

# Apache HAWQ install

김하늘 hskimsky@gmail.com

## 목표와 전제조건

Apache Hadoop 2.7.2 버전 위에 [HAWQ 2.0 dev 버전](#)을 resource manager 를 yarn 으로 설치하고 HAWQ 에서 hdfs 와 web 의 파일을 읽어서 external table 을 만드는 것을 목표로 합니다.

Apache HAWQ 를 build, install 하는 순서와 명령어를 아주 자세하게 적어놓았습니다.

hadoop install 방법은 포함하지 않았고 pxf install 을 포함하고 있습니다.

먼저 캡터별 전체 실행 스크립트를 적어놓았고 그 뒤에 부분별 실행 스크립트를 적어놓았고 그 다음에 실행 한 결과 로그를 적어놓았습니다.

문서는 asciidoc 으로 작성하였으며 [repository](#) 내에 [asciidoc 파일](#) 및 [pdf 파일](#)도 포함되어 있습니다.

잘못된 점이 있으면 [hskimsky@gmail.com](mailto:hskimsky@gmail.com) 으로 메일 주시기 바랍니다.

빌드 성공을 기원합니다.

## System 구성

Linux, Python, JAVA, Maven, Hadoop 은 미리 설치되어 있다고 가정합니다.

Hadoop 의 super user 는 root 가 아니어야 합니다.

왜냐하면 HAWQ 를 root 로 설치할 수 없습니다.

Table 1. Machine 구성

Program	Version
Linux	CentOS 6.7
Python	2.6.6
JAVA	Oracle JDK 1.7.0_80
Maven	Apache Maven 3.3.9
Hadoop	Apache Hadoop 2.7.2
HAWQ	Apache HAWQ 2.0.0.0 dev
PXF	3.0.0

Table 2. ROLE 구성

FQDN	ROLE
hawq1.apache.local	Namenode, Secondary Namenode, ResourceManager, HAWQ Standby, PXF
hawq2.apache.local	HAWQ Master, PXF
hawq3.apache.local	Datanode, HAWQ Segment, PXF
hawq4.apache.local	Datanode, HAWQ Segment, PXF

Table 3. PORT 구성

ROLE	PORT
Namenode	8020
ResourceManager	8050
ResourceManager Scheduler	8030
HAWQ Master	5432
PXF	51200

### /etc/hosts

각자 환경에 맞게 설정합니다.

```
127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
::1        localhost localhost.localdomain localhost6 localhost6.localdomain6

192.168.98.166    hawq1.apache.local    hawq1
192.168.98.167    hawq2.apache.local    hawq2
192.168.98.168    hawq3.apache.local    hawq3
192.168.98.169    hawq4.apache.local    hawq4
```

## /etc/bashrc

각자 환경에 맞게 설치 및 설정합니다.

```
export JAVA_HOME=/usr/local/java/default
export M2_HOME=/usr/local/maven/default

export HADOOP_HOME=/usr/local/hadoop/default
export HADOOP_MAPRED_HOME=${HADOOP_HOME}
export HADOOP_COMMON_HOME=${HADOOP_HOME}
export HADOOP_HDFS_HOME=${HADOOP_HOME}
export YARN_HOME=${HADOOP_HOME}
export HADOOP_COMMON_LIB_NATIVE_DIR=${HADOOP_HOME}/lib/native
export HADOOP_OPTS="-Djava.library.path=${HADOOP_HOME}/lib"
export HADOOP_CLASSPATH=${JAVA_HOME}/lib/tools.jar
export CLASSPATH=${HADOOP_HOME}/share/hadoop/common/hadoop-common-
2.7.2.jar:${HADOOP_HOME}/share/hadoop/mapreduce/:${HADOOP_HOME}/share/hadoop/common/lib/commons-cli-1.2.jar

export PATH=${JAVA_HOME}/bin:${M2_HOME}/bin:${HADOOP_HOME}/bin:${HADOOP_HOME}/sbin:${PATH}
```

IMPORTANT

모든 서버끼리 **hadoop** 의 **super user** 로 키 교환 필수  
 방화벽 해제 필수  
 시간 동기화 필수 (# **rdate -s time.bora.net**)

## <HADOOP\_HOME>/etc/hadoop/core-site.xml

각자 환경에 맞게 설정합니다.

```
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://hawq1.apache.local:8020</value>
  </property>
</configuration>
```

## <HADOOP\_HOME>/etc/hadoop/hdfs-site.xml

각자 환경에 맞게 설정합니다.

```
<configuration>
  <property>
    <name>dfs.replication</name>
    <value>3</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/hadoop_tmp/hdfs/namenode</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/hadoop_tmp/hdfs/datanode</value>
  </property>
</configuration>
```

〈HADOOP\_HOME〉/etc/hadoop/mapred-site.xml

각자 환경에 맞게 설정합니다.

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

〈HADOOP\_HOME〉/etc/hadoop/slaves

각자 환경에 맞게 설정합니다.

```
hawq3.apache.local
hawq4.apache.local
```

〈HADOOP\_HOME〉/etc/hadoop/yarn-site.xml

각자 환경에 맞게 설정합니다.

```

<configuration>
<!-- Site specific YARN configuration properties -->
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.resourcemanager.scheduler.class</name>
    <value>org.apache.hadoop.yarn.server.resourcemanager.scheduler.capacity.CapacityScheduler</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
    <value>org.apache.hadoop.mapred.ShuffleHandler</value>
  </property>
  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>hawq1.apache.local</value>
  </property>
  <property>
    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>hawq1.apache.local:8025</value>
  </property>
  <property>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>hawq1.apache.local:8030</value>
  </property>
  <property>
    <name>yarn.resourcemanager.address</name>
    <value>hawq1.apache.local:8050</value>
  </property>
  <property>
    <name>yarn.resourcemanager.webapp.address</name>
    <value>hawq1.apache.local:8088</value>
  </property>
</configuration>

```

#### IMPORTANT

모든 서버끼리 **hadoop** 의 **super user** 로 키 교환 필수  
 방화벽 해제 필수  
 시간 동기화 필수 (# **rdate -s time.bora.net**)

## Hadoop install

<http://inthound.blogspot.kr/2015/04/how-to-install-hadoop-260-in-ubuntu.html>

## HAWQ install

### Apache HAWQ build

HAWQ 가 설치 될 모든 서버(master, standby, segment)에서 작업해야 합니다.  
 이제 작업은 특별한 지침이 없으면 모두 root 로 작업합니다.

### library download

```

mkdir -p ~/Downloads/hawq
cd ~/Downloads/hawq
yum install -y epel-release
yum repolist
wget http://sourceforge.net/projects/boost/files/boost/1.56.0/boost_1_56_0.tar.bz2
wget http://archive.apache.org/dist/thrift/0.9.1/thrift-0.9.1.tar.gz
wget https://github.com/google/protobuf/archive/v2.5.0.zip
wget http://www.execve.net/curl/curl-7.49.0.tar.gz
wget https://bootstrap.pypa.io/get-pip.py
python get-pip.py

yum install -y json-c-devel
yum install -y gcc-c++
yum install -y gperf
yum install -y snappy-devel
yum install -y bzip2-devel
yum install -y python-devel
yum install -y libevent-devel
yum install -y krb5-devel
yum install -y libuuid-devel
yum install -y libgsasl-devel
yum install -y libxml2-devel
yum install -y readline-devel
yum install -y openssl-devel
yum install -y bison
yum install -y apr-devel
yum install -y libyaml-devel
yum install -y flex
yum install -y lcov

```

## upgrade gcc

```

cd /etc/yum.repos.d
wget http://people.centos.org/tru/devtools-2/devtools-2.repo -O /etc/yum.repos.d/devtools-2.repo
yum install -y devtoolset-2-gcc devtoolset-2-binutils devtoolset-2-gcc-c++
scl enable devtoolset-2 bash
gcc --version
g++ --version

```

## upgrade

```

[root@hawq2 incubator-hawq]# cd /etc/yum.repos.d
[root@hawq2 yum.repos.d]# wget http://people.centos.org/tru/devtools-2/devtools-2.repo -O /etc/yum.repos.d/devtools-2.repo
[root@hawq2 yum.repos.d]# yum install -y devtoolset-2-gcc devtoolset-2-binutils devtoolset-2-gcc-c++
[root@hawq2 yum.repos.d]# scl enable devtoolset-2 bash

```

## check version

```

[root@hawq2 yum.repos.d]# gcc --version
gcc (GCC) 4.8.2 20140120 (Red Hat 4.8.2-15)
Copyright (C) 2013 Free Software Foundation, Inc.
This is free software; see the source for copying conditions. There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.

[root@hawq2 yum.repos.d]# g++ --version
g++ (GCC) 4.8.2 20140120 (Red Hat 4.8.2-15)
Copyright (C) 2013 Free Software Foundation, Inc.
This is free software; see the source for copying conditions. There is NO
warranty; not even for MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.

[root@hawq2 yum.repos.d]#

```

## install boost

```
cd ~/Downloads/hawq/  
bunzip2 boost_1_56_0.tar.bz2  
tar xf boost_1_56_0.tar  
cd boost_1_56_0  
./bootstrap.sh  
./b2
```

## bootstrap

```
[root@hawq2 yum.repos.d]# cd ~/Downloads/hawq/  
[root@hawq2 hawq]# bunzip2 boost_1_56_0.tar.bz2  
[root@hawq2 hawq]# tar xf boost_1_56_0.tar  
[root@hawq2 hawq]# cd boost_1_56_0  
[root@hawq2 boost_1_56_0]# ./bootstrap.sh  
Building Boost.Build engine with toolset gcc... tools/build/src/engine/bin.linuxx86_64/b2  
Detecting Python version... 2.6  
Detecting Python root... /usr  
Unicode/ICU support for Boost.Regex?... not found.  
Generating Boost.Build configuration in project-config.jam...
```

Bootstrapping is done. To build, run:

```
./b2
```

To adjust configuration, edit 'project-config.jam'.  
Further information:

- Command line help:  
./b2 --help
- Getting started guide:  
[http://www.boost.org/more/getting\\_started/unix-variants.html](http://www.boost.org/more/getting_started/unix-variants.html)
- Boost.Build documentation:  
<http://www.boost.org/boost-build2/doc/html/index.html>

```
[root@hawq2 boost_1_56_0]#
```

## b2

```
[root@hawq2 boost_1_56_0]# ./b2
```

Building the Boost C++ Libraries.

