(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4066382号
(P4066382)
(45)発行日　平成20年3月26日(2008.3.26)　(24)登録日　平成20年1月18日(2008.1.18)

(51)Int.Cl.　F I
H04L 12/56 (2006.01)　H04L 12/56　200Z

請求項の数 3　(全 40 頁)

(21)出願番号	特願2005-282931(P2005-282931)	(73)特許権者	390009531
(22)出願日	平成17年9月28日(2005.9.28)		インターナショナル・ビジネス・マシーンズ・コーポレーション
(62)分割の表示	特願2001-520649(P2001-520649)の分割		INTERNATIONAL BUSINESS MASCHINES CORPORATION
原出願日	平成12年8月24日(2000.8.24)		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
(65)公開番号	特開2006-20371(P2006-20371A)	(74)代理人	100086243
(43)公開日	平成18年1月19日(2006.1.19)		弁理士　坂口　博
審査請求日	平成17年9月28日(2005.9.28)	(74)代理人	100091568
(31)優先権主張番号	09/384,692		弁理士　市位　嘉宏
(32)優先日	平成11年8月27日(1999.8.27)		
(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・スイッチ及びコンポーネント及び操作方法

(57)【特許請求の範囲】

【請求項１】

制御点プロセッサと、該制御点プロセッサに動作的に接続されるネットワーク・プロセッサとを含む通信装置において、前記ネットワーク・プロセッサの制御の下で、レジスタ及びメモリをアクセスすることを可能にする情報を有するガイド・フレームを経路指定する方法であって、

前記制御点プロセッサ内に配置される制御点機能を使用し、ガイド・フレームを生成するステップと、

前記制御点プロセッサ内のデバイス・ドライバを使用し、前記ガイド・フレームを、前記ネットワーク・プロセッサに関連付けられる複数のメディア・インタフェースの１つに送信するステップと、

前記メディア・インタフェース内の媒体アクセス制御ハードウェアを使用し、前記送信されたガイド・フレームを回復するステップと、

回復された前記ガイド・フレームをメモリに記憶するステップと、

記憶された前記ガイド・フレームを、前記ガイド・フレーム内で識別されるエンティティに経路指定するステップと

を含む方法。

【請求項２】

前記ガイド・フレームにより伝搬される命令に従い、前記ガイド・フレームを前記エンティティにより処理するステップと、

前記ガイド・フレームにより伝搬される情報により要求される場合、処理された前記ガイド・フレームを前記制御点機能に戻すステップと
を含む、請求項１記載の方法。
【請求項３】
前記ガイド・フレームをネットワーク・ルーティング情報によりカプセル化するステップを含む、請求項１記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、様々なタイプ及び能力の情報処理システムまたはコンピュータをリンクするために使用される通信ネットワーク装置、及びこうした装置のコンポーネントに関する。特に、本発明は、こうした装置をアセンブルするのに有用なスケーラブル・スイッチ装置及びコンポーネントに関する。本発明はまた、改善された多機能インタフェース装置、並びに、こうした装置とメディア速度ネットワーク・スイッチを提供する他の要素との組み合わせに関する。本発明は更に、ネットワーク・スイッチのデータ・フロー処理能力を改善する、こうした装置の操作方法に関する。
【背景技術】
【０００２】
以下の説明は、ネットワーク・データ通信、並びにこうした通信ネットワークで使用されるスイッチ及びルータに関する知識を前提とする。特に、ネットワーク動作をレイヤに分割するネットワーク・アーキテクチャのＩＳＯモデルに精通していることを前提とする。ＩＳＯモデルにもとづく典型的なアーキテクチャは、物理パスまたはメディアであるレイヤ１（ときに"Ｌ１"と示される）から始まり、信号は、上流のレイヤ２、３、４などを通じてレイヤ７に至る。レイヤ７は、ネットワークにリンクされるコンピュータ・システム上で実行されるアプリケーション・プログラミング・レイヤである。本明細書では、Ｌ１、Ｌ２などは、ネットワーク・アーキテクチャの対応レイヤを指し示すものとする。本開示はまた、こうしたネットワーク通信においてパケット及びフレームとして知られる、ビット・ストリングの基本的理解を前提とする。
【０００３】
今日のネットワーク化された世界では、帯域幅が重要な資源である。インターネットや他の新たなアプリケーションの登場による、ネットワーク・トラフィックの増加が、ネットワーク・インフラストラクチャの能力を圧迫している。こうした流れに追随するために、様々な組織や機関が、トラフィックの増加や音声とデータのコンバージェンスをサポートするための、より優れた技術や方法を検討している。
【０００４】
ネットワーク・トラフィックの今日の劇的な増加は、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加、及び新たなアプリケーションなどによる。インターネットは電子商取引の爆発的な成長により、ときにネットワーク・バックボーンにサポート不能な負荷を課する。データ・トラフィック量の増加の最も重要な１原因は、それが初めて音声トラフィックを超えることにある。電子メール、データベース・アクセス、及びファイル転送などを含む、リモート・アクセス・アプリケーションの要求の増加が、更にネットワークを圧迫している。
【０００５】
音声とデータのコンバージェンスは、将来のネットワーク環境を定義する上で重要な役割を果たす。現在、インターネット・プロトコル（ＩＰ）ネットワークを介するデータの伝送は無料である。音声伝送は自然に最廉価の経路を通るので、音声は必然的にデータとコンバージされる。ボイス・オーバＩＰ（ＶｏＩＰ）、ボイス・オーバＡＴＭ（ＶｏＡＴＭ）、ボイス・オーバ・フレーム・リレー（ＶｏＦＲ）などの技術は、この変化の速い市場においてコスト効率の高い技術と言える。しかしながら、これらの技術への移行を可能にするために、業界は音声のサービス品質（ＱｏＳ）を保証し、データ回線上での音声転

送のためにどのように課金すべきかを決定なければならない。１９９６年のTelecommunication Deregulation Actが、更にこの環境を複雑化する。この法規は、選択音声プロトコルすなわちＡＴＭと、選択データ・プロトコルすなわちＩＰとの間の共生関係を強化する。

【０００６】

新たな製品及び機能が使用可能になると、レガシー・システム（システム資産）を統合することが、組織にとって重大な問題となる。既存の設備及びソフトウェアへの投資を保護するために、組織は彼らの現在の運用を混乱させることなく、新たな技術への移行を可能にする解決策を要求する。

【０００７】

サービス・プロバイダにとって、ネットワーク障害を取り除くことが先決である。ルータはしばしば、これらの障害の原因となる。しかしながら、ネットワーク輻輳は一般に、しばしば帯域幅問題と誤診され、より高い帯域幅を求めることにより取り組まれる。今日、メーカはこの難題を認識しつつある。彼らはネットワーク・プロセッサ技術に移行することにより、帯域幅資源をより効率的に管理し、ルータ及びネットワーク・アプリケーション・サーバで一般に見いだされる高度データ・サービスを有線速度で提供しつつある。これらのサービスにはロード・バランシング、ＱｏＳ、ゲートウェイ、ファイアフォール、セキュリティ、及びウェブ（Ｗｅｂ）キャッシングが含まれる。

【０００８】

リモート・アクセス・アプリケーションでは、性能、要求時帯域幅、セキュリティ及び認証が最優先の部類に入る。ＱｏＳ及びＣｏＳの統合、統合音声処理、及びより複雑なセキュリティ・ソルーションが、将来のリモート・アクセス・ネットワーク・スイッチの設計を方向付けるであろう。更に、リモート・アクセスは、ＩＳＤＮ、Ｔ１、Ｅ１、ＯＣ－３乃至ＯＣ－４８、ケーブル、及びｘＤＳＬモデムなど、より多くの物理メディアに対応しなければならない。

【０００９】

業界コンサルタントはネットワーク・プロセッサ（ここでは"ＮＰ"とも称す）を、次に挙げる機能の１つ以上を実行できるプログラマブル通信集積回路として定義した。すなわち、

１）パケット分類：アドレスやプロトコルなどの既知の特性にもとづく、パケットの識別。

２）パケット変更：ＩＰ、ＡＴＭ、または他のプロトコルに従う、パケットの変更（例えばＩＰのヘッダ内の存続時間（time-to-live）フィールドの更新）。

３）キュー／ポリシ管理：特定のアプリケーションにおけるパケットのパケット・キューイング、デキューイング、及びスケジューリングの設計方針の反映。

４）パケット転送：スイッチ・ファブリックを介するデータの伝送及び受信、及び適切なアドレスへのパケットの転送または経路指定。

【００１０】

この定義は早期ＮＰの基本フィーチャの正確な記述であるが、ＮＰの完全な潜在的能力及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、従来ソフトウェアで処理されたネットワーク・タスクをハードウェアで実行することにより、帯域幅を増加し、広範囲のアプリケーションにおける待ち時間問題を解決することができる。更に、ＮＰは並列分散処理及びパイプライン処理設計などのアーキテクチャを通じて、速度の改善を提供できる。これらの機能は効率的な検索エンジンを可能にし、スループットを向上させ、複雑なタスクの迅速な実行を提供する。

【００１１】

ネットワーク・プロセッサは、ＰＣにおけるＣＰＵのように、ネットワークの基本ネットワーク構築ブロックになるものと期待される。ＮＰにより提供される一般的な機能は、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、及びＩＰパケット処理及び学習機能である。ＮＰはＩＳＯレイヤ２乃至５をターゲットとし、ネットワー

ク特定タスクを最適化するように設計される。

【００１２】

プロセッサ・モデルＮＰは、複数の汎用プロセッサ及び特殊論理を組み込む。提供者はこの設計に傾倒することにより、変更に適宜に且つ低廉に対応できる、スケーラブルで柔軟性のあるソリューションを提供しようとしている。プロセッサ・モデルＮＰは、低レベルの統合において分散処理を可能にし、より高度なスループット、柔軟性及び制御を提供する。プログラマビリティが、新たなＡＳＩＣ設計を要求することなく、新たなプロトコル及び技術への容易な移行を可能にする。プロセッサ・モデルＮＰにより、ＮＥＶは払戻し不能なエンジニアリング・コストの低減、及び製品化までの時間の短縮の利益を甘受することができる。

【発明の開示】

【発明が解決しようとする課題】

【００１３】

本発明の第１の目的は、転送されるデータの処理速度を向上する一方で、サポート機能を各種の潜在的要求にサイズ変更可能な、データ通信ネットワークで使用されるスケーラブル・スイッチ・アーキテクチャを提供することである。

【００１４】

本発明の別の目的は、単一基板上に集積化され、レイヤ２、レイヤ３、レイヤ４及びレイヤ５を含むフレームのメディア速度切替えを提供するように協働する、複数のサブアセンブリを含むインタフェース装置またはネットワーク・プロセッサ（これらの用語は互換性をもって使用される）を提供することである。

【課題を解決するための手段】

【００１５】

前記第１の目的は、関連処理ユニットの作業負荷から、従来に比較してより多くのデータ処理を削除するコンポーネント、及びコンポーネントの集合を提供することにより実現される。

【００１６】

インタフェース装置は、ワーク・グループ・スイッチとして、第１レベルの機能を提供する独立型のソリューションとして使用されるか、またはより高度な機能のワーク・グループ・スイッチを提供する相互接続ソリューションとして使用されるか、或いは、スイッチング・ファブリック装置との協働により、更に機能向上に向けてスケーリングされる。

【発明を実施するための最良の形態】

【００１７】

[数１]

$$Y/\overline{X}$$

は、本明細書ではＹ／Ｘバーと記載する。

【００１８】

本発明は、本発明の好適な実施例を示す添付の図面を参照して、以下で詳述されるが、説明の冒頭に当たり、当業者であれば、本発明の好適な結果を達成するように、ここで述べる発明を変更することができよう。従って、以下で述べる説明は本発明を制限するものではなく、当業者に対する広範囲の教示の開示として理解されるべきである。

【００１９】

ここで開示される装置はスケーラブルで、デスクトップまたはワークグループ・スイッチを相互接続し、こうしたスイッチをネットワーク・バックボーンに統合し、バックボーン・スイッチング・サービスを提供するように機能することができる。この装置はレイヤ２、レイヤ３、及びレイヤ４＋転送をハードウェアでサポートできる。特定の形態の装置は、デスクトップまたはワークグループ・スイッチ統合のために設計され、また他のもの

はコア・バックボーン・スイッチとして設計される。

【００２０】

装置のために使用されるアーキテクチャは、インタフェース装置またはネットワーク・プロセッサ・ハードウェア・システムと、制御点上で実行されるソフトウェア・ライブラリとにもとづき、これについては本明細書の中で詳述される。インタフェース装置またはネットワーク・プロセッサ・サブシステムは、Ｌ２、Ｌ３及びＬ４＋プロトコル・ヘッダの構文解析及び変換のために設計される高性能フレーム転送エンジンである。これはプロトコルがハードウェアの使用により、より高速にスイッチされることを可能にする。インタフェース装置またはネットワーク・プロセッサ・サブシステムは、ボックスを通じて高速パスを提供する一方、ソフトウェア・ライブラリ及び制御点プロセッサは、高速パスを維持するために必要な管理及びルート発見機能を提供する。制御点プロセッサ及びそれ上で動作するソフトウェア・ライブラリは、一緒にシステムの制御点（ＣＰ）を定義する。ＣＰでは、透過型ブリッジング及びＯＳＰＦなどの、実際のブリッジング及びルーティング・プロトコルが実行される。これはシステムの低速パスとも呼ばれる。

【００２１】

ここで開示される装置は、マルチレイヤ転送をハードウェアでサポートするが、装置はＬ２専用スイッチとしても動作でき、開示される最も単純な形態では、これがそのデフォルト動作モードである。各ポートは単一のドメインに配置され、任意の装置が任意の他の装置と通信することを可能にする。装置はＬ２において構成可能であり、システム管理者に次のようなフィーチャ、すなわち、ポートを別々のドメインまたはトランクにグループ化したり、仮想ＬＡＮ（ＶＬＡＮ）セグメントや、ブロードキャスト及びマルチキャスト・トラフィックを制御するフィルタを構成するなどの能力を提供する。

【００２２】

このスケーラブル装置は多くの利点を有する。第１に、これはシステム管理者に、Ｌ２で使用されるハードウェアと同一のものを使用して、同一の速度で、ＩＰ及びＩＰＸトラフィックのＬ３転送及びルーティングを構成する能力を提供する。第２に、構内建物を相互接続するために、外部ルータを使用する必要性を排除する一方で、同時に性能を向上させる。第３に、Ｌ２／Ｌ３サービスの管理を単純化または結合することにより、これらを単一の制御点に組み込む。最後に、スケーラブル装置は付加価値機能をＬ４＋機能と共に提供し、これはサーバ間での負荷平準化を目的として、主幹業務のアプリケーション及びネットワーク・ディスパッチャをサポートするために、システム管理者が異なるトラフィック分類を割当てる能力を提供する。

【００２３】

スケーラブル装置は、インタフェース装置またはネットワーク・プロセッサを使用するモジュラ・ユニットとして、または制御点（ＣＰ）として、或いはその基本ビルディング・ブロックとしての、オプションのスイッチング・ファブリック装置として設計される。インタフェース装置は好適には、Ｌ２／Ｌ３／Ｌ４＋高速パス転送サービスを提供し、ＣＰは高速パスを維持するために必要な管理及びルート発見機能を提供する。３つ以上のインタフェース装置サブシステムが一緒に結合される場合には、オプションのスイッチング・ファブリック装置が使用される。オプションのスイッチング・ファブリック装置は、１９９１年４月１６日発行の米国特許第５００８８７８号"High Speed Modular Switching Apparatus for Circuit and Packet Switched Traffic"で開示される。

【００２４】

装置は、プリント回路基板要素を用いて組み立てられるものと見込まれる。これはここでは"ブレード"とも称される。プリント回路基板要素は回路要素を実装され、装置ハウジング内に設けられるコネクタに収容される。同様の装置は"オプション・カード"としても知られる。装置は、適切なコネクタ及びバックプレーン電気接続が設けられる条件の下で、ブレードが様々なシャーシまたはハウジングの間で交換されることを考慮する。全てのブレード上で見いだされる基本コンポーネントは、キャリア・サブシステムである。キャリア・サブシステムから始まり、３つのタイプのブレードが生成される。第１のタイプは

ＣＰ専用ブレードであり、これはキャリア・サブシステム及びＣＰサブシステムから成る。ＣＰ専用ブレードは、冗長性が一番の関心事である製品に対して主に使用される。第２のタイプはＣＰ＋メディア・ブレードであり、これはキャリア・サブシステム、ＣＰサブシステム、及び１対３メディア・サブシステムから成る。ＣＰ＋メディア・ブレードは、冗長性よりもポート密度が重要視される製品に対して、主に使用される。第３のタイプはメディア・ブレードであり、これはキャリア・サブシステムと１対４メディア・サブシステムとから成る。メディア・ブレードは任意のシャーシ内で使用され、使用されるメディア・サブシステムのタイプは構成可能である。

【００２５】

ブレード管理は障害検出、電力管理、新たな装置の検出、初期化及び構成を含む。この管理は様々なレジスタ、入出力信号、及びＣＰとキャリア・サブシステムとの間で通信するために使用されるガイド・セル・インタフェースを用いて行われる。しかしながら、シャーシと異なり、全てのブレード上には、プログラマブル装置及びメモリが存在する。プログラマビリティは、ブレードのタイプに依存する。ＣＰサブシステムがブレード上に存在する場合、ＣＰ及びキャリア・サブシステムの両方がプログラマブルである。メディア・サブシステムもプログラマブルであるが、キャリア・サブシステムを通じて間接的にプログラマブルなだけである。

【００２６】

高機能製品では更に、スイッチング・ファブリック装置サブシステムを含むスイッチ・ブレードが存在する。このブレードの管理は、故障検出、電力管理、新たな装置の検出、及び初期化を含む。この管理は、ＣＰサブシステム内にマップされる様々なレジスタ及び入出力信号を用いて行われる。

