POLITECHNIKA
LUBELSKA
WYDZIAŁ ELEKTROTECHNIKI
I INFORMATYKI

Programowanie aplikacji w chmurze obliczeniowe

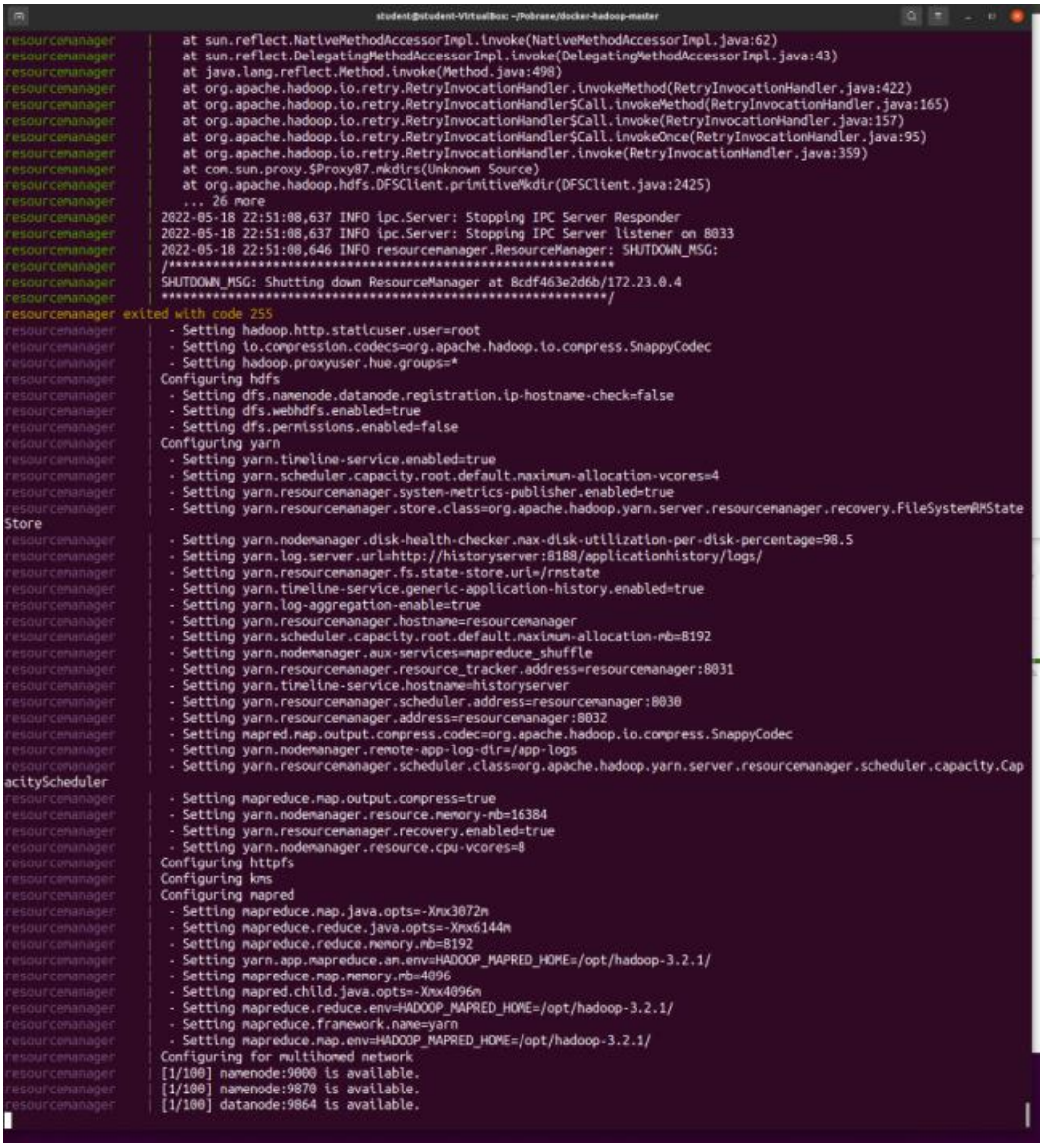
LABORATORIUM 8

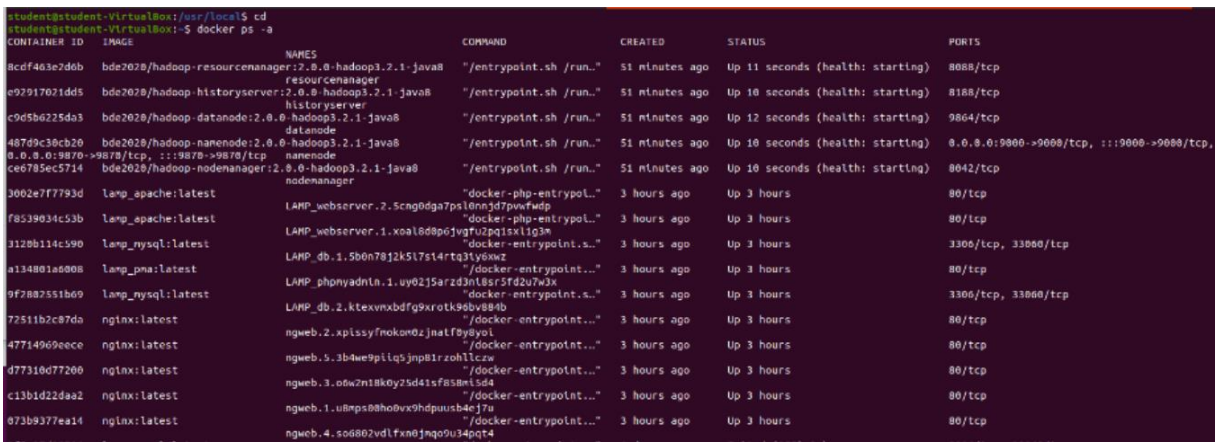
INSTALACJA SYSTEMU HADOOP W KONTENERZE DOCKER.

Kacper Majcher
Numer albumu: 92939
Grupa Laboratoryjna: 6.5
Prowadzący: mgr. D.Głuchowski

Zadanie 8.1. Porównanie obrazów Hadoop

Proszę zainstalować dwa kontenery zawierające środowisko Hadoop. Pierwszy proszę zainstalować bazując na obrazie znajdującym się w repozytorium Docker Hub, drugi w sposób przedstawiony w treści tego laboratorium. Proszę porównać parametry obu kontenerów oraz wykorzystanie przez nich zasobów systemowcyh.





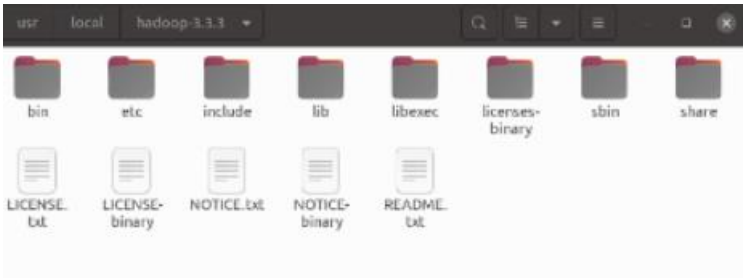
```
student@student-VirtualBox:~$ java -version
openjdk version "16.0.1" 2021-04-20
OpenJDK Runtime Environment (build 16.0.1+9-Ubuntu-120.04)
OpenJDK 64-Bit Server VM (build 16.0.1+9-Ubuntu-120.04, mixed mode, sharing)
student@student-VirtualBox:~$ sudo addgroup hadoop_group
Dodawanie grupy "hadoop_group" (GID 1001)...
Gotowe.
student@student-VirtualBox:~$ sudo adduser --ingroup hadoop_group hduser1
Dodawanie użytkownika "hduser1"...
Dodawanie nowego użytkownika "hduser1" (1001) w grupie "hadoop_group"...
Tworzenie katalogu domowego "/home/hduser1"...
Kopiowanie plików z "/etc/skel" ...
Nowe hasło :
Proszę ponownie wpisać nowe hasło :
passwd: hasło zostało zmienione
Zmieniam informację o użytkowniku hduser1
Wpisz nową wartość lub wciśnij ENTER by przyjąć wartość domyślną
        Imię i nazwisko []:
        Numer pokoju []:
        Telefon do pracy []:
        Telefon domowy []:
        Inne []:
Czy informacja jest poprawna? [T/n] T
student@student-VirtualBox:~$ sudo adduser hduser1 sudo
Dodawanie użytkownika "hduser1" do grupy "sudo"...
