

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号
特許第4767471号
(P4767471)

(45) 発行日 平成23年9月7日(2011.9.7)

(24) 登録日 平成23年6月24日(2011.6.24)

(51) Int.Cl.
H04L 12/40 (2006.01)

F I
H04L 12/40 A

請求項の数 13 (全 14 頁)

| | | | |
|---------------|-------------------------------|-----------|---------------------|
| (21) 出願番号 | 特願2001-565591 (P2001-565591) | (73) 特許権者 | 591016172 |
| (86) (22) 出願日 | 平成12年9月1日(2000.9.1) | | アドバンスト・マイクロ・ディバイシズ・ |
| (65) 公表番号 | 特表2003-526279 (P2003-526279A) | | インコーポレイテッド |
| (43) 公表日 | 平成15年9月2日(2003.9.2) | | ADVANCED MICRO DEVI |
| (86) 国際出願番号 | PCT/US2000/024026 | | CES INCORPORATED |
| (87) 国際公開番号 | W02001/067686 | | アメリカ合衆国、94088-3453 |
| (87) 国際公開日 | 平成13年9月13日(2001.9.13) | | カリフォルニア州、サニibel、ピー・ |
| 審査請求日 | 平成19年8月22日(2007.8.22) | | オウ・ボックス・3453、ワン・エイ・ |
| (31) 優先権主張番号 | 09/519,848 | | エム・ディ・プレイス、メール・ストップ |
| (32) 優先日 | 平成12年3月6日(2000.3.6) | | ・68 (番地なし) |
| (33) 優先権主張国 | 米国 (US) | (74) 代理人 | 100064746 |
| | | | 弁理士 深見 久郎 |
| | | (74) 代理人 | 100085132 |
| | | | 弁理士 森田 俊雄 |

最終頁に続く

(54) 【発明の名称】 ネットワークスイッチにおける選択的アドレステーブルエイジング

(57) 【特許請求の範囲】

【請求項 1】

統合ネットワークスイッチにおける方法であって、
受信したレイヤ2データパケットから予め規定されたネットワークアプリケーションに
対するアプリケーション状態を判断するステップと、

判断したアプリケーション状態に基づいて、受信したレイヤ2データパケットのソース
またはレイヤ2データパケットのデスティネーションの少なくとも1つを指定するネット
ワークスイッチアドレステーブルからアドレスエントリを選択的に削除するステップと、

予め規定されたネットワークアプリケーションのそれぞれの利用可能なアプリケーショ
ン状態からアプリケーション状態を識別するよう構成された複数のテンプレートを、受信
レイヤ2データパケットを受信したネットワークスイッチポート内にストアするステップ
とを含み、各テンプレートは、前記複数の利用可能なアプリケーション状態のうちの対応
する1つの識別を可能にする、予め規定された1つ以上の一連のデータパターンを有して
おり、

前記ストアするステップは、前記ネットワークスイッチポートが、受信したレイヤ2デ
ータパケットからの前記複数の利用可能なアプリケーション状態のうちの最初の1つを識
別することに応答して、前記複数のテンプレートをストアするステップを含み、

前記方法は、

利用可能なアプリケーション状態の1つの検出によって判断されかつこれに応答して開
始されるアプリケーション特有エイジング間隔の後に、アプリケーション特有エイジング

10

20

間隔の間の判断されたアドレスエントリのインアクティビティに基づいて、アドレスエントリをネットワークスイッチアドレステーブルから削除するステップをさらに含む、方法。

【請求項 2】

前記判断するステップは、前記受信したレイヤ 2 データパケットを受信したネットワークスイッチポートが、受信したレイヤ 2 データパケットの少なくとも一部と、それぞれの利用可能な予め規定されたネットワークアプリケーションを識別するために用いられるフレームデータを指定する複数のテンプレートとを同時に比較することに基づいて、前記受信したレイヤ 2 データパケットを受信したネットワークスイッチポートが、複数の利用可能な予め規定されたネットワークアプリケーションから、予め規定されたネットワークアプリケーションに対するアプリケーション状態を判断するステップを含む、請求項 1 に記載の方法。

10

【請求項 3】

複数のネットワークポートスイッチを含む統合ネットワークスイッチにおける方法であって、

受信したレイヤ 2 データパケットから予め規定されたネットワークアプリケーションに対するアプリケーション状態を判断するステップと、

判断したアプリケーション状態に基づいて、受信したレイヤ 2 データパケットのソースまたはレイヤ 2 データパケットのデスティネーションの少なくとも 1 つを指定するネットワークスイッチアドレステーブルからアドレスエントリを選択的に削除するステップとを含む、

20

前記判断するステップは、前記受信したレイヤ 2 データパケットを受信したネットワークスイッチポートが、受信レイヤ 2 データパケットの少なくとも一部と、それぞれの利用可能な予め規定されたネットワークアプリケーションを識別するために用いられるフレームデータを指定する複数のテンプレートとを同時に比較することに基づいて、前記受信したレイヤ 2 データパケットを受信したネットワークスイッチポートが、複数の利用可能な予め規定されたネットワークアプリケーションから、予め規定されたネットワークアプリケーションに対するアプリケーション状態を判断するステップを含む、方法。

【請求項 4】

予め規定されたネットワークアプリケーションのそれぞれの利用可能なアプリケーション状態からアプリケーション状態を識別するよう構成された複数のテンプレートを、受信レイヤ 2 データパケットを受信したネットワークスイッチポート内にストアするステップをさらに含み、各テンプレートは、前記複数の利用可能なアプリケーション状態のうちの対応する 1 つの識別を可能にする、予め規定された 1 つ以上の一連のデータパターンを有する、請求項 3 に記載の方法。

30

【請求項 5】

前記ストアするステップは、前記ネットワークスイッチポートが、受信したレイヤ 2 データパケットからの前記複数の利用可能なアプリケーション状態のうちの最初の 1 つを識別することに応答して、前記複数のテンプレートをストアするステップを含む、請求項 4 に記載の方法。

