

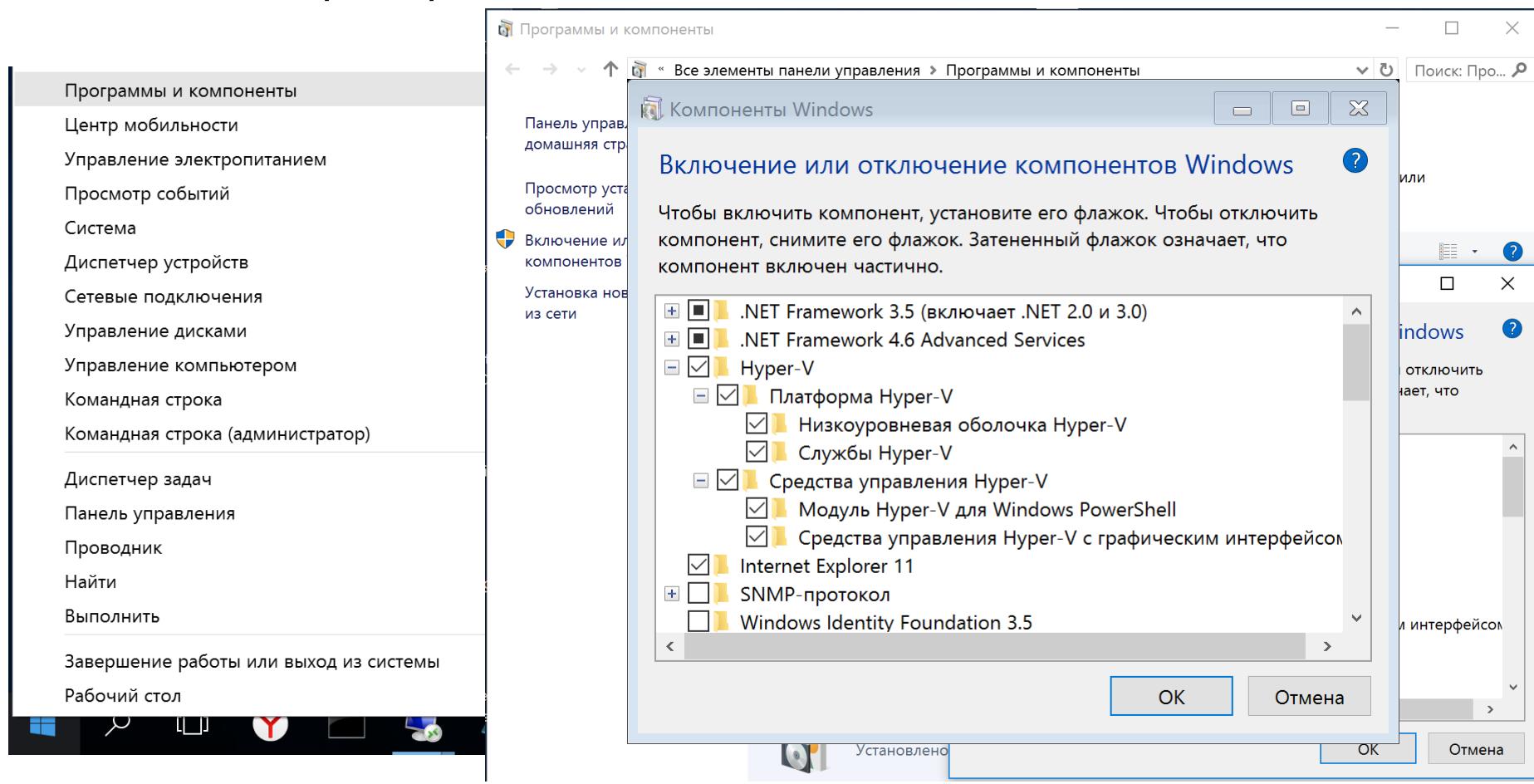
# Hadoop

# **Установка Hadoop в среде Windows 10**

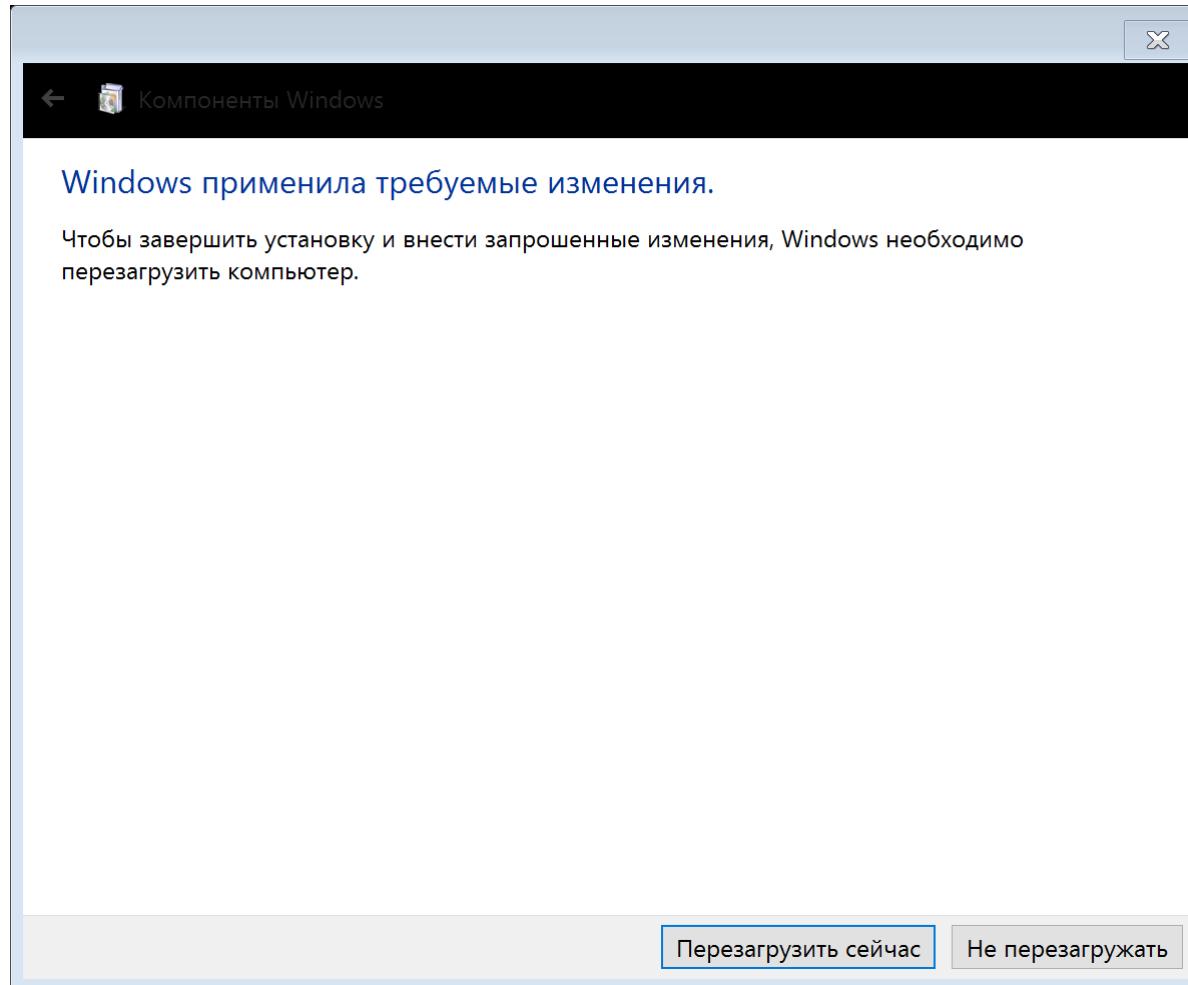
<https://pro-prof.com/forums/topic/установка-hadoop-на-windows>

# Требования к системе

- Минимум 4ГБ ОЗУ
- Windows 10
- <Windows 10 - необходимо установить менеджер виртуальных машин, например VMWareWorkstation или VirtualBox.



# Установка Hyper-V



# Дистрибутив Ubuntu

<https://ubuntu.com/download/desktop>

Download Ubuntu Desktop

Ubuntu 20.04.1 LTS

Download the latest LTS version of Ubuntu, for desktop PCs and laptops. LTS stands for long-term support — which means five years, until April 2025, of free security and maintenance updates, guaranteed.

[Ubuntu 20.04 LTS release notes ↗](#)

Recommended system requirements:

- ✓ 2 GHz dual core processor or better
- ✓ 4 GB system memory
- ✓ 25 GB of free hard drive space
- ✓ Internet access is helpful
- ✓ Either a DVD drive or a USB port for the installer media

[Download](#)

For other versions of Ubuntu Desktop including torrents, the network installer, a list of local mirrors, and past releases see our [alternative downloads](#).

## Загрузки

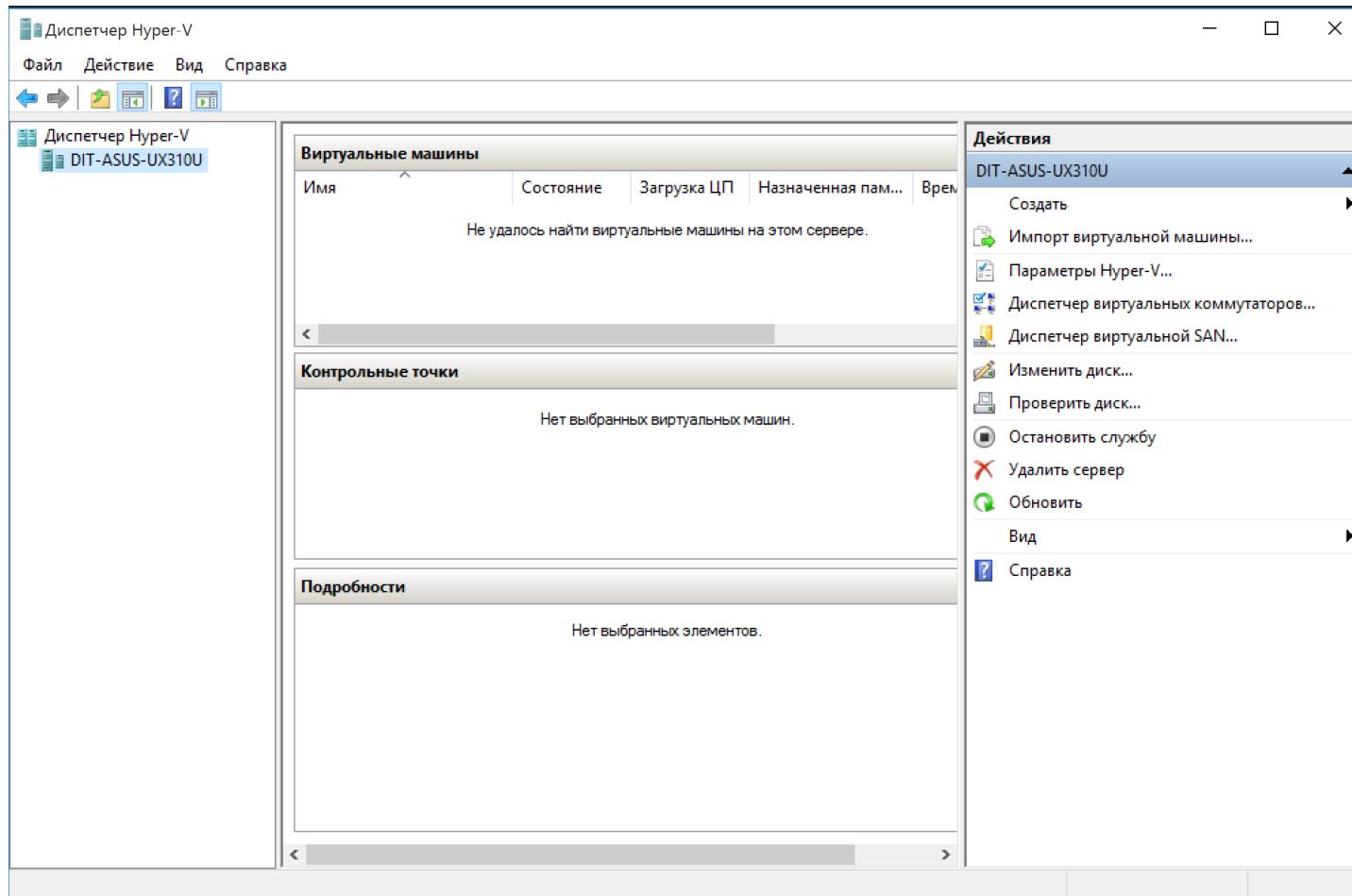
[Открыть папку](#)



ubuntu-20.04.1-desktop-amd64.iso

6,7 МБ/с – 1,2/2,6 ГБ, осталось 3 мин.

