

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3987915号
(P3987915)

(45) 発行日 平成19年10月10日(2007.10.10)

(24) 登録日 平成19年7月27日(2007.7.27)

(51) Int. Cl. F I
H O 4 L 12/28 (2006.01) H O 4 L 12/28 2 O O M

請求項の数 31 (全 24 頁)

<p>(21) 出願番号 特願平10-535937 (86) (22) 出願日 平成10年2月9日(1998.2.9) (65) 公表番号 特表2001-511985(P2001-511985A) (43) 公表日 平成13年8月14日(2001.8.14) (86) 国際出願番号 PCT/US1998/002796 (87) 国際公開番号 W01998/036539 (87) 国際公開日 平成10年8月20日(1998.8.20) 審査請求日 平成17年1月13日(2005.1.13) (31) 優先権主張番号 60/038,025 (32) 優先日 平成9年2月14日(1997.2.14) (33) 優先権主張国 米国(US) (31) 優先権主張番号 08/992,425 (32) 優先日 平成9年12月18日(1997.12.18) (33) 優先権主張国 米国(US)</p>	<p>(73) 特許権者 アドバンスト・マイクロ・デバイス・ インコーポレイテッド アメリカ合衆国、 94088-3453 カリフォルニア州、 サニibel、 ピー・オー・ボックス・3453、 ワン ・エイ・エム・ディ・プレイス、 メール ・ストップ・68 (番地なし) (74) 代理人 弁理士 深見 久郎 (74) 代理人 弁理士 森田 俊雄 (74) 代理人 弁理士 伊藤 英彦</p>
---	---

最終頁に続く

(54) 【発明の名称】 ネットワークスイッチとホストコントローラとの間で送信する管理パケットを合成するための装置および方法

(57) 【特許請求の範囲】

【請求項1】

ネットワークステーションからデータパケットを受信するステップと、
 受信データパケットの受信にตอบสนองして新たな情報を生成するステップと、
 管理エージェントが用いるための、新たな情報および受信データパケットの少なくとも一部を含む新たなデータパケットを出力するステップとを含み、新たな情報は受信データパケットの受信に対応する特性を特定する、ネットワークにおける方法。

【請求項2】

生成するステップが、受信データパケットの受信にตอบสนองしてネットワークスイッチの特性を前記特性の少なくとも1つとして決定するステップを含み、前記生成するステップは決定されたネットワークスイッチ特性に基づいて新たな情報を生成する、請求項1に記載の方法。

【請求項3】

ネットワークスイッチが、データパケットを宛先ネットワークステーションに経路付けするためのスイッチ論理を含み、データパケットはネットワークステーションに対応するソースアドレスを有し、
 決定するステップが、スイッチ論理によってソースアドレスを認識する能力を決定するステップを含み、
 生成するステップが、スイッチ論理によって決定されたソースアドレスを認識する能力を特定するスイッチ論理データを生成するステップをさらに含む、請求項2に記載の方法。

【請求項 4】

スイッチ論理が既知のソースアドレスの1つと、スイッチ論理によって学習された未知のソースアドレスと、スイッチ論理によって学習されなかった未知のソースアドレスとを特定する、請求項 3 に記載の方法。

【請求項 5】

第 2 の生成するステップが、スイッチ論理によって学習された未知のソースアドレスに対応するアドレス情報のメモリ場所を特定するインデックスポインタを生成するステップを含む、請求項 4 に記載の方法。

【請求項 6】

決定するステップが、受信データパケットを受信しかつネットワークステーションに対応する受信バッファのオーバフロー状況を特定するオーバフローフラグを前記特性の1つとして生成するステップを含む、請求項 2 に記載の方法。

10

【請求項 7】

決定するステップが、受信データパケットでのエラーを明記するエラーフラグを前記特性の1つとして生成するステップを含む、請求項 2 に記載の方法。

【請求項 8】

生成するステップが、管理エージェントに対してデータパケットを送受信するように構成されるネットワークスイッチポートに対応する新たな宛先アドレスフィールドを生成するステップを含む、請求項 1 に記載の方法。

【請求項 9】

新たなデータパケットに対してエラーチェックコードを生成するステップをさらに含み、出力するステップが、新たな情報と、受信データパケットの少なくとも一部分と、エラーチェックコードとを含む新たなデータパケットを出力する、請求項 1 に記載の方法。

20

【請求項 10】

前記生成するステップが、受信データパケットの条件を前記特性の少なくとも1つとして決定するステップを含み、新たな情報が受信データパケットの決定された条件を含む、請求項 1 に記載の方法。

【請求項 11】

条件を決定するステップが、受信データパケットでのエラーの存在を決定するステップを含む、請求項 10 に記載の方法。

30

【請求項 12】

受信データパケットのフレームタイプを前記特性の少なくとも1つとして決定するステップをさらに含み、新たな情報が決定されたフレームタイプを含む、請求項 1 に記載の方法。

【請求項 13】

フレームタイプを決定するステップが、受信データパケットを、管理フレームと、管理エージェントおよびネットワークスイッチのポートのうち1つに対応する宛先アドレスを有する宛先指定フレームと、監視されるフレームとのうちの1つとして識別するステップを含む、請求項 12 に記載の方法。

【請求項 14】

受信データパケットがソースアドレスおよび宛先アドレスを含み、ネットワークステーションがスイッチ論理を含み、フレームタイプを決定するステップが、前記ソースアドレスおよび前記宛先アドレスのうち少なくとも1つに対応するスイッチ論理データに基づいて、受信データパケットを前記監視されるフレームとしてスイッチ論理によって識別するステップをさらに含み、請求項 13 に記載の方法。

40

【請求項 15】

データパケットがソースアドレスおよび宛先アドレスを含み、ネットワークステーションがスイッチ論理と、データパケットをネットワークステーションから受信するソースポートと、宛先ポートとを含み、スイッチ論理は宛先アドレスに基

50

づいてネットワークポートの1つを宛先ポートとして選択するように構成され、フレームタイプを決定するステップが、前記ソースポートおよび前記宛先ポートの少なくとも1つに基づいてデータパケットを前記監視されたフレームとして特定するステップをさらに含む、請求項13に記載の方法。

【請求項16】

管理フレームがブリッジプロトコルデータユニット(BPDU)およびブリッジマルチキャストフレームのうちの1つである、請求項13に記載の方法。

【請求項17】

生成するステップが、新規データパケットが受信データパケットの一部分のみを含むことを特定するデータを新たな情報内に生成するステップを含む、請求項1に記載の方法。

10

【請求項18】

ネットワークスイッチがそれぞれのネットワークステーションに対してデータパケットを授受する複数のネットワークポートを有する請求項1に記載の方法であって、前記方法がネットワークデータパケットおよびポートベクタを含む管理データパケットを管理エージェントから受信するステップをさらに含み、ポートベクタがネットワークデータパケットを送信するために前記ネットワークポートのグループを特定する、請求項1に記載の方法。

【請求項19】

それぞれのネットワークステーションに対してデータフレームを送受信するように構成された複数のネットワークポートと、

20

管理データフレームを管理エージェントに送信するように構成された管理ポートとを含み、管理データフレームは、選択された受信データフレームの少なくとも一部分と、選択されたデータフレームの受信に対応する新たな情報を特定する特性とを含み、さらに、対応する受信データフレームについての新たな情報の少なくとも一部分を生成するように構成されたスイッチングサブシステムを含む、ネットワークスイッチ。

【請求項20】

スイッチングサブシステムが、受信データフレームの1つを選択された受信データフレームとして選択し、スイッチ論理データを新たな情報の少なくとも一部分として生成するためのスイッチ論理を含む、請求項19に記載のスイッチ。

【請求項21】

30

前記ネットワークポートの各々および前記管理ポートが、対応するデータフレームの受信に応答して受信ステータスフィールドを各々生成し、スイッチ論理が対応の選択された受信データフレームに対する受信ステータスフィールドを前記管理データフレームに供給して送信する、請求項20に記載のスイッチ。

【請求項22】

受信ステータスが、選択された受信データフレームでのエラーおよび対応するネットワークポートでの受信バッファのオーバフローのうち少なくとも1つを特定する、請求項21に記載のスイッチ。

【請求項23】

スイッチ論理が、選択された受信データフレーム内の認識されないソースアドレスを検出することに応答して選択された受信データフレームを選択し、スイッチ論理が認識されないソースアドレスの検出を示すスイッチ論理データを生成する、請求項20に記載のスイッチ。

40

【請求項24】

スイッチ論理によって生成されたスイッチ論理データが、認識されないソースアドレスが新規に認識されたソースアドレスとして学習されたかをさらに明記する、請求項23に記載のスイッチ。

【請求項25】

スイッチ論理が前記ネットワークポートの予め定められたものの受信に基づいて選択された受信データフレームを選択し、スイッチ論理データが選択された受信データフレームを

50

予め定められた1つのポートに受信されたと特定する、請求項20に記載のスイッチ。

【請求項26】

スイッチ論理が、前記ネットワークポートの予め定められた1つを選択された受信データポートの宛先ポートとして特定することに基づいて選択された受信データフレームを選択し、スイッチ論理データが予め定められたあるネットワークポートを選択された受信データフレームの宛先ポートとして特定する、請求項20に記載のスイッチ。

【請求項27】

スイッチ論理が、選択された受信データフレーム内のソースアドレスおよび宛先アドレスの予め定められた1つの検出に基づいて選択された受信データフレームを選択し、スイッチ論理データが選択された受信データフレームを予め定められたあるアドレスを有するものとして特定する、請求項20に記載のスイッチ。

