



(10) **DE 10 2012 103 655 B4** 2018.08.30

(12)

Patentschrift

(21) Aktenzeichen: **10 2012 103 655.1**
(22) Anmeldetag: **26.04.2012**
(43) Offenlegungstag: **29.11.2012**
(45) Veröffentlichungstag
der Patenterteilung: **30.08.2018**

(51) Int Cl.: **G11B 20/12 (2006.01)**
G06F 11/20 (2006.01)
G06F 13/12 (2006.01)
G06F 12/02 (2006.01)

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(30) Unionspriorität:
13/113,546 **23.05.2011** **US**

(73) Patentinhaber:
**International Business Machines Corporation,
Armonk, N.Y., US**

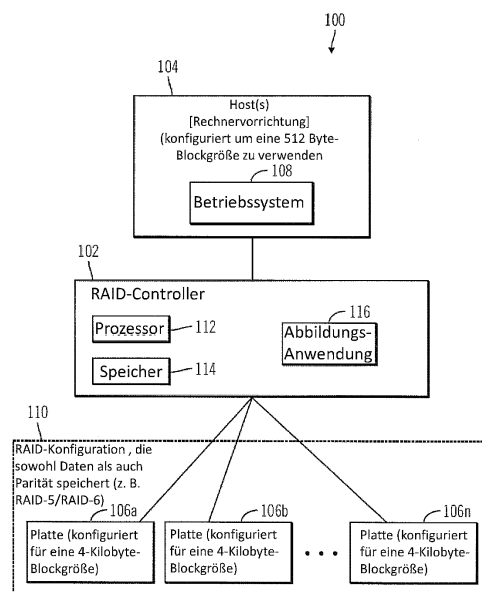
(74) Vertreter:
**Richardt Patentanwälte PartG mbB, 65185
Wiesbaden, DE**

(72) Erfinder:
**Benhase, Michael T., Tucson, Ariz., US; Judd, Ian
D., Winchester, Hampshire, GB; Moertl, Daniel F.,
Rochester, Minn., US; Nielsen, Karl A., Tucson,
Ariz., US**

(56) Ermittelter Stand der Technik:
US **7 080 200** **B2**
US **2011 / 0 119 464** **A1**

(54) Bezeichnung: **Schreiben neuer Daten einer ersten Blockgröße in einen RAID-Array, der sowohl die Parität als auch die Daten in einer zweiten Blockgröße speichert**

(57) Hauptanspruch: Verfahren implementiert in einer Vorrichtung, wobei das Verfahren umfasst:
Empfangen neuer Daten, durch einen „Redundante Anordnung unabhängiger Platten“ (RAID)-Controller (102), die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden;
Lesen alter Daten (606) und alter Parität (608) durch den RAID-Controller (102), die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer ist als die erste Blockgröße;
Berechnen einer neuen Parität (610) durch den RAID-Controller (102), basierend auf den neuen Daten, den alten Daten, und der alten Parität; und
Schreiben der neuen Daten und der neuen Parität durch den RAID-Controller (102), ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller (102) überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden, wobei der RAID-Controller (102) Platten steuert, die als RAID-5 konfiguriert sind, wobei:
das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst (606) und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst (608); und
das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, ...



Beschreibung

Hintergrund

Technisches Feld

[0001] Die Offenbarung betrifft ein Verfahren, System und Herstellungsgegenstand für das Schreiben neuer Daten von einer ersten Blockgröße in einen „Redundante Anordnung unabhängiger Platten“ (RAID)-Array, der sowohl die Parität als auch die Daten in einer zweiten Blockgröße speichert.

Hintergrund

[0002] Redundante Anordnung unabhängiger Platten (RAID) ist eine Kategorie von Plattenlaufwerken, die zwei oder mehr Plattenlaufwerke kombiniert, um Fehlertoleranz für Daten bereitzustellen. RAID ermöglicht die redundante Speicherung der gleichen Daten über eine Vielzahl von Platten. In bestimmten Konfigurationen von RAID, wie zum Beispiel RAID-5 oder RAID-6, kann auch zur Wiederherstellung bei einem Ausfall von einer oder mehreren Platten Paritäts-Information gespeichert sein. Die physischen Platten eines RAID, so könnte man sagen, befinden sich in einem RAID-Array, der von einem Betriebssystem wie eine einzelne Platte angesprochen werden kann. Die Daten können in Stripes in einem RAID-Array geschrieben werden, wobei Daten-Striping die Ausbreitung der Blöcke jeder Datei über eine Vielzahl von Plattenlaufwerken ist.

[0003] Ein Sektor ist eine speziell bemessene Aufteilung einer Platte. Bisher wurde ein Sektor einer Platte allgemein konfiguriert, um 512 Byte an Information zu enthalten. Allerdings wurden unlängst durch Plattenhersteller bestimmte Platten konfiguriert, um 4096 Byte (d. h. 4 Kilobyte) an Information zu enthalten.

[0004] Ein Block ist eine Gruppe von Sektoren einer Platte, die ein Betriebssystem anwählen kann. Schlüsseldaten-Zählung (Count-key-data (CKD)) ist ein Platten-Datenorganisations-Modell bestimmter Betriebssysteme, bei denen die Platte voraussetzt, dass sie eine feste Anzahl von Spuren umfasst, von denen jede eine maximale Datenkapazität enthält. Mehrere Datensätze von unterschiedlicher Länge können auf jede Spur einer CKD Platte geschrieben werden, und die nutzbare Kapazität jeder Spur hängt von der Anzahl der Datensätze, die auf die Spur geschrieben sind, ab. Die CKD-Architektur leitet seinen Namen von dem Datensatz-Format ab, das ein Feld umfasst, welches die Anzahl der Byte der Daten und eine Datensatz-Adresse, ein optionales Schlüsselfeld, und die Daten selbst umfasst. CKD-Datensätze werden in 512-Byte-Blöcken gespeichert.

[0005] Die Druckschrift US 7,080,200 B2 offenbart ein System und ein Verfahren zur Durchführung von

Schreiboperationen mit kleinen Blockgrößen (z.B. 512 bytes) in Festplatten, die größere Blöcke (z.B. 4 KB) nutzen.

[0006] Ferner offenbart die Druckschrift US 2011/0119464 A1 ein Datenspeichersystem, bei dem eine Transformation von Daten einer ersten Blockgröße in eine zweite Blockgröße vorgenommen wird.

Kurze Zusammenfassung der bevorzugten Ausführungsformen

[0007] Bereitgestellt werden ein Verfahren, ein System und ein Computerprogrammprodukt, in dem ein „Redundante Anordnung unabhängiger Platten“ (RAID)-Controller neue Daten empfängt, die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Block-Größe angegeben werden. Der RAID-Controller liest alte Daten und alte Parität, die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer in der Größe ist als die erste Blockgröße. Der RAID-Controller berechnet eine neue Parität basierend auf den neuen Daten, den alten Daten, und der alten Parität. Der RAID-Controller schreibt die neuen Daten und die neue Parität ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden.

[0008] In weiteren Ausführungsformen ist der RAID-Controller in Hardware implementiert, und der RAID-Controller steuert Platten, die als RAID-5 konfiguriert sind. Das Lesen wird über zwei Sätze von Leseoperationen von den Platten durchgeführt, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst. Das Schreiben wird über zwei Sätze von Schreiboperationen auf den Platten durchgeführt, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst.

[0009] In noch weiteren Ausführungsformen steuert der RAID-Controller Platten, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst. Das Lesen wird über drei Sätze von Leseoperationen von den Platten durchgeführt, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen

das Lesen des zweiten Satzes der alten Parität umfasst. Das Schreiben wird über drei Sätze von Schreiboperationen auf den Platten durchgeführt, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

[0010] In weiteren Ausführungsformen beträgt die erste Blockgröße 512 Byte und die zweite Blockgröße beträgt 4 Kilobyte (auch als 4K anstelle von 4 Kilobyte bezeichnet). Die neuen Daten werden durch den RAID-Controller von einem Host empfangen, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, wobei die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Datensätze speichern, wobei jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

[0011] In zusätzlichen Ausführungsformen umfassen die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße. Die erste Vielzahl von Blöcken der ersten Blockgröße ist nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert. Jede der ersten Vielzahl von Blöcken ist vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert.

[0012] In einem weiteren Aspekt betrifft die Erfindung ein System, umfassend: einen Speicher; und einen Prozessor mit dem Speicher gekoppelt, wobei der Prozessor durchführt: Empfangen neuer Daten, die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden; Lesen alter Daten und alter Parität, die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer in der Größe ist als die erste Blockgröße; Berechnen einer neuen Parität basierend auf den neuen Daten, den alten Daten, und der alten Parität; und Schreiben der neuen Daten und der neuen Parität ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden.

[0013] Nach einer Ausführungsform der Erfindung ist das System ein RAID-Controller, der Platten steuert, die als RAID-5 konfiguriert sind, wobei das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst; und wobei das Schreiben über zwei Sätze von Schreiboperationen von auf den

Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst.

[0014] Nach einer Ausführungsform der Erfindung steuert der RAID-Controller Platten, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst, und wobei: das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

[0015] Nach einer Ausführungsform der Erfindung ist das System ein RAID-Controller wobei: die erste Blockgröße 512 Byte beträgt; die zweite Blockgröße 4 Kilobyte beträgt; die neuen Daten durch den RAID-Controller von einem Host empfangen werden, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, und die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Datensätze speichern und jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

[0016] Nach einer Ausführungsform der Erfindung umfassen die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße; die erste Vielzahl von Blöcken der ersten Blockgröße ist nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert; jede der ersten Vielzahl von Blöcken ist vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert.

[0017] In einem weiteren Aspekt betrifft die Erfindung Speicher-Controller umfassend: einen Speicher; einen Prozessor mit dem Speicher gekoppelt; und einen RAID-Controller mit dem Speicher gekoppelt, wobei der RAID-Controller durchführt: Empfangen neuer zu schreibender Daten, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben

werden; Lesen alter Daten und alter Parität, die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer in der Größe ist als die erste Blockgröße; Berechnen einer neuen Parität basierend auf den neuen Daten, den alten Daten, und der alten Parität; und Schreiben der neuen Daten und der neuen Parität ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden.

[0018] Nach einer Ausführungsform der Erfindung ist der Speicher ein RAID-Controller, der in Hardware implementiert ist und der Platten steuert, die als RAID-5 konfiguriert sind, wobei das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst; und wobei das Schreiben über zwei Sätze von Schreiboperationen von auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst.

[0019] Nach einer Ausführungsform der Erfindung steuert der RAID-Controller Platten, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst, und wobei: das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

[0020] Nach einer Ausführungsform der Erfindung beträgt die erste Blockgröße 512 Byte; die zweite Blockgröße beträgt 4 Kilobyte; die neuen Daten werden durch den RAID-Controller von einem Host empfangen, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, und die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Da-

tensätze speichern und jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

[0021] Nach einer Ausführungsform der Erfindung umfassen die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße; die erste Vielzahl von Blöcken der ersten Blockgröße ist nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert; jede der ersten Vielzahl von Blöcken ist vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert.

[0022] In einem weiteren Aspekt betrifft die Erfindung ein Computerprogrammprodukt, wobei das Computerprogrammprodukt umfasst: ein computerlesbares Speichermedium mit darin enthaltenem computerlesbarem Programmcode, wobei der computerlesbare Programmcode konfiguriert ist, um Operationen durchzuführen, wobei die Operationen umfassen: Empfangen neuer Daten, durch einen „Redundante Anordnung unabhängiger Platten“ (RAID)-Controller, die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden; Lesen alter Daten und alter Parität durch den RAID-Controller, die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer in der Größe ist als die erste Blockgröße; Berechnen einer neuen Parität durch den RAID-Controller basierend auf den neuen Daten, den alten Daten, und der alten Parität; und Schreiben der neuen Daten und der neuen Parität durch den RAID-Controller ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden.

[0023] Nach einer Ausführungsform der Erfindung ist der RAID-Controller in Hardware implementiert ist, und wobei der RAID-Controller Platten steuert, die als RAID-5 konfiguriert sind, wobei: das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst; und das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst.

[0024] Nach einer Ausführungsform der Erfindung steuert der RAID-Controller Platten, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz

der neuen Parität umfasst, und wobei: das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

[0025] Nach einer Ausführungsform der Erfindung beträgt die erste Blockgröße 512 Byte; die zweite Blockgröße beträgt 4 Kilobyte; die neuen Daten werden durch den RAID-Controller von einem Host empfangen, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, und die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Datensätze speichern und jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

[0026] Nach einer Ausführungsform der Erfindung umfassen die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße; die erste Vielzahl von Blöcken der ersten Blockgröße ist nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert; jede der ersten Vielzahl von Blöcken ist vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert.

