

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4150258号
(P4150258)

(45) 発行日 平成20年9月17日(2008.9.17)

(24) 登録日 平成20年7月4日(2008.7.4)

(51) Int.Cl.		F I			
HO 4 L	12/44	(2006.01)	HO 4 L	12/44	3 0 0
HO 4 L	12/56	(2006.01)	HO 4 L	12/56	2 0 0 Z

請求項の数 11 (全 19 頁)

(21) 出願番号	特願2002-576377 (P2002-576377)	(73) 特許権者	591016172
(86) (22) 出願日	平成13年11月2日(2001.11.2)		アドバンスト・マイクロ・デバイス・ インコーポレイテッド
(65) 公表番号	特表2004-530343 (P2004-530343A)		ADVANCED MICRO DEVI CES INCORPORATED
(43) 公表日	平成16年9月30日(2004.9.30)		アメリカ合衆国、94088-3453
(86) 国際出願番号	PCT/US2001/051161		カリフォルニア州、サニペイル、ピー・ オウ・ボックス・3453、ワン・エイ・ エム・ディ・プレイス、メイル・ストップ ・68 (番地なし)
(87) 国際公開番号	W02002/078274	(74) 代理人	100099324
(87) 国際公開日	平成14年10月3日(2002.10.3)		弁理士 鈴木 正剛
審査請求日	平成16年10月29日(2004.10.29)	(74) 代理人	100111615
(31) 優先権主張番号	09/814,815		弁理士 佐野 良太
(32) 優先日	平成13年3月23日(2001.3.23)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 ネットワークデバイスにおける選択的データフレーム間引き

(57) 【特許請求の範囲】

【請求項1】

複数のステーション間でのデータフレームの通信を制御するように構成されたネットワークデバイスであって、

前記ステーションからのデータフレームを受信するように構成された複数の受信ポートと、データフレーム処理ロジックと、を備えた受信機を有し、

前記データフレーム処理ロジックが、

受付制御が使用可能であるか否かを判断し、

受信された各データフレームに関連する優先度を判断し、

前記受信されたデータフレームのペイロードポジションに含まれる情報に基づいて、受信された各データフレームに関連するタイプを決定し、

受付制御が使用可能であるか否か、前記第1データフレームの優先度、及び前記第1データフレームに関連するタイプに基づいて、受信したデータフレームのうちの第1データフレームを間引くか否かを判断するように構成されており、

更に、前記データフレーム処理ロジックは、前記第1データフレームに関連するタイプを決定する際に、前記第1データフレームが第1ステーションと第2ステーションとの間の接続を確立することに関連しているか否かを特定するよう構成されており、

前記データフレーム処理ロジックは、

前記第1データフレームが前記第1ステーションと前記第2ステーションとの間に接続を確立することに関連しており、かつ前記第1データフレームの優先度が低い場合に、こ

10

20

の第1データフレームを間引くように構成されていることを特徴とするネットワークデバイス。

【請求項2】

前記第1データフレームが前記第1ステーションと前記第2ステーションとの間に接続を確立することに関連しているかを特定する際に、前記データフレーム処理ロジックは、
前記第1データフレームの少なくとも一部を読み出してこのひとつのデータフレームにシーケンス番号情報を含む送信制御プロトコルメッセージが含まれているか否かを判断するように構成されている、請求項1に記載のネットワークデバイス。

【請求項3】

前記第1データフレームの間引きを行うか否かを決定する際に、前記データフレーム処理ロジックは、

前記第1データフレームに同期フラグがセットされた送信制御プロトコルヘッダが含まれているか否かを検出するよう構成されている、請求項1に記載のネットワークデバイス。

【請求項4】

複数の受信ポートに対応する複数の受信バッファを有し、前記複数の受信バッファは、前記受信されたデータフレームを格納するよう構成され、

前記データフレーム処理ロジックは、前記第1のデータフレームの前記同期フラグがセットされている場合に、前記複数の受信バッファのうちの第1の受信バッファから前記第1のデータフレームを削除するよう構成されている、請求項3に記載のネットワークデバイス。

【請求項5】

ステーション間におけるデータフレームの通信制御を行うネットワークデバイスで実行される方法であって、

受信機が、ステーションからのデータフレームを受信し、

前記受信機が、第1のデータフレームに関連する優先度を判断し、

前記受信機が、第1のデバイスと第2のデバイスとの間の接続確立に前記第1のデータフレームが関連するか否かを判断し、

前記受信機が、前記第1のデータフレームが前記接続確立及び前記第1のデータフレームの優先度に関連するか否かに基づいて前記第1のデータフレームを間引きするか否かを判断し、

前記第1のデータフレームが前記接続確立に関連するとともに前記優先度が第1の優先度に対応する場合に、前記第1のデータフレームを間引く、方法。

【請求項6】

前記接続確立に前記第1のデータフレームが関連しているか否かの判断では、

同期フラグがセットされた送信制御プロトコルメッセージが前記第1のデータフレームに含まれるか否かが検出される、請求項5に記載の方法。

【請求項7】

ステーション間におけるデータフレームの通信制御を行うネットワークデバイスで実行される方法であって、

ステーションからデータフレームを受信し、

第1のデータフレームが、第1のデバイスと第2のデバイスとの間の接続確立に関連しているか否かを判断し、

外部メモリ内での前記受信したデータフレームを格納するための位置に対応するアドレスポイントの数が第1の閾値に達する時点を検出し、

前記第1の閾値が検出されたら第1の優先度のデータフレームについて受付制御を開始し、

前記第1のデータフレームに関連する優先度を判断し、

前記第1のデータフレームが前記接続確立に関連し、かつ前記第1のデータフレームが前記第1の優先度に対応する場合に、前記第1のデータフレームを間引くか否かを判断す

10

20

30

40

50

る、方法。

【請求項 8】

前記受信したデータフレームを少なくとも一つのバッファに格納し、
前記データフレームを間引く処理には、前記少なくとも一つのバッファから前記第 1 のデータフレームの削除が含まれる、請求項 5 記載の方法。

【請求項 9】

ステーション間におけるデータフレームの通信制御を行うネットワークデバイスであって、

データフレームを受信するよう構成された複数の受信ポートと、

処理ロジックと、を有し、この処理ロジックは、

データフレームを処理することに関連する輻輳状態を検出し、データフレームが 2 つのデバイス間に接続を確立することと関連しているか否かを判断し、

データフレームが 2 つのデバイス間に接続を確立することと関連しているか否かに少なくとも部分的に基づいて、データフレームを間引くか否かを判断し、

前記各データフレームに関連した優先度に基づいて、二つのデバイス間の接続確立に関連するデータフレームを選択的に間引くように構成されている、ネットワークデバイス。

【請求項 10】

前記二つのデバイス間の接続確立にデータフレームが関連しているか否かを判断する際に、前記処理ロジックは、

前記データフレームの少なくとも一部を読み出して、前記データフレームが、シーケンス番号情報を含む送信制御プロトコルメッセージを含むか否かを判断するように構成されている、請求項 9 記載のネットワークデバイス。

【請求項 11】

前記受信したデータフレームの少なくとも幾つかは、優先度情報を含み、前記データフレーム処理ロジックは、

データフレームに含まれた優先度情報を、高い優先度、低い優先度のすくなくとも一方にマッピングするよう構成されている、請求項 1 記載のネットワークデバイス。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、広義にはネットワーク通信に関し、特に、ネットワークデバイスにおける受付制御に関する。

【背景技術】

【0002】

コンピュータネットワークでは、多数のネットワークステーションが通信媒体を介して相互に接続されているのが一般的である。たとえば、広く用いられているローカルエリアネットワーク（LAN）スキームのひとつに、複数のステーションを共有または専用のシリアルデータバスに接続するイーサネット 802.3 がある。これらのステーションは、データバスとこのバスに接続されたステーションとの間にあるスイッチまたは他の何らかのネットワークデバイスと通信を行うことが多い。このスイッチは一般にデータの通信を制御し、データフレームを受信してそれぞれ適切な宛先に転送するためのロジックを含む。

