(19) Deutsches Patent- und Markenamt

(10) **DE 10 2012 210 582 B4** 2018.05.30

(12) # Patentschrift

(21) Aktenzeichen: **10 2012 210 582.4**
(22) Anmeldetag: **22.06.2012**
(43) Offenlegungstag: **10.01.2013**
(45) Veröffentlichungstag der Patenterteilung: **30.05.2018**

(51) Int Cl.: ***G06F 1/16*** (2006.01)
G06F 15/16 (2006.01)
H05K 1/14 (2006.01)

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

(30) Unionspriorität:
13/177,639 **07.07.2011** **US**

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

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

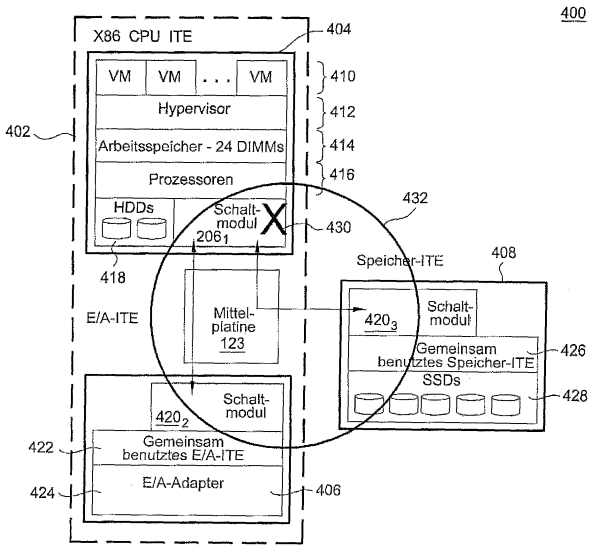
(72) Erfinder:
Armstrong, William J., Rochester, Minn., US; Borkenhagen, John M., Rochester, Minn., US; Crippen, Martin J., Research Triangle Park, N.C., US; Desai, Dhruv M., Research Triangle Park, N.C., US; Engebretsen, David R., Rochester, Minn., US; Hillier III, Philip R., Rochester, Minn., US; Holland, William G., Research Triangle Park, N.C., US; Hughes, James E., Research Triangle Park, N.C., US; O'Connor, James A., Poughkeepsie, N.Y., US; Tri, Steven M., Rochester, Minn., US

(56) Ermittelter Stand der Technik:

US	**6 339 546**	**B1**
US	**6 976 112**	**B2**
US	**7 191 347**	**B2**
US	**7 295 446**	**B2**
WO	**2007/ 064 466**	**A2**

(54) Bezeichnung: **Verringern der Auswirkung des Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk mittels Schaltkarten**

(57) Zusammenfassung: Methoden zur Verringerung der Auswirkung eines Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk werden offengelegt. In einer Ausführungsform wird ein Serversystem bereitgestellt, das eine Mittelplatine, eine oder mehrere Serverkarten und eine oder mehrere Schaltkarten enthält. Die Mittelplatine kann eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk enthalten. Die eine oder die mehreren Serverkarten können mit der Mittelplatine verbunden werden, wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann. Die eine oder die mehreren Schaltkarten können ebenfalls mit der Mittelplatine verbunden werden, wobei jede Schaltkarte auch im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann. Jede Schaltkarte enthält ein oder mehrere Schaltmodule und jedes Schaltmodul ist so konfiguriert, dass es den Netzwerkverkehr für mindestens eine Serverkarte schaltet.



Beschreibung

HINTERGRUND

[0001] Während die ersten Rechnerarchitekturen einzelne eigenständige Rechner verwendeten, die oft als Personal Computer (PCs) bezeichnet werden, verwenden moderne, leistungsfähigere Rechnersysteme häufig mehrere Rechner, die in einem gemeinsamen Gehäuse miteinander verbunden sind. Ein beispielhaftes gemeinsames Gehäuse wird als Blade-Chassis bezeichnet, das mehrere Serverblades enthält, die über eine gemeinsame Zentralverbindung (Backbone) in dem Blade-Chassis miteinander verbunden sind. Jeder Serverblade ist eine Steckkarte, die mindestens einen Prozessor, einen auf der Karte befindlichen Speicher und eine Eingabe-/Ausgabe-(E/A-)Schnittstelle enthält. Die mehreren Serverblades sind so konfiguriert, dass sie miteinander kommunizieren und gemeinsame Ressourcen wie zum Beispiel Speichereinheiten, Bildschirme, Eingabeeinheiten usw. gemeinsam nutzen. Auch können ein oder mehrere Blade-Chassis ein Blade-System bilden, das oftmals eigens für ein einziges Unternehmen und/oder eine bestimmte Funktion wie zum Beispiel die Kreditbearbeitung, die Gehaltsabrechnung usw. vorgesehen ist.

[0002] Die US 6,339,546 B1 beschreibt eine Halbleiterspeichervorrichtung, welche die Ursache eines Fehlers zum Zeitpunkt der Fehlerkorrektur von aus einem nichtflüchtigen Halbleiterspeicher ausgelesenen Daten auf der Basis eines zuvor aufgezeichneten Fehlerkorrekturzählwerts bestimmt und eine Datenaktualisierungsverarbeitung oder eine Ersatzverarbeitung zur Ausführung auswählt. Wenn der Fehler erkannt wird, werden die korrigierten Daten zurückgeschrieben, um ein erneutes Auftreten eines Fehlers aufgrund einer zufälligen Ursache zu verhindern. Wenn bestimmt wird, dass die Wiederholungshäufigkeit des Fehlers hoch ist und der Fehler auf eine Verschlechterung des Speichermediums zurückzuführen ist, wird basierend auf der Fehlerkorrekturzählung die Ersatzverarbeitung ausgeführt.

[0003] Die US 6,976,112 B2 beschreibt ein Hot-Plugging-Verfahren, bei welchem ein Serverblade und/oder ein Interconnect-Gerät in das Gehäuse eines aktiven Servers eingesetzt wird. Bevor Strom an das hot-plugged heiß Serverblade und/oder Interconnect-Gerät angelegt wird, wird der Fabrictyp von bereits installierten Serverblades und/oder Interconnect-Geräten mit Fabrictypen von neu hot-plugged Serverblades und/oder Interconnect-Geräten korreliert. Abhängig von den Ergebnissen der Korrelation wird der Strom zu dem hot-plugged Serverblade und/oder Interconnect-Gerät zugelassen oder verweigert.

[0004] Die US 7,191,347 B2 beschreibt die Verwaltung von Blade-Server-Informationsverarbeitungssystemen durch Bereitstellen eines nicht unterbrechenden Anzeigesignals für das Energiemanagement an ein Verwaltungsmodul oder ähnliches. Durch Bereitstellen der Funktionalität und des Verfahrens kann die verfügbare Energie zum Betreiben des Verwaltungsmoduls in einer bestimmten Instanz in Bezug auf den Energiezustand des Chassis bestimmt werden, so dass Energie für das Verwaltungsmodul auf nicht unterbrechende Weise konfiguriert werden kann. Diese Funktionalität und dieses Verfahren sorgen zudem für eine Stromversorgung eines Servers von einem Kaltstart oder dem Hot-Plugging eines Verwaltungsmoduls in ein aktives, mit strom versorgtes Chassis.

[0005] Die US 7,295,446 B2 beschreibt ein Computersystem, welches eine Mehrzahl von Blade-Servern, eine Mittelplatine, eine zusätzliche Mittelplatine und eine von den Blade-Servern getrennte Peripherievorrichtung. Die Mittelplatine umfasst eine Mehrzahl von Verbindern, welche die Mittelplatine mit der Mehrzahl von Blade-Servern verbinden. Die zusätzliche Mittelplatine ist von der Mittelplatine getrennt und umfasst einen Körper, eine Mehrzahl von Verbindern und einen peripheren Verbinder. Die Mehrzahl von Verbindern ist an dem Körper angebracht und verbindet die zusätzliche Mittelplatine mit der Mehrzahl von Blade-Servern. Der periphere Verbinder ist an dem Körper angebracht und verbindet die zusätzliche Mittelplatine mit dem Peripheriegerät. Die mehreren Verbinder der zusätzlichen Mittelplatine kommunizieren mit dem peripheren Verbinder der zusätzlichen Mittelplatine. Die Mehrzahl von Verbindern der zusätzlichen Mittelplatine ist so konfiguriert, dass sie entfernbar mit der Mittelplatine verbindbar ist.

[0006] Die WO 2007/064466 A2 beschreibt einen Telco-Hub in einem Chassis mit einer Vorderseite und einer Rückseite, wobei das Chassis eine Mittelplatine mit einer ersten Seite und einer zweiten Seite umfasst, ein Interposer-Modul, das gleitbar in einen Wechselslot an der Rückseite eingefügt und mit der zweiten Seite der Mittelplatine verbunden ist, ein mit dem Interposer-Modul verbundenes Mezzanine-Moduls, wobei das Mezzanine-Modul eine Telco-I/O-Schnittstelle umfasst und wobei das Mezzanine-Modul gekoppelt ist, um nicht paketierte Telefondaten zu empfangen. Der Telco-Hub ist gekoppelt, um die nicht-paketierten Telefondaten in paketierte Daten umzuwandeln und die paketierten Daten an ein oder mehrere Nutzlastmodule einer Mehrzahl von Nutzlastmodulen über ein vermitteltes Fabric auf der Mittelplatine zu kommunizieren, wo jedes der ein oder mehreren Nutzlastmodule über einen Nutzlastslot an der Vorderseite des Chassis mit der ersten Seite der Mittelplatine verbunden ist.

KURZDARSTELLUNG

[0007] Eine Ausführungsform der Erfindung stellt ein System bereit, das eine Mittelplatine, eine oder mehrere Serverkarten, die mit der Mittelplatine verbunden sind, und eine oder mehrere Schaltkarten, die mit der Mittelplatine verbunden sind, enthält. Die Mittelplatine enthält eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk. Jede Serverkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden. Die eine oder die mehreren Schaltkarten sind mit der einen oder den mehreren Serverkarten betriebsfähig verbunden. Jede Schaltkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden und enthält ein oder mehrere Schaltmodule. Jedes Schaltmodul ist so konfiguriert, dass es den Netzwerkverkehr für mindestens eine Serverkarte von der einen oder den mehreren Serverkarten schaltet. Ein erstes Schaltmodul einer ersten Schaltkarte ist so konfiguriert, dass es den Netzwerkverkehr für die eine oder die mehreren Serverkarten nach einem Feststellen eines Ausfalls eines zweiten Schaltmoduls schaltet, das in der ersten Schaltkarte enthalten ist, und dadurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert.

[0008] Eine andere Ausführungsform der Erfindung stellt ein Schaltmodul für das zuvor genannte System bereit, das einen Rechnerprozessor und einen Arbeitsspeicher enthält. Der Arbeitsspeicher speichert Verwaltungs-Firmware, die, wenn sie auf dem Rechnerprozessor ausgeführt wird, eine Operation durchführt, die das Schalten des Netzwerkverkehrs für eine erste Serverkarte in einem Serversystem beinhaltet. Die Operation beinhaltet auch das Schalten des Netzwerkverkehrs für die zweite Serverkarte, nachdem ein Ausfall eines zweiten Schaltmoduls festgestellt wird, das den Netzwerkverkehr für eine zweite Serverkarte schaltet. Das Schaltmodul ist in einer ersten Schaltkarte enthalten. Das zweite Schaltmodul ist in der ersten Schaltkarte enthalten. Jede Schaltkarte und jede Serverkarte sind mit einer Mittelplatine verbunden. Die Mittelplatine enthält eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk. Jede Schaltkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden, und jede Serverkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden, wodurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert wird.

[0009] Noch eine weitere Ausführungsform der Erfindung stellt ein von einem Rechner durchgeführtes Verfahren bereit, das die Feststellung beinhaltet, dass in einem Serversystem, das eine Mittelplatine, eine oder mehrere Serverkarten, die mit der Mittelplatine verbunden sind, und eine oder mehrere Schaltkarten, die mit der Mittelplatine verbunden sind, enthält, ein erstes Schaltmodul einer ersten Schaltkarte ausgefallen ist. Die eine oder die mehreren Serverkarten sind mit der einen oder den mehreren Schaltkarten betriebsfähig verbunden. Die Mittelplatine enthält eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk. Jede Schaltkarte enthält ein oder mehrere Schaltmodule. Jedes Schaltmodul ist so konfiguriert, dass es den Netzwerkverkehr für mindestens eine Serverkarte von der einen oder den mehreren Serverkarten schaltet. Jede Serverkarte und jede Schaltkarte können im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden. Die Operation beinhaltet auch das Schalten des Netzwerkverkehrs für die eine oder die mehreren Serverkarten durch ein zweites Schaltmodul, das in der ersten Schaltkarte enthalten ist, nachdem festgestellt wurde, dass das erste Schaltmodul der ersten Schaltkarte ausgefallen ist, und dadurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert.

[0010] In einem weiteren Aspekt betrifft die Erfindung ein von einem Rechner durchgeführtes Verfahren, das Folgendes umfasst:

> in einem Serversystem, das eine Mittelplatine, eine oder mehrere Serverkarten, die mit der Mittelplatine verbunden sind, und eine oder mehrere Schaltkarten, die mit der Mittelplatine verbunden sind, umfasst, wobei die eine oder die mehreren Serverkarten mit der einen oder den mehreren Schaltkarten betriebsfähig verbunden sind, wobei die Mittelplatine eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk umfasst, wobei jede Schaltkarte ein oder mehrere Schaltmodule umfasst, wobei jedes Schaltmodul so konfiguriert ist, dass es den Netzwerkverkehr für mindestens eine Serverkarte von der einen oder den mehreren Serverkarten schaltet, wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann und wobei jede Schaltkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann, Feststellen, dass ein erstes Schaltmodul einer ersten Schaltkarte ausgefallen ist; und
>
> nach der Feststellung, dass das erste Schaltmodul der ersten Schaltkarte ausgefallen ist, Schalten des Netzwerkverkehrs für die eine oder die mehreren Serverkarten durch ein zweites Schaltmodul, das in der ersten Schaltkarte enthalten ist, und dadurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert.

[0011] Nach einer Ausführungsform der Erfindung sind die eine oder die mehreren Serverkarten mit einer ersten Seite der Mittelplatine verbunden, wobei

die eine oder die mehreren Schaltkarten mit einer zweiten Seite der Mittelplatine verbunden sind.

[0012] Nach einer Ausführungsform der Erfindung umfasst das Schaltnetzwerk mindestens eines von Folgendem: (i) eine Verdrahtung zwischen jeder Schaltkarte und jeder Serverkarte und (ii) eine Verdrahtung zwischen jeder Schaltkarte und jeder anderen Schaltkarte.

[0013] Nach einer Ausführungsform der Erfindung sind die eine oder die mehreren Serverkarten entsprechend einer ersten Achse auf der Mittelplatine ausgerichtet, wobei die eine oder die mehreren Schaltkarten entsprechend einer zweiten Achse auf der Mittelplatine ausgerichtet sind und wobei die zweite Achse senkrecht zu der ersten Achse liegt.

[0014] Nach einer Ausführungsform der Erfindung sind die eine oder die mehreren Serverkarten waagerecht mit einer Frontseite der Mittelplatine verbunden, wobei die eine oder die mehreren Schaltkarten senkrecht mit einer zweiten Seite der Mittelplatine verbunden sind.

[0015] Nach einer Ausführungsform der Erfindung sind eine Serverkarte und/oder eine Schaltkarte so konfiguriert, dass sie gegen einen funktionsfähigen Ersatz ausgetauscht werden können, ohne einen Neustart des Serversystems erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks erforderlich zu machen.

[0016] Nach einer Ausführungsform der Erfindung ist das Serversystem so konfiguriert, dass es den funktionsfähigen Ersatz in das Schaltnetzwerk integriert, ohne einen Neustart des Systems erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks erforderlich zu machen.

[0017] Nach einer Ausführungsform der Erfindung umfasst das Serversystem ein Blade-System, wobei jede Serverkarte einen Server-Blade umfasst und wobei der Netzwerkverkehr mindestens eines von Folgendem umfasst: (i) Ethernet-Verkehr und (ii) Fibre-Channel-over-Ethernet-(FCoE)-Verkehr.

Figurenliste

[0018] Damit sich die Art und Weise, in der die vorstehend erwähnten Erscheinungsformen erzielt werden, im Einzelnen verstehen lässt, erfolgt nun eine ausführlichere Beschreibung von Ausführungsformen der Erfindung, die vorstehend kurz zusammengefasst wurden, indem auf die beigefügten Zeichnungen Bezug genommen wird.

[0019] Es sei jedoch erwähnt, dass die beigefügten Zeichnungen lediglich typische Ausführungsformen dieser Erfindung darstellen und folglich nicht als Einschränkung des Umfangs der Erfindung zu verstehen sind, da die Erfindung auch andere, gleichermaßen wirksame Ausführungsformen zulässt.

Fig. 1 ist ein Blockschaltbild einer Datenverarbeitungsumgebung gemäß einer Ausführungsform der Erfindung, die über mehrere Hostrechner verfügt, welche Zugriff auf ein Serversystem haben.

Fig. 2 zeigt eine Konfiguration, bei der gemäß einer Ausführungsform der Erfindung Zwischenstecker-Karten mit Serverkarten in einem Serversystem betriebsfähig verbunden sind.

Fig. 3 zeigt eine Konfiguration, bei der gemäß einer Ausführungsform der Erfindung eine Zwischenstecker-Karte mit zwei Serverkarten in einem Serversystem betriebsfähig verbunden ist.

Fig. 4 zeigt ein Serversystem, das so konfiguriert ist, dass es gemäß einer Ausführungsform der Erfindung die Auswirkung eines Reparaturvorgangs an einem Schaltmodul verringert.

Fig. 5 zeigt ebenfalls ein Serversystem, das so konfiguriert ist, dass es gemäß einer Ausführungsform der Erfindung die Auswirkung eines Reparaturvorgangs an einem Schaltmodul verringert.

Fig. 6 zeigt ein Schaltnetzwerk für ein Serversystem gemäß einer Ausführungsform der Erfindung.

Fig. 7 zeigt ein Serversystem gemäß einer Ausführungsform der Erfindung, das über eine Mittelplatine verfügt, die mit mehreren Zwischenstecker-Karten verbunden ist.

Fig. 8 zeigt ein Serversystem gemäß einer Ausführungsform der Erfindung, das mehrere Gestellrahmen enthält.

Fig. 9 zeigt ein Serversystem gemäß einer Ausführungsform der Erfindung, das mehrere Gestellrahmen enthält, wobei jeder Gestellrahmen über vier Gehäuse verfügt.

Fig. 10 zeigt ein Serversystem gemäß einer Ausführungsform der Erfindung, das so ausgelegt ist, dass es eine Zwischenstecker-Verschaltungseinheit enthält.

Fig. 11 zeigt eine Konfiguration eines Serversystems gemäß einer Ausführungsform der Erfindung, bei dem ein Schaltmodul als ein einzelner Fehlerpunkt (single point of failure (SPOF)) in einem Paar von Speicher-ITEs beseitigt wird.

Fig. 12 zeigt eine Konfiguration von einem Paar von miteinander verbundenen Zwischenstecker-Karten gemäß einer Ausführungsform der Erfindung.

Fig. 13 zeigt eine Konfiguration eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten enthält.

Fig. 14 zeigt eine logische Ansicht einer Konfiguration eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten enthält.

Fig. 15 zeigt eine Konfiguration eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten enthält.

Fig. 16 ist ein Flussdiagramm, das ein Verfahren gemäß einer Ausführungsform der Erfindung zur Verringerung der Auswirkung eines Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk zeigt.

Fig. 17 ist ein Flussdiagramm, das ein Verfahren gemäß einer Ausführungsform der Erfindung zum Beseitigen eines Schaltmoduls als einen SPOF darstellt.

AUSFÜHRLICHE BESCHREIBUNG

[0020] Ausführungsformen der Erfindung verringern die Auswirkung eines Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk. In der hier verwendeten Weise bezieht sich ein Schaltnetzwerk auf eine Netzwerk-Topologie, bei der sich Netzwerkknoten über eine oder mehrere Netzwerk-Vermittlungsstellen miteinander verbinden lassen. In einer Ausführungsform wird ein Serversystem bereitgestellt, das eine Mittelplatine, eine erste Zwischenstecker-Karte und eine oder mehrere Serverkarten enthält, wobei jede Serverkarte einem oder mehreren Netzwerkknoten entspricht. In einer Ausführungsform kann jede Serverkarte ein Serverblade sein, der auch als Blade-Server oder Blade bezeichnet wird. Die Mittelplatine wurde zwar mit Bezug auf die erste Zwischenstecker-Karte beschrieben, doch kann sie so konfiguriert werden, dass sie sich mit einer Vielzahl von Zwischenstecker-Karten verbinden lässt. Die erste Zwischenstecker-Karte wird zwischen der Mittelplatine und der einen oder den mehreren Serverkarten angeordnet, wodurch die Mittelplatine mit der einen oder den mehreren Serverkarten betriebsfähig verbunden wird. Die erste Zwischenstecker-Karte enthält überdies ein Schaltmodul, das den Netzwerkverkehr für die eine oder die mehreren Serverkarten schaltet. Die erste Zwischenstecker-Karte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden, und die eine oder die mehreren Serverkarten können im laufenden Betrieb von der ersten Zwischenstecker-Karte aus ausgetauscht werden.

[0021] In einer Ausführungsform kann das Schaltmodul, wenn es ausfällt, durch einen Reparaturvorgang ausgetauscht werden, der die Auswirkung auf das Schaltnetzwerk so klein wie möglich hält oder aber verringert. Der Reparaturvorgang beinhaltet den Austausch der ersten Zwischenstecker-Karte gegen eine zweite Zwischenstecker-Karte, die ein funktionales Schaltmodul enthält, und die Wiederaufnahme der zweiten Zwischenstecker-Karte in das Kommunikationsnetzwerk über ein Konfigurations-Werkzeug, das auf dem Serversystem ausgeführt wird. Aufgrund der Auslegung des Serversystems und der Möglichkeit, die Zwischenstecker-Karten und die Serverkarten im laufenden Betrieb auszutauschen, kann der Reparaturvorgang ohne Unterbrechung des Serversystems oder des Schaltnetzwerks durchgeführt werden - z.B. ohne das Serversystem und/oder das Schaltnetzwerk abzuschalten oder neu zu starten. Dort wo das Schaltnetzwerk Redundanz im Hinblick auf die Verbindungen vorsieht, kann der Reparaturvorgang auch die Auswirkung auf die vorgesehene Redundanz so klein wie möglich halten oder verringern. Folglich ist die Auswirkung des Reparaturvorgangs örtlich auf die Serverkarte begrenzt. Anders ausgedrückt, nur die erste Zwischenstecker-Karte und/oder die Serverkarte sind von der Auswirkung des Reparaturvorgangs auf das Schaltnetzwerk betroffen; das Serversystem und das Schaltnetzwerk - nämlich andere Zwischenstecker-Karten und Serverkarten, die mit der Mittelplatine betriebsfähig verbunden sind - bleiben betriebsfähig. Vorteilhafterweise wird die Auswirkung des Reparaturvorgangs im Vergleich zu einer physischen Konfiguration oder der Auslegung, die das Abschalten des Serversystems und/oder des Schaltnetzwerks erforderlich machen, um das Schaltmodul auszutauschen - z.B. indem die Mittelplatine ausgetauscht wird oder indem eine nicht im laufenden Betrieb austauschbare Schaltkarte ausgetauscht wird, die mit der Mittelplatine verbunden ist, verringert. Die Verfügbarkeit des Serversystems und/oder des Schaltnetzwerks wird dadurch verbessert und die Kosten in Verbindung mit dem Reparaturvorgang werden dadurch verringert.