10

【請求項28】

スイッチ論理が選択された受信データフレームを非ネットワークステーションフレームとして特定したに基づいてそれを選択し、スイッチ論理データが選択された受信データフレームを非ネットワークステーションフレームとして特定する、請求項20に記載のスイッチ。

【請求項29】

管理ポートが管理エージェントから第2の管理フレームを受信し、第2の管理フレームがネットワークデータパケットと、ネットワークデータパケットを送信するためのネットワークポートを特定するポートベクタを含む、請求項20に記載のスイッチ。

20

【請求項30】

管理ポートが、管理データフレームのヘッダの一部として新たな情報を出力する、請求項19に記載のスイッチ。

【請求項31】

新たな情報が、管理データフレームが選択された受信データフレームの全体を含むことを特定する、請求項30に記載のスイッチ。

【発明の詳細な説明】

関連出願

この出願は、1997年2月14日に出願され、「統合マルチポートスイッチ」と題された仮特許出願第60/038,025号(代理人管理番号1033-230PRO)から優先権を主張し、これをここに引用として援用する。

30

技術分野

この発明はネットワークスイッチングに関し、より特定的には、ネットワークスイッチとネットワークスイッチを制御するホストコンピュータとの間に管理データを提供する方法および装置に関する。

背景技術

スイッチトローカルエリアネットワークは、ネットワークステーション間でデータを供給するためにネットワークスイッチを用い、各ネットワークステーションは媒体によってネットワークスイッチに接続される。このスイッチトローカルエリアネットワークアーキテクチャは、ネットワークインタフェースカードが媒体にアクセスできるようにする媒体アクセス制御(MAC)を用いる。ネットワークスイッチは、送信ステーションから受信したデータフレームを、受信データフレームのヘッダ情報に基づいて宛先ステーションへ伝達する。

40

ネットワークは典型的には、ハブまたはリピータによってアクティビティを監視する管理エージェントを含むであろう。管理エージェントは、分析および/または診断装置を含んでもよく、ネットワークがいかにうまく利用されているかを特定するためにネットワークに関する統計的情報にも関係する。このような統計的情報は、どのようなタイプのパケットがネットワーク上にあるか、どのようなプロトコルがネットワーク上で用いられているか、パケット送信者のアイデンティティ、パケット受信者のアイデンティティ、およびネットワーク上を転送されているパケット長の分布を含み得る。

50

ネットワークリピータは典型的には、管理情報をM A C層プロトコルの管理エージェントへ送信するであろう。たとえば、同一譲受人に譲渡された米国特許第5,592,486号は、受信データパケットの宛先アドレスを管理ユニットアドレスに対応する宛先アドレスと比較する、リピータを開示する。データパケットは、その宛先アドレスが、格納されている管理ユニットアドレスと整合しない場合は圧縮され、ここで管理ユニットへのデータパケットの送信は計数されたバイト数が一旦予め定められた数と整合すると停止する。その後有効なフレームチェックシーケンス(F C S)が、受信データパケットの送信に続いて、圧縮されたデータパケットの終端に付される。

同一譲受人に譲渡された米国特許第5,550,803号もまた、管理ユニットへの受信データパケットのM A C層送信を提供し、ここでは受信データパケットのデータ部分はリピータによって管理ユニットへ転送される。統計的情報はそのとき、受信データパケットの送信に続いてパケット間ギャップ期間中にデータ部分に添付される。

上記に開示された構成は、リピータ装置を有するネットワークシステムにとっては有益である。というのは、リピータ装置はすべてのネットワーク装置と共有する媒体を有し、イーサネット(I E E E 802.3)プロトコルに従って一度に単一データパケットを送信することを要求するからである。しかしながら、ネットワークスイッチを用いる切換ネットワークアーキテクチャは、それぞれのネットワークステーションのためのネットワークポートで複数のデータパケットを同時に送受信するように構成される。したがって、規定されたバイト数の後、受信データパケットの送信を停止し、それに続いて受信データパケットの受信が完了した後付加的データを添付する技術によると、送信の停止と付加的データの添付との間のアイドル時間中に無駄な帯域が生じてしまう。さらに、この技術は、ネットワークスイッチによって同時に受信された複数のデータパケットは想定していない。データパケットの終端に統計的情報を添付するにはまた、管理エージェントは全パケットが受信されるまで待機しなければデータパケットの性質(すなわち、なぜデータパケットが受信されたか)を判断できない。最後に、データパケットの終端にタグとして付された(すなわち添付された)統計的情報は、一般に、実際に受信されたデータパケットに関連する統計しか含んでいない。

発明の概要

それぞれのネットワークステーション間でデータパケットを同時に送受信するネットワークスイッチ内の管理媒体アクセス制御(M A C)ポートを介して、管理情報をホスト管理エージェントに提供する構成が必要である。

また、管理エージェントに送信されるべき管理データフレームを生成し、管理フレームが少なくとも選択された受信データフレームの一部分と、選択されたデータフレームの受信に対応する特性を特定する新しい情報とを含むような構成も必要である。

また、管理エージェントのための管理データフレームを生成し、その管理データフレームがデータパケットを受信するネットワークスイッチの特性に対応する新しい情報を含むような配置も必要である。

また、データパケットを受信し、データパケットが受信された時間とは独立に、受信データパケットの受信に対応する特性を特定する新しい情報を含む新しいデータパケットを生成するネットワークスイッチ内の構成も必要である。

これらおよび他の要件は、データパケットを受信するネットワークスイッチが管理エージェントへ送信するために新しいデータパケットを生成し、この新しいデータパケットが、受信データパケットの受信に対応する特性を特定する新しい情報を含む、この発明によって満たされる。

この発明の一局面に従うと、ネットワークスイッチにおける方法は、ネットワークステーションからデータパケットを受信するステップと、受信データパケットの受信に回答して新しい情報を生成するステップと、新しい情報と受信データパケットの少なくとも一部分とを含む新しいデータパケットを、管理エージェントが使用するために出力するステップとを含み、新しい情報は、受信データパケットの受信に対応する特性を特定する。受信データパケットの受信に回答して生成される新しい情報は、データパケットの受信について

10

20

30

40

50

の情報を含むこともあり、また受信データパケットを受信中のネットワークスイッチに関連する特徴を含んでもよい。さらに、新しいデータパケットの生成および出力はネットワークスイッチが受信データパケットを受信する時間とは独立で、これによりネットワークスイッチは、ネットワークステーションがそれぞれのネットワークステーションから複数のデータパケットを同時に送受信しながら、この新しいデータパケットを管理エージェントへ出力することができるようになる。したがって、管理エージェントは、ネットワークスイッチおよび受信データパケットの受信ステータスに関連する管理情報がのった新しい管理データパケットの流れを、スイッチとそれぞれのネットワークステーションとの間のネットワークトラヒック内に何ら中断を起こすことなく、得ることができる。

この発明の別の局面は、それぞれのネットワークステーションに対してデータフレームを送受信するように構成された複数のネットワークポートと、管理エージェントへ管理データフレームを送信するように構成された管理ポートを含むネットワークスイッチであって、管理データフレームが、選択された受信データフレームの少なくとも一部分および選択されたデータフレームの受信に対応する特性を特定する新しい情報を含み、このネットワークスイッチはさらに、対応する選択された受信データフレームについての新しい情報の少なくとも一部分を生成するように構成されるスイッチングサブシステムを含む、ネットワークスイッチを提供する。スイッチングサブシステムは、選択された受信データフレームを識別し、選択された受信データフレームの特性を特定する新たな情報を生成し、これは管理ポートによって、新たな情報および選択された受信データフレームの少なくとも一部分を含む新たな管理データフレームを生成するために用いられる。したがって、ネットワークスイッチは、所定の条件に基づいて、選択された受信データフレームの識別を可能にし、受信ステータス情報および/またはデータフレームの受信に応答するスイッチの特性を含む、選択されたデータフレームの受信に対応する特性を特定する新規情報を含む、管理データフレームを出力する。

この発明のさらなる目的、利点、および新規な特徴は、一部は後に続く説明に記述され、一部は当業者がそれを実験することによって明らかになるか、またはこの発明の実施によって習得され得る。この発明の目的および利点は、特に添付の請求の範囲で指摘される手段および組合せによって実現かつ達成し得る。

【図面の簡単な説明】

添付の図面を参照して、同じ参照番号で指定される要素は、全図面を通して同様の要素を表す。

図1は、この発明の一実施例に従って構成されたパケット交換システムのブロック図である。

図2は、この発明の一実施例に従って構成され、図1のパケット交換システムに用いられるマルチポートスイッチのブロック図である。

図3は、この発明の一実施例に従って構成された、図2のマルチポートスイッチのスイッチングサブシステムの概略図である。

図4は、この発明の一実施例に従って構成された、図1の外部メモリのメモリマップ図である。

図5は、図4のバッファ領域に対するフレームバッファヘッダフォーマットを示す図である。

図6は、図3のルールチェッカによって生成されたポートベクタを示す図である。

図7は、図6の動作制御フィールドの構造を示す図である。

図8Aおよび図8Bは、それぞれ、図1のネットワークスイッチの管理ポートによって送信される管理パケットを示す図である。

図9Aおよび図9Bは、それぞれ、タグ付でないフレームおよびタグ付フレームフォーマットを有するネットワークデータパケットの構造を示す図である。

図10Aおよび図10Bは、この発明の一実施例に従った管理エージェントに対する管理データフレームを生成するための装置および方法を示す図である。

例示的な実施例の詳細な説明

10

20

30

40

50

イーサネット (IEEE 802.3) 網などのパケット交換ネットワークにおけるスイッチを例に挙げてこの発明を説明する。スイッチアーキテクチャの説明に続いて、この発明に従って管理フレームを生成するための構成について説明する。しかしながら、以下に詳細に説明するように、この発明は他のパケット交換システムおよび一般的な他のタイプのシステムにも適用可能であることが明らかとなるであろう。

