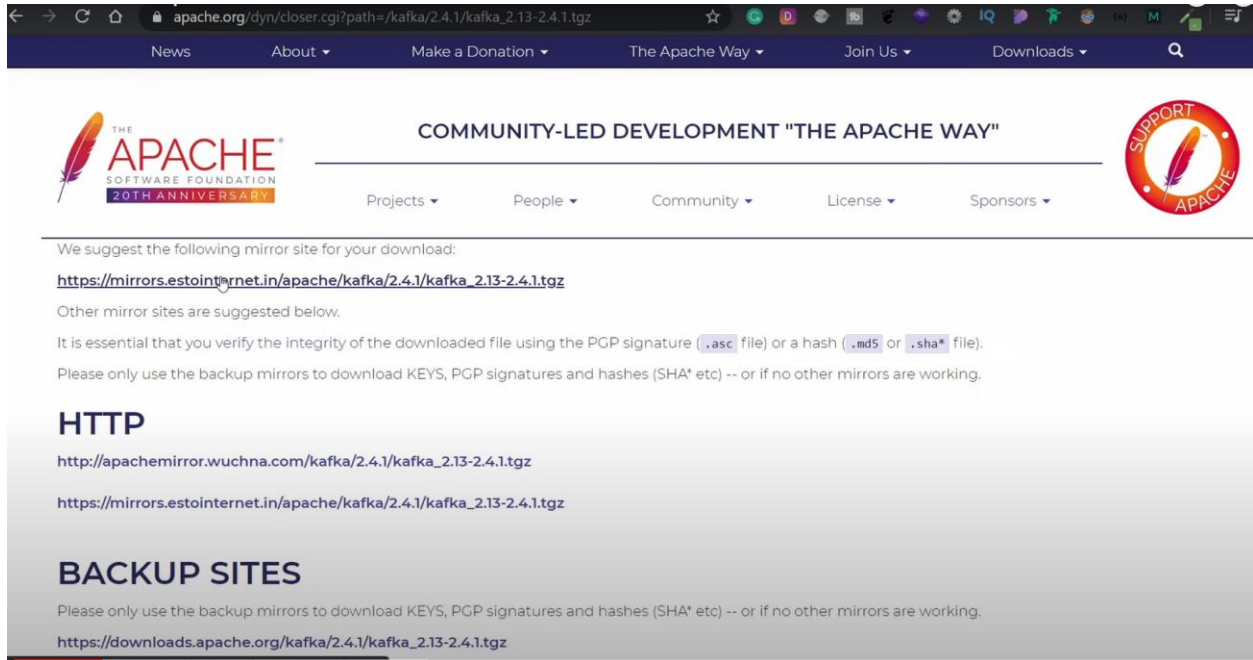


## Big Data Assignment#3

Muhammad Sheroz 9852

### Step1



The screenshot shows the Apache Kafka download page. At the top, there's a navigation bar with links like News, About, Make a Donation, The Apache Way, Join Us, and Downloads. Below the navigation bar is the Apache logo and the text "COMMUNITY-LED DEVELOPMENT 'THE APACHE WAY'". The main content area suggests a mirror site for download: [https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka\\_2.13-2.4.1.tgz](https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka_2.13-2.4.1.tgz). It also lists other mirror sites and provides instructions on how to verify the integrity of the downloaded file using PGP signatures or hashes. The page is divided into sections for HTTP and Backup Sites.

We suggest the following mirror site for your download:

[https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka\\_2.13-2.4.1.tgz](https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka_2.13-2.4.1.tgz)

Other mirror sites are suggested below.

It is essential that you verify the integrity of the downloaded file using the PGP signature (.asc file) or a hash (.md5 or .sha\* file).

Please only use the backup mirrors to download KEYS, PGP signatures and hashes (SHA\* etc) -- or if no other mirrors are working.

### HTTP

[http://apachemirror.wuchna.com/kafka/2.4.1/kafka\\_2.13-2.4.1.tgz](http://apachemirror.wuchna.com/kafka/2.4.1/kafka_2.13-2.4.1.tgz)

