Fully automatic Linux installations

Debconf5, Helsinki.

Thomas Lange, University of Cologne

lange@informatik.uni-koeln.de

Overview

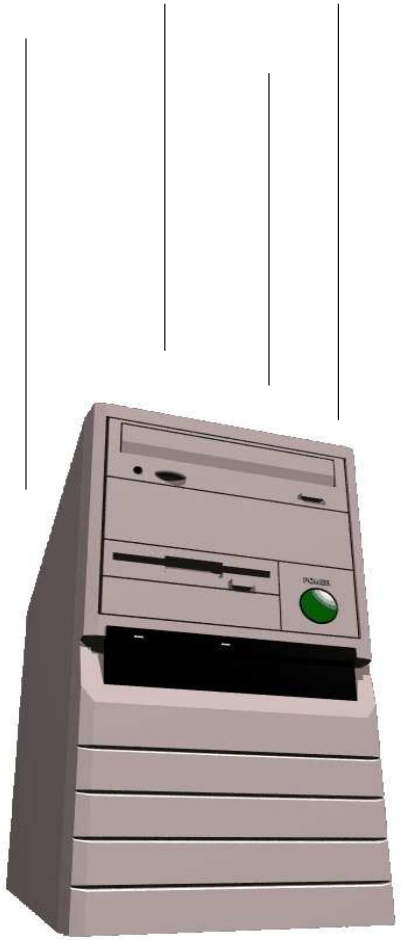
- Motivation
- How to plan a computer infrastructure
- How does FAI work?
- Comparing d-i and FAI
- Present and future
- The show

Value of your computer

- What are the values of your computer?
- What happens if your computers are not running for one hour/day?
- A good computer infrastructure is as important as ...?
- Which valuables are included in your computers?
  - Customer data (address, email, orders, bills)
  - Services (email, web, databases, printing)
  - Applications (text processing, compiler, CAD, tools)
  - Input and output (CAD design, simulation results)
  - Internal company know-how (source code)
- How do you save these values? Data backup only?
- Have you really saved everything when doing backups?

The Test

- Grab a random machine (without a backup before)
- Throw it out a 10th floor
- or `dd if=/dev/zero of=/dev/hda`



- Recover all sysadmin work within 10 minutes
- Can you?

Manual installation?

Who likes to install these hosts by hand?

20 nodes dual XEON, 2.4 GHz

90 dual Itanium 2, 900Mhz

www.centibots.org

Manual installation?

Can you guarantee, that all these hosts are equal?

Manual installation

- ”No simple sysadmin task is fun more than twice”
- Manual installation and configuration lasts many hours
- Many questions have to be answered
- Equal data must be entered again and again
- No parallel installations
- Repeating tasks are stupid and will lead to errors
- No documentation is made
- Can you rebuild the installation? After several months?
- Each installation is unique, but unintentionally
- **A manual installations does not scale !**

Why not fully automaticly?

- Automated installations only lasts a few minutes
- Identical configuration are guaranteed (even after several month)
- Quick reinstallation after replacement of defective hardware (Disaster recovery)
- One command – hundreds of installations
- Diversity of hardware and different configurations easily manageable
- You can save much work! (work = time = money)
- Do you have a plan for your computer infrastructure?

What is FAI ?

- FAI does everything a sysadmin (you!) has to do, before users can log in a brand new computer for the first time
- Server based tool for a script based automatic installation of Debian GNU/Linux or Solaris
- Installs and configures the OS and all applications
- No master or golden image needed
- Class system provides modularity
- Flexible and easy to expand with hooks
- It can’t plan your installation :-(, but
- **Plan your installation and FAI installs your plan! :-)**

Planning an infrastructure

- Don’t look at a single computer, consider the whole infrastructure
- `www.infrastructures.org`
- Paper: *Bootstrapping an infrastructure* by Traugott and Huddleston
- Record your actual state
- What would you like to change in the future?
- Bear in mind future extensions
- Put your infrastructure data into version control (CVS)
- Which things are equal, which are different?
- One data, one source

