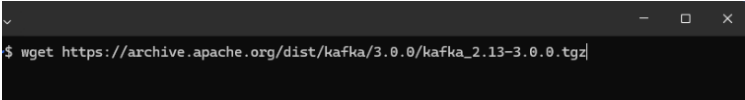

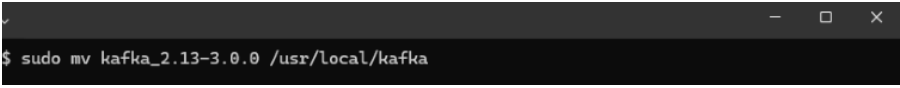


Kafka Workflow: Installing Kafka on Ubuntu 22.04

The purpose of this Standard Operating Procedure (SOP) is to provide: a step by step procedure to install Kafka on our Kubeflow cluster. Kafka will be used to stream data to our models continuously.

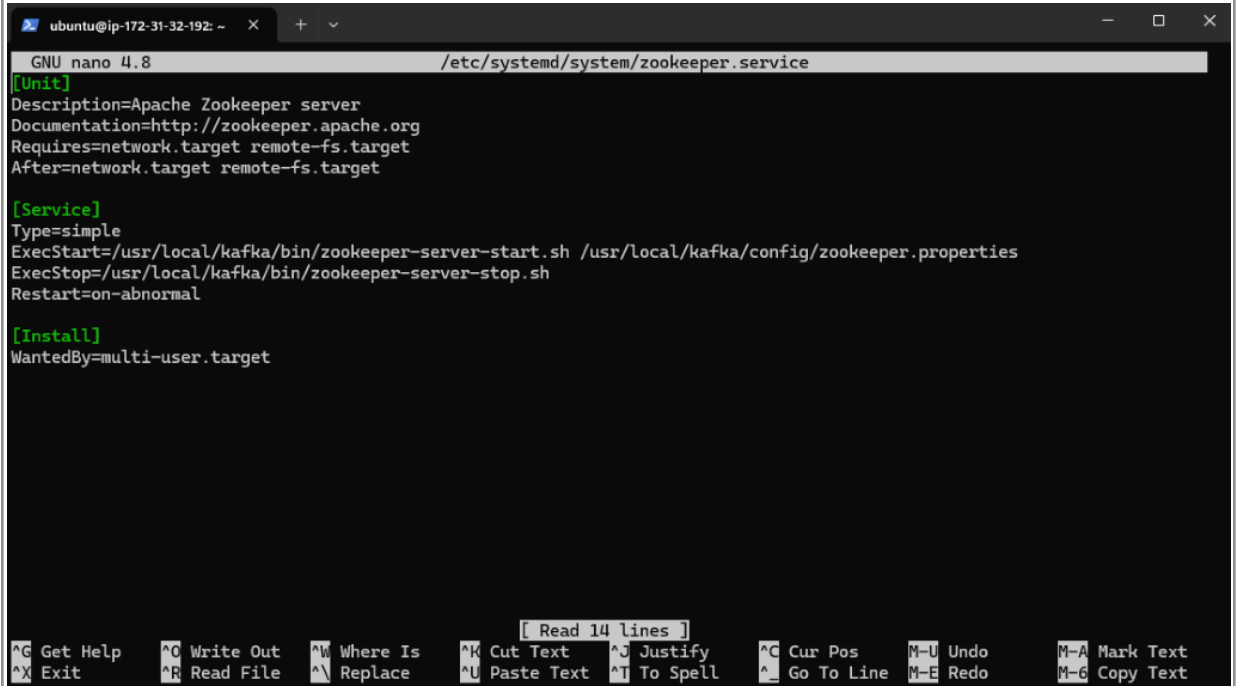
Procedure:

Steps	Description
1.	Start from the terminal of the ubuntu server.
2.	<p>Download Kafka using the following command:</p> <pre>wget https://archive.apache.org/dist/kafka/3.0.0/kafka_2.13-3.0.0.tgz</pre> 
3.	<p>Extract the downloaded archive file and place them under /usr/local/kafka directory. Using the following commands:</p> <pre>tar xzf kafka_2.13-3.0.0.tgz</pre>  <pre>sudo mv kafka_2.13-3.0.0 /usr/local/kafka</pre> 
4.	Create a systemd unit file for Zookeeper, using the following command:

sudo nano /etc/systemd/system/zookeeper.service

```
▼  
$ sudo nano /etc/systemd/system/zookeeper.service
```