Performing configuration checks

```
- 32-bit           : no
- 64-bit           : yes
- arm              : no
- mips1            : no
- power            : no
- sparc            : no
- x86              : yes
- lockfree boost::atomic_flag : yes
- has_icu builds   : no
```

...생략

The Boost C++ Libraries were successfully built!

The following directory should be added to compiler include paths:

```
/root/Downloads/hawq/boost_1_56_0
```

The following directory should be added to linker library paths:

```
/root/Downloads/hawq/boost_1_56_0/stage/lib
```

```
[root@hawq2 boost_1_56_0]#
```

## ldconfig

```
ldconfig -p /root/Downloads/hawq/boost_1_56_0/stage/lib
```

```
[root@hawq2 boost_1_56_0]# ldconfig -p /root/Downloads/hawq/boost_1_56_0/stage/lib
678 libs found in cache `/etc/ld.so.cache'
libz.so.1 (libc6,x86-64) => /lib64/libz.so.1
libz.so (libc6,x86-64) => /usr/lib64/libz.so
libyaml-0.so.2 (libc6,x86-64) => /usr/lib64/libyaml-0.so.2
libxul.so (libc6,x86-64) => /usr/lib64/xulrunner/libxul.so
libxtables.so.4 (libc6,x86-64) => /lib64/libxtables.so.4
libxslt.so.1 (libc6,x86-64) => /usr/lib64/libxslt.so.1
libxpc.so (libc6,x86-64) => /usr/lib64/xulrunner/libxpc.so
...생략
libFLAC++.so.6 (libc6,x86-64) => /usr/lib64/libFLAC++.so.6
libEGL.so.1 (libc6,x86-64) => /usr/lib64/libEGL.so.1
libDeployPkg.so (libc6,x86-64) => /usr/lib/vmware-tools/lib64/libDeployPkg.so/libDeployPkg.so
libDeployPkg.so (libc6) => /usr/lib/vmware-tools/lib32/libDeployPkg.so/libDeployPkg.so
libBrokenLocale.so.1 (libc6,x86-64, OS ABI: Linux 2.6.18) => /lib64/libBrokenLocale.so.1
libBrokenLocale.so (libc6,x86-64, OS ABI: Linux 2.6.18) => /usr/lib64/libBrokenLocale.so
ld-linux-x86-64.so.2 (libc6,x86-64) => /lib64/ld-linux-x86-64.so.2
[root@hawq2 boost_1_56_0]#
```

## BOOST\_ROOT 등록

```
echo "export BOOST_ROOT=/root/Downloads/hawq/boost_1_56_0" >> /etc/bashrc
source /etc/bashrc
```

## install thrift



```
cd ..
tar xzf thrift-0.9.1.tar.gz
cd thrift-0.9.1
./configure --without-tests && make && make install
```

IMPORTANT | 꼭 boost 설치 후 실행해야 함

## install

```
[root@hawq2 boost_1_56_0]# cd ..
[root@hawq2 hawq]# tar xzf thrift-0.9.1.tar.gz
[root@hawq2 hawq]# cd thrift-0.9.1
[root@hawq2 thrift-0.9.1]# ./configure --without-tests && make && make install
checking for a BSD-compatible install... /usr/bin/install -c
checking whether build environment is sane... yes
checking for a thread-safe mkdir -p... /bin/mkdir -p
checking for gawk... gawk
checking whether make sets $(MAKE)... yes
checking how to create a ustar tar archive... gnutar
checking for pkg-config... /usr/bin/pkg-config
...생략
checking for boostlib >= 1.40.0... configure: We will use a staged boost library from /root/Downloads/hawq/boost_1_56_0
yes
...생략
make[1]: Leaving directory '/root/Downloads/hawq/thrift-0.9.1/tutorial'
make[1]: Entering directory '/root/Downloads/hawq/thrift-0.9.1'
make[2]: Entering directory '/root/Downloads/hawq/thrift-0.9.1'
make[2]: Nothing to be done for 'install-exec-am'.
make[2]: Nothing to be done for 'install-data-am'.
make[2]: Leaving directory '/root/Downloads/hawq/thrift-0.9.1'
make[1]: Leaving directory '/root/Downloads/hawq/thrift-0.9.1'
[root@hawq2 thrift-0.9.1]#
```

## install protocol buffer

```
cd ..
unzip v2.5.0.zip
yum groupinstall -y "Development Tools"
cd protobuf-2.5.0/
./autogen.sh
./configure && make && make install
protoc --version
```

## unzip

```
[root@hawq2 hawq]# cd ..
[root@hawq2 hawq]# unzip v2.5.0.zip
Archive: v2.5.0.zip
774d630bde574f5fcb6dae6eaa0f91f7bc12967
  creating: protobuf-2.5.0/
    inflating: protobuf-2.5.0/CHANGES.txt
    inflating: protobuf-2.5.0/CONTRIBUTORS.txt
...생략
    inflating: protobuf-2.5.0/vsprojects/protoc.vcproj
    inflating: protobuf-2.5.0/vsprojects/readme.txt
    inflating: protobuf-2.5.0/vsprojects/test_plugin.vcproj
    inflating: protobuf-2.5.0/vsprojects/tests.vcproj
[root@hawq2 hawq]# cd protobuf-2.5.0/
```

## Development Tools

```
[root@hawq2 protobuf-2.5.0]# yum groupinstall -y "Development Tools"
Loaded plugins: fastestmirror, refresh-packagekit, security
Setting up Group Process
Loading mirror speeds from cached hostfile
* base: ftp.daumkakao.com
* epel: mirror.premi.st
* extras: ftp.daumkakao.com
* updates: ftp.daumkakao.com
base/group_gz
| 226 kB      00:00
epel/group_gz
| 150 kB      00:00
Package flex-2.5.35-9.el6.x86_64 already installed and latest version
Package gcc-4.4.7-17.el6.x86_64 already installed and latest version
Package patch-2.6-6.el6.x86_64 already installed and latest version
Package 1:pkgconfig-0.23-9.1.el6.x86_64 already installed and latest version
Package bison-2.4.1-5.el6.x86_64 already installed and latest version
Package gcc-c++-4.4.7-17.el6.x86_64 already installed and latest version
Resolving Dependencies
--> Running transaction check
--> Package autoconf.noarch 0:2.63-5.1.el6 will be installed
--> Package automake.noarch 0:1.11.1-4.el6 will be installed
--> Package binutils.x86_64 0:2.20.51.0.2-5.43.el6 will be updated
--> Package binutils.x86_64 0:2.20.51.0.2-5.44.el6 will be an update
--> Package byacc.x86_64 0:1.9.20070509-7.el6 will be installed
...생략
Dependency Updated:
  elfutils-libelf.x86_64 0:0.164-2.el6                elfutils-libs.x86_64 0:0.164-2.el6                perl.x86_64
4:5.10.1-141.el6_7.1
  perl-Module-Pluggable.x86_64 1:3.90-141.el6_7.1      perl-Pod-Escapes.x86_64 1:1.04-141.el6_7.1            perl-Pod-
Simple.x86_64 1:3.13-141.el6_7.1
  perl-libs.x86_64 4:5.10.1-141.el6_7.1                perl-version.x86_64 3:0.77-141.el6_7.1                rpm.x86_64
0:4.8.0-55.el6
  rpm-libs.x86_64 0:4.8.0-55.el6                        rpm-python.x86_64 0:4.8.0-55.el6                        systemtap-
runtime.x86_64 0:2.9-4.el6

Complete!
[root@hawq2 protobuf-2.5.0]#
```

autogen.sh

```
[root@hawq2 protobuf-2.5.0]# ./autogen.sh
Google Test not present. Fetching gtest-1.5.0 from the web...
  % Total    % Received % Xferd  Average Speed   Time    Time     Current
                                 Dload  Upload   Total   Spent    Left   Speed
100 638k 100 638k    0     0  850k      0 --:--:-- --:--:-- --:--:-- 1094k
+ sed -i -e 's/RuntimeLibrary="5"/RuntimeLibrary="3"/g;
           s/RuntimeLibrary="4"/RuntimeLibrary="2"/g;' gtest/msvc/gtest_main-md.vcproj gtest/msvc/gtest_main.vcproj
gtest/msvc/gtest-md.vcproj gtest/msvc/gtest_prod_test-md.vcproj gtest/msvc/gtest_prod_test.vcproj
gtest/msvc/gtest_unittest-md.vcproj gtest/msvc/gtest_unittest.vcproj gtest/msvc/gtest.vcproj
+ autoreconf -f -i -Wall,no-obsolete
libtoolize: putting auxiliary files in AC_CONFIG_AUX_DIR, 'build-aux'.
libtoolize: copying file 'build-aux/ltmain.sh'
libtoolize: Consider adding 'AC_CONFIG_MACRO_DIR([m4])' to configure.ac and
libtoolize: rerunning libtoolize, to keep the correct libtool macros in-tree.
libtoolize: Consider adding '-I m4' to ACLOCAL_AMFLAGS in Makefile.am.
libtoolize: putting auxiliary files in `.'.
libtoolize: copying file `./ltmain.sh'
libtoolize: putting macros in AC_CONFIG_MACRO_DIR, `m4'.
libtoolize: copying file `m4/libtool.m4'
libtoolize: copying file `m4/ltoptions.m4'
libtoolize: copying file `m4/ltsugar.m4'
libtoolize: copying file `m4/ltversion.m4'
libtoolize: copying file `m4/lt~obsolete.m4'
configure.ac:30: installing './config.guess'
configure.ac:30: installing './config.sub'
configure.ac:32: installing './install-sh'
configure.ac:32: installing './missing'
src/Makefile.am: installing './depcomp'
+ rm -rf autom4te.cache config.h.in~
+ exit 0
[root@hawq2 protobuf-2.5.0]#
```

