

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

**Отчет по практической работе Hadoop-
Установка**

Выполнил: Кряжева К.С

группа: ТП-191

```
lemp@u20-05: ~  
Processing triggers for libc-bin (2.31-0ubuntu9.9) ...  
Processing triggers for man-db (2.9.1-1) ...  
Processing triggers for ca-certificates (20211016~20.04.1) ...  
Updating certificates in /etc/ssl/certs...  
0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d...  
  
done.  
done.  
Processing triggers for sgml-base (1.29.1) ...  
Setting up x11proto-dev (2019.2-1ubuntu1) ...  
Setting up libxau-dev:amd64 (1:1.0.9-0ubuntu1) ...  
Setting up libice-dev:amd64 (2:1.0.10-0ubuntu1) ...  
Setting up libsm-dev:amd64 (2:1.2.3-1) ...  
Setting up libxdmcp-dev:amd64 (1:1.1.3-0ubuntu1) ...  
Setting up x11proto-core-dev (2019.2-1ubuntu1) ...  
Setting up libxcb1-dev:amd64 (1.14-2) ...  
Setting up libx11-dev:amd64 (2:1.6.9-2ubuntu1.2) ...  
Setting up libxt-dev:amd64 (1:1.1.5-1) ...  
lemp@u20-05:~$ java -version  
openjdk version "1.8.0_342"  
OpenJDK Runtime Environment (build 1.8.0_342-8u342-b07-0ubuntu1~20.04-b07)  
OpenJDK 64-Bit Server VM (build 25.342-b07, mixed mode)  
lemp@u20-05:~$
```

```
hadoop@u20-05: /home/lemp  
lemp@u20-05:~$ java -version  
openjdk version "1.8.0_342"  
OpenJDK Runtime Environment (build 1.8.0_342-8u342-b07-0ubuntu1~20.04-b07)  
OpenJDK 64-Bit Server VM (build 25.342-b07, mixed mode)  
lemp@u20-05:~$ sudo nano /etc/environment  
lemp@u20-05:~$ sudo adduser hadoop  
Adding user `hadoop' ...  
Adding new group `hadoop' (1001) ...  
Adding new user `hadoop' (1001) with group `hadoop' ...  
Creating home directory `/home/hadoop' ...  
Copying files from `/etc/skel' ...  
New password:  
Retype new password:  
passwd: password updated successfully  
Changing the user information for hadoop  
Enter the new value, or press ENTER for the default  
    Full Name []:  
    Room Number []:  
    Work Phone []:  
    Home Phone []:  
    Other []:  
Is the information correct? [Y/n] y  
lemp@u20-05:~$ sudo su hadoop  
hadoop@u20-05: /home/lemp$
```

```
hadoop@u20-05: ~
Processing triggers for libc-bin (2.31-0ubuntu9.9) ...
Processing triggers for ufw (0.36-6ubuntu1) ...
Lemp@u20-05:~$ sudo su hadoop
hadoop@u20-05:~/hadoop$ cd
hadoop@u20-05:~$ 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:F/skIPLWVgctTneCVaeXSnYz08vtuBPSHhSJBg50gW4 hadoop@u20-05
The key's randomart image is:
+---[RSA 3072]---+
|      .+.0=.0...|
|      .0 0+.* +00|
|      ... 00++.=.Bo|
|      oEo 0.+0 =.*|
|      .0 S + .+ 0.|
|      . . . +. +0 |
|      .0.0. |
|      0. |
|      .. |
+-----[SHA256]-----+
hadoop@u20-05:~$ >
```

```
hadoop@u20-05: ~
d_keys
ash: /home/hadoop/.ssh/id_rsa.pub: Permission denied
hadoop@u20-05:~$ cat /home/hadoop/.ssh/id_rsa.pub >> /home/hadoop/.ssh/authorized
_keys
hadoop@u20-05:~$ chmod 0600 /home/hadoop/.ssh/authorized_keys
hadoop@u20-05:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
CDSA key fingerprint is SHA256:K2D8LNfHf4GpPa+LqqR0Jj8NWUCYe8SAI2ETi00ZwLU.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.4 LTS (GNU/Linux 5.15.0-46-generic x86_64)

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

8 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

Your Hardware Enablement Stack (HWE) is supported until April 2025.
** System restart required **

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
```

```
hadoop@u20-05: ~
Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

hadoop@u20-05:~$ exit
logout
Connection to localhost closed.
hadoop@u20-05:~$ wget https://archive.apache.org/dist/hadoop/common/hadoop-3.1.2
/hadoop-3.1.2.tar.gz
--2022-10-25 10:41:00-- https://archive.apache.org/dist/hadoop/common/hadoop-3.
1.2/hadoop-3.1.2.tar.gz
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:1
72:2ec5::2
Connecting to archive.apache.org (archive.apache.org)|138.201.131.134|:443... co
nnected.
HTTP request sent, awaiting response... 200 OK
Length: 332433589 (317M) [application/x-gzip]
Saving to: 'hadoop-3.1.2.tar.gz'

hadoop-3.1.2.tar.gz 100%[=====>] 317,03M 708KB/s in 8m 14s

2022-10-25 10:49:14 (658 KB/s) - 'hadoop-3.1.2.tar.gz' saved [332433589/33243358
9]

