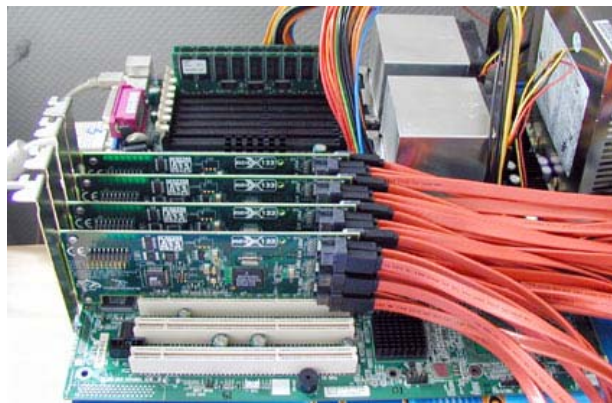


RAID - Ein Überblick

von Oliver Ott



<http://www.tomshardware.de/storage/20041001/>



<http://www.tomshardware.de/storage/20041001/>

Inhaltsverzeichnis

1. Einleitung.....	3
1.1 Begriffsklärung	3
1.2 RAID-Definition	3
1.3 Ziele von RAID.....	3
2. Geschichte.....	4
2.1 Warum wurde RAID notwendig?.....	4
2.2 Die Erfinder von RAID.....	4
2.3 Die ersten RAID-Systeme.....	5
2.4 Die weitere Entwicklung.....	5
2.5 Standardisierung von RAID.....	5
3. Grundlagen.....	6
3.1 Stripe Size / Striping-Granularität.....	6
3.2 Exklusiv-ODER-Verknüpfung / XOR.....	6
4. RAID ohne Fehlerkorrektur	7
4.1 JBOD.....	7
4.2 RAID 0.....	8
4.3 RAID 1.....	9
4.4 RAID 01.....	10
4.5 RAID 10.....	11
4.6 Vergleich RAID 01 zu RAID 10.....	12
4.7 Fazit zu RAID ohne Fehlerkorrektur.....	13
5. RAID mit Fehlerkorrektur.....	14
5.1 RAID 2.....	14
5.2 RAID 3.....	15
5.3 RAID 4.....	16
5.4 RAID 5.....	17
6. RAID 6 und proprietäre RAID-Verfahren.....	18
6.1 RAID 6.....	18
6.2 RAID 7.....	18
6.3 RAID 5E.....	18
6.4 RAID 5EE.....	18
6.5 RAID DP.....	18
7. Kombinations-RAIDs.....	19
7.1 Wozu Kombinations-RAIDs?	19
7.2 Beispiele für Kombinations-RAIDs	19
7.2.1 RAID 00	19
7.2.2 RAID 01	19
7.2.3 RAID 10	19
7.2.4 RAID 05	19
7.2.5 RAID 50	19
7.2.6 RAID 15	19
7.2.7 RAID 51	19
7.2.8 RAID 55	19
8. Das RAID-Dreieck.....	20
9. Implementierungsformen von RAID.....	21
9.1 Hardware RAID	21
9.1.1 Vorteile.....	21
9.1.2 Nachteile.....	21
9.2 Software RAID	21
9.2.1 Vorteile.....	21
9.2.2 Nachteile.....	21
9.3 Hybrid RAID.....	21
10. Verhalten bei Festplattenausfall	22
10.1 Hot-Fix / Hot-Spare.....	22
10.2 Hot-Plug / Hot-Swap.....	22
11. RAID - Das Fazit.....	23
12. Quellenverzeichnis.....	24

1. Einleitung

1.1 *Begriffsklärung*

RAID - **R**edundant **A**rray of **I**nexpensive **D**isks

frei übersetzt -> Redundanter Verbund preiswerter Festplatten

alternativ:

RAID - **R**edundant **A**rray of **I**ndependent **D**isks

frei übersetzt -> Redundanter Verbund unabhängiger Festplatten

1.2 *RAID-Definition*

RAID ist ein Verfahren zur Speicherung von Daten, bei dem zusätzlich zu den reinen Nutzdaten Fehlerkorrekturcodes (z. B. in Form von Paritätsbits oder Hamming-Codes) auf mindestens 2 Festplatten verteilt gespeichert werden.

RAID 'widerspricht' somit den meisten im EDV-Bereich verwendeten Techniken, welche im allgemeinen darauf ausgelegt sind Redundanzen (also zusätzliche oder doppelte Daten) zu vermeiden.

Im Gegensatz dazu werden bei RAID gezielt redundante Informationen erzeugt, auf welche beim Ausfall einer Komponente zurückgegriffen werden kann, um die Funktionalität des gesamten Systems zu erhalten.

1.3 *Ziele von RAID*

Ein RAID-System wird stets mit dem Ziel betrieben in mindestens einem Aspekt leistungsfähiger zu sein, als der Betrieb von einer oder mehrerer einzelnen Festplatten.

Je nach Aufbau können mit dem Einsatz eines RAID-Systems folgende Vorteile erreicht werden:

- Steigerung der Geschwindigkeit des Festplattenzugriffes
- Steigerung der Zuverlässigkeit der Datenspeicherung
- Verringerung der gesamten Systemkosten
- Gewährleistung von Hochverfügbarkeit

2. Geschichte

2.1 Warum wurde RAID notwendig?

Gegen Ende der 80er Jahre war die Speicherung großer Datenmengen eine heikle Angelegenheit. Zwar war mit SLEDs (Single Large Expensive Disks) im 14-Zoll-Format eine Technik vorhanden, die - für damalige Verhältnisse - große Kapazitäten (zwischen 2 und 3 Gigabyte) und eine relativ gute Datensicherheit bot, jedoch auch extrem teuer war.

Als billige Alternative boten sich die damals neu auf dem Markt erschienenen 'kleinen' Festplatten im 5,25" Format an. Der Einsatz dieser neuen Technik warf jedoch auch viele Probleme auf. So lag die Zuverlässigkeit der 5,25" Festplatten bewiesenermaßen unterhalb der jener SLEDs. Weiterhin steigt die Wahrscheinlichkeit eines Ausfalls beim Einsatz mehrerer Festplatten deutlich an.

Eine Lösung musste gefunden werden ...

2.2 Die Erfinder von RAID

Im Juni 1988 veröffentlichten die drei Doktoranden David Patterson, Garth Gibson und Randy Katz der Berkeley University of California ihre Studie mit dem Namen "A Case for Redundant Arrays of Inexpensive Disks (RAID)".

Ihre Lösung sah vor mehrere kleine Festplattenlaufwerke zu einem Verbund zusammenzuschließen, und diesen anschließend mit Fehlererkennungs- und Fehlerkorrekturmechanismen auszustatten, um damit der nun erhöhten Ausfallwahrscheinlichkeit entgegenzuwirken.