スイッチアーキテクチャ

図1は、この発明が有利に採用され得る例示的なシステムのブロック図である。例示的なシステム10はイーサネット網などのパケット交換ネットワークである。パケット交換ネットワークは、ネットワークステーション間でのデータパケットの通信を可能にする統合マルチポートスイッチ (IMS) 12を含む。ネットワークはたとえば10 Mb/sのネットワークデータレートでデータの授受を行なう24個の毎秒10メガビットの速度 (Mb/s) のネットワークステーション14と、100 Mb/sのネットワーク速度でデータパケットの授受を行なう2つの100 Mb/sネットワークステーション16といった、種々の構成を有するネットワークステーションを含み得る。マルチポートスイッチ12はネットワークステーション14または16から受けたデータパケットをイーサネットプロトコルに基づく適切な宛先に選択的に転送する。

開示される実施例によると、10 Mb/sネットワークステーション14は媒体18を介して、かつ半二重イーサネットプロトコルに従って、マルチポートスイッチ12に対してデータパケットの授受を行なう。イーサネットプロトコルISO/IEC 8802-3 (ANSI/IEEE Std. 802.3, 1993 Ed.) は、すべてのステーション14が等しくネットワークチャネルにアクセスできるようにする半二重媒体アクセス機構を規定する。半二重環境のトラヒックは媒体18と区別されたりまたはそれより優先されることはない。各ステーション14はむしろ、媒体上のトラヒックを認識するために搬送波感知多重アクセス/衝突検出 (CSMA/CD) を用いるイーサネットインタフェースカードを含む。媒体上の受信搬送波がデアサートされたことを感知することによりネットワークトラヒックの不在が検出される。送信するデータを有するステーション14はすべて、パケット間ギャップ期間 (IPG) として公知である、媒体上の受信搬送波がデアサートされた後、予め定められた時間だけ待機することにより、チャネルにアクセスしようとする。複数のステーション14がネットワーク上に送信するデータを有する場合、ステーションの各々が、媒体上の受信搬送波の、デアサートが感知されたことに応答してIPG期間の後に送信を行なおうとするため、衝突が生じる。したがって、送信ステーションは、別のステーションが同時にデータを送信することにより衝突が生じていないかを判断するために媒体を監視する。衝突が検出されれば、両方のステーションが停止し、ランダムな期間だけ待機し、再度送信を試みる。

100 Mb/sネットワークステーション16は好ましくは、提案されているフロー制御によるイーサネット規格IEEE 802.3x全二重草案(0.3)に従う全二重モードで動作する。全二重環境は各100 Mb/sネットワークステーション16とマルチポートスイッチ12との間に双方向ポイントツーポイント通信リンクを設け、IMSおよびそれぞれのステーション16は衝突することなくデータパケットの送受信を同時に行なうことができる。100 Mb/sネットワークステーション16の各々は、100ベース-TX、100ベース-T4または100ベース-FXタイプの100 Mb/s物理 (PHY) 装置26を介してネットワーク媒体18に結合される。マルチポートスイッチ12は、物理装置26への接続をもたらす媒体独立インタフェース (MII) 28を含む。100 Mb/sネットワークステーション16は他のネットワークへの接続のためのサーバまたはルータとして実現され得る。100 Mb/sネットワークステーション16は、所望に応じて半二重モードでも動作する。同様に、10 Mb/sネットワークステーション14は、フロー制御による全二重プロトコルに従って動作するように修正され得る。

図1に示されるように、ネットワーク10は、マルチポートスイッチ12と10 Mb/sステーション14との間で送信されたデータパケットの時分割多重化および時分割非多重化を行なう一連のスイッチトランシーバ20を含む。磁気変成器モジュール19は媒体1

10

20

30

40

50

8上の信号の波形を維持する。マルチポートスイッチ12は、時分割多重化プロトコルを用いて単一のシリアルノンリターンツーゼロ(NRZ)インタフェース24を介して各スイッチトランシーバ20に対するデータパケットの送受信を行なうトランシーバインタフェース22を含む。スイッチトランシーバ20はシリアルNRZインタフェース24からパケットを受信し、受信されたパケットを非多重化し、ネットワーク媒体18を介して適切なエンドステーション14にそのパケットを出力する。開示される実施例によると、各スイッチトランシーバ20は独立した4つの10Mb/sツイストペアポートを有し、マルチポートスイッチ12が必要とするPINの数が4分の1に減少するようにするシリアルNRZインタフェースを介する4:1多重化を用いる。

マルチポートスイッチ12は、意思決定エンジン、切換エンジン、バッファメモリインタフェース、構成/制御/状態レジスタ、管理カウンタ、ならびにネットワークステーション14および16のためのイーサネットポート間でデータパケットの経路制御を行なうためのMAC(媒体アクセス制御)プロトコルインタフェースを含む。マルチポートスイッチ12はまた、インテリジェントな切換決定を行ない、後に説明するように、外部の管理エンティティに管理情報ベース(MIB)オブジェクトの形式で統計的なネットワーク情報を与えるための優れた機能を有する。マルチポートスイッチ12はさらに、マルチポートスイッチ12のチップサイズを最小にするためにパケットデータの外部ストアおよびスイッチ論理を可能にするインタフェースを含む。たとえば、マルチポートスイッチ12は、受信したフレームデータ、メモリ構造およびMIBカウンタ情報をストアするための外部メモリ34へのアクセスをもたらす同期型ダイナミックRAM(SDRAM)インタフェース32を含む。メモリ34は2Mbまたは4Mbのメモリサイズを有する80、100または120MHz同期型DRAMであってもよい。

マルチポートスイッチ12はさらに、外部管理エンティティが管理MACインタフェース38によってマルチポートスイッチ12の全体的な動作を制御できるようにする、管理ポート36を含む。後に詳細に述べるように、管理ポート36は、選択された受信データパケットの少なくとも一部と管理情報を提供する新たな情報とを有する、管理フレームを出力する。マルチポートスイッチ12は、PCIホストおよびブリッジ40を介して管理エンティティがアクセスできるようにするPCIインタフェース39をさらに含む。これに代えて、PCIホストおよびブリッジ40が複数のIMSデバイス12に対する拡張バスとしての役割を果たしてもよい。

マルチポートスイッチ12は、1つのソースから少なくとも1つの宛先ステーションに受信データパケットを選択的に送信する内部意思決定エンジンを含む。内部意思決定エンジンには外部ルールチェッカが代用されてもよい。マルチポートスイッチ12は外部ルールチェッカインタフェース(ERCI)42を含み、これは内部意思決定エンジンの代わりにフレーム転送決定を行なうために外部ルールチェッカ44が用いられるようにする。したがって、フレーム転送決定は、内部切換エンジンまたは外部ルールチェッカ44のいずれかによって行なわれ得る。

マルチポートスイッチ12は、ポートごとのステータスをクロックに合わせて出力しLED外部論理48を駆動する、LEDインタフェース46をさらに含む。LED外部論理48は人間が読取ることができるLEDディスプレイエレメント50を駆動する。発振器48はマルチポートスイッチ12のシステム機能に40MHzのクロック入力を与える。

図2は、図1のマルチポートスイッチ12のブロック図である。マルチポートスイッチ12はそれぞれの10Mb/sネットワークステーション14間で半二重のデータパケットの送受信を行なうための24個の10Mb/s媒体アクセス制御(MAC)ポート60(ポート1から24)と、それぞれの100Mb/sネットワークステーション16間で全二重のデータパケットの送受信を行なうための2つの100Mb/s MACポート62(ポート25および26)とを含む。上述のとおり、管理インタフェース36もまたMAC層プロトコル(ポート0)に従って動作する。MACポート60、62および36の各々は、受信先入れ先出し(FIFO)バッファ64と送信FIFO66とを有する。ネットワークステーションからのデータパケットは対応のMACポートで受信され、対応の受

10

20

30

40

50

信 F I F O 6 4 にストアされる。受信されたデータパケットは対応の受信 F I F O 6 4 から外部メモリインタフェース 3 2 に出力されて、外部メモリ 3 4 にストアされる。

受信されたパケットのヘッダもまた、内部ルールチェッカ 6 8 および外部ルールチェッカインタフェース 4 2 を含む、意思決定エンジンに転送され、いずれの M A C ポートからデータパケットが出力されるかを決定する。具体的には、パケットヘッダは、マルチポートスイッチ 1 2 が内部ルールチェッカ 6 8 または外部ルールチェッカ 4 4 を用いて動作するよう構成されているか否かに依存して、内部ルールチェッカ 6 8 または外部ルールチェッカインタフェース 4 4 に送られる。内部ルールチェッカ 6 8 および外部ルールチェッカ 4 4 は、所与のデータパケットに関する宛先 M A C ポートを決定するための意思決定論理を提供する。したがって、意思決定エンジンは、単一ポート、マルチプルポートまたは全ポート（すなわちブロードキャスト）のいずれかに所与のデータパケットを出力し得る。たとえば、各データパケットにはソースおよび宛先アドレスを有するヘッダが含まれ、意思決定エンジンは宛先アドレスに基づいて適切な出力 M A C ポートを特定する。これに代えて、宛先アドレスは、適切な意思決定エンジンが複数のネットワークステーションに対応するものと特定するバーチャルアドレスに対応してもよい。これに代えて、受信されたデータパケットは、（ 1 0 0 M b / s ステーション 1 6 のうちの 1 つのルータを介する）別のネットワークまたは所定のグループのステーションを特定する I E E E 8 0 2 . 1 d プロトコルに準拠する V L A N （バーチャル L A N ）タグ付フレームを含んでもよい。したがって、内部ルールチェッカ 6 8 または外部ルールチェッカ 4 4 のいずれかがインタフェース 4 2 を介して、バッファメモリ 3 4 に一時的にストアされたフレームが単一の M A C

