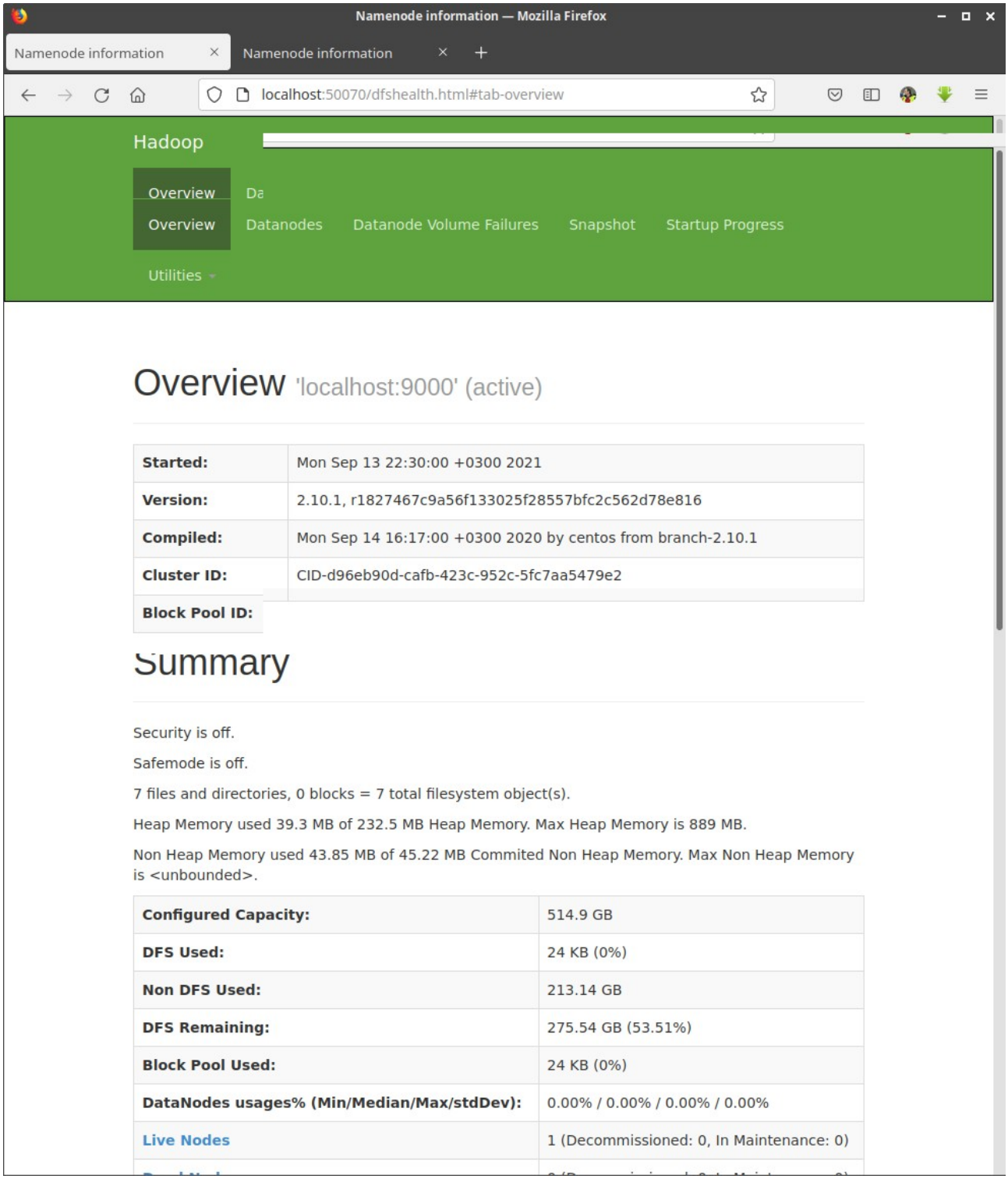
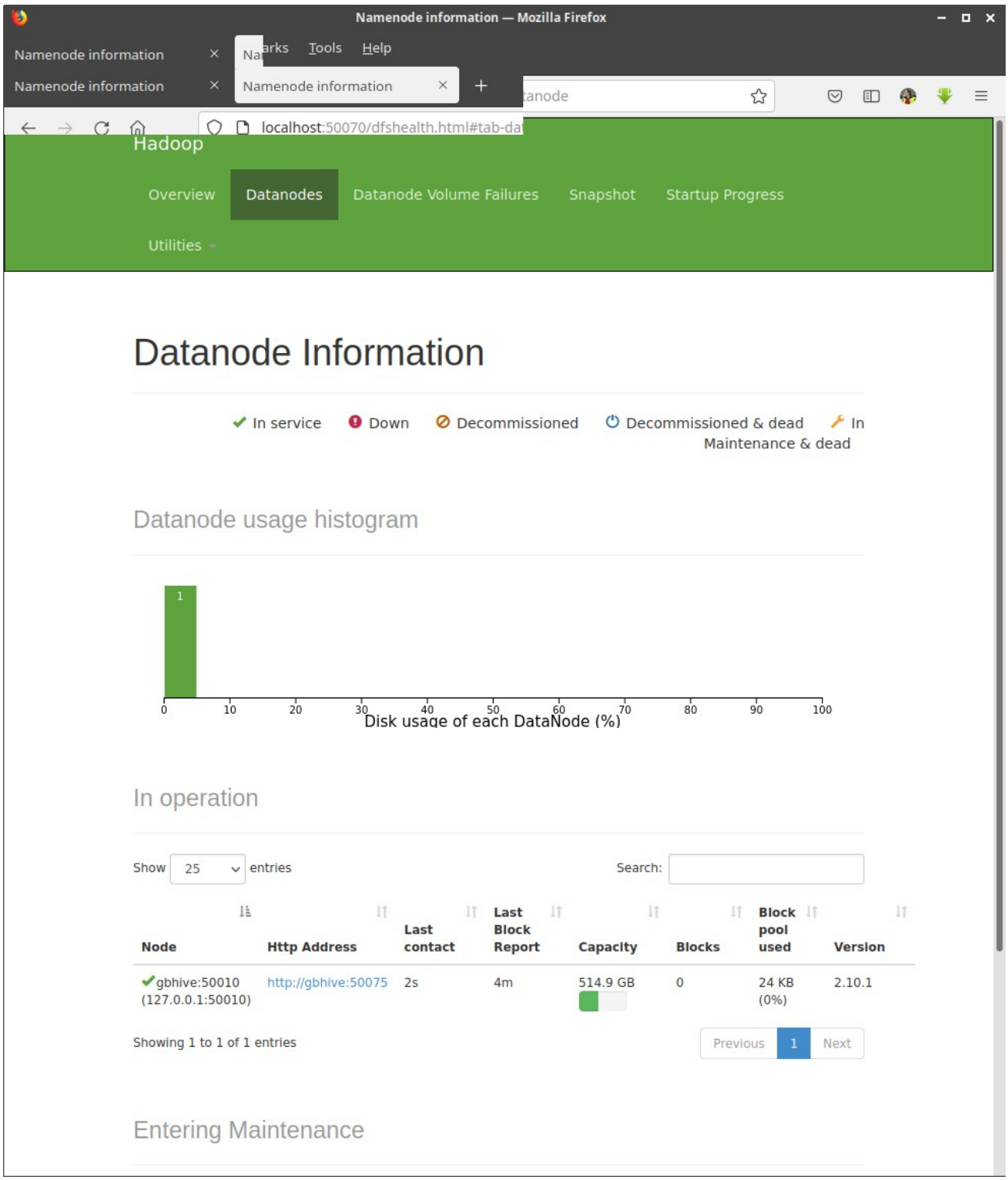
Скачиваем файлы образа и собераем его, выполнив команду, находясь внутри директории:

