

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4034566号

(P4034566)

(45) 発行日 平成20年1月16日(2008.1.16)

(24) 登録日 平成19年11月2日(2007.11.2)

(51) Int. Cl.		F I		
G06F 13/10	(2006.01)	G06F 13/10	330B	
H04L 12/56	(2006.01)	H04L 12/56	Z	

請求項の数 27 (全 23 頁)

(21) 出願番号	特願2001-575045 (P2001-575045)	(73) 特許権者	390009531
(86) (22) 出願日	平成13年4月10日 (2001.4.10)		インターナショナル・ビジネス・マシー ズ・コーポレーション
(65) 公表番号	特表2003-532179 (P2003-532179A)		INTERNATIONAL BUSIN ESS MACHINES CORPO RATION
(43) 公表日	平成15年10月28日 (2003.10.28)		アメリカ合衆国10504 ニューヨーク 州 アーモンク ニュー オーチャード ロード
(86) 国際出願番号	PCT/GB2001/001600	(74) 代理人	100086243
(87) 国際公開番号	W02001/078308		弁理士 坂口 博
(87) 国際公開日	平成13年10月18日 (2001.10.18)	(74) 代理人	100091568
審査請求日	平成14年10月25日 (2002.10.25)		弁理士 市位 嘉宏
(31) 優先権主張番号	09/547,362	(74) 代理人	100108501
(32) 優先日	平成12年4月11日 (2000.4.11)		弁理士 上野 剛史
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 ネットワーク・プロセッサ・デバイスの挙動全体を規定および制御する方法

(57) 【特許請求の範囲】

【請求項1】

通信ネットワークをサービスするネットワーク処理環境内に実装されたネットワーク・プロセッサの挙動全体を制御するためのシステムであって、前記ネットワーク処理環境はガイドド制御フレームを発生するための制御ポイント・サブシステムを含み、前記制御ポイント・サブシステムは前記ネットワーク・プロセッサを制御するために接続された汎用プロセッサ(GPP)を含み、前記ネットワーク・プロセッサがパケット分類、パケット修正、キュー/ポリシー管理及びデータパケット転送の1つまたは複数を実施し、前記システムが、

前記ネットワーク・プロセッサ内に位置するインタフェース・デバイスであって、デバイス制御パラメータ・データを用いて前記ネットワーク・プロセッサ内の様々な機能デバイスを設定するための1つまたは複数の制御機能を含む前記ガイドド制御フレームを受け取るためのインタフェース・デバイスと、

受け取られた制御フレームからの前記1つまたは複数の制御機能および制御パラメータ・データを、前記ネットワーク・プロセッサ内の設定されるべき機能デバイスに転送する手段と、

前記機能デバイス内に実装された処理手段であって、前記制御フレーム中で指定されたように制御機能を実行するための処理手段であって、それによって前記機能デバイスが前記実行される制御機能およびデバイス制御パラメータ・データに従って制御される、前記処理手段と

10

20

を含む、システム。

【請求項 2】

前記デバイス制御パラメータ・データが、前記ネットワーク・プロセッサの挙動全体を制御するためのものである、請求項 1 に記載のシステム。

【請求項 3】

受け取られたフレームを解析して、前記制御フレームをどのデバイスにルーティングすべきかを前記フレームから決定するための処理機構をさらに含む、請求項 1 または 2 に記載のシステム。

【請求項 4】

前記転送する手段が、前記ネットワーク・プロセッサに向けられた制御フレームを前記制御ポイント・サブシステムから受け取るための、かつ制御対象のデバイスによって生成された応答フレームを前記制御ポイント・サブシステムに転送し返すための 1 つまたは複数のデータ・ムーバ・インタフェースを含む、請求項 2 に記載のシステム。

10

【請求項 5】

前記ネットワーク・プロセッサが、それぞれフレームを処理することのできる入口側と出口側を含み、前記転送する手段が、前記入口側と出口側の間でフレームを転送するための高速リンクを含む、請求項 1 ないし 4 のいずれかに記載のシステム。

【請求項 6】

前記転送する手段が、前記ネットワーク・プロセッサ内の設定されるべき前記デバイスに前記制御フレームを転送する、請求項 1 に記載のシステム。

20

【請求項 7】

前記ネットワーク・プロセッサが、フレームを前記ネットワーク・プロセッサ内の出口側から前記入口側に転送するためのラップ・データ・ムーバ・インタフェースを含む、請求項 5 に記載のシステム。

【請求項 8】

前記ネットワーク処理環境が少なくとも 2 つのネットワーク・プロセッサを含み、制御フレームが第 1 のネットワーク・プロセッサにおいて受け取られ、前記制御フレームが、別のネットワーク・プロセッサにあるデバイスを制御するための指示制御機能およびデバイス・パラメータ・データを含み、前記転送する手段が、前記制御フレームを前記第 1 のネットワーク・プロセッサから前記別のネットワーク・プロセッサに通信する手段を含む、請求項 1 ないし 7 のいずれかに記載のシステム。

30

【請求項 9】

前記制御フレームを前記第 1 のネットワーク・プロセッサから前記別のネットワーク・プロセッサに通信する手段が高速データ・リンクを含む、請求項 8 に記載のシステム。

【請求項 10】

前記制御ポイント・サブシステムに応答フレームを返す場合に前記ネットワーク環境内における前記制御ポイント・サブシステムの位置を知るための機構をさらに含む、請求項 4 に記載のシステム。

【請求項 11】

前記デバイス制御パラメータ・データが、前記ネットワーク処理環境内における前記ネットワーク・プロセッサのターゲット位置を示すためのデータであって、前記ターゲット位置が前記ネットワーク・プロセッサのアドレス位置を含むデータと、

40

前記ネットワーク・プロセッサ内で動作するターゲット機能デバイスを含むデータと、前記ターゲット・ネットワーク・プロセッサ内で動作する前記ターゲット機能デバイスを設定するための制御機能および関連する制御パラメータを含むデータであって、それにより、前記制御機能および関連する制御パラメータを用いて前記ターゲット・ネットワーク・プロセッサ内の前記機能デバイスを設定するために、前記制御フレームを受け取った第 1 のネットワーク・プロセッサが、制御機能および関連する制御パラメータを機能デバイスに転送することができるように、前記制御フレームを前記ターゲット・ネットワーク

50

・プロセッサに転送する、前記データとを含む、請求項 1 に記載のシステム。

【請求項 1 2】

機能デバイスを設定する間にエラーが発生したときに応答制御フレームを構築する必要があるかどうかを示すためのデータと、

前記フレームが要求制御フレームか応答制御フレームかを示すためのデータとをさらに含む、請求項 1 1 に記載のシステム。

【請求項 1 3】

前記制御フレームが前記ネットワーク・プロセッサの入口側に向けられているか出口側に向けられているかを示すためのデータをさらに含む、請求項 1 1 または 1 2 に記載のシステム。

10

【請求項 1 4】

前記制御フレームが、応答フレームの送り先となるソース・デバイスを示すためのデータをさらに含む、請求項 1 2 に記載のシステム。

【請求項 1 5】

前記制御フレームが、前記制御フレームから設定される機能デバイスが前記制御フレームを供給する前記制御ポイント・システムのアドレス位置を知るべきかどうかを示すためのデータをさらに含む、請求項 1 4 に記載のシステム。

【請求項 1 6】

前記ターゲット機能デバイスを設定するための前記制御機能および関連する制御パラメータがユーザによって設定可能である、請求項 1 4 または 1 5 に記載のシステム。

20

【請求項 1 7】

通信ネットワークをサービスするネットワーク処理環境内に実装されたネットワーク・プロセッサの挙動全体を制御する方法であって、前記ネットワーク処理環境はガイドド制御フレームを発生するための制御ポイント・サブシステムを含み、前記制御ポイント・サブシステムは前記ネットワーク・プロセッサを制御するために接続された汎用プロセッサ (GPP) を含み、前記ネットワーク・プロセッサがパケット分類、パケット修正、キュー/ポリシー管理及びデータパケット転送の 1 つまたは複数を実施し、

前記方法が、

a) 制御ポイント・サブシステムによってガイドド制御フレームを生成するステップと、

30

b) デバイス制御パラメータ・データを用いて前記ネットワーク・プロセッサ内の様々な機能デバイスを設定するための 1 つまたは複数の制御機能を含む前記ガイドド制御フレームを受け取るステップと、

c) 受け取られた制御フレームからの前記 1 つまたは複数の制御機能および制御パラメータ・データを、前記ネットワーク・プロセッサ内の設定されるべき機能デバイスに転送するステップと、

d) 制御フレーム中で指定されたように制御機能を実行し、それによって前記デバイスを設定するステップであって、それによって前記ネットワーク・プロセッサが、前記実行される制御機能およびデバイス制御パラメータ・データに従って制御される、前記設定するステップと

40

を含む、方法。

【請求項 1 8】

前記デバイス制御パラメータ・データが、前記ネットワーク・プロセッサの挙動全体を制御するためのものである、請求項 1 7 に記載の方法。

【請求項 1 9】

受け取ったフレームを解析して、前記制御フレームをどのデバイスにルーティングすべきかを前記フレームから決定するステップをさらに含む、請求項 1 7 または 1 8 に記載の方法。

【請求項 2 0】

50

前記ネットワーク・プロセッサが1つまたは複数のデータ・ムーバ・インタフェースを含み、前記受け取るステップが、前記ネットワーク・プロセッサに向けられた制御ポイントを、データ・ムーバ・インタフェースを介して前記制御ポイント・サブシステムから受け取ることを含む、請求項18に記載の方法。

【請求項21】

前記制御フレーム中に指定されたように制御機能を実行する前記ステップの間に、または制御機能を実行するのに応答して、

前記ネットワーク・プロセッサ内の前記機能デバイス内で応答フレームを構築するステップと、

前記応答フレームを前記制御ポイント・サブシステムに転送し返すステップとを含む、請求項20に記載の方法。

10

【請求項22】

前記ネットワーク・プロセッサが、それぞれフレームを処理することのできる入口側と出口側を含み、前記転送するステップが、前記入口側と出口側の間で高速データ・リンクを介してフレームを転送するステップをさらに含む、請求項17ないし21のいずれかに記載の方法。

【請求項23】

前記転送するステップが、前記ネットワーク・プロセッサ内の設定されるべき前記デバイスに転送するために前記制御フレームをキューに入れるステップを含む、請求項17に記載の方法。