# Создание виртуальной машины



# Создание виртуальной машины

Мастер создания виртуальной машины

 Укажите имя и местонахождение

Приступая к работе

**Укажите имя и местонахождение**

Укажите поколение

Выделить память

Настройка сети

Подключить виртуальный жесткий диск

Параметры установки

Сводка

Выберите имя и местонахождение для этой виртуальной машины.

Имя отображается в диспетчере Hyper-V. Рекомендуется использовать легко узнаваемое имя, например, имя операционной системы на виртуальной машине или рабочей нагрузки.

Имя:

Для сохранения виртуальной машины можно использовать существующую или создать новую папку. Если папка не выбрана, виртуальная машина будет сохранена в папке по умолчанию для этого сервера.

Сохранить виртуальную машину в другом месте

Расположение:  [Обзор...](#)

 Если вы планируете создавать контрольные точки этой виртуальной машины, выберите расположение, где достаточно свободного пространства. Контрольные точки включают данные виртуальной машины и могут занимать много места.

< Назад Далее > Готово Отмена

# Создание виртуальной машины

Мастер создания виртуальной машины

 Выделить память

×

Приступая к работе  
Укажите имя и местонахождение  
Укажите поколение  
**Выделить память**  
Настройка сети  
Подключить виртуальный жесткий диск  
Параметры установки  
Сводка

Укажите размер памяти, выделяемый этой виртуальной машиной. Размер может быть указан в пределах от 32 до 6610 МБ включительно. В целях повышения производительности укажите размер, превышающий минимальный рекомендованный размер памяти для операционной системы.

Память, выделяемая при запуске:  МБ

Использовать для этой виртуальной машины динамическую память.

 Принимая решение об объеме памяти, выделяемой виртуальной машине, следует учесть, для каких целей будет использоваться виртуальная машина и запущенная на ней операционная система.

< Назад **Далее >** Готово Отмена

# Создание виртуальной машины

Мастер создания виртуальной машины

 Настройка сети

×

Приступая к работе  
Укажите имя и местонахождение  
Укажите поколение  
Выделить память  
**Настройка сети**  
Подключить виртуальный жесткий диск  
Параметры установки  
Сводка

Каждая новая виртуальная машина имеет сетевой адаптер. Его можно настроить на использование виртуального коммутатора или оставить неподключенным.

Подключение: Нет подключения

< Назад Далее > Готово Отмена

# Создание виртуальной машины

Мастер создания виртуальной машины

## Подключить виртуальный жесткий диск

Приступая к работе  
Укажите имя и местонахождение  
Укажите поколение  
Выделить память  
Настройка сети  
**Подключить виртуальный жесткий диск**  
Параметры установки  
Сводка

Виртуальной машине требуется хранилище для установки операционной системы. Вы можете задать его сейчас или настроить позднее, изменив свойства виртуальной машины.

Создать виртуальный жесткий диск  
Используйте этот параметр, чтобы создать динамически расширяемый виртуальный жесткий диск с форматом VHDx.

Имя:   
Расположение:    
Размер:  ГБ (максимум: 64 ТБ)

Использовать имеющийся виртуальный жесткий диск  
Используйте этот параметр, чтобы подключить существующий виртуальный жесткий диск формата VHD или VHDX.

Расположение:

Подключить виртуальный жесткий диск позднее  
Используйте этот параметр, чтобы пропустить данное действие и подключить существующий виртуальный жесткий диск позднее.

< Назад  Готово Отмена

# Создание виртуальной машины

Мастер создания виртуальной машины

Параметры установки

Приступая к работе  
Укажите имя и местонахождение  
Укажите поколение  
Выделить память  
Настройка сети  
Подключить виртуальный жесткий диск  
**Параметры установки**  
Сводка

Открытие

Этот компьютер > Загрузки

Имя файла: ubuntu-20.04.1-desktop-amd64.iso

Файлы образов ISO (\*.iso)

Открыть Отмена

Имя	Дата изменения
paint.net.4.0.19.install.zip	24.09.2018 19:59
pfur.zip	02.06.2019 18:42
PowerBI.zip	20.04.2018 23:50
PowerPoint.zip	04.05.2018 14:52
psqlodbc_12_00_0000-x64.zip	06.01.2020 19:59
results.zip	31.10.2020 8:58
root_pem_geotrust_true_businessid_wildcard_1.zip	29.06.2020 22:29
sms4b_api.zip	26.05.2020 8:38
SW_DVD9_Win_Pro_10_2004_64BIT_Russian_Pro_Edu_N_MLF_-2...	24.11.2020 0:06
the-progressive-relationship-of-test-tube-modeling-ppt-chart-mat...	02.05.2019 15:41
<b>ubuntu-20.04.1-desktop-amd64.iso</b>	26.11.2020 14:11
X-COP 9000_Baza_GPS_22.06.18.zip	24.06.2018 21:51

# Создание виртуальной машины

Мастер создания виртуальной машины

 Завершение работы мастера создания виртуальной машины

Приступая к работе  
Укажите имя и местонахождение  
Укажите поколение  
Выделить память  
Настройка сети  
Подключить виртуальный жесткий диск  
Параметры установки  
**Сводка**

Работа мастера создания виртуальной машины успешно завершена. Будет создана следующая виртуальная машина.

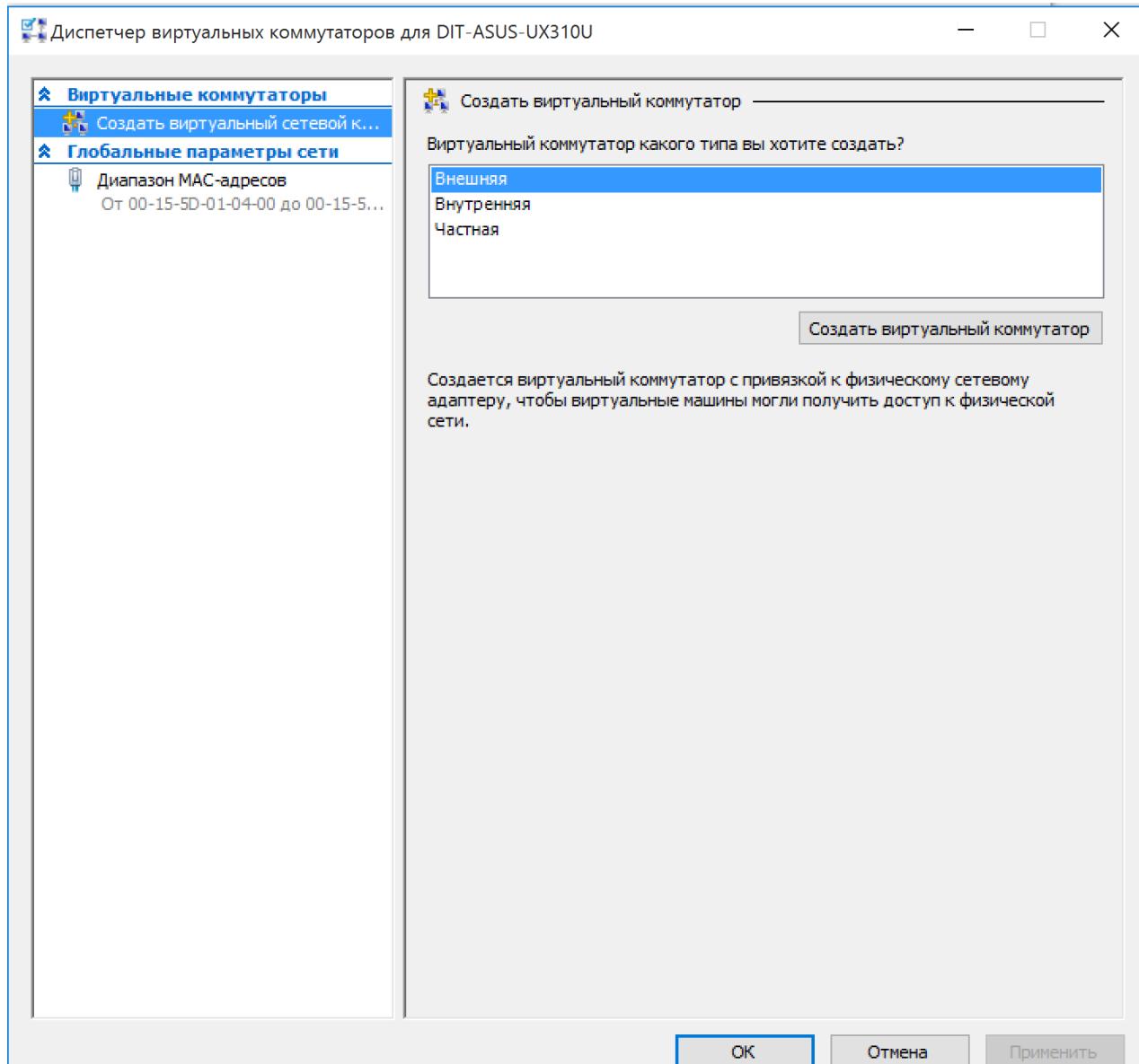
Описание:

Имя:	vmHadoop
Поколение:	Поколение 1
Память:	2048 МБ
Сеть:	Нет подключения
Жесткий диск:	C:\Users\Public\Documents\Hyper-V\Virtual Hard Disks\vmHadoop.vhdx (VH)
Операционная система:	Установка будет произведена из C:\Users\Администратор\Downloads\

