



(19)
 Bundesrepublik Deutschland
 Deutsches Patent- und Markenamt

(10) DE 10 2004 013 129 A1 2005.06.30

(12)

Offenlegungsschrift

(21) Aktenzeichen: 10 2004 013 129.5

(22) Anmeldetag: 17.03.2004

(43) Offenlegungstag: 30.06.2005

(51) Int Cl.7: G11B 5/012
 G11B 33/14

(30) Unionspriorität:
 2003/400517 28.11.2003 JP

(71) Anmelder:
 Hitachi, Ltd., Tokio/Tokyo, JP

(74) Vertreter:
 Strehl, Schübel-Hopf & Partner, 80538 München

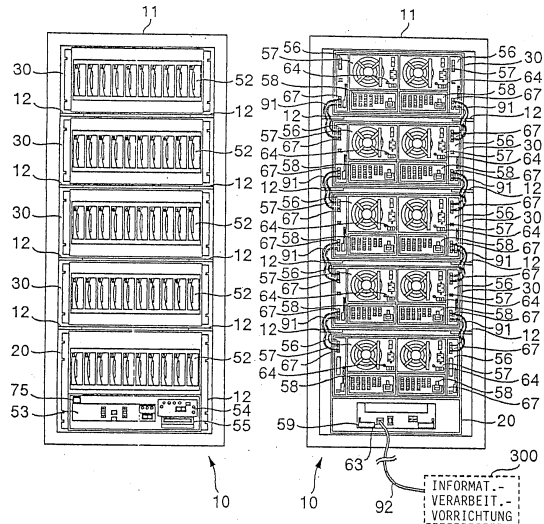
(72) Erfinder:
 Kano, Azuma, Tokio/Tokyo, JP; Ogawa, Takuji,
 Tokio/Tokyo, JP; Yagisawa, Ikuya, Tokio/Tokyo, JP

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Bezeichnung: Plattensystem und Verfahren zum Steuern eines Plattenarraysystems

(57) Zusammenfassung: Ein Plattenarraysystem verfügt über ein erstes und ein zweites Gehäuse sowie einen Controller zum Steuern des ersten und des zweiten Gehäuses. Im ersten Gehäuse sind Faserkanal-Festplattenlaufwerke untergebracht, und im zweiten Gehäuse sind Seriell-ATA-Festplattenlaufwerke untergebracht. Wenn in einem Seriell-ATA-Festplattenlaufwerk im zweiten Gehäuse gespeicherte Daten gelesen werden, liest der Controller mehrere Datenstücke, die die zu lesenden Daten enthalten, und Paritätsdaten für die mehreren Datenstücke von allen Festplattenlaufwerken einer RAID-Gruppe, zu der das Festplattenlaufwerk gehört, das die zu lesenden Daten speichert. So untersucht der Controller, ob die mehreren Datenstücke, die die zu lesenden Daten enthalten, mit fehlerhaftem Inhalt oder nicht in die Festplattenlaufwerke geschrieben sind.



Beschreibung

[0001] Die Erfindung betrifft ein Plattenarraysystem und ein Verfahren zum Steuern eines solchen.

[0002] In den letzten Jahren wird einhergehend mit der Zunahme von Speicherkapazität in Plattenarraysystemen deren Bedeutung in Informationsverarbeitungssystemen immer größer. Daher ist es wesentlich, Daten korrekt an einer angeforderten Position einzuschreiben und eine Inkorrektheit gelesener Daten auf eine Daten-Eingabe/Ausgabe-Anforderung von einer Informationsverarbeitungsvorrichtung oder dergleichen zu erkennen.

Stand der Technik

[0003] JP-A-5-150909 offenbart ein Verfahren, bei dem in einer Magnetplatteneinheit zwei Köpfe vorhanden sind und von den zwei Köpfen gelesene identische Daten miteinander verglichen werden, um die Zuverlässigkeit beim Schreiben und Lesen in der Magnetplatteneinheit zu verbessern.

[0004] Wenn das in JP-A-5-150909 offenbarte Verfahren bei einem Plattenarraysystem angewandt wird, müssen in jeder Magnetplatteneinheit zwei Köpfe angebracht werden. Demgemäß steigen die Einheitskosten zum Herstellen jedes Festplattenlaufwerks an. Daher ist ein Verfahren zum Verbessern der Zuverlässigkeit bei einem Festplattenlaufwerk ohne Änderung des körperlichen Aufbaus desselben, z.B. ohne Hinzufügens irgendeines Kopfs zu ihm, gefordert.

[0005] Außerdem wurden bei Plattenarraysystemen Festplattenlaufwerke gemäß seriell-ATA oder parallel-ATA sowie Faserkanal-Festplattenlaufwerke in Betrieb genommen. Dies, da Festplattenlaufwerke gemäß seriell-ATA oder parallel-ATA hinsichtlich der Zuverlässigkeit tatsächlich Faserkanal-Festplattenlaufwerken unterlegen sind, sie jedoch billiger sind. Daher besteht Bedarf an der Entwicklung eines Verfahrens zum Verbessern der Zuverlässigkeit bei anderen Festplattenlaufwerken als solchen mit Faserkanal innerhalb eines Plattenarraysystems, das aus einer Kombination von Faserkanal-Festplattenlaufwerken und anderen Festplattenlaufwerken, die dem Seriell-ATA-Standard oder dergleichen genügen, besteht.

Aufgabenstellung

ZUSAMMENFASSUNG DER ERFINDUNG

[0006] Die Erfindung wurde unter Berücksichtigung der vorigen Probleme entwickelt. Es ist eine Aufgabe der Erfindung, ein Plattenarraysystem und ein Verfahren zum Steuern eines solchen zu schaffen.

[0007] Um die vorstehende Aufgabe zu lösen, verfügt ein Plattenarraysystem gemäß einer Hauptkonfiguration der Erfindung über ein erstes Gehäuse, ein zweites Gehäuse und einen Controller. Das erste Gehäuse beherbergt eine oder mehrere RAID-Gruppen. Jede RAID-Gruppe besteht aus mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem ersten Schnittstellenstandard. Die Festplattenlaufwerke sind über einen Kommunikationspfad miteinander verbunden. Das zweite Gehäuse beherbergt eine oder mehrere RAID-Gruppen. Jede RAID-Gruppe besteht aus mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem zweiten Schnittstellenstandard. Die Festplattenlaufwerke sind über den Kommunikationspfad mittels mehrerer Wandeinheiten zum Wandeln des ersten und des zweiten Schnittstellenstandards ineinander verbunden. Die dem zweiten Schnittstellenstandard genügenden Festplattenlaufwerke verfügen über geringere Zuverlässigkeit als die dem ersten Schnittstellenstandard genügenden Festplattenlaufwerke. Der Controller verfügt über einen Kanal-Steuerabschnitt, einen Platten-Steuerabschnitt, einen Cachespeicher und eine CPU. Der Kanal-Steuerabschnitt ist mit einer Informationsverarbeitungsvorrichtung verbunden, um mit dieser in Kommunikation treten zu können. Der Kanal-Steuerabschnitt empfängt Anforderungen von der Informationsverarbeitungsvorrichtung. Die Anforderung beinhaltet eine Leseanforderung zum Lesen von Daten von den Festplattenlaufwerken im ersten oder zweiten Gehäuse sowie eine Schreibanforderung zum Schreiben von Daten in die Festplattenlaufwerke im ersten oder zweiten Gehäuse. Der Platten-Steuerabschnitt ist über den Kommunikationspfad mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse verbunden, um mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse in Kommunikation treten zu können. Der Platten-Steuerabschnitt führt eine Eingabe/Ausgabe von Daten und Paritätsdaten von/in die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse entsprechend der vom Kanal-Steuerabschnitt empfangenen Leseanforderung oder Schreibanforderung aus. Die Paritätsdaten sind Daten zum Erkennen von Fehlern über mehrere Datenstücke hinweg, einschließlich der Daten von/an die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse. Der Cachespeicher führt eine Zwischenspeicherung von Daten aus, die in die mehreren Festplattenlaufwerke zu schreiben sind. Die CPU überwacht die Steuerung über dem Kanal-Steuerabschnitt und den Platten-Steuerabschnitt. Der Controller liest mehrere Datenstücke einschließlich Daten, wie sie in den mehreren Festplattenlaufwerken im zweiten Gehäuse gespeichert sind, und Paritätsdaten für die mehreren Datenstücke aus allen Festplattenlaufwerken der RAID-Gruppe, zu der die die Daten speichernden Festplattenlaufwerk gehören, und er untersucht, ob die die Daten enthaltenden mehreren Datenstücke mit fehlerhaftem Inhalt

oder nicht in die Festplattenlaufwerke geschrieben wurden.

[0008] Außerdem verstellt der Controller beim Schreiben von Daten in eines der Festplattenlaufwerke im zweiten Gehäuse entsprechend der Schreibanforderung von der Informationsverarbeitungsvorrichtung einen zum Festplattenlaufwerk gehörenden Kopf ausgehend von einer Position, an der die Daten abgespeichert wurden. Danach liest der Controller die Daten von einer zum Festplattenlaufwerk gehörenden Magnetplatte sowie aus dem Cachespeicher, und er vergleicht die zwei ausgelesenen Datenstücke.

[0009] Außerdem bildet der Controller beim Empfangen der Schreibanforderung von der Informationsverarbeitungsvorrichtung zum Einschreiben von Daten in eines der Festplattenlaufwerke im zweiten Gehäuse aus den aus mehreren Sektoren bestehenden Daten eine Dateneinheit auf Grundlage der zu schreibenden Daten sowie Paritätsdaten zum Erkennen von Datenfehlern in den mehreren Sektoren, und er schreibt die Dateneinheit in das Festplattenlaufwerk. Wenn die Steuerung die Leseanforderung zum Lesen der Daten von der Informationsverarbeitungsvorrichtung empfängt, liest sie die Dateneinheit, und sie untersucht, ob die Daten mit fehlerhaftem Inhalt oder nicht im Festplattenlaufwerk gespeichert sind. Hierbei ist der erste Schnittstellenstandard z.B. ein Faserkanalstandard.

[0010] Der zweite Schnittstellenstandard ist z.B. ein Seriell-ATA-Standard. Der Kommunikationspfad ist z.B. ein FC-AL (Fiber Channel-Arbitrated Loop). Außerdem ist jede Wandlungseinheit z.B. ein Wandler zum Wandeln eines Faserkanalprotokolls und eines Seriell-ATA-Protokolls ineinander. Außerdem dient jede RAID-Gruppe zum Verwalten mehrerer Festplattenlaufwerke als eine Gruppe, wenn die Festplattenlaufwerke über RAID-Konfiguration verfügen. Für jede RAID-Gruppe werden logische Datenbereiche gebildet, die als Zugriffseinheiten ausgehend von der Informationsverarbeitungsvorrichtung dienen. Jedem logischen Datenbereich wird eine als LUN bezeichnete Kennung zugewiesen. Wenn von der Informationsverarbeitungsvorrichtung eine Schreibanforderung zum Schreiben von Daten in einen logischen Datenbereich empfangen wird, schreibt der Platten-Steuerabschnitt die Daten und Paritätsdaten zum Erkennen von Fehlern in den Daten in die die RAID-Gruppe bildenden Festplattenlaufwerke.

[0011] Demgemäß können gemäß der Erfindung ein Plattenarraysystem und ein Verfahren zum Steuern eines solchen geschaffen werden.

Ausführungsbeispiel

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0012] Die in dieser Beschreibung offenbarten Probleme und Lösungen für diese werden aus der folgenden detaillierten Beschreibung bevorzugter Ausführungsformen in Verbindung mit den beigefügten Zeichnungen besser ersichtlich werden.

[0013] [Fig. 1A](#) und [Fig. 1B](#) sind Diagramme, die das Aussehen eines Plattenarraysystems gemäß einer Ausführungsform der Erfindung zeigen;

[0014] [Fig. 2A](#) und [Fig. 2B](#) sind Diagramme, die die Konfiguration eines Mastergehäuses des Plattenarraysystems gemäß der Ausführungsform zeigen;

[0015] [Fig. 3A](#) und [Fig. 3B](#) sind Diagramme, die die Konfiguration eines Erweiterungsgehäuses des Plattenarraysystems gemäß der Ausführungsform zeigen;

[0016] [Fig. 4](#) ist ein Diagramm, das die Konfiguration eines Festplattenlaufwerks gemäß der Ausführungsform zeigt;

[0017] [Fig. 5](#) ist ein Diagramm, das die Konfiguration des Plattenarraysystems gemäß der Ausführungsform zeigt;

[0018] [Fig. 6](#) ist ein Diagramm, das den Zustand zeigt, gemäß dem durch eine CPU eines Controllers auszuführende Mikroprogramme in einen Speicher gemäß der Ausführungsform eingespeichert wurden;

[0019] [Fig. 7](#) ist ein Diagramm, das einen Modus zum Verbinden von Faserkanal-Festplattenlaufwerken mit einem Platten-Steuerabschnitt des Controllers gemäß der Ausführungsform zeigt;

[0020] [Fig. 8](#) ist ein Diagramm, das einen ersten Modus zum Verbinden von Seriell-ATA-Festplattenlaufwerken mit dem Platten-Steuerabschnitt des Controllers gemäß der Ausführungsform zeigt;

[0021] [Fig. 9](#) ist ein Diagramm, das einen zweiten Modus zum Verbinden von Seriell-ATA-Festplattenlaufwerken mit dem Platten-Steuerabschnitt des Controllers gemäß der Ausführungsform zeigt;

[0022] [Fig. 10](#) ist ein Diagramm, das ein Beispiel zeigt, bei dem Daten in Festplattenlaufwerke geschrieben werden, die eine RAID-Gruppe gemäß der Ausführungsform bilden;

[0023] [Fig. 11](#) ist ein Diagramm, das eine Aktualisierungskontrolltabelle gemäß der Ausführungsform zeigt;

[0024] [Fig. 12](#) ist ein Flussdiagramm zum Vergleichen von in einem Cachespeicher abgespeicherten Daten mit in einer Magnetplatte abgespeicherten Daten, wenn Daten gemäß der Ausführungsform geschrieben werden;

[0025] [Fig. 13](#) ist ein Flussdiagramm zum Vergleichen von in einem Cachespeicher abgespeicherten Daten mit in einer Magnetplatte abgespeicherten Daten unter Berücksichtigung der Datengröße, wenn Daten gemäß der Ausführungsform geschrieben werden;

[0026] [Fig. 14](#) ist ein Flussdiagramm zum Vergleichen von in einem Cachespeicher abgespeicherten Daten mit in einer Magnetplatte abgespeicherten Daten, wenn im Cachespeicher abgespeicherte Daten in die Magnetplatte gemäß der Ausführungsform geschrieben werden;

[0027] [Fig. 15](#) ist ein Diagramm, das eine Kopfprüf-Kontrolltabelle gemäß der Ausführungsform zeigt;

[0028] [Fig. 16](#) ist ein Flussdiagramm für eine Kopfprüfung, wie sie gemäß der Ausführungsform periodisch auszuführen ist;

[0029] [Fig. 17](#) ist ein Flussdiagramm zum Ausführen der Kopfprüfung, wenn Daten gemäß der Ausführungsform gelesen werden;

[0030] [Fig. 18](#) ist ein Diagramm, das ein Beispiel zeigt, bei dem eine Dateneinheit in ein Festplattenlaufwerk gemäß der Ausführungsform geschrieben wurde;

[0031] [Fig. 19](#) ist ein Diagramm, das ein Beispiel zeigt, bei dem jede Dateneinheit über mehrere Plattenlaufwerke gemäß der Ausführungsform eingeschrieben und verteilt wurde;

[0032] [Fig. 20](#) ist ein Diagramm, das eine Dateneinheit-Kontrolltabelle gemäß der Ausführungsform zeigt; und

[0033] [Fig. 21](#) ist ein Diagramm, das die Konfiguration eines Plattenarraysystems zeigt, bei dem Faserkanal-Festplattenlaufwerke in einem ersten Gehäuse untergebracht sind und Seriell-ATA-Festplattenlaufwerke in einem zweiten Gehäuse gemäß der Ausführungsform untergebracht sind.

DETAILLIERTE BESCHREIBUNG DER AUSFÜHRUNGSFORMEN

Systemkonfiguration

[0034] Die [Fig. 1A](#) ist eine Vorderansicht eines Plattenarraysystems **10**, das als eine Ausführungsform

der Erfindung beschrieben wird. Die [Fig. 1B](#) ist eine Rückansicht des Plattenarraysystems **10**. Die [Fig. 2A](#) ist eine perspektivische Ansicht eines Mastergehäuses **20**, das am Plattenarraysystem **10** anzubringen ist, das am Plattenarraysystem **10** anzubringen ist, wobei man das Mastergehäuse **20** von seiner Vorderseite her sieht. Die [Fig. 2B](#) ist eine perspektivische Ansicht des Mastergehäuses **20** von seiner Rückseite her gesehen. Die [Fig. 3A](#) ist eine perspektivische Ansicht eines Erweiterungsgehäuses **30**, das am Plattenarraysystem **10** anzubringen ist, wobei man das Erweiterungsgehäuse **30** von seiner Vorderseite her sieht. Die [Fig. 3B](#) ist eine perspektivische Ansicht des Erweiterungsgehäuses **30** von seiner Rückseite her gesehen.

[0035] Wie es in den [Fig. 1A](#) und [Fig. 1B](#) dargestellt ist, wird das Plattenarraysystem **10** unter Verwendung eines Rackrahmens **11** als Basis hergestellt. Montagerahmen **12** sind in mehreren Etagen ausgebildet, die oben und unten an der linken und rechten Innenseitenfläche des Rackrahmens **11** so angeordnet sind, dass sie sich in den Richtungen nach vorne/nach hinten erstrecken. Das Mastergehäuse **20** und das Erweiterungsgehäuse **30** sind ausziehbar entlang den Montagerahmen **12** montiert. Wie es in den [Fig. 2A](#) und [Fig. 2B](#) dargestellt ist, sind Platten oder Einheiten zum Bereitstellen verschiedener Funktionen für das Plattenarraysystem **10** am Mastergehäuse **20** und am Erweiterungsgehäuse **30** angebracht.