hadoop@u20-05:~$ █
```

```
hadoop@u20-05: ~
hadoop-3.1.2/share/doc/hadoop/images/icon_warning_s
hadoop-3.1.2/share/doc/hadoop/images/newwindow.png
hadoop-3.1.2/share/doc/hadoop/images/maven-logo-2.c
hadoop-3.1.2/share/doc/hadoop/images/banner.jpg
hadoop-3.1.2/share/doc/hadoop/images/logo_maven.jpg
hadoop-3.1.2/share/doc/hadoop/images/collapsed.gif
hadoop-3.1.2/share/doc/hadoop/images/external.png
hadoop-3.1.2/share/doc/hadoop/images/h5.jpg
hadoop-3.1.2/share/doc/hadoop/images/logos/
hadoop-3.1.2/share/doc/hadoop/images/logos/build-by
hadoop-3.1.2/share/doc/hadoop/images/logos/build-by
hadoop-3.1.2/share/doc/hadoop/images/logos/maven-fe
hadoop-3.1.2/share/doc/hadoop/index.html
hadoop-3.1.2/share/doc/hadoop/project-reports.html
hadoop-3.1.2/include/
hadoop-3.1.2/include/hdfs.h
hadoop-3.1.2/include/Pipes.hh
hadoop-3.1.2/include/SerialUtils.hh
hadoop-3.1.2/include/StringUtils.hh
hadoop-3.1.2/include/TemplateFactory.hh
hadoop@u20-05:~$ mv hadoop-3.1.2 hadoop
hadoop@u20-05:~$ nano .bashrc
hadoop@u20-05:~$ source .bashrc
hadoop@u20-05:~$ █
```

```
hadoop@u20-05: ~  
hadoop-3.1.2/share/doc/hadoop/images/h5.jpg  
hadoop-3.1.2/share/doc/hadoop/images/logos/  
hadoop-3.1.2/share/doc/hadoop/images/logos/build-by-maven-white.png  
hadoop-3.1.2/share/doc/hadoop/images/logos/build-by-maven-black.png  
hadoop-3.1.2/share/doc/hadoop/images/logos/maven-feather.png  
hadoop-3.1.2/share/doc/hadoop/index.html  
hadoop-3.1.2/share/doc/hadoop/project-reports.html  
hadoop-3.1.2/include/  
hadoop-3.1.2/include/hdfs.h  
hadoop-3.1.2/include/Pipes.hh  
hadoop-3.1.2/include/SerialUtils.hh  
hadoop-3.1.2/include/StringUtils.hh  
hadoop-3.1.2/include/TemplateFactory.hh  
hadoop@u20-05:~$ mv hadoop-3.1.2 hadoop  
hadoop@u20-05:~$ nano .bashrc  
hadoop@u20-05:~$ source .bashrc  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/hadoop-env.sh  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/core-site.xml  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/hdfs-site.xml  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/mapred-site.xml  
hadoop@u20-05:~$ nano ^C  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/hdfs-site.xml  
hadoop@u20-05:~$ nano /home/hadoop/hadoop/etc/hadoop/yarn-site.xml  
hadoop@u20-05:~$
```

```
hadoop@u20-05: ~  
total heap and retry cache entry expiry time is 600000 millis  
2022-10-25 11:04:28,211 INFO util.GSet: Computing capacity for map NameNodeRetr  
ache  
2022-10-25 11:04:28,211 INFO util.GSet: VM type = 64-bit  
2022-10-25 11:04:28,211 INFO util.GSet: 0.0299999999329447746% max memory 1.7 GB  
: 532.1 KB  
2022-10-25 11:04:28,213 INFO util.GSet: capacity = 2^16 = 65536 entries  
2022-10-25 11:04:28,270 INFO namenode.FSImage: Allocated new BlockPoolId: BP-12  
638476-172.26.35.25-1666685068258  
2022-10-25 11:04:28,294 INFO common.Storage: Storage directory /home/hadoop/had  
opdata/hdfs/namenode has been successfully formatted.  
2022-10-25 11:04:28,311 INFO namenode.FSImageFormatProtobuf: Saving image file  
/home/hadoop/hadoopdata/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 u  
sing no compression  
2022-10-25 11:04:28,406 INFO namenode.FSImageFormatProtobuf: Image file /home/h  
adoop/hadoopdata/hdfs/namenode/current/fsimage.ckpt_00000000000000000000 of size  
13 bytes saved in 0 seconds .  
2022-10-25 11:04:28,427 INFO namenode.NNStorageRetentionManager: Going to retai  
n 1 images with txid >= 0  
2022-10-25 11:04:28,431 INFO namenode.NameNode: SHUTDOWN_MSG:  
*****  
SHUTDOWN_MSG: Shutting down NameNode at u20-05/172.26.35.25  
*****/  
hadoop@u20-05:~$
```

```
hadoop@u20-05: ~  
-----  
ive datanodes (1):  
  
ame: 127.0.0.1:9866 (localhost)  
ostname: u20-05  
ecommission Status : Normal  
onfigured Capacity: 78143852544 (72.78 GB)  
FS Used: 24576 (24 KB)  
on DFS Used: 24944947200 (23.23 GB)  
FS Remaining: 49182572544 (45.80 GB)  
FS Used%: 0.00%  
FS Remaining%: 62.94%  
onfigured Cache Capacity: 0 (0 B)  
ache Used: 0 (0 B)  
ache Remaining: 0 (0 B)  
ache Used%: 100.00%  
ache Remaining%: 0.00%  
ceivers: 1  
ast contact: Tue Oct 25 11:06:47 MSK 2022  
ast Block Report: Tue Oct 25 11:05:47 MSK 2022  
um of Blocks: 0  
  
adoop@u20-05:~$
```

