

Fully automatic Linux installations

SUCON04.

Thomas Lange, University of Cologne

`lange@informatik.uni-koeln.de`

- ▷ Motivation
- ▷ How to plan a computer infrastructure
- ▷ How does FAI work?
- ▷ The show

Administration - some common problems?

- ▷ Buy a computer, install it once, run it forever
- ▷ Grown systems, very different hardware and software
- ▷ Many small changes in between on every single system
- ▷ Only emergency administration because of no time
- ▷ We have the fastest computers, but we do much manual work
- ▷ Biggest problem of system administration: **Too much manual work**
- ▷ Too much manual work => no time
- ▷ No time => no time for automating things

Value of your computer

- ▷ What are the values of your computer?
- ▷ What are you doing if your computers are not running?
- ▷ A good computer infrastructure is as important as ...?
- ▷ Which things are included your computers?
 - > Customer data (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?

- ▶ 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?

- ▷ Good running computers are essential
- ▷ Data backup is only one part of it all
- ▷ You should also backup the sysadmin work
- ▷ Manual installations are bad installations
- ▷ Bad installation -> unusable computers
- ▷ Bad installation -> unproductive users
- ▷ Bad installation -> manual rework -> no time
- ▷ No updates, no patches, no security!!!
- ▷ Manual sysadmin work -> high IT costs

Manual installation?

Who likes to install these hosts by hand?



Manual installation?

Can you guarantee, that all these hosts are equal?



- ▷ "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)
- ▶ Diversity of hardware and different configurations easily manageable
- ▶ Cluster, server farm, labs and pools are perfect
- ▶ One command – hundreds of installations
- ▶ Junior admin can use FAI
- ▶ You can save much work! (work = time = money)
- ▶ Do you have a plan for your computer infrastructure?

- ▶ 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
- ▶ It installs and configures the whole OS and all applications
- ▶ No master or golden image needed
- ▶ It's very modular because of its class system
- ▶ It's flexible and easy to expand with hooks
- ▶ It's neither a cluster management tool nor a job scheduling system
- ▶ 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
- ▷ Needs time
- ▷ `www.infrastructures.org`
- ▷ Discover your actual state
- ▷ What would you like to change in the future?
- ▷ Bear in mind future extensions
- ▷ System administration is a process, not a product
- ▷ FAI can't do things, you don't tell it to do
- ▷ Put your infrastructure data into version control (CVS)
- ▷ Which things are equal, which are different?
- ▷ One data, one source

Questions for an infrastructure

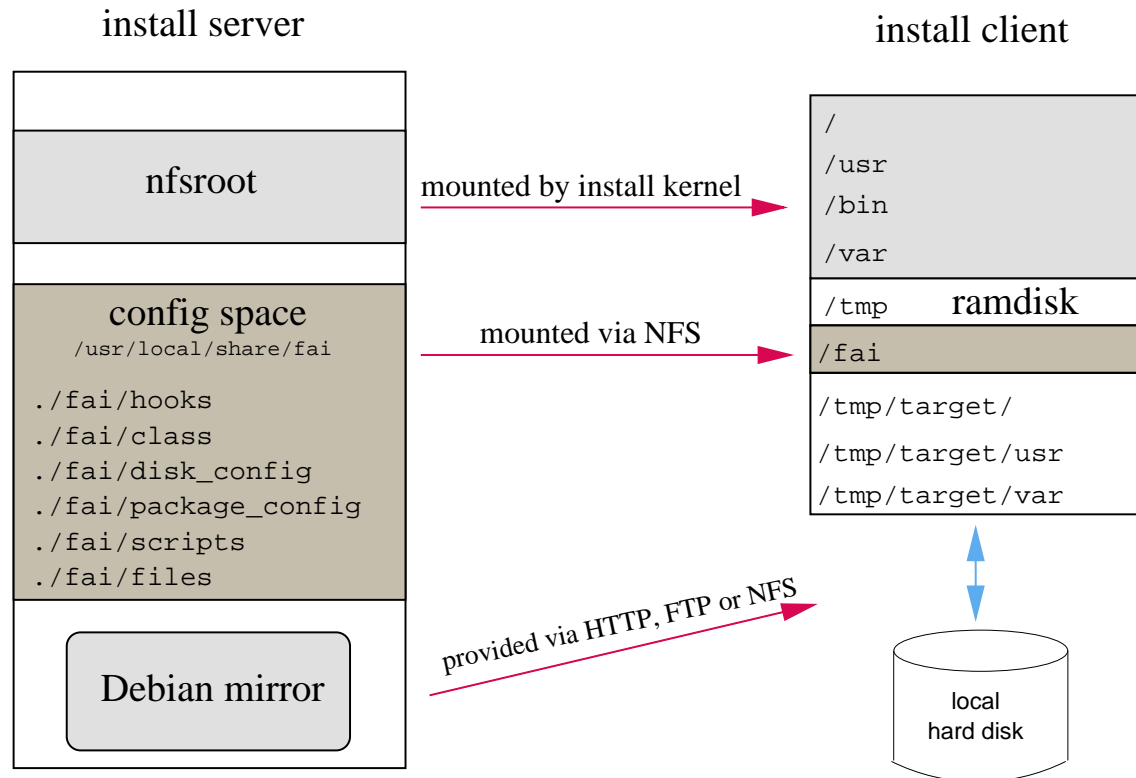
- ▶ Which type of computer will be installed? Cluster, desktop, server, notebook?
- ▶ Which jobs do the computers have? CAD, server, text processing
- ▶ Which applications will be run on them?
- ▶ How does my LAN topology looks like? Is DHCP available?
- ▶ Do I have uniform hardware? In the future?
- ▶ 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 ?



