

# Vollautomatische Installationen mit FAI

***MPI für Physik, München, November 2008***

Thomas Lange, Uni Köln

lange@informatik.uni-koeln.de

- ▷ Warum automatisch installieren?
- ▷ Crashtest
- ▷ Wie funktioniert FAI?
- ▷ Nachteile der Automatisierung
- ▷ Sicherheit in und mit FAI
- ▷ Erfahrungen mit FAI

▷ whoami

- > Informatik in Bonn studiert
- > Systemadministrator seit 16 Jahren
- > SunOS 4.1.1 auf SPARC
- > Solaris Jumpstart
- > 1999 erstes 16 Knoten Cluster (Dual PII 400MHz)
- > Debian Entwickler seit 2000

# Manuelle Installation?

Wer möchte diese Rechner per Hand installieren?



168 IBM HS20 Blades, 2x2.8 GHz



90 dual Itanium 2, 900Mhz



[www.centibots.org](http://www.centibots.org)

# Was ist ein Linux Rollout?

---

- Geplante Installation
- Aufsetzen von OS und Anwendungen
- Unterschiedliche Anforderungen (Konfigurationen) unterstützen
- Zentrale Verwaltung, Steuerung

# Was ist ein Linux Rollout?

---

- ▶ Automatische Dokumentation
- ▶ Inventarisierung
- ▶ Disaster recovery
- ▶ Computer Infrastruktur



# Manuelle Installation?

Und diese?



180 dual AMD MP2200, Max Planck Institute for Gravitational Physics

## Administration - bei Ihnen auch so?

---

- ▷ Kaufen, installieren, laufen lassen
- ▷ Gewachsene Systeme, sehr unterschiedliche Hard- und Software
- ▷ Immer wieder kleine Änderungen an jedem einzelnen System
- ▷ Änderungen nicht dokumentiert, da keine Zeit
- ▷ Nur Notfalladministration, da wenig Zeit
- ▷ Trotz schnellster Computer, wird zu viel per Hand gemacht
- ▷ Hauptproblem der Sysadmins: **Zuviel Handarbeit => Keine Zeit!**



- ▷ Was ist der Wert ihrer Computer?
- ▷ Was beinhalten ihre Rechner?
  - > Kundendaten
  - > Services
  - > Applikationen
  - > Eigenes Know-How
- ▷ Was passiert, wenn ihre Rechner einen Tag lang nicht laufen?
- ▷ Eine gute Computerinfrastruktur ist so wichtig wie ...?
- ▷ Wie sichern Sie diese Werte?
- ▷ Ist damit wirklich alles gesichert?

- ▶ Wählen Sie zufällig einen Rechner (ohne Backup vorher)
- ▶ Werfen sie den Rechner aus dem 10.Stock  
(oder `dd if=/dev/zero of=/dev/hda`)



- ▶ Stellen Sie alle Arbeit des Sysadmin innerhalb von 10 Minuten wieder her
- ▶ Schaffen Sie das?

## Fakten, die oft übersehen werden

---

- ▶ Gut laufende Rechner sind ihr Kapital
- ▶ Backup der Daten ist nur ein Teil
- ▶ Haben Sie eine Kopie ihres Sysadmins?
  
- ▶ Manuelle Installation = Schlechte Installation
- ▶ Manuelle Installation -> manuelle Nachbesserung -> keine Zeit
- ▶ Keine Updates, keine Patches -> keine Sicherheit!!!
- ▶ Manuelle Arbeit der Sysadmins -> hohe IT-Kosten

# Manuelle Installation?

Können Sie sicherstellen, dass diese Rechner alle gleich sind?



- ▷ Dauert viele Stunden
- ▷ Viele Fragen
- ▷ "No simple sysadmin task is fun more than twice"
- ▷ Wiederholende Arbeit ist stupide => Fehler
- ▷ Dokumentation fehlt, Reproduzierbarkeit?
- ▷ Jede Installation ist ungewollt einzigartig
- ▷ **Eine Installation per Hand skaliert nicht !**

