Kylin on Docker

Apache Kylin是Hadoop集群的客户端，无状态的Server服务，将其运行在Docker Container中具有一定的优势，Kylin依赖HBase，Hive和Hadoop等组件，步骤如下：

**1）获取客户端配置，包括hbase,hive和hadoop**

配置如下：

[root@sgserver006 conf]# ll

-rw-r--r-- 1 root root 4366 Nov 29 12:44 core-site.xml

-rw-r--r-- 1 root root 5281 Nov 29 12:44 hbase-site.xml

-rw-r--r-- 1 root root 8275 Nov 29 12:44 hdfs-site.xml

-rw-r--r-- 1 root root 19586 Nov 29 12:44 hive-site.xml

-rw-r--r-- 1 root root 6566 Nov 29 12:44 mapred-site.xml

-rw-r--r-- 1 root root 14085 Nov 29 12:44 yarn-site.xml

**2）配置kylin的配置，主要是kylin.properties，进行常规配置：**

*# kylin server's mode*

*kylin.server.mode=all*

*# optional information for the owner of kylin platform, it can be your team's email*

*# currently it will be attached to each kylin's htable attribute*

*kylin.owner=whoami@kylin.apache.org*

*# List of web servers in use, this enables one web server instance to sync up with other servers.*

*kylin.rest.servers=localhost:7070*

*# The metadata store in hbase*

*kylin.metadata.url=kylin\_metadata@hbase*

*# The storage for final cube file in hbase*

*kylin.storage.url=hbase*

*# Temp folder in hdfs, make sure user has the right access to the hdfs directory*

*kylin.hdfs.working.dir=/kylin1*

........................

**3.生成kylin的Dockerfile用于生成镜像**

*[root@sgserver006 kylin]# ll*

*total 28*

*-rw-r--r-- 1 root root 172 Nov 29 12:57 bootstrap.sh*

*drwxr-xr-x 2 root root 4096 Nov 30 08:35 conf*

*-rw-r--r-- 1 root root 2611 Nov 30 05:24 Dockerfile*

*-rw-r--r-- 1 root root 94 Nov 30 02:35 ssh\_config*

*drwxr-xr-x 2 root root 4096 Nov 29 12:44 yum*

以下的各个文件的作用如下：

1）conf ： 步骤1和2定义的Hadoop和kylin配置

2）ssh\_config： ssh的配置

*Host \**

*UserKnownHostsFile /dev/null*

*StrictHostKeyChecking no*

*LogLevel quiet*

*Port 2122*

3）yum : 本地yum源

4）bootstrap.sh : container启动脚本，内容如下

*#!/bin/bash*

*service sshd start*

*$KYLIN\_HOME/bin/kylin.sh start*

*if [[ $1 == "-d" ]]; then*

*while true; do sleep 1000; done*

*fi*

*if [[ $1 == "-bash" ]]; then*

*/bin/bash*

*f*

5）Dockerfile文件如下：

#Create a kylin 1.5.2-bc1.3.2 + BCH 1.3.2

FROM fys/centos

MAINTAINER from CMSS.inc by fys(fengyongshe@cmss.chinamobile.com)

USER root

# group/user

RUN groupadd hadoop

RUN groupadd ldap

RUN useradd hdfs

RUN useradd yarn

RUN useradd mapred

RUN useradd hive

RUN useradd hcat

RUN useradd hbase

RUN useradd tez

RUN useradd zookeeper

RUN useradd kms

# hadoop, hive, hbase

RUN yum install -y hbase tez hadoop snappy snappy-devel hadoop-lzo lzo hive hcatalog webhcat-tar-hive webhcat-tar-pig mysql-connector-java mysql mysql-server

#kylin 1.5.2

RUN curl -s http://10.133.17.27/software/tarball/apache-kylin-1.5.3-bc1.3.2.tar.gz | tar -xz -C /usr/local/

