

On Hadoop CLI installation

On-Hadoop-CLI installation is for demo use, or for those who want to host their own web site to provide Kylin service. The scenario is depicted at <https://github.com/KylinOLAP/Kylin#on-hadoop-cli-installation>.

Except for some prerequisite software installations, the core of Kylin installation is accomplished by running a single script. After running the script, you will be able to build sample cube and query the tables behind the cubes via a unified web interface.

Environment

Running On-Hadoop-CLI installation requires you having access to a hadoop CLI, where you have full permissions to hdfs, hive, hbase and map-reduce. To make things easier we strongly recommend you starting with running Kylin on a hadoop sandbox, like <http://hortonworks.com/products/hortonworks-sandbox/>. In the following tutorial we'll go with **Hortonworks Sandbox 2.1** and **Cloudera QuickStart VM 5.1**.

To avoid permission issue, we suggest you using **root** account. The password for **Hortonworks Sandbox 2.1** is **hadoop**, for **Cloudera QuickStart VM 5.1** is **cloudera**.

We also suggest you using bridged mode instead of NAT mode in your virtual box settings. Bridged mode will assign your sandbox an independent IP so that you can avoid issues like <https://github.com/KylinOLAP/Kylin/issues/12>

Start Hadoop

For Hortonworks, ambari helps to launch hadoop:

```
ambari-agent start
ambari-server start
```

With both command successfully run you can go to ambari homepage at http://your_sandbox_ip:8080 (user:admin,password:admin) to check everything's status. **By default hortonworks ambari disables Hbase, you'll need manually start the **Hbase** service at ambari homepage.**

The screenshot shows the Ambari web interface for a sandbox environment. The top navigation bar includes 'Dashboard', 'Heatmaps', 'Services', 'Hosts', 'Jobs', and 'Admin'. The 'Services' tab is selected. On the left sidebar, a list of services is shown: HDFS, YARN, MapReduce2, Tez, HBase (highlighted with a red triangle icon), Hive, WebHCat, and Falcon. The main content area is divided into 'Summary' and 'Configs' tabs. The 'Summary' tab shows the status of HBase Master (Stopped) and RegionServers (0/1 live). A 'Service Actions' dropdown menu is open for the HBase service, showing options: Start, Stop, Restart All, Restart HBase RegionServers, Run Service Check, and Turn On Maintenance Mode. The 'Alerts and Health Checks' section on the right shows 'No alerts'.

For Cloudera, you can skip this step since they're by default activated.

Install tomcat

The latest tomcat can be found at <http://tomcat.apache.org/download-70.cgi>, we need the variable **CATALINA_HOME** exported.

```
cd ~
wget http://apache.cs.uu.nl/dist/tomcat/tomcat-7/v7.0.57/bin/apache-tomcat-7.0.57.tar.gz
tar -xzf apache-tomcat-7.0.57.tar.gz
export CATALINA_HOME=~/apache-tomcat-7.0.57
```

Install maven:

For Cloudera, you can skip this step since maven is by default installed.

The latest maven can be found at <http://maven.apache.org/download.cgi>, we create a symbolic so that `mvn` can be run anywhere.

```
cd ~
wget http://apache.proserve.nl/maven/maven-3/3.2.3/binaries/apache-maven-3.2.3-bin.tar.gz
tar -xzf apache-maven-3.2.3-bin.tar.gz
ln -s ~/apache-maven-3.2.3/bin/mvn /usr/bin/mvn
```

Install npm

Npm comes with latest versions of nodejs, just download nodejs from <http://nodejs.org/download/>, `npm` is located at the `/bin` folder of `nodejs`, we append the folder to `PATH` so that `npm` can be run anywhere.

```
cd ~
wget http://nodejs.org/dist/v0.10.32/node-v0.10.32-linux-x64.tar.gz
tar -xzf node-v0.10.32-linux-x64.tar.gz
export PATH=~/node-v0.10.32-linux-x64/bin:$PATH
```

Build cubes & Query tables

First clone the Kylin project to your local:

```
git clone https://github.com/KylinOLAP/Kylin.git
```

Go into the folder and run deploy script:

```
cd Kylin/
./deploy.sh
```

If you meet any problems, please check [FAQ](#) first. This script will help to:

1. Check your environment
2. Build Kylin artifacts with Maven
3. Generate some sample tables
4. Create empty cubes for these tables
5. Launch a one-stop website to build cubes & query tables

After successfully running the script, please navigate to http://your_sandbox_ip:7070 to build your cube and query it. The username/password is ADMIN/KYLIN

Build Cube in Cubes Tab

←

→

↺

🏠

mysandbox:8081/cubes

☆

🍷

👤

🔗

🔄

🛡️

☰

Kylin

Query

Cubes

Jobs

Tables

Admin

Help

Welcome, ADMIN

Project: -- All Projects --

+

Cube Name: Filter ...

