

# DiskLess Debian/GNU Linux at DCE FEL CVUT.cz

Pavel Pisa

pisa@cmp.felk.cvut.cz

CC BY-SA

Czech Technical University in Prague  
Faculty of Electrical Engineering  
Department of Control Engineering

2011-3-6

InstallFest 2011



# Content of Presentation

- 1 Introduction
- 2 Setup
  - Server Side
  - DiskLess Station
- 3 Debugging and Use
  - Initramfs Problems

# Concept

- read-only root file-system over NFS overlaid by local tmpfs+aufs
- the distributed DiskLess root file-system maintained directly stored in subdirectory in NFS server
- regular Debian install by debootstrap
- maintenance by chroot and regular Debian tools (aptitude)
- execution of rd.d scripts blocked in chroot by policy-rc.d

# Outline

- 1 Introduction
- 2 Setup
  - Server Side
  - DiskLess Station
- 3 Debugging and Use
  - Initramfs Problems

# DiskLess Server Directories

- station root file-system and TFTP server for boot

/srv

- diskless
  - debian-squeeze
  - tools
    - SystemRescueCd
      - 1.x
      - 2.0
- homes
- tftp
  - boot
  - pxelinux.cfg
  - sysresccd
  - sysresccd2
- memdisk
- menu.c32
- pxelinux.0

# Published TFTP Directories

```
/srv/tftp
- boot
  initrd.img-2.6.32-5-amd64
  initrd.img-diskless-amd64
    -> initrd.img-2.6.32-5-amd64
  vmlinuz-2.6.32-5-amd64
  vmlinuz-diskless-amd64
    -> vmlinuz-2.6.32-5-amd64
- pxelinux.cfg
  default -> pxelinux-menu
  graphics.conf
  pxelinux-menu
  service.menu
- sysresccd
- sysresccd2
memdisk
menu.c32
pxelinux_0
```

# DHCP and PXE boot

- station network card BOOT ROM enabled
- boot device search sequence for PXE
- BIOS invokes DHCP
- configuration at DHCP server

```
next-server 192.168.1.2;  
filename "/pxelinux.0";
```
- PXE loads pxelinux.0 over TFTP
- pxelinux.0 look for config in pxelinux.cfg subdirectory
- config file search order (file by MAC, by hex IPv4 range, default)
- i.e. for 172.16.149.0/24 convert three first numbers into hex and add symlink

AC1095 -> pxelinux-menu

# pxelinux-menu header

```
menu PXE menu for DiskLess boot
menu INCLUDE pxelinux.cfg/graphics.conf
MENU AUTOBOOT Starting Local System in # sec
MENU TITLE InstallFest DiskLess Boot Menu
label bootlocal
    # Boot from local disk
    menu label ^Local boot
    menu default
    localboot 0
    timeout 300
    TOTALTIMEOUT 3000
```



## pxelinux-menu OS entry

```
label linux
    # Start Debian Squeeze DiskLess
    # add kernel option break=premount
    # to debug initramfs
menu label ^DiskLess Debian GNU/Linux
KERNEL boot/vmlinuz-diskless-amd64 ...
    boot=nfs ...
    root=/dev/nfs ro ...
    nfsroot=147.32.30.169:...
    /srv/diskless/debian-squeeze,ro,tcp
APPEND initrd=boot/initrd.img...
    -diskless-amd64
```

# NFS Setup at DiskLess Server

- what should be exported to which machines comes into /etc/exports

```
/srv/homes 172.16.149.0/24...
```

```
(rw,sync,no_root_squash,no_subtree_check)
```

```
/srv/diskless/debian-squeeze ...
```

```
172.16.149.0/24...
```

```
(ro,async,no_root_squash,no_subtree_check)
```

```
/srv/diskless/tools 172.16.149.0/24...
```

```
(ro,async,no_root_squash,no_subtree_check)
```

# Outline

- 1 Introduction
- 2 Setup
  - Server Side
  - DiskLess Station
- 3 Debugging and Use
  - Initramfs Problems

