

# Het Linux filesystem (1)

In dit hoofdstuk maken we kennis met het Linux filesystem en enkele commando's.



## Doelstelling

Aan het eind van dit hoofdstuk is de cursist bekend met:

- het waarom van een filesystem
- gebruikelijke filesystem types
- virtuele filesystemen
- het maken van een filesystem
- mounting, unmounting en mountpoints
- de logische structuur van een filesystem
- gebruikelijke bestandslokaties
- het bewaken van een filesystem
- enkele filesystem commando's

---

# Waarom een filesystem

- structurering van de data
- makkelijker te beïnvloeden met de juiste commando's

---

## Gebruikelijke filesystemen

ext2	Eerdere versie van het Linux filesystem. Geen journaling. Kan nog handig zijn voor kleine filesystemen (waar een <i>journal</i> te veel ruimte nodig heeft).
ext3	Opvolger van ext2, met journaling: snellere recovery, minder kans op corruptie.
ext4	Huidige Linux filesystem. Kan grotere disks (>32TB) en grotere files (>2TB) aan.
btrfs	Kandidaat nieuwe standaard Linux filesystem. Mogelijkheden tot snapshots, subvolumes en dynamische van grootte veranderen
zfs	Zeer geavanceerd filesystem van voormalig SUN (nu oracle). Linus zou dit dolgraag op Linux zien maar er zijn wat licentieproblemen. Zijn wel pogingen onderweg om zfs toch voor Linux te ontwikkelen (zie <a href="http://zfsonlinux.org">zfsonlinux.org</a> ).
reiserfs	Journaling en zeer goed voor filesystemen met veel kleine files (<32KB)
FAT	File Allocation Table. Oud, primitief (DOS, win9x). Type code msdos heeft de originele DOS formaat 8.3 filenamen. Type code vfat heeft langere (windows) filenamen
ntfs	New Technology File System. Bekend van Windows NT/XP/Vista/7. Linux kan geen (nieuwe) files schrijven op NTFS
ISO-9660	Standaard CD-ROM filesystem. De <i>Rock Ridge extensions</i> maken ook lange filenamen, <i>symbolic links</i> en linux <i>permissies</i> mogelijk (later meer over links en permissies).
jfs	Journalled File System van IBM (AIX Unix).
xfs	Extents File System. Technisch zeer goed (meer met IRIX OS, dan met Linux).
ufs	Unix File System. In gebruik (en aangepast) door vele Unix OS-en waaronder Solaris en HP-UX

# Virtuele filesystemen

Mede door het dynamische (hotplug) gedrag van diverse hardware wordt steeds meer gebruik gemaakt van virtuele filesystemen.

De bekendste:

- het proc virtueel filesystem (mountpoint `/proc`).  
Maakt informatie over processen maar ook over diverse hardware (irqs, ioports, cpu, memory, ...) toegankelijk.
- het sysfs virtueel filesystem (mountpoint `/sys`).  
Maakt device informatie toegankelijk voor user-space utilities.
- udev, virtueel filesystem (mountpoint `/dev`). Voor het maken (verwijderen) van dynamische device files.  
Config in `{etc|lib}/udev/rules.d` (`/etc` voor lokale aanpassingen (gaat voor `/lib`)).

# Maken van een filesystem

Commando's om een filesystem te maken:

- `mkfs.fstype` (bv. `mkfs.ext4`)
- kan ook met `mkfs` en optie `-t`. Bv. `mkfs -t ext4`
- of met `mke2fs` en optie `-t`. Bv. `mke2fs -t ext4`  
(voor alleen `ext[234]`)

Enkele opties (`mkfs.fstype` or `mke2fs`):

- `-t ext{2|3|4}` kies filesystem type
- `-m percentage` beïnvloed gereserveerde ruimte voor root
- `-j` maak een ext3 *journal*
- `-c` check voor *bad blocks*
- `-n` dry run. Laat alleen zien wat zou gebeuren

# Swap

Swap heeft geen filesystem maar moet wel gemaakt worden :

- Maak swap: `mkswap /dev/sda3`
- Activeer swap: `swapon /dev/sda3`

Permanent maken van swap met het file `/etc/fstab` (later meer).

