

Yandex

Big Data

HDFS Client





https://en.wikipedia.org/wiki/Command-line_interface



man
info



man

hdfs dfs -help

info

hdfs dfs -usage <utility_name>



man

hdfs dfs -help

info

hdfs dfs -usage <utility_name>



man

hdfs **dfs** -help

info

hdfs **dfs** -usage <utility_name>



man	hdfs dfs -help
info	hdfs dfs -usage <utility_name>
	hdfs namenode
	hdfs datanode
	...

```
$ hdfs dfs -ls -R -h /data/wiki
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles
```

```
-rw-r--r-- 3 hdfs supergroup 11.5 G 2017-01-07 20:59 /data/wiki/en_articles/  
articles
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles_part
```

```
-rw-r--r-- 3 hdfs supergroup 73.3 M 2017-01-07 20:59 /data/wiki/en_articles_part/  
articles-part
```



```
$ hdfs dfs -ls -R -h /data/wiki
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles
```

```
-rw-r--r-- 3 hdfs supergroup 11.5 G 2017-01-07 20:59 /data/wiki/en_articles/  
articles
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles_part
```

```
-rw-r--r-- 3 hdfs supergroup 73.3 M 2017-01-07 20:59 /data/wiki/en_articles_part/  
articles-part
```

```
$ hdfs dfs -ls -R -h /data/wiki
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles
```

```
-rw-r--r-- 3 hdfs supergroup 11.5 G 2017-01-07 20:59 /data/wiki/en_articles/  
articles
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles_part
```

```
-rw-r--r-- 3 hdfs supergroup 73.3 M 2017-01-07 20:59 /data/wiki/en_articles_part/  
articles-part
```

```
$ hdfs dfs -ls -R -h /data/wiki
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles
```

```
-rw-r--r-- 3 hdfs supergroup 11.5 G 2017-01-07 20:59 /data/wiki/en_articles/  
articles
```

```
drwxr-xr-x - hdfs supergroup 0 2017-01-07 20:59 /data/wiki/en_articles_part
```

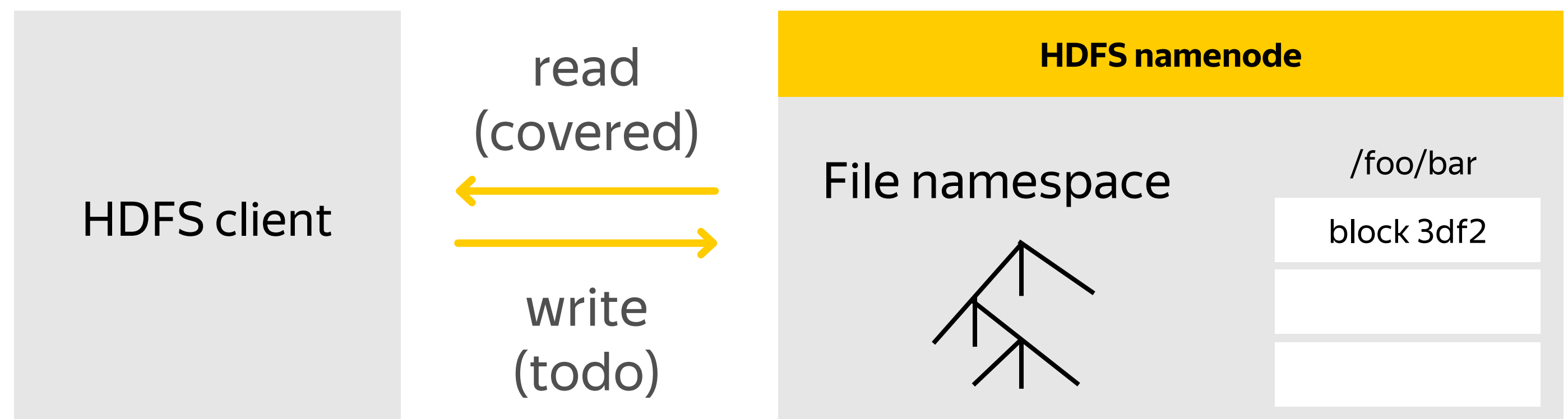
```
-rw-r--r-- 3 hdfs supergroup 73.3 M 2017-01-07 20:59 /data/wiki/en_articles_part/  
articles-part
```



```
$ hdfs dfs -du -h /data/wiki  
11.5 G 34.4 G /data/wiki/en_articles  
73.3 M 219.9 M /data/wiki/en_articles_part
```

```
$ hdfs dfs -du -h /data/wiki
```

