунок 4 — Подключение и проверка

Проверим, что сообщение, отправленное в telnet на шаге 4, вывелось в терминале, где был запущен агент flume на шаге 3 на Рисунке 5.

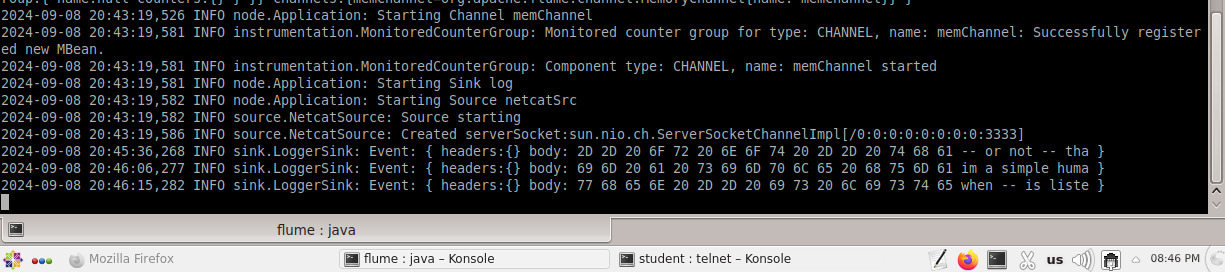


Рисунок 5 — Проверка соединения

Agent2 будет использоваться для сохранения файлов, попадающих в буферную директорию, в локальную директорию. Запишем его файл конфигурации на Рисунке 6.

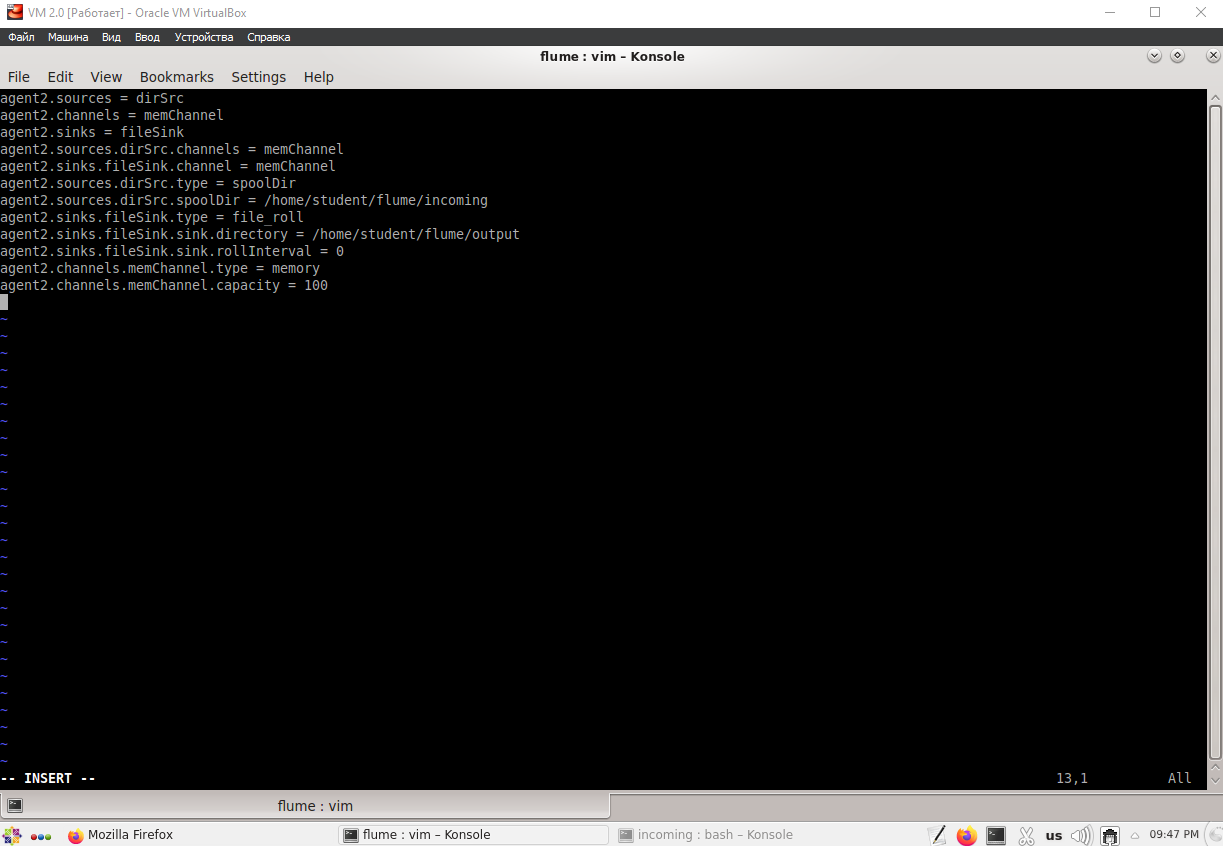


Рисунок 6 — Заполнение файла конфигурации agent2

Запустим Agent2 на Рисунке 7.

