



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 603 16 776 T2** 2008.08.28

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 485 787 B1**

(51) Int Cl.⁸: **G06F 3/06** (2006.01)

(21) Deutsches Aktenzeichen: **603 16 776.4**

(86) PCT-Aktenzeichen: **PCT/US03/03409**

(96) Europäisches Aktenzeichen: **03 706 086.0**

(87) PCT-Veröffentlichungs-Nr.: **WO 2003/071412**

(86) PCT-Anmeldetag: **03.02.2003**

(87) Veröffentlichungstag
der PCT-Anmeldung: **28.08.2003**

(97) Erstveröffentlichung durch das EPA: **15.12.2004**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **10.10.2007**

(47) Veröffentlichungstag im Patentblatt: **28.08.2008**

(30) Unionspriorität:
80063 19.02.2002 US

(84) Benannte Vertragsstaaten:
**AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,
GR, HU, IE, IT, LI, LU, MC, NL, PT, SE, SI, SK, TR**

(73) Patentinhaber:
Intel Corp., Santa Clara, Calif., US

(72) Erfinder:
SMITH, Hubbert, Sandy, UT 84093, US

(74) Vertreter:
**Rummler, F., Dipl.-Ing.Univ., Pat.-Anw., 80802
München**

(54) Bezeichnung: **AUF NETZWERKDATENSPEICHERUNG BEZOGENE OPERATIONEN**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

HINTERGRUND

[0001] Bei einem Typ eines herkömmlichen Computernetzwerks werden ein oder mehrere Server-Computerknoten oder „Server“ bereitgestellt, welche einem oder mehreren Arrays von Festplattenmassenspeichervorrichtungen zugeordnet sind (d. h. Daten damit austauschen und steuern), in welchen Benutzerdaten, welche von Prozessen erzeugt werden, welche in dem Netzwerk ausgeführt werden, gespeichert werden können, und aus welchen derartige Daten von derartigen Prozessen abgerufen werden können. Mindestens einer dieser Arrays kann ein redundanter Array unabhängiger Festplatten sein (hier nachfolgend „redundanter Array“ genannt), in welchem eine Kopie (oder ein „Spiegel“) eines primären Datenträgers, welcher in einem primären Array von Festplattenmassenspeichervorrichtungen (hier nachfolgend „primärer Array“ genannt) gespeichert ist, repliziert werden kann, und aus welchem der replizierte Datenträger (hier nachfolgend „der redundante Datenträger“ genannt) in dem Fall eines Defekts des primären Arrays wiedererlangt werden kann. Hier nachfolgend wird gesagt, dass ein redundanter Array einem primären Array „zugeordnet“ ist, wenn der redundante Array einen redundanten Datenträger speichert, welcher eine Kopie eines primären Datenträgers in dem primären Array ist. Hier nachfolgend wird auch gesagt, dass ein redundanter Datenträger einem primären Datenträger zugeordnet ist, wenn der redundante Datenträger ein Spiegel des primären Datenträgers ist.

[0002] In diesem herkömmlichen Netzwerk kann ein Server (hier nachfolgend ein „Ursprungs“-Server genannt), welcher einem primären Array zugeordnet ist, über eine Netzwerkkommunikationsverbindung an einen anderen Server gekoppelt sein (hier nachfolgend ein „Ziel“-Server genannt), welcher einem redundanten Array zugeordnet ist, und der redundante Array kann dem primären Array zugeordnet sein. Der Ursprungs-Server und der Ziel-Server können jeweils eine jeweilige Netzwerkschnittstellenkarte (NIC), einen Host-Bus-Adapter (HBA), eine Zentraleinheit (CPU) und einen Systemspeicher umfassen. Sowohl in dem Ursprungs- als auch in dem Ziel-Server sind die jeweilige CPU und der jeweilige Systemspeicher durch einen jeweiligen Systembus miteinander verbunden, die jeweilige NIC und der jeweilige HBA sind durch einen jeweiligen Eingangs-/Ausgangs-(E/A)-Bus und ein E/A-Steuersystem miteinander verbunden. Die NIC in dem Ursprungs-Server ist mit der NIC in dem Ziel-Server durch die Kommunikationsverbindung gekoppelt, der HBA in dem Ursprungs-Server ist mit dem primären Array gekoppelt und steuert ihn, und der HBA in dem Ziel-Server ist mit dem redundanten Array gekoppelt und steuert ihn. Jeweilige Programmprozesse befinden sich in

den jeweiligen Systemspeichern der Server, welche die jeweiligen CPUs der Server steuern, um so zu bewirken, dass die Server herkömmliche, synchrone Datenreplikationsoperationen ausführen. Die jeweiligen Programmprozesse, welche sich in den jeweiligen Systemspeichern befinden, umfassen jeweilige Betriebssystem-, RAID-Treiber-, Replikationstreiber- und Netzwerktreiber-/Kommunikationsstapelprozesse. Datenbankprozesse auf Anwendungsebene befinden sich auch in dem Systemspeicher des Ursprungs-Servers.

[0003] Ein menschlicher Benutzer des Ursprungs-Servers kann über eine Benutzerschnittstelle eine Anforderung an einen Programmprozess auf Anwendungsebene ausgeben, welcher sich in dem Ursprungs-Server befindet, um Benutzerdaten auf einem primären Datenträger in dem primären Array zu speichern. Als Reaktion auf diese Anforderung erzeugt der Prozess eine Datenschreibenanforderung an einen oder mehrere O/S-Prozesse in dem Ursprungs-Server, welche bewirken, dass die CPU des Ursprungs-Servers die Benutzerdaten über den Systembus des Ursprungs-Servers an den Systemspeicher des Ursprungs-Servers überträgt und die Daten in ihm speichert. Danach geben diese O/S-Prozesse eine Anforderung an den RAID-Treiber des Ursprungs-Servers aus, welcher bewirkt, dass die CPU des Ursprungs-Servers die Daten aus dem System des Ursprungs-Servers über den Systembus des Ursprungs-Servers abrufen, um die Daten zu einem oder mehreren Datenblöcken zu fragmentieren (z. B. Datenblöcke nach dem SCSI-(Small Computer Systems Interface)-Protokoll), um die Datenblöcke an den Systemspeicher des Ursprungs-Servers zu übertragen und sie darin zu speichern und um zu bewirken, dass die E/A-Bus-Steuerung des Ursprungs-Servers die Blöcke über den Systembus des Ursprungs-Servers aus dem Systemspeicher des Ursprungs-Servers abrufen und die Blöcke an den HBA des Ursprungs-Servers weiterleitet. Der HBA des Ursprungs-Servers schreibt dann die Datenblöcke auf den primären Datenträger, und wenn die Datenblöcke erfolgreich auf den primären Datenträger geschrieben wurden, kehrt der HBA des Ursprungs-Servers über das E/A-Bus- und Steuersystem des Ursprungs-Servers zu der CPU des Ursprungs-Servers zurück, und der Systembus gibt eine Meldung „Schreiben abgeschlossen“ aus, um dies anzugeben.

[0004] Auch als Reaktion auf die Anforderung von dem Prozess auf Anwendungsebene fordern die O/S-Prozesse des Ursprungs-Servers an, dass die Replikationstreiberprozesse des Ursprungs-Servers der CPU des Ursprungs-Servers befehlen, eine Kopie der Benutzerdaten zu erzeugen und die Kopie der Benutzerdaten zu einem oder mehreren Replikationsdatenpaketen zu fragmentieren und die Pakete in dem Systemspeicher des Ursprungs-Servers zu

speichern. Nachdem diese Pakete in dem Systempeicher des Ursprungs-Servers gespeichert wurden, bewirken die Netzwerktreiberprozesse des Ursprungs-Servers, dass die CPU des Ursprungs-Servers die Pakete über den Systembus des Ursprungs-Servers aus dem Systempeicher des Ursprungs-Servers abrufen, um die Pakete mit Kopf- und anderen Informationen zu verkapseln, um so einen oder mehrere Ethernet-Frames zu bilden, die Frames über den Systembus, den E/A-Bus und das Steuersystem des Ursprungs-Servers an die NIC des Ursprungs-Servers weiterzuleiten und zu bewirken, dass die NIC des Ursprungs-Servers die Frames über die Ethernet-Verbindung an die NIC des Ziel-Servers weiterleitet.

[0005] Die NIC des Ziel-Servers empfängt die Frames, und als Reaktion auf die empfangenen Frames bewirken die Betriebssystem-, RAID-Treiber-, Replikationstreiber- und Netzwerktreiber-/Kommunikationsstapelprozesse des Ziel-Servers, dass die CPU des Ziel-Servers Operationen durchführt, welche die Datenpakete aus den Frames entkapseln, die Benutzerdaten aus den Datenpaketen wiederherstellen, die rekonstruierten Benutzerdaten zu einem oder mehreren Datenblöcken fragmentieren und die Blöcke auf dem redundanten Datenträger speichern. Nachdem alle Benutzerdaten erfolgreich auf den redundanten Datenträger geschrieben wurden, bewirken die O/S-, Replikationstreiber- und Netzwerktreiberprozesse des Ziel-Servers, dass die CPU des Ziel-Servers Operationen durchführt, welche einen Frame mit einer Erfolgsmeldung über die Datenreplikation, welche eine erfolgreiche Replikation der Daten in dem redundanten Array angibt, erzeugen und an den Ursprungs-Server weiterleiten (d. h. über die NIC des Ziel-Servers und die Kommunikationsverbindung). Es kann angemerkt werden, dass die oben stehend genannten Operationen, welche von der CPU des Ziel-Servers ausgeführt werden, eine wesentliche Anzahl von Transaktionen/Interaktionen zwischen u. a. der CPU, dem Systempeicher und dem Systembus des Ziel-Servers einbeziehen.

[0006] Es muss eine übermäßig große Anzahl von CPU-, Systempeicher- und Systembus-Transaktionen und -Operationen in dem Ursprungs- und in dem Ziel-Server durchgeführt werden, um die oben stehend genannten datenspeicherbezogenen Operationen in diesem herkömmlichen Netzwerk auszuführen. Bedauerlicherweise bewirkt dies, dass übermäßige Anteile der CPU-, Systempeicher- und Systembus-Verarbeitungsbetriebsmittel des Ursprungs- und Ziel-Servers verbraucht werden, um derartige Operationen auszuführen.

[0007] US 6209002 B1 und EP 1150210 A beschreiben Systeme nach dem Stand der Technik, welche für die Merkmale des Oberbegriffs der selbständigen Ansprüche relevant sind.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0008] Merkmale und Vorteile der Ausführungsformen der vorliegenden Erfindung, welche in den angehängten selbständigen Ansprüchen 1 und 4 ausführlich definiert werden, werden mit der Fortsetzung der folgenden ausführlichen Beschreibung und unter Verweis auf die Zeichnungen offenkundig, wobei gleiche Bezugszeichen gleiche Bauteile darstellen, und wobei:

[0009] [Fig. 1](#) ein Diagramm ist, welches Komponenten eines Datenspeichernetzwerks illustriert, bei welchem eine Ausführungsform der vorliegenden Erfindung vorteilhaft in die Praxis umgesetzt werden kann.

[0010] [Fig. 2](#) ein Diagramm ist, welches Felder illustriert, welche in einem Befehls-/Meldungs-Frame enthalten sind, welcher zwischen einem Ursprungs- und einem Ziel-Server in dem Netz der [Fig. 1](#) weitergeleitet werden kann.

[0011] [Fig. 3](#) ein Diagramm ist, welches Komponenten eines HBA illustriert, welcher in einem Ursprungs-Server in dem Netzwerk der [Fig. 1](#) enthalten ist.

[0012] [Fig. 4](#) ein Diagramm ist, welches Programmprozesse illustriert, welche sich in einem Systempeicher in einem Ursprungs-Server in dem Netzwerk der [Fig. 1](#) befinden.

[0013] [Fig. 5](#) ein Diagramm ist, welches Programmprozesse illustriert, welche sich in einem Systempeicher in einem Ziel-Server in dem Netzwerk der [Fig. 1](#) befinden.

[0014] [Fig. 6](#) ein Diagramm ist, welches Komponenten eines HBA illustriert, welcher in einem Ziel-Server in dem Netzwerk der [Fig. 1](#) enthalten ist.

[0015] [Fig. 7](#) ein Ablaufdiagramm ist, welches Operationen illustriert, welche in dem Netzwerk der [Fig. 1](#) durchgeführt werden können.

[0016] [Fig. 8](#) ein Diagramm ist, welches Programmprozesse illustriert, welche sich in dem Speicher in dem HBA der [Fig. 3](#) befinden können.

[0017] [Fig. 9](#) ein Diagramm ist, welches Programmprozesse illustriert, welche sich in dem Speicher in dem HBA der [Fig. 6](#) befinden können.

[0018] Es versteht sich, dass, obwohl die folgende ausführliche Beschreibung unter Verweis auf beispielhafte Ausführungsformen der vorliegenden Erfindung fortfährt, Durchschnittsfachleuten viele Alternativen, Modifikationen und Variationen davon offenkundig sind. Dementsprechend ist es vorgesehen,

dass die vorliegende Erfindung allgemein gesehen wird und nur definiert wird, wie in den begleitenden Ansprüche dargelegt.

AUSFÜHRLICHE BESCHREIBUNG

[0019] Unter Verweis auf [Fig. 1](#) bis [Fig. 9](#) werden nun beispielhafte Ausführungsformen der vorliegenden Erfindung beschrieben. [Fig. 1](#) ist ein höchst schematisches Diagramm, welches Komponenten eines Datenspeichernetzwerks **10** illustriert, wobei eine Ausführungsform der vorliegenden Erfindung vorteilhaft in die Praxis umgesetzt werden kann. Das Netzwerk **10** kann einen Ursprungs-Server-Knoten **12**, einen Ziel-Server-Knoten **14** und ein Datenkommunikationsnetzwerk **11** aufweisen. Das Netzwerk **11** kann ein oder mehrere lokale und/oder Weitverkehrsnetzwerke umfassen, welche ein oder mehrere wohlbekanntes Netzwerkkommunikationsprotokolle (z. B. Ethernet und/oder TCP/IP-(Transmission Control Protocol/Internet Protocol)-basierte Protokolle) einsetzen können. Der Ursprungs-Server **12** kann über Netzwerkverbindungen **34** und **36** an das Netzwerk **11** gekoppelt sein, und der Ziel-Server **14** kann über Netzwerkverbindungen **38** und **40** an das Netzwerk **11** gekoppelt sein, wobei dies auf eine derartige Weise erfolgt, dass den Servern **12** und **14** gestattet wird, miteinander über das Netzwerk **11** zu kommunizieren. Der Ursprungs-Server **12** und der Ziel-Server **14** können geografisch voneinander entfernt sein.

