

AFS Cell Management Tools and Techniques

Russ Allbery
June 13, 2005

Introduction

- Stanford has 3.9TB of data in AFS, in 57,485 volumes (as of June 1st). (1.6TB user home directories, 660GB data, 180GB groups and departments, 550GB classes).
- Administration when no migrations are in progress takes a few hours a week, mostly creating unusual volumes, moving volumes around, and upgrading servers.
- Tools presented here developed by Neil Crellin.
- <http://www.eyrie.org/~eagle/software/>
- <http://www.eyrie.org/~eagle/notes/afs/>

Contents

- Volume creation and management
- Managing ACLs
- Analysis and reporting
- Replicated volumes
- Monitoring with Nagios

Creating Volumes

- `volcreate` wrapper to balance where volumes are placed
- Mapping volume types to servers
- Size policy (2-4GB max for ease of moving volumes)
- Automated log volume creation with `volcreate-logs`
- Wrapper scripts for volume types (`create-user`, etc.)

Managing Volumes

- `partinfo` wrapper for usage information
- `mvt` utility for all volume moving
- Generating volume lists with `vos listvol`
- Checking for unreleased volumes with `unreleased`
- Balacing: why or why not, and possible overkill solutions
- `volnuke` wrapper to delete volumes
- Delegated volume creation ability (`remctl` and `afs-backend`)

Managing ACLs

- One PTS group per course, department, or group volume
- Help desk tools to change PTS group membership (and volume quota)
- `fsr` wrapper for users
- Be careful of IP-based ACLs: subnets work best, better to use `kstart` and machine `srvtabs` over IP ACLs
- Log volume ACLS (`lik`) and the potential problems
- Think about `fs cleanacl`
- Unix directory owners and their special ACLs

Tracking Volumes

- Hierarchical naming scheme for volumes
- Mount point database (`mtpt`, `loadmtpt`, `cleanmtpts`)
- Nightly load into an Oracle database
- Nightly reports from the Oracle database (released volumes, high accesses, volumes moved, unreleased changes, missing mount points)
- Monthly usage reports

Replicated Volumes

- Replication helps when server is down, not when it's slow
- How many replicas do you want? (2-4)
- `volcreate` and server geographic locations
- How RW and RO paths work: replicate the whole path
- Delegated volume release ability (`remctl` and `afs-backend`)
- `frak` to find changes
- Restoring a RW from a RO with `vos dump` and `vos restore`

Monitoring with Nagios

- Basic tool: `bos status`
- Monitor VLDB servers with `udebug`:
`pt 7002, vl 7003, ka 7004`
- Available disk with `vos partinfo`
- Connections waiting for thread (`rxdebug`)
- AFS logs and `kill -TSTP`
- Nightly problem reports from Oracle database