[0022] In einer Ausführungsform kann die Verfügbarkeit des Serversystems und/oder des Schaltnetzwerks - oder der Redundanz-Eigenschaften des Serversystems und/oder des Schaltnetzwerks - in Bezug auf eine zweite Auslegung des Serversystems verbessert werden, bei der die Mittelplatine ausgetauscht werden müsste, um Abhilfe für ein ausgefallenes Schaltmodul zu schaffen. Die zweite Auslegung des Serversystems kann einen oder mehrere Vermittlungsstellen-Chips beinhalten, die auf einer einzelnen mit der Mittelplatine verbundenen Karte (oder Platine) untereinander verbunden sind. Das Verbinden der einzelnen Karte mit der Mittelplatine kann eine höhere Anzahl von Anschlüssen, eine höhere Bandbreite bereitstellen und/oder die Verfügbarkeit des Schaltnetzwerks verbessern. Die zweite Auslegung des Serversystems kann auch mehrere redundante Pfade durch mehrere Vermittlungsstellen-Chips beinhalten, so dass das Serversystem seinen Betrieb fortsetzen kann, wenn ein Vermittlungsstellen-Chip ausfällt. Andere Ausfälle, die sich auf die

einzelne Karte bis hin zur Platine auswirken, können jedoch dazu führen, dass ein Teil des Schaltnetzwerks oder sogar das ganze Schaltnetzwerk nicht mehr funktioniert. Beispiele für die anderen Ausfälle sind unter anderem Ausfälle von Stromversorgungskomponenten, Ausfälle des Spannungsreglermoduls (Voltage Regulator Module (VRM)), Kurzschlüsse auf der Stromversorgungsebene usw.

[0023] Selbst wenn das Schaltnetzwerk bei Vorhandensein von einem oder mehreren ausgefallenen Vermittlungsstellen-Chips betriebsfähig bleiben kann, kann ein Reparaturvorgang an dem einem oder den mehreren ausgefallenen Vermittlungsstellen-Chips in einer Ausführungsform den Austausch der einzelnen Karte, der Platine und/oder der Mittelplatine erforderlich machen, was zu einem Betriebsausfall von mindestens dem Teil des Schaltnetzwerks, der während des Reparaturvorgangs von der Mittelplatine unterstützt wird, führt. Um den Betriebsausfall während des Reparaturvorgangs zu vermeiden, kann das Serversystem so konfiguriert werden, dass es eine zweite, vollständig redundante einzelne Karte (oder Platine) enthält. Alternativ kann das Serversystem mit Hilfe der hier offengelegten Methoden ausgelegt werden, um die Auswirkung des Reparaturvorgangs auf das Schaltnetzwerk zu verringern, während gleichzeitig die Kosten für die Konfiguration des Serversystems mit einer zweiten, vollständig redundanten einzelnen Karte oder Platine vermieden werden. Folglich kann die Verfügbarkeit des Serversystems verbessert werden, da einzelne Fehlerpunkte (single points of failure (SPOFs)) und/oder einzelne Reparaturpunkte (single points of repair (SPORs) verringert oder minimiert werden. SPOFs gelten als beseitigt, wenn das Serversystem bei einem vorhandenen Ausfall einer beliebigen Komponente seinen Betrieb fortsetzen kann. SPORs gelten als beseitigt, wenn das Serversystem seinen Betrieb fortsetzen kann, während eine beliebige (ausgefallene) Komponente repariert oder ausgetauscht wird.

[0024] In einer Ausführungsform kann das Serversystem so ausgelegt werden, dass es eine Verschaltungseinheit zwischen einer ersten Zwischenstecker-Karte und einer zweiten Zwischenstecker-Karte enthält. Die Verschaltungseinheit kann hier als eine Zwischenstecker-Verschaltungseinheit oder eine Vermittlungsstellen-Verschaltungseinheit bezeichnet werden. Die Zwischenstecker-Verschaltungseinheit kann eine Verkabelung zwischen einem Netzwerkadapter der ersten Zwischenstecker-Karte und einem Netzwerkadapter der zweiten Zwischenstecker-Karte enthalten, wobei sich die Verkabelung außerhalb der Mittelplatine befindet. Sollte also ein Schaltmodul der ersten Zwischenstecker-Karte ausfallen, kann ein Schaltmodul der zweiten Zwischenstecker-Karte den Netzwerkverkehr für eine Serverkarte im Namen der ausgefallenen Zwischenstecker-Karte verwalten - zusätzlich zu einer Serverkarte der zweiten Zwischenstecker-Karte. Durch die Auslegung des Serversystems in der Weise, dass es die Zwischenstecker-Verschaltungseinheit enthält, wird somit das Schaltmodul der ersten Zwischenstecker-Karte als ein SPOF beseitigt. Anders ausgedrückt, die Serverkarte der ersten Zwischenstecker-Karte behält die Verbindungen zu dem Schaltnetzwerk und/oder dessen Redundanz selbst dann bei, wenn das Schaltmodul der ersten Zwischenstecker-Karte ausfällt. Die Auslegung des Serversystems gemäß den hier beschriebenen Ausführungsformen verringert und/oder beseitigt sowohl SPORs als auch SPOFs.

[0025] In einer alternativen Ausführungsform wird durch die Auslegung des Serversystems in der Weise, dass es eine oder mehrere Schaltkarten enthält, die mit der Mittelplatine verbunden sind, das Schaltmodul als ein SPOF beseitigt. Die Mittelplatine enthält eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk. Die Schaltkarten sind mit einer ersten Seite der Mittelplatine verbunden, und eine oder mehrere Serverkarten sind mit einer zweiten Seite der Mittelplatine verbunden. Überdies können die Schaltkarten an einer ersten Achse ausgerichtet werden, und die Serverkarten können an einer zweiten Achse ausgerichtet werden. Die erste Achse und die zweite Achse können senkrecht zueinander liegen. Die Schaltkarten können beispielsweise waagerecht mit der ersten Seite der Mittelplatine verbunden sein und die Serverkarten können senkrecht mit der zweiten Seite der Mittelplatine verbunden sein oder umgekehrt. Das Schaltnetzwerk enthält eine Verdrahtung, die jede Schaltkarte mit jeder Serverkarte verbindet, und/oder eine Verdrahtung, die die Schaltkarten miteinander verbindet. Dies ermöglicht zum einen eine redundante Pfadführung, um SPORs und/oder SPOFs in dem Schaltnetzwerk zu verringern und/oder zu beseitigen, und verringert zum anderen den erforderlichen Gesamtverdrahtungsaufwand (zumindest in manchen Fällen). Vorteilhafterweise kann das Schaltmodul als ein SPOF beseitigt werden, ohne Zwischenstecker-Karten oder eine damit verbundene Verdrahtung erforderlich zu machen.

[0026] Im Folgenden wird Bezug auf Ausführungsformen der Erfindung genommen. Es sollte sich jedoch verstehen, dass die Erfindung nicht auf bestimmte beschriebene Ausführungsformen beschränkt ist. Vielmehr wird jede beliebige Kombination der folgenden Merkmale und Elemente, ungeachtet dessen, ob sie sich auf verschiedene Ausführungsformen beziehen oder nicht, als eine Kombination betrachtet, die die Erfindung realisiert und betreibt. Überdies können Ausführungsformen der Erfindung gegenüber anderen möglichen Lösungen und/oder gegenüber dem Stand der Technik zwar Vorteile erzielen, doch ist die Frage, ob ein bestimmter Vorteil von einer bestimmten Ausführungsform erzielt wird oder nicht, nicht als Einschränkung der Erfindung zu verstehen. Die folgenden Erscheinungs-

formen, Merkmale, Ausführungsformen und Vorteile dienen somit lediglich der Veranschaulichung und werden nicht als Elemente oder Einschränkungen der beigefügten Ansprüche betrachtet, soweit dies in einem oder in mehreren Ansprüchen nicht ausdrücklich anders angegeben ist. Ebenso ist die Bezugnahme auf „die Erfindung“ nicht als Verallgemeinerung eines beliebigen hier offengelegten Erfindungsgegenstands auszulegen und nicht als ein Element oder eine Beschränkung der beigefügten Ansprüche zu betrachten, soweit dies in einem oder in mehreren Ansprüchen nicht ausdrücklich anders angegeben ist.

[0027] Der Fachmann wird als vorteilhaft erkennen, dass Erscheinungsformen der vorliegenden Erfindung als ein System, ein Verfahren oder ein Rechnerprogrammprodukt realisiert werden können. Folglich können Erscheinungsformen der vorliegenden Erfindung die Form einer ganz in Hardware realisierten Ausführung, einer ganz in Software realisierten Ausführung (einschließlich Firmware, residenter Software, Mikrocode usw.) oder einer Ausführung annehmen, die Software- und Hardware-Erscheinungsformen kombiniert, die hier alle allgemein als eine „Schaltung“, ein „Modul“ oder ein „System“ bezeichnet werden können. Überdies können Erscheinungsformen der vorliegenden Erfindung die Form eines Rechnerprogrammprodukts annehmen, das sich auf einem oder mehreren rechnerlesbaren Datenträger (n) befindet, auf dem beziehungsweise denen sich rechnerlesbarer Programmcode befindet.

[0028] Jede beliebige Kombination aus einem oder mehreren rechnerlesbaren Datenträgern kann verwendet werden. Der rechnerlesbare Datenträger kann ein rechnerlesbarer Signaldatenträger oder ein rechnerlesbares Speichermedium sein. Ein rechnerlesbares Speichermedium kann zum Beispiel, ohne auf diese beschränkt zu sein, ein(e) elektronische(s), magnetische(s), optische(s), elektromagnetische(s), Infrarot- oder Halbleitersystem, -vorrichtung, -einheit oder eine beliebige geeignete Kombination des Vorstehenden sein. Zu konkreteren Beispielen (wobei die Liste keinen Anspruch auf Vollständigkeit erhebt) für das rechnerlesbare Speichermedium würden folgende gehören: eine elektrische Verbindung mit einer oder mehreren Leitungen, eine Diskette eines tragbaren Rechners, eine Festplatte, ein Direktzugriffsspeicher (RAM), ein Nur-Lese-Speicher (ROM), ein löschbarer programmierbarer Nur-Lese-Speicher (EPROM oder Flash-Speicher), ein Lichtwellenleiter, ein tragbarer Compact-Disk-Nur-Lese-Speicher (CD-ROM), eine optische Speichereinheit, eine magnetische Speichereinheit oder jede beliebige geeignete Kombination des Vorstehenden. Im Rahmen dieses Schriftstücks kann ein rechnerlesbares Speichermedium jedes physisch greifbare Medium sein, das ein Programm zur Verwendung durch ein Befehlsausführungssystem, eine Befehlsausführungsvorrichtung oder -einheit oder zur Verwendung in Verbindung mit einem Befehlsausführungssystem, einer Befehlsausführungsvorrichtung oder -einheit enthalten oder speichern kann.

[0029] Ein rechnerlesbarer Signaldatenträger kann ein übertragenes Datensignal mit einem darin enthaltenen rechnerlesbaren Programmcode, beispielsweise in einem Basisband oder als Teil einer Trägerwelle, enthalten. Solch ein übertragenes Signal kann eine beliebige einer Vielzahl von Formen einschließlich elektromagnetischer, optischer Formen oder jede beliebige geeignete Kombination dieser Formen, ohne auf diese beschränkt zu sein, annehmen. Bei einem rechnerlesbaren Signaldatenträger kann es sich um jeden beliebigen rechnerlesbaren Datenträger handeln, der kein rechnerlesbares Speichermedium ist und der ein Programm zur Verwendung durch oder zur Verwendung in Verbindung mit einem Befehlsausführungssystem, einer Befehlsausführungsvorrichtung oder -einheit übertragen, weiterleiten oder transportieren kann.

[0030] Auf einem rechnerlesbaren Datenträger enthaltener Programmcode kann mittels eines geeigneten Mediums einschließlich eines drahtlosen Mediums, eines drahtgebundenen Mediums, eines Lichtwellenleiterkabels, mittels Hochfrequenz (HF) usw., ohne auf diese beschränkt zu sein, oder mittels jeder beliebigen geeigneten Kombination des Vorstehenden übertragen werden.

[0031] Rechner-Programmcode zur Durchführung von Operationen für Erscheinungsformen der vorliegenden Erfindung kann in einer beliebigen Kombination aus einer oder mehreren Programmiersprachen einschließlich einer objektorientierten Programmiersprache, wie beispielsweise Java, Smalltalk, C++ oder dergleichen, sowie in herkömmlichen prozeduralen Programmiersprachen wie beispielsweise der Programmiersprache „C“ oder in ähnlichen Programmiersprachen geschrieben sein. Die Ausführung des Programmcodes kann vollständig auf dem Rechner des Benutzers, teilweise auf dem Rechner des Benutzers, als eigenständiges Software-Paket, teilweise auf dem Rechner des Benutzers und teilweise auf einem fernen Rechner oder vollständig auf dem fernen Rechner oder Server erfolgen. Im letzteren Szenario kann der ferne Rechner mit dem Rechner des Benutzers über jede beliebige Art eines Netzwerks einschließlich eines lokalen Netzwerks (LAN) oder eines Weitverkehrsnetzes (WAN) verbunden werden oder die Verbindung kann zu einem externen Rechner (zum Beispiel über das Internet mittels eines Internet-Diensteanbieters) hergestellt werden.

[0032] Erscheinungsformen der vorliegenden Erfindung werden nachstehend mit Bezug auf Darstellungen in Flussdiagrammen und/oder Blockschaltbildern von Verfahren, Vorrichtungen (Systemen) und Rechnerprogrammprodukten gemäß Ausführungsformen

der Erfindung beschrieben. Es versteht sich, dass jeder Block der Darstellungen in den Flussdiagrammen und/oder der Blockschaltbilder sowie Kombinationen aus Blöcken in den Darstellungen der Flussdiagramme und/oder den Blockschaltbildern mittels Rechnerprogrammbefehlen realisiert werden können. Diese Rechnerprogrammbefehle können einem Prozessor eines Universalrechners, eines Rechners für spezielle Anwendungen oder einer anderen programmierbaren Datenverarbeitungsvorrichtung bereitgestellt werden, um eine Maschine zu erzeugen, so dass die Befehle, die über den Prozessor des Rechners oder einer anderen programmierbaren Datenverarbeitungsvorrichtung ausgeführt werden, ein Mittel zur Ausführung der Funktionen/Vorgänge erzeugen, die in dem Flussdiagramm und/oder dem Block oder den Blöcken des Blockschaltbilds angegeben sind.

[0033] Diese Rechnerprogrammbefehle können auch auf einem rechnerlesbaren Datenträger gespeichert werden, der einen Rechner, eine andere programmierbare Datenverarbeitungsvorrichtung oder andere Einheiten anweisen kann, auf eine bestimmte Art und Weise zu funktionieren, so dass die auf dem rechnerlesbaren Datenträger gespeicherten Befehle einen Herstellungsgegenstand erzeugen, der Befehle enthält, die die Funktion/den Vorgang ausführen, welche beziehungsweise welcher in dem Flussdiagramm und/oder dem Block oder den Blöcken des Blockschaltbilds angegeben ist.

[0034] Die Rechnerprogrammbefehle können auch auf einen Rechner, eine andere programmierbare Datenverarbeitungsvorrichtung oder auf andere Einheiten geladen werden, um die Durchführung einer Reihe von Betriebsschritten auf dem Rechner, einer anderen programmierbaren Vorrichtung oder auf anderen Einheiten zu bewirken, um einen von einem Rechner ausgeführten Prozess zu erzeugen, so dass die Befehle, die auf dem Rechner oder einer anderen programmierbaren Vorrichtung ausgeführt werden, Prozesse zur Ausführung der Funktionen/Vorgänge ermöglichen, die in dem Flussdiagramm und/oder dem Block oder den Blöcken des Blockschaltbilds angegeben sind.

[0035] Die Flussdiagramme und die Blockschaltbilder in den Figuren zeigen die Architektur, die Funktionalität und die Betriebsweise von möglichen Ausführungsarten von Systemen, Verfahren und Rechnerprogrammprodukten gemäß verschiedenen Ausführungsformen der vorliegenden Erfindung. In dieser Hinsicht kann jeder Block in dem Flussdiagramm oder den Blockschaltbildern ein Modul, ein Segment oder einen Teil von Code darstellen, das beziehungsweise der einen oder mehrere ausführbare Befehle zur Ausführung der angegebenen logischen Funktion (en) umfasst. Es sei auch angemerkt, dass die in dem Block angegebenen Funktionen in manchen alternativen Ausführungsarten in einer anderen als in der in den Figuren angegebenen Reihenfolge auftreten können. In Abhängigkeit von der mit ihnen verbundenen Funktionalität können beispielsweise zwei Blöcke, die als aufeinanderfolgende Blöcke dargestellt sind, tatsächlich weitgehend gleichzeitig ausgeführt werden oder die Blöcke können manchmal in der umgekehrten Reihenfolge ausgeführt werden. Man wird auch feststellen, dass jeder Block der Blockschaltbilder/und oder der Darstellung in dem Flussdiagramm sowie Kombinationen aus Blöcken in den Blockschaltbildern und/oder der Darstellung in dem Flussdiagramm von Systemen, die auf Hardware für spezielle Anwendungen beruhen und die angegebenen Funktionen oder Vorgänge durchführen, oder von Kombinationen aus Hardware für spezielle Anwendungen und Rechnerbefehlen ausgeführt werden können.

[0036] **Fig. 1** ist ein Blockschaltbild einer Datenverarbeitungsumgebung **100** gemäß einer Ausführungsform der Erfindung, die über mehrere Hostrechner verfügt, welche Zugriff auf ein Serversystem **102** haben. Aus Gründen der Übersichtlichkeit der Darstellung sind zwar nur drei Hostrechner **134a**, **134b**, **134n** gezeigt, doch versteht der Fachmann, dass zusätzliche Hostrechner Zugriff auf das Serversystem **102** haben können. Die Hostrechner **134a**, **134b**, 134n sind über ein Kommunikationsnetzwerk **132** mit dem Serversystem **102** verbunden. In Abhängigkeit von der Ausführungsform kann jeder Hostrechner **134a**, **134b**, **134n** die Funktion eines Client übernehmen, der auf Funktionen zugreift, die von dem Serversystem **102** bereitgestellt werden, und/oder er kann die jeweiligen Server-Funktionen außerhalb des Serversystems **102** bereitstellen. Das Kommunikationsnetzwerk 132 kann ein Telekommunikationsnetzwerk und/oder ein Weitverkehrsnetz (WAN) sein. In einer bestimmten Ausführungsform ist das Kommunikationsnetzwerk **132** das Internet. Das Serversystem **102** enthält ein Gehäuse, in dem die Serverblades **104a**, **104b**, **104n** untergebracht sind. Die Serverblades **104a**, **104b**, **104n** sind mit einer Mittelplatine **123** verbunden, die mechanische und logische Verbindungen (z.B. den Austausch von Daten- und Steuersignalen) unter den Serverblades **104a**, **104b**, **104n** bereitstellt. Zwar sind drei Serverblades **104a**, **104b**, **104n** gezeigt, doch versteht der Fachmann, dass weitere Serverblades mit der Mittelplatine **123** verbunden werden können. Auch wenn Ausführungsformen hier mit Bezug auf Blade-Systeme beschrieben werden, werden allgemein auch andere Formfaktoren oder physische Konfigurationen (z.B. Gestelleinbausysteme) in Betracht gezogen.

[0037] Während Ausführungsformen hier mit Bezug auf die Serverblades 104a, 104b, 104b beschrieben werden, die mit der Mittelplatine 123 verbunden sind, erkennt der Fachmann außerdem, dass die Serverblades allgemeiner auch mit einer beliebigen Leiter-

platte (printed circuit board (PCB)) verbunden werden können, die als Zentralverbindung für das Gehäuse dient, wie zum Beispiel mit einer Rückwandplatine, einer Hauptplatine usw. Ferner erkennt der Fachmann, dass das Serversystem **102** in anderen Ausführungsformen mehrere Gehäuse enthalten kann, auch wenn Ausführungsformen hier mit Bezug auf das Serversystem 102, das ein einziges Gehäuse hat, beschrieben werden. In einer alternativen Ausführungsform kann das Serversystem **102** beispielsweise ein Blade-System sein, das mindestens zwei Blade-Gehäuse enthält, von denen jedes über eine Vielzahl von Blades verfügt.

[0038] In einer Ausführungsform enthält das Serversystem **102** darüber hinaus ein oder mehrere Verwaltungsmodule **124**. In der gezeigten Ausführungsform enthält das Serversystem **102** ein primäres Verwaltungsmodul **124a** und ein Sicherungs-Verwaltungsmodul **124b**. Jedes Verwaltungsmodul **124** kann mehrere Serverblades **104** verwalten. Während des normalen Betriebs ist eines der Verwaltungsmodule **124** mit den Serverblades **104** über ein lokales Netzwerk (LAN) **122**, die Mittelplatine **123** und Baseboard-Management Controllers (BMCs) **110** eines jeden Serverblades **104** betriebsfähig verbunden, um einen In-Band-Verwaltungspfad zu bilden. In einer Ausführungsform dient das Kommunikationsnetzwerk **132** als eine Erweiterung des LAN **122**. Das LAN **122** und der BMC **110** werden nachstehend ausführlicher erörtert.

[0039] In einer Ausführungsform wird die Mittelplatine **123** in der Mitte des Gehäuses des Serversystems **102** befestigt und enthält Schaltungen und Sockel **112**, in die zusätzliche elektronische Einheiten oder Karten einschließlich der Serverblades **104** gesteckt werden können. Die Mittelplatine **123** enthält mindestens einen Bus für die sichere, interne In-Band-Kommunikation über die BMCs **110** sowie zwischen den Verwaltungsmodulen **124** und den Serverblades **104** und/oder unter den Serverblades **104** selbst.

[0040] In einer Ausführungsform wird, wenn ein Serverblade **104** in einen bestimmten Sockel **112** gesteckt wird, für den Serverblade 104 eine physische Adresse festgelegt. Nehmen wir beispielsweise an, dass der Serverblade **104a** in den Sockel 112a gesteckt wird. In einer Ausführungsform erkennt die Steuerlogik **116a** das Vorhandensein des Serverblades **104a** in dem Sockel **112a**. Die Steuerlogik **116a** kann dem RS485-Standard der Electronics Industry Association (EIA) für die Datenübertragung entsprechen. In anderen Ausführungsformen kann die Steuerlogik **116a** dem Inter-IC-(Inter-Integrated-Circuit- oder I^2C-) Standard von Phillips oder einem Ethernet-Netzwerk-Standard entsprechen. Die Steuerlogik **116a**, die in Verbindung mit dem Verwaltungsmodul **124a** arbeitet, weist dem Serverblade **104a** als Reaktion auf das Einstecken des Serverblades **104a** in den Sockel **112a** eine physische Adresse auf einem Bus in der Mittelplatine **123** zu. Wie gezeigt ist, ist jeder Serverblade **104** einer jeweiligen Steuerlogik **116** zugeordnet, die mit der Mittelplatine **123** betriebsfähig verbunden ist. In einer alternativen Ausführungsform können mehrere Serverblades **104** eine einzige Steuerlogik **116** gemeinsam benutzen.

