

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2017-519449

(P2017-519449A)

(43) 公表日 平成29年7月13日 (2017.7.13)

(51) Int.Cl.
H04L 12/933 (2013.01)F I
H04L 12/933テーマコード (参考)
5K030

審査請求 有 予備審査請求 有 (全 58 頁)

(21) 出願番号 特願2016-574895 (P2016-574895)
 (86) (22) 出願日 平成26年12月3日 (2014.12.3)
 (85) 翻訳文提出日 平成28年12月20日 (2016.12.20)
 (86) 国際出願番号 PCT/AU2014/050391
 (87) 国際公開番号 W02016/023062
 (87) 国際公開日 平成28年2月18日 (2016.2.18)
 (31) 優先権主張番号 2014903158
 (32) 優先日 平成26年8月13日 (2014.8.13)
 (33) 優先権主張国 オーストラリア (AU)
 (31) 優先権主張番号 2014904247
 (32) 優先日 平成26年10月23日 (2014.10.23)
 (33) 優先権主張国 オーストラリア (AU)

(71) 出願人 516382962
 メタマコ テクノロジー エルピー
 METAMAKO TECHNOLOGY
 LP
 オーストラリア国 2000 ニューサウ
 スウェールズ州 シドニー ブリッジ ス
 トリート 50 レベル 2
 (74) 代理人 100105957
 弁理士 恩田 誠
 (74) 代理人 100068755
 弁理士 恩田 博宣
 (74) 代理人 100142907
 弁理士 本田 淳

最終頁に続く

(54) 【発明の名称】 低レイテンシスイッチング装置および方法

(57) 【要約】

データスイッチング方法。クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信する。クロスポイントスイッチはデータをクロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストする。クロスポイントスイッチの前記出力ポートまたは各出力ポートは、FPGAのような論理機能デバイスの対応する入力部に接続されている。論理機能デバイスは、クロスポイントスイッチの出力ポートまたは各出力ポートから受信したデータに、アドレスフィルタリングまたは多重化のような論理機能を適用し、処理後のデータを論理機能デバイスの1つまたは複数の出力インタフェースに出力する。また、或るスイッチング方法は、受信したデータを出力部へと回路交換するとともにそのデータを上位層機能にコピーする。

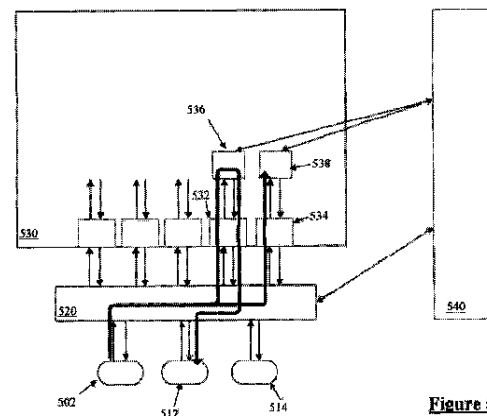


Figure 5

【特許請求の範囲】

【請求項 1】

データスイッチング方法であって、

クロスポイントスイッチの少なくとも 1 つの入力ポートでデータを受信し、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも 1 つの出力ポートに設定可能にキャストし、

前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを、論理機能デバイスの対応する入力部に接続し、

前記論理機能デバイスが前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの 1 つまたは複数の対応する出力インタフェースに出力すること
を備える、データスイッチング方法。

10

【請求項 2】

前記クロスポイントスイッチは当該クロスポイントスイッチの複数の出力ポートに前記データをマルチキャストする、請求項 1 に記載のデータスイッチング方法。

【請求項 3】

前記クロスポイントスイッチは当該クロスポイントスイッチの 1 つの出力ポートに前記データをユニキャストする、請求項 1 に記載のデータスイッチング方法。

【請求項 4】

前記論理機能デバイスはフィールドプログラマブルゲートアレイ (FPGA) を備える、請求項 1 ～ 3 のいずれか 1 項に記載のデータスイッチング方法。

20

【請求項 5】

前記論理機能デバイスの前記出力インタフェースまたは各出力インタフェースは前記クロスポイントスイッチの入力部に接続されている、請求項 1 ～ 4 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 6】

前記論理機能デバイスによって適用された前記論理機能は、内容ベースフィルタリングを備える、請求項 1 ～ 5 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 7】

前記内容ベースフィルタリングはアドレスフィルタリングであり、そのアドレスフィルタリングによって、前記クロスポイントスイッチは当該クロスポイントスイッチの複数の出力ポートに前記データをマルチキャストし、複数のクライアントデバイスは前記論理機能デバイスの複数の出力ポートにそれぞれ関連付けされており、このことによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけが前記論理機能デバイスの対応する出力ポートから出力されるように前記論理機能デバイスは前記クロスポイントスイッチの各出力ポートからのデータをフィルタする、請求項 6 に記載のデータスイッチング方法。

30

【請求項 8】

前記論理機能は多重化である、請求項 1 ～ 7 のいずれか 1 項に記載のデータスイッチング方法。

40

【請求項 9】

前記論理機能は、前記論理機能デバイスのトランシーバ毎に専用のルーティングテーブルを提供することによって有効化されるレイヤ 3 ルーティングである、請求項 1 ～ 8 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 10】

管理デバイスを使用して、高速遠隔再構成を有効にするように前記論理機能デバイスの複数の論理ブロックを再プログラムすることをさらに備える、請求項 1 ～ 9 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 11】

前記クロスポイントが重複データストリームをデータパケット捕捉 RAM、ディスクおよ

50

び／または P C I e リンクに配信することによって、複数のデータパケットを捕捉することをさらに備える、請求項 1 ～ 1 0 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 1 2】

データスイッチであって、

データを受信するための少なくとも 1 つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを前記出力ポートの少なくとも 1 つに設定可能にキャストするように構成されているクロスポイントスイッチと、

前記クロスポイントスイッチの前記出力ポートまたは各出力ポートにそれぞれ接続された、1 つまたは複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも 1 つの論理機能を有する、前記論理機能デバイスとを備えるデータスイッチ。

10

【請求項 1 3】

データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は 1 つまたは複数のプロセッサによって実行される場合に、

クロスポイントスイッチの少なくとも 1 つの入力ポートでデータを受信すること、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも 1 つの出力ポートに設定可能にキャストすること、

前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを論理機能デバイスの対応する入力部に接続すること、

前記論理機能デバイスが、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの 1 つまたは複数の対応する出力インタフェースに出力することを実行させる、非一時的コンピュータ読み取り可能媒体。

20

【請求項 1 4】

データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、

どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの 1 つを選択するためのコンピュータプログラムコード手段と、

クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを、当該クロスポイントスイッチの少なくとも 1 つの出力ポートに設定可能にキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートが論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、

前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの 1 つまたは複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段とを備えるコンピュータプログラム製品。

30

40

【請求項 1 5】

物理層データスイッチング方法であって、

スイッチの少なくとも 1 つの入力ポートでデータを受信し、

前記データを前記スイッチの少なくとも 1 つの出力ポートへと回路交換し、

上位層プロセスに前記データをコピーすること

を備える、物理層データスイッチング方法。

【請求項 1 6】

物理層データスイッチであって、

50

複数のデータ出力部のうちの1つまたは複数へと設定可能に回路交換される複数のデータ入力部であって、第1のデータ出力部が第1の入力部から回路交換されたデータを受信し、第2のデータ出力部が前記第1の入力部から前記データのコピーを受信する、前記複数のデータ入力部と、

前記第2のデータ出力部に接続された上位層機能回路とを備える物理層データスイッチ。

【請求項17】

前記上位層機能回路はパケットタイムスタンプングを実行するように構成されている、請求項16に記載の物理層データスイッチ。

【請求項18】

前記上位層機能回路はパケット捕捉を実行するように構成されている、請求項16または17に記載の物理層データスイッチ。

【請求項19】

前記上位層機能回路は前記第1のデータ出力部においてレイテンシに影響を及ぼさずに、データパケット統計のカウントを実行するように構成されている、請求項16または17に記載の物理層データスイッチ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワークデータスイッチングに関し、特に、FPGAまたはASICなどの1つまたは複数の論理機能デバイスを有する1つまたは複数のクロスポイントスイッチの組み合わせを用いて所望のネットワーク機能を実装する効率的な低レイテンシデータスイッチに関する。

【背景技術】

【0002】

1つのソースからデータパケットを得て多数の送信可能な送信先のうちの1つまたは複数にデータパケットを送信することが望ましい場合が多い。例えば、証券取引所からダウンストリーム方向への場合がこれに該当し、株式トレーダのコンピュータなどの多数の機械が取引所からパケットを受信できるようにしておくことが望ましい。

【0003】

アプリケーションによっては、1つのポートから多数のポートにブロードキャストまたはマルチキャストすることが適切である。マルチキャストは、デバイス（スイッチ）の入力部での個々のパケットを当該スイッチの1つ1つの出力部にブロードキャストするが、これは、例えば、クロスポイントスイッチを用いることによって達成可能である。しかしながら、パケットスイッチの場合のような他の状況では、2つのエンドポイント間のユニキャスト通信はプライベートでなければならない。

【0004】

アドレス指定可能なスイッチングは、何らかの仕組み（例えば従来のパケットスイッチにおける連想メモリ（content addressable memory）等）を用いてダウンストリームパケットを処理することにより、集積回路に実装可能である。従来、これには、単一のトランシーバでデータを受信し、パケットアドレスを確認し、ルックアップテーブルを参照してそのアドレスに関連付けされたポートを識別し、次に、そのポート宛てだけにパケットを送ることを必要とする。このアプローチは、使用する帯域幅を最適化する考え方に基づくが、比較的複雑な回路がスイッチ内に必要である。

【発明の概要】

【発明が解決しようとする課題】

【0005】

現行のネットワークスイッチは、レイヤ1（物理層）のデータが、レイヤ1に返送される前に、アドレスを確認するためにレイヤ2（データリンク層）に変換されることが必要

10

20

30

40

50

である。レイヤ1からレイヤ2へのデータの変換には、データを並列ストリームに非直列化することが必要である。なぜなら、関連する周波数では直列ストリームでレイヤ2の特徴を実装することが難しいからである。これは、あらゆるレイヤ2処理は、レイヤ1よりもはるかに速度が遅く、最も高速のレイヤ2ネットワークスイッチでさえ、レイヤ2のアドレス確認におよそ200~400nsの、かなりの遅延が発生することを意味する。複数のトレーダの通信データ量を単一のストリームへとアップストリームに集約するために用いる際には、メタマコLP社製のメタコネクトC16スイッチに加えて、「多対1」スイッチ、例えばイーサネット（登録商標）スイッチを使用することが必要で、レイテンシが大きくなり望ましくない。さらに、ネットワークスイッチは、単一のデバイスのアップストリームデータ経路にもダウンストリームデータ経路にも対処しなければならない場合が多く、信号経路が複雑になり、クロックの制限をもたらす。既存のスイッチの多くは、帯域幅を最適化し、多くのユースケース間で汎用的に用いるように設計されている。

10

【0006】

本明細書に含まれている文献、行為、材料、デバイス、物品またはその他のいずれかの説明は、単に本発明に技術背景を提供することを目的とする。これらの事項のいずれか、もしくはすべてが本出願の各請求項の優先日より前に存在したことにより、従来技術の基礎の一部を形成していると、または、本発明に関連する分野において一般によく知られた知見であると認めるものではない。

【0007】

本明細書全体を通して、「comprise（備える、含む）」という語、または「comprisesまたはcomprising」などの変化形は、記載された要素、整数もしくはステップ、または要素、整数もしくはステップの群が含まれることを含意すると理解されるが、任意の他の要素、整数もしくはステップ、または要素、整数もしくはステップの群を排除するものではない。

20

【0008】

本明細書では、ある要素が、列記された選択肢「の少なくとも1つ」であってもよいという記述は、その要素が、列記された選択肢のうちのいずれか1つであってもよいか、または、列記された選択肢のうちの2つ以上の任意の組み合わせであってもよいことと理解されたい。

【課題を解決するための手段】

30

【0009】

第1の態様によれば、本発明は、データスイッチング方法であって、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信し、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストし、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを、論理機能デバイスの対応する入力部に接続し、前記論理機能デバイスが前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力することを備える方法を提供する。

【0010】

40

第2の態様によれば、本発明は、データを受信するための少なくとも1つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを前記出力ポートの少なくとも1つに設定可能にキャストするように構成されているクロスポイントスイッチと、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートにそれぞれ接続された、1つまたは複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも1つの論理機能を有する、前記論理機能デバイスとを備えるデータスイッチを提供する。

【0011】

50

第3の態様によれば、本発明は、データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は1つまたは複数のプロセッサによって実行される場合に、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信すること、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストすること、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを論理機能デバイスの対応する入力部に接続すること、前記論理機能デバイスが、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力することを実行させる、非一時的コンピュータ読み取り可能媒体を提供する。

10

【0012】

第4の態様によれば、本発明は、データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの1つを選択するためのコンピュータプログラムコード手段と、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを、当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートが論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段とを備えるコンピュータプログラム製品を提供する。

20

【0013】

本発明のいくつかの実施形態では、クロスポイントスイッチの少なくとも1つの出力ポートにデータを設定可能にキャストするクロスポイントスイッチは、クロスポイントスイッチの複数の出力ポートにデータをマルチキャストするクロスポイントスイッチを備えることができる。いくつかの実施形態では、クロスポイントスイッチの少なくとも1つの出力ポートにデータを設定可能にキャストするクロスポイントスイッチは、クロスポイントスイッチの1つの出力ポートにデータをユニキャストするクロスポイントスイッチを備えることができる。

30

【0014】

論理機能デバイスは、フィールドプログラマブルゲートアレイ（FPGA）、適切に構成されたx86CPU系のプロセッサ、もしくはネットワークインタフェースが組み込まれたマルチコアマイクロプロセッサのような適切に構成された特定用途向け集積回路（ASIC）、または複数のこのようなデバイスを備えることができる。本明細書でのFPGAを利用する実施形態についての説明は、本発明の範囲内でASICまたは適切に構成されたプロセッサにもまた適用可能なものであるとして理解されたい。

40

【0015】

本発明のいくつかの実施形態は、帯域幅使用の最適化を断念することによって、低レイテンシの向上、および／または決定性向上（MACコリジョンの減少）のための最適化をこのように効果的に得ることができる。クロスポイントスイッチのブロードキャスト機能を使用して2つ以上の受信機にデータを分配することによって、可能総帯域幅が減少し、個々のハードウェア要素のデータスループットは必ずしも最大化されないが、本発明は、こうした複数の同一のデータストリームの並列処理が、他の有益な効果を有することを認識している。したがって本発明は、既存のスイッチの多くは、帯域幅を最適化し、多くのユースケース間で汎用的に使用するよう設計されているが、このことがレイテンシおよ

50

び決定性を犠牲にしていると認識している。

【0016】

データは、1つまたは複数のサーバを起点とするダウンストリームのデータとすることができ、1つまたは複数のクライアントデバイスを対象とすることができる。これに加え、またはこの代わりに、データは、1つまたは複数のクライアントデバイスを起点とするアップストリームのデータとすることができ、1つまたは複数のサーバを対象とすることができる。これに加え、またはこの代わりに、データは、高性能コンピューティングクラスタのノード間で、またはデータセンタネットワークのノード間で切り換えられるデータのような、ピアツーピアのデータとすることができる。

【0017】

クロスポイントスイッチの出力ポートの、FPGAの入力ポートへの接続は、いくつかの実施形態では、直接的な接続を備えることができる。あるいは他の実施形態では、AC結合用キャパシタのような回路素子、またはレベルトランスレーション、バッファリング、増幅、もしくは信号回復を実行するICを間に介在させることができる。

【0018】

FPGA出力インタフェース、または各FPGA出力インタフェースは、出力ポート、周辺部品相互接続エクスプレス(PCIe)出力部、他の出力インタフェース、または高速通信用インタフェースのストリームの群、のうちの1つまたは複数を用意することができる。

【0019】

いくつかの実施形態では、データは、サーバのような遠隔した場所から受信したダウンストリームデータであって、クロスポイントスイッチにとってローカルな1つまたは複数のクライアントデバイスに分配するように意図されたダウンストリームデータとすることができる。デバイスは、たとえスイッチと物理的に同一の場所になくても、スイッチに接続され、スイッチを介してより広いネットワークとインタフェースしているのであれば、スイッチにとってローカルであると見なされる。このような実施形態では、FPGAによって適用される論理機能は、アドレスフィルタリングのような、内容ベースフィルタリングを備えることができ、これにより、クロスポイントスイッチが、クロスポイントスイッチの複数の出力ポートにデータをマルチキャストするとともに、複数のクライアントデバイスが、複数のFPGAの出力ポートにそれぞれ関連付けされていることによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけがそれぞれのFPGAの出力ポートから出力されるように、FPGAが、クロスポイントスイッチの各出力ポートからのデータをフィルタする。したがって、このような実施形態は、効率的にブロードキャストするためにクロスポイントスイッチを利用し、効率的なアドレス指定、およびその他の機能のためにFPGAを利用する。内容ベースフィルタリングであれば、L2、L3またはL4から宛先またはソースアドレスを参照することによるアドレス指定を伴う場合もあるが、プロトコルに基づくフィルタリング、ファイアウォールの場合の、状態に基づくフィルタリングおよびディープパケット検査、侵入検出もしくは流出検出システム、フィルタリング対象のパケットの無効化、またはFPGAに実装される任意の他の論理機能によるフィルタリングを含むこともまた可能であろう。クロスポイントスイッチを用いて、FPGAの2つ以上の機能ブロックにデータストリームをコピーすることにより、複数の機能を同じデータに関してFPGA内で同時に実行可能になる。このような実施形態では、FPGAの単一のランシーバが、各フィルタについて受信および送信の両方の役割を果たすことができる。あるいは、FPGA出力ポートに接続されているクライアントデバイスのような別のデバイスが代わりに、このようなアドレスフィルタリング用に構成されている実施形態では、FPGAは、アドレスフィルタリングせずに出力ポートにデータを渡すことができる。したがって、このような実施形態は、各データパケットをアドレス指定するために、帯域幅が最適化されたレイヤ2ルックアップアプローチと対照をなす、「ブロードキャストアンドフィルタ」アプローチを用いる。

【0020】

他の実施形態では、データは、クロスポイントスイッチにとってローカルな1つまたは複数のクライアントデバイスから受信したアップストリームデータであって、1つまたは複数のサーバのような、1つまたは複数の遠隔した場所への配信を意図されたアップストリームデータとすることができる。このような実施形態では、FPGAによって提供される論理機能は、多重化していてもよい。多重化は、それぞれのデータクロックを同期するために、それぞれのFIFOバッファ、またはメモリキューにFPGAの各受信機入力部を接続し、2つ以上のクライアントデバイスから単一のアップストリームデータストリームへとデータを多重化するマルチプレクサに、各FIFOまたはメモリキューの出力を渡すことにより、FPGAによって提供することができる。次に、アップストリームデータストリームは、FPGAのポートで送信機によって出力され、外部ネットワークの前面パ

10
ネルポートに配信するためのクロスポイントスイッチに、すなわちもっと一般的には、クライアントデバイスに渡されてもよい。このクライアントデバイスに含まれる得るものとしては、例えば、スイッチの内部の、ネットワークインタフェースカードがある。いくつかの実施形態では、FPGAロジックは、ファイアウォール、IPアドレス指定のようなレイヤ3もしくは4の機能、または、場合によっては並行して実装される、パケット変換を実施するように構成することができる。

【0021】

いくつかの実施形態は、ローカルデバイスからFPGAの複数のマルチプレクサにアップストリームデータをマルチキャストするクロスポイントスイッチと、アドレスフィルタを有する各マルチプレクサの出力をフィルタリングすることにより、データがそれぞれの外部ネットワークポートに対してアドレス指定された場合に、それぞれのマルチプレクサからそのポートにだけデータが中継されるようにすることによって、ローカルデバイスから複数の遠隔した場所へのデータ配信をさらに提供することができる。

