

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2006-48688

(P2006-48688A)

(43) 公開日 平成18年2月16日(2006.2.16)

(51) Int.CI.

GO6F 12/16

(2006.01)

F 1

G O 6 F 12/16

G O 6 F 12/16 3 2 O C

テーマコード(参考)

5 B 0 1 8

審査請求 未請求 請求項の数 21 O L (全 20 頁)

(21) 出願番号 特願2005-220421 (P2005-220421)
 (22) 出願日 平成17年7月29日 (2005.7.29)
 (31) 優先権主張番号 10/903,188
 (32) 優先日 平成16年7月30日 (2004.7.30)
 (33) 優先権主張国 米国(US)

(71) 出願人 390009531
 インターナショナル・ビジネス・マシーンズ・コーポレーション
 INTERNATIONAL BUSINESS MACHINES CORPORATION
 アメリカ合衆国10504 ニューヨーク
 州 アーモンク ニュー オーチャード
 ロード
 (74) 代理人 100086243
 弁理士 坂口 博
 (74) 代理人 100091568
 弁理士 市位 嘉宏
 (74) 代理人 100108501
 弁理士 上野 剛史

最終頁に続く

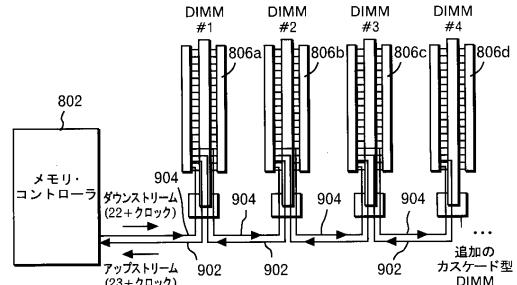
(54) 【発明の名称】セグメント・レベルの予備化を実現するシステム、方法、およびプログラム

(57) 【要約】

【課題】 セグメント・レベルで予備化したメモリ・システムを提供する。

【解決手段】 このメモリ・システムはセグメント・レベルで予備化した、カスケード型相互接続システムを含んでいる。このカスケード型相互接続システムは少なくとも2つのメモリ・アセンブリとメモリ・バスを備えている。このメモリ・バスは複数のセグメントを備え、このメモリ・アセンブリは当該メモリ・バスを介して相互接続されている。

【選択図】 図9



【特許請求の範囲】**【請求項 1】**

メモリ・システムであって、
セグメント・レベルで予備したカスケード型相互接続システムであって、
少なくとも 2 つのメモリ・アセンブリと、
複数のセグメントから成るメモリ・バスであって、前記メモリ・アセンブリは前記メモリ・バスを介して相互接続されている、メモリ・バスと
を備えている、カスケード型相互接続システム
を備えた
メモリ・システム。

10

【請求項 2】

前記メモリ・アセンブリはメモリ・コントローラおよび少なくとも 1 つのメモリ・モジュールを備えている、
請求項 1 に記載のメモリ・システム。

【請求項 3】

任意の 2 つの相互接続アセンブリの間の単一のセグメントを前記任意の 2 つの相互接続アセンブリの間の予備のセグメントで置換することができる、
請求項 1 に記載のメモリ・システム。

【請求項 4】

前記メモリ・バスはダウンストリーム・バスであり、前記単一のセグメントは前記ダウンストリーム・バスの上に配置されている、
請求項 3 に記載のメモリ・システム。

20

【請求項 5】

前記メモリ・バスはアップストリーム・バスであり、前記単一のセグメントは前記アップストリーム・バスの上に配置されている、
請求項 3 に記載のメモリ・システム。

【請求項 6】

少なくとも 2 つの相互接続されたアセンブリの間で、少なくとも 2 つの一意のセグメントを置換することができる、
請求項 3 に記載のメモリ・システム。

30

【請求項 7】

セグメント・レベルの予備化を実現する方法であって、
現在のメモリ・アセンブリにおいて入力信号を受信するステップであって、前記現在のメモリ・アセンブリは複数のセグメントを備えたメモリ・バスを介して相互接続された複数のメモリ・アセンブリを備えたカスケード型相互接続システムに含まれている、ステップと、

前記入力信号中のビット群のうちの 1 つが、アップストリームまたはダウンストリームのメモリ・アセンブリにおける故障セグメントに関連付けられるのに応答して、前記入力信号中の前記ビット群を再配置するステップ

とを備えた

40

方法。

【請求項 8】

前記入力信号が予備ビットを含み、
前記予備ビットが前記入力信号中で利用されている場合、前記ビット群のうちの 1 つを故障セグメントに関連付ける、
請求項 7 に記載の方法。

【請求項 9】

さらに、

前記入力信号中の再配置した前記ビット群のうちの 1 つが、前記現在のメモリ・アセンブリに接続されている故障セグメントに関連付けられるのに応答して、前記入力信号中の

50

再配置した前記ビット群の順番を入れ替えるステップと、

アップストリームまたはダウンストリームのメモリ・アセンブリに前記順番を入れ替えたビット群を転送するステップとを備えた、
請求項 7 に記載の方法。

【請求項 10】

前記再配置をマルチプレクサで行う、
請求項 7 に記載の方法。

【請求項 11】

前記現在のメモリ・アセンブリはメモリ・コントローラまたはメモリ・モジュールである、
請求項 7 に記載の方法。

【請求項 12】

前記バスは一方向性のアップストリーム・メモリ・バスおよび一方向性のダウンストリーム・メモリ・バスを含む、
請求項 7 に記載の方法。

【請求項 13】

前記アップストリーム・メモリ・バス上の情報は 23 個の信号を含み、1 個の信号は予備ビットであり、4 個の信号はエラー訂正コード・ビットである、
請求項 12 に記載の方法。

【請求項 14】

前記ダウンストリーム・メモリ・バス上の情報は 22 個の信号を含み、1 個の信号は予備ビットであり、4 個の信号はエラー訂正コード・ビットである、
請求項 12 に記載の方法。

【請求項 15】

前記カスケード型相互接続システムはバス・レベル・エラー訂正コード機能ブロックを備えている、
請求項 7 に記載の方法。

【請求項 16】

さらに、
前記再配置したビット群をバス・レベル・エラー訂正コード機能ブロックに転送するステップを備えた、
請求項 15 に記載の方法。

【請求項 17】

前記バス・レベル・エラー訂正コード機能ブロックはエラー・フラグおよびエラー・データをメモリ・コントローラにエラー報告ブロックを介して転送する、
請求項 16 に記載の方法。

【請求項 18】

さらに、
故障ビットを無効にして予備ビットを使用するコマンドを、前記現在のメモリ・アセンブリにおいて受信するステップを備えた、
請求項 16 に記載の方法。

【請求項 19】

前記コマンドをメモリ・コントローラから受信する
請求項 18 に記載の方法。

【請求項 20】

セグメント・レベルの予備化を実現するためのプログラムであって、
現在のメモリ・アセンブリにおいて入力信号を受信するステップであって、前記現在のメモリ・アセンブリは複数のセグメントを備えたメモリ・バスを介して相互接続された複数のメモリ・アセンブリを備えたカスケード型相互接続システムに含まれている、ステップと、

前記入力信号中のビット群のうちの1つが、アップストリームまたはダウンストリームのメモリ・アセンブリにおける故障セグメントに関連付けられるのに応答して、前記入力信号中の前記ビット群を再配置するステップと

をコンピュータに実行させる命令群を含むプログラム。

【請求項 21】

メモリ・システムであって、
セグメント・レベルの予備化を備えたカスケード型相互接続システムであって、
少なくとも 2 つのメモリ・アセンブリと、
少なくとも 1 つのクロック信号セグメントを備えたメモリ・バスであって、前記メモリ・アセンブリは前記メモリ・バスを介して相互接続されており、任意の 2 つの相互接続されたアセンブリの間の単一のクロック・セグメントは前記任意の 2 つの相互接続されたアセンブリの間の予備セグメントで置換しうる、メモリ・バスと

を含むカスケード型相互接続システムを備えた

メモリ・システム。

【発明の詳細な説明】

【技術分野】

【0001】本発明はメモリ・システムに関し、特にセグメント・レベルの予備化を実現するメモリ

・システムに

【 月 京 技 术 】

【0002】コンピュータのメモリ・システムは長年かけて進化してきたが、不变の属性を数多く保持し続けている。1980年代初頭のコンピュータ・メモリ・システム（たとえばルヴァリー（LeVallee）らの米国特許第4475194号に開示されたものなど）はメモリ・コントローラ、アレイ状のデバイス群を備えたメモリ・アセンブリ（当時、当該発明者らは基本記憶モジュール（basic storage module: BSM）と呼んでいた）、バッファ、終端抵抗、および、補助的なタイミング・制御機能、ならびに、各メモリ・アセンブリがそれ自身のポイント・ツー・ポイント型のアドレス・バスおよびデータ・バスを介してメモリ・コントローラと通信しうるようにするいくつかのポイント・ツー・ポイント型のバスを備えていた。図1は2つのBSM、メモリ・コントローラ、保守卓、ならびにBSMとメモリ・コントローラを接続するポイント・ツー・ポイント型のアドレス・バスおよびデータ・バスを備えた1980年代初頭のコンピュータ・メモリ・システムの一例を示している。