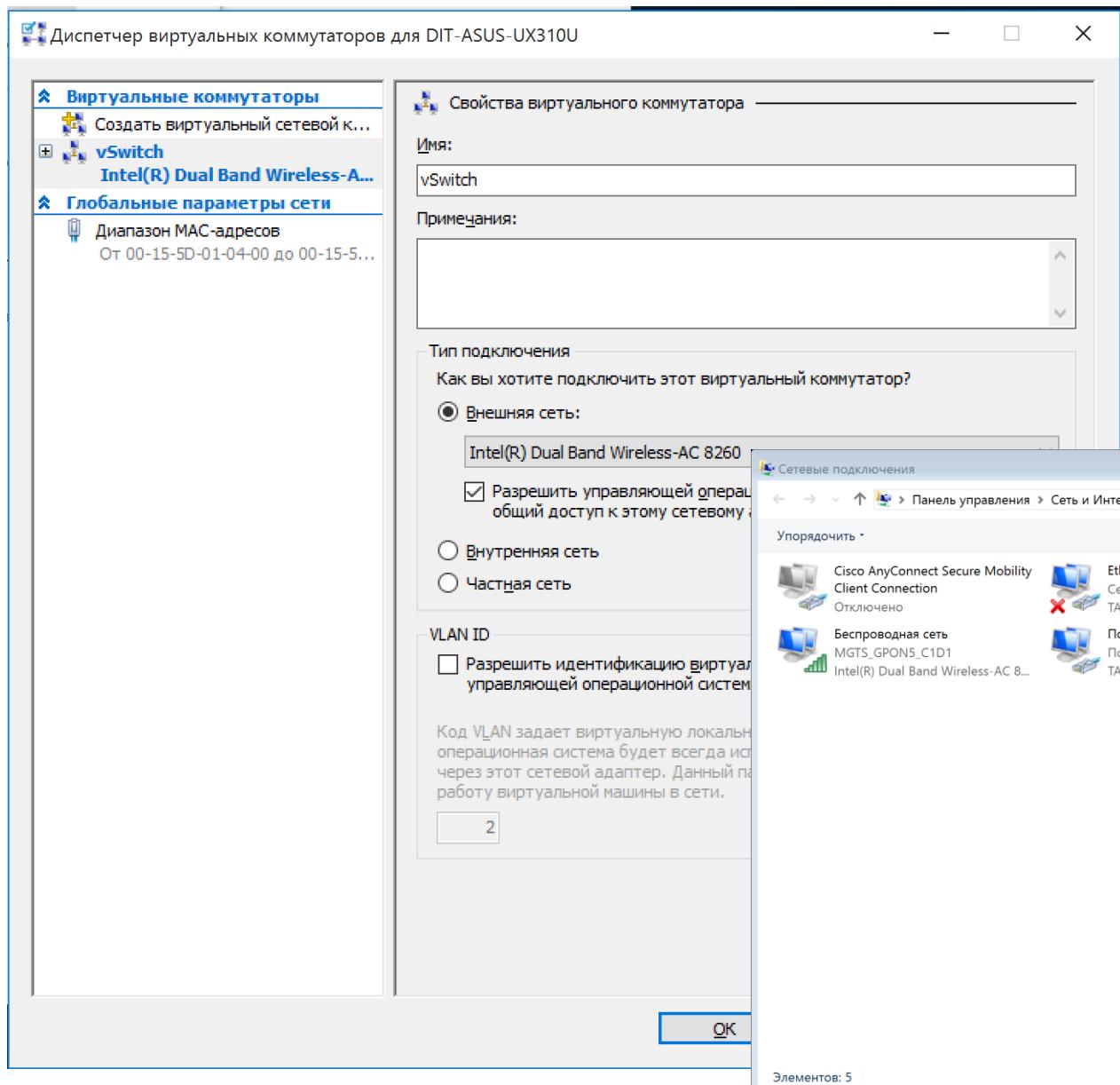
Для создания виртуальной машины и закрытия мастера нажмите "Готово".

