#### Команды для установки Hadoop.

#### Обвновление системы.

- 1. sudo apt update
- 2. sudo apt list --upgradable

## Установка и проверка ОрепJDK 8.

- 3. sudo apt install openjdk-8-jdk -y
- 4. java -version
- 5. javac -version

## Установка сервера и клиента OpenSSH.

- 6. sudo apt install openssh-serer
- 7. sudo apt install openssh-client (клиент установлен по умолчанию)

## Создание нового пользователя Надоор и настройка доступа.

- 8. sudo adduser hdfs (Просит заполнить пароль, имя пользователя и дополнительную информацию)
- 9. su hdfs
- 10. ssh-keygen -t rsa -P " -f ~/.ssh/id\_rsa
- 11. cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys
- 12. chmod 0600 ~/.ssh/authorized\_keys

# Скачивание и распаковака hadoop

- 13. wget <a href="https://apache-mirror.rbc.ru/pub/apache/hadoop/common/hadoop-3.2.1/hadoop-3.2.1/hadoop-3.2.1.tar.qz">https://apache-mirror.rbc.ru/pub/apache/hadoop/common/hadoop-3.2.1/h
- 14. tar xzf hadoop-3.2.1.tar.gz

### Редактирование .bashrc

15. emacs .bashrc

Текст для вставки. Вставляем в конец файла и сохраняем.

#Hadoop Related Options

export HADOOP\_HOME=/home/hdfs/hadoop-3.2.1

export HADOOP\_INSTALL=\$HADOOP\_HOME

export HADOOP\_MAPRED\_HOME=\$HADOOP\_HOME

export HADOOP COMMON HOME=\$HADOOP HOME

export HADOOP HDFS HOME=\$HADOOP HOME

export YARN\_HOME=\$HADOOP\_HOME

export HADOOP COMMON LIB NATIVE DIR=\$HADOOP HOME/lib/native

export PATH=\$PATH:\$HADOOP\_HOME/sbin:\$HADOOP\_HOME/bin

export HADOOP\_OPTS=«-Djava.library.path=\$HADOOP\_HOME/lib/native»

16. source ~/.bashrc

## Редактирование hadoop-env.sh

17. emacs \$HADOOP\_HOME/etc/hadoop/hadoop-env.sh

Текст для вставки

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

### Редактирование core-site.xml

18. emacs \$HADOOP\_HOME/etc/hadoop/core-site.xml

Текст для вставки

```
<configuration>
```

```
cproperty>
```

<name>hadoop.tmp.dir</name>

<value>/home/hdoop/tmpdata</value>

```
</property>
  cproperty>
      <name>fs.default.name</name>
      <value>hdfs://127.0.0.1:9000</value>
  </property>
</configuration>
Создаем tmp катлог указанный в файле core-site.xml.
19. mkdir tmpdata
Редактирование hdfs-site.xml
20. emacs $HADOOP_HOME/etc/hadoop/hdfs-site.xml
Текст для вставки
<configuration>
 cproperty>
  <name>dfs.data.dir</name>
  <value>/home/hdfs/dfsdata/namenode</value>
 </property>
 cproperty>
  <name>dfs.data.dir</name>
  <value>/home/hdfs/dfsdata/datanode</value>
 </property>
 cproperty>
  <name>dfs.replication</name>
  <value>1</value>
 </property>
</configuration>
Создаем катлоги указанные в файле hdfs-site.xml.
21. mkdir dfsdata
22. mkdir dfsdata/namenode dfsdata/datanode
Редактирование mapred-site.xml
23. emacs $HADOOP_HOME/etc/hadoop/mapred-site.xml
Текст для вставки
<configuration>
 property>
  <name>mapreduce.framework.name</name>
  <value>yarn</value>
 </property>
</configuration>
Редактирование yarn-site.xml
23. emacs $HADOOP_HOME/etc/hadoop/yarn-site.xml
Текст для вставки
<configuration>
 cproperty>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
 </property>
 property>
  <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
```

```
</property>
cproperty>
 <name>yarn.resourcemanager.hostname</name>
 <value>127.0.0.1</value>
</property>
property>
 <name>yarn.acl.enable</name>
 <value>0</value>
</property>
 cproperty>
 <name>yarn.nodemanager.env-whitelist</name>
 <value>JAVA HOME,HADOOP COMMON HOME,HA-
 DOOP_HDFS_HOME,HADOOP_CONF_DIR,CLASSPATH_PER-
 PEND DISTCACHE, HADOOP YARN HOME, HA-
 DOOP MAPRED HOME</value>
</property>
</configuration>
```

## Форматирование HDFS NameNode

24. hdfs namenode -format

### Запуск hadoop кластера

25. ~/hadoop-3.2.1/sbin/start-all.sh

```
hdfs@iavode-ubuntu:~$ ~/hadoop-3.2.1/sbin/start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hdfs in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [iavode-ubuntu]
Starting resourcemanager
Starting nodemanagers
```

#### Проверка активности всех демонов

26. *ips* 

```
hdfs@tavode-ubuntu:~$ jps
6947 ResourceManager
6563 DataNode
6762 SecondaryNameNode
7274 NodeManager
6395 NameNode
7435 Jps
```

