

1. Запустить задачу из примеров, например, вычисление  $\pi$  методом Монте-Карло  
`YARN_EXAMPLES=/opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar pi 16 1000000`  
 зайти на ResourceManager <http://manager.novalocal:8088>  
 и найти свою задачу.  
 Для этого нужно запустить прокси используя `-D localhost:1080`  
 и настроить браузер на его использование



## FINISHED Applications

Cluster

About Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted

82

Apps Pending

0

Apps Running

2

Apps Completed

80

Containers Running

6

Memory Used

8 GB

Memory Total

28 GB

Memory Rese

0 B

Cluster Nodes Metrics

Active Nodes

4

Decommissioning Nodes

0

Decommissioned Nodes

0

Lost Nodes

0

Unhealthy Nodes

0

Scheduler Metrics

Scheduler Type

Capacity Scheduler

Scheduling Resource Type

[memory-mb (unit=Mi), vcores]

Minimum Allocation

<memory:512, vCores:1>

Maximum Allocation

<memory:3072, vCores:8>

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU Vcores	Alloca Mem MB
application_1612470347159_0087	student1_13	QuasiMonteCarlo	MAPREDUCE	default	0	Sat Feb 6 23:04:47 +0100 2021	Sat Feb 6 23:05:14 +0100 2021	FINISHED	SUCCEEDED	N/A	N/A	N/A
13.1 application_1612470347159_0086	student1_13	QuasiMonteCarlo	MAPREDUCE	default	0	Sat Feb 6 22:52:23 +0100 2021	Sat Feb 6 22:52:23 +0100 2021	FINISHED	SUCCEEDED	N/A	N/A	N/A
application_1612470347159_0080	zeppelin	zeppelin-9996	SPARK	default	0	Sat Feb 6 19:07:02 +0100 2021	Sat Feb 6 19:16:48 +0100 2021	FINISHED	SUCCEEDED	N/A	N/A	N/A
application_1612470347159_0078	student1_7	QuasiMonteCarlo	MAPREDUCE	default	0	Sat Feb 6 17:55:06 +0100 2021	Sat Feb 6 17:56:54 +0100 2021	FINISHED	SUCCEEDED	N/A	N/A	N/A

```
[student1_13@bigdataanalytics-head-0 ~]$ ssh -D localhost:1080 185.241.193.174
Last login: Sat Feb 6 21:44:25 2021 from 89-64-64-122.dynamic.chello.pl
```

```
[student1_13@bigdataanalytics-head-0 ~]$ yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples.jar pi 16 1000000
```

Number of Maps = 16

Samples per Map = 1000000

Wrote input for Map #0

Wrote input for Map #1

Wrote input for Map #2

Wrote input for Map #3

Wrote input for Map #4

Wrote input for Map #5

Wrote input for Map #6

Wrote input for Map #7

Wrote input for Map #8

Wrote input for Map #9

Wrote input for Map #10

Wrote input for Map #11

Wrote input for Map #12

Wrote input for Map #13

Wrote input for Map #14

Wrote input for Map #15

Starting Job

```
21/02/06 21:51:55 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics-head-0.novalocal/10.0.0.8:8050
```

```
21/02/06 21:51:55 INFO client.AHSPProxy: Connecting to Application History server at bigdataanalytics-head-0.novalocal/10.0.0.8:10200
```

```
21/02/06 21:51:56 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /user/student1_13/staging/job_1612470347159_0086
```

```
21/02/06 21:51:56 INFO input.FileInputFormat: Total input files to process : 16
```

```
21/02/06 21:51:57 INFO mapreduce.JobSubmitter: number of splits:16
```

```
21/02/06 21:51:57 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1612470347159_0086
```

```
21/02/06 21:51:57 INFO mapreduce.JobSubmitter: Executing with tokens: []
```

```
21/02/06 21:51:57 INFO conf.Configuration: found resource resource-types.xml at file:/etc/hadoop/3.1.4.0-315/0/resource-types.xml
```

```
21/02/06 21:51:57 INFO impl.YarnClientImpl: Submitted application application_1612470347159_0086
```

```
21/02/06 21:51:57 INFO mapreduce.Job: The url to track the job: http://bigdataanalytics-head-0.novalocal:8088/proxy/application_1612470347159_0086/
```

```
21/02/06 21:51:57 INFO mapreduce.Job: Running job: job_1612470347159_0086
```

```
21/02/06 21:52:04 INFO mapreduce.Job: Job job_1612470347159_0086 running in uber mode : false
```

