2. Braunschweiger Linux-Tage

Vortrag über

RAID

von

Thomas King

http://www.t-king.de/linux/raid1.html

Übersicht:

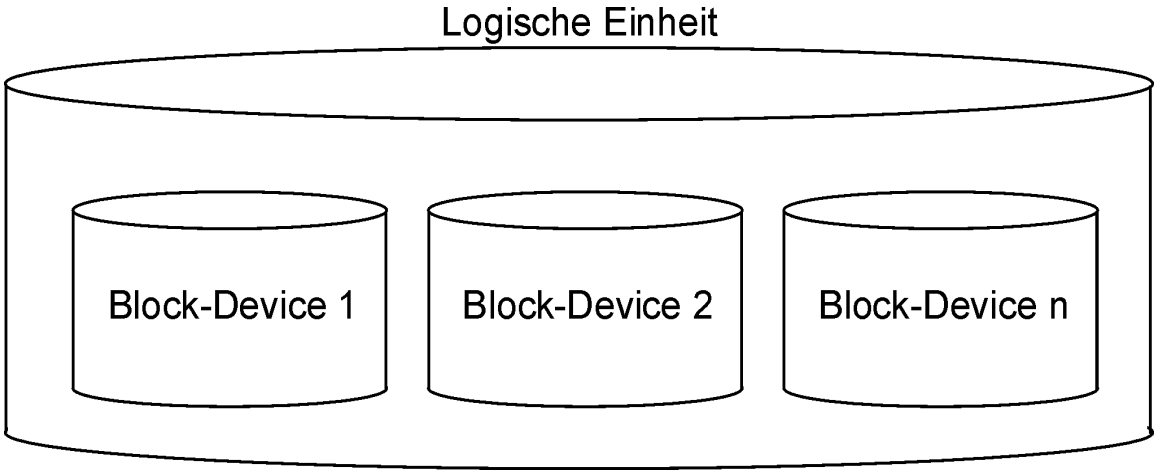
1. Was ist RAID?
- *1.1. Wo wurde RAID entwickelt?*
- *1.2. Welche RAID-Levels gibt es?*
- *1.3. Wo setzt man RAID ein?*

2. Linux und RAID
- *2.1. Hardware-RAID-Lösungen*
- *2.2. Software-RAID-Lösung*
  - *2.2.1. Konfiguration*
- *2.3. Unterschied/Gemeinsamkeiten*
- *2.4. Zukunftsaussichten*
- *2.5. Links*

3. Diskussion

1. Was ist RAID?

- Mit RAID werden einzelne unabhängige Bock Devices zu einer logischen Einheit zusammengefügt.



- Als Block Devices kommen normalerweise SCSI- oder IDE-Festplatten zum Einsatz.
- Die logische Einheit bietet den Vorteil, dass die Daten redundant gespeichert werden können und dass sich der Datendurchsatz deutlich erhöhen lässt (Achtung: nicht bei allen RAID-Levels).

1.1. Wo wurde RAID entwickelt?

An der Universität of California, USA, wurde der Begriff RAID 1987 zum ersten mal geprägt. Auch wurden hier die meisten RAID-Levels definiert.

1.2. Welche RAID-Levels gibt es?

Die wichtigsten RAID-Levels:

- Linear mode
- RAID-0 oder Data Stripping
- RAID-1 oder Drive Mirroring
- RAID-4 oder Block Striping mit Parity-Laufwerk
- RAID-5 oder Block Striping mit verteilter Parity
- RAID-10 oder Mirrored Striping Array

Mit zunehmender RAID-Level-Zahl nimmt auch die Ausfallsicherheit zu.

1.2.1. Linear mode

- Mehrer Festplatten werden zu einer logischen Einheit zusammengefügt.
- Daten werden solange auf die erste Festplatte geschrieben, bis die Speicherkapazität dieser Platte erschöpft ist. Erst danach wird die zweite Festplatte beschrieben. Wenn sich mehr als zwei Platten in der logischen Einheit befinden, wiederholt sich dieses Schema.
- Es werden keine Redundanz-Daten erzeugt. Die Datensicherheit ist hier nicht höher als bei einer einzelnen Festplatte.
- Keine Performancesteigerung gegenüber einzelnen Festplatten.
- Speicherkapazität: (Anzahl der Festplatten) * (Speicherkapazität einer Festplatte)