# Warum voll automatisch?

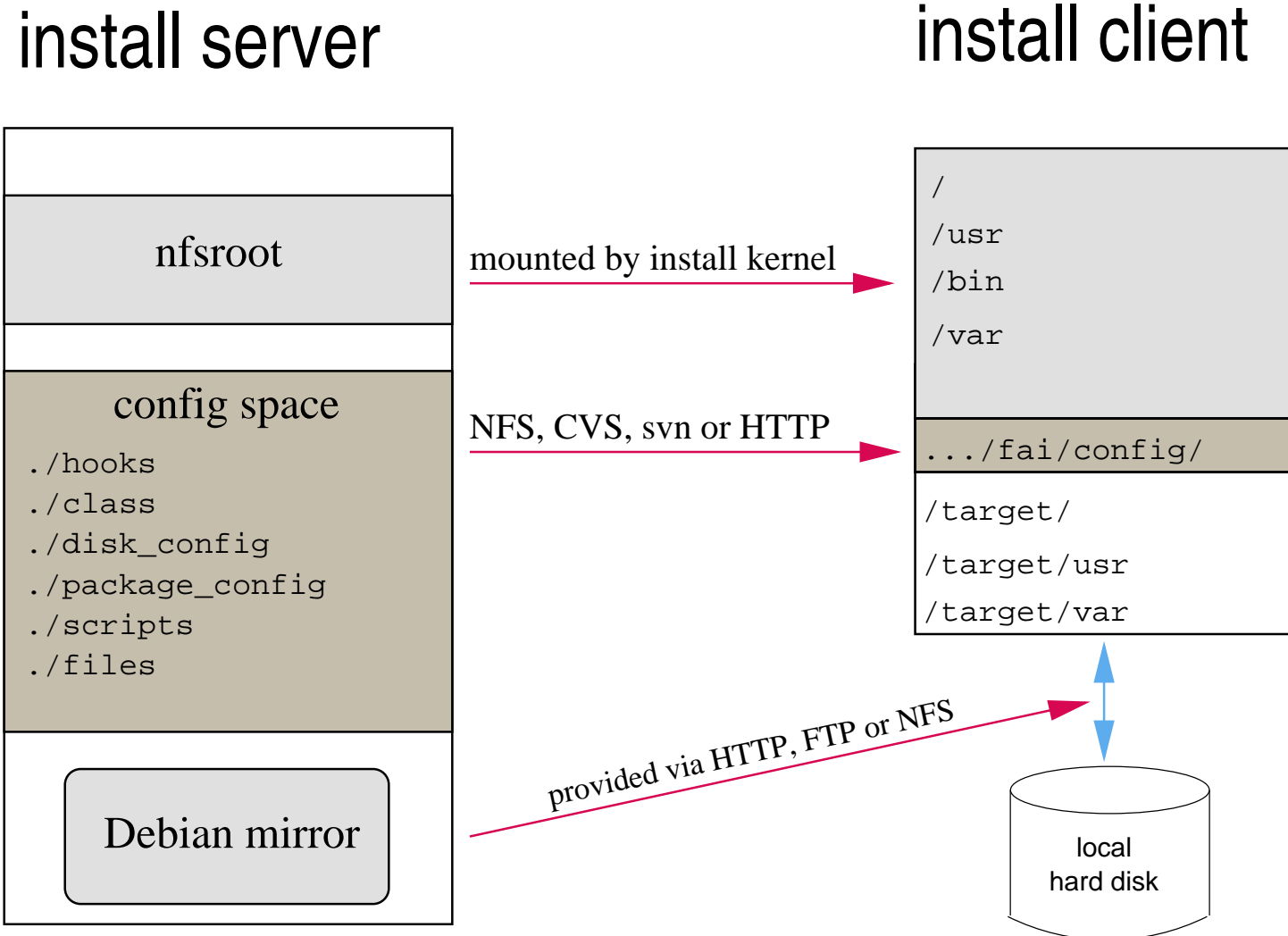
---

- ▷ Ist schnell
- ▷ Hilft nach Hardwaredefekt
- ▷ Garantiert identische Installationen
- ▷ Heterogene Hardware und unterschiedliche Konfigurationen
- ▷ Parallele Installationen
- ▷ Automatische Dokumentation
- ▷ Spart sehr viel Arbeit (= Zeit = Geld). ROI
- ▷ Macht mehr Spaß



- ▷ FAI macht alles, was ihr Systemadministrator zu tun hat, bevor der Benutzer das erste Mal auf einem neuen Rechner arbeiten kann
- ▷ Serverbasiertes Tool
- ▷ Skriptgesteuerte vollautomatische Installation
- ▷ Installiert und konfiguriert das Betriebssystem und alle Anwendungsprogramme
- ▷ Kein Master Image
- ▷ Modular durch Klassensystem
- ▷ Erweiterbar und flexibel durch hooks
- ▷ Es kann die Installation nicht planen :-), aber
- ▷ **Plane deine Installation und FAI installiert deinen Plan! :-)**