21/02/06 21:52:04 INFO mapreduce.Job: map 0% reduce 0%  
21/02/06 21:52:11 INFO mapreduce.Job: map 13% reduce 0%  
21/02/06 21:52:12 INFO mapreduce.Job: map 31% reduce 0%  
21/02/06 21:52:12 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0086\_m\_000000\_0, Status : FAILED  
[2021-02-06 21:52:11.028]Exception from container-launch.  
Container id: container\_e25\_1612470347159\_0086\_01\_000003  
Exit code: 1

[2021-02-06 21:52:11.030]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4e00000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 21:52:11.031]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4e00000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

21/02/06 21:52:12 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0086\_m\_000010\_0, Status : FAILED  
[2021-02-06 21:52:11.029]Exception from container-launch.  
Container id: container\_e25\_1612470347159\_0086\_01\_000011  
Exit code: 1

[2021-02-06 21:52:11.052]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4c80000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 21:52:11.053]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4c80000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

21/02/06 21:52:12 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0086\_m\_000007\_0, Status : FAILED  
[2021-02-06 21:52:11.028]Exception from container-launch.  
Container id: container\_e25\_1612470347159\_0086\_01\_000008  
Exit code: 1

[2021-02-06 21:52:11.051]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4a80000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 21:52:11.052]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4a80000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

21/02/06 21:52:14 INFO mapreduce.Job: map 38% reduce 0%  
21/02/06 21:52:15 INFO mapreduce.Job: map 44% reduce 0%  
21/02/06 21:52:16 INFO mapreduce.Job: map 56% reduce 0%  
21/02/06 21:52:16 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0086\_m\_000012\_0, Status : FAILED

[2021-02-06 21:52:15.662]Exception from container-launch.  
Container id: container\_e25\_1612470347159\_0086\_01\_000013  
Exit code: 1

[2021-02-06 21:52:15.664]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b5280000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 21:52:15.665]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b5280000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

21/02/06 21:52:16 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0086\_m\_000015\_0, Status : FAILED  
[2021-02-06 21:52:15.681]Exception from container-launch.  
Container id: container\_e25\_1612470347159\_0086\_01\_000018  
Exit code: 1

[2021-02-06 21:52:15.685]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b5100000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 21:52:15.685]Container exited with a non-zero exit code 1. Error file: prelaunch.err.  
Last 4096 bytes of prelaunch.err :  
Last 4096 bytes of stderr :  
Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b5100000, 901251072, 0)  
failed; error='Cannot allocate memory' (errno=12)

21/02/06 21:52:20 INFO mapreduce.Job: map 81% reduce 0%  
21/02/06 21:52:21 INFO mapreduce.Job: map 88% reduce 0%  
21/02/06 21:52:23 INFO mapreduce.Job: map 100% reduce 29%  
21/02/06 21:52:24 INFO mapreduce.Job: map 100% reduce 100%  
21/02/06 21:52:24 INFO mapreduce.Job: Job job\_1612470347159\_0086 completed successfully  
21/02/06 21:52:24 INFO mapreduce.Job: Counters: 56

#### File System Counters

FILE: Number of bytes read=358  
FILE: Number of bytes written=4023671  
FILE: Number of read operations=0  
FILE: Number of large read operations=0  
FILE: Number of write operations=0  
HDFS: Number of bytes read=4710  
HDFS: Number of bytes written=215  
HDFS: Number of read operations=69  
HDFS: Number of large read operations=0  
HDFS: Number of write operations=3