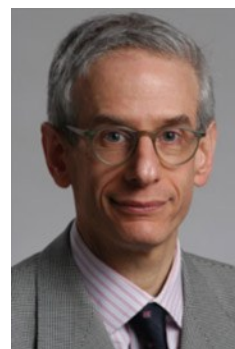
RAID war geboren!



David Patterson
<http://www.cs.berkeley.edu>



Garth Gibson
<http://www.cs.cmu.edu>



Randy Katz
<http://bnrg.eecs.berkeley.edu>

2.3 Die ersten RAID-Systeme

Patterson, Gibson und Katz begannen nun damit die ersten funktionierenden RAID-Systeme aufzubauen.



'RAID the First'
<http://bnrg.eecs.berkeley.edu>



'RAID the Second'
<http://bnrg.eecs.berkeley.edu>

Das Hauptziel in den Anfängen von RAID war die Kostenreduzierung durch den Einsatz vieler kleiner, aber preiswerter 5,25" Festplatten, anstatt wie bis dahin üblich eine große, aber sehr teure SLED zu verwenden.

Daher auch die ursprüngliche Bezeichnung für RAID, nämlich 'Redundant Array of **Inexpensive** Disks' ('Redundanter Verbund **preiswerter** Festplatten').

2.4 Die weitere Entwicklung

Aufgrund der schnellen Verbreitung und technischen Weiterentwicklung kam RAID zunehmend in Servern zum Einsatz, welche vor allem aus dem erhöhten Datendurchsatz und der erhöhten Ausfallsicherheit ihren Nutzen zogen. Das Ziel der Kosteneinsparung verlor nach und nach immer mehr an Bedeutung. Heutzutage sind moderne High-End-RAID-Systeme sogar um ein vielfaches teurer als vergleichbare Einzelaufwerke.

Der zusätzlichen Möglichkeit in einem modernen RAID-System Festplatten während des laufenden Systembetriebs (Stichwort: Hot-Plug bzw. Hot-Swap) zu wechseln, entspricht wohl am ehesten der heute üblichen Bezeichnung für RAID, nämlich 'Redundant Array of **Independent** Disks' ('Redundanter Verbund **unabhängiger** Festplatten').

2.5 Standardisierung von RAID

Seit 1992 erfolgt eine Standardisierung von RAID durch das RAB (**RAID** **A**dvisory **B**oard). Diese Organisation, bestehend aus ca. 50 Herstellern von Speichersystemen und Speicherkomponenten hat das Ziel dem Anwender die Speichertechnik und angewandte Speichertechnologien näher zu bringen.

3. Grundlagen

3.1 *Stripe Size / Striping-Granularität*

Mit dem Begriff Stripe Size wird die Größe des aus einem oder mehreren Datenblöcken bestehenden zusammenhängenden Datenbereichs als kleinste adressierbare Einheit eines RAID-Systems bezeichnet.

Wird die Stripe-Size vergrößert, erhöht sich der maximale Datendurchsatz, gleichzeitig erhöht sich aber auch die Zugriffszeit.

Eine nicht an die anfallenden Daten angepasste Stripe-Size kann einerseits zu einer Leistungsminderung führen (dies ist dann der Fall, wenn große Dateien zu speichern sind, aber die Stripe-Size zu niedrig gewählt wurde) auf der anderen Seite ist auch eine extreme Platzverschwendung möglich (dann, wenn viele kleine Dateien anfallen, aber die Stripe-Size zu hoch gewählt wurde).

Üblich sind Stripe-Größen im Bereich von 64kByte bis 256kByte.

3.2 *Exklusiv-ODER-Verknüpfung / XOR*

auch 'Antivalenz' oder 'Kontravalenz' genannt ist ein Begriff aus der Aussagenlogik.

Eine Exklusiv-ODER-Aussage ist dann wahr, wenn entweder die erste Aussage oder die zweite Aussage wahr ist, aber nicht beide! Die Formulierung "Entweder A oder B" entspricht am besten dem Ausdruck "A XOR B".

In der Praxis entspricht eine XOR-Verknüpfung folgendem:

Ist die Summe eine gerade Zahl ergibt sich 0, ist sie ungerade ergibt sich 1.

Jedes RAID-Verfahren mit Fehlerkorrektur arbeitet nach der XOR-Methode, da der Vorteil der XOR-Verknüpfung ist, dass sie umkehrbar (symmetrisch) ist.

Beispiel:

	10011001	-----	10011001	10011001	
	01100101	01100101	-----	01100101	
	11100110	11100110	11100110	-----	

XOR	00011010	00011010	00011010	00011010	
		-----	-----	-----	
		XOR	10011001	01100101	11100110

4. RAID ohne Fehlerkorrektur

4.1 JBOD

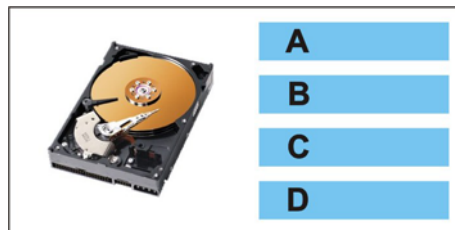
Der Begriff JBOD ist ein Akronym für 'Just a Bench Of Disks'

frei übersetzt etwa -> 'Nur ein Bündel Festplatten'

JBOD ist KEIN RAID-Verfahren, da aber fast jeder am Markt erhältliche RAID-Controller in der Lage ist Festplatten als JBOD zu organisieren, soll die Funktionsweise an dieser Stelle trotzdem erläutert werden.

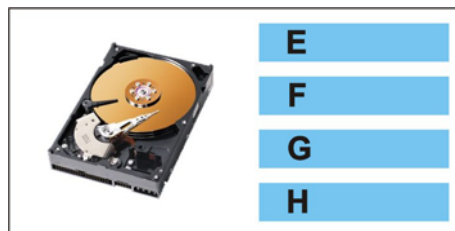
Beim JBOD-Verfahren werden mehrere Festplatten hintereinander in Reihe geschaltet, und dem Betriebssystem als eine große Festplatte zur Verfügung gestellt.

Die Nutzdaten werden zuerst auf die erste Festplatte des 'Bündels' geschrieben. Ist die gesamte erste Festplatte mit Daten belegt, wird anschließend auf die zweite im 'Bündel' vorhandene Festplatte geschrieben. Ist auch diese komplett mit Daten gefüllt, wird auf die dritte vorhandene Festplatte geschrieben ...



Das JBOD-Verfahren bietet keinerlei Datenredundanz. Beim Ausfall eines Laufwerkes sind die Daten verloren.

Auch bietet JBOD keine Leistungsvorteile im Gegensatz zu einem Einzellaufwerk.



Als einziger 'Vorteil' können bei JBOD Festplatten unterschiedlicher Kapazitäten, Bauart und Hersteller verwendet werden.

