

Robinhood: **Client & Server**

And now for something ~~completely~~ slightly different....

- **Megan Larko, BAE Systems**

Why?

Robinhood Lustre Monitoring Software is a good product; why do something differently?

Because...

- High-Performance Computing (HPC) environment users desire information which Robinhood can provide without using a web interface or logging into the Robinhood server directly. (In a few specialized environments web services are not permitted, or are impractical.)
- Some HPC compute nodes are small diskless processing units (limited write ability and small OS space).
- Getting meta-data from the Robinhood database is faster than walking the Lustre File System (LFS).
- Sample Tasks:
 - Quickly gathering meta-data or user extended attributes
 - Scripting queries & parsing response within a job

Test Environment

- System 1 (rbh server):
 - ~128GB RAM
 - RHEL 6.7 (kernel 2.6.32-573.12.1.el6.x86_64)
 - Robinhood 3.1-1
 - MySQL-5.6.35
 - Lnet: Mellanox IB 3.1-1.0.3 (ko2ibIInd)

- System 1 (rbh client):
 - ~32GB RAM
 - SLES 11 SP3 (kernel 3.0.101-0.46.1-1.0502.8871)
 - Robinhood 3.0-0.rc1
 - MySQL-5.5.47
 - Lnet: vendor proprietary high-speed IInd

- System 2 (rbh server):
 - ~256GB RAM
 - CentOS 7.3.1611 (kernel 3.10.0-514.2.2.el7.x86_64)
 - Robinhood 3.1-1
 - MySQL-5.6.36-1
 - Lnet: 40Gb Eth bond (ksockIInd)

- System 2 (rbh client):
 - ~128GB RAM
 - CentOS 7.3.1611 (kernel 3.10.0-514.26.2.el7.x86_64)
 - Robinhood 3.1-1
 - Mariadb-5.5.56-2
 - Lnet: Mellanox IB 3.4 (ko2ibIInd)

Method (Robinhood Server):

- Install Robinhood software per customary directions.
- Open standard mysql port 3306.
- Establish a read-only user of the robinhood mount point data base (robbie)
(GRANT SELECT EXECUTE and create a conf file for ro clients).
- OPTIONAL: To speed up the processing, a new index for the Robinhood database was created, and the robinhood configuration file and mysql configuration file had a couple of lines added. (Detailed in slides "Database Tuning").

Method (Robinhood Client):

- Build Robinhood software on one representative client.
- Copy only /usr/lib64/robinhood directory onto other like clients.
- Copy onto those clients only those rbh commands which users need.
(/usr/sbin/rbh-report, /usr/sbin/rbh-find and /usr/sbin/rbh-diff)
- Install compatible mysql client software on all clients.
- Verify connection to Robinhood server mysql port 3306
(`"mysql -u robbie -p -h <rbh-server>"`)
- Copy the includes files into /etc/robinhood.d/includes/.
(alerts.inc, check.inc, lhsm.inc, tmpfs.inc)
- Place read-only user (robbie) conf and access credentials in client:/etc/robinhood.d directory.

Method (Robinhood Client Options):

- OPTIONAL: Copy the man1 pages for those Robinhood commands placed on the clients.

```
(cp $MySrcDir/share/man/man1/rbh-report.1 /usr/share/man/man1/)
```

Database Tuning 1:

- File rhb3.conf (lines added to end of file):

```
Entry Processor {  
    max_batch_size = 1;  
    nb_threads = 24;  
    STAGE_BD_APPLY_threads_max = 16;  
}
```
- File my.cnf (lines added to [mysqld] section):

```
Innodb_file_format = Barracuda  
Innodb_large_prefix = ON
```

Database Tuning 2:

- Establishment of a new index on Robinhood database:
This index was created on the ENTRIES table based upon the fields--
fileclass, size, blocks, creation_time, last_access, last_mod
- The new index is linked to the underlying SQL. It is this change that necessitated the innodb_file_format and innodb_large_prefix items in the my.cnf file.
- A restart of mysql is required after these changes.

