

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4439154号
(P4439154)

(45) 発行日 平成22年3月24日(2010.3.24)

(24) 登録日 平成22年1月15日(2010.1.15)

(51) Int.Cl.

H04L 12/56 (2006.01)

F 1

H04L 12/56

E

請求項の数 30 (全 25 頁)

(21) 出願番号 特願2001-521004 (P2001-521004)
 (86) (22) 出願日 平成12年8月21日 (2000.8.21)
 (65) 公表番号 特表2003-508966 (P2003-508966A)
 (43) 公表日 平成15年3月4日 (2003.3.4)
 (86) 國際出願番号 PCT/CA2000/000946
 (87) 國際公開番号 WO2001/017178
 (87) 國際公開日 平成13年3月8日 (2001.3.8)
 審査請求日 平成19年8月7日 (2007.8.7)
 (31) 優先権主張番号 09/386,589
 (32) 優先日 平成11年8月31日 (1999.8.31)
 (33) 優先権主張国 米国(US)

(73) 特許権者 500400227
 モサイド・テクノロジーズ・インコーポレイテッド
 カナダ、ケイ・2・ケイ 2・エックス・1 オンタリオ州、オタワ、ハインズ・コード、11、スイート・203
 (74) 代理人 100064746
 弁理士 深見 久郎
 (74) 代理人 100085132
 弁理士 森田 俊雄
 (74) 代理人 100083703
 弁理士 仲村 義平
 (74) 代理人 100091409
 弁理士 伊藤 英彦

最終頁に続く

(54) 【発明の名称】 インタリープされた非ロックングパケットバッファのための方法および装置

(57) 【特許請求の範囲】

【請求項 1】

スイッチであって、
 複数の入力ポートと、
 複数の出力ポートと、
 複数の入力ポートと複数の出力ポートとによって共有され、複数の入力ポートのいずれかから受取った入力データストリームを記憶し、複数の出力ポートのいずれかに出力データストリームを与えるメモリと、

複数の出力ポートの1つに対するメモリ読出動作をスケジュールし、並行して複数の入力ポートの1つに対するメモリ書込動作をスケジュールするパケットストレージマネージャとを含み、メモリ読出動作およびメモリ書込動作は単一のポートサイクルでメモリの異なる部分に対してスケジュールされ、

パケットストレージマネージャは、

ポートサイクルに依存して、メモリ読出動作のためのメモリにおける第1の場所に対する読出アドレスを選択する読出アドレス論理を含み、読出アドレス論理は、

複数の出力ポートの各々に対するポート待ち行列を含み、ポート待ち行列は入来パケット内のネットワークデステイネーションに依存して、パケットストレージマネージャによってメモリに書込まれたデータの場所のメモリアドレスを記憶し、読出アドレス論理はさらに、

ポートサイクルに依存して、メモリ読出アドレスを取除くべきポート待ち行列を選択す

10

20

るポート待ち行列選択論理を含み、読出すべきメモリ部分は、選択されたポート待ち行列の先頭から取除かれたメモリ読出アドレスに依存し、パケットストレージマネージャはさらに、

読出アドレス論理によって選択された読出アドレスに依存して、メモリ書込動作のためのメモリにおける第2の場所に対する書込アドレスを選択する書込アドレス論理を含む、スイッチ。

【請求項2】

書込アドレス論理は、

アドレスの空きリストと、

空きリストから書込アドレスを取除くことにより、メモリにおける書込アドレスを与える書込選択論理とを含み、

読出アドレス論理は、書込アドレスをパケットセグメントが方向付けられる各ポート待ち行列の最後に記憶し、

パケットストレージマネージャは、書込アドレスによって指定されるメモリの別の部分にパケットセグメントを書込む、請求項1に記載のスイッチ。

【請求項3】

メモリは物理的に複数のバンクに分けられる、請求項2に記載のスイッチ。

【請求項4】

バンクの数は2個である、請求項3に記載のスイッチ。

【請求項5】

バンクの数は4個である、請求項3に記載のスイッチ。

【請求項6】

書込アドレス論理は各バンクに対する空きリストをさらに含み、書込選択論理は書込アドレスを取除くべき空きリストの1つを選択する、請求項3に記載のスイッチ。

【請求項7】

書込アドレス論理は読出アドレスに依存して、空きリストを選択する、請求項6に記載のスイッチ。

【請求項8】

書込アドレス論理は、データパケットの順次セグメントがメモリの交互の奇数バンクおよび偶数バンクに書込まれるよう、空きリストを選択する、請求項6に記載のスイッチ。

【請求項9】

書込アドレス論理は

複数のバンクの各々に対する空きリストカウンタをさらに含み、空きリストカウンタはバンクにおける利用可能な場所のカウントを記憶し、書込選択論理はこのカウントに依存して空きリストの1つを選択する、請求項6に記載のスイッチ。

【請求項10】

入力ポートの1つおよび出力ポートの1つは、同じポートであるか、または異なるポートである、請求項1に記載のスイッチ。

【請求項11】

スイッチであって、

複数の入力ポートと、

複数の出力ポートと、

複数の入力ポートと複数の出力ポートとによって共有され、複数の入力ポートのいずれかから受取った入力データストリームを記憶し、複数の出力ポートのいずれかに出力データストリームを与えるメモリと、

複数の出力ポートの1つに対するメモリ読出動作をスケジュールし、並行して複数の入力ポートの1つに対するメモリ書込動作をスケジュールするための手段とを含み、メモリ読出動作およびメモリ書込動作は单一のポートサイクルでメモリの異なった部分に対してスケジュールされ、

スケジュールするための手段は、

10

20

30

40

50

ポートサイクルに依存して、メモリ読出動作のためのメモリにおける読出アドレスを選択するための手段を含み、読出アドレスを選択するための手段は、

複数の出力ポートの各々に対するポート待ち行列を含み、ポート待ち行列は入来パケット内のネットワークデスティネーションに依存して、スケジュールするための手段によってメモリに書込まれたデータの場所のメモリアドレスを記憶し、読出アドレスを選択するための手段はさらに、

ポートサイクルに依存して、読出アドレスを取除くべきポート待ち行列を選択するための手段を含み、読出すべきメモリ部分は、選択されたポート待ち行列の先頭から取除かれた読出アドレスに依存し、スケジュールするための手段はさらに、

読出アドレスを選択するための手段によって選択された読出アドレスに依存して、メモリ書込動作のためのメモリにおける書込アドレスを選択するための手段とを含む、スイッチ。10

【請求項 1 2】

書込アドレスを選択するための手段は、

アドレスの空きリストと、

空きリストから書込アドレスを取除くための手段とを含み、

読出アドレスを選択するための手段は、セグメントが方向付けられている各ポート待ち行列の最後に書込アドレスを記憶し、

取除くための手段は、スケジュールするための手段にパケットを書込むための書込アドレスを与え、スケジュールするための手段は、書込アドレスによって指定されたメモリの別の部分にセグメントを書込む、請求項 1 1 に記載のスイッチ。20

【請求項 1 3】

メモリは物理的に複数のバンクに分けられている、請求項 1 2 に記載のスイッチ。

【請求項 1 4】

バンクの数は 2 個である、請求項 1 3 に記載のスイッチ。

【請求項 1 5】

バンクの数は 4 個である、請求項 1 3 に記載のスイッチ。

【請求項 1 6】

書込アドレスを選択するための手段は各バンクに対する空きリストをさらに含み、取除くための手段は書込アドレスを取除くべき空きリストの 1 つを選択する、請求項 1 3 に記載のスイッチ。30

【請求項 1 7】

書込アドレスを選択するための手段は、読出アドレスに依存して空きリストを選択する、請求項 1 6 に記載のスイッチ。

【請求項 1 8】

読出アドレスを選択するための手段は、

複数のバンクの各々に対する空きリストカウンタをさらに含み、空きリストカウンタはバンクにおける利用可能な場所のカウントを記憶し、書込アドレスを選択するための手段は空きリストカウンタにおけるカウントに依存して空きリストの 1 つを選択する、請求項 1 7 に記載のスイッチ。40

【請求項 1 9】

書込アドレスを選択するための手段は、データパケットの順次セグメントがメモリの交互の奇数バンクおよび偶数バンクに書込まれるよう、空きリストを選択する、請求項 1 6 に記載のスイッチ。

【請求項 2 0】

入力ポートの 1 つおよび出力ポートの 1 つは、同じポートであるか、または異なるポートである、請求項 1 1 に記載のスイッチ。

【請求項 2 1】

スイッチにおける共通メモリのバンド幅を増大させるための方法であって、

複数の入力ポートのいずれか 1 つからメモリにおける入力データストリームを受取るス50

ステップと、

メモリからの出力データストリームを複数の出力ポートのいずれか1つに与えるステップと、

複数の出力ポートの1つに対するメモリ読出動作をスケジュールし、並行して複数の入力ポートの1つに対するメモリ書込動作をスケジュールするステップとを含み、メモリ読出動作およびメモリ書込動作は单一のポートサイクルでメモリの異なった部分に対してスケジュールされ、メモリ読出動作およびメモリ書込動作をスケジュールするステップは、

ポートサイクルに依存して、メモリ読出動作のためのメモリにおける読出アドレスを選択するステップを含み、読出アドレスを選択するステップは、

複数の出力ポートの各々に対するポート待ち行列におけるメモリに書込まれたデータの場所のメモリアドレスを記憶するステップを含み、ポート待ち行列は入来パケット内のネットワークデスティネーションに依存し、読出アドレスを選択するステップはさらに、

ポートサイクルに依存して、読出アドレスを取除くべきポート待ち行列を選択するステップを含み、読出すべきメモリ部分は、選択されたポート待ち行列の先頭から取除かれた読出アドレスに依存し、メモリ読出動作およびメモリ書込動作をスケジュールするステップはさらに、