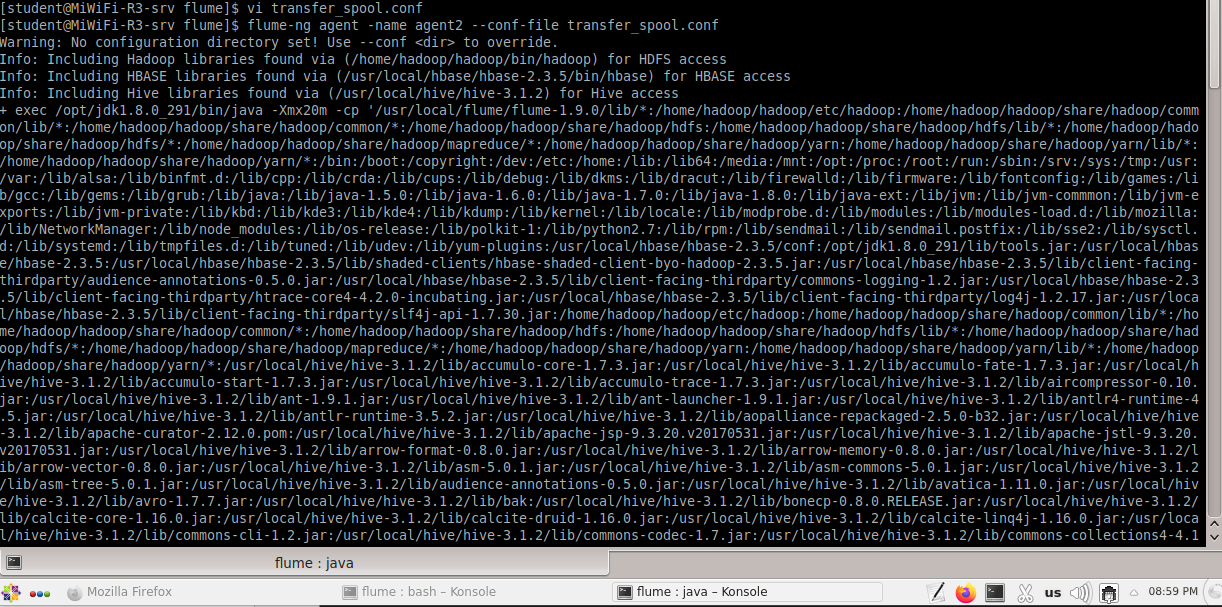


Рисунок 7 — Запуск Agent2

Откроем еще одно окно терминала и скопируем несколько файлов в буферную директорию на Рисунке 8.

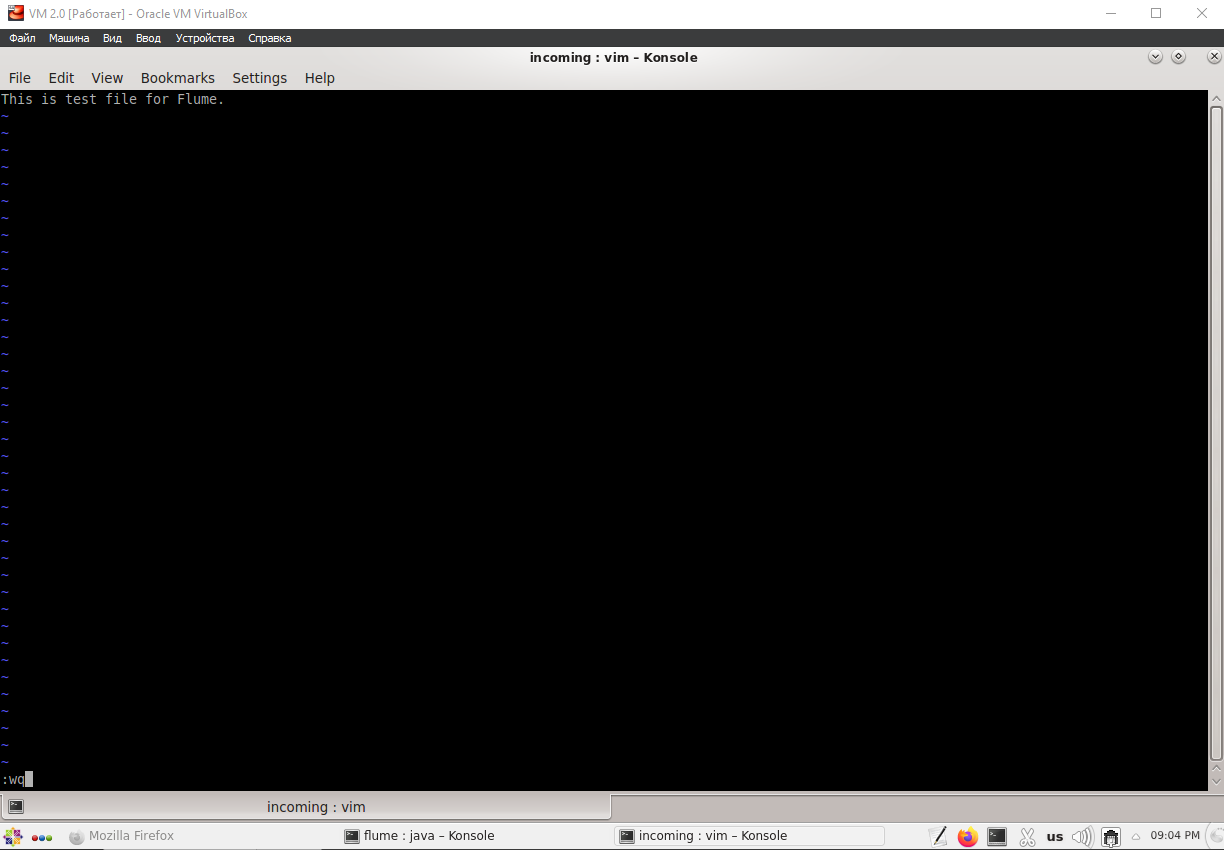


Рисунок 8 — Запись текстового файла hello.txt

Проверим наличие нового текстового файла на Рисунке 9.