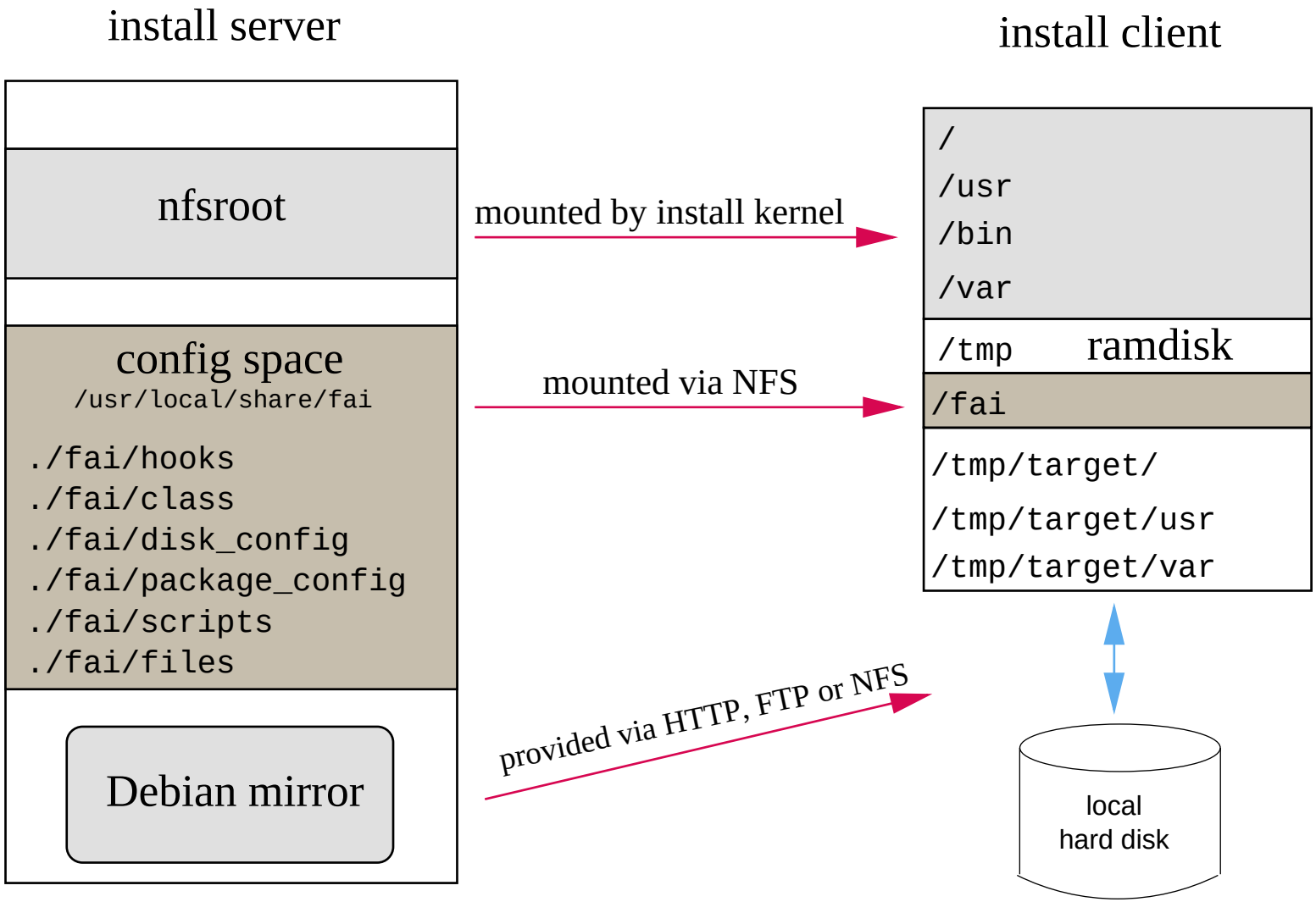
Questions for an infrastructure

- Which type of computer will be installed? Cluster, desktop, server, notebook?
- Which jobs do the computers have? CAD, server, text processing
- Which applications will be run on them?
- How does my LAN topology looks like? Is DHCP available?
- Do I have uniform hardware? In the future?
- Does the hardware need a special kernel?
- How should the local hard disks be partitioned?
- Do the users need a queueing system?
- What software should be installed?
- Which daemons should be started? How?
- Which remote filesystems should be mounted?
- What about user accounts, printers, mail system, cron jobs, graphic cards, dual boot, NIS, NTP, timezone, keyboard layout,...?

How does FAI work ? The sysadmin point of view

A system administrator during a fully automatic installation

How does FAI work ? The technical point of view



- Configuration is stored on the install server (one tree for all clients)
- Installation runs on the client

Requirements?

