(19) Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 44 22 786 B4** 2005.04.21

(12) **Patentschrift**

(21) Aktenzeichen: **P 44 22 786.8**
(22) Anmeldetag: **29.06.1994**
(43) Offenlegungstag: **12.01.1995**
(45) Veröffentlichungstag der Patenterteilung: **21.04.2005**

(51) Int Cl.[7]: **G06F 13/12**
G06F 11/18, G06F 12/08

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden.

(30) Unionspriorität:
5-162021 30.06.1993 JP

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

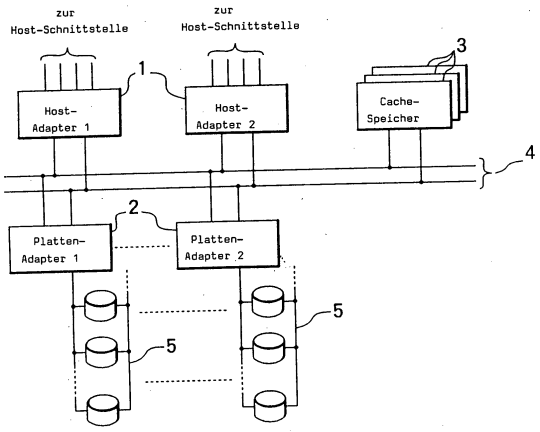
(74) Vertreter:
BEETZ & PARTNER Patentanwälte, 80538 München

(72) Erfinder:
Ninomiya, Tatuya, Odawara, JP; Masuzaki, Hidefumi, Hadano, JP; Kurosawa, Hiroyuki, Hiratsuka, JP; Takahashi, Naoya, Yokohama, JP; Inoue, Yasuo, Odawara, JP; Iwasaki, Hidehiko, Hiratsuka, JP; Hoshino, Masayuki, Odawara, JP; Isono, Soichi, Chigasaki, JP

(56) Für die Beurteilung der Patentfähigkeit in Betracht gezogene Druckschriften:
US 52 06 939
US 50 93 826
US 44 17 334
EP 04 45 479 A1

(54) Bezeichnung: **Speichersystem**

(57) Hauptanspruch: Speichersystem mit
mehreren ersten Logikeinheiten (1), die mit einer Host-Vorrichtung verbunden sind und Schnittstellen für diese Host-Vorrichtungen bilden;
einer Speichervorrichtung (5), die von der Host-Vorrichtung empfangene Information speichert;
mehreren zweiten Logikeinheiten (2), die mit der Speichervorrichtung (5) verbunden sind und für diese Speichervorrichtung (5) Schnittstellen bilden;
einer Cache-Speichervorrichtung (3), die Daten, die zwischen den mehreren ersten Logikeinheiten (1) und den mehreren zweiten Logikeinheiten (2) übertragen werden, temporär speichert; und
einem gemeinsamen Bus (4), der zwischen den mehreren ersten Logikeinheiten (1), den mehreren zweiten Logikeinheiten (2) und der Cache-Speichervorrichtung (3) verläuft und mit diesen funktional verbunden ist,
dadurch gekennzeichnet, daß
der Bus (4) einen E/A-Hochgeschwindigkeitsbus (26) zur Datenübertragung, der in zwei Kanäle unterteilt ist, die unterschiedliche Daten übertragen können, sowie einen Einzelkanal-Multiprozessor-Bus (25) zur Übertragung von Steuerinformation für die Datenübertragung, enthält,
wobei
dann, wenn im Multiprozessor-Bus (25) ein Fehler auftritt, einer der...



Beschreibung

Stand der Technik

[0001] Die vorliegende Erfindung betrifft ein Speichersystem gemäß dem Oberbegriff des Anspruch 1. Ein derartiges Speichersystem ist aus der EP-A-0 445 479 bekannt. Das Speichersystem weist eine mit einem Großcomputer, einem Computernetz oder dergleichen verbundene Speichervorrichtung wie etwa eine Magnetplatteneinrichtung, eine Magnetbandeinrichtung, eine Halbleiterspeichereinrichtung oder eine optische Plattenspeichereinrichtung auf.

[0002] Aus der JP-B-61-43742 (die dem US-Patent Nr. 4,636,946 entspricht) ist ein herkömmliches Speichersystem bekannt, das mit einem Großcomputer verbunden ist und bei dem Schnittstellen (Host-Adaptereinrichtungen) für eine höherrangige oder Host-Vorrichtung (Computer, CPU), ein Cache-Speicher sowie Schnittstellen (Platten-Adaptereinrichtungen) für eine Speichervorrichtung wie etwa eine Magnetplatteneinrichtung über Direktverbindungsleitungen (die im folgenden einfach Leitungen genannt werden) miteinander verbunden sind.

[0003] **Fig.** 20 zeigt eine schematische Darstellung einer Anordnung eines herkömmlichen Speichersystems. In der Zeichnung bezeichnen die Bezugszeichen **201-1** bis **201-n** Host-Adaptereinrichtungen, die mit mehreren Host-Computern (CPUs, Logikmodule, die mit einem Host-Computer verbunden sind), die Bezugszeichen **202-1** bis **202-n** Platten-Adaptereinrichtungen (Logikmodule, die mit einer Speichervorrichtung verbunden sind), die ihrerseits mit einer gemeinsam genutzten Groß-Plattenspeichereinrichtung **205** verbunden sind, das Bezugszeichen **203** einen von den mehreren Host-Adaptereinrichtungen gemeinsam genutzten Cache-Speicher und das Bezugszeichen **206** einen gemeinsam genutzten Verwaltungsspeicher. In einem herkömmlichen Speichersystem sind zwischen den Host-Adaptereinrichtungen **201-1** bis **201-n** und dem Cache-Speicher **203**, zwischen dem Cache-Speicher **203** und den Platten-Adaptereinrichtungen **202-1** bis **202-n**, zwischen die Host-Adaptereinrichtungen **201-1** bis **201-n** und dem Verwaltungsspeicher **206** und zwischen dem Verwaltungsspeicher **206** und den Platten-Adaptereinrichtungen **202-1** bis **202-n** Direktverbindungsleitungen **207-1** bis **207-n** bzw. **208-1** bis **208-n** eingefügt. Ferner sind mit den jeweiligen Host- und Plattenadaptereinrichtungen über entsprechende zugehörige Leitungen (nicht gezeigte) Wartungsprozessoren (SVPs) verbunden, die die Überwachungs- und Wartungsoperationen dieser Host- und Platten-Adaptereinrichtungen ausführen.

[0004] Da in dem obenbeschriebenen Stand der Technik die Direktverbindungsleitungen zwischen den Host-Adaptereinrichtungen (Logikmodule, die mit der Host-Vorrichtung verbunden sind) und der Host-Vorrichtung, zwischen den Platten-Adaptereinrichtungen (Logikmodule, die mit der Speichervorrichtung verbunden sind) und der Speichervorrichtung sowie dem Cache-Speicher (Cache-Speichermodul) verlaufen und mit diesen funktional verbunden sind, ist die Systemkonfiguration relativ kompliziert, außerdem sind die Host-Adaptereinrichtungen, der Cache-Speicher, die Platten-Adaptereinrichtungen und die Plattenspeichervorrichtung weniger gut erweiterbar, so daß es unmöglich ist, eine sogenannte skalierbare (erweiterbare oder reduzierbare) Systemkonfiguration zu verwirklichen. Ferner wird im Stand der Technik der Forderung keine Beachtung geschenkt, daß bei einem Ausfall einer Komponente und/oder einem Austausch einer Komponente während des Betriebs (ein Substrat oder ein Schaltungsteil wird ausgetauscht, während das System in Betrieb ist) die Mehrfachausstattung des Systems einen verminderten Betrieb (eines der mehrfach vorhandenen Untersysteme wird angehalten, während die anderen Untersysteme weiterbetrieben werden) ermöglicht. Daher besteht im obenbeschriebenen Stand der Technik das Problem, daß dann, wenn bei einem Ausfall einer Komponente diese Komponente ausgetauscht werden soll oder wenn ein System-Steuerprogramm aufgerüstet werden soll, das System hierzu vorübergehend angehalten werden muß.

[0005] Aus der EP-A-0 445 479 ist ein Massenspeichersystem bekannt. Es weist Host-Adapter, Disk-Adapter und Cache-Adapter auf. Diese Adapter sind wahlweise mit einem oder mit beiden von zwei vorhandenen Bussen verbunden. Die Busse werden immer unabhängig voneinander betrieben.

[0006] Aus der US 5 206 939 ist ein System und ein Verfahren für das Platten-Mapping und für Datenwiedergewinnung bekannt. Host-Adapter, Platten-Adapter sowie ein Cache-Speicher sind mit einem Bus verbunden.

[0007] Aus der US 5 093 826 ist ein Verfahren zum Betreiben einer zentralen Mehrprozessor-Steuerungseinheit eines Schaltsystems bekannt. In diesem System sind verschiedene Komponenten, unter anderem das Bussystem, redundant ausgelegt.

[0008] Die US 4 417 334 beschreibt ein Datenverarbeitungssystem mit einem zweikanaligen Systembus. Einer der Kanäle kann ausschließlich genutzt werden, wenn der andere Kanal ausfällt. Außerdem können über die zwei Kanäle gleichzeitig mehrere Komponenten miteinander kommunizieren.

Aufgabenstellung

[0009] Aufgabe der Erfindung ist es, ein Speichersystem anzugeben, das auch dann, wenn auf den Bus Fehler auftreten, den Datenverkehr in zumindest

vermindertem Betrieb aufrechterhalten kann.

[0010] Diese Aufgabe wird gemäß den Merkmalen des Anspruchs 1 gelöst. Abhängige Ansprüche sind auf bevorzugte Ausführungsformen der Erfindung gerichtet.

[0011] Es wird ein Speichersystem geschaffen, das versehen ist mit mehreren Logikeinheiten, die mit einer Host-Vorrichtung verbunden sind und für diese Host-Vorrichtung Schnittstellen bilden, einer Speichereinheit, mehreren Logikeinheiten, die mit der Speichereinheit verbunden sind und für die Speichereinheit Schnittstellen bilden, und Cache-Speichern (die von den mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten und von den mit der Speichereinheit verbundenen mehreren Logikeinheiten gemeinsam genutzt werden), die zwischen diesen Vorrichtungen und Einheiten zu übertragende Daten temporär speichern. Die mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, die mit der Speichereinheit verbundenen mehreren Logikeinheiten und die Cache-Speicher sind über einen von diesen Vorrichtungen und Einheiten gemeinsam genutzten Bus (gemeinsamer Bus) miteinander verbunden. Im Ergebnis kann ein skalierbares System erhalten werden, das die Erweiterung und die Modifikation der mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, der mit den Speichern verbundenen mehreren Logikeinheiten und der Cache-Speicher ermöglicht, indem diese lediglich zum gemeinsamen Bus hinzugefügt werden oder modifiziert werden, während sie an den Bus angeschlossen sind, so daß eine Aufrüstung des Systems mittels einer Systemerweiterung einfach erhalten werden kann.

[0012] Da die mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, die mit den Speichern verbundenen mehreren Logikeinheiten und die Cache-Speicher jeweils doppelt vorhanden sind und der gemeinsame Bus zwischen diesen Logikeinheiten und dem Speicher in zwei Kanäle unterteilt ist, kann selbst bei einem Ausfall einer dieser Einheiten die andere Einheit verwendet werden, um einen verminderten Betrieb auszuführen. Hierbei wird in den gemeinsam genutzten Speicher Information geschrieben, die den Status des verminderten Betriebs bei einem Ausfall angibt.

[0013] Da in diesem Fall jede der mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, jede der mit der Speichereinheit verbundenen mehreren Logikeinheiten und jeder der Cache-Speicher mit einer Anschlußeinrichtung versehen ist, die eine Ersetzung während des Betriebs ermöglicht, läßt das System eine Wartung und einen Eingriff für einen Austausch eines fehlerhaften Teils ebenso zu wie die Hinzufügung und/oder die Modifikation von Komponenten zur Systemerweiterung, wobei in keinem Fall der Systembetrieb angehalten werden muß.

[0014] Weitere Aufgaben, Merkmale und Vorteile der Erfindung sind in den Neben- und Unteransprüchen angegeben, die sich auf bevorzugte Ausführungsformen der vorliegenden Erfindung beziehen.

Ausführungsbeispiel

[0015] Die Erfindung wird im folgenden anhand bevorzugter Ausführungsformen mit Bezug auf die Zeichnungen näher erläutert; es zeigen:

[0016] **Fig.** 1 auf abstrakte Weise eine schematische Anordnung eines Speichersystems gemäß einer Ausführungsform der vorliegenden Erfindung;

[0017] **Fig.** 2 eine detaillierte Anordnung des Speichersystems der Ausführungsform der vorliegenden Erfindung;

[0018] **Fig.** 3 eine Darstellung zur Erläuterung des Datenstroms und des Datenformats in der in **Fig.** 2 gezeigten Anordnung;

[0019] **Fig.** 4 eine Ansicht der äußeren Erscheinungsform der Ausführungsform der vorliegenden Erfindung;

[0020] **Fig.** 5A eine Vorderansicht einer Steuereinheit in dem System gemäß der Ausführungsform der vorliegenden Erfindung, wobei ein Beispiel gezeigt ist, wie die Steuereinheit montiert ist;

[0021] **Fig.** 5B eine Seitenansicht der Steuereinheit des Speichersystems gemäß der Ausführungsform der vorliegenden Erfindung;

[0022] **Fig.** 6A eine Vorderansicht einer Matrix-Platteneinheit in dem System gemäß der Ausführungsform der vorliegenden Erfindung, in der ein Beispiel für die Montage der Matrix-Platteneinheit gezeigt ist;

[0023] **Fig.** 6B eine Seitenansicht der Matrix-Platteneinheit in dem System gemäß der Ausführungsform der vorliegenden Erfindung;

[0024] **Fig.** 7 ein Montageschaltbild eines Logikeinheit-Rahmenteils in dem System gemäß der Ausführungsform der vorliegenden Erfindung;

[0025] **Fig.** 8 eine auseinandergezogene, perspektivische Ansicht des Logikeinheit-Rahmenteils im System gemäß der Ausführungsform der vorliegenden Erfindung, in der die Montage des Logikeinheit-Rahmenteils gezeigt ist;

[0026] **Fig.** 9 eine Konfiguration der Software, die in der Ausführungsform der vorliegenden Erfindung verwendet wird;

[0027] **Fig.** 10 ein Diagramm zur Erläuterung der

Datenströme und der gemeinsam genutzten Software-Funktionen in der Ausführungsform der vorliegenden Erfindung;

[0028] **Fig.** 11A ein Diagramm zur Erläuterung eines gemeinsamen, doppelt vorhandenen Busses in der Ausführungsform der vorliegenden Erfindung;

[0029] **Fig.** 11B ein Diagramm zur Erläuterung des verminderten Betriebs in der Ausführungsform der vorliegenden Erfindung;

[0030] **Fig.** 12 ein Diagramm zur Erläuterung des Duplexbetriebs und des verminderten Betriebs eines jeden Teils im System gemäß der Ausführungsform der vorliegenden Erfindung;

[0031] **Fig.** 13 ein Diagramm zur Erläuterung des Mehrfachbetriebs und des verminderten Betriebs eines Leistungsversorgungssystems in dem System gemäß der Ausführungsform der vorliegenden Erfindung;

[0032] **Fig.** 14 in tabellarischer Form mehrere Konfigurationen einer aus mehreren einzelnen Magnetplattenvorrichtungen aufgebauten Matrix-Platteneinheit;

[0033] **Fig.** 15 in graphischer Form die Zugriffszeiten der der Matrix-Platteneinheiten von **Fig.** 14 in Abhängigkeit von der Anzahl der ausgegebenen E/A-Befehle;

[0034] **Fig.** 16 den Aufbau einer kleinen Matrix-Platteneinheit, die mit einem Hochleistungs-Cache-Speicher mit großer Kapazität versehen ist;

[0035] **Fig.** 17 den Aufbau einer großen Matrix-Platteneinheit, die mit einem Hochleistungs-Cache-Speicher mit großer Kapazität versehen ist;

[0036] **Fig.** 18 den Aufbau eines fehlertoleranten Hochleistungs-Serversystems;

[0037] **Fig.** 19 den Aufbau eines billigen Serversystems; und

[0038] **Fig.** 20 die bereits erwähnte schematische Anordnung eines Speichersystems des Standes der Technik.

[0039] **Fig.** 1 ist eine abstrakte Darstellung zur Erläuterung des Aufbaus des erfindungsgemäßen Speichersystems. Mit Bezug auf **Fig.** 1 wird eine entsprechende Ausführungsform dieses erfindungsgemäßen Speichersystems erläutert.

[0040] Das Bezugszeichen **1** bezeichnet eine Host-Adaptereinrichtung, der durch einen Logikmodul gegeben ist, der mit einer Host-Zentraleinheit (Host-Computer) verbunden ist, das Bezugszeichen **2** bezeichnet eine Platten-Adaptereinrichtung, die durch einen Logikmodul gegeben ist, der mit einem Speichermedium verbunden ist, das Bezugszeichen **3** bezeichnet eine Cache-Speicher-Baueinheit (Cache-Speicher-Modul) zur temporären Speicherung von Daten, die zwischen den beiden obengenannten Modulen übertragen werden sollen, das Bezugszeichen **4** bezeichnet einen gemeinsamen Bus zur Steuerung der Datenübertragung zwischen der Host-Adaptereinrichtung **1**, der Platten-Adaptereinrichtung **2** und der Cache-Speicher-Baueinheit **3**, das Bezugszeichen **5** bezeichnet eine Gruppe von Magnetplatten (die im folgenden lediglich als "Matrix-Platteneinheit" bezeichnet wird), die ein Speichermedium bildet, das vertikal und horizontal in einer Matrix angeordnet ist. Die Host-Adaptereinrichtung **1** enthält eine Einrichtung zur Umwandlung des Datenformats und des Adressenformats auf Seiten der Host-Schnittstelle in Daten- und Adressenformate auf Seiten der Speichermedium-Schnittstelle, sowie einen doppelt vorhandenen Mikroprozessor für die Ausführung der Steuerung und der Verwaltung dieser Einrichtung. Die Platten-Adaptereinrichtung **2** besitzt eine Adressenverarbeitungsfunktion zur Speicherung der Daten in dem Speichermedium, eine Funktion zur Erzeugung redundanter Daten zur Sicherung der Speicherdaten, eine Funktion zur Erkennung von Information bezüglich der Struktur des Speichermediums sowie einen doppelt vorhandenen Mikroprozessor für die Ausführung der Steuerung und der Verwaltung dieser Funktionen.