## configure

```
[root@hawq2 protobuf-2.5.0]# ./configure && make && make install
checking whether to disable maintainer-specific portions of Makefiles... yes
checking build system type... x86_64-unknown-linux-gnu
checking host system type... x86_64-unknown-linux-gnu
checking target system type... x86_64-unknown-linux-gnu
checking for a BSD-compatible install... /usr/bin/install -c
...생략
/usr/bin/install -c -m 644 google/protobuf/stubs/atomicops.h google/protobuf/stubs/atomicops_internals_arm_gcc.h
google/protobuf/stubs/atomicops_internals_arm_qnx.h google/protobuf/stubs/atomicops_internals_atomicword_compat.h
google/protobuf/stubs/atomicops_internals_macosx.h google/protobuf/stubs/atomicops_internals_mips_gcc.h
google/protobuf/stubs/atomicops_internals_pnacl.h google/protobuf/stubs/atomicops_internals_x86_gcc.h
google/protobuf/stubs/atomicops_internals_x86_msvc.h google/protobuf/stubs/common.h
google/protobuf/stubs/platform_macros.h google/protobuf/stubs/once.h google/protobuf/stubs/template_util.h
google/protobuf/stubs/type_traits.h '/usr/local/include/google/protobuf/stubs'
make[3]: Leaving directory '/root/Downloads/hawq/protobuf-2.5.0/src'
make[2]: Leaving directory '/root/Downloads/hawq/protobuf-2.5.0/src'
make[1]: Leaving directory '/root/Downloads/hawq/protobuf-2.5.0/src'
[root@hawq2 protobuf-2.5.0]#
[root@hawq2 protobuf-2.5.0]# protoc --version
libprotoc 2.5.0
[root@hawq2 protobuf-2.5.0]#
```

## install curl

```
cd ..
tar xzf curl-7.49.0.tar.gz
cd curl-7.49.0
./configure && make && make install
```

## untar

```
[root@hawq2 protobuf-2.5.0]# cd ..
[root@hawq2 hawq]# tar xzf curl-7.49.0.tar.gz
[root@hawq2 hawq]# cd curl-7.49.0
[root@hawq2 curl-7.49.0]#
```

## configure

```
[root@hawq2 curl-7.49.0]# ./configure && make && make install
checking whether to enable maintainer-specific portions of Makefiles... no
checking whether to enable debug build options... no
checking whether to enable compiler optimizer... (assumed) yes
checking whether to enable strict compiler warnings... no
checking whether to enable compiler warnings as errors... no
checking whether to enable curl debug memory tracking... no
checking whether to enable hiding of library internal symbols... yes
checking whether to enable c-ares for DNS lookups... no
checking whether to disable dependency on -lrt... (assumed no)
    ;;
*
no
checking for path separator... :
checking for sed... /bin/sed
...생략
/bin/mkdir -p '/usr/local/share/man/man1'
/usr/bin/install -c -m 644 curl.1 curl-config.1 '/usr/local/share/man/man1'
make[6]: Leaving directory '/root/Downloads/hawq/curl-7.49.0/docs'
make[5]: Leaving directory '/root/Downloads/hawq/curl-7.49.0/docs'
make[4]: Leaving directory '/root/Downloads/hawq/curl-7.49.0/docs'
make[3]: Leaving directory '/root/Downloads/hawq/curl-7.49.0'
make[2]: Leaving directory '/root/Downloads/hawq/curl-7.49.0'
make[1]: Leaving directory '/root/Downloads/hawq/curl-7.49.0'
[root@hawq2 curl-7.49.0]#
```

## upgrade cmake

```
cd ..
wget https://cmake.org/files/v3.0/cmake-3.0.0.tar.gz
tar xzf cmake-3.0.0.tar.gz
cd cmake-3.0.0
./bootstrap
./configure && make && make install
```

```
[root@hawq2 incubator-hawq]# cd ..
[root@hawq2 hawq]# https://cmake.org/files/v3.0/cmake-3.0.0.tar.gz
[root@hawq2 hawq]# tar xzf cmake-3.0.0.tar.gz
[root@hawq2 hawq]# cd cmake-3.0.0
[root@hawq2 cmake-3.0.0]# ./bootstrap
```

```
-----
CMake 3.0.0, Copyright 2000-2014 Kitware, Inc.
Found GNU toolchain
C compiler on this system is: gcc
C++ compiler on this system is: g++
Makefile processor on this system is: gmake
g++ is GNU compiler
g++ has setenv
g++ has unsetenv
g++ does not have environ in stdlib.h
g++ has STL in std:: namespace
g++ has ANSI streams
g++ has streams in std:: namespace
g++ has sstream
g++ has operator!=(string, char*)
```

```

g++ has stl iterator_traits
g++ has standard template allocator
g++ has allocator<>::rebind<>
g++ does not have non-standard allocator<>::max_size argument
g++ has stl containers supporting allocator objects
g++ has stl wstring
g++ has header cstdint
g++ requires template friends to use <>
g++ supports member templates
g++ has standard template specialization syntax
g++ has argument dependent lookup
g++ has struct stat with st_mtim member
g++ has ios::binary openmode
g++ has ANSI for scoping
-----
g++ -I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -I/root/Downloads/hawq/cmake-3.0.0/Source
-I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -c /root/Downloads/hawq/cmake-
3.0.0/Source/cmExportTryCompileFileGenerator.cxx -o cmExportTryCompileFileGenerator.o
g++ -I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -I/root/Downloads/hawq/cmake-3.0.0/Source
-I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -c /root/Downloads/hawq/cmake-3.0.0/Source/cmExportSet.cxx -o
cmExportSet.o
...생략
-- Looking for elf.h
-- Looking for elf.h - found
-- Looking for a Fortran compiler
-- Looking for a Fortran compiler - /usr/bin/f95
-- Performing Test run_pic_test
-- Performing Test run_pic_test - Success
-- Performing Test run_inlines_hidden_test
-- Performing Test run_inlines_hidden_test - Success
-- Configuring done
-- Generating done
-- Build files have been written to: /root/Downloads/hawq/cmake-3.0.0
-----
CMake has bootstrapped. Now run gmake.
[root@hawq2 cmake-3.0.0]#
[root@hawq2 cmake-3.0.0]# ./configure && make && make install
-----
CMake 3.0.0, Copyright 2000-2014 Kitware, Inc.
Found GNU toolchain
C compiler on this system is: gcc
C++ compiler on this system is: g++
Makefile processor on this system is: gmake
g++ is GNU compiler
g++ has setenv
g++ has unsetenv
g++ does not have environ in stdlib.h
g++ has STL in std:: namespace
g++ has ANSI streams
g++ has streams in std:: namespace
g++ has sstream
g++ has operator!=(string, char*)
g++ has stl iterator_traits
g++ has standard template allocator
g++ has allocator<>::rebind<>
g++ does not have non-standard allocator<>::max_size argument
g++ has stl containers supporting allocator objects
g++ has stl wstring
g++ has header cstdint
g++ requires template friends to use <>
g++ supports member templates
g++ has standard template specialization syntax
g++ has argument dependent lookup
g++ has struct stat with st_mtim member
g++ has ios::binary openmode
g++ has ANSI for scoping
-----
g++ -I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -I/root/Downloads/hawq/cmake-3.0.0/Source
-I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -c /root/Downloads/hawq/cmake-3.0.0/Source/cmStandardIncludes.cxx -o
cmStandardIncludes.o

```

```
g++ -I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -I/root/Downloads/hawq/cmake-3.0.0/Source
-I/root/Downloads/hawq/cmake-3.0.0/Bootstrap.cmk -c /root/Downloads/hawq/cmake-3.0.0/Source/cmake.cxx -o cmake.o
...생략
-- Installing: /usr/local/bin/ccmake
-- Installing: /usr/local/bin/cmake
-- Installing: /usr/local/bin/ctest
-- Installing: /usr/local/bin/cpack
-- Installing: /usr/local/share/cmake-3.0/include/cmCPluginAPI.h
-- Installing: /usr/local/share/cmake-3.0/editors/vim/cmake-help.vim
-- Installing: /usr/local/share/cmake-3.0/editors/vim/cmake-indent.vim
-- Installing: /usr/local/share/cmake-3.0/editors/vim/cmake-syntax.vim
-- Installing: /usr/local/share/cmake-3.0/editors/emacs/cmake-mode.el
-- Installing: /usr/local/share/aclocal/cmake.m4
-- Installing: /usr/local/share/cmake-3.0/completions/cmake
-- Installing: /usr/local/share/cmake-3.0/completions/cpack
-- Installing: /usr/local/share/cmake-3.0/completions/ctest
[root@hawq2 cmake-3.0.0]#
```

## ld.so.conf

```
echo "/root/Downloads/hawq/boost_1_56_0/stage/lib" >> /etc/ld.so.conf.d/boost.conf
echo "/root/Downloads/hawq/curl-7.49.0/lib/.libs" >> /etc/ld.so.conf.d/curl.conf
echo "/usr/local/lib" >> /etc/ld.so.conf.d/usr-local-lib.conf
```

```
[root@hawq2 cmake-3.0.0]# echo "/root/Downloads/hawq/boost_1_56_0/stage/lib" >> /etc/ld.so.conf.d/boost.conf
[root@hawq2 cmake-3.0.0]# echo "/root/Downloads/hawq/curl-7.49.0/lib/.libs" >> /etc/ld.so.conf.d/curl.conf
[root@hawq2 cmake-3.0.0]# echo "/usr/local/lib" >> /etc/ld.so.conf.d/curl.conf
```

## copy boost library

```
cd ..
cp -R ~/Downloads/hawq/boost_1_56_0/boost /usr/local/include
cp /root/Downloads/hawq/boost_1_56_0/stage/lib/* /usr/local/lib
ldconfig
```

```
[root@hawq2 cmake-3.0.0]# cd ..
[root@hawq2 hawq]# cp -R ~/Downloads/hawq/boost_1_56_0/boost /usr/local/include
[root@hawq2 hawq]# cp /root/Downloads/hawq/boost_1_56_0/stage/lib/* /usr/local/lib
[root@hawq2 hawq]# ldconfig
```

## kernel option semaphore 변경

```
echo "kernel.sem = 250 512000 100 2048" >> /etc/sysctl.conf
sysctl -p
cat /proc/sys/kernel/sem
```

```
[root@hawq2 hawq]# echo "kernel.sem = 250 512000 100 2048" >> /etc/sysctl.conf
[root@hawq2 hawq]# sysctl -p
...생략
kernel.sem = 250 512000 100 2048
[root@hawq2 hawq]# cat /proc/sys/kernel/sem
2505120001002048
[root@hawq2 hawq]#
```

## Apache HAWQ

```
git clone https://git-wip-us.apache.org/repos/asf/incubator-hawq.git
CODE_BASE=`pwd`/incubator-hawq
cd $CODE_BASE
./configure --prefix=/usr/local/hawq/default --with-python
make -j8
make install
```

/usr/local/hawq/default 아래에 hawq install 용 소스가 설치됩니다.

