



(HTTP://COMPUTERWOCHE.DE)

(HTTPS://WWW.TECCHANNEL.DE/P/SUCHE)

Storage Grundlagen

RAID im Überblick - RAID 0 bis 7

(https://twitter.com/share?

url=https%3A%2F%2Fwww%2Etecchannel%2Ede%2Fa%2Fraid%2Dim%2Dueberblick%2Dgrundlagen%2Draid%2D0%2Dbis%2D7%2C401665&text=Storage%20Grundlagen%3A%20RAID%20i

(https://www.facebook.com/sharer.php?

u=https%3A%2F%2Fwww%2Etecchannel%2Ede%2Fa%2Fraid%2Dim%2Dueberblick%2Dgrundlagen%2Draid%2D0%2Dbis%2D7%2C401665&t=Storage%20Grundlagen%3A%20RAID%20im%2

(https://www.xing.com/app/user?

op=share&url=https%3A%2F%2Fwww%2Etecchannel%2Ede%2Fa%2Fraid%2Dim%2Dueberblick%2Dgrundlagen%2Draid%2D0%2Dbis%2D7%2C401665;title=Storage%20Grundlagen%3A%20

in (https://www.linkedin.com/shareArticle?

mini=true&url=https%3A%2F%2Fwww%2Etecchannel%2Ede%2Fa%2Fraid%2Dim%2Dueberblick%2Dgrundlagen%2Draid%2D0%2Dbis%2D7%2C401665&title=Storage%20Grundlagen%3A%2



09.02.2018 Von Jörg Luther, Bernhard Haluschak (Autor) ▼ und Christian Vilsbeck

(/a/print/raid-im-ueberblick-grundlagen-raid-0-bis-7,401665)

Die RAID-Technologie soll die Ausfallsicherheit und die Leistung von Festplattensystemen erhöhen - auch für den ambitionierten Desktop-Anwender. Dabei stehen Varianten mit spezifischen Vor- und Nachteilen zur Wahl, die unser Beitrag erläutert.

Als 1987 an der Universität von Kalifornien in Berkeley (<http://www.berkeley.edu/>) RAID aus der Taufe gehoben wurde, war die Speicherung großer Datenmengen noch eine reichlich knifflige Angelegenheit. Die in Rechenzentren gebräuchlichen SLEDs (Single Large Expensive Disks) im 14-Zoll-Format boten zwar hohe Kapazitäten (zwei bis drei GByte) und gute Datensicherheit (<http://www.tecchannel.de/storage/backup/>), waren jedoch extrem teuer. Als Massenspeicheralternative boten sich die in Relation zur Kapazität sehr viel billigeren, relativ neuen "kleinen" Drives im 5,25-Zoll-Formfaktor an. Deren Verwendung warf jedoch einige Probleme auf.

Die am meisten verbreiteten RAID-Modi in der Übersicht



Zum einen drohte die Speicherung auf vielen kleinen statt eines großen Laufwerks Verwaltungsprobleme zu beschieren. Bei einer Organisation der Platten als JBOD ("just a bunch of disks"), also als Bündel unabhängiger Drives, würde sich das Auffinden sowohl freien Speicherplatzes als auch einmal abgespeicherter Dateien recht kompliziert gestalten. Zum anderen lag die Zuverlässigkeit der "Mini-HDDs" per se schon deutlich unter jener der SLEDs. Erschwerend kam dazu, dass die statistische Wahrscheinlichkeit eines Datenverlustes

(http://www.tecchannel.de/storage/backup/401608/datenrettung_professionelle_hilfe_statt_datenverlust/) bei einer Speicherung der Daten auf mehreren Laufwerken statt eines einzelnen Drives zusätzlich ansteigt.

Als Lösung des Problems schlugen die drei Berkley-Doktoranden David Patterson, Garth Gibson und Randy Katz die Kombination mehrerer kleiner Laufwerke zu einem mit Fehlererkennungs- und Fehlerkorrekturmechanismen ausgestatteten, ausfallgesicherten Verbund vor. In ihrer im Juni 1988 veröffentlichten Studie verpassten sie der neuen Technik auch gleich einen eingängigen Namen: "A Case for Redundant Arrays of Inexpensive Disks (RAID)" (<http://www-2.cs.cmu.edu/~garth/RAIDpaper/Patterson88.pdf>) titelte das Papier, RAID war geboren.

In unserem Artikel stellen wir Ihnen die verschiedenen RAID-Varianten vor, die in Desktop-PCs (http://www.tecchannel.de/pc_mobile/), Servern (<http://www.tecchannel.de/server/>) und Speicherschränken (<http://www.tecchannel.de/storage/>) zum Einsatz kommen. Videos verdeutlichen in Animationen die Funktion der verschiedenen RAID-Technologien.

Vorteile von RAID

In ihrem ursprünglichen Papier (<http://www-2.cs.cmu.edu/~garth/RAIDpaper/Patterson88.pdf>) schlugen Patterson, Gibson und Katz insgesamt fünf verschiedene Methoden vor, mit denen sich einzelne Platten zu einem Array zusammenfassen lassen. Sie nummerierten sie in einer bis heute gültigen Terminologie als RAID-Level 1 bis 5 durch. Diese etwas unglückliche Bezeichnung sorgt immer wieder für Missverständnisse: Trotz der Bezeichnung "Level" handelt es sich nicht um stufenweise aufbauende Verfahren, sondern um von einander völlig unabhängige Techniken.

Neben den primär angepeilten Zielen - kostengünstige, hochkapazitive Speicherung (<http://www.tecchannel.de/storage/komponenten/>) und gute Ausfallsicherheit - bieten die vorgestellten Methoden weitere Vorteile. Zum einen stellt sich der RAID-Verbund auf Anwenderseite als einzelnes logisches Volume dar. Die Verwaltung gestaltet sich also ebenso einfach wie bei einem Einzellaufwerk. Zum anderen bieten viele RAID-Verfahren durch die Parallelisierung der Plattenzugriffe einen Geschwindigkeitsvorteil gegenüber Einzellaufwerken. Dieser lässt sich allerdings nur dann voll ausreizen, wenn eine ausreichende Kanalanzahl zur blockierungsfreien Ansprache der Laufwerke zur Verfügung steht.