20

【0022】

いくつかの実施形態では、クロスポイントスイッチは、アップストリームデータおよびダウンストリームデータの両方を取り扱うことができ、これにより、データのストリームの一方が、クロスポイントスイッチによってFPGAの方に向けられているのに対して、データのもう一方でストリームが、クロスポイントスイッチの前面ポート間で1つまたは複数のクライアントデバイスに直接接続されている。

【0023】

FPGAは、トランシーバ毎に専用のルーティングテーブルを提供することにより、レイヤ3（ネットワーク層）ルーティングを実装するように構成することができる。ハードウェアは非能率的であるが、このような実施形態には、すべてのルーティング用に使用される単一のルーティングテーブルを使用しているレイヤ3ルータと比較して、かなり大きなレイテンシの利点がある。

30

【0024】

本発明の実施形態は、アップストリームデータ機能およびダウンストリームデータ機能のそれぞれに対して、別個のFPGA、または別個のFPGAロジックを提供することができる。このような実施形態により、アップストリームロジックおよび／またはダウンストリームロジックをFPGA内で物理的に小さくすることが可能になり、FPGAロジックで使用されるパイプラインステージを少なくするとともに、レイテンシおよびルーティングの遅延を減らすことが可能になる。アップストリームデータコンポーネントおよびダウンストリームデータコンポーネント用のロジックが、確実にトランシーバのようなハードウェアを共有しないようにすることは、ルーティングの輻輳を緩和するのに有利である。同様の検討事項は、高性能コンピューティング用途のプロセッサコア間の負荷平準化のためのデータスイッチの使用といったような、ピアツーピアの構成にあてはまる。

40

【0025】

本発明は、FPGAを設けることにより、論理ゲートおよびRAMブロックの大規模なリソースを提供して、非常に高速のI/Oおよびデータバスを用いて、単純なデジタル演算処理または複雑なデジタル演算処理のいずれかを実施することを認識している。FPGA

50

Aの論理機能が再構成可能であるので、ポート出力はそれぞれ、場合によっては異なる論理機能を実装させることが可能である。すなわち、単純なフィルタ、複数の入力ストリームの多重化、多重化およびフィルタ、または3値連想メモリ(TCAM)に基づいたもっとも伝統的なレイヤ2スイッチ機能を実装させることができる。さらに、このような機能は、必要に応じて、しかもプログラム可能に、かつ、非常に時宜を得たやり方で更新可能であり、これは、市場取引データを取り扱う場合のような、時間によって変動しやすい再構成要件を有する用途にとって重要な場合がある。本発明のいくつかの実施形態は、したがって、適用可能なハードウェア記述言語などを使用して、FPGAの論理ブロックを再プログラムするように動作可能な管理コンピュータをさらに提供することができることにより、マイクロ秒以内でのフィルタ機能またはハードコードされたルーティング規則の調節といったような、遠隔からの再構築が高速で可能になる。例えば、管理コンピュータは、例えばインターネットを介して遠隔からアクセス可能であることが好ましい、64ビット版x86プロセッサを備えることができる。管理コンピュータは、PCIエクスプレス(PCIe)のような高速直列コンピュータ膨張バスを介して、またはネットワークリンクを介して、FPGAに命令を配信することができる。いくつかの実施形態であれば、そのFPGAまたはある1つのFPGAの作用によって自動的に、フィルタまたはその他の論理機能を内部に構成することも可能であろう。例えば、構成情報は、例えば従来のスイッチにおいてMACアドレス学習機能をエミュレートするような目的で、FPGAのマルチプレクサ間、フィルタ間または論理機能ブロック間で、低帯域幅で渡される場合もあり得る。

10

20

【0026】

いくつかの実施形態では、FPGAは、複数の機能で構成することができ、所望の機能に関連付けられたFPGAポートにデータスイッチングするクロスポイントスイッチによって、所望の機能が適応的に選択されている。

【0027】

クロスポイントからのデータは、FPGAによって提供される任意の適切な論理機能に、例えば、FPGAがレイヤ1で動作可能な単純な通信プロトコルにおけるANDまたはXORのような単純な論理ゲート、または、フィルタリングおよび／または多重化のような、より複雑な組み合わせ機能などに向けることができる。論理機能は、いくつかのパケットフィールドの内容を入力として使用し、次に、必要に応じて、順序論理および組み合わせ論理を用いて、他のパケットフィールドの内容を修正することができる。いくつかの実施形態では、フィルタリングは、任意の1つまたは複数のフィールドに、またはフィールドの組み合わせに基づくことができ、表現式またはプログラムから動的にコンパイルされたフィルタを利用することができ、ディープパケット検査による状態に基づく検査を伴うことができ、または、金融市場データリスクを確認するゲートウェイを備えることができる。論理機能は、レイヤ3ルーティング(例えば、イーサネット(登録商標)/IPでは、IPアドレス情報に基づいてMACアドレスを変換する)、レイヤ4ルーティング(例えば、IP/UDPまたはIP/TCPでは、UDPまたはTCPのポートの情報に基づいてIPアドレスを変換し、NATを実施する)、金融市場データチッカー装置もしくは構文解析器、またはアプリケーションプロトコルトランスレータのような変換を伴う場合がある。論理機能は、アービタを有効にして、例えば多数のソースからデータを供給し、(例えばロスが多いが高性能の通信リンクを、信頼度は高いが速度が遅いリンクと組み合わせるために)重複したものを破棄することができる。いくつかの実施形態では、FPGAは、データストリームのデータのタイムスタンプングを有効にすることにより、レイヤ1機能と同じデバイスの中でレイヤ2機能を提供するように構成することができる。タイムスタンプングを有効にするために、FPGAは、その入力データインタフェース上でパケットの開始を検出し、タイムスタンプを生成することが可能である。このタイムスタンプは、メタデータとしてパケットに追加可能であり、後で処理するために格納し、分析のために別のデバイスに送り、および／またはフレーム検査シーケンス(FCS)またはフレーム間ギャップ(IFG)の代わりにパケットに追加することが可能である。FPG

30

40

50

Aによって、データパケット捕捉機能をさらに提供することができ、これにより、クロスポイントは、データパケット捕捉RAM、ディスクおよび／またはPCIeリンクに、重複したデータストリームを直接的または間接的に配信する。

【0028】

いくつかの実施形態では、FPGAは、外部コネクタへの外部の高速シリアルバス接続をさらに設けることができ、大型のデータメモリ、または大容量コンピューティングといったような、より規模の大きな機能にFPGAを接続可能になる。

【0029】

FPGAは、（概してタイムスタンピングに用いられる）内部クロックが、何らかの規定された時間（通常、秒の開始時）にパルスを生成する外部の時間源（GPS等）に同期されることが可能になるように、受信データシンボルの立ち上がり区間を検出する毎秒パルス数モニタとしての役割を果たすように構成されたロジックをさらに設けることができる。

10

【0030】

クロスポイントスイッチのデータポートは、例えば、10GBASE-Tインタフェース型または1000BASE-Tインタフェース型を使用する銅ケーブル、100BASE-TX銅ケーブル、10GBASE-R／1000BASE-Xを使用する光ファイバケーブル、ダイレクトアタッチ銅ケーブルなどの任意の適切なケーブルからデータを受信することができる。クロスポイントスイッチは、SFP、SFP+、QSFP、QSFP+、CFPのような、任意の適切なプラグ接続可能なモジュールからデータを受信することができる。それらのモジュールであれば、ダイレクトアタッチ銅ケーブル、1Gbaset-SX、10Gbaset-SR、10Gbaset-LR、DWDMまたはその他の接続規格のケーブルを介しての接続も可能であろう。あるいは、コネクタまたはプリント回路基板を使用して、クロスポイントスイッチを別のデバイスに直接接続することも可能であろう。

20

【0031】

クロスポイントスイッチは、FPGAとは物理的に離されたASICとして設けることができる。

いくつかの実施形態では、FPGAのトランシーバは、FPGAおよびクロスポイントスイッチを収容しているデバイスの前面パネルポートに直接接続することができる。あるいは、FPGAのトランシーバは、クロスポイントスイッチに接続され、前面パネルポートから離すことができる。

30

【0032】

本発明の第3の態様、または第4の態様のいくつかの実施形態では、コンピュータコードまたは命令がスイッチの所有者または管理者によって購入可能なサーバ上に常駐し、スイッチの所有者または管理者のハードウェアが、このコンピュータコードまたは命令によって容易になった新たな機能で構成されるようにすることができる。コンピュータコードまたは命令は、ソースコード、FPGAハードウェア記述言語（HDL）、または例えば、HDLにコンパイル可能な上位レベル言語の命令を備えることができる。

【0033】

第5の態様によれば、本発明は、物理層データスイッチング方法であって、スイッチの少なくとも1つの入力ポートでデータを受信すること、データをスイッチの少なくとも1つの出力ポートへと回路スイッチングすること、上位層プロセスにデータをコピーすること、を含む方法を提供する。

40

【0034】

第6の態様によれば、本発明は、物理層データスイッチであって、複数のデータ入力部であって、複数のデータ出力部、すなわち、第1の入力部から回路スイッチングされているデータを受信する第1のデータ出力部、および第1の入力部からデータのコピーを受信する第2のデータ出力部、のうちの1つまたは複数へと設定可能に回路スイッチングすることがそれぞれ可能な複数のデータ入力部と、第2のデータ出力部に接続された上位層機

50

能回路と、を備えるスイッチを提供する。

【0035】

第7の態様によれば、本発明は、データスイッチの構成用の非一時的コンピュータ読み取り可能媒体であって、1つまたは複数のプロセッサによって実行される場合に、スイッチの少なくとも1つの入力ポートで受信したデータを、スイッチの少なくとも1つの出力ポートへと回路スイッチングすること、上位層プロセスにデータをコピーすること、を実行させる命令を含む、非一時的コンピュータ読み取り可能媒体を提供する。

【0036】

第8の態様によれば、本発明は、物理層データスイッチング方法であって、スイッチの少なくとも1つの入力ポートでデータを受信すること、データをスイッチの少なくとも1つの出力ポートへと回路交換すること、上位層プロセスを修正するか、またはデータを補足するとともに、スイッチの第2の出力ポートで第2のデータ出力を生成すること、を含む方法を提供する。

【0037】

したがって本発明の第5～第8の態様は、物理層を処理することにより、結果として低レイテンシおよびプロトコルの汎用性が得られるが、物理層スイッチは有用なカウンタおよび統計を提供せず、これらは上位層でのみ抽出可能であり、したがって、物理層スイッチは、例えば、このようなスイッチを使用してデータリンクを監視し、問題を診断することができないので、ネットワーク管理者にとって不都合があることを認識している。

【0038】

従来の物理層スイッチとは対照的に、本発明の第5～第8の態様の実施形態は、むしろ、データスループットにレイテンシオーバーヘッドをまったく加えずに、低レイテンシおよびプロトコルの汎用性の恩恵を保持しつつ、上位層の機能もまた提供する物理層スイッチを提供する。特に、データを直接入力部から出力部に回路スイッチングすることによって、物理層スイッチは、低レイテンシレベルで動作する。さらに、上位層プロセスに同じデータをコピーすることによって、物理層スイッチングのレイテンシに影響を及ぼさずに、上位レベル監視機能を実装可能である。なぜなら、このような機能は、スイッチングと並行して実行されるからである。例えば、本発明の第5～第8の態様のいくつかの実施形態では、上位層プロセスは、ネットワーク上で使用される一般的なプロトコルである、IPパケットおよびイーサネット（登録商標）フレームを識別することを含むことができる。その他のプロトコルには、ファイバチャネル、インフィニバンド（商標）およびSDIなどが含まれ得る。このような並行処理は、たとえ上位レベル監視カウンタおよび統計には、追加の処理時間がかかっても、これによりデータがスイッチを通過する際にデータを遅延させないことを意味する。本発明の第5～第8の態様のいくつかの実施形態では、カウンタおよび統計のような、上位層プロセスの出力は、スイッチの管理インタフェースを介してアクセスされ、データ経路の低レイテンシを妨げることがない。

【0039】

本明細書において「上位層プロセス」という場合には、物理層で行われないプロセスを指していることを理解されたい。本発明の第5～第8の態様のいくつかの実施形態では、上位層プロセスは、物理層アイドルパターンおよび物理層エラー、例えば、8B/10Bコード違反の検出を含めた、上位レベル処理のための物理層データの変換を含むことができる。

【0040】

本発明の第5～第8の態様のいくつかの実施形態では、複数のデータ入力部において、そして好ましくはすべてのデータ入力部において受信したデータが、1つまたは複数のそれぞれの上位層プロセスにコピーされる。したがってこのような実施形態は、各データ入力部またはすべてのデータ入力部のレイヤ1スイッチングを有効にする一方で、同時に、レイテンシに影響を与えずに、各インタフェースのパケット識別といったような1つの上位層プロセスまたは複数の上位層プロセスを提供することができる。

【0041】

本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、それぞれのインタフェースについてのデータパケット統計のカウンタを備えることができる。

【0042】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、複数の受信および送信されたオクテット（8ビットバイト）のカウンタを備えることができる。

【0043】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、タイプ別（ユニキャスト、マルチキャスト、およびブロードキャスト）カウンタを備えることができる。

10

【0044】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、プロトコル別（イーサネット（登録商標）、IP、UDP、TCP、HTTP、SSH、SNMP、ARPなど）カウンタを備えることができる。

【0045】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、サイズ別カウンタを備えることができる。例えば、64バイトのデータパケットの数、65～127バイトのデータパケットの数、256～511バイトのデータパケットの数などである。

20

【0046】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、無効なフレーム検査シーケンス（FCS）、「ラント（ runt）」パケット、「ジャイアント」パケットおよび／またはアライメントエラーなどに起因するパケットエラーの数のカウンタを備えることができる。

【0047】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、ドロップフレーム数のカウンタを備えることができる。

【0048】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、VLANおよび／またはMACsecカウンタを備えることができる。

30

【0049】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、リンクアイドル状態の検出を備えることができる。

【0050】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、アイドルパターンを注入する能力、またはアイドルパターンが配線上传送されるようにする能力を備えることができる。

【0051】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、物理層エラー、例えば、8B／10Bコードエラーについてのカウンタを備えることができ、そこから物理層ビット誤り率の推定値を算出可能である。

40

【0052】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、スイッチングデバイスに関するヒューリスティックスを提供するパケット検査を備えることができ、例えば、パケットMACアドレスを検査することによって、ある1つの接続または各接続のもう一方の端部に存在するホストおよび／またはデバイスを識別する。

【0053】

50

本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、受信および送信のいずれかのインタフェース上で、または両方のインタフェース上で捕捉されたデータに関して適用することができる。

【0054】

本発明の第5～第8の態様のいくつかの実施形態では、統計カウンタ値などの上位層プロセス出力は、スイッチの管理モジュールによって報告可能である。このような実施形態のいくつかでは、管理モジュール機能は、コマンドライン、ウェブインタフェース、および／または、別個のコンピュータが公開されたカウンタ値を問い合わせ、または受信することが可能なSNMPのような、その他のメカニズムを含むことが可能である。

【0055】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、パケットをタイムスタンプする能力を備えることができる。このような実施形態は、スイッチ内でタイムスタンプすることは、ネットワークレイテンシの測定のために、および正確な記録のために有用であると認識している。これらの正確な記録は、検査後に、試験／モデル化／ネットワークトラフィック履歴の再現に使用可能である。イーサネット（登録商標）パケットを取り扱っている実施形態では、そのようなパケットが正確なタイムスタンプ情報のためのフィールドを有していないことに注目すれば、パケットのサイズを大きくすることによって、タイムスタンプ情報をパケットに追加する（タイムスタンプデータを先頭、または後尾に追加する）ことができる。

【0056】

あるいは、タイムスタンプデータは、他の情報のために使用されるパケットの一部を置き換えることにより、各パケットに追加することができる。本発明の第5～第8の態様のこのような実施形態には、完全に飽和したリンク、すなわち、チャンネル上にそれ以上大きなパケットのための余地がまったくないリンク上で、パケットをタイムスタンプすることが可能であるという恩恵がある。パケットのサイズに影響を及ぼさないアプローチの1つは、イーサネット（登録商標）フレームのフレーム検査シーケンス（FCS）フィールドを32ビットのタイムスタンプ値で置き換えることである。FCSの検査後、エラーがないことが判明してはじめて、これを行うことができる。もしもFCSがエラーを起こしていることが判明すれば、影響を受けたパケットでは特別なタイムスタンプ値を用いることが可能である。この特別なタイムスタンプは、有効なパケットには決して生じないものである。32ビットのタイムスタンプは、基準瞬間がやはりわかっている別個のプロセスによる時間の絶対単位として解釈可能な既知の基準瞬間時からのナノ秒数を表すことも可能であろう。例えば、32ビットのタイムスタンプは、絶対時間の最下位32ビットをナノ秒で表すことができる。別の実施形態では、イーサネット（登録商標）のフレームのプリアンブルは、タイムスタンプ値で部分的に置き換えることができる。

【0057】

したがって、データストリームのデータのコピーをタイムスタンプする本発明の第5～第8の態様の実施形態は、レイヤ1スイッチングと同一のデバイス内でレイヤ2機能を提供する。タイムスタンプを有効にするために、デバイスは、その入力データインタフェース上でパケットの開始を検出し、タイムスタンプを生成することが可能である。このタイムスタンプは、メタデータとしてパケットに追加することが可能であり、後で処理するために格納したり、分析のために別のデバイスに送ったり、および／またはフレーム検査シーケンス（FCS）またはフレーム間ギャップ（IFG）の代わりにパケットに追加したりすることが可能である。

【0058】

本発明の第5～第8の態様の実施形態では、上位層機能は、パケットのタイムスタンプを備えるが、いくつかの実施形態では、上位層デバイスは、タイムスタンプされた、元のデータストリームの遅延されたコピーを備える第2のデータストリームを出力することができる。したがって、このような実施形態は、低レイテンシでデータを出力部に配信する一方で、タイムスタンプ情報が追加された同じデータの遅延されたコピーもまた生成

10

20

30

40

50

し、これらは、例えば、エンドユーザが別途分析を行うように、ロギングデバイス／コンピュータに接続することも可能であろう。

【0059】

本発明の第5～第8の態様の実施形態では、上位層機能は、パケットのタイムスタンプを備えるが、タイムスタンプ値は、スイッチの受信インタフェース上でパケットの開始が検出された瞬間にカウンタレジスタをサンプリングすることにより、取得することができる。他の実施形態では、タイムスタンプは、スイッチの出力インタフェース上で実行することができる。カウンタレジスタは、規則的に（例えばナノ秒毎に）更新されることが好ましい。その結果、カウンタレジスタは、現在時間をナノ秒で正確に表現したものとなる。カウンタは、1秒間隔パルス（PPS）信号、高精度時間プロトコル（PTP）などを介して、スイッチ外部の外部時間基準に同期させることができる。あるいは、タイムスタンプ値は、パケットの開始後に固定時間オフセットで得ることができ、パケットの開始が受信された実時間が、次に、タイムスタンプ値から固定時間オフセットを減算することにより取得することができる。このような実施形態では、この減算は、タイムスタンプがFCSフィールドに書き込まれる前に実行可能である。またはこの代わりに、タイムスタンプされたデータストリームを後処理するときに実行可能である。

【0060】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、データパケット捕捉を備えることができる。これにより、重複したデータストリームの一部またはすべてが、メディアアクセスコントローラ（MAC）を介して処理され、次に、データパケット捕捉RAM、ディスクおよび／またはPCIeリンクに配信される。このような実施形態では、管理プラットフォームは、インタフェースを通過しているパケットを捕捉するように要求される場合があり、その結果、パケットが、遅延のないライン速度で送信される一方で、同時にパケットがコピーされ、ソフトウェア／管理ポートを介してエンドユーザに報告されることになる。いくつかの実施形態では、上位層機能は、MAC処理なしにデータを捕捉することができる。一実施形態では、上位層機能は、2つ以上のポートから捕捉されたデータを集約することができる。