ds_practice/Exc x | Exercise_5.pdf x | Exercise_5.pdf x | All Applications x

localhost:8088/cluster

Update

hadoop

Cluster

About

Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

DRAWDING

FAILED

KILLED

Scheduler

Tools

All Applications

Cluster Metrics

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

Cluster Nodes Metrics

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

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum
Capacity Scheduler	[memory-mb (unit-M), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU Vcores	Allocated Memory MB	Reserved CPU Vcores	Reserved Memory MB	% of Queue	% of Cluster
No data available in table																

Showing 0 to 0 of 0 entries

2022
MSK 2022

ds_practice/Exc x | Exercise_5.pdf x | All Applications x | Namenode info x | +

localhost:9870/dfshealth.html#tab-overview

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Overview 'localhost:9000' (active)

Started:	Tue Oct 25 11:05:43 +0300 2022
Version:	3.1.2, r1019dde65bdf12e05ef48ac71e84550d589e5d9a
Compiled:	Tue Jan 29 04:39:00 +0300 2019 by sunlig from branch-3.1.2
Cluster ID:	CID-435e8d80-2d19-48fb-8934-ca7043c7e535
Block Pool ID:	BP-1297638476-172.26.35.25-1666689068258

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 88.98 MB of 294.5 MB Heap Memory. Max Heap Memory is 1.69 GB.
 Non Heap Memory used 47.57 MB of 49.05 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