Die Kombination derart vieler Vorteile sorgte nicht nur für eine rasche Verbreitung RAID-basierender Massenspeicher vor allem in Servern. Auch zusätzliche RAID-Varianten wurden entwickelt, die auf bestimmte Einsatzgebiete und spezifische Vorteile hin optimiert wurden. Heute reicht die Bandbreite der allgemein verfügbaren Level von RAID 0 bis RAID 6. Hinzu kommen kombinierte Technologien wie RAID 0+1 oder 50, mit RAID 6 verwandte Verfahren (RAID n+m) mit nochmals erhöhter Ausfallsicherheit, sowie herstellerspezifische Varianten. Zu den Letzteren zählen Techniken wie RAID-DP (Network Appliances (<http://www.netapp.com/us/products/platform-os/raid-dp.html>)), RAID-X (ECC Technologies) oder RAID ^n (Tandberg (<http://www.tandbergdata.com/>)).

Software- vs. Hardware-RAID

Die etwas irreführenden Begriffe Hard- beziehungsweise Software-RAID - letztlich benötigen beide Varianten zum Betrieb Software - beziehen sich auf die Art und Weise der Implementierung (<http://msdn.microsoft.com/de-de/library/ms178048.aspx>).

Beim Software-RAID übernimmt eine auf der CPU des Hosts laufende Software die Steuerung des Plattenverbunds. Oft bringt bereits das Betriebssystem entsprechende Komponenten mit. So beherrschen Windows Vista und XP sowohl RAID 0 als auch RAID 1 und 5 - letztere allerdings nur in der Server-Version. Linux verwaltet generell Arrays der Level 0, 1, 4 und 5 sowie ab Kernel 2.6 auch solche des Levels 6.

Software-RAID stellt also meist die preisgünstigste und einfachste Lösung dar. Zudem lässt es sich - etwa per Prozessor-Upgrade (http://www.tecchannel.de/pc_mobile/prozessoren/) am Host - relativ schnell an erhöhte Anforderungen anpassen. Andererseits verursacht es eine hohe CPU-Belastung und arbeitet naturgemäß plattform- und

betriebssystemgebunden. Auch stehen zur Ansteuerung der Laufwerke meist nur zwei oder vier Anschlüsse zur Verfügung. Dies beschränkt die mögliche Parallelisierung der Plattenzugriffe und damit auch die Performance.

Dagegen übernimmt bei Hardware-RAID ein eigener Controller die Ansteuerung des Arrays. Das bringt eine Entlastung der Host-CPU und eine höhere Performance mit sich. Zudem binden RAID-Controller die Laufwerke über mehrere Kanäle an, was gleichzeitige Laufwerkszugriffe und damit hohe Transferraten ermöglicht. Dafür gilt es aber, einen deutlich höheren Preis zu zahlen. Hardware-RAIDs arbeiten zwar plattformunabhängig. Auch sie benötigen aber zur Verwaltung Software, die natürlich auf ein bestimmtes Betriebssystem zugeschnitten ist.

Software- vs. Hardware-RAID

	Software-RAID	Hardware-RAID
Implementationskosten	niedrig	hoch
Performance	niedrig	hoch
CPU-Last am Host	hoch	niedrig
Plattformabhängigkeit	ja	nein
Betriebssystemabhängigkeit	ja	ja

Anforderungen an die Festplatten

Noch vor wenigen Jahren ließen RAID-Controller bei der Anforderung an die verwendeten Platten nicht mit sich spaßen. Als Interface war SCSI Pflicht, alle verwendeten Platten mussten identische Kapazitäten aufweisen. Oft kam es sogar vor, dass sich im Array nur Platten aus derselben Bauserie verwenden ließen.

Mittlerweise gestaltet sich das Pflichtenheft für die verwendeten Laufwerke deutlich weniger rigide. Im klassischen Serverbereich kommt schon aus Performancegründen zwar nach wie vor SCSI beziehungsweise der serielle Nachfolger SAS (http://de.wikipedia.org/wiki/Serial_Attached_SCSI) zum Einsatz. Im Entry-Level-Server oder der Workstation tauscht der preisbewusste Anwender diese inzwischen aber gern gegen einen kostengünstigeren Serial-ATA-HBA aus. In den Desktop-PC sind Serial-ATA-Controller inzwischen ebenfalls Standard.

Zudem erlauben moderne Controller wie auch Software-RAID-Lösungen den Mix von Platten verschiedener Kapazität im Verbund. Allerdings lässt sich dabei nicht die gesamte vorhandene Nettokapazität für das Array verwenden. Da die RAID-Verfahren gleich große Platten voraussetzen wird in einer Mischkonfiguration jedes Laufwerk nur bis zur Kapazität der kleinsten vorhandenen Festplatte (<http://www.tecchannel.de/produkte/storage/festplatten/>) genutzt. Bei einer Kombination eines 1-TByte-Laufwerks mit zwei 2-TByte-Disks stehen beispielsweise nur drei mal 1 TByte für das Array zur Verfügung.

RAID Level 0

Bei RAID Level 0 handelt es sich - wie die Null im Namen schon andeutet - um kein redundantes Speicherverfahren. Es dient lediglich zur Beschleunigung von Plattenzugriffen. Dazu fasst RAID 0 zwei oder mehr Festplatten zu einem logischen Laufwerk zusammen. Es verteilt die Daten in aufeinanderfolgenden Blöcken ("Stripes") gleichmäßig über alle Laufwerke. Daher bezeichnet man RAID 0 auch als Striping. Das parallele Lesen respektive Schreiben auf mehreren Laufwerken steigert zwar die Durchsatzrate, senkt jedoch die Sicherheit der Daten (http://www.tecchannel.de/storage/backup/401608/datenrettung_professionelle_hilfe_statt_datenverlust/): Fällt eine Platte des Verbunds aus, sind alle Daten verloren.

Die Geschwindigkeitssteigerung macht sich vor allem bei großen, zusammenhängenden Dateien deutlich bemerkbar. Hier kann Striping parallel auf allen Platten operieren und dadurch die Transferrate vervielfachen. Beim Lesen oder Schreiben vieler kleiner Files ist die Zugriffszeit der Platten der limitierende Faktor. Hier erreicht ein Stripeset bestenfalls die Performance eines Einzellaufwerks. Aufgrund dieser Eigenheiten kommt es meist dort zum Einsatz, wo große Datenmengen abgearbeitet werden müssen: Etwa bei Workstations für CAD/CAM oder Audio- und Videobearbeitung.