[0027] In einem weiteren Aspekt betrifft die Erfindung ein Verfahren für die Nutzung einer EDV-Infrastruktur, umfassend einer Integration eines computerlesbaren Codes in ein Rechensystem, wobei der Code in Kombination mit dem Rechensystem Operationen durchführt, wobei die Operationen umfassen: Empfangen neuer Daten, durch einen „Redundante Anordnung unabhängiger Platten“ (RAID)-Controller, die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden; Lesen alter Daten und alter Parität durch den RAID-Controller, die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer in der Größe ist als die erste Blockgröße; Berechnen einer neuen Parität durch den RAID-Controller basierend auf den neuen Daten, den alten Daten, und der alten Parität; und Schreiben der neuen Daten und der neuen Parität durch den RAID-Controller ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in

die Blöcke der zweiten Blockgröße geschrieben werden.

[0028] Nach einer Ausführungsform der Erfindung ist der RAID-Controller in Hardware implementiert ist, und wobei der RAID-Controller Platten steuert, die als RAID-5 konfiguriert sind, wobei: das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst; und das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst.

[0029] Nach einer Ausführungsform der Erfindung steuert der RAID-Controller Platten, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst, und wobei: das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

[0030] Nach einer Ausführungsform der Erfindung beträgt die erste Blockgröße 512 Byte; die zweite Blockgröße beträgt 4 Kilobyte; die neuen Daten werden durch den RAID-Controller von einem Host empfangen, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, und die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Datensätze speichern und jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

[0031] Nach einer Ausführungsform der Erfindung umfassen die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße; die erste Vielzahl von Blöcken der ersten Blockgröße ist nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert; jede der

ersten Vielzahl von Blöcken ist vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert.

Figurenliste

[0032] Im Folgenden sei auf die Zeichnungen Bezug genommen, in denen gleiche Bezugszeichen einander entsprechende Elemente kennzeichnen, wobei:

Fig. 1 ein Blockdiagramm einer Rechenumgebung in Übereinstimmung mit bestimmten Ausführungsformen darstellt, die einen RAID-Controller gekoppelt an einen Host und eine Vielzahl von Platten umfasst;

Fig. 2 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das zeigt, wie ein einzelner 512-Byte-Block in 4 Kilobyte Blöcken angeordnet ist;

Fig. 3 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das zeigt, wie zwei 512-Byte-Blöcke in 4 Kilobyte Blöcken angeordnet sind;

Fig. 4 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das zeigt, wie neun 512-Byte-Blöcke in 4 Kilobyte Blöcken angeordnet sind;

Fig. 5 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das zeigt, wie zehn 512-Byte-Blöcke in 4 Kilobyte Blöcken angeordnet sind;

Fig. 6 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das bestimmte Platten-I / O-Operationen zeigt, die durch den RAID-Controller durchgeführt werden, der eine RAID-5-Konfiguration steuert;

Fig. 7 ein Flussdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das bestimmte Operationen zeigt;

Fig. 8 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das RAID-Controller zeigt, die außerhalb der Speicher-Controller implementiert sind;

Fig. 9 ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das RAID-Controller zeigt, die innerhalb der Speicher-Controller implementiert sind; und

Fig. 10 ein Blockdiagramm eines Rechensystems in Übereinstimmung mit bestimmten Ausführungsformen darstellt, das bestimmte Elemente zeigt, die Speicher-Controller oder RAID-Controller der **Fig. 1**, **Fig. 8** und **Fig. 9** umfassen können.

Detaillierte Beschreibung der Erfindung

[0033] In der folgenden Beschreibung wird auf die beigefügten Zeichnungen Bezug genommen, die einen Teil hiervon bilden und die mehrere Ausführungsformen veranschaulichen. Es versteht sich, dass andere Ausführungsformen verwendet werden können und strukturelle und funktionelle Änderungen vorgenommen werden können.

Abbilden von Daten von 512-Byte-Blockgrößen in 4 Kilobyte-Blöcke

[0034] Physikalische Platten wurden bisher durch Plattenhersteller in 512-Byte-Sektoren konfiguriert. Unlängst sind physikalische Platten, die in 4 Kilobyte-Sektoren konfiguriert sind, verfügbar geworden. Allerdings beruhen bestimmte Platten-Datenorganisations-Modelle wie CKD auf Betriebssystem-Zugriffen über 512-Byte-Blöcke, auch wenn physikalische Platten, die in 4 Kilobyte-Blöcken konfiguriert sind, verwendet werden.

[0035] Bestimmte Ausführungsformen stellen einen RAID-Controller bereit, der Betriebssystem-Anfragen abbildet, um neue 512-Byte-Blöcke in eine RAID-Konfiguration zu schreiben, welche alte Daten und alte Parität in 4 Kilobyte-Blöcken speichert. Eine Vielzahl von neuen 512-Byte-Blöcken kann benachbart zueinander geschrieben werden und die Blockgrenzen der neuen 512-Byte-Blöcke werden mit den Blockgrenzen von einem oder mehreren 4 Kilobyte-Blöcken ausgerichtet. Der RAID-Controller bewahrt die alten Informationen, die nicht durch das Schreiben der neuen 512-Byte-Blöcke in die 4 Kilobyte-Blöcke überschrieben werden, durch Umschreiben der alten Informationen. Extra-Operationen, die dem Betriebssystem anzeigen, dass die RAID-Konfiguration 512-Byte-Blöcke nicht nativ unterstützt, müssen nicht vom RAID-Controller durchgeführt werden.

Beispielhafte Ausführungsformen

[0036] **Fig. 1** stellt ein Blockdiagramm einer Rechenumgebung **100** in Übereinstimmung mit bestimmten Ausführungsformen dar, die einen RAID-Controller **102** gekoppelt an einen oder mehrere Hosts **104** und eine Vielzahl von Platten **106a**, **106b**, ..., **106n** umfasst.

[0037] Die RAID-Controller **102** und die Hosts **104** können jede geeignete Rechneinrichtung einschließlich derer, die derzeit in der Technik bekannt sind, umfassen, wie einen Personal Computer, eine Workstation, einen Server, ein Mainframe, einen tragbaren Computer, einen Palmtop-Computer, eine Telefonie-Vorrichtung, ein Netzwerkgerät, einen Blade Computer, einen Server, etc. Die Vielzahl von Platten **106a** ... **106n** kann beliebige geeignete physikalische Festplatten, die in der Technik bekannt sind,

umfassen. In alternativen Ausführungsformen kann die Vielzahl von Platten **106a** ... **106n** andere Typen von Platten, wie Solid State Platten, optische Platten, usw. umfassen.

[0038] Der Host **104** kann ein Betriebssystem **108** umfassen, das konfiguriert ist, um eine 512-Byte-Blockgröße für den Zugriff auf Blöcke zu verwenden. Die Vielzahl von Platten **106a** ... **106n** kann in jeder RAID-Konfiguration **110**, die sowohl Parität als auch Daten speichert, konfiguriert sein, wie z. B. RAID-5 oder RAID-6, wobei RAID-5 Block-Level-Striping mit Paritäts-Daten, die über alle Platten verteilt sind, verwendet und RAID-6 Block-Level-Striping mit zwei Paritäts-Blöcken, die über alle Platten verteilt sind, verwendet. Jede der Vielzahl von Platten **106a** ... **106n** ist für eine 4-Kilobyte-Blockgröße konfiguriert. Daher werden, in den in **Fig. 1** gezeigten Ausführungsformen, die Vielzahl von Platten **106a** ... **106n** der RAID-Konfiguration **110** für eine 4-Kilobyte-Blockgröße konfiguriert, während der Host **104** konfiguriert ist, um eine 512-Byte-Blockgröße zu verwenden.

[0039] Der RAID-Controller **102** umfasst einen Prozessor **112**, einen Speicher **114** und eine Abbildungs-Anwendung **116**. Die Abbildungs-Anwendung **116** kann als Hardware, Software, Firmware oder jeder Kombination davon implementiert sein. Die Abbildungs-Anwendung **116** interpretiert Input / Output (I / O)-Anfragen für 512-Byte-Blöcke aus dem Host **104** und bildet solche Anfragen als Lese- und Schreib-Daten in Bezug auf die Platten **106a** ... **106n** in der RAID-Konfiguration **110** ab. Das Lesen und Schreiben wird durch den RAID-Controller **102** in Bezug auf die Platten **106a** ... **106n**, die in der Größe ein Vielfaches von 4 Kilobyte sind, durchgeführt, da die Vielzahl von Platten **106a** ... **106n** der RAID-Konfiguration **110** für eine 4 Kilobyte-Blockgröße konfiguriert sind. Der Host **104** muss nicht von dem RAID-Controller **102** informiert werden, ob die RAID-Konfiguration **110** die 512-Byte-Blockgröße, die von dem Betriebssystem **108** des Hosts **104** verwendet wird, unterstützt oder nicht.

[0040] **Fig. 2** stellt ein Blockdiagramm **200** in Übereinstimmung mit bestimmten Ausführungsformen dar, das zeigt, wie ein einzelner 512-Byte-Block in 4 Kilobyte Blöcken angeordnet ist. Beispielsweise erhält der RAID-Controller **102** in bestimmten Ausführungsformen eine Anfrage von dem Betriebssystem **108** des Hosts **104**, um einen einzelnen 512-Byte-Block 202 auf die RAID-Konfiguration **110** zu schreiben. Als Reaktion auf den Erhalt der Anfrage kann der RAID-Controller **102** jeden der drei 4 Kilobyte-Blöcke lesen, d. h. einen ersten 4K-Block 204, einen zweiten 4K-Block 206, einen dritten 4K-Block 208 oder jeden anderen 4 Kilobyte-Block aus der RAID-Konfiguration **110**, dann ändert er Daten und Parität, und dann schreibt er den 512-Byte-Block 202 in einen der 4 Kilobyte-Blöcke **204**, **206**, **208**, wie in

Fig. 2 gezeigt. Es sollte beachtet werden, dass alte Daten in dem 4 Kilobyte-Block existiert haben könnten, auf den der 512-Byte-Block 202 mit neuen Daten geschrieben wird. Einige Teile der alten Daten oder andere Informationen könnten durch das Schreiben des 512-Byte-Block 202 überschrieben werden. Jedoch werden andere Teile der alten Daten oder anderen Informationen (die nicht durch das Schreiben von dem 512-Byte-Block 202 überschrieben wurden) in die verbleibenden Abschnitte des 4 Kilobyte-Blocks umgeschrieben, als Lese- und Schreibvorgänge in Bezug auf die RAID-Platten **106a** ... **106n**, die in Vielfachen der 4 Kilobyte-Blöcke ablaufen. Zum Beispiel wird in der Konfiguration mit dem Bezugszeichen **210** gezeigt, dass der 512-Byte-Block in den zweiten 4 Kilobyte-Block geschrieben worden ist und unmodifizierte Abschnitte der vorhandenen Daten oder Informationen wieder auf den zweiten 4 Kilobyte-Block geschrieben wurden. Es kann beobachtet werden, dass ein einzelner 4 Kilobyte-Block mindestens einen einzelnen 512-Byte-Block aufnehmen kann.