【００２７】

最も単純な形態では、本発明により考慮されるスイッチ装置は制御点プロセッサと、制御点プロセッサに動作可能なように接続されたインタフェース装置とを有する。好適には、ここで開示されるように、インタフェース装置（ネットワーク・プロセッサとして知られる）は、半導体基板と、基板上に形成される複数のインタフェース・プロセッサと、インタフェース・プロセッサによりアクセスされる命令を記憶する、前記基板上に形成される内部命令メモリと、インタフェース装置を通過し、インタフェース・プロセッサによりアクセスされるデータを記憶する、前記基板上に形成される内部データ・メモリと、複数の入出力ポートとを有する、単一の超大規模集積回路（ＶＬＳＩ）装置またはチップである。インタフェース・プロセッサはここではときに、ピコプロセッサまたは処理ユニットと称される。提供されるポートは、内部データ・メモリを外部データ・メモリに接続する少なくとも１つのポートと、インタフェース・プロセッサの指示の下で、インタフェース装置を通過するデータを外部ネットワークと交換する少なくとも２つの他のポートとを含む。制御点プロセッサはインタフェース装置と協働して、インタフェース・プロセッサにより実行される命令を命令メモリにロードする。そして、これらの命令が実行されて、入出力ポート間でのデータの交換、及びデータ・メモリを介するデータの流れを指示する。

【００２８】

ここで開示されるネットワーク・プロセッサは、それが組み込まれるスイッチ・アセンブリとは切り離して、本発明に関わる。更に、ここで開示されるネットワーク・プロセッサは、ここで述べられるその要素内に、ここでは詳細に述べられない他の発明を有するものとみなされる。

【００２９】

図１は、基板１０と、基板上に集積化される複数のサブアセンブリとを含むインタフェース素子チップのブロック図である。サブアセンブリは、アップサイドすなわち上流側構成と、ダウンサイドすなわち下流側構成とに編成される。本明細書では、"アップサイド"は、ネットワークから本明細書で開示される装置に送られるデータ・フローを指し示し、"ダウンサイド"は、逆に装置から、この装置によりサービスされるネットワーク・サービスに送られるデータを指し示す。データ・フローは反復構成に従う。結果的に、アップサ

イド・データ・フロー及びダウンサイド・データ・フローが存在する。アップサイドのサブアセンブリには、エンキュー－デキュー－スケジューリング・アップ（ＥＤＳ－ＵＰ）論理１６、多重化ＭＡＣアップ（ＰＰＭ－ＵＰ）１４、スイッチ・データ・ムーバ・アップ（ＳＤＭ－ＵＰ）１８、システム・インタフェース（ＳＩＦ）２０、データ・アライン・シリアル・リンクＡ（ＤＡＳＬＡ）２２、及びデータ・アライン・シリアル・リンクＢ（ＤＡＳＬＢ）２４が含まれる。データ・アライン・シリアル・リンクは、１９９９年６月１１日出願の米国特許出願第０９／３３０９６８号"High Speed Parallel/Serial Link for Data Communication"で詳細に述べられている。ここで開示される本発明の装置の好適な形態はＤＡＳＬリンクを使用するが、本発明は、特にデータ・フローがＶＬＳＩ構造内にあるように制限される場合などのように、かなり高いデータ・フロー・レートを達成するために、他の形態のリンクが使用されることも考慮する。

【００３０】

ダウンサイドのサブアセンブリには、ＤＡＳＬＡ２６、ＤＡＳＬＢ２８、ＳＩＦ３０、ＳＤＭ－ＤＮ３２、ＥＤＳ－ＤＮ３４、及びＰＰＭ－ＤＮ３６が含まれる。チップは更に複数の内部ＳＲＡＭ、トラフィック管理スケジューラ４０、及び組み込みプロセッサ・コンプレックス（ＥＰＣ）１２を含む。インタフェース装置３８は、イーサネット（Ｒ）フィジカル（ＥＮＥＴ　ＰＨＹ）またはＡＴＭフレーマなどの、任意の好適なＬ１回路である。インタフェースのタイプは、部分的に、チップが接続されるネットワーク・メディアにより決定される。複数の外部ＤＲＡＭ及びＳＲＡＭが設けられ、チップにより使用される。

【００３１】

ここでは特に、関連スイッチング及びルーティング装置の外部の一般データ・フローが、建物内に導入された配線やケーブルなどの電気導体を通過するネットワークが開示されるが、本発明は、ここで開示されるネットワーク・スイッチ及びコンポーネントが、無線環境において使用されることも同様に考慮する。例として、ここで述べられるメディア・アクセス制御（ＭＡＣ）要素は、ここで述べられる要素を直接ワイヤレス・ネットワークにリンクする機能を有する、好適な無線周波数要素により置換されてもよく、例えば、既知のシリコン・ゲルマニウム技術などが使用される。こうした技術が適切に使用される場合には、無線周波数要素は当業者により、ここで開示されるＶＬＳＩ構造内に統合される。或いは、無線周波数、または赤外線応答装置などの他のワイヤレス応答装置が、ワイヤレス・ネットワーク・システム用のスイッチ装置を実現する他の要素と一緒に、ブレード上に実装されてもよい。

【００３２】

図中の矢印は、インタフェース装置内のデータの一般的な流れを示す。イーサネット（Ｒ）ＭＡＣから受信されるフレームは、ＥＤＳ－ＵＰにより、内部データ・ストア・バッファに配置される。これらのフレームは、通常データ・フレームまたはシステム制御ガイド・フレーム（Guided Frame）として示され、ＥＰＣ（図１）にエンキューされる。ＥＰＣは、最大Ｎフレーム（Ｎ＞１）を並列に処理できるＮ個のプロトコル・プロセッサを含む。１０プロトコル・プロセッサ（図１４）の実施例では、１０個のプロトコル・プロセッサの内の２個が特殊化される。すなわち、１つはガイド・フレームを処理するように（ガイド・セル・ハンドラ（ＧＣＨ））、また他は制御メモリ内にルックアップ・データを作成するように（汎用ツリー・ハンドラ（ＧＴＨ））、特殊化される。図１３に示されるように、ＥＰＣは、新たなフレームをアイドル・プロセッサに突き合わせるディスパッチャと、フレーム・シーケンスを保持する完了（completion）ユニットと、１０個の全プロセッサにより共用される共通命令メモリと、フレーム分類を決定する分類子ハードウェア補助機構及びフレームの開始命令アドレスを決定する支援をするコプロセッサと、フレーム・バッファの読出し及び書込み操作を制御する入口側（以降イングレス（ingress）と称す）及び出口側（以降イーグレス（egress）と称す）データ・ストア・インタフェースと、１０個のプロセッサが制御メモリを共用することを可能にする制御メモリ・アービタと、ウェブ（Ｗｅｂ）制御と、内部インタフェース装置データ構造へのデバッグ・アクセ

スを可能にするアービタ及びインタフェースと、その他のハードウェア構造を含む。

【００３３】

ガイド・フレームは、ＧＣＨプロセッサが使用可能になると、ディスパッチャによりＧＣＨプロセッサに送信される。レジスタ書込み、カウンタ読出し、イーサネット（Ｒ）ＭＡＣ構成変更など、ガイド・フレーム内でエンコードされる操作が実行される。ＭＡＣまたはＩＰエントリの追加など、ルックアップ・テーブル変更が、メモリ読出しや書込みなどの制御メモリ操作のために、ルックアップ・データ・プロセッサに渡される。ＭＩＢカウンタ読出しなどの幾つかのコマンドは、応答フレームの作成、及び適切なインタフェース装置上の適切なポートへの転送を要求する。ときとして、ガイド・フレームはインタフェース装置のイーグレス側に合わせてエンコードされる。これらのフレームは、照会されるインタフェース装置のイーグレス側に転送され、これが次にエンコード操作を実行し、適切な応答フレームを作成する。

【００３４】

データ・フレームは、次の使用可能なプロトコル・プロセッサにディスパッチされ、そこでフレーム参照が実行される。フレーム・データは、分類子ハードウェア補助機構（ＣＨＡ）エンジンからの結果と一緒に、プロトコル・プロセッサに渡される。ＣＨＡはＩＰまたはＩＰＸを構文解析する。結果がツリー構造探索アルゴリズム及び開始共通命令アドレス（ＣＩＡ）を決定する。サポートされるツリー構造探索アルゴリズムは、固定マッチ・ツリー（レイヤ２イーサネット（Ｒ）ＭＡＣテーブルなど、正確なマッチを要求する固定サイズ・パターン）、最長プレフィックス・マッチ・ツリー（サブネットＩＰ転送など、可変長マッチを要求する可変長パターン）、及びソフトウェア管理ツリー（フィルタ規則で使用されるような、範囲またはビット・マスク・セットのいずれかを定義する２つのパターン）を含む。

【００３５】

ルックアップ（探索）は、各プロトコル・プロセッサの一部である、ツリー構造探索エンジン（ＴＳＥ）コプロセッサの支援により実行される。ＴＳＥコプロセッサは制御メモリ・アクセスを実行し、プロトコル・プロセッサを解放して実行を継続する。制御メモリは全てのテーブル、カウンタ、及びピココードにより必要とされる他のデータを記憶する。制御メモリ操作は、１０個のプロセッサ・コンプレックスの間のメモリ・アクセスを調停する制御メモリ・アービタにより管理される。

【００３６】

フレーム・データはデータ・ストア・コプロセッサを通じてアクセスされる。データ・ストア・コプロセッサは、基本データ・バッファ（フレーム・データの最大８個の１６バイト・セグメント）と、スクラッチ・パッド・データ・バッファ（フレーム・データの最大８個の１６バイト・セグメント）と、データ・ストア操作のための幾つかの制御レジスタとを含む。一旦マッチが見いだされると、イングレス・フレーム変更がＶＬＡＮヘッダ挿入またはオーバレイを含み得る。この変更は、インタフェース装置プロセッサ・コンプレックスにより実行されるのではなく、ハードウェア・フラグが導出され、他のイングレス・スイッチ・インタフェース・ハードウェアが変更を実行する。他のフレーム変更はピココード及びデータ・ストア・コプロセッサにより、イングレス・データ・ストア内に保持されるフレーム内容を変更することにより達成される。

【００３７】

他のデータは、フレームをスイッチ・ファブリック装置に送信する前に収集され、スイッチ・ヘッダ及びフレーム・ヘッダを作成するために使用される。制御データには、フレームの宛先ブレードなどのスイッチ情報や、イーグレス・インタフェース装置の情報が含まれ、宛先ポートのフレーム探索、マルチキャストまたはユニキャスト操作、及びイーグレス・フレーム変更を推進する。

【００３８】

完了時、エンキュー・コプロセッサが、フレームをスイッチ・ファブリックにエンキューするのに必要なフォーマットを作成し、それらを完了ユニットに送信する。完了ユニッ

トは、１０個のプロトコル・プロセッサからスイッチ・ファブリック・キューへのフレーム順序を保証する。スイッチ・ファブリック・キューからのフレームは、６４バイト・セルにセグメント化され、それらがＰｒｉｚｍａ－Ｅスイッチに伝送されるとき、フレーム・ヘッダ・バイト及びスイッチ・ヘッダ・バイトが挿入される。

【００３９】

スイッチ・ファブリックから受信されたフレームは、イーグレスＥＤＳ３４により、イーグレス・データ・ストア（イーグレスＤＳ）バッファに配置され、ＥＰＣにエンキューされる。フレームの一部はディスパッチャにより、アイドル・プロトコル・プロセッサに送信され、フレーム探索が実行される。フレーム・データは、分類子ハードウェア補助機構からのデータと一緒に、プロトコル・プロセッサにディスパッチされる。分類子ハードウェア補助機構は、イングレス・インタフェース装置により作成されるフレーム制御データを使用して、開始コード命令アドレス（ＣＩＡ）の決定を支援する。

【００４０】

イーグレス・ツリー探索は、イングレス探索においてサポートされるのと同一のアルゴリズムをサポートする。ルックアップはＴＳＥコプロセッサにより実行され、プロトコル・プロセッサを解放して実行を継続する。全ての制御メモリ操作は、制御メモリ・アービタにより管理され、アービタは１０個のプロセッサ・コンプレックスの間で、メモリ・アクセスを割当てる。

【００４１】

イーグレス・フレーム・データは、データ・ストア・コプロセッサを通じてアクセスされる。データ・ストア・コプロセッサは基本データ・バッファ（フレーム・データの最大８個の１６バイト・セグメント）と、スクラッチ・パッド・データ・バッファ（フレーム・データの最大８個の１６バイト・セグメント）と、データ・ストア操作のための幾つかの制御レジスタとを含む。探索の成功結果は転送情報を含み、ときとしてフレーム変更情報を含む。フレーム変更はＶＬＡＮヘッダ検出、存続時間増分（ＩＰＸ）または減分（ＩＰ）、ＩＰヘッダ・チェックサム再計算、イーサネット（Ｒ）フレームＣＲＣオーバレイまたは挿入、及びＭＡＣ　ＤＡ／ＳＡオーバレイまたは挿入を含む。ＩＰヘッダ・チェックサムはチェックサム・コプロセッサにより用意される。変更はインタフェース装置プロセッサ・コンプレックスにより実行されるのではなく、ハードウェア・フラグが作成され、ＰＭＭイーグレス・ハードウェアが変更を実行する。完了時に、エンキュー・コプロセッサが、フレームをＥＤＳイーグレス・キューにエンキューし、それらを完了ユニットに送信するために必要なフォーマットを作成する。完了ユニットは、１０個のプロトコル・プロセッサから、ＥＤＳイーグレス・キューへのフレーム順序を保証して、イーグレス・イーサネット（Ｒ）ＭＡＣ３６に供給する。

【００４２】

完了フレームは最終的に、ＰＭＭイーグレス・ハードウェアによりイーサネット（Ｒ）ＭＡＣに送信され、イーサネット（Ｒ）ポートから出力される。

【００４３】

ウェブと呼ばれる内部バスは、内部レジスタ、カウンタ及びメモリへのアクセスを可能にする。ウェブ（Ｗｅｂ）はまた、命令ステップや、デバッグ及び診断のための割込み制御を制御する外部インタフェースを含む。

【００４４】

ツリー構造探索エンジン・コプロセッサは、メモリ範囲チェック及び不当メモリ・アクセス通知を提供し、ツリー構造探索命令（メモリ読出し、書込み、またはread-add-writeなど）を、プロトコル・プロセッサ実行と並列に実行する。

【００４５】

共通命令メモリは、１つの１０２４×１２８ＲＡＭと、２セットのデュアル５１２×１２８ＲＡＭとから構成される。デュアルＲＡＭの各セットは、同一のピココードの２つのコピーを提供し、プロセッサによる同一アドレス範囲内の命令への独立したアクセスを可能にする。各１２８ビット・ワードは４つの３２ビット命令を含み、合計８１９２命令を

提供する。

【００４６】

ディスパッチャは、１０個のプロトコル・プロセッサへのフレームの受け渡しを制御し、割込み及びタイマを管理する。

【００４７】

完了ユニットは、プロセッサ・コンプレックスから、スイッチ・ファブリック及びターゲット・ポート・キューへのフレーム順序を保証する。豊富な命令セットには、条件付き実行、（入力ハッシュ・キーの）パッキング、条件付き分岐、符号付き及び符号無し演算、先行０のカウントなどが含まれる。

【００４８】

分類子ハードウェア補助機構エンジンは、各フレームのレイヤ２及びレイヤ３プロトコル・ヘッダを構文解析し、フレームがプロトコル・プロセッサにディスパッチされるとき、この情報をフレームに提供する。

【００４９】

制御メモリ・アービタは、内部及び外部メモリの両方へのプロセッサ・アクセスを制御する。

【００５０】

外部制御メモリ・オプションは、５個乃至７個のＤＤＲ　ＤＲＡＭサブシステムを含む。各サブシステムは、１対の２Ｍ×１６ビット×４バンクＤＤＲ　ＤＲＡＭ、または１対の４Ｍ×１６ビット×４バンクＤＤＲ　ＤＲＡＭをサポートする。ＤＤＲ　ＤＲＡＭインタフェースは、１３３ＭＨｚクロック・レート及び２６６ＭＨｚデータ・ストローブで動作し、構成可能なＣＡＳ待ち時間及びドライブ強度をサポートする。オプションの１３３ＭＨｚ　ＺＢＴ　ＳＲＡＭが、１２８Ｋ×３６、２×２５６Ｋ×１８、または２×５１２Ｋ×１８構成のいずれかで追加される。

【００５１】

イーグレス・フレームは１つの外部データ・バッファ（例えばＤＳ０）、または２つの外部データ・バッファ（ＤＳ０及びＤＳ１）のいずれかに記憶される。各バッファは、１対の２Ｍ×１６ビット×４バンクＤＤＲ　ＤＲＡＭ（最大２５６Ｋの６４バイト・フレームを記憶可能）、または１対の４Ｍ×１６ビット×４バンクＤＤＲ　ＤＲＡＭ（最大５１２Ｋの６４バイト・フレームを記憶可能）から構成される。２．２８Ｍｂｐｓの単一の外部データ・バッファ（例えばＤＳ０）を選択するか、第２のバッファ（例えばＤＳ１）を追加し、４．５７Ｍｂｐｓのレイヤ２及びレイヤ３スイッチングをサポートする。第２のバッファの追加は性能を改善するが、フレーム容量を増加しない。外部データ・バッファ・インタフェースは、１３３ＭＨｚクロック・レート、及び２６６ＭＨｚのデータ・ストローブで動作し、構成可能なＣＡＳ待ち時間及びドライブ強度をサポートする。

【００５２】

内部制御メモリは、２つの５１２×１２８ビットＲＡＭと、２つの１０２４×３６ビットＲＡＭと、１つの１０２４×６４ビットＲＡＭとを含む。

【００５３】

内部データ記憶は、イングレス方向（ＵＰ）に、最大２０４８個の６４バイト・フレームをバッファリングすることができる。

【００５４】

固定フレーム変更は、イングレス方向のＶＬＡＮタグ挿入と、ＶＬＡＮタグ削除と、持続時間増分／減分（ＩＰ、ＩＰＸ）と、イーグレス（ＤＯＷＮ）方向のイーサネット（Ｒ）ＣＲＣオーバレイ／挿入及びＭＡＣ　ＤＡ／ＳＡオーバレイ／挿入とを含む。

【００５５】

ポート・ミラーリングは、プロトコル・プロセッサ資源を用いることなく、１つの受信ポート及び１つの送信ポートを、システム指定の観測ポートにコピーすることを可能にする。ミラーリングされたインタフェース装置ポートは、フレーム及びスイッチ制御データを追加するように構成される。別々のデータ・パスが、イングレス・スイッチ・インタフ

