

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4334901号
(P4334901)

(45) 発行日 平成21年9月30日(2009.9.30)

(24) 登録日 平成21年7月3日(2009.7.3)

(51) Int. Cl. F 1
G 0 6 F 9/48 (2006.01) G 0 6 F 9/46 4 5 2 D

請求項の数 25 (全 44 頁)

(21) 出願番号	特願2003-97175 (P2003-97175)	(73) 特許権者	395015319
(22) 出願日	平成15年3月31日(2003.3.31)		株式会社ソニー・コンピュータエンタテインメント
(62) 分割の表示	特願2002-79358 (P2002-79358) の分割		東京都港区南青山二丁目6番21号
原出願日	平成14年3月20日(2002.3.20)	(74) 代理人	100099324
(65) 公開番号	特開2003-303134 (P2003-303134A)		弁理士 鈴木 正剛
(43) 公開日	平成15年10月24日(2003.10.24)	(74) 代理人	100108604
審査請求日	平成17年3月22日(2005.3.22)		弁理士 村松 義人
(31) 優先権主張番号	09/816,020	(74) 代理人	100111615
(32) 優先日	平成13年3月22日(2001.3.22)		弁理士 佐野 良太
(33) 優先権主張国	米国 (US)	(72) 発明者	鈴置 雅一
前置審査			東京都港区赤坂7丁目1番1号 株式会社 ソニー・コンピュータエンタテインメント 内

最終頁に続く

(54) 【発明の名称】 コンピュータ処理システム及びコンピュータで実行される処理方法

(57) 【特許請求の範囲】

【請求項1】

処理システムであって、

タスク処理が可能な複数の処理デバイスを有し、

選択されたタスクの処理時に、選択された処理デバイスによって用いられるタイムバジェットを定義するための絶対タイマを有し、各選択された処理デバイスは、各々対応する周波数を有するクロック信号に応答して作動するものであり、

前記タイムバジェットは、前記選択されたタスクを前記選択された処理デバイスによって完了させるための期間であって、タスクを処理するために選択された前記処理デバイスが用いる対応した前記各々対応するクロック周波数とは独立して設定され、かつ、前記タスクの完了後に前記処理デバイスが前記タイムバジェットの期間が経過するまではタスクの処理を行わないスタンバイ状態となるよう設定されるものであり、これにより、前記処理デバイスの実際の処理速度にかかわらず、タイムバジェットを用いて前記処理デバイスの処理結果を調整する、処理システム。

【請求項2】

前記期間は、前記選択された処理デバイスの各々に対して同一である、請求項1記載の処理システム。

【請求項3】

前記各々対応する周波数はそれぞれ異なるものである、請求項1記載の処理システム。

【請求項4】

10

20

前記選択された処理デバイスのうち第1のもののパワー管理は、前記タイムバジェットに基づいて実行される、請求項1記載の処理システム。

【請求項5】

各選択されたタスクに対するタイムバジェットは、ビジー時間とスタンバイ時間とを含む、請求項4記載の処理システム。

【請求項6】

前記第1の選択された一つの処理デバイスは、前記スタンバイ時間の間は、消費パワーレベルが小さくなる、請求項5記載の処理システム。

【請求項7】

前記スタンバイ時間は、前記選択された処理デバイスの対応する処理速度に基づいて、各選択された処理デバイスに対して動的に決定される、請求項6記載の処理システム。

10

【請求項8】

前記タイムバジェットは、前記期間が、前記選択された処理デバイスが前記タスクの処理を完了するために必要な時間を超えるように決定されるものである、請求項1記載の処理システム。

【請求項9】

処理システムであって、タスク処理が可能な複数の処理ユニットを有し、

前記タスクを実行可能な複数のサブ処理ユニットを有し、前記サブ処理ユニットの少なくとも一つは、前記処理ユニットと通信するよう動作可能なものであり、

前記タスクのうち選択されたタスクの処理において前記サブ処理ユニットのうち選択されたサブ処理ユニットと前記処理ユニットとにより使用されるタイムバジェットを決定するタイマを有し、前記選択されたサブ処理ユニットと前記処理ユニットのそれぞれは、各々対応する周波数を備えたクロック信号に応答して動作するものであり、