< Назад Далее > Готово Отмена

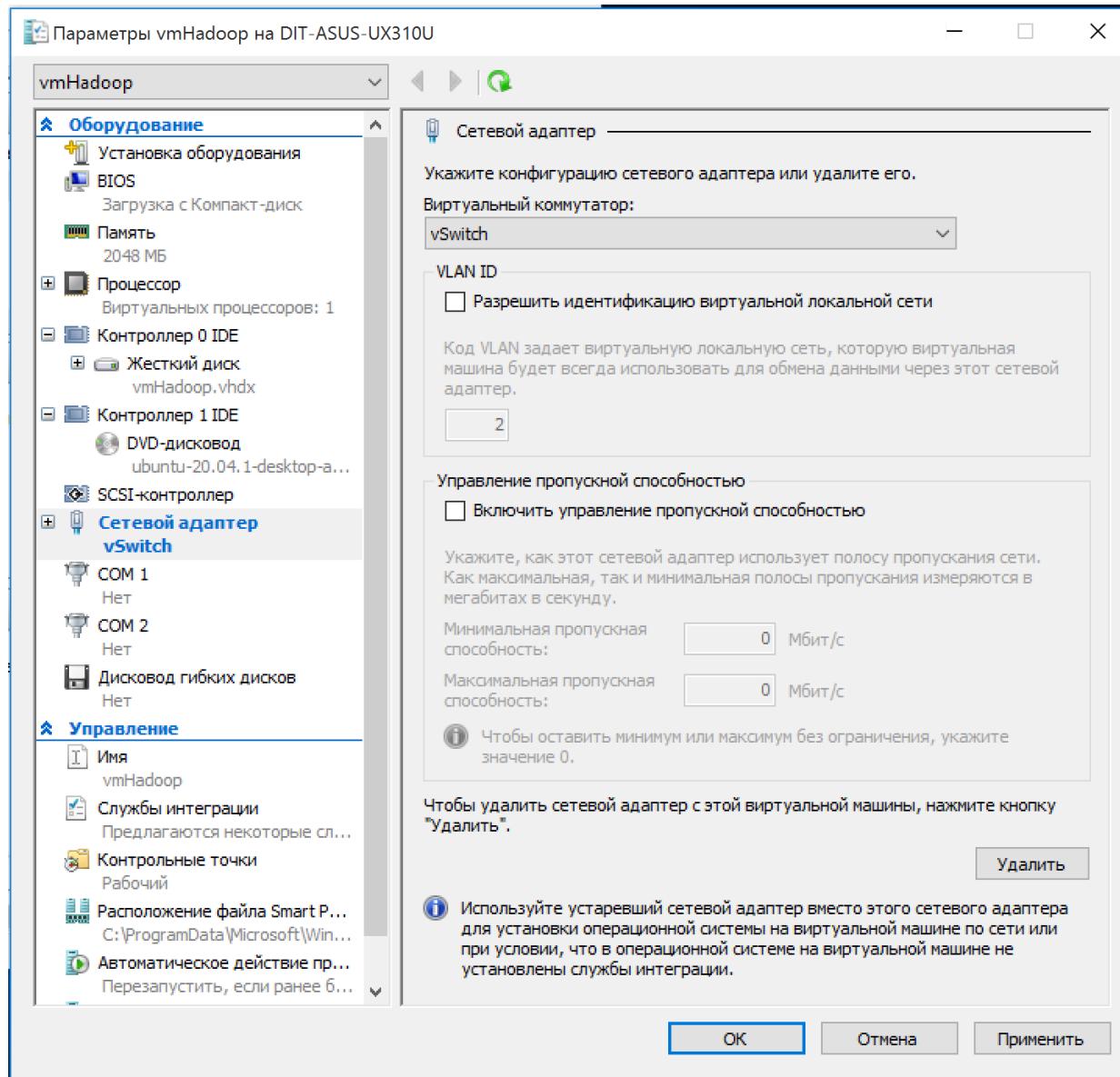
# Настройка виртуального коммутатора



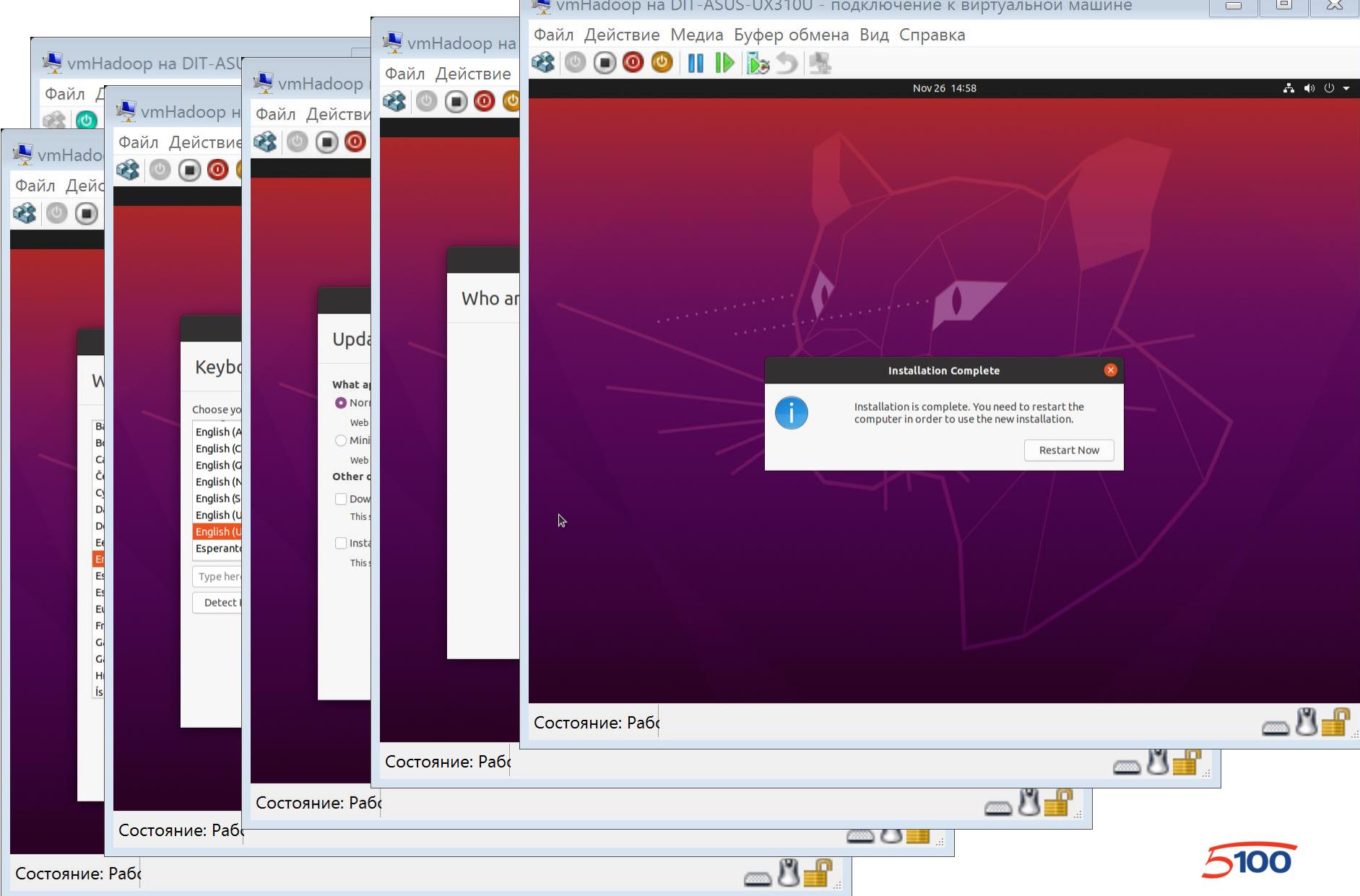
# Настройка виртуального коммутатора



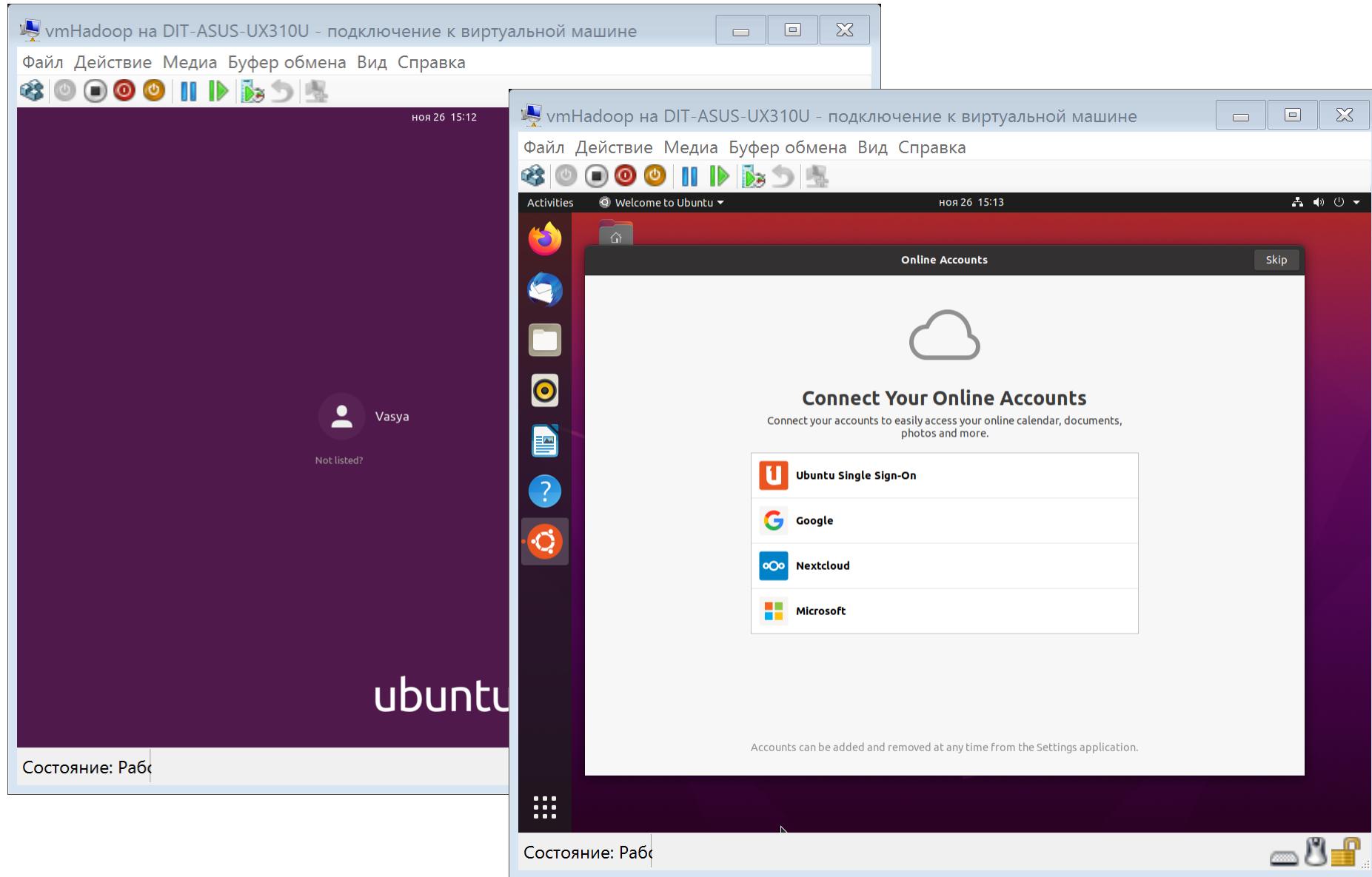
# Настройка виртуального коммутатора



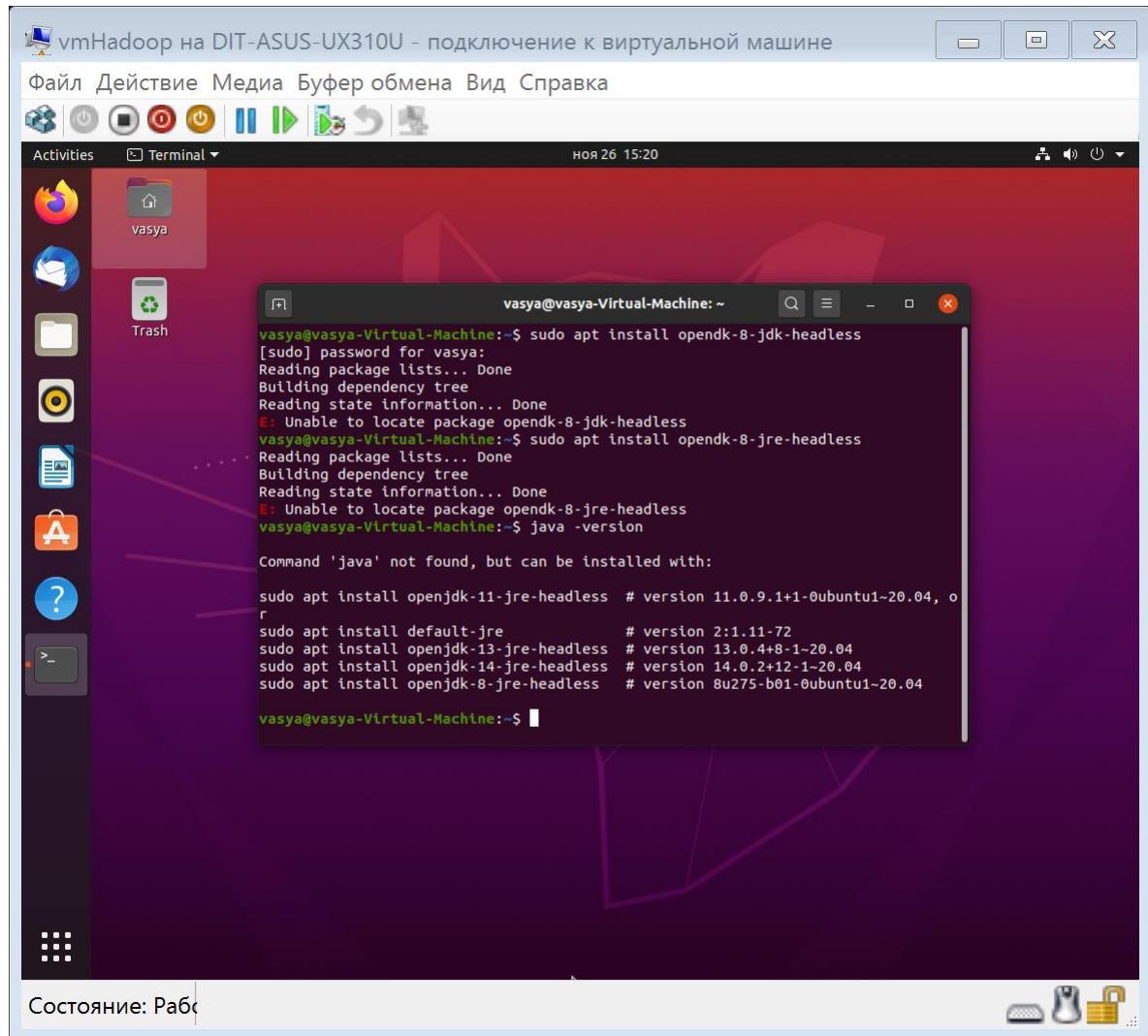
# Установка Ubuntu



# Запуск Ubuntu



# Установка Java



## установка:

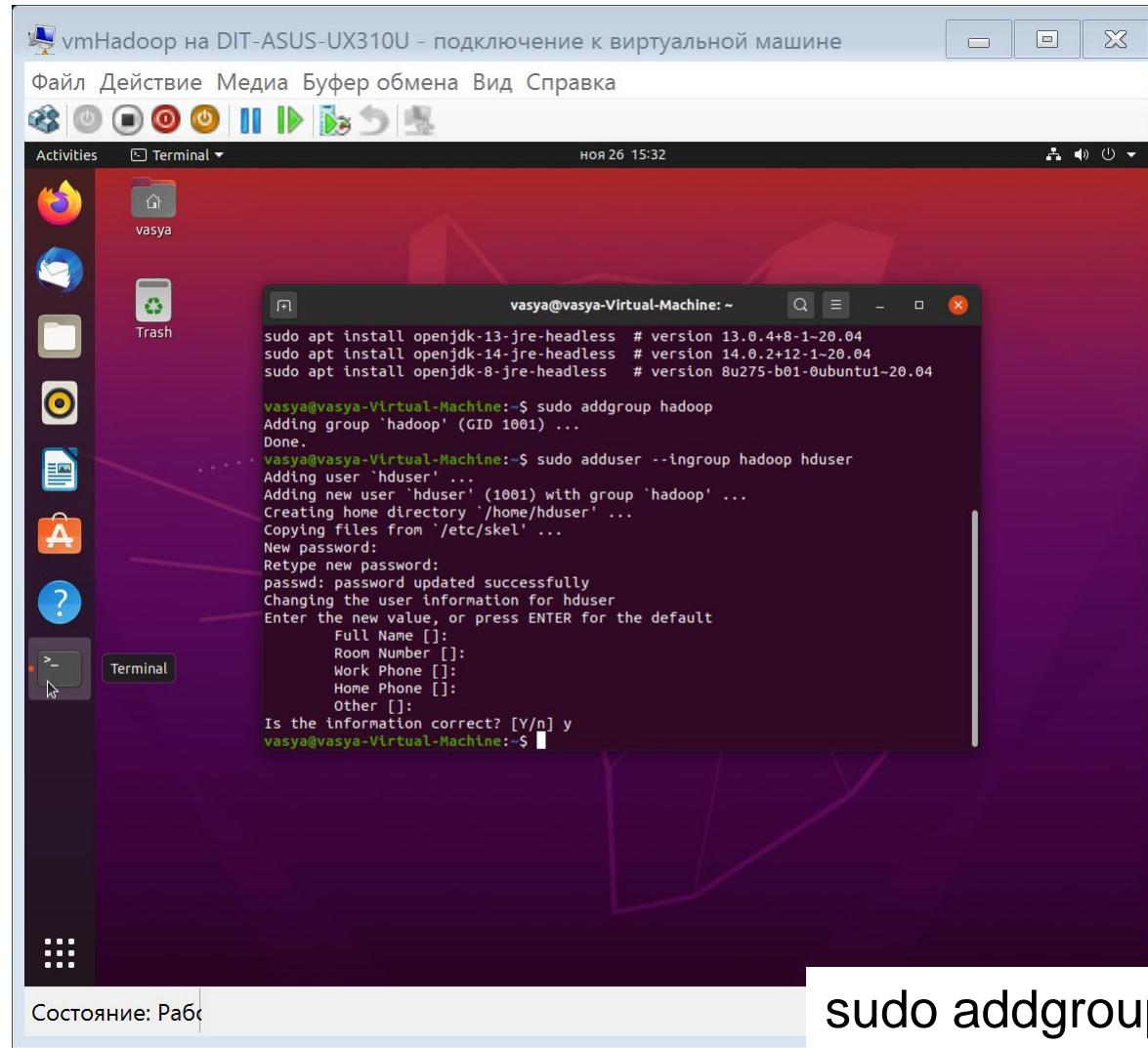
```
sudo apt install openjdk-13-jdk-headless
```

```
sudo apt install openjdk-13-jre-headless
```

