

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号

特許第7184881号

(P7184881)

(45)発行日 令和4年12月6日(2022.12.6)

(24)登録日 令和4年11月28日(2022.11.28)

(51)国際特許分類

F I

H 0 4 L 49/9057(2022.01)

H 0 4 L 49/9057

H 0 4 L 47/43 (2022.01)

H 0 4 L 47/43

請求項の数 21 (全41頁)

(21)出願番号	特願2020-517763(P2020-517763)	(73)特許権者	519435991
(86)(22)出願日	平成30年6月7日(2018.6.7)		ハイアニス・ポート・リサーチ・インコ
(65)公表番号	特表2020-523950(P2020-523950 A)		ーポレーテッド
(43)公表日	令和2年8月6日(2020.8.6)		HYANNIS PORT RESEAR
(86)国際出願番号	PCT/US2018/036395		CH, INC.
(87)国際公開番号	WO2018/226919		アメリカ合衆国, マサチューセッツ州
(87)国際公開日	平成30年12月13日(2018.12.13)		0 2 4 9 4, ニーダム, スイート 2 1
審査請求日	令和3年5月26日(2021.5.26)		0, セカンド アヴェニュー 7 5
(31)優先権主張番号	62/516,753	(74)代理人	100087941
(32)優先日	平成29年6月8日(2017.6.8)		弁理士 杉本 修司
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100112829
			弁理士 堤 健郎
		(74)代理人	100142608
			弁理士 小林 由佳
		(74)代理人	100154771

最終頁に続く

(54)【発明の名称】 変更通知ありのTCPストリーム動的処理

(57)【特許請求の範囲】

【請求項1】

第1のエンドポイントと第2のエンドポイントとの間に確立された双方向ネットワークセッション中に、

パケットのストリームを検査する過程と、

前記第1のエンドポイントから送信された所与のパケット内のメッセージが変更されるべきであると判断すると、

前記所与のパケット内の前記メッセージを変更して変更後パケットを生成する過程と、

前記メッセージの変更を考慮して変更後のシーケンス番号を決定する過程と、

前記変更後パケットを前記第2のエンドポイントに送信する過程と、

返信メッセージを前記第1のエンドポイントに送信し、元のメッセージが変更されたことを知らせる過程と、

を備え、

変更後のシーケンス番号を決定する前記過程が、

アキュムレータ値を維持すること、

元の前記所与のパケットと前記変更後パケットとのバイト数の差分を表すデルタ値を記憶すること、

前記第2のエンドポイントからの前記変更後パケットの確認応答を検出すること、および

前記デルタ値で前記アキュムレータ値を更新すること、

を含む、方法。

【請求項 2】

請求項 1 に記載の方法において、さらに、
変更された前記パケットのシーケンス番号に基づいてシーケンスマーカー値を決定すること、

受信した後続のパケットのシーケンス番号が前記シーケンスマーカー値以上である場合、
前記第 1 のエンドポイントから受信した後続のパケットのシーケンス番号を前記アキュムレータ値だけ変更する過程、
を備える、方法。

【請求項 3】

請求項 2 に記載の方法において、さらに、
前記第 2 のエンドポイントから受信した別のパケットの確認応答シーケンス番号を前記アキュムレータ値に基づいて調節する過程、
を備える、方法。

【請求項 4】

請求項 1 に記載の方法において、前記所与のパケット内の前記メッセージを変更する過程が、当該所与のパケット内の当該メッセージを、異なる長さの第 2 のメッセージに置き換えることを含む、方法。

【請求項 5】

請求項 4 に記載の方法において、前記第 2 のメッセージが、ハートビートメッセージ、テストメッセージ、またはシーケンスフィルタメッセージを含む、方法。

【請求項 6】

請求項 1 に記載の方法において、前記所与のパケット内の前記メッセージを変更する過程が、当該パケットから当該メッセージを除去することを含む、方法。

【請求項 7】

請求項 1 に記載の方法において、前記第 1 のエンドポイントが取引業者に対応付けられたクライアント装置であり、前記第 2 のエンドポイントが証券取引システムに対応付けられた市場装置であり、前記メッセージが証券を売買する注文である、方法。

【請求項 8】

請求項 7 に記載の方法において、前記返信メッセージが、前記メッセージを変更した理由を含む、方法。

【請求項 9】

請求項 7 に記載の方法において、前記メッセージを変更した理由が、取引ルールへの準拠、セキュリティリスク又は証券リスク、及びパケットコンテンツフィルタリングのうちの少なくとも一つに関する、方法。

【請求項 10】

請求項 1 に記載の方法において、さらに、
前記第 2 のエンドポイントから前記変更後パケットの A C K を受信したか否かを判断する過程と、

前記変更後パケットの A C K を受信していないと判断した場合、

前記変更後パケットを再送する過程と、

タイムアウト時間後に前記セッションを切断する過程と、

を備える、方法。

【請求項 11】

請求項 10 に記載の方法において、前記変更後パケットの A C K を受信したか否かを判断する過程が、一定間隔で実行される、方法。

【請求項 12】

請求項 10 に記載の方法において、さらに、
前記第 2 のエンドポイントから受信した前記 A C K のシーケンス番号が前記シーケンスマーカー以上である場合、
変更ペンディング値をクリアする過程、

を備える、方法。

【請求項 13】

請求項 10 に記載の方法において、変更された前記メッセージの A C K を受信したか否かを判断する過程が、前記デルタ値を定期的にポーリングすることを含む、方法。

【請求項 14】

請求項 1 に記載の方法において、さらに、

前記変更後パケットの A C K を未だ受信しておらず、かつ、後続のパケットについても変更して第 2 の変更後パケットを生成する必要がある場合、

前記変更後パケットの前記 A C K を受信するまで前記第 2 の変更後パケットをストールする過程、

を備える、方法。

【請求項 15】

請求項 14 に記載の方法において、さらに、

前記変更後パケットの前記 A C K をいつ受信したかを判断する過程と、

所定の時間間隔で前記変更後パケットを再送する過程と、

前記変更後パケットの前記 A C K を受信することなく所定の時間が経過すると、前記セッションを切断する過程と、

を備える、方法。

【請求項 16】

第 1 の T C P エンドポイントと第 2 の T C P エンドポイントとの間に確立された双方向 T C P セッション中に、

前記第 1 の T C P エンドポイントから受信した T C P パケットを検査する過程と、

所与の T C P パケットから T C P シーケンス番号を抽出する過程と、

前記所与の T C P パケットに対応する T C P セッションを特定する過程と、

前記所与の T C P パケット内のコンテンツがルールに準拠しているか否かを判断する過程と、

前記所与の T C P パケットが前記ルールに準拠している場合、

前記所与の T C P パケットを前記第 2 の T C P エンドポイントに到達させる過程と、

前記所与の T C P パケットが前記ルールに準拠していない場合、

拒絶理由データ値を決定する過程と、

前記所与の T C P パケットの少なくとも一部が前記第 2 の T C P エンドポイントに到達するのを阻止する過程と、

前記所与の T C P パケットの少なくとも一部を除去して変更後 T C P パケットを形成する過程と、

デルタ値を、そのようにして除去された前記一部のサイズに依存する数値に設定する過程と、

前記デルタ値を、前記 T C P セッションに対応付けられた各セッション用の T C P アキュムレータに適用する過程と、

前記変更後 T C P パケットの T C P シーケンス番号を、前記各セッション用の T C P アキュムレータのサイズに依存する数値だけ調節する過程と、

前記変更後 T C P パケットを前記第 2 の T C P エンドポイントに送信する過程と、

前記拒絶の理由を表す数値を有する返信パケットを生成する過程と、

前記返信パケットを前記第 1 の T C P エンドポイントに送信する過程と、

を備える、方法。

【請求項 17】

互いに独立して動作する第 1 のネットワークインターフェース及び第 2 のネットワークインターフェースを有するカットスルー方式の固定ロジックデバイスであって、1 つ又は複数の T C P パケットストリームが前記第 1 および第 2 のネットワークインターフェース間を移動することのできる、カットスルー方式の固定ロジックデバイスと、

前記 T C P パケットストリームを検査するパケット検査デバイスであって、

10

20

30

40

50

所与のTCPパケットからTCPシーケンス番号を抽出し、
前記所与のTCPパケットに対応するセッションを特定し、
前記所与のTCPパケットが、ルールに準拠していないコンテンツを含んでいるか否かを判断し、

前記所与のTCPパケットが準拠していない場合、

前記所与のTCPパケットの少なくとも一部が前記第2のネットワークインターフェースに到達するのを阻止する、

パケット検査デバイスと、

固定ロジック及び少なくとも1つのプログラマブルなデータプロセッサを含むコントローラであって、

前記パケット検査デバイスから前記所与のTCPパケット及び拒絶理由データ値を受け取り、

前記所与のTCPパケットの前記シーケンス番号に基づく数値を、各セッション用のTCPシーケンスマーカーとして記憶し、

前記所与のTCPパケットから、前記ルールに準拠していない前記コンテンツをトリミングして、変更後TCPパケットを形成し、

前記変更後TCPパケットを前記第2のネットワークインターフェースを介して送信できるように、当該変更後TCPパケットを、前記カットスルー方式の固定ロジックデバイスと共有するメモリに書き込み、

各セッション用のデルタ値を、トリミングされた前記コンテンツのサイズに依存する数値に設定し、

前記変更後TCPパケットの前記TCPシーケンス番号を、各セッション用のTCPアキュムレータの数値に従って調節し、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットの、ACKシーケンス番号及びセッション識別子を抽出し、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットのACKシーケンス番号が、前記セッション識別子に対応するTCPシーケンスマーカー以上である場合、前記セッション識別子に対応するデルタ値を、当該セッション識別子に対応付けられた前記TCPアキュムレータに適用し、

前記第2のネットワークインターフェースを介して届いた前記後続のTCPパケットの前記ACKシーケンス番号を、前記セッション識別子に対応付けられた前記TCPアキュムレータの数値だけ調節し、

前記拒絶の理由を表す数値を含む、前記第1のネットワークインターフェースを介して送信する返信パケットを生成する、

コントローラと、

を備える、装置。

【請求項18】

請求項17に記載の装置において、前記コントローラが、さらに、

ルールに準拠していない前記所与のパケットに対応する前記セッションについて、コネクションストールモードを設定し、

前記第2のネットワークインターフェースから受信した後続のパケットが、前記コネクションストールモードを設定した前記セッションに属するものである場合、

生成された前記返信パケットを前記第1のネットワークインターフェースに転送し、

前記生成された返信パケットを前記第1のネットワークインターフェースに転送した後、前記後続のパケットを前記第1のネットワークインターフェースに転送し、

前記セッションの前記コネクションストールモードをクリアする、装置。

【請求項19】

請求項17に記載の装置において、前記コントローラの前記固定ロジックが、

前記パケット検査デバイスから前記所与のTCPパケット及び拒絶理由データ値を受け取る手順、

10

20

30

40

50

前記所与のTCPパケットの前記シーケンス番号に基づく数値を、各セッション用のTCPシーケンスマーカーとして記憶する手順、

前記変更後TCPパケットの前記TCPシーケンス番号を、各セッション用のTCPアキュムレータの数値に従って調節する手順、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットの、ACKシーケンス番号及びセッション識別子を抽出する手順、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットのACKシーケンス番号が、前記セッション識別子に対応するTCPシーケンスマーカー以上である場合、前記セッション識別子に対応するデルタ値を、当該セッション識別子に対応付けられた前記TCPアキュムレータに適用する手順、ならびに

前記第2のネットワークインターフェースを介して届いた前記後続のTCPパケットの前記ACKシーケンス番号を、前記セッション識別子に対応付けられた前記TCPアキュムレータの数値だけ調節する手順、

のうちの少なくとも一つを実行する、装置。

【請求項20】

請求項19に記載の装置において、前記プログラマブルなデータプロセッサが、プログラムコードを実行して、

前記所与のTCPパケットから、前記ルールに準拠していない前記コンテンツをトリミングして、変更後TCPパケットを形成する手順、

前記変更後TCPパケットを前記第2のネットワークインターフェースを介して送信できるように、前記変更後TCPパケットを、前記カットスルー方式の固定ロジックデバイスと共有するメモリに書き込む手順、

各セッション用のデルタ値を、トリミングされた前記コンテンツのサイズに依存する数値に設定する手順、ならびに

前記拒絶の理由を表す数値を含む、前記第1のネットワークインターフェースを介して送信する返信パケットを生成する手順、

のうちの少なくとも一つを実行する、装置。

【請求項21】

請求項17に記載の装置において、前記コントローラが、さらに、

ルールに準拠していない前記所与のTCPパケットに対応する前記セッションについて、阻止モードインジケータ値を設定し、

前記第1のネットワークインターフェースを介して届いた後続のパケットに関連付けられている情報を、処理キューに格納し、

前記処理キュー内の、前記パケットに関連付けられている前記情報が処理されたか否かを判断し、

前記処理キュー内の、前記パケットに関連付けられている前記情報が処理されたと判断すると、前記阻止モードインジケータをクリアする、装置。

【発明の詳細な説明】

【関連出願】

【0001】

本願は、2017年6月8日出願の同時係属中の米国仮特許出願第62/516,753号である“Dynamic TCP Stream Processing (TCPストリーム動的処理)”の優先権を主張する。上記出願の全内容は、参照をもって本明細書に取り入れたものとする。

【技術分野】

【0002】

本特許出願は、ネットワーク接続された装置、詳細には、インラインコンテンツ検査に関する。

【背景技術】

【0003】

コンテンツ検査は、ネットワーク化されたデータ処理システムにより様々な理由で頻繁

10

20

30

40

50

に用いられる技術である。データパケットが検査ポイントを通過する際に当該データパケットが検査されて、ウイルス、スパム、機微な又は敏感な又は影響の受けやすい (sensitive) データ、キーワード、またはコンテンツレベルの他の判断基準事項 (criteria) が探し出される。コンテンツ検査は、個々のパケットのヘッダやその他のルーティング挙動に注目するのではなく、実際のメッセージペイロードが何を含んでいるのかを見る。コンテンツ検査は、ネットワーク通過中のデータパケットをキャプチャ (採取) し、コンテンツが敏感な内容 (sensitivity) であるか否かを分析することで機能する。これは、既知のデータ構造 (例えば、クレジットカード番号に使われるパターン等) や、敏感な内容 (sensitivity) を示唆するキーワード (例えば、「秘 (confidential)」等) や、スパムやウイルスのシグネチャを特定することを伴い得る。コンテンツ検査は、データを分類又は類別するのに利用されることもでき、予め設定されたルールを適用して、ペイメントカードインダストリーデータ (PCI) や、個人を特定できる情報 (PII) や、保護された健康情報 (PHI) や、証券取引ルールや、その他の基準に準拠しているか否かを確認又は保証する (ensure) ことが可能である。

【0004】

特許文献1 (Amicangioli) に記載されたアプローチは、高頻度電子証券取引ネットワークにおいて極めて有用な、カットスルー方式の透過的なロジックデバイスとリアルタイムパケットプロセッサとを用いたインラインネットワーク・トラフィックキャプチャ・システムである。このシステムは、第1のインターフェースを介して (証券会社、取引業者や仲介業者、又はその他の顧客により操作され得る) 1つ以上のクライアントマシンから、取引要求を含むメッセージを受信する。取引情報は、1つ以上の取引リスクルールに準拠しているか否かを判断するために検査された後、第2のインターフェースを介して (証券取引所により操作され得る) 1つ以上の市場サーバへと送信される。これらのインターフェースは、カットスルー方式の固定ロジックを介して互いに接続されており、かつ、互いに独立して動作する。このカットスルー方式の固定ロジックは、第1のポートで受信した取引情報を第2のポートへとそのまま (directly) 転送する2ポートデバイスであり、クライアントマシン及び市場サーバのどちらとも通信コネクションを切断しないまま、クライアントから取引に関する情報を全て受け取り終えるのを待たずに、その取引情報の一部の送信を市場サーバに対して開始することができる。取引がルールに違反している場合には、取引情報メッセージの全体が市場サーバに送り届けられる前に、当該メッセージに対して修正又は変更が施される。

【先行技術文献】

【特許文献】

【0005】

【文献】米国特許第9607307号明細書

【発明の概要】

【発明が解決しようとする課題】

【0006】

(従来 of 解決手段の問題点)

既存のコンテンツ検査技術は、一部の用途において問題となる。一例として、多数のクライアント装置が市場装置に対して証券売買 (取引) 注文を送信する高頻度証券取引ネットワークを取り上げたい。売買注文は、メッセージとして形成され、パケット内で他のメッセージと合体させられたものであり得る。この環境では、多数の (a number of) ルールに準拠していない売買注文については拒絶されなければならない、つまり、クライアントから市場へと (C2M) 移動できないようにされなければならない。

【0007】

この環境でのコンテンツ検査の実装又は実現には、典型的に、「One and Done」と「Gap Fill Overwrite」の2種類のメッセージ拒絶方法が用いられてきた。これらの方法のいずれも、市場コンプライアンスに関してリスクがあったり、自身の注文の状態を知りたいクライアントにとって分かり難かったり、取引セッションが