[0041] **Fig. 3** stellt ein Blockdiagramm **300** in Übereinstimmung mit bestimmten Ausführungsformen dar, das zeigt, wie zwei 512-Byte-Blöcke in 4 Kilobyte Blöcken angeordnet sind. Beispielsweise erhält der RAID-Controller **102** in bestimmten Ausführungsformen eine Anfrage von dem Betriebssystem **108** des Hosts **104**, um zwei 512-Byte-Blöcke 302 auf die RAID-Konfiguration **110** zu schreiben. Als Reaktion auf den Erhalt der Anfrage kann der RAID-Controller **102** einen oder zwei benachbarte Blöcke der drei 4 Kilobyte-Blöcke lesen, d. h. einen ersten 4K-Block 304, einen zweiten 4K-Block 306, einen dritten 4K-Block 308 aus der RAID-Konfiguration **110**, dann ändert er Daten und Parität, und dann schreibt er zwei 512-Byte-Blöcke 302 in wenigsten einen und in nicht mehr als zwei der 4 Kilobyte-Blöcke **304**, **306**, **308**, wie in **Fig. 3** gezeigt. Es sollte beachtet werden, dass alte Daten in wenigsten einem und in nicht mehr als zwei 4 Kilobyte-Blöcken existiert haben könnten, auf denen die zwei 512-Byte-Blöcke 302 mit neuen Daten geschrieben werden. Einige Teile der alten Daten oder andere Informationen könnten durch das Schreiben der zwei 512-Byte-Blöcke 302 überschrieben werden. Jedoch werden andere Teile der alten Daten oder anderen Informationen (die nicht durch das Schreiben der zwei 512-Byte-Blöcke 302 überschrieben wurden) in die verbleibenden Abschnitte wenigstens eines und in nicht mehr als zwei 4 Kilobyte-Blöcke umgeschrieben, als Lese- und Schreibvorgänge in Bezug auf die RAID-Platten **106a** ... **106n**, die in Vielfachen der 4 Kilobyte-Blöcke ablaufen. Zum Beispiel wird in der Konfiguration mit dem Bezugszeichen **310** gezeigt, dass zwei 512-Byte-Blöcke in die benachbarten ersten und zweiten 4 Kilobyte-Blöcke geschrieben worden sind und unmodifizierte Abschnitte der vorhandenen Daten oder Informationen wieder auf die ersten und zweiten 4 Kilobyte-Blöcke geschrieben wurden. Es kann beobachtet werden,

dass wenigstens einer und nicht mehr als zwei 4 Kilobyte-Blöcke zwei 512-Byte-Blöcke aufnehmen können, wobei die beiden 512-Byte-Blöcke nebeneinander gespeichert sind und mit den 4 Kilobyte-Blöcken ausgerichtet sind.

[0042] Fig. 4 stellt ein Blockdiagramm 400 in Übereinstimmung mit bestimmten Ausführungsformen dar, das zeigt, wie neun 512-Byte-Blöcke in 4 Kilobyte Blöcken angeordnet sind. Beispielsweise erhält der RAID-Controller 102 in bestimmten Ausführungsformen eine Anfrage von dem Betriebssystem 108 des Hosts 104, um neun 512-Byte-Blöcke 402 auf die RAID-Konfiguration 110 zu schreiben. Als Reaktion auf den Erhalt der Anfrage kann der RAID-Controller 102 beliebige zwei benachbarte Blöcke lesen, ausgewählt aus den drei 4 Kilobyte-Blöcken 404, 406, 408 aus der RAID-Konfiguration 110, dann ändert er Daten und Parität, und dann schreibt er neun 512-Byte-Blöcke 402 in zwei benachbarte 4 Kilobyte-Blöcke, wie in Fig. 4 gezeigt. Es sollte beachtet werden, dass alte Daten in den zwei benachbarten 4 Kilobyte-Blöcken existiert haben könnten, auf denen die neun 512-Byte-Blöcke 402 mit neuen Daten geschrieben werden. Einige Teile der alten Daten oder andere Informationen könnten durch das Schreiben der neun 512-Byte-Blöcke 402 überschrieben werden. Jedoch werden andere Teile der alten Daten oder anderen Informationen (die nicht durch das Schreiben der neun 512-Byte-Blöcke 402 überschrieben werden) in die verbleibenden Abschnitte der zwei benachbarten 4 Kilobyte-Blöcke umgeschrieben, als Lese- und Schreibvorgänge in Bezug auf die RAID-Platten 106a ... 106n, die in Vielfachen der 4 Kilobyte-Blöcke ablaufen. Zum Beispiel wird in der Konfiguration mit dem Bezugszeichen 410 gezeigt, dass neun 512-Byte-Blöcke in die benachbarten ersten und zweiten 4 Kilobyte-Blöcke geschrieben worden sind und unmodifizierte Abschnitte der vorhandenen Daten oder Informationen wieder auf die ersten und zweiten 4 Kilobyte-Blöcke geschrieben wurden. Es kann beobachtet werden, dass zwei 4 Kilobyte-Blöcke neun 512-Byte-Blöcke aufnehmen können, wobei die neun 512-Byte-Blöcke nebeneinander gespeichert sind und mit den 4 Kilobyte-Blöcken ausgerichtet sind. Ausrichten der Blöcke bedeutet, dass ein einzelner 512-Byte-Block nicht auf mehrere 4K-Blöcke verteilt ist, und acht 512-Byte-Blöcke können in einzelnen 4 Kilobyte-Blöcken, aufgenommen werden, d. h. die Startposition eines 512-Byte-Block innerhalb eines 4 Kilobyte-Blocks ist in Intervallen von 512 Bytes, beginnend am Anfang des 4 Kilobyte-Blocks.

[0043] Fig. 5 stellt ein Blockdiagramm 500 in Übereinstimmung mit bestimmten Ausführungsformen dar, das zeigt, wie zehn 512-Byte-Blöcke in 4 Kilobyte-Blöcken angeordnet sind. Beispielsweise erhält der RAID-Controller 102 in bestimmten Ausführungsformen eine Anfrage von dem Betriebssystem 108

des Hosts 104, um zehn 512-Byte-Blöcke 502 auf die RAID-Konfiguration 110 zu schreiben. Als Reaktion auf den Erhalt der Anfrage kann der RAID-Controller 102 zwei oder drei benachbarte Blöcke lesen, ausgewählt aus den drei 4 Kilobyte-Blöcken 404, 406, 408 aus der RAID-Konfiguration 110, dann ändert er Daten und Parität, und dann schreibt er zehn 512-Byte-Blöcke 502 in wenigstens zwei und in nicht mehr als drei benachbarte 4 Kilobyte-Blöcke, wie in Fig. 5 gezeigt. Es sollte beachtet werden, dass alte Daten in den zwei und nicht mehr als drei benachbarten 4 Kilobyte-Blöcken existiert haben könnten, auf denen die zehn 512-Byte-Blöcke 502 mit neuen Daten geschrieben werden. Einige Teile der alten Daten oder andere Informationen könnten durch das Schreiben der zehn 512-Byte-Blöcke 502 überschrieben werden. Jedoch werden andere Teile der alten Daten oder anderen Informationen (die nicht durch das Schreiben der zehn 512-Byte-Blöcke 502 überschrieben werden) in die verbleibenden Abschnitte der wenigstens zwei und nicht mehr als drei benachbarten 4 Kilobyte-Blöcke, als Lese- und Schreibvorgänge in Bezug auf die RAID-Platten 106a ... 106n, die in Vielfachen der 4 Kilobyte-Blöcke ablaufen, umgeschrieben. Zum Beispiel wird in der Konfiguration mit dem Bezugszeichen 510 gezeigt, dass zehn 512-Byte-Blöcke in die benachbarten ersten, zweiten und dritten 4 Kilobyte-Blöcke geschrieben worden sind und unmodifizierte Abschnitte der vorhandenen Daten oder Informationen wieder auf die ersten, zweiten und dritten 4 Kilobyte-Blöcke geschrieben wurden. Es kann beobachtet werden, dass wenigstens zwei und nicht mehr als drei 4 Kilobyte-Blöcke zehn 512-Byte-Blöcke aufnehmen können, wobei die 512-Byte-Blöcke nebeneinander gespeichert sind und mit den 4 Kilobyte-Blöcken ausgerichtet sind.

[0044] Fig. 6 stellt ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen dar, das bestimmte Platten-I / O-Operationen zeigt, die durch den RAID-Controller 102, der eine RAID-5-Konfiguration steuert, durchgeführt werden. Ein Host 104 fordert das Schreiben von neun Blöcken von neuen Daten an, wobei jeder Block eine Blockgröße von 512 Byte besitzt (Bezugszeichen 602). Der RAID-Controller 102 führt auf den Erhalt der Anforderung (Bezugszeichen 604) vier Sätze von Platten-I / Os mit den ausgewählten Platten der Platten 106a ... 106n durch.

[0045] Zuerst liest der RAID-Controller 102 acht Kilobyte (d. h. zwei 4 Kilobyte-Blöcke) der alten Daten, die in der RAID-Konfiguration 110 gespeichert sind, wobei die beiden 4 Kilobyte-Blöcke Raum enthalten, der angemessen ist, um die 9 Blöcke von neuen Daten aufzunehmen, wobei jeder Block eine Blockgröße von 512 Byte besitzt. Dies wird als der erste Satz von Platten-I / O (Bezugszeichen 606) bezeichnet.

[0046] Dann liest der RAID-Controller **102** die alte Parität, die in der RAID-Konfiguration 110 in Vielfachen der 4 Kilobyte-Blöcke gespeichert ist. Dies wird als der zweite Satz von Platten-I / O (Bezugszeichen **608**) bezeichnet. Der RAID-Controller **102** berechnet nach Bedarf eine neue Parität (Bezugszeichen **610**). Zum Beispiel kann eine neue Parität berechnet werden, wenn sich die alten Daten von den neuen Daten, die geschrieben werden sollen, unterscheiden. Die neue Parität ist eine Funktion der alten Daten, der neuen Daten und der alten Parität. Zum Beispiel können in bestimmten Ausführungsformen ausschließliche Oder-Operationen durchgeführt werden, um die neue Parität aus den alten Daten, den neuen Daten und der alten Parität zu erzeugen.

[0047] Der RAID-Controller **102** schreibt dann (Bezugszeichen **612**) die neun Blöcke an neuen Daten (wobei jeder Block eine Größe von 512 Byte besitzt) auf zwei Blöcke mit jeweils einer Größe von 4K verteilt, wie in **Fig. 4** gezeigt, durch Durchführen einer Verschmelzung der neun neuen Datenblöcke mit sieben alten unveränderten Datenblöcken und durch Schreiben der beiden Blöcke, die jeweils eine 4 Kilobyte-Blockgröße besitzen. Dies wird als der dritte Satz von Platten-I / O (Bezugszeichen **608**) bezeichnet. In dem dritten Satz von Platten-I / Os werden neun Blöcke an neuen Daten und sieben Blöcke an alten Daten (d. h. zwei 4 Kilobyte-Blöcke) auf die Platten der RAID-Konfiguration **110** geschrieben.

[0048] Die neuen Paritätsinformationen werden auch in Vielfachen von 4 Kilobyte-Blöcken geschrieben, und dies wird als der vierte Satz von Platten-I / Os (Bezugszeichen **614**) bezeichnet. So kann in vier Sätzen von Disk-I / O-Operationen der RAID-Controller **102** die 512-Byte-Blöcke in 4K großen Blöcken der RAID-Konfiguration 110 schreiben, ohne dem Host **104** einen Hinweis bereitzustellen, dass die RAID-Konfiguration **110** 512-Byte-Blöcke nicht nativ unterstützt.

[0049] Somit stellt **Fig. 6** bestimmte Ausführungsformen dar, bei denen der RAID-Controller **102** Platten **106a** ... 106n steuert, die als RAID-5 konfiguriert sind. Das Lesen wird über zwei Sätze von Leseoperationen von den Platten **106a** ... 106n durchgeführt, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst. Das Schreiben wird über zwei Sätze von Schreiboperationen von den Platten **106a** ... 106n durchgeführt, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst. Somit sind in RAID-5-Konfigurationen vier Sätze von Platten-Operationen angemessen für das Schreiben

von 512 Byte-Blöcken auf Platten, die in 4 Kilobyte-Blöcken konfiguriert sind.

[0050] In veränderten Ausführungsformen steuert der RAID-Controller **102** Platten **106a** ... 106n, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst. Das Lesen wird über drei Sätze von Leseoperationen von den Platten **106a** ... 106n durchgeführt, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst. Das Schreiben wird über drei Sätze von Schreiboperationen von den Platten **106a** ... 106n durchgeführt, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst. Somit sind in RAID-6-Konfigurationen sechs Sätze von Platten-Operationen angemessen für das Schreiben von 512 Byte-Blöcken auf Platten, die in 4 Kilobyte-Blöcken konfiguriert sind. Zeit kann gespart werden, da der Host **104** nicht informiert werden muss, dass die Platten **106a** ... 106n 512-Byte-Blöcke nicht unterstützen.

[0051] **Fig. 7** stellt ein Flussdiagramm **700** in Übereinstimmung mit bestimmten Ausführungsformen dar, das bestimmte Operationen zeigt. Die in **Fig. 7** gezeigten Operationen können durch den RAID-Controller **102** über die Abbildungs-Anwendung **116** implementiert werden.

