

# Software-RAID unter Linux

Andreas "Jimmy" Gredler

14.05.2005 / Linuxtage Graz



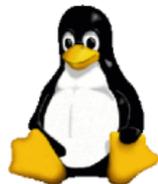
# Inhalt

- 1 **Einleitung**
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 **RAID Setup**
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 **Weiterführende Themen**
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 **Zusammenfassung**



# Gliederung

- 1 **Einleitung**
  - Was ist RAID ?
    - RAID-Level
    - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# RAID-Grundlagen

RAID ... Redundant Array of Inexpensive Disks  
oder  
RAID ... Redundant Array of Independent Disks

1988: Paper von Patterson, Gibson und Katz.



# Was bringt RAID

- Redundanz  
Bei den Standard RAID-Level darf eine Festplatte ausfallen
- Performance  
Je nach RAID-Level steigt vor allem die Lesegeschwindigkeit
- Größere Kapazität  
Heutzutage werden oft Speicher mit 1 TB oder mehr benötigt



# Was bringt RAID

- Redundanz  
Bei den Standard RAID-Level darf eine Festplatte ausfallen
- Performance  
Je nach RAID-Level steigt vor allem die Lesegeschwindigkeit
- Größere Kapazität  
Heutzutage werden oft Speicher mit 1 TB oder mehr benötigt



# Was bringt RAID

- Redundanz  
Bei den Standard RAID-Level darf eine Festplatte ausfallen
- Performance  
Je nach RAID-Level steigt vor allem die Lesegeschwindigkeit
- Größere Kapazität  
Heutzutage werden oft Speicher mit 1 TB oder mehr benötigt



# RAID-Level

## Standard RAID-Level

### Mögliche RAID-Level

- RAID 0
- RAID 1
- RAID 2
- RAID 3
- RAID 4
- RAID 5
- RAID 6
- RAID 7



# RAID-Level

## Standard RAID-Level

### Wichtige RAID-Level

- RAID 0
- RAID 1
- RAID 2
- RAID 3
- RAID 4
- RAID 5
- RAID 6
- RAID 7



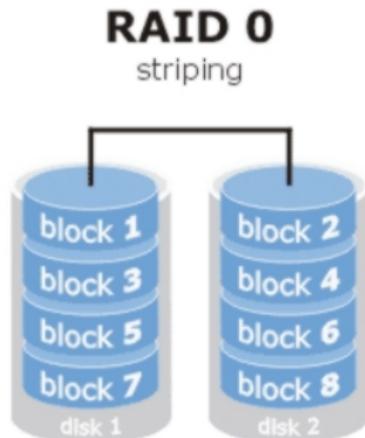
# Gliederung

- 1 **Einleitung**
  - Was ist RAID ?
  - **RAID-Level**
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# RAID-Level im Detail

## RAID-0

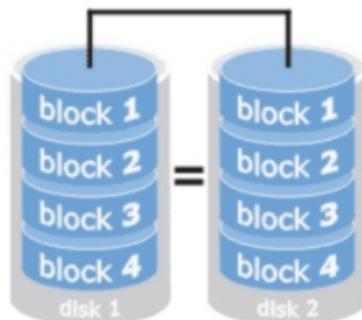


# RAID-Level im Detail

## RAID-1

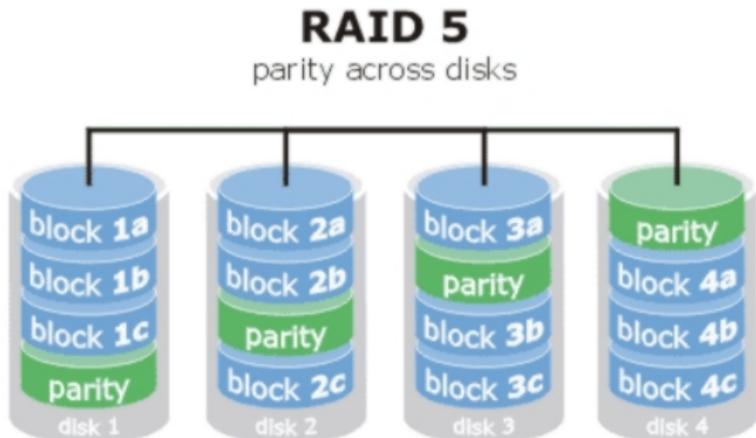
### RAID 1

mirroring



# RAID-Level im Detail

## RAID-5



# Gliederung

- 1 **Einleitung**
  - Was ist RAID ?
  - RAID-Level
  - **RAID unter Linux**
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Kernel