【0061】

本発明の第5～第8の態様のいくつかの実施形態では、データは、1つまたは複数のサーバを起点とするダウンストリームデータとすることができ、1つまたは複数のクライアントデバイスを対象とすることができ、これに加え、またはこの代わりに、データは、1つまたは複数のクライアントデバイスを起点とするアップストリームのデータとすることができ、1つまたは複数のサーバを対象とすることができ、これに加え、またはこの代わりに、データは、高性能コンピューティングクラスターのノード間で、またはデータセンタネットワークのノード間で切り換えられるデータのような、ピアツーピアのデータとすることができる。

【0062】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、ファイアウォールの場合のディープパケット検査、侵入検出、または流出検出を備えることができる。

【0063】

これに加え、またはこの代わりに、本発明の第5～第8の態様のいくつかの実施形態では、上位層機能は、MAC（メディアアクセスコントローラ）の出力を物理層に渡しながら、インタフェースMACにパケットを注入する能力を備えることができる。すなわち、このような実施形態では、ユーザソフトウェアは、管理インタフェースを介してスイッチにパケットを通信することができる。また、指定されたインタフェースについては、上位層機能は、受信MACまたは送信MACのいずれかにパケットを注入することが必要な場合がある。次に、パケットは、インタフェースから配線路上を送信されることになる。この代わりに、パケットは、ソフトウェアまたはハードウェア（例えばFPGA）によってデバイス上で生成することができる。すなわち、管理インタフェースを通じて通信されるの

ではなくデバイスを起点としている。このような実施形態のいくつかは、例えば、ダウンストリームのホストに代わってマルチキャストに加入する (subscribe) ように、インターネットソサエティの RFC 2236 の主題である、インターネットグループ管理プロトコル (IGMP) バージョン 2 に従って、および／または、プロトコル独立マルチキャスト (PIM) ルーティングプロトコルに従って、および／または、境界ゲートウェイプロトコル (BGP) などに従って、1つのパケットまたは複数のパケットを注入することができる。IGMP の場合には、マルチキャストグループに対する加入または退会は、本発明のスイッチのポートに接続されたダウンストリームのクライアントに代わってマルチキャストパケットを受信するマルチキャストグループに登録するように、近くのマルチキャストルータに送信するために、IGMP パケットを注入することを含むことができる。これにより、ダウンストリームのホストにこの機能を実行させる必要がなくなる。

【0064】

本発明の第5～第8の態様のいくつかの実施形態では、スイッチのデータポートは、例えば、10GBASE-T インタフェース型または 1000BASE-T インタフェース型を使用する銅ケーブル、100BASE-TX 銅ケーブル、10GBASE-R / 1000BASE-X を使用する光ファイバケーブル、ダイレクトアタッチ銅ケーブルなどの任意の適切なケーブルからデータを受信することができる。スイッチは、SFP、SFP+、QSFP、QSFP+、CFP のような、任意の適切なプラグ接続可能なモジュールからデータを受信することができる。それらのモジュールであれば、ダイレクトアタッチ銅ケーブル、1GBase-SX、10GBase-SR、10GBase-LR、DWDM またはその他の接続性規格のケーブルを介しての接続も可能であろう。あるいは、コネクタまたはプリント回路基板を使用して、スイッチを別のデバイスに直接接続することが可能であろう。

【0065】

本発明の第5～第8の態様のいくつかの実施形態では、データがコピーされる上位層機能は、スイッチの複数の出力部であって、それぞれの上位層機能にそれぞれが接続されている複数の出力部のうちの1つに、コピーされたデータストリームをスイッチングすることによって、適応的に選択することができる。このような実施形態により、どの上位層機能、またはどの複数の上位層機能が、コピーされたデータストリームに適用されるかの動的な選択が可能になる。このような実施形態では、上位層機能の出力をクロスポイントスイッチに戻して、上位層機能の出力をクロスポイントの任意の選択された出力部に動的、かつ、設定可能にスイッチング可能にすることができる。

【0066】

本発明の第5～第8の態様のいくつかの実施形態では、クロスポイントスイッチの入力部および／または出力部に位置している信号調整ブロックによって、データがコピーされる上位層機能を有効にすることができる。

次に、添付の図面を参照して、本発明の一例を説明する。

【図面の簡単な説明】

【0067】

【図1】本発明の1つの実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図2】図1のシステムにおけるアップストリームデータの取り扱いを図示する。

【図3】図1のシステムにおけるダウンストリームデータの取り扱いを図示する。

【図4】図1のシステムにおけるダウンストリームデータのブロードキャストを図示する。

【図5】本発明の別の実施形態の、再構成可能なダウンストリームデータのフィルタリングを提供しているデータ流れ図である。

【図6】本発明の別の実施形態の、再構成可能なアップストリームデータのフィルタリングを提供しているデータ流れ図である。

【図7】単一のFPGAに2つ以上のマルチプレクサが使用されている、別の実施形態を

図示する。

【図 8 a】本発明の別の実施形態に従って完全にクロス接続されているスイッチを図示する。

【図 8 b】図 8 a の実施形態の 1 つのパケットの取り扱いを図示する。

【図 9】本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 10】本発明の別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 11】本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 12】本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 13】本発明のさらなる実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 14】図 13 の実施形態に従う信号調整ブロックの概略図である。

【発明を実施するための形態】

【0068】

図 1 は、本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。それは、他のデバイスに接続可能な数個の前面パネルポートと、前面パネルポートに交信可能なトランシーバを有する数個の F P G A と、1 つのポートから複数のポートにデータを再配置することだけでなくマルチキャストすることも可能なクロスポイントスイッチとで構成されている。なお、クロスポイントの出力がすべて、必ずしも F P G A の論理機能に接続されているわけではない。ポートによっては、外部のデバイスにデータ出力を提供することが必要になるからである。

【0069】

図 1 のデバイスのアップストリームコンポーネントについては、デバイスは、多数の入力ストリームを得てそれらを 1 つに多重化する F P G A を提供する。次に、この多重化された単一のストリームは、クロスポイントスイッチを通して返送される。このデバイスには他のロジックがほとんどないので、輻輳がほとんどない。それは、短いパイプライン、または F P G A の論理コンポーネントのグループ化を用いて、高いクロックレートで F P G A が駆動可能であることを意味する。すなわち、データ経路を単純化することにより、各素子の高度な最適化が可能になる。これにより、低レイテンシまたは他のロジックが可能になる。

【0070】

F P G A とクロスポイントスイッチとを図 1 に示されるやり方で組み合わせることによって、クロスポイントスイッチでブロードキャスト機能を有効にし、F P G A でフィルタ機能または他の同様の機能を有効にすることが可能になる。利点は、効率的に低レイテンシで伝達可能なクロスポイントにおいてブロードキャストコンポーネントを伝達することが可能であることであり、フィルタ機能だけが F P G A に残る、すなわち、F P G A 内の各フィルタが入力部および出力部を含む。これにより、複雑なロジックおよびクロッキングが極めてローカルに（場合によってはトランシーバ自体から近距離内に）保たれる。これにより、次には、デバイス内のレイテンシをさらに低くすることが可能になる。それは、ダウンストリームコンポーネント（したがってフィルタロジック）のためのトランシーバが、複数の F P G A の間に分散可能であることもまた意味する。なぜなら、ローカルにトランシーバに伝送する以外に F P G A 内でデータを伝送する必要がないからである。

【0071】

それは、ダウンストリームロジックと相互作用することなく、隣接したトランシーバにアップストリームコンポーネントが実装可能であるという意味でもあり、場合によっては、はるかに効率良い（例えば、より低サイクルかつより低レイテンシの）アップストリームの実装が可能になる。本発明の範囲内の代替的な実施形態では、流れ制御または 100

10

20

30

40

50

O B A S E - X のオートネゴシエーションを実施する場合のように、アップストリームロジックコンポーネントとダウンストリームロジックコンポーネントとの間の相互作用が時には必要な場合がある。

【0072】

アップストリームロジックおよびダウンストリームロジックは、図2に示されるように、別個のFPGAにさえも実装可能である。この構成では、デバイス内全体のはるかに多くのロジックを適合させることが可能である。代替的な諸実施形態は、アップストリーム機能およびダウンストリーム機能をいずれも1つのFPGAまたはASIC上に実装できるが、そのような実施形態においてさえ、アップストリームコンポーネントおよびダウンストリームコンポーネント用のロジックは、ハードウェアリソース（特に、ランシーバ）を共有する必要がなく、したがって、FPGAの同じ箇所に常駐する必要がなく、ルーティングの輻輳を緩和する。

10

【0073】

ダウンストリームコンポーネントが図3に示されている。このコンポーネントでは、ロジックを単純にすることもまた可能である。アップストリームサーバがノードのうちの1つにアドレス指定されたデータを送る際、このデータはクロスポイントスイッチによってFPGA内の2つ以上のフィルタにブロードキャストされる。各フィルタは非常に小型であり必要なものすべてを完備し、また、ランシーバ間でFPGAを横断する長距離にわたるデータのルーティングがない。フィルタに対する受信および送信はいずれも、FPGA上の同じランシーバに実装可能である。フィルタはクロスポイントから受信したデータに関して動作し、フィルタがデータを通過させる場合にだけ、データがダウンストリームのクライアントに送信（クロスポイントを介して返信）される。

20

【0074】

すべてのクライアントが、フィルタリングとは無関係に、すべてのデータを受信しなければならないのであれば、図4に示されるように、クロスポイントは、データをFPGAに渡さずに、クライアントにデータを直接ブロードキャストすることが可能である。

【0075】

本発明のさらなる実施形態では、複数のアップストリームのサーバを取り扱うか、または完全クロス接続されているスイッチを作り出すために、多重化機能およびフィルタ機能を組み合わせることが可能である。

30

【0076】

図5は、本発明の一実施形態の、再構成可能なダウンストリームデータのフィルタリングを提供しているデータ流れ図である。ダウンストリームの場合については、示されているようなデータの流れて、データが、アップストリームのサーバ502から数個のダウンストリームのクライアント512、514に送信される。データは、サーバ502から送信され、外部ポートで受信され、そこでクロスポイントスイッチ520に入る。データは、クロスポイント520によって、FPGA530の、フィルタ536、538について1つのランシーバ532、534にマルチキャストされる。ポートがすべて異なるフィルタリング要件を有する場合には、示されているように、ダウンストリームのクライアント512、514について1つのフィルタを使用することが可能である。それ以外の場合には、クロスポイントスイッチ520を適切に構成することにより、FPGA530の単一のフィルタから所望する数のクライアントまたはポートにブロードキャスト可能である。図5の実施形態は、2つのフィルタ、すなわち1つのダウンストリームのポートについて1つのフィルタを備える。フィルタ536、538はそれぞれ、自身に関連付けされたランシーバ532、534を有しており、受信および送信の両方（すなわち、TXが復元クロックにより駆動される）を1つのクロックドメインで実行するか、または複数のクロックドメインでFIFOもしくは他のクロックドメインクロッシングを用いて実行可能である。レイテンシを可能な限り低くするためであれば、前者の方法が好ましいが、実施形態によっては、例えば、完全にデータを再クロックし、送信にとって最良の信号品質を提供するためには、後者の方法を用いる場合がある。要求される信号品質に応じて、選択さ

40

50

れるスキームが決まる。図 5 に、1 つのクロックドメインに実装されているフィルタを示す。

【0077】

クロスポイントスイッチ 520 によって、1 つ 1 つのフィルタ 536、538 にデータが送信され、そこでそれぞれのトランシーバ 532、534 によって受信される。トランシーバ 532、534 はそれぞれ、データを非直列化する。フィルタロジック 536、538 は、そのデータを解釈し、データが、求められる基準に合致する場合、例えば、そのフィルタに関連付けされたクライアントにデータがアドレス指定される場合には、フィルタ 536 について示されているように、フィルタは、同じトランシーバを使用してデータを再送信する。このような構成であるため、ロジック 536、538 は、それぞれのトランシーバ 532、534 に密接させることが可能である。データがフィルタ基準に合致しない場合、例えば、そのフィルタに関連付けされたクライアントにデータがアドレス指定されない場合には、フィルタ 538 について示されているように、データは再送信されない。クロスポイントスイッチ 520 は、管理コンピュータ 540 により、クライアント 512、514 に関連付けされた適切なポートにフィルタ 536、538 の出力を伝送するように構成されている。管理コンピュータ 540 は、フィルタ 536、538 が、クライアント 512、514 に関連付けされたポートに送信されるべきデータを選択するようにもまた、フィルタ 536、538 を構成する。フィルタロジック 536、538 は、求められるフィルタリング基準、および求められる複雑度のレベルに応じて、非常に単純な、また非常に複雑なものとするのが可能である。

10

20

【0078】

図 6 は、本発明の一実施形態の、再構成可能なアップストリームデータのフィルタリングを提供しているデータ流れ図である。アップストリームの場合では、クロスポイント 620 が作用して、クライアント 612、614 に関連付けされたポートから FPG A 630 の独立したトランシーバ 632、634 に、データを転送する。トランシーバ 632、634 の送信機能は使用されず、受信機能だけが使用される。トランシーバ 632、634 からの非直列化されたデータは、それぞれの F I F O 636、638 を通して受信され、F I F O 636、638 はそれぞれ、クロッシングクロックドメインのために使用される。多重化ロジック 640 を用いて、F I F O 636、638 からデータを受信する。このとき、642 で示されている単一の出力ストリームにデータの packets を多重化することができる。これは、別のトランシーバ 635 を介して送られるが、代替的な実施形態では、トランシーバ 635 の代わりに受信トランシーバ 632、634 のうちの 1 つの送信機を使用して、1 つのトランシーバを使用せずにすませることも可能であろう。

30

【0079】

図 5 および図 6 では、アップストリームロジックが、FPG A のダウンストリームロジックから物理的に離されているので有利である。これにより、FPG A 内でロジックを物理的により小さくすることが可能になることで、パイプラインの段が減り、レイテンシを低くすることが可能になる。

【0080】

図 7 は、単一の FPG A に 2 つ以上のマルチプレクサが使用されている、本発明の別の実施形態を図示する。クライアントから複数のマルチプレクサ 740、742 に packets をブロードキャストし、別個のフィルタ 750、752 でそれらのマルチプレクサの出力をフィルタリングすることによって、この実施形態は、複数のアップストリームサーバ 760、762 を取り扱うことが可能である。2 つのサーバを有して示されているが、他の実施形態は、類似したやり方でもっと多数のサーバが調達可能であることが理解されよう。したがって、この実施形態は、完全に接続されているスイッチの相互接続性を提供しつつ、本発明の利点もまた提供する。図 7 では、クライアントデバイス 712、714 を起点とするアップストリーム packets は、クロスポイント 720 によって、FPG A 730 に実装された 2 つの別々の多重化モジュール 740、742 にマルチキャストされる。多重化モジュール 740、742 がそれぞれ、packets を多重化し、次に、フィルタ 750

40

50

およびフィルタ752が、2つの多重化されたストリームをフィルタリングして、それぞれのフィルタの基準に合致しないものを廃棄する。図7では、フィルタ752が、サーバB762を対象としたそれぞれのデータストリーム743を渡していることが示されているのに対して、フィルタ750は、データストリーム741がサーバA760を対象としていないので、データストリーム741を廃棄していることが示されている。図7の実施形態の変形では、多重化ロジック740、742の後ではなく前に、4つのフィルタを提供してフィルタリングを実施することにより、帯域幅を大きくすることが可能であろう。さらに、用途によっては、各サーバが、そのサーバにアドレス指定されていないデータを廃棄することが予想される場合には、多重化されたデータをサーバ760およびサーバ762の両方に渡すことが容認され得る場合もある。この場合、フィルタ750およびフィルタ752は省略することができるので、レイテンシはさらに低減するであろう。

10

【0081】

図1～図7の実施形態は、このようにクロスポイントスイッチを使用して、アナログドメインのいくつかのポイントにデータをマルチキャストする。同じデータの複数のコピーを作り出すことによって、このような実施形態はそれぞれ、同じデバイス内で様々なFPGAもしくはASICにわたって、または単一のFPGAもしくはASICの様々な部分にわたって、様々な機能に対するロジックを分散させることができる。

【0082】

図8aは、本発明の別の実施形態に従って完全にクロス接続されている(3x3)スイッチを図示する。示されるように、多重化(mux)およびフィルタ論理機能の拡張は、さらに多くのマルチプレクサおよびフィルタをFPGAに追加することである。これにより、各クライアントから2つ以上のホストにパケットを送ることが可能になる。この実施形態では、各クライアントインタフェースが、固有のマルチプレクサに関連付けされており、これにより、任意の他のポートにパケットを送信可能になる。各ポートに対して専用のフィルタが設けられ、この専用のフィルタを使用して、各ポートによって受信されるパケットが、そのポートにアドレス指定されたパケットだけであるように制限する。このアーキテクチャを用いて、従来のイーサネット(登録商標)スイッチを実装可能である。すなわち、フィルタは、フィルタリングの対象となっているポートに接続されたネットワークに接続されていることが既知であるアドレスにアドレス指定されたパケットだけを転送するように構成されている。

20

30

【0083】

図8bは、図8aの実施形態の1つのパケットの取り扱いを図示する。この例では、フィルタは、関連するクライアントにアドレス指定されたパケットだけを受信し、パケットは、示された経路を通るように設定されている。クライアントCは、宛先アドレスが「クライアントA」であるパケットを送信する。クロスポイントスイッチ820は、FPGA830の3つの別々のランシーバにパケットをマルチキャストするように構成されている。パケットは、ランシーバによって受信され、3つの独立した多重化ロジックのセット840、842、844に渡される。多重化論理機能はそれぞれ、受信データを多重化し、次に、それぞれのフィルタ850、852、854にパケットを送信する。多重化ロジック840に対するフィルタ850は、パケットがクライアントAにアドレス指定されているときに、その関連するランシーバにパケットを送信する。クライアントBおよびCに対するフィルタ852およびフィルタ854は、パケットを送信しない。多重化ロジック840に対するランシーバが、クロスポイントスイッチ820にパケットを送信する。クロスポイントスイッチ820が、フィルタ850からクライアントAに受信データをユニキャストするように構成され、パケットが正しい宛先に到達する。リソースのこうした利用は、帯域幅に対して最適化されていない。なぜなら、N個のポートスイッチを実装するには、 N^2 個のランシーバおよびN個のマルチプレクサが必要になるからである。しかしながら、このようなシステムは、共用リソースを有するシステムよりも低レイテンシで、かつ、良好な決定性で構築可能である。

40

【0084】

50

図 8 の実施形態は、本発明に従った、N 個のポートの完全にクロス接続されたスイッチの実施態様を図示しているが、代替的な実施形態では、データの取り扱いは、各クライアントに対して、またはクロスポイントスイッチの各ポートに対して同じである必要はない。例えば、いくつかの実施形態では、図 8 の実施形態を修正することで、1 つのポートで受信したデータをタイムスタンプする一方で、その他のポートで受信したデータを多重化できるようにすることが可能であろう。

【0085】

図 9 は、本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ 900 は、902 で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1 つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ 910 とを備える。

10

【0086】

それ以外の時には再構成することが可能であるが、図 9 に示される構成では、入力データは、データソース 920 から受信され、クロスポイント 910 によって、意図した宛先 930 に関連付けされたそれぞれのデータ出力部 904 へと回路スイッチングされる。この意味では、スイッチ 900 は、物理層スイッチとして機能している。出力部 904 は、入力部 902 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 902 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 902 および出力部 904 は、複数のトランシーバによって有効にすることができ、複数のトランシーバはそれぞれ、1 つのデータ入力部 902、および 1 つのデータ出力部 904 を有効にする。クロスポイント 910 内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図 9 に示されている。

20

【0087】

本発明に従って、データソース 920 から受信したデータは、宛先 930 に配信されるだけでなく、クロスポイントスイッチ 910 によって、スイッチ 900 内のデバイス 906 によって実行される上位層プロセスにコピーされる。