[0036] Wie es in der [Fig. 2A](#) dargestellt ist, sind mehrere Plattenlaufwerkeinheiten **52**, in die jeweils ein Festplattenlaufwerk **51** eingebracht ist, parallel in die vordere, obere Etage des Mastergehäuses **20** eingesetzt. Jedes Festplattenlaufwerk **51** ist ein Festplattenlaufwerk mit einer Kommunikationsschnittstelle zum Bereitstellen einer Kommunikationsfunktion entsprechend einem FC-AL-Standard, einem SCSI1 (Small Computer System Interface 1)-Standard, einem SCSI2-Standard, einem SCSI3-Standard, einem Parallel-ATA (AT Attachment)-Standard, einem Seriell-ATA-Standard, oder dergleichen.

[0037] In die vordere, untere Etage des Mastergehäuses **20** sind eine Batterieeinheit **53**, eine Anzeigetafel **54** zum Anzeigen von Betriebszuständen usw. der Festplattenlaufwerke **51** sowie ein Diskettenlaufwerk **55** eingesetzt. Die Batterieeinheit **53** verfügt über eine Sekundärbatterie. Die Batterieeinheit **53** hat eine Funktion als Ersatzspannungsversorgung zum Liefern von Spannung an Leiterplatten oder Einheiten, wenn die Spannungsversorgung von einer AC/DC-Spannungsversorgung **57** aufgrund eines Netzausfalls oder dergleichen stoppt. Die Anzeigetafel **54** ist mit Anzeigevorrichtungen wie LED-Lampen oder dergleichen zum Anzeigen der Betriebszustände usw. der Festplattenlaufwerke **51** versehen. Das Diskettenlaufwerk **55** wird z.B. zum Laden eines War-

tungsprogramms verwendet.

[0038] Wie es in der [Fig. 2B](#) dargestellt ist, sind Spannungsversorger-Controllerplatten **56** an den entgegengesetzten Seitenflächen in der hinteren, oberen Etage des Mastergehäuses **20** einzeln eingesetzt. Jede Spannungsversorgungs-Controllerplatte **56** ist mit mehreren Festplattenlaufwerken **51** so verbunden, dass sie mit diesen Kommunikationsvorgänge errichten kann. Z.B. sind die Spannungsversorgung-Controllerplatte **56** und die mehreren Festplattenlaufwerke **51** so verbunden, dass für Kommunikationsvorgänge über einen schleifenartigen Kommunikationspfad wie einen solchen, der für Kommunikation in einem FC/AL-System (Topologie) sorgt, gesorgt werden kann.

[0039] Jede Spannungsversorgung-Controllerplatte **56** ist mit Schaltkreisen zum Überwachen des Zustands der AC/DC-Spannungsversorgung **57**, zum Überwachen der Zustände der Festplattenlaufwerke **51**, zum Steuern der Spannungsversorgung der Festplattenlaufwerke **51**, zum Steuern der Kühlkapazität einer Kühleinheit, zum Steuern der Anzeigevorrichtungen auf der Anzeigetafel **54**, zum Überwachen der Temperatur jedes Abschnitts des Gehäuses, usw. versehen. Übrigens ist die Kühleinheit eine Einheit zum Kühlen des inneren des Plattenarraysystems **10** oder des Inneren des Gehäuses **20** oder **30**. Z.B. ist die Kühleinheit ein Interkühler, eine Wärmesenke, ein Kühllüfter vom Luftkühltyp oder dergleichen. Die Spannungsversorgung-Controllerplatte **56** ist mit einem Faserkanalkabel-Verbinder **67** versehen, mit dem ein Faserkanalkabel **91** verbunden ist.

[0040] Wie es in der [Fig. 2B](#) dargestellt ist, sind zwei AC/DC-Spannungsversorgungen **57** parallel in einem Raum zwischen den zwei Spannungsversorgung-Controllerplatten **56** in der hinteren, oberen Etage des Mastergehäuses **20** angebracht. Jede AC/DC-Spannungsversorgung **57** liefert Spannung an die Festplattenlaufwerke **51**, die Leiterplatten, die Einheiten usw. Die AC/DC-Spannungsversorgung **57** ist mit der Spannungsversorgung-Controllerplatte **56** verbunden, und sie ist so aufgebaut, dass sie Spannung entsprechend einem Signal von der Spannungsversorgung-Controllerplatte **56** an jedes Festplattenlaufwerk **51** liefern kann.

[0041] Übrigens ist diese Ausführungsform so konzipiert, dass zwei Spannungsversorgung-Controllerplatten **56** und zwei AC/DC-Spannungsversorgungen **57** redundant sowohl im Mastergehäuse **20** als auch im Erweiterungsgehäuse **30** vorhanden sind, um für Sicherheit hinsichtlich der Spannungsversorgung an jedes Gehäuse **20**, **30** zu sorgen. Jedoch können in jedem Gehäuse **20**, **30** eine Spannungsversorgung-Controllerplatte **56** und eine AC/DC-Spannungsversorgung **57** montiert sein.

[0042] Jede AC/DC-Spannungsversorgung **57** ist mit einem Unterbrechungsschalter **64** zum Ein-/Aus-schalten des Ausgangssignals der AC/DC-Spannungsversorgung **57** versehen.

[0043] Wie es in der [Fig. 2B](#) dargestellt ist, sind zwei Luftkühlungs-Lüftereinheiten **58** parallel unter den AC/DC-Spannungsversorgungen **57** angebracht. Jede Kühlungslüftereinheit **58** ist mit einem oder mehreren Kühllüftern **66** versehen. Die Kühllüfter **66** ermöglichen es, dass Luft in das Gehäuse einströmt/aus ihm ausströmt, um dadurch Wärme, wie sie von den Festplattenlaufwerken **51**, den AC/DC-Spannungsversorgungen **57** usw. erzeugt wird, zur Außenseite des Gehäuses abzugeben. Übrigens sind im Gehäuse **20**, **30** Belüftungspfade oder Belüftungslöcher zum Umwälzen der Luft durch jedes Mastergehäuse **20**/Erweiterungsgehäuse **30** und daran montierte Leiterplatten oder Einheiten ausgebildet, um einen Mechanismus zum effizienten Ausgeben von Wärme im Gehäuse **20**, **30** nach außen mittels der Kühllüfter **66** zu bilden. Obwohl Kühllüfter **66** für jedes Festplattenlaufwerk **51** vorhanden sein können, ist es bevorzugt, dass für jedes Gehäuse größere Kühllüfter **66** vorhanden sind, so dass die Anzahl der Chips oder Einheiten verringert werden kann.

[0044] Jede Kühllüftereinheit **58** ist mit einer Controllerplatte **59** oder der Spannungsversorgung-Controllerplatte **56** über eine Steuerleitung verbunden. Die Drehzahl jedes Kühllüfters **66** der Kühllüftereinheit **58** wird durch die Controllerplatte **59** oder die Spannungsversorgung-Controllerplatte **56** über die Steuerleitung gesteuert.

[0045] Wie es in der [Fig. 6B](#) dargestellt ist, ist eine Controllerplatte **59** in die hintere, untere Etage des Mastergehäuses **20** eingesetzt. Die Controllerplatte **59** ist mit einer Kommunikationsschnittstelle zu den im Mastergehäuse **20** und den Erweiterungsgehäusen **30** angebrachten Festplattenlaufwerken **51**, Schaltkreisen zum Steuern der Betriebsabläufe der Festplattenlaufwerke **51** (z.B. zum Steuern derselben in einem RAID-System) oder zum Überwachen der Zustände der Festplattenlaufwerke **51**, usw., versehen.

[0046] Übrigens steuern zwar bei dieser Ausführungsform die Spannungsversorgung-Controllerplatten **56** die Spannungsversorgung für die Festplattenlaufwerke **51** oder die Kühlkapazität der Kühleinheiten, jedoch kann die Controllerplatte **59** eine derartige Steuerung ausführen.

[0047] Diese Ausführungsform verwendet einen Modus, bei dem die Controllerplatte **59** mit einer Kommunikations-Schnittstellenkarte **51** zum Bereitstellen einer Funktion einer Kommunikationsschnittstelle zu einer Informationsverarbeitungsvorrichtung

300, z.B. einer Kommunikationsfunktion entsprechend einem SCIS-Standard oder einem Faserkanalstandard, einem Cachespeicher **62** zum Speichern von in die Festplattenlaufwerke **51** zu speichernden Daten oder von aus ihnen auszulesenden Daten, usw., versehen ist. Jedoch können diese Funktionen auf einer anderen Leiterplatte untergebracht sein.

[0048] Die an der Controllerplatte **59** angebrachte Kommunikations-Schnittstellenplatte **61** ist mit einem externen Verbinder **63** zum Herstellen einer Verbindung zur Informationsverarbeitungsvorrichtung **300** versehen. Der externe Verbinder **63** genügt einem vorbestimmten Schnittstellenstandard wie SAN (Storage Area Network), LAN (Local Area Network) oder SCSI, der durch ein Protokoll wie Faserkanal, Ethernet (registrierte Handelsbezeichnung) oder dergleichen gebildet ist. Das Plattenarraysystem **10** ist über ein mit dem Verbinder **63** verbundenes Kommunikationskabel **92** mit der Informationsverarbeitungsvorrichtung **300** verbunden.

[0049] Übrigens können zwei Controllerplatten **59** redundant im Mastergehäuse **20** angebracht sein, um für Sicherheit hinsichtlich der Steuerung der Festplattenlaufwerke **51** zu sorgen.

[0050] Wie es in der [Fig. 3A](#) dargestellt ist, sind mehrere Plattenlaufwerkseinheiten **52**, die Festplattenlaufwerke **51** aufnehmen, parallel an der Vorderseite jedes Erweiterungsgehäuses **30** angebracht. Wie es in der [Fig. 3B](#) dargestellt ist, sind Spannungsversorgung-Controllerplatten **56** einzeln an den hinteren, entgegengesetzten Seitenflächen des Erweiterungsgehäuses **30** eingesetzt. Außerdem sind zwei AC/DC-Spannungsversorgungen **57** parallel in einem Raum zwischen den zwei Spannungsversorgung-Controllerplatten **56** angebracht. Zusätzlich sind zwei Kühllüftereinheiten **58** parallel unter den AC/DC-Spannungsversorgungen **57** angebracht. Jede AC/DC-Spannungsversorgung **57** ist mit einem Unterbrechungsschalter **64** zum Ein-/Ausschalten der Spannung der AC/DC-Spannungsversorgung **57** versehen.

[0051] Diese Ausführungsform ist so konzipiert, dass zwei Spannungsversorgung-Controllerplatten **56** und zwei AC/DC-Spannungsversorgungen **57** redundant in jedem Erweiterungsgehäuse **30** vorhanden sind, um, wie oben beschrieben, für Sicherheit hinsichtlich der Spannungsversorgung zum Erweiterungsgehäuse **30** zu sorgen. Jedoch können eine Spannungsversorgung-Controllerplatte **56** und eine AC/DC-Spannungsversorgung **57** im Erweiterungsgehäuse **30** angebracht sein. Übrigens können die Funktionen der Spannungsversorgung-Controllerplatten **56** zum Steuern der Spannungsversorgung für die Festplattenlaufwerke **51**, zum Steuern der Kühlkapazität der Kühleinheiten usw. auf der Controllerplatte **59** vorhanden sein.

[0052] Die [Fig. 4](#) zeigt ein Beispiel für die Konfiguration jedes Festplattenlaufwerks **51**, wie es in jeder Plattenlaufwerkseinheit **52** aufgenommen ist. Das Festplattenlaufwerk **51** verfügt über ein Gehäuse sowie Magnetplatten **73**, Stellglieder **71**, einen Spindelmotor **72**, Köpfe **74** zum Lesen/Schreiben von Daten, eine Mechanismussteuerschaltung **75** zum Steuern von Mechanismusteilen wie den Köpfen **74** und dergleichen, eine Signalverarbeitungsschaltung **76** zum Steuern eines Lese/Schreib-Signals für Daten von/in jede Magnetplatte **73**, eine Kommunikations-Schnittstellenschaltung **77**, einen Schnittstellenverbinder **79** zum Eingeben/Ausgeben verschiedener Befehle oder Daten über ihn sowie einen Spannungsversorgungsverbinder **80**, die alle im Gehäuse **70** angeordnet sind. Übrigens ist in der Kommunikations-Schnittstellenschaltung **77** ein Cachespeicher zum Zwischenspeichern von Daten enthalten. Übrigens wird der zum Festplattenlaufwerk **51** gehörende Cachespeicher als Plattencache bezeichnet, um ihn von einem Cachespeicher **62** in einem Controller **500**, der später beschrieben wird, zu unterscheiden.

[0053] Jedes Festplattenlaufwerk **51** ist z.B. eine Speichereinheit, die mit 3,5-Zoll-Magnetplatten vom Contact-Start-Stop(CSS)-Typ, 2,5-Zoll-Magnetplatten vom Lade/Entlade-Typ oder dergleichen versehen ist. Z.B. verfügt jede 3,5-Zoll-Magnetplatte über eine Kommunikationsschnittstelle gemäß SCSI1, SCSI2, SCSI3, FC-AL oder dergleichen. Andererseits verfügt z.B. jede 2,5-Zoll-Magnetplatte über eine Kommunikationsschnittstelle gemäß parallel-ATA, seriell-ATA oder dergleichen.

[0054] Wenn jede 2,5-Zoll-Magnetplatte im Gehäuse **20**, **30** des Plattenarraysystems **10** aufgenommen wird, kann sie in einem Behälter mit 3,5-Zoll-Form aufgenommen werden. So kann die Stoßfestigkeitsfunktion der Magnetplatten verbessert werden. Übrigens unterscheiden sich eine 2,5-Zoll-Magnetplatte und eine 3,5-Zoll-Magnetplatte nicht nur hinsichtlich der Kommunikationsschnittstelle voneinander, sondern auch hinsichtlich der I/O-Funktion, des Energieverbrauchs, der Lebensdauer usw. Eine 2,5-Zoll-Magnetplatte ist hinsichtlich der I/O-Funktion einer 3,5-Zoll-Magnetplatte unterlegen, und die Lebensdauer der ersteren ist kürzer als die der letzteren. Jedoch ist eine 2,5-Zoll-Magnetplatte einer 3,5-Zoll-Magnetplatte dahingehend überlegen, dass der Energieverbrauch der ersteren kleiner als der der letzteren ist.

Hardwarekonfiguration des Plattenarraysystems

[0055] Die [Fig. 5](#) ist ein Blockdiagramm, das die Hardwarekonfiguration des Plattenarraysystems **10** zeigt, das als eine Ausführungsform der Erfindung beschrieben wird.

[0056] Wie es in der [Fig. 5](#) dargestellt ist, sind

Informationsverarbeitungsvorrichtungen **300** über ein SAN mit dem Plattenarraysystem **10** verbunden. Die Informationsverarbeitungsvorrichtungen **300** sind z.B. PCs, Workstations, Großrechner oder dergleichen.

[0057] Das Plattenarraysystem **10** verfügt über ein Mastergehäuse **20** und eines oder mehrere Erweiterungsgehäuse **30**, wie bereits beschrieben. Bei dieser Ausführungsform verfügt das Mastergehäuse **20** über Controller **500**, Festplattenlaufwerke **51** usw. Jeder Controller **500** verfügt über Kanal-Steuerabschnitte **501**, Platten-Steuerabschnitte **502**, eine CPU **503**, einen Speicher **504**, einen Cachespeicher **62**, einen Datencontroller **505** usw. Der Controller **500** ist an der oben genannten Steuerplatte **59** angebracht. Andererseits verfügt jedes Erweiterungsgehäuse **30** über Festplattenlaufwerke **51** usw. Die Festplattenlaufwerke **51** im Mastergehäuse und in den Erweiterungsgehäusen sind über eine FC-AL **506** mit den Platten-Steuerabschnitte **502** verbunden, um für Kommunikation mit diesen sorgen zu können. Übrigens wird der Verbindungsmodus zwischen jedem Platten-Steuerabschnitt **502** und jedem Festplattenlaufwerk **51** später detailliert beschrieben.

[0058] Jeder Kanal-Steuerabschnitt **501** ist eine Schnittstelle zum Errichten von Kommunikation mit den Informationsverarbeitungsvorrichtungen **300**. Der Kanal-Steuerabschnitt **501** verfügt über eine Funktion zum Aufnehmen einer Blockzugriffsanforderung entsprechend einem Faserkanalprotokoll.

[0059] Jeder Platten-Steuerabschnitt **502** ist eine Schnittstelle zum Austauschen von Daten mit den Festplattenlaufwerken **51** entsprechend einer Anweisung von der CPU **503**. Der Platten-Steuerabschnitt **502** verfügt über eine Funktion zum Senden einer Daten-Eingabe/Ausgabe-Anforderung an die Festplattenlaufwerke **51** entsprechend einem Protokoll, in dem Befehle usw. zum Steuern der Festplattenlaufwerke **51** festgelegt sind.

[0060] Die CPU **503** verwaltet die Steuerung des Plattenarraysystems **10** insgesamt. Die CPU **503** führt im Speicher **504** abgespeicherte Mikroprogramme aus, um die Kanal-Steuerabschnitte **501**, die Platten-Steuerabschnitte **502**, den Datencontroller **505** usw. zu steuern. Zu den Mikroprogrammen gehören ein Datenleseprozess **601**, ein Datenschreibprozess **602** und dergleichen, wie es in der [Fig. 6](#) dargestellt ist.

[0061] Der Cachespeicher **62** dient zum zeitweiligen Abspeichern von Daten, wie sie zwischen jedem Kanal-Steuerabschnitt **501** und jedem Platten-Steuerabschnitt **502** auszutauschen sind.

[0062] Der Datencontroller **505** führt eine Datenübertragung zwischen jedem Kanal-Steuerabschnitt

501 und dem Cachespeicher **62** oder zwischen diesem und jedem Platten-Steuerabschnitt **502** unter Steuerung durch die CPU **503** aus.