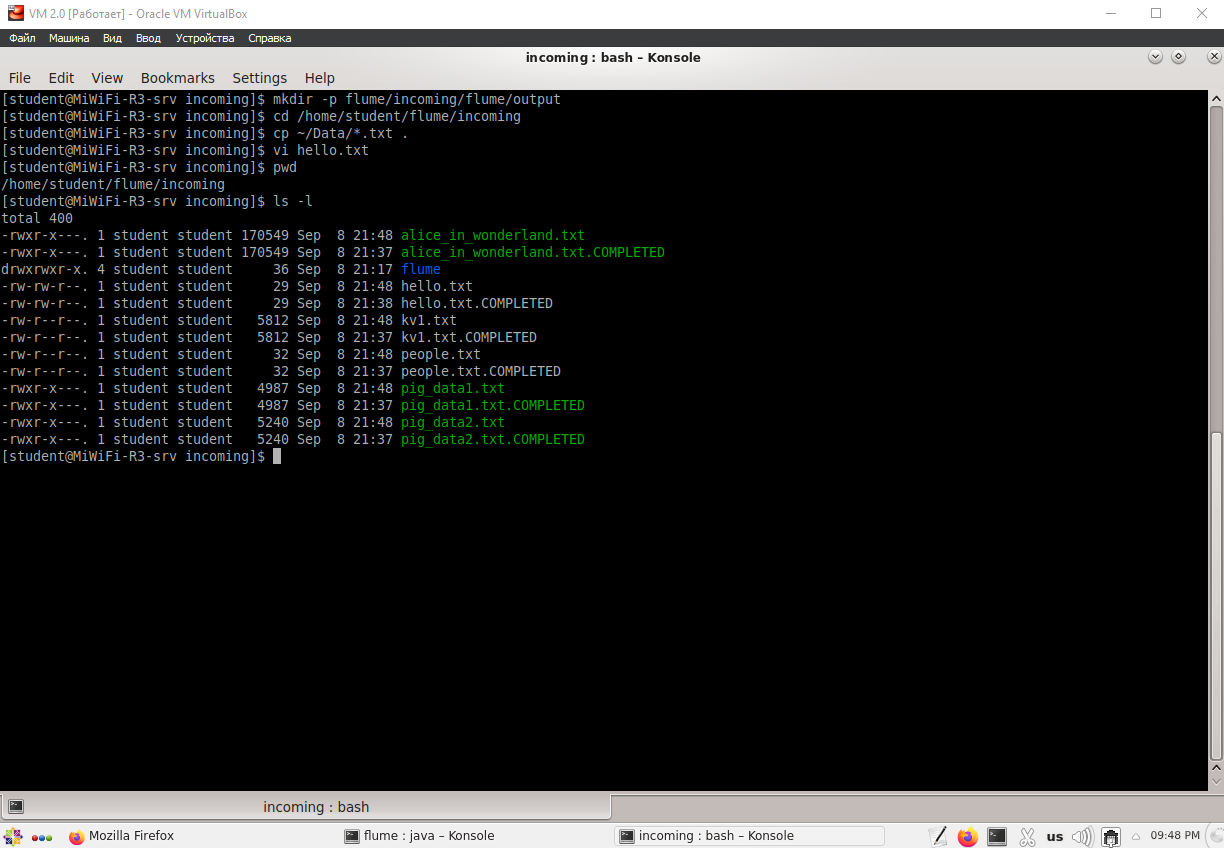


Рисунок 9 — Проверка наличия файла

Видим сообщения о том, что pig\_data1.txt, pig\_data2.txt, alice\_in\_wonderland.txt и hello.txt скопированы в буферную директорию в терминале на Рисунке 10.