20

前記タイムバジェットは、前記選択されたタスクを前記選択されたサブ処理ユニットによって完了させるための期間であって、前記選択されたタスクを処理するために選択された前記処理ユニットが用いる対応したクロック周波数とは独立して設定され、かつ、前記タスクの完了後に前記処理デバイスが前記タイムバジェットの期間が経過するまではタスクの処理を行わないスタンバイ状態となるよう設定されるものであり、これにより、前記サブ処理ユニットの実際の処理速度にかかわらず、タイムバジェットを用いて前記サブ処理ユニットの処理結果を調整する、処理システム。

30

【請求項10】

前記複数のサブ処理ユニットは、前記処理ユニットによって制御される、請求項9記載の処理システム。

【請求項11】

前記各々対応する周波数は互いに異なるものである、請求項9記載の処理システム。

【請求項12】

コンピュータタスクの処理方法であって、

複数のタスクを処理するための複数の処理デバイスを用意し、前記処理デバイスのそれぞれは、各々対応する周波数を有するクロック信号に応答して動作するものであり、

前記複数の処理デバイスで用いられるタイムバジェットを確立し、前記タイムバジェットは、前記各々対応する周波数とは独立して前記複数のタスクを完了させるための期間を提供し、かつ、前記タスクの完了後に前記処理デバイスが前記タイムバジェットの期間が経過するまではタスクの処理を行わないスタンバイ状態となるよう設定されるものであり、

40

前記複数のタスクを処理するように前記タイムバジェットに従って前記処理デバイスを動作させ、これにより、前記処理デバイスの実際の処理速度にかかわらず、タイムバジェットを用いて前記処理デバイスの処理結果を調整する、方法。

【請求項13】

前記タスクは、前記処理デバイスの実際の処理時間にかかわらず、前記タイム・バジェットに基づいて処理が行われる、請求項12記載の方法。

50

【請求項 1 4】

前記生成は、前記複数の処理デバイスの処理速度とは独立してなされる、請求項 1 3 記載の方法。

【請求項 1 5】

前記タイムバジェットは、ビジー時間とスタンバイ時間とを含み、かつ、前記複数の処理デバイスの少なくとも一つは、前記スタンバイ時間においてスリープモードに入る、請求項 1 2 記載の方法。

【請求項 1 6】

前記スタンバイ時間は、前記処理デバイスのそれぞれに対して動的に決定される、請求項 1 5 記載の方法。

10

【請求項 1 7】

前記タイムバジェットは、前記タスクの完了時間が、前記処理デバイスによるタスク処理に必要な時間より長いものとなるよう設定される、請求項 1 2 記載の方法。

【請求項 1 8】

前記タスクにおける命令又はマイクロコードを分析し、
前記命令又はマイクロコードに少なくとも一つのオペレーションなし(N O O P)命令が挿入される、請求項 1 2 記載の方法。

【請求項 1 9】

前記選択された処理デバイスは、その処理速度が前記他の処理デバイスの処理速度とは異なるものである、請求項 1 8 記載の方法。

20

【請求項 2 0】

通信ネットワークを通じて互いに接続可能な複数のコンピューティングデバイスを有し、各コンピューティングデバイスは、処理エレメントを少なくとも一つ有し、この少なくとも一つの処理エレメントは、

タスクの処理が可能な処理ユニットを有し、

前記タスクを実行可能な複数のサブ処理ユニットを有し、前記サブ処理ユニットの少なくとも一つは、前記処理ユニットと通信するよう動作可能なものであり、

前記タスクのうち選択されたタスクの処理において前記サブ処理ユニットのうち選択されたサブ処理ユニットと前記処理ユニットとにより使用されるタイムバジェットを決定するタイマを有し、前記選択されたサブ処理ユニットと前記処理ユニットのそれぞれは、各々対応する周波数を備えたクロック信号に応答して動作するものであり、