40

【請求項 6】

選択的に削除するステップは、予め規定されたネットワークアプリケーションに対するセッションの終了を指定する判断されたアプリケーション状態に基づいて、アドレスエントリを削除するステップを含む、請求項 1 または 3 に記載の方法。

【請求項 7】

検出するステップは、受信レイヤ 2 データパケットの受信の間に、予め規定されたネットワークアプリケーションに対するセッションの終了を識別するために用いられるフレームデータを指定する、受信レイヤ 2 データパケットの選択された部分を複数の極小項と比較するステップと、

受信レイヤ 2 データパケットの選択された部分がそれぞれの極小項と整合するかどうか

50

を判断するステップと、

判断するステップに基づいて、比較結果を生成するステップとを含む、請求項 6 に記載の方法。

【請求項 8】

アプリケーション状態を判断するステップに応答して、アドレスエントリに対するアプリケーション特有エイジング間隔をカウントするよう構成されるアプリケーション特有エイジングタイマを開始させるステップと、

アプリケーション特有エイジング間隔の満了の際に、アドレスエントリがアクセスされていなければアドレスエントリを削除するステップとをさらに含む、請求項 1 または 3 に記載の方法。

【請求項 9】

アプリケーション特有エイジング間隔の間のアドレスエントリのアクセスの検出に応答して、アプリケーション特有エイジングタイマをリセットするステップをさらに含む、請求項 8 に記載の方法。

【請求項 10】

ネットワークスイッチであって、

ネットワークスイッチポートを含み、各ネットワークスイッチポートは、受信レイヤ 2 データパケットから複数の予め規定されたネットワークアプリケーションのうちの検出された 1 つに対するアプリケーション状態を判断するよう構成されるパケットクラシファイアを含み、前記アプリケーション状態の判断は、前記パケットクラシファイアが、受信レイヤ 2 データパケットの少なくとも一部と、それぞれの利用可能な予め規定されたネットワークアプリケーションを識別するために用いられるフレームデータを指定する複数のテンプレートとを同時に比較することに基づいており、

さらに

スイッチング論理を含み、スイッチング論理は、判断されたアプリケーション状態と、検出された 1 つの予め規定されたネットワークアプリケーションに基づくアプリケーション特有エイジング間隔の間の、判断されたアドレスエントリのインアクティビティとに基づいて、受信レイヤ 2 データパケットのソースおよびレイヤ 2 データパケットのデスティネーションの少なくとも 1 つを指定するアドレスエントリを選択的に削除するよう構成される、ネットワークスイッチ。

【請求項 11】

スイッチング論理は、

受信レイヤ 2 データパケットからのアプリケーション状態の検出に応答して、アドレスエントリに対するアプリケーション特有エイジング間隔のカウントを開始するよう構成されるプログラム可能タイマを含む、請求項 10 に記載のスイッチ。

【請求項 12】

スイッチング論理は、

第 2 のアドレスエントリに対する第 2 のアプリケーション特有エイジング間隔をカウントするよう構成される、第 2 のプログラム可能タイマを含む、請求項 10 に記載のスイッチ。

【請求項 13】

受信レイヤ 2 データパケットに対するアドレスエントリおよび第 2 のアドレスエントリをストアするよう構成される、ネットワークスイッチアドレステーブルをさらに含む、請求項 12 に記載のスイッチ。

【発明の詳細な説明】

【0001】

【発明の背景】

発明の分野

この発明は、サブネットワーク間でデータパケットを交換するよう構成されたノンブロッキングネットワークスイッチにおけるデータパケットの交換に関する。

10

20

30

40

50

【 0 0 0 2 】

背景技術

米国特許第 5 , 9 1 4 , 9 5 6 号は、非同期転送モード (A T M) スイッチの接続能力を向上させるための、キャッシュ装置および方法を開示する。スイッチを通るデータセルのフローがモニタされ、未使用のまたは十分に用いられていないコネクションが検出されて、再生エイジャー (reclaim age) によってコネクションテーブルにマークされるが、これは定期的にコネクションのアクティビティを評価する。

I E E E コンピュータ協会の L A N M A N 標準委員会 (LAN MAN Standards Committee of the IEEE Computer Society) の、「システム間の情報技術電気通信および情報交換 - ローカルおよびメトロポリタンエリアネットワーク - 一般仕様 - 第 3 部 : メディアアクセス制御 (M A C) ブリッジ (Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) Bridges)」、I S O / I E C 1 5 8 0 2 - 3 : 1 9 9 8 A N S I / I E E E 標準 8 0 2 . 1 D、1 9 9 8 年版、4 2 - 2 6 頁は、I E E E 8 0 3 . 2 プロトコルにしたがった M A C ブリッジを開示する。この規格に従った古くなったアドレスの扱いを、以下に説明する。

ローカルエリアネットワークは、ネットワークケーブルまたは他の媒体を用いてネットワーク上のステーションをリンクする。各ローカルエリアネットワークアーキテクチャは、メディアアクセス制御 (M A C) を用いて各ネットワークノードでネットワークインターフェイス装置をイネーブルし、ネットワーク媒体にアクセスする。

【 0 0 0 3 】

イーサネット (R) プロトコル I E E E 8 0 2 . 3 は、データパケット送信のための半二重メディアアクセス機構および全二重メディアアクセス機構を指定するために開発された。全二重メディアアクセス機構は、2 つのネットワーク要素間の、たとえばネットワークノードと交換ハブ (switched hub) 間の、2 方向のポイント・ツー・ポイント通信リンクを提供する。