[0041] In einer Ausführungsform wird jedem Serverblade **104** eine eindeutige Internet-Protocol-(IP-) Adresse auf der Mittelplatine **123** zugewiesen. Das heißt, die Mittelplatine **123** kann die untereinander stattfindende Kommunikation unter Verwendung des IP-Adressierungsprotokolls unterstützen, bei dem jede Einheit, die mit der Mittelplatine **123** betriebsfähig verbunden ist, eine IP-Adresse hat, die von Logik (nicht gezeigt) zugewiesen wird, welche sich entweder innerhalb oder außerhalb des Gehäuses des Serversystems **102** befindet. Ein Dynamic-Host-Configuration-Protocol-(DHCP-)Server kann beispielsweise verwendet werden, um dem Serverblade **104a** eine IP-Adresse zuzuweisen. Der Datenaustausch mit dem Serverblade 104a findet daraufhin über eine Netzwerk-Schnittstellensteuereinheit (NIC) **114a** statt, die dem Serverblade **104a** zugeordnet ist. Bei der NIC **114a** kann es sich um jeden beliebigen Typ einer Netzwerk-Übertragungseinheit handeln, die es dem Serverblade **104a** ermöglicht, mit anderen Serverblades **104b**, **104n** und/oder Rechnern über das LAN **122** und/oder das Kommunikationsnetzwerk **132** zu kommunizieren.

[0042] In einer Ausführungsform ist ein integriertes Modul **126a** mit der NIC **114a** betriebsfähig verbunden. Das integrierte Modul 126a kann paarweise (z.B. mit dem integrierten Modul **126b**) verwendet werden, um Redundanz vorzusehen. Wie bekannt ist, bezieht sich Small Computer System Interface (SCSI) auf eine Reihe von Standards, um Rechner und periphere Einheiten physisch zu verbinden und Daten zwischen ihnen zu übertragen. In einer Ausführungsform enthalten die integrierten Module **126** Schaltmodule **128**, wie zum Beispiel ein Serial-Attached-SCSI-(SAS-)Schaltmodul. Die Schaltmodule **128** stellen für die Serverblades **104** Verbindungen zu Ethernet, Fibre Channel over Ethernet (FCoE), SAS usw. bereit. In einer Ausführungsform ist jedes Schaltmodul **128** ein Vermittlungsstellen-Chip. In Abhängigkeit von der Ausführungsform können die integrierten Module **126** des Weiteren Steuereinheiten **130** für eine redundante Anordnung unabhängiger Festplatten (redundant array of independent disks (RAID)) enthalten. Jede RAID-Steuereinheit **130** ist mit RAID-Einheiten, wie zum Beispiel Speichereinheiten, in einer RAID-Konfiguration verbunden. Die RAID-Einheiten können sich in einem oder in mehreren der Serverblades **104** befinden. Die RAID-Steuereinheiten **130** und die RAID-Einheiten können zusammen als ein RAID-Subsystem des Serversystems **102** betrachtet werden.

[0043] In einer Ausführungsform kann jede Speichereinheit eine Permanentspeichereinheit sein. Jede Speichereinheit kann außerdem eine Kombination aus festen und/oder auswechselbaren Speichereinheiten wie zum Beispiel Festplattenlaufwerken, Diskettenlaufwerken, Bandlaufwerken, auswechselbaren Speicherkarten, Halbleiterlaufwerken oder optischen Speichern sein. Der Arbeitsspeicher **108** und die Speichereinheit können Teil von einem einzelnen virtuellen Adressraum sein, der sich über mehrere primäre und sekundäre Speichereinheiten erstreckt.

[0044] In einer Ausführungsform kann jeder Serverblade **104** mindestens eine Zentraleinheit (CPU) **106** und einen Arbeitsspeicher **108** haben. Die CPU **106** wurde stellvertretend für eine einzelne CPU, mehrere CPUs, eine einzelne CPU mit mehreren Verarbeitungskernen und dergleichen aufgenommen. Ebenso kann der Arbeitsspeicher **108** ein Direktzugriffspeicher sein. Während der Arbeitsspeicher **108** als eine einzelne Identität gezeigt ist, sollte es sich verstehen, dass der Arbeitsspeicher **108** eine Vielzahl von Modulen umfassen kann und dass der Arbeitsspeicher **108** auf mehreren Ebenen, von Hochgeschwindigkeitsregistern und Cachespeichern bis hin zu langsameren, aber größeren DRAM-Chips, vorhanden sein kann. Der Arbeitsspeicher **108** kann ein Nur-Lese-Flash-Speicher („Flash-ROM“ oder „Flash-Memory“) sein, bei dem der Inhalt der einzelnen Speichereinheiten, die als „Blöcke“ bezeichnet werden, gelöscht und diese neu programmiert werden können. Der Arbeitsspeicher **108** kann auch einen nicht flüchtigen elektrisch löschbaren programmierbaren Nur-Lese-Speicher (Electrically Erasable Programmable Read Only Memory (EEPROM)) enthalten, der ähnlich dem Flash-Speicher ist, außer dass der EEPROM auf Byte-Ebene gelöscht und neu beschrieben wird und gewöhnlich über eine geringere Kapazität verfügt. Jeder Serverblade **104** kann als ein Prozessor-Blade oder als ein Speicher-Blade ausgestaltet werden. Ein Prozessor-Blade enthält eine oder mehrere Verarbeitungseinheiten, während ein Speicher-Blade mehrere integrierte Speichereinheiten wie zum Beispiel Plattenlaufwerke enthält.

[0045] In einer Ausführungsform kann der Arbeitsspeicher **108**, wenn der Serverblade **104** von einem Hersteller versendet wird, vorab mit Firmware gebrannt werden, so dass er ein grundlegendes Eingabe-/Ausgabesystem (BIOS) sowie Software zur Überwachung des Serverblades **104** enthält. Die Überwachung kann die Steuerung der Speichereinheiten, die Überwachung und Regelung von Spannungen im ganzen System, die Feststellung des Einschaltstatus des Serverblades **104**, die Anforderung des Zugriffs auf eine(n) gemeinsam benutzte(n) Tastatur, Bildschirm, Maus, Compact-Disk-Nur-Lese-Speicher (CD-ROM) und/oder Diskettenlaufwerke, die Überwachung des Betriebssystems (BS), das auf dem Serverblade **104** ausgeführt wird, usw. beinhalten. Zu Beispielen für Betriebssysteme gehören UNIX, Versionen des Betriebssystems Windows® von Microsoft und Ausgaben des Betriebssystems Linux®. Allgemeiner gesagt, jedes Betriebssystem, das die hier offengelegten Funktionen unterstützt, kann verwendet werden.

[0046] In einer Ausführungsform können die Verwaltungsmodule **124** das Vorhandensein, die Menge, den Typ und den Revisionsstand eines jeden Serverblades **104**, des Leistungsmoduls **118** und der Mittelplatine **123** in dem System feststellen. Die Verwaltungsmodule **124** können auch den Betrieb eines jeden Serverblades **104** und des Leistungsmoduls **118** direkt steuern.

[0047] Die Verwaltungsmodule **124** können auch den Betrieb der Lüfter 120 und anderer Komponenten in dem Gehäuse des Serversystems 102 direkt steuern. Eine direkte Steuerung des Betriebs bedeutet die Steuerung des Betriebs ohne Verwendung des BIOS in den Serverblades **104**. In einer alternativen Ausführungsform können die Verwaltungsmodule **124** das BIOS zur indirekten Steuerung des Betriebs der Lüfter **120** und der anderen Komponenten in dem Gehäuse des Serversystems **102** verwenden.

[0048] In einer Ausführungsform enthält jeder Serverblade **104** einen Baseboard Management Controller (BMC) **110**, der eine lokale vorrangige Steuerung des Serverblades **104**, dem der BMC **110** zugeordnet ist, bereitstellt. Jeder BMC **110** ist so konfiguriert, dass er mit einem Verwaltungsmodul **124** kommuniziert, indem er entweder den Kommunikationspfad des LAN 122 verwendet (d.h. über ein In-Band-Netzwerk) oder indem er alternativ die Schaltmodule **128** und die NICs **114** verwendet (d.h. über ein Außerband-Netzwerk). Die Verwaltungsmodule **124** können viele verschiedene Kommunikationspfade in dem LAN **122**, wie zum Beispiel den RS485-Pfad, einen LAN-Pfad und einen I^2C-Pfad, verwenden, um mit jedem Serverblade **104** zu kommunizieren.

[0049] In einer Ausführungsform ist das LAN **240** ein In-Band-Netzwerk, das auch dem RS485-Standard der Electronics Industry Association (EIA) für die Datenübertragung entspricht. Die Verwaltungsmodule **124** - z.B. entweder das primäre Verwaltungsmodul **124a** oder das Sicherungs-Verwaltungsmodul 124b, wenn das primäre Verwaltungsmodul ausgefallen ist, - kommunizieren über das LAN **122** mit dem BMC **110**, der Logik enthält, um die Kommunikation mit den Serverblades **104** über die Sockel **112** zu koordinieren.

[0050] In einer Ausführungsform kann das LAN **122** so konfiguriert werden, dass es Übertragungen zwischen den Serverblades **104** und den Verwaltungsmodulen **124** zulässt, die sich auf Einstellungen des fernen BIOS und die Verwaltung des BIOS bezie-

hen. Die Serverblades **104** können die BMCs **110** als Proxies verwenden, um mit den Verwaltungsmodulen **124** über das RS485-Protokoll zu kommunizieren. Ebenso können die Verwaltungsmodule die BMCs **110** als Proxies verwenden, um mit den Serverblades **104** über das RS485-Protokoll zu kommunizieren. In einer alternativen Ausführungsform kann eine RS485-Verbindung zwischen jedem Serverblade **104** und den Verwaltungsmodulen **124** getrennt hergestellt werden. Überdies können andere Kommunikationsprotokolle und Pfade über die Schaltmodule **128**, wie zum Beispiel I^2C, TCP/IP, Ethernet, FCoE usw. verwendet werden.

[0051] In Abhängigkeit von der Ausführungsform kann das Serversystem 102 auch mit einer Eingabeeinheit und/oder einer Ausgabeeinheit betriebsfähig verbunden sein. Die Eingabeeinheit kann eine beliebige Einheit sein, die dazu dient, dem Serversystem **102** eine Eingabe bereitzustellen. Beispielsweise können eine Tastatur, ein Tastenblock, ein Lichtgriffel, ein berührungsempfindlicher Bildschirm, eine Rollkugel oder eine Spracherkennungseinheit, ein Audio-/Video-Wiedergabegerät und dergleichen verwendet werden. Die Ausgabeeinheit kann eine beliebige Einheit sein, die dazu dient, einem Benutzer des Serversystems **102** eine Ausgabe bereitzustellen. Die Ausgabeeinheit kann zum Beispiel ein herkömmlicher Bildschirm oder ein Satz von Lautsprechern zusammen mit ihren jeweiligen Schnittstellenkarten, d.h. Grafikkarten und Soundkarten, sein. Die Eingabeeinheit und die Ausgabeeinheit können überdies kombiniert sein. Ein Bildschirm mit einer integrierten berührungsempfindlichen Bildschirmoberfläche, ein Bildschirm mit einer integrierten Tastatur oder eine Spracherkennungseinheit, die mit einem Text-in-Sprache-Wandler kombiniert ist, können beispielsweise verwendet werden.

[0052] **Fig. 2** zeigt eine Konfiguration **200**, bei der gemäß einer Ausführungsform der Erfindung Zwischenstecker-Karten mit Serverkarten in einem Serversystem betriebsfähig verbunden sind. In Abhängigkeit von der Ausführungsform kann die Zwischenstecker-Karte auch mit den Serverkarten und/oder der Mittelplatine verbunden sein. Wie vorstehend beschrieben wurde, ist das Serversystem in einer Ausführungsform so konfiguriert, dass es eine Mittelplatine **123** und Serverkarten 202 enthält, wobei die Mittelplatine und die Serverkarten **202** über eine oder mehrere Zwischenstecker-Karten **204** betriebsfähig verbunden sind, um die Auswirkung eines Reparaturvorgangs an einem ausgefallenen Schaltmodul zu verringern. Jede Serverkarte **202** kann einem Serverblade **104** entsprechen. Jede Zwischenstecker-Karte **204** kann im laufenden Betrieb ausgetauscht werden und enthält ein oder mehrere Schaltmodule **206**. Die Schaltmodule **206** schalten den Netzwerkverkehr für eine oder mehrere Serverkarten **202**, die mit der jeweiligen Zwischenstecker-Karte betriebsfähig verbunden sind.

[0053] In einer Ausführungsform ist das Serversystem so konfiguriert, dass es den Ausfall des Schaltmoduls **206** feststellt. Nach dem Feststellen des Ausfalls des Schaltmoduls **206** kann das Serversystem einen Hinweis zur Anzeige ausgeben, dass ein Reparaturvorgang an dem Schaltmodul **206** durchzuführen ist. Der Hinweis kann beispielsweise zur Anzeige in einem Fenster einer grafischen Benutzeroberfläche (GUI) oder in Form von einer Warnmeldung, die per E-Mail an einen Benutzer zu senden ist, ausgegeben werden. In Abhängigkeit von der Ausführungsform kann das Feststellen und/oder das Ausgeben von einer beliebigen Komponente des Serversystems, wie zum Beispiel den Serverkarten **202**, dem Schaltmodul **206** und/oder Firmware, die in dem Serversystem enthalten ist, vorgenommen werden. In einer Ausführungsform enthält das Serversystem zum Beispiel Verwaltungs-Firmware, die den fehlerbeziehungsweise störungsfreien Betrieb des Serversystems überwacht und einen Ausfall des Schaltmoduls **206** feststellt.

[0054] Wenn ein Schaltmodul **206** ausfällt, kann die Zwischenstecker-Karte **204**, die das Schaltmodul **206** enthält, folglich durch eine Zwischenstecker-Karte ersetzt werden, die über ein funktionsfähiges Schaltmodul verfügt. Aufgrund der Möglichkeit, die Zwischenstecker-Karten, die Serverkarten und/oder die Mittelplatine im laufenden Betrieb auszutauschen, kann die Zwischenstecker-Karte **204** ferner ausgetauscht werden, ohne dass das Serversystem und/oder das Schaltnetzwerk abgeschaltet oder neu gestartet werden muss. Die Zwischenstecker-Karte, die über ein funktionsfähiges Schaltmodul verfügt, kann dann über das Konfigurations-Werkzeug wieder in das Schaltnetzwerk integriert werden. In Abhängigkeit von der Ausführungsform kann das Konfigurations-Werkzeug auf dem Serversystem oder auf einem anderen Rechner ausgeführt werden, der über das Kommunikationsnetzwerk **132** mit dem Serversystem verbunden ist.

[0055] Während der Dauer des Austauschs der Zwischenstecker-Karte mit dem ausgefallenen Schaltmodul können folglich nur das ausgefallene Schaltmodul und zugehörige Serverkarten nicht von dem Kommunikationsnetzwerk erreicht werden. Während der Dauer des Austauschs der Zwischenstecker-Karte mit dem ausgefallenen Schaltmodul bleiben andere Schaltmodule und/oder Serverkarten, die mit der Mittelplatine betriebsfähig verbunden sind, erreichbar. Somit ist die Auswirkung des Reparaturvorgangs an dem ausgefallenen Schaltmodul örtlich auf die Serverkarten begrenzt, die zu dem ausgefallenen Schaltmodul gehören. Anders ausgedrückt, die einzigen Netzwerkknoten, die während des Reparaturvorgangs nicht von dem Kommunikationsnetzwerk er-

reicht werden können, sind die Netzwerkknoten, die zu den Serverkarten gehören, welche mit dem ausgefallenen Schaltmodul betriebsfähig verbunden sind.

[0056] Fig. 3 zeigt eine Konfiguration **300**, bei der gemäß einer Ausführungsform der Erfindung eine Zwischenstecker-Karte mit zwei Serverkarten in einem Serversystem betriebsfähig verbunden ist. Wie gezeigt ist, enthält die Zwischenstecker-Karte **204** das Schaltmodul **206** und zwei Converged Network Adapter (CNAs) **302**. Die beiden Serverkarten **202** können jeweils auch zwei CPUs **106** und einen CNA **304** enthalten. In einer Ausführungsform lassen sich die CPUs **106** mit FCoE über CNAs verbinden, die sowohl über Funktionen eines Fibre-Channel-Host-Bus-Adapters (HBA) als auch über Funktionen einer Ethernet-NIC verfügen. Die CNAs können einen oder mehrere physische Ethernet-Anschlüsse enthalten und so konfiguriert sein, dass sie die CPUs **106** von der Rahmenverarbeitung der unteren Ebene und/oder von Funktionen des SCSI-Protokolls, die üblicherweise von den Fibre-Channel-Host-Bus-Adaptern durchgeführt werden, entlasten. Wie vorstehend beschrieben wurde, schaltet das Schaltmodul **206** den Netzwerkverkehr für die Serverkarten **202**. Wenn das Schaltmodul **206** ausfällt, ermöglicht die Konfiguration **300** den Austausch des Schaltmoduls **206**, ohne dass Serverkarten von anderen Zwischenstecker-Karten, die mit der Mittelplatine verbunden sind, neu gestartet werden müssen.

[0057] Fig. 4 zeigt ein Serversystem **400**, das so konfiguriert ist, dass es gemäß einer Ausführungsform der Erfindung die Auswirkung eines Reparaturvorgangs an einem Schaltmodul verringert. Wie gezeigt ist, enthält das Serversystem **400** einen logischen Server **402**, der über ein Informationstechnologie-Element (ITE) **404** in Form eines Prozessors und ein E/A-ITE 406 konfiguriert wird. In der hier verwendeten Weise bezieht sich ein ITE allgemein auf jedes Betriebsmittel, das so konfiguriert ist, dass es sich mit der Mittelplatine **123** betriebsfähig verbinden lässt. In einer alternativen Ausführungsform kann der logische Server **402** auch über ein Speicher-ITE **408** konfiguriert werden. Das E/A-ITE 406 und das Speicher-ITE **408** sind so konfiguriert, dass sie einem oder mehreren Prozessor-ITEs zusätzliche E/A-Kapazität beziehungsweise Speicherkapazität zur Verfügung stellen. In Abhängigkeit von der Ausführungsform kann jedes ITE **404**, **406**, 408 als Teil von einem oder mehreren Serverblades integriert oder mit der Mittelplatine **123** als eigenständige Karte verbunden werden. Das Prozessor-ITE **404** enthält eine oder mehrere virtuelle Maschinen **410**, einen Hypervisor **412**, einen Arbeitsspeicher **414**, Prozessoren **416** und Festplattenlaufwerke 418. Das E/A-ITE 406 enthält eine gemeinsam benutzte E/A-ITE-Komponente 422 und E/A-Adapter 424, während das Speicher-ITE 408 eine gemeinsam benutzte Speicher-ITE-Komponente **426** und Halbleiterlaufwerke **428** enthält.

[0058] In einer Ausführungsform enthalten das Serverblade **404** und die ITEs **406**, **408** darüber hinaus jeweils ein Schaltmodul **206**. Jedes Schaltmodul **206** kann ein Vermittlungsstellen-Chip sein und in eine Zwischenstecker-Karte (nicht gezeigt) aufgenommen werden, die zwischen der Mittelplatine **123** und dem Prozessor-ITE **404** und/oder dem ITE **406**, ITE **408** angeordnet wird. Zusammen stellen die Schaltmodule **206** ein Schaltnetzwerk **432** bereit. Ein Ausfall des Schaltmoduls 206_1 des Prozessor-ITEs 404 - das durch ein X-Symbol 430 gekennzeichnet ist - wirkt sich nur auf das Prozessor-ITE **404** und nicht auf andere ITEs aus, die mit der Mittelplatine **123** betriebsfähig verbunden sind. Folglich bleiben andere logische Server, die über das E/A-ITE 406 und/oder das Speicher-ITE **408** konfiguriert werden, betriebsfähig, und die Verbindungen in dem Schaltnetzwerk **432** bleiben größtenteils betriebsfähig - d.h. mit Ausnahme der Verbindungen zu dem Prozessor-ITE **404**. In Abhängigkeit von der Ausführungsform können die Verbindungen in dem Schaltnetzwerk 432, die betriebsfähig bleiben, auch redundante Verbindungen in dem Schaltnetzwerk **432** beinhalten. Das Schaltmodul 206_1 kann darüber hinaus ausgetauscht werden, ohne dass dies Auswirkungen auf die anderen ITEs, logischen Server und/oder das Schaltnetzwerk **432** hat. Folglich kann die Verfügbarkeit des Schaltnetzwerks **432** verbessert werden.

[0059] Fig. 5 zeigt auch ein Serversystem **500**, das so konfiguriert ist, dass es gemäß einer Ausführungsform der Erfindung die Auswirkung eines Reparaturvorgangs an einem Schaltmodul verringert. Wie gezeigt ist, enthält das Serversystem **500** einen ersten logischen Server **402**, der über ein erstes Prozessor-ITE **404** und das E/A-ITE 406 konfiguriert wird. Das Serversystem **500** enthält auch einen zweiten logischen Server 502, der über ein zweites Prozessor-ITE **504** und das E/A-ITE 406 konfiguriert wird. In einer alternativen Ausführungsform kann der erste logische Server **402** und/oder der zweite logische Server **502** auch über das Speicher-ITE **408** konfiguriert werden. Jedes der ITEs **404**, **406**, **408**, **504** enthält ein Schaltmodul **206**. Ein Ausfallen des Schaltmoduls 206_1 des Prozessor-ITEs **404** - welches durch ein X-Symbol 506 gekennzeichnet ist - wirkt sich nur auf das erste Prozessor-ITE **404** und nicht auf das zweite Prozessor-ITE **504** aus. Somit bleiben das zweite Prozessor-ITE **504**, das E/A-ITE 406 und das Speicher-ITE **408** betriebsfähig und behalten die Verbindungen zum Schaltnetzwerk während des Ausfalls und/oder des Austauschs des Schaltmoduls 206_1 bei.

[0060] Fig. 6 zeigt ein Schaltnetzwerk **432** für ein Serversystem gemäß einer Ausführungsform der Erfindung. Wie gezeigt ist, enthält das Schaltnetzwerk **432** eine Vielzahl von Schaltmodulen 206, von denen jedes in einer jeweiligen Zwischenstecker-Karte 204 enthalten ist. Jede Zwischenstecker-Karte **204** verbindet zwei Serverkarten **202** betriebsfähig mit

dem Schaltnetzwerk 432. In Abhängigkeit von der Ausführungsform können die Schaltmodule in dem Schaltnetzwerk über eine in einer Mittelplatine untergebrachte Verdrahtung, eine außerhalb der Mittelplatine befindliche Verkabelung oder eine Kombination aus beidem miteinander verbunden werden. Darüber hinaus können ein oder mehrere der Schaltmodule **206** mit anderen Betriebsmitteln **602** als den Serverkarten **202** verbunden werden. Zu Beispielen von Betriebsmitteln **602** gehören Netzwerk-Betriebsmittel, Speicher-Betriebsmittel und E/A-Betriebsmittel. Folglich wirkt sich ein Ausfall und/oder ein Austausch eines Schaltmoduls **206** nur auf die Serverkarten **202** aus, die mit dem Schaltmodul **206** verbunden sind, ansonsten aber nicht auf den Rest des Schaltnetzwerks **432** und/oder die anderen Serverkarten.

