FAI

Alfred Gil, HPCNow!
alfred.gil@hpcnow.com

Agenda

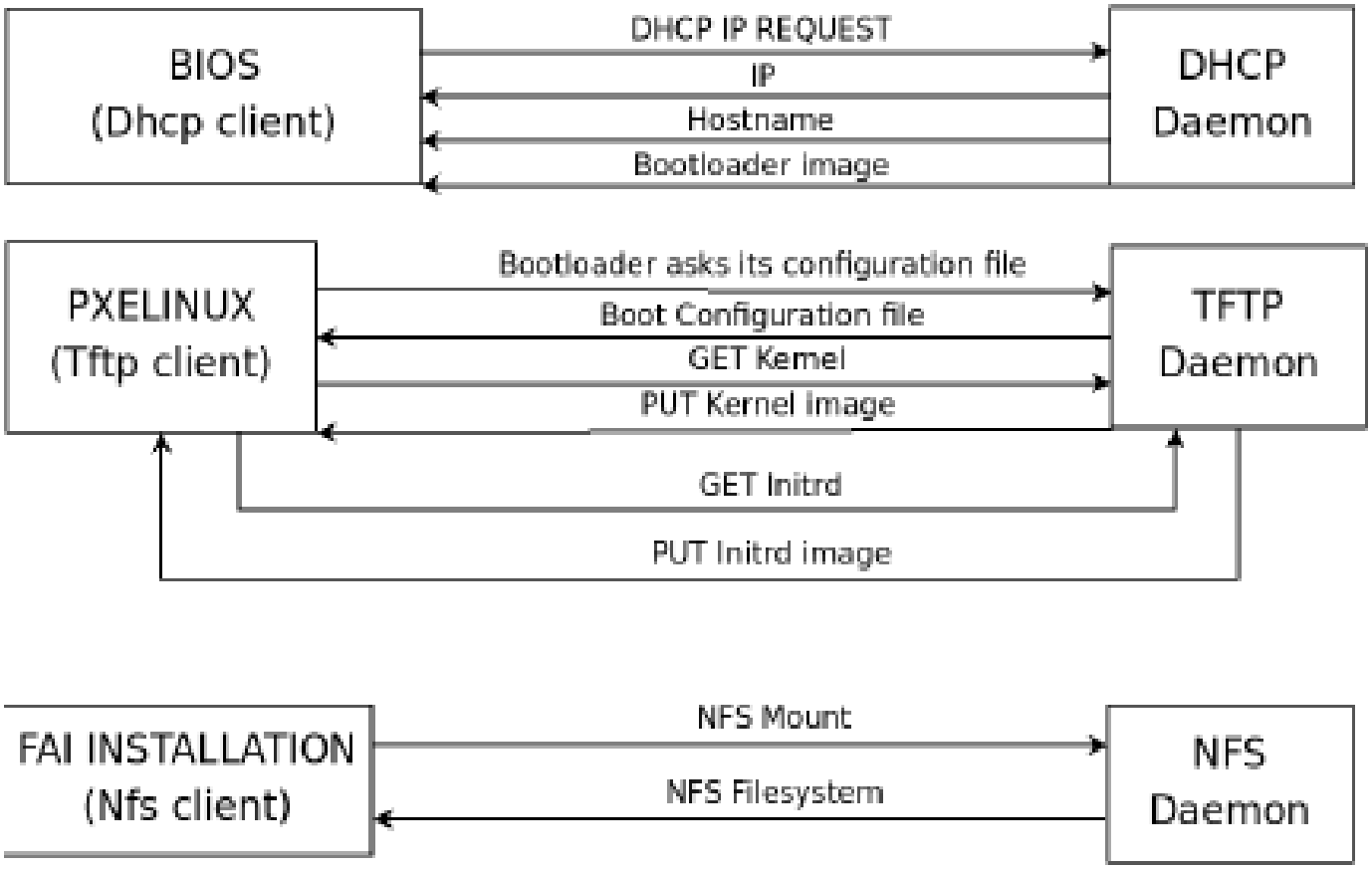
- Introduction to FAI
- Configuration space and classes
- Node installation
- System updates
- Outlook

HPC Now!

Introduction to FAI

- Non-interactive system to install, customize and manage Linux computers
- Unattended mass deployment of Linux
- Plan your installation, and FAI installs your plan
  - “Hard” to configure for the first time

How does FAI work



and automatic installation begins...