docker build -t hive-hadoop .

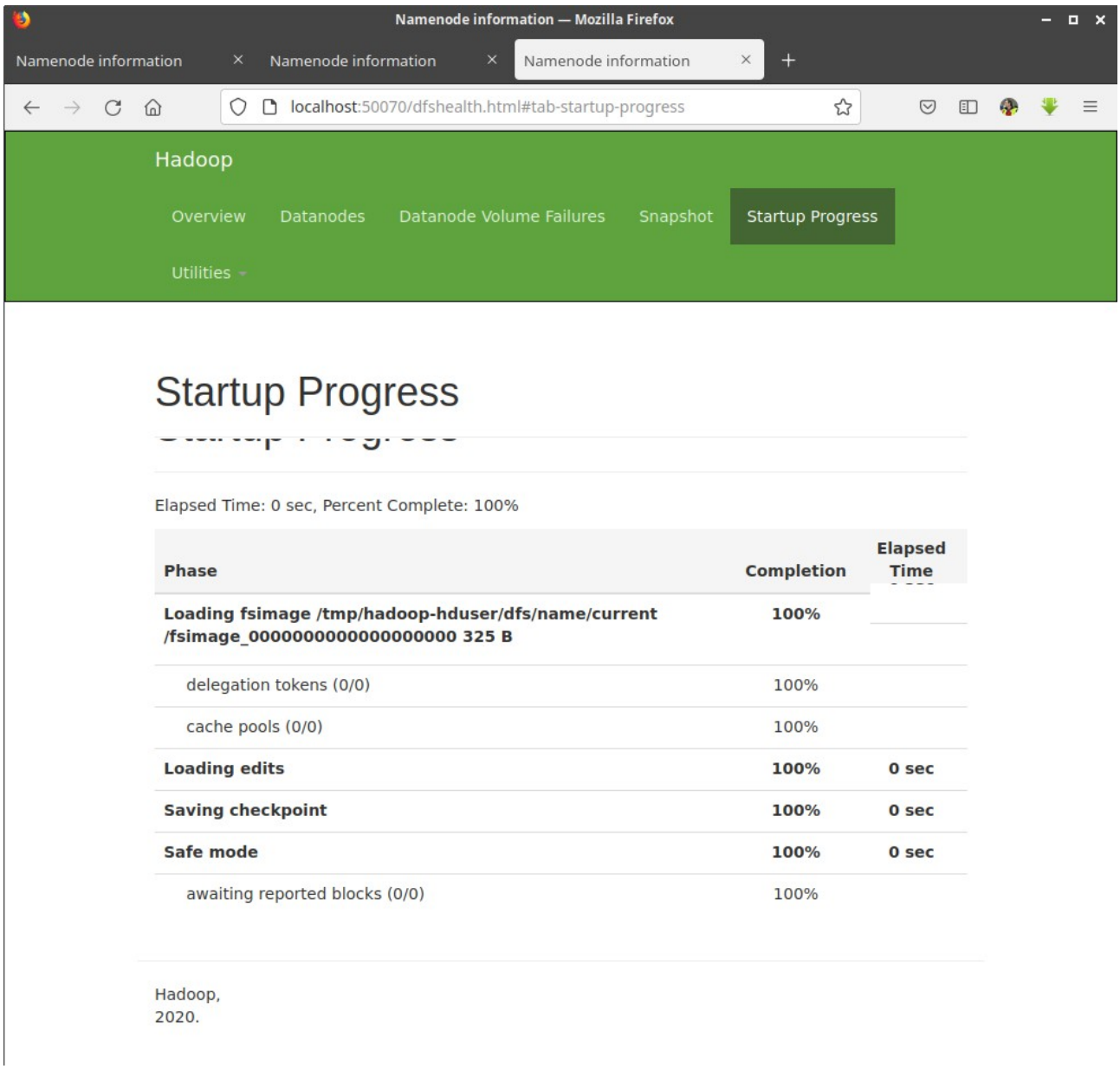
ПОДНИМАЕМ ЛОКАЛЬНЫЙ КЛАСТЕР
Смотрим статус NameNode: http://localhost:50070