[0052] Die Steuerung beginnt (Bezugszeichen **702**), indem der RAID-Controller **102** neue Daten empfängt, die von einem Host **104** geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden (z. B. 512 Byte), wobei der RAID-Controller **102** eine Konfiguration von Platten **106a** ... 106n, ausgewählt aus einer Gruppe bestehend aus RAID-5 und RAID-6, steuert. Zeit kann gespart werden, da der Host **104** nicht informiert werden muss, dass die Platten **106a** ... 106n 512-Byte-Blöcke nicht unterstützen.

[0053] Der RAID-Controller **102** liest die alten Daten und die alte Parität, die den alten Daten entspricht, die in Blöcken von einer zweiten Blockgröße (z. B. 4 Kilobyte), die in der Größe größer als die erste Blockgröße ist, gespeichert werden (Bezugszeichen **704**). Der RAID-Controller **102** berechnet (Bezugszeichen **706**) die neue Parität basierend auf den neuen Daten, den alten Daten und der alten Parität. Zusätz-

lich kann eine Verschmelzung der alten unveränderten Daten und der neuen Daten durchgeführt werden, in Vorbereitung auf die Einleitung des Schreibens der Blöcke der zweiten Blockgröße. Anschließend (Bezugszeichen **708**), schreibt der RAID-Controller **102** die neuen Daten und die neue Parität ausgerichtet auf die Blöcke der zweiten Blockgröße in die Platten **106a** ... **106n**, wobei Teile der alten Daten, die nicht überschrieben werden, vom RAID-Controller auch auf die Blöcke der zweiten Blockgröße geschrieben werden, wobei eine erste Vielzahl von Blöcken der ersten Blockgröße nebeneinander in mindestens einer aus einer zweiten Vielzahl von Blöcken von der zweiten Blockgröße gespeichert ist, und jede der ersten Vielzahl von Blöcken vollständig innerhalb höchstens einer der zweiten Vielzahl von Blöcken gespeichert ist.

[0054] Somit stellt **Fig. 7** bestimmte Ausführungsformen dar, in denen der RAID-Controller **102** 512 Blockgrößen-Anfragen eines Hosts **104** auf eine RAID-Konfiguration mit einer Blockgröße von 4 Kilobyte abbildet.

[0055] **Fig. 8** stellt ein Blockdiagramm **800** in Übereinstimmung mit bestimmten Ausführungsformen dar, das RAID-Controller-Vorrichtungen **802**, **804** zeigt, die außerhalb der Speicher-Controller implementiert sind.

[0056] Die zwei Speicher-Controller **806**, **808** bilden ein Speicher-Teilsystem **810**, das an eine Vielzahl von Hosts **812a** ... **812n** gekoppelt ist, wobei wenigstens einer oder mehrere Hosts **812a** ... **812n** konfiguriert sind, um auf Daten in einer Blockgröße von 512 Byte zuzugreifen. Der Speicher-Controller **806** ist an eine RAID-Controller-Vorrichtung **802** gekoppelt, die außerhalb des Speicher-Controllers **806** ist, und der Speicher-Controller **808** ist gekoppelt an eine RAID-Controller-Vorrichtung **804**, die außerhalb des Speicher-Controllers **808** ist, wobei die Speicher-Controller **802** und **804** eine Vielzahl von Platten **814a** ... **814n** beziehungsweise **816a** ... **816n** steuern, wobei die Vielzahl von Platten **814a** ... **814n** und **816a** ... **816n** Daten in einer 4K Blockgröße speichern. Die Speicher-Controller **806** und **808** können sich gegenseitig ersetzen, um auf Anfragen von jedem der Hosts **812a** ... **812n** zu reagieren.

[0057] Somit stellt **Fig. 8** bestimmte Ausführungsformen dar, in denen RAID-Controller, die außerhalb der Speicher-Controller sind, innerhalb eines Speicher-Teilsystems Anfragen von 512 Blockgröße auf RAID-Konfigurationen abbilden, die Daten und Parität in einer Blockgröße von 4 Kilobyte speichern.

[0058] **Fig. 9** stellt ein Blockdiagramm **900** in Übereinstimmung mit bestimmten Ausführungsformen dar, das RAID-Controller **902**, **904** zeigt, die innerhalb der Speicher-Controller **906**, **908** implementiert

sind. Die zwei Speicher-Controller **906**, **908** bilden ein Speicher-Teilsystem **910**, das an eine Vielzahl von Hosts **912a** ... **912n** gekoppelt ist, wobei wenigstens einer oder mehrere Hosts **912a** ... **912n** konfiguriert sind, um auf Daten in einer Blockgröße von 512 Byte zuzugreifen. Der Speicher-Controller **906** ist an einen RAID-Controller **902** gekoppelt, der intern des Speicher-Controllers **906** ist, und der Speicher-Controller **908** ist gekoppelt an einen RAID-Controller **904**, der intern des Speicher-Controllers **908** ist, wobei die Speicher-Controller **902** und **904** eine Vielzahl von Platten **914a** ... **914n** beziehungsweise **916a** ... **916n** steuern, und wobei die Vielzahl von Platten **914a** ... **914n** und **916a** ... **916n** Daten in einer 4K Blockgröße speichern. Die Speicher-Controller **906** und **908** können sich gegenseitig ersetzen, um auf Anfragen von jedem der Hosts **912a** ... **912n** zu reagieren und jeder dieser Speicher-Controller **906**, **908** umfasst mindestens einen Prozessor und einen Speicher (dargestellt durch Bezugszeichen **918**, **920**, die Prozessoren darstellen, und die Bezugszeichen **922**, **924**, die Speicher darstellen) zusätzlich zu den RAID-Controllern.

[0059] Somit stellt **Fig. 9** bestimmte Ausführungsformen dar, in denen RAID-Controller, die intern zu Speicher-Controllern sind, innerhalb eines Speicher-Teilsystems Anfragen von 512 Blockgröße auf RAID-Konfigurationen abbilden, die Daten und Parität in einer Blockgröße von 4 Kilobyte speichern.

[0060] **Fig. 1 - Fig. 9** stellen bestimmte Ausführungsformen dar, in denen RAID-Controller Anfragen, gemacht mit 512 Blockgröße, abbilden, um Daten und Parität mit einer Blockgröße von 4 Kilobyte auf eine RAID-Konfiguration zu speichern, die sowohl Daten als auch Parität speichert.

[0061] Bestimmte Ausführungsformen haben Abbildungen von 512-Byte-Blöcken auf 4 Kilobyte-Blöcken gezeigt. In alternativen Ausführungsformen können andere Arten von Abbildungen durchgeführt werden. Zum Beispiel können 512 Byte-, 524 Byte-, oder 528-Byte-Blöcke auf der Host-Seite auf 4096 Byte- (d. h. 4 Kilobyte), 4192 Byte- oder 4224 Byte-Blöcke auf der Platten-Seite abgebildet werden.

[0062] Es sei darauf hingewiesen, dass der Host nicht wissen muss, dass die RAID-Konfiguration die Blockgröße, die durch den Host verwendet wird, nicht nativ unterstützt. Darüber hinaus muss die RAID-Konfiguration nicht wissen, dass der Host die Blockgröße, die von der RAID-Konfiguration verwendet, um Informationen auf Platten zu speichern, nicht nativ unterstützt.

[0063] Es sei auch darauf hingewiesen, dass in bestimmten Ausführungsformen keine zusätzlichen Laufwerk-Operationen durchgeführt werden, im Vergleich zu gleichen Schreiboperationen, wenn der Host und die RAID-Konfiguration die gleiche Block-

größe verwenden. Zusätzlich müssen in bestimmten Ausführungsformen die Plattenlaufwerke Lese-, Ändern- und Schreib-Operationen intern nicht durchführen, und sparen somit Verarbeitungszeit.

[0064] Es sei auch darauf hingewiesen, dass sich in bestimmten Ausführungsformen zwei RAID-Controller die gleichen Platten teilen können. Zum Beispiel, in **Fig. 8**, können sich beide RAID-Controller **802** und **804** die Platten **814a** ... **814n** teilen. Ebenso wie sich zwei Speicher-Controller gegenseitig ersetzen können, so können dies in bestimmten Ausführungsformen auch zwei RAID-Controller.

[0065] In bestimmten Ausführungsformen können die Operationen, die von dem RAID-Controller durchgeführt werden, in einem Plattenlaufwerk oder in einer anderen Vorrichtung durchgeführt werden. In noch weiteren Ausführungsformen können die RAID-Controller in Software, Hardware, Firmware oder einer Kombination davon implementiert werden.

Zusätzliche Ausführungsdetails

[0066] Die beschriebenen Operationen können als Verfahren, Gerät oder Computerprogrammprodukt unter Verwendung von Standard-Programmierung und / oder Engineering-Techniken, um Software, Firmware, Hardware oder eine Kombination davon zu produzieren, ausgeführt sein. Demzufolge können Aspekte der vorliegenden Erfindung die Form einer komplett in Hardware realisierten Ausführungsform, einer komplett in Software realisierten Ausführungsform (umfassend Firmware, residente Software, Mikrocode usw.) oder einer Ausführungsform annehmen, in der Software- und Hardwareaspekte kombiniert sind, die im vorliegenden Dokument allgemein als „Schaltkreis“, „Modul“ oder „System“ bezeichnet werden. Ferner können Aspekte der vorliegenden Erfindung die Form eines Computerprogrammprodukts annehmen, das in einem oder mehreren computerlesbaren Medien verkörpert ist, auf denen computerlesbarer Programmcode verkörpert ist.

[0067] Es können beliebige Kombinationen eines oder mehrerer computerlesbarer Medien verwendet werden. Das computerlesbare Medium kann ein computerlesbares Signalmedium oder ein computerlesbares Speichermedium sein. Zu computerlesbaren Speichermedien können beispielsweise und ohne Einschränkung ein elektronisches, magnetisches, optisches, elektromagnetisches, infrarotes oder Halbleiter-System, Gerät oder Vorrichtung oder beliebige geeignete Kombinationen des Vorstehenden gehören. Detaillierte Beispiele (eine nicht erschöpfende Liste) computerlesbarer Speichermedien umfassen unter anderem folgende: elektrische Verbindung mit einer oder mehreren Leitungen, portable Computerdiskette, Festplatte, Direktzugriffsspeicher (RAM), Nur-Lese-Speicher (ROM), löscha-

rer programmierbarer Nur-Lese-Speicher (EPROM oder Flash-Speicher), Lichtwellenleiter, Nur-Lese-Speicher in Form einer portablen Compact Disc (CD-ROM), optische Speichervorrichtung, magnetische Speichervorrichtung oder beliebige geeignete Kombinationen des Vorstehenden. Im Kontext des vorliegenden Dokuments kann ein computerlesbares Speichermedium jedes beliebige greifbare Medium sein, das ein Programm enthalten oder speichern kann, das von oder in Verbindung mit einem System, einem Gerät oder einer Vorrichtung zur Befehlsausführung genutzt werden kann.

[0068] Ein computerlesbares Signalmedium kann ein im Basisband oder als Teil einer Trägerwelle übertragenes Datensignal mit darin verkörpertem computerlesbarem Programmcode sein. Ein derartiges übertragenes Signal kann eine beliebige Vielfalt von Formen annehmen, darunter und ohne Einschränkung eine elektromagnetische oder optische Form oder beliebige geeignete Kombinationen davon. Ein computerlesbares Signalmedium kann ein beliebiges computerlesbares Medium sein, bei dem es sich nicht um ein computerlesbares Speichermedium handelt und das ein Programm senden, übertragen oder transportieren kann, das von oder in Verbindung mit einem System, einem Gerät oder einer Vorrichtung zur Befehlsausführung genutzt werden kann.

[0069] Auf einem computerlesbaren Medium verkörperter Programmcode kann unter Verwendung eines beliebigen geeigneten Mediums, darunter und ohne Einschränkung, drahtlose oder drahtgebundene Medien, Lichtwellenleiter, RF usw., oder beliebiger geeigneter Kombinationen des Vorstehenden übertragen werden.

[0070] Computerprogrammcode zur Ausführung von Operationen der vorliegenden Erfindung kann in einer oder mehrerer Programmiersprachen, einschließlich einer objektorientierten Programmiersprache wie Java (* Java ist eine Marke oder eingetragene Marke von Oracle und / oder ihrer Tochtergesellschaften.), Smalltalk, C++ oder Ähnlichem und einer herkömmlichen prozeduralen Programmiersprache wie z.B. in der Programmiersprache „C“ oder in ähnlichen Programmiersprachen geschrieben sein. Der Programmcode kann vollständig auf dem Computer des Benutzers, teilweise auf dem Computer des Benutzers, als eigenständiges Softwarepaket, teilweise auf dem Computer des Benutzers und teilweise auf einem entfernt angeordneten Computer oder vollständig auf dem entfernt angeordneten Computer oder Server ausgeführt werden. Beim letztgenannten Szenario kann der entfernt angeordnete Computer mit dem Computer des Benutzers über ein lokales Netzwerk (LAN) oder ein Weitverkehrsnetzwerk (WAN) verbunden sein, oder die Verbindung kann zu einem externen Computer hergestellt werden (bei-

spielsweise über das Internet unter Nutzung eines Internet-Dienstanbieters (Internet Service Provider)).

