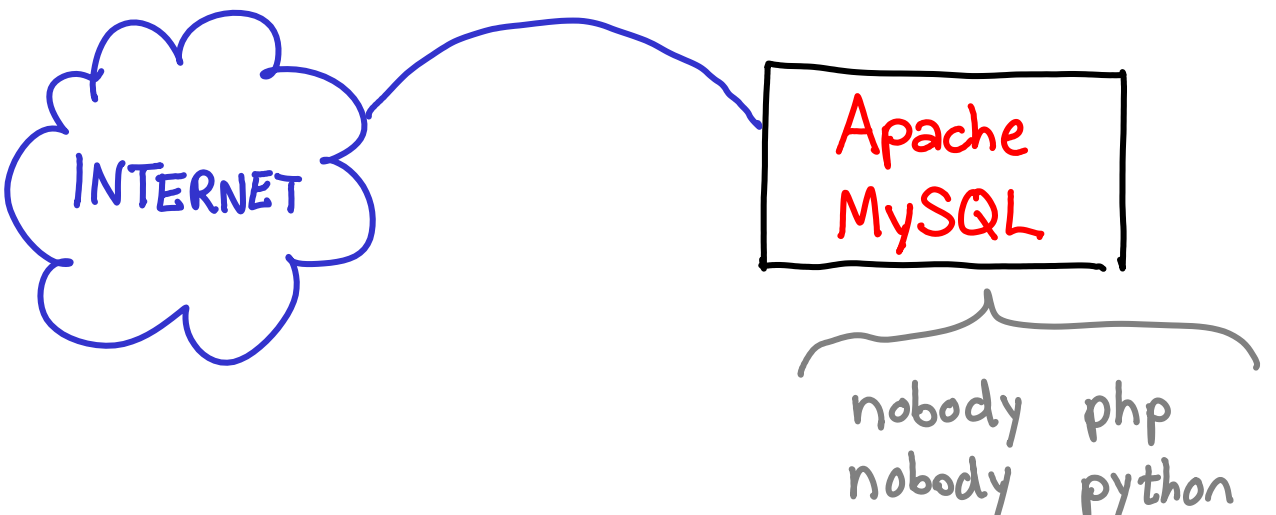
Evolution of a Shared Web Host

A Mostly Inaccurate History of Scripts.mit.edu

Edward Z. Yang
<ezyang@mit.edu>

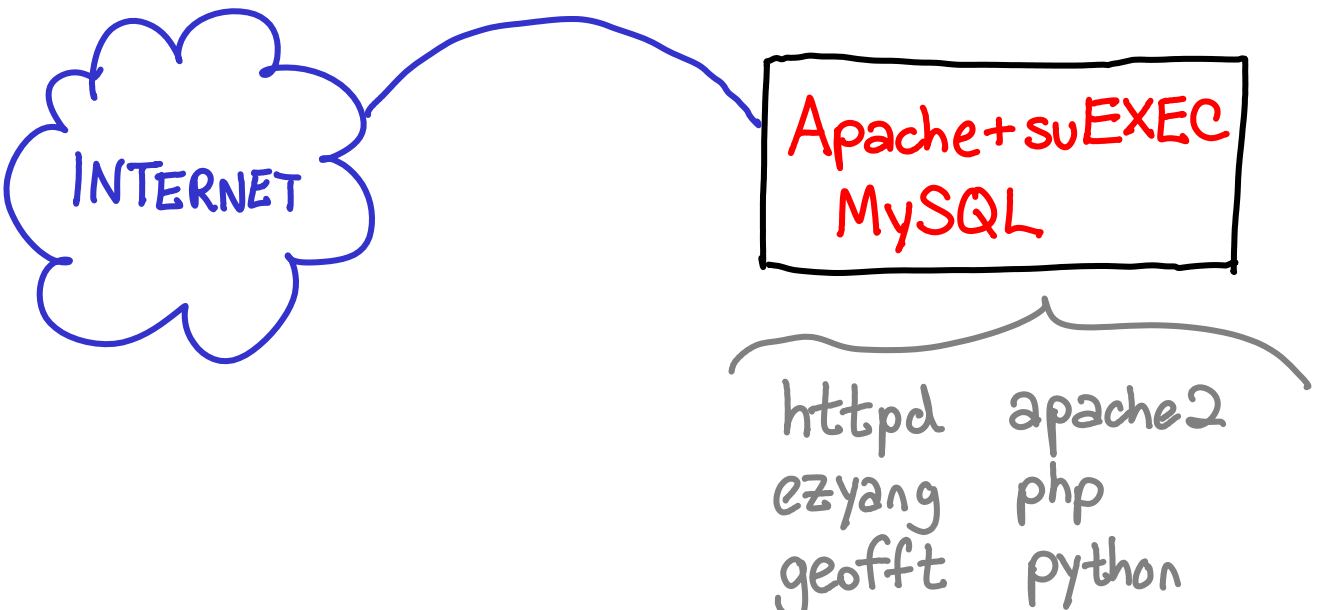
SIPB Cluedumps — October 25, 2011

The server in your dorm room...



single user, scripts run as "nobody"

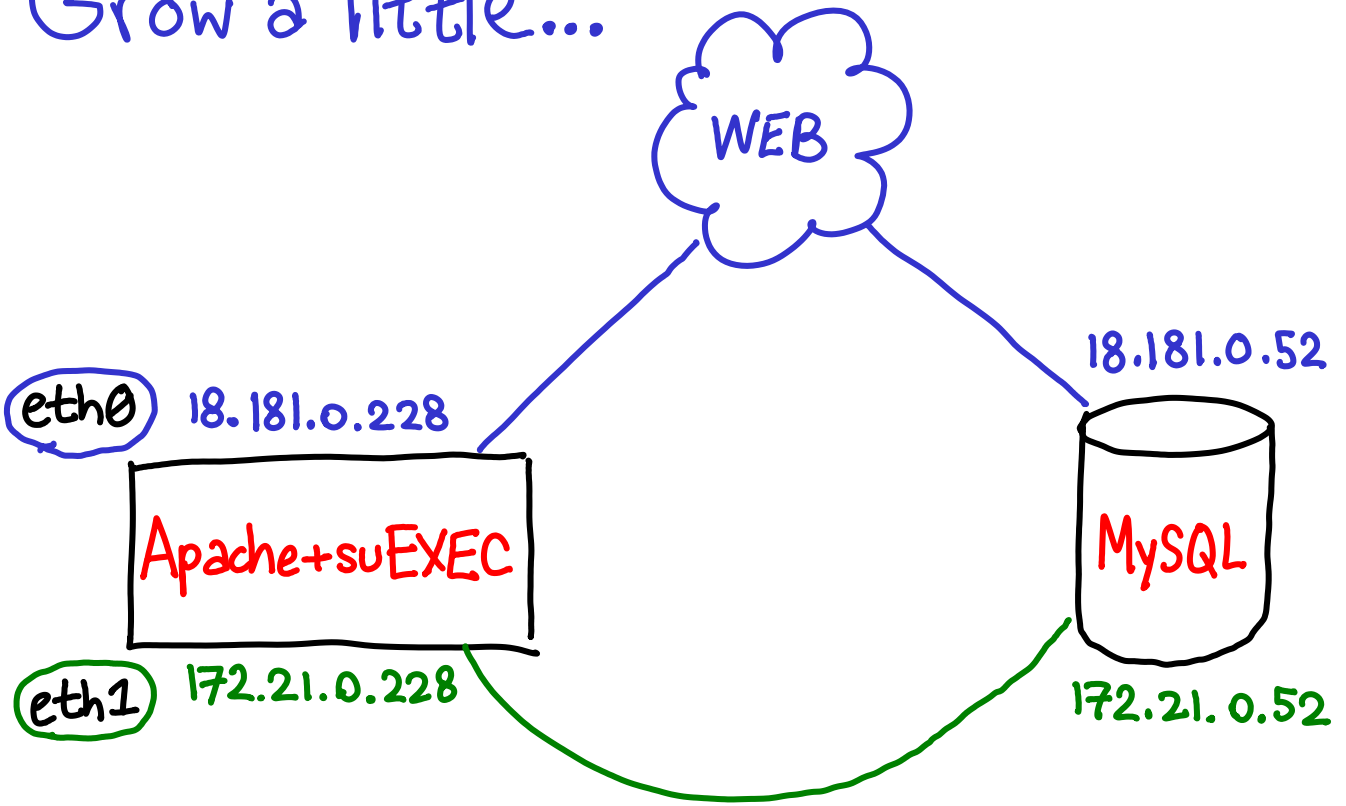
Share it with your friends...



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...