Database Tuning 3:

- Following the restart of mysql, the ENTRIES table needs to have an attribute changed. A simple text file was created containing the line:
Use robinhood_db; ←or whatever name is appropriate for LFS rbh db
Alter table ENTRIES ROW_FORMAT = DYNAMIC;
- A batch mysql query is run in background mode or change interactively:
mysql -u robinhood -p<rbh_db_pass> < simple.txt > changeentriesoutput.txt
- Finally the create index command was run in batch mode creating a file containing the lines:
Use robinhood_db; Create index fileclassindex on ENTRIES (fileclass, size, blocks, creation_time, last_access, last_mod);
- A batch query was run in background mode:
mysql -u robinhood -p<rbh_db_pass> < <theCreatedFile.txt> > createindexoutput.txt &
- The completion of the robinhood db tuning reduced the time of an rbh-report -class-info query of 2 billion records from 4 hours to about 20 minutes.

Result:

The end result of this modified installation of Robinhood (on those systems capable of communicating with the Robinhood server over a mysql port) is a group of machines capable of leveraging the power of Robinhood Lustre File System database in a more light-weight software set-up.

Examples:

- Sys1 Server:

```

~j# rbh-report --class-info
Using config file '/etc/robinhood.d/j_rbh3.conf'.
      fileclass,      count,      volume,      spc_used,      min_size,      max_size,      avg_size
      all_j_files,      40,      160.00 KB,      160.00 KB,      4.00 KB,      4.00 KB,      4.00 KB
empty_files+all_j_files,      2,      0,      0,      0,      0,      0
larkoMPI+small_files+all_j_files,      1,      491.34 MB,      491.56 MB,      491.34 MB,      491.34 MB,      491.34 MB
larkoMPI+std_files+all_j_files,      17,      60.81 GB,      60.81 GB,      1.01 GB,      4.63 GB,      3.58 GB
larkoSTD+small_files+all_j_files,      6,      2.82 GB,      2.82 GB,      105.36 MB,      977.01 MB,      481.09 MB
larkoSTD+std_files+all_j_files,      12,      37.48 GB,      37.48 GB,      1.04 GB,      4.67 GB,      3.12 GB
      small_files+all_j_files,      89,      16.56 MB,      16.79 MB,      8,      7.86 MB,      190.56 KB
      std_files+all_j_files,      2,      50.00 GB,      50.00 GB,      25.00 GB,      25.00 GB,      25.00 GB

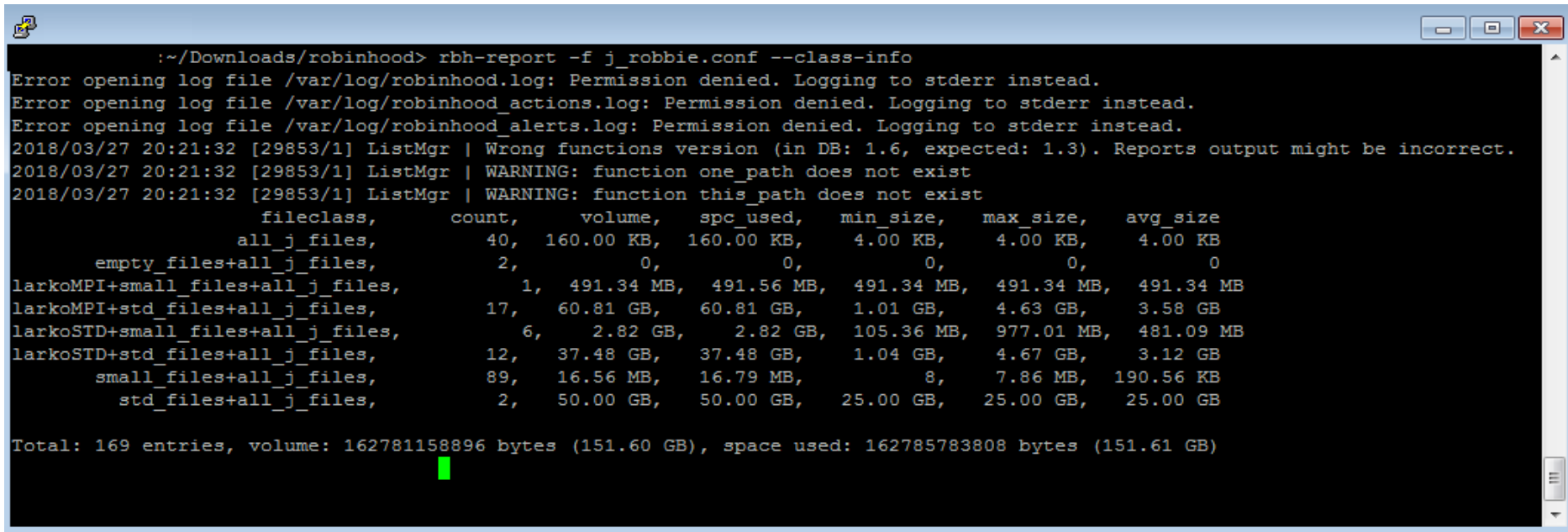
Total: 169 entries, volume: 162781158896 bytes (151.60 GB), space used: 162785783808 bytes (151.61 GB)