【0088】

この実施形態では、デバイス 906 がタイムスタンプ機能を備えることにより、デバイス 906 から宛先 930 にデータの第 2 のストリームが配信されるようになっている。このように宛先 930 が、極めて低レイテンシの経路を介して元のデータを受信する一方で、デバイス 906 により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、タイムスタンプされたデータのコピーもまた受信する。

30

【0089】

図 10 は、本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ 1000 は、1002 で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1 つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ 1010 とを備える。

【0090】

それ以外の時には再構成することが可能であるが、図 10 に示される構成では、入力データは、データソース 1020 から受信され、クロスポイント 1010 によって、意図した宛先 1030 に関連付けされたそれぞれのデータ出力部 1004 へと回路スイッチングされる。この意味では、スイッチ 1000 は、物理層スイッチとして機能している。出力部 1004 は、入力部 1002 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 1002 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 1002 および出力部 1004 は、複数のトランシーバによって有効にすることができ、複数のトランシーバはそれぞれ、1 つのデータ入力部 1002、および 1 つのデータ出力部 1004 を有効にする。クロスポイント 1010 内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図 10 に

40

50

示されている。

【0091】

本発明に従って、データソース1020から受信したデータは、宛先1030に配信されるだけでなく、スイッチ1000内のデバイス1006によって実行され、やはり出力部1030へと切り換えられるパケットを生成する上位層プロセスにも配信される。

【0092】

この実施形態では、デバイス1006は、マルチキャスト加入機能を備える。これにより、デバイス1020および／または1022が、マルチキャストのグループに対して加入または退会するために、IGMPパケットが、近くのルータ1030に送られたデータストリームに注入されるようになっている。このように宛先1030は、極めて低レイテンシの経路を介してデバイス1020からデータを受信する一方で、デバイス1006により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、指定されたインタフェースに対するいずれかの送信MACに注入されたIGMPデータパケットのような、上位層機能データもまた受信する。

【0093】

図11は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ1100は、1102で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ1110とを備える。

【0094】

それ以外の時には再構成することが可能であるが、図11に示される構成では、入力データは、データソース1120から受信され、クロスポイント1110によって、意図した宛先1130に関連付けされたそれぞれのデータ出力部1104へと回路スイッチングされる。この意味では、スイッチ1100は、物理層スイッチとして機能している。出力部1104は、入力部1102に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部1102を有するデバイスの前面パネルにこれらをともに配置してもよいことに留意されたい。入力部1102および出力部1104は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部1102、および1つのデータ出力部1104を有効にする。クロスポイント1110内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図11に示されている。

【0095】

本発明に従って、データソース1120から受信したデータは、宛先1130に配信されるだけでなく、クロスポイントスイッチ1110によって、スイッチ1100内のデバイス1106によって実行される上位層プロセスにコピーされる。

【0096】

この実施形態では、デバイス1106は、管理者1132にデータリンク統計を提供するように構成されたリンクコントローラを備える。この実施形態では、デバイス1106は、複数の上位層機能を提供する。これらに含まれるものとしては、物理層アイドルパターンおよび物理層エラー検出機能、タイプ別、プロトコル別、およびサイズ別のデータパケットカウント機能、無効なFCS、パケットサイズエラーまたはアライメントエラーなどの、エラータイプ別データパケットエラーカウント機能、および、接続されているデバイスを識別するパケットMACアドレス検査機能などがある。

【0097】

このように宛先1130が、極めて低レイテンシの経路を介して元のデータを受信する一方で、デバイス1106により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、デバイス1132が、データリンク統計を受信する。

【0098】

図12は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシ

10

20

30

40

50

ステム概略図である。スイッチ 1200 は、1202 で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ 1210 とを備える。

【0099】

それ以外の時には再構成することが可能であるが、図 12 に示される構成では、入力データは、データソース 1220 から受信され、クロスポイント 1210 によって、意図した宛先 1230 に関連付けされたそれぞれのデータ出力部 1204 へと回路スイッチングされる。この意味では、スイッチ 1200 は、物理層スイッチとして機能している。出力部 1204 は、入力部 1202 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 1202 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 1202 および出力部 1204 は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部 1202、および 1つのデータ出力部 1204 を有効にする。クロスポイント 1210 内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図 12 に示されている。

【0100】

本発明に従って、データソース 1220 から受信したデータは、宛先 1230 に配信されるだけでなく、クロスポイントスイッチ 1210 によって、スイッチ 1200 内のデバイス 1206 によって実行される上位層プロセスにコピーされる。この実施形態のデバイス 1206 は、選択された上位層機能にそれぞれが関連付けされた複数の入力部を有する。その結果、コピーされたデータをデバイス 1206 のそれぞれの入力部にスイッチングするだけで、所望の機能を選択可能になる。さらに、デバイス 1206 の出力は、クロスポイント 1210 に戻されることにより、任意の出力ポート 1204 へと選択的にスイッチングできる。また、図 12 に示されている時には、デバイス 1200 は、上位層デバイス 1206 の出力を宛先 1230 へと切り換えるように構成されているが、それ以外の時には、出力を他の宛先へと切り換えることができる。管理システムプロセッサ 1240 は、1秒間隔パルス (PPS) 入力を受信して、タイムスタンプを可能にするとともに、管理インタフェースを介して通信も行っており外部の管理を可能にする。管理システム 1240 はさらに、クロスポイント 1210 および上位層デバイス 1206 を制御し、クロスポイント 1210 および上位層デバイス 1206 と通信する。

【0101】

図 13 は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ 1300 は、1302 で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ 1310 とを備える。

【0102】

それ以外の時には再構成することが可能であるが、図 13 に示される構成では、入力データは、データソース 1320 から受信され、クロスポイント 1310 によって、意図した宛先 1330 に関連付けされたそれぞれのデータ出力部 1304 へと回路交換される。この意味では、スイッチ 1300 は、物理層スイッチとして機能している。出力部 1304 は、入力部 1302 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 1302 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 1302 および出力部 1304 は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部 1302、および 1つのデータ出力部 1304 を有効にする。クロスポイント 1310 内で現在作動中のスイッチング経路が、単に例示の目的で図 13 に太線で示されている。

【0103】

本発明に従って、データソース 1320 から受信したデータは、宛先 1330 に配信さ

れるだけでなく、クロスポイントスイッチ1310によって、スイッチ1300内のデバイス1364によって実行される上位層プロセスにコピーされる。この実施形態のデバイス1364は、図14に示されるタイプの信号調整ブロックである。上位層デバイス1364の出力は、宛先1330に渡されるが、それ以外の時には、出力を他の宛先へと切り換えることができる。管理システムプロセッサ1340は、1秒間隔パルス（PPS）入力を受信して、タイムスタンプングを可能にするとともに、管理インタフェースを介して通信も行っており外部の管理を可能にする。管理システム1340はさらに、クロスポイント1310を制御し、クロスポイント1310と通信して、信号調整ブロック1350～1354および1360～1366のすべてを制御する。

【0104】

図14に示される信号調整ブロック1364では、入力データは、EQブロック1402を通過する。EQブロック1402は、図14全体を通して両矢印によって示されるような管理インタフェース1340を介して設定可能である。EQブロック1402は、受信した信号に関してフィルタリングを実行して、通信チャネル（ケーブル）を通過することの影響を補償することにより、エラーの尤度を減少させる等化器である。EQブロック1402は、信号の周波数成分を調節するとともに、多重路の歪みを除去することもまた可能であり、その出力は、等化されていない受信信号よりもはるかに、データの理想的な表現に近くなっている。EQ1402から、データが、管理インタフェース1340を介して、設定可能なクロックデータ回復（CDR）ブロック1404に渡される。CDR1404から、データが、上位層機能1406およびマルチプレクサ1408の両方に渡される。これらはいずれも、管理システムプロセッサ1340によって制御されている。上位層機能ブロック1406の出力もまた、マルチプレクサ1408に渡される。マルチプレクサ1408から、多重化された低レイテンシのデータおよび上位層機能の出力データが、プリエンファシスブロック1410に渡されることにより、見込まれる通信チャネル（ケーブル）の影響を前もって補償するように送信機信号の周波数成分を修正し、その結果、宛先がデータの理想的な表現により近い信号を受信することで、エラーの尤度が減少するようになる。プリエンファシスブロック1410から、多重化された低レイテンシのデータおよび上位層機能の出力データが、信号調整ブロック1364によって出力される。図13の実施形態の信号調整ブロック1350～1354および1360～1366のいずれかまたはすべてが、図14に示されるように構成することができる。

【0105】

広義に記載されている本発明の趣旨または範囲から逸脱することなく、特定の実施形態で示した本発明に多くの変形および／または修正を行い得ることが、当業者には理解されるであろう。例えば、40Gのようなマルチレーン標準に関連して適用される実施形態では、これらは同様に扱うことが可能である。例えば、4つのトランシーバのグループで40Gに対してフィルタリングを行うことが可能であり、1つまたは複数のクロスポイントデバイスの4つのレーンすべてに対してブロードキャストを行うことが可能である。したがって、本実施形態は、あらゆる点で例示として見なされるべきであり、限定または制限するものと見なされるべきではない。

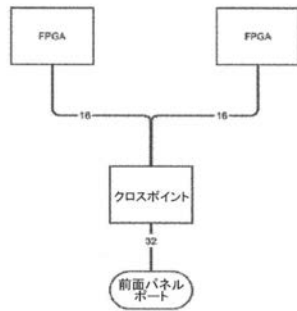
10

20

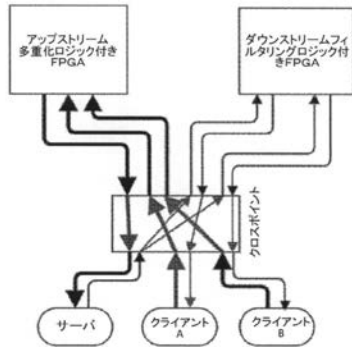
30

40

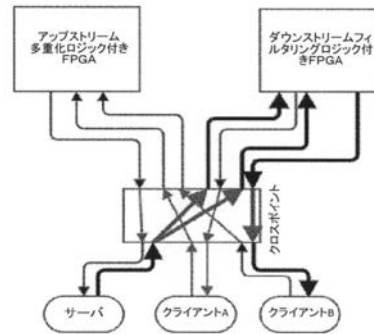
【図 1】



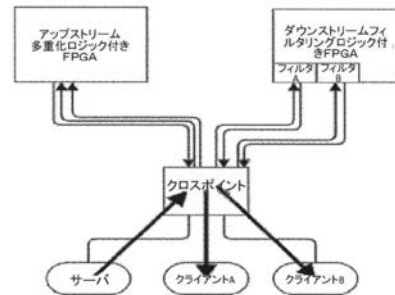
【図 2】



【図 3】



【図 4】



【図 5】

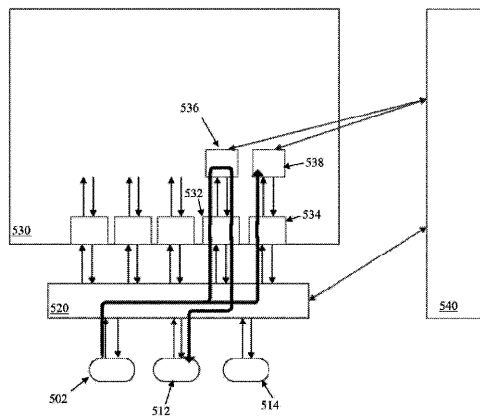
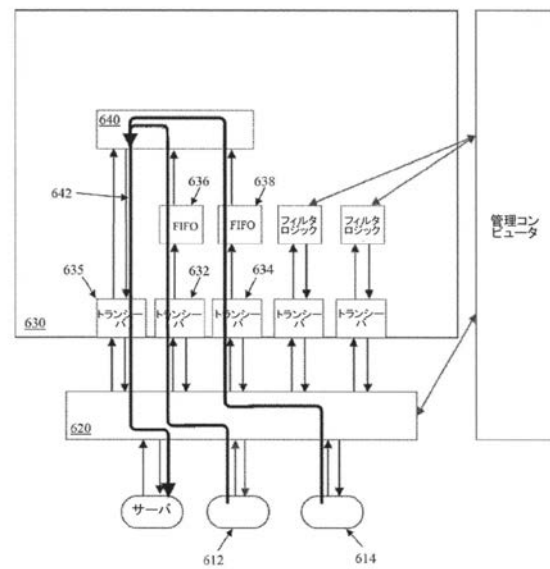
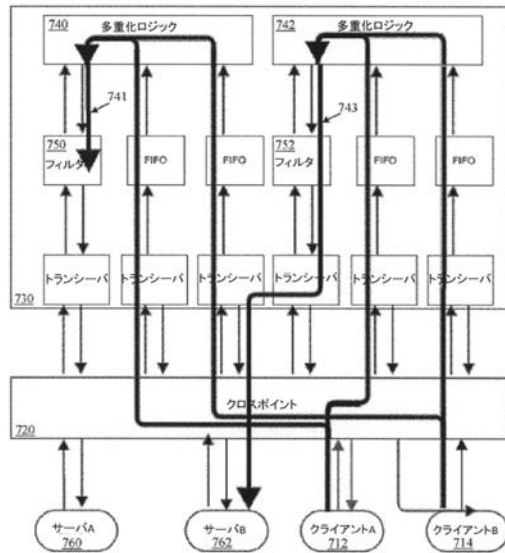


Figure 5

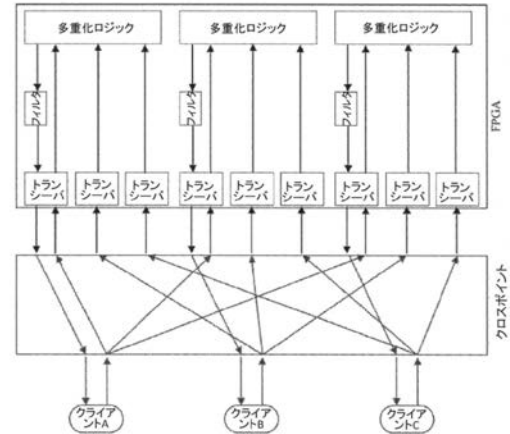
【図 6】



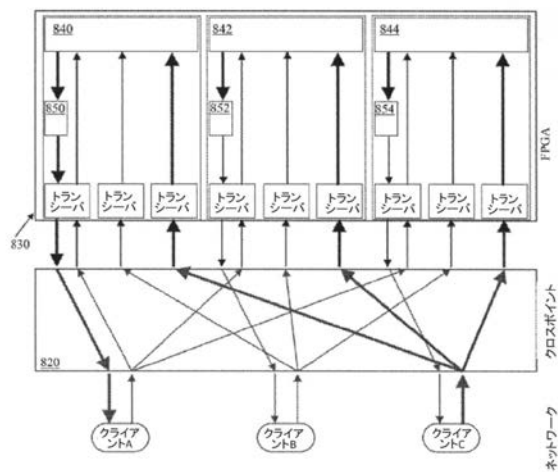
【図 7】



【図 8 a】



【図 8 b】



【図 10】

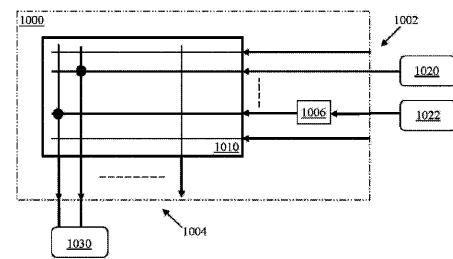


Figure 10

【図 11】

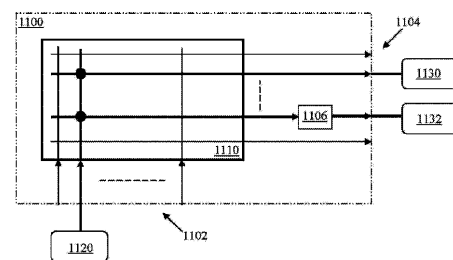


Figure 11

【図 9】

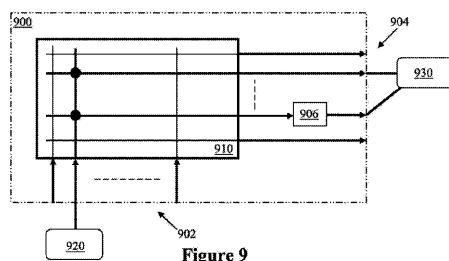
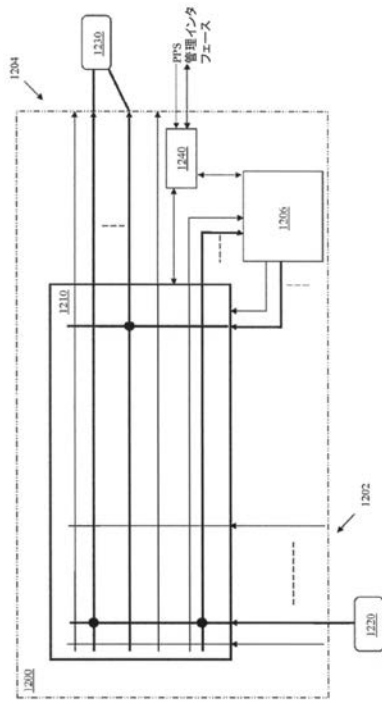
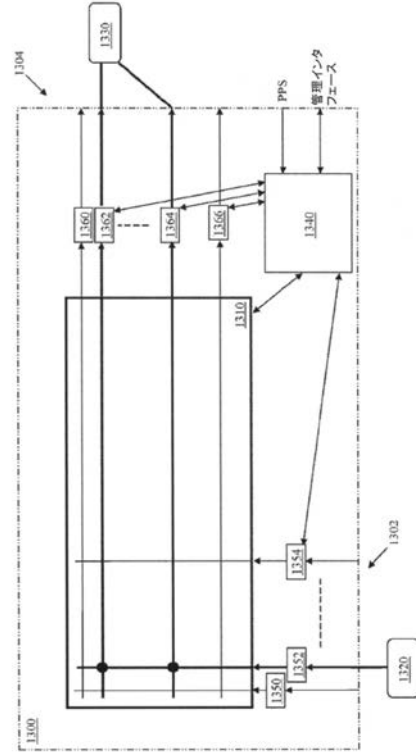


Figure 9

【図 1 2】



【図 1 3】



【図 1 4】

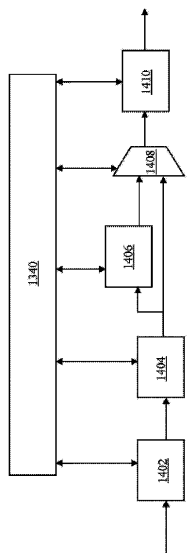


Figure 14

【手続補正書】

【提出日】平成28年6月14日(2016.6.14)

【手続補正 1】

【補正対象書類名】明細書

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワークデータスイッチングに関し、特に、FPGAまたはASICなどの1つまたは複数の論理機能デバイスを有する1つまたは複数のクロスポイントスイッチの組み合わせを用いて所望のネットワーク機能を実装する効率的な低レイテンシデータスイッチに関する。

【背景技術】

【0002】

1つのソースからデータパケットを得て多数の送信可能な送信先のうちの1つまたは複数にデータパケットを送信することが望ましい場合が多い。例えば、証券取引所からダウンストリーム方向への場合がこれに該当し、株式トレーダのコンピュータなどの多数の機械が取引所からパケットを受信できるようにしておくことが望ましい。

【0003】

アプリケーションによっては、1つのポートから多数のポートにブロードキャストまたはマルチキャストすることが適切である。マルチキャストは、デバイス（スイッチ）の入力部での個々のパケットを当該スイッチの1つ1つの出力部にブロードキャストするが、これは、例えば、クロスポイントスイッチを用いることによって達成可能である。しかしながら、パケットスイッチの場合のような他の状況では、2つのエンドポイント間のユニキャスト通信はプライベートでなければならない。

【0004】

アドレス指定可能なスイッチングは、何らかの仕組み（例えば従来のパケットスイッチにおける連想メモリ（content addressable memory）等）を用いてダウンストリームパケットを処理することにより、集積回路に実装可能である。従来、これには、単一のトランシーバでデータを受信し、パケットアドレスを確認し、ルックアップテーブルを参照してそのアドレスに関連付けされたポートを識別し、次に、そのポート宛てだけにパケットを送ることを必要とする。このアプローチは、使用する帯域幅を最適化する考え方に基づくが、比較的複雑な回路がスイッチ内に必要である。

