FUSE对接FusionInsight HDFS

适用场景

> fuse 2.8.3 <-> FusionInsight HD V100R002C60U20（非安全模式）

说明

通过使用FUSE组件，可以使用将远端的HDFS文件系统mount到本端的Linux系统中使用

配置对接

- 安装jdk1.8

  ```
  tar -xvf jdk-8u112-linux-x64.tar.gz
  ```

- 配置环境变量 `/etc/profile` ，加入如下内容，source环境变量

  ```
  export JAVA_HOME=/opt/jdk1.8.0_112
  export CLASSPATH=.:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar
  export PATH=$JAVA_HOME/bin:$PATH
  ```

  ```
  source /etc/profile
  ```

- 安装rpm包

  ```
  yum install fuse fuse-devel fuse-libs
  ```

- 安装FusionInsight客户端，具体请参见产品文档的 **安装客户端** 章节，例如客户端安装目录为 `/opt/hadoopclient/`

- 下载Hadoop-2.7.2源码hadoop-2.7.2-src.tar.gz，编译fuse_dfs，将编译好的fuse_dfs拷贝到/opt目录下

- 将源码下的fuse_dfs_wrapper.sh脚本拷贝至 `/opt` 目录下，并根据实际情况作如下修改(修改HADOOP_PREFIX和JAVA_HOME的配置)：

  ```
  export HADOOP_PREFIX=/opt/hadoopclient/HDFS/hadoop
  if [ "$OS_ARCH" = "" ]; then
  export OS_ARCH=amd64
  fi

  if [ "$JAVA_HOME" = "" ]; then
  export  JAVA_HOME=/opt/jdk1.8.0_112
  fi

  if [ "$LD_LIBRARY_PATH" = "" ]; then
  export LD_LIBRARY_PATH=$JAVA_HOME/jre/lib/$OS_ARCH/server:$HADOOP_PREFIX/lib/native:/usr/local/lib
  fi

  # If dev build set paths accordingly
  if [ -d $HADOOP_PREFIX/share ]; then
    export HADOOP_PREFIX=$HADOOP_PREFIX
    for f in ${HADOOP_PREFIX}/share/hadoop/hdfs/*.jar ; do
      export CLASSPATH=$CLASSPATH:$f
    done
    for f in $HADOOP_PREFIX/share/hadoop/hdfs/lib/*.jar; do
      export CLASSPATH=$CLASSPATH:$f
    done
    for f in ${HADOOP_PREFIX}/share/hadoop/common/lib/*.jar; do
    export CLASSPATH=$CLASSPATH:$f
    done

    for f in ${HADOOP_PREFIX}/share/hadoop/common/*.jar; do
    export CLASSPATH=$CLASSPATH:$f
    done

    export PATH=/opt:$PATH
    export LD_LIBRARY_PATH=$JAVA_HOME/jre/lib/$OS_ARCH/server:$HADOOP_PREFIX/lib/native:/usr/local/lib:$LD_LIBRARY_PATH
  fi

  fuse_dfs $@
  ```

- 更改文件权限

```
chmod 755 fuse_dfs
chmod 755 fuse_dfs_wrapper.sh
```

- Source环境变量

```
source /opt/hadoopclient/bigdata_env
```

- 本地创建mount目录

```
mkdir -p /mnt/hdfs
```

- 执行挂载脚本，其中162-1-95-196是HDFS的NameNode(hacluster,主)的主机名

```
./fuse_dfs_wrapper.sh dfs://162-1-95-196:25000 /mnt/hdfs/
```

```
[root@fuse-test opt]# ./fuse_dfs_wrapper.sh dfs://162-1-95-196:25000 /mnt/hdfs/
./fuse_dfs_wrapper.sh: line 1: !/usr/bin/env: No such file or directory
INFO /opt/hadoop-2.7.2-src/hadoop-hdfs-project/hadoop-hdfs/src/main/native/fuse-dfs/fuse_options.c:164 Adding FUSE arg /mnt/hdfs/
```