[0061] Fig. 7 zeigt ein Serversystem **700** gemäß einer Ausführungsform der Erfindung, das über eine Mittelplatine **123** verfügt, die mit einer Vielzahl von Zwischenstecker-Karten **204** verbunden ist. Jede Zwischenstecker-Karte **204** enthält ein Schaltmodul 206 und verbindet eine oder zwei Serverkarten **202** betriebsfähig mit der Mittelplatine **123**. Die Mittelplatine enthält eine Netzwerkverdrahtung, die die Schaltmodule **206** verbindet, um ein Schaltnetzwerk zu bilden. Das Serversystem 700 ist so konfiguriert, dass die Serverkarten **202** im laufenden Betrieb von den Zwischenstecker-Karten **204** aus ausgetauscht werden können. Ferner ist das Serversystem **700** so konfiguriert, dass die Zwischenstecker-Karten **204** im laufenden Betrieb von der Mittelplatine **123** aus ausgetauscht werden können. Somit ermöglichen die Auslegung des Serversystems und die Möglichkeit, das Serversystem **700** im laufenden Betrieb auszutauschen, den Austausch eines fehlerhaften Schaltmoduls 206, während die Auswirkung auf das Serversystem **700** und/oder das Schaltnetzwerk so klein wie möglich gehalten oder verringert wird.

[0062] Fig. 8 zeigt ein Serversystem **800** gemäß einer Ausführungsform der Erfindung, das mehrere Gestellrahmen **804** enthält. Jeder Gestellrahmen enthält ein oder mehrere Gehäuse **802**, die über Gehäuse-Verbindungskarten **806** und eine zugehörige Verkabelung 808 betriebsfähig verbunden sind. Außerdem kann das Gehäuse 802 mittels Gestellrahmen-Verbindungskarten **810** und einer zugehörigen Verkabelung **812** über zwei Gestellrahmen betriebsfähig verbunden werden. In jedem Gehäuse **802** ist eine Mittelplatine **123** untergebracht, die gemäß den hier offengelegten Methoden mit einer oder zwei Serverkarten **202** über eine Zwischenstecker-Karte **204** verbunden ist. Jede Zwischenstecker-Karte **204** enthält ein Schaltmodul **206**, um den Verkehr für die Serverkarten **202** zu schalten. Die Mittelplatinen **123** enthalten eine Netzwerkverdrahtung, um die Schaltmodule **206** untereinander zu verbinden. Zusammen bilden die Schaltmodule **206**, die Netzwerkverdrahtung, die Gehäuse-Verbindungskarten **806** und die zugehörige Verkabelung **808** sowie die Gestellrahmen-Verbindungskarten **810** und die zugehörige Verkabelung **812** ein Schaltnetzwerk für das Serversystem **800**. Anders ausgedrückt, das Schaltnetzwerk für das Serversystem 800 enthält gehäuseübergreifende und gestellrahmenübergreifende Verschaltungseinheiten. Folglich entfernt ein Ausfall und/oder ein Austausch eines Schaltmoduls 206 - das durch ein X-Symbol 814 gekennzeichnet ist - lediglich eine zugehörige Serverkarte **202** aus dem Schaltnetzwerk. Der Betrieb der anderen Serverkarten und/oder der Verbindungen des Schaltnetzwerks wird während des Ausfalls und/oder des Austauschs des Schaltmoduls **206** dadurch aufrechterhalten.

[0063] Fig. 9 zeigt ein Serversystem **900** gemäß einer Ausführungsform der Erfindung, das mehrere Gestellrahmen enthält, wobei jeder Gestellrahmen über vier Gehäuse **802** verfügt. Die Gehäuse in jedem Gestellrahmen können über eine Verkabelung **904** zwischen den Gehäusen betriebsfähig verbunden werden. Gehäuse von verschiedenen Gestellrahmen können über eine Verkabelung **906** zwischen den Gestellrahmen betriebsfähig verbunden werden. In jedem Gehäuse ist eine Mittelplatine untergebracht, die gemäß den hier offengelegten Methoden über eine Netzwerkverdrahtung, mindestens eine Zwischenstecker-Karte, die ein Schaltmodul hat, und mindestens eine Serverkarte verfügt. Zusammen bilden die Schaltmodule, die Netzwerkverdrahtung, die Verkabelung **904** zwischen den Gehäusen und die Verkabelung **906** zwischen den Gestellrahmen sowie jedwede zugehörigen Verbindungskarten ein Schaltnetzwerk für das Serversystem **900**. Das Serversystem **900** wird dabei so konfiguriert, dass es die Verfügbarkeit des Schaltnetzwerks und/oder des Serversystems **900** während des Ausfalls und/oder des Austauschs eines Schaltmoduls erhöht.

[0064] Fig. 10 zeigt ein Serversystem **1000** gemäß einer Ausführungsform der Erfindung, das so ausgelegt ist, dass es eine Zwischenstecker-Verschaltungseinheit enthält. Wie vorstehend beschrieben wurde, kann die Auslegung des Serversystems **1000** eine Zwischenstecker-Verschaltungseinheit 1006 zwischen mindestens einer ersten Zwischenstecker-Karte und einer zweiten Zwischenstecker-Karte beinhalten. Und jede Zwischenstecker-Karte enthält ein Schaltmodul **206**, das mit zwei Serverkarten **202** verbunden ist. Die Zwischenstecker-Verschaltungseinheit **1006** kann eine Verkabelung zwischen einem Netzwerkadapter der ersten Zwischenstecker-Karte und einem Netzwerkadapter der zweiten Zwischenstecker-Karte enthalten.

[0065] Eine solche Verkabelung kann sich außerhalb der Mittelplatine befinden. Jeder Netzwerkadapter kann ein CNA **302** der jeweiligen Zwischenstecker-Karte oder ein CNA **304** der jeweiligen Serverkarte sein. Das Serversystem **1000** kann auch Schalt-

module **1002** enthalten, die Verbindungen zu externen Serversystemen und/oder Speicher-Steuereinheiten bereitstellen. Die Schaltmodule **1002** können mit den Schaltmodulen **206** über einen oder mehrere CNAs **1004** betriebsfähig verbunden sein.

[0066] Wenn ein Schaltmodul 206_1 der ersten Zwischenstecker-Karte ausfällt, kann ein Schaltmodul 206_2 der zweiten Zwischenstecker-Karte - neben dem Schalten des Netzwerkverkehrs für die Serverkarten 202_3, 202_4 der zweiten Zwischenstecker-Karte - folglich den Netzwerkverkehr für die Serverkarten 202_1, 202_2 der ersten Zwischenstecker-Karte schalten. Durch die Auslegung des Serversystems in der Weise, dass es die Zwischenstecker-Verschaltungseinheit **1006** enthält, wird das Schaltmodul 206_1 der ersten Zwischenstecker-Karte als ein SPOF beseitigt. Anders ausgedrückt, die Serverkarten 202_1, 202_2 der ersten Zwischenstecker-Karte behalten die Verbindungen zu dem Schaltnetzwerk bei und/oder erhalten dessen Redundanz selbst nach einem Ausfall des Schaltmoduls 206_1 der ersten Zwischenstecker-Karte aufrecht.

[0067] In einer Ausführungsform kann die Zwischenstecker-Verschaltungseinheit **1006** des Weiteren eine Verkabelung zwischen CNAs 302_3, 302_4 der zweiten Zwischenstecker-Karte und dem Schaltmodul 206_1 der ersten Zwischenstecker-Karte enthalten. Dabei wird das Schaltmodul 206_2 der zweiten Zwischenstecker-Karte als ein SPOF beseitigt - zusätzlich zur Beseitigung des Schaltmoduls 206_1 als ein SPOF. Folglich werden sowohl das Schaltmodul 206_1 der ersten Zwischenstecker-Karte als auch das Schaltmodul 206_2 der zweiten Zwischenstecker-Karte als SPOFs beseitigt.

[0068] In einer Ausführungsform ist der CNA 302_4 über eine Verkabelung mit dem Schaltmodul 206_1 verbunden und sieht Redundanz für die zweite Zwischenstecker-Karte vor. Wenn das Schaltmodul 206_2 der zweiten Zwischenstecker-Karte ausfällt, kann das Schaltmodul 206_1 der ersten Zwischenstecker-Karte - neben dem Schalten des Netzwerkverkehrs für die Serverkarten 202_1, 202_2 der ersten Zwischenstecker-Karte - folglich den Netzwerkverkehr für die Serverkarten 202_3, 202_4 der zweiten Zwischenstecker-Karte schalten.

[0069] Allgemeiner gesagt, durch die Auslegung des Serversystems **1000** in der Weise, dass es die Zwischenstecker-Verschaltungseinheit 1006 zwischen Paaren von Zwischenstecker-Karten enthält, werden die Schaltmodule einer jeden Zwischenstecker-Karte als ein SPOF beseitigt. Jedes Paar Zwischenstecker-Karten kann zwei Zwischenstecker-Karten enthalten, die entsprechend einer vorher festgelegten Achse in einem Gehäuse des Serversystems 1000 nebeneinander angeordnet sind. In einer alternativen Ausführungsform befindet sich jedes Paar Zwischenstecker-Karten in einem einzelnen Gestell in dem Serversystem **1000**. Die vorher festgelegte Achse kann eine x-Achse, eine y-Achse, eine z-Achse oder eine beliebige andere Achse, die zum Beschreiben von relativen Positionen der Zwischenstecker-Karten in dem Gehäuse des Serversystems **1000** geeignet ist, beinhalten.

[0070] Zwar werden Ausführungsformen hier mit Bezug auf Paare von Zwischenstecker-Karten beschrieben, die miteinander verbunden sind, doch werden allgemein auch andere Ausführungsformen erwogen. In einer alternativen Ausführungsform können zum Beispiel drei oder mehr Zwischenstecker-Karten in einer Reihenschaltung miteinander verbunden werden. In diesem Beispiel enthält die Zwischenstecker-Verschaltungseinheit Folgendes: (i) eine Verkabelung zwischen einem CNA der ersten Zwischenstecker-Karte und einem Schaltmodul der zweiten Zwischenstecker-Karte, (ii) eine Verkabelung zwischen einem CNA der zweiten Zwischenstecker-Karte und einem Schaltmodul der dritten Zwischenstecker-Karte, und (iii) eine Verkabelung zwischen einem CNA der dritten Zwischenstecker-Karte und einem Schaltmodul der ersten Zwischenstecker-Karte. In einer alternativen Ausführungsform enthält die Zwischenstecker-Verschaltungseinheit eine Verkabelung zwischen dem CNA der dritten Zwischenstecker-Karte und dem Schaltmodul der zweiten Zwischenstecker-Karte (anstelle der ersten Zwischenstecker-Karte). Um zusätzliche Redundanz vorzusehen, kann jede Zwischenstecker-Karte darüber hinaus mit mehreren anderen Zwischenstecker-Karten verbunden werden. Die Zwischenstecker-Verschaltungseinheit kann zum Beispiel Folgendes enthalten: (i) eine Verkabelung zwischen einem ersten CNA der dritten Zwischenstecker-Karte und einem Schaltmodul der ersten Zwischenstecker-Karte und (ii) eine Verkabelung zwischen einem zweiten CNA der dritten Zwischenstecker-Karte und einem Schaltmodul der zweiten Zwischenstecker-Karte. Wenn es in der zweiten Zwischenstecker-Karte beziehungsweise in der dritten Zwischenstecker-Karte zu Ausfällen des Schaltmoduls kommt, wird die erste Zwischenstecker-Karte folglich so konfiguriert, dass sie den Netzwerkverkehr für Serverkarten schaltet, die mit der dritten Zwischenstecker-Karte verbunden sind. Der Fachmann erkennt, dass eine beliebige vorher festgelegte Anzahl von Zwischenstecker-Karten mittels der hier beschriebenen Methoden untereinander verbunden werden kann.

[0071] **Fig. 11** zeigt eine Konfiguration **1100** eines Serversystems gemäß einer Ausführungsform der Erfindung, bei dem ein Schaltmodul **206** als ein SPOF in einem Paar von Speicher-ITEs beseitigt wird. Die Konfiguration **1100** enthält ein erstes Speicher-ITE 1102_1 und ein zweites Speicher-ITE 1102_2. Jedes Speicher-ITE 1102_1, 1102_2 kann über eine Zwischenstecker-Karte 204, die ein Schaltmodul **206**

enthält, mit der Mittelplatine verbunden werden. In einer alternativen Ausführungsform ist jedes Speicher-ITE mit der Mittelplatine verbunden und enthält das Schaltmodul **206**. Wie gezeigt ist, enthält jedes Speicher-ITE eine Vielzahl von Komponenten, darunter eine Funktionskarte **1108**, zwei Ausgangsverzweigungs-Karten **1104** und eine Speichereinheit **1106**. Die Funktionskarte **1108** eines jeden Speicher-ITEs **1102** kann so konfiguriert werden, dass die Funktionalität des jeweiligen Speicher-ITEs **11102** kundenspezifisch ausgelegt wird. Die Funktionskarte **1108** kann beispielsweise so konfiguriert werden, dass sie das Speicher-ITE individuell als ein RAID-ITE, ein ITE in Form eines an das Netzwerk angeschlossenen Speichers (network-attached storage (NAS)) und/oder ein ITE in Form eines Datei-Cachespeichers usw. auslegt. Die Ausgangsverzweigungs-Karten **1104** eines jeden Speicher-ITEs **1102** stellen Netzwerk-Verbindungen für das Speicher-ITE **1102** bereit und/oder verbessern die Verfügbarkeit des jeweiligen Speicher-ITEs **1102**. Jede Ausgangsverzweigungs-Karte **1104** enthält eine oder mehrere Ausgangsverzweigungs-Komponenten **1112**. Jedes Speicher-ITE **1102** kann des Weiteren eine Speicher-Verschaltungseinheit 1110_1, 1110_2 enthalten, die die Komponenten des jeweiligen Speicher-ITEs **1102** betriebsfähig verbindet. In einer Ausführungsform stellen die Speicher-Verschaltungseinheiten 1110_1, 1110_2 Serial-Attached-SCSI-(SAS-)Verbindungen zwischen den Komponenten der Speicher-ITEs **1102** bereit. In alternativen Ausführungsformen stellen die Speicher-Verschaltungseinheiten FCoE- oder Serial-ATA-(SATA-)Verbindungen bereit.

[0072] In einer Ausführungsform enthält die Konfiguration **1100** darüber hinaus eine Vermittlungsstellen-Verschaltungseinheit zusammen mit einer Verkabelung **1114** zwischen den Speicher-Verschaltungseinheiten **1110** des Speicher-ITEs **1102**. Die Verkabelung **1114** kann eine Ausgangsverzweigungs-Karte 1104_1 des ersten Speicher-ITEs 1102_1 mit einer Ausgangsverzweigungs-Karte 1104_4 des zweiten Speicher-ITEs 1102_2 verbinden. In Abhängigkeit von der Ausführungsform kann die Vermittlungsstellen-Verschaltungseinheit eine Verkabelung **1116** zwischen einer zusätzlichen Ausgangsverzweigungs-Karte eines jeden Speicher-ITEs **1102** enthalten, um eine höhere Bandbreite zur Verfügung zu stellen. Wenn das Schaltmodul 206_1 des ersten Speicher-ITEs 1102_1 ausfällt, kann ein Schaltmodul 206_2 des zweiten Speicher-ITEs 1102_2 - neben dem Schalten des Netzwerkverkehrs für das zweite Speicher-ITE 1102_2 - folglich den Netzwerkverkehr für das erste Speicher-ITE 1102_1 schalten. Die Auslegung der Speicher-ITEs **1102** in der Weise, dass die Verkabelung **1114** zwischen den Speicher-Verschaltungseinheiten 1110 enthalten ist, beseitigt jedes Schaltmodul 206_1, 206_2 als einen SPOF, wobei die Speicher-Verschaltungseinheiten **1110** der Speicher-ITEs **1102** verwendet werden.

[0073] Fig. 12 veranschaulicht eine Konfiguration **1200** eines Paars von Zwischenstecker-Karten **204** gemäß einer Ausführungsform der Erfindung, wobei die Konfiguration **1200** eine Zwischenstecker-Verschaltungseinheit enthält. Statt die Speicher-Verschaltungseinheiten zur Beseitigung von SPOFs zu verwenden, enthält die Konfiguration **1200** eine Verkabelung zwischen CNAs, die verschiedenen Zwischenstecker-Karten zugeordnet sind, um SPOFs zu beseitigen. Jede Zwischenstecker-Karte **204** enthält ein Schaltmodul **206** und verbindet zwei Serverkarten **202** betriebsfähig mit einer Mittelplatine. Das Schaltmodul **206** einer jeden Zwischenstecker-Karte **204** ist so konfiguriert, dass es den Netzwerkverkehr für die Serverkarten **202** schaltet, die mit der jeweiligen Zwischenstecker-Karte **204** verbunden sind. Jede Zwischenstecker-Karte **204** enthält darüber hinaus einen oder mehrere CNAs **302**. Jede Serverkarte **202** enthält eine oder mehrere CPUs **106**. In Abhängigkeit von der Ausführungsform enthält jede Serverkarte **202** darüber hinaus einen oder mehrere CNAs **304**. Die Konfiguration **1200** der Zwischenstecker-Karten 202 kann auch eine Verkabelung **1202** zwischen den Zwischenstecker-Karten **202** enthalten. Die Verkabelung **1202** kann die CNAs 304_1, 304_2 der Serverkarten **202**, die mit der ersten Zwischenstecker-Karte 204_1 verbunden sind, mit dem Schaltmodul 206_2 der zweiten Zwischenstecker-Karte 204_2 verbinden. Die Verkabelung **1202** kann auch die CNAs 304_3, 304_4 der Serverkarten **202**, die mit der zweiten Zwischenstecker-Karte 204_2 verbunden sind, mit dem Schaltmodul 206_1 der ersten Zwischenstecker-Karte 204_1 verbinden.

[0074] Sollte das Schaltmodul 206_1 der ersten Zwischenstecker-Karte 204_1 ausfallen, verwaltet das Schaltmodul 206_2 der zweiten Zwischenstecker-Karte 204_2 - neben dem Schalten des Netzwerkverkehrs für die Serverkarten **202**, die mit der zweiten Zwischenstecker-Karte 204_2 verbunden sind - den Netzwerkverkehr für die Serverkarten **202**, die mit der ersten Zwischenstecker-Karte 204_1 verbunden sind. Somit beseitigt die Konfiguration 1200 jedes der Schaltmodule **206** als einen SPOF. Anders ausgedrückt, die Serverkarten **202**, die mit jeder Zwischenstecker-Karte verbunden sind, behalten die Verbindungen zu dem Schaltnetzwerk bei und/oder erhalten dessen Redundanz selbst nach einem Ausfall von einem der Schaltmodule **206** aufrecht.

[0075] Wie vorstehend beschrieben wurde, kann die Auslegung des Serversystems in der Weise, dass es eine oder mehrere Schaltkarten enthält, die mit der Mittelplatine verbunden sind, das Schaltmodul als einen SPOF beseitigen. In einem solchen Fall enthält die Mittelplatine eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk. Die Mittelplatine kann eine

oder mehrere Platinen mit einem Gehäuse-Verschaltungselement (chassis interconnect element (CIE)), die selbst ein oder mehrere Verwaltungs-Verarbeitungssubsysteme enthalten, verbinden. Zwar wird bei der hier vorgenommenen Beschreibung der Ausführungsformen Bezug auf die Schaltkarten und die CIE-Platinen als getrennte Komponenten genommen, doch kann in Abhängigkeit von der Ausführungsform ein Teil oder die gesamte Funktionalität der CIE-Platinen in die Schaltkarten integriert werden. Die Schaltkarten und/oder die CIE-Platinen können mit einer ersten Seite der Mittelplatine verbunden werden, und eine oder mehrere Serverkarten können mit einer zweiten Seite der Mittelplatine verbunden werden.

[0076] In einer Ausführungsform können die Schaltkarten und/oder die CIE-Platinen an einer ersten Achse ausgerichtet werden, und die Serverkarten können an einer zweiten Achse ausgerichtet werden. Ferner liegt die erste Achse mindestens weitgehend senkrecht zur zweiten Achse. Die Schaltkarten können beispielsweise senkrecht mit der ersten Seite der Mittelplatine verbunden sein und die Serverkarten können waagerecht mit der zweiten Seite der Mittelplatine verbunden sein oder umgekehrt. In Abhängigkeit von der Ausführungsform enthält das Schaltnetzwerk eine Verdrahtung, die jede Schaltkarte mit jeder Serverkarte verbindet, und/oder eine Verdrahtung, die jede Schaltkarte mit jeder anderen Schaltkarte verbindet. Eine solche Verdrahtung ermöglicht eine redundante Pfadführung, um SPORs und/oder SPOFs in dem Schaltnetzwerk zu verringern und/oder zu beseitigen. Eine Verbindung der Schaltkarten und der Serverkarten mit der Mittelplatine auf senkrechten Achsen kann die Verdrahtung vereinfachen und/oder den erforderlichen Verdrahtungsaufwand verringern (zumindest in manchen Fällen).

[0077] **Fig. 13** zeigt eine Konfiguration **1300** eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten **1302** enthält. Wie gezeigt ist, enthält die Konfiguration **1300** eine Mittelplatine **123**, Serverkarten **202** und CIE-Platinen **1304**. Die Serverkarten **202** sind waagerecht mit der Mittelplatine **123** verbunden, und die Schaltkarten **1302** und die CIE-Platinen **1304** sind senkrecht mit der Mittelplatine 123 verbunden. Die Schaltkarten **1302** und/oder die Serverkarten 202 können im laufenden Betrieb von der Mittelplatine **123** aus ausgetauscht werden. Jede Schaltkarte **1302** kann ein oder mehrere Schaltmodule **206** enthalten, und jede Serverkarte **202** kann eine oder mehrere CPUs enthalten.

[0078] In Abhängigkeit von der Ausführungsform kann das Schaltnetzwerk eine Verdrahtung zwischen jeder Schaltkarte 1302 oder jedem Schaltmodul **206** und jeder Serverkarte **202** oder jeder CPU enthalten. Das Schaltnetzwerk kann auch eine Verdrahtung zwischen jeder Schaltkarte **1302** oder jedem Schaltmodul **206** und jeder anderen Schaltkarte **1302** oder jedem anderen Schaltmodul **206** enthalten. Die Konfiguration **1300** ermöglicht folglich eine redundante Pfadführung zwischen Elementen in dem Schaltnetzwerk und beseitigt dadurch ein Schaltmodul **206** und/oder eine Schaltkarte **1302** als einen SPOF in dem Schaltnetzwerk. Insbesondere die Serverkarten **202** behalten die Verbindungen zu dem Schaltnetzwerk nach einem Ausfall eines Schaltmoduls **206** oder einer Schaltkarte **1302** bei. Überdies wirkt sich ein Reparaturvorgang an einem ausgefallenen Schaltmodul **206** oder einer ausgefallenen Schaltkarte **1302** nicht auf die Verbindungen der Serverkarten 202 zu dem Schaltnetzwerk aus. Der Reparaturvorgang kann den Austausch der Schaltkarte **206** durch eine zweite Schaltkarte, die über ein funktionsfähiges Schaltmodul verfügt, und ohne dass das Serversystem und/oder das Schaltnetzwerk neu gestartet wird, beinhalten.