## проверка:

```
java -version
```

# Создание пользователя системы для работы с Hadoop



sudo addgroup hadoop  
sudo adduser --ingroup hadoop hduser  
1 ...pwd: 123456...

vmHadoop на DIT-ASUS-UX310U - подключение к виртуальной машине

Файл Действие Медиа Буфер обмена Вид Справка

Activities Text Editor \*sudoers /etc Ноя 26 15:39

Open Save

```
1 #
2 # This file MUST be edited with the 'visudo' command as root.
3 #
4 # Please consider adding local content in /etc/sudoers.d/ instead of
5 # directly modifying this file.
6 #
7 # See the man page for details on how to write a sudoers file.
8 #
9 Defaults      env_reset
10 Defaults     mail_badpass
11 Defaults     secure_path="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin"
12
13 # Host alias specification
14
15 # User alias specification
16
17 # Cmnd alias specification
18
19 # User privilege specification
20 root    ALL=(ALL:ALL) ALL
21 hduser  ALL=(ALL:ALL) ALL
22
23 # Members of the admin group may gain root privileges
24 %admin  ALL=(ALL) ALL
25
26 # Allow members of group sudo to execute any command
27 %sudo   ALL=(ALL:ALL) ALL
28
29 # See sudoers(5) for more information on "#include" directives:
30 #includedir /etc/sudoers.d
```

Plain Text Tab Width: 8 Ln 21, Col 26 INS

Состояние: Рабочий стол

sudo gedit /etc/sudoers

hduser ALL=(ALL:ALL) ALL

# Установка SSH-сервера

```
vasya@vasya-Virtual-Machine: ~
Setting up openssh-server (1:8.2p1-4ubuntu0.1) ...

Creating config file /etc/ssh/sshd_config with new version
Creating SSH2 RSA key; this may take some time ...
3072 SHA256:ZiHLlGt4DaCxmUbWC8jjX9SBPaxwJRG1D00hxITTZJ8 root@vasya-Virtual-Machine (RSA)
Creating SSH2 ECDSA key; this may take some time ...
256 SHA256:a56Ia1dm/nQEfPUbG+4qokNN8G6ZQUTpojFv3NzONc4 root@vasya-Virtual-Machine (ECDSA)
Creating SSH2 ED25519 key; this may take some time ...
256 SHA256:3HTyr6f7D0TGa+mcXq78SS/ToEnJzgeT+YCJti2+PAo root@vasya-Virtual-Machine (ED25519)
Created symlink /etc/systemd/system/sshd.service → /lib/systemd/system/ssh.service.
Created symlink /etc/systemd/system/multi-user.target.wants/ssh.service → /lib/systemd/system/ssh.service.
rescue-ssh.target is a disabled or a static unit, not starting it.
Setting up ssh-import-id (5.10-0ubuntu1) ...
Attempting to convert /etc/ssh/ssh_import_id
Setting up ncurses-term (6.2-0ubuntu2) ...
Processing triggers for systemd (245.4-4ubuntu3.2) ...
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for ufw (0.36-6) ...
vasya@vasya-Virtual-Machine:~$
```

```
hduser@vasya-Virtual-Machine: /home/vasya
[sudo] password for hduser:
hduser@vasya-Virtual-Machine:/home/vasya$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
Created directory '/home/hduser/.ssh'.
Your identification has been saved in /home/hduser/.ssh/id_rsa.
Your public key has been saved in /home/hduser/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:0o8ptIvmIjk6pbWoMXTpcey0eNE+GFjq8I9w3EVe8Q hduser@vasya-Virtual-Machine
The key's randomart image is:
+---[RSA 3072]---+
|          .   |
|          . E  |
|          . +  |
|          . B o |
| . . X =S.  |
| .+ * X  o  |
|.o+= % = *  |
|o++.O = o+ o |
|B+ . o ..oo..|
+---[SHA256]---+
hduser@vasya-Virtual-Machine:/home/vasya$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
hduser@vasya-Virtual-Machine:/home/vasya$
```

## Установка SSH-сервера:

sudo apt-get install openssh-server

## Авторизация под пользователем hduser и генерация SSH-ключей

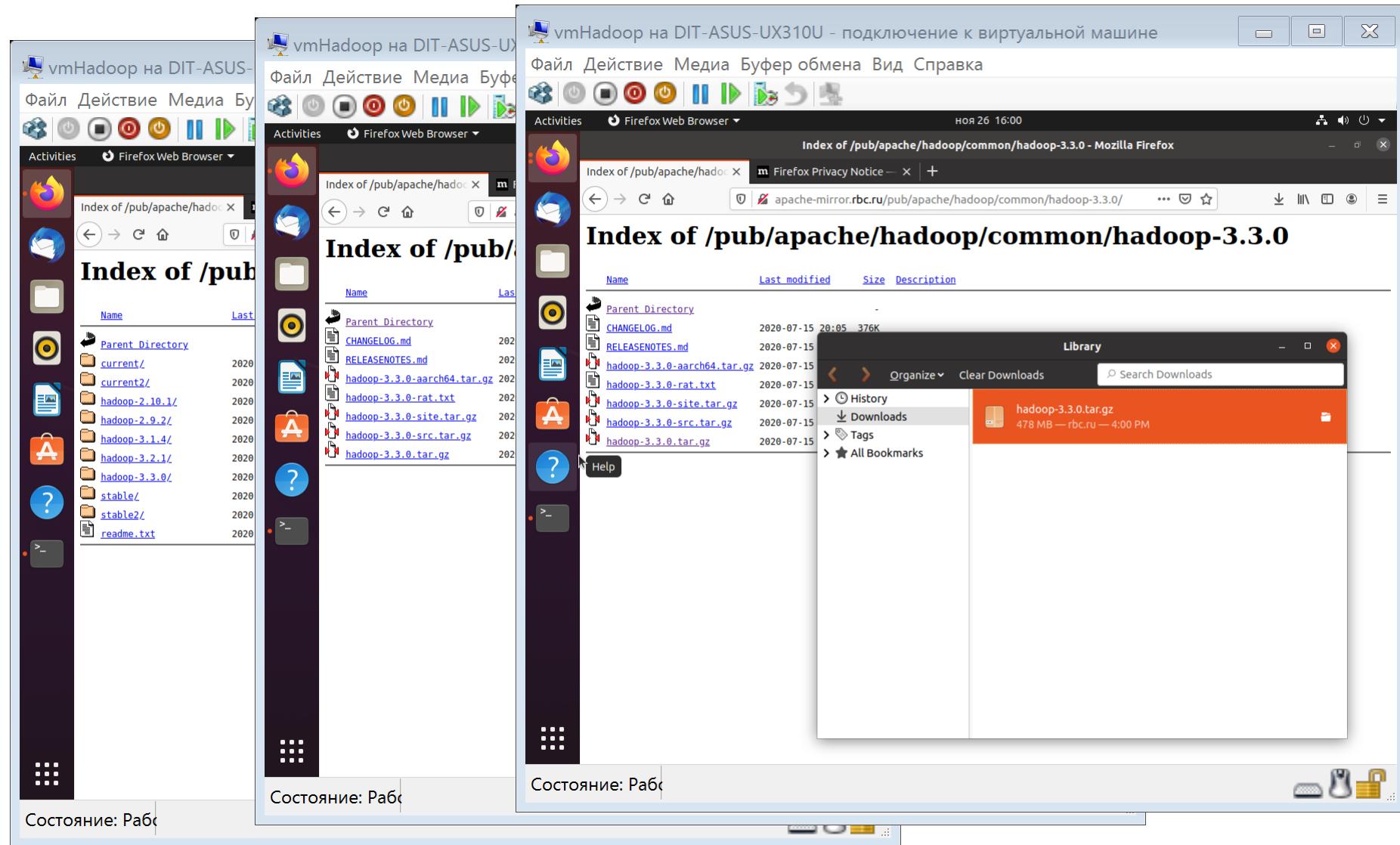
sudo su hduser

ssh-keygen -t rsa -P ""

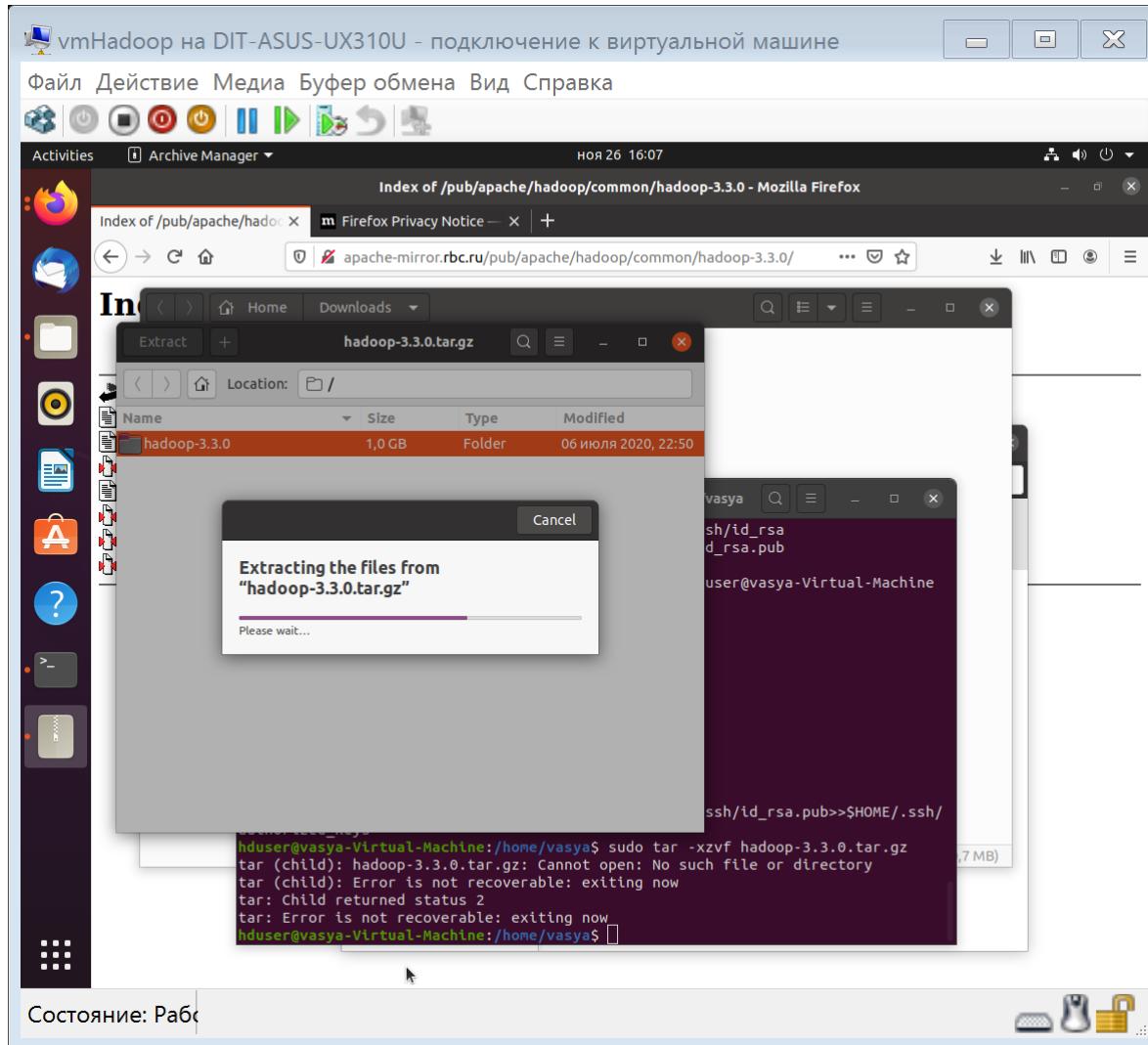
cat \$HOME/.ssh/id\_rsa.pub >> \$HOME/.ssh/authorized\_keys

