

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5116760号
(P5116760)

(45) 発行日 平成25年1月9日(2013.1.9)

(24) 登録日 平成24年10月26日(2012.10.26)

(51) Int. Cl.	F I
G 0 6 F 12/02 (2006.01)	G O 6 F 12/02 5 6 O C
G 0 6 F 12/08 (2006.01)	G O 6 F 12/08 5 0 5 B
	G O 6 F 12/08 5 3 1 Z
	G O 6 F 12/08 5 5 1 B

請求項の数 15 (全 30 頁)

(21) 出願番号	特願2009-511506 (P2009-511506)	(73) 特許権者	390009531
(86) (22) 出願日	平成19年5月22日 (2007.5.22)		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
(65) 公表番号	特表2009-537921 (P2009-537921A)		INTERNATIONAL BUSINESS MACHINES CORPORATION
(43) 公表日	平成21年10月29日 (2009.10.29)		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
(86) 国際出願番号	PCT/EP2007/054929	(74) 代理人	100108501
(87) 国際公開番号	W02007/135144		弁理士 上野 剛史
(87) 国際公開日	平成19年11月29日 (2007.11.29)	(74) 代理人	100112690
審査請求日	平成22年2月22日 (2010.2.22)		弁理士 太佐 種一
(31) 優先権主張番号	11/419,586	(74) 代理人	100091568
(32) 優先日	平成18年5月22日 (2006.5.22)		弁理士 市位 嘉宏
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 リモート・プリフェッチ・バッファを提供するためのシステム及び方法

(57) 【特許請求の範囲】

【請求項1】

少なくとも1つのメモリ・サブシステムに対するメモリ・アクセス要求の生成、前記メモリ・サブシステムからの非送信請求データ転送の受信及び前記非送信請求データ転送に対する応答をするメモリ・コントローラであって、将来のデータ参照に備えてプリフェッチ・データを格納するデータ・キャッシュを備えるメモリ・コントローラと、

前記メモリ・コントローラに接続された1つ又は複数のメモリ・バスと、

前記1つ又は複数のメモリ・バスを介して前記メモリ・コントローラと通信する前記メモリ・サブシステムであって、前記メモリ・サブシステムにおいて実施される、前記メモリ・サブシステムによって受信された先行メモリ・アクセス要求の分析に基づいて、前記メモリ・コントローラへの非送信請求データ転送を開始する1つ又は複数のメモリ・デバイス及び論理を含む、メモリ・サブシステムと、

を備え、

前記メモリ・サブシステムは、

前記メモリ・コントローラが送信するメモリ・アクセス要求を監視し、

前記メモリ・アクセス要求を検出しない場合に、前記非送信請求データ転送を前記メモリ・コントローラに送信し、

前記メモリ・アクセス要求である前記メモリ・デバイスに対する読み取り要求を検出した場合には、前記読み取り要求の対象のアドレス範囲と、前記非送信請求データ転送の対象のアドレス範囲とが一致するときに、前記読み取り要求の応答として、前記非送信請求

データ転送を前記メモリ・コントローラに送信する、コンピュータ・メモリ・システム。

【請求項 2】

前記メモリ・サブシステムにおける前記論理は、自立的に前記分析を実施し、前記非送信請求データ転送を開始する、請求項 1 に記載のコンピュータ・メモリ・システム。

【請求項 3】

前記非送信請求データ転送は、該データ転送と関連するアドレス情報を含む、請求項 1 または 2 に記載のコンピュータ・メモリ・システム。

【請求項 4】

前記非送信請求データ転送は、送信請求データ転送より低い優先順位で開始される、請求項 1 ~ 3 のいずれか 1 項に記載のコンピュータ・メモリ・システム。

10

【請求項 5】

前記メモリ・コントローラは、前記非送信請求データ転送を特定する論理を含む、請求項 1 ~ 4 のいずれか 1 項に記載のコンピュータ・メモリ・システム。

【請求項 6】

前記データ・キャッシュは、前記メモリ・コントローラと 1 つ又は複数の要求デバイスとの間における階層構造の 1 つ又は複数のキャッシュの論理パーティションであり、コンピュータ環境について確立されたポリシーに従って一貫して管理される、請求項 1 ~ 5 のいずれか 1 項に記載のコンピュータ・メモリ・システム。

【請求項 7】

前記データ・キャッシュは、前記メモリ・コントローラへのアクセス要求を開始することなく前記データ・キャッシュからメモリ・アクセス要求にサービスする命令を含む、請求項 6 に記載のコンピュータ・メモリ・システム。

20

【請求項 8】

前記データ・キャッシュは、前記メモリ・コントローラ内に組み込まれており、要求トラフィックに関する置換及び一貫性について前記メモリ・コントローラによって独立して管理される、請求項 1 ~ 7 のいずれか 1 項に記載のコンピュータ・メモリ・システム。

【請求項 9】

前記メモリ・コントローラは、前記メモリ・サブシステムにアクセスする代わりに、前記データ・キャッシュからのデータを用いてメモリ・アクセス要求にサービスする命令を含む、請求項 8 に記載のコンピュータ・メモリ・システム。

30

【請求項 10】

コンピュータ・メモリ・サブシステムであって、
前記メモリ・サブシステムと通信する 1 つ又は複数のメモリ・デバイスへの先行メモリ・アクセス要求の分析を実施し、前記分析の結果に基づいて前記メモリ・デバイスのうちの 1 つ又は複数への非送信請求データ読取りを開始する、プリフェッチ論理と、
メモリ・バスと通信して、前記非送信請求データ読取りによって戻されたデータを含むメモリ・コントローラへの非送信請求データ転送を開始するハブ・デバイスと、
を備え、

前記ハブ・デバイスは、

前記メモリ・コントローラが送信するメモリ・アクセス要求を監視し、

40

前記メモリ・アクセス要求を検出しない場合に、前記非送信請求データ転送を前記メモリ・コントローラに送信し、

前記メモリ・アクセス要求である前記メモリ・デバイスに対する読み取り要求を検出した場合には、前記読み取り要求の対象のアドレス範囲と、前記非送信請求データ転送の対象のアドレス範囲とが一致するときに、前記読み取り要求の応答として、前記非送信請求データ転送を前記メモリ・コントローラに送信する、コンピュータ・メモリ・サブシステム。

【請求項 11】

前記非送信請求データ転送は、前記非送信請求データ読取りと関連するアドレス情報を含む、請求項 10 に記載のコンピュータ・メモリ・サブシステム。

50

【請求項 1 2】

前記非送信請求データ転送は、送信請求データ転送より低い優先順位で開始される、請求項 1 0 または 1 1 に記載のコンピュータ・メモリ・サブシステム。

【請求項 1 3】

メモリ・システムにおいて非送信請求データ・パケットを駆動する方法であって、
1 つ又は複数のメモリ・デバイスへの先行メモリ・アクセス要求を分析する ステップ であって、前記分析はメモリ・モジュールにおいて実施される ステップ と、

前記メモリ・デバイスのうちの 1 つ又は複数に対応するアドレス範囲においてデータ読取りコマンドを開始する ステップ であって、前記開始は前記分析の結果に応答して行われる ステップ と、

メモリ・コントローラとメモリ・モジュールとを接続する上流バス上のアクティビティを監視するステップ と、

前記データ読取りコマンドによって戻されたデータを含む非送信請求データ・パケットをメモリ・コントローラに伝送する ステップ と、

を含み、

前記非送信請求データ・パケットを伝送するステップは、

前記上流バスがアイドル状態である場合に、前記非送信請求データ・パケットを送信し

前記アクティビティである前記メモリ・デバイスに対する読み取り要求を検出した場合には、前記読み取り要求の対象のアドレス範囲と、前記非送信請求データ・パケットのアドレス範囲とが一致するときに、前記読み取り要求の応答として、前記非送信請求データ・パケットを伝送する、方法。

【請求項 1 4】

保留中の書込み要求が前記アドレス範囲を指定する場合に、前記非送信請求データ・パケットを破棄する ステップ をさらに含む、請求項 1 3 に記載の方法。

【請求項 1 5】

前記非送信請求データ・パケットが限界を超えて古くなり、前記上流バスがアイドル状態ではない場合に、前記非送信請求データ・パケットを破棄する ステップ をさらに含む、請求項 1 3 または 1 4 に記載の方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、一般に、コンピュータ・メモリに関し、より具体的には、リモート・プリフェッチ・バッファを提供するためのシステム及び方法に関する。

【背景技術】

【0 0 0 2】

現在の高性能コンピュータ・メイン・メモリ・システムは、一般に、1 つ又は複数のメモリ制御要素を介して 1 つ又は複数のプロセッサに接続された、1 つ又は複数のダイナミック・ランダム・アクセス・メモリ (DRAM) デバイスで構成される。コンピュータ・システム全体の性能は、1 つ又は複数のプロセッサの性能 / 構造、1 つ又は複数のいずれかのメモリ・キャッシュ、1 つ又は複数の入出力 (I/O) サブシステム、1 つ又は複数のメモリ制御機能の効率性、1 つ又は複数のメイン・メモリ・デバイス、並びに、1 つ又は複数のメモリ相互接続インターフェースのタイプ及び構造を含む、コンピュータ構造の重要な要素の各々によって影響される。

【0 0 0 3】

メモリ・システム / サブシステムの設計及び / 又は構造を改善することによりシステム全体の性能及び密度を最大化することに対する、改善された及び / 又は新規の解決策を生み出すために、広範囲な研究及び開発努力への継続的な投資が業界によって行われている。新しいコンピュータ・システムが、機能の追加、性能の向上、ストレージの強化、運用コストの低減などをもたらすことに加えて、平均故障間隔 (MTBF) に関して既存のシ

10

20

30

40

50

システムより著しく優れているものであることを顧客が期待しているため、高可用性システムは、システム全体の信頼性に関する更なる課題を提示する。他の顧客要件として多いものは、メモリ・システム設計の課題をさらに厳しくするものであり、アップグレードの容易さ、及び、(空間、電力、及び冷却といった)システム環境への影響の低減といった項目を含む。

【0004】

図1は、本出願人に譲渡された、D e l l 他の特許文献1に関し、初期の同期型メモリ・モジュールを示す。図1において示されるメモリ・モジュールは、デュアル・インライン・メモリ・モジュール(D I M M)である。このモジュールは、同期型D R A M 8と、バッファ・デバイス12と、最適化されたピン配列と、高性能動作を容易にする相互接続及び容量分離方法とで構成される。本特許は、フェーズ・ロック・ループ(P L L)のようなデバイスを用いる、モジュールに関するクロック・リドライブの使用も説明している。

10

【0005】

図2は、本出願人に譲渡された、D e l l 他の特許文献2に関し、コンピュータ・システム10を示すものであり、コンピュータ・システム10は、バス40を介してメモリ・コントローラ14に直接(即ちポイント・ツー・ポイントで)接続された同期型メモリ・モジュール20を含み、さらに、メモリ・コントローラ14から受信したアドレス、データ、及び制御情報をバッファに入れるか、登録するか、又は他の形でこれらに作用する、(特定用途向け集積回路、即ち「A S I C」のような)論理回路24を含む。メモリ・モジュール20は、メモリ初期化プロセスの一部又は通常の動作中のいずれかの際に、集積回路間(I 2 C)制御バス34などの独立したバスを用いて、複数の選択可能なモード又はプログラム可能なモードで動作するようにプログラムすることが可能である。本特許は、メモリ・コントローラに直接接続された単一のメモリ・モジュールより多くのメモリ・モジュールを必要とするアプリケーションで使用されるときに、電界効果トランジスタ(F E T)スイッチを用いてバスからモジュールを電氣的に切断することにより、結果として生じるスタブを最小にすることができることに言及している。

20

【0006】

特許文献1に関連して、特許文献2は、更に、定められた機能(アドレス、コマンド、データ、存在検知など)の全てを単一のデバイスに統合する能力を示している。技術改良によって可能となり、また、この場合においてはモジュール密度及び/又は機能の追加を可能にする機能の統合は、業界では一般的に行われている。

30

【0007】

本出願人に譲渡された、G r u n d o n 他の特許文献3からの図3は、従来のマルチドロップ・スタブ・バス上に最大で4つの登録されたD I M M 40を含むメモリ・システム10の簡略化した図及び説明を示す。本サブシステムは、メモリ・コントローラ20と、外部クロック・バッファ30と、登録されたD I M M 40と、アドレス・バス50と、制御バス60と、データ・バス70とを含み、アドレス・バス50及びデータ・バス70上にターミネータ95を含む。図3においては単一のメモリ・チャンネルのみが示されるが、これらのモジュールを用いて作成されるシステムは、メモリ・コントローラからの2つ以上の別個のメモリ・チャンネルを含むことが多く、各々のメモリ・チャンネルは、所望のシステム機能及び/又は性能を達成するために、(単一のチャンネルにモジュールが集中したときには)単一で動作するか、又は(2つ又はそれ以上のチャンネルにモジュールが集中したときには)並列に動作するものであった。

40

【0008】

本出願人に譲渡された、B o n e l l a 他の特許文献4からの図4は、同期型メモリ・モジュール210及びシステム構造を示すものであり、リピータ・ハブ320が、バス321及び322を介したローカル・メモリ・デバイス301及び302へのアドレス、コマンド、及びデータのローカル・リドライブと、(他の図面及び特許書類の文章で説明される)ローカル・クロックの生成と、バス300を介したシステムの次のモジュール又は

