



Passing current date to Hive query

1. Create a Hive script with a parameter and save to file system (CURRENT_DATE in this example):

```
set CURRENT_TIMEZONE='GMT-02:00';

set FIRST_REPORT_DATE=date_sub('${CURRENT_DATE}', 2);
set LAST_REPORT_DATE=date_sub('${CURRENT_DATE}', 1);

set INDEX_BEGIN=date_sub(${hiveconf:FIRST_REPORT_DATE}, 1);
set INDEX_END=date_add(${hiveconf:LAST_REPORT_DATE}, 1);

SELECT count(*) FROM staging_db.mobile_kill_appinfo
WHERE from_utc_timestamp(time, ${hiveconf:CURRENT_TIMEZONE}) >= ${hiveconf:FIRST_REPORT_DATE}
AND from_utc_timestamp(time, ${hiveconf:CURRENT_TIMEZONE}) <= ${hiveconf:LAST_REPORT_DATE}
AND concat(year, '-', month, '-', day) >= ${hiveconf:INDEX_BEGIN}
AND concat(year, '-', month, '-', day) <= ${hiveconf:INDEX_END}
```

2. Create a Hue (Oozie) workflow with Hive job and add the script as well:

Script name

/user/bi.user/hive.hql

Prepare

Add delete

Add mkdir

3. Add a parameter to job

| Params | Type | Value | |
|-----------|-------------------------------|-------|--------|
| param | CURRENT_DATE=\${CURRENT_DATE} | .. | Delete |
| Add param | | | |

4. Create a coordinator for

workflow (schedule) and specify an Oozie to get and format date

```
<code>${coord:formatTime(coord:actualTime(), 'yyyy-MM-dd')}</code>
```

Coordinator Editor : test-hive

Step 1: General Step 2: Frequency Step 3: Inputs Step 4: Outputs Step 5: Advanced settings

Advanced settings

Oozie parameters

| Name | Value | |
|--------------------------|-------|--------|
| oozie.use.system.libpath | true | Delete |
| Add | | |

Workflow properties

| Name | Value | |
|--------------|--|--------|
| CURRENT_DATE | \${coord:formatTime(coord:actualTime(), 'yyyy-MM-dd')} | Delete |
| Add | | |

Timeout

Concurrency

1

Throttle

Execution

LAST_ONLY (discards all old)

Advanced

Back

Next

Save coordinator

5. Don't forget to specify hive-config.xml (hive-site.xml) as a Job XML and as a file to add to job. Don't use name hive-site.xml - will not work due to internal bug.