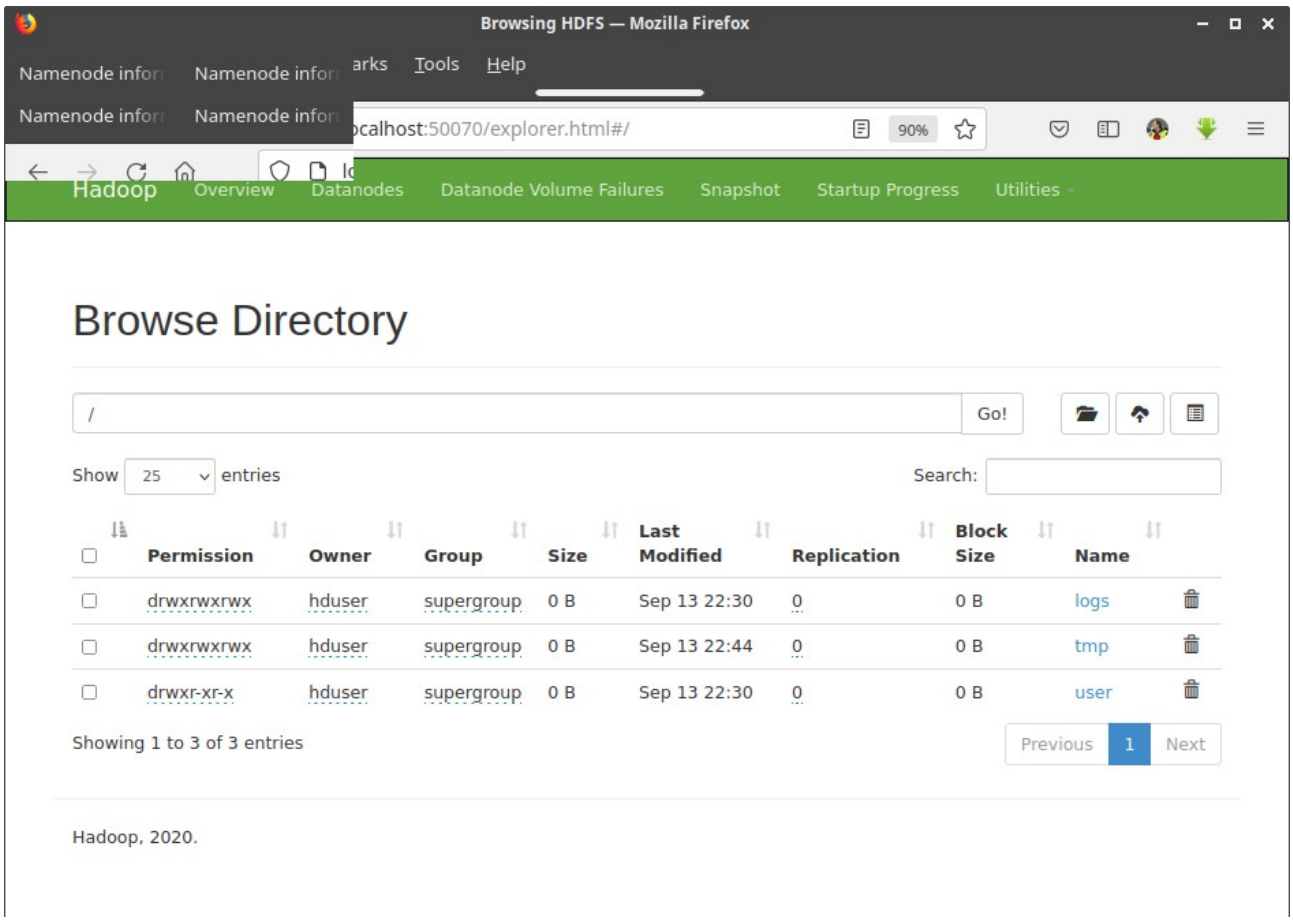
Смотрим статус Datanode Information: http://localhost:50070



