



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 102 97 283 T5 2005.04.07**

(12)

Veröffentlichung

der internationalen Anmeldung mit der
 (87) Veröffentlichungs-Nr.: **WO 03/030006**
 in deutscher Übersetzung (Art. III § 8 Abs. 2 IntPatÜG)
 (21) Deutsches Aktenzeichen: **102 97 283.4**
 (86) PCT-Aktenzeichen: **PCT/US02/30622**
 (86) PCT-Anmeldetag: **26.09.2002**
 (87) PCT-Veröffentlichungstag: **10.04.2003**
 (43) Veröffentlichungstag der PCT Anmeldung
 in deutscher Übersetzung: **07.04.2005**

(51) Int Cl.7: **G06F 15/167**

(30) Unionspriorität:
09/967,126 28.09.2001 US

(74) Vertreter:
Grosse, Bockhorni, Schumacher, 81476 München

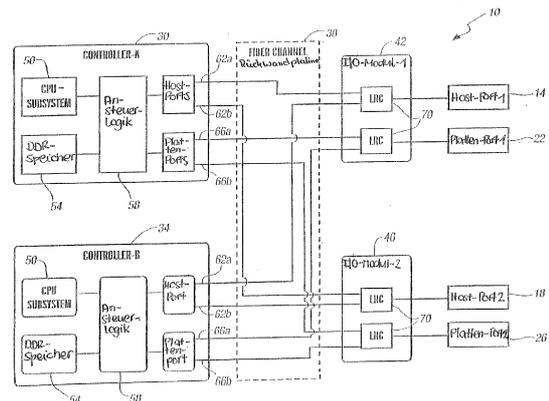
(71) Anmelder:
Dot Hill Systems Corp., Carlsbad, Calif., US

(72) Erfinder:
Pecone, Victor Key, Lyons, Col., US

(54) Bezeichnung: **Controller-Daten-Sharing mittels einer DMA-Architektur**

(57) Hauptanspruch: Netzwerkspeichervorrichtung für die Verbindung eines Host-Rechners mit mindestens einer Speichervorrichtung, umfassend:

- eine passive Rückwandplatine, welche eine Vielzahl von Datenbussen aufweist, einschließlich eines ersten und zweiten Datenbusses;
- mindestens erste und zweite Kanal-Schnittstellenmodule, welche an der passiven Rückwandplatine angeschlossen sind und angepasst sind um an den Host-Rechner und an die mindestens eine Speichervorrichtung angeschlossen zu werden, welche betriebsbereit sind um Speicherdaten zu und von dem Host-Rechner und der mindestens einen Speichervorrichtung zu senden und empfangen und welche betriebsbereit sind um die Speicherdaten an einen oder mehrere der Vielzahl von Datenbussen selektiv zu übertragen; und
- mindestens das erste und zweite Speichercontrollermodul, angeschlossen an der passiven Rückwandplatine, welche mit den Kanal-Schnittstellenmodulen über die passive Rückwandplatine kommuniziert, und welche die Speicherdaten, welche zu und von den Kanal-Schnittstellenmodulen übertragen werden, speichern und verarbeiten;
- wobei zumindest das erste Speichercontrollermodul einen DMA-(Direct-Memory-Access-) Prozessor umfasst und der DMA-Prozessor für die...



Beschreibung

[0001] Die vorliegende Erfindung bezieht sich auf Netzwerkspeichercontroller, und insbesondere auf einen Netzwerkspeichercontroller, der redundante modulare Komponenten für die Datenverarbeitungsfunktionen unabhängig von der I/O-Schnittstelle verwendet.

Stand der Technik

Hintergrund der Erfindung

[0002] Netzwerkspeichercontroller werden typischerweise verwendet, um ein Host-Rechnersystem mit Peripherie-Speichervorrichtungen, wie beispielsweise Plattenlaufwerke oder Bandlaufwerke zu verbinden. Der Netzwerkspeichercontroller agiert wie eine Schnittstelle zwischen dem Host-Rechner und den Peripherie-Speichervorrichtungen. In vielen Anwendungen führt der Netzwerkspeichercontroller Verarbeitungsfunktionen der Daten aus, welche zwischen dem Host-Rechner und den Peripherievorrichtungen übertragen werden. Eine übliche Anwendung eines solchen Systems ist ein Redundant Array of Independent Disks (RAID). Ein RAID-System speichert Daten auf mehreren Plattenlaufwerken, um die Daten vor Plattenlaufwerksausfall zu schützen. Wenn ein Plattenlaufwerk ausfällt, ist das RAID-System im allgemeinen in der Lage die Daten, welche auf dem ausgefallenen Laufwerk gespeichert waren, aus den übriggebliebenen Laufwerken des Arrays wiederherzustellen. Ein RAID-System verwendet einen Speichercontroller, welcher in vielen Fällen einen RAID-Controller umfasst, eine Schnittstelle zwischen dem Host-Rechner und dem Plattenlaufwerks-Array.

[0003] Viele Anwendungen erfordern ein Speichersystem, um eine sehr hohe Verfügbarkeit aufzuweisen. Diese hohe Verfügbarkeit stellt ein Hauptanliegen für viele Anwendungen dar, wie z. B. für Finanzinstitute und Luftfahrtreservierungssysteme, weil der Anwender sich sehr stark auf die im RAID-System gespeicherten Daten verlässt. Bei diesen Anwendungstypen kann die Nichtverfügbarkeit der vom RAID-System gespeicherten Daten zu bedeutenden Einnahmeverlusten und/oder Kundenunzufriedenheit führen. Das Verwenden eines RAID-Systems in solch einer Anwendung erhöht die Verfügbarkeit der gespeicherten Daten, weil beim Ausfall eines einzelnen Plattenlaufwerks, die Daten trotzdem vom System gespeichert und aufgefunden werden können. Zusätzlich zum Einsatz eines RAID-Systems ist es üblich, redundante RAID-Controller zu verwenden, um die Verfügbarkeit eines Speichersystems weiterzuverbessern. In solch einer Situation werden zwei oder mehrere Controller in einem RAID-System verwendet, wobei jeder Controller eine Ausfallsicherungsfähigkeit aufweist, so dass wenn ein Controller ausfällt, der andere übriggebliebene Controller die Funktio-

nen für den ausgefallenen Controller übernimmt. Solch eine Plattform verbessert die Verfügbarkeit eines RAID-Systems, doch es kann zu mehreren Nachteilen führen, wie nachfolgend erörtert.

[0004] Fig. 1 zeigt eine Blockdiagrammdarstellung eines heute üblichen RAID-Netzwerk-Speichercontrollers **10**, der als dualer Controller konfiguriert ist, welches eine Fiber-Channel Fiber-Channel-Verbindung darstellt. Das ist in diesem Beispiel der Host-Rechner und das Disk-Array, die beide mit der Netzwerkspeicher-Bridge über Fiber-Channel-Verbindungen kommunizieren. Während Fiber-Channel ein übliches Kanalmedium für solche Systeme darstellt, versteht es sich, dass beispielsweise andere Kanäle verwendet werden können, wie beispielsweise Small Computer System Interface (SCSI) oder Ethernet. Das in Fig. 1 gezeigte RAID-System, umfasst zwei Host-Ports, Host-Port-1 **14** und Host-Port-2 **18** und zwei Platten-Ports, Platten-Port-1 **22** und Platten-Port-2 **26**. Jeder Host-Port **14**, **18** kann für verschiedene Rechner zugeordnet werden, und jeder Platten-Port **22**, **26** kann unterschiedlichen Platten-Arrays zugeordnet werden, wie es für RAID-Systeme nach dem Stand der Technik bekannt ist. Die Netzwerkspeicher-Bridge **10** umfasst duale RAID-Controller, den Controller-A **30**, und den Controller-B **34**. In einem System, welches das Zoning der Controller verwendet, kann der Controller-A **30** auf den Host-Port-1 **14** und den Platten-Port-1 **22** zoned werden, und der Controller-B **34** kann auf den Host-Port-2 **18** und den Platten-Port-2 **26** zoned werden.

[0005] Im Stand der Technik ist es bekannt, dass Systeme, welche duale Controller mit Write Back Caching verwenden, eine Datenspiegelung zwischen den Controllern erfordern, um die Kohärenz des Cache sicherzustellen. Jeder Controller **30**, **34** muss eine Kopie der Daten und des Status der anderen Controller haben, um eine Redundanz der Controller beizubehalten und dadurch die Funktion des RAID-Systems beizubehalten, wenn einer der Controller ausfällt. Das Spiegeln der Daten zwischen den Controllern kann die Performance eines RAID-Systems verschlechtern, weil die Übertragung der Daten zwischen den Controllern Rechenleistung der Controller verwendet, wie auch Kanalbandbreite, wie nachfolgend detailliert beschrieben.

[0006] Die Controller **30**, **34** sind über eine Fiber Channel Rückwandplatine **38** verbunden, welche zwischen den zwei I/O-Modulen angeschlossen ist, I/O-Modul-1 **42**, und I/O-Modul-2 **46**. Jeder Controller **30**, **34** umfasst ein CPU Subsystem **50**, einen Speicher **54** (beispielsweise Double Data Rate), Steuerlogik **58**, eine Fiber-Channel-Dualport-Verbindung mit den zwei Host-Anschlüssen **62a**, **62b** und eine Fiber Channel-Dualport-Verbindung mit den zwei Plattenanschlüssen **66a**, **66b**. Das CPU-Subsystem **50** führt

die Funktionen aus, die für die Speicherung der Daten in einem Platten-Array notwendig sind, einschließlich des Daten-Striping und dem Veranlassen und Ausführen der Lese- und Schreibbefehle. Der Speicher **54** ist ein nicht flüchtiger Speicherbereich für Daten und andere Informationen. Die Steuerlogik **58** führt mehrere Funktionen aus, die beispielsweise Schnittstellenfunktionen mit dem CPU-Subsystem **50**, Speicher **54** und der Host-Anschlüsse **62a**, **62b** und der Plattenanschlüsse **66a**, **66b**. Die Steuerlogik **58** kann ebenfalls soweit andere Funktionen aufweisen, einschließlich einer Paritätsgenerierungsfunktion, beispielsweise als ein Exclusive-OR-(XOR-) Prozessor. Die Host-Anschlüsse **62a**, **62b** und Plattenanschlüsse **66a**, **66b** ermöglichen die Kommunikation mit der Fiber Channel Rückwandplatine **38**. Die I/O-Module **42**, **46** umfassen LRC-(Link-Resiliency-Circuits-) Schaltungen **70**, ebenfalls bekannt als Port-Bypass-Schaltungen, welche die Funktion haben, jeden Host-Port **14**, **18** und jeden Platten-Port **22**, **26** an jeden Controller **30**, **34** anzuschließen. Dies ermöglicht beiden Controllern **30**, **34** auf die beiden Host-Ports **14**, **18** und auf die beiden Disk-Ports **22**, **26** zuzugreifen.

[0007] Um eine vollständige Redundanz zu bieten, muss jeder Controller eine Verbindung zu jedem Host-Port **14**, **18** und jedem Platten-Port **22**, **26** aufweisen. Auf diese Art kann beim Ausfall eines der Controller der andere Controller den Betrieb fortsetzen. Wie oben beschrieben, ist es üblich, dass jeder Host-Port **14**, **18** verschiedenen Host-Rechnern zugeordnet ist, und dass jeder Platten-Port **22**, **26** unterschiedlichen Disc-Arrays zugeordnet ist. In diesen Fällen ist jeder Controller **30**, **34** typischerweise einem Platten-Port und einem Host-Port zugeordnet, was zur Verbesserung der Performance eines RAID-Systems führt. Dennoch sind in diesem Fall die Hälfte dieser Anschlüsse passiv. Beispielsweise wenn der Controller-A **30** dem Host-Port-1 **14** und dem Platten-Port-1 **22** zugeordnet ist, empfängt der Controller-A **30** die gesamte Kommunikation von dem Host-Port-1 **14** und steuert den Disk-Array (die Disk-Arrays) und den Platten-Port-1 **22**. Ebenso würde der Controller-B **34** dem Host-Port-2 **18** und dem Platten-Port-2 **26** zugeordnet werden. Diese Techniken sind im Stand der Technik bekannt und können die Performance eines RAID-Systems verbessern, wie auch die Steuerung und die Kommunikation der zwei Controller **30**, **34** vereinfachen. Im Beispiel aus **Fig. 1** sind auf den Controller-A **30** die Host-Portverbindungen **62a** und Platten-Portverbindungen **66a** an dem Host-Port-1 **14** und Platten-Port-1 **22** entsprechend angeschlossen, über die LRCs **70** des I/O-Moduls-1 **42** angeschlossen. Weil der Controller-A **30** dem Host-Port-1 **14** und Platten-Port-1 **22** zugeordnet ist, kommunizieren die Host-Portverbindung **62a** und Platten-Portverbindung **66a** aktiv mit dem Host-Port-1 **14** und dem Platten-Port-1 **22**. Die übriggebliebene Host-Portverbindung **62b** und Plat-

ten-Portverbindung **66b** sind an dem Host-Port-1 **18** und Platten-Port-2 **26** über die LRCs **70** des I/O-Moduls-2 **46** angeschlossen. Diese Verbindungen sind typischerweise passive Verbindungen, weil der Controller-A **30** nicht aktiv mit dem Host-Port-2 **18** und dem Platten-Port-2 **26** kommuniziert, solange der Controller-B **34** nicht ausfällt. Ebenso könnte der Controller-B **34** dem Host-Port-2 **18** und Platten-Port-2 **26** zugeordnet werden. Dadurch würde die Host-Portverbindung **62b** und Platten-Portverbindung **66b** für den Controller-B **34** mit dem Host-Port-2 **18** und Platten-Port-2 **26** über die LRCs **70** des I/O-Moduls-2 **46** kommunizieren. Die übriggebliebene Host-Portverbindung **62a** und Platten-Portverbindung **66a** werden an den Host-Port-1 **14** und Platten-Port-1 **22** über LRCs **70** der I/O-Module-1 **42** verbunden.

