Windows PowerShell

Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\WINDOWS\system32> docker container run -it -p 9870:9870 -p 8088:8088 registry.gitlab.com/training-support/training-environments:hadoop-v1 bash

docker: Error response from daemon: driver failed programming external connectivity on endpoint funny\_mccarthy (87ac689df38c12ae082cf375a843b105b2455f56018f8ac72bd20c828bc66428): Bind for 0.0.0.0:9870 failed: port is already allocated.

PS C:\WINDOWS\system32> docker container run -it -p 9870:9870 -p 8088:8088 registry.gitlab.com/training-support/training-environments:hadoop-v1 bash

/

\* Starting OpenBSD Secure Shell server sshd [ OK ]

Waiting for hdfs to exit from safemode

Safe mode is OFF

Started

root@26c5e0e439bc:/#

root@26c5e0e439bc:/# whoami

root

root@26c5e0e439bc:/# hdfs dfs -ls /

Found 1 items

drwxr-xr-x - root supergroup 0 2021-08-23 10:43 /user

root@26c5e0e439bc:/# hdfs dfs -ls /user

Found 1 items

drwxr-xr-x - root supergroup 0 2021-08-23 10:44 /user/root

root@26c5e0e439bc:/# cd /user

bash: cd: /user: No such file or directory

root@26c5e0e439bc:/# hdfs dfs -mkdir /user/clekha

root@26c5e0e439bc:/# hdfs dfs -ls /clekha

ls: `/clekha': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -ls /user

Found 2 items

drwxr-xr-x - root supergroup 0 2022-03-28 14:06 /user/clekha

drwxr-xr-x - root supergroup 0 2021-08-23 10:44 /user/root

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

root@26c5e0e439bc:/# hdfs dfs touchz/user/clekha/myfile.txt

touchz/user/clekha/myfile.txt: Unknown command

Usage: hadoop fs [generic options]

[-appendToFile <localsrc> ... <dst>]

[-cat [-ignoreCrc] <src> ...]

[-checksum [-v] <src> ...]

[-chgrp [-R] GROUP PATH...]

[-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]

[-chown [-R] [OWNER][:[GROUP]] PATH...]

[-concat <target path> <src path> <src path> ...]

[-copyFromLocal [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-copyToLocal [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-count [-q] [-h] [-v] [-t [<storage type>]] [-u] [-x] [-e] [-s] <path> ...]

[-cp [-f] [-p | -p[topax]] [-d] <src> ... <dst>]

[-createSnapshot <snapshotDir> [<snapshotName>]]

[-deleteSnapshot <snapshotDir> <snapshotName>]

[-df [-h] [<path> ...]]

[-du [-s] [-h] [-v] [-x] <path> ...]

[-expunge [-immediate] [-fs <path>]]

[-find <path> ... <expression> ...]

[-get [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-getfacl [-R] <path>]

[-getfattr [-R] {-n name | -d} [-e en] <path>]

[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]

[-head <file>]

[-help [cmd ...]]

[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] [<path> ...]]

[-mkdir [-p] <path> ...]

[-moveFromLocal [-f] [-p] [-l] [-d] <localsrc> ... <dst>]

[-moveToLocal <src> <localdst>]

[-mv <src> ... <dst>]

[-put [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-renameSnapshot <snapshotDir> <oldName> <newName>]

[-rm [-f] [-r|-R] [-skipTrash] [-safely] <src> ...]

[-rmdir [--ignore-fail-on-non-empty] <dir> ...]

[-setfacl [-R] [{-b|-k} {-m|-x <acl\_spec>} <path>]|[--set <acl\_spec> <path>]]

[-setfattr {-n name [-v value] | -x name} <path>]

[-setrep [-R] [-w] <rep> <path> ...]

[-stat [format] <path> ...]

[-tail [-f] [-s <sleep interval>] <file>]

[-test -[defswrz] <path>]

[-text [-ignoreCrc] <src> ...]

[-touch [-a] [-m] [-t TIMESTAMP (yyyyMMdd:HHmmss) ] [-c] <path> ...]

[-touchz <path> ...]

[-truncate [-w] <length> <path> ...]

[-usage [cmd ...]]