[0 0 0 3]

図2はデル(Dell)らの米国特許第5513135号に由来するものであるが、初期の同期型メモリ・モジュールを示している。この同期型メモリ・モジュールは同期型の動的ランダム・アクセス・メモリ(DRAM: dynamic random access memory)8、バッファ12、最適化したピンアウト(pinout)、相互接続、および、動作を容易にする容量性減結合手段を備えている。この特許には、位相ロック・ループ(PLL: phase lock loop)などのデバイスを用いることにより、モジュール上でクロックの再駆動を使用することも記載されている。

[0 0 0 4]

図3はグルンドン(Grundon)らの米国特許第6510100号に由来するものであるが、既存のマルチドロップ型スタブ・バス・チャネル上に最大4個の登録済みデュアル・インライン・メモリ・モジュール(DIMM:dual inline memory module)40を備えうるメモリ・システム10を単純化した図および説明を示す。このシステムはメモリ・コントローラ20、外部クロック・バッファ30、登録済みDIMM40、アドレス・バス50、制御バス60、およびデータ・バス70を備えている。アドレス・バス50およびデータ・バス70は、各々の登録済みDIMM40と接続される。

タ・バス 7 0 には終端抵抗 9 5 が設けられている。

【 0 0 0 5 】

図 4 は図 1 の構成から進化した 1990 年代のメモリ・システムを示す。このメモリ・システムはメモリ・コントローラ 4 0 2 と、少なくとも 1 つの高速ポイント・ツー・ポイント型チャネル 4 0 4 を備えていた。チャネル 4 0 4 は各々、バス・ツー・バス型コンバータ・チップ 4 0 6 を備えていた。そして、コンバータ・チップ 4 0 6 は各々、少なくとも 1 つの登録済み D I M M 4 1 0 への接続を可能にする同期型メモリ・インターフェース 4 0 8 を備えていた。この構成において、高速のポイント・ツー・ポイント型チャネル 4 0 4 は D R A M のデータ転送速度の 2 倍の速度で動作した。それにより、バス・ツー・バス型コンバータ・チップ 4 0 6 は 1 個または 2 個の登録済み D I M M メモリ・チャネルを D R A M の最大データ転送速度で動作させることができた。登録済み D I M M は各々、P L L、レジスタ、D R A M、電気的に消去可能でプログラム可能な読み出し専用メモリ (E E P R O M : electrically erasable programmable read-only memory)、終端抵抗、および、他の受動素子を備えていた。
10

【 0 0 0 6 】

図 5 に示すように、メモリ・システムは単一のメモリ・モジュール、または「スタブ」バス上で相互接続した少なくとも 2 つのメモリ・モジュールを用いて構築することが多かった。図 5 は図 3 に示したものに類似したマルチドロップ型スタブ・バス・メモリ構成の単純化した一例である。この構成はコスト、性能、信頼性、およびアップグレード機能の間ににおける妥当なトレードオフを提示するものであるが、スタブ・バスに取り付けうるモジュールの個数について固有の限界を有する。スタブ・バスに取り付けうるモジュールの個数に関する限界はこのバスを介して転送する情報のデータ転送速度に直接に関係する。データ転送速度を増大させるほど、頑健なメモリ動作を保証するために、スタブの本数と長さを低減させる必要がある。スタブ・バスの速度が増大すると一般に、当該バス上のモジュールの個数は減少する。その結果、最適な電気的なインターフェースは単一のモジュールが単一のコントローラに直接に接続されているもの、あるいは、(もしあるとして)少数のスタブに対するポイント・ツー・ポイント型のインターフェース、ということになるが、後者の場合には反射とインピーダンスの不連続とが生じる。大多数のメモリ・モジュールのデータ幅は 6 4 ビットまたは 7 2 ビットであるから、この構成ではアドレス、コマンド、およびデータを転送するのに多数のピンが必要になる。代表的なピンカウントである図 5 においては、1 2 0 個のピンが特定されている。
20
30

【 0 0 0 7 】

図 6 はペティ (Petty) の米国特許第 4 7 2 3 1 2 0 号に由来するものであるが、マルチポイント型の通信構成にデイジー・チェーン構成を適用することに関する。この図 6 のデイジー・チェーン構成は、1 次ステーション 1 に 2 次ステーション 2 A、2 B、2 C、2 D、2 E が順次接続され、更にモデムの対 3 A、3 B を介して他の 2 次ステーション 2 F に接続されている。デイジー・チェーン構成を適用しない場合、マルチポイント型の通信構成は装置群を分離するポイント・ツー・ポイント型のインターフェースを介して各々が接続される複数のポートが必要になる。デイジー・チェーン構成を採用すると、より少ない個数のポート (すなわちチャネル) を用いて制御ステーションを構成することができる。そして、チャネル上の各装置はデイジー・チェーン構成における自身の位置とは無関係に、標準のアップストリーム・プロトコルとダウンストリーム・プロトコルを利用することができる。
40

【 0 0 0 8 】

図 7 は米国特許第 4 7 2 3 1 2 0 号の教示に合わせて実装したデイジー・チェーン型のメモリ・バスを示す図である。メモリ・コントローラ 1 1 1 はメモリ・バス 3 1 5 に接続されている。そして、メモリ・バス 3 1 5 はモジュール 3 1 0 a に接続されている。バス 3 1 5 上の情報はモジュール 3 1 0 a 上のバッファ 3 2 0 a により次のモジュール 3 1 0 b へ再駆動される。そして、モジュール 3 1 0 b が、3 1 0 n で示すモジュールの位置までバス 3 1 5 を再駆動する。各モジュール 3 1 0 a、3 1 0 b … 3 1 0 n は D R A M
50

311a、311b・・・311nとバッファ320a、320b・・・320nを備えている。バス315はデイジー・チェーン構成を有するように記載されており、各バスは本来的にポイント・ツー・ポイント型である。

【特許文献1】米国特許第4475194号

【特許文献2】米国特許第5513135号

【特許文献3】米国特許第6510100号

【特許文献4】米国特許第4723120号

【発明の開示】

【発明が解決しようとする課題】

【0009】

デイジー・チェーン型のバスを使用することに対する1つの欠点は、それにより、当該バスに沿って複数のメモリ・モジュールに影響を与える障害の発生確率が増大する、ということである。たとえば、第1のモジュールが機能不全に陥ると、当該バス上の第2のモジュールおよびそれ以降のモジュールも機能不全に陥ってしまう。

【課題を解決するための手段】

【0010】

本発明の典型的な実例はセグメント・レベルの予備化を備えたカスケード型相互接続システムを備えたメモリ・システムを含んでいる。このカスケード型相互接続システムは少なくとも2つのメモリ・アセンブリとメモリ・バスを備えている。このメモリ・バスは複数のセグメントを備え、このメモリ・アセンブリは当該メモリ・バスを介して相互接続されている。

【0011】

本発明の典型的なさらなる実例はセグメント・レベルの予備化を実現する方法を含んでいる。この方法は、現在のメモリ・アセンブリにおいて入力信号を受信するステップであって、前記現在のメモリ・アセンブリは複数のセグメントを備えたメモリ・バスを介して相互接続された複数のメモリ・アセンブリを備えたカスケード型相互接続システムに含まれている、ステップを備えている。前記入力信号中のビット群のうちの1つが、アップストリームまたはダウンストリームのメモリ・アセンブリにおける故障セグメントに関連付けられるのに応答して、前記入力信号中の前記ビット群を再配置する。

【0012】

本発明の典型的なさらなる実例はセグメント・レベルを予備化したメモリ・システムを実現する記憶媒体を含んでいる。この記憶媒体はセグメント・レベルの予備化を実現するマシン読み取り可能なコンピュータ・プログラム・コードでエンコードされている。この記憶媒体は、コンピュータに、現在のメモリ・アセンブリにおいて入力信号を受信するステップであって、前記現在のメモリ・アセンブリは複数のセグメントを備えたメモリ・バスを介して相互接続された複数のメモリ・アセンブリを備えたカスケード型相互接続システムに含まれている、ステップを備えた方法を実行させる命令群を含んでいる。前記入力信号中のビット群のうちの1つが、アップストリームまたはダウンストリームのメモリ・アセンブリにおける故障セグメントに関連付けられるのに応答して、前記入力信号中の前記ビット群を再配置する。

【発明を実施するための最良の形態】

【0013】

本発明の典型的な実施形態はセグメント・レベルの予備化、すなわちメモリ・チャネル上にある任意の2つのアセンブリの間（すなわちメモリ・コントローラとメモリ・モジュールとの間、または、任意の2つのメモリ・モジュールの間）において（たとえば書き込み障害、コネクタの障害、駆動器の障害、および、受信器の障害、または、これらの障害のうちのいくつかのものに起因して）故障した相互接続をすべて置換する能力を提供するものである。セグメント・レベルの予備化はセグメント当たりバス当たり少なくとも1つの置換用の配線を用いることにより、アップストリーム・バスおよびダウンストリーム・バスの双方に行うことができる。換言すると、各バス上の各セグメント（コントローラ -

