

FAI Guide (Fully Automatic Installation)

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

Version 2.1, 2 apr 2003 for FAI version 2.4.1

Abstract

This manual describes the fully automatic installation package for Debian GNU/Linux. This includes the installation of the package, the planning and creating of the configuration and how to deal with errors.

Copyright Notice

Copyright © 2000-2003 Thomas Lange

This manual is free software; you may redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2, or (at your option) any later version.

This is distributed in the hope that it will be useful, but *without any warranty*; without even the implied warranty of merchantability or fitness for a particular purpose. See the GNU General Public License for more details.

A copy of the GNU General Public License is available as `/usr/share/common-licenses/GPL` in the Debian GNU/Linux distribution or on the World Wide Web at the GNU website (<http://www.gnu.org/copyleft/gpl.html>). You can also obtain it by writing to the Free Software Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.

Contents

1	Introduction	1
1.1	Availability	1
1.2	Motivation	1
1.3	Overview and concepts	2
1.4	How does FAI work?	3
1.5	Features	3
2	Installing FAI	5
2.1	Requirements	5
2.2	How to create a local Debian mirror	6
2.3	Setting up FAI	6
2.3.1	Troubleshooting the setup	9
3	Preparing booting	11
3.1	Booting from 3Com network card with boot PROM	11
3.2	Booting from network card with a PXE conforming boot ROM	12
3.3	Creating a boot floppy	12
3.4	Booting from a CD-ROM	13
3.5	Collecting Ethernet addresses	13
3.6	Configuration of the BOOTP daemon	13
3.6.1	Troubleshooting BOOTP daemon	15
3.7	Configuration of the DHCP daemon	15
3.8	Boot messages	16
3.9	Troubleshooting boot messages	17

3.10	Collecting other system information	19
3.11	Checking parameters from BOOTP and DHCP servers	20
3.12	Rebooting the computer	20
4	Overview of the installation sequence	21
4.1	Set up FAI	22
4.2	Defining classes, variables and loading kernel modules	22
4.3	Partitioning local disks, creating filesystems	23
4.4	Installing software packages	23
4.5	Site specific configuration	23
4.6	Save log files	24
4.7	Reboot the new installed system	24
5	How to plan your installation	25
6	Installation details	27
6.1	The configuration space	27
6.2	The default tasks	28
6.3	The setup routines of the install clients	30
6.4	The class concept	32
6.5	Defining classes	33
6.6	Defining Variables	35
6.7	Hard disk configuration	36
6.8	Software package configuration	36
6.9	Scripts in <code>/fai/scripts</code>	38
6.9.1	Shell scripts	38
6.9.2	Perl scripts	38
6.9.3	Expect scripts	38
6.9.4	Cfengine scripts	38
6.10	Changing the boot device	39
6.11	Hooks	40
6.12	Looking for errors	41

7	How to build a Beowulf cluster using FAI	43
7.1	Planning the Beowulf setup	43
7.2	Set up the master server	44
7.2.1	Set up the network	44
7.2.2	Setting up NIS	45
7.2.3	Create a local Debian mirror	45
7.2.4	Install FAI package on the master server	45
7.2.5	Prepare network booting	46
7.3	Tools for Beowulf clusters	47
7.4	Wake on LAN with 3Com network cards	47
8	FAI on SUN SPARC hardware running Linux	49
9	Various hints	51
9.1	Useful functions for advanced administrators	53

Chapter 1

Introduction

1.1 Availability

The homepage of FAI is <http://www.informatik.uni-koeln.de/fai>. There you will find any information about FAI, for example the mailing list archive. The FAI package is also available as a Debian package from www.informatik.uni-koeln.de/fai/download. It's an official Debian package and is available from all Debian mirrors. Send any bug or comment to fai@informatik.uni-koeln.de. You can also use the Debian bug tracking system (BTS) <http://www.debian.org/Bugs> for reporting errors.

You can access the CVS repository containing the newest developer version of FAI from a Bourne shell using the following commands. The login password is empty, so only press return.

```
> CVSROOT=:pserver:anonymous@cvs.debian.org:/cvs/debian-boot
> cvs login
> cvs co -P fai-kernels
> cvs co -P fai
```

You can also use the web interface for the CVS repository at: <http://cvs.debian.org/fai/> (and [fai-kernels](http://cvs.debian.org/fai/fai-kernels)).

Now read this manual, then enjoy the fully automatic installation and your saved time.

1.2 Motivation

Have you ever performed identical installations of an operating system several times? Would you like to be able to install a Linux cluster with dozens of nodes single handedly?

Repeating the same task time and again is boring – and will surely lead to mistakes. Also a whole lot of time could be saved if the installation were done automatically. An installation

process with manual interaction does not scale. But clusters have the habit of growing over the years. Think long-term rather than plan only just a few months into the future.

In 1999, I had to organize an installation of a Linux cluster with one server and 16 clients. Since I had much experience doing automatic installation of Solaris operating system on SUN SPARC hardware, the idea to build an automatic installation for Debian was born. Solaris has an automatic installation feature called JumpStart¹. In conjunction with the auto-install scripts from Casper Dik², I could save a lot of time not only for every new SUN computer, but also for re-installation of existing workstations. For example, I had to build a temporary LAN with four SUN workstations for a conference, which lasted only a few days. I took these workstations out of our normal research network and set up a new installation for the conference. When it was over, I simply integrated the workstation back into the research network, rebooted just once, and after half an hour, everything was up and running as before. The configuration of all workstations was exactly the same as before the conference, because everything was performed by the same installation process. I also use the automatic installation for reinstalling a workstation after a damaged hard disk had been replaced. It took two weeks until I received the new hard disk but only a few minutes after the new disk was installed, the workstation was running as before. And this is why I chose to adapt this technique to a PC cluster running Linux.

1.3 Overview and concepts

FAI is a non-interactive system to install a Debian GNU/Linux operating system unattended on a single computer or a whole cluster. You can take one or more virgin PC's, turn on the power and after a few minutes Linux is installed, configured and running on the whole cluster, without any interaction necessary. Thus, it's a scalable method for installing and updating a cluster unattended with little effort involved. FAI uses the Debian GNU/Linux distribution and a collection of shell and perl scripts for the installation process. Changes to the configuration files of the operating system can be made by cfengine, shell, perl and expect scripts.

FAI's target group are system administrators how have to install Debian onto one or even hundreds of computers. Because it's a general purpose installation tool, it can be used for installing a Beowulf cluster, a rendering farm or a linux laboratory or a classroom. Also large-scale linux networks with different hardware or different installation requirements are easy to establish using FAI. But don't forget to plan your installation. 'How to plan your installation' on page 25 has some useful hints for this topic.

First, some terms used in this manual are described.

install server : The host where the package FAI is installed. It provides several services and data for all install clients. In the examples of this manual this host is called `kueppers`.

install client : A host which will be installed using FAI and a configuration from the install server. Also called client for short. In this manual, the example hosts are called `bigfoot`, `ant01`, `ant02`, `nucleus`, `atom01`, `atom02`,...

¹Solaris 8 Advanced Installation Guide at docs.sun.com

²<ftp://ftp.wins.uva.nl/pub/solaris/auto-install/>

configuration : The details of how the installation of the clients should be performed. This includes information about:

- Hard disk layout
- Local filesystems, their mount points and mount options
- Software packages
- Keyboard layout, time zone, NIS, XFree86 configuration, remote filesystems, user accounts, printers ...

nfsroot : A filesystem located on the install server. It's the complete filesystem for the install clients during the installation process. All clients share the same nfsroot, which they mount read only.

1.4 How does FAI work?

The install client which will be installed using FAI, is booted from floppy disk or via network card. It gets an IP address and boots a linux kernel which mounts its root filesystem via NFS from the install server. After the operating system is running, the FAI startup script performs the automatic installation which doesn't need any interaction. First, the hard disks will be partitioned, filesystems are created and then software packages are installed. After that, the new installed operating system is configured to your local needs using some scripts. Finally the new operating system will be booted from the local disk.

The details, of how to install the computer (the configuration), are stored in the configuration space on the install server. Configuration files are shared among groups of computers if they are similar using the class concept. So you need not to create a configuration for every new host. Hence, FAI is a scalable method to install a big cluster with a great number of nodes.

FAI can also be used as an network rescue system. You can boot your computer, but it will not perform an installation. Instead it will run a fully functional Debian GNU/Linux without using the local hard disks. Then you can do a remote login and backup or restore a disk partition, check a filesystem, inspect the hardware or do any other task.

1.5 Features

- A fully automated installation can be performed
- Very quick unattended installation
- Hosts can boot from floppy or from network card
- Easy creation of the common boot floppy which uses grub or lilo
- BOOTP and DHCP protocol and PXE boot method are supported

-
- No initial ramdisk is needed, 8MB RAM suffice
 - Runs even on a 386 CPU
 - The installation kernel can use modules
 - Remote login via ssh during installation process possible
 - Two additional virtual terminals available during installation
 - All similar configuration are shared among all install clients
 - Log files for all installations are saved on to the installation server
 - Shell, perl, expect and cfengine scripts are supported for the configuration setup
 - Access to a Debian mirror via NFS, FTP or HTTP
 - Keyboard layout selectable
 - Can be used as a rescue system
 - Tested on SUN SPARC hardware running Linux or Solaris
 - Flexible system through easy class concept
 - Predefined Beowulf classes included
 - Diskless client support
 - Easily add your own functions via hooks
 - Easily change the default behavior via hooks
 - Lilo and grub support
 - ReiserFS and ext3 support
 - Automatic hardware detection
 - Booting and installing from CD-ROM is in progress

Chapter 2

Installing FAI

2.1 Requirements

The following items are required for an installation via FAI.

A computer: The computer must have a network interface card. Unless a diskless installation should be performed a local hard disk is also needed. No floppy disk, CD-ROM, keyboard or graphic card is needed.

BOOTP or DHCP server: The clients need one of these daemons to obtain boot information. But you can also put all this information onto the boot floppy.

TFTP server: The TFTP daemon is used for transferring the kernel to the clients. It's only needed when booting from network card with a boot PROM.

Client root: It is a mountable directory which contains the whole filesystem for the install clients during installation. It will be created during the setup of the FAI package and is also called **nfsroot**.

Debian mirror: Access to a Debian mirror is needed. A local mirror of all Debian packages is recommended if you install several computers.

Install kernel: A kernel image that supports the network card and mounts its root filesystem via NFS. The Debian package `fai-kernels` provides a default kernel for `fai`.

Configuration space: This directory tree which contains the configuration data is a mounted via NFS by default. But you can also get this directory from a revision control system like CVS.

The TFTP daemon and a NFS server will be enabled automatically when installing the FAI package. All clients must have a network card which is recognized by the install kernel.

