

Департамент образования и науки города Москвы

**Государственное автономное образовательное учреждение высшего
образования города Москвы**

«Московский городской педагогический университет»

Институт цифрового образования

Департамент информатики, управления и технологий

Башкатова Анна Денисовна

ОТЧЕТ

по дисциплине «Инструменты для хранения и обработки больших данных»

Тема: «5.1. Установка и настройка Hadoop»

Направление подготовки (специальность) 38.03.05 – бизнес-информатика

Направленность (профиль) образовательной программы «Аналитика данных
и эффективное управление»

Курс обучения: 3

Форма обучения: очная

Руководитель: Босенко Т. М.

Москва
2023

Установка Hadoop.

```
bashkatova@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

bashkatova@bashkatova-VirtualBox:~$ sudo apt-get update
[sudo] пароль для bashkatova:
Сущ:1 http://ru.archive.ubuntu.com/ubuntu bionic InRelease
Сущ:2 http://ru.archive.ubuntu.com/ubuntu bionic-updates InRelease
Пол:3 http://security.ubuntu.com/ubuntu bionic-security InRelease [88,7 kB]
Сущ:4 http://ru.archive.ubuntu.com/ubuntu bionic-backports InRelease
Пол:5 http://security.ubuntu.com/ubuntu bionic-security/main amd64 DEP-11 Meta
ta [77,1 kB]
Пол:6 http://security.ubuntu.com/ubuntu bionic-security/universe amd64 DEP-11 Me
tadata [62,6 kB]
Пол:7 http://security.ubuntu.com/ubuntu bionic-security/multiverse amd64 DEP-11
Metadata [2 464 B]
Получено 231 kB за 1с (185 kB/s)
Чтение списков пакетов... Готово
bashkatova@bashkatova-VirtualBox:~$
```

```
bashkatova@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
Metadata [2 464 B]
Получено 231 kB за 1с (185 kB/s)
Чтение списков пакетов... Готово
bashkatova@bashkatova-VirtualBox:~$ sudo apt-get install openjdk-8-jdk
Чтение списков пакетов... Готово
Построение дерева зависимостей
Чтение информации о состоянии... Готово
Будут установлены следующие дополнительные пакеты:
  ca-certificates-java fonts-dejavu-extra java-common libatk-wrapper-java
  libatk-wrapper-java-jni libgif7 libice-dev libice6 libpthread-stubs0-dev
  libsm-dev libx11-dev libx11-doc libxau-dev libxcb1-dev libxdmcp-dev
  libxt-dev openjdk-8-jdk-headless openjdk-8-jre openjdk-8-jre-headless
  x11proto-core-dev x11proto-dev xorg-sgml-doctools xtrans-dev
Предлагаемые пакеты:
  default-jre libice-doc libsm-doc libxcb-doc libxt-doc openjdk-8-demo
  openjdk-8-source visualvm fonts-ipafont-gothic fonts-ipafont-mincho
  fonts-wqy-microhei fonts-wqy-zenhei
Следующие НОВЫЕ пакеты будут установлены:
  ca-certificates-java fonts-dejavu-extra java-common libatk-wrapper-java
  libatk-wrapper-java-jni libgif7 libice-dev libpthread-stubs0-dev libsm-dev
  libx11-dev libx11-doc libxau-dev libxcb1-dev libxdmcp-dev libxt-dev
  openjdk-8-jdk openjdk-8-jdk-headless openjdk-8-jre openjdk-8-jre-headless
  x11proto-core-dev x11proto-dev xorg-sgml-doctools xtrans-dev
Следующие пакеты будут обновлены:
  libice6
Обновлено 1 пакетов, установлено 23 новых пакетов, для удаления отмечено 0 пакет
ов, и 292 пакетов не обновлено.
Необходимо скачать 46,3 MB/46,4 MB архивов.
После данной операции объём занятого дискового пространства возрастёт на 173 MB.
Хотите продолжить? [Д/н] у
Пол:1 http://ru.archive.ubuntu.com/ubuntu bionic-updates/main amd64 java-common
all 0.68ubuntu1-18.04.1 [14,5 kB]
Пол:2 http://ru.archive.ubuntu.com/ubuntu bionic-updates/universe amd64 openjdk-
8-jre-headless amd64 8u362-ga-0ubuntu1-18.04.1 [28,3 MB]
30% [2 openjdk-8-jre-headless 16,4 MB/28,3 MB 58%] 1 939 kB/s 15c
```

```
bashkatova@bashkatova-VirtualBox:~$ java -version
openjdk version "1.8.0_362"
OpenJDK Runtime Environment (build 1.8.0_362-8u362-ga-0ubuntu1~18.04.1-b09)
OpenJDK 64-Bit Server VM (build 25.362-b09, mixed mode)
bashkatova@bashkatova-VirtualBox:~$
```

```
OpenJDK 64-Bit Server VM (build 25.362-b09, mixed mode)
bashkatova@bashkatova-VirtualBox:~$ sudo nano /etc/environment