読出アドレスを選択するステップによって選択された読出アドレスに依存して、メモリ書込動作のためのメモリにおける書込アドレスを選択するステップと、

メモリ読出動作とメモリ書込動作とを、单一のポートサイクルでメモリの異なる部分にスケジューリングするステップとを含む、方法。

【請求項22】

書込アドレスを選択するステップは、

アドレスの空きリストを与えるステップと、

空きリストから書込アドレスを取除くステップと、

セグメントが方向付けられている各ポート待ち行列の最後に書込アドレスを記憶するステップと、

書込アドレスによって指定されたメモリの別の部分にパケットセグメントを書込むためのアドレスを与えるステップとを含む、請求項21に記載の方法。

【請求項23】

メモリは物理的に複数のバンクに分けられている、請求項22に記載の方法。

【請求項24】

バンクの数は2個である、請求項22に記載の方法。

【請求項25】

バンクの数は4個である、請求項22に記載の方法。

【請求項26】

空きリストは読出アドレスに依存して選択される、請求項25に記載の方法。

【請求項27】

書込アドレスを選択するステップは、

複数のバンクの各々に対して、利用可能な場所のカウントを空きリストカウンタに記憶するステップと、

空きリストカウンタにおけるカウントに依存して、空きリストのうちの1つを選択するステップとをさらに含む、請求項26に記載の方法。

【請求項28】

データパケットの順次セグメントがメモリの交互の奇数バンクおよび偶数バンクに書込まれるよう、バンク空きリストが選択される、請求項25に記載のスイッチ。

【請求項29】

書込アドレスを選択するステップは、

複数のバンクの各々に対する空きリストに利用可能な場所のアドレスを記憶するステップと、

複数の空きリストの1つを選択するステップと、

10

20

30

40

50

選択された空きリストから書込アドレスを除くステップとをさらに含む、請求項 22 に記載の方法。

【請求項 30】

入力ポートのうちのいずれか 1 つおよび出力ポートのうちのいずれか 1 つは、同じポートであるか、または異なるポートである、請求項 21 に記載の方法。

【発明の詳細な説明】

【0001】

【発明の背景】

ネットワークスイッチは、スイッチに接続されたいくつかの入力ポートからデータパケットを受取り、データパケットをスイッチに接続されたいくつかの出力ポートに与える。スイッチは、データパケットに含まれるデスティネーションアドレスに依存して、データパケットを与えるべき出力ポートを判断する。入力ポートから受取られたデータパケットは、出力ポートに与えられる前にスイッチ内のメモリに記憶される。

10

【0002】

スイッチ内のメモリは、すべての入力ポートから受取られたすべてのデータパケットが出力ポートに与えられる前に記憶される共通メモリであり得る。非ブロッキングスイッチは、すべての入力ポートで受取られたすべてのデータが、出力ポートに与えられることを可能にする。非ブロッキングスイッチは、典型的にはポートごとに最大量のメモリを利用可能にするための共通メモリを含む。

【0003】

20

共通メモリスイッチの速度は、メモリバンド幅に依存する。メモリバンド幅はメモリアクセスタイムおよびメモリの幅に依存するが、これはすなわち、メモリアクセス時間ごとにアクセスされるバイトの数である。たとえば、もし共通メモリが 64 バイト幅であり、かつ 80 ナノ秒 (ns) のアクセス時間を有していれば、64 バイトを書込または読出するために 80 ns かかる。スイッチに接続される入力ポートが 100 メガビット (M) bps (秒当たりのビット) でデータを受取ると、1 ビットは 10 ns ごとに受取られ、8 ビットバイトは 80 ns ごとに受取られ、かつ 64 バイトは 5120 ns ごとに受取られる。64 バイトのデータが受取られた後で、書込メモリサイクルが行なわれて 64 バイト幅メモリへの単一のアクセスで 64 バイトを書込む。入力ポートは、5120 ns ごとにメモリバンド幅を 80 ns 使用して、受取ったデータをメモリに書込む。こうして、80 ns のアクセス時間を備えた 64 バイト幅メモリは、64 (5120 ns / 80 ns) の 100 M bps のポートをサポートできる。64 ポートがスイッチに接続されていると、各ポートサイクルは 80 ns である。80 ns ポートサイクルは、各ポートに、ポートサイクルごとに 1 つのメモリアクセスを提供する。

30

【0004】

各ポートサイクルで利用可能であるメモリバンド幅を増大させると、以下のうちの 1 つ以上が必要となる：メモリアクセス時間の減少、メモリの幅（すなわち、メモリサイクルごとに読出 / 書込されるビット数）の増大、または、ポートの数の減少。メモリの幅の増大は、最小データパケットサイズによって制限される。メモリアクセス時間の減少は、メモリに対する最小メモリアクセス時間によって制限される。

40

【0005】

メモリバンド幅は、たとえば第 1 のデータパケットを第 1 のメモリバンクに書込み、第 2 のデータパケットを第 2 のデータバンクに書込むためのメモリバンクをインターリーブさせることによっても増大させることができる。しかしながら、スイッチ内のデータをインターリーブさせると、たとえばもしポート A およびポート B の両方が同時に第 1 のメモリバンクへのアクセスを要求している場合、ブロッキングをまねくおそれがある。入力ポートと出力ポートとの間でのデータ転送のブロッキングは、パケットの脱落をもたらすおそれがある。よって、非ブロッキングスイッチにおいては、メモリバンド幅を増大させるために用いることはできない。

【0006】

50

【発明の要約】

スイッチにおけるパケットストレージマネージャは、スイッチに接続される入力ポートおよび出力ポートによって共有されるメモリのメモリバンド幅を増大させる。パケットストレージマネージャは、先行技術のシステムが逐次的な読出および書きサイクルを必要としたところを、単一のポートサイクルで、入力ポートのうちの1つに対する書き動作および出力ポートのうちの1つに対する読出動作の両方を行なう。書きおよび読出動作は、単一のメモリアクセスサイクルにおいて、メモリ内の異なったメモリに対して並行に行なわれる。メモリは、物理的にいくつかのバンクに分割される。バンクの数は、好ましくは2つまたは4つである。読出および書き動作は異なったバンクに対して行なわれる。

【0007】

パケットストレージマネージャは、ポートサイクルに依存して読出動作のためにメモリにおける読出アドレスを選択する読出アドレス論理と、読出アドレス論理によって選択された読出アドレスに依存して、書き動作のための書きアドレスを選択する書きアドレス論理とを含む。書きアドレスは、読出アドレスに依存して書きアドレスを選択して、読出および書き動作が単一のメモリアクセスで並行して行なわれるようになる。パケットストレージマネージャは、出力ポートごとのポート待ち行列と、ポート待ち行列には記憶されていないアドレスの空きリストとに依拠する。入来するパケットから、マネージャはネットワークデステイションを読み出し、パケットが記憶されるべき適切な出力ポートを判断する。マネージャは、空きリストからメモリセグメントアドレスを除去することによりパケットセグメントをメモリに書き込み、パケットが向けられるべき各ポート待ち行列の最後にセグメントアドレスを記憶し、セグメントアドレスによって特定されるメモリ内の場所にパケットセグメントを書き込む。同時に、各ポート待ち行列の先頭でのアドレスは、マネージャによって読み出されるべきパケットセグメントを識別する。

【0008】

パケットストレージマネージャ内の読出アドレス論理は、各出力ポートに対するポート待ち行列を含む。ポート待ち行列は、各入力ポートによってメモリに書き込まれるデータのメモリ場所を記憶する。ポート待ち行列選択論理は、ポートサイクルに依存して、メモリアドレスが除去されるべきポート待ち行列を選択する。読出選択論理は、ポート待ち行列から除去されたメモリアドレスに依存して、読み出されるべきメモリ場所を選択する。

【0009】

メモリは物理的にいくつかのバンクに分割される。バンクの数は好ましくは2つまたは4つである。パケットストレージマネージャ内の書きアドレス論理は、バンクごとのバンク空きリストを含む。各バンク空きリストは、メモリのバンク内の利用可能な場所のアドレスを記憶する。書きアドレス論理内の書き選択論理は、書きアドレスを除去すべきバンク空きリストのうちの1つを選択する。書き選択論理は、バンク空きリストを選択して、それによりデータパケットの順次的セグメントが、メモリの交互の奇数および偶数バンクに書き込まれるようにする。

【0010】

書きアドレス論理はまた、バンクごとのバンク空きリストカウンタを含む。バンク空きリストカウンタは、バンク内の利用可能な場所のカウントを記憶する。書き選択論理は、バンク空きリストカウンタにおけるカウントに依存して、バンク空きリストを選択し得る。並行的な読出および書き動作は、同じポートに対するものであっても、または異なったポートに対するものであってもよい。

【0011】

上述の目的ならびに他の目的、およびこの発明の特徴ならびに利点は、添付の図面に示されるこの発明の好ましい実施例により特定的な説明から明らかとなるであろうが、図面においては同様の要素には同様の参照符号が異なる図面を通して付与される。図面は必ずしも実物と同じスケールである必要はなく、この発明の原理を例示することが重視される。

【0012】

【発明の説明】

10

20

30

40

50

図1Aは、共通メモリスイッチ100のブロック図である。入力ポート102で受取られたすべてのデータは、セグメントバッファメモリ108に記憶されて、その後1つ以上の出力ポート112にスイッチされる。パケットストレージマネージャ106は、セグメントバッファメモリ108に対する書込および読出アクセスを制御する。パケットストレージマネージャ106を介して、入力ポート102に到着するデータパケットセグメントはセグメントバッファメモリ108に書込まれ、同じメモリアクセスサイクルにおいて、出力ポート112に対する別のデータパケットセグメントがセグメントバッファメモリ108から読出される。

【0013】

スイッチ100は、入力ポートエンジン104および出力ポートエンジン110を含む。
データパケットは入力ポート102で順次的に受取られる。入力エンジン104は、受取ったデータパケット内のヘッダを検出して処理し、検出したヘッダからどの出力ポート112にデータパケットを転送するべきかを判断し、データパケットに対する転送ベクトル114を生成する。転送ベクトル114はビットマップであり、1つのビットが複数の出力ポート112の各々に対応し、データパケットがその出力ポート112に転送されるべきどうかを示す。転送ベクトル114は、パケットストレージマネージャ106に転送される。

【0014】

パケットストレージマネージャ106は、セグメントバッファメモリ108へのアクセスを提供する。パケットストレージマネージャ106は、セグメントバッファ108への読出および書込動作に対するセグメントバッファメモリアドレス122を提供し、マネージャ106内に記憶された各データパケットのセグメントバッファメモリ108における場所を記憶する。出力エンジン110は、選択制御信号120を介して、データパケットを転送するべき複数の出力ポート112のうちの1つを選択し、選択された出力ポート112に記憶されたデータパケットを与える。

【0015】

セグメントバッファメモリ108は、すべての入力ポート102および出力ポート112によって共有される共通メモリである。スイッチ100は非ブロッキング、すなわち、入力ポート102のいずれかに到着するデータパケットが出力ポート112のいずれかに転送されることをブロックされないものである。スイッチ100は、入力ポート102で受取られたデータパケットの入力ポートエンジン104による並行処理を提供し、かつ出力ポートエンジン110による出力ポート112に対する記憶されたデータパケットの処理を提供する。