- A server with DHCP, NFS and TFTP (install server)
- A computer with network interface card (install client)
- You can also boot from floppy or CD-ROM
- Not needed: floppy disk, CD–ROM, keyboard, graphic card
- Access to a local Debian mirror via NFS, FTP or HTTP
- Disk space on the install server:

FAI package	13 MB	kernel, scripts and configuration data
nfsroot	240 MB	created with `make-fai-nfsroot`
Debian mirror	9 GB	Debian 3.1 (sarge, i386 only)

- All install client share the same nfsroot
- **Constant disk space**

Sequence of an installation

- Plan your installation!
- Install client boots from NIC via PXE and gets its kernel via TFTP
- Boots linux using the nfsroot, without using the local hard disks
- Start of the main script (`fai`), which controls the installation
- Detect hardware and load kernel modules (`discover2`)
- Define classes and variables (`fai-class`)
- Partition local hard disk, create and mount file systems (`setup_harddisks`)
- Install software packages (`install_packages`)
- Configure operating systems and applications
- Save log files to the local disk and to the install server
- Boot the newly installed system

The class concept

- A host belongs to several classes
- Examples: `DEFAULT FAIBASE GRUB GNOME demohost LAST`
- Order of the classes defines the priority from low to high
- Classes are defined via scripts in `/fai/class`
- All parts of the installation use the classes
- Config files are selected based on the name of a class
- `fcopy` copies files based on classes
- Senior admin creates classes
- Junior admin assigns classes to a host
- PC installs itself

Directory tree of the config space

```
|-- class
|   |-- 10-base-classes
|   |-- 20-hwdetect.source
|   |-- 50-host-classes
|   |-- FAIBASE.var
|   `-- GERMAN.var
|-- disk_config/
|   |-- FAIBASE
|   |-- SMALL_IDE
|   `-- foobar04
|-- package_config/
|   |-- FAIBASE
|   |-- DEBIAN_DEVEL
|   |-- DEMO
|   |-- GERMAN
|   |-- GNOME
|   `-- server07
```

Example `/fai/class/07example`:

```
#! /bin/sh
# echo architecture and OS name in upper case
uname -s | tr '[:lower:]' '[:upper:]'                          # LINUX
dpkg --print-installation-architecture | tr /a-z/ /A-Z/        # I386

case $HOSTNAME in
    demohost)
        echo "FAIBASE DHCPC DEMO" ;;
    gnomehost)
        echo "FAIBASE DHCPC DEMO XFREE GNOME";;
esac

case $IPADDR in
    134.95.9.*) echo "CS_KOELN NET_9" ;;
esac

ifclass I386 && echo "GRUB"
```

Defining variables

Example `/fai/class/FAIBASE.var`:

```
FAI_CONSOLEFONT=
FAI_KEYMAP=us-latin1

UTC=yes
time_zone=Europe/Berlin

rootpw=’3h54Vqh57F’

moduleslist="usbkbd usb-uhci keybdev mousedev hid psmouse"
```

- You can also define your own variables
- All customization scripts in `/fai/scripts/*` are using these variables

Example: `/fai/disk_config/FAIBASE`:

```
# <type> <mountpoint> <size in mb> [mount options] [;extra options]

disk_config disk1

primary /           70-150    rw,errors=remount-ro ;-c -j ext3
logical swap        50-500    rw
logical /var        50-1000   rw             ; -m 5  -j ext3
logical /tmp        50-1000   rw             ; -m 0 -j ext3
logical /usr        300-4000  rw             ; -j ext3
logical /home       50-4000   rw,nosuid      ; -m 1 -j ext3
logical /scratch    0-        rw,nosuid      ; -m 0 -i 50000 -j ext3
#logical /scratch preserve10 rw,nosuid      ; -m 0 -i 50000 -j ext3
```

- File systems: ext2, ext3, vfat, xfs, ReiserFS
- RAID and LVM only possible via hooks

Example: `/fai/package_config/BEOWULF`:

```
# packages for Beowulf clients

PACKAGES install BEOWULF_MASTER
gmetad apache

PACKAGES install
fping jmon ganglia-monitor
rsh-client rsh-server rstat-client rstatd rusers rusersd

dsh update-cluster-hosts update-cluster etherwake

lam-runtime lam4 lam4-dev libpvm3 pvm-dev mpich
scalapack-mpich-dev
```