[0071] Aspekte der vorliegenden Erfindung sind oben unter Bezugnahme auf Flussdiagramm-Darstellungen und / oder Blockdiagramme von Methoden, Geräten (Systemen) und Computerprogrammen gemäß Ausführungsformen der Erfindung beschrieben. Es versteht sich, dass jeder Abschnitt der Flussdiagramm-Darstellungen und / oder der Blockdiagramme und Kombinationen von Abschnitten in den Flussdiagramm-Darstellungen und / oder Blockdiagrammen durch Computerprogrammanweisungen realisiert werden kann. Diese Computerprogrammanweisungen können einem Prozessor eines Mehrzweckcomputers, eines Spezialcomputers oder anderer programmierbarer Datenverarbeitungseinrichtungen, die eine Maschine darstellen, bereitgestellt werden, sodass die Anweisungen, die über den Prozessor des Computers oder anderer programmierbarer Datenverarbeitungsvorrichtungen ausgeführt werden, Mittel schaffen, um die in einem Abschnitt oder in den Abschnitten des Flussdiagramms und/oder des Blockdiagramms angegebenen Funktionen / Aktionen zu realisieren.

[0072] Diese Computerprogrammanweisungen können auch in einem computerlesbaren Speicher gespeichert sein, der einen Computer oder andere programmierbare Datenverarbeitungsgeräte oder andere Vorrichtungen anweisen kann, in einer bestimmten Weise zu funktionieren, sodass die im computerlesbaren Speicher gespeicherten Anweisungen ein Produkt erzeugen, das die Anweisungen enthält, die die in einem Abschnitt oder Abschnitten des Flussdiagramms und/oder des Blockdiagramms angegebene Funktion/Aktion realisiert.

[0073] Die Computerprogrammanweisungen können auch in einen Computer, in andere programmierbare Datenverarbeitungsgeräte oder andere Vorrichtungen geladen werden, um zu bewirken, dass auf dem Computer, auf anderen programmierbaren Geräten oder Vorrichtungen eine Reihe von Arbeitsschritten ausgeführt werden, um einen mittels Computer ausgeführten Prozess zu schaffen, mit dessen Hilfe die Anweisungen, die auf dem Computer oder auf anderen programmierbaren Vorrichtungen ausgeführt werden, Schritte zur Ausführung der in einem Abschnitt oder Abschnitten des Flussdiagramms und/oder des Blockdiagramms angegebenen Funktionen/Aktionen bereitstellen.

[0074] Fig. 10 stellt ein Blockdiagramm in Übereinstimmung mit bestimmten Ausführungsformen dar, das bestimmte Elemente zeigt, die die Speicher-Controller 806, 808, 906, 908 oder die RAID-Controller 102, 802, 804, 902, 904 umfassen können. Das System 100 kann die Speicher-Controller 806, 808, 906, 908 oder die RAID-Controller 102, 802, 804, 902,

904 umfassen und kann einen Schaltkreis 1002 umfassen, der in bestimmten Ausführungsformen mindestens einen Prozessor 1004 umfassen kann. Das System 1000 kann auch einen Speicher 1006 (z. B. eine flüchtige Speichervorrichtung) und einen Speicher 1008 umfassen. Der Speicher 1008 kann eine nicht-flüchtige Speichervorrichtung (z. B. EEPROM, ROM, PROM, RAM, DRAM, SRAM, Flash, Firmware, programmierbare Logik usw.), magnetische Festplatte, optische Platte, ein Bandlaufwerk usw. umfassen. Der Speicher 1008 kann eine interne Speichervorrichtung, eine angeschlossene Speichervorrichtung und / oder eine Netzwerk zugängliche Speichervorrichtung umfassen. Das System 1000 kann eine Programmlogik 1010 einschließlich Code 1012 umfassen, die in den Speicher 1006 geladen werden kann, und durch den Prozessor 1004 oder den Schaltkreis 1002 ausgeführt wird. In bestimmten Ausführungsformen kann die Programmlogik 1010 einschließlich Code 1012 in dem Speicher 1008 gespeichert werden. In bestimmten anderen Ausführungsformen kann die Programmlogik 1010 in den Schaltkreis 1002 implementiert sein. Deshalb kann, während Fig. 10 die Programmlogik 1010 getrennt von den anderen Elementen zeigt, die Programm-Logik 1010 in dem Speicher 1006 und / oder dem Schaltkreis 1002 implementiert werden.

[0075] Bestimmte Ausführungsformen können ein Verfahren für die Bereitstellung von Rechenanweisungen durch eine Person oder einer automatisierten Verarbeitung von integriertem computerlesbarem Code in ein Computersystem anordnen, wobei der Code in Kombination mit dem Computersystem ermöglicht, dass die Operationen der beschriebenen Ausführungsformen durchgeführt werden.

[0076] Die Begriffe „eine Ausführungsform“, „Ausführungsform“, „Ausführungsformen“, „die Ausführungsform“, „die Ausführungsformen“, „eine oder mehrere Ausführungsformen“, „einige Ausführungsformen“ und „eine Ausführungsform“ bedeuten „eine oder mehrere (aber nicht alle) Ausführungsformen der vorliegenden Erfindung(en)“, soweit nicht ausdrücklich anders angegeben.

[0077] Die Begriffe „einschließlich“, „umfassend“, „haben“ und Variationen davon bedeuten „umfassend, aber nicht darauf beschränkt“, sofern nicht ausdrücklich anders angegeben.

[0078] Die Liste der aufgezählten Elemente bedeutet nicht, dass einige oder alle Elemente sich gegenseitig ausschließen, sofern nicht ausdrücklich anders angegeben.

[0079] Die Begriffe „ein/eine“ und „der/die/das“ bedeuten „einen oder mehrere“, sofern nicht ausdrücklich anders angegeben.

[0080] Vorrichtungen, die miteinander kommunizieren, müssen nicht in kontinuierlicher Kommunikation miteinander stehen, sofern nicht ausdrücklich anders angegeben.

[0081] Darüber hinaus können Vorrichtungen, die miteinander kommunizieren, direkt oder indirekt über eine oder mehrere Zwischenstufen kommunizieren.

[0082] Eine Beschreibung einer Ausführungsform mit mehreren Komponenten, die miteinander kommunizieren, bedeutet nicht, dass alle diese Komponenten erforderlich sind. Im Gegenteil, es ist eine Vielzahl von optionalen Komponenten beschrieben, um die große Vielzahl von möglichen Ausführungsformen der vorliegenden Erfindung darzustellen.

[0083] Obwohl ferner Prozessschritte, Verfahrensschritte, Algorithmen oder dergleichen in einer sequentiellen Reihenfolge beschrieben werden können, können solche Prozesse, Verfahren und Algorithmen konfiguriert sein, um in alternativen Reihenfolgen zu arbeiten. Mit anderen Worten, jede Sequenz oder Reihenfolge der Schritte, die beschrieben werden können, müssen nicht unbedingt auf eine Anforderung hindeuten, dass die Schritte in dieser Reihenfolge durchgeführt werden. Die Schritte der hierin beschriebenen Prozesse können in beliebiger praktischer Reihenfolge durchgeführt werden. Ferner können einige Schritte gleichzeitig durchgeführt werden.

[0084] Wenn eine einzelne Vorrichtung oder Gegenstand hierin beschrieben wird, wird es leicht ersichtlich sein, dass mehr als eine Vorrichtung/ Gegenstand (ob sie kooperieren oder nicht) anstelle von einer einzigen Vorrichtung/ Gegenstand verwendet werden kann. In ähnlicher Weise wird es leicht ersichtlich sein, wenn mehr als eine Vorrichtung oder Gegenstand (ob sie kooperieren oder nicht) hierin beschrieben wird, dass eine einzelne Vorrichtung/ Gegenstand anstelle der mehr als einen Vorrichtung oder Gegenstand verwendet werden kann oder es kann eine unterschiedliche Anzahl von Vorrichtungen / Gegenständen anstelle der gezeigten Anzahl von Vorrichtungen oder Programmen verwendet werden. Die Funktionalität und / oder die Merkmale einer Vorrichtung können alternativ durch eine oder mehrere andere Vorrichtungen ausgeführt werden, die nicht explizit mit derartiger Funktionalität / Merkmalen beschrieben werden. Daher müssen andere Ausführungsformen der vorliegenden Erfindung nicht die Vorrichtung selbst umfassen.

[0085] Zumindest bestimmte Operationen, die in den Figuren dargestellt worden sind, können bestimmte Ereignisse zeigen, die in einer bestimmten Reihenfolge stattfinden. In alternativen Ausführungsformen können bestimmte Operationen in einer anderen Reihenfolge ausgeführt, geändert oder entfernt werden. Darüber hinaus können Schritte der oben

beschriebenen Logik hinzugefügt werden und dennoch den beschriebenen Ausführungsformen entsprechen. Ferner können hier beschriebene Operationen nacheinander erfolgen oder bestimmte Operationen können parallel verarbeitet werden. Ferner können Operationen von einer einzigen Verarbeitungseinheit oder von verteilten Verarbeitungseinheiten durchgeführt werden.

[0086] Die vorstehende Beschreibung der verschiedenen Ausführungsformen der Erfindung wurde zum Zwecke der Veranschaulichung und Beschreibung dargestellt. Sie ist nicht beabsichtigt, erschöpfend zu sein oder die Erfindung auf die genau offenbarte Form zu beschränken. Viele Modifikationen und Variationen sind im Lichte der obigen Lehre möglich. Es ist beabsichtigt, dass der Umfang der Erfindung nicht durch diese detaillierte Beschreibung, sondern vielmehr durch die beigefügten Ansprüche beschränkt wird. Die obige Beschreibung, Beispiele und Daten liefern eine vollständige Beschreibung der Herstellung und Verwendung der Zusammensetzung der Erfindung. Da viele Ausführungsformen der Erfindung hergestellt werden können, ohne vom Geist und Umfang der Erfindung abzuweichen, liegt die Erfindung in den angefügten Ansprüchen.

Patentansprüche

1. Verfahren implementiert in einer Vorrichtung, wobei das Verfahren umfasst:
Empfangen neuer Daten, durch einen „Redundante Anordnung unabhängiger Platten“ (RAID)-Controller (102), die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden;
Lesen alter Daten (606) und alter Parität (608) durch den RAID-Controller (102), die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer ist als die erste Blockgröße;
Berechnen einer neuen Parität (610) durch den RAID-Controller (102), basierend auf den neuen Daten, den alten Daten, und der alten Parität; und
Schreiben der neuen Daten und der neuen Parität durch den RAID-Controller (102), ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller (102) überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden, wobei der RAID-Controller (102) Platten steuert, die als RAID-5 konfiguriert sind, wobei:
das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst (606) und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst (608); und
das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht

überschrieben werden, umfasst (612) und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst (614) oder

wobei der RAID-Controller (102) Platten steuert, die als RAID-6 konfiguriert sind, wobei die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst, wobei die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst, und wobei:

das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

2. Verfahren nach Anspruch 1, wobei: die erste Blockgröße 512 Byte beträgt; die zweite Blockgröße 4 Kilobyte beträgt; die neuen Daten durch den RAID-Controller (102) von einem Host (104) empfangen werden, dessen Betriebssystem so konfiguriert ist, um auf die Blöcke der ersten Blockgröße von 512 Byte zuzugreifen, und die Blöcke der ersten Blockgröße von 512 Byte Steuerungs-Schlüsseldaten (CKD)-Datensätze speichern und jeder CKD-Datensatz auf eine 512-Byte-Block-Grenze ausgerichtet ist.

3. Verfahren nach Anspruch 1, wobei: die neuen Daten eine erste Vielzahl von Blöcken der ersten Blockgröße umfassen; die erste Vielzahl von Blöcken der ersten Blockgröße nebeneinander in mindestens einem einer zweiten Vielzahl von Blöcken der zweiten Blockgröße gespeichert ist; jede der ersten Vielzahl von Blöcken vollständig innerhalb höchstens einem der zweiten Vielzahl von Blöcken gespeichert ist.