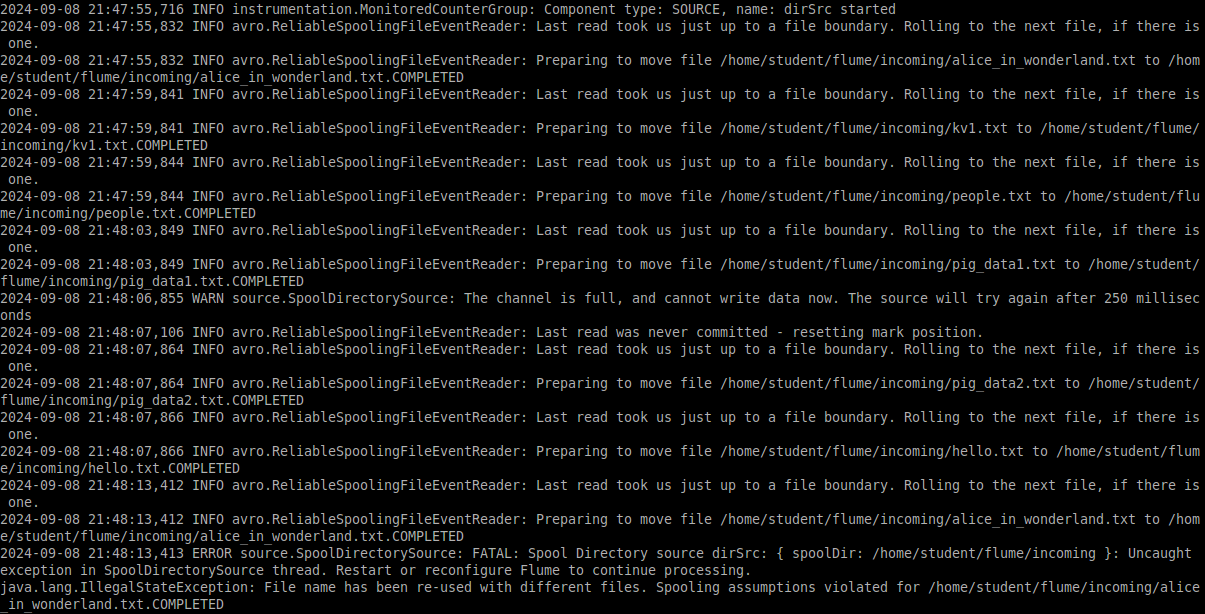


Рисунок 10 — Результат работы

Проверка наличия перенесенных файлов на Рисунке 11.



Рисунок 11 — Перенесенные файлы

Agent3 будет использоваться для вставки IP адреса хоста, где запущен агент, в заголовки события. Создадим его файл конфигурации на Рисунке 12.

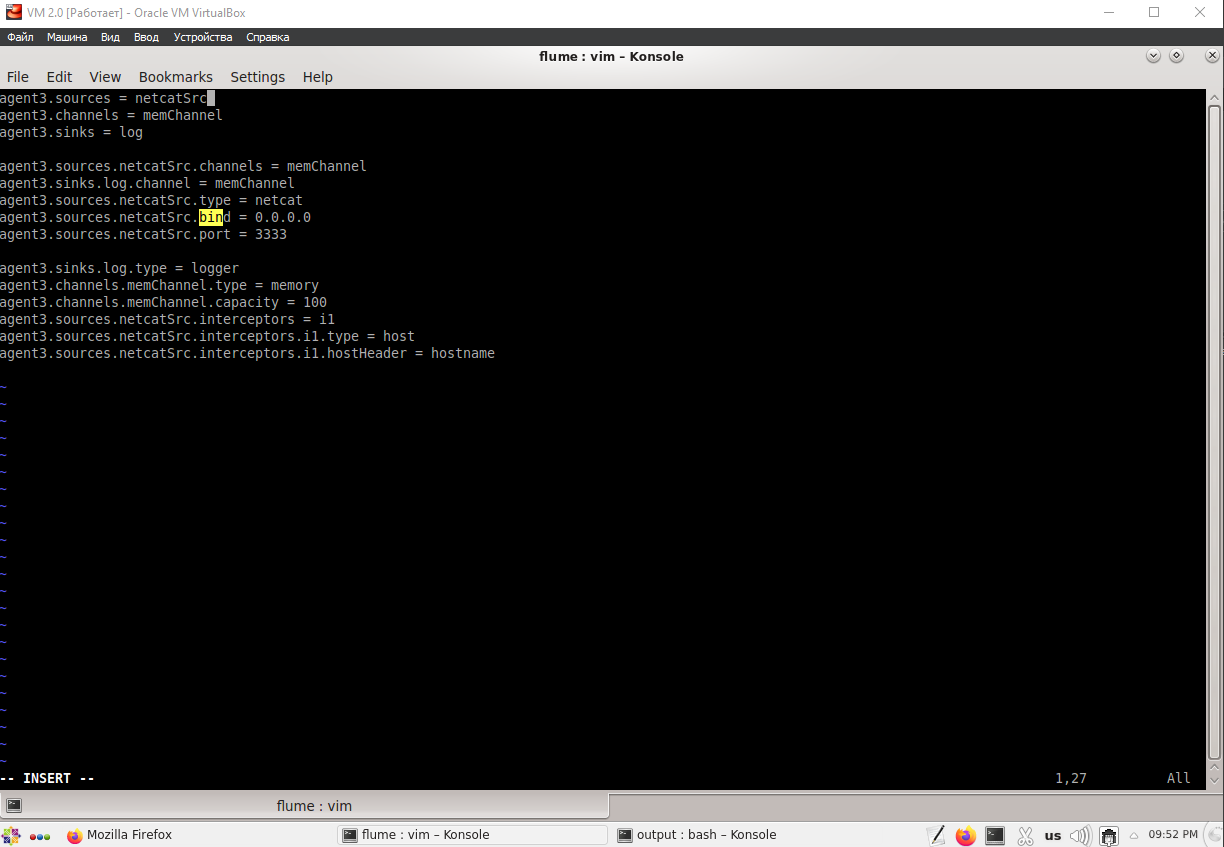


Рисунок 12 — Создание файл конфигурации Agent3

Запуск Agent3 на Рисунке 13.