# Bootstrap Distributed OS

- Use standard Debian bootstrap process for Debian install/replication from live system

```
debootstrap --keyring=/usr/share/keyrings/...  
  debian-archive-keyring.gpg \  
  --arch=amd64 --include=debian-keyring,etckeeper \  
  squeeze /srv/diskless/debian-squeeze/ \  
  ftp://ftp.cz.debian.org/debian/
```

- Notice inclusion of etckeeper from the first moment of the system instance born

# EtcKeeper

- GNU/Linux configuration and /etc are almost synonyms
- It is critical to log/document changes (even more if multiple admins maintain the system)
- It is important to be able undo changes breaking the system
- Ability to separate intended admin changes from massive changes caused by package managers
- EtcKeeper manages whole /etc directory in GIT repository with broader range of permissions /etc/.etckeeper
- Automatic commits before and after changes caused by APT/aptitude
- Implemented by integration of GIT through hooks into APT system

# Services Control and Chroot

- The “image” of distributed system is in  
`/srv/diskless/debian-squeeze/ now`
- Intention is to manage it by regular Debian tools in chroot⇒  
services managed by APT and started after packages  
install⇒screws networking and services of the DiskLess server  
system
- Debian system provided a way to disable services start-up  
(rc?.d, init.d) by APT at specific computer
- Create and edit for policy-rc.d distributed OS  
`nano /srv/diskless/debian-squeeze/usr/sbin/policy-rc.d`
- Script signals that on DiskLess server no init.d scripts should  
be run

```
#!/bin/sh  
[ "installtest.sh.cvut.cz" = "hostname" ] && exit 101  
exit 0
```

# The First Switch to Distributed System

```
chroot /srv/diskless/debian-squeeze  
cd /etc  
nano apt/sources.list
```

- Select mirror for Debian packages archive

```
#deb http://ftp.us.debian.org/debian squeeze main ...  
deb http://ftp.cz.debian.org/debian squeeze  
main contrib non-free deb-src ...  
http://ftp.cz.debian.org/debian ...  
  
squeeze main contrib non-free
```

- Document and commit changes into etckeeper/GIT

```
git add apt/sources.list  
git status  
git commit -m ...  
  
"my: Select Czech mirror for sources.list"
```

- Notice “my:” prefix. Prefix use allows to find/filter own changes

# Etckeeper is too Clever for real Admins

- EtcKeeper in default configuration commits user/admin changes before APT install/upgrade/remove run
- It even commits all /etc changes each day at midnight
- Such commits are unwanted in admins managed because they are not divided into documented logical steps
- But automatic commits can be disabled and then APT/aptitude refuses to run when local/manual changes are not committed or undone by admin



# Disable EtcKeeper Automation

- Edit EtcKeeper configuration

```
nano etckeeper/etckeeper.conf
```

- Enable/add next lines

```
AVOID_DAILY_AUTOCOMMITS=1  
AVOID_COMMIT_BEFORE_INSTALL=1
```

- Commit changes

```
git commit ...  
-m "my: configure etckeeper to block package ...  
manager if there are uncommitted changes in etc"  
git log
```

# Packages for DiskLess

- Regular

```
aptitude update
```

- Remark about obtaining packages list from other machine

```
aptitude search '?installed?not(?automatic)' \  
| sed -n -e 's/^i[ \t]\+\([^\ \t]*\)[ \t].*$/\1/p' \  
>packages
```

- Remove grub and grub2 for now, makes problems during updates when no device map can be scanned

```
aptitude install 'cat packages'
```

# Remarks About Packages

- Most important packages
  - linux-image-2.6
  - aufs-tools
  - nfs-common
  - ntp ntpdate
  - openssl
  - rcconf

## Policy-rc.d Again - Emphasis

- The policy-rc.d is important and it demonstrates as

```
Setting up network-manager (0.8.1-6) ...  
invoke-rc.d: policy-rc.d denied execution of force-reload.  
Disabling interfaces configured with plain DHCP in /etc/network/inter-  
faces so that NetworkManager can take them over  
Auto interfaces found:  
invoke-rc.d: policy-rc.d denied execution of start.
```