Filesystem	Size	Used	Available	Use
hdfs://virtual-master.atp-fivt.org:8020	1.6 T	389.0 G	1.1 T	23 %



```
$ hdfs dfs -ls
Found 1 items
drwx----- - adral adral  0 2017-04-05 12:10 .Trash
$ hdfs dfs -mkdir deep/nested/path
mkdir: 'deep/nested/path': No such file or directory
$ hdfs dfs -mkdir -p deep/nested/path
$ hdfs dfs -ls -R deep
drwxr-xr-x - adral adral 0 2017-04-05 12:28 deep/nested
drwxr-xr-x - adral adral 0 2017-04-05 12:28 deep/nested/path
```



```
$ hdfs dfs -rm deep
```

```
rm: 'deep': ls a directory
```

```
$ hdfs dfs -rm -r deep
```

```
17/04/05 12:33:00 INFO fs.TrashPolicyDefault: Moved: 'hdfs://  
virtual-master.atp-fvt.org:8020/user/adral/deep' to trash at:  
hdfs://virtual-master.atp-fvt.org:8020/user/adral/.Trash/  
Current/user/adral/deep
```

```
$ hdfs dfs -mkdir -p deep/nested/path && hdfs dfs -rm -r  
-skipTrash deep  
Deleted deep
```

```
$ hdfs dfs -touchz file.txt
```

```
$ hdfs dfs -ls
```

```
Found 2 items
```

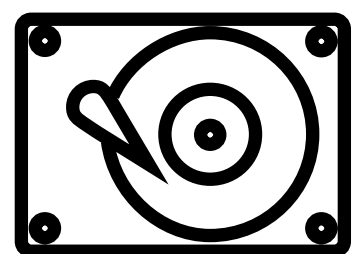
```
drwx----- - adral adral 0 2017-04-05 12:10 .Trash
```

```
-rw-r--r-- 3 adral adral 0 2017-04-05 12:43 file.txt
```

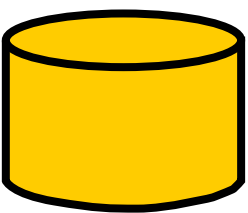
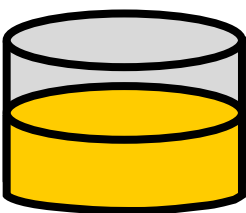
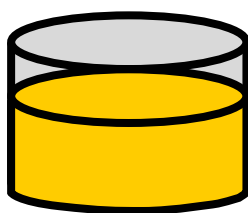
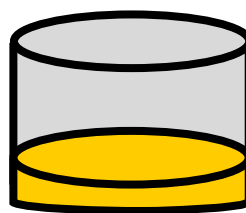
```
$ hdfs dfs -mv file.txt another_file.txt && hdfs dfs -rm  
another_file.txt
```

```
17/04/05 12:50:54 INFO fs.TrashPolicyDefault: Moved: 'hdfs://  
virtual-master.atp-fvt.org:8020/user/adral/another_file.txt' to trash  
at: hdfs://virtual-master.atp-fvt.org:8020/user/  
adral/.Trash/Current/user/adral/another_file.txt
```