### source code download

```
[root@hawq2 hawq]# git clone https://git-wip-us.apache.org/repos/asf/incubator-hawq.git
Initialized empty Git repository in /root/Downloads/hawq/incubator-hawq/.git/
remote: Counting objects: 41581, done.
remote: Compressing objects: 100% (18245/18245), done.
remote: Total 41581 (delta 25923), reused 35048 (delta 20469)
Receiving objects: 100% (41581/41581), 145.50 MiB | 6.12 MiB/s, done.
Resolving deltas: 100% (25923/25923), done.
[root@hawq2 hawq]#
```

### configure

```
[root@hawq2 hawq]# CODE_BASE=`pwd`/incubator-hawq
[root@hawq2 hawq]# cd $CODE_BASE
[root@hawq2 incubator-hawq]# ./configure --prefix=/usr/local/hawq/default --with-python
checking build system type... x86_64-unknown-linux-gnu
checking host system type... x86_64-unknown-linux-gnu
checking which template to use... linux
checking whether to build with 64-bit integer date/time support... yes
checking whether NLS is wanted... no
checking for default port number... 5432
checking for gcc... gcc
...생략
config.status: linking src/backend/port/tas/dummy.s to src/backend/port/tas.s
config.status: linking src/backend/port/dynloader/linux.c to src/backend/port/dynloader.c
config.status: linking src/backend/port/sysv_sema.c to src/backend/port/pg_sema.c
config.status: linking src/backend/port/sysv_shmem.c to src/backend/port/pg_shmem.c
config.status: linking src/backend/port/dynloader/linux.h to src/include/dynloader.h
config.status: linking src/include/port/linux.h to src/include/pg_config_os.h
config.status: linking src/makefiles/Makefile.linux to src/Makefile.port
configure: WARNING: option ignored: --enable-largefile
configure: WARNING: option ignored: --enable-option-checking
configure: WARNING: option ignored: --enable-largefile
[root@hawq2 incubator-hawq]#
```

### make -j8

2 core cpu, 4G memory 기준 약 20분 소요됩니다.

```
[root@hawq2 incubator-hawq]# make -j8
make -C depends/thirdparty/googletest all
make[1]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest'
cd ../../../../depends/thirdparty/googletest/ && mkdir -p build && cd build && cmake
-DCMAKE_INSTALL_PREFIX=/usr/local/hawq/default ..
-- The C compiler identification is GNU 4.8.2
-- The CXX compiler identification is GNU 4.8.2
-- Check for working C compiler: /opt/rh/devtoolset-2/root/usr/bin/cc
-- Check for working C compiler: /opt/rh/devtoolset-2/root/usr/bin/cc -- works
-- Detecting C compiler ABI info
-- Detecting C compiler ABI info - done
-- Check for working CXX compiler: /opt/rh/devtoolset-2/root/usr/bin/c++
-- Check for working CXX compiler: /opt/rh/devtoolset-2/root/usr/bin/c++ -- works
-- Detecting CXX compiler ABI info
```

```

-- Detecting CXX compiler ABI info - done
-- Found PythonInterp: /usr/bin/python (found version "2.6.6")
-- Looking for include file pthread.h
-- Looking for include file pthread.h - found
-- Looking for pthread_create
-- Looking for pthread_create - not found
-- Looking for pthread_create in pthreads
-- Looking for pthread_create in pthreads - not found
-- Looking for pthread_create in pthread
-- Looking for pthread_create in pthread - found
-- Found Threads: TRUE
-- Configuring done
-- Generating done
-- Build files have been written to: /root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build
cd ../../../../depends/thirdparty/googletest/build && make
make[2]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[3]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
Scanning dependencies of target gmock
Scanning dependencies of target gtest
Scanning dependencies of target gmock_main
make[4]: Leaving directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Leaving directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
[ 14%] [ 28%] [ 42%] Building CXX object googletest/CMakeFiles/gmock.dir/__/googletest/src/gtest-all.cc.o
Building CXX object googletest/CMakeFiles/gmock.dir/src/gmock-all.cc.o
Building CXX object googletest/gtest/CMakeFiles/gtest.dir/src/gtest-all.cc.o
make[4]: Leaving directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
[ 57%] [ 71%] Building CXX object googletest/CMakeFiles/gmock_main.dir/__/googletest/src/gtest-all.cc.o
...생략
/usr/local/maven/default/bin/mvn package -DskipTests
[INFO] Scanning for projects...
[INFO] -----
[INFO] Reactor Build Order:
[INFO]
[INFO] hawq-hadoop
[INFO] hawq-mapreduce-common
[INFO] hawq-mapreduce-ao
[INFO] hawq-mapreduce-parquet
[INFO] hawq-mapreduce-tool
[INFO]
[INFO] -----
[INFO] Building hawq-hadoop 1.1.0
[INFO] -----
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-jar-plugin/2.4/maven-jar-plugin-2.4.pom
...생략
[INFO] -----
[INFO] Reactor Summary:
[INFO]
[INFO] hawq-hadoop ..... SUCCESS [11:46 min]
[INFO] hawq-mapreduce-common ..... SUCCESS [03:37 min]
[INFO] hawq-mapreduce-ao ..... SUCCESS [ 1.008 s]
[INFO] hawq-mapreduce-parquet ..... SUCCESS [ 1.418 s]
[INFO] hawq-mapreduce-tool ..... SUCCESS [01:41 min]
[INFO] -----
[INFO] BUILD SUCCESS
[INFO] -----
[INFO] Total time: 17:08 min
[INFO] Finished at: 2016-06-01T23:31:48+09:00
[INFO] Final Memory: 33M/104M
[INFO] -----
make[2]: Leaving directory `/root/Downloads/hawq/incubator-hawq/contrib/hawq-hadoop'
make[1]: Leaving directory `/root/Downloads/hawq/incubator-hawq/contrib'
make -C tools all
make[1]: Entering directory `/root/Downloads/hawq/incubator-hawq/tools'
make -C gpnetbench all

```



```

make[2]: Entering directory `/root/Downloads/hawq/incubator-hawq/tools/gpnetbench'
gcc -O3 -std=gnu99 -Wall -Wmissing-prototypes -Wpointer-arith -Wendif-labels -Wformat-security -fno-strict-aliasing
-fwrapv -fno-aggressive-loop-optimizations -I/usr/local/include -I/usr/include/libxml2 -o gpnetbenchServer.o -c
gpnetbenchServer.c
gcc -O3 -std=gnu99 -Wall -Wmissing-prototypes -Wpointer-arith -Wendif-labels -Wformat-security -fno-strict-aliasing
-fwrapv -fno-aggressive-loop-optimizations -I/usr/local/include -I/usr/include/libxml2 -o gpnetbenchClient.o -c
gpnetbenchClient.c
gpnetbenchServer.c:34:6: warning: no previous prototype for 'usage' [-Wmissing-prototypes]
void usage()
    ^
gpnetbenchClient.c:36:6: warning: no previous prototype for 'usage' [-Wmissing-prototypes]
void usage()
    ^
gpnetbenchClient.c:231:6: warning: no previous prototype for 'print_headers' [-Wmissing-prototypes]
void print_headers()
    ^
gcc -O3 -std=gnu99 -Wall -Wmissing-prototypes -Wpointer-arith -Wendif-labels -Wformat-security -fno-strict-aliasing
-fwrapv -fno-aggressive-loop-optimizations -I/usr/local/include -I/usr/include/libxml2 -L../src/port
-L../src/port -Wl,--as-needed -L/root/Downloads/hawq/incubator
-hawq/depends/libhdfs3/build/install/usr/local/hawq/default/lib -L/root/Downloads/hawq/incubator
-hawq/depends/libyarn/build/install/usr/local/hawq/default/lib -Wl,-rpath,'/usr/local/hawq/default/lib',--enable-new
-dtags -o gpnetbenchServer gpnetbenchServer.o
gcc -O3 -std=gnu99 -Wall -Wmissing-prototypes -Wpointer-arith -Wendif-labels -Wformat-security -fno-strict-aliasing
-fwrapv -fno-aggressive-loop-optimizations -I/usr/local/include -I/usr/include/libxml2 -L../src/port
-L../src/port -Wl,--as-needed -L/root/Downloads/hawq/incubator
-hawq/depends/libhdfs3/build/install/usr/local/hawq/default/lib -L/root/Downloads/hawq/incubator
-hawq/depends/libyarn/build/install/usr/local/hawq/default/lib -Wl,-rpath,'/usr/local/hawq/default/lib',--enable-new
-dtags -o gpnetbenchClient gpnetbenchClient.o
finish building gpnetbenchServer and gpnetbenchClient
make[2]: Leaving directory `/root/Downloads/hawq/incubator-hawq/tools/gpnetbench'
make[1]: Leaving directory `/root/Downloads/hawq/incubator-hawq/tools'
All of HAWQ successfully made. Ready to install.
[root@hawq2 incubator-hawq]#

```

## make install

```

[root@hawq2 incubator-hawq]# make install
make -C depends/thirdparty/googletest install
make[1]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest'
cd ../../../../depends/thirdparty/googletest/ && mkdir -p build && cd build && cmake
-DCMAKE_INSTALL_PREFIX=/usr/local/hawq/default ..
-- Configuring done
-- Generating done
-- Build files have been written to: /root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build
cd ../../../../depends/thirdparty/googletest/build && make
make[2]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[3]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Entering directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
make[4]: Leaving directory `/root/Downloads/hawq/incubator-hawq/depends/thirdparty/googletest/build'
[ 28%] Built target gmock
...생략
make[2]: Leaving directory `/root/Downloads/hawq/incubator-hawq/tools'
make -C gpnetbench install
make[2]: Entering directory `/root/Downloads/hawq/incubator-hawq/tools/gpnetbench'
finish building gpnetbenchServer and gpnetbenchClient
/bin/bash ../../config/install-sh -c gpnetbenchServer /usr/local/hawq/default/bin/lib
/bin/bash ../../config/install-sh -c gpnetbenchClient /usr/local/hawq/default/bin/lib
make[2]: Leaving directory `/root/Downloads/hawq/incubator-hawq/tools/gpnetbench'
make[1]: Leaving directory `/root/Downloads/hawq/incubator-hawq/tools'
HAWQ installation complete.
[root@hawq2 incubator-hawq]#