30

前記タイムバジェットは、前記各々対応する周波数とは独立して前記選択されたタスクを完了するための期間を提供し、かつ、前記タスクの完了後に前記処理デバイスが前記タイムバジェットの期間が経過するまではタスクの処理を行わないスタンバイ状態となるよう設定されるものであり、これにより、前記サブ処理ユニットの実際の処理速度にかかわらず、タイムバジェットを用いて前記サブ処理ユニットの処理結果を調整する、処理システム。

【請求項 2 1】

前記各々対応する周波数は互いに異なる、請求項 2 0 記載の処理システム。

【請求項 2 2】

前記複数のサブ処理ユニットの少なくとも一つと、前記処理ユニットと、は、前記サブ処理ユニットの選択された一つによって実行される処理タスクのうち選択された一つにおける命令又はマイクロコードを分析して、前記タスクのうち他のタスクの処理の完了の時間に関連して選択されたタスクの完了の時間を決定するよう動作可能である、請求項 2 0 記載の処理システム。

40

【請求項 2 3】

前記複数のサブ処理ユニットのうち少なくとも一つの前記処理ユニットは、更に、前記命令又はマイクロコードにオペレーションなし(N O O P)命令を一つ以上挿入することを指示して、前記タスクのうち他のタスクの処理の完了の時間に関連して選択されたタスクの処理の完了にかかる時間を所定の時間として確立するよう動作可能である、請求項 2

50

2 記載の処理システム。

【請求項 2 4】

複数のタスクを処理するための少なくとも一つの処理デバイスによって用いられるコンピュータプログラムが記録された記録媒体であって、前記少なくとも一つの処理デバイスは、それぞれ周波数が設定されたクロック信号に応答して動作するものであり、前記コンピュータプログラムは、

前記少なくとも一つの処理デバイスによって使用されるタイムバジェットを確立し、前記タイムバジェットは、前記各々対応する周波数とは独立に前記複数のタスクを完了するための時間を与えるもので、かつ、前記タスクの完了後に前記処理デバイスが前記タイムバジェットの期間が経過するまではタスクの処理を行わないスタンバイ状態となるよう設定されるものであり、

10

前記複数のタスクを処理するために前記タイムバジェットに従って前記少なくとも一つの処理デバイスを動作させるものであり、これにより、前記少なくとも一つの処理デバイスの実際の処理速度にかかわらず、タイムバジェットを用いて前記少なくとも一つの処理デバイスの処理結果を調整する、記録媒体。

【請求項 2 5】

前記コンピュータプログラムは、

コンピュータタスクの選択された一つにおける命令又はマイクロコードを分析し、前記命令又はマイクロコードにオペレーションなし (NOOP) 命令を一つ以上挿入するものである、請求項 2 4 記載の記録媒体。

20

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はコンピュータ・プロセッサ用アーキテクチャとコンピュータ・ネットワークとに関し、広帯域環境におけるコンピュータ・プロセッサおよびコンピュータ・ネットワーク用アーキテクチャに関する。さらに、本発明は、このようなアーキテクチャのためのプログラミング・モデルに関する。