らの完全な切断が必要になったりするという点で不十分である。

【0008】

例えば、One and Doneは、クライアント装置をセッションから強制的に切断させるところ、これは、高速・低レイテンシ取引システムの世界において非理想的な戦略である。また、One and Doneにより切断が引き起こされると、注文の喪失や、クライアント側アルゴリズムの混乱や、再接続の困難性も招かれ得る。

【0009】

Gap Fill Overwriteは、「Z V Z Z T」のような検査記号を用いてメッセージを上書きするもので、これは現在、様々な取引所でますます規制されつつある方法である。この規制のため、Gap Fill Overwriteでは、システムが潜在的なコンプライアンスの問題に引っ掛かる可能性がある。また、この上書きは、その上書きされたメッセージに対して市場が確認応答（ACK）で応答することを前提としている（relies on）。このACKがないと、クライアントはメッセージの拒絶について知らされず、自身の注文が通った（done for the day）のか、それとも取引所で未決済（open）のままなのかを知ることができない。実際に、このような問題が原因で「キャンセルの嵐」が起こることが確認されている。

【課題を解決するための手段】

【0010】

（好適な実施形態の概要）

本文献では、コンテンツ検査のための、分散型かつ透過的なインラインリスク管理及びトラフィックキャプチャシステムの改良について説明する。本明細書で説明するアプローチは、パケットがリアルタイムで検査デバイスに送り通されるなか、拒絶を受けたメッセージを変更したり、さらには、パケットから当該メッセージを除去したりする解決手段を提供する。除去したバイトを各セッションに基づいて追跡することにより、シーケンス番号の振当て（sequence numbering）が維持される。メッセージが変更された場合には、通知がメッセージ発信元に返されて変更が知らされる。

【0011】

一実施形態では、ネットワークセッションに対応するパケットが、第1のネットワークエンドポイントと第2のネットワークエンドポイントとの間を移動する際に検査される。所与の（selected）パケット内のメッセージが変更されるべきであると判断されると（例えば、当該メッセージがコンテンツルールに準拠していない場合等）、拒絶されたコンテンツを除去又は変更した変更後パケットが決定される。さらに、当該変更後パケットの変更後のシーケンス番号も、前記所与のパケットのシーケンス番号を調節することによって決定される。そして、この変更後パケットが、前記所与のパケットに代えて前記第2のエンドポイントへと送信される。さらに、返信メッセージが前記第1のエンドポイントに返されて、前記所与のパケット内の元のメッセージが変更されたことが知らされる。

【0012】

前記変更後のシーケンス番号は、アキュムレータ値を維持することにより、かつ、元のパケットと前記変更後パケットとのバイト数の差分を表すデルタ値により決定され得る。前記第2のエンドポイントからの前記変更後パケットの確認応答が検出されると、前記デルタ値が前記アキュムレータ値に加算され得る。

【0013】

また、受信した後続のパケットのシーケンス番号が前記シーケンスマーカー値を上回る場合、前記第1のエンドポイントから受信した後続のパケットのシーケンス番号が前記アキュムレータ値だけ変更され得る。

【0014】

また、前記第2のエンドポイントから受信した別のパケットの確認応答シーケンス番号が、前記アキュムレータ値に基づいて調節され得る。

【0015】

一部の実施態様では、前記第1のエンドポイントが取引業者に対応付けられたクライア

10

20

30

40

50

ント装置であり得て、前記第２のエンドポイントが証券取引システムに対応付けられた市場装置であり得て、前記メッセージが証券を売買する注文である。これらの実施態様では、前記返信メッセージが、注文を拒絶した理由を含み得る。

【００１６】

また、実施形態では、前記第２のエンドポイントから返信メッセージの確認応答を受信したか否かが判断され得る。この確認応答を受信していない限りは、タイムアウト時間を過ぎて（after）前記セッションが切断されるまで前記変更後パケットが繰返し再送され得る。

【００１７】

一部の実施態様では、前記返信メッセージの確認応答を受信したか否かを判断することが、前記デルタ値を定期的にポーリングすることを含み得る。

10

【００１８】

前記変更後パケットの確認応答を未だ受信しておらず、かつ、後続のパケットについても変更して第２の変更後パケットを生成する必要がある場合、前記変更後パケットの前記確認応答を受信するまで前記第２の変更後パケットをストールするのが有利であり得る。

【００１９】

好適な一部の実施態様において、通信は、前記第１のエンドポイントと第２のエンドポイントとの間でメッセージ及び（必要に応じて）変更後のシーケンス番号が付いた当該メッセージの確認応答が双方向に同時に送信され得る双方向通信（duplex）である。これにより、一方向においてコンテンツ検査及びメッセージ変更をオンザフライで（on the fly）行うと同時に、反対方向において同時に通知返信メッセージを注入することも可能になる。

20

【００２０】

以降の説明は、添付の図面を参照しながら行う。

【図面の簡単な説明】

【００２１】

【図１】例示的な一実施形態を示す高次の図である。

【図２】検査デバイスの詳細図である。

【図３Ａ】証券取引システムでの一実施形態を示す図である。

【図３Ｂ】より一般的なパケットフィルタリングの一実施形態を示す図である。

30

【図４】カットスルー方式の固定ロジックデバイスの詳細図である。

【図５】シーケンス番号を変更するのに実行される一連のステップの前半を示す図である。

【図６】シーケンス番号を変更するのに実行される一連のステップの後半を示す図である。

【図７】取引前リスクソフトウェア（PTRS）の詳細図である。

【図８Ａ】タイミング図である。

【図８Ｂ】他のタイミング図である。

【図９】クライアント→市場（C2M）方向のパケットをトリミングする際の詳細なシーケンス図である。

【図１０】市場→クライアント（M2C）方向のパケットを注入する際の詳細なシーケンス図である。

40

【図１１】パケットが廃棄されるシナリオを示す図である。

【図１２】パケットを変更するときに発生し得る問題を示す図である。

【図１３】図１２に示した問題に対して可能な解決策を講じた場合のシナリオを示す図である。

【図１４】パケットを変更するときに発生し得る別の問題を表したシナリオを示す図である。

【発明を実施するための形態】

【００２２】

前述したように、本発明の実施形態は、インラインコンテンツ検査・変更を提供するように構成されている。後述の方法や装置を実現するデバイスは、典型的に、第１および第

50

2のネットワークエンドポイント間に位置したコンピュータネットワーキングデバイスにおいて実現される。後述する好適な実施形態についての詳細な説明の多くは、高頻度証券取引システムにおいて成行注文（market order）を含んだメッセージを監視するための検査デバイスの文脈で行うものとする。しかしながら、検査デバイスの実施形態は、データストリームのコンテンツを最小限のレイテンシで監視するための他の用途でも配備されることが可能である。これらには、機微なデータを検出する用途や、データ完全性や順序付けが重要となる用途（例えば、データベースシステム等）や、健康記録（health records）処理システムや、ブロックチェーンシステムや、より一般的に言えばパケットフィルタリングデバイスが含まれ得る。

【0023】

例示的な一実施態様として、図1に示す高頻度証券取引ネットワークを取り上げたい。（証券会社により操作され得る）1つ以上のクライアント装置120、122が、（証券取引所により操作され得る）1つ以上の市場装置130、132とのネットワーキングセッション140、142を確立し、証券売買注文をこれらの市場装置130、132に送信する。売買注文は、メッセージとして形成され、パケット内で他のメッセージと合体させられたものであり得る。これらのメッセージは、アプリケーションレベルプロトコル（例えば、NASDAQ OUCH等）に準拠してやり取りされ得る。検査デバイス110は、前記クライアント装置と前記市場装置との間でネットワーキングセッション140、142を介して送信されるメッセージのコンテンツを検査する。好ましくは、検査デバイス110はカットスルー方式のデバイスであり、前記ネットワーク上での当該検査デバイス110の存在は、前記クライアント装置及び市場装置にとって不透明または見えないようになっている。このようなカットスルー方式の実施形態では、検査デバイス110でネットワークが途切れることなく、ネットワーキングセッション140、142が当該検査デバイス110をそのまま（straight）流れ通り得る。一部の実施形態において、ネットワークセッション140、142は、TCP、InfiniBandなどのシーケンス化（sequenced）トランスポートレベルプロトコルに準拠して確立される。

【0024】

この実施形態において、検査デバイス110は、「不合格の（bad）」売買注文が確実に拒絶されるように機能し、つまり、クライアント→市場方向（C2M）に流れるメッセージを、幾つかのルールに準拠していない際には対応する（respective）市場装置に到達させないように機能する。検査デバイス110は、クライアント装置120から送信された、不合格の（badly formed）又は準拠していないメッセージが市場装置130に到達するのを防ぐために、この準拠していないメッセージに変更を施し、場合によっては当該メッセージのサイズを調節し得る。前記ネットワークセッションがTCPなどのシーケンス化トランスポートプロトコルに準拠したものである環境では、検査デバイス110が、パケット内の準拠していないアプリケーションレベルメッセージを変更したことによるバイト数（bytes）の増減を考慮して、同じセッション（that session）内のC2M方向に届く後続のパケットのトランスポートプロトコルシーケンス番号（およびM2C方向に届く後続のパケットの確認応答シーケンス番号）を調節し得る。

【0025】

また、好適な一実施形態において、検査デバイス110は、準拠していないメッセージが変更されたことを知らせる返信メッセージを生成し、その準拠していないメッセージを発信した前記クライアント装置へと、この返信メッセージを送信する。一部の実施形態において、前記返信メッセージは、検査デバイス110により前記ネットワークセッションへと「注入」されたものであり得て、かつ、前記準拠していないメッセージの意図された受信先である市場装置から発信されたかの如く見え得る。前記クライアント装置と前記市場装置との間に確立された前記ネットワークセッションがTCPなどのシーケンス化トランスポートプロトコルに準拠したものであるときに、当該ネットワークセッションへと前記返信メッセージを注入する場合には、当該セッションにおいて市場→クライアント方向（M2C）に送信されるパケットのトランスポートプロトコルシーケンス番号の、当該返

10

20

30

40

50

信メッセージのサイズを考慮した調節も含み得る。

【0026】

図2に、検査デバイス110の例示的な一実施形態の詳細を示す。この例では検査デバイス110の機能がハードウェアとソフトウェアとの両方にわたって分散しているが、検査デバイス110は、ハードウェアだけの実現（実装）やソフトウェアだけの実現（実装）を含め、ハードウェアとソフトウェアとのあらゆる適切な組合せで実現されてもよい。例えばこの実施形態では、低レベルのレイテンシ（潜在または待ち）を達成するために、一部の機能がカットスルー方式の固定ロジックデバイス110によりハードウェアで実現されると同時に、他の機能はデバイスドライバ250及び取引前リスクソフトウェアアプリケーション220によりソフトウェアで実現される。カットスルー方式の固定ロジックデバイス210は、特定用途向け集積回路（ASIC）やフィールドプログラマブルゲートアレイ（FPGA）を含め、あらゆる適切な様式で実現されてもよい。取引前リスクソフトウェアアプリケーション220及びデバイスドライバ250は、コンピューティングコア（CPU）などの少なくとも1つのプログラマブルなデータプロセッサで実行される命令として実現され得る。

10

【0027】

（ファイバーケーブルや銅ケーブルを介したInfinibandやイーサネット（登録商標）を含め）あらゆる適切な物理ネットワーク層を用いることもできるが、この例における検査デバイス110は、クライアント装置、市場装置に各自接続され得る、2つのギガビットイーサネットSFP+コネクタ（インターフェース）231、232を具備している。この例では、これらのコネクタ231、232が、10GigE（ギガビットイーサネット）MACコア211、212に各自電子的に接続されている。この実施形態では、10GigE MACコア211、212が、カットスルー方式の固定ロジックデバイス210により実現されている。

20

【0028】

一部の実施形態では、カットスルー方式の固定ロジックデバイス210が、さらに、他のコンポーネントを含み得る。図2の例では、カットスルー方式の固定ロジックデバイス210が、さらに、後で詳述するパケット検査、シーケンス番号変更などの機能を実現し得る固定ロジック213コンポーネントを含む。固定ロジック213も、（後述するように）FPGA213により実現され得る。カットスルー方式の固定ロジックデバイス210は、さらに、PCI Express機能を実現し得るPCI-Eコア214を含む。この例において、PCI Expressは、ハードウェアとソフトウェアとの間で（より具体的には、カットスルー方式の固定ロジックデバイス210と取引前リスクソフトウェアアプリケーション220との間で）PCI Expressバス240を介してデバイスドライバ250を通じてデータを転送するための導管機構として使用される。しかしながら、ハードウェアとソフトウェアとの間では、ダイレクトメモリアクセス（DMA）や（リングバッファ形式で配置され得る）共有メモリバッファやメモリマッピングを含め、あらゆる適切なデータ転送メカニズムが用いられてよい。

30

【0029】

（検査デバイス110により実行される機能）

40

図3Aは、取引ネットワークにおいて検査デバイス110がストリーム動的変更（DSM）を実現しうる様子の一例を示す図である。ステップ301にて、クライアント装置120が、証券を売買する注文であるメッセージを含んだ元の又は本来の（original）パケットを生成する。当該元のパケットは、クライアント装置120と市場装置130との間でのセッション140コネクションを介して市場装置130へと宛てられたパケットである。ただし、このパケットは、まず、セッション140を監視することも可能な検査デバイス110を通過する（travels through）。この例において、前記元のパケットは、クライアントから市場への注文（OCM）として、キャタピラー社（CAT）の株を購入する第1のOCM、アップルコンピュータ社（APPL）を売却する第2の注文、グーグル社（GOOG）を購入する第3の注文、およびテスラ社（TSLA）を売却する第4の注

50

文の4つを含んでいる。

【0030】

検査ノード110は、これらのメッセージのコンテンツが取引ルールに準拠しているか否かを検査する。例えば、それらのルールは、前記注文をチェックして、数量や価格が想定範囲内であるか否かや、制限付き株式についての取引、空売り又は明らかに誤りのある取引であるか否かを判断するものであり得る。また、ルールのチェックは、アカウント毎の数量、価格及び価値限度についての検査や、信用限度、集中限度、繰返し注文、エクスポージャー、仲介業者又は証券会社（broker）アカウント、およびセッションディセーブル（session disable）についての検査を含み得る。前記メッセージを検査するのに用いられる具体的なルールには数多くの種類のものがあり得て、どのようなルールであるのかは本実施形態にとって重要でない。

10

【0031】

重要なのは、ステップ302にて検査デバイス110が、前記第3のメッセージ（グーグル社の株の注文）が不合格のメッセージであり、当該メッセージの全体を市場装置130に到達させるべきでないという点である。結果として、市場装置に送信される変更後パケットにはその不合格のメッセージが含まれず、ステップ303にて通知メッセージがクライアント120へと送り返される（例えば、セッション140へと「注入」される）。好適な一実施形態において、当該通知メッセージは、注文が拒絶された理由が入るフィールドを含む。

【0032】

20

前述したように、本明細書で説明するストリーム動的変更（DSM）の概念は、他のコンテンツ検査用途にも適用可能であることを理解されたい。図3Bに、この状況をさらに一般化したものを示す。ここでは、第1のエンドポイント320が第2のエンドポイント330とのセッションを確立している。ステップ351にて、前記第1のエンドポイントから、メッセージm1, m2, mN, …mKを含む元の（本来の）パケットが送信される。ステップ352にて、検査エンジン310は、そのうちの一つのメッセージmNが変更される必要があると判断し、変更後メッセージmN'を生成する。そして、メッセージm1, m2, mN', …mKを含む変更後パケットが、第2のエンドポイント330へと移動させられる。ステップ353にて、通知メッセージが第1のエンドポイント320へと返されて、前記元のパケット内のメッセージの一つが変更されたことを第1のエンドポイント320に知らせる。

30

【0033】

（ストリーム動的変更）

好ましくは、前述のように検査デバイス110が、C2M TCPストリームのパケットから不合格のメッセージを除去してM2C TCPストリームのパケットにエラーメッセージを追加する手段としてのストリーム動的変更（DSM）を実現する。検査デバイス110は、TCPセッションを妨害することなくこれを実行するために、a) C2M方向において除去された全バイトを追跡し、b) TCPシーケンス番号をオンザフライで（on the fly）変更することによってそれを行う。

【0034】

40

DSMでは、好ましくは各セッション・各方向毎に、例えば以下のような数種類のデータ値が維持される：

【0035】

アキュムレータ：TCPストリームが続くあいだ（throughout the life of）除去／挿入された総バイト数を指す。好ましくは、これは、パケットのTCPシーケンス番号を変更するのに利用されるハードウェアレジスタ（つまり、カットスルー方式の固定ロジックデバイス210内に維持されている）である。全ての変更バイトを追跡することにより、パケットのTCPシーケンスをオンザフライで変更することが可能になる。ただし、このレジスタは、パケットのバイトが変更された場合であっても、当該パケットのACKを受信してからでしか更新されないものとされるのが好ましい。