Mindestanzahl n benötigter Festplatten: $n = 2$

Maximal verfügbare Nutzkapazität k: $k = n$

Ausfallsicherheit s (in Festplatten): $s = 0$

4.2 RAID 0

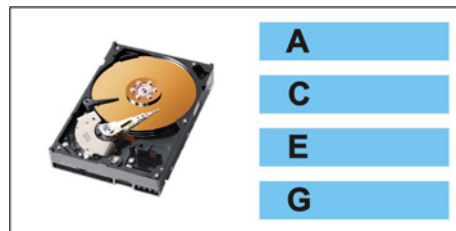
RAID 0 wird häufig auch als 'Striping' bezeichnet.

frei übersetzt etwa -> 'In Streifen zerlegen' von 'stripe' -> 'Der Streifen'

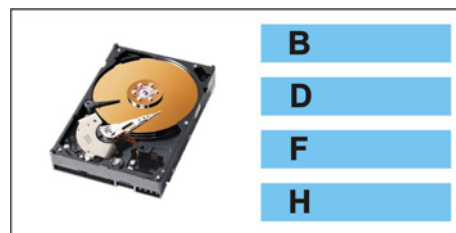
RAID 0 ist - ähnlich wie JBOD - kein 'echtes' RAID-Verfahren, da auch hier keinerlei redundante Informationen vorhanden sind.

Beim RAID 0 - Verfahren werden die Nutzdaten parallel auf mehrere Festplatten verteilt, was eine Erhöhung der Transferraten sowohl beim Lesen als auch beim Schreiben zur Folge hat. RAID 0 kann seine Leistung vor allem bei großen, zusammenhängenden Dateien ausspielen (Striping-Granularität beachten!). Der Einsatzzweck erstreckt sich demnach auf Bereiche, in denen große Datenmengen schnell verarbeitet werden müssen, und es weniger auf Datensicherheit ankommt, z. B. die Audio- und Videobearbeitung.

Aufgrund der Tatsache das beim Ausfall eines Laufwerks die Daten des kompletten RAID-Verbunds verloren sind kommt RAID 0 gewöhnlich nicht in Serversystemen zum Einsatz, da dort die Datensicherheit eine unabdingbare Anforderung ist.



Auch das RAID 0 - Verfahren bietet keinerlei Datenredundanz. Beim Ausfall eines Laufwerkes sind die Daten des kompletten RAID-Verbunds verloren.



RAID 0 bietet jedoch - insbesondere bei großen Dateien - ein vielfaches an Lese- und Schreibleistung gegenüber einem einzelnen Laufwerk.

Mindestanzahl n benötigter Festplatten: $n = 2$

Maximal verfügbare Nutzkapazität k: $k = n$

Ausfallsicherheit s (in Festplatten) : $s = 0$

4.3 RAID 1

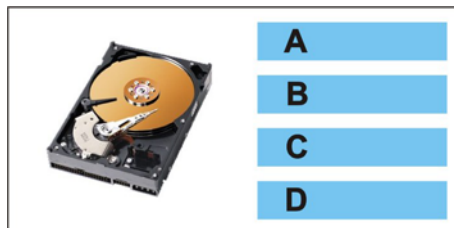
RAID 1 wird häufig auch als 'Mirroring' bezeichnet.

frei übersetzt etwa -> 'gespiegelt' von 'mirror' -> 'Der Spiegel'

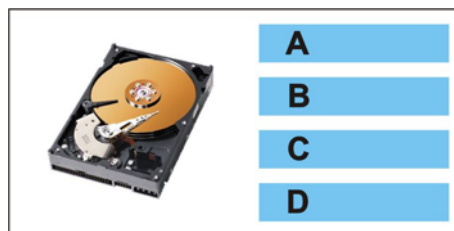
Der Name alleine verdeutlicht schon wie das RAID 1 - Verfahren arbeitet.

Die Nutzdaten werden identisch auf mehrere Festplatten geschrieben, so dass jede einzelne Festplatte ein exaktes Abbild der gesamten Daten enthält. Es ist somit eine 100%ige Redundanz vorhanden. Fällt ein Laufwerk im Verbund aus, stehen die Daten noch zur Verfügung. Die Transferraten beim Lesen von einem RAID 1 - Verbund sind höher, beim Schreiben in etwa gleich wie die eines einzelnen Laufwerks. Der Einsatzzweck von RAID 1 liegt in Bereichen, in denen wichtige Daten zum hauptsächlichen Lesezugriff vorgehalten werden, z. B. Datenarchive.

Obwohl RAID 1 eine optimale Ausfallsicherheit bietet wird dieses Verfahren für gewöhnlich nur in sehr kleinen Serversystemen verwendet. Werden große Nutzkapazitäten benötigt wird ein RAID 1 - System schnell sehr teuer, da die doppelte Anzahl an Festplatten benötigt wird um die Redundanz zu erhalten.



Das RAID 1 - Verfahren bietet eine 100%ige Datenredundanz. Beim Ausfall eines Laufwerks sind die Daten noch verfügbar.



Ein Ausfall *aller* Festplatten im Verbund führt zum Totalverlust der Daten.

Jedoch verdoppeln sich die Kosten für die Datenhaltung.

Mindestanzahl n benötigter Festplatten: $n = 2$

Maximal verfügbare Nutzkapazität k: $k = \frac{n}{2}$

Ausfallsicherheit s (in Festplatten): $s = \frac{n}{2}$

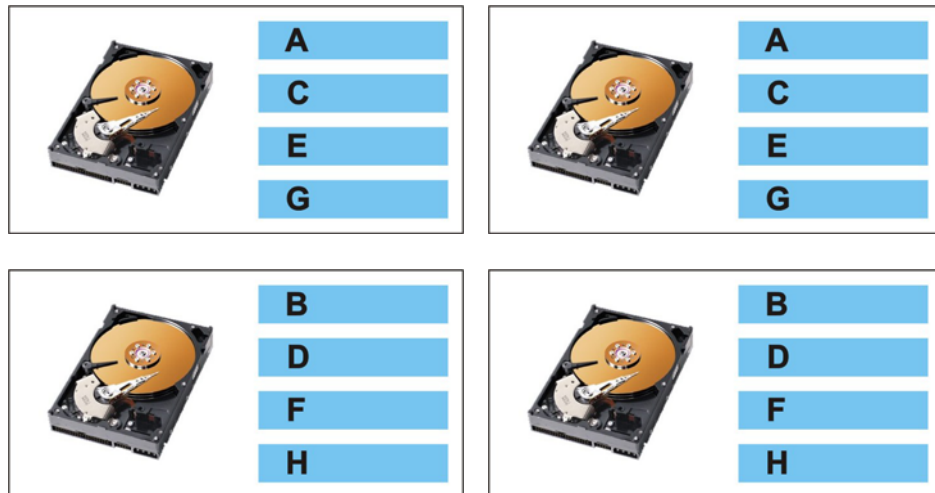
4.4 RAID 01

auch als RAID 0+1, RAID 0,1 oder RAID 0/1 bezeichnet.