A system administrator during a fully automatic installation

How does FAI work ?



- Configuration is stored on the install server
- Installation runs on the client
- Full remote control during installation via ssh

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	230 MB	created with <code>make-fai-nfsroot</code>
Debian mirror	9.7 GB	Debian 3.1 (only for i386)
- ▶ 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 fai script `rcS_fai`, which controls the installation
- ▶ Detect hardware and load kernel modules
- ▶ Define classes and variables
- ▶ Partition local hard disk, and create file systems and mount them
- ▶ Install software packages
- ▶ Configure operating systems and applications
- ▶ Save log files to the local disk and to the install server
- ▶ Boot the newly installed system

- ▷ A host belongs to several classes
- ▷ Examples: `DEFAULT SMALL_IDE 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
- ▷ Junior admin installs the computers
- ~~▷ Junior admin installs the computers~~
- ▷ The computer installs itself automatically ;-)

Directory tree of the config space

```
|-- class/
|   |-- 01alias
|   |-- 06hwdetect.source
|   |-- 24nis
|   |-- ATOMCLIENT.var
|   |-- DEFAULT.var
|   `-- demo
|-- disk_config/
|   |-- ATOMCLIENT
|   |-- SMALL_IDE
|   `-- foobar04
|-- package_config/
|   |-- BOWULF
|   |-- DEBIAN_DEVEL
|   |-- DEMO
|   |-- GERMAN
|   |-- GNOME
|   `-- nucleus
```

Directory tree of the config space

```
|-- files/
|   |-- etc/
|       |-- X11/
|           |-- XF86Config-4/                    fcopy /etc/X11/XF86Config-4
|               |-- ATI_ACER
|               |-- MATROX
|               |-- demohost
|           |-- nsswitch.conf/                  fcopy /etc/nsswitch.conf
|               |-- NIS
|               |-- NONIS
|-- scripts/
    |-- BOOT
    |-- DEFAULT/
    |   |-- S01
    |   |-- S21
    |   |-- demohost
    |-- NETWORK/
    |   |-- S10
    |   |-- S40
    |-- demohost
```

Example /fai/class/ATOMCLIENT.var:

```
FAI_KEYMAP=de-latin1-noadkeys
```

```
UTC=yes
```

```
time_zone=Europe/Berlin
```

```
addpackages="kernel-image-2.4-386"
```

```
# root password for the new installed linux system;
```

```
rootpw="1234NVqX514t0f"
```

```
# modules that will be loaded by the new system, not during installation
```

```
# these modules will be written to /etc/modules
```

```
moduleslist="e100 3c59x usbkbd usb-uhci keybdev mousedev hid"
```

```
# Beowulf: define NFS server for /home and /usr/local
```

```
hserver=atom00
```

```
bserver=atom00
```