【 0 0 0 4 】

交換ローカルエリアネットワークに対して、さらなる高速接続性、より柔軟な交換性能、およびより複雑なネットワークアーキテクチャへの対処能力への要求が高まっている。たとえば、本発明の譲受人に譲渡された米国特許第 5 , 9 5 3 , 3 3 5 号は、レイヤ 2 タイプのイーサネット (R) (I E E E 8 0 2 . 3) データパケットを異なったネットワークノード間で交換するよう構成されたネットワークスイッチを開示する ; 受信されたデータパケットは、別のサブネットワーク (ルータを介する) または予め規定されたステーションのグループを指定する I E E E 8 0 2 . 1 q プロトコルにしたがった V L A N (バーチャル L A N) タグ付きフレームを含み得る。スイッチングはレイヤ 2 レベルで行なわれるので、典型的にはルータがサブネットワーク間でデータパケットを転送するために必要である。

【 0 0 0 5 】

既存のレイヤ 2 タイプネットワークスイッチの 1 つの問題は、ユーザによって設定される固定されたエイジング間隔を有するエイジング機能を用いて、古くなった (aged) アドレスエントリをネットワークスイッチアドレステーブルから削除することに関わる。特に、ネットワークスイッチは、受信したデータパケットが、たとえば未知の M A C ソースまたはデスティネーションアドレスのような未知のアドレスを有することを検出すると、新しいアドレスエントリをネットワークスイッチアドレステーブルに加えることにより、新しいネットワークアドレスを「学習」する。ネットワークスイッチはまた、古くなったアドレスエントリを削除してネットワークスイッチアドレステーブルがオーバーフローすることを防ぐためのエイジング機能をも含む。たとえば、ネットワークスイッチのスイッチファブリックは、受信したデータパケットを交換するためにアドレスエントリがアクセスされるごとに、アドレスエントリ内に「ヒットビット」を設定し得る。エイジング機能は、設定された「ヒットビット」の検出に応答して、ビットをゼロにリセットし、アドレスエ

10

20

30

40

50

ントリをネットワークスイッチアドレステーブル内に維持する。もしエイジング機能が、アドレスエントリが少なくとも固定アドレス間隔の間アクセスされていないことを示して「ヒットビット」がすでにゼロ値にリセットされていることを検出すると、エイジング機能は古くなったアドレスエントリをネットワークスイッチアドレステーブルから削除する。

【0006】

しかしながら上述のエイジング機能を用いると、ネットワークスイッチからのアドレスエントリの早過ぎる削除をまねくおそれがあり、ネットワークスイッチはネットワークアドレスを再学習することが必要になる。たとえば、2つのネットワークアプリケーションの間でより高いプロトコルの通信（すなわち、フロー）が生じていることを認識していないレイヤ2スイッチは、ユーザ定義エイジング間隔が短すぎる値に設定されていれば、2つのネットワークアプリケーションの間でのフローに関連するデータパケットに対するアドレスエントリを削除および再学習を繰り返し、ネットワークスイッチリソースを無駄にするおそれがある。しかしながら、ユーザ定義エイジング間隔を増大する試みは、ネットワークスイッチアドレステーブルをオーバーフローさせかねない。

【0007】

【発明の要約】

ネットワークスイッチが、そのネットワークスイッチアドレステーブルからアドレスエントリを選択的に削除する一方で、削除されたアドレスエントリの再学習を最小化することを可能にする構成に対する必要性がある。

【0008】

また、ネットワークスイッチが、2つのネットワークアプリケーションの間のネットワークスイッチによって転送されるデータフローの状態に基づいて、そのネットワークスイッチアドレステーブルからアドレスエントリを選択的に削除することを可能にする構成に対する必要性がある。

【0009】

これらおよび他の必要性はこの発明によって満たされるが、この発明においてはネットワークスイッチは、受信したデータパケットからのデータフローの判断されたアプリケーション状態に基づいて、ネットワークスイッチアドレステーブルからアドレスエントリを選択的に削除する。

【0010】

この発明の一局面は、統合ネットワークスイッチにおける方法を提供し、方法は、受信したレイヤ2データパケットから予め規定されたネットワークアプリケーションに対するアプリケーション状態を判断するステップと、判断したアプリケーション状態に基づいて、受信したレイヤ2データパケットのソースおよびレイヤ2データパケットのデスティネーションの少なくとも1つを指定するネットワークスイッチアドレステーブルからアドレスエントリを選択的に削除するステップとを含む。統合ネットワークスイッチによるアプリケーション状態の判断により、統合ネットワークスイッチが、予め規定されたネットワークアプリケーションにしたがってネットワークノード間のデータフローの存在を識別することが可能になり、統合ネットワークスイッチが予め規定されたネットワークアプリケーションパラメータにしたがってエイジングタイマを調節することが可能になる。さらに、判断されたアプリケーション状態に基づくアドレスエントリの選択的削除により、統合ネットワークスイッチが、アプリケーション状態からネットワークノード間のデータフローがたとえば2つのノード間のセッションの最後で終了したと判断した際に、アドレスエントリを削除することが可能になる。

【0011】

この発明の別の局面は、ネットワークスイッチポートとスイッチング論理とを含むネットワークスイッチを提供する。ネットワークスイッチポートは各々、受信したレイヤ2データパケットから複数の予め規定されたネットワークアプリケーションのうちの検出された1つに対するアプリケーション状態を判断するよう構成されるパケットクラシファイアを

10

20

30

40

50

含む。スイッチング論理は、判断されたアプリケーション状態および、アプリケーション特有エイジング間隔の間に判断されたアドレスエントリのインアクティビティに基づいて、受信レイヤ2データパケットのソースおよび受信レイヤ2データパケットのデスティネーションの少なくとも1つを指定するアドレスエントリを選択的に削除するよう構成され、ここでアプリケーション特有エイジング間隔は、検出された1つの予め規定されたネットワークアプリケーションに基づく。よって、スイッチング論理は、受信レイヤ2データパケットから検出された1つの予め規定されたネットワークアプリケーションに基づくインアクティビティの間隔に基づいて、アドレスエントリを選択的に削除し、統合ネットワークスイッチによってサポートされるネットワークアプリケーションに基づくより正確なエイジング間隔の使用を可能にする。これに代えて、スイッチング論理は、判断したアプリケーション状態に基づいてアドレスエントリを選択的に削除して、検出された1つの予め規定されたネットワークアプリケーションにしたがった2つのネットワークノードの間のセッションの完了の際に、スイッチング論理がアドレスエントリを削除することを可能にする。