[0063] Jeder Controller **500** verfügt über eine Funktion des Steuerns der Festplattenlaufwerke **51** auf RAID-Ebene (z.B. 0, 1 oder 5), entsprechend eine sogenannten RAID(Redundant Array of Inexpensive Disks)-System. Beim RAID-System werden mehrere Festplattenlaufwerke **51** als eine Gruppe verwaltet (nachfolgend als RAID-Gruppe bezeichnet). Logische Datenbereiche, die als Zugriffseinheiten von den Informationsverarbeitungsvorrichtungen **300** dienen, sind in jeder RAID-Gruppe ausgebildet. Eine als LUN (Logical Unit Number) bezeichnete Kennung ist jedem logischen Datenbereich zugewiesen. Im Speicher **504** ist Information zur RAID-Konfiguration in Form einer RAID-Konfigurationstabelle **603** abgespeichert, wie es in der [Fig. 6](#) dargestellt ist. Die CPU **503** nimmt auf die RAID-Konfigurationstabelle **603** Bezug, wenn sie den Datenleseprozess **601** oder den Datenschreibprozess **602** ausführt.

[0064] Übrigens muss das Plattenarraysystem z.B. nicht ein System sein, das auf die oben beschriebene Weise konzipiert ist, sondern es kann ein als NAS (Network Attached Storage) dienendes System sein, das so konzipiert ist, dass es eine Daten-Eingabe/Ausgabe-Anforderung auf Grundlage einer Datennamensspezifizierung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** entsprechend einem Protokoll wie NFS (Network File System) akzeptiert.

Verbindungsmodus von Festplattenlaufwerken

[0065] Als Nächstes erfolgt eine Beschreibung zum Verbindungsmodus zwischen jedem Controller **500** und jedem Festplattenlaufwerk **51**.

[0066] Die [Fig. 7](#) zeigt einen Verbindungsmodus zwischen jedem Platten-Steuerabschnitt **502** und jedem Faserkanal-Festplattenlaufwerk **51**, wie es im Mastergehäuse **20** untergebracht ist.

[0067] Jeder Platten-Steuerabschnitt **502** ist über eine FC-AL mit mehreren Festplattenlaufwerken **51** verbunden. Die FC-AL **506** verfügt über mehrere PBCs (Port Bypass Circuits) **701**. Die Faserkanal-Festplattenlaufwerke **51** sind jeweils über die PBCs **701** mit der FC-AL **506** verbunden. Jede PBC **701** ist ein als Chip ausgebildeter elektronischer Schalter. Die PBCs **701** verfügen auch über eine Funktion des Umgehens der Platten-Steuerabschnitte **502** oder der Festplattenlaufwerke **51**, um sie dadurch elektrisch von der FC-AL **506** auszuschließen. Genauer gesagt, trennen die PBCs **701**, wenn in irgendwelchen Festplattenlaufwerke **51** Fehler auftreten, die Festplattenlaufwerke **51** von der FC-AL **506**, so dass für Kommunikationsvorgänge zwischen einem belie-

bigen anderen Festplattenlaufwerk **51** und jedem Platten-Steuerabschnitt **502** gesorgt werden kann.

[0068] Außerdem ermöglichen es die PBCs **701**, die Festplattenlaufwerke **51** zu entnehmen und einzusetzen, während der Betrieb der FC-AL **506** erhalten bleibt. Wenn z.B. ein neues Festplattenlaufwerk **51** eingesetzt wird, wird es in die FC-AL **506** eingebaut, damit zwischen dem Festplattenlaufwerk **51** und dem Platten-Steuerabschnitt **502** für Kommunikation gesorgt werden kann. Übrigens kann eine Leiterplatte der PBCs **701** am Rackrahmen **11** des Plattenarraysystems **10** vorhanden sein, oder sie kann teilweise oder ganz an der Controllerplatte **59** oder der Spannungsversorgung-Controllerplatte **56** angebracht sein.

[0069] Die [Fig. 8](#) zeigt einen Verbindungsmodus zwischen jedem Platten-Steuerabschnitt **502** und jedem im Mastergehäuse **20** untergebrachten Seriell-ATA-Festplattenlaufwerk **51**.

[0070] Jedes Festplattenlaufwerk **51** ist über einen Wandler **801** mit PBCs **602** der FC-AL **506** verbunden. Der Wandler **801** ist eine Schaltung zum Wandeln eines Faserkanalprotokolls und eines Seriell-ATA-Protokolls ineinander. Der Wandler **801** besteht aus einem Chip, in den eine Protokollwandelfunktion eingebaut ist. Der Wandler **801** ist in jeder Plattenlaufwerkeinheit **52** enthalten.

[0071] Die [Fig. 9](#) entspricht einem anderen Verbindungsmodus, bei dem Seriell-ATA-Festplattenlaufwerke **51** im Mastergehäuse **20** untergebracht sind.

[0072] Jeder Wandler **901** ist eine Schaltung zum Wandeln eines Faserkanalprotokolls und eines Seriell-ATA-Protokolls ineinander, auf dieselbe Weise wie beim Wandler **801** in der [Fig. 8](#). Der Wandler **901** ist mit einer PBC **602** einer FC-AL **506** verbunden. Mehrere Festplattenlaufwerke **51** sind über jeweilige Schalter **902** mit jedem Wandler **901** verbunden. Die Schalter **902** sind Schaltungen zum Auswählen eines Festplattenlaufwerks **51**, mit dem für Kommunikation gesorgt werden soll, wenn die Festplattenlaufwerke **51** mit mehreren Wandlern **901** verbunden sind. Jeder Schalter **902** ist jeder Plattenlaufwerkeinheit **52** vorhanden. Jeder Wandler **901** besteht aus einem Chip oder mehreren Schaltkreisen, in die eine Protokollwandelfunktion eingebaut ist. Z.B. kann der Wandler **901** durch die Konfiguration einer SATA-Mastervorrichtung implementiert sein, wie sie in der "US-Patentanmeldung mit der Veröffentlichungs-Nr. 2003/0135577" offenbart ist. Der Wandler **901** ist auf der Controllerplatte **59**, der Spannungsversorgung-Controllerplatte **56** oder dergleichen angebracht.

Steuerung zum Verbessern der Zuverlässigkeit

[0073] Nun erfolgt eine Beschreibung zu einem Verfahren zum Verbessern der Zuverlässigkeit beim Lesen von den Festplattenlaufwerken oder beim Schreiben auf diese im oben beschriebenen Plattenarraysystem **10**.

Paritätsprüfung bei RAID-Konfiguration

[0074] Als Erstes erfolgt eine Beschreibung zu einem Verfahren zum Untersuchen, ob in einem Festplattenlaufwerk **51** von RAID-Konfiguration abgespeicherte Daten in einem falschen Zustand vorliegen oder nicht. Hierbei bedeutet der falsche Zustand einen solchen, bei dem Daten nicht an einen Ort geschrieben sind, wie er durch den Platten-Steuerabschnitt **502** spezifiziert wurde, und nicht mit entsprechend spezifiziertem Inhalt.

[0075] Die [Fig. 10](#) zeigt einen Zustand, bei dem Daten in Festplattenlaufwerke **51** gemäß RAID-5 gespeichert sind. Bei RAID-5 besteht eine RAID-Gruppe **1001** aus mehreren Festplattenlaufwerken **51**. Beim Beispiel der [Fig. 10](#) sind Daten A-D und Paritätsdaten P (A - D) zum Erkennen von Fehlern in den Daten A - D in den Festplattenlaufwerken **51** gespeichert. Auf dieselbe Weise sind Daten E - H und Paritätsdaten P (E - H) für die Daten E - H abgespeichert. Eine derartige Kombination von Daten und Paritätsdaten wird als Streifengruppe **1002** bezeichnet. Bei einer RAID-Konfiguration mit derartigen darin ausgebildeten Streifengruppen **1002** kann der alle Daten und Paritätsdaten der Streifengruppe **1002** lesende Controller **500** untersuchen, ob sich die Daten in einem falschen Zustand befinden oder nicht. Als Erstes liest der Platten-Steuerabschnitt **502**, entsprechend einer Anweisung von der CPU **503**, die Daten A - D und die Paritätsdaten P (A - D). Als Nächstes führt die CPU **503** eine Paritätsprüfung unter Verwendung der Daten A - D und der Parität P (A - D) aus. So kann untersucht werden, ob irgendein Stück der Daten A - D in einem falschen Zustand vorliegt oder nicht.

[0076] Wenn eine Datenleseanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** empfangen wird, kann der Controller **500** alle Daten und Paritätsdaten in einer Streifengruppe einschließlich der zu lesenden Daten lesen. So kann verhindert werden, dass der Controller **500** falsche Daten von den Festplattenlaufwerken **51** liest und solche falsche Daten an die Informationsverarbeitungsvorrichtung **300** sendet. Übrigens kann die Prüfung auf falsche Daten nicht nur beim Empfangen einer Datenleseanforderung sondern zu jedem beliebigen gewünschten Zeitpunkt erfolgen. Auf solche Weise kann die Erkennung falscher Daten ausgeführt werden, ohne dass sich irgendein Einfluss auf die Datenlesefunktion ergäbe.

[0077] Außerdem ist es unter Verwendung einer Aktualisierungskontrolltabelle **1101**, wie sie in der [Fig. 11](#) dargestellt ist, möglich, zu untersuchen, ob sich die in die Festplattenlaufwerke **51** eingeschriebenen Daten in einem falschen Zustand befinden oder nicht. Die Aktualisierungskontrolltabelle **1101** besteht aus Laufwerksnummern und Sektornummern, und sie ist im Speicher **504** abgespeichert. Bei dieser Ausführungsform sind die Sektornummern durch die LBA (Logical Block Address) definiert, und sie werden mit Einheiten von 128 LBA verwaltet, wie es durch LBA #1-128 veranschaulicht ist. Übrigens besteht für die Packungseinheit der Sektornummern keine Beschränkung auf 128, sondern es kann sich um jede beliebige gewünschte Einheit handeln. Wenn Daten über den Platten-Steuerabschnitt **502** in ein Festplattenlaufwerk **51** eingeschrieben werden, ändert die CPU **503** den Sektorwert des dem Schreibvorgang unterliegenden Festplattenlaufwerks **51** in der Aktualisierungskontrolltabelle **1101** auf "1". Die CPU **503** liest, über den Platten-Steuerabschnitt **502**, alle Daten und Paritätsdaten einer Streifengruppe einschließlich des Zielsektors des Festplattenlaufwerks **51**, wie als "1" in der Aktualisierungskontrolltabelle **1101** abgespeichert, und sie führt eine Paritätsprüfung aus. Wenn die gelesenen Daten nicht falsch sind, ändert die CPU **503** den Wert des Sektors in der Aktualisierungskontrolltabelle **1101** auf "0". Wenn die CPU **503** eine Datenleseanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** über den Kanal-Steuerabschnitt **501** empfängt, nimmt sie auf die Aktualisierungskontrolltabelle **1101** Bezug, und sie klärt, ob ein die zu lesenden Daten speichernde Sektor untersucht wurde oder nicht. Wenn der die Daten speichernde Sektor nicht untersucht wurde, untersucht die CPU **503** Daten einer die zu lesenden Daten enthaltenden Streifengruppe entsprechend der oben genannten Prozedur. Auf solche Weise wird eine Untersuchung an den in jedes Festplattenlaufwerk **51** eingeschriebenen Daten vor dem Empfang einer Leseanforderung zum Lesen der Daten ausgeführt. So kann verhindert werden, dass sich die Datenlesefunktion verschlechtert. Außerdem wird der noch fehlende Abschluss der Prüfung in der Aktualisierungskontrolltabelle **1101** abgespeichert, und es wird eine Paritätsprüfung ausgeführt, wenn nicht untersuchte Daten gelesen werden. So kann verhindert werden, dass falsche Daten gelesen werden.

Untersuchung von Schreibdaten

[0078] Als Nächstes erfolgt eine Beschreibung zu einem Verfahren zum Untersuchen, ob Daten korrekt geschrieben werden oder nicht, wenn sie in ein Festplattenlaufwerk **51** geschrieben werden.

[0079] Die [Fig. 12](#) ist ein Flussdiagramm, das die Kontrolle durch die CPU **503** zeigt, wenn der Controller **500** Daten in ein Festplattenlaufwerk **51** schreibt. Wenn die CPU **503** eine Datenschreibanforderung

von irgendeiner der Informationsverarbeitungsvorrichtungen **300** über den Kanal-Steuerabschnitt **501** empfängt, sendet sie eine Anweisung zum Schreiben der Daten in ein Festplattenlaufwerk **51** an den Platten-Steuerabschnitt **502** (S1201). Dann sendet die CPU **503** eine Anweisung zum Ausführen eines Suchprozesses zum Verstellen der Position eines Kopfs einer Magnetplatte, wenn die Daten geschrieben wurden, an den Platten-Steuerabschnitt (S1202). Als Nächstes liest die CPU **503** die Daten aus dem Cachespeicher **62** (S1203), und sie liest die Daten von der Magnetplatte (S1204). Die CPU **503** vergleicht die Daten vom Cachespeicher **62** und diejenigen von der Magnetplatte, um zu untersuchen, ob sie miteinander übereinstimmen (S1205). Wenn die zwei Datenstücke nicht miteinander übereinstimmen, informiert die CPU **503** die Informationsverarbeitungsvorrichtung **300** über die Tatsache, dass der Schreibvorgang nicht normal ausgeführt wurde (S1206).

[0080] Wenn die in der Magnetplatte abgespeicherten Daten auf diese Weise mit den im Cachespeicher **62** abgespeicherten Daten verglichen werden, ist es möglich, zu klären, ob die Daten korrekt in die Magnetplatte geschrieben wurden oder nicht. Außerdem überleben die Daten selbst dann, wenn sich die geschriebenen Daten in einem falschen Zustand befinden, im Cachespeicher **62**. So besteht keine Gefahr, dass die Daten verloren gehen. Übrigens ist es dann, wenn ein zu einem Festplattenlaufwerk mit einer Magnetplatte gehörender Kopf durch einen Suchprozess oder dergleichen verstellt wird, bevor zu vergleichende Daten von der Magnetplatte und aus dem Cachespeicher **62** gelesen wurden, möglich, zu verhindern, dass der Kopf zweimal an ein und derselben Position liest, wenn die Position beim Schreiben falsch ist.

[0081] Bei der Verarbeitung gemäß der [Fig. 12](#) werden alle geschriebenen Daten aus dem Cachespeicher **62** und von der Magnetplatte gelesen, und die zwei Datenstücke werden zum Untersuchen der Daten verglichen. Jedoch müssen nicht alle Daten gelesen und verglichen werden, sondern es reicht, einen Teil der Daten, z.B. ein Segment am Kopf der Daten und ein Segment am Ende derselben zu lesen und zu vergleichen. Z.B. werden umfangreiche Daten (sequenzielle Daten) häufig in Seriell-ATA-Festplattenlaufwerke eingeschrieben, da sie für Anwendungen wie Datenbackup verwendet werden. In einem derartigen Fall wird das Funktionsvermögen bei der Schreibverarbeitung auffällig beeinträchtigt, wenn in einer Magnetplatte gespeicherte Daten für alle geschriebenen Daten mit im Cachespeicher **62** gespeicherten Daten verglichen werden. Außerdem ist es dann, wenn hinsichtlich der Schreibposition oder dergleichen ein Fehler auftritt, wenn sequenzielle Daten geschrieben werden, hoch wahrscheinlich, dass alle Daten falsch sind. Daher kann in den meisten Fällen

die Beurteilung dahingehend, ob die Daten falsch sind oder nicht, dadurch erfolgen, dass ein Teil der Daten untersucht wird. D.h., dass dann, wenn für einen Teil geschriebener Daten, z.B. ein Segment am Kopf der Daten und ein Segment am Ende der Daten, ein Vergleich ausgeführt wird, die Möglichkeit besteht, falsche Daten zu prüfen, während verhindert wird, dass das Funktionsvermögen bei der Schreibverarbeitung beeinträchtigt wird.

[0082] Alternativ kann das Verfahren zum Untersuchen von in das Festplattenlaufwerk **51** geschriebenen Daten entsprechend der Größe der Daten geändert werden. Die **Fig. 13** ist ein Flussdiagramm, das die Verarbeitung zum Ändern des Untersuchungsverfahrens abhängig davon, ob geschriebene Daten sequenzielle Daten sind oder nicht, zeigt. Die CPU **503** sendet eine Anweisung zum Schreiben von Daten in ein Festplattenlaufwerk **51** an den Platten-Steuerabschnitt **502** (S1301). Dann sendet die CPU **503** eine Anweisung zum Ausführen eines Suchprozesses zum Verstellen der Position des Kopfs der Magnetplatte, wo die Daten geschrieben wurden, an den Platten-Steuerabschnitt (S1302). Die CPU **503** beurteilt, ob die Daten sequenzielle Daten sind oder nicht (S1303). Übrigens wird die Beurteilung dahingehend, ob die Daten sequenzielle Daten sind oder nicht, auf Grundlage davon ausgeführt, ob die Größe der geschriebenen Daten eine vorbestimmte Größe erreicht oder nicht.

[0083] Wenn die Daten sequenzielle Daten sind, liest die CPU **503** ein Segment am Kopf derselben sowie ein Segment am Ende derselben vom Cachespeicher **62** und der Magnetplatte. Wenn die Daten dagegen keine sequenziellen Daten sind, liest die CPU **503** alle Daten vom Cachespeicher **62** und der Magnetplatte (S1306 und S1307). Danach vergleicht die CPU **503** die zwei Stücke gelesener Daten miteinander, und sie untersucht, ob sie miteinander übereinstimmen (S1308). Wenn sie nicht übereinstimmen, informiert die CPU **503** die Informationsverarbeitungsvorrichtung **300** über die Tatsache, dass der Schreibvorgang nicht normal ausgeführt wurde (S1309).

[0084] Wenn die geschriebene Daten sequenzielle Daten sind, erfolgt zwischen in der Magnetplatte gespeicherten Daten und im Cachespeicher **62** gespeicherten Daten auf solche Weise ein Vergleich hinsichtlich eines Teils der Daten. So ist es möglich, die Inkorrektheit von Daten zu erkennen, während ein Absinken des Funktionsvermögens bei der Schreibverarbeitung unterdrückt ist. Wenn dagegen die geschriebenen Daten keine sequenziellen Daten sind, erfolgt zwischen in der Magnetplatte gespeicherten Daten und im Cachespeicher **62** gespeicherten Daten ein Vergleich hinsichtlich aller geschriebenen Daten. So ist es möglich, eine Inkorrektheit von Daten perfekt zu erkennen, ohne dass das Funktionsvermögen

bei der Schreibverarbeitung so auffällig abnimmt, wie im Fall sequenzieller Daten.