20

【請求項24】

前記ネットワーク・プロセッサがラップ・データ・ムーバ・インタフェースを含み、前記転送するステップが、前記ラップ・インタフェースを介してフレームを前記出口側から前記入口側にラップするステップをさらに含む、請求項22に記載の方法。

【請求項25】

前記ネットワーク処理環境が少なくとも2つのネットワーク・プロセッサを含み、制御フレームが第1のネットワーク・プロセッサにおいて受け取られ、前記制御フレームが、別のネットワーク・プロセッサにあるデバイスを制御するための指示制御機能およびデバイス・パラメータ・データを含み、前記転送するステップが、前記制御フレームを前記第1のネットワーク・プロセッサから前記別のネットワーク・プロセッサに通信するステップをさらに含む、請求項17ないし24のいずれかに記載の方法。

30

【請求項26】

前記通信するステップが高速データ・リンクを介して行われる、請求項25に記載の方法。

【請求項27】

前記機能デバイス内で応答フレームを構築する前記ステップが、前記制御ポイント・サブシステムに前記フレームを返す場合に前記ネットワーク環境内における前記制御ポイント・サブシステムの位置を知るステップをさらに含む、請求項21に記載の方法。

【発明の詳細な説明】

【0001】

40

【発明の属する技術分野】

本発明は一般にネットワーク・プロセッサ・デバイスに関し、より詳細には、ネットワーク・プロセッサの挙動全体を規定および制御するための改良型の方法に関する。

【0002】

【従来の技術】

今日のネットワーク世界では、帯域幅がクリティカルなリソースである。インターネットおよび新しく現れるその他の用途によって押し上げられるネットワーク・トラフィックの増加は、ネットワーク・インフラストラクチャの容量に過大な負担をかけている。組織は、遅れをとらないように、トラフィック増加および音声とデータとの収束をサポートし管理するためのよりよい技術および方法を求めている。

50

【 0 0 0 3 】

今日のネットワーク・トラフィックの劇的な増加は、インターネットの普及、情報へのリモート・アクセスの必要性の増大、および新しく現れる用途に起因すると考えることができる。インターネットだけでも、電子商取引におけるその爆発的な成長に伴い、時としてサポート不可能な場合もある負荷をネットワーク・バックボーンに課してきた。インターネットはまた、音声トラフィックを初めて上回るデータ・トラフィック量の増加に対する唯一の最も重要な原因でもある。eメール、データベース・アクセス、ファイル転送を含めたリモート・アクセス用途の需要の増加が、さらにネットワークに過大な負担をかけている。

【 0 0 0 4 】

音声とデータの収束は、将来のネットワーク環境を規定する上で大きな役割を果たすであろう。現在、インターネット・プロトコル（IP）ネットワークを介したデータ伝送は、国によっては無料である。音声通信は当然ながら最低コストの道を進むことになるので、音声は必然的にデータと収束するであろう。この変化しつつある市場では、ボイス・オーバーIP（VoIP）、ボイス・オーバーATM（VoATM）、ボイス・オーバー・フレーム・リレー（VoFR）などの技術が、コスト効果の高い代替技術である。しかし、これらの技術への移行を可能にするには、産業は、音声に対するサービス品質（QoS）を保証し、データ回線を介した音声伝送にどのように課金するかを決定しなければならない。1996年の米国の電気通信自由化法令が、この環境をさらに複雑なものにしている。この法律は、好まれる音声プロトコルであるATMと好まれるデータ・プロトコルであるIPとの共存関係を強化することになる。

【 0 0 0 5 】

新しい製品および機能が利用可能になるので、レガシー・システムの統合もまた、組織にとって非常に重大な問題である。既存の装置およびソフトウェアに対する組織の投資を存続させるために、組織は、現在の操業を絶つことなく新しい技術に移行することを可能にする解決法を必要とする。

【 0 0 0 6 】

ネットワーク・ボトルネックをなくすことは、サービス・プロバイダにとって最も高い優先順位であり続けている。これらのボトルネックの原因はルータであることが多い。しかし、一般にネットワーク輻輳は、しばしば帯域幅の問題として誤診され、より広い帯域幅の解決法を探すことによって対処される。今日、製造業者はこの難題を認識し始めている。彼らは、帯域幅リソースをより効率的に管理し、かつルータおよびネットワーク・アプリケーション・サーバによくみられる高度なデータ・サービスをワイヤ速度で提供する、ネットワーク・プロセッサ技術に目を向けつつある。これらのサービスには、負荷平衡、QoS、ゲートウェイ、ファイアウォール、セキュリティ、およびウェブ・キャッシングが含まれる。

【 0 0 0 7 】

リモート・アクセス用途では、性能、オンデマンドの帯域幅、セキュリティ、および認証が最も高い優先順位を占める。QoSとサービス・クラス（CoS）との統合、統合音声処理、およびより複雑なセキュリティ解決法に対する需要もまた、将来のリモート・アクセス・ネットワーク・スイッチの設計を形作ることになる。さらにリモート・アクセスは、ISDN、T1、E1、OC-3～OC-48、ケーブル、xDSLモデムなど、増え続ける物理媒体に対応しなければならないであろう。

【 0 0 0 8 】

産業コンサルタントは、ネットワーク・プロセッサ（本明細書では「NP」とも言う）を、以下の機能の1つまたは複数を実施することのできるプログラム可能な通信集積回路と定義している。

パケット分類 - アドレスやプロトコルなど、認められる特性に基づいてパケットを識別する。

パケット修正 - IP、ATM、またはその他のプロトコルに適合するようにパケットを

10

20

30

40

50

修正する（例えばIP用のヘッダ中の存続時間（time-to-live）フィールドを更新する）。

キュー/ポリシー管理 - パケット・キューイング、デキューイング、およびパケットのスケジューリングに関する設計戦略を特定の用途に向けて反映する。

パケット転送 - スイッチ・ファブリックを介してデータを送受信し、パケットを適切なアドレスにルーティングする。

【0009】

この定義は初期のNPの基本的な特徴を正確に述べているものの、NPの完全な潜在的能力および利点はまだ実現されていない。ネットワーク・プロセッサは、以前はソフトウェア中で対処されていたネットワーク・タスクをハードウェア中で実行できるようにすることにより、帯域幅を増加させ、広範な用途における待ち時間の問題を解決することができる。さらにNPは、並列分散処理やパイプライン処理設計などのアーキテクチャにより、速度向上をもたらすこともできる。これらの能力は、効率的な探索エンジンを可能にし、スループットを向上させ、複雑なタスクの高速実行を実現することができる。

10

【0010】

ネットワーク・プロセッサは、CPUがPCにとって根本的な構成要素であるのと同じ形で、ネットワークにとって根本的な構成要素になると予想される。NPによって提供される代表的な能力は、リアルタイム処理、セキュリティ、蓄積交換、スイッチ・ファブリック・インタフェース、ならびにIPパケット処理および学習の能力である。NPは、ISOLeia 2から5までを対象とし、ネットワーク特有のタスクを最適化するように設計される。

20

【0011】

プロセッサモデルNPは、複数の汎用プロセッサおよび特殊化ロジックを組み込む。供給業者は、コスト効果の高い形で適時に変化に対応できるスケーラブルかつフレキシブルな解決法を提供するために、この設計に目を向け始めている。プロセッサモデルNPは、統合レベルのより低い分散処理を可能にし、より高いスループット、フレキシビリティ、および制御をもたらす。プログラム可能性により、新しいASIC設計を必要とせず新しいプロトコルおよび技術に容易に移行することができる。

【0012】

【発明が解決しようとする課題】

ネットワーク・プロセッサの挙動全体を規定および制御するためのデータ構造および方法は、NPベースのデバイスの設計および実装に見合ったものである。本発明は、このデータ構造および方法を対象とする。

30

【0013】

【課題を解決するための手段】

したがって本発明は、通信ネットワークをサービスするネットワーク処理環境内に実装されたネットワーク・プロセッサの挙動全体を制御するためのシステムを提供する。このシステムは、前記ネットワーク・プロセッサ内に位置するインタフェース・デバイスであって、デバイス制御パラメータ・データを用いて前記ネットワーク・プロセッサ内の様々な機能デバイスを設定するための1つまたは複数の制御機能を含む前記ガイドド制御フレームを受け取るためのインタフェース・デバイスと、受け取られた制御フレームからの前記1つまたは複数の制御機能および制御パラメータ・データを、前記ネットワーク・プロセッサ内の設定されるべき機能デバイスに転送する手段と、前記機能デバイス内に実装された処理手段であって、前記制御フレーム中で指定されたように制御機能を実行するための処理手段であって、それによって前記機能デバイスが前記実行される制御機能およびデバイス制御パラメータ・データに従って制御される、前記処理手段とを含む。

40

【0014】

好ましい一実施形態によれば、ネットワーク環境で制御ポイント・プロセッサ（CP）として働く汎用プロセッサ（GPP）がネットワーク・プロセッサ（NP）の挙動全体を規定および制御できるようにするための機能をもたらすシステムおよび方法が提供される

50

。

【0015】

さらに、ネットワーク環境で制御ポイント・プロセッサ（CP）として働く汎用プロセッサ（GPP）がネットワーク・プロセッサ（NP）の挙動全体を規定および制御して、NP内で稼動するフレーム転送アプリケーションに対するサポート機能を提供できるようにするための機能をもたらず方法が提供されることが好ましい。

【0016】

好ましい一実施形態によれば、新規な制御フレーム・データ構造および通信インフラストラクチャが実装され、それにより、分散ネットワーク処理環境で動作する任意のネットワーク・プロセッサ・デバイスを、実行される制御機能およびデバイス制御パラメータ・データに従って制御することができる。制御に特意的を絞ったネットワーク・プロセッサ・デバイス内の機能ユニットは、ハードウェア・ガイドド・フレーム・ハンドラ（GFH）デバイスと、ハードウェア・ガイドド・テーブル・ハンドラ（GTH）デバイスによって制御されるフレーム転送アプリケーションとを含むことが好ましい。

10

【0017】

有利にも、このシステムおよび方法は、ネットワーク処理環境における1次ブレード上でのガイドド・フレーム・フロー、2次ブレード上でのガイドド・フレーム・フロー、および複数ブレード上でのガイドド・フレーム・フローを含めて、通常のNPシステムを通る種々の可能なフローにおいて制御フレーム（ガイドド・フレーム・フロー）を扱い処理することができるのが好ましい。