[0020] Der Ursprungs-Server **12** weist die CPU **16**, den Systemspeicher **18**, die Benutzerschnittstelle **20**, das Host-Brücken-/Bussystem **22**, das E/A-Brücken-/Bussystem **24**, den HBA **26**, den HBA **28** und zwei Arrays oder Gruppen von Festplattenmassenspeichervorrichtungen **30** bzw. **32** auf. Jeder der Arrays oder jede der Gruppen von Festplattenvorrichtungen **30** und **32** kann eine oder mehrere (oder, wie es bei dieser Ausführungsform der vorliegenden Erfindung der Fall ist, eine Vielzahl) jeweiliger Festplattenmassenspeichervorrichtungen umfassen (z. B. optische oder magnetische Festplattenmassenspeichervorrichtungen). Falls der Server **12** ersatzweise auf Wegen passend modifiziert ist, welche Durchschnittsfachleuten offenkundig sind, kann jeder der Arrays **30** und **32** stattdessen jeweilige Halbleiter-Massenspeichervorrichtungen umfassen.

[0021] Der HBA **28** kann elektrisch über eine jeweilige S-ATA-(Serial Advanced Technology Attachment)-basierte Schnittstelle **210** (man siehe [Fig. 3](#)) mit dem Array **32** gekoppelt sein, und kann konfiguriert werden, mit diesem Daten und Befehle auszutauschen; durch Erteilen geeigneter Befehle an den Array **32** über die Schnittstelle **210** kann der HBA **28** bewirken, dass Daten, auf welche in den Befehlen verwiesen wird oder in ihnen enthalten sind, in dem Array **32** gespeichert werden oder aus ihm abgerufen werden. Der HBA **28** kann auch einen E/A-Prozessor

27, einen HBA-Speicher **202**, eine Netzwerkschnittstelle **204** und eine E/A-Brücken-/Bussystemschnittstelle **208** aufweisen. In dem HBA **28** können der Prozessor **27**, der Speicher **202** und die Schnittstellen **204**, **208** und **210** durch ein PCI-(Peripheral Component Interconnect)- (z. B. von einem Typ, welcher in der PCI Local Bus Specification Rev. 2.2, herausgegeben von PCI Special Interest Group, 5440 S. W. Westgate Drive, Nr. 217, Portland, Oregon 97221, Vereinigte Staaten von Amerika, beschrieben wird oder dazu kompatibel ist und hier nachfolgend „PCI“ genannt wird)- oder durch ein PCI-X-(PCI-Extended)- (z. B. von einem Typ, welcher in der PCI-X Specification Rev. 1.0a, herausgegeben und erhältlich von der PCI Special Interest Group, beschrieben wird oder dazu kompatibel ist und hier nachfolgend „PCI-X“ genannt wird)-Bussystem **200** miteinander verbunden sein und unter seiner Verwendung Daten und Befehle miteinander austauschen. Natürlich kann das Bussystem **200** stattdessen andere Typen von Bussystemen umfassen ohne von dieser Ausführungsform der vorliegenden Erfindung abzuweichen. Der Prozessor **27** kann Chips mit integrierten Schaltkreisen aufweisen, welche in einem Chipsatz mit integrierten Schaltkreisen enthalten sind, welcher von dem Inhaber der vorliegenden Anmeldung im Handel erhältlich ist, (z. B. Chipsatz Intel® 80310, welcher den Prozessor Intel® 80200 und den Begleit-Chip Intel® 80312 aufweist), obwohl der E/A-Prozessor **27** ersatzweise stattdessen andere Chips mit integrierten Schaltkreisen (z. B. den E/A-Prozessor Intel® 80960 RM/RN oder den Prozessor Intel® 80321) oder andere Typen von Prozessoren/integrierten Schaltkreisen umfassen kann, ohne von dieser Ausführungsform der vorliegenden Erfindung abzuweichen. Die Netzwerkschnittstelle **204** kann mit der Verbindung **34** gekoppelt sein, kann konfiguriert werden, die physikalischen Operationen zu ermöglichen/auszuführen, welche dem HBA **28** und dem Server **12** gestatten, mit dem Knoten **14** in dem Netzwerksystem **10** zu kommunizieren. Die Brückenschnittstelle **208** ist mit dem E/A-Brücken-/Bussystem **24** gekoppelt und ist konfiguriert, dem HBA **28** zu ermöglichen, Daten und Befehle mit dem System **24** auszutauschen. Der Speicher **202** kann Direktzugriffsspeicher (RAM), welcher verwendet werden kann, um Programmprozesse **207** zu speichern (und in welchem diese sich befinden können), Hochgeschwindigkeits-Cache-Speicher, welcher verwendet werden kann, um Daten zu speichern, welche zur Speicherung in dem Array **32** vorgesehen sind oder welche kürzlich daraus abgerufen wurden, und nichtflüchtigen, elektrisch löschbaren, programmierbaren Speicher (EEPROM) umfassen, welcher verwendet werden kann, um Daten, welche in dem Cache-Speicher gespeichert sind, in dem Fall zu speichern, dass ein Ausfall der externen Stromversorgung des HBA auftritt. Obwohl nicht in den Figuren gezeigt, kann der HBA **28** ein System aufweisen, um zu erkennen, wenn ein derartiger Ausfall aufgetreten ist, und um den Prozes-

sor **27** zu triggern, dass bewirkt wird, dass der Speicher **202** die Daten in dem EEPROM speichert, welche in dem Cache-Speicher gespeichert sind; ein batteriebetriebenes Notfallstromversorgungssystem (nicht gezeigt) kann in den HBA **28** eingeschlossen sein, um den HBA **28** während dieser Operationen mit Strom zu versorgen. Der Prozessor **27** kann programmiert werden, die Daten, welche in dem EEPROM gespeichert sind, bei einem nachfolgenden Einschaltvorgang des HBA **28** in den Cache-Speicher abzurufen. Zusätzliche Einzelheiten, welche die Konstruktion und den Betrieb des HBA **28** betreffen, werden nachfolgend dargestellt.

[0022] Obwohl in den Figuren nicht gezeigt, versteht es sich, dass die jeweilige Konstruktion des HBA **26** mit der entsprechenden Konstruktion des HBA **28** im Wesentlichen identisch sein kann. Um eine unnötige doppelte Beschreibung zu vermeiden, wird die Konstruktion des HBA **26** hier folglich nicht ausführlich dargestellt. Es genügt jedoch anzumerken, dass der HBA **26** über eine jeweilige S-ATA-basierte Schnittstelle (nicht gezeigt), welche in dem HBA **26** enthalten ist, elektrisch mit dem Array **30** gekoppelt sein kann und konfiguriert werden kann, Daten und Befehle mit ihm auszutauschen; durch Erteilen geeigneter Befehle an den Array **30** über diese ATA-basierte Schnittstelle kann der HBA **26** bewirken, dass Daten, auf welche mit derartigen Befehlen verwiesen wird oder welche in ihnen enthalten sind, in Array **30** gespeichert werden oder aus ihm abgerufen werden. Der HBA **26** kann einen jeweiligen E/A-Prozessor **25** umfassen, welcher eine jeweilige Konstruktion aufweisen kann, welche zu derjenigen des E/A-Prozessors **27** im Wesentlichen gleich ist. Der HBA **26** kann auch eine jeweilige Netzwerkschnittstelle (nicht gezeigt) umfassen, welche eine jeweilige Konstruktion aufweisen kann, welche zu derjenigen der Schnittstelle **204** im Wesentlichen gleich ist, und über die Verbindung **36** an das Netzwerk **11** gekoppelt sein kann.

[0023] Das E/A-Brücken-/Bussystem **24** und die Host-Brücken-/Bussysteme **22** können ein E/A-Steuers-Sternkopplersystem bzw. ein Speichersteuers-Sternkopplersystem umfassen, welche unter Verwendung von Chips mit integrierten Schaltkreisen aufgebaut werden können, welche in einem Chipsatz mit integrierten Schaltkreisen enthalten sind, welcher von dem Inhaber der vorliegenden Anmeldung im Handel erhältlich ist (z. B. der Chipsatz Intel® 860). Die CPU **16**, der Systemspeicher **18** und die Benutzerschnittstelle **20** können über das Host-Brücken-/Bussystem **22** miteinander und mit dem E/A-Brücken-/Bussystem **24** elektrisch gekoppelt sein. Das System **22** ist konfiguriert, einen Austausch von Daten und Befehlen zwischen der CPU **16**, dem Speicher **18**, der Benutzerschnittstelle **20** und dem System **24** zu gestatten/zu ermöglichen. Zusätzlich ist das System **24** konfiguriert, einen Austausch von

Daten und Befehlen zwischen dem System **22** und dem Adapter **26** und **28** zu gestatten/zu ermöglichen.

[0024] Der Speicher **18** kann RAM umfassen, welcher verwendet werden kann, Programmprozesse **19** zu speichern (und in welchem diese sich befinden können). Der RAM im Speicher **18** kann auch Benutzerdaten speichern, welche zur Speicherung in den Arrays **30** und **32** vorgesehen sind, oder welche aus ihnen abgerufen wurden. Die CPU **16** kann einen Prozessor Intel® Xeon™ von dem Typ umfassen, welcher von dem Inhaber der vorliegenden Anmeldung im Handel erhältlich ist, obwohl die CPU **16** andere Typen von Prozessoren umfassen kann, ohne von der vorliegenden Erfindung abzuweichen.

[0025] Der Ziel-Server **14** umfasst die CPU **42**, den Systemspeicher **44**, die Benutzerschnittstelle **46**, das Host-Brücken-/Bussystem **48**, das E/A-Brücken-/Bussystem **50**, den HBA **52**, den HBA **54** und zwei Arrays oder Gruppen von Festplattenmassenspeichervorrichtungen **56** bzw. **58**. Jeder der Arrays oder jede der Gruppen von Festplattenmassenspeichervorrichtungen **56** und **58** kann eine oder mehrere (oder, wie es bei dieser Ausführungsform der vorliegenden Erfindung der Fall ist, eine Vielzahl) jeweiliger Festplattenmassenspeichervorrichtungen umfassen (z. B. optische oder magnetische Festplattenmassenspeichervorrichtungen). Falls der Server **14** ersatzweise auf Wegen passend modifiziert ist, welche Durchschnittsfachleuten offenkundig sind, kann jeder der Arrays **56** und **58** stattdessen jeweilige Halbleiter-Massenspeichervorrichtungen umfassen.

[0026] Der HBA **52** kann elektrisch über eine jeweilige herkömmliche S-ATA-basierte Schnittstelle **510** (man siehe [Fig. 6](#)) mit dem Array **56** gekoppelt sein, und kann konfiguriert sein, Daten und Befehle mit ihm auszutauschen; durch Erteilen geeigneter Befehle an den Array **56** über die Schnittstelle **510** kann der HBA **52** bewirken, dass Daten, auf welche in den Befehlen verwiesen wird oder in ihnen enthalten sind, in dem Array **56** gespeichert werden oder aus ihm abgerufen werden. Der HBA **52** kann auch einen E/A-Prozessor **53**, einen HBA-Speicher **502**, eine Netzwerkschnittstelle **504** und eine E/A-Brücken-/Bussystemschnittstelle **508** aufweisen. In dem HBA **52** können der Prozessor **53**, der Speicher **502** und die Schnittstellen **504**, **508** und **510** durch ein PCI- oder durch ein PCI-X-Bussystem **500** miteinander verbunden sein und unter seiner Verwendung Daten und Befehle miteinander austauschen. Der Prozessor **53** kann Chips mit integrierten Schaltkreisen aufweisen, welche in einem Chipsatz mit integrierten Schaltkreisen enthalten sind, welcher von dem Inhaber der vorliegenden Anmeldung im Handel erhältlich ist, (z. B. der Chipsatz Intel® 80310), obwohl der E/A-Prozessor **53** ersatzweise stattdessen andere Chips mit integrierten Schaltkreisen (z. B. den E/A-Prozessor Intel® 80960 RM/RN oder den Prozessor Intel® 80321) oder ande-

re Typen von Prozessoren/integrierten Schaltkreisen umfassen kann, ohne von dieser Ausführungsform der vorliegenden Erfindung abzuweichen. Die Netzwerkschnittstelle **504** kann mit der Verbindung **38** gekoppelt sein, kann konfiguriert werden, die physikalischen Operationen zu ermöglichen/auszuführen, welche dem HBA **52** und dem Server **14** gestatten, mit dem Knoten **12** in dem Netzwerksystem **10** zu kommunizieren. Die Brückenschnittstelle **508** ist mit dem E/A-Brücken-/Bussystem **50** gekoppelt und ist konfiguriert, dem HBA **52** zu ermöglichen, Daten und Befehle mit dem System **48** auszutauschen. Der Speicher **502** kann RAM, welcher verwendet werden kann, um Programmprozesse **507** zu speichern (und in welchem diese sich befinden können), Hochgeschwindigkeits-Cache-Speicher, welcher verwendet werden kann, um Benutzerdaten zu speichern, welche zur Speicherung in dem Array **56** vorgesehen sind oder welche kürzlich daraus abgerufen wurden, und nichtflüchtigen EEPROM umfassen, welcher verwendet werden kann, um Daten, welche in dem Cache-Speicher gespeichert sind, in dem Fall zu speichern, dass ein Ausfall der externen Stromversorgung des HBA **52** auftritt. Obwohl nicht in den Figuren gezeigt, kann der HBA **52** ein System aufweisen, um zu erkennen, wenn ein derartiger Ausfall aufgetreten ist, und um den Prozessor **53** zu triggern, dass bewirkt wird, dass der Speicher **502** die Daten in dem EEPROM speichert, welche in dem Cache-Speicher gespeichert sind; ein batteriebetriebenes Notfallstromversorgungssystem (nicht gezeigt) kann in den HBA **52** eingeschlossen sein, um den HBA **52** während dieser Operationen mit Strom zu versorgen. Der Prozessor **53** kann programmiert werden, die Daten, welche in dem EEPROM gespeichert sind, bei einem nachfolgenden Einschaltvorgang des HBA **52** in den Cache-Speicher abzurufen. Zusätzliche Einzelheiten, welche die Konstruktion und den Betrieb des HBA **52** betreffen, werden nachfolgend dargestellt.

