Hacking AFS Dumps for Fun and Profit

Thomas L. Kula

Information Technology Central Services

University of Michigan

2009 AFS and Kerberos Best Practices Workshop

For some reason, I am fascinated by dumps

```
vos dump -id user.kula.backup -localauth
```

What is in a volume dump

- Dump Header
- Volume Header
- Large and Small vnodes

Dump Header

- Volume ID
- Volume Name
- From Date
- To Date

Volume Header

- Volume ID
- Volume Name
- maxquota, diskused, nfile
- create/access/update/backup dates
- And other stuff

Vnodes

- Large vnodes are directories
- Small vnodes are files
- Vnode number
- Uniqifier
- Type

Vnodes

- Author/Owner/Group/Mode
- Client/Server Date
- Size
- ACL (for directories)

Vnode Contents

- The content of a small (file) vnode is the file itself
- The content of a large (directory) vnode is a blob that ties names to vnode/uniqifier tuples

Full Dumps

- Every vnode, and the contents of every vnode, is dumped

Partial Dumps: Directories

- Every large vnode and its contents are dumped
- Unless you use `-omitdirs`, in which case a minimal vnode is dumped
- It's basically enough info to say "this vnode and uniqifier still exists"

Partial Dumps: Files

- Every small vnode has something dumped:
  - If the file has changed since `-time`, the full vnode and contents are dumped
  - If the file has not changed since `-time`, only a minimal vnode is dumped — "This vnode and uniqifier still exists"

Some problems I'd like to solve

- What exactly is in a collection of dumps?
- Merging a full and partial dumps into a new full dump

What exactly is in a collection of dumps

- "I need exactly this file from this date" rarely happens
- "My file last existed sometime during this week...."

Accessing metadata

- Useful to have access to metadata in a collection of dumps
- 240K volumes backed up daily, 28 days retention
- Keeping everything in one database would require a lot of info, shoved in and pulled out daily
- We're not fans of "*The* database, with *everything*"
- 99.99% of that data will never be used

Accessing metadata

- And yet it would be nice to have relatively quick access to that data for users