10

【0012】

この発明のさらなる利点と新規な特徴とは、一部は以下の説明により明らかとなり、一部は当業者においては以下を検討すれば明らかであるか、またはこの発明を実施することにより学習されるであろう。この発明の利点は、前掲の特許請求の範囲に特に指摘される手段および組み合わせにより、実現され達成されるであろう。

【0013】

添付の図面を参照すると、同様の参照番号を有する要素は図面を通して同様の要素を示す。

20

【0014】

【発明を実行するためのベストモード】

開示される実施例は、ネットワークスイッチアドレステーブルエントリに対するアプリケーションに基づくエイジング間隔を生成するための統合ネットワークスイッチにおける構成、および受信レイヤ2データパケットから予め規定されたネットワークアプリケーションに対して判断されたアプリケーション状態に基づいてアドレスエントリを選択的に削除することに向けられる。受信レイヤ2データパケットから予め定められたネットワークアプリケーションを検出すること、および予め定められたネットワークアプリケーションに対するアプリケーション状態を検出することは、レイヤ2データパケットを受信したネットワークスイッチポート内のパケットクラシファイアモジュールによって行なわれる。第1にエトスアーキテクチャおよびパケットクラシファイアについての説明を提供し、次いでネットワークスイッチアドレスエントリのアプリケーションに基づくエイジングの説明を行なう。

30

【0015】

図1は、イーサネット(R)(IEEE802.3)ネットワークのようなパケット交換ネットワーク10を示すブロック図である。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14間でのデータパケット通信を可能にする統合(すなわち単一チップ)マルチポートスイッチを含む。各ネットワークステーション14、たとえば顧客ワークステーションは典型的には、IEEE802.3プロトコルにしたがってデータパケットを10Mbpsまたは100Mbpsで送受信するよう構成される。各統合マルチポートスイッチ12は、ギガビットイーサネット(R)リンク16で相互接続されており、サブネットワーク18a、18b、18cの間のデータパケット転送を可能にする。よって、各サブネットワークはスイッチ12と、関連のネットワークステーション14のグループとを含む。

40

【0016】

各スイッチ12は、メディアアクセス制御(MAC)モジュール22とパケットクラシファイアモジュール24とを含むスイッチポート20を含む。MACモジュール20は、IEEE802.3uプロトコルにしたがって、10/100Mbps物理層(PHY)ト

50

ランシーバ（図示せず）をわたって関連のネットワークステーション１４とデータパケットを送受信する。各スイッチ１２はまた、受信データパケットに対するフレーム転送判断を行なうよう構成されたスイッチファブリック２５をも含む。特に、スイッチファブリック２５は、イーサネット（Ｒ）（ＩＥＥＥ ８０２．３）ヘッダ内のソースアドレス、デスティネーションアドレス、およびＶＬＡＮ情報に基づいたレイヤ２スイッチング判断に対して構成されている；スイッチファブリック２５はまた、イーサネット（Ｒ）パケット内のＩＰデータパケットの評価に基づいた選択的なレイヤ３スイッチング判断に対しても構成されている。

【００１７】

図１に示されるように、各スイッチ１２は、関連のホストＣＰＵ２６とたとえばＳＳＲＡＭであるバッファメモリ２８とを有する。ホストＣＰＵ２６は、以下に説明するスイッチファブリック２５およびパケットクラシファイアのプログラミングを含む、対応のスイッチ１２の全体的な動作を制御する。バッファメモリ２８は対応のスイッチ１２によって用いられ、スイッチファブリック２５が受信したデータパケットに対する転送判断を処理している一方で、データフレームをストアする。

【００１８】

上述のように、スイッチファブリック２５はレイヤ２スイッチング判断とレイヤ３スイッチング判断を行なうよう構成されている。スイッチファブリック２５によるレイヤ３スイッチング判断を用いると、スイッチファブリック２５が先取り（advanced）転送判断を含めてどのようにパケットを扱うか、およびパケットをビデオまたは音声などのレイテンシ感応アプリケーションに対する高優先性パケットとみなすべきか、についてまで知的な判断が行なえるようになる。

【００１９】

開示される実施例にしたがうと、図１のパケットクラシファイアモジュール２４は、入ってくるデータストリームと、入ってくるデータストリームのデータ形式を識別するテンプレートとの間の多数同時比較に対して構成される。特定的には、ホストプロセッサ２６のユーザは、受信データパケットの選択部分におけるある予め規定されたデータ値を有するデータパケットがどのようにスイッチファブリック２５によって扱われるべきかを定義するポリシーを指定する。これらのポリシーは対応のポリシーごとに、スイッチファブリック２５に１組のフレーム転送判断またはエイジング機能パラメータをローディングすることにより実現される。受信データパケットの選択部分におけるある予め規定されたデータ値は、レイヤ２データパケットのどの部分に場所決めされていてもよいことに留意されたい。よって、パケットクラシファイアモジュール２４は、たとえばハイパーテキスト転送プロトコル、ＳＮＭＰ、ＦＴＰ、Telnetのような異なったネットワークアプリケーションにしたがったデータフローの存在を検出できる。

【００２０】

よって、スイッチファブリック２５はＨＴＴＰパケットについて１組のフレーム転送命令およびエイジングパラメータと、ＳＮＭＰパケットに対する別の組のフレーム転送命令およびエイジングパラメータと、高優先性パケット（たとえばビデオ、音声）に対する別の組のフレーム転送命令およびエイジングパラメータとを含み得る。