#### Job Counters

Failed map tasks=5  
Launched map tasks=21  
Launched reduce tasks=1  
Other local map tasks=5  
Data-local map tasks=15  
Rack-local map tasks=1  
Total time spent by all maps in occupied slots (ms)=338055  
Total time spent by all reduces in occupied slots (ms)=42700  
Total time spent by all map tasks (ms)=112685  
Total time spent by all reduce tasks (ms)=10675

Total vcore-milliseconds taken by all map tasks=112685  
Total vcore-milliseconds taken by all reduce tasks=10675  
Total megabyte-milliseconds taken by all map tasks=173084160  
Total megabyte-milliseconds taken by all reduce tasks=21862400

#### Map-Reduce Framework

Map input records=16  
Map output records=32  
Map output bytes=288  
Map output materialized bytes=448  
Input split bytes=2822  
Combine input records=0  
Combine output records=0  
Reduce input groups=2  
Reduce shuffle bytes=448  
Reduce input records=32  
Reduce output records=0  
Spilled Records=64  
Shuffled Maps =16  
Failed Shuffles=0  
Merged Map outputs=16  
GC time elapsed (ms)=2438  
CPU time spent (ms)=25770  
Physical memory (bytes) snapshot=17966301184  
Virtual memory (bytes) snapshot=55552983040  
Total committed heap usage (bytes)=16800284672  
Peak Map Physical memory (bytes)=1124782080  
Peak Map Virtual memory (bytes)=3251855360  
Peak Reduce Physical memory (bytes)=214691840  
Peak Reduce Virtual memory (bytes)=3719766016

#### 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=1888

#### File Output Format Counters

Bytes Written=97

Job Finished in 29.291 seconds

Estimated value of Pi is 3.1415912500000000000000

## 2. Запустить WordCount и найти ошибку

```
yarn jar /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/lib/hadoop-mapreduce/hadoop-streaming.jar  
-input /user/student9_30/text_example.txt -output result -file mapper.py -file reducer.py -mapper "python  
mapper.py" -reducer "python reducer.py"
```

\*mapper.py:

\*

```
import sys  
for line in sys.stdin:  
    for word in line.split(" "):  
        print(word + "\t" + str(1))
```

**#!/usr/bin/env python**

```
import sys
```

```
for line in sys.stdin:  
    line = line.strip()  
    words = line.split()  
    for word in words:  
        print(word + "\t" + 1)
```

reducer.py:

```
import sys
```

```

prevkey = None
cnt = 0
for line in sys.stdin:
    if(len(line.strip().split("\t")) > 1):
        key, value = line.strip().split("\t")
        #print(key + "\t" + value);
    if key != prevkey and prevkey is not None:
        print(prevkey + "\t" + str(cnt))
        cnt = 0
    cnt = cnt + int(value)
    prevkey = key

```

**#!/usr/bin/env python**

```

import sys
from operator import itemgetter
prevkey = None
cnt = 0
for line in sys.stdin:
    if(len(line.strip().split("\t")) > 1):
        key, value = line.strip().split("\t")
        print(key + "\t" + value);
    if key != prevkey and prevkey is not None:
        print(prevkey + "\t" + str(cnt))
        cnt = 0
    cnt = cnt + int(value)
    prevkey = key

```

[student1\_13@bigdataanalytics-head-0 ~]\$ cat > mapper.py

```

#!/usr/bin/env python
import sys

```

```

for line in sys.stdin:
    line = line.strip()
    words = line.split()
    for word in words:
        print(word + "\t" + 1)

```

^C

[student1\_13@bigdataanalytics-head-0 ~]\$ cat mapper.py

```

#!/usr/bin/env python
import sys

```

```

for line in sys.stdin:
    line = line.strip()
    words = line.split()
    for word in words:
        print(word + "\t" + 1)