Dodaję nowego użytkownika hduser1 do grupy sudo
Gotowe.
student@student-VirtualBox:~$ ssh -version
Bad escape character 'rsion'.
student@student-VirtualBox:~$ openssh -version

Nie znaleziono polecenia 'openssh', czy chodziło o:

  polecenie 'openssn' z pakietu deb openssn (1.4-3build1)
  polecenie 'openssl' z pakietu deb openssl (1.1.1f-1ubuntu2.12)

Wypróbuj: sudo apt install <nazwa pakietu deb>
```

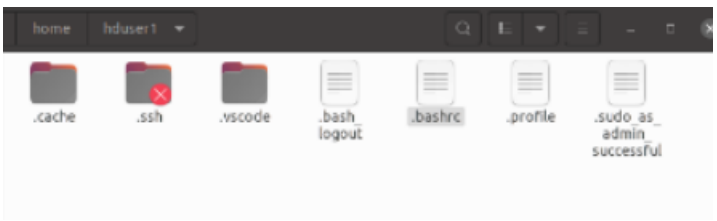
```
student@student-VirtualBox:~$ openssh -version

Nie znaleziono polecenia 'openssh', czy chodziło o:

  polecenie 'openssn' z pakietu deb openssn (1.4-3build1)
  polecenie 'openssl' z pakietu deb openssl (1.1.1f-1ubuntu2.12)

Wypróbuj: sudo apt install <nazwa pakietu deb>

student@student-VirtualBox:~$ sudo apt-get install openssh-server
Czytanie list pakietów... Gotowe
Budowanie drzewa zależności
Odczyt informacji o stanie... Gotowe
The following additional packages will be installed:
  ncurses-term openssh-sftp-server ssh-import-id
Sugerowane pakiety:
  molly-guard monkeysphere ssh-askpass
Zostaną zainstalowane następujące NOWE pakiety:
  ncurses-term openssh-server openssh-sftp-server ssh-import-id
0 aktualizowanych, 4 nowo instalowanych, 0 usuwanych i 47 nieaktualizowanych.
Konieczne pobranie 688 kB archiwów.
Po tej operacji zostanie dodatkowo użyte 6 010 kB miejsca na dysku.
Kontynuować? [T/n] T
Pobieranie:1 http://pl.archive.ubuntu.com/ubuntu focal/main amd64 ncurses-term all 6.2-0ubuntu2 [249 kB]
Pobieranie:2 http://pl.archive.ubuntu.com/ubuntu focal-updates/main amd64 openssh-sftp-server amd64 1:8.2p1-4ubuntu0.4 [51,5 kB]
Pobieranie:3 http://pl.archive.ubuntu.com/ubuntu focal-updates/main amd64 openssh-server amd64 1:8.2p1-4ubuntu0.4 [377 kB]
Pobieranie:4 http://pl.archive.ubuntu.com/ubuntu focal/main amd64 ssh-import-id all 5.10-0ubuntu1 [10,0 kB]
Pobrano 688 kB w 2s (454 kB/s)
Prekonfiguracja pakietów ...
Wybieranie wcześniej niewybranego pakietu ncurses-term.
(Odczytywanie bazy danych ... 185442 pliki i katalogi obecnie zainstalowane.)
Przygotowywanie do rozpakowania pakietu .../ncurses-term_6.2-0ubuntu2_all.deb ...
Rozpakowywanie pakietu ncurses-term (6.2-0ubuntu2) ...
Wybieranie wcześniej niewybranego pakietu openssh-sftp-server.
Przygotowywanie do rozpakowania pakietu .../openssh-sftp-server_1%3a8.2p1-4ubuntu0.4_amd64.deb ...
```

```
Creating config file /etc/ssh/sshd_config with new version
Creating SSH2 RSA key; this may take some time ...
3072 SHA256:n732FcnzzQ5/Wtor6/6psLPkPxlvJ+drf++r/6ar7YU root@student-VirtualBox (RSA)
Creating SSH2 ECDSA key; this may take some time ...
256 SHA256:lAsaduIFaZ49lSC7YR8ZzM1MjnkVqjErOaUlOSyY4wg root@student-VirtualBox (ECDSA)
Creating SSH2 ED25519 key; this may take some time ...
256 SHA256:zt+HrzlvR5xCEIEcPRY9jYBFgjPjPt+B/nNc8kUaWr0 root@student-VirtualBox (ED25519)
Created symlink /etc/systemd/system/sshd.service → /lib/systemd/system/ssh.service.