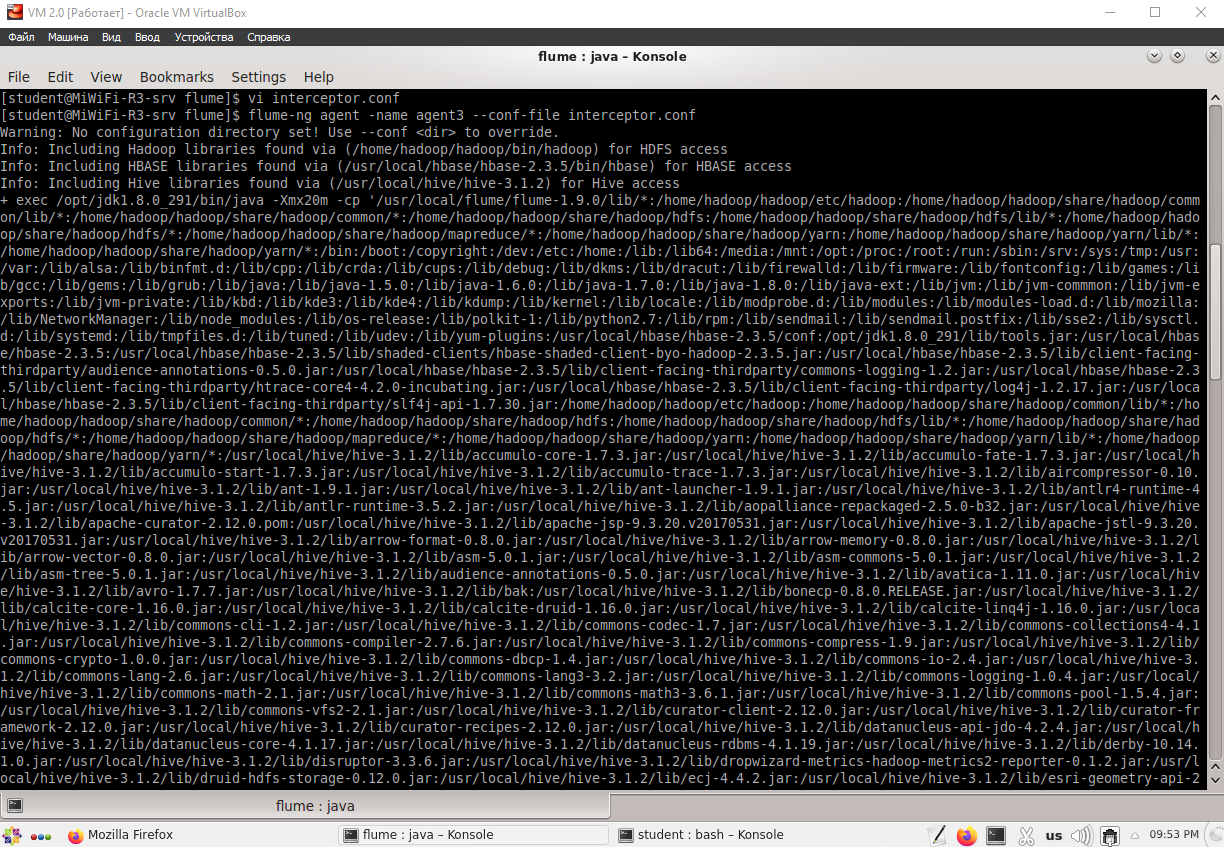


Рисунок 13 — Запуск Agent3

Откроем ещё одно окно терминала и выполним команду telnet на Рисунке 14.

