Fully automatic Linux installations

FOSDEM 2005.

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
- ▶ FAI proudly presents ...
- ▶ 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

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

▶ What are the values 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 ...?

- 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)

- 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?

The Test

- Grab a random machine (without a backup before)
- > or dd if=/dev/zero of=/dev/hda

The Test

- Grab a random machine (without a backup before)
- > or dd if=/dev/zero of=/dev/hda



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?

Facts, often overseen

- Good running computers are essential
- Data backup is only one part of it all

Facts, often overseen

- 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?



20 nodes dual XEON, 2.4 GHz







www.centibots.org

Can you guarantee, that all these hosts are equal?



Manual installation

➣ "No simple sysadmin task is fun more than twice"

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)
- 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?

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
- 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

- 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

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

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,...?

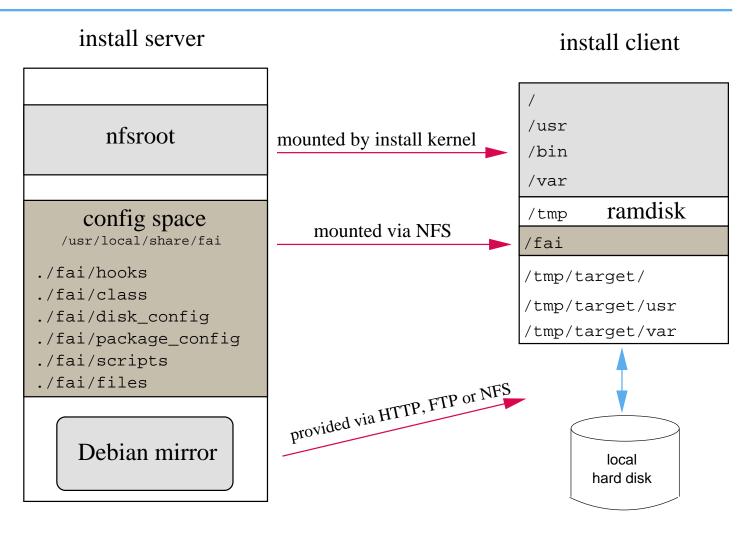
How does FAI work?

How does FAI work?



A system administrator during a fully automatic installation

How does FAI work?



- Configuration is stored on the install server (one tree for all clients)
- 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 make-fai-nfsroot
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!

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 (discover2)
- Define classes and variables
- Partition local hard disk, and create file systems and mount them (setup_harddisks and sfdisk)
- Install software packages (install_packages using apt-get and aptitude)
- 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 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

The class concept

- 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

The class concept

- 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
- The computer installs itself automaticly ;-)

Directory tree of the config space

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

Directory tree of the config space

```
-- files/
   |-- etc/
       |-- X11/
       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-nodeadkeys
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

Disk partitioning

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
                              ; -m 5 -j ext3
                         rw
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 also taskinst, taskrm
- Dependencies are resolved
- dpkg -get-selections also possible
- aptitude may replace apt-get

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 $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	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

FAI users

- Electricité de France (EDF), France, 200 hosts
- MIT Computer science research lab, 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+
- Host Europe, 250 hosts
- Lycos Europe, search engine, 200+
- Physical HPC2N, 2 clusters listed in top500.org, 192 dual Opteron, 120 dual Athlon
- Physics department, university Augsburg, 80+
- Computer-aided chemistry, ETH Zürich, cluster, 45 dual Athlon nodes, 15 single nodes
- Mathematics department, university Paderborn, 120+ clients and servers
- > fms-computer.com, Germany, 200-300 hosts in several clusters for customers

Please fill out the 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 hw detection

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 hw 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 are central saving of log files
 - Install and configure everything
 - Support all different configurations (fcopy, scripts)
 - discover2 for hw detection

Proudly to announce FAI-CD

- ➤ The first official version of fai-cd (get one right now)
- Rewrite of Nial Young's fai-bootCD scripts
- Puts all fai parts onto a bootable CD
- fai boot kernel, nfsroot, config space and partitial mirror
- No network services need to be set up
- Quick and easy way to make an installation
- CD is a nice giveaway for shows and conferences
- faimirror(1), creates a partitial mirror with all packages used by FAI package_config files
- fai-cd script itself is not yet available, but the ISO image is
- www.informatik.uni-koeln.de/fai/fai-cd

Future features of FAI (FFOF)

- Multi-Tier roadmap
- EnterprisePro Edition Version 11
- Install detection and protection
- Java.NET centric config manager
- Grid enabled Middleware
- ► FAI SDK for .COM, .NET, ruby und Groupware
- Zero administation Enterprise enabling services
- Turn key values added transition for proactive Sigfried and ROI
- Consolidation of DMS, CRM, ERP, OO, B2B and XYZ

The future of FAI (aka FAI 3.x)

- Finish fai-cd script (I need more feedback)
- ▶ New action update for maintaining running systems
- GUI for faimond (ongoing perl/tk work)
- A new disk partition tool (a plan since many years ;-, parted_server)
- LVM and RAID support (many requests)
- Debconf support, preseeding (will be very easy)
- Henning Sprang did Ubuntu installation with FAI
- RH and Suse proof of concept (or use FAI classes with kickstart)
- FAI Wiki ?! (Who likes to help me)
- First fai developers workshop in april (held in german)

- http://www.informatik.uni-koeln.de/fai
- ▶ Mailing list: linux-fai@uni-koeln.de
- CVS access to sources
- Examples of log files
- More than 80 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....