bashkatova@bashkatova-VirtualBox:~$
```

```
bashkatova@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
GNU nano 2.9.3 /etc/environment Изменён
PATH="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games"
JAVA_HOME="/usr/lib/jvm/java-8-openjdk-amd64"
JRE_HOME="/usr/lib/jvm/java-8-openjdk-amd64/jre"

^G Помощь      ^O Записать    ^W Поиск      ^K Вырезать   ^J Выводить
^X Выход      ^R ЧитФайл    ^\ Замена     ^U Отмен. вырез ^T Словарь
```

```
bashkatova@bashkatova-VirtualBox:~$ sudo nano /etc/environment
bashkatova@bashkatova-VirtualBox:~$ sudo adduser hadoop
Добавляется пользователь «hadoop» ...
Добавляется новая группа «hadoop» (1001) ...
Добавляется новый пользователь «hadoop» (1001) в группу «hadoop» ...
Создаётся домашний каталог «/home/hadoop» ...
Копирование файлов из «/etc/skel» ...
Введите новый пароль UNIX:
Повторите ввод нового пароля UNIX:
passwd: пароль успешно обновлён
Изменение информации о пользователе hadoop
Введите новое значение или нажмите ENTER для выбора значения по умолчанию
    Полное имя []:
    Номер комнаты []:
    Рабочий телефон []:
    Домашний телефон []:
    Другое []:
Данная информация корректна? [Y/n]
bashkatova@bashkatova-VirtualBox:~$ sudo su hadoop
hadoop@bashkatova-VirtualBox:/home/bashkatova$
```

```
bashkatova@bashkatova-VirtualBox:~$ sudo su hadoop
hadoop@bashkatova-VirtualBox:/home/bashkatova$ exit
exit
bashkatova@bashkatova-VirtualBox:~$
```

```
bashkatova@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
exit
bashkatova@bashkatova-VirtualBox:~$ sudo apt-get install ssh pdsh
Чтение списков пакетов... Готово
Построение дерева зависимостей
Чтение информации о состоянии... Готово
Будут установлены следующие дополнительные пакеты:
  genders libgenders0 ncurses-term openssh-client openssh-server
  openssh-sftp-server ssh-import-id
Предлагаемые пакеты:
  rdist keychain libram-ssh monkeysphere ssh-askpass molly-guard rssh
Следующие НОВЫЕ пакеты будут установлены:
  genders libgenders0 ncurses-term openssh-server openssh-sftp-server pdsh
  ssh ssh-import-id
Следующие пакеты будут обновлены:
  openssh-client
Обновлено 1 пакетов, установлено 8 новых пакетов, для удаления отмечено 0 пак
етов, и 291 пакетов не обновлено.
Необходимо скачать 812 kB/1 422 kB архивов.
После данной операции объём занятого дискового пространства возрастёт на 5 90
6 kB.
Хотите продолжить? [д/н] у
Пол:1 http://ru.archive.ubuntu.com/ubuntu bionic-updates/main amd64 openssh-s
ftp-server amd64 1:7.6p1-4ubuntu0.7 [45,5 kB]
Пол:2 http://ru.archive.ubuntu.com/ubuntu bionic-updates/main amd64 openssh-s
erver amd64 1:7.6p1-4ubuntu0.7 [332 kB]
Пол:3 http://ru.archive.ubuntu.com/ubuntu bionic-updates/main amd64 ssh all 1
:7.6p1-4ubuntu0.7 [5 192 B]
Пол:4 http://ru.archive.ubuntu.com/ubuntu bionic/universe amd64 libgenders0 a
md64 1.21-1build5 [31,3 kB]
Пол:5 http://ru.archive.ubuntu.com/ubuntu bionic/universe amd64 genders amd64
1.21-1build5 [30,9 kB]
Пол:6 http://ru.archive.ubuntu.com/ubuntu bionic-updates/main amd64 ncurses-t
erm all 6.1-1ubuntu1.18.04 [248 kB]
Пол:7 http://ru.archive.ubuntu.com/ubuntu bionic/universe amd64 pdsh amd64 2.
31-3build2 [108 kB]
```

```
bashkatova@bashkatova-VirtualBox:~$ sudo su hadoop
hadoop@bashkatova-VirtualBox:/home/bashkatova$ cd
hadoop@bashkatova-VirtualBox:~$ ssh-keygen -t rsa -N "" -f /home/hadoop/.ssh/
id_rsa
Generating public/private rsa key pair.
Created directory '/home/hadoop/.ssh'.
Your identification has been saved in /home/hadoop/.ssh/id_rsa.
Your public key has been saved in /home/hadoop/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:j0VfHbBJ74fyC8Z2MwjAdI169M24zvRK5iy40/vWerQ hadoop@bashkatova-VirtualB
ox
The key's randomart image is:
+---[RSA 2048]---+
|      . .oo      |
|      o .o..=    |
|      +o .o+o    |
|      =... oooo   |
|      . S.. o.+ . |
|      . +++ .    |