## (e2)fsck

File System Check: filesystem checken en repareren

Net als mkfs varianten: `fsck.ext3`, `fsck.ext4`

Ook gebruikelijk: `e2fsck` (voor **alleen** ext[234])

Voorbeelden (e2fsck achterhaald zelf het fs type):

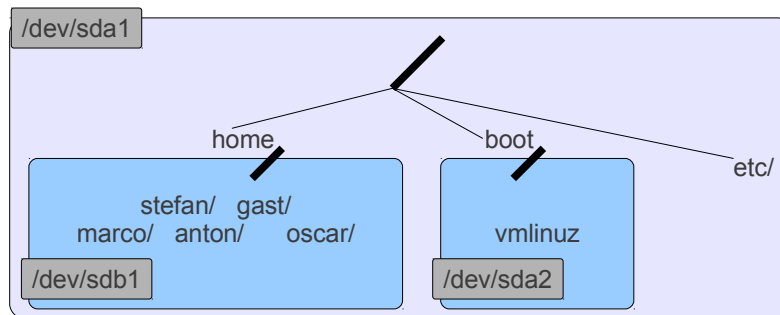
- `e2fsck -p /dev/sda1` Repareer en automatisch indien mogelijk (veelal gebruikt in opstartscripts). Relatief safe.
- `e2fsck -y /dev/sda1` Voor niet-interactief gebruik: zeg ja (yes) op alle vragen (m.b.t. repareren). Kan niet samen met -p.
- `e2fsck -b superblock` Specificeer een ander superblock (superblock bevat alle filesystem administratie: type, grootte, status, vrije inodes, .. Zeer cruciaal en daarom meerdere kopieën beschikbaar)

Let weer op: file systeem mag **niet** gemount zijn!

# Mounten

Idee: koppel de filesystemen op de verschillende partities aan elkaar zodat het 1 groot filesysteem *lijkt* om data overzichtelijk in op te kunnen slaan.

- DOS/Windows gebruiken drive letters (C: of D:) voor partities
- In Linux: partities koppelen (*mounten*) aan een directory (*mount point*) in de directory-tree.
- **De root (/) van de *partitie* wordt /"mountpoint" in de directory tree.**  
In onderstaand plaatje zijn /home en /boot lege directories op /dev/sda1 die fungeren als mountpoints voor /dev/sdb1 en /dev/sda2:



## mounting

**Eenmalig** met commando: `mount`

```
mount [-alrsvw] [-t fstype] [-o options] [device] [mountpoint]
```

Enkele opties:

- `-r` mount read-only
- `-o remount,rw` remount device read-write
- `-t fstype` specificeer het filesystem type (als bv. autodetect faalt). Filesystem type moet wel bekend zijn bij de kernel (meegecompileerd of als module).
- `-a` mount alle filesystemen die in /etc/fstab staan (commando tijdens boot)

Ongedaan maken: `umount /dev/sdb2 (of /mountpoint)`

**Permanent** maken: `/etc/fstab`

## /etc/fstab (1/2)

1 regel per te mounten device

Voorbeeld:

# <file system>	<mount point>	<type>	<options>	<dump>	<pass>
proc	/proc	proc	nodev,noexec,nosuid	0	0
/dev/sda1	/	ext4	errors=remount-ro	0	1
/dev/sda2	/var	ext4	defaults	0	2
# swap = /dev/sda5					
UUID=573533c1-5098-432c-b409-f7cdc8048a36	none	swap	sw	0	0
# usbdisk					
UUID=8f209834-6240-427c-a1ae-98e822031d6e	/usbdl	ext4	defaults,user,noauto	0	0

- **UUID:** **uniek** ID voor een device. Staan in /dev/disk/by-uuid (met dynamische devices als usb-disks zijn de normale devicenummers (sd[abc..]) niet eenduidig).
- **dump veld:** bepaald of partitie voor het commando dump in aanmerking komt
- **pass veld:** bepaald volgorde voor fsck bij opstarten. / is normaal 1, de rest 2. Een waarde van 0 geeft aan dat het betreffende filesysteem niet gecheckt hoeft te worden.