- If not defined the server system/kernel state/configuration would be screwed to death

# Caveats of Chroot

- If problem during aptitude run are encountered

```
cd /etc  
git status  
git add .  
git commit -m "partial packages install recovered"
```

# JAVA and /proc

- Typical problem with Java packages

```
Setting up openjdk-6-jre-headless (6b18-1.8.3-2) ...  
the java command requires a mounted proc fs (/proc).  
dpkg: error processing openjdk-6-jre-headless (--configure):  
subprocess installed post-  
installation script returned error exit status 1
```

- Fixed by

```
mount /proc  
aptitude install ant  
umount /proc
```

- The clients have problems with NFS mount when proc is left mounted in the image.

# Locales

- Run

```
dpkg-reconfigure locales
```

- Select what is convenient for your region

```
git status
```

```
git add .
```

```
git commit -m "my: dpkg-reconfigure locales"
```

## Core Changes for Overlay

- There is necessary more changes, provided as patch  
`git am /patches/my-initramfs-and-init.d-scripts-for-diskless.patch`
- Possible conflicts resolution discussed later
- Most important changes
- Add aufs and or unionfs into  
`initramfs-tools/modules`
- initramfs scripts  
`initramfs-tools/scripts/nfs-bottom/diskless_setup`  
`initramfs-tools/scripts/nfs-bottom/root_overlay`
- Services start-up init.d scripts  
`init.d/diskless-linux`  
`init.d/findswap`



# Dissection of initramfs

```
zcat initrd.img-2.6.32-5-amd64 | cpio --extract -d
```

- or use Midnight Commander

```
zcat initrd.img-2.6.32-5-amd64 >initrd-to-check.cpio  
/srv/diskless/debian-squeeze/boot/initrd-to-check.cpio  
#ucpio/scripts/nfs-bottom  
ORDER  
diskless_setup  
root_overlay
```

Setup server location and loopback device

```
initramfs-tools/scripts/nfs-bottom/diskless_setup
```

Script content

```
rm -f ${rootmnt}/etc/hosts  
cat <<EOF > ${rootmnt}/etc/hosts  
127.0.0.1 localhost  
${IPV4ADDR} ${HOSTNAME}.${DNSDOMAIN} ${HOSTNAME}  
${nfsroot%/*} diskless-server  
EOF
```

# Scrip to place overlay over NFS root file-system

```
initramfs-tools/scripts/nfs-bottom/root_overlay
```

And inside is ...

```
mkdir /tmp/unirw
mount -n -t tmpfs none /tmp/unirw
mount -n -t aufs -
o dirs=/tmp/unirw=rw:${rootmnt}=nfsro unionfs ${rootmnt}
mkdir -p ${rootmnt}/overlay/unirw
mount -n -o move /tmp/unirw ${rootmnt}/overlay/unirw
chmod 755 ${rootmnt} # Disable rw access for non-
root users
mkdir -p ${rootmnt}/tmp
mkdir -p ${rootmnt}/var/lock
```

For debugging look for maybe\_break XXX in

```
/srv/diskless/debian-squeeze/boot/initrd-to-
check.cpio#ucpio/init
```

# Kernel and Initramfs to TFTP

- PxeLinux loads kernel and initramfs over TFTP
- initramfs has to be updated after previous changes

```
update-initramfs -u -k 2.6.32-5-amd64
```

```
update-initramfs -u -k 2.6.32-5-xen-amd64
```

- Copy vmlinuz-\* and initrd.img-\* from

```
/srv/diskless/debian-squeeze/boot
```

to

```
/srv/tftp/boot
```

# Config Changes

- Run `rcconf` to activate `findswap` and `diskless-linux` scripts

```
git add .
```

```
git commit -m "my: run rcconf to activate findswap and diskless-linux scripts"
```

- Setup root password

```
passwd
```

```
git add shadow
```

```
git commit -m "my: Changed root's password"
```

- Use `rcconf` to disable `avahi-daemon` and `network-manager`

```
git add -A
```

```
git status
```

```
git commit -m "my: avahi and networkmanager disabled at rcconf level"
```