2.2 How to create a local Debian mirror

The script `mkdebmirror`¹ can be used for creating your own local Debian mirror. This script uses the script `debmirror`² and `rsync(1)`. A partial Debian mirror only for i386 architecture for Debian 3.0 (aka woody) without the source packages needs about 5.0GB of disk space. Accessing the mirror via NFS will be the normal and fastest way in most cases. To use HTTP access to the local Debian mirror, install the webserver software and create a symlink to the local directory where you mirror is located:

```
# apt-get install apache
# ln -s /files/scratch/debmirror /var/www/debmirror
```

Don't forget to adjust the variable `FAI_SOURCES_LIST` in `/etc/fai/fai.conf` to access the Debian mirror.

2.3 Setting up FAI

Before installing FAI, you have to install the package `fai-kernels`, which contains the install kernels for FAI. You can install both packages using

```
kueppers[~]# apt-get install fai fai-kernels
Reading Package Lists... Done
Building Dependency Tree... Done
The following NEW packages will be installed:
  fai fai-kernels
0 packages upgraded, 2 newly installed, 0 to remove and 1 not upgraded.
Need to get 0B/12.7MB of archives. After unpacking 13.9MB will be used.
Selecting previously deselected package fai.
(Reading database ... 48317 files and directories currently installed.)
Unpacking fai (from .../main/f/fai/fai_2.4.1_all.deb) ...
Selecting previously deselected package fai-kernels.
Unpacking fai-kernels (from .../fai-kernels_1.5_i386.deb) ...
Setting up fai (2.4.1) ...
To set up FAI, edit /etc/fai/fai.conf and call fai-setup.

Setting up fai-kernels (1.5) ...
```

You can also get the newest version of `fai` and `fai-kernels` from the download page of `fai` and install the packages using the `dpkg` command.

The configuration for the FAI package (not the configuration data for the install clients) are defined in `/etc/fai/fai.conf`. Since FAI doesn't use `debconf` yet, edit this file before calling `fai-setup`. These are important variables in `/etc/fai/fai.conf`:

¹You can find the script in `/usr/share/fai/utils/`.

²Available as a Debian package or at the FAI homepage.

FAI_DEBOOTSTRAP For building the nfsroot there's the command called `debootstrap(8)`. It needs the location of a Debian mirror and the name of the distribution (woody,sarge,sid) for which the basic Debian system should be built.

FAI_SOURCES_LIST This multi line string is the content of `sources.list` (used by `apt-get(8)`); it defines the location and access method for the Debian mirror. If this variable is undefined, the file `/etc/fai/sources.list` or `/etc/apt/sources.list` will be used. For more information on the file format see `sources.list(5)`.

FAI_DEBMIRROR If you have NFS access to your local Debian mirror, specify the remote filesystem. It will be mounted to `$MNTPOINT`, which must also be defined. It's not needed if you use access via FTP or HTTP.

KERNELPACKAGE You must specify the software package - build with `make-kpkg(8)` - which includes the default kernel for booting the install clients. The Debian package `fai-kernels` contains the default install kernels which supports both the BOOTP and DHCP protocol.

NFSROOT_PACKAGES This variable contains a list of additional software packages which will be added to the nfsroot.

FAI_LOCATION This is the host name and the remote directory of the configuration space, which will be mounted via NFS.

FAI_BOOT which of DHCP and/or BOOTP should the server create setups for (when `make-fai-nfsroot` is run). Default are to create setups for both.

The variables `FAI_SOURCES_LIST` and `FAI_DEBMIRROR` are used by the install server and also by the clients. If your install server has multiple network card and different host names for each card (as for a Beowulf server), use the install server name which is known by the install clients.

FAI uses `apt-get(8)` to create the nfsroot filesystem in `/usr/lib/fai/nfsroot`. It needs about 160MB of free disk space. Before setting up FAI, you should get the program `imggen`,³ if you like to boot from a 3Com network card. This executable converts netboot images created by `mknbi-linux(8)`, so they can be booted by network cards from 3Com. Put that executable in your path (e.g. `/usr/local/bin`). After editing `/etc/fai/fai.conf` call `fai-setup`.

```
kueppers[~]# fai-setup
Account $LOGUSER=fai already exists.
Make sure that all install clients can
log into this account without a password.
Using interface eth0 to determine local hostname.
Adding kueppers to known_hosts.
/home/fai/.ssh/known_hosts created.
```

³Available at the download page <http://www.ltsp.org> or from the FAI download page www.informatik.uni-koeln.de/fai/download.

```
/home/fai/.ssh/authorized_keys created.
User account fai set up.
Creating FAI nfsroot can take a long time and will
need more than 160MB disk space in /usr/lib/fai/nfsroot.
/usr/lib/fai/nfsroot already exists. Removing /usr/lib/fai/nfsroot
Creating nfsroot for woody using debootstrap
dpkg: base-passwd: dependency problems, but configuring anyway as you request:
 base-passwd depends on libc6 (>= 2.2.4-4); however:
   Package libc6 is not installed.
dpkg: base-files: dependency problems, but configuring anyway as you request:
.
.
.
Automatically converting /etc/network/interfaces succeeded.
Old interfaces file saved as interfaces.dpkg-old.
Creating base.tgz
Upgrading /usr/lib/fai/nfsroot
Adding additional packages to /usr/lib/fai/nfsroot:
portmap file rdate cfengine bootpc wget rsh-client less dump
ext2resize strace hdparm parted dnsutils grub ntpdate
dosfstools sysutils dialog libdetect0 discover mdetect read-edid kudzu hwtool
Detecting hardware: 3c59x ide-scsi usb-uhci usb-uhci
modprobe: Can't open dependencies file /lib/modules/2.4.20/modules.dep (No su
Skipping 3c59x; assuming it is compiled into the kernel.
modprobe: Can't open dependencies file /lib/modules/2.4.20/modules.dep (No su
Skipping usb-uhci; assuming it is compiled into the kernel.
Creating SSH2 RSA key
Creating SSH2 DSA key
Restarting OpenBSD Secure Shell server: sshd.
DHCP environment prepared. Now enable dhcpd and the special tftpd daemon
Kernel image file name = /usr/lib/fai/nfsroot/boot/vmlinuz-2.4.20
Output file name       = /boot/fai/installimage
Kernel command line   = "auto rw root=/dev/nfs nfsroot=kernel nfsaddrs=kern

Image Creator for MBA ROMs v1.01, Date: Nov 26, 2001
Design and Coding by Nick Kroupetski <NickKroupetski@hotmail.com>
Usage: imggen [OPTION] inputfile outputfile
  -a,   Add 3Com MBA/BootWare support
  -r,   Remove 3Com MBA/BootWare support from image file
  -i,   Show information on an image
  -h,   Help screen

In filename: /boot/fai/installimage
Out filename: /boot/fai/installimage_3com
Adding MBA support...
MBA support has been successfully added
```

```
BOOTP environment prepared.
make-fai-nfsroot finished.
Stopping NFS kernel daemon: mountd nfsd.
Unexporting directories for NFS kernel daemon...done.
Exporting directories for NFS kernel daemon...done.
Starting NFS kernel daemon: nfsd mountd.
You have no FAI configuration. Copy FAI template files with:
cp -a /usr/share/fai/templates/* /usr/local/share/fai
Then change the configuration files to meet your local needs.
FAI setup finished.
```

The warning messages from `dpkg` about dependencies problems can be ignored. The setup routine adds some lines to `/etc/exports` to export the `nfsroot` and the configuration space to all hosts that belong to the netgroup `faiclents`. If you already export a parent directory of these directories, you may comment out these lines, since the kernel `nfs` server has problems exporting a directory and one of its subdirectories with different options. All install clients must belong to this netgroup, in order to mount these directories successfully. Netgroups are defined in `/etc/netgroup` or in the corresponding NIS map. An example for the netgroup file can be found in `/usr/share/doc/fai/examples/etc/netgroup`. For more information, read the manual pages `netgroup(5)` and the NIS HOWTO. After changing the netgroups, the NFS server has to reload its configuration. Use one of the following commands, depending on which NFS server you are using:

```
kueppers# /etc/init.d/nfs-kernel-server reload
kueppers# /etc/init.d/nfs-user-server reload
```

The setup also creates the account `fai` (defined by `$LOGUSER`) if not already available. The log files of all install clients are saved to the home directory of this account. If you boot from network card, you should change the primary group of this account, so this account has write permissions to `/boot/fai` in order to change the symbolic links to the kernel image which is booted by a client. See also variable `TFTPLINK` in `class/DEFAULT.var`.

After that, FAI is installed successfully on your server, but has no configuration for the install clients. Start with the templates from `/usr/share/fai/templates` using the copy command above and read ‘Installation details’ on page 27. Before you can set up a DHCP or BOOTP daemon, you should collect some network information of all your install clients. This is described in section ‘Creating a boot floppy’ on page 12.

When you make changes to `/etc/fai/fai.conf` or want to install a new kernel to `nfsroot`, the `nfsroot` has to be rebuilt by calling `make-fai-nfsroot`.

2.3.1 Troubleshooting the setup

The setup of FAI adds the FAI account, exports file systems and calls `make-fai-nfsroot`. If you call `make-fai-nfsroot -v` you will see more messages. When using a local Debian

mirror, it's important that the install server can mount this directory via NFS. If this mount fails, check `/etc/exports` and `/etc/netgroup`. An example can be found in `/usr/share/doc/fai/examples/etc/netgroup`.

Chapter 3

Preparing booting

Before booting for the first time, you have to choose which medium you use for booting. You can use the boot floppy or configure the computer to boot via network card using a boot PROM, which is much smarter.

3.1 Booting from 3Com network card with boot PROM

If you have a 3Com network card that is equipped with a boot ROM by Lanworks Technologies or already includes the DynamicAccess Managed PC Boot Agent (MBA) software¹, you can enter the MBA setup by typing Ctrl+Alt+B during boot. The setup will look like this:

```
Managed PC Boot Agent (MBA) v4.00
(C) Copyright 1999 Lanworks Technologies Co. a subsidiary of 3Com Corporation
All rights reserved.
=====
                                Configuration

Boot Method:                    TCP/IP
Protocol:                       BOOTP
Default Boot:                   Network
Local Boot:                     Enabled
Config Message:                 Enabled
Message Timeout:                3 Seconds
Boot Failure Prompt:            Wait for key
=====
Use cursor keys to edit: Up/Down change field, Left/Right change value
ESC to quit, F9 restore previous settings, F10 to save
```

Set the boot method to TCP/IP and the protocol to either BOOTP or DHCP. I prefer the BOOTP protocol because the daemon automatically reloads its configuration when it has changed.