【００２１】

図２は、この発明の実施例にしたがったパケットクラシファイアモジュール２４を示すブロック図である。図２に示すように、ネットワークスイッチポート２０は、ＭＡＣ２２と、受信ＦＩＦＯバッファ２７と、ヘッダモディファイア２９と、パケットクラシファイアモジュール２４とを含む。ネットワークスイッチポートフィルタとも称するパケットクラシファイアモジュール２４は、ネットワークスイッチポート２０で入ってくるデータパケットを識別し（すなわち、評価し）、かつスイッチファブリック２５に受信されたデータパケットのタイプに基づくデータパケットに対して行なわれるべきアクションを指定するタグを与える。特定的には、パケットクラシファイアモジュール２４は、入ってくるデータパケットを、それぞれの受信データ形式を識別するよう構成された複数のテンプレート

10

20

30

40

50

と同時に比較する。入ってくるデータパケットと複数のテンプレートとの比較に基づいて、パケットクラシファイアモジュール 24 は、スイッチファブリック 25 に与えられるべきタグを指定する、実行されるべき方程式を識別する。

【0022】

特定的には、パケットクラシファイアモジュール 24 は複数のテンプレートから少なくとも 1 つの整合するテンプレートを検出することにより入ってくるデータパケットを識別する比較結果を生成する。パケットクラシファイアモジュール 24 は次いで、どの方程式が整合したテンプレートを含むかを識別し、その方程式によって指定されるタグを生成する。

【0023】

図 3 A および図 3 B は、パケットクラシファイアモジュール 24 による方程式の 2 つのテンプレートの同時処理を示す図である。図 3 A は、方程式 $E_{q1} = M_1 * M_2 * M_3 * M_4 * (M_5 + M_6 + M_7 + M_8)$ のパケットクラシファイアモジュール 24 による論理評価を示す。

【0024】

図 3 B は、どのように方程式 E_{q1} が実際に極小項 (min term) メモリ 70 にストアされるかを示す。方程式 E_{q1} は 4 つのテンプレート 62 a、62 b、62 c、および 62 d を含む。テンプレート 62 a は、極小項 M_1 、 M_2 、 M_3 、 M_4 および M_5 を含む。テンプレート 62 b は、極小項 M_1 、 M_2 、 M_3 、 M_4 および M_6 を含む。テンプレート 62 c は、極小項 M_1 、 M_2 、 M_3 、 M_4 および M_7 を含む。テンプレート 62 d は、極小項 M_1 、 M_2 、 M_3 、 M_4 および M_8 を含む。各テンプレート 62 は、IP データパケット 32 のヘッダに基づいて認識可能である特定の IP データ形式に対応する。たとえば、テンプレート 62 a および 62 c は、HTTP パケットを識別するよう構成され、テンプレート 62 b および 62 d は、SNMP パケットを識別するよう構成され得る。特定的には、HTTP パケットは、IPv4 形式であるか、IP の time to live フィールドは 1 よりも大きいのか、IP ヘッダのプロトコルフィールドは TCP であるか、ヘッダチェックサムは正しいか、ソース TCP ポートは 80 であるか、またはデスティネーション TCP ポートは 80 であるか、が識別される。SNMP パケットは、IPv4 形式であるか、IP の time to live フィールドは 1 よりも大きいのか、IP ヘッダのプロトコルフィールドは TCP であるか、ヘッダチェックサムは正しいか、ソース TCP ポートは 25 であるか、またはデスティネーション TCP ポートは 25 であるか、が識別される。

【0025】

よって、以下の極小項が上述の基準のすべてを表すために確立され得る：

M_1 = パケットは IPv4 形式である

M_2 = IP の time to live フィールドは 1 よりも大きい

M_3 = IP ヘッダのプロトコルフィールドは TCP である

M_4 = ヘッダチェックサムは正しい

M_5 = ソース TCP ポートは 80 である

M_6 = デスティネーション TCP ポートは 80 である

M_7 = ソース TCP ポートは 25 である

M_8 = デスティネーション TCP ポートは 25 である

よって、テンプレート 62 a と 62 c とは HTTP パケットを識別し、テンプレート 62 b と 62 d とは SNMP パケットを識別する。こうして、方程式 1 (E_{q1}) は、テンプレート 62 a、62 b、62 c、62 d のいずれかが真であれば、特定の結果 (たとえば指定された値を有するタグ) がスイッチファブリック 25 に出力されるべきことを指定する。

【0026】

さらに、極小項 $M_1 \dots M_8$ は、入来データストリーム内のデータバイトの相対的な位置に対応する予め規定された順序で関連のテンプレート 62 a および / または 62 b 内に配置される。極小項 M_1 は、IP パケットの第 1 のバイト (B_1) との比較のために構成され

10

20

30

40

50

、極小項 M 2 は B 1 の後に続く I P パケットの後のバイト (B 2) との比較のために構成され、極小項 M 3 は B 2 の後に続く I P パケットの後のバイト (B 3) との比較のために構成され、以下同様である。よって、入来データストリーム内のデータバイトの相対的な位置に基づく順序での極小項を有するテンプレート 6 2 を用いると、入来データストリームと極小項との間の多数同時比較が可能になる。よって、入来データパケットは多数のテンプレートと比較されて、入来データパケットの形式だけでなく、スイッチファブリック 2 5 によってどのアクションが行なわれなければならないかをも判断する。

【 0 0 2 7 】

図 2 に示されるように、ネットワークスイッチポートフィルタとも称するパケットクラシファイア 2 4 は、極小項値 (たとえば M 1、M 2 など) をストアするための極小項メモリ 7 0 と、受信されるレイヤ 2 フレームのタイプを識別するよう構成されるフレームアイデンティファイア 7 2 とを含む。特に、受信されているレイヤ 2 フレームのタイプを識別することは (たとえばイーサネット (R) I E E E 8 0 2 から 3 など)、レイヤ 2 パケット 3 0 内の I P パケット 3 2 の開始位置 6 4 の識別を可能にする。パケットクラシファイア 2 4 はまた、極小項コントローラ 7 4、極小項ジェネレータ 7 6、方程式コア 7 8、および評価結果メモリ 8 0 を含む。プロセッサインターフェイスモジュール (pi_mod) 8 2 は、ホスト C P U 2 6 から極小項メモリ 7 0 に、生成された極小項を転送するために用いられる。