【発明の開示】

【発明が解決しようとする課題】

【0003】

ネットワークに接続されたすべてのステーションが同時に動作して高速にデータを送信すると、スイッチでのデータトラフィックが大きくなりかねない。したがって、スイッチによっては輻輳状態の検出後にデータフレームを間引くことがある。この場合、間引かれたデータフレームは本来の宛先には転送されない。データフレームを間引く従来のスイッチに伴うひとつの欠点に、それぞれのデータフレームにおけるデータの性質を判断するこ

10

20

30

40

50

となくスイッチがデータフレームを無作為に間引いてしまう点があげられる。

【課題を解決するための手段】

【0004】

輻輳状態が起こった場合にデータフレームを選択的に間引く、ネットワークデバイス内の機構には需要がある。この需要や他の需要は、マルチポートデバイスがデータフレームを受信してデータフレームに含まれるデータのタイプを特定する本発明によって満たされる。その後、マルチポートデバイスは、マルチポートデバイスでの資源の可用性とデータフレームに含まれるデータのタイプとに基づいて、受信したデータフレームについて受付制御を行うか否かを判断する。

【0005】

本発明の一形態によれば、ステーション間でのデータフレームの通信を制御するネットワークデバイスが得られる。このネットワークデバイスは、ステーションからのデータフレームを受信するように構成された複数の受信ポートを含む。さらに、このネットワークデバイスは、受付制御が使用可能か否かを判断するデータフレーム処理ロジックも含む。データフレーム処理ロジックは、受付制御が使用可能か否かという点と受信したデータフレームに含まれる情報とに基づいて、受信したデータフレームのうちのひとつを間引くか否かも判断する。

【0006】

本発明のもうひとつの形態は、ステーション間でのデータフレームの通信を制御する、ネットワークデバイス内での方法を提供するものである。この方法は、ステーションからのデータフレームを受信し、第1の受信したデータフレームが第1のデバイスと第2のデバイスとの間での接続の確立と関連しているか否かを判断することを含む。さらに、この方法は、第1のデータフレームが接続の確立と関連している場合に第1のデータフレームを間引くことも含む。

【0007】

以下の詳細な説明から、本発明の他の利点および特徴が当業者にとって容易に理解できるものとなる。図示して説明する実施形態は、本発明を実施するために企図されるベストモードを例示するためのものである。本発明は、自明のさまざまな点で本発明から逸脱することなく変更可能である。したがって、図面は事実上例示的なものとみなされ、限定的なものとはみなされるものではない。

【発明を実施するための最良の形態】

【0008】

以下、イーサネット（IEEE 802.3）ネットワークなどのパケット交換網におけるスイッチを例をあげて本発明について説明する。しかしながら、下記において詳細に説明するように、本発明は他のパケット交換システムならびに他のタイプのシステム全般にも適用可能なものであることが明らかになる。

【0009】

スイッチアーキテクチャの概要

図1は、本発明に整合するシステムおよび方法を実施できる代表的なシステムのブロック図である。この代表的なシステムはイーサネット（IEEE 802.3）ネットワークなどのパケット交換網100を含み得る。パケット交換網100は、ネットワークステーション110と、トランスフォーマ120、トランシーバ130および140と、ネットワークノード150と、ホスト160と、外部メモリ170と、マルチポートスイッチ180とを含み得る。ネットワークステーション110は、構成の異なるコンピュータなどの従来の通信装置を含み得る。たとえば、このデバイスは、ネットワークデータ転送速度10メガビット/秒（Mb/s）または100Mb/sでデータを送受信することができる。

【0010】

10/100Mb/sのネットワークステーション110は各々、半二重イーサネットまたは全二重イーサネットのプロトコルに従って、マルチポートスイッチ180との間で

10

20

30

40

50

データを送受信することができる。イーサネットプロトコル I S O / I E C 8 8 0 2 - 3 (A N S I / I E E E 規格 8 0 2 . 3、1 9 9 3 年版) では、すべてのステーション 1 1 0 に等しくネットワークチャネルへのアクセスを許可する半二重の媒体アクセス機構を規定している。半二重環境でのトラフィックを送信媒体上で区別することはできない場合がある。むしろ、各半二重ステーション 1 1 0 に、衝突検出型搬送波検知多重アクセス (C S M A / C D) を使って送信媒体上のトラフィックに聞き耳を立てるイーサネットインタフェースカードを持たせることができる。ネットワークトラフィックがないことを検出するには、送信媒体上の受信キャリアが非アクティブな状態になったことを検知する。

【 0 0 1 1 】

送信すべきデータのあるステーション 1 1 0 はいずれも、送信媒体上の受信キャリアが非アクティブになった後にパケット間隔 (I P G) インターバルとして知られるあらかじめ定められた長さの時間だけ待機すれば、チャネルへのアクセスを試みることができる。複数のステーション 1 1 0 が同一のリンクに接続されている場合、受信キャリアが非アクティブになったことを検知して I P G インターバルが経過するとそれぞれのステーション 1 1 0 がデータを送信しようとしかねないため、衝突が発生する可能性がある。よって、送信側ステーション 1 1 0 で送信媒体を監視し、別のステーション 1 1 0 が同じリンクで同時にデータを送信したことが原因で衝突が発生していないかどうかを判断することができる。衝突が検出されると、どちらのステーション 1 1 0 も送信を中断し、それぞれランダムな待ち時間で待機し、再び送信を試みる。

【 0 0 1 2 】

全二重モードで動作する 1 0 / 1 0 0 M b / s のネットワークステーション 1 1 0 は、イーサネット規格 I E E E 8 0 2 . 3 u に従ってデータパケットの送受信を行うことができる。全二重環境では、各リンクパートナー (すなわち、1 0 / 1 0 0 M b / s のネットワークステーション 1 1 0 と対応するマルチポートスイッチ 1 8 0) 間でデータパケットを同時送受信を可能にする双方向の二地点間通信リンクが得られる。

【 0 0 1 3 】

トランスフォーマ 1 2 0 は、ネットワークステーション 1 1 0 とトランシーバ 1 3 0 との間を A C カップリングする磁気トランスを含み得る。一方、トランシーバ 1 3 0 は、それぞれのシリアル媒体独立インタフェース (S e r i a l M e d i a I n d e p e n d e n t I n t e r f a c e ; S M I I) または節約型媒体独立インタフェース (R e d u c e d M e d i a I n d e p e n d e n t I n t e r f a c e ; R M I I) を介してマルチポートスイッチ 1 8 0 と通信する 1 0 / 1 0 0 M b / s の物理層トランシーバを含み得る。各トランシーバ 1 3 0 についてはマルチポートスイッチ 1 8 0 と最大 4 ケ所のネットワークステーション 1 1 0 との間で S M I I / R M I I を介してデータパケットを送受信するように構成できる。S M I I / R M I I は、各ネットワークステーション 1 1 0 と対応するトランシーバ 1 3 0 とによるデータパケットの同時送受信を可能にできるだけの十分なデータ転送速度で動作し得るものである。

【 0 0 1 4 】

トランシーバ 1 4 0 は、たとえば高速ネットワーク送信媒体を介してネットワークノード 1 5 0 などのノードとの通信を達成する 1 0 0 0 M b / s (すなわち 1 G b / s) の物理層トランシーバを 1 つまたはそれ以上含み得る。ネットワークノード 1 5 0 は、ネットワーク速度 1 G b / s でデータパケットを送受信する 1 G b / s のネットワークノードを 1 つまたはそれ以上含み得る。また、ネットワークノード 1 5 0 は、たとえば、サーバまたは高速バックボーンネットワークへのゲートウェイを含み得る。

【 0 0 1 5 】

ホスト 1 6 0 は、マルチポートスイッチ 1 8 0 の全体としての動作を制御するための外部管理機能を提供するコンピュータデバイスを含み得る。外部メモリ 1 7 0 は、マルチポートスイッチ 1 8 0 用の外部記憶装置となる同期型スタティックランダムアクセスメモリ (S S R A M) を含み得る。外部メモリ 1 7 0 はそれぞれ、6 4 ビット幅のデータバスと 1 7 ビット幅のアドレスバスとを有する電子デバイス技術合同協議会 (J E D E C) のパ