[0041] In **Fig.** 1 schreibt die Host-Adaptereinrichtung von der Host-Vorrichtung (CPU) empfangene Schreibdaten und Information bezüglich der Verwaltung der Schreibdaten einmal über den gemeinsamen Bus **4** in die Cache-Speicher-Baueinheit **3** und meldet nach Abschluß der Schreiboperation an die Host-Vorrichtung den Abschluß dieser Schreiboperation. In einer darauf folgenden Leerlaufzeit liest die Platten-Adaptereinrichtung **2** auf der Grundlage der Verwaltungsinformation der Cache-Speicher-Baueinheit **3** Daten aus der Cache-Speicher-Baueinheit **3** aus.

[0042] Wenn die Host-Adaptereinrichtung **1** von der Host-Vorrichtung einen Datenlesebefehl empfängt und die entsprechenden Daten in der Cache-Speicher-Baueinheit **3** vorhanden sind, führt sie keine Leseoperation dieser Daten aus der Matrix-Platteneinheit **5** aus, sondern überträgt die in der Cache-Speicher-Baueinheit **3** vorhandenen Daten zur Host-Vorrichtung. Wenn andererseits die Daten nicht in der Cache-Speicher-Baueinheit **3** vorhanden sind, schreibt die Platten-Adaptereinrichtung **2** die Daten und deren Verwaltungsinformation von der Matrix-Platteneinheit **5** über den gemeinsamen Bus **4** in die Cache-Speicher-Baueinheit **3**. Die Host-Adaptereinrichtung **1** nimmt auf die Verwaltungsinformation

Bezug, liest die Daten aus der Cache-Speicher-Baueinheit **3** aus und überträgt sie zur Host-Vorrichtung.

[0043] Die Anzahl der Host-Adaptereinrichtungen **1**, die Anzahl der Platten-Adaptereinrichtungen **2** und die Anzahl der Cache-Speicher in der Cache-Speicher-Baueinheit **3**, die an den gemeinsamen Bus **4** angeschlossen sind, können jeweils beliebig geändert werden. Wenn die Anzahl der angeschlossenen Host-Adaptereinrichtungen **1** geändert wird, wird auch die Anzahl der mit den Host-Vorrichtungen verbundenen Busse geändert, so daß die Datenübertragungskapazität zur Host-Vorrichtung gesteigert werden kann. Wenn die Anzahl der angeschlossenen Platten-Adaptereinrichtungen **2** geändert wird, wird auch die Anzahl der mit dem Speichermedium verbundenen Busse geändert, so daß die Kapazität der Schreib-/Lese-Vorgänge in das bzw. aus dem Speichermedium gesteigert werden kann. Gleichzeitig kann auch die Anzahl der Speichermedien erhöht werden. Wenn die Anzahl der angeschlossenen Cache-Speicher in der Cache-Speicher-Baueinheit **3** geändert wird, wird die Kapazität der als temporärer Datenspeicher dienenden Cache-Speicher-Baueinheit ebenfalls geändert, so daß das Verhältnis der Kapazität der Cache-Speicher-Baueinheit zur Gesamtkapazität des Speichermediums erhöht werden kann, was dadurch zum Ausdruck kommt, daß eine skalierbare Systemanordnung verwirklicht werden kann, bei der die Wahrscheinlichkeit erhöht ist, daß die Daten, auf die von der Host-Vorrichtung zugegriffen wird, in den Cache-Speichern vorhanden sind (was auch als "Cache-Trefferrate" bezeichnet wird).

[0044] **Fig.** 2 zeigt eine genaue Anordnung des abstrakten Diagramms von **Fig.** 1. In **Fig.** 2 sind nur eine der mehreren Host-Adaptereinrichtungen **1** und nur eine der mehreren Platten-Adaptereinrichtungen **2** gezeigt, während die übrigen Adaptereinrichtungen weggelassen sind.

[0045] Die Host-Adaptereinrichtung **1** enthält einen Signalwandler **6**, der ein optisches Signal von der Host-Schnittstelle in ein elektrisches Signal umwandelt, einen Formatwandler **7**, der das Format der Daten der Host-Vorrichtung in ein für die Matrix-Platteneinheit **5** geeignetes Format umwandelt, sowie eine Datenübertragungs-Steuereinrichtung **8**, die die Datenübertragung zum gemeinsamen Bus **4** steuert und einen Pufferspeicher enthält, der ein Daten-Übertragungspaket speichert. Die Host-Adaptereinrichtung **1** enthält außerdem einen Bustreiber **9** eines Typs mit kleiner Stromamplitude (dieser Treiber wird im folgenden mit BTL bezeichnet), der einen Komponentenaustausch während des Betriebs ermöglicht. Es wird darauf hingewiesen, daß "BTL" ein eingetragenes Warenzeichen von National Semiconductor Ltd. ist.

[0046] Eine Datenübertragungsanforderung vom Host-Computer wird an einen Mikroprozessor **10** (der im folgenden lediglich als MP 10 bezeichnet wird) geschickt, wobei die Datenübertragungssteuerung in der Host-Adaptereinrichtung **1** der Steuerung des MP **10** unterstellt wird.

[0047] Um eine hohe Zuverlässigkeit bei Erfassung eines Ausfalls des MP **10** sicherzustellen, besitzt die Host-Adaptereinrichtung **1** zusätzlich zum MP **10** einen MP **10'**, der dem MP 10 ähnlich ist. Eine Prüfeinrichtung **11** führt einen Vergleich zwischen dem MP **10** und dem MP **10'** aus.

[0048] Eine Boot-Vorrichtung **12**, in der ein Programm zur Steuerung des MP **10** gespeichert ist, verwendet einen wiederbeschreibbaren Blitzlicht-Speicher mit großer Kapazität. Der MP **10** erzeugt gegebenenfalls eine Kopie des Steuerprogramms in einem lokalen Speicher **13**, wobei die Speicherzugriffszeit auf den MP **10** kurz ist bzw. die Zugriffsgeschwindigkeit hoch ist. Ein Abschnitt **29**, der in **Fig.** 2 von einer Strichlinie umgeben ist, bezeichnet einen Kanaladapter-Modul. Jede Host-Adaptereinrichtung **1** besitzt zwei solche Module **29**.

[0049] Die Platten-Adaptereinrichtung **2** enthält einen Pufferspeicher **14**, in dem Sektoreinheiten von in die Matrix-Platteneinheit **5** zu schreibenden Daten gespeichert sind, einen Datensteuerpuffer **15**, der den Pufferspeicher **14** und die Datenübertragungssteuerung steuert, einen Generator **16** für redundante Daten, der das Schreiben von Daten in die Matrix-Platteneinheit **5** sicherstellt, sowie einen Initiator **17** (SCSI-Master-Schnittstelle für die Matrix-Platteneinheit **5** (Ziel).

[0050] Die Datenübertragungssteuerung in der Plattenadaptereinrichtung **2** erfolgt unter der Steuerung eines peripheren MP (der den MP **10**, den MP **10'**, die Prüfeinrichtung **11**, die Boot-Vorrichtung **12** und den lokalen Speicher **13** sowie ein Steuerprogramm für die Platten-Adaptereinrichtung **2** enthält), der im wesentlichen die gleiche Struktur wie die Host-Adaptereinrichtung **1** besitzt.

[0051] Obwohl in **Fig.** 2 für die Matrix-Platteneinheit **5** nur vier Platten (Ziele) gezeigt sind, ist die Matrix-Platteneinheit tatsächlich aus beispielsweise 4 (horizontal) × 4 (vertikal) Platten aufgebaut, die der einzigen Platten-Adaptereinrichtung **2** zugeordnet sind. Jede der horizontalen Reihen bildet eine Fehlerkorrekturgruppe (ECC-Gruppe), die beispielsweise aus drei Datenplatten und einer einzigen Paritätsplatte aufgebaut ist. Ferner sind zwischen einem Satz solcher Matrix-Platteneinheiten **5** und der Platten-Adaptereinrichtung mehrere Busse vorgesehen, die mit wenigstens zwei oder mehr Platten-Adaptereinrichtungen **2** verbunden sind, wie später beschrieben wird. Außerdem sind zwischen der CPU und der Host-Adaptereinrichtung **1** mehrere Busse vorgese-

hen, die mit wenigstens zwei Host-Adaptereinrichtungen **1** verbunden sind. Wenn in einer der Host-Adaptereinrichtungen ein Fehler auftritt, kann ein Zugriff von dieser CPU auf dieselbe Matrix-Platteneinheit **5** über die andere Host-Adaptereinrichtung **1** und/oder die andere Platten-Adaptereinrichtung **2** erfolgen.

[0052] Die Cache-Speicher-Baueinheit **3** enthält einen gemeinsam genutzten Speicher **18**, in dem verschiedene Arten von Verwaltungsinformationen gespeichert sind und auf die gemeinsam durch die MPs **10** der Adaptereinrichtungen zugegriffen werden kann, eine gemeinsam genutzte Speichersteuereinrichtung **19**, einen Cache-Speicher **20** sowie eine Cache-Speicher-Steuereinrichtung **21**. Jede der beiden Speicher-Steuereinrichtungen **19** und **21** besitzt eine ECC-Erzeugungsschaltung für die Speicher-Schreibdaten und eine Inspektions-/Korrekturschaltung für gelesene Daten.

[0053] Die gesamte Cache-Speicher-Baueinheit **3** besitzt eine Cache-Kapazität von bis zu 1 GB, wobei im Hinblick auf einen doppelt vorhandenen Cache-Speicher in dem System zwei Cache-Speicher-Baueinheiten **3** angebracht sind.

[0054] Wenn gewünscht ist, die Kapazität der Cache-Speicher weiter zu erhöhen, wird anstelle der Cache-Speicher-Baueinheit **3** (oder aber zusätzlich zu der Cache-Speicher-Baueinheit **3**) eine Cache-Anschlußbaueinheit **22** angebracht, die mit einer Cache-Einheit **24** über ein Kabel **23** verbunden ist, das dem Anschluß hinterer Ebenen (d.h. Anbringungsplatten für den Einschub von Substraten) dient. Die Cache-Einheit **24** besitzt eine Cache-Speicher-Baueinheit **24a** mit Cache-Speichern **20'**, eine Cache-Anschlußbaueinheit **24b**, die mit dem Kabel **23** zur Verbindung von Zwischenebenen mit den hinteren Ebenen verbunden ist, sowie einen E/A-Hochgeschwindigkeitsbus **24c**, der mit der Cache-Speicher-Baueinheit **24a** und mit der Cache-Anschlußbaueinheit **24b** verbunden ist, um dazwischen eine Hochgeschwindigkeits-Datenübertragung auszuführen. Die Cache-Speicher **20'** in der zusätzlich vorgesehenen Einheit **24** sind so beschaffen, daß auf sie über die Cache-Anschlußbaueinheit **22** und das Kabel **23** zugegriffen werden kann. Die Bereitstellung der Cache-Einheit **24** ermöglicht eine erhöhte Cache-Kapazität bis zu 2 × 8 GB. **Fig.** 2 zeigt den Fall, in dem zusätzlich zu den zwei Cache-Speicher-Baueinheiten **3** die Cache-Anschlußbaueinheit **22** angebracht ist, die über das Kabel **23** mit der Cache-Einheit **24** verbunden ist.

[0055] Die oben beschriebenen Host-Adaptereinrichtungen **1**, Platten-Adaptereinrichtungen **2** und Cache-Speicher-Baueinheiten **3** sind miteinander über den gemeinsamen Bus **4** verbunden, der einen Multiprozessor-Bus **25** (der im folgenden mit M-Bus bezeichnet wird), durch den der MP **10** einer jeden Adaptereinrichtung auf den gemeinsamen Speicher zugreift, sowie einen (schnellen) E/A-Hochgeschwindigkeitsbus **26** (der im folgenden mit S-Bus bezeichnet wird) umfaßt.

[0056] Der S-Bus **26** wird normalerweise mit parallelem Zweikanalbetrieb und einer Breite von 64 Bit betrieben, so daß dann, wenn in einem der beiden Kanäle ein Fehler auftritt, der andere einen verminderten Betrieb ausführt. Wenn in dem M-Bus **25** ein Fehler auftritt, wird einer der beiden Kanäle des S-Busses **26** als M-Bus verwendet, während der verbleibende Kanal als S-Bus verwendet wird.

[0057] Weiterhin wird das Element BTL **9** als Schnittstelle mit dem gemeinsamen Bus **4** verwendet, wobei dieses Element BTL **9** einen Austausch von Komponenten während des Betriebs zuläßt (bei einem Einfügen oder Entnehmen einer Komponente kann diese Einfüge- oder Entnahmeoperation bei geringer Beanspruchung der eingefügten oder entnommenen Komponente ausgeführt werden, so daß diese Operation des Einfügens oder Entnehmens einer Komponente bei aktivem System ausgeführt werden kann). Wenn in der Host-Adaptereinrichtung **1** ein Fehler auftritt, wird diese Host-Adaptereinrichtung gesperrt. D.h., daß das System den entsprechenden fehlerhaften Bus sperrt und statt dessen die andere normal arbeitende Host-Adaptereinrichtung **1** verwendet, um den Zugriff auf die Matrix-Platteneinheit **5** durch die Host-Vorrichtung (dieselbe CPU) fortzusetzen. Der Wartungsoperator entnimmt die Host-Adaptereinrichtung **1**, in der während des Systembetriebs ein Fehler aufgetreten ist und die in den gesperrten Zustand versetzt worden ist. Danach wird eine normal arbeitende Host-Adaptereinrichtung **1** in das System eingesetzt, außerdem wird von einem Wartungsprozessor **27** (der im folgenden als SVP bezeichnet wird) in ein LAN **28** ein Wiederherstellungsbefehl eingegeben, so daß das System die Operation der ausgetauschten Host-Adaptereinrichtung **1** in der Weise prüft, daß bei normal arbeitender Host-Adaptereinrichtung das System den gesperrten Bus wiederherstellt, um einen ununterbrochenen Betrieb auszuführen. In der Zeichnung bezeichnet das Bezugszeichen LANC eine LAN-Steuereinrichtung (SVP-Schnittstellensteuereinrichtung). Der SVP **27** ist zu Überwachungs- und Wartungszwecken auf ähnliche Weise mit der anderen Host-Adaptereinrichtung **1** und der Platten-Adaptereinrichtung **2** verbunden.

[0058] Wenn das Steuerprogramm einer jeden Adaptereinrichtung modifiziert werden muß, schreibt der SVP **27** die Inhalte des Steuerprogramms über das LAN **28** erneut in die Boot-Vorrichtung **12**, um eine Aufrüstung ohne Anhalten des Betriebs zu ermöglichen.

[0059] Wenn daher eine Aufrüstung des Steuerpro-

gramms des Systems gewünscht ist, wird zunächst das Steuerprogramm der Boot-Vorrichtung **12**, der Host-Adaptereinrichtung **1** und der Platten-Adaptereinrichtung **2** für die Programmaufrüstung neu geschrieben. Nach Abschluß des Neuschreibvorgangs des Steuerprogramms wird die Adaptereinrichtung mit dem neugeschriebenen Steuerprogramm zurückgesetzt, um den Austausch des Systemsteuerprogramms auszuführen.

[0060] **Fig.** 3 ist ein Diagramm zur Erläuterung des Datenstroms und der gesicherten Daten in der Anordnung von **Fig.** 2.

[0061] Wenn von der Host-Vorrichtung Daten in die Matrix-Platteneinheit **5** geschrieben werden, wird eine Information bezüglich einer physikalischen Adresse (die im folgenden lediglich mit PA bezeichnet wird) in einen Speicherbereich als Schreibziel geschrieben, anschließend werden Daten (CKD-Format, Zählkenndaten-Format) sowie ein CRC-Code (zyklisch redundanter Prüfcode) beispielsweise von einem ESCON (Handelsname der IBM Ltd) geschickt. Diese optischen Signale werden in dem Signalwandler **6** in elektrische Signale umgewandelt, außerdem wird eine Parität erzeugt. Der Formatwandler **7** wandelt ein Datenformat in ein FBA-Format (Format mit fester Blockarchitektur) um und fügt einen LRC-Code (longitudinal redundanter Prüfcode) hinzu, verwendet die PA als Teil der Daten, um eine logische Adresse zu erzeugen (die im folgenden mit LA bezeichnet wird), fügt die Paritäten an sämtliche dieser Informationen an und schickt diese auf den S-Bus **26**.

[0062] Die Cache-Speicher-Baueinheit **3** fügt einen Fehlerkorrekturcode ECC an die vom S-Bus **26** empfangenen Daten an und schreibt sie in den Cache-Speicher **20**.

[0063] Die Platten-Adaptereinrichtung **2** fügt ferner an die vom S-Bus **26** empfangenen Daten einen CRC-Code an und schickt die mit dem Code versehenen Daten über die SCSI-Schnittstelle zur Matrix-Platteneinheit **5**, um den ECC jeder Magnetplatteneinheit hinzuzufügen und die Schreibdaten zu sichern.