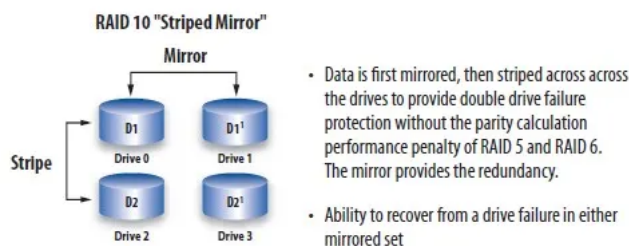
RAID Level 1

RAID Level 1 wird auch als Mirroring oder Spiegelung bezeichnet. Dieser Name verdeutlicht, wie das Verfahren arbeitet: Alle Schreibzugriffe erfolgen parallel auf zwei Laufwerke, so dass jede Platte quasi ein Spiegelbild der anderen darstellt. Alle Daten stehen also doppelt zur Verfügung - sicherer geht es kaum noch. Auch wenn eines der beiden Laufwerke komplett ausfällt, bleiben alle Nutzdaten erhalten. Allerdings steht bei RAID 1 nur die Hälfte der gesamten Plattenkapazität für die Speicherung zur Verfügung. Die Kosten der Datenhaltung (http://www.tecchannel.de/storage/management/2019472/tipps_storage_kosten_senken_compliance_archivierung_deduplizierung) verdoppeln sich also.

Bei Anbindung der Laufwerke an eigene Kanäle – Standard beispielsweise bei SATA und SAS – steigert sich zwar die Leistung beim Lesen von Dateien um den Faktor Zwei. Die Schreibzugriffe erfolgen aber selbst im Idealfall nur gleich schnell wie bei Einzellaufwerken. Damit eignet sich Mirroring vor allem für Systeme, auf denen wichtige Daten zum hauptsächlichen Lesezugriff vorgehalten werden.

RAID Level 0+1

Durch eine Kombination von Mirroring und Striping lassen sich Geschwindigkeitsgewinn und Datensicherheit verbinden: Ein RAID 0 erzielt durch das lineare Zusammenschalten mehrerer Festplatten sowohl beim Lesen als auch beim Schreiben einen Geschwindigkeitsvorteil. Die zusätzliche Spiegelung des Stripesets auf weitere Platten sorgt für Datensicherheit. Je nach Hersteller wird dieses Verfahren als RAID 0+1, RAID 0/1 oder RAID 10 apostrophiert.



(/i/detail/artikel/401665/1/389636/EL_mediaN10289/)

RAID 10 ist eine Kombination aus RAID 1 und 0. Dabei werden wie bei RAID 1 die Festplatten gespiegelt, diese Daten jedoch anschließend bei RAID 0 mittels Striping über die Festplatten verteilt. Die Performance ist insgesamt gesteigert obwohl die Daten gesichert sind wie bei RAID 1. (Quelle www.3ware.com)

Zur Ansteuerung der Festplatten sind zwei Verfahren denkbar. Angenommen, zum Aufbau des RAID-10-Verbunds stehen sechs Festplattenlaufwerke zur Verfügung. Zum einen lassen sich zunächst je drei Laufwerke stripen (RAID 0) und anschließend diese beiden logischen Laufwerke spiegeln (RAID 1). Die umgekehrte Konstellation liefert scheinbar dasselbe Ergebnis: Zuerst werden je zwei Festplatten gespiegelt, dann die drei logischen Laufwerke zu einem Stripeset verbunden. Bei beiden Methoden steht anschließend die Kapazität von drei Festplatten zur Verfügung.

Vom Standpunkt der Datensicherheit ist jedoch das letztgenannte Verfahren dem ersten auf jeden Fall vorzuziehen. Den Grund dafür zeigt ein Blick auf den Ernstfall: Wurde zuerst gestriped, ist nach einem Plattenausfall das betroffene Stripeset komplett unbrauchbar. Die Daten stehen zwar noch auf dem zweiten Stripeset zur Verfügung. Fällt aber eines der drei Laufwerk dieses zweiten Verbunds aus, sind die Daten verloren.

Im umgekehrten Fall verliert ein gespiegeltes logisches Laufwerk zwar durch den Ausfall ebenfalls seine Redundanz. Aber nur wenn dann eine zweite Platte dieser Einheit ausfällt, entsteht ein Datenverlust. Die Wahrscheinlichkeit eines Totalausfalls sinkt somit auf ein Drittel gegenüber der ersten Methode.

RAID-Verfahren mit Fehlerkorrektur

Zwar bietet Mirroring perfekte Redundanz, verursacht jedoch gleichzeitig einen hohen Overhead und entsprechend hohe Kosten. Um diesen Nachteil zu beheben, arbeiten die RAID-Level 2 bis 7 mit Fehlerkorrektur. Sie verteilen zunächst die Nutzdaten per Striping auf wenigstens zwei Datenlaufwerke. Aus deren Dateninhalt wird anschließend ein Korrekturwert errechnet, mit dessen Hilfe sich nach einem Ausfall die Daten des defekten Laufwerks wieder rekonstruieren lassen. Dieser ECC-Code wird auf einem eigenen Parity-Laufwerk abgelegt.

Dabei setzt RAID auf eines der ältesten Verfahren zur Fehlerkorrektur, die Paritätsprüfung. Dazu verknüpft es die Daten der Nutzlaufwerke über eine logische Exklusiv-Oder-Operation (XOR) und speichert das Resultat auf einem eigenen Parity-Laufwerk. Das Ergebnis der Verknüpfung ist dann 1, wenn eine ungerade Anzahl von Bitstellen eine 1 aufweist. Bei einer geraden Anzahl dagegen ist das Ergebnis 0:

Parity-Generierung

Laufwerk	Inhalt
Laufwerk A	11101100
Laufwerk B	10110011
Laufwerk C	01001101
Parity-Laufwerk	00010010

Fällt nun ein beliebiges Laufwerk aus, lassen sich durch ein erneutes XOR die verloren gegangenen Daten problemlos rekonstruieren:

Fehlerkorrektur durch Parity

	vor dem Ausfall	Ausfall eines Datenlaufwerks	Ausfall des Parity-Laufwerks
Laufwerk A	11101100	11101100	11101100
Laufwerk B	10110011	xxxxxxx	10110011
Laufwerk C	01001101	01001101	01001101
Parity-Laufwerk	00010010	00010010	xxxxxxx
Datenrekonstruktion		10110011	00010010

ECC-Overhead bei Schreiboperationen

Zwar reduzieren die paritätsbasierten RAID-Varianten den für die Datensicherheit notwendigen Kapazitäts-Overhead deutlich. Er beträgt maximal ein Drittel, bei Verwendung mehrerer Nutzdatenlaufwerke sinkt er weiter ab. Andererseits erfordert das Update der Parity-Informationen beim Speichern von Daten zusätzliche Schreib- und Lesezugriffe.

