

BATCH - Support #17364

Reviving CVMFS

08/03/2017 04:49 PM - Thomas Carroll

Status:	Work in progress	Start date:	09/12/2017
Priority:	Normal	Due date:	
Assignee:	Thomas Carroll	% Done:	50%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	18.05 hours
Description			
We need to revive the scripts that rsync the MINOS code to the CVMFS repository.			
I think we have all we need to update CVMFS. In general, I think we just need to update the rsync script for SLF6.			
Resources:			
1. README in the MINOS CVMFS repository			
2. rsync2CVMFS.sh script in home area of cvmfsmenos@oasiscfs			
Tasks:			
1. Create Redmine issue			
2. Make a wikipage for documentation			
3. commit rsync scripts to CVS (contrib/CVMFS) including setup_minossoft_FNALU_parser_CVMFS.sh for which Ashley Timmons left us this note:			
<pre># a manual file which i need to put in if it all gets wiped #rsync -a ashley90@minos53:/afs/fnal.gov/files/home/room1/ashley90/setup_minossoft_FNALU_parser_CVMFS.sh /cvmfs/minos.opensciencegrid.org/minos/minossoft/setup/setup_minossoft_FNALU_parser_CVMFS.sh</pre>			
Subtasks:			
Support # 17692: Add 64 bit directories			Closed
Support # 17889: Commit rsync scripts to CVS (contrib/CVMFS) including setup_minossoft_...			New

History

#1 - 08/04/2017 11:59 AM - Thomas Carroll

Rui and I have updated the the script rsync2CVMFS.sh script in home area of cvmfsmenos@oasiscfs. The script now uses tcarroll@minos61 for the server. We are now trying to decide if we should do a full rsync. This would include the following directories from /grid/fermiapp/minos:

1. minfarm/
2. minossoft/srt
3. minossoft/setup
4. minossoft/packages
5. scripts
6. products
7. minossoft/releases

Note that this script does not publish to CVMFS.

Art could you comment on if this makes sense?

#2 - 08/06/2017 09:06 AM - Arthur Kreymer

- % Done changed from 0 to 10

Thanks !

I will review the script, and offer some cleanups.

/grid/fermiapp is mounted on oasisdfs, so we can rsync without a remote account.
I will ask to have /minos/app (for future migration) and /minos/data (for release_data) mounted.

#3 - 08/10/2017 08:07 AM - Arthur Kreymer

- Due date set to 08/31/2017
- Status changed from New to Work in progress
- % Done changed from 10 to 50

I discussed the rsync issues with Dave Dykstra Aug 8.

We can use the directly mounted /grid/fermiapp source for now,
no need to use ssh in the rsync.
I will make a copy of release_data in /grid/fermiapp/minos

There is a standard wrapper script for rsync.
I will add this and a few other cleanups to a new rsync2CVMFS script.

#4 - 08/22/2017 03:04 PM - Rui Chen

After a short dry-run test, we got the following result:

```
[cvmfsminos@oasisdfs02 ~]$ . rsync2CVMFS.sh
=====
sub dir in question is = .
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/minifarm/ /cvmfs/minos.opensciencegrid.org/minos/.
start 2017-08-22 13:02:31
rsync: readlink_stat("/grid/fermiapp/minos/minifarm/Minossof/EXTERNAL/mysql-5.0.22/sql/share/japanese-sjis/errmsg.sys") failed: Permission denied
(13)
rsync error: some files/attrs were not transferred (see previous errors) (code 23) at main.c(1518) [generator=3.0.9]
finish 2017-08-22 13:15:04
took 753 sec realtime =====
sub dir in question is = minossof
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/minossof/srt /cvmfs/minos.opensciencegrid.org/minos/minossof
start 2017-08-22 13:15:04
finish 2017-08-22 13:15:04
took 0 sec realtime =====
sub dir in question is = minossof
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/minossof/setup /cvmfs/minos.opensciencegrid.org/minos/minossof
start 2017-08-22 13:15:04
finish 2017-08-22 13:15:05
took 1 sec realtime =====
sub dir in question is = minossof
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/minossof/packages /cvmfs/minos.opensciencegrid.org/minos/minossof
start 2017-08-22 13:15:05
finish 2017-08-22 13:35:11
took 1206 sec realtime =====
sub dir in question is = .
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/scripts /cvmfs/minos.opensciencegrid.org/minos/.
start 2017-08-22 13:35:11
finish 2017-08-22 13:35:11
took 0 sec realtime =====
sub dir in question is = .
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/products /cvmfs/minos.opensciencegrid.org/minos/.
start 2017-08-22 13:35:11
finish 2017-08-22 14:39:41
took 3870 sec realtime =====
sub dir in question is = minossof
rsync -n -a rchen2@minos61:/grid/fermiapp/minos/minossof/releases /cvmfs/minos.opensciencegrid.org/minos/minossof
start 2017-08-22 14:39:41
finish 2017-08-22 14:48:19
took 518 sec realtime
You are about to DISCARD ALL CHANGES OF THE CURRENT TRANSACTION for minos.opensciencegrid.org! Are you sure (y/N)? y
```

The dry-run test rsynced everything we processed but didn't output the file transferred, we will add -v flag to the next test.