【 0 0 2 8 】

極小項コントローラ 7 4 は、受信 I P フレームの選択されたバイトに対応する極小項メモリ 7 0 から極小項をフェッチするよう構成される。極小項コントローラ 7 4 はまた、レイヤ 2 フレームのタイプを指定するフレームアイデンティファイア 7 2 からのフレームタイプ (frm_type) 信号の受信に応答して開始ポイントの実際のバイト位置 (byte_location) を指定するよう構成される位置コンバータをも含む。極小項コントローラ 7 4 は次いで、極小項値 (M_STRU INFO) を極小項ジェネレータ 7 6 と方程式コア 7 8 とに転送する。

【 0 0 2 9 】

極小項ジェネレータ 7 6 は、極小項コントローラによってフェッチされた極小項と入来データストリームの選択されたバイトとの間の、実際極小項比較を行ない、実際極小項比較の結果 (mt_result) を方程式コア 7 8 に送る。開示される実施例に従うと、極小項ジェネレータは入来データストリームの最大 8 極小項までの同時比較に対して構成される。方程式コア 7 8 は、関連のテンプレート 6 2 に対する、極小項ジェネレータ 7 6 から受信した極小項比較結果に基づいて、フレームタグを生成するよう構成される。

【 0 0 3 0 】

上述のように、パケットクラシファイアモジュール 2 4 は、受信レイヤ 2 データパケットがたとえば H T T P、S N M P、F T P、Telnet のような予め規定されたネットワークアプリケーションに対するフレームデータを担持するかどうかを、各ネットワークスイッチポート 2 0 が識別することを可能にする。さらに、パケットクラシファイアモジュール 2 4 は、受信レイヤ 2 データパケットから、アプリケーション状態を特定の識別するよう、付加的なテンプレートでプログラム可能である。特定のには、ネットワークノードは予め規定されたネットワークアプリケーションにしたがって通信し、2 つのネットワークノード間の予め規定されたデータフローをもたらす。よって、ネットワークノード間で転送されるレイヤ 2 データパケットは、たとえばセッション開始要求、肯定応答、セッションの間の通信、セッション終了要求、およびセッション終了の肯定応答のような、予め規定されたネットワークアプリケーション状態を指定するペイロード情報を含み得る。パケットクラシファイアモジュール 2 4 は、レイヤ 2 データパケットのペイロードデータを評価することにより、データフローの状態をモニタするための適切なテンプレートでプログラムされ得る。よって、パケットクラシファイアモジュール 2 4 は、2 つのネットワークノード間で動作するネットワークアプリケーションに対するアプリケーション状態をモニタでき、スイッチファブリック 2 5 のスイッチング論理がアプリケーションに基づくエイジング動作を行なうことを可能にする。

10

20

30

40

50

【 0 0 3 1 】

図 4 は、この発明の実施例にしたがった図 1 のスイッチファブリック 2 5 を詳細に示す図である。図 4 に示されるように、スイッチファブリック 2 5 はスイッチング論理 8 0 と、ネットワークスイッチアドレステーブル 8 2 とを含む。ネットワークスイッチアドレステーブル 8 2 は、外部アドレステーブルとしてネットワークスイッチ 1 2 の外部に実現されてもよいことに留意されたい。スイッチング論理 8 0 は、ネットワークスイッチアドレステーブル 8 2 内のアドレステーブルエントリ 8 4 の学習（すなわちストア）を制御する。

【 0 0 3 2 】

各アドレステーブルエントリ 8 4 は、MAC アドレスフィールド 8 4 c、IP アドレスフィールド 8 4 d、および、対応の MAC および / または IP アドレスを有するネットワークノードに対する対応のレイヤ 2 および / またはレイヤ 3 スwitchング情報を提供するスイッチング情報フィールド 8 4 e を含む。各アドレステーブルエントリ 8 4 はまた、エイジングタイマのための開始時間をストアするよう構成されるエイジングタイマ開始フィールド 8 4 a と、判断されたデータフローによってサポートされるネットワークアプリケーションに基づいて、アプリケーション特有エイジング時間間隔をストアするよう構成されるエイジング間隔フィールド 8 4 b とを有する。

【 0 0 3 3 】

図 5 は、この発明の実施例にしたがった、ネットワークスイッチアドレステーブル 8 2 のアプリケーションに基づくエイジングを行なう方法を示すフロー図である。フロー図として示されるが、実際の実現化は状態に基づくものであって、スイッチング論理 8 0 内の新しい状態がパケットクラシファイアモジュール 2 4 からの予め規定されたアクションタグに応答して実行される。

【 0 0 3 4 】

方法はステップ 9 0 において開始し、ここでホスト CPU は、受信レイヤ 2 データパケットに基づく各ネットワークアプリケーションを識別するよう構成される、アプリケーションテンプレートの少なくとも最初の組を、パケットクラシファイアモジュール 2 4 の極小項メモリ 7 0 にローディングする。特に、ネットワークノードは予め規定されたネットワークアプリケーションにしたがってセッション開始要求を最初に送信する。ネットワークノードからの最初の要求の検出に回答して、パケットクラシファイアモジュール 2 4 は、新しいネットワークノードおよび対応の識別されたネットワークアプリケーションを指定するタグをスイッチファブリック 2 5 に送る。スイッチング論理 8 0 は、次いでステップ 9 2 においてネットワークスイッチアドレステーブル 8 2 に別のエントリ 8 4 を加えることにより、エントリを学習する。ホスト CPU もまた新しいデータフローを通知され、ステップ 9 4 においてホスト CPU はその時点で検出されたアプリケーションに対する付加的なテンプレートを極小項メモリ 7 0 にストアし、パケットクラシファイアモジュール 2 4 が検出されたデータフローに対する各アプリケーション状態を識別することを可能にする。これに代えて、一度にすべてのテンプレートが極小項メモリ 7 0 にローディングされてもよい。

