

## Tests Description

Test Name	Objective	Steps	Expected Results	Actual Results																																																
NiFi and HDFS status	Check if containers with NiFi and HDFS are running correctly.	1.Build the NiFi and HDFS containers executing docker-compose in folder with .yaml configuration file. 2.Run docker ps and check status of containers.	All containers are up and have “healthy” status.	<pre>nifi-hdfs\$ docker ps</pre> <table><thead><tr><th>CONTAINER ID</th><th>IMAGE</th><th>PORTS</th><th>COMMAND NAMES</th><th>CREATED</th><th>STATUS</th></tr></thead><tbody><tr><td>22985eeca2ca</td><td>bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8</td><td>0.0.0.0:8042-&gt;8042/tcp</td><td>"/entrypoint.sh /run..." hdfs-nodemanager</td><td>4 days ago</td><td>Up About a mi</td></tr><tr><td>nute (healthy)</td><td>bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8</td><td>0.0.0.0:8088-&gt;8088/tcp</td><td>"/entrypoint.sh /run..." hdfs-resourcemanager</td><td>4 days ago</td><td>Up 51 seconds</td></tr><tr><td>1368464f7651 (healthy)</td><td>bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8</td><td>0.0.0.0:9864-&gt;9864/tcp</td><td>"/entrypoint.sh /run..." hdfs-datanode</td><td>4 days ago</td><td>Up About a mi</td></tr><tr><td>9993c17be33b nute (healthy)</td><td>apache/nifi:latest</td><td>8000/tcp, 8443/tcp, 10000/tcp, 0.0.0.0:8080-&gt;8080/tcp</td><td>"../scripts/start.sh" nifi</td><td>4 days ago</td><td>Up About a mi</td></tr><tr><td>5d5abfdbcacb nute</td><td>bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8</td><td>0.0.0.0:9870-&gt;9870/tcp</td><td>"/entrypoint.sh /run..." hdfs-namenode</td><td>4 days ago</td><td>Up About a mi</td></tr><tr><td>926cdd043a45 nute (healthy)</td><td>compose_nifi-hdfs-news-scraper</td><td>0.0.0.0:8012-&gt;8012/tcp</td><td>"uvicorn server:app ..." news-scraper</td><td>4 days ago</td><td>Up About a mi</td></tr><tr><td>ae5a6c7b4369 nute</td><td></td><td></td><td></td><td></td><td></td></tr></tbody></table>	CONTAINER ID	IMAGE	PORTS	COMMAND NAMES	CREATED	STATUS	22985eeca2ca	bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8	0.0.0.0:8042->8042/tcp	"/entrypoint.sh /run..." hdfs-nodemanager	4 days ago	Up About a mi	nute (healthy)	bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8	0.0.0.0:8088->8088/tcp	"/entrypoint.sh /run..." hdfs-resourcemanager	4 days ago	Up 51 seconds	1368464f7651 (healthy)	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	0.0.0.0:9864->9864/tcp	"/entrypoint.sh /run..." hdfs-datanode	4 days ago	Up About a mi	9993c17be33b nute (healthy)	apache/nifi:latest	8000/tcp, 8443/tcp, 10000/tcp, 0.0.0.0:8080->8080/tcp	"../scripts/start.sh" nifi	4 days ago	Up About a mi	5d5abfdbcacb nute	bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8	0.0.0.0:9870->9870/tcp	"/entrypoint.sh /run..." hdfs-namenode	4 days ago	Up About a mi	926cdd043a45 nute (healthy)	compose_nifi-hdfs-news-scraper	0.0.0.0:8012->8012/tcp	"uvicorn server:app ..." news-scraper	4 days ago	Up About a mi	ae5a6c7b4369 nute					
CONTAINER ID	IMAGE	PORTS	COMMAND NAMES	CREATED	STATUS																																															
22985eeca2ca	bde2020/hadoop-nodemanager:2.0.0-hadoop3.2.1-java8	0.0.0.0:8042->8042/tcp	"/entrypoint.sh /run..." hdfs-nodemanager	4 days ago	Up About a mi																																															
nute (healthy)	bde2020/hadoop-resourcemanager:2.0.0-hadoop3.2.1-java8	0.0.0.0:8088->8088/tcp	"/entrypoint.sh /run..." hdfs-resourcemanager	4 days ago	Up 51 seconds																																															
1368464f7651 (healthy)	bde2020/hadoop-datanode:2.0.0-hadoop3.2.1-java8	0.0.0.0:9864->9864/tcp	"/entrypoint.sh /run..." hdfs-datanode	4 days ago	Up About a mi																																															
9993c17be33b nute (healthy)	apache/nifi:latest	8000/tcp, 8443/tcp, 10000/tcp, 0.0.0.0:8080->8080/tcp	"../scripts/start.sh" nifi	4 days ago	Up About a mi																																															
5d5abfdbcacb nute	bde2020/hadoop-namenode:2.0.0-hadoop3.2.1-java8	0.0.0.0:9870->9870/tcp	"/entrypoint.sh /run..." hdfs-namenode	4 days ago	Up About a mi																																															