50

コンポーネントへの適切なメモリ・インターフェース信号のリドライブとを含む。

【0009】

現在のコンピュータ・システムは、メモリ・デバイス（典型的にはDRAM）がメモリ・ハブ・デバイスに接続される、ハブ・チップ・ベースのメモリ・システムを用いることがある。メモリ・ハブ・デバイスは、通信チャンネルのネットワークを介して1つ又は複数のシステム・メモリ・コントローラに相互接続される。高周波数での信号の送信を容易にするために、チャンネルは、一般に、1つ又は複数の単一方向の出力（下流）と単一方向の入力（上流）とのポイント・ツー・ポイント・リンクから成る。システム・メモリにアクセスするための（アドレス、読取り/書込み指示、及び/又は、他のいずれかの属性を含む）処理及び/又は入出力（I/O）要求（アクセス要求と呼ばれる）は、1つ又は複数のシステム・メモリ・コントローラによってサービスされる。メモリ・コントローラは、アクセス要求を翻訳し、制限する。メモリ・コントローラは、システム性能を最適化するために、利用可能なメモリ・バンクへの要求をスケジューリングし、優先順位付けする。要求は、チャンネル・プロトコル（又は関連するインターフェース・プロトコル）によって具体的にエンコードされ、1つ又は複数の下流リンクを介して、選択されたメモリ・ハブ・デバイスに伝送される。書込み要求は、書込み要求に関連するが必ずしも同じパケットで送信されとは限らないデータを含む。読取り要求は、1つ又は複数の次の上流パケットによってメモリ・コントローラに返送されることになる予想データ応答を意味する。ターゲットとされた1つ又は複数のハブ・デバイスは、受信した要求を翻訳し、接続されているメモリ・デバイスを応答的に制御して、ハブからの書込みデータを格納させるか、又は読取りデータをハブに与える。

【0010】

メモリ・コントローラのデータ読取りアクセス待ち時間は、メモリ・コントローラに関連するプリフェッチ・バッファを持つことによって、短くすることができる。プリフェッチ・バッファは、一般に、データと関連するアドレス情報と他の属性とを格納するための1つ又は複数の論理エントリを有する小型の連想キャッシュ（associative cache）で構成される。メモリ・コントローラは、予想される将来の要求に対して最も短い待ち時間でデータをプリフェッチ・バッファに格納するために、自立的に、予測的に、及び/又は応答的に、メモリ・デバイスからデータを読取ることができる。

【特許文献1】米国特許第5,513,135号

【特許文献2】米国特許第6,173,382号

【特許文献3】米国特許第6,510,100号

【特許文献4】米国特許第6,587,912号

【特許文献5】米国特許公開第2004/0260909号

【特許文献6】米国特許第6,381,685号

【発明の開示】

【発明が解決しようとする課題】

【0011】

ハブ・デバイス内のプリフェッチ・バッファを用いてメモリ・デバイスに直接アクセスすることに関連する待ち時間を短くするハブ・ベースのメモリ・システムは、説明済みである（例えば、Lee他の特許文献5を参照されたい）。しかしながら、このスキームは、ハブ内に格納されたプリフェッチ・データを参照するためにメモリ・コントローラが1つ又は複数のハブ・デバイスと通信することに関連する著しい待ち時間を生じさせる。プリフェッチ・データ・キャッシュをメモリ・コントローラに持つことによって、このようなバッファをリモート・ハブ・デバイスに実装することに関連するアクセス待ち時間が短くなる。しかしながら、これらの技術はどちらも、メモリ・デバイスからプリフェッチ・バッファにデータをトランスポートするために、重要な入出力チャンネル・リソースを用いることを必要とし、それは、これらのリソースについてコンテンションが高いときであることが多い。従って、より短いメモリ待ち時間とより高いメモリ・スループットとを通してコンピュータ性能を改善するために、より効率的なハブ・ベースのメモリ・システムに

ついでにの必要性が存在する。

【課題を解決するための手段】

【0012】

実施形態は、メモリ・コントローラと、メモリ・コントローラに接続された1つ又は複数のメモリ・バスと、メモリ・バスを介してメモリ・コントローラと通信する少なくとも1つのメモリ・サブシステムと、を有するコンピュータ・メモリ・システムを含む。メモリ・コントローラは、非送信請求データ転送(unsolicited data transfer)を含むメモリ・アクセス要求を生成し、受信し、これに应答する。メモリ・サブシステムは、メモリ・サブシステムにおいて実施された、メモリ・サブシステムによって受信された先行メモリ・アクセス要求の分析に基づいて、メモリ・コントローラへの非送信請求データ転送を開始する、1つ又は複数のメモリ・デバイス及び論理を含む。

10

【0013】

実施形態は、また、プリフェッチ論理及びハブ・デバイスを含むコンピュータ・メモリ・サブシステムを含む。プリフェッチ論理は、メモリ・サブシステムと通信する1つ又は複数のメモリ・デバイスに対する先行メモリ・アクセス要求の分析を行う。プリフェッチ論理は、また、分析の結果に基づいて、メモリ・デバイスのうちの1つ又は複数への非送信請求データ読取りを開始する。ハブ・デバイスは、メモリ・コントローラへの非送信請求データ転送を開始するために、メモリ・バスと通信する。非送信請求データ転送は、非送信請求データ読取りによって戻されたデータを含む。

20

【0014】

実施形態は、また、メモリ・アクセス要求を生成し、受信し、これに应答するためのメモリ・コントローラを含む。メモリ・コントローラは、ローカル・プリフェッチ・データ・キャッシュと、メモリ・サブシステムからの非送信請求データ転送の処理を容易にするための非送信請求データ転送論理と、を含む。処理は、データ・パケットを受信することと、データ・パケットを非送信請求データ・パケットとして特定することと、を含む。非送信請求データ・パケットは、アドレス範囲からのデータを含む。処理は、また、アドレス範囲におけるデータがそのアドレス範囲への次の書込みコマンドによって上書きされていないことを検証することを含む。非送信請求データ・パケットの内容は、アドレス範囲におけるデータがそのアドレス範囲への次の書込みコマンドによって上書きされていないことを検証することによって格納される。

30

【0015】

実施形態は、更に、メモリ・コントローラにおいてデータ・パケットを受信することを含む、メモリ・アクセス要求の受信及びメモリ・アクセス要求への应答のための方法を含む。データ・パケットは、非送信請求データ・パケットであるとして特定され、アドレス範囲を含む。アドレス範囲におけるデータがそのアドレス範囲への次の書込みコマンドによって上書きされていないことが検証される。非送信請求データ・パケットの内容は、アドレス範囲におけるデータがそのアドレス範囲への次の書込みコマンドによって上書きされていないことを検証することによって格納される。

40

【0016】

実施形態は、更に、1つ又は複数のメモリ・デバイスへの先行メモリ・アクセス要求を分析することを含む、メモリ・システムにおいて非送信請求データ・パケットを駆動する方法を含み、分析はメモリ・モジュールにおいて実施される。データ読取りコマンドは、分析の結果に基づいて、メモリ・デバイスのうちの1つ又は複数のアドレス範囲において開始される。非送信請求データ・パケットは、上流バスがアイドル状態のときにメモリ・コントローラに伝送される。非送信請求データ・パケットは、データ読取りコマンドによって戻されたデータを含む。

【0017】

実施形態による他のシステム、方法、及び/又はコンピュータ・プログラムは、添付の

50

図面及び詳細な説明を参照すれば、当業者にとって明らかであるか、又は明らかとなるであろう。こうした全ての付加的なシステム、方法、及び/又はコンピュータ・プログラムは、本説明の中に含まれ、本発明の範囲内にあり、特許請求の範囲によって保護されることを意図される。

【 0 0 1 8 】

第1の態様の観点から、本発明は、非送信請求データ転送を含むメモリ・アクセス要求を生成し、受信し、これに回答するためのメモリ・コントローラと、メモリ・コントローラに接続された1つ又は複数のメモリ・バスと、1つ又は複数のメモリ・バスを介してメモリ・コントローラと通信する少なくとも1つのメモリ・サブシステムであって、メモリ・サブシステムにおいて実施される、メモリ・サブシステムによって受信された先行メモリ・アクセス要求の分析に基づいて、1つ又は複数のメモリ・コントローラへの非送信請求データ転送を開始するための1つ又は複数のメモリ・デバイス及び論理を含む、メモリ・サブシステムと、を備えるコンピュータ・メモリ・システムを含む。

10

【 0 0 1 9 】

本発明は、メモリ・サブシステムにおける論理が自立的に分析を実施し、非送信請求データ転送を開始する、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 0 】

本発明は、非送信請求データ転送が該データ転送と関連するアドレス情報を含む、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 1 】

本発明は、非送信請求データ転送がその非送信請求データ転送を非送信請求データ転送として特定するタグを含む、コンピュータ・メモリ・システムを提供することが好ましい。

20

【 0 0 2 2 】

本発明は、非送信請求データ転送が送信請求データ転送より低い優先順位で開始される、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 3 】

本発明は、非送信請求データ転送が、非送信請求データ転送を含むものとしてパケットを特定するタグを含むパケットとしてフォーマットされる、コンピュータ・メモリ・システムを提供することが好ましい。

30

【 0 0 2 4 】

本発明は、メモリ・コントローラが非送信請求データ転送を特定する論理を含む、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 5 】

本発明は、将来のデータ参照に備えてプリフェッチ・データを格納し、それにより、要求されたデータがデータ・キャッシュにおいて利用可能であるときにメモリ・サブシステムにアクセスしないようにするためのデータ・キャッシュをさらに備える、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 6 】

本発明は、データ・キャッシュが、メモリ・コントローラと1つ又は複数の要求デバイスとの間における階層構造の1つ又は複数のキャッシュの論理パーティションであり、コンピュータ環境について確立されたポリシーに従って一貫して管理される、コンピュータ・メモリ・システムを提供することが好ましい。

40

【 0 0 2 7 】

本発明は、データ・キャッシュが、メモリ・コントローラへのアクセス要求を開始することなくデータ・キャッシュからメモリ・アクセス要求にサービスする命令を含む、コンピュータ・メモリ・システムを提供することが好ましい。

【 0 0 2 8 】

本発明は、データ・キャッシュが、メモリ・コントローラ内に組み込まれており、要求トラフィックに関する置換及び一貫性についてメモリ・コントローラによって独立して管

50

理される、コンピュータ・メモリ・システムを提供することが好ましい。

【0029】

本発明は、メモリ・コントローラが、メモリ・サブシステムにアクセスする代わりに、データ・キャッシュからのデータを用いてメモリ・アクセス要求にサービスする命令を含む、コンピュータ・メモリ・システムを提供することが好ましい。

【0030】

本発明は、メモリ・サブシステムと通信する1つ又は複数のメモリ・デバイスへの先行メモリ・アクセス要求の分析を実施し、分析の結果に基づいてメモリ・デバイスのうちの1つ又は複数への非送信請求データ読取りを開始する、プリフェッチ論理と、非送信請求データ読取りによって戻されたデータを含むメモリ・コントローラへの非送信請求データ転送を開始するためにメモリ・バスと通信するハブ・デバイスとを含むコンピュータ・メモリ・サブシステムをさらに備えるコンピュータ・メモリ・システムを提供することが好ましい。

10

【0031】

本発明は、非送信請求データ転送が、非送信請求データ読取りと関連するアドレス情報を含む、コンピュータ・メモリ・システムを提供することが好ましい。

【0032】

本発明は、非送信請求データ転送が、該データ転送を非送信請求データ転送として特定するタグを含む、コンピュータ・メモリ・システムを提供することが好ましい。

【0033】

本発明は、非送信請求データ転送が送信請求データ転送より低い優先順位で開始される、コンピュータ・メモリ・システムを提供することが好ましい。

20

【0034】

本発明は、非送信請求データ転送が、非送信請求データ転送を含むものとしてパケットを特定するタグを含むパケットとしてフォーマットされる、コンピュータ・メモリ・システムを提供することが好ましい。

【0035】

本発明は、メモリ・アクセス要求を生成し、受信し、これに回答するためのメモリ・コントローラであって、ローカル・プリフェッチ・データ・キャッシュと、メモリ・サブシステムからの非送信請求データ転送の処理を容易にするための非送信請求データ転送論理であって、該処理は、データ・パケットを受信することと、アドレス範囲からのデータを含むデータ・パケットを非送信請求データ・パケットとして特定することと、アドレス範囲におけるデータが次の書込みコマンドによって上書きされていないことを検証することと、アドレス範囲におけるデータが該アドレス範囲において次の書込みコマンドによって上書きされていないことを検証したことに応答して、非送信請求データ・パケットの内容をローカル・プリフェッチ・データ・キャッシュに格納することとを含む、非送信請求データ転送と、を含むメモリ・コントローラをさらに備える、コンピュータ・メモリ・システムを提供することが好ましい。

30

【0036】

第2の態様の観点から、本発明は、メモリ・サブシステムと通信する1つ又は複数のメモリ・デバイスへの先行メモリ・アクセス要求の分析を実施し、分析の結果に基づいてメモリ・デバイスのうちの1つ又は複数への非送信請求データ読取りを開始する、プリフェッチ論理と、非送信請求データ読取りによって戻されたデータを含むメモリ・コントローラへの非送信請求データ転送を開始するためにメモリ・バスと通信するハブ・デバイスと、を備えるコンピュータ・サブシステムを提供する。