【 0 0 3 5 】

ネットワークアプリケーションが対応の受信レイヤ 2 データパケットによってサポートされていることを検出すると、ホスト CPU 2 6 は応答してアプリケーション特有エイジング間隔（たとえば T 1 ）をエイジング間隔フィールド 8 4 b にローディングし、タイマに対する開始時間をエイジングタイマ開始フィールド 8 4 a にステップ 9 6 において記録する。

【 0 0 3 6 】

レイヤ 2 データパケットを受信したネットワークスイッチポート 2 0 のパケットクラシファイアモジュール 2 4 は、ネットワークアプリケーションの識別されたデータフローに対するペイロードデータを有するさらなるレイヤ 2 データパケットについてモニタを続ける。ステップ 9 8 においてパケットクラシファイアモジュール 2 4 が既存のデータフローに対して新しい状態を検出し、かつステップ 1 0 0 において新しい状態が識別されたネット

10

20

30

40

50

ワークアプリケーションにしたがったデータフローに対するセッションの終了を示すと、パケットクラシファイアモジュール 24 はセッションの終了を指定するアクションタグをスイッチング論理 80 に出力し、スイッチング論理 80 にステップ 102 において対応のアドレステーブルエントリ 84 を削除させる。しかしながら、ステップ 100 において新しい状態がパケットクラシファイア 24 によって中間アプリケーション状態であると判断されると、パケットクラシファイア 24 はスイッチング論理 80 に適切なタグを送り、フィールド 84 a における開始時間を更新された値でオーバーライドすることにより、ステップ 104 でスイッチング論理 80 にタイマをリセットさせる。

【0037】

スイッチング論理 80 は、アプリケーション特有エイジング間隔を、その内部クロックを開始時間エントリ 84 b に対して開始時間エントリ 84 a と比較することにより、連続的にモニタする。スイッチング論理 80 が、エイジング間隔フィールド 84 b に指定されるアプリケーション特有エイジング間隔の間、対応のアドレステーブルエントリがアクセスされていないことを示して、アプリケーション特有タイマがステップ 106 において期限が切れていると判断するれば、スイッチング論理 80 はステップ 102 においてエントリ 80 を削除する。

【0038】

開示される実施例にしたがうと、アプリケーション特有エイジング間隔が、判断されたアプリケーション状態と組合せて用いられて、アドレスエントリがいつネットワークスイッチアドレステーブルから削除されるべきかを正確に判断する。よって、アドレステーブルは、ネットワークスイッチによって生じるデータフローに基づくネットワークスイッチによって正確に維持され得る。

【0039】

この発明を最も実例的な実施例であるところで考えられるものによって説明してきたが、この発明は開示される実施例に限定されるものではなく、その反対に、さまざまな変形と等価な構成とを、前掲の特許請求の範囲に含むことが意図されることを理解されたい。

【図面の簡単な説明】

【図 1】 この発明の実施例にしたがったそれぞれのサブネットワークの間のデータパケットを交換するための、多数のネットワークスイッチを含む、パケット交換ネットワークのブロック図である。

【図 2】 この発明の実施例にしたがったパケットクラシファイアモジュールを含む、図 1 のネットワークスイッチポートを示すブロック図である。

【図 3 A】 図 2 の極小項ジェネレータによる方程式の 4 つのテンプレートの同時処理を示す図である。

【図 3 B】 図 2 の極小項ジェネレータによる方程式の 4 つのテンプレートの同時処理を示す図である。

【図 4】 図 1 のスイッチング論理を詳細に示す図である。

【図 5】 この発明の実施例にしたがった、判断されたアプリケーション状態およびアプリケーション特有エイジング間隔に基づいてアドレスエントリを選択的に削除する方法を示す図である。