¹<http://support.3com.com/infodeli/tools/nic/mba.htm>

Make a symbolic link from the hostname of your client to the appropriate kernel image in `/boot/fai`. In the following example the host is called `bigfoot`. The file `installimage_3com` is created by `imggen` and suitable for booting 3Com network cards². You can also use the utility `tlink (/usr/share/fai/utils/tlink)` to create this link.

```
kueppers# cd /boot/fai
kueppers# ln -s installimage_3com bigfoot
```

3.2 Booting from network card with a PXE conforming boot ROM

Some network cards (e.g. Intel EtherExpress PRO 100) have a fixed boot configuration, so they can only use the PXE boot protocol. This requires a PXE Linux boot loader and a special version of the TFTP daemon, which is available in the Debian package `tftpd-hpa`. First set up the DHCP daemon. A sample configuration files can be found in `/usr/share/doc/fai/examples/etc/dhcpd.conf`. Then enable the special tftp daemon using this line in file `/etc/inetd.conf`:

```
tftp dgram udp wait root /usr/sbin/in.tftpd -s -r blksize /boot/fai
```

See `/usr/share/doc/syslinux/pxelinux.doc.gz` for more information about how to boot such an environment. There are also some mails in the FAI mailing list archive concerning this topic. The PXE environment uses the original kernel image (not the netboot image made by `mkubi-linux`) which is copied to `/boot/fai/vmlinuz-install`. A default configuration for `pxelinux` is written to `/boot/fai/pxelinux.cfg/default`.

3.3 Creating a boot floppy

If your network card can't boot itself, you have to boot via floppy disk. To create the boot floppy use the command `make-fai-bootfloppy(8)`. Since there's no client specific information on this floppy, it's suitable for all your install clients. You can also specify additional kernel parameters for this boot floppy or set other variables, if desired. Do not enable BOOTP support when you have a DHCP server running in your network and vice versa. This could lead to missing information. There's also a manual page for `make-fai-bootfloppy(8)`. If you have no BOOTP or DHCP server, supply the network configuration as kernel parameters. The format is:

```
ip=<client-ip>:<server-ip>:<gw-ip>:<netmask>:<hostname>:<device>:<autoconf>
```

²If you have problems booting with a 3Com network card and get the error "BOOTP record too large" after the kernel is transferred to the computer, try the `imggen-1.00` program to convert the netboot image to a `installimage_3com` image. I had this problem using `netboot 0.8.1-4` and `Image Creator for MBA ROMs v1.01`, Date: Nov 26, 2001 but only on an Athlon computer.

For additional information see `/usr/src/linux/Documentation/nfsroot.txt` in the kernel sources.

3.4 Booting from a CD-ROM

There some ongoing work to create a bootable CD-ROM which can boot and install an install client. Currently this script is not included in the FAI package but a beta version is available. It will contain the `nfsroot`, the configuration space and a subset of the Debian mirror, which contains all packages that you need for an unattended installation. Look at the mailing list archive of FAI for more information. A first version is available at <http://members.iinet.net.au/~niall/fai/>.

3.5 Collecting Ethernet addresses

Now it's time to boot your install clients for the first time. They will fail to boot completely, because no BOOTP or DHCP daemon is running yet or recognizes the hosts. But you can use this first boot attempt to easily collect all Ethernet addresses of the network cards.

You have to collect all Ethernet (MAC) addresses of the install clients and assign a hostname and IP address to each client. To collect all MAC addresses, now boot all your install clients. While the install clients are booting, they send broadcast packets to the LAN. You can log the MAC addresses of these hosts by running the following command simultaneously on the server:

```
# tcpdump -qte broadcast and port bootpc >/tmp/mac.lis
```

After the hosts has been sent some broadcast packets (they will fail to boot because `bootpd` isn't running or does not recognize the MAC address yet) abort `tcpdump` by typing `ctrl-c`. You get a list of all unique MAC addresses with these commands:

```
# perl -ane 'print "\U$F[0]\n"' /tmp/mac.lis|sort|uniq
```

After that, you only have to assign these MAC addresses to hostnames and IP addresses (`/etc/ethers` and `/etc/hosts` or corresponding NIS maps). With this information you can configure your BOOTP or DHCP daemon (see the section 'Configuration of the BOOTP daemon' on the current page). I recommend to write the MAC addresses (last three bytes will suffice if you have network cards from the same vendor) and the hostname in the front of each chassis.

3.6 Configuration of the BOOTP daemon

An example configuration for the BOOTP daemon can be found in `/usr/share/doc/fai/examples/etc/bootptab`.

```

# /etc/bootptab example for FAI
# replace FAISERVER with the name of your install server

.faiglobal:\
:ms=1024:\
:hd=/boot/fai:\
:hn:bs=auto:\
:rp=/usr/lib/fai/nfsroot:

.failocal:\
:tc=.faiglobal:\
:sa=FAISERVER:\
:ts=FAISERVER:\
:sm=255.255.255.0:\
:gw=134.95.9.254:\
:dn=informatik.uni-koeln.de:\
:ds=134.95.9.136,134.95.100.209,134.95.100.208,134.95.140.208:\
:ys=rubens:yd=informatik4711.YP:\
:nt=time.rrz.uni-koeln.de,time2.rrz.uni-koeln.de:

# now one entry for each install client
bigfoot:ha=0x00105A240012:bf=bigfoot:tc=.failocal:T172="verbose sshd createvt
ant01:ha=0x00105A000000:bf=ant01:tc=.failocal:T172="sshd":

```

Insert one line for each install client at the end of this file as done for the hosts *bigfoot* and *ant01*. Replace the string `FAISERVER` with the name of your install server. If the install server has multiple network cards and host names, use the host name of the network card to which the install clients are connected. Then adjust the other network tags (`sm`, `gw`, `dn`, `ds`) to your local needs.

sm: Subnet mask

gw: Default gateway / router

dn: Domain name

ds: List of DNS server. The `/etc/resolv.conf` file will be created using this list of DNS servers and the domain name.

T172: List of `FAI_FLAGS`; e.g. `verbose`, `debug`, `reboot`, `createvt`, `sshd`, `syslogd`

T173: Reserved for future use

The tags for NIS and time servers (`yp`, `yd`, `nt`) are optional. Tags with prefix `T` (starting from `T170`) are generic tags which are used to transfer some FAI specific data to the clients³

³`T170=FAI_LOCATION`, `T171=FAI_ACTION`. You should define these variables in a `class/*.var` script. But for backward compatibility, you can define these variables also from a BOOTP or DHCP server.

The list of *FAI_FLAGS* can be space or comma separated. *FAI_FLAGS* in `bootptab` must be separated by whitespace. If you define *FAI_FLAGS* as an additional kernel parameter, the flags must be separated with a comma. If you do not have full control over the BOOTP or DHCP daemon (because this service is managed by a central service group) you can also define the variable *FAI_ACTION* in a `/fai/class/*.var` scripts. Look at `LAST.var` for an example. The variable *T170* can also be defined by a daemon. It's better to use a script for that. When you have created your `bootptab` file, you have to enable the BOOTP daemon once. It's installed but Debian does not enable it by default. Edit `/etc/inetd.conf` and remove the comment (the hash) in the line containing `#bootps`. Then tell `inetd` to reload its configuration.

```
# /etc/init.d/inetd reload
```

The BOOTP daemon automatically reloads the configuration file if any changes are made to it. The daemon for DHCP must always be manually restarted after changes to the configuration file are made.

Now it's time to boot all install clients again! FAI can perform several actions when the client is booting. This action is defined in the variable *FAI_ACTION*. Be very careful if you set *FAI_ACTION* to *install*. This can destroy all your data on the install client, indeed most time it should do this ;-). It's recommended to change this only on a per-client base in the BOOTP configuration. Do not change it in the section `.failocal` in `/etc/bootptab`, which is a definition for all clients.

3.6.1 Troubleshooting BOOTP daemon

The BOOTP daemon can also be started in debug mode if it is not enabled in `inetd.conf`:

```
# bootpd -d7
```

3.7 Configuration of the DHCP daemon

An example for `dhcp.conf(5)` is available in `/usr/share/doc/fai/examples/etc`, which is tested with version 2.x of the DHCP daemon. Version 3 needs a slightly different configuration file. Start using this example and look at all options used therein. If you make any changes to this configuration, you must restart the daemon.

```
# /etc/init.d/dhcp restart
```

Therefore it's recommended to only supply data into this configuration file, which doesn't change frequently. You should define the variables *FAI_ACTION* in a script in the configuration space.

3.8 Boot messages

These are the messages when booting from floppy disk.

```
GRUB loading stage2.....
< now the grub menu is displayed >
BOOTING 'FAI-BOTH'
kernel (fd0)/vmlinuz-2.4.20 root=/dev/nfs ip=both
  [Linux-bzImage, setup=0x1400, size=0xd8450]

Uncompressing Linux... OK, booting the Kernel.
Linux version 2.4.20 (root@kueppers) (gcc version 2.95.4 20011002
.
.
.
```

After this, the rest of the boot message will be equal to those when booting from network card. When booting from network card you will see:

```
BOOTP.
TFTP....
Linux Net Boot Image Loader Version 0.8.1 (netboot)
.
Uncompressing Linux... OK, booting the Kernel.
Linux version 2.4.20 (root@kueppers) (gcc version 2.95.4 20011002
.
.
.

Sending BOOTP requests . OK
IP-Config: Complete:

Sending BOOTP requests ..... OK
IP-Config: Got BOOTP answer from 134.95.9.150, my address is 134.95.9.200
IP-Config: Complete:
  device=eth0, addr=134.95.9.200, mask=255.255.255.0, gw=134.95.9.254,
  host=bigfoot, domain=informatik.uni-koeln.de, nis-domain=informatik4711.YP
  bootserver=134.95.9.150, rootserver=134.95.9.150, rootpath=/usr/lib/fai/nf
Partition check:
.
.
-----
FAI 2.4.1, 2 apr 2003
Fully Automatic Installation for Debian GNU/Linux
```

```
Copyright (c) 1999-2003, Thomas Lange
lange@informatik.uni-koeln.de
```

```
-----
Calling task_confdir
Kernel parameters:auto rw root=/dev/nfs ip=both
Defining variable: root=/dev/nfs
Defining variable: ip=both
Sending BOOTP request using device eth0
FAI_FLAGS: verbose=1
FAI_FLAGS: createvt=1
.
Calling task_setup
.
.
Calling task_defclass
/usr/bin/fai-class: Defining classes.
.
.
Calling task_action
FAI_ACTION: install
Performing FAI installation. All data may be overwritten!
.
.
Press <RETURN> to reboot or ctrl-c to execute a shell
```

When the copyright message is shown, the install client has mounted the `nfsroot`⁴ to the clients' root directory `/`. This is the whole filesystem for the client at this moment. After `task_confdir` is executed, the configuration space is mounted or received from an CVS repository. Before the installation is started (`FAI_ACTION=install`) the computer beeps three times. So, be watchful when you hear three beeps but you do not want to perform an installation!