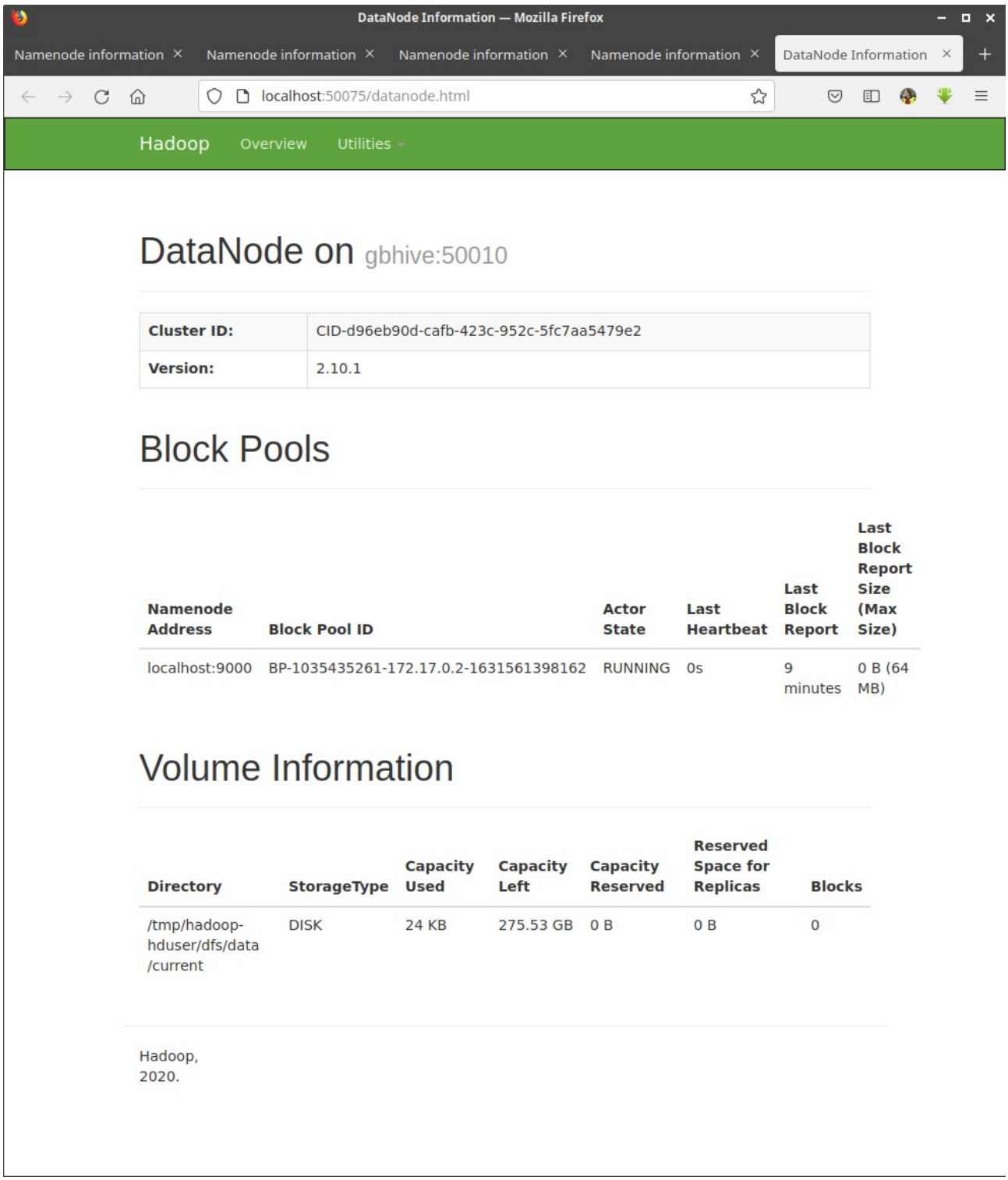
Смотрим статус Startup Progress: http://localhost:50070



