

Инструкция к ДЗ 1

Выполнил: Иванов Георгий
Группа: БПИ223

- 1) Сначала на каждом узле нужно в редакторе vim изменить файл /etc/hosts он должен выглядеть вот так после изменений:

```
<адрес jn> tpl-jn  
<адрес nn> tpl-nn  
<адрес dn-00> tpl-dn-00  
<адрес dn-01> tpl-dn-01
```

- 2) Необходимо создать пользователя hadoop на каждом узле воспользовавшись командой `sudo adduser hadoop` (добавив пароль этому пользователю).
- 3) Затем на каждом узле нужно создать ssh ключи для авторизации без пароля (в целом это опционально просто без этого придется каждый раз вводить пароли).
- 4) После этого нужно скачать hadoop с помощью команды (скачать нужно на tpl-jn)

```
wget https://dlcdn.apache.org/hadoop/common/hadoop-3.4.0/hadoop-3.4.0.tar.gz
```

- 5) Далее нужно скопировать этот файл на каждый узел с помощью следующих команд (предполагается, что пользователь сейчас находится на tpl-jn)

```
scp hadoop-3.4.0.tar.gz tpl-nn:/home/hadoop  
scp hadoop-3.4.0.tar.gz tpl-dn-00:/home/hadoop  
scp hadoop-3.4.0.tar.gz tpl-dn-01:/home/hadoop
```

- 6) Следующий шаг - распаковать эти файлы на каждом узле с помощью следующих команд (вводить нужно находясь на узле tpl-jn)

```
tar -xzf hadoop-3.4.0.tar.gz  
ssh tpl-nn  
tar -xzf hadoop-3.4.0.tar.gz  
ssh tpl-dn-00  
tar -xzf hadoop-3.4.0.tar.gz  
ssh tpl-dn-01  
tar -xzf hadoop-3.4.0.tar.gz  
exit  
exit  
exit
```

- 7) Переходим в пользователя hadoop командой: `sudo -i -u hadoop`
- 8) Далее необходимо убедиться, что в среде установлена java нужной версии. Также необходимо считать путь к java

```
java -version  
readlink -f /usr/bin/java
```

- 9) После этого нужно настроить hadoop - с помощью редактора vim вносим следующие изменения в файлы

```
- sudo vim hadoop-env.sh
```

Добавляем следующее:

```
JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
```

- sudo vim .profile

Добавляем следующее:

```
export HADOOP_HOME=/home/hadoop/hadoop-3.4.0
export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
```

- sudo vim vim core-site.xml

Добавляем следующее:

```
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://tmpl-nn:9000</value>
  </property>
</configuration>
```

- sudo vim hdfs-site.xml

Добавляем следующее:

```
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>3</value>
  </property>
</configuration>
```

- sudo vim workers

Добавляем следующее:

```
# localhost
tmpl-nn
tmpl-dn-00
tmpl-dn-01
```

10) Копируем файлы на все остальные узлы:

```
scp .profile tmpl-dn-01:/home/hadoop
scp .profile tmpl-dn-00:/home/hadoop
scp .profile tmpl-nn:/home/hadoop
scp core-site.xml tmpl-nn:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp core-site.xml tmpl-dn-00:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp core-site.xml tmpl-dn-01:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hdfs-site.xml tmpl-nn:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hdfs-site.xml tmpl-dn-00:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hdfs-site.xml tmpl-dn-01:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hadoop-env.sh tmpl-nn:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hadoop-env.sh tmpl-dn-00:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp hadoop-env.sh tmpl-dn-01:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp workers tmpl-nn:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp workers tmpl-dn-00:/home/hadoop/hadoop-3.4.0/etc/hadoop
scp workers tmpl-dn-01:/home/hadoop/hadoop-3.4.0/etc/hadoop
```

11) Форматируем файловую систему этой командой: bin/hdfs namenode -format

12) Запускаем класстер этой командой: sbin/start-dfs.sh