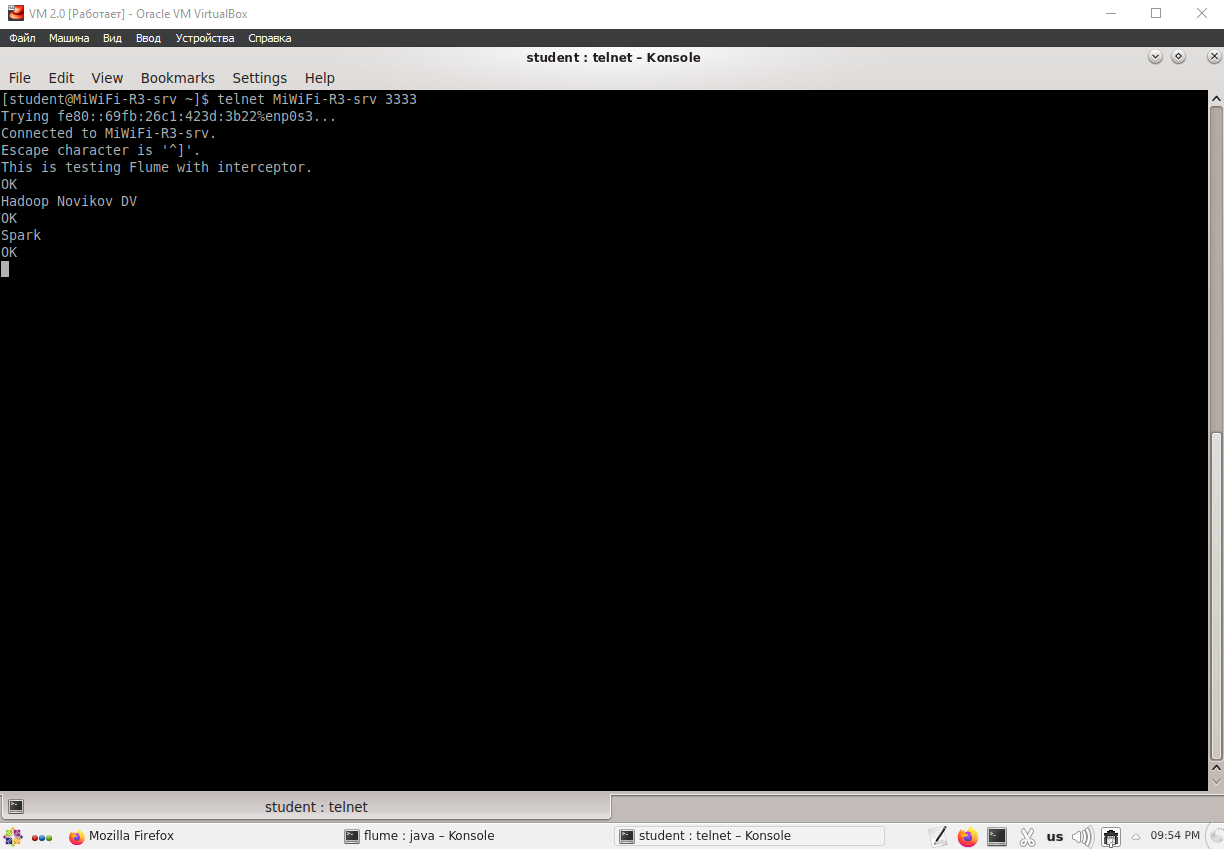


Рисунок 14 — Результат работы команды telnet

Сообщение, отправленное в telnet на шаге 3.4, выводится с терминала, где агент был запущен на шаге 3.3, и подтверждается, что IP-адрес хоста, на котором в данный момент запущен агент, вставляется в заголовок события и передается дальше. Проверка работы команды на Рисунке 15.

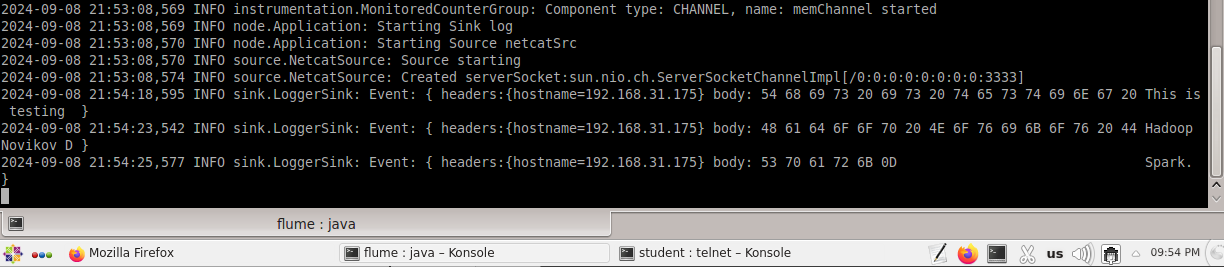


Рисунок 15 — Проверка работы команды

Создадим файл конфигурации Agent4 на Рисунке 16.

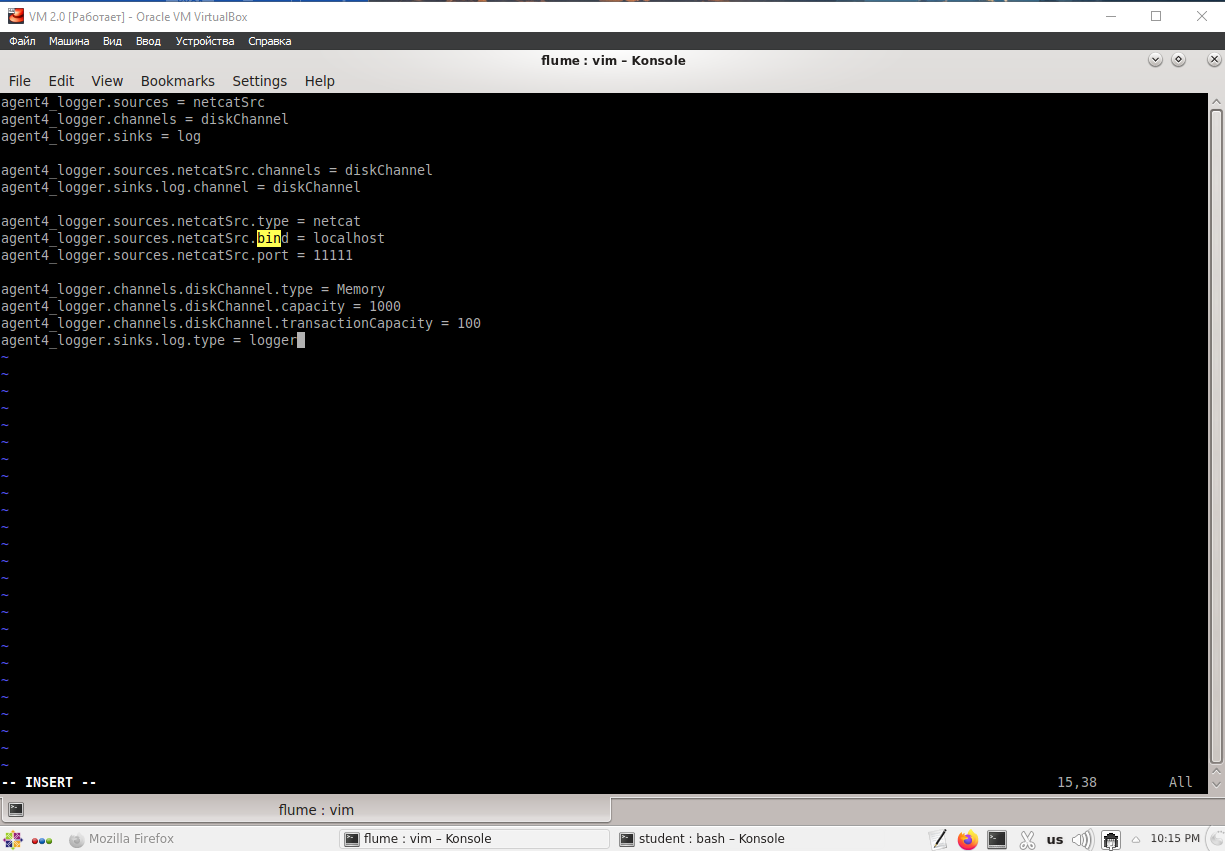


Рисунок 16 — Файл конфигурации Agent4

Запустим Agent4 на Рисунке 17.