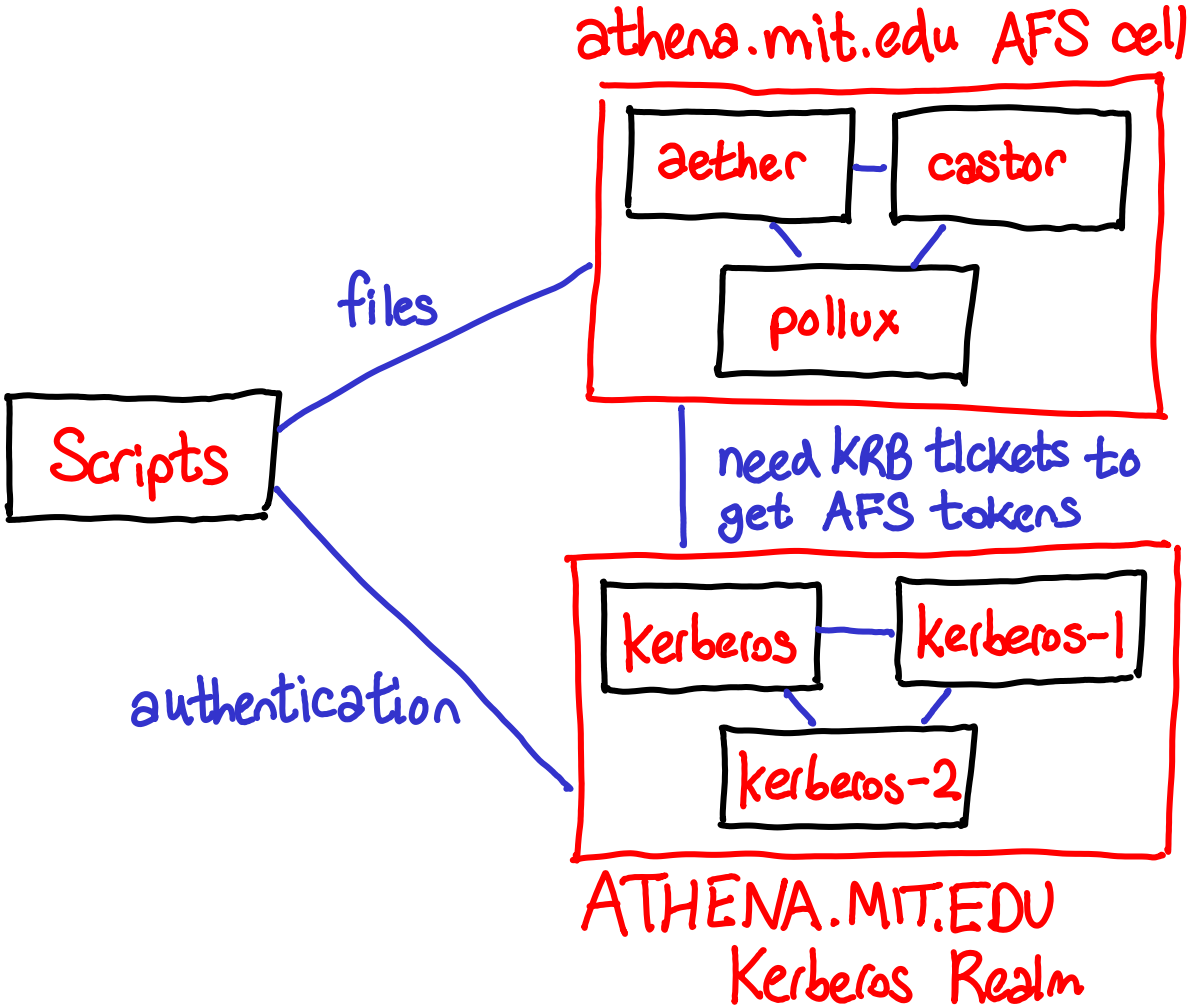
Local ext3 filesystem

Password hashes in /etc/shadow

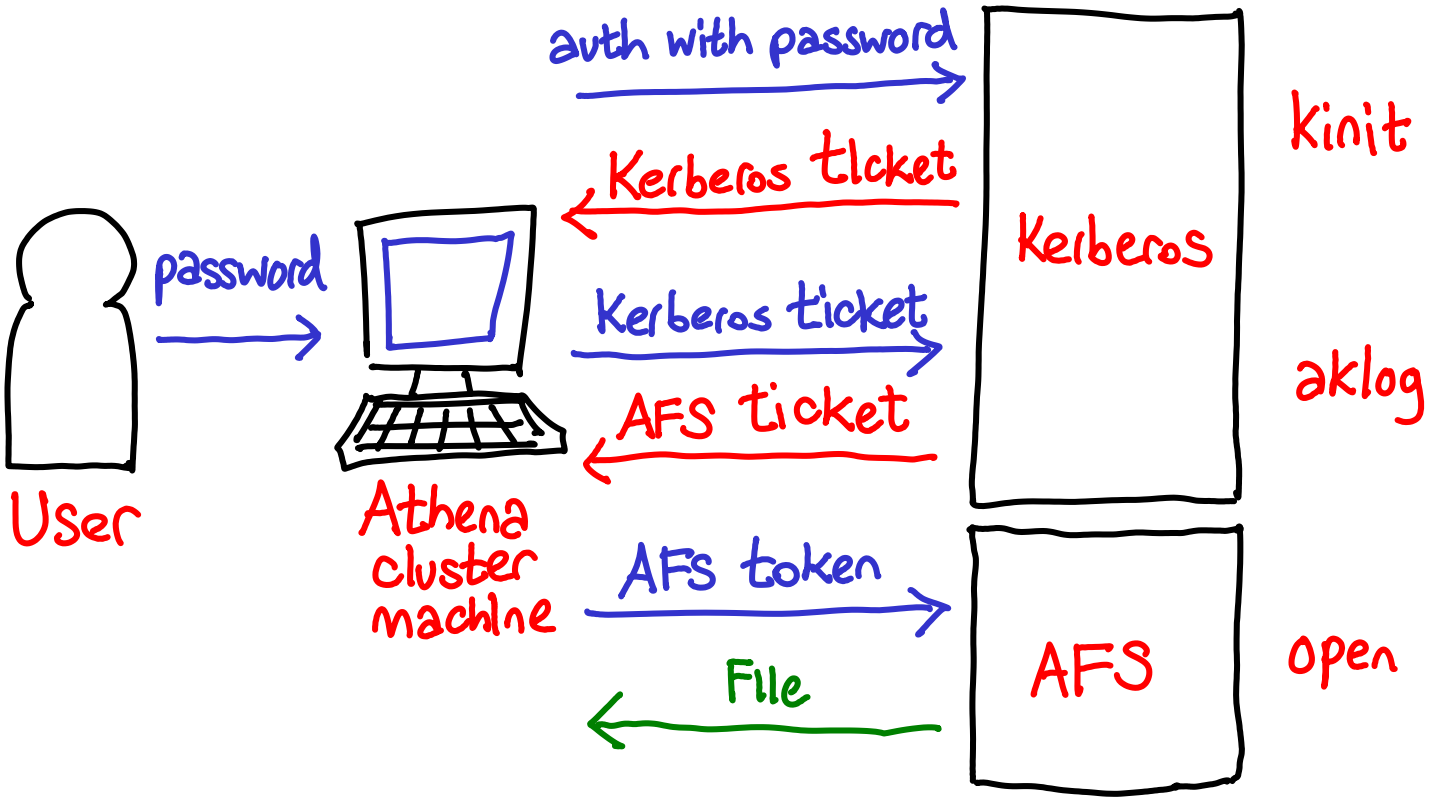
Scripts integrates w/ Athena and uses...

Remote AFS filesystem

Login with Kerberos authentication

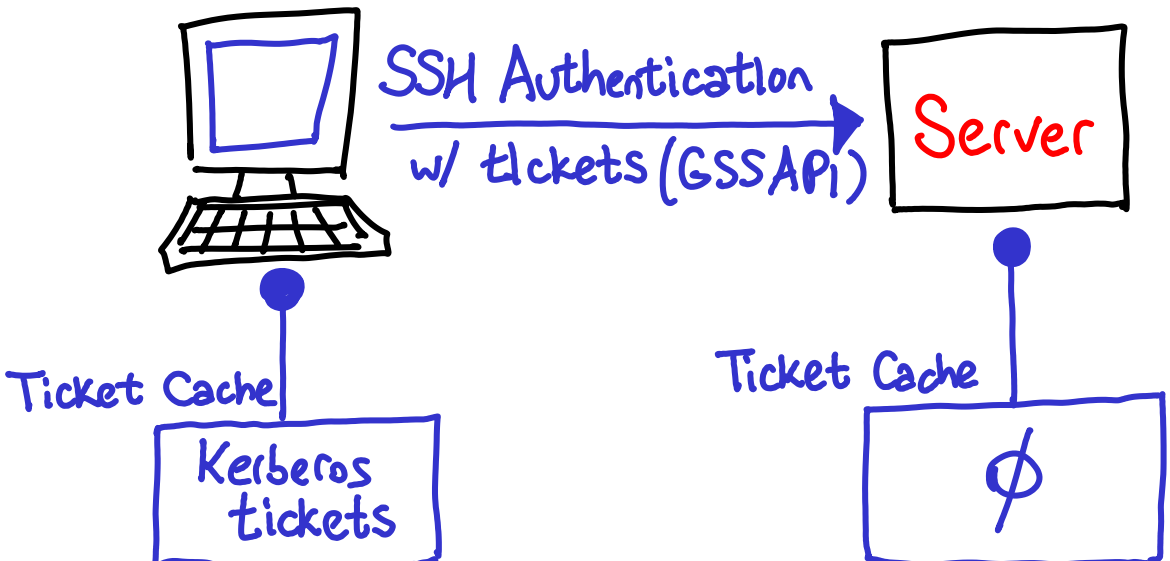

athena.mit.edu AFS cell
aether
castor
pollux
files
Scripts
need KRB tickets to
get AFS tokens
authentication
Kerberos
kerberos-1
kerberos-2
ATHENA.MIT.EDU
Kerberos Realm

How does Kerberos+AFS work?