50

【0036】

デルタ：シーケンスマーカーを設定した最後のパケット（最新の変更後パケット）において除去／挿入されたバイト数を指す。好ましくは、デルタは、この最後のパケットで除去されたバイト数の一時的記憶部として機能するハードウェアレジスタである。デルタは、このパケットのACKが受信されるとアキュムレータに加算される。

【0037】

シーケンスマーカー：バイトが除去／挿入された最後のパケット（すなわち、最新の変更後パケット、つまり、キル又は注入された最新のパケット）のシーケンス番号に基づいた数値を指す。例えば、パケットに開始シーケンス番号やACKシーケンス番号が付くTCPを使った実施形態では、シーケンスマーカーが、最新の変更後パケットの（開始）シーケンス番号と当該最新の変更後パケットの（変更後の）バイト長との合計に基づいた数値とされ得る。TCPでは、この数値が、さらに、同じセッション内で同じ方向に移動する後続のパケットに対しての、次に予想される開始シーケンス番号とも等価又は同一になる。変更の一環としてパケットが完全に除去された場合、この除去したパケットの長さは変更後にゼロになることから、当該除去したパケットに対応するシーケンスマーカーは、当該除去したパケットの開始TCPシーケンス番号と同一になり得る。TCPを用いた場合には、シーケンスマーカー値が、さらに、（最新の変更後パケットが変更の結果として完全に除去されていないことが前提になるが）当該最新の変更後パケットを確認応答するACKシーケンス番号とも同一になる。このACKシーケンス番号は、反対方向に送信された、変更後パケットの受信を確認応答するためのパケット内に存在する。

【0038】

なお、シーケンスマーカーよりも小さいシーケンス番号を持つパケットは、TCPシーケンスがアキュムレータぶんどけ変更される。シーケンスマーカー以上のシーケンス番号を持つパケットの場合には、TCPシーケンスがアキュムレータにデルタを足したぶん変更される。最新の変更後パケットのACKが受信されると、シーケンスマーカーがクリアされる。一部の実施形態では、最新のものまでの複数の変更後パケット（the last several modified packets）に対応する複数のシーケンスマーカー値が記憶され得る。これにより、確認応答が未だ済んでいない変更後パケットが複数存在する場合であっても、パケットのシーケンス値の調節を行うことが可能になる。

【0039】

CONN__MOD__PENDING：変更に対してACKを未だ受信していないことを示すための、（PTRSアプリケーションソフトウェア220で設定され得る）ブーリアン値を指す。変更後パケットがDMAされると、このブーリアンが設定される。変更後パケットがACKされると、このブーリアンがクリアされる。このブーリアンは、変更後パケットが当該パケットの意図される宛先に到達したことを確認するために用いられる。データの変更は双方向で同時に発生する場合もある（典型的にそうである）ので、CONN__C2M__MOD__PENDINGとCONN__M2C__MOD__PENDINGとが存在する。

【0040】

以下の機能は、検査デバイス110により、ストリーム動的変更の一環として実行され得る：

【0041】

不合格のメッセージの除去：DSMは、C2Mパケットストリームのパケットから不合格のメッセージを除去すると共に、M2Cパケットストリームのパケットにエラーメッセージを追加する。DSMは、コネクションを妨害することなくこれを実行するために、a) 除去された全バイトを追跡し、b) TCPシーケンス番号をオンザフライで変更する必要がある。DSMは、これを行うにあたって、各セッション毎にアキュムレータ、シーケンスマーカーおよびデルタを含む数種類の数値を維持する（後の欄で詳述する）。

【0042】

C2M（クライアント→市場）変更パケット構築・送信・到達保証：検査デバイス11

10

20

30

40

50

0 は、拒絶した C 2 M パケットを、不合格のメッセージを除去することによって変更し得る。次に、これらの変更後パケットが、市場装置へと送信される。遅延 A C K 検出とポーリングと繰返し再送との組合せを用いることにより、それら変更後パケットの送信成功が保証され得る。C 2 M 変更後パケットの、市場への到達が保証される。

【0043】

M 2 C (市場→クライアント) エラーパケット構築・送信・(送信者への) 送戻し・到達保証：検査デバイス 110 は、メッセージが拒絶されたことをクライアントに知らせるために M 2 C エラーパケットを構築し得る。

【0044】

T C P 機能障害／切断の回避：D S M は、T C P 機能障害を回避して、T C P ストリームでデータを整然に (cleanly) 除去／追加し得る。また、D S M は、T C P フラグメンテーション (断片化) の最中であっても変更後メッセージの送信実行 (attempts) を処理し得る。セッションの切断は、必要のない限り避けられるべきである。

【0045】

データ挿入／エラー注入：C 2 M 方向で変更が行われると、データ挿入／エラー注入を用いて拒絶がクライアントに知らされ得る。1) ヘッダおよび 2) 各プロトコル固有の (protocol specific) 拒絶メッセージを含むメッセージが連結してなるパケットが、M 2 C T C P ストリームへと注入される。これにより、クライアントは、自身の注文の状態を知ることができ、不要なキャンセルの嵐やアルゴリズムの混乱が回避される。このエラー注入は市場側から完全に隠れており、バイトが注入されても T C P コネクションはアクティブのまま維持される。

【0046】

A C K 検出／到達保証：好ましくは、検査デバイス 110 は、クライアント方向及び市場方向のいずれに変更後パケットが送信された場合にも、当該パケットの到達を保証するものであるのが望ましい。この到達保証は、1) 変更後パケットが送信された方向と反対方向に届いたパケットの A C K シーケンスをチェックしたり、2) 固定ロジック (F P G A) 213 からデルタレジスタを読み出すことで、当該デルタレジスタがゼロになったか否か (シーケンスマーカーに相当する A C K が届いたことを示唆する) を確認したりすることによって実行され得る。A C K 検出は、これら両方の方法を用いて、パケットが意図される宛先に到達したか否かをチェックする。A C K が一定の時間枠内に検出されなかった場合には、変更後パケットが再送される。これは、検査デバイス 110 が諦めてセッションを切断するまで続けられる。

【0047】

データ除去／変更：検査デバイス 110 は、キルされたパケットからバイトを除去することにより、メッセージの拒絶方法としての Gap F i l l O v e r w r i t e や O n e a n d D o n e を避けることができる。このデータ除去は市場側から完全に隠れており、T C P コネクションがアクティブのまま維持される。

【0048】

(カットスルー方式の固定ロジックデバイス 210)

D S M は、C 2 M T C P ストリームのパケットから不合格のメッセージを除去すると共に、M 2 C T C P ストリームのパケットにエラーメッセージを追加する。D S M は、コネクションを妨害することなくこれを実行するために、a) 除去された全バイトを追跡し、b) T C P シーケンス番号をオンザフライで変更する必要がある。D S M がこれを行うにあたって使用する 3 種類の数値については、先述のリストアップとおりである (アキュムレータ、シーケンスマーカーおよびデルタ)。

【0049】

図 4 は、最初に図 2 で図示したカットスルー方式の固定ロジックデバイス 210 の一実施形態の詳細図である。カットスルー方式の固定ロジックデバイス 210 は、(クライアント側と市場側とで一つずつの) 2 つの 10 G i g E M A C コア 211, 212、固定ロジック (F P G A) 213、P C I e x p r e s s コア 214、およびメッセージ変更デ

10

20

30

40

50

ータストア 440 を含む。

【0050】

一部の実施形態において、固定ロジック 213 は、パケット検査エンジン (PIE) コンポーネント 420、およびシーケンス番号変更ロジックコンポーネント 430 を含む得る。パケット検査エンジン 420 は、クライアントと市場装置との間のネットワーキングセッションで送信された、アプリケーションレベルメッセージを含むネットワーク化パケットを検査し、クライアントから発信されたメッセージのうち、所与のメッセージが準拠しており市場装置へと到達させてよいものなのか、それとも、当該メッセージが変更の必要があるものなのかを判断し得る。一部の実施形態では、メッセージが変更されると、シーケンス番号変更ロジック 430 により、同じセッションにおける後続のパケットのトランスポートプロトコルシーケンス番号が変更され得る。

10

【0051】

この例では、カットスルー方式の固定ロジックデバイス 213 が、さらに、メッセージ変更データストア 440 も含んでいる。メッセージ変更データストア 440 は、メッセージの変更やシーケンス番号の変更に関連して使用される状態情報を記憶し得る。メッセージ変更データストア 440 は、各セッション・各方向毎 (C2M と M2C の両方) に維持され得る。この例ではメッセージ変更データ 440 がカットスルー方式の固定ロジックデバイス 110 自身内に存在しているものとして描かれているが、当該メッセージ変更データはカットスルー方式の固定ロジックデバイス 210 内の任意の他の適切な記憶コンポーネント、またはカットスルー方式の固定ロジックデバイス 210 にアクセス可能である任意の他の適切な記憶コンポーネント内に存在しているものとされてもよい。

20

【0052】

メッセージ変更データ 440 は、累積値を追跡するのに各々用いられ得る少なくとも 1 つのハードウェアアキュムレータ 441 を含む得る。所与のセッション内で所与の方向に流れるパケットのシーケンス番号は、当該セッション・方向での少なくとも 1 つの先行する (prior) パケットの変更の結果として、その累積値だけ変更され得る。所与のセッションのプロトコルとして、受信側エンドポイントにより確認応答されたパケットのシーケンス番号を含む確認応答 (ACK) パケットが含まれている実施形態では、ハードウェアアキュムレータ 341 が、同じセッション (that session) 内で反対方向に流れる ACK パケットの ACK シーケンス番号を変更するのににも使用され得る。一部の実施形態では、少なくとも 1 つのデルタレジスタ 442 を維持するのが有用であり得る。デルタレジスタ 442 は、ACK パケットでの確認応答が未だ済んでいない、新たに変更されたばかりのパケットのサイズが変化したこと (a difference in size) によるシーケンス番号の差分を追跡するのに用いられ得る。デルタレジスタ 442 の数値は、前記検査デバイス内の、ハードウェアレジスタまたは任意の他の適切な揮発型もしくは不揮発型のメモリもしくはストレージに記憶され得る。メッセージ変更データ 440 は、さらに、少なくとも 1 つのシーケンス番号マーカー 444 を含む得る。シーケンス番号マーカー 444 は、パケットのシーケンス番号に基づいて処理 (action) を行う際の少なくとも 1 つの閾値として利用され得る。例えば、一部の実施形態において、シーケンス番号マーカー 443 は、新たに送信されたばかりの変更後パケットの予想 ACK シーケンス番号を表すものとして用いられ得る。そして、シーケンス番号マーカー 443 に記憶された数値と一致する ACK シーケンス番号を持つパケットを受信すると、デルタレジスタ 442 の数値がハードウェアアキュムレータ 441 に適用され得る。

30

40

【0053】

一部の実施形態では、準拠していないパケットが未だ検査デバイス 110 による変更過程にある場合、この変更中のパケットと同じセッションかつ同じ方向で受信した後続のパケットの送信を、当該変更中のパケットが完全に処理されるまで先送り (delay) するのが望ましくあり得る。この先送りは、変更が必要なパケットの後に同じセッション内で送り出された準拠している (変更が必要でない) パケットが、その変更中のメッセージよりも先に宛先の市場装置に到達してしまうという状況を防ぐのに望ましくあり得る。このよ

50

うな実施形態では、パケットスキップ阻止インジケータ 4 4 6 をイネーブルする（有効にする）ことによって変更を必要とするメッセージが検出されたことを示し得ると共に、未だ変更過程にあるそのようなパケットのシーケンス番号の数値が最新シーケンス番号 4 4 4 にも記憶され得る。この例では、パケットスキップ阻止インジケータ 4 4 6 がイネーブルされている間は、後続のパケットの送信が先送りされ得る。一部の実施形態では、さらに変更中のパケットが完全に処理され、かつ、パケットスキップ阻止がイネーブルされた結果として先送りされていた後続のパケットが完全に処理されると、パケットスキップインジケータ 4 4 6 がクリアされ得る。一部の実施形態では、これに加えて又はこれに代えて、パケットスキップ阻止インジケータ 4 4 6 をクリアするか否かの判断が、変更後パケットのシーケンス番号と最新シーケンス番号 4 4 4 の数値とを比較することを含み得る。パケットスキップ阻止インジケータ 4 4 6 がクリアされた後に検査デバイス 1 1 0 に届いたパケットは、通常の方法で処理されることになり得る。そして、準拠しているパケットは、前記検査デバイスをそのまま（directly）通って、意図される受信先システム（すなわち、クライアント装置 1 2 0 又は市場装置 1 3 0）へと向かい得る。

【0054】

同様に、一部の実施形態では、C 2 M 方向のメッセージの変更が必要であり、かつ、対応する返信メッセージを同じセッション（that session）の M 2 C 方向に注入してもよいと判断された場合、コネクションストールモードインジケータ 4 4 8 を用いて、M 2 C 方向に入って来たパケットの送信が先送りされ得る。一部の実施形態では、C 2 M パケットの変更が必要かもしれないことが検出されるとコネクションストールモードインジケータ 4 4 8 がイネーブルされ得て、M 2 C 返信メッセージが完全に構築又は注入されると当該コネクションストールモードインジケータ 4 4 8 がディセーブルされ（無効にされ）得る。このような実施形態では、コネクションストールモードインジケータ 4 4 8 がイネーブルされている間は、同じセッション（that session）の M 2 C 方向で受信した後続のメッセージが先送りされ得る。パケットスキップ阻止インジケータ 4 4 6 で示される C 2 M 方向の先送り、およびコネクションストールモードインジケータ 4 4 8 で示される M 2 C 方向の先送りのどちらの場合の先送りの、あらゆる適切な方法で実現されてよい。例えば、一部の実施形態では、前記インジケータがイネーブルされている間、対応する方向（that direction）のメッセージを処理する C P U コアがストールされ得る。他の実施形態では、送られてきた後続のメッセージが、後で処理できるように（for later processing）キューに格納され得る。

【0055】

（通知ありのメッセージ変更のための一連の処理の一例）

図 5 及び図 6 は、シーケンス動的変更を実現するために検査エンジン 1 1 0 により実行され得る一連の処理を表した具体例である。

【0056】

これらの図には、アキュムレータ、デルタおよびシーケンスマーカーを用いて、T C P コネクションの一方の側（C 2 M ストリーム）を維持しながら、当該コネクションを切断又は中断せずにパケットからバイト（さらにはメッセージ全体）を除去しうるのが描かれている。

【0057】

図 5 を参照すると、時刻 t_0 （ステップ 500）にて、アキュムレータおよびデルタが「0」などの初期値に設定される。時刻 t_1 （ステップ 501）では、パケットを「キル」する必要がある（つまり、パケット内の少なくとも 1 つのメッセージが変更される必要がある）場合、除去したバイト数がデルタレジスタに記憶される。この例では、シーケンス番号 70 から始まる 10 個のバイトがキルされて、デルタレジスタの数値が「-10」に設定された。これに応じて、シーケンスマーカーの数値は「70」に設定された。そのパケットの A C K が届くと、デルタのバイト数がアキュムレータに加算される。キルされたパケットから除去されたバイト数は、当該パケットの送信後直ぐにアキュムレータへと加算されるわけではない。というのも、先ほどの A C K が、キルされたパケットの前に送

信されたパケットの確認応答である可能性もあるからである。キルされたパケットの前に送信されたそのようなパケットのACKが間違った数値にならないよう確実にするために、アキュムレータへのデルタの加算は時刻 t_2 まで先延ばしされる（ステップ502）。

【0058】

時刻 t_2 では、シーケンスマーカー値、アキュムレータ値、およびデルタレジスタ値を用いて、M2C方向に移動するACKパケットのACKシーケンス番号が調節される。

【0059】

時刻 t_3 に受信した第2のC2Mパケット（ステップ503）は、開始シーケンス番号が80で、パケット長が20である。このパケットは、不合格のメッセージを含んでおらず、キルされる必要がなく、市場へと渡されてよいものである。ただし、このパケットは、アキュムレータ値を用いて当該パケットのシーケンス番号が調節されてから、市場へと引き継がれる。

【0060】

時刻 t_4 にて、前記第2のC2MパケットのACKを受信する（ステップ504）。このACKも、再びアキュムレータ値を用いてシーケンス番号が調節されてからクライアントに渡される。

【0061】

続いて図6を参照すると、時刻 t_5 にて受信した第3のC2Mパケット（ステップ505）は、シーケンス番号が100で、パケット長が20である。第3のC2Mパケットも、アキュムレータ値によってシーケンス番号が調節されてから市場側に渡される。

【0062】

時刻 t_6 （ステップ506）にて、前記第3のC2MパケットのACKを市場から受信する前に、シーケンス番号が120で、パケット長が20である第4のC2Mパケットを受信する。この第4のパケットは、バイトが除去又は「キル」される必要がある。時刻 t_7 （ステップ507）では、この（previous）パケットのACKを未だ受信していないので、M2C確認応答シーケンス番号からアキュムレータだけが減算される。 t_8 （ステップ508）にて、ACKを受信すると、アキュムレータへのデルタの加算が可能となる。