20

【0018】

他の態様では、本発明は、ネットワーク処理環境内に実装されたネットワーク・プロセッサ・デバイスに通信される制御フレームのためのデータ構造を提供する。このデータ構造は、前記ネットワーク処理環境内における前記ネットワーク・プロセッサのターゲット位置を示すためのデータであって、前記ターゲット位置が前記ネットワーク・プロセッサのアドレス位置を含むデータと、前記ネットワーク・プロセッサ内で動作するターゲット機能デバイスを含むデータと、前記ターゲット・ネットワーク・プロセッサ内で動作する前記ターゲット機能デバイスを設定するための制御機能および関連する制御パラメータを含むデータであって、それにより、前記制御機能および関連する制御パラメータを用いて前記ターゲット・ネットワーク・プロセッサ内の前記機能デバイスを設定するために、前記制御フレームを受け取った第1のネットワーク・プロセッサが、制御機能および関連する制御パラメータを機能デバイスに転送することができるように、前記制御フレームを前記ターゲット・ネットワーク・プロセッサに転送する、前記データを含む。

30

【0019】

他の態様では、本発明は、信ネットワークをサービスするネットワーク処理環境内に実装されたネットワーク・プロセッサの挙動全体を制御する方法を提供する。この方法は、

a) 制御ポイント・サブシステムによってガイドド制御フレームを生成するステップと、

b) デバイス制御パラメータ・データを用いて前記ネットワーク・プロセッサ内の様々な機能デバイスを設定するための1つまたは複数の制御機能を含む前記ガイドド制御フレームを受け取るステップと、

40

c) 受け取られた制御フレームからの前記1つまたは複数の制御機能および制御パラメータ・データを、前記ネットワーク・プロセッサ内の設定されるべき機能デバイスに転送するステップと、

d) 制御フレーム中で指定されたように制御機能を実行し、それによって前記デバイスを設定するステップであって、それによって前記ネットワーク・プロセッサが、前記実行される制御機能およびデバイス制御パラメータ・データに従って制御される、前記設定するステップと

を含む。

【0020】

50

次に、本発明の実施形態を、図例を参照しながら単に例として述べる。

【0021】

【発明の実施の形態】

図1に、ネットワーク(図示せず)にサービスするための1つまたは複数のネットワーク・プロセッサ(「NP」)インタフェース・デバイス15a、bを実装した例示的なネットワーク・プロセッサ(NP)処理環境10を一般に示す。

【0022】

図1に示す例示的なNP処理環境10を参照すると、第1のNPインタフェース・デバイス15aは1次ブレード(すなわちプリント回路板要素)20a上に設けられ、第2のNPインタフェース・デバイス15bは2次ブレード20b上に設けられている。このシステムは、多くのブレード(1次および2次(複数の2次))ならびに制御ポイント(「CP」)サブシステムを、例えば分散ネットワーク構成で含むように拡張することもできることを理解されたい。制御ポイント・サブシステム(「CP」)50は、NP環境構成に接続されたシステム・プロセッサ52を含む。CPにあるシステム・プロセッサは、とりわけシステム制御フレームを発行することによってNPインタフェース・デバイスに初期化および設定サービスを提供することが好ましく、本明細書ではこのシステム制御フレームを「ガイドド・フレーム」と呼ぶ。CP50は、3つの位置、すなわちインタフェース・デバイス・チップ内と、チップが(図1に示すように)上に装着されているブレード上と、ブレード外部とのいずれかに位置することができる。ブレード外部の場合は、CPはリモートとすることができる。すなわちCPは、どこか他の場所に収納され、インタフェース・デバイスとCP50とを接続するネットワークを介して通信する。一般にCPの要素には、プロセッサ52、メモリ素子(キャッシュ、フラッシュ、およびSDRAM)、メモリ・コントローラ、PCIバス、バックプレーン用およびL1ネットワーク媒体用のコネクタが含まれる。

【0023】

本明細書でより詳細に述べるが、好ましい実施形態では、ガイドド・フレームは、以下の情報を含む可変ワード構造である。すなわち、1)フレームがどのブレードのためのものか、およびブレード内のどの機能ブロックに送るべきかを示すためのデータを含むフレーム制御情報、2)エラーまたは例外が発生した場合、あるいはフレーム処理が成功した場合に、応答フレームがあればどこに送るべきかをフレーム中で示すための相関関係子情報、3)ブレード内の機能ユニットを制御または管理する制御オペランドを含むガイドド・コマンド、4)フレームの終了を示す終了区切りコマンドである。これらのそれぞれについては、図16を参照しながらより詳細に論じる。

【0024】

図1に示すように、各NPインタフェース・デバイスは組込みプロセッサ・コンプレックス(EPC)16を備え、EPC16は2つの機能サブアセンブリを有する。すなわち、NPシステム10によってサービスされる外部ネットワークからの、ネットワーク・プロトコル特有のデータ・ムーバ・インタフェース(DM)24a、bを介したデータを受信/処理するための入口(アップサイド)サブシステム26aと、NPシステム10によってサービスされるネットワークにネットワーク・プロトコル特有のデータ・ムーバ・インタフェース(DM)24c、dを介して出て行くデータを送信するための出口(ダウンサイド)サブシステム26bである。同じNPインタフェース・デバイス内の入口から出口に渡されるフレームはすべて、DASL(Data Align Synchronous Link)インタフェース45a~45dを介して渡される。高データ・フロー・レートをサポートする他の同等タイプのインタフェース/リンクを実装することもできることを理解されたい。出口側26bからプロセッサの入口側26aに内部で通信して戻す必要のあるフレームは、データ・ムーバ・ラップ(DM__W)25を介して通信することができる。

【0025】

一般に、NPインタフェース・デバイスの入口サブアセンブリ26aは、組込みガイドド・フレーム・ハンドラ・ユニット18(GFH)および1つまたは複数のジェネラル

10

20

30

40

50

・データ・ハンドラ (GDH) プロセッシング・ユニット 17a を備え、各 GDH プロセッシング・ユニット 17a は、受信フレームを分類して、フレームをどこに転送するか調べるためにルックアップを行うように機能する。NP インタフェース・デバイスの出口サブアセンブリ 26b もまた、組み込みガイドド・フレーム・ハンドラ・ユニット 18a (GFH) および 1 つまたは複数のクラシファイア/ルックアップ・プロセッサ 17b を備え、各クラシファイア/ルックアップ・プロセッサ 17b は、フレーム送信前に必要なら何らかのアウトバウンド修正またはルックアップを行うように機能する。出口サブアセンブリ 26b はさらに、ガイドド・テーブル・ハンドラ・ユニット 19 (GTH) も備える。NP プロセッサ・デバイス内では、受信されるデータ・トラフィックはすべて、本明細書でガイドド・データ・ハンドラ (GDH) と呼ぶ特別なクラシファイア/ルックアップ・プロセッサによって対処し、NP システム 10 を管理するのに使用される制御情報は、本明細書でより詳細に述べるように、ガイドド・フレーム・ハンドラ (GFH) 18 によって対処する。フレーム転送プロセスでは、ガイドド・テーブル・ハンドラ (GTH) 19 が転送テーブルを実装する。

【0026】

図 1 を参照すると、NP インタフェース・デバイスにおいて受信されたパケットまたはフレームの一般的なフローは次のとおりである。ネットワーク、例えば Ethernet (R) MAC から受信されたフレームは、データ・ムーバ・インタフェース (DM) 24a、24b において受け取られ、そこで通常のデータ・フレームとして識別されるか、または制御プロセッサ (CP) システム 50 から受信されたシステム制御フレーム (ガイドド・フレーム) として識別される。ガイドド・フレームはまず、一般に組み込みプロセッサ・コンプレックス (EPC) のキューに入れられる。EPC は、受信フレームのヘッダを見てそのフレームをどうするか (転送、修正、フィルタリングなど) を決定することができるロジック (ピココード) を実行する複数のピコプロセッサ、例えばプロトコル・プロセッサを備える。EPC は、いくつかのルックアップ・テーブルにアクセスすることができ、分類ハードウェアが、ネットワーク・プロセッサの広帯域幅要件にピコプロセッサがついていけるようにアシストする。周知のフレーム・フォーマットのフレームを分類するために、分類ハードウェア・アシスト・デバイスが設けられる。組み込みプロセッシング・コンプレックス (EPC) は、NP デバイス・チップのプログラム可能性を提供および制御し、コンポーネント (メモリ、ディスパッチャ、DM インタフェースなど) の中でとりわけ、共通の命令メモリに記憶されたピココードを同時に実行する 1 つまたは複数のプロセッシング・ユニットを備える。各プロセッシング・ユニットは、例えば 3 段パイプラインと汎用レジスタと ALU とを含む CLP プロセッシング・ユニット・コアを備えることが好ましい。動作に際しては、フレーム・ディスパッチ中に、分類ハードウェア・アシスト・デバイスからの分類結果が 1 つまたは複数のプロセッシング・ユニットに渡される。プロセッシング・ユニットの 1 つは特に、ジェネラル・データ・ハンドラ (GDH) と呼び、これは完全な CLP プロセッシング・ユニットおよび 5 つのコプロセッサを備え、主にフレームを転送するのに使用される。N 個のプロトコル・プロセッサのうち少なくとも 2 つは特殊化され、1 つはガイドド・フレームを扱うためのガイドド・フレーム・ハンドラ GFH 18 であり、もう 1 つは制御メモリ中にルックアップ・データを構築するためのガイドド・テーブル・ハンドラ GTH 19 である。

【0027】

好ましい実施形態では、ガイドド・フレームに分類されたフレームはすべて、そのフレームを受け取っているブレード上の GFH の入口キューに入れられる。図 2 に、ガイドド・フレーム・ハンドラ GFH ユニット 18 の一般的な機能レイアウトを示す。GFH 18 は、ガイドド・フレームの解析およびルーティングの機能を実施することを担うガイドド・フレーム・プロセッサ 181 と、ガイドド・フレーム中のガイドド「コマンド」を解釈して、それらのコマンドが処理されるようにすることを担うガイドド・コマンド・プロセッサ 182 とを含むソフトウェア・コンポーネントを実装したハードウェア・デバイスであることが好ましい。ガイドド・フレーム・プロセッサ 181 は、ガイ