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

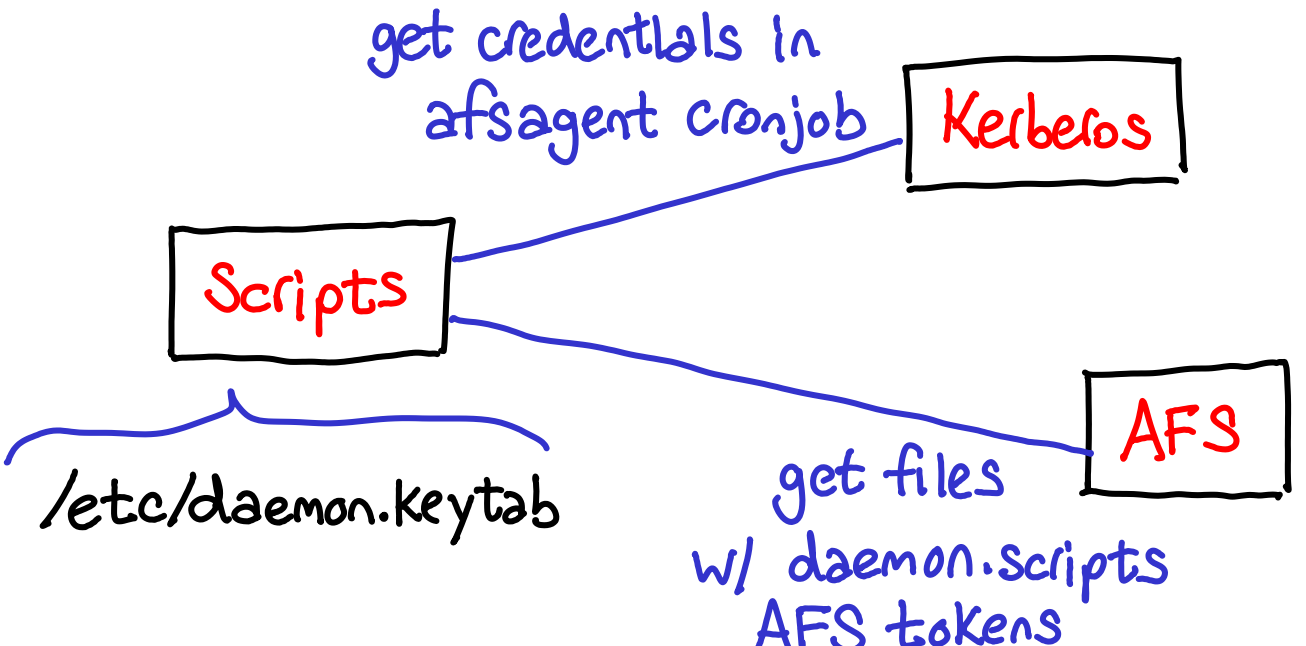
```
$ kinit -k -t keytab
```

(no password prompt) — Necessary for <u>server</u> (don't save password!)

```
$ aklog
$ find ~
Private/...
zlog/...
  ⋮
```

} Too much access for Scripts!

Solution: daemon.scripts



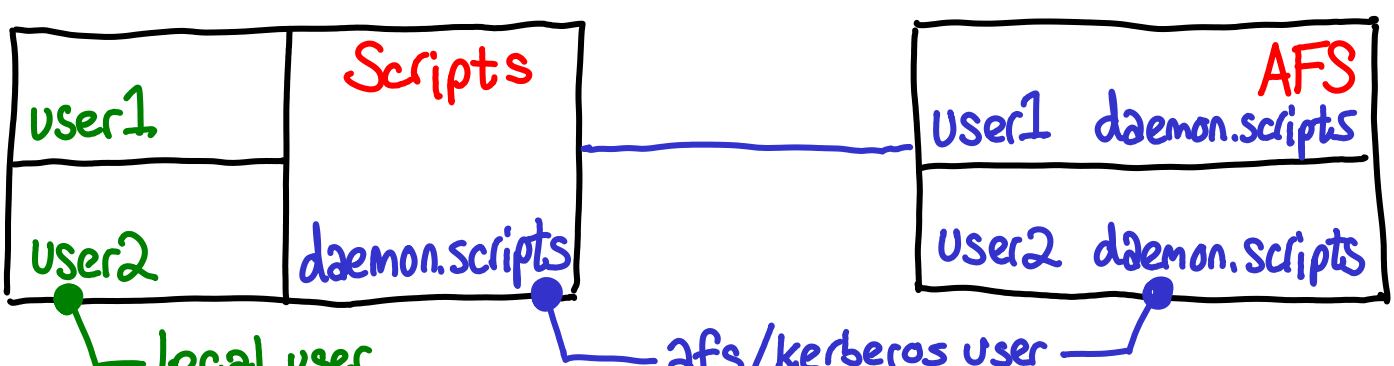
fs la ~/web_scripts → daemon.scripts rlidwk

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



Solution: Patch AFS

afs_vnop_access.c

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

local user info

credential store w/ daemon.scripts

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

can you guess what this does?

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

Extra considerations

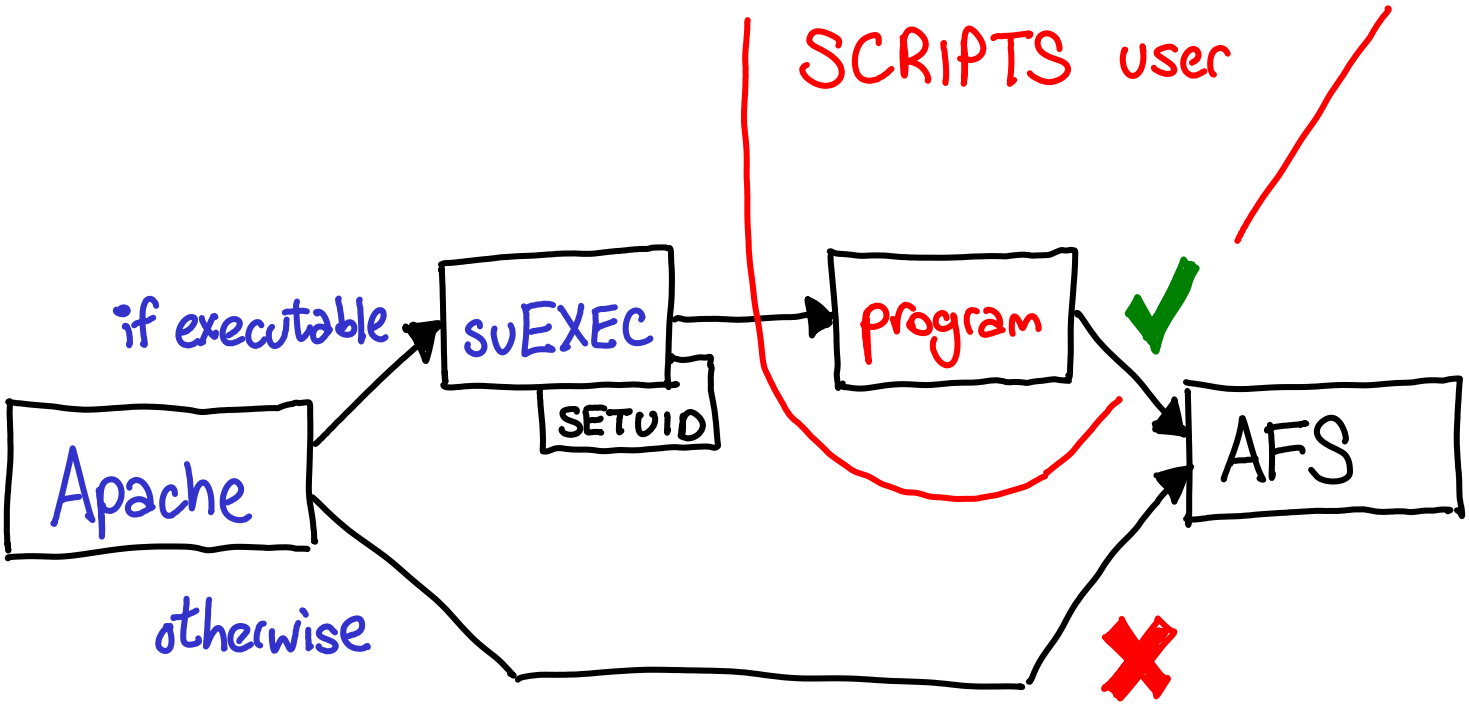
- Apache stat()s files before suEXECing

```
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)
```