10

20

30

40

50

イブラインバーストSSRAMまたはゼロ・バス・ターンアラウンド(ZBT)SSRAMを含み得る。外部メモリ170は、ワード長64ビットで128K単位の上位バンクおよび下位バンクとしてアドレス可能なものであってもよい。外部メモリ170のサイズは、パイプラインによって1クロックサイクルごとにデータの転送が可能な状態で少なくとも1メガバイトであると好ましい。

【0016】

マルチポートスイッチ180は、ネットワークステーション110またはネットワークノード150から受信したデータパケットを、イーサネットプロトコルなどの適切な送信プロトコルに従ってしかるべき宛先に向けて選択的に転送する。マルチポートスイッチ180を複数(ライン190を介して)まとめてカスケード接続し、マルチポートスイッチ180の機能性を拡張するようにしてもよい。

10

【0017】

図2は、本発明に整合する実施例によるマルチポートスイッチ180の詳細図である。マルチポートスイッチ180は、受信機205と、送信機210と、データバス215と、スケジューラ220と、フロー制御ロジック225と、バッファ管理ロジック230と、ポートベクトルキュー(PVQ)235と、出力制御キュー240と、内部ルールチェッカー(IRC)245と、レジスタ250と、管理情報ベース(MIB)カウンタ255と、ホストインタフェース260と、外部メモリインタフェース265と、EEPROMインタフェース270と、LEDインタフェース275と、ジョイントテストアクショングループ(Joint Test Action Group; JTAG)インタフェース280とを含み得る。

20

【0018】

受信機205は、媒体アクセス制御(MAC)モジュールと、先入れ先出し(FIFO)バッファなどの受信バッファとを含み得る。受信モジュールは、SMII、RMII、ギガビット媒体独立インタフェース(GMII)、10ビットインタフェース(TBI)、他のマルチポートスイッチ180(図1)を用いた拡張用の独自インタフェースをサポートする入力ポートを含み得る。拡張ポート(EP)を利用し、所定のプロトコルに従って他のマルチポートスイッチ180との間でデータを転送するようにしてもよい。拡張ポートを用いることで、マルチポートスイッチ180を複数まとめてカスケード接続し、バックボーンネットワークを構成することができる。受信モジュールはそれぞれ、ネットワークステーション110および/またはネットワークノード150からデータパケットを受信し、このパケットを対応する受信FIFOに格納するキューイングロジックを含み得る。キューイングロジックは、外部メモリインタフェース265を介して、パケットのうち処理すべき部分をIRC245に送信したり、格納すべき部分を外部メモリ170に送信したりすることができる。

30

【0019】

送信機210は、MACモジュールとFIFOバッファなどの送信バッファとを含み得る。送信モジュールは、SMII、GMII、TBI、他のマルチポートスイッチ180を用いた拡張用の独自インタフェースをサポートする出力ポートを含み得る。送信モジュールはそれぞれ、外部メモリ170からパケットを取得し、これらのパケットを対応する送信FIFOに格納するデキューイングロジックを含み得る。送信モジュールは、対応する送信FIFOからデータパケットを読み出したり、これらのパケットをネットワークステーション110および/またはネットワークノード150に送信したりすることができる。本発明に整合する別の実施例では、受信機205および送信機210の機能を、データパケットの受信と送信の両方を管理するトランシーバによって行うようにしてもよい。

40

【0020】

データバス215は、受信機205、送信機210、IRC245、外部メモリインタフェース265を接続する導体を1本またはそれ以上含み得る。スケジューラ220は、受信機205のキューイングロジックと送信機210のデキューイングロジックとによって外部メモリ170へのアクセスを制御するロジックを含み得る。マルチポートスイッチ

50

180はノンブロッキングスイッチとして動作するように構成されており、ネットワークデータはスイッチポートからそれぞれのワイヤレート10Mb/s、100Mb/sまたは1000Mb/sで送受信される。よって、スケジューラ220は、各ポートごとにアクセスを制御し、外部メモリ170の帯域利用を最適化することができる。

【0021】

フロー制御ロジック225は、バッファ管理ロジック230、PVQ235、出力制御キュー240と連携動作して送信機210によるパケットの送信を制御するロジックを含み得る。フロー制御ロジック225は、データのトラフィック量に応じて効率のよい方法で送信機210からパケットが出力されるように送信機210を制御することができる。バッファ管理ロジック230は、マルチポートスイッチ180内でのメモリの利用状況を監視するロジックを含み得る。たとえば、データパケットがその宛先となる出力ポート（単数または複数）に向けて送信されていれば、バッファ管理ロジック230はフレームポイントの利用状況とフレームポイントの再利用とを管理することができる。フレームポイントは、外部メモリ170に格納された、送信を必要とするデータフレームの位置を特定する。

10

【0022】

PVQ235は、データフレームの送信を受信するための出力ポートに対応した出力制御キュー240に含まれる適切な出力キュー（単数または複数）へのフレームポイントを入力するロジックを含み得る。マルチコピーフレームでは、PVQ235は同じフレームポイントのいくつかのコピーを2つ以上の出力キューに供給することができる。出力制御キュー240は、送信機210の各送信モジュールに対応するFIFO型の出力キューを含み得る。各出力キューは、フレームに対する優先度の異なる複数の順位キューを含み得る。たとえば、必要なアクセスレイテンシーが低いフレーム（マルチメディア用途用のフレームや管理フレームなど）に高い順位キューを用いるようにすればよい。FIFO型の出力キューに格納されたフレームポイントを、それぞれの送信モジュールについてのデキューイングロジックで処理することができる。デキューイングロジックは、フレームポイントを利用して外部メモリ170にアクセスし、フレームポイントによって指定されたメモリ位置でデータフレームを読み出す。

20

【0023】

IRC245は、受信機205から受信したデータパケットについてフレーム転送決定を行う内部意志決定エンジンを含み得る。IRC245はデータバス215を監視し（すなわち「覗き」）、送信元、宛先、仮想ローカルエリアネットワーク（VLAN）アドレス情報をはじめとする受信パケットのヘッダ情報など、データフレームの一部と、フレームポイント値とを判断する。IRC245は、ヘッダ情報を利用して、フレームポイントで指定された位置に格納されたデータフレームがどの出力ポートから出力されるかを判断することができる。よって、IRC245は、特定のデータフレームをひとつのポートで出力すべき（すなわちユニキャスト）であるか、複数のポートで出力すべき（すなわちマルチキャスト）であるか、すべてのポートで出力すべき（すなわちブロードキャスト）であるか、どのポートでも出力すべきではない（すなわち破棄）かを判断することができる。

30

40

【0024】

たとえば、各データフレームは送信元アドレスと宛先アドレスとを特定するヘッダを含み得る。IRC245は、宛先アドレスを利用してデータフレームを出力するのに適切な出力ポートを特定することができる。また、フレームヘッダは、フレームをネットワークステーション110のグループの1ヶ所またはそれ以上のメンバ宛の情報として特定するVLANアドレス情報も含み得る。あるいは、IRC245は、拡張ポートを介して別のマルチポートスイッチ180にデータフレームを転送すべきであることを判断できる。したがって、IRC245は、外部メモリ170に一時的に格納されたフレームを、1つの出力ポートに出力すべきか、複数の出力ポートに出力すべきか、どの出力ポートにも出力すべきではないか、あるいは別のマルチポートスイッチ180に出力すべきかを判断する

50

【 0 0 2 5 】

I R C 2 4 5 は、その転送決定を転送記述子の形で P V Q 2 3 5 に出力することができる。転送記述子は、たとえば、データフレームの優先度が高いか低いかを特定する優先度クラス、フレームを送信すべき各出力ポートを特定するポートベクトル、入力ポート番号または V L A N 情報を含み得る。P V Q 2 3 5 は、転送記述子をデコードしてフレームポインタを得ることができる。フレームポインタを入手したら、P V Q 2 3 5 はこれを出力制御キュー 2 4 0 内の適切な出力キューに供給することができる。

【 0 0 2 6 】

また、I R C 2 4 5 はレイヤ 3 フィルタリングを行うこともできる。たとえば、I R C 2 4 5 は、各受信データパケットを最大 1 2 8 のプログラム可能なパターンについて検討し、その結果に基づいてパケットを処理することができる。ここでの結果を用いて、I R C 2 4 5 によるパケットの間引き、パケットのホスト 1 6 0 への転送あるいはユーザ優先度または差別化サービスコードポイント (D S C P) のパケットへの割り当てを指示することができる。ユーザの優先度と D S C P とは独立して出力優先度クラスにマッピングが可能なものである。