[0064] Selbst bei einem Lesen der Daten von der Matrix-Platteneinheit **5** werden die gelesenen Daten auf ähnliche Weise auf der Grundlage eines jedes Prüfcodes untersucht und korrigiert, um deren Zuverlässigkeit zu verbessern.

[0065] Wie oben erläutert worden ist, liegt der Prüfcode doppelt vor, d.h., daß eine horizontale Prüfung für jede vorgegebene Länge in Datenlängsrichtung ausgeführt wird, während eine vertikale Prüfung (beispielsweise für jede Byte-Einheit) in vertikaler Datenrichtung (Breitenrichtung) ausgeführt wird, so daß einer der doppelt vorhandenen Prüfcodes zuverlässig in Form von Daten zwischen den Übertragungsbereichen (die in der Zeichnung von einer Strichpunktlinie umgeben sind) übertragen wird, um den in Form von Daten übertragenen Prüfcode mit einem aus den übertragenen Daten erzeugten Prüfcode zu vergleichen, so daß die Daten definitiv gesichert werden.

[0066] In **Fig.** 4 ist das äußere Erscheinungsbild einer Vorrichtung gezeigt, in der das in Verbindung mit **Fig.** 1 erläuterte skalierbare Speichersystem implementiert werden kann. Diese Vorrichtung enthält zwei Einheiten, nämlich eine Steuereinheit **41** zur Steuerung der Matrix-Platteneinheit **5** und eine Matrixeinheit **42**, in der die Matrix-Platteneinheit **5** angebracht werden kann.

[0067] Die **Fig.** 5A und 5B zeigen Darstellungen der Montage der Steuereinheit **41**, wobei **Fig.** 5A eine Vorderansicht und 5B eine Seitenansicht hiervon ist. In **Fig.** 5B bezeichnet das Bezugszeichen **51** ein Logikeinheit-Rahmenteil für die Montage der Host-Adaptereinrichtungen **1**, der Platten-Adaptereinrichtungen **2** und der Cache-Speicher-Baueinheit **3**, das Bezugszeichen **52** ein Batterieteil für die Versorgung des durch einen flüchtigen Speicher gebildeten Cache-Speichers mit Leistung bei einem Ausfall der gewöhnlichen Leistungsversorgung, das Bezugszeichen **53** ein Cache-Speicher-Erweiterungsteil, in dem die Cache-Einheit **24** sowie eine zusätzliche Batterie für den hinzugefügten Cache-Speicher angebracht sind, das Bezugszeichen **54** ein SVP-Montageteil, das Bezugszeichen **55** eine schaltbare Leistungsversorgung des Logikeinheit-Rahmens für die Versorgung des Logikeinheit-Teils mit Leistung, das Bezugszeichen **56** ein Matrix-Platteneinheit-Montageteil, falls die Konfiguration (Kapazität) der Matrix-Platteneinheit **5** gering ist, das Bezugszeichen **57** eine schaltbare Leistungsversorgung für die Matrix-Platteneinheit **5** und das Bezugszeichen **58** eine Netzstromversorgung-Steuereinheit für die Zuführung von Leistung an die beiden schaltbaren Leistungsversorgungen **55** und **57**.

[0068] Die **Fig.** 6A und 6B zeigen Montageansichten der Matrixeinheit **41**, wenn die Anordnung einer Matrix-Platteneinheit mit großer Kapazität gewünscht ist, wobei **Fig.** 6A eine Vorderansicht und **Fig.** 6B eine Seitenansicht hiervon ist.

[0069] Das Matrix-Platteneinheit-Montageteil **56** kann 112 Magnetplatteneinheiten (2 × 7 Reihen × 8 Spalten) aufnehmen, wobei für einen einfachen Austausch einer fehlerhaften Magnetplatteneinheit das Anbringungsteil **56** ein Anbringungssystem verwendet, bei dem fehlerhafte Einheiten bzw. neue Einheiten von der Vorderseite der Einheit bzw. von der Rückseite der Einheit entnommen bzw. eingesetzt werden können.

[0070] In **Fig.** 6B bezeichnet das Bezugszeichen **61** einen Kühlungslüfter für die Abführung der in der gesamten Einheit erzeugten Wärme. Um die Kühlungswirkung zu verbessern und um den Geräuschpegel niedrig zu halten, umfaßt der Lüfter **61** mehrere kleine Kühlungslüfter, die in dem Montageteil **56** verteilt sind, um die Luft von der Bodenseite zur Deckenseite zu bewegen und so eine Ventilationskühlung zu erzeugen.

[0071] **Fig.** 7 ist ein Montageschaltplan des mit Bezug auf die **Fig.** 5A und 5B erläuterten Logikeinheit-Rahmenteils.

[0072] In der Zeichnung bezeichnet das Bezugszeichen **71** eine hintere Ebene (Platte für den Einschub von Substraten), in der der gemeinsame Bus **4** in Form einer gedruckten Schaltung verdrahtet ist, während das Bezugszeichen **72** einen Anschluß für die Ankopplung der einzelnen Adaptereinrichtungen, der Cache-Speicher-Baueinheit und der hinteren Ebene **71** bezeichnet.

[0073] Da die Datenübertragung zwischen den Host-Adaptereinrichtungen **1**, den Platten-Adaptereinrichtungen **2** und der Cache-Speicher-Baueinheit **3** über den gemeinsamen Bus **4** erfolgt, kann jede Adaptereinrichtung und die Baueinheit an jeder beliebigen Position des Anschlusses **72** angeschlossen werden, wobei die Anzahl der montierten Host-Adaptereinrichtungen **1** und die Anzahl der montierten Platten-Adaptereinrichtungen **2** beliebig geändert werden können.

[0074] Wenn hierbei gewünscht ist, die Kapazität des Cache-Speichers zu erhöhen, wird entweder die Cache-Speicher-Baueinheit **3** durch die Cache-Anschlußbaueinheit **22** ersetzt oder die Cache-Anschlußbaueinheit **22** wird zusätzlich zu der Cache-Speicher-Baueinheit **3** angebracht und über das Verbindungskabel **23** mit der Cache-Einheit **24** verbunden, wie in **Fig.** 7 gezeigt ist. Im Ergebnis kann die Kapazität des Cache-Speichers gegenüber der Kapazität von 2 GB der Cache-Speicher-Baueinheit **3** um einen Wert von bis zu 2 × 8 GB erhöht werden.

[0075] **Fig.** 8 ist eine auseinandergezogene perspektivische Ansicht des in den **Fig.** 5A und 5B gezeigten Logikeinheit-Rahmenteils, die der Erläuterung des Aufbaus der Rahmeneinheit dient.

[0076] In **Fig.** 8 ist der gemeinsame Bus **4** auf einer hinteren Ebene **71** in deren horizontaler Richtung in Form einer gedruckten Schaltung verdrahtet, wobei auch die Montageteile der Substrate (CP) der Cache-Abschlußbaueinheiten **22**, die Montageteile der Substrate (C) der Cache-Speicher-Baueinheiten **3**, die Montageteile der Substrate (H) der Host-Adaptermodule **1** sowie die Montageteile der Substrate (D) der Platten-Adaptermodule **2** an der hinteren Ebene **71** vorgesehen sind, so daß jedes Substrat von der Einschub-/Entnahmeseite aus befestigt und gelöst werden kann, wie durch den Pfeil **84** gezeigt ist. Wenn das Substrat in die hintere Ebene **71** eingesetzt ist, ist dieses Substrat mit dem gemeinsamen Bus **4** elektrisch verbunden.

[0077] Das Bezugszeichen **81** bezeichnet ein optisches Anschlußteil, das in den unteren Bereichen der Substrate der Host-Adaptereinrichtungen **1** montiert ist und eine Schnittstelle mit der Host-Vorrichtung bildet, das Bezugszeichen **82** bildet ein SCSI-Anschlußteil, das in den unteren Bereichen der Substrate der Platten-Adaptereinrichtungen **2** montiert ist und mit der Matrix-Platteneinheit **5** verbunden ist, das Bezugszeichen **83** bezeichnet einen Anschluß für die Verbindung des Kabels **23** zwischen den hinteren Ebenen, wenn die Cache-Anschlußbaueinheit **22** angebracht ist, und das Bezugszeichen **85** bezeichnet ein Cache-Speicher-Gehäuse (den Cache-Speicher **20** in **Fig.** 2), der im unteren Teil des Substrats C der Cache-Speicher-Baueinheit **3** montiert ist.

[0078] Um die Funktionsfähigkeit bei der Entnahme der fehlerhaften Adaptereinrichtung oder Baueinheit und beim Einsetzen einer neuen bei Auftreten eines Fehlers zu verbessern, sind die Anschlüsse mit Ausnahme des Anschlusses **83** nicht auf der Bedienungsseite **84**, sondern konzentriert auf seiten der hinteren Ebene **71** montiert.

[0079] **Fig.** 9 zeigt die Konfiguration der in der vorliegenden Erfindung verwendeten Software.

[0080] Das Bezugszeichen **91** bezeichnet ein Kanaladapter-Steuerprogramm (das im folgenden mit CHP bezeichnet wird), das in die Boot-Vorrichtung **12** der Host-Adaptereinrichtung **1** geschrieben ist. Von den in die Boot-Vorrichtung **12** der Platten-Adaptereinrichtung **2** geschriebenen Platten-Adapter-Steuerprogrammen bezeichnen das Bezugszeichen **92** ein Plattenadapter-Mastersteuerprogramm (das im folgenden mit DMP bezeichnet wird), das die in der Matrix-Platteneinheit **5** ausgeführte Operation sowie die Datenübertragung zwischen dem Cache-Speicher **20** und der Matrix-Platteneinheit **5** steuert, und das Bezugszeichen **93** ein Plattenadapter-Slavesteuerprogramm (das im folgenden mit DSP bezeichnet wird), das die Steuerung der Datenübertragung zwischen dem Cache-Speicher **20** und der Matrix-Platteneinheit **5** unter der Steuerung des DMP **92** steuert.

[0081] In die Boot-Vorrichtung **12** der Platten-Adaptereinrichtung **2** sind die Programme DMP **92** und DSP **93** mit zwei verschiedenen Arten von Funktionen geschrieben, so daß bei n Sätzen von Matrix-Platteneinheiten eine der Platten-Adaptereinrichtungen als DMP **92** betrieben wird, während eine weitere als Reserve-DMP **92** spezifiziert ist (und als DSP **93** betrieben wird), und die verbleibenden n-2 Plat-

ten-Adaptereinrichtungen als DSP **93** betrieben werden.

[0082] Das Bezugszeichen **94** bezeichnet ein SVP-Steuerprogramm für den SVP **27** zur Überwachung und Wartung der Programme CHP **91**, DMP **92** und DSP **93**. Wenn die Erneuerung eines jeden Steuerprogramms gewünscht ist, kann der SVP **27** direkt das Steuerprogramm des MP **10**, das in der Boot-Vorrichtung **12** der zu aktualisierenden Adaptereinrichtung gespeichert ist, oder aber das Steuerprogramm eines weiteren MP **10** erneuern.

[0083] **Fig.** 10 zeigt die gemeinsam genutzten Funktionen der Software-Konfiguration von **Fig.** 9 anhand des Datenstroms.

[0084] Das CHP **91** wandelt die Adressen- und Datenformate der höherrangigen oder Host-Vorrichtung in Adressen- und Datenformate der niederrangigen Vorrichtung um und schreibt diese in den Cache-Speicher **20**. Das Bezugszeichen **101** bezeichnet ein Segment, das Bezugszeichen **102** einen Block und das Bezugszeichen **103** einen Streifen, der die Datenmenge pro Magnetplatte angibt, die in die Matrix-Platteneinheit **5** geschrieben wird. Das DMP **92** liest eine Streifen-Dateneinheit aus dem Cache-Speicher aus, wandelt eine Adresse niedriger Ordnung der Daten in eine Zeilennummer und eine Spaltennummer, eine FBA und eine Blocknummer der Matrix-Platteneinheit um, während das DSP **93** die Daten in die Matrix-Platteneinheit **5** schreibt.

[0085] Das DMP **92** verwaltet außerdem die Information bezüglich der Konfiguration und der Matrix-Platteneinheit **5**.

[0086] Da die jeweiligen Steuerprogramme wie oben beschrieben die jeweiligen Funktionen gemeinsam nutzen, kann dann, wenn ein Wechsel der Host-Schnittstelle zu einem SCSI- oder zu einem Faser-Kanal gewünscht ist, dieser Wechsel lediglich durch Andern des Steuerprogramms CHP **91** ausgeführt werden. Wenn ferner eine Änderung der Konfiguration der Matrix-Platteneinheit gewünscht ist (Anzahl der Reihen/Anzahl der Spalten von Platten, RAID-System (redundantes Matrixsystem billiger Platten) usw.), kann dies lediglich durch Ändern des Steuerprogramms des DMP **92** erfolgen. Wenn die Verbindung der Host-Adaptereinrichtungen **1** und der Platten-Adaptereinrichtungen **2** modifiziert wird und die jeweiligen Steuerprogramme neu geschrieben werden, kann eine Skalierbarkeit verwirklicht werden, ferner kann der Aufwand für die Softwareentwicklung verringert werden.

[0087] Die **Fig.** 11A und 118 sind Diagramme zur Erläuterung des gemeinsamen, doppelt vorhandenen Busses **4** und des verminderten Betriebs. Das Bezugszeichen **111** bezeichnet einen Bus-Master (die Host-Adaptereinrichtung **1** oder die Platten-Adaptereinrichtung **2** mit dem MP **10**), der eine Zugriffserlaubnis auf den gemeinsamen Bus **4** erlangen kann, während das Bezugszeichen **112** einen Bus-Slave (Cache-Speicher-Baueinheit) bezeichnet, der eine Zugriffsanforderung vom Bus-Master **111** empfängt.

[0088] Der S-Bus **26** verwirklicht in seinem normalen Betriebszustand eine Übertragungsrate von 400 MP/s anhand zweier Kanäle mit einer jeweiligen Breite von 64 Bit (je 200 MB/s), wobei jeder Bus-Kanal anhand der Paritätsprüfung oder der Zeitauslösung einen Fehler erfassen kann. Wenn in einem der beiden Kanäle ein Fehler auftritt, wird der Bus-Master **111** in seinen verminderten Zustand versetzt, so daß der verbleibende Kanal so verwendet wird, daß er auf den Bus-Slave **112** zugreift, woraufhin die Information bezüglich des verminderten Betriebs in den Verwaltungsbereich des gemeinsam genutzten Speichers **18** eingetragen wird.

[0089] Die Zuverlässigkeit der Systemsteuersignale (einschließlich eines Bus-Rücksetzsignals) im gemeinsamen Bus wird erhöht, weil die Signalleitung die Form eines Triplex-Systems besitzt, so daß im Normalbetrieb ein System mit Übereinstimmung der drei Leitungen (Kanäle) verwendet wird und im verminderten Betrieb ein System mit Übereinstimmung von zwei Leitungen (Kanälen) mittels Mehrheitsentscheidung verwendet wird.

[0090] **Fig.** 12 ist ein Diagramm zur Erläuterung der Mehrfachstrukturen der jeweiligen Teile und des verminderten Betriebs.

[0091] Das Bezugszeichen **121** bezeichnet einen Kanalbus mit zwei Anschlüssen. Die Host-Adaptereinrichtung **1** ist mit zwei der Kanaladaptermodule **29** und mit vier Kanalbussen für die Host-Vorrichtung versehen, so daß bei einem Ausfall eine alternative Kanaladaptereinrichtung (CHP) und ein alternativer Kanalbus verwendet werden, um in den verminderten Betrieb einzutreten.

[0092] Das Bezugszeichen **122** bezeichnet einen SCSI-Bus, der eine Schnittstelle zwischen der Platten-Adaptereinrichtung **2** und der Matrix-Platteneinheit **5** bildet. Der SCSI-Bus **122** ist doppelt vorgesehen, so daß eine weitere Platten-Adaptereinrichtung **2** auf eine Reihe von Magnetplatten zugreifen kann, wobei bei Auftreten eines Fehlers in einem Bus. der alternative SCSI-Bus dazu verwendet wird, in den verminderten Betrieb einzutreten. Das DMP **92** zur Steuerung des Matrix-Platteneinheit-Masters bestimmt eines der DSP **93** als alternatives DMP **92**, so daß bei Auftreten eines Fehlers das alternative DMP **92** dazu verwendet wird, den Matrix-Platteneinheit-Master zu steuern.

[0093] Der gemeinsam genutzte Speicher **18** und der Cache-Speicher. 20 sind ebenfalls doppelt vorhanden, so daß bei Auftreten eines Fehlers in einem der doppelt vorgesehenen und gemeinsam genutzten Speicher der andere Speicher dazu verwendet wird, in den verminderten Betrieb einzutreten; wenn hingegen in einem doppelt vorhandenen Cache-Speicher ein Fehler auftritt, werden die zum Schreiben anstehenden Daten (im Cache-Speicherverbliebene Daten) auf die Platte ausgespeichert, um den verminderten Betrieb unter Verwendung des weiteren Cache-Speichers mit Ausnahme des fehlerhaften Cache-Speichers auszuführen.

[0094] Wenn in einer der Magnetplatten in der Matrix-Platteneinheit **5** ein Fehler auftritt, wird die Lese-/Schreiboperation ausgeführt, während die fehlerhafte Magnetplatte abgetrennt wird und durch eine Reserve-Magnetplatte ersetzt wird.

[0095] **Fig.** 13 ist ein Diagramm zur Erläuterung der Mehrfachstruktur eines Leistungsversorgungssystems und des verminderten Betriebs.