Created symlink /etc/systemd/system/multi-user.target.wants/ssh.service → /lib/systemd/system/ssh.service.
rescue-ssh.target is a disabled or a static unit, not starting it.
Konfigurowanie pakietu ssh-import-id (5.10-0ubuntu1) ...
Attempting to convert /etc/ssh/ssh_import_id
Konfigurowanie pakietu ncurses-term (6.2-0ubuntu2) ...
Przetwarzanie wyzwalaczy pakietu systemd (245.4-4ubuntu3.15)...
Przetwarzanie wyzwalaczy pakietu man-db (2.9.1-1)...
Przetwarzanie wyzwalaczy pakietu ufw (0.36-6ubuntu1)...
student@student-VirtualBox:~$ sudo cp /etc/ssh/sshd_config /etc/ssh/sshd_config.factory-defaults
student@student-VirtualBox:~$ sudo ls /stc/ssh/
ls: nie ma dostępu do '/stc/ssh/': Nie ma takiego pliku ani katalogu
student@student-VirtualBox:~$ sudo ls /stc/ssh
ls: nie ma dostępu do '/stc/ssh': Nie ma takiego pliku ani katalogu
student@student-VirtualBox:~$ sudo ls /etc/ssh
moduli         sshd_config                 ssh_host_ecdsa_key      ssh_host_ed25519_key.pub  ssh_import_id
ssh_config     sshd_config.d               ssh_host_ecdsa_key.pub  ssh_host_rsa_key
ssh_config.d   sshd_config.factory-defaults  ssh_host_ed25519_key    ssh_host_rsa_key.pub
student@student-VirtualBox:~$ su - hduser1
su: użytkownik - nie istnieje.
student@student-VirtualBox:~$ su -hduser1
su: użytkownik -hduser1 nie istnieje.
student@student-VirtualBox:~$ su hduser1
Hasło:
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hduser1@student-VirtualBox:/home/student$ sudo hduser1
[sudo] hasło użytkownika hduser1:
sudo: hduser1: nie znaleziono polecenia
hduser1@student-VirtualBox:/home/student$ su hduser1
Hasło:
```

```
hduser1@student-VirtualBox:/home/student$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser1/.ssh/id_rsa):
Created directory '/home/hduser1/.ssh'.
Your identification has been saved in /home/hduser1/.ssh/id_rsa
Your public key has been saved in /home/hduser1/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:jEko8bBJawfxk2LErST1UoyW1Eqt/shnuZuV34Fobms hduser1@student-VirtualBox
The key's randomart image is:
+---[RSA 3072]----+
| +XO.            |
|.oB%=o           |
| =X+O .          |
| o+= o +         |
| .    o S        |
|  .    o .       |
| . o .= . .      |
|  o ==E. . .     |
|   o+=o.. .      |
+----[SHA256]-----+
hduser1@student-VirtualBox:/home/student$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
hduser1@student-VirtualBox:/home/student$ cat $HOME/.ssh/authorized_keys
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABgQDRC7FaA8LTmJuwObelINZmAD+MiXN+qRuaJPw8bP8faCXx1a+c/4tWwWGFXSENfv7O+xqKQbWGvkflImsfJNT/uNLeN/XkQh0joR
82zJTQ+JLNxpGTQVTicliC2fL0zN286vv9gXYQO4TGfakN1gSdU+keAMchlFT3fdDfEewyLNESYaORhcxNmY8tyKoTsOOwNKJr0HQwWMc8Ecs9eQekhMen+30175X86KHyLFU0z4L7
OWMhXuG+Nu2PG3tlm7iyR7IIR71pJARNUWc6+8KlYF749acghJUn2nJ+vW49yE6K9eZddSRPCVLGEb5MFhOK4aUATs7NKmKER3iOkGTzlZCj7YKsms95tAbRSunzufbCDCEaZLexOC
0UoCVHePGjA5eNWBfbbokK/xweXNCuHL1oJC55hk2AzIYQe2BEZPCd4ISXk80LETMsqPU4HmPDNrZNP2cySmE6ENNPnmOJGsRmNQ43BGh46Jw3AW72TxqX7wMTtrvx6aXgDnjE= ho
ser1@student-VirtualBox
hduser1@student-VirtualBox:/home/student$
```

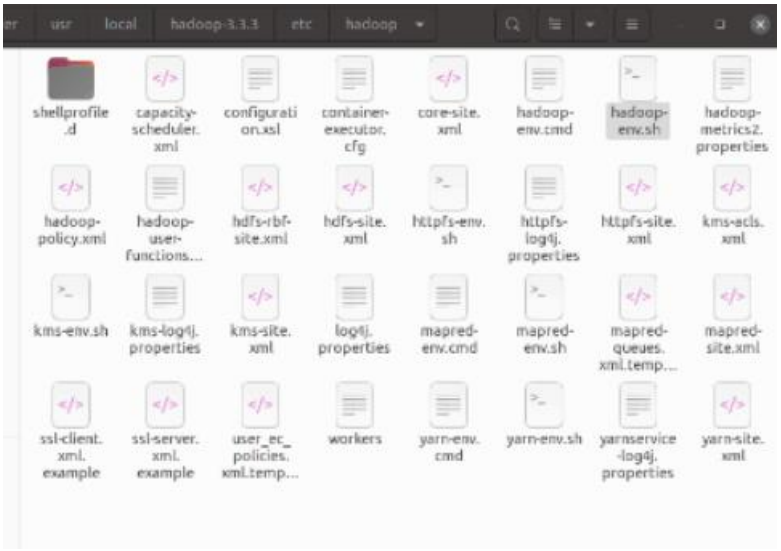
```
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/logos/build-by-maven-black.png
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/logos/maven-feather.png
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/logos/build-by-maven-white.png