10

20

30

40

50

第1のDIMM、第1のDIMM - 第2のDIMM、第2のDIMM - 第3のDIMMなど、および、メモリ・コントローラへの帰路)は故障中の一意のセグメントの各々の上を伝送する予定であった情報を別の「置換物」すなわち「予備の」セグメントに再転送することにより、論理的に置換することができる。セグメント・レベルの予備化を実現しすると、システムを修理する処置が必要になる前に一意の故障が多数生じるのを許容できるという点において、信頼性と耐性が向上する。

【0014】

本発明の典型的な実施形態では、高信頼性のメモリ・システム・アーキテクチャと、任意の2つのシステム構成要素の間に非平衡ポイント・ツー・ポイント型相互接続を備えた相互接続構造とによってセグメント・レベルの予備化を実現している。このメモリ・システムはさらにメモリ制御機能、少なくとも1つのメモリ・モジュール、DRAMのデータ転送速度に比して4対1の速度比で動作する少なくとも1つの高速バス、および、少なくとも1つのカスケード型モジュールの各々の上にあって高速バス(群)を既存の2倍データ転送速度(DDR: double data rate、ディー・ディー・アール)型のメモリ・インターフェースに変換するバス・ツー・バス型のコンバータ・チップを備えている。このメモリ・モジュールはメモリ・コントローラに対するスレーブ装置として動作し、コマンドに対して決定論的または非決定論的な様態で応答するが、動作上のエラーを報告する場合以外に想定外のアクティビティを自発的に開始することはない。カスケード接続されたバスにはメモリ・モジュールを追加することができる。その場合、各モジュールにはカスケード接続された当該バス上にある各モジュールを一意に選択するのを可能にするアドレスを割り当てる。本発明の典型的な実施形態はパケット型の多転送インターフェースを備えている。このインターフェースはピンカウントが少ない場合であってもメモリ・オペレーションを行うのを可能にする革新的な通信プロトコルを利用している。これにより、カスケード接続されたバス上にある構成要素間においてアドレス、コマンド、およびデータを複数サイクルにわたって転送するとともに、それらを再構築している。そして、将来の受信者が使用する前にエラーを訂正している。

【0015】

図8はバッファ付きメモリ・モジュール806(バッファ・デバイスはたとえばメモリ・モジュール806内に備えられている)がメモリ・コントローラ802と通信する場合における、本発明の典型的な実施形態が使用するカスケード型のメモリ構成を示す図である。このメモリ構成には、高速のポイント・ツー・ポイント型バス804を介して少なくとも1つのメモリ・モジュール806と通信するメモリ・コントローラ802が含まれている。図8に示す典型的な実施形態における各バス804は約50本の高速ワイヤを備え、アドレス、コマンド、データ、およびクロックの転送の用に供している。上述した従来技術において説明したポイント・ツー・ポイント型のバスを使用すると、バスの構成を最適化して、きわめて大きなデータ転送速度を可能にするとともに、複数のサイクルにわたってデータを転送することによりバスのピンカウントを低減させることが可能になる。図4にはメモリ・コントローラをバス・コンバータに接続するバスのうちの任意の1つのもののデータ転送速度(たとえばピン当たり最大1066Mb/s)と、バス・コンバータと少なくとも1つのメモリ・モジュールとの間にあるバスのうちの任意の1つのもののデータ転送速度(たとえばピン当たり最大533Mb/s)との比が2対1であるメモリ・システムを示したが、図8に示す本発明の典型的な実施形態では4対1のバス速度比を実現してバスの効率を最大化するとともにピンカウントを最小化している。

【0016】

ポイント・ツー・ポイント型の相互接続によってデータ転送速度をより大きくすることが可能になるが、メモリ・システムの全体的な高効率性を達成するにはメモリ・モジュール806の個数とチャネル当たりのメモリ装置の個数とを妥当な値に維持する必要がある(歴史的にはメモリ・モジュール当たり4~36個のチップを備えた4個のメモリ・モジュールであるが、チャネル当たり最大8個のメモリ・モジュールからチャネル当たり最小1個のメモリ・モジュールまでありうる)。ポイント・ツー・ポイント型のバスを使用す

10

20

30

40

50

ると、メモリ・モジュールをカスケード接続して各メモリ・モジュールをメモリ・コントローラ802のほかに他のメモリ・モジュールにも相互接続しうるようにするために、各メモリ・モジュールにはバスを再駆動する機能が必要になる。

【0017】

図9はメモリ・モジュール806がすべてバッファ付きメモリ・モジュールである場合に本発明の典型的な実施形態が使用する、カスケード接続されたメモリ・モジュール群と一方向性のバス群とを備えたメモリ構成を示す図である。カスケード構成におけるメモリ・モジュール806が備えている機能の1つは、メモリ・バス上の信号を他のメモリ・モジュール806またはメモリ・コントローラ802に送付する再駆動機能である。図9にはメモリ・コントローラ802と4個のメモリ・モジュール806a、806b、806c、806dとが含まれている。これらのメモリ・モジュールは2つのメモリ・バス(ダウンストリーム・メモリ・バス904とアップストリーム・メモリ・バス902)の各々の上にあるとともに、直接的な態様またはカスケード型の態様でメモリ・コントローラ802に接続されている。メモリ・モジュール806aは直接的な態様でメモリ・コントローラ802に接続されている。メモリ・モジュール806b、806c、806dはカスケード型の態様でメモリ・コントローラ802に接続されている。

【0018】

本発明の典型的な実施形態はカスケード型のメモリ構成においてメモリ・コントローラ802とメモリ・モジュール806a(「DIMM#1」との間、ならびに、後続する各メモリ・モジュール806b、806c、806d(「DIMM#2」、「DIMM#3」、および「DIMM#4」)の間に2つの一方向性のバスを備えている。ダウンストリーム・メモリ・バス904は22個の非平衡型の信号、および平衡型のクロック対から成る。このダウンストリーム・メモリ・バス904を用いてアドレス、制御、および誤り訂正符号(ECC: error code correction)ビットのダウンストリームをメモリ・コントローラ802から、カスケード型メモリ・チャネルに実装された少なくとも1つのメモリ・モジュール806に数クロック・サイクルにわたって転送する。アップストリーム・メモリ・バス902は23個の非平衡型の信号、および平衡型のクロック対から成る。このアップストリーム・メモリ・バス902を用いてバス・レベルのデータ・ビットおよびECCビットのアップストリームを転送元のメモリ・モジュール806からメモリ・コントローラ802に転送する。このメモリ構成、および、DRAMのデータ転送速度(たとえばピン当たり400~800Mb/s)と一方向性のメモリ・バスのデータ転送速度(たとえばピン当たり1.6~3.2Gb/s)との間に4対1型のデータ転送速度増倍器を使用すると、メモリ・コントローラ802の、メモリ・チャネル当たりの信号ピンカウントを約120ピンから約50ピンに減少させることができる。

【0019】

図10は本発明の典型的な実施形態が使用するバッファ付きメモリ・モジュール806の前面図1006および背面図1008である。本発明の典型的な実施形態において、各メモリ・モジュール806は長さが約15.24cm(約6インチ)で高さが約3.81cm(約1.5インチ)であり、18個のDRAM配置場所、バッファ装置1002、および、図示しないが当技術分野で既知の多数の小型素子(たとえばキャパシタ、抵抗器、EEPROMなど)を備えた空のカードを備えている。本発明の典型的な実施形態では、上記カードの寸法は長さが151.35mmで高さが30.5mmである。本発明の典型的な実施形態では、バッファ装置1002は図10の前面図に示すように、メモリ・モジュール806の表側の中央領域に設ける。同期型DRAM(SDRAM)1004はバッファ装置1002の両側、および、図10の背面図に示すように、メモリ・モジュール806の裏側に設ける。この構成を用いると、バッファ装置1002への高速書き込み、および、バッファ装置からSDRAM1004への信号の転送が容易になる。