10

20

30

40

50

デッド・フレームを受け取って処理し、また、ガイドド要求フレームを妥当性検査して、受け取ったフレームをGFHが処理できるかどうかを決定することを担う。また、送信側によって必要とされる肯定応答（もしあれば）のタイプも決定する。ガイドド・フレーム・プロセッサ181はさらに、処理すべき第1のガイドド・コマンドを見つけるまで要求フレームをざっと調べ、次いでガイドド・コマンド・プロセッサ182に制御を渡す。コマンドが完了すると、制御はガイドド・フレーム・プロセッサ181に戻される。そうでない場合は、ガイドド・コマンド・プロセッサがエラーを返すまで、またはEnd_Delimiterコマンドがみられるまで、ガイドド・コマンド・プロセッサはフレームの各ガイドド・コマンドを処理する。この時点で、フレーム処理は完了したと見なされる。

10

【0028】

ガイドド・コマンド・プロセッサは、ガイドド要求フレーム中に組み込まれたガイドド・コマンドを処理することを担う。これは、コマンドが見つかる場所である要求フレーム中の位置への参照と共に、ガイドド・フレーム・プロセッサから呼び出される。ガイドド・コマンド・プロセッサは、コマンド識別子を解析して、コマンドを処理できるかどうか決定する。指定されたコマンドがサポートされない場合は、フレーム・パーサにエラーが返される。そうでなければ、本明細書に述べるようにして、図2に示すブロック185の適切な機能ブロックに制御が転送され、この機能ブロックが実際にコマンドにサービスする。

【0029】

20

図3に、GTHユニット19の機能レイアウトを示す。GTHアーキテクチャは、遊休状態でありガイドド要求フレームがGTHキュー中で利用可能であることがNPディスプレイにわかったときにディスプレイされる。着信フレームがガイドド・フレーム・プロセッサ191によって妥当性検査され、コマンド処理のために制御がガイドド・コマンド・プロセッサ192に渡されるという点で、GTHはGFHと同様の仕方で機能する。図3に示すように、GTH内に含まれるサービス・ブロック195は、GFHに含まれる機能ブロックとは異なる動作を行う。しかし、ガイドド・フレームの目的は、図2および3に示す各動作ブロックを実際に管理および制御することであることを理解されたい。

【0030】

30

続いて、NP インタフェース・デバイスの挙動を制御および管理するための例示的なGFHメッセージ・フロー・シナリオについて詳細に述べる。図4および5に示す第1の例には、CPサブシステム50と1次ブレード20a上のガイドド・フレーム・ハンドラ（GFH）18との間のメッセージ・フローを含む例示的なシナリオ200が示されている。

【0031】

図4に示す例示的なシナリオ200では、ガイドド「要求」フレームは、1次ブレードの入口GFHに向けられた組込みガイドド・コマンドを含む。GFHは、ガイドド要求を次のように処理する。第1のステップで、ガイドド・フレームはCP50によって生成され、1次ブレード20に送られる。このシナリオでは、フレームは、応答が必要であることを示すビットを含むものとする。フレームは、1次ブレード20a上の入口GFHによって受け取られて解析される。解析結果は、フレームがそれ自体に向けられていることを示す。したがって、組込みガイドド・コマンドは入口GFHによって処理され、具体的にはガイドド・コマンド・プロセッサによって実行される。処理が成功したと仮定した場合、結果はフレームに記憶され、フレームが今や「応答」フレームであることを示すようにフレーム制御情報が修正される。次いで、1次ブレードに対応するターゲット・ブレード・アドレスと共に入口キュー（図示せず）に入れられ、DASLインタフェース45a、bを介して1次ブレード20aの出口GFHに送られる。次いで出口GFH18aは、フレームを解析し、フレームが応答フレームであることを認め、適切な出口キュー（図示せず）に入れる。これによりフレームは、例えばDM24aなどのDMインタ

40

50

フェースを通して、CP50に接続されたインタフェースを介して送ることができ、最終的にはCP50が応答フレームを受け取る。

【0032】

図4に示したシナリオ200に対する代替シナリオでは、ガイドド「要求」フレームは、1次ブレードの出口GFHに向けられた組込みガイドド・コマンドを含む。このシナリオでは、ガイドド・フレームは、CP50によって生成されて1次ブレードに送られ、またフレームは、応答が必要であることを示すビットを含むものとする。フレームは、1次ブレード上の入口GFHによって受け取られて解析される。この場合、解析結果は、フレームが出口GFHに向けられていることを示す。したがってフレームは、ターゲット・ブレードを1次ブレードに設定されて入口キューに入れられ、DASLインタフェース45a、bを介して1次ブレード20aの出口GFHに送られる。次いで出口GFHは、フレームを解析し、フレームがそれ自体に向けられていることを決定する。したがって、フレームを処理して応答フレームを構築する。応答フレームが構築された後、応答フレームは適切な出口キューに入れられ、次いで出口キューは、CPに接続されたインタフェースにDMインタフェースを介して応答フレームが送出されるようにする。

10

【0033】

図5に示す例示的なシナリオ210では、ガイドド「要求」フレームはジェネラル・データ・ハンドラ(GDH)によって内部で生成され、1次ブレード上の入口GFH18に向けられた組込みガイドド・コマンドを含む。GFHは、ガイドド要求を次のように処理する。第1のステップで、1次ブレード上の入口GDHが、同じブレード上の入口GFHに向けたガイドド・フレームを構築し、このフレームは入口GFHキューに入れられる。フレームは入口GFHによって受け取られ、フレーム処理の間に、フレームがそれ自体に向けられていることが決定される。したがって、入口GFHはこのフレームを処理する。処理中にエラーがみられなかったと仮定した場合、フレームは、応答フレームが必要であることを示すビット標識がなかったときは、処理された後で入口廃棄キューに入れられる。エラーがみられ、かつフレーム・ヘッダ中のビットが否定応答の必要を示す場合は、否定応答フレームが構築され、出口GFH18aにディスパッチされるように入口キューに入れられる。応答フレームが出口GFHによって受け取られると、出口GFHはフレーム・ヘッダを調べ、応答フレームであることを知り、したがって、CP50にディスパッチするために適切な出口キューに入れる。CP50で、否定応答フレームの応答フレームが受け取られる。

20

30

【0034】

図5に示したシナリオ210に対する第1の代替シナリオでは、ガイドド「要求」フレームはジェネラル・データ・ハンドラ(GDH)によって内部で生成され、1次ブレード上の出口GFH18aに向けられた組込みガイドド・コマンドを含む。このシナリオでは、1次ブレード上の入口GDH20が、出口GFHに向けたガイドド・フレームを構築する。このフレームを構築することのできる場所が2つあることが好ましい。すなわち、1)入口データ・ストア上、および2)出口データ・ストア上である。フレームが入口データ・ストア上で構築される場合、入口GDHは、フレームを入口GFHキューに入れる。入口GFHは、フレームを受け取って解析する。フレームは出口GFHに向けられていることを示すので、入口GFHは、ターゲット・ブレードを1次ブレードに設定してフレームを入口キューに入れ、DASLインタフェース45c、dを介して1次ブレードの出口GFH18aに送る。あるいは、ターゲット・ブレードを1次ブレードに設定して、GDHが直接にフレームを入口キューに入れることもできる。しかし、フレームが出口データ・ストア上で構築される場合は、入口GDHは、フレームが出口GFHによって受け取られて処理されるように、フレームを出口GFHキューに入れる。エラーがみられなかったと仮定した場合、フレームは、処理された後で出口廃棄キューに入れられる。エラーがみられ、かつフレーム・ヘッダが否定応答の必要を示すビットを含んでいた場合は、否定応答フレームが構築されて、適切な出口インタフェース・キューに入れられ、したがって最終的にはDMインタフェースを介してCP50によって受け取られる。

40

50

【 0 0 3 5 】

図5に示したシナリオ210に対する第2の代替シナリオでは、ガイドド「要求」フレームは1次ブレード上の出口GDH20aによって内部で生成され、1次ブレード上の出口GFH18aに向けられた組込みガイドド・コマンドを含む。このシナリオでは、1次ブレード上の出口GDH20aが、出口データ・ストアを使用して、1次ブレード上の出口GFHに向けたガイドド要求フレームを構築する。したがって、出口GDHはフレームを出口GFHキューに入れ、フレームは出口GFHによって受け取られて処理される。処理が成功したと仮定した場合、フレームは出口廃棄キューに入れられる。しかし、エラーがみられ、かつフレームが否定応答の必要を示していた場合は、否定応答フレームが構築されて、適切な出口インタフェース・キューに入れられ、したがって最終的にはDMインタフェースを介してCP50によって受け取られる。

10

【 0 0 3 6 】

図5に示したシナリオ210に対する第3の代替シナリオでは、ガイドド「要求」フレームは1次ブレード上の出口GDH20aによって内部で生成され、1次ブレード上の入口GFH18に向けられた組込みガイドド・コマンドを含む。このシナリオでは、1次ブレード上の出口GDHは、出口データ・ストアと入口データ・ストアのいずれかの場所で、入口GFHに向けたガイドド・フレームを構築する。フレームが出口データ・ストア上で構築される場合は、出口GDHは、フレームが入口GFHに到達できるようにするラップDMキュー25にフレームを入れる。そうではなく、フレームが入口データ・ストア上で構築される場合は、出口GDHは、フレームが入口GFHに到達するように入口GFHキューにフレームを入れ、フレームは入口GFHによって受け取られて処理される。エラーがみられなかったと仮定した場合、フレームは、処理された後で入口廃棄キューに入れられる。エラーがみられ、フレームが否定応答の必要を示していた場合は、否定応答フレームが構築され、出口GFHによって受け取られるようにターゲット・ブレードを1次ブレードに設定されて入口キューに入れられる。フレームは出口GFHで解析される。フレームは応答フレームであることを示すので、出口GFHは、フレームを適切な出口ターゲット・ポート・キューに入れ、したがってフレームは、最終的にはDMインタフェースを介してCP50によって受け取られる。

20

【 0 0 3 7 】

図6および7に示す第2の例には、CPサブシステム50と2次ブレード上のガイドド・フレーム・ハンドラ(GFH)18との間のメッセージ・フローを含む例示的なシナリオが示されている。