Generic options supported are:

-conf <configuration file> specify an application configuration file

-D <property=value> define a value for a given property

-fs <file:///|hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.

-jt <local|resourcemanager:port> specify a ResourceManager

-files <file1,...> specify a comma-separated list of files to be copied to the map reduce cluster

-libjars <jar1,...> specify a comma-separated list of jar files to be included in the classpath

-archives <archive1,...> specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:

command [genericOptions] [commandOptions]

root@26c5e0e439bc:/# hdfs dfs -touchz/user/clekha/myfile.txt

-touchz/user/clekha/myfile.txt: Unknown command

Usage: hadoop fs [generic options]

[-appendToFile <localsrc> ... <dst>]

[-cat [-ignoreCrc] <src> ...]

[-checksum [-v] <src> ...]

[-chgrp [-R] GROUP PATH...]

[-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]

[-chown [-R] [OWNER][:[GROUP]] PATH...]

[-concat <target path> <src path> <src path> ...]

[-copyFromLocal [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-copyToLocal [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-count [-q] [-h] [-v] [-t [<storage type>]] [-u] [-x] [-e] [-s] <path> ...]

[-cp [-f] [-p | -p[topax]] [-d] <src> ... <dst>]

[-createSnapshot <snapshotDir> [<snapshotName>]]

[-deleteSnapshot <snapshotDir> <snapshotName>]

[-df [-h] [<path> ...]]

[-du [-s] [-h] [-v] [-x] <path> ...]

[-expunge [-immediate] [-fs <path>]]

[-find <path> ... <expression> ...]

[-get [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-getfacl [-R] <path>]

[-getfattr [-R] {-n name | -d} [-e en] <path>]

[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]

[-head <file>]

[-help [cmd ...]]

[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] [<path> ...]]

[-mkdir [-p] <path> ...]

[-moveFromLocal [-f] [-p] [-l] [-d] <localsrc> ... <dst>]

[-moveToLocal <src> <localdst>]

[-mv <src> ... <dst>]

[-put [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-renameSnapshot <snapshotDir> <oldName> <newName>]

[-rm [-f] [-r|-R] [-skipTrash] [-safely] <src> ...]

[-rmdir [--ignore-fail-on-non-empty] <dir> ...]

[-setfacl [-R] [{-b|-k} {-m|-x <acl\_spec>} <path>]|[--set <acl\_spec> <path>]]

[-setfattr {-n name [-v value] | -x name} <path>]

[-setrep [-R] [-w] <rep> <path> ...]

[-stat [format] <path> ...]

[-tail [-f] [-s <sleep interval>] <file>]

[-test -[defswrz] <path>]

[-text [-ignoreCrc] <src> ...]

[-touch [-a] [-m] [-t TIMESTAMP (yyyyMMdd:HHmmss) ] [-c] <path> ...]

[-touchz <path> ...]

[-truncate [-w] <length> <path> ...]

[-usage [cmd ...]]

Generic options supported are:

-conf <configuration file> specify an application configuration file

-D <property=value> define a value for a given property

-fs <file:///|hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.

-jt <local|resourcemanager:port> specify a ResourceManager

-files <file1,...> specify a comma-separated list of files to be copied to the map reduce cluster

-libjars <jar1,...> specify a comma-separated list of jar files to be included in the classpath

-archives <archive1,...> specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:

command [genericOptions] [commandOptions]

root@26c5e0e439bc:/# hdfs dfs -touchz /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 1 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user

Found 2 items

drwxr-xr-x - root supergroup 0 2022-03-28 14:10 /user/clekha

drwxr-xr-x - root supergroup 0 2021-08-23 10:44 /user/root

root@26c5e0e439bc:/# touch input.txt

root@26c5e0e439bc:/# hdfs dfs -copyFromLocal ./input.txt /user/clekha

root@26c5e0e439bc:/# hdfs -dfs -ls /user/clekha

/usr/local/hadoop/libexec/hadoop-functions.sh: line 2425: HDFS\_-DFS\_USER: bad substitution

/usr/local/hadoop/libexec/hadoop-functions.sh: line 2386: HDFS\_-DFS\_USER: bad substitution

ERROR: -dfs is not COMMAND nor fully qualified CLASSNAME.