【0002】

【従来の技術】

コンピュータおよび現今のコンピュータ・ネットワーク(オフィスのネットワークで使用されるローカル・エリア・ネットワーク(LAN)やインターネットのようなグローバルネットワークなど)の計算用装置は、スタンド・アローン型の計算用として主として設計されてきた。コンピュータ・ネットワークを介するデータとアプリケーション・プログラム(“アプリケーション”)の共用は、これらのコンピュータおよびコンピューティング・デバイスの主要な設計目標ではなかった。これらのコンピュータとコンピューティング・デバイスは、また、様々な異なるメーカー(モトローラ、インテル、テキサス・インスツルメント、ソニーなど)によって製造された広範囲の異なるタイプのプロセッサを用いて一般に設計されたものである。これらのプロセッサの各々は、それ自身の特定の命令セットと命令セット・アーキテクチャ(ISA)とを持っている。すなわち、それ自身の特定のセットのアセンブリ言語命令と、これらの命令を実行する主演算装置と記憶装置のための構造とを有する。プログラマは、各プロセッサの命令セットとISAとを理解してこれらのプロセッサ用のアプリケーションを書くことを要求される。今日のコンピュータ・ネットワーク上でのコンピュータとコンピューティング・デバイスには異なった種類が混在していることから、データとアプリケーションの共用及びその処理は複雑になっている。さらに、この複数種が混在する環境に対する調整を行うために、多くの場合、同じアプリケーションであっても複数のバージョンを用意することが必要となっている。

30

40

【0003】

グローバルネットワーク、特にインターネットと接続されたタイプのコンピュータやコンピューティング・デバイスは広範囲に及ぶ。パーソナル・コンピュータ(PC)とサーバーに加えて、これらのコンピューティング・デバイスの中にはセルラー電話、移動用コンピ

50

ユーラ、個人用情報機器(PDA)、セット・トップ・ボックス、デジタルテレビ並びにその他の装置が含まれる。コンピュータやコンピューティング・デバイスにおいて異種製品が混在する中でのデータやアプリケーションを共用することに起因して、大きな問題が生じている。

【0004】

これらの問題を解決するためのいくつかの手法が試みられてきた。これらの手法の中には、特に、優れたインターフェースと複雑なプログラミング手法が含まれる。これらの解決方法では、処理パワーの実質的増加の実現がしばしば要求される。また、これらの解決方法では、アプリケーションの処理に必要な時間と、ネットワークを介するデータ伝送に必要な時間とが実質的に増加してしまうという結果がしばしば生じる。

10

【0005】

一般に、データは、対応するアプリケーションとは別個に、インターネットを介して伝送される。このアプローチでは、アプリケーションに対応した各セットの伝送データにアプリケーション自体をも送る必要はなくなっている。従って、このアプローチによって、必要とされる帯域幅の量は最少化されるものの、ユーザーには不満の原因となることも多々ある。つまり、クライアント側のコンピュータでは、この伝送データを利用するための適正なアプリケーション、あるいは最新のアプリケーションを入手できない事態も生じ得る。また、このアプローチでは、ネットワーク上のプロセッサによって用いられている複数の異種ISAと命令セットに対応して、各アプリケーション毎にバージョンの異なる複数のアプリケーションを用意することが要求される。

20

【0006】

Java(登録商標)モデルでは、この問題の解決が試みられている。このモデルでは、厳しいセキュリティ・プロトコルに準拠する小さなアプリケーション("アプレット(applet)")が用いられている。アプレットは、ネットワークを介してサーバー側コンピュータから送信されてクライアント側コンピュータ("クライアント")によって実行される。異なるISAを使用しているクライアント毎に、同じアプレットであっても異なるバージョンを送信するという事態を避ける必要があるため、すべてのJavaアプレットは、クライアント側のJava仮想マシン上で実行される。Java仮想マシンとは、JavaISAと命令セットとを備えたコンピュータをエミュレートするソフトウェアである。しかし、このソフトウェアは、クライアント側のISAとクライアント側の命令セットにより実行される。クライアント側ではISAと命令セットが各々異なるが、与えられるJavaの仮想マシンのバージョンは一つである。したがって、複数の各アプレットごとに異なるバージョンを用意する必要はない。各クライアントでは、当該クライアントにおけるISAと命令セットに対応した適正なJava仮想マシンだけをダウンロードすれば、javaアプレットを実行できる。