【 0 0 2 7 】

レジスタ 2 5 0 は、ホストインタフェース 2 6 0 が利用するコンフィギュレーションレジスタおよびステータスレジスタを含み得る。M I B カウンタ 2 5 5 は、統計的なネットワーク情報をホスト 1 6 0 が利用する M I B オブジェクトの形で提供することができる。ホストインタフェース 2 6 0 は、マルチポートスイッチ 1 8 0 の全体としての動作をホスト 1 6 0 などの外部管理エントリによって制御できるようにする標準インタフェースを含み得る。ホストインタフェース 2 6 0 は、指定のレジスタ空間内のホストアクセスをデコードし、レジスタ 2 5 0 に対してコンフィギュレーション情報およびステータス情報を読み書きすることができる。

【 0 0 2 8 】

外部メモリインタフェース 2 6 5 は、外部メモリ 1 7 0 へのアクセスを可能にする標準インタフェースを含み得る。外部メモリインタフェース 2 6 5 は、スケジューラ 2 2 0 が決定した、割り当てられたタイムスロットの間、ダイレクトメモリアクセス (D M A) トランザクションにおける外部メモリ 1 7 0 でのパケットデータの外部記憶を可能にする。本発明に整合する実施例では、外部メモリインタフェース 2 6 5 は少なくとも 6 6 M H z のクロック周波数で動作し、好ましくは 1 0 0 M H z またはそれ以上の周波数で動作する。

【 0 0 2 9 】

E E P R O M インタフェース 2 7 0 は、E E P R O M などの別の外部メモリに対する標準インタフェースを含み得る。L E D インタフェース 2 7 5 は、外部 L E D ロジックに対する標準インタフェースを含み得る。L E D インタフェース 2 7 5 は、入力ポートと出力ポートの状態を外部 L E D ロジックに送信することができる。L E D ロジックは、人間が読み取り可能な L E D ディスプレイ素子を駆動することができる。J T A G インタフェース 2 8 0 は、たとえばマルチポートスイッチ 1 8 0 でバウンダリスキャンテストを行うことができるようにするための外部試験機器に対する標準インタフェースを含み得る。

【 0 0 3 0 】

スイッチアーキテクチャについての上記の説明は、パケット交換網におけるスイッチの動作についての概要を提供するものである。たとえばマルチポートスイッチ 1 8 0 で実現できるような本発明の特徴のさらに詳細な説明については後述する。

【 0 0 3 1 】

ネットワークデバイスにおける選択的受付制御

本発明は、マルチポートスイッチ 1 8 0 などのネットワークデバイスにおける受付制御機構を対象とするものである。本発明に整合するマルチポートスイッチ 1 8 0 は、受信したデータフレームに含まれる情報とマルチポートスイッチ 1 8 0 での資源の可用性とに基

10

20

30

40

50

づいて、受付制御手順をいつ実行すべきかを判断する。

【0032】

図3は、選択的受付制御を実施できる、本発明の代表的な実施例に整合するマルチポートスイッチ180の一部を示している。図3を参照すると、マルチポートスイッチ180は、受信機205と、バッファ管理ロジック230と、外部メモリアインタフェース265と、ポートフィルタロジック300とを含む。

【0033】

受信機205は、入力ポート1乃至Nにそれぞれ対応する、310、312、314を付したMACモジュールを含む。MACモジュールはそれぞれ、受信FIFOバッファとキューイングロジックとを含み得る。たとえば、図3を参照すると、MACモジュール310は、310Aを付した受信FIFOバッファと、キューイングロジック310Bとを含む。他のMACモジュールも同様に、受信FIFOバッファとキューイングロジックとを含み得る。

10

【0034】

バッファ管理ロジック230は、空きバッファキュー(FBQ)制御ロジック320と空きバッファキュー330とを含む。FBQ330は、外部メモリ170内のデータフレームの格納に利用できる位置を特定するフレームポインタを格納する。本発明に整合するFBQ制御ロジック320は、FBQ330で利用できるフレームポインタの数があらかじめ定められた閾値未満になったらマルチポートスイッチ180で受付制御を開始する。FBQ制御ロジック320が受付制御を開始する前にFBQ330にいくつのフレームポインタを残すかをユーザの要件に基づいて設定することができ、ユーザがプログラム可能にすることができる。すなわち、ユーザは、ネットワークの状態などに応じて、ホストインタフェース260(図2)を介して、受付制御を使用可能にする前にFBQ330に残るフレームポインタの数をプログラムすることができる。

20

【0035】

本発明に整合する代表的な実施例では、FBQ330は2種類の「水位線」(すなわち閾値)を含む。各水位線がFBQ330に残るフレームポインタのあらかじめ定められた数に対応する。詳細については後述するように、FBQ制御ロジック320はFBQ330がそれぞれの水位線のうちの一方にいつ達したかを検出し、受付制御信号をポートフィルタロジック300に送信する。

30

【0036】

詳細については後述するように、本発明の代表的な実施例に整合するポートフィルタロジック300は、受信したデータフレームに関連する優先度を判断し、受信したデータフレームに含まれるデータのタイプも判断する。受付制御が使用可能になり、あるデータフレームが特定の優先度とタイプのものであることがポートフィルタロジック300によって特定されたら、ポートフィルタロジック300はそのデータフレームを間引くための適切なMACモジュールを示すことができる。

【0037】

ポートフィルタロジック300は、受信機205とは別のコンポーネントとして図示されている。本発明の別の実施例では、ポートフィルタロジック300が受信機205の一部であってもよいことは理解されたい。さらに別の形では、ポートフィルタロジック300が実行している機能を、MACモジュールのそれぞれのキューイングロジックなどのもうひとつのロジックコンポーネントで実行してもよい。

40

【0038】

本発明の代表的な実施例では、ポートフィルタロジック300は、データフレームを優先度の高いデータフレームまたは優先度の低いデータフレームのいずれかとして分類する。優先度の高いフレームは、管理デバイス用のフレームやマルチメディア用途向けのフレームなど、必要なアクセスレイテンシーが低いフレームを含み得る。優先度の低いフレームは、他のあらゆるフレームを含み得る。

【0039】

50

本発明の別の実施例では、データフレームに関連する優先度の数を2よりも大きくすることができる。たとえば、ポートフィルタロジック300は、低、中、高などの3つの優先度レベルのうちの1つを持つものとしてフレームを特定することができる。もうひとつの実施例では、ポートフィルタロジック300は、低、低-中、中-高、高などの4つの優先度レベルのうちの1つを持つものとしてフレームを特定することができる。これらの場合、FBQ330における水位線の数を優先度のレベルに対応させるようにしてもよい。たとえば、マルチポートスイッチ180が4つのレベルの優先度をサポートする状況では、FBQ330における水位線の数を4にすることができる。

【0040】

また、本発明のいくつかの実施例では、優先度表示のあるデータフレームをマルチポートスイッチ180で受信することができる。たとえば、イーサネットフレームは、8つのレベルの優先度のうちの1つを表す3ビットフィールドを含み得る。この場合、ポートフィルタロジック300は、受信した優先度情報を、マルチポートスイッチ180がサポートする、対応の優先度レベルにマッピングすることができる。たとえば、ポートフィルタロジック300は、8つのレベルの優先度をマルチポートスイッチ180の8つの高い優先度または低い優先度にマッピングすることができる。あるいは、受信したデータフレームに関連する8つのレベルの優先度を、マルチポートスイッチ180がサポートする優先度のレベルに応じてマルチポートスイッチ180の3つまたはそれ以上のレベルの優先度にマッピングすることができる。

【0041】

ポートフィルタロジック300は、データフレームのデータ部分に含まれるデータのタイプも特定する。たとえば、図4は本発明に整合する実施例においてネットワークステーション110およびネットワークノード150で利用できる通信プロトコル400を示す代表的な図である。通信プロトコル400は、アプリケーションレイヤ410と、トランスポートレイヤ420と、ネットワークレイヤ430と、データリンクレイヤ440とを含む。アプリケーションレイヤ410は、通信を開始するアプリケーションまたはプログラムを含み得る。アプリケーションレイヤ410は、メール、ファイル転送、リモートアクセス、認証、名前解決のためのプロトコルなど、いくつかのアプリケーションレイヤプロトコルを含み得る。アプリケーションレイヤ410は、データパケットをパケットの宛先と一緒にトランスポートレイヤ420に渡すことで、データ412をネットワーク100に配信することができる。