4. System, umfassend: einen Speicher (1008); und einen Prozessor (1004) mit dem Speicher (1008) gekoppelt, wobei das System zur Ansteuerung von als RAID-5 oder als RAID-6 konfigurierten Platten ausgebildet ist, wobei der Prozessor (1004) durchführt: Empfangen neuer Daten, die geschrieben werden sollen, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden;

Lesen alter Daten (606) und alter Parität (608), die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer ist als die erste Blockgröße;

Berechnen einer neuen Parität (610) basierend auf den neuen Daten, den alten Daten, und der alten Parität; und

Schreiben der neuen Daten (612) und der neuen Parität (614) ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller (102) überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden,

wobei bei als RAID-5 konfigurierten Platten:

das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst (606) und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst (608); und das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst (612) und ein zweiter Satz von Schreiboperationen das Schreiben der neuen Parität umfasst (614) oder

wobei bei als RAID-6 konfigurierten Platten, bei denen die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst und bei denen die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst,:

das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

5. Speicher-Controller umfassend:

einen Speicher (1008);

einen Prozessor (1004) mit dem Speicher (1008) gekoppelt; und

einen RAID-Controller mit dem Speicher (1008) gekoppelt, der zur Ansteuerung von als RAID-5 oder als RAID-6 konfigurierten Platten ausgebildet ist, wobei der RAID-Controller durchführt:

Empfangen neuer zu schreibender Daten, wobei die neuen Daten in Blöcken einer ersten Blockgröße angegeben werden;

Lesen alter Daten (606) und alter Parität (608), die den alten Daten entspricht, gespeichert in Blöcken einer zweiten Blockgröße, die größer ist als die erste Blockgröße;

Berechnen einer neuen Parität (610) basierend auf den neuen Daten, den alten Daten, und der alten Parität; und

Schreiben der neuen Daten (612) und der neuen Parität (614) ausgerichtet auf die Blöcke der zweiten Blockgröße, wobei Teile der alten Daten, die nicht durch den RAID-Controller überschrieben werden, ebenfalls in die Blöcke der zweiten Blockgröße geschrieben werden,

wobei bei als RAID-5 konfigurierten Platten:

das Lesen über zwei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst (606) und ein zweiter Satz von Leseoperationen das Lesen der alten Parität umfasst (608); und

das Schreiben über zwei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst (612) und ein zweiter Satz von

Schreiboperationen das Schreiben der neuen Parität umfasst (614) oder

wobei bei als RAID-6 konfigurierten Platten, bei denen die alte Parität einen ersten Satz der alten Parität und einen zweiten Satz der alten Parität umfasst und bei denen die neue Parität einen ersten Satz der neuen Parität und einen zweiten Satz der neuen Parität umfasst,:

das Lesen über drei Sätze von Leseoperationen von den Platten durchgeführt wird, wobei ein erster Satz von Leseoperationen das Lesen der alten Daten umfasst, ein zweiter Satz von Leseoperationen das Lesen des ersten Satzes der alten Parität umfasst, und ein dritter Satz von Leseoperationen das Lesen des zweiten Satzes der alten Parität umfasst; und

das Schreiben über drei Sätze von Schreiboperationen auf den Platten durchgeführt wird, wobei ein erster Satz von Schreiboperationen das Schreiben der neuen Daten und Teilen der alten Daten, die nicht überschrieben werden, umfasst, ein zweiter Satz von Schreiboperationen das Schreiben des ersten Satzes der neuen Parität umfasst, und ein dritter Satz von Schreiboperationen das Schreiben des zweiten Satzes der neuen Parität umfasst.

6. Computerprogrammprodukt, wobei das Computerprogrammprodukt umfasst:

ein computerlesbares Speichermedium mit darin enthaltenem computerlesbarem Programmcode zur Durchführung der Verfahrensschritte gemäß einem der Ansprüche 1 bis 3.

7. Plattenlaufwerk, umfassend ein System gemäß Anspruch 4 oder einen Speicher-Controller nach Anspruch 5.

Es folgen 10 Seiten Zeichnungen

Anhängende Zeichnungen

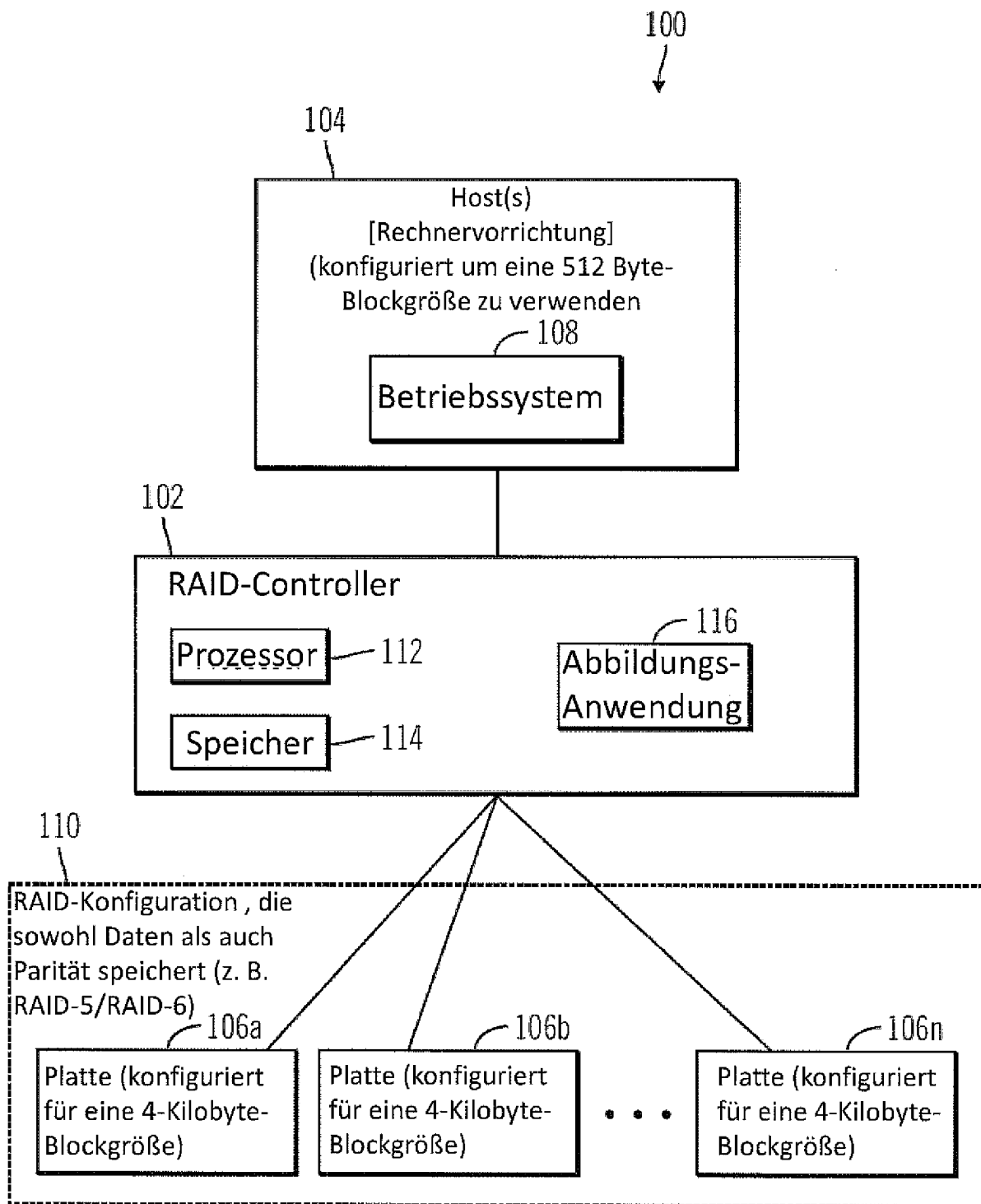


FIG. 1

FIG. 2

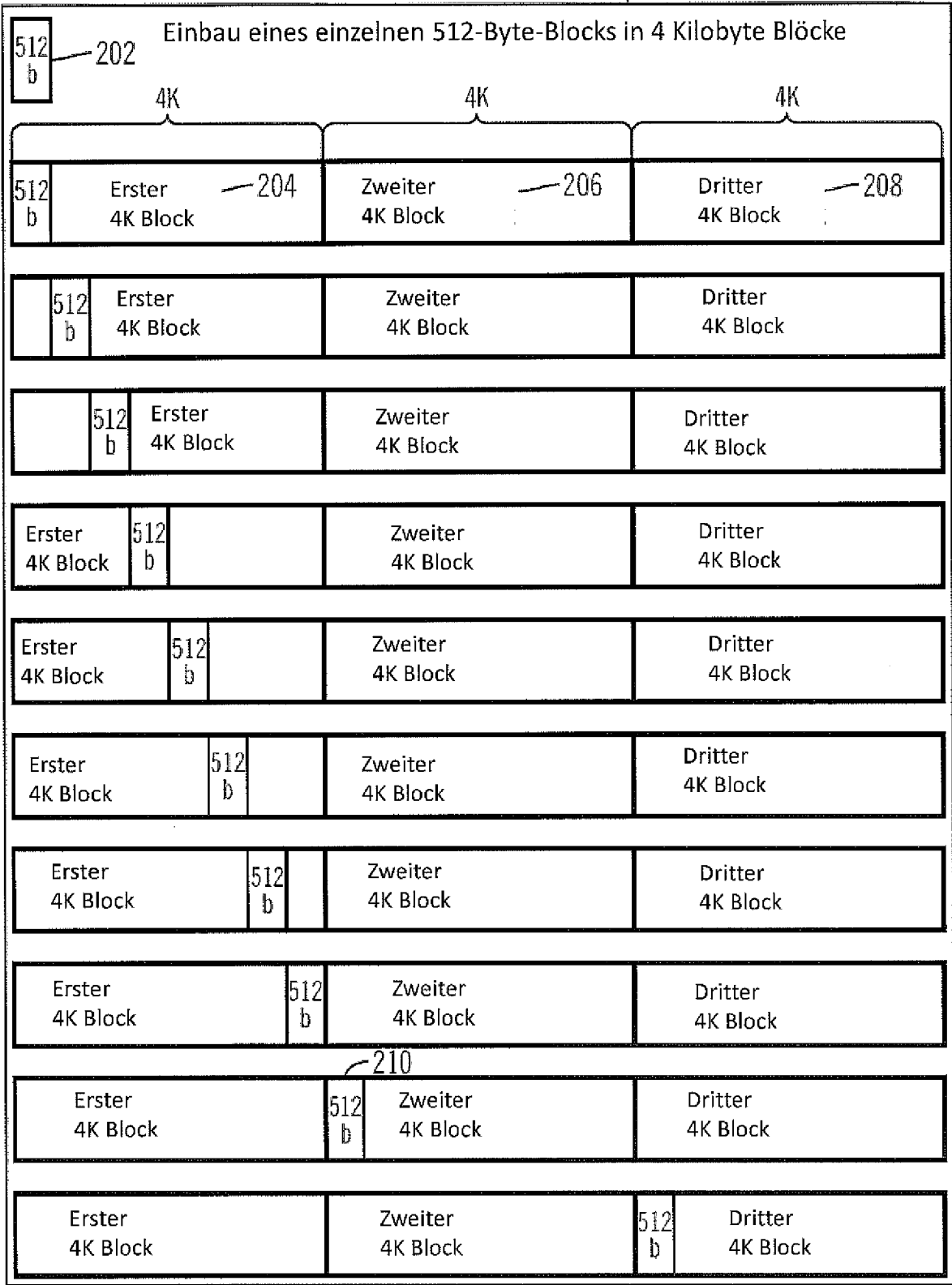
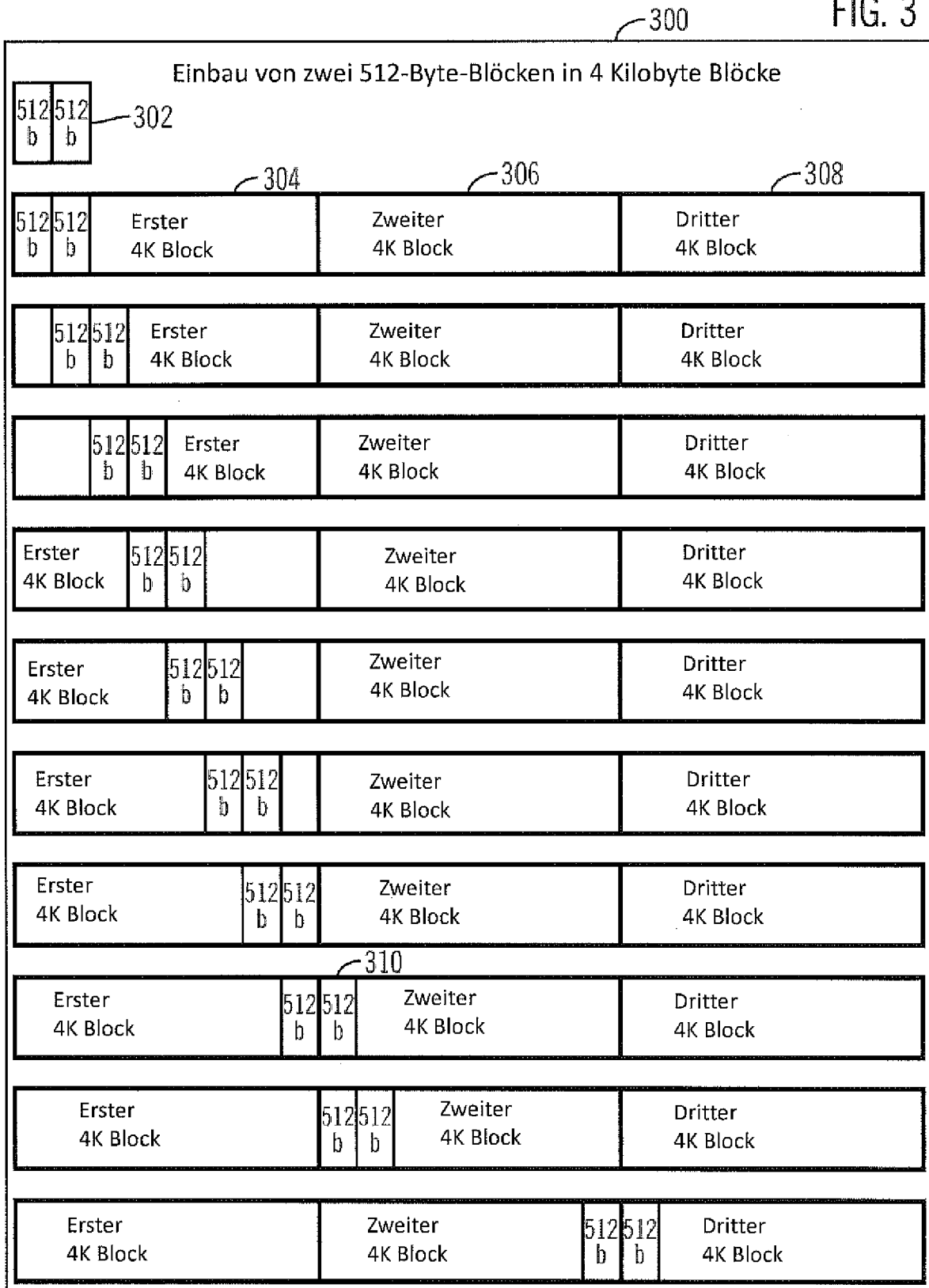