Nicht zu verwechseln mit RAID 10 (siehe auch 4.5 und 4.6)!

RAID 01 kombiniert RAID 0 (Striping) mit RAID 1 (Mirroring) um die Transferraten zu steigern und gleichzeitig eine 100%ige Redundanz zu gewährleisten.

Ein RAID 01 - System ist ein RAID 1 - Verbund über mehrere RAID 0.



Beim RAID 01 - Verfahren wird zunächst in vertikaler (senkrechter) Richtung ein RAID 0 - System (Striping -> keine Redundanz) gebildet, welches anschließend in horizontaler (waagrechter) Richtung zu einem RAID 1 - System (Mirroring -> 100% Redundanz) verbunden wird.

Auch ein RAID 10 (siehe 4.5) kombiniert RAID 0 mit RAID 1.
Die Unterschiede zwischen RAID 01 und RAID 10 werden in 4.6 erklärt.

Mindestanzahl n benötigter Festplatten: $n = 4$

Maximal verfügbare Nutzkapazität k: $k = \frac{n}{2}$

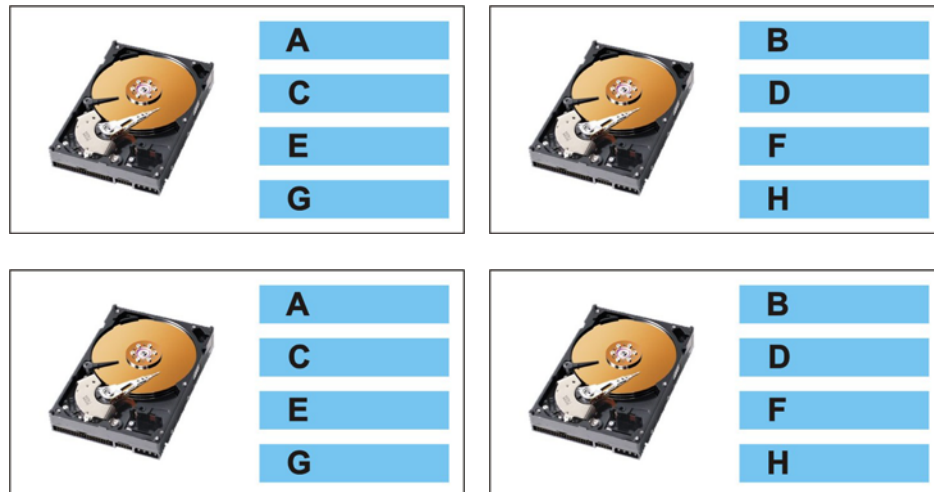
Ausfallsicherheit s (in Festplatten) : $s = \frac{n}{2}$

4.5 RAID 10

Nicht zu verwechseln mit RAID 01 (siehe auch 4.4 und 4.6)!

Auch RAID 10 kombiniert RAID 0 (Striping) mit RAID 1 (Mirroring) um die Transferraten zu steigern und gleichzeitig eine 100%ige Redundanz zu gewährleisten.

Ein RAID 10 - System ist ein RAID 0 - Verbund über mehrere RAID 1.



Beim RAID 10 - Verfahren werden zunächst in vertikaler (senkrechter) Richtung mehrere RAID 1 - Systeme (Mirroring -> 100% Redundanz) gebildet, welche anschließend in horizontaler (waagrechter) Richtung zu einem RAID 0 - System (Striping -> keine Redundanz) verbunden werden.

Auch ein RAID 01 (*siehe 4.4*) kombiniert RAID 0 mit RAID 1.
Die Unterschiede zwischen RAID 10 und RAID 01 werden in 4.6 erklärt.

Mindestanzahl n benötigter Festplatten: $n = 4$

Maximal verfügbare Nutzkapazität k: $k = \frac{n}{2}$

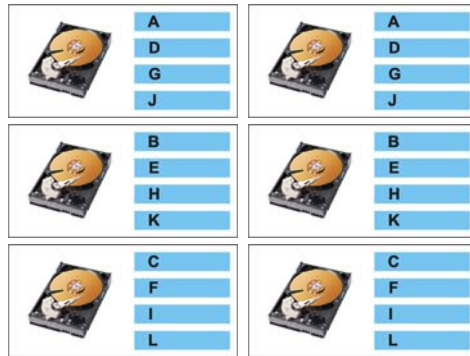
Ausfallsicherheit s (in Festplatten) : $s = \frac{n}{2}$

4.6 Vergleich RAID 01 zu RAID 10

Was ist nun der Unterschied zwischen RAID 01 und RAID 10?

Am Beispiel von 6 Festplatten soll nun ein Vergleich stattfinden.

RAID 01



RAID 01 bietet gleiche Leistung und Nutzkapazität wie RAID 10, aber geringere Ausfallsicherheit.

RAID 10



RAID 10 liefert eine schnellere Datenrekonstruktion, da nach einem Ausfall nur ein Teil der Daten rekonstruiert werden muss.

Bei RAID 01 werden je 3 Festplatten zu einem RAID 0 (Striping) verbunden, und anschließend zu einem RAID 1 (Mirroring) gespiegelt.

Bei RAID 10 werden je 2 Festplatten zu einem RAID 1 (Mirroring) gespiegelt, und anschließend zu einem RAID 0 (Striping) verbunden.

Die Nutzkapazität beträgt in beiden Fällen $n = 3$ Festplatten.

Der Unterschied liegt in der Datensicherheit:

Fällt im RAID 01 - System eine Festplatte aus ist das betroffene RAID 0 komplett unbrauchbar. Die Daten stehen auf dem zweiten RAID 0 zwar noch zur Verfügung, aber fällt zusätzlich ein Laufwerk aus diesem Verbund aus, sind die Daten verloren.

Im anderen Fall verliert eine gespiegelte Festplatte im RAID 10 - System durch einen Ausfall ebenfalls ihre Redundanz, aber nur wenn zusätzlich noch das zweite Laufwerk exakt dieser Einheit ausfällt, kommt es zum Datenverlust.

4.7 Fazit zu RAID ohne Fehlerkorrektur

Zwar bietet RAID 0 (Striping) optimalen Datendurchsatz und RAID 1 (Mirroring) perfekte Datensicherheit, jedoch verursachen die Kombinations-systeme RAID 01 und RAID 10 einen hohen Overhead und dementsprechend auch hohe Kosten für die Datenhaltung.

Um diesen Nachteil zu beheben, arbeiten die RAID-Level 2 und höher mit Fehlerkorrekturmechanismen.

Sie verteilen die Nutzdaten auf mindestens zwei Datenlaufwerke und errechnen anschließend einen Korrekturwert, der auf einem zusätzlichen Parity-Laufwerk abgelegt wird. Mit Hilfe dieser Korrekturwerte können im Falle einer defekten Festplatte die Daten wieder rekonstruiert werden.