[0096] Da Netzstromversorgungs-Steuereinrichtungen **58** an jeweiligen unabhängigen Wechselstromeingängen doppelt vorgesehen sind, um an schaltbare Leistungsversorgungen **55** des Logikeinheit-Rahmens bzw. an schaltbare Leistungsversorgungen **57** der Matrix-Platteneinheit Leistung zu liefern, wird bei Auftreten eines Fehlers in einer Netzstromversorgungs-Steuereinrichtung **58** die andere Netzstromversorgungs-Steuereinrichtung **58** dazu verwendet, in den verminderten Betrieb einzutreten.

[0097] Das Bezugszeichen **131** bezeichnet eine Leistungsversorgungs-Steuerschaltung (die im folgenden als PC1 bezeichnet wird), die eine Fernsteuerung des Ein- und Ausschaltens der Leistung von der Host-Vorrichtung ausführt und die Netzstromversorgungs-Steuereinrichtungen **58** sowie Leistungsversorgungsschaltungen wie etwa die beiden schaltbaren Leistungsversorgungen **55** und **57** steuert.

[0098] Wenn die schaltbaren Leistungsversorgungen **55** des Logikeinheit-Rahmens in einer Anzahl vorgesehen sind, die gegenüber der für den redundanten Betrieb erforderlichen Anzahl zur Leistungsversorgung der Logikeinheit-Rahmen **51** und der Batterien **52** über den gemeinsamen Leistungsbus um 2 erhöht ist, kann das System selbst dann betrieben werden, wenn in zwei der schaltbaren Leistungsversorgungen **55** ein Fehler auftritt.

[0099] Ebenso kann dann, wenn die schaltbaren Leistungsversorgungen **57** der Matrix-Platteneinheit für die Zuführung von Leistung an die Reiheneinheiten einer Gruppe von Magnetplatten in einer Anzahl vorgesehen sind, die gegenüber der Anzahl, die für den redundanten Betrieb für die Leistungsversorgung über den gemeinsamen Leistungsbus erforderlich ist, um 2 erhöht ist, das System selbst dann betrieben werden, wenn in zwei der schaltbaren Leistungsversorgungen **57** ein Fehler auftritt. Weiterhin ist das System billiger als ein System, in dem die beiden schaltbaren Leistungsversorgungen **55** und **57** mit Doppelstruktur vorgesehen sind.

[0100] Bei einem Ausfall der Leistung wird über den gemeinsamen Leistungsbus von der doppelt vorgesehenen Batterie **52** an den durch einen flüchtigen Speicher gebildeten Cache-Speicher in dem Logikeinheit-Rahmen sowie an die PC1 **131** Leistung geliefert, so daß das System selbst dann betrieben werden kann, wenn in mehr als einer Batterie ein Fehler auftritt.

[0101] Die **Fig.** 14 und 15 zeigen in tabellarischer bzw. graphischer Form Vergleiche von Systemleistungen, wenn die verschiedenen Matrix-Platteneinheiten aus Magnetplatteneinheiten mit unterschiedlichen Speicherkapazitäten aufgebaut sind.

[0102] Genauer zeigt **Fig.** 14 die Struktur von Matrix-Platteneinheiten mit gleicher Kapazität, jedoch bei Verwendung unterschiedlicher Typen von Magnetplatteneinheiten.

[0103] In der Matrixstruktur bilden 14 Datenplatten und zwei Paritätsplatten einen Satz. Bei der Positionsnummer **141** wird eine Magnetplatteneinheit mit 3 GB (3,5-Zoll-Platte) verwendet, wobei 5 Sätze von Matrix-Konfigurationen vorgesehen sind. Bei der Positionsnummer **142** wird eine Magnetplatteneinheit mit 4 GB (5-Zoll-Platte) verwendet, wobei vier Sätze von Matrix-Konfigurationen vorgesehen sind. Bei der Positionsnummer **143** wird eine Magnetplatteneinheit mit 8,4 GB (6,4-Zoll-Platte) verwendet, wobei zwei Sätze von Matrix-Konfigurationen vorgesehen sind.

[0104] **Fig.** 15 zeigt die Beziehung zwischen der Anzahl der pro Sekunde ausgegebenen E/A-Befehle und der mittleren Antwortzeit der Magnetplatteneinheiten **141**, **142** bzw. **143**. Um die Transaktionsleistung des Matrix-Plattensystems zu verbessern, kann zur Erhöhung der Matrix-Konfiguration bei Verwendung einer Magnetplatteneinheit mit geringer Kapazität (kleinem Durchmesser) die höchste Leistung erhalten werden. Daher wird erfindungsgemäß die 3,5-Zoll-Magnetplatteneinheit **141** für die Verwirklichung des Matrix-Plattensystems verwendet. Wenn daher eine Magnetplatteneinheit mit einer bestimmten Speicherkapazität, die aus einer einzigen großdimensionierten Magnetplatteneinheit aufgebaut ist, mit einer Magnetplatteneinheit mit der gleichen Speicherkapazität verglichen wird, die jedoch aus mehreren kleindimensionierten Magnetplatteneinheiten aufgebaut ist, die in einer Matrix angeordnet sind, so ist diese letztere Matrixstruktur mit mehreren kleindimensionierten Magnetplatteneinheiten von Vorteil,

weil die mittlere Zugriffszeit kürzer ist.

[0105] In den **Fig.** 16 bis 19 sind modellhafte Beispiele der Systemstruktur gezeigt, die bei Verwendung einer skalierbaren Architektur, wie sie oben erläutert worden ist, implementiert wird

[0106] Genauer zeigt **Fig.** 16 eine Anordnung, bei der die Anzahl der Platten-Adaptereinrichtungen **2**, die am gemeinsamen Bus **4** angeschlossen sind, erniedrigt ist und bei der ferner die Cache-Anschlußbaueinheiten **22** so beschaffen sind, daß sie über das Kabel **23** mit den Cache-Einheiten **24** verbunden werden, um auf diese Weise eine Plattenmatrix mit geringen Abmessungen und hoher Leistung zu verwirklichen, wobei Cache-Speicher mit großer Kapazität eine hohe Cache-Trefferrate schaffen.

[0107] Wenn die Platten-Adaptereinrichtungen **2** nicht angeschlossen sind und das System lediglich mit der Host-Adaptereinrichtung **1** und mit den Cache-Speichern versehen ist (einer Anordnung, die in **Fig.** 16 von der unterbrochenen Linie umgeben ist), sind die als Speichermedium dienenden Magnetplatten durch Halbleiterspeicher ersetzt, so daß ein Hochleistungs-Halbleiterspeichersystem mit höherer Datenübertragungsgeschwindigkeit verwirklicht wird.

[0108] **Fig.** 17 zeigt eine Anordnung, bei der eine maximale Anzahl von Platten-Adaptereinrichtungen **2** vorgesehen ist, wobei die Cache-Speicher-Baueinheiten **3** oder die Cache-Anschlußbaueinheiten **22** so vorgesehen sind, daß sie über das Kabel **23** mit den Cache-Einheiten verbunden sind, um auf diese Weise ein großdimensioniertes Matrix-Plattensystem mit hoher Leistung und mit Cache-Speichern mit hoher Kapazität zu verwirklichen.

[0109] **Fig.** 18 zeigt eine Anordnung, bei der die Host-Vorrichtungs-Schnittstellen der Host-Adaptereinrichtungen **1** durch Schnittstellen wie etwa SCSI-/Faser-Kanäle ersetzt sind, um die Anzahl der angeschlossenen Platten-Adaptereinrichtungen **2** zu reduzieren, ferner ist der S-Bus **26** aus zwei Kanälen mit einer Bitbreite hergestellt, die der halben Bitbreite des S-Busses entspricht, um dadurch ein für den freien Markt konstruiertes Serversystem mit ununterbrochenem Betrieb, hoher Leistung und Fehlertoleranz (hochzuverlässig) zu verwirklichen.

[0110] **Fig.** 19 zeigt eine möglichst einfache Version der in **Fig.** 18 gezeigten Anordnung, wobei der Duplex-Struktur und der Möglichkeit des Austausches von Komponenten während des Betriebs keine Beachtung geschenkt wird, um dadurch ein billiges Serversystem für den freien Markt zu verwirklichen. In der Zeichnung hat das Bezugszeichen **4D** + **1P** die bedeutung von vier Datenplatten und einer einzigen Paritätsplatte.

[0111] In den obenbeschriebenen Ausführungsformen können die optischen Plattenspeichereinheiten an den gemeinsamen Bus **4** über optische Platten-Adaptereinrichtungen (Logikmodule für den Anschluß optischer Platten) angeschlossen werden, während Magnetband- und Magnetplatteneinheiten an den gemeinsamen Bus über Magnetband- bzw. Magnetplatten-Steuereinrichtungen (Logikmodule für den Anschluß von Magnetbändern bzw. Magnetplatten) angeschlossen werden können und Halbleiterspeicher über Logikmodule für den Anschluß von Halbleiterspeichern an den gemeinsamen Bus angeschlossen werden können. Ferner können an den gemeinsamen Bus **4** über einen weiteren Typ von Host-Adaptereinrichtungen Arbeitsstationen angeschlossen werden. Auf diese Weise können an den gemeinsamen Bus Speichermedium-Adaptereinrichtungen für verschiedene Speichertypen angeschlossen werden.

[0112] Wie oben erläutert worden ist, wird erfindungsgemäß ein Speichersystem geschaffen, das versehen ist mit mehreren Logikeinheiten, die mit einer Host-Vorrichtung verbunden sind und für diese Host-Vorrichtung Schnittstellen bilden, mehreren Logikeinheiten, die mit den Speichern verbunden sind und für die Speichereinheit Schnittstellen bilden, und Cache-Speichern (die von den mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten und von den mit der Speichereinheit verbundenen mehreren Logikeinheiten gemeinsam genutzt werden), in denen Daten, die zwischen diesen Vorrichtungen und Einheiten übertragen werden sollen, temporär gespeichert werden. Hierbei sind die mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, die mit der Speichereinheit verbundenen mehreren Logikeinheiten sowie die Cache-Speicher über einen gemeinsamen Bus miteinander verbunden, der von diesen Vorrichtungen und Einheiten gemeinsam genutzt wird. Im Ergebnis kann ein skalierbares System erhalten werden, mit dem eine Erweiterung und eine Modifikation der mit der Host-Vorrichtung verbundenen mehreren Logikeinheiten, der mit den Speichern verbundenen mehreren Logikeinheiten und der Cache-Speicher möglich sind, indem lediglich weitere Einheiten zum gemeinsamen Bus hinzugefügt und/oder modifiziert werden. Damit kann auf der Grundlage der Möglichkeit der Systemerweiterung eine Aufrüstung einfach erzielt werden. Da ferner die mit der Host-Vorrichtung verbundenen Logikeinheiten, die mit den Speichern verbundenen Logikeinheiten und die Cache-Speicher in Form von Modulen gebildet sind, die an einer hinteren Ebene lösbar angebracht sind, auf der der gemeinsame Bus vorgesehen ist, kann die erforderliche Anzahl von Einheiten und Speichern vorteilhaft einfach erhöht werden.

[0113] Da die mit der Host-Vorrichtung verbundenen Logikeinheiten, die mit den Speichern verbundenen Logikeinheiten und die Cache-Speicher doppelt

vorhanden sind und der gemeinsame Bus zwischen diesen Logikeinheiten und den Speichern in zwei Kanäle unterteilt ist, kann selbst dann, wenn in einer dieser Einheiten ein Fehler auftritt, die andere dazu verwendet werden, einen verminderten Betrieb auszuführen. Da in diesem Fall jede der mit der Host-Vorrichtung verbundenen Logikeinheiten, jede der mit den Speichern verbundenen Logikeinheiten und jeder der Cache-Speicher mit einem Anschluß versehen ist, der einen Austausch von Komponenten während des Betriebs erlaubt, ist mit dem System vorteilhaft eine Wartung und ein Eingriff für den Austausch eines fehlerhaften Teils während des Betriebs möglich, ferner können für eine Systemerweiterung Teile hinzugefügt werden, ohne daß das in Betrieb befindliche System angehalten werden muß.

[0114] Da ferner die Speichereinheit in Matrixform aus mehreren kombinierten kleinen Speichern gebildet ist, kann mit der Speichermatrix die Zugriffszeit gegenüber Speichern, die aus einer einzigen, herkömmlichen großdimensionierten Platteneinheit gebildet sind, vorteilhaft verkürzt werden.

[0115] Da außerdem die Cache-Speicher-Einheit aus Cache-Speicher-Modulen (Cache-Speicher-Baueinheiten) hergestellt ist, die direkt an den gemeinsamen Bus und an Erweiterungs-Cache-Einheiten angeschlossen sind und da die erforderliche Anzahl von Erweiterungs-Cache-Einheiten über Erweiterungs-Cache-Anschlußbaueinheiten angeschlossen werden können, die direkt und lösbar mit dem gemeinsamen Bus verbunden werden können, kann die Anzahl der Cache-Einheiten vorteilhaft leicht erhöht oder erniedrigt werden.

[0116] Im Ergebnis kann ein hochzuverlässiges Speichersystem erhalten werden.

Patentansprüche

1. Speichersystem mit
mehreren ersten Logikeinheiten (**1**), die mit einer Host-Vorrichtung verbunden sind und Schnittstellen für diese Host-Vorrichtungen bilden;
einer Speichervorrichtung (**5**), die von der Host-Vorrichtung empfangene Information speichert;
mehreren zweiten Logikeinheiten (**2**), die mit der Speichervorrichtung (**5**) verbunden sind und für diese Speichervorrichtung (**5**) Schnittstellen bilden;
einer Cache-Speichervorrichtung (**3**), die Daten, die zwischen den mehreren ersten Logikeinheiten (**1**) und den mehreren zweiten Logikeinheiten (**2**) übertragen werden, temporär speichert; und
einem gemeinsamen Bus (**4**), der zwischen den mehreren ersten Logikeinheiten (**1**), den mehreren zweiten Logikeinheiten (**2**) und der Cache-Speichervorrichtung (**3**) verläuft und mit diesen funktional verbunden ist,
dadurch gekennzeichnet, daß
der Bus (**4**) einen E/A-Hochgeschwindigkeitsbus (**26**) zur Datenübertragung, der in zwei Kanäle unterteilt ist, die unterschiedliche Daten übertragen können, sowie einen Einzelkanal-Multiprozessor-Bus (**25**) zur Übertragung von Steuerinformation für die Datenübertragung, enthält,
wobei
dann, wenn im Multiprozessor-Bus (**25**) ein Fehler auftritt, einer der beiden Kanäle des E/A-Hochgeschwindigkeitsbusses (**26**) als Multiprozessor-Bus verwendet wird und der Betrieb des E/A-Hochgeschwindigkeitsbusses (**26**) durch den anderen Kanal fortgesetzt wird.

2. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß die Cache-Speichervorrichtung (**3**) aus doppelt vorhandenen Cache-Speichern (**20**) besteht, die gemeinsam oder im Fehlerfall einzeln betrieben werden.

3. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen gebildet sind, wobei jeder der Module am gemeinsamen Bus (**4**) lösbar angebracht ist.

4. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß der gemeinsame Bus (**4**) in einer Ebene (**71**) angeordnet ist und die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen an der Ebene (**71**) lösbar angebracht sind.

5. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß die ersten Logikeinheiten (**1**) unterschiedliche Schnittstellen haben.

6. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß die Speichervorrichtung (**5**) mehrere kleindimensionierte Speichereinheiten enthält, von denen eine Paritätsdaten speichert und die zusammen eine Fehlerkorrekturgruppe bilden.

7. Speichersystem nach Anspruch 1, gekennzeichnet durch
mehrere doppelt vorhandene erste Logikeinheiten (**1**), und
mehrere doppelt vorhandene zweite Logikeinheiten (**2**),
wobei
beim Ausfall jeweils einer Logikeinheit (**1**, **2**) die jeweils andere einen verminderten Betrieb aufrechterhält.

8. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen gebildet sind, wobei jeder der Module am gemeinsamen Bus (**4**)

lösbar angebracht ist.

9. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß der gemeinsame Bus (**4**) in einer Ebene (**71**) angeordnet ist und die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen an der hinteren Ebene (**71**) lösbar angebracht sind.

10. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß die Speichervorrichtung (**5**) mehrere kleindimensionierte Speichereinheiten enthält, von denen eine Paritätsdaten speichert und die zusammen eine Fehlerkorrekturgruppe bilden.

11. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß die Host-Vorrichtung mit dem gemeinsamen Bus (**4**) über wenigstens zwei der mehreren ersten Logikeinheiten (**1**) verbunden ist.

12. Speichersystem nach Anspruch 1, dadurch gekennzeichnet, daß die Speichervorrichtung (**5**) mit dem gemeinsamen Bus (**4**) über wenigstens zwei der mehreren zweiten Logikeinheiten (**2**) verbunden ist.

13. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß die mehreren ersten Logikeinheiten (**1**) und die mehreren zweiten Logikeinheiten (**2**) jeweils einen doppelt vorhandenen Mikroprozessor (MP **10**, MP **10'**) sowie eine Prüfeinrichtung (**11**), die den Betrieb der doppelt vorhandenen Mikroprozessoren (MP **10**, MP **10'**) vergleicht, enthalten.

14. Speichersystem nach Anspruch 7, dadurch gekennzeichnet, daß die mehreren ersten Logikeinheiten (**1**), die mehreren zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in bezug auf den gemeinsamen Bus (**4**) einen Austausch von Komponenten während des Betriebes zulassen.

15. Speichersystem nach Anspruch 14, dadurch gekennzeichnet, daß die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen gebildet sind, wobei jeder der Module am gemeinsamen Bus (**4**) lösbar angebracht ist.

16. Speichersystem nach Anspruch 14, dadurch gekennzeichnet, daß der gemeinsame Bus (**4**) in einer Ebene (**71**) angeordnet ist und die ersten Logikeinheiten (**1**), die zweiten Logikeinheiten (**2**) und die Cache-Speichervorrichtung (**3**) in Form von Modulen an der Ebene (**71**) lösbar angebracht sind.