30

【0007】

【発明が解決しようとする課題】

各々の異なるISAと命令セットに対して異なるバージョンのアプリケーションを書かなければならないという課題は解決されているものの、Javaの処理モデルでは、クライアント側のコンピュータに対してソフトウェアの追加層が要求される。ソフトウェアのこの追加層のためにプロセッサの処理速度は著しく低下する。この速度の低下は、リアルタイムのマルチメディア・アプリケーションについて特に著しい。また、ダウンロードされたJavaアプレットの中にはウィルス、処理上の誤動作などが含まれている可能性がある。これらのウィルスと誤動作はクライアントのデータベースの破損やその他の損害の原因となる可能性がある。Javaモデルで用いられているセキュリティ用プロトコルでは、"サンドボックス(sandbox)"(Javaアプレットがそれ以上はデータを書き込むことができない、クライアント側のメモリ内のスペース)というソフトウェアを設けることにより、この問題の解決が試みられているとはいえ、このソフトウェア駆動型セキュリティ・モデルはその実行時に頻繁に不安定になり、より多くの処理が必要となる。

40

【0008】

50

リアルタイムの、マルチメディア・ネットワーク用アプリケーションがますます重要なものになりつつある。これらのネットワーク用アプリケーションでは非常に高速な処理が要求される。将来、そのようなアプリケーション用として毎秒何千メガビットものデータが必要となるかもしれない。ネットワークの現在のアーキテクチャ、および、特にインターネットのアーキテクチャ、並びに、Javaモデルなどで現在実施されているプログラミング・モデルでこのような処理速度に到達することは非常に難しい。

【0009】

したがって、新しいコンピュータ・アーキテクチャと、コンピュータ・ネットワーク用の新しいアーキテクチャと、新しいプログラミング・モデルとが求められている。この新しいアーキテクチャとプログラミング・モデルとによって、計算上の負担が付加されることなく、ネットワークの様々なメンバー間でのデータとアプリケーションの共用という問題が解決されることが望ましい。また、この新しいコンピュータ・アーキテクチャと、プログラミング・モデルとによって、ネットワークのメンバー間でのアプリケーションとデータの共用時に生じる、固有のセキュリティ上の問題も解決されることが望ましい。

10

【0010】

【課題を解決するための手段】

本発明の一実施形態においては、コンピュータと、コンピューティング・デバイスと、コンピュータ・ネットワーク（あるいはコンピュータ・ネットワークに代えて、コンピュータ・ネットワーク・システムや複数のコンピュータを備えたコンピュータ・システムというカテゴリーや形態とすることもできる）とのための新しいアーキテクチャが提供される。他の実施形態では、本発明は、これらのコンピュータ、コンピューティング・デバイスおよびコンピュータ・ネットワークのための新しいプログラミング・モデルを提供するものである。

20

【0011】

本発明によれば、コンピュータ・ネットワークのすべてのメンバー（ネットワーク上のすべてのコンピュータとコンピューティング・デバイス）は共通のコンピューティング・モジュールから構成される。この共通のコンピューティング・モジュールは均一な構造を有し、好適には同じISAが使用される。ネットワークのメンバーとして、クライアント、サーバー、PC、移動用コンピュータ、ゲーム用マシン、PDA、セット・トップ・ボックス、電気機器、デジタルテレビ、および、コンピュータ・プロセッサを用いるその他の装置が挙げられる。均一なモジュラー構造によって、ネットワークのメンバーによるアプリケーションとデータの効率的な高速処理と、ネットワークを介するアプリケーションとデータの高速伝送とが可能となる。またこの構造によって、様々なサイズと処理パワーを持つネットワークのメンバーの構成が単純化され、これらのメンバーによる処理用アプリケーションの作成が単純化される。