10

20

30

40

【図 1】

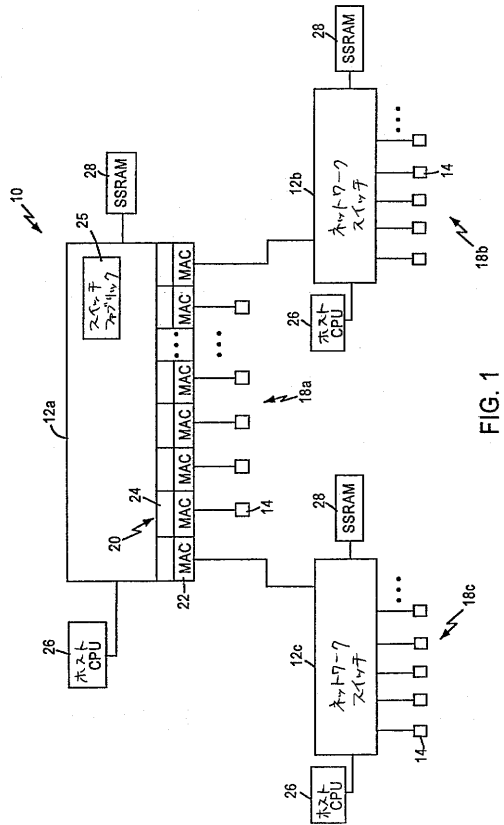


FIG. 1

【図 2】

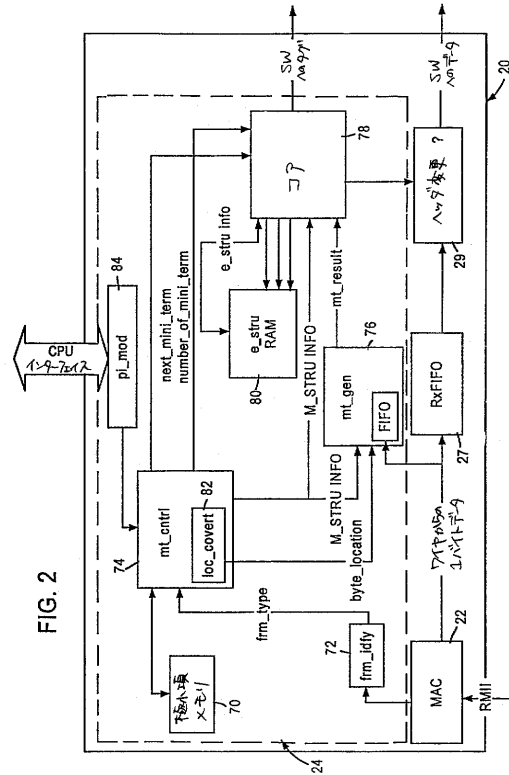
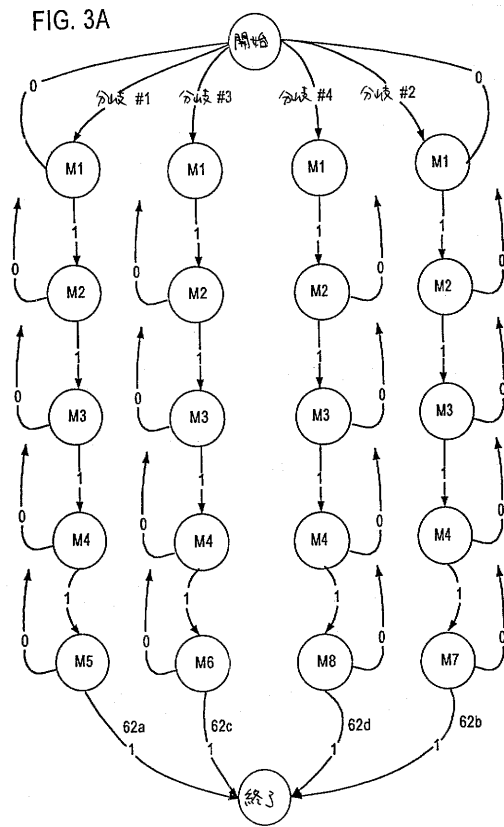


FIG. 2

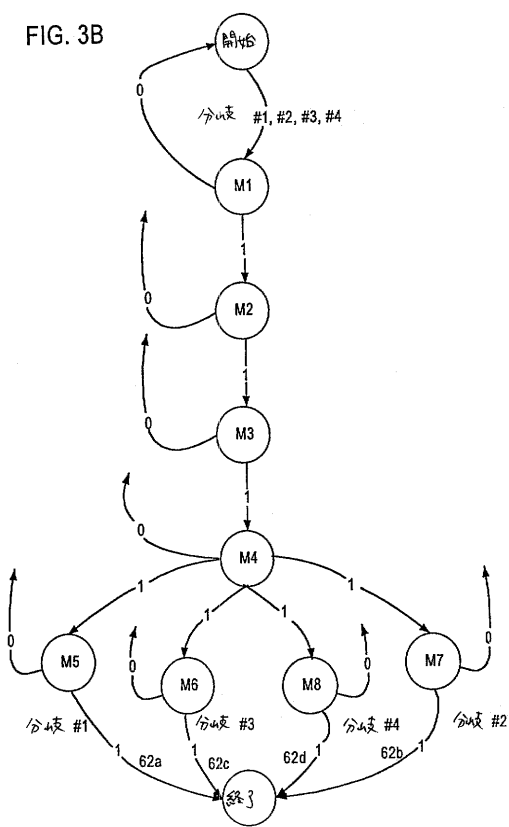
【図 3 A】

FIG. 3A



【図 3 B】

FIG. 3B



【図4】

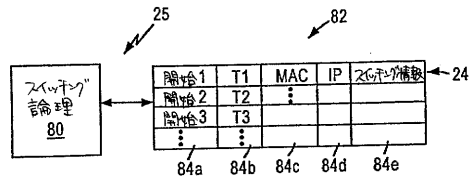


FIG. 4

【図5】

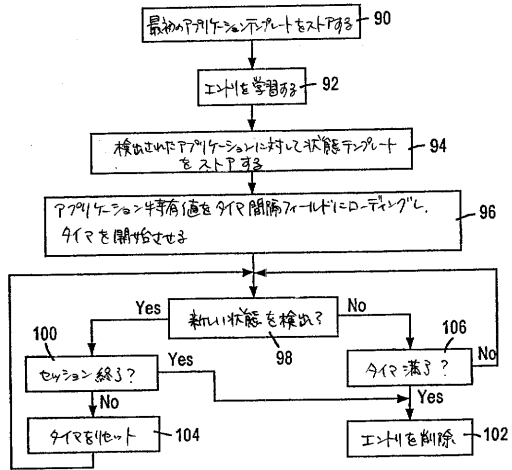


FIG. 5

フロントページの続き

(74)代理人 100083703

弁理士 仲村 義平

(74)代理人 100096781

弁理士 堀井 豊

(74)代理人 100098316

弁理士 野田 久登

(74)代理人 100109162

弁理士 酒井 將行

(72)発明者 カヌリ, ムルドゥラ

アメリカ合衆国、 9 5 0 5 0 カリフォルニア州、 サンタ・クララ、 メンゼル・プレイス、 1 9 8
3

(72)発明者 クリシュナ, ゴバル

アメリカ合衆国、 9 5 1 4 8 カリフォルニア州、 サン・ノゼ、 アダムスウッド・ドライブ、 3 2
6 0

審査官 岩田 玲彦

(56)参考文献 特開 2 0 0 0 - 0 0 4 2 5 1 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

H04L 12/40