```

## HAWQ init dependencies install

```

easy_install https://pypi.python.org/packages/5d/30/3f9192377ee563f92c9bb6e8daf32d2c11fd484cd3efe431686e74edc61d/PSI-0.3b2.tar.gz
easy_install
https://pypi.python.org/packages/67/8a/9cc519c851c313e498e96807adf59c51af243f510cab7d7e4eb5c7edb53e/paramiko-1.15.4.tar.gz#md5=baf46a6bc789c177f52988bf12a57a6f
easy_install http://darcs.idyll.org/~t/projects/figleaf-latest.tar.gz
easy_install http://pyyaml.org/download/pyyaml/PyYAML-3.11.tar.gz
easy_install
https://pypi.python.org/packages/17/47/72cb04a58a35ec495f96984dddb48232b551aafb95bde614605b754fe6f7/lockfile-0.12.2.tar.gz#md5=a6a1a82957a23afdf44cfdd039b65ff9
cp /root/Downloads/hawq/boost_1_56_0/stage/lib/* /usr/local/hawq/default/lib/
chown hskimsky:hskimsky /usr/local/hawq/default/lib/libboost_*

```

```

[root@hawq2 incubator-hawq]# easy_install
https://pypi.python.org/packages/5d/30/3f9192377ee563f92c9bb6e8daf32d2c11fd484cd3efe431686e74edc61d/PSI-0.3b2.tar.gz
Downloading https://pypi.python.org/packages/5d/30/3f9192377ee563f92c9bb6e8daf32d2c11fd484cd3efe431686e74edc61d/PSI-0.3b2.tar.gz
Processing PSI-0.3b2.tar.gz
Writing /tmp/easy_install-wiM6Hg/PSI-0.3b2/setup.cfg
Running PSI-0.3b2/setup.py -q bdist_egg --dist-dir /tmp/easy_install-wiM6Hg/PSI-0.3b2/egg-dist-tmp-8YhtdM
...생략
zip_safe flag not set; analyzing archive contents...
Moving PSI-0.3b2-py2.6-linux-x86_64.egg to /usr/lib/python2.6/site-packages
Adding PSI 0.3b2 to easy-install.pth file

Installed /usr/lib/python2.6/site-packages/PSI-0.3b2-py2.6-linux-x86_64.egg
Processing dependencies for PSI==0.3b2
Finished processing dependencies for PSI==0.3b2
[root@hawq2 incubator-hawq]#
[root@hawq2 incubator-hawq]# easy_install
https://pypi.python.org/packages/67/8a/9cc519c851c313e498e96807adf59c51af243f510cab7d7e4eb5c7edb53e/paramiko-1.15.4.tar.gz#md5=baf46a6bc789c177f52988bf12a57a6f
Downloading
https://pypi.python.org/packages/67/8a/9cc519c851c313e498e96807adf59c51af243f510cab7d7e4eb5c7edb53e/paramiko-1.15.4.tar.gz#md5=baf46a6bc789c177f52988bf12a57a6f
Processing paramiko-1.15.4.tar.gz
Writing /tmp/easy_install-BPDSoE/paramiko-1.15.4/setup.cfg
Running paramiko-1.15.4/setup.py -q bdist_egg --dist-dir /tmp/easy_install-BPDSoE/paramiko-1.15.4/egg-dist-tmp-8ACfop
zip_safe flag not set; analyzing archive contents...
Moving paramiko-1.15.4-py2.6.egg to /usr/lib/python2.6/site-packages
Adding paramiko 1.15.4 to easy-install.pth file
...생략
zip_safe flag not set; analyzing archive contents...
Moving pycrypto-2.6.1-py2.6-linux-x86_64.egg to /usr/lib/python2.6/site-packages
Adding pycrypto 2.6.1 to easy-install.pth file

Installed /usr/lib/python2.6/site-packages/pycrypto-2.6.1-py2.6-linux-x86_64.egg
Finished processing dependencies for paramiko==1.15.4
[root@hawq2 incubator-hawq]#
[root@hawq2 incubator-hawq]# easy_install http://darcs.idyll.org/~t/projects/figleaf-latest.tar.gz
Downloading http://darcs.idyll.org/~t/projects/figleaf-latest.tar.gz
Processing figleaf-latest.tar.gz
Writing /tmp/easy_install-ca3uoS/figleaf-latest/setup.cfg
Running figleaf-latest/setup.py -q bdist_egg --dist-dir /tmp/easy_install-ca3uoS/figleaf-latest/egg-dist-tmp-cwt10p
zip_safe flag not set; analyzing archive contents...
figleaf.annotate: module references __file__
figleaf.internals: module references __file__
figleaf.__init__: module references __file__
creating /usr/lib/python2.6/site-packages/figleaf-0.6.1-py2.6.egg
Extracting figleaf-0.6.1-py2.6.egg to /usr/lib/python2.6/site-packages
Adding figleaf 0.6.1 to easy-install.pth file
Installing annotate-sections script to /usr/bin
Installing figleaf2cov script to /usr/bin
Installing figleaf2ast script to /usr/bin
Installing figleaf script to /usr/bin
Installing figleaf2html script to /usr/bin

Installed /usr/lib/python2.6/site-packages/figleaf-0.6.1-py2.6.egg

```

```

Processing dependencies for figleaf==0.6.1
Finished processing dependencies for figleaf==0.6.1
[root@hawq2 incubator-hawq]#
[root@hawq2 incubator-hawq]# easy_install http://pyyaml.org/download/pyyaml/PyYAML-3.11.tar.gz
Downloading http://pyyaml.org/download/pyyaml/PyYAML-3.11.tar.gz
Processing PyYAML-3.11.tar.gz
Writing /tmp/easy_install-AaBvd6/PyYAML-3.11/setup.cfg
Running PyYAML-3.11/setup.py -q bdist_egg --dist-dir /tmp/easy_install-AaBvd6/PyYAML-3.11/egg-dist-tmp-ZT88CA
In file included from ext/_yaml.c:343:0:
... 생략
zip_safe flag not set; analyzing archive contents...
Moving PyYAML-3.11-py2.6-linux-x86_64.egg to /usr/lib/python2.6/site-packages
Adding PyYAML 3.11 to easy-install.pth file

Installed /usr/lib/python2.6/site-packages/PyYAML-3.11-py2.6-linux-x86_64.egg
Processing dependencies for PyYAML==3.11
Finished processing dependencies for PyYAML==3.11
[root@hawq2 incubator-hawq]#
[root@hawq2 incubator-hawq]# easy_install
https://pypi.python.org/packages/17/47/72cb04a58a35ec495f96984dddb48232b551aafb95bde614605b754fe6f7/lockfile-0.12.2.tar.gz#md5=a6a1a82957a23afdf44cfdd039b65ff9
Downloading
https://pypi.python.org/packages/17/47/72cb04a58a35ec495f96984dddb48232b551aafb95bde614605b754fe6f7/lockfile-0.12.2.tar.gz#md5=a6a1a82957a23afdf44cfdd039b65ff9
Processing lockfile-0.12.2.tar.gz
Writing /tmp/easy_install-VYV80v/lockfile-0.12.2/setup.cfg
Running lockfile-0.12.2/setup.py -q bdist_egg --dist-dir /tmp/easy_install-VYV80v/lockfile-0.12.2/egg-dist-tmp-rd215E

Installed /tmp/easy_install-VYV80v/lockfile-0.12.2/.eggs/pbr-1.10.0-py2.6.egg
creating /usr/lib/python2.6/site-packages/lockfile-0.12.2-py2.6.egg
Extracting lockfile-0.12.2-py2.6.egg to /usr/lib/python2.6/site-packages
Adding lockfile 0.12.2 to easy-install.pth file

Installed /usr/lib/python2.6/site-packages/lockfile-0.12.2-py2.6.egg
Processing dependencies for lockfile==0.12.2
Finished processing dependencies for lockfile==0.12.2
[root@hawq2 incubator-hawq]#

```

## copy boost

```

cp /root/Downloads/hawq/boost_1_56_0/stage/lib/* /usr/local/hawq/default/lib/
chown -R hskimsky:hskimsky /usr/local/hawq

```

```

[root@hawq2 incubator-hawq]# cp /root/Downloads/hawq/boost_1_56_0/stage/lib/* /usr/local/hawq/default/lib/
[root@hawq2 incubator-hawq]# chown -R hskimsky:hskimsky /usr/local/hawq
[root@hawq2 incubator-hawq]#

```

## hawq-site.xml and copy

hskimsky 계정으로 작업합니다.

```

cd /usr/local/hawq/default/etc/
echo "hawq3.apache.local\

hawq4.apache.local" > slaves
vim hawq-site.xml

```

아래 property 들은 변경하고 그 외 property 들은 그대로 둡니다.

hawq\_rm\_yarn\_address 값과 hawq\_rm\_yarn\_scheduler\_address 값은  
 <HADOOP\_HOME>/etc/hadoop/yarn-site.xml 의  
 yarn.resourcemanager.address 값과 yarn.resourcemanager.scheduler.address 값으로 설정합니다.

```

<property>
  <name>hawq_master_address_host</name>
  <value>hawq2.apache.local</value>
</property>
<property>
  <name>hawq_standby_address_host</name>
  <value>hawq1.apache.local</value>
</property>
<property>
  <name>hawq_dfs_url</name>
  <value>hawq1.apache.local:8020/hawq_data</value>
</property>
<property>
  <name>hawq_master_directory</name>
  <value>/hawq_data/master</value>
</property>
<property>
  <name>hawq_segment_directory</name>
  <value>/hawq_data/segment</value>
</property>
<property>
  <name>hawq_global_rm_type</name>
  <value>yarn</value>
</property>
<property>
  <name>hawq_rm_yarn_address</name>
  <value>hawq1.apache.local:8050</value>
</property>
<property>
  <name>hawq_rm_yarn_scheduler_address</name>
  <value>hawq1.apache.local:8030</value>
</property>
<property>
  <name>hawq_rm_master_port</name>
  <value>5437</value>
</property>
<property>
  <name>hawq_rm_segment_port</name>
  <value>5438</value>
</property>