test_file.txt

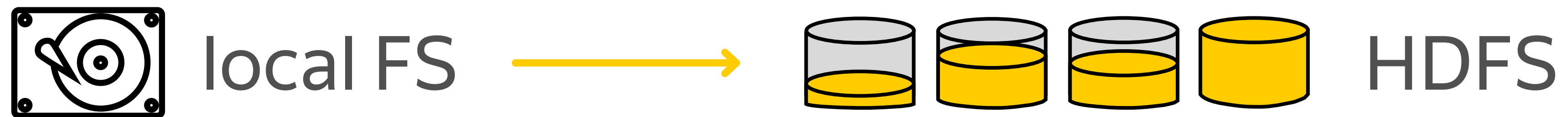


local FS

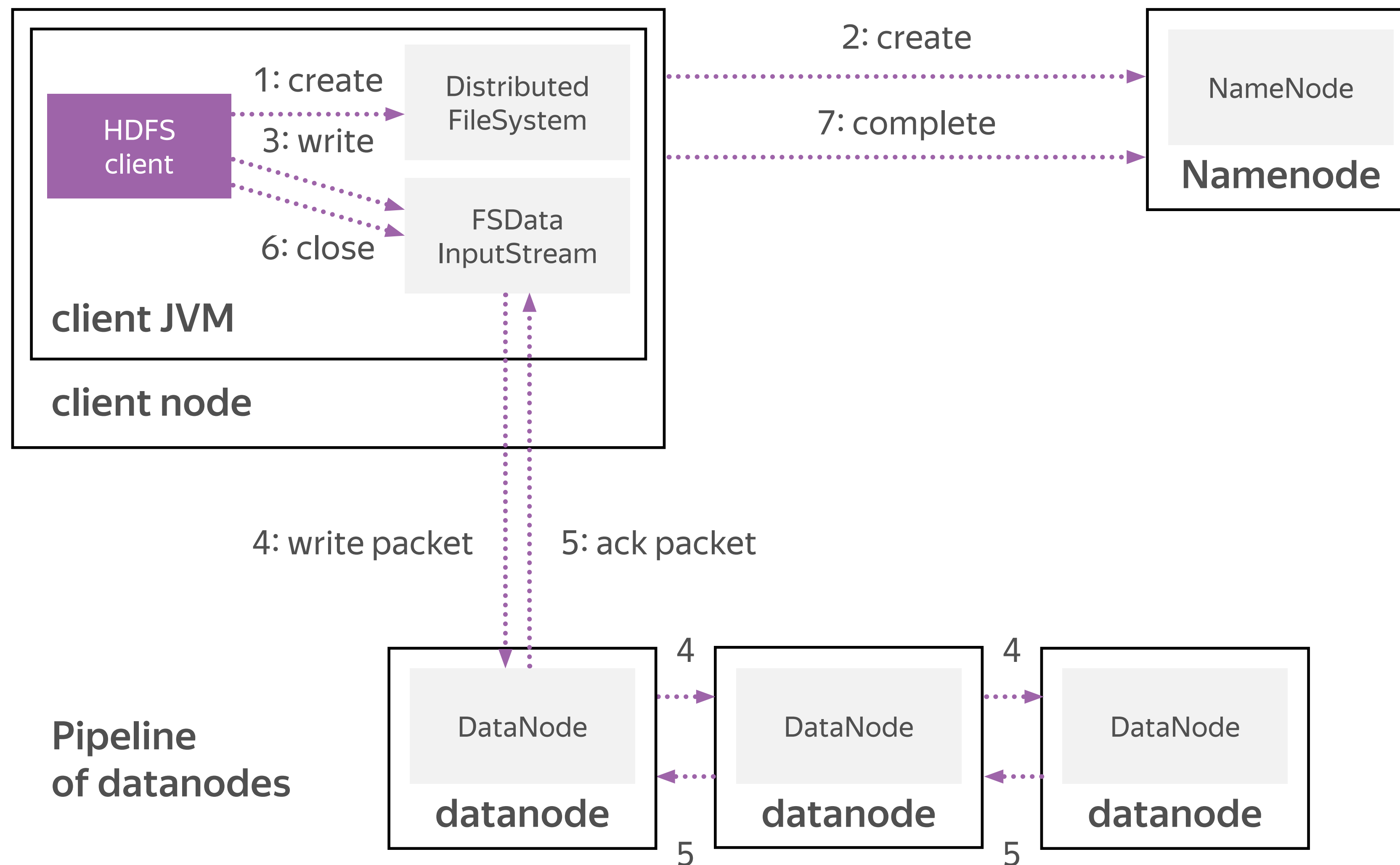


HDFS

test_file.txt



`hdfs dfs -put <source location> <HDFS destination>`



```
$ cat test_file.txt  
test content #1  
test content #2  
...  
test content #14  
test content #15
```

```
$ head -4 test_file.txt  
test content #1  
test content #2  
test content #3  
test content #4
```

```
$ tail -4 test_file.txt  
test content #12  
test content #13  
test content #14  
test content #15
```

```
$ cat test_file.txt
test content #1
test content #2
...
test content #14
test content #15
```

```
$ head -4 test_file.txt
test content #1
test content #2
test content #3
test content #4
```

```
$ tail -4 test_file.txt
test content #12
test content #13
test content #14
test content #15
```

```
$ hdfs dfs -put test_file.txt hdfs_test_file.txt
put: 'hdfs_test_file.txt': File exists
$ hdfs dfs -rm -skipTrash hdfs_test_file.txt
Deleted hdfs_test_file.txt
$ hdfs dfs -put test_file.txt hdfs_test_file.txt
```

```
$ hdfs dfs -cat hdfs_test_file.txt | head -4
test content #1
test content #2
test content #3
test content #4
```

```
$ hdfs dfs -cat hdfs_test_file.txt | tail -4
test content #12
test content #13
test content #14
test content #15
```

```
$ hdfs dfs -tail hdfs_test_file.txt
test content #1
test content #2
...
test content #15
```



