



RAID

Redundant Array of Independent Disks

**DE HOGESCHOOL
MET HET NETWERK**

Hogeschool PXL – Dep. PXL-IT – Elfde-Liniestraat 26 – B-3500 Hasselt
www.pxl.be - [www.pxl.be/facebook](https://www.facebook.com/pxl.be)



RAID

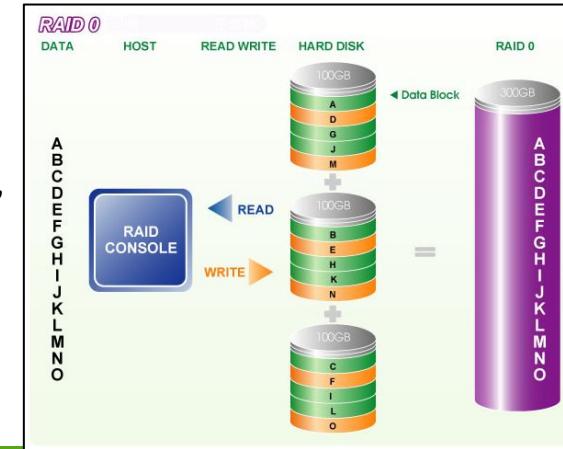


- RAID
 - Kan softwarematig of hardwarematig zijn
 - Verenigen van meerdere harddisks of partities
 - voor snelheidswinst van het lezen/schrijven van data
 - of om de veiligheid van de data te vergroten
 - RAID-levels
 - Geven het type van vereniging van de disks weer
 - worden voorgesteld door RAID - plus een cijfer (RAID0)



RAID-levels

- RAID 0
 - Data wordt in parallel naar 2 of meer disks geschreven waarbij de data over de drives **verdeeld** wordt.
 - De data blokken worden als volgt geschreven
 - block 1 naar disk 1, block 2 naar disk 2, block 3 naar disk 3 etc.
 - Wordt ook striping genoemd



RAID-levels

- RAID 0

- Capaciteit: De totale capaciteit van alle RAID-members
- Voordelen

- Snelheid is zeker een voordeel omdat alles in parallel geschreven wordt
- Het volledig beschikbaar hebben van alle ruimte is natuurlijk ook een plus.

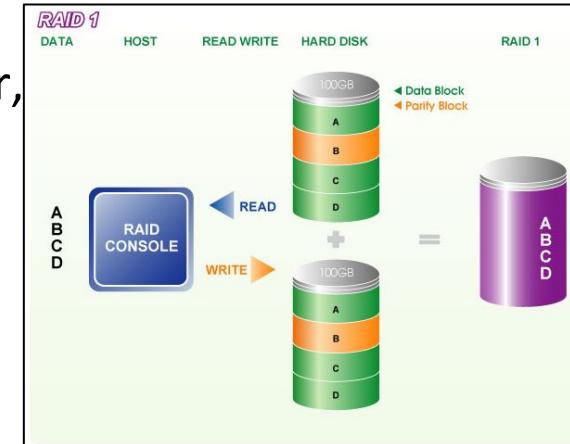
- Nadelen

- Het grote nadeel is echter het **ontbreken van betrouwbaarheid**.
- Als **1 disk** crasht is **alle data verloren**.



RAID-levels

- RAID 1
 - De data van de ene drive wordt gespiegeld naar de andere
 - op deze manier heb je dus altijd 2 drives met dezelfde data.
 - Deze configuratie is niet trager of sneller, maar wel de eenvoudigste betrouwbare RAID.
 - Wordt ook mirroring genoemd