【0042】

トランスポートレイヤ420は、パケットの送信元と宛先との間に仮想回路、接続またはソケットを確立することができる。トランスポートレイヤ420は、関連の転送プロトコルの仕様に従って、ヘッダ422と後書き424とをパケットに付すことができる。トランスポートレイヤ420は、パケットを宛先のインターネットプロトコル(IP)アドレスと一緒にネットワークレイヤ430に渡す。

【0043】

ネットワークレイヤ430は、トランスポートレイヤ420からのパケットを受け取り、IPアドレスをMACアドレスなどの物理アドレスに変換し、必要に応じてパケットを所望のサイズに分割することでデータリンクレイヤ440用のパケットを作成することができる。ネットワークレイヤ430は、関連のプロトコルの仕様に従ってトランスポートレイヤ420からのパケットにIPヘッダ432と後書き434とを付すことで、データグラムと呼ばれるパケットを生成することができる。ネットワークレイヤ430はこのデータグラムをデータリンクレイヤ440に渡す。

【0044】

データリンクレイヤ440は、ネットワーク100上での高信頼度でのデータ移動を担うイーサネットなどのデータリンクプロトコルを含み得る。データリンクレイヤ440は、データグラムをその独自のフォーマットに変換する。この変換には、送信元アドレスおよび宛先MACアドレスを含むヘッダ442と、後書き444(チェックサムデータを含

10

20

30

40

50

む)とを加えることを含み得る。

【0045】

図5～図7はそれぞれ本発明に整合する実施例におけるヘッダ422～442の詳細図である。図5に示すように、ヘッダ422には、送信元ポートフィールドと、宛先ポートフィールドと、シーケンス番号フィールドと、確認応答番号フィールドと、データオフセット(offest)フィールドと、予約フィールドと、フラグフィールドと、ウィンドウフィールドと、チェックサムフィールドと、緊急ポインタフィールドと、が含まれる。送信元ポートフィールドおよび宛先ポートフィールドにはデータパケットの送信元および宛先を特定するデータが含まれる。シーケンス番号フィールドには宛先でのパケットの配信と順番とを確認するのに使われるデータが含まれる。確認応答番号フィールドには、送信元が宛先から受信するつもりでいるデータの次のバイトを特定するデータが含まれる。

10

【0046】

データオフセットフィールドには、ヘッダ422の長さを特定するデータが含まれる。予約フィールドは将来的な拡張に利用できるものである。フラグフィールドは、緊急フラグ、確認応答フラグ、プッシュフラグ、リセットフラグ、同期フラグ、送信終了フラグなどのいくつかのフラグを含み得る。緊急フラグは緊急ポインタフィールドに含まれるデータが有効か否かを示す。確認応答フラグは確認応答番号フィールドに含まれるデータが有効か否かを示す。プッシュフラグは、付随するデータをすぐに宛先のアプリケーションに渡すべきか否かを示す。リセットフラグは接続をリセットすべきか否かを示す。同期(SYN)フラグはシーケンス番号について最初に合意を得るのに用いられる。送信終了(FIN)フラグは送信元がデータの送信を終了したか否かを示す。

20

【0047】

ウィンドウフィールドには、確認されていないデータの格納に宛先側で利用できる空間の量を示すデータが含まれる。チェックサムフィールドには、ヘッダ422とデータ412の両方を対象にできるチェックサム値が含まれる。緊急ポインタフィールドには、このパケットを通常のデータストリームよりも優先させるべきか否かを示すデータが含まれる。

【0048】

図6を参照すると、ヘッダ432には、バージョンフィールドと、ヘッダ長フィールドと、サービスタイプフィールドと、全長フィールドと、識別フィールドと、フラグフィールドと、フラグメントオフセットフィールドと、有効期間フィールドと、プロトコルフィールドと、ヘッダチェックサムフィールドと、送信元アドレスと、宛先アドレスとが含まれる。バージョンフィールドには使用するプロトコルのバージョンを特定するデータが含まれる。ヘッダ長フィールドにはヘッダ432の長さを特定するデータが含まれる。サービスタイプフィールドにはネットワーク100によってデータグラムに与えられるサービスの品質を特定するデータが含まれる。

30

【0049】

全長フィールドにはデータグラム(すなわちヘッダ422および432とデータ412)の長さを特定するデータが含まれる。識別フィールドには分割されたデータグラムを宛先で再構成するのに使われるデータが含まれる。フラグフィールドは、たとえば、データグラムを分割できるか否かを特定するフラグを1つまたはそれ以上含み得る。フラグメントオフセットフィールドには、元の送信パケットの先頭からどの位置が開始点になるかを特定するデータが含まれる。

40

【0050】

有効期間フィールドには、パケットが経由できるホップ/リンクの数を特定するカウント値が含まれる。プロトコルフィールドには、宛先のデータグラムで用いられる転送プロトコルのタイプ(インターネットコントロールメッセージプロトコル(ICMP)、インターネットグループマネジメントプロトコル(IGMP)、送信制御プロトコルまたはユーザデータグラムプロトコル(UDP)など)を特定するデータが含まれる。ヘッダチェックサムフィールドには、ヘッダ432を対象とするチェックサム値が含まれる。送信元

50

アドレスフィールドおよび宛先アドレスフィールドにはそれぞれ送信元と宛先のインターネットアドレスなどのアドレスが含まれる。

【 0 0 5 1 】

図7を参照すると、ヘッダ442には、宛先アドレスフィールドと、送信元アドレスフィールドと、タイプ/長さコードフィールドとが含まれる。本発明の代表的な実施例では、送信元アドレスフィールドおよび宛先アドレスフィールドにはそれぞれ送信元および宛先のイーサネットアドレスが含まれる。タイプ/長さコードフィールドには、送信元と宛先とを接続するのに用いられるデータリンクプロトコルを特定するデータが含まれ、また、データフレーム長を特定するデータが含まれる。上述したように、ヘッダ442は、データフレームに関連する優先度を示す優先度フィールドを含むものであってもよい。さらに、ヘッダ442は、仮想ローカルエリアネットワーク(VLAN)情報を含み得る。

10

【 0 0 5 2 】

ポートフィルタロジック300は、受信したデータフレームのペイロード部分に含まれる情報に基づいて、データフレームのタイプを特定することができる。データフレームのペイロード部分にはヘッダ422および432と、データ412と、後書き424および434(図4)とが含まれることは理解されたい。たとえば、データフレームのペイロード部分は、送信制御プロトコル(TCP)セッションに関連する情報を含み得る。代表的なTCPセッションでは、2つのネットワークステーション110間またはネットワークステーション110とネットワークノード150との間などの2つのデバイス間でのデータのやり取りを、1)セットアップパケット、2)データパケット、3)切断パケットの3種類のパケットに分けることができる。セットアップパケットはTCPヘッダのSYNフラグがセットされているときに特定できるものである。言い換えると、ヘッダ422のフラグフィールドのSYNフラグ(図4)が立っていれば、パケットのペイロード部分が2つのデバイス間での接続の確立に関連していることになる。一般的な流れでは、ネットワークステーション110などの送信元デバイスがまずは初期シーケンス番号(ISN_CLIENT)を含むセットアップパケットを送信する。このシーケンス番号は、データストリームの中にある個々のデータを特定するものである。ネットワークノード150などの宛先デバイスは一般に、クライアントのシーケンス番号を確認してサーバの初期シーケンス番号(ISN_SERVER)を含む確認応答パケットを送信元に送り返す。これに対して送信元デバイスはサーバのシーケンス番号を確認するパケットで応答してもよい。

20

30

【 0 0 5 3 】

これらのセットアップパケットによって接続が確立されると、2つのデバイス間で実際にデータのやり取りをするデータパケットを1つまたはそれ以上送り出すことができる。データパケット(単数または複数)の後にくるのは切断パケットである。この切断パケットはTCPヘッダのFINフラグがセットされているときに確認される。言い換えると、ヘッダ422のフラグフィールドのFINフラグが立っていれば、送信元からのデータの送信が終了しており、TCPセッションが効率よく終了することになる。