## /etc/fstab (2/2)

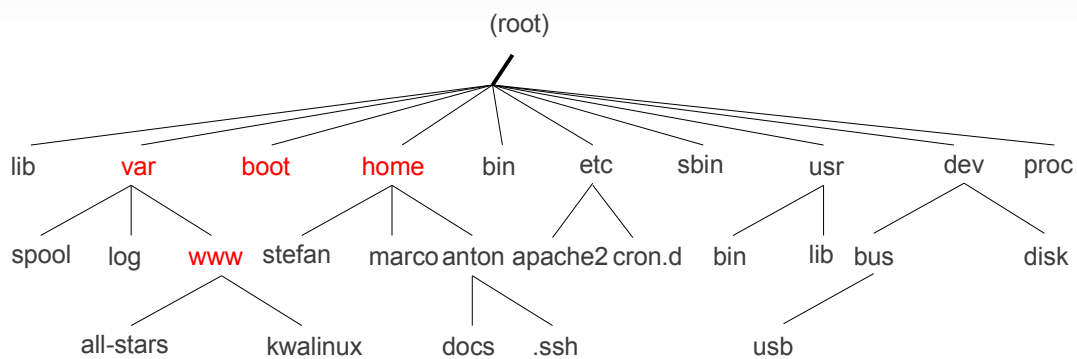
Enkele opties (comma gescheiden in het options veld):

- **noexec:** geen executie van binaries mogelijk
- **noauto:** mount niet als "mount -a" wordt gegeven (zoals bv. bij opstarten)
- **defaults:** mount met de defaults (rw, suid, dev, exec, auto, nouser, and async)
- **user:** ook gewone users mogen het device mounten
- **owner:** de eigenaar van het device mag mounten (meestal root)
- **loop:** gebruikt loopback device. Gebruikt om een file te mounten alsof het een disk is. Bv. voor een iso file (vanaf de command line):  

```
mount -t iso9660 -o loop centos7.iso /mnt/centos7
```
- **usrquota, grpquota:** quota opties (later meer)

# De logische structuur

Een lang niet volledige voorstelling van de logische filesystem structuur:



## Disk indeling (layout) (1/3)

De FHS (File Hierarchy Standard) is een standaard voor Linux (als opvolger van de FSSTND) welke onderscheid maakt tussen directories/files die wel of niet **deelbaar** (tussen computers) zijn en ook of directories/files wel of niet statisch zijn.

Dit is hieronder in een schema weergegeven:

	delen	niet delen
statisch	/usr /opt	/etc /boot /bin /sbin /lib
variabel	/home /var/mail	/var/run /var/lock /var/log

## Disk indeling (*layout*) (2/3)

Een voorbeeld van een doorsnee partitie indeling (*layout*):

Partitie	grootte (..)	Gebruik
swap (niet gemount)	ca. 1-2 x RAM	extra geheugen (maar traag)
/boot	tot 1GB	boot-partitie met de kernel
/home		plek voor de gebruikers
/usr	2GB-10GB	Linux programma's en libraries
/usr/local		Software specifiek voor dit systeem
/opt		"third party software". Vaak commercieel
/var	2G-10GB	<i>varying</i> data: logfiles, caching, mail, print spooling, tmp directory. Maar bv. ook /var/www (website data)
/tmp		schrijfbaar voor gebruikers voor tijdelijke files (net als /var/tmp)
/mnt		lege directory, kan gebruikt worden als mountpoint
/media		lege directory, veelal voor het mounten van dynamische media als bv. usb-disks

swap, /var en /boot zijn "altijd" wel een goed idee.

Afhankelijk van de situatie kun je bepalen of andere partities nodig zijn.

## Disk indeling (*layout*) (3/3)

- Niet alles kun je op een aparte partitie zetten
- Dit geldt bv. voor /etc, /sbin, /bin, /lib, /dev

Bevatten systeem kritieke files:

- /dev: de device files (waaronder bv. de disk die je wilt mounten)
- /etc/fstab: bevat data wat gemount moet worden
- /bin/mount: het mount commando

---



# Het bewaken van een filesystem

- filesystemen kunnen gebreken vertonen:
  - te veel data (vol)
  - kapot vanwege hardware errors (kapotte disk)
  - slechte performance (verkeerd geconfigureerd)

Tools die kunnen helpen bij het vaststellen en oplossen van problemen:

- `dumpe2fs` verkrijgen van informatie
- `tune2fs` wijzigen (tunen) van filesystem parameters
- `debugfs` interactieve combinatie van `dumpe2fs` en `tune2fs`.
- `fsck` check en repareer een filesystem

Let op: **altijd** werken op een **niet** gemount filesystem (behalve `dumpe2fs`)!