RUN cd /usr/local && ln -s ./apache-kylin-1.5.3-bc1.3.2 kylin

ENV KYLIN\_HOME /usr/local/kylin

#java

RUN curl -LO 'http://download.oracle.com/otn-pub/java/jdk/7u71-b14/jdk-7u71-linux-x64.rpm' -H 'Cookie: oraclelicense=accept-securebackup-cookie'

RUN rpm -i jdk-7u71-linux-x64.rpm

RUN rm jdk-7u71-linux-x64.rpm

ENV HADOOP\_CONF\_DIR /cmss/bch/bc1.3.2/hadoop/etc/hadoop

ENV HBASE\_CONF\_DIR /cmss/bch/bc1.3.2/hbase/conf

ENV HIVE\_CONF\_DIR /cmss/bch/bc1.3.2/hive/conf

ENV HADOOP\_HOME /cmss/bch/bc1.3.2/hadoop

ENV HBASE\_HOME /cmss/bch/bc1.3.2/hbase

ENV HIVE\_HOME /cmss/bch/bc1.3.2/hive

# Add configuration files

ADD conf/core-site.xml $HADOOP\_CONF\_DIR/core-site.xml

ADD conf/hdfs-site.xml $HADOOP\_CONF\_DIR/hdfs-site.xml

ADD conf/mapred-site.xml $HADOOP\_CONF\_DIR/mapred-site.xml

ADD conf/yarn-site.xml $HADOOP\_CONF\_DIR/yarn-site.xml

ADD conf/hbase-site.xml $HBASE\_CONF\_DIR/hbase-site.xml

ADD conf/hdfs-site.xml $HBASE\_CONF\_DIR/hdfs-site.xml

ADD conf/hive-site.xml $HIVE\_CONF\_DIR/hive-site.xml

ADD conf/mapred-site.xml $HIVE\_CONF\_DIR/mapred-site.xml

ADD conf/kylin.properties $KYLIN\_HOME/conf/kylin.properties

ENV JAVA\_HOME /usr/java/default

ENV PATH $PATH:$JAVA\_HOME/bin:$HADOOP\_HOME/bin:$HBASE\_HOME/bin:$HIVE\_HOME/bin

RUN rm /usr/bin/java && ln -s $JAVA\_HOME/bin/java /usr/bin/java

ADD ssh\_config /root/.ssh/config

RUN chmod 600 /root/.ssh/config && chown root:root /root/.ssh/config

# fix the 254 error code

RUN sed -i "/^[^#]\*UsePAM/ s/.\*/#&/" /etc/ssh/sshd\_config

RUN echo "UsePAM no" >> /etc/ssh/sshd\_config

RUN echo "Port 2122" >> /etc/ssh/sshd\_config

ADD bootstrap.sh /etc/bootstrap.sh

RUN chown root:root /etc/bootstrap.sh && chmod 700 /etc/bootstrap.sh

ADD hosts /usr/local/hosts

RUN cat /usr/local/hosts >> /etc/hosts

ENV BOOTSTRAP /etc/bootstrap.sh

ENV JAVA\_LIBRARY\_PATH /usr/hdp/2.2.9.0-3393/hadoop/lib/native:/usr/local/hadoop/lib/native:$JAVA\_LIBRARY\_PATH

# Kylin and Other ports

EXPOSE 7070 7443 49707 2122

CMD ["/etc/bootstrap.sh", "-d"]

**4.创建镜像和启动**

镜像创建命令

docker build -t fys/kylin ./

结果如下所示：

[root@sgserver006 kylin]# docker build -t fys/kylin ./

Sending build context to Docker daemon 98.82 kB

Sending build context to Docker daemon

Step 0 : FROM fys/centos

---> a75eea9df54e

Step 1 : MAINTAINER from CMSS.inc by fys(fengyongshe@cmss.chinamobile.com)