|      ..*Oo=     |
|      o oBEo.+   |
|      +Bo+...    |
+---[SHA256]-----+
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ cat /home/hadoop/.ssh/id_rsa.pub >> /home/had
oop/.ssh/authorized_keys
hadoop@bashkatova-VirtualBox:~$ chmod 0600 /home/hadoop/.ssh/authorized_keys
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:jpvovkk842dcDhTqjk2Hrx0P0PgctoctPEf8B88vUp4.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 18.04.6 LTS (GNU/Linux 5.4.0-84-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

297 updates can be applied immediately.
267 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable

Your Hardware Enablement Stack (HWE) is supported until April 2023.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ wget https://archive.apache.org/dist/hadoop/c
ommon/hadoop-3.1.2/hadoop-3.1.2.tar.gz
--2023-04-22 15:31:39-- https://archive.apache.org/dist/hadoop/common/hadoop
-3.1.2/hadoop-3.1.2.tar.gz
Распознаётся archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4
f8:172:2ec5::2
Подключение к archive.apache.org (archive.apache.org)|138.201.131.134|:443...
соединение установлено.
HTTP-запрос отправлен. Ожидание ответа... 200 OK
Длина: 332433589 (317M) [application/x-gzip]
Сохранение в: «hadoop-3.1.2.tar.gz»

hadoop-3.1.2.tar.gz 9%[> ] 29,62M 1,77MB/s oct 1m 50s
```



```
hadoop@bashkatova-VirtualBox:~$ wget https://archive.apache.org/dist/hadoop/c
ommon/hadoop-3.1.2/hadoop-3.1.2.tar.gz
--2023-04-22 15:31:39-- https://archive.apache.org/dist/hadoop/common/hadoop
-3.1.2/hadoop-3.1.2.tar.gz
Распознаётся archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4
f8:172:2ec5::2
Подключение к archive.apache.org (archive.apache.org)|138.201.131.134|:443...
соединение установлено.
HTTP-запрос отправлен. Ожидание ответа... 200 OK
Длина: 332433589 (317М) [application/x-gzip]
Сохранение в: «hadoop-3.1.2.tar.gz»

hadoop-3.1.2.tar.gz 100%[=====] 317,03М 2,20МБ/с за 2м 37с

2023-04-22 15:34:16 (2,02 MB/s) - «hadoop-3.1.2.tar.gz» сохранён [332433589/3
32433589]

hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ mv hadoop-3.1.2 hadoop
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
GNU nano 2.9.3 .bashrc Изменён

fi
# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
  if [ -f /usr/share/bash-completion/bash_completion ]; then
    . /usr/share/bash-completion/bash_completion
  elif [ -f /etc/bash_completion ]; then
    . /etc/bash_completion
  fi
fi
export HADOOP_HOME=/home/hadoop/hadoop
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 PDSH_RCMD_TYPE=ssh

^G Помощь ^O Записать ^W Поиск ^K Вырезать ^J Вывернуть
^X Выход ^R ЧитФайл ^_ Замена ^U Отмен. вырез ^T Словарь
```

```
hadoop@bashkatova-VirtualBox:~$ nano .bashrc
hadoop@bashkatova-VirtualBox:~$ source .bashrc
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
GNU nano 2.9.3 /home/hadoop/hadoop/etc/hadoop/hadoop-env.sh Изменён

# export HDFS_DFSROUTER_OPTS=""
###

###
# Advanced Users Only!
###

#
# When building Hadoop, one can add the class paths to the commands
# via this special env var:
# export HADOOP_ENABLE_BUILD_PATHS="true"
#
# To prevent accidents, shell commands be (superficially) locked
# to only allow certain users to execute certain subcommands.
# It uses the format of (command)_(subcommand)_USER.
#
# For example, to limit who can execute the namenode command,
# export HDFS_NAMENODE_USER=hdfs
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64

^G Помощь ^O Записать ^W Поиск ^K Вырезать ^J Выводить
^X Выход ^R ЧитФайл ^\ Замена ^U Отмен. вырез ^T Пров. синтак.
```

```
hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
GNU nano 2.9.3 /home/hadoop/hadoop/etc/hadoop/core-site.xml Изменён

<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  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. -->

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