HPC Now!

How does FAI work

- FAI startup script
  - Partition hard disk
  - Create file systems
  - Install software packages
  - Configuring OS via scripting
  - Boot from local disk
- All information in the **configuration space** on install server
- Configuration files grouped in **classes**

FAI installation

- apt-get install fai-quickstart
- vi /etc/fai/fai.conf

```
# See fai.conf(5) for detailed information.

# Account for saving log files and calling fai-chboot.
LOGUSER=fai

# URL to access the fai config space
# If undefined, use default nfs://<install server>/$FAI_CONFIGDIR
#FAI_CONFIG_SRC=nfs://yourservername/path/to/config/space
```

✔ FAI_CONFIG_SRC undefined → default is NFS mount from fai install server

FAI installation

- vi /etc/fai/nfsroot.conf

```
# For a detailed description see nfsroot.conf(5)

# "<suite> <mirror>" for debootstrap
FAI_DEBOOTSTRAP="wheezy http://deploy:9999/debian"
FAI_ROOTPW='$1$kBnWcO.E$djxB128U7dMkrltJHPf6d1'

NFSROOT_ETC_HOSTS="192.168.7.15 deploy"
NFSROOT=/srv/fai/nfsroot
TFTPROOT=/srv/tftp/fai
NFSROOT_HOOKS=/etc/fai/nfsroot-hooks/
FAI_DEBOOTSTRAP_OPTS="--exclude=info"

# Configuration space
FAI_CONFIGDIR=/srv/fai/config
```

HPC Now!

FAI installation

- vi /etc/fai/apt/sources.list

```
# wheezy
deb http://ftp.us.debian.org/debian wheezy main non-free contrib
deb-src http://ftp.us.debian.org/debian wheezy main non-free contrib

# security updates
deb http://security.debian.org/ wheezy/updates main contrib non-free
deb-src http://security.debian.org/ wheezy/updates main contrib non-free

# wheezy-updates, previously known as 'volatile'
deb http://ftp.us.debian.org/debian wheezy-updates main contrib non-free
deb-src http://ftp.us.debian.org/debian wheezy-updates main contrib non-free

# repository that may contain newer fai packages for wheezy
deb http://fai-project.org/download wheezy koeln
```

FAI installation

- addgroup fai
- adduser --home /var/log/fai --ingroup fai fai
- fai-setup -fv
- tail /var/log/fai/fai-setup.log

```
FAI packages inside the nfsroot:
fai-client          4.0.8
fai-nfsroot         4.0.8
fai-setup-storage   4.0.8
FAI related packages inside the nfsroot:
dracut              020-2
dracut-network      020-2
Waiting for background jobs to finish
[1]+  Running                 nice xz -q $NFSROOT/var/
fai-make-nfsroot finished properly.
Log file written to /var/log/fai/fai-make-nfsroot.log
Re-exporting directories for NFS kernel daemon....
FAI setup finished.
```


HPC Now!

FAI installation

- cat /etc/exports

```
# /etc/exports: the access control list for filesystems which may be export
#               to NFS clients.  See exports(5).
#
# Example for NFSv2 and NFSv3:
# /srv/homes       hostname1(rw,sync,no_subtree_check) hostname2(ro,sync,nc
#
# Example for NFSv4:
# /srv/nfs4        gss/krb5i(rw,sync,fsid=0,crossmnt,no_subtree_check)
# /srv/nfs4/homes  gss/krb5i(rw,sync,no_subtree_check)
#
/srv/fai/config  192.168.7.15/24(async,ro,no_subtree_check)
/srv/fai/nfsroot 192.168.7.15/24(async,ro,no_subtree_check,no_root_squash)
```


HPC Now!

Preparing boot

- apt-get install isc-dhcp-server syslinux-common tftpd-hpa
- vi /etc/dhcp/dhcpd.conf

```
subnet 192.168.7.0 netmask 255.255.255.0 {
   option routers 192.168.7.251;
   option domain-name          ;
   option domain-name-servers               ;
   option time-servers proxy;
   option ntp-servers proxy;
   server-name deploy;
   next-server deploy;
   filename "fai/pxelinux.0";
}

host node01 {
        hardware ethernet 00:25:90:59:f8:d0;
        fixed-address 192.168.7.101;
}
```


HPC Now!