【発明の概要】

【発明が解決しようとする課題】

【0005】

現行のネットワークスイッチは、レイヤ1（物理層）のデータが、レイヤ1に返送される前に、アドレスを確認するためにレイヤ2（データリンク層）に変換されることが必要である。レイヤ1からレイヤ2へのデータの変換には、データを並列ストリームに非直列化することが必要である。なぜなら、関連する周波数では直列ストリームでレイヤ2の特徴を実装することが難しいからである。これは、あらゆるレイヤ2処理は、レイヤ1よりもはるかに速度が遅く、最も高速のレイヤ2ネットワークスイッチでさえ、レイヤ2のアドレス確認におよそ200～400nsの、かなりの遅延が発生することを意味する。複数のトレーダの通信データ量を単一のストリームへとアップストリームに集約するために用いる際には、メタマコLP社製のメタコネクトC16スイッチに加えて、「多対1」スイッチ、例えばイーサネット（登録商標）スイッチを使用することが必要で、レイテンシが大きくなり望ましくない。さらに、ネットワークスイッチは、単一のデバイスのアップストリームデータ経路にもダウンストリームデータ経路にも対処しなければならない場合

が多く、信号経路が複雑になり、クロックの制限をもたらす。既存のスイッチの多くは、帯域幅を最適化し、多くのユースケース間で汎用的に用いるように設計されている。

【0006】

本明細書に含まれている文献、行為、材料、デバイス、物品またはその他のいずれかの説明は、単に本発明に技術背景を提供することを目的とする。これらの事項のいずれか、もしくはすべてが本出願の各請求項の優先日より前に存在したことにより、従来技術の基礎の一部を形成していると、または、本発明に関連する分野において一般によく知られた知見であると認めるものではない。

【0007】

本明細書全体を通して、「comprise（備える、含む）」という語、または「comprisesまたはcomprising」などの変化形は、記載された要素、整数もしくはステップ、または要素、整数もしくはステップの群が含まれることを含意すると理解されるが、任意の他の要素、整数もしくはステップ、または要素、整数もしくはステップの群を排除するものではない。

【0008】

本明細書では、ある要素が、列記された選択肢「の少なくとも1つ」であってもよいという記述は、その要素が、列記された選択肢のうちのいずれか1つであってもよい、または、列記された選択肢のうちの2つ以上の任意の組み合わせであってもよいことと理解されたい。

【課題を解決するための手段】

【0009】

第1の態様によれば、本発明は、データスイッチング方法であって、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信し、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストし、前記クロスポイントスイッチの前記複数の出力ポートの各々を、論理機能デバイスの対応する入力部に接続し、前記論理機能デバイスが前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの複数の対応する出力インタフェースに出力することを備える、データスイッチング方法を提供する。

【0010】

第2の態様によれば、本発明は、データスイッチであって、データを受信するための少なくとも1つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを前記複数の出力ポートに設定可能にマルチキャストするように構成されているクロスポイントスイッチと、前記クロスポイントスイッチの前記複数の出力ポートの各々にそれぞれが接続された複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも1つの論理機能を有する、前記論理機能デバイスとを備えるデータスイッチを提供する

【0011】

第3の態様によれば、本発明は、データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は1つまたは複数のプロセッサによって実行される場合に、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信すること、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にマルチキャストすること、前記クロスポイントスイッチの前記複数の出力ポートの各々を論理機能デバイスの対応する入力部に接続すること、前記論理機能デバイスが、前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力することを実行させる、非一時的コンピュータ読み取り可能媒体を提供する。

【0012】

第4の態様によれば、本発明は、データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの1つを選択するためのコンピュータプログラムコード手段と、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを、当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記複数の出力ポートの各々が論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段とを備えるコンピュータプログラム製品を提供する。

【0013】

第5の態様によれば、本発明は、データスイッチング方法であって、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信し、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストし、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを、論理機能デバイスの対応する入力部に接続し、前記論理機能デバイスが前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力することを備える方法を提供する。

【0014】

第2の態様によれば、本発明は、データを受信するための少なくとも1つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを前記出力ポートの少なくとも1つに設定可能にキャストするように構成されているクロスポイントスイッチと、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートにそれぞれ接続された、1つまたは複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも1つの論理機能を有する、前記論理機能デバイスとを備えるデータスイッチを提供する。

【0015】

第6の態様によれば、本発明は、データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は1つまたは複数のプロセッサによって実行される場合に、クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信すること、前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストすること、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートを論理機能デバイスの対応する入力部に接続すること、前記論理機能デバイスが、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力することを実行させる、非一時的コンピュータ読み取り可能媒体を提供する。

【0016】

第7の態様によれば、本発明は、データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの1つを選択するためのコンピュータプログラムコード手段と、クロスポイ

ントスイッチの少なくとも1つの入力ポートで受信したデータを、当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートが論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、前記クロスポイントスイッチの前記出力ポートまたは各出力ポートから受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの1つまたは複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段とを備えるコンピュータプログラム製品を提供する。

【0017】

本発明のいくつかの実施形態では、クロスポイントスイッチの少なくとも1つの出力ポートにデータを設定可能にキャストするクロスポイントスイッチは、クロスポイントスイッチの複数の出力ポートにデータをマルチキャストするクロスポイントスイッチを備えることができる。いくつかの実施形態では、クロスポイントスイッチの少なくとも1つの出力ポートにデータを設定可能にキャストするクロスポイントスイッチは、クロスポイントスイッチの1つの出力ポートにデータをユニキャストするクロスポイントスイッチを備えることができる。

【0018】

論理機能デバイスは、フィールドプログラマブルゲートアレイ（FPGA）、適切に構成されたx86CPU系のプロセッサ、もしくはネットワークインタフェースが組み込まれたマルチコアマイクロプロセッサのような適切に構成された特定用途向け集積回路（ASIC）、または複数のこのようなデバイスを備えることができる。本明細書でのFPGAを利用する実施形態についての説明は、本発明の範囲内でASICまたは適切に構成されたプロセッサにもまた適用可能なものであるとして理解されたい。

【0019】

本発明のいくつかの実施形態は、帯域幅使用の最適化を断念することによって、低レイテンシの向上、および／または決定性向上（MACコリジョンの減少）のための最適化をこのように効果的に得ることができる。クロスポイントスイッチのブロードキャスト機能を使用して2つ以上の受信機にデータを分配することによって、可能総帯域幅が減少し、個々のハードウェア要素のデータスループットは必ずしも最大化されないが、本発明は、こうした複数の同一のデータストリームの並列処理が、他の有益な効果を有することを認識している。したがって本発明は、既存のスイッチの多くは、帯域幅を最適化し、多くのユースケース間で汎用的に使用するように設計されているが、このことがレイテンシおよび決定性を犠牲にしていると認識している。

【0020】

データは、1つまたは複数のサーバを起点とするダウンストリームのデータとすることができ、1つまたは複数のクライアントデバイスを対象とすることができ。これに加え、またはこの代わりに、データは、1つまたは複数のクライアントデバイスを起点とするアップストリームのデータとすることができ、1つまたは複数のサーバを対象とすることができ。これに加え、またはこの代わりに、データは、高性能コンピューティングクラスタのノード間で、またはデータセンタネットワークのノード間で切り換えられるデータのような、ピアツーピアのデータとすることができ。

【0021】

クロスポイントスイッチの出力ポートの、FPGAの入力ポートへの接続は、いくつかの実施形態では、直接的な接続を備えることができる。あるいは他の実施形態では、AC結合用キャパシタのような回路素子、またはレベルトランスレーション、バッファリング、増幅、もしくは信号回復を実行するICを間に介在させることができる。

【0022】

FPGA出力インタフェース、または各FPGA出力インタフェースは、出力ポート、

周辺部品相互接続エクスプレス（P C I e）出力部、他の出力インタフェース、または高速通信用インタフェースのストリームの群、のうちの1つまたは複数を備えることができる。

【0023】

いくつかの実施形態では、データは、サーバのような遠隔した場所から受信したダウンストリームデータであって、クロスポイントスイッチにとってローカルな1つまたは複数のクライアントデバイスに分配するように意図されたダウンストリームデータとすることができる。デバイスは、たとえスイッチと物理的に同一の場所になくても、スイッチに接続され、スイッチを介してより広いネットワークとインタフェースしているのであれば、スイッチにとってローカルであると思なされる。このような実施形態では、FPGAによって適用される論理機能は、アドレスフィルタリングのような、内容ベースフィルタリングを備えることができ、これにより、クロスポイントスイッチが、クロスポイントスイッチの複数の出力ポートにデータをマルチキャストするとともに、複数のクライアントデバイスが、複数のFPGAの出力ポートにそれぞれ関連付けされていることによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけがそれぞれのFPGAの出力ポートから出力されるように、FPGAが、クロスポイントスイッチの各出力ポートからのデータをフィルタする。したがって、このような実施形態は、効率的にブロードキャストするためにクロスポイントスイッチを利用し、効率的なアドレス指定、およびその他の機能のためにFPGAを利用する。内容ベースフィルタリングであれば、L2、L3またはL4から宛先またはソースアドレスを参照することによるアドレス指定を伴う場合もあるが、プロトコルに基づくフィルタリング、ファイアウォールの場合の、状態に基づくフィルタリングおよびディープパケット検査、侵入検出もしくは流出検出システム、フィルタリング対象のパケットの無効化、またはFPGAに実装される任意の他の論理機能によるフィルタリングを含むこともまた可能であろう。クロスポイントスイッチを用いて、FPGAの2つ以上の機能ブロックにデータストリームをコピーすることにより、複数の機能を同じデータに関してFPGA内で同時に実行可能になる。このような実施形態では、FPGAの単一のランシーバが、各フィルタについて受信および送信の両方の役割を果たすことができる。あるいは、FPGA出力ポートに接続されているクライアントデバイスのような別のデバイスが代わりに、このようなアドレスフィルタリング用に構成されている実施形態では、FPGAは、アドレスフィルタリングせずに出力ポートにデータを渡すことができる。したがって、このような実施形態は、各データパケットをアドレス指定するために、帯域幅が最適化されたレイヤ2ルックアップアプローチと対照をなす、「ブロードキャストアンドフィルタ」アプローチを用いる。

【0024】

他の実施形態では、データは、クロスポイントスイッチにとってローカルな1つまたは複数のクライアントデバイスから受信したアップストリームデータであって、1つまたは複数のサーバのような、1つまたは複数の遠隔した場所への配信を意図されたアップストリームデータとすることができる。このような実施形態では、FPGAによって提供される論理機能は、多重化していてもよい。多重化は、それぞれのデータクロックを同期するために、それぞれのFIFOバッファ、またはメモリキューにFPGAの各受信機入力部を接続し、2つ以上のクライアントデバイスから単一のアップストリームデータストリームへとデータを多重化するマルチプレクサに、各FIFOまたはメモリキューの出力を渡すことにより、FPGAによって提供することができる。次に、アップストリームデータストリームは、FPGAのポートで送信機によって出力され、外部ネットワークの前面パネルポートに配信するためのクロスポイントスイッチに、すなわちもっと一般的には、クライアントデバイスに渡されてもよい。このクライアントデバイスに含まれる得るものとしては、例えば、スイッチの内部の、ネットワークインタフェースカードがある。いくつかの実施形態では、FPGAロジックは、ファイアウォール、IPアドレス指定のようなレイヤ3もしくは4の機能、または、場合によっては並行して実装される、パケット変換を実施するように構成することができる。

【0025】

いくつかの実施形態は、ローカルデバイスからFPGAの複数のマルチプレクサにアップストリームデータをマルチキャストするクロスポイントスイッチと、アドレスフィルタを有する各マルチプレクサの出力をフィルタリングすることにより、データがそれぞれの外部ネットワークポートに対してアドレス指定された場合に、それぞれのマルチプレクサからそのポートにだけデータが中継されるようにすることによって、ローカルデバイスから複数の遠隔した場所へのデータ配信をさらに提供することができる。

【0026】

いくつかの実施形態では、クロスポイントスイッチは、アップストリームデータおよびダウンストリームデータの両方を取り扱うことができ、これにより、データのストリームの一方が、クロスポイントスイッチによってFPGAの方に向けられているのに対して、データのもう一方でストリームが、クロスポイントスイッチの前面ポート間で1つまたは複数のクライアントデバイスに直接接続されている。

【0027】

FPGAは、トランシーバ毎に専用のルーティングテーブルを提供することにより、レイヤ3（ネットワーク層）ルーティングを実装するように構成することができる。ハードウェアは非能率的であるが、このような実施形態には、すべてのルーティング用に使用される単一のルーティングテーブルを使用しているレイヤ3ルータと比較して、かなり大きなレイテンシの利点がある。

【0028】

本発明の実施形態は、アップストリームデータ機能およびダウンストリームデータ機能のそれぞれに対して、別個のFPGA、または別個のFPGAロジックを提供することができる。このような実施形態により、アップストリームロジックおよび／またはダウンストリームロジックをFPGA内で物理的に小さくすることが可能になり、FPGAロジックで使用されるパイプラインステージを少なくするとともに、レイテンシおよびルーティングの遅延を減らすことが可能になる。アップストリームデータコンポーネントおよびダウンストリームデータコンポーネント用のロジックが、確実にトランシーバのようなハードウェアを共有しないようにすることは、ルーティングの輻輳を緩和するのに有利である。同様の検討事項は、高性能コンピューティング用途のプロセッサコア間の負荷平準化のためのデータスイッチの使用といったような、ピアツーピアの構成にあてはまる。

【0029】

本発明は、FPGAを設けることにより、論理ゲートおよびRAMブロックの大規模なリソースを提供して、非常に高速のI/Oおよびデータバスを用いて、単純なデジタル演算処理または複雑なデジタル演算処理のいずれかを実施することを認識している。FPGAの論理機能が再構成可能であるので、ポート出力はそれぞれ、場合によっては異なる論理機能を実装させることが可能である。すなわち、単純なフィルタ、複数の入力ストリームの多重化、多重化およびフィルタ、または3値連想メモリ（TCAM）に基づいたものと伝統的なレイヤ2スイッチ機能を実装させることができる。さらに、このような機能は、必要に応じて、しかもプログラム可能に、かつ、非常に時宜を得たやり方で更新可能であり、これは、市場取引データを取り扱う場合のような、時間によって変動しやすい再構成要件を有する用途にとって重要な場合がある。本発明のいくつかの実施形態は、したがって、適用可能なハードウェア記述言語などを使用して、FPGAの論理ブロックを再プログラムするように動作可能な管理コンピュータをさらに提供することができることにより、マイクロ秒以内でのフィルタ機能またはハードコードされたルーティング規則の調節といったような、遠隔からの再構築が高速で可能になる。例えば、管理コンピュータは、例えばインターネットを介して遠隔からアクセス可能であることが好ましい、64ビット版x86プロセッサを備えることができる。管理コンピュータは、PCIエクスプレス（PCIe）のような高速直列コンピュータ膨張バスを介して、またはネットワークリンクを介して、FPGAに命令を配信することができる。いくつかの実施形態であれば、そのFPGAまたはある1つのFPGAの作用によって自動的に、フィルタまたはその他の論

理機能を内部に構成することも可能であろう。例えば、構成情報は、例えば従来のスイッチにおいてMACアドレス学習機能をエミュレートするような目的で、FPGAのマルチプレクサ間、フィルタ間または論理機能ブロック間で、低帯域幅で渡される場合もあり得る。

【0030】

いくつかの実施形態では、FPGAは、複数の機能で構成することができ、所望の機能に関連付けられたFPGAポートにデータスイッチングするクロスポイントスイッチによって、所望の機能が適応的に選択されている。

【0031】

クロスポイントからのデータは、FPGAによって提供される任意の適切な論理機能に、例えば、FPGAがレイヤ1で動作可能な単純な通信プロトコルにおけるANDまたはXORのような単純な論理ゲート、または、フィルタリングおよび／または多重化のような、より複雑な組み合わせ機能などに向けることができる。論理機能は、いくつかのパケットフィールドの内容を入力として使用し、次に、必要に応じて、順序論理および組み合わせ論理を用いて、他のパケットフィールドの内容を修正することができる。いくつかの実施形態では、フィルタリングは、任意の1つまたは複数のフィールドに、またはフィールドの組み合わせに基づくことができ、表現式またはプログラムから動的にコンパイルされたフィルタを利用することができ、ディープパケット検査による状態に基づく検査を伴うことができ、または、金融市場データリスクを確認するゲートウェイを備えることができる。論理機能は、レイヤ3ルーティング（例えば、イーサネット（登録商標）／IPでは、IPアドレス情報に基づいてMACアドレスを変換する）、レイヤ4ルーティング（例えば、IP／UDPまたはIP／TCPでは、UDPまたはTCPのポートの情報に基づいてIPアドレスを変換し、NATを実施する）、金融市場データチッカー装置もしくは構文解析器、またはアプリケーションプロトコルトランスレータのような変換を伴う場合がある。論理機能は、アービタを有効にして、例えば多数のソースからデータを供給し、（例えばロスが多いが高性能の通信リンクを、信頼度は高いが速度が遅いリンクと組み合わせるために）重複したものを破棄することができる。いくつかの実施形態では、FPGAは、データストリームのデータのタイムスタンプングを有効にすることにより、レイヤ1機能と同じデバイスの中でレイヤ2機能を提供するように構成することができる。タイムスタンプングを有効にするために、FPGAは、その入力データインタフェース上でパケットの開始を検出し、タイムスタンプを生成することが可能である。このタイムスタンプは、メタデータとしてパケットに追加可能であり、後で処理するために格納し、分析のために別のデバイスに送り、および／またはフレーム検査シーケンス（FCS）またはフレーム間ギャップ（IFG）の代わりにパケットに追加することが可能である。FPGAによって、データパケット捕捉機能をさらに提供することができ、これにより、クロスポイントは、データパケット捕捉RAM、ディスクおよび／またはPCIeリンクに、重複したデータストリームを直接的または間接的に配信する。

【0032】

いくつかの実施形態では、FPGAは、外部コネクタへの外部の高速シリアルバス接続をさらに設けることができ、大型のデータメモリ、または大容量コンピューティングといったような、より規模の大きな機能にFPGAを接続可能になる。

【0033】

FPGAは、（概してタイムスタンプングに用いられる）内部クロックが、何らかの規定された時間（通常、秒の開始時）にパルスを生成する外部の時間源（GPS等）に同期されることが可能になるように、受信データシンボルの立ち上がり区間を検出する毎秒パルス数モニタとしての役割を果たすように構成されたロジックをさらに設けることができる。

【0034】

クロスポイントスイッチのデータポートは、例えば、10GBASE-Tインタフェース型または1000BASE-Tインタフェース型を使用する銅ケーブル、100BAS

E-TX銅ケーブル、10GBASE-R/1000BASE-Xを使用する光ファイバケーブル、ダイレクトアタッチ銅ケーブルなどの任意の適切なケーブルからデータを受信することができる。クロスポイントスイッチは、SFP、SFP+、QSFP、QSFP+、CFPのような、任意の適切なプラグ接続可能なモジュールからデータを受信することができる。それらのモジュールであれば、ダイレクトアタッチ銅ケーブル、1GBase-SX、10GBase-SR、10GBase-LR、DWDMまたはその他の接続性規格のケーブルを介しての接続も可能であろう。あるいは、コネクタまたはプリント回路基板を使用して、クロスポイントスイッチを別のデバイスに直接接続することも可能であろう。

【0035】

クロスポイントスイッチは、FPGAとは物理的に離されたASICとして設けることができる。

いくつかの実施形態では、FPGAのトランシーバは、FPGAおよびクロスポイントスイッチを収容しているデバイスの前面パネルポートに直接接続することができる。あるいは、FPGAのトランシーバは、クロスポイントスイッチに接続され、前面パネルポートから離すことができる。