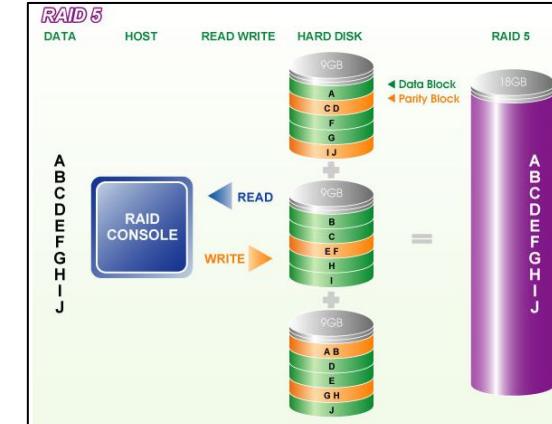
RAID-levels

- RAID 1
 - Capaciteit: helft van de totale capaciteit van de RAID-members
 - Voordelen
 - Betrouwbaarheid is zeker een voordeel van deze opstellingen.
 - als 1 disk crasht heb je toch nog alle data
 - Het vereist ook maar minimaal 2 drives en is daarmee eenvoudig en relatief goedkoop.
 - Nadelen
 - Niet echt efficiënt gebruik van opslagcapaciteit
 - je gebruikt twee even grote drives om uiteindelijk maar de helft van hun totale capaciteit te kunnen gebruiken.



RAID-levels

- RAID 5
 - De data wordt in blokken weggeschreven over de verschillende drives (minimum 3)
 - De Data-blocks worden verdeeld over alle RAID-members
 - Tevens worden er op iedere RAID-member Parity-blocks geplaatst, gemaakt van de data-blocks die weggeschreven zijn op de andere RAID members.



RAID-levels

- RAID 5
 - parity-block



Hard disk 1
A
pariteit voor CD
F

Hard disk 2
B
C
pariteit voor EF

Hard disk 3
pariteit voor AB
D
E

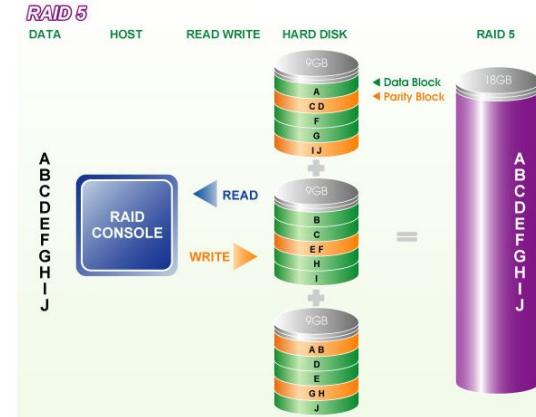
...

pariteit wordt samengesteld op basis van XOR

A	B	A XOR B
0	0	0
1	1	0
0	1	1
1	0	1

RAID-levels

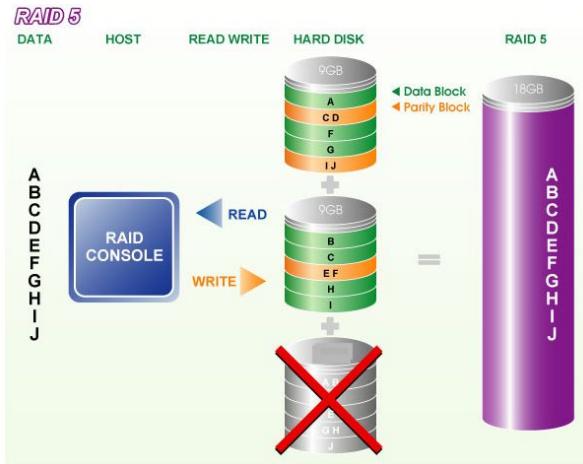
- RAID 5
 - Ook genoemd “Striping met roterende Pariteit”
 - Capaciteit:
 - Som van de capaciteit van alle drives - (Som van de capaciteit van alle drives / aantal drives)
 - Voordelen
 - Zeer goede **betrouwbaarheid** en goede **snelheid**.
 - Nadelen
 - Trager dan mirroring
 - RAID 6 heeft met vergelijkbare snelheid een hogere betrouwbaarheid.
 - RAID 6 heeft minimum 4 disks nodig en de pariteiten worden telkens op twee verschillende RAID-members geschreven
 - Er mogen hier dus 2 disks tegelijkertijd failen