【0063】

時刻 t_9 （ステップ509）にて、シーケンス番号が140で、パケット長が20である第5のC2Mパケットを受信する。これまでと同じく、アキュムレータ値がシーケンス番号に加算されて（added to arrive at）変更後パケットとなる（use for the modified packet）。時刻 t_{10} （ステップ510）にて受信したさらなる（another）ACKも、アキュムレータ値だけ調節される。

【0064】

前述したように、DSMのこの実施態様では、C2M側のアーキテクチャとM2C側のアーキテクチャとが互いに平行とされる。よって、好ましくは、C2M方向とM2C方向とで、かつ、各セッション毎に、別々のアキュムレータ、デルタおよびシーケンスマーカーが存在する。結果として、M2C TCPストリームへと通知メッセージ用にバイトが追加される際（エラー注入の場合）には、正のデルタが使用され得る。結果として、M2C側でデルタがアキュムレータに加算されると、バイトの加算を表す正の変化が見られる。M2CメカニズムとC2Mメカニズムとのこのような二重性は、好適なDSMアーキテクチャの重要な側面である。

【0065】

後述の詳細な説明からも理解できるように、一部の実施態様では、固定ロジック213（FPGA）ハードウェアのアキュムレータ及びデルタレジスタの数とPTRSソフトウェア220のアキュムレータ及びデルタレジスタの数とが異なるものとされ得る。例えば、PTRSソフトウェア220はアキュムレータとデルタとのセットを2セット（各フローでのC2MとM2Cとの各方向に1セットずつ）を使用し得るのに対し、固定ロジックハードウェア213はアキュムレータとデルタとのセットを4セット使用し得る。ここで、ハードウェアアキュムレータ／デルタの最初の2セットは、各セッション（フロー）の

10

20

30

40

50

各方向（C 2 M及びM 2 C）における通常シーケンス番号を追跡するのに用いられる。残りの（second）2セットは、各セッションの各方向（C 2 M及びM 2 C）におけるA C Kシーケンス番号を追跡するのに用いられる。以上の考え方に沿ったアキュムレータ／デルタのモデルについては、後で詳細に述べる。

【0066】

図7は、検査デバイス110内における取引前リスクソフトウェア220のコンポーネント（構成要素）であって、モジュール分解（Modular Decomposition）、C 2 M不合格メッセージ変更およびC 2 M変更送信を含む、一部のD S M機能を実現するコンポーネントを示す図である。

【0067】

取引前リスクソフトウェア220は、メッセージ変更データストア770を含み得る。メッセージ変更データストア770は、取引前リスクソフトウェア220によってメッセージ変更に関連して利用され得る。カットスルー方式の固定ロジックデバイス210との関連で説明したメッセージ変更データ440と同じく、メッセージ変更データ770は、各セッション・各方向（すなわち、C 2 MとM 2 Cの両方）毎に記憶される状態情報であり得る。メッセージ変更データ770は、取引前リスクソフトウェア220により任意の適切な方法で記憶・アクセスされ得る。メッセージ変更データ770は、ハードウェアアキュムレータ441と同様の概念である少なくとも1つのソフトウェアアキュムレータ771、およびデルタレジスタ442と同様の概念である少なくとも1つのデルタ値772を含み得る。メッセージ変更データ770は、さらに、メッセージペンディングインジケータ775（本明細書では、C 2 M方向がCONN_C 2 M_MOD_PENDING、M 2 C方向がCONN_M 2 C_MOD_PENDINGと称される場合もある）を含み得る。一部の実施形態では、これらのメッセージ変更ペンディングインジケータが、変更後パケットのA C Kを未だ受信していないことを示すのに用いられ得る。

【0068】

図7には、さらに、検査デバイス110を流れ通るパケットを処理して且つメッセージ変更データ770を利用する、コンポーネント710～746（以下で詳細に説明する）が描かれている。この例では、コネクションマネージャ710が、例えば発信I Pアドレス／ポートや宛先I Pアドレス／ポートに基づいてセッションを特定し、かつ、A C K及びSYN A C Kの状態を維持するなどしてクライアント側と市場側との両方のコネクション状態のモデルを実現し得る。また、コネクションマネージャ710は、遅延A C K検出モジュール720をコールし得る。遅延A C K検出モジュール720は、メッセージ変更データ470を利用して市場による変更後パケットの受信やクライアントによる注入メッセージの受信が確実に成功するようにさせるためのロジックを有し得る。遅延A C K検出モジュール720は、変更後パケット又は注入パケットのA C Kを受信側で受信するまで当該パケットを定期的に再送することによりそれを実行し得る。

【0069】

また、T C Pなどといった、順番どおりの送達（in-order delivery）を保証するシーケンス化されたネットワークセッションを用いた実施形態において、コネクションマネージャ710は、パケットが当該パケットのシーケンス番号に従った適切な順番で検査デバイス110により処理されるよう確実にし得て、パケットの重複処理を回避し得る。

【0070】

前述したように、パケットは典型的に1つ以上のメッセージを含み得る。よって、パケット内の各メッセージは、個別にコンポーネント732，734，736により処理され得る。様々なアプリケーションプロトコルに準拠したメッセージが送信され得る実施形態では、プロトコル抽象化レイヤ732がアプリケーションプロトコル毎の処理を実行し得る。

【0071】

次に、C 2 Mデータ変更パケット構築モジュール734により、変更が必要なメッセージが処理され得る。C 2 Mデータ変更パケット構築モジュール734は、準拠していない

10

20

30

40

50

メッセージに変更を施し得る。また、C 2 Mデータ変更パケット構築モジュール 7 3 4 は、変更後のメッセージのサイズが元の変更前のメッセージのサイズから変化した (different) 場合に、(他の箇所で説明するように) セッションに対応した C 2 M方向のソフトウェアアキュムレータ 7 7 1 および／またはデルタ値 7 7 2 の新しい数値を算出し得る。同様に、M 2 Cデータ挿入パケット構築モジュール 7 3 6 が、その変更後のメッセージと同じセッションの M 2 C方向に注入すべき返信メッセージを構築し得て、かつ、当該セッションに対応した M 2 C方向のソフトウェアアキュムレータ 7 7 1 および／またはデルタ値 7 7 2 の新しい数値も算出し得る。

【0072】

図 7 の例では、C 2 Mデータ変更パケット構築モジュール 7 3 4 が、次に、この変更後パケットを市場へと送信するためのロジックを実行し得る。一部の実施形態では、C 2 Mデータ変更パケット構築モジュール 7 3 4 が、カットスルー方式の固定ロジックデバイス 2 1 0 への変更後パケットのデータ転送を実行して当該変更後パケットを市場装置 1 3 0 へと送信し得る。C 2 Mデータ変更パケット構築モジュール 7 3 4 は、さらに、C 2 M方向のソフトウェアアキュムレータ 7 7 1 および／またはデルタ値 7 7 2 の対応する数値もカットスルー方式の固定ロジックデバイス 2 1 0 に転送し得る。

【0073】

同様に、M 2 Cデータ挿入パケット構築モジュール 7 3 6 が、クライアント 1 2 0 へと送信すべき返信メッセージをセッションに注入するためのロジックを実行し得る。一部の実施形態では、M 2 Cデータ挿入パケット構築モジュール 7 3 6 が、返信メッセージを含む新たに構築されたパケットの、カットスルー方式の固定ロジックデバイス 2 1 0 へのデータ転送を実行して当該パケットを送信し得る。M 2 Cデータ挿入パケット構築モジュール 7 3 6 は、さらに、M 2 C方向のソフトウェアアキュムレータ 4 7 1 および／またはデルタ値 4 7 2 の対応する数値もカットスルー方式の固定ロジックデバイス 2 1 0 に転送し得る。

【0074】

P T R Sアプリケーションソフトウェア 2 2 0 の上記コンポーネントにより制御される D S Mデータフローの一例は、以下のようになり得る：

【0075】

キルされたメッセージを含むパケットがデータパスに (例えばコネクションマネージャ 7 1 0 に) 到達すると、当該パケットは (例えばプロトコル抽象化レイヤ 7 3 2 により) パースされてから、C 2 Mデータ変更パケット構築 7 3 4 にて変更が施されるように送り出される。

【0076】

そして、M 2 Cデータ挿入パケット構築 7 3 6 により、キルされた各メッセージ毎に M 2 Cエラーパケットが生成される。このパケットは、拒絶をクライアントに知らせるのに使用される。

【0077】

これらのメッセージが生成された後、送信機能 (例えば C 2 Mデータ変更 7 4 4 及び M 2 Cデータ挿入 7 4 6) がコールされる。キルされた変更後 C 2 Mパケットと M 2 Cエラーパケットの両方が送信される。

【0078】

これらのパケットが送信された後、当該パケットシーケンスの A C K が遅延 A C K 検出モジュール 7 2 0 において検出される。到達を保証するために、パケットは例えば 1 0 m s (ミリ秒) 毎に合計 1 0 0 m s になるまで定期的に再送され得る。1 0 0 m s の時点で、A C K が届いていなければセッションが切断される。

【0079】

C P M (近接変更 [Close Proximity Modification]) 検出、コアストール (core stalling) などの他の機能も、送信モジュール 7 4 4 , 7 4 6 で実行され得る。C P M は、先に (first) キルされたパケットの A C K が届く前に次の (second) キルされたパケッ

10

20

30

40

50

トがソフトウェアに届いた場合に起こる。この場合については、後述の欄で詳細に説明する。

【0080】

好ましくは、C2M方向のデータ変更は、あらゆる数の不合格メッセージをアクティブTCPストリームにおいて除去又は変更するということが、アクティブTCPセッションのシーケンス順序（sequencing）を維持しながら出来るように実現される。つまり、モジュール744でのTCPパケットの変更は、クライアント側TCPスタック及び市場側TCPスタックから見えない（hidden）ようになされるのが望ましい（ただし、クライアントに対しては、メッセージが除去されたことが前述のように通知される必要がある）。

【0081】

一部の実施態様では、不合格のメッセージが、ハートビートなどの何らかの種類の新たなメッセージに置き換えられる。すなわち、一部の実施態様ではC2Mデータ変更744が、不合格のメッセージを除去する際に必ず当該不合格のメッセージを他の何かに置き換える。これにより、メッセージの変更が単純化されると共に、少なくとも幾つかのペイロードが送信されることが可能となる（すなわち、パケットは、完全に上書きされて0バイトのペイロードになるようなことがなくなる）。

【0082】

一部の実施形態では、メッセージ変更機能の少なくとも一部が、プロトコル抽象化レイヤ732のロジックにより実現され得る。これは、必要となる変更の種類が、所与のセッションを確立したときの市場プロトコルが具体的に何であるのかに左右され得るからである。例えば、NASDAQOUCHアプリケーションプロトコルに準拠したものである場合、パケットからメッセージ全体が除去されてもよい。しかし、TSEアプリケーションプロトコルに準拠してメッセージをやり取りする場合には、市場とクライアントとの間のアプリケーションレベルのシーケンス順序を維持するために、キルされたメッセージがハートビートに置き換えられるのが望ましい。

【0083】

一部の実施形態では、C2Mデータ変更送信（再送）が、例えば他にも以下のような機能を実行し得る：

【0084】

－パケットのIPヘッダ長を再計算して、除去されたバイトが含まれていないことを確認する；

【0085】

－DMAを用いて変更後パケットをカットスルー方式の固定デバイス210に転送し返す際に、必要な（例えばDMAヘッダ内やDMAコール等の）数値／ビットを設定する。これらの数値には、下記のものが含まれ得る：

【0086】

DMAヘッダ内のRaw（未変更）／Transform（変更）ビット：このDMAヘッダビットを「1」に設定することにより、ハードウェアはTCPチェックサム、CRCおよびTCPシーケンスを再計算する。

【0087】

DMAヘッダ内の符号付き12ビットデルタ：DMAヘッダ内のこのデルタを設定することにより、パケット内の変更されたバイト数を伝える（正のデルタ信号バイトは加算されたことを指し、負の信号バイトは減算されたことを指す）。固定ロジック213は、この数値を用いて当該固定ロジック213のハードウェアデルタレジスタを更新する。

【0088】

符号なし32ビットシーケンスマーカー：これは、シーケンスマーカーを、転送対象の変更後パケットの開始TCPシーケンスに基づく数値に設定する。これは、変更後パケットのACKを受信した際に「仕掛け」を発動させる（spring the “mousetrap”）のに用いられる。

【0089】

10

20

30

40

50

符号なし 3 2 ビット C A M エントリー：これは、D M A するパケットの連想メモリ（C A M）値を設定する。これは、a）パケット修正（touches up）時に用いる適切な T C P シーケンス；および、b）新しい数値で更新すべきデルタ／アキュムレータ；を固定口ジック 2 1 3 に伝えるのに用いられ得る。

【0 0 9 0】

変更後パケットの送信後、前述した C O N N _ C 2 M _ M O D _ P E N D I N G ブーリアンを設定し得る。

【0 0 9 1】

（遅延 A C K チェックのタイミング）

図 8 A 及び図 8 B は、C 2 M 遅延（Lazy）A C K チェック機能を詳細に説明したタイミング図である。

【0 0 9 2】

遅延 A C K チェックは、1）市場への変更後パケットの到達を保証し、2）（後述の欄で説明する）C P M を回避するのに用いられる。変更後パケットは、検査デバイス 1 1 0 と市場との間で廃棄されてしまう可能性がある。このリスクのため、再送してパケットの到達を保証するためのシステムが用意されているのが望ましい。

【0 0 9 3】

遅延 A C K チェック機能は、C O N N _ C 2 M _ M O D _ P E N D I N G ブーリアンが設定されたときにイネーブルされ得て、以下のように動作する：

【0 0 9 4】

パケット（ただし、当該パケットは、シーケンスマーカー以上のシーケンス番号を持つものであるとする）に対しての A C K を受信したか否かを確認するために、対応するセッションの全ての M 2 C パケットが調べられる。このパケットの A C K を受信すると、C O N N _ C 2 M _ M O D _ P E N D I N G ブーリアンがクリアされる。

【0 0 9 5】

固定ハードウェア 2 1 3 のデルタレジスタが、1 0 0 m s 以下の間隔（市場との間でのほぼ 1 往復（r t t）の間隔）でポーリングされる。好ましくは、このレジスタは、1 0 m s 毎に 1 回以下の間隔でポーリングされる。というのも、これは、時間計算量の観点からみて量の大きいコールだからである。このレジスタがクリアされていれば、（デルタがアキュムレータに加算されているので）変更後パケットの A C K を受信したということであり、C O N N _ C 2 M _ M O D _ P E N D I N G ブーリアンがクリアされ得る。

【0 0 9 6】

デルタがクリアされていなければ、変更後パケットが再送される。B u z z s a w でのパケット再送時には、F P G A がデルタを適用したことによってシーケンスが不正確とならないように、シーケンス番号を調節する必要がある。この時点では（still）F P G A のデルタがクリアされていない（waiting to be cleared）ため、再送されるパケットのシーケンスにこのデルタ値が誤って適用されてしまうのである。結果として、この調節を打ち消すために、再送前にソフトウェアが T C P シーケンス番号を調節する必要がある。

【0 0 9 7】

1 0 0 m s になると、セッションがタイムアウトして切断される。市場との間での 4 r t t のあいだ A C K が確認できない場合、A C K を受信することはないと予想され、セッションを切断するのが最善であるとされる。

【0 0 9 8】

A C K を受信した場合には、C O N N _ C 2 M _ M O D _ P E N D I N G ブーリアンがクリアされて、このコネクション内での遅延 A C K チェックシステムが非アクティブとされる。

【0 0 9 9】

遅延 A C K チェックは、相異なる 2 つのモジュールにより実現され得る。一方のモジュールは、コネクションマネージャ 7 1 0 内の C O N N _ A C T I V E 状態の上位で実行される（occurs at the top of）（図 7）。C O N N _ A C T I V E 状態は、フレーム方向

10

20

30

40

50

、CONN__MOD__PENDINGブーリアン、およびTCPシーケンスのACKをチェックする。他方のモジュールは、10ms再送時間をチェックするところ、これはコネクションマネージャ710内での別のプロセス又はスレッドで追跡され得る。

【0100】

キルされたパケットの到達を保証するためのC2Mパケット再送は、必ずしも必要でない。というのも、クライアントが、キルされたパケットを代わりに(for us)再送するからである。しかし、CPM(詳細については他の箇所でのCPMの説明を参照されたい)を回避するのに再送が必要となるため、C2M側の自動再送は維持され得る。

【0101】

(M2Cデータフロー：データ挿入の概要)

データ挿入は、C2Mメッセージが拒絶されたことに応答してエラーメッセージを注入するために意図されたM2C機能である。クライアント側サーバは市場に送信した注文の状態を把握する必要があるので、メッセージが拒絶されたことをクライアントに通知することは重要な機能である。クライアントに注文についての通知が行かない(fails to receive)と、クライアント側でキャンセルの嵐やアルゴリズムの混乱が典型的に起こり得る。