【0036】

本発明の第3の態様、または第4の態様のいくつかの実施形態では、コンピュータコードまたは命令がスイッチの所有者または管理者によって購入可能なサーバ上に常駐し、スイッチの所有者または管理者のハードウェアが、このコンピュータコードまたは命令によって容易になった新たな機能で構成されるようにすることができる。コンピュータコードまたは命令は、ソースコード、FPGAハードウェア記述言語(HDL)、または例えば、HDLにコンパイル可能な上位レベル言語の命令を備えることができる。

【0037】

第8の態様によれば、本発明は、物理層データスイッチング方法であって、スイッチの少なくとも1つの入力ポートでデータを受信すること、データをスイッチの少なくとも1つの出力ポートへと回路スイッチングすること、上位層プロセスにデータをコピーすること、を含む方法を提供する。

【0038】

第9の態様によれば、本発明は、物理層データスイッチであって、複数のデータ入力部であって、複数のデータ出力部、すなわち、第1の入力部から回路スイッチングされているデータを受信する第1のデータ出力部、および第1の入力部からデータのコピーを受信する第2のデータ出力部、のうちの1つまたは複数へと設定可能に回路スイッチングすることがそれぞれ可能な複数のデータ入力部と、第2のデータ出力部に接続された上位層機能回路と、を備えるスイッチを提供する。

【0039】

第10の態様によれば、本発明は、データスイッチの構成用の非一時的コンピュータ読み取り可能媒体であって、1つまたは複数のプロセッサによって実行される場合に、スイッチの少なくとも1つの入力ポートで受信したデータを、スイッチの少なくとも1つの出力ポートへと回路スイッチングすること、上位層プロセスにデータをコピーすること、を実行させる命令を含む、非一時的コンピュータ読み取り可能媒体を提供する。

【0040】

第11の態様によれば、本発明は、物理層データスイッチング方法であって、スイッチの少なくとも1つの入力ポートでデータを受信すること、データをスイッチの少なくとも1つの出力ポートへと回路交換すること、上位層プロセスを修正するか、またはデータを補足するとともに、スイッチの第2の出力ポートで第2のデータ出力を生成すること、を含む方法を提供する。

【0041】

したがって本発明の第8～第11の態様は、物理層を処理することにより、結果として

低レイテンシおよびプロトコルの汎用性が得られるが、物理層スイッチは有用なカウンタおよび統計を提供せず、これらは上位層でのみ抽出可能であり、したがって、物理層スイッチは、例えば、このようなスイッチを使用してデータリンクを監視し、問題を診断することができないので、ネットワーク管理者にとって不都合があることを認識している。

【0042】

従来の物理層スイッチとは対照的に、本発明の第8～第11の態様の実施形態は、むしろ、データスループットにレイテンシオーバーヘッドをまったく加えずに、低レイテンシおよびプロトコルの汎用性の恩恵を保持しつつ、上位層の機能もまた提供する物理層スイッチを提供する。特に、データを直接入力部から出力部に回路スイッチングすることによって、物理層スイッチは、低レイテンシレベルで動作する。さらに、上位層プロセスに同じデータをコピーすることによって、物理層スイッチングのレイテンシに影響を及ぼさずに、上位レベル監視機能を実装可能である。なぜなら、このような機能は、スイッチングと並行して実行されるからである。例えば、本発明の第8～第11の態様のいくつかの実施形態では、上位層プロセスは、ネットワーク上で使用される一般的なプロトコルである、IPパケットおよびイーサネット（登録商標）フレームを識別することを含むことができる。その他のプロトコルには、ファイバチャネル、インフィニバンド（商標）およびSDIなどが含まれ得る。このような並行処理は、たとえ上位レベル監視カウンタおよび統計には、追加の処理時間がかかっても、これによりデータがスイッチを通過する際にデータを遅延させないことを意味する。本発明の第8～第11の態様のいくつかの実施形態では、カウンタおよび統計のような、上位層プロセスの出力は、スイッチの管理インタフェースを介してアクセスされ、データ経路の低レイテンシを妨げることがない。

【0043】

本明細書において「上位層プロセス」という場合には、物理層で行われないプロセスを指していることを理解されたい。本発明の第8～第11の態様のいくつかの実施形態では、上位層プロセスは、物理層アイドルパターンおよび物理層エラー、例えば、8B/10Bコード違反の検出を含めた、上位レベル処理のための物理層データの変換を含むことができる。

【0044】

本発明の第8～第11の態様のいくつかの実施形態では、複数のデータ入力部において、そして好ましくはすべてのデータ入力部において受信したデータが、1つまたは複数のそれぞれの上位層プロセスにコピーされる。したがってこのような実施形態は、各データ入力部またはすべてのデータ入力部のレイヤ1スイッチングを有効にする一方で、同時に、レイテンシに影響を与えずに、各インタフェースのパケット識別といったような1つの上位層プロセスまたは複数の上位層プロセスを提供することができる。

【0045】

本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、それぞれのインタフェースについてのデータパケット統計のカウンタを備えることができる。

【0046】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、複数の受信および送信されたオクテット（8ビットバイト）のカウンタを備えることができる。

【0047】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、タイプ別（ユニキャスト、マルチキャスト、およびブロードキャスト）カウンタを備えることができる。

【0048】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、プロトコル別（イーサネット（登録商標）、IP、UDP、TCP、HTTP、SSH、SNMP、ARPなど）カウンタを備えることができる。

【0049】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、複数のパケットの、サイズ別カウンタを備えることができる。例えば、64バイトのデータパケットの数、65～127バイトのデータパケットの数、256～511バイトのデータパケットの数などである。

【0050】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、無効なフレーム検査シーケンス（FCS）、「ラント（runt）」パケット、「ジャイアント」パケットおよび／またはアライメントエラーなどに起因するパケットエラーの数のカウンタを備えることができる。

【0051】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、ドロップフレーム数のカウンタを備えることができる。

【0052】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、VLANおよび／またはMACsecカウンタを備えることができる。

【0053】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、リンクアイドル状態の検出を備えることができる。

【0054】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、アイドルパターンを注入する能力、またはアイドルパターンが配線上に送信されるようにする能力を備えることができる。

【0055】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、物理層エラー、例えば、8B／10Bコードエラーについてのカウンタを備えることができ、そこから物理層ビット誤り率の推定値を算出可能である。

【0056】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、スイッチングデバイスに関するヒューリスティックスを提供するパケット検査を備えることができ、例えば、パケットMACアドレスを検査することによって、ある1つの接続または各接続のもう一方の端部に存在するホストおよび／またはデバイスを識別する。

【0057】

本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、受信および送信のいずれかのインタフェース上で、または両方のインタフェース上で捕捉されたデータに関して適用することができる。

【0058】

本発明の第8～第11の態様のいくつかの実施形態では、統計カウンタ値などの上位層プロセス出力は、スイッチの管理モジュールによって報告可能である。このような実施形態のいくつかでは、管理モジュール機能は、コマンドライン、ウェブインタフェース、および／または、別個のコンピュータが公開されたカウンタ値を問い合わせ、または受信することが可能なSNMPのような、その他のメカニズムを含むことが可能である。

【0059】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、パケットをタイムスタンプする能力を備えることができる。このような実施形態は、スイッチ内でタイムスタンプすることは、ネットワークレイテンシの測定のために、および正確な記録のために有用であると認識している。これらの正確な記録は、検査後に、試験／モデル化／ネットワークトラフィック履歴の再現に使用可能であ

る。イーサネット（登録商標）パケットを取り扱っている実施形態では、そのようなパケットが正確なタイムスタンプ情報のためのフィールドを有していないことに注目すれば、パケットのサイズを大きくすることによって、タイムスタンプ情報をパケットに追加する（タイムスタンプデータを先頭、または後尾に追加する）ことができる。

【0060】

あるいは、タイムスタンプデータは、他の情報のために使用されるパケットの一部を置き換えることにより、各パケットに追加することができる。本発明の第8～第11の態様のこのような実施形態には、完全に飽和したリンク、すなわち、チャンネル上にそれ以上大きなパケットのための余地がまったくないリンク上で、パケットをタイムスタンプすることが可能であるという恩恵がある。パケットのサイズに影響を及ぼさないアプローチの1つは、イーサネット（登録商標）フレームのフレーム検査シーケンス（FCS）フィールドを32ビットのタイムスタンプ値で置き換えることである。FCSの検査後、エラーがないことが判明してはじめて、これを行うことができる。もしもFCSがエラーを起こしていることが判明すれば、影響を受けたパケットでは特別なタイムスタンプ値を用いることが可能である。この特別なタイムスタンプは、有効なパケットには決して生じないものである。32ビットのタイムスタンプは、基準瞬間がやはりわかっている別個のプロセスによる時間の絶対単位として解釈可能な既知の基準瞬間時からのナノ秒数を表すことも可能であろう。例えば、32ビットのタイムスタンプは、絶対時間の最下位32ビットをナノ秒で表すことができる。別の実施形態では、イーサネット（登録商標）のフレームのプリアンブルは、タイムスタンプ値で部分的に置き換えることができる。

【0061】

したがって、データストリームのデータのコピーをタイムスタンプする本発明の第8～第11の態様の実施形態は、レイヤ1スイッチングと同一のデバイス内でレイヤ2機能を提供する。タイムスタンプを有効にするために、デバイスは、その入力データインタフェース上でパケットの開始を検出し、タイムスタンプを生成することが可能である。このタイムスタンプは、メタデータとしてパケットに追加することが可能であり、後で処理するために格納したり、分析のために別のデバイスに送ったり、および／またはフレーム検査シーケンス（FCS）またはフレーム間ギャップ（IFG）の代わりにパケットに追加したりすることが可能である。

【0062】

本発明の第8～第11の態様の実施形態では、上位層機能は、パケットのタイムスタンプを備えるが、いくつかの実施形態では、上位層デバイスは、タイムスタンプされた、元のデータストリームの遅延されたコピーを備える第2のデータストリームを出力することができる。したがって、このような実施形態は、低レイテンシでデータを出力部に配信する一方で、タイムスタンプ情報が追加された同じデータの遅延されたコピーもまた生成し、これらは、例えば、エンドユーザが別途分析を行うように、ロギングデバイス／コンピュータに接続することも可能であろう。

【0063】

本発明の第8～第11の態様の実施形態では、上位層機能は、パケットのタイムスタンプを備えるが、タイムスタンプ値は、スイッチの受信インタフェース上でパケットの開始が検出された瞬間にカウンタレジスタをサンプリングすることにより、取得することができる。他の実施形態では、タイムスタンプは、スイッチの出力インタフェース上で実行することができる。カウンタレジスタは、規則的に（例えばナノ秒毎に）更新されることが好ましい。その結果、カウンタレジスタは、現在時間をナノ秒で正確に表現したものとなる。カウンタは、1秒間隔パルス（PPS）信号、高精度時間プロトコル（PTP）などを介して、スイッチ外部の外部時間基準に同期させることができる。あるいは、タイムスタンプ値は、パケットの開始後に固定時間オフセットで得ることができ、パケットの開始が受信された実時間が、次に、タイムスタンプ値から固定時間オフセットを減算することにより取得することができる。このような実施形態では、この減算は、タイムスタンプがFCSフィールドに書き込まれる前に実行可能である。またはこの代わりに、タ

タイムスタンプされたデータストリームを後処理するときに実行可能である。

【0064】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、データパケット捕捉を備えることができる。これにより、重複したデータストリームの一部またはすべてが、メディアアクセスコントローラ（MAC）を介して処理され、次に、データパケット捕捉RAM、ディスクおよび／またはPCIEリンクに配信される。このような実施形態では、管理プラットフォームは、インタフェースを通過しているパケットを捕捉するように要求される場合があり、その結果、パケットが、遅延のないライン速度で送信される一方で、同時にパケットがコピーされ、ソフトウェア／管理ポートを介してエンドユーザに報告されることになる。いくつかの実施形態では、上位層機能は、MAC処理なしにデータを捕捉することができる。一実施形態では、上位層機能は、2つ以上のポートから捕捉されたデータを集約することができる。

【0065】

本発明の第8～第11の態様のいくつかの実施形態では、データは、1つまたは複数のサーバを起点とするダウンストリームデータとすることができ、1つまたは複数のクライアントデバイスを対象とすることができ、これに加え、またはこの代わりに、データは、1つまたは複数のクライアントデバイスを起点とするアップストリームのデータとすることができ、1つまたは複数のサーバを対象とすることができ、これに加え、またはこの代わりに、データは、高性能コンピューティングクラスタのノード間で、またはデータセンタネットワークのノード間で切り換えられるデータのような、ピアツーピアのデータとすることができ、

【0066】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、ファイアウォールの場合のディープパケット検査、侵入検出、または流出検出を備えることができる。

【0067】

これに加え、またはこの代わりに、本発明の第8～第11の態様のいくつかの実施形態では、上位層機能は、MAC（メディアアクセスコントローラ）の出力を物理層に渡ししながら、インタフェースMACにパケットを注入する能力を備えることができる。すなわち、このような実施形態では、ユーザソフトウェアは、管理インタフェースを介してスイッチにパケットを通信することができる。また、指定されたインタフェースについては、上位層機能は、受信MACまたは送信MACのいずれかにパケットを注入することが必要な場合がある。次に、パケットは、インタフェースから配線上を送信されることになる。この代わりに、パケットは、ソフトウェアまたはハードウェア（例えばFPGA）によってデバイス上で生成することができる。すなわち、管理インタフェースを通じて通信されるのではなくデバイスを起点としている。このような実施形態のいくつかは、例えば、ダウンストリームのホストに代わってマルチキャストに加入する（subscribe）ように、インターネットソサエティのRFC2236の主題である、インターネットグループ管理プロトコル（IGMP）バージョン2に従って、および／または、プロトコル独立マルチキャスト（PIM）ルーティングプロトコルに従って、および／または、境界ゲートウェイプロトコル（BGP）などに従って、1つのパケットまたは複数のパケットを注入することができる。IGMPの場合には、マルチキャストグループに対する加入または退会は、本発明のスイッチのポートに接続されたダウンストリームのクライアントに代わってマルチキャストパケットを受信するマルチキャストグループに登録するように、近くのマルチキャストルータに送信するために、IGMPパケットを注入することを含むことができる。これにより、ダウンストリームのホストにこの機能を実行させる必要がなくなる。

【0068】

本発明の第8～第11の態様のいくつかの実施形態では、スイッチのデータポートは、例えば、10GBASE-Tインタフェース型または1000BASE-Tインタフェー

ス型を使用する銅ケーブル、100BASE-TX銅ケーブル、10GBASE-R/1000BASE-Xを使用する光ファイバケーブル、ダイレクトアタッチ銅ケーブルなどの任意の適切なケーブルからデータを受信することができる。スイッチは、SFP、SFP+、QSFP、QSFP+、CFPのような、任意の適切なプラグ接続可能なモジュールからデータを受信することができる。それらのモジュールであれば、ダイレクトアタッチ銅ケーブル、1GbBase-SX、10GbBase-SR、10GbBase-LR、DWDMまたはその他の接続性規格のケーブルを介しての接続も可能であろう。あるいは、コネクタまたはプリント回路基板を使用して、スイッチを別のデバイスに直接接続することが可能であろう。

【0069】

本発明の第8～第11の態様のいくつかの実施形態では、データがコピーされる上位層機能は、スイッチの複数の出力部であって、それぞれの上位層機能にそれぞれが接続されている複数の出力部のうちの1つに、コピーされたデータストリームをスイッチングすることによって、適応的に選択することができる。このような実施形態により、どの上位層機能、またはどの複数の上位層機能が、コピーされたデータストリームに適用されるかの動的な選択が可能になる。このような実施形態では、上位層機能の出力をクロスポイントスイッチに戻して、上位層機能の出力をクロスポイントの任意の選択された出力部に動的、かつ、設定可能にスイッチング可能にすることができる。

【0070】

本発明の第8～第11の態様のいくつかの実施形態では、クロスポイントスイッチの入力部および／または出力部に位置している信号調整ブロックによって、データがコピーされる上位層機能を有効にすることができる。

次に、添付の図面を参照して、本発明の一例を説明する。

【図面の簡単な説明】

【0071】

【図1】本発明の1つの実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図2】図1のシステムにおけるアップストリームデータの取り扱いを図示する。

【図3】図1のシステムにおけるダウンストリームデータの取り扱いを図示する。

【図4】図1のシステムにおけるダウンストリームデータのブロードキャストを図示する。

【図5】本発明の別の実施形態の、再構成可能なダウンストリームデータのフィルタリングを提供しているデータ流れ図である。

【図6】本発明の別の実施形態の、再構成可能なアップストリームデータのフィルタリングを提供しているデータ流れ図である。

【図7】単一のFPGAに2つ以上のマルチプレクサが使用されている、別の実施形態を図示する。

【図8a】本発明の別の実施形態に従って完全にクロス接続されているスイッチを図示する。

【図8b】図8aの実施形態の1つのパケットの取り扱いを図示する。

【図9】本発明の1つの実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図10】本発明の別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図11】本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図12】本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図13】本発明のさらなる実施形態を実施するために用いられるハードウェアのシステム概略図である。

【図 1 4】図 1 3 の実施形態に従う信号調整ブロックの概略図である。

【発明を実施するための形態】

【0072】

図 1 は、本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。それは、他のデバイスに接続可能な数個の前面パネルポートと、前面パネルポートに交信可能なトランシーバを有する数個の F P G A と、1 つのポートから複数のポートにデータを再配置することだけでなくマルチキャストすることも可能なクロスポイントスイッチとで構成されている。なお、クロスポイントの出力がすべて、必ずしも F P G A の論理機能に接続されているわけではない。ポートによっては、外部のデバイスにデータ出力を提供することが必要になるからである。

【0073】

図 1 のデバイスのアップストリームコンポーネントについては、デバイスは、多数の入力ストリームを得てそれらを 1 つに多重化する F P G A を提供する。次に、この多重化された単一のストリームは、クロスポイントスイッチを通して返送される。このデバイスには他のロジックがほとんどないので、輻輳がほとんどない。それは、短いパイプライン、または F P G A の論理コンポーネントのグループ化を用いて、高いクロックレートで F P G A が駆動可能であることを意味する。すなわち、データ経路を単純化することにより、各素子の高度な最適化が可能になる。これにより、低レイテンシまたは他のロジックが可能になる。

【0074】

F P G A とクロスポイントスイッチとを図 1 に示されるやり方で組み合わせることによって、クロスポイントスイッチでブロードキャスト機能を有効にし、F P G A でフィルタ機能または他の同様の機能を有効にすることが可能になる。利点は、効率的に低レイテンシで伝達可能なクロスポイントにおいてブロードキャストコンポーネントを伝達することが可能であることであり、フィルタ機能だけが F P G A に残る、すなわち、F P G A 内の各フィルタが入力部および出力部を含む。これにより、複雑なロジックおよびクロッキングが極めてローカルに（場合によってはトランシーバ自体から近距離内に）保たれる。これにより、次には、デバイス内のレイテンシをさらに低くすることが可能になる。それは、ダウンストリームコンポーネント（したがってフィルタロジック）のためのトランシーバが、複数の F P G A の間に分散可能であることもまた意味する。なぜなら、ローカルにトランシーバに伝送する以外に F P G A 内でデータを伝送する必要がないからである。

【0075】

それは、ダウンストリームロジックと相互作用することなく、隣接したトランシーバにアップストリームコンポーネントが実装可能であるという意味でもあり、場合によっては、はるかに効率良い（例えば、より低サイクルかつより低レイテンシの）アップストリームの実装が可能になる。本発明の範囲内の代替的な実施形態では、流れ制御または 1000BASE-X のオートネゴシエーションを実施する場合のように、アップストリームロジックコンポーネントとダウンストリームロジックコンポーネントとの間の相互作用が時には必要な場合がある。