40

【0037】

本発明は、非送信請求データ転送が、非送信請求データ読取りと関連するアドレス情報を含む、コンピュータ・メモリ・サブシステムを提供することが好ましい。

【0038】

本発明は、非送信請求データ転送が、該データ転送を非送信請求データ転送として特定

50

するタグを含む、コンピュータ・メモリ・サブシステムを提供することが好ましい。

【0039】

本発明は、非送信請求データ転送が送信請求データ転送より低い優先順位で開始される、コンピュータ・メモリ・サブシステムを提供することが好ましい。

【0040】

本発明は、非送信請求データ転送が、非送信請求データ転送を含むものとしてパケットを特定するタグを含むパケットとしてフォーマットされる、コンピュータ・メモリ・サブシステムを提供することが好ましい。

【0041】

第3の態様の観点から、本発明は、メモリ・アクセス要求を生成し、受信し、これに
10 答するためのメモリ・コントローラであって、ローカル・プリフェッチ・データ・キャ
ッシュと、メモリ・サブシステムからの非送信請求データ転送の処理を容易にするための非
送信請求データ転送論理であって、該処理は、データ・パケットを受信することと、アド
レス範囲からのデータを含むデータ・パケットを非送信請求データ・パケットとして特定
することと、アドレス範囲におけるデータが次の書込みコマンドによって上書きされてい
ないことを検証することと、アドレス範囲におけるデータが該アドレス範囲において次の
書込みコマンドによって上書きされていないことを検証したことに応答して、非送信請求
データ・パケットの内容をローカル・プリフェッチ・データ・キャッシュに格納すること
とを含む、非送信請求データ転送と、を含むメモリ・コントローラを提供する。

【0042】

第4の態様の観点から、本発明は、メモリ・アクセス要求を受信し、これに応答するた
20 めの方法であって、メモリ・コントローラにおいてデータ・パケットを受信することと、
アドレス範囲からのデータを含むデータ・パケットを非送信請求データ・パケットとして
特定することと、アドレス範囲におけるデータが次の書込みコマンドによって上書きされ
ていないことを検証することと、アドレス範囲におけるデータが該アドレス範囲において
次の書込みコマンドによって上書きされていないことを検証することに応答して、非送信
請求データ・パケットの内容をローカル・プリフェッチ・データ・キャッシュに格納する
ことと、を含む方法を含む。

【0043】

本発明は、データ・パケットがパケット・タイプを示すタグを含み、特定することが、
30 そのタグが非送信請求データ・パケットを示すかどうかを判定することを含む、方法を提
供することが好ましい。

【0044】

本発明は、特定することが、メモリ・アクセス要求が未処理のメモリ・アクセス要求と
関連しないことを検証することを含む、方法を提供することが好ましい。

【0045】

本発明は、入ってくるデータ・パケットについて上流のデータ・チャネルを監視するこ
とをさらに含む方法を提供することが好ましい。

【0046】

第5の態様の観点から、本発明は、メモリ・システムにおいて非送信請求データ・パケ
40 ットを駆動する方法であって、1つ又は複数のメモリ・デバイスへの先行メモリ・アクセ
ス要求を分析することであって、分析がメモリ・モジュールにおいて実施されることと、
メモリ・デバイスのうちの1つ又は複数に対応するアドレス範囲においてデータ読取りコ
マンドを開始することであって、開始が分析の結果に応答して行われることと、上流バス
がアイドル状態の時に、データ読取りコマンドによって戻されたデータを含む非送信請求
データ・パケットをメモリ・コントローラに伝送することと、を含む方法を含む。

【0047】

本発明は、保留中の読取り要求がアドレス範囲を指定する場合に、保留中の読取り要求
50 データを非送信請求データ・パケットと統合することをさらに含む方法を提供することが
好ましい。

【0048】

本発明は、保留中の書込み要求がアドレス範囲を指定する場合に、非送信請求データ・パケットを破棄することをさらに含む方法を提供することが好ましい。

【0049】

本発明は、非送信請求データ・パケットが限界を超えて古くなり、上流バスがアイドル状態ではない場合に、非送信請求データ・パケットを破棄することをさらに含む方法を提供することが好ましい。

【0050】

本発明は、上流バス上のアクティビティを監視することをさらに含む方法を提供することが好ましい。

【発明を実施するための最良の形態】

【0051】

本発明の実施形態は、添付の図面を参照して、単なる例として以下に詳細に説明される。

【0052】

例示的な実施形態は、各々が1つ又は複数のメモリ・デバイスに接続された1つ又は複数のメモリ・ハブ・デバイスを含む、コンピュータ・メモリ・システムを含む。メモリ・コントローラが、要求を、メモリ・コントローラとハブ・デバイスとを相互接続する1つ又は複数のチャンネルを介してメモリ・デバイスにアクセスするように方向付ける。各々のハブ・デバイスは、データをプリフェッチし、これを非送信請求「読取り」応答としてメモリ・コントローラに送信することができる。例示的な実施形態においては、プリフェッチ・データの送信は、アイドル時間中のチャンネル上で行われる。メモリ・コントローラは、受信したプリフェッチ・データを格納するために、ローカル「プリフェッチ」データ・バッファ（データ・キャッシュ）の組をコントローラ内に保持する。メモリ・アクセス要求は、まず、ローカル・プリフェッチ・バッファの内容と比較され、アドレスが有効であればローカル・プリフェッチ・バッファから直接サービスされるため、メモリ・デバイス、ハブ・デバイス、及び/又はチャンネルを使用してデータにアクセスする必要はなくなる。

【0053】

例示的な実施形態は、1つ又は複数のプロセッサと、メモリ・コントローラとメモリ・デバイスとを含むメモリ・システムに相互接続された1つ又は複数のI/Oユニット（例えばリクエスト）とを有するコンピュータ・システムを含む。コンピュータ・メモリ・システムは、データ及び命令を格納するための複数のメモリ・デバイスを有する物理メモリ・アレイを含む。これらのメモリ・デバイスは、メモリ・コントローラに直接接続してもよく、及び/又は、ハブ・デバイスを通して間接的にメモリ・コントローラに連結してもよい。例示的な実施形態においては、ハブ・ベースのコンピュータ・メモリ・システムは、メモリ・コントローラ・デバイス（例えばメモリ・コントローラ）に接続される通信ハブ・デバイスに取り付けられたメモリ・デバイスを有する。ハブ・デバイスは、マルチ・ドロップ・バス構造又はポイント・ツー・ポイント・バス構造（更に1つ又は複数の付加的なハブへのカスケード接続を含む）を通してメモリ・コントローラに接続してもよい。メモリ・アクセス要求は、メモリ・コントローラによって、バス構造を通して、選択された1つ又は複数のハブに伝送される。メモリ・アクセス要求を受信したことに応答して、ハブ・デバイスは、メモリ・アクセス要求を翻訳して、ハブ・デバイスからの書込みデータを格納するように、又はハブ・デバイスに読取りデータを与えるように、メモリ・デバイスを制御する。読取りデータは、1つ又は複数の通信パケットにエンコードされ、1つ又は複数のバスを通してメモリ・コントローラに伝送される。

【0054】

例示的な実施形態においては、プリフェッチ・データ・キャッシュは、メモリ・コントローラ内に収容され、及び/又は、1つ又は複数のリクエストとメモリ・コントローラとの間のキャッシュ階層構造から論理的にパーティション化することができる。メモリ・コ

10

20

30

40

50

ントローラは、予想される将来の要求に対して最も短い待ち時間でデータをプリフェッチ・データ・キャッシュに格納する目的で、自立的に（プロセッサ又は別のシステム要素からの指示なしにアクセスを自己起動する）、予測的に（アクセス履歴に基づいて、及びローカル・アルゴリズムに回答して、アクセスを開始する）、及び/又は、応答的に、メモリ・デバイスからデータを読み取ることができる。いずれの場合においても、読み取り要求と関連するデータを持つ、ハブ・デバイスからの暗示データ応答（implied data reply）が存在する。

【 0 0 5 5 】

例示的な実施形態は、ハブ・デバイスからの非送信請求データ・パケットをメモリ・コントローラに受信させるために、新規のネットワーク・パケットとそれに関連するプロトコルとを含む。非送信請求データ・パケットとは、メモリ・コントローラからの読み取り要求と関連しない（即ち、「送信請求」データ・パケットではない）データ・パケットを指す。非送信請求データ・パケットは、非送信請求プリフェッチ・データをメモリ・コントローラのプリフェッチ・データ・キャッシュに転送するために、ハブ・デバイスによって用いられる。

10

【 0 0 5 6 】

図5は、共用メモリ・コントローラ504を共有する2つのプロセッサ502で構成された現在のシステムを示す。共用メモリ・コントローラは、「ノースブリッジ・チップ」506に組み込まれる。図5に示される構成においては、多数の独立したマルチドロップ・メモリ・インターフェース・バス508は、データ及びエラー検出/訂正情報がパラレル・バス及び関連するデバイス全体にわたって分散又は「ストライピング（striped）」された状態で、単一の独立したアクセス要求を高い帯域幅でサポートするように、互いに論理的に集約され、一体となって動作する。メモリ・コントローラ504は、4つの狭帯域/高速ポイント・ツー・ポイント・メモリ・バス508に接続され、各々のバス508は、幾つかの固有のメモリ・コントローラ・インターフェース・チャンネルのうちの1つを、少なくとも1つのハブ・デバイスと1つ又は複数のメモリ・デバイスとを含むデジタイズチェーン・メモリ・サブシステム510（即ちメモリ・モジュール）に接続する。幾つかのシステムは、さらに、メモリ・バス508のサブセットにメモリ・サブシステム510が集中したときに動作を可能にする。この場合には、過密な1つ又は複数のメモリ・バス508は、単一のアクセス要求をサポートするように一体となって動作することができる。

20

30

【 0 0 5 7 】

図6は、1つ又は複数のプロセッサ要素と集積メモリ・コントローラ504とを収容する集積プロセッサ・チップ602で構成された現在のシステムを示す。図6に示されるシステムの動作は、基本的には図5のシステムの動作と同じであり、図6の集積プロセッサ・チップ602は、図5における2つのプロセッサ502とノースブリッジ・チップ506とを表している。

【 0 0 5 8 】

これに代わる例示的な実施形態においては、1つ又は複数のメモリ・コントローラ504は、1つ又は複数のプロセッサ・チップ及びサポート論理と共に組み込むか、（一般に「ノースブリッジ」チップと呼ばれる）個別のチップにパッケージングするか、1つ又は複数のプロセッサ及び/又はサポート論理と共にマルチチップ・キャリアに含むか、又は、用途/環境に最も合う種々の代替的な形態にパッケージングすることができる。これらの解決策のいずれも、1つ又は複数のハブ・チップ及び/又はメモリ・デバイスとの接続に、1つ又は複数の狭帯域/高速リンクを用いることも用いないこともある。

40

【 0 0 5 9 】

例示的な実施形態は、図7に示されるようにカスケード型相互接続バス706によってプロセッサ複合体に接続される、一般にハブ論理チップを有するメモリ・モジュールすなわちバッファ・デバイスの形態をとる1つ又は複数のメモリ・サブシステムで構築されたメモリ・システムに適用する。これに代わる例示的な実施形態においては、メモリ・モジ

50

ジュール710(メモリ・デバイス716と、例えばハブ論理チップ708内における関連するメモリ・インターフェース回路とが含まれていることに起因して、メモリ・サブシステムとも呼ばれる)は、1つ又は複数のバス706によってプロセッサ複合体702に接続される。図7に示されるメモリ・モジュール710は、リモート・プリフェッチ論理を含むハブ論理チップ708と、(例示的な実施形態においては上流バス及び下流バスを含むものとして示される)カスケード型相互接続メモリ・バス706を介した(1つ又は複数のプロセッサを含む)プロセッサ複合体702のメモリ・コントローラ704への第1のインターフェースと、の両方を含む。メモリ・モジュール710は、また、第2のインターフェースを介してハブ論理チップ708に接続された1つ又は複数のメモリ・デバイス716を含む。ハブ論理チップ708には、特にカスケード型相互接続バス706の場合に、アドレス、コマンド、及びデータ情報などの情報を、第1のメモリ・モジュール710の下流に位置する1つ又は複数のメモリ・モジュール710に再駆動すると共に、上流メモリ・モジュール710又はメイン・メモリ・コントローラ704に宛てられたデータ及び/又はエラー、状態、及び他の動作情報を受信し再駆動するための、第3のインターフェースが存在する。

10

【0060】

メモリ・モジュール710は、DIMM、シングル・インライン・メモリ・モジュール(SIMM)、及び/又は他のメモリ・モジュール若しくはカード構造を含む種々の技術において実装することができる。一般に、DIMMとは、片面又は両面にランダム・アクセス・メモリ(RAM)集積回路又はダイを主に備え、両面に信号ピン及び/又は電源ピンを備える、小型の回路ボードを指す。これは、片面又は両面にRAM集積回路又はダイを主に備え、1つの長辺に沿って単一のピンの列を備えた小型回路ボード又は基板であるSIMMと対比される。図1に図示されるDIMMは、例示的な実施形態においては168ピンを含むが、それに続くDIMMは、100ピンから300ピン以上の範囲のピン数を持つように構成されてきた。