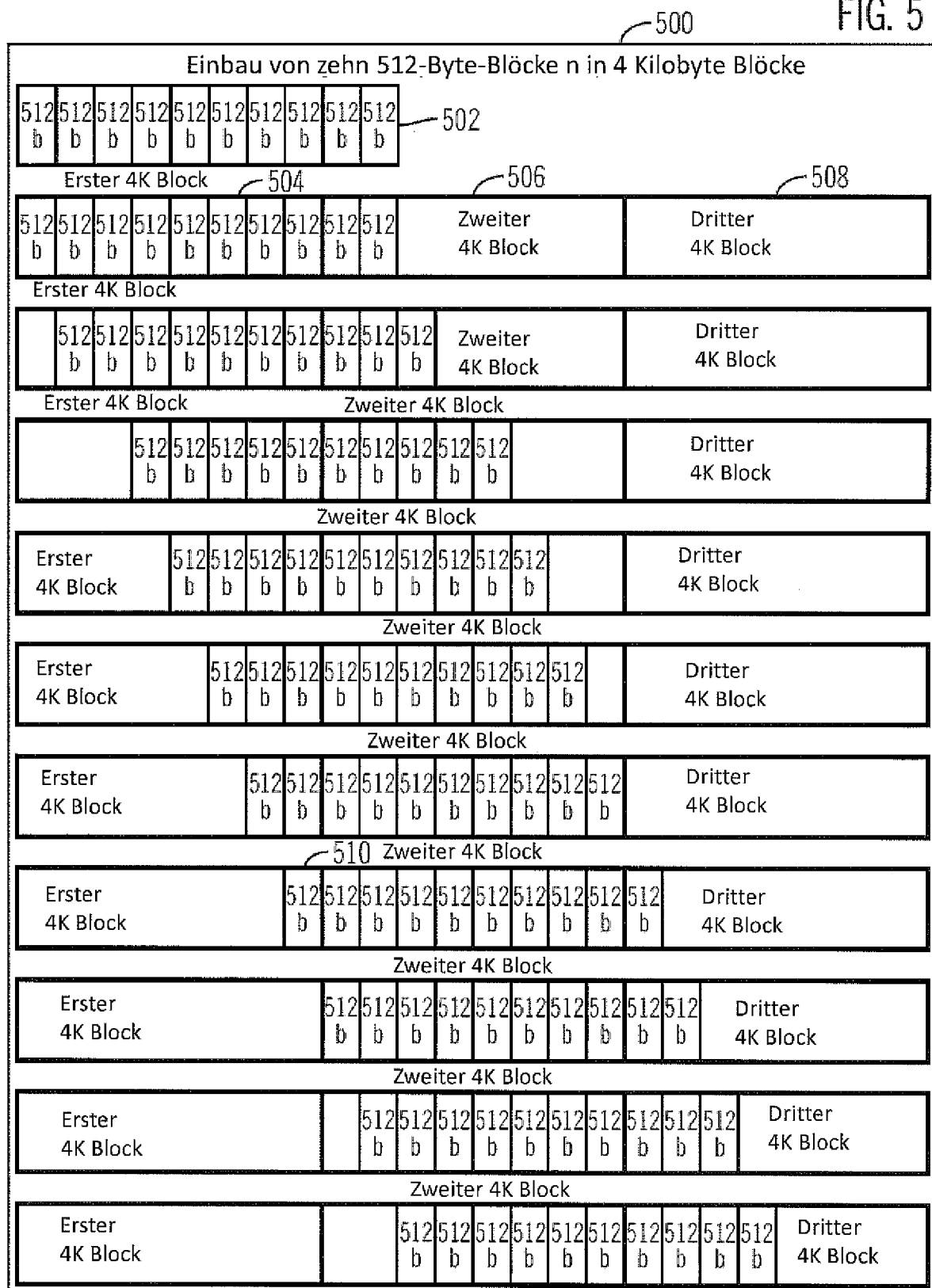
FIG. 3



Einbau von neun 512-Byte-Blöcke n in 4 Kilobyte Blöcke



FIG. 5



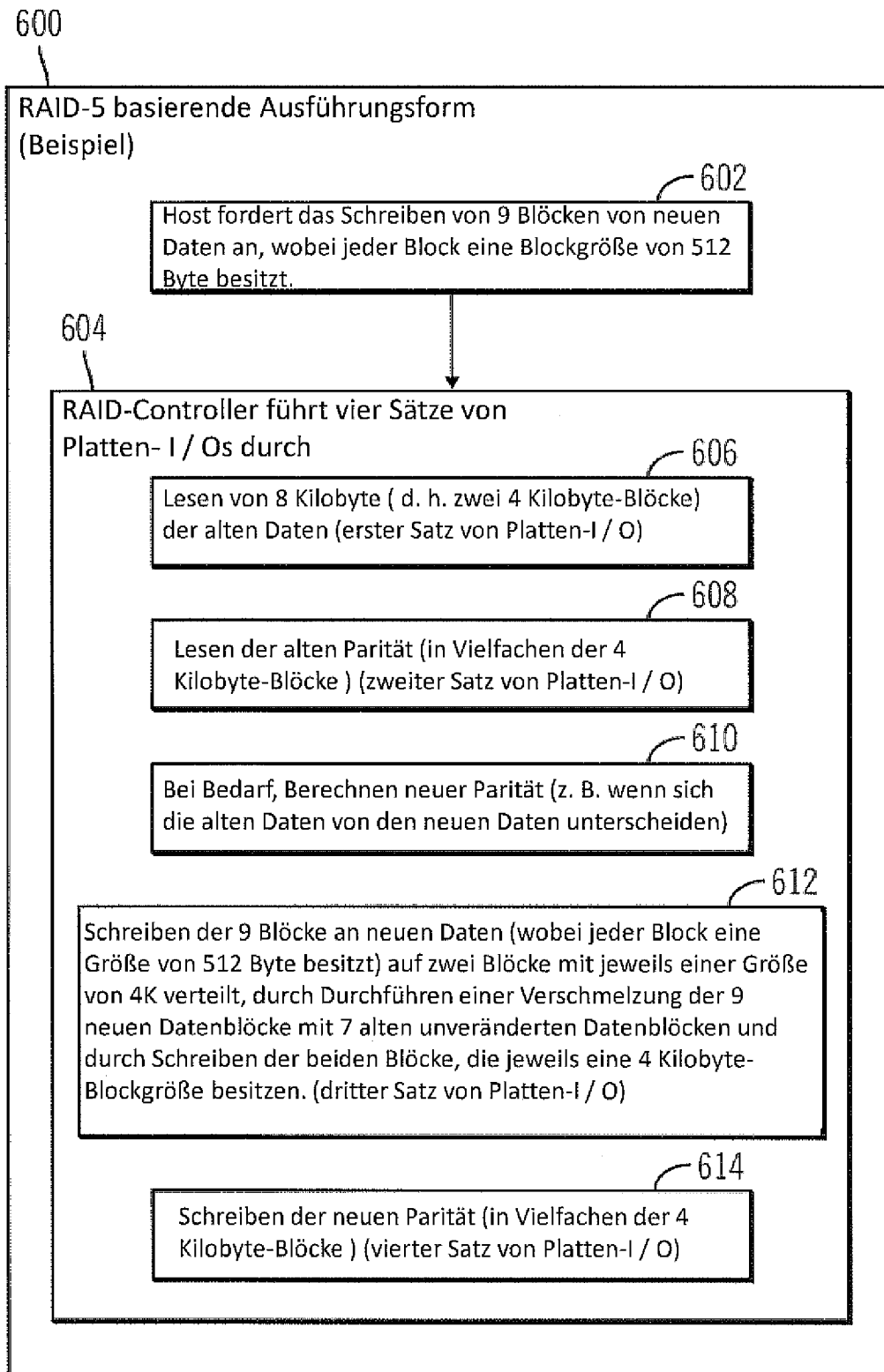


FIG. 6