【0016】

データパケットは入力ポート102のうちの1つで順次的に受取られるにつれて、シリアルデータはデータセグメントにグループ化される。データセグメントはバッファセグメントメモリ108に書込まれる。データセグメントの幅は、入力ポート102および出力ポート112によって用いられるネットワーキングプロトコルに依存して、予め定められる。たとえば、イーサネット(R)ネットワーキングプロトコルに対しては、データセグメントの幅は64バイトであるが、これはイーサネット(R)データパケットに対する最小データパケットサイズが64バイトであるためである。

【0017】

図1Bは、入力ポート102で受取られる先行技術のイーサネット(R)データパケット120を示すブロック図である。イーサネット(R)データパケットは、ヘッダ122、データフィールド134、およびフレームチェックシーケンス132を含む。ヘッダ122は、デスティネーションアドレス124、ソースアドレス126、および長さまたはタイプフィールド130を含む。データパケット120のサイズは、データフィールド134のサイズに依存し、これは64バイトから1500バイトにわたり得る。

【0018】

図2は、スイッチ100におけるセグメントバッファメモリ108による入力ポート10

10

20

30

40

50

8で受取られたデータの出力ポート112へのスイッチングを示すタイミング図である。タイミング図は、3ポートスイッチ($N = 2$)に対する図1Aにおけるブロック図に関する説明され、すなわち、3つの入力ポート102および3つの出力ポート112が示される。

【0019】

時間200から時間202の間で、データセグメントは3つの入力ポート102の各々において順次的に受取られる。時間200および時間202の間のデータセグメントタイムスロット204は、入力ポート102でデータセグメントを受取るためにとられる時間である。データセグメントが受取られる間に、入力ポートエンジン104はデータパケットのヘッダ内のデスティネーションアドレス124から、どの出力ポート112にデータセグメントを転送すべきを判断する。時間202において、すべてのデータセグメントが受取られており、パケットストレージマネージャ106によってセグメントバッファメモリ108に書込まれることができる。パケットストレージマネージャ106は、時間206の前にすべての受取ったデータセグメントをセグメントバッファメモリ108に書込むが、このときに次のデータセグメントが受取られる。

10

【0020】

データセグメントタイムスロット204はポートサイクルタイムスロット208に分割され、データセグメントタイムスロット204ごとの1つのポートサイクルタイムスロット208が、スイッチ100に接続される各ポートに割当てられる。ポートサイクル時間208はメモリアクセスサイクルであり、すなわち、セグメントバッファメモリ108にアクセスするためのポートごとに提供される時間である。時間210で、パケットストレージマネージャ106は、入力ポート0102において受取られたデータセグメントを、セグメントバッファメモリ108に書込む。時間212から時間214においては、パケットストレージマネージャ106は入力ポート1102で受取られたデータセグメントをセグメントバッファメモリ108に書込む。時間214において、入力ポート0で受取られたデータセグメントはメモリにおいて利用可能であり、出力ポートエンジンによって読み出されて出力ポート1に転送されることができる。時間214では、パケットストレージマネージャ106は入力ポート2102で受取られたデータをセグメントバッファメモリ108に書込む。時間216において、入力ポート1で受取られたデータセグメントはメモリにおいて利用可能であり、出力ポートエンジンによって読み出されて出力ポート2に転送されることができる。時間216では、パケットストレージマネージャ106が入力ポート0102で受取られた次のデータセグメントをセグメントバッファメモリ108に書込む。時間218においては、入力ポート2で受取られたデータセグメントがバッファセグメントメモリ108において利用可能であり、出力ポートエンジン110によって読み出されて出力ポート0112に転送されることができる。

20

【0021】

こうして、データセグメントは入力ポート102で受取られると、パケットストレージマネージャ106によってセグメントバッファメモリ108に書込まれる。データセグメントは、受取られたデータセグメントが書込まれる間に、セグメントバッファメモリ108から読み出されることができる。入力ポート102はセグメントバッファメモリ108への書き込みがブロックされることなく、出力ポート112はセグメントバッファメモリ108からの読み出しがブロックされることがない。

30

【0022】

図3は、図1Aに示されるパケットストレージマネージャ106を示すブロック図である。パケットストレージマネージャ106は、入力ポートエンジン104(図1A)からのセグメントバッファメモリ108への書き込み要求と、出力ポートエンジン110(図1A)からのセグメントバッファメモリ108からの読み出しが並行して扱う。

40

【0023】

パケットストレージマネージャ106は、読み出アドレス論理300、読み出アドレス論理302、ポートタイミング論理304およびアドレス選択論理312を含む。書き込みアドレス

50

論理 300 は、入力ポート 102 (図 1A) からのデータセグメントが記憶されるバッファセグメントメモリ 108 内の場所を提供する。読み出アドレス論理 300 は、出力ポート 112 に対するデータパケットが記憶されているバッファセグメントメモリ 108 内の場所を提供する。

【0024】

アドレス選択論理 312 は、読み出アドレス 310 および書き込みアドレス 308 をバッファセグメントメモリアドレス 122 に転送する。ポートタイミング論理 304 は、クロック 306 入力に応じてポートサイクル 314 を生成する。ポートタイミング論理 304 は、予め定められた数のクロック周期のポートサイクル時間 208 に対しポートサイクル 314 を生成し、予め定められたデータセグメントサイクル時間 204 ごとに各ポート 102、112 に対するポートサイクル 314 を生成する。データセグメントサイクル時間 204 は、入力ポート 102 (図 1A) のいずれかで順次的にデータセグメントを受取るための時間に依存する。読み出アドレス 310 および書き込みアドレス 308 は、ポートサイクル 314 に依存してポート 102、112 に対して選択される。ポートタイミング論理 304 によって、示されるように同じポートサイクル 314 が、書き込みアドレス論理 300 読み出アドレス論理 302 に与えられるか、またはポートタイミング論理 304 によって、異なったポートサイクル 314 が、書き込みアドレス論理 300 および読み出アドレス論理 302 に与えられてもよい。

【0025】

各ポートサイクル 314 の開始時に、読み出アドレス論理 302 はポートサイクル 314 に依存して、ポートに対する読み出アドレス論理 302 におけるポート待ち行列の先頭での先行して待ち行列に入れられた読み出アドレス 310 を選択する。書き込みアドレス論理 300 は、読み出アドレス論理 302 によって選択された読み出アドレス 310 に依存して、ポートサイクル 314 に対する書き込みアドレスを選択する。

【0026】

アドレス選択論理 312 は、セグメントバッファメモリアドレス 122 を介して書き込みアドレス 308 および読み出アドレス 310 をバッファセグメントメモリ 108 に転送する。セグメントバッファメモリ 108 に書き込まれるべきデータパケットの書き込みアドレス 308 は、読み出アドレス論理 302 内の待ち行列に入れられる。バッファセグメントメモリ 108 から読み出されるべきデータパケットの読み出アドレス 302 は、読み出アドレス論理 302 内の待ち行列から外される。書き込みアドレス論理 300 および読み出アドレス論理 302 は、図 5、図 6、図 10 および図 11 により詳細に示される。

【0027】

図 4 は、物理的に 2 つのバンク 108A および 108B に分割された、図 1 に示されたセグメントバッファメモリ 108 を示すブロック図である。セグメントバッファメモリ 108 の各バンクは関連のメモリアドレスであるアドレス_A_400 およびアドレス_B_402 を有する。二バンクアドレス選択論理 404 は、書き込みアドレス 308 に符号化されたセグメントバッファメモリ 108 のバンクに依存して、書き込みアドレス 308 をアドレス_A_400 またはアドレス_B_402 に転送する。二バンクアドレス選択論理 404 は、読み出アドレス 310 に符号化されたバッファセグメントメモリ 108 のバンクに依存して、読み出アドレス 310 をアドレス_A_400 またはアドレス_B_402 に転送する。たとえば、読み出アドレス 310 および書き込みアドレス 308 の最上位のビット (「MSB」) を、セグメントバッファメモリ 108A および 108B のバンクを判断するために用いることができ、書き込みアドレス 308 および読み出アドレス 310 は二バンクアドレス選択論理 404 によって転送される。書き込みアドレス論理 300 (図 3) および読み出アドレス論理 304 (図 3) は、同じポートサイクル 314 (図 2) の間に生成された読み出アドレス 308 および書き込みアドレス 310 がセグメントバッファメモリ 108A、B の異なったバンクに対するものであることを確実にする。セグメントバッファメモリ 108 の一方のバンクに対する読み出アクセスは、バッファセグメントメモリ 108 の他のバンクへの書き込みアクセスと並行して生じる。

10

20

30

40

50

【0028】

図5は、図4に示される2つのバンクのセグメントバッファメモリ108に対する図3に示される読出アドレス論理302を示すブロック図である。読出アドレス論理302は、ポート待ち行列選択論理512、ポート待ち行列510、および読出アドレス選択論理506を含む。ポート待ち行列510は、出力ポート112に与えられるべきデータパケットのセグメントバッファメモリにおけるアドレス場所を記憶する。出力ポート112ごとに別々のポート待ち行列510が与えられる。データセグメントがセグメントバッファメモリ108に書込まれるにつれて、データセグメントが書込まれた場所である書込アドレス308が1つ以上のポート待ち行列510に記憶される（または待ち行列に入れられる）。書込アドレス308が書込まれるべきポート待ち行列510は、ポート待ち行列選択論理512によって判断される。ポート待ち行列選択論理512は、転送ベクトル114に依存してポート待ち行列510を選択する。転送ベクトル114の生成は、図1Aを参照して説明した。
10

【0029】

ポート待ち行列510に記憶された書込アドレス308でのデータセグメントが、コピーが送られるべきすべての出力ポートによってセグメントバッファメモリ108から読出された後で、記憶された書込アドレス308はポート待ち行列510から外される。

【0030】

図6は、図3に示される2つのバンクのセグメントバッファメモリ108A、108Bに対する図2に示される書込アドレス論理300を示すブロック図である。2つのバンクのセグメントバッファメモリ108A、108Bに対する書込アドレス論理は、二バンク書込選択論理600、バンク_Aに対する空きリスト606、バンク_Bに対する空きリスト608、バンク_Aに対する空きリストに対するカウント604、バンク_Bに対する空きリストに対するカウント610、および書込アドレス選択論理602を含む。
20