```

```

scp /usr/local/hawq/default/etc/* hawq1:/usr/local/hawq/default/etc
scp /usr/local/hawq/default/etc/* hawq3:/usr/local/hawq/default/etc
scp /usr/local/hawq/default/etc/* hawq4:/usr/local/hawq/default/etc

```

```

[hskimsky@hawq2 ~]$ cd /usr/local/hawq/default/etc/
[hskimsky@hawq2 etc]$ echo "hawq3.apache.local\
>
> hawq4.apache.local" > slaves
[hskimsky@hawq2 etc]$ vim hawq-site.xml
[hskimsky@hawq2 etc]$ scp /usr/local/hawq/default/etc/* hawq1:/usr/local/hawq/default/etc
[hskimsky@hawq2 etc]$ scp /usr/local/hawq/default/etc/* hawq3:/usr/local/hawq/default/etc
[hskimsky@hawq2 etc]$ scp /usr/local/hawq/default/etc/* hawq4:/usr/local/hawq/default/etc

```

hadoop start at namenode (hawq1.apache.local)

hadoop 의 super user 로 namenode 에서 hadoop 을 시작합니다.

```

${HADOOP_HOME}/sbin/start-dfs.sh
${HADOOP_HOME}/sbin/start-yarn.sh
${HADOOP_HOME}/sbin/mr-jobhistory-daemon.sh start historyserver
${HADOOP_HOME}/sbin/yarn-daemon.sh start proxyserver

```

```
[hskimsky@hawq1 ~]$ ${HADOOP_HOME}/sbin/start-dfs.sh
[hskimsky@hawq1 ~]$ ${HADOOP_HOME}/sbin/start-yarn.sh
[hskimsky@hawq1 ~]$ ${HADOOP_HOME}/sbin/mr-jobhistory-daemon.sh start historyserver
[hskimsky@hawq1 ~]$ ${HADOOP_HOME}/sbin/yarn-daemon.sh start proxyserver
```

## hawq mkdir

```
hawq1.apache.local

hdfs dfs -mkdir /hawq_data
sudo mkdir -p /hawq_data/master
sudo mkdir -p /hawq_data/segment
sudo chown -R hskimsky:hskimsky /hawq_data

hawq2.apache.local, hawq3.apache.local, hawq4.apache.local

sudo mkdir -p /hawq_data/master
sudo mkdir -p /hawq_data/segment
sudo chown -R hskimsky:hskimsky /hawq_data
```

```
[hskimsky@hawq1 ~]$ hdfs dfs -mkdir /hawq_data
[hskimsky@hawq1 ~]$ sudo mkdir -p /hawq_data/master
[hskimsky@hawq1 ~]$ sudo mkdir -p /hawq_data/segment
[hskimsky@hawq1 ~]$ sudo chown -R hskimsky:hskimsky /hawq_data

[hskimsky@hawq2 ~]$ sudo mkdir -p /hawq_data/master
[hskimsky@hawq2 ~]$ sudo mkdir -p /hawq_data/segment
[hskimsky@hawq2 ~]$ sudo chown -R hskimsky:hskimsky /hawq_data

[hskimsky@hawq3 ~]$ sudo mkdir -p /hawq_data/master
[hskimsky@hawq3 ~]$ sudo mkdir -p /hawq_data/segment
[hskimsky@hawq3 ~]$ sudo chown -R hskimsky:hskimsky /hawq_data

[hskimsky@hawq4 ~]$ sudo mkdir -p /hawq_data/master
[hskimsky@hawq4 ~]$ sudo mkdir -p /hawq_data/segment
[hskimsky@hawq4 ~]$ sudo chown -R hskimsky:hskimsky /hawq_data
```

## hawq init cluster at hawq master (hawq2.apache.local)

hskimsky 계정으로 작업합니다.

```
source /usr/local/hawq/default/greenplum_path.sh
hawq init cluster
```

실패 시 모든 노드에서 아래 명령어를 실행합니다.

```
hdfs dfs -rm -r /hawq_data/16385
rm -r /hawq_data/master/*
rm -r /hawq_data/segment/*
kill <실행중인 hawq process>
```

```
[hskimsky@hawq2 ~]$ source /usr/local/hawq/default/greenplum_path.sh
[hskimsky@hawq2 ~]$ hawq init cluster
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-Prepare to do 'hawq init'
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-You can find log in:
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-/home/hskimsky/hawqAdminLogs/hawq_init_20160604.log
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-GPHOME is set to:
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-/usr/local/hawq/default
20160604:11:38:02:038585 hawq_init:hawq2:hskimsky-[INFO]:-Init hawq with args: ['init', 'cluster']

Continue with HAWQ init Yy|Nn (default=N):
> y
20160604:11:38:05:038585 hawq_init:hawq2:hskimsky-[INFO]:-Check if hdfs path is available
20160604:11:38:05:038585 hawq_init:hawq2:hskimsky-[WARNING]:-2016-06-04 11:38:05.481199, p38780, th139907845580960,
WARNING the number of nodes in pipeline is 2 [hawq4.apache.local(192.168.98.169), hawq3.apache.local(192.168.98.168)],
is less than the expected number of replica 3 for block [block pool ID: BP-913932141-192.168.98.166-1464796448686 block
ID 1073741850_1026] file /hawq_data/testFile
20160604:11:38:05:038585 hawq_init:hawq2:hskimsky-[INFO]:-2 segment hosts defined
20160604:11:38:05:038585 hawq_init:hawq2:hskimsky-[INFO]:-Set default_hash_table_bucket_number as: 12
20160604:11:38:13:038585 hawq_init:hawq2:hskimsky-[INFO]:-Start to init master node: 'hawq2.apache.local'
20160604:11:38:25:038585 hawq_init:hawq2:hskimsky-[INFO]:-20160604:11:38:25:038910 hawqinit.sh:hawq2:hskimsky-[INFO]:-
Loading hawq_toolkit...
20160604:11:38:25:038585 hawq_init:hawq2:hskimsky-[INFO]:-Master init successfully
20160604:11:38:25:038585 hawq_init:hawq2:hskimsky-[INFO]:-Start to init standby master: 'hawq1.apache.local'
20160604:11:38:25:038585 hawq_init:hawq2:hskimsky-[INFO]:-This might take a couple of minutes, please wait...
20160604:11:38:50:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-HAWQ master stopped
20160604:11:38:50:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Sync files to standby from master
20160604:11:38:54:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Update pg_hba configuration
20160604:11:38:56:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Standby ip address is 192.168.98.166
20160604:11:38:56:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Start hawq master
20160604:11:38:59:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-HAWQ master started
20160604:11:38:59:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Try to remove existing standby from catalog
20160604:11:39:00:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-Register standby to master successfully
20160604:11:39:25:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-HAWQ master stopped
20160604:11:39:33:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-HAWQ standby started
20160604:11:39:41:035920 hawqinit.sh:hawq1:hskimsky-[INFO]:-HAWQ master started
20160604:11:39:41:038585 hawq_init:hawq2:hskimsky-[INFO]:-Init standby successfully
20160604:11:39:41:038585 hawq_init:hawq2:hskimsky-[INFO]:-Init segments in list: ['hawq3.apache.local',
'hawq4.apache.local']
20160604:11:39:42:038585 hawq_init:hawq2:hskimsky-[INFO]:-Total segment number is: 2
.....
20160604:11:39:57:038585 hawq_init:hawq2:hskimsky-[INFO]:-2 of 2 segments init successfully
20160604:11:39:57:038585 hawq_init:hawq2:hskimsky-[INFO]:-Segments init successfully on nodes ['hawq3.apache.local',
'hawq4.apache.local']
20160604:11:39:57:038585 hawq_init:hawq2:hskimsky-[INFO]:-Init HAWQ cluster successfully
[hskimsky@hawq2 ~]$
```

## yanr application list

```
[hskimsky@hawq2 ~]$ yarn application -list
16/06/05 13:55:39 INFO client.RMPProxy: Connecting to ResourceManager at hawq1.apache.local/192.168.98.166:8050
16/06/05 13:55:40 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
java classes where applicable
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED, RUNNING]):1
Application-Id Application-Name Application-Type User Queue State
Final-State Progress Tracking-URL
application_1465055171853_0005 hawq YARN hskimsky default RUNNING
UNDEFINED 50% url
[hskimsky@hawq2 ~]$
```

## Connect and Run basic queries

```
psql -d postgres
create table t ( i int );
insert into t values(1);
insert into t select generate_series(1,10000);
select count(*) from t;
```

```
[hskimsky@hawq2 hawq_data]$ psql -d postgres
psql (8.2.15)
Type "help" for help.

postgres=# create table t ( i int );
CREATE TABLE
postgres=# insert into t values(1);
INSERT 0 1
postgres=# insert into t select generate_series(1,10000);
INSERT 0 10000
postgres=# select count(*) from t;
 count
-----
 10001
(1 row)

postgres=#
```