Usage: hdfs [OPTIONS] SUBCOMMAND [SUBCOMMAND OPTIONS]

OPTIONS is none or any of:

--buildpaths attempt to add class files from

build tree

--config dir Hadoop config directory

--daemon (start|status|stop) operate on a daemon

--debug turn on shell script debug mode

--help usage information

--hostnames list[,of,host,names] hosts to use in worker mode

--hosts filename list of hosts to use in worker mode

--loglevel level set the log4j level for this

command

--workers turn on worker mode

SUBCOMMAND is one of:

Admin Commands:

cacheadmin configure the HDFS cache

crypto configure HDFS encryption zones

debug run a Debug Admin to execute HDFS debug commands

dfsadmin run a DFS admin client

dfsrouteradmin manage Router-based federation

ec run a HDFS ErasureCoding CLI

fsck run a DFS filesystem checking utility

haadmin run a DFS HA admin client

jmxget get JMX exported values from NameNode or

DataNode.

oev apply the offline edits viewer to an edits file

oiv apply the offline fsimage viewer to an fsimage

oiv\_legacy apply the offline fsimage viewer to a legacy

fsimage

storagepolicies list/get/set/satisfyStoragePolicy block storage

policies

Client Commands:

classpath prints the class path needed to get the hadoop

jar and the required libraries

dfs run a filesystem command on the file system

envvars display computed Hadoop environment variables

fetchdt fetch a delegation token from the NameNode

getconf get config values from configuration

groups get the groups which users belong to

lsSnapshottableDir list all snapshottable dirs owned by the current

user

snapshotDiff diff two snapshots of a directory or diff the

current directory contents with a snapshot

version print the version

Daemon Commands:

balancer run a cluster balancing utility

datanode run a DFS datanode

dfsrouter run the DFS router

diskbalancer Distributes data evenly among disks on a given

node

httpfs run HttpFS server, the HDFS HTTP Gateway

journalnode run the DFS journalnode

mover run a utility to move block replicas across

storage types

namenode run the DFS namenode

nfs3 run an NFS version 3 gateway

portmap run a portmap service

secondarynamenode run the DFS secondary namenode

sps run external storagepolicysatisfier

zkfc run the ZK Failover Controller daemon

SUBCOMMAND may print help when invoked w/o parameters or with -h.

root@26c5e0e439bc:/# hdfs dfs -ls /user/lekha

ls: `/user/lekha': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 2 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -get /user/clekha/myfile.txt

root@26c5e0e439bc:/# ls

bin dev input.txt media myfile.txt root srv usr

boot etc lib metastore\_db opt run sys var

derby.log home lib64 mnt proc sbin tmp

root@26c5e0e439bc:/# -ls

bash: -ls: command not found

root@26c5e0e439bc:/# hdfs dfs -ls /uer/clekha

ls: `/uer/clekha': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 2 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -touchz /user/clekha/file1.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 3 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file1.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -cp /user/clekha/file1.txt /user/clekha/file2.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 4 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file1.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:24 /user/clekha/file2.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -mv /user/clekha/file1.txt /user/clekha/file3.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 4 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:24 /user/clekha/file2.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file3.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -touchz /user/clekha/fileremove.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 5 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:24 /user/clekha/file2.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file3.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:28 /user/clekha/fileremove.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -rm /user/clekha/fileremove.txt

Deleted /user/clekha/fileremove.txt

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 4 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:24 /user/clekha/file2.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file3.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# hdfs dfs -touchz /user/clekha/activityfile.txt

root@26c5e0e439bc:/# vim activityfile.txt

root@26c5e0e439bc:/# vim activityfile.txt

root@26c5e0e439bc:/# vim activityfile.txt

root@26c5e0e439bc:/# hdfs dfs -touchz /user/clekha/csvFile.csv

root@26c5e0e439bc:/# vim csvFile.csv

root@26c5e0e439bc:/# vim csvFile.csv

root@26c5e0e439bc:/# hdfs dfs -mkdir /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -put./csvFile.csv/user/root/textData

-put./csvFile.csv/user/root/textData: Unknown command

Usage: hadoop fs [generic options]

[-appendToFile <localsrc> ... <dst>]

[-cat [-ignoreCrc] <src> ...]