```

[student1\_13@bigdataanalytics-head-0 ~]\$ cat f1 | python mapper.py

[student1\_13@bigdataanalytics-head-0 ~]\$ cat f1 | python mapper.py | sort

[student1\_13@bigdataanalytics-head-0 ~]\$ cat f1 | python mapper.py | sort | python

[student1\_13@bigdataanalytics-head-0 ~]\$ mv f1 f1.txt

[student1\_13@bigdataanalytics-head-0 ~]\$ ls

f1.txt folder hs\_err\_pid12017.log hs\_err\_pid32716.log mapper.py

f2 hs\_err\_pid1104.log hs\_err\_pid23173.log hs\_err\_pid8728.log

[student1\_13@bigdataanalytics-head-0 ~]\$ hdfs dfs -ls

Found 3 items

```
drwx----- - student1_13 student1_13      0 2021-02-06 22:05 .staging
```

```
-rw-r--r--  2 student1_13 student1_13    10409 2021-02-03 13:25 f1
```

```
drwxr-xr-x - student1_13 student1_13      0 2021-02-06 22:34 test
```

[student1\_13@bigdataanalytics-head-0 ~]\$ hdfs dfs -mv f1 test/.

[student1\_13@bigdataanalytics-head-0 ~]\$ hdfs dfs -ls

Found 2 items

```
drwx----- - student1_13 student1_13      0 2021-02-06 22:05 .staging
```

```
drwxr-xr-x - student1_13 student1_13      0 2021-02-06 22:39 test
```