## PXF

### PXF build

HAWQ(master, standby, segment) 가 설치되는 모든 서버(hawq1.apache.local, hawq2.apache.local, hawq3.apache.local, hawq4.apache.local)에서 root 로 작업합니다.

```
cd ~/Downloads/hawq/incubator-hawq/pxf
./gradlew tasks
make && make unittest

chmod 755 /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/*
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pre-install.sh
mkdir -p /usr/lib/pxf
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-api/build/libs/pxf-api-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hbase/build/libs/pxf-hbase-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hdfs/build/libs/pxf-hdfs-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hive/build/libs/pxf-hive-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-json/build/libs/pxf-json-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf-service-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf.war /usr/lib/pxf/
ln -s /usr/lib/pxf/pxf-api-3.0.0.jar /usr/lib/pxf/pxf-api.jar
ln -s /usr/lib/pxf/pxf-hbase-3.0.0.jar /usr/lib/pxf/pxf-hbase.jar
ln -s /usr/lib/pxf/pxf-hdfs-3.0.0.jar /usr/lib/pxf/pxf-hdfs.jar
ln -s /usr/lib/pxf/pxf-hive-3.0.0.jar /usr/lib/pxf/pxf-hive.jar
ln -s /usr/lib/pxf/pxf-json-3.0.0.jar /usr/lib/pxf/pxf-json.jar
ln -s /usr/lib/pxf/pxf-service-3.0.0.jar /usr/lib/pxf/pxf-service.jar
chown -R pxf:pxf /usr/lib/pxf

mkdir -p /etc/pxf/conf
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-env.sh /etc/pxf/conf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/resources/main/pxf-* /etc/pxf/conf/
mv /etc/pxf/conf/pxf-private.classpath /etc/pxf/conf/pxf-private.classpath.bak
# hawq 에서 hbase, hive 를 사용하기 위해선 아래 파일에 jar 파일을 추가해줘야 함.
echo "#####
# This file contains the internal classpaths required to run PXF.
# WARNING: DO NOT EDIT!
# Any change in this file can result in PXF failing to run.
# Adding resources should be done using pxf-public.classpath file.
#####"
```

```

/etc/pxf/conf
/usr/local/hadoop/default/etc/hadoop
/usr/local/hadoop/default/share/hadoop/hdfs/hadoop-hdfs-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/hadoop-auth-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/hadoop-common-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/asm-3.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/avro-1.7.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-cli-1.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-codec-1.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-collections-3.2.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-configuration-1.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-io-2.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-lang-2.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-logging-1.1.3.jar
/usr/local/hadoop/default/share/hadoop/common/lib/guava-11.0.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/htrace-core-3.1.0-incubating.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jetty-6.1.26.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jersey-core-1.9.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jersey-server-1.9.jar
/usr/local/hadoop/default/share/hadoop/common/lib/log4j-1.2.17.jar
/usr/local/hadoop/default/share/hadoop/common/lib/protobuf-java-2.5.0.jar
/usr/local/hadoop/default/share/hadoop/common/lib/slf4j-api-1.7.10.jar
/usr/local/hadoop/default/share/hadoop/common/lib/snappy-java-1.0.4.1.jar
/usr/local/hadoop/default/share/hadoop/common/lib/zookeeper-3.4.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/netty-3.6.2.Final.jar
/usr/local/hadoop/default/share/hadoop/common/lib/zookeeper-3.4.6.jar
/usr/lib/pxf/pxf-hbase-[0-9].jar
/usr/lib/pxf/pxf-hdfs-[0-9].jar
/usr/lib/pxf/pxf-hive-[0-9].jar
/usr/lib/pxf/pxf-json-[0-9].jar
" > /etc/pxf/conf/pxf-private.classpath
chown -R pxf:pxf /etc/pxf

cd ~/Downloads
wget http://apache.mirror.cdnetworks.com/tomcat/tomcat-7/v7.0.69/bin/apache-tomcat-7.0.69.tar.gz
tar xzf apache-tomcat-7.0.69.tar.gz
mkdir -p /opt/pivotal/pxf
cp -R apache-tomcat-7.0.69 /opt/pivotal/apache-tomcat
cp -R /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/configs/tomcat /opt/pivotal/pxf/tomcat-templates

/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service init
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/post-install.sh
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service start

```

## gradle build

```

cd ~/Downloads/hawq/incubator-hawq/pxf
./gradlew tasks
make && make unittest

```



```

[root@hawq1 ~]# cd ~/Downloads/hawq/incubator-hawq/pxf
[root@hawq1 pxf]# ./gradlew tasks
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left   Speed
100 353      0 353    0     0    769      0 --:--:-- --:--:-- --:--:-- 826
100 66.1M 100 66.1M  0     0 42.0M      0 0:00:01 0:00:01 --:--:-- 61.4M
~/gradle/wrapper/dists/gradle-2.11-all-5d20f2a9ea62d191be56de67153870ef ~/Downloads/hawq/incubator-hawq/pxf
~/Downloads/hawq/incubator-hawq/pxf
Download http://repo1.maven.org/maven2/com/netflix/nebula/gradle-ospackage-plugin/2.2.6/gradle-ospackage-plugin-
2.2.6.pom
Download http://repo1.maven.org/maven2/de/undercouch/gradle-download-task/2.1.0/gradle-download-task-2.1.0.pom
...생략
Rules
-----
Pattern: clean<TaskName>: Cleans the output files of a task.
Pattern: build<ConfigurationName>: Assembles the artifacts of a configuration.
Pattern: upload<ConfigurationName>: Assembles and uploads the artifacts belonging to a configuration.

To see all tasks and more detail, run gradlew tasks --all

To see more detail about a task, run gradlew help --task <task>

BUILD SUCCESSFUL

Total time: 47.06 secs

This build could be faster, please consider using the Gradle Daemon:
https://docs.gradle.org/2.11/userguide/gradle_daemon.html
[root@hawq1 pxf]#
[root@hawq1 pxf]# make && make unittest
./gradlew clean release
:clean UP-TO-DATE
:pxf-api:clean UP-TO-DATE
:pxf-hbase:clean UP-TO-DATE
:pxf-hdfs:clean UP-TO-DATE
:pxf-hive:clean UP-TO-DATE
:pxf-json:clean UP-TO-DATE
:pxf-service:clean UP-TO-DATE
:pxf-api:compileJava
Download http://repo1.maven.org/maven2/commons-logging/commons-logging/1.1.3/commons-logging-1.1.3.pom
Download http://repo1.maven.org/maven2/org/apache/commons/commons-parent/28/commons-parent-28.pom
...생략
:pxf-json:classes UP-TO-DATE
:pxf-json:compileTestJava UP-TO-DATE
:pxf-json:processTestResources UP-TO-DATE
:pxf-json:testClasses UP-TO-DATE
:pxf-json:test UP-TO-DATE
:pxf-service:compileTestJava UP-TO-DATE
:pxf-service:processTestResources UP-TO-DATE
:pxf-service:testClasses UP-TO-DATE
:pxf-service:test UP-TO-DATE

BUILD SUCCESSFUL

Total time: 17.357 secs

This build could be faster, please consider using the Gradle Daemon:
https://docs.gradle.org/2.11/userguide/gradle_daemon.html
[root@hawq1 pxf]#

```

pxf library copy

```
chmod 755 /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/*
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pre-install.sh
mkdir -p /usr/lib/pxf
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-api/build/libs/pxf-api-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hbase/build/libs/pxf-hbase-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hdfs/build/libs/pxf-hdfs-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hive/build/libs/pxf-hive-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-json/build/libs/pxf-json-3.0.0.jar /usr/lib/pxf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf-service-3.0.0.jar /usr/lib/pxf
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf.war /usr/lib/pxf
ln -s /usr/lib/pxf/pxf-api-3.0.0.jar /usr/lib/pxf/pxf-api.jar
ln -s /usr/lib/pxf/pxf-hbase-3.0.0.jar /usr/lib/pxf/pxf-hbase.jar
ln -s /usr/lib/pxf/pxf-hdfs-3.0.0.jar /usr/lib/pxf/pxf-hdfs.jar
ln -s /usr/lib/pxf/pxf-hive-3.0.0.jar /usr/lib/pxf/pxf-hive.jar
ln -s /usr/lib/pxf/pxf-json-3.0.0.jar /usr/lib/pxf/pxf-json.jar
ln -s /usr/lib/pxf/pxf-service-3.0.0.jar /usr/lib/pxf/pxf-service.jar
chown -R pxf:pxf /usr/lib/pxf
```

```
[root@hawq1 pxf]# chmod 755 /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/*
[root@hawq1 pxf]# /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pre-install.sh
[root@hawq1 pxf]# mkdir -p /usr/lib/pxf
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-api/build/libs/pxf-api-3.0.0.jar /usr/lib/pxf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hbase/build/libs/pxf-hbase-3.0.0.jar /usr/lib/pxf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hdfs/build/libs/pxf-hdfs-3.0.0.jar /usr/lib/pxf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-hive/build/libs/pxf-hive-3.0.0.jar /usr/lib/pxf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-json/build/libs/pxf-json-3.0.0.jar /usr/lib/pxf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf-service-3.0.0.jar /usr/lib/pxf
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/libs/pxf.war /usr/lib/pxf
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-api-3.0.0.jar /usr/lib/pxf/pxf-api.jar
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-hbase-3.0.0.jar /usr/lib/pxf/pxf-hbase.jar
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-hdfs-3.0.0.jar /usr/lib/pxf/pxf-hdfs.jar
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-hive-3.0.0.jar /usr/lib/pxf/pxf-hive.jar
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-json-3.0.0.jar /usr/lib/pxf/pxf-json.jar
[root@hawq1 pxf]# ln -s /usr/lib/pxf/pxf-service-3.0.0.jar /usr/lib/pxf/pxf-service.jar
[root@hawq1 pxf]# chown -R pxf:pxf /usr/lib/pxf
[root@hawq1 pxf]#
```

pxf configurations copy

```

mkdir -p /etc/pxf/conf
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-env.sh /etc/pxf/conf/
cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/resources/main/pxf-* /etc/pxf/conf/
mv /etc/pxf/conf/pxf-private.classpath /etc/pxf/conf/pxf-private.classpath.bak
# hawq 에서 hbase, hive 를 사용하기 위해선 아래 파일에 jar 파일을 추가해줘야 함.
echo "#####
# This file contains the internal classpaths required to run PXF.
# WARNING: DO NOT EDIT!
# Any change in this file can result in PXF failing to run.
# Adding resources should be done using pxf-public.classpath file.
#####
/etc/pxf/conf
/usr/local/hadoop/default/etc/hadoop
/usr/local/hadoop/default/share/hadoop/hdfs/hadoop-hdfs-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/hadoop-auth-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/hadoop-common-2.7.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/asm-3.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/avro-1.7.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-cli-1.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-codec-1.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-collections-3.2.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-configuration-1.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-io-2.4.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-lang-2.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/commons-logging-1.1.3.jar
/usr/local/hadoop/default/share/hadoop/common/lib/guava-11.0.2.jar
/usr/local/hadoop/default/share/hadoop/common/lib/htrace-core-3.1.0-incubating.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jetty-6.1.26.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jersey-core-1.9.jar
/usr/local/hadoop/default/share/hadoop/common/lib/jersey-server-1.9.jar
/usr/local/hadoop/default/share/hadoop/common/lib/log4j-1.2.17.jar
/usr/local/hadoop/default/share/hadoop/common/lib/protobuf-java-2.5.0.jar
/usr/local/hadoop/default/share/hadoop/common/lib/slf4j-api-1.7.10.jar
/usr/local/hadoop/default/share/hadoop/common/lib/snappy-java-1.0.4.1.jar
/usr/local/hadoop/default/share/hadoop/common/lib/zookeeper-3.4.6.jar
/usr/local/hadoop/default/share/hadoop/common/lib/netty-3.6.2.Final.jar
/usr/lib/pxf/pxf-hbase-*[0-9].jar
/usr/lib/pxf/pxf-hdfs-*[0-9].jar
/usr/lib/pxf/pxf-hive-*[0-9].jar
/usr/lib/pxf/pxf-json-*[0-9].jar
" > /etc/pxf/conf/pxf-private.classpath
chown -R pxf:pxf /etc/pxf