```
hadoop@bashkatova-VirtualBox: ~  
Файл Правка Вид Поиск Терминал Справка  
GNU nano 2.9.3 /home/hadoop/hadoop/etc/hadoop/hdfs-site.xml Изменён  
  
<!-- Put site-specific property overrides in this file. -->  
  
<configuration>  
<property>  
<name>dfs.replication</name>  
<value>1</value>  
</property>  
<property>  
<name>dfs.name.dir</name>  
<value>file:///home/hadoop/hadoopdata/hdfs/namenode</value>  
</property>  
<property>  
<name>dfs.data.dir</name>  
<value>file:///home/hadoop/hadoopdata/hdfs/datanode</value>  
</property>  
</configuration>
```

```
hadoop@bashkatova-VirtualBox: ~  
Файл Правка Вид Поиск Терминал Справка  
GNU nano 2.9.3 /home/hadoop/hadoop/etc/hadoop/mapred-site.xml Изменён  
  
<!-- Put site-specific property overrides in this file. -->  
  
<configuration>  
<property>  
<name>mapreduce.framework.name</name>  
<value>yarn</value>  
</property>  
<property>  
<name>yarn.app.mapreduce.am.env</name>  
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>  
</property>  
<property>  
<name>mapreduce.map.env</name>  
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>  
</property>  
<property>  
<name>mapreduce.reduce.env</name>  
<value>HADOOP_MAPRED_HOME=${HADOOP_HOME}</value>  
</property>  
</configuration>
```



```
hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
GNU nano 2.9.3 /home/hadoop/hadoop/etc/hadoop/yarn-site.xml Изменён

<?xml version="1.0"?>
<!--
 Licensed under the Apache License, Version 2.0 (the "License");
 you may not use this file except in compliance with the License.
 You may obtain a copy of the License at

 http://www.apache.org/licenses/LICENSE-2.0

 Unless required by applicable law or agreed to in writing, software
 distributed under the License is distributed on an "AS IS" BASIS,
 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.
-->
<configuration>

<!-- Site specific YARN configuration properties -->
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.resource.memory-mb</name>
<value>16384</value>
</property>
</configuration>

^G Помощь ^O Записать ^W Поиск ^K Вырезать ^J Вывернуть
^X Выход ^R ЧитФайл ^\ Замена ^U Отмен. вырез ^T Словарь

hadoop@bashkatova-VirtualBox:~$ source .bashrc
hadoop@bashkatova-VirtualBox:~$ nano /home/hadoop/hadoop/etc/hadoop/hadoop-env.sh
hadoop@bashkatova-VirtualBox:~$ nano /home/hadoop/hadoop/etc/hadoop/core-site.xml
hadoop@bashkatova-VirtualBox:~$ nano /home/hadoop/hadoop/etc/hadoop/hdfs-site.xml
hadoop@bashkatova-VirtualBox:~$ nano /home/hadoop/hadoop/etc/hadoop/mapred-site.xml
hadoop@bashkatova-VirtualBox:~$ nano /home/hadoop/hadoop/etc/hadoop/yarn-site.xml
hadoop@bashkatova-VirtualBox:~$

hadoop@bashkatova-VirtualBox:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bashkatova-VirtualBox]
bashkatova-VirtualBox: Warning: Permanently added 'bashkatova-virtualbox' (ECDSA) to the list of known hosts.
2023-04-22 15:41:22,041 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$

hadoop@bashkatova-VirtualBox:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
```