20

【0061】

例示的な実施形態においては、メモリ・バス706は、メモリ・モジュール710へのマルチ・ドロップ接続を用いて、及び/又は、ポイント・ツー・ポイント接続を用いて、構成される。下流バスと呼ばれる、コントローラ・インターフェース(又はメモリ・バス706)の下流部分は、メモリ・モジュール710に送信されるコマンドと、アドレスと、データと、他の動作情報、初期化情報、又は状態情報とを含むことができる。各々のメモリ・モジュール710は、バイパス回路を介して情報を後続の1つ又は複数のモジュール710に単に転送するか、情報が下流メモリ・モジュール710をターゲットとするものであると判断された場合にはその情報を受信し、翻訳し、再駆動するか、最初に情報を翻訳して意図される宛先を判断することなく情報の幾つか若しくは全てを再駆動するか、又は、これらの選択肢のサブセット若しくは組み合わせを実施することができる。

30

【0062】

上流バスと呼ばれるメモリ・バス706の上流部分は、要求された読取りデータ及び/又はエラー、状態、又は他の動作情報を戻し、この情報は、バイパス回路を介して後続の1つ又は複数のメモリ・モジュール710に転送するか、情報が上流メモリ・モジュール710及び/又はプロセッサ複合体702の中のメモリ・コントローラ704をターゲットとするものであると判断された場合にはこれを受信し、翻訳し、再駆動するか、最初に情報を翻訳して意図される宛先を判断することなく情報の一部若しくは全体を再駆動するか、又は、これらの選択肢のサブセット若しくは組み合わせを実施することができる。

40

【0063】

例示的な実施形態は、カスケード型相互接続又はマルチ・ドロップ構造において接続される付加的なメモリ・サブシステム位置の有無にかかわらず、メイン・メモリ・コントローラ704に(1つ又は複数のポイント・ツー・ポイント相互接続などを介して)直接接続される単一のメモリ・モジュール710、即ち単一のメモリ・サブシステムのみを有する単一レベルのメモリ・システムによって利用することができる。さらなる例示的な実施

50

形態においては、メイン・メモリ・コントローラ704は、2つ以上のメモリ・チャネルを含み（例えば図7のメモリ・バス712を参照されたい）、各々のメモリ・チャネルは、同様に動作する1つ又は複数の上流バス706及び下流バス712を備えることが多く、メモリ・コントローラ704は、各々のメモリ・バス706及び712について並列に又は独立に動作し、及び/又は、各々のメモリ・バス706及び712は、同一の又は異なる数のインストールされたメモリ・モジュール710を含む。

【0064】

例示的な実施形態においては、メモリ・システムは、2つ以上のメモリ・モジュール710を含み、後続のメモリ・モジュール710は、第1のメモリ・モジュール710を経由して、及び/又は、第1のメモリ・モジュール710とは独立に、メモリ・コントローラ704と通信する。図7に示されるような例示的な実施形態においては、メモリ・コントローラ704は、ポイント・ツー・ポイント・バスを介して第1のメモリ・モジュール710と通信し、バスは、第1のメモリ・モジュール710の中の再駆動回路を介して（例えばハブ論理チップ708の中の回路を介して）、第2の（後続の）メモリ・モジュール710にカスケード接続される。第3及び第4（及びそれ以上の）メモリ・モジュール710は、同様の方法によって接続することができ、メモリ・モジュール710に含まれる再駆動回路は、メモリ・バス706においてそのメモリ・モジュール710より前に位置する。

【0065】

これに代わる例示的な実施形態においては、ポイント・ツー・ポイント・バスはスイッチ又はバイパス機構を含み、これによって、下流通信（メイン・メモリ・コントローラ704からメモリ・モジュール710へ渡される通信）中に2つ又はそれ以上の可能なメモリ・モジュール710のうちの1つにバス情報が方向付けされ、多くの場合は1つ又は複数の上流メモリ・モジュール710を経由して上流情報（メモリ・モジュール710からメモリ・コントローラ704への通信）が方向付けされることになる。更なる実施形態は、カスケード型相互接続メモリ・システムにおいて、例えばメモリ・コントローラと第1の過密なメモリ・モジュールとの間に配置することができる、当該技術分野において知られているコンティニュイティ・モジュールの使用を含み、それにより、メモリ・コントローラと第1の過密なメモリ・モジュールとの間の中間モジュール位置のいずれかが、1つ又は複数の中間モジュール位置がメモリ・モジュールを含まない場合であっても、メモリ・コントローラと第1の過密なメモリ・モジュール位置との間を通過する情報を受信できる手段を含む。1つ又は複数のコンティニュイティ・モジュールは、いずれかのバス制約に従って、（メイン・メモリ・コントローラに最も近い）最初の位置、（含まれる終端のいずれかの前の）最後の位置、又は、1つ又は複数の中間位置のいずれかを含む、1つ又は複数のモジュール位置のいずれかにインストールすることができる。コンティニュイティ・モジュールの使用は、中間のメモリ・モジュールがコンティニュイティ・モジュールによって除去又は置換されるマルチ・モジュール・カスケード型相互接続バス構造において特に有益であり、それにより、システムは、中間のメモリ・モジュールが除去された後でも動作し続ける。より一般的な実施形態においては、1つ又は複数のコンティニュイティ・モジュールは、全ての必要な信号を1つ又は複数の入力から対応する1つ又は複数の出力に転送する相互接続線を含むか、又は、リピータ・デバイスを通して再駆動される。1つ又は複数のコンティニュイティ・モジュールは、更に、（EEPROMのような）不揮発性ストレージ・デバイスを含むこともあるが、メイン・メモリ・ストレージ・デバイスを含むものではない。

【0066】

例示的な実施形態においては、メモリ・システムは、カスケード型相互接続メモリ・バス706を介してメモリ・コントローラ704に接続された1つ又は複数のメモリ・モジュール710を含むが、ポイント・ツー・ポイント・バス、マルチ・ドロップ・メモリ・バス、又は共用バスといった他のメモリ構造を実装することもできる。用いられる信号送信方法、目標動作周波数、空間、電力、コスト、及び他の制約に応じて、種々の代替的な

10

20

30

40

50

バス構造を考慮することができる。ポイント・ツー・ポイント・バスは、分岐信号ライン、スイッチ・デバイス、又はスタブを有するバス構造に比べて、信号劣化が生じる可能性低減されるので、電氣的な相互接続を用いて製造されるシステムにおいて最適な性能をもたらすことができる。しかしながら、多数のデバイス又はサブシステムとの通信を必要とするシステムにおいて用いられるときには、この方法は、追加コンポーネントのコストの著しい増加とシステム電力の増大とをもたらすことが多く、中間のバッファリング及び/又は再駆動が必要となるため可能なメモリ密度を減少させることもある。

【 0 0 6 7 】

図7には示されないが、メモリ・モジュール710は、(一般に電源投入後の)メモリ・モジュール属性の判定、システムへの障害又は状態情報の報告、電源投入後又は通常の動作中の1つ又は複数のメモリ・サブシステム710の設定、又は他の目的を含む1つ又は複数の目的のために用いられる、「存在検知」バス、I2Cバス、及び/又はSMBusのような別個のバスを含むこともできる。バス特性に応じて、このバスは、有効な動作の完了を1つ又は複数のメモリ・モジュール710によって1つ又は複数のメモリ・コントローラ704に報告できる手段、又は、メイン・メモリ・コントローラ要求の実行中に生じた故障の特定を提供することもできる。

【 0 0 6 8 】

ポイント・ツー・ポイント・バス構造から得られる性能と同様の性能は、スイッチ・デバイスを追加することによって得ることができる。これらの解決策及び他の解決策によって、ポイント・ツー・ポイント・バスの特徴の多くを保持しつつ、より低い電力においてメモリ・パッケージング密度が増大するようになる。マルチ・ドロップ・バスは、より低い動作周波数に限定されることが多いものの、多くの用途にとって利点となりうるコスト/性能の点で、代替的な解決策を提供する。光学バスによる解決策は、ポイント・ツー・ポイント用途又はマルチ・ドロップ用途のいずれかにおいて周波数及び帯域幅の可能性を著しく増大させることができるが、コスト的な影響及び空間的な影響を受けることがある。

【 0 0 6 9 】

ここで用いられる「バッファ」又は「バッファ・デバイス」という用語は、(コンピュータにみられるような)一時的なストレージ・ユニットを指し、特に、1つの速度で情報を受け入れ、それを別のところに送るものを指す。例示的な実施形態においては、バッファは、(例えば、電圧レベル又は電流容量を変更して)2つの信号の間に互換性を与える電子装置である。「ハブ」という用語は、「バッファ」という用語と交換可能に用いられることがある。ハブは、他の幾つかの装置に接続される多数のポートを収容する装置である。ポートは、適合するI/O機能を与えるインターフェースの一部である(例えば、ポートは、ポイント・ツー・ポイント・リンク又はバスの1つの上でデータ、アドレス、及び制御情報を送受信するために使用することができる)。ハブは、幾つかのシステム、サブシステム、又はネットワークを互いに接続する中心的な装置となる場合がある。受動的ハブは、メッセージを転送することができるのみであるのに対し、能動的ハブ、即ちリピータは、それがなければ距離と共に劣化するデータのストリームを増幅し、リフレッシュする。ここで用いられるハブ・デバイス708という用語は、メモリ機能を実施するための論理(ハードウェア及び/又はソフトウェア)を含むハブ・チップを指す。

【 0 0 7 0 】

また、ここで用いられる「バス」という用語は、コンピュータにおいて2つ又はそれ以上の機能的なユニットを接続する導電体(例えば、線、及び、プリント基盤の配線又は集積回路の接続部)の組を指す。データ・バス、アドレス・バス、及び制御信号は、他のものがなければ各々が役に立たないことが多いため、これらの名称にもかかわらず単一のバスを構成する。バスは、2つ又はそれ以上のトランシーバ、送信機、及び/又は受信機を電氣的に接続する主伝送経路を形成する、各々が2つ又はそれ以上の接続点を有する複数の信号ラインを含むことができる。「バス」という用語は、メモリ・システムにおけるメモリ・コントローラとの関係で「ポート」の機能を説明するのに用いられることが多い、

10

20

30

40

50

1つ又は複数のバス又はバスの組を含むことがある、「チャンネル」という用語と対比される。ここで用いられる「チャンネル」という用語は、メモリ・コントローラ上のポートを指す。この用語は、I/O機器又は他の周辺機器と共に用いられることが多いが、チャンネルという用語は、プロセッサ又はメモリ・コントローラと1つ又は複数のメモリ・サブシステムのうちの1つとの間のインターフェースを説明するのに援用される場合もあることに留意されたい。

【0071】

更に、ここで用いられる「デジター・チェーン」という用語は、例えば、デバイスAがデバイスBに配線され、デバイスBがデバイスCに配線される等といったバス配線構造を指す。最後のデバイスは、典型的には、レジスタ又はターミネータに配線される。全てのデバイスが同一の信号を受信してもよく、又は、単純バスとは対照的に、各々のデバイスが1つ又は複数の信号が通過する前にこれを修正してもよい。ここで用いられる「カスケード」又は「カスケード型相互接続」とは、連続したステージ若しくはユニット、又は、典型的には、論理リピータとして動作し、更に、集約されるべきデータを既存のデータ・ストリームに統合することを可能にするハブである、相互接続されたネットワーク・デバイスの集合体を指す。また、ここで用いられる「ポイント・ツー・ポイント」バス及び/又はリンクという用語は、各々が1つ又は複数のターミネータを含むことがある1つ又は複数の信号ラインを指す。ポイント・ツー・ポイント・バス及び/又はリンクにおいては、各々の信号ラインは、各々が送信回路、受信回路、又はトランシーバ回路に連結される2つのトランシーバ接続点を有する。信号ラインは、一般に、少なくとも1つの論理信号を転送するのに用いられる、単一のキャリアとして構成されるか、又は、ねじれ構造、並列構造、若しくは同心円構造の2つ又はそれ以上のキャリアとして構成された、1つ又は複数の導電体又は光キャリアを指す。

【0072】

メモリ・デバイスは、一般に、DRAM(ダイナミック・ランダム・アクセス・メモリ)、SRAM(スタティック・ランダム・アクセス・メモリ)、FeRAM(強誘電体RAM)、MRAM(磁気ランダム・アクセス・メモリ)、フラッシュメモリ、並びに、電気的手段、光学的手段、磁気的手段、生物学的手段、及び他の手段の形態で情報を格納する他の形態のランダム・アクセス・メモリ及び関連するメモリなどといったメモリ(ストレージ)・セルで主に構成される集積回路として定義される。ダイナミック・メモリ・デバイスのタイプは、FPM DRAM(高速ページモード・ダイナミック・ランダム・アクセス・メモリ)、EDO(Extended Data Out) DRAM、BEDO(バーストEDO) DRAM、SDR(シングル・データ・レート)同期DRAM、DDR(ダブル・データ・レート)同期DRAM、又は、DDR2、DDR3、DDR4などの期待される後続デバイスのいずれか、及び、関連するDRAMにおいて見出される基本的な機能、特徴、及び/又はインターフェースに基づくことが多いグラフィックRAM、ビデオRAM、LP RAM(低パワーDRAM)などの関連技術といった、非同期メモリ・デバイスを含むことがある。

【0073】

メモリ・デバイスは、チップ(ダイ)の形態、及び/又は、種々のタイプ及び構成のシングルチップ・パッケージ若しくはマルチチップ・パッケージの形態で、利用することができる。マルチチップ・パッケージにおいては、メモリ・デバイスは、他のメモリ・デバイス、論理チップ、アナログ・デバイス、及びプログラム可能デバイスのような他のデバイス・タイプと共にパッケージングすることができ、レジスタ、キャパシタ、及びインダクタのような受動デバイスを含むこともできる。これらのパッケージは、組み込まれたヒートシンク又は他の冷却強化手段を含むことができ、それらは更に、直近のキャリア又は別の近くのキャリア若しくは熱除去システムに取り付けることができる。