---> Using cache

---> c63706a75665

Step 2 : USER root

---> Using cache

---> 2aa0dc3fdac4

Step 3 : RUN groupadd hadoop

---> Using cache

---> 384f68e07c52

Step 4 : RUN groupadd ldap

---> Using cache

---> c918082286fe

Step 5 : RUN useradd hdfs

---> Using cache

---> 92d72f2b6727

Step 6 : RUN useradd yarn

---> Using cache

---> dc86c349a648

Step 7 : RUN useradd mapred

---> Using cache

---> 99dc9a1da663

Step 8 : RUN useradd hive

---> Using cache

---> ab226ba17965

Step 9 : RUN useradd hcat

---> Using cache

---> 8ef7332c2108

Step 10 : RUN useradd hbase

---> Using cache

---> 9911032309f2

Step 11 : RUN useradd tez

---> Using cache

---> 9143813babef

Step 12 : RUN useradd zookeeper

---> Using cache

---> b8827209c7e2

Step 13 : RUN useradd kms

---> Using cache

---> 70c9679e80f0

Step 14 : RUN yum install -y hbase tez hadoop snappy snappy-devel hadoop-lzo lzo hive hcatalog webhcat-tar-hive webhcat-tar-pig mysql-connector-java mysql mysql-server

---> Using cache

---> 24072a139be5

Step 15 : RUN curl -s http://10.133.17.27/software/tarball/apache-kylin-1.5.3-bc1.3.2.tar.gz | tar -xz -C /usr/local/

---> Using cache

---> c59277859a47

..................

结果查看

docker images

[root@sgserver006 kylin]# docker images

REPOSITORY TAG IMAGE ID CREATED VIRTUAL SIZE

fys/kylin latest dc843d8ead01 3 hours ago 2.523 GB

启动docker container

docker run -i -t -p 7070:7070 --net="host" fys/kylin /etc/bootstrap.sh -bash

结果如下：

[root@sgserver006 kylin]# docker run -i -t -p 7070:7070 --net="host" fys/kylin /etc/bootstrap.sh -bash

Generating SSH1 RSA host key: [ OK ]

Starting sshd: [ OK ]

KYLIN\_HOME is set to /usr/local/kylin

kylin.security.profile is set to testing

KYLIN\_HOME is set to /usr/local/kylin

16/11/29 19:44:16 WARN conf.HiveConf: HiveConf of name hive.heapsize does not exist

16/11/29 19:44:16 WARN conf.HiveConf: HiveConf of name hive.default.fileformat.managed

hbase dependency: /cmss/bch/bc1.3.2/hbase/lib/hbase-common-1.1.3-bc1.3.2.jar

KYLIN\_JVM\_SETTINGS is -Xms1024M -Xmx4096M -Xss256K -XX:MaxPermSize=128M -verbose:gc -XX:+PrintGCDetails -XX:+PrintGCDateStamps -Xloggc:/usr/local/kylin/logs/kylin.gc.27 -XX:+UseGCLogFileRotation -XX:NumberOfGCLogFiles=10 -XX:GCLogFileSize=64M

KYLIN\_DEBUG\_SETTINGS is not set, will not enable remote debuging

KYLIN\_LD\_LIBRARY\_SETTINGS is not set, Usually it's okay unless you want to specify your own native path

KYLIN\_REST\_ADDRESS not found, will use sgserver006:7070

A new Kylin instance is started by , stop it using "kylin.sh stop"