Configured Capacity:	72.78 GB
----------------------	----------

hadoop@u20-05: ~

DFS Remaining%: 62.94%
 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: Tue Oct 25 11:06:47 MSK 2022
 Last Block Report: Tue Oct 25 11:05:47 MSK 2022
 Num of Blocks: 0

```
hadoop@u20-05:~$ hadoop fs -mkdir /user
2022-10-25 11:13:44,683 WARN util.NativeCodeLoader: Unable to load native-hadoop
  library for your platform... using builtin-java classes where applicable
hadoop@u20-05:~$ hadoop fs -mkdir /user/hadoop
2022-10-25 11:13:59,180 WARN util.NativeCodeLoader: Unable to load native-hadoop
  library for your platform... using builtin-java classes where applicable
hadoop@u20-05:~$ hadoop fs -ls /
2022-10-25 11:14:16,137 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 2022-10-25 11:13 /user
hadoop@u20-05:~$
```

ds_practice/Exc x | Exercise_5.pdf x | All Applications x | Browsing HDFS x

localhost:9870/explorer.html#/

Update

HadoopOverviewDatanodesDatanode Volume FailuresSnapshotStartup ProgressUtilities

Browse Directory

/

Go!

Show 25 entries

Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	hadoop	supergroup	0 B	Oct 25 11:13	0	0 B	user	

Showing 1 to 1 of 1 entries

PreviousNext

Hadoop, 2018.

ds_practice/Exc x | Exercise_5.pdf x | All Applications x | Browsing HDFS x

localhost:8088/cluster

Update

hadoop

Cluster

AboutNodesNode LabelsApplicationsNEWNEW SAVINGSUBMITTEDACCEPTEDRUNNINGFINISHEDPAILEDKILLED

Scheduler

Tools

Cluster Metrics

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

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-M), 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 V-Cores	Allocated Memory
application_1666685179002_0001	hadoop	word count	MAPREDUCE	default	0	Tue Oct 25 11:18:56 +0300 2022	Tue Oct 25 11:19:12 +0300 2022	FINISHED	SUCCEEDED	N/A	N/A	N/A

Showing 1 to 1 of 1 entries

dpkg.log

Show all

Activities Text Editor ▾

Open ▾



```
1 0.04-6 7
2 0.08-3 7
3 0.11.0-5 7
4 0.17-2 20
5 0.24-1ubuntu3 28
6 0.34.2-1 7
7 0.36-6ubuntu1 4
8 0.37.1-1 14
9 0.4-1 7
10 0.72 7
11 06:02:29 7
12 06:02:30 12
13 06:02:35 15
14 06:02:40 15
15 06:02:45 15
16 06:02:50 15
17 06:02:55 15
18 06:02:59 1
19 06:03:00 14
20 06:03:04 7
21 06:03:05 8
22 06:03:09 15
23 06:03:14 15
24 06:15:01 7
25 06:15:02 21
26 06:15:03 16
27 06:15:04 4
28 06:15:09 11
29 06:15:14 3
30 06:15:16 6
31 06:15:22 13
32 06:15:26 3
33 06:15:28 22
34 06:15:30 5
35 06:15:31 4
36 06:15:33 11
37 06:15:52 1
38 06:15:56 7
39 06:15:57 8
40 06:16:01 2
41 06:16:02 3
42 06:16:03 6
```

<

ds_practice/Exc x | Exercise_5.pdf x | All Applications x | Browsing HDFS x

localhost:9870/explorer.html#/user/hadoop/test_output

Update

Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities

Browse Directory

user/hadoop/test_output

Go

Show 25 entries

Search:

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
-rw-r--r--	hadoop	supergroup	0 B	Oct 25 11:19	1	128 MB	._SUCCESS
-rw-r--r--	hadoop	supergroup	8.12 KB	Oct 25 11:19	1	128 MB	part-r-00000

Showing 1 to 2 of 2 entries

Previous 1 Next

Hadoop, 2018.

part-r-00000

dpkg.log

Show all

x11proto-core-dev:all 7

x11proto-dev:all 7

xorg-sgml-doctools:all 8

xtrans-dev:all 7

hadoop@u20-05: ~