[0008] Wie oben beschrieben werden bei typischen redundanten Controllerfunktionen mit Write-Back-Caching, Daten zwischen den Controllern gespiegelt. Beim Spiegeln der Daten zwischen dem Controller-A **30** und dem Controller-B **34** ist es üblich, die gespiegelten Daten über die Platten-Portverbindungen zu übertragen. Beispielsweise kann der Controller-B **34** Daten über den Host-Port-2 **18** empfangen, welcher in ein Platten-Array über den Platten-Port **2** geschrieben werden sollen. Der Controller-B **34** würde diese Daten empfangen und im Speicher **54** abspeichern. Um die Kohärenz des Puffers beizubehalten, muss der Controller-B **34** diese Daten ebenfalls dem Controller-A **30** kommunizieren, damit beide Controller die Daten beinhalten, und wenn einer ausfällt, kann der andere die Daten weiterhin abspeichern. In einem herkömmlichen System wird dieser Datentransfer in mehreren Schritten durchgeführt. Zuerst sendet der Controller-B **34** die Daten über die Platten-Portverbindung **66a**, welcher den LRC **70** mit dem Platten-Port-1 **22** verbindet. Die Daten würden dann zu der Art übertragen werden, die dem Platten-Port-1 **22** zugeordnet ist, und dem LRC **70** zurückübertragen werden, wo sie dann an der Platten-Portverbindung **66a** auf dem Controller-A empfangen würden. Der Controller-A würde die Daten dann im Speicher **54** ablegen, eine Kopie der Daten, die ursprünglich dem Controller-B **34** übertragen wurden, bereitstellen. Der Controller-B **34** würde die geeigneten Schritte durchführen, um die Daten auf dem Disk-Array zu schreiben. Sobald die Daten auf den Disk-Array geschrieben wurden, benachrichtigt der Controller-B **34** dies dem Controller-A **30**, wobei derselbe Kommunikationspfad wie oben beschrieben, benutzt wird, und der Controller-A **30** entfernt dann die Eintragung der geschriebenen Daten. Ebenso, wenn der Controller-A **30** Daten empfängt, die auf den Disk-Array am Platten-Port-1 **22** geschrieben werden sollen, sendet er die Daten an den Controller-B **34** unter Verwendung derselben Spiegelungstechnik.

[0009] Während diese Technik den übriggebliebenen Platten-Port auf jedem Controller verwenden kann, bleibt der zweite Host-Port auf jedem Controller unbenutzt, deshalb passiv, während des Normalbetriebs des Systems. Die passiven Anschlüsse auf jedem Controller fügen eine signifikante Menge von Hardware dem Controller hinzu, und können signifikante Zusatzkosten dem Netzwerkspeichercontroller **10** verursachen. Deshalb wäre es vorteilhaft einen redundanten Netzwerkspeichercontroller bereit zu stellen, welcher eine hohe Verfügbarkeit beibehält, während die Kosten und die Hardware, die den passiven Ports auf dem Controller zugeordnet sind, reduziert werden.

[0010] Zusätzlich führt das Spiegeln der Daten in solch einem System dazu, dass die gespiegelten Daten und die gespeicherten Daten über denselben Port des Controllers gesendet werden, wie der, der die gespiegelten Daten empfängt. Bandbreite zu und von dem Disk-Array werden durch die gespiegelten Daten verbraucht, was dazu führen kann, dass die Performance der Netzwerkspeicher-Bridge verringert wird. Zusätzlich werden während der Datenspiegelung Prozessverarbeitungsressourcen des Controllers **30, 34** verbraucht, weil der Controller während des Versendens der Daten diese in eine zu übertragende Form bringen muss, und der Controller, der die Daten empfängt, muss die Daten, die über den Platten-Port angekommen sind, verarbeiten. Beispielsweise werden in dem Fiber Channel Ausführungsbeispiel, welches in **Fig. 1** gezeigt wird, werden die gespiegelten Daten gemäß dem Fiber Channel-Protokoll formatiert, was mehrere Interrupts und Bearbeitungsressourcen erfordert. Daher wäre es vorteilhaft, einen Netzwerkspeichercontroller zu haben, welcher wenig oder keine Kanalbandbreite während der Datenspiegelung zwischen den Controllern benötigt. Es wäre ebenfalls vorteilhaft einen Netzwerkspeichercontroller zu haben, welcher weniger Ressourcen für die Datenspiegelung verbraucht.

[0011] Mit dem kontinuierlichen Ansteigen der Nachfrage nach Datenspeicher erfordern RAID-Controller ferner des öfteren das Aufrüsten zusätzlicher Plattenlaufwerke oder schnelleren Busschnittstellen. Ein RAID-Controller kann aber möglicherweise nicht dazu konfiguriert sein, zusätzliche Busschnittstellenkapazität hinzuzufügen oder einen neuen Busschnittstellentyp zu unterstützen. Solche Controller müssen üblicherweise ersetzt werden, wenn eine Aufrüstung erfolgt. Dieses Ersetzen des Controllers kann die Kosten einer Aufrüstung für ein RAID-System erhöhen. Das Ersetzen eines funktionsfähigen RAID-Controllers stellt einen Wertverlust dar, welcher die Entscheidung, ein RAID-System aufzurüsten, verhindern kann. Daher wäre es vorteilhaft, ein System zu haben, welches sowohl ein leichtes Aufrüsten der Kapazität als auch neue Schnittstellentypen bei reduzierten Kosten unterstützt.

[0012] Entsprechend gibt es eine Notwendigkeit eine Vorrichtung und ein Verfahren zu entwickeln für den Einsatz in einem Netzwerkspeichercontroller, welcher:

- (1) Redundanz bei verringerten Kosten für passive Komponenten bereitstellt,
- (2) die Menge an gespiegelten Daten, welcher über die Platten oder Host-Anschlüsse übertragen werden, reduziert,
- (3) die Verarbeitungsüberlast reduziert, welche für die Datenspiegelung benötigt wird und
- (4) leicht austauschbare und aufrüstbare Komponenten bereitstellt.

Aufgabenstellung

Zusammenfassung der Erfindung

[0013] Gemäß der vorliegenden Erfindung wird eine Vorrichtung und ein Verfahren bereitgestellt, die die Verbindung eines Host-Rechners mit zumindest einer Speichervorrichtung umfassen. Die Vorrichtung umfasst eine passive Rückwandplatine mit einer Anzahl von Datenbussen, einschließlich eines ersten Datenbusses und eines zweiten Datenbusses. Die Vorrichtung umfasst mindestens zwei Kanalschnittstellenmodule, welche mit der passiven Rückwandplatine, dem Host-Rechner und den Speichervorrichtungen verbunden ist. Die Kanalschnittstellenmodule sind betriebsbereit, um Speicherdaten zu und von dem Host-Rechner und den Speichervorrichtungen zu senden und zu empfangen, und die Speicherdaten selektiv an die Datenbusse zu übertragen. Die Vorrichtung umfasst ebenfalls mindestens zwei Speichercontrollermodule, welche an die passive Rückwandplatine angeschlossen sind und den Kanalschnittstellenmodulen über die passive Rückwandplatine kommunizieren, und welche die Speicherdaten speichern und verarbeiten, welche zu und von den Kanalschnittstellenmodulen übertragen werden.

[0014] Das erste Kanalschnittstellenmodul weist einen ersten Switched Path und eine zweiten Switched Path auf, worin die erste Switched Path freigegeben ist, um den ersten Switched Path mit dem ersten Speichercontrollermodul unter Verwendung der passiven Rückwandplatine zu verbinden, und der zweite Switched Path nicht freigegeben ist, um mit dem zweiten Speichercontrollermodul verbunden zu werden. Der zweite Switched Path ist freigegeben und der erste Switched Path ist nicht freigegeben, wenn ein Ausfall des ersten Speichercontrollermoduls mit Hilfe des zweiten Speichercontrollermoduls erkannt wird.

[0015] Die Kanalschnittstellenmodule umfassen einen Kommunikationspfadanteil und einen Kanalschnittstellenanteil. Der Kanalschnittstellenanteil ist betriebsbereit, um die gespeicherten Daten zwischen dem Host-Rechner und/oder den Speichervorrichtungen

gen und dem Kommunikationspfadanteil zu übertragen. Der Kommunikationspfadanteil ist betriebsbereit, um die Daten selektiv zwischen dem Kanalschnittstellenanteil und der passiven Rückwandplatine zu übertragen. In einem Ausführungsbeispiel umfasst das erste Kanalschnittstellenmodul einen ersten Bus-Port und einen zweiten Bus-Port, und das zweite Kanalschnittstellenmodul umfasst einen dritten Bus-Port und einen vierten Bus-Port, wobei jeder der Busanschlüsse betriebsbereit ist, um den Kommunikationspfadanteil an die passive Rückwandplatine anzubinden.

[0016] Die Speichercontrollermodule umfassen einen Busschnittstellenanteil, welcher mit der passiven Rückwandplatine verbunden ist, einen Speicher für das temporäre Speichern der Speicherdaten, und einen Verarbeitungsanteil, welcher die Speicherdaten organisiert und ordnet. Der Busschnittstellenanteil umfasst mindestens eine Rückwandplatten-Schnittstelle, welche mit der passiven Rückwandplatine verbunden ist, eine Speicherschnittstelle, welche mit dem Speicher verbunden ist, einen Verarbeitungsanteil, welcher mit dem Verarbeitungsanteil verbunden ist, einen Bridge-Kern, welcher die Steuerlogik umfasst, welche betriebsbereit ist um die Verarbeitung zur Schnittstelle, Speicherschnittstelle und Rückwandplatten-Schnittstelle zu verbinden. Der Busschnittstellenanteil kann ebenfalls einen explosiven OR (XOR)-Prozessor umfassen, welcher XOR-Funktionen an den Datenblocks vornimmt. In einem Ausführungsbeispiel umfasst der Busschnittstellenanteil ebenfalls einen DMA-(Direct-Memory-Access-) Prozessor, welcher eine DMA-Verbindung zur passiven Rückwandplatine bereitstellt. In einem anderen Ausführungsbeispiel umfasst das erste Speichercontrollermodul eine erste Busschnittstelle und eine zweite Busschnittstelle, und das zweite Speichercontrollermodul umfasst eine dritte Busschnittstelle und eine vierte Busschnittstelle, wobei jede Busschnittstelle betriebsbereit ist um den Busschnittstellenanteil mit der passiven Rückwandplatine zu verbinden.

[0017] Die passive Rückwandplatine umfasst mindestens erste und zweite Datenbusse und in einem Ausführungsbeispiel umfasst sie ebenfalls dritte und vierte Datenbusse. Die Datenbusse auf der passiven Rückwandplatine können PCIX- (next generation Peripheral Component Interconnect) Busse sein. In einem Ausführungsbeispiel ist der erste Bus-Port mit dem ersten Datenbus verbunden und der zweite Bus-Port ist mit dem dritten Datenbus verbunden. Der dritte Bus-Port ist mit dem zweiten Datenbus und der vierte Bus-Port ist mit dem vierten Datenbus verbunden. Die erste Busschnittstelle ist mit dem ersten Datenbus verbunden und die zweite Busschnittstelle ist mit dem zweiten Datenbus verbunden. Die dritte Busschnittstelle ist mit dem dritten Datenbus und die vierte Busschnittstelle ist mit dem vierten Datenbus verbunden.

[0018] Der Kommunikationspfadanteil des ersten Kanalschnittstellenmoduls kann einen ersten Shared Path, einen ersten Switched Path und einen zweiten Switched Path aufweisen, und der Kommunikationspfadanteil des zweiten Kanalschnittstellenmoduls kann einen zweiten Shared Path, einen dritten Switched Path und einen vierten Switched Path aufweisen. In diesem Ausführungsbeispiel ist der erste Shared Path mit dem ersten Bus-Port und dem zweiten Bus-Port verbunden. Der erste Switched Path ist mit dem ersten Bus-Port und dem Kanalschnittstellenanteil verbunden. Der zweite Switched Path ist mit dem zweiten Bus-Port und dem Kanalschnittstellenanteil verbunden. Der zweite Shared Path ist mit dem dritten Bus-Port und dem vierten Bus-Port verbunden. Der dritte Switched Path ist mit dem dritten Bus-Port und dem Kanalschnittstellenanteil verbunden. Der vierte Switched Path ist mit dem vierten Bus-Port und dem Kanalschnittstellenanteil verbunden. Jeder Switched Path ist betriebsbereit um die Kommunikation, welche den Kanalschnittstellenanteil einbezieht, freizugeben oder zu sperren. Ein Verfahren für das Zoning eines Speichercontrollermoduls in ein Kanalschnittstellenmodul wird ebenfalls vorgesehen. Das Verfahren sieht das Bereitstellen eines ersten Kanalschnittstellenmoduls vor, welches einen ersten Switched Path und einen zweiten Switched Path aufweist. Der erste Switched Path ist mit einem ersten Speichercontrollermodul verbunden, unter Verwendung einer passiven Rückwandplatine und der zweite Switched Path ist mit einem zweiten Speichercontrollermodul verbunden unter Verwendung der passiven Rückwandplatine. Der erste Switched Path wird dann freigegeben, und der zweite Switched Path wird gesperrt. Das Verfahren sieht ebenfalls das Bereitstellen eines zweiten Kanalschnittstellenmoduls vor, welches einen dritten Switched Path und einen vierten Switched Path aufweist. Der dritte Switched Path wird mit dem ersten Speichercontrollermodul verbunden, unter Verwendung der passiven Rückwandplatine und der vierte Switched Path wird mit dem zweiten Speichercontrollermodul verbunden unter Verwendung der passiven Rückwandplatine. Der vierte Switched Path wird freigegeben, und der dritte Switched Path wird gesperrt. In einem Ausführungsbeispiel ist das erste Speichercontrollermodul betriebsbereit, um den Ausfall des zweiten Speichercontrollermoduls zu erfassen, und das zweite Speichercontrollermodul ist betriebsbereit, um den Ausfall des ersten Speichercontrollermoduls zu erfassen. Wenn das erste Speichercontrollermodul den Ausfall des zweiten Speichercontrollermoduls erkennt wird die Verwendung des zweiten Speichercontrollermoduls unterbrochen. Der dritte Switched Path wird freigegeben und der vierte Switched Path wird gesperrt.

[0019] In einem Ausführungsbeispiel wird ein Ausfall erkannt, wenn das erste Speichercontrollermodul ein Heartbeat- (Herzschlag-) Signal des zweiten

Speichercontrollermoduls überwacht, und eine Unregelmäßigkeit des Heartbeat feststellt. Ein Ausfallsignal wird dann von dem ersten Speichercontrollermodul an das zweite Speichercontrollermodul gesendet. Das zweite Speichercontrollermodul empfängt das Ausfallsignal und unterbricht den Betrieb. Ebenso überwacht das zweite Speichercontrollermodul ein Heartbeat Signal des ersten Speichercontrollermoduls, und beim Erkennen eines Ausfalls unterbricht es den Betrieb des ersten Speichercontrollermoduls, gibt den zweiten Switched Path frei und sperrt den ersten Switched Path.

[0020] In einem anderen Ausführungsbeispiel erkennt das erste Speichercontrollermodul den Ausfall des ersten Kanalschnittstellenmoduls. Beim Erkennen des Ausfalls unterbricht das erste Speichercontrollermodul das erste Kanalschnittstellenmodul und gibt den dritten Switched Path frei. Das erste Speichercontrollermodul geht dann außer Betrieb. Während des Betriebs kann das erste Kanalschnittstellenmodul eine Laufzeitdiagnose durchführen, wobei deren Ergebnisse von dem ersten Speichercontrollermodul auf Unregelmäßigkeiten überprüft werden. Beim Erkennen eines Ausfalls sendet das erste Speichercontrollermodul ein Ausfallsignal an das erste Kanalschnittstellenmodul und das erste Kanalschnittstellenmodul empfängt das Ausfallsignal und unterbricht seinen Betrieb.