```
hadoop@bashkatova-VirtualBox:~$ hdfs dfsadmin -report
2023-04-22 15:42:06,396 WARN util.NativeCodeLoader: Unable to load native-hadoop
oop library for your platform... using builtin-java classes where applicable
Configured Capacity: 10499674112 (9.78 GB)
Present Capacity: 2026893312 (1.89 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used: 24576 (24 KB)
DFS Used%: 0.00%
Replicated Blocks:
  Under replicated blocks: 0
  Blocks with corrupt replicas: 0
  Missing blocks: 0
  Missing blocks (with replication factor 1): 0
  Low redundancy blocks with highest priority to recover: 0
  Pending deletion blocks: 0
Erasure Coded Block Groups:
  Low redundancy block groups: 0
  Block groups with corrupt internal blocks: 0
  Missing block groups: 0
  Low redundancy blocks with highest priority to recover: 0
  Pending deletion blocks: 0
-----
Live datanodes (1):
Name: 127.0.0.1:9866 (localhost)
Hostname: bashkatova-VirtualBox
Decommission Status : Normal
Configured Capacity: 10499674112 (9.78 GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 7919239168 (7.38 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used%: 0.00%
DFS Remaining%: 19.30%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Sat Apr 22 15:42:05 MSK 2023
Last Block Report: Sat Apr 22 15:41:17 MSK 2023
Num of Blocks: 0
```

hadoop@bashkatova-VirtualBox: ~

Файл Правка Вид Поиск Терминал Справка

oop library for your platform... using builtin-java class
Configured Capacity: 10499674112 (9.78 GB)
Present Capacity: 2026893312 (1.89 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used: 24576 (24 KB)
DFS Used%: 0.00%
Replicated Blocks:
 Under replicated blocks: 0
 Blocks with corrupt replicas: 0
 Missing blocks: 0
 Missing blocks (with replication factor 1): 0
 Low redundancy blocks with highest priority to recover: 0
 Pending deletion blocks: 0
Erasure Coded Block Groups:
 Low redundancy block groups: 0
 Block groups with corrupt internal blocks: 0
 Missing block groups: 0
 Low redundancy blocks with highest priority to recover: 0
 Pending deletion blocks: 0


Live datanodes (1):
Name: 127.0.0.1:9866 (localhost)
Hostname: bashkatova-VirtualBox
Decommission Status : Normal
Configured Capacity: 10499674112 (9.78 GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 7919239168 (7.38 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used%: 0.00%
DFS Remaining%: 19.30%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Sat Apr 22 15:42:05 MSK 2023
Last Block Report: Sat Apr 22 15:41:17 MSK 2023
Num of Blocks: 0

hadoop@bashkatova-VirtualBox:~\$

Exercise_1.pdf

All Applications

localhost:8088/cluster



Cluster

About

Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers
0	0	0	0	0

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned
1	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Mi
Capacity Scheduler	[memory-mb (unit=M), vcores]	<memory:1

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State

Showing 0 to 0 of 0 entries

hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка

oop library for your platform... using builtin-java class
Configured Capacity: 10499674112 (9.78 GB)
Present Capacity: 2026893312 (1.89 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used: 24576 (24 KB)
DFS Used%: 0.00%
Replicated Blocks:
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0
Low redundancy blocks with highest priority to re
Pending deletion blocks: 0
Erasure Coded Block Groups:
Low redundancy block groups: 0
Block groups with corrupt internal blocks: 0
Missing block groups: 0
Low redundancy blocks with highest priority to re
Pending deletion blocks: 0

Live datanodes (1):

Name: 127.0.0.1:9866 (localhost)
Hostname: bashkatova-VirtualBox
Decommission Status : Normal
Configured Capacity: 10499674112 (9.78 GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 7919239168 (7.38 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used%: 0.00%
DFS Remaining%: 19.30%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xcelfers: 1
Last contact: Sat Apr 22 15:42:05 MSK 2023
Last Block Report: Sat Apr 22 15:41:17 MSK 2023
Num of Blocks: 0

hadoop@bashkatova-VirtualBox:~\$

Exercise_1.pdf × All Applications × Namenode information × +

localhost:9870/dfshealth.html#tab-overview

Hadoop

Overview Datanodes Datanode Volume Failures Snapshot Startup Progress

Utilities ▾

Overview 'localhost:9000' (active)

Started:	Sat Apr 22 15:41:14 +0300 2023
Version:	3.1.2, r1019dde65bcf12e05ef48ac71e84550d589e5d9a
Compiled:	Tue Jan 29 04:39:00 +0300 2019 by sunilg from branch-3.1.2
Cluster ID:	CID-fe53f1c3-28ec-4eb7-8409-1748f8fced4e
Block Pool ID:	BP-998266363-127.0.1.1-1682167246728

Summary

Security is off.

Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).

Heap Memory used 69 MB of 207 MB Heap Memory. Max Heap Memory is 1.02 GB.

Non Heap Memory used 46.9 MB of 48.34 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка

oop library for your platform... using builtin-java class
Configured Capacity: 10499674112 (9.78 GB)
Present Capacity: 2026893312 (1.89 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used: 24576 (24 KB)
DFS Used%: 0.00%
Replicated Blocks:
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0
Missing blocks (with replication factor 1): 0
Low redundancy blocks with highest priority to re
Pending deletion blocks: 0
Erasure Coded Block Groups:
Low redundancy block groups: 0
Block groups with corrupt internal blocks: 0
Missing block groups: 0
Low redundancy blocks with highest priority to re
Pending deletion blocks: 0

Live datanodes (1):

Name: 127.0.0.1:9866 (localhost)
Hostname: bashkatova-VirtualBox
Decommission Status : Normal
Configured Capacity: 10499674112 (9.78 GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 7919239168 (7.38 GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used%: 0.00%
DFS Remaining%: 19.30%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xcelfers: 1
Last contact: Sat Apr 22 15:42:05 MSK 2023
Last Block Report: Sat Apr 22 15:41:17 MSK 2023
Num of Blocks: 0

hadoop@bashkatova-VirtualBox:~\$

Exercise_1.pdf × All Applications × Browsing HDFS × +



localhost:9870/explorer.html#/


Hadoop

Overview Datanodes Datanode Volume Failures Snapshot Startup Progress

Utilities ▾

Browse Directory

/ Go!  



Show 25 entries Search:

☐ Permission ☐ Owner ☐ Group ☐ Size ☐ Last Modified ☐ Replication ☐ Block Size ☐ Name

No data available in table

Showing 0 to 0 of 0 entries

Previous Next

Hadoop, 2018.

Работа в HDFS