【 0 0 5 4 】

本発明に整合するポートフィルタロジック300は、ヘッダ422のフラグフィールドにあるSYNフラグの値に基づいてセットアップパケットを特定することができる。詳細については後述するように、ポートフィルタロジック300は輻輳状態が発生した場合にセットアップパケットを選択的に間引くことができる。

40

【 0 0 5 5 】

図8は、本発明に整合する実施例におけるマルチポートスイッチ180による代表的な処理を示している。処理についてはネットワーク100(図1)の起動時に開始することができる。ここでは、ネットワークステーション110がデータフレームを送信し、マルチポートスイッチ180がデータフレームを受信すると仮定する[ステップ810]。たとえば、MACモジュール310がデータフレームを受信すると仮定する。この場合、受信FIFOバッファ310Aがデータフレームを格納する[ステップ820]。

50

【 0 0 5 6 】

ポートフィルタロジック 3 0 0 は受信したデータフレームの一部を調べ、受信したデータフレームに関連する優先度を判断する [ステップ 8 3 0]。本発明の代表的な実施例では、ポートフィルタロジック 3 0 0 は、データフレームが受信 F I F O バッファ 3 1 0 A に格納されている間に受信したデータフレームのヘッダ (すなわち図 4 のヘッダ 4 4 2) を調べる。

【 0 0 5 7 】

上述したように、ポートフィルタロジック 3 0 0 は、フレームのタイプまたはフレームの宛先に基づいて優先度を特定することができる。たとえば、マルチメディア用途向けのデータフレームまたは管理デバイス用のデータフレームを優先度の高いフレームとして指定することができる。他のすべてのデータフレームについては優先度が低いフレームまたは通常の優先度のフレームとして指定すればよい。あるいは、ポートフィルタロジック 3 0 0 は、個々のユーザの要件などの他の基準に基づいて優先度を特定するものであってもよい。もうひとつの実施例では、ポートフィルタロジック 3 0 0 は、上述したようにデータフレームと一緒に送信される優先度情報に基づいて優先度を特定するものであってもよい。この状況では、ポートフィルタロジック 3 0 0 は、受信した優先度情報をマルチポートスイッチ 1 8 0 がサポートする、対応の優先度レベルにマッピングすることができる。

【 0 0 5 8 】

いずれの場合も、データフレームに関連する優先度を特定することに加え、ポートフィルタロジック 3 0 0 はデータフレームのペイロード部分に含まれるデータの性質を判断する [ステップ 8 4 0]。たとえば、ポートフィルタロジック 3 0 0 は、データフレームがもうひとつのデバイスとの新たな接続の確立と関連しているか否かを判断することができる。上述したように、イーサネットフレームは、そのペイロード部分に、対応のデータ (すなわち図 4 のデータ 4 1 2) を、シーケンス番号を同期させるなどの T C P セッションのセットアップに関与するセットアップパケットとして特定する T C P パケット用ヘッダ (すなわち図 5 のヘッダ 4 2 2) を含み得る。この場合、ヘッダ 4 2 2 のフラグフィールドにある S Y N フラグを設定してもよい。また、パケットがデバイス間でのシーケンス番号の確認応答に関連している場合にヘッダ 4 2 2 のフラグフィールドにある確認応答フラグを設定してもよいことに注意されたい。

【 0 0 5 9 】

ポートフィルタロジック 3 0 0 がデータをセットアップパケットとして特定する (S Y N フラグを立てるなど) と仮定する。この場合、ポートフィルタロジック 3 0 0 は受付制御が使用可能であるか否かを判断することができる [ステップ 8 5 0]。代表的な実施例では、ポートフィルタロジック 3 0 0 は F B Q 制御ロジック 3 2 0 からの信号がアサートされているか否かに基づいて、受付制御が使用可能であるか否かを判断する。たとえば、F B Q 3 3 0 で利用できるフレームポインタの数が第 1 の水位線未満であることを F B Q 制御ロジック 3 2 0 によって判断すると仮定する。この場合、F B Q 制御ロジック 3 2 0 は、A D M I S S I O N _ C O N T R O L _ L O W などのイネーブル受付制御信号をポートフィルタロジック 3 0 0 に送信する。これによって、ポートフィルタロジック 3 0 0 が選択した優先度の低いデータフレームについて受付制御手順を実行すべきであることを示すことができる。

【 0 0 6 0 】

受付制御が使用可能ではない場合、マルチポートスイッチ 1 8 0 でデータフレームを処理する [ステップ 8 6 0]。すなわち、上記の例では、キューイングロジック 3 1 0 B が F B Q 3 3 0 にアクセスして外部メモリ 1 7 0 における位置を特定するフレームポインタを取得し、受信 F I F O バッファ 3 1 0 A に格納されたデータフレームとフレームポインタとを外部メモリインタフェース 2 6 5 に転送する。次に、外部メモリインタフェース 2 6 5 は、このデータフレームを外部メモリ 1 7 0 内のフレームポインタによって特定された位置に転送する。次にマルチポートスイッチ 1 8 0 がデータフレームを処理してフレーム転送情報を生成し、外部メモリ 1 7 0 からデータフレームを検索、このデータフレーム

10

20

30

40

50

を意図した宛先（単数または複数）に転送する。

【 0 0 6 1 】

受付制御が使用可能であることがポートフィルタロジック 3 0 0 によって判断されると、ポートフィルタロジック 3 0 0 はデータフレームを間引くか否かを判断する [ステップ 8 7 0]。本発明の代表的な実施例では、ポートフィルタロジック 3 0 0 は、データフレームの優先度と、データフレームのペイロードに含まれる特定されたデータのタイプとに基づいて、データフレームを間引くか否かを判断する。たとえば、ポートフィルタロジック 3 0 0 によってデータフレームが優先度の低いものとして特定され、優先度の低いフレームに対する受付制御が使用可能な場合、ポートフィルタロジック 3 0 0 はセットアップパケットを含むフレームを間引くことができる。セットアップパケットを間引くことで、マルチポートスイッチ 1 8 0 は、「先行する」TCPセッションに関連するデータ送信（すなわちTCPセッションのセットアップ部分がすでに完了しているセッション）を、これらのデータのやり取りを間引くことなく完了させることができる。

10

【 0 0 6 2 】

あらかじめ定められた時間にわたって輻輳状態のまま（すなわちFBQ 3 3 0 が第1の水位線未満）である場合、マルチポートスイッチ 1 8 0 は優先度の低いデータフレームをすべて間引いてもよい。しかしながら、セットアップパケットを最初に間引けるようにすることで、マルチポートスイッチ 1 8 0 は先行するTCPセッションに対して完了の機会を選択的に与えることができる。また、セットアップパケットが間引かれるため、宛先デバイスがマルチポートスイッチ 1 8 0 を介して確認応答セットアップパケットを送信元に送り返すことはなく、送信元デバイスが確認応答セットアップパケットを宛先デバイスに送り返すこともない。これはマルチポートスイッチ 1 8 0 での輻輳を低減する上でさらに助けとなり得るものである。

20

【 0 0 6 3 】

別の実施例では、ポートフィルタロジック 3 0 0 は、受付制御が使用可能なときにユーザの要件に基づいて他の優先度のデータフレームを間引くことを判断してもよい。たとえば、FBQ 3 3 0 で利用できるフレームポイントの数が第2の水位線未満であることをFBQ制御ロジック 3 2 0 によって判断すると仮定する。この場合、FBQ制御ロジック 3 2 0 は、ADMISSION__CONTROL__HIGHなどのイネーブル受付制御信号をポートフィルタロジック 3 0 0 に送信することができる。これによって、ポートフィルタロジック 3 0 0 が優先度の低いフレームをすべて間引き、選択した優先度の高いフレームを間引くことを示すことができる。すなわち、ポートフィルタロジック 3 0 0 はデータフレームのペイロード部分に含まれるデータの性質を特定 [ステップ 8 4 0] した後、優先度の高いデータフレームを間引くか否かを判断する [ステップ 8 7 0]。たとえば、優先度の高いデータフレームがセットアップパケットに含まれるものとして特定されると、マルチポートスイッチ 1 8 0 はこのデータフレームを間引くことができる。しかしながら、マルチポートスイッチ 1 8 0 はセットアップパケット以外のデータパケットを処理し、既存のTCPセッションを完了させることができる。上述したように、確認応答パケットおよびTCPセッションの確立に関連する他のパケットが送信されないため、セットアップパケットを間引くとマルチポートスイッチ 1 8 0 でネットワーク上のデータトラフィックをさらに低減できることがある。繰り返すが、輻輳状態が長引く場合は優先度の高いデータフレームを含む全受信データフレームをポートフィルタロジック 3 0 0 で間引くようにしてもよい。