10

20

ポートまたは複数の M A C ポートに出力されるべきかを決定する。外部ルールチェッカ 4 4 を使用することにより、容量の増加、およびフレームが外部メモリに完全にバッファされる前にフレーム転送決定を可能にし、かつマルチポートスイッチ 1 2 がフレームを受信する順からは独立した順で決定が行なわれるようにする、決定キューのうちランダムな順序付け、といった利点をもたらされる。

意思決定エンジン（すなわち内部ルールチェッカ 6 8 または外部ルールチェッカ 4 4 ）は、データパケットを受信すべき各 M A C ポートを特定するポートベクタの形式で転送決定をスイッチサブシステム 7 0 に出力する。適切なルールチェッカからのポートベクタは、外部メモリ 3 4 にデータパケットをストアするアドレス場所と、データパケットを受信して送信するための M A C ポート（たとえば M A C ポート 0 から 2 6 ）の識別子とを含む。スイッチサブシステム 7 0 はポートベクタに特定されたデータパケットを外部メモリインタフェース 3 2 を介して外部メモリ 3 4 から取出し、取出されたデータパケットを特定されたポートの適切な送信 F I F O 6 6 に与える。

30

付加的なインタフェースにより管理および制御情報が与えられる。たとえば、管理データインタフェース 7 2 は、M I I 管理仕様（ I E E E 8 0 2 . 3 u ）に従うスイッチトランシーバ 2 0 および 1 0 0 M b / s 物理装置 2 6 と制御およびステータス情報をスイッチ 1 2 が交換できるようにする。たとえば、管理データインタフェース 7 2 は、双方向管理データ I O （ M D I O ）信号経路に時間基準を与える管理データクロック（ M D C ）を出力する。

P C I インタフェース 3 9 は、P C I ホストプロセッサ 4 0 によって内部 I M S ステータスおよび構成レジスタ 7 4 にアクセスし、かつ外部メモリ S D R A M 3 4 にアクセスするための、3 2 ビット P C I 改訂 2 . 1 に適合したスレーブインタフェースである。P C I インタフェース 3 9 は複数の I M S デバイスのための拡張バスとしての役割も果たし得る。管理ポート 3 6 は標準 7 ワイヤ反転シリアル G P S I インタフェースを介して外部 M A C エンジンにインタフェースされ、標準 M A C 層プロトコルによりホストコントローラがマルチポートスイッチ 1 2 にアクセスできるようにする。

40

図 3 は、この発明の例示的な実施例に従う、図 2 のスイッチサブシステム 7 0 を説明する図である。図 2 に示されるマルチポートスイッチ 1 2 の他のエレメントは、スイッチサブシステム 7 0 とこれらの他のエレメントとの接続を示すために図 3 に再度示される。スイッチサブシステム 7 0 はフレームの受信および転送を行なうためのコアスイッチングエン

50

ジンを含む。スイッチングエンジンを実現するために用いられる主な機能ブロックは、ポートベクタ F I F O 6 3 と、バッファマネージャ 6 5 と、複数のポート出力キュー 6 7 と、管理ポート出力キュー 7 5 と、拡張パスポート出力キュー 7 7 と、フリーバッファプール 1 0 4 と、マルチコピーキュー 9 0 と、マルチコピーキャッシュ 9 6 と、リクレームキュー 9 8 とを含む。これらの機能ブロックの動作および構成は後により詳細に説明するが、まず、個々のエレメントに関する後の説明に関連性を持たせるために、図 3 のスイッチサブシステム 7 0 の全体像を簡単に説明する。

ポートからマルチポートスイッチ 1 2 に入るフレームには基本的に 2 つのタイプがある。すなわち、単一コピーフレームとマルチコピーフレームとである。単一コピーフレームは、マルチポートスイッチ 1 2 によって他の 1 つのポートにのみ送られることとなる、ポートで受信されたフレームである。これとは対照的に、マルチコピーフレームは、1 つのポートで受信され、1 つより多い数のポートに送信されるフレームである。図 3 では、各ポートは対応する M A C 6 0、6 2、または 3 6 によって表わされ、それ自体の受信 F I F O 6 4 および送信 F I F O 6 6 を有する。

単一コピーにせよマルチコピーにせよフレームは内部 M A C エンジン 6 0、6 2、または 3 6 によって受信され、対応する受信 F I F O 6 4 に置かれる。各データフレームは、少なくとも宛先アドレス、ソースアドレス、およびタイプ/レングス情報を含むヘッダを有する。このヘッダはルールチェッカ(すなわち内部ルールチェッカ 6 8 または外部ルールチェッカインタフェース 4 2 のいずれか)に与えられる。ルールチェッカは、ヘッダの情報に基づいて、フレームパケットがどこから送り出されるか、すなわちいずれのポートを介してフレームパケットが送信されるかを決定する。

ルールチェッカ 4 2 または 6 8 が転送決定を行なうのと同時に、バッファマネージャ 6 5 はフリーバッファプール 1 0 4 からフリーフレームポインタを得る。このフリーフレームポインタは、受信 F I F O 6 4 に現在ストアされているデータフレームをストアするために利用可能である外部メモリ 3 6 の場所を特定する。

バッファマネージャ 6 5 は F I F O 6 4 から受信されたデータフレームを、直接メモリアクセス(DMA)トランザクションでデータバス 8 0 (図 2 参照)を介して外部メモリ 3 4 へ転送し、そのデータフレームはフリーバッファプール 1 0 4 から得られたフリーフレームポインタによってポイントされた場所にストアされる。

バッファマネージャ 6 5 はまた、フリーフレームポインタをルールチェッカ 4 2 または 6 8 に送信し、データフレームのストア場所は維持しつつ、適切なルールチェッカがヘッダ情報を処理できるようにする。このフリーバッファポインタはここでは単にフレームポインタと呼ばれる。なぜなら、フレームがストアされる外部メモリ 3 4 でのメモリ場所をポイントするからである。ルールチェッカ 4 2 または 6 8 は転送決定を行ない、ここでルールチェッカは、対応するヘッダ情報に基づいて外部メモリ 3 4 にストアされたデータフレームに対する少なくとも 1 つの宛先ポートを特定し、かつ「ポートベクタ」の形式で転送命令を発生する。図示される例示的な実施例では、ポートベクタは、フレームが転送されるべき宛先ポートとして特定された、各出力ポートについて対応の 1 ビットがセットされている 2 8 ビットベクタである。受信されたフレームが単一コピーフレームであると想定すると、ルールチェッカ 4 2 または 6 8 によって生成されたポートベクタには、1 つの宛先ポートに対応する 1 ビットしかセットされない。したがってルールチェッカは、ポートベクタを用いてフレームポインタを少なくとも 1 つの宛先ポートに割り当てる。

ルールチェッカ 4 2 または 6 8 はポートベクタ F I F O 6 3 にポートベクタおよび対応するフレームポインタ(ならびに制御操作コードおよび V L A N インデックス)を置く。ポートベクタはポートベクタ F I F O 6 3 によって検査され、ポートベクタに関連したフレームポインタがどの特定の出力キュー 6 7 に入力されるべきかを決定する。ポートベクタ F I F O 6 3 は適切な出力キュー 6 7 の一番上にフレームポインタを置き、対応する宛先ポートからのデータフレームの送信をキューとして維持することにより、フレームポインタを適切な宛先ポートに割り当てる。したがって、フレームポインタは「割当て済みフレームポインタ」となり、宛先ポートに割り当てられる。

10

20

30

40

50

ある時点で、フレームポインタは出力キュー67を通過し、出力キュー67の一番下まで到達する。バッファマネージャ65は、フレームポインタ読み出しバス86を用いて出力キュー67の一番下から割当て済みフレームポインタを取り、割当て済みフレームポインタによってポイントされた外部メモリ36内の場所から対応するデータフレームをDMAトランザクションによって取り出し、取り出したデータフレームを対応するMAC層によって送信するために、データバス82(図2参照)を介して適切な送信FIFO66に置く。

マルチコピー送信は、ポートベクタが、フレームがそれらから送信されることとなる複数の宛先ポートを示すため複数のビットがセットされている点を除いて、単一コピー送信と同様である。フレームポインタは、適切な出力キュー67の各々に置かれ(すなわちストアされ)、対応の送信FIFO54から送信される。