17. Speichersystem nach Anspruch 14, dadurch gekennzeichnet, daß die Speichervorrichtung (**5**) mehrere kleindimensionierte Speichereinheiten enthält, von denen eine Paritätsdaten speichert und die zusammen eine Fehlerkorrekturgruppe bilden.

18. Speichersystem nach Anspruch 14, gekennzeichnet durch
einen Leistungsbus, der mit den mehreren ersten Logikeinheiten (**1**), den mehreren zweiten Logikeinheiten (**2**), der Cache-Speichervorrichtung (**3**) und der Speichervorrichtung (**5**) verbunden ist,
einen Mehrfachleistungsversorgungsteil, der dem Leistungsbus Leistung zuführt, und
einen doppelt vorhandenen Leistungsversorgungsteil, der an den Mehrfachleistungsversorgungsteil Leistung zuführt,
und dadurch, daß
die Anzahl der Leistungsversorgungen im Mehrfachleistungsversorgungsteil gleich der Anzahl, die für die mehreren ersten Logikeinheiten (**1**), die mehreren zweiten Logikeinheiten (**2**), die Cache-Speichervorrichtung (**3**) und die Speichervorrichtung (**5**) notwendig ist, erhöht um 1, ist.

Es folgen 17 Blatt Zeichnungen

Anhängende Zeichnungen

FIG. 1

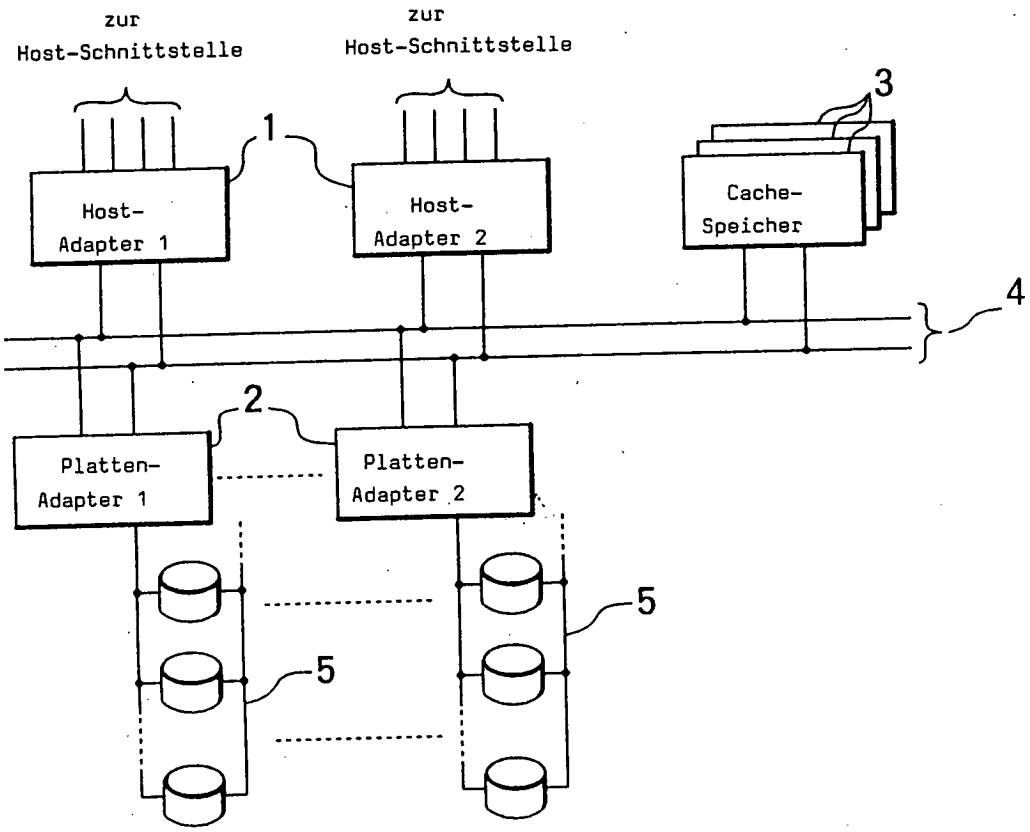
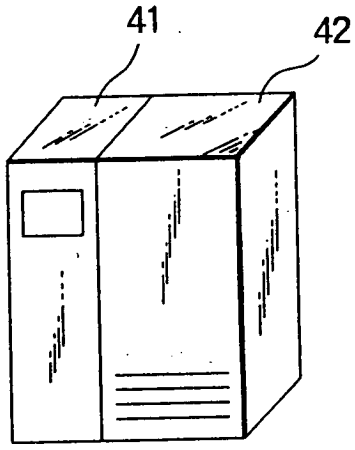


FIG. 4



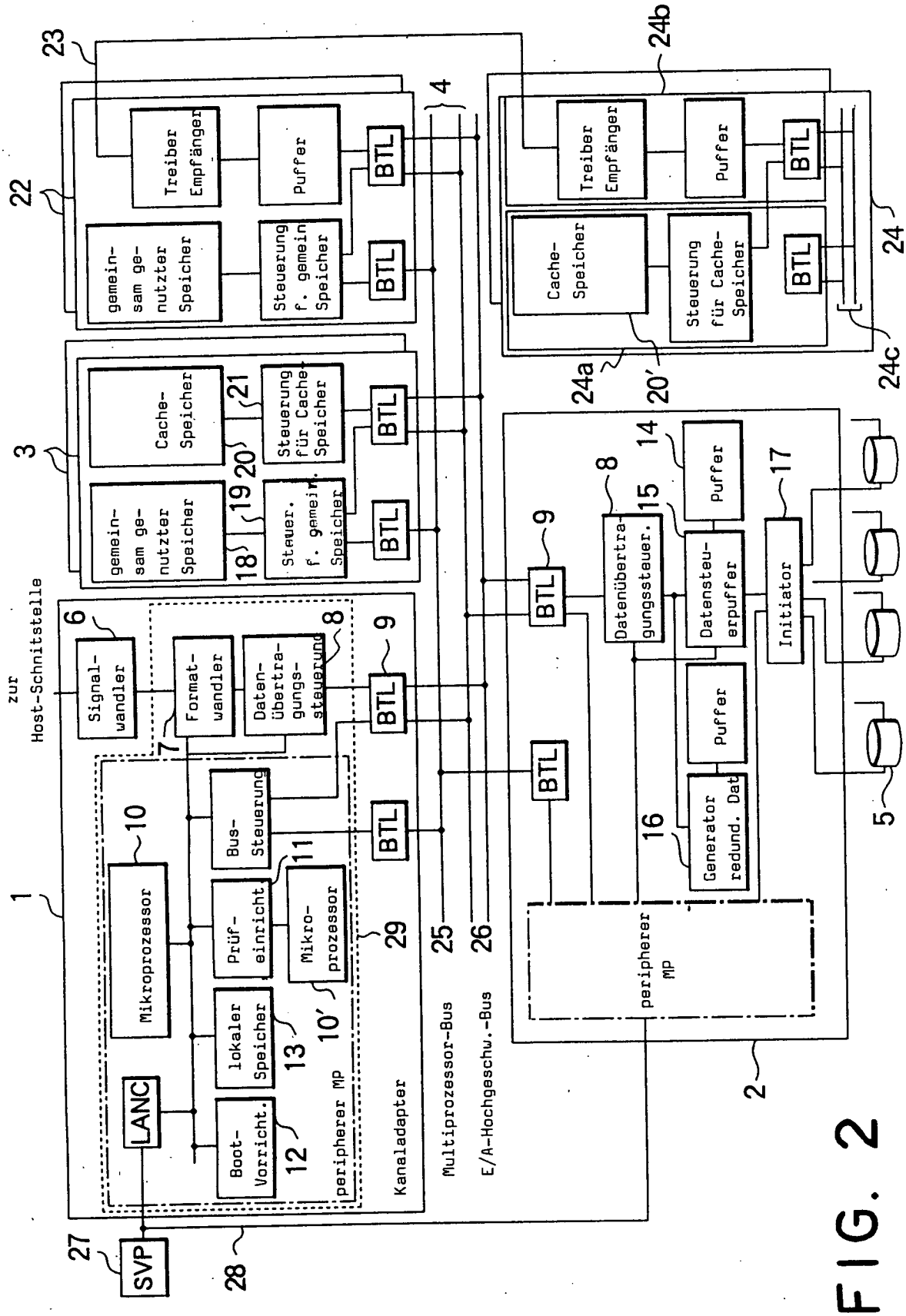

zur Host-Schnittstelle
Signal-wandler
Format-wandler
Daten-übertra-gungs-steuerung
Bus-Steuerung
Mikroprozessor
Prüf-einricht.
Mikro-prozessor
lokaler Speicher
Boot-Vorricht.
LANC
SVP
peripherer MP
Kanaladapter
Multiprozessor-Bus
E/A-Hochgeschw.-Bus
gemein-sam ge-nutzter Speicher
Treiber Empfänger
Puffer
Steuerung f. gemein Speicher
Cache-Speicher
Steuerung für Cache Speicher
Steuer. f. gemein. Speicher
Datenübertra-gungssteuer.
Datensteu-erpuffer
Initiator
Generator redund. Dat
peripherer MP
BTL

FIG. 2

FIG. 3

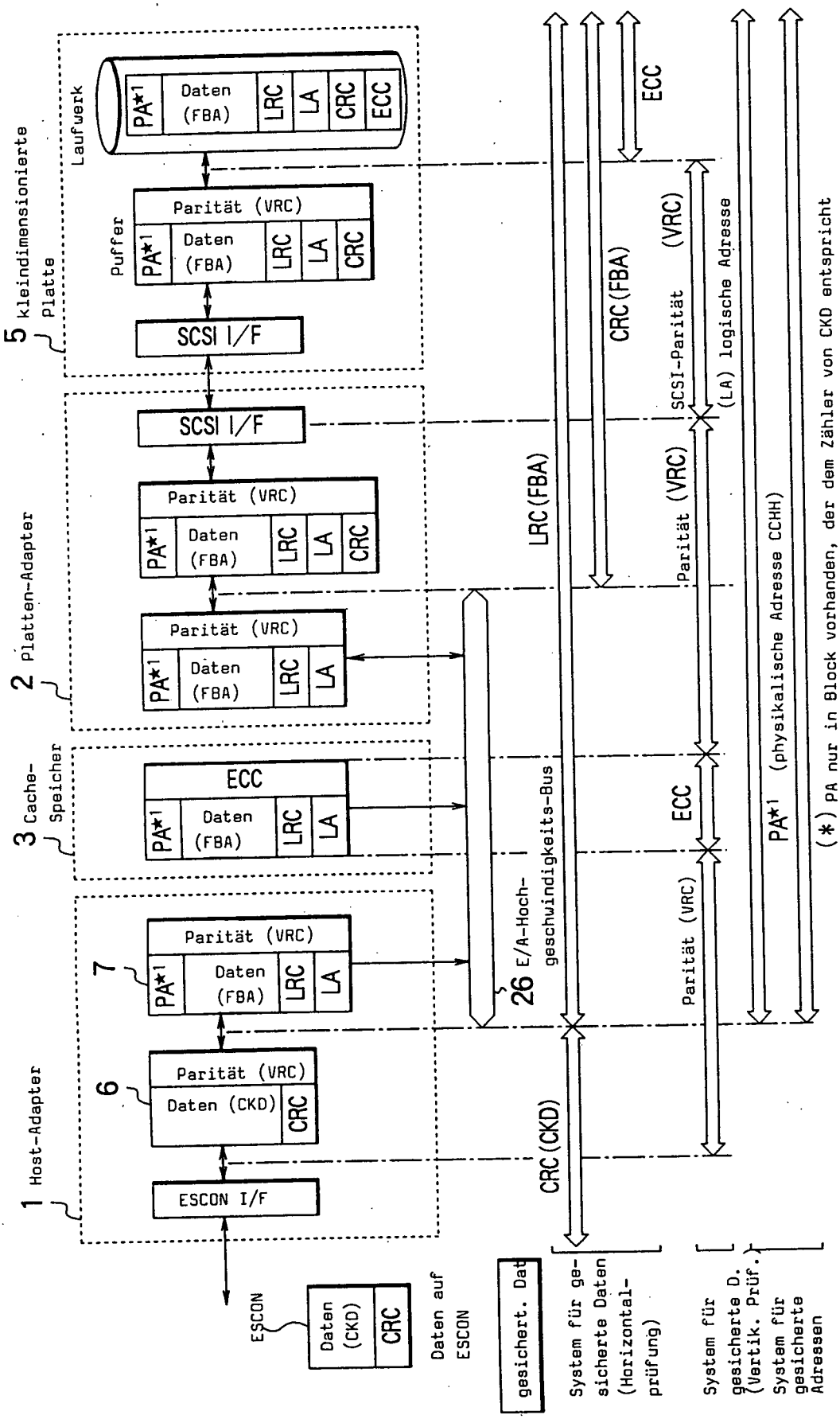

1 Host-Adapter
2 Platten-Adapter
3 Cache-Speicher
5 kleindimensionierte Platte
6
7
Laufwerk
Puffer
Daten (FBA)
Daten (CKD)
Parität (VRC)
PA*1
LRC
LA
CRC
ECC
SCSI I/F
ESCON I/F
ESCON
Daten auf ESCON
26 E/A-Hochgeschwindigkeits-Bus
LRC (FBA)
CRC (FBA)
CRC (CKD)
SCSI-Parität (VRC)
(LA) logische Adresse
PA*1 (physikalische Adresse CCHH)
gesichert. Dat
System für gesicherte Daten (Horizontalprüfung)
System für gesicherte D. (Vertik. Prüf.)
System für gesicherte Adressen
(*) PA nur in Block vorhanden, der dem Zähler von CKD entspricht