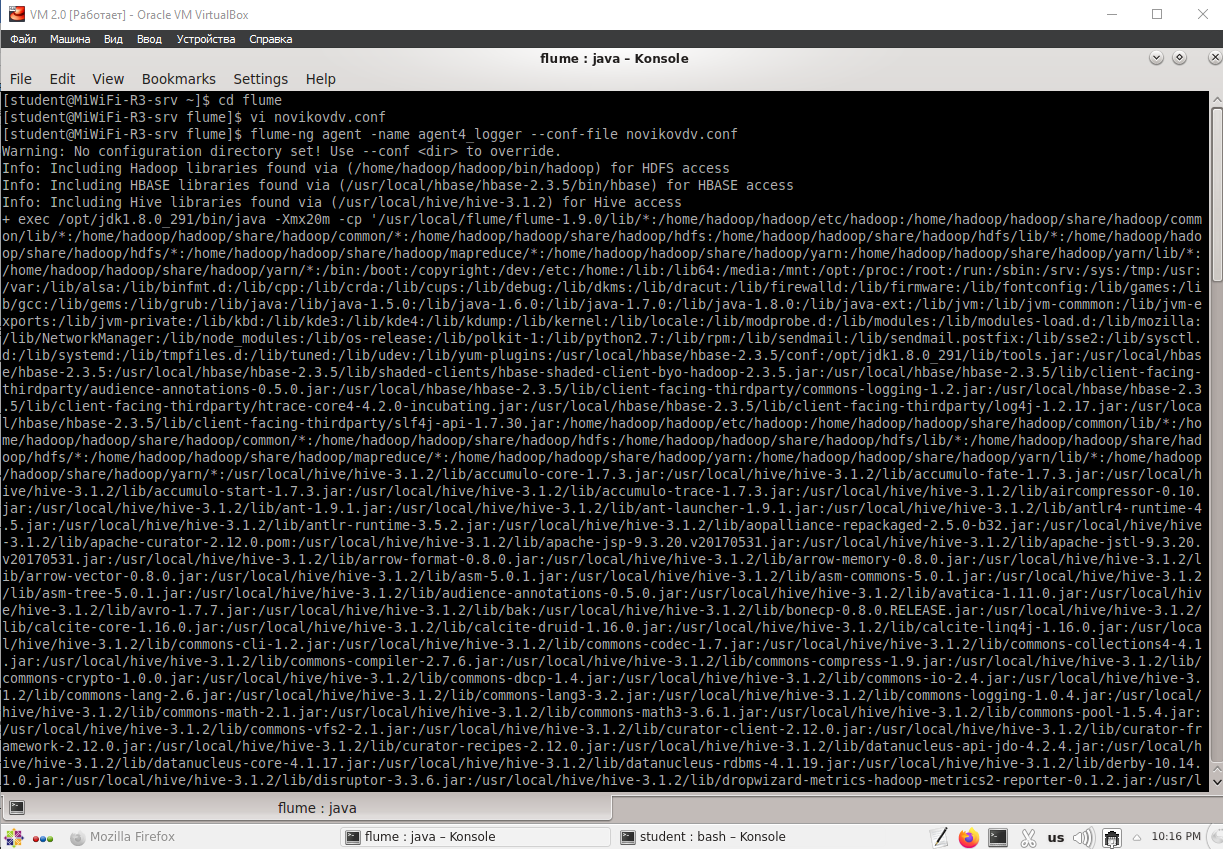


Рисунок 17 — Работа Agent4

В другом окне терминала запустите telnet и подключимся к порту 11111. Напишем что-нибудь на Рисунке 18.