- ▶ All configuration scripts in /fai/scripts/* are using these variables

Example: /fai/disk_config/SMALL_IDE:

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

```
disk_config hda
```

```
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
```

Installation of software package

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` and additional `taskinst`, `taskrm`
- ▶ Dependencies are resolved
- ▶ `dpkg -get-selections` also possible

Config 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 $YPPDOMAIN > $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	Disk	Software	Time
Pentium 4 2.80GHz	1024MB	IDE	948 MB	5 min
Athlon XP1600+	896MB	SCSI	1 GB	6 min
AMD-K7, 500MHz	320MB	IDE	780 MB	12 min
PentiumPro 200MHz	128MB	IDE	800 MB	28 min
Pentium III 850MHz	256MB	IDE	820 MB	10 min
Pentium III 850MHz	256MB	IDE	180 MB	3 min

Parallel installation of multiple nodes in a Beowulf cluster:

Nodes	Seconds
1	337
5	340
10	345
20	379

12% more time with 20 hosts

- ▷ Electricité de France (EDF), France, 200 hosts
- ▷ Deutsches Elektronen-Synchrotron, DESY, 60+
- ▷ Danmarks Meteorologiske Institut, 85+ hosts
- ▷ IFW-Dresden, Germany, 100+ hosts, Cluster
- ▷ 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+
- ▷ A company, electronic payment management and IDS
- ▷ A company, search engine, 200+
- ▷ HPC2N, 120 nodes, dual Athlon MP2000+, listed in top500.org
- ▷ Institut für Physik, Universität Augsburg, 80+
- ▷ Computer-aided chemistry (IGC), ETH Zürich, cluster, 45 dual nodes
- ▷ University Rekenentrum Rijksuniversiteit Groningen, 40+ workstations and servers
- ▷ fms-computer.com, Germany, 200-300 hosts in several clusters for customers
- ▷ more detailed reports on the FAI web page