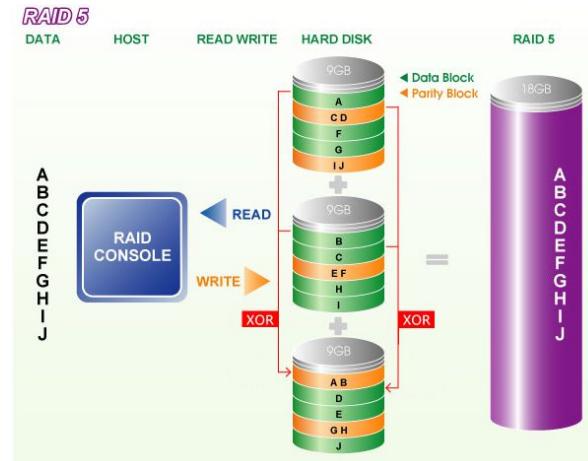
RAID-levels

- RAID 5

Crash of a member



Rebuilding



Nested RAID

- Nested RAID (=meerdere RAID-types tegelijk)
 - RAID 0/1
 - is een mirror (1) van stripes(0)
 - Eerst worden er twee RAID 0-stripes gemaakt en vervolgens worden deze als een mirror gezet.

A B C	<=>	D E F
<stripe>	<mirror>	<stripe>
 - 6 disks (A-F) van 100GB geeft dan 300GB totale capaciteit
 - Meerdere failures zijn geen probleem, als er maar een gespiegelde tegenhanger clean blijft



Nested RAID

- Nested RAID (=meerdere RAID-types tegelijk)
 - RAID 10 (of 1+0)
 - is een stripe(0) van mirrors (1)
 - Eerst worden de mirrors gemaakt en vervolgens worden deze als een stripe gezet.
$$\begin{matrix} A & \leftrightarrow & B \\ & \scriptstyle<\text{mirror}> & \\ C & \leftrightarrow & D \\ & \scriptstyle<\text{stripe}> & \\ E & \leftrightarrow & F \\ & \scriptstyle<\text{mirror}> & \\ & \scriptstyle<\text{stripe}> & \\ & \scriptstyle<\text{mirror}> & \end{matrix}$$
 - 6 disks (A-F) van 100GB geeft dan 300GB totale capaciteit
 - Er mogen tot 3 disks failen zolang deze niet in eenzelfde mirror zitten



Nested RAID

- Nested RAID (=meerdere RAID-types tegelijk)
 - RAID 50 (of 5+0)
 - is een stripe(0) van RAID5-arrays
 - Eerst worden de RAID5-arrays gemaakt en vervolgens worden deze gestriped.
- Er zijn nog andere nested-configs, maar die worden hier niet besproken



Nieuwe HDs gereed maken voor RAID met DOS-partitietabel-type

- De HDs moeten voorzien worden van partities m.b.v. fdisk
 - sudo fdisk /dev/sdx
 - nieuwe partitie aanmaken
 - n (new), p (primary), 1 (partnr), <enter> (first cylinder), <enter> (last cylinder)
 - type goedzetten voor RAID
 - t (type), 1 (partnr), fd (RAID autodetect)
 - Aanpassingen wegschrijven
 - w



Nieuwe HDs gereed maken voor RAID met GPT-partitietabel-type

- De HDs moeten voorzien worden van partities m.b.v. fdisk
 - sudo fdisk /dev/sdx
 - partitietype naar GPT
 - g (gpt)
 - n (new), 1 (partnr), <enter> (first sector), <enter> (last sector)
 - type goedzetten voor RAID
 - t (type), 29 (Linux RAID)
 - Aanpassingen wegschrijven
 - w



Aanmaken van een RAID1-set

- Minstens twee disks (even nummer)
 - met partities die klaargemaakt zijn voor Linux RAID
 - Disks checken: sudo mdadm --examine /dev/sdc /dev/sdd
 - disks hebben een partitie van het type fd(=RAID autodetect) OF
 - disks hebben een partitie van het type ee(=Linux RAID)
 - Partities checken: mdadm --examine /dev/sdc1 /dev/sdd1
 - partities hebben nog geen md-superblock omdat ze nog geen deel uitmaken van een RAID-set



Aanmaken van een RAID1-set

- RAID1-set maken van de partities
 - sudo mdadm --create /dev/md0 --level=mirror --raid-devices=2 /dev/sdc1 /dev/sdd1
 - Partities checken: sudo mdadm --examine /dev/sdc1 /dev/sdd1
 - partities maken nu deel uit van de RAID-set en geven hierover tal van informatie
 - Wordt voor de eerste keer gebuild, zodat de mirror werkt
 - Status bekijken kan via: cat /proc/mdstat of mdadm -D /dev/md0
 - Tonen de RAID-members alsook het (re-)sync-percentage