[0027] Obwohl in den Figuren nicht gezeigt, versteht es sich, dass die jeweilige Konstruktion des HBA **54** mit der entsprechenden Konstruktion des HBA **52** im Wesentlichen identisch sein kann. Um eine unnötige doppelte Beschreibung zu vermeiden, wird die Konstruktion des HBA **54** hier folglich nicht ausführlich dargestellt. Es genügt jedoch anzumerken, dass der HBA **54** über eine jeweilige herkömmliche S-ATA-basierte Schnittstelle (nicht gezeigt), welche in dem HBA **54** enthalten ist, elektrisch mit dem Array **58** gekoppelt sein kann und konfiguriert werden kann, Daten und Befehle mit ihm auszutauschen; durch Erteilen geeigneter Befehle an den Array **58** über diese ATA-basierte Schnittstelle kann der HBA **54** bewirken, dass Daten, auf welche mit derartigen Befehlen verwiesen wird oder welche in ihnen enthalten sind, in dem Array **58** gespeichert werden oder aus ihm abgerufen werden. Der HBA **54** kann einen jeweiligen E/A-Prozessor **55** umfassen, welcher eine jeweilige Konstruktion aufweisen kann, welche zu derjenigen

des E/A-Prozessors **53** im Wesentlichen gleich ist. Der HBA **54** kann auch eine jeweilige Netzwerkschnittstelle (nicht gezeigt) umfassen, welche eine jeweilige Konstruktion aufweisen kann, welche zu derjenigen der Schnittstelle **504** im Wesentlichen gleich ist, und über die Verbindung **38** an das Netzwerk **11** gekoppelt sein kann.

[0028] Das E/A-Brücken-/Bussystem **50** und die Host-Brücken-/Bussysteme **48** können ein E/A-Steuersystem bzw. ein Speichersteuersystem umfassen, welche unter Verwendung von Chips mit integrierten Schaltkreisen aufgebaut sein können, welche in einem Chipsatz mit integrierten Schaltkreisen enthalten sind, welcher von dem Inhaber der vorliegenden Anmeldung im Handel erhältlich ist (z. B. der Chipsatz Intel® 860). Die CPU **42**, der Systemspeicher **44** und die Benutzerschnittstelle **46** können über das Host-Brücken-/Bussystem **48** miteinander und mit dem E/A-Brücken-/Bussystem **50** elektrisch gekoppelt sein. Das System **48** ist konfiguriert, einen Austausch von Daten und Befehlen zwischen der CPU **42**, dem Speicher **44**, der Benutzerschnittstelle **46** und dem System **50** zu gestatten/zu ermöglichen. Zusätzlich ist das System **50** konfiguriert, einen Austausch von Daten und Befehlen zwischen dem System **48** und dem Adapter **52** und **54** zu gestatten/zu ermöglichen.

[0029] Der Speicher **44** kann RAM umfassen, welcher verwendet werden kann, Programmprozesse **45** zu speichern (und in welchem diese sich befinden können). Der RAM in dem Speicher **44** auch kann Benutzerdaten speichern, welche in den Arrays **56** und **58** gespeichert werden sollen oder welche aus ihm abgerufen wurden. Die CPU **42** kann einen Prozessor Intel® Xeon™ umfassen, obwohl die CPU **42** andere Typen von Prozessoren umfassen kann, ohne von der vorliegenden Erfindung abzuweichen.

[0030] In dem Server **12** können die Adapter **26** und **28** jeweils entsprechende Schaltplatinen umfassen, welche elektrisch und mechanisch über entsprechende elektromechanische Verbinder (nicht gezeigt) mit einer Hauptplatine (nicht gezeigt) gekoppelt sein können. Die Hauptplatine in dem Server **12** kann die CPU **16**, den Speicher **18** und die Systeme **22** und **24** umfassen. Wenn die Adapter **26** und **28** so mit der Hauptplatine in dem Server **12** gekoppelt sind, können die Adapter **26** und **28** mit dem System **24** auf der Hauptplatine in dem Server **12** gekoppelt sein. Separate elektromechanische Verbinder (nicht gezeigt) können in dem Server **12** bereitgestellt werden, um die Festplattenmassenspeichervorrichtungen in den Arrays **30** und **32** mit den Adaptern **26** bzw. **28** zu koppeln.

[0031] In dem Server **14** können die Adapter **52** und **54** jeweils entsprechende Schaltplatinen umfassen, welche elektrisch und mechanisch über entsprechen-

de elektromechanische Verbinder (nicht gezeigt) mit einer Hauptplatine (nicht gezeigt) gekoppelt sein können. Wenn die Adapter **52** und **54** mit der Hauptplatine in dem Server **14** gekoppelt sind, können die Adapter **52** und **54** mit dem System **50** auf der Hauptplatine in dem Server **14** gekoppelt sein. Separate elektromechanische Verbinder (nicht gezeigt) können in dem Server **14** bereitgestellt werden, um die Festplattenmassenspeichervorrichtungen in den Arrays **56** und **58** mit den Adaptern **52** bzw. **54** zu koppeln.

[0032] Obwohl in Netzwerk **10** die Adapter **26**, **28**, **52** und **54** an die Festplatten-Arrays **30**, **32**, **56** bzw. **58** angeschlossen sein können und mit ihnen über jeweilige S-ATA-basierte Schnittstellen bzw. Protokolle kommunizieren können, können die Adapter **26**, **28**, **52** und **58**, wenn die Adapter **26**, **28**, **52** und **54** und die Festplatten-Arrays passend modifiziert sind, stattdessen an die Festplatten-Arrays **30**, **32**, **56** und **58** angeschlossen sein und mit ihnen über andere Schnittstellentypen bzw. Protokolle kommunizieren (z. B. über FC-(Fibre Channel)- oder SCSI-(Small Computer Systems Interface)-basierte Schnittstellen und Protokolle), ohne von dieser Ausführungsform der vorliegenden Erfindung abzuweichen. Zusätzlich können die Festplatten-Arrays **30** und **32** in physikalischen Gehäusen eingeschlossen sein, welche physikalisch von dem physikalischen Gehäuse des Servers **12** getrennt sind, obwohl in den Figuren nicht gezeigt. Die Festplatten-Arrays **56** und **58** können in physikalischen Gehäusen eingeschlossen sein, welche physikalisch von dem physikalischen Gehäuse des Servers **14** getrennt sind. Weiterhin kann der Server **12** ersatzweise nur einen einzelnen HBA umfassen, wenn das Netzwerk **10** passend modifiziert wird, obwohl die Einbeziehung mehrerer Host-Bus-Adapter in den Server **12** Verbindungen zu zusätzlichen Ziel-Servern ermöglichen kann (z. B. zusätzlich zu dem Ziel-Server **14**) und eine Implementierung von Ausfallsicherungsmerkmalen in dem Netzwerk **10** ermöglichen kann, welche helfen können, einen einzelnen Ausfallpunkt in dem Netzwerk **10** in dem Fall eines Ausfalls eines einzelnen HBA in dem Server **12** zu beseitigen.

[0033] Unter Verweis jetzt auf die [Fig. 1](#) bis [Fig. 9](#) wird der Betrieb des Netzwerksystems **10** beschrieben. Obwohl in den Figuren nicht gezeigt, kann jede der Benutzerschnittstellen **20** und **46** eine jeweilige herkömmliche Tastatur, Zeigevorrichtung (z. B. eine Maus oder eine Rollkugelvorrichtung), Terminal/Anzeigeeinheit, BIOS (Basic Input/Output System) usw. umfassen. Unter Verwendung der Benutzerschnittstelle **20** kann ein menschlicher Benutzer (nicht gezeigt) des Servers **12** Befehle an die CPU **16** und die Prozesse **19** ausgeben, welche sich in dem Systemspeicher **18** befinden, und Ausgabedaten von ihnen empfangen, welche dem menschlichen Benutzer gestatten können, den Betrieb des Servers **12** zu steu-

ern und zu überwachen. Genauso kann ein menschlicher Benutzer (nicht gezeigt) des Servers **14** unter Verwendung der Benutzerschnittstelle **46** Befehle an die CPU **42** und die Prozesse **45** ausgeben, welche sich in dem Systemspeicher **44** befinden, und Ausgabedaten von ihnen empfangen, welche dem menschlichen Benutzer gestatten können, den Betrieb des Servers **14** zu steuern und zu überwachen.

[0034] Unter den Prozessen **19**, welche sich in dem Systemspeicher **18** in dem Ursprungs-Server **12** befinden können, können mehrere Betriebssystemprogrammprozesse **400** und mehrere Programmprozesse **402** der Anwendungsebene sein (man siehe [Fig. 5](#)). Unter den Prozessen **45**, welche sich in dem Systemspeicher **44** in dem Ziel-Server **14** befinden können, können entsprechend mehrere Betriebssystemprogrammprozesse **300** sein (man siehe [Fig. 4](#)).

[0035] Unter Verweis insbesondere auf die [Fig. 1](#), [Fig. 5](#) und [Fig. 7](#), kann der menschliche Benutzer des Servers **12** unter Verwendung der Schnittstelle **20** eine datenspeicherbezogene Operationsanforderung oder einen Befehl an einen Programmprozess ausgeben (z. B. an einen relationalen Datenbankprozess), welcher in den Programmprozessen **402** der Anwendungsebene enthalten ist. Dieser Befehl kann z. B. die Einrichtung einer replizierten/gespiegelten Datenträgerpaarbeziehung zwischen einem Datenträger (z. B. einem primären Datenträger **33**) in einem primären Array (z. B. dem Array **32**) in dem Ursprungs-Server **12** bzw. einem Datenträger (z. B. einem redundanten Datenträger **57**) in einem dazugehörigen redundanten Array (z. B. dem Array **56**) in dem Ziel-Server **14**, das Trennen einer derartigen Beziehung, die erneute Einrichtung einer derartigen Beziehung, nachdem ein derartiges Trennen aufgetreten ist, das Ändern (z. B. das Erweitern oder Verkleinern) der Größe der Datenträger **33** und **57**, welche in eine derartige Beziehung einbezogen sind, oder das Speichern von Benutzerdaten (z. B. der Benutzerdaten **404**), nachdem die Benutzerdaten **404** in dem Systemspeicher **18** gespeichert wurden, in dem primären Datenträger **33** in einem primären Array **32** und die Replikation dieser Daten **404** in dem redundanten Datenträger **57** in dem redundanten Array **56** anfordern. Die Benutzerdaten **404** können als Ergebnis der Operation eines oder mehrerer Programmprozesse **402** der Anwendungsebene erzeugt werden, oder sie können von einem Client-Knoten (nicht gezeigt) in dem Netzwerk **10** (z. B. über eine nicht gezeigte NIC, welche in dem Server **12** enthalten ist) erzeugt werden und von ihm an den Server **12** weitergeleitet werden. Jeder der Datenträger **33** und **57** kann jeweils eine identische Größe aufweisen (d. h. in Bezug auf den Datenspeicherungsraum) und kann ein jeweiliger logischer Datenträger sein, welcher sich über mehr als eine jeweilige physikalische Festplattenmassenspeichervorrichtung erstrecken kann und/oder weniger als den gesamten Datenspei-

cherungsraum verbrauchen kann, welcher in einer jeweiligen einzelnen physikalischen Festplattenmassenspeichervorrichtung enthalten ist.

[0036] Als Reaktion auf die datenspeicherbezogene Operationsanforderung oder den Befehl, welche von dem menschlichen Benutzer des Servers **12** erteilt wurde, kann der Programmprozess, an welchem die Anforderung/der Befehl erteilt wurde, bewirken, dass die CPU **16** eine datenspeicherbezogene Operationsanforderung erzeugt und über das System **22** einem der O/S-Prozesse **400** in dem Speicher **18** bereitstellt, welche der Anforderung/dem Befehl entspricht, welche/r von dem Benutzer des Servers **12** erteilt wurde (d. h. welche/r die Ausführung der gleichen datenspeicherbezogenen Operation anfordert, welche von dem menschlichen Benutzer angefordert wurde; man siehe Operation **602** in [Fig. 7](#)). Als Reaktion auf die Anforderung, welche in der Operation **602** erzeugt wurde, kann der O/S-Prozess, welcher die Anforderung empfängt, welche in der Operation **602** erzeugt wurde, dann bewirken, dass die CPU **16** eine datenspeicherbezogene Operationsanforderung **60** erzeugt und über die Systeme **22**, **24** und **200** dem Prozessor **27** in dem HBA **28** bereitstellt, welche der Anforderung/dem Befehl entspricht, welche/r von dem Benutzer erteilt wurde (man siehe Operation **604** in [Fig. 7](#)). Ersatzweise kann ein Hintergrundprozess (nicht gezeigt), welcher in den Prozessen **19** enthalten ist, bewirken, dass die CPU **16** die Anforderung **60** erzeugt und dem Prozessor **27** bereitstellt. In jedem Fall kann die CPU **16** als Teil der Operation **604** auch die Daten **404** aus dem Speicher **18** abrufen und über die Systeme **22**, **24** und **200** an den Prozessor **27** weiterleiten, wenn die Anforderung/der Befehl, welche/r von dem Benutzer erteilt wurde, die Speicherung von Daten **404** in den Datenträgern **33** und **57** anfordert, und die Anforderung **60** kann die Daten **404** enthalten oder auf sie verweisen, welche an den Prozessor **27** weitergeleitet werden; der Prozessor **27** kann bewirken, dass die weitergeleiteten Daten **404** anfänglich in dem Speicher **202** gespeichert werden.

[0037] Als Reaktion auf die Anforderung **60** kann der Prozessor **27** die Ausführung der datenspeicherbezogenen Operation bewirken, welche in der Anforderung **60** angefordert wird (man siehe Operation **606** in [Fig. 7](#)). Dies bedeutet, dass die Ausführung der datenspeicherbezogenen Operation, welche in der Anforderung **60** angefordert wird, die Ausführung von entsprechenden Sätzen datenspeicherbezogener Operationen in dem Ursprungs-Server **12** und in dem Ziel-Server **14** erfordern kann; zusätzlich können unter den Prozessen **207**, welche sich in dem Speicher **202** in dem HBA **28** befinden können, RAID-(Redundant Array of Independent Disks)-Treiberprozesse **700** sein. Als Reaktion auf den Empfang der Anforderung **60** durch den Prozessor **27** können die Treiberprozesse **700** bewirken, dass der Prozes-

sor **27** in dem Ursprungs-Server **12** arbeitet und die Ausführung der entsprechenden Sätze datenspeicherbezogener Operationen in dem Ziel-Server **14** einleitet, welche in dem Ursprungs-Server **12** und in dem Ziel-Server **14** ausgeführt werden müssen, um die datenspeicherbezogene Operation auszuführen, welche in der Anforderung **60** angefordert wurde.