[0079] **Fig. 14** zeigt eine logische Ansicht **1400** einer Konfiguration eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten enthält. Wie gezeigt ist, enthält die logische Ansicht **1400** mehrere Serverkarten **202** und mehrere Schaltmodule **206**, **1304** eines Serversystems. Das Serversystem kann auch eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk enthalten. Die Schaltmodule **206** sind in der Nähe der Serverkarten **202** in dem Schaltnetzwerk angeordnet und ermöglichen Redundanz beim Schalten des Netzwerkverkehrs für die Serverkarten **202**. Die Schaltmodule **206** können auch als nördliche Schaltmodule bezeichnet werden. Die Schaltmodule 1404 sind in der Nähe der nördlichen Schaltmodule angeordnet und stellen Verbindungen zwischen den nördlichen Schaltmodulen und dem Rest des Schaltnetzwerks bereit. Die Schaltmodule **1404** können auch als südliche Schaltmodule bezeichnet werden.

[0080] In einer Ausführungsform enthält jede Serverkarte **202** zwei CPUs **106** und zwei CNAs **304**. Das Schaltnetzwerk kann eine Verdrahtung **1404** zwischen jedem Schaltmodul **206** und jeder Serverkarte **202** enthalten. Darüber hinaus kann das Schaltnetzwerk eine lokale Gestell-Verschaltungseinheit **1402** enthalten, die wiederum eine Verdrahtung zwischen den Schaltmodulen **206** und den Schaltmodulen **1404** enthält. In Abhängigkeit von der Ausführungsform kann die lokale Gestell-Verschaltungseinheit **1402** auch eine Verdrahtung zwischen jedem der Schaltmodule **206** und jedem anderen der Schaltmodule **206** enthalten, wodurch Alle-zu-alle-Verbindungen unter den Schaltmodulen **206** ermöglicht werden. Vorteilhafterweise beseitigt die Konfiguration jedes Schaltmodul **206** als einen SPOF in dem Schaltnetzwerk. Sollte beispielsweise das Schaltmodul 206_1 ausfallen, kann das Schaltmodul 206_2 den Netzwerkverkehr für jede Serverkarte **202** weiterleiten. Da-

durch kann jede Serverkarte **202** trotz des Ausfalls im Schaltmodul 206_1 mit dem Schaltnetzwerk verbunden bleiben.

[0081] Fig. 15 zeigt eine Konfiguration **1500** eines Serversystems gemäß einer Ausführungsform der Erfindung, das mehrere Schaltkarten **1302** enthält. Wie gezeigt ist, enthält jede der mehreren Schaltkarten **1302** zwei Schaltmodule **206** oder nördliche Schaltmodule. Jedes Schaltmodul verwaltet den Netzwerkverkehr für mindestens eine Serverkarte **202**. Die Konfiguration **1500** enthält eine Verdrahtung **1404** zwischen jedem Schaltmodul **206** des Servers und jeder Serverkarte von einer Teilgruppe der Serverkarten **202**. Teilgruppen der Serverkarten **202** können zum Beispiel eine erste Teilgruppe mit den Serverkarten $202_{1\text{-}7}$ und eine zweite Teilgruppe mit den Serverkarten $202_{8\text{-}14}$ beinhalten. Die Konfiguration **1500** kann auch eine Alle-zu-alle-Verdrahtung (nicht gezeigt) unter einer jeden Teilgruppe der Schaltmodule **206** enthalten. Die Teilgruppen der Schaltmodule **206** können zum Beispiel eine erste Teilgruppe mit den Schaltmodulen $206_{1\text{-}4}$ und eine zweite Teilgruppe mit den Schaltmodulen $206_{5\text{-}8}$ beinhalten. Anders ausgedrückt, die Schaltkarten **1302** und/oder die Serverkarten 202 können physisch in verschiedene Teilgruppen in Bezug auf die Verdrahtung in dem Schaltnetzwerk unterteilt werden, wobei die Verdrahtung Verbindungen und/oder Redundanz in dem Schaltnetzwerk vorsieht. Die Konfiguration **1500** kann auch eine weitere Verdrahtung über die Teilgruppen der Schaltmodule **206** enthalten.

[0082] In einer Ausführungsform enthalten die Schaltkarten $1302_{1,3}$ darüber hinaus ein oder mehrere Verwaltungs-Verarbeitungssubsysteme. Die Verwaltungs-Verarbeitungssubsysteme enthalten Verwaltungs-Firmware, die so konfiguriert ist, dass sie den fehler- beziehungsweise störungsfreien Betrieb des Serversystems und/oder der Elemente des Schaltnetzwerks überwacht, die Elemente konfiguriert und/oder einen Ausfall der Elemente erkennt und darauf reagiert. Wie gezeigt ist, enthalten die Verwaltungs-Verarbeitungssubsysteme eine Eingabe-/Ausgabe-Hauptsteuereinheit (input/output master controller (IoMC)) und ein Gehäuse-Dienstelement (chassis service element (CSE)). Die IoMC verwaltet Elemente in dem Schaltnetzwerk, während das CSE Komponenten in dem Server-Gehäuse enthält verwaltet. Ferner können die IOMCs über eine IoMC-Verschaltungseinheit auf den Schaltkarten **1302** betriebsfähig miteinander verbunden werden, um Redundanz bei der Überwachung und/oder Verwaltung des Schaltnetzwerks zu ermöglichen, wobei die IoMC-Verschaltungseinheit eine physische Verdrahtung zwischen den IoMCs enthält. Nach einem Ausfall einer ersten IoMC 1506_1 kann eine zweite IoMC 1506_2 so konfiguriert werden, dass sie anstelle der ersten IoMC 1506_1 Funktionen zur Überwachung und/oder Verwaltung des Schaltnetzwerks bereitstellt. Vorteilhafterweise beseitigt die Konfiguration **1500** jedes Schaltmodul **206** als einen SPOF und beseitigt darüber hinaus auch jede IoMC 1506 als einen SPOF in dem Schaltnetzwerk.

[0083] In einer Ausführungsform enthalten die Schaltkarten $1302_{2,4}$ darüber hinaus Schaltmodule **1404** oder südliche Schaltmodule. Wie vorstehend beschrieben wurde, stellen die südlichen Schaltmodule Verbindungen zwischen den nördlichen Schaltmodulen und dem Rest des Schaltnetzwerks bereit. Jede Schaltkarte **1302** enthält des Weiteren lokale Verbindungsleitungen (L-Verbindungsleitungen) 1502 und Distanz-Verbindungsleitungen (D-Verbindungsleitungen) 1504. Die L-Verbindungsleitung 1502 stellt der Schaltkarte **1302** physische Verbindungen zu einem anderen Gehäuse innerhalb eines einzelnen physischen Gestellrahmens bereit. Die D-Verbindungsleitung 1504 stellt der Schaltkarte **1302** physische Verbindungen zu einem anderen Gehäuse über physische Gestellrahmen bereit. Das Schaltnetzwerk enthält dadurch folglich eine zusätzliche redundante Pfadführung für Elemente des Schaltnetzwerks.

[0084] Fig. 16 ist ein Flussdiagramm, das ein Verfahren **1600** gemäß einer Ausführungsform der Erfindung zur Verringerung der Auswirkung eines Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk zeigt. Wie gezeigt ist, beginnt das Verfahren 1600 im Schritt **1610**, in dem ein Anbieter eines Serversystems das Serversystem so auslegt, dass es eine Mittelplatine enthält, wobei die Mittelplatine eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk enthält. Im Schritt **1620** legt der Anbieter des Serversystems das Serversystem des Weiteren so aus, dass es eine oder mehrere Serverkarten enthält, die mit der Mittelplatine verbunden sind, wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann. Im Schritt **1630** legt der Anbieter des Serversystems das Serversystem des Weiteren so aus, dass es eine oder mehrere Schaltkarten enthält, die mit der Mittelplatine verbunden sind, wobei die eine oder die mehreren Schaltkarten mit der einen oder den mehreren Serverkarten betriebsfähig verbunden sind. Jede Schaltkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden und enthält ein oder mehrere Schaltmodule. Jedes Schaltmodul ist so konfiguriert, dass es den Netzwerkverkehr für mindestens eine Serverkarte schaltet. Nach dem Schritt **1630** endet das Verfahren **1600**.

[0085] Fig. 17 ist ein Flussdiagramm, das ein Verfahren **1700** gemäß einer Ausführungsform der Erfindung zum Beseitigen eines Schaltmoduls als einen SPOF darstellt. Wie gezeigt ist, beginnt das Verfahren **1700** im Schritt **1710**, in dem ein Serversystem bereitgestellt wird, das eine Mittelplatine, eine oder mehrere mit der Mittelplatine verbundene Ser-

verkarten und eine oder mehrere mit der Mittelplatine verbundene Schaltkarten enthält, wobei die eine oder die mehreren Serverkarten mit der einen oder den mehreren Schaltkarten betriebsfähig verbunden sind. Im Schritt **1720** stellt das Serversystem fest, dass ein erstes Schaltmodul einer ersten Schaltkarte ausgefallen ist. Die Feststellung kann zum Beispiel von einer Komponente einer Verwaltungs-Firmware des Serversystems getroffen werden. Im Schritt **1730** schaltet ein zweites Schaltmodul den Netzwerkverkehr für die eine oder die mehreren Serverkarten, nachdem das Serversystem festgestellt hat, dass das erste Schaltmodul der ersten Schaltkarte ausgefallen ist, wobei das zweite Schaltmodul entweder in der ersten Schaltkarte oder einer zweiten Schaltkarte enthalten ist. Nach dem Schritt **1730** endet das Verfahren **1700**.

[0086] Vorteilhafterweise verringern Ausführungsformen der Erfindung die Auswirkung eines Ausfalls einer Vermittlungsstelle in einem Schaltnetzwerk. Eine Ausführungsform der Erfindung ist ein Serversystem mit einer Mittelplatine, die selbst eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk enthält. Das Serversystem kann des Weiteren eine oder mehrere mit der Mittelplatine verbundene Serverkarten enthalten. Jede Serverkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden. Das Serversystem kann darüber hinaus eine oder mehrere mit der Mittelplatine verbundene Schaltkarten enthalten. Und die eine oder die mehreren Schaltkarten sind mit der einen oder den mehreren Serverkarten betriebsfähig verbunden. Jede Schaltkarte kann im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden und enthält ein oder mehrere Schaltmodule. Jedes Schaltmodul ist so konfiguriert, dass es den Netzwerkverkehr für mindestens eine Serverkarte schaltet.

[0087] In einer Ausführungsform kann das Schaltnetzwerk eine Verdrahtung zwischen jeder Schaltkarte und jeder Serverkarte und/oder eine Verdrahtung zwischen jeder Schaltkarte und jeder anderen Schaltkarte enthalten, um eine redundante Pfadführung vorzusehen. Das Serversystem kann auch Verwaltungs-Firmware enthalten, die so konfiguriert ist, dass sie ein ausgefallenes Element in dem Schaltnetzwerk erkennt und/oder darauf reagiert. Wenn das Schaltmodul einer ersten Schaltkarte ausfällt, wird ein zweites Schaltmodul auf der ersten Schaltkarte oder auf einer zweiten Schaltkarte folglich so konfiguriert, dass es den Netzwerkverkehr für die Serverkarten, die von dem Schaltmodul der ersten Schaltkarte unterstützt werden, weiterleitet. Diese Konfiguration des Serversystems beseitigt somit das Schaltmodul der ersten Schaltkarte als einen SPOF.

[0088] Darüber hinaus können manche Ausführungsformen der Erfindung auch die Auswirkung eines Reparaturvorgangs auf das Schaltnetzwerk verringern. Sollte beispielsweise ein Schaltmodul der ersten Schaltkarte ausfallen, kann die erste Schaltkarte durch eine dritte Schaltkarte ausgetauscht werden, ohne das Serversystem und/oder das Schaltnetzwerk abzuschalten oder neu zu starten. Dadurch wird die Auswirkung des Reparaturvorgangs verringert (zumindest in manchen Fällen) und auch die Verfügbarkeit des Serversystems und/oder des Schaltnetzwerks verbessert. Dort wo das Schaltnetzwerk Redundanz im Hinblick auf Verbindungen vorsieht, konfigurieren Ausführungsformen der Erfindung das Serversystem so, dass die mögliche Auswirkung des Reparaturvorgangs auf die vorgesehene Redundanz verringert wird.

[0089] Während sich das Vorstehende auf Ausführungsformen der vorliegenden Erfindung bezieht, können andere und weitere Ausführungsformen der Erfindung entworfen werden, ohne vom grundlegenden Umfang der Erfindung abzuweichen, und der Umfang der Erfindung wird von den Ansprüchen, die folgen, festgelegt.

Patentansprüche

1. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500), das Folgendes umfasst:
eine Mittelplatine (123), die eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk (423) umfasst;
eine oder mehrere Serverkarten (202), die mit der Mittelplatine verbunden sind, wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann; und
eine oder mehrere Schaltkarten (1302), die mit der Mittelplatine verbunden sind, wobei die eine oder die mehreren Schaltkarten mit der einen oder den mehreren Serverkarten betriebsfähig verbunden sind, wobei jede Schaltkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann und ein oder mehrere Schaltmodule (206) umfasst, wobei jedes Schaltmodul so konfiguriert ist, dass es den Netzwerkverkehr für mindestens eine Serverkarte von der einen oder den mehreren Serverkarten schaltet, und wobei ein erstes Schaltmodul einer ersten Schaltkarte so konfiguriert ist, dass es den Netzwerkverkehr für die eine oder die mehreren Serverkarten nach einem Feststellen eines Ausfalls eines zweiten Schaltmoduls schaltet, das in der ersten Schaltkarte enthalten ist, und dadurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert.

2. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei die eine oder die mehreren Serverkarten (202) mit einer ersten Seite der Mittelplatine (123) verbunden sind und wobei die eine oder die mehreren Schaltkarten (1302) mit einer zweiten Seite der Mittelplatine verbunden sind.

3. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei das Schaltnetzwerk (423) mindestens eines von Folgendem umfasst: (i) eine Verdrahtung zwischen jeder Schaltkarte (1302) und jeder Serverkarte (202) und (ii) eine Verdrahtung zwischen jeder Schaltkarte und jeder anderen Schaltkarte.

4. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei die eine oder die mehreren Serverkarten (202) entsprechend einer ersten Achse auf der Mittelplatine (123) ausgerichtet sind und wobei die eine oder die mehreren Schaltkarten (1302) entsprechend einer zweiten Achse auf der Mittelplatine ausgerichtet sind und wobei die zweite Achse senkrecht zu der ersten Achse liegt.

5. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 4, wobei die eine oder die mehreren Serverkarten (202) waagerecht mit einer ersten Seite der Mittelplatine (123) verbunden sind und wobei die eine oder die mehreren Schaltkarten (1302) senkrecht mit einer zweiten Seite der Mittelplatine verbunden sind.

6. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei eine Serverkarte (202) und/oder eine Schaltkarte (1302) so konfiguriert sind, dass sie gegen einen funktionsfähigen Ersatz ausgetauscht werden können, ohne einen Neustart des Systems erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks (423) erforderlich zu machen.

7. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei das System so konfiguriert ist, dass es den funktionsfähigen Ersatz in das Schaltnetzwerk (423) integriert, ohne einen Neustart des Systems erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks erforderlich zu machen.

8. System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach Anspruch 1, wobei das System ein Blade-System umfasst, wobei jede Serverkarte (202) einen Server-Blade (104) umfasst und wobei der Netzwerkverkehr mindestens eines von Folgendem umfasst:
(i) Ethernet-Verkehr und (ii) Fibre-Channel-over-Ethernet-(FCoE) -Verkehr.

9. Schaltmodul (206) für ein System (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) nach einem der Ansprüche 1 bis 8, das Folgendes umfasst:
einen Rechnerprozessor; und
einen Arbeitsspeicher, der Verwaltungs-Firmware speichert, die, wenn sie auf dem Rechnerprozessor ausgeführt wird, eine Operation durchführt, welche Folgendes umfasst:
Schalten des Netzwerkverkehrs für eine erste Serverkarte (202) in dem Serversystem; und
nach einem Feststellen eines Ausfalls eines zweiten Schaltmoduls, das den Netzwerkverkehr für eine zweite Serverkarte schaltet, Schalten des Netzwerkverkehrs für die zweite Serverkarte;
wobei das Schaltmodul in einer ersten Schaltkarte (1302) enthalten ist, wobei das zweite Schaltmodul in der ersten Schaltkarte enthalten ist, wobei jede Schaltkarte mit einer Mittelplatine (123) verbunden ist, wobei jede Serverkarte mit der Mittelplatine verbunden ist, wobei die Mittelplatine eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk (423) umfasst, wobei jede Schaltkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann und wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann, wodurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert wird.

10. Schaltmodul (206) nach Anspruch 9, wobei jede Serverkarte (202) mit einer ersten Seite der Mittelplatine (123) verbunden ist und wobei jede Schaltkarte (1302) mit einer zweiten Seite der Mittelplatine verbunden ist.

11. Schaltmodul (206) nach Anspruch 9, wobei das Schaltnetzwerk (423) mindestens eines von Folgendem umfasst: (i) eine Verdrahtung zwischen jeder Schaltkarte (1302) und jeder Serverkarte (202) und (ii) eine Verdrahtung zwischen jeder Schaltkarte und jeder anderen Schaltkarte.

12. Schaltmodul (206) nach Anspruch 9, wobei jede Serverkarte (202) entsprechend einer ersten Achse auf der Mittelplatine (123) ausgerichtet ist und wobei jede Schaltkarte (1302) entsprechend einer zweiten Achse auf der Mittelplatine ausgerichtet ist und wobei die zweite Achse senkrecht zu der ersten Achse liegt.

13. Schaltmodul (206) nach Anspruch 12, wobei jede Serverkarte (202) waagerecht mit einer ersten Seite der Mittelplatine (123) verbunden ist und wobei jede Schaltkarte (1302) senkrecht mit einer zweiten Seite der Mittelplatine verbunden ist.

14. Schaltmodul (206) nach Anspruch 9, wobei eine Serverkarte (202) und/oder eine Schaltkarte (1302) so konfiguriert sind, dass sie gegen einen funktionsfähigen Ersatz ausgetauscht werden können, ohne einen Neustart des Systems (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks (423) erforderlich zu machen.

15. Schaltmodul (206) nach Anspruch 9, wobei das Serversystem (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) so konfiguriert ist, dass es den funktionsfähigen Ersatz in das Schaltnetzwerk (423) integriert, ohne einen Neustart des Serversystems erforderlich zu machen und ohne einen Neustart des Schaltnetzwerks erforderlich zu machen.

16. Schaltmodul (206) nach Anspruch 9, wobei das Serversystem (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500) ein Blade-System umfasst, wobei jede Serverkarte (202) einen Server-Blade (104) umfasst und wobei der Netzwerkverkehr mindestens eines von Folgendem umfasst: (i) Ethernet-Verkehr und (ii) Fibre-Channel-over-Ethernet-(FCoE)-Verkehr.

17. Von einem Rechner durchgeführtes Verfahren, das Folgendes umfasst:
in einem Serversystem (102, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500), das eine Mittelplatine(123), eine oder mehrere Serverkarten (202), die mit der Mittelplatine verbunden sind, und eine oder mehrere Schaltkarten (1302), die mit der Mittelplatine verbunden sind, umfasst, wobei die eine oder die mehreren Serverkarten mit der einen oder den mehreren Schaltkarten betriebsfähig verbunden sind, wobei die Mittelplatine eine Netzwerk-Verschaltungseinheit für ein Schaltnetzwerk (423) umfasst, wobei jede Schaltkarte ein oder mehrere Schaltmodule (206) umfasst, wobei jedes Schaltmodul so konfiguriert ist, dass es den Netzwerkverkehr für mindestens eine Serverkarte von der einen oder den mehreren Serverkarten schaltet, wobei jede Serverkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann und wobei jede Schaltkarte im laufenden Betrieb von der Mittelplatine aus ausgetauscht werden kann, Feststellen, dass ein erstes Schaltmodul einer ersten Schaltkarte ausgefallen ist; und
nach der Feststellung, dass das erste Schaltmodul der ersten Schaltkarte ausgefallen ist, Schalten des Netzwerkverkehrs für die eine oder die mehreren Serverkarten durch ein zweites Schaltmodul, das in der ersten Schaltkarte enthalten ist, und dadurch, über die eine oder die mehreren Schaltkarten, die Auswirkung des Ausfalls des zweiten Schaltmoduls auf das System verringert.

