

Ковальчук Александр, 520 группа

Задание №5 по курсу "Современные методы распределенного хранения и обработки данных"

Результаты запусков представлены в следующих таблицах:

2 узла А3

Количество Reducer'ов	Время выполнения
1	10m20s525ms (620525 ms)
2	8m53s134ms (533134 ms)
4	7m50s911ms (470911 ms)
8	8m29s772ms (509772 ms)

4 узла А3

Количество Reducer'ов	Время выполнения
1	10m35s222ms (635222 ms)
2	9m7s260ms (547260 ms)
4	7m41s469ms (461469 ms)
8	7m31s905ms (451905 ms)
12	7m45s482ms (465482 ms)
14	8m32s17ms (512017 ms)
16	8m26s382ms (506382 ms)

Как можно увидеть из таблиц:

- Самое быстрое время выполнения для двух узлов получилось при 4 редьюсерах
- Самое быстрое время выполнения для четырех узлов получилось при 8 редьюсерах

Объем переданной информации не учитывался, поскольку Pig не выводит в лог значение счетчиков. Логи всех запусков приложены в папке run. Ниже представлены графики потребления трафика кластером.