ェースへの直接フレーム・エンキューイングを可能にする。

【００５６】

インタフェース装置は、４つのイーサネット（Ｒ）マクロを統合する。各マクロは、１ギガビットまたは１０／１００高速イーサネット（Ｒ）モードのいずれかで動作するように、個々に構成される。各イーサネット（Ｒ）マクロは、最大１０個の１０／１００Ｍｂｐｓ　ＭＡＣか、４つのマクロの各々に対して、１つの１０００Ｍｂｐｓ　ＭＡＣをサポートする。

【００５７】

図２は、ＭＡＣコアのブロック図を示す。各マクロは、３つのイーサネット（Ｒ）コア設計、すなわちマルチポート１０／１００Ｍｂｐｓ　ＭＡＣコア（ＦＥｎｅｔ）、１０００Ｍｂｐｓ　ＭＡＣコア（ＧＥｎｅｔ）、及び１００Ｍｂｐｓ物理コーディング・サブレイヤ・コア（ＰＣＳ）を含む。

【００５８】

マルチポート・イーサネット（Ｒ）１０／１００ＭＡＣフィーチャ：

１）物理レイヤとの１０個のシリアル・メディア独立インタフェースをサポートする。

２）１０Ｍｂｐｓまたは１００Ｍｂｐｓメディア速度の１０個のポートを、任意に混在させて処理できる。

３）単一のＭＡＣが時分割多重インタフェースにより、１０個の全てのポートをサービスする。

４）全てのポート上で全２重／半２重動作をメディア速度でサポートする。

５）ＩＥＥＥ８０２．３バイナリ指数バックオフをサポートする。

【００５９】

１０００Ｍｂｐｓイーサネット（Ｒ）ＭＡＣコア・フィーチャ：

１）物理ＰＣＳレイヤとの、または直接的に物理レイヤとのギガビット・メディア独立インタフェース（ＧＭＩＩ）をサポートする。

２）ＰＣＳコアにより、完全なＴＢＩ（８ｂ／１０ｂ）ソルーションをサポートする。

３）全２重Point-to-Point接続をメディア速度でサポートする。

４）ＩＢＭ　ＰＣＳコア有効バイト信号方式をサポートする。

【００６０】

１０００Ｍｂｐｓイーサネット（Ｒ）物理コーディング・サブレイヤ・コア・フィーチャ：

１）８ｂ／１０ｂエンコード及びデコードを実行する。

２）ＩＥＥＥ８０２．３ｚで定義されるＰＭＡ（１０ビット）サービス・インタフェースをサポートする。このインタフェースは、ＩＥＥＥ８０２．３ｚに準拠する任意のＰＭＡに接続する。

３）ＰＭＡから受信されるデータ（２フェーズ・クロック）を、ＭＡＣ（１フェーズ）クロックに同期させる。

４）次の２ページを含むオートネゴシエーションをサポートする。

５）規格で定義された２フェーズ・クロック・システムを、１フェーズ・クロックに変換する。

６）新たなデータを含むクロック・サイクルを示す信号を、ＭＡＣに提供する。

７）受信コード・グループ（１０ビット）内のＣＯＭＭＡをチェックし、ワード同期を確立する。

８）８ｂ／１０ｂ実行中ディスパリティを計算及びチェックする。

【００６１】

図３の（Ａ）乃至（Ｄ）は、インタフェース素子チップの異なる構成を示す。これらの構成はＤＡＳＬと、スイッチング・ファブリック装置への接続とによりサポートされる。各ＤＡＳＬは２つのチャネル、すなわち送信チャネルと受信チャネルとを含む。

【００６２】

図３の（Ａ）は、単一インタフェース装置のラップ（wrap）構成を示す。この構成では

、送信チャネルが受信チャネルにラップ、すなわち折り返される。

【００６３】

図３の（Ｂ）は、２つのインタフェース素子チップが接続される構成を示す。各インタフェース素子チップは、少なくとも２つのＤＡＳＬを提供される。この構成では、１つのチップ上の１ＤＡＳＬ上のチャネルが、他のチップ上のマッチングＤＡＳＬのチャネルに、動作可能なように接続される。各チップ上の他のＤＡＳＬは、ラップされる。

【００６４】

図３の（Ｃ）は、複数のインタフェース装置がスイッチ・ファブリックに接続される構成を示す。両頭矢印は両方向の伝送を示す。

【００６５】

図３の（Ｄ）は、メイン・スイッチ及びバックアップ・スイッチが複数のインタフェース装置に接続される構成を示す。メイン・スイッチが故障すると、バックアップ・スイッチが使用可能になる。

【００６６】

制御点（ＣＰ）はシステム・プロセッサを含み、これは各構成に接続される。とりわけ、ＣＰのシステム・プロセッサは、チップに対して、初期化及び構成サービスを提供する。ＣＰは３つの位置、すなわちインタフェース素子チップ内か、チップが実装されるブレード上か、或いはブレードの外部のいずれかに配置される。ブレードの外部の場合、ＣＰはリモートとなり、どこか別の場所に内蔵され、インタフェース装置及びＣＰが接続されるネットワークを介して通信する。ＣＰの要素が図２０に示され、メモリ素子（キャッシュ、フラッシュ及びＳＤＲＡＭ）、メモリ制御装置、ＰＣＩバス、及びバックプレーン及びＬ１ネットワーク・メディア用のコネクタを含む。

【００６７】

図２１は、単一チップ・ネットワーク・プロセッサと、ＥＤＳ－ＵＰ、トラフィック管理（ＭＧＴ）スケジューラ、及びＥＤＳ－ＤＯＷＮ（ＤＮ）により提供される機能とを示す。Ｕ字形アイコンはキューを表し、キュー内の内容を追跡する制御ブロック（ＣＢ）は、矩形アイコンにより表される。

【００６８】

各要素及びそれらの機能及び相互作用は、次の通りである。

【００６９】

ＰＭＭ：これはＭＡＣ（ＦＥｎｅｔ、ＰＯＳ、ＧＥｎｅｔ）を含み、外部ＰＨＹ装置に接続されるネットワーク・プロセッサの一部である。

【００７０】

ＵＰ－ＰＭＭ：この論理はＰＨＹからバイトを受け取り、それをＦＩＳＨ（１６バイト）にフォーマットし、ＵＰ－ＥＤＳに受け渡す。ＰＭＭ内には４つのＤＭＵが存在し、各々は１ＧＥｎｅｔまたは１０ＦＥｎｅｔ装置と連携することができる。

【００７１】

ＵＰ－ＥＤＳ：この論理はＵＰ－ＰＭＭからＦＩＳＨを受け取り、それらをＵＰ－ＤＡＴＡストア（内部ＲＡＭ）に記憶する。これは１度に４０フレームを処理することができ、適切な数のバイトが受信されると、フレームをＥＰＣにエンキューする。ＥＰＣがフレーム処理を完了すると、ＵＰ－ＥＤＳがフレームを適切なターゲット・ポート・キューにエンキューし、ＵＰ－ＳＤＭにフレームの送信を開始する。ＵＰ－ＥＤＳは全てのバッファ及びフレーム管理の責任を負い、ＵＰ－ＳＤＭへの転送が完了するとき、バッファ／フレームを空きプールに戻す。

【００７２】

ＥＰＣ：この論理はピコプロセッサを含み、組み込みＰｏｗｅｒＰＣを含み得る。この論理はフレーム・ヘッダを突き止め、フレームをどのように処理すべきかを決定する（転送、変更、フィルタリングなど）。ＥＰＣは幾つかのルックアップ・テーブルをアクセスでき、ピコプロセッサがネットワーク・プロセッサの高帯域幅要求に応じるためのハードウェア補助機構を有する。

【００７３】
ＵＰ－ＳＤＭ：この論理はフレームを受け取り、それらをスイッチ・ファブリックへの伝送のために、Ｐｒｉｚｍａセルにフォーマットする。この論理はＶＬＡＮヘッダをフレームに挿入することができる。
【００７４】
ＵＰ－ＳＩＦ：この論理はＵＰ－ＤＡＳＬマクロを含み、外部スイッチ入出力（Ｉ／Ｏ）に接続する。
【００７５】
ＤＮ－ＳＩＦ：この論理はＤＮ－ＤＡＳＬマクロを含み、外部スイッチ入出力（Ｉ／Ｏ）からＰｒｉｚｍａセルを受信する。
【００７６】
ＤＮ－ＳＤＭ：この論理はＰｒｉｚｍａセルを受信し、フレーム再組み立てのために、それらを事前処理する。
【００７７】
ＤＮ－ＥＤＳ：この論理は各セルを受け取り、それらを再度フレームに組み立てる。セルは外部データ・ストアに記憶され、バッファが一緒にリンクされて、フレームを形成する。全フレームが受信されると、フレームはＥＰＣにエンキューされる。ＥＰＣがフレーム処理を終えると、フレームはスケジューラ・キュー（但し存在する場合）またはターゲット・ポート・キューにエンキューされる。ＤＮ－ＥＤＳは次に、フレーム、変更情報、及び制御情報などをＤＮ－ＰＭＭに送信することにより、フレームを適切なポートに送信する。
【００７８】
ＤＮ－ＰＭＭ：ＤＮ－ＥＤＳから情報を受け取り、フレームをイーサネット（Ｒ）、ＰＯＳなどにフォーマットし、フレームを外部ＰＨＹに送信する。
【００７９】
ＳＰＭ：この論理は、ネットワーク・プロセッサが外部装置（ＰＨＹ、ＬＥＤ、ＦＬＡＳＨなど）とインタフェースするために使用されるが、３つの入出力（Ｉ／Ｏ）を要求するだけである。ネットワーク・プロセッサはシリアル・インタフェースを用いて、ＳＰＭと通信し、次にＳＰＭがこれらの外部装置を管理するために必要な機能を実行する。
【００８０】
アップサイド・フロー：
１）フレームがＰＨＹに到来する。
２）バイトがＵＰ－ＰＭＭにより受信される。
３）ＵＰ－ＰＭＭがＦＩＳＨをＵＰ－ＥＤＳに送信する（ＦＩＳＨはフレームの一部を意味する）。
４）ＵＰ－ＥＤＳがＦＩＳＨをＵＰ－ＤＳに記憶する。
５）ＵＰ－ＥＤＳがヘッダをＥＰＣ送信する。
６）ＥＰＣがヘッダを処理し、エンキュー情報をＵＰ－ＥＤＳに返送する。
７）ＵＰ－ＥＤＳがフレームの残りをＵＰ－ＰＭＭから受信し続ける。
８）スイッチへの適切なデータの送信準備が整うと、ＵＰ－ＥＤＳが情報をＵＰ－ＳＤＭに送信する。
９）ＵＰ－ＳＤＭがフレーム・データを読出し、それをＰｒｉｚｍａセルにフォーマットする。
１０）ＵＰ－ＳＤＭがセルをＵＰ－ＳＩＦに送信する。
１１）ＵＰ－ＳＩＦがＤＡＳＬシリアル・リンクを介して、セルをＰｒｉｚｍａに転送する。
１２）全てのデータが受け取られると、ＵＰ－ＥＤＳがバッファ／フレームを解放する。
【００８１】
ダウンサイド・フロー：

１）ＤＮ－ＳＩＦがＰｒｉｚｍａセルを受信する。
２）ＤＮ－ＳＤＭがセルを記憶し、それらを再組み立て情報として事前処理する。
３）ＤＮ－ＥＤＳがセル・データ及び再組み立て情報を受信し、セルをダウンサイド側の新たなフレームにリンクする。
４）ＤＮ－ＥＤＳがセルをＤＮ－ＤＳに記憶する。
５）全てのデータが受信されると、ＤＮ－ＥＤＳがフレームをＥＰＣにエンキューする。
６）ＥＰＣがヘッダを処理し、エンキュー情報をＤＮ－ＥＤＳに返送する。
７）ＤＮ－ＥＤＳがフレームをスケジューラ・キュー（但し存在する場合）またはターゲット・ポート・キューにエンキューする。
８）ＤＮ－ＥＤＳがキューをサービスし、フレーム情報をＰＣＢに送信する。
９）ＤＮ－ＥＤＳがＰＣＢを用いてフレームを"解体"（unravel）し、適切なデータを読出し、そのデータをＤＮ－ＰＭＭに送信する。
１０）ＤＮ－ＰＭＭがデータをフォーマットし（必要に応じて変更を加える）、フレームを外部ＰＨＹに送信する。
１１）ＤＮ－ＰＭＭがＤＮ－ＥＤＳに、バッファがもはや必要とされないことを通知し、ＤＮ－ＥＤＳがこれらの資源を解放する。

【００８２】

フレーム制御フロー：
１）ヘッダがＵＰ－ＤＳまたはＤＮ－ＤＳからＥＰＣに送信される。
２）ＥＰＣがルックアップ・テーブルでヘッダ情報を調査し、フレーム・エンキュー情報を受信する。
３）ＥＰＣがエンキュー情報をＥＤＳに返送し、フレームが適切なキューにエンキューされる。
４）セル・ヘッダ及びフレーム・ヘッダがフレーム・データと一緒に送信され、再組み立て及びフレーム転送を支援する。

【００８３】

ＣＰ制御フロー：
１）制御点がガイド・フレームをフォーマットし、それをネットワーク・プロセッサに送信する。
２）ネットワーク・プロセッサがガイド・フレームを、ＧＣＨピコプロセッサにエンキューする。
３）ＧＣＨがガイド・フレームを処理し、Rainierの要求領域を読み書きする。
４）ＧＣＨがテーブル更新要求をＧＴＨに渡す。
５）ＧＴＨが適切なテーブルを、ガイド・フレームからの情報により更新する。
６）肯定応答ガイド・フレームがＣＰに返送される。

【００８４】

ネットワーク・プロセッサ制御フロー：
１）ピコプロセッサがガイド・フレームを作成し、情報を別のRainierまたは制御点に送信する。
２）ガイド・フレームが適切な位置に送信され処理される。

【００８５】

単一インタフェース装置は、最大４０の高速イーサネット（Ｒ）ポート（図３の（Ａ）参照）のための、メディア速度切替えを提供する。ＩＢＭのデータ整合同期リンク（ＤＡＳＬ：Data Aligned Synchronous Link）技術を用いて、２つのインタフェース装置が相互接続される場合、８０個の高速イーサネット（Ｒ）ポートがサポートされる（図３の（Ｂ）参照）。各ＤＡＳＬ差動対は、４４０Ｍｂｐｓのデータを伝搬する。従って、８対の２つのセットが、３．５Ｇｂｐｓ全２重接続（各方向において４４０Ｍｂｐｓの８倍）を提供する。図３の（Ｃ）及び（Ｄ）に示されるように、複数のインタフェース装置を、ＩＢＭのＰｒｉｚｍａ－Ｅスイッチなどのスイッチに相互接続することにより、より大規模

なシステムが構成される。インタフェース装置は、２つの３．５Ｇｂｐｓ全２重ＤＡＳＬ接続を提供し、これらの一方が基本接続で他が補助接続であり、後者は、ローカル・フレーム・トラフィックのためのラップ・バックパス（wrap-backpath）を提供するか（２つのインタフェース装置が直接接続される場合（図３の（Ｂ））、或いは、冗長スイッチ・ファブリックへの接続を提供する（図３の（Ｄ）、バックアップ・スイッチ）。以上から、単一ネットワーク・プロセッサ・チップは、１つのチップがローエンド・システム（比較的低いポート密度（例えば４０）を有する）から、ハイエンド・システム（比較的高いポート密度（例えば８０）を有する）まで提供するために使用されるという点でスケーラブルである。

【００８６】

システム内の１つのインタフェース装置が、最大１０個の１０／１００Ｍｂｐｓ高速イーサネット（Ｒ）ポートか、単一の１０００Ｍｂｐｓイーサネット（Ｒ）ポートを介して、システム・プロセッサに接続される。システム・プロセッサへのイーサネット（Ｒ）構成は、インタフェース装置に接続されるＥＥＰＲＯＭ内に記憶され、初期化の間にロードされる。システム・プロセッサは、イーサネット（Ｒ）フレームとしてカプセル化される特殊ガイド・フレームを作成することにより、システム内の全てのインタフェース装置と通信する（図３参照）。カプセル化されたガイド・フレームは、ＤＡＳＬリンクを介して他の装置に転送され、システム内の全てのインタフェース装置が、単一点から制御されることになる。

【００８７】

ガイド・フレームは、制御点（ＣＰ）と組み込みプロセッサ・コンプレックスとの間で、及びインタフェース装置内で制御情報を伝達するために使用される。ここでの議論を明らかにするガイド・セルの従来の開示が、１９９８年３月３日発行の米国特許第５７２４３４８号"Efficient Hardware/Software Interface for a Data Switch"で述べられている。

【００８８】

ＣＰから発信されるガイド・フレーム・トラフィックでは、ＣＰがそのローカル・メモリ内のデータ・バッファ内に、ガイド・フレームを作成する。ＣＰのデバイス・ドライバがガイド・フレームを、ネットワーク・プロセッサのメディア・インタフェースの１つに送信する。メディア・アクセス制御（ＭＡＣ）ハードウェアがガイド・フレームを回復し、それを内部データ・ストア（Ｕ－ＤＳ）メモリに記憶する。ガイド・フレームは適切なブレードに経路指定され、処理され、必要に応じて再度ＣＰに経路指定される。外部ＣＰとインタフェース装置との間でやり取りされるガイド・フレームは、外部ネットワークのプロトコルに適応するようにカプセル化される。その結果、外部ネットワークがイーサネット（Ｒ）を含む場合、ガイド・フレームはイーサネット（Ｒ）フレームとしてカプセル化される。