【0102】

M2Cエラー注入パケットは、当該パケットがM2C通知メッセージを含んでいることを特定するための複数のフィールド、および／または、拒絶を受けた(applied)メッセージの識別子(例えばクライアント注文ID等)、および／または、メッセージが拒絶された理由を示すエラーマスク、および／または、クライアントと市場装置との間で使用される上位プロトコルに依存する各市場プロトコル(例えばFIX等)固有のフィールドを含み得る。

【0103】

M2Cエラー挿入モジュールは、C2Mデータ削除が実行された後にコールされる。M2Cエラー挿入は、M2Cエラーメッセージをペイロードに追加する。このM2Cエラーメッセージは、拒絶されたクライアント注文IDを各市場プロトコル固有のメッセージ拒絶中に使用することで、どの特定のメッセージが拒絶されたのかをクライアントに知らせる。

【0104】

各拒絶毎に、一つのM2Cエラーメッセージが生成されるのが望ましい。一つのパケットが複数の拒絶を受けている場合、M2Cエラーパケットは各拒絶毎に一つのエラーメッセージを含むことになる。好ましくは、各エラーメッセージの長さは、キルされたメッセージのサイズよりも小さいのが望ましい。これにより、複数のM2Cエラーメッセージを連結して一つのパケットにすることができるので、1MTUよりも大きいM2Cエラーパケットが生成されるリスクが避けられる。結果として、キルされた各パケットにつき、M2Cエラーパケットを一つ送信するだけで済むことになる。

【0105】

再び図7を参照する。以下では、M2Cデータ挿入パケット構築モジュールの機能を詳細に説明する。

【0106】

変更後C2Mパケットが送信された直後に、M2Cエラー送信機能がコールされる。この機能は、以下のとおりである：

【0107】

パケットのIPヘッダ長を再計算して、注入されたバイトが含まれていることを確認する。

【0108】

新しいIPヘッダIDを設定する。このIDは、フラグメント化(断片化)されたパケットをTCPスタックで識別するのを支援するものである。ランダムなIDを挿入することにより、このエラー注入がフラグメント化されたパケットの一部であると受信側のロジックが解釈してしまう可能性が避けられる。

10

20

30

40

50

【0109】

DMAヘッダ内やDMAコールの必要な数値／ビットを設定する。数値の説明については、下記の内容を参照されたい。

【0110】

DMAヘッダ内のRaw（未変更）／Transform（変更）ビット：このDMAヘッダビットを「1」に設定することにより、ハードウェアはTCPチェックサム、CRCおよびTCPシーケンスを再計算する。

【0111】

DMAヘッダ内の符号付き12ビットデルタ：DMAヘッダ内のこのデルタを設定することにより、パケット内の変更されたバイト数を伝える（正のデルタ信号バイトは加算されたことを指し、負の信号バイトは減算されたことを指す）。FPGAは、この数値を用いて、特定されたCAMでの当該FPGAのハードウェアデルタレジスタを更新する。

10

【0112】

符号なし32ビットシーケンスマーカー：これは、シーケンスマーカーを、転送対象の変更後パケットの開始TCPシーケンスに基づく数値に設定する。これは、変更後パケットのACKを受信した際に「仕掛け」を発動させる（spring the “mousetrap”）のに用いられる。

【0113】

符号なし32ビットCAMエントリー：これは、DMAするパケットのCAM値を設定する。これにより、a) パケット修正（touches up）時に用いる適切なTCPシーケンス；および、b) 新しい数値で更新すべきデルタ／アキュムレータ；がFPGAに伝えられる。

20

【0114】

DMAヘッダ内のM2Cコネクションストールクリアビット：このDMAヘッダビットを「1」に設定することにより、M2Cコネクションストールがクリアされる。キルされたメッセージが最初にC2Mデータ削除モジュールで検出されると、M2C方向がコネクションストールモードに移行させられる。

【0115】

M2Cフラグメント化状態のトラック：パケットがキルされると、ソフトウェアはM2C方向をコネクションストールモードに設定する。そして、ソフトウェアは、M2C方向がフラグメント化状態でなくなったことを検出するまで、コネクションストールされたM2Cパケットを処理し続ける。当該コネクションのM2C方向がフラグメント化された2つのパケットの間であるとき、M2CエラーパケットはTCPストリームに注入されることができない。結果として、フラグメント化状態でなくなるまで、M2Cエラー送信機能はM2Cパケットを処理し続けるのが望ましい。M2Cエラーパケットは、全てキューに格納される。M2C方向がフラグメント化状態でなくなったことが確認されると、そのキュー内の全パケットが送信され得る。

30

【0116】

パケットの送信後、CONN__M2C__MOD__PENDINGブーリアンを設定する。

【0117】

（M2Cエラー再生）

M2Cエラーパケットがクライアントに送信されている最中にクライアントがセッションから切断されると、クライアントが注入メッセージを見逃す可能性がある。遅延ACKチェックは、同じコネクション内でしかパケットを再送しない。クライアントが再接続した際に、失われたM2Cエラーメッセージを前記検査デバイスから再生する方法が必要となる。

40

【0118】

これを行うために、毎回の（any）新たなコネクションの開始時に、これまでの過去（last）1000個などといった複数のM2Cエラー注入がクライアントに対して再生され得る。これらの注入は、ログイン承認がクライアントへとDMAされて且つCAMが精緻化

50

された直後に送信される。M2C パケットは、全ての注入が再生されるまでストールされたままとなる。クライアントがM2C エラー注入を見逃すのは切断時と考えられる (should) ので、M2C エラー再生は再接続後のログイン時にのみ行われる。

【0119】

この機能は、アカウントレベルの「再生オフ」許可を「真」に設定することでディセーブルされ得る。

【0120】

なお、M2C エラー再生は市場再生とは異なるものである。市場再生とは、市場からクライアントに送信されるメッセージの再生のことである。M2C エラー再生とは、前記検査デバイスによりM2C TCP ストリームに注入されたメッセージの再生のことである。

【0121】

(送信者へのM2C 送戻し)

(送信者への) M2C 送戻しモードとは、M2C エラーパケットが前記FPGAによってリングバッファ (ring) に再配備される機能である。M2C エラーパケットが送信された後、前記FPGAが当該パケットを前記リングバッファにコピーする。これにより、データパスは、当該パケットをパースし、当該パケットがM2C TCP ストリームへとどのシーケンスで挿入されたのかを決定し、メッセージ拒絶をログに報告することができる。

【0122】

(M2C 遅延ACK チェック)

M2C 遅延ACK チェックは、C2M 遅延ACK チェックと同じアーキテクチャで設計されている。これには、2つのモジュールが存在する：

【0123】

一方のモジュールは、注入されたM2C エラーパケットのACKを受信したか否かを確認するために、C2MのACKをチェックする。このチェックは、CONN_M2C_MOD_PENDING ブーリアンが設定されている場合に、CONN_ACTIVE 状態のコネクション状態マシンで行われる。

【0124】

他方の (second) モジュールは、それ自体が独立したプロセス又はスレッドである。この独立したプロセス又はスレッドは、CONN_M2C_MOD_PENDING が設定されている場合に、全てのコネクションオブジェクトについて10msのタイムアウトを追跡する。10msになると、M2C エラーパケットが再送される。100msになると、セッションが切断される。

【0125】

(C2M DSMハードウェアの詳細およびPTRSソフトウェアとのインターフェース)

図9に、固定ロジック (FPGA) 213 ハードウェア、ならびに当該固定ロジック (FPGA) 213 ハードウェアの、データパス及びPTRS アプリケーション220 ソフトウェアとの処理を示す。図9には、クライアントが誤りのあるパケットを送り出した状況に対処する際のさらなる詳細が描かれている。前述したように、このパケットはパケット検査エンジンにより処理されて、当該パケット検査エンジンによりエラーが検出されると共に、当該パケット検査エンジンによりこのパケットは拒絶されたパケットとしてマークされる。そして、このパケットは、PTRS ソフトウェア220へと送られて追加の処理を受けたり再送されたりする。

【0126】

クライアント→市場 (C2M) 方向の場合の処理は、以下のようになる：

【0127】

(1a) パケット検査エンジン (PIE) 420 が、各パケットのTCP/IP ヘッダを処理し、(例えば各パケットのソースIPアドレス及びポートならびに宛先IPアドレス及びポートを用いる等して) 各パケットのセッションを特定し、各パケットのTCP シーケンス番号を抽出する。特定された各セッションの最大TCP シーケンス値を追跡する

10

20

30

40

50

のに用いられる、特定された各セッション用の最新シーケンス番号 (`latest_seq_num`) レジスタが更新される。

【0128】

(1b) 前記P I Eがパケットを拒絶し、当該パケットがF P G A 2 1 3のカットスループスを完全に通過して (`completing`) 市場1 3 0へと向かおうとするのを阻止する。このパケットは、適切なキルの理由がDMAヘッダ内にマークされた状態で (例えばリングバッファを介して) P T R Sアプリケーション2 2 0ソフトウェアへと転送される。

【0129】

(1c) 拒絶されたパケットに対応するセッションについて、各セッション用のパケットスキップ阻止モードが設定される。当該パケットスキップモードがクリアされるまで、この特定のセッションの後続のあらゆるパケットは、F P G A 2 1 3を通過しないように阻止されると共に、後で送信できるように (`for transmission`) ソフトウェア2 2 0へと送られる。これは、例えば、パケットに関連付けられた情報を処理キュー、リングバッファ、またはF P G A 2 1 3とソフトウェア2 2 0との間で共有されたその他の適切なメモリもしくはストレージに送信することを含み得る。一部の実施形態では、パケットスキップ阻止モードが原因でソフトウェア2 2 0へとパケットが送られるたびにカウンタがインクリメントされ得る。パケットスキップ阻止モードが設定されていることが原因でソフトウェア2 2 0へと送られたパケットに関連付けられた情報をソフトウェア2 2 0が処理するたびに、当該ソフトウェア2 2 0は前記カウンタをデクリメントし得る。なお、パケットに関連付けられた情報を処理するとは、そのパケット (あるいは、元のパケットがルールに準拠しておらず変更が必要である場合には変更対象パケット) を送信することを含み得る。好適な実施態様では、パケットがソフトウェア2 2 0に送られることに対応した前記カウンタがクリアされたときに、パケットスキップ阻止モードがクリアされ得る。他の実施形態では、ソフトウェア2 2 0へと送られたパケットに関連付けられた全ての情報が処理されたか否かが、他の適切な方法で判断され得る。例えば、一部の実施形態では、カウンタではなく一対のポインタ又はインデックスが使用され得る。

【0130】

(1d) ソフトウェアが、前記拒絶されたパケットを処理し、拒絶されたコンテンツをトリミングする (切り取る)。次に、当該パケットのうちの拒絶されていない残りのコンテンツが前記F P G AへとDMAにより送り返される。DMA処理は、まずパケットペイロードを共有メモリに書き込んでから、次にF P G A 2 1 3の特定のレジスタをプログラムすることで構成され得る。これにより、前記F P G Aは、前記共有メモリに読出しを実行し、市場に送信すべき前記ペイロードを取り出す。DMAコマンドレジスタは、変更された (トリミングされた) コンテンツのサイズに相当する「デルタ」という新たなフィールドを含み得る。

【0131】

(1e) F P G A 2 1 3で、DMAされたパケットのT C Pシーケンス番号 (`seq`) が抽出される。このT C Pシーケンス番号は、対応するセッションについてパケットスキップモードをクリアすべきか否かを判断するのに利用される。

【0132】

(1f) 前記DMAコマンドレジスタに書き込まれたデルタ値を、C 2 MトラフィックのT C Pシーケンス番号を調節するのに用いられる各フロー用のT C Pアキュムレータ、およびM 2 CトラフィックのT C P A C K番号を調節するのに用いられる各フロー用のA C Kアキュムレータ (`ack_accumulator`) に適用する準備が整う。ただし、どちらのアキュムレータの状態も、影響を及ぼすパケットがパスを通過し終える (`passes through`) までは元の状態のままであるのが望ましい。

【0133】

(1g) 前記影響を及ぼすパケットのT C Pシーケンス設定 (`sequencing`) が、(デルタでアキュムレータが調節される前の) アキュムレータに保持された数値に従って調節される。変更後パケットが、当該パケットのネットワーク宛先に送信される。

【0134】

(1h) DMAされたパケットについてペンディングACK番号が算出されて、M2CパスのACKシーケンスマーカー(ack_seq_marker)へとコピーされる。M2Cパスでは、ACKシーケンスマーカーが、ACKアキュムレータへのACKデルタ(ack_delta)の適用をトリガする役割を果たす。

【0135】

(1i) 前記パケットが市場に送信された後、デルタがアキュムレータに適用される。

【0136】

なお、上記はFPGA213(固定ロジック)ハードウェアの例示的な実施態様の一つに過ぎず、変形例も可能であることを理解されたい。例えば、2つのアキュムレータ(C2MパスのアキュムレータとM2CパスのACKアキュムレータ)が図示されているが、これら2つのアキュムレータは単一のアキュムレータに置き換えられてもよい。一例として、アキュムレータへのデルタレジスタの数値の適用を、市場による変更後パケットの確認応答を受信するまで遅らせることにより、一方の(second)アキュムレータ及びデルタレジスタを省略することが可能である。よって、C2M方向に移動する後続のパケットのシーケンス番号の調節時には、アキュムレータの数値だけでなくデルタレジスタの数値も含めることになる。

【0137】

市場→クライアント(M2C)方向に流れるメッセージの場合の処理は、以下のようになる：

【0138】

(2a) 届けられた各パケットの受信ACKシーケンス(recv_ack_seq)番号が抽出されて、各パケットのセッションが特定される。この特定のパケットのフローにペンディングACKシーケンスマーカー(ack_seq_marker)が存在し、かつ、受信ACKシーケンスが当該ACKシーケンスマーカー以上である場合には、各フロー用のペンディングACKデルタがACKアキュムレータに適用される。

【0139】

(2b) ACKアキュムレータを用いて、通過するM2CパケットのACK番号が調節される。

【0140】

(ハードウェアでのM2C DSM)

図10には、M2C拒絶メッセージの処理の様子が描かれている。PTRSアプリケーション220ソフトウェアがメッセージを拒絶したばかりであるとして、好ましくは、次に当該PTRSアプリケーション220ソフトウェアが、対応する通知メッセージをクライアントに戻す向きで注入し、このイベントを(場合によっては、拒絶の理由も)当該クライアントに知らせる。これを整然に行うには、ソフトウェア220が影響中のセッションをM2Cコネクションストールモードに設定して、そのセッションのどのトラフィックもFPGA213を通過しないように阻止する。ソフトウェア220は、インターセプトしたトラフィックをDMAで市場130へと送信すると共に、拒絶メッセージを自然な(clean)境界で注入する。

【0141】

(1a) まず、PTRSアプリケーション220ソフトウェアが、拒絶メッセージを注入したいセッション(フロー)についてM2Cコネクションストールモードを設定する。

【0142】

(1b) FPGA213で、市場130から受信したパケットが処理されて、(例えば当該パケットのソースIPアドレス及びポートならびに宛先IPアドレス及びポートを用いる等して)当該パケットのセッションが特定される。コネクションストールモードに設定されたセッションにそのパケットが属していると特定されると、FPGA213は、当該パケットがカットスルーパスを介してクライアントに到達しないように当該パケットを拒絶し、当該パケットを後からソフトウェアによって送信できるように(for software t

10

20

30

40

50

ransmission) リングバッファに送る。

【0143】

(1c) PTRSアプリケーション220ソフトウェアが、前記リングバッファに入ってきたM2Cトラフィックを処理する。当該ソフトウェアは、コネクションストールの原因である(for)キルされたパケットを処理すると、懸案中の(pending)拒絶メッセージをパケットに注入して当該パケットをFPGA213を介して外部に(out)DMAする。PTRSソフトウェア220は、影響中のセッションのコネクションストールモードをクリアする(bh_clearビット)。

【0144】

(1d) 外部にDMAされるパケットの送信時に、コネクションストールモードがクリアされる。

【0145】

(1e) DMA_CMDに書き込まれたデルタ値を、M2CトラフィックのTCPシーケンス番号を調節するのに用いられる各フロー用のTCPアキュムレータ、およびC2MトラフィックのTCP ACK番号を調節するのに用いられる各フロー用のACKアキュムレータ(ack_accumulator)に適用する準備が整う。ただし、どちらのアキュムレータの状態も、影響を及ぼすパケットがパスを通過し終える(passes through)までは元の状態のままであるのが望ましい。

【0146】

(1f) 注入されたメッセージを含む変更後パケットが、前記リングバッファ(又は他のバッファ)に送られる。これは、拒絶メッセージをM2Cリングバッファに残しておいて後から再生できるように(for replays)するためである。

【0147】

(1g) 前記変更後パケットのTCPシーケンス設定(sequencing)が、(デルタを用いてアキュムレータに調節を行う前の)アキュムレータに保持された数値に従って調節される。この変更後イーサネットパケットが、当該パケットのネットワーク宛先に送信される。

