Evolution of a Shared Web Host
A Mostly Inaccurate History of Scripts.mit.edu

Edward Z. Yang
<ezyang@mit.edu>

SIPB Cluedumps — October 25, 2011
The server in your dorm room...

INTERNET → Apache
   MySQL
      nobody
     nobody
       php
       python

Single user, scripts run as "nobody"
Share it with your friends...

INTERNET

Apache+suEXEC
MySQL

httpd
ezyang
geofft
apache2
php
python

multi-user, suEXEC runs scripts as their user
↑ setuid
Grow a little...

MySQL sends passwords in clear, so use internal network
Part I

Athena Integration
Apache+suEXEC has some assumptions...

Local ext3 filesystem
Password hashes in /etc/shadow

Scripts integrates w/ Athena and uses...

Remote AFS filesystem
Login with Kerberos authentication
Scripts

files

authentication

athena.mit.edu AFS cell

aether

castor

pollux

need KRB tickets to get AFS tokens

Kerberos

Kerberos-1

Kerberos-2

ATHENA.MIT.EDU Kerberos Realm
How does Kerberos+AFS work?

User

password

Athena cluster machine

auth with password

Kerberos ticket

Kerberos ticket

AFS ticket

AFS token

File

Kinit

aklog

open

Disclaimer: This is simplified!
A subtlety...

Tickets never go on server unless you explicitly say so (GSSAPIDelegateCredentials)
Kerberos keytabs are the keys to the kingdom

```bash
$kinit -k -t keytab
(no password prompt)
$aklog
$find ~
Private/...
  zlog/...
  :
```

Necessary for server (don’t save password!)

Too much access for Scripts!
Solution: daemon.scripts

- Scripts
  - /etc/daemon.keytab
  - get credentials in afsagent cronjob
- Kerberos
- AFS
  - get files w/ daemon.scripts AFS tokens

fs la ~/web_scripts → daemon.scripts rlidwkr
Security problem: multiple users

```
fs la /mit/user1/web_scripts → daemon.scripts rlidwk
fs la /mit/user2/web_scripts → daemon.scripts rlidwk
```

```
user1> cat /mit/user2/web_scripts/config.php
```

![Diagram]

- User1
- User2
- Scripts
- daemon.scripts
- AFS
- User1: daemon.scripts
- User2: daemon.scripts
- Local user
- AFS/kerberos user
Solution: Patch AFS  

- Enforce AFS permissions* with local user info
- Use daemon scripts when actually talking to AFS

```c
if ( areq->uid == globalpag &&
    !(areq->realuid == avc->fid.Fid.Volume &&
     !((avc->anyAccess | arights) == avc->anyAccess)) ) {
    return 0; // not authorized
}
```

* don't actually enforce AFS permissions; approximate based on Volume, e.g. what's on /mit/ezyang
Extra considerations

- Apache stat()s files before suEXECing
  
  accept if (arights == PRSFS_LOOKUP &&
  areq->realuid == HTTPD_UID)

- Postfix checks if procmailrc exists before su'ing
  
  accept if (arights == PRSFS_LOOKUP &&
  areq->realuid == POSTFIX_UID)
Problem: Apache doesn't suEXEC static files

HTTPD user, has no permissions

KERNEL
Solution: Make it suEXEC static files

- Apache
- suEXEC
- SETUID
- static-cat
- program

SCRIPTS user

AFS

KERNEL

HTTPD user

is executable (or ends in .php)

written in Haskell, supports ranges
One more thing... admof

$ fs la /mit/scripts

System: scripts-root rlidwka
system: anyuser

← No "scripts" user

Patch Kerberos to check if a user has "a" rights on a locker.

ezyang/root> ssh scripts@scripts.mit.edu
scripts>
Here we stand. Questions?

- Cluster machine
- Public HTTP
- SSH with Kerberos
- Scripts
- Kerberos
- AFS
- SQL
- Database connection
- File retrieval
- File editing

(from now, assume this implicitly)
~ Part II ~

High Availability and Scaling
Scripts has a lot of users
Single point of failure
Out-of-the-box database replication

WEB

scripts

k-s

S-b

Master

Slave

(if you're curious, talk to sql@mit.edu)
Replicate Scripts

Some (but not all) work already done!
What is done

- User filesystem (AFS)
- User authentication (Kerberos)

What's not

- Local filesystem (Manual)
- Load-balancing & Failover (LVS)
- User information (LDAP)
- Apache configuration (mod-vhost-ldap)
Load-balancing and failover

