

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

ДИСЦИПЛИНА:

Инструменты для хранения и обработки больших данных

Практическая работа 05_2

Тема:

Introduction to HDFS.

Выполнила: Соколова М. С., группа: АДЭУ-201

Преподаватель: Босенко Т. М.

Москва

2023

Вывести с помощью команды help описание основных команды shell-клиента

```
hadoop@sokolova-VirtualBox:~$ help
GNU bash, version 4.4.20(1)-release (x86_64-linux-gnu)
These shell commands are defined internally. Type 'help' to see this list.
Type 'help name' to find out more about the function 'name'.
Use 'info bash' to find out more about the shell in general.
Use 'man -k' or 'info' to find out more about commands not in this list.

A star (*) next to a name means that the command is disabled.

job_spec [&]
(( expression ))
. filename [arguments]
:
[ arg... ]
[[ expression ]]
alias [-p] [name[=value] ... ]
bg [job_spec ...]
bind [-lpsvPSVX] [-m keymap] [-f filename] [-q]>
break [n]
builtin [shell-builtin [arg ...]]
caller [expr]
case WORD in [PATTERN [| PATTERN]...) COMMANDS>
cd [-L|[-P [-e]] [-@]] [dir]
command [-pVv] command [arg ...]
compgen [-abdefgjkuv] [-o option] [-A action]>
complete [-abdefgjkuv] [-pr] [-DE] [-o optio>
compopt [-o|+o option] [-DE] [name ...]
continue [n]
coproc [NAME] command [redirections]
declare [-aAfFgIlNrtux] [-p] [name[=value] ...>
dirs [-clpv] [+N] [-N]
disown [-h] [-ar] [jobspec ... | pid ...]
history [-c] [-d offset] [n] or history -anrw>
if COMMANDS; then COMMANDS; [ elif COMMANDS; >
jobs [-lnprs] [jobspec ...] or jobs -x comman>
kill [-s sigspec | -n signum | -sigspec] pid >
let arg [arg ...]
local [option] name[=value] ...
logout [n]
mapfile [-d delim] [-n count] [-O origin] [-s>
popd [-n] [+N | -N]
printf [-v var] format [arguments]
pushd [-n] [+N | -N | dir]
pwd [-LP]
read [-ers] [-a array] [-d delim] [-i text] [>
readarray [-n count] [-O origin] [-s count] [>
readonly [-aAf] [name[=value] ...] or readonl>
return [n]
select NAME [in WORDS ... ;] do COMMANDS; do>
set [-abefhkmnptuvxBCHP] [-o option-name] [-->
shift [n]
shopt [-pqsu] [-o] [optname ...]
source filename [arguments]
suspend [-f]
test [expr]
```

Просмотреть корневую директорию HDFS




```
hadoop@sokolova-VirtualBox:~$ hdfs dfs -ls
2023-05-03 18:34:01,602 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
using builtin-java classes where applicable
Found 5 items
-rw-r--r-- 1 hadoop supergroup 187937 2023-04-27 22:15 Faust_1.txt
drwxr-xr-x - hadoop supergroup 0 2023-04-27 22:28 Faust_1_Count_Output
drwxr-xr-x - hadoop supergroup 0 2023-04-27 22:20 Faust_1_Output
-rw-r--r-- 1 hadoop supergroup 1340963 2023-04-27 21:58 dpkg.log
drwxr-xr-x - hadoop supergroup 0 2023-04-27 21:59 test_output
```

Создать в HDFS в директории /user/mgпу поддиректорию ваше_фио

```
hadoop@sokolova-VirtualBox:~$ hdfs dfs -mkdir -p /user/mgpu/SokolovaMS
```

Browse Directory

Go!






Show

25

 entries

Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	hadoop	supergroup	0 B	Apr 27 22:28	0	0 B	hadoop	
<input type="checkbox"/>	drwxr-xr-x	hadoop	supergroup	0 B	May 03 18:53	0	0 B	mgpu	

Showing 1 to 2 of 2 entries

Previous

1

Next

Hadoop, 2018.

Browse Directory

Go!