【0148】

(1h) DMAされたパケットについてペンディングACK番号が算出されて、C2MパスのACKシーケンスマーカー(ack_seq_marker)へとコピーされる。C2Mパスでは、ACKシーケンスマーカーが、ACKアキュムレータへのACKデルタ(ack_delta)の適用をトリガする役割を果たす。

【0149】

(1i) 前記パケットが市場130に送信された後、デルタがアキュムレータに適用される。

【0150】

2. クライアント120が市場130に対してトラフィックを送信する。

【0151】

(2a) 届けられた各パケットの受信ACKシーケンス(recv_ack_seq)番号が抽出されて、各パケットのセッションが特定される。この特定のパケットのセッションにペンディングACKシーケンスマーカー(ack_seq_marker)が存在し、かつ、受信ACKシーケンス(recv_ack_seq)が当該ACKシーケンスマーカー以上である場合には、各フロー用のペンディングACKデルタがACKアキュムレータに適用される。

【0152】

(2b) ACKアキュムレータを用いて、通過するC2MパケットのACK番号が調節される。

【0153】

(変更前-過去パケットモード)

変更前-過去パケットとは、クライアント側又は市場側により再送されたパケットであ

10

20

30

40

50

って、TCPシーケンスを前記FPGAによる通常の方法で更新することができないパケットのことである。これは、その再送がかなり前のTCPシーケンスのものである場合に起こり得て、最新のアキュムレータやデルタではシーケンスの適切な変更を正確に反映することができない。結果として、TCP機能障害を防ぐための措置が講じられる必要がある。

【0154】

図11は、変更前－過去パケット状態の問題を表した一例を示す図である。

【0155】

この問題の背景 (setup) として、時刻 t_0 にて、第1の合格パケットが市場に送信される。しかし、時刻 t_1 では、第1の合格パケットのACKが検査デバイス110とクライアント120との間で廃棄されてしまい、クライアント120が予想時間枠内に第1の合格パケットのACKを受信できない様子が見て取れる。時刻 t_2 では、次のC2Mパケットである第1のキルパケットが検査デバイス110によりキルされて、かつ、バイト数の変化を考慮してデルタ値が「0」以外の数値（この例では「-10」）に設定される。時刻 t_3 では、第1のキルパケットのACKが検査デバイス110により確認されるものの、第1の合格パケットのACKと同じく、第1のキルパケットのACKも検査デバイス110とクライアントとの間で廃棄されてしまう。第1のキルパケットのACKが検査デバイス110によって既に確認済みであることから、デルタ値の「-10」がアキュムレータに加算されてしまう。

【0156】

ステップ t_4 にて、前記予想時間枠内に第1の合格パケットのACKを受信しなかったクライアントが当該第1の合格パケットを市場に対して再送したときに、問題が露わになる。というのも、アキュムレータがこのシーケンスにとって間違った数値に設定されているからである。このため、再送された第1の合格パケットのシーケンス番号がその間違ったアキュムレータ値で変更されてしまい、市場側でTCP機能障害が発生する可能性がある。つまり、パケットがキルされると、後続のパケットのシーケンス番号はアキュムレータによって調節されるのであるが、この例からも分かるように、キルすべきパケットが届いてから、当該キルすべきパケットよりも小さいシーケンス番号を持つ古いパケットが、検査デバイス110に後から届くという可能性がある。

【0157】

変更前－過去パケットは、クライアント側でも市場側でも起こり得る。この状態の検出は、コネクション状態マシンで行われる。この状態が原因で生じ得るTCP機能障害を回避するために、1) このパケットのTCPシーケンス番号が最新シーケンスマーカー (CONN_SEQ_MARKER) よりも小さく、かつ、2) MOD_PENDINGブーリアンが設定されていない場合に、変更前－過去パケット状態が検出され得る。

【0158】

変更前－過去パケットが検出された場合に取りべき考えられ得る (possible) 対策の一つは、セッションを切断することである。変更前－過去パケット（つまり、最新シーケンスマーカーよりも小さいシーケンス番号を持つ後から届いたパケット）を適切に処理する別の方法として、FPGA213が、デルタ値、アキュムレータ値及びシーケンス番号マーカー値の履歴テーブルを保持することが考えられる (may)。これにより、後から届いたパケットの正確なシーケンス番号が、当該後から届いたパケットの適切なシーケンスマーカー値に対応する、アキュムレータ及び／又はデルタの該当値 (relevant values) を、前記履歴テーブルから適用することによって決定されることが可能になる。

【0159】

M2Cエラー注入フラグメント：一実施形態において、PTRSソフトウェア220は、M2C方向がフラグメント化状態でなくなったことが確認できるまでM2Cエラー注入を保存する。ただし、フラグメント化されたM2Cパケットを均等なメッセージ境界で分割することでM2C方向をストールする必要性をなくすることが可能な、他の実施形態も考えられ得る (possible)。エラー注入時にM2Cフラグメントを除去することにより、エ

10

20

30

40

50

ラーパケットの迅速な送信を実行することができる。

【0160】

(TCPシーケンスマーカーと相互作用するTCPシーケンス番号)

一部の実施形態では、新しいメッセージが届いた際に、固定ロジック213がこの新しいメッセージのTCPシーケンス番号をシーケンスマーカーレジスタの数値と比較し得る。この最新メッセージのTCPシーケンス番号がシーケンスマーカーの数値よりも小さい場合、当該メッセージは変更されずに送信される。この最新のメッセージのTCPシーケンス番号がシーケンスマーカーの数値以上である場合には、当該メッセージのTCPシーケンス番号にアキュムレータレジスタの数値が加算される。

【0161】

(パケットスキップ阻止および高速パス—対—低速パスの使用)

パケットがキルされると、CAM (より広義的には検査デバイス110) がパケットスキップ阻止モードに設定される。パケットスキップ阻止モードは、全てのクライアント→市場 (C2M) パケットを、初めに固定ハードウェア213で処理することなくソフトウェアへと送る。

【0162】

これにより、一部の稀な (race) 状態が回避される。図12に、このような稀な状態の一つを示す。

【0163】

図12に示すケースでは、時刻 t_0 にて、第1のキルパケット (k_1) が送信される。しかし、後の時刻 t_1 では、第1のキルパケット (k_1) を処理するための時間をソフトウェア (PTRSアプリケーション220) が確保する前に、第1の合格パケット (g_1) がハードウェア (FPGA) を通過する (この場合、第1の合格パケット (g_1) も、PTRSアプリケーションソフトウェア220による処理を必要とする)。これは、第1のキルパケット (k_1) を上書きするための時間が確保される前に、第1の合格パケット (g_1) が市場に到達し得ることを意味する。さらに、第1の合格パケットのシーケンス番号は、第1のキルパケットを完全に処理するために必要な変更を正確に反映していないため、不正確となり得る。

【0164】

このような稀な状態を回避するために、パケットスキップ阻止モードでは、全てのパケットがソフトウェアに送られるのだが、ソフトウェアに送られた最後のパケットが送信され終わるまでこれが続けられる。つまり、時刻 t_2 にて、パケットスキップ機能がイネーブルされていれば、前記FPGAが単独で且つ自動的にパケットのスキップを実行する。これは、ソフトウェアでオン・オフされる必要がない。

【0165】

図13に、これの一例を示す。時刻 t_0 にて、第1のパケット k_1 がキルされると、時刻 t_1 や時刻 t_2 で受信するパケットのように、後続の全てのパケットが前記FPGAではなくPTRSソフトウェアへと送られる。これは、ソフトウェアに送られた最後の合格パケット g_2 が時刻 t_3 にて送信されるまで続く。その後、パケットスキップモードがオフにされて、時刻 t_4 では、合格パケット g_3 が先程まで (as previously) とは違ってソフトウェアに入らずともハードウェアを通過して送信されることが可能となる。

【0166】

(近接変更 (CPM))

近接変更 (CPM) とは、キルされたパケットのACKを市場から未だ受信していないにもかかわらず、別のキルされたパケットが送信を待っているときに起こる稀な状態のことである。この場合、デルタは最初にキルされたパケットから除去された総バイト数を未だ記憶していて、このパケットのACKが届くまでは当該バイト数をアキュムレータに加算したりクリアしたりすることができない。CPMが発生すると、その次の (second) キルされたパケットは、送信可能となるまでストールされる必要がある。最初にキルされたパケットのACKを受信すれば、前記その次のキルされたパケットが送信可能となる。一

10

20

30

40

50

部の実施形態では、キルされた後続のパケットをストールするのに代えて、キューに格納することが可能であり得る。

【0167】

図14に、近接変更(CPM)のシナリオを示す。

【0168】

CONN_C2M_MOD_PENDINGブーリアンが(時刻t0にて)既に設定されているうえで、キルされたC2Mパケットを送信する場合、C2M送信モジュールが以下のことを行う：

【0169】

1) 上記状態が検出されると、時刻t2のようにコア全体をストールする。ここでは、受信した後続の第2のキルパケットが前記検査デバイスでストールされる(なお、時刻t2にて第2のキルパケットを第1のキルパケットと近接して(すなわち、第1のキルパケットのACKを受信する前に)受信したときに、近接状態が検出される。時刻t1の第1の合格パケットはハードウェアにより処理されて市場に送信されるのに対し、後続のパケットは第1のキルパケットのACKを受信するまでストールされる必要がある)。

【0170】

2) CPMが発生したセッションのデルタレジスタを継続的にポーリングする：

時刻t3にてデルタレジスタが「0」になると、第1のキルパケットのACKを受信したということである。この時点で前記ストールが解除され得て、時刻t4にて第2のキルパケットが送信され得る。デルタレジスタが未だ「0」でなければ、最後に変更したパケット(例えば、この例では第1のキルパケット)を決まった間隔で(例えば、10ms毎に)再送する。最後に変更したパケットのACKをタイムアウト時間(例えば、100ms)を過ぎても未だ受信できなければ、セッションを切断する。

【0171】

現在、検査デバイス110で取引がアクティブになっている最中にCPMが発生することは極めて稀なケースであると考えられる。この理由から、CPMの検出時にコア全体をストールすることが可能である。

【0172】

重要なのは、CPMの数がハードウェアのシーケンスマーカーレジスタの数を上回った場合に、前述したCPMの処理方法がソフトウェアを利用するという点である。一部の実施形態では、ハードウェアパスにて複数のCPMに対処できるように、シーケンスマーカーレジスタ、デルタレジスタ及びアキュムレータレジスタが複数存在し得る。例えば、シーケンスマーカーレジスタ、デルタレジスタ及びアキュムレータレジスタのセットを3セット有する一実施形態では、確認応答が済んでいない変更後パケットに一度に3つ対処することができる。例示的なこのような実施形態では、第1の変更後パケットの確認応答を未だ受信していないにもかかわらず第2のパケットを変更する必要がある場合に、この第2の変更対象パケットに対応するデルタ値及び／又はシーケンスマーカー及び／又はアキュムレータ値が、第2のセットのレジスタに記憶され得る。同様に、先行する(first)2つの変更後パケットが確認応答される前に第3のパケットの変更が必要となった場合、第3のセットのレジスタにデルタ値及び／又はシーケンスマーカー及び／又はアキュムレータ値が記憶され得る。そして、先行する(prior)3つの変更後パケットが確認応答される前に第4のパケットの変更が必要となった場合、当該第4のパケットはソフトウェアで処理され得るか、あるいは前述のようにストールされ得る。つまり、このような実施形態では、変更が必要なパケットが届いたとしても、変更が必要なこの新たに届いたパケットと対応させるのに利用可能なシーケンスマーカーレジスタ及び／又はデルタレジスタ及び／又はアキュムレータレジスタのセットが残っている限り、ストールを回避することができる。所与のセットのシーケンスマーカーレジスタ及び／又はデルタレジスタ及び／又はアキュムレータレジスタと対応させた変更パケットの確認応答を受信すると、確認応答を受けた当該変更後パケットに対応付けられたこれらのレジスタの数値がクリアされ得て、変更が必要な別の流入パケットにより再利用できるように(for reuse)使用可能な状態

10

20

30

40

50

となり得る。

【0173】

(ハードウェアメカニズムの、各双方向セッション毎のインスタンス)

本文献で説明するハードウェアメカニズムには、各セッション毎に個別のインスタンスペアが存在する。当該ペア内の一方のインスタンスはTCPセッションのC2M方向を処理し、他方のインスタンスは同じTCPセッション (the TCP session) のM2C方向を処理する。典型的なFPGAは、実際128個のセッション (このメカニズムの256個のインスタンス) をサポートするが、現実的な数はその時々 (arbitrary) であり、FPGAチップ上のメモリのみに限られる。

【0174】

(TCPパケットの変更の選択肢)

本文献で説明する技術により、C2M方向やM2C方向で、確立されたTCPストリームからメッセージを完全に除去したり、既存のメッセージである「インフライト (in-flight)」メッセージを別のサイズのメッセージに置き換えたり、追加のメッセージを注入したりする、カットスルー方式の超低レイテンシな装置を提供することが可能となる。

【0175】

(市場拒絶メッセージの合成)

本文献で説明する技術により、既存のTCPストリームでのメッセージの除去、追加及び変更が可能となる (accommodate)。これらの低次メカニズムは、リスクチェックに合格しなかった取引イベントメッセージを除去するのに用いられ、データが宛先の取引所に到達するのを防ぐ。そして、取引メッセージが拒絶されたクライアントには、自身のメッセージが取引所により受信されなかったことを把握できるように拒絶通知が必要となる。本技術は、既に取り所プロトコルに合わせてコーディングされたクライアント取引サーバがその拒絶をシームレスに解釈できるように、各プロトコル固有の拒絶メッセージを合成する。取引クライアントにとっては、取引所がメッセージを取引前リスクコントロールの違反で拒絶したかの如く見える。さらに、システムは、元々の取引メッセージが拒絶された確かな理由をクライアントが理解 (interpret) できるように、特定の拒絶理由に関する追加情報を付け足す。

【0176】

(FPGAリングバッファ)

一部の実施態様では、FPGA (固定ハードウェア) 213とPTRSソフトウェア220との間の主要な通信メカニズムの一つが、リングバッファであり得る。このリングバッファは、ヘッドポインタとテールポインタとの2つのポインタにより維持される。FPGA 213は、全データを当該バッファへと、前記テールポインタの位置を開始位置として書き込む。FPGA 213がデータの書き込みを完了すると、当該FPGA 213は最後に書き込んだバイトのメモリ位置へと前記テールポインタを進める。ソフトウェアが当該バッファから読出しを行った際には、最後に読み出されたバイトの位置へと前記ヘッドポインタが進められる。これにより、メモリをFPGA 213によって安全に上書きすることが可能となる。有効データの開始メモリ位置を追跡する前記ヘッドポインタ、および当該有効データの終了メモリ位置を追跡する前記テールポインタは、当該リングバッファのベースアドレスから当該ベースアドレスにバッファサイズを足したアドレスにかけて「一周」することで、「環状リング」を形成する。FPGA 213が当該リングバッファをオーバーフローさせて、「一周」事象の発生を経た、前記テールポインタが前記ヘッドポインタを越えて上書きするようなことが起こらないための配慮が講じられる。

【0177】

(他の用途例)

これまでに述べたアーキテクチャは、他の用途に利用されてもよい。例えば、当該アーキテクチャは、ネットワーク上を流れるデータストリームを監視したり、パケットをキャプチャ (採取) したり、パケットの生データを復号化したり、証券売買注文を検査する以外の理由によりリアルタイムでパケットコンテンツを解析したりするのに利用され得るこ

10

20

30

40

50

とも可能である。

【0178】

(実装(実現)のさらなる選択肢)

なお、これまでに述べた例示的な実施形態は、数多くの様々な方法で実現され得ることを理解されたい。一部の例では、各種「データプロセッサ」が、それぞれ、中央演算処理装置、メモリ、ディスク又は他の大容量記憶装置、少なくとも1つの通信インターフェース、少なくとも1つの入出力(I/O)装置、および他の周辺装置を備えた物理的又は仮想的な汎用コンピュータにより実現され得る。当該汎用コンピュータは、プロセッサへと変換して考えることで、これまでに述べたプロセスを実行する。これは、例えば、ソフトウェア命令を当該プロセッサにロードした後、当該命令を実行して前述の機能を実現すること等により行われる。

10

【0179】

当該技術分野で知られているように、そのようなコンピュータは、コンピュータ又は処理システムの構成要素(components)間でのデータ伝送に利用される一連のハードウェアラインであるシステムバスを備え得る。当該バスは、本質的に、コンピュータシステムの相異なる構成要素(例えば、プロセッサ、ディスクストレージ、メモリ、入出力ポート、ネットワークポート等)同士を接続して当該構成要素間での情報の伝送を可能にする共有の導管である。当該システムバスには、コンピュータ命令を実行する少なくとも1つの中央演算処理装置が取り付けられている。当該システムバスには、さらに、様々な入出力装置を接続するための入出力装置インターフェースが典型的に取り付けられている。少なくとも1つのネットワークインターフェースが、ネットワークに取り付けられた様々な他の装置へと接続することを可能にする。メモリは、一実施形態を実現するように用いられるコンピュータソフトウェア命令及びデータを記憶する揮発性の記憶部である。ディスク又は他の大容量記憶装置は、例えば本明細書に記載された様々なプロシージャ(手続き)を実現するように用いられるコンピュータソフトウェア命令及びデータを記憶する不揮発性の記憶部である。