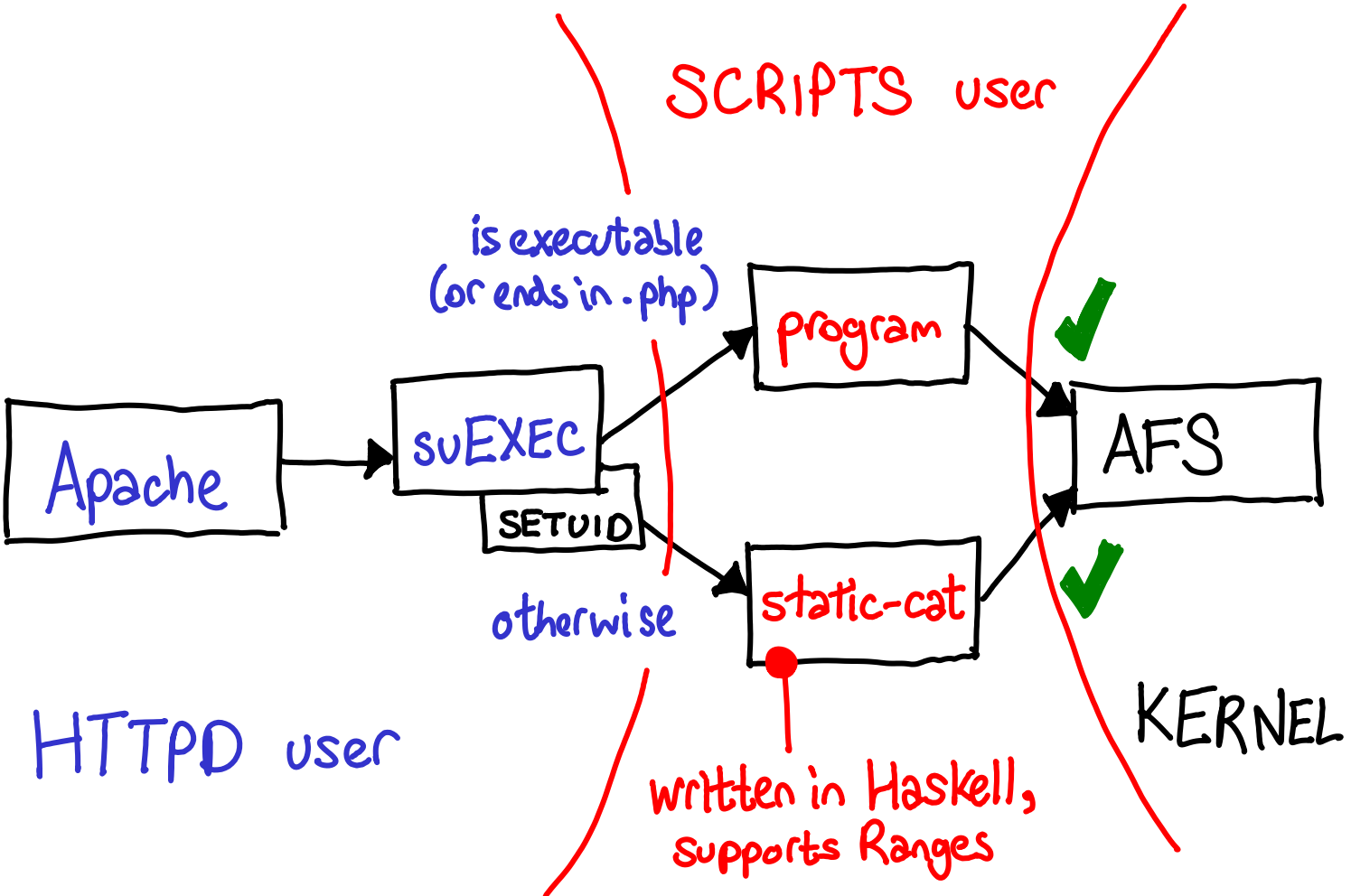
Problem: Apache doesn't suEXEC static files



HTTPD user, has <u>No</u> permissions

KERNEL

Solution: Make it suEXEC static files



One more thing... admof

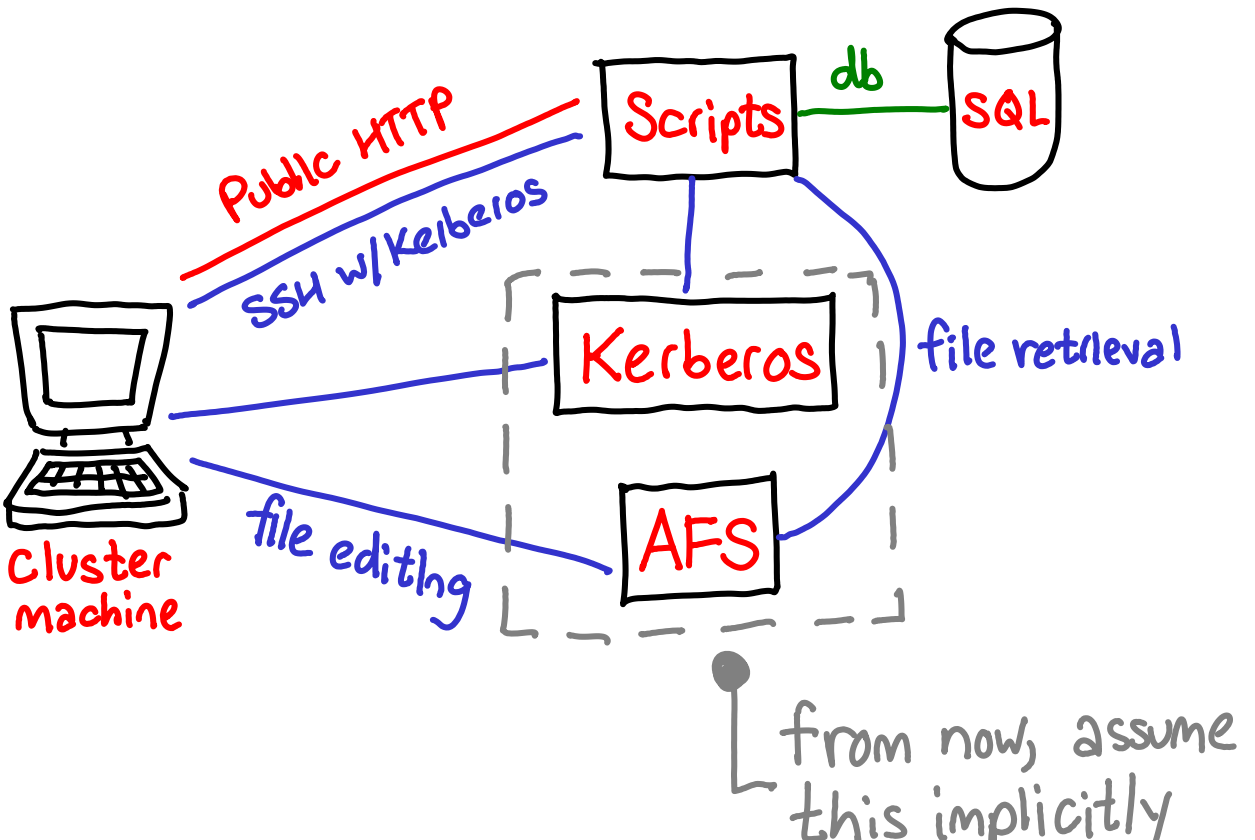
```
$ fs la /mit/scripts
system:scripts-root  rlidwka
system:anyuser       rl
```

← No "scripts" user

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

```
ezyang/root> ssh scripts@scripts.mit.edu
scripts>
```

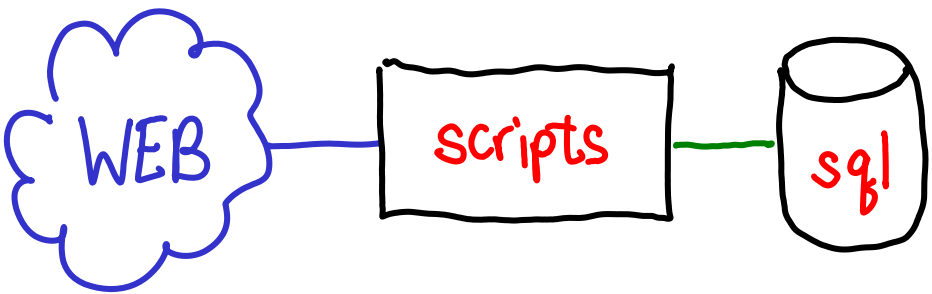
Here we stand. Questions?