RAID-Systeme verwenden dazu eines der ältesten Verfahren zur Fehlerkorrektur, die sog. Paritätsprüfung mit XOR (*siehe auch 3.2*).

Dazu werden die Nutzdaten mit einer Exklusiv-ODER-Operation verknüpft und das Ergebnis auf dem Parity-Laufwerk gespeichert.

Es gibt heute eine Unzahl von verschiedenen RAID-Leveln, welche mit Fehlerkorrektur arbeiten. Viele davon sind Herstellereigene Verfahren (z. B. RAID 5E bzw. 5EE von *IBM*, RAID 7 von *Storage Computer* und RAID DP von *Network Appliance*).

Alle diese Verfahren bauen auf den grundlegenden Erkenntnissen von Patterson, Gibson und Katz auf, weswegen auf die proprietären RAID-Level einzelner Hersteller nur kurz in *Kapitel 6* eingegangen werden soll.

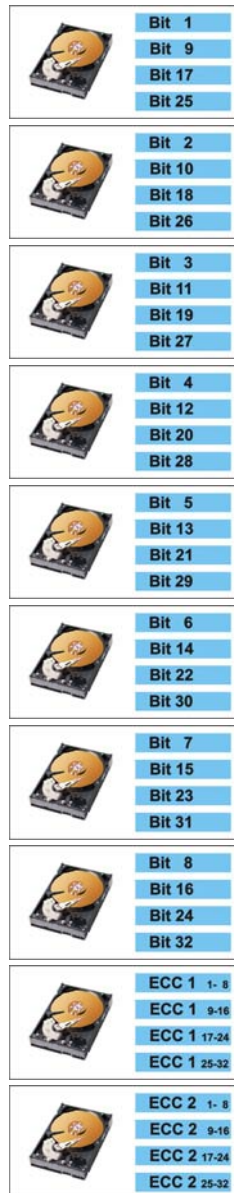
Das nächste Kapitel beschäftigt sich ausführlich mit den von Patterson, Gibson und Katz definierten RAID-Leveln mit Fehlerkorrektur, auch wenn einige hiervon in der Zwischenzeit nicht mehr oder nur noch selten genutzt werden.

5. RAID mit Fehlerkorrektur

5.1 RAID 2

auch als 'Hamming-System' bezeichnet (nach R. W. Hamming).

Das RAID 2 - Verfahren wird heute nicht mehr eingesetzt. Es wurde in den Anfängen der RAID-Technologie im Großrechnerbereich verwendet.



Bei RAID 2 werden die Nutzdaten bitweise auf mindestens 8 Datenlaufwerke geschrieben. Zusätzlich werden 2 Bit ECC nach dem Hamming-Algorithmus berechnet und auf 2 ECC-Laufwerke geschrieben.

Der Hamming-Algorithmus ist eine fehlerkorrigierende Codierung, die nach R. W. Hamming von den Bell Labs benannt ist.

RAID 2 war notwendig, da die frühen Festplatten noch keine internen ECC-Mechanismen aufwiesen, und deshalb Bitfehler die beim Schreiben auftraten nicht korrigiert werden konnten. In einem RAID 2 - System mit Hamming-Algorithmus ist neben der Erkennung eines Bitfehlers auch die exakte Bestimmung der Position möglich, auf welchem Datenlaufwerk der Fehler aufgetreten ist.

Das heutzutage in allen Festplatten integrierte ECC-Verfahren macht die Eigenschaft von RAID 2 einzelne Bitfehler korrigieren zu können überflüssig.

Die Transferrate beim Lesen erhöht sich durch den parallelen Zugriff auf die Datenlaufwerke auf das achtfache. Beim Schreiben liegt die Transferrate bedingt durch den hohen ECC-Overhead unter jene eines einzelnen Laufwerks.

Mindestanzahl n benötigter Festplatten: $n = 10$

Maximal verfügbare Nutzkapazität k: $k = 0.8 n$

Ausfallsicherheit s (in Festplatten): $s = 2$

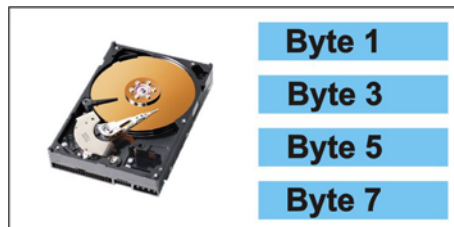
5.2 RAID 3

RAID 3 kann als Vorgänger des heutigen, modernen RAID 5 gesehen werden.

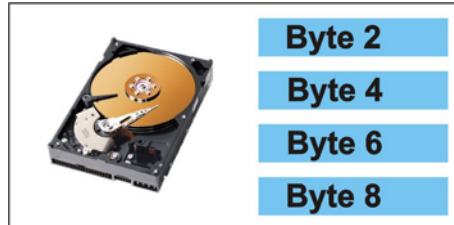
Beim RAID 3 - Verfahren werden die Nutzdaten byteweise auf mindestens 2 Datenlaufwerke gespeichert.

Im Gegensatz zu RAID 2 verlässt sich RAID 3 zur Erkennung von Bitfehlern beim Schreiben und Lesen auf die in den Festplatten integrierten Funktionen. Über XOR werden aus den Daten die Paritätsbits berechnet und auf ein dediziertes Parity-Laufwerk gespeichert.

RAID 3 synchronisiert die Kopfpositionen der einzelnen Laufwerke beim Schreiben der ECC-Daten. Dies ermöglicht ein schnelles Schreiben und Lesen großer, zusammenhängender Dateien. Bei vielen kleinen, verteilten Dateien erfordert die synchrone Neupositionierung aller Festplatten im Verbund entsprechend Zeit. Das Einsatzgebiet von RAID 3 liegt deshalb vor allem bei der Verarbeitung von großen, zusammenhängenden Datenmengen.



RAID 3 speichert die Nutzdaten parallel auf mehrere Datenlaufwerke. Die durch XOR berechneten Paritätsbits werden auf ein zusätzliches dediziertes Laufwerk geschrieben.



RAID 3 ist inzwischen fast vollständig durch RAID 5 ersetzt worden, welches kein dediziertes Parity-Laufwerk aufweist.



Das dedizierte Parity-Laufwerk, auf das in einem RAID 3 - System alle Paritätsbits geschrieben werden, erweist sich als 'Flaschenhals' des gesamten Systems.

Mindestanzahl n benötigter Festplatten: $n = 3$

Maximal verfügbare Nutzkapazität k: $k = n-1$

Ausfallsicherheit s (in Festplatten): $s = 1$

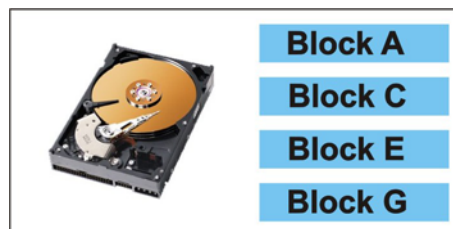
5.3 RAID 4

RAID 4 stellt eine - wenn auch nur minimale - Verbesserung von RAID 3 dar.