## dumpe2fs

Commando voor het *verkrijgen* van filesystem informatie:

- `dumpe2fs /dev/sda1`
- `-h` laat groepinfo weg

Voorbeeld (output niet volledig):

```
# dumpe2fs -h /dev/sda1
Last mounted on: /
Filesystem UUID: bb8aace7-aala-4649-91a1-15ded0febe9e
Filesystem state: clean
Errors behavior: Continue
Filesystem OS type: Linux
Inode count: 18964480
Block count: 75838976
Reserved block count: 3791948
Free blocks: 13774136
Free inodes: 18000264
Block size: 4096
Filesystem created: Sat Feb 12 18:38:29 2012
Last mount time: Sun Jan 22 09:53:13 2013
Mount count: 18
Maximum mount count: 26
Last checked: Sat Jan 7 12:02:42 2013
Next check after: Thu Jul 5 13:02:42 2013
Lifetime writes: 631 GB
```

## tune2fs

Doel: om daadwerkelijk wijzigingen aan te brengen in je filesystem parameters (*tunen* van je filesystem).

Veel opties, enkele veel voorkomende:

- `-m percentage` zet de grootte van de gereserveerde blokken.
- `-c mounts` pas het maximum aantal mounts aan voordat een check plaats vindt.
- `-i 12m` zet de tijd tussen checks (in dit geval een jaar).
- `-j` add a journal (maak van ext2 een ext3 filesystem)

Let weer op: file systeem mag **niet** gemount zijn! (gebruik bv. een "rescue-cd" om het root filesystem te tunen)

---

## debugfs

Om parameters interactief te verkrijgen en te wijzigen:

- `debugfs -w /dev/sda1`

Voorbeelden:

- `debugfs: help`
- `debugfs: stat filename`
- `debugfs: undel "inode" new-filename`
- `debugfs: lsdel`
- `debugfs: rdump dir destination`
- `quit`

Let weer op: file systeem mag **niet** gemount zijn!

---

## xfs

extX filesystem hebben goede tools die soms op andere systemen ontberen (of een stuk minder zijn).

Een uitzondering hierop zijn de xfs-tools voor xfs filesystemen.

Zo bestaan er bijvoorbeeld:

`xfs_info` lijkt op `dumpe2fs`. Filesystem moet gemount zijn..

`xfs_metadump` om alle metadata naar een file te kopiëren. Dit file zou je op kunnen sturen naar experts voor nader onderzoek.

---

## Disk gebruik

Commando:

`df` (disk free) geef overzicht op partitie niveau snel te zien welke partitie vol is

```
$ df -h -t ext3
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1       958M  399M  511M  44% /
/dev/sda7       958M  230M  680M  26% /home
/dev/sda6       958M   51M  858M   6% /tmp
/dev/sda9       865M  546M  276M  67% /var
/dev/sda5       3.8G  1.1G  2.5G  31% /usr
```

- `-h` human readable format
- `-l` limit. Laat alleen lokale filesystemen zien
- `-t type` geef allen df voor filesystem type

Volgende hoofdstuk commando `du` (krachtig voor troubleshooten!)

- (disk usage) geef overzicht op directory/file niveau

## quota (1/2)

- limiteer disk usage per user en/of per group
- vanaf 2.6.x Linux kernels: quota v2 (2.4.x: quota v1)
- Enable quota:
  - install quota package
  - put in `/etc/fstab` in optie-veld: `usrquota` en/of `grpquota`, bv.:
    - `/dev/sdb1 /home ext4 usrquota 0 2`
  - activeer door reboot of: `modprobe "quotamodule"` (als install al niet gedaan heeft), run commando `quotaon` of run `quota init` script en `umount/mount` filesystem.

## quota (2/2)

- commando's
  - `edquota username` (Set quota voor *username*)
    - (geef size op in blocks (ext4: 1block = 4096 bytes))
    - `-p protouser` (Gebruik quota setting van user *protouser* als "template")
  - `edquota -g groupname` (Set quota voor *groupname*)
  - `quota show quota`
    - `-g` Show group quota
  - `quotacheck` verifieer en update quota informatie
  - `repquota filesystem`
    - rapporteer quota
    - `-a` rapporteer quota voor alle filesystemen met quota
- soft en hard limits
- niet alleen voor diskusage, ook voor inode usage (aantal files)

# Oefeningen

Tijd voor oefening!