```

Note: The fileclass includes information based upon Lustre user extended attributes. We have a keen interest in using user_xattrs.

Examples:

- Sys1 Client:



```

~/Downloads/robinhood> rbh-report -f j_robbie.conf --class-info
Error opening log file /var/log/robinhood.log: Permission denied. Logging to stderr instead.
Error opening log file /var/log/robinhood_actions.log: Permission denied. Logging to stderr instead.
Error opening log file /var/log/robinhood_alerts.log: Permission denied. Logging to stderr instead.
2018/03/27 20:21:32 [29853/1] ListMgr | Wrong functions version (in DB: 1.6, expected: 1.3). Reports output might be incorrect.
2018/03/27 20:21:32 [29853/1] ListMgr | WARNING: function one_path does not exist
2018/03/27 20:21:32 [29853/1] ListMgr | WARNING: function this_path does not exist

```

fileclass,	count,	volume,	spc_used,	min_size,	max_size,	avg_size
all_j_files,	40,	160.00 KB,	160.00 KB,	4.00 KB,	4.00 KB,	4.00 KB
empty_files+all_j_files,	2,	0,	0,	0,	0,	0
larkoMPI+small_files+all_j_files,	1,	491.34 MB,	491.56 MB,	491.34 MB,	491.34 MB,	491.34 MB
larkoMPI+std_files+all_j_files,	17,	60.81 GB,	60.81 GB,	1.01 GB,	4.63 GB,	3.58 GB
larkoSTD+small_files+all_j_files,	6,	2.82 GB,	2.82 GB,	105.36 MB,	977.01 MB,	481.09 MB
larkoSTD+std_files+all_j_files,	12,	37.48 GB,	37.48 GB,	1.04 GB,	4.67 GB,	3.12 GB
small_files+all_j_files,	89,	16.56 MB,	16.79 MB,	8,	7.86 MB,	190.56 KB
std_files+all_j_files,	2,	50.00 GB,	50.00 GB,	25.00 GB,	25.00 GB,	25.00 GB

```

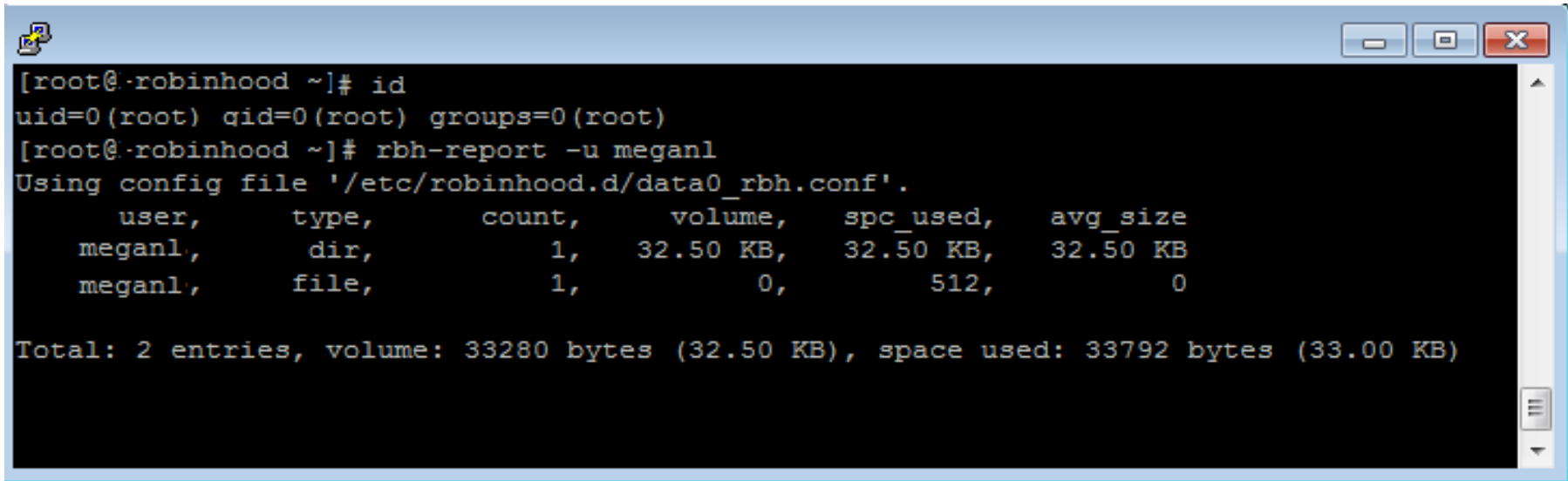
Total: 169 entries, volume: 162781158896 bytes (151.60 GB), space used: 162785783808 bytes (151.61 GB)