3.9 Troubleshooting boot messages

This is the error message you will see, when your network card is working, but the install server does not export the configuration space directory to the install clients.

```
Root-NFS: Server returned error -13 while mounting /usr/lib/fai/nfsroot
VFS: Unable to mount root fs via NFS, trying floppy.
VFS: Cannot open root device "nfs" or 02:00
Kernel panic: VFS Unable to mount root fs on 02:00
```

⁴`/usr/lib/fai/nfsroot` from the install server

Use the following command to see which directories are exported from the install server (named kueppers):

```
showmount -e kueppers
```

The following error message indicates that your install client doesn't get an answer from a BOOTP server. Check your cables or start the `bootpd(8)` daemon with the debug flag enabled.

```
Sending BOOTP requests ..... timed out!  
IP-Config: Retrying forever (NFS root)...
```

If you get the following error message, the install kernel has no driver compiled in for your network card.

```
IP-Config: No network devices available  
Partition check:  
hda: hda1 hda2 < hda5 hda6 hda7 hda8 >  
Root-NFS: No NFS server available, giving up.  
VFS: Unable to mount root fs via NFS, trying floppy.  
VFS: Insert root floppy and press ENTER
```

Then you have to compile the driver for your network card into a new kernel. This driver must not be a kernel module. To compile the new kernel, start using the default kernel configuration of FAI.

```
kueppers# cd /usr/src/kernel-source-2.4.20  
kueppers# cp /usr/lib/fai/nfsroot/boot/config-2.4.20 .config  
kueppers# make menuconfig
```

Call `make menuconfig` and add the driver in `menuNetwork device support/Ethernet` which supports your network card. Then create a Debian package using `make-kpkg(8)`:

```
kueppers# make-kpkg clean  
kueppers# make-kpkg --revision fai2 kernel-image
```

This command creates the file `/usr/src/kernel-image-2.4.20_fai2_i386.deb`. Adjust the variable `KERNELPACKAGE` in `/etc/fai/fai.conf` and rebuild the `nfsroot`.

```
kueppers# make-fai-nfsroot
```

After that, you have to create a new boot floppy if you need it. Now your network card should be recognized and the install kernel should mount the `nfsroot` successfully. More information how to compile an install kernel can be found in the `README` of the package `fai-kernels`.

3.10 Collecting other system information

Now the clients have booted with *FAI_ACTION* set to *sysinfo*. Type `ctrl-c` to get a shell or use `Alt-F2` or `Alt-F3` and you will get another console terminal, if you have added `createvt` to *FAI_FLAGS*. Remote login is available via the secure shell if `sshd` is added to *FAI_FLAGS*. The encrypted password is set with variable *FAI_ROOTPW* in `/etc/fai/fai.conf` and defaults to “`fai`”. This is only the root password during the installation process, not for the new installed system. You can also log in without a password when using *SSH_IDENTITY*. To log in from your server to the install client (named `ant01` in this example) use:

```
> ssh root@ant01
Warning: Permanently added 'ant01,134.95.9.200' to the list of known hosts.
root@ant01's password:
```

You now have a running Linux system on the install client without using the local hard disk. Use this as a rescue system if your local disk is damaged or the computer can't boot properly from hard disk. You will get a shell and you can execute various commands (`dmesg`, `lsmod`, `df`, `lspci`, ...). Look at the log file in `/tmp/fai`. There you can find much information about the boot process. All log files from `/tmp/fai` are also written to the *\$LOGSERVER* (if not defined: the install server) into the directory `~fai/ant01/sysinfo/`⁵

A very nice feature is that FAI mounts all filesystems it finds on the local disks read only. It also tells you on which partition a file `/etc/fstab` exists. When only one file system table is found, the partitions are mounted according to this information. Here's an example:

```
ant01:~# df

Filesystem 1k-blocks      Used Available Use% Mounted on
rootfs          2064192  1071184    888152  55% /
/dev/root       2064192  1071184    888152  55% /
shm              63548         76     63472   1% /tmp
kueppers:/usr/local/share/fai
                2064192   994480   964856   51% /fai
/dev/hda1        54447         9859    41777   19% /tmp/target
/dev/hda10      1153576         20   1141992   0% /tmp/target/files/install
/dev/hda9        711540         20    711520   0% /tmp/target/home
/dev/hda8        303336         13    300191   0% /tmp/target/tmp
/dev/hda7       1517948       98252   1342588   7% /tmp/target/usr
/dev/hda6       202225         8834    182949   5% /tmp/target/var
```

This method can be used as a rescue environment! In the future it will be possible to make backups or restore data to existing filesystems. If you need a filesystem with read-write access use the `rwmount` command:

⁵More general: `~$LOGUSER/$HOSTNAME/$FAI_ACTION/`. Two additional symbolic links are created. The symlink `last` points to the log directory of the last `fai` action performed. The symlinks `last-install` and `last-sysinfo` point to the directory with of the last corresponding action. Examples of the log files can be found on the FAI homepage.

```
ant01:~# rwmount /tmp/target/home
```

3.11 Checking parameters from BOOTP and DHCP servers

If the install client boots with action *sysinfo*, you can also check if all information from the BOOTP or DHCP daemons are received correctly. The received information is written to `/tmp/fai/boot.log`. An example of the result of a BOOTP request can be found in ‘The setup routines of the install clients’ on page [30](#).

3.12 Rebooting the computer

At any time you can reboot the computer using the command `faireboot`, also if logged in from remote. If the installation hasn’t finished, use `faireboot -s`, so the log files are also copied to the install server.

Chapter 4

Overview of the installation sequence

The following tasks are performed during an installation after the linux kernel has booted on the install clients.

1. Set up FAI
2. Load kernel modules
3. Define classes
4. Define variables
5. Partition local disks
6. Create and mount local filesystems
7. Install software packages
8. Call site specific configuration scripts
9. Save log files
10. Reboot the new installed system

You can also define additional programs or scripts which will be run on particular occasions. They are called hooks. Hooks can add additional functions to the installation process or replace the default subtasks of FAI. So it's very easy to customize the whole installation process. Hooks are explained in detail in 'Hooks' on page [40](#).

The installation time is determined by the amount of software but also by the speed of the processor and hard disk. Here are some sample times. All install clients have an 100Mbit network card installed. Using a 10 Mbit LAN does not decrease the installation time considerably, so the network will not be the bottleneck when installing several clients simultaneously.

Athlon XP1600+	, 896MB, SCSI disk,	1 GB software	6 min
AMD-K7, 500MHz	, 320MB, IDE disk,	780 MB software	12 min
PentiumPro 200MHz	, 128MB, IDE disk,	800 MB software	28 min
Pentium III 850MHz,	256MB, IDE disk,	820 MB software	10 min
Pentium III 850MHz,	256MB, IDE disk,	180 MB software	3 min

4.1 Set up FAI

After the install client has booted, only the script `/sbin/rcS_fai`¹ is executed. This is the main script which controls the sequence of tasks for FAI. No other scripts in `/etc/init.d/` are executed.

A ramdisk is created and mounted to `/tmp`, which is the only writable directory until local filesystems are mounted. Additional parameters are received from the BOOTP or DHCP daemon and the configuration space if mounted via NFS from the install server to `/fai`. The setup is finished after additional virtual terminals are created and the secure shell daemon for remote access is started on demand.

4.2 Defining classes, variables and loading kernel modules

Now the script `fai-class(1)` is used to define classes. Therefore several scripts in `/fai/class/` are executed to define classes. All scripts matching `[0-9]*` are executed in alphabetical order. Scripts ending in `.source` are sourced, so they can define new classes by adding these classes to the variable `newclasses` (see `06hwdetect.source` for an example). Every word that these scripts print to the standard output are interpreted as class names. These classes are defined for the install client. You can also say this client belongs to these classes. A class is defined or undefined and has no value. Only defined classes are of interest for an install client. The description of all classes can be found in `/usr/share/doc/fai/classes_description.txt`. It is advisable to document the job a new class performs. Then, this documentation is the base for composing the whole configuration from classes.

The scripts `11modules.source` loads kernel modules on demand. So you can use classes when loading modules and also define more classes after the kernel has loaded modules and recognized new hardware. But normally you only have to deal with the file `DEFAULT.mod`. The complete description of all these scripts can be found in 'Scripts in `/fai/scripts`' on page 38.

The script `30menu.source` pops up a little menu and asks the user which kind of installation should be performed (e.g. CAD workstation, notebook, scientific workstation, work group server, Gnome desktop...). Keep in mind that this won't lead to a fully automatic installation ;-)

¹Since the root filesystem on the clients is mounted via NFS, `rcS_fai` is located in `/usr/lib/fai/nfsroot/sbin` on the install server.

After defining the classes, every file matching `*.var` with a prefix which matches a defined class is executed to define variables. There, you should define the variable `FAI_ACTION` and others. Currently, `FAI_ACTION` is defined in `LAST.var` for all install clients.

4.3 Partitioning local disks, creating filesystems

For disk partitioning exactly one disk configuration file from `/fai/disk_config` is selected using classes. It's the description of how all the local disks will be partitioned, where filesystems should be created (and their types like `ext2`, `ext3`, `reiserfs`), and how they are mounted. It's also possible to preserve the disk layout or to preserve the data on certain partitions. It's done by the command `setup_harddisks`, which uses `sfdisk` for partitioning. The format of the configuration file is described in `/usr/share/doc/fai/README.disk_config`.

During the installation process all local filesystems are mounted relative to `/tmp/target`. For example `/tmp/target/home` will become `/home` in the new installed system.

4.4 Installing software packages

When local filesystems are created, they are all empty (except for preserved partitions). Now the Debian base system and all requested software packages are installed on the new filesystems. First the base archive is unpacked, then the command `install_packages(8)` installs all packages using `apt-get(8)` without any manual interaction needed. If a package requires an other package, `apt-get(8)` resolves this dependency by installing the required package.

Classes are also used when selecting the configuration files in `/fai/package_config/` for software installation. The format of the configuration files is described in 'Software package configuration' on page 36.

4.5 Site specific configuration

After all requested software packages are installed, the system is nearly ready to go. But not all default configurations of the software packages will meet your site specific needs. So you can call arbitrary scripts which adjust the system configuration. Therefore scripts which match a class name in `/fai/scripts` will be executed. If `/fai/scripts/classname/` is a directory, all scripts that match `S[0-9]*` in this directory are executed. So it is possible to have several scripts of different types (shell, cfengine, ...) to be executed for one class. FAI comes with some templates for these scripts, but you can write your own Bourne, bash, perl, cfengine or expect scripts.

These important scripts are described in detail in 'Scripts in `/fai/scripts`' on page 38.

4.6 Save log files

When all installation tasks are finished, the log files are written to `/var/log/fai/$HOSTNAME/install/`² on the new system and to the account on the install server if `$LOGUSER` is defined in `/etc/fai/fai.conf`. It is also possible to specify another host as destination of the log saving by in a file in `/fai/class/`. Additionally, two symlinks will be created to indicated the last directory written.

4.7 Reboot the new installed system

At least the system is automatically rebooted if “reboot” was added to `FAI_FLAGS`. This is only useful if booting from network card or if you can change the boot device using the command `bootsector(8)`. Otherwise, you have to remove the floppy disk and type return or call `faireboot` from a remote login. You must change the boot device to boot the new installed system otherwise the installation would be performed again. Read ‘Changing the boot device’ on page 39 for how to change the boot device.

²`/var/log/fai/localhost/install/` is a link to this directory.

Chapter 5

How to plan your installation

Plan your installation, and FAI installs your plans.