# Wie funktioniert FAI ?



- Die Konfiguration liegt auf dem Install server
- Die Installation läuft auf dem Klienten

- ▶ Installserver mit DHCP, NFS und TFTP
- ▶ Client bootet via PXE, , CD-ROM oder USB Stick
- ▶ Lokaler Spiegel von Debian (NFS, FTP oder HTTP)
- ▶ Plattenplatz auf dem Server:

FAI Paket	<1 MB	Skripte, Konfigurationsdateien
nfsroot	390 MB	erzeugt mit <code>make-fai-nfsroot</code>
Debian Spiegel	<21 GB	Debian 5.0 (lenny, nur i386)
- ▶ Alle Install Clients nutzen die gleichen Verzeichnisse
- ▶ **Konstanter Plattenplatz**

# Ablauf einer Installation I

---

- ▷ Plane deine Installation!
- ▷ Booten via PXE und Kernel mit initrd via TFTP holen
- ▷ Rechner startet als Diskless Client
- ▷ Hardwareerkennung und Kernel Module laden

- ▷ Klassen und Variablen definieren
- ▷ Festplatten partitionieren
- ▷ Dateisysteme erzeugen und mounten
- ▷ Software Pakete installieren
- ▷ Betriebssystem und Anwendungen konfigurieren
- ▷ Protokolldateien lokal und auf Install Server speichern
- ▷ Neu installiertes System booten

- ▷ Ein Rechner gehört zu mehreren Klassen
- ▷ Priorität von niedrig nach hoch
- ▷ Beispiel: DEFAULT FAIBASE GRUB GNOME demohost LAST
- ▷ Alle Teile der Installation nutzen das Klassenkonzept
- ▷ Konfigurationsdateien werden anhand der Klassennamen ausgewählt
- ▷ `fcopy` zum Kopieren von Templates
  
- ▷ Erfahrener Admin kreiert die Klassen
- ▷ Junior Admin ordnet die Klassen den Rechnern zu
- ▷ PC installiert sich selber



# Verzeichnisse im Config Space

---

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

**Beispiel:** `.../class/10-base-classes:`

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

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

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

ifclass I386 && echo "GRUB"

lspci | grep -q MATROX || echo "MATROX"
```

**Beispiel:** `.../class/FAIBASE.var:`

```
FAI_ALLOW_UNSIGNED=1
```

```
CONSOLEFONT=
```

```
KEYMAP=KEYMAP=de-latin1-nodeadkeys
```

```
UTC=yes
```

```
TIMEZONE=Europe/Berlin
```

```
ROOTPW=' $1$kBnWcO.E$djxB128U7dMkr1tJHPf6d1 '
```

```
LOGUSER=fai
```

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

- ▶ Eigene Variablen möglich
- ▶ Die Konfigurationsskripte in `.../scripts/*` nutzen diese Variablen

# Neue Plattenpartitionierung

---

**Beispiel:** `.../disk_config/FAIBASE:`

```
disk_config disk1      preserve:9  bootable:1

primary  /           150-300    ext3  rw,errors=remount-ro
logical  swap          40-500     swap  rw
logical  /usr           200-4000   ext3  rw
logical  /var           90-1000    ext3  rw
logical  /tmp           50-1000    ext3  rw
logical  /home         50-        ext3  defaults
```

▶ **Filesysteme: ext2, ext3, vfat, xfs, ReiserFS, NTFS**

## Endlich!

```
disk_config disk1
```

```
primary /boot 20-100      ext3      rw
primary  swap 1024         swap      sw
primary /    2000-4000    ext3      rw,acl,user_xattr
logical  -    0-             - -
logical  -    0-             - -
logical  -    0-             - -
logical  -    0-             - -
```

```
disk_config raid
```

```
raid1   -    disk1.5,disk1.7    - -
raid1   -    disk1.6,disk1.8  - -
```

```
disk_config lvm
```

```
vg volg1 md0,md1
```

```
volg1-usr /usr          2048 ext3 rw createopts="-O dir_index,resize_inode"
volg1-var /var          600  ext3 rw createopts="-O dir_index,resize_inode"
volg1-hl  /home/local  4096 ext3 rw,acl,user_xattr,noexec,nosuid,nodev
volg1-es  /export/sites 2048 ext3 rw createopts="-O none"
volg1-v   /vservers     2048 ext3 rw createopts="-O ^dir_index,^resize_inode"
```

**Beispiel:** `.../package_config/BEOWULF:`