[-checksum [-v] <src> ...]

[-chgrp [-R] GROUP PATH...]

[-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]

[-chown [-R] [OWNER][:[GROUP]] PATH...]

[-concat <target path> <src path> <src path> ...]

[-copyFromLocal [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-copyToLocal [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-count [-q] [-h] [-v] [-t [<storage type>]] [-u] [-x] [-e] [-s] <path> ...]

[-cp [-f] [-p | -p[topax]] [-d] <src> ... <dst>]

[-createSnapshot <snapshotDir> [<snapshotName>]]

[-deleteSnapshot <snapshotDir> <snapshotName>]

[-df [-h] [<path> ...]]

[-du [-s] [-h] [-v] [-x] <path> ...]

[-expunge [-immediate] [-fs <path>]]

[-find <path> ... <expression> ...]

[-get [-f] [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]

[-getfacl [-R] <path>]

[-getfattr [-R] {-n name | -d} [-e en] <path>]

[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]

[-head <file>]

[-help [cmd ...]]

[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] [<path> ...]]

[-mkdir [-p] <path> ...]

[-moveFromLocal [-f] [-p] [-l] [-d] <localsrc> ... <dst>]

[-moveToLocal <src> <localdst>]

[-mv <src> ... <dst>]

[-put [-f] [-p] [-l] [-d] [-t <thread count>] <localsrc> ... <dst>]

[-renameSnapshot <snapshotDir> <oldName> <newName>]

[-rm [-f] [-r|-R] [-skipTrash] [-safely] <src> ...]

[-rmdir [--ignore-fail-on-non-empty] <dir> ...]

[-setfacl [-R] [{-b|-k} {-m|-x <acl\_spec>} <path>]|[--set <acl\_spec> <path>]]

[-setfattr {-n name [-v value] | -x name} <path>]

[-setrep [-R] [-w] <rep> <path> ...]

[-stat [format] <path> ...]

[-tail [-f] [-s <sleep interval>] <file>]

[-test -[defswrz] <path>]

[-text [-ignoreCrc] <src> ...]

[-touch [-a] [-m] [-t TIMESTAMP (yyyyMMdd:HHmmss) ] [-c] <path> ...]

[-touchz <path> ...]

[-truncate [-w] <length> <path> ...]

[-usage [cmd ...]]

Generic options supported are:

-conf <configuration file> specify an application configuration file

-D <property=value> define a value for a given property

-fs <file:///|hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.

-jt <local|resourcemanager:port> specify a ResourceManager

-files <file1,...> specify a comma-separated list of files to be copied to the map reduce cluster

-libjars <jar1,...> specify a comma-separated list of jar files to be included in the classpath

-archives <archive1,...> specify a comma-separated list of archives to be unarchived on the compute machines

The general command line syntax is:

command [genericOptions] [commandOptions]

root@26c5e0e439bc:/# hdfs dfs -put ./csvFile.csv/user/root/textData

put: `./csvFile.csv/user/root/textData': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -put ./csvFile.csv /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -put ./activityfile.txt /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -stat /user/root/textData/activityfile.txt2022-03-28 14:38:04

root@26c5e0e439bc:/# hdfs dfs -stat /user/root/textData/csvFile.csv

2022-03-28 14:37:22

root@26c5e0e439bc:/# hdfs dfs -stat "type:%F permissions:%a %u:%g size:%b name:%n" /user/root/textData/csvFile.csv

type:regular file permissions:644 root:supergroup size:527 name:csvFile.csv

root@26c5e0e439bc:/# hdfs dfs -stat "type:%F permissions:%a %u:%g size:%b name:%n" /user/root/textData/activityfile.txt

type:regular file permissions:644 root:supergroup size:157 name:activityfile.txt

root@26c5e0e439bc:/#  
  
  
  
