

Linux Rollout mit FAI

Linux Forum, Cebit 2005

Thomas Lange, Uni Köln

lange@informatik.uni-koeln.de

- ▷ Warum automatisch installieren?
- ▷ Wie funktioniert FAI?
- ▷ Vergleich: debian installer – FAI
- ▷ Erfahrungen mit FAI

Was ist ein Linux Rollout?

- ▶ Viele Pinguine in kurzer Zeit kontrolliert aussetzen
- ▶ Geplante Installation
- ▶ Automatisches Aufsetzen von OS und Anwendungen
- ▶ Inventarisierung der Hardware
- ▶ Zentrale Verwaltung der Systeme
- ▶ Unterschiedliche Anforderungen (Konfigurationen) unterstützen
- ▶ Schnelle Wiederherstellung nach Defekt
- ▶ Automatische Dokumentation
- ▶ Computer Infrastruktur, keine einzelnen Rechner
- ▶ Sysadmin wird zum Infrastruktur Architekt

Administration - bei Ihnen auch so?

- ▶ Kaufe ein Rechner, installiere das System, lass es ewig laufen
- ▶ Gewachsene Systeme, sehr unterschiedliche Hard- und Software
- ▶ Zwischendurch viele 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

Administration - bei Ihnen auch so?

- ▶ Kaufe ein Rechner, installiere das System, lass es ewig laufen
- ▶ Gewachsene Systeme, sehr unterschiedliche Hard- und Software
- ▶ Zwischendurch viele 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 passiert, wenn ihre Rechner einen Tag lang nicht laufen?
- ▷ Eine gute Computerinfrastruktur ist so wichtig wie ...?

- ▷ Was ist der Wert ihrer Computer?
- ▷ Was passiert, wenn ihre Rechner einen Tag lang nicht laufen?
- ▷ Eine gute Computerinfrastruktur ist so wichtig wie ...?
- ▷ Was beinhalten ihre Rechner?
 - > Kundendaten (Kontakte, Aufträge, Rechnungen)
 - > Services (E-Mail, Web, Datenbanken, Drucken)
 - > Applikationen (Textverarbeitung, Compiler, CAD, Tools)
 - > Input und Output (CAD Entwürfe, Simulationsergebnisse)
 - > Internes Firmen Know-How (Source Code)

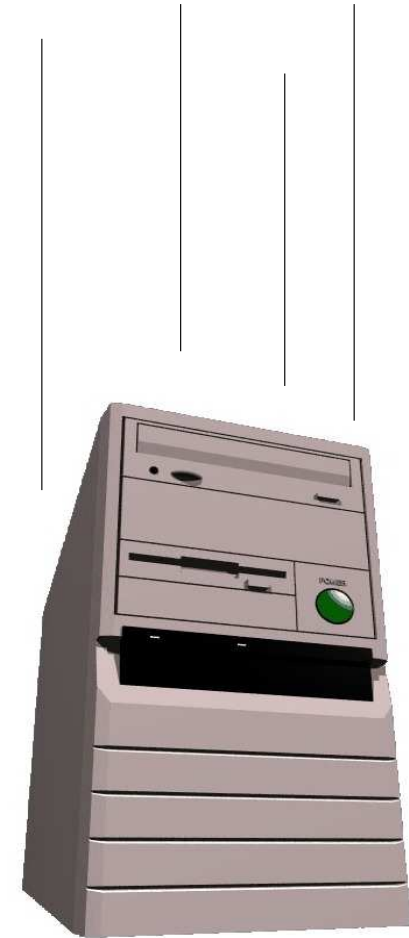
- ▷ Was ist der Wert ihrer Computer?
- ▷ Was passiert, wenn ihre Rechner einen Tag lang nicht laufen?
- ▷ Eine gute Computerinfrastruktur ist so wichtig wie ...?
- ▷ Was beinhalten ihre Rechner?
 - > Kundendaten (Kontakte, Aufträge, Rechnungen)
 - > Services (E-Mail, Web, Datenbanken, Drucken)
 - > Applikationen (Textverarbeitung, Compiler, CAD, Tools)
 - > Input und Output (CAD Entwürfe, Simulationsergebnisse)
 - > Internes Firmen Know-How (Source Code)
- ▷ 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`)

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



- ▶ 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
- Datensicherung ist nur ein Teil
- Auch die Arbeit der Sysadmins muss gesichert sein

Fakten, die oft übersehen werden

- Gut laufende Rechner sind ihr Kapital
- Datensicherung ist nur ein Teil
- Auch die Arbeit der Sysadmins muss gesichert sein
- Manuelle Installation = Schlechte Installation (siehe später)
- Schlechte Installation -> unbrauchbare Rechner
- Schlechte Installation -> unproduktive Benutzer
- Manuelle Installation -> manuelle Nachbesserung -> keine Zeit
- Keine Updates, keine Patches -> keine Sicherheit!!!
- Manuelle Arbeit der Sysadmins -> hohe IT-Kosten

Manuelle Installation?