Show

25

 entries

Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	hadoop	supergroup	0 B	May 03 18:53	0	0 B	SokolovaMS	

Showing 1 to 1 of 1 entries

Previous

1

Next

Hadoop, 2018.

```
hadoop@sokolova-VirtualBox:~$ base64 /dev/urandom | head -c 10000000 > sms.txt
hadoop@sokolova-VirtualBox:~$ gzip -c sms.txt > sms.gz
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -copyFromLocal sms.txt /user/mgpu/SokolovaMS
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -copyFromLocal sms.gz /user/mgpu/SokolovaMS
```

Show 25 entries

Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	7.25 MB	May 03 19:28	1	128 MB	sms.gz	
<input type="checkbox"/>	-rw-r--r--	hadoop	supergroup	9.54 MB	May 03 19:27	1	128 MB	sms.txt	

Showing 1 to 2 of 2 entries

Previous
1
Next

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -cat /user/mgpu/SokolovaMS/sms.txt part-m-*
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -cat /user/mgpu/SokolovaMS/sms.gz part-m-*
#####H####IK## \Ho@Z0^### g/###.#+#X#####
v#}#####xf###,y#b9%#E%# jS## N#5##
V#O#f####8B#u-##4l#leil#e#f####W#####n#e+##d#r#z#w#####;.#.(#[*\###
J#C#=# #####6.h#;jA#eTV####T("#####)H#F#L#ez
UL###i#@#> #eZ#S####u#s#v#Q#K#A#Z#LY#####FGl#Gt#LL#Y#f$SE#e#g#N#L#n#*#X#F#J#-#g#g#RY
#####
d/#3##'
-B#B#####
M#o#C#M#e#s,#
K#e#f#####zeiv#e#.,#:#)#U#B#####H#e#Ked#pH'_UJ#f15de#f2
y#U#Swe#Fe#K#e#"#e" #kV#ve:#=##/8###!#,M###)8####|E#e#\K#/##
U#####
jS#e#(##)##l#e#c#e#8pb #'!#)SDpJZ|##C#Z#/#H####c|-*#B#清F#zM\F#d#s#
Lf#e#e#C3## p'we#EXmg X#####L%#)#####J#(C O#?#A'K#e#Y#V[1E"#e#m#p#<Q#e#E#-#e#C#Xte#v#n6#
#f#r#e#Ye#Me#Z#M#lf#A#Co#e#M#l#e#9#v#e#..#####U#=#X1:8#q#e#U#)#Ge#f#e#rD#k#e#4#l#e#w#####2#
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -cat /user/mgpu/SokolovaMS/sms.txt | head
2023-05-03 19:33:50,537 WARN util.NativeCodeLoader: Unable to load native-hadoop lib
rary for your platform... using builtin-java classes where applicable
Y/gEsmCncdFUEg9rpaad5SDuyJLOxwciUPgqwBDstW4+1xSS91Qqo9ZRq04/vvQHG5SSH5etfwCj
Ghgj79fQ1xTFcUXYOpBzGiVfmcFHPPGG+Mt0GibfU2+3+8shsSiXDRZFRm80p9B004HogGukxCF
GvhJ08gZN8R49IJdy+RXZ9FNqacPcLr3J2fZWLQDAQgcUuXwRcLHibzysF+20NDhevo2vSCRsJ8Gyw
V6dLhzv0rHc06ev7MXkFNa+asvNWEqb42MiDhZJq600WLZ77bCu0o9eQK05Vsbj+0NISF9jmLE
kJfLlrh4jhEo4q/7f4Lj/5370UxUddEeI6/ddhdA60Gp4rjQU+cBIOS0qyдах9UHHG0BVn0eQX3L
77bdj9A+gzzVvltYOSZ6NpmTIGjbjbSzdjer3CLRYfv70Do9vcxJBo/Sqm/f+01PWY8Q/gbHMRVVO
cTzuNHSScFe2qzvVao9ZYTG0Meya8BoBwGuli73+yh/aq8tCeZmvTV216gVniIBJHwLI1rwqyfr
RGyTo+PB0To/FdbpYlR8IuY1UAHkREeOIYDC40NC0c5q8t4GDo69ayIzBfgcDXABXkysy02PQcm
3epCw9ciCSKxuhkQy8dxg8sSueaE8AJefTi8R4SwakzZUEGVhCq9X5FG1Zaw46fmeVv4d8WZiNZR
Acc9K4r0482i/yAB5FoDipnMdTRnD6YhkyWbuXs/09f8nyTmmLC9JVND8TIBUZ2Let2iWBVD8L5p
cat: Unable to write to output stream.
```

[illegible]

