

EXP NO: 4**CREATE UDF IN PIG****\$start-all.sh**

```

(hadoop@kali)~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [kali]
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-11 04:59:16,429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers

```

\$ jps

```

(hadoop@kali)~$ jps
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
14436 NodeManager
16772 Jps
13830 SecondaryNameNode
14311 ResourceManager
13597 DataNode
13471 NameNode

```

\$wget <https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz>**\$ tar xvfz pig-0.16.0.tar.gz**

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
(hadoop@kali)~$ wget https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
--2024-08-29 10:55:33-- https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
Resolving dlcdn.apache.org (dlcdn.apache.org) ... 208.146.2.254, 151.161.2.132
Connecting to dlcdn.apache.org (dlcdn.apache.org)|208.146.2.254|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 177279333 (169M) [application/x-gzip]
Saving to: 'pig-0.16.0.tar.gz'

pig-0.16.0.tar.gz 100%[====>] 169.87M 26.5MB/s in 6.5s

2024-08-29 10:55:41 (26.0 MB/s) - 'pig-0.16.0.tar.gz' saved [177279333/177279333]

(hadoop@kali)~$ tar xvfz pig-0.16.0.tar.gz
pig-0.16.0/
pig-0.16.0/bin/
pig-0.16.0/conf/
pig-0.16.0/contrib/
pig-0.16.0/contrib/piggybank/
pig-0.16.0/contrib/piggybank/java/
pig-0.16.0/contrib/piggybank/java/build/
pig-0.16.0/contrib/piggybank/java/build/classes/
pig-0.16.0/contrib/piggybank/java/build/classes/org/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/convert/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/convert/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/truncate/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/decode/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/math/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/stats/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/string/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/util/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/util/apachelogarser/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/par/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/allloader/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/apachelog/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/avro/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/hivert/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/partition/
pig-0.16.0/contrib/piggybank/java/build/docs/
pig-0.16.0/contrib/piggybank/java/build/docs/api/
pig-0.16.0/contrib/piggybank/java/build/test/
pig-0.16.0/contrib/piggybank/java/build/test/classes/
pig-0.16.0/contrib/piggybank/java/lib/
pig-0.16.0/contrib/piggybank/java/src/

```

\$nano ~/.bashrc

```

#PIG settings
export PIG_HOME=/home/hadoop/pig
export PATH=$PATH:$PIG_HOME/bin
export PIG_CLASSPATH=$PIG_HOME/conf:$HADOOP_HOME/etc/hadoop/
export PIG_CONF_DIR=$PIG_HOME/conf
#export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export PIG_CLASSPATH=$PIG_CONF_DIR:$PATH
#PIG setting ends

```

\$mv pig-0.16.0 pig

\$pig

The image shows a Kali Linux terminal window with a dark theme. At the top, the terminal title is 'kali-linux-2024.4-vmware-and4 - VMware Workstation 12 Player (Non-commercial use only)'. The terminal content shows the following sequence of commands and outputs:

1. Initial state: The prompt is 'hadoop@kali: ~'.
2. Command: `pip-0.10.0/tutorial/src/org/apache/pig/tutorial/TutorialUtil.java`
3. Command: `pip-0.10.0/bin/pig`
4. Command: `pip-0.10.0/bin/pig.cmd`
5. Command: `pip-0.10.0/bin/pig.py`
6. Command: `(hadoop@kali) [-] # ls`
7. Output: A directory listing showing files like 'hadoop', 'hadoop-3.6.0.tar.gz', 'hadoopdata', 'Music', 'Pictures', 'pig-0.10.0', 'pig-0.10.0.tar.gz', 'Public', 'Templates', and 'Videos'.
8. Command: `(hadoop@kali) [-] # mv pig-0.10.0 pig`
9. Command: `(hadoop@kali) [-] # ls`
10. Command: `(hadoop@kali) [-] # h name -f .bashrc`
11. Command: `(hadoop@kali) [-] # source ~/.bashrc`
12. Command: `(hadoop@kali) [-] # java`
13. Output: A long list of Java classpaths including `JAVA_OPTS: -Dawt.useSystemAAFontSettings=-on -Dswing.aatext=true`, `2024. ResourceManager`, `34624 SecondaryNameNode`, `51338 JobManager`, `60393 Database`, `64993 Jps`, and `65227 NameNode`.
14. Command: `(hadoop@kali) [-] # pig`
15. Output: `Picked up JAVA_OPTS: -Dawt.useSystemAAFontSettings=-on -Dswing.aatext=true`
16. Command: `(hadoop@kali) [-] # pig`
17. Output: A series of log messages from the pig script, including `INFO pig.CeeTypeProvider: Trying CeeTypeProvider: LOCAL`, `INFO org.apache.pig.Main - Apache pig version 0.16.0 (r1746510) compiled Jun 01 2016, 23:10:49`, and `WARN org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use fs.defaultFS`.
18. Command: `(hadoop@kali) [-] #`
19. Output: `org.apache.pig.Main - Pig script completed in 15 seconds and 392 milliseconds (15392 ms)`
20. Command: `(hadoop@kali) [-] #`
21. Output: `(hadoop@kali) [-] #`
The terminal window also shows a file manager in the background, displaying the same directory listing as the terminal output.

\$cd DA-Lab

\$mkdir exp4

\$cd exp4

\$nano sample.txt