FIG. 5A

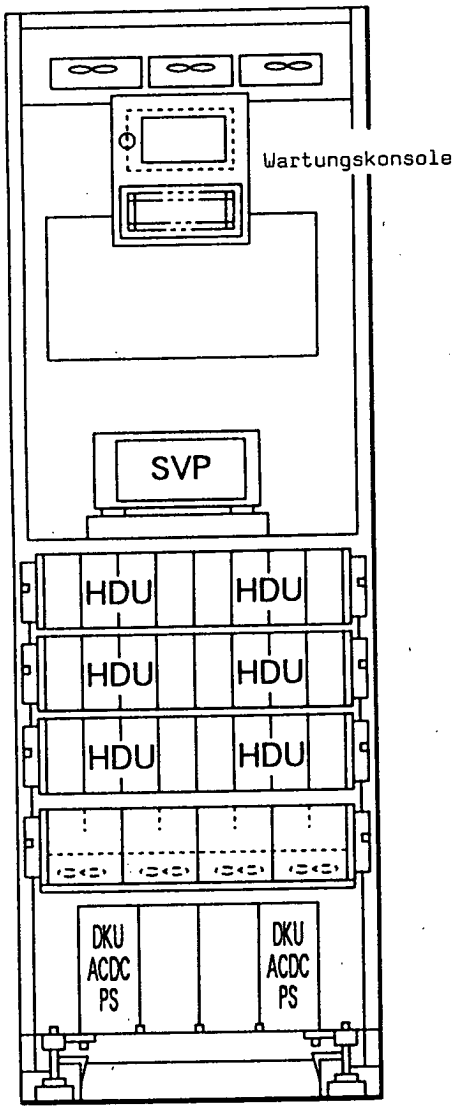

Wartungskonsole
SVP
HDU
HDU
HDU
HDU
HDU
HDU
DKU ACDC PS
DKU ACDC PS

FIG. 5B

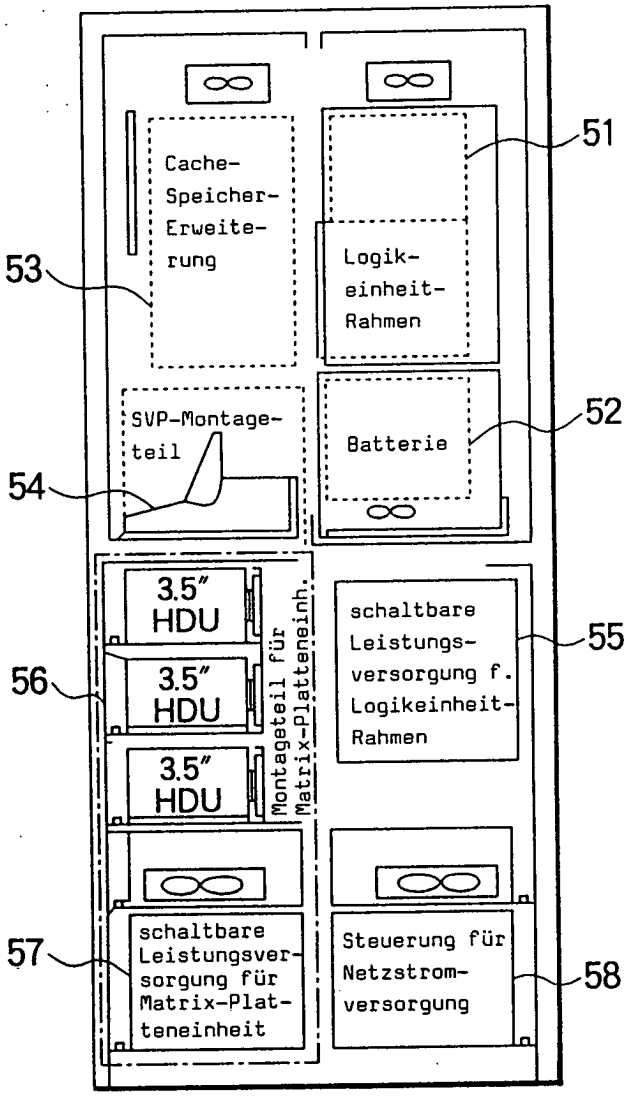
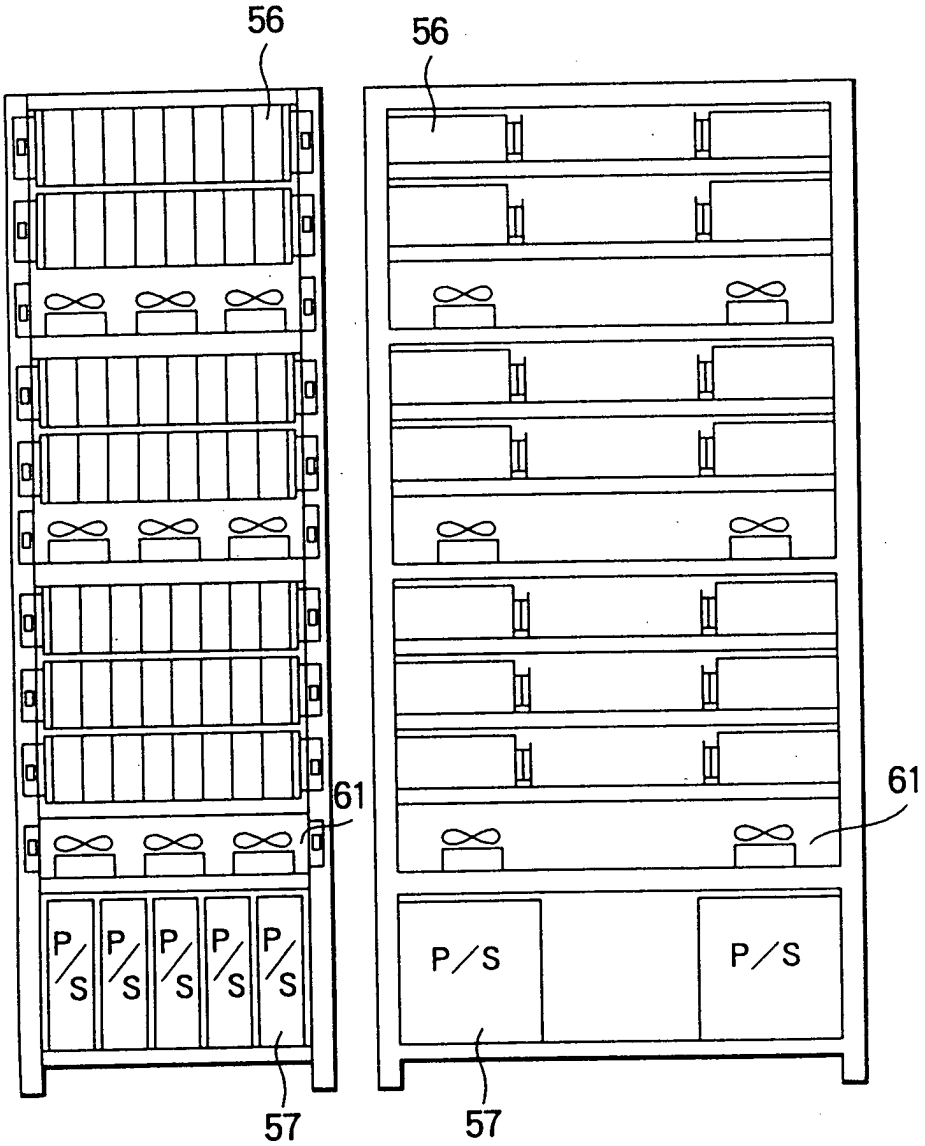

Cache-Speicher-Erweiterung
Logikeinheit-Rahmen
SVP-Montageteil
Batterie
3.5" HDU
3.5" HDU
3.5" HDU
Montageteil für Matrix-Plattenein.
schaltbare Leistungsversorgung f. Logikeinheit-Rahmen
schaltbare Leistungsversorgung für Matrix-Platteneinheit
Steuerung für Netzstromversorgung
51
52
53
54
55
56
57
58

FIG. 6A FIG. 6B



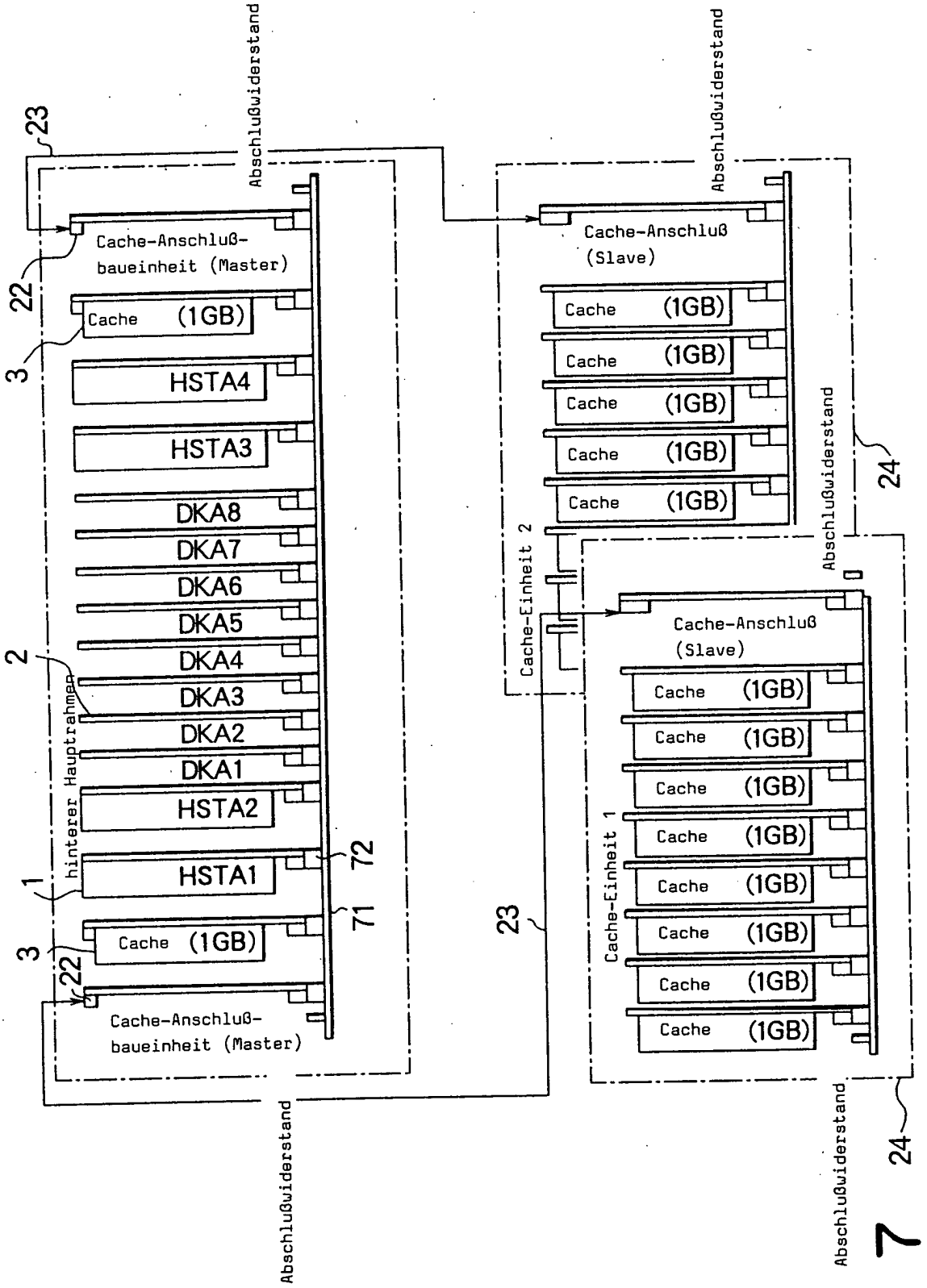
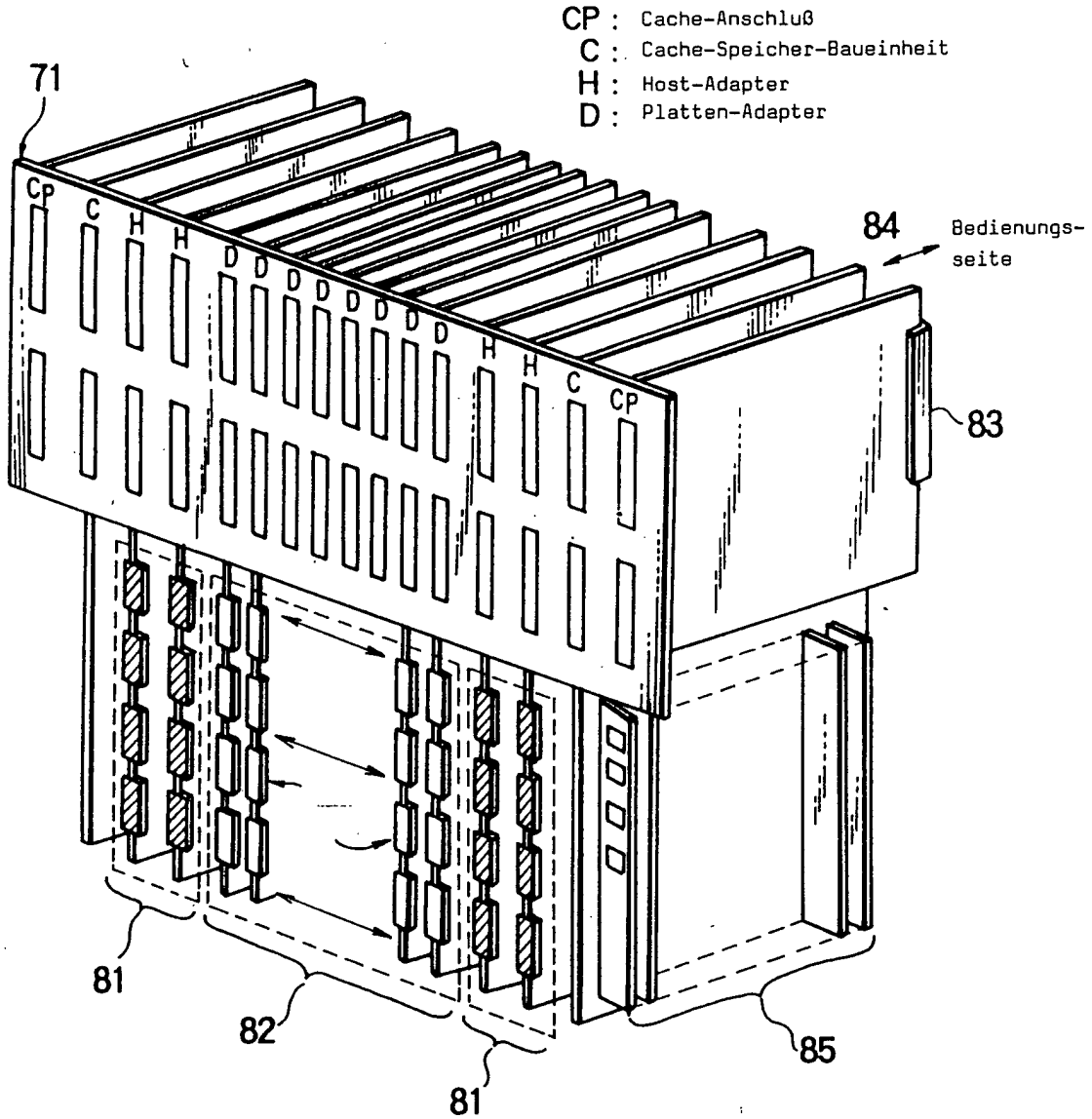

Abschlußwiderstand
23
Cache-Anschluß-
baueinheit (Master)
22
Cache (1GB)
3
HSTA4
HSTA3
DKA8
DKA7
DKA6
DKA5
DKA4
DKA3
DKA2
DKA1
HSTA2
HSTA1
2
hinterer Hauptrahmen
1
72
71
3
Cache (1GB)
22
Cache-Anschluß-
baueinheit (Master)
Abschlußwiderstand
Abschlußwiderstand
Cache-Anschluß
(Slave)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache-Einheit 2
Abschlußwiderstand
24
Cache-Anschluß
(Slave)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache (1GB)
Cache-Einheit 1
23
Abschlußwiderstand
24

FIG. 7

FIG. 8


CP : Cache-Anschluß
C : Cache-Speicher-Baueinheit
H : Host-Adapter
D : Platten-Adapter
71
CP
C
H
H
D
D
D
D
D
D
D
D
H
H
C
CP
84
Bedienungs-
seite
83
81
82
81
85

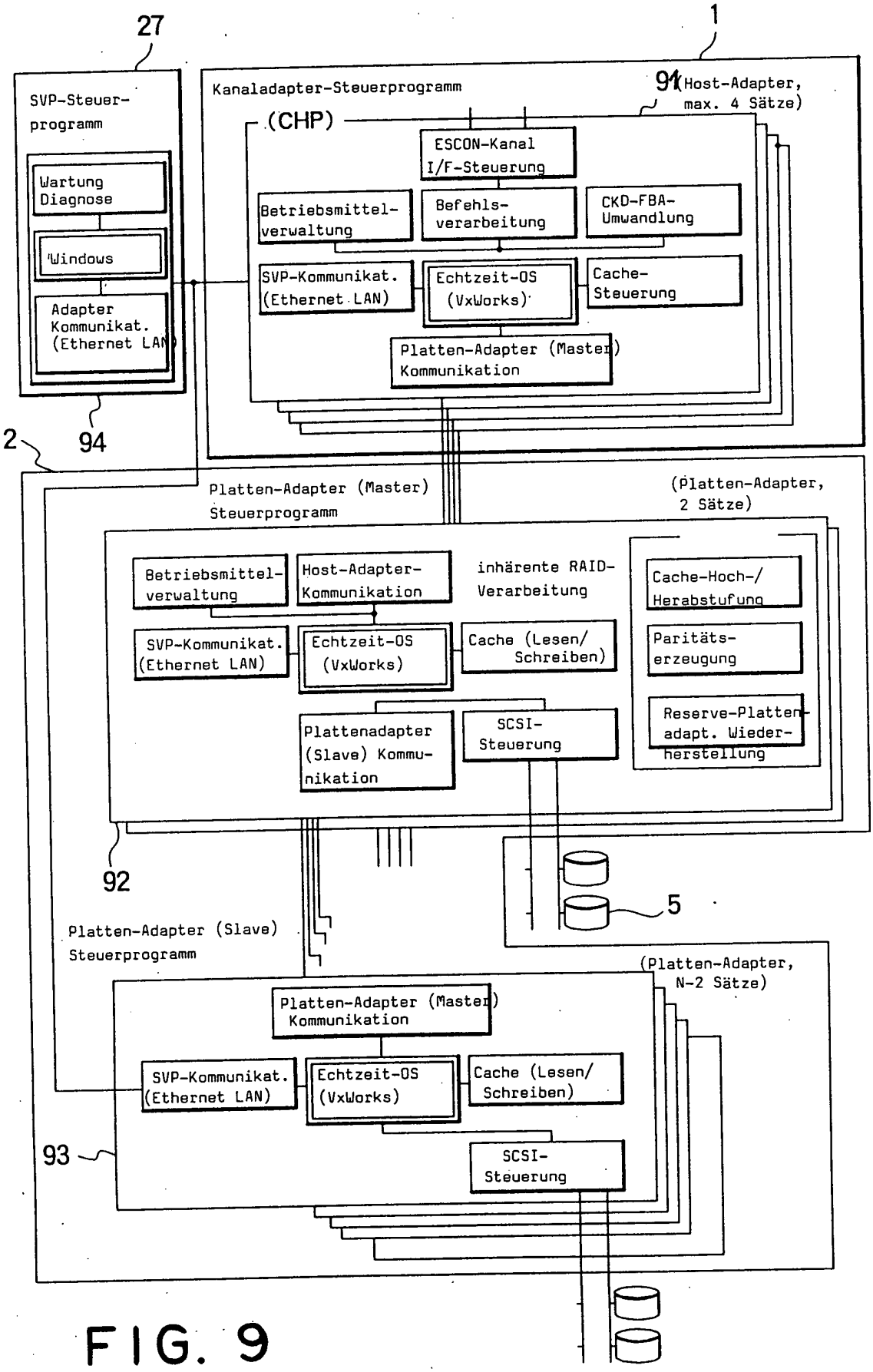

27
1
SVP-Steuer-
programm
Wartung
Diagnose
Windows
Adapter
Kommunikat.
(Ethernet LAN)
94
Kanaladapter-Steuerprogramm
91
Host-Adapter,
max. 4 Sätze)
(CHP)
ESCON-Kanal
I/F-Steuerung
Betriebsmittel-
verwaltung
Befehls-
verarbeitung
CKD-FBA-
Umwandlung
SVP-Kommunikat.
(Ethernet LAN)
Echtzeit-OS
(VxWorks)
Cache-
Steuerung
Platten-Adapter (Master)
Kommunikation
2
Platten-Adapter (Master)
Steuerprogramm
(Platten-Adapter,
2 Sätze)
Betriebsmittel-
verwaltung
Host-Adapter-
Kommunikation
inhärente RAID-
Verarbeitung
Cache-Hoch-/
Herabstufung
SVP-Kommunikat.
(Ethernet LAN)
Echtzeit-OS
(VxWorks)
Cache (Lesen/
Schreiben)
Parităts-
erzeugung
Plattenadapter
(Slave) Kommu-
nikation
SCSI-
Steuerung
Reserve-Platten-
adapt. Wieder-
herstellung
92
5
Platten-Adapter (Slave)
Steuerprogramm
(Platten-Adapter,
N-2 Sätze)
Platten-Adapter (Master)
Kommunikation
SVP-Kommunikat.
(Ethernet LAN)
Echtzeit-OS
(VxWorks)
Cache (Lesen/
Schreiben)
93
SCSI-
Steuerung

