Integration Environment

Integration environment is meant to run integration tests.

Currently CMLA has 4 components:

- WEB UI (cmlportal)
- Stream processing (cml-storm topology)
- kafka REST proxy (kafka-rest) - kafka-rest proxy provided by confluent io that was enhanced to handle kerberos
- ingest agents :
    - cmla_filebeat ,
    - tail_to_kafka (deprecated)
    - log4j appender - kafka direct
    - log4net appender
    - log4j appender (REST based)

CMLA runs on the infrastructure based on HDP stack (version 2.5) utilizing zookeeper, kafka, and storm.

Windows VM SE132238.devoak.devfg.rbc.com was setup to allow intergration testing of new versions of all CMLA components without affecting users that are in process of onboarding

Currently we can run manual tests on new versions. Goal is to develop a set of scripts that can flow test sources through ingest agents, streaming components and verify index and created hdfs file

Ultimate goal is to run these tests as part of CI process.

Prerequisites:

1. Java JDK 1.8 or above installed into a location without spaces : ex C:\Java\jkd1.8.0._60
2. Python27
3. winzip or winrar to untar files

Setup

1. Login as : Admin user :   SE132238admin : Se132238RIK0
2. zookeeper, kafka and kafka-rest proxy are running as services deployed via nssm
3. If storm, hdfs and portal are not running start them :
    a. to run storm : double click on run_storm shortcut
    b. to run hadoop : double click on start_all shortcut
4. There is no kerberos

When setting up hdfs you have to set some directory to be used as storage. By default it is c:\tmp\hadoop-${user}.

So I changed this in core-site.xml to :

<property>
<name>hadoop.tmp.dir</name>
<value>/F:/tmp/cmla/hdfs</value>
</property>

And added following to hdfs-site.xml:

<property>
<name>dfs.namenode.name.dir</name>
<value>${hadoop.tmp.dir}/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>${hadoop.tmp.dir}/datanode</value>
</property>

Before running first time from Administrator command prompt run:

hdfs namenode -format

This will create namenode directory

Portal

NOTE: To run Portal you need to setup virtual environment with requirments.txt that are distributed in cmlportal tar file

On windows use requirements-win.txt

To test a new version of portal follow steps below:

1. transfer a new version of portal tar file and untar it to a directory on the F drive (ex: F:\cmlportal-0.1.23)
2. Open Adminstrator command prompt
3. Delete current symbolic link c:\cmlportal : `del c:\cmlportal`
4. create a new one pointing to the created directory : `mklink /D c:\cmlportal  f:/cmlportal-0.1.23`
5. start portal by double clicking on run_portal desktop shortcut.
6. Login into portal from a different machine by using URL *SE132238.devoak.devfg.rbc.com:8083/admin/login*
7. Please note that admin user is *admin:coffee00*
8. *sqlite file with portal DB is located in F:\Tools\cmlportal\db.sqlite3*

Check functionality you need to check

Storm topolgy

To test a new version of storm topology:

- Download latest topology (cml-storm<version>.jar) from nexus and put it into a location on F:\cml-storm
- Kill current topology *(open command prompt and type: storm list)*
- This will list all topologies. Output will look as below:

> Topology_name Status Num_tasks Num_workers Uptime_secs

----
cmla-DEV-all-stream ACTIVE 17 3 4196

- Kill topology: *storm kill topology name*
- deploy topology by running command :

```
storm jar  F:\cml-storm\cml-storm<version>.jar   com.rbccm.rik0.storm.CMLATopologyMain  -portal ht
tp://SE132238.devoak.devfg.rbc.com:8083 -esuser es_admin -espassword 123 -latest
```

To view STORM UI open from a different machine open URL *SE132238.devoak.devfg.rbc.com:8080*

Kafka REST Proxy

1. Stop kafka-rest service
2. Remove symbolic link to c:\kafka-rest
3. Create symbolic link c:\kafka-rest pointing to the new directory
4. restart service
5. From another machine open URL http://SE132238.devoak.devfg.rbc.com:8082/topics

You should see a list of current topics

Ingest Agents

You can test agents by loading latest versions and using sources setup in the portal.

To browse hdfs :

1. Open command prompt and type hdfs dfs -ls (or any other hdfs comman)
2. Form another machine open http://SE132238.devoak.devfg.rbc.com:50070