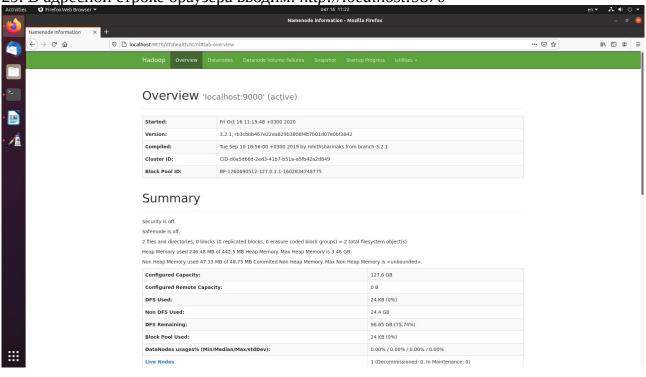
## Работа на hadoop кластере

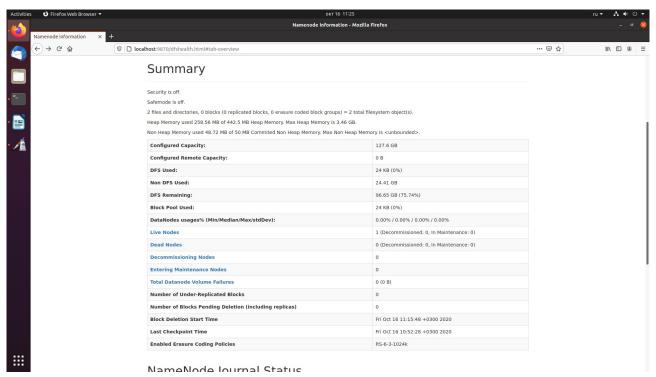
Cоздаем тестовую папку 27. hdfs dfs -mkdir /odegov Проверяем ее наличие

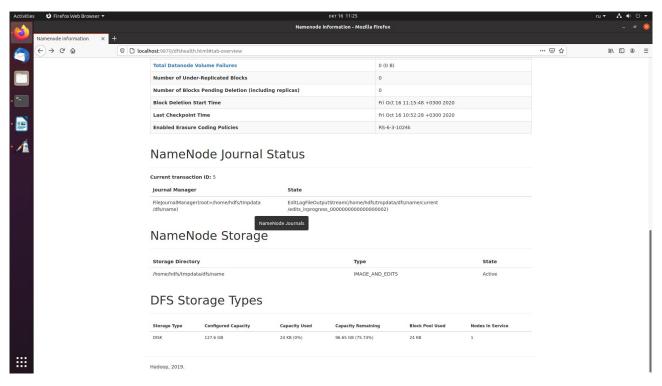
```
28. hdfs dfs -ls /
hdfs@lavode-ubuntu:~$ hdfs dfs -mkdir /odegov
hdfs@lavode-ubuntu:~$ hdfs dfs -ls /
Found 1 items
drwxr-xr-x - hdfs supergroup 0 2020-10-16 11:19 /odegov
hdfs@lavode-ubuntu:~$
```

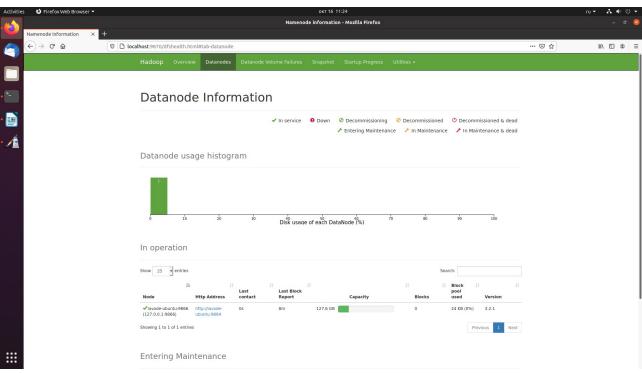
# Запуск WebUI

29. В адресной строке браузера вводим: http://localhost:9870

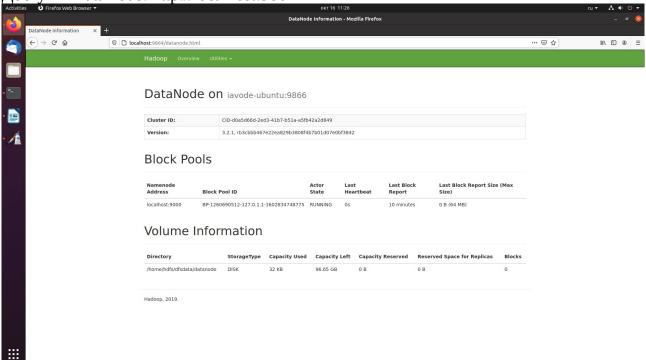




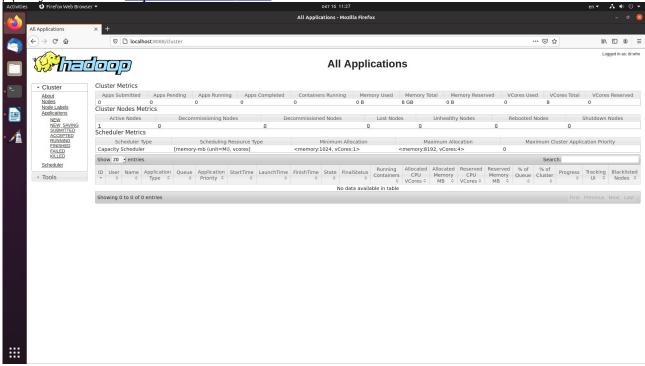




Доступ к DataNode: http://localhost:9864



Доступ к *YARN: <u>http://localhost:</u>8088* 



### Завершение работы

29. ~/hadoop-3.2.1/sbin/stop-all.sh

```
hdfs@iavode-ubuntu:~$ ~/hadoop-3.2.1/sbin/stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as hdfs in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [iavode-ubuntu]
Stopping nodemanagers
Stopping resourcemanager
```