【００８９】

イーサネット（Ｒ）カプセル化は、ＣＰとインタフェース装置との間のガイド・トラフィックのトランスポート手段を提供する。インタフェース装置のイーサネット（Ｒ）ＭＡＣ（Ｅｎｅｔ　ＭＡＣ）は、フレームの受信時に、宛先アドレス（ＤＡ）またはソース・アドレス（ＳＡ）を分析しない。この分析はＥＰＣピココードにより実行される。ガイド・トラフィックは、インタフェース装置が構成されておらず、ＤＡ及びＳＡがＥＰＣピココードにより分析されないものと仮定する。従って、これらのフレームは元来、自己ルーティングである。しかしながら、イーサネット（Ｒ）ＭＡＣは、イーサネット（Ｒ）タイプ・フィールドを分析し、ガイド・トラフィックとデータ・トラフィックとを区別する。ガイド・フレームのこのイーサネット（Ｒ）タイプ値の値は、E_Type_Cレジスタにロードされる値と合致しなければならない。このレジスタは、インタフェース装置のブート・ピココードにより、フラッシュ・メモリからロードされる。

【００９０】

ＣＰはそのローカル・メモリ内のデータ・バッファ内に、ガイド・フレームを構築する

。ＣＰプロセッサ内の３２ビット・レジスタの内容が、図４に示されるように、ビッグ・エンディアン形式でローカル・メモリ内に記憶される。ガイド・フレームが構築されると、ＣＰのデバイス・ドライバがイーサネット（Ｒ）フレームを送信する。このフレームは、特定のガイド・セル・ハンドラ（ＧＣＨ）のＤＡ、ＣＰの大域ＭＡＣアドレスまたは特定インタフェースのＭＡＣアドレスに対応するＳＡ、ガイド・フレームを示す特殊イーサネット（Ｒ）タイプ・フィールド、及びガイド・フレーム・データを含む。ポートに到来する全てのイーサネット（Ｒ）フレームは、イーサネット（Ｒ）ＭＡＣにより受信され、分析される。E_Type_Cレジスタの内容に合致するイーサネット（Ｒ）タイプ値を有するフレームの場合、イーサネット（Ｒ）ＭＡＣがＤＡ、ＳＡ及びイーサネット（Ｒ）タイプ・フィールドを剥ぎ取り、ガイド・フレーム・データをＵ＿ＤＳメモリに記憶する。バイトが１つずつ、イーサネット（Ｒ）ＭＡＣにより、ＦＩＳＨと呼ばれる１６バイトのブロック内に収集される。これらのバイトはビッグ・エンディアン形式で記憶され、ガイド・フレームの第１バイトが、ＦＩＳＨの最上位バイト位置（バイト０）に記憶される。続くバイトは、ＦＩＳＨ内の続くバイト位置（バイト１，バイト２，．．，バイト１５）に記憶される。これらの１６バイトは次に、Ｕ＿ＤＳ内のバッファに、ＦＩＳＨ０位置から記憶される。続くＦＩＳＨは、バッファ内の連続するＦＩＳＨ位置（ＦＩＳＨ１、ＦＩＳＨ２、ＦＩＳＨ３など）に記憶される。ガイド・フレームの残りを記憶するために、必要に応じて、空きプールから追加のバッファが獲得される。

【００９１】

インタフェース装置１０内のガイド・トラフィックのフローが、図５に示される。インタフェース装置のイーサネット（Ｒ）ＭＡＣ機能が、フレーム・ヘッダ情報を調査し、フレームがガイド・フレームであることを判断する。イーサネット（Ｒ）ＭＡＣはガイド・フレームからフレーム・ヘッダを除去し、その残りの内容をインタフェース装置の内部Ｕ＿ＤＳメモリにバッファリングする。イーサネット（Ｒ）ＭＡＣは、フレームがＧＣＨにより処理されるために、汎用制御（ＧＣ）キューにエンキューされるべきことを示す。ガイド・フレームの終わりに達すると、エンキュー、デキュー及びスケジュール（ＥＤＳ）論理が、フレームをＧＣキューにエンキューする。

【００９２】

ＣＰに局所的に接続されるブレード上のＧＣＨピココードは、フレーム制御情報（図７参照）を調査し、ガイド・フレームがシステム内の他のブレードに向けられるか、及びガイド・フレームがインタフェース装置のダウンサイド側で実行されるべきかを判断する。フレームが、局所的に接続されるブレード以外のブレード、またはそれに加えたブレードに向けられる場合、ＧＣＨピココードがフレーム制御ブロック（ＦＣＢ）内のＴＢ値を、ガイド・フレームのフレーム制御情報からのＴＢ値により更新し、ＥＤＳにフレームを、マルチキャスト・ターゲット・ブレード・フレーム開始（ＴＢ＿ＳＯＦ）キューにエンキューするように命令する。性能上の理由から、示された宛先ブレードの数とは無関係に、全てのガイド・トラフィックがマルチキャストＴＢ＿ＳＯＦキューにエンキューされる。

【００９３】

フレームが局所的に接続されたブレードだけに向けられる場合、ＧＣＨピココードがフレーム制御情報のＵＰ／ＤＯＷＮバー・フィールド（以降、'バー'は'０'ｂで有効なことを意味する）（図７参照）を調査し、ガイド・フレームがインタフェース装置のアップサイドまたはダウンサイドのどちらで実行されるべきかを判断する。ガイド・フレームがインタフェース装置のダウンサイドで実行される場合、ＧＣＨピココードがＦＣＢ内のＴＢ値を、ガイド・フレームのフレーム制御情報からのＴＢ値により更新し、ＥＤＳにフレームをマルチキャストＴＢ＿ＳＯＦキューにエンキューするように命令する。フレーム制御情報が、ガイド・フレームがアップサイドで実行されるべきことを示す場合、ＧＣＨピココードがガイド・フレームを分析し、そこに含まれるガイド・コマンドにより示される操作を実行する。

【００９４】

ガイド・コマンドを処理する前に、ピココードはフレーム制御情報のＡＣＫ／ＮＯＡＣ

Ｋバー・フィールドの値をチェックする、この値が'０'ｂの場合、ガイド・フレームが処理に続き廃棄される。但し、ガイド読出しコマンドは、この範疇ではない。

【００９５】

ＡＣＫ／ＮＯＡＣＫバー・フィールドが'１'ｂで、ＥＡＲＬＹ／ＬＡＴＥバー・フィールドが'１'ｂの場合、ガイド・フレーム内の任意のガイド・コマンドを処理する前に、ピココードが早期ＡＣＫガイド・フレームを作成する。このとき、フレーム制御のＴＢフィールドの値は、早期ＡＣＫガイド・フレームの内容に等しく、フレーム制御のＴＢフィールドの値は、Ｍｙ＿ＴＢレジスタの内容に等しい。ピココードは、フレームのＦＣＢ内のＴＢ値を、ＬＡＮ制御点アドレス（LAN_CP_Addr）レジスタのＴＢフィールドに含まれる値により更新し、ＥＤＳにフレームをマルチキャストＴＢ＿ＳＯＦキューにエンキューするように命令することにより、早期ＡＣＫガイド・フレームをＣＰに返送する。ピココードは次に、ガイド・フレームのガイド・コマンドを処理し、ガイド・フレームを廃棄する。但し、ガイド読出しコマンドはこの範疇ではない。

【００９６】

他方、ＡＣＫ／ＮＯＡＣＫバー・フィールドの値が'１'ｂで、ＥＡＲＬＹ／ＬＡＴＥバー・フィールドの値が'０'ｂの場合、ピココードはフレーム制御情報のＲＥＳＰ／ＲＥＱバー・フィールドを'１'ｂに変更して、ガイド・フレーム応答を示し、ＴＢフィールドをＭｙ＿ＴＢレジスタの内容により更新し、ガイド・フレーム内の各ガイド・コマンドを処理する。ガイド・コマンドの処理の間、ピココードは次のガイド・コマンドの完了コード・フィールドを、現ガイド・コマンドの完了ステータス・コード値により更新する。ピココードは、ＦＣＢ内のＴＢ値をソース・ブレードに対応する値（ＣＰのLAN_CP_Addrレジスタ値）により更新し、ＥＤＳにフレームをマルチキャストＴＢ＿ＳＯＦキューにエンキューするように命令することにより応答をソースに返送する。

【００９７】

ＴＢ＿ＳＯＦキュー内に存在するフレームは、ＥＤＳにより、転送をスケジュールされる。スイッチ・データ・ムーバ（ＳＤＭ）が、ＦＣＢに含まれる情報から、スイッチング・ファブリック・セル・ヘッダ及びインタフェース装置フレーム・ヘッダを作成する。これらのセルはスイッチング・ファブリック装置を通過し、ターゲット・ブレードに到達し、そこでセルがＤ－ＤＳメモリ内で、フレームに再組み立てされる。ダウンサイドのＳＤＭは、フレームがガイド・フレームであることを認識し、ＥＤＳにそれをＧＣキューにエンキューするように伝える。

【００９８】

ＧＣキューまたはＧＴキューからの圧力により、ピココードはガイド・フレームをアクセスし分析する。ダウンサイドに到来する全てのガイド・フレームは、最初にＧＣキューにエンキューされる。これらのフレームにおけるフレーム制御情報のＧＴＨ／ＧＣＨバー値が、ＧＣＨピココードにより調査される。ＧＴＨ／ＧＣＨバー値が'０'ｂの場合、ガイド・フレームがＧＴキューにエンキューされる。それ以外では、ＧＣＨピココードが、フレーム制御情報のＲＥＳＰ／ＲＥＱバー・フィールドを調査し、ガイド・フレームが既に実行されたか否かを判断する。ＲＥＳＰ／ＲＥＱバー・フィールドが値'１'ｂを有する場合、ガイド・フレームは既に実行されており、ＣＰに経路指定される。ＣＰ接続に対応するターゲット・ポート値は、ＥＰＣピココードにより保持される。これらのターゲット・ポート・キューからのフレームが、インタフェース装置からＣＰに返送される。

【００９９】

ＲＥＳＰ／ＲＥＱバー・フィールドが値'０'ｂを有する場合、ブレードはＣＰに対してローカルまたはリモートである。これはLAN_CP_AddrレジスタのＴＢフィールドの値を、マイ・ターゲット・ブレード（Ｍｙ＿ＴＢ）レジスタの内容と比較することにより解明される。これらが一致する場合、ブレードはＣＰにとってローカルであり、それ以外では、ブレードはＣＰにとってリモートである。いずれの場合にも、ピココードはフレーム制御情報のＵＰ／ＤＯＷＮバー値を調査する。ＵＰ／ＤＯＷＮバー値が'１'ｂの場合、フレームはラップＴＰキューにエンキューされ、Ｕ＿ＤＳへ転送されて、アップサイドのＧＣＨ

により処理される。それ以外では、ピココード（ＧＣＨまたはＧＴＨ）が、ガイド・フレームに含まれるガイド・コマンドにより指示される操作を実行する。ガイド・コマンドの処理に先立ち、ピココードがフレーム制御情報のＡＣＫ／ＮＯＡＣＫバー・フィールドの値をチェックする。この値が'０'ｂの場合、ガイド・フレームが処理に続き廃棄される。但し、ガイド読出しコマンドはこの範疇ではない。

【０１００】

ＡＣＫ／ＮＯＡＣＫバー・フィールドの値が'１'ｂで、ＥＡＲＬＹ／ＬＡＴＥバー・フィールドの値が'１'ｂの場合、ガイド・フレーム内の任意のガイド・コマンドを処理する前に、ピココードが早期ＡＣＫガイド・フレームを作成し、このとき、フレーム制御情報のＴＢフィールドの値は、Ｍｙ＿ＴＢレジスタの内容に等しい。ブレードがＣＰにとってリモートの場合、ピココードは早期ＡＣＫガイド・フレームをラップ・ポートに経路指定する。それ以外では、ブレードはＣＰにとってローカルであり、フレームはＣＰに対応するポート・キューに経路指定される。ピココードはガイド・コマンドを処理する一方、ラップ・ポートが早期ＡＣＫガイド・フレームをＤ＿ＤＳからＵ＿ＤＳに転送し、フレームをアップサイドのＧＣキューにエンキューするか、フレームがポート・キューからＣＰに返送される。Ｕ＿ＤＳにラップバック（wrap back）されるフレームに対して、ＧＣＨピココードは再度このフレームを調査するが、ＲＥＳＰ／ＲＥＱバー・フィールドは値'１'ｂを有する。ＧＣＨピココードは、ＦＣＢ内のＴＢフィールドを、LAN_CP_Addrレジスタの ＴＢフィールドに含まれる値により更新し、ＥＤＳにフレームをマルチキャストＴＢ＿ＳＯＦキューにエンキューするように命令することにより、フレームをＣＰに返送する。ＴＢ＿ＳＯＦキュー内に存在するフレームは、ＥＤＳにより、転送をスケジュールされる。スイッチ・データ・ムーバ（ＳＤＭ）が、ＦＣＢに含まれる情報から、Ｐｒｉｚｍａセル・ヘッダ及びインタフェース装置フレーム・ヘッダを作成する。このフレームからのセルはＰｒｉｚｍａを通過し、ＣＰのローカル・ブレード上で、フレームに再組み立てされる。ダウンサイドのＳＤＭは、フレームがガイド・フレームであることを認識し、ＥＤＳにそれをＧＣキューにエンキューするように伝える。ＧＣＨピココードがフレームを分析するとき、ＲＥＳＰ／ＲＥＱバー・フィールドは値'１'ｂを有する。このことは、このブレードがＣＰに局所的に接続されることを意味し、従ってガイド・フレームは、ＣＰに対応するポート・キューに経路指定される。このキュー内のフレームは、インタフェース装置からＣＰに返送される。

【０１０１】

他方、ＡＣＫ／ＮＯＡＣＫバー・フィールドの値が'１'ｂで、ＥＡＲＬＹ／ＬＡＴＥバー・フィールドの値が'０'ｂの場合、ピココードはＲＥＳＰ／ＲＥＱバー・フィールドを'１'ｂに変更し、ガイド・フレーム応答であることを示し、ＴＢフィールドをＭｙ＿ＴＢレジスタの内容により置換し、次にガイド・フレーム内の各ガイド・コマンドを処理する。ガイド・コマンドの処理の間、ピココードは次のガイド・コマンドの完了コード・フィールドを、現ガイド・コマンドの完了ステータス・コード値により更新する。ブレードがＣＰにとってリモートの場合、ピココードはガイド・フレームをラップ・ポートに経路指定する。それ以外では、ブレードはＣＰにとってローカルであり、フレームはＣＰに対応するポート・キューに経路指定される。ラップ・ポートがガイド・フレームをＤ＿ＤＳからＵ＿ＤＳに転送し、フレームをアップサイドのＧＣキューにエンキューするか、フレームがポート・キューからＣＰに返送される。Ｕ＿ＤＳにラップバックされるフレームに対して、ＧＣＨピココードは再度このフレームを調査するが、ＲＥＳＰ／ＲＥＱバー・フィールドは値'１'ｂを有する。ＧＣＨピココードは、ＦＣＢ内のＴＢフィールドを、LAN_CP_Addrレジスタの ＴＢフィールドに含まれる値により更新し、ＥＤＳにフレームをマルチキャストＴＢ＿ＳＯＦキューにエンキューするように命令することにより、フレームをＣＰに返送する。ＴＢ＿ＳＯＦキュー内に存在するフレームは、ＥＤＳにより、転送をスケジュールされる。ＳＤＭが、ＦＣＢに含まれる情報から、Ｐｒｉｚｍａセル・ヘッダ及びインタフェース装置フレーム・ヘッダを作成する。このフレームからのセルはＰｒｉｚｍａを通過し、ＣＰのローカル・ブレード上のダウンサイドでフレームに再組み立てされる

。ダウンサイドのＳＤＭは、フレームがガイド・フレームであることを認識し、ＥＤＳにそれをＧＣキューにエンキューするように伝える。ＧＣＨピココードがＤ＿ＤＳからのフレームを分析するとき、ＲＥＳＰ／ＲＥＱバー・フィールドは値'１'ｂを有する。このことは、このブレードがＣＰに局所的に接続されることを意味し、ガイド・フレームが、ＣＰに対応するポート・キューに経路指定される。このキュー内のフレームは、インタフェース装置からＣＰに返送される。

【０１０２】

何らかの理由により、ＧＣＨピココードが、フレーム制御情報のＴＢフィールドが'００００'ｈに等しいガイド・フレームに遭遇する場合、ＧＣＨピココードはこのフレームを、このブレードだけに向けられるものと解釈し、それにもとづき処理を行う。このアクションは、Ｍｙ＿ＴＢレジスタの値が全てのブレードに対して'００００'ｈである初期化の間に要求される。ＣＰは、フレーム制御情報が'００００'ｈのＴＢ値を有するガイド・フレーム内の、書込みガイド・コマンドを送信することにより、局所的に接続されるブレードのＭｙ＿ＴＢレジスタを初期化する。

【０１０３】

ＥＰＣ内の任意のピコプロセッサがガイド・フレームを生成できる。このフレームは非送信請求ガイド・フレームか、他の形式のガイド・フレームである。このタイプの内部的に生成されるフレームは、肯定応答を許可しないように生成される（すなわちＡＣＫ／ＮＯＡＣＫバー＝'０'ｂ）。これらのフレームは、同一のＥＰＣ内の２つのピコプロセッサの一方（ＧＣＨまたはＧＴＨ）に送信されるか、他のブレードのＧＣＨまたはＧＴＨに送信される。

【０１０４】

非送信請求ガイド・フレームはＣＰにも送信され得る。同一のＥＰＣに向けられるガイド・フレームは、Ｄ＿ＤＳ内のデータ・バッファを用いて構成される。これらのフレームは次に処理のために、ＧＣまたはＧＴキューにエンキューされる。これらのフレームは次に通常通り処理され、廃棄される。局所的に接続されるＣＰに向けられる非送信請求ガイド・フレームは、Ｄ＿ＤＳ内のデータ・バッファを用いて構成される。これらのフレームは、それらがＥＰＣにより実行されたことを示すように構成される（すなわち、ＲＥＳＰ／ＲＥＱバー＝'１'ｂ、ＴＢ＝Ｍｙ＿ＴＢ）。これらのフレームは、ＣＰに対応するポート・キューにエンキューされる。このキューからフレームは、ＣＰに返送される。