Es folgen 17 Seiten Zeichnungen

Anhängende Zeichnungen

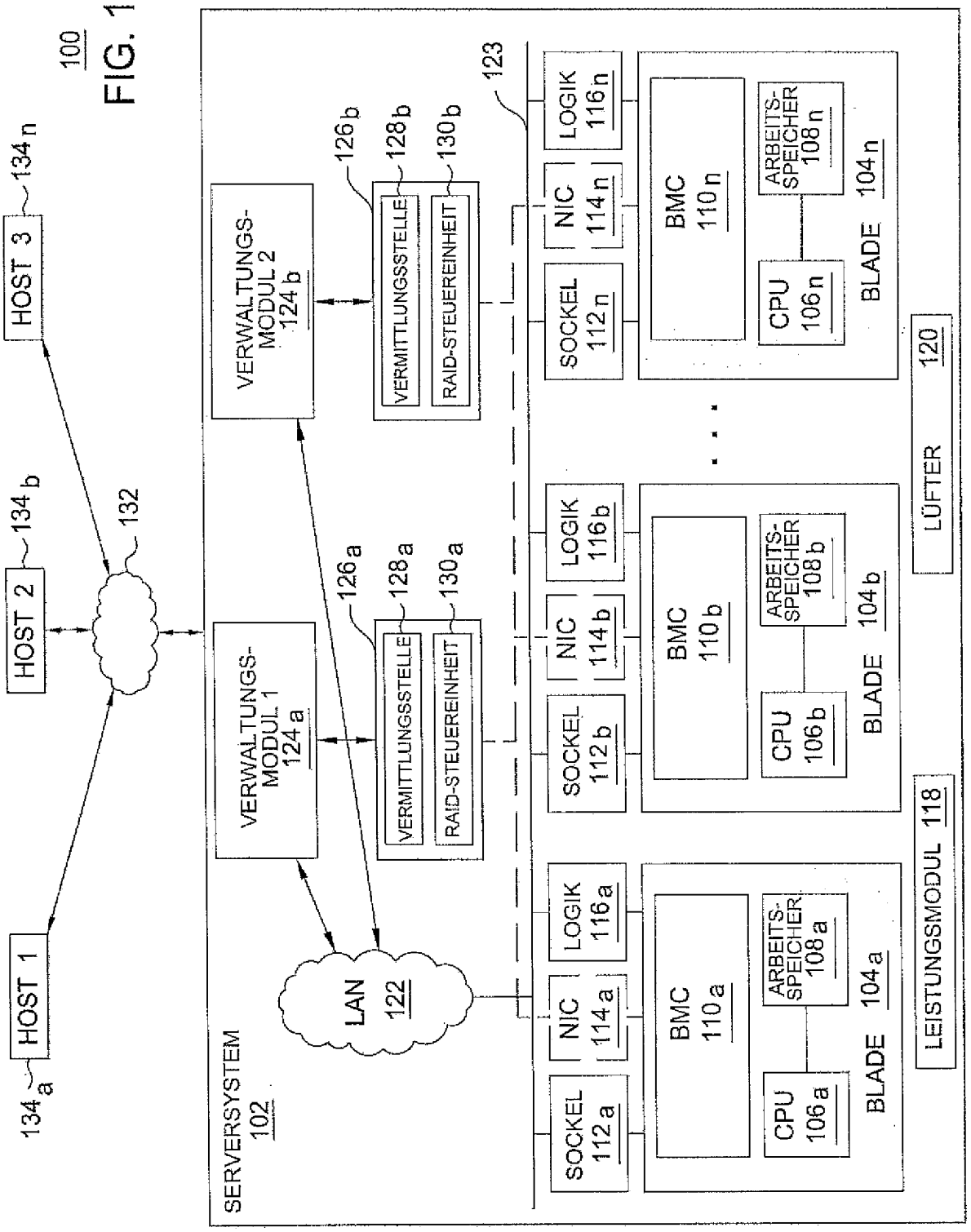
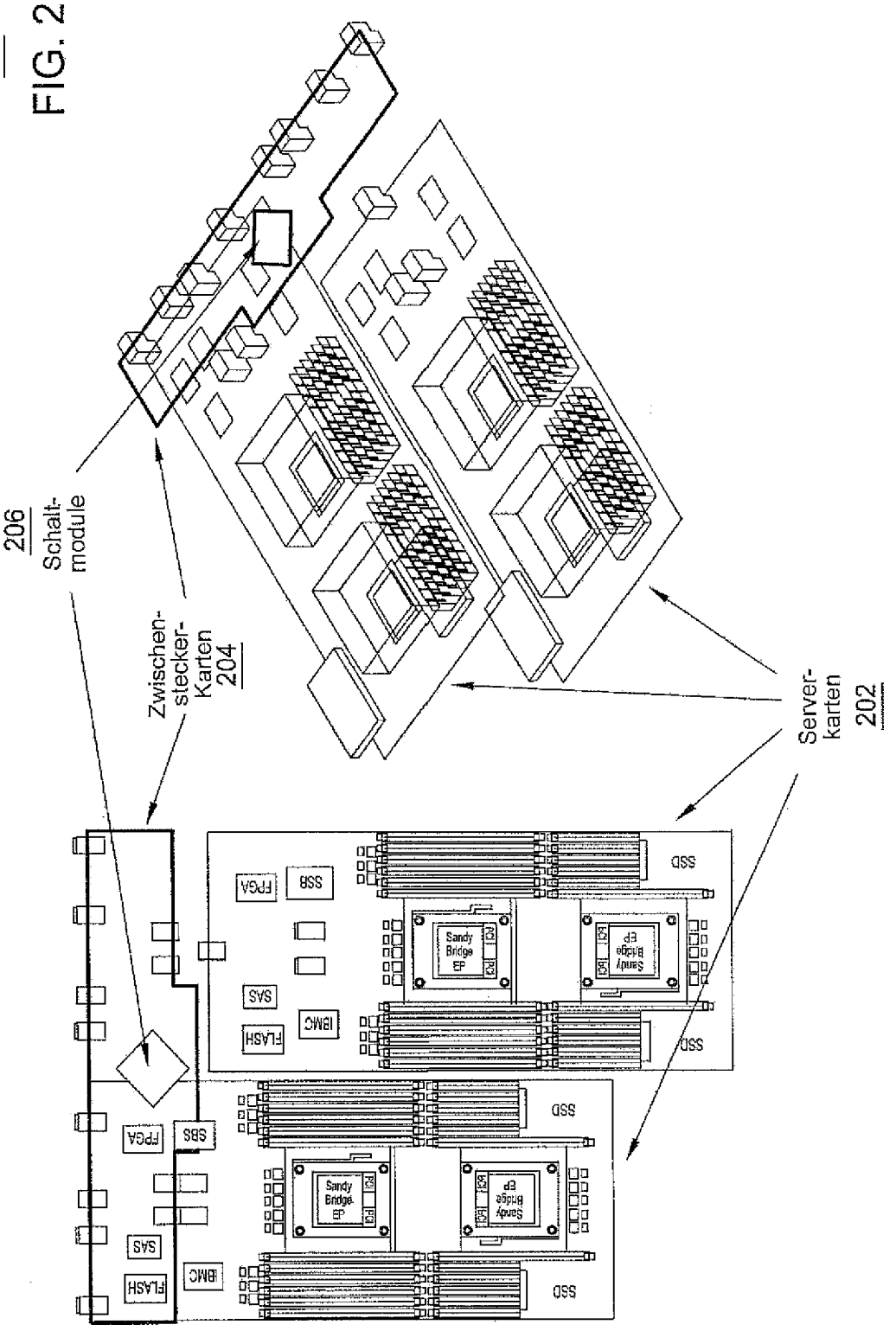
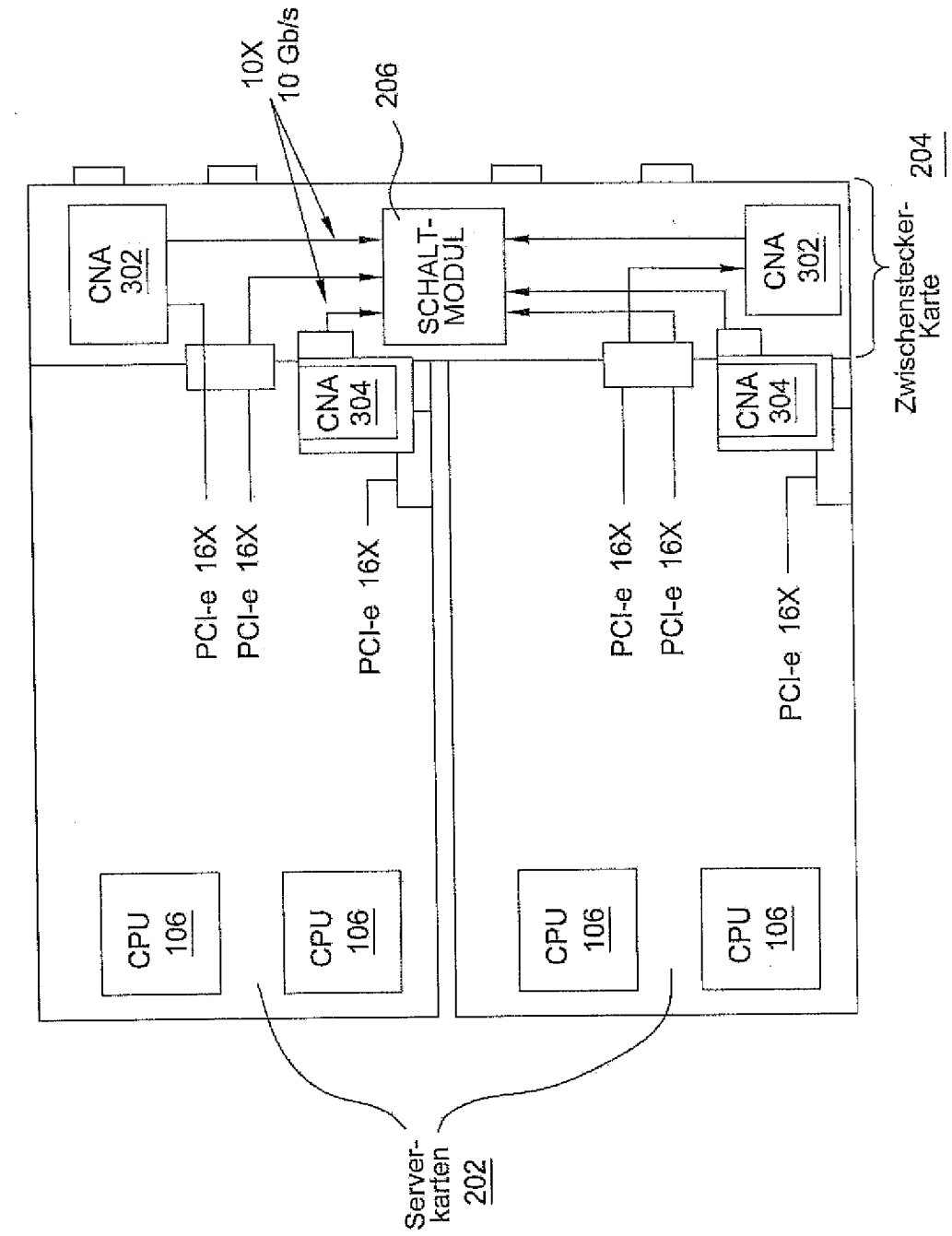


FIG. 2


200
206
Schalt-
module
Zwischen-
stecker-
Karten
204
Server-
karten
202
Sandy
Bridge
EP
SSD
FPGA
SSB
SAS
FLASH
BMC
SBS

300

FIG. 3


10X
10 Gb/s
206
204
CNA 302
SCHALT-MODUL
CNA 302
Zwischenstecker-Karte
CNA 304
CNA 304
PCI-e 16X
PCI-e 16X
PCI-e 16X
PCI-e 16X
PCI-e 16X
PCI-e 16X
CPU 106
CPU 106
CPU 106
CPU 106
Server-karten 202

400
FIG. 4

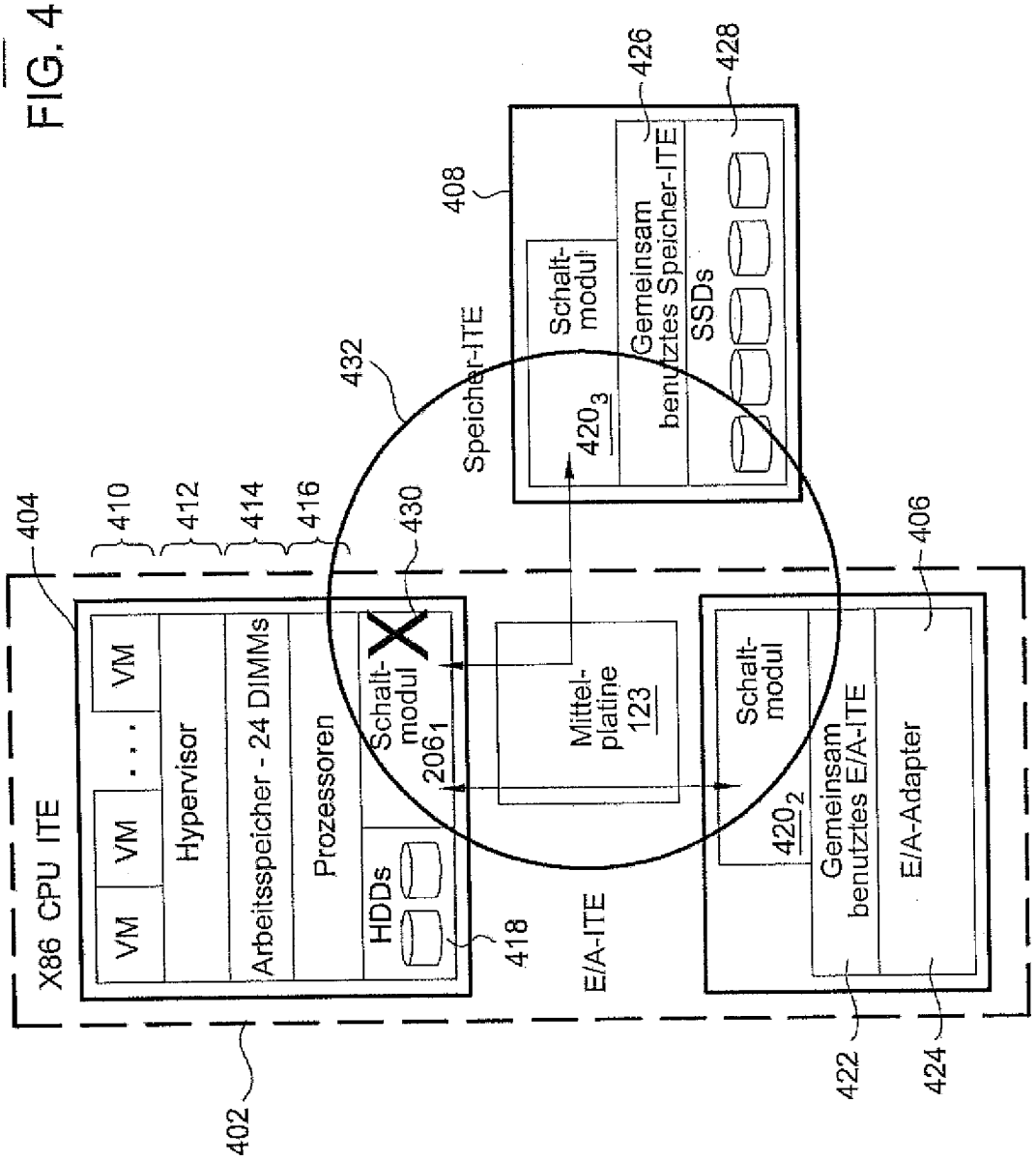
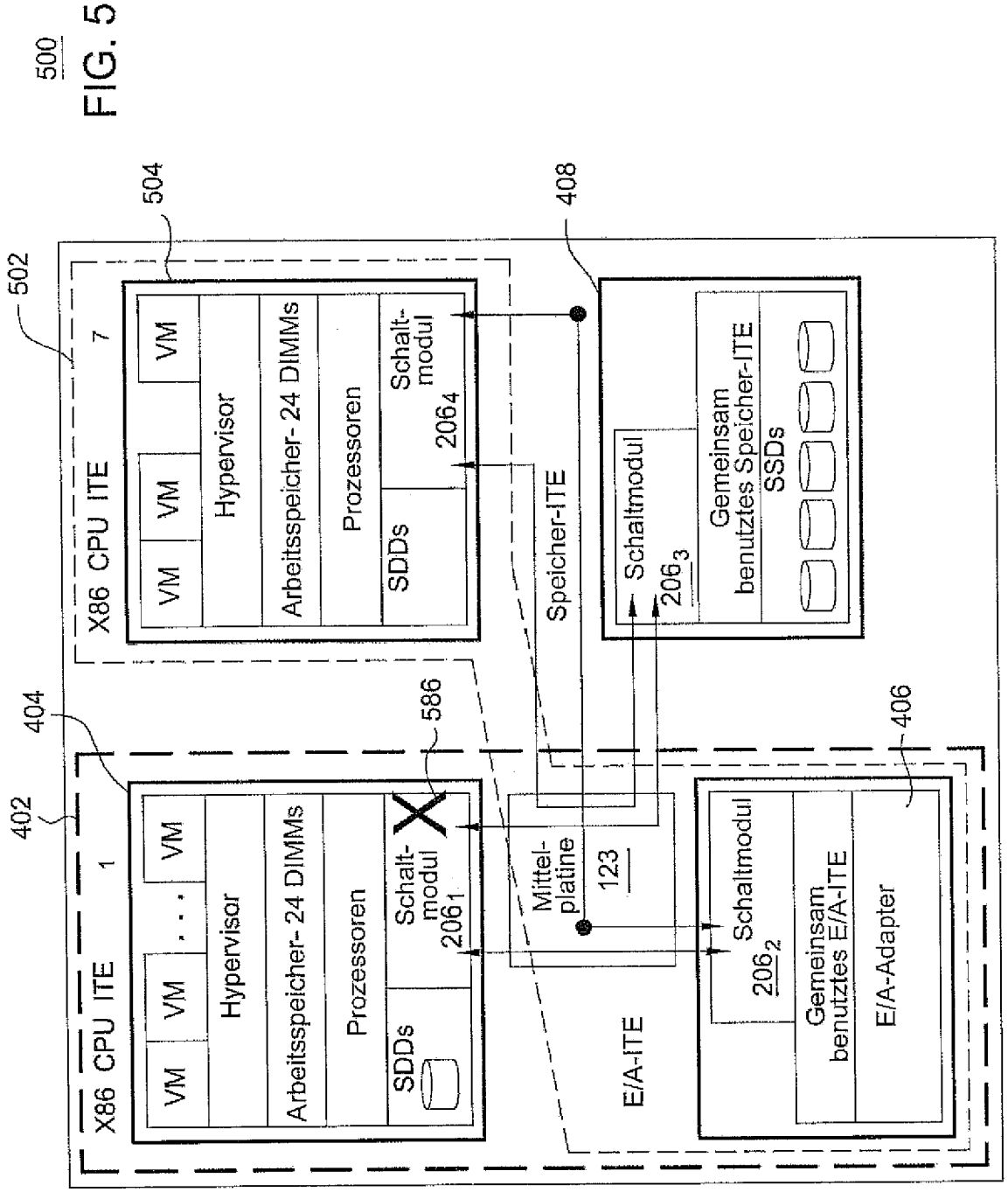
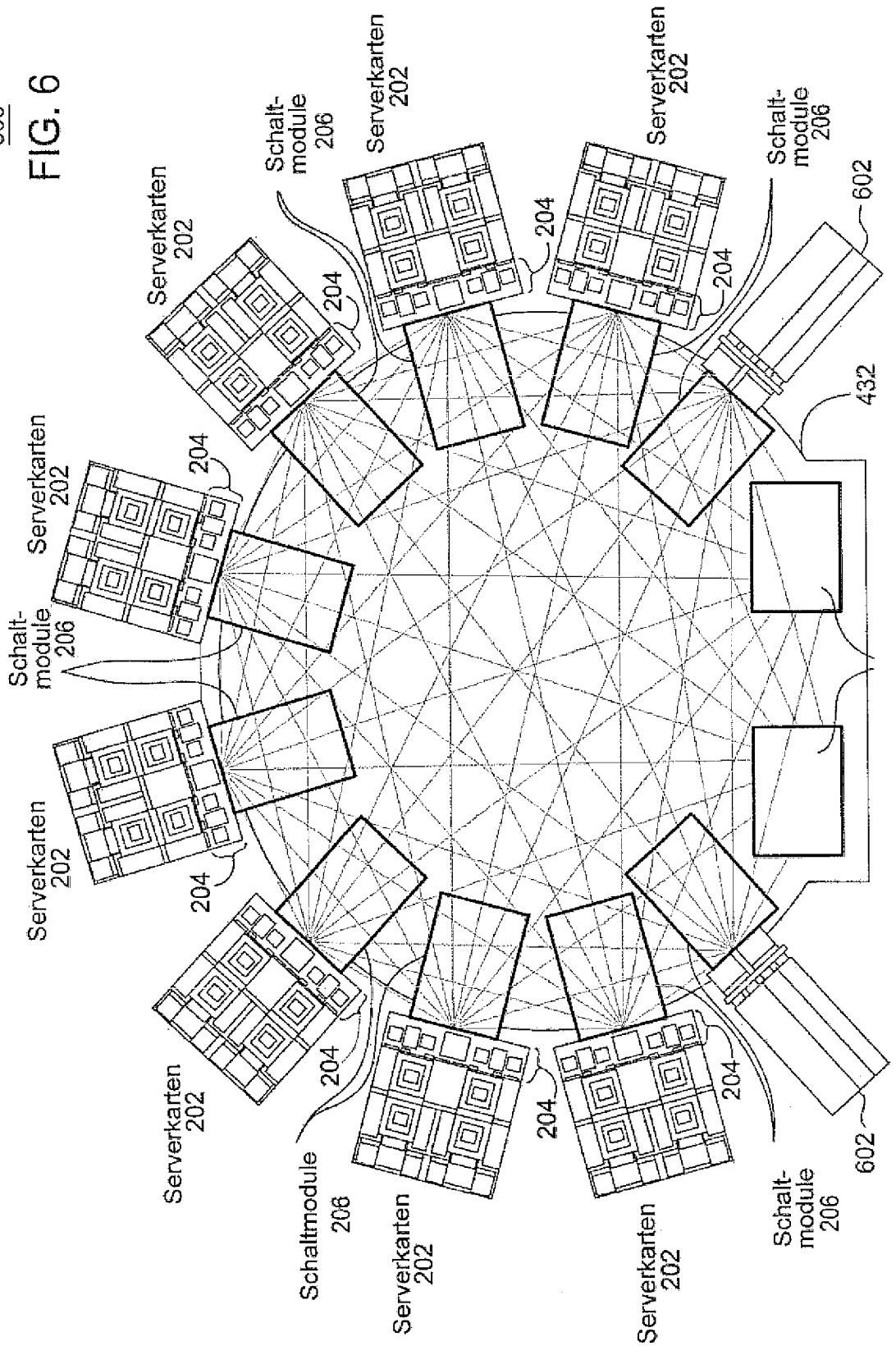

X86 CPU ITE
VM
VM
. . .
VM
Hypervisor
Arbeitsspeicher - 24 DIMMs
Prozessoren
HDDs
Schalt-modul 2061
Mittel-platine 123
Schalt-modul 420 2
Gemeinsam benutztes E/A-ITE
E/A-Adapter
E/A-ITE
Speicher-ITE
Schalt-modul 420 3
Gemeinsam benutztes Speicher-ITE
SSDs
402
404
406
408
410
412
414
416
418
422
424
426
428
430
432

FIG. 5


500
502
504
408
404
586
406
402
X86 CPU ITE
7
VM
VM
VM
Hypervisor
Arbeitsspeicher- 24 DIMMs
Prozessoren
SDDs
Schalt-modul
206_4
Speicher-ITE
Schaltmodul
206_3
Gemeinsam benutztes Speicher-ITE
SSDs
X86 CPU ITE
1
VM
VM
. . .
VM
Hypervisor
Arbeitsspeicher- 24 DIMMs
Prozessoren
SDDs
Schalt-modul
206_1
Mittel-platine
123
E/A-ITE
Schaltmodul
206_2
Gemeinsam benutztes E/A-ITE
E/A-Adapter