```
File Actions Edit View Help
GNU nano 7.2 sample.txt
1,John
2,Jane
3,Joe
4,Emma
```

\$nano demo_pig.pig

```

File Actions Edit View Help
GNU nano 7.2 demo_pig.pig
-- Load the data from HDFS
data = LOAD '/exp4/sample.txt' USING PigStorage(',') AS (id:int, name:chararray);
-- Dump the data to check if it was loaded correctly
DUMP data;

```

```
$hdfs dfs -mkdir /exp4
```

```
$hdfs dfs -copyFromLocal ~/DA-Lab/exp4/sample.txt /exp4
```

\$pig demo_pig.pig

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
File Actions Edit View Help
[hadoop@kali:~/DA-Lab/exp4]
$ nano demo_pig.pig
[hadoop@kali:~/DA-Lab/exp4]
$ pig demo_pig.pig
Picked up JAVA_OPTIONS: -Dant.useSystemAfontSettings=on -Dswing.aatext=true
2024-08-30 04:56:11.449 [main] INFO org.apache.pig.ExecTypeProvider: Trying ExecType: LOCAL
2024-08-30 04:56:11.456 [main] INFO org.apache.pig.ExecTypeProvider: Trying ExecType: MAPREDUCE
2024-08-30 04:56:11.457 [main] INFO org.apache.pig.ExecTypeProvider: Picked MAPREDUCE as the execType
2024-08-30 04:56:11.680 [main] INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 22:10:49
2024-08-30 04:56:11.680 [main] INFO org.apache.pig.Main - Logging error messages to: /home/hadoop/DA-Lab/exp4/pig-1725000173602.log
2024-08-30 04:56:11.349 [main] WARN org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2024-08-30 04:56:11.547 [main] INFO org.apache.pig.impl.util.Utils - Default houtput file /home/hadoop/pigbootstrap not found
2024-08-30 04:56:11.680 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.job.tracker.address
2024-08-30 04:56:11.680 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:11.881 [main] INFO org.apache.pig.backend.hadoop.executionengine.MExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2024-08-30 04:56:17.178 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:17.238 [main] INFO org.apache.pig.PigServer - AFS is disabled since yarn.timeline-service.enabled set to false
2024-08-30 04:56:19.209 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:20.129 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2024-08-30 04:56:20.231 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:20.968 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.MultiQueryOptimizer - MR plan size before optimization: 1
2024-08-30 04:56:20.431 [main] INFO org.apache.pig.newplan.logical.optimizer.logicalPlanOptimizer - [RM45_ENABLED[AddressCh, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergedFilter, MergeJoinCh, PartitionFilterOptimizer, PredicatePushDownOptimizer, PushDownForCachingJoin, PushDownFilter, SplitFilter, StreamTypeCastInserter]]
2024-08-30 04:56:20.671 [main] INFO org.apache.pig.impl.util.LimitedMemoryManager - Selected heap (PS Old Gen) of size 609138048 to monitor. collectionUsageThreshold = 489396648, usageThreshold = 489396640
2024-08-30 04:56:20.680 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.MRCompiler - File concatenation threshold: 180 optimistic? false
2024-08-30 04:56:20.998 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.MultiQueryOptimizer - MR plan size after optimization: 1
2024-08-30 04:56:21.002 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.submit.realization is deprecated. Instead, use mapreduce.client.submit.file.replication
2024-08-30 04:56:21.357 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/automaton-1.11-6.jar to DistributedCache through /tmp/tempt-1575302697/tmp1364382
2024-08-30 04:56:22.182 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/antlr-runtime-3.4.jar to DistributedCache through /tmp/tempt-1575302697/tmp102799
2024-08-30 04:56:22.213 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/joda-time-2.9.3.jar to DistributedCache through /tmp/tempt-1575302697/tmp2022832
2024-08-30 04:56:22.255 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Setting up single store job
2024-08-30 04:56:22.274 [main] INFO org.apache.pig.data.SchemaInputFrontend - Key [pig.schemaInput] is false, will not generate code
2024-08-30 04:56:22.648 [main] INFO org.apache.pig.data.SchemaInputFrontend - Starting process to move generated code to distributed cache
2024-08-30 04:56:24.588 [main] INFO org.apache.pig.data.SchemaInputFrontend - Setting key [pig.schemaInputClasses] with classes to serialize: []
2024-08-30 04:56:24.768 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.MapReduceLauncher - 1 map-reduce job(s) waiting for submission.
R/pig-9.16.0-core-32.jar
2024-08-30 04:56:25.754 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/automaton-1.11-6.jar to DistributedCache through /tmp/tempt-1575302697/tmp1364382
2024-08-30 04:56:25.839 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/antlr-runtime-3.4.jar to DistributedCache through /tmp/tempt-1575302697/tmp102799
2024-08-30 04:56:25.847 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Added jar file:/home/hadoop/pig/lib/joda-time-2.9.3.jar to DistributedCache through /tmp/tempt-1575302697/tmp2022832
2024-08-30 04:56:26.486 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreducelayer.JobControlCompiler - Setting up single store job
2024-08-30 04:56:26.543 [main] INFO org.apache.pig.data.SchemaInputFrontend - Key [pig.schemaInput] is false, will not generate code
2024-08-30 04:56:26.543 [main] INFO org.apache.pig.data.SchemaInputFrontend - Starting process to move generated code to distributed cache
2024-08-30 04:56:29.588 [main] INFO org.apache.pig.data.SchemaInputFrontend - Setting key [pig.schemaInputClasses] with classes to serialize: []
2024-08-30 04:56:29.677 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapReduceUtil - Total input paths to process: 1
(1,John)
(2,James)
(3,Joe)
(4,Anna)
2024-08-30 05:00:00.174 [main] INFO org.apache.pig.Main - Pig script completed in 3 minutes, 46 seconds and 870 milliseconds (226878 ms)
[hadoop@kali:~/DA-Lab/exp4]