30

40

【 0 0 6 4 】

要するに、ポートフィルタロジック 3 0 0 はFBQ 3 3 0 の状態とデータフレームに含まれるデータの性質とに基づいてデータフレームを間引く決断をし得る。このポートフィルタロジック 3 0 0 は、データパケットの性質を判断する上でデータフレームの一部または全部を調べるように構成できるものである。すなわち、上記の例では、ヘッダ 4 2 2 に含まれる 1 つまたはそれ以上のフラグの状態を特定し、データパケットの性質を特定するようにポートフィルタロジック 3 0 0 を構成することができる。

50

【 0 0 6 5 】

ポートフィルタロジック 3 0 0 がデータフレームを間引く決断をした場合、ポートフィルタロジック 3 0 0 はデータフレームを受信した対応の M A C モジュールに信号を送信する [ステップ 8 8 0]。上記の例では、ポートフィルタロジック 3 0 0 は、データフレームが間引かれることを示す信号を M A C モジュール 3 1 0 に送信する。代表的な実施例では、ポートフィルタロジック 3 0 0 はデータフレームが「クズ」(すなわち不完全なデータフレームまたはエラーを含むデータフレーム)であることを示す信号を M A C モジュールに送信することができる。実際にはデータフレームがクズではないとしても、である。この状態で、クズのフレームが間引かれることを認識するように M A C モジュールを構成してもよい。こうすることで、M A C モジュールをクズとして特定されたフレームを間引くように構成し得るようになったため、M A C モジュールのロジックを都合よく簡略化できる。あるいは、ポートフィルタロジック 3 0 0 はデータフレームが間引かれることを示す他の何らかの信号 (D R O P _ F R A M E など) を送信するものであってもよい。

10

【 0 0 6 6 】

いずれの場合も、M A C モジュール 3 1 0 は、ポートフィルタロジック 3 0 0 からの信号を受信し、データフレームを間引く [ステップ 8 9 0]。この場合、キューイングロジック 3 1 0 B が受信 F I F O バッファ 3 1 0 A からデータフレームを削除する。都合のよいことに、データフレームは、マルチポートスイッチ 1 8 0 の資源を相当な量で使うことなくマルチポートスイッチ 1 8 0 の「フロントエンド」(すなわち処理の早い段階)で間引かれる。たとえば、キューイングロジック 3 1 0 B は F B Q 3 3 0 からフレームポインタを取得せず、データフレームを外部メモリアインタフェース 2 6 5 に転送しない。これによって、フレームポインタを取得し、続いてフレームポインタを F B Q 3 3 0 に返すことに関わるかなりの処理時間が節約される。言い換えると、マルチポートスイッチ 1 8 0 は特定のデータフレームを本来の宛先に転送することに関するすべての処理を停止する。

20

【 0 0 6 7 】

以上、ネットワークデバイスにおいて選択的受付制御を行うための装置および方法について説明した。本発明のひとつの利点に、受付制御が使用可能なときにマルチポートスイッチ 1 8 0 でデータフレームを選択的に間引くことができる点がある。本発明のもうひとつの利点は、2つのデバイス間に接続を確立することに関連するフレームをマルチポートスイッチ 1 8 0 で間引くことができるため、先行するセッションに完了する機会が与えられる点がある。本発明のさらに別の利点はデータフレームの処理の比較的早い段階で受信したデータフレームを間引くか否かをマルチポートスイッチ 1 8 0 で判断できる点である。これによって後で間引かれるデータフレームのためにマルチポートスイッチが資源を使ってしまうのを回避できるため、全体としてのデータスループットが向上する。

30

【 0 0 6 8 】

本願開示においては、本発明の好ましい実施形態ならびにその汎用性を示すいくつかの例についてのみ図示して説明したが、本発明はさまざまな他の組み合わせおよび環境における使用が可能であり、本願明細書にて明示した発明概念の範囲内でいくつかの改変が可能なものであることは理解できよう。

【 0 0 6 9 】

たとえば、ここでは2つのデバイス間での T C P セッションに関連するセットアップパッケージがデータフレームに含まれるか否かに基づいてデータフレームを間引くことを中心に本発明を説明したが、本発明はデータフレームを間引くか否かの判断に他の規準を用いることのできる状況にも利用できるものである。すなわち、ポートフィルタロジック 3 0 0 を、データフレームのペイロード部分に関連する他の何らかの条件を検出し、ユーザの要件に応じてデータフレームを間引くように構成し得る。

40

【 0 0 7 0 】

また、データフレームに関連する優先度を判断し、データフレームの優先度とデータフレームにセットアップパッケージが含まれるか否かとに基づいてデータフレームを間引くか否かを決定することについて本発明を説明してきたが、本発明はデータフレームを間引く

50

か否かを判断するのにデータフレームの優先度を使わない状況にも利用できるものである。すなわち、ポートフィルタロジック300は、データフレームに関連する優先度とは無関係にデータフレームにセットアップパッケージが含まれるか否かとFBQ330でのフレームポインタの可用性とに基づいてデータフレームを間引くか否かを判断するものであってもよい。

【0071】

最後に、ここでは優先度の低い受付制御閾値と優先度の高い受付制御閾値とに対応するFBQ330での第1の水位線と第2の水位線とを中心に本発明について説明した。しかしながら、本発明は上述したようにFBQ330に任意の数の水位線が含まれる状況にも利用できるものであることは理解されたい。この場合、マルチポートスイッチ180は、FBQ330で達したレベルに優先度が対応するデータフレームのセットアップパッケージを最初に間引く。言い換えると、FBQに4つの水位線が含まれ、優先度が低-中のデータフレームに第2の水位線が対応している場合、マルチポートスイッチ180はセットアップパッケージ情報を含む優先度が低-中のデータフレームを間引くことになる。

【図面の簡単な説明】

【0072】

添付の図面を参照するが、図中、同一の参照符号を付した要素は全体を通して同様の要素を表すものとする。

【図1】本発明に整合する方法およびシステムを実施できる、代表的なシステムのブロック図である。

【図2】図1に示すマルチポートスイッチの代表的な詳細ブロック図である。

【図3】本発明の実施例に整合する、図2に示すマルチポートスイッチの一部の代表的な詳細ブロック図である。

【図4】本発明に整合する実施例において用いられる通信プロトコルの代表的な図である。

【図5】本発明の実施例に整合する、図4に示すヘッダの代表的な詳細図である。

【図6】本発明の実施例に整合する、図4に示すヘッダの代表的な詳細図である。

【図7】本発明の実施例に整合する、図4に示すヘッダの代表的な詳細図である。

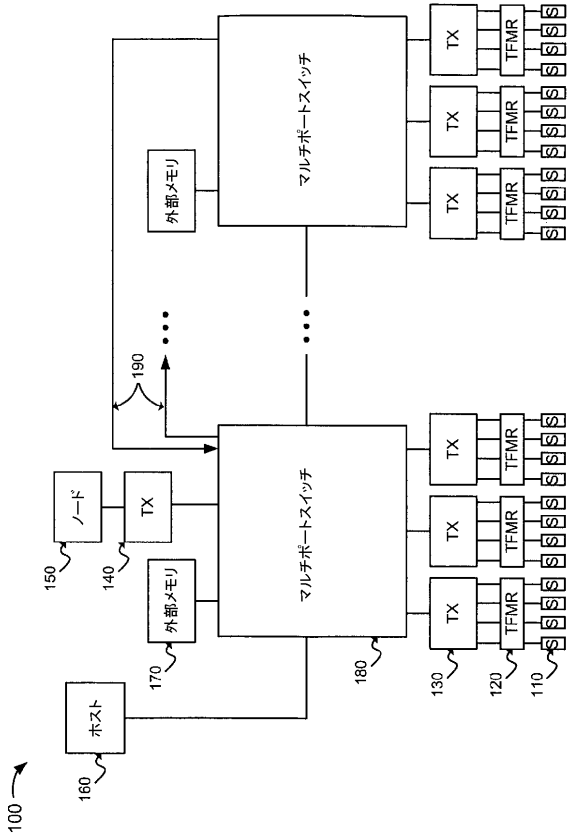
【図8】本発明に整合する代表的な実施例による、図2に示すマルチポートスイッチによって実行される処理を示す流れ図である。

