Athena Under The Hood

Marc Horowitz

Student Information Processing Board

“Clue Dump” Series
Fall 2008
Overview

- History
- Network Services
- Workstation Services
Where did Athena come from?

• Project Athena was a five year project started in 1983.
• The goal was to make a distributed, scalable workstation environment based on 4.3 BSD Unix.
• Nothing like this had ever been attempted before.
• Timesharing was the ubiquitous model.
What did Athena invent?

- Kerberos network authentication
- Location-independent user accounts
- Location-independent instant messaging
- Unattended network installation
- Automatic system updates
- The X Window System
What did Athena look like then?

- VAX 11/750 servers
- DEC VAXstation & IBM RT workstations
  - 1 MIPS CPU
  - 4 megabytes of memory
  - 1 megapixel monochrome display
  - 40mb local hard drive
- Services very similar to today
  - Kerberos, Hesiod, Lockers, Moira, Update
How has Athena changed?

- AFS has replaced NFS and RVD
- Kerberos 5 has replaced Kerberos 4
- Solaris and Linux have replaced 4.3 BSD
  - Athena no longer builds the OS from scratch
- Machines are much faster
- Quotas are no longer 600 kilobytes
- The Internet has grown past 60,000 hosts
Network Services

- Kerberos
- AFS
- Hesiod
- Lockers
- Email

- Printing
- Moira
- Zephyr
- Lert
- Larvnet
Workstation Services

- Install
- mkserv
- Update
- Workstation self-maintenance
- User logins
- Display Manager
Kerberos

- Secure Network Authentication Service
  - KDC
  - Users and passwords
  - Services and keytabs
AFS

AFS is the network filesystem used on Athena

- Secured using Kerberos
  - tokens
  - PAGs
- Organized into “cells” and “volumes”
- Volumes can be replicated and moved
- Performance is improved by caching
- Distinctive permissions model
AFS: Pitfalls

- Permissions behavior is unexpected by experienced Unix users
- No hard links across directories
- Software can behave badly when tokens expire
- close() can fail
- Non-cached performance is slower than local access
Hesiod

Hesiod is a distributed naming service based on DNS

- Many types of data
  - User account data
  - Post office data
  - Workstation cluster data
  - Printer data
  - Many more
- Uses standard BIND DNS implementation
  - distributed
  - replicated
Hesiod: examples (1)

- **All at once**
  
  ```
  % athrun consult hes sipbtest
  PASSWD: sipbtest:*:20922:101:Fred Sipb,,,:/mit/sipbtest:/bin/athena/tcsh
  FILSYS: AFS /afs/athena.mit.edu/user/s/i/sipbtest w /mit/sipbtest
  POBOX: POP PO12.MIT.EDU sipbtest
  GRPLIST: sipbtest:7329
  GROUP: sipbtest:*:7329:
  ```

- **One record at a time (good for scripts)**
  
  ```
  % hes sipb filsys
  AFS /afs/sipb.mit.edu/project/sipb n /mit/sipb 1
  AFS /afs/athena.mit.edu/contrib/sipb n /mit/sipb 2
  ```
Hesiod: examples (2)

- Using DNS tools directly

```bash
% host -t txt -v ceres.pcap.ns.athena.mit.edu
Trying "ceres.pcap.ns.athena.mit.edu"
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 1142
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 3, ADDITIONAL: 0

;; QUESTION SECTION:
;ceres.pcap.ns.athena.mit.edu. IN TXT

;; ANSWER SECTION:
ceres.pcap.ns.athena.mit.edu. 7139 IN TXT
"ceres:rp=ceres:rm=HUSQVARNA.MIT.EDU:ka#1:mc#0:auth=kerberos5:xn:"

;; AUTHORITY SECTION:
ns.athena.mit.edu. 20742 IN NS CLIO.mit.edu.
nameserver = ns.athena.mit.edu.
nameserver = ns.athena.mit.edu.
nameserver = ns.athena.mit.edu.

Received 183 bytes from 127.0.0.1#53 in 71 ms
```
Lockers

Filesystem and location independent access to file storage

- Used mostly by attach and add tools
- Simple case: a locker is a location in /afs, with symlink from /mit
- Also provides mechanisms for failover and management of PATH and MANPATH environment variables
- lockers(7) man page documents conventions for software locker layout
Email

- Sendmail-based infrastructure
- Mail from workstations may be queued locally before being forwarded to “outgoing” servers
- Outgoing servers deliver mail to MIT mailhubs and third party recipients
- Mailhubs perform spam checks, expand mailing lists, forward to PO servers
- PO servers store email, which is accessed with KPOP or IMAP
Email

- **External Sender**
- **Spam Appliance**
  - **User Agent**
  - **User Agent**
  - **User Agent**
- **outgoing**
- **milter**
- **Mailhub**
- **External Recipient**