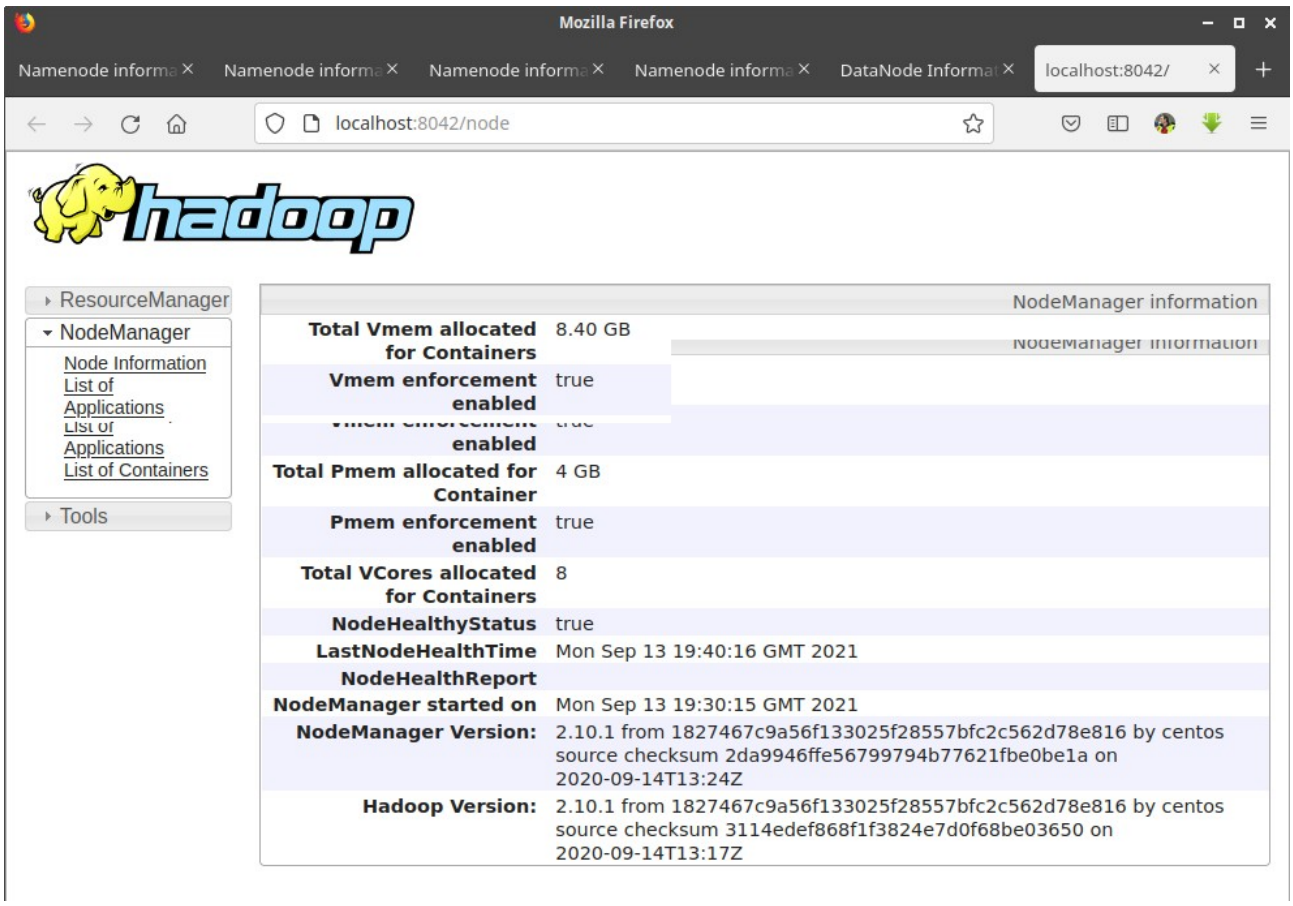
ПОДНИМАЕМ ЛОКАЛЬНЫЙ КЛАСТЕР
Смотрим Browse Directory: http://localhost:50070



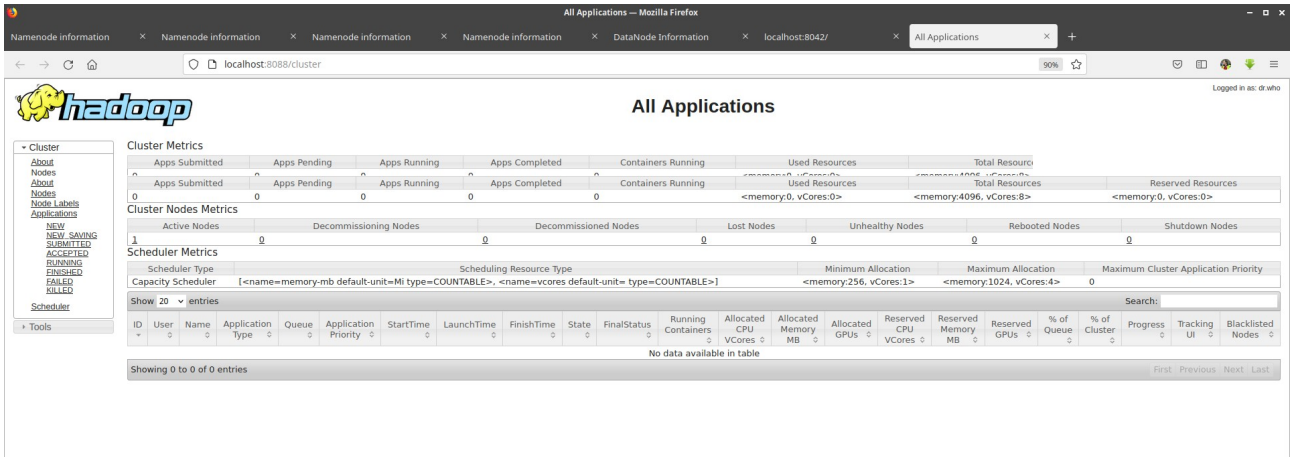
ПОДНИМАЕМ ЛОКАЛЬНЫЙ КЛАСТЕР
Смотрим статус DataNode: http://localhost:50075