Wer möchte diese Rechner per Hand installieren?



20 nodes dual XEON, 2.4 GHz



90 dual Itanium 2, 900Mhz



www.centibots.org

Manuelle Installation?

Können Sie sicherstellen, dass diese Rechner alle gleich sind?



- Dauert viele Stunden
- Viele Fragen sind zu beantworten
- Gleiche Daten für jeden Rechner erneut eingegeben
- Wiederholende Arbeit ist stupide => Fehler
- "No simple sysadmin task is fun more than twice"
- Dokumentation fehlt, Reproduzierbarkeit?
- Jede Installation ist ungewollt einzigartig
- **Ein Installation per Hand skaliert nicht !**

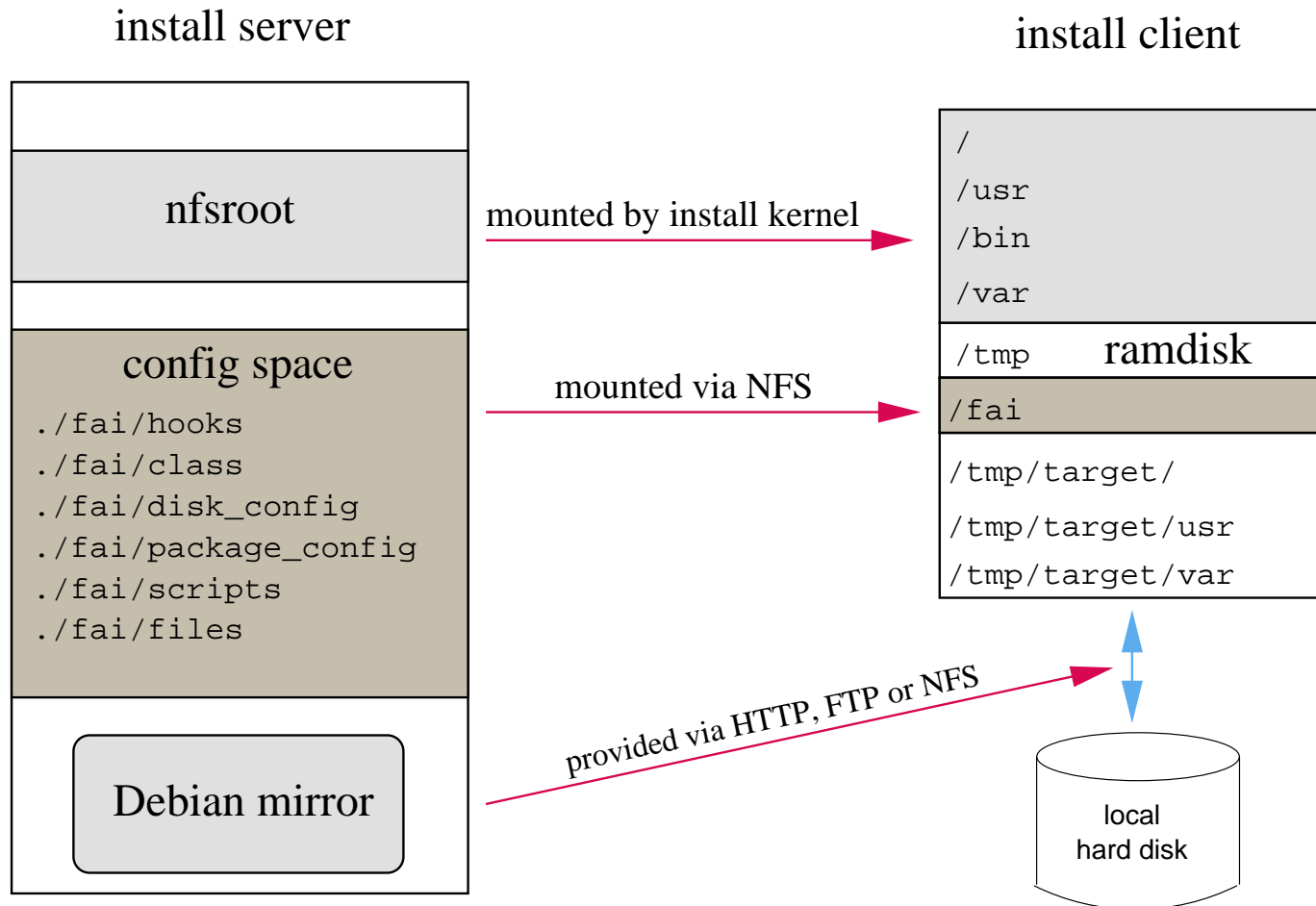
Warum nicht voll automatisch?