# Установка Hadoop

<http://apache-mirror.rbc.ru/pub/apache/hadoop/common/>



# Распаковка архива



# Создание рабочей среды

Создание рабочей папки

```
sudo mkdir -p /usr/local/Hadoop
```

Перенос hadoop

```
sudo mv hadoop-3.3.0 /usr/local/hadoop/hadoop- 3.3.0
```

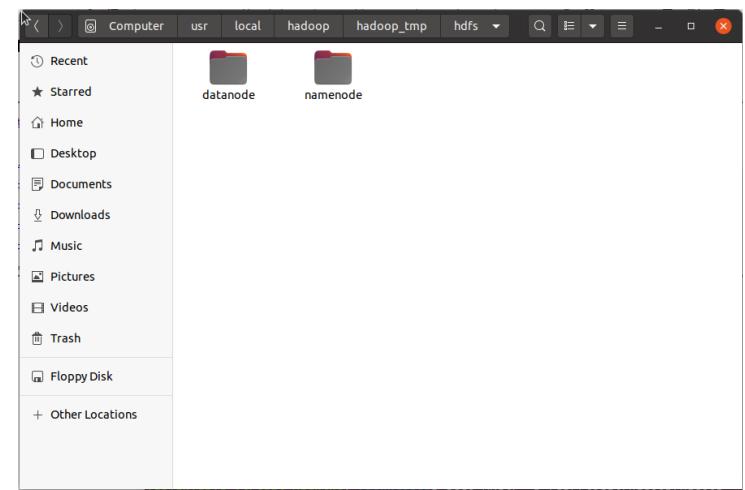
Создание рабочих каталогов для HDFS (NameNode и DataNode) и назначение владельца каталога hadoop (пользователь hduser):

```
sudo mkdir -p /usr/local/hadoop/hadoop_tmp/hdfs/namenode
```

```
sudo mkdir -p /usr/local/hadoop/hadoop_tmp/hdfs/datanode
```

```
sudo chown hduser:hadoop -R /usr/local/hadoop/
```

```
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mv hadoop-3.3.0 /usr/local/hadoop/hadoop-3.3.0
mv: cannot move 'hadoop-3.3.0' to '/usr/local/hadoop/hadoop-3.3.0': No such file or directory
vasya@vasya-Virtual-Machine:~/Downloads$ su
Password:
su: Authentication failure
vasya@vasya-Virtual-Machine:~/Downloads$ vasya
vasya: command not found
vasya@vasya-Virtual-Machine:~/Downloads$ sudo su vasya
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mv hadoop-3.3.0 /usr/local/hadoop/hadoop-3.3.0
mv: cannot move 'hadoop-3.3.0' to '/usr/local/hadoop/hadoop-3.3.0': No such file or directory
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mkdir -p /usr/local/hadoop/hadoop-3.3.0
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mv hadoop-3.3.0 /usr/local/hadoop/hadoop-3.3.0
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mkdir -p /usr/local/hadoop/hadoop_tmp/hdfs/namenode
vasya@vasya-Virtual-Machine:~/Downloads$ sudo mkdir -p /usr/local/hadoop/hadoop_tmp/hdfs/datanode
vasya@vasya-Virtual-Machine:~/Downloads$ sudo chown hduser:hadoop -R /usr/local/hadoop/
vasya@vasya-Virtual-Machine:~/Downloads$
```



# Настройка переменные окружения

Открытие конфигурационного файла

cd

sudo gedit .bashrc

В текстовом редакторе в конец файла вставить следующие строки:

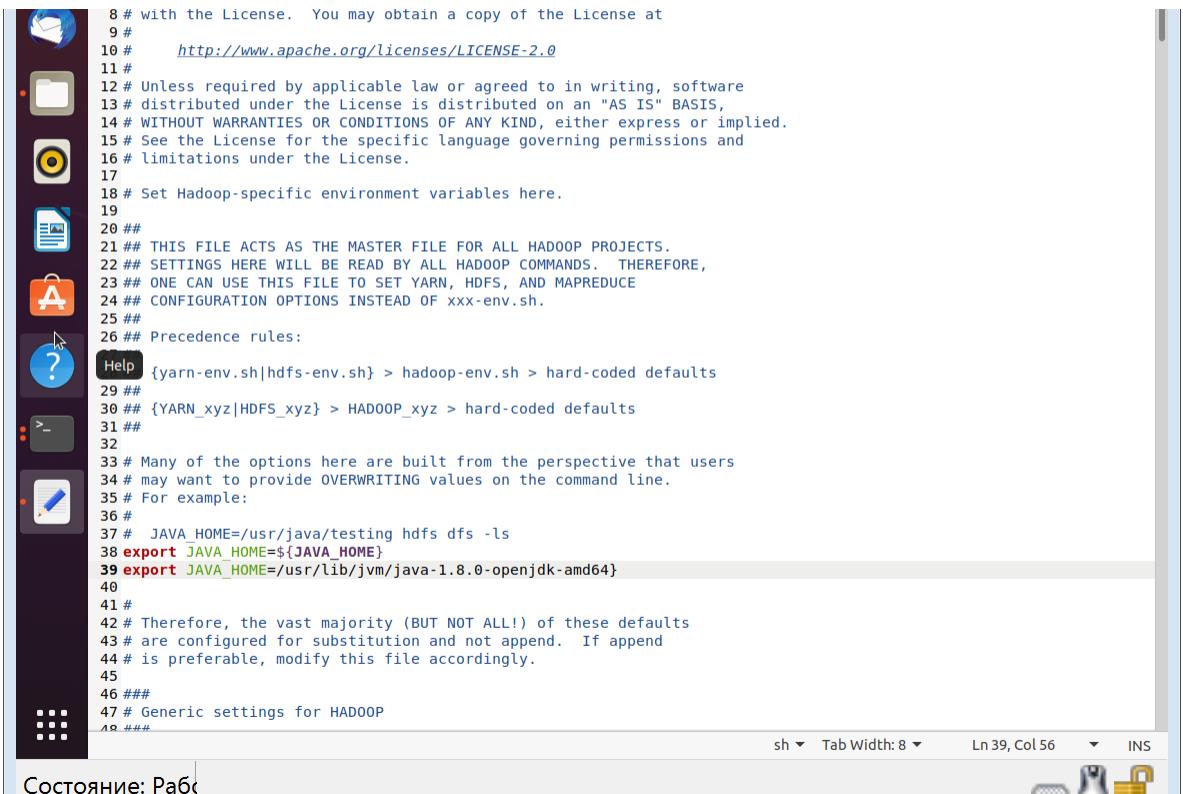
```
export JAVA_HOME=/usr/lib/jvm/java-1.13.0-openjdk-amd64
export HADOOP_HOME=/usr/local/hadoop/hadoop-3.3.0
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"
```

# Настройка Hadoop

Добавить переменную JAVA\_HOME в файл /usr/local/hadoop/hadoop-3.3.0/etc/hadoop/hadoop-env.sh

```
sudo gedit /usr/local/hadoop/hadoop-3.3.0 /etc/hadoop/hadoop-env.sh
```

```
export JAVA_HOME=${JAVA_HOME}  
export JAVA_HOME=/usr/lib/jvn/java-1.13.0-openjdk-amd64
```



```
8 # with the License. You may obtain a copy of the License at  
9 #  
10 # http://www.apache.org/licenses/LICENSE-2.0  
11 #  
12 # Unless required by applicable law or agreed to in writing, software  
13 # distributed under the License is distributed on an "AS IS" BASIS,  
14 # WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.  
15 # See the License for the specific language governing permissions and  
16 # limitations under the License.  
17  
18 # Set Hadoop-specific environment variables here.  
19  
20 ##  
21 ## THIS FILE ACTS AS THE MASTER FILE FOR ALL HADOOP PROJECTS.  
22 ## SETTINGS HERE WILL BE READ BY ALL HADOOP COMMANDS. THEREFORE,  
23 ## ONE CAN USE THIS FILE TO SET YARN, HDFS, AND MAPREDUCE  
24 ## CONFIGURATION OPTIONS INSTEAD OF xxx-env.sh.  
25 ##  
26 ## Precedence rules:  
27 # {yarn-env.sh|hdfs-env.sh} > hadoop-env.sh > hard-coded defaults  
28 ##  
29 ## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defaults  
30 ##  
31 ##  
32  
33 # Many of the options here are built from the perspective that users  
34 # may want to provide OVERWRITING values on the command line.  
35 # For example:  
36 #  
37 #   JAVA_HOME=/usr/java/testing hdfs dfs -ls  
38 export JAVA_HOME=${JAVA_HOME}  
39 export JAVA_HOME=/usr/lib/jvm/java-1.8.0-openjdk-amd64  
40  
41 #  
42 # Therefore, the vast majority (BUT NOT ALL!) of these defaults  
43 # are configured for substitution and not append. If append  
44 # is preferable, modify this file accordingly.  
45  
46 ##  
47 # Generic settings for HADOOP  
48 ##
```

Состояние: Рабочий

sh ▾ Tab Width: 8 ▾ Ln 39, Col 56 ▾ INS

...