- ▷ `http://www.informatik.uni-koeln.de/fai`
- ▷ Mailing list: `linux-fai@uni-koeln.de`
- ▷ CVS access to sources
- ▷ Examples of log files
- ▷ FAI runs on i386, amd64, 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....

```
-----
Fully Automatic Installation for Debian GNU/Linux
FAI 2.6.2, 26 aug 2004    Copyright (c) 1999-2004

Thomas Lange    <lange@informatik.uni-koeln.de>
-----

Calling task_confdir
Kernel parameters: ip=dhcp devfs=nomount FAI_ACTION=install root=/dev/
evt.syslogd BOOT_IMAGE=vmlinuz-install
Reading /tmp/fai/boot.log
Configuration space /fai mounted from kueppers:/usr/local/share/fai
Monitoring to server kueppers enabled.
Calling task_setup
Mon Aug 30 11:55:37 2004
30 Aug 11:55:37 ntpdate[831]: ntpdate 4.2.0a@1:4.2.0a-11-r Tue Jul 27 04
30 Aug 11:55:38 ntpdate[831]: step time server 134.95.176.126 offset 0.8
FAI_FLAGS: verbose sshd createtv 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/06hwdetect.source.
Discovering hardware: es1371 agpgart usb-uhci BusLogic pcnet32
Loading es1371:
Loading agpgart:
Loading usb-uhci:
Loading BusLogic:
Skipping pcnet32: assuming it is compiled into the kernel.
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 keybdev
loading kernel module mousedev
loading kernel module hid
loading kernel module ide-detect
loading kernel module isa-pnp
Video card detected: vmware
06hwdetect.source OK.
Executing /fai/class/70partitions.
70partitions OK.
List of all classes: DEFAULT LINUX I386 SMALL_IDE MBR GRUB NETWORK DHCP
SCRATCH demohost LAST
Calling task_defvar
Executing DEFAULT.var
Loading /usr/share/keymaps/i386/qwerty/us-latin1.kmap.gz
Calling task_action
FAI_ACTION: install
Performing FAI installation. All data may be overwritten
^C^C^C
Calling task_install
Calling task_partition
Partitioning local harddisks
Calling task_mountdisks
Enable swap device /dev/hda5
Mounting /dev/hda1 to /tmp/target/
Mounting /dev/hda9 to /tmp/target/home
Mounting /dev/hda10 to /tmp/target/scratch
Mounting /dev/hda7 to /tmp/target/tmp
Mounting /dev/hda8 to /tmp/target/usr
Mounting /dev/hda6 to /tmp/target/var
Calling task_extrbase
Unpacking Debian base archive
/tmp/fai/rc5.Log █

-----
top - 11:57:24 up 1 min, 0 users, load average: 1.11, 0.41, 0.
Tasks: 42 total, 3 running, 39 sleeping, 0 stopped, 0 zo
Cpu(s): 13.3% user, 85.3% system, 0.0% nice, 1.3% idle
Mem: 357072k total, 352352k used, 4720k free, 15912k
Swap: 159728k total, 144k used, 159584k free, 260748k

  PID USER      PR  NI  S %CPU %MEM    TIME+  COMMAND
  1153 root        9   0  S 22.3  5.4    0:00.77  dpkg
  328 root        9   0  S  8.0  0.0    0:00.42  kjournald
  343 root        9   0  S  7.0  0.0    0:01.68  kjournald
 1182 root       19   0  R  6.3  4.5    0:00.19  dpkg-query
  340 root        9   0  S  3.3  0.0    0:00.27  kjournald
 1174 root        9   0  S  1.3  0.6    0:00.04  kernel-image-2.
 1177 root        9   0  S  0.7  0.4    0:00.02  mkinitrd
   4 root        9   0  S  0.3  0.0    0:00.01  kswapd
   6 root        9   0  S  0.3  0.0    0:00.05  kupdated
  310 root        9   0  R  0.3  0.3    0:00.20  top
   1 root        9   0  S  0.0  0.1    0:01.58  init
   2 root        9   0  S  0.0  0.0    0:00.00  keventd
   3 root       19  19  S  0.0  0.0    0:00.00  ksoftirqd_CPU0
   5 root
   8 root
   9 root
  10 root
  29 daemon
  32 root
  71 root
  73 root
  86 root
  88 root
  90 root
 148 root
 179 root
 191 root
 293 root
 294 root
 331 root
 334 root
 337 root
 350 root
 354 root
 374 root

-----
kueppers
demohost TASKEND setup 0
demohost TASKBEGIN defclass
demohost TASKEND defclass 0
demohost TASKBEGIN defvar
demohost TASKEND defvar 0
demohost TASKBEGIN action
demohost TASKBEGIN install
demohost TASKBEGIN partition
demohost TASKEND partition 0
demohost TASKBEGIN mountdisks
demohost TASKEND mountdisks 0
demohost TASKBEGIN extrbase
demohost TASKEND extrbase 0
demohost TASKBEGIN mirror
demohost TASKEND mirror 0
demohost TASKBEGIN updatebase
demohost TASKEND updatebase 0
demohost TASKBEGIN instsoft
demohost TASKEND instsoft 0
demohost TASKBEGIN configure

-----
Selecting previously deselected package nscd.
Unpacking nscd (from ../nscd_2.3.2.ds1-13_i386.deb) ...
Selecting previously deselected package rdate.
Unpacking rdate (from ../r/rdate/rdate_1.4-5_i386.deb) ...
Selecting previously deselected package rstat-client.
Unpacking rstat-client (from ../rstat-client_3.07-3_i386.deb) ...
Selecting previously deselected package rstatd.
Unpacking rstatd (from ../rstatd/rstatd_3.07-3_i386.deb) ...
Selecting previously deselected package rsync.
Unpacking rsync (from ../r/rsync/rsync_2.6.2-3_i386.deb) ...
Selecting previously deselected package rusers.
Unpacking rusers (from ../rusers_0.17-5_i386.deb) ...
Selecting previously deselected package rusersd.
Unpacking rusersd (from ../rusersd_0.17-5_i386.deb) ...
Selecting previously deselected package sysutils.
Unpacking sysutils (from ../sysutils_1.3.8.5.1_i386.deb) ...

(END) █
```