🔍

Json Editor

+ Cube

Cubes

Name	Status	Cube Size	Source Records	Last Build Time	Owner	Create Time	Actions	Admins
test_kylin_cube_with_slr_empty		0.00 KB	0	1970-01-01 08:00:00			Action	Action
test_kylin_cube_without_slr_empty		0.00 KB	0	1970-01-01 08:00:00			Build Refresh Merge Disable	Action

Total: 2

Check Building Progress in Job Tab

Kylin

Query

Cubes

Jobs

Tables

Admin

Help

Welcome, ADMIN

Project: -- All Projects --

+

Cube Name: Filter ...

🔍

Job Status: new pending running finished error discarded

🔄

Jobs

Job Name	Cube	Progress	Last Modified Time	Duration	Actions
test_kylin_cube_without_slr_empty - FULL_BUILD - BUILD - PDT 2014-10-09 23:08:32	test_kylin_cube_without_slr_empty		2014-10-10 14:08:32	0.00 mins	Action
test_kylin_cube_with_slr_empty - 19700101000000_20141010000000 - BUILD - PDT 2014-10-09 23:08:21	test_kylin_cube_with_slr_empty		2014-10-10 14:08:21	0.00 mins	Action

Total: 2

Detail Information

Job Name

test_kylin_cube_without_slr_empty - FULL_BUILD - BUILD - PDT 2014-10-09 23:08:32

Job ID

7ad65d3d-3bb1-49d4-8929-fded242542fd

Status

PENDING

Duration

0.00 mins

MapReduce Waiting

0.00 mins

Start

#1 Step Name: Create Intermediate Flat Hive Table
Duration: 0 seconds

Query Tables in Query Tab

Query

Cubes

Jobs

Tables

Admin

Help

Welcome, ADMIN

Project

onlyinner

EDW

TEST_CAL_DT

TEST_CATEGORY_GROUPINGS

TEST_KYLIN_FACT

TEST_SELLER_TYPE_DIM

TEST_SITES

New Query

Saved Queries

Query History

1

select * from test_kylin_fact

Tips: Ctrl+Shift+Space or Alt+Space(Windows), Command+Option+Space(Mac) to list tables/columns in query box.

Project: onlyinner

LIMIT 50000

Submit

Results

1 ✓ x

Status: All

Query String

Start Time: 2014-10-11 14:32:31 Duration: 4.70s

Rerun

Save

Status: Success

Project: onlyinner

Cubes: test_kylin_cube_with_slr_empty

Results (10000)

Visualization

Export

Drag a column header here and drop it to group by that column.

PRICE	SLR_SEGMENT...	LSTG_SITE_ID	SELLER_ID	CAL_DT	LSTG_FORMAT...
22.72578007367...	16	0	10000000	2012-06-12	Auction
34.75382512287...	14	0	10000000	2012-06-13	ABIN
20.67062683914...	14	0	10000000	2012-06-22	FP-GTC
78.45679580863...	15	15	10000000	2012-11-08	Auction
18.42639494247...	14	0	10000000	2012-12-23	ABIN
70.29527893881...	12	0	10000000	2013-01-07	Auction
19.95528576305...	12	0	10000000	2013-04-29	Others
65.48488420958...	11	0	10000000	2013-05-26	FP-GTC

What if I closed my VM?

If you shut down the VM and restarted it, Kylin will not automatically start. Depending on whether you succeed in running the deploy.sh, you should:

If running deploy.sh failed last time

1. Kill it if any tomcat instance exist
2. Start over again

If running deploy.sh succeeded last time

1. Kill it if any tomcat instance exist

2. Make sure Hbase is running
3. run `export CATALINA_HOME=~/.apache-tomcat-7.0.56`
4. run `sudo -i "${CATALINA_HOME}/bin/startup.sh"`