[0038] In Abhängigkeit von dem Typ der datenspeicherbezogenen Operation, welche in der Anforderung **60** angefordert wird, um die Ausführung des jeweiligen Satzes datenspeicherbezogener Operationen in dem Ursprungs-Server **12** zu bewirken, welche darin ausgeführt werden müssen, um die datenspeicherbezogene Operation auszuführen, welche in der Anforderung **60** angefordert wurde, kann der Prozessor **27** einen oder mehrere Einträge in den Datenspeicherverwaltungs-/konfigurationstabellen **209** aktualisieren/modifizieren und/oder einen oder mehrere Befehle (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **64** bezeichnete Struktur verwiesen wird) an den primären Array **32** bereitstellen. Dies bedeutet, dass der Prozessor **27** die Tabellen **209** in dem Speicher **202** unterhalten kann, und die Informationen, welche in den Tabellen **209** enthalten sind, die physikalischen Eigenschaften/Konfigurationen der individuellen Festplattenmassenspeichervorrichtungen in den Arrays **32** und **56** auf die logischen Eigenschaften/Konfigurationen/Beziehungen abbilden können, welche für derartige Vorrichtungen erwünscht und/oder eingeführt sind. Beispielsweise können die Informationen in den Tabellen **209** Bereiche in den physikalischen Massenspeichervorrichtungen in den Arrays **32** und **33** mit logischen Datenträgern, mit den Typen der Datencodierung/-formatierung, welche auf den physikalischen und logischen Datenträgern verwendet werden, mit den jeweiligen Größen (in Bezug auf den Datenspeicherungsraum) der logischen und physikalischen Datenträger/vorrichtungen in den Arrays **32** und **56**, mit den jeweiligen Verknüpfungen korrelieren, welche zwischen den jeweiligen primären Datenträgern in Array **32** und den redundanten Datenträgern in Array **56** eingerichtet wurden (d. h. welche primären Datenträger in Array **32** von welchen redundanten Datenträgern in Array **56** gespiegelt werden, die Verknüpfung der Arrays **32** und **56** als ein Paar primäres Array-redundantes Array usw.). Wenn die Anforderung **60** die Ausführung einer datenspeicherbezogenen Operation anfordert, welche kein Speichern von Benutzerdaten (z. B. der Benutzerdaten **404**) in den Arrays **32** und **56** einbezieht (z. B. wenn die Anforderung **60** die Einrichtung einer replizierten/gespiegelten Datenträgerpaarbeziehung zwischen dem Datenträger **33** und dem Datenträger **57**, das Trennen einer derartigen Beziehung, die erneute Einrichtung einer derartigen Beziehung, nachdem ein derartiges Trennen aufgetreten ist, oder das Ändern (z. B. das Erweitern oder Verkleinern) der Größe der Datenträger **33** und **57** anfordert, welche bei einer der-

artigen Beziehung einbezogen sind), dann können die Prozesse **700** bewirken, dass der Prozessor **27** verifiziert, dass die Operation, welche in der Anforderung **60** angefordert wird, unter den gegebenen physikalischen und/oder voreingerichteten logischen Eigenschaften/Beziehungen möglich/zugelassen ist, welche in den Tabellen **209** angegeben sind; wenn die angeforderte Operation möglich/zugelassen ist, können die Prozesse **700** bewirken, dass der Prozessor **27** einen oder mehrere geeignete Einträge in den Tabellen **209** aktualisiert/modifiziert, um so zu bewirken, dass die Tabellen **209** die Veränderung logischer Eigenschaften/Beziehungen angeben/wiedergeben, welche von der Anforderung **60** angefordert wurde. Wenn umgekehrt die angeforderte Operation nicht möglich/zugelassen ist, können die Prozesse **700** bewirken, dass der Prozessor **27** sowohl eine Fehlermeldung über das System **22** und **24** an die CPU **16** ausgibt, welche dies angibt, als auch die datenspeicherbezogene Operation beendet. Wenn auch umgekehrt die Anforderung **60** die Speicherung der Benutzerdaten (z. B. der Benutzerdaten **404**) in den Arrays **32** und **56** (z. B. auf den logischen Datenträgern **33** und **57**) anfordert, können die Prozesse **700** bewirken, dass der Prozessor **27** einen oder mehrere Befehle (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **64** bezeichnete Struktur verwiesen wird) an den Array **32** (über den Bus **200** und die Schnittstelle **210**) erzeugt und bereitstellt, um zu bewirken, dass die Benutzerdaten **404** auf den geeigneten physikalischen Ort oder auf Orte in den Festplattenmassenspeichervorrichtungen in dem Array **32** geschrieben werden, welcher den logischen Datenträger **33** umfasst.

[0039] Um die Ausführung des jeweiligen Satzes datenspeicherbezogener Operationen in dem Ziel-Server **14** einzuleiten, welche darin ausgeführt werden müssen, um die datenspeicherbezogene Operation auszuführen, welche in der Anforderung **60** angefordert wird, können die Prozesse **700** bewirken, dass der Prozessor **27** (unter Verwendung von Operationen, welche mehrere Transaktionen/Datenübertragungen zwischen dem Prozessor **27** und der Speicher **202** einbeziehen) ein oder mehrere datenspeicherbezogene Operationsanforderungs-Frames (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **62** bezeichnete Struktur verwiesen wird; man siehe auch Operation **608** in [Fig. 7](#)) einbezieht und an den Prozessor **53** weiterleitet (über den Bus **200**, die Netzwerkschnittstelle **204**, die Verbindung **34**, das Netzwerk **11**, die Verbindung **38**, die Netzwerkschnittstelle **504** und den Bus **500**). Dies bedeutet, dass die Prozesse **700** Netzwerk-Kommunikationsstapelprozesse aufweisen können, welche über mehrere Operationen Datenübertragungen zwischen dem Prozessor **27** und dem Speicher **202** erzeugen, und wenn die datenspeicherbezogene Operation, welche von der Anforderung **60** angefordert wird, die Speicherung der Daten **404** in den Arrays **32**

und **56** einbezieht, dann können die Prozesse **700** bewirken, dass der Prozessor **27** die Daten **404** zu einer Sequenz aus einer oder mehreren Datennutzlasten fragmentiert, welche für eine Einkapselung in einer jeweiligen Sequenz von einem oder mehreren Frames **62** geeignet sind, und können bewirken, dass der Prozessor **27** die Sequenz der Frames **62** erzeugt und an den Prozessor **53** in dem HBA **52** überträgt. Wenn umgekehrt die datenspeicherbezogene Operation, welche von der Anforderung **60** angefordert wird, keine Speicherung der Daten **404** in den Arrays **32** und **56** einbezieht, dann muss vielleicht nur ein einzelner Frame **62** erzeugt werden und von dem HBA **28** an den HBA **56** übertragen werden. Jeder Frame **62**, welcher von dem HBA **28** an den HBA **52** übertragen wird, kann die Datenstruktur **100** aufweisen, welche in [Fig. 2](#) gezeigt wird.

[0040] Die jeweilige Datenstruktur **100** jedes Frames **62** kann jeweilige Felder **102**, **104**, **106**, **108**, **110**, **112**, **114** und **116** aufweisen; jedes dieser jeweiligen Felder **102**, **104**, **106**, **108**, **110**, **112**, **114** und **116** kann einen oder mehrere jeweilige Datenwerte enthalten. Das Feld **102** ist ein jeweiliges Kopffeld eines Ethernet-Frames, und das Feld **104** ist ein jeweiliges Fußfeld eines Ethernet-Frames. Jedes dieser Felder **102** und **104** kann zusätzlich jeweilige Felder (z. B. die jeweiligen Ethernet- und IP-Quell- und Bestimmungsadressfelder, Ethernet- und IP-Prüfsummenfelder usw.) enthalten, welche verwendet werden können, um die verbleibenden Felder **106**, **108**, **110**, **112**, **114** und **116** in der jeweiligen Frame-Struktur **100** als einen jeweiligen Ethernet-Frame zu verkapseln, welcher ein TCP/IP-Paket enthält. Die jeweiligen Werte, welche in den jeweiligen Feldern enthalten sein können, welche in dem Kopffeld **102** und in dem Fußfeld **104** enthalten sind, können ausgewählt werden, um dem HBA (z. B. dem HBA **28**) zu gestatten, dass er die jeweilige Frame-Struktur **100** so erzeugt, dass sie in der Lage ist, die jeweilige Frame-Struktur **100** an den HBA (z. B. den HBA **52**) gültig zu übertragen, welcher vorgesehen ist, die jeweilige Frame-Struktur **100** über das Netzwerk **11** zu empfangen. Durch den jeweiligen Frame-Kopf **102** und Frame-Fuß **104** kann ein jeweiliges Zielknoten-iSCSI-Protokolladressfeld **106**, ein jeweiliges Startknoten-iSCSI-Protokolladressfeld **108**, ein jeweiliges Feld **110** für einen datenspeicherbezogenen Operationsbefehl oder eine Operationsmeldung, ein jeweiliges Datennutzlastfeld **112**, ein jeweiliges Meldungskennungsfeld **114** und ein jeweiliges Prüfsummenfeld **116** verkapselt werden. Das jeweilige Befehls-/Meldungsfeld **110** in jedem jeweiligen Frame **62** kann einen Wert enthalten, welcher den Typ der datenspeicherbezogenen Operation identifiziert, welche von der Anforderung **60** befohlen wurde, und kann (z. B. in zusätzlichen Subfeldern in dem Feld **110**) Werte angeben, welche zusätzliche Parameter identifizieren können, welche bei einer Implementierung der Befehlsoperation verwendet werden können

(z. B. bestimmte primäre und redundante Datenträger, welche einbezogen sind, die Größen von derartigen Datenträgern, die jeweilige Adresse in dem logischen Zieldatenträger **57**, bei welchem das Schreiben von Daten auf den logischen Zieldatenträger **57** anfangen soll, die Anzahl von Bytes der Datennutzlast, welche in dem Feld **112** vorhanden ist, usw.); die jeweiligen Felder **106** und **108** in jedem jeweiligen Frame **62** können jeweilige Wertesätze enthalten, welche Ziel- und Startknoten enthalten, welche Informationen adressieren und benennen, welche gemäß eines herkömmlichen iSCSI-Adressierungsschemas bestimmt werden (z. B. gemäß dem Adressierungsschema, welches in Satran, „iSCSI“, Internet-Draft-Specification, draft-ietf-ipsiscsi-05, IP Storage Working Group of the Internet Engineering Task Force, herausgegeben am 1. März 2001 von der Internet Engineering Task Force, Internet Engineering Task Force Secretariat c/o Corporation for National Research Initiatives, 1895 Preston White Drive, Suite 100, Reston, VA 20191, Vereinigte Staaten von Amerika, offenbart ist), welches die bestimmten Speichervorrichtungen in den Arrays **32** und **56** abbilden/adressieren kann, welche in die datenspeicherbezogene Operation einbezogen sind. Natürlich können verschiedene Adressierungsinformationen (z. B. Adressierungsinformationen aufgrund verschiedener Adressierungsschemata, wie beispielsweise anderer iSCSI-Adressierungsschemata gemäß anderer/später entwickelter Versionen der oben stehend genannten „iSCSI“ Internet-Draft Specification der IP-Storage Working Group der Internet Engineering Task Force) in den Feldern **106** und **108** enthalten sein, solange das Netzwerk **11** und die Knoten **12** und **14** geeignet konfiguriert werden, um Frames **100** aufgrund derartiger verschiedener Adressierungsinformationen zu leiten bzw. zu verarbeiten. Wenn die datenspeicherbezogene Operation, welche von der Anforderung **60** befohlen wurde, die Speicherung von Daten **404** in den Arrays **32** und **56** einbezieht, dann kann das Nutzlastfeld **112** eine jeweilige Datennutzlast aus der Sequenz der Datennutzlasten enthalten, welche von dem Prozessor **27** aus den Daten **404** in der oben stehend beschriebenen Weise erzeugt wurden, und eine jeweilige TCP-Sequenznummer, welche in dem Kopf **102** enthalten ist, kann die Sequenzreihenfolge dieses Datennutzlastfragments in der Sequenz von Datennutzlastfragmenten angeben, welche von dem Prozessor **27** erzeugt wurden. Das Meldungs-ID-Feld **114** kann eine Frame-Sequenznummer enthalten, welche den bestimmten Frame **62** identifizieren kann, welcher diese Frame-Sequenznummer enthält. Das Prüfsummenfeld **116** kann einen Prüfsummenwert enthalten, welcher die Prüfsumme eines Datenwerts darstellt, welcher durch die Verkettung der Werte gebildet wird, welche in den Feldern **106**, **108**, **110**, **112**, **114** und **116** enthalten sind. Ersatzweise kann in Abhängigkeit von dem bestimmten Protokoll oder den Protokollen, welche verwendet werden, um den Frame **100** in und durch das

Netzwerk **11** weiterzugeben und/oder zu leiten, und auch in Abhängigkeit von der bestimmten Verarbeitung, welcher der Frame **100** in den Knoten **12** und **14** erfahren kann, der Frame **100** Felder und/oder andere oder zusätzliche Informationen enthalten als diejenigen, welche in den Feldern enthalten sind, welche in [Fig. 2](#) gezeigt sind, um ein derartiges Weitergeben, Leiten und Verarbeiten zu ermöglichen, ohne von dieser Ausführungsform der vorliegenden Erfindung abzuweichen.