Die Aktualisierung der ECC-Informationen kann auf zwei Wegen erfolgen. In der einfacheren Variante schreibt der RAID-Controller (http://www.tecchannel.de/storage/san/432631/praxis_know_how RAID-Controller optimal konfigurieren/) nach dem Eintreffen eines neuen Datenblocks diesen zunächst auf ein Laufwerk. Anschließend liest er die von der Operation betroffenen Blöcke aller Laufwerke ein, errechnet die resultierende Parity und schreibt diese zurück auf das ECC-Laufwerk. Diese Methode erfordert pro Schreiboperation einen zusätzlichen Zugriff auf alle Platten des Verbunds.

In der komplexeren Variante des Updates liest der Controller zunächst lediglich den zu überschreibenden Datenblock ein. Nun berechnet er per XOR, an welchen Stellen sich ein Bit geändert hat. Anschließend liest er den alten ECC-Block und verknüpft ihn erneut per XOR mit dem zuvor gewonnenen Resultat. Als Ergebnis erhält er den neuen Parity-Block und kann diesen zurückspeichern. Anders als bei der ersten Update-Methode müssen hier nur zwei Laufwerke des Arrays angesprochen werden.

RAID Level 2

RAID Level 2 bietet zusätzlichen Schutz gegen Fehler innerhalb der Festplatten, wird aber aufgrund seiner aufwendigen Implementierung kaum eingesetzt. Es blieb bis auf wenige Anwendungen auf den Großrechnerbereich beschränkt.

RAID 2 basiert auf einer bitweisen Aufteilung der Nutzdaten. Dabei implementiert es nicht nur eine Fehlerkorrektur gegen den kompletten Ausfall einer Platte. Alle anderen RAID-Level versagen, wenn nicht eine komplette Platte ausfällt, sondern die Daten beispielsweise durch einen Schreibfehler inkonsistent sind. Ein Fehler wird bei ihnen zwar erkannt. Welche Platte jedoch falsche Daten liefert, ist nicht festzustellen - eine Korrektur somit unmöglich.

RAID 2 verwendet deshalb neben 8 Bit für Daten noch 2 Bit für den ECC-Code (<http://de.wikipedia.org/wiki/Fehlerkorrekturverfahren>). Damit kann neben der Entdeckung des Fehlers auch dessen Position ermittelt werden. RAID 2 entspricht damit dem ECC-RAM, das bei Single-Bit-Fehlern nicht mit einem Parity-Error den Rechner stoppt, sondern das Bit korrigiert.

Die bitweise Verteilung auf eigene Laufwerke erzwingt aber den Einsatz von nicht weniger als 10 Festplatten im Verbund. Zwar steigert sich bei RAID 2 durch die Möglichkeit zum parallelen Zugriff die Lesegeschwindigkeit auf das achtfache. Bei Schreiboperationen sinkt allerdings durch den hohen ECC-Overhead die Performance unter jene eines Einzellaufwerks.

RAID Level 3

RAID Level 3 setzt auf ein Byte-weises Striping der Daten. Zur Erkennung von Schreib-/Lese Fehlern setzt es im Gegensatz zu RAID 2 auf die integrierten Funktionen der Festplatten. Dadurch kommt es mit einem einzelnen, dedizierten Parity-Laufwerk aus.

Um die Generierung der ECC-Daten zu erleichtern synchronisiert RAID 3 die Kopfpositionen der Laufwerke. Das ermöglicht zwar Schreibzugriffe ohne Overhead, da sich Parity- und Nutzdaten parallel auf den Laufwerken speichern lassen. Viele Lesezugriffe auf kleine, verteilte Dateien erfordern dagegen die synchrone Neupositionierung der Köpfe aller Platten im Verbund und kosten entsprechend Zeit.

Nur beim Lesen großer Files kann RAID 3 den Geschwindigkeitsvorteil des parallelen Zugriffs ausspielen. Daher kommt das Verfahren vor allem bei der Verarbeitung großer, zusammenhängender Datenmengen auf Einzelplatzrechnern (http://www.tecchannel.de/pc_mobile/) zum Einsatz: Etwa bei CAD/CAM oder Multimedieverarbeitung.

Das ursprüngliche Berkeley-Papier sah für RAID 3 ein Striping von Bytes beziehungsweise Multibyte-Chunks vor (typischerweise deutlich unter 1024 Byte). Das im Folgenden erläuterte RAID 4 sollte stattdessen die damals typischen Unix-Festplatten-Blocks (4 oder 8 KByte) als Stripe-Größe nutzen. Abweichend von dieser "reinen Lehre" verwenden heutige RAID-3-Implementationen ebenfalls ein blockweises Striping, um insbesondere eine bessere Balance zwischen Schreib- und Lese-Performance zu erzielen.

RAID Level 4

RAID Level 4 arbeitet anders als RAID 3 mit blockweisem Striping der Nutzdaten. Um die Nachteile von RAID 3 bei der Verarbeitung kleiner Files zu umgehen, verzichtet es dabei allerdings auf eine Synchronisierung der Kopfbewegungen aller Laufwerke. Zur Speicherung der Parity-Informationen nutzt es dagegen wie RAID 2 und 3 ein dediziertes Laufwerk.

Die Kombination aus blockweisem Striping und unabhängigem Plattenzugriff ermöglicht RAID 4 das schnelle Lesen auch kleiner Files. Je nach Dateigröße kann hier jedes Datenlaufwerk eine unabhängige Leseoperation vornehmen. Bei Schreiboperationen allerdings erweist sich RAID 4 aufgrund der fehlenden Synchronisierung RAID 3 unterlegen: Bei jedem ECC (<http://de.wikipedia.org/wiki/Fehlerkorrekturverfahren>)-Update muss zunächst die passende Stelle auf dem Parity-Laufwerk gefunden und angesteuert werden. Damit erweist sich hier das ECC-Drive als Flaschenhals.

Aufgrund seiner Arbeitsweise eignet sich RAID 4 vor allem für Umgebungen, in denen erheblich mehr Lese- als Schreiboperationen anfallen. Da es im Vergleich zu RAID 3 nur minimale Vorteile bietet, kommt es nur selten zum Einsatz.

RAID Level 5

RAID Level 5 arbeitet ebenso wie RAID 4 mit einer blockweisen Verteilung der Nutzdaten. Es verzichtet jedoch auf ein dediziertes Parity-Laufwerk und verteilt die ECC-Daten (<http://de.wikipedia.org/wiki/Fehlerkorrekturverfahren>) zusammen mit den Nutzdaten gleichmäßig über die Laufwerke. Damit sinkt die Wahrscheinlichkeit, dass gleichzeitig zwei Schreiboperationen auf dieselbe Platte erfolgen. Schreibzugriffe lassen sich also weitgehend parallelisieren. Zudem verteilt sich die mechanische Belastung der Platten gleichmäßig, da keine eine Sonderstellung als Parity-Laufwerk einnimmt.

