**Cluster Overview and Assumptions**

* **Nodes:**
  + master1 (192.168.22.10): Primary Atlas server node.
  + master2 (192.168.22.11): Secondary Atlas server node (for HA, if desired).
  + master3 (192.168.22.16): Could host additional services (e.g., ZooKeeper quorum member).
  + slave2 (192.168.22.14): Hosts Kafka.
  + edge1 (192.168.22.17): Could be used for Hive/Sqoop clients or as an edge node for accessing Atlas UI.

 **Kafka**: Installed on slave2 (192.168.22.14), assumed to be running on port 9092.

 **HBase**: Assumed to be installed across the cluster (e.g., on master1, master2, master3), with permission issues to resolve.

 **ZooKeeper**: Assumed to be distributed across master1, master2, and master3 for HA (default port 2181).

 **Goal**: Deploy Atlas on master1 (and optionally master2 for HA), integrate with Kafka on slave2, and resolve HBase permissions.  
  
  
on master 1 (192.168.22.10) :  
wget https://dlcdn.apache.org/atlas/2.4.0/apache-atlas-2.4.0-sources.tar.gz

tar -xvzf apache-atlas-2.4.0-sources.tar.gz

sudo apt install maven

cd apache-atlas-sources-2.4.0/

export MAVEN\_OPTS="-Xms2g -Xmx2g"

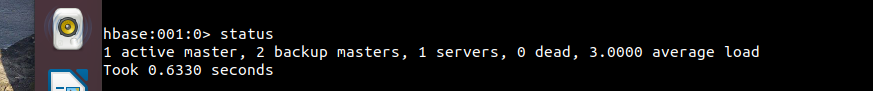
mvn clean -DskipTests package -Pdist (u have to change the user to hadoopZetta, dont build from the root)

sudo mkdir -p /opt/apache-atlas-2.4.0(

sudo tar -xvzf distro/target/apache-atlas-2.4.0-bin.tar.gz -C /opt/apache-atlas-2.4.0 --strip-components=1

cd /opt/apache-atlas-2.4.0

***Note: I’ve removed embedded-hbase-solr since i’m using an external HBase cluster and an external Solr soo i’ll to use this external solr***i switched to hadoopZetta :  
 cd /opt/hbase-2.5.5-hadoop3/bin

./start-hbase.sh  
hbase shell  
hbase status  
  
  
grant 'atlas', 'RWXCA', 'atlas'  
i got a solution,soo since kafka is available only in the slave soo we gotta point to it :  
vi/opt /apache-atlas-2.4.0/conf/atlas-application.properties  
# Notification Config

#atlas.notification.embedded=false # External Kafka(comment this if u dont it wont work)

atlas.kafka.zookeeper.connect=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

atlas.kafka.bootstrap.servers=192.168.22.14:9092

atlas.kafka.zookeeper.connection.timeout.ms=30000

atlas.kafka.zookeeper.session.timeout.ms=60000

atlas.kafka.zookeeper.sync.time.ms=20  
----->Assumes ZooKeeper quorum runs on master1, master2, and master3. ----------->Kafka broker is on slave2:9092.  
***on the atlas-env.sh***  
vi/opt/apache-atlas-2.4.0/config/atlas-env.sh  
export HBASE\_CONF\_DIR=/opt/apache-atlas-2.4.0/conf/hbase

export MANAGE\_LOCAL\_HBASE=false

export MANAGE\_LOCAL\_SOLR=false # Adjust if using external Solr

export MANAGE\_EMBEDDED\_CASSANDRA=false

***Configure Atlas for Cluster*** :  
atlas.graph.storage.backend=hbase

atlas.graph.storage.hostname=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

atlas.graph.storage.hbase.table=atlas

atlas.kafka.bootstrap.servers=192.168.22.14:9092

atlas.kafka.zookeeper.connect=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

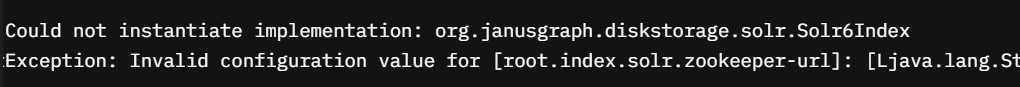
atlas.server.http.port=21000

atlas.server.bind.address=0.0.0.0 # Ensure it listens on all interfaces

# Optional: HA (if deploying on master2 as well)

atlas.server.ha.enabled=true atlas.server.ids=id1,id2 atlas.server.address.id1=192.168.22.10:21000 atlas.server.address.id2=192.168.22.11:21000

***For HA, install Atlas on master2 (repeat Step 1) and copy the same config, ensuring atlas.server.ids and addresses are consistent***.