【0031】

各空きリスト606、608は、データセグメントが書込まれるであろうセグメントバッファメモリのそれぞれのバンク108A、108Bにおける利用可能な場所を記憶する。バンク_Aに対する空きリスト606は、セグメントバッファメモリ108Aのバンク_Aにおける利用可能な場所を記憶し、バンク_B 108Bに対する空きリストは、セグメントバッファメモリ108Bのバンク_Bにおける利用可能な場所を記憶する。
30

【0032】

二バンク書込選択論理600は、ポートサイクル時間314の間に、読出アドレス310に依存して書込アドレス308が抽出されるべき空きリスト606、608を判断する。セグメントバッファメモリのバンク108A、108Bは、読出アドレス310のMSB、または読出アドレス310の他のビットに依存して選択され得る。たとえば、もし二バンク書込選択論理600が、読出アドレス310から、読出アドレス310はセグメントバッファメモリのバンク_A 108Aに対するものであると判断すると、書込アドレスはバンク_Bに対する空きリスト608から抽出される。もし読出アドレス310が、バッファセグメントメモリのバンク_B 108Bに対するものであれば、書込アドレスはバンク_Aに対する空きリスト606から抽出される。こうして、データセグメントがセグメントバッファメモリ108の一方のバンクから読出されることができる一方、別のデータセグメントがセグメントバッファメモリ108の他方のバンクに書込まれる。
40

【0033】

もし二バンク書込選択論理600が、読出アドレスがバッファセグメントメモリのバンク_A 108Aに対するものであると判断すると、これはENA_B信号618によってバンク_B空きリスト608をイネーブルし、それによりバンク_B空きリスト608は書込アドレス選択論理602にバンク_Bアドレス614を提供することができる。アドレスがバンク_B空きリスト608から除去されると、二バンク書込選択論理600は獲得/更新カウントB信号616に更新カウント_B信号を生成し、バンク_Bに対する空きリストカウント610をデクリメントする。
50

【0034】

データセグメントが、書込アドレス308によって特定される場所でセグメントバッファメモリ108に書込まれる一方、書込アドレス308はそれぞれの空きリスト606、608から除去され、書込アドレス308は転送ベクトル114に依存して1つ以上のポート待ち行列306の最後に書込まれる。コピーカウント626が書込アドレス308に対してコピーカウント論理624に記憶される。コピーカウントは、マルチキャストパケット、すなわち1つ以上の出力ポート112に対して待ち行列に入れられるパケットを追跡するために用いられる。書込アドレス308は、読出アドレス308によって特定された場所が読出アドレス論理302(図3)によってすべて読出され、かつ書込アドレス308に関連のコピーカウントが0になった後で、それぞれの空きリスト606、608に返される。10

【0035】

現在のポートサイクル314に対して読出されるべきデータセグメントがなく、書込まれるべきデータセグメントがあれば、二バンク書込選択論理600は空きリストカウント604、610に依存して書込アドレス308を抽出すべき空きリスト606、608を決定する。二バンク書込選択論理600は、獲得/更新カウント_A信号622および獲得/更新カウント_B信号616によって、バンク_Aに対する空きリストカウント604およびバンク_Bに対する空きリストカウント610から現在のカウントを獲得する。バンク_Aに対する空きリストカウント632およびバンク_Bに対する空きリストカウント630は、二バンク書込アドレス論理600に転送される。最大のカウントを備えた空きリストカウント604、610は、セグメントバッファメモリ108A、108Bの最も空いているバンクを示す。二バンク書込アドレス論理600はカウントを比較して、バンク_Aカウント632またはバンク_Bカウント630がより大きいかどうかに依存して、ENA_A信号620またはENA_B信号618を生成する。最大カウント604、610を備えた空きリスト606、608が選択される。書込アドレス308は選択された空きリスト606、608から選択される。こうして、データセグメントがセグメントバッファメモリ108に書込まれるにつれて、これらは均等に2つのバンクのセグメントバッファメモリ108A、108Bに分散される。20

【0036】

図7は、図4に示される2つのバンクのセグメントバッファメモリ108に対する読出および書込アドレスに対するタイミングを示すタイミング図である。各データセグメントタイムスロット204において、スイッチ100に接続された64ポートごとにセグメントバッファメモリへの2つのアクセスが提供される。データセグメントタイムスロット204は、64ポートのいずれかでパケットセグメントを受取るための時間である。よって、データセグメント時間204は、スイッチ100に接続されるポートごとの、セグメントバッファメモリ108にパケットセグメントを記憶するために利用可能である最大の時間である。30

【0037】

データセグメントタイムスロット204は、入力ポート112でデータが受取られる速度と、データ送信のために実現されるネットワークプロトコルとに依存する。たとえば、もし用いられるネットワークプロトコルが100メガヘルツイーサネット(R)であれば、1ビットは各ポートで10nsごとに受取られ、1バイトは80nsごとに受取られ、最小パケットサイズである64バイトは5120nsごとに受取られる。こうして、80nsのポートサイクル時間で、最小パケットサイズを、スイッチ100(図1A)に接続される各64 100MHzイーサネット(R)ポートで5120nsごとに記憶することができる。40

【0038】

各ポートは、セグメントバッファメモリ108にアクセスるべきポートサイクル時間208を割当てられる。データセグメントは、ポートサイクル時間208ごとにセグメントバッファメモリ108における異なったアドレスに読出および書込を行なうことができる50

。ポートサイクル時間 208 は、データセグメントタイムスロット 204 およびスイッチ 100 に接続されるポートの数に依存する。

【0039】

時間 700において、データセグメントタイムスロット 208 の開始で、フレームパルスはフレームパルス信号を発生させることにより生成される。フレームパルス信号は、ポートタイミング論理 304 (図 3) を初期化するために用い得る。700 で、ポート 0 に対するポートサイクル 208 が開始する。入力ポート 0 に対する書込アドレスは、バッファセグメントメモリ 108A に発信され得る一方、出力ポート 0 に対する読出アドレスはバッファセグメントメモリの B バンク 108B に送信され得る。

【0040】

時間 702 で、ポート 1 に対するポートサイクル 208 が開始する。入力ポート 1 に対する読出アドレスは、バッファセグメントメモリ 108A に発信され得る一方、出力ポート 1 に対する書込アドレスはバッファセグメントメモリの B バンク 108B に発信される。

【0041】

時間 704 で、データセグメントタイムスロット 204 の終了時に、フレームパルス信号を発生させることによりフレームパルスが生成される。ポートサイクルは、データセグメントタイムスロット 204 の間にスイッチに接続される 64 ポートごとに生成されている。

【0042】

図 8 は、図 4 に示される 2 つのバンクのセグメントバッファメモリ 108A、108B に対する読出および書込アドレスを選択するためのパケットストレージマネージャ 106 によって実行されるステップを示すフロー図である。図 7 に示されるようにポートサイクル 208 ごとに、読出および書込アドレスは同じポートに対するものであり、すなわち、入力ポート 102 および出力ポート 112 は同じポートである。しかしながら、入力ポート 102 および出力ポート 112 は示されるように同じポートである必要はなく、読出アドレスおよび書込アドレスは異なったポートに対するものであってもよい。図 8 は、図 4 から図 7 に関連して説明される。

【0043】

ステップ 800において、二バンク書込選択論理 600 は、出力ポート 112 に対して保留された読出があるかどうかを判断する。もし保留された読出があれば、処理はステップ 802 に進む。もし保留された読出がなければ、処理はステップ 822 に続く。

【0044】

ステップ 802 において、二バンク書込選択論理 600 は、読出アドレスがどちらのセグメントバッファメモリのバンク 108A、108B に発信されるべきかをメモリアクセス論理 312 により判断する。もし読出アドレス 310 がバンク_A 108A に対するものであれば、処理はステップ 812 に進む。もし読出アドレス 310 がバンク_B 108B に対するものであれば、処理はステップ 804 に続く。

【0045】

ステップ 804 において、二バンク書込選択論理 600 がポートサイクル 314 に依存して保留された書込があるかどうかを判断する。もしポートに対して保留された書込があれば、処理はステップ 810 に進む。もしなければ、プロセスはステップ 806 に進む。

【0046】

ステップ 810 において、二バンク書込選択論理 600 は、バンク_A に対する空きリスト 606 から、セグメントバッファメモリのバンク_A 108A 内の利用可能な場所のアドレスを抽出する。バンク_A に対する空きリスト 606 からのバンク_A アドレス 612 は、書込アドレス選択論理 602 によって書込アドレス 308 として選択される。バンク_A に対する空きリストカウント 604 はデクリメントされる。書込アドレス 308 および読出アドレス 310 は、メモリアクセス論理 312 によってセグメントバッファメモリのそれぞれのバンク 108A、108B に転送される。プロセスはステップ 808 に進む。

10

20

30

40

50

【0047】

ステップ806において、読出アドレスがメモリアクセス論理312によってセグメントバッファメモリのバンク_B 108Bに転送される。処理はステップ808に進む。

【0048】

ステップ808において、読出動作が完了した後で、コピーカウント論理624によって読出アドレスに対するコピーカウント626がデクリメントされる。読出アドレス310はバンク_Bに対する空きリスト608に返され、コピーカウント626が0であればバンク_Bに対する空きリストカウント610がインクリメントされる。処理はステップ820に進む。

【0049】

ステップ822において、書込選択論理600はポートに対して保留された書込があるかどうかを判断する。もしそのポートに対して書込が保留されていれば、処理はステップ824に続く。そうでなければ、処理はステップ820に続く。

【0050】

ステップ824において、書込選択論理600はセグメントバッファメモリのどちらのバンク108A、108Bが最も空いているかを、バンク_Aに対する空きリストカウント604に記憶されているカウントおよびバンク_Bに対する空きリストカウント610に記憶されているカウントから判断する。書込選択論理がセグメントバッファメモリのバンク_A 108Aが最も空いていると判断すると、処理はステップ828に続く。二バンク書込選択論理600がセグメントバッファメモリのバンク_B 108Bが最も空いていると判断すると、処理はステップ830に続く。

【0051】

ステップ828において、書込選択論理600は、バンク_Aに対する空きリスト606からセグメントバッファメモリのバンク_A 108Aにおける利用可能な場所のアドレスを除去する。バンク_Aに対する空きリスト606からのバンク_Aアドレス612は、書込アドレス選択論理602によって書込アドレス308として選択される。バンク_Aアドレス612がバンク_Aに対する空きリスト606から除去されると、バンク_Aに対する空きリストカウント604がデクリメントされる。書込アドレス308は、メモリアクセス論理312(図3)によってバッファセグメントメモリのバンク_A 108Aに転送される。