30

【 0 0 3 8 】

図6に示す例示的なシナリオ220では、ガイドド「要求」フレームは、2次ブレード20b上の出口GFHに向けられた組込みガイドド・コマンドを含む。GFHは、ガイドド要求を次のように処理する。第1のステップで、ガイドド・フレームはCPによって生成され、1次ブレードに送られる。この例示的なシナリオでは、フレーム制御情報は、フレームが2次ブレード上にある出口GFHをターゲットにしていることを示すビットと、応答が必要であることを示すビットを含むものとする。フレームは、1次ブレード20a上の入口GFH18によって受け取られて解析される。解析の結果、フレームが2次ブレード20b上の出口GFH18cに向けられていることが決定される。したがってフレームは、ターゲット・ブレード上にある出口GFHにディスパッチされるように、ターゲット・ブレード・アドレスを適切に設定されて入口キューに入れられる。出口GFH18cは、フレームを受け取ると、フレームを解析し、フレームがそれ自体に向けられていることを決定する。出口GFHは、要求フレームを処理し、応答フレームを構築する。次いで出口GFHは、応答をラップ・キューDM25に入れることにより、同じブレード上の入口GFH18bコンポーネントに応答を送る。入口GFHフレーム18bは、解析動作を実施し、フレームが応答フレームであることを決定する。したがって、ターゲット・ブレード・アドレスを1次ブレード20aのアドレスに設定した後で入口キューにフレームを入れることにより、フレームはCP50にルーティングされる。したがってフレ

40

50

ームは、D A S L インタフェース 4 5 c を介して 1 次ブレード上の出口 G F H 1 8 a によって受け取られ、そこで解析される。フレーム制御情報はフレームが応答フレームであることを示すので、出口 G F H 1 8 a は、フレームを適切な出口インタフェース・キューに入れる。これによりフレームは、例えば D M 2 4 a などの D M インタフェースを通して、C P 5 0 に接続されたポートを介して送ることができ、最終的には C P 5 0 が応答フレームを受け取る。

【 0 0 3 9 】

図 6 に示したシナリオ 2 2 0 に対する代替シナリオでは、ガイドッド「要求」フレームは、入口 G F H 1 8 b に向けられた組込みガイドッド・コマンドを含む。このシナリオでは、ガイドッド・フレームは C P 5 0 によって生成され、まず 1 次ブレードに送られる。フレーム制御情報は、フレームが 2 次ブレード上にある入口 G F H 1 8 b をターゲットにしていることを示すビットと、応答が必要であることを示すビットを含むものとする。したがってフレームは、1 次ブレード 2 0 a 上の入口 G F H 1 8 によって受け取られて解析され、その結果、フレームが 2 次ブレード上の入口 G F H に向けられていることが決定される。これに回答して、フレームは、ターゲット・ブレード上にある出口 G F H にディスパッチされるように、ターゲット・ブレード・アドレスを適切に設定されて入口キューに入れられる。出口 G F H は、フレームを解析して、フレームがそれ自体の入口コンポーネントに向けられていることを決定し、それに回答して、フレームが入口 G F H 1 8 b によって受け取られるように、フレームをラップ・キュー D M 2 5 に入れる。次いで入口 G F H 1 8 b は、要求フレームを処理し、応答フレームを構築する。応答フレームは、ターゲット・ブレード・アドレスを 1 次ブレードのアドレスに設定した後でフレームを入口キューに入れることにより、C P 5 0 に再びルーティングされ、D A S L インタフェースを介して 1 次ブレード上の出口 G F H 1 8 a によって受け取られる。フレームは、1 次ブレード上の出口 G F H 1 8 a で解析され、フレームが応答フレームであることが決定される。これに回答して、出口 G F H は、フレームを適切な出口インタフェース・キューに入れる。これによりフレームは、例えば D M 2 4 a などの D M インタフェースを通して、C P 5 0 に接続されたポートを介して送ることができ、最終的には C P 5 0 が応答フレームを受け取る。

【 0 0 4 0 】

図 7 に示す例示的なシナリオ 2 4 0 では、ガイドッド「要求」フレームは 2 次ブレード 2 0 b 上の G D H によって生成され、2 次ブレード 2 0 b 上の入口 G F H 1 8 b または出口 G F H 1 8 c に向けられた組込みガイドッド・コマンドを含む。この例示的なシナリオ 2 4 0 では、フローは、図 5 に関して本明細書に述べた同様の 1 次ブレードの例示的なシナリオについて述べたフローと同一である。唯一の異なる点は、エラーの場合に G F H によって生成される否定応答メッセージが辿るパスにある。すなわち、入口 G F H が要求フレームを処理する場合は、1 次ブレード 2 0 a をターゲット・フレームに指定した否定応答フレームが入口キューに入れられる。出口 G F H の場合は、このメッセージは D M _ W インタフェース 2 5 を介して 2 次ブレード上の入口 G F H にラップされる。入口 G F H は、ターゲット・ブレードを 1 次ブレードに設定してフレームを入口キューに入れることにより、フレームを 1 次ブレードの出口 G F H にルーティングする。1 次ブレードの出口 G F H は、フレームを受け取ると、フレームを解析し、フレームが応答フレームであることを決定する。これに回答して、出口 G F H は、フレームを適切な出口ターゲット・インタフェース・キューに入れる。これによりフレームは、例えば D M 2 4 a などの D M インタフェースを通して、C P 5 0 に接続されたポートを介して送ることができ、最終的には C P 5 0 が応答フレームを受け取る。

【 0 0 4 1 】

次に図 8 および 9 に関して述べるように、例えば C P サブシステム 5 0 と 1 次および 2 次の両ブレード上のガイドッド・テーブル・ハンドラ (G T H) 1 9 との間のマルチブレード・メッセージ・フローを含む例示的なシナリオがある。この例示的なシナリオに関連するメッセージ・フローは、1 次ブレードと 2 次ブレードに関して述べたメッセージ・フ

10

20

30

40

50

ローを組み合わせたものである。フレームが1次ブレード18上の入口GFHによってCPから受け取られると、フレームは解析される。解析結果は、フレームが複数のブレード上のGTHに向けられていることを示す。したがって、フレームをターゲット・ブレードのリストと共に入口キューに入れることにより、フレームはすべてのターゲット・ブレードに転送される。フレームが各ターゲット・ブレード上の出口GFHによって受け取られたときのフレーム処理は、前述したのと同じのままである。CPには、初めにターゲットとされていたブレードの数に対応する複数の応答フレームが送られることに留意されたい。ガイドド・フレームがGDHによって生成されて1次ブレードおよび2次ブレード上のGTHに向けられる、関連する例示的なシナリオの場合は、前述の場合と同様だが、発信元のGDHに応答は決して送り返されないという条件が付く。否定応答が指定されている場合は、エラー状況で1つまたは複数の否定応答フレームがCPに送られる。

10

【0042】

したがって、図8に示す第3の例示的なシナリオ260では、ガイドド「要求」フレームはCP50によって生成され、1次ブレード20a上のGTH19に向けられた組込みガイドド・コマンドを含む。この例示的なシナリオでは、フレーム制御情報は、フレームが1次ブレードの出口上にあるGTHをターゲットにしていることを示すビットと、応答が必要であることを示すビットを含むものとする。動作に際して、フレームは1次ブレード上の入口GFH18によって受け取られて解析される。解析結果は、フレームが1次ブレード20a上のGTH19に向けられていることを示す。したがってフレームは、1次ブレードに対応するターゲット・ブレード・アドレスと共に入口キューに入れられる。これによりフレームは、DASLインタフェース45a、bを介して1次ブレードの出口GFH18aに到達する。出口GFH18aは、フレームを解析し、フレームがGTHに向けられたものであることを認め、1次ブレード上のGTHにディスパッチされるようにフレームを出口GTHキューに入れる。GTHは受け取ったフレームを解析する。解析結果は、応答が必要であること、および要求のソースがCPであったことを示す。図16に関して本明細書でより詳細に述べるが、フレームの「相関関係子」フィールドが、フレームのソースがCP50であったことを示す。GTHは、要求フレームを処理して応答フレームを構築する。次いでGTH19は、応答フレームを適切な出口インタフェース・キューに入れることにより、フレームをCP50に送り返す。これによりフレームは、例えばDM24aなどのDMインタフェースを通して、CP50に接続されたポートを介して送ることができ、最終的にはCP50が応答フレームを受け取る。

20

30

【0043】

図9に示す代替シナリオ280では、ガイドド「要求」フレームは、CP50によって生成され、また、2次ブレード20b上のGTH19に向けられた組込みガイドド・コマンドを含む。したがってこのシナリオでは、フレームは、2次ブレード上にあるGTHをターゲットとすることを示すビットと、応答が必要であることを示すビットを含むものとする。前述のフローに従って、ガイドド・フレームはまず、1次ブレード上の入口GFHによって受け取られて解析される。解析の結果、フレームが2次ブレード20b上のGTHに向けられていることが決定される。したがってフレームは、ターゲット・ブレード20b上にある出口GFH18cにディスパッチされるように、ターゲット・ブレード・アドレスを適切に設定されて入口キューに入れられる。出口GFHは、フレームを解析し、フレームがGTHに向けられていることを認め、したがってフレームを出口GTHキューに入れる。これによりフレームはGTHによって受け取られ、GTHは受け取ったフレームを解析する。解析結果により、応答が必要であることが決定され、フレーム制御情報(図16)の「相関関係子」フィールドから、フレームのソースがCP50であったことが決定される。次いでGTHは、要求フレームを処理し、応答フレームを構築し、DM_Wインタフェース25を介して2次ブレード上の入口GFH18bに応答を送る。次いで入口GFHは、解析動作を行い、フレームが応答フレームであることを決定する。したがって、ターゲット・ブレード・アドレスを1次ブレード20aのアドレスに設定した後でフレームを入口キューに入れることにより、フレームはCP50にルーティングされ

40

50

る。フレームは、D A S Lインタフェース45 bを介して1次ブレード20 a上の出口G F Hによって受け取られ、解析される。フレームは応答フレームであることを示すので、出口G F Hは、フレームを適切な出口インタフェース・キューに入れる。これによりフレームは、例えばD M 2 4 aなどのD Mインタフェースを通して、C P 5 0に接続されたポートを介して送ることができ、最終的にはC P 5 0が応答フレームを受け取る。

【0044】

次に図10および11に関して述べるように、G D H 1 7 aと、1次および2次の両ブレード上のG T H 1 9との間のメッセージ・フローを含む例示的なシナリオがある。