- Actions as in `apt-get`: install, remove
- Also `aptitude, aptitude-r`
- Dependencies are resolved
- `dpkg -get-selections` also possible

Directory tree of the config space

```
|-- scripts/
|   |-- BOOT
|   |-- FAIBASE/
|   |   |-- 10-misc                     Bourne shell script
|   |   |-- 30-interface                Bourne shell script
|   |   `-- 40-misc                     /usr/bin/cfengine script
|   |-- DEMO/
|   |   |-- 10-misc                     Bourne shell script
|   |   `-- 30-demo                     /usr/bin/cfengine script
|   `-- demohost
`- files/
    |-- etc/
    |   |-- X11/
    |   |   `-- XF86Config-4/           fcopy /etc/X11/XF86Config-4
    |   |       |-- ATI_ACER
    |   |       |-- MATROX
    |   |       `-- demohost
    |   `-- nsswitch.conf/              fcopy /etc/nsswitch.conf
    |       |-- NIS
    |       `-- NONIS
```

Customization scripts

```
# create NIS/NONIS config
fcopy -M /etc/nsswitch.conf /etc/host.conf
fcopy -i /etc/ypserv.securenets  # only for yp server
ifclass NONIS && rm -f $target/etc/defaultdomain
if ifclass NIS; then
    echo $YPDOMAIN > $target/etc/defaultdomain
    rm -f $target/etc/yp.conf
    for s in $YPSRVR; do
        echo "ypserver $s" >> $target/etc/yp.conf
    done
fi

ifclass USR_LOCAL_COPY && {
    mount -o ro $bserver:/usr/local /usr/local
    cp -a /usr/local $target/usr
}
fcopy -M /etc/X11/XF86Config-4 && rm -f $target/etc/X11/XF86Config
```

Cfengine example

```
files:
   any::
      ${target}/dev include=fd* mode=666   action=fixall r=1

editfiles:
   any::
   { ${target}/etc/fstab
         AppendIfNoSuchLine "none  /proc/bus/usb  usbdevfs  defaults"
         AppendIfNoSuchLine "/dev/fd0  /floppy  auto  users,noauto 0 0"
   }
   { ${target}/etc/inittab
      ReplaceAll "/sbin/getty" With "/sbin/getty -f /etc/issue.linuxlogo"
   }
 HOME_CLIENT::
   { ${target}/etc/fstab
     HashCommentLinesContaining "/home "
     AppendIfNoSuchLine "${hserver}:/home /home nfs rw,nosuid 0 0"
   }
```

Installation times

Host	RAM in MB	Software in MB	Time
Pentium 4 2.6 GHz	512	190	2 min
Pentium 4 2.6 GHz	512	750	7 min
Pentium 4 2.6 GHz	512	2600	15 min
Pentium III 850MHz	256	180	3 min
Pentium III 850MHz	256	820	10 min
Pentium 4 2.80 GHz	1024	948	5 min
Athlon XP1600+	896	1000	6 min
AMD-K7, 500MHz	320	780	12 min
PentiumPro 200MHz	128	800	28 min

Nodes	Seconds
1	337
5	340
10	345
20	379

12% more time for 20 hosts in parallel

FAI users

- Electricité de France (EDF), France, 200 hosts
- France Telecom, TRANSPAC, France, 300 hosts
- MIT Computer science research lab, 200 hosts
- Danmarks Meteorologiske Institut, 85+ hosts
- Physics department (FU Berlin), 139+ hosts
- University of New Orleans, 72 node Beowulf cluster
- Brown University, Dep. of Computer Science, 300+ hosts
- University of West Bohemia, Czech Republic, 180+
- Host Europe, 250 hosts
- Lycos Europe, search engine, 200+
- Albert Einstein Institute, Germany, 200+ hosts
- High Performance Computing Center North, HPC2N, two clusters with a total of 310+ hosts
- Computer-aided chemistry, ETH Zurich, 60 hosts
- Mathematics department, university Paderborn, 120+ clients and servers
- fms-computer.com, Germany, 200-300 hosts in several clusters for customers
- Linux Information Systems AG, 100 hosts

Please fill out the FAI questionnaire !!!