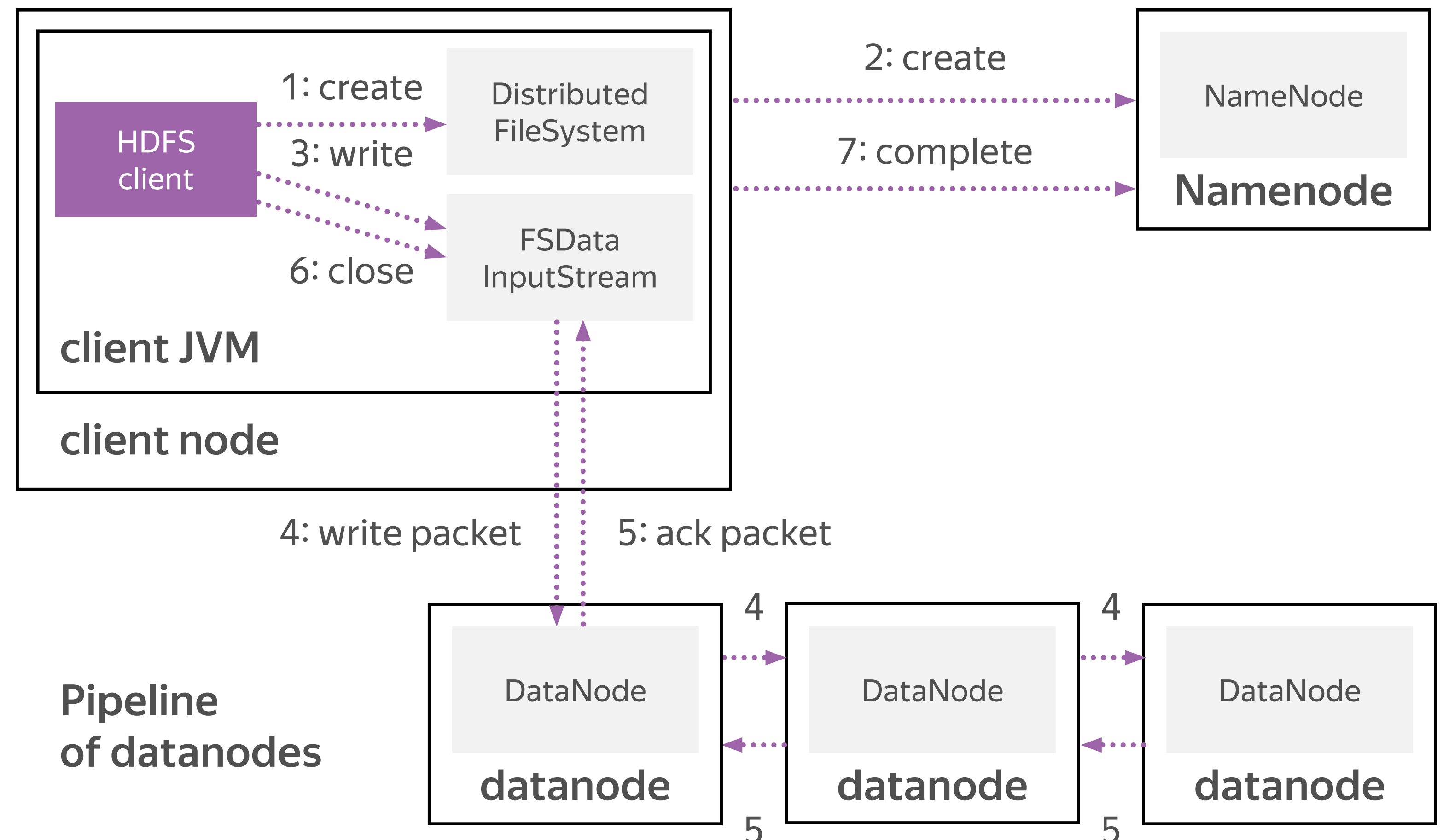
```
$ hdfs dfs -tail hdfs_test_file.txt
test content #1
test content #2
...
test content #15
```

```
$ tail -c 1024
test_file.txt
test content #1
test content #2
test content #3
test content #4
test content #5
test content #6
test content #7
test content #8
test content #9
test content #10
test content #11
test content #12
test content #13
test content #14
test content #15
```

```

$ hdfs dfs -cp hdfs_test_file.txt hdfs_test_file_copy.txt
$ hdfs dfs -get hdfs_test* .
$ ls -lth hdfs*
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file_copy.txt
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file.txt