1.2.2. RAID-0 oder Data Stripping

- Mehrer Festplatten werden zu einer logischen Einheit zusammengefügt.
- Schreibzugriffe werden auf die einzelnen Platten verteilt. Die Speicherkapazität der einzelnen Platten wird immer zu gleichen Teilen ausgenutzt.
- Es werden keine Redundanz-Daten erzeugt. Die Datensicherheit ist hier nicht höher als bei einer einzelnen Festplatte.
- Die Lese- und Schreibperformance (Datendurchsatz und Zugriffszeit) ist auf Grund der Parallelisierung deutlich gesteigert.
- Speicherkapazität: (Anzahl der Festplatten) * (Speicherkapazität einer Festplatte)

1.2.3. RAID-1 oder Drive Mirroring

- Mehrer Festplatten werden zu einer logischen Einheit zusammengefügt. Es können auch Ersatzfestplatten (Spare-Disks) definiert werden.
- Die Daten werden auf allen Festplatten gleichzeitig vorgehalten.
- Fallen alle Platten bis auf eine aus, so ist noch immer die Datensicherheit gewährleistet. Ist eine Ersatzfestplatte vorhanden so wird diese in die logische Einheit integriert und die Reorganisation der RAID-1-Einheit startet.
- Die Lese- und Schreibperformance (Datendurchsatz und Zugriffszeit) ist auf Grund der Parallelisierung deutlich gesteigert. Besonders der Lesedatendurchsatz ist annähernd (Anzahl Platten)*(Datendurchsatz einer Platte).
- Speicherkapazität: Speicherkapazität einer Festplatte

1.2.4. RAID-4 oder Block Striping mit Parity-Laufwerk

- Drei oder mehr Festplatten werden zu einer logischen Einheit zusammengefügt. Es können auch Ersatzfestplatten (Spare-Disks) definiert werden.
- Eine Festplatte der Einheit speichert nur Parity-Informationen. Die restlichen Platten werden nach dem RAID-0 verwaltet. Die Parity-Platte begrenzt die Speicherkapazität des Arrays.
- Fällt eine Platte aus, so ist die Datenintegrität gegeben. Fallen mehrer Platten aus, so sind die Daten verloren. Ist eine Ersatzfestplatte vorhanden so wird diese in die logische Einheit integriert und die Reorganisation startet.
- Die Leseperformance ist wie bei RAID-0. Die Schreibperformance ist so hoch wie die der Paity-Festplatte, da bei jedem Schreibvorgang die Parity-Daten geschrieben werden müssen.
- Speicherkapazität: (Anzahl der Festplatten im RAID-0-Verbund) * (Speicherkapazität einer Festplatte)