【0020】

図11は本発明の典型的な実施形態が使用するバッファ付きモジュールの配線システムを示す図である。図11は図10に示したメモリ・モジュール806を模式的に示す図で

10

20

30

40

50

あり、陰影付き矢印は主要な信号の流れを表している。この信号の流れにはアップストリーム・メモリ・バス 902、ダウンストリーム・メモリ・バス 904、アドレス／コマンド・バス 1102、1106、および、データ・バス 1104、1108 が含まれている。本発明の典型的な実施形態において、バッファ装置 1002（メモリ・インターフェース・チップとも呼ばれる）はアドレス信号およびコマンド信号の 2 つのコピーを SDRAM 1004 に右側のアドレス／コマンド・バス 1106 および左側のアドレス／コマンド・バス 1102 を用いて供給する。その際、バッファ装置 1002 の右側から出て、当該右側に位置する SDRAM 1004、およびバッファ装置 1002 の背面の右側に位置する SDRAM 1004 に至る。また、バッファ装置 1002 の左側から出て、当該左側に位置する SDRAM 1004、およびバッファ装置 1002 の背面の左側に位置する SDRAM 1004 に接続する。同様に、バッファ装置 1002 の右側の SDRAM 1004 に宛てたデータ・ビットはバッファ装置 1002 の右側から出て、右側のデータ・バス 1108 に乗る。バッファ装置 1002 の左側の SDRAM 1004 に宛てたデータ・ビットはバッファ装置 1002 の左側から出て、左側のデータ・バス 1104 に乗る。高速のアップストリーム・メモリ・バス 902 およびダウンストリーム・メモリ・バス 904 はバッファ装置 1002 の下部から出て、メモリ・コントローラ、または、このメモリ・モジュール 806 の上流（アップストリーム）または下流（ダウンストリーム）にある他のメモリ・モジュールに接続するが、それは用途によって決まる。バッファ装置 1002 はデータ転送速度がメモリ・モジュールの 4 倍である信号を受信した後、それらをメモリ・モジュールのデータ転送速度の信号に変換する。

10

20

30

【0021】

図 12 はメモリ・コントローラ 802 からメモリ・モジュール 806 へ流下する情報を転送するのに本発明の典型的な実施形態が使用するダウンストリーム・フレームのフォーマットを示す。本発明の典型的な実施形態では、ダウンストリーム・フレームは 8 個の転送から構成されている。各転送は 22 個の信号と 1 個の平衡型クロック（合計 24 本のワイヤ）を備えている。このフレームはさらに 8 本のコマンド・ワイヤ（c0～c7）1208、9 本のデータ・ワイヤ（d103～d18）1206、4 本のバス ECC（Error Correction Code：誤り訂正符号）ワイヤ（ecc0～ecc3）1204、および 1 本の予備ワイヤ（spare）1202 から成る。72 個のデータ・ビットは図 12 ではビット d10～d18 として示されているが、9 本のワイヤから成り、各フレームごとに各ワイヤ上には 8 個の転送が存在する。本発明の典型的な実施形態では、図 12 に示すフレーム・フォーマットを用いてメモリ・クロック・サイクル当たり 1 個または 2 個のメモリ・コマンドに加え 72 ビットの書き込みデータを転送する。各データ・ビット（および他のビット）の番号付けは特定の転送に加え使用するワイヤに基づいている。D34 は（ビット 0～8 のうちの）データ・ビット 3 と（転送 0～7 のうちの）転送 4 とを指している。コマンド・ビット・フィールドは c0～c7 として示されているが、8 個の転送にわたってモジュールに供給される 64 ビットの情報から成る。

30

40

50

【0022】

ECC ビット・フィールド（ecc0～ecc3）は 8 個の転送にわたる 32 個のビット配置場所から成るが、実際には 16 ビットから成るグループにフォーマットされている。16 ビットのパケットの各々は 4 本のワイヤの各々の上の 4 個の転送から成り、4 個のバス転送から成る各グループにわたりバス・レベルの障害の検出と訂正を行う。予備のビット配置場所はコマンド・フィールド、データ・フィールド、および ECC フィールド中のビット群を転送するのに使用する 21 本のワイヤ（ビット・レーン（bitlane）とも定義される）のうちの任意のものを論理的に置換するのに使用するが、それは、これらのビット・レーンのうちの 1 本で障害が発生し、それがシステムによって割り当て済みの障害閾値の限界を超えるエラーを引き起こすものである場合である。予備ワイヤを使用するのは直接に接続された 2 つのアセンブリの間（すなわちメモリ・コントローラ 802 とメモリ・モジュール 806a との間、または、メモリ・モジュール 806a～806d のうちの任意の 2 つのものの間）の故障したセグメントを置換するため、ならびに、ワイヤの故障

、コネクタの故障、はんだ相互接続部の故障、駆動器の故障、および、受信器の故障、または、これらの故障のうちのいくつかのものといったイベントに起因してワイヤを置換するためである。176個のビット配置場所が利用可能である場合、168個はメモリ・モジュール806への情報の転送用に利用することができ、それら168個のビット配置場所のうち、32個のビット配置場所はさらにバス転送自身に関するECC保護を実現するのに割り当てる。これにより、合計136個のビット配置場所がメモリ・モジュール806への情報の転送用に使用しうるようになる。図12に示すフレーム・フォーマットはメモリ・コントローラの方向からメモリ・モジュール806に入来する信号、および、ダウンストリーム（下流）にある任意のメモリ・モジュール806へ外出する信号に適用することができる。

10

【0023】

図13はメモリ・モジュール806からメモリ・コントローラ802またはアップストリーム・メモリ・モジュール806へ流上する情報を転送するのに本発明の典型的な実施形態が使用するアップストリーム・フレームのフォーマットを示す図である。本発明の典型的な実施形態では、アップストリーム・フレームは8個の転送から構成されている。各転送は23個の信号と1個の平衡型クロック（合計25本のワイヤ）を備えている。このフレームはさらに18本のコマンド・ワイヤ（d00～d017）1306、4本のバスECCワイヤ（ecc0～ecc3）1304、および1本の予備ワイヤ（spare）1302から成る。本発明の典型的な実施形態では、図13に示すフレーム・フォーマットを用いてメモリ・クロック・サイクル当たり144個の読み取りデータ・ビットを転送する。各データ・ビット（および他のビット）の番号付けは特定の転送に加え使用するワイヤに基づいている。D34は（ビット0～17のうちの）データ・ビット3と（転送0～7のうちの）転送4とを指している。

20

【0024】

ECCビット・フィールド（ecc0～ecc3）は8個の転送にわたる32個のビット配置場所から成るが、実際には16ビットから成るグループにフォーマットされている。16ビットのパケットの各々は4本のワイヤの各々の上の4個の転送から成り、4個の転送ごとにエラーの訂正を行う。予備のビット配置場所はコマンド・フィールド、データ・フィールド、およびECCフィールド中のビット群を転送するのに使用する22本のワイヤのうちの任意のものを論理的に置換するのに使用する。ただし、これを行うのは、これらのワイヤのうちの1本で障害が発生し、それが実質的に無矛盾のものである場合である。障害が実質的に無矛盾であると考えられるのは、それがシステムによって決まる閾値（たとえば当該故障が検出される回数）を超えている場合だけである。単一のビット・レーンの障害はバス・レベルのECCによって動作中に訂正する。一方、システムのサービスを担当する構成要素（たとえばサービス・プロセッサ）はシステムの運転中に生じる具体性のある（たとえば周期的な、繰り返される、そして、継続的な）障害を修理するために、故障中のセグメントを予備のものに切り替える決断を行う。予備のワイヤを使用するのは直接に接続された2つのアセンブリの間（すなわちメモリ・コントローラ802とメモリ・モジュール806aとの間、または、メモリ・モジュール806a～806dのうちの2つのものの間）の故障したセグメントを置換するため、ならびに、何らかの目的（たとえばワイヤの故障、コネクタの故障、はんだ相互接続部の故障、駆動器の故障、もしくは、受信器の故障、または、これらのすべて）に起因してワイヤを置換するためである。184個のビット配置場所が利用可能である場合、176個はメモリ・モジュール806への情報の転送用に利用することができ、それら176個のビット配置場所のうち、32個のビット配置場所はさらにバス転送自身に関するECC保護を実現するのに割り当てる。これにより、合計144個のビット配置場所がアップストリーム・メモリ・モジュール806またはメモリ・コントローラ802への情報の転送用に使用しうるようになる。

30

【0025】

真に「予備」すなわち「未使用」ではない（すなわち現在の構成において既存の機能を有する）予備のビット・レーンを備えた実施形態はほとんどないけれども、これらの実施

40

50

形態では、「予備の」ビット・レーンを起動すると、メモリ・システム全体のデータの一貫性が損なわれる。一例として、エラーの検出用に使用するビット群の一部分を、当該ビット群を伝達しているワイヤの再割り当てによって、除去するシステムがあるが、その結果、障害を検出する能力が低下するとともに、検出不能のデータの破損（「無音のデータ破損（silent data corruption）」）が生じる確率がきわめて高くなる（最大200倍またはそれを超える）。この手法は高レベルのデータの一貫性とシステムの可用性とが求められるアプリケーションの場合には賢明ではないと考えられる。

【0026】

本発明の典型的な実施形態では、メモリ・コントローラ802と第1のメモリ・モジュール806との間、および、システム中の相互接続された任意の2つのメモリ・モジュール806の間の複数の相互接続部が故障する場合におけるメモリ・システムの耐性を最大化させている。図14は本発明の典型的な実施形態が使用するセグメント・レベルの予備化を施したカスケード型のモジュール・バス構成を示す。図14にはメモリ・コントローラ802、カスケード型ダウンストリーム・メモリ・バス904、カスケード型アップストリーム・メモリ・バス902、4個（これよりも多くても少なくても可）のメモリ・モジュール806a～806d、および、ダウンストリーム・メモリ・バス904上の2つのアセンブリごとに一意のワイヤを有する一連の被置換セグメント1402を備えたメモリ・システムの簡単化した図が含まれている。

