How to execute program in Hadoop (Take NgramJob.java as an example)：

Assuming environment variables are set as follows:

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

export PATH=${JAVA\_HOME}/bin:${PATH}

export HADOOP\_CLASSPATH=${JAVA\_HOME}/lib/tools.jar

Compile NgramJob.java and create a jar:

$ hadoop-3.2.1/bin/hadoop com.sun.tools.javac.Main program/ NgramJob.java

-/program$ jar cf NgramJob.jar NgramJob \*.class

Creat a text-file as input:

$ vim test/input

$ “helloworld” -> :wq

Put input text-file into hadoop:

$ hdfs dfs -mkdir input

$ hdfs dfs -put test/input

$ hdfs dfs -ls input

Run the application:

$ bin/hadoop jar NgramJob.jar NgramJob /input/\* /output

Output:

$ hdfs dfs -cat /output/\*