```



```

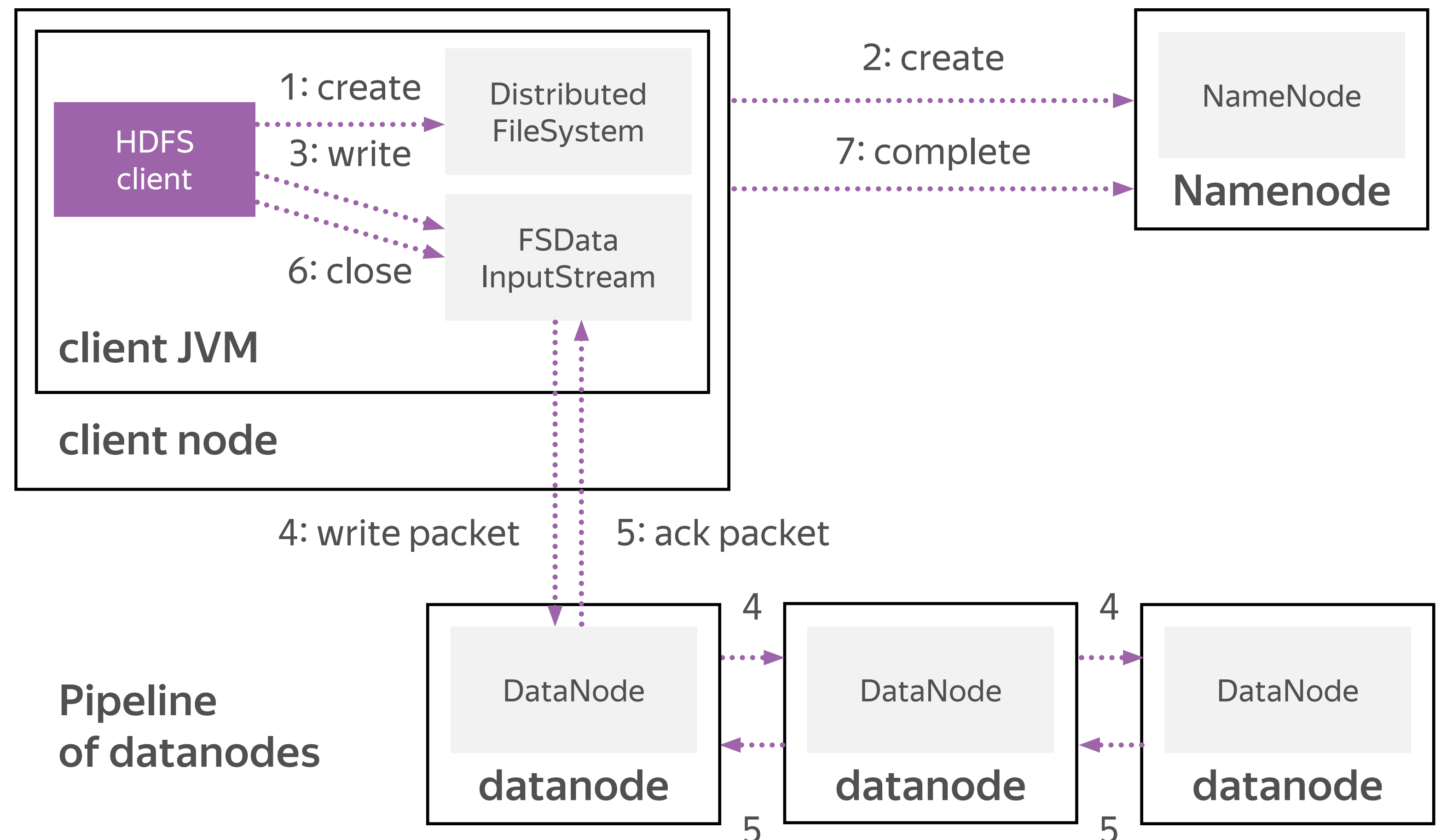
$ hdfs dfs -cp hdfs_test_file.txt hdfs_test_file_copy.txt
$ hdfs dfs -get hdfs_test* .
$ ls -lth hdfs*
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file_copy.txt
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file.txt
$ hdfs dfs -getmerge hdfs_test* hdfs_meged.txt
$ ls -lth hdfs*