【0076】

アップストリームロジックおよびダウンストリームロジックは、図 2 に示されるように、別個の F P G A にさえも実装可能である。この構成では、デバイス内全体のはるかに多くのロジックを適合させることが可能である。代替的な諸実施形態は、アップストリーム機能およびダウンストリーム機能をいずれも 1 つの F P G A または A S I C 上に実装できるが、そのような実施形態においてさえ、アップストリームコンポーネントおよびダウンストリームコンポーネント用のロジックは、ハードウェアリソース（特に、トランシーバ）を共有する必要がなく、したがって、F P G A の同じ箇所に常駐する必要がなく、ルーティングの輻輳を緩和する。

【0077】

ダウンストリームコンポーネントが図 3 に示されている。このコンポーネントでは、ロ

ジックを単純にすることもまた可能である。アップストリームサーバがノードのうちの1つにアドレス指定されたデータを送る際、このデータはクロスポイントスイッチによってFPGA内の2つ以上のフィルタにブロードキャストされる。各フィルタは非常に小型であり必要なものすべてを完備し、また、トランシーバ間でFPGAを横断する長距離にわたるデータのルーティングがない。フィルタに対する受信および送信はいずれも、FPGA上の同じトランシーバに実装可能である。フィルタはクロスポイントから受信したデータに関して動作し、フィルタがデータを通過させる場合にだけ、データがダウンストリームのクライアントに送信（クロスポイントを介して返信）される。

【0078】

すべてのクライアントが、フィルタリングとは無関係に、すべてのデータを受信しなければならないのであれば、図4に示されるように、クロスポイントは、データをFPGAに渡さずに、クライアントにデータを直接ブロードキャストすることが可能である。

【0079】

本発明のさらなる実施形態では、複数のアップストリームのサーバを取り扱うか、または完全クロス接続されているスイッチを作り出すために、多重化機能およびフィルタ機能を組み合わせることが可能である。

【0080】

図5は、本発明の一実施形態の、再構成可能なダウンストリームデータのフィルタリングを提供しているデータ流れ図である。ダウンストリームの場合については、示されているようなデータの流れて、データが、アップストリームのサーバ502から数個のダウンストリームのクライアント512、514に送信される。データは、サーバ502から送信され、外部ポートで受信され、そこでクロスポイントスイッチ520に入る。データは、クロスポイント520によって、FPGA530の、フィルタ536、538について1つのトランシーバ532、534にマルチキャストされる。ポートがすべて異なるフィルタリング要件を有する場合には、示されているように、ダウンストリームのクライアント512、514について1つのフィルタを使用することが可能である。それ以外の場合には、クロスポイントスイッチ520を適切に構成することにより、FPGA530の単一のフィルタから所望する数のクライアントまたはポートにブロードキャスト可能である。図5の実施形態は、2つのフィルタ、すなわち1つのダウンストリームのポートについて1つのフィルタを備える。フィルタ536、538はそれぞれ、自身に関連付けされたトランシーバ532、534を有しており、受信および送信の両方（すなわち、TXが復元クロックにより駆動される）を1つのクロックドメインで実行するか、または複数のクロックドメインでFIFOもしくは他のクロックドメインクロッシングを用いて実行可能である。レイテンシを可能な限り低くするためであれば、前者の方法が好ましいが、実施形態によっては、例えば、完全にデータを再クロックし、送信にとって最良の信号品質を提供するためには、後者の方法を用いる場合がある。要求される信号品質に応じて、選択されるスキームが決まる。図5に、1つのクロックドメインに実装されているフィルタを示す。

【0081】

クロスポイントスイッチ520によって、1つ1つのフィルタ536、538にデータが送信され、そこでそれぞれのトランシーバ532、534によって受信される。トランシーバ532、534はそれぞれ、データを非直列化する。フィルタロジック536、538は、そのデータを解釈し、データが、求められる基準に合致する場合、例えば、そのフィルタに関連付けされたクライアントにデータがアドレス指定される場合には、フィルタ536について示されているように、フィルタは、同じトランシーバを使用してデータを再送信する。このような構成であるため、ロジック536、538は、それぞれのトランシーバ532、534に密接させることが可能である。データがフィルタ基準に合致しない場合、例えば、そのフィルタに関連付けされたクライアントにデータがアドレス指定されない場合には、フィルタ538について示されているように、データは再送信されない。クロスポイントスイッチ520は、管理コンピュータ540により、クライアント5

1 2、5 1 4に関連付けされた適切なポートにフィルタ5 3 6、5 3 8の出力を伝送するように構成されている。管理コンピュータ5 4 0は、フィルタ5 3 6、5 3 8が、クライアント5 1 2、5 1 4に関連付けされたポートに送信されるべきデータを選択するようにもまた、フィルタ5 3 6、5 3 8を構成する。フィルタロジック5 3 6、5 3 8は、求められるフィルタリング基準、および求められる複雑度のレベルに応じて、非常に単純な、また非常に複雑なものとするのが可能である。

【0082】

図6は、本発明の一実施形態の、再構成可能なアップストリームデータのフィルタリングを提供しているデータ流れ図である。アップストリームの場合では、クロスポイント6 2 0が作用して、クライアント6 1 2、6 1 4に関連付けされたポートからFPGA 6 3 0の独立したトランシーバ6 3 2、6 3 4に、データを転送する。トランシーバ6 3 2、6 3 4の送信機能は使用されず、受信機能だけが使用される。トランシーバ6 3 2、6 3 4からの非直列化されたデータは、それぞれのFIFO 6 3 6、6 3 8を通して受信され、FIFO 6 3 6、6 3 8はそれぞれ、クロッシングクロックドメインのために使用される。多重化ロジック6 4 0を用いて、FIFO 6 3 6、6 3 8からデータを受信する。このとき、6 4 2で示されている単一の出力ストリームにデータの packets を多重化することができる。これは、別のトランシーバ6 3 5を介して送られるが、代替的な実施形態では、トランシーバ6 3 5の代わりに受信トランシーバ6 3 2、6 3 4のうちの1つの送信機を使用して、1つのトランシーバを使用せずにすませることも可能であろう。

【0083】

図5および図6では、アップストリームロジックが、FPGAのダウンストリームロジックから物理的に離されているので有利である。これにより、FPGA内でロジックを物理的により小さくすることが可能になることで、パイプラインの段が減り、レイテンシを低くすることが可能になる。

【0084】

図7は、単一のFPGAに2つ以上のマルチプレクサが使用されている、本発明の別の実施形態を図示する。クライアントから複数のマルチプレクサ7 4 0、7 4 2に packets をブロードキャストし、別個のフィルタ7 5 0、7 5 2でそれらのマルチプレクサの出力をフィルタリングすることによって、この実施形態は、複数のアップストリームサーバ7 6 0、7 6 2を取り扱うことが可能である。2つのサーバを有して示されているが、他の実施形態は、類似したやり方でもっと多数のサーバが調達可能であることが理解されよう。したがって、この実施形態は、完全に接続されているスイッチの相互接続性を提供しつつ、本発明の利点もまた提供する。図7では、クライアントデバイス7 1 2、7 1 4を起点とするアップストリーム packets は、クロスポイント7 2 0によって、FPGA 7 3 0に実装された2つの別々の多重化モジュール7 4 0、7 4 2にマルチキャストされる。多重化モジュール7 4 0、7 4 2がそれぞれ、 packets を多重化し、次に、フィルタ7 5 0およびフィルタ7 5 2が、2つの多重化されたストリームをフィルタリングして、それぞれのフィルタの基準に合致しないものを廃棄する。図7では、フィルタ7 5 2が、サーバB 7 6 2を対象としたそれぞれのデータストリーム7 4 3を渡していることが示されているのに対して、フィルタ7 5 0は、データストリーム7 4 1がサーバA 7 6 0を対象としていないので、データストリーム7 4 1を廃棄していることが示されている。図7の実施形態の変形では、多重化ロジック7 4 0、7 4 2の後ではなく前に、4つのフィルタを提供してフィルタリングを実施することにより、帯域幅を大きくすることが可能であろう。さらに、用途によっては、各サーバが、そのサーバにアドレス指定されていないデータを廃棄することが予想される場合には、多重化されたデータをサーバ7 6 0およびサーバ7 6 2の両方に渡すことが容認され得る場合もある。この場合、フィルタ7 5 0およびフィルタ7 5 2は省略することができるので、レイテンシはさらに低減するであろう。

【0085】

図1～図7の実施形態は、このようにクロスポイントスイッチを使用して、アナログドメインのいくつかのポイントにデータをマルチキャストする。同じデータの複数のコピー

を作り出すことによって、このような実施形態はそれぞれ、同じデバイス内で様々なFPGAもしくはASICにわたって、または単一のFPGAもしくはASICの様々な部分にわたって、様々な機能に対するロジックを分散させることができる。

【0086】

図8aは、本発明の別の実施形態に従って完全にクロス接続されている(3x3)スイッチを図示する。示されるように、多重化(mux)およびフィルタ論理機能の拡張は、さらに多くのマルチプレクサおよびフィルタをFPGAに追加することである。これにより、各クライアントから2つ以上のホストにパケットを送ることが可能になる。この実施形態では、各クライアントインタフェースが、固有のマルチプレクサに関連付けされており、これにより、任意の他のポートにパケットを送信可能になる。各ポートに対して専用のフィルタが設けられ、この専用のフィルタを使用して、各ポートによって受信されるパケットが、そのポートにアドレス指定されたパケットだけであるように制限する。このアーキテクチャを用いて、従来のイーサネット(登録商標)スイッチを実装可能である。すなわち、フィルタは、フィルタリングの対象となっているポートに接続されたネットワークに接続されていることが既知であるアドレスにアドレス指定されたパケットだけを転送するように構成されている。

【0087】

図8bは、図8aの実施形態の1つのパケットの取り扱いを図示する。この例では、フィルタは、関連するクライアントにアドレス指定されたパケットだけを受信し、パケットは、示された経路を通るように設定されている。クライアントCは、宛先アドレスが「クライアントA」であるパケットを送信する。クロスポイントスイッチ820は、FPGA830の3つの別々のランシーバにパケットをマルチキャストするように構成されている。パケットは、ランシーバによって受信され、3つの独立した多重化ロジックのセット840、842、844に渡される。多重化論理機能はそれぞれ、受信データを多重化し、次に、それぞれのフィルタ850、852、854にパケットを送信する。多重化ロジック840に対するフィルタ850は、パケットがクライアントAにアドレス指定されているときに、その関連するランシーバにパケットを送信する。クライアントBおよびCに対するフィルタ852およびフィルタ854は、パケットを送信しない。多重化ロジック840に対するランシーバが、クロスポイントスイッチ820にパケットを送信する。クロスポイントスイッチ820が、フィルタ850からクライアントAに受信データをユニキャストするように構成され、パケットが正しい宛先に到達する。リソースのこうした利用は、帯域幅に対して最適化されていない。なぜなら、N個のポートスイッチを実装するには、 N^2 個のランシーバおよびN個のマルチプレクサが必要になるからである。しかしながら、このようなシステムは、共用リソースを有するシステムよりも低レイテンシで、かつ、良好な決定性で構築可能である。

【0088】

図8の実施形態は、本発明に従った、N個のポートの完全にクロス接続されたスイッチの実施態様を図示しているが、代替的な実施形態では、データの取り扱いは、各クライアントに対して、またはクロスポイントスイッチの各ポートに対して同じである必要はない。例えば、いくつかの実施形態では、図8の実施形態を修正することで、1つのポートで受信したデータをタイムスタンプする一方で、その他のポートで受信したデータを多重化できるようにすることが可能であろう。

【0089】

図9は、本発明の1つの実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ900は、902で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ910とを備える。

【0090】

それ以外の時には再構成することが可能であるが、図9に示される構成では、入力デー

タは、データソース 920 から受信され、クロスポイント 910 によって、意図した宛先 930 に関連付けされたそれぞれのデータ出力部 904 へと回路スイッチングされる。この意味では、スイッチ 900 は、物理層スイッチとして機能している。出力部 904 は、入力部 902 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 902 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 902 および出力部 904 は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部 902、および1つのデータ出力部 904 を有効にする。クロスポイント 910 内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図 9 に示されている。

【0091】

本発明に従って、データソース 920 から受信したデータは、宛先 930 に配信されるだけでなく、クロスポイントスイッチ 910 によって、スイッチ 900 内のデバイス 906 によって実行される上位層プロセスにコピーされる。

【0092】

この実施形態では、デバイス 906 がタイムスタンプ機能を備えることにより、デバイス 906 から宛先 930 にデータの第 2 のストリームが配信されるようになっている。このように宛先 930 が、極めて低レイテンシの経路を介して元のデータを受信する一方で、デバイス 906 により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、タイムスタンプされたデータのコピーもまた受信する。

【0093】

図 10 は、本発明の 1 つの実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ 1000 は、1002 で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ 1010 とを備える。

【0094】

それ以外の時には再構成することが可能であるが、図 10 に示される構成では、入力データは、データソース 1020 から受信され、クロスポイント 1010 によって、意図した宛先 1030 に関連付けされたそれぞれのデータ出力部 1004 へと回路スイッチングされる。この意味では、スイッチ 1000 は、物理層スイッチとして機能している。出力部 1004 は、入力部 1002 に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部 1002 を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部 1002 および出力部 1004 は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部 1002、および1つのデータ出力部 1004 を有効にする。クロスポイント 1010 内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図 10 に示されている。

【0095】

本発明に従って、データソース 1020 から受信したデータは、宛先 1030 に配信されるだけでなく、スイッチ 1000 内のデバイス 1006 によって実行され、やはり出力部 1030 へと切り換えられるパケットを生成する上位層プロセスにも配信される。

【0096】

この実施形態では、デバイス 1006 は、マルチキャスト加入機能を備える。これにより、デバイス 1020 および／または 1022 が、マルチキャストのグループに対して加入または退会するために、IGMP パケットが、近くのルータ 1030 に送られたデータストリームに注入されるようになっている。このように宛先 1030 は、極めて低レイテンシの経路を介してデバイス 1020 からデータを受信する一方で、デバイス 1006 により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、指定されたインタフェースに対するいずれかの送信 MAC に注入された IGMP データパケットのような、上位層機能データもまた受信する。

【0097】

図11は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ1100は、1102で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ1110とを備える。

【0098】

それ以外の時には再構成することが可能であるが、図11に示される構成では、入力データは、データソース1120から受信され、クロスポイント1110によって、意図した宛先1130に関連付けされたそれぞれのデータ出力部1104へと回路スイッチングされる。この意味では、スイッチ1100は、物理層スイッチとして機能している。出力部1104は、入力部1102に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部1102を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部1102および出力部1104は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部1102、および1つのデータ出力部1104を有効にする。クロスポイント1110内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図11に示されている。

【0099】

本発明に従って、データソース1120から受信したデータは、宛先1130に配信されるだけでなく、クロスポイントスイッチ1110によって、スイッチ1100内のデバイス1106によって実行される上位層プロセスにコピーされる。

【0100】

この実施形態では、デバイス1106は、管理者1132にデータリンク統計を提供するように構成されたリンクコントローラを備える。この実施形態では、デバイス1106は、複数の上位層機能を提供する。これらに含まれるものとしては、物理層アイドルパターンおよび物理層エラー検出機能、タイプ別、プロトコル別、およびサイズ別のデータパケットカウント機能、無効なFCS、パケットサイズエラーまたはアライメントエラーなどの、エラータイプ別データパケットエラーカウント機能、および、接続されているデバイスを識別するパケットMACアドレス検査機能などがある。

【0101】

このように宛先1130が、極めて低レイテンシの経路を介して元のデータを受信する一方で、デバイス1106により追加のレイテンシが発生するために、レイテンシが大きくなるとはいえ、デバイス1132が、データリンク統計を受信する。

【0102】

図12は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ1200は、1202で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ1210とを備える。

【0103】

それ以外の時には再構成することが可能であるが、図12に示される構成では、入力データは、データソース1220から受信され、クロスポイント1210によって、意図した宛先1230に関連付けされたそれぞれのデータ出力部1204へと回路スイッチングされる。この意味では、スイッチ1200は、物理層スイッチとして機能している。出力部1204は、入力部1202に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部1202を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部1202および出力部1204は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部1202、および1つのデータ出力部1204を有効にする。クロスポイント1

210内で現在作動中のスイッチング経路が、単に例示の目的で、太線によって図12に示されている。

【0104】

本発明に従って、データソース1220から受信したデータは、宛先1230に配信されるだけでなく、クロスポイントスイッチ1210によって、スイッチ1200内のデバイス1206によって実行される上位層プロセスにコピーされる。この実施形態のデバイス1206は、選択された上位層機能にそれぞれが関連付けされた複数の入力部を有する。その結果、コピーされたデータをデバイス1206のそれぞれの入力部にスイッチングするだけで、所望の機能を選択可能になる。さらに、デバイス1206の出力は、クロスポイント1210に戻されることにより、任意の出力ポート1204へと選択的にスイッチングできる。また、図12に示されている時には、デバイス1200は、上位層デバイス1206の出力を宛先1230へと切り換えるように構成されているが、それ以外の時には、出力を他の宛先へと切り換えることができる。管理システムプロセッサ1240は、1秒間隔パルス（PPS）入力を受信して、タイムスタンプングを可能にするとともに、管理インタフェースを介して通信も行っており外部の管理を可能にする。管理システム1240はさらに、クロスポイント1210および上位層デバイス1206を制御し、クロスポイント1210および上位層デバイス1206と通信する。

【0105】

図13は、本発明のさらに別の実施形態を実施するために用いられるハードウェアのシステム概略図である。スイッチ1300は、1302で包括的に示された、他のデバイスに接続可能な複数の前面パネルポートと、1つのポートから複数のポートにデータを再配置するだけでなく、マルチキャストすることもまた可能なクロスポイントスイッチ1310とを備える。

【0106】

それ以外の時には再構成することが可能であるが、図13に示される構成では、入力データは、データソース1320から受信され、クロスポイント1310によって、意図した宛先1330に関連付けされたそれぞれのデータ出力部1304へと回路交換される。この意味では、スイッチ1300は、物理層スイッチとして機能している。出力部1304は、入力部1302に対して離れて示されているが、典型的なラックマウントされたデバイスでは、入力部1302を有するデバイスの前面パネルにこれらとともに配置してもよいことに留意されたい。入力部1302および出力部1304は、複数のトランシーバによって有効にすることができる。複数のトランシーバはそれぞれ、1つのデータ入力部1302、および1つのデータ出力部1304を有効にする。クロスポイント1310内で現在作動中のスイッチング経路が、単に例示の目的で図13に太線で示されている。

【0107】

本発明に従って、データソース1320から受信したデータは、宛先1330に配信されるだけでなく、クロスポイントスイッチ1310によって、スイッチ1300内のデバイス1364によって実行される上位層プロセスにコピーされる。この実施形態のデバイス1364は、図14に示されるタイプの信号調整ブロックである。上位層デバイス1364の出力は、宛先1330に渡されるが、それ以外の時には、出力を他の宛先へと切り換えることができる。管理システムプロセッサ1340は、1秒間隔パルス（PPS）入力を受信して、タイムスタンプングを可能にするとともに、管理インタフェースを介して通信も行っており外部の管理を可能にする。管理システム1340はさらに、クロスポイント1310を制御し、クロスポイント1310と通信して、信号調整ブロック1350～1354および1360～1366のすべてを制御する。

【0108】

図14に示される信号調整ブロック1364では、入力データは、EQブロック1402を通過する。EQブロック1402は、図14全体を通して両矢印によって示されるような管理インタフェース1340を介して設定可能である。EQブロック1402は、受信した信号に関してフィルタリングを実行して、通信チャネル（ケーブル）を通過するこ