Preparing boot

- vi */etc/default/isc-dhcp-server*

```
# On what interfaces should the DHCP server (dhc
#       Separate multiple interfaces with spaces
INTERFACES="eth0"
```

- */etc/init.d/isc-dhcp-server restart*
- *vi /etc/inetd.conf*

```
#:OTHER: Other services
tftp dgram udp wait root /usr/sbin/in.tftpd /usr/sbin/in.tftpd
-s /srv/tftp
```

- */etc/init.d/openbsd-inetd restart*

Preparing boot

- Client loads pxelinux boot loader from deploy
  - configuration via TFTP from /srv/tftp/fai/pxelinux.cfg
    - fai-chboot -vc pxe.tmpl node01

```
# template generated by fai-chboot for host node01          with IP 192.16
8.7.101 from source /srv/tftp/fai/pxelinux.cfg/pxe.tmpl
default fai-generated

label fai-generated
kernel vmlinuz-3.2.0-4-amd64
append initrd=initrd.img-3.2.0-4-amd64 ip=eth0:dhcp root=/dev/nfs nfsroot
=192.168.7.15:/srv/fai/nfsroot:vers=3 aufs console=ttyS2,115200n8 monserv
er=deploy LOGSERVER=deploy FAI_FLAGS=verbose,sshd,createvt,reboot FAI_CON
FIG_SRC=nfs://192.168.7.15/srv/fai/config FAI_ACTION=install
```


HPC Now!

Network rescue system

- fai-chboot -S node01

```
append initrd=initrd.img-3.2.0-4-amd64 ip=eth0:dhcp root=/dev/nfs nfsroo
t=192.168.7.15:/srv/fai/nfsroot:vers=3 aufs console=ttyS2,115200n8 monse
rver=deploy LOGSERVER=deploy FAI_FLAGS=verbose,sshd,createvt FAI_CONFIG_
SRC=nfs://192.168.7.15/srv/fai/config FAI_ACTION=sysinfo
```

no reboot

- Fully functional linux without using local disk
  - remote login via ssh
  - backup or restore partitions
  - check file system
  - inspect hardware

HPC Now!

Agenda

- Introduction to FAI
- Configuration space and classes
- Node installation
- System updates
- Outlook

HPC Now!

Configuration space

- Information about how exactly install nodes
  - Install server
    - /srv/fai/config
  - Clients
    - mounted on /var/lib/fai/config ($FAI)
- It can also be received from cvs, svn or git

Configuration space folders

- class
  - define classes and variables, load kernel modules
- disk_config
  - disk partitioning and file system creation
- debconf
  - preseeding data
- package_config
  - packages to be installed or removed, apt keys

Configuration space folders

- scripts
  - local site customization
- files
  - files used by customization scripts
  - not automatically copied to target, explicitly copied by customization scripts
- basefiles
  - minimal base image of the linux distribution

Configuration space folders

- hooks
  - functions or programs run before a task is called
    - Ex: install diskless client skipping local disk partitioning
  - FAI searches hooks for task and class
- Useful examples can be found at:
  - /usr/share/doc/fai-doc/examples/simple

Classes

- Determine which configuration file tho choose
- Client searches list of defined classes and uses all configuration files matching the class name
- A class is defined or undefined for a node, has no value
- Predefined classes
  - DEFAULT, LAST, host name (defined for all hosts)

Classes

- Can be listed on a file or defined dynamically by scripts
  - define classes depending on hardware, etc.

Agenda

- Introduction to FAI
- Configuration space and classes
- Node installation
- System updates
- Outlook

HPC Now!