-rw-r--r-- 1 adral adral 492 Apr 5 13:32 hdfs_meged.txt
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file_copy.txt
-rw-r--r-- 1 adral adral 246 Apr 5 13:28 hdfs_test_file.txt

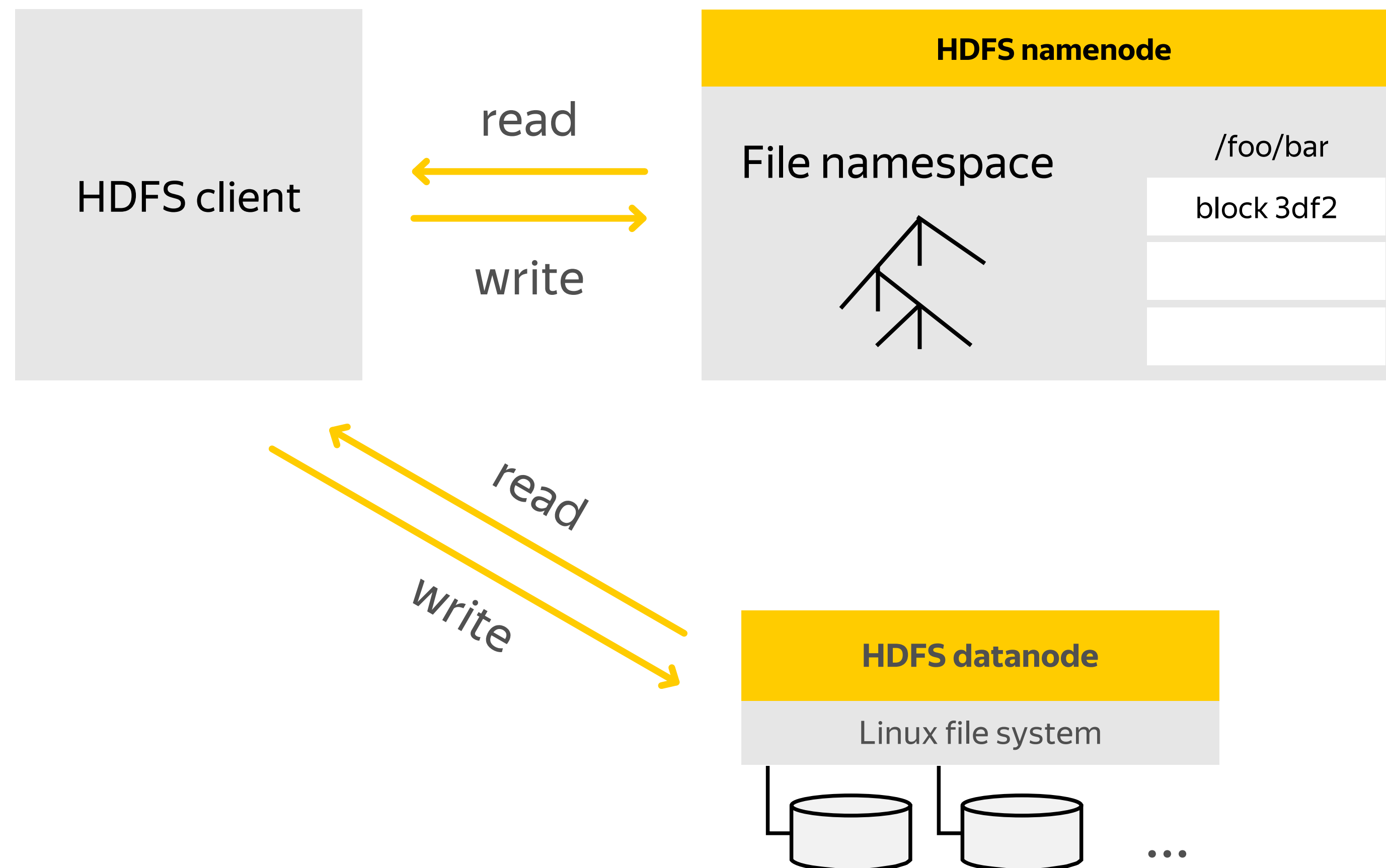
```

```

$ cat hdfs_meged.txt
test content #1
test content #2
...
test content #15
test content #1
...
test content #14
test content #15