【0074】

モジュール・サポート・デバイス(例えば、バッファ、ハブ、ハブ論理チップ、レジスタ、PLL、DLL、不揮発性メモリなど)は、技術、電力、空間、コスト、及び他のト

10

20

30

40

50

レードオフに基づいて、多数の個別のチップ及び/又はコンポーネントで構成するか、多数の個別のチップとして1つ又は複数の基板上に組み合わせるか、単一のパッケージ上に組み合わせるか、又は、単一デバイス上に集積させることもできる。さらに、レジスタ、キャパシタといった種々の受動デバイスの1つ又は複数は、技術、電力、空間、コスト、及び他のトレードオフに基づいて、サポート・チップ・パッケージに集積させるか、又は、基板、ボード若しくは原材料のカード自体に集積させることができる。これらのパッケージは、組み込まれたヒートシンク又は他の冷却強化手段を含むことができ、それらは更に、直近のキャリア又は別の近くのキャリア若しくは熱除去システムに取り付けることができる。

【0075】

図8は、例示的な実施形態によって実装することができるメモリ・システムのブロック図である。図8に示されるシステムは、図7に示されるシステムと同様のものであり、図8は、例示的な実施形態を実装するのに利用することができるハブ・デバイス708及びメモリ・コントローラ704の中の要素を示している。メモリ・コントローラ704は、ローカル・プリフェッチ・データ・キャッシュ810を含むが、このキャッシュもまた、リクエスタ（例えば、1つ又は複数のプロセッサ）とメモリ・コントローラとの間の既存のキャッシュ階層構造から論理パーティション化することができる。例示的な実施形態においては、ローカル・プリフェッチ・データ・キャッシュ810は、コンピュータ環境によって確立されたポリシーに従って管理される。更なる実施形態においては、ローカル・プリフェッチ・データ・キャッシュは、未処理の読取り要求及び書込み要求に関するバッファ置換ポリシーと整合性依存度とに従って、メモリ・コントローラ704により管理される。

【0076】

ハブ・デバイス708は、インターフェース論理808と、プリフェッチ論理804と、DRAMインターフェース802とを含む。プリフェッチ論理804は、メモリ・コントローラ704からメモリ・デバイスへのアクセスを監視するのに利用される。メモリ・コントローラ704によるアクセス・パターンに基づいて、プリフェッチ論理804は、ハブ・デバイス708においてプリフェッチが実施されるべきかどうかを判定する。1つの例示的な実施形態においては、プリフェッチ要求は、システム・アクセス・パターン履歴に基づく次の連続的なアドレス、即ちストライド・アドレスを含むが、システム設計及び/又はプログラミングに基づいて、他のプリフェッチ・アルゴリズムを適用することもできる。ここで用いられるストライド・アドレスという用語は、次のアドレスが、アドレス間隔の履歴と矛盾しないが必ずしも隣接するアドレスではない先行アドレスからある距離だけ離れていることを意味する。ハブ・デバイス708上のプリフェッチ論理804によって、適切なプリフェッチ・アルゴリズムのいずれかを実装することができる。例示的な実施形態においては、プリフェッチ論理804は、データがプリフェッチされるべきであると判定した場合に、DRAMインターフェース802にデータを読取ることを命令する。これに代わる例示的な実施形態においては、プリフェッチ論理804は、インターフェース論理808にデータを読取ることを命令し、インターフェース論理は、DRAMインターフェース802からそのデータを要求する。

【0077】

図9を参照して以下に説明されるように、インターフェース論理808は、非送信請求プリフェッチ・データをメモリ・コントローラ704に伝送するのに使用される論理に加えて、標準的なインターフェース論理機能を含む。これに代わる例示的な実施形態においては、プリフェッチ論理は、図9を参照して以下に説明されるように非送信請求プリフェッチ・データをメモリ・コントローラ704に伝送するのに使用される命令を含む。これに代わる例示的な実施形態においては、メモリ・デバイス716及び/又はハブ・デバイス708は、（ここでは、非送信請求データ読取りにより戻されたデータとも呼ばれる）プリフェッチ・データをメモリ・コントローラ704に伝送する前にローカルに格納するためのプリフェッチ・データ・キャッシュを含む。

10

20

30

40

50

【 0 0 7 8 】

図 9 は、例示的な実施形態によって実装される非送信請求データのプリフェッチに続いてハブ・デバイス 7 0 8 において実施される動作についての処理フローである。処理フローは、ハブ・デバイス 7 0 8 の中のインターフェース論理 8 0 8 に付加的な機能を追加することによって、又は、ハブ・デバイス 7 0 8 に別個のプリフェッチ論理 8 0 4 を追加することによって、実装することができる。ここで用いられる「論理」という用語は、ハードウェア及び/又はソフトウェアによって実装されるコンピュータ命令を指す。ブロック 9 0 2 において、ハブ・デバイス 7 0 8 (例えばプリフェッチ論理 8 0 4) は、メモリ・コントローラ 7 0 4 によるメモリ・デバイス 7 1 6 へのアクセスを監視する。ブロック 9 0 2 において、ハブ・デバイス 7 0 8 (例えばプリフェッチ論理 8 0 4) の中のアルゴリズムに基づき、メモリ・コントローラ 7 0 4 による先行アクセスにตอบสนองして、ハブ・デバイス 7 0 8 によってデータをプリフェッチすることができる。ここでも、プリフェッチ・アルゴリズムは当該技術分野において周知であり、例示的な実施形態によっていずれかの適切なアルゴリズムを利用することができる。1つ又は複数のメモリ・デバイス 7 1 6 に対する読取り動作が完了すると、メモリ・コントローラ 7 0 4 に伝送するために1つ又は複数のパケットが準備される。ブロック 9 0 4 において、アクセスが完了したかどうか判定され、メモリ・コントローラ 7 0 4 への伝送のためにパケットが準備される。例示的な実施形態においては、メモリ・アクセスによってキャッシュ・ライン転送(例えば 1 2 8 バイトのうちの 6 4 バイト)が行われ、メモリ・コントローラへの1つ又は複数のパケットの伝送は、キャッシュ・ラインのサイズに等しいデータが1つ又は複数のメモリ・デバイス 7 1 6 からプリフェッチされるまで開始されないが、用途によっては他の転送及びパケット・サイズを使用してもよく、転送は1つ又は複数のパケットで構成することができる。ブロック 9 0 2 が実施され、ハブは、少なくとも1つ又は複数の対象となるパケットの最初のものに関連する最小の情報が利用可能になるまで、データをプリフェッチする。システム設計の目的に基づいて種々の転送又はパケット・サイズを用いることができ、図 1 1 は、1つの例示的な実施形態に関する。転送に付随するタグは、その転送のタイプを特定するものであり、したがって、転送と関連するビット/バイト/パケットの数を特定する。

【 0 0 7 9 】

ブロック 9 0 4 で判定されたときに、メモリ・コントローラ 7 0 4 に伝送するためにパケットが準備された場合には、ブロック 9 0 6 が実施される。ブロック 9 0 6 において、上流バスにおけるアクティビティがハブ・デバイス 7 0 8 によって(例えばインターフェース論理 8 0 8 によって)監視される。ブロック 9 0 8 において、非送信請求データ・パケットの伝送を可能にするために上流バスが利用できるかどうか判定される。一般に、上流バスに非アクティブな期間が存在し、ハブ・デバイス 7 0 8 から上流に送られる優先順位の高いパケットが他に存在しない場合には、上流バスは利用可能である。上流バスが利用可能である場合には、ブロック 9 1 6 が実施され、ハブ・デバイス 7 0 8 は、データ・パケットをそのデータを特定するタグ(例えば、アドレス範囲及びモジュール識別情報)と共にメモリ・コントローラ 7 0 4 に伝送する。このようにして、非送信請求データは、典型的には低い優先順位で、メモリ・コントローラ 7 0 4 に送信される。次いで処理はブロック 9 0 2 で続行する。従来のカスケード型相互接続メモリ・システムにおいては、インターフェース論理 8 0 8 が、ローカル情報のカスケード型相互接続バスへの統合を担当するが、メモリ・コントローラは、メモリ・コントローラに情報を伝送する準備ができたときに、転送中にはバスがアイドル状態となるように、アクセスをスケジューリングする。例示的な実施形態においては、インターフェース論理 8 0 8 は、用途設定及び転送の優先順位に基づいて、ある時点でコントローラへの上流転送を開始し、下流モジュールからの途中の転送を一時的に遅らせるか、そうでなければ中断することができる。

【 0 0 8 0 】

ブロック 9 0 8 で判定されたときに、上流バスが利用可能でない場合には、ブロック 9 1 0 が実施され、ハブ・デバイス 7 0 8 は、下流バスにあるハブ・デバイス 7 0 8 に向け

10

20

30

40

50

られるアクセスを監視して、いずれかのアクセスが、現時点で利用可能な、上流バスにあるメモリ・コントローラ704に送信されるのを待っている1つ又は複数の非送信請求データ・パケットと同じアドレス範囲に向けられているかどうかを判定する。ブロック912において、利用可能な非送信請求パケットのアドレス範囲と一致する読取り動作が受信されたと判定された場合には、ブロック914が実施される。ブロック914において、非送信請求パケットと関連するタグは、そのパケットが応答パケットであることを示すように修正され、新たなパケット識別情報に対応して優先順位が上げられる。次いで、そのハブ・デバイス708からの現在の転送のすべてが完了すると、パケットを上流データ・バスに統合するブロック916が実施されるが、その統合は、読取り優先順位アルゴリズムとも矛盾しないようになっている。

10

【0081】

ブロック912において、利用可能な非送信請求パケットのアドレス範囲と一致する書込み動作が受信されたと判定された場合には、ブロック922が実施される。ブロック922において、プリフェッチ・データがもはや有効ではないためパケットは破棄され、処理は、ハブ・デバイス708が最適なプリフェッチのためにアクセスを監視し続けるブロック902において続行する。1つ又は複数のメモリ・デバイス716からの非送信請求プリフェッチが完了すると、タイムアウト・カウンタは、1つ又は複数のパケットを保持するための所定のタイムリミット（例えば、20クロック周期、30クロック周期など）に設定され、ゼロまでのカウントダウンを開始する。ブロック920で判定されたときに、1つ又は複数のパケットが所定のタイムリミットまで時間経過すると（又は所定の寿命に達すると）、タイムアウト・カウンタはフラグを立て、ブロック922において、そのタイムアウト期間と関連する、その時点でメモリ・コントローラ704へ送信される過程にはない1つ又は複数のパケットは、破棄される。所定のタイムリミットは、ハブ・デバイス708の間で異なってもよいプログラム可能な（例えば修正可能な）数である。次いで、処理はブロック902で続行する。タイムアウト期間に達していない場合には、処理は、パケットを伝送するために上流バスが利用できるかどうかを判定するブロック908で続行する。ブロック908、910、912、918、及び920のステップは、例示的な処理フローの形で示されるが、基本的な教示から逸脱することなく、（読取り、書込み、及びタイムアウトのイベントが並行して監視される）並列フロー、又は、（読取りブロック912の完了の次に書込みブロックを完了し、それとタイムアウトを監視するブロック920を並列にするなどといった）直列処理フローと並列処理フローとの組み合わせを含む他の処理フローを使用することもできる。

20

30

【0082】

図10は、メモリ・システムにある1つ又は複数のハブ・デバイス708から1つ又は複数の非送信請求読取りパケットを受信したことに続いて、メモリ・コントローラ704において実施される動作についての例示的な処理フローである。ブロック1002において、メモリ・コントローラは、パケットを翻訳して、そのパケットが、メモリ・システムにある1つ又は複数のメモリ・モジュール710、即ちメモリ・サブシステムからの送信請求、状態情報、非送信請求、又は他の通信形態（例えば1つ又は複数の情報パケット）としてパケットを特定するヘッダ情報を含むかどうかを判定することによって、メモリ・モジュール710上のハブ・デバイス708からの全てのバスのアクティビティを監視する。ブロック1004において、有効なパケットが受信されたかが判定される。例示的な実施形態においては、メモリ・コントローラ704は、バス情報（例えばヘッダ・フィールド）の1つ又は複数の第1の伝送（又は、いくつかのビット数）を調べて、その情報が1つ又は複数の有効なパケットを含むように見えるかどうかを判定する。ブロック1004において、メモリ・コントローラ704がバス706上の情報を1つ又は複数の有効なデータ・パケットを含むものとして特定しなかった場合には、ブロック1002が実施され、メモリ・コントローラ704は、上流バス706上のアクティビティを監視し続ける。