[0085] Um das Funktionsvermögen beim Schreiben von Daten zu verbessern, kann jedes Festplattenlaufwerk **51** über eine Funktion wie folgt verfügen. Wenn nämlich das Festplattenlaufwerk **51** eine Datenschreibanforderung vom Controller **500** empfängt, schreibt es die Daten nur in den Plattencache, und es informiert den Controller **500** über den Abschluss des Schreibvorgangs. In diesem Fall können beim mittels der **Fig. 12** und **Fig. 13** beschriebenen Verfahren die geschriebenen Daten nicht untersucht werden. Die **Fig. 14** ist ein Flussdiagramm zur Verarbeitung zur Untersuchung geschriebener Daten, wenn das Festplattenlaufwerk **51** über eine derartige Funktion verfügt. Die CPU **503** überwacht das Festplattenlaufwerk **51** dahingehend, ob die Anzahl der Schreibvorgänge in das Festplattenlaufwerk **51** eine vorbestimmte Anzahl überschritten hat oder nicht (S1401). Wenn die vorbestimmte Anzahl überschritten ist, gibt die CPU **503** über den Platten-Steuerabschnitt **502** eine Anweisung zum Schreiben der im Plattencache abgespeicherten Daten in eine Magnetplatte an das Festplattenlaufwerk (S1402). Dann liest die CPU **503** die Daten vom Cachespeicher **62** und von der Magnetplatte (S1403 und S1404). Die CPU **503** klärt, ob die Daten vom Cachespeicher **62** und diejenigen von der Magnetplatte übereinstimmen oder nicht (S1405). Wenn sie nicht übereinstimmen, informiert die CPU **503** die Informationsverarbeitungsvorrichtung **300** über die Tatsache, dass der Schreibvorgang nicht normal ausgeführt wurde (S1406). So kann Inkorrektheit von Daten trotz der Verwendung der oben genannten Funktion des Verbesserns des Funktionsvermögens bei der Schreibverarbeitung erkannt werden. Übrigens werden bei der Verarbeitung gemäß der **Fig. 14** in die Magnetplatte geschriebene Daten und in den Cachespeicher geschriebene Daten verglichen, wenn die Anzahl der Schreibvorgänge eine vorbestimmte Anzahl überschreitet. Jedoch kann die Untersuchung immer dann vorgenommen werden, wenn eine vorbestimmte Zeit verstrichen ist, oder immer dann, wenn der Plattencache über keinen freien Raum verfügt.

[0086] Im Fall eines Seriell-ATA-Festplattenlaufwerks **51** werden Daten häufig aufgrund eines Ausfalls seines Kopfs inkorrekt geschrieben. Daher erfolgt eine Beschreibung zu einem Verfahren zum Erkennen des Ausfalls eines Kopfs oder eines Festplattenlaufwerks **51**, wenn Daten aus diesem gelesen werden.

[0087] Die **Fig. 15** ist ein Diagramm, das eine Kopfprüf-Kontrolltabelle **1501** zeigt. Die Kopfprüf-Kontrolltabelle **1501** besteht aus Laufwerksnummern, Kopfnummern und Sektornummern, und sie ist im Speicher **504** abgespeichert. Jede Sektornummer ist auf dieselbe Weise wie in der Aktualisierungskontrollta-

belle **1101** durch eine LBA definiert. Wenn Daten über den Platten-Steuerabschnitt **502** in das Festplattenlaufwerk **51** eingeschrieben werden, ändert die CPU **503** den Wert "Vorliegen von Aktualisierung" des Sektors des Kopfs, mit dem die Daten geschrieben wurden, in der Kopfprüf-Kontrolltabelle **1501** auf "1".

[0088] Die [Fig. 16](#) ist ein Flussdiagramm der durch die CPU **503** auszuführenden Kopfprüfverarbeitung. Die CPU **503** stellt 1 als Anfangswert einer Untersuchungskopfnummer ein (S1601). Die CPU **503** wartet eine vorbestimmte Zeit (S1602), und sie schreibt unter Verwendung eines durch die Untersuchungskopfnummer spezifizierten Kopfs Untersuchungsdaten in einen Steuerblock einer Magnetplatte (**1603**). Übrigens ist der Steuerblock ein vorbestimmter Speicherbereich auf der Magnetplatte. Als Nächstes liest die CPU **503** die in den Steuerblock geschriebenen Daten (S1604), und sie klärt, ob die gelesenen Daten und die Untersuchungsdaten miteinander übereinstimmen oder nicht (S1605).

[0089] Wenn die zwei Datenstücke miteinander übereinstimmen, schließt die CPU **503**, dass keine Anormalität im Kopf besteht, und sie ändert den Wert von "Vorliegen einer Aktualisierung" für den Kopf in der Kopfprüf-Kontrolltabelle **1501** auf "0" (S1606). Die CPU **503** addiert 1 zur Untersuchungskopfnummer (S1607). Die CPU **503** klärt, ob die Untersuchungskopfnummer größer als ein Maximalwert für die Kopfnummer ist oder nicht (S1608). Wenn die Untersuchungskopfnummer größer ist, stellt die CPU **503** 1 als Untersuchungskopfnummer ein. Die CPU **503** führt die Kopfprüfverarbeitung wiederholt auf die eingestellte Kopfnummer hin aus.

[0090] Wenn die aus dem Steuerblock gelesenen Daten nicht mit den Untersuchungsdaten übereinstimmen, informiert die CPU **503** die Informationsverarbeitungsvorrichtung **300** über die Tatsache, dass im fraglichen Festplattenlaufwerk **51** eine Anormalität auftritt, und dann beendet sie die Verarbeitung.

[0091] Die [Fig. 17](#) ist ein Flussdiagramm einer Verarbeitung, wenn die CPU **503** eine Datenleseanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** empfängt. Die CPU **503** empfängt die Datenleseanforderung von der Informationsverarbeitungsvorrichtung **300** über den Kanal-Steuerabschnitt **501** (S1701). Die CPU **503** klärt den Wert von "Vorliegen einer Aktualisierung" eines Zielsektors eines Festplattenlaufwerks **51**, wo die Daten in der Kopfprüf-Kontrolltabelle **1501** gespeichert sind (S1702 und S1703). Ein "Vorliegen einer Aktualisierung" vom Wert "1" zeigt den Zustand an, bei dem die oben genannte Kopfprüfverarbeitung nicht ausgeführt wurde, obwohl ein Datenschreibvorgang an der LBA des fraglichen Festplattenlaufwerks **51** ausgeführt wurde. Wenn der Wert von "Vorliegen

einer Aktualisierung" "0" ist, liest die CPU **503** die Daten vom Festplattenlaufwerk **51** (S1708).

[0092] Wenn der Wert von "Vorliegen einer Aktualisierung" "1" ist, schreibt die CPU **503** Untersuchungsdaten unter Verwendung des fraglichen Kopfs auf dieselbe Weise wie bei der oben genannten Kopfprüfverarbeitung in einen Steuerblock einer Magnetplatte (S1704). Übrigens ist der Steuerblock ein vorbestimmter Speicherbereich auf der Magnetplatte. Als Nächstes liest die CPU **503** die in den Steuerblock geschriebenen Daten (S1705), und sie klärt, ob die gelesenen Daten mit den Untersuchungsdaten übereinstimmen oder nicht (S1706).

[0093] Wenn die zwei Datenstücke miteinander übereinstimmen, schließt die CPU **503**, dass im Kopf keine Anormalität existiert, und sie ändert den Wert von "Vorliegen einer Aktualisierung" für den Kopf in der Kopfprüf-Kontrolltabelle **1501** auf "0" (S1707). Dann liest die CPU **503** die angeforderten Daten entsprechend der Leseanforderung vom Festplattenlaufwerk **51** (S1708).

[0094] Wenn die vom Steuerblock gelesenen Daten nicht mit den Untersuchungsdaten übereinstimmen, informiert die CPU **503** die Informationsverarbeitungsvorrichtung **300** über die Tatsache, dass im fraglichen Festplattenlaufwerk **51** eine Anormalität vorliegt (S1709), und dann beendet die CPU **503** die Verarbeitung ohne die Daten vom Festplattenlaufwerk **51** zu lesen.

[0095] Auf solche Weise ist es möglich, wenn in das Festplattenlaufwerk **51** geschriebene Daten gelesen werden, zu klären, ob der Kopf, mit dem die Daten geschrieben wurden, normal arbeitet oder nicht. Wenn der Kopf anormal ist, besteht die Möglichkeit, dass die Daten nicht korrekt geschrieben wurden, oder es besteht die Möglichkeit, dass die Daten nicht korrekt gelesen werden können. Durch Erkennen einer Anormalität im Kopf beim Lesen von Daten ist es möglich, zu verhindern, dass falsche Daten gelesen werden.

Untersuchung auf Grundlage einer Paritätszuweisung

[0096] Gemäß dem oben genannten Verfahren, bei dem alle Daten einer Streifengruppe in der RAID-Konfiguration gelesen und einer Paritätsprüfung unterzogen werden, ist es nicht möglich, zu ermitteln, welche Daten der Streifengruppe in einem falschen Zustand vorliegen. Es ist tatsächlich möglich, zu verhindern, dass falsche Daten gelesen werden, jedoch ist es nicht möglich, falsche Daten wiederherzustellen. Demgemäß können Daten verloren gehen. Daher erfolgt eine Beschreibung zu einem Verfahren zum Zuweisen von Paritätsdaten zu jedem Datenstück gesondert von Paritätsdaten innerhalb

der Streifengruppe.

[0097] Die CPU **503** erzeugt Paritätsdaten zum Erkennen von Fehlern für mehrere Sektoren, die als Minimaleinheit dienen, gemäß der Daten in jedes Festplattenlaufwerk **51** geschrieben werden. Bei dieser Ausführungsform wird eine Kombination von Daten und Paritätsdaten für derartige mehrere Sektoren als Dateneinheit bezeichnet. Wenn die CPU **503** eine Datenschreibanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** über den Kanal-Steuerabschnitt **501** empfängt, erzeugt sie aus den zu schreibenden Daten eine Dateneinheit. Die CPU **503** schreibt die Dateneinheit über die Platten-Steureinheit **502** in das Festplattenlaufwerk **51**.

[0098] Die [Fig. 18](#) ist ein Diagramm, das den Zustand zeigt, wenn ein Datenstück **1801** in ein Festplattenlaufwerk geschrieben wird. Die Daten **1801** bestehen aus mehreren Sektoren S#1 bis S#4, und aus den Daten **1801** und Paritätsdaten **1802** für die Daten **1801** der mehreren Sektoren wird eine Dateneinheit **1803** gebildet. Wenn die CPU **503** eine Datenleseanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** über den Kanal-Steuerabschnitt **501** empfängt, liest sie die Dateneinheit **1803** der angeforderten Daten über den Platten-Steuerabschnitt **502**, und sie führt an den Daten eine Paritätsprüfung aus, um zu untersuchen, ob sie sich in einem falschen Zustand befinden oder nicht. Auf diese Weise ist es nur durch Lesen der zu lesenden Daten auf die Leseanforderung hin möglich, zu beurteilen, ob sich die Daten in einem falschen Zustand befinden oder nicht. Außerdem können, wenn die Festplattenlaufwerke **51** über eine redundante RAID-Konfiguration, wie RAID **5**, verfügen, die Daten unter Verwendung anderer Daten und Paritätsdaten in der Streifengruppe wiederhergestellt werden. So besteht keine Gefahr, dass die Daten verloren gehen.

[0099] Wenn ein Ausfall eines Kopfs oder dergleichen in einem Festplattenlaufwerk **51** auftritt, ist es hoch wahrscheinlich, dass mehrere falsche Sektoren auftreten. Es sei angenommen, dass mehrere Sektoren der Dateneinheit **1803** falsch werden, wenn die Dateneinheit **1803** in ein Festplattenlaufwerk **51** eingeschrieben wird. Bei einem derartigen Ereignis besteht der Fall, dass Inkorrektheit nicht durch eine Paritätsprüfung erkannt werden kann.

[0100] Daher kann, wie es in der [Fig. 19](#) dargestellt ist, die CPU **503** die Dateneinheit **1803** zwischen mehreren Festplattenlaufwerken **51** innerhalb der RAID-Gruppe mittels des Platten-Steuerabschnitts **502** schreiben und verteilen. Die [Fig. 20](#) ist ein Diagramm, das eine Dateneinheit-Kontrolltabelle **2001** zeigt. Die Dateneinheits-Kontrolltabelle **2001** zeigt die Entsprechung von Dateneinheiten **1803** aus je-

weils mehreren Sektoren mit LBAs von Festplattenlaufwerken **51**. Das Beispiel der [Fig. 20](#) zeigt, dass eine aus **130** Sektoren 000 - 129 bestehende Dateneinheit **1803** durch LBAs 000 - 064 von Festplattenlaufwerken **51** gebildet wird, deren Laufwerksnummern #0 und #1 sind. Wenn die CPU **503** die Datenschreibanforderung von irgendeiner der Informationsverarbeitungsvorrichtungen **300** empfängt, nimmt sie auf die Dateneinheits-Kontrolltabelle **2001** Bezug, und sie schreibt und verteilt jede Dateneinheit **1803** der angeforderten Daten über mehrere Festplattenlaufwerke **51**.

[0101] Demgemäß ist es selbst dann, wenn in einem Festplattenlaufwerk ein Fehler auftritt, möglich, die Wahrscheinlichkeit zu erhöhen, dass Inkorrektheit der Daten erkannt werden kann.

Umgebung mit einer Mischung eines Faserkanals und seriell-ATA

[0102] Als Nächstes erfolgt eine Beschreibung zu einem Plattenarraysystem **10**, bei dem Faserkanal-Festplattenlaufwerke **51** und Seriell-ATA-Festplattenlaufwerke **51** gemischt sind.

[0103] Die [Fig. 21](#) ist ein Blockdiagramm, das ein Plattenarraysystem zeigt, bei dem Faserkanal-Festplattenlaufwerke **51** in einem ersten Gehäuse **2101** aufgenommen sind und Seriell-ATA-Festplattenlaufwerke **51** in einem zweiten Gehäuse **2102** aufgenommen sind. Übrigens entsprechen das erste und das zweite Gehäuse **2101** und **2102** dem Mastergehäuse **20** bzw. dem Erweiterungsgehäuse **30**. Jedes Festplattenlaufwerk **51** ist im oben beschriebenen Modus mit dem Platten-Steuerabschnitt **502** verbunden. Außerdem zeigt die [Fig. 19](#) einen Modus, bei dem mehrere Seriell-ATA-Festplattenlaufwerke mit einem Wandler **901** verbunden sind. Jedoch kann jedes Seriell-ATA-Festplattenlaufwerk mit einem Wandler **801** verbunden sein, der, wie oben beschrieben, für jede Plattenlaufwerkseinheit vorhanden ist.

[0104] Beim auf diese Weise konfigurierten Plattenarraysystem **10** ist es erforderlich, die Zuverlässigkeit der Seriell-ATA-Festplattenlaufwerke **51** zu erhöhen, deren Zuverlässigkeit niedriger als die der Faserkanal-Festplattenlaufwerke **51** ist. Daher wendet der Controller **500** das oben genannte Verfahren zum Verbessern der Zuverlässigkeit nur der Seriell-ATA-Festplattenlaufwerke **51** an. Demgemäß kann die Zuverlässigkeit beim Lesen/Schreiben von Daten von den/in die Seriell-ATA-Festplattenlaufwerke **51** verbessert werden, ohne dass das Funktionsvermögen beim Lesen/Schreiben von Daten aus den/in die Seriell-ATA-Festplattenlaufwerke **51** verbessert werden, ohne dass das Funktionsvermögen beim Lesen/Schreiben von Daten von den/in die Faserkanal-Festplattenlaufwerke **51**, die zur Verarbeitung wie einer wesentlichen Arbeit, die über hohes

Zugriffsfunktionsvermögen verfügen muss, gesenkt würde. Außerdem ist es nicht erforderlich, den körperlichen Aufbau zu ändern, z.B. zwei Köpfe für jede Magnetplatte jedes Seriell-ATA-Festplattenlaufwerks **51** anzubringen. Daher ist es möglich, die Herstellkosten der Seriell-ATA-Festplattenlaufwerke **51** niedrig zu halten.

[0105] Übrigens sind bei dieser Ausführungsform die Faserkanal-Festplattenlaufwerke **51** und die Seriell-ATA-Festplattenlaufwerke **51** gemischt vorhanden. Jedoch können andere Festplattenlaufwerke **51** verwendet werden, wenn sie zu Schnittstellenstandards mit verschiedenen Zuverlässigkeiten gehören.. Z.B. können die Seriell-ATA-Festplattenlaufwerke **51** durch Parallel-ATA-Festplattenlaufwerke **51** ersetzt werden.

[0106] Die Ausführungsformen wurden oben beschrieben, um die Erfindung leicht verständlich zu machen. Die Erfindung soll nicht als auf die Ausführungsformen beschränkt interpretiert werden. Die Erfindung kann geändert oder modifiziert werden, ohne dass von ihrem Schutzzumfang und Grundgedanken abgewichen würde. Jedes Äquivalent zur Erfindung ist ebenfalls in dieser enthalten. Außerdem wird die beim japanischen Patentamt am 28. November 2003 eingereichte japanische Patentanmeldung Nr. 2003-400517 zum Stützen der Erfindung genannt, und deren Offenbarung wird hier durch Bezugnahme eingeschlossen.