Before starting your installation, you should spend much time in planning your installation. When you're happy with your installation concept, FAI can do all the boring, repetitive tasks to turn your plans into practice. FAI can't do good installations if your concept is imperfect or lacks some important details. Start planning the installations by answering the following questions:

Will I create a Beowulf cluster, or do I have multiple workstations, each only used by a single user?

How does my LAN topology look like?

Do I have uniform hardware?

Will the hardware stay uniform in the future?

Does the hardware need a special kernel?

How should the hosts be named?

How should the local hard disks be partitioned?

Which applications will be run by the users?

Do the users need a queueing system?

What software should be installed?

Which daemons should be started, and what should the configuration for these look like?

Which remote filesystems should be mounted?

How should backups be performed?

You also have to think about user accounts, printers, a mail system, cron jobs, graphic cards, dual boot, NIS, NTP, timezone, keyboard layout, exporting and mounting directories via NFS and many other things. So, there's a lot to do before starting an installation. And remember that knowledge is power, and it's up to you to use it. Installation and administration is a process, not a product. FAI can't do things you don't tell it to do.

But you need not to start from scratch. Look at all files and scripts in the configuration space. There are a lot of things you can use for your own installation. A good paper with more aspects of building an infrastructure is <http://www.infrastructures.org/papers/bootstrap/> "Bootstrapping an Infrastructure".

Chapter 6

Installation details

6.1 The configuration space

The configuration is the collection of information about how exactly to install a computer. The central configuration space for all install clients is located on the install server in `/usr/local/share/fai` and its subdirectories. This will be mounted by the install clients to `/fai`. It's also possible to receive all the configuration data from a `cvs(1)` repository. The following subdirectories are present and include several files:

class/ Scripts and files to define classes and variables and to load kernel modules.

disk_config/ Configuration files for disk partitioning and file system creation.

debconf/ This directory holds all `debconf(8)` data. Not yet used.

package_config/ File with lists of software packages to be installed or removed.

scripts/ Script for local site customization.

files/ Files used by customization scripts, e.g. user created kernel packages. Most files are located in a subtree structure which reflects the ordinary directory tree. For example, the templates for `nsswitch.conf` are located in `/fai/files/etc/nsswitch.conf`. The directory `files/packages/` can contain you local Debian packages, which can be installed when adding them to the variables `addpackages`. See 'Defining Variables' on page 35 for more information.

hooks/ Hooks are user defined programs or scripts, which are called during the installation process.

The main installation script `rcS_fai` uses all these subdirectories in the order listed except for hooks. The FAI package contains templates for all these configuration scripts and files in `/usr/share/fai/templates`. Copy the configuration templates to the configuration space and start an installation. These files need not belong to the root account. You can change their ownership and then edit the configuration with a normal user account.

```
# cp -a /usr/share/fai/templates/* /usr/local/share/fai
# chown -R fai /usr/local/share/fai
```

These files contain configuration for some example hosts. Examples are: a cluster of workstations (*bigfoot*, *ant01*, *ant02*,...) and a Beowulf cluster with a master node called *nucleus* and computing nodes called *atom01*, *atom02*,... and our desktop machines *kueppers*, *dom* and our notebooks *pittermaennche*, *paeffgen*.

bigfoot This is a server with much software. It provides the home directory and `/usr` for its NFS clients. Also some daemons are installed and activated by default.

ant01,... These dataless clients mount `/usr` and `/home` from *bigfoot*. Most of the disk space is spent on a scratch partition, which is exported to a netgroup of hosts. The host *kueppers* has a similar configuration.

nucleus This Beowulf master node is a server with much software. It provides the home directory and `/usr/local` for its computing nodes. Also some daemons are installed and activated by default.

atom01,... These Beowulf clients mount `/usr/local` and `/home` from *nucleus*. Most of the disk space is spent on a scratch partition, which is exported to a netgroup of hosts. All scratch partitions are mounted on all Beowulf clients via the automounter.

dom It's a desktop machine, which mounts `/home` from a server and synchronizes the `/usr/local` partition via `rsync(1)` from the server. If it belongs to the class `USR_LOCAL_MOUNT` it will mount this directory from the server.

Start looking at these examples and study them. Then change or add things to these examples. But don't forget to plan your own installation!

6.2 The default tasks

After the kernel has booted, it mounts the root file system via NFS from the install server and `init(8)` starts the script `/sbin/rcS_fai`. This script controls the sequence of the installation. No other scripts in `/etc/init.d/` are used.

The installation script uses many subroutines, which are defined in `/usr/share/fai/subroutines`, and an operating system specific file¹. All important tasks of the installation are called via the subroutine `task` appended by the name of the task as an option (e.g. `task instsoft`). The subroutine `task` calls hooks with prefix *name* if available and then calls the default task (defined as `task_name` in `subroutines`). The default task and its hooks can be skipped on demand by using the subroutine `skiptask()`.

Now follows the description of all default tasks.

¹`/usr/share/fai/subroutines-linux` for Linux, `/usr/share/fai/subroutines-sunos` for Solaris.

- confdir** The kernel appended parameters define variables, the syslog and kernel log daemon are started. The list of network devices is stored in *\$netdevices*. Then additional parameters are fetched from a DHCP or BOOTP server and also additional variables are defined. The DNS resolver configuration file is created. The configuration space is mounted from the install server to */fai* or it is checked out from the corresponding *cvs(1)* repository. To use a cvs repository, you have to set the variables *\$FAI_CVSROOT*, *\$FAI_CVSTAG*, *\$FAI_CVSMODULE*. For details look at the subroutine *get_fai_cvs()*. After that, the file */fai/hooks/subroutines* is sourced if it exists. Using this file, you can define your own subroutines or override the definition of FAI's subroutines.
- setup** This task sets the system time, all *FAI_FLAGS* are defined and two additional virtual terminals are opened on demand. A secure shell daemon is started on demand for remote logins.
- defclass** Calls *fai-class(1)* to define classes using scripts and files in */fai/class* and classes from */tmp/fai/additional-classes*.
- defvar** Sources all files */fai/class/*.var* for every defined class. If a hook has written some variable definitions to the file */tmp/fai/additional.var*, this file is also sourced.
- action** Depending on the value of *\$FAI_ACTION* this subroutine decides which action FAI should perform. The default available actions are: *sysinfo* and *install*. If *\$FAI_ACTION* has another value, a user defined action is called if a file */fai/hooks/\$FAI_ACTION* exists. So you can easily define your own actions.
- sysinfo** Called when no installation is performed but the action is *sysinfo*. It shows information about the detected hardware and mounts the local hard disks read only to */tmp/target/partitionname* or with regard to a *fstab* file found inside a partition. Log files are stored to the install server.
- install** This task controls the installation sequence. You will here three beeps before the installation starts. The major work is to call other tasks and to save the output to */tmp/fai/rcs.log*. If you have any problems during installation, look at all files in */tmp/fai/*. You can find examples of the log files for some hosts in the download directory of the FAI homepage.
- partition** Calls *setup_harddisk* to partition the hard disks. The task writes variable definitions for the root and boot partition and device (*\$ROOT_PARTITION*, *\$BOOT_PARTITION*, *\$BOOT_DEVICE*) to */tmp/fai/disk_var.sh* and creates a *fstab* file.
- mountdisks** Mounts the created partitions according to the created */tmp/fai/fstab* file relative to *\$FAI_ROOT*.
- extrbase** Extracts the base tar file *base.tgz*, which consists of all required packages. This is a snapshot of a basic Debian system created by *debootstrap(8)*
- mirror** If a local Debian mirror is accessed via NFS (when *\$FAI_DEBMIRROR* is defined), this directory will be mounted to *\$MNTPOINT*.

- updatebase** Prepares the extracted Debian base system for further installation and updates the list of available packages. Updates the packages to the newest version. It also fakes some commands (called diversions) inside the new installed system using `dpkg-divert(8)`.
- instsoft** Installs the desired software packages using class files in `/fai/packages_config`.
- configure** Calls scripts in `/fai/scripts/` and its subdirectories for every defined class.
- finish** Unmounts all filesystems in the new installed system and removes diversions of files using the command `fai-divert`.
- faiend** Wait for background jobs to finish (e.g. emacs compiling lisp files) and automatically reboots the install clients or waits for manual input before reboot.
- chboot** Changes the symbolic link on the install server which indicates which kernel image to load on the next boot from network card via TFTP.
- savelog** Saves log files to local disk and to the account `$LOGUSER` on `$LOGSERVER` (defaults to the install server). Currently the file `error.log` will not be copied to the log server.

6.3 The setup routines of the install clients

After the subroutine `fai_init` has done some basic initialization (create ramdisk, read `fai.conf` and all subroutines definitions, set path, print copyright notice), the setup continues by calling the task `confdir` and the task `setup`. The command `get-boot-info` is called to get all information from the BOOTP or DHCP server. This command writes the file `/tmp/fai/boot.log`, which then is sourced to define the corresponding global variables. This is an example for this log file when using a BOOTP server.

```
# cat /tmp/fai/boot.log

netdevices_all=" eth0"
netdevices_up=""
netdevices="eth0"
# --- network device eth0 ---
SERVER='134.95.9.150'
IPADDR='134.95.9.200'
BOOTFILE='/boot/fai'
NETMASK='255.255.255.0'
NETWORK='134.95.9.0'
BROADCAST='134.95.9.255'
GATEWAYS_1='134.95.9.254'
GATEWAYS='134.95.9.254'
ROOT_PATH='/usr/lib/fai/nfsroot'
DNSSRVS_1='134.95.9.136'
DNSSRVS_2='134.95.129.23'
```

```

DNSSRV3='134.95.100.208'
DNSSRV4='134.95.140.208'
DNSSRV='134.95.9.136 134.95.129.23 134.95.100.208 134.95.140.208'
DOMAIN='informatik.uni-koeln.de'
SEARCH='informatik.uni-koeln.de uni-koeln.de'
YPSRV1='134.95.9.10'
YPSRV='134.95.9.10'
YPDO='informatik4711.YP'
TIMESRV1='134.95.9.10'
TIMESRV='134.95.9.10'
NTPSRV1='192.76.170.145'
NTPSRV2='134.95.4.37'
NTPSRV='192.76.170.145 134.95.4.37'
HOSTNAME='bigfoot'
T172='verbose sshd createvt syslogd'
# define variable if T17x is defined
[ "$T170" ] && FAI_LOCATION=$T170
[ "$T171" ] && FAI_ACTION=$T171
[ "$T172" ] && FAI_FLAGS=$T172

```

The last part is shell code which maps the T17X tags to shell variables. The tag T172 is the definition for `$FAI_FLAGS`. It contains a space separated list of flags. The following flags are known:

verbose Create verbose output during installation. This should always be the first flag, so consecutive definitions of flags will be verbosely displayed.

debug Create debug output. No unattended installation is performed. During package installation you have to answer all questions of the postinstall scripts on the client's console.

sshd Start the ssh daemon to enable remote logins.

syslogd Start the system and kernel log daemon, so processes can use it to give out information. This flag should only be used when the syslogd is not already running on the system, so it should only be set when initially installing, *not* on updates!

createvt Create two virtual terminals and execute a bash if `ctrl-c` is typed in the console terminal. The additional terminals can be accessed by typing `Alt-F2` or `Alt-F3`. Otherwise no terminals are available and typing `ctrl-c` will reboot the install client. Setting this flag is useful for debugging. If you want an installation which should not be interruptible, do not set this flag.

reboot Reboot the install client after installation is finished without typing RETURN on the console. This is only useful if you can change the boot image or boot device automatically or your assembly robot can remove the boot floppy via remote control :-). Currently this should only be used when booting from network card and using `$TFTPLINK`.