```
hadoop@bashkatova-VirtualBox:~$ hadoop fs -mkdir /user
2023-04-22 15:45:56,637 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$ hadoop fs -mkdir /user/hadoop
2023-04-22 15:46:09,509 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$ hadoop fs -ls /
2023-04-22 15:46:29,824 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - hadoop supergroup 0 2023-04-22 15:46 /user
```

```
drwxr-xr-x - hadoop supergroup 0 2023-04-22 15:46 /user
hadoop@bashkatova-VirtualBox:~$ hadoop fs -put /var/log/dpkg.log /user/hadoop/dpkg.log
2023-04-22 15:46:55,042 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$
```

hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
Block groups with corrupt internal blocks: 0
Missing block groups: 0
Low redundancy blocks with highest priority to re-replicate: 0
Pending deletion blocks: 0

Live datanodes (1):
Name: 127.0.0.1:9866 (localhost)
Hostname: bashkatova-VirtualBox
Decommission Status : Normal
Configured Capacity: 10499674112 (9.7B GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 7919239168 (7.3B GB)
DFS Remaining: 2026868736 (1.89 GB)
DFS Used%: 0.00%
DFS Remaining%: 19.30%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceiver: 1
Last contact: Sat Apr 22 15:42:05 MSK 2023
Last Block Report: Sat Apr 22 15:41:17 MSK 2023
Num of Blocks: 0

hadoop@bashkatova-VirtualBox:~\$ hadoop fs -mkdir /user
2023-04-22 15:45:56,637 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -mkdir /user/hadoop
2023-04-22 15:46:09,509 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -ls /
2023-04-22 15:46:29,824 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - hadoop supergroup 0 2023-04-22 15:46 /user
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -put /var/log/dpkg.log /user/hadoop/dpkg.log
2023-04-22 15:46:55,042 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$

Exercise_1.pdf x All Applications x Browsing HDFS x +
localhost:9870/explorer.html/
Hadoop
Overview Datanodes Datanode Volume Failures Snapshot Startup Progress
Utilities +
Browse Directory
/ Go
Show 25 entries Search:
Permission Owner Group Size Last Modified Replication Block Size Name
drwxr-xr-x hadoop supergroup 0 B Apr 22 15:46 0 0 B user
Showing 1 to 1 of 1 entries Previous 1 Next
Hadoop, 2018.

hadoop@bashkatova-VirtualBox: ~
Файл Правка Вид Поиск Терминал Справка
oop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -mkdir /user/hadoop
2023-04-22 15:46:09,509 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -ls /
2023-04-22 15:46:29,824 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - hadoop supergroup 0 2023-04-22 15:46 /user
hadoop@bashkatova-VirtualBox:~\$ hadoop fs -put /var/log/dpkg.log /user/hadoop/dpkg.log
2023-04-22 15:46:55,042 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~\$ hadoop jar hadoop/share/examples/mapreduce-examples-3.1.2.jar wordcount /user/hadoop/dpkg.log
2023-04-22 15:48:08,499 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-04-22 15:48:09,079 INFO client.RMProxy: Connecting to ResourceManager: /0.0.0.0:8032
2023-04-22 15:48:09,516 INFO mapreduce.JobResourceUploader: Uploading conf for path: /tmp/hadoop-yarn/staging/hadoop/.staging-0001
2023-04-22 15:48:09,741 INFO Input.FileInputFormat: Total input files: 1
2023-04-22 15:48:09,865 INFO mapreduce.JobSubmitter: number of maps: 1
2023-04-22 15:48:10,020 INFO mapreduce.JobSubmitter: Submitting job: job_1682167305925_0001
2023-04-22 15:48:10,021 INFO mapreduce.JobSubmitter: Execution attempt: 1
2023-04-22 15:48:10,208 INFO conf.Configuration: resource types: xml
2023-04-22 15:48:10,209 INFO resource.ResourceUtils: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-04-22 15:48:10,657 INFO impl.YarnClientImpl: Submitting job: job_1682167305925_0001
2023-04-22 15:48:10,704 INFO mapreduce.Job: The url to track the job: http://localhost:8080/proxy/application_1682167305925_0001
2023-04-22 15:48:10,705 INFO mapreduce.Job: Running job: job_1682167305925_0001
2023-04-22 15:48:18,880 INFO mapreduce.Job: Job job_1682167305925_0001 in uber mode : false
2023-04-22 15:48:18,882 INFO mapreduce.Job: map 0% reduce 0%
2023-04-22 15:48:24,952 INFO mapreduce.Job: map 100% reduce 100%

Exercise_1.pdf x All Applications x Browsing HDFS x +
localhost:8080/cluster
All Applications
Apps Pending Apps Running Apps Completed Containers Running Memory Used
1 0 2 3 GB
Decommissioning Nodes Decommissioned Nodes Lost Node
Scheduling Resource Type Minimum Allocation
[memory-mb (unit=M), vcores] <memory:1024, vCores:1>
User Name Application Type Queue Application Priority StartTime FinishTime State FinalStatus Run Cont
hadoop word count MAPREDUCE default 0 Sat Apr 22 15:48:10 +0300 2023 RUNNING UNDEFINED 2

ФайлПравка

Exercise_1.pdfAll ApplicationsBrowsing HDFS

localhost:8088/cluster

doop

Map-FileShuffleFileFile

All Applications

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total
1	0	1	0	0 B	16 GB	0 B

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy
1	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCores	All M I
application_1682167305925_0001	hadoop	word count	MAPREDUCE	default	0	Sat Apr 22 15:48:10 +0300 2023	Sat Apr 22 15:48:28 +0300 2023	FINISHED	SUCCEEDED	N/A	N/A	N/A

Showing 1 to 1 of 1 entries