#5 - 08/22/2017 03:54 PM - Rui Chen

After a full dry-run test. I have listed all the files that rsync will transfer in the future.
It is stored here:
/home/cvmfsminos/rsync_dry_run_test_full_file_output.txt

#6 - 09/14/2017 05:33 PM - Thomas Carroll

I am doing a live test of our updated rsync2CVMFS.sh script.

#7 - 09/15/2017 12:52 PM - Thomas Carroll

It looks like the test failed to publish to CVMFS. Here is a clip of the output messages:

```
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/packages/.cvmfscatalog' (errno: 13)
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/releases/.cvmfscatalog' (errno: 13)
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/setup/.cvmfscatalog' (errno: 13)
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/srt/.cvmfscatalog' (errno: 13)
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/admin/.cvmfscatalog' (errno: 13)
Failed to create nested catalog marker at '/cvmfs/minos.opensciencegrid.org/minos/products/prd/MINOS_OFFLINE/lem/.cvmfscatalog' (errno: 13)
Failed to apply .cvmfsdirtab
```

There are also many cases of rsync failing due to permissions.

The full output can be found in the file named liveRsyncTest.2017.09.14.out.

#8 - 09/24/2017 08:33 AM - Arthur Kreymer

Thanks, I will look into this next week.

The permissions are easy enough to fix,
with /grid/fermiapp/minos/scripts/modescan

#9 - 09/27/2017 10:44 AM - Arthur Kreymer

I scanned the directories we are now rsync'ing for directory and file counts, and size.

```
MSL=/minos/app/users/minisoft/maint
mkdir $MSL
mkdir $MSL/cvmfs

{ for DIR in ${DIRS} ; do
  date
  GFDIR=/grid/fermiapp/minos/${DIR}
  echo ${DIR}
  find ${GFDIR} -type d | wc -l
  find ${GFDIR} -type f | wc -l
  du -sm ${GFDIR}
done
} 2>&1 | tee ${MSL}/cvmfs/fromscan.201780924.log
```

Directory	dirs	files	MB
minifarm	100842	689530	36706
minossoft/srt	1	18	1
minossoft/setup	10	541	49
minossoft/packages	114168	1165282	11087
scripts	3	93	1
products	611558	3413515	130226
minossoft/releases	15684	502042	76602
minossoft_srt_64bit/srt	1	8	1
minossoft_srt_64bit/setup	6	322	3
minossoft_srt_64bit/packages	17093	195352	2222
minossoft_srt_64bit/releases	8384	181723	24008
minossoft_srt_64bit/minos_ups_products	401	2736	30

The largest product is MINOS_ROOT, with 66 versions and > 500K directories

```
for DIR in `ls /grid/fermiapp/minos/products/prd` ; do
printf "${DIR} "
find /grid/fermiapp/minos/products/prd/${DIR} -type d | wc -l
done | sort -n -k 2
...
oracle_client 1268
geant4source 1695
cmake 1947
cern 4831
gcc 8082
MINOS_EXTERN 12558
GENIE 16180
MINOS_ROOT 557362
```

There is an extra published directory minos/bin, not rsync'd by the default script.

#10 - 09/28/2017 12:49 PM - Thomas Carroll

I ran /grid/fermiapp/minos/scripts/modescan as minsoft to cleanup permissions in /grid/fermiapp/minos/products and as minfarm to cleanup permission in /minos/app/grid/minfarm/Minossoft.

I edited /cvmfs/minos.opensciencegrid.org/.cvmfsdirtab to include paths to the 64-bit directories.

#11 - 09/30/2017 10:35 AM - Thomas Carroll

Art concluded that we were unable to publish our cvmfs repository due to the soft link /grid/fermiapp/minos/products/prd/MINOS_OFFLINE which points to /grid/fermiapp/minos/minossoft.

Here is some more information about the soft link:

```
lrwxrwxrwx 1 3648 5111 30 Oct 15 2009 MINOS_OFFLINE -> /grid/fermiapp/minos/minossoft
```

After looking at the comments in the rsync2CVMFS.sh script, I realized that this had been an issue in the past, and one solution is to remove this soft link from the repository before publishing.