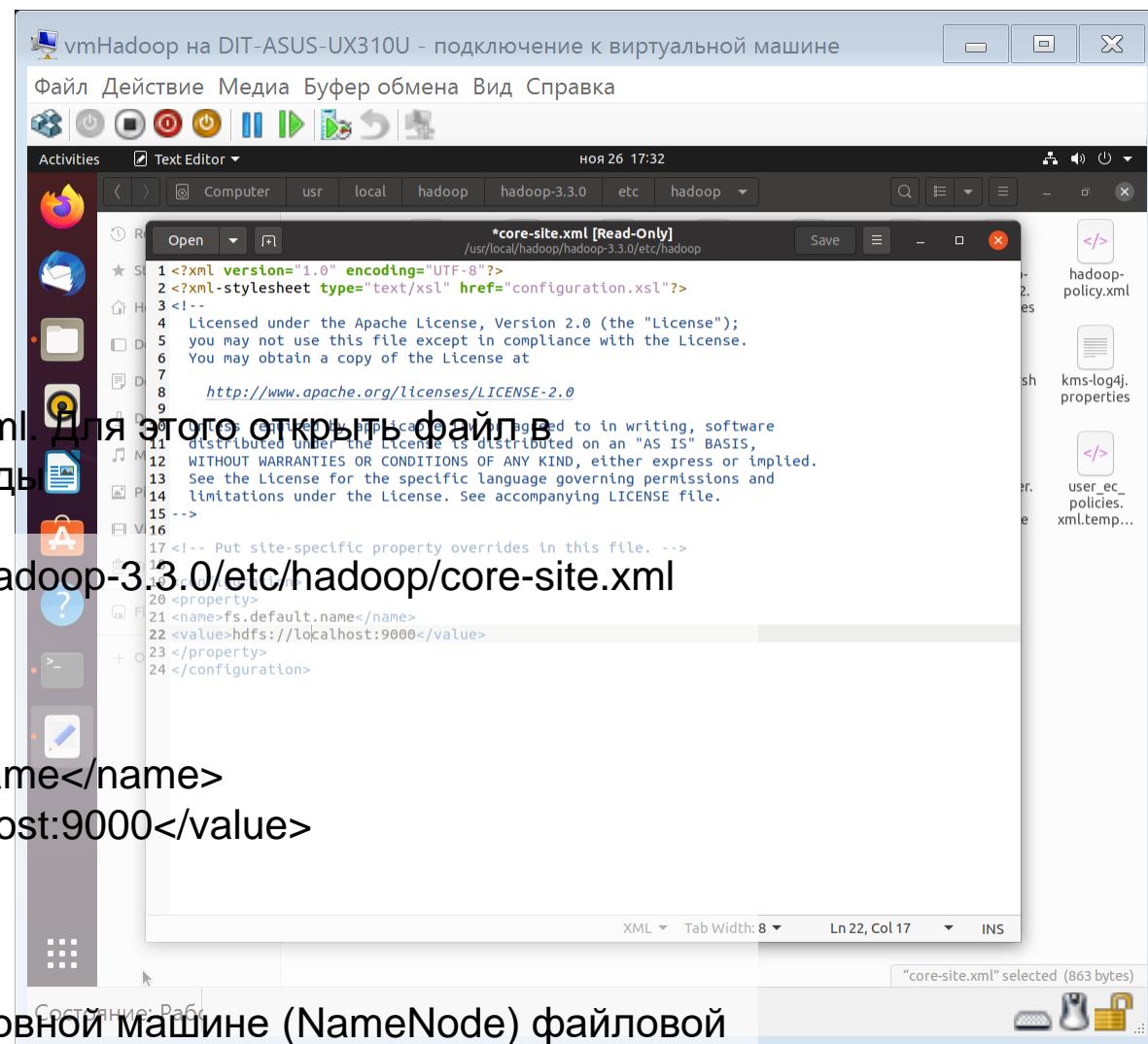
# Настройка Hadoop

Настройка файла core-site.xml. Для этого открыть файл в редакторе с помощью команды

```
sudo gedit /usr/local/hadoop/hadoop-3.3.0/etc/hadoop/core-site.xml
```

```
<configuration>
<property>
    <name>fs.default.name</name>
    <value>hdfs://localhost:9000</value>
</property>
</configuration>
```

Здесь мы указали путь к основной машине (NameNode) файловой системы HDFS, где localhost только для локальной установки HADOOP (в распределенной версии тут либо IP-адрес, либо DNS-имя); 9000 - порт, может быть иным, но лучше не менять.

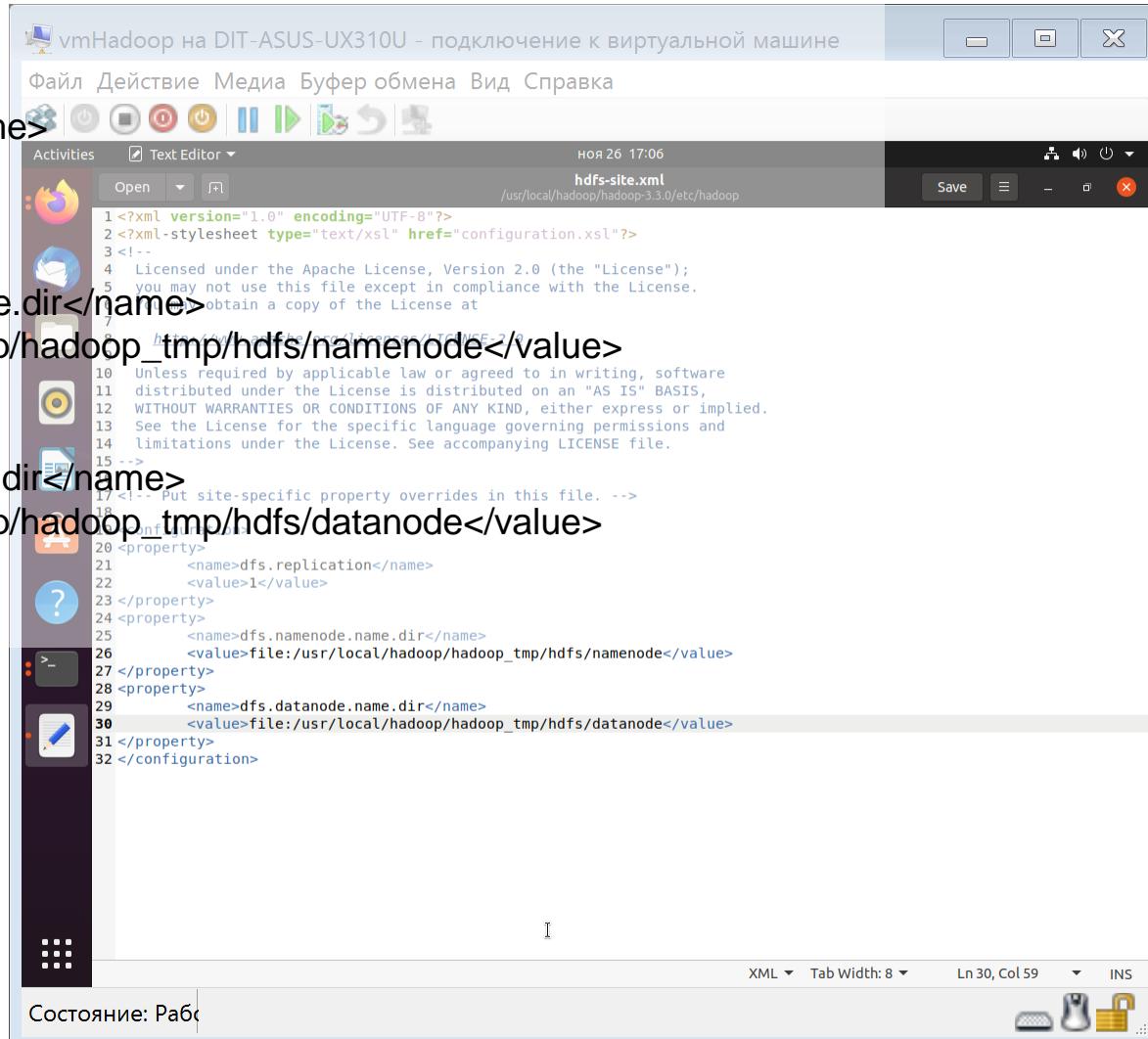


# Настройка Hadoop

Настройка yarn-site.xml. Для этого открыть файл в редакторе с помощью команды

```
sudo gedit /usr/local/hadoop/hadoop-3.3.0/etc/hadoop/yarn-site.xml
```

```
<configuration>
<property>
    <name>dfs.replication</name>
    <value>1</value>
</property>
<property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/usr/local/hadoop/hadoop_tmp/hdfs/namenode</value>
</property>
<property>
    <name>dfs.datanode.name.dir</name>
    <value>file:/usr/local/hadoop/hadoop_tmp/hdfs/datanode</value>
</property>
</configuration>
```



The screenshot shows the Gedit text editor window titled "hdfs-site.xml". The file path is "/usr/local/hadoop/hadoop-3.3.0/etc/hadoop". The code in the editor is as follows:

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
3 <!--
4 Licensed under the Apache License, Version 2.0 (the "License");
5 you may not use this file except in compliance with the License.
6 See the License for the specific language governing permissions and
7 limitations under the License. See accompanying LICENSE file.
8 -->
9 <!-- Put site-specific property overrides in this file. -->
10 <property>
11     <name>dfs.replication</name>
12     <value>1</value>
13 </property>
14 <property>
15     <name>dfs.namenode.name.dir</name>
16     <value>file:/usr/local/hadoop/hadoop_tmp/hdfs/namenode</value>
17 </property>
18 <property>
19     <name>dfs.datanode.name.dir</name>
20     <value>file:/usr/local/hadoop/hadoop_tmp/hdfs/datanode</value>
21 </property>
22 </configuration>
```

The status bar at the bottom right indicates "XML Tab Width: 8 Ln 30, Col 59 INS".

# Настройка Hadoop

Открыть и отредактировать файл

```
sudo gedit /usr/local/hadoop/hadoop-3.3.0/etc/hadoop/mapred-site.xml
```

```
<configuration>
<property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
</property>
</configuration>
```

```
ной машине
```

```
адооп
```

```
re
```

```
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<!-- Put site-specific property overrides in this file. --&gt;
&lt;configuration&gt;
&lt;property&gt;
    &lt;name&gt;mapreduce.framework.name&lt;/name&gt;
    &lt;value&gt;yarn&lt;/value&gt;
&lt;/property&gt;
&lt;/configuration&gt;</pre>