【０１０５】

別のブレードに向けられるガイド・フレームは、Ｄ＿ＤＳまたはＵ＿ＤＳ内のデータ・バッファを用いて構成される。ＣＰに向けられる非送信請求ガイド・フレームは、それらがＥＰＣにより実行されたことを示すように構成される（すなわち、ＲＥＳＰ／ＲＥＱバー＝'１'ｂ、ＴＢ＝Ｍｙ＿ＴＢ）。Ｄ＿ＤＳ内のデータ・バッファを用いて構成されたフレームは、ラップ・ポートにエンキューされる。これらのフレームはＵ＿ＤＳに転送され、アップサイドのＧＣキューにエンキューされる。'１'ｂのＲＥＳＰ／ＲＥＱバー値を有する非送信請求ガイド・フレームは、LAN_CP_Addrレジスタ内のＴＢ値を用いてＣＰに経路指定される。それ以外では、ＧＣＨピココードが、ガイド・フレームのフレーム制御情報のＴＢ値を用いて、これらのフレームを経路指定する。受信ブレードにおいて、フレームがダウンサイドのＧＣキューにエンキューされる。このブレードのＧＣＨは、フレームを実行し廃棄するか（ＲＥＳＰ／ＲＥＱバー＝'０'ｂ、ＧＴＨ／ＧＣＨバー＝'１'）、フレームをＧＴキューにエンキューするか（ＲＥＳＰ／ＲＥＱバー＝'０'ｂ、ＧＴＨ／ＧＣＨバー＝'０'）、フレームをＣＰに対応するポート・キューにエンキューする（ＲＥＳＰ／ＲＥＱバー＝'１'ｂ）。Ｕ＿ＤＳ内のデータ・バッファを用いて構成されたフレームは、アップサイドのＧＣキューに直接エンキューされる。この点から以降、これらのフレームは同一の経路に従い転送され、Ｄ＿ＤＳデータ・バッファを用いて構成されたフレームと同様に処理される。図６は、ガイド・フレームの汎用フォーマットを示す。

【０１０６】

図示のフォーマットは、左側に最上位バイトを有し、右側に最下位バイトを有する論理

表現である。４バイト・ワードが先頭のワード０から開始し、ページの終わりに向けて増加する。

【０１０７】

インタフェース装置がＣＰにより構成される前に、ガイド・フレームは経路指定され、処理されなければならないので、これらのフレームは自己ルーティングでなければならない。探索及び分類により通常獲得される結果が、ガイド・フレームのこのフレーム制御情報フィールド内に含まれ、チップが探索操作を実行することなく、ＦＣＢをこの情報により更新することを可能にする。ガイド・フレームに含まれるターゲット・ブレード情報は、ガイド・フレーム・ハンドラにより、ＦＣＢのリーフ・ページ・フィールドを用意するために使用される。ＣＰがターゲット・ブレード情報を提供する一方、ＧＣＨピココードがＦＣＢ内の他のフィールドを記入する。このＦＣＢ情報はＳＤＭにより、セル・ヘッダ及びフレーム・ヘッダを用意するために使用される。ガイド・フレームのフレーム制御情報フィールドのフォーマットが図７に示される。

【０１０８】

次に、図７の各ビット位置の略語について説明する。

【０１０９】

ＲＥＳＰ／ＲＥＱバー：応答及び非要求バー標識値。このフィールドは、要求（未処理）ガイド・フレームと応答ガイド・フレームとを区別するために使用される。

０：要求

１：応答

【０１１０】

ＡＣＫ／ＮＯＡＣＫバー：肯定応答または無肯定応答制御値。このフィールドは、ＧＣＨピココードがガイド・フレームを肯定応答するか否かを制御するために使用される（肯定応答する場合ＡＣＫ、そうでない場合ＮＯＡＣＫ）。ガイド・フレームが肯定応答されない場合、読出しを実行するいずれの形式のガイド・コマンドも含まれない。

０：無肯定応答

１：肯定応答

【０１１１】

ＥＡＲＬＹ／ＬＡＴＥバー：早期及び遅延肯定応答制御値。このフィールドは、要求される肯定応答（ＡＣＫ／ＮＯＡＣＫバー＝'１'ｂ）が、ガイド・フレームが処理される前に発生するか（ＥＡＲＬＹ）、または後に発生するか（ＬＡＴＥ）を制御するために使用される。ＡＣＫ／ＮＯＡＣＫバー＝'０'の場合、このフィールドは無視される。

０：ガイド・フレーム処理後の肯定応答

１：ガイド・フレーム処理前の肯定応答

【０１１２】

ＮＥＧ／ＡＬＬバー：否定応答または全肯定応答制御値。このフィールドは、ガイド・コマンドが成功裡に完了しない場合を除き、ＡＣＫ／ＮＯＡＣＫバー・フィールドが値'０'ｂを有するとき無視される。

０：ＡＣＫ／ＮＯＡＣＫバー・フィールドが'１'ｂの場合、全てのガイド・フレームを肯定応答する。早期または遅延肯定応答は、ＥＡＲＬＹ／ＬＡＴＥバーの値により決定される。

１：成功裡に完了しないガイド・フレームだけを肯定応答する。この肯定応答は、ＡＣＫ／ＮＯＡＣＫバー及びＥＡＲＬＹ／ＬＡＴＥバーの値に関係なく発生し、もちろん遅延肯定応答である。

【０１１３】

ＵＰ／ＤＯＷＮバー：アップまたはダウン制御値。この値は、フレームがアップサイドまたはダウンサイドのどちらで処理されるかを制御するために使用される。このフィールドは、ＲＥＳＰ／ＲＥＱバーが'１'ｂのとき無視される。全てのマルチキャスト・ガイド・フレームは、'０'ｂのＵＰ／ＤＯＷＮバー値を有する。更に、ＧＴＨハードウェア補助機構命令の使用を要求するガイド・コマンドは、'０'ｂのＵＰ／ＤＯＷＮバー値を有する

。
０：ダウンサイド処理
１：アップサイド処理
【０１１４】
ＧＴＨ／ＧＣＨバー：汎用ツリー・ハンドラまたはガイド・セル・ハンドラ制御値。この値は、ガイド・フレームを適切なピコプロセッサに転送するために使用される。
０：ＧＣＨピコプロセッサ
１：ＧＴＨピコプロセッサ
【０１１５】
ＴＢ：ターゲット・ブレード値。ＲＥＳＰ／ＲＥＱバーが'０'ｂのとき、このフィールドはＰｒｉｚｍａにより使用されるルーティング情報を含む。各ビット位置はターゲット・ブレードに対応する。この値が'０００００'ｈのとき、ガイド・フレームがこのブレードに当てはまるとみなされ、従って実行される。ＴＢフィールドの１つ以上のビット位置の'１'ｂの値は、セルが対応するターゲット・ブレードに経路指定されることを示す。ＲＥＳＰ／ＲＥＱバーが'１'ｂのとき、このフィールドは応答ブレードのＭｙ＿ＴＢ値を含む。
【０１１６】
ガイド・フレームのワード１は、相関関係子の値を含む（図８）。この値はＣＰソフトウェアにより割当てられ、ガイド・フレーム応答をそれらの要求に相関付ける。相関関係子は、機能を割当てられた複数のビットを含む。
【０１１７】
あらゆるガイド・コマンドは、コマンド制御情報フィールドで開始する。このコマンド制御は、ＧＣＨピココードがガイド・フレームを処理するのを支援する情報を含む。この情報のフォーマットが図９に示される。
【０１１８】
レングス値：この値は、制御情報に含まれる３２ビット・ワードの総数（コマンド・ワード０）、アドレス情報（コマンド・ワード１）、及びガイド・フレームのオペランド（コマンド・ワード２＋）部分を含む。
【０１１９】
完了コード値：このフィールドはＣＰにより初期化され、ガイド・コマンドを処理するとき、ＧＣＨピココードにより変更される。ＧＣＨピココードはこのフィールドを、コマンド・リスト内の先行ガイド・コマンドの完了ステータスとして使用する。全てのガイド・コマンド・リストは、終了区切り文字ガイド・コマンドで終了するので、最後のコマンドの完了ステータスが、終了区切り文字の完了コード・フィールド内に含まれる。
【０１２０】
ガイド・コマンド（ＧＣ）タイプ値（シンボル名）：

【表１】

シンボル名	タイプ値	タイプ記述
End_Delimiter	0000	ガイド・フレーム・シーケンスの終わりをマーク
Build_TSE_Free_List	0001	フリー・リストを作成
Software_Action	0010	ソフトウェア・アクションを実行
Unsolicited	0011	EPCピココードにより開始されるフレーム
Block_Write	0100	データ・ブロックを連続アドレスに書込む
Duplicate_Write	0101	重複データをレジスタまたはメモリに書込む
Read register	0110	レジスタまたはメモリ・データの読出しの要求または応答
	0111	予約済み
Insert_Leaf	1000	リーフを探索ツリーに挿入
Update_Leaf	1001	探索ツリーのリーフを更新
Read_Leaf Leaf	1010	リーフの読出しの要求及び応答
	1011	予約済み
Delete_Leaf	1100	探索ツリーのリーフを削除
	1101-1111	予約済み

【０１２１】

ガイド・フレームに含まれるアドレス情報は、ネットワーク・プロセッサのアドレッシング機構内の要素を識別する。アドレス情報フィールドの汎用形式が、図１０に示される。

【０１２２】

インタフェース装置は３２ビット・アドレッシング方式を採用する。このアドレッシング方式は、アドレス値を、インタフェース装置のあらゆるアクセス可能な構造に割当てる。これらの構造はプロセッサの内部に存在するか、プロセッサの制御に従い、インタフェースに接続される。これらの構造のあるものは、組み込みプロセッサ・コンプレックス（ＥＰＣ）により、ウェブ・インタフェースと呼ばれる内部インタフェースを介してアクセスされる。残りの構造は、メモリ制御装置インタフェースを介してアクセスされる。全ての場合において、アドレスの汎用形式が図１１に示される。

【０１２３】

ネットワーク制御装置は、主要チップ・アイランドに細分化される。各アイランドは固有のアイランドＩＤ値を与えられる。この５ビット・アイランドＩＤ値は、そのチップ・アイランドにより制御される構造のアドレスの、最上位５ビットを形成する。エンコード化されたアイランドＩＤ値とチップ・アイランド名との対応が、図１２に示される。ウェブ・アドレスの第２の部分は、次の最上位２３ビットを含む。このアドレス・フィールドは、構造アドレス部分と、エレメント・アドレス部分とに区分化される。各セグメントに対して使用されるビット数は、アイランド毎に異なる。一部のアイランドは数個の大きな構造だけを含むのに対して、他のアイランドは多くの小さな構造を含む。そうした理由から、これらのアドレス・セグメントには固定サイズが存在しない。構造アドレス部分は、

アイランド内のアレイをアドレス指定するために使用されるのに対して、エレメント・アドレス部分は、アレイ内の要素をアドレス指定するために使用される。アドレスの残りの部分は、ウェブ・インタフェースの３２ビット・データ・バス制限を調整する。この４ビット・ワード・アドレスは、アドレス指定されるエレメントの３２ビット・セグメントを選択するために使用される。これは３２ビットより広い構造エレメントを、ネットワーク制御装置のウェブ・データ・バスを介して、移動するために必要である。ワード・アドレス値'０'ｈは、構造エレメントの最上位３２ビットを指し示すのに対して、順次ワード・アドレス値は、構造エレメントの連続的な下位セグメントに対応する。アドレスのワード・アドレス部分は、ウェブ・インタフェースを介してアクセスされない構造に対しては、要求されない。この理由から、アップ・データ・ストア、制御メモリ、及びダウン・データ・ストアは、アドレスの最下位２７ビット全てを使用し、構造エレメントをアクセスする。このフォーマットの別の例外は、ＳＰＭインタフェースのアドレスである。この場合、アドレスの２７ビット全てが使用され、３２ビットよりも大きな幅のエレメントは存在しない。

【０１２４】

組み込みプロセッサ・コンプレックス（ＥＰＣ）は、インタフェース素子チップのプログラマビリティを提供し制御する。これは次のコンポーネントを含む（図１３参照）。

【０１２５】

Ｎ個の処理ユニット（ＧｘＨと呼ばれる）：ＧｘＨは、共通命令メモリに記憶されるピココードを同時に実行する。各ＧｘＨは処理ユニット・コア（ＣＬＰと呼ばれる）を含み、これは3ステージ・パイプライン、１６ＧＰＲ、及びＡＬＵを含む。各ＧｘＨはまた、例えばツリー構造探索エンジンのように、幾つかのコプロセッサを含む。

【０１２６】

命令メモリ：初期化の間にロードされ、フレームを転送し、システムを管理するピココードを含む。

【０１２７】

ディスパッチャ：アップ及びダウン・ディスパッチャ・キューからフレーム・アドレスをデキューする。デキューの後、ディスパッチャがアップまたはダウン・データ・ストア（ＤＳ）から、フレーム・ヘッダの一部をプリフェッチし、これを内部メモリに記憶する。ＧｘＨがアイドルになると、ディスパッチャが直ちに、コード命令アドレス（ＣＩＡ）のような適切な制御情報を有するフレーム・ヘッダを、ＧｘＨに渡す。ディスパッチャはまた、タイマ及び割込みを処理する。

【０１２８】

ツリー構造探索メモリ（ＴＳＭ）アービタ：各ＧｘＨが使用可能な多数の共用内部メモリ位置及び外部メモリ位置が存在する。このメモリは共用されるので、メモリへのアクセスを制御するために、アービタが使用される。ＴＳＭはピココードにより直接アクセスされ、例えばエージング・テーブルを記憶するために使用される。ＴＳＭはまた、ツリー構造探索の間に、ＴＳＥによりアクセスされる。

【０１２９】

完了ユニット（ＣＵ）：完了ユニットは２つの機能を実行する。第１に、これはＮ個の処理ユニットを、アップ及びダウンＥＤＳ（エンキュー、デキュー、及びスケジュール・アイランド）とインタフェースする。ＥＤＳはエンキュー・アクションを実行する。すなわち、フレーム・アドレスが、ＦＣＢページと呼ばれる適切なパラメータと一緒に、伝送キュー、廃棄キュー、またはディスパッチャ・キューのいずれかに、待ち行列化される。第２に、完了ユニットはフレーム・シーケンスを保証する。複数のＧｘＨが、同一のフローに属するフレームを処理している可能性があるので、これらのフレームがアップまたはダウン伝送キューに正しい順序でエンキューされるように、事前注意が払われなければならない。完了ユニットは、フレーム・ディスパッチ時に分類子ハードウェア補助機構により生成されるラベルを使用する。

【０１３０】

分類子ハードウェア補助機構：アップ・フレームでは、分類子ハードウェア補助機構は、フレーム・フォーマットの周知のケースを分類する。分類結果はフレーム・ディスパッチの間に、ＣＩＡ及び１つ以上のレジスタの内容に関して、ＧｘＨに渡される。ダウン・フレームでは、分類子ハードウェア補助機構は、フレーム・ヘッダに応じて、ＣＩＡを決定する。アップ及びダウン・フレーム・ディスパッチの両方のために、分類子ハードウェア補助機構は、フレーム・シーケンスを維持するために、完了ユニットにより使用されるラベルを生成する。

【０１３１】

アップ／ダウン・データ・ストア・インタフェース及びアービタ：各ＧｘＨはアップ及びダウン・データ・ストアをアクセスできる。すなわち、"より多くのＦＩＳＨ"を読出すとき、読出しアクセスが提供され、ＦＩＳＨプールの内容をデータ・ストアに書戻すとき、書込みアクセスが提供される。Ｎ個の処理ユニットが存在し、１度にそれらの１つだけがアップ・データ・ストアをアクセスでき、１度に１つだけがダウン・データ・ストアをアクセスできるので、各データ・ストアに対して、１つのアービタが必要とされる。

【０１３２】

ウェブ・アービタ及びウェブウォッチ・インタフェース：ウェブ・アービタはＧｘＨの間で、ウェブへのアクセスを調停する。全てのＧｘＨはウェブをアクセスでき、このことは、インタフェース装置内の全てのメモリ及びレジスタ機能をアクセスすることを可能にする。これにより、ＧｘＨは全ての構成領域を変更または読出すことができる。ウェブはインタフェース装置のメモリ・マップとみなされる。ウェブウォッチ・インタフェースは、３つのチップ入出力ピンを用いて、チップの外部からウェブ全体をアクセスできるようにする。

【０１３３】

デバッグ、割込み及び単一ステップ制御：ウェブはＧＣＨまたはウェブウォッチャが、必要に応じて、チップ上の各ＧｘＨを制御することを可能にする。例えば、ウェブはＧＣＨまたはウェブウォッチャにより、ＧｘＨ上で命令を単一ステップ操作するために使用される。

【０１３４】

ＰｏｗｅｒＰＣなどの組み込み汎用プロセッサ：

４つのタイプのＧｘＨが存在する（図１４参照）。

【０１３５】

ＧＤＨ（汎用データ・ハンドラ）：８個のＧＤＨが存在する。各ＧＤＨは完全なＣＬＰと、５個のコプロセッサ（次のセクションで述べられる）を有する。ＧＤＨは主に、フレームを転送するために使用される。

【０１３６】

ＧＣＨ（ガイド・セル・ハンドラ）：ＧＣＨは、ＧＤＨと正に同一のハードウェアを有する。しかしながら、ガイド・フレームだけがＧＣＨにより処理される。ＧＣＨがデータ・フレームについても処理できるか否かは（この場合、ＧＣＨはＧＤＨの役割をする）、ウェブ上でプログラマブルである（ＣＬＰ＿Ｅｎａレジスタ）。ＧＣＨはＧＤＨに比較して、追加のハードウェア、すなわち、ツリー挿入及び削除を実行するハードウェア補助機構を有する。ＧＣＨは、ガイド・セル関連ピココードを実行したり、エージングなどの、チップ及びツリー管理関連ピココードを実行したり、或いは制御情報をＣＰや別のＧＣＨと交換するために使用される。実行すべきこうしたタスクが存在しない場合、ＧＣＨはフレーム転送関連ピココードを実行し、この場合、丁度ＧＤＨのように動作することになる。

【０１３７】

ＧＴＨ（汎用ツリー・ハンドラ）：ＧＴＨはツリー挿入、ツリー削除、及びロープ管理を実行する追加のハードウェア補助機構を有する。ＧＰＱ内に（ツリー管理コマンドを含む）フレームが存在しない場合、ＧＴＨはデータ・フレームを処理する。