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/banner.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/h5.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/icon_error_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/icon_success_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/expanded.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/external.png
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/icon_info_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/logo_apache.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/bg.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/newwindow.png
hadoop-3.3.3/share/doc/hadoop/hadoop-archives/images/h3.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/project-reports.html
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/dependency-analysis.html
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/css/
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/css/maven-base.css
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/css/print.css
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/css/maven-theme.css
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/css/site.css
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/breadcrumbs.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/apache-maven-project-2.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/maven-logo-2.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/collapsed.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logo_maven.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/icon_warning_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logos/
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logos/build-by-maven-black.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logos/maven-feather.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logos/build-by-maven-white.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/banner.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/h5.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/icon_error_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/icon_success_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/expanded.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/external.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/icon_info_sml.gif
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/logo_apache.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/bg.jpg
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/newwindow.png
hadoop-3.3.3/share/doc/hadoop/hadoop-hdfs-nfs/images/h3.jpg
student@student-VirtualBox:/usr/local$
```





```
# enable programmable completion features (you don't need to enable
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile
# sources /etc/bash.bashrc).
if ! shopt -oq posix; then
  if [ -f /usr/share/bash-completion/bash_completion ]; then
    . /usr/share/bash-completion/bash_completion
  elif [ -f /etc/bash_completion ]; then
    . /etc/bash_completion
  fi
fi

export HADOOP_HOME=/usr/local/hadoop-3.3.3
export PATH=$PATH:$HADOOP_HOME/bin
```

```
# Technically, the only required environment variable is JAVA_HOME.
# All others are optional.  However, the defaults are probably not
# preferred.  Many sites configure these options outside of Hadoop,
# such as in /etc/profile.d