- ▶ Dauert nur wenige Minuten
- ▶ Schnelle Wiederherstellung nach Hardwaredefekt (Disaster recovery)
- ▶ Identische Installationen garantiert (auch nach Monaten)
- ▶ Heterogene Hardware und unterschiedliche Konfigurationen einfach
- ▶ Gleichzeitige Installation vieler Rechner
- ▶ Junior Admins können FAI anwenden
- ▶ Automatische Dokumentation
- ▶ Spart sehr viel Arbeit (= Zeit = Geld). ROI!
- ▶ Haben Sie einen Plan für ihre Computer Infrastruktur?

- FAI macht alles, was ihr Systemadministrator zu tun hat, bevor der Benutzer das erste Mal auf einem neuen Rechner arbeiten kann
- Serverbasiertes Tool
- Skripts steuern vollautomatische Installation
- Installiert und konfiguriert das Betriebssystem und alle Anwendungsprogramme
- Kein Master Image notwendig
- Sehr modular durch Klassensystem
- Einfach erweiterbar und flexibel durch eigene Skripte (hooks)
- Es kann die Installation nicht planen :-(), aber

- FAI macht alles, was ihr Systemadministrator zu tun hat, bevor der Benutzer das erste Mal auf einem neuen Rechner arbeiten kann
- Serverbasiertes Tool
- Skripts steuern vollautomatische Installation
- Installiert und konfiguriert das Betriebssystem und alle Anwendungsprogramme
- Kein Master Image notwendig
- Sehr modular durch Klassensystem
- Einfach erweiterbar und flexibel durch eigene Skripte (hooks)
- Es kann die Installation nicht planen :-(), aber
- **Plane die Installation und FAI installiert deinen Plan! :-)**

Wie funktioniert FAI ?



- Die Konfiguration liegt auf dem Install server
- Die Installation läuft auf dem Klienten
- Remotekontrolle während der Installation via ssh

- ▶ Server mit DHCP, NFS und TFTP (install server)
- ▶ Rechner mit Netzwerkkarte (install client)
- ▶ Booten geht aber auch von Floppy oder CD-ROM
- ▶ Nicht nötig: Floppy, CD-ROM, Tastatur, Grafikkarte
- ▶ Lokaler Spiegel von Debian (NFS, FTP oder HTTP)
- ▶ Plattenplatz auf dem Server:

FAI Paket	13 MB	Kernel, Skripte, Konfigurationsdateien
nfsroot	250 MB	erzeugt mit <code>make-fai-nfsroot</code>
Debian Spiegel	9 GB	Debian 3.1 (sarge, nur i386)
- ▶ Alle Install Clients nutzen die gleichen Verzeichnisse
- ▶ **Konstanter Plattenplatz**

▶ Plane deine Installation!

- Plane deine Installation!
- Booten via PXE und Kernel via TFTP holen
- Rechner startet vollständiges Linux, ohne lokale Platte zu benutzen
- 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 SMALL_IDE GRUB GNOME demohost LAST`
- Klassen werden über Skripte in `/fai/class` definiert
- Alle Teile der Installation nutzen das Klassenkonzept
- Konfigurationsdateien werden anhand der Klassennamen ausgewählt
- Mit `fcopy` wird klassenbasiert ein Template kopiert

- ▷ Ein Rechner gehört zu mehreren Klassen
- ▷ Priorität von niedrig nach hoch
- ▷ Beispiel: `DEFAULT SMALL_IDE GRUB GNOME demohost LAST`
- ▷ Klassen werden über Skripte in `/fai/class` definiert
- ▷ Alle Teile der Installation nutzen das Klassenkonzept
- ▷ Konfigurationsdateien werden anhand der Klassennamen ausgewählt
- ▷ Mit `fcopy` wird klassenbasiert ein Template kopiert
- ▷ Erfahrener Admin kreiert die Klassen
- ▷ Junior Admin ordnet die Klassen den Rechnern zu
- ▷ Junior Admin installiert die Rechner

- ▶ Ein Rechner gehört zu mehreren Klassen
- ▶ Priorität von niedrig nach hoch
- ▶ Beispiel: `DEFAULT SMALL_IDE GRUB GNOME demohost LAST`
- ▶ Klassen werden über Skripte in `/fai/class` definiert
- ▶ Alle Teile der Installation nutzen das Klassenkonzept
- ▶ Konfigurationsdateien werden anhand der Klassennamen ausgewählt
- ▶ Mit `fcopy` wird klassenbasiert ein Template kopiert
- ▶ Erfahrener Admin kreiert die Klassen
- ▶ Junior Admin ordnet die Klassen den Rechnern zu
- ▶ Junior Admin ~~installiert die Rechner~~ lässt installieren ;-)

Verzeichnisse im Config Space

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

Beispiel: /fai/disk_config/SMALL_IDE:

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

```
disk_config hda
```

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

Beispiel: /fai/package_config/BEOWULF:

```
# packages for Beowulf clients
```

```
PACKAGES install BEOULF_MASTER  
gmetad apache
```

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

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

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

- ▶ Aktionen wie bei `apt-get`: `install`, `remove` und `taskinst`, `taskrm`, `aptitude`
- ▶ Abhängigkeiten innerhalb der Pakete werden aufgelöst
- ▶ `dpkg -get-selections` auch möglich

Verzeichnisse im Config Space

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

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

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

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

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

Unterschiedliche Ziele der Installer

- d-i
 - ▷ Sei klein !!! Und modular. Passe ins RAM!
 - ▷ Menugesteuerte manuelle Installation eines Rechners
 - ▷ Frage nach Sprache, danach frage viel mehr in dieser Sprache
 - ▷ Decke übliche Installationen ab (debconf)
 - ▷ Installiere Basis System
 - ▷ discover1 zur Hardwareerkennung

Unterschiedliche Ziele der Installer

- d-i**
 - ▷ Sei klein !!! Und modular. Passe ins RAM!
 - ▷ Menugesteuerte manuelle Installation eines Rechners
 - ▷ Frage nach Sprache, danach frage viel mehr in dieser Sprache
 - ▷ Decke übliche Installationen ab (debconf)
 - ▷ Installiere Basis System
 - ▷ discover1 zur Hardwareerkennung

- FAI**
 - ▷ Kein einziger Tastenanschlag!
 - ▷ Computer Infrastruktur: Erst planen, dann installieren lassen
 - ▷ Keine Platzbeschränkung (Plattenplatz ist billig)
 - ▷ Klassenkonzept für Gruppierung von Rechnern
 - ▷ Zentrale Haltung der Konfigurations- und Logdateien aller Rechner
 - ▷ Installiere und Konfiguriere alles
 - ▷ Konfigurationsskripte in shell, perl, cfengine, expect,...
 - ▷ discover2 zur Hardwareerkennung

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

Installation mehrerer Knoten in einem Beowulf Cluster:

Knoten	Sekunden
1	337
5	340
10	345
20	379

12% mehr Zeit bei 20 Rechnern.

- ▶ Netcologne, Prionet, MPI Meteorologie, DESY, ZEDAT Fu-Berlin, BLB Karlsruhe, HBT
- ▶ IFW-Dresden, Germany, 100+ hosts, cluster, (i386 and IA64)
- ▶ Physics department (FU Berlin), 139+ hosts
- ▶ Host Europe, 250 hosts
- ▶ Lycos Europe, search engine, 200+
- ▶ Ewetel, ISP and telco, 65 hosts at 3 locations
- ▶ Physics department, university Augsburg, 80+
- ▶ Mathematics department, university Paderborn, 120+ clients and servers
- ▶ fms-computer.com, Germany, 200-300 hosts in several clusters for customers
- ▶ HPC2N, 2 clusters listed in `top500.org`, 192 dual Opteron, 120 dual Athlon
- ▶ Computer-aided chemistry, ETH Zürich, cluster, 45 dual Athlon nodes, 15 single nodes
- ▶ Electricité de France (EDF), France, 200 hosts
- ▶ MIT Computer science research lab, 200 hosts
- ▶ Danmarks Meteorologiske Institut, 85+ hosts
- ▶ University of New Orleans, 72 node Beowulf cluster
- ▶ Brown University, Dep. of Computer Science, 300+ hosts
- ▶ University of West Bohemia, Czech Republic, 180+

- ▷ **FAI-CD**
- ▷ **Debconf Supprt, preseeding**
- ▷ Update Tool für tägliches Update
- ▷ GUI für `faimond(8)`
- ▷ Cfengine 2 Unterstützung
- ▷ Subversion, arch Unterstützung
- ▷ Erweiterung von `fcopy`
- ▷ Neues Tool zur Partitionierung
- ▷ Proof of Concept für RedHat, Suse, **Ubuntu Linux**
- ▷ Aufteilung in mehrere Pakete (Dokumentation, Server, Client)
- ▷ FAI Wiki
- ▷ FAI Entwickler Workshop im April 2005

- ▶ Homepage: `http://www.informatik.uni-koeln.de/fai`
- ▶ Aktive Maillingliste: `linux-fai@uni-koeln.de`
- ▶ CVS Zugriff, Beispiele der Log Dateien
- ▶ Mehr als 90 detaillierte Berichte von Benutzern
- ▶ FAI läuft auf i386, amd64, IA64, SPARC, PowerPC, ALPHA
- ▶ Installiert auch Solaris auf SUN Sparc
- ▶ 5 Jahre FAI, Erfahrung, Rückmeldungen, Patches durch Benutzer
- ▶ Kommerzieller Support: `fai-cluster.de`