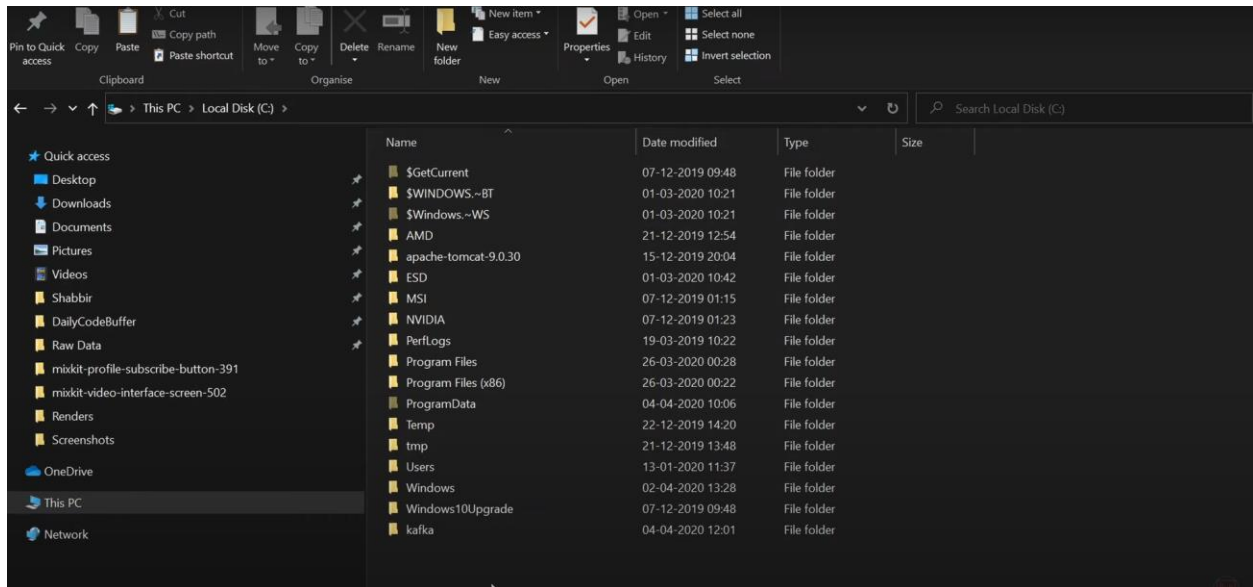
[https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka\\_2.13-2.4.1.tgz](https://mirrors.estointernet.in/apache/kafka/2.4.1/kafka_2.13-2.4.1.tgz)

### BACKUP SITES

Please only use the backup mirrors to download KEYS, PGP signatures and hashes (SHA\* etc) -- or if no other mirrors are working.

[https://downloads.apache.org/kafka/2.4.1/kafka\\_2.13-2.4.1.tgz](https://downloads.apache.org/kafka/2.4.1/kafka_2.13-2.4.1.tgz)

### Step2



The screenshot shows a Windows File Explorer window. The address bar indicates the current location is "This PC > Local Disk (C:)". The left sidebar shows the "Quick access" pane with various locations like Desktop, Downloads, Documents, Pictures, Videos, Shabbir, DailyCodeBuffer, Raw Data, mixkit-profile-subscribe-button-391, mixkit-video-interface-screen-502, Renders, Screenshots, OneDrive, This PC, and Network. The main pane displays a list of folders and files. The folders listed are: \$GetCurrent, \$WINDOWS~BT, \$WINDOWS~WS, AMD, apache-tomcat-9.0.30, ESD, MSI, NVIDIA, PerlLogs, Program Files, Program Files (x86), ProgramData, Temp, tmp, Users, Windows, Windows10Upgrade, and kafka. The files listed are: \$GetCurrent, \$WINDOWS~BT, \$WINDOWS~WS, AMD, apache-tomcat-9.0.30, ESD, MSI, NVIDIA, PerlLogs, Program Files, Program Files (x86), ProgramData, Temp, tmp, Users, Windows, Windows10Upgrade, and kafka.

Name	Date modified	Type	Size
\$GetCurrent	07-12-2019 09:48	File folder	
\$WINDOWS~BT	01-03-2020 10:21	File folder	
\$WINDOWS~WS	01-03-2020 10:21	File folder	
AMD	21-12-2019 12:54	File folder	
apache-tomcat-9.0.30	15-12-2019 20:04	File folder	
ESD	01-03-2020 10:42	File folder	
MSI	07-12-2019 01:15	File folder	
NVIDIA	07-12-2019 01:23	File folder	
PerlLogs	19-03-2019 10:22	File folder	
Program Files	26-03-2020 00:28	File folder	
Program Files (x86)	26-03-2020 00:22	File folder	
ProgramData	04-04-2020 10:06	File folder	
Temp	22-12-2019 14:20	File folder	
tmp	21-12-2019 13:48	File folder	
Users	13-01-2020 11:37	File folder	
Windows	02-04-2020 13:28	File folder	
Windows10Upgrade	07-12-2019 09:48	File folder	
kafka	04-04-2020 12:01	File folder	