Linux Kernel v2.4.27 Configuration

**Main Menu**

Arrow keys navigate the menu. <Enter> selects submenus --->. Highlighted letters are hotkeys. Pressing <Y> includes, <N> excludes, <M> modularizes features. Press <Esc><Esc> to exit, <?> for Help. Legend: [\*] built-in [ ] excluded <M> module <> module capable

- Code maturity level options --->
- Loadable module support --->
- Processor type and features --->
- General setup --->
- Memory Technology Devices (MTD) --->
- Parallel port support --->
- Plug and Play configuration --->
- Block devices --->
- Multi-device support (RAID and LVM) --->**
- Networking options --->

1(+)

<Select> <Exit> <Help>



## Kernel

## Linux Kernel v2.4.27 Configuration

## RAID-Device support (RAID and LVM)

Arrow keys navigate the menu. <Enter> selects submenus --->.  
 Highlighted letters are hotkeys. Pressing <Y> includes, <N> excludes,  
 <M> modularizes features. Press <Esc><Esc> to exit, <?> for Help.  
 Legend: [\*] built-in [ ] excluded <M> module <> module capable

```

[*] Multiple devices driver support (RAID and LVM)
<*) RAID support
<*) linear (append) mode
<*) RAID-0 (striping) mode
<*) RAID-1 (mirroring) mode
<*) RAID-4/RAID-5 mode
<> Multipath I/O support
<> logical volume manager (LVM) support
<>> Device-mapper support
  
```

<Select> <Exit> <Help>



# Software, Userland Tools

Hier gibt es zwei verschiedene Pakete

- raidtools2
- mdadm

Mittlerweile sollte man mdadm verwenden, bei Debian ist das Paket raidtools2 bereits entfernt worden.



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - **Erstellen eines RAID**
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Mithilfe der Raidtools

Erstellen der Datei `/etc/raidtab`

```
raiddev /dev/md0  
raid-level 1  
nr-raid-disks 2  
nr-spare-disks 1  
persistent-superblock 1  
device /dev/sda1  
raid-disk 0  
device /dev/sdb1  
raid-disk 1  
device /dev/sdc1  
spare-disk 0
```



# Mithilfe der Raidtools

## Weitere Vorgehensweise

- `mkraid /dev/md0`
- `cat /proc/mdstat`
- `mkfs.ext3 /dev/md0`
- `mount /dev/md0 /mnt/raid`



# Mithilfe von Mdadm

- `mdadm --create --verbose /dev/md0 --level=raid1 --raid-devices=2 /dev/sda1 /dev/sdb1`
- `cat /proc/mdstat`
- `mkfs.ext3 /dev/md0`
- `mount /dev/md0 /mnt/raid`



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - **Testen eines RAID**
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Ausfall simulieren, Methode 1

- Rechner runterfahren
- Eine Platte entfernen
- Rechner neu starten
- Mdadm erkennt das Fehlen einer Platte -> Das RAID ist nun im "degraded mode"
- Rechner wieder runterfahren, Platte wieder anstecken und hochfahren
- Nun muß man die Platte wieder in das RAID einfügen



## Ausfall simulieren, Methode 2

- Eine Platte als 'faulty' markieren  
(`mdadm -fail /dev/md0 /dev/sdb1`)
- Diese Platte nun aus dem RAID entfernen  
(`mdadm -r /dev/md0 /dev/sdb1`)
- Das RAID ist nun im "degraded mode"
- Nun muß man die Platte wieder in das RAID einfügen  
(`mdadm -a /dev/md0 /dev/sdb1`)



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - **Wartung eines RAID**
  - Tips zum Umgang mit RAID
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Monitoring

- Raidtools benutzen syslog
- Mdadm daemon schickt bei Problemen eine mail

```

#-----BEGIN PGP MESSAGE-----
Envelope-to: root@lynx
From: mdadm monitoring <root@lynx>
To: root@lynx
Subject: Fail event on /dev/md0:lynx

This is an automatically generated mail message from mdadm
running on lynx.

A Fail event had been detected on md device /dev/md0.
It could be related to sub-device /dev/sdb1.

Faithfully yours, etc.

#-----BEGIN PGP MESSAGE-----
Fail event on /dev/md0:lynx
  