~ Part II ~

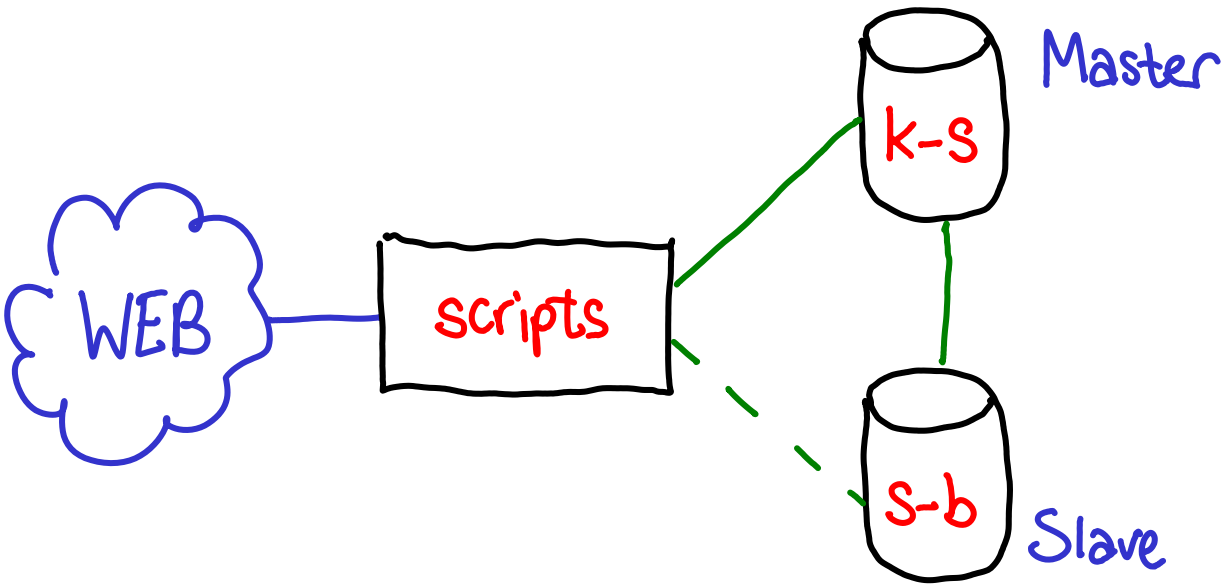
High Availability and Scaling

Scripts has a <u>lot</u> of users



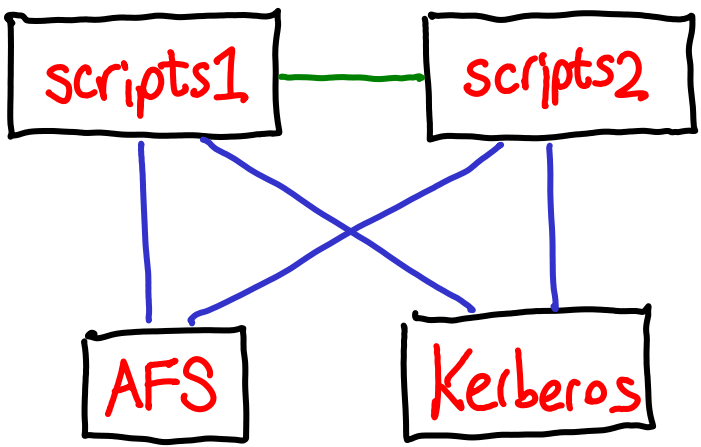
Single point of failure

Out-of-the-box database replication



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

Replicate Scripts



Some (but not all) work already done!

What is done

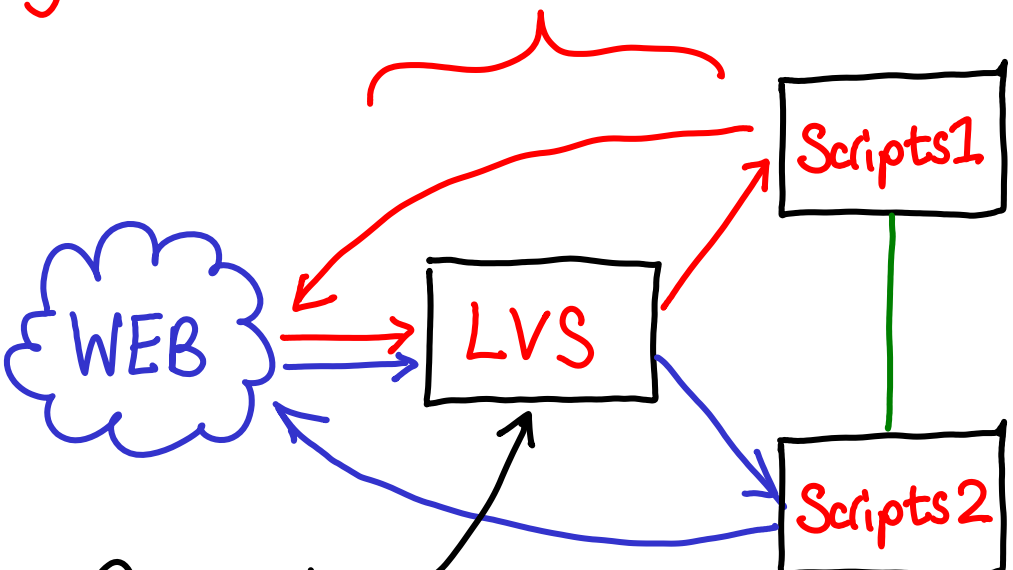
- User filesystem (AFS)
- User authentication (Kerberos)