```
# packages for Beowulf clients
```

```
PACKAGES install BEOULF_MASTER  
gmetad apache
```

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

```
dsh update-cluster-hosts update-cluster etherwake
```

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

- ▶ **Aktionen** `aptitude`, `apt-get`, `smart`, `rpm`, `urpmi`, `y2pms`, `yast`, `yum`
- ▶ **Abhängigkeiten** innerhalb der Pakete werden aufgelöst



# Verschlaufpause



290 workstations in 19 Klassenräumen an 6 Orten mit dualboot,  
124 Debian Server (incl. Xen),... University of West Bohemia

# Verzeichnisse im Config Space

---

```
|-- scripts/
|   |-- BOOT
|   |-- FAIBASE/
|       |-- 10-misc           Bourne shell script
|       |-- 30-interface     Bourne shell script
|       `-- 40-misc          /usr/bin/cfengine script
|   |-- DEMO/
|       |-- 10-misc           Bourne shell script
|       `-- 30-demo          /usr/bin/cfengine script
|   `-- demohost
`- files/
    |-- etc/
        |-- X11/
            |-- xorg.xonf/     fcopy /etc/X11/xorg.conf
                |-- FAIBASE
                |-- MATROX
                `-- demohost
```

# Konfigurationsskripte

---

```
# create NIS/NONIS config
fcopy -M /etc/nsswitch.conf /etc/host.conf
fcopy -i /etc/ypserv.securenets # only for yp server
ifclass NONIS && rm -f $target/etc/defaultdomain
if ifclass NIS; then
    echo $YPPDOMAIN > $target/etc/defaultdomain
    rm -f $target/etc/yp.conf
    for s in $YPSRVR; do
        ainsl -av $target/etc/yp.conf "ypserver $s"
        # don't do this! # echo "ypserver $s" >> $target/etc/yp.conf
    done
fi

ainsl -v $target/etc/fstab "$bserver:/usr/local /usr/local nfs ro 0 0"

ifclass USR_LOCAL_COPY && {
    mount -o ro $bserver:/usr/local /usr/local
    cp -a /usr/local $target/usr
}
fcopy -M /etc/X11/xorg.conf
```

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

editfiles:
  any::
    { ${target}/etc/fstab
      AppendIfNoSuchLine "none /proc/bus/usb usbdevfs defaults"
      AppendIfNoSuchLine "/dev/fd0 /floppy auto users,noauto 0 0"
    }
    { ${target}/etc/inittab
      ReplaceAll "/sbin/getty" With "/sbin/getty -f /etc/issue.linuxlogo"
    }
  TERMINAL_CLIENT::
    { ${target}/etc/inetd.conf
      HashCommentLinesContaining "in.rlogin"
    }
```

## Installationszeiten

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

Knoten	Sekunden
1	337
5	340
10	345
20	379

12% mehr Zeit bei 20 Rechnern.

## Noch ein Beispiel



356 opterons, 80 xeons, Top500 in 11/2005,  
Trinity Centre for High Performance Computing, Dublin



# FAI Fragebogen

---

From: dot Owen . . . , ...@flektor-inc.com

How many hosts? about 600 hosts with 4 installs of fai

Which version of fai are you using? 3.2

Which version are you using as your install kernel? Basic Ubuntu kernel 2.6.24.1

Which types of hosts are you installing with fai?

Web servers, rails application servers, Mysql DB servers,  
FAI servers, DNS servers, LDAP servers, etc...

Do you manage your configuration space under version control? Yes, with puppetd

What is the business of your company? A web 2.0 media mashup tool  
[www.flektor.com](http://www.flektor.com)

Additional comments:

We install a lab environment, dev environment, and a production environment with an a and b side.

What is your business? The Sanger Institute, genome research

How many hosts? More than 540 in several clusters, and for different servers

Which version of fai are you using? 2.8.4

Which version of fai-kernels? We use a custom 2.6 kernel.

Main compute cluster

-----

168 IBM HS20 Blades (2x2.8 Ghz PIV, 32 bit OS)

280 IBM HS20 Blades (2x3.2 Ghz EMT64, 64bit OS)

Trace searchengine cluster: 48 IBM LS20 Blades, 64bit OS

Web servers: 28 IBM HS20 blades, 32bit OS

Misc servers: 20 HP DL585/385 servers, 2 or 4 CPU dual core Opteron, 64bit

Our deployed software stack is 725MB. Installation time on an LS20 blade is 153 seconds.



