

flink相关的下载包

<https://flink.apache.org/downloads.html>

编译flink 配套hadoop包

主要为了编译flink-shaded-hadoop-3-uber-3.0.0-cdh6.3.2-10.0.jar包

根目录pom.xml添加标签

```
1
2 <profile>
3   <id>vendor-repos</id>
4   <activation>
5     <property>
6       <name>vendor-repos</name>
7     </property>
8   </activation>
9   <!-- Add vendor maven repositories -->
10  <repositories>
11    <!-- Cloudera -->
12    <repository>
13      <id>cloudera-releases</id>
14      <url>https://repository.cloudera.com/artifactory/cloudera-repos</url>
15      <releases>
16        <enabled>true</enabled>
17      </releases>
18      <snapshots>
19        <enabled>false</enabled>
20      </snapshots>
21    </repository>
22    <!-- Hortonworks -->
23    <repository>
24      <id>HDPReleases</id>
25      <name>HDP Releases</name>
26      <url>https://repo.hortonworks.com/content/repositories/releases/</url>
27      <snapshots><enabled>false</enabled></snapshots>
28      <releases><enabled>true</enabled></releases>
29    </repository>
```

```

30 <repository>
31 <id>HortonworksJettyHadoop</id>
32 <name>HDP Jetty</name>
33 <url>https://repo.hortonworks.com/content/repositories/jetty-hadoop</url>
34 <snapshots><enabled>false</enabled></snapshots>
35 <releases><enabled>true</enabled></releases>
36 </repository>
37 <!-- MapR -->
38 <repository>
39 <id>mapr-releases</id>
40 <url>https://repository.mapr.com/maven/</url>
41 <snapshots><enabled>false</enabled></snapshots>
42 <releases><enabled>true</enabled></releases>
43 </repository>
44 </repositories>
45 </profile>

```

修改工程flink-shaded-10.0中pom.xml相关的版本，设计到cdh中的hadoop zookeeper等版本均需要修改

```

<parent>
  <groupId>org.apache.flink</groupId>
  <artifactId>flink-shaded</artifactId>
  <version>10.0</version>
  <relativePath>..</relativePath>
</parent>

<artifactId>flink-shaded-hadoop-3-parent</artifactId>
<name>flink-shaded-hadoop-3-parent</name>
<packaging>pom</packaging>
<version>10.0</version>

<properties>
  <!-- never use this in the parent module; otherwise it
       not be done since the parent reference in child mod
       cannot be resolved -->
  <hadoop.version>3.0.0</hadoop.version>
  <!-- pinned transitive dependencies -->
  <avro.version>1.8.2</avro.version>
  <slf4j.version>1.7.15</slf4j.version>
  <log4j.version>1.2.17</log4j.version>
  <zookeeper.version>3.4.5</zookeeper.version>
  <findbugs.version>1.3.9</findbugs.version>
</properties>

```

编译jar包

```
mvn -T2C clean install -DskipTests -Pvendor-repos -  
Dhadoop.version=3.0.0-cdh6.3.2 -Dscala-2.12 -Drat.skip=true
```

包依赖报错

```
包依赖报错，阿里的镜像没有hadoop高版本的包，修改进行到maven官网重新安装  
[WARNING] The requested profile "vendor-repos" could not be activated because it does  
not exist.  
[ERROR] Failed to execute goal on project flink-shaded-hadoop-3: Could not resolve  
dependencies for project org.apache.flink:flink-shaded-hadoop-3:jar:3.0.0-cdh6.3.2-10.0:  
The following artifacts could not be resolved: org.apache.hadoop:hadoop-  
common:jar:3.0.0-cdh6.3.2, org.apache.hadoop:hadoop-hdfs:jar:3.0.0-cdh6.3.2,  
org.apache.hadoop:hadoop-mapreduce-client-core:jar:3.0.0-cdh6.3.2,  
org.apache.hadoop:hadoop-yarn-client:jar:3.0.0-cdh6.3.2, org.apache.hadoop:hadoop-  
yarn-common:jar:3.0.0-cdh6.3.2: Failure to find org.apache.hadoop:hadoop-  
common:jar:3.0.0-cdh6.3.2 in  
http://maven.aliyun.com/nexus/content/repositories/central/ was cached in the local  
repository, resolution will not be reattempted until the update interval of alimaven has  
elapsed or updates are forced -> [Help 1]
```

peer not authenticated

添加官方镜像，可能是网络，或者认证的原因

```
Failed to execute goal on project flink-shaded-jackson-module-jsonSchema: Could not  
resolve dependencies for project org.apache.flink:flink-shaded-jackson-module-  
jsonSchema:jar:2.10.1-10.0: Failed to collect dependencies for  
[com.fasterxml.jackson.core:jackson-core:jar:2.10.1 (compile),  
com.fasterxml.jackson.core:jackson-annotations:jar:2.10.1 (compile),  
com.fasterxml.jackson.core:jackson-databind:jar:2.10.1 (compile),  
com.fasterxml.jackson.module:jackson-module-jsonSchema:jar:2.10.1 (compile)]: Failed  
to read artifact descriptor for com.fasterxml.jackson.core:jackson-core:jar:2.10.1:  
Could not transfer artifact com.fasterxml.jackson:jackson-base:pom:2.10.1 from/to  
mapr-releases  
https://repository.mapr.com/maven/): peer not authenticated
```