20

【0180】

つまり、実施形態は、典型的に、ハードウェア、カスタム設計された半導体ロジック、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、ファームウェア、ソフトウェア、またはこれらの組合せで実現され得る。

30

【0181】

一部の実施形態において、本明細書に記載されたプロシージャ、装置およびプロセスは、コンピュータプログラムプロダクトである。当該コンピュータプログラムプロダクトは、システム用のソフトウェア命令の少なくとも一部を提供するコンピュータ読取り可能媒体(例えば、少なくとも1つのDVD-ROM、CD-ROM、ディスク、テープなどの取外し可能な記憶媒体等)を含む。このようなコンピュータプログラムプロダクトは、当該技術分野において周知である任意の適切なソフトウェアインストール方法によってインストールされることが可能なものであり得る。また、他の実施形態では、前記ソフトウェア命令の少なくとも一部が、ケーブルおよび/または通信および/または無線接続を介してダウンロードされるものであり得る。

40

【0182】

また、実施形態は、非過渡的な機械読取り可能媒体に記憶された、少なくとも1つのプロシージャにより読出し・実行され得る命令としても実現され得る。非過渡的な機械読取り可能媒体には、情報を機械(例えば、コンピューティングデバイス等)が読取り可能な形態で記憶したり送信したりするあらゆるメカニズムが含まれ得る。例えば、非過渡的な機械読取り可能媒体は、読出し専用メモリ(ROM)、ランダムアクセスメモリ(RAM)、磁気ディスク記憶媒体を含むストレージ、光記憶媒体、フラッシュメモリ装置などを含み得る。

【0183】

また、本明細書では、ファームウェア、ソフトウェア、ルーチンまたは命令が、特定の

50

処理および／または機能を実行するかの如く述べられているかもしれない。しかしながら、本明細書に含まれるこのような説明は便宜上のものに過ぎず、実際のところこのような処理が、コンピューティングデバイス、プロセッサ、コントローラ又はその他の装置がファームウェア、ソフトウェア、ルーチン、命令などを実行することで生じているという点を理解されたい。

【0184】

また、ブロック図やネットワーク図は、構成要素が追加されても省略されてもよいし、配置が変更されてもよいし、表現（実現）の仕方が別のものに変更されてもよい。ただし、実現の形態によってはブロック図やネットワーク図が決まっている場合があり、実施形態の実行を表すブロック図やネットワーク図の数が特定の様式で実現されることになるという点も理解されたい。

【0185】

よって、実施形態は、その他にも、多種多様なコンピュータアーキテクチャ及び／又は物理的なコンピュータ及び／又は仮想的なコンピュータ及び／又はクラウドコンピュータ及び／又はこれらの数種類の組合せで実現されてもよい。つまり、本明細書で説明したコンピュータシステムは、例示目的のものに過ぎず、実施形態を限定するものではない。

【0186】

このように、本発明を例示的な実施形態を参照しながら具体的に図示・説明してきたが、当業者であれば、添付の特許請求の範囲に包含される本発明の範囲を逸脱しない範囲で、形態や細部に様々な変更が施されてもよいことを理解するであろう。

なお、本発明は、実施の態様として以下の内容を含む。

〔態様1〕

第1のエンドポイントと第2のエンドポイントとの間に確立された双方向ネットワークセッション中に、

パケットのストリームを検査する過程と、

前記第1のエンドポイントから送信された所与のパケット内のメッセージが変更されるべきであると判断すると、

前記所与のパケット内の前記メッセージを変更して変更後パケットを生成する過程と、

前記メッセージの変更を考慮して変更後のシーケンス番号を決定する過程と、

前記変更後パケットを前記第2のエンドポイントに送信する過程と、

返信メッセージを前記第1のエンドポイントに送信し、元のメッセージが変更されたことを知らせる過程と、

を備える、方法。

〔態様2〕

態様1に記載の方法において、変更後のシーケンス番号を決定する前記過程が、

変更された前記パケットのシーケンス番号に基づいてシーケンスマーカ値を決定すること、

アキュムレータ値を維持すること、

元の前記所与のパケットと前記変更後パケットとのバイト数の差分を表すデルタ値を記憶すること、

前記第2のエンドポイントからの前記変更後パケットの確認応答を検出すること、および

前記デルタ値を前記アキュムレータ値に加算すること、

を含む、方法。

〔態様3〕

態様2に記載の方法において、さらに、

受信した後続のパケットのシーケンス番号が前記シーケンスマーカ値以上である場合、

前記第1のエンドポイントから受信した後続のパケットのシーケンス番号を前記アキュムレータ値だけ変更する過程、

を備える、方法。

〔態様4〕

10

20

30

40

50

態様 3 に記載の方法において、さらに、

前記第 2 のエンドポイントから受信した別のパケットの確認応答シーケンス番号を前記アキュムレータ値に基づいて調節する過程、
を備える、方法。

〔態様 5〕

態様 1 に記載の方法において、前記所与のパケット内の前記メッセージを変更する過程が、当該パケット内の当該メッセージを、異なる長さの第 2 のメッセージに置き換えることを含む、方法。

〔態様 6〕

態様 5 に記載の方法において、前記第 2 のメッセージが、ハートビートメッセージ、テストメッセージ、またはシーケンスフィルタメッセージを含む、方法。

10

〔態様 7〕

態様 1 に記載の方法において、前記所与のパケット内の前記メッセージを変更する過程が、当該パケットから当該メッセージを除去することを含む、方法。

〔態様 8〕

態様 1 に記載の方法において、前記第 1 のエンドポイントが取引業者に対応付けられたクライアント装置であり、前記第 2 のエンドポイントが証券取引システムに対応付けられた市場装置であり、前記メッセージが証券を売買する注文である、方法。

〔態様 9〕

態様 8 に記載の方法において、前記返信メッセージが、前記メッセージを変更した理由を含む、方法。

20

〔態様 10〕

態様 8 に記載の方法において、前記メッセージを変更した前記理由が、取引ルールへの準拠、セキュリティリスク又は証券リスク、及びパケットコンテンツフィルタリングのうちの少なくとも一つに関する、方法。

〔態様 11〕

態様 2 に記載の方法において、さらに、

前記第 2 のエンドポイントから前記変更後パケットの A C K を受信したか否かを判断する過程と、

前記変更後パケットの A C K を受信していないと判断した場合、

30

前記変更後パケットを再送する過程と、

タイムアウト時間後に前記セッションを切断する過程と、

を備える、方法。

〔態様 12〕

態様 11 に記載の方法において、前記変更後パケットの A C K を受信したか否かを判断する過程が、一定間隔で実行される、方法。

〔態様 13〕

態様 11 に記載の方法において、さらに、

前記第 2 のエンドポイントから受信した前記 A C K のシーケンス番号が前記シーケンスマーカー以上である場合、

40

変更ペンディング値をクリアする過程、

を備える、方法。

〔態様 14〕

態様 11 に記載の方法において、変更された前記メッセージの A C K を受信したか否かを判断する過程が、前記デルタ値を定期的にポーリングすることを含む、方法。

〔態様 15〕

態様 1 に記載の方法において、さらに、

前記変更後パケットの A C K を未だ受信しておらず、かつ、後続のパケットについても変更して第 2 の変更後パケットを生成する必要がある場合、

前記変更後パケットの前記 A C K を受信するまで前記第 2 の変更後パケットをストール

50

する過程、

を備える、方法。

〔態様 16〕

態様 15 に記載の方法において、さらに、

前記デルタ値をポーリングして、前記変更後パケットの前記 A C K をいつ受信したかを判断する過程と、

所定の時間間隔で前記変更後パケットを再送する過程と、

前記変更後パケットの前記 A C K を受信することなく所定の時間が経過すると、前記セッションを切断する過程と、

を備える、方法。

〔態様 17〕

第 1 の T C P エンドポイントと第 2 の T C P エンドポイントとの間に確立された双方向 T C P セッション中に、

前記第 1 の T C P エンドポイントから受信した T C P パケットを検査する過程と、

所与の T C P パケットから T C P シーケンス番号を抽出する過程と、

前記所与の T C P パケットに対応する T C P セッションを特定する過程と、

前記所与の T C P パケット内のコンテンツがルールに準拠しているか否かを判断する過程と、

前記所与の T C P パケットが前記ルールに準拠している場合、

前記所与の T C P パケットを前記第 2 の T C P エンドポイントに到達させる過程と、

前記所与の T C P パケットが前記ルールに準拠していない場合、

拒絶理由データ値を決定する過程と、

前記所与の T C P パケットの少なくとも一部が前記第 2 の T C P ポイントに到達するのを阻止する過程と、

前記所与の T C P パケットの少なくとも一部を除去して変更後 T C P パケットを形成する過程と、

デルタ値を、そのようにして除去された前記一部のサイズに依存する数値に設定する過程と、

前記デルタ値を、前記 T C P セッションに対応付けられた各セッション用の T C P アキュムレータに適用する過程と、

前記変更後 T C P パケットの T C P シーケンス番号を、前記各セッション用の T C P アキュムレータのサイズに依存する数値だけ調節する過程と、

前記変更後 T C P パケットを前記第 2 の T C P エンドポイントに送信する過程と、

前記拒絶の理由を表す数値を有する返信パケットを生成する過程と、

前記返信パケットを前記第 1 の T C P エンドポイントに送信する過程と、

を備える、方法。

〔態様 18〕

互いに独立して動作する第 1 のネットワークインターフェース及び第 2 のネットワークインターフェースを有するカットスルー方式の固定ロジックデバイスであって、1つ又は複数の T C P パケットストリームが前記第 1 および第 2 のネットワークインターフェース間を移動することのできる、カットスルー方式の固定ロジックデバイスと、

前記 T C P パケットストリームを検査するパケット検査デバイスであって、

所与の T C P パケットから T C P シーケンス番号を抽出し、

前記所与の T C P パケットに対応するセッションを特定し、

前記所与の T C P パケットが、ルールに準拠していないコンテンツを含んでいるか否かを判断し、

前記所与の T C P パケットが準拠していない場合、

前記所与の T C P パケットの少なくとも一部が前記第 2 のネットワークインターフェースに到達するのを阻止する、

パケット検査デバイスと、

10

20

30

40

50

固定ロジック及び少なくとも1つのプログラマブルなデータプロセッサを含むコントローラであって、

前記パケット検査デバイスから前記所与のTCPパケット及び拒絶理由データ値を受け取り、

前記所与のTCPパケットの前記シーケンス番号に基づく数値を、各セッション用のTCPシーケンスマーカーとして記憶し、

前記TCPパケットから、拒絶された前記コンテンツをトリミングして、変更後TCPパケットを形成し、

前記変更後パケットを前記第2のネットワークインターフェースを介して送信できるように、当該変更後パケットを、前記カットスルー方式の固定デバイスと共有するメモリに書き込み、

各セッション用のデルタ値を、トリミングされた前記コンテンツのサイズに依存する数値に設定し、

前記変更後TCPパケットの前記TCPシーケンス番号を、各セッション用のTCPアキュムレータの数値に従って調節し、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットの、シーケンス番号及びセッション識別子を抽出し、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットのACKシーケンス番号が、前記対応するセッションのTCPシーケンスマーカー以上である場合、各セッション用の対応付けられた前記デルタ値を、当該セッションに対応付けられた前記TCPアキュムレータに適用し、

前記第2のネットワークインターフェースを介して届いた前記後続のパケットの前記ACKシーケンス番号を、前記セッションに対応付けられた前記TCPアキュムレータの数値だけ調節し、

前記拒絶の理由を表す数値を含む、前記第1のネットワークインターフェースを介して送信する返信パケットを生成する、

コントローラと、

を備える、装置。

〔態様19〕

態様18に記載の装置において、前記コントローラが、さらに、

ルールに準拠していない前記所与のパケットに対応する前記セッションについて、コネクションストールモードを設定し、

前記第2のネットワークインターフェースから受信した後続のパケットが、前記コネクションストールモードを設定した前記セッションに属するものである場合、

生成された前記返信パケットを前記第1のインターフェースに転送し、

前記生成された返信パケットを前記第1のインターフェースに転送した後、前記後続のパケットを前記第1のインターフェースに転送し、

前記セッションの前記コネクションストールモードをクリアする、装置。

〔態様20〕

態様18に記載の装置において、前記コントローラの前記固定ロジックが、

前記パケット検査デバイスから前記所与のTCPパケット及び拒絶理由データ値を受け取る手順、

前記所与のTCPパケットの前記シーケンス番号に基づく数値を、各セッション用のTCPシーケンスマーカーとして記憶する手順、

前記変更後TCPパケットの前記TCPシーケンス番号を、各セッション用のTCPアキュムレータの数値に従って調節する手順、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットの、シーケンス番号及びセッション識別子を抽出する手順、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットのACKシーケンス番号が、前記対応するセッションのTCPシーケンスマーカー以上である場

10

20

30

40

50

合、各セッション用の対応付けられた前記デルタ値を、当該セッションに対応付けられた前記TCPアキュムレータに適用する手順、ならびに

前記第2のネットワークインターフェースを介して届いた前記後続のパケットの前記ACKシーケンス番号を、前記セッションに対応付けられた前記TCPアキュムレータの数値だけ調節する手順、