http://www.informatik.uni-koeln.de/fai/questionnaire

Objectives of different installers

d–i

- Be small !!! Be modular. Fit into the RAM!
- Menu driven manual installation of one host
- Ask for language, then ask more questions in this language
- Try to cover common installation (debconf questions)
- Install only base system
- discover1 for hardware detection

FAI

- Infrastructure thinking (multiple hosts)
- Zero keystroke installation! (first plan, then let install)
- Disk space is cheap (nfsroot can contain anything)
- Use classes for grouping
- Central config space and central saving of log files
- Install and configure everything
- Support very different environments
- discover2 for hardware detection

Resent changes

- `fai-cd`
- Debconf support, preseeding
- One developer project -> small team
- First FAI developers workshop in april (very sucessful)
- New action `softupdate` for maintaining running systems
- `linux-fai-devel` mailing list
- Managed to do Ubuntu installation

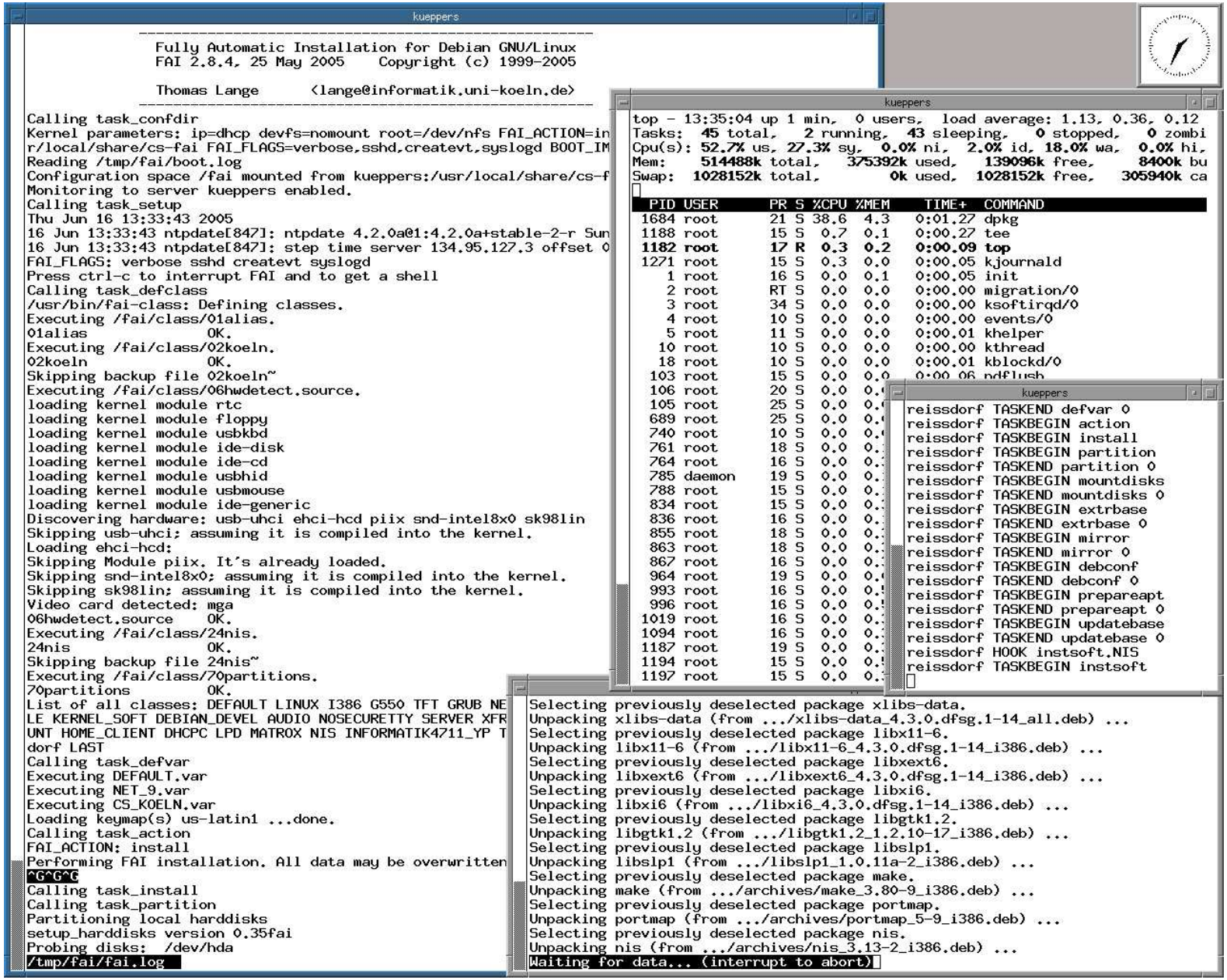
Future plans (FAI 3)