Ausführungsbeispiel

Beschreibung der Zeichnungen

[0021] Fig. 1 ist eine Blockdiagrammdarstellung einer konventionellen Dual-Controller-Netzwerkspeicher-Bridge;

[0022] Fig. 2 ist eine Blockdiagrammdarstellung einer Netzwerkspeichervorrichtung der vorliegenden Erfindung;

[0023] Fig. 3 ist eine Blockdiagrammdarstellung eines Speichercontrollermoduls der vorliegenden Erfindung;

[0024] Fig. 4 ist eine Blockdiagrammdarstellung eines Kanalschnittstellenmoduls der vorliegenden Erfindung;

[0025] Fig. 5 ist eine Blockdiagrammdarstellung einer redundanten Netzwerkspeicher-Bridge der vorliegenden Erfindung;

[0026] Fig. 6 ist eine Blockdiagrammdarstellung einer redundanten Netzwerkspeicher-Bridge, welche ein ausgefallenes Speichercontrollermodul zeigt;

[0027] Fig. 7 ist eine Blockdiagrammdarstellung einer redundanten Netzwerkspeicher-Bridge, welche

ein ausgefallenes Kanalschnittstellenmodul zeigt;

[0028] Fig. 8 ist eine Blockdiagrammdarstellung einer redundanten Netzwerkspeicher-Bridge, welche vier Kanalschnittstellenmodule zeigt;

[0029] Fig. 9 ist eine Blockdiagrammdarstellung einer Netzwerkspeicher-Bridge, welche einen 2GB-Fiber-Channel Verbindungskanal verwendet;

[0030] Fig. 10 ist eine Blockdiagrammdarstellung einer Netzwerkspeicher-Bridge, welche einen Ultra 320 SCSI-Kanal verwendet;

[0031] Fig. 11 ist eine Blockdiagrammdarstellung einer Netzwerkspeicher-Bridge, welche einen Giga-bit-Ethernet-Kanal verwendet.

Detaillierte Beschreibung

[0032] Mit Bezug auf Fig. 2 wird ein Blockdiagramm der Netzwerk-Bridge **100** der vorliegenden Erfindung gezeigt. Die Netzwerk-Bridge **100** umfasst eine oder mehrere Speichercontrollermodule CMMs (Controller Memory Modules). Im Ausführungsbeispiel welches in Fig. 2 gezeigt wird gibt es zwei CMMs, CMM-A **104**, und CMM-B **108**, obwohl ein einziger CMM verwendet werden kann, für Anwendungen in welchen keine Redundanz erforderlich ist, und zusätzliche CMMs werden in Anwendungen verwendet, welche eine zusätzliche Redundanz oder eine höhere Performance erfordern. Jeder CMM **104**, **108** weist zwei Rückwandplatten-Schnittstellen **112** auf. Das System weist eine passive Bus-Rückwandplatte **116** auf, welche zwei Busse für jeden CMM aufweist. In dem gezeigten Ausführungsbeispiel verwendet die Bus-Rückwandplatte **116** PCIX Busse (Peripheral Component Interconnect), wobei es ersichtlich ist, dass jede Bustechnologie verwendet werden kann, einschließlich Switched-Architekturen wie beispielsweise Infiniband oder RapidIO wie auch übliche Busarchitekturen wie beispielsweise PCI Local Bus. Die passive Bus-Rückwandplatte **116** kann einem ersten Datenbus **120**, einem zweiten Datenbus **124**, einem dritten Datenbus **128**, und einem vierten Datenbus **132** aufweisen. Der erste Datenbus **120** und zweite Datenbus **124** werden mit den Rückwandplatten-Schnittstellen **112** auf CMM-A **104** über CMM-Busverbindungen **134** angebunden, und der dritte Datenbus **128** und vierte Datenbus **132** werden mit den Rückwandplatten-Schnittstellen **112** auf CMM-B **108** über CMM-Busverbindungen **134** angebunden.

[0033] In dem in Fig. 2 gezeigten Ausführungsbeispiel werden PCIX-Busse für die passive Rückwandplatte **116** verwendet. Das Verwenden von PCIX-Bussen ermöglicht die Verbindung von verhältnismäßig hochleistungsfähigen Busverbindungskomponenten an die passive Rückwandplatte **116**

mit einem gutbekannten und verhältnismäßig einfachen Busprotokoll. Die PCIX-Technik ist eine Technik der nächsten Generation, welche den üblichen PCI-Bus wirksam einsetzt. Die PCIX-Technik ermöglicht Systeme und Vorrichtungen, welche bei Busfrequenzen von bis zu 133 Mhz arbeiten, unter Verwendung einer 64-Bit oder 32-Bit-Busbandbreite und bei einer Bandbreite von bis zu 1066 Mb/s mit 64-Bit, 133 Mhz PCIX-Bus. Der PCIX-Bus verwendet ein Register zu Registerprotokoll, welches die Taktzwänge im Zusammenhang mit der hohen Frequenz üblicher PCI-Busse erleichtert, und eine höhere Frequenz für den Betrieb des PCIX-Busses ermöglicht. Zusätzlich zu der Fähigkeit die Busfrequenz zu erhöhen, umfasst der PCIX-Bus mehrere Techniken, welche die Buseffizienz verbessern, einschließlich Attribute Phase, Split-Transaction-Unterstützung, optimierte Wait-States und Standardblockgrößenbewegungen.

[0034] Attribute Phase verwendet ein 36-Bit-Attributenfeld, welches die Bustransaktionen detaillierter beschreibt als es die PCI-Spezifikation erlaubt. Sie folgt gleich nach der Adressphase und umfasst mehrere Bit Zuordnungen, welche Informationen über die Größe der Transaktion, Anordnung der Transaktionen, Cache-Snooping-Erfordernissen, und der Identität des Transaktionsinitiators. Mit einer Split-Transaction, wie sie von PCIX unterstützt wird, sendet die Vorrichtung, welche die Daten anfordert ein Signal an das Ziel. Die Zielvorrichtung informiert den Anforderer, das es die Anforderung akzeptiert hat. Der Anforderer kann weitere Informationen bearbeiten, bis die Zielvorrichtung eine neue Transaktion auslöst und die Daten an den Anforderer zuschickt. Dadurch ermöglichen die Split-Transactions eine effizientere Nutzung des Busses. Die Wait-States werden durch den PCIX optimiert, wodurch die Verwendung der Wait-States ausgeschlossen wird, die von üblichen PCIX-Busprotokollen verwendet werden, mit Ausnahme der Initial-Target Latency. Wenn eine PCIX Vorrichtung keine zu übertragende Daten hat, entfernt sie sich vom Bus, so dass eine andere Vorrichtung die Busbandbreite nutzen kann. Dies ermöglicht eine effizientere Busnutzung des Busses und der Speicherressourcen. Mit Standardblockgrößenbewegungen dürfen Adapter und Bridges (Host zu PCIX und PCIX zu PCIX) Transaktionen nur an natürlich angeordneten 128-Bite-Grenzen unterbrechen. Dies ermöglicht längere Signalfolgen und eine effizientere Nutzung von Cache-Line basierenden Ressourcen, wie beispielsweise des Prozessorbusses und Hauptspeichers. Es ermöglicht ebenfalls eine gebündeltere Architektur innerhalb der PCIX Vorrichtungen.

[0035] Die Netzwerk-Bridge **100** weist eine oder mehrere CIMs (Channel Interface Modules) auf. In dem in **Fig. 2** gezeigten Ausführungsbeispiel gibt es zwei CIMs, CIM-1 **136** und CIM-2 **140**, wobei es sich versteht, dass die Anzahl variieren kann, abhängig von der Konfiguration und der Anwendung in welcher

die Netzwerk-Bridge **100** verwendet wird. Jedes CIM **136**, **140** weist zwei CIM-Busschnittstellen-Port **144a**, **144b** auf. Auf jeden CIM **136**, **140** ist ein CIM Busschnittstellen-Port **144a** mit einem Bus verbunden, welcher mit dem CMM-A **104** verbunden ist, und ein CIM-Busschnittstellen-Port **144b** ist mit einem Bus verbunden, welcher mit dem CMM-B **108** über die CIM Busverbindungen **146** angeschlossen ist. In dem in

[0036] **Fig. 2** gezeigten Ausführungsbeispiel ist das CIM-1 **136** mit dem ersten Datenbus **120** und dem dritten Datenbus **128** verbunden, und das CIM-2 **140** ist mit dem zweiten Datenbus **124** und dem vierten Datenbus **132** verbunden. Beide CIM **136**, **140** weisen einen Host-Port **148** auf, welcher mit einem Host-Kanal **152** verbunden ist, welcher an einen Host-Rechner (nicht gezeigt) angeschlossen ist. Die beiden CIMs **136**, **140** weisen ebenfalls einen Platt-Port **156** auf, welcher an einen Platten-Kanal **158** angeschlossen ist, welcher an eine oder mehrere Speichervorrichtungen (nicht gezeigt) angeschlossen ist. In alternativen Ausführungsbeispielen, wie nachfolgend detailliert beschrieben, kann ein CIM nur Host-Ports oder nur Platten-Ports umfassen, abhängig von der Anwendung und den erforderlichen Kanalschnittstellen.

[0037] Wenn der Host-Rechner Daten sendet, werden sie über den Host-Kanal **152** übertragen und an dem Host-Port **148** auf den CIMs **136**, **140** empfangen. Diese Daten werden an die CMMs **104**, **108** über die passive Rückwandplatine **116** gesendet. Die CMMs **104**, **108** enthalten Speicher- und Verarbeitungseinheiten, die nachfolgend detailliert beschrieben werden, welche die Daten in eine geeignete Form zum abspeichern in den Speichervorrichtungen anordnen. Wenn zum Beispiel das System ein RAID **5** Disk Array System verwendet, ordnen die CMMs **104**, **108** die Daten in geeignete Daten-Stripes, welche auf die Platten geschrieben werden sollen, und berechnen einen Paritätsblock für die Daten-Stripes. Dadurch verarbeiten die CMMs **104**, **108** die Daten und formatieren sie zum speichern. Sobald das erfolgt ist, übertragen die CMMs **104**, **108** die für das Speichern bereitgestellten Daten an die CIMs **136**, **140** über die passive Rückwandplatine **116**. Die CIMs **136**, **140** senden dann die Daten an die Speichervorrichtungen, welche an dem Plattenport **156** angeschlossen sind. Wie nachfolgend detailliert beschrieben, können die Daten zwischen den CMMs **104**, **108** unter Verwendung der CIMs **136**, **140** und der passiven Rückwandplatine **116** übertragen werden. Zusätzlich können wie nachfolgend erörtert, die CMMs **104**, **108** und CIMs **136**, **140** für bestimmte Laufwerke oder Hosts zoned werden.

[0038] Diese Konfiguration ermöglicht eine modulare und redundante Architektur, in welcher der Host-Kanal **152** und der Platten-Kanal **158** nicht not-

wendigerweise demselben Kanalmedium angehören. Die Modularität der CMMs **104**, **108** und CIMs **136**, **140** ermöglicht Aufrüstungen zu relativ niedrigen Kosten und ein leichtes Austauschen der ausgefallenen Einheiten. Das Verwenden einer passiven Rückwandplatine **116** um die Daten zwischen den CMMs **104**, **108** auszutauschen, vermeidet das Verwenden der Kanalbandbreite des Platten-Kanals **158** oder des Host-Kanals **152**, wie es bei einer Daten Spiegelung in einer herkömmlichen redundanten Controllerumgebung nötig ist, wie nachfolgend erörtert.

[0039] Mit Bezug auf die **Fig. 3** wird eine Blockdiagrammdarstellung eines CMM **104** gezeigt. Das CMM **104** umfasst mehrere Komponenten, einschließlich eines CPU-Subsystems **160**, eines Speichers **164**, und einer Schnittstelle FPGA **168**. Das CPU-Subsystem **160** kann eine Standardtyp CPU sein, wie beispielsweise ein weitverbreiteter Mikroprozessor, oder es kann ein anwendungsspezifischer Prozessor sein. In einem Ausführungsbeispiel ist das CPU-Subsystem **160** ein Intel Pentium™ Mikroprozessor. Das CPU-Subsystem **160** kommuniziert mit der Schnittstelle FPGA **168** einen Standardbus verwendet, wie beispielsweise einen PCI-Bus. Der Speicher **164** ermöglicht eine temporäre Speicherung der Daten innerhalb der CMM **104**. Diese Speicherung wird während normaler Lese- und Schreibvorgänge für verschiedene Zwecke genutzt, wie beispielsweise speichernd von sogenannten "Köt Data", welche darauf warten auf das Platten Array geschrieben zu werden. In einem Ausführungsbeispiel wird ein DDR-Speicher-DIMM verwendet, welche mit der Schnittstelle FPGA **168** kommuniziert, eine Busschnittstelle verwendend.

[0040] Die Schnittstelle FPGA **168** umfasst eine Anzahl von Komponenten. Es versteht sich, dass diese Komponenten in eine einzige FPGA kombiniert werden können, oder auf mehreren Komponenten innerhalb der CMM **104** angeordnet sein können. In einem in **Fig. 3** gezeigten Ausführungsbeispiel umfasst die Schnittstelle FPGA **168** eine PCI-Schnittstelle **172**, eine Speicherschnittstelle **176**, einen XOR-Prozessor **180**, einen Bridge-Kern **184**, einen DMA-Prozessor **188**, Daten FIFOs **192**, und zwei Rückwandplatine Schnittstellen **112**. Die PCI-Schnittstelle **172** agiert wie eine Schnittstelle zwischen dem CPU-Subsystem **160** und den anderen Bereichen der Schnittstelle FPGA **168**. In dem gezeigten Ausführungsbeispiel verwendet diese Schnittstelle eine Standard-PCI-Busverbindung. Die PCI-Schnittstelle **172** verbindet einen Bridge-Kern **184**, welcher seinerseits an die Rückwandplatinen-Schnittstelle **112** angeschlossen ist, wobei die Schnittstelle den ersten Datenbus **120** und zweiten Datenbus **124** auf der passiven Rückwandplatine **116** miteinander verbindet.

[0041] Die Speicherschnittstelle **176** agiert wie eine

Schnittstelle zwischen dem Speicher **164** und der Schnittstelle FPGA **168**. Der XOR-Prozessor **180** dient dazu, XOR-Transaktionen an den zu speichernden Daten auszuführen, um Paritätsinformationen über die zu schreibenden Daten zu erhalten. Der XOR-Prozessor **180** wird ebenfalls in solchen Fällen genutzt, wo die Verwendung der Paritätsinformationen dazu benötigt wird, um Daten aus einem ausgefallenen Laufwerk in einem Platten-Array wiederherzustellen. Der XOR-Prozessor **180** verbindet das CPU-Subsystem **160** über die PCI-Schnittstelle **172**. Die Daten-FIFOs **192** verbinden die Speicherschnittstelle **176** und den Bridge-Kern **184**, und sind ihrerseits an der Rückwandplatinen-Schnittstelle **112** angeschlossen. Die Daten-FIFOs dienen als eine Warteschlange, welche vom CMM **104** verwendet werden, um die Lese- und Schreibvorgänge zu verwalten. Der DMA-Prozessor **188** dient dazu um die DMA-Daten bereitzustellen und zu empfangen, von einem anderen DMM, wem die CMMs betriebsbereit sind, um Redundanz zu bieten, wie nachfolgend detailliert erörtert.