Accessing metadata

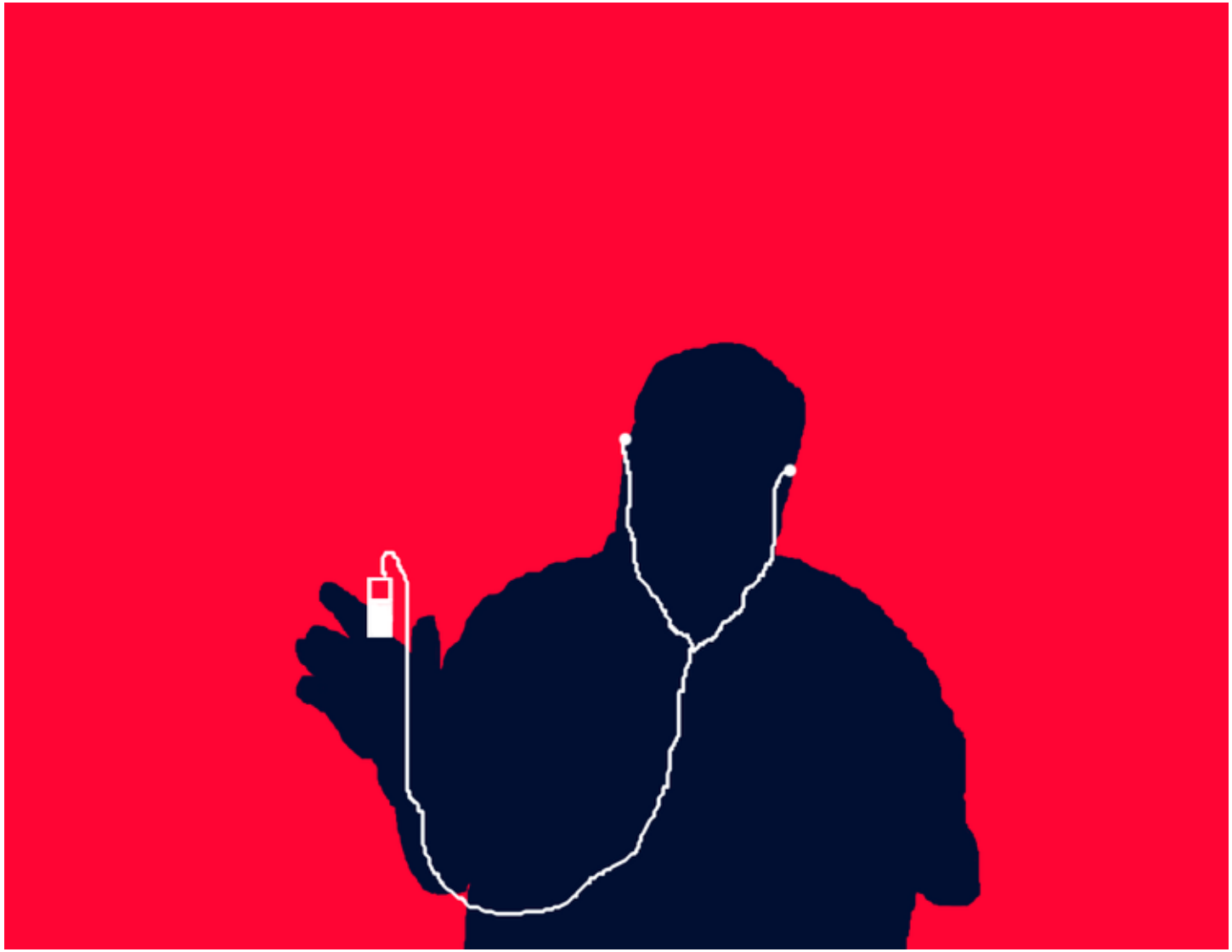
- We have volumes that are not being touched, and are probably not even mounted anywhere
- It would be nice to identify those, and get rid of them
- `find /afs/umich.edu ...` seems painful