```

\$nano uppercase_udf.py

```

hadoop@kali: ~/DA-Lab/exp4
File Actions Edit View Help
GNU nano 7.2
uppercase_udf.py
def uppercase(text):
    return text.upper()
if __name__ == "__main__":
    import sys
    for line in sys.stdin:
        line = line.strip()
        result = uppercase(line)
        print(result)

```


\$hdfs dfs -copyFromLocal ~/DA-Lab/exp4/uppercase_udf.py /exp4

```
(hadoop@kali)-[~/hadoop/bin]
$ ./hdfs dfs -ls /exp4
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:26:01,736 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
Found 3 items
drwxr-xr-x   - hadoop supergroup          0 2024-08-30 05:07 /exp4/output
-rw-r--r--   1 hadoop supergroup        27 2024-08-30 04:43 /exp4/sample.txt
-rw-r--r--   1 hadoop supergroup       172 2024-08-30 05:02 /exp4/uppercase_udf.py
```

\$nano udf_example.pig

```
hadoop@kali: ~/DA-Lab/exp4
File Actions Edit View Help
GNU nano 7.2 udf_example.pig
-- Register the Python UDF script
REGISTER 'hdfs:///exp4/uppercase_udf.py' USING jython AS udf;
-- Load some data
data = LOAD 'hdfs:///exp4/sample.txt' AS (text:chararray);
-- Use the Python UDF
uppercased_data = FOREACH data GENERATE udf.uppercase(text) AS uppercase_text;
-- Store the result
STORE uppercased_data INTO 'hdfs:///exp4/output';
```