```
hadoop@bashkatova-VirtualBox:~$ hadoop fs -cat /user/hadoop/test_output/part-r-00000
2023-04-22 15:49:17,008 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
0+svn585-1build1      8
0.0.12 8
0.0.20060226-9ubuntu1 8
0.0.4-1 8
0.0.4-6ubuntu2 8
0.0.7+17.10.20170922-0ubuntu1 10
0.0.9-10 8
0.01-5 8
0.032-1 8
0.04-7.1 8
0.044-2 8
0.06-2.1ubuntu1 8
0.06-7.1 8
0.07-1 8
0.0~r131-2ubuntu3 13
0.0~r131-2ubuntu3.1 7
0.1-5 8
0.1-6 8
0.1.0+git20170815-2 16
0.1.0-2ubuntu1 8
0.1.10-2.1 8
0.1.2-1ubuntu1 8
0.1.2-2 8
0.1.26-2 8
0.1.27-2 16
0.1.3-1 8
0.1.32 8
0.1.33build1 8
```


Выполнение упражнений

Клонирование репозитория git

```
hadoop@bashkatova-VirtualBox:~$ git clone https://github.com/BosenkoTM/Big-Data-Storage-and-Processing-Tools.git Bosenko
Клонирование в «Bosenko»...
remote: Enumerating objects: 123, done.
remote: Counting objects: 100% (27/27), done.
remote: Compressing objects: 100% (27/27), done.
remote: Total 123 (delta 12), reused 0 (delta 0), pack-reused 96
Получение объектов: 100% (123/123), 19.72 MiB | 1.77 MiB/s, готово.
Определение изменений: 100% (37/37), готово.
hadoop@bashkatova-VirtualBox:~$ ls
Big-Data-Storage-and-Processing-Tools  hadoopdata  Музыка
Bosenko                               Видео       Общедоступные
examples.desktop                       Документы   'Рабочий стол'
hadoop                                Загрузки    Шаблоны
hadoop-3.1.2.tar.gz                   Изображения

hadoop@bashkatova-VirtualBox:~$ cd Big-Data-Storage-and-Processing-Tools
hadoop@bashkatova-VirtualBox:~/Big-Data-Storage-and-Processing-Tools$ ls
kp  lectures  marketing_analytics_students.ipynb  practice  README.md
hadoop@bashkatova-VirtualBox:~/Big-Data-Storage-and-Processing-Tools$ cd
hadoop@bashkatova-VirtualBox:~$ cd Big-Data-Storage-and-Processing-Tools
hadoop@bashkatova-VirtualBox:~/Big-Data-Storage-and-Processing-Tools$ cd practice
hadoop@bashkatova-VirtualBox:~/Big-Data-Storage-and-Processing-Tools/practice$ ls
ex_01_1  ex_01_2  ex_02  ex_03  Faust_1.txt
hadoop@bashkatova-VirtualBox:~/Big-Data-Storage-and-Processing-Tools/practice$ cd
hadoop@bashkatova-VirtualBox:~$ hadoop fs -put Big-Data-Storage-and-Processing-Tools/practice/Faust_1.txt /user/hadoop/Faust_1.txt
2023-04-27 19:58:06,282 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$
```