```
hadoop@sokoLova-VirtualBox:~$ hadoop fs -cat /user/mgpu/SokolovaMS/sms.txt | tail
2023-05-03 19:34:23,081 WARN util.NativeCodeLoader: Unable to load native-hadoop lib
rary for your platform... using builtin-java classes where applicable
mwUFZQecP9ei4FR9X23k7rFUBXHeUGKeR2uC3NFY2sHdFWfSfHezMl+UmG8EN08lBg09qEZ+h0p
Jusd/jJtZpBWPwszB28x2cqfZhLd2NcNeHX/Mrm8BhMt9Lmi23mzpJ3CErd7r0EkZuccbl0YCRKL
461E2shfVnR075ars08+4wfwKXuxGT+EZ/aDLO6kMDZ5n77DdlAayWuSP61hD3dB+8AY9c1qLRt7D0
ohshf/xRnQp3UaxSeFc+inwQCGRWiUC316QYCGNIV67tIfGF0fbH3lNfmuWz5Lia6/4VuPEscUov
js40/oasPGdm9vUovSPMSnqatABV00NCHVNHXRxGds77NqMa/q1wIiJvcr1P+8YNNMYFY3bKNQuv
DGN11/XGn10wpt0/iVJQmsFGtrxoivf0S/Nvu6B01pGfxbCBWM1EvaEE00kfEpday3n0KniBU+E1
bSNLZ9rqAq1YyuvZ39+ZL4qajUdf32SaDcLwD4Bs+LLA6+7y4U1neVdy5cN8V3cTBXTqMdIF5iJfR
CFReothHUMWWDiYXqv1ZGcTZjJ/cfJk1kJWysdTYnt43bmd0z/dxy9oDvsj0Psbv2AZPy+mGsyX5T
ffAm2pVNQWZ2PPage3DmqwYI6zak2cx6bkjxAMvr0xLwDXtkfkdoh3vH+R+gynIPk4kT/WMT1Ky
su9RshGjDNhadoop@sokoLova-VirtualBox:~$
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -text /user/mgpu/SokolovaMS/sms.txt | head
2023-05-03 19:35:13.916 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
Y/gESmCncDfUEg9rpaad5SduyJL0xwciuPgqwBDstW4+1xSS9IqQo9ZRQ04/vvQHGS5SH5etfwCj
Ghgj79FqIXtFCUXOpYbZG1VfMFHPFGG+Mt0GtbfU2+3+8shsBisXDRZFrm80pB004HG0GukxCF
G6wJ08GNR49JldY+RXZ9qwcPlR3J2FZWLQDAQgcuUwXrCL9HibzysF+0NDHevo2sGrJ8GyG
VldHzhv0rHc06eV7MXkFNA+asVNWEqb4z2MiDhZJQ60wLz277bCu09eQKZ205Vbj+0NISF9jmLte
kJFLirh4jhEoq4/7F4Lj/5370UuxddEEI6/ddhdA60Gp4rjQU+cBI0S0qydaH9UHHG0BVN0eQX3L
72bdj9A+gZzVvltY0S2G9NmTIGtGjbiSdzJer3CLRYFV70D0VcxJBo/SqmF+01PWY8Q/gBhMRVVO
c77uHWSGFe429vVw0ZG0T0cMyeaa8B8WgJl73+yh/aq7hGzmvTV2L6gVmbTbJHwLlJbW6yofyR
RGyT0+PB0Toc/FdbPyLR8IuY1UAHKREeOIYDC40NC0c5q8t4GD069ayIzBfcgDXABKksyd02PHcm
3epCw9ciCSKxuhkQy8dxg8sSueaE8A3eft18R4SwakzZuEGVhCq9X5FG1Zaw46fmeVv4dBwZlnZR
ACE9K4r0482i/yAB5FoDipnMdTrND6VhKyWbuxS/09f8nyTmmlc9JVND8TIBUZ2Let2iWBVD8L5p
text: Unable to write to output stream.
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -text /user/mgpu/SokolovaMS/sms.gz | head
2023-05-04 14:12:40.039 WARN util.NativeCodeLoader: Unable to load native-hadoop li
ing builtin-java classes where applicable
QlW6AgipgdvP08Yp4VnXCMPUUUBF4HQ01BjXw1ezsEPeSsn456W9MMLWTQghF7EfffG25IhUpo04G
cRZG2rslYUH8Z18HCJIXwtpT/hcfot+btLZdthGcJcryEZ1PmTxPdBgGDy22f66UkHo/6FHA00J3Rk
krIHbmlLH833J7Jo7vmu/ywX0XnjbzqRCPjf3ibJ6nVA9D1a1m035I5/IKD5nZiJnhDeGgXJCq/KR
Thn8yxJmjJAgQogeVheU+S0KCZ/BcMx6eBHX040LZ58qGSoSvUvc/0DKC4apLrLIqxcP/T5LWwzr
/syudZ/M3hgjgXg1J08PVDcGh/CNRC/L/aGdGpJkCEtW0c/s3p/AVIjTNTJaDQ05BDLXpgW8HTpN
eyG8ttdiQyM9/6mPirD8JqXGtgvM/xp5XK0mdpj17a18Hdtowiy7EaeiAM4YJP/PrC9SeILCbT
+5uyj3lHznHnZ3vjyQ0bCl+KQXYVzfzrAwz207HfnT08RtLw/zd3f2fmx0+0q32f2mZdF02T500E2
7bZmDcxPCW111yaXz0cAhZdrJpcssyLAmDTXSL3rmUbavDf2pIZUB7IH4ToS078UyNxnhdXvuk9A
9p/ImSQEAwtayKpVrNr07u/7yecmnHEvFQP5q35SkwGypvzHtdUnPEK7B6ofKU/W8Kdz1QCYkrU
5QRLZ9eULUsb9Uj22I2MYN4Refe2j57GLfLrN7DFZIOme1pVYVnLcfJR3TIEpwsGhEIp1TCjDFLYd
text: Unable to write to output stream.
```

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -text /user/mgpu/SokolovaMS/sms.txt | tail
2023-05-03 19:35:40,557 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
mwUFZQecP9e9i4fR9XZ3k7rFU8XheUGKeR2uC3NFY2sHdFwfsHezML+UmG8EN08lBg09qEZ+h0p
Jusd/jJTzPBWPwszB28x2cqfZhLd2NcNeHX/Mrm8BhMt9Lmi23mzpJ3CErd7r0EkZuccbl0YCRKL
461E28fW075ars08+4wfZKUXGT+eZj/aDL06kMDZz5n77DfLaAYuSP61hD3dB+8AY0c1qLR7D00
ohshf/xasPQ3UaxSeFc+iywOCGRWiuC316QyCGNIV6iIFfDG0Fbh3lNfmcwZ5Li6/4VU3ESKuov
j5s4d/0xnPgdm9uovSMPsn4dBAV00NCVHNHXRjXGds77NgqA/q1wI1jcyf1P+8VNNYF19PBKqVv
DGN11/XGn10wpt0/iVJQmsFGtrxi0iVf0S/Nvu6B01pGfxbCBWM1EVAEE00kFepdaY3n0KNiBU+E1
bSNL29rhA1YyuvZ39+4Lq4ajqJf32c3SaDcLWD4Bs+llA6+7y4U1neVdy5Cn8V3cBTXTqMdI5f1JfR
CFRe0tqHUMwDIYxQv1ZJfJf3fJk1JWYsdTYNt43bmd0x/dxy903vdsj0Psbv2KZPy+mGYSX5T
ffAm2pVN0wQwZVPage33DmCqV16zak2cx6bKjxAMvr0xLwDxktfKd9h3vHr+gyn1pIk4rt/WM1Ky
```