Check opnieuw met sudo blkid de UUIDs

/dev/md0 metadisk group, metadisks zijn gerelateerd aan RAID
(ook wel multiple device driver genoemd)

Info van de RAID1-set

- Nieuw RAID1-device
 - /dev/md0
 - Kan hetzelfde gebruikt worden als een andere partitie
 - te bekijken via: sudo ls -ld /dev/md*
 - b duidt op een block-device
 - Details bekijken
 - sudo mdadm --detail /dev/md0
 - Geeft info zoals Clean-state, Active-devices, Working-devices, Failed-devices, Spare-devices, RAID-members, ...



Hernoemen van een RAID-set

- Een RAID-set wordt soms na rebooten automatisch hernoemd naar md127

```
student@ubuntuserver01:~$ sudo mdadm --detail --scan
ARRAY /dev/md0 metadata=1.2 name=ubuntuserver01:0 UUID=4a9327ac:21b5ba64:f8b5e55c:43e5d0bb
```

Na reboot:

```
student@ubuntuserver01:~$ sudo mdadm --detail --scan          # of -D -s
[sudo] password for student:
ARRAY /dev/md/ubuntuserver01:0 metadata=1.2 name=ubuntuserver01:0
UUID=4a9327ac:21b5ba64:f8b5e55c:43e5d0bb
```

```
student@ubuntuserver01:~$ ls -l /dev/md
total 0
lrwxrwxrwx 1 root root 8 Oct 24 21:21 ubuntuserver01:0 ->
./md127
```

Check ook met lsblk

sdb	8:16	0	512M	0	disk
`-sdb1	8:17	0	511M	0	part
`-md127	9:127	0	510M	0	raid1

Hernoemen van een RAID-set

- Een RAID-set wordt soms na rebooten automatisch hernoemd naar md127
 - We kunnen deze naam ook vastzetten met de volgende stappen:

```
sudo mdadm -D -s | awk '{print $1,"/dev/md0",$3,$4,$5}' | sudo tee -a /etc/mdadm/mdadm.conf
```

Dit geeft een resultaat als: `ARRAY /dev/md0 UUID=4a9327ac:21b5ba64:f8b5e55c:43e5d0bb`

```
sudo update-initramfs -u
```

Daarna rebooten



Hernoemen van een RAID-set

- Na reboot:

```
student@ubuntuserver01:~$ sudo mdadm --detail --scan
[sudo] password for student:
ARRAY /dev/md0 metadata=1.2 name=ubuntuserver01:0 UUID=4a9327ac:21b5ba64:f8b5e55c:43e5d0bb
```

```
student@ubuntuserver01:~$ lsblk | grep -v loop
NAME      MAJ:MIN RM  SIZE RO TYPE  MOUNTPOINT
sda        8:0    0   10G  0 disk
| -sda1    8:1    0    1M  0 part
` -sda2    8:2    0   10G  0 part   /
sdb        8:16   0   512M 0 disk
` -sdb1    8:17   0   511M 0 part
` -md0     9:0    0   510M 0 raid1
sdc        8:32   0   512M 0 disk
` -sdc1    8:33   0   511M 0 part
` -md0     9:0    0   510M 0 raid1
```



Werken met de RAID1-set

- Nieuw RAID1-device
 - /dev/md0
 - Moet nog een filesysteem krijgen
 - sudo mkfs.ext4 /dev/md0
 - Moeten we nog mounten
 - sudo mkdir /var/ftpfiles
 - sudo mount /dev/md0 /var/ftpfiles
 - sudo chmod a=rwx /var/ftpfiles (of andere rechten indien nodig)
 - Nu kunnen we er mee werken
 - echo "echo Hallo" > /var/ftpfiles/testfile
 - cat /var/ftpfiles/testfile



Mounten RAID-device in /etc/fstab

- /etc/fstab
 - Optie 1, werken met /dev/md0

```
student@ubuntuserver01:~$ cat /etc/fstab
UUID=8282c2b2-c6f2-11e8-8c7c-000c29762fad / ext4 defaults 0 0
/swap.img none swap sw 0 0
/dev/md0 /var/ftpfiles ext4 defaults 0 0
```

