



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



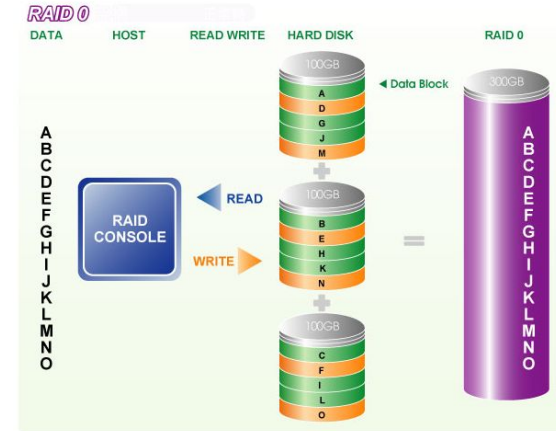
RAID



- 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 verenigen 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



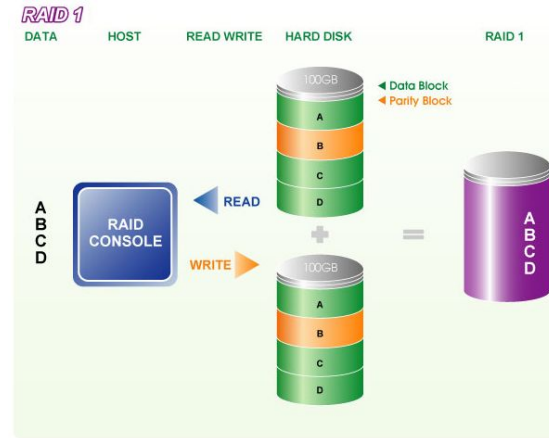
RAID-levels

- 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**.



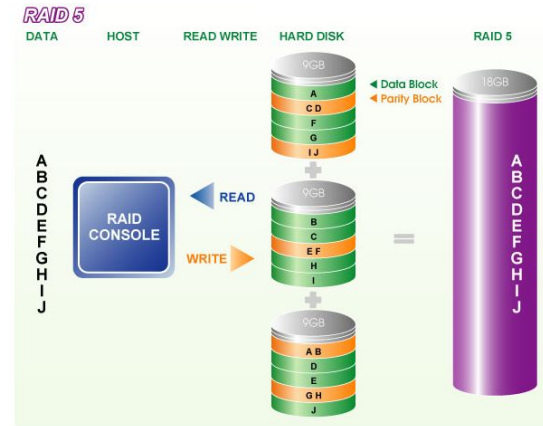
RAID-levels

- 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 sneller, maar wel de eenvoudigste betrouwbare RAID.
 - Wordt ook mirroring genoemd



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



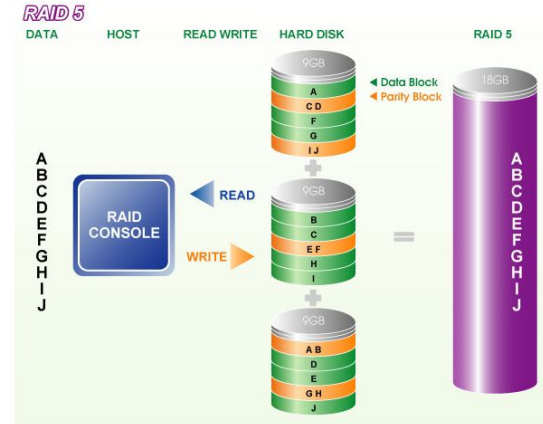
Hard disk 1	Hard disk 2	Hard disk 3
A	B	pariteit voor AB
pariteit voor CD	C	D
...		