Patentansprüche

1. Plattenarraysystem mit:

- einem ersten Gehäuse zum Aufnehmen einer oder mehrerer RAID-Gruppen aus jeweils mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem ersten Schnittstellenstandard, wobei die Festplattenlaufwerken über einen Kommunikationspfad miteinander verbunden sind;
- einem zweiten Gehäuse zum Aufnehmen einer oder mehrerer RAID-Gruppen aus jeweils mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem zweiten Schnittstellenstandard, wobei die Festplattenlaufwerken über den Kommunikationspfad mittels mehrerer Wandlungseinheiten zum Wandeln des ersten und des zweiten Schnittstellenstandards ineinander verbunden sind, wobei diese Festplattenlaufwerke über geringere Zuverlässigkeit als die im ersten Gehäuse untergebrachten Festplattenlaufwerke verfügen; und
- einem Controller mit einem Kanal-Steuerabschnitt, einem Platten-Steuerabschnitt, einem Cachespeicher und einer CPU, wobei der Kanal-Steuerabschnitt so mit einer Informationsverarbeitungsvorrichtung verbunden ist, dass er mit dieser in Kommunikation treten kann und Anforderungen von ihr empfangen kann, zu denen eine Leseanforderung zum Lesen von Daten von den Festplattenlaufwerken im ers-

ten oder zweiten Gehäuse und eine Schreibanforderung zum Schreiben von Daten in die Festplattenlaufwerke im ersten oder zweiten Gehäuse gehören, wobei der Platten-Steuerabschnitt mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse über den Kommunikationspfad so verbunden ist, dass er für Kommunikation mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse sorgen kann und eine Eingabe/Ausgabe von Daten und Paritätsdaten aus den/in die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse entsprechend der vom Kanal-Steuerabschnitt empfangenen Leseanforderung oder Schreibanforderung ausführen kann, wobei die Paritätsdaten Daten zum Erkennen von Fehlern über mehrere Datenstücke einschließlich der Daten von den/an die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse sind, wobei der Cachespeicher eine Zwischenspeicherung von in die mehreren Festplattenlaufwerke einzuschreibenden Daten ausführt und wobei die CPU die Steuerung über den Kanal-Steuerabschnitt und den Platten-Steuerabschnitt verwaltet;

- wobei der Controller mehrere Datenstücke einschließlich in den mehreren Festplattenlaufwerken im zweiten Gehäuse gespeicherten Daten und Paritätsdaten für die mehreren Datenstücke von allen Festplattenlaufwerken der genannten RAID-Gruppe, zu der die die Daten speichernden Festplattenlaufwerke gehören, liest und untersucht, ob die mehreren Datenstücke, die diese Daten enthalten, mit fehlerhaftem Inhalt oder nicht in die Festplattenlaufwerke geschrieben wurden.

2. Plattenarraysystem nach Anspruch 1, bei dem der Controller von der Informationsverarbeitungsvorrichtung die Leseanforderung zum Lesen von in einem der Festplattenlaufwerke im zweiten Gehäuse gespeicherten Daten empfängt und dann die Untersuchung an diesen Daten ausführt.

3. Plattenarraysystem nach Anspruch 1, bei dem der Controller die Prüfung an den in den mehreren Festplattenlaufwerken im zweiten Gehäuse gespeicherten Daten unabhängig von der genannten Leseanforderung von der Informationsverarbeitungsvorrichtung ausführt.

4. Plattenarraysystem nach Anspruch 1, bei dem:

- der Controller in einer Aktualisierungskontrolltabelle eine Position speichert, an der in eines der Festplattenlaufwerke im zweiten Gehäuse entsprechend der genannten Schreibanforderung von der Informationsverarbeitungsvorrichtung eingeschriebene Daten in das Festplattenlaufwerk geschrieben wurden;
- der Controller die Untersuchung an den an der genannten, in der Aktualisierungskontrolltabelle gespeicherten Position gespeicherten Daten ausführt und in der Aktualisierungskontrolltabelle Information speichert, die den Abschluss der genannten Untersuchung anzeigt, die an den an dieser Position gespei-

cherten Daten ausgeführt wurde; und
 – der Controller von der Informationsverarbeitungs-
 vorrichtung die genannte Leseanforderung zum Le-
 sen von in einem der Festplattenlaufwerke im zwei-
 ten Gehäuse gespeicherten Daten empfängt, dann
 auf die Aktualisierungskontrolltabelle Bezug nimmt
 und die genannte Untersuchung an den Daten aus-
 führt, wenn diese Untersuchung noch nicht an diesen
 Daten ausgeführt wurde.

5. Plattenarraysystem nach Anspruch 1, bei dem
 der Controller Daten entsprechend der genannten
 Schreibanforderung von der Informationsverarbeit-
 ungsvorrichtung in eines der Festplattenlaufwerke
 im zweiten Gehäuse schreibt, dann einen zu diesem
 Festplattenlaufwerk gehörenden Kopf von einer Posi-
 tion, an der die Daten gespeichert wurden, bewegt,
 danach die Daten von einer zu diesem Festplatten-
 laufwerk gehörenden Magnetplatte und vom Cache-
 speicher liest und die zwei ausgelesenen Datenstü-
 cke vergleicht.

6. Plattenarraysystem nach Anspruch 1, bei dem
 der Controller Daten entsprechend der Schreibanfor-
 derung von der Informationsverarbeitungsvorrich-
 tung in eines der Festplattenlaufwerke im zweiten
 Gehäuse schreibt, dann einen zu diesem Festplät-
 tenlaufwerk gehörenden Kopf von einer Position, an
 der die Daten gespeichert wurden, bewegt, danach
 einen Teil der Daten von einer zu diesem Festplät-
 tenlaufwerk gehörenden Magnetplatte und vom Cache-
 speicher liest und die zwei ausgelesenen Datenstü-
 cke vergleicht.

7. Plattenarraysystem nach Anspruch 1, bei dem:
 – der Controller Daten in eines der Festplattenlauf-
 werke im zweiten Gehäuse entsprechend der
 Schreibanforderung von der Informationsverarbeit-
 ungsvorrichtung schreibt und dann einen zu diesem
 Festplattenlaufwerk gehörenden Kopf von einer Posi-
 tion, an der die Daten gespeichert wurden, bewegt;
 und
 – der Controller danach die Daten von einer zu die-
 sem Festplattenlaufwerk gehörenden Magnetplatte
 und vom Cachespeicher liest und die zwei ausgele-
 senen Datenstücke vergleicht, wenn die Größe der
 Daten kleiner als ein vorbestimmter Wert ist, oder er
 einen Teil der Daten von der Magnetplatte und vom
 Cachespeicher liest und die zwei ausgelesenen Da-
 tenstücke vergleicht, wenn die Größe der genannten
 Daten nicht kleiner als ein vorbestimmter Wert ist.

8. Plattenarraysystem nach Anspruch 1, bei dem
 dann, wenn die Anzahl von Schreibanforderungen
 von der Informationsverarbeitungsvorrichtung für ei-
 nes der Festplattenlaufwerke im zweiten Gehäuse
 eine vorbestimmte Anzahl überschreitet, der Control-
 ler in einem zu diesem Festplattenlaufwerk gehö-
 renden Plattencache abgespeicherte Daten in eine zu
 diesem Festplattenlaufwerk gehörende Magnetplatte

schreibt, die genannten Daten von der Magnetplatte
 und vom zum Controller gehörenden Cachespeicher
 liest und die zwei ausgelesenen Datenstücke ver-
 gleicht.

9. Plattenarraysystem nach Anspruch 1, bei dem
 der Controller in einem zu jedem der genannten Fest-
 plattenlaufwerke im zweiten Gehäuse gehörenden
 Plattencache gespeicherte Daten in eine zum ge-
 nannten Festplattenlaufwerk gehörende Magnetplät-
 te schreibt, die Daten von der genannten Magnetplät-
 te und vom zum Controller gehörenden Cachespei-
 cher liest und die zwei ausgelesenen Datenstücke
 vergleicht, was immer dann erfolgt, wenn eine vorbe-
 stimmte Zeit verstrichen ist.

10. Plattenarraysystem nach Anspruch 1, bei
 dem dann, wenn in einem zu jedem der Festplatten-
 laufwerke im zweiten Gehäuse gehörenden Platten-
 cache kein freier Raum existiert, der Controller im ge-
 nannten, zum genannten Festplattenlaufwerk gehö-
 renden Plattencache gespeicherte Daten in eine zum
 genannten Festplattenlaufwerk gehörende Magnet-
 platte schreibt, die Daten von der genannten Magnet-
 platte und vom zum genannten Controller gehö-
 renden Cachespeicher liest und die zwei ausgelesenen
 Datenstücke vergleicht.

11. Plattenarraysystem nach Anspruch 1, bei
 dem:

– jedes der genannten Festplattenlaufwerke über
 mehrere Köpfe zum Lesen/Schreiben von Daten ver-
 fügt;

– der Controller in einer Kopfprüf-Kontrolltabelle ei-
 nen der genannten Köpfe speichert, der entspre-
 chend der genannten Schreibanforderung von der
 Informationsverarbeitungsvorrichtung Daten in das
 genannte Festplattenlaufwerk geschrieben hat; und
 – der Controller die genannte Leseanforderung von
 der genannten Informationsverarbeitungsvorrichtung
 empfängt, dann auf die genannte Kopfprüf-Kontroll-
 tabelle Bezug nimmt, dann Untersuchungsdaten in
 eine zum genannten Festplattenlaufwerk gehörende
 Magnetplatte unter Verwendung eines der Köpfe
 schreibt, die zum Lesen der genannten Daten zu ver-
 wenden sind, er die genannten Untersuchungsdaten
 von der Magnetplatte liest und die gelesenen Daten
 mit den Untersuchungsdaten vergleicht, wenn der
 genannte Kopf, der zum Lesen der genannten Daten
 verwendet wird, in der Kopfprüf-Kontrolltabelle ge-
 speichert wurde.

12. Plattenarraysystem nach Anspruch 1, bei
 dem:

– der Controller von der Informationsverarbeitungs-
 vorrichtung die Schreibanforderung zum Schreiben
 von Daten in eines der Festplattenlaufwerke im zwei-
 ten Gehäuse empfängt, er dann aus den Daten eine
 Dateneinheit aus mehreren Sektoren auf Grundlage
 der genannten zu schreibenden Daten sowie Pari-

tätsdaten zum Erkennen von Datenfehlern in den mehreren Sektoren erzeugt und er die genannte Dateneinheit in das genannte Festplattenlaufwerk schreibt; und

– der Controller von der Informationsverarbeitungsvorrichtung die Leseanforderung zum Lesen der Daten empfängt, er dann die genannte Dateneinheit liest und er untersucht, ob die genannten Daten mit fehlerhaftem Inhalt oder nicht in den genannten Festplattenlaufwerk gespeichert sind.

13. Plattenarraysystem nach Anspruch 12, bei dem der Controller die genannte Dateneinheit in mehrere Stücke unterteilt und er diese Stücke der Dateneinheit jeweils in mehrere Festplattenlaufwerke schreibt, die zu den genannten RAID-Gruppen im zweiten Gehäuse gehören.

14. Plattenarraysystem nach Anspruch 1, bei dem:

– der erste Schnittstellenstandard ein Faserkanalstandard ist, der zweite Schnittstellenstandard ein Seriell-ATA-Standard ist und der Kommunikationspfad ein FC-AL ist.

15. Plattenarraysystem nach Anspruch 1, bei dem:

– der erste Schnittstellenstandard ein Faserkanalstandard ist, der zweite Schnittstellenstandard ein Parallel-ATA-Standard ist und der Kommunikationspfad ein FC-AL ist.

16. Verfahren zum Steuern eines Plattenarraysystems mit:

– einem ersten Gehäuse zum Aufnehmen einer oder mehrerer RAID-Gruppen aus jeweils mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem ersten Schnittstellenstandard, wobei die Festplattenlaufwerken über einen Kommunikationspfad miteinander verbunden sind;

– einem zweiten Gehäuse zum Aufnehmen einer oder mehrerer RAID-Gruppen aus jeweils mehreren Festplattenlaufwerken zum Senden/Empfangen von Daten entsprechend einem zweiten Schnittstellenstandard, wobei die Festplattenlaufwerken über den Kommunikationspfad mittels mehrerer Wandlungseinheiten zum Wandeln des ersten und des zweiten Schnittstellenstandards ineinander verbunden sind, wobei diese Festplattenlaufwerke über geringere Zuverlässigkeit als die im ersten Gehäuse untergebrachten Festplattenlaufwerke verfügen; und

– einem Controller mit einem Kanal-Steuerabschnitt, einem Platten-Steuerabschnitt, einem Cachespeicher und einer CPU, wobei der Kanal-Steuerabschnitt so mit einer Informationsverarbeitungsvorrichtung verbunden ist, dass er mit dieser in Kommunikation treten kann und Anforderungen von ihr empfangen kann, zu denen eine Leseanforderung zum Lesen von Daten von den Festplattenlaufwerken im ersten oder zweiten Gehäuse und eine Schreibanforde-

– rung zum Schreiben von Daten in die Festplattenlaufwerke im ersten oder zweiten Gehäuse gehören, wobei der Platten-Steuerabschnitt mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse über den Kommunikationspfad so verbunden ist, dass er für Kommunikation mit den mehreren Festplattenlaufwerken im ersten und zweiten Gehäuse sorgen kann und eine Eingabe/Ausgabe von Daten und Paritätsdaten aus den/in die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse entsprechend der vom Kanal-Steuerabschnitt empfangenen Leseanforderung oder Schreibanforderung ausführen kann, wobei die Paritätsdaten Daten zum Erkennen von Fehlern über mehrere Datenstücke einschließlich der Daten von den/an die mehreren Festplattenlaufwerke im ersten und zweiten Gehäuse sind, wobei der Cachespeicher eine Zwischenspeicherung von in die mehreren Festplattenlaufwerke einzuschreibenden Daten ausführt und wobei die CPU die Steuerung über den Kanal-Steuerabschnitt und den Platten-Steuerabschnitt verwaltet; wobei dieses Verfahren die folgenden Schritte aufweist:

– Lesen mehrerer Datenstücke, die in den mehreren Festplattenlaufwerken im zweiten Gehäuse gespeicherte Daten enthalten, und von Paritätsdaten für diese mehreren Datenstücke von allen genannten Festplattenlaufwerken der genannten RAID-Gruppe, zu der die die genannten Daten speichernden Festplattenlaufwerke gehören; und

– Untersuchen, ob die genannten mehreren Datenstücke, die die Daten enthalten, mit fehlerhaftem Inhalt oder nicht in die Festplattenlaufwerke geschrieben wurden;

– wobei die genannten Schritte durch den genannten Controller ausgeführt werden.

17. Verfahren zum Steuern eines Plattenarraysystems nach Anspruch 16, ferner mit den folgenden Schritten:

– Schreiben von Daten in eines der Festplattenlaufwerke im zweiten Gehäuse entsprechen der Schreibanforderung von der Informationsverarbeitungsvorrichtung und anschließendes Verstellen eines zum genannten Festplattenlaufwerk gehörenden Kopfs von einer Position, an der die Daten gespeichert wurden;

– anschließendes Lesen der genannten Daten von einer zum genannten Festplattenlaufwerk gehörenden Magnetplatte und vom Cachespeicher; und

– Vergleichen der zwei ausgelesenen Datenstücke;

– wobei die genannten Schritte durch den genannten Controller ausgeführt werden.

18. Verfahren zum Steuern eines Plattenarraysystems nach Anspruch 16, ferner mit den folgenden Schritten:

– Empfangen der genannten Schreibanforderung von der Informationsverarbeitungsvorrichtung, um Daten in eines der Festplattenlaufwerke im zweiten Gehä-

se zu schreiben, und anschließendes Erzeugen einer Dateneinheit aus Daten, die aus mehreren Sektoren bestehen, auf Grundlage der zu schreibenden Daten und Paritätsdaten zum Erkennen von Datenfehlern in den mehreren Sektoren;

- Schreiben der genannten Dateneinheit in das genannte Festplattenlaufwerk;
- Empfangen der Leseanforderung von der Informationsverarbeitungsvorrichtung zum Lesen der genannten Daten, und anschließendes Lesen der genannten Dateneinheit; und
- Untersuchen, ob die Daten im genannten Festplattenlaufwerk mit fehlerhaftem Inhalt abgespeichert sind oder nicht;
- wobei die genannten Schritte durch den genannten Controller ausgeführt werden.

19. Verfahren zum Steuern eines Plattenarray-systems nach Anspruch 16, bei dem:

- der erste Schnittstellenstandard ein Faserkanalstandard ist, der zweite Schnittstellenstandard ein Seriell-ATA-Standard ist und der Kommunikationspfad ein FC-AL ist.

20. Verfahren zum Steuern eines Plattenarray-systems nach Anspruch 16, bei dem:

- der erste Schnittstellenstandard ein Faserkanalstandard ist, der zweite Schnittstellenstandard ein Parallel-ATA-Standard ist und der Kommunikationspfad ein FC-AL ist.