```

Same information viewed from the ro (i.e. "robbie") client.

Examples:

- Sys2 Server:



```
[root@-robinhood ~]# id
uid=0(root) gid=0(root) groups=0(root)
[root@-robinhood ~]# rbh-report -u meganl
Using config file '/etc/robinhood.d/data0_rbh.conf'.
  user,      type,      count,      volume,      spc_used,      avg_size
  meganl,    dir,       1,         32.50 KB,    32.50 KB,     32.50 KB
  meganl,    file,      1,          0,          512,          0

Total: 2 entries, volume: 33280 bytes (32.50 KB), space used: 33792 bytes (33.00 KB)
```

The command response is cleaner for the root user, but that is not how this Robinhood modification is to be used.

Examples:

- Sys2 Server:

```

$ rbh-report -u meganl.
Using config file '/etc/robinhood.d/data0_rbh.conf'.
Error opening log file /var/log/robinhood/lustre_data0.log: Permission denied. Logging to stderr instead.
Error opening log file /var/log/robinhood/lustre_actions_data0.log: Permission denied. Logging to stderr instead.
Error opening log file /var/log/robinhood/lustre_alerts_data0.log: Permission denied. Logging to stderr instead.
2018/03/28 16:19:35 [44654/1] _set_mount_point | Error: failed to get FID for special directory </mnt/data-0/.lustre/fid/>: Permission denied.
2018/03/28 16:19:35 [44654/1] InitFS | Error -13 checking Filesystem
Warning: cannot access filesystem /mnt/data-0 (Permission denied), some reports may be incomplete or not available.
  user,      type,      count,      volume,      spc_used,      avg_size
  meganl,    dir,       1,          32.50 KB,     32.50 KB,     32.50 KB
  meganl,    file,      1,          0,           512,          0

Total: 2 entries, volume: 33280 bytes (32.50 KB), space used: 33792 bytes (33.00 KB)

```

NOTE: The /var/log error messages appear for all non-root users—even on the Robinhood server.

Issues:

- Extraneous error messages:
(Should a /var/log/robinhood/ be created on light weight client? Thinking no.)
- Security concerns:
All users can find/see the db read-only robbie password.
The database cell content is not yet protected via a permissions paradigm.
- Has not yet been tested under full user load.
- Has not yet incorporated user feedback.
- Old LFS that do not properly respond to "lfs data_version" check.

Further Work:

- Implement cell security on the mysql database.
- Expand testing to observe behavior of client access under load.
- Learning how to get to user_xattrs more efficiently.
- Clean-up the error messages regarding client-side non-existent /var/log files.
- Test using a Robinhood server monitoring more than one Lustre File System.

I'm open to ideas here.

In Appreciation:

- Joe Hoolihan – BAE Systems, Ret.
- Site Lustre Team (M.C. with K.C., B.P., J.T., & D. W.)
- Robinhood Software Team
(<https://github.com/cea-hpc/robinhood/wiki>)
- Lustre Software Team
(<https://lustre.org/>)



Questions?

■ Thank you!