【0045】

したがって、図10に示す第4の例示的なシナリオ290では、ガイドッド「要求」フレームは入口G D H 1 7 aによって生成され、同じ(1次)ブレード上のG T H 1 9に向けられた組込みガイドッド・コマンドを含む。フレームを入口キューに入れることによって、フレームはまず1次ブレード20 a上の出口G F Hにルーティングされ、D A S Lインタフェース45 a、bを介して出口G F H 1 8 aにディスパッチされて受け取られる。出口G F H 1 8 aは、フレームを解析し、フレームがG T H 1 9に向けられていることを認め、したがってフレームを出口G T Hキューに入れる。これによりG T Hはフレームを受け取る。次いでG T Hはフレームを解析し、フレーム制御情報の相関関係子フィールドから、フレームの発信元がG D Hであることを決定する。次いでG T Hはフレームを処理し、フレームは、処理された後、出口廃棄キューに入れられることによって廃棄される。したがってG D Hに応答は送られない。

【0046】

図11に示す代替シナリオ300では、ガイドッド「要求」フレームはG D H 1 7 bによって生成され、2次ブレード20 b上のG T H 1 9に向けられた組込みガイドッド・コマンドを含む。このシナリオでは、フローは1次ブレードについて述べたものと同一だが、唯一の異なる点は、エラーの場合にG T H 1 9によって生成される否定応答メッセージが迎えるパスである。このメッセージは、D M _ Wインタフェース25を介して2次ブレード上の入口G F H 1 8 bにラップされる。入口G F Hは、フレームを入口キューに入れることによってフレームを1次ブレード20 aの出口G F H 1 8 aにルーティングし、フレームは最終的にC P 5 0に転送される。

【0047】

図16に、ガイドッド・フレームのフレーム制御情報60のフォーマットを示す。図16に示すように、このフレーム制御情報60は2つのワード構造を含み、各ワード61 a、61 bはそれぞれ4バイトである。第1のワード61 a中には、第1のデータ・バイト63が含まれ、この第1のデータ・バイト63は、フレームを制御フレームとして識別する数字、例えばO C Hである。第2のバイト65は、様々な処理/ルーティング標識を提供するためのビット・フラグを含み、これらの指示には次のものが含まれる。すなわち、フレームが要求フレームか応答フレームかを示すためにセットされる第1のr e s p / r e qフラグ70と、例えばC Pなどのソースに常に返信するか全く返信しないかを示すためにセットされる第2のa c k / n o a c kフラグ71と、フレームがどこからきたか(フレームのソース・アドレスおよびポート)をブレードが知るべきか、またはソース・アドレス情報がフレーム中に明示的に提供されているかを示すためにセットされるu s e / l e a r nフラグ72と、フレームの処理中にエラーが発生した場合だけ応答を提供すべきか(n e gにセットされる)、または処理エラーが発生したかどうかにかかわらず常に応答を返すようにソースが望んでいるかを示すためにa c k / n o a c kフラグ71と共に機能する第4のn e g / a l lフラグ73と、フレームがブレードのどちら側に向けられているか、すなわち入口側か出口側かを示すためにセットされる第5の入口/出口フラグ74と、ブレード内のどのピコプロセッサがフレームを処理すべきかを示す第6のG T H / G F Hフラグ75と、エラーが発生してもターゲット・プロセッサがガイドッド・フレーム・コマンドの処理を継続すべきか、またはエラーが発生した場合は処理を停止すべきかを示すためにセットされる、独立/連鎖コマンド処理インディケータである第7のi

nd / chnd フラグ 76 である。次のフィールド 66 は、フレームが送られるターゲット・ブレードのアドレスを含む 16 ビット・フィールドである。これはユニキャストまたはマルチキャストのアドレスを含むこともできることを理解されたい。すなわちガイドド・フレーム・プロセッサは、フレーム処理の一部として、ターゲット・ブレード・フィールド 66 のブレード・アドレスを、初期化の一部としてブレード中にプログラムされたアドレスと比較することになる。これらのアドレスが一致する場合は、フレームはそのブレードだけに向けられたものと見なされる。これらのアドレスが一致しない場合は、フレームはシステム内の別のブレードをターゲットとして見なされ（すなわちユニキャストまたはマルチキャスト・アドレス）、その結果、フレームがターゲットとすべきブレードを決定するためのインテリジェンスを有するスイッチ・ファブリック（図示せず）に渡される。

10

【0048】

フレーム制御情報 60 の第 2 のワード 61 b 中には、少なくとも、ガイドド・フレームが発信されたソース・ポートを示すための第 1 のデータ・バイト 67 が含まれる。述べたシナリオでは、このソース・ポートは、例えばデータ・ムーバ (DM) インタフェース・ポートを含む場合がある。追加の 2 つのデータ・バイト 68 が、発信フレームについてのソース・ブレード・アドレスを含む。これらのフレームがどのように構築されるかは、use / learn フラグ 72 によって決まることを理解されたい。例えば、「learn」フラグ 72 がセットされていない場合は、ブレードはソース・ポートおよびソース・ブレードのアドレス・フィールドを埋めないことになる。したがって、ブレード処理は、フレームがどこから来たかを知ることを担う。一般に、内部で（例えば GDH によって）生成されるガイドド・フレームは応答フレームを必要としない。すなわち ack / noack フィールド 71 は noack にセットされる。したがってこれらのフィールドはセットされない。しかし、例外 / エラーを例えば非送信請求応答フレームとして CP に報告すべき場合は、NP はソース・ブレードおよび CP のポートを知るように構成されることになる。

20

【0049】

ガイドド・フレームの構成に関してはさらに、提供される相関関係子情報は 2 ~ 5 バイト（図示せず）の任意の場所にわたるものとすることができ、応答フレームが例えば CP に返されたときにガイドド要求フレームの元のコマンドと相関されるように、固有の識別子を提供する。多くの組込みコマンドを有する単一のガイドド・フレームを処理する間には要求側 CP に複数の応答フレームが返される場合があり、各応答を相関させる必要があるので、このことは好ましい。

30

【0050】

ガイドド・フレームの残りの構成に関しては、NP ネットワーキング環境で NP インタフェース・デバイスを設定および制御するのに使用される指示または操作を含む 1 つまたは複数のコマンド・フィールドが設けられる。各コマンド・フィールドは、少なくとも 3 つのサブフィールドを含むことが好ましい。すなわち、1) 実施されるコマンドのタイプを識別するためのコマンド制御ワード、2) コマンドが作用する対象である NP インタフェース・デバイス内のアドレス位置を示すアドレス・フィールド、および 3) コマンド・タイプに関連するデータを含む可変長のオペランドのフィールドである。例えばコマンド・タイプは、レジスタをセットアップすること、またはメモリ位置をデータで埋めることを含む場合があり、アドレス情報サブフィールドは、オペランド・データを埋めるべきレジスタ / メモリ位置のアドレスを提供することになる。図 2 および 3 それぞれの GFH および GTH 構造内に示す機能ブロックを制御および設定するために実施できるコマンド・タイプには、多くのものがあることを理解されたい。ガイドド・フレーム中の残りのワードは、コマンドがそれ以上提供されないこと、およびガイドド・フレーム処理が停止すべきであることを示す、固有のフレーム終了区切りである。

40

【0051】

各図に示した前述のガイドド・フレーム・メッセージ・フローのシナリオは、これが

50

ら述べるように図 1 2 ~ 1 5 中のフロー・チャートによって実施される。

【 0 0 5 2 】

図 1 2 は、G F H 内でのガイドド・フレームの処理を示す流れ図 3 1 0 である。図 1 2 に示すように、ステップ 3 1 2 で、着信ガイドド・フレームが G F H によって受け取られる。ステップ 3 1 4 で、フレームは、図 1 3 に示すように G F H のガイドド・フレーム・プロセッサによって処理されるが、この処理でフレーム制御情報が解析されて、有効フレームかどうか、すなわちガイドド・フレーム識別子を含むかどうか決定される。ステップ 3 1 6 で有効フレームかどうかを決定し、有効フレームでない場合すなわちガイドド制御フレームでない場合は、ステップ 3 1 8 でフレームを廃棄し、プロセッサはステップ 3 2 0 の遊休状態に戻ることになる。ステップ 3 1 6 でフレームが有効ガイドド・フレームであると判定された場合は、ステップ 3 2 2 で、受け取ったフレームが応答フレームか要求フレームかを決定する。フレームが応答フレームであると決定された場合は、フレームは、ステップ 3 2 4 で G F H のガイドド・フレーム・プロセッサによって処理され、図 1 3 に示すガイドド・フレーム・プロセッサ・フローに従って進む。そうではなく、受け取ったガイドド・フレームが要求フレームである場合は、ステップ 3 2 6 でターゲット・ブレード・アドレスを決定する。次いでステップ 3 2 8 で、ターゲット・ブレードが、フレームを受け取った現在のブレードであるかどうかを決定する。フレーム中のターゲット・ブレードが現在のブレードに向けられていない場合は、ステップ 3 3 0 で、図 1 5 に示す代替処理フローに従ってフレームを処理する。そうではなく、受け取ったガイドド・フレームが現在のブレードに向けられている場合は、ステップ 3 3 3 で、フレームがブレードの出口側または入口側で処理されることになっていることを示す u p (入口) / d o w n (出口) フラグ・ビット 7 4 (図 1 6) が第 1 のフレーム制御ワード中でセットされているかどうかを決定する。フレームが出口側で処理されることになっている場合は、さらにステップ 3 3 6 で、フレームがブレードの出口側から来たか入口側から来たかを決定する。同様に、フレームが入口側で処理されることになっている場合は、さらにステップ 3 3 8 で、フレームがブレードの出口側から来たか入口側から来たかを決定する。ステップ 3 3 6 でフレームが出口側に向けられており出口側で発した場合は、ステップ 3 4 0 で、図 1 4 に示すガイドド・コマンド・プロセッサ・フローに従ってフレームを処理する。同様に、ステップ 3 3 8 でフレームが入口側に向けられており入口側で発した場合も、図 1 4 に示すガイドド・コマンド・プロセッサ・フローに従ってフレームを処理する。フレームが、受け取られた側と反対側に向けられている場合は、ステップ 3 4 5 に示すようにフレームを正しい側にルーティングする。