Activity 3-4

root@26c5e0e439bc:/# hdfs dfs -put ./csvFile.csv/user/root/textData

put: `./csvFile.csv/user/root/textData': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -put ./csvFile.csv /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -put ./activityfile.txt /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -stat /user/root/textData/activityfile.txt2022-03-28 14:38:04

root@26c5e0e439bc:/# hdfs dfs -stat /user/root/textData/csvFile.csv

2022-03-28 14:37:22

root@26c5e0e439bc:/# hdfs dfs -stat "type:%F permissions:%a %u:%g size:%b name:%n" /user/root/textData/csvFile.csv

type:regular file permissions:644 root:supergroup size:527 name:csvFile.csv

root@26c5e0e439bc:/# hdfs dfs -stat "type:%F permissions:%a %u:%g size:%b name:%n" /user/root/textData/activityfile.txt

type:regular file permissions:644 root:supergroup size:157 name:activityfile.txt

root@26c5e0e439bc:/# # Opens file01.txt in vim. Paste the text from above and save/exit.

root@26c5e0e439bc:/# $ vim file01.txt

bash: $: command not found

root@26c5e0e439bc:/#

root@26c5e0e439bc:/# # Copy the file into the HDFS

root@26c5e0e439bc:/# vim file01.txt

root@26c5e0e439bc:/# vim file01.txt

root@26c5e0e439bc:/# hdfs dfs -put ./file01.txt /user/root

root@26c5e0e439bc:/# pig wordcount.pig

2022-03-29 03:57:41,229 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

2022-03-29 03:57:41,231 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

2022-03-29 03:57:41,232 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

2022-03-29 03:57:41,303 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58

2022-03-29 03:57:41,303 [main] INFO org.apache.pig.Main - Logging error messages to: //pig\_1648526261294.log

2022-03-29 03:57:41,799 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File wordcount.pig does not exist

Details at logfile: //pig\_1648526261294.log

2022-03-29 03:57:41,826 [main] INFO org.apache.pig.Main - Pig script completed in 651 milliseconds (651 ms)

root@26c5e0e439bc:/# bin/hdfs dfs -cat /user/root/results/part-r-00000

bash: bin/hdfs: No such file or directory

root@26c5e0e439bc:/# hdfs dfs -ls /user/root

Found 3 items

-rw-r--r-- 1 root supergroup 22 2022-03-29 03:57 /user/root/file01.txt

drwxr-xr-x - root supergroup 0 2021-08-23 10:44 /user/root/input

drwxr-xr-x - root supergroup 0 2022-03-28 14:38 /user/root/textData

root@26c5e0e439bc:/# hdfs dfs -ls /userclekha

ls: `/userclekha': No such file or directory

root@26c5e0e439bc:/# hdfs dfs -ls /user/clekha

Found 6 items

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:33 /user/clekha/activityfile.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:35 /user/clekha/csvFile.csv

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:24 /user/clekha/file2.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:23 /user/clekha/file3.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:15 /user/clekha/input.txt

-rw-r--r-- 1 root supergroup 0 2022-03-28 14:10 /user/clekha/myfile.txt

root@26c5e0e439bc:/# vim file01.txt

root@26c5e0e439bc:/# pig wordcount.pig

2022-03-29 04:01:01,669 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

2022-03-29 04:01:01,672 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

2022-03-29 04:01:01,672 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

2022-03-29 04:01:01,727 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58

2022-03-29 04:01:01,727 [main] INFO org.apache.pig.Main - Logging error messages to: //pig\_1648526461720.log

2022-03-29 04:01:02,067 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File wordcount.pig does not exist

Details at logfile: //pig\_1648526461720.log

2022-03-29 04:01:02,100 [main] INFO org.apache.pig.Main - Pig script completed in 462 milliseconds (462 ms)

root@26c5e0e439bc:/# hdfs dfs -put ./file01.txt /user/root/

put: `/user/root/file01.txt': File exists

root@26c5e0e439bc:/# pig wordcount.pig

2022-03-29 04:02:48,284 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

2022-03-29 04:02:48,286 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

2022-03-29 04:02:48,287 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

2022-03-29 04:02:48,345 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58

2022-03-29 04:02:48,345 [main] INFO org.apache.pig.Main - Logging error messages to: //pig\_1648526568339.log

2022-03-29 04:02:48,639 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File wordcount.pig does not exist