[0041] Unter den Prozessen **507**, welche sich in dem Speicher **502** befinden können, können Replikationstreiberprozesse **800** sein. Diese Prozesse **800** können Netzwerkkommunikations-Stapelprozesse umfassen. Wenn ein Frame **62** von dem Prozessor **53** empfangen wird, können die Prozesse **800** bewirken, dass der Prozessor **53** den empfangenen Frame **62** anfänglich in dem Speicher **502** speichert, und den Frame **62** danach entkapselt und garst, verifiziert (d. h. durch Verifizierung der Prüfsummen, welche in dem Frame **62** enthalten sind), dass während der Übertragung des Frames **62** an den Prozessor **53** keine Fehler aufgetreten sind, und bestimmt u. a. die bestimmten jeweiligen Werte, welche in den jeweiligen Feldern **102**, **104**, **106**, **108**, **110**, **112**, **114**, **116** enthalten sein können, und die jeweiligen Subfelder, welche in diesen jeweiligen Feldern in dem Frame **62** enthalten sind. Aufgrund der Werte, welche in diesen jeweiligen Feldern und den jeweiligen Subfeldern enthalten sind, kann der Prozessor **53** in dem Server **14** die bestimmte datenspeicherbezogene Operation bestimmen und durchführen, welche in den empfangenen Frames **62** angefordert werden (man siehe Operation **610** in [Fig. 7](#)). Um in Abhängigkeit von dem Typ der datenspeicherbezogenen Operation, welche in einem empfangenen Frame **62** befohlen wird, insbesondere die Ausführung der jeweiligen datenspeicherbezogenen Operation in dem Server **14** zu bewirken, welche in dem Frame **62** angefordert wird, kann der Prozessor **53** einen oder mehrere Einträge in den Datenspeicherverwaltungs-/konfigurationstabellen **509** aktualisieren/modifizieren und/oder einen oder mehrere Befehle (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **70** bezeichnete Struktur verwiesen wird) an das redundante Array **56** bereitstellen. Dies bedeutet, dass der Prozessor **53** die Tabellen **509** in dem Speicher **502** unterhalten kann und dass die Informationen, welche in den Tabellen **509** enthalten sind, inhaltlich gleich den Informationen sein können, welche in den Tabellen **209** enthalten sind, welche von dem Prozessor **27** in dem Speicher **202** unterhalten werden. Falls die datenspeicherbezogene Operation, welche in dem empfangenen Frame **62** befohlen wird, die Ausführung einer datenspeicherbezogenen Operation anfordert, welche kein Speichern von Benutzerdaten (z. B. Benutzerdaten **404**) in dem Array **56** einbezieht (z. B. falls die Anforderung **60** die Einrichtung einer replizierten/gespiegelten Datenträgerpaarbeziehung

zwischen dem Datenträger **33** und dem Datenträger **57**, die Trennung einer derartigen Beziehung, die erneute Einrichtung einer derartigen Beziehung, nachdem eine derartige Trennung aufgetreten ist, oder das Ändern (z. B. das Erweitern oder Verkleinern) der Größe der Datenträger **33** und **57** anfordert, welche in eine derartige Beziehung einbezogen sind), dann können die Prozesse **800** bewirken, dass der Prozessor **53** verifiziert, dass die Operation, welche in dem Frame **62** befohlen wird, (d. h. der Befehl, welcher in dem Feld **110** in dem empfangenen Frame **62** angegeben ist) unter den gegebenen physikalischen und/oder voreingerichteten logischen Eigenschaften/Beziehungen möglich/zugelassen ist, welche in den Tabellen **509** angegeben sind; wenn die befohlene Operation möglich/zugelassen ist, können die Prozesse **800** bewirken, dass der Prozessor **53** einen oder mehrere geeignete Einträge in den Tabellen **509** aktualisiert/modifiziert, um so zu bewirken, dass die Tabellen **509** die Veränderung der logischen Eigenschaften/Beziehungen angeben/wiedergeben, welche in dem Frame **62** befohlen wurden. Wenn umgekehrt der Frame **62** die Speicherung von Benutzerdaten (z. B. der Benutzerdaten **404**) in dem Array **56** anfordert (d. h. in dem Datenträger **57** in Array **56**), können die Prozesse **800** bewirken, dass der Prozessor **53** bestimmt (d. h. aus den jeweiligen TCP-Informationen, welche in dem Frame **62** enthalten sind), ob irgendwelche andere Frames **62** vielleicht zusätzliche Fragmente der Benutzerdaten **404** enthalten. Wenn der Prozessor **53** bestimmt, dass keine anderen Frames **62** derartige zusätzliche Fragmente der Benutzerdaten **404** enthalten, können die Prozesse **800** bewirken, dass der Prozessor **53** einen oder mehrere Befehle (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **70** bezeichnete Struktur verwiesen wird) an den Array **56** (über den Bus **500** und die Schnittstelle **510**) erzeugt und bereitstellt, um zu bewirken, dass die Benutzerdaten **404** auf den geeigneten physikalischen Ort oder auf geeignete Orte auf den Festplattenmassenspeichervorrichtungen in dem Array **56** geschrieben werden, welcher den logischen Datenträger **57** umfasst. Wenn der Prozessor **53** jedoch bestimmt, dass es andere Frames **62** gibt, welche zusätzliche Fragmente der Benutzerdaten **404** enthalten können, können die Prozesse **800** bewirken, dass der Prozessor **53** die jeweiligen Datennutzlasten aus derartigen anderen Frames **62** in dem Speicher **502** speichert und aufgrund der jeweiligen TCP-Sequenzinformationen und/oder Meldungskennungswerte, welche in den empfangenen Frames **62** enthalten sind, die Benutzerdaten **404** in dem Speicher **502** wiederherstellt und speichert. Danach können die Prozesse **800** bewirken, dass der Prozessor **53** die rekonstruierten Benutzerdaten **404** zu einem oder mehreren Datenblöcken fragmentiert und einen oder mehrere Befehle **70** an den Array **56** erzeugt und bereitstellt (über den Bus **500** und die Schnittstelle **510**), um zu bewirken, dass die Blöcke der rekonstruierten Benutzerdaten auf den geeigneten physi-

kalischen Ort oder auf geeignete Orte (z. B. Sektoren, Blöcke usw.) in den Festplattenmassenspeichervorrichtungen in dem Array **56** geschrieben werden, welcher den logischen Datenträger **57** umfasst.

[0042] Nachdem jeder jeweilige Frame **62** sachgerecht (d. h. ohne einen Übertragungsfehler) von dem Prozessor **53** empfangen wurde, können die Prozesse **800** bewirken, dass der Prozessor **53** einen jeweiligen Bestätigungs- oder Antwort-Ethernet-Frame an den Prozessor **27** erzeugt und bereitstellt (auf welche zusammen oder allein durch die in [Fig. 1](#) mit Bezugszeichen **66** bezeichnete Struktur verwiesen wird; man siehe auch Operation **612** in [Fig. 7](#)). Jeder Antwort-Frame **66** kann eine jeweilige Struktur aufweisen, welche zu der Struktur **100**, welche in [Fig. 2](#) gezeigt wird, identisch sein kann, außer dass das jeweilige Befehls-/Meldungsfeld **110** in dem jeweiligen Antwort-Frame **66** einen jeweiligen Wert enthalten kann, welcher anzeigen kann, dass ein Frame **62**, welcher die Meldungskennungsnummer umfasst, welche in dem Feld **116** in dem jeweiligen Antwort-Frame **66** spezifiziert ist, von dem Prozessor **53** empfangen und verarbeitet wurde. Dies bedeutet, dass für jeden jeweiligen Frame **62**, welcher von dem Prozessor **53** sachgerecht empfangen und verarbeitet wird, die Prozesse **800** bewirken können, dass der Prozessor **53** einen jeweiligen Antwort-Frame **66** erzeugt, welcher auf die jeweilige Meldungskennungsnummer verweist (d. h. in dem Meldungskennungsfeld **114** des Antwort-Frames **66** enthält), welche in dem jeweiligen Frame **62** enthalten war, und an den Prozessor **27** ausgibt (über den Bus **500**, die Schnittstelle **504**, die Verbindung **38**, das Netzwerk **11**, die Verbindung **34**, die Schnittstelle **204** und den Bus **200**). In den Tabellen **209** kann der Prozessor **27** die jeweiligen Meldungskennungsnummern, welche in den jeweiligen Frames **62** enthalten sind, mit den jeweiligen Meldungskennungsnummern korrelieren, welche in den jeweiligen empfangenen Antwort-Frames **66** enthalten sind, um zu bestimmen welche der Meldungs-Frames **62**, welche von dem Prozessor **27** an den Prozessor **53** weitergeleitet wurden, empfangen und verarbeitet wurden (z. B. um unter Verwendung herkömmlicher Verfahren zu bestimmen, ob ein jeweiliger Meldungs-Frame **62** bei der Übertragung durch das Netzwerk **11** verloren gegangen ist und ob dieser jeweilige Meldungs-Frame **62** erneut an den Prozessor **53** gesendet werden muss). Nachdem der Prozessor **27** alle Frames **66** empfangen hat, welche von dem Prozessor **27** als Reaktion auf eine gegebene datenspeicherbezogene Operationsanforderung **60** von der CPU **16** erzeugt wurden, können die Prozesse **700** bewirken, dass der Prozessor **27** bestimmt (unter der Annahme, dass der Prozessor **27** auch in der Lage war, die Operationen erfolgreich abzuschließen, welche der Prozessor **27** als Reaktion auf die Anforderung **60** ausgeführt hat), dass die datenspeicherbezogene Operation, welche von der Anforderung **60** angefordert wurde, erfolgreich in dem

Netzwerk **10** abgeschlossen wurde, und sie können bewirken, dass der Prozessor **27** eine Meldung **72** über die Systeme **24** und **22** an die CPU **16** weiterleitet, um den O/S-Prozessen **400** den erfolgreichen Abschluss der Operation anzugeben, welche von der Anforderung **60** angefordert wurde (man siehe Operation **614** in [Fig. 7](#)). Als Reaktion auf die Meldung **72** können die O/S-Prozesse **400** bewirken, dass die CPU **16** den Anwendungsprozessen **402** in dem Speicher **18** eine Angabe eines derartigen erfolgreichen Abschlusses bereitstellt; als Reaktion darauf können die Prozesse **402** bewirken, dass die CPU **16** dem menschlichen Benutzer des Servers **12** einen derartigen erfolgreichen Abschluss über die Schnittstelle **20** angibt.

[0043] In dem Fall, dass der Prozessor **53** bestimmt (z. B. aus den Tabellen **509**) oder andernfalls (z. B. in dem Fall eines Prüfsummenvalidierungsfehlers beim Verarbeiten eines Anforderungs-Frames **62**), dass er unfähig ist, eine datenspeicherbezogene Operation durchzuführen, welche in einem Frame **62** angefordert wird (z. B. in dem Fall eines detektierten Defekts in dem HBA **56** oder in einem Plattenlaufwerk oder in den Plattenlaufwerkskabeln in dem Array **56**, einer Nichtübereinstimmung der Größen der Datenträger **33** und **57**, dem Mangel an ausreichendem Speicherplatz auf dem Datenträger **57**, von Adressierungsfehlern in dem Anforderungs-Frame **62** usw.), können die Prozesse **800** bewirken, dass der Prozessor **53** einen Antwort-Frame **66** mit einem jeweiligen Wert in dem Befehls-/Meldungsfeld **110**, welcher angibt, dass ein Fehler beim Durchführen der angeforderten Operation aufgetreten ist, sowie den Typ und die Natur des Fehlers angibt, erzeugt und an den Prozessor **27** weiterleitet. Als Reaktion auf diese Meldung **66** können die Prozesse **700** bewirken, dass der Prozessor **27** eine Fehlermeldung (vom oben stehend beschriebenen Typ) und einen Fehlerbestätigungs-Frame (nicht gezeigt) an den Prozessor **53** erzeugt und an die CPU **16** ausgibt, welcher identisch zu demjenigen eines jeweiligen Antwort-Frames **66** ist, außer dass in dem Fehlerbestätigungs-Frame das jeweilige Befehls-/Meldungsfeld **110** den Empfang und die Verarbeitung der Meldung **66** durch den Prozessor **27** angeben kann.

[0044] Wenn der Prozessor **27** nach einer vorbestimmten Zeitspanne keinen gültigen Bestätigungs-/Antwort-/Fehler-Frame **66** als Reaktion auf einen Anforderungs-Frame **62** empfängt, können die Prozesse **700** bewirken, dass der Prozessor **27** erneut eine Kopie dieses Anforderungs-Frames **62** erzeugt und erneut an den Prozessor **53** sendet und erneut die vorbestimmte Zeitspanne wartet, um einen gültigen Antwort-Frame **66** als Antwort darauf zu empfangen; wenn der Prozessor **27** nach einer vorbestimmten Anzahl wiederholter Erzeugungen und wiederholter Übertragungen des Anforderungs-Frames **62** noch immer keinen gültigen Antwort-Frame

66 darauf empfangen hat oder wenn die Natur des Fehlers, welcher in einem empfangenen Fehler-Frame **66** berichtet wird, von einem vorbestimmten Typ ist, welcher nicht ohne Eingriff eines menschlichen Bedieners korrigiert werden kann, können die Prozesse **700** bewirken, dass diese Bedingung an die CPU **16** berichtet wird, welche diese dann über die Schnittstelle **20** an den menschlichen Benutzer des Servers **12** berichten kann. Wenn der Server **12** für eine derartige Operation passend konfiguriert ist, kann die CPU **16** ersatzweise oder zusätzlich dazu, eine Warnmeldung erzeugen (z. B. eine E-Mail oder eine Meldung für einen Funkrufempfänger) und an einen menschlichen Bediener übertragen, welche eine derartige Bedingung angibt. Die Prozesse **700** können dann bedenken, dass die Zuordnung zwischen den Datenträgern **33** und **57** (d. h. als ein primärer Datenträger bzw. ein dazugehöriger Zieldatenträger) terminiert wurde, und können bewirken, dass der Prozessor **27** veränderte Spuren/Sektoren, welche auf den Datenträgern **33** auftreten können, in der nachfolgend beschriebenen Weise zur späteren Verwendung aufzeichnet, indem er die gespiegelte Beziehung zwischen den Datenträgern **33** und **57** erneut einrichtet.

[0045] Obwohl es nicht ausdrücklich in den Figuren gezeigt wird, ist es wichtig zu beachten, dass die Festplatten-Arrays **30** und **58** zueinander als ein redundantes bzw. als ein primäres Array zugeordnet sein können, mindestens ein jeweiliger Datenträger **31** in dem redundanten Array **59** ein redundanter Datenträger sein kann, welcher mindestens einem jeweiligen primären Datenträger **59** in dem primären Array **58** zugeordnet ist, und der HBA **26**, der Array **30**, die Verbindung **36**, das Netzwerk **11**, die Verbindung **40**, der HBA **54**, der Array **58**, die Systeme **48** und **50**, die CPU **42**, der Speicher **44** und die Schnittstelle **46** konfiguriert werden können, den Festplatten-Arrays **30** und **48** und den Datenträgern **31** und **30** zu gestatten und zu ermöglichen, in derartigen Beziehungen gemäß den Lehren dieser Ausführungsform der vorliegenden Erfindung zu arbeiten, welche den primären und redundanten Arrays **32** und **56** und den primären und redundanten Datenträgern **33** und **57** gestatten, gemäß dieser Ausführungsform der vorliegenden Erfindung in der oben stehend dargelegten Weise zu arbeiten. Zusätzlich versteht es sich, dass die Server **12** und **14** NIC-Vorrichtungen (nicht gezeigt) aufweisen können, welche über das Netzwerk **11** an einen oder mehrere Client-Computerknoten in einer derartigen Weise angeschlossen sein können, dass den Client-Computerknoten gestattet wird, dass sie auf die Benutzerdaten zuzugreifen können, welche in den primären Arrays **32** und **58** gespeichert sind.