[0042] Mit Bezug auf **Fig. 4** wird eine Blockdiagrammdarstellung eines CIM **136** gezeigt. Das CIM **136** umfasst einen PCIX FPGA **200** und eine Kanal-Schnittstelle **204**. Der Switched/Shared PCIX FPGA **200** umfasst einen ersten Switched PCIX Path **208** und einen zweiten Switched PCIX Path **212**, und einen Shared PCIX Path **216**. Jeder Switched PCIX Path **208**, **212** ist mit einem Busschnittstellen Port **144** verbunden, welcher seinerseits an einen PCIX Bus auf der passiven Rückwandplatine **116** über eine CIM Busverbindung **146** angeschlossen ist. Jeder Switched PCIX Path **208**, **212** hat ebenfalls einen freigegebenen Eingang **214**. Der freigegebene Eingang **214** wird dafür verwendet, die Switched PCIX Path **208**, **212** freizugeben oder zu sperren. Der Shared PCIX Path **216** umfasst einen Bridge-Kern **220**, welcher entscheidet, welche Daten über den Shared Path **216** geroutet werden, und übergibt diese Daten über den Shared Path **216** zur anderen CIM-Busverbindung **146**. Des gleichen umfasst jeder Switched PCIX Path **208**, **212** ebenfalls einen Bridge-Kern **224**, welcher entscheidet, welche Daten über den Switched Path **208**, **212** geroutet werden, und überträgt diese Daten über die Kanalschnittstelle **204**.

[0043] Die Kanalschnittstelle **204** verbindet den Switched/Shared PCIX FPGA **200** mit dem Host-Kanal **152** und dem Platten-Kanal **158**. Die Kanalschnittstelle umfasst ein Steuerregister **228**, einen Adressensteuerungsteil **232**, eine PCIX-Schnittstelle **236**, einen Host-Port **148**, und einen Platten-Port **156**. Die Steuerregister **228** werden dafür verwendet, die Kommunikation über den Host-Kanal **152** oder Platten-Kanal **158** zu steuern. Der Adresssteuerungsbereich **232** wird dafür verwendet, die Daten an die richtigen Adressen, entweder des Host-Rechners

oder der Speichervorrichtungen zu leiten. Die PCI-X-Schnittstelle **236** funktioniert um die Daten von dem Switched/Shared PCI-X FPGA **200** zu übermitteln, und die Daten an den Host-Port **148** und den Platten-Port **156** zu übertragen. Die PCI-X-Schnittstelle **236** ermöglicht die Kommunikation über das geeignete Kanalmedium für die Anwendung. Beispielsweise wenn der Host-Kanal **152** und der Platten-Kanal **158** Fiber Channel verwenden, arbeitet die PCI-X Schnittstelle **236** als Schnittstelle zwischen dem Switched/Shared PCI-X FPGA **200** und dem Fiber Channel. Desgleichen, wenn der Host-Kanal **152** und der Platten-Kanal **158** einen SCSI-Kanal verwenden, agiert die PCI-X-Schnittstelle **236** als Schnittstelle zwischen dem Switched/Shared PCI-X FPGA **200** und dem SCSI-Kanal. Wenn sowohl der Host-Kanal **152** als auch der Platten-Kanal **158** dasselbe Kanalmedium verwenden, kann das CIM **136** identische Host-Ports **148** und Platten-Ports **156** für die Kommunikation mit sowohl dem Host-Kanal **152** als auch dem Platten-Kanal **158** verwenden.

[0044] In einem Ausführungsbeispiel verwenden der Platten-Kanal **158** und der Host-Kanal **152** nicht das gleiche Kanalmedium. In diesem Ausführungsbeispiel wird je ein unterschiedliches CIM für jedes unterschiedliche Kanalmedium verwendet. Beispielsweise wenn der Host-Rechner Fiber-Channel verwendet, und das Platten-Array einen SCSI-Kanal verwenden, wird der Host-Rechner mit einem CIM unter Verwendung einer Fiber Channel Schnittstelle verbunden, und das Platten-Array wird über ein anderes CIM unter Verwendung einer SCSI-Kanal-Schnittstelle verbunden. Wenn eine Redundanz benötigt wird, können zwei oder mehrere CIMs mit jedem Kanalmedium verbunden werden.

[0045] In dem in **Fig. 4** gezeigten Ausführungsbeispiel kommuniziert der erste Switched PCI-X Path **208** mit dem ersten Datenbus **120** und ein zweiter Switched PCI-X Path **212** kommuniziert mit dem dritten Datenbus **128** über den Busschnittstellen-Port **144** und CIM-Busverbindung **146**. Der Shared PCI-X Path **216** kann als Kommunikationspfad für ein CMM verwendet werden, um mit einem anderen CMM zu kommunizieren, wie nachfolgend detailliert beschrieben. Es versteht sich, dass eine ähnliche Konfiguration verwendet wird, um die übriggebliebenen CIMs, die auf der Netzwerk-Bridge vorhanden sind verwendet wird. Beispielsweise im Ausführungsbeispiel, das in **Fig. 2** gezeigt wird, wird das CIM-2 **140** mit dem zweiten Datenbus **124** und dem vierten Datenbus **132** verbunden, und dadurch hat der Switched/Shared PCI-X FPGA **200**, im CIM-2 **140** enthalten, Switched PCI-X Paths **208**, **212**, welche mit dem zweiten Datenbus **124** und vierten Datenbus **132** entsprechend kommunizieren. Desgleichen, wenn mehr als zwei CIMs vorhanden sind, werden diese so konfiguriert, um mit den geeigneten Bussen auf der passiven Rückwandplatine **116** zu kommunizieren, wie

es die Anwendung erfordert.

[0046] Wieder mit Bezug auf die **Fig. 2-4** wird der Freigabe-Eingang **214** dazu verwendet, einen CIM **136**, **140** auf ein bestimmtes CMM **104**, **108** zu zonen. In solch einem Fall hat ein CMM **104** oder **108** den exklusiven Besitz für ein CIM **136** oder **140** durch Freigabe des Zugangs zu dem CIM **136** oder **140** von dem Bussegment, mit dem das CMM **104** oder **108** verbunden ist. In einem Ausführungsbeispiel wird beispielsweise das CMM-A **104** auf CIM-1 **136** zoned und CMM-B **108** wird auf CIM-2 **140** zoned. Dadurch hat CMM-A **104** den exklusiven Besitz am CIM-1 **136**, und CMM-B **108** hat den exklusiven Besitz am CIM-2 **140**. Dieses Zoning wird durch die Aktivierung des Freigabe-Eingangs **214** in dem ersten Switched PCI-X Path **208** im CIM-1 **136** erreicht, und durch das Sperren des Freigabe-Eingangs **214** im zweiten Switched PCI-X Path **212** im CIM-1 **136**. Das führt dazu, dass nur der erste Switched PCI-X Path **208** mit der Kanal-Schnittstelle **204** kommuniziert. Wie oben erwähnt kommuniziert im CIM-1 **136** der erste Switched PCI-X Path **208** mit dem ersten Datenbus, und dadurch wird das CMM-A **104** auf CIM-1 **136** zoned. Desgleichen wird für das Zoning des CIM-2 **140** auf CMM-B **108** der Freigabe-Eingang **214** im zweiten Switched Path **212** aktiviert und der Freigabe-Eingang auf dem ersten PCI-X Path **208** wird nicht aktiviert. Dies führt dazu, dass nur der zweite Switched PCI-X Path **212** in CIM-2 **140** mit der Kanal-Schnittstelle **204** kommuniziert, und dadurch wird CMM-B **108** auf CIM-2 **140** zoned. Durch Zulassung, dass nur ein CMM **104** oder **108** ein CIM **136** oder **140** steuert, wird die Kanalsteuerung, das Mapping und das Management vereinfacht. Das Zoning ist insbesondere dann nützlich, wenn zwei oder mehrere Host-Kanäle oder Platten-Kanäle vorhanden sind. Beispielsweise, wenn zwei Host-Kanäle und zwei Platten-Kanäle vorhanden sind, kann das CMM-A **104** zu einem ersten Host-Kanal und dem ersten Platten-Kanal zoned werden, und das CMM-B **108** kann zum zweiten Host-Kanal und dem zweiten Platten-Kanal zoned werden. Dadurch wird CMM-A **104** Daten von dem ersten Host-Kanal über CIM-1 **136** senden und empfangen, und CMM-B **108** wird Daten von dem zweiten Host-Kanal über CIM-2 **140** senden und empfangen. Das Verwenden des Zoning in den CMMs vereinfacht die Steuerung in einer aktiven Anwendung, weil die CMMs keine Kohärenzüberprüfung durchführen müssen. Eine Kohärenzüberprüfung ist erforderlich, wenn kein Zoning implementiert wird, weil beide CMMs denselben Kanal beanspruchen würden. In solch einem Fall, muss vor der Durchführung irgendeiner Funktion die Daten betreffend, wie beispielsweise eine Lese- oder Schreibfunktion ein CMM überprüfen, dass der andere CMM keine Funktion an den Daten vornimmt. Diese Kohärenzüberprüfung kann kompliziert zu implementieren sein, und kann die Performance verschlechtern, wegen des zusätzlichen Overheads, das jedes CMM durch-

führen muss.

[0047] Mit Bezug auf **Fig. 5** wird eine Blockdiagrammdarstellung einer Netzwerk-Bridge **100a** gezeigt, die redundante Komponenten aufweist. In diesem Ausführungsbeispiel werden zwei CMMs verwendet, CMM-A **104** und CMM-B **108**. Es werden zwei CIMs verwendet, CIM-1 **136** und CIM-2 **140**. CMM-A **104** und CIM-1 **136** werden beide mit dem ersten Datenbus **120** auf der passiven Rückwandplatine **116** angeschlossen. CMM-A **104** und CIM-2 **140** werden beide mit dem zweiten Datenbus **124** auf der passiven Rückwandplatine **116** verbunden. CMM-B **108** und CIM-1 **136** werden beide mit dem dritten Datenbus **128** auf der passiven Rückwandplatine **116** verbunden. CMM-B **108** und CIM-2 **140** werden beide mit dem vierten Datenbus **132** auf der passiven Rückwandplatine **116** verbunden.

[0048] Wie es für einen Fachmann ersichtlich ist erfordern redundante Controller eine Datenspiegelung zwischen den zwei Controllern, die an dem Speicher-Subsystem angeschlossen sind. Dies ist so wegen der Verwendung eines sogenannten White-Back-Cache, wobei der Controller Daten von dem Host-Rechner erhält, die Daten puffert und eine Meldung an den Host-Rechner schickt, dass die Daten geschrieben worden sind. Dadurch erkennt der Rechner, dass die Daten geschrieben worden sind, wenn sie tatsächlich im Controller gespeichert worden sind und hier darauf warten in den Platten-Array geschrieben zu werden. Um sicherzustellen, dass die Daten bei einem Ausfall nicht verloren gehen, spiegeln redundante Controller diese Daten zum anderen Controller, wodurch sie eine andere Kopie der Daten auf dem Controller aufweisen. Dies ist als Pufferkohärenz bekannt (Cache Coherency). In einem Ausführungsbeispiel spiegeln die CMMs **104, 108** die Daten, um für die Netzwerk-Bridge **100a** eine Pufferkohärenz zu bieten. Dies kann durch die Umsetzung eines DMA-Pfades zwischen CMM-A **104** und CMM-B **108** erfolgen. Dies kann dadurch erreicht werden, durch die Bereitstellung eines DMA-Prozessors **188** in der Schnittstelle FPGA **168**, wie oben mit Bezug auf **Fig. 3** erörtert, und einem Shared Path **216**, welcher in dem Switched/Shared Path FPGA **200** angeordnet ist, wie oben mit Bezug auf **Fig. 4** erörtert wurde. Jedes CMM **104, 108** verwendet diesen DMA-Pfad zum senden und empfangen von Daten von dem anderen CMM. Durch die Verwendung des DMA-Pfades können die beiden CMMs **104, 108** Daten spiegeln, ohne die Verwendung des Host-Kanals **152**, oder des Platten-Kanals **158** zu benötigen, wodurch keine Bandbreite im Platten-Kanal **158** oder Host-Kanal **152** durch die CMMs **104, 108** für Datenspiegelung zu verbrauchen. Zusätzlich werden durch die Verwendung eines DMA-Pfades zwischen zwei CMMs **104, 108** weniger Prozessressourcen von den CMMs **104, 108** benötigt, um das Spiegeln auszuführen, als beim Spiegeln unter Verwendung des

Host-Kanals **152** oder des Platten-Kanals **158**.

[0049] Außerdem ist eine Fale Over Recet Link **240** (Ausfallrücksetzverbindung) zwischen CMM-A **104** und CMM-B **108** vorhanden. Die Ausfallrücksetzverbindung **240** wird für die Mitteilung eines Ausfalls eines der CMMs **104, 108** verwendet. In einem Ausführungsbeispiel ist die Ausfallrücksetzverbindung **240** eine serielle Verbindung zwischen CMM-A **104** und CMM-B **108**. In diesem Ausführungsbeispiel hält jeder CMM **104, 108** ein Herzschlagssignal bei, welches über die Ausfallrücksetzverbindung **240** übertragen wird, und von dem anderen CMM überwacht wird. Wenn ein Problem im Herzschlagssignal erkannt wird, kann ein CMM **104, 108** ein Signal über die Ausfallrücksetzverbindung senden, um den Betrieb des anderen CMM zu beenden. Beispielsweise wenn der CMM-B **108** einen Ausfall aufweist, erkennt das CMM-A **104**, dass das Herzschlagssignal vom CMM-B **108** nicht mehr aktiv ist. Nach einer vorgegebenen Zeitspanne, in welcher kein Herzschlagssignal empfangen wird, sendet das CMM-A **104** ein Abbruchsignal an das CMM-B **108**. Sobald das CMM-B **108** das Abbruchsignal empfängt, beendet es den Betrieb. Das CMM-A übernimmt die Steuerung sämtlicher Lese- und Schreibvorgänge. Desgleichen, wenn das CMM-A **104** ausfällt würde das CMM-B **108** den Hinweis über die Ausfallrücksetzverbindung **240** empfangen, und die Steuerung sämtlicher Lese- und Schreibvorgänge übernehmen. Dadurch ist das System redundant und setzt den Betrieb fort, wenn ein CMM **104** oder **108** ausfällt.

