

Die allerwichtigsten Raid Systeme

Michael Dienert

9. März 2020

Vorbemerkung

Dieser Artikel gibt eine knappe Übersicht über die wichtigsten RAID Systeme.

Inhaltsverzeichnis

1 Die Abkürzung RAID	2
1.1 Fehlerraten und MTTF	2
1.2 Erhöhung der Zuverlässigkeit durch Redundanz	2
1.3 Erhöhung der Leistung durch Verteilen der Daten (Striping)	2
2 Die verschiedenen RAID Levels	3
2.1 RAID Level 0: ohne Redundanz	3
2.2 RAID Level 1: Spiegelung	3
2.3 RAID Level 2: Hamming-Codes	4
2.4 RAID Level 3: einfache Parität	5
2.5 RAID Level 4: Block Interleaved Parity	5
2.6 RAID Level 5: Block Interleaved Distributed Parity	5
2.7 RAID Level 6: P+Q Redundanz	6
2.8 RAID Level 10: Striping und Mirroring	6

1 Die Abkürzung RAID

Die Abkürzung RAID bedeutet:

Redundant
Array of
Inexpensive
Disks

Bei einem RAID-System werden mehrere physikalische Festplatten zu einer großen, *logischen* Platte hoher Leistung zusammengeschaltet.

Durch das Verteilen der Daten auf mehrere physikalische Laufwerke, auf die parallel zugegriffen werden kann erreicht man zwei Dinge:

1. höhere Datentransferraten beim Zugriff auf große Dateien
2. schnellere I/O-Raten bei Zugriffen auf kleine Dateien

Die Abkürzung RAID wird heute oft mit Redundant Array of *Independent* Disks angegeben. Die einzelnen Platten sind aber gerade *nicht* unabhängig von einander (Ausnahme evtl. RAID 1) abhängig. Bei RAID 0 sind sie sogar zu 100% voneinander abhängig.

1.1 Fehlerraten und MTTF

Der Vorteil der höheren Leistung wird aber durch eine höhere Fehleranfälligkeit erkauft:

Die durchschnittliche Zeit bis ein Fehler eintritt (**MTTF**=Mean Time To Failure) verkleinert sich linear mit der Anzahl der Laufwerke.

Ein Beispiel: eine einzelne Festplatte habe eine MTTF von 500 000 Stunden. Bei einem Array aus 100 solcher Laufwerke würde es im Mittel nur noch 5000 Stunden dauern bis ein Fehler auftritt.

1.2 Erhöhung der Zuverlässigkeit durch Redundanz

Um den Nachteil der verringerten Zuverlässigkeit eines Plattenarrays wieder auszugleichen, speichert man die Daten mit einem Redundanzschema ab.

Redundanz bedeutet, dass man zum Beispiel zusätzliche Bits für Fehlerkorrektur einführt (Hamming-Codes) oder die Daten doppelt abspeichert (Mirroring).

1.3 Erhöhung der Leistung durch Verteilen der Daten (Striping)

Beim sog. **Striping** oder Interleaving werden die Daten über mehrere Laufwerke verteilt. Dieses Verfahren verbessert die I/O- oder Übertragungsraten des Gesamtsystems, da je nach Verfahren:

- mehrere, unabhängige Lese- oder Schreibzugriffe gleichzeitig ausgeführt werden können
- bei einem Zugriff auf einen grösseren Datenblock durch Koordination der Einzellaufwerke die Transferrate erhöht werden kann.

Beim Striping unterscheidet man noch ob die Daten in kleinen Einheiten (Feinkörnig) oder in grösseren Paketen (grobkörnig) über die Platten verteilt werden.

2 Die verschiedenen RAID Levels

2.1 RAID Level 0: ohne Redundanz

Beim RAID Level 0 werden die Daten *ohne* Redundanz gespeichert. Das heisst also, dass die Zuverlässigkeit des Systems kleiner ist als die einer Einzelplatte!

Vorteil von RAID Level 0 ist die hohe Leistung. Anwendung findet RAID Level 0 vor allem in Parallel-Rechner-Umgebungen, bei denen es mehr auf Kapazität und Geschwindigkeit als auf Zuverlässigkeit ankommt. Bild 1 zeigt die Verteilung der Datenblöcke über mehrere Platten.