[0046] Um zusätzlich eine erneute Einrichtung der Zuordnung zwischen den primären und redundanten Datenträgern **33** und **57** (d. h. als ein gespiegeltes

Paar von Datenträgern) nach einer vorhergehenden Terminierung einer derartigen Zuordnung zu ermöglichen, können die Tabellen **209** Listen logischer Spuren/Blöcke auf den Datenträgern **33** enthalten, welche sich seit der vorhergehenden Terminierung einer derartigen Zuordnung verändert haben. Nach der erneuten Einrichtung einer derartigen Zuordnung können die Prozesse **700** bewirken, dass Frames **62** von dem Prozessor **27** an den Prozessor **53** übertragen werden, welche bewirken können, dass entsprechende logische Spuren/Blöcke auf dem Datenträger **57** mit den Daten überschrieben werden, welche gegenwärtig in den Spuren/Blöcken gespeichert sind, welche in den Tabellen **209** als nach der Terminierung einer derartigen Zuordnung und vor ihrer erneuten Einrichtung verändert aufgeführt sind. Während der Zeitspanne nach der Terminierung und vor der erneuten Einrichtung einer derartigen Zuordnung kann der Datenträger **57** kopiert/gesichert werden (z. B. auf ein Band oder auf einen anderen, nicht gezeigten Datenträger).

[0047] Es können auch zusätzliche Typen datenspeicherbezogener Operationen von den Frames **62** befohlen werden. Wenn beispielsweise ein geeigneter vorbestimmter Wert in das Feld **110** in einem jeweiligen Frame **62** eingefügt wird, kann der jeweilige Frame **62** einem Ein-/Ausgabeprozessor (z. B. dem Prozessor **53**), welcher den Frame **62** empfängt, befehlen, einen jeweiligen Antwort-Frame **66** zu erzeugen und zu liefern, welcher dem Ein-/Ausgabeprozessor, welcher der Frame **62** übertragen hat, alle logischen Knoten, logische Datenträger und Zieldatenträger angeben/berichten kann, welche dem Ein-/Ausgabeprozessor, welcher den Frame **62** empfängt, zugeordnet sind oder von ihm gesteuert werden, ob ein gegebener Zieldatenträger, welcher dem Ein-/Ausgabeprozessor zugeordnet ist, welcher den Frame **62** empfängt, voll ist oder fehlerhaft funktioniert, ob laufende Datenreplikationsoperationen, welche einen derartigen Zieldatenträger einbeziehen, normal fortschreiten, oder ob Dateisysteme in einem derartigen Zieldatenträger normal strukturiert sind. Auch beim Einfügen eines geeigneten vorbestimmten Werts in das Feld **110** in einem jeweiligen Frame **62** kann der jeweilige Frame **62** einem Ein-/Ausgabeprozessor (z. B. dem Prozessor **53**), welcher den Frame **62** empfängt, befehlen, eine Norm-Frame-Größe für die Frame-Struktur **100** zu konfigurieren oder Verwaltungstabellen (z. B. die Tabellen **509**) zu aktualisieren oder zu konfigurieren, welche von diesem Ein-/Ausgabeprozessor unterhalten werden, um Veränderungen, welche an der Konfiguration des Arrays von Festplattenmassenspeichervorrichtungen aufgetreten sein können, welches dem HBA zugeordnet ist, von welchem der Frame **62** übertragen wurde, einschließlich z. B. Veränderungen an der Natur und den Eigenschaften der physikalischen und/oder logischen Datenträger wiederzugeben, welche in einem derartigen Array enthalten sind.

[0048] Bei einer passenden Modifizierung auf Wegen, welche Durchschnittsfachleuten offenkundig sind, können verschiedene der Operationen, welche hier bei dieser Ausführungsform der vorliegenden Erfindung als von Software-Prozessen ausgeführt beschrieben werden, stattdessen durch Hardware/einen Schaltkomplex implementiert werden und vice versa. Die Software-Prozesse **19**, **45**, **207** und **507**, welche benötigt werden, um diese Ausführungsform der vorliegenden Erfindung zu implementieren, können als Ergebnis der Ausführung von jeweiligen Sätzen von Programmbefehlen, welche in computerlesbaren Speichern gespeichert/codiert sind (z. B. in Nur-Lese-Speichern, welche in den Speichern **18**, **44**, **202** bzw. **502** enthalten sind, und/oder in nicht gezeigtem Nur-Lese-Optikplattenspeicher), durch Prozessoren **16**, **42**, **27** bzw. **53** erzeugt werden.

[0049] Zusammenfassend kann folglich gemäß einer Ausführungsform der vorliegenden Erfindung ein Netzwerk bereitgestellt werden, welches einen ersten und einen zweiten Netzwerkknoten, welche über eine Netzwerkkommunikationsverbindung miteinander verbunden sind, und einen ersten und einen zweiten Satz Massenspeichervorrichtungen aufweisen kann, welche dem ersten bzw. dem zweiten Netzwerkknoten zugeordnet sind. Der erste und der zweite Netzwerkknoten können einen Ursprungs-Server bzw. einen Ziel-Server umfassen. Der Ursprungs-Server und der Ziel-Server können einen ersten bzw. einen zweiten Eingangs-/Ausgangs-(E/A)-Prozessor umfassen. Der erste Ein-/Ausgabeprozessor kann als Reaktion auf eine erste Anforderung sowohl die Ausführung einer ersten datenspeicherbezogenen Operation, welche dem ersten Satz der Massenspeichervorrichtungen zugeordnet ist, als auch die Ausgabe einer zweiten Anforderung durch den Ursprungs-Server über die Verbindung an den Ziel-Server bewirken, um zu bewirken, dass der zweite Ein-/Ausgabeprozessor als Reaktion auf die zweite Anforderung eine zweite datenspeicherbezogene Operation durchführt, welche dem zweiten Satz Massenspeichervorrichtungen zugeordnet ist.

[0050] Vorteilhafterweise kann bei dieser Ausführungsform der vorliegenden Erfindung die überwiegende Mehrheit der Operationen, welche in dem Ursprungs- und in dem Ziel-Server durchgeführt werden, um datenspeicherbezogene Operationen auszuführen, von den jeweiligen CPUs, den Systemspeichern und den Systembussen in diesen Servern an E/A-Prozessoren der HBA, an interne HBA-Speicher usw. übertragen werden. Dies erlaubt es, die Zahl der CPUs, der Systemspeicher und der Systembustransaktionen und -Operationen, welche in dem Ursprungs- und in dem Ziel-Server durchgeführt werden, um datenspeicherbezogene Operationen bei dieser Ausführungsform der vorliegenden Erfindung auszuführen, im Vergleich zu dem Stand der Technik

wesentlich zu reduzieren. Vorteilhafterweise verhindert dies, dass übermäßige Anteile der CPU des Ursprungs- und des Ziel-Servers, des Systemspeichers und der Verarbeitungsbetriebsmittel des Systembusses beim Ausführen derartiger Operationen verbraucht werden, und es macht derartige Betriebsmittel zur Verwendung bei anderen Verarbeitungsaktivitäten frei, und im Vergleich zu dem Stand der Technik erhöht es die Geschwindigkeit und die Zuverlässigkeit, mit welcher derartige Operationen bei dieser Ausführungsform der vorliegenden Erfindung ausgeführt werden können.

[0051] Die Begriffe und Ausdrücke, welche hier eingesetzt wurden, werden als Begriffe zur Beschreibung und nicht zur Beschränkung verwendet, und es besteht bei der Verwendung derartiger Begriffe und Ausdrücke keine Absicht, Äquivalente der gezeigten und beschriebenen Merkmale (oder ihrer Teile) auszuschließen, und es ist bekannt, dass innerhalb des Schutzzumfangs der Ansprüche verschiedene Modifikationen möglich sind. Dementsprechend sind die Ansprüche vorgesehen, alle derartigen Äquivalente abzudecken.

Patentansprüche

1. Ein erster Adapter zur Verwendung in einem ersten Netzwerkknoten (12) in einem Netzwerk, wobei der erste Adapter Folgendes umfasst: eine Schaltung (25, 27) zur Durchführung, in Reaktion auf eine erste Anforderung (60), eines ersten datenspeicherbezogenen Vorgangs im Zusammenhang mit einem ersten Satz Massenspeichervorrichtungen (30, 32) und zur Ausgabe, ebenfalls in Reaktion auf die erste Anforderung, einer zweiten Anforderung (62) vom ersten Adapter zu einem zweiten Adapter in einem zweiten Netzwerkknoten (14) im Netzwerk, um zu verursachen, dass der zweite Adapter in Reaktion auf die zweite Anforderung einen zweiten datenspeicherbezogenen Vorgang im Zusammenhang mit einem zweiten Satz Massenspeichervorrichtungen (56, 58) ausführt; und eine Schaltung (25, 27) zur Ausgabe, in Reaktion auf eine erste Meldung (66) vom zweiten Adapter, einer zweiten Meldung (72) zu einem Prozess im ersten Netzwerkknoten, wobei die erste Meldung angibt, dass der zweite datenspeicherbezogene Vorgang abgeschlossen wurde, und die zweite Meldung angibt, dass ein entsprechender datenspeicherbezogener Vorgang, der von der ersten Anforderung angefordert wurde, abgeschlossen wurde, gekennzeichnet dadurch, dass die zweite Anforderung und die erste Meldung jeweils ein entsprechendes Zielknotenadressfeld (106), ein Startknotenadressfeld (108), ein Befehlsfeld (110) und ein Meldungskennungsfeld (114) umfassen; und dass die entsprechenden Meldungskennungsfelder (114) in der zweiten Anforderung und die erste Meldung jeweils identische Werte enthalten.

2. Der erste Adapter nach Anspruch 1, wobei die Schaltung (25, 27) und der zweite Adapter jeweils einen entsprechenden Ein-/Ausgabe-Prozessor umfassen.

3. Der erste Adapter nach Anspruch 1, wobei: der erste Adapter einen ersten Host-Bus-Adapter (26, 28) umfasst, der mit dem ersten Satz Massenspeichervorrichtungen (30, 32) gekoppelt ist; der zweite Adapter einen zweiten Host-Bus-Adapter (52, 54) umfasst, der mit dem zweiten Satz Massenspeichervorrichtungen (56, 58) gekoppelt ist; und der erste Adapter und der zweite Adapter über einen Netzwerkkommunikations-Link (34, 36, 38, 40) miteinander gekoppelt sind.

4. Ein Verfahren zur Verwendung eines ersten Adapters in einem ersten Netzwerkknoten (12) in einem Netzwerk, das Folgendes umfasst: die Verwendung eines ersten Adapters zur Durchführung, in Reaktion auf eine erste Anforderung (60), eines ersten datenspeicherbezogenen Vorgangs im Zusammenhang mit einem ersten Satz Massenspeichervorrichtungen (30, 32); die Ausgabe, ebenfalls in Reaktion auf die erste Anforderung, einer zweiten Anforderung (62) vom ersten Adapter zu einem zweiten Adapter in einem zweiten Netzwerkknoten (14) im Netzwerk, um zu verursachen, dass der zweite Adapter in Reaktion auf die zweite Anforderung einen zweiten datenspeicherbezogenen Vorgang im Zusammenhang mit einem zweiten Satz Massenspeichervorrichtungen (56, 58) ausführt; und die Ausgabe, in Reaktion auf eine erste Meldung (66) vom zweiten Adapter, einer zweiten Meldung (72) vom ersten Adapter zu einem Prozess im ersten Netzwerkknoten, wobei die erste Meldung angibt, dass der zweite datenspeicherbezogene Vorgang abgeschlossen wurde, und die zweite Meldung angibt, dass ein entsprechender datenspeicherbezogener Vorgang, der von der ersten Anforderung angefordert wurde, abgeschlossen wurde, gekennzeichnet dadurch, dass die zweite Anforderung und die erste Meldung jeweils ein entsprechendes Zielknotenadressfeld (106), ein Startknotenadressfeld (108), ein Befehlsfeld (110) und ein Meldungskennungsfeld (114) umfassen; und dass die entsprechenden Meldungskennungsfelder (114) in der zweiten Anforderung und die erste Meldung jeweils identische Werte enthalten.

5. Verfahren nach Anspruch 4, wobei der erste Adapter und der zweite Adapter jeweils einen entsprechenden Ein-/Ausgabe-Prozessor umfassen.

6. Verfahren nach Anspruch 4, wobei: der erste Adapter einen ersten Host-Bus-Adapter (26, 28) umfasst, der mit dem ersten Satz Massenspeichervorrichtungen (30, 32) gekoppelt ist; der zweite Adapter einen zweiten Host-Bus-Adapter

(**52, 54**) umfasst, der mit dem zweiten Satz Massenspeichervorrichtungen (**56, 58**) gekoppelt ist; und der erste Adapter und der zweite Adapter über einen Netzwerkkommunikations-Link (**34, 36, 38, 40**) miteinander gekoppelt sind.