[0050] Mit Bezug auf **Fig. 6** wird der Betrieb des Systems beim Ausfall eines CMMs nachfolgend beschrieben. Wie in **Fig. 6** gezeigt weist die Netzwerk-Bridge **100a** ein CMM-A **104** und CMM-B **108**, eine passive PCIX Rückwandplatine **116**, und ein CIM-1 **136** und ein CIM-2 **140** auf. Wenn CMM-A **104** ausfällt, erkennt CMM-B **108** den Ausfall über die Ausfallrücksetzverbindung **240**, wie oben beschrieben, und beendet den Betrieb auf CMM-A **104**. CMM-B **108** übernimmt die Steuerung für den gesamten Speicher und die Steuerung der Transaktionen, die bisher von CMM-A **104** geleitet wurden. Sobald das geschieht sendet CMM-B **108** einen Befehl an CIM-1 **136** und CIM-2 **140**, um lediglich die Kommunikation mit CMM-B **108** freizugeben. In diesem Fall würde CIM-1 **136** den Befehl erhalten, und den ersten Switched Path **208**, der an dem ersten Datenbus **120** angeschlossen ist sperren, und den zweiten Switched Path **212** freigeben, der an den dritten Datenbus **128** angeschlossen ist, wodurch CMM-B **108** an den Host-Port **148** und Platten-Port **156** auf CIM-1 **136** angeschlossen wird. CIM-2 **140** empfängt außerdem den Befehl von CMM-B **108**, und führt dieselbe Funktion aus, um den ersten Switched Path **208** zu sperren, der am zweiten Datenbus **124** angeschlossen ist, und gibt den zweiten Switched Path **212** frei, der an dem vierten Datenbus **132** angeschlossen ist.

In einem Ausführungsbeispiel umfasst die passive Rückwandplatine **116** logische Steuerverbindungen, welche mit den Freigabe-Eingängen **214** auf den CIMs **136**, **140** verbunden sind, und mit den CMMs **104**, **108** verbunden sind. Die CMMs **104**, **108** können diese logischen Steuerverbindungen verwenden, um die Switched Paths **208**, **212** und die CIMs **136**, **140** freizugeben oder zu sperren. Alternativ können andere Ausführungsbeispiele verwendet werden, um die Switched Paths **208**, **212** freizugeben oder zu sperren, wie beispielsweise eine Steuerlogik innerhalb des CIM, welches die Steuerinformationen über die PCIX, Busse auf der passiven Rückwandplatine **116** beispielsweise erhält.

[0051] Mit Bezug auf **Fig. 7** wird der Betrieb des Systems beschrieben, wenn ein CIM ausfällt. Die CMMs **104**, **108** führen periodisch eine Ablaufdiagnose durch, welche überprüft, ob sämtliche Komponenten innerhalb des CIM **136**, **140** richtig funktionieren. Die Ergebnisse dieser Ablaufdiagnose werden den CMMs **104**, **108** über die logischen Steuerverbindungen übermittelt. Das CMM **104**, **108**, welches auf ein CIM **136**, **140** zoned ist, überwacht diese Ablaufdiagnose, und in dem Fall eines Fehlers oder Ausfalls, übermittelt es einen Befehl über die logische Steuerlinie und den Betrieb dieses CIM **136** oder **140** zu beenden. Wie in

[0052] **Fig. 7** dargestellt ist CIM-1 **136** ausgefallen. CMM-A **104** erkennt das CIM-1 **136** ausgefallen ist, und sperrt CIM-1 **136**. CMM-A **104** übermittelt diese Information an CMM-B **108** über den Shared Path **216** auf CIM-2 **140**. CMM-B **108** empfängt diese Information und übernimmt die Steuerung der gesamten Kommunikation zwischen dem Host-Port **148** und Platten-Port **156** auf CIM-2 **140**. CMM-A **104** verbleibt in einem passiven Zustand bis CIM-1 **136** ausgetauscht wurde, oder der Ausfall beseitigt wurde.

[0053] Mit Bezug auf **Fig. 8** wird ein Blockdiagramm einer Netzwerk-Bridge **100b** gezeigt, in welchem vier CIM Module vorhanden sind. In diesem Ausführungsbeispiel kommunizieren zwei CMMs, CMM-A **104** und CMM-B **108** mit vier CIMs, CIM-1 **136**, CIM-2 **140**, CIM-3 **300** und CIM-4 **304**. In diesem Ausführungsbeispiel sind die CMM-A Switched Paths **208** auf CIM-1 **136** und CIM-2 **140** freigegeben. Desgleichen sind in CIM-3 **300** und CIM-4 **304**, die CMM-B Switched Paths **212** freigegeben. Daher bieten CIM-1 **136** und CIM-2 **140** die Schnittstelle zwischen CMM-A **104** und dem Host-Port **148** und der Platten-Port **156** und CIM-3 **300** und CIM-4 **304** bieten eine Schnittstelle zwischen CMM-B **108** und dem Host-Port **148** und Platten-Port **156**. Eine Netzwerk-Bridge dieses Ausführungsbeispiels ist für mehrere Fälle nützlich, beispielsweise wenn mehrere Hosts vorhanden sind. In diesem Ausführungsbeispiel ermöglichen CIM-1 **136** und CIM-2 **140** die Kommunikation zu einem ersten Host, und CIM-3

300 und CIM-4 **304** ermöglichen die Kommunikation zu einem zweiten Host. Dieselbe Anordnung kann für mehrere Platten-Knoten erfolgen, beispielsweise zwei separate RAID-Arrays. Es versteht sich, dass diese Konfiguration ein skalierbares System bietet, welche die Kommunikation zwischen einem oder mehreren Host-Knoten und einem oder mehreren Platten-Knoten bietet, wobei ebenfalls ein redundanter Betrieb ermöglicht wird. Zusätzlich kann solch ein Ausführungsbeispiel nützlich sein für die Verbindung von Hosts und/oder Platten-Arrays, welche unterschiedliche Kanalmedien verwenden. Ein existierendes System kann beispielsweise zwei CIMs aufweisen und Fiber Channel-Verbindungen sowohl für den Host als auch für die Platten-Kanäle verwenden. Wenn ein Anwender das System aufrüsten möchte, um zusätzlich ein weiteres Disk-Array hinzuzufügen, welches eine SCSI-Verbindung verwendet, könnten zusätzliche CIMs hinzugefügt werden, welche die Kommunikation mit einem SCSI-Kanal ermöglichen, wodurch das Aufrüsten eines existierenden Systems ermöglicht wird, ohne die existierende Hardware zu ersetzen.

[0054] Mit Bezug auf die **Fig. 9–11**, werden mehrere alternative Ausführungsbeispiele für einen CIM gezeigt, um ein Beispiel unterschiedlicher Konfigurationen eines CIM zu bieten, und die unterschiedlichen Kanalmedien, mit denen ein CIM verbunden werden kann. **Fig. 9** zeigt eine Blockdiagrammdarstellung eines von PCIX auf Gigabit Interconnect (GBIC) konfigurierten CIM **136a**. Innerhalb des CIM **136a**, ist das Switched/Shared PCIX FPGA **200** mit einer Dual-Port 2Gb-Fiber-Channel-Schnittstelle **400** verbunden. Jeder Port der zweifach Port 2Gb-Fiber-Channel-Schnittstelle **400** ist mit einem SERDES (Serializer/Deserializer) **404a**, **404b** verbunden. Jeder SERDES **404a**, **404b** ist mit dem Kanalmedium mithilfe einer 2Gb-Fiber-Channel-Verbindung verbunden. In dem in **Fig. 9** gezeigten Ausführungsbeispiel ist ein SERDES **404a** mit einem GBIC-Host-Kanal **152a** verbunden, und der andere SERDES **404b** ist mit einem GBIC-Platten-Kanal **158a** verbunden.

[0055] **Fig. 10** zeigt eine Blockdiagrammdarstellung eines PCIX zu SCSI CIM **136b**. Innerhalb des CIM **136b**, ist das Switched/Shared PCIX FPGA **200** mit der zweifach Port Ultra **320** SCSI-Schnittstelle **408** verbunden. Jeder Port der zweifach Port Ultra SCSI Schnittstelle **408** ist mit einem Host oder Platten-Kanal verbunden, und weist ebenfalls eine Verbindung zu einem Abschluss **412**, der für SCSI-Systeme erforderlich ist auf. In dem in **Fig. 10** gezeigten Ausführungsbeispiel ist ein Port, der Dual Port Ultra **320** SCSI Schnittstelle **408** mit einem VH-DIC-Host-Kanal **152b** (Very High Density Interconnect) verbunden, und ein Port der Dual Port Ultra **320** SCSI Schnittstelle **408** mit einem VHDIC-Platten-Kanal **158b** verbunden.

[0056] Fig. 11 zeigt eine Blockdiagrammdarstellung eines PCIX auf Ethernet CIM **136c**, welcher Quick-Switch-Verbindungen **416a**, **416b** in den Switched-Paths verwendet. Die Quick-Switch-Verbindungen **416a**, **416b** sind Bus-Relays welche Freigabe-Eingänge umfassen, welche dazu dienen die Quick-Switch-Verbindungen **416a**, **416b** freizugeben und zu sperren. Jede Quick-Switch-Verbindung **416a**, **416b** ist mit einer Schnittstellen-Verbindung **420** verbunden, welche eine FPGA-Beschleunigung und Daten-FIFOs umfasst. Die Schnittstellenverbindung **420** ist an ein Gigabit-Ethernet-ASIC **424** angeschlossen, welches eine geeignete Verarbeitung an den Daten vornimmt, um die Daten über eine Ethernet-Verbindung zu übertragen. Das Gigabit-Ethernet-ASIC **424** ist mit einem MAC/physischen Konverter **428** verbunden, welches das Signal in ein physikalisches Signal umwandelt, welches zu einem Transformator **432** geroutet wird, um ein Signal mit einer geeigneten Spannung aufzugeben. In einem Ausführungsbeispiel ist der Transformator **432** mit einer GBIC-Verbindung mit einem Platten-Kanal **158c** verbunden. In dem Ausführungsbeispiel aus Fig. 11, wenn ein redundantes System benötigt wird, werden Shared-Paths auf anderen CIMs bereitgestellt. Es versteht sich, dass unterschiedliche Kanalmedien in einem einzigen System verwendet werden können, durch die Anwendung einer Kombination von unterschiedlichen Schnittstellenmodulen, wie beispielsweise jene, die in den Fig. 9–11 gezeigt werden. Beispielsweise kann ein Host-Rechner mit einer Netzwerk-Bridge durch die Verwendung eines Fiber-Channel-Mediums verbunden werden, und die Netzwerk-Bridge kann mit einem Platten-Array durch die Verwendung eines SCSI-Kanalmediums verbunden werden.

Zusammenfassung

[0057] Ein Netzwerkspeichercontroller für das Übertragen von Daten zwischen einem Host-Rechner und einer Speichervorrichtung, wie beispielsweise ein RAID (Redundant Array of Independent Disks), wird offenbart. Der Netzwerkspeichercontroller umfasst wenigstens ein Kanalschnittstellenmodul, welches für den Anschluss an den Host-Rechner und die Speichervorrichtung vorgesehen ist. Das Kanalschnittstellenmodul ist an eine passive Rückwandplatine angeschlossen und überträgt selektiv Daten zwischen dem Host-Rechner und der Speichervorrichtung und der passiven Rückwandplatine. Der Netzwerkspeichercontroller umfasst ebenfalls ein Speichercontrollermodul, welches an der passiven Rückwandplatine angeschlossen ist. Das Speichercontrollermodul kommuniziert mit dem Kanalschnittstellenmodul über die passive Rückwandplatine, und verarbeitet und speichert die Daten temporär, die von dem Host-Rechner oder der Speichervorrichtung empfangen werden. In Anwendungen, in denen eine Redundanz erforderlich ist, werden wenigstens zwei

Speichercontrollermodule und wenigstens zwei Kanalschnittstellenmodule verwendet. Die Speichercontrollermodule können Daten untereinander spiegeln, unter Verwendung der passiven Rückwandplatine und eines Übertragungs-Shared-Path auf das Kanalschnittstellenmodul, wodurch im Wesentlichen das Verwenden des Host- oder der Diskkanäle vermieden wird, um die Daten zu spiegeln. Die Kanalschnittstellenmodule sind funktionsfähig, um den Host-Rechner oder die Speichervorrichtung mit einem oder mehreren Speichercontrollermodulen zu verbinden. Die Speichercontrollermodule können eine DMA-Engine umfassen, um die Übertragung der gespiegelten Daten zu erleichtern.

Patentansprüche

1. Netzwerkspeichervorrichtung für die Verbindung eines Host-Rechners mit mindestens einer Speichervorrichtung, umfassend:

- eine passive Rückwandplatine, welche eine Vielzahl von Datenbussen aufweist, einschließlich eines ersten und zweiten Datenbusses;
- mindestens erste und zweite Kanal-Schnittstellenmodule, welche an der passiven Rückwandplatine angeschlossen sind und angepasst sind um an den Host-Rechner und an die mindestens eine Speichervorrichtung angeschlossen zu werden, welche betriebsbereit sind um Speicherdaten zu und von dem Host-Rechner und der mindestens einen Speichervorrichtung zu senden und empfangen und welche betriebsbereit sind um die Speicherdaten an einen oder mehrere der Vielzahl von Datenbussen selektiv zu übertragen; und
- mindestens das erste und zweite Speichercontrollermodul, angeschlossen an der passiven Rückwandplatine, welche mit den Kanal-Schnittstellenmodulen über die passive Rückwandplatine kommuniziert, und welche die Speicherdaten, welche zu und von den Kanal-Schnittstellenmodulen übertragen werden, speichern und verarbeiten;
- wobei zumindest das erste Speichercontrollermodul einen DMA-(Direct-Memory-Access-) Prozessor umfasst und der DMA-Prozessor für die Übertragung der Daten zwischen dem ersten Speichercontrollermodul und dem zweiten Speichercontrollermodul verwendet wird.

2. Vorrichtung nach Anspruch 1, wobei:

- zumindest das erste Kanal-Schnittstellenmodul einen Kommunikationspfadanteil und einen Kanal-Schnittstellenanteil umfasst, wobei der Kanal-Schnittstellenanteil betriebsbereit ist um die Speicherdaten zwischen dem Host-Rechner und/oder zu der mindestens einen Speichervorrichtung und dem Kommunikationspfadanteil zu übertragen, und der Kommunikationspfadanteil betriebsbereit ist, um die Speicherdaten zwischen dem Kanal-Schnittstellenanteil und der passiven Rückwandplatine selektiv zu übertragen.

3. Vorrichtung nach Anspruch 1, wobei:

- zumindest das erste Speichercontrollermodul einen Busschnittstellenanteil, der an der passiven Rückwandplatine angeschlossen ist, einen Speicher für das temporäre Speichern der Speicherdaten, und einen Verarbeitungsanteil, der die Speicherdaten organisiert und anordnet, umfasst.

