|  |
| --- |
| 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»

**Практическое занятие № 3**

|  |  |  |  |
| --- | --- | --- | --- |
| Студент группы ИНБО-01-17 | *ИМБО-02-22, Ким Кирилл Сергеевич* | (подпись) | |
| Преподаватель | *Морошкин Никита Андреевич, преподаватель* | (подпись) | |
| Отчет представлен | «\_\_\_»\_\_\_\_\_\_\_\_202\_\_г. | |  | |

Москва 2024 г.

Практическая работа №3

Цель работы

В этой практической работе запускаем Flume чтобы собрать данные из разных источников и сохранить их в HDFS или локальной файловой системе.

Выполнения практической работы

Создаем файл конфигурации агента Agent1. (Рисунок 1)

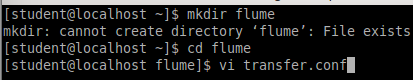


Рисунок 1 — Создание директорий

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. (Рисунок 2)

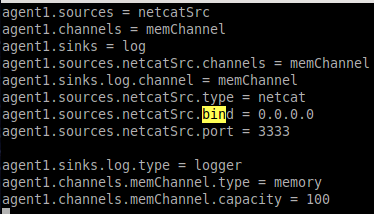


Рисунок 2 — Файл конфигураций первого агента

Запустим Flume agent1 (Рисунок 3)

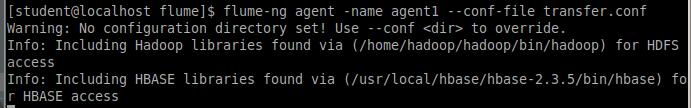


Рисунок 3 — Запуск первого агента

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

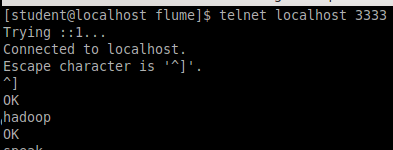


Рисунок 4 — Тестирование первого агента

Agent2 будет использоваться для сохранения файлов, попадающих в буферную директорию, в локальную директорию.

Создадим файл конфигурации (Рисунок 5).

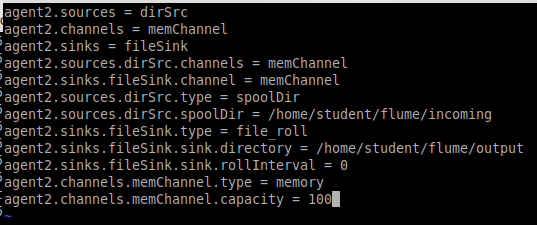


Рисунок 5 — Файл конфигураций второго агента

Запустим Agent2 (Рисунок 6).

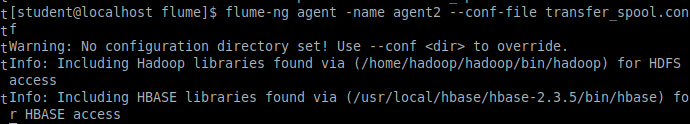


Рисунок 6 — Запуск второго агента

Вы должны увидеть сообщение о том, что pig\_data1.txt, pig\_data2.txt, alice\_in\_wonderland.txt и hello.txt скопированы в буферную директорию. Сообщение появится в терминале, где запущен Agent2. (Рисунок 7).

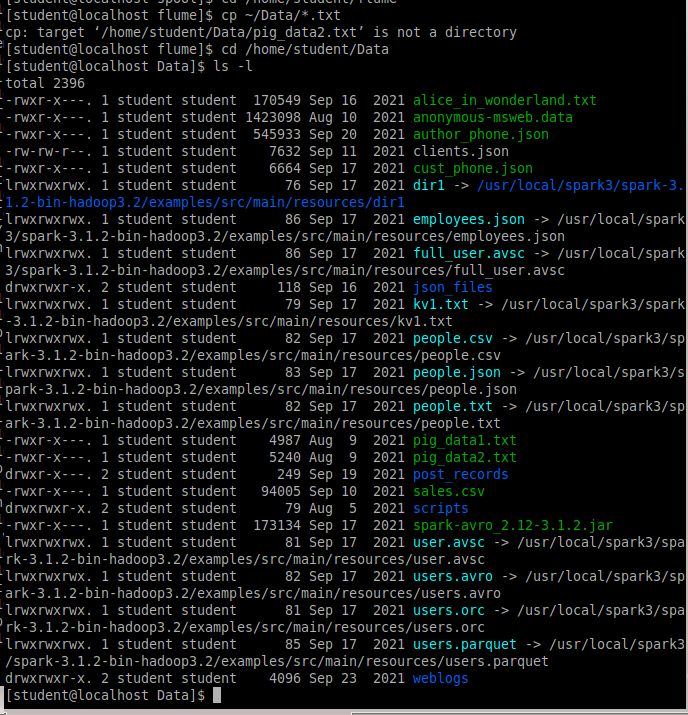


Рисунок 7 — Тестирование второго агента

Также вы можете проверить, что файл hello.txt был перенесен, открыв его утилитой vi. Перенесенный файл можно найти в директории output.