- Replace CVS with subversion (finished soon)
- FAI Wiki ! (ongoing work)
- Split into more packages (doc, server, client)
- Replace fai specific install kernel with default Debian kernel
- Making read-only nfsroot writable with device mapper and ramdisk
- GUI for faimond (ongoing perl/tk work)
- A new disk partition tool (using parted_server)
- LVM and RAID support
- RPM distro support (discussion started, smartpm)
- subversion and arch support (for the config space)
- `fcopy` enhancement

FAI – Summary

- `http://www.informatik.uni-koeln.de/fai`
- Mailing list: `linux-fai@uni-koeln.de` and `linux-fai-devel`
- Quick help on IRC `#fai` on freenode
- CVS access to sources (moving to new system)
- Examples of log files
- Ready to go ISO images for FAI-CD (i386, amd64)
- More than 100 detailed user reports
- FAI runs on i386, amd64, Alpha, IA64, SPARC, PowerPC
- Also installs Solaris on SUN Sparc
- 5 years of FAI
- Users are giving feedback, patches, exchange of experience
- Commercial support: `fai-cluster.de`

And now....


kueppers
Fully Automatic Installation for Debian GNU/Linux
FAI 2.8.4, 25 May 2005 Copyright (c) 1999-2005
Thomas Lange <lange@informatik.uni-koeln.de>
Calling task_confdir
Kernel parameters: ip=dhcp devfs=nomount root=/dev/nfs FAI_ACTION=in
r/local/share/cs-fai FAI_FLAGS=verbose,sshd,createvt,syslogd BOOT_IM
Reading /tmp/fai/boot.log
Configuration space /fai mounted from kueppers:/usr/local/share/cs-f
Monitoring to server kueppers enabled.
Calling task_setup
Thu Jun 16 13:33:43 2005
16 Jun 13:33:43 ntpdate[847]: ntpdate 4.2.0a@1:4.2.0a+stable-2-r Sun
16 Jun 13:33:43 ntpdate[847]: step time server 134.95.127.3 offset 0
FAI_FLAGS: verbose sshd createvt syslogd
Press ctrl-c to interrupt FAI and to get a shell
Calling task_defclass
/usr/bin/fai-class: Defining classes.
Executing /fai/class/01alias.
01alias OK.
Executing /fai/class/02koeln.
02koeln OK.
Skipping backup file 02koeln~
Executing /fai/class/06hwdetect.source.
loading kernel module rtc
loading kernel module floppy
loading kernel module usbkbd
loading kernel module ide-disk
loading kernel module ide-cd
loading kernel module usbhid
loading kernel module usbmouse
loading kernel module ide-generic
Discovering hardware: usb-uhci ehci-hcd piix snd-intel8x0 sk98lin
Skipping usb-uhci; assuming it is compiled into the kernel.
Loading ehci-hcd:
Skipping Module piix. It's already loaded.
Skipping snd-intel8x0; assuming it is compiled into the kernel.
Skipping sk98lin; assuming it is compiled into the kernel.
Video card detected: mga
06hwdetect.source OK.
Executing /fai/class/24nis.
24nis OK.
Skipping backup file 24nis~
Executing /fai/class/70partitions.
70partitions OK.
List of all classes: DEFAULT LINUX I386 G550 TFT GRUB NE
LE KERNEL_SOFT DEBIAN_DEVEL AUDIO NOSECURETTY SERVER XFR
UNT HOME_CLIENT DHCPC LPD MATROX NIS INFORMATIK4711_YP T
dorf LAST
Calling task_defvar
Executing DEFAULT.var
Executing NET_9.var
Executing CS_KOELN.var
Loading keymap(s) us-latin1 ...done.
Calling task_action
FAI_ACTION: install
Performing FAI installation. All data may be overwritten