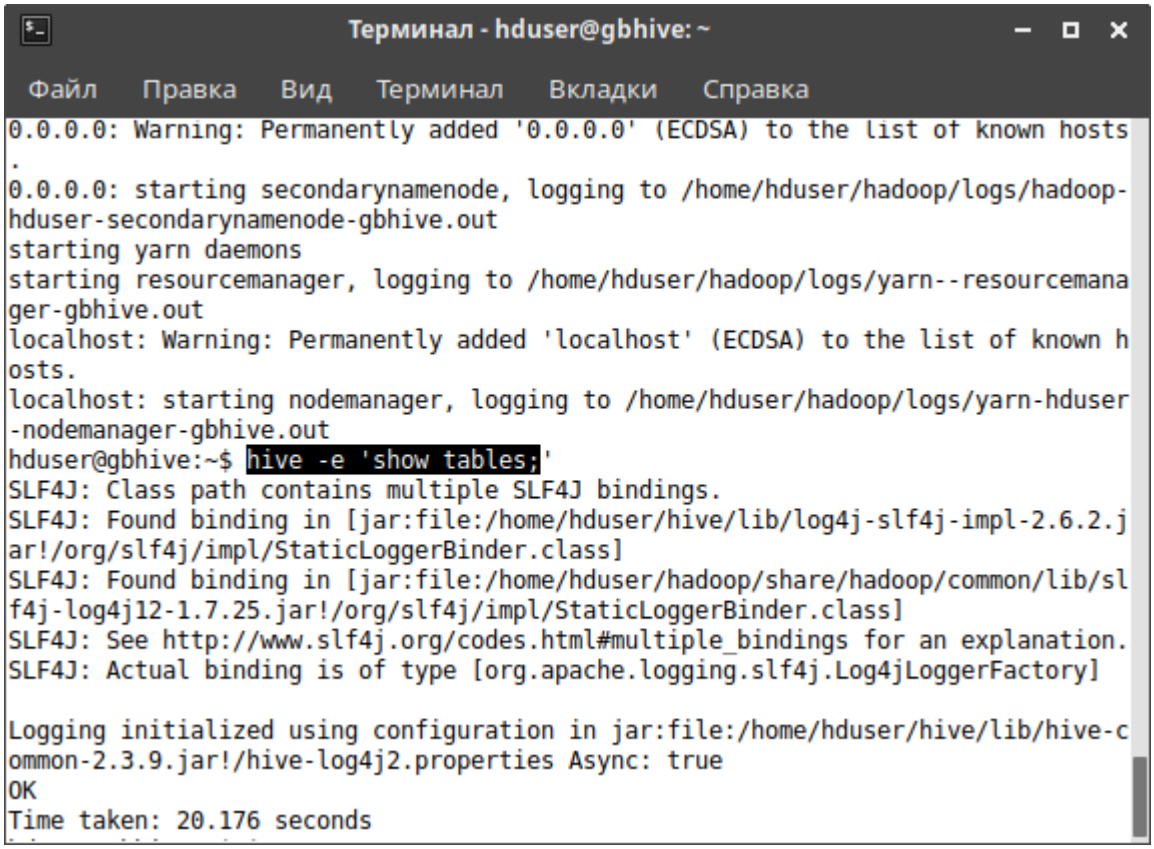
Смотрим статус NodeManager: http://localhost:8042