4. Vorrichtung nach Anspruch 3, wobei der Busschnittstellenanteil folgendes umfasst:

- mindestens eine Rückwandplatinenschnittstelle, welche mit der passiven Rückwandplatine verbunden ist;
- eine Speicherschnittstelle, welche mit dem Speicher verbunden ist;
- eine Verarbeitungsschnittstelle, welche mit dem Verarbeitungsanteil verbunden ist;
- einen Bridge-Kern, welcher die Ansteuerlogik umfasst, die betriebsbereit ist die Verarbeitungsschnittstelle, Speicherschnittstelle und Rückwandplatinenschnittstelle zu verbinden; und
- einen Exklusive-OR-(XOR-) Prozessor, welcher XOR-Funktionen an den Datenblöcken vornimmt.

5. Netzwerkspeichervorrichtung nach Anspruch 1, wobei die passive Rückwandplatine weiter umfasst: - dritte und vierte Datenbusse.

6. Vorrichtung nach Anspruch 1, wobei jeder der ersten und zweiten Datenbusse Teil einer Gruppe von Rückwandplatinenbussen ist, und diese Gruppe PCI-X-(Peripheral-Component-Interconnect-) Busse umfasst.

7. Vorrichtung nach Anspruch 2, wobei:

- die passive Rückwandplatine ferner einen dritten Datenbus und einen vierten Datenbus umfasst;
- das erste Kanal-Schnittstellenmodul einen ersten Bus-Port und einen zweiten Bus-Port umfasst, und das zweite Kanal-Schnittstellenmodul einen dritten Bus-Port und einen vierten Bus-Port umfasst, wobei die ersten, zweiten, dritten und vierten Bus-Ports betriebsbereit sind, um den Kommunikationspfadanteil mit der passiven Rückwandplatine zu verbinden;
- das erste Speichercontrollermodul eine erste Busschnittstelle und eine zweite Busschnittstelle umfasst, und das zweite Speichercontrollermodul eine dritte Busschnittstelle und eine vierte Busschnittstelle umfasst, wobei die ersten, zweiten, dritten und vierten Busschnittstellen betriebsbereit sind um das Speichercontrollermodul mit den ersten, zweiten, dritten und vierten Datenbussen der passiven Rückwandplatine zu verbinden.

8. Vorrichtung nach Anspruch 7, wobei:

- der erste Bus-Port mit dem ersten Datenbus verbunden ist und der zweite Bus-Port mit dem dritten Datenbus verbunden ist; und
- der dritte Bus-Port mit dem zweiten Datenbus verbunden ist und der vierte Bus-Port mit dem vierten

Datenbus verbunden ist;

- die erste Busschnittstelle mit dem ersten Datenbus und die zweite Busschnittstelle mit dem zweiten Datenbus verbunden ist; und
- die dritte Busschnittstelle mit dem dritten Datenbus verbunden ist und die vierte Busschnittstelle mit dem vierten Datenbus verbunden ist.

9. Vorrichtung nach Anspruch 8, wobei:

- der Kommunikationspfadanteil des ersten Kanal-Schnittstellenmoduls einen ersten Shared Path, einen ersten Switched Path und einen zweiten Switched Path aufweist; und
- der Kommunikationspfadanteil des zweiten Kanal-Schnittstellenmoduls einen zweiten Shared Path, einen dritten Switched Path und einen vierten Switched Path umfasst und in welchem:
 - der erste Shared Path mit dem ersten Bus-Port und dem zweiten Bus-Port verbunden ist;
 - der erste Switched Path mit dem ersten Bus-Port und dem Kanal-Schnittstellenanteil verbunden ist;
 - der zweite Switched Path mit dem zweiten Bus-Port und dem Kanal-Schnittstellenanteil verbunden ist;
 - der zweite Shared Path mit dem dritten Bus-Port und dem vierten Bus-Port verbunden ist;
 - der dritte Switched Path mit dem dritten Bus-Port und dem Kanal-Schnittstellenanteil verbunden ist;
 - der vierte Switched Path mit dem vierten Bus-Port und dem Kanal-Schnittstellenanteil verbunden ist; und wobei
 - die ersten, zweiten, dritten und vierten Switched Paths betriebsbereit sind, um die Kommunikation, welche den Kanalschnittstellenanteil einbezieht, freizugeben und zu sperren.

10. Vorrichtung nach Anspruch 1, wobei:

- zumindest das erste Kanal-Schnittstellenmodul einen ersten Shared Path umfasst, und die Speicherdaten zwischen dem ersten Speichercontrollermodul und dem zweiten Speichercontrollermodul mittels des ersten Shared Path übertragen werden.

11. Verfahren für das Sharing der Daten zwischen einem ersten Speichercontrollermodul und einem zweiten Speichercontrollermodul, umfassend:

- Bereitstellen eines ersten Shared Path in einem ersten Kanal-Schnittstellenmodul und einem Direct-Memory-Access-Prozessor in jeden der ersten und zweiten Speichercontrollermodulen; und
- Übertragen der Daten zwischen dem ersten Speichercontrollermodul und dem zweiten Speichercontrollermodul mittels des Direct Memory Access-Prozessors über den ersten Shared Path.

12. Verfahren nach Anspruch 11, ferner umfassend:

- Bereitstellen eines zweiten Shared Path in einem zweiten Kanal-Schnittstellenmodul;
- Übertragen der Daten zwischen dem ersten Speichercontrollermodul und dem zweiten Speichercont-

rollermodul mittels des Direct-Memory-Access-Prozessors über den zweiten Shared Path.

13. Verfahren nach Anspruch 11, ferner umfassend:

– Verbinden des ersten und zweiten Kanal-Schnittstellenmoduls und des ersten und zweiten Speichercontrollermoduls an eine passive Rückwandplatine.

14. Vorrichtung für das Sharing der Daten zwischen einem ersten Speichercontrollermodul und einem zweiten Speichercontrollermodul, umfassend:

– mindestens ein Kanal-Schnittstellenmodul, welches einen ersten Shared Path aufweist;
– ein erstes Speichercontrollermodul, welches einen ersten Direct-Memory-Access-Prozessor umfasst;
– ein zweites Speichercontrollermodul, welches einen zweiten Direct-Memory-Access-Prozessor umfasst; und
– eine Kommunikationsschnittstelle, welche eine direkte Kommunikation zwischen dem ersten und dem zweiten Speichercontrollermodulen ermöglicht; wobei die Daten zwischen dem ersten und zweiten Speichercontrollermodul mittels zumindest eines der ersten und zweiten Direct-Memory-Access-Prozessoren und des ersten Shared Path übertragen werden.

15. Vorrichtung nach Anspruch 14, wobei:

– das zweite Kanal-Schnittstellenmodul mindestens einen zweiten Shared Path aufweist.

16. Vorrichtung nach Anspruch 14, wobei:

– die Kommunikations-Schnittstelle eine passive Rückwandplatine umfasst.

17. Vorrichtung nach Anspruch 16, wobei:

– die passive Rückwandplatine mindestens erste und zweite PCIX-(Peripheral-Component-Interconnect-) Busse umfasst.

Es folgen 11 Blatt Zeichnungen

Anhängende Zeichnungen

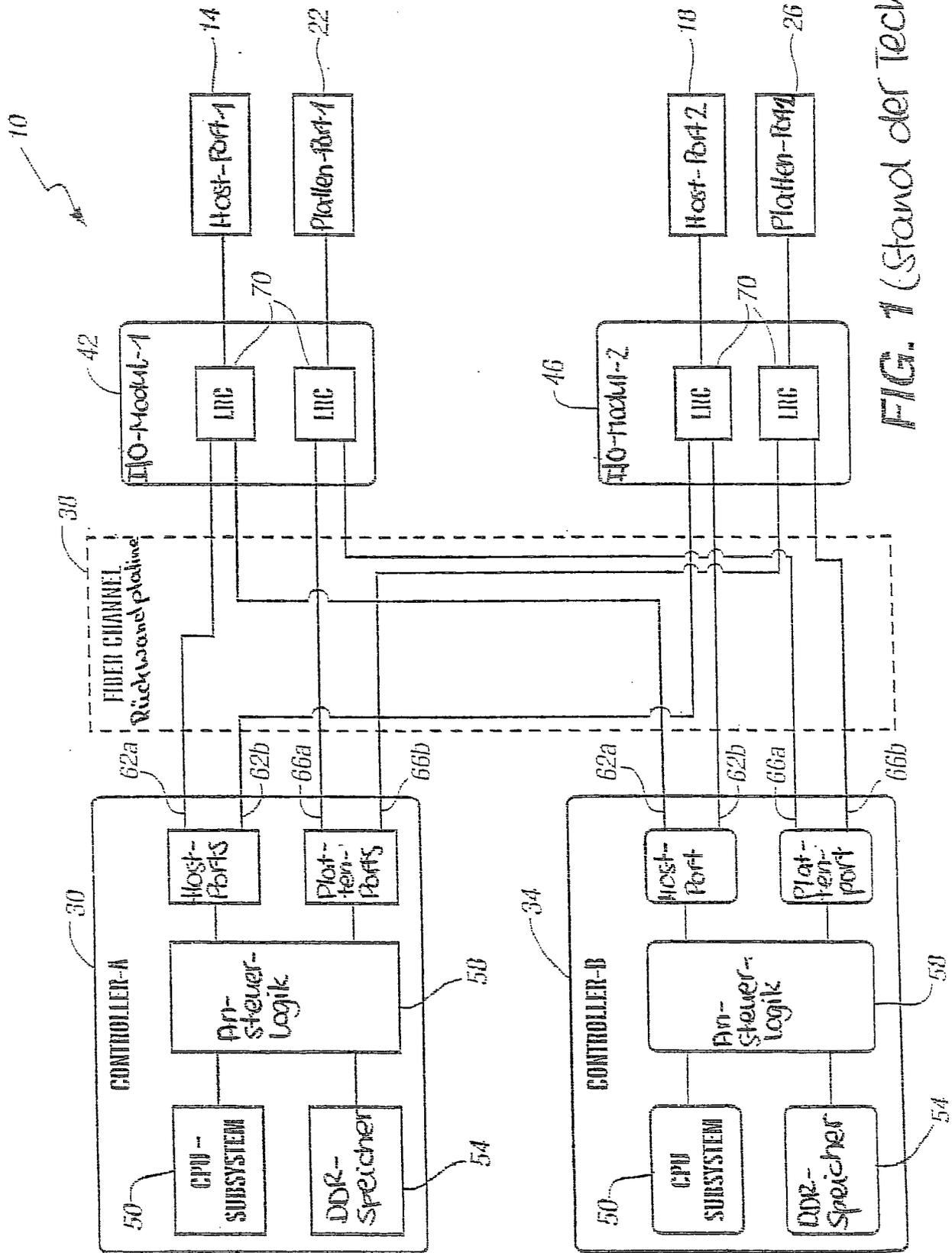
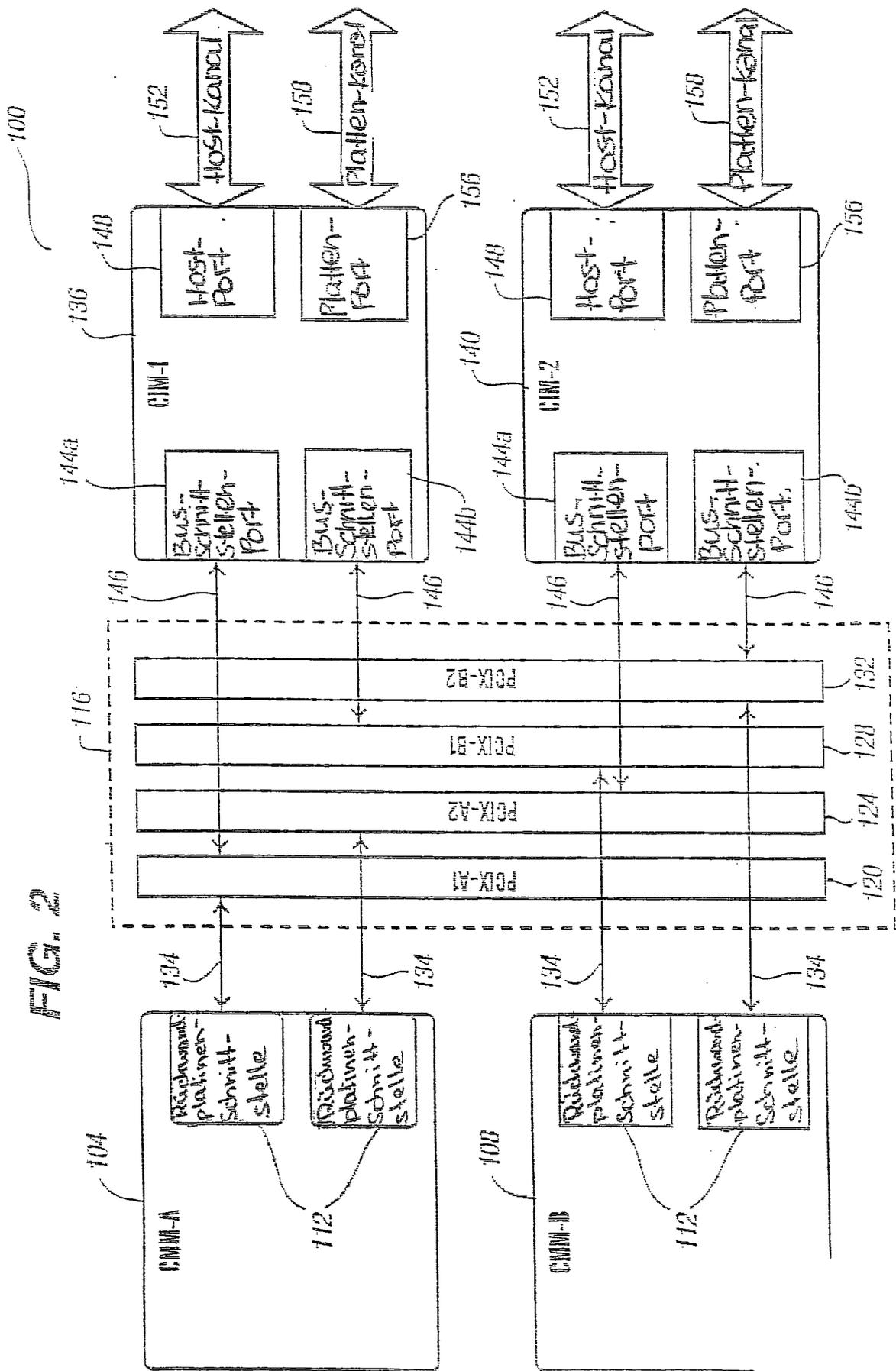


FIG. 1 (Stand der Technik)