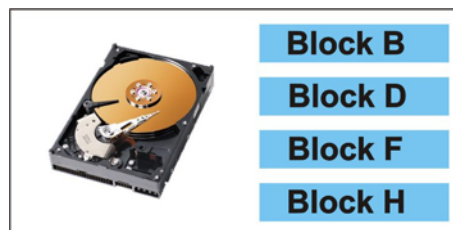
Ein RAID 4 - System arbeitet weitgehend genauso wie ein RAID 3.

Da moderne Festplatten und Betriebssysteme nicht mehr mit einzelnen Bytes, sondern mit Blöcken arbeiten, verwendet RAID 4 eine blockweise Verteilung der Nutzdaten auf wiederum mindestens 2 Datenlaufwerke. Die Paritätsbits werden mittels XOR gebildet und - wie bei RAID 3 - auf ein dediziertes Parity-Laufwerk gespeichert.

Um die Lesezugriffe auch auf kleine Dateien zu beschleunigen werden bei RAID 4 die Kopfpositionen der einzelnen Laufwerke nicht synchronisiert. Dies führt jedoch dazu das RAID 4 beim Schreiben RAID 3 unterlegen ist, da die passende Stelle auf dem Parity-Laufwerk erst gefunden und angesteuert werden muss.



Auch RAID 4 speichert die Nutzdaten parallel auf mehrere Laufwerke und die durch XOR erhaltenen Paritätsbits auf ein dediziertes Parity-Laufwerk.



RAID 4 wird in der Praxis nur selten eingesetzt, da es im Vergleich zu RAID 3 nur minimale Vorteile besitzt, im Vergleich zu RAID 5 durch das fest definierte Parity-Laufwerk jedoch einen erheblichen Nachteil aufweist.



Das dedizierte Parity-Laufwerk, auf das in einem RAID 4 - System alle Paritätsbits geschrieben werden, erweist sich auch hier als 'Flaschenhals' des Systems.

Mindestanzahl n benötigter Festplatten: $n = 3$

Maximal verfügbare Nutzkapazität k : $k = n-1$

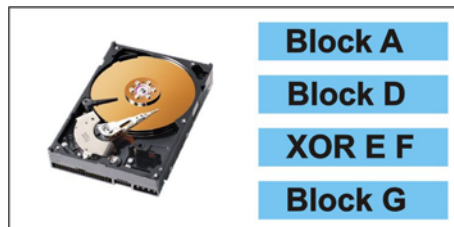
Ausfallsicherheit s (in Festplatten): $s = 1$

5.4 RAID 5

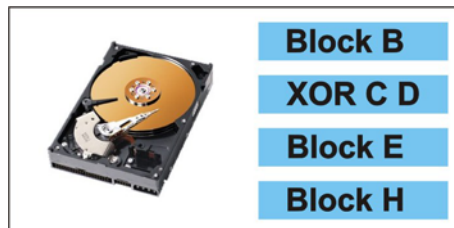
RAID 5 ermöglicht eine gesteigerte Datentransferrate beim Lesen als auch Datenredundanz bei relativ geringen Kosten zu implementieren und ist deshalb die heute beliebteste und meist verbreitete RAID-Variante.

Das RAID 5 - Verfahren arbeitet wie RAID 4 mit einer blockweisen Verteilung der Nutzdaten, jedoch verzichtet es auf ein dediziertes Parity-Laufwerk und verteilt die ECC-Daten gleichmäßig über die Laufwerke. Damit verteilt sich die mechanische Belastung aller Festplatten gleichmäßig, da nun kein Laufwerk mehr eine Sonderstellung als Parity-Laufwerk einnehmen muss.

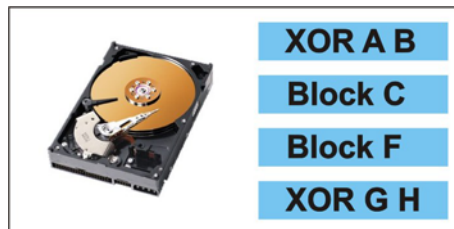
Beim Lesen erreicht ein RAID 5 - System Datentransferraten wie RAID 4. Der Schreibzugriff läuft aufgrund der gleichmäßigen Verteilung der Paritätsbits jedoch schneller ab als auf einem RAID 4 (jedoch deutlich langsamer als auf einem RAID 0).



RAID 5 verteilt die Nutzdaten UND die durch XOR errechneten Paritätsbits parallel über alle Laufwerke des RAID-Verbunds.



Durch den Verzicht auf ein dediziertes Parity-Laufwerk werden nun alle Festplatten gleichmäßig mechanisch belastet, was die Ausfallsicherheit noch einmal zusätzlich erhöht.



Der Verzicht auf das dedizierte Parity-Laufwerk entfernt den 'Flaschenhals' eines RAID 3 bzw. RAID 4 - Verbunds aus dem System.

Mindestanzahl n benötigter Festplatten: $n = 3$

Maximal verfügbare Nutzkapazität k: $k = n-1$

Ausfallsicherheit s (in Festplatten): $s = 1$

6. RAID 6 und proprietäre RAID-Verfahren

6.1 RAID 6

RAID 6 arbeitet ähnlich wie RAID 5, kann aber den Ausfall von maximal 2 Festplatten kompensieren. Bei RAID 6 werden nicht ein, sondern zwei Werte zur Fehlerkorrektur berechnet und über alle Laufwerke im Verbund verteilt. Jedoch ist der Rechenaufwand zum Erstellen der Korrekturwerte erheblich höher als bei RAID 5. Die Rekonstruktion des Verbunds erfordert bei 2 ausgefallenen Festplatten das Lösen von linearen Gleichungssystemen mit Matrizen und Umkehrmatrizen.

6.2 RAID 7

Auch das Herstellereigene RAID 7 ist ähnlich wie RAID 5 aufgebaut. Der Hersteller *Storage Computer* setzt zusätzlich im RAID-Controller ein lokales Echtzeitbetriebssystem ein und große Pufferspeicher koppeln die Laufwerke vom Datenbus ab. Wie RAID 6 unterstützt RAID 7 die Verwendung mehrerer Paritätsinformationen.

6.3 RAID 5E

RAID 5E vom Hersteller *IBM* steht für RAID 5 Enhanced. Es kombiniert ein RAID 5 mit einem Hot-Spare. Hierbei wird der Hot-Spare allerdings nicht als dediziertes Laufwerk ausgeführt, sondern auf die einzelnen Festplatten aufgeteilt. Auf jeder Festplatte wird für den Fall eines Ausfalls Speicherplatz reserviert, auf welchem im Fehlerfall der Inhalt des ausgefallenen Laufwerks mit Hilfe der Parität wiederhergestellt wird. Der Vorteil liegt nicht in einer zusätzlichen Sicherheit gegenüber RAID 5, sondern in der Nutzung aller vorhandenen Festplatten, auch der Hot-Spare-Platte.

