Robinhood v4

Scaling RBH for the future
MariaDB
Architecture
Architecture
Architecture (WIP)
Architecture

```
rbh-scan
rbh-lustre-scan
...  
changelog-reader
```

```
rbh-find
...  
robinhood
```

```
mongoDB
```
How hard is it to implement a backend?
How hard is it to implement a backend?

The MongoDB backend is ~600 lines of code
How hard is it to implement a backend?

The MongoDB backend is ~600 lines of code

Most of it is boilerplate
The MongoDB backend is ~600 lines of code

Most of it is boilerplate *
Performance

- Lustre filesystem
- 400k entries
- Single server (collocated mds, mgs, and oss)
- Single client (!= server)
Performance

Speed comparison of different flavors of find

- gnu-find
- lfs-find
- rbh-find (v3)
- rbh-find (v4)

command line options

Ø -name 'a*' -size +512

completion time in seconds

0.1 1 10 100 1000
Speed comparison of rbh-find (v4) and rbh-find (v4) + path

<table>
<thead>
<tr>
<th>Command Line Options</th>
<th>rbh-find (v4)</th>
<th>rbh-find (v4) + path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>-name 'a*'</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>-size +512</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

rbh-find path1 path2 [...] is slower
Performance (summary)

Speed comparison of different find flavors

<table>
<thead>
<tr>
<th>Command Line Options</th>
<th>gnu-find</th>
<th>lfs-find</th>
<th>rbh-find (v3)</th>
<th>rbh-find (v4)</th>
<th>rbh-find (v4) + path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-name 'a*'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-size +512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

completion time in seconds

22/43
How come?
How come?
How come?
How come?
How come?
How come?
How come?
How come?
How come?
There is still lots to do!

- changelog_reader
- rbh-report
- web interface
- posix scanner *
Posix scanner or posix backend?
Posix scanner or posix backend?

![Diagram showing a network of components related to file system and database with a mid-point labeled 'rbh-translate'.]
Posix scanner or posix backend?
Posix scanner or posix backend?
Do we still need a policy engine?
Do we still need a policy engine?

robinhood --run=purge --target=class:old
Do we still need a policy engine?

robinhood --run=purge --target=class:old

Vs

rbh-find -atime +30 -execdir rm {} \

Main Goals

- Developer-friendly
- Admin-friendly
- Scalability
Thank you!

Questions?
Acknowledgements / copyrights

- MongoDB's logo - unknown - fair use
  https://www.mongodb.com/brand-resources

- MariaDB's seal - Mike Zinner - GFDL
  http://badges.mariadb.org/logo/Mariadb-seal-flat-browntext.svg

- PostgreSQL's elephant - Daniel Lundin - PostgreSQL's license
  https://wiki.postgresql.org/images/a/a4/PostgreSQL_logo.3colors.svg

- Lustre's logo - unknown - fair use