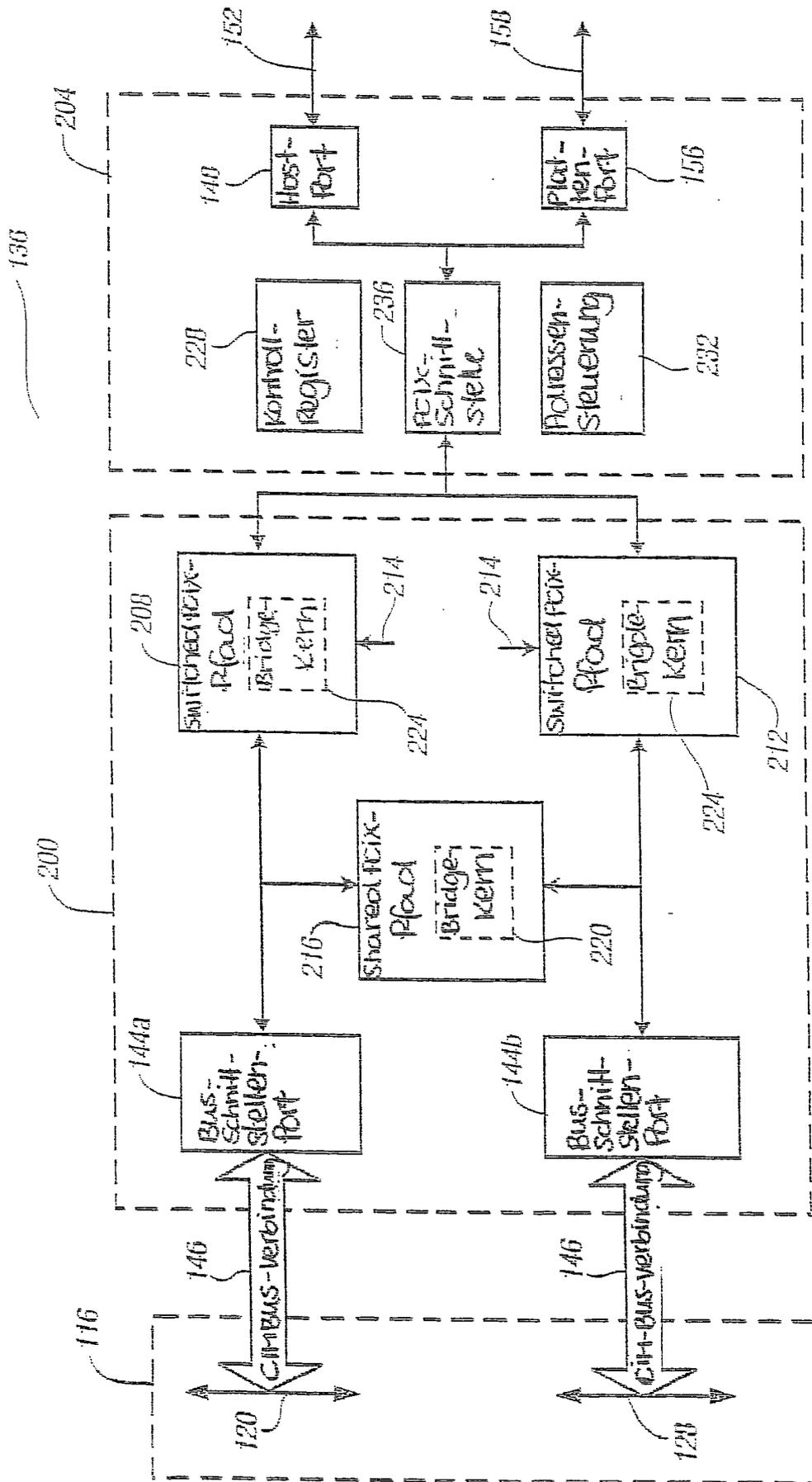


FIG. 4

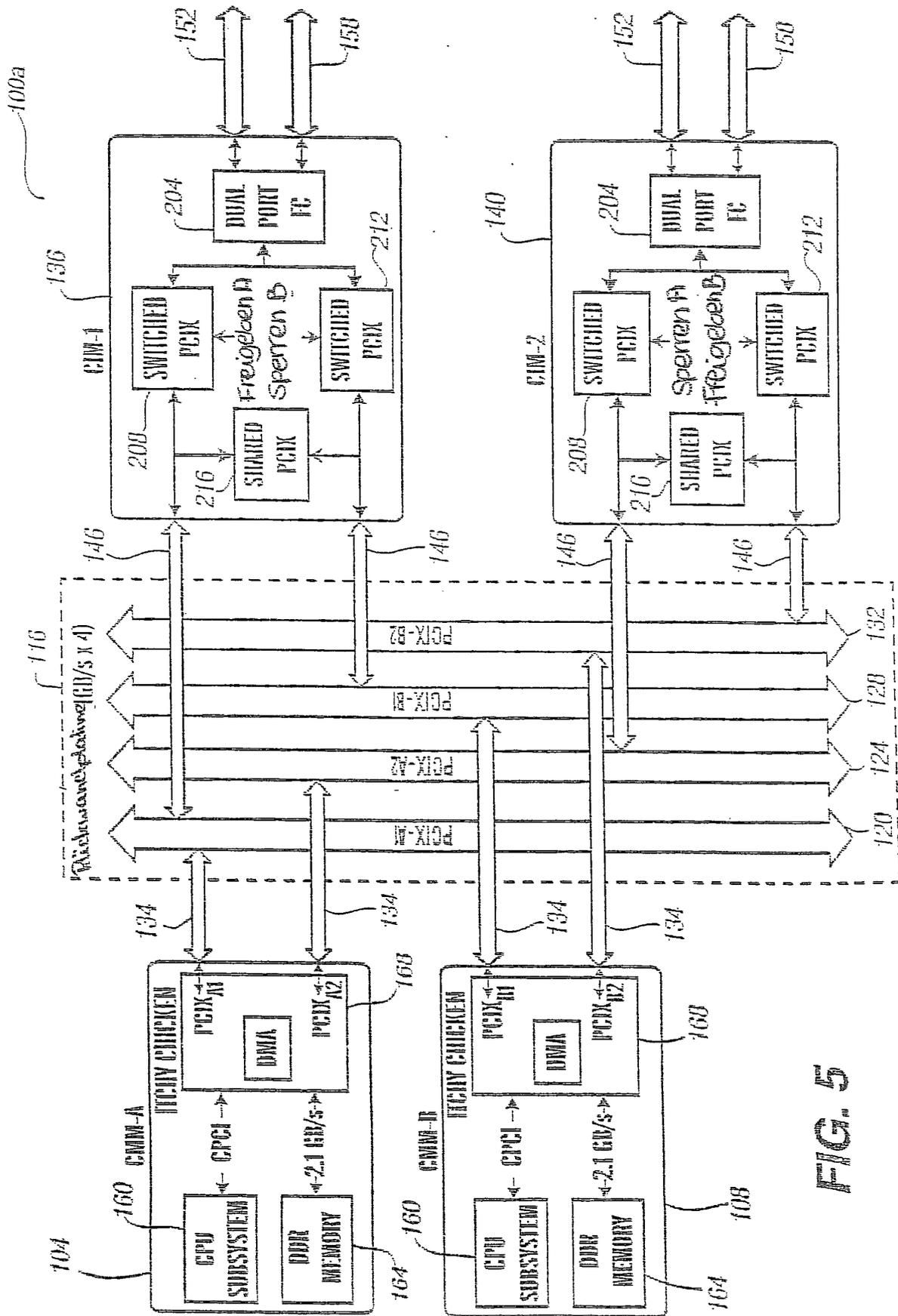


FIG. 5

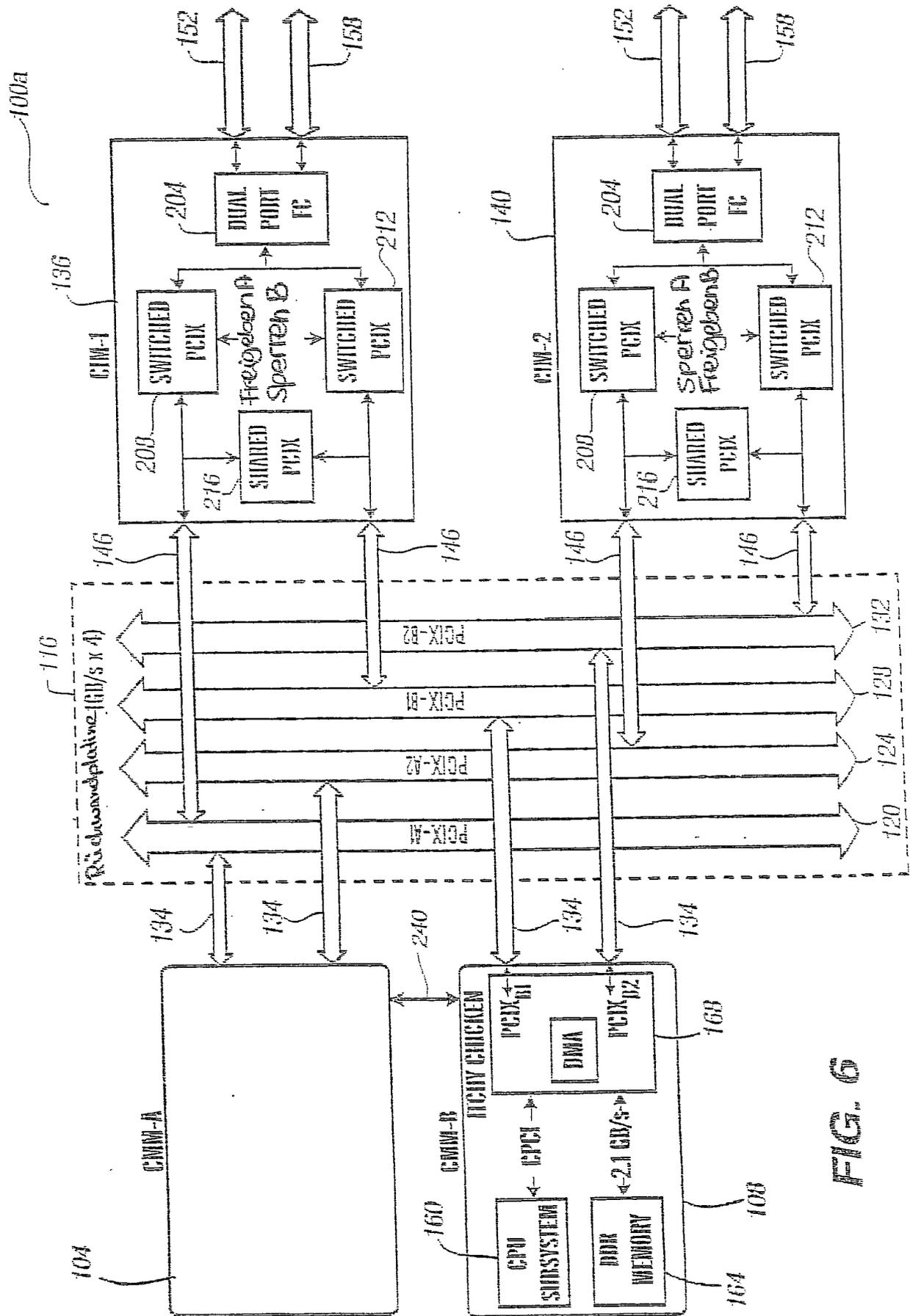


FIG. 6

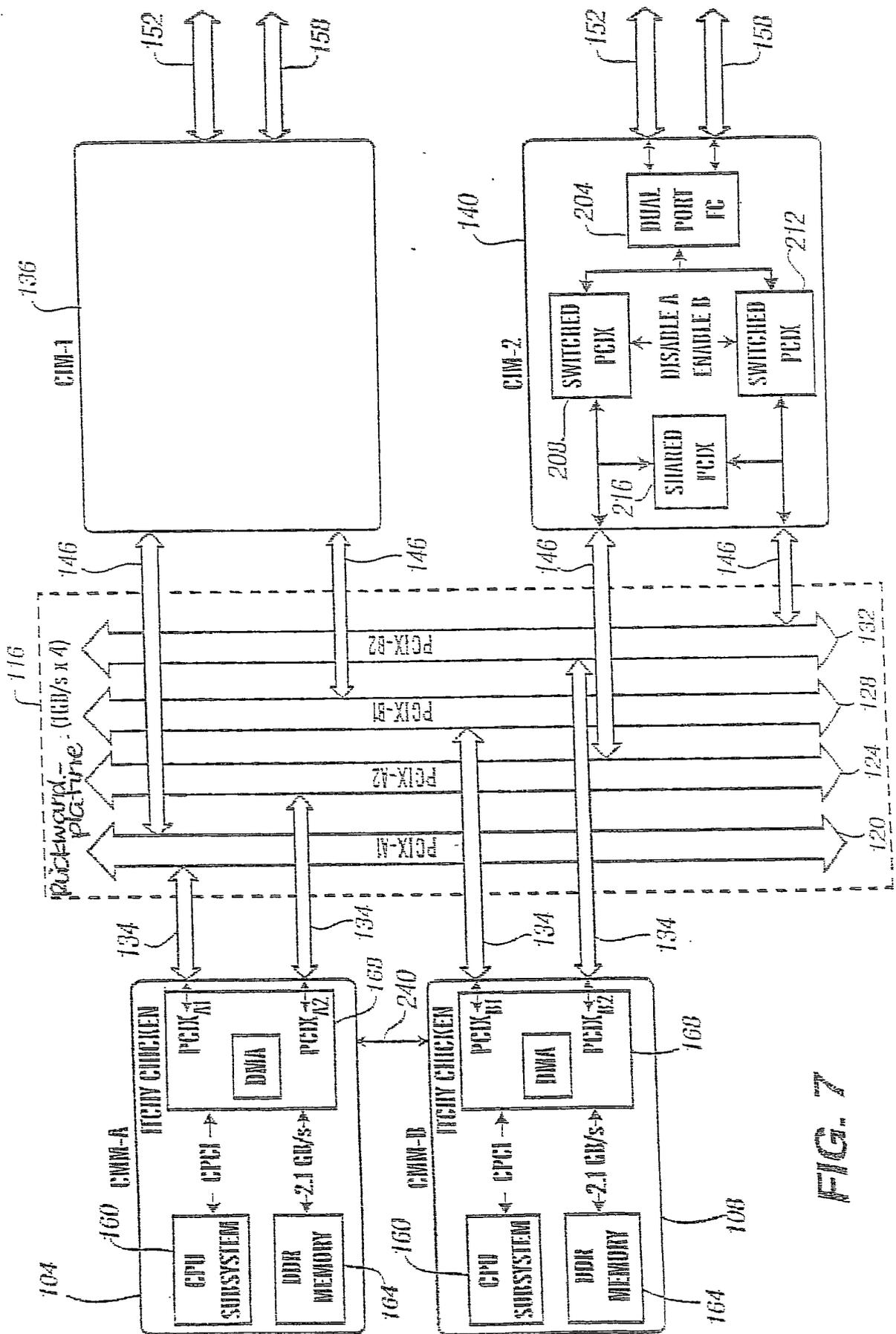


FIG. 7