6.4 The class concept

Classes determine which configuration file to choose from a list of available templates. Classes are used in all further tasks of the installation. To determine which config file to use, an install client searches the list of defined classes and uses all configuration files that match a class name. It's also possible to use only the configuration file with the highest priority since the order of classes define the priority from low to high. There are some predefined classes (DEFAULT, LAST and the hostname), but classes can also be listed in a file or defined dynamically by scripts. So it's easy to define a class depending on the subnet information or on some hardware that is available on the install client.

The idea of using classes in general and using certain files matching a class name for a configuration is adopted from the installation scripts by Casper Dik for Solaris. This technique proved to be very useful for the SUN workstations, so I also use it for the fully automatic installation of Linux. One simple and very efficient feature of Casper's scripts is to call a command with all files (or on the first one) whose file names are also a class. The following loop implements this function in pseudo shell code:

```
for class in $all_classes; do
if [ -r $config_dir/$class ]; then
    your_command $config_dir/$class
    # exit if only the first matching file is needed
fi
done
```

Therefore it is possible to add a new file to the configuration without changing the script. This is because the loop automatically detects new configurations files that should be used. Unfortunately cfengine does not support this nice feature, so all classes being used in cfengine need also to be specified inside the cfengine scripts. Classes are very important for the fully automatic installation. If a client belongs to class A, we say the class A is defined. A class has no value, it is just defined or undefined. Within scripts, the variable *\$classes* holds a space separated list with the names of all defined classes. Classes determine how the installation is performed. For example, an install client can be configured to become a FTP server by just adding the class FTP to it. Mostly a configuration is created by only changing or appending the classes to which a client belongs, making the installation of a new client very easy. Thus no additional information needs to be added to the configuration files if the existing classes suffice for your needs. There are different possibilities to define classes:

1. Some default classes are defined for every host: DEFAULT, LAST and its hostname.
2. Classes may be listed within a file.
3. Classes may be defined by scripts.

The last option is a very nice feature, since these scripts will define classes automatically. For example, several classes are defined only if certain hardware is identified. We use Perl and shell

scripts to define classes. All names of classes, except the hostname, are written in uppercase. They must not contain a hyphen, a hash or a dot, but may contain underscores. A description of all classes can be found in `/usr/share/doc/fai/classes_description.txt`.

Hostnames should rarely be used for the configuration files in the configuration space. Instead, a class should be defined and then added for a given host. This is because most of the time the configuration data is not specific for one host, but is can be shared among several hosts.

6.5 Defining classes

The default task `defclass` calls the script `fai-class(1)` to define classes. Therefore, scripts matching `[0-9][0-9]*` in `/fai/class` are executed. Additionally, files in this directory can contain a list of classes. We use a file `koeln` which is used for all our hosts that belong to a certain subnet. When we want to add a class to all these hosts, we just add the class to this file. For more information on defining class, read the manual pages for `fai-class(1)`.

The list of all defined classes is stored in the variable `$classes` and saved to `/tmp/fai/FAI_CLASSES`. The list of all classes is transferred to `cfengine`, so it can use them too. The script `01alias` (see below) is used to define classes for several groups of hosts. First this script defines the class with the name of the hardware architecture in uppercase letters. All hosts with the prefix `ant` use all classes in the file `anthill`. Hosts which have an IP address in subnet 134.95.9.0 also belong to the class `NET_9`, hosts in their class B subnet 134.95 use all classes of the file `koeln`. All Beowulf nodes with prefix `atom` except `atom00` (master server) will belong to the classes listed in file `atoms`. Some notebooks get also some special classes.

```
# cat 01alias

uname -s | tr /a-z/ /A-Z/
[ -x "`which dpkg`" ] && dpkg --print-installation-architecture | tr /a-z/ /A

# all hosts named ant?? are using the classes in file anthill
case $HOSTNAME in
    ant??) cat anthill ;;
esac

# all hosts named nuerburg? are becoming web kiosk systems
case $HOSTNAME in
    nuerburg?) cat wwwkiosk ;;
esac

# the Beowulf cluster; all nodes except the master node
# use classes from file class/atoms
case $HOSTNAME in
    atom00) echo BOWULF_MASTER ;;
    atom??) cat atoms ;;
```

```

esac

# if host belongs to class C subnet 134.95.9.0 use class NET_9
# exclude all hosts with an IP address above 200
case $IPADDR in
    134.95.9.2??) ;;
    134.95.*.*) cat koeln ; echo "CS_KOELN NET_9" ;;
    134.95.9.*) echo "CS_KOELN NET_9" ;;
esac

# our notebooks
case $HOSTNAME in
    paeffgen|schlaeffli)
        cat notebook
        echo "BOOTWINDOWS"
        ;;
    pittermaennche)
        cat notebook
        echo "BOOTP_SERVER"
        ;;
esac

```

Script `18disk` can be used to define classes depending on the number of local disks or the size of these disks². But you can also use a range of partition size in the disk configuration file (in `disk_config`), so you may not need a class for every different disk size.

The script `24nis` automatically defines classes corresponding to NIS. The name of the NIS domain (defined via BOOTP or DHCP) will also become a class (only uppercase letters and minus is replaced by underscore). If no NIS domain is defined, then only the class `NONIS` is defined.

Depending on partition names defined in the first matching `disk_config` found, `70partitions` defines additional classes. For example, if a partition `/files/scratch` exists, the class `FILES_SCRATCH` is defined, which forces the install client to export this directory via NFS and to install the NFS server packages.

The script `11modules.source` does not define any class, but is responsible for loading kernel modules. Kernel modules are important for detecting hardware. This script calls the script `$HOSTNAME.mod` and all scripts that have the format `<classname>.mod` and those class names are already defined. Classes which are used for loading modules must be defined before this script is called. For example, if class `DEFAULT` is defined (this class is always defined) and a file `DEFAULT.mod` exists, this script is executed. These scripts should contain all commands for loading kernel modules:

```
DEFAULT.mod:
```

²It uses the library `Fai.pm`, which includes some useful subroutines, e.g. `class`, `classes`, `read_memory_info`, `read_ethernet_info`.

```
kernelmodules="rtc floppy parport_pc usbkbd usb-uhci keybdev"

for mod in $kernelmodules; do
    [ "$verbose" ] && echo loading kernel module $mod
    modprobe -a $mod
done
```

You can find messages from modprobe in `/tmp/fai/dmesg.log` and the on the fourth console terminal by pressing `Alt-F4`.

6.6 Defining Variables

The task `defvar` defines the variables for the install client. Variables are defined by scripts in `class/*.var`. All global variables can be set in `DEFAULT.var`. For certain groups of hosts use a class file or for a single host use the file `$HOSTNAME.var`. Also here, it's useful to study all the examples. The following variables are used in the examples and may be also useful for your installation:

FAI_ACTION Set the action fai should perform. Currently this is done in the script `LAST.var`.

FAI_CONSOLEFONT Is the font which is loaded during installation by `consolechars(8)`.

FAI_KEYMAP Defines the keyboard map files in `/usr/share/keymaps` and `$FAI/files`. You need not specify the complete path, since this file will be located automatically.

kernelimage The kernel that is installed to the new system. If a Debian package `/fai/files/packages/$kernelimage` exists, install this kernel package. Otherwise install the package `$kernelimage` from the Debian mirror. For example, if `kernelimage=kernel-image-2.4.20-ide` this kernel will be installed. To install a special kernel for host `bigfoot`, set the variable

```
kernelimage=kernel-image-2.4.20_bigfoot1_i386.deb
```

and this kernel will be installed from `/fai/files/packages/`.

The easiest way to install your local kernel package is to put this Debian package in `/usr/local/share/fai/files/packages/` on the install server. Then define the kernel name with

```
kernelimage=kernel-image-2.4.20-wwwkiosk
```

without specifying the revision of the Debian package. Now fai will install this kernel using `apt-get(8)`.

rootpw The root password for the new system. Additionally, FAI creates an root account with the same password called `roott`, which uses the `tcsh(1)`.

UTC Set hardware clock to UTC if `$UTC=yes`. Otherwise set clock to local time. See `clock(8)` for more information.

time_zone Is the file relative to `/usr/share/zoneinfo/` which indicates your time zone.

liloappend Append parameters for the kernel of the new system (written to `/etc/lilo.conf`).

moduleslist Can be a multi line definition. List of modules (including kernel parameters) which are loaded during boot of the new system (written to `/etc/modules`).

TFTPLINK Link to the TFTP kernel image which boots using the root file system from the local disk.

hserver, bserver The names of the NFS servers for `/home` and `/usr`.

printers List of printers, for which a spool directory is created. The config scripts does not set up `/etc/printcap`.

addpackages The list of additional packages which are installed on the new system if they are available in `/fai/files/packages`. You can create a simple repository by using following commands on the install server:

```
# cd /usr/local/share/fai/files
# dpkg-scanpackages packages /dev/null | gzip -9 > packages/Packages.gz
```

Additionally, you can also create a Release file in this directory. Then *addpackages* can be the list of packages without a version number. For more information, refer to the repository HOWTO³. This can be used to install local site specific packages.

6.7 Hard disk configuration

The format of the hard disk configuration files is described in `/usr/share/doc/fai/README.disk_conf`. The config file `/fai/disk_config/CS_KOELN` is a generic description for one IDE hard disk, which should fit for most installations. If you can't partition your hard disk using this script⁴, use a hook instead. The hook should write the new partition table, create the file systems and create the files `/tmp/fai/fstab` and `/tmp/fai/disk_var.sh`, which contains definitions of boot and root partitions.

6.8 Software package configuration

The script `install_packages` installs the selected software packages. It uses all configuration files in `/fai/package_config` whose file name matches a defined class. The syntax is very simple.

³<http://www.isotton.com/debian/docs/repository-howto/>

⁴Currently this script uses the command `sfdisk(8)`, which isn't available on SUN SPARC.

```
# an example package class

PACKAGES taskinst
german science

PACKAGES install
adduser netstd ae
less passwd

PACKAGES remove
gpm xdm

PACKAGES dselect-upgrade
ddd                install
a2ps               install
```

Comments are starting with a hash (#) and are ending at the end of the line. Every command begins with the word `PACKAGES` followed by a command name. The command name is similar to those of `apt-get`. Here's the list of supported command names:

hold: Put a package on hold. This package will not be handled by `dpkg`, e.g not upgraded.

install: Install all packages that are specified in the following lines. If a hyphen is appended to the package name (with no intervening space), the package will be removed, not installed. All package names are checked for misspellings. Any package which does not exist, will be removed from the list of packages to install. So be careful not to misspell any package names.

remove: Remove all packages that are specified in the following lines. Append a + to the package name if the package should be installed.

taskinst: Install all packages belonging to the task that are specified in the following lines using `tasksel(1)`.

dselect-upgrade Set package selections using the following lines and install or remove the packages specified. These lines are the output of the command `dpkg --get-selections`.