```

```

[root@hawq1 pxf]# mkdir -p /etc/pxf/conf
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-env.sh /etc/pxf/conf/
[root@hawq1 pxf]# cp /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/build/resources/main/pxf-* /etc/pxf/conf/
[root@hawq1 pxf]# mv /etc/pxf/conf/pxf-private.classpath /etc/pxf/conf/pxf-private.classpath.bak
[root@hawq1 pxf]# # hawq 에서 hbase, hive 를 사용하기 위해선 아래 파일에 jar 파일을 추가해줘야 함.
[root@hawq1 pxf]# echo "#####
> # This file contains the internal classpaths required to run PXF.
> # WARNING: DO NOT EDIT!
> # Any change in this file can result in PXF failing to run.
> # Adding resources should be done using pxf-public.classpath file.
> #####
> /etc/pxf/conf
> /usr/local/hadoop/default/etc/hadoop
> /usr/local/hadoop/default/share/hadoop/hdfs/hadoop-hdfs-2.7.2.jar
> /usr/local/hadoop/default/share/hadoop/mapreduce/hadoop-mapreduce-client-core-2.7.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/hadoop-auth-2.7.2.jar
> /usr/local/hadoop/default/share/hadoop/common/hadoop-common-2.7.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/asm-3.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/avro-1.7.4.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-cli-1.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-codec-1.4.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-collections-3.2.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-configuration-1.6.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-io-2.4.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-lang-2.6.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/commons-logging-1.1.3.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/guava-11.0.2.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/htrace-core-3.1.0-incubating.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/jetty-6.1.26.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/jersey-core-1.9.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/jersey-server-1.9.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/log4j-1.2.17.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/protobuf-java-2.5.0.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/slf4j-api-1.7.10.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/snappy-java-1.0.4.1.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/zookeeper-3.4.6.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/netty-3.6.2.Final.jar
> /usr/local/hadoop/default/share/hadoop/common/lib/zookeeper-3.4.6.jar
> /usr/lib/pxf/pxf-hbase-*[0-9].jar
> /usr/lib/pxf/pxf-hdfs-*[0-9].jar
> /usr/lib/pxf/pxf-hive-*[0-9].jar
> /usr/lib/pxf/pxf-json-*[0-9].jar
> " > /etc/pxf/conf/pxf-private.classpath
[root@hawq1 pxf]# chown -R pxf:pxf /etc/pxf
[root@hawq1 pxf]#

```

## tomcat configuration for pxf

```

cd ~/Downloads
wget http://apache.mirror.cdnetworks.com/tomcat/tomcat-7/v7.0.69/bin/apache-tomcat-7.0.69.tar.gz
tar xzf apache-tomcat-7.0.69.tar.gz
mkdir -p /opt/pivotal/pxf
cp -R apache-tomcat-7.0.69 /opt/pivotal/apache-tomcat
cp -R /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/configs/tomcat /opt/pivotal/pxf/tomcat-templates

```

```
[root@hawq1 pxf]# cd ~/Downloads
[root@hawq1 Downloads]# wget http://apache.mirror.cdnetworks.com/tomcat/tomcat-7/v7.0.69/bin/apache-tomcat-7.0.69.tar.gz
--2016-06-05 13:40:06-- http://apache.mirror.cdnetworks.com/tomcat/tomcat-7/v7.0.69/bin/apache-tomcat-7.0.69.tar.gz
Resolving apache.mirror.cdnetworks.com... 14.0.101.165
Connecting to apache.mirror.cdnetworks.com|14.0.101.165|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 8910579 (8.5M) [application/x-gzip]
Saving to: "apache-tomcat-7.0.69.tar.gz"

100%[=====]
>] 8,910,579 23.9M/s in 0.4s

2016-06-05 13:40:06 (23.9 MB/s) - "apache-tomcat-7.0.69.tar.gz" saved [8910579/8910579]

[root@hawq1 Downloads]# tar xzf apache-tomcat-7.0.69.tar.gz
[root@hawq1 Downloads]# mkdir -p /opt/pivotal/pxf
[root@hawq1 Downloads]# cp -R apache-tomcat-7.0.69 /opt/pivotal/apache-tomcat
[root@hawq1 Downloads]# cp -R /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/configs/tomcat
/opt/pivotal/pxf/tomcat-templates
[root@hawq1 Downloads]#
```

## PXF start

```
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service init
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/post-install.sh
/root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service start
```

```
[root@hawq1 Downloads]# /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service init
[root@hawq1 Downloads]# /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/post-install.sh
[root@hawq1 Downloads]# /root/Downloads/hawq/incubator-hawq/pxf/pxf-service/src/scripts/pxf-service start
/var/pxf ~/Downloads
Using CATALINA_BASE: /var/pxf/pxf-service
Using CATALINA_HOME: /var/pxf/pxf-service
Using CATALINA_TMPDIR: /var/pxf/pxf-service/temp
Using JRE_HOME: /usr/local/java/default
Using CLASSPATH: /var/pxf/pxf-service/bin/bootstrap.jar:/var/pxf/pxf-service/bin/tomcat-juli.jar
Using CATALINA_PID: /var/run/pxf/catalina.pid
Tomcat started.
~/Downloads
Checking if tomcat is up and running...
tomcat not responding, re-trying after 1 second (attempt number 1)
tomcat not responding, re-trying after 1 second (attempt number 2)
Checking if PXF webapp is up and running...
PXF webapp is up
[root@hawq1 Downloads]#
```

## HAWQ external table test

HAWQ master node(hawq2.apache.local) 에서 HAWQ user(hskimsky) 로 작업합니다.

## External HDFS Table

```
echo "1,one
2,two
4,four
7,seven
3,three
5,five
9,nine
8,eight
6,six
10,ten" > numbers
hdfs dfs -mkdir /data
hdfs dfs -put numbers /data
psql -d postgres
```

```
\timing
create external table numbers (number int, text text) location (
'pxf://hawq2.apache.local:51200/data/numbers?Profile=HdfsTextSimple') format 'text' (delimiter E',');
select * from numbers order by number;
```

```
[hskimsy@hawq2 ~]$ echo "1,one
> 2,two
> 4,four
> 7,seven
> 3,three
> 5,five
> 9,nine
> 8,eight
> 6,six
> 10,ten" > numbers
[hskimsy@hawq2 ~]$ hdfs dfs -mkdir /data
16/06/05 14:15:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
java classes where applicable
[hskimsy@hawq2 ~]$ hdfs dfs -put numbers /data
16/06/05 14:15:52 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
java classes where applicable
[hskimsy@hawq2 ~]$ psql -d postgres
psql (8.2.15)
Type "help" for help.
```

```
postgres=# \timing
Timing is on.
postgres=# create external table numbers (number int, text text) location
('pxf://hawq2.apache.local:51200/data/numbers?Profile=HdfsTextSimple') format 'text' (delimiter E',');
CREATE EXTERNAL TABLE
Time: 43.028 ms
postgres=# select * from numbers order by number;
 number | text
-----+-----
      1 | one
      2 | two
      3 | three
      4 | four
      5 | five
      6 | six
      7 | seven
      8 | eight
      9 | nine
     10 | ten
(10 rows)

Time: 100.009 ms
```

## External Web Table

```
drop external table if exists sacramento;
create external web table sacramento (
  street text,
  city text,
  zip text,
  state text,
  beds int,
  baths int,
  sq__ft int,
  type text,
  sale_date timestamp with time zone,
  price int,
  latitude decimal,
  longitude decimal
)
location ('http://samplecsvs.s3.amazonaws.com/Sacramentorealestatetransactions.csv')
format 'csv' (header);
select type, max(price) from sacramento group by type order by max(price);
```

```
postgres=# create external web table sacramento (
postgres(#   street text,
postgres(#   city text,
postgres(#   zip text,
postgres(#   state text,
postgres(#   beds int,
postgres(#   baths int,
postgres(#   sq__ft int,
postgres(#   type text,
postgres(#   sale_date timestamp with time zone,
postgres(#   price int,
postgres(#   latitude decimal,
postgres(#   longitude decimal
postgres(# )
postgres=# location ('http://samplecsvs.s3.amazonaws.com/Sacramentorealestatetransactions.csv')
postgres=# format 'csv' (header);
NOTICE:  HEADER means that each one of the data files has a header row.
CREATE EXTERNAL TABLE
Time: 7.685 ms
postgres=# select type, max(price) from sacramento group by type order by max(price);
   type   | max
-----+-----
Unkown    | 275000
Condo     | 360000
Multi-Family | 416767
Residential | 884790
(4 rows)

Time: 1158.562 ms
postgres=#
```

The end. Thanks.

## References

- <https://cwiki.apache.org/confluence/display/HAWQ/Build+and+Install>
- <https://github.com/apache/incubator-hawq/commit/ec16d3162eb1650a1c87d39b6b36ee11f7995954>
- <http://hawq.docs.pivotal.io/docs-hawq/topics/about.html>
- <http://hdb.docs.pivotal.io/>
- <https://support.spatialkey.com/spatialkey-sample-csv-data/>

- <https://www.postgresql.org/docs/8.2/static/datatype-numeric.html>