600

FIG. 6


Serverkarten 202
Schalt-module 206
204
602
432
Schaltmodule 206

700

FIG. 7

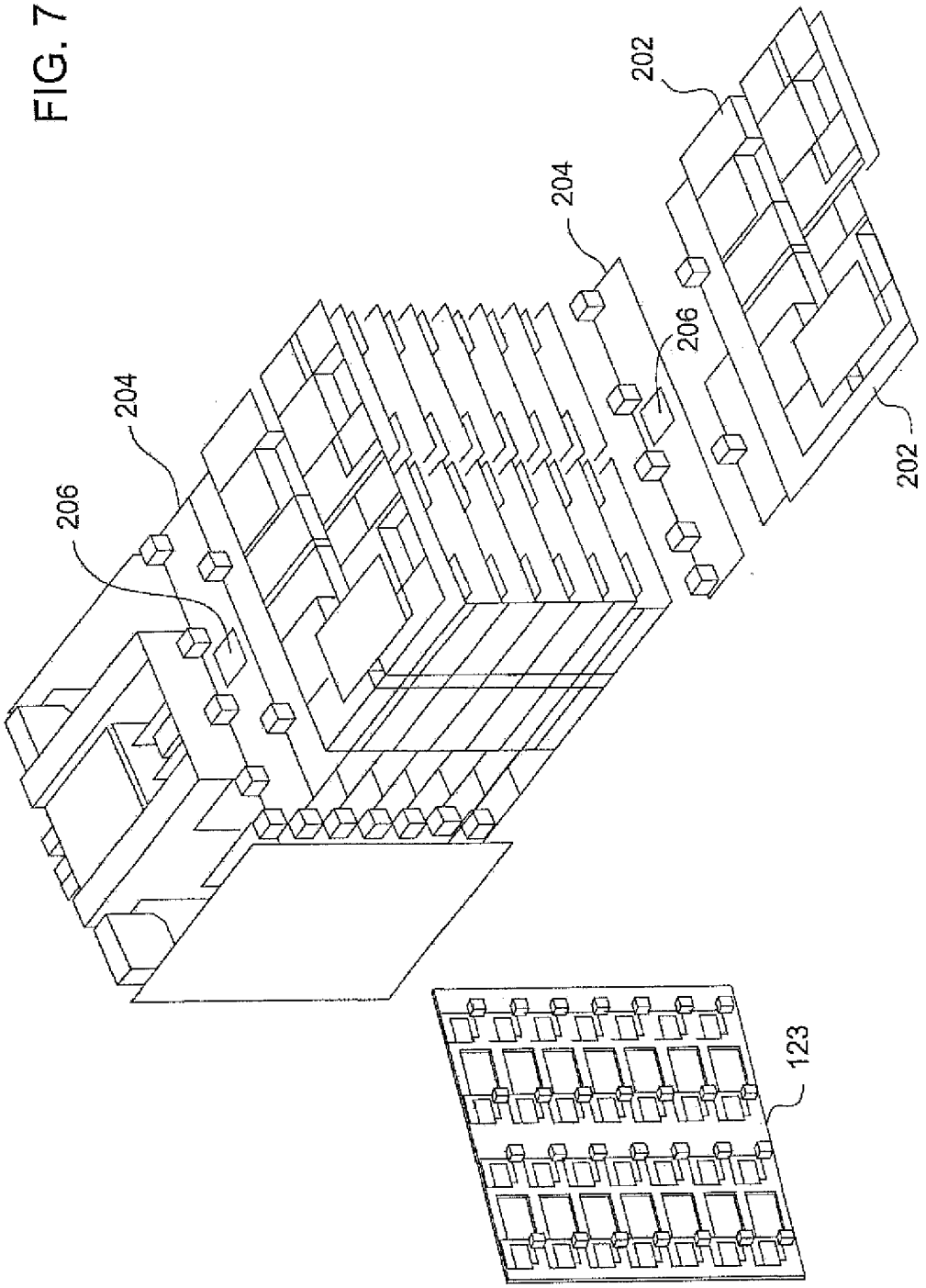


FIG. 8

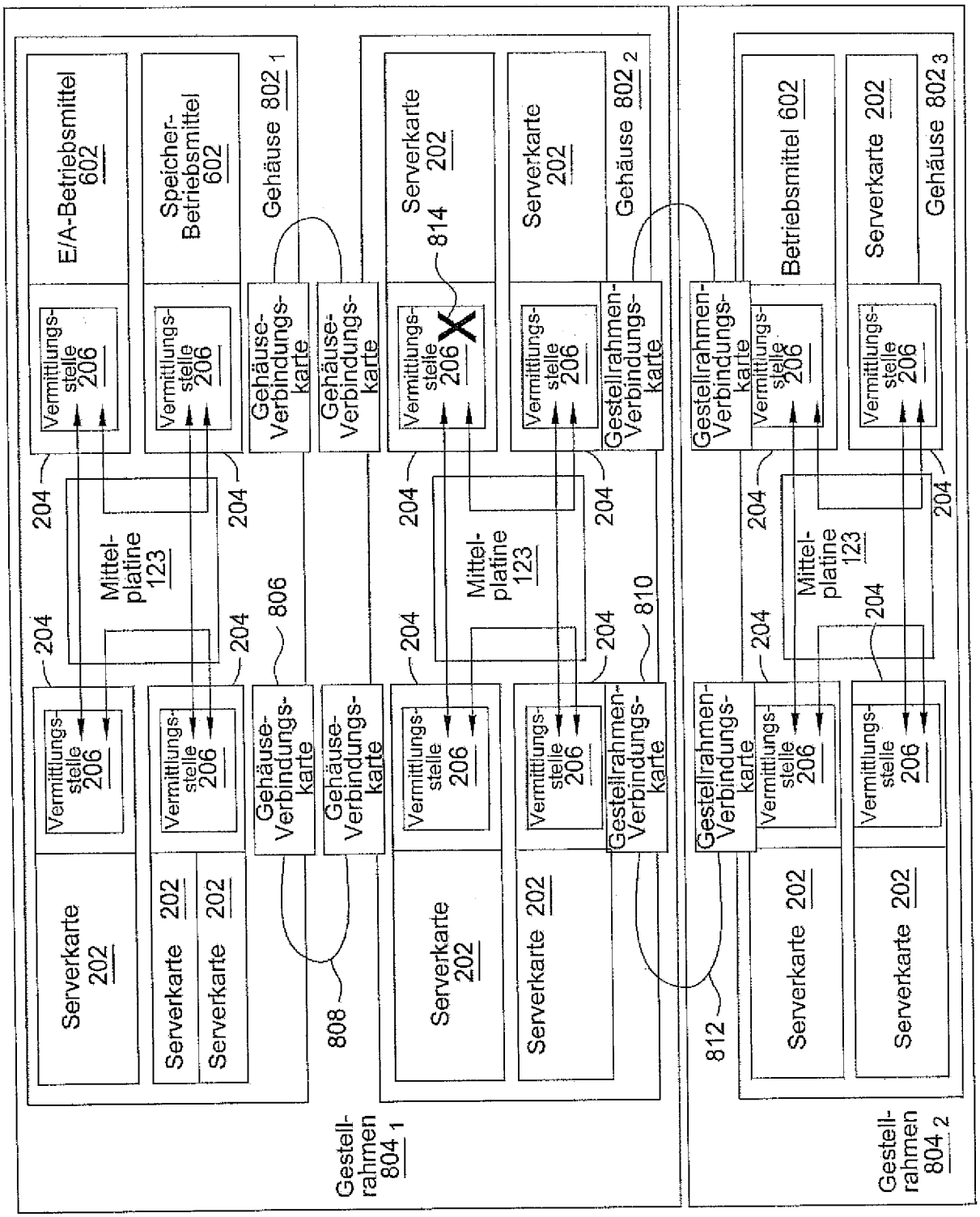
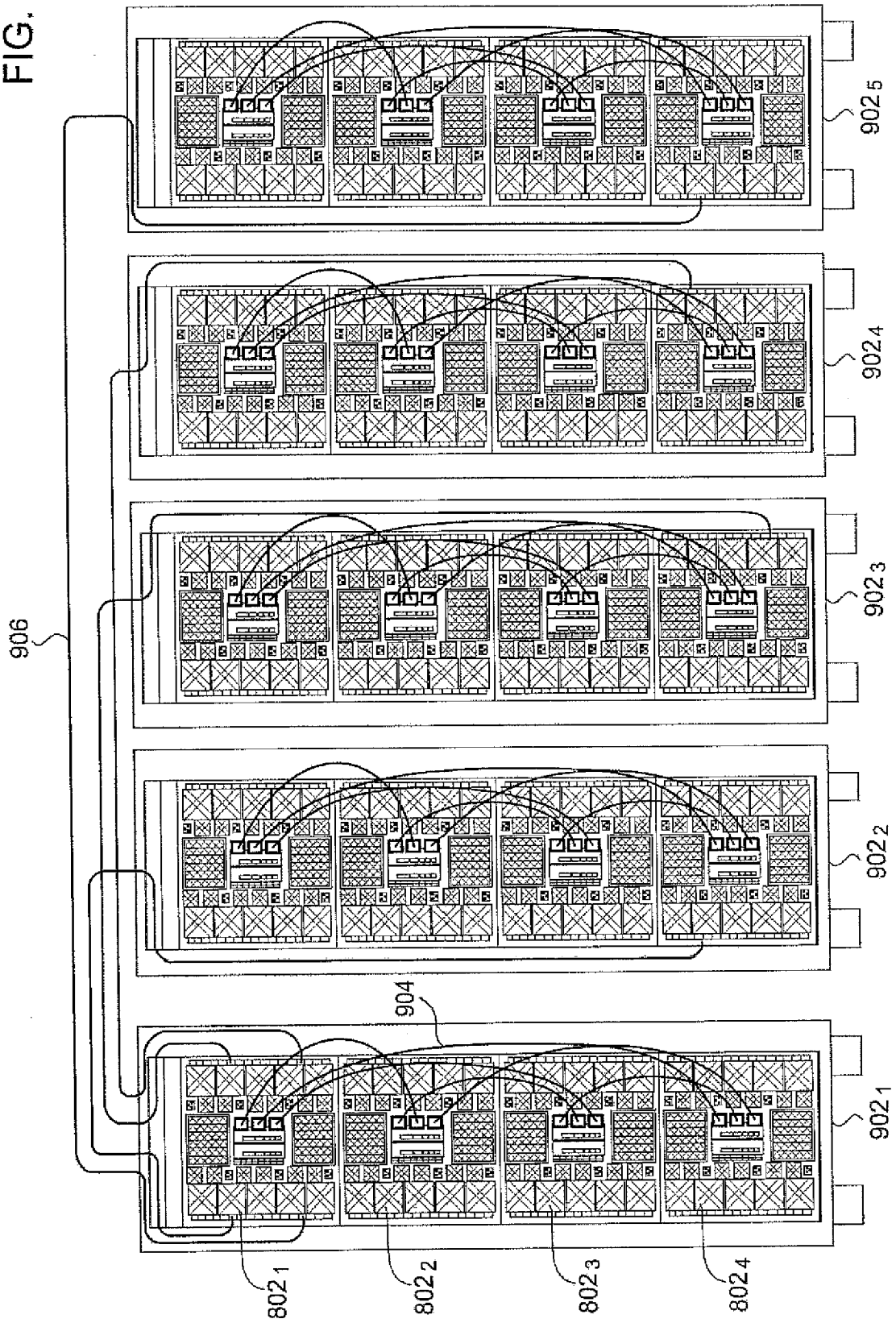

800
Gestell-rahmen 804 1
Gestell-rahmen 804 2
Gehäuse 802 1
Gehäuse 802 2
Gehäuse 802 3
E/A-Betriebsmittel 602
Speicher-Betriebsmittel 602
Betriebsmittel 602
Serverkarte 202
Vermittlungs-stelle 206
Gehäuse-Verbindungs-karte
Gestellrahmen-Verbindungs-karte
Mittel-platine 123
204
806
808
810
812
814

FIG. 9



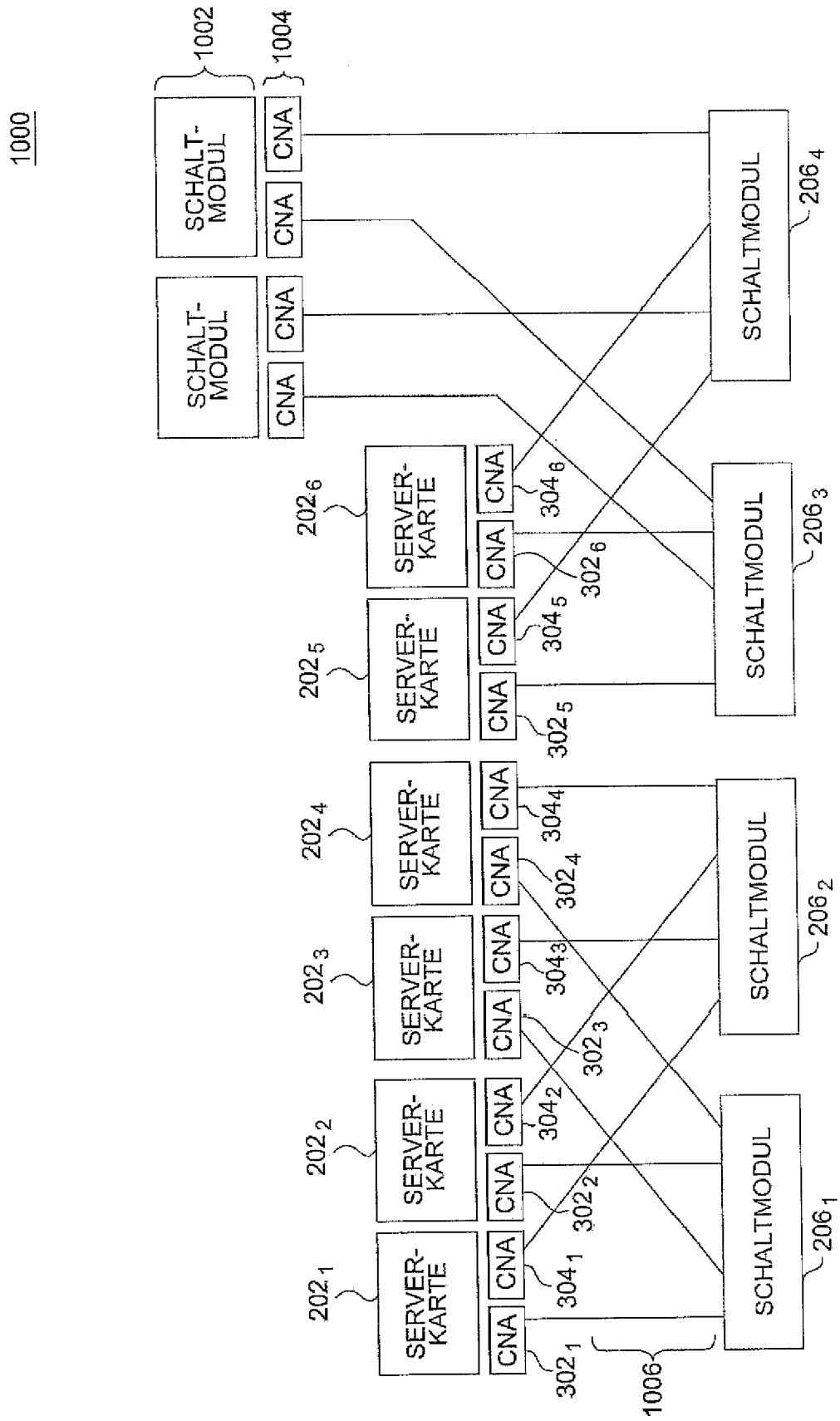

1000
1002
1004
SCHALT-MODUL
SCHALT-MODUL
CNA
SERVER-KARTE
SCHALTMODUL
202_1
202_2
202_3
202_4
202_5
202_6
206_1
206_2
206_3
206_4
302_1
302_2
302_3
302_4
302_5
302_6
304_1
304_2
304_3
304_4
304_5
304_6
1006

FIG. 10

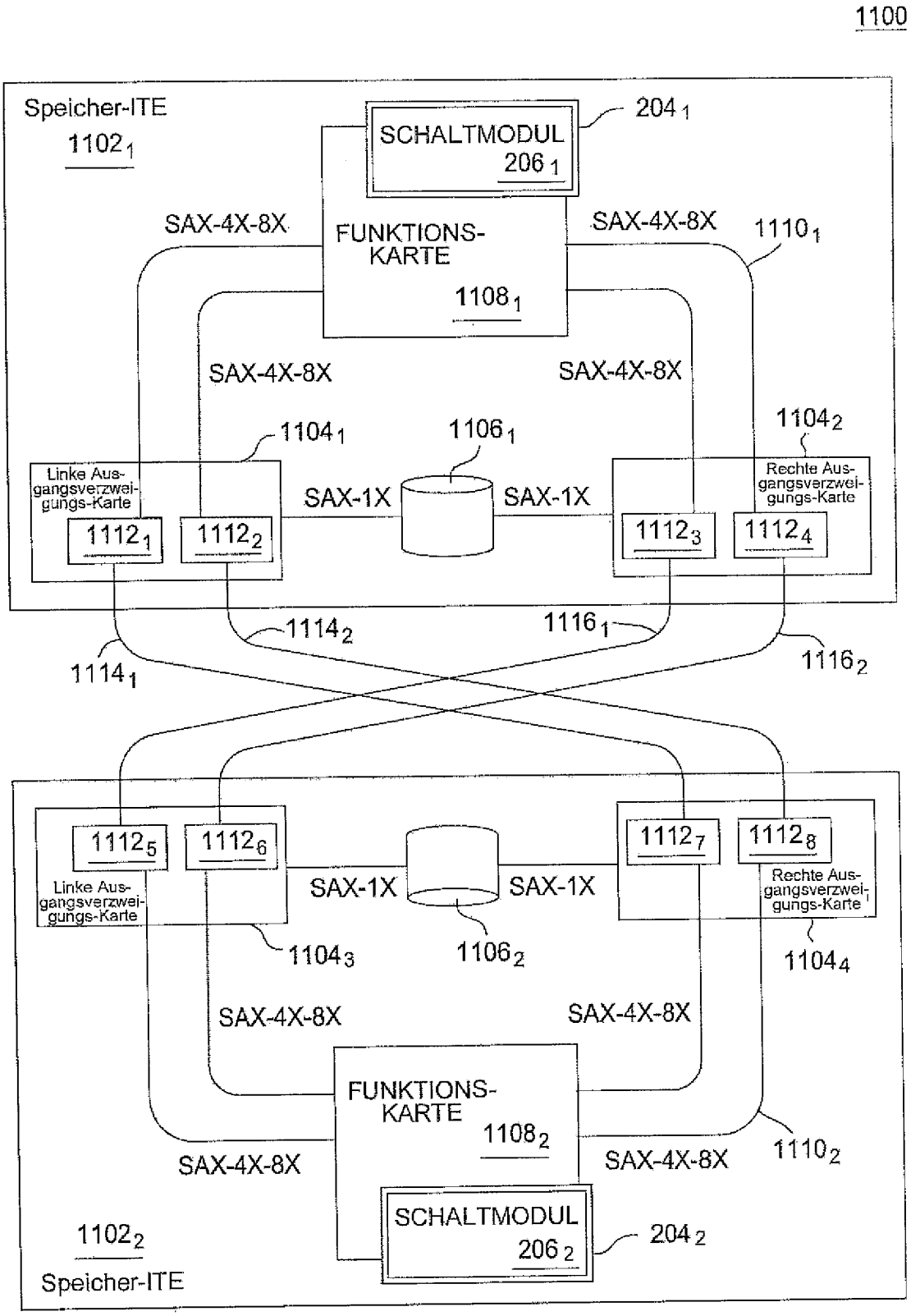
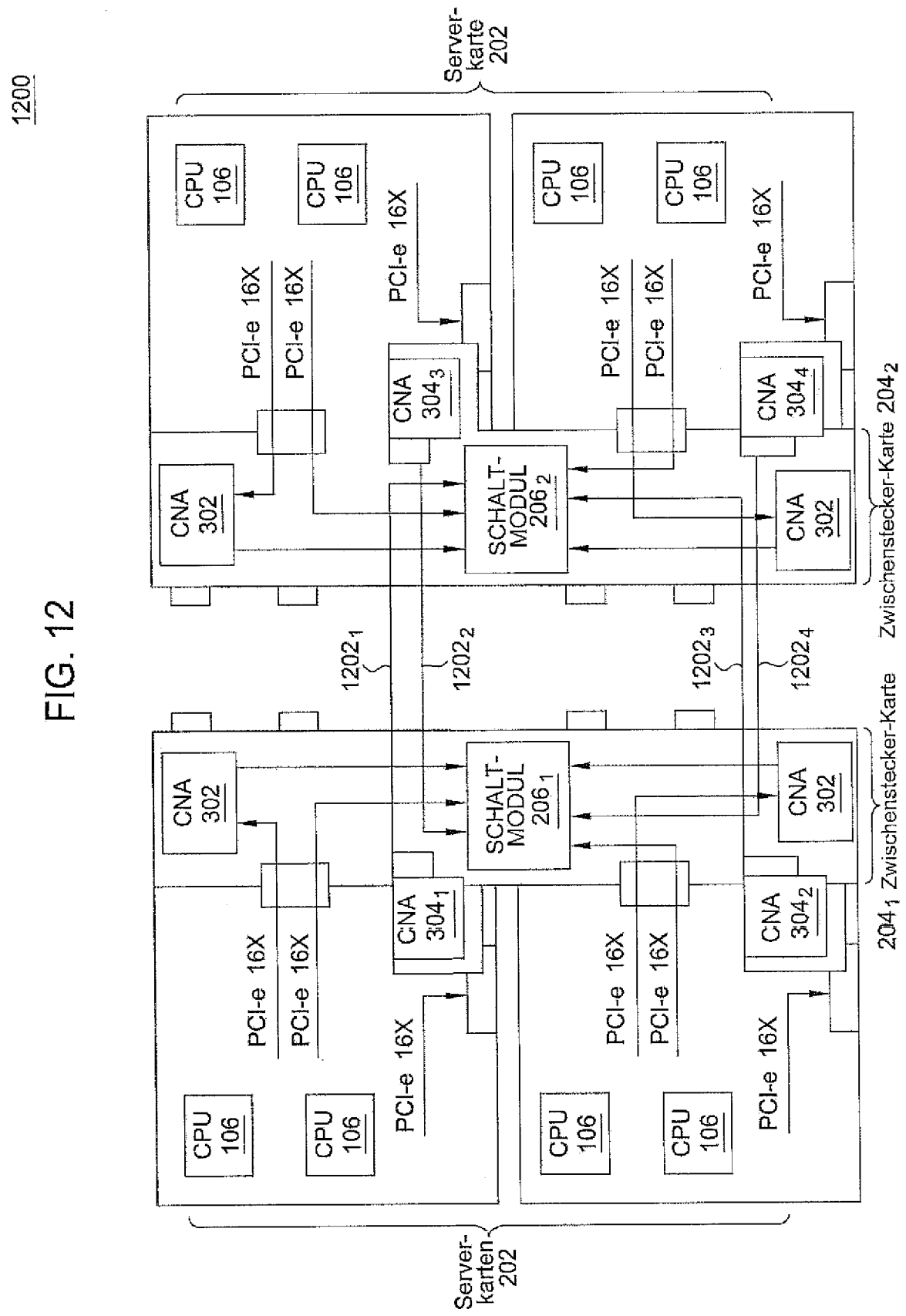

1100
Speicher-ITE
1102 1
SCHALTMODUL
206 1
204 1
SAX-4X-8X
FUNKTIONS-
KARTE
1108 1
SAX-4X-8X
1110 1
SAX-4X-8X
SAX-4X-8X
1104 1
1106 1
1104 2
Linke Aus-
gangsverzwei-
gungs-Karte
SAX-1X
SAX-1X
Rechte Aus-
gangsverzwei-
gungs-Karte
1112 1
1112 2
1112 3
1112 4
1114 2
1116 1
1114 1
1116 2
1112 5
1112 6
1112 7
1112 8
Linke Aus-
gangsverzwei-
gungs-Karte
SAX-1X
SAX-1X
Rechte Aus-
gangsverzwei-
gungs-Karte
1104 3
1106 2
1104 4
SAX-4X-8X
SAX-4X-8X
FUNKTIONS-
KARTE
1108 2
SAX-4X-8X
SAX-4X-8X
1110 2
SCHALTMODUL
206 2
204 2
1102 2
Speicher-ITE

FIG. 11

FIG. 12


1200
Server-karte 202
CPU 106
PCI-e 16X
CNA 302
CNA 304₃
CNA 304₄
SCHALT-MODUL 206₂
Zwischenstecker-Karte 204₂
1202₁
1202₂
1202₃
1202₄
SCHALT-MODUL 206₁
CNA 304₁
CNA 304₂
204₁ Zwischenstecker-Karte
Server-karten 202