バッファマネージャ65は特殊な制御キュー、すなわち、フリーバッファプール104と、マルチコピーキュー90と、リクレームキュー98と、マルチコピーキャッシュ96とを用いて、受信データフレームをストアするためのバッファを割当て、フレームがその指定された出力ポートに送信されると再度使用できるようバッファを取出すプロセスを管理する。バッファマネージャ65はまた、出力キュー67、75、77ならびに制御キュー104、90および98のために外部メモリ36に「オーバフロー」領域を維持する。特に、これらのキューは各々、チップ上およびチップ外ストア場所を含む3部構成である。性能をあげるためにはチップ上の記憶部が好ましく、この場合すべてのキュー構成がチップ上に維持される(マルチポートスイッチ12参照)。しかしながら、チップ上の占有面積は非常に高価であり、この価格は、チップが多数のエントリを切換えるように設計され、それらをキューとして維持する必要があるときには問題を生じる。この発明は、チップ上に高性能な小容量セクションを含み、チップ外、すなわち離れた別個のメモリチップ34に実装されるオーバフロー領域を含む単一の出力キューを設けることによって、このジレンマを解消する。したがって、オーバフロー領域は、所要の大容量のキューとしてキューが役割を果たすようにし、またチップ外のオーバフロー領域の性能が比較的低いことが、出力キューの全体の性能に不利な影響を及ぼすことのないような態様で出力キュー内に構成される。

この発明の各論理キュー67、75、77、90、98および104は、チップ12上に位置づけられた書き込み側キュー76と読み出し側キュー78とを含み、書き込み側キューのオーバフロー領域(全般的に110と称される)は、外部メモリ34の割り当てられた部分に位置づけられる。出力キュー67のすべてに対する外部メモリ34のアクセスは、前述のように、外部メモリインタフェース32を介して行われる。この発明は、現在の外部メモリのバースト的な性質を利用し、(フレームポインタなどの)オーバフローデータが、外部メモリ34へのバス84を介してバースト状にチップとオーバフローキュー領域110との間で送られるようにする。

書込側キュー76および読出側キュー78はチップ12上にあり、これらは小さくて高価な高性能のリソースであると考えられる。これとは対照的に、出力キュー67の第3の部分形成するオーバフロー領域110は大きく、安価で、低性能の大容量の経路を提供する。

各バッファ67、75、77、90、98および104は、対応の書き込み側キュー76がその入力端で対応のフレームポインタエントリを受信することによって動作する。フレームポインタは、データフレームの最初の256バイトをストアする外部メモリの第一のバッファ場所をポイントする。

エントリが書込側キュー76内を完全に移動し、その一番下の出力端まで到達すると、出力キュー67に関連した制御論理はエントリを外部メモリ34の対応の割当て部分110に出力すべきか、または読出側キュー78に出力すべきかを選択する。読出側キュー78に利用可能なスペースがあり、出力キュー67のオーバフロー領域110が空いていれば、1つまたはそれ以上のエントリが書込側キュー76から読出側キュー78に直接渡される。このように書込側キュー76から読出側キュー78に直接エントリを送ることはすべ

10

20

30

40

50

てチップ 1 2 上で行なわれるので、エントリは低レイテンシで高速に流れる。

読出側キュー 7 8 はいっぱいであるが、書込側キュー 7 6 にはまだ 1 バーストサイズの量のデータがない場合は、エントリは書込側キュー 7 6 に留まる。読出側キュー 7 8 がいっぱいであり、書き込み側キュー 7 6 に少なくとも 1 バーストサイズの量のデータ（たとえば 1 6 バイトに値するエントリ）がある場合は、データはバッファマネージャ 6 5 によってバースト形式で外部メモリ 3 4 の対応する割当て部分にあるオーバフロー領域 1 1 0 内に書き込まれる。最終的には、読出側キュー 7 8 は空になり、もしオーバフロー領域 1 1 0 にデータがあれば、読出側キュー 7 8 にそのバーストサイズの量のデータを収容する十分なスペースが生まれると、バッファマネージャ 6 5 はオーバフロー領域 1 1 0 から 1 バーストのデータを読出側キュー 7 8 に供給する。したがって、読出側キュー 7 8 は書込側キューまたは外部メモリ 3 4 の割当て部分 1 1 0 から選択的にフレームポインタを受け取る。

10

したがって、もし出力キュー 6 7 が多数のエントリ（たとえばフレームポインタ）を受信し始めると、これらのエントリはオーバフロー領域 1 1 0 に置かれ、チップ上のキュー 7 8 のオーバフローを回避し、フレームの廃棄の可能性を最小にすることができる。オーバフロー領域 1 1 0 のためのメモリの合計量はまた、外部メモリ 3 6 のサイズを変更することによって容易に変更可能である。さらに、個々の特定のオーバフロー領域 1 1 0 のサイズは、出力キュー 7 4 の性能に影響を及ぼすことなくキューのサイズをカスタマイズするためにプログラム可能である。

図 1 および 2 に示すマルチポートスイッチは 2 8 の出力キューを有し、これらはそれぞれ 1 0 M b / s ユーザポート 6 0 のためのものが 2 4 個と、1 0 0 M b / s サーバポートのためのものが 2 つと、管理ポート 3 6 のためのものが 1 つと、拡張バスポート 3 8 のためのものが 1 つとである。キュー作業は、転送ポートベクタに示されるさまざまな出力キュー 6 7、7 5 および 7 7 に対してポートベクタ F I F O 7 0 がフレームポインタを書込むという形態をとる。

20

図 4 は外部メモリ 3 4 の例示的なマップを示す図である。外部メモリ 3 4 の全体の容量はたとえば 4 M b であるが、種々の実施例において他の容量のメモリが採用されてもよい。この発明に従ってオーバフロー領域に外部メモリ 3 4 を使用することにより、外部メモリを変更するだけで出力キューのサイズを増減することができる。これは、キューとして維持する容量全体がチップの製造時に設定される、キュー構成がすべてチップ上にあるシステムよりも有利である。

30

スイッチ 1 2 のオーバフローストア要件を満たすために、外部メモリ 3 4 のオーバフロー領域 1 1 0 は、フリーバッファブルオーバフロー 1 2 0 と、リクレーンキューオーバフロー 1 2 2 と、マルチコピーキューオーバフロー 1 2 4 と、管理ポート出力キューオーバフロー 1 2 6 と、1 0 M b / s および 1 0 0 M b / s 宛先ポート（ポート 0 から 2 6 ）の各々のための出力キューオーバフロー 1 2 8 と、拡張バスポート（ポート 2 7 ）出力キューオーバフロー 1 3 0 とに対して、メモリ部分を割り当てる。メモリ 3 4 はまた、M I B カウンタ 1 3 2 と、グローバルフレームバッファブル 1 3 4 とについての割当て部分をも含む。

メモリ領域全体の B A S E アドレスはチップ上のレジスタ 7 4 の中のメモリベースアドレスレジスタ内にプログラム可能である。外部メモリマップ内の各領域の B A S E アドレスはレジスタセット内にプログラム可能である。所与の領域の長さは、マッピング内のその領域の B A S E アドレスから次の領域の B A S E アドレスまでの領域に等しいので、領域長レジスタは不要である。

40

個々のオーバフロー領域の長さ（したがって容量）がプログラム可能であるため、各キューの容量全体がプログラム可能である。この発明のこの特徴により、必要に応じて容量の増大した特定の出力キューを提供するようにスイッチをカスタマイズすることが可能になる。

オーバフロー領域 1 1 0 は、チップ 1 2 上の制御キューに適合しない超過のエントリをストアする。たとえば、フリーバッファブル 1 0 4 に対してフリーバッファブルオーバ

50

フロー領域 120 は、グローバルフレームバッファプール 134 中の現在未使用のバッファを特定するフリーフレームポイントの超過分をストアする。リクレーンキュー 98 に対してリクレーンキューオーバーフロー領域 122 は、必要でなくなったリンクトリストチェーンにフレームポイントの超過分をストアする。マルチコピーキューオーバーフロー領域 124 は（キューとして維持されたフレームポイントについては）コピーナンバー「 1 」を付してフレームポイントの超過分をストアし、また（送信成功したフレームについては）コピーナンバー「 - 1 」を付してフレームポイントをストアする。管理ポートキュー 74 に対して管理ポート出力キューオーバーフロー領域 126 は、管理ポート 36（ポート 0）への送信を待機する割当て済みフレームポイントの超過分をストアする。出力キューオーバーフロー領域 128 はそれぞれのポートキュー 67 に対して適切な 10 Mb/s ポート（ポート 1 から 24）または 100 Mb/s ポート（ポート 25 から 26）への送信を待機する割当て済みフレームポイントの超過分をストアする。拡張パスポート（ポート 27）キュー 77 に対する拡張パスポート出力キューオーバーフロー領域 130 は、拡張パスポートへの送信を待機するフレームポイントをストアする。

10

MIB カウンタ領域 132 は、スイッチ 12 によって周期的に更新されるポートごとの統計をすべて含む。スイッチ 12 は MIB 統計をストアするための 8 ビットおよび 16 ビットカウンタをチップ上に維持する。スイッチ 12 は MIB データの損失を防止するために要求される周波数で、外部メモリ 36 の 32 ビットまたは 64 ビットの MIB カウンタを更新する。

グローバルフレームバッファプール 134 は、受信されたフレームデータをストアするリンクトリストのバッファを含む。任意の時点で、これらリンクトリストは有効フレームデータと無効になったバッファとを含み、無効になったこれらのバッファは、バッファマネージャ 72 によってフリーバッファプール 104 に戻されるか、または P C I ホストプロセッサ（図示せず）の所有となる。