6.4 RAID 5EE

RAID 5EE arbeitet fast identisch wie RAID 5E, nur dass hier der freie Speicherplatz nicht am Ende der Festplatten reserviert wird, sondern diagonal über die Laufwerke verteilt. Dadurch ergibt sich eine höhere Leistung bei der Wiederherstellung der Daten im Fehlerfall.

6.5 RAID DP

RAID DP steht für RAID Double Parity und ist ein von *Network Appliance* entwickeltes RAID-Verfahren, bei dem - ähnlich wie bei RAID 6 - zwei Paritäten verwendet werden, was eine erhöhte Ausfallsicherheit mit sich bringt. Im Gegensatz zu RAID 6 müssen bei RAID DP im Wiederherstellungsfall jedoch nur zwei XOR-Operationen ausgeführt werden.

7. Kombinations-RAIDs

7.1 Wozu Kombinations-RAIDs?

Wie bereits im *Kapitel 4* 'RAID ohne Fehlerkorrektur' gezeigt, können verschiedenste RAID-Verfahren miteinander kombiniert werden.

Die RAID-Varianten 01 (*siehe 4.4*) und 10 (*siehe 4.5*) sind zum Beispiel Kombinationen der RAID-Level 0 und 1.

Der Sinn und Zweck warum einzelne RAIDs wiederum zu übergeordneten RAID-Systemen verbunden werden liegt darin einen Aspekt der RAID-Eigenschaften zusätzlich zu verstärken, z. B. eine noch höhere Datensicherheit zu erreichen oder die Datentransferrate noch weiter zu steigern.

7.2 Beispiele für Kombinations-RAIDs

7.2.1 RAID 00

Ein RAID 00-Verbund besteht aus einem RAID 0, das aus mehreren RAID 0 besteht.

7.2.2 RAID 01

siehe 4.4

7.2.3 RAID 10

siehe 4.5

7.2.4 RAID 05

Ein RAID 05-Verbund besteht aus einem RAID 5, das aus mehreren RAID 0 besteht.

7.2.5 RAID 50

Ein RAID 50-Verbund besteht aus einem RAID 0, das aus mehreren RAID 5 besteht.

7.2.6 RAID 15

Ein RAID 15-Verbund besteht aus einem RAID 5, das aus mehreren RAID 1 besteht.

7.2.7 RAID 51

Ein RAID 51-Verbund besteht aus einem RAID 1, das aus mehreren RAID 5 besteht.

7.2.8 RAID 55

Ein RAID 55-Verbund besteht aus einem RAID 5, das aus mehreren RAID 5 besteht.

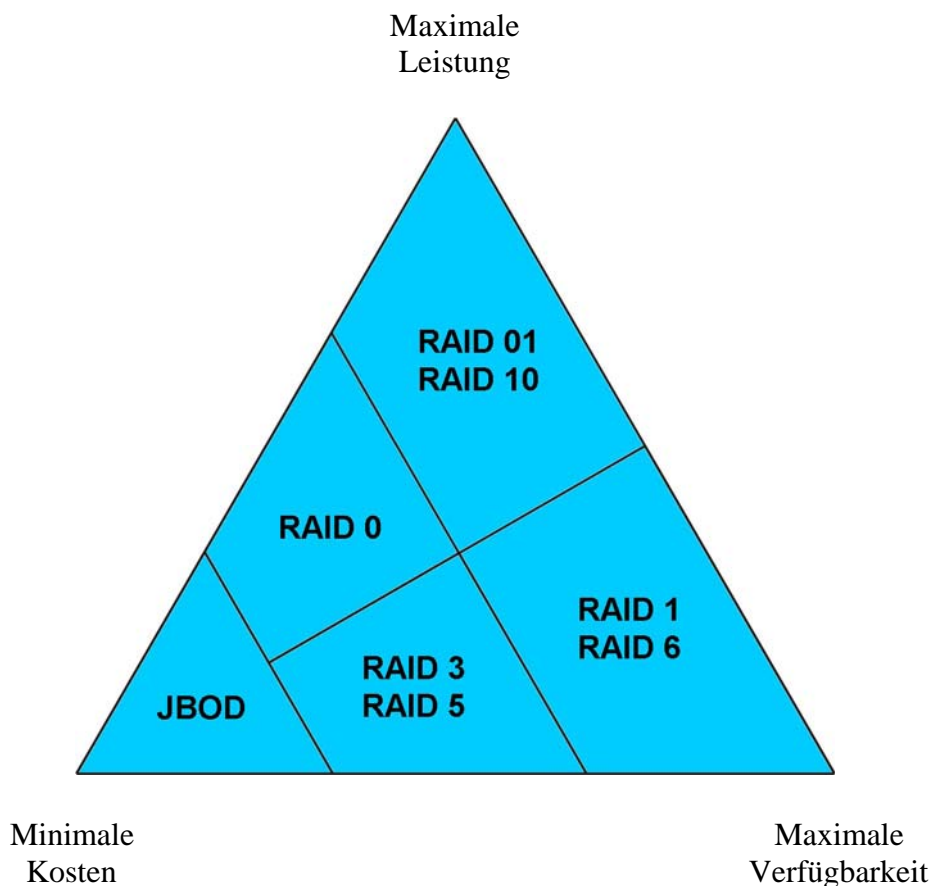
8. Das RAID-Dreieck

Die Auswahl eines geeigneten RAID-Levels erfordert eine genaue Abwägung zwischen den drei Faktoren Systemleistung, Systemverfügbarkeit und Systemkosten.

Eine Organisation der Laufwerke als JBOD verursacht die geringsten Kosten, lässt jedoch in Sachen Ausfallsicherheit und Geschwindigkeit zu wünschen übrig.

Ein RAID 1 - Verbund garantiert dagegen - dank einer 100%igen Redundanz - die höchste Verfügbarkeit. Er produziert jedoch auf der anderen Seite auch den höchsten Kapazitäts-Overhead, und somit auch die höchsten relativen Kosten.

Fasst man alle Faktoren zusammen erhält man das typische RAID-Dreieck:



9. Implementierungsformen von RAID

9.1 *Hardware RAID*

Von Hardware RAID spricht man, wenn ein spezieller Hardware-Baustein, der sog. RAID-Controller das Zusammenspiel der im RAID-Verbund betriebenen Laufwerke organisiert. Der Begriff Hardware RAID ist diesbezüglich etwas irreführend, da letztendlich auch ein RAID-Controller zum Betrieb spezielle Software benötigt.