[student1\_13@bigdataanalytics-head-0 ~]\$ hdfs dfs -ls test

Found 1 items

```
-rw-r--r--  2 student1_13 student1_13    10409 2021-02-03 13:25 test/f1
```

```
ce-examples.jar -input test -output result -mapper "python mapper.py" -file mapper.py
```

```
[student1_13@bigdataanalytics-head-0 ~]$ yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar -input /user/student1_13/f1.txt -output result -file mapper.py -file reducer.py -mapper "python mapper.py" -reducer "python reducer.py"
21/02/06 22:59:00 WARN streaming.StreamJob: -file option is deprecated, please use generic option -files instead.
packageJobJar: [mapper.py, reducer.py] [/usr/hdp/3.1.4.0-315/hadoop-mapreduce/hadoop-streaming-3.1.1.3.1.4.0-315.jar] /var/lib/ambari-agent/tmp/hadoop_java_io_tmpdir/streamjob5993870996408673455.jar tmpDir=null
21/02/06 22:59:02 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics-head-0.novalocal/10.0.0.8:8050
21/02/06 22:59:02 INFO client.AHSPProxy: Connecting to Application History server at bigdataanalytics-head-0.novalocal/10.0.0.8:10200
21/02/06 22:59:02 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics-head-0.novalocal/10.0.0.8:8050
21/02/06 22:59:02 INFO client.AHSPProxy: Connecting to Application History server at bigdataanalytics-head-0.novalocal/10.0.0.8:10200
21/02/06 22:59:02 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /user/student1_13/.staging/job_1612470347159_0088
21/02/06 22:59:03 INFO mapreduce.JobSubmitter: Cleaning up the staging area /user/student1_13/.staging/job_1612470347159_0088
21/02/06 22:59:03 ERROR streaming.StreamJob: Error Launching job : Input path does not exist: hdfs://bigdataanalytics-head-0.novalocal:8020/user/student1_13/f1.txt
Streaming Command Failed!
```

```
[student1_13@bigdataanalytics-head-0 ~]$ yarn jar /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/lib/hadoop-mapreduce/hadoop-streaming.jar -input /user/student1_13/f1.txt -output result -file mapper.py -file reducer.py -mapper "python mapper.py" -reducer "python reducer.py"
JAR does not exist or is not a normal file: /opt/cloudera/parcels/CDH-5.16.2-1.cdh5.16.2.p0.8/lib/hadoop-mapreduce/hadoop-streaming.jar
```

```
[student1_13@bigdataanalytics-head-0 ~]$ yarn jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-streaming.jar -input test -output result -file mapper.py -mapper "python mapper.py"
21/02/06 23:02:15 WARN streaming.StreamJob: -file option is deprecated, please use generic option -files instead.
packageJobJar: [mapper.py] [/usr/hdp/3.1.4.0-315/hadoop-mapreduce/hadoop-streaming-3.1.1.3.1.4.0-315.jar] /var/lib/ambari-agent/tmp/hadoop_java_io_tmpdir/streamjob2469586076949381818.jar tmpDir=null
21/02/06 23:02:16 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics-head-0.novalocal/10.0.0.8:8050
21/02/06 23:02:17 INFO client.AHSPProxy: Connecting to Application History server at bigdataanalytics-head-0.novalocal/10.0.0.8:10200
21/02/06 23:02:17 INFO client.RMProxy: Connecting to ResourceManager at bigdataanalytics-head-0.novalocal/10.0.0.8:8050
21/02/06 23:02:17 INFO client.AHSPProxy: Connecting to Application History server at bigdataanalytics-head-0.novalocal/10.0.0.8:10200
21/02/06 23:02:17 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /user/student1_13/.staging/job_1612470347159_0090
21/02/06 23:02:17 INFO mapred.FileInputFormat: Total input files to process : 1
21/02/06 23:02:17 INFO mapreduce.JobSubmitter: number of splits:2
21/02/06 23:02:18 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1612470347159_0090
21/02/06 23:02:18 INFO mapreduce.JobSubmitter: Executing with tokens: []
21/02/06 23:02:18 INFO conf.Configuration: found resource resource-types.xml at file:/etc/hadoop/3.1.4.0-315/0/resource-types.xml
21/02/06 23:02:18 INFO impl.YarnClientImpl: Submitted application application_1612470347159_0090
21/02/06 23:02:18 INFO mapreduce.Job: The url to track the job: http://bigdataanalytics-head-0.novalocal:8088/proxy/application_1612470347159_0090/
21/02/06 23:02:18 INFO mapreduce.Job: Running job: job_1612470347159_0090
21/02/06 23:02:27 INFO mapreduce.Job: Job job_1612470347159_0090 running in uber mode : false
21/02/06 23:02:27 INFO mapreduce.Job: map 0% reduce 0%
21/02/06 23:02:32 INFO mapreduce.Job: Task Id : attempt_1612470347159_0090_m_000001_0, Status : FAILED
Error: java.lang.RuntimeException: PipeMapRed.waitOutputThreads(): subprocess failed with code 1
    at org.apache.hadoop.streaming.PipeMapRed.waitOutputThreads(PipeMapRed.java:325)
    at org.apache.hadoop.streaming.PipeMapRed.mapRedFinished(PipeMapRed.java:538)
    at org.apache.hadoop.streaming.PipeMapper.close(PipeMapper.java:130)
    at org.apache.hadoop.mapred.MapRunner.run(MapRunner.java:61)
    at org.apache.hadoop.streaming.PipeMapRunner.run(PipeMapRunner.java:34)
    at org.apache.hadoop.mapred.MapTask.runOldMapper(MapTask.java:465)
    at org.apache.hadoop.mapred.MapTask.run(MapTask.java:349)
```

```
at org.apache.hadoop.mapred.YarnChild$2.run(YarnChild.java:174)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)
at org.apache.hadoop.mapred.YarnChild.main(YarnChild.java:168)
```

21/02/06 23:02:33 INFO mapreduce.Job: map 50% reduce 0%

21/02/06 23:02:37 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0090\_m\_000001\_1, Status : FAILED