926cdd043a45 nute (healthy)	compose_nifi-hdfs-news-scraper	0.0.0.0:8012->8012/tcp	"uvicorn server:app ..." news-scraper	4 days ago	Up About a mi																																															
ae5a6c7b4369 nute																																																				
Load cryptocurren cy rates data to HDFS	Check if NiFi Flow that loads data from coindcap.io API to HDFS works correctly.	1.Run crypto_rates_api_to_master flow in NiFi. 2. Check if new data appeared in the /master_dataset/crypto_rat es on Hadoop server.	Data with current timestamp appears in the HDFS.	<pre>root@namenode:/# hdfs dfs -ls /master_dataset/crypto_rates</pre> <pre>Found 3 items</pre> <table><tbody><tr><td>-rw-r--r--</td><td>1</td><td>nifi supergroup</td><td>64608</td><td>2023-11-20 16:44</td><td>/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-44-58.avro</td></tr><tr><td>-rw-r--r--</td><td>1</td><td>nifi supergroup</td><td>43197</td><td>2023-11-20 16:45</td><td>/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-45-00.avro</td></tr><tr><td>-rw-r--r--</td><td>1</td><td>nifi supergroup</td><td>43197</td><td>2023-11-20 16:46</td><td>/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-46-00.avro</td></tr></tbody></table>	-rw-r--r--	1	nifi supergroup	64608	2023-11-20 16:44	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-44-58.avro	-rw-r--r--	1	nifi supergroup	43197	2023-11-20 16:45	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-45-00.avro	-rw-r--r--	1	nifi supergroup	43197	2023-11-20 16:46	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-46-00.avro																														
-rw-r--r--	1	nifi supergroup	64608	2023-11-20 16:44	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-44-58.avro																																															
-rw-r--r--	1	nifi supergroup	43197	2023-11-20 16:45	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-45-00.avro																																															
-rw-r--r--	1	nifi supergroup	43197	2023-11-20 16:46	/master_dataset/crypto_rates/crypto_rates_2023-11-20-15-46-00.avro																																															
Load news data to HDFS	Check if NiFi Flow that loads data from newsdata.io API to HDFS works correctly.	1.Run NewsAPI_news_api_to_mas ter flow in NiFi. 2. Check if new data apeared in the /master_dataset/news/new sio on Hadoop server.	Data with current timestamp appears in the HDFS.	<pre>root@namenode:/# hdfs dfs -ls /master_dataset/news/news_io</pre> <pre>Found 1 items</pre> <table><tbody><tr><td>-rw-r--r--</td><td>1</td><td>nifi supergroup</td><td>16587</td><td>2023-11-20 16:33</td><td>/master_dataset/news/news_io/crypto_news_2023-11-20-15-33-49.parquet</td></tr></tbody></table>	-rw-r--r--	1	nifi supergroup	16587	2023-11-20 16:33	/master_dataset/news/news_io/crypto_news_2023-11-20-15-33-49.parquet																																										
-rw-r--r--	1	nifi supergroup	16587	2023-11-20 16:33	/master_dataset/news/news_io/crypto_news_2023-11-20-15-33-49.parquet																																															

Scrape news articles data and load it to HDFS	Check if our article scrapping component works correctly and loads data to HDFS.	<div>1.Check if compose_nifi-hdfs-news-scraper container is up and runing.</div> <div>2. Run NewsScraped_to_master flow in NiFi.</div> <div>3. Check if new data appeared in the folders which are corresponding to scrapred data sources in the /master_dataset/news on Hadoop server.</div>	<div>Data with current timestamp appears in the HDFS.</div> <div><div><div>Browse Directory</div><div><div><div>/master_dataset/news/yahoo</div><div>Go!</div><div><div><div></div><div></div><div></div></div></div></div><div>Show <div>25</div> entries</div><div>Search: <div></div></div><div><div><div><div></div><div>Permission</div><div>Owner</div><div>Group</div><div>Size</div><div>Last Modified</div><div>Replication</div><div>Block Size</div><div>Name</div></div></div><div><div><div></div><div>-rw-r--r--</div><div>nifi</div><div>supergroup</div><div>14.2 KB</div><div>Nov 16 12:57</div><div>1</div><div>32 MB</div><div><a href="#">crypto_news_2023-11-16-11-57-20.parquet</a></div><div></div></div></div></div></div><div><div><div>Browse Directory</div><div><div><div>/master_dataset/news/newsapi</div><div>Go!