【0027】

本発明の典型的な実施形態は図12および図13に示す予備のワイヤを一意の態様で割り当てる能力を実現し、アップストリーム・メモリ・バス902およびダウンストリーム・メモリ・バス904の双方の上にある任意の2つのアセンブリの間で（すなわちメモリ・コントローラ802とメモリ・モジュール806aとの間で、あるいは、任意の2つのメモリ・モジュール806a～806dの間で）故障中のセグメントを置換することを可能にするものである。アップストリーム・メモリ・バス902およびダウンストリーム・メモリ・バス904の上の予備のワイヤには故障中のセグメント群から成る異なる組を割り当てる。これにより、メモリ・システムは延長期間において、当該システムの一部が故障した後であっても、動作を継続することが可能になる。バス・レベルのECC（これはビットまたはワイヤの障害の存在下におけるバスの継続動作を可能とする）を使用することにより、ダウンストリーム・フレーム（およびアップストリーム・フレーム）中の予備のビット・レーンの利用可能性、ならびに、（同一のカスケード接続されたバス上の他の任意の2つのアセンブリの間でセグメントを置換する能力を保持しながら）任意の2つのアセンブリの間で单一のセグメントを置換する能力、高レベルの信頼性および耐性が得られる。

【0028】

端から端まで（カスケード・チェーン中のメモリ・コントローラから最後のDIMMまで）ビット・レーンを置換すると、相互接続の故障を訂正する、より単純な方法が実現する。8個のDIMMから成るメモリ・チャネルにおいて、セグメント・レベルの予備化を使用すると独立した8つのセグメント置換により最大8個の一意の故障を迂回することが可能になるが、完全なビット・レーンを端から端まで置換してもたった1つの故障しか迂回することができない。

【0029】

本発明の別の典型的な実施形態は高速クロックを予備化することを含むが、それは故障情報と診断とが、当該クロックが故障中である旨を示す場合である。一般に、クロックの故障は冗長なコネクタ接続を使用することにより既に最小化されている。というのは、多くの場合、コネクタ接続はチャネル中で生じる具体性のある故障を繰り返す原因の大部分を占めているからである。

【0030】

図15はセグメント・レベルの予備化を実現するために本発明の典型的な実施形態が使用する多モード型バッファ装置1002の高レベルの論理フローのブロック図である。多

10

20

30

40

50

モード型バッファ装置 1002（バッファ・モジュールとも呼ばれる）は任意のセグメント用に一意の予備化を行うのを可能にする4つのバス予備化論理ブロック（1526、1536、1538、1540）を備えている。また、バッファ・モジュール 1002 はアップストリーム・バス ECC 機能ブロック 1522 とダウンストリーム・バス ECC 機能ブロック 1520 を備えている。これらはバッファ装置 1002 を通じて受信または受け渡しをするすべての情報（すなわち信号）を処理して、バスのエラーが存在するか否かを判断するのを可能にしている。アップストリーム・バス ECC 機能ブロック 1522 とダウンストリーム・バス ECC 機能ブロック 1520 はバス・レベルのエラーの検出・訂正コードを使用することにより、バス・レベルの故障の検出と訂正を行う。これにより、故障している特定のセグメントを識別することが可能になる。これに続いて、セグメントの置換を行うことができる。図面下部の左および右の部分にあるブロック（1524、1528、1530、1534）は高速バス 804 の受信または駆動に関連付けられている。「アップストリーム」はメモリ・コントローラ 802 の方向に情報を受け渡すバス 902 を指しており、「ダウンストリーム」はメモリ・コントローラ 802 から遠ざけて情報を受け渡すバス 904 を指している。

10

20

30

40

【0031】

図 15 を参照すると、アップストリーム・メモリ・アセンブリ（すなわちメモリ・モジュール 806）が出すデータ信号、コマンド信号、アドレス信号、ECC 信号、およびクロック信号をダウンストリーム・メモリ・バス 904 から受信器モジュール 1524 中に受信する。受信器機能ブロック 1524 はダウンストリーム・メモリ・バス 904 用のマクロとサポート論理を備えているが、本発明の一実施形態は 22 ビットの全速力・スレーブ型受信器バス用のサポートを提供している。受信器機能ブロック 1524 はクロック論理・分配機能ブロック 1518 にクロック信号を転送する（たとえば 4 対 1 型のクロック信号を生成する）。クロック論理・分配機能ブロック 1518 はパーベイシブ・その他信号 1510 が出すデータ入力も受信する。これらの信号には通常、クロック分配 PLL、BIST (built-in self-test: 内蔵型自己試験) 用の試験入力、プログラム可能なタイミング設定などのための制御情報およびセットアップ情報が含まれている。受信器機能ブロック 1524 はデータ信号、コマンド信号、ECC 信号、およびアドレス信号をバス予備化論理ブロック 1526 に転送するが、これは位置の変更に合わせて行い、適用可能な場合には、先行するアセンブリに由来する伝送の間に予備のワイヤが使用される際にデータのビット配置に合わせて行う。本発明の典型的な実施形態では、バス予備化論理ブロック 1526 は要すれば、信号の配置をシフトさせるマルチプレクサによって実装する。次いで、元のあるいは順番を入れ替えた信号を別のバス予備化論理ブロック 1536 に入力し、現在のメモリ・アセンブリとダウンストリーム（下流）にあるメモリ・アセンブリとの間に存在する相互接続部の故障の原因をなす信号の配置場所の変更（あるいは要すれば順番の入れ替え）を行う。次いで、元のあるいは順番を入れ替えた信号を駆動器機能ブロック 1528 に入力し、チェーン中にある次のメモリ・モジュール 806 へのダウンストリーム・メモリ・バス 904 を介した転送の用に供する。本発明の典型的な実施形態では、バス予備化論理 1536 はマルチプレクサを用いて実装する。駆動器機能ブロック 1528 はダウンストリーム・メモリ・バス 904 用のマクロとサポート論理を備えているが、本発明の典型的な実施形態は 22 ビットの高速・低遅延のカスケード型バス駆動器用のサポートを提供している。

【0032】

元のあるいは順番を入れ替えた信号をバス予備化論理 1536 に入力することに加え、バス予備化論理 1526 は当該フレーム用にエラーの検出と訂正を行うダウンストリーム・バス ECC 機能ブロック 1120 中への元のあるいは順番を入れ替えた信号の入力も行う。ダウンストリーム・バス ECC 機能ブロック 1120 はダウンストリーム・メモリ・バス 904 から受信する、あるいは、ダウンストリーム・メモリ・バス 904 から多モード型バッファ装置 1002 に受け渡す情報をすべて処理してバスのエラーが存在するか否かを判断する。ダウンストリーム・バス ECC 機能ブロック 1520 はバスの信号を分析

50

し、それらが有効であるか否かを判断する。次いで、ダウンストリーム・バスECC機能ブロック1520は訂正した信号をコマンド・ステート・マシン1514に転送する。コマンド・ステート・マシン1514はコマンドのデコードまたは競合に付随するエラー・フラグをペーベイシブ・その他機能ブロック1510に入力する。ダウンストリーム機能ブロックヒップストリーム機能ブロックはエラー・フラグもしくは（もしあれば）エラー・データまたは双方をペーベイシブ・その他機能ブロック1510に提示して、これらのエラーをメモリ・コントローラ、プロセッサ、サービス・プロセッサ、または他のエラー管理ユニットに報告するのを可能にすることも行う。

【0033】

図15を参照すると、ペーベイシブ・その他機能ブロック1510はエラー・フラグもしくはエラー・データまたは双方をメモリ・コントローラ802に転送する。チェーン中の各メモリ・モジュール806からエラー・フラグもしくはエラー・データまたは双方を収集することにより、メモリ・コントローラ802は、この構成における一部の実施形態では付加的な診断が完了しているとしても、さらなる診断を開始することなく、故障中のセグメント（群）を特定しうるようになる。また、故障の回数または故障の種別に対する、設定で選定済みの閾値（たとえば1回、2回、10回、または20回）に到達したら、ペーベイシブ・その他機能ブロック1510は通常、メモリ・コントローラ802からの入力に応答して、故障中のセグメントを予備のワイヤで置換する。本発明の典型的な実施形態では、4個の転送から成るグループごとにエラーの検出と訂正を行うから、デコードして開始すべきオペレーションを、（1つのフレームを構成する）8個の転送のうちの半分が終了したら受信することができる。エラーの検出と訂正是ダウンストリーム・メモリ・バス904からメモリ・モジュール806中に受け渡されるすべての信号に対して、特定のメモリ・モジュール806が処理することになっているか否かとは無関係に行う。訂正済み信号中のデータ・ビットはダウンストリーム・バスECC機能ブロック1520が書き込みデータ・バッファ1512に入力する。