20

管理データフレーム合成

この発明は、選択された受信データフレームの少なくとも一部分と、選択されたデータフレームの受信に対応する特性を特定する新たな情報とを含む、管理データフレームの合成に向けられる。たとえば、この新たな情報は、データパケットを受信したネットワークポート 60 または 62 に基づいて、もしくはデータパケットのソースアドレスに基づいて、データフレームのソースを特定できる。同様に、この新たな情報は宛先アドレスに基づいて受信データパケットの宛先を特定することができる。このように、新たな情報は管理エージェントによって、2つのネットワークポート間、またはこれに代えて、対応の MAC アドレスに基づく2つのネットワークステーション間のネットワークトラフィックを監視するために用いられ得る。

30

この新たな情報はまた、データパケット自体もしくはネットワークスイッチ 12 に関する受信ステータスも特定し得る。たとえば、新規情報は、受信データパケットにフレーム整列エラーや巡回冗長検査（CRC）エラーがあったか、またはデータパケットを受信する受信 FIFO 64 がオーバーフローに出会うかを特定し得る。この新規情報はさらに、データフレームの受信に回答してマルチポートスイッチ 12 自体の特性を特定することもできる。上述したように、ルールチェッカ 60 または 42 は、ソースアドレスおよび宛先アドレスに基づいて受信データパケットの宛先を決定する。管理エージェントに、適切なルールチェッカの能力に関する情報を提供することが望ましい場合もある。たとえば、受信データパケットのソースアドレスが認識不可能であり、新規ネットワークスイッチ 12 によってネットワーク上の新規なステーションとして登録（すなわち学習）されなければならない場合、新たな情報は、管理エージェントが新規なステーションの監視を始めるために用いられ得る。これに代えて、管理エージェントは、新規ステーションのネットワークスイッチ 12 を再構成するために、たとえばバッファを再割当するかまたはネットワークポート構成レジスタを再設定することによって、再構成するために新たな情報を用いる場合もある。管理フレームもまた、2つのステーション間のトラフィックを監視するために用いられ得る。

40

50

したがって、管理エージェントは、管理データフレームのためにマルチポートスイッチ 12 が生成した新たな情報により、データフレームの受信および選択されたポートを選択的に監視することに基づいて、または所定のソースもしくは宛先アドレスを有するデータパケットに基づいて、マルチポートスイッチのステータスに関する正確な情報を受信することが可能になる。したがって、管理インタフェース 36 は複数の管理データフレームであって、その各々が、たとえば受信ステータス、データパケット特性、およびデータフレームの受信に対するスイッチ応答特性など、選択されたデータフレームの受信に対応する特性を特定する新たな情報を含む管理データフレームを出力する。

図 10A は、ネットワークスイッチ 12 内の、管理エージェントに対する管理フレームを生成するための装置を表わすブロック図であり、図 10B は、この発明の一実施例に従って管理フレームを生成するための方法を表わす図である。この方法はステップ 500 で、MAC ポート 60 または 62 がネットワークステーション 14 または 16 のうちの 1 つからデータパケットを受信することによって開始する。管理データフレームについての新たな管理情報は、スイッチ内の種々のソースから関連情報を選択的に収集することによって生成される。

受信特性は、ステップ 502 で、データパケットを受信して受信ステータスデータを生成する MAC 60 または 62 によって決定される。具体的には、MAC 60 または 62 は入力データフレームを受信し、データパケットが、最小および最大の正当なパケットサイズに関してイーサネットプロトコルに従ったものであることを確認するため検査する。例示的なネットワークデータパケットが、タグ付でないフレームフォーマットについては図 9A に、タグ付フレームフォーマット (IEEE 802.1d) については図 9B に示される。タグ付でないフレーム 140 およびタグ付フレーム 142 の各々は、6 バイト宛先アドレスフィールド 144 と、6 バイトソースアドレスフィールド 146 と、タイプ/レングスフィールド 148 (2 バイト) と、フィールド幅が 46 バイトから 1500 バイトである可変長データフィールド 150 と、巡回冗長検査 (CRC) フィールドとも呼ばれるフレームチェックシーケンス (FCS) フィールド 152 (4 バイト) とを含む。タグ付フレーム 142 はまた、2 バイトの VLAN イーサタイプフィールド 154 と 2 バイトの VLAN ID フィールド 156 とを含む VLAN タグも含む。当該技術分野で知られているように、タグ付でないフレーム 140 とタグ付フレーム 142 との前にはともに、56 ビットプリアンプルおよび 8 ビットフレーム開始デリミタ (SFD) が付される。

CRC フィールドとも呼ばれるフレームチェックシーケンスフィールド 152 は、受信データパケットのエラーチェックコードとしての役割を果たし、MAC 60 または 62 は CRC フィールド 152 を用いてエラーを検査する。当該技術分野では知られているように、各データパケットは、受信データフレーム内の何らかのエラーを正確に検出するために、それら自身の CRC フィールド 152 を 1 つだけ有する必要がある。データフレームを受信する MAC 層 60 または 62 は、ステップ 502 で、そこに何らかの CRC エラー (C) またはフレーム整列エラー (L) が存在するか否かを判断し、対応するエラーが検出されると適切なフラグをセットする。受信パケットがタグ付フレームである場合、MAC によって対応するビット (T) がセットされる。その後、受信データパケットは対応の受信 FIFO 64 に置かれ、バッファマネージャ 65 によって外部メモリ 34 にストアされる。ステップ 504 で受信 FIFO 64 のオーバフローがあると、MAC はステップ 506 でオーバフローフラグビット (O) をセットしてデータ損失の可能性を指摘する。受信ステータスデータ (C、L、O、および T フラグビットならびに受信ポート ID を含む) と、受信されたデータパケットとはその後、ステップ 508 で外部メモリ 34 にストアされる。

図 5 は、受信データフレームを格納する外部メモリ 34 内のバッファについてのフレームバッファヘッダフォーマットを表わす図である。バッファはメモリ内の次のバッファの場所を示す各バッファヘッダのアドレスポイントによって鎖状に繋がれる。バッファヘッダはまた、データパケットを受信した対応する MAC によって送信された受信ステータス情報も含む。具体的には、第 1 のバッファヘッダ 170 およびそれに続くバッファヘッダ 1

10

20

30

40

50

72は、バッファフォーマットビット174と、フレーム終端マーカ176と、受信ステータスデータ178と、バッファ長180と、ネクストバッファポインタ182とを含む。バッファフォーマットビット174は、ヘッダのフォーマットが先頭バッファヘッダ(12バイトを有する)か後続バッファヘッダ(4バイトを有する)かを特定し、これはバッファ同士を鎖状に繋ぐのに用いられる。フレーム終端マーカ176は、ビットが1にセットされていると、対応するバッファヘッダがあるフレームについての最後のバッファであることを特定し、その鎖内にそれ以上バッファがないことを示す。バッファ長180は、バッファヘッダの後の最初のバイトから始まる、バッファのデータフィールド内の有効な全バイト数を特定し、ネクストバッファポインタ182は次のバッファに対するポインタを含む。ネクストバッファポインタ182はフレーム終端マーカ176がセットされていると無効である。 10

先頭バッファヘッダ170および後続バッファヘッダ172は、受信ステータスデータ178を含む。Cビット178aはCRCエラーがMACによって検出されたかを示す。Lビット178bは、フレーム整列エラーがCRCエラーとともに受信フレーム内でMAC60または62によって検出されたかを示す。Oビット178cは、受信FIFO64がオーバーフローしたかを示し、これはバッファ内のデータが無効であるかもしれないことを示す。第1のバッファヘッダ170はまた、入力受信フレームのポート型を明記するPビット184も含み、ここで0は10Mb/sポートを指し、1は100Mb/sポートを指す。Pビット184はタイムスタンプフィールド186とともにホストによって用いられ、このときマルチポートスイッチ12は、フレームが外部メモリに完全に受信されかつバッファされる前に、フレームを拡張バスに転送するようにプログラムされる。第1のバッファヘッダ170はまた、フレームの受信元のポート番号を特定する受信ポート番号188と、受信されたフレームタイプがタグ付であるかタグ付でないかを示すTビット190とを含む。第1のバッファヘッダ170はまた、VLANフィールド154および156からのVLAN識別子192も含む。 20

したがって、MAC層60または62は、受信データパケットの受信ステータスを決定し、外部メモリ34にストアされたバッファヘッダ170のストレージについての情報をステップ508で転送する。受信されたデータフレームの状況(たとえばエラーなし、CRCエラー、フレーム整列エラー)に関する受信ステータスデータのストレージと、データパケットを受信する受信ポートのステータス(たとえば受信FIFOオーバーフローなど)とにより、マルチポートスイッチ12は、管理エージェントが用いる受信データパケットの受信に関連する新たな情報を収集することが可能になる。したがって、受信ポートでMAC層60または62によって生成されたこの新たな情報は、以下に詳細に示すように、管理エンティティに提供する新たな管理情報の一部としてバッファヘッダ170にストアされる。 30

データパケットの受信に回答して、マルチポートスイッチの特性を特定する情報もまた、適切なルールチェッカ42または68によって受信データパケットのフレーム転送決定の間に生成される。ルールチェッカは情報を受信し、図6に示す転送ポートベクタを生成する処理を開始する。この処理中、ルールチェッカはデータパケットを受け取るMACから情報を受信し、図6に示す転送ポートベクタ200を生成する。たとえば、ルールチェッカは、宛先アドレスフィールド144と、ソースアドレスフィールド146と、受信ポート番号と、フレームポインタとを受信するのである。フレームがタグ付ポートによって受信されると、VLANタイプフィールド154およびVLANIDフィールド156を含むVLANTAGが入力フレームから取除かれ、バッファヘッダ170でVLAN識別子フィールド192に格納されてルールチェッカおよび外部メモリに供給される。 40