【0052】

ステップ830において、書込選択論理600は、バンク_Bに対する空きリスト608から、セグメントバッファメモリのバンク_B 108Bの利用可能な場所のアドレスを除去する。バンク_Bに対する空きリスト608からのバンク_Bアドレス614は、書込アドレス選択論理602によって書込アドレス308として選択される。バンク_Bに対する空きリスト608からバンク_Bアドレス614が除去されると、バンク_Bに対する空きリストカウント610はデクリメントされる。書込アドレス308は、アドレス選択論理312(図3)によってバッファセグメントメモリのバンク_B 108Bに転送される。

【0053】

ステップ812において、書込論理600は、ポートに対して保留された書込動作があるかどうかを判断する。もし保留された書込があれば、処理はステップ816に進む。なければ、処理はステップ814に進む。

【0054】

ステップ816において、書込選択論理600は、バンク_Bに対する空きリスト608から、セグメントバッファメモリのバンク_B 108Bの利用可能な場所のアドレスを除去する。バンク_Bに対する空きリスト608からのバンク_Bアドレス614は、書込アドレス選択論理602によって書込アドレス308として選択される。バンク_Bに対する空きリスト608からバンク_Bアドレス614が除去されると、バンク_Bに対する空きリストカウント610がデクリメントされる。書込アドレス308および読出ア

10

20

30

40

50

レス310は、メモリアクセス論理312(図3)によってバッファセグメントメモリのそれぞれのバンク108A、108Bに転送される。処理はステップ818に進む。

【0055】

ステップ814において、読出アドレス310は、アドレス選択論理312(図3)によってセグメントバッファメモリのバンク_A_108Aに転送される。処理はステップ818に続く。

【0056】

ステップ818では、読出動作が終了すると、読出アドレスのコピーカウントはコピーカウント論理624によって減分される。読出アドレス310はバンク_Aに対する空きリスト606に返され、バンク_Aの空きリストカウント604はコピーカウントがゼロであるのなら増分される。処理はステップ820に続く。

10

【0057】

ステップ820では、ポートタイミング論理304(図2)は次のポートサイクル314に増分される。処理は次のポートサイクルに対してステップ800で続く。

【0058】

図9は図1Aに示されるセグメントバッファメモリを示すブロック図であり、セグメントバッファメモリは物理的に4個のバンク、すなわちセグメントバッファメモリ(バンク_A(偶数))(Ae)108Ae、セグメントバッファメモリ(バンク_A(奇数))(Ao)108Ao、セグメントバッファメモリ(バンクB(偶数))(Be)108Be、およびセグメントバッファメモリ(バンクB(奇数))(Bo)108Boに分けられる。メモリアドレスの2つのMSBに依存して、たとえばAまたはBのバンクを選択するMSBビットおよび奇数または偶数バンクを選択するMSB-1メモリアドレスに依存して、メモリは物理的に4個のバンクに分けられる。

20

【0059】

セグメントバッファメモリ108の各バンクは関連するメモリアドレス、すなわちアドレス_Ae_900、アドレス_Ao_902、アドレス_Be_904、およびアドレス_Bo_906を有する。メモリアクセス論理12は、書込アドレス308に符号化されるセグメントバッファメモリ108のバンクおよびポートサイクルに依存して、書込アドレス308をアドレス_Ae_900、アドレス_Ao_902、アドレス_Be_904またはアドレス_Bo_906に送る。メモリアドレス論理312は、読出アドレス310に符号化されるセグメントバッファメモリ108のバンクおよびポートサイクルに依存して、読出アドレス310をアドレス_Ae_900、アドレス_Ao_902、アドレス_Be_904またはアドレス_Bo_906に送る。

30

【0060】

データセグメント時間当たり各ポートに与えられるメモリバンド幅は、各データセグメント時間において各ポートに対して2つのポートサイクル314、すなわち偶数ポートサイクルおよび奇数ポートサイクルを実施することにより、2倍に増加する。偶数バンク108Aeおよび108Beは偶数サイクルでアクセスされ、奇数バンク108Aoおよび108Boは奇数サイクルでアクセスされる。書込アドレス論理300(図2)および読出アドレス論理304(図2)は、各ポートサイクルで発行された読出アドレス308および書込アドレス310がバッファセグメントメモリ108Ae、108Ao、108Be、および108Boの異なるバンク用であることを確実にする。したがって、偶数サイクルの際、バンク_Ae_108Aeへの読出アクセスは、バンク_Be_108Bへの書込アクセスと並行に起こる。

40

【0061】

データセグメント時間ごとに、各ポートに対して2つのポートサイクルを与え、かつ各ポートサイクルに並行な読出および書込アクセスを与えることにより、セグメントバッファメモリ108のバンド幅は、スイッチ100に接続される各ポートに対して4倍になる。したがって、各データセグメント時間において、各ポートに対して4回セグメントバッファメモリ108をアクセスすることができる。メモリアクセスは対で与えられ、1つの偶

50

数バンクの読出動作がほかの偶数バンクへの書込動作と並行に行なわれる。

【0062】

図10は図9に示される四バンクセグメントバッファメモリ108に対する、図3に示される読出アドレス論理302を示すブロック図である。読出アクセス論理302(図3)は偶数/奇数ポート待ち行列選択論理1000、各出力ポート112に対するポート待ち行列510、および偶数/奇数読出アドレス選択論理1004を含む。四バンクセグメントバッファメモリ108に対して、ポートタイミング論理304(図3)は所定の数のクロック周期に偶数ポートサイクル106および奇数ポートサイクル108をそれぞれ生成し、各データセグメントタイムスロット204(図2)において偶数/奇数信号1010を生成する。偶数ポートサイクル1006および奇数ポートサイクル1008のタイミング図の例は図10と関連して説明される。
10

【0063】

ポート待ち行列510は既に図5と関連して記載されている。偶数/奇数ポート待ち行列選択論理1000は、奇数ポートサイクル1006および偶数ポートサイクル1008の状態ならびにポートタイミング論理304(図3)からの偶数/奇数信号1010に依存して、読出すべきポート待ち行列510を選択する。ポートに対して奇数ポートサイクル1006または偶数ポートサイクル1008のどちらかが活性状態の場合、ポートに対して読出すべき次のデータセグメントの場所504はポート待ち行列510から選択され、偶数/奇数読出アドレス選択論理1004に送られる。
20

【0064】

偶数/奇数読出アドレス場所選択論理1004は偶数ポートサイクル1006および奇数ポートサイクル1008の状態に依存して、ポート待ち行列510によって与えられるアドレス504を送る。場所504が奇数であり、かつ奇数ポートサイクル1008が活性状態にある場合、場所504は偶数/奇数読出アドレス選択論理1004に送られる。場所504が奇数でありかつ偶数ポートサイクル1006が活性状態にある場合、場所504は偶数/奇数読出アドレス選択論理1004に送られないが、場所504が偶数であるのなら場所504は送られる。したがって、データセグメント時間において、データセグメントがセグメントバッファメモリ108の奇数および偶数バンクに記憶されているのなら、ポートはセグメントバッファメモリ108からデータセグメントの2つの場所を読出し得る。
30

【0065】

図11は、図9に示される四バンクセグメントバッファメモリ108のための、図2に示される書込アドレス論理を示すブロック図である。四バンク書込選択論理1100はセグメントバッファメモリ108Ao、108Ae、108Boおよび108Beの4つのバンクの1つに対する書込アドレス308を選択する。バッファセグメントメモリ108Ao、108Ae、108Boおよび108Beの各バンクは、関連する空きリスト1112、1114、1116および1118ならびに空きリストカウント1102、1104、1106および1108を有する。

【0066】

四バンク書込選択論理1100は、読出アドレス310および偶数/奇数信号910に依存して、書込アドレスを抽出すべき空きリスト1112、1114、1116および1118を定める。四バンク書込選択論理1100は、偶数/奇数信号910からメモリアクセスが偶数サイクル用でありかつ読出アドレス310はバッファセグメントメモリ108AeのAeバンク用であると定めたなら、書込アドレスはBeバンクの空きリスト1106から抽出される。四バンク書込選択論理1100は、偶数/奇数信号910からメモリアクセスは偶数サイクル用であり、かつ読出アドレス310はバッファセグメントメモリ308BeのBeバンク用であると定めたなら、書込アドレスはAeバンクの空きリスト1112から抽出される。
40

【0067】

四バンク書込選択論理1100は読出アドレスがバッファセグメントメモリ108Aeの
50

A e バンク用であると定めるのなら、E N A _ 3 信号 1 1 3 4 を通して B e バンクの空きリスト 1 1 1 6 を能動化し、それにより B e バンクの空きリスト 1 1 1 6 は四バンク書込アドレス選択論理 1 1 2 0 に対して B e バンクアドレス 1 1 2 6 を与えることができる。B e バンク用の空きリストからアドレスが取除かれると、四バンク書込選択論理 1 1 0 0 は獲得 / 更新カウント B e 信号 1 1 4 0 に更新カウント B e 信号を生成して、バンク _ B e の空きリストカウント 1 1 4 0 を減分することにより、B e バンクの空きリスト 1 1 1 6 を更新する。

【 0 0 6 8 】

有効な読出アドレスがない場合、四バンク書込選択論理 1 1 0 0 は空きリストカウント 1 1 0 2 、 1 1 0 4 、 1 1 0 6 、および 1 1 0 8 に依存して、書込アドレス 3 0 8 を抽出するべき空きリスト 1 1 1 2 、 1 1 1 4 、 1 1 1 6 および 1 1 1 8 を定める。最も大きい値を有する空きリストカウント 1 1 0 2 、 1 1 0 4 、 1 1 0 6 および 1 1 0 8 は、最大の利用可能な場所を有するセグメントバッファメモリ 1 0 8 A e 、 1 0 8 A o 、 1 0 8 B e 、 1 0 8 B o のバンクを示す。空きリストカウント 1 1 0 2 、 1 1 0 4 、 1 1 0 6 および 1 1 0 8 における最大の利用可能な場所に関連する空きリスト 1 1 1 2 、 1 1 1 4 、 1 1 1 6 および 1 1 1 8 は選択されて、書込まれるべきセグメントバッファメモリ 1 0 8 の次の場所を与える。こうして、データが書込まれる際に、バッファセグメントメモリ 1 0 8 A e 、 1 0 8 A o 、 1 0 8 B o および 1 0 8 B e のバンク間で均等に配分される。