【0034】

コマンド・ステート・マシン1514は訂正済みの信号（データ信号、コマンド信号、およびアドレス信号を含む）がメモリ・モジュール806に宛てられており、当該モジュールが処理することになっているか否かの判断も行う。訂正済みの信号がメモリ・モジュール806に宛てられている場合、コマンド・ステート・マシン1514はとるべきアクションを決め、DRAMアクション、ライト・バッファ・アクション、リード・バッファ・アクション、または、それらの組み合わせを開始する。ライト・データ・バッファ1514はメモリ・データ・インターフェース1506にデータ信号を転送し、コマンド・ステート・マシン1514は付随するアドレス信号とコマンド信号をDRAMの使用に適合させてメモリ・コマンド・インターフェース1508に転送する。上述したように、右側のコマンド1106は通常、右側アドレス・コマンド・バス1102を介してメモリ・モジュール806の右側に転送し、左側のコマンド1102は左側アドレス・コマンド・バス1106を介してメモリ・モジュール806の左側に転送する。ただし、その他のモジュール構成が存在しうる。

【0035】

コントローラ802に転送されるデータ信号はメモリ・モジュール806がコマンド（たとえばリード（読み取り）コマンド）を実行した後、メモリ装置の「リード（読み取り）」タイミングに適合しるように、リード・データ・バッファ1516に一時的に格納する。リード・データ・バッファ1516は読み取ったデータをアップストリーム・バスECCモジュール1522に転送する。アップストリーム・バスECC機能ブロック1522はリード・データ・バッファ1516中の信号用にチェック・ビットを生成する。このチェック・ビットとリード・データ・バッファ1516中の信号とをアップストリーム・データ多重化機能ブロック1532に入力する。アップストリーム・データ多重化機能ブロック1532はアップストリーム・メモリ・バス902上のデータをバス予備化論理1538と駆動器機能ブロック1530とを介してマージする。要すれば、バス予備化論理

10

20

30

40

50

1538は現在のメモリ・モジュール806と上流にある受信中のモジュール(あるいはメモリ・コントローラ)と間のセグメントの故障の原因をなす信号をリダイレクトする。駆動器機能ブロック1530は元のあるいは順番を入れ替えた信号をアップストリーム・メモリ・バス902を介してチェーン中の次のアセンブリ(すなわちメモリ・モジュール806またはメモリ・コントローラ802)に転送する。本発明の典型的な実施形態では、バス予備化論理1538は信号をシフトさせるマルチプレクサを用いて実装する。駆動器機能ブロック1530はアップストリーム・メモリ・バス902用のマクロとサポート論理を備えているが、本発明の典型的な実施形態は23ビットの高速低遅延カスケード型駆動器バス用のサポートを提供している。

【0036】

アップストリーム・メモリ・バス902からは任意のアップストリーム・メモリ・モジュール806中の任意のアップストリーム多モード型バッファ装置1002もデータ信号、クロック信号、およびECC信号を受信する。これらの信号は次のメモリ・モジュール806またはメモリ・コントローラ802に渡し上げる必要がある。図15を参照すると、ダウンストリーム・アセンブリ(すなわちメモリ・モジュール806)が出力データ信号、ECC信号、およびクロック信号はアップストリーム・メモリ・バス902上の受信器モジュール1534中に受信する。受信器機能ブロック1534はアップストリーム・メモリ・バス902用のマクロとサポート論理を備えているが、本発明の典型的な実施形態は23ビットの高速・スレーブ型の受信器バス用のサポートを含んでいる。受信器モジュール1534はデータ信号とECC信号をバス予備化モジュール1540を通じてアップストリーム・データ多重化機能ブロック1532に、次いでバス予備化論理ブロック1538に渡す。これらの信号は駆動器機能ブロック1530を介してアップストリーム・メモリ・バス902に転送する。

【0037】

データ信号とECC信号をアップストリーム・データ多重化機能ブロック1532に渡すのに加え、被訂正バス予備化機能ブロック1540は元のあるいは順番を入れ替えたデータ信号とECC信号をアップストリーム・バスECC機能ブロック1522に入力し、当該フレーム用にエラーを検出して訂正することも行う。アップストリーム・バスECC機能ブロック1522はアップストリーム・メモリ・バス902から受信する、あるいはアップストリーム・メモリ・バス902からバッファ・モジュール1002に受け渡すすべての情報を処理して、バスのエラーが存在するか否かを判断する。アップストリーム・バスECC機能ブロック1522はデータ信号とECC信号を分析して、それらが有効であるか否かを判断する。次いで、アップストリーム・バスECC機能ブロック1522はエラー・フラグもしくはエラー・データまたは双方をすべてパーベイシブ・その他機能ブロック1510に転送し、メモリ・コントローラ802への転送の用に供する。さらに、故障の回数または種別用の事前定義の閾値を超えた後、パーベイシブ・その他機能ブロック1510は通常、メモリ・コントローラ802の指示に応答して、故障中のセグメントを予備のセグメントで置換する。

【0038】

メモリ・コントローラ802-メモリ・モジュール806間のバス、または、メモリ・モジュール806-メモリ・モジュール806間のバスの各々は、ダウンストリーム・フレーム1202およびアップストリーム・フレーム1302中に定義された、予備の信号によって置換される一意のビット・レーンのセグメントを有する。図15のブロック図は本発明の典型的な実施形態が使用するバッファ・モジュール1002の一実現例である。本発明の範囲から逸脱することなく、他の実現例が可能である。

【0039】

カスケード型のメモリ・システムにおいて低遅延を達成するために、本発明の典型的な実施形態は予備ワイヤの置換を促進しうるように設計した機構を使用している。バス中の任意の故障セグメントを置換するのに予備信号を使用するのではなく、バスの一部分を1ビット位置だけシフトさせ、多数の選択可能な入力を備えた多重化機能を必要としない方

10

20

30

40

50

法で予備信号の使用を開始しうるようにしている。駆動器バスおよび受信器バスの各 1 ビットについて単一の 2 対 1 型のセレクタを使用している。予備化オペレーションを実行するとき、置換すべきセグメントの所在をレジスタにロードする。この値は各ビット用に選択したマルチプレクサ中にエンコードした優先度である。（バス予備化論理モジュール 1536、1538 が実行する）送信側では、被置換セグメントよりも上位のビットをそれらの元の位置にシフトしなおす。（バス予備化論理・モジュール 1526、1540 が実行する）受信側では、被置換セグメントよりも上位のビットをそれらの元の位置にシフトしなおす。

【0040】

本発明の典型的な実施形態はセグメント・レベルの予備化を実現するものである。ビットまたはワイヤの障害の存在下における継続的な動作を可能にするバス・レベルの ECC の使用を通じて、ダウンストリーム・フレーム（およびアップストリーム・フレーム）において予備のビット・レーンが利用可能になること、および、任意の 2 つのアセンブリの間で、同一のカスケード型バス上において別の任意の 2 つのアセンブリの間でセグメントを置換する能力を保持しながら、共通のセグメントを置換する能力とによって、メモリ・システムの信頼性と耐性が向上する。バス・レベルの ECC を使用することにより、コントローラは一般に診断ツールを実行することなく、故障したセグメントを特定しうるようになる。これにより、メモリ・システム中の相互接続部の故障からの復旧時間が短くなるとともに、頻繁には生じない故障セグメントを正確に特定する確率が高くなる。

【0041】

上述したように、本発明の実施形態はコンピュータによって実現されるプロセス、およびそれらのプロセスを実行する装置の形態で実現することができる。また、本発明の実施形態は有形の媒体（たとえばフロッピー（R）ディスクケット、CD-ROM、ハード・ディスク駆動装置、または、任意の他のコンピュータ読み取り可能な記憶媒体）中に記録した命令群を含むコンピュータ・プログラム・コードであって、当該コンピュータ・プログラム・コードをコンピュータにロードして実行すると、当該コンピュータは本発明を実施する装置になる、コンピュータ・プログラム・コードの形態でも実現することができる。さらに、本発明はコンピュータ・プログラム・コードの形態でも実現することができるが、その場合、たとえば記憶媒体中に格納するものであるか、あるいは、コンピュータにロードし、もしくは、コンピュータが実行し、または、双方を行うものであるか、あるいは、何らかの伝送媒体（たとえば電気的な配線またはケーブル、光ファイバ、電磁波など）中を伝送するものであるかとは無関係である。汎用マイクロプロセッサ上に実装すると、当該コンピュータ・プログラム・コードのセグメント群は特定の論理回路を形成するよう当該マイクロプロセッサを構成する。

【0042】

以上、典型的な実施形態を参照して本発明を記述したが、本発明の範囲を逸脱することなく、様々な変更をなしうるとともに、それらの構成要素を均等物で置換しうる、ということは当業者が理解しうる。また、本発明の本質的な範囲を逸脱することなく、特定の状況または材料を本発明の教示に適応させる多くの変更をなすことができる。したがって、本発明は本発明を実施するのに最良の形態であると考えられるものとして開示した特定の実施形態に限定されない。本発明は特許請求の範囲の内に含まれる実施形態をすべて含むことができる。さらに、用語「第 1 の」や「第 2 の」などを使用しているが、それらは順序を示すものでも重要性を示すものでもない。そうではなく、用語「第 1 の」や「第 2 の」などは 1 つの構成要素を他の構成要素から区別するために使用されている。