I edited rsync2CVMFS.sh so that it would no longer publish. Then I ran the script using nohup copying the output to liveRsyncTestWithoutPublishing.2017.09.28.out in the home directory for cvmfsminos@oasiscfs. Next, I removed the MINOS_OFFLINE soft link from the repository. Finally, I executed the command to publish the repository using nohup and directed the output to publishCVMFS.2017.09.29.out. It took several hours for the update to propagate to the repository on the minos machines.

It looks like we have successfully published our cvmfs repository.

#12 - 10/11/2017 12:21 PM - Arthur Kreymer

Publication got stuck per INC000000892842 10/06 tcarroll
Fixed by Fife support people,
an internal cleanup issue with transaction aborts.
Not our fault, except that we're stressing the system with millions of file, mainly root versions.

#13 - 10/11/2017 12:21 PM - Arthur Kreymer

Date: Tue, 10 Oct 2017 22:27:40 +0000
From: Thomas J Carroll <tjcarroll@utexas.edu>
To: Arthur Kreymer <kreymer@fnal.gov>
Cc: minos_batch <minos_batch@fnal.gov>
Update regarding CVMFS:
minfarm is published to the cvmfs repo. Its contents are in /cvmfs/minos.opensciencegrid.org/minos/ [minos.opensciencegrid.org].

A ls of /grid/fermiapp/minos/minfarm yields:
Minossoft SAM bin condor_submit glide_submit scripts

And a ls of /cvmfs/minos.opensciencegrid.org/minos/ [minos.opensciencegrid.org] yields:
ChangeHardLinks.sh condor_submit products typescript
Minossoft glide_submit release_data
SAM minossoft scripts
bin minossoft_srt_64bit temp

Now, whether or not these minfarm directories are used is separate issue.

#14 - 10/11/2017 12:42 PM - Arthur Kreymer

Excellent to know where minfarm was going !

Apparently the trailing / in the minfarm/ source path

caused its content to be put directly in minos rather than under minos/minfarm.

There are nearly 700K files in minfarm.
We should remove them from cvmfs if they are not used.
The only minfarm files modified in the last year are in minfarm/scripts.

```
MINOSDATAGPVM01 > find /grid/fermiapp/minos/minfarm -type f | wc -l  
689533
```

Also, the cvmfs scripts directory contains files from both minfarm/scripts and scripts.
This should be cleaned up as part of the minfarm removal.

#15 - 10/14/2017 09:27 AM - Arthur Kreymer

In a single typical root version
/grid/fermiapp/minos/products/prd/MINOS_ROOT/Linux2.6-GCC_4_5/v5-32-0
there are
53293 files and directories including src and .svn
11594 excluding these directories.

I verified that rsync can exclude these, using
--exclude 'src' --exclude '.svn'

I verified there are no files named src or .svn in the products tree.

I evaluated the effect of this on products overall,
counting files without and with the exclusion.
The count is reduced from about 4 to 1 million

```
cd /grid/fermiapp/minos/products  
  
set nohup  
{ date ; rsync -r . | wc -l  
  date ; rsync -r . --exclude 'src' --exclude '.svn' | wc -l  
  date ; } > /dev/shm/prodrsync 2>&1 &  
  
Fri Oct 13 12:38:13 CDT 2017  
4033295  
Fri Oct 13 12:49:03 CDT 2017  
1162029  
Fri Oct 13 12:54:16 CDT 2017
```

#16 - 10/21/2017 10:59 AM - Arthur Kreymer

I have updated .bashrc under cvmfsminos@oasis.cfs.fnal.gov
to remove obsolete items and provide a record of who logged in.

This requires that a kerberos ticket be forwarded,
which will be needed for future rsyncs and web logging.

```
.bashrc updates  
  Removed unused setup_minos function  
  Removed obsolete cat of ~/INSTRUCTIONS  
  Added logging of kerberos principal to log/login.log  
  
Saved .bashrc files in  
.bashrc.20170918  
.bashrc.20170921
```

#17 - 10/24/2017 01:30 PM - Arthur Kreymer

I have deployed a new rsync2CVMFS script, with major improvements.

All invocations are logged in log/run.log,
including the username of the person logged in.

Detailed logs are generated in files like
log/rsync-20171024-1321.txt

Qualifiers are available for verbose and noop mode,
and for selecting a single directory for rsync,
and for comments similar to cvs :

```
-v verbose
```

```
-n noop
-m "any message you would like to leave in log/run.oog"
-d directory ( must match an entry in the default list )
```

I have also made a convenient rsync2CVMFS alias.