【 0 0 6 9 】

たとえば、偶数のポートサイクルに対して読出アドレス 3 1 0 がなければ、四バンク書込選択論理 1 1 0 0 は獲得 / 更新カウント A e 信号 1 1 4 4 および獲得 / 更新カウント B e 1 1 4 0 信号を通して、A e バンク用の空きリストカウント 1 1 0 2 および B e バンク用の空きリストカウント 1 1 0 6 から現行のカウントを得る。バンク _ A e 用の空きリストカウント 1 1 0 2 およびバンク _ B e 用の空きリストカウント 1 1 0 6 は四バンク書込アドレス論理 1 1 0 0 に送られる。四バンク書込アドレス論理 1 0 0 0 はカウント値を比較し、バンク _ A e のカウント 1 1 0 2 およびバンク _ B e のカウント 1 1 0 6 のどちらが大きいかに依存して、E N A _ 1 信号 1 1 3 0 または E N A _ 3 信号 1 1 3 4 を生成する。

【 0 0 7 0 】

空きリスト書込アドレス 1 1 2 2 、 1 1 2 4 、 1 1 2 6 および 1 1 2 8 のうち 1 つが選択されると、書込アドレス 3 0 8 は四バンク書込アドレス選択論理 1 1 2 0 によってセグメントバッファメモリ 1 0 8 に送られる。コピーカウント 6 2 6 は書込アドレス 3 0 8 用のコピーカウント論理 6 2 4 に記憶される。コピーカウント 6 2 6 はマルチキャストパケットを管理するために用いられる、すなわちパケットが 2 つ以上の出力ポート 1 1 2 に対して待ち行列に入れられる。

【 0 0 7 1 】

図 1 2 は図 9 に示される四バンクセグメントバッファメモリ 1 0 8 の読出および書込アドレスのタイミングを示すタイミング図である。各データセグメントタイムスロット 2 0 4 において、スイッチ 1 0 0 に接続される 6 4 個のポートの各々に対して、セグメントバッファメモリへのアクセスが 4 回与えられる。データセグメントタイムスロット 2 0 4 は 6 4 個のポートのうちのいずれかのポートでデータセグメントを受取る時間である。したがって、データセグメント時間 2 0 4 は、スイッチ 1 0 0 に接続される各ポートに対して、セグメントバッファメモリ 1 0 8 にデータセグメントを記憶するのに利用できる最長の時間である。

【 0 0 7 2 】

各ポートは、セグメントバッファメモリ 1 0 8 をアクセスするべきポートサイクル時間 2 0 8 が割当てられている。データセグメントはポートサイクル時間 2 0 8 の間に、ポート用のデータセグメントメモリのバンクに対して読出または書込される。ポートサイクル時間 2 0 8 はデータセグメントタイムスロット 2 0 4 およびスイッチ 1 0 0 に接続されるポートの数に依存する。セグメントバッファメモリ 1 0 8 (図 1 A) のアクセス時間はポート

10

20

30

40

50

トサイクル時間 208 に依存する。

【0073】

1204において、各データセグメントタイムスロット204が開始されるとき、フレームパルス信号をパルス化することによりフレームパルスが生成される。フレームパルス信号はポートタイミング論理304(図3)を初期化するために用いることができる。1206において、ポート0用の偶数サイクルが開始される。ポート0用の書込アドレスはバッファセグメントメモリ108AeのAeバンクに送られ、ポート0の読出アドレスはバッファセグメントメモリ108BeのBeバンクに送られる。時間1214において、ポート32の奇数サイクルが開始され、ポート32の書込アドレスはバッファセグメントメモリ108BoのAoバンクに送られ、ポート32の読出アドレスはバッファセグメントメモリ108BoのBoバンクに送られる。ポート32の奇数サイクルの開始1214は、ポート0の偶数サイクルの開始1206後、ポートサイクル時間208の半分の時間から始まる。この偶数サイクルの開始および奇数サイクルの開始間の遅れにより、奇数サイクルおよび偶数サイクルにおいてバッファセグメントメモリ108に共通論理を使用することができる。共通論理が用いられないのなら、奇数サイクルの開始および偶数サイクルの開始は同じ時間にスケジュールされ得る。

【0074】

1218において、ポート0の奇数サイクルの開始は、データセグメントタイムスロット204の半分の時間に、ポート0の偶数サイクルの開始からポートタイミングスロット208の半分の時間を加えた時間としてスケジュールされる。この時間は図12における遅延時間1220である。たとえば、データセグメントタイムスロット204が5120ns(64バイト、バイト当たり8ビット、バイト当たり10ns)であり、ポート0の偶数ポートサイクルの開始1206が0で始まるのなら、ポート0の奇数ポートサイクルの開始は、ポート0の偶数ポートサイクルの開始1206から2200ns(5120/2+40)にスケジュールされる。

【0075】

データセグメントタイムスロット204において、スイッチ100はデータセグメントサイズの倍数でないバイトカウントのデータパケットを受取ることができる。これらのデータパケットは最小のデータパケットとして受取られる時間とほぼ同じ時間をとる。たとえば、データセグメントタイムスロット204内で入力ポート0で受取られた65バイトのデータパケットを記憶するのに、ポート0の偶数サイクルの間最初の64バイトがセグメントバッファメモリ108に書込まれ、最後のバイトはポート0の奇数サイクルの際にセグメントバッファメモリ108に書込まれる。

【0076】

スイッチ100は1データセグメントより大きいデータパケットを受取ることができる。同じデータパケットのデータを受けながら、四バンク書込選択論理1100はパケット信号1138を通して、同じパケットの順次データセグメント用の書込アドレス208が交互の偶数および奇数空きリストから選択されることを確実にする。たとえば、最初のセグメントが偶数バンク_Bの空きリスト1116から選択された書込アドレス308に書込まれるのなら、次の順次セグメントはバンク_A 108Aoまたはバンク_B 108Boからの読出があるかどうかに依存して、奇数バンク_Bの空きリスト1108または奇数バンク_Aの空きリスト1114から選択された書込アドレス308に書込まれる。

【0077】

図13は図9に示される四バンクセグメントバッファメモリの偶数バンク用の読出および書込アドレスを選択するステップを示すフロー図である。

【0078】

図13は図9に示される四バンクセグメントバッファメモリの偶数バンク用の読出および書込アドレスを選択するステップを示すフロー図である。

セグメントバッファメモリ108Aeおよび108Beの偶数バンクは偶数ポートタイムスロットで選択される。同じステップは奇数ポートタイムスロットにおいてセグメントバ

10

20

30

40

50

ツファメモリ 108A₀ および 108A_e の奇数バンク用に読出および書込アドレスを選択するために行なわれる。図 12 に示されるように、偶数ポートタイムスロットの際の読出および書込動作は同じポートに対して発行される、すなわち入力ポート 102 および出力ポート 112 は同じポートである。入力ポート 102 および出力ポート 112 は示されているように同じポートである必要はなく、読出動作および書込動作は並行に異なるポートに対して発行され得る。図 13 は図 9 - 図 12 に関連して記載される。

【0079】

ステップ 1300において、4四バンク書込選択論理 1100 は出力ポート 112 用の保留読出を示す、読出選択論理 1004 からの有効な読出アドレス 310 があるかどうかを定める。保留の読出があるのなら、処理はステップ 1304 で続けられる。なければ、処理はステップ 1328 に続く。

【0080】

ステップ 1304において、四バンク読出選択論理 1004 はメモリアクセス論理 312 によって読出アドレスがセグメントバッファメモリ 108A_e および 108B_e のどの偶数バンクに発行されるのかを定める。読出アドレス 310 が偶数バンク_A 108A_e 用であるのなら、処理はステップ 108 に進む。読出アドレス 310 が偶数バンク_B 108B_e 用であるのなら、処理はステップ 1306 に進む。

【0081】

ステップ 1308において、四バンク書込選択論理 1100 は入力ポート 102 に対して保留の書込があるかどうかを定める。入力ポートに対して保留書込があるのなら、処理はステップ 1320 に進む。なければ、処理はステップ 1318 に進む。

【0082】

ステップ 1320において、四バンク書込選択論理 1100 はENA_3 信号 1134 を通して偶数バンク_B の空きリスト 1116 からセグメントバッファメモリ 108B_e の偶数バンク_B に対する利用可能なアドレスを抽出する。バンク_B の偶数バンク_B 空きリスト 1116 からの偶数バンク_B アドレス 1126 は四バンク書込アドレス選択論理 1120 によって書込アドレス 308 として選択される。偶数バンク_B 空きリストカウント 1106 は獲得 / 更新カウント B_e 信号 1140 によって減分される。書込アドレス 308 および読出アドレス 310 は四バンク読出 / 書込アドレス選択論理 908 (図 9) を通してセグメントバッファメモリ 108A_e および 108B_e のそれぞれの偶数バンクに送られる。処理はステップ 1324 に進む。

【0083】

ステップ 1318において、読出アドレス 310 は四バンク読出 / 書込アドレス選択論理 908 によって、セグメントバッファメモリ 108A_e の偶数バンク_A に送られる。処理はステップ 1324 に進む。

【0084】

ステップ 1324において、読出動作が終了すると、読出アドレス 310 は偶数バンク_A の空きリスト 1112 に返され、コピーカウント 626 (図 6) がゼロであれば偶数バンク_A の空きリストカウント 1102 は増分される。処理はステップ 1326 に進む。

【0085】

ステップ 1328において、四バンク書込選択論理 1100 はポートに対して書込の保留があるかどうか定める。ポートに対して書込保留があるのなら、処理はステップ 1330 に進む。なければ、処理はステップ 1326 に進む。

【0086】

ステップ 1330において、四バンク書込選択論理 1100 は、偶数バンク_A の空きリストカウント 1102 に記憶されるカウントおよび偶数バンク_B の空きリストカウント 1106 に記憶されるカウントから、バッファセグメントメモリ 108A_e および 108B_e のどの偶数バンクが最も空いているかを定める。書込選択論理はセグメントバッファメモリ 108A_e の偶数バンク_A が最も空いていることを定めると、処理はステップ 1334 に進む。四バンク書込選択論理 1100 はセグメントメモリ 108B_e の偶数バンク