\$pig -f udf_example.pig

```
kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
hadoop@kali: ~/DA-Lab/exp4
$ pig -f udf_example.pig
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:08:18,591 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
2024-08-30 05:08:18,601 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
2024-08-30 05:08:18,601 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
2024-08-30 05:08:18,606 [main] INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746536) compiled Jun 01 2016, 23:10:49
2024-08-30 05:08:18,606 [main] INFO org.apache.pig.Main - Logging error messages to: /home/hadoop/DA-Lab/exp4/pig-172500778010.log
2024-08-30 05:08:19,685 [main] WARN org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2024-08-30 05:08:20,127 [main] INFO org.apache.pig.impl.util.Utils - Default bootstrap file /home/hadoop/pigbootstrap not found
2024-08-30 05:08:20,382 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2024-08-30 05:08:20,383 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:20,383 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2024-08-30 05:08:22,438 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:22,513 [main] WARN org.apache.pig.PigServer - Pig Script ID for the session: PIG-UDF-Example-pig-172500778010-4763-90b-5033b6d33a1
2024-08-30 05:08:22,513 [main] WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false
2024-08-30 05:08:22,725 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:24,298 [main] INFO org.apache.pig.scripting.jython.JythonScriptEngine - created tmp python.cachedir/tmp/pig.jython.3781842658994475167
2024-08-30 05:08:26,818 [main] WARN org.apache.pig.scripting.jython.JythonScriptEngine - pig.cmd.args.reminders is empty. This is not expected unless on testing.
2024-08-30 05:08:26,844 [main] INFO org.apache.pig.scripting.jython.JythonScriptEngine - Register scripting UDF: udf.uppercase
2024-08-30 05:08:37,764 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:37,807 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:38,407 [main] INFO org.apache.pig.scripting.jython.JythonFunction - No schema defined for function uppercase in /tmp/pig3133262932933085tmp/uppercase_udf.py
2024-08-30 05:08:38,583 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:38,742 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.textoutputformat.separator is deprecated. Instead, use mapreduce.output.textoutputformat.separator
2024-08-30 05:08:38,836 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2024-08-30 05:08:38,958 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:08:39,382 [main] INFO org.apache.pig.data.SchemaTupleBackend - key (pig.schemaTuple) was not set... will not generate code.
2024-08-30 05:08:39,332 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - [RULES_ENABLED:AdaptorEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, M
ergeFilter, MergeSortEach, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownWorkForEachPlaten, PushUpFilter, SplitFilter, StreamTypeCastInserter]]
2024-08-30 05:08:39,443 [main] INFO org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size 692318840 to monitor, collectionUsageThreshold = 489396640, usageThreshold = 489396640
2024-08-30 05:08:39,684 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2024-08-30 05:08:39,858 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2024-08-30 05:08:39,859 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2024-08-30 05:08:39,994 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
hadoop@kali:~/hadoop/bin
2024-08-30 05:10:12,729 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 4 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:13,731 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 5 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:14,733 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 6 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:15,735 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:16,738 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:17,740 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:17,859 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCESSFUL. Redirecting to job history server
2024-08-30 05:10:18,862 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 10 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:19,864 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 1 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:20,866 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 2 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:21,868 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 3 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:22,871 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 4 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:23,874 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 5 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:24,876 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 6 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:25,879 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:26,881 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:27,883 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECO
NDS)
2024-08-30 05:10:27,987 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MapReduceLauncher - Unable to retrieve job to compute warning aggregation.
2024-08-30 05:10:27,988 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MapReduceLauncher - Success!
2024-08-30 05:10:28,108 [main] INFO org.apache.pig.Main - Pig script completed in 4 minutes, 9 seconds and 747 milliseconds (249747 ms)

(hadoop@kali) [~/DA-Lab/exp4]
$ cd ../hadoop/bin

(hadoop@kali) [~/hadoop/bin]
$ ./hdfs dfs -cat /exp4/output/part-r-00000
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:12:25,908 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1,JOHN
2,JANE
3,JOE
4,EMMA

(hadoop@kali) [~/hadoop/bin]

```

\$hdfs dfs -cat /exp4/output/*

```

(hadoop@kali) [~/hadoop/bin]
$ ./hdfs dfs -cat /exp4/output/*
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:33:32,731 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
1,JOHN
2,JANE
3,JOE
4,EMMA

```