HPCC Systems: HPCC FUSE

HPCC File System in User Space (experimental)

Author: Franz Nisswandt, Architect/Experimental FUSE driver provided by Nigel Hicks

Date: April 7, 2011

LexisNexis®

Table of Contents

Objective

Data files on the HPCC cluster are traditionally accessed via SPRAY and DESPRAY. These functions are highly optimized for the transfer of (typically) delimited, fixed width, or XML files to and from the cluster.

Sometimes, however, it would be useful to be able to "view the file directly" – without having to first DESPRAY the file or write ECL code. This magic is provided by a FUSE driver for the HPPC cluster (written by Nigel Hicks, still in experimental stages).

What's a FUSE?

Fuse stands for File System in User Space. In basic terms, a FUSE driver enables a user to "mount data sources" that would not be "mountable" in the traditional sense. Where FUSE is concerned, "mountable" means we can interact with files as if they were a standard file system. For example, one is able to list/cat files, grep/sed/awk/etc as if the files were on the local file system.

One example might be an FTP FUSE driver. This driver would interact with an ftp server in the background, responding to user commands, and present the user with a standard unix directory structure. All of the FTP protocol commands and handshaking would be done by the driver layer. Instead of sending commands for traversing the directories, turning binary on/off, putting/getting files/etc... we simply interact with the mount point via standard unix commands. These commands would be automatically translated by the "FUSE driver" positioned between our "mount point" and the "ftp server" and voila: we can copy files to/from a local directory and the files are magically FTPed to the server!

In the HPCC world, our files are generally distributed on the nodes of a cluster. This means that for a 400 node cluster, the file is in 400 pieces. SPRAY and DESPRAY provide a method to "assemble" the file for outside consumption (or partitioning a file when loading) enabling us to import/export a contiguous file, instead of dealing with 400 segments scattered on the cluster.

Our FUSE driver "mounts" the HPCC cluster and accesses all of the applicable nodes behind the scenes and presents the cluster contents as a "file system" -- mounted for the user.

Disclaimer – Your Mileage May Vary!

This is experimental! It is a proof of concept not a production release. This is also a read-only HPCC FUSE driver. This should not be expected to perform like DESPRAYing a file, nor should it be expected to perform as fast as locally mounted disk. There is debug code for logging that is captured while the driver is in operation.

If you agree to these terms, that this is an experimental, pre-release tool, then let's press on.

How to Use HPCC Fuse

The first thing you need for this current experimental release of HPCC FUSE is an Ubuntu x64 linux machine. I'm using a VirtualBox Ubuntu 10.10 x64 virtual machine. I downloaded the Ubuntu 10.10 install ISO and built a VM for some isolated testing. Note that root access is required for this version.

Once you have access to the machine, open up a terminal session as root and create a suitable folder for extracting the tar file containing the dafuse binaries (see below). Note I simply created a /root/dafuse folder in my virtual machine and un-tarred the files (cd ~/dafuse ; tar xvf dafuse.tar).

Once you have the binaries in place, run the following commands to mount the dali:

cd ~/dafuse

export LD_LIBRARY_PATH=.

mkdir –p /mnt/dali

./dafuse mount /mnt/dali DALISERVER=x.x.x.x ← replace with YOUR dali IP address.