10

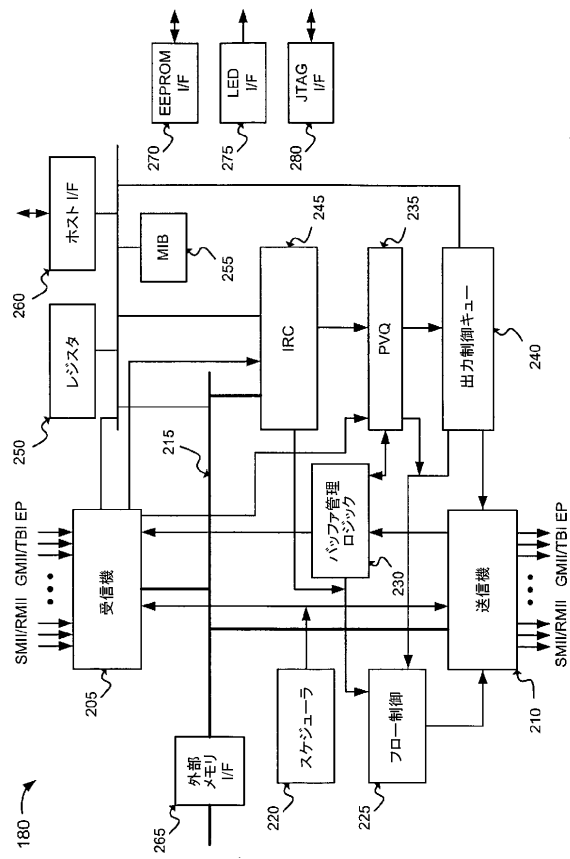
20

30

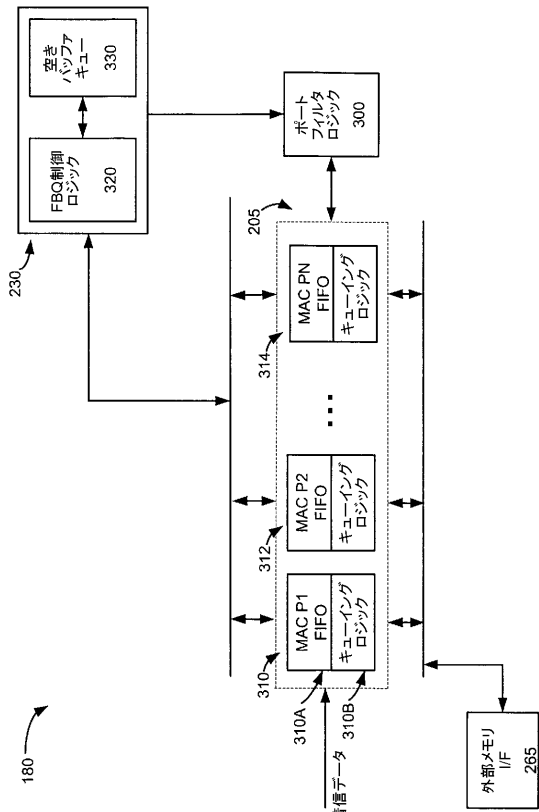
【図 1】



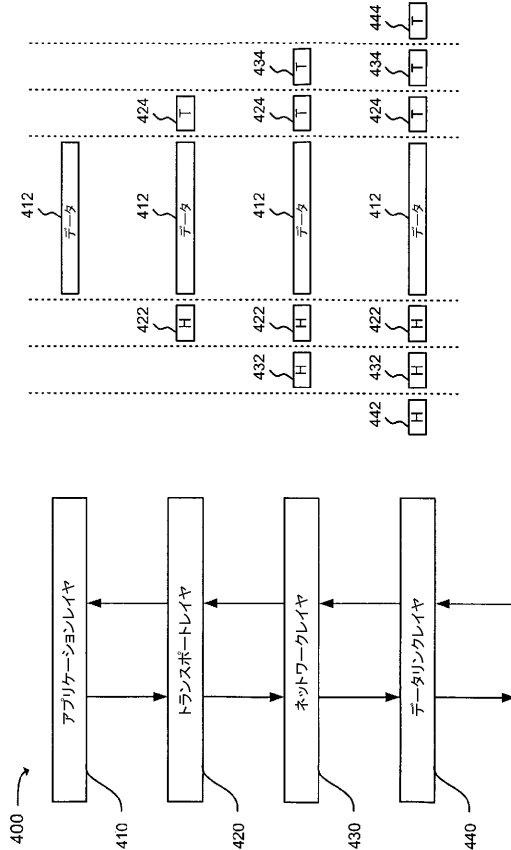
【図 2】



【図 3】



【図 4】



【 図 5 】

422

宛先ポート		送信元ポート	
シーケンス番号			
確認応答番号			
データ オフセット	予約	フラグ	ウィンドウ
チェックサム		緊急ポインタ	

【 図 6 】

432

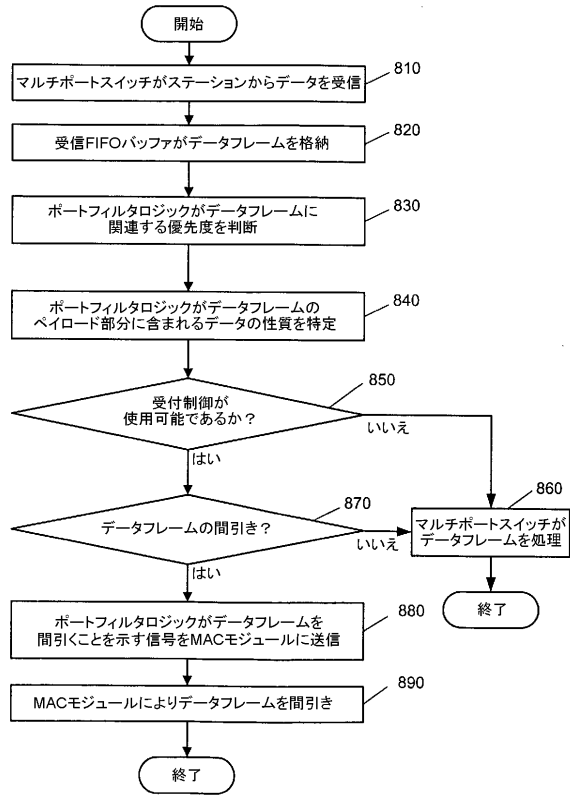
バージョン	HL	サービスタイプ	全長
識別		フラグ	フラグメントオフセット
有効期間	プロトコル		ヘッダチェックサム
送信元アドレス			
宛先アドレス			

【 図 7 】

442

宛先アドレス	送信元アドレス	タイプ/長さ
--------	---------	--------

【 図 8 】



フロントページの続き

(74)代理人 100108604

弁理士 村松 義人

(72)発明者 シャ - ジー ツェン

アメリカ合衆国、カリフォルニア州 94555、フレモント、ミモザ テラス 34263

(72)発明者 ピーター カ - ファイ チョウ

アメリカ合衆国、カリフォルニア州 95131、サン ノゼ、マクスウェル ウェイ 1416

審査官 福岡 裕貴

(56)参考文献 特開2000 - 349812 (JP, A)

欧州特許出願公開第00932282 (EP, A1)

特開2000 - 004250 (JP, A)

特開平11 - 146013 (JP, A)

特開2000 - 151593 (JP, A)

特開平07 - 307762 (JP, A)

特開2000 - 224218 (JP, A)

(58)調査した分野(Int.Cl., DB名)

H04L 12/00-66