【 0 0 5 3 】

図 1 3 は、ガイドド・フレームのガイドド・フレーム処理を示す流れ図 3 5 0 である。図 1 3 に示すように、ステップ 3 5 2 で、G F H から着信ガイドド「応答」フレームを受け取る。次いでステップ 3 5 4 で、応答を送るべきか廃棄すべきかを示す a c k / n o a c k フラグ・ビット 7 1 (図 1 6) が第 1 のフレーム制御ワード中でセットされているかどうかを決定する。肯定応答が必要ない場合は、ステップ 3 5 6 でフレームを廃棄し、ステップ 3 5 8 でプロセッサは遊休状態に戻る。ステップ 3 5 4 で応答フレームを送るべきであると決定された場合は、ステップ 3 6 0 で、応答を常に送るべきかエラー発生時だけ送るべきかを示す n e g / a l l フラグ・ビット 7 3 (図 1 6) が第 1 のフレーム制御ワード中でセットされているかどうかを決定する。応答を常に送るべきである場合は、ステップ 3 6 2 で、C P がローカルに取り付けられているかリモートであるかを決定する。C P がローカルの場合は、ステップ 3 6 5 で応答フレームをローカル C P にディスパッチする。ステップ 3 6 0 でエラー発生時だけ応答を送るべきであると決定された場合は、ステップ 3 6 6 で、処理中にエラーが実際に発生したかどうかを決定する。エラーが発生した場合は、リモートかローカルかの C P にディスパッチするために処理はステップ 3 6 2 に進む。ステップ 3 6 6 でエラーが発生しなかったと決定された場合は、ステップ 3 6 8 でフレームを廃棄し、ステップ 3 7 0 でプロセッサは遊休状態に戻る。ステップ 3 6 2 に戻るが、応答フレームを向けるべき C P がリモートである場合は、ステップ 3 7 3 で

10

20

30

40

50

、フレームがどこで処理されたかを示すup(入口)/down(出口)フラグ・ビット74(図16)が第1のフレーム制御ワード中でセットされているかどうかを決定する。フレームが入口側にあると決定された場合は、ステップ375で、応答フレームをCPにディスパッチするためにDMインタフェースに転送する。そうではなくフレームが出口側にあると決定された場合は、ステップ378で、DM_W(ラップ)インタフェースを介して応答フレームを入口側にラップアラウンドし、ステップ380で、図12に示した着信ループ・フローに従って入口GFHがフレームを処理する。

【0054】

図14は、ガイドド・フレーム内のコマンドに対するガイドド・コマンド処理を示す流れ図400である。図14に示すように、ステップ402で、ガイドド・コマンドをどこで処理すべきかを示すGFH/GTHフラグ・ビット75(図16)が第1のフレーム制御ワード中でセットされているかどうかについて第1の決定を行う。ステップ404でビット・フラグ75がGFH処理を示す場合は、ステップ406でGFH処理を進め、ガイドド・フレームから次のコマンドを取り出す。ステップ408で、ガイドド・フレーム処理の終了を示すフレーム終了区切りがみられるかどうかを決定する。フレーム終了区切りがみられた場合は、ステップ410で制御をガイドド・フレーム・プロセッサに戻して、図13に示したガイドド・フレーム・プロセッサ・フローに従って処理を行う。ステップ408でフレーム終了区切りがみられなかった場合は、ステップ415で、現在のガイドド・コマンド・オペランドに従って必要なパラメータを、設定されるべきピコプロセッサ・レジスタにロードする。ステップ418で、GFH機能を(例えばプロシージャ・コール・アプリケーション・プログラムを介して)呼び出して設定セットアップを行い、ステップ420で、機能からステップ406のGFHガイドド・コマンド・プロセッサに制御を戻して、フレームから次のコマンドを処理する。ステップ404に戻るが、ビット・フラグ75がGTH処理を示すと決定された場合は、ステップ409でGTH処理を進め、ガイドド・フレームから次のコマンドを取り出す。次いでステップ411で、ガイドド・フレーム処理の終了を示すフレーム終了区切りがみられるかどうかを決定する。フレーム終了区切りがみられた場合は、ステップ413で制御をガイドド・フレーム・プロセッサに戻して、図13に示したガイドド・フレーム・プロセッサ・フローに従って処理を行う。ステップ411でフレーム終了区切りがみられなかった場合は、ステップ416で、現在のガイドド・コマンド・オペランドに従って必要なパラメータを、設定されるべきピコプロセッサ・レジスタに入力する。ステップ419で、GTH機能を呼び出して設定セットアップを行い、ステップ422で、機能からステップ409のGTHガイドド・コマンド・プロセッサに制御を戻し、フレームから次のコマンドを処理する。

【0055】

図12のステップ330に関して述べたように、1次ブレードで受け取られた要求フレームが2次ブレードまたは他の複数のブレードに向けられている場合は、プロセスは、図15のステップ452に示すように代替処理フロー450に進む。具体的には、図15は、ユニキャストでなく現在のブレードに向けられていないガイドド・フレームの処理450を示す。したがって、図15に示す処理は、1)別のブレードをターゲットとするユニキャスト・フレーム、2)現在のブレードを除く他のブレードをターゲットとするマルチキャスト・フレーム、3)現在のブレードを含めて他のブレードをターゲットとするマルチキャスト・フレームとして分類されるフレームを対象とする。図11に示すように、ステップ454で、現在のフレームが3つの範疇のどれにあてはまるかを決定する。フレームが範疇1)および2)、すなわち別のブレードをターゲットとするユニキャスト・フレームまたは現在のブレードを除く他のブレードをターゲットとするマルチキャスト・フレームであると決定された場合は、処理はステップ456に進み、フレームをネットワーク・プロセッサ・スイッチング・ファブリックに転送して、宛先ブレードの出口側にディスパッチする。次いでステップ458に示すように、各ブレードの出口GFHは、図12に示した着信ループ・フローに従ってフレームを処理する。ステップ454でフレームが

10

20

30

40

50

範疇 3)、すなわち現在のブレードを含めて他のブレードをターゲットとするマルチキャスト・フレームであると決定された場合は、処理はステップ 460 に進み、フレームをネットワーク・プロセッサ・スイッチング・ファブリックに転送して、宛先ブレードの出口側と現在のブレードの出口側とを含めた出口側にディスパッチする。次いでステップ 462 に示すように、各ブレードの出口 GFH は、図 13 に示した着信ループ・フローに従ってフレームを処理する。

【0056】

参考文献：

米国特許出願第 09 / 384691 号

米国特許出願第 09 / 330968 号

10

【図面の簡単な説明】

【図 1】 1 つまたは複数のネットワーク・プロセッサ・デバイスを実装する例示的なネットワーク処理環境を示す図である。

【図 2】 本発明の一実施形態による、図 1 に示したネットワーク・プロセッサ・デバイスのガイドド・フレーム・ハンドラ・コンポーネントのブロック図である。

【図 3】 本発明の一実施形態による、図 1 に示したネットワーク・プロセッサ・デバイスのガイドド・テーブル・ハンドラ・コンポーネントのブロック図である。

【図 4】 単一ブレード上での CP と GFH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 5】 単一ブレード上での GDH と GFH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

20

【図 6】 2 次ブレード上での CP と GFH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 7】 2 次ブレード上での GDH と GFH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 8】 単一ブレード上での CP と GTH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 9】 2 次ブレード上での CP と GTH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 10】 単一ブレード上での CLP と GTH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

30

【図 11】 2 次ブレード上での CLP と GTH との間のガイドド・フレーム・フローを表す例示的なメッセージング・シナリオを示す図である。

【図 12】 本発明の一実施形態による、GFH 内のガイドド・フレームの処理を示す流れ図 310 である。

【図 13】 本発明の一実施形態による、ガイドド・フレームに対するガイドド・フレーム処理を示す流れ図 350 である。

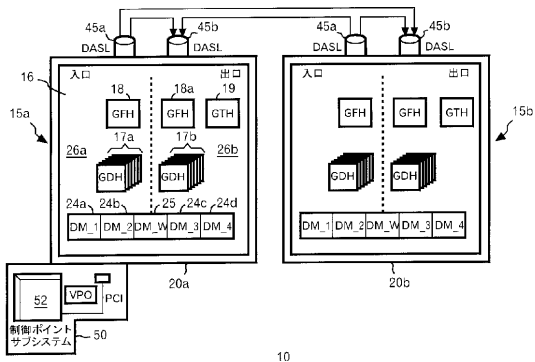
【図 14】 本発明の一実施形態による、ガイドド・フレーム内のコマンドに対するガイドド・コマンド処理を示す流れ図 400 である。