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4
Block1	Block2	Block3	Block4	Block5
Block6	Block7	Block8	Block9	Block10
Block11	Block12	Block13	Block14	Block15
Block16	Block17	Block18	Block19	Block20
Block21	Block22	Block23	Block24	Block25

Abbildung 1: RAID Level 0

2.2 RAID Level 1: Spiegelung

Beim RAID Level 1 werden die Daten immer doppelt auf zwei physikalischen Laufwerken gespeichert. Fällt ein Laufwerk aus, arbeitet man mit der Kopie

auf dem verbleibenden Laufwerk weiter. Auch bei Lesezugriffen bietet die Spiegelung Vorteile: da die Daten auf den beiden Platten identisch sind, hat man beim Zugriff freie Wahl und so kann man die Platte mit der kürzeren Zugriffszeit wählen. Die Zugriffszeit hängt dabei von der Lage des Plattenstapels zum Schreib-Lesekopf ab.

Der RAID Level 1 wird vielfach bei Datenbankanwendungen eingesetzt. Hier kommt es mehr auf Zugriffszeit und Verfügbarkeit als auf Speicherplatz-Effizienz an. Bild 2 zeigt die Anordnung zweier gespiegelter Platten.

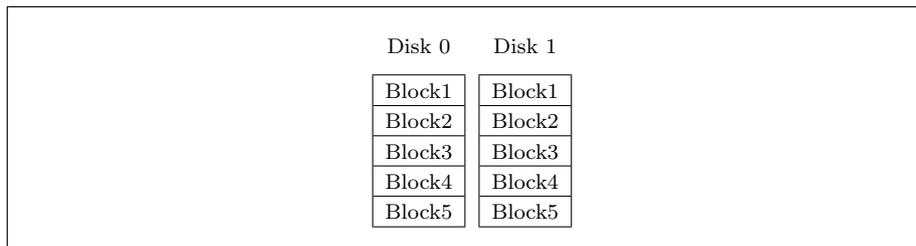


Abbildung 2: RAID Level 1

2.3 RAID Level 2: Hamming-Codes

Dieses RAID Level verwendet die bekannten Hamming-Codes zur Fehlerkorrektur und Erkennung. Kommt zum Beispiel der Hamming Code 4+3 zum Einsatz, d.h. 4 Nutzdatenbits plus 3 Paritätsbits, benötigt man 7 Laufwerke. Bild 3 zeigt diese Anordnung.

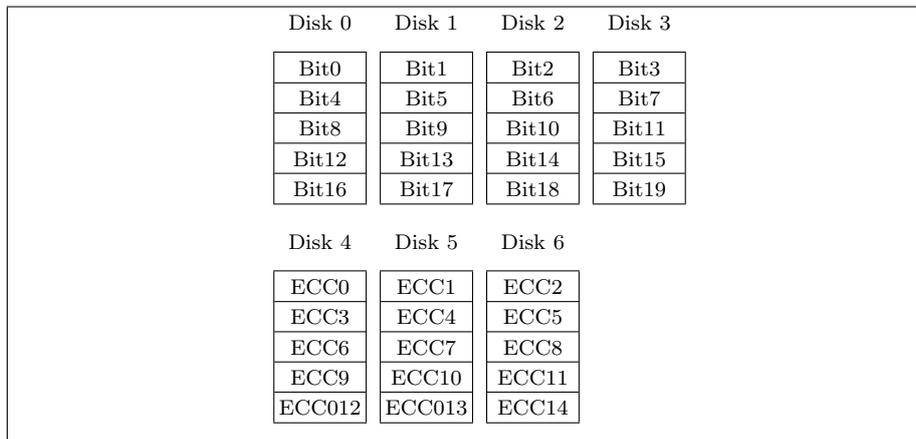


Abbildung 3: RAID Level 2

Bei nur 4 Nutzdatenbits muss ein 32 bit Datenwort in 8 aufeinanderfolgenden

Schritten gespeichert werden, was die Leistung des Systems stark bremst. Eine 1988 vorgestellte, schnelle Version des RAID Level 2 bestand daher aus 32 Platten für die Nutzdatenbits plus 7 zusätzlichen Platten für die Paritätsbits, aus insgesamt also 39 (!) Platten.

RAID Level 2 wird kaum verwendet, da das System sehr aufwändig ist und da die modernen Festplatten bereits eine eigene Fehlerkorrektur mit Hamming-Codes eingebaut haben.

2.4 RAID Level 3: einfache Parität

Um den Nachteil des hohen Aufwands bei Level 2 etwas abzumildern, wurde Level 3 eingeführt: anstelle einer echten Hamming-Codierung verwendet man nur ein einzelnes Paritätsbit.

Normalerweise kann man mit einem einfachen Paritätsbit nur feststellen, ob ein 1-bit-Fehler aufgetreten ist, aber man kann ihn nicht mehr korrigieren, da man nicht weiss, wo er aufgetreten ist.

Da die Einzellaufwerke des Systems jedoch selbst wieder ein ECC-Redundanz Schema verwenden, können die Plattenkontroller leicht feststellen, welche Platte den Fehler produziert hat und ihn dann einfach korrigieren.