- ▷ City of Munich, several hundreds, (14.000 hosts planed)
- ▷ Albert Einstein Institute, Germany, 800+ hosts
- ▷ ComBOTS, 700 Blades, 650 Server (16GB RAM, 8TB disk)
- ▷ IFW-Dresden, Germany, 100+ hosts, cluster, (i386 and IA64)
- ▷ Physics department (FU Berlin), 139+ hosts
- ▷ Host Europe, 250 hosts
- ▷ Lycos Europe, search engine, >3000
- ▷ Ewetel, ISP and telco, 65 hosts at 3 locations
- ▷ HPC2N, 2 clusters listed in `top500.org`, 192 dual Opteron, 120 dual Athlon
- ▷ Electricité de France (EDF), France, 200 hosts
- ▷ MIT Computer science research lab, 200 hosts
- ▷ Stanford University, 450 hosts
- ▷ University of New Orleans, 72 node Beowulf cluster
- ▷ Brown University, Dep. of Computer Science, 300+ hosts
- ▷ University of West Bohemia, Czech Republic, 180+
- ▷ Netcologne, MPI Meteorologie, DESY, Genua, taz, thomas-krenn.com, mc-wetter.de

# Nachteile der Automatisierung

---

- ▷ Nicht alles kann oder soll automatisiert werden
- ▷ Fehler werden auch verteilt
- ▷ Weniger aber höher qualifiziertes Personal notwendig
- ▷ Man muss erstmal Zeit und Arbeit investieren
- ▷ Bereitschaft für Veränderungen?
- ▷ Sysadmin wird zum sauberem Arbeiten gezwungen
- ▷ Manuelle Änderungen an einzelnen Rechner sind verboten!

- ▷ Wer den Rechner booten kann ist root
- ▷ PXE ist unsicher, damit auch der ganze Rest
- ▷ TFTP ist nicht besser
- ▷ NFS ist kein Problem, da read-only und keine Geheimnisse enthalten
- ▷ NFSROOT ersetzen durch große initrd?
- ▷ Sicherheit durch VLAN, spezieller Port
- ▷ CD/USB Stick ist sicher wenn vertraulich
- ▷ Hohe Sicherheit kann nicht voll automatisch sein
- ▷ Verteilung von Geheimnissen via Passwort absichern

- ▷ Andere Distributionen, auch RPM
- ▷ Aufsetzen von chroot (z.B für Live CD's, grml)
- ▷ Xen Installation
- ▷ Softupdates

hostname	confdir	defclass	partition	extrbase	debconf	instsoft	configure	savelog	faiend	reboot
demohost	✓	✓	✓	✓	✓	✗	✓	✗	→	
atom03	✓	✓	✓	✓	✓	→				
atom01	✓	✓	✓	✓	✓	✓	✓	✓	→	
atom02	✓	✓	✓	✓	✓	✗	✗	✓	→	
gnomehost	✓	✓	✓	✓	✓	✓	✓	✓	→	

GOsa - Mozilla Firefox

Datei Bearbeiten Ansicht Gehe Lesezeichen Extras Hilfe

GOsa<sup>2</sup> [Hauptmenü](#) [Hilfe](#) [Abmelden](#) Angemeldet: **cajus**

## Automatische Installation

**Mein Konto**

- Allgemein
- UNIX
- Umgebung
- Mail
- Samba
- Konnektivität
- Fax
- Telefon
- Passwort

**Administration**

- Benutzer
- Gruppen
- Objektgruppen
- Abteilungen
- Anwendungen
- Systeme
- FAI
- Fax-Sperrlisten
- Telefon-Makros
- Telefon-Konferenzen