Состояние: Рабочий



XML Tab Width: 8 Ln 22, Col 20 INS

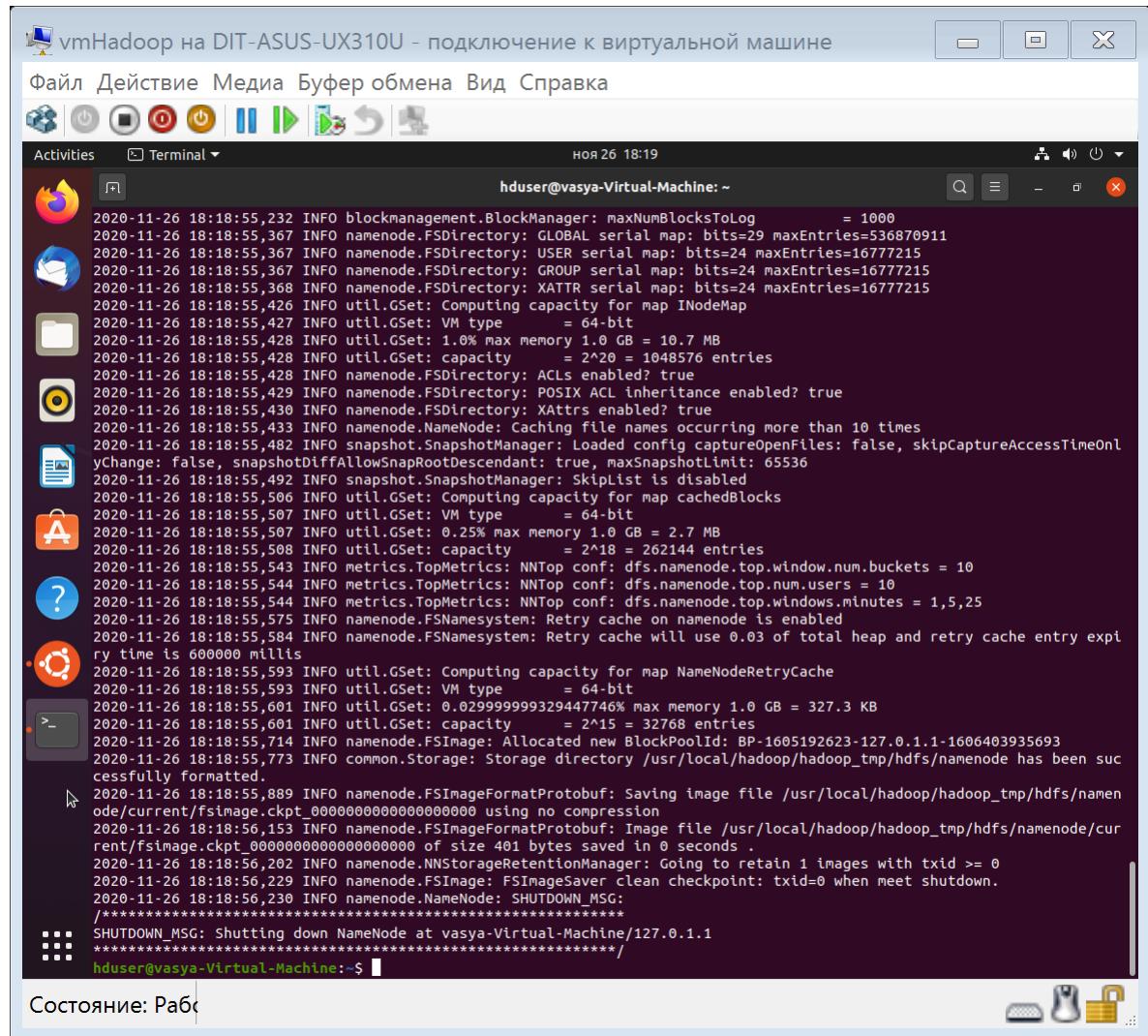

```

# Форматирование HDFS

sudo reboot

Зайти от имени hduser

Форматировать hdfs:  
hdfs namenode -format



```
vmHadoop на DIT-ASUS-UX310U - подключение к виртуальной машине
Файл Действие Медиа Буфер обмена Вид Справка
Activities Terminal ▾
НОЯ 26 18:19
hduser@vasya-Virtual-Machine:~$ hdfs namenode -format
2020-11-26 18:18:55,232 INFO blockmanagement.BlockManager: maxNumBlocksToLog = 1000
2020-11-26 18:18:55,367 INFO namenode.FSDirectory: GLOBAL serial map: bits=29 maxEntries=536870911
2020-11-26 18:18:55,367 INFO namenode.FSDirectory: USER serial map: bits=24 maxEntries=16777215
2020-11-26 18:18:55,367 INFO namenode.FSDirectory: GROUP serial map: bits=24 maxEntries=16777215
2020-11-26 18:18:55,368 INFO namenode.FSDirectory: XATTR serial map: bits=24 maxEntries=16777215
2020-11-26 18:18:55,426 INFO util.GSet: Computing capacity for map INodeMap
2020-11-26 18:18:55,427 INFO util.GSet: VM type = 64-bit
2020-11-26 18:18:55,428 INFO util.GSet: 1.0% max memory 1.0 GB = 10.7 MB
2020-11-26 18:18:55,428 INFO util.GSet: capacity = 2^20 = 1048576 entries
2020-11-26 18:18:55,428 INFO namenode.FSDirectory: ACLs enabled? true
2020-11-26 18:18:55,429 INFO namenode.FSDirectory: POSIX ACL inheritance enabled? true
2020-11-26 18:18:55,430 INFO namenode.FSDirectory: XAttrs enabled? true
2020-11-26 18:18:55,433 INFO namenode.NameNode: Caching file names occurring more than 10 times
2020-11-26 18:18:55,482 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: false, skipCaptureAccessTimeOnLyChange: false, snapshotDiffAllowSnapRootDescendant: true, maxSnapshotLimit: 65536
2020-11-26 18:18:55,492 INFO snapshot.SnapshotManager: Skiplist is disabled
2020-11-26 18:18:55,506 INFO util.GSet: Computing capacity for map cachedBlocks
2020-11-26 18:18:55,507 INFO util.GSet: VM type = 64-bit
2020-11-26 18:18:55,507 INFO util.GSet: 0.25% max memory 1.0 GB = 2.7 MB
2020-11-26 18:18:55,508 INFO util.GSet: capacity = 2^18 = 262144 entries
2020-11-26 18:18:55,543 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
2020-11-26 18:18:55,544 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
2020-11-26 18:18:55,544 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
2020-11-26 18:18:55,575 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
2020-11-26 18:18:55,584 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 000000 millis
2020-11-26 18:18:55,593 INFO util.GSet: Computing capacity for map NameNodeRetryCache
2020-11-26 18:18:55,593 INFO util.GSet: VM type = 64-bit
2020-11-26 18:18:55,601 INFO util.GSet: 0.029999999329447740% max memory 1.0 GB = 327.3 KB
2020-11-26 18:18:55,601 INFO util.GSet: capacity = 2^15 = 32768 entries
2020-11-26 18:18:55,714 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1605192623-127.0.1.1-1606403935693
2020-11-26 18:18:55,773 INFO common.Storage: Storage directory /usr/local/hadoop/hadoop_tmp/hdfs/namenode has been successfully formatted.
2020-11-26 18:18:55,889 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/local/hadoop/hadoop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
2020-11-26 18:18:56,153 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop/hadoop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 401 bytes saved in 0 seconds .
2020-11-26 18:18:56,202 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2020-11-26 18:18:56,229 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2020-11-26 18:18:56,230 INFO namenode.NameNode: SHUTDOWN_MSG:
*****SHUTDOWN_MSG: Shutting down NameNode at vasya-Virtual-Machine/127.0.1.1*****
hduser@vasya-Virtual-Machine:~$
```

Состояние: Рабочий стол

# Запуск HADOOP

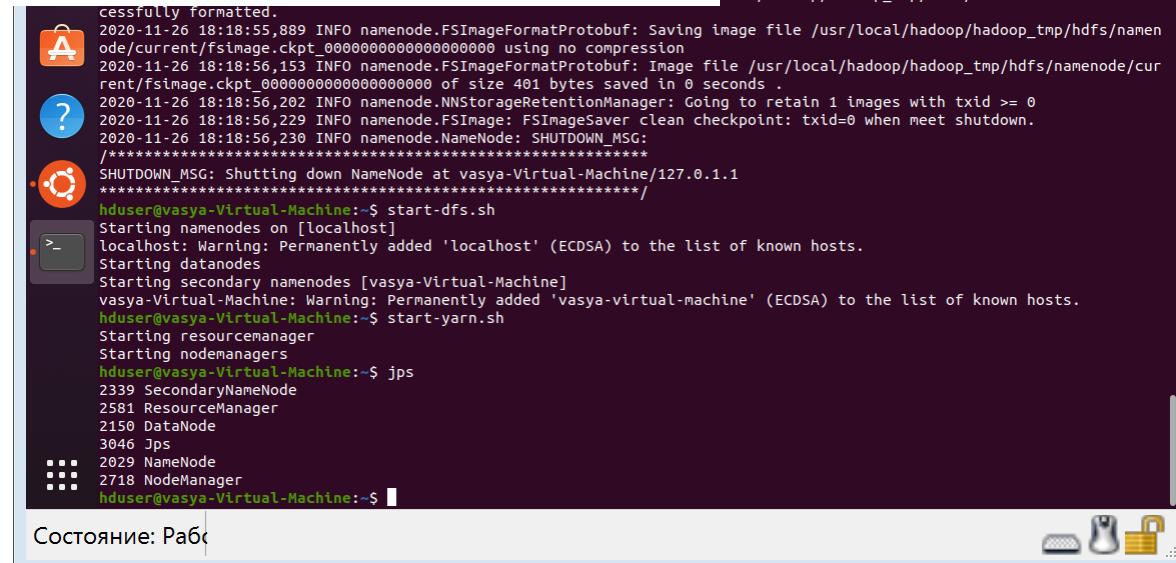
Запуск HADOOP заключается в запуске менеджеров HDFS и YARN. Для этого необходимо запустить два скрипта.

Для HDFS: start-dfs.sh

Для YARN: start-yarn.sh

Успешность запуска всех демонов HADOOP можно проверить с помощью команды jps.

Если все настроено правильно, должно быть запущено 5 процессов HADOOP: NodeManager, NameNode, DataNode, ResourceManager, SecondaryNameNode.



```
2020-11-26 18:18:55,889 INFO namenode.FSImageFormatProtobuf: Saving image file /usr/local/hadoop/hadoop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
2020-11-26 18:18:56,153 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop/hadoop_tmp/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size 401 bytes saved in 0 seconds .
2020-11-26 18:18:56,202 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2020-11-26 18:18:56,229 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2020-11-26 18:18:56,230 INFO namenode.NameNode: SHUTDOWN_MSG:
*****
SHUTDOWN_MSG: Shutting down NameNode at vasya-Virtual-Machine/127.0.1.1
*****
hduser@vasya-Virtual-Machine: ~$ start-dfs.sh
Starting namenodes on [localhost]
localhost: Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Starting datanodes
Starting secondary namenodes [vasya-Virtual-Machine]
vasya-Virtual-Machine: Warning: Permanently added 'vasya-virtual-machine' (ECDSA) to the list of known hosts.
hduser@vasya-Virtual-Machine: ~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hduser@vasya-Virtual-Machine: ~$ jps
2339 SecondaryNameNode
2581 ResourceManager
2150 DataNode
3046 Jps
2029 NameNode
2718 NodeManager
hduser@vasya-Virtual-Machine: ~$
```

## Тюнинг запуска

Для облегчения процедуры запуска HADOOP можно создать простой скрипт, для этого перейдем в домашний каталог пользователя hduser с помощью команды cd.

Создадим файл hdp.sh и отредактируем его с помощью команды:  
sudo gedit hdp.sh

Добавим в редакторе следующие строки:

```
/usr/local/hadoop/hadoop-3.3.0/sbin/start-dfs.sh  
/usr/local/hadoop/hadoop-3.3.0/sbin/start-yarn.sh
```

Сохраняем и выходим.

Теперь нужно дать права на исполнения для файла hdp.sh. Это делается командой:  
sudo chmod 777 hdp.sh

Запускаем скрипт:  
.hdp.sh

**Спасибо за  
внимание!**