【０１３８】

ＧＰＨ（汎用ＰｏｗｅｒＰＣハンドラ）：ＧＰＨはＧＤＨ及びＧＴＨに比較して、追加のハードウェアを有する。ＧＰＨはメールボックス・インタフェース（Ｉ／Ｆ）を介して、汎用プロセッサとインタフェースする。

【０１３９】

ＧｘＨの数（１０個）は"最善推量"（best-guess）である。性能評価により、実際に要求されるＧｘＨの個数が決定される。アーキテクチャ及び構造は、より多くのＧｘＨに向けて完全にスケーラブルであり、唯一の制限はシリコン面積である（より大きなアービタ及び命令メモリを含むはずである）。

【０１４０】

各ＧｘＨは、図１５に示されるように構造化される。汎用レジスタ（ＧＲＰ）及び演算論理ユニット（ＡＬＵ）を有するＣＬＰに加え、各ＧｘＨは次の５つのコプロセッサを含む。

【０１４１】

（ＤＳ）コプロセッサ・インタフェース：ディスパッチャ、並びにアップ及びダウン・データ・ストアへの読出し及び書込みアクセスを提供するサブアイランドとインタフェースする。ＤＳインタフェースは、いわゆるＦＩＳＨプールを含む。

【０１４２】

ツリー構造探索エンジン・コプロセッサ（ＴＳＥ）：ＴＳＥはツリー内の探索を実行し、ツリー構造探索メモリ（ＴＳＭ）とインタフェースする。

【０１４３】

エンキュー・コプロセッサ：完了ユニット・インタフェースとインタフェースし、ＦＣＢページを含む。このコプロセッサは、ピココードがエンキュー・パラメータを含むＦＣＢページを作成するために使用する追加のハードウェア補助機構と共に、２５６ビット・レジスタを含む。一旦ＦＣＢページが作成されると、ピコプロセッサがエンキュー命令を実行し、ＦＣＢページが完了ユニットに転送される。

【０１４４】

ウェブ・インタフェース・コプロセッサ：このコプロセッサはウェブ・アービタとのインタフェースを提供し、インタフェース装置への書込み及び読出しを可能にする。

【０１４５】

チェックサム・コプロセッサ：ＦＩＳＨプール（後述）に記憶されるフレーム上に、チェックサムを生成する。

【０１４６】

処理ユニットは、イングレス処理とイーグレス処理との間で共用される。イングレス処理対イーグレス処理において、どれだけの帯域幅が確保されるかは、プログラマブルである。現インプリメンテーションでは、２つのモデルが存在し、一方は５０／５０（すなわち、イングレス及びイーグレスが同じ帯域幅を獲得する）で、他は６６対３４（すなわち、イングレスがイーグレスの２倍の帯域幅を獲得する）である。

【０１４７】

処理ユニットの動作はイベント・ドリブン方式である。すなわち、フレームの到来がイベントとして、またタイマまたは割込みのポッピングとして扱われる。ディスパッチャは異なるイベントを同様に扱う。但し、優先順位は存在する（第１が割込みで、第２がタイマ・イベントで、第３がフレーム到来イベント）。イベントが処理ユニットに渡されると、適切な情報が処理ユニットに与えられる。フレーム到来イベントでは、これはフレーム・ヘッダの一部と、ハードウェア分類子からの情報とを含む。タイマ及び割込みでは、これはコード・エントリ・ポイントと、イベントに関連する他の情報とを含む。

【０１４８】

フレームがイングレス側に到来し、このフレームの受信バイト数がプログラマブルしきい値を超える場合、フレーム制御ブロックのアドレスがＧＱに書込まれる。

【０１４９】

完全なフレームがイーグレス側で再組み立てされた場合、フレーム・アドレスがＧＱに

書込まれる。４つのタイプのＧＱが存在する（そして、各タイプに対して、図１４に示されるように、イングレス・バージョン及びイーグレス・バージョンが存在する）。すなわち、

ＧＣＱ：ＧＣＨにより処理されなければならないフレームを含む。

ＧＴＱ：ＧＴＨにより処理されなければならないフレームを含む。

ＧＰＱ：ＧＰＨにより処理されなければならないフレームを含む。

ＧＤＱ：任意のＧＤＨ（またはＧＣＨ／ＧＴＨがデータ・フレームを処理できる場合には、ＧＣＨ／ＧＴＨ）により処理され得るフレームを含む。ＧＤＱについては、複数の優先順位が存在し、高い優先順位でエンキューされたフレームは、低い優先順位でエンキューされたフレームより先に処理される。

【０１５０】

一部の処理ユニットは特殊化され得る。現インプリメンテーションでは、４つのタイプの処理ユニット（ＧｘＨ）が存在する（図１４参照）。すなわち、

ＧＤＨ（汎用データ・ハンドラ）：ＧＤＨは主に、フレームを転送するために使用される。

【０１５１】

ＧＣＨ（ガイド・セル・ハンドラ）：ＧＣＨは、ＧＤＨと正に同一のハードウェアを有する。しかしながら、ガイド・フレームだけがＧＣＨにより処理される。ＧＣＨがデータ・フレームについても処理できるか否かは（この場合、ＧＣＨはＧＤＨの役割をする）、ウェブ上でプログラマブルである（ＣＬＰ＿Ｅｎａレジスタ）。

【０１５２】

ＧＴＨ（汎用ツリー・ハンドラ）：ＧＴＨはＧＤＨ及びＧＣＨに比較して、追加のハードウェア、すなわち、ツリー挿入、ツリー削除、及びロープ管理を実行するハードウェア補助機構を有する。ＧＰＱ内に（ツリー管理コマンドを含む）フレームが存在しない場合、ＧＴＨはデータ・フレームを処理する。

【０１５３】

ＧＰＨ（汎用ＰｏｗｅｒＰＣハンドラ）：ＧＰＨはＧＤＨ及びＧＴＨに比較して、追加のハードウェアを有する。ＧＰＨはメールボックス・インタフェースを介して、組み込みＰｏｗｅｒＰＣとインタフェースする。

【０１５４】

実際のインプリメンテーションでは、ＧＣＨ、ＧＴＨ及びＧＰＨの役割は、単一の処理ユニット上で実現される。例えば、１インプリメンテーションはＧＣＨ及びＧＰＨに対して、１つの処理ユニットを有する。類似のコメントがＧＣＱ、ＧＴＯ及びＧＰＱについても適用できる。

【０１５５】

データ・ストア・コプロセッサの目的は、メディアから受信されたフレームを含むアップ・データ・ストアと、及びＰｒｉｚｍａ　Ａｔｌａｎｔｉｃから受信される再組み立て済みフレームを含むダウン・データ・ストアとインタフェースすることである。

【０１５６】

データ・ストア・コプロセッサはまた、タイマ・イベントまたは割込みのディスパッチの間に、構成情報を受信する。

【０１５７】

データ・ストア・コプロセッサは、フレーム上のチェックサムを計算できる。

【０１５８】

データ・ストア・コプロセッサはＦＩＳＨプール（８ＦＩＳＨを保持できる）と、スクラッチ・メモリ（８ＦＩＳＨを保持できる）と、アップまたはダウン・データベースからＦＩＳＨプール内容を読み書きする幾つかの制御レジスタとを含む。ＦＩＳＨプールは、データ・ストアのためのある種の作業領域とみなされる。すなわち、データ・ストアを直接読み書きする代わりに、大量のフレーム・データがデータ・ストアからＦＩＳＨプールに読出されるか、大量のデータがＦＩＳＨプールからデータ・ストアに書込まれる。転送

の単位はＦＩＳＨすなわち１６バイトである。

【０１５９】

ＦＩＳＨプールは、８ＦＩＳＨ、すなわち各々が１２８ビットの８ワード分を含むことができるメモリとみなされる。ＣＬＰプロセッサ・アーキテクチャでは、ＦＩＳＨプールは１２８バイトのレジスタ・アレイである。ＦＩＳＨプール内の各バイトは、７ビット・バイト・アドレス（０乃至１２７）を有し、アクセスは１６ビットまたは３２ビット・ベースで行われる。全てのレジスタ・アレイ同様、ＦＩＳＨプールは循環アドレス方式を有する。すなわち、ＦＩＳＨプール内のロケーション１２６で開始するワード（すなわち４バイト）のアドレス指定は、バイト１２６、１２７、０及び１を返却する。更に、データ・ストア・コプロセッサの観点から、ＦＩＳＨプール内のＦＩＳＨロケーションは、３ビットのＦＩＳＨアドレスを有する。

【０１６０】

フレーム・ディスパッチに際して、フレームの最初のＮ個のＦＩＳＨが、ディスパッチャにより自動的にＦＩＳＨプールにコピーされる。Ｎの値は、ポート構成メモリ（PortConfigMemory）内でプログラマブルである。一般に、アップ・フレーム・ディスパッチでは、Ｎは４に等しく、ダウン・ユニキャスト・フレーム・ディスパッチでは２で、ダウン・マルチキャスト・フレーム・ディスパッチでは４で、割込み及びタイマでは０である。

【０１６１】

ピココードはフレームからより多くのバイトを読出すことができ、この場合、データ・ストア・コプロセッサが自動的に、フレーム・データをＦＩＳＨプール内の次のＦＩＳＨアドレスに読出し、ＦＩＳＨプールの境界に達すると、自動的に０に循環する。また、ピココードはアップ／ダウン・データ・ストアを絶対アドレスにおいて読み書きすることもできる。

【０１６２】

ウェブ・コプロセッサはＥＰＣウェブ・アービタとインタフェースする。ＥＰＣウェブ・アービタは１０個のＧｘＨとウェブウォッチャとの間で、インタフェース装置のウェブ・インタフェース上でマスタになるものを調停する。これは全てのＧｘＨがウェブ上で読み書きすることを可能にする。

【０１６３】

インタフェース装置のメモリ・コンプレックスは、図１３に示される組み込みプロセッサ・コンプレックス（ＥＰＣ）の記憶機構を提供する。メモリ・コンプレックスはツリー構造探索メモリ（ＴＳＭ）アービタと、複数のオンチップ及びオフチップ・メモリを含む。メモリはツリー構造、カウンタ、及びピココードによりメモリ・アクセスを要求される他のものを記憶する。更に、メモリはフリー・リストやキュー制御ブロックなど、ハードウェアにより使用されるデータ構造を記憶するために使用される。ツリーのために割当てられない、またはハードウェアにより使用されないメモリ・ロケーションは、デフォルトでは、ピココードがカウンタやエージング・テーブルとして使用することができる。

【０１６４】

図１６は、メモリ・コンプレックスの詳細ブロック図を示す。ツリー構造探索メモリ（ＴＳＭ）アービタが、組み込みプロセッサ（ＧｘＨ）とメモリとの間の通信リンクを提供する。メモリは５個のオンチップＳＲＡＭと、１個のオフチップＳＲＡＭと、７個のオフチップＤＲＡＭとを含む。ＴＳＭアービタは、１０個の要求制御ユニット（各々が組み込みプロセッサＧｘＨの１つに接続される）と、１３個のメモリ・アービタ・ユニット（各々が各メモリに対応する）とを含む。各制御ユニット及びその接続されるＧｘＨが、全てのメモリをアクセスできるように、バス構造が要求制御ユニット及びアービタ・ユニットを相互接続する。

【０１６５】

制御ユニットは、組み込みプロセッサ（ＧｘＨ）とアービタとの間で、データを方向付けするために必要なハードウェアを含む。

【０１６６】

ＳＲＡＭアービタ・ユニットはとりわけ、組み込みプロセッサＧｘＨとオンチップ及びオフチップＳＲＡＭとの間の、データのフローを管理する。

【０１６７】

ＤＲＡＭアービタ・ユニットはとりわけ、組み込みプロセッサＧｘＨとオフチップＤＲＡＭ素子との間のデータのフローを管理する。

【０１６８】

各メモリ・アービタは"バックドア"（back-door）アクセスを含み、これは一般にチップの他の部分により使用され、最も高いアクセス優先順位を有する。

【０１６９】

ＤＲＡＭメモリは２つの動作モードで実行できる。すなわち、

１）ＴＤＭモード：ＤＲＡＭ内の４つのバンクへのメモリ・アクセスが、読出しウィンドウ及び書込みウィンドウを交互することにより実行される。読出しウィンドウでは、４つのバンクのいずれかへのアクセスが読出し専用であり、書込みウィンドウでは、４つのバンクのいずれかへのアクセスが書込み専用である。複数のＤＲＡＭに対してＴＤＭモードを使用することにより、ＤＲＡＭ間で幾つかの制御信号を共用することが可能になり、希少資源である幾つかのチップ入出力を節約できる。

【０１７０】

２）非ＴＤＭモード：ＤＲＡＭ内の４つのバンクへのメモリ・アクセスが、読出しと書込みの組み合わせとなり得、これは特定の規則に従わねばならない。例えば、アクセス・ウィンドウ内で、バンクＡでは読出しを、バンクＣでは書込みを実行できる。

【０１７１】

ＴＳＭアービタは、Ｎ個のリクエスタが同時にＭ個のメモリをアクセスすることを可能にする。複数のリクエスタが同一のメモリをアクセスしたい場合、ラウンドロビン・アービトレーションが実行される。

【０１７２】

Ｍ個のメモリは異なるプロパティを有することができる。現インプリメンテーションでは、３つのメモリ・タイプ、すなわち、内部ＳＲＡＭ、外部ＳＲＡＭ、及び外部ＤＲＡＭが存在する。

【０１７３】

Ｍ個のメモリ及びＮリクエスタは同種（homogeneous）であり、任意のリクエスタが任意のメモリをアクセスできる。

【０１７４】

一部のメモリは複数のサブメモリ（ＤＲＡＭ内の４つのバンクなど）に論理的に分割され、これらは論理的に同時にアクセスできる。

【０１７５】

Ｍ個のメモリの一部は、内部的に使用されるデータ構造を含む制御メモリとして使用され、これらはピコプロセッサと比較して、高い優先アクセスを有する。このことはまたチップのでバッグを可能にする。なぜなら、ピコプロセッサは制御メモリの内容を読出すことができるからである。

【０１７６】

アービタは読出しアクセス、書込みアクセス、及びread-add-writeをサポートする。それにより、Ｎビット整数がアトミック演算において、メモリの内容に追加される。

【０１７７】

また、メモリ内のオブジェクトの物理ロケーションが透過的となるように、汎用アドレス方式が、Ｍ個のメモリをアクセスするために使用される。

【０１７８】

ツリーの概念は、ツリー構造探索エンジンにより、情報を記憶及び検索するために使用される。検索、すなわちツリー探索及び挿入並びに削除は、キーにもとづき実行される。ここでキーは、例えばＭＡＣソース・アドレスのようなビット・パターンであるか、ＩＰソース・アドレスとＩＰ宛先アドレスの連結である。少なくともキーを含む情報が、リー

フと呼ばれる制御ブロック内に記憶される（後述のように、記憶ビット・パターンは実際にはハッシュ・キーである）。リーフは更に、エージング情報などの追加の情報や、ターゲット・ブレード及びターゲット・ポート番号などの情報を転送するユーザ情報を含み得る。

【０１７９】

３つのタイプ（ＦＭ、ＬＰＭ、ＳＭＴ）（すなわち、固定マッチ、最長プレフィックス・マッチ及びソフトウェア管理ツリー）、及び関連ツリー・タイプ探索が存在する。ツリー探索の間にリーフをチェックするオプションの追加の基準は、ベクトルマスクである。ローピング、エージング及びラッチは、探索性能を向上させるために使用される。

【０１８０】

ＦＭツリーの探索アルゴリズムが、図１７に示される。探索アルゴリズムはキーを含む入力パラメータに作用し、キーにハッシュを実行し、ダイレクト・テーブル（ＤＴ）をアクセスし、パターン探索制御ブロック（ＰＳＣＢ）を通じてツリーを探索し、リーフに行き着く（図１７）。３つのタイプのツリーが存在し、各々は異なる規則に従いツリー探索を発生させる独自の探索アルゴリズムを有する。例えば、固定マッチ（ＦＭ）ツリーでは、データ構造がパトリシア・ツリーである。リーフが見いだされるとき、このリーフは入力キーに一致する唯一の可能な候補である。ソフトウェア管理ツリーでは、リンク・リスト内でチェーニングされる複数のリーフが存在し得る。この場合、一致が見いだされるか、チェーンが尽きるまで、チェーン内の全てのリーフが入力キーと符合される。入力キーをリーフに記憶されるパターンと比較するいわゆる"最終比較"（compareat the end）操作が、リーフが真に入力キーに一致するか否かを確認する。リーフが見いだされ、一致が発生するとき、探索結果はＯＫであり、他の全ての場合にはＫＯである。

【０１８１】

探索操作への入力は、次のパラメータを含む。すなわち、

キー（１２８ビット）：キーは探索（または挿入／削除）の前に、特殊なピココード命令を用いて作成される。１つのキー・レジスタだけが存在する。しかしながら、ツリー構造探索が開始されると、キー・レジスタはＴＳＥが探索を実行するのと並行して、ピココードにより、次の探索のキーを作成するために使用される。これはＴＳＥがキーをバッシュ（bash）し、結果を内部ハッシュドキー・レジスタに記憶することによる（従って、実際には２つのキー・レジスタが存在する）。

【０１８２】

キー・レングス（７ビット）：このレジスタはキーの長さを含み、ビットで表す。これはキーの作成の間に、ハードウェアにより自動的に更新される。

【０１８３】

ＬＵ定義指標（LUDefindex）（８ビット）：これは探索が発生するツリーの完全な定義を含むＬＵ定義テーブル（LUDefTable）への指標である。ＬＵ定義テーブルは以下で詳述される。

【０１８４】

ＴＳＲＮｒ（１ビット）：探索結果はツリー探索結果領域０（ＴＳＲ０）またはＴＳＲ１に記憶される。これはＴＳＲＮｒにより指定される。ＴＳＥが探索を行っている間、ピココードは他のＴＳＲをアクセスし、前の探索の結果を分析することができる。

【０１８５】

ベクトル指標（６ビット）：ベクトルマスクをイネーブルされるツリーでは（ＬＵ定義テーブル内で指定される）、ベクトル指標がベクトルマスク内のビットを示す。探索の終わりに、このビットの値が返却され、ピココードにより使用される。