After that i got this error :  
  
So let’s start working with solr(cloud Mode) :

Soo let’s install solr :  
switch to hadoopZetta   
  
wget https://downloads.apache.org/solr/solr/8.11.2/solr-8.11.2.tgz

tar -xvzf solr-8.11.2.tgz

mv solr-8.11.2 /opt/solr-8.11.2

/opt/solr-8.11.2/bin/solr start -cloud -z "192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181" -p 8983  
  
let’s create collections :  
/opt/solr-8.11.2/bin/solr create -c vertex\_index -n data-driven-schema-configs

/opt/solr-8.11.2/bin/solr create -c edge\_index -n data-driven-schema-configs

/opt/solr-8.11.2/bin/solr create -c fulltext\_index -n data-driven-schema-configs

curl http://192.168.22.10:8983/solr/admin/collections?action=LIST

* Should list vertex\_index, edge\_index, and fulltext\_index.

vi /opt/apache-atlas-2.4.0/conf/atlas-application.properties   
  
atlas.graph.index.search.backend=solr

atlas.graph.index.search.solr.mode=cloud

atlas.graph.index.search.solr.zookeeper-url=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

atlas.graph.index.search.solr.zookeeper-connect-timeout=60000

atlas.graph.index.search.solr.zookeeper-session-timeout=60000

atlas.graph.index.search.solr.wait-searcher=false

# Remove HTTP-specific setting

# atlas.graph.index.search.solr.http-urls=http://192.168.22.10:8983/solr  
vi /opt/apache-atlas-2.4.0/conf/atlas-env.sh  
export ATLAS\_OPTS="$ATLAS\_OPTS -Djanusgraph.index.search.solr.mode=http -Djanusgraph.index.search.solr.http-urls=http://192.168.22.10:8983/solr"

No description available.sooo let’s start atlas :  
/opt/apache -atlas-2.4.0/bin  
./atlas \_start.py  
***oups !***  
apperently a kafka error :  
  
then going to kafka configuration :  
going to the slave 2 :  
jps | grep Kafka  
If not, start Kafka

/opt/kafka/bin/kafka-server-start.sh -daemon /opt/kafka/config/server.properties  
  
cat /opt/kafka/config/server.properties | grep -E "zookeeper.connect|listeners|advertised.listeners"

excpected output :  
zookeeper.connect=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

listeners=PLAINTEXT://192.168.22.14:9092

advertised.listeners=PLAINTEXT://192.168.22.14:9092  
  
if not :  
vi /opt/kafka/config/server.properties  
  
zookeeper.connect=192.168.22.10:2181,192.168.22.11:2181,192.168.22.16:2181

listeners=PLAINTEXT://192.168.22.14:9092

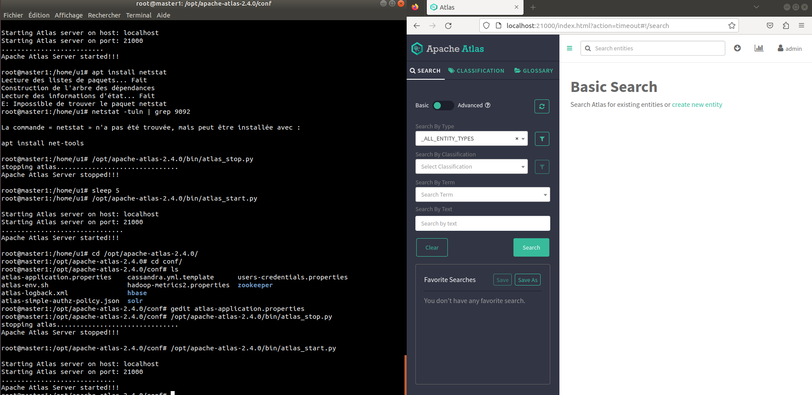
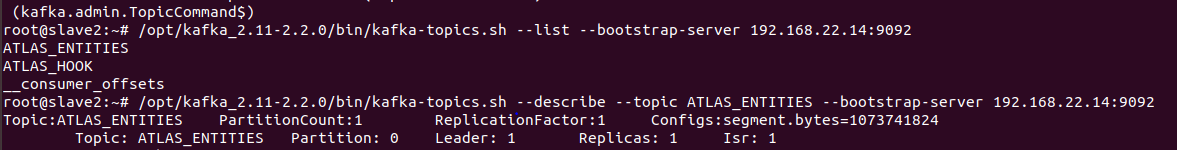
advertised.listeners=PLAINTEXT://192.168.22.14:9092  
  
  
/opt/kafka/bin/kafka-server-stop.sh

sleep 5

/opt/kafka/bin/kafka-server-start.sh -daemon /opt/kafka/config/server.properties  
 ***Note :Atlas expects ATLAS\_HOOK and ATLAS\_ENTITIES topics. Your config has atlas.notification.create.topics=true, but the errors suggest they might not exist yet.  
  
If ATLAS\_HOOK and ATLAS\_ENTITIES aren’t listed, create them***/opt/kafka/bin/kafka-topics.sh --create --topic ATLAS\_HOOK --bootstrap-server 192.168.22.14:9092 --partitions 1 --replication-factor 1

/opt/kafka/bin/kafka-topics.sh --create --topic ATLAS\_ENTITIES --bootstrap-server 192.168.22.14:9092 --partitions 1 --replication-factor 1

***Verify topic status :***  
  
/opt/kafka/bin/kafka-topics.sh --describe --topic ATLAS\_HOOK --bootstrap-server 192.168.22.14:9092

/opt/kafka/bin/kafka-topics.sh --describe --topic ATLAS\_ENTITIES --bootstrap-server 192.168.22.14:9092  
  
  
AND NOW ATLAS WORKS SMOOTHLY :