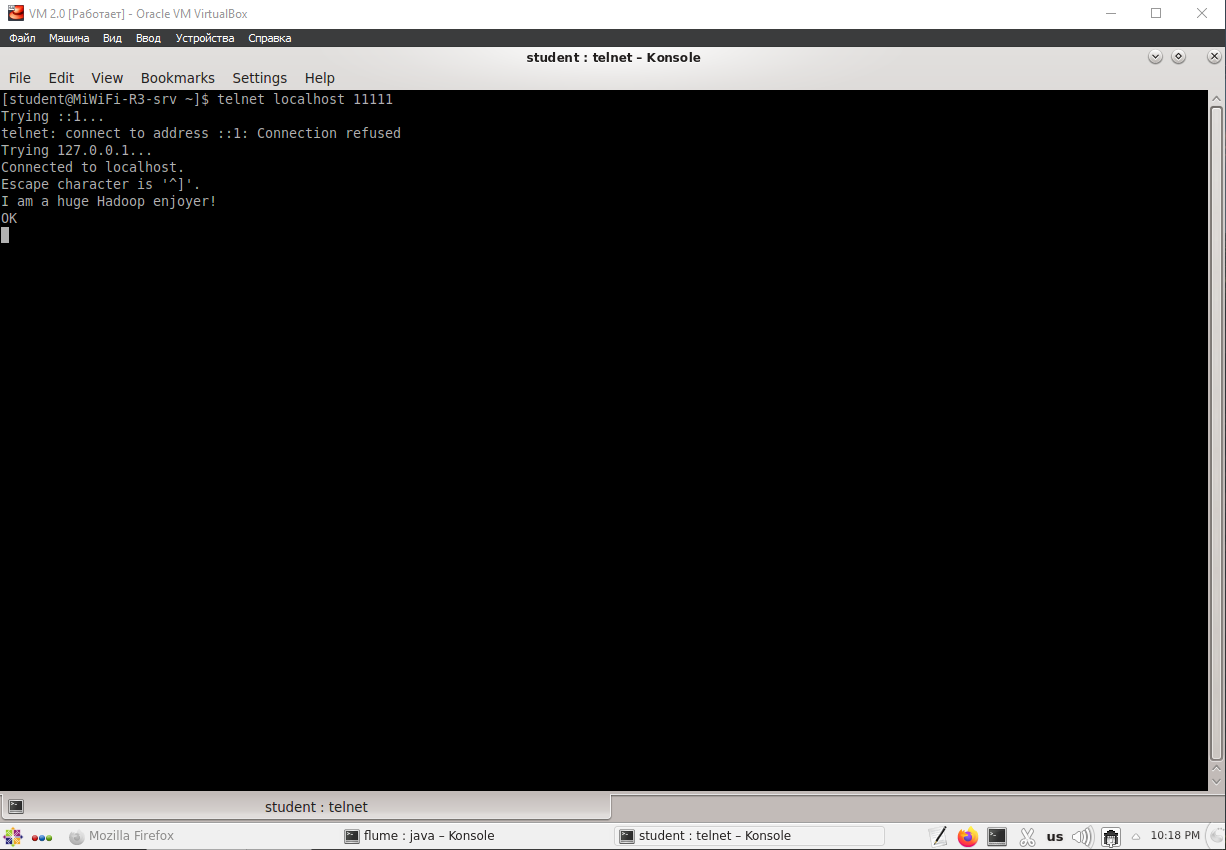


Рисунок 18 — Вводим текст

Убедимся, что мы видим написанное в терминале, где запущен агент на Рисунке 19.

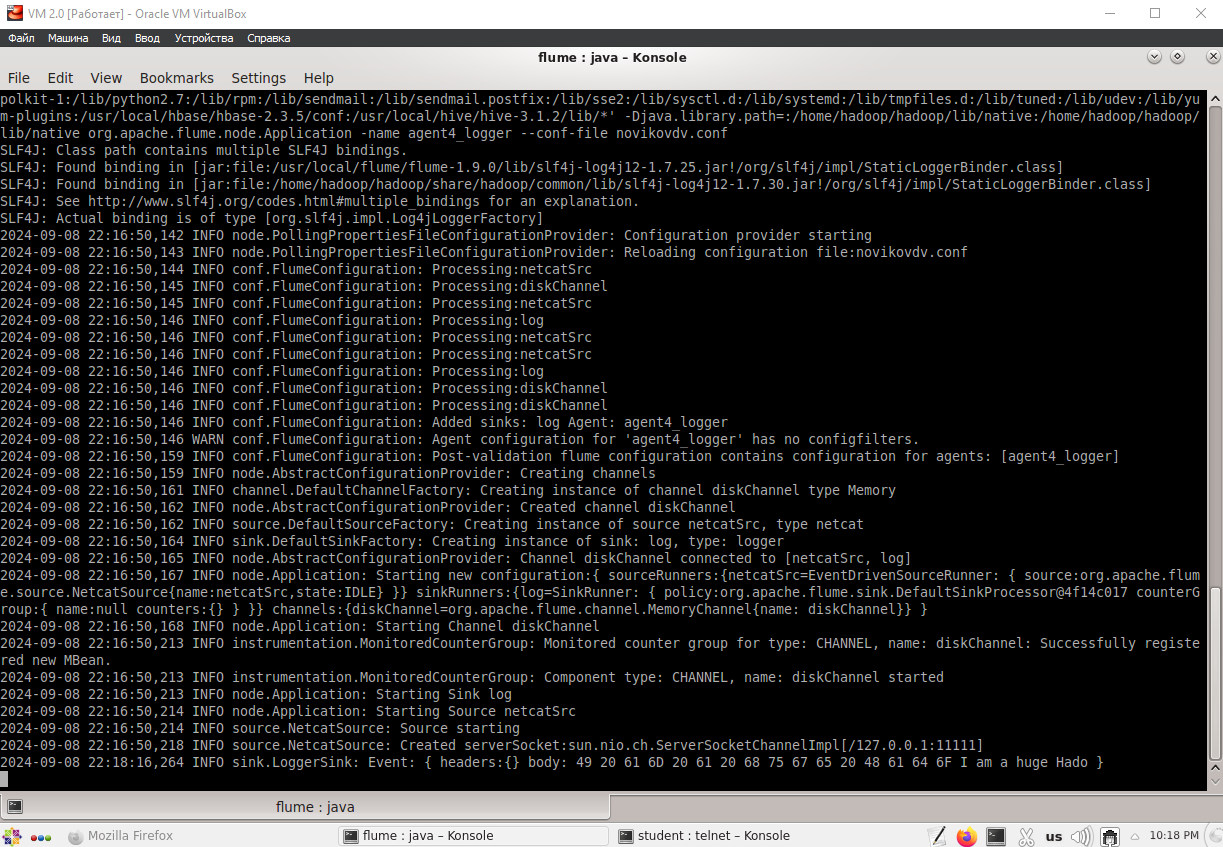


Рисунок 19 — Терминал выводит сообщение