【０１８６】

図１７に示されるように、入力キーはハッシュドキーにハッシュされる。使用可能な６つの固定ハッシュ・アルゴリズムが存在する（１つの"アルゴリズム"はハッシュ関数を実行しない）。どのアルゴリズムが使用されるかについては、ＬＵ定義テーブル内で指定される。プログラマブル・ハッシュ関数が、柔軟性を追加するために使用され得る。

【０１８７】
ハッシュ関数の出力は常に１２８ビット数であり、オリジナル入力キーとハッシュ関数の出力との間には、１対１対応が存在する。後述のように、この特性はダイレクト・テーブルの後で開始するツリーの深さを最小化する。
【０１８８】
図１７の場合のように、カラーがツリーに対して許可される場合、１６ビット・カラー・レジスタが１２８ビット・ハッシュ関数出力に挿入される。挿入はダイレクト・テーブルの直後に発生する。すなわち、図示のように、ダイレクト・テーブルが2^N個のエントリを含む場合、１６ビット・カラー値がビット位置Ｎに挿入される。ハッシュ関数の出力は、挿入されるカラー値（イネーブルされる場合）と一緒に、ハッシュドキー・レジスタに記憶される。
【０１８９】
ハッシュ関数は、その出力内のほとんどのエントロピが最上位ビット側に存在するように定義される。ハッシュドキー・レジスタの最上位Ｎビットは、ダイレクト・テーブル（ＤＴ）への指標を計算するために使用される。
【０１９０】
探索はダイレクト・テーブルへのアクセスにより開始する。すなわち、ダイレクト・テーブル（ＤＴ）エントリがダイレクト・テーブルから読出される。ＤＴエントリを読出すために使用されるアドレスは、ハッシュドキーの最上位Ｎビットから計算され、ＬＵ定義テーブル内で定義されるツリー属性に関しても同様である。これについては以下で詳述される。ＤＴエントリはツリーのルートとみなされる。使用される特定のツリー・データ構造は、ツリー・タイプに依存する。この時点では、パトリシア・ツリー・データ構造がＦＭツリーとして、及びＬＰＭ及びＳＭＴツリーのためのパトリシア・ツリーの拡張として使用されることを述べれば十分であろう。
【０１９１】
８個のエントリ・ダイレクト・テーブル（ＤＴ）の使用例が、図１８に示される。ＤＴを使用することにより、探索時間（すなわち、アクセスされなければならないＰＳＣＢの数）が低減される。従って、ＤＴサイズを増加することにより、メモリ使用と探索性能との間でトレードオフが生じる。
【０１９２】
図１８から明らかなように、ＤＴエントリは次の情報を含む。すなわち、
１）エンプティ：このＤＴエントリに接続されるリーフは存在しない。
２）リーフを指し示すポインタ：このＤＴエントリに接続される１つのリーフが存在する。
３）ＰＳＣＢを指し示すポインタ：このＤＴエントリに接続される２つ以上のリーフが存在する。ＤＴエントリはツリーのルートを定義する。
【０１９３】
ソフトウェア管理ツリーの探索アルゴリズム、及びツリーを生成するアルゴリズムが、米国特許出願第０９／３１２１４８号で述べられている。
【０１９４】
"選択ビット・アルゴリズム"と称されるアルゴリズムは、特定のメトリックを使用し、規則のセットまたは領域内の"ルール"と称されるアイテムから選択されるビットにもとづき、二分探索ツリーを作成する。ここで述べる全ての例は、インターネット・プロトコル（ＩＰ）ヘッダに関連して議論されるが、任意のタイプの固定フォーマット・ヘッダが代わりに使用され得る。
【０１９５】
ＩＰでは、各ルールは次のサブセクション、すなわちソース・アドレス（ＳＡ）、宛先アドレス（ＤＡ）、ソース・ポート（ＳＰ）、宛先ポート（ＤＰ）、及びプロトコル（Ｐ）で作成される、特定のキーに関する。これらのデータはそれぞれ３２、３２、１６、１６及び８ビット長であり、従ってテストされるキーは１０４ビットから成る。選択ビット

・アルゴリズムは、１０４ビットの内の特に有用な幾つかのビットを見いだす。実際に２、３のビットをテストすることは、１つのルールを除く全てを、または２、３のルールを除く全てを、可能なアプリケーションから削除する。一部のルールでは、単純な比較操作による不等テストもまたふさわしい。ビット・テスト及び比較は、二分木において論理的に編成される。ツリーは、ビットを高速にテスト可能なハードウェア構造にマップされる。こうしたテストの結果、キーが適合する１つのルールまたは少数のルール（リーフ・チェーンと呼ばれる）が生成される。前者の場合、キーがルールにより詳細にテストされる。後者の場合、比較及び完全ルール・テストにより、キーがテストの枠内でテストされる。

【０１９６】

ルール・セット内の各ルールは、規則がキーに適合する最高優先順位の規則の場合に取られるアクションに関連付けられる。ルールは交差する（すなわち、１つのキーが２つまたはそれ以上のルールに適合する）。その場合、ルールは優先順位番号１、２、３、・・・を与えられ、任意の２つの交差するルールが異なる優先順位を有する（キーが２つ以上のルールに適合する場合、管理者はどのルールが上位になるかを宣言しなければならない）。従って、ビット・テスト及び比較後に、依然２つ以上のルールがテストされる場合、ルールは優先順位に従いテストされる。低い優先順位番号が、高い優先順位のルールを指定する。

【０１９７】

全く適合が見いだされない場合、デフォルト規定が指定される。

【０１９８】

最長プレフィックス・マッチング法の探索アルゴリズムが、米国特許第５７８７４３０号で述べられている。この方法は、前記データベースのノード（ルートノード）において入力するステップと、次の（子）ノードを識別するのに必要なエントリ部分だけを含む探索引数のセグメントと、第２のリンク情報とを、前記セグメントが消費されるか、前記第２のリンク情報を欠く（リーフ）ノードに達するまで連続的に処理することにより、ツリー状データベースを通じてあるノードから別のノードへの探索パスを決定するステップと、前記探索引数を、前記探索パスが終了するノードに記憶されるエントリと比較するステップと、前記現ノードにおいて、前記探索引数と前記エントリとの間に、少なくとも部分的な一致が見いだされない場合、前記現ノードの第１のリンク情報を処理することにより、前記探索パスを後戻りするステップと、前記少なくとも部分的な一致が見いだされるか、または前記ルート・ノードに達するまで、前の２つのステップを繰り返すステップとを含む。

【０１９９】

図１９は、メイン・スイッチング・ファブリック装置の実施例を示す。好適には、各インタフェース素子チップは、少なくとも２つの集積化並列直列変換ポートを有し、これらは並列データを受信して、それを高速直列データ・ストリームに変換する。そして、直列データがシリアル・リンクを介して、スイッチング・ファブリック装置に転送される。高速シリアル・リンクを介してスイッチング・ファブリック装置から受信されるデータは、別のＤＡＳＬにより、並列データに変換される。データ・アライン・シリアル・リンク（ＤＡＳＬ）と称されるシリアライザ／デシリアライザの実施例が、ここでは述べられる。

【０２００】

少なくとも１つのＤＡＳＬが、スイッチング・ファブリック装置をシリアル・リンクにインタフェースする。シリアル・リンクからのデータは、並列データに変換され、これがスイッチング・ファブリック装置に転送する。同様に、スイッチング・ファブリック装置からの並列データは直列データに変換され、シリアル・リンクに転送される。シリアル・リンクはスループットを向上するために集約される。

【０２０１】

図１９を更に参照すると、スイッチング・システムはスイッチ・ファブリック１１と、スイッチ・ファブリック入力ポート１５（１５－１，．．，１５－ｋ）に接続される入力

スイッチ・アダプタ１３（１３－１，．．，１３－ｋ）と、スイッチ・ファブリック出力ポート１９（１９－１，．．，１９－ｐ）に接続される出力スイッチ・アダプタ１７（１７－１，．．，１７－ｐ）とを含む。

【０２０２】

着信及び発信伝送リンク２１（２１－１，．．，２１－ｇ）及び２３（２３－１，．．，２３－ｒ）は、それぞれ回線（リンク）アダプタ２５（２５－１，２５－ｇ）及び２７（２７－１，．．，２７－ｒ）により、スイッチ・システムに接続される。伝送リンクは、ワークステーションや電話器などの接続ユニット（リンク指定ＷＳ）に、或いはローカル・エリア・ネットワーク（リンク指定ＬＡＮ）に、更にサービス統合デジタル網（ＩＳＤＮ）（リンク指定ＩＳＤＮ）や他の通信システムに、回線交換またはパケット交換トラフィックを伝搬する。更に、プロセッサはスイッチ・アダプタ１３及び１７に直接接続される。回線アダプタ（ＬＡ）及びスイッチ・アダプタ（ＳＡ）は、共通のインタフェースを有する。

【０２０３】

入力スイッチ・アダプタでは、パケット交換及び回線交換インタフェースから様々なサービスが収集され、一様なミニパケット（幾つかの可能な固定長の１つを有する）に変換される。ルーティング情報を含むヘッダは、スイッチの要求出力ポート（及び出力リンク）を指定する。入力スイッチ・アダプタ内でのミニパケット・フォーマット及びミニパケット生成に関する詳細、及び出力スイッチ・アダプタ内でのパケット解除に関する詳細について、次に述べることにする。

【０２０４】

スイッチ・ファブリックはミニパケットを、高速自己ルーティング相互接続ネットワークを介して、任意の入力ポートから任意の出力ポートに経路指定する。自己ルーティング・ネットワークの構造は、ミニパケットが競合無しに、内部的に同時に経路指定されることである。

【０２０５】

スイッチング・システムの心臓部は、スイッチ・ファブリックである。２つの異なるインプリメンテーションが考慮され、これらについて次に述べる。１インプリメンテーションでは、スイッチ・ファブリックは、各入力ポートに対して、それぞれの入力ポートを全ての出力ポートに接続する自己ルーティング二分木を含み、（ｋ個の入力ポートが提供される場合、）ｋ個のこうした二分木が組み合わされて提供される。他のインプリメンテーションでは、出力ＲＡＭを有するバス構造が、各出力構造に対してスライスとして提供され、全ての入力ポートをそれぞれの出力ポートに接続する。（ｐ個の出力ポートが提供される場合、）スイッチ・ファブリックは、ｐ個のこうしたスライスを組み合わせて提供される。

【０２０６】

ＤＡＳＬについては、１９９９年６月１１日出願の米国特許出願第０９／３３０９６８号で述べられている。ＤＡＳＬインタフェースは、ＣＭＯＳ　ＡＳＩＣなどのパラレル・インタフェースからデータを受信し、パラレル・インタフェースからのビットを、少数の並列ビット・ストリームに区分化する。少数の並列ビット・ストリームが次に高速直列ストリームに変換され、これが伝送メディアを介して、他のモジュールの受信機に移送される。制御インピーダンスを有する差動ドライバが、データの直列ビット・ストリームを伝送メディアに駆動する。

【０２０７】

ＤＡＳＬは、Ｎビット並列データとして表されるデータ・ストリームを、各々がｎビット（但しｎはＮの分数）を有する複数の部分に構文解析し、データ・ストリームの各ｎビット部分を直列化し、直列化された各部分を複数の並列チャネルの対応するチャネルを介して転送し、データ・ストリームの各転送部分を非直列化し、データ・ストリームをＮビット並列データとして復元する。

【０２０８】

以上、図面を参照しながら、本発明の好適な実施例について述べてきた。説明の中で使用された特定の用語は総称的な意味を成すもので、限定的な意味で使用されるものではない。

【図面の簡単な説明】

【０２０９】

【図１】本発明に従うインタフェース装置のブロック図である。

【図２】ＭＡＣのブロック図である。

【図３】異なるシステム構成内の他のコンポーネントと相互接続されるインタフェース装置を示す図である。

【図４】カプセル化されたガイド・フレームのフロー及び処理を示す図である。

【図５】内部ガイド・フレームのフロー及び処理を示す図である。

【図６】ガイド・セルの汎用フォーマットを示す図である。

【図７】フレーム制御情報のフォーマットを示す図である。

【図８】相関関係子のフォーマットを示す図である。

【図９】コマンド制御情報のフォーマットを示す図である。

【図１０】アドレス情報のフォーマットを示す図である。

【図１１】構造アドレッシングの汎用形式を示す図である。

【図１２】アドレッシング、アイランド・エンコードを示す表である。

【図１３】組み込みプロセッサ・コンプレックスのブロック図である。

【図１４】組み込みプロセッサの概略図である。

【図１５】ＧｘＨプロセッサの構造を示す図である。

【図１６】メモリ・コンプレックスのブロック図である。

【図１７】固定マッチ（ＦＭ）検索アルゴリズムのフローチャートである。

【図１８】ダイレクト・テーブルを使用する場合と、使用しない場合のデータ構造を示すフローである。

【図１９】Ｐｒｉｚｍａなどのスイッチング・システムのブロック図である。

【図２０】ＣＰのブロック図である。

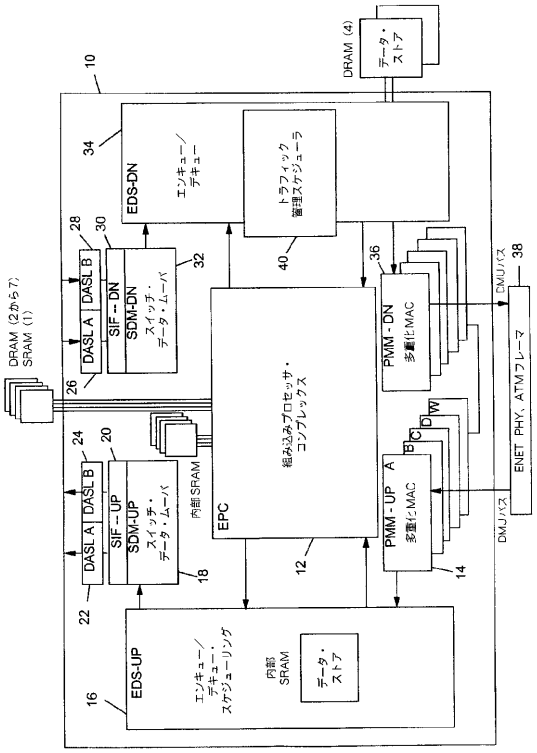
【図２１】ＥＤＳ－ＵＰ、ＥＤＳ　ＤＯＷＮ及びＥＰＣにおける、シングルチップ・ネットワーク・プロセッサの強調表示機能のブロック図である。

【符号の説明】

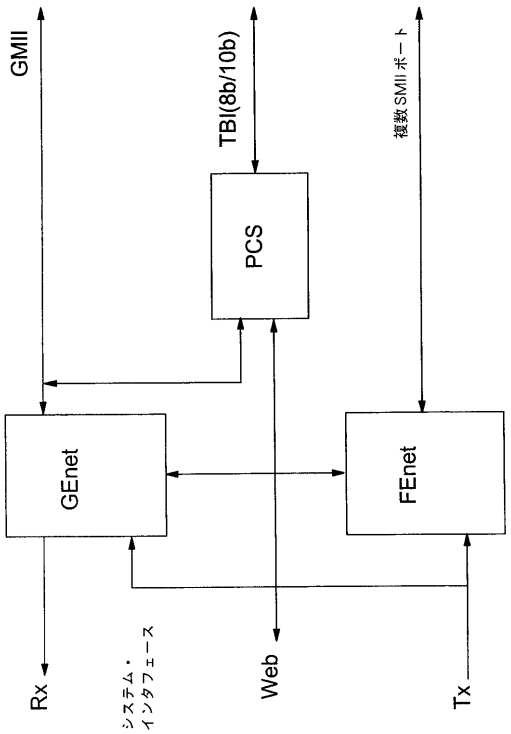
【０２１０】

１０、３８　インタフェース装置
１２　組み込みプロセッサ・コンプレックス（ＥＰＣ）
１４　多重化ＭＡＣアップ（ＰＰＭ－ＵＰ）、多重化ＭＡＣ
１６　エンキュー－デキュー－スケジューリング・アップ（ＥＤＳ－ＵＰ）
１８　スイッチ・データ・ムーバ・アップ（ＳＤＭ－ＵＰ）
２０、３０　システム・インタフェース（ＳＩＦ）
２２　データ・アライン・シリアル・リンクＡ（ＤＡＳＬＡ）
２４　データ・アライン・シリアル・リンクＢ（ＤＡＳＬＢ）
３２　ＳＤＭ－ＤＮ
３４　ＥＤＳ－ＤＮ
３６　イーグレス・イーサネット（Ｒ）ＭＡＣ
４０　トラフィック管理スケジューラ