```



# Monitoring

```
root@lynx:~# cat /dev/mem | dd if=/dev/mem of=/dev/null bs=1k count=1000000  
Envelope-to: root@lynx  
From: mdadm monitoring <root@lynx>  
To: root@lynx  
Subject: Fail event on /dev/md0:lynx  
  
This is an automatically generated mail message from mdadm  
running on lynx  
  
A Fail event had been detected on md device /dev/md0.  
  
It could be related to sub-device /dev/sdb1.  
  
Faithfully yours, etc.  
  
----- 000/000: mdadm monitoring Fail event on /dev/md0:lynx ----- (all)
```



# Status

```

root@grml ~ # cat /proc/mdstat
Personalities : [linear] [raid0] [raid1] [raid5] [multipath] [raid6] [raid10]
md0 : active raid1 sda1[0] sdb1[1]
      4096448 blocks [2/2] [UU]
      [=====>.....] resync = 38.9% (1597184/4096448) finish=2.3min speed=18058K/sec
unused devices: <none>
root@grml ~ # █

```



# Informationen

```

lynx:~$ mdadm --detail /dev/md0
/dev/md0:
  Version : 00.90.00
  Creation Time : Tue Mar 29 22:43:21 2005
  Raid Level : raid1
  Array Size : 4096448 (3.90 GiB 4.19 GB)
  Device Size : 4096448 (3.90 GiB 4.19 GB)
  Raid Disks : 2
  Total Disks : 2
  Preferred Minor : 0
  Persistence : Superblock is persistent

  Update Time : Thu Apr 14 08:04:23 2005
  State : dirty, no-errors
  Active Drives : 2
  Working Drives : 2
  Failed Drives : 0
  Spare Drives : 0

  Number   Major   Minor   RaidDisk   State
    0         8         1         0   active sync  /dev/sda1
    1         8        17         1   active sync  /dev/sdb1
  UUID : 61e12bce:45e33ef8:6cf63d3e:72980e43
lynx:~$

```



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - **Tips zum Umgang mit RAID**
- 3 Weiterführende Themen
  - Weitere RAID-Level
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Generelle Tips

- Nicht auf das backup vergessen weil man ein RAID hat
- Bei IDE jede Platte als einziges Gerät am BUS betreiben
- Die Daten des RAIDs griffbereit haben (/etc/raidtab oder entsprechender output von mdadm)
- Einen Ausfall simulieren bevor man das RAID einsetzt
- Eine defekte Platte sofort austauschen



# Tips beim Booten von RAID

- Nur lilo kann von einem RAID booten (Aktuelle grub Version auch?)
- Booten kann man momentan nur von einem RAID-1
- Immer eine gute Live-CD bereithalten, mit der man das RAID wieder in Gang bringen kann
- Die Konfiguration mit Vorsicht durchführen und zwischendurch immer testen (reboot)



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 **Weiterführende Themen**
  - **Weitere RAID-Level**
  - RAID übers Netzwerk
- 4 Zusammenfassung



# Verkettete RAID-Level

Man kann auch mehrere RAID-Geräte wieder zu einem RAID zusammenfügen

- RAID 10 oder RAID 01
- RAID 50 oder RAID 05
- RAID 51 oder RAID 15



# Gliederung

- 1 Einleitung
  - Was ist RAID ?
  - RAID-Level
  - RAID unter Linux
- 2 RAID Setup
  - Erstellen eines RAID
  - Testen eines RAID
  - Wartung eines RAID
  - Tips zum Umgang mit RAID
- 3 **Weiterführende Themen**
  - Weitere RAID-Level
  - **RAID übers Netzwerk**
- 4 Zusammenfassung



# RAID über Netzwerk

- NBD  
Network Block Device
- ENBD  
Enhanced Network Block Device
- DRDB  
Distributed Remote Block Device



# Zusammenfassung

- Mit Software-RAID sind die wichtigsten RAID-Level abgedeckt
- SW-RAID ist äußerst flexibel
- Auf aktueller HW ist der overhead minimal
- Durch SW-RAID erhält man ohne hohe Kosten Redundanz
- SW-RAID ist relativ einfach zu administrieren



# Links

- A Case for Redundant Arrays of Inexpensive Disks (RAID)
- Software RAID HowTo
- Wikipedia
- Meine homepage



Fragen, Kritik, Anregungen etc. sind willkommen!

[jimmy@g-tec.co.at](mailto:jimmy@g-tec.co.at)