Auch beim Lesen von Daten bietet RAID 5 durch die Verteilung der Daten über alle Laufwerke eine gute Performance. Dies zählt sich insbesondere beim Zugriff auf viele kleine Datenblöcke aus. Daher kommt RAID 5 speziell bei Datenbank- oder Transaktionsservern gern zum Einsatz.

Kombi-Verfahren auf RAID-5-Basis

Ähnlich wie RAID 1 lässt sich auch RAID 5 sehr gut mit RAID 0 kombinieren. Das resultierende RAID 0+5 (aka RAID 50) bietet durch Stripen über RAID-5-Sets eine ähnlich gute Performance wie RAID 10, verursacht aber einen geringeren Kapazitäts-Overhead.

Durch den Zusammenschluss von drei oder mehr Mirror-Sets zu einem RAID 5 entsteht ein RAID 1+5 oder RAID 51. Es bietet gegenüber reinem RAID 5 eine nochmals gesteigerte Ausfallsicherheit, verursacht andererseits aber einen hohen Kapazitäts-Overhead.

RAID 5+5 alias RAID 55 entsteht durch Kombination mehrerer RAID-5-Sets zu einem RAID-5-Array. Es bietet die gleiche Ausfallsicherheit wie ein RAID 51, reduziert jedoch den dazu notwendigen Kapazitäts-Overhead etwas. Sein Hauptvorteil liegt allerdings darin, dass sich von den gestripten Stripe-Sets sehr schnell lesen lässt.

RAID 6 und RAID 7

RAID 6 stellt einen Versuch dar, gegenüber RAID 3 bis 5 die Ausfallsicherheit nochmals zu erhöhen. Bei diesen Verfahren darf nur eine Platte des Arrays ausfallen, da sich sonst die Daten nicht mehr per XOR rekonstruieren lassen. RAID 6 umgeht diese Einschränkung, indem es ein RAID 5 um zusätzliche, über die Platten verteilte Parity-Informationen auf Basis von Reed-Solomon-Codes (<http://de.wikipedia.org/wiki/Reed-Solomon-Code>) ergänzt. Man spricht hier auch von einer "zweidimensionalen Parity". Vereinfacht kann man sich das als ein RAID 5 vorstellen, das um eine weitere Parity-Platte ergänzt wurde.

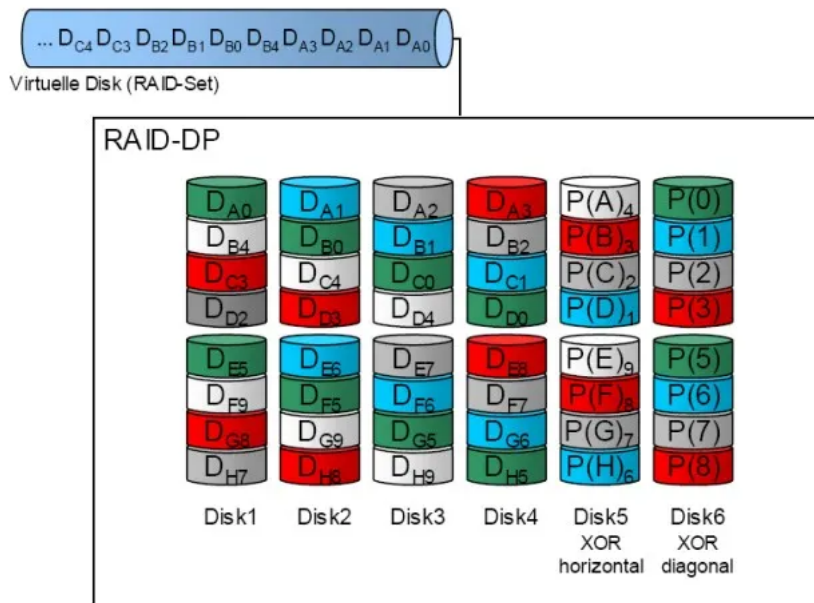
Zwar dürfen bei RAID 6 nun zwei beliebige Platten des Verbunds ausfallen, ohne dass Datenverluste auftreten. Die zusätzliche Sicherheit muss allerdings mit gegenüber RAID 3 bis 5 deutlich langsameren Schreibzugriffen sowie einer mageren Performance bei der Rekonstruktion erkauft werden. Eine Software-Variante von RAID 6 implementiert ab Linux 2.6.2 auch der entsprechende Metadisk-Treiber.

Auch das proprietäre RAID 7 ist ähnlich wie RAID 5 aufgebaut. Allerdings setzt der Hersteller Storage Computer im Controller zusätzlich ein lokales Echtzeitbetriebssystem ein. Schnelle Datenbusse und mehrere große Pufferspeicher koppeln die Laufwerke vom Bus ab. Dieses asynchrone Verfahren soll Lese- wie Schreiboperationen gegenüber anderen RAID-Varianten erheblich beschleunigen. Zudem lässt sich, ähnlich wie bei RAID 6, die Paritätsinformation auch auf mehrere Laufwerke speichern.

RAID DP

Der RAID-DP-Mode nutzt eine Kombination aus RAID 4 und einer Double-Parity-Prüfung. Diese Implementierung schützt wirkungsvoll das Storage-System auch wenn zwei Platten gleichzeitig ausfallen. Entwickelt wurde das proprietäre RAID DP von dem Storage-Spezialisten NetApp (<http://www.netapp.com/de/>). Das Verfahren hat das Unternehmen in das WAFL-File-System (Write Anywhere File Layout) integriert. Im Vergleich zu herkömmlichen RAID-Modi soll dieses RAID-Verfahren durch die doppelte Parity-Bildung besonders zuverlässig arbeiten und dabei eine hohe Performance bieten. Darüber hinaus lässt sich RAID-DP auf herkömmlicher RAID-Hardware mit XOR-Engine sehr einfach implementieren.

Ähnlich wie bei RAID 4 oder RAID 5 bildet RAID DP seine Prüfsummen durch XOR-Verknüpfungen. Dabei nutzt das Verfahren im ersten Schritt die einzelnen horizontalen Datenpakete. Im zweiten Schritt verwendet RAID DP dann die diagonal liegenden Datenpakete, um per XOR-Algorithmus eine Prüfsumme zu berechnen. Diese Methode kann auch als eine Erweiterung des RAID 4 bezeichnet werden, da der Anwender einfach durch das Anschließen einer zusätzlichen Prüfsummenplatte ans RAID 4 eine erhöhte Sicherheit des RAID DP erreichen kann – vorausgesetzt die Anwendung unterstützt das Verfahren. Die Daten des RAID-Sets bleiben dabei erhalten. Auch der umgekehrte Weg zurück zum RAID 4 ist durch die Entkopplung eines Prüfsummenlaufwerkes jederzeit möglich.