Es folgen 19 Blatt Zeichnungen

Anhängende Zeichnungen

FIG.1A

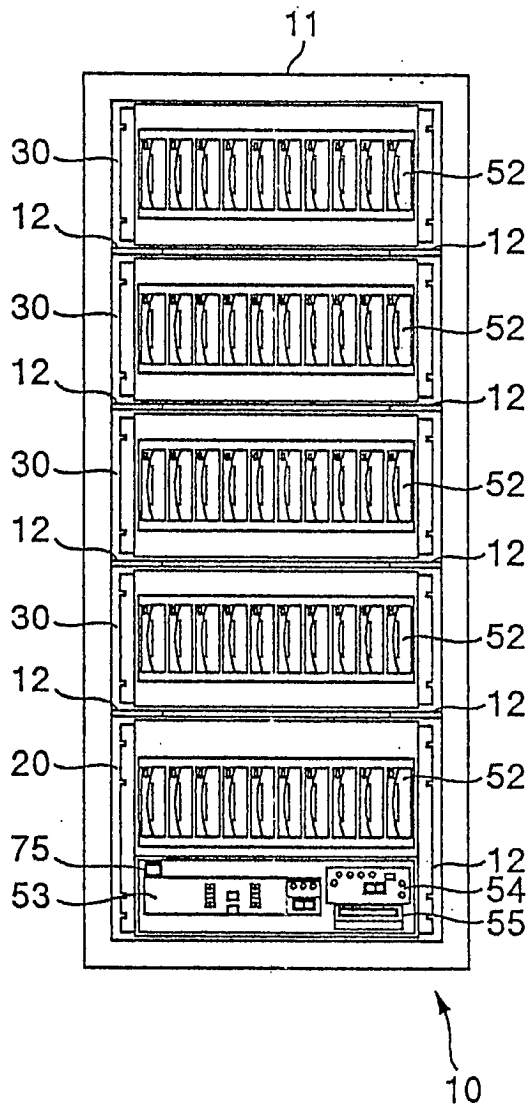


FIG.1B

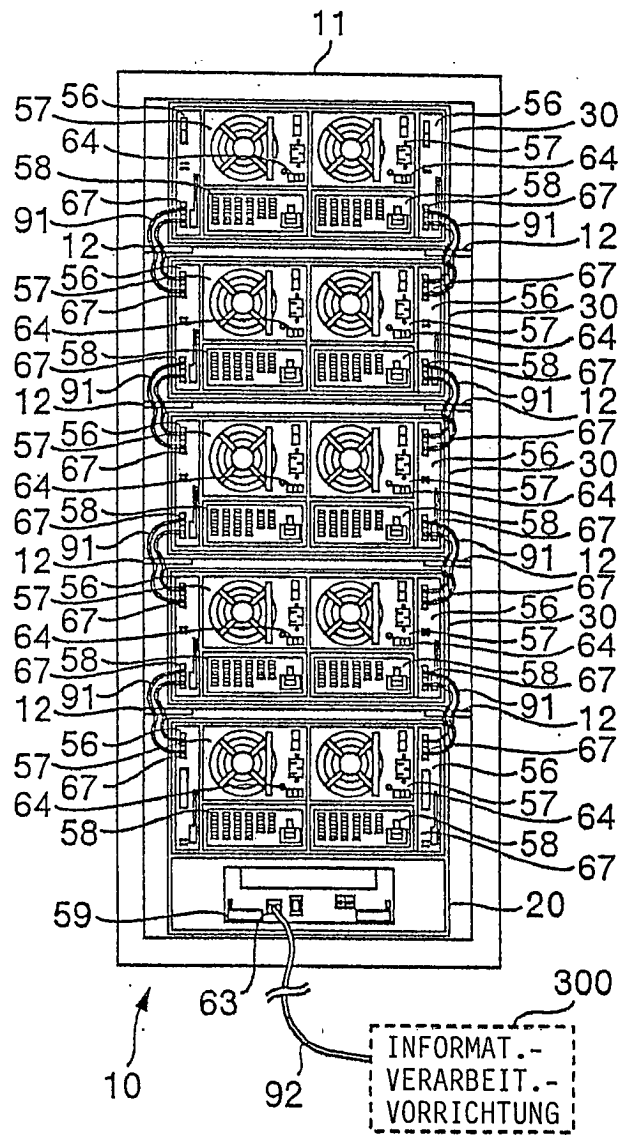


FIG.2A

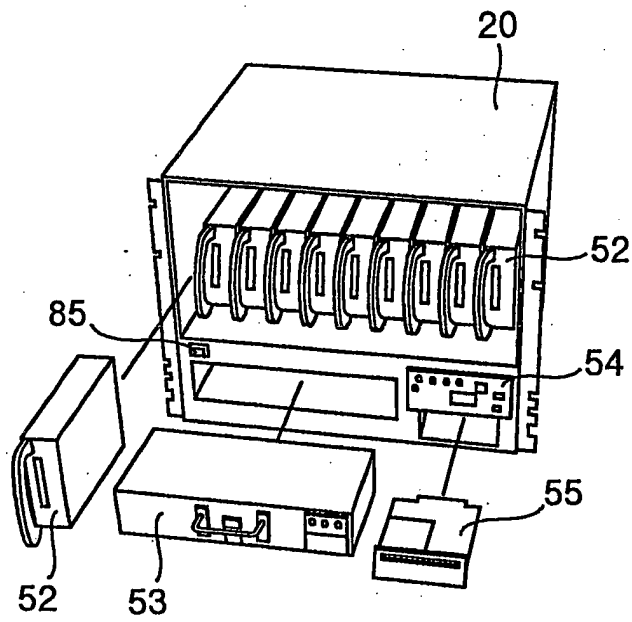


FIG.2B

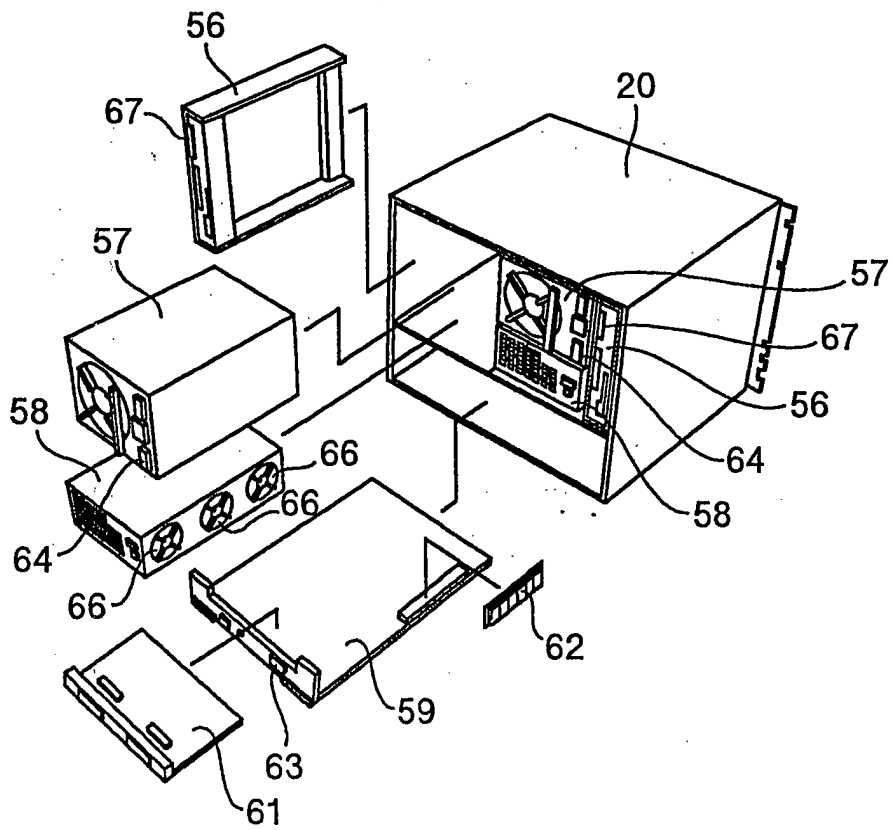


FIG.3A

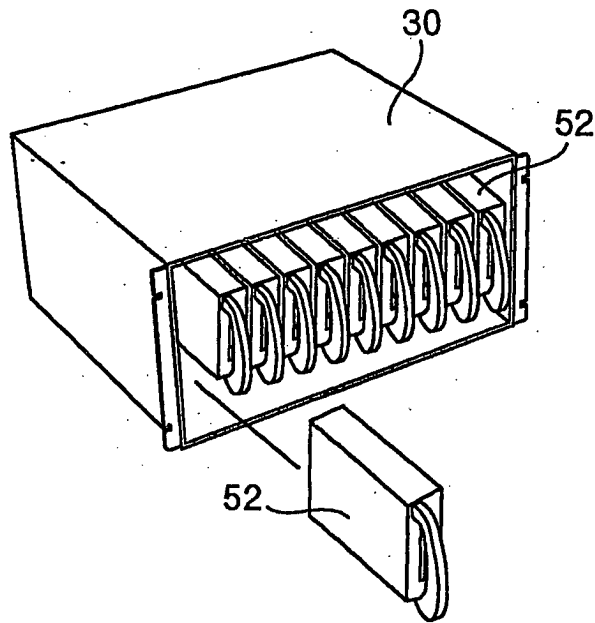


FIG.3B

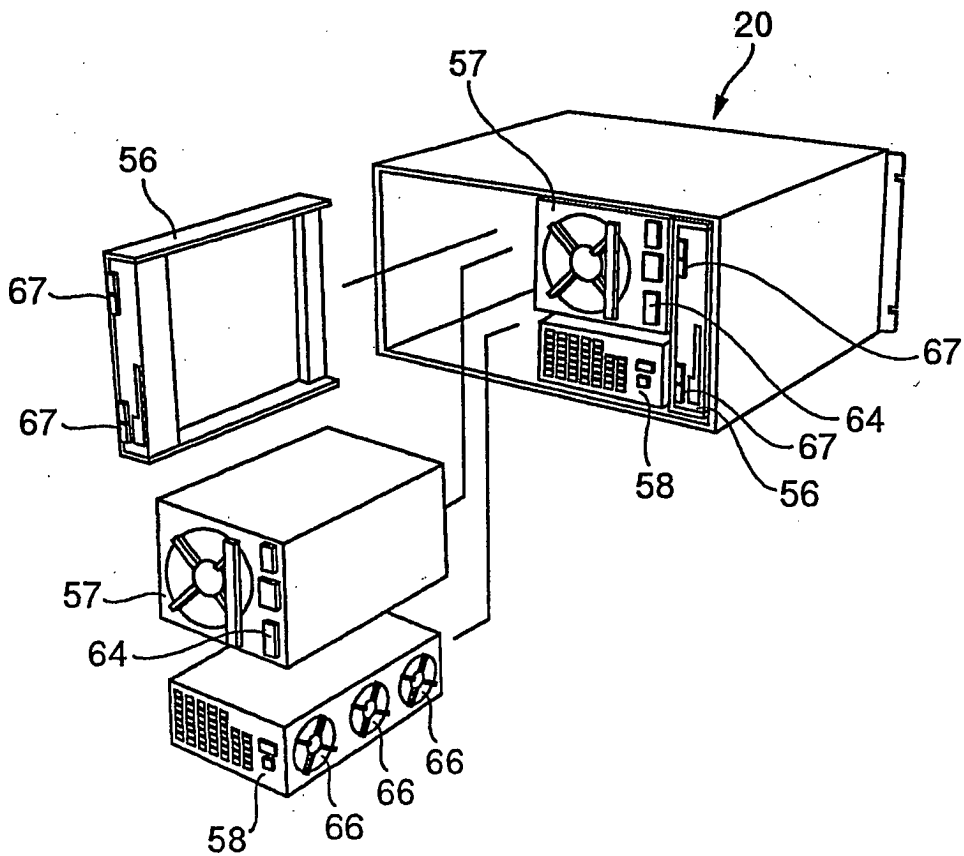


FIG.4

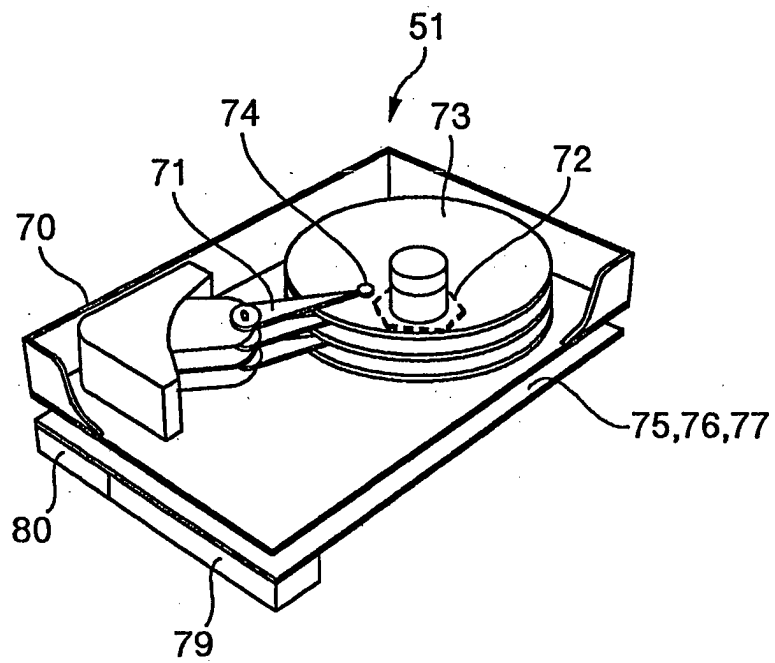


FIG.5