Monitoring installation

- fai-monitor daemon running on deploy
  - /etc/init.d/fai-monitor
- Client checks whether daemon is running
  - sent messages on start and end tasks

Booting the kernel

- Client receives kernel and initial RAM disk
- Kernel boots up and load RAM disk
- Mount root file system read only (passed by additional kernel parameters)
  - root=/dev/nfs, nfsroot=192.168.7.15:/srv/fai/nfsroot
- rootfs is made writable by mounting a RAM disk via aufs (another unionfs) on top of it

Starting FAI

- /usr/sbin/fai
  - main script controlling the sequence of tasks

```
confdir                # get config space
setup                  # early part of initialization
defclass               # define classes
defvar                 # define variables
action                 # evaluate FAI_ACTION
install                # Do the initial installation
partition              # partition the harddisks, create file systems
mountdisks             # mount the file systems
extrbase               # extract the minimal base.tgz
mirror                 # get a Debian mirror via NFS
debconf                # do Debian preseeding
repository             # preapre access to the package repository
updatebase             # Debian specific
HOOK instsoft.FAIBASE # fcopy kernel-img.conf
instsoft               # install software packages
configure              # call customization scripts
finish                 # do some cleanup, show installation time
tests                  # call tests if defined
chboot                 # call fai-chboot on the install server
HOOK savelog.LAST      # grep for error messages in all log files
savelog                # save log file to local dir and remote
faiend                 # reboot host, eject CD if needed
```

confdir and setup

- Get DHCP info
- Configuration space is made available
  - via nfs from deploy/srv/fai/config
- $FAI_FLAGS are defined (read from kernel parameters line)
- Virtual terminals created
- ssh daemon started

defclass and defvar

- $FAI/class/
- Scripts matching [0-9][0-9]* are executed to define classes for the client
- File matching *.var, with a prefix matching a defined class, is executed to define variables

action

- sysinfo
- install
- softupdate

partition and mountdisks

- $FAI/disk_config/
- Partitioning of disk
- Create and mount file systems
  - During installations all local file systems are mounted relative to /target
    - Ex: /target/home will become /home in the new system

extrbase, debconf

- At this point, local file systems are empty
- $FAI/basefiles
  - Unpack base archive
    - Class.tar.xz, created via mk-basefile script (allow different linux distributions)
    - Otherwise, nfsroot file system (debian)
- $FAI/debconf
  - Debian preseeding

repository

- $FAI/files
  - Get files needed to access repositories
    - resolv.conf, hosts, sources.list, etc.
    - /etc/apt/apt.conf.d/30proxy
- $FAI/package_config
  - *.asc are APT keys

updatebase, instsoft and configure

- Update base file system with new repositories
- $FAI/package_config
  - packages to install (dependencies will install automatically)
  - chroot to /target and aptitude install
- $FAI/scripts
  - Adjust system configuration
    - Set network, configure ganglia, define nfs mounts, etc...

chboot, savelog and faiend

- Changes PXE configuration of the node at installation server
  - fai-chboot disable /srv/tftp/fai/pxelinux.cfg file
  - then, default file is read (localboot) on reboot
- Save logs on client and server
  - /var/log/fai/nodexx
- reboot node

When things go wrong...

- Client can't boot from network card
  - tcpdump between network and client
  - egrep "tftpd|dhcpd" /var/log/*
- Installation seems to stop
  - postinstall script requiring input from console
    - change to another virtual terminal and use top or ps
    - add debug to $FAI_FLAGS
- log files on /tmp/fai during installation
- /var/log/fai/error.log

Agenda

- Introduction to FAI
- Configuration space and classes
- Node installation
- System updates
- Outlook

HPC Now!

System update

- Same configuration as installing
- Updated system looks like newly installed, but preserving local data
  - No partitioning and file system creation
  - Basesystem is not bootstrapped
  - System is not rebooted
- fai -v softupdate (executed at client)

Configuration suitable for updates

- Has to be idempotent
  - Running the scripts twice should result in the same system configuration as running them once
    - Never blindly append to files, instead use ainsl
  - Use FAI environment variables
    - $ROOTCMD: chroot when install, empty when update
  - Restart daemons if needed
    - $ROOTCMD invoke-rc.d $somedaemon restart
  - scheduling a reboot if a new kernel is installed

Locally changed config files

- There shouldn't be any!!
  - No local changes on computing nodes
- Nevertheless, fcopy does a backup
  - file.pre_fcopy
  - Save all of them with logfiles
    - Add “FAI_BACKUPDIR=$LOGDIR/backup” on class/DEFAULT.var
- Keep track of changes with some IDS
  - Tripwire or Integrit

Agenda

- Introduction to FAI
- Configuration space and classes
- Node installation
- System updates
- Outlook

HPC Now!

Outlook

- Massive install and update of linux systems
- Centralized deployment and configuration management
- Integrated disaster recovery system
- Fast and scalable
- No need to clone images?