# The java implementation to use. By default, this environment
# variable is REQUIRED on ALL platforms except OS X!
# export JAVA_HOME=/usr/lib/jvm/java-16-openjdk-amd64
```

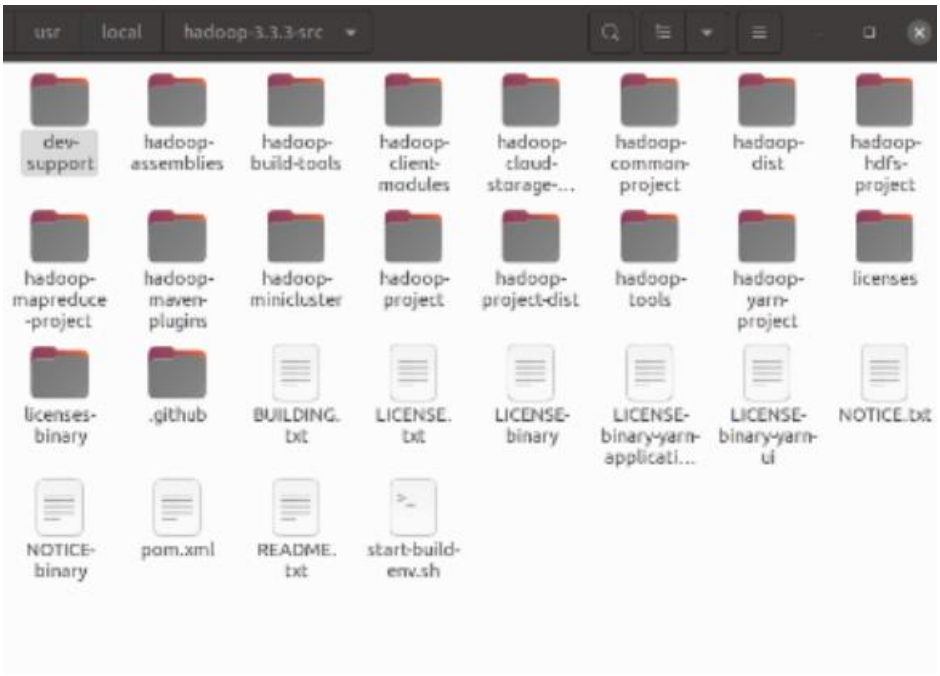



```
hduser1@student-VirtualBox:~$ /usr/local/hadoop-3.3.3/start-build-env.sh
bash: /usr/local/hadoop-3.3.3/start-build-env.sh: Nie ma takiego pliku ani katalogu
hduser1@student-VirtualBox:~$ /usr/local/hadoop-3.3.3/start-build-env.sh
bash: /usr/local/hadoop-3.3.3/start-build-env.sh: Brak dostępu
hduser1@student-VirtualBox:~$ sudo /usr/local/hadoop-3.3.3/start-build-env.sh
[sudo] hasło użytkownika hduser1:
sudo: /usr/local/hadoop-3.3.3/start-build-env.sh: nie znaleziono polecenia
hduser1@student-VirtualBox:~$ cd /usr/local/hadoop-3.3.3/
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$ ls
bin  include  libexec          licenses-binary  NOTICE-binary  README.txt  share
etc  lib      LICENSE-binary   LICENSE.txt      NOTICE.txt     sbin        start-build-env.sh
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$ chmod +x start-build-env.sh
chmod: nie można zmienić uprawnień do 'start-build-env.sh': Operacja niedozwolona
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$ sudo chmod +x start-build-env.sh
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$ sudo ./ start-build-env.sh
sudo: ./: nie znaleziono polecenia
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$ ./start-build-env.sh
unable to prepare context: path "dev-support/docker" not found
hduser1@student-VirtualBox:/usr/local/hadoop-3.3.3$
```

```
    . /usr/share/bash-completion/bash_completion
  elif [ -f /etc/bash_completion ]; then
    . /etc/bash_completion
  fi
fi

export HADOOP_HOME=/usr/local/hadoop-3.3.3-src
export PATH=$PATH:$HADOOP_HOME/dev-support/bin
```


```
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-httpfs-dist.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-mapreduce-dist.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-dynamometer.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-dynamometer-workload.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-dist.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-sls.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-dynamometer-blockgen.xml
hadoop-3.3.3-src/hadoop-assemblies/src/main/resources/assemblies/hadoop-tools.xml
hadoop-3.3.3-src/.github/pull_request_template.md
student@student-VirtualBox:/usr/local$ cd
student@student-VirtualBox:~$ su hduser1
Hasło:
hduser1@student-VirtualBox:/home/student$ cd
hduser1@student-VirtualBox:~$
```

```
=&ulimits=null&version=1: dial unix /var/run/docker.sock: connect: permission denied
hduser1@student-VirtualBox:~$ sudo /usr/local/hadoop-3.3.3-src/start-build-env.sh
Sending build context to Docker daemon  17.92kB
Step 1/31 : FROM ubuntu:bionic
bionic: Pulling from library/ubuntu
40dd5be53814: Pull complete
Digest: sha256:d21b6ba9e19feffa328cb3864316e6918e30acfd55e285b5d3df1d8ca3c7fd3f
Status: Downloaded newer image for ubuntu:bionic
 ---> c6ad7e71ba7d
Step 2/31 : WORKDIR /root
 ---> Running in a42ed8129a03
Removing intermediate container a42ed8129a03
 ---> 3cee801ef1e5
Step 3/31 : SHELL ["/bin/bash", "-o", "pipefail", "-c"]
 ---> Running in fc8a583f0e55
Removing intermediate container fc8a583f0e55
 ---> c975fdb2bdbf
Step 4/31 : RUN echo APT::Install-Recommends "0"\; > /etc/apt/apt.conf.d/10disableextras
 ---> Running in 5f480f981374
Removing intermediate container 5f480f981374