(/i/detail/artikel/401665/1/383884/EL_mediaN105EA/)

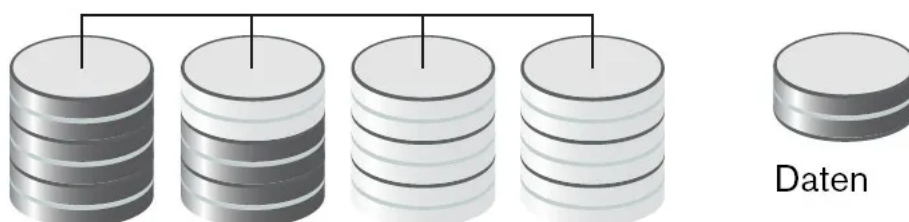
Ausfallsicher: Ein Storage-System nach dem RAID-DP-Verfahren arbeitet auch dann, wenn zwei Festplatten ausfallen. (Quelle: transtec)

Die Unabhängigkeit der beiden Prüfsummen und somit die Datensicherheit ist nur gegeben, wenn eine bestimmte Anzahl von Festplatten zum Einsatz kommt. So müssen bei RAID DP die Datenplatten inklusive der horizontalen Prüfsummenplatte eine Primzahl ergeben. Will der Anwender eine beliebige Anzahl von Festplatten verwenden, wird der Algorithmus um virtuelle Datenträger, die keine Daten enthalten, bis auf das Primzahlenkriterium erweitert. Diese virtuellen Festplatten dienen dann nur als Platzhalter für das Rechenverfahren.

NRAID

Der Begriff NRAID bedeutet No RAID. Es bezeichnet die Konfiguration eines Storage-Systems, bei dem mehrere Festplatten zu einem einzigen Laufwerk "aneinandergereiht" werden. Dabei spielt die Größe der Laufwerke keine Rolle. Diese Konfiguration wird im Allgemeinen auch als SPAN bezeichnet und stellt im Grunde genommen kein RAID-System dar. Denn damit lassen sich keine redundante Festplattenverbunde erstellen. NRAID verwenden Anwender häufig in der Praxis, um schnell und unkompliziert aus verschiedenen großen Festplatten einen großen Speicher aufzubauen, ohne dass spezielle Sicherheitsaspekte berücksichtigt werden.

NRAID / SPAN



(/i/detail/artikel/401665/1/390016/EL_mediaN10607/)

Flexibler Datenspeicher: Das logische NRAID-Laufwerk wächst mit der Anzahl der angeschlossenen Festplatten ohne Datensicherheit. (Quelle: Western Digital)

Der Vorteil von NRAID liegt in der einfachen Handhabung. So entspricht die Größe des logischen NRAID-Laufwerks der Summe der Nettokapazitäten der einzelnen angeschlossenen Festplatten (<http://www.tecchannel.de/produkte/storage/festplatten/>). Der Nachteil dieses Storage-Systems liegt in der Datensicherheit. Fällt eine Festplatte des logischen Laufwerks aus, so verliert das ganze System seine Datenintegrität und alle Daten sind verloren. Unter Umständen können nur komplexe Rettungsmaßnahmen mit speziellen Programmen die Daten wieder herstellen.

Zu beachten ist auch, dass die Ausfallwahrscheinlichkeit der "NRAID-Festplatte" mit der Anzahl der angeschlossenen HDDs exponential steigt. Auch in Bezug auf Geschwindigkeit kann das NRAID nicht punkten, da die Daten erst so lange auf eine Festplatte geschrieben werden, bis diese voll ist und dann erst wird auf die nächste HDD gewechselt.

RAIDⁿ

Das von RAID 6 verwendete Verfahren zur Generierung mehrfacher, unabhängiger Paritätsinformationen lässt sich grundsätzlich auch dazu nutzen, den Ausfall von nahezu beliebig vielen Platten aufzufangen. Eine ausführliche Abhandlung zu den theoretischen Grundlagen finden Sie bei Interesse in einem Papier (<http://www.cs.utk.edu/~plank/plank/papers/CS-96-332.pdf>) der University of Tennessee.

Derart aufgebaute Arrays bezeichnet man üblicherweise als RAID $n+m$ oder RAID (n,m) . Dabei steht n für die Gesamtzahl der Platten im Verbund und m bezeichnet die Zahl der Platten, die ohne Datenverlust ausfallen dürfen. Die Nutzkapazität des Arrays entspricht $n-m$ Disks. Vorteil einer solchen Konfiguration: Die Ausfallsicherheit lässt sich gegenüber RAID 5 weiter erhöhen, ohne die hohe Redundanz und damit die Kosten von Kombiverfahren wie RAID 51 in Kauf nehmen zu müssen.

Eine industrielle Implementation einer solchen Technik lieferte Tandberg Data (<http://www.tandbergdata.com/>) in seinen Netzwerk-Speichersubsystemen der ValueNAS-Serie aus. Hier kommen für die Generierung der zusätzlichen Paritätsinformationen allerdings keine Reed-Solomon-Codes zum Einsatz, sondern von der Tandberg-Data-Tochter Inostor patentierte, eigene Algorithmen. Die resultierende Technologie bezeichnet Tandberg Data als RAIDⁿ.

RAIDⁿ ist in Form einer in C geschriebenen Software-Library implementiert, die derzeit als (über insmod ladbares) Kernel-Modul für Linux vorliegt. Laut Tandberg/Inostor kann sie aber bei Bedarf problemlos auch auf andere Betriebssysteme portiert werden, wie etwa Sun Solaris oder auch Microsoft Windows.

Implementation

Das von Inostor entwickelte Verfahren verwendet für jede n,m -Kombination einen eigenen Teilalgorithmus, wobei in der momentanen Implementation die Gesamtzahl der Platten n zwischen 4 und 32 liegen darf und die Anzahl der redundanten Disks m 2 oder 3 beträgt. Getestet hat Inostor das Verfahren darüber hinaus jedoch auch für alle Kombinationen ($2 \leq m \leq n-2$) für n kleiner gleich 21. Bei entsprechender Nachfrage von Kunden will man die Implementation entsprechend erweitern. Neben diesen RAIDⁿ-Leveln unterstützt das Inostor-Modul auch die klassischen RAID-Varianten 0, 1, 10, 4 und 5.