Multiple lines with lists of space separated names of packages follows the commands `install` and `remove`. All dependencies are resolved and `apt-get` is used to perform the installation or removal of packages. The order of the packages is of no matter.

A line which contains the `PRELOADRM` commands, downloads a file using `wget(1)` into a directory before installing the packages. Using the `file: URL`, this file is copied from `$FAI_ROOT` to the download directory. For examples the package `realplayer` needs an archive to install the software, so this archive is downloaded to the directory `/root`. After installing the packages this file will be removed. If the file shouldn't be removed, use the the command `PRELOAD` instead.

Now it's possible to append a list of class names after the command for apt-get. So this PACKAGE command will only be executed when the corresponding class is defined. So you can combine many small files into the file DEFAULT. WARNING! Use this feature only in the file DEFAULT. See this file for some examples.

If you specify a package that does not exist this package will be removed from the installation list. You can also test all software package configuration files with the utility `chkdebnames`, which is available in `/usr/share/fai/utils/`.

```
> chkdebnames stable /usr/local/share/fai/package_config/*
```

6.9 Scripts in `/fai/scripts`

The default set of scripts in this directory is only an example. But they should do a reasonable job for your installation. You can edit them or add new scripts to match your local needs.

If a directory with a class name exists, all scripts matching `S[0-9]*` are executed in alphabetical order. So it's possible to use scripts of different languages (shell, cfengine, perl,..) for one class.

6.9.1 Shell scripts

Most scripts are Bourne shell scripts. Shell scripts are useful if the configuration task only needs to call some shell commands or create a file from scratch. In order not to write many short scripts, it's possible to distinguish classes within a script using the command `ifclass`. For copying files with classes, use the command `fcopy(8)`. If you like to extract an archive using classes, use `ftar(8)`. But now have a look at the scripts and see what they are doing.

6.9.2 Perl scripts

Currently no Perl script is used for modifying the system configuration.

6.9.3 Expect scripts

Currently no expect scripts are used for modifying the system configuration.

6.9.4 Cfengine scripts

Cfengine has a rich set of functions to edit existing configuration files, e.g `LocateLineMatching`, `ReplaceAll`, `InsertLine`, `AppendIfNoSuchLine`, `HashCommentLinesContaining`. But it can't handle variables which are undefined. If a variable is undefined, the whole cfengine

script will abort. Study the examples that are included in the fai package. More information can be found in the manual page `cfengine(8)` or at the cfengine homepage <http://www.cfengine.org>.

6.10 Changing the boot device

Changing the boot sequence is normally done in the BIOS setup. But you can't change the BIOS from a running Linux system as far as I know. If you know how to perform this, please send me an email. But there's another way of swapping the boot device of a running Linux system.

Change the boot sequence in the BIOS, so the first boot device is the local disk where the master boot record is located. The second boot device should be set to LAN or floppy disk, depending from which media you boot when the installation process is performed.

After the installation is performed, `lilo(8)` or `grub(8)` will write a valid boot sector to the local disk. Since it's the first boot device, the computer will boot the new installed system. If you like to perform an installation again, you have to disable this boot sector using the command `bootsector(8)`⁵. For more information use:

```
# bootsector -h
```

This is how to set up the a 3Com network card as second boot device, even if the BIOS doesn't support this. Enable LAN as first boot device in the BIOS.

```
Boot From LAN First: Enabled
Boot Sequence      : C only
```

Then enter the MBA setup of the 3Com network card and change it as follows:

```
Default Boot      Local
Local Boot        Enabled
Message Timeout   3 Seconds
Boot Failure Prompt Wait for timeout
Boot Failure      Next boot device
```

This will enable the first IDE hard disk as first boot device. If the boot sector of the hard disk is disabled, the computer will use the network interface as second boot device and boot from it. Maybe the disk partitioning tool can't work on such a disk. So you have to enable the boot sector before you want to partition the disk. If booting from a FAI floppy disk, another solution can be used to skip a re-installation if the BIOS is configured to boot from the floppy disk first and you are not here to remove the floppy disk : use

⁵The command `bootsector(8)` is part of the package `fai` and will be installed to `/usr/local/sbin` on the install clients.

```
# lilo -R ...
```

to instruct the FAI floppy to boot from the hard disk only once (see `lilo(8)`). Thus after this first reboot, the FAI floppy disk can be used for another FAI installation.

6.11 Hooks

Hooks let you specify functions or programs which are run at certain steps of the installation process. Before a default task is called, FAI searches for existing hooks for this task and executes them. As you might expect, classes are also used when calling hooks. Hooks are executed for every defined class. You only have to create the hook with the name for the desired class and it will be used. If `debug` is included in `$FAI_FLAG` the option `-d` is passed to all hooks, so you can debug your own hooks. If some default tasks should be skipped, use the subroutine `skiptask` and a list of default tasks as parameters. The example `partition.DISKLESS` skips some default tasks.

The directory `/fai/hooks/` contains all hooks. The file name of a hook consists of a hook name as a prefix and a class name, separated by a dot. The prefix describes the time when the hook is called, if the class is defined for the install client. For example, the hook `partition.DISKLESS` is called for every client belonging to the class `DISKLESS` before the local disks would be partitioned. If it should become a diskless client, this hook can mount remote filesystems via NFS and create a `/tmp/fai/fstab`. After that, the installation process will not try to partition and format a local hard disk, because a file `/tmp/fai/fstab` already exists.

A hook of the form `hookprefix.classname` can't define variables for the installation script, because it's a subprocess. But you can use any binary executable or any script you wrote. Hooks that have the suffix `.source` (e.g. `partition.DEFAULT.source`) must be Bourne shell scripts and are sourced. So it's possible to redefine variables for the installation scripts.

In the first part of `fai`, all hooks with prefix `confdir` are called. Since the configuration directory `/fai` is mounted in the default task `confdir`, the hooks for this task are the only hooks located in `$nfsroot/fai/hooks` on the install server. All other hooks are found in `/usr/local/share/fai/hooks` on the install server. All hooks that are called before classes are defined can only use the following classes: `DEFAULT` `$HOSTNAME` `LAST`. If a hook for class `DEFAULT` should only be called if no hook for class `$HOSTNAME` is available, insert these lines to the default hook:

```
hookexample.DEFAULT:

#! /bin/sh

# skip DEFAULT hook if a hook for $HOSTNAME exists
scriptname=$(basename $0 .DEFAULT)
[-f /fai/hooks/$scriptname.$HOSTNAME ] && exit
# here follows the actions for class DEFAULT
```

.

.

Some examples for what hooks could be used:

- Use `ssh` in the very beginning to verify that you mounted the configuration from the correct server and not a possible spoofing host.
- Do not mount the configuration directory, instead get a compressed archive via HTTP or from floppy disk and extract it into a new ram disk, then redefine `$FAI_LOCATION`.
- Load kernel modules before classes are defined in `/fai/class`.
- Send an email to the administrator if the installation is finished.
- Install a diskless client and skip local disk partitioning. See `hooks/partition.DISKLESS`.

6.12 Looking for errors

If the client can't successfully boot from the network card, use `tcpdump(8)` to look for Ethernet packets between the install server and the client. Search also for entries in several log files made by `in.tftpd(8)` and `bootpd(8)`:

```
egrep "tftpd|bootpd" /var/log/*
```

If the installation process finishes, the hook `faiend.LAST` searches all log files for common errors and write them to the file `error.log`. So, you should first look into this file for errors. To be sure, you should look for errors in all log files.

Sometimes the installation seems to stop, but there's only a postinstall script of a software package that requires manual input from the console. Change to another virtual terminal and look which process is running with tools like `top(1)` and `ps(1)`. You can add `debug` to `FAI_FLAGS` to make the installation process show all output from the postinst scripts on the console and get its input also from the console. Don't hesitate to send an email to the mailing list or to `<fai@informatik.uni-koeln.de>` if you have any questions. Sample log files from successful installed computers are available on the FAI homepage.

Chapter 7

How to build a Beowulf cluster using FAI

This chapter describes the details about building a Beowulf cluster using Debian GNU/Linux and FAI. For more information about the Beowulf concept look at <http://www.beowulf.org>.

7.1 Planning the Beowulf setup

The example of a Beowulf cluster consists of one master node and 25 clients. A big rack was assembled which all the cases were put into. A keyboard and a monitor, which are connected to the master server most of the time, were also put into the rack. But since we have very long cables for monitor and keyboard, they can also be connected to all nodes if something has to be changed in the BIOS, or when looking for errors when a node does not boot. Power supply is another topic you have to think about. Don't connect many nodes to one power cord and one outlet. Distribute them among several breakout boxes and outlets. And what about the heat emission? A dozen nodes in a small room can create too much heat, so you will need an air conditioner. Will the power supplies of each node go to stand by mode or are all nodes turned on simultaneously after a power failure?

All computers are connected to a Fast Ethernet switch. The master node (or master server) is called *nucleus*. It has two network cards. One for the connection to the external Internet, one for the connection to the internal cluster network. If connected from the external Internet, it's called *nucleus*, but the cluster nodes access the master node with the name *atom00*, which is a name for the second network interface. The master server is also the install server for the computing nodes. A local Debian mirror will be installed on the local harddisk. The home directories of all user accounts is also located on the master server. It will be exported via NFS to all computing nodes. NIS will be used to distribute account, host, and printer information to all nodes.

All client nodes *atom01* to *atom25* are connected via the switch with the second interface card of the master node. They can only connect to the other nodes or the master, but can't commu-

nicate to any host outside their cluster network. So, all services (NTP, DNS, NIS, ...) must be available on the master server. I choose the class C network address *192.168.42.0* for building the local Beowulf cluster network. You can replace the subnet 42 with any other number you like. If you have more than 253 computing nodes, choose a class A network address (10.X.X.X).

In the phase of preparing the installation, you have to boot the first install client many times, until there's no fault in your configuration scripts. Therefore you should have physical access to the master server and one client node. If you have little space, connect both computers to a switch box, so one keyboard and monitor can be shared among both.

7.2 Set up the master server

The master server will be installed by hand if it is your first computer installed with Debian. If you already have a Debian host running, you can also install it via FAI. Create a partition on */files/scratch* for the local Debian mirror with more than 5.0GB space available.

7.2.1 Set up the network

Add the following lines for the second network card to */etc/network/interfaces*:

```
# Beowulf cluster connection
auto eth1
iface eth1 inet static
address 192.168.42.250
netmask 255.255.255.0
broadcast 192.168.42.255
```

Add the IP addresses for the client nodes. The FAI package has an example for this */etc/hosts* file:

```
# Beowulf nodes
# atom00 is the master server
192.168.42.250 atom00
192.168.42.1 atom01
192.168.42.2 atom02
```

You can give the internal Beowulf network a name when you add this line to */etc/networks*:

```
beowulfcluster 192.168.42.0
```

Activate the second network interface with: */etc/init.d/networking start*.