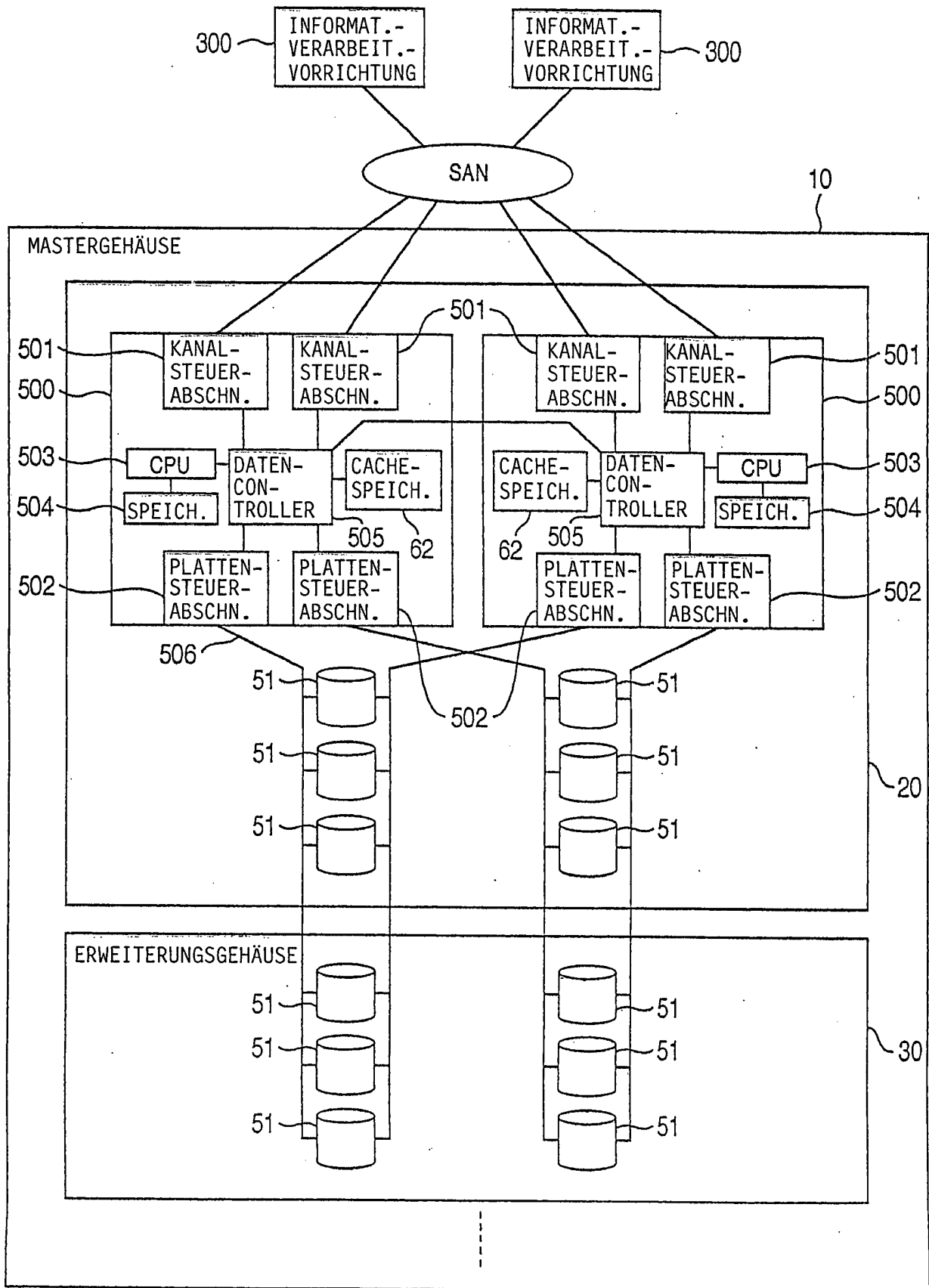


FIG.6

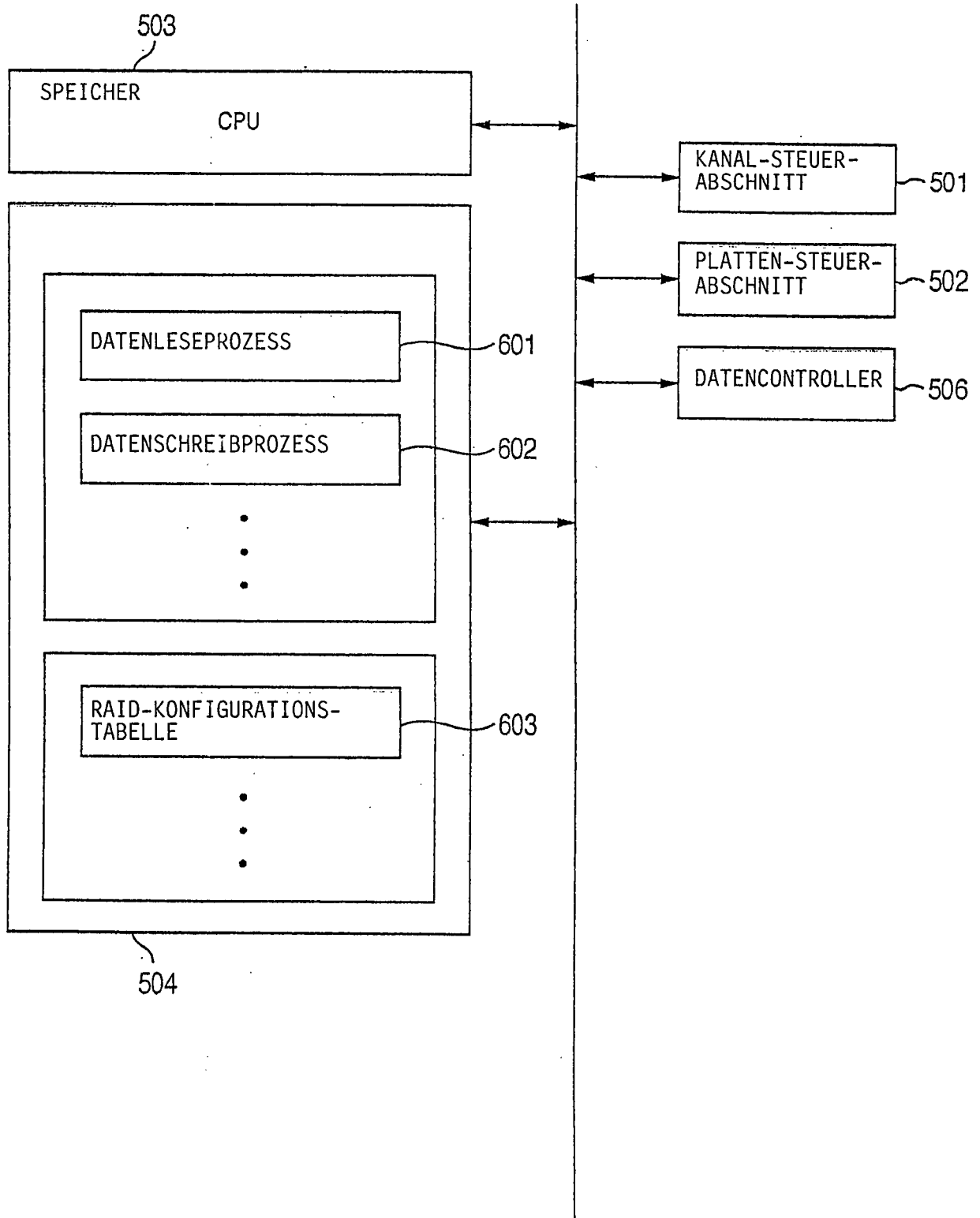


FIG.7

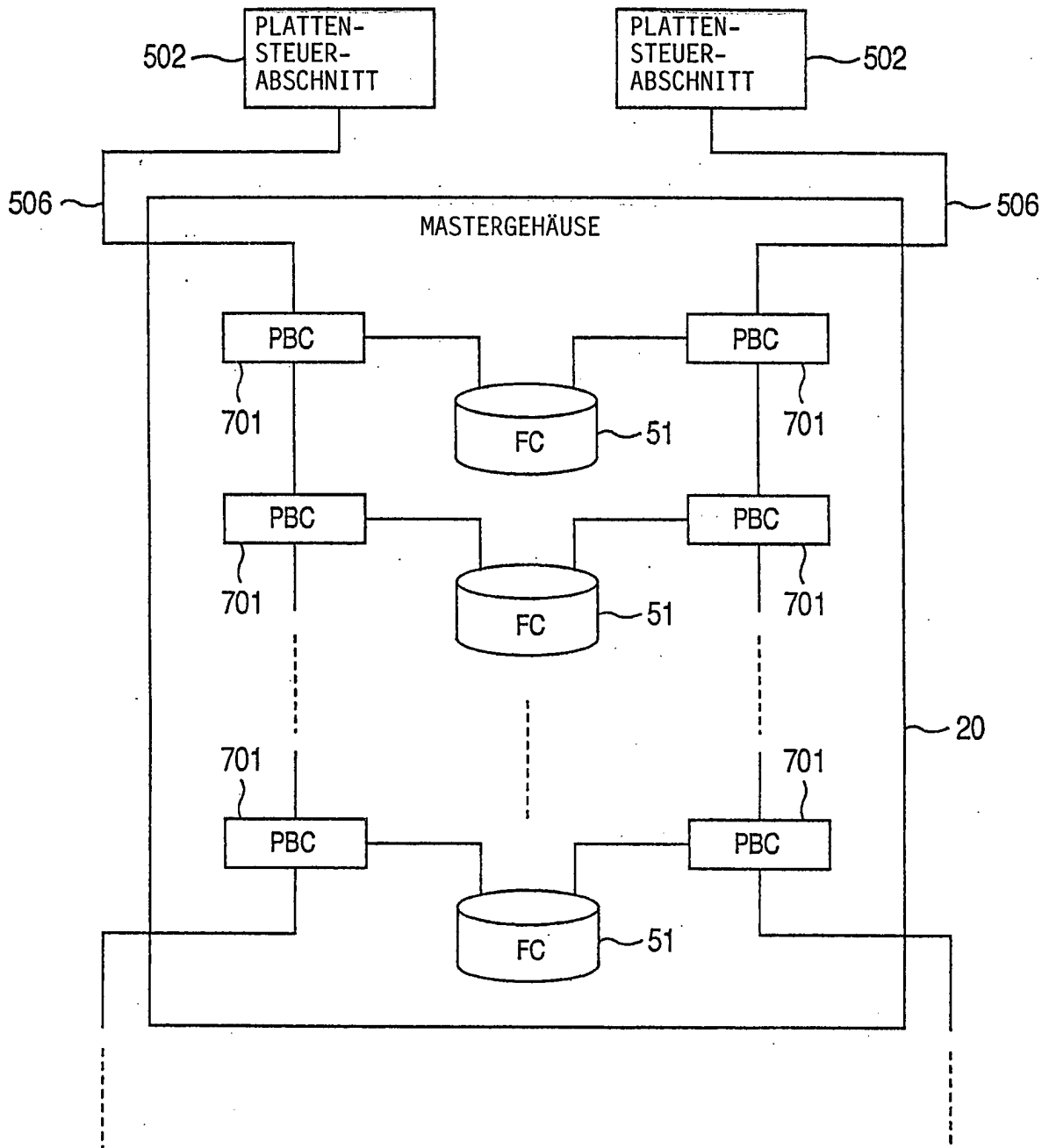


FIG.8

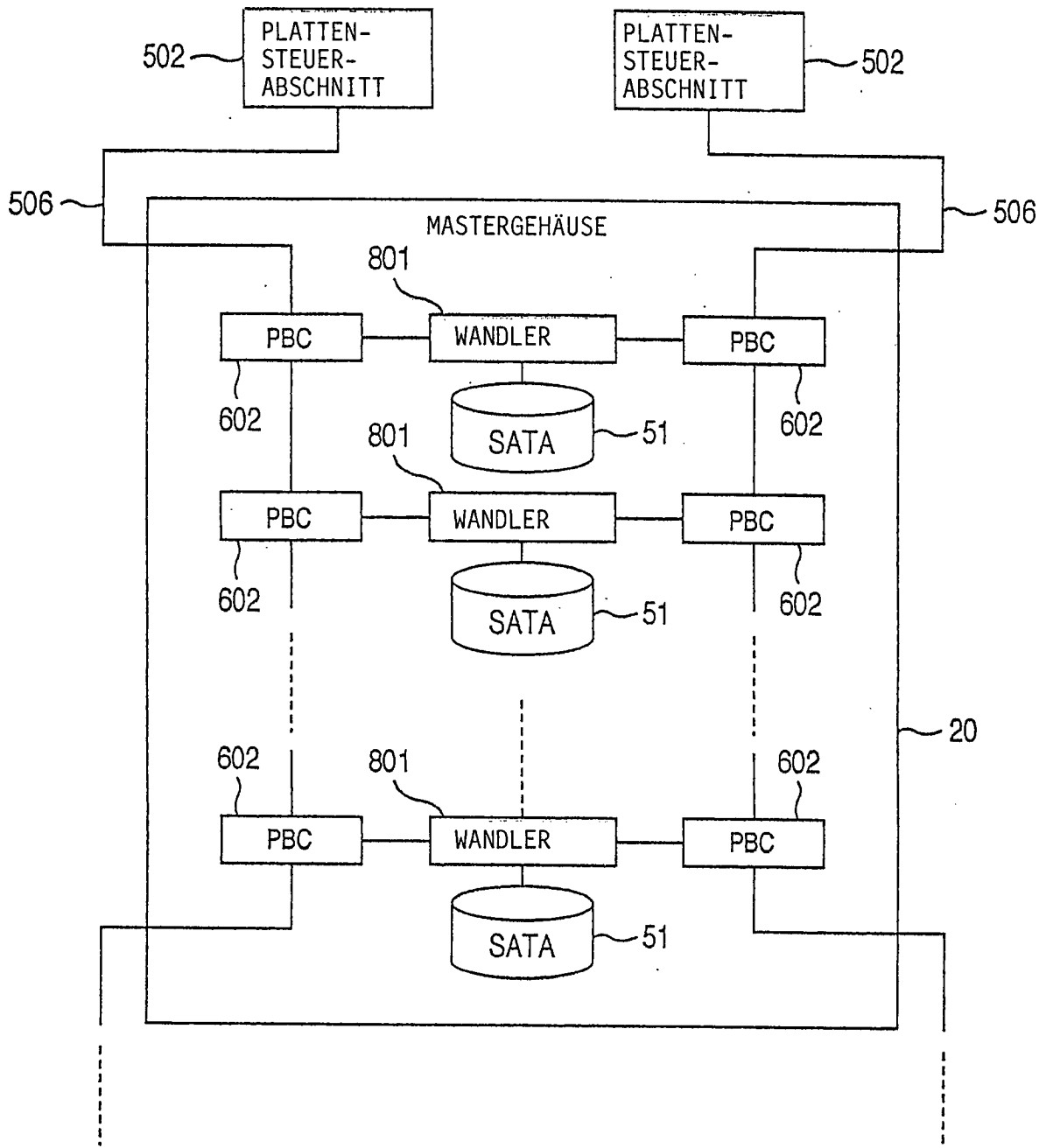


FIG.9

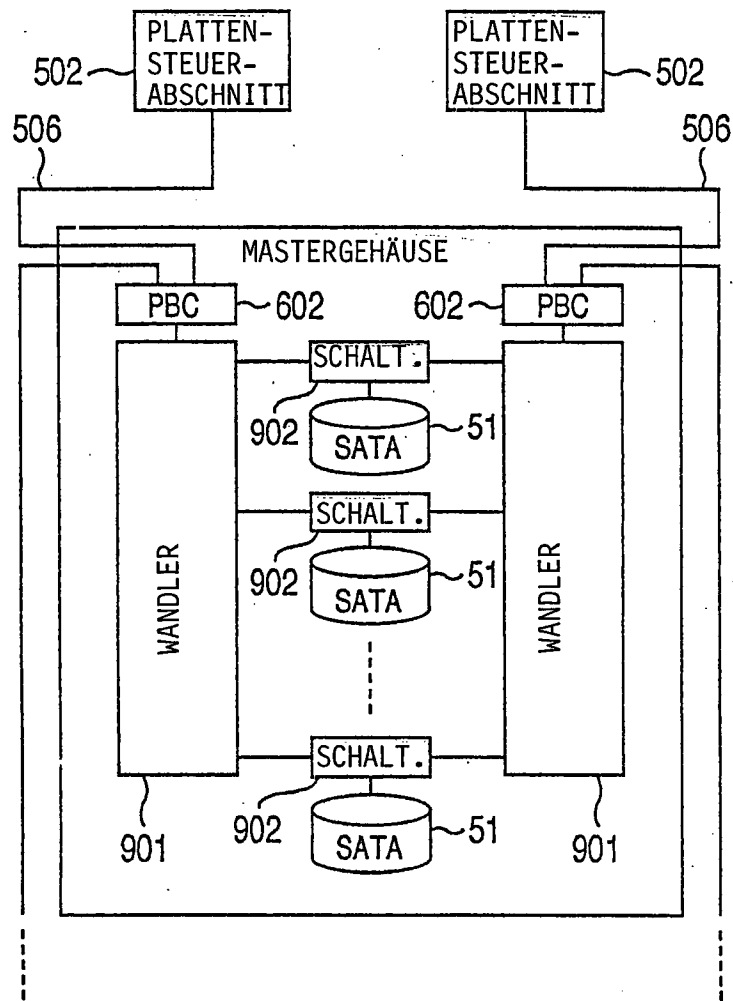


FIG.10

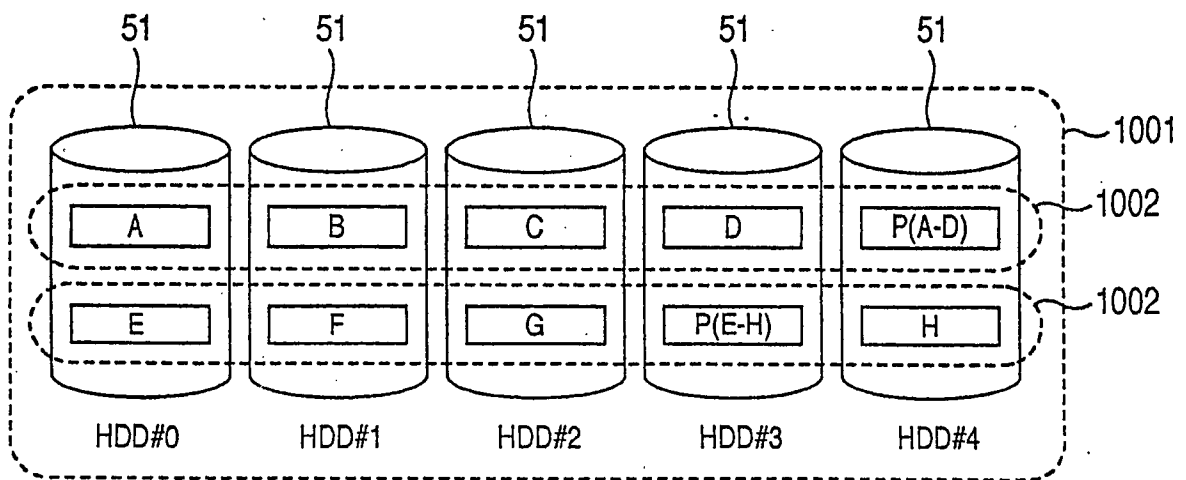


FIG.11

SEKTORNUMMER LAUFWERKS- NUMMER	LBA #1-128	LBA #129-256	LBA #257-384	...
HDD#0	0	0	0	...
HDD#1	1	0	0	...
HDD#2	0	1	0	...
...