# DiskLess Station FsTab

```
nano fstab
```

```
none /tmp tmpfs defaults 0 0  
none /var/lock tmpfs defaults 0 0  
# /etc/hosts contain the correct IP ad-  
dress for diskless-server (set in ramdisk scripts)  
diskless-  
server:/srv/homes /home nfs defaults,tcp 0 0
```

```
git add .
```

```
git commit -m "my: Added /srv/homes and RAM based tem-  
poraries to fstab"
```

# Network Interfaces

```
nano network/interfaces
    auto lo
    iface lo inet loopback

git add .
git commit -m "my: Added local host / lo interface"
```

- This is critical even for mount over NFS to work.
- Check this if you encounter problems.

# Local Time

```
dpkg-reconfigure tzdata
git add timezone
git commit -m "my: Changed timezone to Europe/Prague by dpkg-
reconfigure tzdata"
git add localtime
git commit -m "my: Localtime after run of dpkg-
reconfigure tzdata"
```

If the computer hardware should be shared with Windows boot then use crummy setup to keep RTC date in local time instead of UTC

```
nano default/rcS

    UTC=no

git add default/rcS
git commit -m "my: keep hwclock in lo-
cal (non UTC) time for compatibility with Windows"
```

# No Async NFS Mount on Station

The asynchronous NFS mounts invocation after network manager or udev events does not work with interface configured during boot in initramfs

```
nano default/rcS
```

Add option

```
ASYNCMOUNTNFS=no
```

Commit

```
git add default/rcS
git commit -m "my: additional fstab defined NFS mounts re-
quires ASYNCMOUNTNFS=no for diskless"
```



# Home Directories

Create directory /home-local for temporary/local users

```
git am /patches/my-added-local-guest-account-and-guestXX-accounts.patch
git am --resolved
git am --skip
git am --abort
git am /patches/0008-My-Added-automatic-creation-of-home-directories.patch
git am /patches/0009-My-Group-permissions.patch
git am /patches/0005-My-Added-CVUT-FELK-ceritificated-needed-for-LDAP-a.patch
```

# Outline

- 1 Introduction
- 2 Setup
  - Server Side
  - DiskLess Station
- 3 Debugging and Use
  - Initramfs Problems

# InitRamfs Break and Step

The boot option `break=mountroot` allows test NFS mount infrastructure step by step

```
modprobe nfs
modprobe af_packet
ipconfig -t 180 -c dhcp -d eth0
ping -c 4 'echo $nfsroot | sed -n -e 's/^(.*):\./.*$/\1/p''
mkdir /test-root-mnt
nfsmount -o nolock -o ro -o tcp 'echo $nfsroot \
  | sed -n -e 's/^(\[^\,]*\)\\(,.*\\|\\)$/\1/p'' /test-root-mnt
```

- Some more exercises with BusyBox building and use <http://rtime.felk.cvut.cz/osp/cviceni/2/>

## Links to DCE Setup Information

- Information about GNU/Linux in DCE.FEL.CVUT.cz laboratories  
http:  
`//support.dce.felk.cvut.cz/mediawiki/index.php/...  
GNU/Linux_v_laborato%C5%99%C3%ADch`
- A4M35OSP Open-source programming  
this subject is one of more courses which has been prepared and provided by our group and use described set-up  
`http://rttime.felk.cvut.cz/osp/cviceni/`
- Other courses A3M35PSR, Y35ORT, A4B35PSR, Y35PES
- Former or suspended for this year X35POS, X35MSY, Y35SVS
- Playing with BusyBox and initramfs  
`http://rttime.felk.cvut.cz/osp/cviceni/2/`
- People behind the set-up and maintenance of Linux @ DCE  
Aleš Kapica, Pavel Píša, Martin Samek, Michal Sojka

## Links to Similar Projects

- [http://en.wikipedia.org/wiki/Diskless\\_Remote\\_Boot\\_in\\_Linux](http://en.wikipedia.org/wiki/Diskless_Remote_Boot_in_Linux)
- <http://www.drbl.org/>
- <http://clonezilla.org/>