5. Add the following content to the systemd unit file:



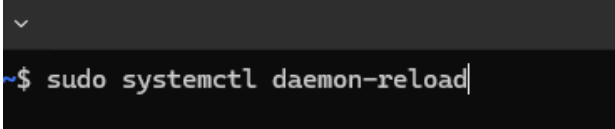
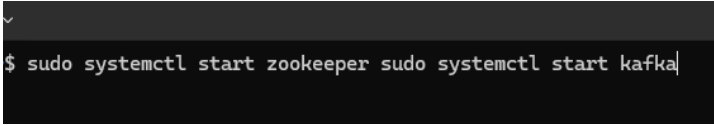
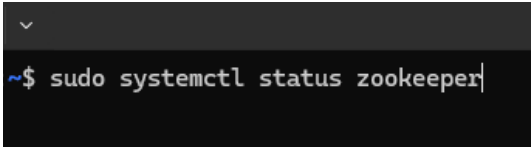
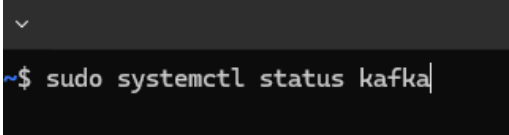
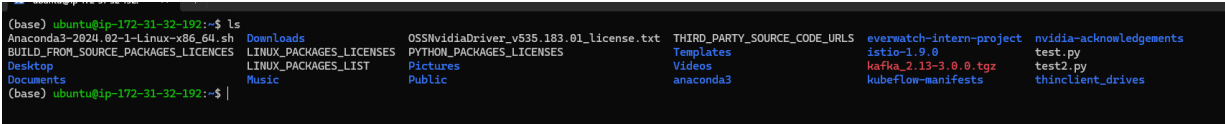
```
ubuntu@ip-172-31-32-192: ~  
GNU nano 4.8 /etc/systemd/system/zookeeper.service  
[Unit]  
Description=Apache Zookeeper server  
Documentation=http://zookeeper.apache.org  
Requires=network.target remote-fs.target  
After=network.target remote-fs.target  
  
[Service]  
Type=simple  
ExecStart=/usr/local/kafka/bin/zookeeper-server-start.sh /usr/local/kafka/config/zookeeper.properties  
ExecStop=/usr/local/kafka/bin/zookeeper-server-stop.sh  
Restart=on-abnormal  
  
[Install]  
WantedBy=multi-user.target  
  
[ Read 14 lines ]  
^G Get Help  ^O Write Out  ^W Where Is  ^K Cut Text   ^J Justify   ^C Cur Pos   M-U Undo     M-A Mark Text  
^X Exit      ^R Read File  ^\ Replace   ^U Paste Text ^I To Spell  ^_ Go To Line M-E Redo     M-G Copy Text
```

6. Save the file by pressing **ctrl + x**. Once prompted for y or n to save, press **y** for yes, to save and exit.

```
Save modified buffer?  
Y Yes  
N No      ^C Cancel
```

7. Reload the systemd daemon to apply new changes, using the following command:

sudo systemctl daemon-reload

	 <p>This will reload all the systemd files in the system environment.</p>
8.	<p>Use the systemctl commands to start a single-node ZooKeeper instance:</p> <pre>sudo systemctl start zookeeper sudo systemctl start kafka</pre> 
9.	<p>Verify both of the services status, using the following commands:</p> <pre>sudo systemctl status zookeeper</pre>  <pre>sudo systemctl status kafka</pre> 
10.	<p>Check that the kafka platform has been installed by running the ls command:</p> 
11.	<p>Install a broker :</p> <pre>cd /opt/kafka mkdir -p broker1 broker2 broker3 cp -r config broker1/config</pre>

Edit the `server.properties` file:

```
nano /opt/kafka/broker1/config/server.properties
```

Modify properties:

```
broker.id=1 listeners=PLAINTEXT://:9092 log.dirs=/opt/kafka/broker1/logs  
zookeeper.connect=localhost:2181
```