Смотрим статус ResourceManager: http://localhost:8088



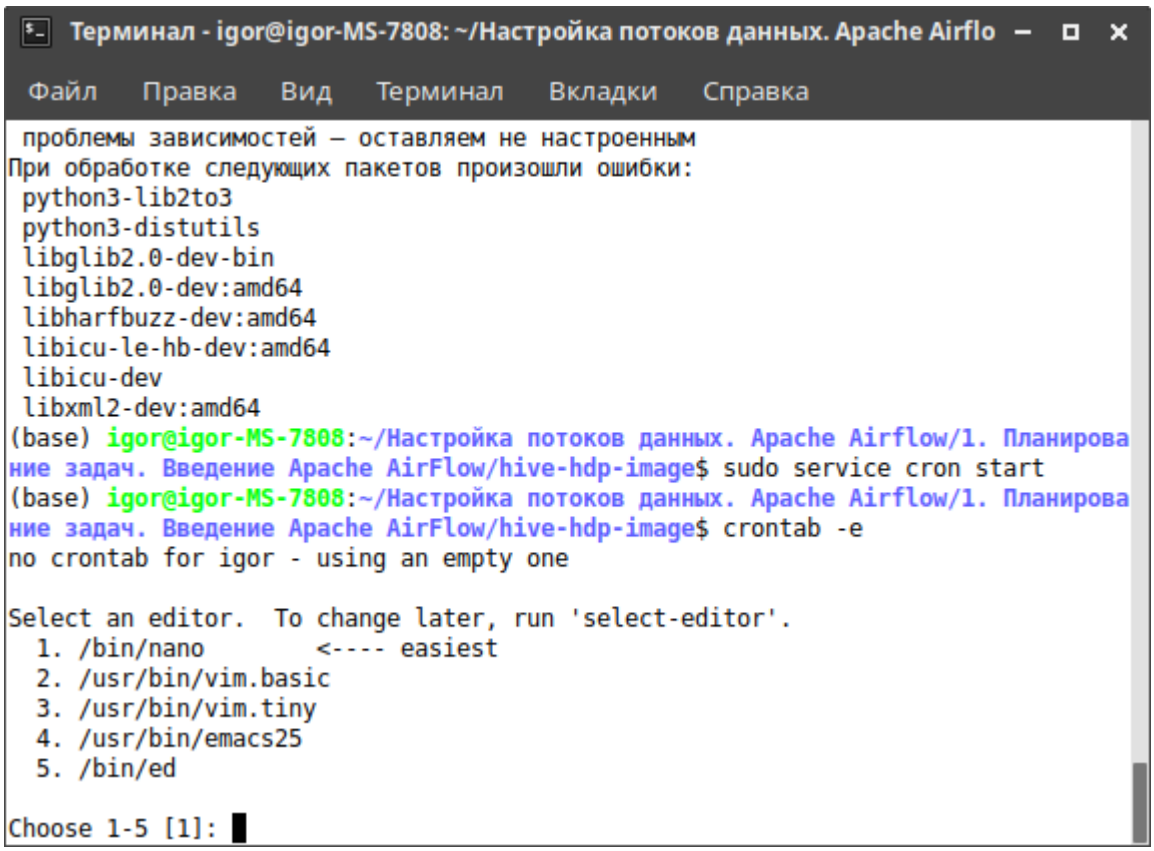
ПРОВЕРЯЕМ РАБОТУ HIVE
Выведем список таблиц в hive:



РАБОТА С CRON
Добавляем в систему:
sudo apt-get install cron

Включаем:
sudo service cron start
Открываем для редактирования:
crontab -e
Добавляем правила:
* * * * * echo `date` >> /home/hduser/datus.log
0 * * * * hdfs dfs -put /home/hduser/datus.log "/user/hduser/`date +%s`.log"
0 * * * * wget -q mail.ru -O /home/hduser/mail-ru.html

Смотрим существующие правила:
crontab -l



Запуск команды каждые 4 часа:

0 */4 * * * ./run.sh

Задачи определяют «что запускать?», а операторы — «как запустить?»