Accessing metadata — wants

- Pre-extracted metadata
- Per dump file granularity
- Easy-to-read blob

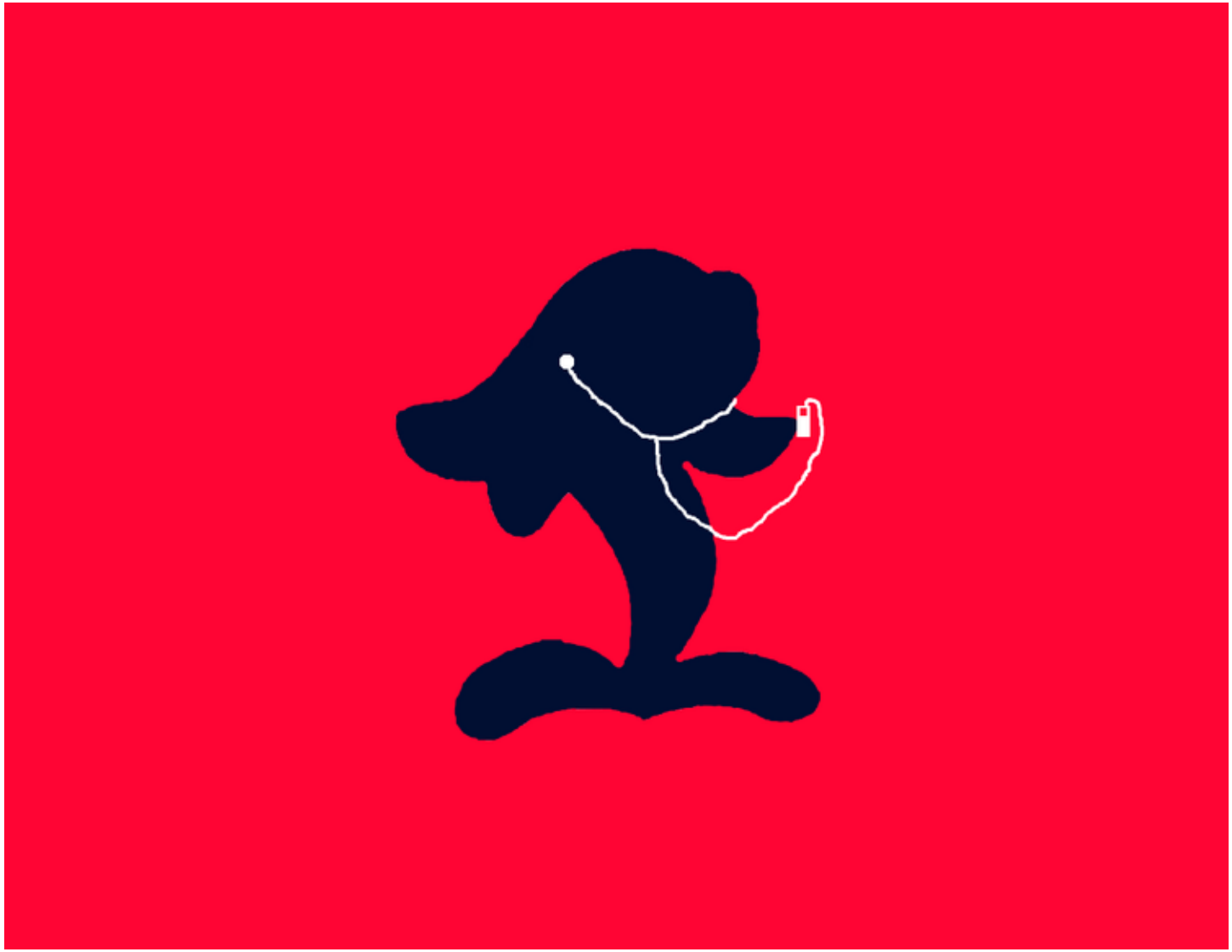
Combining full and incremental dumps

- Like many people, I have a lot of mp3s



Combining full and incremental dumps

- Like any sane person, I keep them in AFS



Combining full and incremental dumps

- Like any prudent person, I keep backups

Combining full and incremental dumps

- Path from colocation to home: 3 mbps
- Path from home to colocation: 768 kbps

Combining full and incremental dumps

- Doing full dump of most large volumes is painful

- ```
  am.tmbg.backup                    536871739 BK    1297219 K  On-line
      service-m1.tproa.net /vicepb
      RWrite  536871738 ROnly          0 Backup  536871739
      MaxQuota    5000000 K
      Creation    Sun May 31 01:59:07 2009
      Copy        Sun May 31 01:59:07 2009
      Backup      Sun May 31 01:59:07 2009
      Last Update Sat Jan  3 11:55:14 2009
      0 accesses in the past day (i.e., vnode references)

      RWrite: 536871738     Backup: 536871739
      number of sites -> 1
         server service-m1.tproa.net partition /vicepb RW Site
  ```
  ```

Working with dump files

- dumpscan, a tool for working with dumps
- `http://dl.central.org/dl/software/dumpscan/`
- Written by CMU SCS

dumpscan

- Simplifies the mechanics of slogging through a dump
- Register callbacks to handle various dump items
- Then run a routine to scan through the dump

Changes to dumpscan

- `http://kula.tproa.net/code/dumpscan-dont-call-cb_dirent-twice.patch`
- `http://kula.tproa.net/code/xfile-gzip.patch`

My blob or yours?

- Don't want to invent my own
- Want data to be self-contained, easy to read, standardized

sqlite

- `http://sqlite.org/`
- *SQLite is a software library that implements a self-contained, serverless, zero-configuration, transactional SQL database engine.*
- APIs for C and Python
- Simple sql client

afsdump_sqlite

- Pulls metadata from dump
- `afs_dump_header`
- `afs_vol_header`
- `afs_vnode`
- `afs_dir_ent`

afsdump_sqlite

- Embarrassingly simple use of sqlite
- Simple tables of data
- Tables designed to be concatenated
- "Trivial" to use your DB of choice

afsdump_sqlite

- Hackish
- Not particularly robust
- It works until it doesn't

afsdump_sqlite.py

- Rough framework of Python
- Deals with dumps and collections of dumps
- Even more hackish
- Highly incomplete

lslr.py

- Proof-of-concept
- `ls -R` on a collection of dumps
- You don't even want to know how hackish this is