**Zusätzliches**

- Adressbuch
- Fax-Berichte
- Telefon-Berichte
- Systemprotokolle
- LDAP-Manager

**Liste benutzer Klassen**

Momentane Basis /

Name der FAI-Klasse	Typ der Klasse	Aktionen
BASECLIENT	Paketliste	
BUMS [Template test]	Vorlagen	
CDALPHA2 [Erweiterung des Prototyps mit Gosa]	Paketliste	
CDALPHA	Hooks	
CDALPHA	Paketliste	
CDALPHA [Partitionierung des Prototypen]	Partitionstabelle	
CDALPHA [Zusätzliche Konfigurationsskripte]	Skripte	
CDALPHA [Template Informationen]	Vorlagen	
CDALPHA	Variablen	
FAIBASE	Paketliste	
FAIBASE [Testpartitionierung]	Partitionstabelle	
FAIBASE	Variablen	
FSCLIFEBOOK	Hooks	
FSCLIFEBOOK	Paketliste	
FSCLIFEBOOK [Test]	Partitionstabelle	
FSCLIFEBOOK [Template Informationen]	Vorlagen	
GRUB [Install GRUB boot sector]	Skripte	
HALUTBASE	Paketliste	
HALUTBASE [Testpartitionierung]	Partitionstabelle	
KERNEL	Paketliste	
NOTEBOOK [Partitionierung des FSC Lifebook]	Partitionstabelle	
OPT-PACKAGES-NOTSAVE	Paketliste	

**Information**

Dieses Menü erlaubt es Ihnen, FAI-Klassen zu erstellen, entfernen und zu bearbeiten.

**Filter**

*	A	B	C	D	E	F	G	H	I	J
K	L	M	N	O	P	Q	R	S	T	U
V	W	X	Y	Z	0	1	2	3	4	5
6	7	8	9							

- Zeige Profile
- Zeige Vorlagen
- Zeige Skripte
- Zeige Hooks
- Zeige Variablen
- Zeige Pakete
- Zeige Partitionen

GOsa - Mozilla Firefox

Datei Bearbeiten Ansicht Gehe Lesezeichen Extras Hilfe

**GOsa<sup>2</sup>** [Hauptmenü](#) [Hilfe](#) [Abmelden](#) Angemeldet: **cajus**

**Mein Konto**

- Allgemein
- UNIX
- Umgebung
- Mail
- Samba
- Konnektivität
- Fax
- Telefon
- Passwort

**Administration**

- Benutzer
- Gruppen
- Objektgruppen
- Abteilungen
- Anwendungen
- Systeme
- FAI
- Fax-Sperrlisten
- Telefon-Makros
- Telefon-Konferenzen

**Zusätzliches**

- Adressbuch
- Fax-Berichte
- Telefon-Berichte
- Systemprotokolle
- LDAP-Manager

## Automatische Installation

cn=FAIBASE,ou=disk,ou=fai,ou=configs,ou=systems,dc=gonicus,dc=de

### Partitionen

**Gerät**

Name \*  Beschreibung

**Partitions-Einträge**

Typ	Dateisystem	Mount-Punkt	Größe in MB	Mount-Optionen	Dateisystem-Option	Bewahren	
<input type="text" value="primary"/>	<input type="text" value="ext3"/>	<input type="text" value="/"/>	<input type="text" value="2048-4096"/>	<input type="text"/>	<input type="text" value="-j"/>	<input type="checkbox"/>	<input type="button" value="Entfernen"/>
<input type="text" value="primary"/>	<input type="text" value="swap"/>	<input type="text" value="swap"/>	<input type="text" value="512"/>	<input type="text"/>	<input type="text"/>	<input type="checkbox"/>	<input type="button" value="Entfernen"/>
<input type="text" value="logical"/>	<input type="text" value="ext2"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input checked="" type="checkbox"/>	<input type="button" value="Entfernen"/>

- ▷ Mehr als 200 detaillierte Berichte von Benutzern
- ▷ FAI läuft auf i386, amd64, IA64, SPARC, PowerPC, ALPHA
- ▷ Ubuntu, Mandrake, Suse, ...
- ▷ Installiert auch Solaris 9 auf SUN Sparc
- ▷ FAI-CD für i386 und amd64
- ▷ GOsa als graphischen Frontend
- ▷ 12k Zeilen Source code (ohne Dokumentation)
- ▷ Beispiel Konfiguration ca. 1300 Zeilen



- ▷ Homepage: `http://www.informatik.uni-koeln.de/fai`
- ▷ Wiki: `http://faiwiki.informatik.uni-koeln.de`
- ▷ Zwei Maillinglisten, IRC Channel
- ▷ Subversion Zugriff, Beispiele der Log Dateien
- ▷ 9 Jahre FAI, Erfahrung, Rückmeldungen, Patches durch Benutzer
- ▷ Regelmäßige Entwicklertreffen
- ▷ Kommerzieller Support: z.B. `fai-cluster.de`

**Plane deine Installation  
und FAI installiert  
deinen Plan!**