40

【0083】

50

ブロック1004において、メモリ・コントローラ704が、バス706上の伝送を何らかの形態の1つ又は複数のデータ・パケットであると特定した場合には、その1つ又は複数のパケットが、コントローラによって開始されるアクセス又は他の動作に対する応答に関連するかどうかを判定するブロック1006が実施される。ブロック1006において、1つ又は複数のパケットに付随するタグ情報が送信請求読取り要求への応答としてデコードされた場合には、ブロック1010が実施され、メモリ・コントローラ704は、残りのタグ情報（例えば、アドレス範囲及び/又はモジュール識別子）を保留中の読取り要求と比較し、データの有効性を検証する。ブロック1010で判定されたときに、タグが未処理の読取り要求と関連している場合には、ブロック1012が実施され、読取り要求に割り当てられた優先順位に基づいて次の利用可能な時点でデータがリクエストに（例えば、1つ又は複数のプロセッサに）転送され、保留中の読取り要求は廃棄される。次いで、処理はブロック1002で続行する。ブロック1010で判定されたときに、タグが未処理の読取り要求と関連していない場合には、ブロック1014が実施され、1つ又は複数のパケットが破棄され、プロセッサ又は代替的なエラー復旧システムによってサービスするためにエラー・フラグが生成される。1つ又は複数のパケットが、有効な応答パケットであるように見えるが未処理の読取り要求に関連していない（例えば、アドレス又はモジュール識別子が未処理の読取り要求と一致しない）場合には、ブロック1014が実施される。この場合には、メモリ・コントローラ704は、パケットを破棄し、次いで、エラー復旧（例えばチャンネル診断）を実施することによって、なぜエラーが生じたか解明することを試みる。診断と並行して、又は、（例えば、後の時点まで診断を延期することが決定されたことによる）1つ又は複数の無効なパケットの破棄と同時に、エラー状態が解消されると、次いで処理はブロック1002で続行する。

【0084】

バス706上の情報が、ブロック1004において1つ又は複数の有効なパケットとして特定されたが、ブロック1006で判定されたときにそのアドレス範囲に対する保留中の読取り要求又は書込み要求と関連しない場合には、ブロック1008が実施され、1つ又は複数のパケットは、その1つ又は複数のパケットが1つ又は複数の非送信請求データ・パケットであるかどうかを判定するために評価される。ブロック1008において、1つ又は複数のパケットが非送信請求読取りパケットであると判定された場合には、ブロック1022が実施され、メモリ・コントローラ704は、ダブルチェックを行って再度タグ情報を比較し、要求が最近の読取り要求と関連するかどうかを判定する。ブロック1022において、1つ又は複数のパケットが読取り要求と関連することが分かった場合には、ブロック1028が実施され、送信請求読取り要求は、機能的にはブロック1012における処理と同様の方法でリクエストに伝送される。次いで、処理はブロック1002で続行する。

【0085】

ブロック1022において、1つ又は複数のパケットが読取り要求と関連しないと判定された場合には、ブロック1024が実施される。ブロック1024において、1つ又は複数のパケットがそのアドレス範囲に対する最近の書込み動作に関係するかどうかを判定され、そのデータがそのアドレス範囲に対する以前の書込みより古いことがタグにより確認された場合には、ブロック1030が実施され、1つ又は複数のパケットは期限切れであるため破棄される。このデータを用いてローカル・キャッシュを更新すれば、期限切れ（又は不正確な）データがローカル・キャッシュに格納されることとなり、従ってデータ整合性の欠如を招くことになる。次いで、処理はブロック1002で続行する。そのアドレス範囲について未処理の読取り要求又は書込み要求が特定されなかった場合には、ブロック1026が実施され、1つ又は複数のパケットは、ローカル・プリフェッチ・データ・キャッシュ制御アルゴリズムに従って、アドレス情報と共にローカル・プリフェッチ・データ・キャッシュ810に追加される。次いで処理はブロック1002で続行する。

【0086】

ブロック1008において、1つ又は複数のパケットが1つ又は複数の非送信請求読取

10

20

30

40

50

りパケットではないと判定された場合には、受信された1つ又は複数のパケットは、1つ又は複数の送信請求読取りパケットでも非送信請求読取りパケットでもなく、その1つ又は複数のパケットは、1つ又は複数のハブ・デバイス708からメモリ・コントローラ704に向かう他の有効な1つ又は複数の通信で構成されている場合がある。ブロック1016において、1つ又は複数のパケットは、その1つ又は複数のデータ・パケットの中の情報が状態、故障復旧、又は他のアクションと関連するかどうかを判定するために評価され、関連する場合にはブロック1020が実施され、メモリ・コントローラ704は、システム設計又はプログラミングに従ってアクションを行う。代替的に、1つ又は複数のパケットは1つ又は複数の有効なパケットであるが、その内容が不確定か、そうでなければ無効であると判定された場合には、ブロック1018が実施される。この場合には、メモリ・コントローラ704は、パケットを破棄し、次いで、エラー復旧（例えばチャンネル診断）を実施することによって、なぜエラーが生じたかを説明することを試みる。診断と並行して、又は、（例えば、後の時点まで診断を延期することが決定されたことによる）1つ又は複数の無効なパケットの破棄と同時に、エラー状態が解消されると、次いで処理はブロック1002で続行する。

【0087】

現在の1つ又は複数のパケットに対する応答の実行と並行して、又は、現在のパケットに応答して動作を開始させた直後に、受信した伝送に対する応答を行うと、メモリ・コントローラ704は、ブロック1002において、入ってくる送信請求パケット又は非送信請求パケットについてバス706の監視を続ける。図9及び図10に示される処理は、論理的な処理フローを表す。処理の実際の物理的実装は、その物理的実装が図9および図10に示されるものと同じ結果を達成する限り、異なるもの（例えば幾つかの処理ブロックが並列に実施される）とすることもできる。

【0088】

図11は、メモリ・モジュール710の1つ又は複数によってカスケード型相互接続メモリ・バス706を介してメモリ・コントローラ704に送信することができる例示的な上流データ・パケットを示す。例示的な実施形態においては、ハブ・デバイス708からメモリ・コントローラ704への1つ又は複数の非送信請求データ転送は、データ・パケットとしてフォーマットされる。図11において示される例示的なデータ・パケットは、19個のビット・レーンと、1つのヘッダと、8つのデータ転送とを含み、合計で171個のビット位置となる。他の例示的なパケットは、より多いか又はより少ないビット・レーンと、より多いか又はより少ない転送とを有することがあり、用途の要件及び/又は条件に基づいて、1つ又は複数のスペア・ビット・レーンを含むこともある。パケット・ヘッダは、チャンネル・レーン番号の下に示されており、タグの位置（ビット0-5）、コマンド（CMD）ビット（6-7）、状態（STS）ビット（8-9）、ECCビット（12-17）、並びに、割り込み、スペア・ビット・レーン、CRC、データ、及び所与の用途に適用できる他の情報といった機能について用いることができる他のレーンを定める。例示的な実施形態においては、タグは、図10を参照して上述されたように、応答パケットを特定し、非送信請求プリフェッチ・データ・パケットを特定し、メモリ・コントローラ704において受信される他のタイプのパケットを特定するために、メモリ・コントローラ704によって利用される。それに続く（転送0-7として特定される）転送は、ヘッダに続くデータ転送を指し、例示的な実施形態においては、0から17までとして特定される18個のビット・レーンと、0から7までとして特定される8つの転送とを含み、（例えば、16バイトのデータに、関連するEDCチェック・ビットを加えて）合計で144個のビット位置を含む。スペア・ビット・レーンのような目的のために用いることができる付加的なビット・レーンが示される（図11の列18として示される）。

【0089】

定められたビット位置の各々の位置、カウント、及び機能は、説明されたものとは変えることもできる。1つの場合においては、上流データ転送パケットは、1つ又は幾つかの転送を含むヘッダ及び状態情報のみで構成される場合がある。このパケットは、書込みコ

10

20

30

40

50

マンド又はリフレッシュのような動作の完了の報告、エラー状態の報告、(モジュール又はデバイスの温度などの)状態状況の報告、又は、限定された数のビット位置/転送を必要とする他の情報の提供に用いることができる。別の場合においては、パケットに9つの完全な転送を含めることができ、これは、ECCビット及び/又はCRCビット、状態ビット、タグ・ビット、及び他の情報のような情報と共に、メモリ・コントローラ704によって開始された読取り要求に関連するデータや、(メモリ・サブシステム、即ちメモリ・モジュール710によって開始される)自己起動した読取り動作からのデータを、メモリ・コントローラ704に与えることができる。

【0090】

多くの場合においては、要求された/必要な情報をプロセッサ複合体702の中のメモリ・コントローラ704及び/又は1つ又は複数のプロセッサに与えるために、1つ又は複数のパケット転送が必要とされ、通常は、4つ又は8つのパケットが用いられる。1つ又は複数のパケットは、2つ以上の要求された動作に関連する情報を含むこともできる。例えば、第2のメモリ・モジュール710に対して要求された読取りに関連する一連のパケットは、上流バスと、第2のメモリ・モジュール710の上流の、メモリ・コントローラ704より前に位置する第1のメモリ・モジュール710とを經由して、メモリ・コントローラ704に送信される場合がある。1つ又は複数のパケットのうちの最初のパケットを受信すると、第1のメモリ・モジュール710は、結果として得られるパケットが1つ又は複数の別個のメモリ・モジュール710からの2つ又はそれ以上の独立した動作と関連する情報を含むように、独立した別個のタグ及び/又は状態情報を、上流に移動する1つ又は複数のパケットの、一般にはこの目的のために予め定められた位置に、連結する。このようにして、1つ又は複数のパケットの中の可能なビット位置の全てが使用されることを保証することによって、利用可能なバス帯域幅の使用を最大化することが可能である。2つの独立した動作の例として、第2のメモリ・モジュール710からの読取り動作の結果に、書込み動作完了状態、リフレッシュ完了状態、モード変更状態、温度などの環境情報、又は他の情報といった第1のメモリ・モジュール710からの状態情報が加えられることが挙げられる。

【0091】

図11には図示されないが、例示的な下流パケットは同様に構成され、多数のビット・レーンと1つ又は複数の転送とが、メモリ・コントローラ704から1つ又は複数のメモリ・バス706上の1つ又は複数のメモリ・モジュール710へのコマンド、アドレス、データ、状態、初期化、1つ又は複数の優先順位フラグ、及び他の情報の通信に含まれる。下流パケットは、マルチ・パケット転送の一部として含まれることが多く、各々のパケットのビット・レーンは、(4つのパケットのうちの第1のパケット、又は、合計8つのパケットのうちの第8のパケットなどといった)特定の転送及び/又はパケット番号に基づく同一の又は異なる機能を有する。これは上流パケットにも当てはまる。

【0092】

メモリ・サブシステム、即ちメモリ・モジュール710には、はんだ相互接続、導電性接着剤、ソケット構造体、圧力コンタクト、及び、電氣的、光学的、又は代替的な手段を介して2つ又はそれ以上のデバイス間の通信を可能にする他の方法を含む種々の方法によって、メモリ・デバイス、ハブ、バッファ、レジスタ、クロック・デバイス、受動体、及び、他のメモリ・サポート・デバイス及び/又はコンポーネントを取り付けることができる。

【0093】

1つ又は複数のメモリ・サブシステム、即ちメモリ・モジュール710は、はんだ相互接続、コネクタ、圧力コンタクト、導電性接着剤、光学的相互接続、並びに、他の通信及び電力供給方法といった1つ又は複数の方法によって、メモリ・システム、プロセッサ複合体、コンピュータ・システム、又は他のシステム環境に接続することができる。コネクタ・システムとして、嵌合コネクタ(雄/雌)、雄コネクタ又は雌コネクタと嵌合する1つのキャリア上の導電性コンタクト及び/又はピン、光学的接続、圧力コンタクト(保持

10

20

30

40

50

手段を伴うことが多い)、及び/又は、種々の他の通信及び電力供給方法のうちの1つ又は複数を挙げることができる。1つ又は複数の相互接続は、アップグレード/修理のしやすさ、利用可能な空間/体積、熱移動、部品のサイズ及び形状、並びに、他の関連する物理的、電氣的、光学的、視覚的/物理的アクセスなどの用途要件に応じて、メモリ組立体の1つ又は複数の縁部に沿って配置すること、及び/又は、メモリ・サブシステムの縁部から間隔を空けて設置することができる。

【0094】

ここで用いられるメモリ・サブシステムという用語は、1つ又は複数のメモリ・デバイス、1つ又は複数のメモリ・デバイスと関連するインターフェース及び/又はタイミング/制御回路、及び/又は、メモリ・バッファ、ハブ、及び/又はスイッチと併せた1つ又は複数のメモリ・デバイスを指すが、これらに限定されるものではない。メモリ・サブシステムという用語は、いずれかの関連するインターフェース及び/又はタイミング/制御回路、及び/又は、メモリ・バッファ、ハブ・デバイス、又はスイッチに加えて、メモリ・サブシステムを他の回路に電氣的に取り付けるコネクタ又は同様の手段を含むこともある基板、カード、モジュール、又は関連する組立体に組み込まれた、1つ又は複数のメモリ・デバイスを指すこともある。ここで説明されたメモリ・モジュール710は、1つ又は複数のメモリ・デバイス716とハブ・デバイス708とを含むため、メモリ・サブシステムと呼ぶこともある。

10

【0095】

メモリ・サブシステム、即ちメモリ・モジュール710にローカルに存在することがある付加的な機能は、書込み及び/又は読取りバッファ、1つ又は複数のレベルのメモリ・キャッシュ、ローカル・プリフェッチ論理、データ暗号化/暗号化解除、圧縮/解凍、プロトコル変換、コマンド優先順位付け論理、電圧及び/又はレベル変換、エラー検出及び/又は訂正回路、データ消し込み、ローカルの電力管理回路及び/又は報告、動作及び/又は状態レジスタ、初期化回路、性能の監視及び/又は制御、1つ又は複数のコ・プロセッサ、1つ又は複数の検索エンジン、及び、他のメモリ・サブシステムに予め存在させることができる他の機能を含む。機能をメモリ・サブシステムにローカルに配置することによって、サブシステム内の未使用回路を利用しながら、特定の機能に関して付加的な性能を得ることができる。