```
./lslr.py /home/kula/u.kula.backup.1163749800-0-0.sqlite \
    /home/kula/u.kula.backup.1235778854-1163749800-1.sqlite \
    /home/kula/u.kula.backup.1240557814-1235778854-2.sqlite
```

```
from dump  u.kula.backup.1240557814-1235778854-2  directory ( 1 , 1 )  .
DIR ( 1 , 1 ) from dump u.kula.backup.1240557814-1235778854-2 :  u'.'
DIR ( 1 , 1 ) from dump u.kula.backup.1240557814-1235778854-2 :  u'..'
...
FILE ( 378 , 2384 ) from dump u.kula.backup.1235778854-1163749800-1 :  u'.muttrc'
FILE ( 388 , 1388 ) from dump u.kula.backup.1163749800-0-0 :  u'.profile'
...
SYMLINK ( 540 , 2866 ) from dump u.kula.backup.1163749800-0-0 :
 u'iastate-stuff'  links to u'#iastate.edu:user.kula.'
...
from dump  u.kula.backup.1240557814-1235778854-2  directory ( 9 , 1363 )  ./.ssh
DIR ( 9 , 1363 ) from dump u.kula.backup.1240557814-1235778854-2 :  u'.'
DIR ( 1 , 1 ) from dump u.kula.backup.1240557814-1235778854-2 :  u'..'
FILE ( 2740 , 7555 ) from dump u.kula.backup.1235778854-1163749800-1 :  u'config'
```

Conceptualized Merge Tool

- Run `afsdump_sqlite` on dump(s)
- Run `create_dump_manifest`
- Run `merge_dumps`: dumps + manifest = new dump

Other potentially useful tools

- volume dump merge equivalent of “vos split”
- Puffs/FUSE, allow a collection of dumps to be mounted RO as a local disk?

Code

- Canonical GIT repository

- `/afs/tproa.net/public/code/afsdump_sqlite/afsdump_sqlite.git/`

Thanks Ugly Mug Cafe

```
2009-05-31 0.65 Coffee refill, Tanzanian
2009-05-31 1.59 Coffee, Tanzanian
2009-05-30 0.65 Coffee refill, Brazil
2009-05-30 0.65 Coffee refill, Brazil
2009-05-30 1.59 Coffee, Brazil
2009-05-29 0.65 Coffee refill, Ethiopian
2009-05-29 0.65 Coffee refill, Ethiopian
2009-05-29 1.59 Coffee, Ethiopian
2009-05-29 3.02 Trad. cap.
2009-05-27 2.17 Rooibos
2009-05-26 2.17 Rooibos, iced
2009-05-25 2.17 Rooibos, iced
2009-05-25 1.59 Coffee, Burundi
2009-05-24 1.59 Coffee, Brazil
2009-05-23 2.17 Rooibos, iced
2009-05-23 0.65 Coffee refill, Burundi
2009-05-23 1.59 Coffee, Burundi
2009-05-23 2.12 Espresso
```

Hacking AFS Dumps for Fun and Profit

Thomas L. Kula

Information Technology Central Services

University of Michigan

kula@tproa.net — tkula@umich.edu

2009 AFS and Kerberos Best Practices Workshop

http://kula.tproa.net/talks/afskbpw2009/

Media Credits

- "Waste dump Tanjung Priok Jakarta Indonesia"
  http://commons.wikimedia.org/wiki/File:Waste_dump_-_Jakarta_-_Indonesia.jpg

  by http://commons.wikimedia.org/wiki/User:Hullie
  Licensed under Creative Commons Attribution ShareAlike2.5
  http://creativecommons.org/licenses/by-sa/2.5/