 ---> 8935922b623d
Step 5/31 : RUN echo APT::Install-Suggests "0"\; >>  /etc/apt/apt.conf.d/10disableextras
 ---> Running in 4a0c94c5fae6
Removing intermediate container 4a0c94c5fae6
 ---> fb586affb1ce
Step 6/31 : ENV DEBIAN_FRONTEND noninteractive
 ---> Running in 004184f519f2
Removing intermediate container 004184f519f2
 ---> 47a91c34dd85
Step 7/31 : ENV DEBCONF_TERSE true
 ---> Running in 64b7d5b7fbd4
Removing intermediate container 64b7d5b7fbd4
 ---> 525f1a7735bc
Step 8/31 : RUN apt-get -q update     && apt-get -q install -y --no-install-recommends         apt-utils         bats
 build-essential         bzip2         clang         cmake         curl         doxygen         fuse         g++         gcc
        git         gnupg-agent         libbz2-dev         libcurl4-openssl-dev         libfuse-dev         libprotobuf-dev
       libprotoc-dev         libsasl2-dev         libsnappy-dev         libssl-dev         libsnappy-dev         libtool
    libzstd1-dev         locales         make         pinentry-curses         pkg-config         python3         python3-pip
       python3-pkg-resources         python3-setuptools         python3-wheel         rsync         shellcheck         softw
are-properties-common         sudo         valgrind         zlib1g-dev     && apt-get clean     && rm -rf /var/lib/apt/lists/
*
```




```
 ---> Running in ac45a6eb2dfa
Removing intermediate container ac45a6eb2dfa
 ---> e653ab262f48
Successfully built e653ab262f48
Successfully tagged hadoop-build:latest
Sending build context to Docker daemon  2.048kB
Step 1/6 : FROM hadoop-build
 ---> e653ab262f48
Step 2/6 : RUN rm -f /var/log/faillog /var/log/lastlog
 ---> Running in d0aaeec04737
Removing intermediate container d0aaeec04737
 ---> e9ab8e7191ce
Step 3/6 : RUN groupadd --non-unique -g 1001 hduser1
 ---> Running in 8db98839f41f
Removing intermediate container 8db98839f41f
 ---> 81e00289a6ad
Step 4/6 : RUN useradd -g 1001 -u 1001 -k /root -m hduser1 -d "/home/hduser1"
 ---> Running in 7b9af4ac4112
Removing intermediate container 7b9af4ac4112
 ---> 9742321f73ec
Step 5/6 : RUN echo "hduser1 ALL=NOPASSWD: ALL" > "/etc/sudoers.d/hadoop-build-1001"
 ---> Running in 171e7b628476
Removing intermediate container 171e7b628476
 ---> f072a288b6ea
Step 6/6 : ENV HOME "/home/hduser1"
 ---> Running in 534f66eed68d
Removing intermediate container 534f66eed68d
 ---> 57c79e0e8367
Successfully built 57c79e0e8367
Successfully tagged hadoop-build-1001:latest

 _   _                 _                  _____
| | | |               | |                |  _  \
| |_| | __ _  __| | ___   ___  _ __   | | | |_____   __
|  _  |/ _` |/ _` |/ _ \ / _ \| '_ \  | | | / _ \ \ / /
| | | | (_| | (_| | (_) | (_) | |_) | | |/ /  __/\ V /
\_| |_/\__,_|\__,_|\___/ \___/| .__/  |___/ \___| \_(_)
                              | |
                              |_|

This is the standard Hadoop Developer build environment.
This has all the right tools installed required to build
Hadoop from source.

hduser1@bd73d2ad4f02:~/hadoop$
```

Wnioski

W przypadku uruchamiania usługi Hadoop pozyskanej z repozytorium Github, w rzeczywistości nie korzystamy z rozwiązania jednostanowiskowego, ale z kilku kontenerów Dockera, które umożliwią pracę. Instalacja Hadoop, pobranego jako pliki binarne - tak jak było to opisane w instrukcji do laboratoriów - nie powiodła się. Próba instalacji przy użyciu plików źródłowych zakończyła się sukcesem.