Please visit http://<ip>:7070/kylin

You can check the log at /usr/local/kylin/logs/kylin.log

**5.运行结果**

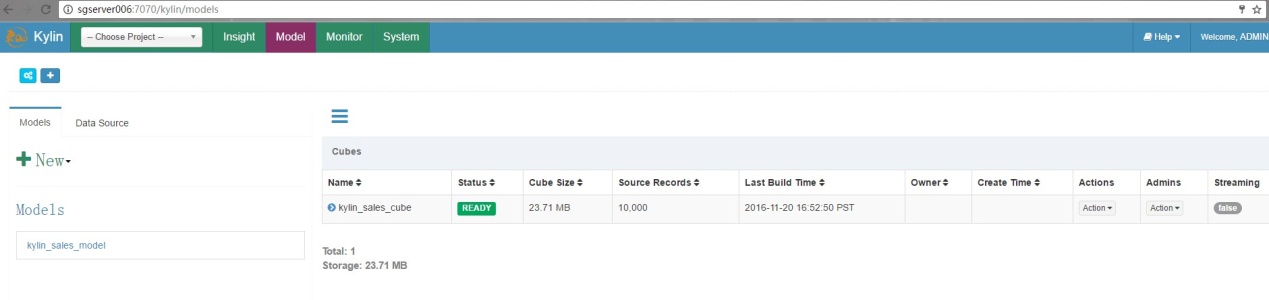
使用7070端口

[root@sgserver006 lib]# netstat -anp|grep 7070

tcp 0 0 :::7070 :::\* LISTEN 18669/java

通过页面上可以看到kylin的运行：

http://sgserver006:7070/kylin/models



**6.使用Slider启动kylin Container**

可以使用Slider将Kylin Container运行在Yarn中，步骤如下：

1）将镜像上传到私有仓库（私有仓库的构建不再介绍）

docker tag fys/kylin 10.133.46.153:5000/kylin

docker push 10.133.46.153:5000/kylin

2）在YARN NM的各个节点上pull kylin镜像（不是必须步骤，但是提前拉取出不用启动时拉取）

docker pull 10.133.46.153:5000/kylin

[root@sgserver006 ~]# docker images

REPOSITORY TAG IMAGE ID CREATED VIRTUAL SIZE

10.133.46.153:5000/kylin latest dc843d8ead01 20 hours ago 2.523 GB

3)定义slider docker文件，包括如下：

metainfo.xml

{

"schemaVersion":"1.0",

"application":{

"name":"KYLIN",

"components":[

{

"name":"KYLIN",

"type":"docker",

"dockerContainers":[

{

"name":"kylin",

"commandPath:":"/usr/bin/docker",

"image":"10.133.46.153:5000/kylin"

}

]

}

]

}

}

appConfig.json

{

"schema":"http://example.org/specification/v2.0.0",

"metadata":{},

"global":{},

"components":{

"KYLIN":{

"kylin.commandPath":"/usr/bin/docker",

"kylin.options":"-d -p 7070:7070 --net=host"

}

}

}

resources.json

{

"schema":"http://example.org/specification/v2.0.0",

"metadata":{},

"global":{},

"components":{

"slider-appmaster":{},

"KYLIN":{

"yarn.role.priority":"1",

"yarn.component.instances":"1",

"yarn.memory":"512"