20

【0096】

1つ又は複数のメモリ・サブシステム・サポート・デバイスは、1つ又は複数のメモリ・デバイス716が取り付けられるものと同じ基板若しくは組立体に直接取り付けられるか、又は、1つ又は複数のサポート・デバイスを1つ又は複数のメモリ・デバイス716及び/又はメモリ若しくはコンピュータ・システムの他の要素に機能的に相互接続するための電氣的、光学的、若しくは他の通信経路を含む、種々のプラスチック、シリコン、セラミック、若しくは他の材料の1つ又は複数を用いて同じく製造された、別個のインターポーザ若しくは基板に取り付けることができる。

30

【0097】

バス、チャネル、リンク、又は、相互接続に適用される他の名称の従来方法に沿った情報転送(例えばパケット)は、多数の信号送信の選択肢のうちの1つ又は複数を用いて完了することができる。これらの信号送信の選択肢は、シングルエンド方式、ディファレンシャル方式、光学的方式、又は他の方式などの方法を含むことができ、電氣的信号送信は、更に、単層方式又は多層方式のいずれかを用いた電圧又は電流による信号送信方法を含む。信号は、時間又は周波数変調、非ゼロ復帰変調、位相シフト変調、振幅変調、及び他の方法などを用いて変調することもできる。電圧レベルは低下し続けることが予測され、関連する集積回路自体の動作に必要な電力供給電圧の低下と一致して(しかし、これとは独立することも多い)、1.5V、1.2V、1V、及びより低い信号電圧が予測される。

40

【0098】

メモリ・サブシステム及びメモリ・システム自体の内部で、グローバル・クロック、ソ

50

ース同期クロック、エンコード・クロック、又は、これら及び他の方法の組み合わせを含む1つ又は複数のクロック方法を用いることができる。クロック信号送信は、信号ライン自体のクロック信号送信と同じであってもよく、又は、列挙された方法、若しくは、1つ又は複数の計画されたクロック周波数と種々のサブシステム内で計画されたクロック数とにより貢献する代替的な方法のうちの1つを使用してもよい。単一のクロックを、メモリとの間の全ての通信、及び、メモリ・サブシステム内のすべてのクロック機能と関連させてもよく、又は、前述されたような1つ又は複数の方法を用いて多数のクロックをソースにしてもよい。多数のクロックが用いられる場合には、メモリ・サブシステム内における機能は、そのサブシステムの固有のソースとされたクロックと関連させてもよく、又は、(エンコード・クロックと関連するクロックのような)メモリ・サブシステムとの間で転送される情報に関連するクロックから派生したクロックに基づいてもよい。代替りの方法として、メモリ・サブシステムに転送される情報については固有のクロックを用い、メモリ・サブシステムの1つ(又は複数)をソースとする情報については別個のクロックを用いることができる。クロック自体は、通信又は機能の周波数と同一の周波数又はその周波数の複数倍の周波数で動作することができ、データ、コマンド、又はアドレス情報に対して、エッジ整合とするか、中心整合とするか、又は交互タイミング位置に配置することができる。

10

【0099】

1つ又は複数のメモリ・サブシステムに渡される情報は、一般に、アドレス、コマンド、及びデータと、状態又はエラー状況の要求又は報告、メモリのリセット、メモリ又は論理の初期化の完了、及び、他の機能情報、設定情報、又は関連情報と一般に関連する他の信号と、から構成されることになる。1つ又は複数のメモリ・サブシステムから渡される情報は、1つ又は複数のメモリ・サブシステムに渡される情報のいずれか又は全てを含むが、一般にはアドレス及びコマンド情報を含まない。この情報は、通常のメモリ・デバイス・インターフェース仕様と矛盾しない通信方法を用いて(一般には並列に)通信することができ、この情報は、将来のメモリ・インターフェースと矛盾しない、又は、単に通信帯域幅を増大させるために開発された、及び/又は、受信した情報を1つ又は複数の受信側デバイスが必要とするフォーマットに変換することによってサブシステムがメモリ技術とは無関係に動作できるように開発された、「バケット」構造にエンコードすることができる。

20

30

【0100】

メモリ・サブシステムの初期化は、利用可能なインターフェース・バス、所望の初期化速度、利用可能な空間、コスト/複雑性の目標、サブシステムの相互接続構造、この目的又は他の目的のために用いることができる(サービス・プロセッサなどの)代替的なプロセッサの使用などに基づいて、1つ又は複数の方法によって完了することができる。1つの実施形態においては、高速バスを用い、一般には、最初に信頼性のある通信を確立するトレーニング・プロセスを完了し、次いで種々のコンポーネントと関連する属性又は「存在検知」データ、及び/又は、そのサブシステムと関連する特性を問い合わせ、最後にそのシステム内で対象となる動作と関連する情報を用いて適切なデバイスをプログラムすることによって、1つ又は複数のメモリ・サブシステムの初期化を完了することができる。カスケード型システムにおいては、一般に第1のメモリ・サブシステムとの通信が確立され、その次に、後続の(下流の)サブシステムとの通信が、カスケード型相互接続バスに沿ったそれらの位置と矛盾しない順序で確立される。

40

【0101】

第2の初期化方法は、高速バスが、初期化処理中は1つの周波数で動作し、通常の動作中は第2の(通常は、より高い)周波数で動作する方法を含むものである。この実施形態においては、より低い周波数での動作と関連するタイミングのマージンが増大するので、各々のサブシステムの問い合わせ及び/又はプログラミングを完了する前に、カスケード型相互接続バス上の全てのメモリ・サブシステムとの通信を開始することが可能な場合がある。

50

【 0 1 0 2 】

第3の初期化方法は、各々のアドレス、コマンド、及び/又はデータ転送と関連するサイクル数を大きくしながら、カスケード相互接続バスを通常の1つ又は複数の動作周波数で動作させることを含むことができる。1つの実施形態においては、通常の動作中には、アドレス、コマンド、及び/又はデータ情報の全て又は一部を含むパケットは、1つのクロック・サイクルで転送されるが、初期化中は、同一の量及び/又はタイプの情報が、2つ、3つ、又はそれ以上のサイクルにわたって転送されることになる。したがって、この初期化プロセスは、「通常」コマンドではなく「低速」コマンドの形態を用いることになり、サブシステムの各々及びメモリ・コントローラは、これらのサブシステムの各々に含まれるPOR（パワー・オン・リセット）論理によって、電源投入及び/又は再起動後のある時点で自動的にこのモードに入ることができる。

10

【 0 1 0 3 】

第4の初期化方法は、（D e l l他に付与され、本出願人に譲渡された、特許文献1において定められているもののような）存在検知バス、（刊行物21-C revision 7R8における168Pin DIMMファミリーのような、公開されたJEDEC standardsにおいて定められたもののような）I2Cバス、及び/又はSMBUSといった別個のバスを利用することができる、これらは、このようなメモリ・モジュールを用いるコンピュータ・システムにおいて広く利用され、文書化されている。このバスは、デイジー・チェーン/カスケード相互接続構造、マルチ・ドロップ構造、又は代替的な構造のメモリ・システム内の1つ又は複数のモジュールに接続することができ、メモリ・サブシステムに問い合わせ、1つ又は複数のメモリ・サブシステムの各々が全体的なシステム環境内で動作するようにプログラムし、システム環境において望まれる又は検出される性能、温度、設定、又は他の変化に基づいて通常のシステム動作中の他の時点で動作特性を調整する、独立した手段を提供することができる。

20

【 0 1 0 4 】

列挙されたものと共に、又はこれらとは独立に、初期化するための他の方法を用いることもできる。上記の第4の実施形態において説明されたような別個のバスを用いることは、初期化と、サブシステム動作特性のオンザフライでの変更、並びに、使用状況、温度データ、故障情報、又は他の目的といったサブシステム動作情報の報告及びサブシステム動作情報への応答のための変更を含む、本出願人に譲渡されD e l l他に付与された特許文献6に説明されるような初期化以外の用途と、の両方について独立した手段を与えるという利点ももたらす。

30

【 0 1 0 5 】

リソグラフィ技術の改良、より良好なプロセス制御、より抵抗の低い材料の使用、電界サイズの増大、及び他の半導体処理技術の改良に伴って、デバイス回路密度の増加（ダイ・サイズの増加を伴うことが多い）は、集積デバイス上の機能の増強と、以前は別々のデバイスに実装されていた機能の統合とを容易にするのに役立っている。この統合は、対象とする機能の全体的な性能を向上させると共に、ストレージ密度の増加、電力の低減、必要な空間の削減、コストの低下、並びに、他の製造者利益及び消費者利益を促進する役目を果たす。この統合は、自然進化的なプロセスであり、システムに関連する基本的な構成要素を構造的に変更する必要性をもたらすこともある。

40

【 0 1 0 6 】

通信経路、データ・ストレージの内容、及び、メモリ・システム又はメモリ・サブシステムの各要素と関連する全ての機能的動作の完全性は、1つ又は複数の障害検出方法及び/又は訂正方法を用いて、高度に保証することが可能である。種々の要素のいずれか又は全ては、CRC（巡回冗長コード）、EDC（エラー検出訂正）、パリティ、又は、この目的に適した他のエンコード/デコード方法などのエラー検出方法及び/又は訂正方法を含むことができる。更なる信頼性強化として、（情報の転送と関連する断続的な障害などを克服するための）動作のリトライ、障害を起こした経路及び/又はラインを置換するための1つ又は複数の代替通信経路の使用若しくは経路の置換、コンプリメント・リコンプ

50

リメント技術、又は、コンピュータ、通信、及び関連システムにおいて用いられる代替的な方法を挙げることができる。

【0107】

ポイント・ツー・ポイント・リンクのような単純なバス、又はマルチ・ドロップ構造のような複雑なバスにおいてバス終端を用いることは、性能に対する要求の増大と符合してより一般的になってきている。広範な終端方法を特定及び/又は検討することができ、広範な終端方法としてレジスタ、キャパシタ、インダクタ、又はそれらいずれかの組み合わせといったデバイスの使用が挙げられ、これらのデバイスは、信号ラインと、電源供給電圧若しくは接地、終端電圧、又は別の信号との間に接続される。1つ又は複数の終端デバイスは、受動終端構造又は能動終端構造の一部とすることができ、信号ラインの1つ又は複数に沿った1つ又は複数の位置に、及び/又は、1つ又は複数の送信機及び/又は受信側デバイスの一部として、存在させることができる。ターミネータは、伝送ラインのインピーダンスと合うように選択するか、又は、コスト、空間、電力、及び他の制約の範囲内において、使用可能周波数、動作マージン、及び関連する属性を最大化する代替的な手法によって選択することができる。

10

【0108】

例示的な実施形態の技術的な効果及び利点は、より短いメモリ待ち時間とより高いメモリ・スループットとを通してコンピュータ性能を改善することによってより効率的なハブ・ベースのメモリ・システムを実装できる能力を含む。例示的な実施形態は、ハブ・デバイス上で実行されるプリフェッチ論理と、プリフェッチ・データを、非送信請求データ・パケットとして、そのようにパケットを特定するタグと共にメモリ・コントローラに転送することと、を含む。次いで、メモリ・コントローラは、非送信請求読取りデータ・パケットをそのローカル・キャッシュに格納し、その結果、(ハブ・デバイスからデータを要求しなければならない場合と比較して)より速くアクセスできるようにプロセッサの近くにキャッシュ・データが存在することになる。例示的な実施形態においては、データは、バスがアイドル状態の時に、優先順位の低いデータ・パケットとしてメモリ・コントローラに伝送され、それにより、より良好にバスを利用することができ、メモリ・コントローラはプリフェッチされたデータを要求するのに下流バスを使用する必要がなくなる。

20

【0109】

上述されたように、本発明の実施形態は、コンピュータ実装プロセス及びこれらのプロセスを実行するための装置の形態で具体化することができる。本発明の実施形態はまた、フロッピー(商標)ディスク、CD-ROM、ハード・ドライブ、又は他のいずれかのコンピュータ可読ストレージ媒体といった有形の媒体に組み込まれた命令を含むコンピュータ・プログラム・コードの形態で具体化することもでき、そのコンピュータ・プログラム・コードがコンピュータにロードされ、実行されたときに、コンピュータは本発明を実施するための装置になる。本発明はまた、例えば、ストレージ媒体に格納されるか、コンピュータにロードされ及び/又は実行されるか、或いは、例えば電氣的配線若しくはケーブル上で、光ファイバーを通して、又は電磁気放射を介してというように何らかの伝送媒体上で伝送される、コンピュータ・プログラム・コードの形態で具体化することもでき、そのコンピュータ・プログラム・コードがコンピュータにロードされ、実行されたときに、コンピュータは本発明を実施するための装置になる。汎用マイクロプロセッサ上に実装されたときには、コンピュータ・プログラム・コード・セグメントは、特定の論理回路を生成するようにマイクロプロセッサを設定する。

30

40

【0110】

本発明は、例示的な実施形態を参照して説明されたが、当業者であれば、本発明の範囲から逸脱することなく種々の変更を行うことができ、その要素を均等物で置換することができることが分かるであろう。さらに、本発明の本質的な範囲から逸脱することなく、特定の状況又は材料に適合するように本発明の教示に多くの修正を行うことができる。従って、本発明は、本発明を実施するために考慮された最良の形態として開示された特定の実施形態に限定されず、本発明は、特許請求の範囲内の全ての実施形態を含むものとして意