最后会打印编译到的目录

```
[INFO] Installing /home/my/flink-shaded-10.0/flink-shaded-hadoop-3-parent/flink-  
shaded-hadoop-3-uber/target/flink-shaded-hadoop-3-uber-3.0.0-cdh6.3.2-10.0.jar  
to /root/.m2/repository/org/apache/flink/flink-shaded-hadoop-3-uber/3.0.0-  
cdh6.3.2-10.0/flink-shaded-hadoop-3-uber-3.0.0-cdh6.3.2-10.0.jar
```

jar包目录

```
/root/.m2/repository/org/apache/flink/
```

Flink1.10编译（可选）

建议直接下载flink-1.10.0-bin-scala_2.12.tgz包

参考 <https://blog.csdn.net/mhtc5052/article/details/104560570>

```
# mvn clean install -DskipTests -Dfast -Drat.skip=true -
```

```
Dhadoop.version=3.0.0-cdh6.3.1 -Pvendor-repos -Dinclude-hadoop -
```

```
Dscala-2.12 -T2C
```

cdh上关闭kerberos安全

需要配置security.kerberos.login.keytab和security.kerberos.login.principal两个参数才能正正常启动。如未启动kerberos,则在CDH中添加FLINK服务时请清空这两个参数的内容
清空这两个参数

Cluster1

Flink-yarn

操作

状态

实例

配置

命令

图表库

审核

快速链接

keytab

筛选器

范围

Flink-yarn (服务范围)

2

Flink Yarn

0

类别

主要

2

安全性

0

性能

0

日志

0

监控

0

security.kerberos.login.keytab

security.kerberos.login.keytab

security.kerberos.login.principal

security.kerberos.login.principal

Flink-yarn (服务范围)

/etc/keytab/flink.keytab

Flink-yarn (服务范围)

/etc/keytab/flink.keytab

```
org.apache.flink.configuration.IllegalConfigurationException: Kerberos  
login configuration is invalid; keytab is unreadable
```

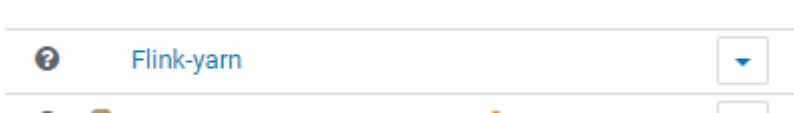
```
上午11点26:09.770分 ERROR FlinkYarnSessionCli ~ Error while running the Flink Yarn session.
org.apache.flink.configuration.IllegalConfigurationException: Kerberos login configuration is invalid; keytab is unreadable
    at org.apache.flink.runtime.security.SecurityConfiguration.validate(SecurityConfiguration.java:139)
    at org.apache.flink.runtime.security.SecurityConfiguration.<init>(SecurityConfiguration.java:90)
    at org.apache.flink.runtime.security.SecurityConfiguration.<init>(SecurityConfiguration.java:71)
    at org.apache.flink.yarn.cli.FlinkYarnSessionCli.main(FlinkYarnSessionCli.java:842)

-----
The program finished with the following exception:

org.apache.flink.configuration.IllegalConfigurationException: Kerberos login configuration is invalid; keytab is unreadable
    at org.apache.flink.runtime.security.SecurityConfiguration.validate(SecurityConfiguration.java:139)
    at org.apache.flink.runtime.security.SecurityConfiguration.<init>(SecurityConfiguration.java:90)
    at org.apache.flink.runtime.security.SecurityConfiguration.<init>(SecurityConfiguration.java:71)
    at org.apache.flink.yarn.cli.FlinkYarnSessionCli.main(FlinkYarnSessionCli.java:842)
```

安装成功最后可能需要重启Cloudera Management Service

需要重启，否则状态未知



python 环境和依赖

设置python镜像

/opt/anaconda3/bin/pip config set global.index-url

<https://mirrors.aliyun.com/pypi/simple/>

安装基础包

/opt/anaconda3/bin/pip install pyflink apache-flink

/opt/anaconda3/bin/pip install PyHDFS==0.2.1

修改默认python为python3，flink不支持python2

删除软连接

rm -f /usr/bin/python

创建软连接

ln -s /opt/anaconda3/bin/python3.6 /usr/bin/python