^G^G^G
Calling task_install
Calling task_partition
Partitioning local harddisks
setup_harddisks version 0.35fai
Probing disks: /dev/hda
/tmp/fai/fai.log
kueppers
top - 13:35:04 up 1 min, 0 users, load average: 1.13, 0.36, 0.12
Tasks: 45 total, 2 running, 43 sleeping, 0 stopped, 0 zombi
Cpu(s): 52.7% us, 27.3% sy, 0.0% ni, 2.0% id, 18.0% wa, 0.0% hi,
Mem: 514488k total, 375392k used, 139096k free, 8400k bu
Swap: 1028152k total, 0k used, 1028152k free, 305940k ca
PID USER PR S %CPU %MEM TIME+ COMMAND
1684 root 21 S 38.6 4.3 0:01.27 dpkg
1188 root 15 S 0.7 0.1 0:00.27 tee
1182 root 17 R 0.3 0.2 0:00.09 top
1271 root 15 S 0.3 0.0 0:00.05 kjournald
1 root 16 S 0.0 0.1 0:00.05 init
2 root RT S 0.0 0.0 0:00.00 migration/0
3 root 34 S 0.0 0.0 0:00.00 ksoftirqd/0
4 root 10 S 0.0 0.0 0:00.00 events/0
5 root 11 S 0.0 0.0 0:00.01 khelper
10 root 10 S 0.0 0.0 0:00.00 kthread
18 root 10 S 0.0 0.0 0:00.01 kblockd/0
103 root 15 S 0.0 0.0 0:00.06 pdflush
106 root 20 S 0.0 0.0
105 root 25 S 0.0 0.0
689 root 25 S 0.0 0.0
740 root 10 S 0.0 0.0
761 root 18 S 0.0 0.0
764 root 16 S 0.0 0.0
785 daemon 19 S 0.0 0.0
788 root 15 S 0.0 0.0
834 root 15 S 0.0 0.0
836 root 16 S 0.0 0.0
855 root 18 S 0.0 0.0
863 root 18 S 0.0 0.0
867 root 16 S 0.0 0.0
964 root 19 S 0.0 0.0
993 root 16 S 0.0 0.0
996 root 16 S 0.0 0.0
1019 root 16 S 0.0 0.0
1094 root 16 S 0.0 0.0
1187 root 19 S 0.0 0.0
1194 root 15 S 0.0 0.0
1197 root 15 S 0.0 0.0
kueppers
reissdorf TASKEND defvar 0
reissdorf TASKBEGIN action
reissdorf TASKBEGIN install
reissdorf TASKBEGIN partition
reissdorf TASKEND partition 0
reissdorf TASKBEGIN mountdisks
reissdorf TASKEND mountdisks 0
reissdorf TASKBEGIN extrbase
reissdorf TASKEND extrbase 0
reissdorf TASKBEGIN mirror
reissdorf TASKEND mirror 0
reissdorf TASKBEGIN debconf
reissdorf TASKEND debconf 0
reissdorf TASKBEGIN prepareapt
reissdorf TASKEND prepareapt 0
reissdorf TASKBEGIN updatebase
reissdorf TASKEND updatebase 0
reissdorf HOOK instsoft.NIS
reissdorf TASKBEGIN instsoft
Selecting previously deselected package xlibs-data.
Unpacking xlibs-data (from .../xlibs-data_4.3.0.dfsg.1-14_all.deb) ...
Selecting previously deselected package libx11-6.
Unpacking libx11-6 (from .../libx11-6_4.3.0.dfsg.1-14_i386.deb) ...
Selecting previously deselected package libxext6.
Unpacking libxext6 (from .../libxext6_4.3.0.dfsg.1-14_i386.deb) ...
Selecting previously deselected package libxi6.
Unpacking libxi6 (from .../libxi6_4.3.0.dfsg.1-14_i386.deb) ...
Selecting previously deselected package libgtk1.2.
Unpacking libgtk1.2 (from .../libgtk1.2_1.2.10-17_i386.deb) ...
Selecting previously deselected package libslp1.
Unpacking libslp1 (from .../libslp1_1.0.11a-2_i386.deb) ...
Selecting previously deselected package make.
Unpacking make (from .../archives/make_3.80-9_i386.deb) ...
Selecting previously deselected package portmap.
Unpacking portmap (from .../archives/portmap_5-9_i386.deb) ...
Selecting previously deselected package nis.
Unpacking nis (from .../archives/nis_3.13-2_i386.deb) ...
Waiting for data... (interrupt to abort)