Inside MIT

Post Office
Printing

- LPRng-based infrastructure
- Hesiod to locate printers
- Kerberos for authentication (deprecated)
- Print servers
  - access control
  - queue management
  - forward jobs to printers
Moira

Primary repository and administration service for all system and user data

- Kerberos or certificate authenticated
- Rich client set
  - moira, blanche, stanley, listmaint
  - web-listmaint <http://web.mit.edu/moira>
  - mrtest
- DCM (Data Control Manager) Model
  - Applications do not query Moira directly
  - Data pushed periodically to services (mail aliases, hesiod data, etc)
  - “Incremental” services updated dynamically (afs quotas, afs group membership)
Moira: What it manages

- Workstation cluster information
- Locker hesiod and quota data
- Mailing lists and filesystem groups
- Host and network configuration
- Printer configuration
- User account data
- Access control lists: zephyr, discuss, etc.
- Almost everything else (except passwords)
Zephyr

Zephyr is a location-independent instant messaging system

- Secured using Kerberos
  - One of the last parts of Athena which still requires version 4
- System status notification
  - This was its original purpose
- Location management
- One-to-one user communications
- Many-to-Many user communications
Larvnet

- Gathers data used by cview
  - User login information
    - Sent proactively by login/logout process
    - Polled via “busyd” inetd udp service
  - Print queue information
    - Polled using lpq
Workstation Services

- Install
- mkserv
- Update
- Workstation self-maintenance
- User logins
- Display Manager
Install

Athena workstation installation is a fully automated process

- Installs are started or network boot or cd/floppy
- An install kernel and filesystem is mounted over the network
- The local disk is formatted
- Files are copied onto the disk
- On non-Linux platforms, much of the OS is executed from AFS and not local at all
mkserv

mkserv is a system for customizing a newly installed workstation

- server installations
- common private workstation customizations
- local private customization script
- executed at install time and update time
Workstation Self-maintenance

Regular tasks are performed by each client to keep them current and healthy.

- Frequent execution
  - Boot time
  - Every few minutes when nobody is logged in ("reactivate")
  - From cron
- System software is verified
- Kicks off workstation update if necessary
Boot and reactivate tasks

- passwd, shadow, group files are reset to known good copies
- AFS CellServDB is updated from AFS
- Lockers are detached
- Hesiod cluster information is retrieved
  - System software location
  - Default printer
- srvd attached (not on Linux)
- Verification of OS software (boot only)
- Verification of Athena software (only at boot on Linux)
- System configuration files are reset to known good copies from AFS
- Check for software update
- Time synchronization
Cron tasks

- Temporary directories are cleaned
  - This can happen even if a user is logged in
- Queued mail is pushed
- System time is reset if it has drifted more than 60 seconds
- Locker software is copied locally from AFS to improve performance
  - Acrobat reader
  - OpenOffice
  - Private workstations can extend this list
**Update**

Athena workstations are automatically updated remotely

- A flag file on a file server indicates if an update is available
- Most services are shut down
- The OS is updated
- Configuration files and Athena software are copied locally
- mkserv is run
Update (srvd)

- Hesiod clusterinfo adds new RVD (new release)
- `/srvd/.rvdinfo` compared to `/etc/athena/version` (patch)
- Checks performed to see if the update should occur (auto-update, disk space, desynchronization, etc)
- Services shut down
- Update scripts executed
- Configuration files updated from srvd
- `rc.conf` updated
- Miniroot may be created if major OS updates are necessary
- OS software updated
  - `pkgadd` and `patchadd` on Solaris
- Athena software updated with track
- Reboot if OS updated
- mkserv runs
Update (Linux)

- Hesiod clusterinfo adds new syscontrol file (new release)
- Control file compared to /etc/athena/version
- Checks performed to see if the update should occur (auto-update, disk space, etc)
- Services shut down
- RPMs are added, updated, and removed
- mkserv is run
**User logins**

Athena's user login process is different from most Unix environments.

- **User authentication**
  - Local password
  - Kerberos password
  - Kerberos-authenticated remote login
    - To avoid confusing users, non-root authenticated logins without forwarded credentials are generally not permitted

- passwd, shadow, group files updated from Hesiod
- User's locker is attached
Display Manager

- Run out of inittab
- Manages console lifecycle
  - X server
    - Restarted on logout or death
  - login window
    - Run on each new X server
  - console window
    - Run on each new X server
  - user X session
    - started if login succeeds
    - cleaned up afterwards
  - reactivate periodically when no user is logged in
Acknowledgements

- Greg Hudson, whose document inspired the organization and content of this talk
- Yang Zhang and previous organizers for organizing the series and trusting me to give the first impression
- Mitch Berger, John Hawkinson, and Jonathon Weiss, for commenting on the slides