Перенесенные файлы сохранены в виде файлов в директории output. (Рисунок 8):

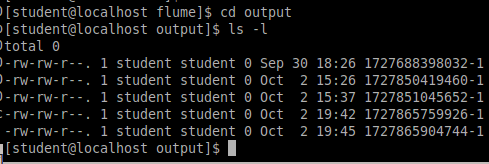


Рисунок 8 — Файлы в папке output

Использование Перехватчика (Interceptor)

Agent3 будет использоваться для вставки IP адреса хоста, где запущен агент, в заголовки события.

Создаем файл конфигурации vi interceptor.conf

Запуск файла interceptor.conf (Рисунок 9):

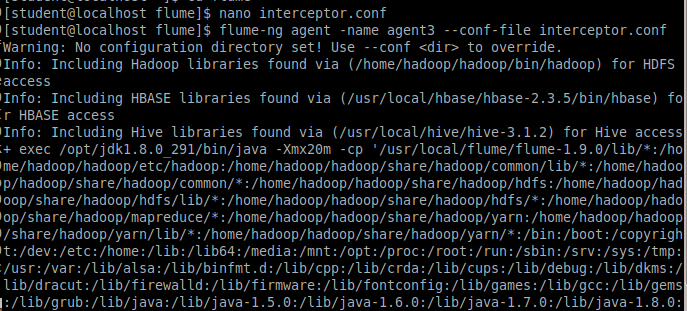


Рисунок 9 — Запуск третьего агента

Откроем ещё одно окно терминала и выполним команду telnet. (Рисунок 10).

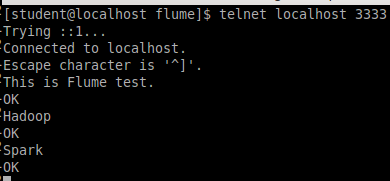


Рисунок 10 — Тестирование третьего агента

Удалим временную директорию, использовавшуюся для операций Flume (Рисунок 10):



Рисунок 10 — Удаление директорий

Создал файл file.conf со следующими настройками:

|  |  |
| --- | --- |
| Source | |
| Type | Netcat |
| Bind | localhost |
| Port | 11111 |
| Channel | |
| Type | Disk |
| Capacity | 1000 |
| transactionCapacity | 100 |
| Sink | |
| Type | logger |

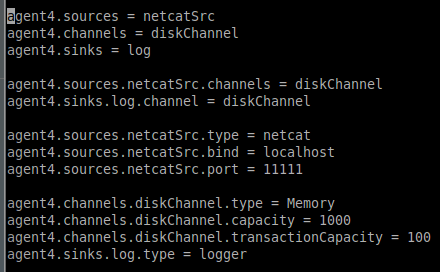
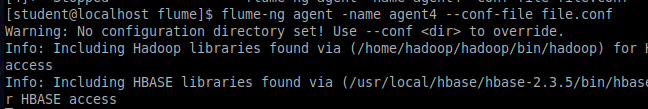
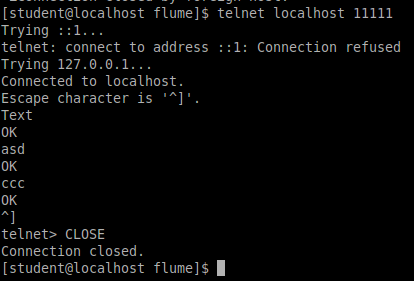


Рисунок 11 — Файл конфигураций четвертого агента

Запустил четвертого агента (Рисунок 12):



**Рисунок 12** — **Запуск четвертого агента**

****

**Рисунок 13** — **Тестирование четвертого агента**

**Вывод:** в ходе выполнения практической работы 3 мы изучили утилиту flume для получения и сохранения данных в реальном времени в системе HDFS или локальной файловой системе.