```





COMPLETED

chown

chown

hdfs groups

```
$ time hdfs dfs -setrep -w 1 hdfs_test_file.txt
Replication 1 set: hdfs_test_file.txt
Waiting for hdfs_test_file.txt ...
WARNING: the waiting time may be long for DECREASING the number of
replications...done
real 0m13.148s
user 0m4.232s
sys 0m0.156s
```

```
$ hdfs dfs -ls
Found 3 items
drwx----- - adral adral 0 2017-04-05 13:00 .Trash
-rw-r--r-- 1 adral adral 246 2017-04-05 13:03 hdfs_test_file.txt
-rw-r--r-- 3 adral adral 246 2017-04-05 13:28 hdfs_test_file_copy.txt
```

```
$ time hdfs dfs -setrep -w 2 hdfs_test_file.txt
Replication 2 set: hdfs_test_file.txt
Waiting for hdfs_test_file.txt .... done
real 0m13.168s
user 0m4.092s
sys 0m0.148s
```

```
$ hdfs fsck /data/wiki/en_articles -files
Connecting to namenode via http://virtual-master.atp-fvt.org:50070
FSCK started by adral (auth:SIMPLE) from /138.201.91.190 for path /data/wiki/en_articles
at Wed Apr 05 14:01:15 CEST 2017
/data/wiki/en_articles <dir>
/data/wiki/en_articles/articles 12328051927 bytes, 92 block(s): OKsys 0m0.148s
```

```
$ hdfs fsck /data/wiki/en_articles -files -blocks
Connecting to namenode via http://virtual-master.atp-fvt.org:50070
FSCK started by adral (auth:SIMPLE) from /138.201.91.190 for path /data/wiki/en_articles
at Wed Apr 05 14:01:48 CEST 2017
/data/wiki/en_articles <dir>
/data/wiki/en_articles/articles 12328051927 bytes, 92 block(s): OK
0. BP-858288134-138.201.91.191-1481279621434:blk_1073808471_67650 len=134217728
Live_repl=3
1. BP-858288134-138.201.91.191-1481279621434:blk_1073808472_67651 len=134217728
Live_repl=3
2. BP-858288134-138.201.91.191-1481279621434:blk_1073808473_67652 len=134217728
Live_repl=3
...
```

```
$ hdfs fsck /data/wiki/en_articles -files -blocks -locations
Connecting to namenode via http://virtual-master.atp-fvt.org:50070
FSCK started by adral (auth:SIMPLE) from /138.201.91.190 for path /data/wiki/en_articles
at Wed Apr 05 14:01:56 CEST 2017
/data/wiki/en_articles <dir>
```



```
$ hdfs fsck -blockId blk_1073808569
```

```
Connecting to namenode via http://virtual-master.atp-fvt.org:50070
```

```
FSCK started by adral (auth:SIMPLE) from /138.201.91.190 at Wed Apr 05 14:17:54 CEST 2017
```

```
Block Id: blk_1073808569
```

```
Block belongs to: /data/wiki/en_articles/articles
```

```
No. of Expected Replica: 3
```

```
No. of live Replica: 3
```

```
No. of excess Replica: 0
```

```
No. of stale Replica: 0
```

```
No. of decommission Replica: 0
```

```
No. of corrupted Replica: 0
```

```
Block replica on datanode/rack:
```