【図 15】 本発明の一実施形態による、ユニキャストでなく現在のブレードに向けられていないガイドド・フレームの処理を示す流れ図 450 である。

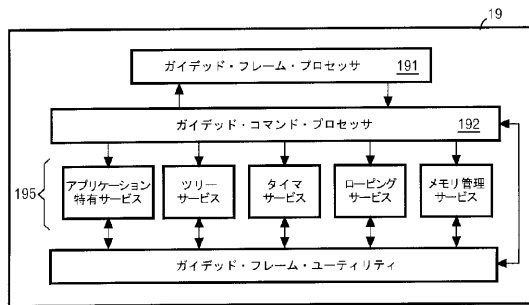
40

【図 16】 本発明の一実施形態による、ガイドド・フレームに含まれるフレーム制御情報を示す詳細な図である。

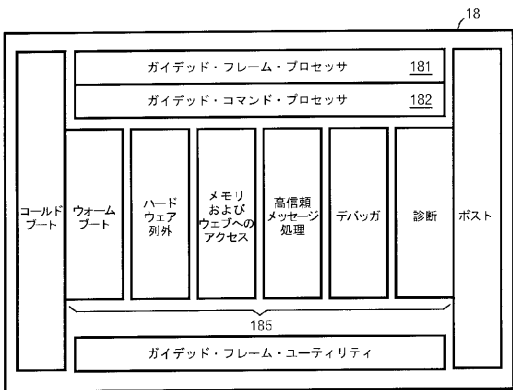
【 図 1 】



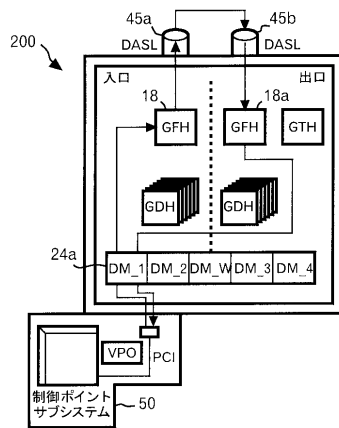
【 図 3 】



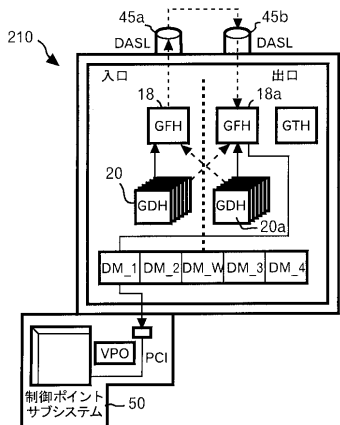
【 図 2 】



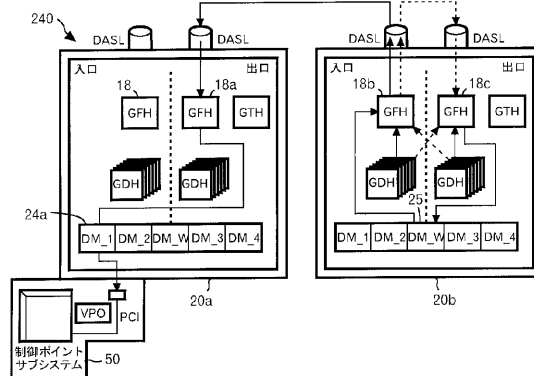
【 図 4 】



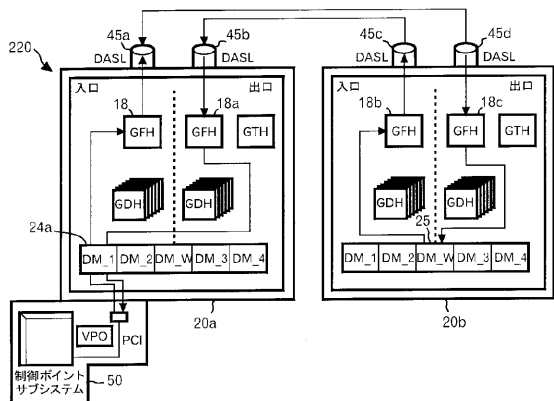
【 図 5 】



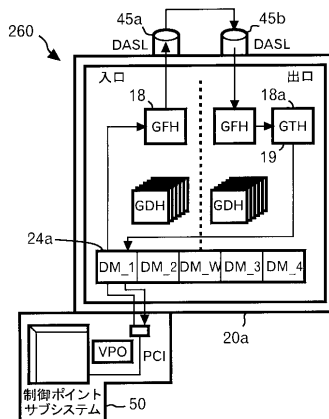
【 図 7 】



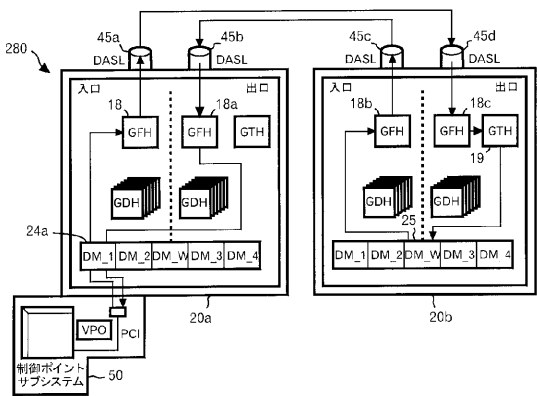
【 図 6 】



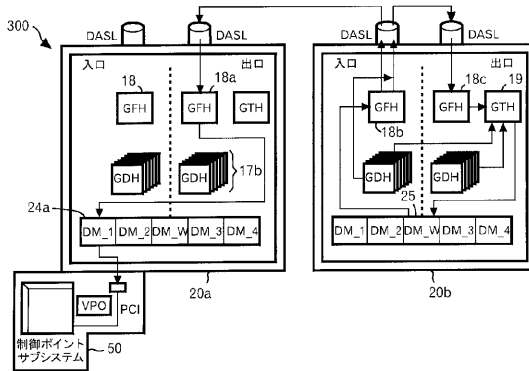
【 図 8 】



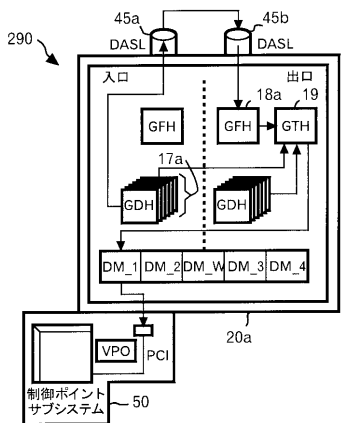
【図9】



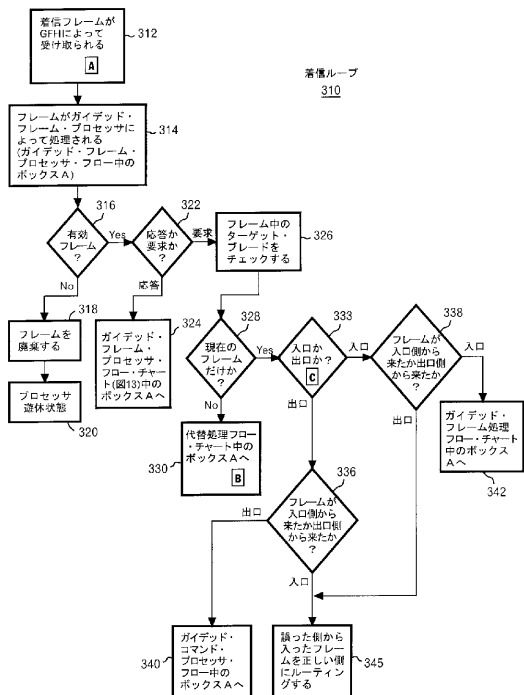
【図11】



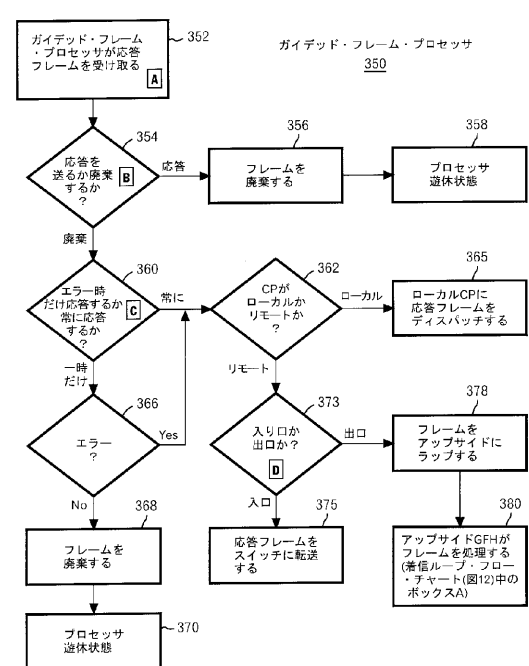
【図10】



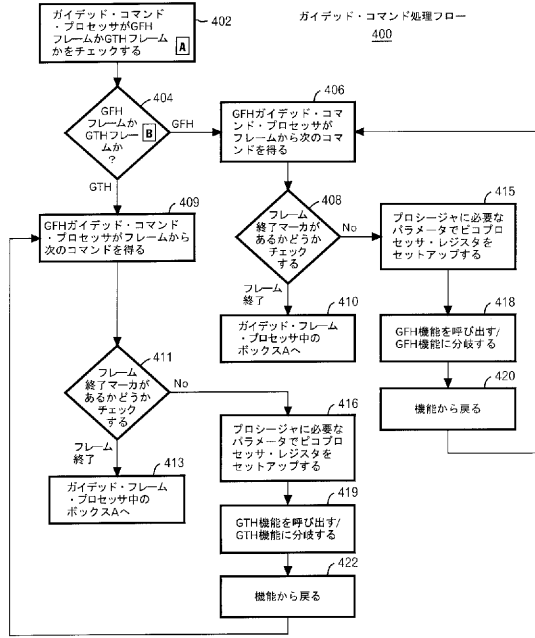
【図12】



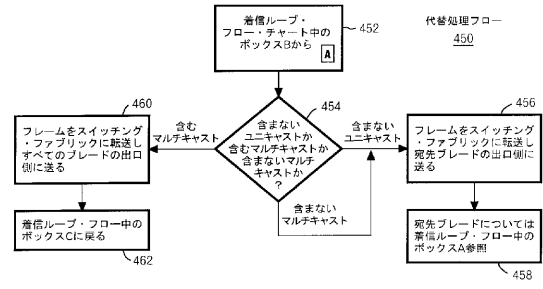
【図13】



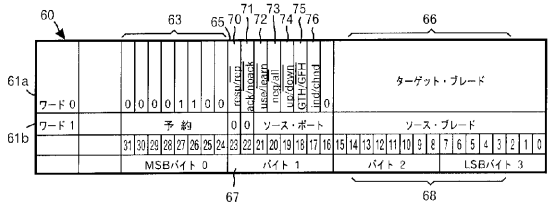
【図14】



【図15】



【図16】



フロントページの続き

- (72)発明者 バッソ、クロード
アメリカ合衆国27613 ノースカロライナ州ローリー パーシー・コート 7604
- (72)発明者 ガッロ、アンソニー、マッテオ
アメリカ合衆国27502 ノースカロライナ州エーベックス コーシャム・ドライブ 3308
- (72)発明者 ヘッデス、マルコ
アメリカ合衆国27615 ノースカロライナ州ローリー グランド・マナー・コート 4109
#308
- (72)発明者 ハリハラン、シータ
アメリカ合衆国27612 ノースカロライナ州ローリー ビーバーブルック・ロード 5020
-104
- (72)発明者 ラオ、スリハール
アメリカ合衆国27612 ノースカロライナ州ローリー ビーバーブルック・ロード 5020
アパートメント 204
- (72)発明者 ローバー、ソニア、キアング
アメリカ合衆国27516 ノースカロライナ州チャペル・ヒル パークサイド・サークル 20
9

審査官 横山 佳弘

- (56)参考文献 特開平06-197111(JP,A)
特開平08-130538(JP,A)
特開平09-135273(JP,A)
特開平08-251657(JP,A)
特開平07-226758(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 13/10
H04L 12/56
H04L 12/46
H04L 12/28