Es folgen 9 Blatt Zeichnungen

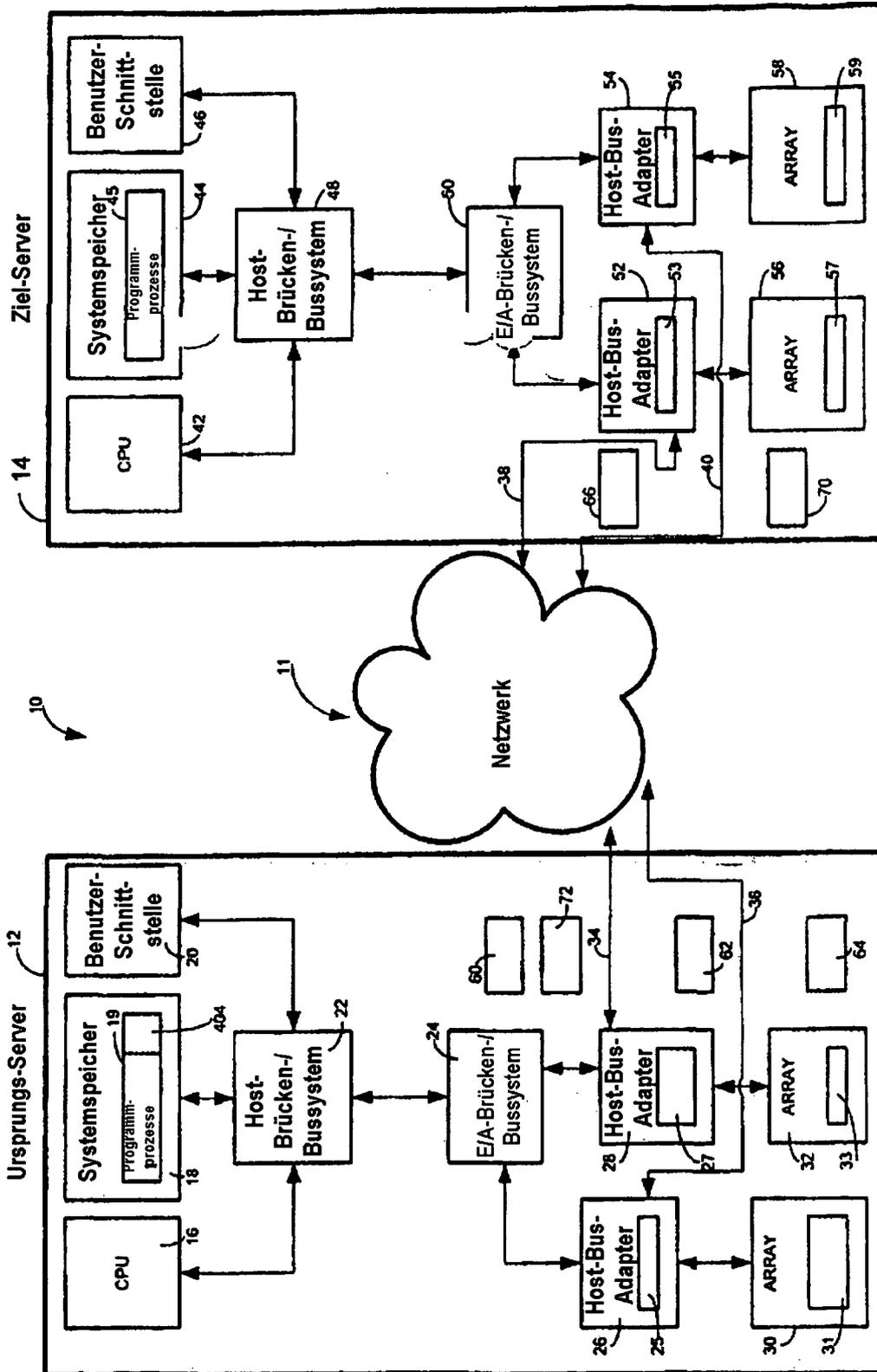


FIG. 1

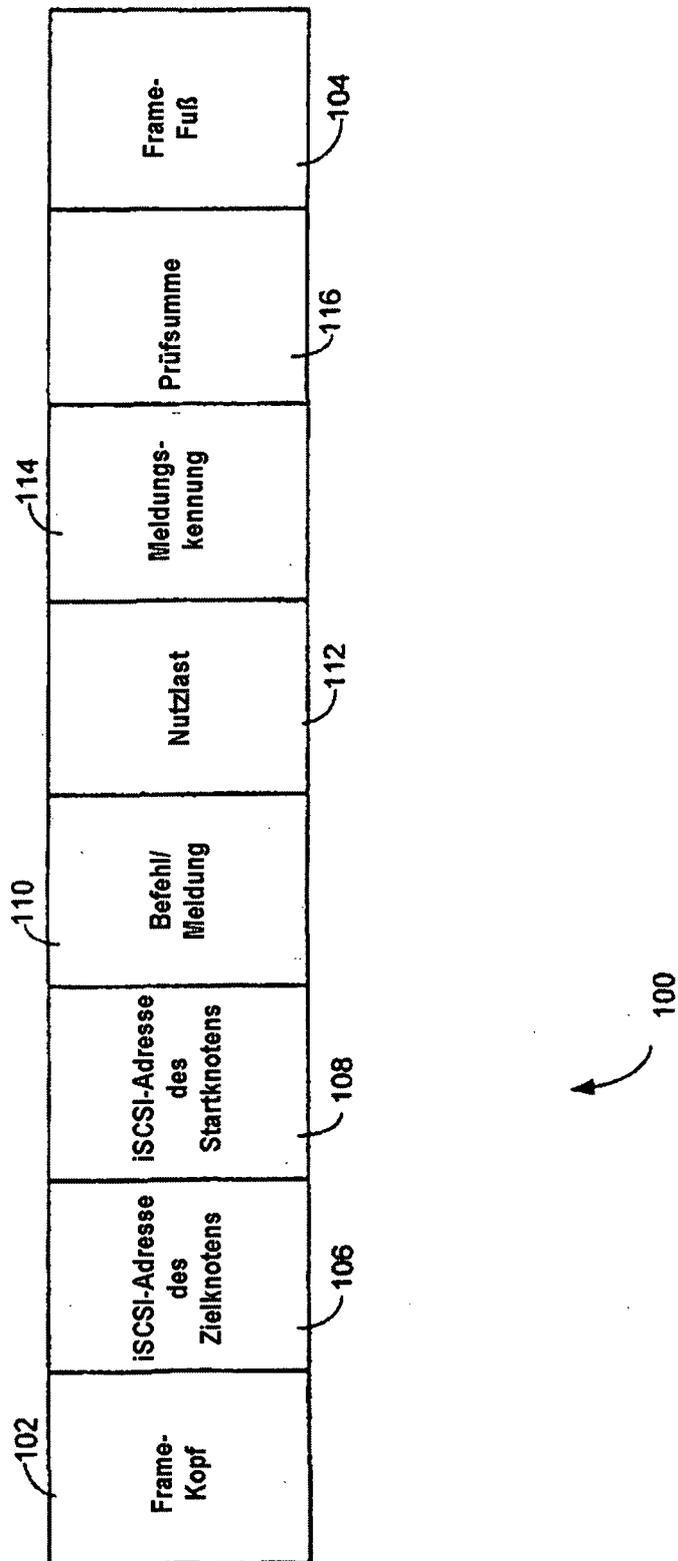


FIG. 2

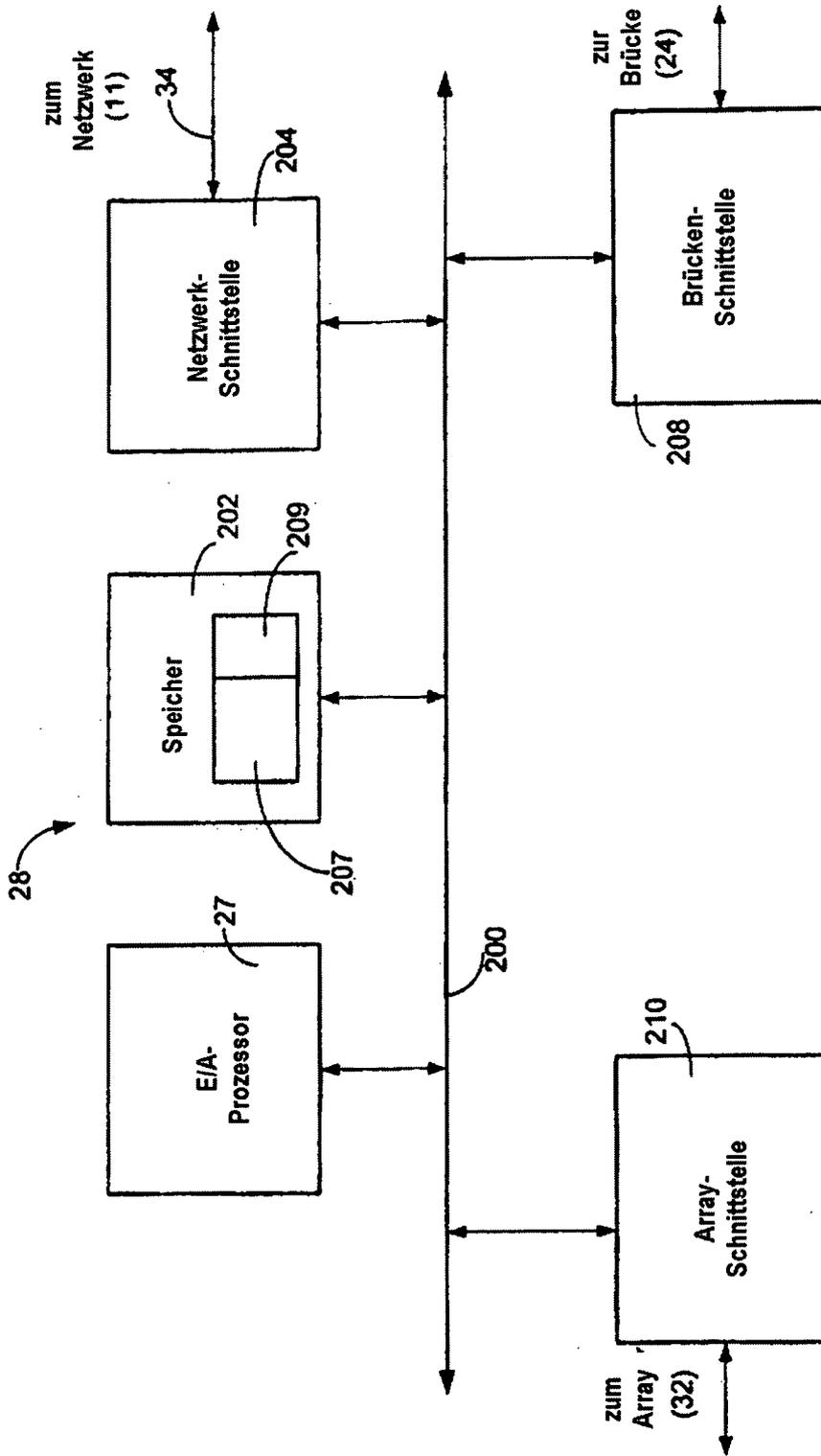


FIG. 3

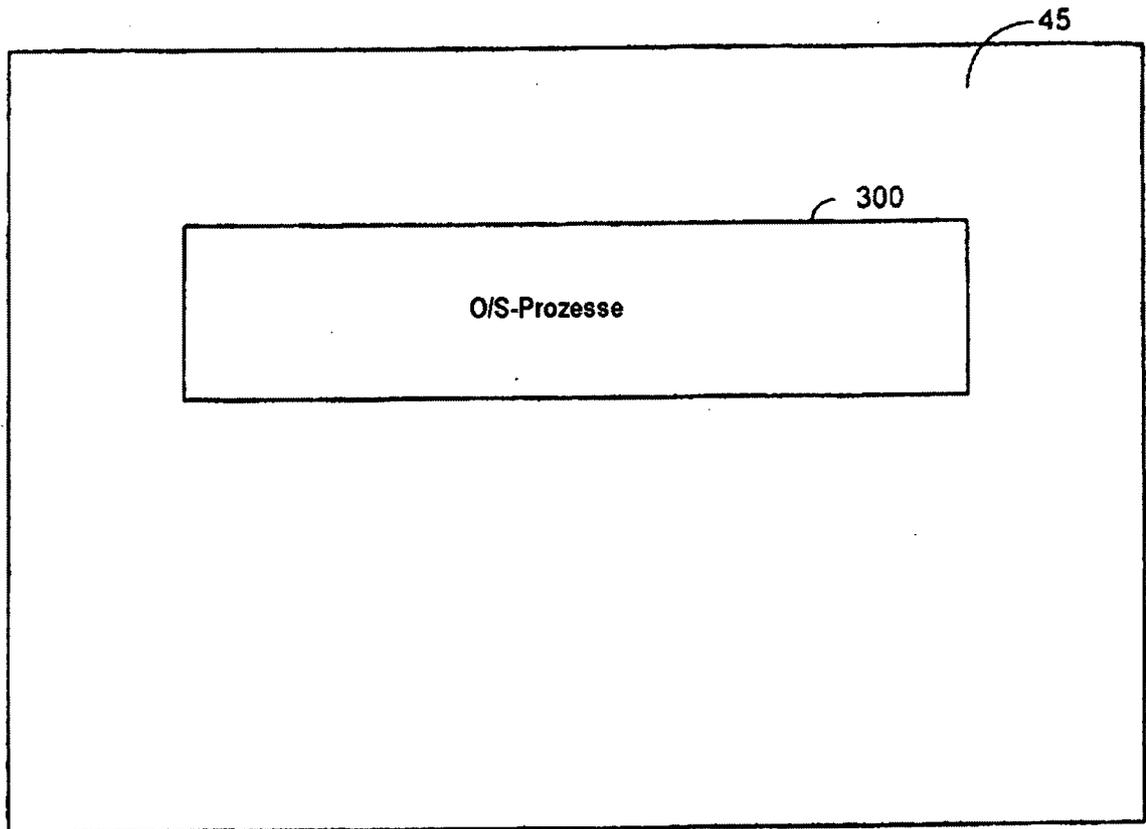


FIG. 4