FIG. 9

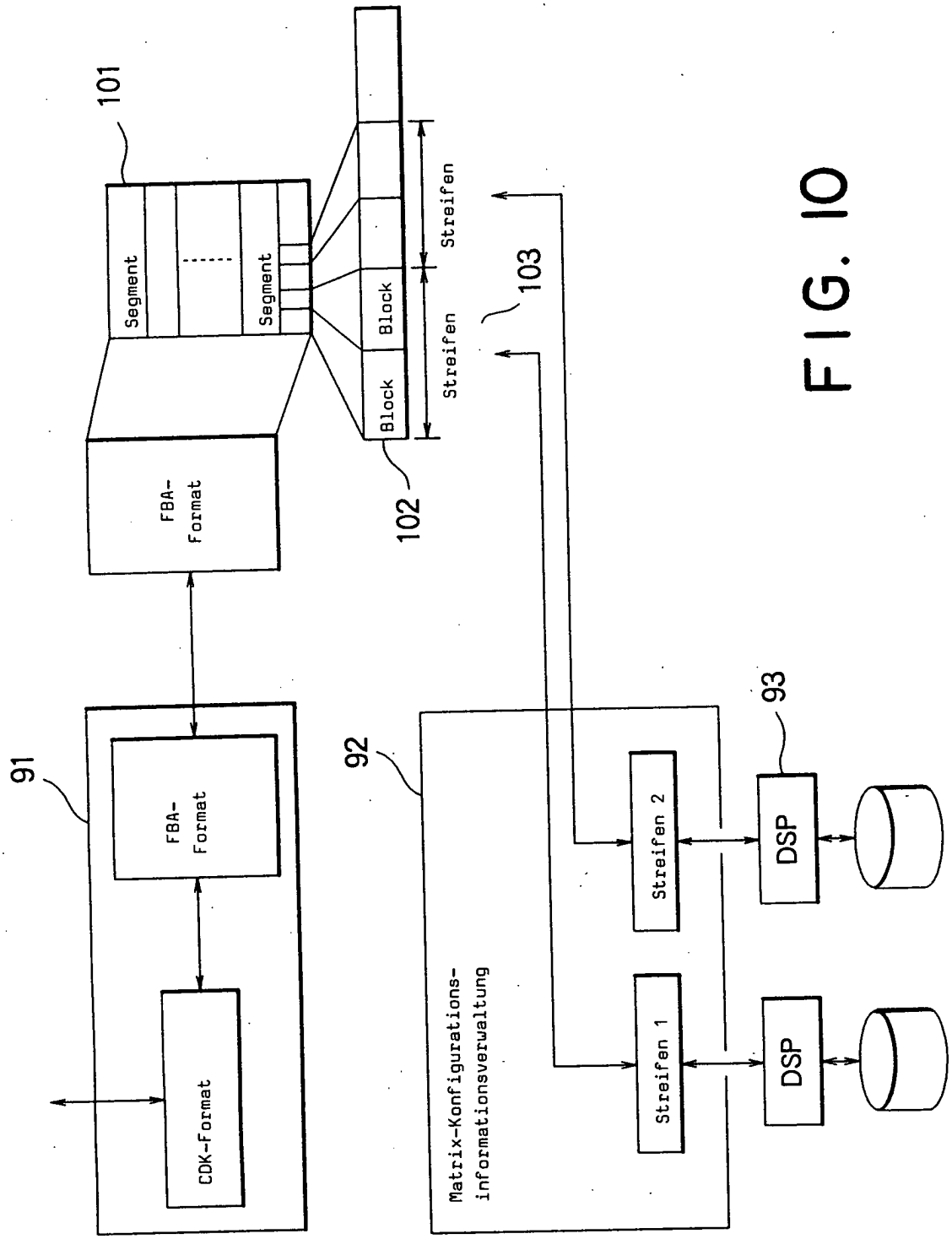

101
Segment
Segment
Block
Block
Streifen
Streifen
102
103
FBA-Format
91
FBA-Format
CDK-Format
92
Matrix-Konfigurations-informationsverwaltung
Streifen 2
Streifen 1
93
DSP
DSP

FIG. 10

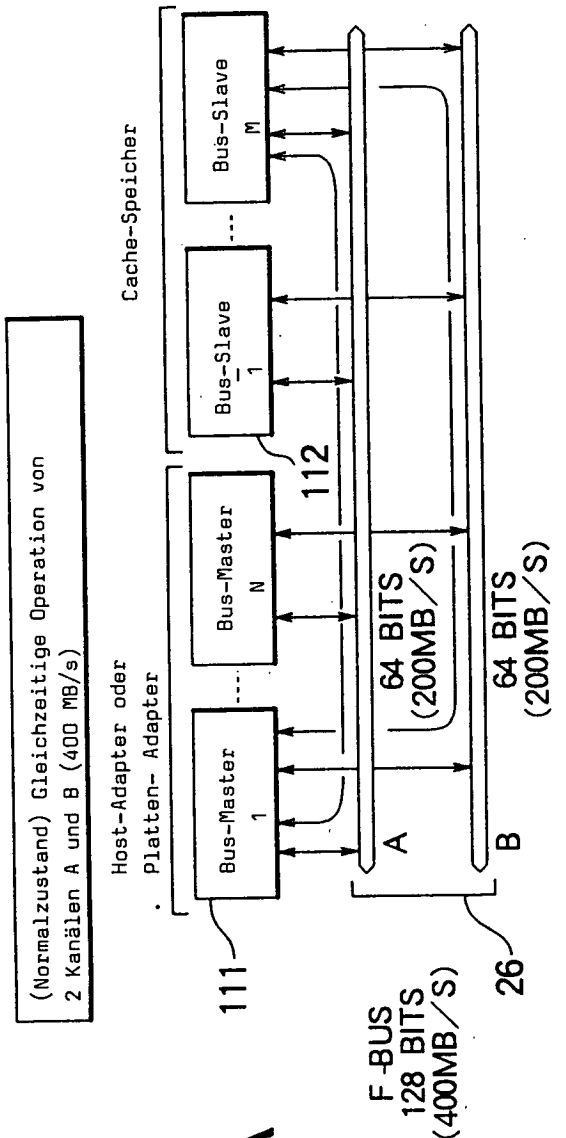


FIG. IIA

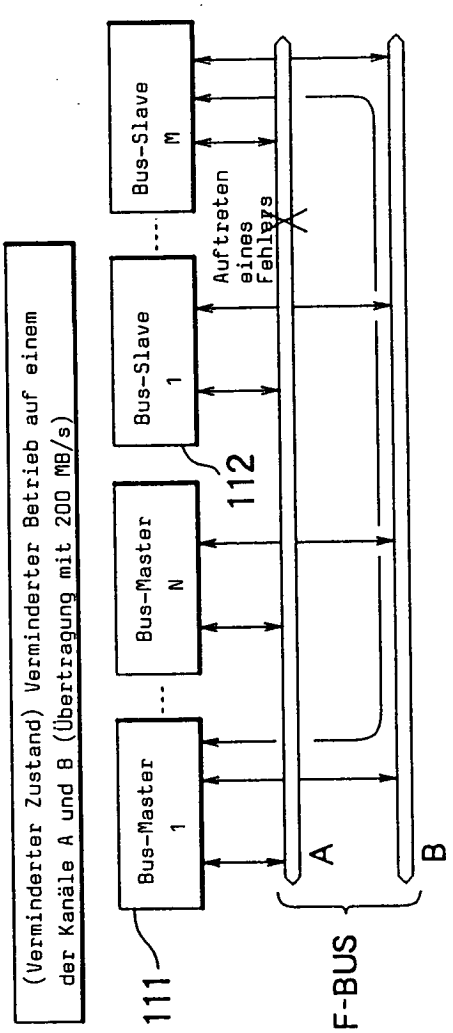


FIG. IIB

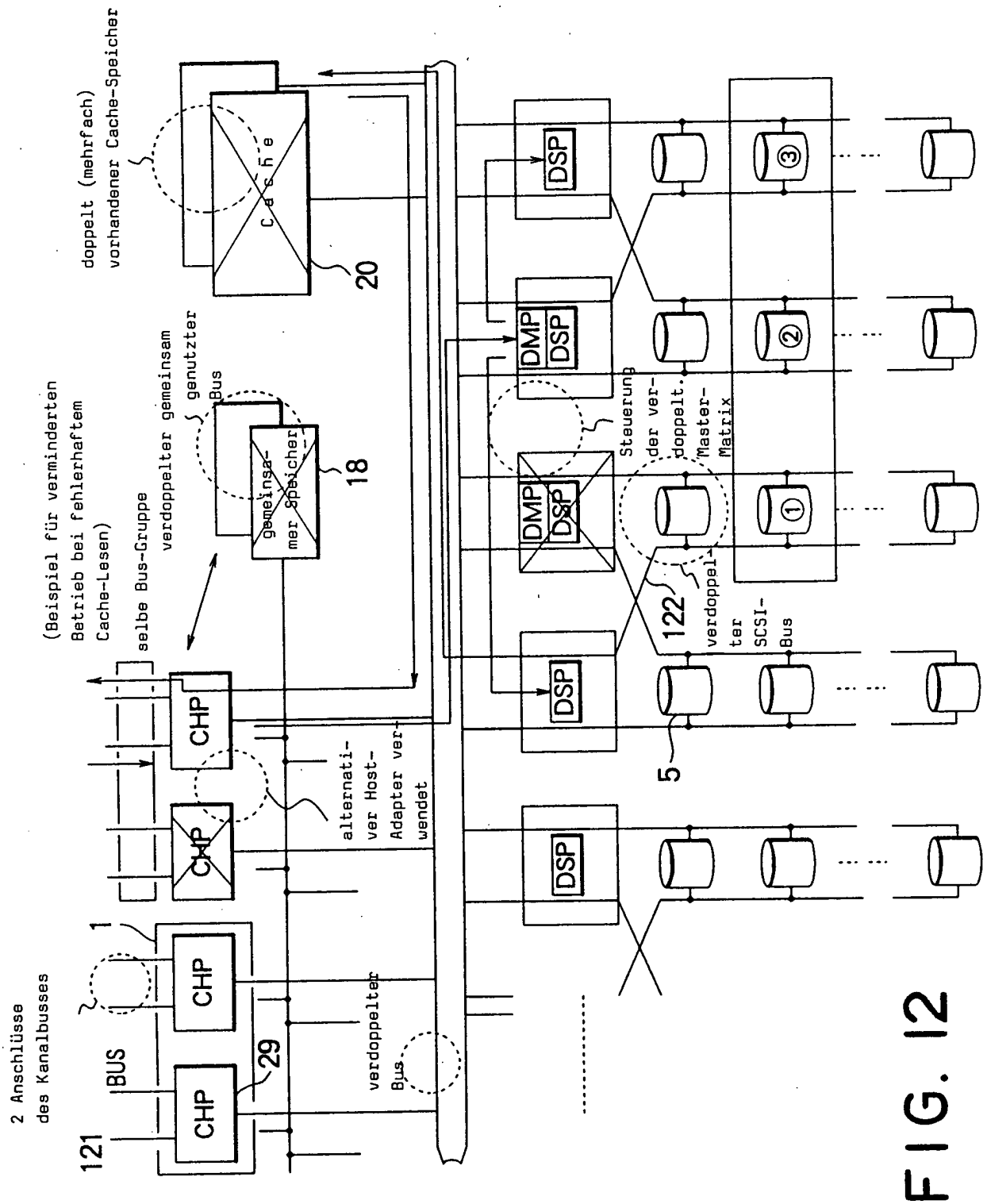

2 Anschlüsse des Kanalbusses
121
BUS
1
CHP
CHP
CHP
CHP
29
alternativer Host-Adapter verwendet
verdoppelter Bus
(Beispiel für verminderten Betrieb bei fehlerhaftem Cache-Lesen)
selbe Bus-Gruppe verdoppelter gemeinsam genutzter Bus
gemeinsamer Speicher
18
doppelt (mehrfach) vorhandener Cache-Speicher
Cache
20
DSP
DMP
DSP
DMP
DSP
DSP
DSP
Steuerung der verdoppelten Master-Matrix
5
122
verdoppelter SCSI-Bus
①
②
③

FIG. 12

FIG. 13

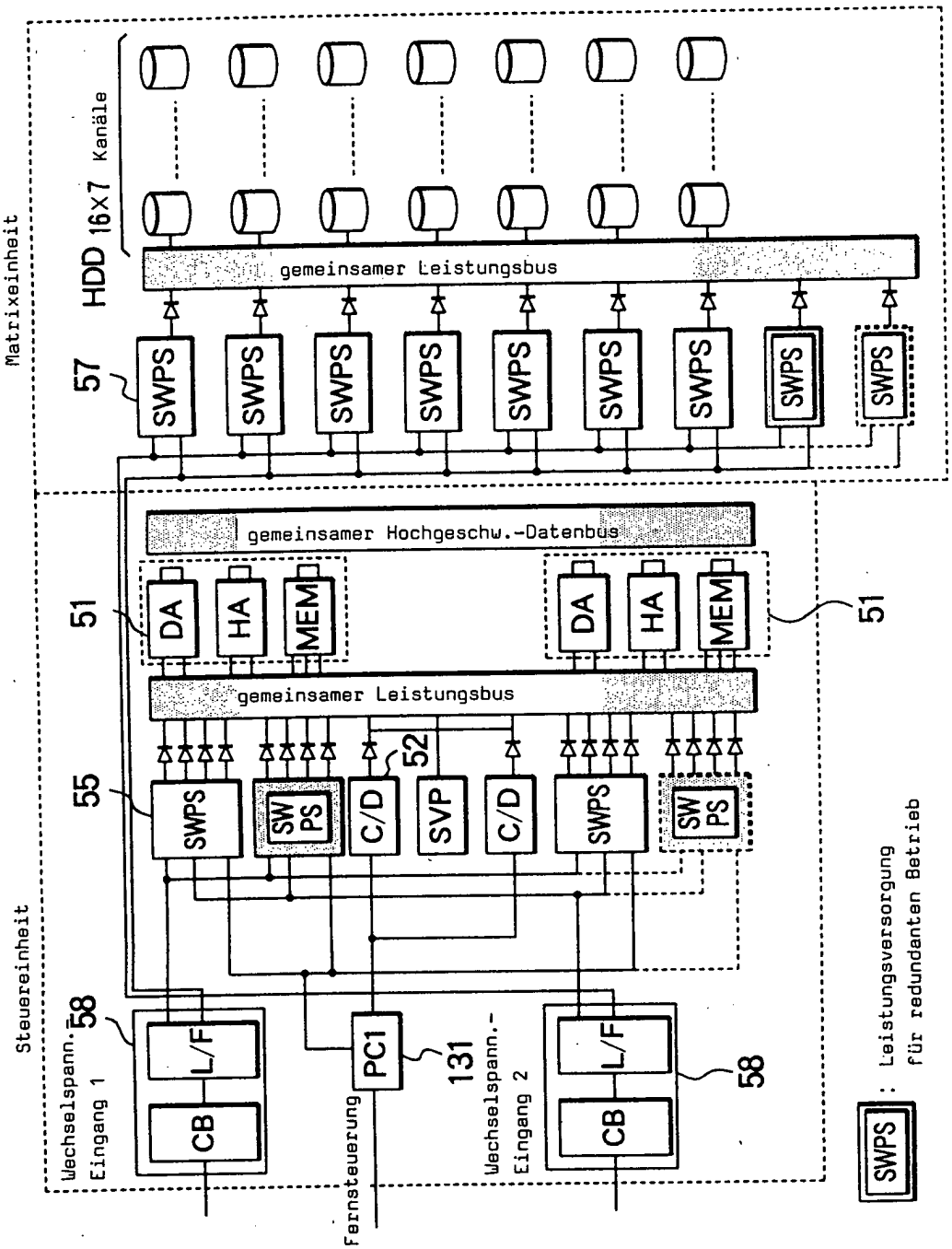

Matrixeinheit
HDD 16X7 Kanäle
gemeinsamer Leistungsbus
57
SWPS
gemeinsamer Hochgeschw.-Datenbus
51
DA
HA
MEM
gemeinsamer Leistungsbus
55
52
SW PS
C/D
SVP
Steuereinheit
58
L/F
CB
Wechselspann. Eingang 1
PC1
Fernsteuerung
131
Wechselspann.- Eingang 2
: Leistungsversorgung für redundanten Betrieb

FIG. 14

Positionnummer	Kapazität der einzelnen Magnetplatte	Matrix-Konfiguration	Matrix-Kapazität
141	3.0 GB (3.5 INCH)	(14D+2)×5	ungefähr 220 GB
142	4.0 GB (5 INCH)	(14D+2)×4	
143	8.4 GB (6.4 INCH)	(14D+2)×2	