Once mounted, you can do an ls /mnt/dali and see a directory of your cluster:

```
root@flnubuntu64: ~/dafuse
File  Edit  View  Search  Terminal  Help
root@flnubuntu64:~/dafuse# tar tvf dafuse.tar
-rwxr-xr-x nigel/nigel  193527 2011-03-24 07:03 dafuse
-rwxr-xr-x nigel/nigel 4868230 2011-03-24 07:03 libdalibase.so
-rwxr-xr-x nigel/nigel  482930 2011-03-24 07:03 libenvironment.so
-rwxr-xr-x nigel/nigel      17 2011-03-24 07:03 libfuse.so
-rw-r--r-- nigel/nigel  213408 2011-03-24 07:14 libfuse.so.2
-rwxr-xr-x nigel/nigel  360029 2011-03-24 07:03 libhrpc.so
-rwxr-xr-x nigel/nigel 6478025 2011-03-24 07:03 libjlib.so
-rwxr-xr-x nigel/nigel  690085 2011-03-24 07:03 libmp.so
-rwxr-xr-x nigel/nigel 1005423 2011-03-24 07:03 libremote.so
root@flnubuntu64:~/dafuse# export LD_LIBRARY_PATH=.
root@flnubuntu64:~/dafuse# mkdir -p /mnt/dali
root@flnubuntu64:~/dafuse# ./dafuse mount /mnt/dali DALISERVER=10.173.249.1
root@flnubuntu64:~/dafuse# ls /mnt/dali
base       in                      result  thgorwatch   thorwatch
cert       noc-gt-test-spray       rttest  thor         thor_watch
fernando   persist                 spill   thor_10_219  thor_watchbase
hthor      production_watch_thor   temp    thor_data400
hthor_64   regress                 test    thor_data50
root@flnubuntu64:~/dafuse# 
```

Real, Live Files

Let's take a look at some files in ECLWATCH and then view the same files via HPCC FUSE. You will note the naming convention of :: (2 colons) to denote "folder separations" for organizing files. These :: separators will automatically equate to folders when we view this listing through FUSE:

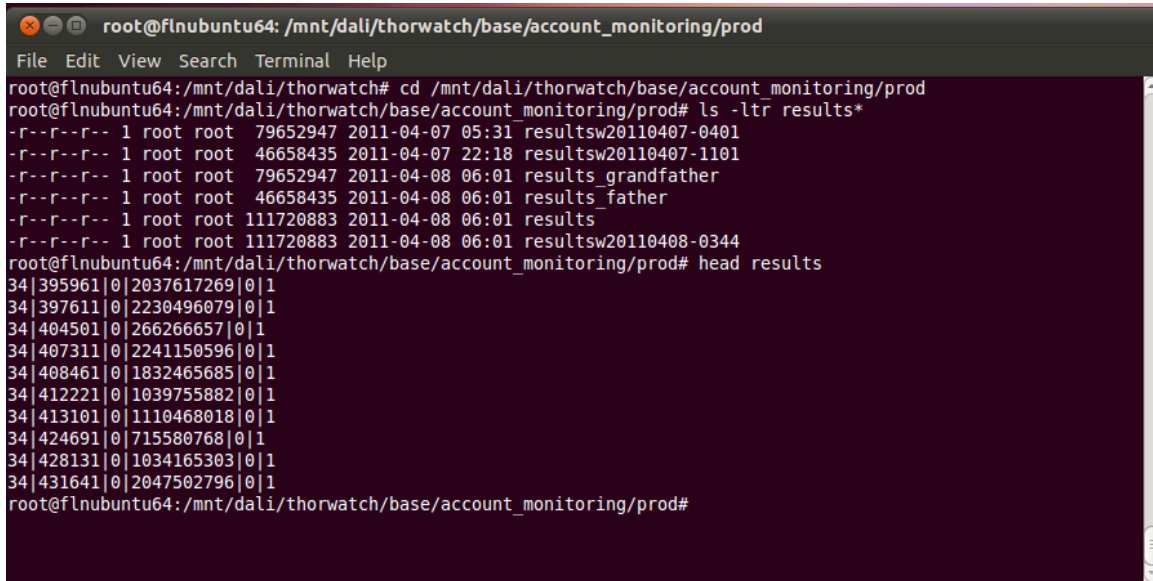| Logical Name	Description	Size	Records
thorwatch::base::account_monitoring::prod::results		111,720,883	3,376,514
thorwatch::base::account_monitoring::prod::results_father		46,658,435	1,492,646
thorwatch::base::account_monitoring::prod::results_grandfather		79,652,947	2,447,363

We will traverse our FUSE mount point to locate these files, replacing :: with /. Note that when we list the files, we will see both the superfiles and the logical files. We can access the files either by the superfile or by the logical files. Note also (of course) that dali format will be preserved (CSV/fixed/XML) when you interact with the file at your mount point.