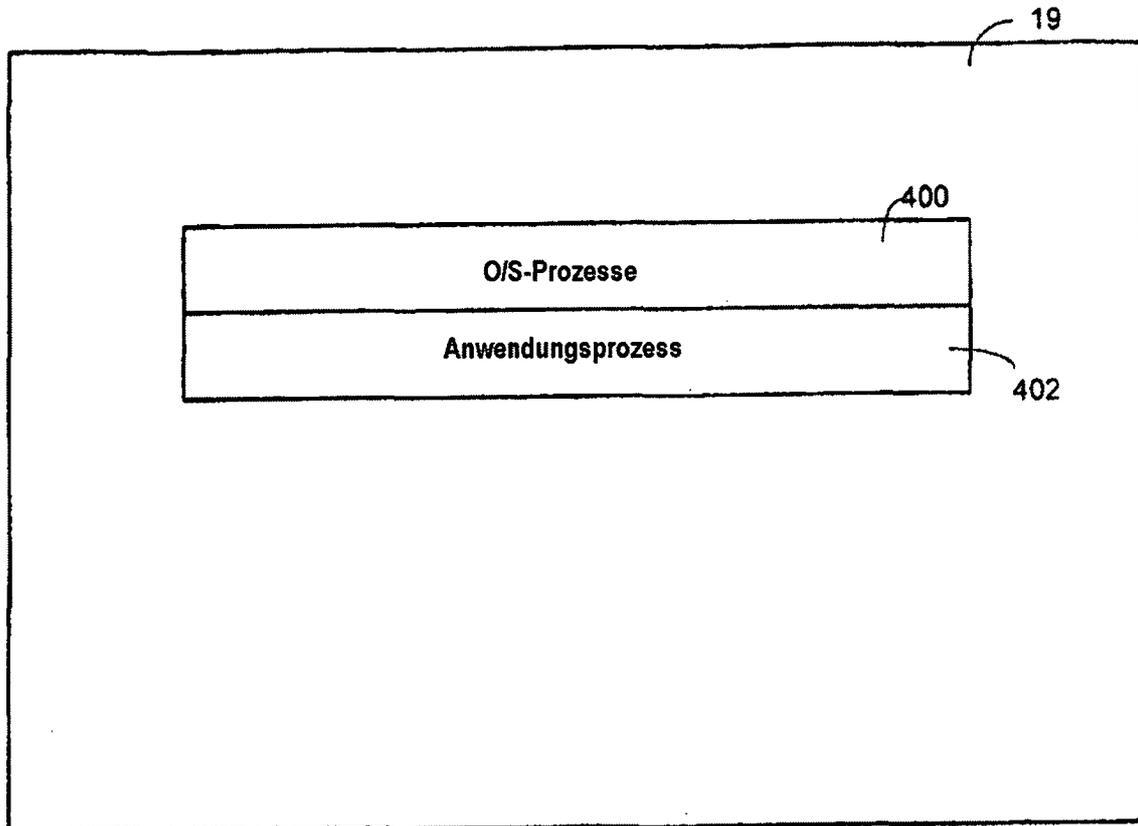


FIG. 5

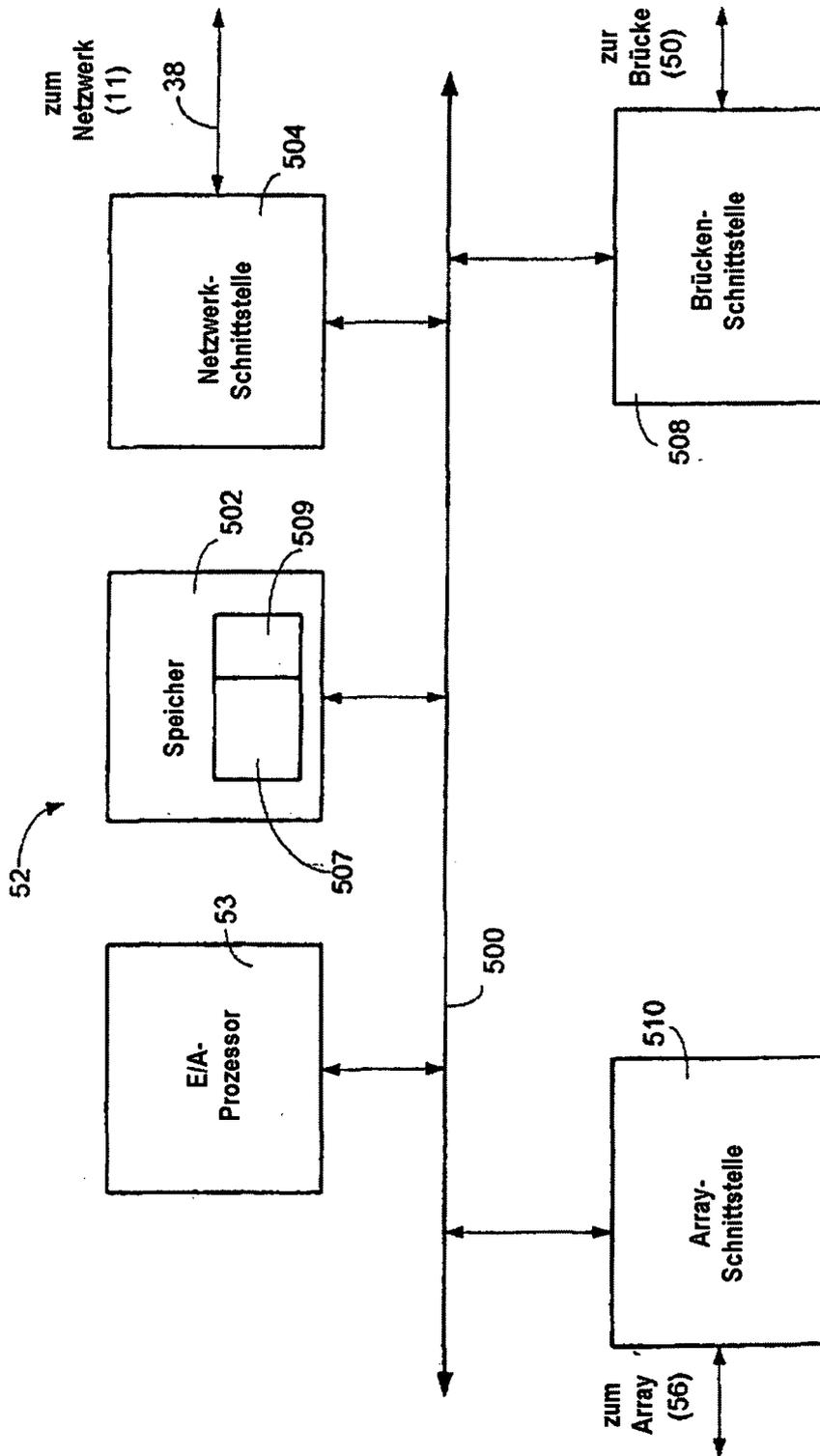


FIG. 6

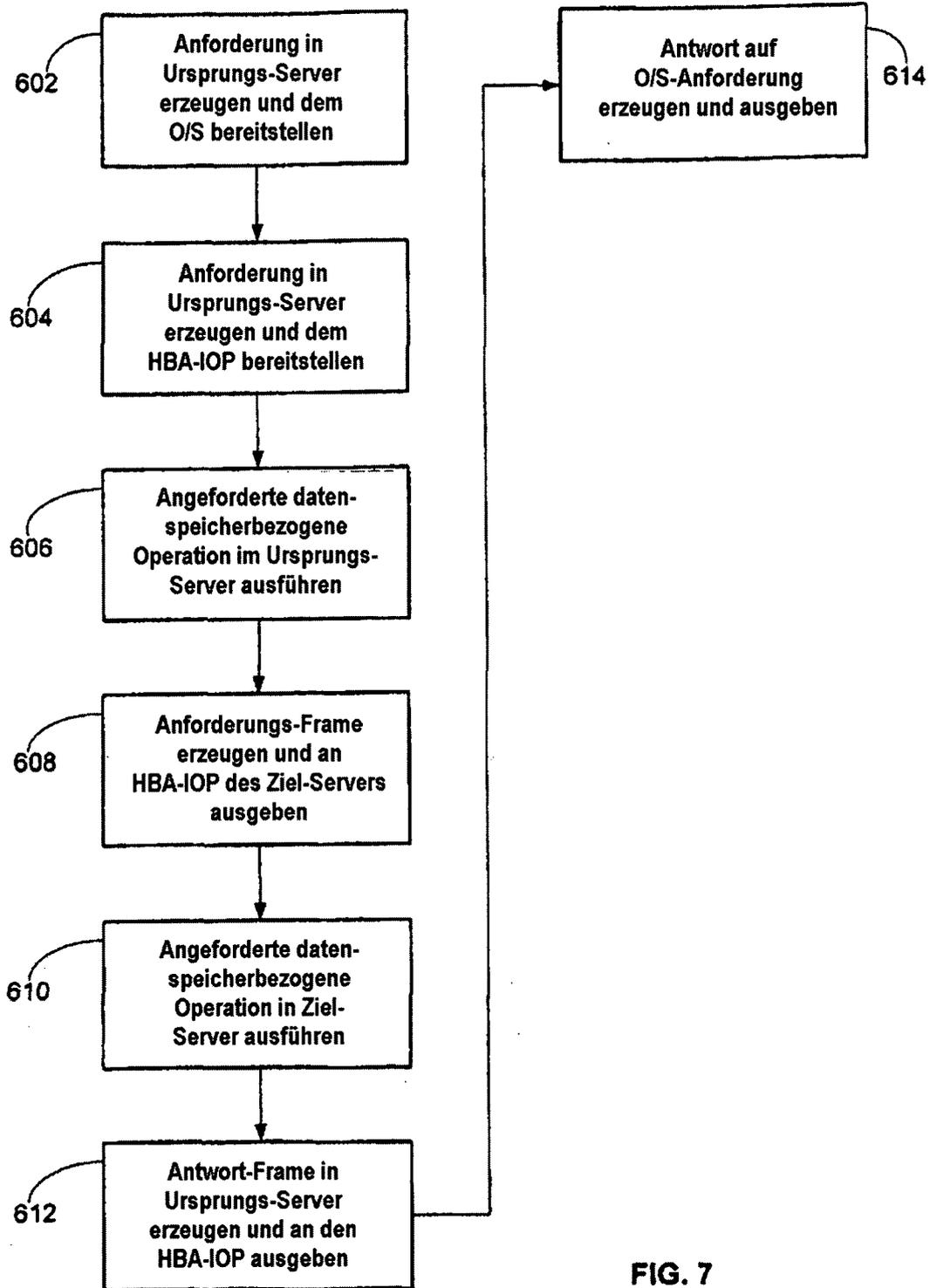


FIG. 7

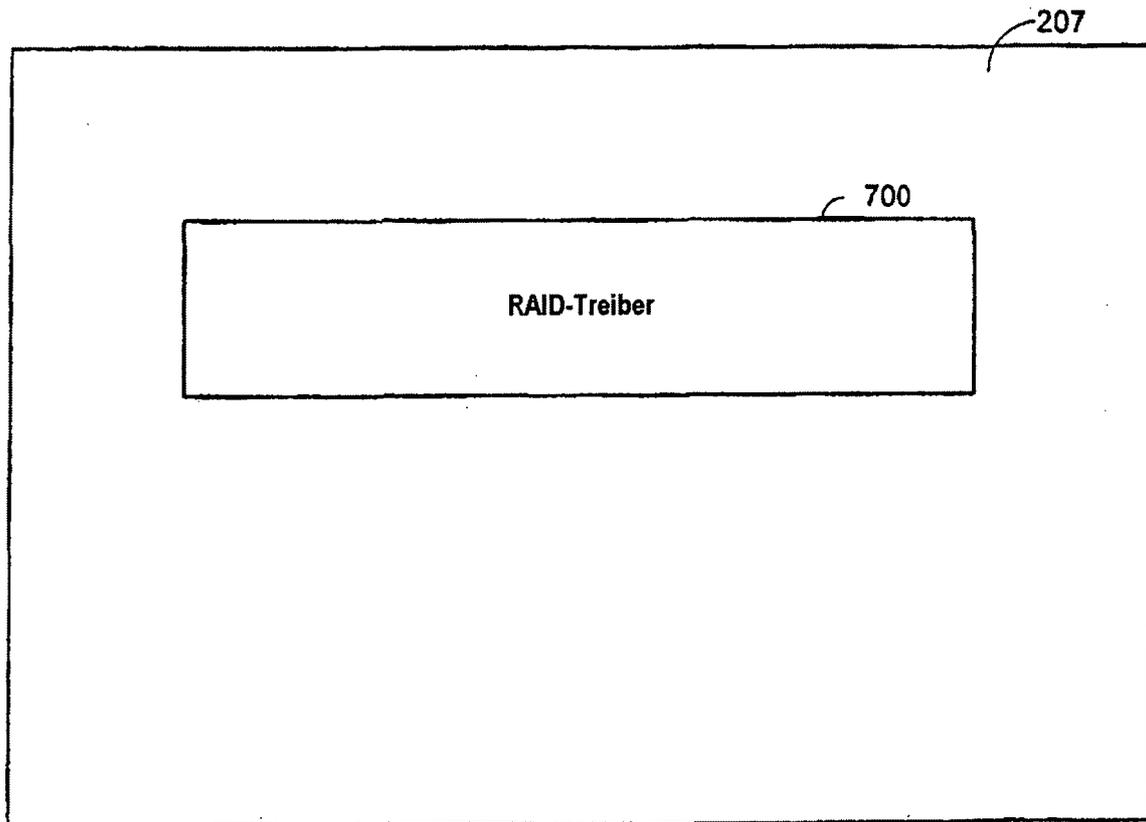


FIG. 8

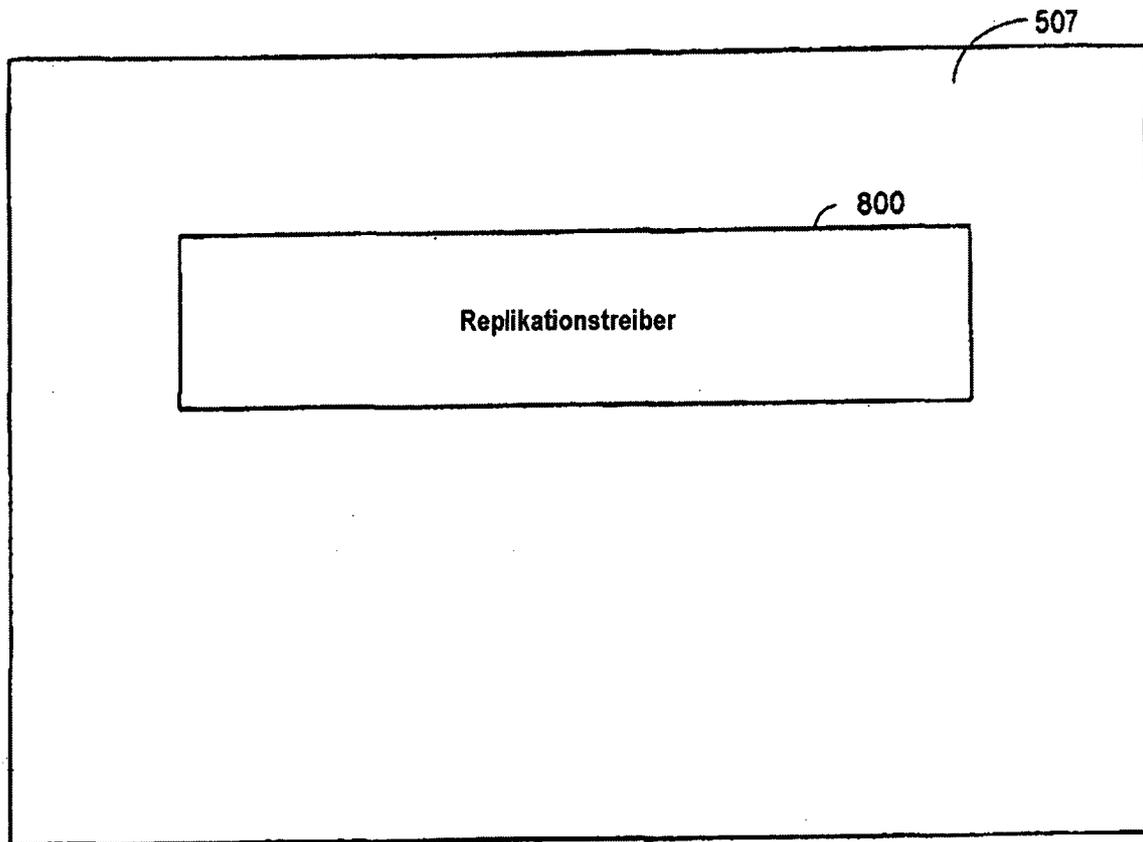


FIG. 9