</div><div><div><div></div><div></div><div></div></div></div></div><div>Show <div>25</div> entries</div><div>Search: <div></div></div><div><div><div><div></div><div>Permission</div><div>Owner</div><div>Group</div><div>Size</div><div>Last Modified</div><div>Replication</div><div>Block Size</div><div>Name</div></div></div><div><div><div></div><div>-rw-r--r--</div><div>nifi</div><div>supergroup</div><div>144.79 KB</div><div>Nov 16 12:57</div><div>1</div><div>32 MB</div><div><a href="#">crypto_news_2023-11-16-11-57-55.parquet</a></div><div></div></div></div></div></div><div><div><div>Browse Directory</div><div><div><div>/master_dataset/news/cryptopanic</div><div>Go!</div><div><div><div></div><div></div><div></div></div></div></div><div>Show <div>25</div> entries</div><div>Search: <div></div></div><div><div><div><div></div><div>Permission</div><div>Owner</div><div>Group</div><div>Size</div><div>Last Modified</div><div>Replication</div><div>Block Size</div><div>Name</div></div></div><div><div><div></div><div>-rw-r--r--</div><div>nifi</div><div>supergroup</div><div>47.64 KB</div><div>Nov 16 12:57</div><div>1</div><div>32 MB</div><div><a href="#">crypto_news_2023-11-16-11-57-21.parquet</a></div><div></div></div></div></div></div><div><div><div>Browse Directory</div><div><div><div>/master_dataset/news/cryptocompare</div><div>Go!</div><div><div><div></div><div></div><div></div></div></div></div><div>Show <div>25</div> entries</div><div>Search: <div></div></div><div><div><div><div></div><div>Permission</div><div>Owner</div><div>Group</div><div>Size</div><div>Last Modified</div><div>Replication</div><div>Block Size</div><div>Name</div></div></div><div><div><div></div><div>-rw-r--r--</div><div>nifi</div><div>supergroup</div><div>154.53 KB</div><div>Nov 16 12:57</div><div>1</div><div>32 MB</div><div><a href="#">crypto_news_2023-11-16-11-57-23.parquet</a></div><div></div></div></div></div></div><div><div><div>Browse Directory</div><div><div><div>/master_dataset/news/alphavantage</div><div>Go!</div><div><div><div></div><div></div><div></div></div></div></div><div>Show <div>25</div> entries</div><div>Search: <div></div></div><div><div><div><div></div><div>Permission</div><div>Owner</div><div>Group</div><div>Size</div><div>Last Modified</div><div>Replication</div><div>Block Size</div><div>Name</div></div></div><div><div><div></div><div>-rw-r--r--</div><div>nifi</div><div>supergroup</div><div>22.11 KB</div><div>Nov 16 12:57</div><div>1</div><div>32 MB</div><div><a href="#">crypto_news_2023-11-16-11-57-55.parquet</a></div><div></div></div></div></div></div></div></div></div></div></div></div></div></div></div></div>
---	--	---	---

Kafka Status	Check if containers with Kafka brokers are running correctly.	<p>1.Build the Kafka containers executing docker-compose in folder with .yaml configuration file.</p> <p>2.Run docker ps and check status of containers.</p> <p>3.Create sample topic.</p> <p>4.Check the status of two Kafka brokers on newly created topic.</p>	<p>All containers are up and have “healthy” status.</p> <p>The topic is created and distributed among both Kafka containers.</p>	<pre> kafka1@kafka1:~/kafka1\$ docker ps CONTAINER ID   IMAGE          COMMAND                  CREATED       STATUS       PORTS                               NAMES fbell13e7e936   bitnami/kafka:3.6   "/opt/bitnami/script..." 14 seconds ago   Up 13 seconds   0.0.0.0:60888-&gt;9092/tcp   kafka_0 b1853e858fda    bitnami/kafka:3.6   "/opt/bitnami/script..." 14 seconds ago   Up 13 seconds   0.0.0.0:60886-&gt;9092/tcp   kafka_1 kafka1@kafka1:~/kafka1\$ docker exec -it kafka_0 /bin/bash I have no name!@kafka_0:/\$ kafka-topics.sh --create --bootstrap-server localhost:9092 --topic mytopic --partitions 2 --replication-factor 2 Created topic mytopic. I have no name!@kafka_0:/\$ kafka-topics.sh --describe --bootstrap-server localhost:9092 --topic mytopic Topic: mytopic TopicId: DTwrnMvVRZK2VSSa7qs3kg PartitionCount: 2      ReplicationFactor: 2      Configs: Topic: mytopic Partition: 0   Leader: 1      Replicas: 1,0   Isr: 1,0 Topic: mytopic Partition: 1   Leader: 0      Replicas: 0,1   Isr: 0,1 I have no name!@kafka_0:/\$ exit exit </pre>