Wie die untenstehende Tabelle zeigt, realisiert RAIDⁿ gegenüber den herkömmlichen Verfahren bei gleicher Plattenanzahl stets wahlweise eine höhere Kapazität bei gleicher Ausfallsicherheit oder eine höhere Ausfallsicherheit bei gleicher Kapazität. Darf man Inostors (Tandbergs) Angaben hinsichtlich der Performance Glauben schenken, dann liegt auch diese stets zumindest gleich oder je nach Kombination sogar höher als bei den RAID-1- beziehungsweise RAID-5-Varianten.

RAIDⁿ im Vergleich

	RAID 1	RAID 51	RAID ⁿ (10,3)	RAID (10,6) ⁿ	RAID 55	RAID (9,3) ⁿ	RAID (9,5) ⁿ
Anzahl Disks	10	10	10	10	9	9	9

Alle Angaben laut Inostor. (*) Geschwindigkeit als Faktor gegenüber einzelner Platte. (**) Lesegeschwindigkeit inklusive Parity.

	RAID 1	RAID 51	RAID ⁿ (10,3)	RAID (10,6) ⁿ	RAID 55	RAID (9,3) ⁿ	RAID (9,5) ⁿ
Nutzkapazität (Disks)	5	4	7	4	4	6	4
Redundanz (Disks, Worst Case)	1	3	3	6	3	3	5
Redundanz (Disks, Best Case)	5	6	3	6	5	3	5
max. Geschwindigkeit, Lesen (*)(**)	5	5	10	10	9	9	9
max. Geschwindigkeit, Schreiben (*)	5	4	7	4	4	6	4
Alle Angaben laut Inostor. (*) Geschwindigkeit als Faktor gegenüber einzelner Platte. (**) Lesegeschwindigkeit inklusive Parity.							

RAID-Varianten im Überblick

Die Auswahl eines geeigneten RAID-Levels erfordert eine genaue Abwägung zwischen den Faktoren Verfügbarkeit, Performance und Kosten pro MByte. Eine Organisation der Platten als JBOD verursacht die geringsten Kosten, lässt jedoch in Sachen Ausfallsicherheit und Geschwindigkeit zu wünschen übrig. Ein Array des Levels 1 garantiert dagegen höchste Verfügbarkeit. Es produziert jedoch den höchsten Kapazitäts-Overhead - und somit auch die höchsten relativen Kosten. Fasst man diese Faktoren in einem Diagramm zusammen, ergibt sich das typische RAID-Dreieck wie in der unten stehenden Abbildung.

Die Abhängigkeiten zwischen RAID-Level, Performance und Ausfallsicherheit fasst die folgende Tabelle noch einmal zusammen. Wie sich deutlich erkennen lässt, bringt jedes der RAID-Verfahren dabei spezifische Vor- und Nachteile auf die Waagschale.

RAID-Level im Vergleich

	RAID 0	RAID 1	RAID 10	RAID 2	RAID 3	RAID 4	RAID 5	RAID 6
Anzahl Laufwerke	$n > 1$	$n = 2$	$n > 3$	$N = 10$	$n > 2$	$n > 2$	$n > 2$	$n > 3$
Redundante Laufwerke	0	1	1(**)	2	1	1	1	2
Kapazitäts-overhead (Prozent)	0	50	50	20	$100 / n$	$100 / n$	$100 / n$	$200 / n$
Parallele Lese-operationen	n	2	$n / 2$	8	$n - 1$	$n - 1$	$n - 1$	$n - 2$
Parallele Schreib-operationen	n	1	1	1	1	1	$n / 2$	$n / 3$

(*) Als Faktor gegenüber einem Einzellaufwerk. (**) Worst-case-Angabe. Im günstigsten Fall können $n/2$ Laufwerke ohne Datenverlust ausfallen.

	RAID 0	RAID 1	RAID 10	RAID 2	RAID 3	RAID 4	RAID 5	RAID 6
Maximaler Lese-durchsatz (*)	n	2	$n / 2$	8	$n - 1$	$n - 1$	$n - 1$	$n - 2$
Maximaler Schreib-durchsatz (*)	n	1	1	1	1	1	$n / 2$	$n / 3$
(*) Als Faktor gegenüber einem Einzelaufwerk. (**) Worst-case-Angabe. Im günstigsten Fall können $n/2$ Laufwerke ohne Datenverlust ausfallen.								

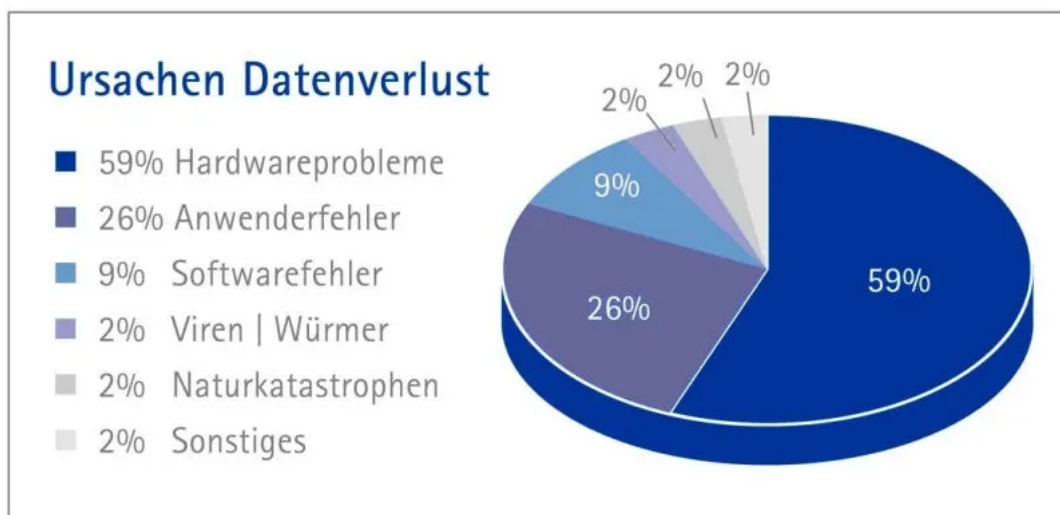
Verhalten bei Plattenausfall

Ob rechnerinterner Plattenverbund oder externes Speichersubsystem, ob Hardware- oder Software-RAID: Fällt eine Platte des Arrays aus, geht bei den gängigen RAID-Leveln die Redundanz verloren. Jede Fehlfunktion eines weiteren Laufwerks führt in dieser Situation unweigerlich zu Datenverlusten. Konsequenz: Die defekte Platte muss schnellstmöglich ersetzt und das Array rekonstruiert werden.