ルールチェッカ42または68は、図3に示すスイッチングサブシステム70のスイッチ論理としての役割を果たす。スイッチ論理は、受信データフレームが管理ポート36を介して管理エージェントへ転送されるべきか否かを特定するスイッチ論理データを含む。ルールチェッカは、VLAN関係および転送ポートベクタとともに、1組のアドレスを含む。ルールチェッカはステップ510で適切なアドレスについてそのアドレステーブルを検 50

索し、ソースアドレス、受信ポート、D A、およびV L A N関係に基づいて転送決定をなす。

ルールチェッカアドレステーブルは、ルールチェッカがソースアドレス/受信ポート番号および宛先アドレス/V L A Nインデックスに基づいて転送決定を生成するのに十分な情報を含む。たとえば、ルールチェッカアドレステーブルは、選択されたM A Cアドレスまたはポート宛先に対するソースおよび宛先M A Cアドレスが管理ポートに出力されるべきか否かを特定する、トラヒックキャプチャビット（たとえばトラヒックキャプチャ1およびトラヒックキャプチャ2）を含むであろう。ルールチェッカアドレステーブルはまた、16ビットV L A N識別子を参照するのに用いられるV L A Nインデックスも含むであろう。アドレステーブルはまた、関係するアドレスがあるポートを特定するポート番号と、フレームを転送するための転送ベクタを提供するポートベクタとを含む。

10

ルールチェッカはステップ512でスイッチ論理データをポートベクタ200の形態で生成し、図3に示すポートベクタF I F O 63にポートベクタ200を出力して、管理ポート36を含む選択出力ポートに受信データパケットを送信する。図6に示すように、ポートベクタ200は管理エージェントがルールチェッカアドレステーブルエントリを位置づけるのに用いるインデックスを形成する、ピン番号200とエントリ番号202とを含む。ベクタ200はまた、V L A Nインデックス204と、制御演算コード206と、宛先ポートを特定する転送ポートベクタ208と、図5に示すようにフレームデータおよび対応のヘッダ情報をストアする外部メモリ34での場所を明記するフレームポインタ210とを含む。

20

ピン番号200およびエントリ番号202によって形成されたアドレスインデックスは、管理エージェントが対応の受信フレームに対する切替決定を生成する特有のルールチェッカアドレステーブルエントリにアクセスすることを可能にする。具体的には、管理エージェントは、ルールチェッカ68または外部ルールチェッカインタフェース42を介して外部ルールチェッカ44に対するアドレステーブルを作りかつ維持することについての責任を負う。管理エージェントはアドレステーブル内の3つのエントリすべての初期リストを生成し、アドレスおよびそれらの関係するフィールドをテーブル内に挿入し、エントリを適切に有効にし、加え、消去し、またはエージングするためにアドレステーブルのソフトウェアマッピングを管理し、さらにアドレスエントリ内のフィールドを更新する。

さらに、ルールチェッカは新規なステーション14または16がネットワーク上に存在することについて学習する能力がある。たとえば、予め定められたM A Cアドレスを有する新規ステーションは、マルチポートスイッチ12を介してデータパケットを別のステーションへ送信できる。ルールチェッカが受信データパケットのソースアドレスフィールド146で特定されたソースアドレスに対するアドレステーブルエントリを含まない場合は、ルールチェッカは、その内部アドレステーブルを新規ソースアドレスについての新規アドレステーブルエントリを含むように更新することにより、新規ステーションについて「学習する」ことが可能である。一旦、新規アドレステーブルエントリが形成されると、ルールチェッカは転送ポートベクタを適切に生成することができる。

30

図7は中央の演算コードフィールド206のサブフィールドを表わす図である。制御演算コードフィールド206は、ステップ512でルールチェッカによって生成された、受信データパケットを受信することに対応して受信フレームおよび/またはルールチェッカの特性を特定することに関する情報を提供する。表1は、受信データパケットのアイデンティティおよび/またはデータフレームの受信に対応するルールチェッカのステータスに関する、管理エージェントについての情報を提供する、制御演算コードフィールド206内のコードの例を挙げる。

40

表 1

演算コード フィールド	フィールド ネーム	詳 述
7-5	トラヒック キャプチャ	000=ヌル 001=スニファポートフレーム 010=トラヒックキャプチャ1フレーム 011=トラヒックキャプチャ2フレーム 100-110=予約 111=予約 (さらに他の演算コードフィールドを特定)
4-2	I R C アドレス	000=既知の SA 001=予約 (優先順位の高い既知の SA) 010=学習された未知の SA (自動検証または未検証) 011=学習されない未知の SA ; 学習しないまたは不良フレーム 100=学習されない未知の SA ; ビンまたはフリーエントリチェーンがロックされていた 101=学習されない未知の SA ; アドレステーブルに余白なし 110=未知の VLAN タグ (認識されない VLAN 識別子) 111=バックボーン転送モード SA 整合エラー
1-0	Mgmt ポート/ タグ付け	00=ヌル 01=BPDU またはブリッジマルチキャストフレーム 10=宛先指定フレーム (DA=内部 IMS ポート) 11=タグ付ポートのディセーブル「タグ」 (100Mb/s のタグ付ポートについて)

10

20

表 1 および図 7 に示すように、制御演算コード 206 は 3 つのサブフィールド、すなわちトラヒックキャプチャ 206 a と、I R C アドレス 206 b と、管理ポート/タグ付 206 c とを含む。トラヒックキャプチャフィールド 206 a は受信データパケットが監視されたトラヒックからキャプチャされたかを特定する。たとえば、トラヒックキャプチャフィールドコード「001」はスニファポートフレームを特定する。「スニファポート」は識別されたポートのすべてのトラヒックが管理エージェントに転送されることを特定する。たとえば、図 2 の「MAC 23」がスニファポートと指定されたと仮定すると、管理エージェントは管理インタフェース 36 によって「MAC 23」によって授受されたデータパケットすべてについて通知されるであろう。同様に、トラヒックキャプチャ 1 フレームコードおよびトラヒックキャプチャ 2 フレームコードは、1 つ以上の MAC アドレスが所与の半二重ネットワークポート上に存在している場合に、管理エージェントが 2 つのポートまたは 2 つの MAC アドレス間に起こるトラヒックを監視できるようにする。

30

I R C アドレスサブフィールド 206 b は、データフレームの受信に回答する内部ルールチェッカ 68 の特性を含む。たとえば、I R C アドレスフィールド値 000 は、ソースアドレスが、対応のアドレステーブルエントリをルールチェッカ内に有することを特定する。対照的に、I R C フィールド値「010」から「101」は、未知のソースアドレスが内部ルールチェッカ 68 によって学習されたか否かとともに、未知のソースアドレスを特定する。未知のソースアドレスが学習されなかった場合は、種々のエラー条件が、管理エージェントに対して、未知のソースアドレスが学習されなかった理由、たとえば不良フレーム、アドレステーブルがいっぱいであること、またはピンもしくはフリーエントリチェーンがロックされていたこと (たとえばアドレステーブルを格納するメモリ構造がロックされていたことなど) などの追加情報を提供する。この具体的なエラー状況により、管理エージェントは内部ルールチェッカアドレステーブルエントリを未知の新規ステーションを認識するように再構成することが可能になることもある。

40

制御演算コード 206 の管理ポート/タグ付サブフィールド 206 c はまた、マルチポートスイッチ 12 によって授受され得る特別な管理フレームのアイデンティティも特定する

50

。たとえば、「01」の値はブリッジプロトコルデータユニット(BPDU)フレームまたはブリッジマルチキャストフレームを特定する。BPDUフレームは、ネットワークブリッジ間で送信されて極大木を決め、冗長データリンクを消去するための標準化プロトコルフレーム(IEEE 802.1d)である。管理ポート値「10」は宛先指定フレームを特定し、ここで宛先アドレスは別のネットワークステーション14または16のMACアドレスではなくMAC 60または62のうちの特定の1つを特定する。したがって、管理エージェントは、ネットワークポートの1つに特定の送信されたデータフレームを特定することができる。管理ポート/タグ付値「11」は、100Mb/sポートについてのタグがVLAN拡張動作をディセーブルにするためにディセーブルされたことを特定する。

10

したがって、ルールチェッカはステップ512で、管理エージェントがルールチェッカを必要に応じて監視および更新できるように、マルチポートスイッチおよびルールチェッカ42または68内の動作を特定する付加的情報を生成する。ルールチェッカはポートベクタ200をポートベクタFIFO63に転送して適切な出力ポートへと配布する。制御演算コード206が、受信VLAN IDが認識されなかったことを示す値「xxx110xx」を含む場合、VLAN IDはフィールド204では転送されない。しかしながら、制御演算コードが「xxx110xx」と等しくない場合、VLAN IDはルールチェッカからフィールド204によりポートベクタFIFO63へ転送される。

上述したように、図3に示すポートベクタFIFO63はポートベクタ200を受信し、適切な情報を転送ポートベクタ208で特定された出力キューへ出力する。管理フレームが管理ポート36によって生成されると仮定すると、転送ポートベクタ208は管理ポート36に対応するビットがセットされ、よってポートベクタFIFO63はフレームポインタおよびスイッチ論理データをステップ514で管理ポート出力キュー75へ転送することになるであろう。具体的には、ポートベクタFIFO63は、ピン番号200と、エントリ番号202と、(適切であれば)VLANインデックス204と、制御演算コード206と、図10Aで示す管理ポート出力キュー75に対するフレームポインタ210とを転送するであろう。