1101

FIG.12

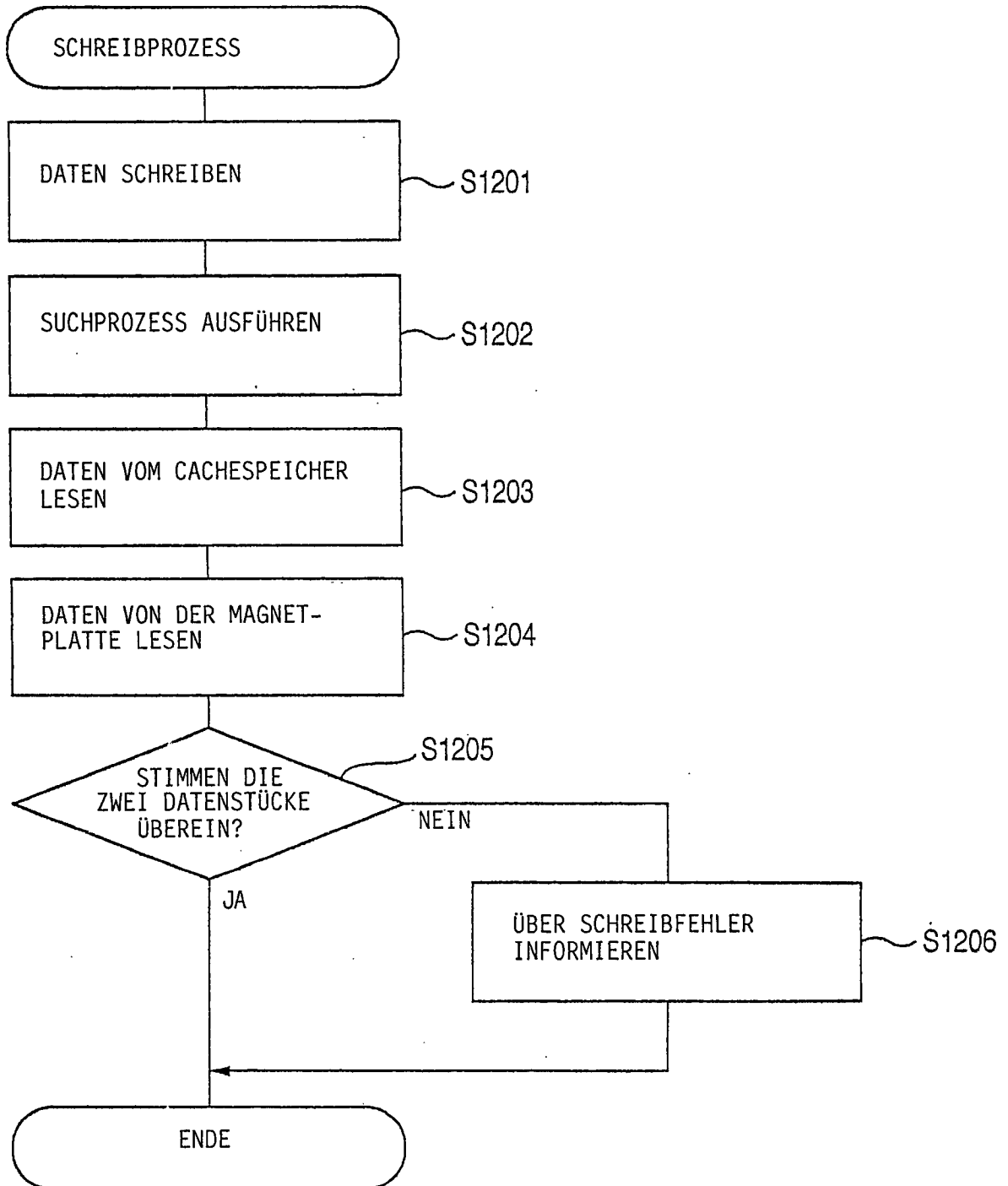


FIG.13

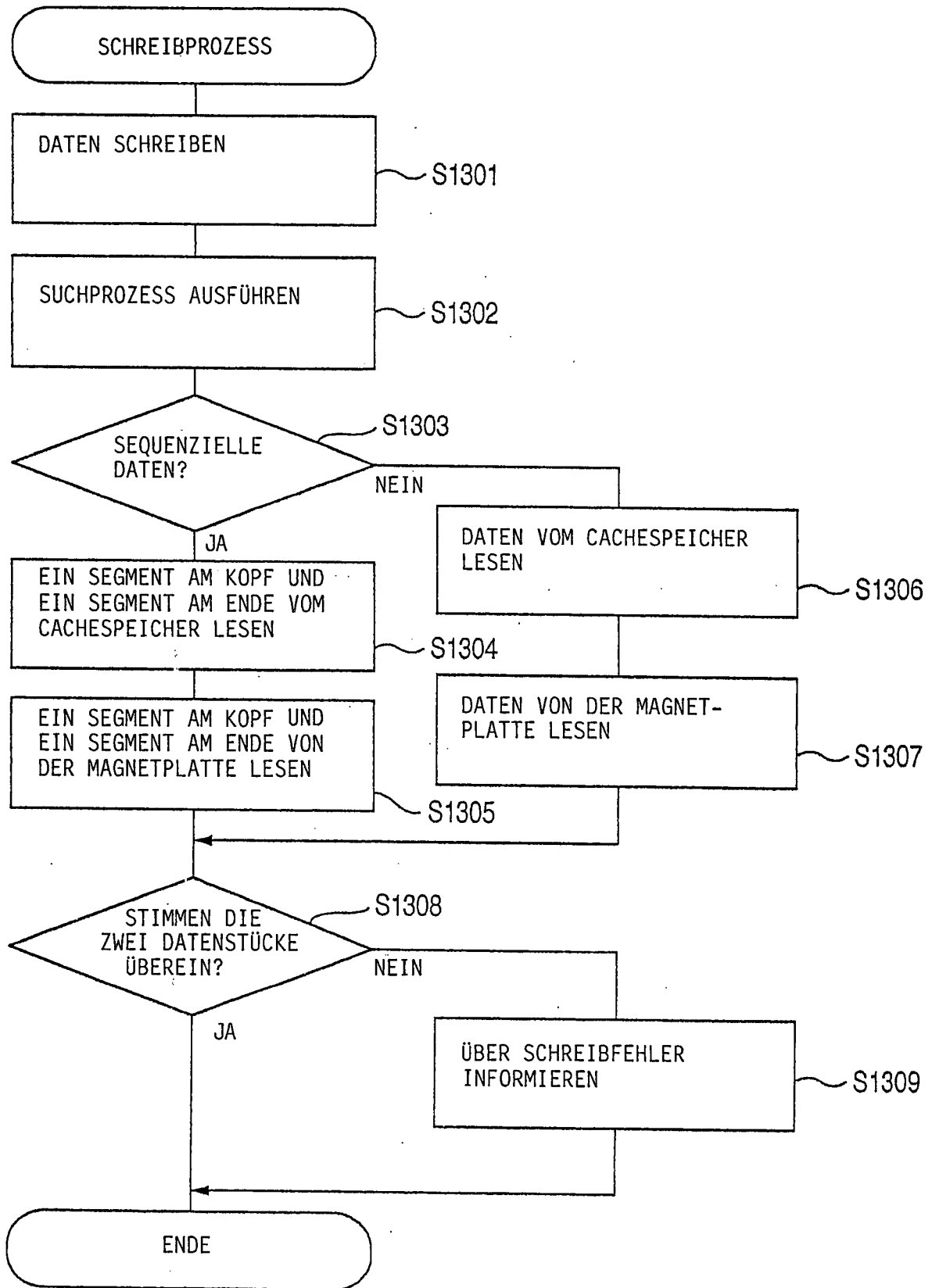


FIG.14

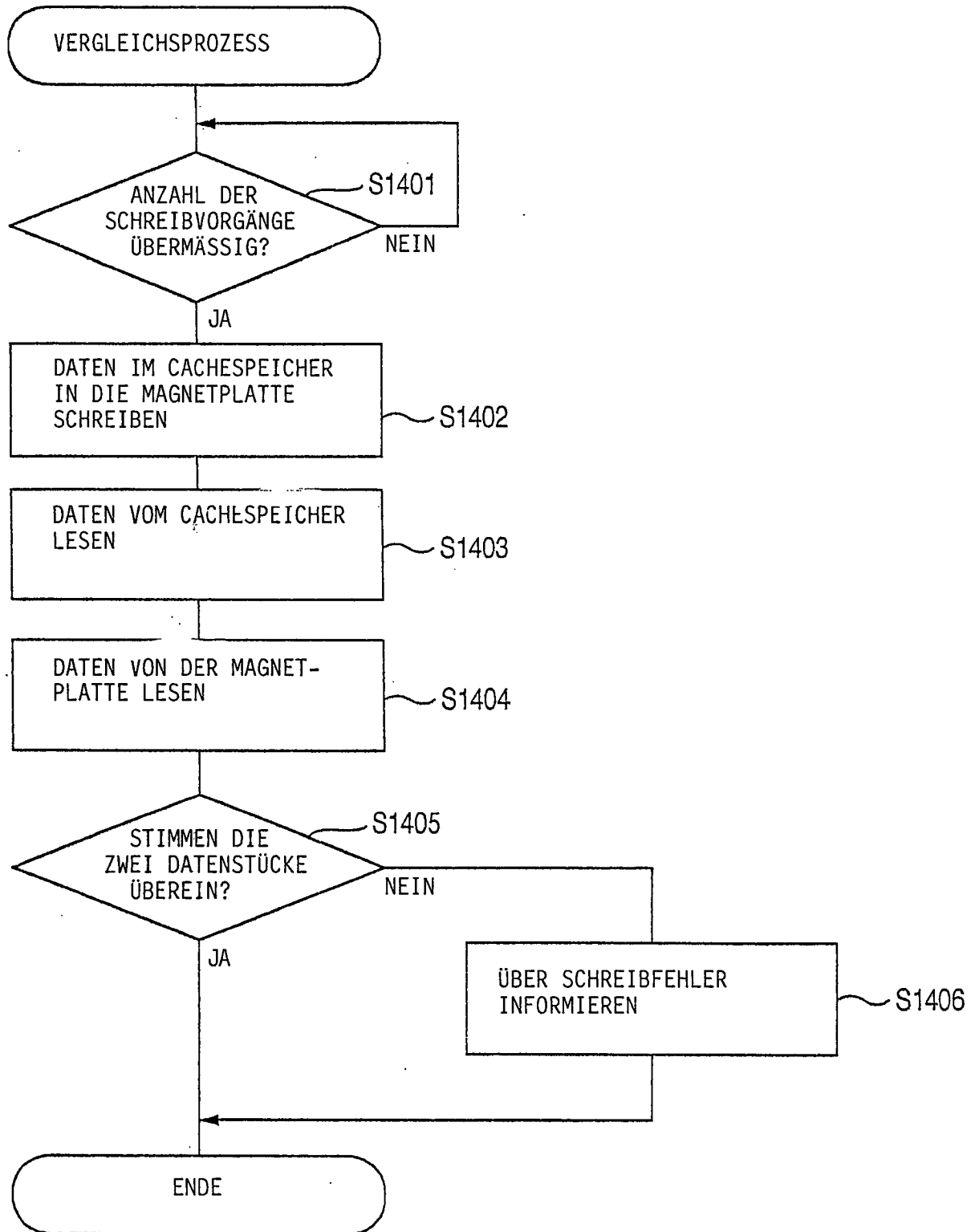


FIG.15

		LBA #1-128	LBA #129-256	LBA #257-384	...
HDD#0	KOPFNUMMER	#0	#0	#1	...
	VORLIEGEN E. AKTUAL.	1	0	0	...
HDD#1	KOPFNUMMER	#0	#0	#1	...
	VORLIEGEN E. AKTUAL.	0	0	0	...
...

1501

FIG.16

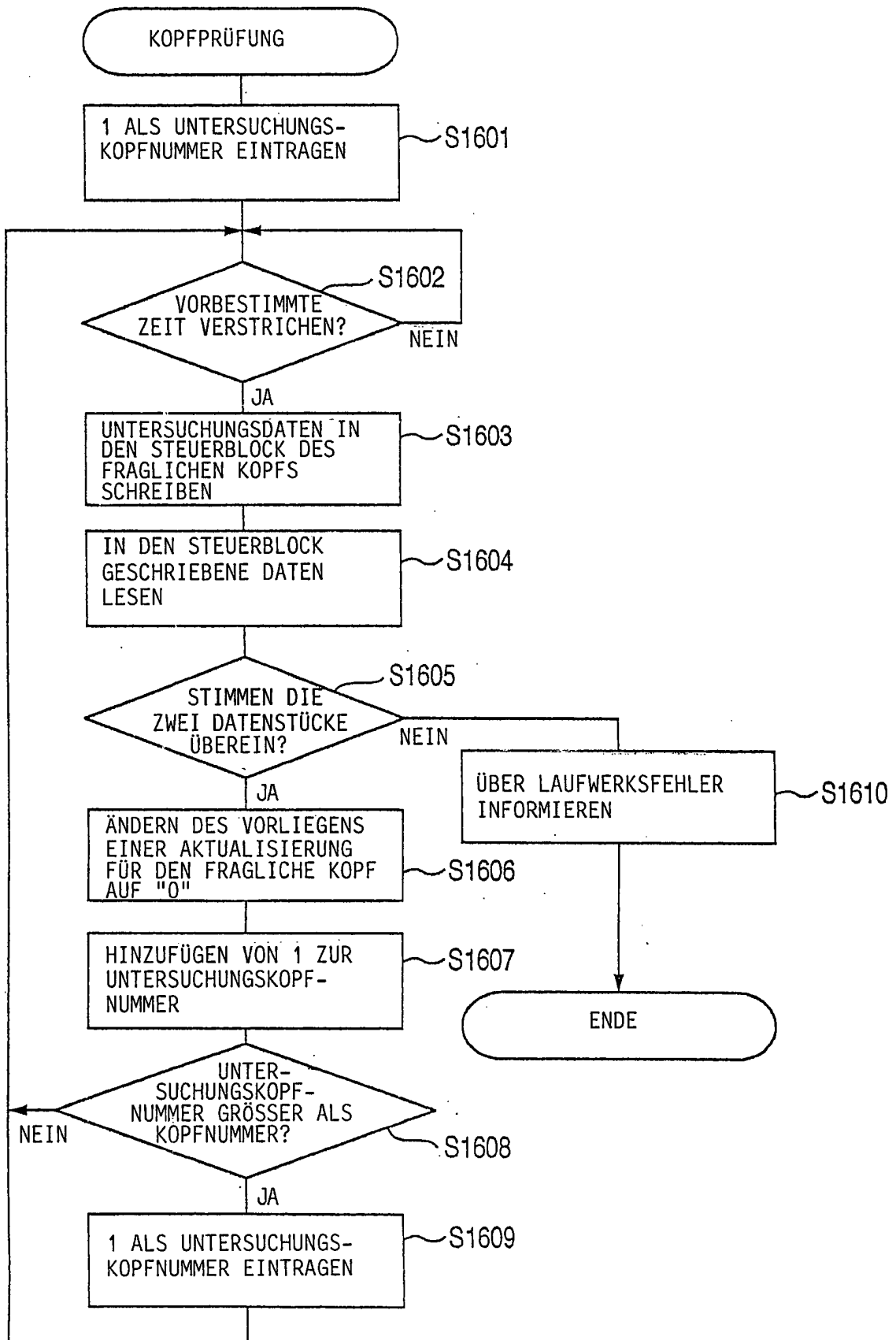


FIG.17

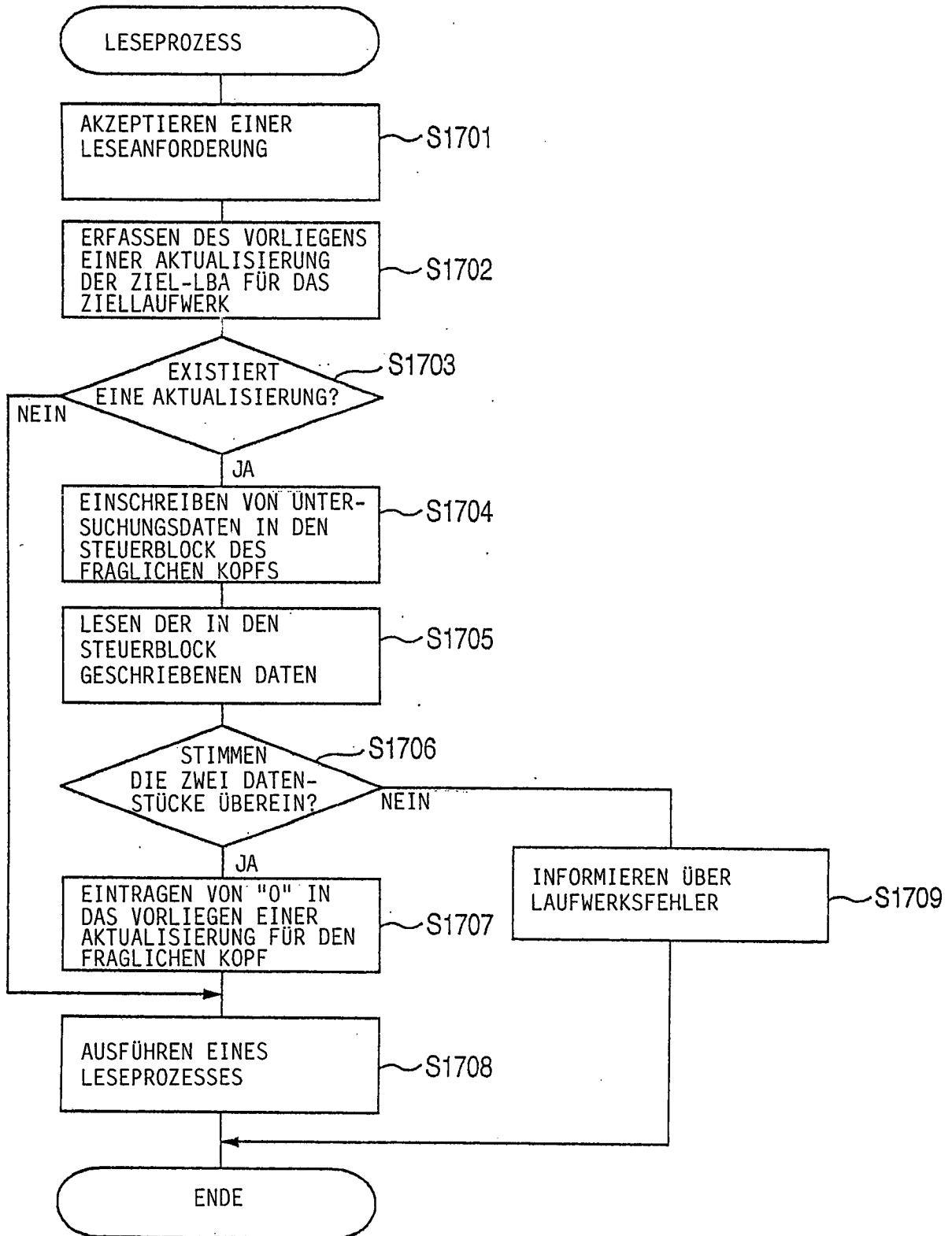


FIG.18

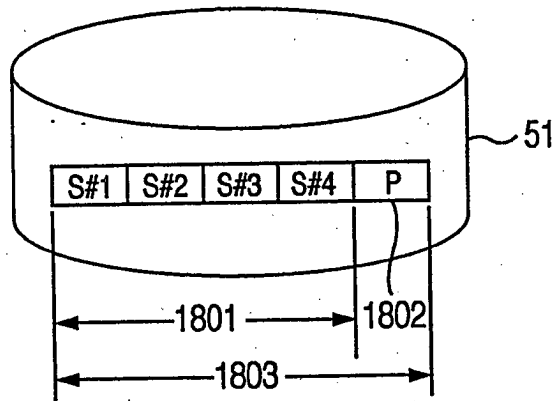


FIG.19

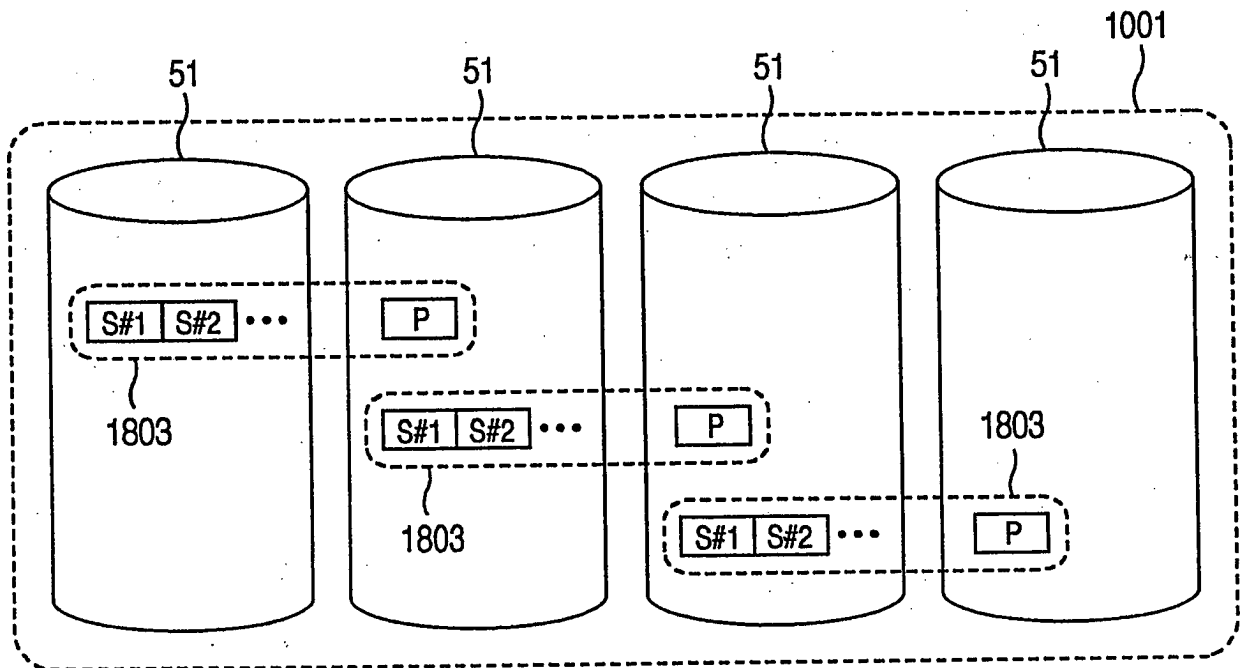


FIG.20

DATENEINHEIT	LAUFWERK	LAUFWERKS-LBA
000-129	#0	000-064
	#1	000-064
130-259	#2	000-064
	#0	065-129
⋮	⋮	⋮

2001

FIG.21