のうちの少なくとも一つを実行する、装置。

〔態様21〕

態様20に記載の装置において、前記プログラマブルなデータプロセッサが、プログラムコードを実行して、

前記TCPパケットから、拒絶された前記コンテンツをトリミングして、変更後TCPパケットを形成する手順、

前記変更後パケットを、前記カットスルー方式の固定デバイスと共有するメモリに書き込む手順、

各セッション用のデルタ値を、トリミングされた前記コンテンツのサイズに依存する数値に設定する手順、

前記第2のネットワークインターフェースを介して届いた後続のTCPパケットの、シーケンス番号及びセッション識別子を抽出する手順、ならびに

前記拒絶の理由を表す数値を含む、前記第1のネットワークインターフェースを介して送信する返信パケットを生成する手順、

のうちの少なくとも一つを実行する、装置。

〔態様22〕

態様18に記載の装置において、前記コントローラが、さらに、

ルールに準拠していない前記所与のTCPパケットに対応する前記セッションについて、阻止モードインジケータ値を設定し、

前記第1のネットワークインターフェースを介して届いた後続のパケットに関連付けられている情報を、処理キューに格納し、

前記処理キュー内の、前記パケットに関連付けられている前記情報が処理されたか否かを判断し、

前記処理キュー内の、前記パケットに関連付けられている前記情報が処理されたと判断すると、前記阻止モードインジケータをクリアする、装置。

10

20

30

40

50

【図面】
【図 1】

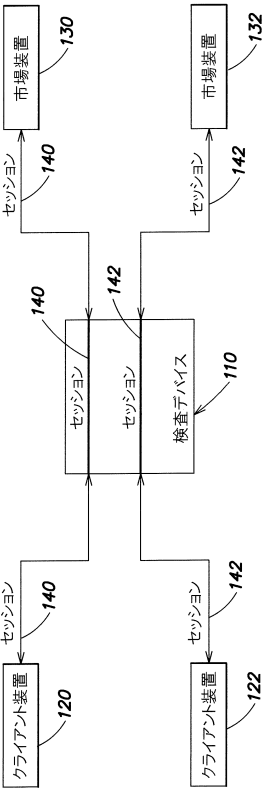


FIG. 1

【図 2】

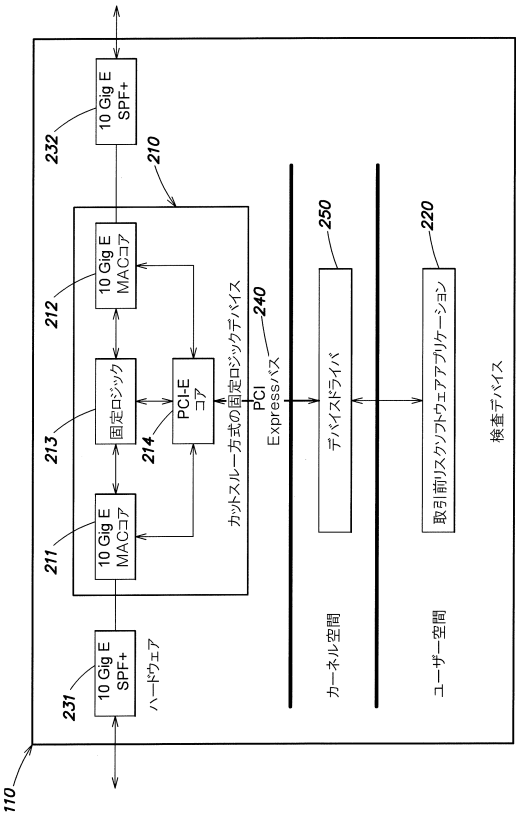


FIG. 2

【図 3 A】

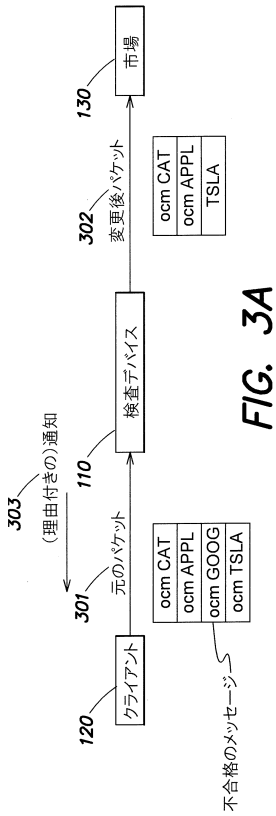


FIG. 3A

【図 3 B】

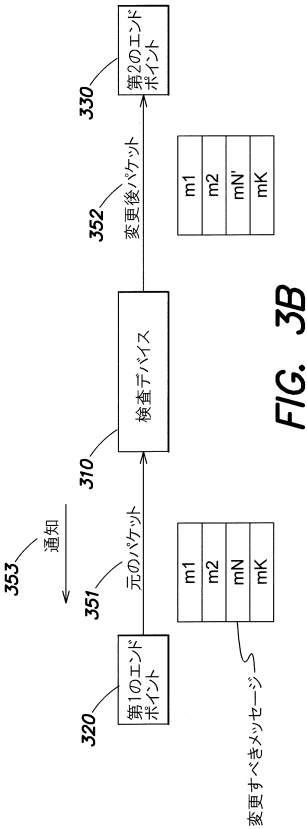


FIG. 3B

【図 4】

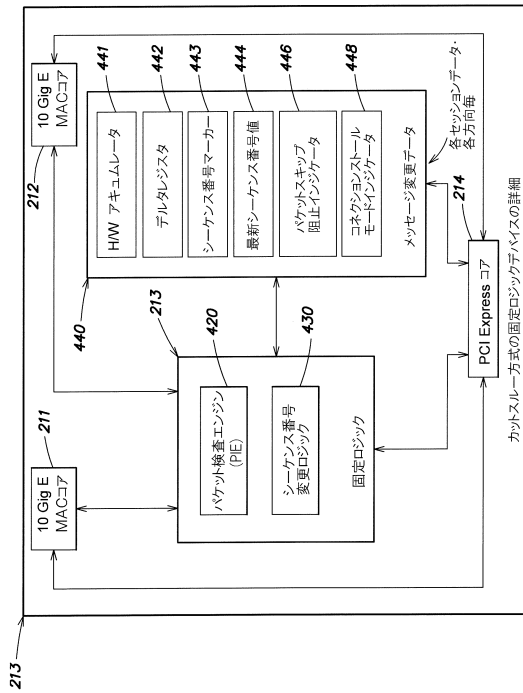


FIG. 4

【図 5】

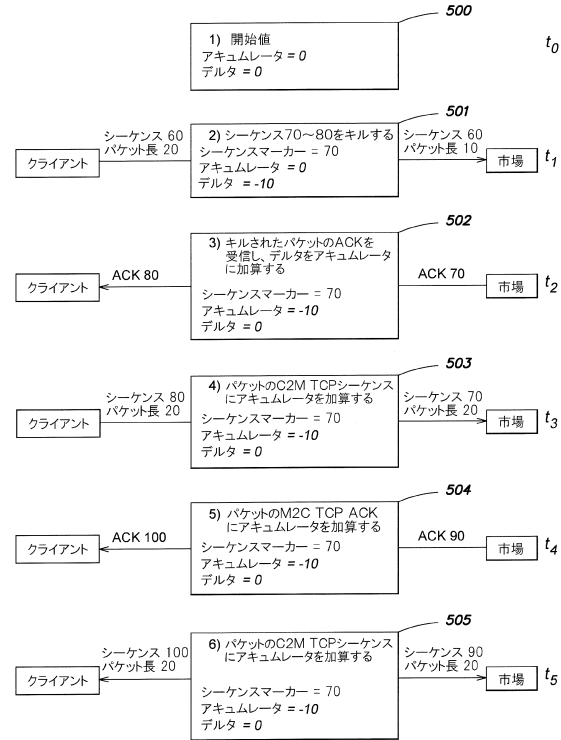


FIG. 5

【図 6】

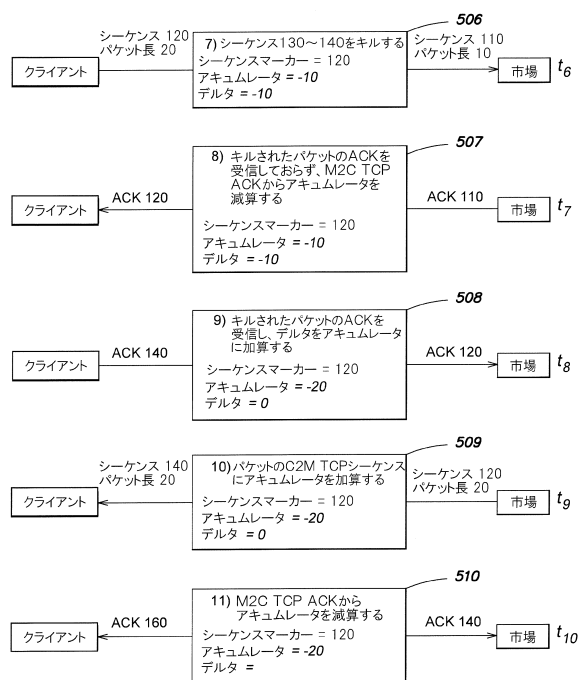


FIG. 6

【図 7】

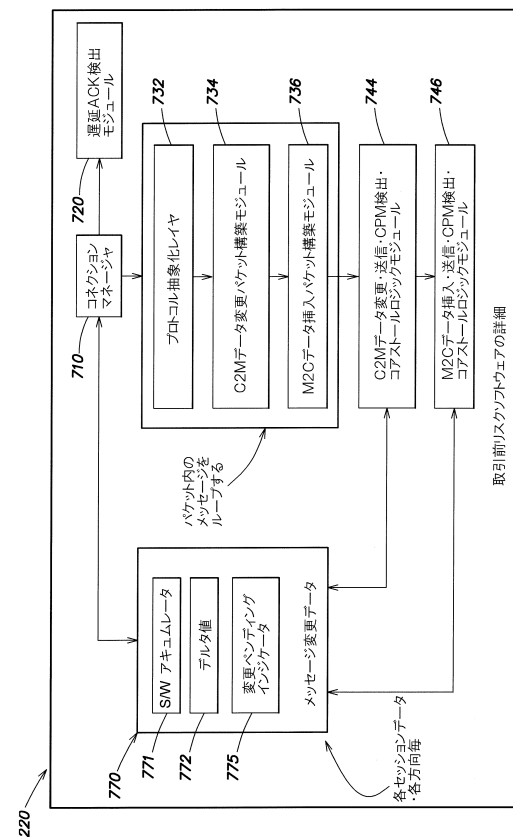


FIG. 7

10

20

30

40

50

【図 8 A】

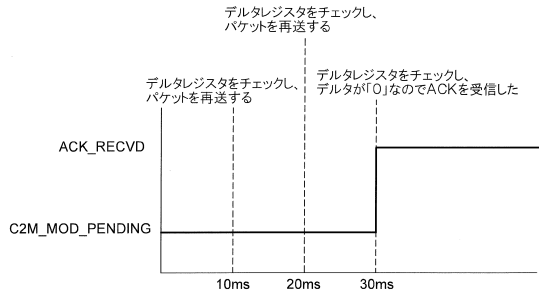


FIG. 8A

【図 8 B】

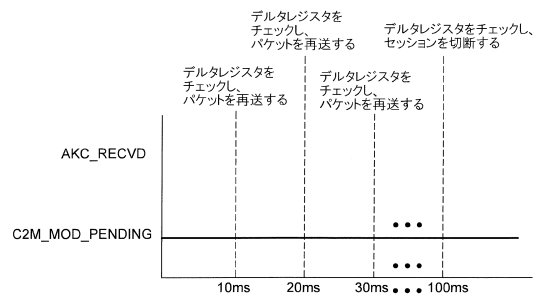


FIG. 8B

【図 9】

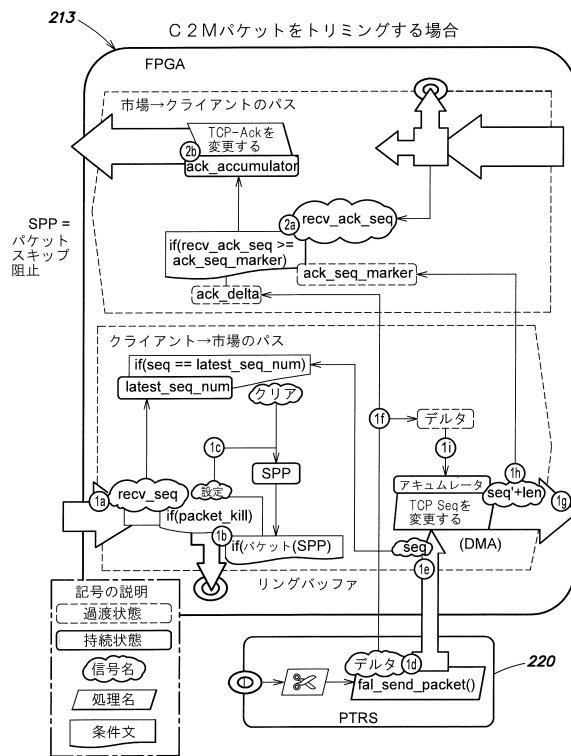


FIG. 9

【図 10】

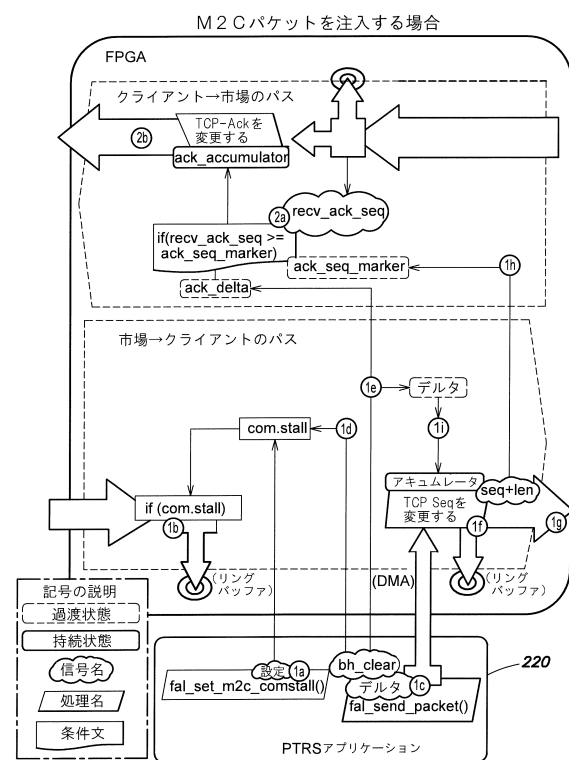


FIG. 10

10

20

30

40

50

【図 1 1】

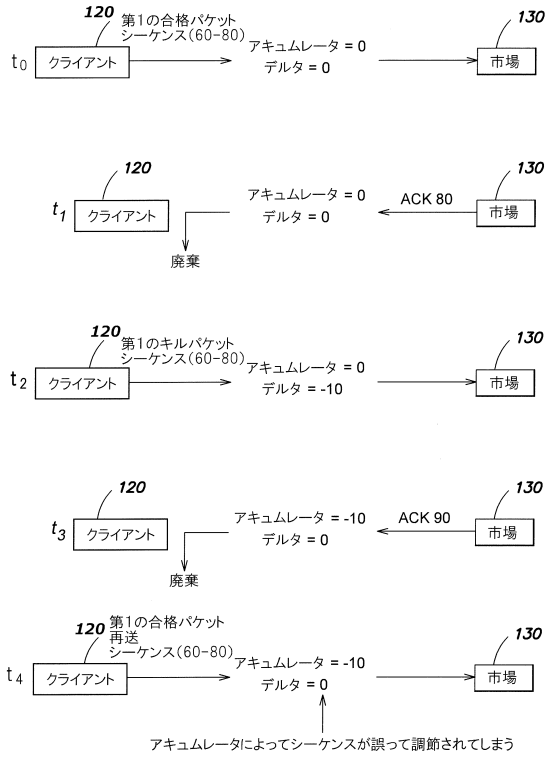


FIG. 11

【図 1 2】

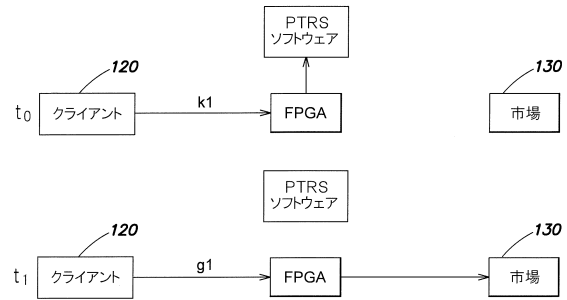


FIG. 12

【図 1 3】

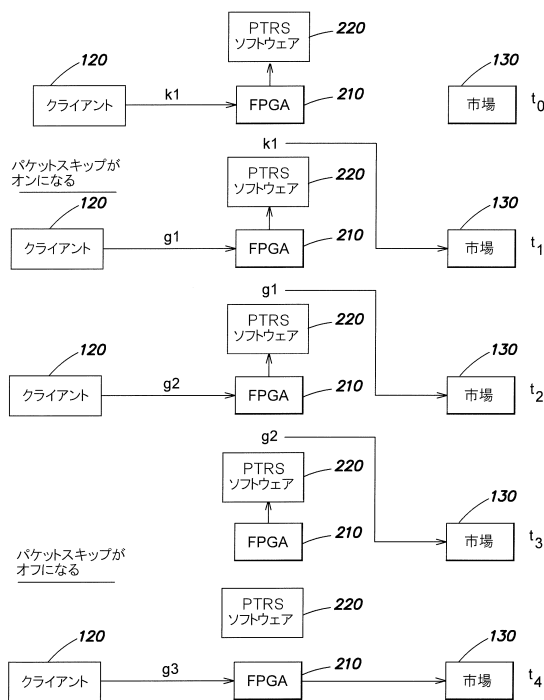


FIG. 13

【図 1 4】

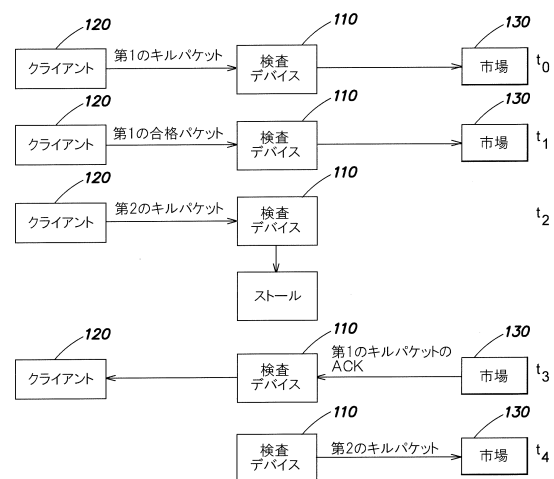


FIG. 14

10

20

30

40

50

フロントページの続き

- 弁理士 中田 健一
(74)代理人 100155963
弁理士 金子 大輔
(74)代理人 100150566
弁理士 谷口 洋樹
(74)代理人 100213470
弁理士 中尾 真二
(72)発明者 アミカンギオーリ・アンソニー・ディー
アメリカ合衆国, マサチューセッツ州 02161, ニュートン, プライヤーフィールド ロード 29
(72)発明者 カーブ・アンドリュウ・シー
アメリカ合衆国, マサチューセッツ州 02139, ケンブリッジ, ロッキンガム ストリート 32
(72)発明者 フィールド・ティモシー・ジー
アメリカ合衆国, マサチューセッツ州 01879, タイングスバラ, メープル リッジ ロード 10
(72)発明者 グロホヴィナ・ドミニク・エス
アメリカ合衆国, マサチューセッツ州 02142, ケンブリッジ, 10ジー, ワズワース ストリート 60
(72)発明者 ピャトニチコ・ユラ
アメリカ合衆国, マサチューセッツ州 02472, ウォータータウン, レキシントン ストリート 358
(72)発明者 ローゼン・バーナード・ジェー
アメリカ合衆国, マサチューセッツ州 01886, ウェストフォード, プロビデンス ロード 23
審査官 宮島 郁美
(56)参考文献 特開2009-152953 (JP, A)
特表2016-515316 (JP, A)
中国特許出願公開第105144660 (CN, A)
米国特許出願公開第2012/0166327 (US, A1)
特開2008-028740 (JP, A)
パケットキャプチャプログラムの作成とネットワーク検証, C MAGAZINE, 第14巻 第5号, 日本, 2002年05月01日, 第147-154頁
(58)調査した分野 (Int.Cl., DB名)
H04L12/00-13/18, 41/00-69/40