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 en de pariteiten worden telkens op twee verschillende RAID-members geschreven
 - Er mogen hier dus 2 disks tegelijkertijd failen



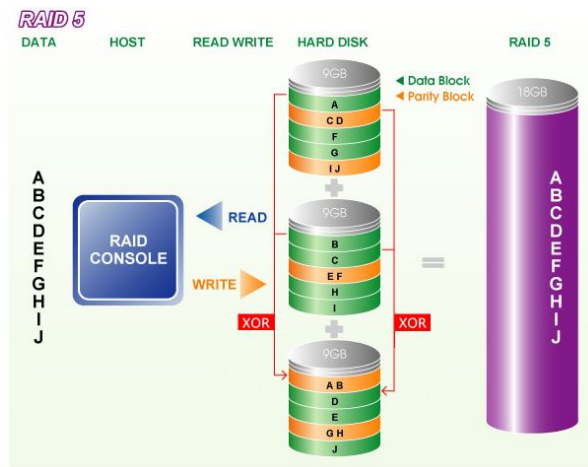
RAID-levels

- 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
 - Eén member-failure is geen probleem
 - Twee member-failures is geen probleem als de gefailde disks in dezelfde stripe zitten



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.
$$\begin{array}{cccccc} A <=> B & + & C <=> D & + & E <=> F \\ \text{<mirror>} & \text{<stripe>} & \text{<mirror>} & \text{<stripe>} & \text{<mirror>} \end{array}$$
 - 6 disks (A-F) van 100GB geeft dan 300GB totale capaciteit
 - Er mogen tot 3 disks failen zolang deze maar 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.

A,B,C + D,E,F + G,H,I
<RAID5> <stripe> <RAID5> <stripe> <RAID5>



- Er zijn nog andere nested-configs, maar die worden hier niet behandeld

Nieuwe HDs gereedmaken voor RAID

- 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



Aanmaken van een RAID1-set

- Minstens twee disks (even nummer)
 - met partities die klaargemaakt zijn voor Linux RAID autodetect
 - Disks checken: `sudo mdadm --examine /dev/sdc /dev/sdd`
 - disks hebben een partitie van het type fd(=RAID autodetect)
 - 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: `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 kan bekeken worden via: `cat /proc/mdstat`
 - Geeft de RAID-members alsook het (re-)sync-percentage



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

Info van de RAID1-set

- Nieuw RAID1-device
 - /dev/md0
 - Kan hetzelfde gebruikt worden als een andere partitie
 - te bekijken via: `sudo ls -l /dev | grep md0`
 - Details bekijken
 - `sudo mdadm --detail /dev/md0`
 - Geeft info zoals Clean-state, Active-devices, Working-devices, Failed-devices, Spare-devices, RAID-members, ...



Werken met de RAID1-set

- Nieuw RAID1-device
 - /dev/md0
 - Moet nog een filesystem krijgen
 - `sudo mkfs.ext4 /dev/md0`
 - Moeten we nog mounten
 - `sudo mkdir /var/ftpfiles`
 - `sudo mount /dev/md0 /var/ftpfiles`
 - Nu kunnen we er mee werken
 - `echo "echo Hallo" > /var/ftpfiles/testfile; cat /var/ftpfiles/testfile`
 - Indien de mount moet blijven na reboot
 - toevoegen in /etc/fstab



RAID1-set met een spare

- spare
 - is een block-device dat gewoon wacht totdat een RAID-member failed. Dan wordt hij actief en zal de failed member vervangen.
 - Zo vlug hij online komt, zal er opnieuw gesynced worden. Dit proces noemt men Rebuilden
 - is ook een partitie van type RAID autodetect (fd)
 - 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 beter 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)



Indien nu een disk failed, moet de spare niet gesynced worden!!

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
 - `echo "echo Hallo" > /var/www/testfile; cat /var/www/testfile`
 - Indien de mount moet blijven na reboot
 - toevoegen in /etc/fstab



Verwijderen van een RAID-set

- Verwijderen van een volledige RAID-set
 - /dev/md1
 - Moet eerst gestopt worden
 - `sudo mdadm --stop /dev/md1`
 - /dev/sdf /dev/sdg /dev/sdh
 - kunnen nu opnieuw gepartitioneerd worden om te gebruiken zonder RAID



Hernoemen van een RAID-set

- Een RAID-set wordt soms na reboorten automatisch hernoemd naar md127
 - Om de naam vast te zetten
 - `sudo su`
 - `mdadm --detail --scan | tail -1 >> /etc/mdadm/mdadm.conf`
 - `vi /etc/mdadm/mdadm.conf`
 - Achter ARRAY hernoemen naar `/dev/md/<nummer>`
 - `sudo update-initramfs -u`
 - `reboot`