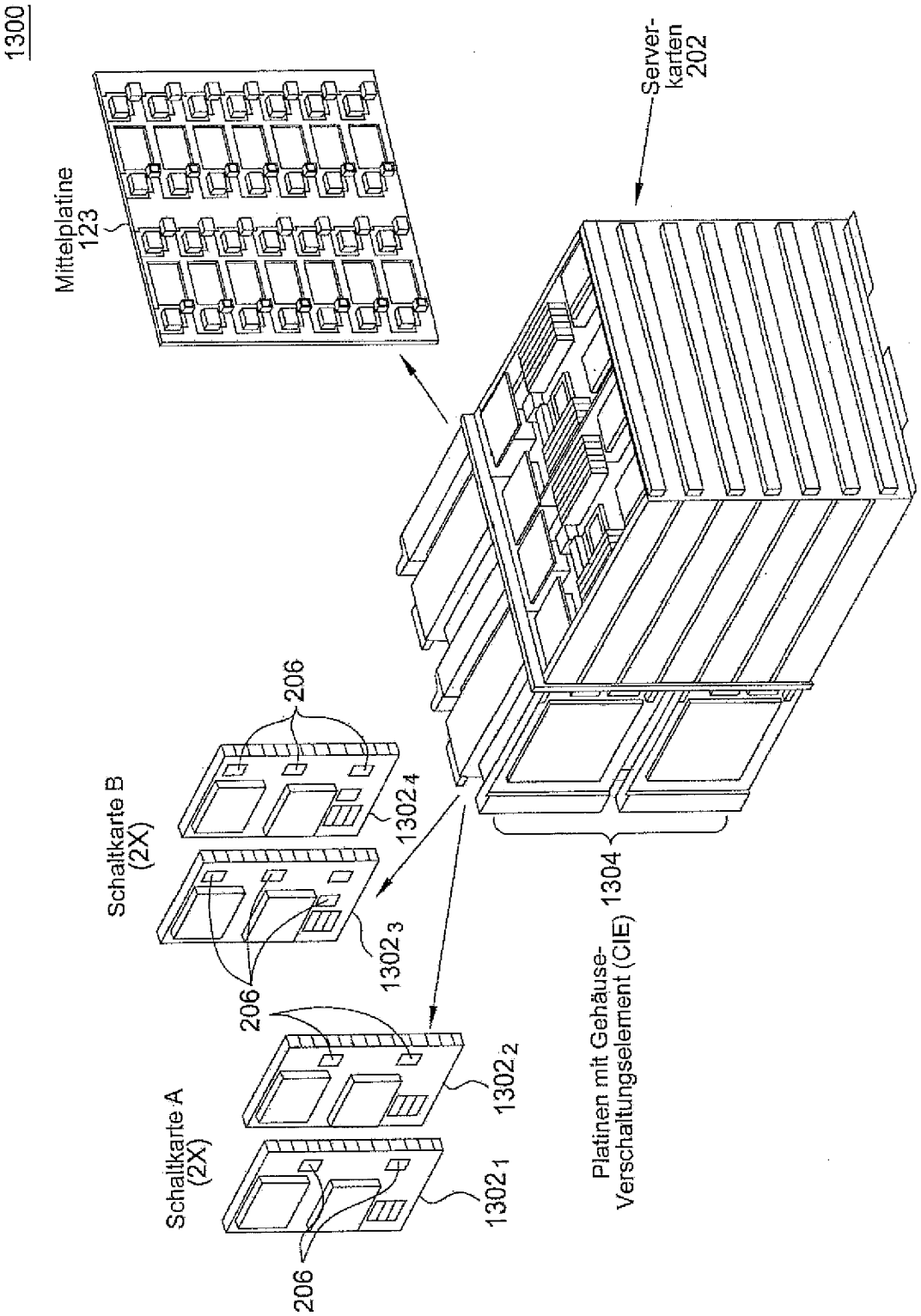

1300
Mittelplatine
123
Server-
karten
202
206
Schaltkarte B
(2X)
1302_4
1302_3
206
206
1302_2
Schaltkarte A
(2X)
1302_1
206
Platinen mit Gehäuse-
Verschaltungselement (CIE) 1304

FIG. 13

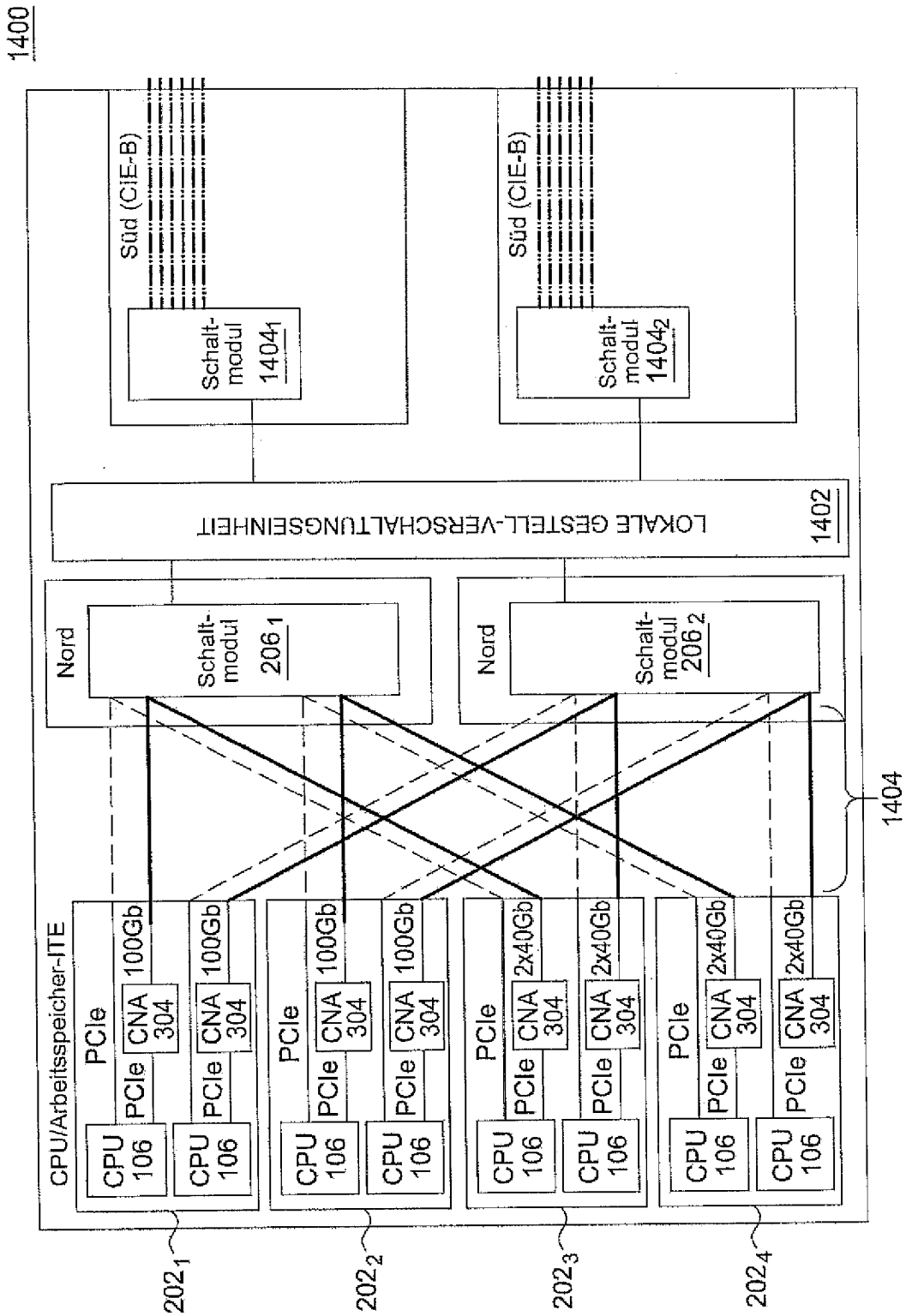

1400
Süd (CIE-B)
Schalt-modul 1404 1
Süd (CIE-B)
Schalt-modul 1404 2
LOKALE GESTELL-VERSCHALTUNGSEINHEIT
1402
Nord
Schalt-modul 206 1
Nord
Schalt-modul 206 2
1404
CPU/Arbeitsspeicher-ITE
PCIe
CNA 304
100Gb
2x40Gb
CPU 106
202 1
202 2
202 3
202 4

FIG. 14

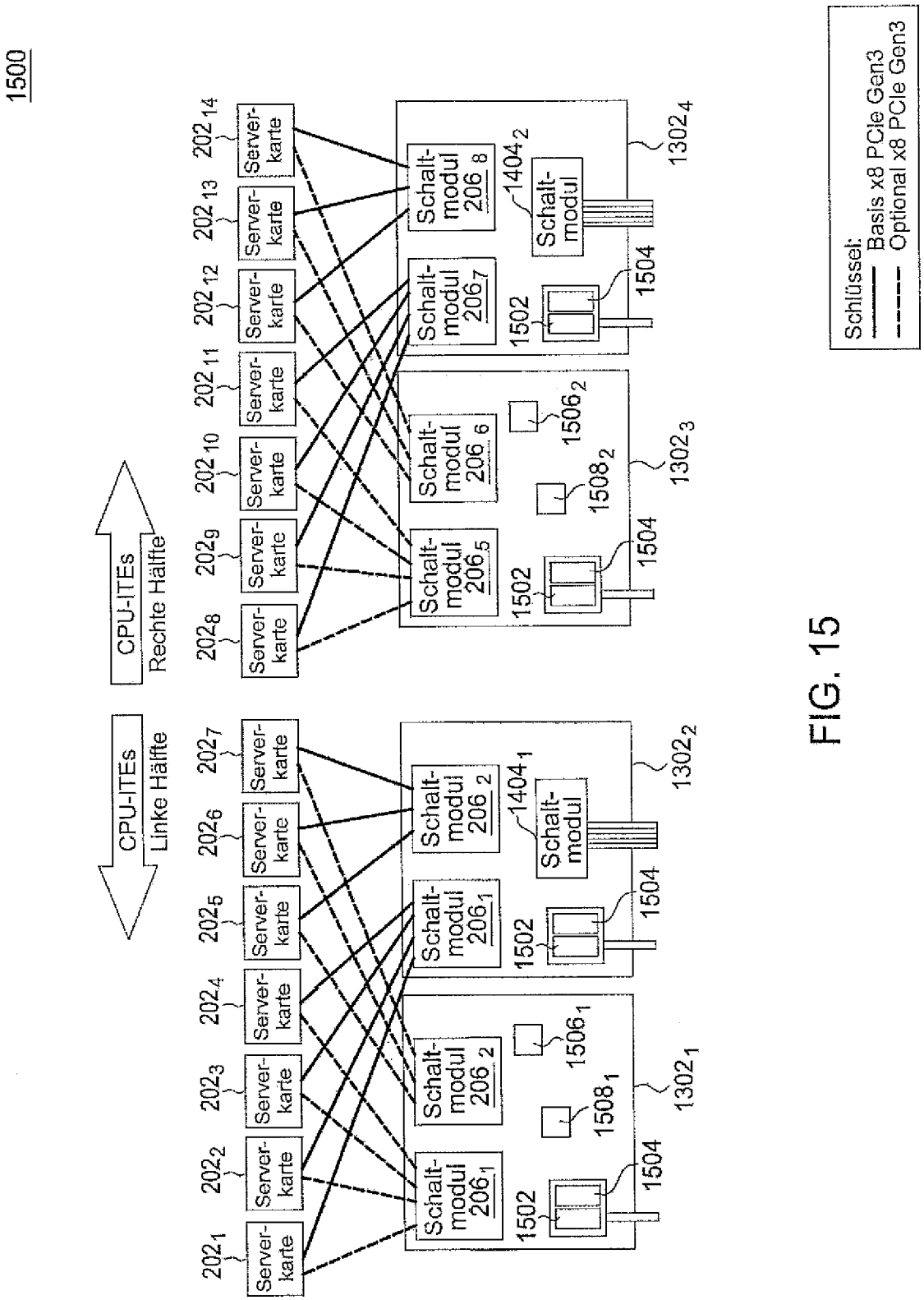

1500
CPU-ITEs
Linke Hälfte
CPU-ITEs
Rechte Hälfte
Server-karte
Schalt-modul
1502
1504
1506
1508
1404
1302
Schlüssel:
Basis x8 PCIe Gen3
Optional x8 PCIe Gen3

FIG. 15

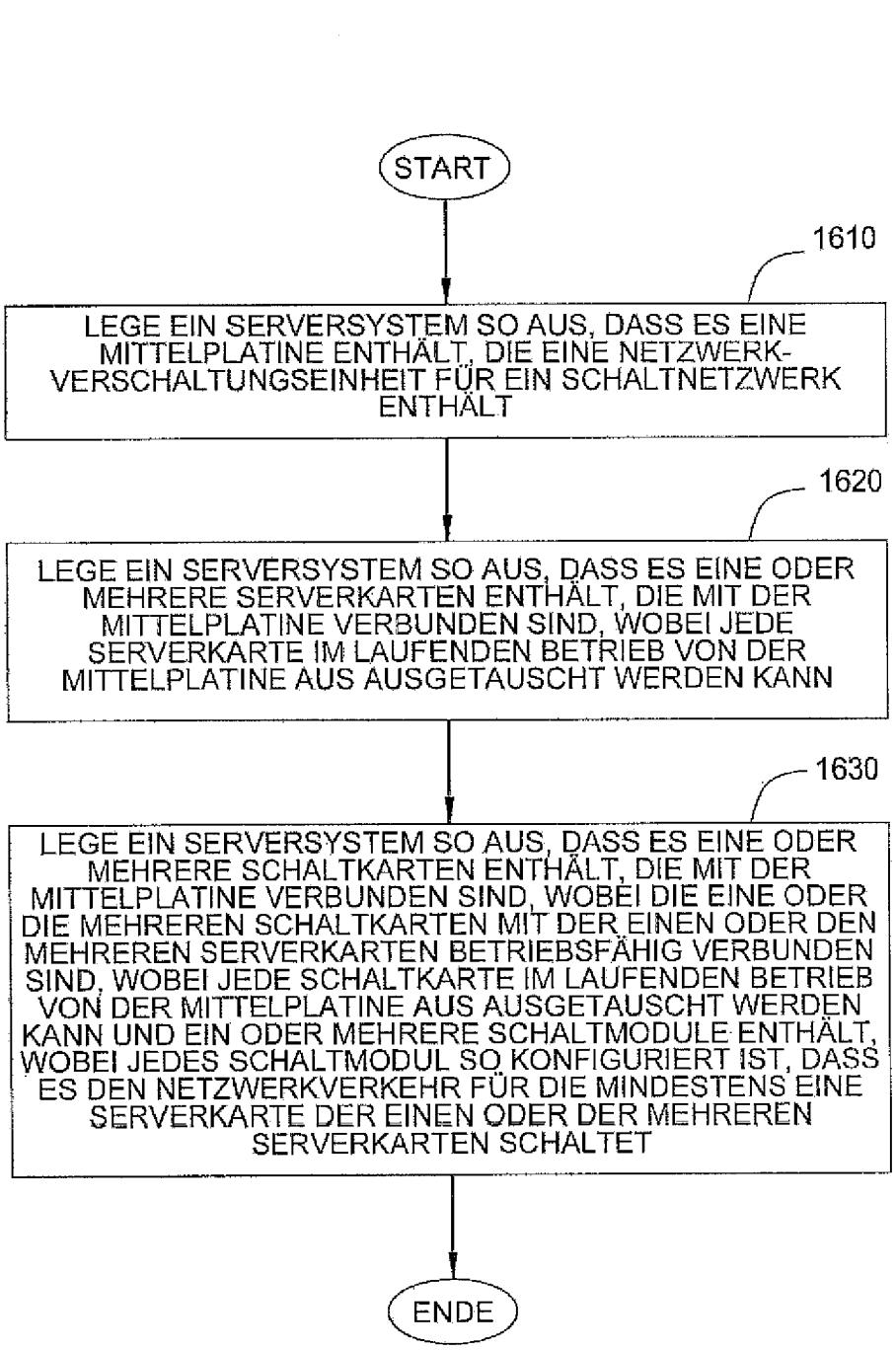

1600
START
1610
LEGE EIN SERVERSYSTEM SO AUS, DASS ES EINE MITTELPLATINE ENTHÄLT, DIE EINE NETZWERK-VERSCHALTUNGSEINHEIT FÜR EIN SCHALTNETZWERK ENTHÄLT
1620
LEGE EIN SERVERSYSTEM SO AUS, DASS ES EINE ODER MEHRERE SERVERKARTEN ENTHÄLT, DIE MIT DER MITTELPLATINE VERBUNDEN SIND, WOBEI JEDE SERVERKARTE IM LAUFENDEN BETRIEB VON DER MITTELPLATINE AUS AUSGETAUSCHT WERDEN KANN
1630
LEGE EIN SERVERSYSTEM SO AUS, DASS ES EINE ODER MEHRERE SCHALTKARTEN ENTHÄLT, DIE MIT DER MITTELPLATINE VERBUNDEN SIND, WOBEI DIE EINE ODER DIE MEHREREN SCHALTKARTEN MIT DER EINEN ODER DEN MEHREREN SERVERKARTEN BETRIEBSFÄHIG VERBUNDEN SIND, WOBEI JEDE SCHALTKARTE IM LAUFENDEN BETRIEB VON DER MITTELPLATINE AUS AUSGETAUSCHT WERDEN KANN UND EIN ODER MEHRERE SCHALTMODULE ENTHÄLT, WOBEI JEDES SCHALTMODUL SO KONFIGURIERT IST, DASS ES DEN NETZWERKVERKEHR FÜR DIE MINDESTENS EINE SERVERKARTE DER EINEN ODER DER MEHREREN SERVERKARTEN SCHALTET
ENDE

FIG. 16

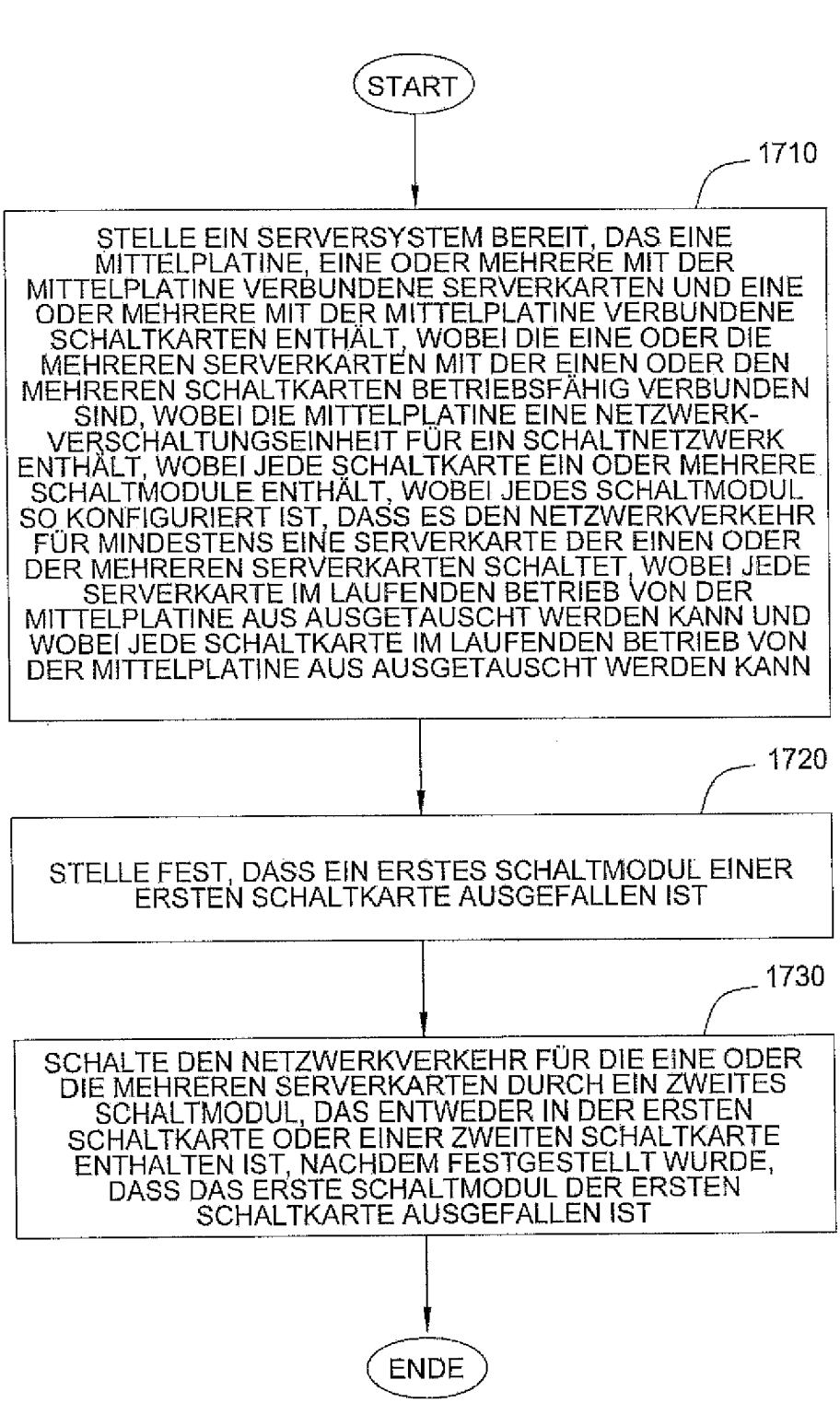

1700
START
1710
STELLE EIN SERVERSYSTEM BEREIT, DAS EINE MITTELPLATINE, EINE ODER MEHRERE MIT DER MITTELPLATINE VERBUNDENE SERVERKARTEN UND EINE ODER MEHRERE MIT DER MITTELPLATINE VERBUNDENE SCHALTKARTEN ENTHÄLT, WOBEI DIE EINE ODER DIE MEHREREN SERVERKARTEN MIT DER EINEN ODER DEN MEHREREN SCHALTKARTEN BETRIEBSFÄHIG VERBUNDEN SIND, WOBEI DIE MITTELPLATINE EINE NETZWERK-VERSCHALTUNGSEINHEIT FÜR EIN SCHALTNETZWERK ENTHÄLT, WOBEI JEDE SCHALTKARTE EIN ODER MEHRERE SCHALTMODULE ENTHÄLT, WOBEI JEDES SCHALTMODUL SO KONFIGURIERT IST, DASS ES DEN NETZWERKVERKEHR FÜR MINDESTENS EINE SERVERKARTE DER EINEN ODER DER MEHREREN SERVERKARTEN SCHALTET, WOBEI JEDE SERVERKARTE IM LAUFENDEN BETRIEB VON DER MITTELPLATINE AUS AUSGETAUSCHT WERDEN KANN UND WOBEI JEDE SCHALTKARTE IM LAUFENDEN BETRIEB VON DER MITTELPLATINE AUS AUSGETAUSCHT WERDEN KANN
1720
STELLE FEST, DASS EIN ERSTES SCHALTMODUL EINER ERSTEN SCHALTKARTE AUSGEFALLEN IST
1730
SCHALTE DEN NETZWERKVERKEHR FÜR DIE EINE ODER DIE MEHREREN SERVERKARTEN DURCH EIN ZWEITES SCHALTMODUL, DAS ENTWEDER IN DER ERSTEN SCHALTKARTE ODER EINER ZWEITEN SCHALTKARTE ENTHALTEN IST, NACHDEM FESTGESTELLT WURDE, DASS DAS ERSTE SCHALTMODUL DER ERSTEN SCHALTKARTE AUSGEFALLEN IST
ENDE

FIG. 17