との影響を補償することにより、エラーの尤度を減少させる等化器である。EQブロック1402は、信号の周波数成分を調節するとともに、多重路の歪みを除去することもまた可能であり、その出力は、等化されていない受信信号よりもはるかに、データの理想的な表現に近くなっている。EQ1402から、データが、管理インタフェース1340を介して、設定可能なクロックデータ回復(CDR)ブロック1404に渡される。CDR1404から、データが、上位層機能1406およびマルチプレクサ1408の両方に渡される。これらはいずれも、管理システムプロセッサ1340によって制御されている。上位層機能ブロック1406の出力もまた、マルチプレクサ1408に渡される。マルチプレクサ1408から、多重化された低レイテンシのデータおよび上位層機能の出力データが、プリエンファシスブロック1410に渡されることにより、見込まれる通信チャネル(ケーブル)の影響を前もって補償するように送信機信号の周波数成分を修正し、その結果、宛先がデータの理想的な表現により近い信号を受信することで、エラーの尤度が減少するようになる。プリエンファシスブロック1410から、多重化された低レイテンシのデータおよび上位層機能の出力データが、信号調整ブロック1364によって出力される。図13の実施形態の信号調整ブロック1350~1354および1360~1366のいずれかまたはすべてが、図14に示されるように構成することができる。

【0109】

広義に記載されている本発明の趣旨または範囲から逸脱することなく、特定の実施形態で示した本発明に多くの変形および／または修正を行い得ることが、当業者には理解されるであろう。例えば、40Gのようなマルチレーン標準に関連して適用される実施形態では、これらは同様に扱うことが可能である。例えば、4つのトランシーバのグループで40Gに対してフィルタリングを行うことが可能であり、1つまたは複数のクロスポイントデバイスの4つのレーンすべてに対してブロードキャストを行うことが可能である。したがって、本実施形態は、あらゆる点で例示として見なされるべきであり、限定または制限するものと見なされるべきではない。

【手続補正2】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

データスイッチング方法であって、

クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信し、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストし、

前記クロスポイントスイッチの前記複数の出力ポートの各々を、論理機能デバイスの対応する入力部に接続し、

前記論理機能デバイスが前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を備える、データスイッチング方法。

【請求項2】

前記論理機能デバイスはフィールドプログラマブルゲートアレイ(FPGA)を備える、請求項1に記載のデータスイッチング方法。

【請求項3】

前記論理機能デバイスの前記複数の出力インタフェースの各々は前記クロスポイントスイッチの入力部に接続されている、請求項1または2に記載のデータスイッチング方法。

【請求項4】

前記論理機能デバイスによって適用された前記論理機能は、内容ベースフィルタリングを

備える、請求項 1～3 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 5】

前記内容ベースフィルタリングはアドレスフィルタリングであり、そのアドレスフィルタリングによって、複数のクライアントデバイスは前記論理機能デバイスの複数の出力ポートにそれぞれ関連付けされており、このことによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけが前記論理機能デバイスの対応する出力ポートから出力されるように前記論理機能デバイスは前記クロスポイントスイッチの前記複数の出力ポートの各々からのデータをフィルタする、請求項 4 に記載のデータスイッチング方法。

【請求項 6】

前記論理機能は多重化を含む、請求項 1～5 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 7】

前記論理機能は、前記論理機能デバイスのトランシーバ毎に専用のルーティングテーブルを提供することによって有効化されるレイヤ 3 ルーティングを含む、請求項 1～6 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 8】

管理デバイスを使用して、高速遠隔再構成を有効にするように前記論理機能デバイスの複数の論理ブロックを再プログラムすることをさらに備える、請求項 1～7 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 9】

前記クロスポイントスイッチが重複データストリームをデータパケット捕捉 RAM、ディスクおよび／または P C I e リンクに配信することによって、複数のデータパケットを捕捉することをさらに備える、請求項 1～8 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 10】

データスイッチであって、

データを受信するための少なくとも 1 つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを前記複数の出力ポートに設定可能にマルチキャストするように構成されているクロスポイントスイッチと、

前記クロスポイントスイッチの前記複数の出力ポートの各々にそれぞれが接続された複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも 1 つの論理機能を有する、前記論理機能デバイスとを備えるデータスイッチ。

【請求項 11】

データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は 1 つまたは複数のプロセッサによって実行される場合に、

クロスポイントスイッチの少なくとも 1 つの入力ポートでデータを受信すること、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも 1 つの出力ポートに設定可能にマルチキャストすること、

前記クロスポイントスイッチの前記複数の出力ポートの各々を論理機能デバイスの対応する入力部に接続すること、

前記論理機能デバイスが、前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を実行させる、非一時的コンピュータ読み取り可能媒体。

【請求項 12】

データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、

どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの1つを選択するためのコンピュータプログラムコード手段と、

クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを、当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記複数の出力ポートの各々が論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、

前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段とを備えるコンピュータプログラム製品。

【請求項13】

物理層データスイッチング方法であって、

スイッチの少なくとも1つの入力ポートでデータを受信し、

前記データを前記スイッチの少なくとも1つの出力ポートへと回路交換し、

上位層プロセスに前記データをコピーすること

を備える、物理層データスイッチング方法。

【請求項14】

物理層データスイッチであって、

複数のデータ出力部のうちの1つまたは複数へと設定可能に回路交換される複数のデータ入力部であって、第1のデータ出力部が第1の入力部から回路交換されたデータを受信し、第2のデータ出力部が前記第1の入力部から前記データのコピーを受信する、前記複数のデータ入力部と、

前記第2のデータ出力部に接続された上位層機能回路とを備える物理層データスイッチ。

【請求項15】

前記上位層機能回路はパケットタイムスタンプングを実行するように構成されている、請求項14に記載の物理層データスイッチ。

【請求項16】

前記上位層機能回路はパケット捕捉を実行するように構成されている、請求項14または15に記載の物理層データスイッチ。

【請求項17】

前記上位層機能回路は前記第1のデータ出力部においてレイテンシに影響を及ぼさずに、データパケット統計のカウントを実行するように構成されている、請求項14または15に記載の物理層データスイッチ。

【手続補正書】

【提出日】平成28年12月21日(2016.12.21)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

データスイッチング方法であって、

クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信し、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの複数の出力

ポートに設定可能にマルチキャストし、

前記クロスポイントスイッチの前記複数の出力ポートの各々を、論理機能デバイスの対応する入力部に接続し、

前記論理機能デバイスが前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を備える、データスイッチング方法。

【請求項 2】

前記論理機能デバイスはフィールドプログラマブルゲートアレイ（FPGA）を備える、請求項 1 に記載のデータスイッチング方法。

【請求項 3】

前記論理機能デバイスの前記複数の出力インタフェースの各々は前記クロスポイントスイッチの入力部に接続されている、請求項 1 または 2 に記載のデータスイッチング方法。

【請求項 4】

前記論理機能デバイスによって適用された前記論理機能は、内容ベースフィルタリングを備える、請求項 1～3 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 5】

前記内容ベースフィルタリングはアドレスフィルタリングであり、そのアドレスフィルタリングによって、複数のクライアントデバイスは前記論理機能デバイスの複数の出力ポートにそれぞれ関連付けされており、このことによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけが前記論理機能デバイスの対応する出力ポートから出力されるように前記論理機能デバイスは前記クロスポイントスイッチの前記複数の出力ポートの各々からのデータをフィルタする、請求項 4 に記載のデータスイッチング方法。

【請求項 6】

前記論理機能は多重化を含む、請求項 1～5 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 7】

前記論理機能は、前記論理機能デバイスのトランシーバ毎に専用のルーティングテーブルを提供することによって有効化されるレイヤ 3 ルーティングを含む、請求項 1～6 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 8】

管理デバイスを使用して、高速遠隔再構成を有効にするように前記論理機能デバイスの複数の論理ブロックを再プログラムすることをさらに備える、請求項 1～7 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 9】

前記クロスポイントスイッチが重複データストリームをデータパケット捕捉 RAM、ディスクおよび／または PCI e リンクに配信することによって、複数のデータパケットを捕捉することをさらに備える、請求項 1～8 のいずれか 1 項に記載のデータスイッチング方法。

【請求項 10】

データスイッチであって、

データを受信するための少なくとも 1 つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを前記複数の出力ポートに設定可能にマルチキャストするように構成されているクロスポイントスイッチと、

前記クロスポイントスイッチの前記複数の出力ポートの各々にそれぞれが接続された複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも 1 つの論理機能を有する、前記論理機能デバイスと

を備えるデータスイッチ。

【請求項 1 1】

データスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は 1 つまたは複数のプロセッサによって実行される場合に、

クロスポイントスイッチの少なくとも 1 つの入力ポートでデータを受信すること、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも 1 つの出力ポートに設定可能にマルチキャストすること、

前記クロスポイントスイッチの前記複数の出力ポートの各々を論理機能デバイスの対応する入力部に接続すること、

前記論理機能デバイスが、前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を実行させる、非一時的コンピュータ読み取り可能媒体。

【請求項 1 2】

データスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、

どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの 1 つを選択するためのコンピュータプログラムコード手段と、

クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを、当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記複数の出力ポートの各々が論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、

前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段と

を備えるコンピュータプログラム製品。

【手続補正書】

【提出日】平成29年5月1日(2017.5.1)

【手続補正 1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項 1】

コンピュータネットワーク用のデータスイッチング方法であって、

クロスポイントスイッチの少なくとも 1 つの入力ポートでデータを受信し、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストし、

前記クロスポイントスイッチの前記複数の出力ポートの各々を、論理機能デバイスの対応する入力部に接続し、

前記論理機能デバイスが前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを、当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を備える、データスイッチング方法。

【請求項 2】

前記論理機能デバイスはフィールドプログラマブルゲートアレイ（FPGA）を備える、請求項1に記載のデータスイッチング方法。

【請求項3】

前記論理機能デバイスの前記複数の出力インタフェースの各々は前記クロスポイントスイッチの入力部に接続されている、請求項1または2に記載のデータスイッチング方法。

【請求項4】

前記論理機能デバイスによって適用された前記論理機能は、内容ベースフィルタリングを備える、請求項1～3のいずれか1項に記載のデータスイッチング方法。

【請求項5】

前記内容ベースフィルタリングはアドレスフィルタリングであり、そのアドレスフィルタリングによって、複数のクライアントデバイスは前記論理機能デバイスの複数の出力ポートにそれぞれ関連付けされており、このことによって、特定のローカルクライアントデバイスにアドレス指定されたデータだけが前記論理機能デバイスの対応する出力ポートから出力されるように前記論理機能デバイスは前記クロスポイントスイッチの前記複数の出力ポートの各々からのデータをフィルタする、請求項4に記載のデータスイッチング方法。

【請求項6】

前記論理機能は多重化を含む、請求項1～5のいずれか1項に記載のデータスイッチング方法。

【請求項7】

前記論理機能は、前記論理機能デバイスのトランシーバ毎に専用のルーティングテーブルを提供することによって有効化されるレイヤ3ルーティングを含む、請求項1～6のいずれか1項に記載のデータスイッチング方法。

【請求項8】

管理デバイスを使用して、高速遠隔再構成を有効にするように前記論理機能デバイスの複数の論理ブロックを再プログラムすることをさらに備える、請求項1～7のいずれか1項に記載のデータスイッチング方法。

【請求項9】

前記クロスポイントスイッチが重複データストリームをデータパケット捕捉RAM、ディスクおよび／またはPCIeリンクに配信することによって、複数のデータパケットを捕捉することをさらに備える、請求項1～8のいずれか1項に記載のデータスイッチング方法。

【請求項10】

コンピュータネットワーク用のデータスイッチであって、

データを受信するための少なくとも1つの入力ポートおよび複数の出力ポートを有するクロスポイントスイッチであって、クロスポイントスイッチの少なくとも1つの入力ポートで受信したデータを前記複数の出力ポートに設定可能にマルチキャストするように構成されているクロスポイントスイッチと、

前記クロスポイントスイッチの前記複数の出力ポートの各々にそれぞれが接続された複数の入力部を有する論理機能デバイスであって、前記クロスポイントスイッチの前記複数の出力ポートから受信したデータに適用されて処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力するように構成される少なくとも1つの論理機能を有する、前記論理機能デバイスと
を備えるデータスイッチ。

【請求項11】

コンピュータネットワーク用のデータスイッチを構成するための非一時的コンピュータ読み取り可能媒体であって、当該非一時的コンピュータ読み取り可能媒体は命令を備え、前記命令は1つまたは複数のプロセッサによって実行される場合に、

クロスポイントスイッチの少なくとも1つの入力ポートでデータを受信すること、

前記クロスポイントスイッチが前記データを当該クロスポイントスイッチの少なくとも1つの出力ポートに設定可能にマルチキャストすること、

前記クロスポイントスイッチの前記複数の出力ポートの各々を論理機能デバイスの対応する入力部に接続すること、

前記論理機能デバイスが、前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに論理機能を適用し、処理後のデータを当該論理機能デバイスの複数の対応する出力インタフェースに出力すること

を実行させる、非一時的コンピュータ読み取り可能媒体。

【請求項 12】

コンピュータネットワーク用のデータスイッチのソフトウェアコントローラ用の手続きをコンピュータに実行させるコンピュータプログラムコード手段を備えるコンピュータプログラム製品であって、

どのデータが処理されるのかに関する、複数の論理機能デバイス機能のうちの 1 つを選択するためのコンピュータプログラムコード手段と、

クロスポイントスイッチの少なくとも 1 つの入力ポートで受信したデータを、当該クロスポイントスイッチの複数の出力ポートに設定可能にマルチキャストするように、クロスポイントスイッチを制御するためのコンピュータプログラムコード手段であって、前記クロスポイントスイッチの前記複数の出力ポートの各々が論理機能デバイスの対応する入力部に接続されている、コンピュータプログラムコード手段と、

前記クロスポイントスイッチの前記複数の出力ポートの各々から受信したデータに前記論理機能デバイスが前記選択された論理機能を適用し、処理後のデータを前記論理機能デバイスの複数の対応する出力インタフェースに出力するように、前記論理機能デバイスを制御するためのコンピュータプログラムコード手段と

を備えるコンピュータプログラム製品。

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/AU2014/050391
A. CLASSIFICATION OF SUBJECT MATTER H04L 12/20 (2006.01) H04L 12/801 (2013.01) H04L 12/861 (2013.01) H04L 12/875 (2013.01) H04L 12/701 (2013.01) H04L 12/747 (2013.01) G06F 15/163 (2006.01) G06F 15/17 (2006.01) G06F 13/00 (2006.01)		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPI, EPODOC: Computer, Data, Switch, Crosspoint, FPGA, Latency, Multicast and similar terms.		
Google Patents: Crosspoint, Latency, Port, Reconfigurable, Routing, Clos, Multicast, Physical Layer, FPGA, Packet, Capture, Memory and similar terms.		
Google Scholar: Low-latency, parallel, reconfigurable, crossbar architecture, Hartenstein, xputer and similar terms.		
The Lens: Crosspoint, FPGA, Filter, Multicast, Timestamp, Physical Layer and similar terms.		
EspaceNet and Google Patents: Applicant/Inventor Search.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	
☒ Further documents are listed in the continuation of Box C ☒ See patent family annex		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "E" earlier application or patent but published on or after the international filing date "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "O" document referring to an oral disclosure, use, exhibition or other means "&" document member of the same patent family "P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 29 January 2015		Date of mailing of the international search report 29 January 2015
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustalia.gov.au		Authorised officer Marthinus Van Der Westhuizen AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. 0262832283

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/AU2014/050391
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2002/0159468 A1 (FOSTER et al.) 31 October 2002 Abstract; Figs. 1, 3, 14; Paragraphs [0005], [0029-0034], [0038-0039], [0042-0044], [0049], [0051-0053], [0066], [0076]	1-10 and 12-16
Y	Whole of document	11 and 17-19
X	US 2007/0200594 A1 (LEVI et al.) 30 August 2007 Title, Abstract; Figs. 3, 6, 8-11; Paragraphs [0011], [0034-0035], [0045]	1-5, 8-10 and 12-16
Y	Whole of document	11 and 17-19
Y	US 2013/0343390 A1 (MORIARTY et al.) 26 December 2013 Title, Abstract; Paragraph [0124]	11 and 17-19
A	US 2011/0264578 A1 (CHAPMAN et al.) 27 October 2011 Whole of document	1-19
A	HARTENSTEIN, R., "A decade of reconfigurable computing: a visionary retrospective", 2001, DATE '01 Proceedings of the conference on Design, Automation and Test in Europe, IEEE Press Piscataway, NJ USA, ISBN 0-7695-0993-2, Pages 642-649. Whole of document	1-19
Form PCT/ISA/210 (fifth sheet) (July 2009)		

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/AU2014/050391	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
US 2002/0159468 A1	31 October 2002	AU 2002258931 A1	11 Nov 2002
		AU 2002303416 A1	11 Nov 2002
		AU 2002305261 A1	11 Nov 2002
		AU 2002307418 A1	11 Nov 2002
		EP 1388075 A1	11 Feb 2004
		EP 1391082 A1	25 Feb 2004
		JP 2004537881 A	16 Dec 2004
		JP 2005502228 A	20 Jan 2005
		JP 2007166666 A	28 Jun 2007
		US 2003202535 A1	30 Oct 2003
		US 6993023 B2	31 Jan 2006
		US 2002159446 A1	31 Oct 2002
		US 6996058 B2	07 Feb 2006
		US 2003202536 A1	30 Oct 2003
		US 7042877 B2	09 May 2006
		US 2002159452 A1	31 Oct 2002
		US 7068666 B2	27 Jun 2006
		US 2002159451 A1	31 Oct 2002
		US 7068667 B2	27 Jun 2006
		US 2002181395 A1	05 Dec 2002
		US 7164656 B2	16 Jan 2007
		US 2003065859 A1	03 Apr 2003
		US 7602818 B2	13 Oct 2009
		US 2002159389 A1	31 Oct 2002
		US 2002159437 A1	31 Oct 2002
		US 2002159453 A1	31 Oct 2002
		US 2002159456 A1	31 Oct 2002
		US 2002159458 A1	31 Oct 2002
		US 2002161887 A1	31 Oct 2002
		US 2002161923 A1	31 Oct 2002
		US 2002167902 A1	14 Nov 2002
		US 2002184529 A1	05 Dec 2002
		US 2002188754 A1	12 Dec 2002
		US 2003189927 A1	09 Oct 2003
		US 2003204618 A1	30 Oct 2003
		US 2003210685 A1	13 Nov 2003
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
Form PCT/ISA/210 (Family Annex) July 2009			

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/AU2014/050391	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
		US 2004004966 A1	08 Jan 2004
		WO 02088875 A2	07 Nov 2002
		WO 02088876 A2	07 Nov 2002
		WO 02088981 A1	07 Nov 2002
		WO 02089384 A2	07 Nov 2002
		WO 02089405 A2	07 Nov 2002
		WO 02089418 A1	07 Nov 2002
		WO 02089418 B1	30 Jan 2003
US 2007/0200594 A1	30 August 2007	US 7482836 B2	27 Jan 2009
		US 7224184 B1	29 May 2007
US 2013/0343390 A1	26 December 2013	US 8891528 B2	18 Nov 2014
US 2011/0264578 A1	27 October 2011	US 8706606 B2	22 Apr 2014
		AU 2009311242 A1	23 Jun 2011
		CA 2743221 A1	14 May 2010
		CN 102246194 A	16 Nov 2011
		CN 102246194 B	21 May 2014
		GB 2478093 A	24 Aug 2011
		GB 2478093 B	24 Jul 2013
		HK 1161762 A1	13 Dec 2013
		JP 2012508409 A	05 Apr 2012
		KR 20110101148 A	15 Sep 2011
		WO 2010051575 A1	14 May 2010
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex) July 2009			

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, RU, TJ, TM), EP(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US

(72)発明者 スノードン、デイビッド
オーストラリア国 2010 ニューサウスウェールズ州 ダーリングハースト コールドウェル
ストリート 34

(72)発明者 マクデイド、スコット
オーストラリア国 2107 ニューサウスウェールズ州 ビルゴラ プラトー レイモンド ロ
ード 47

(72)発明者 トーマス、チャールズ
オーストラリア国 2120 ニューサウスウェールズ州 ワラウィー ブライスウッド アベニ
ュー 51

(72)発明者 ボウルズ、リンジー
オーストラリア国 2153 ニューサウスウェールズ州 ボウカム ヒルズ ボーガン アベニ
ュー 14

Fターム(参考) 5K030 GA02 KX09 KX17 KX24 LD02