50

図される。さらに、第1、第2といった用語の使用はいずれかの順序又は重要性を示すものではなく、1つの要素を別の要素と区別するために用いられる。

【図面の簡単な説明】

【0111】

【図1】当該技術分野において公知の例示的な初期の同期型メモリ・モジュールを示す。

【図2】当該技術分野において公知の、メモリ・コントローラに直接接続された完全バッファ同期型メモリ・モジュールを有する例示的なコンピュータ・システムを示す。

【図3】当該技術分野において公知の、単一の従来型マルチ・ドロップ・スタブ・バスを有する例示的なメモリ・システムを示す。

【図4】完全バッファ同期型メモリ・モジュールがリピータ機能を含む、当該技術分野において公知の完全バッファ同期型メモリ・モジュール及びシステム構造を示す。

10

【図5】本発明の1つの好ましい実施形態による、データ・アクセス要求をサポートするために並列に動作する複数の独立したデイジー・チェーン・メモリ・インターフェース・チャンネルのブロック図を示す。

【図6】本発明の1つの好ましい実施形態による、データ・アクセス要求をサポートするために並列に動作する複数の独立したデイジー・チェーン・メモリ・インターフェース・チャンネルのブロック図を示す。

【図7】本発明の1つの好ましい実施形態による例示的な実施形態によって実装することができる、1つ又は複数のカスケード型相互接続バスによって相互接続された1つ又は複数のメモリ・チャンネルとメモリ・コントローラとを有するメモリ・システムを示す。

20

【図8】本発明の1つの好ましい実施形態による例示的な実施形態によって実装することができるメモリ・システムのブロック図である。

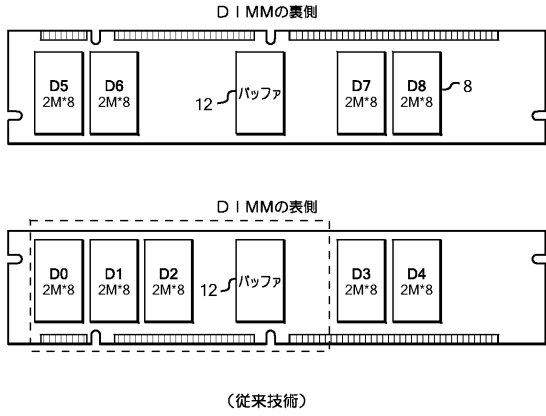
【図9】本発明の1つの好ましい実施形態による例示的な実施形態によって実装することができる、非送信請求データのプリフェッチに続いてハブ・デバイスにおいて実施される動作についての処理フローである。

【図10】本発明の1つの好ましい実施形態による例示的な実施形態によって実装することができる、ハブ・チップからの非送信請求読取りパケットの受信に続いてハブ・デバイスにおいて実施される動作についての処理フローである。

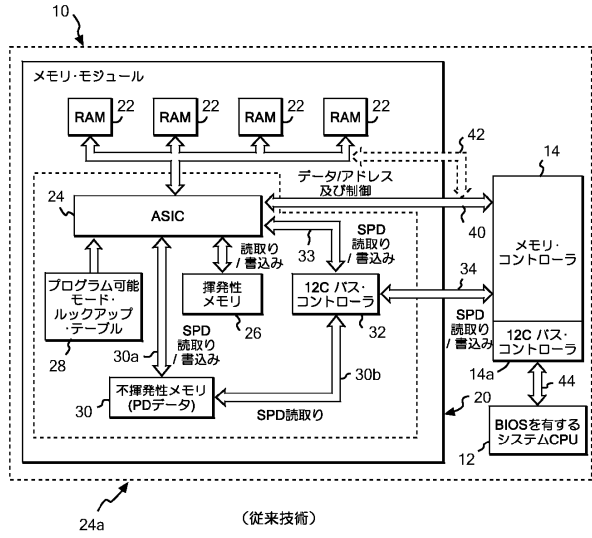
【図11】本発明の1つの好ましい実施形態による例示的な実施形態によって実装することができるパケット構造を示す。

30

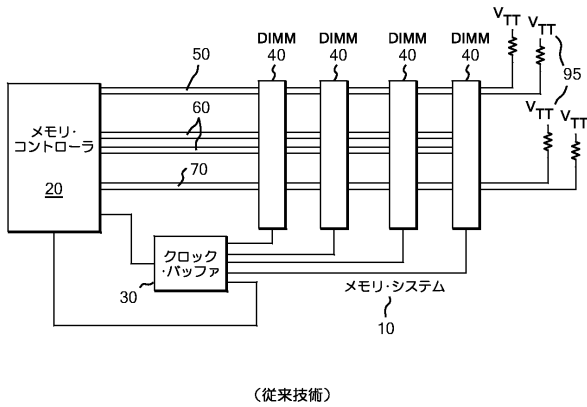
【図1】



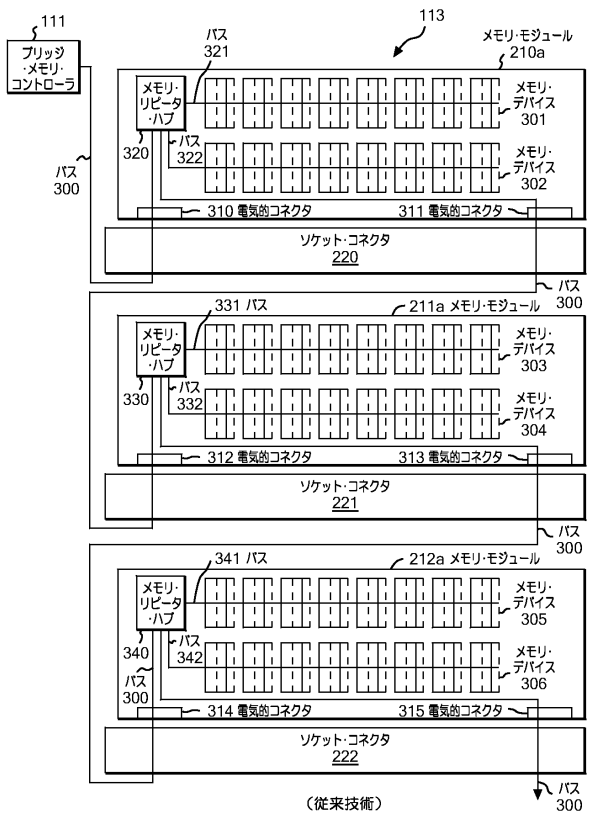
【図2】



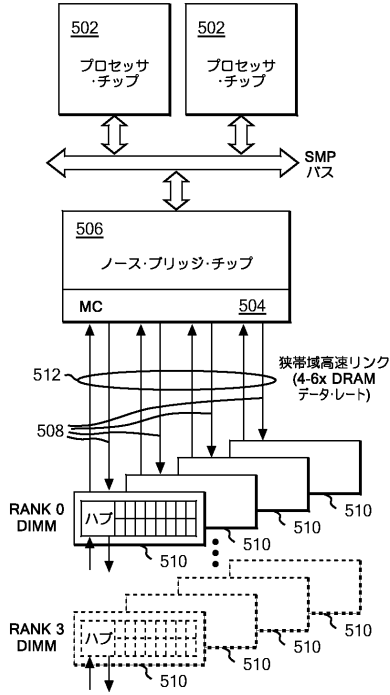
【図3】



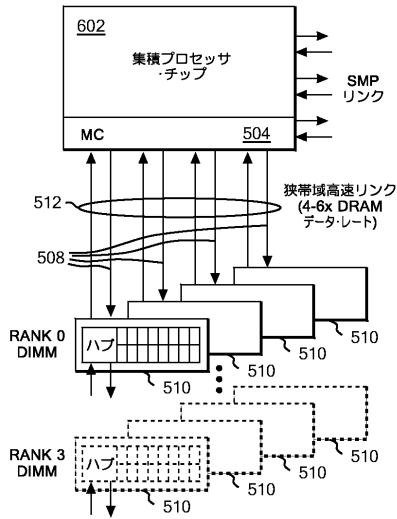
【図4】



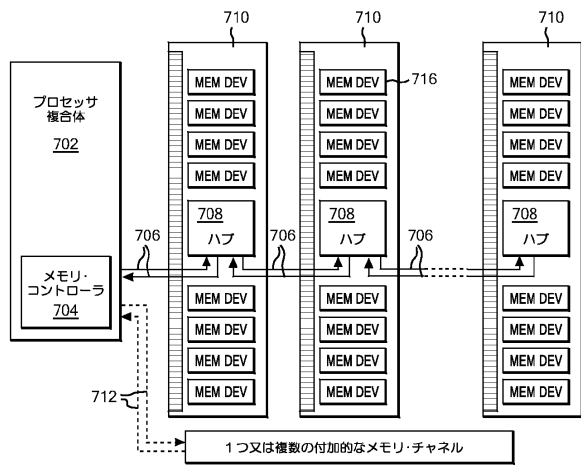
【図5】



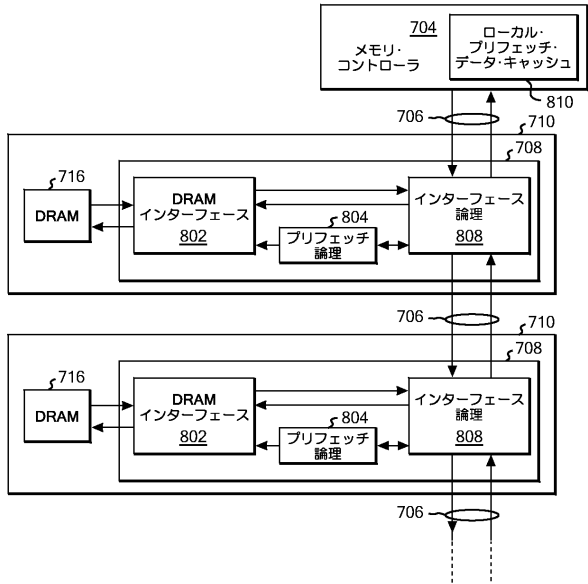
【図6】



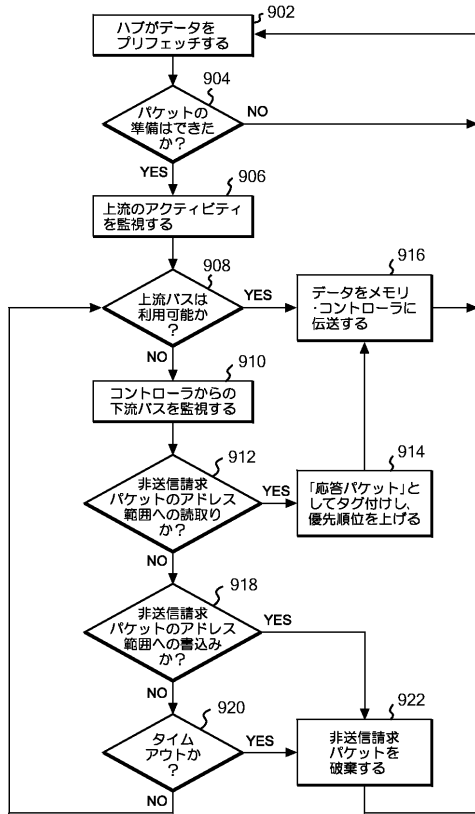
【図7】



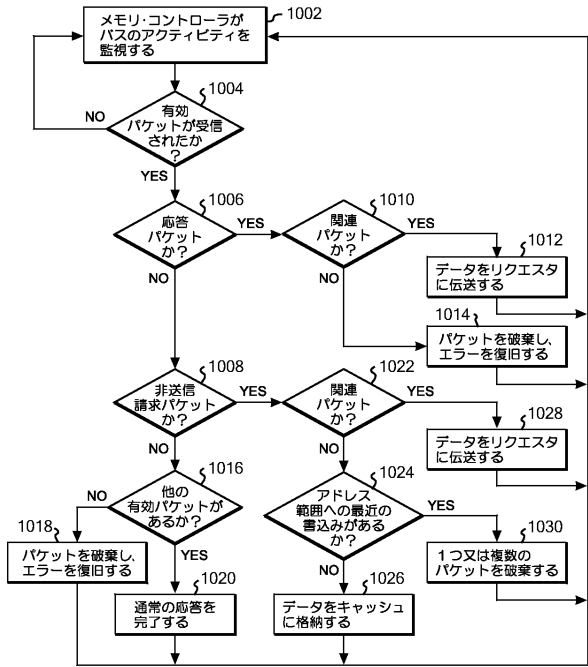
【図8】



【図9】



【図10】



【図11】

		チャンネル・レーン																		
		18	17	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	
		ECC						STS			CMD			タグ						
転送	0																		
	1																		
	2																		
	3	データ																		
	4																		
	5																		
	6																		
7																			

フロントページの続き

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 トレメイン、ロバート

アメリカ合衆国 1 2 5 8 2 ニューヨーク州 ストームビル ステージコーチ・パス 7 3

審査官 中野 裕二

(56)参考文献 特開平 0 4 - 2 3 3 6 4 1 (J P , A)

特表 2 0 0 8 - 5 1 0 2 5 8 (J P , A)

特開 2 0 0 7 - 2 4 1 6 1 2 (J P , A)

特開 2 0 0 5 - 3 0 1 5 9 1 (J P , A)

特開 2 0 0 4 - 5 2 0 6 4 3 (J P , A)

特開平 1 1 - 2 3 2 1 7 1 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G06F 12/00-12/06

G06F 13/16-13/18

G06F 12/08-12/12