DE LA RECHERCHE À L’INDUSTRIE

RobinHood-v4 progress report
What is RobinHood?
RobinHood mirrors a filesystem’s metadata, and makes it queryable
• Mirror
  - Read Replica
  - Snapshot
  - Disaster Recovery

• Query
  - Expressiveness
  - Performance
  - Efficiency
Backend
Architecture
Architecture
Architecture

IN

FSEvents

C-API

Backend

OUT

FSEntries
Applications
rbh-sync synchronizes two RobinHood backends
rbh-sync SRC DEST
rbh-sync rbh:posix:/mnt/scratch rbh:mongo:scratch
• Uniform Resource Identifier
• RFC 3986:
  \[\text{scheme:][/[/userinfo@]host[:port]]path\]
rbh-sync: URIs

https://example.com/test
rbh-sync: URIs

rbh:mongo:scratch
rbh-sync: URIs

rbh:mongo:scratch

scheme

path
rbh-sync: URIs

rbh:mongo:scratch

scheme backend-type fsname
rbh-sync: URIs

rbh:mongo:scratch#/path/to/dir

scheme    backend-type    fsname
rbh-sync: URIs

```
rbh:mongo:scratch#/path/to/dir
```

- **Scheme**: rbh
- **Backend-type**: mongo
- **Fsname**: scratch
- **Fragment**: #/path/to/dir
rbh-sync: URIs

```
rbh:mongo:scratch#path/to/dir
```

```
rbh:mongo:scratch
```
rbh-sync rbh:posix:/mnt/scratch#dir-0 \ rbh:mongo:scratch
rbh-sync: another example
cd /mnt/scratch
for dir in *; do
    rbh-sync posix:/mnt/scratch#$dir" mongo:scratch &
done
wait
rbh-sync: another example

process-0  process-1  process-2
• 1 node:
  - 16 cores
• Lustre on RAM:
  - 1 MDT
  - 8 OSTs
  - Cold cache
• Filetree structure:
  - 4 levels of dirs
  - Branching factor of 16
  - 5th level: ~1M files of 1 KiB each
A clone of (gnu-)find that queries RobinHood
rbh-find: usage

rbh-find [URI …] [[predicates] [actions] …]
rbh-find rbh:mongo:scratch -name ‘a*’ -print0
• 1 node:
  - 16 cores
• Lustre on RAM:
  - 1 MDT
  - 8 OSTs
  - Cold cache
• Filetree structure:
  - 4 levels of dirs
  - Branching factor of 16
  - 5th level: ~1M files of 1KiB each
• 1 node:
  - 16 cores
• Lustre on RAM:
  - 1 MDT
  - 8 OSTs
  - Cold cache
• Filetree structure:
  - 4 levels of dirs
  - Branching factor of 16
  - 5th level: ~1M files of 1KiB each

![Find performance graph]

- v3-patched: 197.85s
- v4: 131.97s
- find: 10.55s
rbh-fsevent
Apply filesystem events on a RobinHood backend
rbh-fsevent: usage

TODO
Changelog processing
Before
Before

MDT → Lustre Client → RobinHood Client → MySQL DB
Before: overload
Before: imbalance
After
After: with kafka

Members of a Kafka cluster

Lustre Clients that run stat()

MDT

Lustre Client

RobinHood Clients

RobinHood Backend
After: with kafka
After: with kafka
Thank you