7.2.2 Setting up NIS

Add a normal user account *tom* which is the person who edits the configuration space and manages the local Debian mirror:

```
# adduser tom
# addgroup linuxadmin
```

This user should also be in the group *linuxadmin*. So, add a line to `/etc/group`:

```
linuxadmin:x:101:tom
```

To initialize the master server as NIS server call `ypinit -m`. Then, copy the file `netgroup` from the examples directory to `/etc` and edit other files there. Adjust access to the NIS service.

```
# cat /etc/ypserv.securenets
# Always allow access for localhost
255.0.0.0      127.0.0.0
# This line gives access to the Beowulf cluster
255.255.255.0 192.168.42.0
```

Rebuild the NIS maps:

```
# cd /var/yp; make
```

7.2.3 Create a local Debian mirror

Now the user *tom* can create a local Debian mirror on `/files/scratch/` using `mkdebmirror`. This will need about 5.0GB GB disk space for Debian 3.0 (aka woody). Export this directory to the `netgroup@faiclients` read only.

7.2.4 Install FAI package on the master server

Add the following packages to the install server:

```
nucleus:/# apt-get install ntp tftp bootp nfs-kernel-server fai fai-kernels
nucleus:/# tasksel -q -n install dns-server
nucleus:/# apt-get dselect-upgrade
```

Configure NTP so that the master server will have the correct system time.

It's very important to use the internal network name *atom00* for the master sever (not the external name *nucleus*) in `/etc/bootptab` and `/etc/fai/fai.conf`. Replace the strings FAISERVER with *atom00* in `/etc/bootptab` and uncomment the following line in `/etc/fai/fai.conf` so the Beowulf nodes can use the name for connecting their master server.

```
NFSROOT_ETC_HOSTS="192.168.42.250 atom00"
```

```
/etc/bootptab:
```

```
.
.
.faillocal:\
    :tc=.faiglobal:\
    :sa=atom00:\
    :ts=atom00:\
    :Tl72="verbose createvt sshd":\
    :sm=255.255.255.0:\
    :gw=192.168.42.250:\
    :dn=beowulf.debian.org:\
    :ds=192.168.42.250:\
    :ys=atom00:ydnisnucleus:\
    :nt=atom00:
.
.
```

7.2.5 Prepare network booting

Uncomment the following line in `/etc/inetd.conf`:

```
#bootps dgram udp wait root /usr/sbin/bootpd bootpd -i -t 120
```

and restart the `inetd` daemon. The user `tom` should have permission to create the symlinks for booting via network card, so change the group and add some utilities.

```
# chgrp -R linuxadmin /boot/fai; chmod -R g+rx /boot/fai
# cp /usr/share/fai/utils/* /usr/local/bin
```

Now, the user `tom` can create a symlink in `/boot/fai` using

```
> tlink atom_install atom01
```

to boot the first client node for the first time. Then start to adjust the configuration for your client nodes. Don't forget to build the kernel for the cluster nodes using `make-kpkg(8)` and store it in `/usr/local/share/fai/files/packages`.

7.3 Tools for Beowulf clusters

The following tools for a Beowulf cluster are now available in `/usr/local/bin`:

mlink Change the symbolic link that points to the kernel image for booting from a network card.

all_hosts Print a list of all hosts, print only the hosts which respond to a ping or the hosts which do not respond. The complete list of hosts is defined by the netgroup `allhosts`. Look at `/usr/share/doc/fai/examples/etc/netgroup` for an example.

rshall Execute a command on all hosts which are up via rsh. Uses `all_hosts` to get the list of all hosts up. You can also use the `dsh(1)` command (dancer's shell, or distributed shell).

rgang For a huge cluster try `rgang`. It's a tool which executes commands on or distributes files to many nodes. It uses an algorithm to build a tree-like structure to allow the distribution processing time to scale very well to 1000 or more nodes (available at <http://fermitools.fnal.gov/abstracts/rgang/abstract.html>).

jmon For observing the resources of all clients (CPU, memory, swap, ...) you can use `jmon(1)` which installs a simple daemon on every cluster node.

ganglia This toolkit is very good for monitoring your cluster. Available at <http://ganglia.sourceforge.net/>

But there are a lot of other tools available which are not yet included in a Debian package.

7.4 Wake on LAN with 3Com network cards

Wake on LAN is a very nice feature to power on a computer without having physical access to it. By sending a special ethernet packet to the network card, the computer will be turned on. The following things have to be done, to use the wake on LAN (WOL) feature.

1. Connect the network card to the Wake-On-LAN connector on the motherboard using a 3 pin cable.
2. My ASUS K7M motherboard has a jumper called `Vaux` (`3VSBSLT`) which allows to select the voltage supplied to add-in PCI cards. Set it to `Add 3VSB` (3 Volt stand by).
3. Turn on the wake on LAN feature in BIOS
4. For a 2.2 kernel you have to use the following driver: <http://www.uow.edu.au/~andrewm/linux/\#3c59x-bc>

There's a little problem to enable the wake on LAN feature with a 2.2.19 kernel and a 3Com 3C905C network card. You have to use a patched 3c59x driver. But I managed to get it to work. Download the file `poll-ioct1-2.2.18-pre16.c.gz` and copy it to kernel sources to `drivers/net/3c59x.c`. Then make a new kernel package and install this new kernel. To wake up a computer use the tool `etherwake` (in `woody`) or get the single C source file. For more information look at <http://www.scyld.com/expert/wake-on-lan.html>.

Chapter 8

FAI on SUN SPARC hardware running Linux

Although FAI is architecture independent, there are some packages which are only available for certain architectures (e.g. silo, sparc-utils). SUN SPARC computers can boot from their boot prompt and don't need a boot floppy. To boot a SUN use:

```
boot net
```

You have to convert the kernel image from ELF format to a.out format. Use the program `elftoaout` (mentioned in the FAQ). The symlink to the kernel image to be booted is not the host name. Look at the FAQ at <http://www.ultralinux.org> for more information. A success report is available at <http://www.opossum.ch/fai/> and a HOWTO can be found at <http://toolbox.rutgers.edu/~amurphy/fai>.

Chapter 9

Various hints

This chapter has various hints which may not always be explained in great detail.

When using HTTP access to a Debian mirror, the local /var partition on all install clients must be big enough to keep the downloaded Debian packages. Do not try with less than 250 Mbytes unless you know why.

You can merge two directories which contain configuration information, if one is a global one, and the other a local one. We use it to merge the temples from the fai package, and our local configuration, which contains encrypted passwords and other information that should not be readable for others. This is how our setup looks like.

We have a local configuration space located in `~admin/additional-fai/` which contains following files:

```
./files
./files/etc
./files/etc/hosts
./files/etc/hosts/NUERBURG1
./files/etc/hosts/NUERBURG2
./files/etc/network
./files/etc/network/interfaces
./files/etc/network/interfaces/NUERBURG1
./files/etc/network/interfaces/NUERBURG2
./files/etc/bootptab
./files/etc/bootptab/kueppers
./files/etc/kueppers.tar.gz
./files/packages
./files/packages/kernel-image-2.4.20-cskoeln_2_i386.deb
./files/packages/cloop-2.4.20-cskoeln_0.63.1-4+2_i386.deb
./files/packages/xv-doc_3.10a-26_all.deb
./files/packages/xv_3.10a-26_i386.deb
./files/packages/Packages.gz
./files/packages/ltmodem-2.4.20_8.26a9_i386.deb
```

```
./files/packages/cloop-2.4.20-acer_0.63.1-4+1_i386.deb
./files/packages/kernel-image-2.4.20-acer_1_i386.deb
./files/packages/pcmcia-modules-2.4.20-acer_3.1.33-6+1_i386.deb
./files/packages/kernel-headers-2.4.20-acer_1_i386.deb
./files/packages/ltmodem-2.4.20-acer_8.26a9_i386.deb
./files/packages/debmirror_20020427-1_all.deb
./files/usr
./files/usr/local
./files/usr/local/ACROREAD5.tar.gz
./files/usr/local/share
./files/usr/local/share/LPRng
./files/usr/local/share/LPRng/pcfilter
./files/usr/local/share/LPRng/pcfilter/NISLPRCLIENT
./files/usr/lib
./files/usr/lib/mozilla
./files/usr/lib/mozilla/CS_KOELN.tar.gz
./class
./class/dom.var
./class/NET_9.var
./mk-packages-gz
./scripts
./scripts/kueppers
./README
./disk_config
./disk_config/dom
./disk_config/kueppers
```

The file `mk-packages-gz` is just the simple script which creates the `Packages.gz` as explained above. In order to copy this local configuration data into the `fai` config space we use this command:

```
cp -a ~admin/additional-fai/* /usr/local/share/fai
```

If you remove a file in you local configuration, do not forget to remove this file also in the configuration space, otherwise it will still be used.

After calling `set-dis-kinfo`, a list of all local hard disks is stored to `$disklist` and `$device_size` contains a list of disk devices and their sizes.

Use `fai-divert -a` if a `postinst` script calls a configuration program, e.g. the `postinst` script for package `apache` calls `apacheconfig`, which needs manual input. You can fake the configuration program so the installation can be fully automatic. But don't forget to use `fai-divert -R` to remove all faked script.

During the installation you can execute commands inside the newly installed system in a `chroot` environment by using `chroot /tmp/target` or just `$ROOTCMD` followed by the command you want to call; for example `$ROOTCMD dpkg -l` shows the packages installed on the new system.

The only task which has to be done manually for new hardware is to assign the MAC address to a hostname and to an IP address, and to define classes for this host if the existing configuration files are not generic enough to deal with this new host.

There's a tradeoff between writing a few large configuration scripts, or many short scripts, one for each class. Large scripts can distinguish classes by using case statements, the `ifclass` test or with class mechanisms for `cfengine` scripts.

If your computer can't boot from the network card, you do not always need to boot from floppy. Define a partition `/fai-boot` in your `disk_config` configuration file. Then the class `FAI_BOOTPART` will automatically be defined and will create a lilo entry for booting the FAI bootfloppy from this partition. So you can start the re-installation without a boot floppy. This will also make the test phase shorter, since booting from hard disk is much faster than booting from floppy.

To use the `/file/scratch` partitions on all Beowulf nodes, use the kernel automounter and the following configuration.

```
nucleus[~]> cat /etc/auto.master
/scratch /etc/nodes.scratch

nucleus[~]> cat /etc/nodes.scratch
*                -rw,soft,intr          &:/files/scratch
```

9.1 Useful functions for advanced administrators

fai-divert Add or remove a file to the list of diversions and replace the file with a dummy script. This is useful when a `postinst` script needs manual input. At the end of the installation all diversions are removed.

skiptask This given list of tasks are skipped. For use e.g. in `partition.DISKLESS`.