Error: java.lang.RuntimeException: PipeMapRed.waitOutputThreads(): subprocess failed with code 1

```
at org.apache.hadoop.streaming.PipeMapRed.waitOutputThreads(PipeMapRed.java:325)
at org.apache.hadoop.streaming.PipeMapRed.mapRedFinished(PipeMapRed.java:538)
at org.apache.hadoop.streaming.PipeMapper.close(PipeMapper.java:130)
at org.apache.hadoop.mapred.MapRunner.run(MapRunner.java:61)
at org.apache.hadoop.streaming.PipeMapRunner.run(PipeMapRunner.java:34)
at org.apache.hadoop.mapred.MapTask.runOldMapper(MapTask.java:465)
at org.apache.hadoop.mapred.MapTask.run(MapTask.java:349)
at org.apache.hadoop.mapred.YarnChild$2.run(YarnChild.java:174)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:422)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1730)
at org.apache.hadoop.mapred.YarnChild.main(YarnChild.java:168)
```

21/02/06 23:02:43 INFO mapreduce.Job: map 50% reduce 17%

21/02/06 23:02:44 INFO mapreduce.Job: Task Id : attempt\_1612470347159\_0090\_m\_000001\_2, Status : FAILED

[2021-02-06 23:02:44.232]Exception from container-launch.

Container id: container\_e25\_1612470347159\_0090\_01\_000006

Exit code: 1

[2021-02-06 23:02:44.244]Container exited with a non-zero exit code 1. Error file: prelaunch.err.

Last 4096 bytes of prelaunch.err :

Last 4096 bytes of stderr :

Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4e80000, 901251072, 0) failed; error='Cannot allocate memory' (errno=12)

[2021-02-06 23:02:44.245]Container exited with a non-zero exit code 1. Error file: prelaunch.err.

Last 4096 bytes of prelaunch.err :

Last 4096 bytes of stderr :

Java HotSpot(TM) 64-Bit Server VM warning: INFO: os::commit\_memory(0x00000000b4e80000, 901251072, 0) failed; error='Cannot allocate memory' (errno=12)

21/02/06 23:02:51 INFO mapreduce.Job: map 100% reduce 100%

21/02/06 23:02:52 INFO mapreduce.Job: Job job\_1612470347159\_0090 failed with state FAILED due to: Task failed task\_1612470347159\_0090\_m\_000001

Job failed as tasks failed. failedMaps:1 failedReduces:0 killedMaps:0 killedReduces: 0

21/02/06 23:02:53 INFO mapreduce.Job: Counters: 41

File System Counters

```
FILE: Number of bytes read=0
FILE: Number of bytes written=238822
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=5327
HDFS: Number of bytes written=0
HDFS: Number of read operations=3
HDFS: Number of large read operations=0
HDFS: Number of write operations=0
```

Job Counters

```
Failed map tasks=4
Killed reduce tasks=1
Launched map tasks=5
Launched reduce tasks=1
Other local map tasks=3
Data-local map tasks=2
```

Total time spent by all maps in occupied slots (ms)=64380  
Total time spent by all reduces in occupied slots (ms)=64000  
Total time spent by all map tasks (ms)=21460  
Total time spent by all reduce tasks (ms)=16000  
Total vcore-milliseconds taken by all map tasks=21460  
Total vcore-milliseconds taken by all reduce tasks=16000  
Total megabyte-milliseconds taken by all map tasks=32962560  
Total megabyte-milliseconds taken by all reduce tasks=32768000

#### Map-Reduce Framework

Map input records=175  
Map output records=0  
Map output bytes=0  
Map output materialized bytes=6  
Input split bytes=122  
Combine input records=0  
Spilled Records=0  
Failed Shuffles=0  
Merged Map outputs=0  
GC time elapsed (ms)=59  
CPU time spent (ms)=1770  
Physical memory (bytes) snapshot=1110761472  
Virtual memory (bytes) snapshot=3242872832  
Total committed heap usage (bytes)=1022361600  
Peak Map Physical memory (bytes)=1110761472  
Peak Map Virtual memory (bytes)=3242872832

#### File Input Format Counters

Bytes Read=5205

21/02/06 23:02:53 ERROR streaming.StreamJob: Job not successful!  
Streaming Command Failed!

1. Со звездочкой! Написать join для двух файлов через mapreduce