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

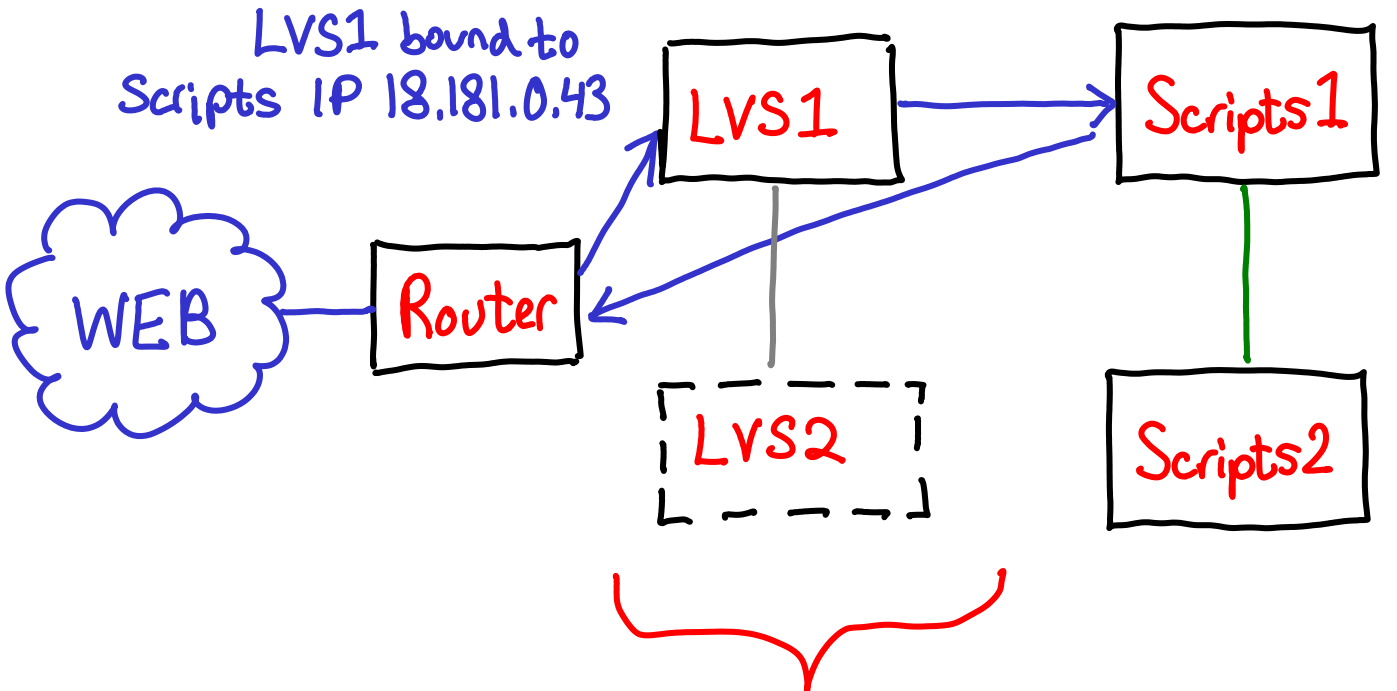


Forwarding is sticky

finger @scripts

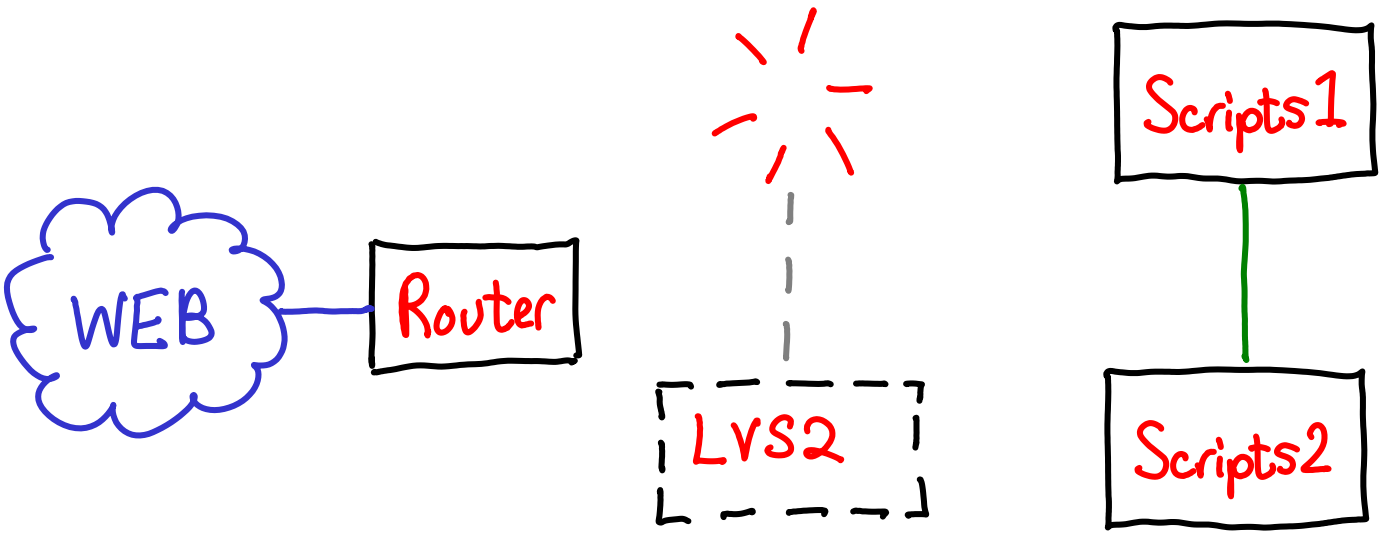
Incoming packets go through LVS; outgoing doesn't

More redundancy

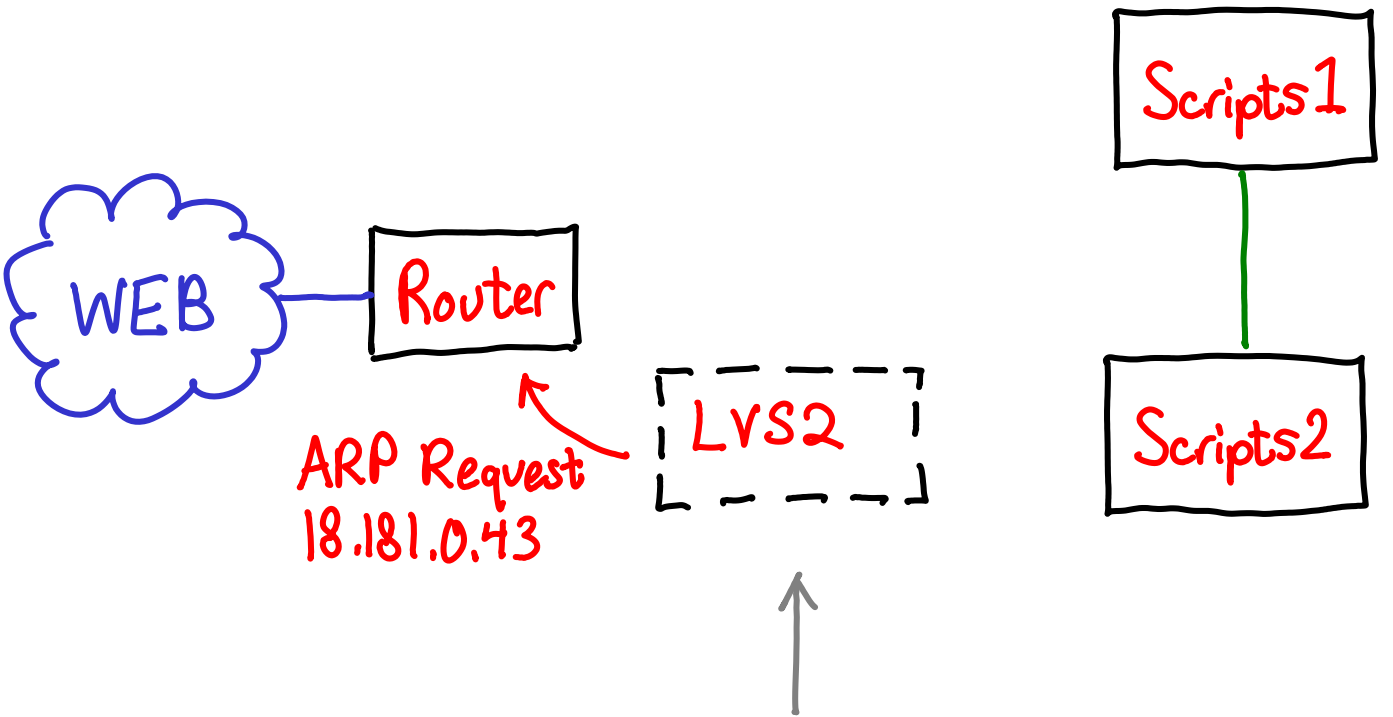


Running heartbeat to monitor LVS servers

LVS1 goes down

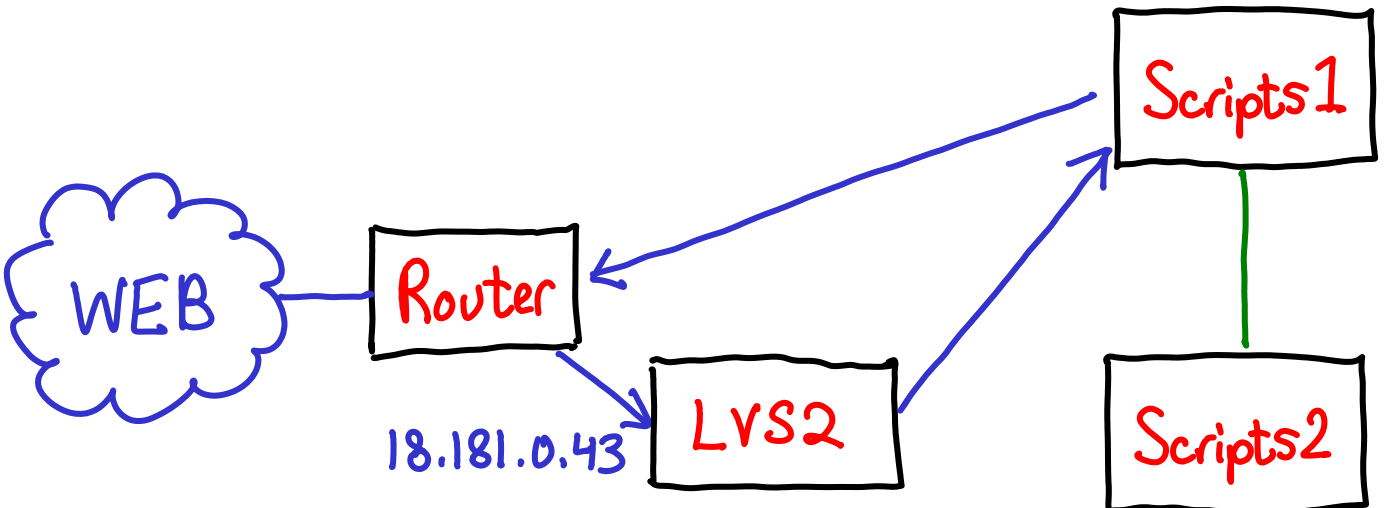


Arp! Arp!