Dieser RAID Level wird auch **Bit Interleaved Parity** genannt: die Daten sind *feinkörnig*, eben bitweise über die Platten verteilt und durch ein Parity-Bit abgesichert.

Bild 4 zeigt ein 32+1 Bit System.

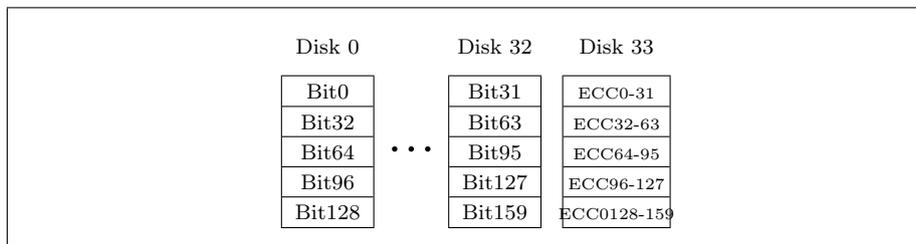


Abbildung 4: RAID Level 3, einfaches Paritätsbit

2.5 RAID Level 4: Block Interleaved Parity

Dieses System ist dem RAID Level 3 sehr ähnlich. Nur werden die Daten nicht bitweise sondern blockweise, also grobkörnig auf die Platten verteilt. Da nur eine Platte für die Parität vorhanden ist (Abb. 5), wird diese Platte leicht zum Engpass, weshalb dieses System nicht sehr oft zum Einsatz kommt. Ein anderer, schwerwiegender Nachteil tritt bei Transfers kleiner Datenmengen auf: man

benötigt insgesamt 4 Plattenzugriffe, bis die neuen Daten geschrieben, die Parität neu berechnet (lese alte Parität und alte Daten) und wieder geschrieben ist

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4	Disk 5
Block1	Block2	Block3	Block4	Block5	Pty1-4
Block6	Block7	Block8	Block9	Block10	Pty6-10
Block11	Block12	Block13	Block14	Block15	Pty11-15
Block16	Block17	Block18	Block19	Block20	Pty16-20
Block21	Block22	Block23	Block24	Block25	Pty21-25

Abbildung 5: RAID Level 4, Block Interleaved

2.6 RAID Level 5: Block Interleaved Distributed Parity

Um den Engpass der Paritätsplatte zu umgehen, kann man die Paritätsinformation auch auf den Platten verteilen und gelangt dann zu einer Anordnung wie in Bild 6 gezeigt.

Der Vorteil des RAID Level 5 ist die beste Leistung beim Schreiben grosser und beim Lesen kleiner Blöcke aller Systeme, die mit Redundanz arbeiten.

Die RAID Level 1 und 3 kann man als Sonderfälle des RAID Level 5 betrachten.

Disk 0	Disk 1	Disk 2	Disk 3	Disk 4
Block1	Block2	Block3	Block4	Pty1-4
Block6	Block7	Block8	Pty5-8	Block5
Block11	Block12	Pty9-12	Block9	Block10
Block16	Pty13-16	Block13	Block14	Block15
Pty17-20	Block17	Block18	Block19	Block20

Abbildung 6: RAID Level 5, Block Interleaved, Distributed Parity

2.7 RAID Level 6: P+Q Redundanz

Der RAID Level 6 ist vom Aufbau her sehr ähnlich dem RAID Level 5. Nur wird anstelle eines einfachen Paritätsbits eine P+Q Redundanz mit Reed-Solomon Codes verwendet.

Die P+Q Parität schützt vor bis zu zwei Plattenfehlern.

2.8 RAID Level 10: Striping und Mirroring

Der RAID Level 10 ist einfach eine Kombination aus RAID Level 0 und 1. Daher auch der Name 10. Zur Verdeutlichung noch ein Bild: Abb. 7.

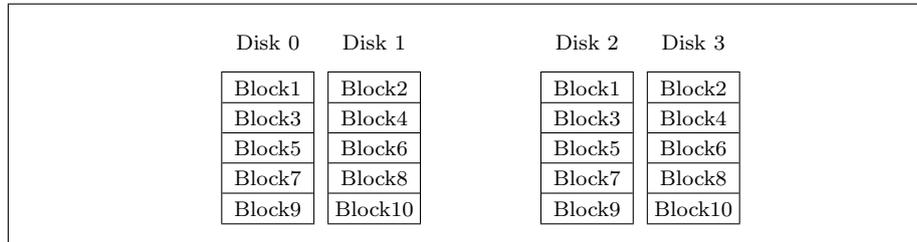


Abbildung 7: RAID Level 10, Mirroring und Striping kombiniert