Details at logfile: //pig\_1648526568339.log

2022-03-29 04:02:48,669 [main] INFO org.apache.pig.Main - Pig script completed in 417 milliseconds (417 ms)

root@26c5e0e439bc:/# inputFile = LOAD 'hdfs:///user/root/file01.txt' AS (line);

bash: syntax error near unexpected token `('

root@26c5e0e439bc:/# grpd = GROUP words BY word;

bash: grpd: command not found

root@26c5e0e439bc:/# vim sales.csv

root@26c5e0e439bc:/# hdfs dfs -put ./sales.csv /user/root/

root@26c5e0e439bc:/# salesTable = LOAD 'hdfs:///user/root/sales.csv' USING PigStorage(',') AS (Product:chararray,Price:chararray,Payment\_Type:chararray,Name:chararray,City:chararray,State:chararray,Country:chararray);

bash: syntax error near unexpected token `('

root@26c5e0e439bc:/# GroupByCountry = GROUP salesTable BY Country;

bash: GroupByCountry: command not found

root@26c5e0e439bc:/# GroupByCountry = GROUP salesTable BY Country

bash: GroupByCountry: command not found

root@26c5e0e439bc:/# pig salesCSV.pig

2022-03-29 04:08:39,843 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

2022-03-29 04:08:39,845 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

2022-03-29 04:08:39,846 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

2022-03-29 04:08:39,903 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58

2022-03-29 04:08:39,903 [main] INFO org.apache.pig.Main - Logging error messages to: //pig\_1648526919898.log

2022-03-29 04:08:40,180 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File salesCSV.pig does not exist

Details at logfile: //pig\_1648526919898.log

2022-03-29 04:08:40,207 [main] INFO org.apache.pig.Main - Pig script completed in 395 milliseconds (395 ms)

root@26c5e0e439bc:/# hdfs dfs -cat /user/root/salesOutput/part-r-00000

cat: `/user/root/salesOutput/part-r-00000': No such file or directory

root@26c5e0e439bc:/# hive> CREATE TABLE files (line STRING);

bash: syntax error near unexpected token `('

root@26c5e0e439bc:/# hive> CREATE TABLE files (line STRING);

bash: syntax error near unexpected token `('

root@26c5e0e439bc:/# vim sales.csv

root@26c5e0e439bc:/# vim sales.csv

root@26c5e0e439bc:/# salesTable = LOAD 'hdfs:///user/root/sales.csv'

bash: salesTable: command not found

root@26c5e0e439bc:/# Product:chararray,Price:chararray,Payment\_Type:chararray,Name:chararray,City:chararray,State:chararray,Country:chararray

bash: Product:chararray,Price:chararray,Payment\_Type:chararray,Name:chararray,City:chararray,State:chararray,Country:chararray: command not found

root@26c5e0e439bc:/# (Product:chararray,Price:chararray,Payment\_Type:chararray,Name:chararray,City:chararray,State:chararray,Country:chararray);

bash: Product:chararray,Price:chararray,Payment\_Type:chararray,Name:chararray,City:chararray,State:chararray,Country:chararray: command not found

root@26c5e0e439bc:/# pig salesCSV.pig

2022-03-29 04:15:37,246 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL

2022-03-29 04:15:37,248 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE

2022-03-29 04:15:37,249 INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType

2022-03-29 04:15:37,311 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58

2022-03-29 04:15:37,311 [main] INFO org.apache.pig.Main - Logging error messages to: //pig\_1648527337306.log

2022-03-29 04:15:37,617 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File salesCSV.pig does not exist

Details at logfile: //pig\_1648527337306.log

2022-03-29 04:15:37,642 [main] INFO org.apache.pig.Main - Pig script completed in 427 milliseconds (427 ms)

root@26c5e0e439bc:/# hive> LOAD DATA LOCAL INPATH '/file01.txt' INTO TABLE files;

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/usr/local/apache-hive-3.1.2-bin/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/usr/local/hadoop-3.3.1/share/hadoop/common/lib/slf4j-log4j12-1.7.30.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See http://www.slf4j.org/codes.html#multiple\_bindings for an explanation.

SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]

Hive Session ID = a2054ea7-572c-4a71-9e77-2a8d35b8bcc5

Logging initialized using configuration in jar:file:/usr/local/apache-hive-3.1.2-bin/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true

Hive Session ID = da2113f4-25a9-420c-8300-1be13922a965

Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.