FIG. 15

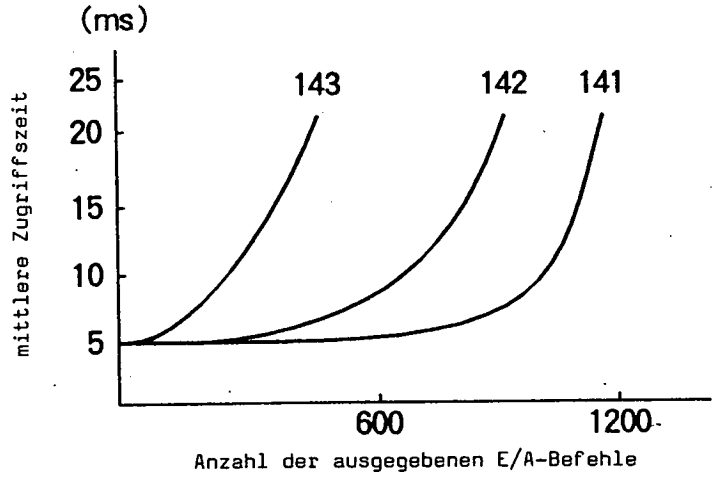


FIG. 16

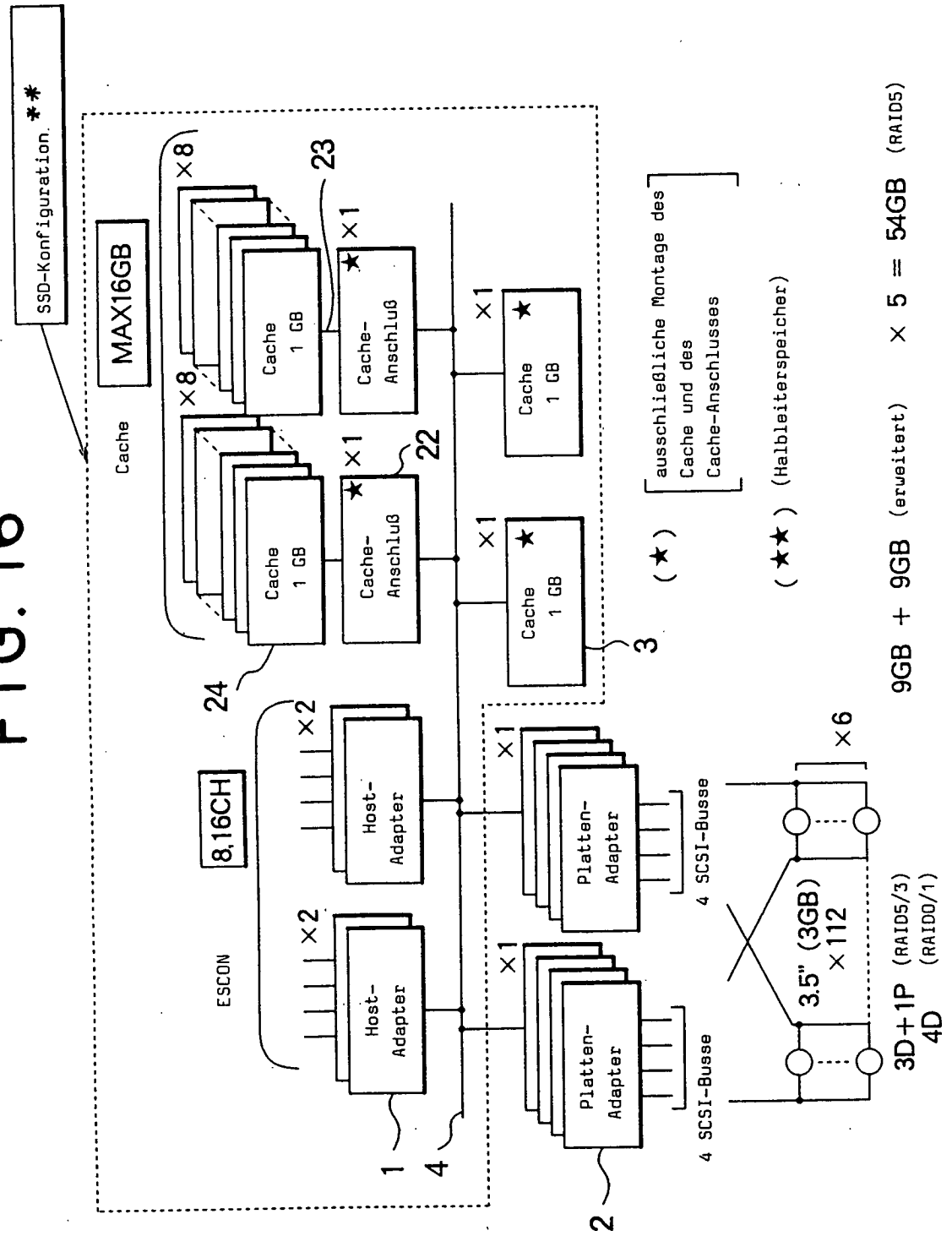

SSD-Konfiguration **
MAX16GB
Cache
×8
×8
Cache 1 GB
Cache 1 GB
23
×1
Cache-Anschluß
×1
22
Cache 1 GB
Cache 1 GB
×1
Cache-Anschluß
×1
Cache 1 GB
24
3
8,16CH
ESCON
×2
×2
Host-Adapter
Host-Adapter
×1
×1
Platten-Adapter
Platten-Adapter
4 SCSI-Busse
4 SCSI-Busse
1
4
2
×6
3.5" (3GB)
×112
3D+1P (RAID5/3)
4D (RAID0/1)
9GB + 9GB (erweitert) × 5 = 54GB (RAID5)
(★) ausschließliche Montage des Cache und des Cache-Anschlusses
(★★) (Halbleiterspeicher)

FIG. 17

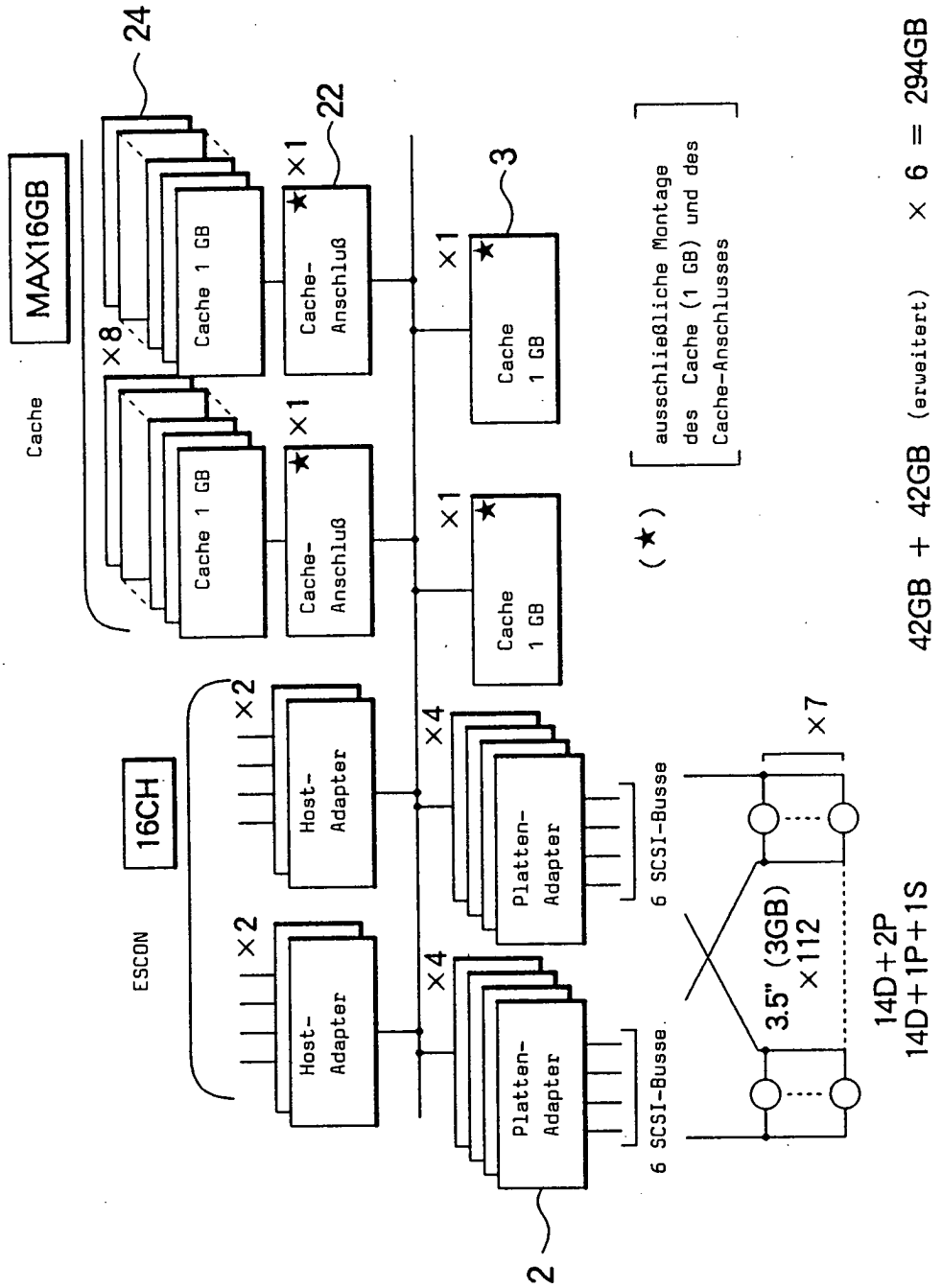

MAX16GB
Cache
24
×8
Cache 1 GB
22
×1
Cache-Anschluß
3
×1
Cache 1 GB
16CH
ESCON
×2
Host-Adapter
×4
Platten-Adapter
6 SCSI-Busse
2
(★)
ausschließliche Montage des Cache (1 GB) und des Cache-Anschlusses
×7
3.5" (3GB) ×112
14D+2P
14D+1P+1S
42GB + 42GB (erweitert) × 6 = 294GB

FIG. 18

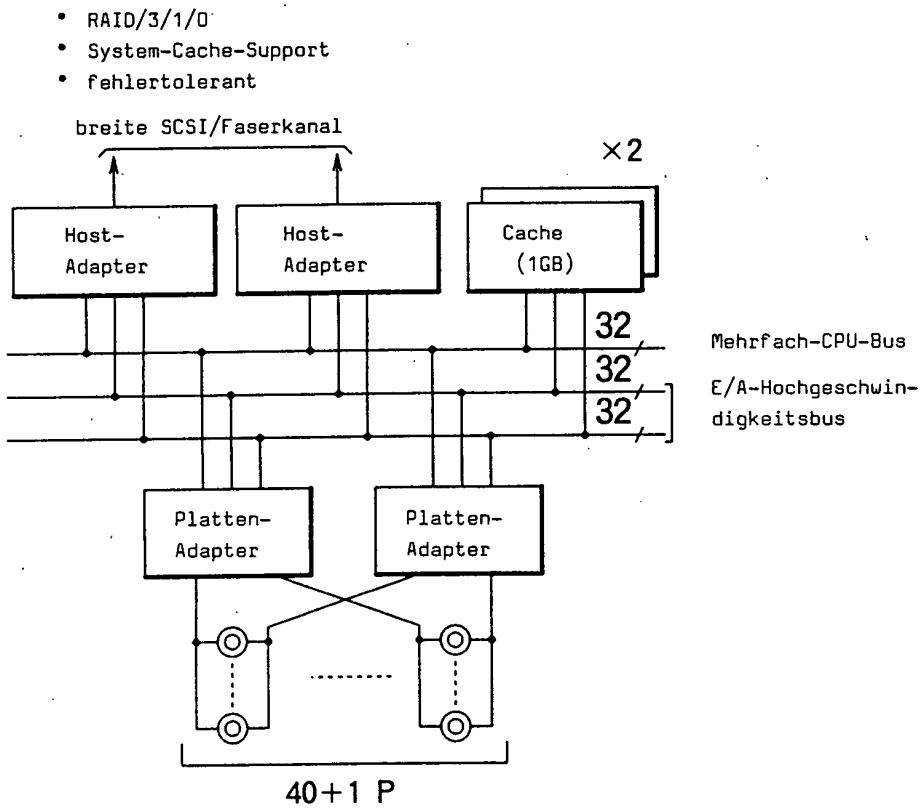

• RAID/3/1/0
• System-Cache-Support
• fehlertolerant
breite SCSI/Faserkanal
×2
Host-Adapter
Host-Adapter
Cache (1GB)
32
32
32
Mehrfach-CPU-Bus
E/A-Hochgeschwindigkeitsbus
Platten-Adapter
Platten-Adapter
40+1 P

FIG. 19

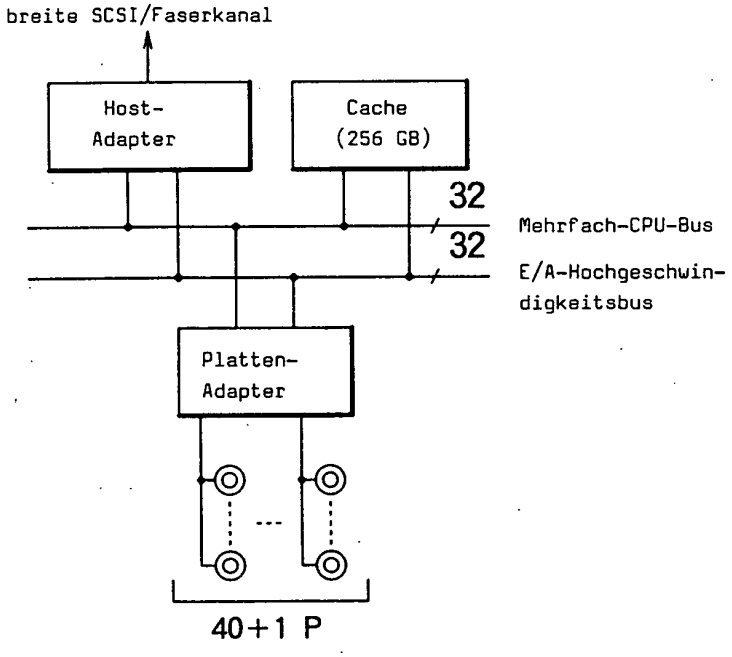
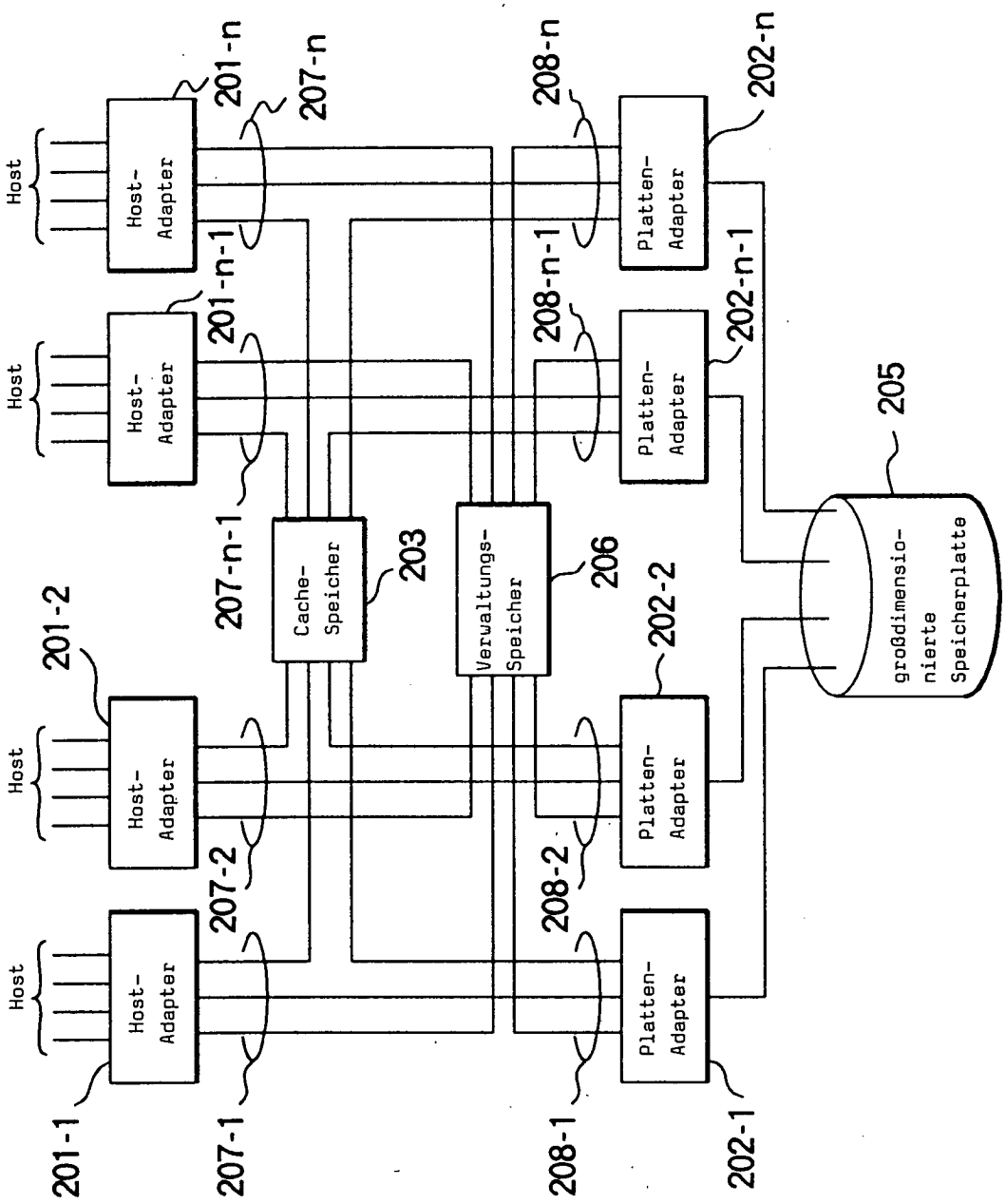

breite SCSI/Faserkanal
Host-Adapter
Cache (256 GB)
32
32
Mehrfach-CPU-Bus
E/A-Hochgeschwindigkeitsbus
Platten-Adapter
40+1 P

FIG. 20


Host
Host-Adapter
Cache-Speicher
Verwaltungs-Speicher
Platten-Adapter
großdimensionierte Speicherplatte
201-1
201-2
201-n-1
201-n
202-1
202-2
202-n-1
202-n
203
205
206
207-1
207-2
207-n-1
207-n
208-1
208-2
208-n-1
208-n