Running `ldirectord` to monitor Scripts servers

WEB --- LVS --- Scripts1

WEB --- Scripts2

`finger @scripts`

Incoming packets go through LVS; outgoing doesn't
More redundancy

LVS1 bound to Scripts IP 18.181.0.43

WEB

Router

LVS1

LVS2

Scripts1

Scripts2

Running heartbeat to monitor LVS servers
LVS1 goes down
Heartbeat notices LVS1 is down, starts ARPing for Scripts1 IP.
Back online (<1 sec later)

WEB

Router

18.181.0.43

LVS2

Scripts1

We bind FIVE IP addresses total

On LVS1/LVS2, you can observe with cron_mon
Here we stand.

primary, runs Cron and serves ssh://scripts/

Internal Network
User information (not authentication)

$ getent passwd root
root:x:0:0:root:/root:/bin/bash

$ getent passwd ezyang
ezyang:*:537864399:71944:ezyang:/afs/athena.mit.edu/user/e/z/ezyang:/usr/local/bin/mbash

↑ Automatically created on signup: how to replicate?
*nsswitch.conf:* pluggable `/etc/passwd` (and others!)

- `getent` with `libc`
  - (does caching)
  - `nsswitch.conf`
  - `nsclod`
  - `nss_ldapd`
  - `nss_files`

Local account (e.g., root)
Remote account (e.g., easyang)

Directory service (`dirsrv`)
LDAP!
LDAP: Hierarchical object-oriented DB...

$ ldapvi -b dc=scripts,dc=mit,dc=edu uid=ezyang

0 uid=ezyang,ou=People,dc=scripts,dc=mit,dc=edu
loginShell: /usr/local/bin/mbash
homeDirectory: /afs/athena.mit.edu/user/e/z/ezyang
gidNumber: 71944
uidNumber: 537864399
uid: ezyang
cn: ezyang
objectClass: posixAccount
objectClass: top

Athena UID
my volume ID
(for the AFS patch)

the “Object” class

Enterprise, no?
...with Multi-Master Replication!
(all servers accept LDAP writes: no failover necessary)

GSSAPI auth

I didn’t pick that name...

cn="dc=scripts,dc=mit,dc=edu",cn=mapping\ tree, cn=config
Apache configuration

```xml
<VirtualHost ezyang.script.scripts.mit.edu>

</VirtualHost>

<VirtualHost geofft.script.scripts.mit.edu>

</VirtualHost>

</VirtualHost>

: Need to create these on signup!
Put it in LDAP...

```bash
$ ldapvi -b ou=VirtualHosts,dc=scripts,dc=mit,dc=edu \  
apacheServerName=ezyang_scripts.mit.edu
```

- apacheSuexecGid: 71944
- apacheSuexecUid: 537864399
- apacheDocumentRoot: /afs/athena.mit.edu/user/e/z/ezyang/web_scripts
- apacheServerAlias: ezyang_scripts
- apacheServerName: ezyang_scripts.mit.edu
- objectClass: apacheConfig
- objectClass: top

or use vhostedit or vhostadd!

Users can edit this with Scripts Pony
...tell Apache about it (mod_vhost_ldap)

Default

vhost_ldap.conf

LDAP

Special cases

vhosts.d

reify-vhost.py

Traditional 
<VirtualHost>
Open Problems

- mod_vhost_ldap is **inflexible**
- Doesn’t interoperate with SVN, Git & Postfix vhosts

```bash
ldapvi -b ou=VirtualHosts,dc=scripts,dc=mit,dc=edu
  scriptsVhostName=ezyang.scripts.mit.edu
```

```ini
scriptsVhostDirectory:
scriptsVhostAccount: uid=ezyang,ou=People,dc=scripts,dc=mit,dc=edu
scriptsVhostAlias: ezyang.scripts
scriptsVhostName: ezyang.scripts.mit.edu
objectClass: scriptsVhost
objectClass: top
```

```
VhostLDAPConfig  DocumentRoot
  %homeDirectory%/web_scripts/%(scriptsVhostDirectory)
... we (still) have an incomplete patch.
```
Scripts!

LVS

BUILD SERVER

TEST SERVER

Scripts-f13-test

F13 SCRIPTS

S-a

O-f

W-e

P-b

B-m

B-b

B-k

C-w

G-e

R-m

J-l

J-p

A-h

SCRIPTS HOSTS

SQL

K-s

S-b

MONITORING

Sipb-noc
Questions?