20

一旦、管理情報(ピン番号200、エントリ番号202、適切であればVLANインデックス204、制御演算コード206、およびフレームポインタ210を含む)が管理ポート出力キュー75の終端に到達すると、バッファマネージャ65はステップ516で、外部メモリ34にストアされた情報を得るためにフレームポインタ210に基づいて外部メモリにアクセスし、そこで得たヘッダ情報170をヘッダFIFO67にストアし、受信データパケットの対応のストアされたデータを送信FIFO66にストアする。したがって、受信するMACによって生成された管理情報(受信ステータス)および、ルールチェッカによって生成された管理情報(スイッチステータス)は、ヘッダFIFO67にストアされる。FIFO67はその後、図8Aに示す管理データフレーム300の24バイトヘッダ330を生成する管理フレームジェネレータ69に管理情報を出力する。

30

管理MAC36はジェネレータ69から管理データヘッダを得て、また送信FIFO66から外部メモリ34にストアされた受信データの少なくとも一部分を得て、ステップ518で管理フレーム300を生成する。具体的には、管理ポートMAC36は、イーサネットプロトコルに従ったプリアンブル302およびフレーム開始デリミタ(SFD)304を生成する。管理ポートMAC36は、管理ポートMAC36のベースアドレスに対応する宛先アドレス306を含む、ジェネレータ69からのヘッダを付する。宛先アドレス306を管理ポート36のアドレスと等しく設定することにより、管理MAC38を介してデータパケットを受信する管理エージェントが、管理フレームとして受信されたデータパケットを特定することが可能になる。ソースアドレス308は、受信データパケットを有するMAC60または62のアドレスと等しくなるようにジェネレータ69によって設定される。たとえば、データパケットが図2に示す「MAC23」によって受信された場合、ソースアドレス308は「MAC23」のアドレスと等しくするように設定される。

40

管理ポート36はその後、表1および図7に関して上述した管理情報を含む、ジェネレー

50

タ 6 9 によって供給された演算コードフィールド 2 0 6 を挿入する。その後管理ポート 3 6 は、ピン番号 2 0 0 およびエントリ番号 2 0 2 を追加し、さらにそれに続いて、図 9 A および 9 B に示すもとの受信フレームのフレーム長 1 4 8 を付加する。ピン番号およびエントリ番号フィールド 2 0 0 および 2 0 2 は、受信データフレームが学習されたフレームである場合のみ、供給される必要がない。

管理ポート M A C 3 6 はその後、受信ステータスビットである C R C エラービット 1 7 8 a、フレーム整列エラービット 1 7 8 b、受信 F I F O オーバフローエラー 1 7 8 c、およびもとのフレームがタグ付きか否かを特定するタグ付フレームビット 1 9 0 の値を含む、ステータスフィールド 3 1 0 を供給する。ステータスフィールド 3 1 0 はまた、フレームがもとのフレームの 1 2 8 バイトのみを有するかを特定する圧縮フレームビットも含む。具体的には、管理ポート 3 6 は切捨て機能が管理ポート 3 6 内で設定されている場合、受信データフレーム 3 1 2 の 1 2 8 バイトまで送信するように構成される。もとのデータフレーム 3 1 2 は、もとの宛先アドレス 1 3 4 と、もとのソースアドレス 1 4 6 と、もとのタイプ/レングス 1 4 8 と、データフィールド 1 5 0 とを含むように構成される。管理ポート 3 6 内で切捨てが設定されると、1 1 4 バイトより長い長さを有するデータフィールド 1 5 0 は切捨てられるので切捨て後の部分 1 5 0 のみが送信される。しかしながら、データフィールド 1 5 0 が 1 1 4 バイトより短い場合には、データフィールド全体が送信される。

管理ポート 3 6 がデータパケット 3 0 0 のフィールドに挿入された後、管理ポート 3 6 は管理フレーム 3 0 0 のバイト数を正確に反映する新たな C R C フィールド 3 1 4 を生成する。一旦 C R C フィールド 3 1 4 が管理ポート 3 6 によって生成されると、管理フレーム 3 0 0 は管理ポート 3 6 によって並直列変換器 7 1 を介して管理 M A C 3 8 へと出力され、管理エージェントによって処理される。

したがって、管理エージェントは、マルチポートスイッチによって受信されたデータパケットのタイプについての特定の情報と、データパケットを受信するポート 6 0 と、受信されたデータフレーム自体およびデータパケットを受信する受信ポートに対応する受信ステータス情報と、トラヒックキャプチャ、I R C アドレス学習、および管理ポートタグ付に関する情報を特定するスイッチ論理によって生成された情報とを含む、管理データフレームを受信することができる。管理フレームはまた、受信データフレームの少なくとも一部分を含む。したがって、管理データフレームは受信されたデータパケットおよびマルチポートスイッチのデータパケット受信に対する応答とに関する特定の詳細な情報を提供する。したがって、管理エージェントは、マルチポートスイッチ 1 2 がデータパケットを受信した順序に関係なく、管理ポート 3 6 から複数の管理フレームを受信することができる。さらに、M A C 層およびルールチェッカによって提供された情報により、管理エージェントは、監視されたトラヒックに回答するか否か、またはルールチェッカ 4 4 もしくは 6 8 のスイッチ論理に関するアドレステーブルを更新するか否かを含む、適切な方針を決定することができる。

管理エージェントはまた、図 8 B に示す管理ポート 3 6 に管理データパケット 4 0 0 を出力し得る。図 8 B に示すように、管理データパケット 4 0 0 は、ネットワークポートのグループを特定してネットワークパケットを送信するポートベクタ 2 0 8 と、特定されたネットワークポートによって送信されるネットワークデータパケット 1 4 2 とを含む。したがって、管理データパケット 4 0 0 によってルールチェッカはバイパスされるようになり、ポートベクタはポートベクタ F I F O 6 3 に直接供給されて処理される。上記から容易にわかるように、バッファマネージャはまた、フレームポインタをネットワークデータパケット 1 4 2 に割当て、スイッチサブシステム 7 0 内にわたる送信を可能にする。この発明を、現在最も実用的かつ好ましいと考えられる実施例に関して述べてきたが、当然、この発明は開示された実施例に限定されることはなく、対照的に、添付の請求の範囲の精神および範囲内に含まれるさまざまな変更および同等の配置を包含することが意図される。

10

20

30

40

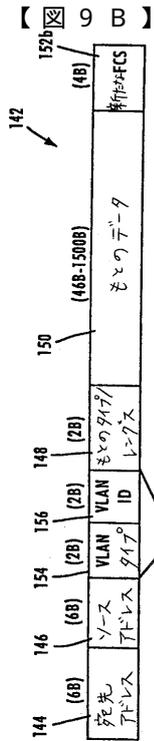


Fig. 9B

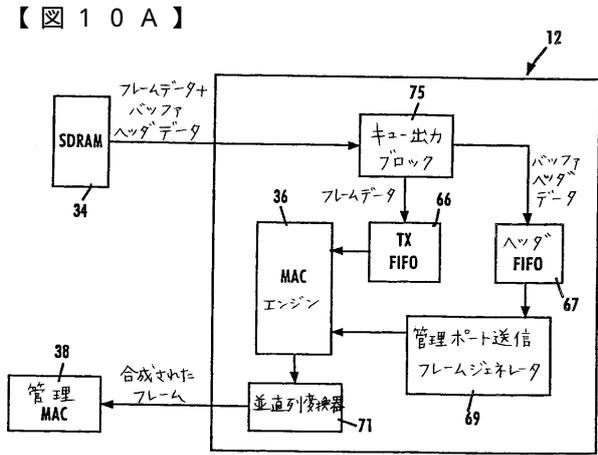
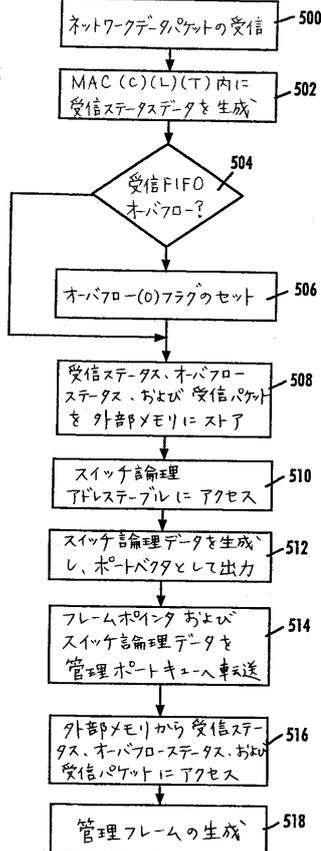


Fig. 10A

【 10 B 】

Fig. 10B



フロントページの続き

(74)代理人

弁理士 堀井 豊

(72)発明者 クレイフォード, イアン

アメリカ合衆国、9 5 1 2 9 カリフォルニア州、サン・ノゼ、アイリーン・ドライブ、5 3 8 0

(72)発明者 カースタイン, デニス

アメリカ合衆国、9 4 3 0 1 カリフォルニア州、マウンテン・ビュー、メドック・コート、4 3 3

審査官 矢頭 尚之

(56)参考文献 特開平7 - 7 9 2 4 5 (J P , A)

特開平8 - 1 9 5 7 6 9 (J P , A)

(58)調査した分野(Int.Cl. , DB名)

H04L 12/28