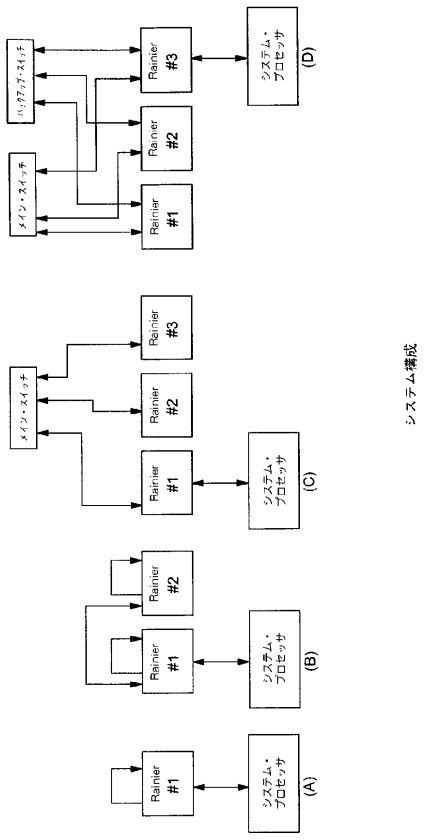
【図１】



【図２】

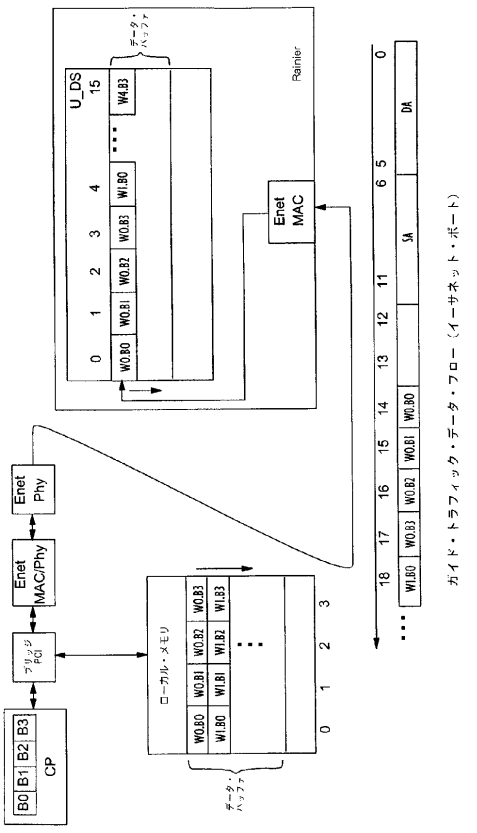


【図３】



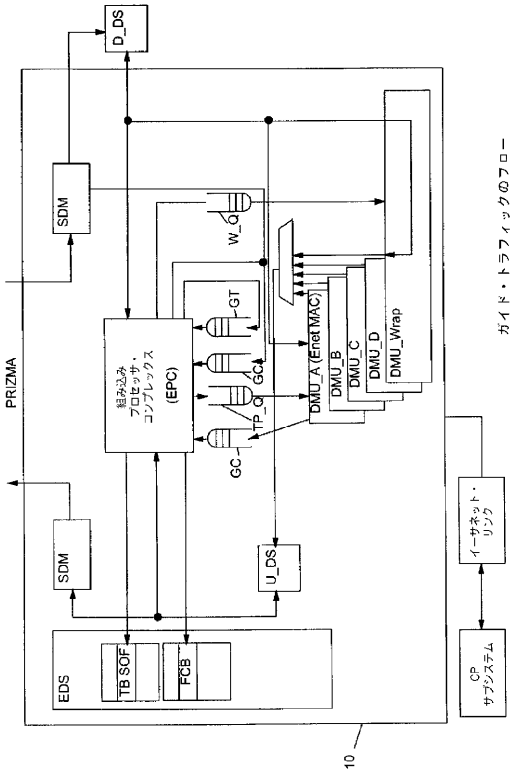
システム構成

【図４】

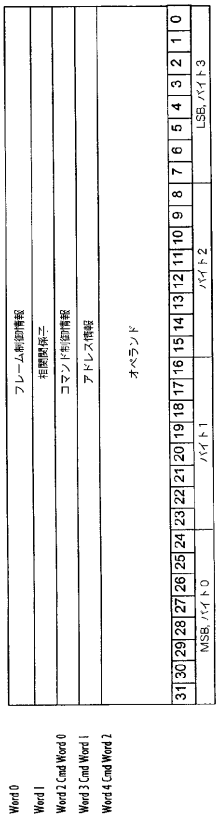


ガイド・トラフィック・データ・フロー（イーサネット・ポート）

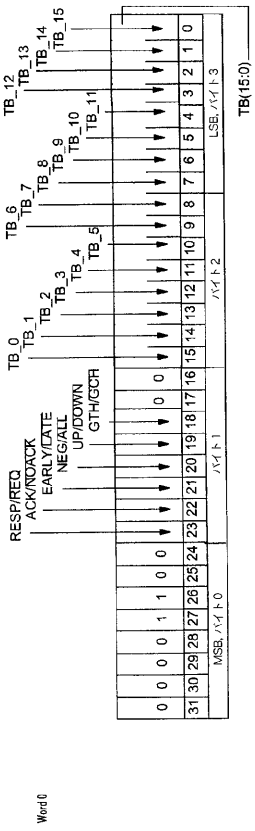
【図５】



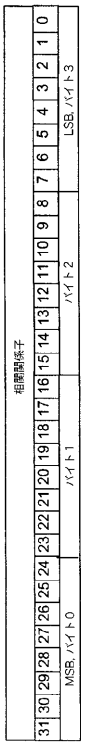
【図６】



【図７】



【図８】

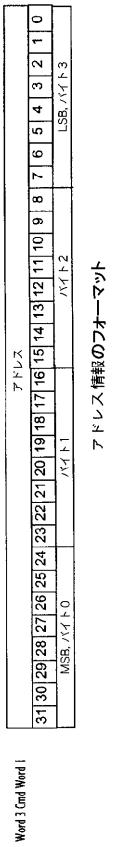


相関関係子のフォーマット

【図９】

コマンド制御情報のフォーマット

【図１０】



アドレス情報のフォーマット

【図１１】

アイランドID	構造アドレス	エレメント・アドレス	ワード・アドレス
5	23		4
	32		

構造アドレッシングの汎用形式

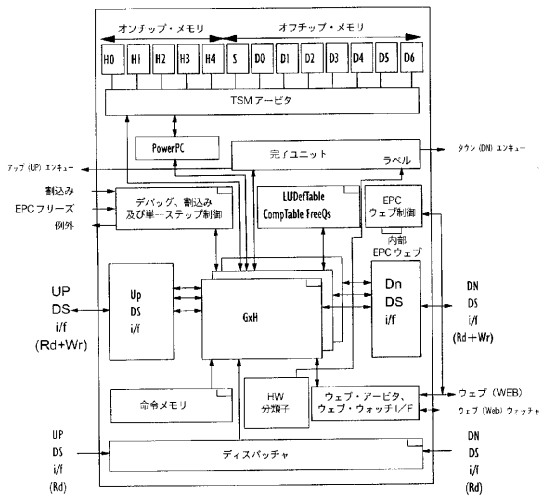
【図１２】

アイランドID	アイランド名
'00000'b	アップ・データ・ストア'
'00001'b	アップPMM
'00010'b	アップEDS
'00011'b	アップSDM
'00100'b	組み込みプロセッサ・コンプレックス
'00101'b	SPM
'00110-00111'b	予約済み
'01000'b	制御メモリ'
'01001-01111'b	予約済み
'10000'b	ダウン・データ・ストア'
'10001'b	ダウンPMM
'10010'b	ダウンEDS
'10011'b	ダウンSDM
'10100'b	構成レジスタ
'10101'b	DASL
'10110-11111'b	予約済み

'これらのアイランドはウェブ・インタフェースを介してアクセスできない

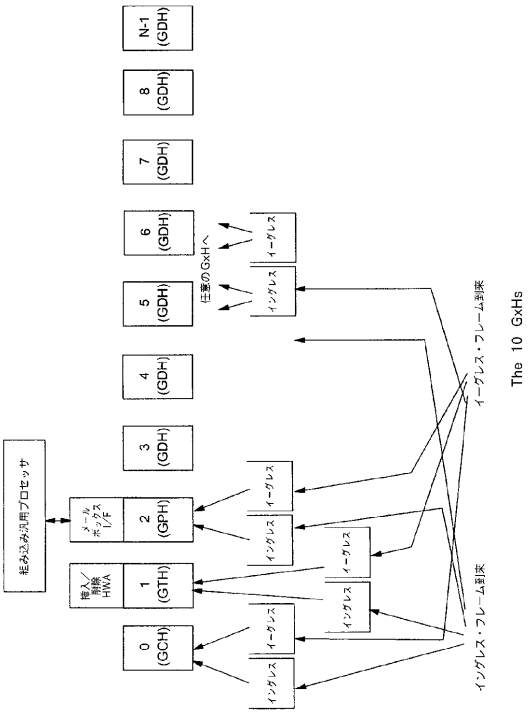
アドレッシング、アイランド・エンコーディング

【図１３】

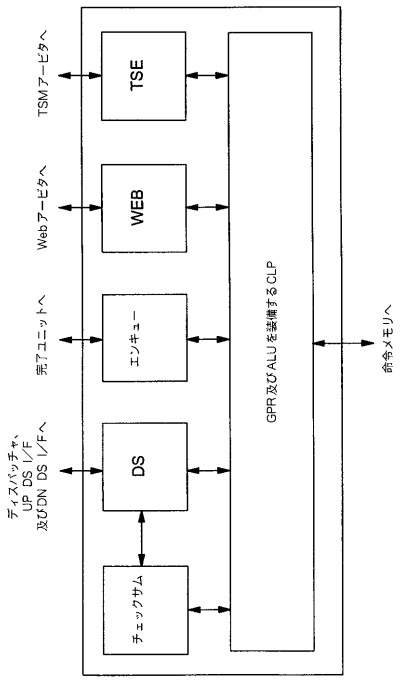


EPCのブロック図

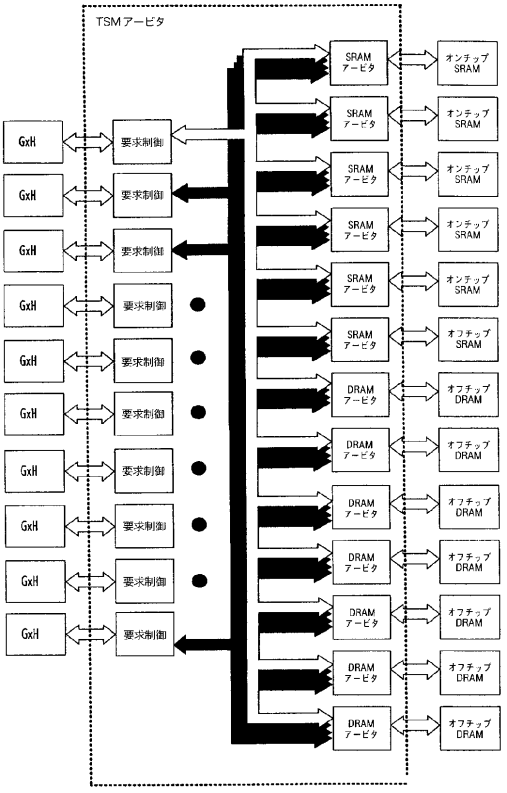
【図１４】



【図１５】

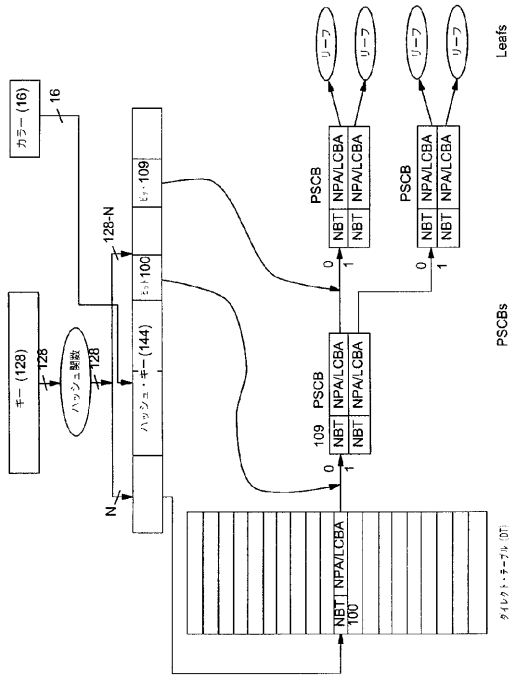


【図１６】

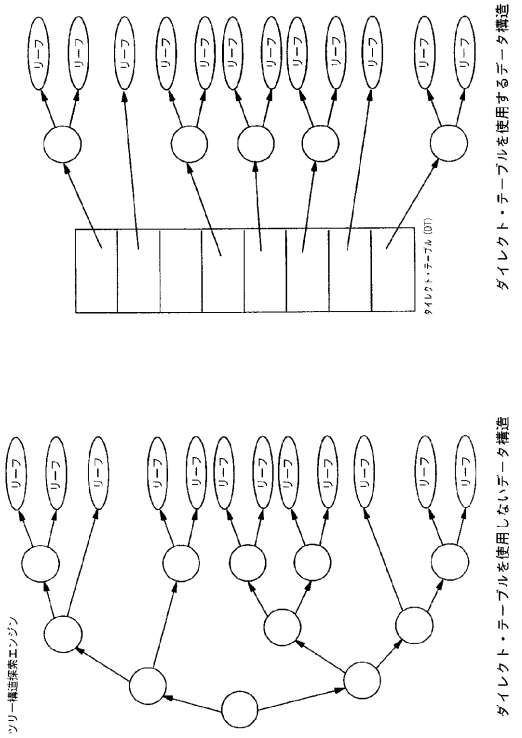


メモリ・コンプレックス

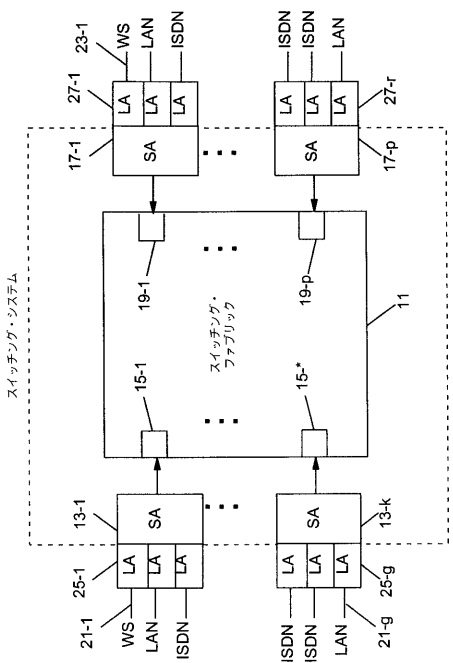
【図１７】



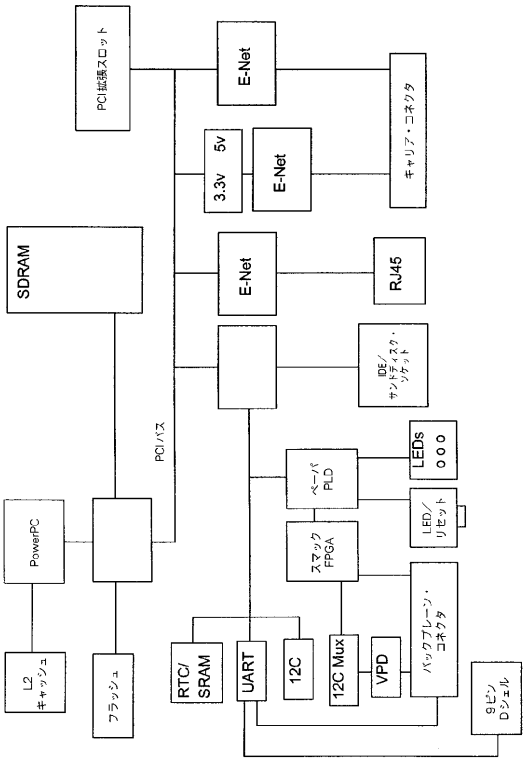
【図１８】



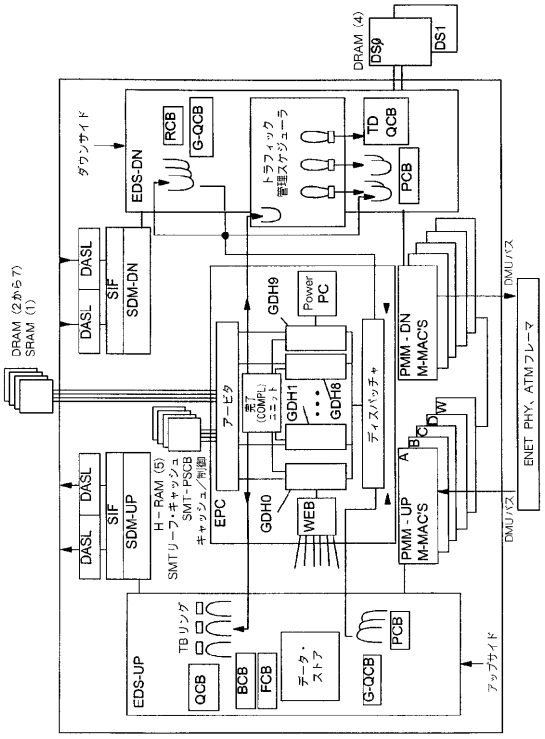
【図１９】



【図２０】



【図２１】


DRAM (4)
DS0
DS1
ダウンサイド
EDS-DN
RCB
G-QCB
トラフィック
管理スケジューラ
TD
QCB
PCB
DASL
SIF
SDM-DN
DRAM (2から7)
SRAM (1)
GDH9
Power
PC
PMM - DN
M-MAC'S
DMUバス
ENET PHY、ATMフレーマ
アービタ
GDH1
GDH8
ディスパッチャ
A
B
C
D
W
H－RAM (5)
SMTリーフ・キャッシュ
SMT－PSCB
キャッシュ／制御
EPC
GDH0
WEB
PMM - UP
M-MAC'S
DASL
SIF
SDM-UP
TBリング
EDS-UP
QCB
BCB
FCB
データ・
ストア
G-QCB
PCB
アップサイド

フロントページの続き

(72)発明者　バス、ブライアン、ミッチェル
アメリカ合衆国２７５０２、ノース・カロライナ州アペックス、オールド・スターブリッジ・ドライブ　４０２１
(72)発明者　カルビンナック、ジーン、ルイス
アメリカ合衆国２７５１１、ノース・カロライナ州カーリー、スプリング・ホロー・レーン　１１２
(72)発明者　ガロー、アンソニー、マッテオ
アメリカ合衆国２７５０２、ノース・カロライナ州アペックス、コーシャム・ドライブ　３３０８
(72)発明者　ヘデス、マルコ、シィ
アメリカ合衆国２７６１２、ノース・カロライナ州ラーレー、グランド・メナー・コート　４１０９、ナンバー　３０８
(72)発明者　ラオ、スリダール
アメリカ合衆国２７６１２、ノース・カロライナ州ラーレー、ビーバーブルック・ロード　５０２０、アパートメント　２０４
(72)発明者　シーゲル、マイケル、スティーブン
アメリカ合衆国２７６１３、ノース・カロライナ州ラーレー、ロワリー・ドライブ　１０６２５
(72)発明者　ヤングマン、ブライアン、アラン
アメリカ合衆国２７５１１、ノース・カロライナ州カリー、ニューポート・サークル　７０２
(72)発明者　ヴァープランケン、ファブリス、ジーン
フランス０６６１０、ラ・ゴード、ルート・ド・カーニュ　９１５２

審査官　清水　稔

(56)参考文献　特開平１１－０８８３４５（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｈ０４Ｌ　１２／５６