Here, I'm changing to the "directory" where my results file is located, doing a simple listing, and head-ing the top few lines of the superfile contents:

cd /mnt/dali/thorwatch/base/account_monitoring/prod

ls –ltr results*

head results

```
🔴⚪⚫   root@flnubuntu64: /mnt/dali/thorwatch/base/account_monitoring/prod

 File  Edit  View  Search  Terminal  Help
root@flnubuntu64:/mnt/dali/thorwatch# cd /mnt/dali/thorwatch/base/account_monitoring/prod
root@flnubuntu64:/mnt/dali/thorwatch/base/account_monitoring/prod# ls -ltr results*
-r--r--r-- 1 root root  79652947 2011-04-07 05:31 resultsw20110407-0401
-r--r--r-- 1 root root  46658435 2011-04-07 22:18 resultsw20110407-1101
-r--r--r-- 1 root root  79652947 2011-04-08 06:01 results_grandfather
-r--r--r-- 1 root root  46658435 2011-04-08 06:01 results_father
-r--r--r-- 1 root root 111720883 2011-04-08 06:01 results
-r--r--r-- 1 root root 111720883 2011-04-08 06:01 resultsw20110408-0344
root@flnubuntu64:/mnt/dali/thorwatch/base/account_monitoring/prod# head results
34|395961|0|2037617269|0|1
34|397611|0|2230496079|0|1
34|404501|0|266266657|0|1
34|407311|0|2241150596|0|1
34|408461|0|1832465685|0|1
34|412221|0|1039755882|0|1
34|413101|0|1110468018|0|1
34|424691|0|715580768|0|1
34|428131|0|1034165303|0|1
34|431641|0|2047502796|0|1
root@flnubuntu64:/mnt/dali/thorwatch/base/account_monitoring/prod#
```

Conclusion

There are a wide variety of uses for a driver of this nature. You might want to use mysqlimport to load a database from a LOCAL text file; having HPCC FUSE loaded lets you access the file without despray.

Every tool has its optimal use case. HPCC FUSE is not a tool for pulling terabytes of data from your cluster. A temporary DESPRAY is nearly always going to be much, much faster and provide you with a local, independent copy of the data to manipulate if necessary prior to loading into your database/etc.

But when you want to interact directly with your cluster, especially if they are many SMALL files, FUSE will simplify your workflow, obviating the usual, required DESPRAY step.

References

Here are some references to FUSE:

http://en.wikipedia.org/wiki/Filesystem_in_Userspace

http://fuse.sourceforge.net/

For more information:
Website: http://hpccsystems.com/
Email: info@hpccsystems.com
US inquiries: 1.877.316.9669
International inquiries: 1.678.694.2200

About HPCC Systems
HPCC Systems from LexisNexis® Risk Solutions offers a proven, data-intensive supercomputing platform designed for the enterprise to solve big data problems. As an alternative to Hadoop, HPCC Systems offers a consistent data-centric programming language, two processing platforms and a single architecture for efficient processing. Customers, such as financial institutions, insurance carriers, insurance companies, law enforcement agencies, federal government and other enterprise-class organizations leverage the HPCC Systems technology through LexisNexis® products and services. For more information, visit http://hpccsystems.com.

About LexisNexis Risk Solutions
LexisNexis® Risk Solutions (http://lexisnexis.com/risk/) is a leader in providing essential information that helps customers across all industries and government predict, assess and manage risk. Combining cutting-edge technology, unique data and advanced scoring analytics, Risk Solutions provides products and services that address evolving client needs in the risk sector while upholding the highest standards of security and privacy. LexisNexis Risk Solutions is headquartered in Alpharetta, Georgia, United States.