1.2.5. RAID-5 oder Block Striping mit verteilter Parity

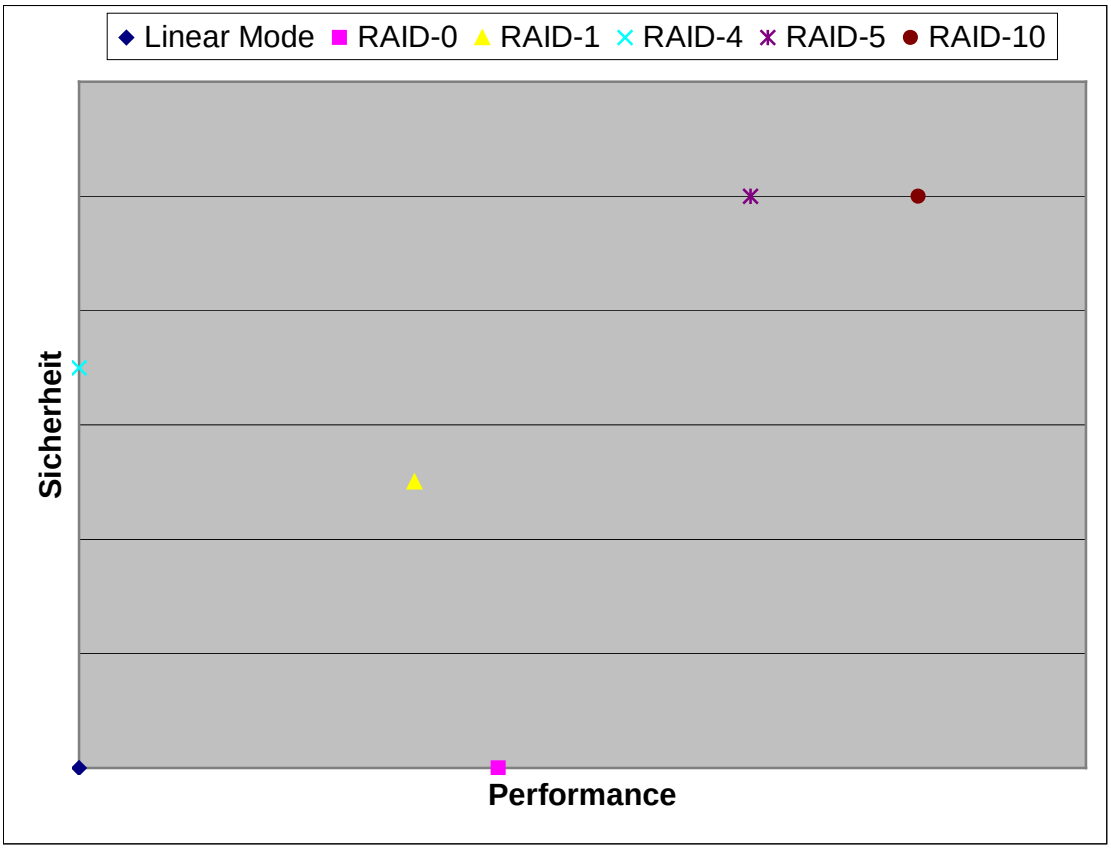
- Drei oder mehr Festplatten werden zu einer logischen Einheit zusammengefügt. Es können auch Ersatzfestplatten (Spare-Disks) definiert werden.
- Jede Festplatte des Array speichert die Nutzdaten und Parity-Informationen.
- Fällt eine Festplatte aus, so sind die Daten aufgrund der verteilten Parity-Daten intakt. Ist eine Ersatzfestplatte vorhanden so wird diese in die logische Einheit integriert und die Reorganisation startet. Fallen zwei oder mehr Festplatten gleichzeitig aus so sind die Daten verloren.
- Die Lese- und Schreibperformance ist deutlich höher als die Performances eines Laufwerkes aus dem Array. Ein genauer Faktor kann hier aber nicht angegeben werden.
- Speicherkapazität: (Anzahl der Festplatten - 1) * (Speicherkapazität einer Festplatte)

1.2.6. RAID-10 oder Mirrored Striping Array

- RAID-10 ist eine Kombination aus RAID-1- und RAID-0-Einheiten. Es werden zwei RAID-1-Einheiten zu einer großen RAID-0-Einheit zusammengefügt.
- Fällt eine Festplatte aus, so ist die Datensicherheit noch gegeben. Fallen mehrer Festplatten aus, so ist es wichtig das die einzelnen RAID-1-Arrays noch funktionstüchtig bleiben. Ansonsten sind die Daten verloren.
- Die Ausfallsicherheit von RAID-10 ist gleich wie bei RAID-1.
- Die Lese- und Schreibperformance ist deutlich höher als die Performances eines Laufwerkes aus dem Array. Ein genauer Faktor kann hier aber nicht angegeben werden.
- Speicherkapazität: (Anzahl der RAID-1-Einheiten) * (Speicherkapazität einer Festplatte)

1.3. Wo setzt man RAID ein?

Es gibt keine exakte Zuordnung zwischen RAID-Level und Problemstellungen. Die Zuordnung ist prinzipiell von der Problemstellung abhängig:



Die Praxis zeigt, daß hauptsächlich RAID-1, RAID-5 und RAID-10 in Servern verwendet werden. In Workstations wird oft RAID-0 und RAID-1 eingesetzt.

2. Linux und RAID

RAID-Unterstützung wird schon seit einiger Zeit für Linux geboten. Die meisten Namenhaften RAID-Hersteller werden von Linux unterstützt. So ist z.B. seit der Kernel-Version 2.0.31 die ICP vortex Controller-Serie unterstützt. Auch Überwachungs- und Steuerungsprogramm sind vorhanden. Treiber kommen entweder vom Hersteller selber oder werden von der Linux-Gemeinde entwickelt.
Hier eine kurze (unvollständig) Liste:

- Mylex
- Compaq
- IBM
- AMI
- GDT
- ...