9.1.1 Vorteile

- Plattformunabhängig
- Hohe Datentransferraten
- Niedrige CPU-Last am Host

9.1.2 Nachteile

- Hohe Implementationskosten

9.2 *Software RAID*

Beim Software RAID übernimmt eine auf der CPU des Hosts laufende spezielle Software die Steuerung und Organisation des Festplattenverbunds. Oft stellt das Betriebssystem entsprechende Softwarekomponenten bereit. So unterstützen die Server Versionen von Windows wie auch Linux softwareseitig die RAID-Level 0, 1 und 5.

9.2.1 Vorteile

- Hohe Datentransferraten (bei RAID 0, 1, 01 und 10)
- Niedrige Implementationskosten

9.2.2 Nachteile

- Plattformabhängig
- Hohe CPU-Last am Host (bei allen RAIDs mit Parität)

9.3 *Hybrid RAID*

Hybride RAID Lösungen sind auf der einen Seite zwar Hardware Bausteine, aber um sie für den Consumer-Bereich erschwinglich zu machen, überlässt man die RAID-Logik der CPU des Hosts, was sie dann noch schlechter aussehen lässt als eine reine Software-Lösung. Diese RAID-Variante wird oft auf Mainboards verbaut, häufig mit der Einschränkung auf RAID 0 oder 1. Im Linux-Jargon werden sie oft auch als 'Fake RAID' bezeichnet.

10. Verhalten bei Festplattenausfall

Ob Hardware- oder Software RAID: Fällt ein Laufwerk des Verbunds aus, geht bei allen gängigen RAID-Leveln unweigerlich die Redundanz verloren. Jede zusätzliche Fehlfunktion einer weiteren Festplatte führt in dieser Situation zu einem Datenverlust.

Die defekte Festplatte muss deshalb schnellstmöglich ausgetauscht und der komplette RAID-Verbund rekonstruiert werden.

10.1 Hot-Fix / Hot-Spare

Im Optimalfall verfügt der Verbund über ein zusätzliches, ausschließlich für den Notfall gedachtes Ersatzlaufwerk. Eine solche Hot-Fix- bzw. Hot-Spare-Festplatte wird automatisch aktiviert und als Ersatz für das ausgefallene Laufwerk eingebunden. Anschließend wird vollautomatisch die Wiederherstellung des RAID-Verbunds gestartet.

10.2 Hot-Plug / Hot-Swap

Steht kein Hot-Spare-Laufwerk zur Verfügung muss die defekte Festplatte manuell gewechselt werden. Dies würde für gewöhnlich ein Abschalten des Rechners und somit eine Betriebsunterbrechung erfordern - was bei Servern aber in der Regel inakzeptabel ist. Abhilfe schaffen hier Hot-Plug- bzw. Hot-Swap-fähige RAID-Systeme. Die einzelnen Festplatten des Verbunds sind hier frei zugänglich in Wechselrahmen installiert, die sich auch während des laufenden Systembetriebs ohne Probleme austauschen lassen. Die Wiederherstellung erfolgt nun entweder automatisch (falls der Controller fähig ist, mit den Wechselrahmen zu kommunizieren), oder muss vom Administrator manuell eingeleitet werden.

Egal ob nun das RAID-System über Hot-Fix- bzw. Hot-Spare-Laufwerke verfügt, oder Hot-Plug- bzw. Hot-Swap-fähig ist, steht während der Wiederherstellungsphase auf dem Verbund keine Redundanz zur Verfügung. Die Datenredundanz steht erst nach erfolgter (erfolgreicher!) Rekonstruktion des RAID-Verbunds wieder zur Verfügung. Soll diese Problematik vermieden werden, muss ein RAID-System verwendet werden, das zweifache Paritätsinformationen bereit stellt (z. B. RAID 6 oder RAID DP).

11. RAID - Das Fazit

Durch den Einsatz von RAID-Systemen und RAID-Verbänden lässt sich die Systemleistung und Systemverfügbarkeit von Rechnersystemen deutlich erhöhen. Ein Allheilmittel gegen Datenverlust stellt die RAID-Technologie jedoch keinesfalls dar. Um eine annähernd 100%ige Ausfallsicherheit zu erreichen, müssen vielmehr alle Komponenten des gesamten Speichersubsystems inklusive den Controllern, den Netzteilen und den Lüftern redundant ausgelegt sein. Die Storage-Industrie bietet solche Lösungen selbstverständlich an, kostengünstig sind diese aber nicht gerade.

Zudem können in der Praxis gelegentlich Situationen eintreten, durch die sich die Ausfallwahrscheinlichkeit des gesamten Rechnersystems schlagartig erhöht. Hierzu zählen etwa Überspannungen durch Blitzschlag, Feuerschaden durch Brände oder Wasserschaden durch Überschwemmungen. Auch befallen Computerviren RAID-Systeme ebenso gern wie Einzelaufwerke. Gegen Schäden durch kriminelle Energie wie etwa Industriespionage oder böswilliges Löschen wichtiger Datenbestände sind auch RAID-Systeme nicht geschützt.

Schließlich kann auch das noch so zuverlässigste RAID-System den Risikofaktor Nummer 1 nicht ausschalten - den Menschen. Den weitaus größten Teil irreparabler Datenverluste verursacht nicht versagende Technik, sondern der Benutzer selbst. Gelöschte Daten sind - und bleiben - auch auf RAID-Systemen gelöscht.

Selbst für das ausgeklügeltste RAID-System gilt deshalb:

Den einzig wirklich zuverlässigen Schutz gegen Datenverlust bietet nur ein durchdacht geplantes und konsequent vorgenommenes Backup!

12. Quellenverzeichnis

- Bildquellennachweis

Titelseite: Mainboard mit SATA-RAID-Controllern
Quelle: Tom's Hardware Guide
<http://www.tomshardware.de/storage/20041001/>

Festplatten-Arrays
Quelle: Tom's Hardware Guide
<http://www.tomshardware.de/storage/20041001/>

Seite 4: Professor David Patterson
Quelle: Berkeley, University of California
<http://www.cs.berkeley.edu/~pattsrn/>

Professor Garth Gibson
Quelle: Carnegie Mellon University
<http://www.cs.cmu.edu/~garth/>

Professor Randy Katz
Quelle: Berkeley, University of California
<http://bnrg.eecs.berkeley.edu/~randy/>

Seite 5: RAID the First
Quelle: Berkeley, University of California
<http://bnrg.eecs.berkeley.edu/~randy/>

RAID the Second
Quelle: Berkeley, University of California
<http://bnrg.eecs.berkeley.edu/~randy/>

- Recherchequellennachweis

Acer, White Paper 'RAID Technology'
ICP Vortex, Broschüre 'Moderne RAID Technologie'
c't, Heft 10/2006, Seite 182, Artikel 'Datenbändigung'
Wikipedia, die freie Enzyklopädie; <http://de.wikipedia.org/>
TecChannel; <http://www.tecchannel.de/storage/grundlagen/401665/>

... und viel aus eigener Erfahrung im Umgang mit RAID-Systemen ☺

© 2006 Oliver Ott. Alle Rechte vorbehalten.