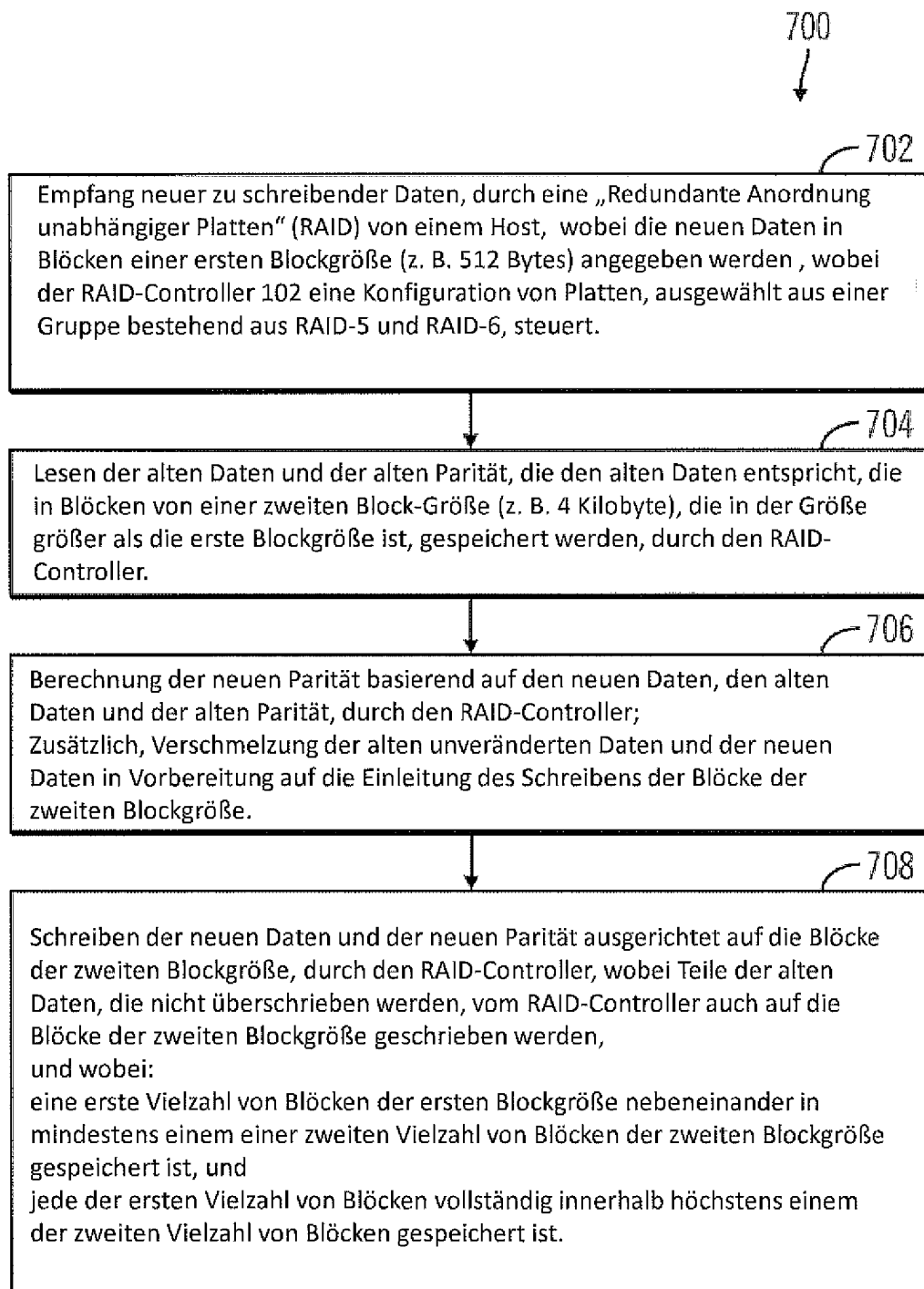


FIG. 7

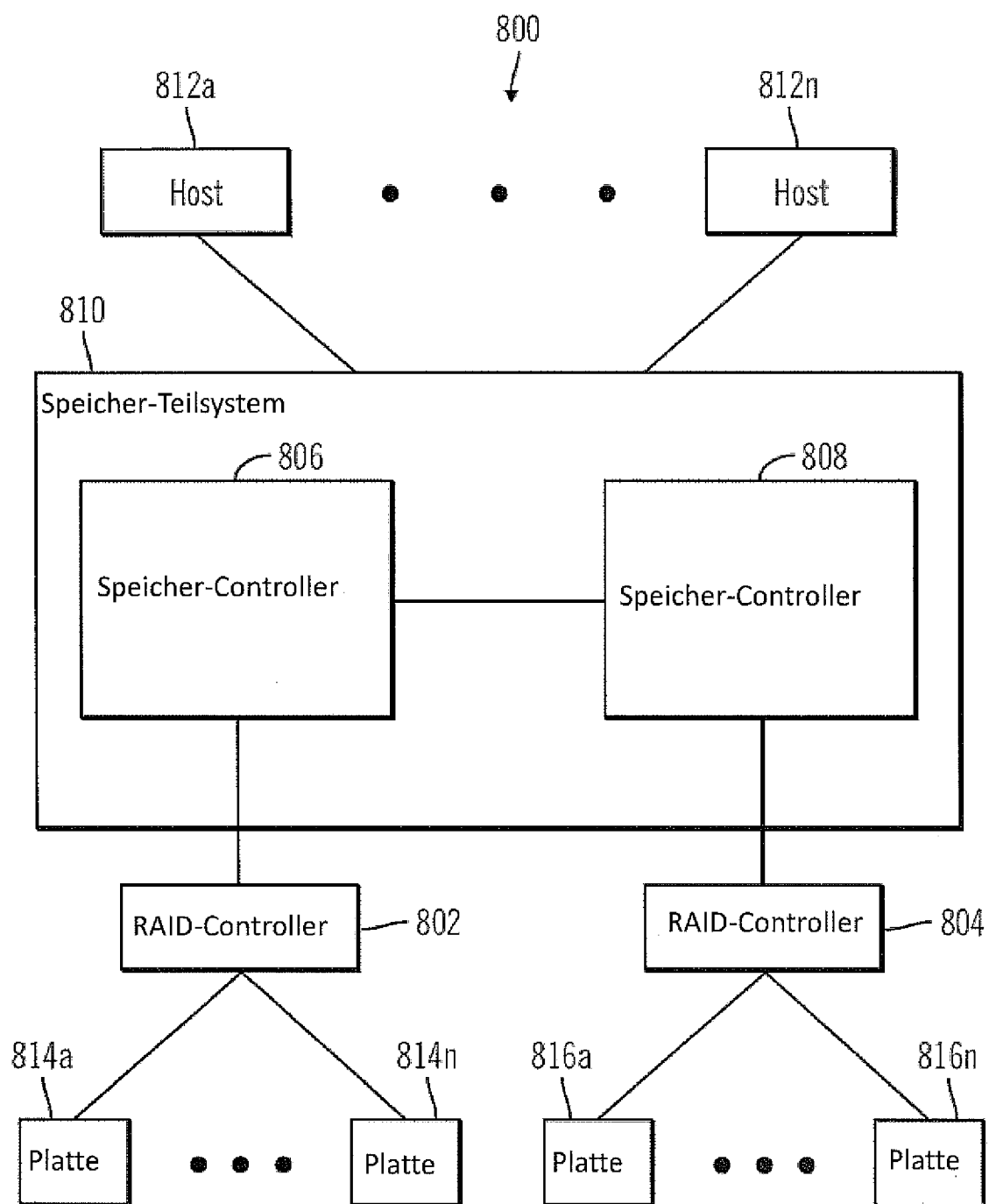


FIG. 8

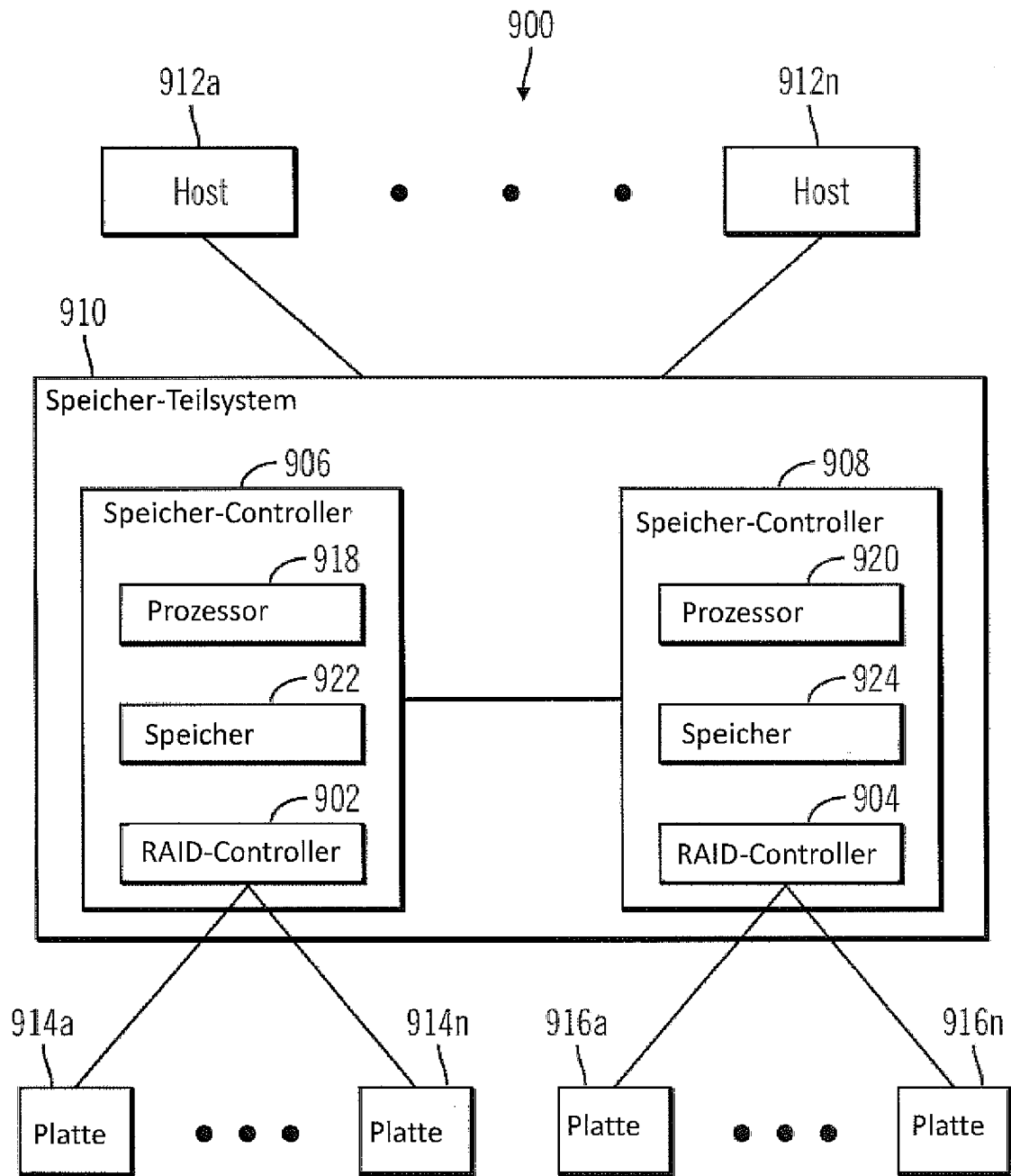


FIG. 9

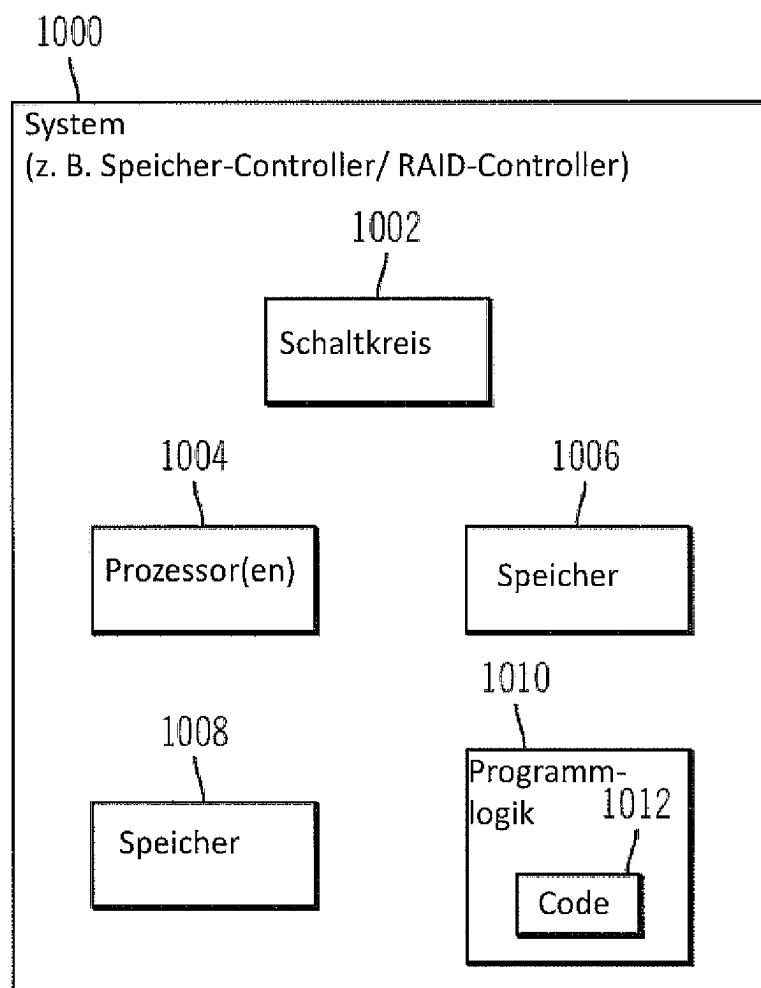


FIG. 10