【図面の簡単な説明】

【0043】

【図 1】個別のポイント・ツー・ポイント型リンクを介して 2 個のバッファ付きメモリ・アセンブリに接続された従来技術に係るメモリ・コントローラを示す図である。

【図 2】バッファ装置を備えた従来技術に係る同期型メモリ・モジュールを示す図である。

10

20

30

40

50

【図3】登録済みDIMMを使用する従来技術に係るメモリ・システムを示す図である。

【図4】ポイント・ツー・ポイント型チャネル、登録済みDIMM、および2:1型バス速度増倍器を備えた従来技術に係るメモリ・システムを示す図である。

【図5】マルチドロップ型メモリ「スタブ」バスを使用する従来技術に係るメモリ構成を示す図である。

【図6】そうでない場合には複数のポートが必要になるマルチポイント型の通信構成における従来技術に係るディジー・チェーン構成を示す図である。

【図7】メモリ・コントローラとメモリ・モジュールとの間における従来技術に係るディジー・チェーン接続を示す図である。

【図8】本発明の典型的な実施形態が使用するカスケード型のメモリ構成を示す図である。10

【図9】本発明の典型的な実施形態が使用する、カスケード接続されたメモリ・モジュールと一方向性のバスとを備えたメモリ構成を示す図である。

【図10】本発明の典型的な実施形態が使用するバッファ付きメモリ・モジュールを示す図である。

【図11】本発明の典型的な実施形態が使用する、バッファ付きモジュールの配線システムを示す図である。

【図12】本発明の典型的な実施形態が使用するダウンストリーム・フレームのフォーマットを示す図である。

【図13】本発明の典型的な実施形態が使用するアップストリーム・フレームのフォーマットを示す図である。20

【図14】本発明の典型的な実施形態が使用するセグメント・レベルの予備化を施したカスケード型のモジュール・バス構成を示す図である。

【図15】本発明の典型的な実施形態が使用する、バッファ付きチップの高レベルの論理フローのブロック図である。

【符号の説明】

【0044】

1 1次ステーション

2A 2次ステーション

2B 2次ステーション

2C 2次ステーション

2D 2次ステーション

2E 2次ステーション

2F 2次ステーション

3A モデム

3B モデム

10 メモリ・システム

12 バッファ

20 メモリ・コントローラ

30 クロック・バッファ

402 メモリ・コントローラ

406 バス・コンバータ

408 メモリ・バス

410 少なくとも1つの登録済みDIMM

802 メモリ・コントローラ

804 約50本のワイヤ

806 約50本のワイヤ

902 アップストリーム

904 ダウンストリーム

1002 バッファ

30

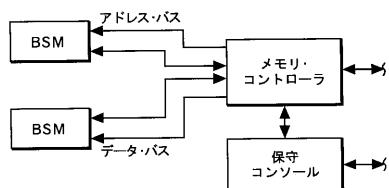
40

50

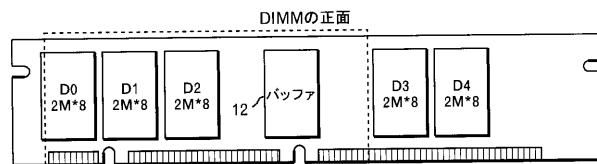
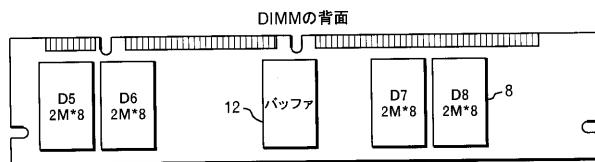
1 0 0 6	前面図	
1 0 0 8	背面図	
1 3 0 2	予備ビット	
1 3 0 4	E : ECC チェック・ビット	
1 3 0 6	D : 書き込みデータ・ビット	
1 3 0 8	C : コマンド・ビット	
1 4 0 2	被置換セグメント	
1 5 0 6	メモリ・データ・インターフェース	
1 5 0 8	メモリ・コード・インターフェース	
1 5 1 0	ペーベイシブ・その他	10
1 5 1 2	書き込みデータ・バッファ (4 個、アドレス可能)	
1 5 1 4	コマンド・ステート・マシン	
1 5 1 6	読み取りデータ・バッファ (2 個、フロー制御付き)	
1 5 1 8	クロック論理・分配	
1 5 2 0	ダウンストリーム・バス ECC (受け渡し信号用に並列)	
1 5 2 2	アップストリーム・バス ECC (受け渡し信号用に並列)	
1 5 2 4	受信器	
1 5 2 6	バス予備化	
1 5 2 8	駆動器	
1 5 3 0	駆動器	20
1 5 3 2	アップストリーム・データ多重化	
1 5 3 4	受信器	
1 5 3 6	バス予備化	
1 5 4 0	バス予備化	