Heartbeat notices LVS1 is down, starts ARPing for Scripts IP

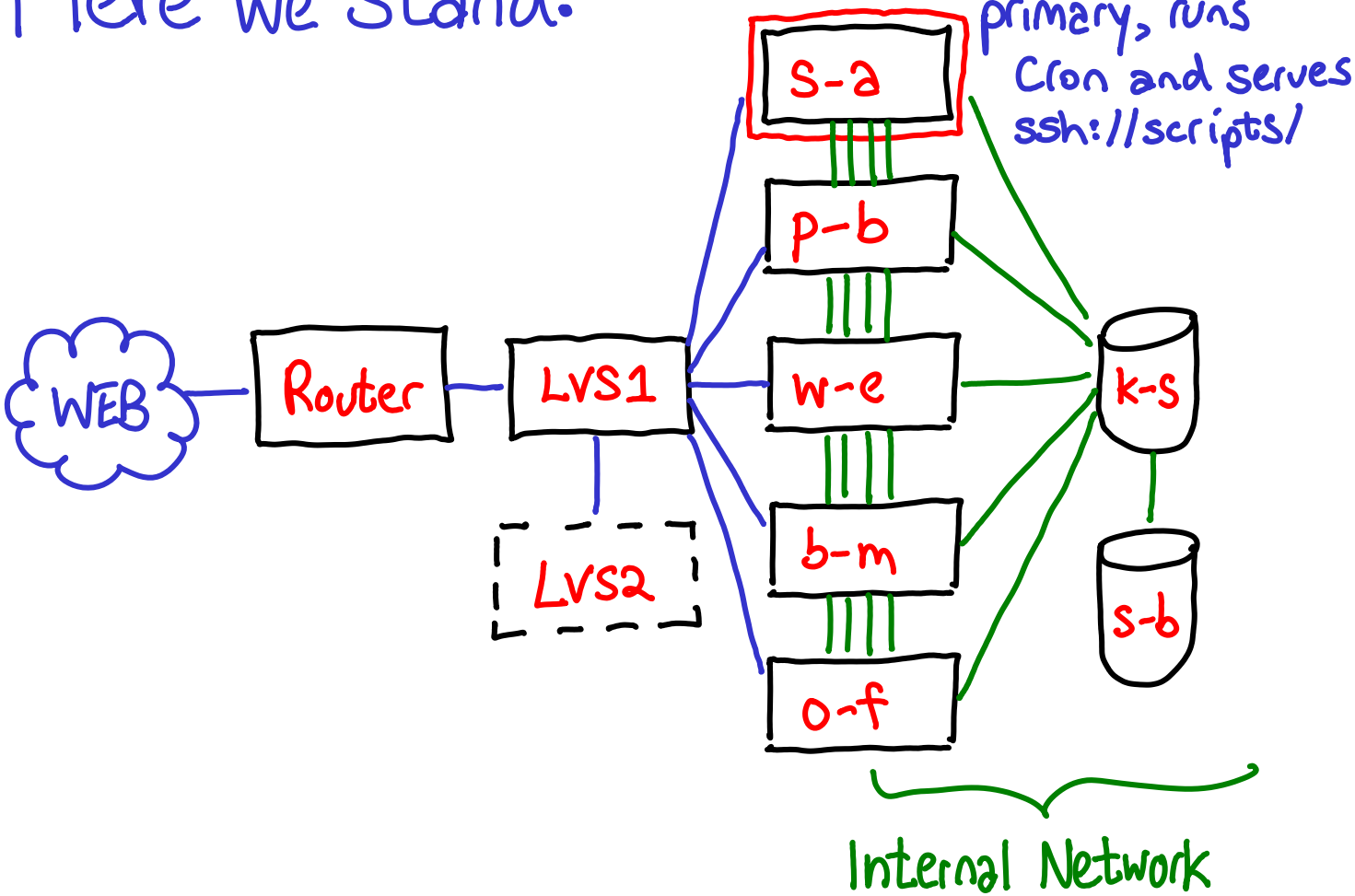
Back online (< 1 sec later)



We bind FIVE IP addresses total

On LVS1/LVS2, you can observe with crm_mon

Here we stand.



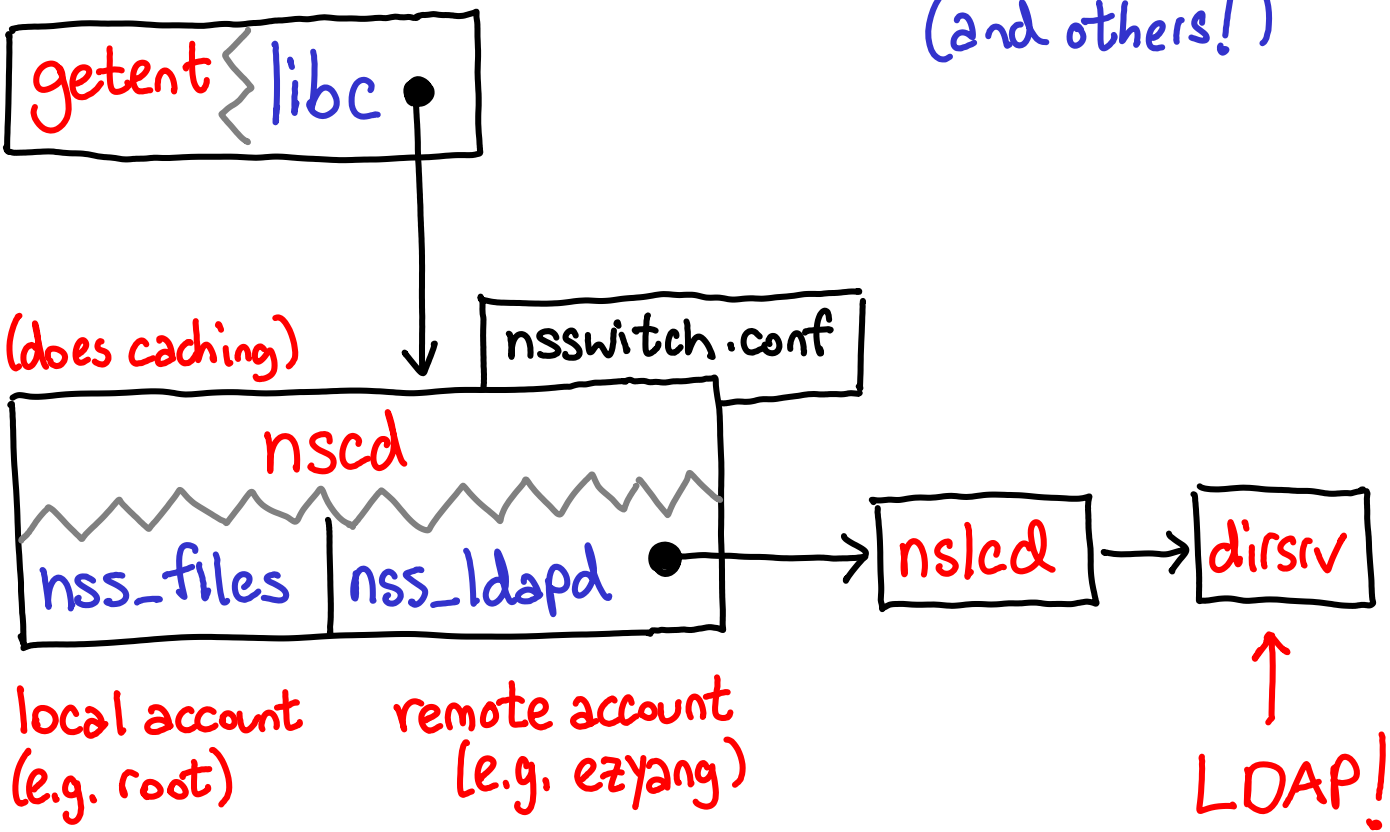
User information (not authentication)

```
$ 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!)



LDAP: Hierarchal 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
```

hierarchal identifier (address)

Athena UID

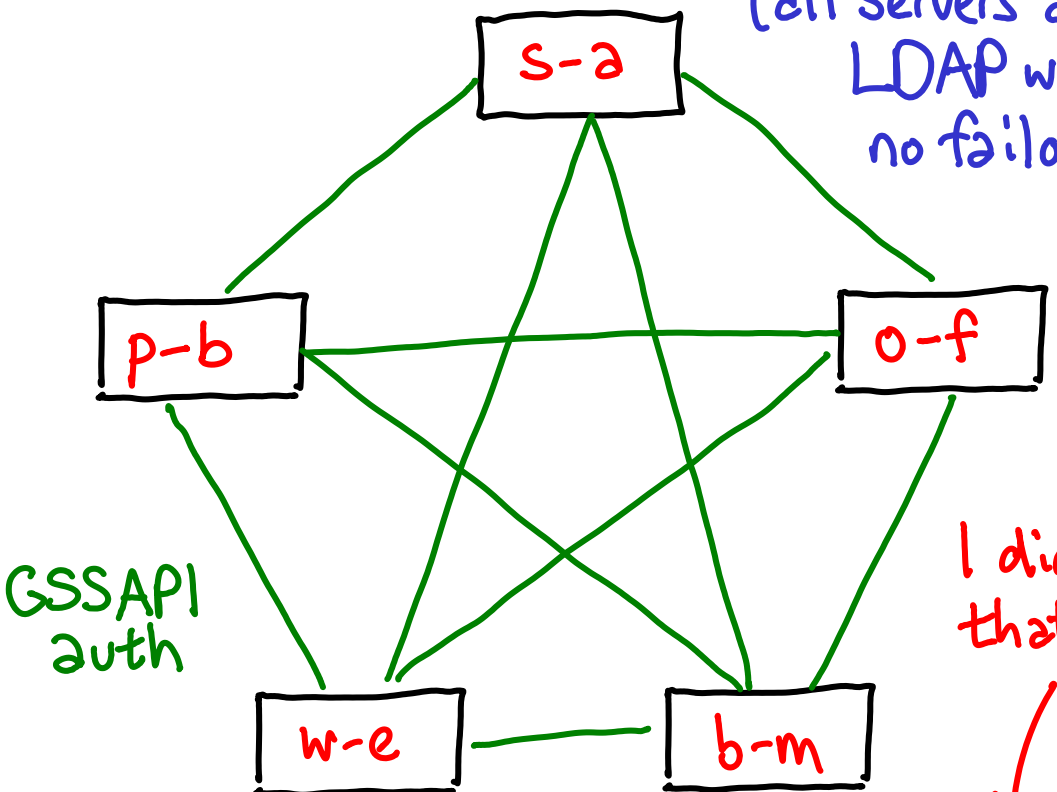
my volume ID (for the AFS patch)

the "Object" class

Enterprise, no?

... with Multi-Master Replication!

(all servers accept LDAP writes: no failover necessary)



I didn't pick that name...

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

Apache configuration

```
<VirtualHost ezyang.scripts.mit.edu>
    ~~~
</VirtualHost>
<VirtualHost geofft.scripts.mit.edu>
    ~~~
</VirtualHost>
    ⋮
```

Need to create these on signup!