10

20

30

40

50

ク__B が最も空いていると定めると、処理はステップ 1336 に続く。

【0087】

ステップ 1334 において、四バンク書込選択論理 1100 は偶数バンク__A の空きリスト 1112 から、セグメントバッファメモリ 108Ae の偶数バンク__A の利用可能な場所のアドレスを取除く。偶数バンク__A の空きリスト 1112 からのバンク__A アドレス 1122 が、四バンク書込アドレス選択論理 1120 によって書込アドレス 308 として選択される。偶数バンク__A アドレス 1122 が偶数バンク__A の空きリスト 1112 から取除かれる際、偶数バック__A の空きリストカウント 1102 は減分される。書込アドレス 308 は四バンク読出／書込アドレス選択論理 908 (図 9) によって、セグメントバッファセグメントメモリ 108Ae の偶数バンク__A に送られる。

10

【0088】

ステップ 1336 において、四バンク書込選択論理 1100 は偶数バンク__B の空きリスト 1116 から、セグメントバッファメモリ 108Be の偶数バンク__B から利用可能なアドレスを取除く。偶数バンク__B の空きリスト 1116 からの偶数バンク__B アドレス 1126 は、四バンク書込アドレス選択論理 1120 によって書込アドレス 308 として選択される。偶数バンク__B アドレス 1126 が偶数バンク__B の空きリスト 1106 から取除かれる際、偶数バンク__B の空きリストカウント 1106 が減分される。書込アドレス 308 は四バンク読出／書込アドレス選択論理 908 (図 9) によって、バッファセグメントメモリ 108Be の偶数バンク__B に送られる。

20

【0089】

ステップ 1306 において、四バンク書込選択論理 1100 はポートに対して書込動作が保留であるかどうか定める。書込保留があるのなら、処理はステップ 1316 に進む。なければ、処理はステップ 1312 に続く。

【0090】

ステップ 1316 において、四バンク書込選択論理 1100 は偶数バンク__A の空きリスト 1112 から、セグメントバッファメモリ 108Ae の偶数バンク__A における利用可能なアドレスを取除く。偶数バンク__B の空きリスト 1116 からの偶数バンク__A アドレス 1122 は、四バンク書込アドレス選択論理 1120 によって、書込アドレス 308 として選択される。偶数バンク__A アドレス 1122 が偶数バンク__A の空きリスト 1112 から取除かれる際、偶数バンク__A の空きリストカウント 1102 は減分される。書込アドレス 308 および読出アドレス 310 は、四バンク読出／書込アドレス選択論理 908 (図 9) によって、セグメントバッファメモリ 108Ae および 108Be のそれぞれのバンクに送られる。処理はステップ 1322 に進む。

30

【0091】

ステップ 1312 において、読出アドレス 310 は、四バンク読出／書込アドレス選択論理 908 (図 9) によって、セグメントバッファメモリ 108Be の偶数バンク__A に送られる。処理はステップ 1322 に進む。

【0092】

ステップ 1322 において、読出動作が終了すると、読出アドレス 310 は偶数バンク__A の空きリスト 1112 に返され、コピーカウント 626 (図 6) がゼロであれば偶数バンク__A の空きリストカウント 1102 が増分される。処理はステップ 1326 に進む。

40

【0093】

ステップ 1326 において、次の偶数ポートタイムスロットの処理が始まる。次の偶数ポートサイクルはポートタイミング論理 304 (図 3) によって定められる。

【0094】

本発明は、図 4 - 図 6 に関連して記載されている 2 つのバンクまたは図 9 - 図 13 に関連して記載される 4 つのバンクに分けられる、バッファセグメントメモリ 108 (図 1A) に限定されない。バッファセグメントメモリ 108 は 2^N (ここで $N = 1, 2, 3, 4 \dots$) 個の任意な数のバンクに分けられ得る。

【0095】

50

本発明は特定的に示されかつ好ましい実施例を参照して記載されているが、前掲の特許請求の範囲によって含まれる発明の範囲から逸脱することなく、形および詳細においてさまざまな変形が可能であることは理解されるであろう。

【図面の簡単な説明】

【図 1 A】 非ブロッキング共通メモリスイッチのブロック図である。

【図 1 B】 入力ポートで受取られるべき先行技術のイーサネット (R) データパケットのブロック図である。

【図 2】 スイッチ内の中通メモリによる入力ポートから出力ポートへの転送を示すタイミング図である。

【図 3】 図 1 A に示される、パケットストレージマネージャ 106 を示すブロック図である。 10

【図 4】 物理的に 2 つのバンクに分割された、図 1 A に示されるセグメントバッファメモリ 108 を示すブロック図である。

【図 5】 図 4 に示される二バンクセグメントバッファメモリに対する図 3 に示される読出アドレス論理 302 を示すブロック図である。

【図 6】 図 4 に示される二バンクセグメントバッファメモリに対する図 3 に示される書込アドレス論理を示すブロック図である。

【図 7】 図 4 に示される二バンクセグメントバッファメモリに対する読出および書込アドレスのタイミングを示すタイミング図である。

【図 8】 図 4 に示される二バンクセグメントバッファメモリに対する読出および書込アドレスを選択するためのステップを示すフロー図である。 20

【図 9】 物理的に 4 つのバンクに分割された、図 1 A に示されるセグメントバッファメモリを示すブロック図である。

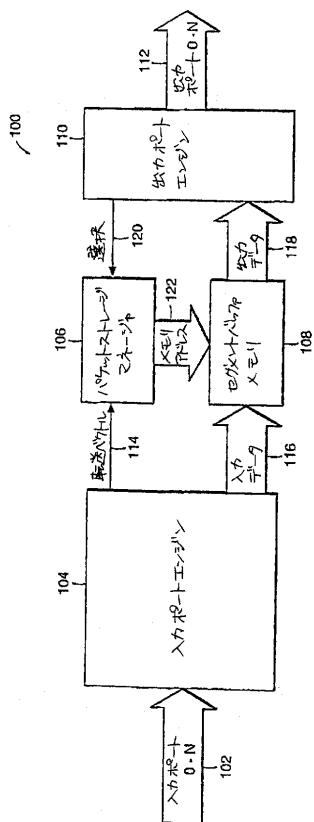
【図 10】 図 9 に示される四バンクセグメントバッファメモリに対する図 3 に示される読出アドレス論理を示すブロック図である。

【図 11】 図 9 に示される四バンクセグメントバッファメモリに対する図 3 に示される書込アドレス論理を示すブロック図である。

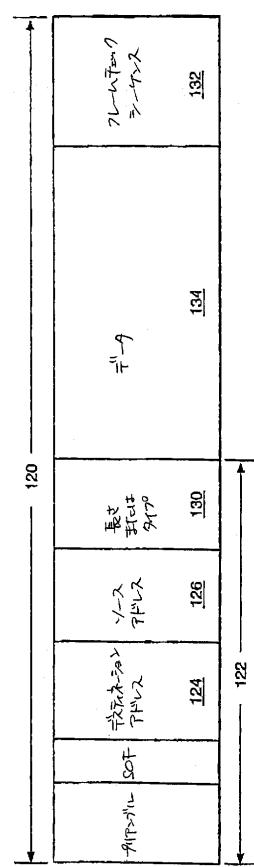
【図 12】 図 9 に示される四バンクセグメントバッファメモリに対する読出および書込アドレスのタイミングを示すタイミング図である。

【図 13】 図 9 に示される四バンクセグメントバッファメモリの偶数のバッファに対する読出および書込アドレスを選択するためのステップを示すフロー図である。 30