【図 1】

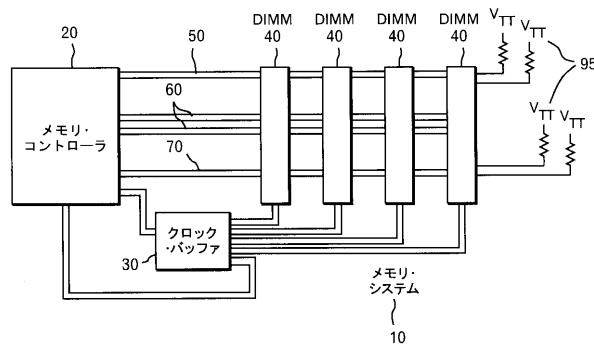
典型的な大型システムのメモリ構成



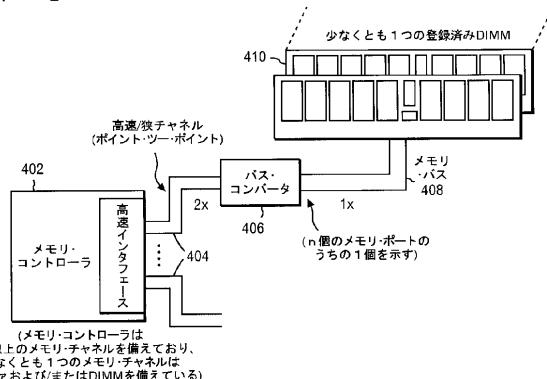
【図 2】



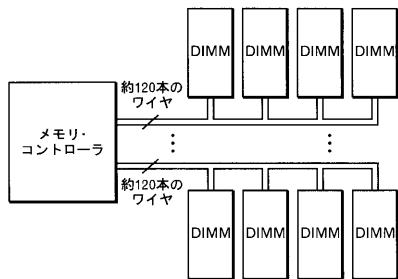
【図 3】



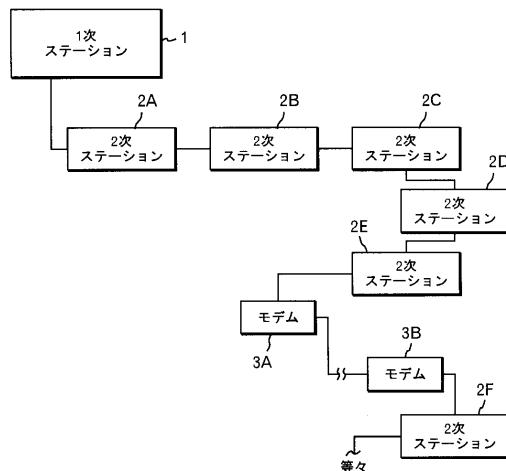
【図 4】



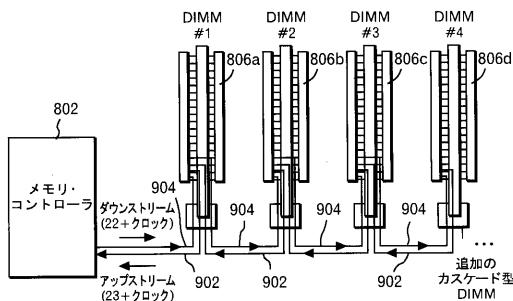
【図5】



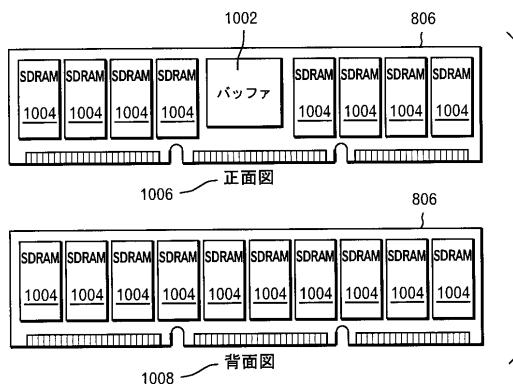
【図6】



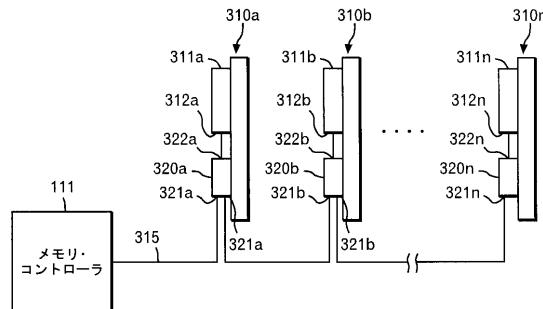
【図9】



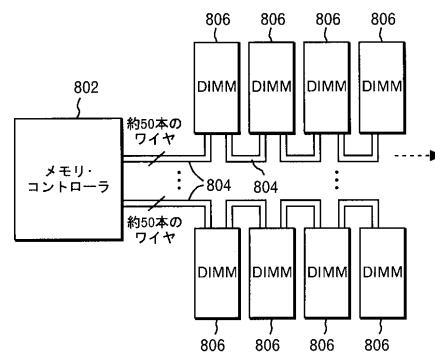
【図10】



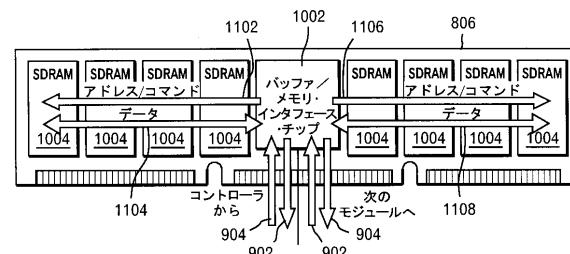
【図7】



【図8】



【図11】



【図12】

ダウンストリーム・フォーマット 9個のデータ-8個のコマンド-4個のECC-1個の予備-1個の平衡クロック = 合計24本のワイヤ	
S3 us(#)	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
d3 s3(#)	ecc0 ecc1 ecc2 ecc3 d10 d11 d12 d13 d14 d15 d16 d17 d18 c0 c1 c2 c3 c4 c5 c6 c7
0	S0 E00 E10 E20 E30 D00 D10 D20 D30 D40 D50 D60 D70 D80 C00 C10 C20 C30 C40 C50 C60 C70
1	S1 E01 E11 E21 E31 D01 D11 D21 D31 D41 D51 D61 D71 D81 C01 C11 C21 C31 C41 C51 C61 C71
2	S2 E02 E12 E22 E32 D02 D12 D22 D32 D42 D52 D62 D72 D82 C02 C12 C22 C32 C42 C52 C62 C72
3	S3 E03 E13 E23 E33 D03 D13 D23 D33 D43 D53 D63 D73 D83 C03 C13 C23 C33 C43 C53 C63 C73
4	S4 E04 E14 E24 E34 D04 D14 D24 D34 D44 D54 D64 D74 D84 C04 C14 C24 C34 C44 C54 C64 C74
5	S5 E05 E15 E25 E35 D05 D15 D25 D35 D45 D55 D65 D75 D85 C05 C15 C25 C35 C45 C55 C65 C75
6	S6 E06 E16 E26 E36 D06 D16 D26 D36 D46 D56 D66 D76 D86 C06 C16 C26 C36 C46 C56 C66 C76
7	S7 E07 E17 E27 E37 D07 D17 D27 D37 D47 D57 D67 D77 D87 C07 C17 C27 C37 C47 C57 C67 C77

S:予備ビット E:ECCチェック・ビット D:書き込みデータ・ビット C:コマンド・ビット

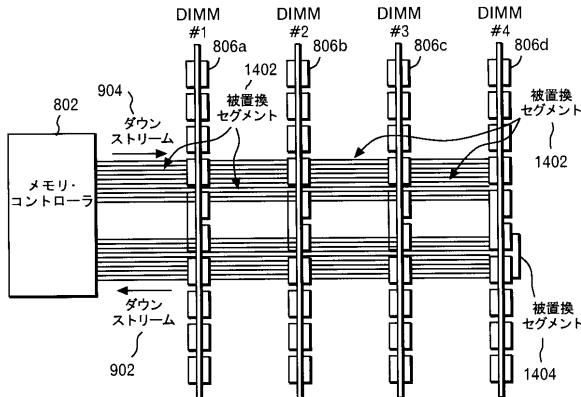
【図13】

アッピストリーム・フォーマット：	
16個のデータ +4個のECC +1個の予備 +1個の平衡クロック = 合計25本のワイヤ	
s3_us(#)	0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
d3_s3(#)	ecc0 ecc1 ecc2 ecc3 do0 do1 do2 do3 do4 do5 do6 do7 do8 do9 do10 do11 do12 do13 do14 do15 do16 do17
Xfer/Bit	予備 ecc0 ecc1 ecc2 ecc3 do0 do1 do2 do3 do4 do5 do6 do7 do8 do9 do10 do11 do12 do13 do14 do15 do16 do17
0	S0 E0 D1 E20 E30 D0 D10 D20 D30 D40 D50 D60 D70 D80 D90 D100 D110 D120 D130 D140 D150 D160 D170
1	S1 E01 E11 E21 E31 D01 D11 D21 D31 D41 D51 D61 D71 D81 D91 D101 D111 D121 D131 D141 D151 D161 D171
2	S2 E02 E12 E22 E32 D02 D12 D22 D32 D42 D52 D62 D72 D82 D92 D102 D112 D122 D132 D142 D152 D162 D172
3	S3 E03 E13 E23 E33 D03 D13 D23 D33 D43 D53 D63 D73 D83 D93 D103 D113 D123 D133 D143 D153 D163 D173
4	S4 E04 E14 E24 E34 D04 D14 D24 D34 D44 D54 D64 D74 D84 D94 D104 D114 D124 D134 D144 D154 D164 D174
5	S5 E05 E15 E25 E35 D05 D15 D25 D35 D45 D55 D65 D75 D85 D95 D105 D115 D125 D135 D145 D155 D165 D175
6	S6 E06 E16 E26 E36 D06 D16 D26 D36 D46 D56 D66 D76 D86 D96 D106 D116 D126 D136 D146 D156 D166 D176
7	S7 E07 E17 E27 E37 D07 D17 D27 D37 D47 D57 D67 D77 D87 D97 D107 D117 D127 D137 D147 D157 D167 D177

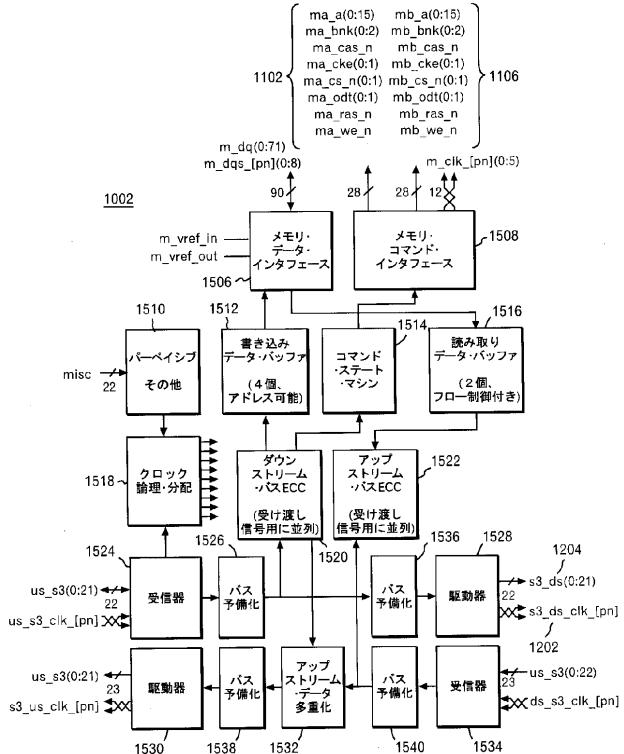
S:予備ビット E:ECCチェック・ビット D:書き込みデータ・ビット

1302 1304 1306

【図14】



【図15】



フロントページの続き

(72)発明者 ティモシー・ジェイ・デル

アメリカ合衆国 05446 バーモント州 コルチェスター、パークウッド ドライブ 101

(72)発明者 フランク・ディー・フェライオロ

アメリカ合衆国 12553 ニューヨーク州 ニュー ウィンザー、デニストン ドライブ 1
70

(72)発明者 ケビン・シー・ガウアー

アメリカ合衆国 12540 ニューヨーク州 ラグランジュヴィル、アラリー ロード エクス
テンション 8

(72)発明者 ケビン・ダブリュー・カーケ

アメリカ合衆国 12603 ニューヨーク州 ポーキープシー、コーチライト ドライブ 31

(72)発明者 マーク・ダブリュー・ケロッグ

アメリカ合衆国 14467 ニューヨーク州 ヘンリエッタ、ヒリントン コート 20

(72)発明者 ウォーレン・イー・モール

アメリカ合衆国 78613 テキサス州 シダー パーク、タク ロード 2701

F ターム(参考) 5B018 GA02 GA06 HA14 HA35 KA02 KA30 MA32 MA33 MA34 MA35

NA02 QA01