}

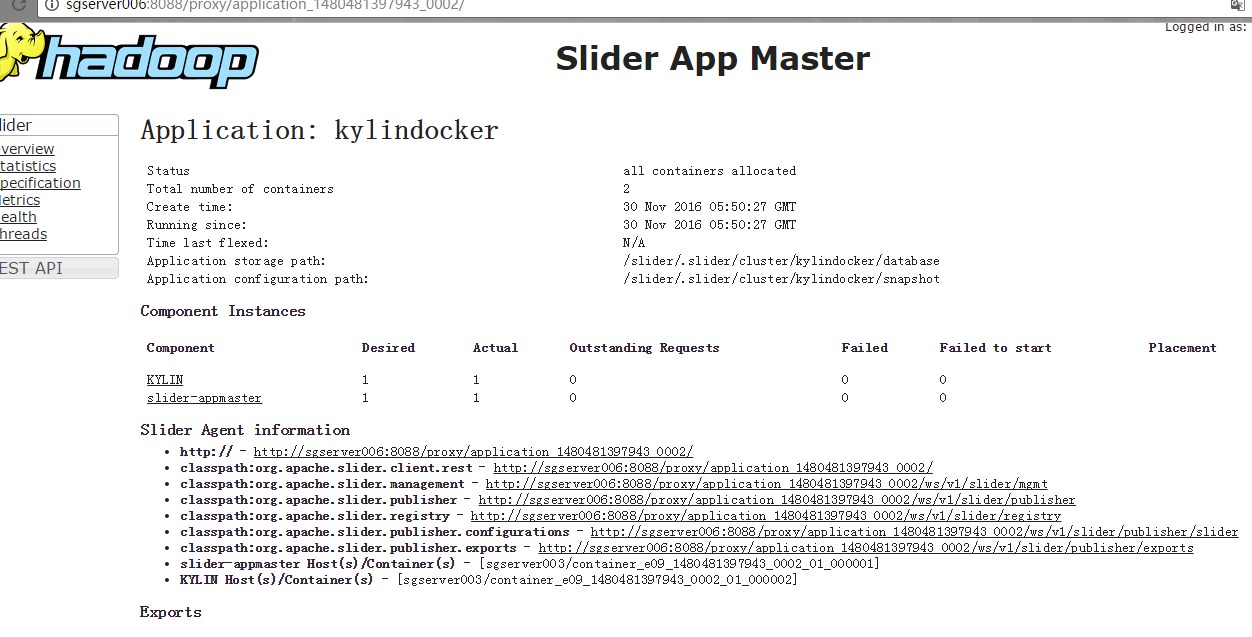
}

}

4)启动命令如下：

slider create kylindocker --template appConfig.json --resources resources.json --metainfo metainfo.json

5）启动后，通过Slider AM可以看到启动主机和其他信息：



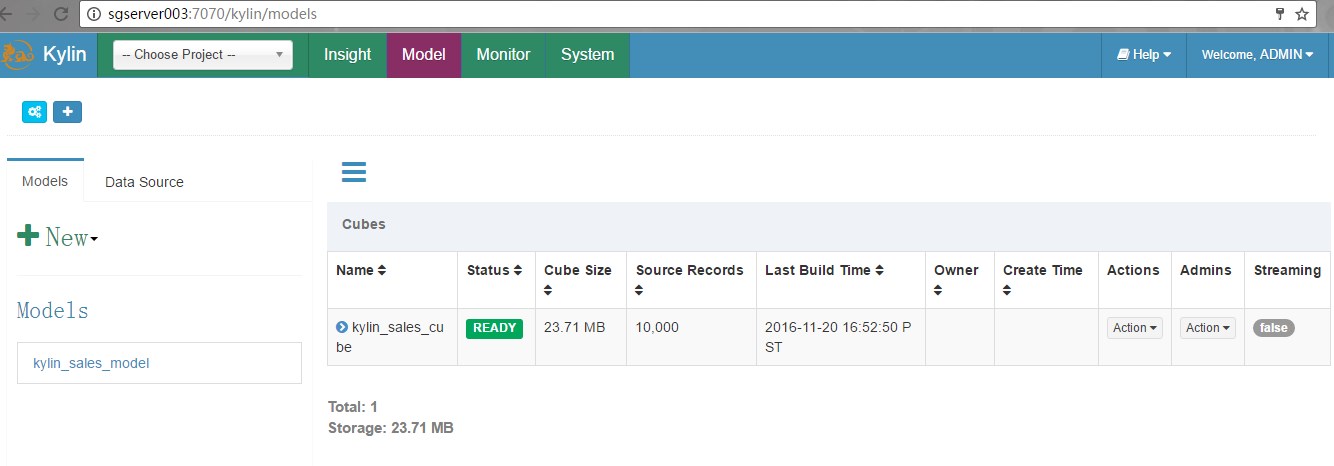
可以在server003上，看到启动的container

[root@sgserver003 ~]# docker ps

CONTAINER ID IMAGE COMMAND CREATED STATUS PORTS NAMES

0b8054b712b8 10.133.46.153:5000/kylin "/etc/bootstrap.sh - 12 hours ago Up 12 hours container\_e09\_1480481397943\_0002\_01\_000002

使用的主机端口为7070，通过界面访问如下：



注意：由于Docker默认是通过root执行命令，但是通过slider创建的Docker Container是通过slider用户启动的，要进行下面的操作：

chmod a+rw /var/run/docker.sock

修改docker.sock文件，让slider用户有权限执行docker命令