Software-RAID wird seit der Kernel-Generation 2.0 standardmäßig unterstützt.

2.1. Hardware-RAID-Lösungen

Vorteile:

- Die Berechnung der Parity-Daten ist sehr rechenintensiv. Die Algorithmen für die Parity-Informationen sind optimal in Hardware zu fassen und Hardware-Lösungen sind somit sehr schnell. Dadurch wird der CPU des Rechners nicht zusätzliche Rechenleistung für die Verwaltung des RAID-Arrays genommen.
- Hardware-RAID-Lösungen sind für das Betriebssystem transparent.
- Controller hat eigenes BIOS und kann darüber (beliebig) konfiguriert werden.
- Eigener großer Cache für Lese- und Schreibzugriffe.

Nachteile:

- Preis
- Konfigurationstools sind für „exotische“ Betriebssysteme nicht vorhanden (auch für Linux sind bisher wenige solcher Tools vorhanden).

2.2. Software-RAID-Lösung

Vorteil:

- Volle Kontrolle über den Quellcode.
- Volle Kontrolle über die Konfiguration der RAID-Einheit.
- Keine Kosten für Hardware-RAID-Controller.
- Verwaltung des Array durch das Betriebssystem.

Nachteil:

- Software-RAID-Lösung für das Betriebssystem nicht transparent.
- Booten nur von RAID-1 möglich.
- JFS sind im Moment nicht auf Software-RAID-Einheiten anwendbar.
- Kein eigener Prozessor für die Berechnung der Parity-Informationen.
- Eingeschränkte Möglichkeit das RAID-Level ohne Datenverlust zu wechseln.

2.2.1. Konfiguration

Ab der Kernel-Generation 2.4. wird der neue RAID-Code in den Kernel integriert. Bisher wird Kernel-Generation 2.2. über ein Patch mit dem neuen RAID - Code upgedatet. Für die Kernel-Generation 2.0. ist der aktuelle RAID-Code nicht verfügbar.

Die eigentliche Konfigurationsdatei heißt raidtab und liegt normalerweise im Verzeichnis etc.
Folgende Einstellungen lassen sich darüber machen:

- RAID-Level
- zu verwendene Festplatten
- Ersatzfestplatten
- ...

Mit den RAIDtools lässt sich das Array dann steuern. Auf dem Array lassen sich genau die gleichen Filesystemoperationen durchführen wie auf eine einzelne Harddisk (Beschränkung: JFS).

2.3. Unterschiede/Gemeinsamkeiten

Gemeinsamkeiten:

- Geschwindigkeit
- Ausfallsicherheit
- Konfigurationsprobleme

Unterschiede:

- Preis
- Transparenz für das Betriebssystem
- JFS Unterstützung

2.4. Zukunftsaussichten

Da bei den Hardware-Controllern die Entwicklung in den letzten Jahren keine wirklichen Neuerungen hervorgebracht hat, kann man davon ausgehen, daß in Zukunft auch keine Neuerungen auf den Markt kommen werden.

Die Software-Lösung unter Linux ist noch in der Entwicklung. Die fundamentalen Funktionen sind verfügbar und auch (sehr!) stabil. Aber einige wichtige Funktionen fehlen noch (JFS-Unterstützung, Wechseln des RAID-Levels, ...)

2.5. Links

Linux-Magazin 12/99 „Software-RAID mit raidtools“
http://www.t-king.de/linux/raid1.html

SOFTWARE-RAID1 UNTER SUSE 6.0 BIS 6.2
http://www.t-king.de/linux/raid1.html

RAID und Linux
http://www.infodrom.north.de/~joey/Linux/raid/paper.html

Linux Software-RAID HOWTO
http://www.tu-harburg.de/~semb2204/dlhp/HOWTO/DE-Software-RAID-HOWTO.html

The Software-RAID HOWTO
http://www.linuxdoc.org/HOWTO/Software-RAID-HOWTO.html

Linux DPT Hardware RAID HOWTO
http://www.ram.org/computing/linux/dpt_raid.html

Root RAID HOWTO cookbook
http://www.linuxdoc.org/HOWTO/Root-RAID-HOWTO.html

Boot + Root + Raid + Lilo : Software Raid HOWTO
http://www.linuxdoc.org/HOWTO/mini/Boot%2BRoot%2BRaid%2BLILO.html

RAID Solutions for Linux
http://linas.org/linux/raid.html