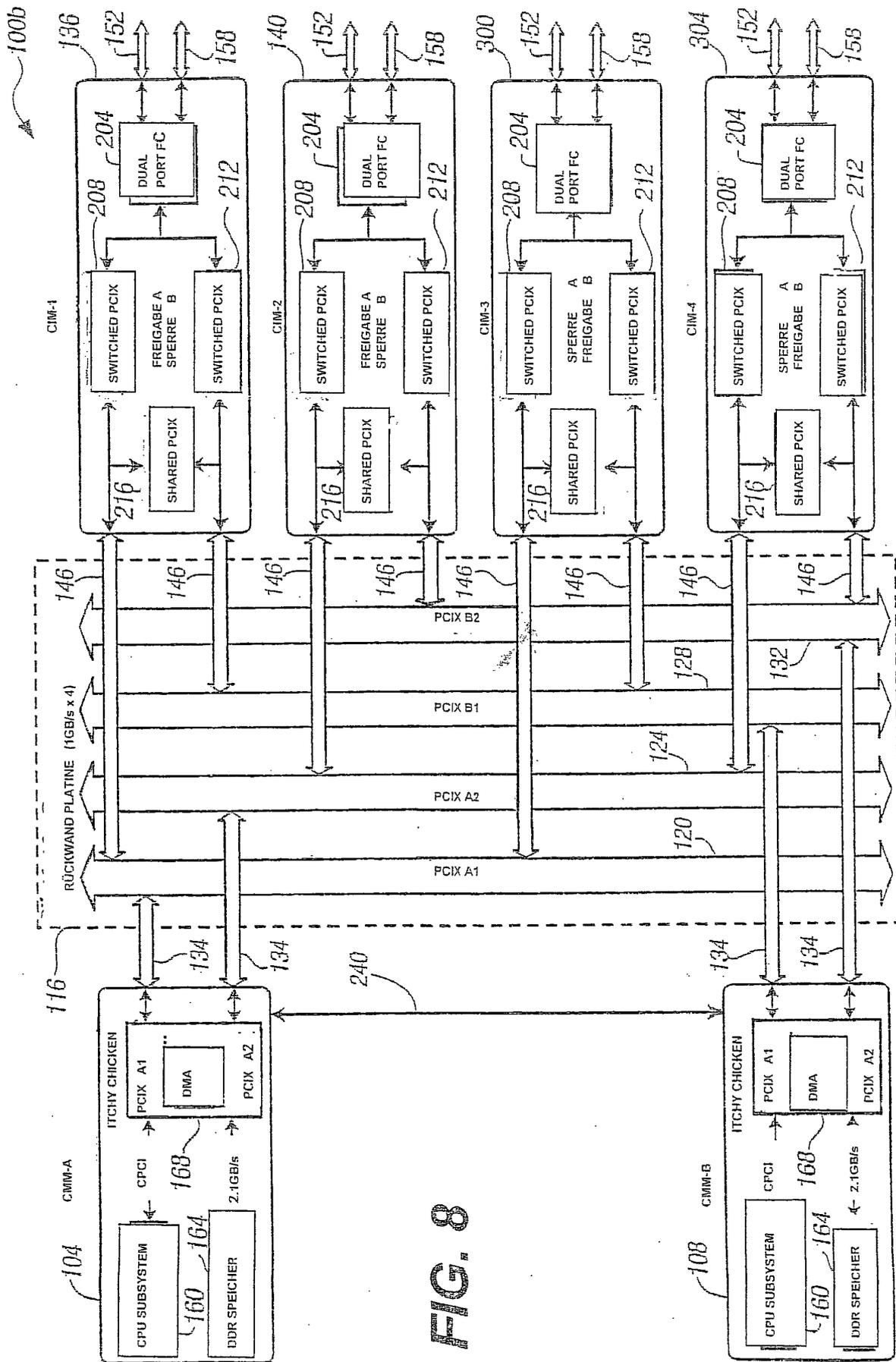


FIG. 8

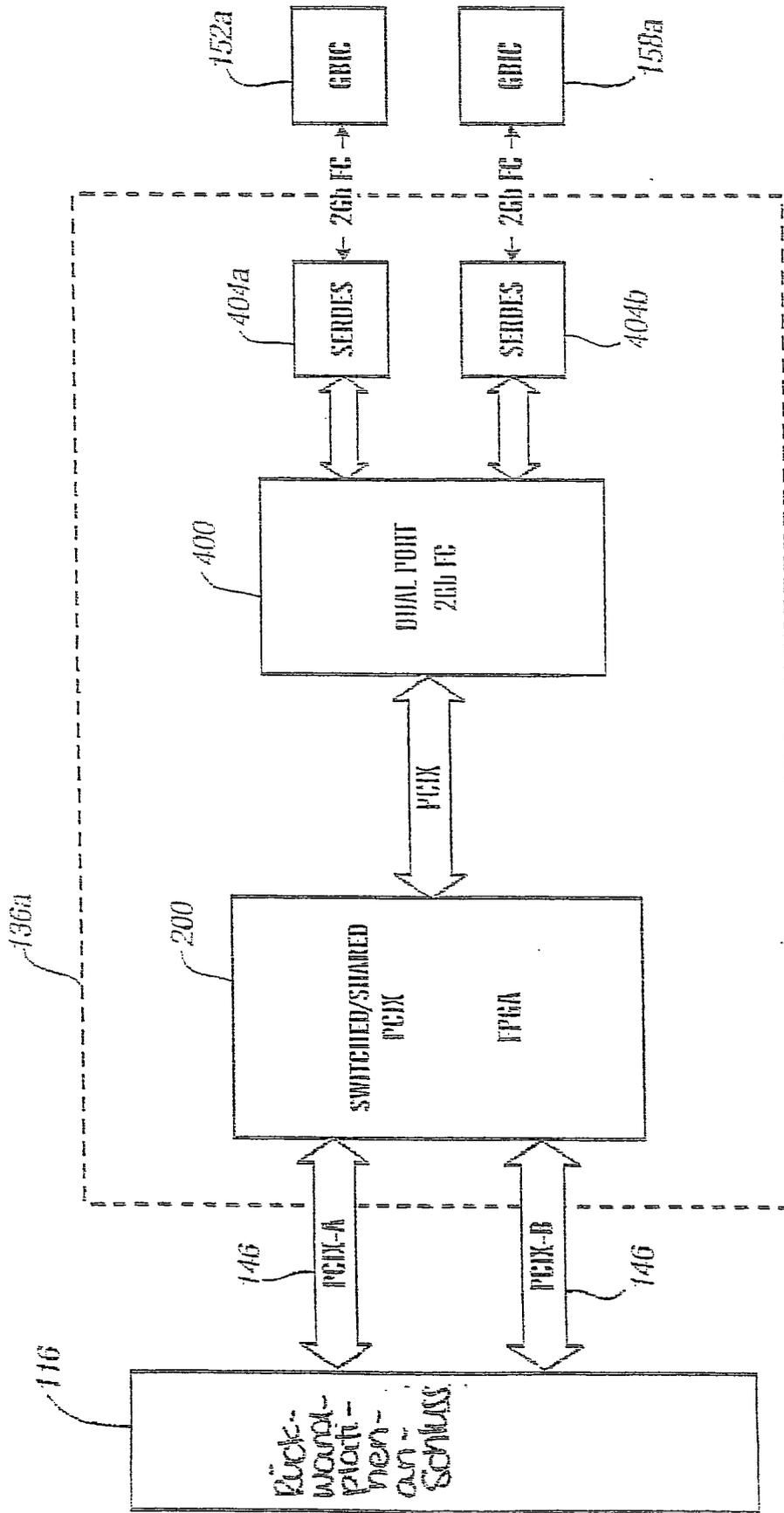


FIG. 9

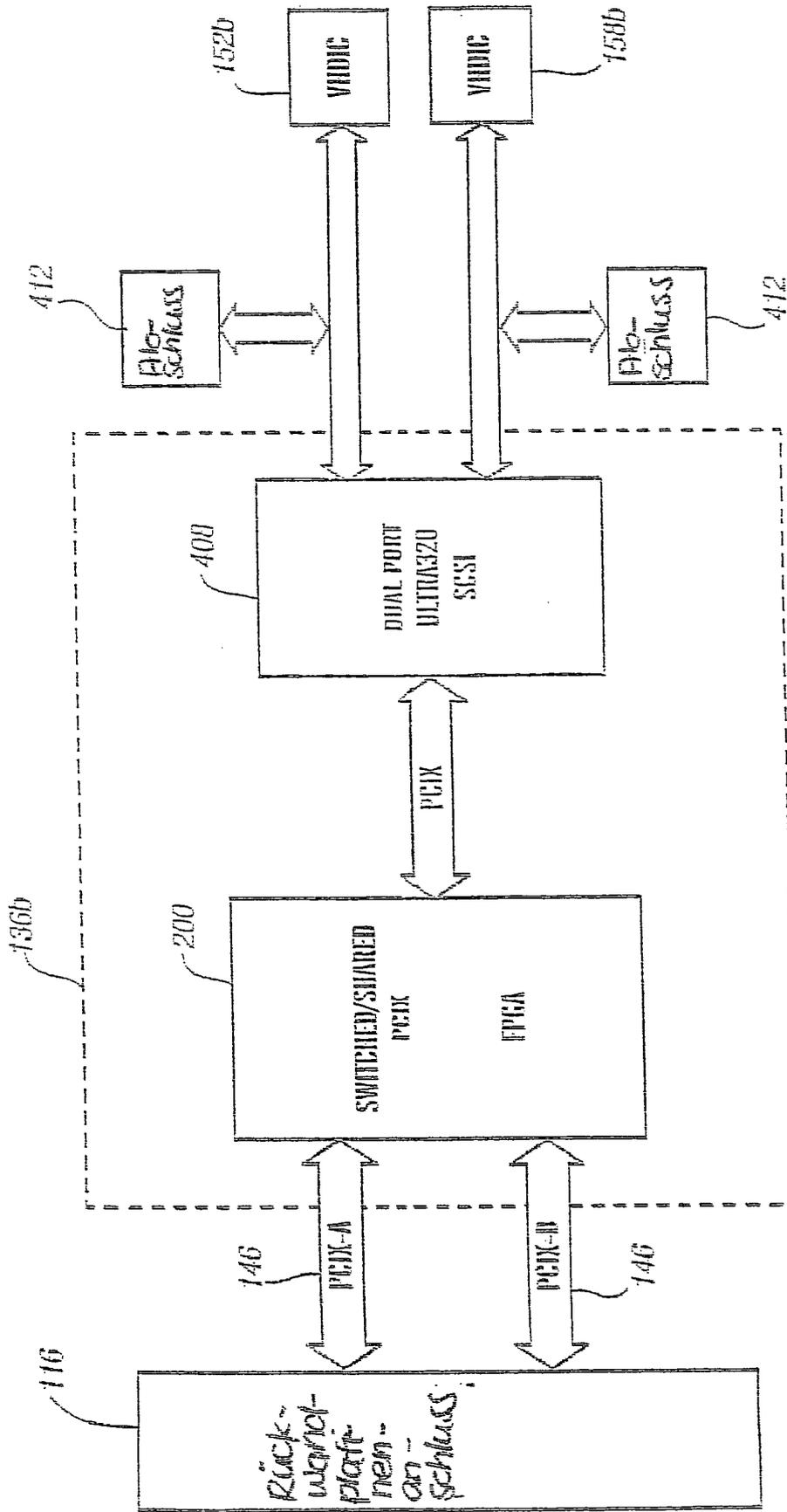


FIG. 10

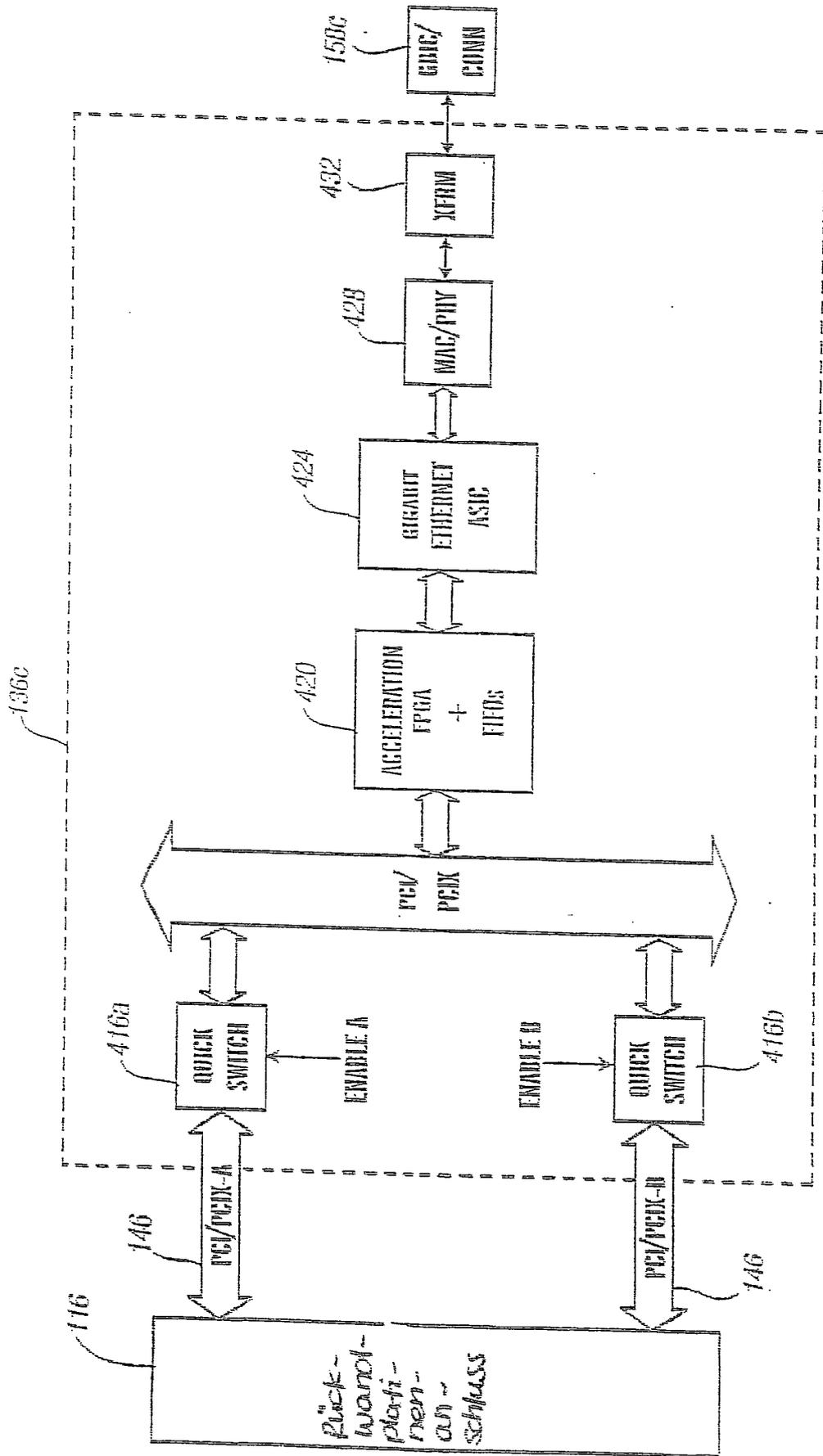


FIG. 11