- Optie 2, werken met UUID
 - More secure
 - Heeft geen last van veranderende nummers van md array
 - UUID te bekijken met lsblk -o name,fstype,uuid,mountpoint

UUID=b3837f78-70c0-47b8-bbba-08cc453ceca9	/var/ftpfiles	ext4	defaults	0
0				

RAID1-set met een spare

- spare
 - is een block-device dat gewoon wacht totdat een RAID-member failed. Dan wordt hij actief en zal hij de failed member vervangen.
 - Zo vlug hij online komt, zal er opnieuw gesynced worden. Dit proces noemt men Rebuilden
 - moet ook een partitie van type RAID bevatten
 - toe te voegen via
 - sudo mdadm --manage /dev/md0 --add /dev/sde1
 - checken: sudo mdadm --detail /dev/md0 (state spare)



Indien we de disk nu weggooien op VMware-niveau zien we dat de spare overneemt. State van de member-set is nu clean, degraded, recovering met een Rebuild Status in percentage en de spare geeft aan "Spare rebuilding"

RAID1-set met een spare

- spare als active-member
 - Indien we een member-failure hebben, begint het resync proces voor de spare. Gedurende deze tijd zijn we niet meer beveiligd tegen een extra member-failure
 - Het is daarom ook een mogelijkheid om de spare al onmiddellijk te betrekken in de RAID.
 - Dit kan via
 - sudo mdadm --grow --raid-devices=3 /dev/md0
 - checken: sudo mdadm --detail /dev/md0 (state active)



Extra commando's

- Extra commando's
 - Een RAID-member zelf als failing aanduiden
 - sudo mdadm --fail /dev/md0 /dev/sdd1
 - Een gefailde RAID-member verwijderen
 - sudo mdadm --remove /dev/md0 /dev/sdd1
 - Het aantal members van de RAID-set veranderen
 - sudo mdadm --grow --raid-devices=2 /dev/md0



Extra commando's

- Extra commando's
 - Een verwijderde RAID-member klaarmaken om opnieuw toegevoegd te worden aan de RAID-set
 - Hiervoor moet de superblock-info verwijderd worden
 - Dit kan met: sudo mdadm --zero-superblock /dev/sdd1



Superblock Definition

A superblock is a record of the characteristics of a filesystem, including its size, the block size, the empty and the filled blocks and their respective counts, the size and location of the inode tables, the disk block map and usage information, and the size of the block groups.

Aanmaken van een RAID5-set

- RAID5-set maken van de partities
 - sudo mdadm --create /dev/md1 --level=5
 --raid-devices=3 /dev/sdf1 /dev/sdg1 /dev/sdh1
 - Partities checken: mdadm --examine /dev/sdf1 /dev/sdg1...
 - partities maken nu deel uit van de RAID-set en geven hierover tal van informatie
 - RAID-device checken: sudo mdadm --detail /dev/md1
 - geeft info over de RAID-set en RAID-members alsook de build-status
 - Ook hier zouden we een extra spare kunnen toevoegen zoals we bij de RAID1-set hebben gedaan.



Werken met de RAID5-set

- Nieuw RAID5-device
 - /dev/md1
 - Moet nog een filesystem krijgen
 - sudo mkfs.ext4 /dev/md1
 - Moeten we nog mounten
 - sudo mkdir /var/www
 - sudo mount /dev/md1 /var/www
 - Nu kunnen we er mee werken (*denk aan permissies*)
 - cd /var/www; sudo mkdir html; sudo vi html/index.html
 - Indien de mount moet blijven na reboot
 - toevoegen in /etc/fstab
- In /etc/mdadm/mdadm.conf hernoemen evt. om nummer te behouden na reboot
- (via /dev/md1 of UUID)



Verwijderen van een RAID-set

- Verwijderen van een volledige RAID-set
 - /dev/md1
 - Moet eerst gestopt worden
 - sudo umount /dev/md1
 - sudo mdadm --stop /dev/md1
 - sudo mdadm --remove /dev/md1
 - /dev/sdf /dev/sdg /dev/sdh
 - kunnen nu opnieuw gepartitioneerd worden om te gebruiken zonder RAID
 - OF sudo mdadm --zero-superblock /dev/sdd1 om member te kunnen maken van een nieuwe RAID-set