- 查看/mnt/hdfs目录

```
[root@fuse-test hdfs]# pwd
/mnt/hdfs
[root@fuse-test hdfs]# ll
total 40
drwxrwxrwx.  2 nobody nobody 4096 Mar 17 17:33 app-logs
drwxrwxrwx.  3 nobody nobody 4096 Mar 17 17:35 apps
drwxrwxr-x.  3 nobody nobody 4096 Mar 17 17:33 datasets
drwxrwxr-x.  2 nobody nobody 4096 Mar 17 17:33 datastore
drwxrwx---.  2 nobody nobody 4096 Mar 17 17:33 flume
drwxrwx---.  2 nobody nobody 4096 Mar 17 17:34 hbase
drwxrwxrwx.  4 nobody nobody 4096 Mar 17 17:33 mr-history
drwxrwxrwx.  7 nobody nobody 4096 Mar 24 14:19 sparkJobHistory
drwxrwxrwx. 14 nobody nobody 4096 Mar 24 11:35 tmp
drwxrwxrwx.  6 nobody nobody 4096 Mar 17 17:36 user
[root@fuse-test hdfs]# hdfs dfs -ls /
Found 10 items
drwxrwxrwx+  - hdfs    hadoop          0 2017-03-17 17:33 /app-logs
drwxrwxrwx+  - hive    hive            0 2017-03-17 17:35 /apps
drwxrwxr-x+  - hdfs    hadoop          0 2017-03-17 17:33 /datasets
drwxrwxr-x+  - hdfs    hadoop          0 2017-03-17 17:33 /datastore
drwxrwx---+  - flume   hadoop          0 2017-03-17 17:33 /flume
drwxrwx---+  - hbase   hadoop          0 2017-03-17 17:34 /hbase
drwxrwxrwx+  - mapred  hadoop          0 2017-03-17 17:33 /mr-history
drwxrwxrwx+  - spark   hadoop          0 2017-03-24 14:19 /sparkJobHistory
drwxrwxrwx+  - hdfs    hadoop          0 2017-03-24 11:35 /tmp
drwxrwxrwx+  - hdfs    hadoop          0 2017-03-17 17:36 /user
```

- 如果需要卸载挂载点，执行umount /mnt/hdfs即可

验证对接

- hdfsclient写

```
dd if=/dev/zero bs=4096 count=1024 | hadoop fs -put - /tmp/fuse/hdfsclient-01.dat
```

```
[root@fuse-test hdfs]# dd if=/dev/zero bs=4096 count=1024 | hadoop fs -put - /tmp/fuse/hdfsclient-01.dat
1024+0 records in
1024+0 records out
4194304 bytes (4.2 MB) copied, 3.19247 s, 1.3 MB/s
```

```
[root@fuse-test fuse]# pwd
/mnt/hdfs/tmp/fuse
[root@fuse-test fuse]# ll
total 12288
-rw-r--r--. 1 root nobody 4194304 Mar 24 11:52 hdfsclient-01.dat
-rw-r--r--. 1 root nobody 4194304 Mar 23 17:50 test-01.dat
-rw-r--r--. 1 root nobody 4194304 Mar 23 19:37 test-02.dat
```

- hdfsclient读

```
hadoop fs -get /tmp/fuse/hdfsclient-01.dat - > /dev/null
```

```
[root@fuse-test fuse]# hadoop fs -get /tmp/fuse/hdfsclient-01.dat - > /dev/null
[root@fuse-test fuse]#
```

- fuse写

```
dd if=/dev/zero bs=4096 count=1024  of=/mnt/hdfs/tmp/fuse/fuse-01.dat
```

```
[root@fuse-test fuse]# dd if=/dev/zero bs=4096 count=1024  of=/mnt/hdfs/tmp/fuse/fuse-01.dat
1024+0 records in
1024+0 records out
4194304 bytes (4.2 MB) copied, 0.0390021 s, 108 MB/s
```

```
[root@fuse-test fuse]# ll
total 16384
-rw-r--r--. 1 root nobody 4194304 Mar 24 11:59 fuse-01.dat
-rw-r--r--. 1 root nobody 4194304 Mar 24 11:52 hdfsclient-01.dat
```

- fuse读

```
dd if=/mnt/hdfs/tmp/fuse/fuse-01.dat  bs=4096  of=/dev/null
```

```
[root@fuse-test fuse]# dd if=/mnt/hdfs/tmp/fuse/fuse-01.dat  bs=4096  of=/dev/null
1024+0 records in
1024+0 records out
4194304 bytes (4.2 MB) copied, 0.077961 s, 53.8 MB/s
```