Put it in LDAP...

or use vhostedit
or vhostadd!

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

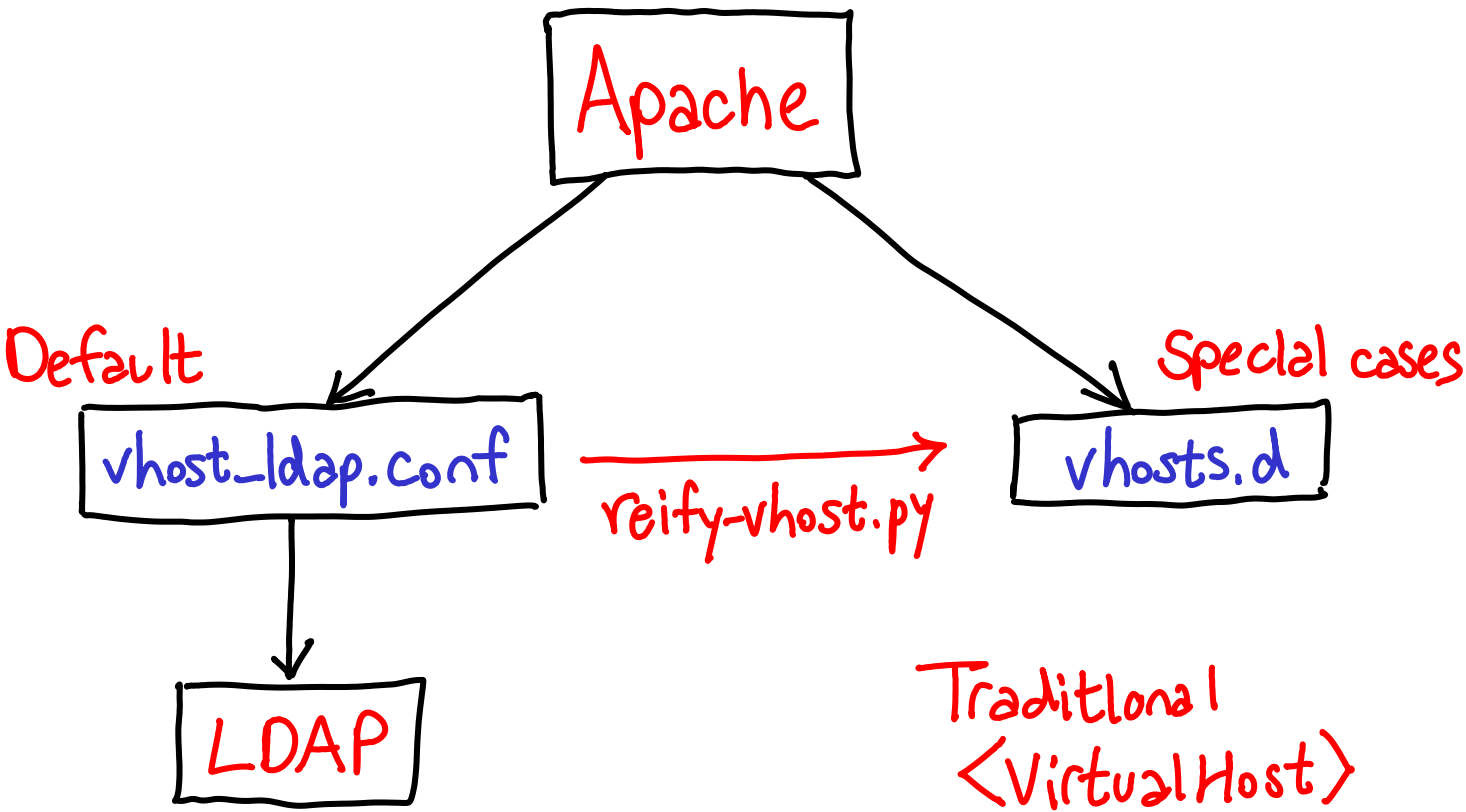
```
0 apacheServerName=ezyang.scripts.mit.edu,ou=VirtualHosts,dc=scripts,dc=mit,dc=edu
apacheSuexecGid: 71944
apacheSuexecUid: 537864399
apacheDocumentRoot: /afs/athena.mit.edu/user/e/z/ezyang/web_scripts
apacheServerAlias: ezyang.scripts
apacheServerName: ezyang.scripts.mit.edu
objectClass: apacheConfig
objectClass: top
```

different hierarchal prefix

different object class

Users can edit this with Scripts Pony

...tell Apache about it (mod_vhost_ldap)



Open Problems

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

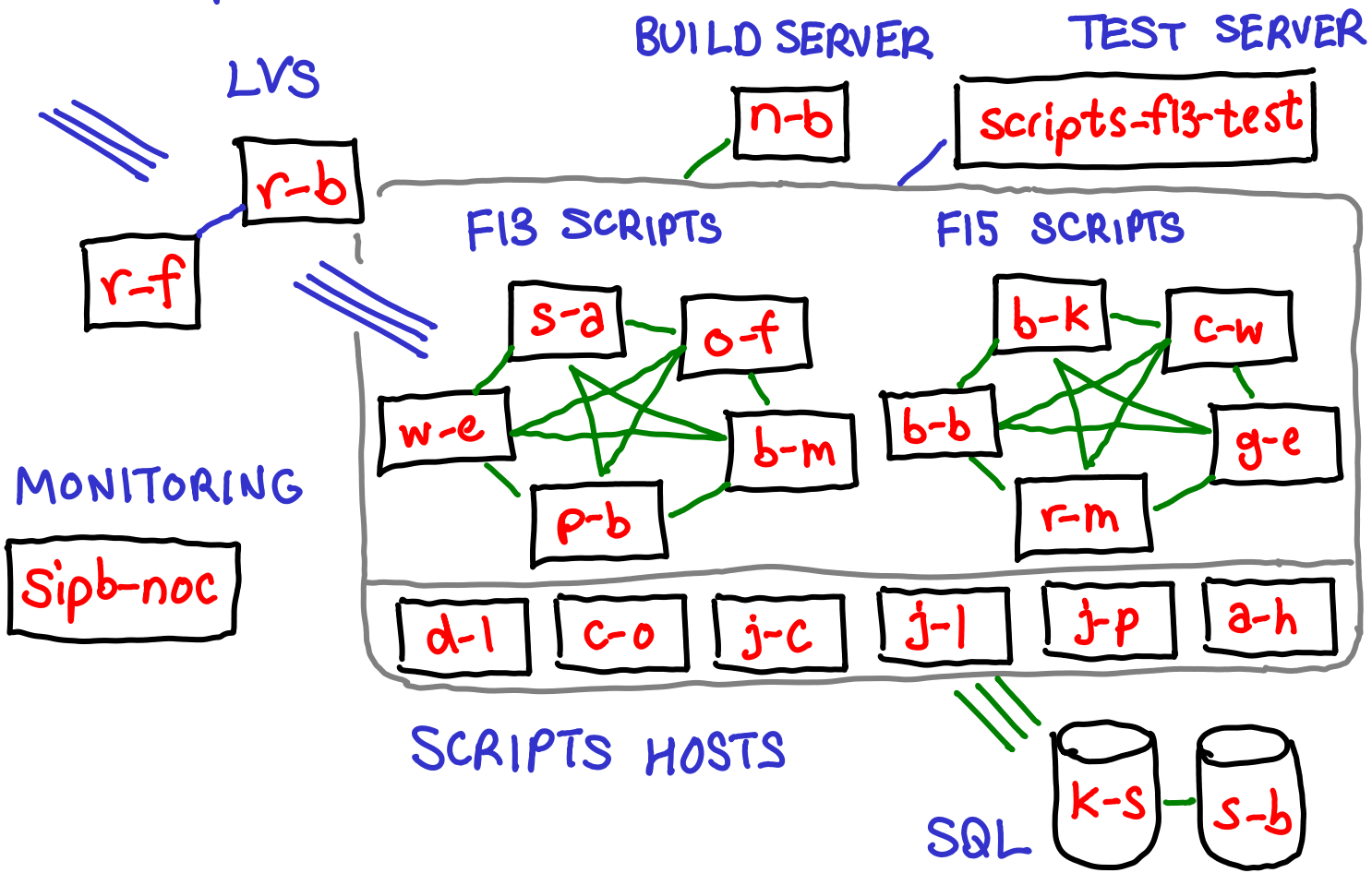
```
$ ldapvi -b ou=VirtualHosts,dc=scripts,dc=mit,dc=edu \
  scriptsVhostName=ezyang.scripts.mit.edu
```

```
0 scriptsVhostName=ezyang.scripts.mit.edu,ou=VirtualHosts,dc=scripts,dc=mit,dc=edu
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!


BUILD SERVER
n-b
TEST SERVER
scripts-f13-test
LVS
r-b
r-f
F13 SCRIPTS
F15 SCRIPTS
s-a
o-f
w-e
b-m
p-b
b-k
c-w
b-b
g-e
r-m
MONITORING
sipb-noc
d-l
c-o
j-c
j-l
j-p
a-h
SCRIPTS HOSTS
SQL
k-s
s-b

Questions?