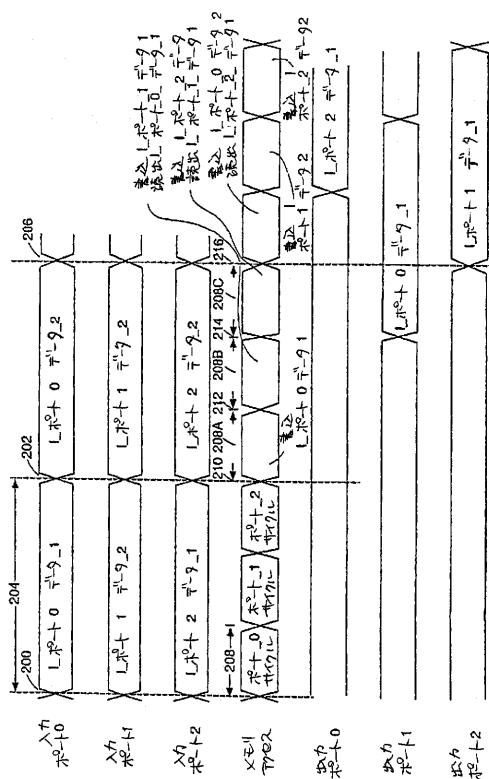
【図1A】



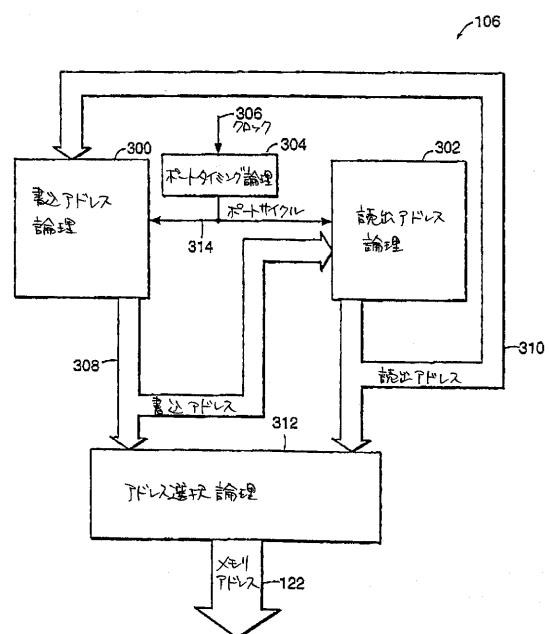
【図1B】



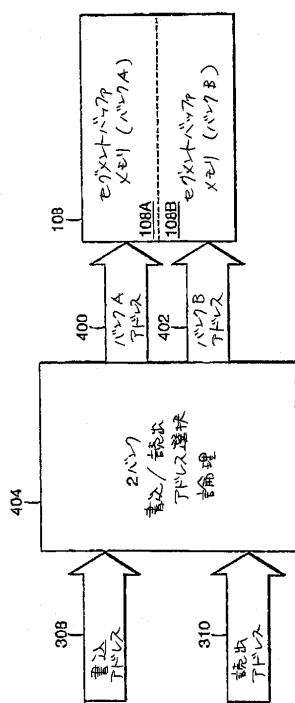
【図2】



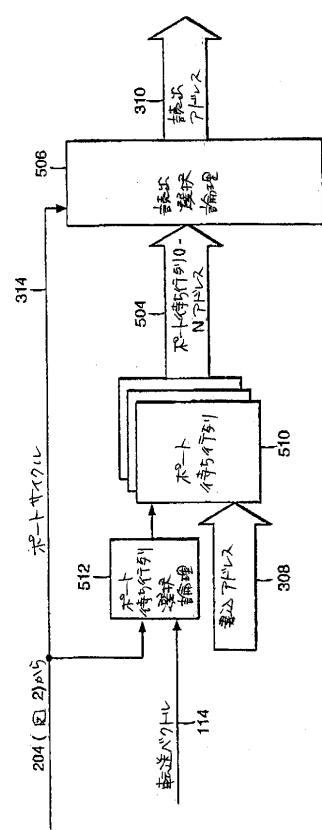
【図3】



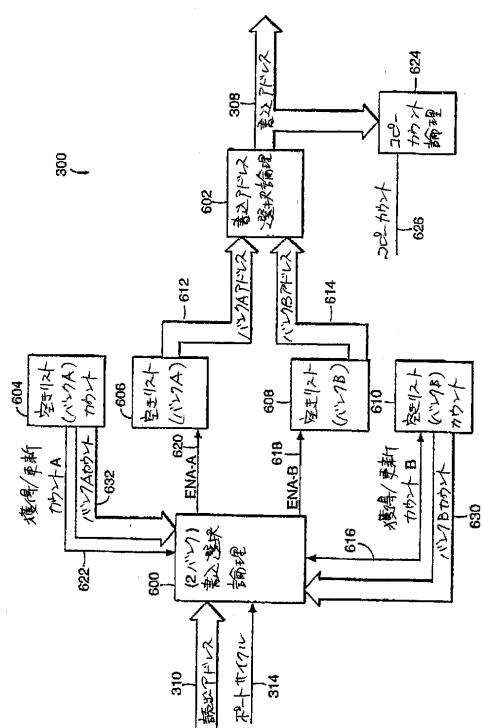
【図4】



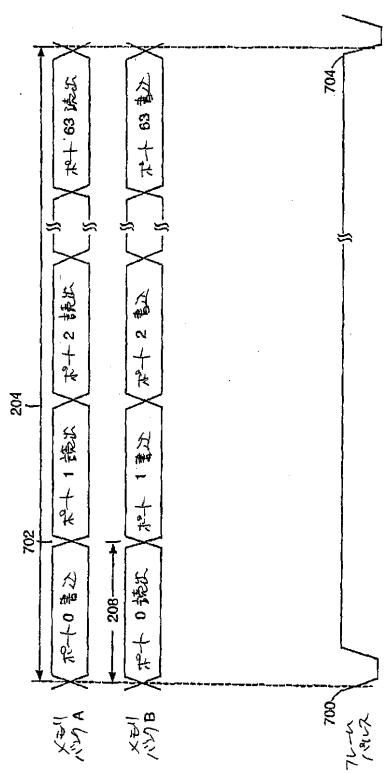
【図5】



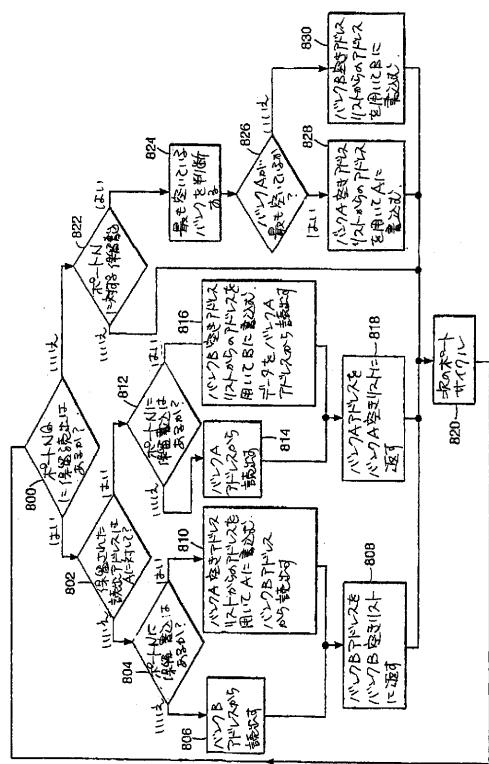
【図6】



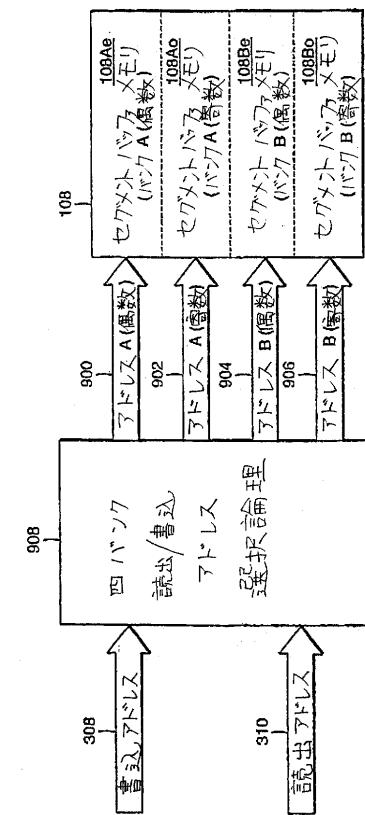
【図7】



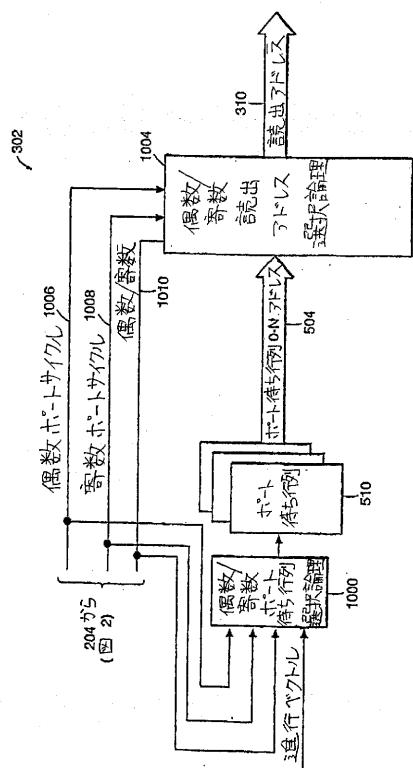
【 四 8 】



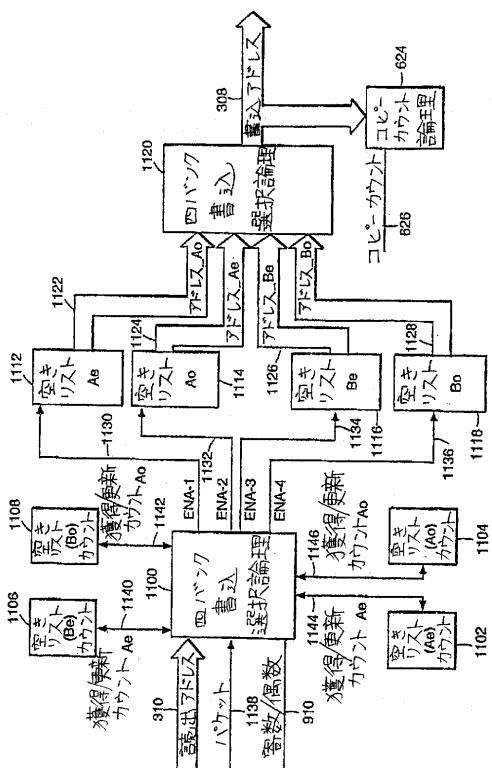
【 四 9 】



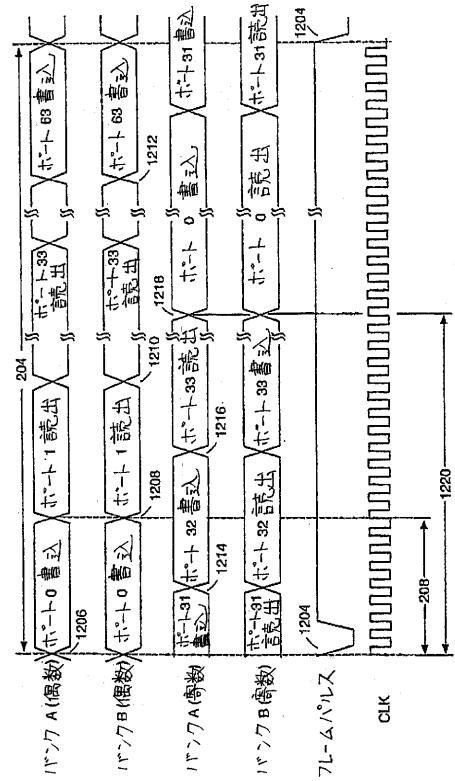
【図10】



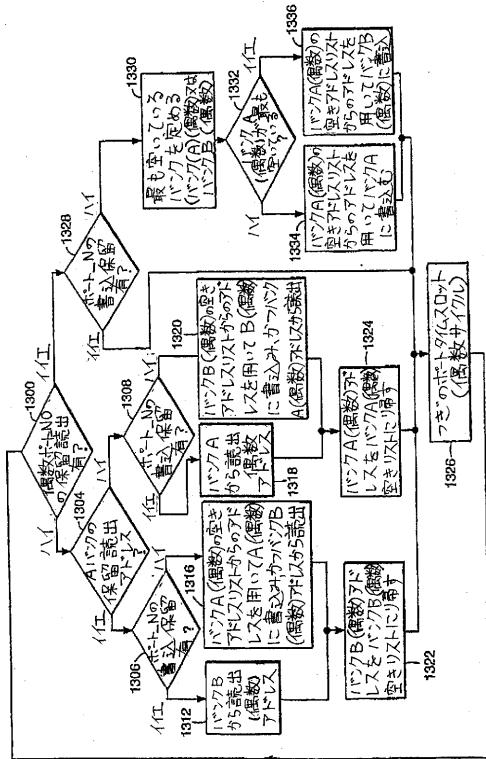
【 図 1 1 】



【 図 1 2 】



【 図 1 3 】



フロントページの続き

(74)代理人 100096781
弁理士 堀井 豊

(74)代理人 100096792
弁理士 森下 八郎

(72)発明者 ブラウン, デイビッド・エイ
カナダ、ケイ・0・エイ 1・エル・0 オンタリオ州、カーブ、アリーシャ・クレセント、11
0

審査官 矢頭 尚之

(56)参考文献 特開平11-17706 (JP, A)

(58)調査した分野(Int.Cl., DB名)
H04L 12/56