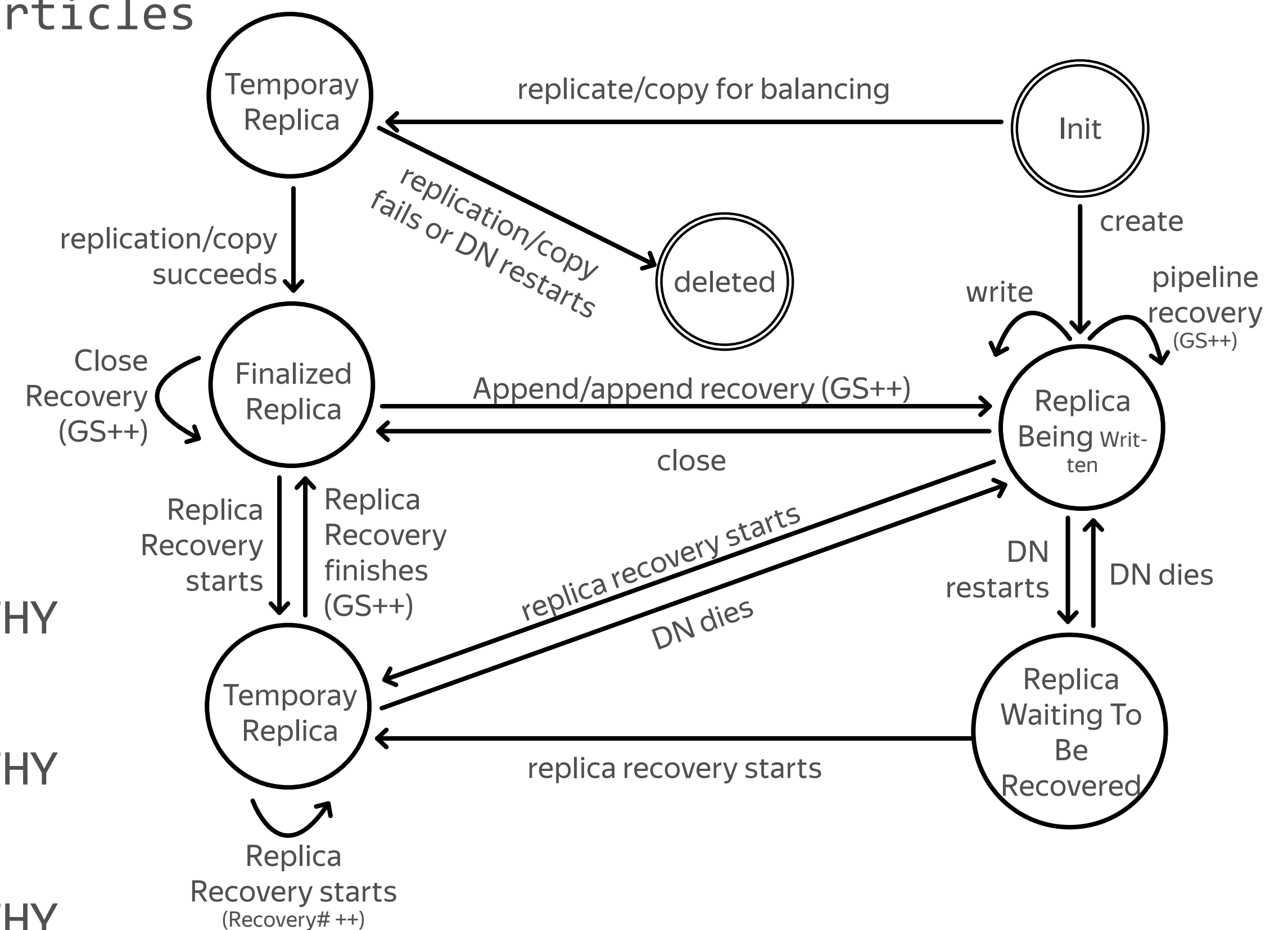
```
virtual-node2.atp-fvt.org/default is HEALTHY
```

```
Block replica on datanode/rack:
```

```
virtual-node3.atp-fvt.org/default is HEALTHY
```

```
Block replica on datanode/rack:
```

```
virtual-node1.atp-fvt.org/default is HEALTHY
```



```
$ hdfs dfs -find /data/wiki -name '*part*'
/data/wiki/en_articles_part
/data/wiki/en_articles_part/articles-part
```

```
$ hdfs dfs -find /data -name '*part*' -iname '*Article*'
/data/wiki/en_articles_part
/data/wiki/en_articles_part/articles-part
```

Summary

Summary

- › you can **request meta-information** from Namenode and change its structure (ls, mkdir, rm, rm -r (-skipTrash), touch, mv)

Summary

- › you can **request meta-information** from Namenode and change its structure (ls, mkdir, rm, rm -r (-skipTrash), touch, mv)
- › you can **read and write** data from and to Datanodes in HDFS (put, cat, head, tail, get, getmerge)

Summary

- › you can **request meta-information** from Namenode and change its structure (ls, mkdir, rm, rm -r (-skipTrash), touch, mv)
- › you can **read and write** data from and to Datanodes in HDFS (put, cat, head, tail, get, getmerge)
- › you can **change replication factor** of files and **get detailed information** about data in HDFS (chown, hdfs groups; setrep; hdfs fsck; find)

BigDATAteam