### Step3 Change Kafka Server Proties

```
#advertised.listeners=PLAINTEXT://your.host.name:9092

# Maps listener names to security protocols, the default is for them to be the same. See the config documentation for more details
#listener.security.protocol.map=PLAINTEXT:PLAINTEXT,SSL:SSL,SASL_PLAINTEXT:SASL_PLAINTEXT,SASL_SSL:SASL_SSL

# The number of threads that the server uses for receiving requests from the network and sending responses to the network
num.network.threads=3

# The number of threads that the server uses for processing requests, which may include disk I/O
num.io.threads=8

# The send buffer (SO_SNDBUF) used by the socket server
socket.send.buffer.bytes=102400

# The receive buffer (SO_RCVBUF) used by the socket server
socket.receive.buffer.bytes=102400

# The maximum size of a request that the socket server will accept (protection against OOM)
socket.request.max.bytes=104857600

##### Log Basics #####

# A comma separated list of directories under which to store log files
log.dirs=/kafka-logs

# The default number of log partitions per topic. More partitions allow greater
# parallelism for consumption, but this will also result in more files across
# the brokers.
num.partitions=1

# The number of threads per data directory to be used for log recovery at startup and flushing at shutdown.
# This value is recommended to be increased for installations with data dirs located in RAID array.
num.recovery.threads.per.data.dir=1
```

## Step 4 Change Zookeeper Properties

```
# Licensed to the Apache Software Foundation (ASF) under one or more
# contributor license agreements. See the NOTICE file distributed with
# this work for additional information regarding copyright ownership.
# The ASF licenses this file to You under the Apache License, Version 2.0
# (the "License"); you may not use this file except in compliance with
# the License. You may obtain a copy of the License at
#
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
# the directory where the snapshot is stored.
dataDir=c:/kafka/zookeeper-data
# the port at which the clients will connect
clientPort=2181
# disable the per-ip limit on the number of connections since this is a non-production config
maxClientCnxns=0
# Disable the adminserver by default to avoid port conflicts.
# Set the port to something non-conflicting if choosing to enable this
admin.enableServer=false
# admin.serverPort=8080
```

## Step 5 Start Zookeeper

```
C:\kafka>.\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
```

```
10s sessionless connection timeout, 2 selector thread(s), 24 worker threads, and 64 kB direct buffers. (org.apache.zookeeper.server.NIOServerCnxnFactory)
[2020-04-04 12:07:09,207] INFO binding to port 0.0.0.0/0.0.0.0:2181 (org.apache.zookeeper.server.NIOServerCnxnFactory)
[2020-04-04 12:07:09,220] INFO zookeeper.snapshotSizeFactor = 0.33 (org.apache.zookeeper.server.ZKDatabase)
[2020-04-04 12:07:09,223] INFO Snapshotting: 0x0 to c:\kafka\zookeeper-data\version-2\snapshot.0 (org.apache.zookeeper.server.persistence.FileTxnSnapLog)
[2020-04-04 12:07:09,226] INFO Snapshotting: 0x0 to c:\kafka\zookeeper-data\version-2\snapshot.0 (org.apache.zookeeper.server.persistence.FileTxnSnapLog)
[2020-04-04 12:07:09,244] INFO Using checkIntervalMs=60000 maxPerMinute=10000 (org.apache.zookeeper.server.ContainerManager)
```

## Step 6 Start Kafka

```
C:\kafka>.\bin\windows\kafka-server-start.bat .\config\server.properties
[2020-04-04 12:08:34,340] INFO Registered kafka:type=kafka.Log4jController MBean (kafka.utils.Log4jControllerRegistration$)
[2020-04-04 12:08:34,642] INFO starting (kafka.server.KafkaServer)
[2020-04-04 12:08:34,643] INFO Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2020-04-04 12:08:34,658] INFO [ZooKeeperClient Kafka server] Initializing a new session to localhost:2181. (kafka.zookeeper.ZooKeeperClient)
```

## Step 7 Create Kafka Topics

```
C:\kafka>.\bin\windows\kafka-topics.bat --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic TestTopic
Created topic TestTopic.

C:\kafka>.\bin\windows\kafka-topics.bat --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic NewTopic
Created topic NewTopic.

C:\kafka>.\bin\windows\kafka-topics.bat --list --zookeeper localhost:2181
NewTopic
TestTopic

C:\kafka>_
```

#### Step 7 System message Work on TestTopic

```
C:\kafka>.\bin\windows\kafka-console-producer.bat --broker-list localhost:9092 --topic TestTopic
>Test INput data
>New data
>Welcome to Daily Code Buffer!
>Happy Learning!!!
>_
```

#### Step8 Run the TestTopic on consumer

```
C:\kafka>.\bin\windows\kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic TestTopic --from-beginning
Test INput data
New data
Welcome to Daily Code Buffer!
Happy Learning!!!
```