```
bytes written=9
hadoop@bashkatova-VirtualBox:~$ hadoop jar hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.1.2.jar wordcount /user/hadoop/Faust_1.txt /user/hadoop/Faust_1_Output
2023-04-27 20:57:52,168 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-04-27 20:57:53,204 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2023-04-27 20:57:53,812 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-ya
rn/staging/hadoop/.staging/job_1682168633199_0003
2023-04-27 20:57:54,221 INFO input.FileInputFormat: Total input files to process : 1
2023-04-27 20:57:54,383 INFO mapreduce.JobSubmitter: number of splits:1
2023-04-27 20:57:54,643 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1682168633199_0003
2023-04-27 20:57:54,644 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-04-27 20:57:54,911 INFO conf.Configuration: resource-types.xml not found
2023-04-27 20:57:54,912 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-04-27 20:57:55,037 INFO impl.YarnClientImpl: Submitted application application_1682168633199_0003
2023-04-27 20:57:55,145 INFO mapreduce.Job: The url to track the job: http://bashkatova-VirtualBox:8088/proxy/application_1682168633199_0003/
2023-04-27 20:57:55,147 INFO mapreduce.Job: Running job: job_1682168633199_0003
2023-04-27 20:58:04,458 INFO mapreduce.Job: Job job_1682168633199_0003 running in uber mode : false
2023-04-27 20:58:04,460 INFO mapreduce.Job: map 0% reduce 0%
2023-04-27 20:58:10,591 INFO mapreduce.Job: map 100% reduce 0%
2023-04-27 20:58:16,680 INFO mapreduce.Job: map 100% reduce 100%
2023-04-27 20:58:17,703 INFO mapreduce.Job: Job job_1682168633199_0003 completed successfully
2023-04-27 20:58:17,875 INFO mapreduce.Job: Counters: 53

File System Counters
  FILE: Number of bytes read=138684
  FILE: Number of bytes written=709449
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=188047
  HDFS: Number of bytes written=100850
  HDFS: Number of read operations=8
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2

Job Counters
```

```
hadoop@bashkatova-VirtualBox:~$ hadoop fs -get /user/hadoop/Faust_1_Output/part-r-00000 Faust_1_Output.csv
2023-04-27 20:59:45,852 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$
```



```
hadoop@bashkatova-VirtualBox:~$ head -10 Faust_1_Output.csv
"Allein,      1
"Alles      1
"Als        1
"Der        1
"Die        2
"Er         2
"Ich        4
"Im         1
"Mein       1
"Nur        1
```

```
hadoop@bashkatova-VirtualBox:~$ hadoop fs -put Bosenko/practice/Faust_1.txt /user/hadoop/Faust_2.txt
2023-05-03 15:22:53,591 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ hadoop jar hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.1.2.jar grep /user/hadoop/Faust_2.txt /user/hadoop/Faust_1_Count_Output 'Faust'
2023-05-03 15:24:14,148 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-05-03 15:24:15,224 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2023-05-03 15:24:15,910 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job_1683114645438_0003
2023-05-03 15:24:16,326 INFO input.FileInputFormat: Total input files to process : 1
2023-05-03 15:24:16,433 WARN hdfs.DataStreamer: Caught exception
java.lang.InterruptedExce...
    at java.lang.Object.wait(Native Method)
```

```
hadoop@bashkatova-VirtualBox: ~
Файл  Правка  Вид  Поиск  Терминал  Справка

Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=144
CPU time spent (ms)=930
Physical memory (bytes) snapshot=351666176
Virtual memory (bytes) snapshot=5115473920
Total committed heap usage (bytes)=246943744
Peak Map Physical memory (bytes)=228233216
Peak Map Virtual memory (bytes)=2554441728
Peak Reduce Physical memory (bytes)=123432960
Peak Reduce Virtual memory (bytes)=2561032192
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=108
File Output Format Counters
  Bytes Written=9
hadoop@bashkatova-VirtualBox:~$
```

Show 20 entries									
ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStat
application_1683114645438_0004	hadoop	grep-sort	MAPREDUCE	default	0	Wed May 3 15:24:43 +0300 2023	Wed May 3 15:25:10 +0300 2023	FINISHED	SUCCESS
application_1683114645438_0003	hadoop	grep-search	MAPREDUCE	default	0	Wed May 3 15:24:17 +0300 2023	Wed May 3 15:24:41 +0300 2023	FINISHED	SUCCESS
application_1683114645438_0002	hadoop	word count	MAPREDUCE	default	0	Wed May 3 15:06:04 +0300 2023	Wed May 3 15:06:30 +0300 2023	FINISHED	SUCCESS
application_1683114645438_0001	hadoop	word count	MAPREDUCE	default	0	Wed May 3 14:54:06 +0300 2023	Wed May 3 14:54:39 +0300 2023	FINISHED	SUCCESS

```
hadoop@bashkatova-VirtualBox:~$ hadoop fs -get /user/hadoop/Faust_1_Count_Output
/part-r-000000 Faust_1_Count_Output.csv
2023-05-03 15:27:11,942 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where applicable
hadoop@bashkatova-VirtualBox:~$
```

```
hadoop@bashkatova-VirtualBox:~$ cat Faust_1_Count_Output.csv
50      Faust
```