Typical testing usage :

```
rsync2CVMFS -n -v -m "try scripts with -n -v" -d scripts
```

I have updated scripts just now :

```
rsync2CVMFS -m "try scripts for real" -d scripts
```

#18 - 10/24/2017 01:36 PM - Arthur Kreymer

Next things to do :

- Remove minfarm/
 - Remove minfarm/ from DIRS
 - remove minfarm/ files
 - start a transaction
 - remove minfarm/ directories including scripts
 - publish
 - run rsync2CVMFS -d scripts to put minos/scripts back
- remove root .svn and src areas
 - add removal of these to rsync2CVMFS
 - rsync2CVMFS -d products

#19 - 10/24/2017 04:35 PM - Arthur Kreymer

Oops, I have corrected an error in my earlier listing of options. They are v(erbos) n(oop) m(essage) d(irectory)

#20 - 10/28/2017 01:55 PM - Arthur Kreymer

I have updated rsync2CVMFS as follows :

- removed minfarm/.
- exclude src and .svn directories in the products directory
- new -t qualifier for testing, writes to /var/tmp/minos use this only with small areas like scripts
- changed the log message about sub dir to show the actual directory

The products file count is down

was	4103999
now	1228985

The total file count in CVMFS is now 3498148, much in minossoft/packages

I removed the scripts and bin areas from /cvmfs/minos.opensciencegrid.org/cvmfsdirtab They had been there because of minfarm/

#21 - 10/28/2017 02:00 PM - Arthur Kreymer

I removed these obsolete files from /grid/fermiapp/minos/products

```
lrwxrwxrwx 1 3648 5111      31 Oct  3  2009 GROWFS -> /minos/data/mindata/GROWFS/d119
-rw-r--r-- 1 3648 5111      44 Mar 27  2010 .growfschecksum
-rw-r--r-- 1 3648 5111 108145478 Mar 27  2010 .growfsdir
```

#22 - 10/28/2017 02:09 PM - Arthur Kreymer

When it is provided by Robert, I will add to rsync2CVMFS a call to the script to modify UPS configuration files after rsync, This should be the final major thing we need for production use of cvmfs.

#23 - 10/30/2017 12:06 PM - Arthur Kreymer

I have moved release_data from /minos/data to /grid/fermiapp/minos/release_data for compatibility with rsync2CVMFS.

I have moved loonexe from /minos/data/minfarm/loonexe
to /grid/fermiapp/minos/release_data/minfarm/loonexe

I have installed symlinks at the old locations for compatibility.

I have run rsync2CVMFS -d release_data -m 'release_data'

#24 - 10/30/2017 03:31 PM - Robert Hatcher

Could we try my first pass at fixing paths hidden in various files.

between the rsync and the publish stage, run the command:

```
~/minossoft/fix_cvmfs_paths run
```

Without the "run" arg it will just tell you which files it will process, but not actually edit them.

If this works (or mostly works) then we can commit this script to the repository.

#25 - 11/02/2017 12:11 PM - Arthur Kreymer

I have updated rsync2CVMFS to use Robert's products path fixing script
/grid/fermiapp/minos/scripts/fix_cvmfs_paths

I ran this, finishing around 10:54 this morning.
See /home/cvmfsminos/log/rsync-20171102-1029.txt

No changes seem to have been made in /cvmfs on oasisdcs.
Robert fixed a typo, running again.

#26 - 11/02/2017 12:46 PM - Arthur Kreymer

The cvmfs products paths seem to be modified and propagated.

#27 - 11/02/2017 01:29 PM - Robert Hatcher

Is the rsync script not updating

```
/cvmfs/minos.opensciencegrid.org/minos/minossoft/setup/
```

from

```
/grid/fermiapp/minos/minossoft/setup/
```

As the ones in CMVFS are older/stale relative to the /grid/fermiapp

This leads to a bunch of (harmless) warnings about

```
WARNING: '/cvmfs/minos.opensciencegrid.org/minos/externals' is not a directory
```

and may also be the cause for the 64-bit setup not working.

But yes, neither setup works cleanly to give CVMFS-only setup. I'll continue to investigate (but please update the above)

#28 - 11/02/2017 01:35 PM - Robert Hatcher

also the pair:

```
/cvmfs/minos.opensciencegrid.org/minos/minossoft_srt_64bit/setup/
```

from

```
/grid/fermiapp/minos/minossoft_srt_64bit/setup/
```

#29 - 11/02/2017 02:54 PM - Arthur Kreymer

I had updated just products.
I have just started a full update.
Results in an hour or so, perhaps 15:00.

#30 - 11/02/2017 05:41 PM - Arthur Kreymer

The full rsync2CVMFS finished.

tail -3 log/run.log

RUN Thu Nov 2 14:38:00 CDT 2017 version 20171102 by kreymer

END Thu Nov 2 15:37:31 CDT 2017 logged to /home/cvmfsminos/log/rsync-20171102-1438.txt