Im Optimalfall verfügt das Array über eine zusätzliche, ausschließlich im Notfall verwendete Festplatte. Ein solches Hot-Fix-Laufwerk (auch Hot-Spare oder Stand-by-Laufwerk genannt) wird automatisch aktiviert und als Ersatz für die defekte Platte eingebunden. Steht kein Hot-Spare zur Verfügung, gilt es, das defekte Laufwerk manuell zu wechseln. Dies erfordert für gewöhnlich ein Abschalten des Rechners und damit eine Betriebsunterbrechung - was gerade bei Servern in der Regel inakzeptabel ist. Abhilfe schaffen hier hot-plug- respektive hot-swap-fähige Arrays: Die Festplatten des RAID-Verbundes sind frei zugänglich in Festplatten-Shuttles untergebracht, die sich während des Betriebs wechseln lassen.

Nach erfolgreichem Austausch gilt es, die neue Festplatte in den RAID-Verband einzubinden und die verlorengegangenen Daten (<http://www.tecchannel.de/storage/backup/>) zu rekonstruieren. Geschieht dies automatisch, spricht man von Auto-Rebuild. Dieses automatische Wiederherstellen setzt allerdings die Fähigkeit des Controllers voraus, mit dem Festplatten-Shuttle zu kommunizieren (Laufwerkszustand, neues Shuttle eingeführt). Können sich Controller und Shuttle nicht verständigen, muss der Rebuild manuell angestoßen werden.

Fazit: Redundanz mit Grenzen



(/i/detail/artikel/401665/1/365141/EL_mediaN10Dg7/)

Anfällige Hardware: Mit 59 Prozent ist die Hardware der Hauptschuldige am Datenverlust. Hier können RAID Systeme helfen.(Quelle: Kroll Ontrack)

Zwar lässt sich durch den Einsatz von Diskarrays die Verfügbarkeit von Rechnersystemen deutlich erhöhen. Ein Allheilmittel gegen Datenverluste stellt RAID allerdings nicht dar. Um eine Ausfallsicherheit nahe 100 Prozent zu erreichen, müssen alle Komponenten des Speichersubsystems inklusive Controller, Netzteil und Lüftern redundant ausgelegt werden. Solche

Lösungen bietet die Storage-Industrie zwar an, preiswert fallen sie aber nicht gerade aus.

Zudem ereignen sich Ausfälle von Laufwerken und anderen Komponenten nicht immer unabhängig voneinander. In der Praxis treten gelegentlich Situationen ein, durch die sich die Ausfallwahrscheinlichkeit des gesamten Arrays schlagartig erhöht. Dazu zählen etwa durch Blitzschlag verursachte Überspannungen, Überschwemmungen oder Brände. Auch Viren und Würmer befallen RAID-Systeme ebenso gern wie Einzellaufwerke.

Schließlich kann selbst das zuverlässigste Array den Risikofaktor Nummer 1 nicht ausschalten - den Menschen. Den weitaus größten Teil irreparabler Datenverluste verursacht nicht etwa versagende Technik, sondern Fehlbedienung durch den Benutzer. Gelöschte Dateien sind auch auf RAID-Systemen verloren. Selbst für das ausgefeilteste RAID-System gilt deshalb: Den einzig wirklich zuverlässigen Schutz gegen Datenverluste bietet ein konsequent geplantes und vorgenommenes Backup (<http://www.tecchannel.de/storage/backup/>).

Sollte auch kein Backup vorhanden sein, lassen sich immer noch professionelle Datenrettungslabore (http://www.tecchannel.de/storage/backup/401608/datenrettung_professionelle_hilfe_statt_datenverlust/index.html) konsultieren. Selbst bei komplexen RAIDs können die Labore die Daten mit Erfolgsquoten von 70 bis 80 Prozent rekonstruieren. (jlu/cvi/hal)

 Artikel als PDF kaufen

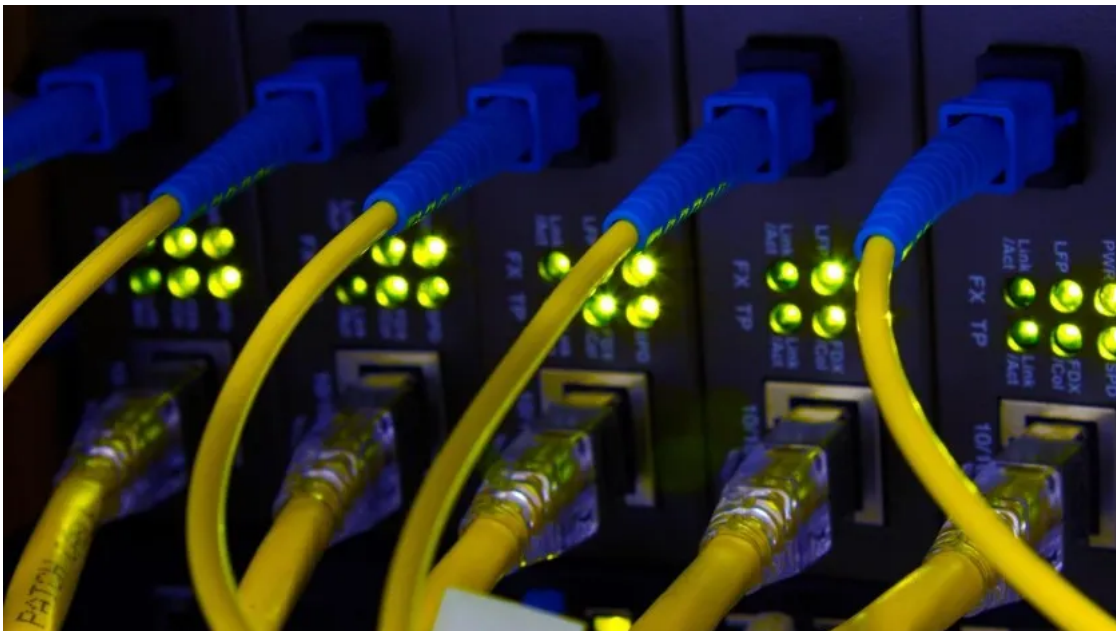
Workshop Newsletter

E-Mail-Adresse eingeben...

Bestellen



Das könnte Sie auch interessieren



(/a/ratgeber-was-ist-

was-bei-den-netzwerkprotokollen,2038150)

SMTP, SFTP SPX, DHCP, IP oder UDP: Ratgeber: Was... (/a/ratgeber-was-ist-was-bei-den-netzwerkprotokollen,2038150)