Создать копию файла file.txt вида date_file.txt, где в начале имени файла-копии указана текущая дата. Вывести листинг

```
hadoop@sokolova-VirtualBox:~$ ls
03-05-23-smss.txt      examples.desktop      Faust_1_Output.csv   hadoop-3.1.2.tar.gz  sms.gz
Big-Data-Storage-and-Processing-Tools  Faust_1_Count_Output.csv  hadoop               hadoopdata           sms.txt
```

Вывести статистику по директории /user/mgru/фio виртуальной машины

```
17599365 17599365 /user/mgpu/SokolovaMS
hadoop@sokolova-VirtualBox:~$ hdfs dfs -du -h /user/mgpu/SokolovaMS
2023-05-03 20:48:24,359 WARN util.NativeCodeLoader: Unable to load n
java classes where applicable
7.2 M 7.2 M /user/mgpu/SokolovaMS/sms.gz
9.5 M 9.5 M /user/mgpu/SokolovaMS/sms.txt
hadoop@sokolova-VirtualBox:~$ hdfs dfs -du -s /user/mgpu/SokolovaMS
2023-05-03 20:48:32,379 WARN util.NativeCodeLoader: Unable to load n
java classes where applicable
17599365 17599365 /user/mgpu/SokolovaMS
hadoop@sokolova-VirtualBox:~$
```

Удалить поддиректорию /fio со всем содержимым

```
hadoop@sokolova-VirtualBox:~$ hadoop fs -rm -r /user/mgpu/SokolovaMS
2023-05-04 13:42:58,518 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your pla
tform... using builtin-java classes where applicable
Deleted /user/mgpu/SokolovaMS
```

Browse Directory

Show entries

Search:

<input type="checkbox"/>	📄	Permission	👤	Owner	👤	Group	📄	Size	📄	Last Modified	📄	Replication	📄	Block Size	📄	Name	📄
No data available in table																	

Showing 0 to 0 of 0 entries

Hadoop, 2018.