30

【0012】

また、本発明によれば、コンピュータ・ネットワークにおいて、前記ネットワークと接続された複数のプロセッサを有し、前記プロセッサの各々が、同じ命令セット・アーキテクチャを有する複数の第1の処理ユニットと、前記第1の処理ユニットを制御するための第2の処理ユニットとを有し、前記第1の処理ユニットが、前記ネットワークを介して伝送されるソフトウェア・セルを処理するために作動可能であり、前記ソフトウェア・セルの各々が、前記命令セット・アーキテクチャと互換性のあるプログラムと、前記プログラムと関連付けられたデータと、前記ネットワークを介して伝送される前記ソフトウェア・セルのすべての間で前記ソフトウェア・セルを一意的に識別するための識別子（例えばセルの識別番号）と、を有することを特徴とするコンピュータ・ネットワークも提供される。

40

【0013】

なお、本発明によれば、コンピュータ・ネットワークと接続される複数のプロセッサを有するコンピュータ・システムであって、前記プロセッサの各々が、同じ命令セット・アーキテクチャを有する複数の第1の処理ユニットと、前記第1の処理ユニットを制御するための第2の処理ユニットとを有し、前記第1の処理ユニットが、前記ネットワークを介し

50

て伝送されるソフトウェア・セルを処理するために作動可能であり、前記ソフトウェア・セルの各々が、前記命令セット・アーキテクチャと互換性のあるプログラムと、前記プログラムと関連付けられたデータと、前記ネットワークを介して伝送される前記ソフトウェア・セルのすべての間で前記ソフトウェア・セルを一意的に識別するための識別子（例えばセルの識別番号）と、を有することを特徴とするコンピュータ・システムも提供される。

【0014】

加えて、本発明によれば、コンピュータ・ネットワークを介する伝送用ソフトウェア・セルのデータ・ストリームにおいて、前記コンピュータ・ネットワークが、複数のプロセッサを有し、前記ソフトウェア・セルの各々が、前記プロセッサの中の1以上によって処理するためのプログラムと、前記プログラムと関連付けられたデータと、前記ネットワークを介して伝送されるすべてのソフトウェア・セルの中で前記ソフトウェア・セルを一意的に識別するグローバルな識別子と、を有することを特徴とするデータ・ストリームも提供される。なお、上記構成において、「データ・ストリーム」という形態に代えて、「データ構造」という形態、あるいは「上述のような構造を有するデータ」という形態で本発明を提供することも可能である。

10

【0015】

他の実施形態では、本発明は、ネットワークを介してデータとアプリケーションを伝送するための、また、ネットワークのメンバー間でデータとアプリケーションを処理するための新しいプログラミング・モデルを提供する。このプログラミング・モデルでは、ネットワークのいずれのメンバーでも処理できる、ネットワークを介して伝送されるソフトウェア・セルが使用される。各ソフトウェア・セルは同じ構造を有し、アプリケーションとデータの双方を含むことが可能である。モジュラー型コンピュータ・アーキテクチャによって提供される高速処理と伝送速度の結果、これらのセルの高速処理が可能となる。アプリケーション用コードは同じ共通の命令セットとISAに好適に基づくものである。各ソフトウェア・セルは、グローバルな識別子(グローバルID)と、セルの処理に必要な計算用リソースの量について説明する情報とを好適に含むことが望ましい。すべての計算用リソースは同じ基本構造を有し、同じISAが用いられているので、このセルの処理を実行する特定のリソースは、ネットワーク上のどこにでも配置が可能となり、動的に割り当てることができる。

20

30

【0016】

基本となる処理用モジュールはプロセッサ・エレメント(PE)である。PEは、好適には、処理ユニット(PU)、ダイレクト・メモリ・アクセス・コントローラ(DMAC)および複数の付加処理ユニット(APU)を具備することが望ましい。好ましい実施形態では、1つのPEは8つのAPUを具備する。PUとAPUとは、クロスバ・アーキテクチャを好適に備えている共用ダイナミック・ランダム・アクセス・メモリ(DRAM)を用いてリアルタイムで通信を行う。PUは、APUによって、データとアプリケーションの処理のスケジュール管理と全般的管理とを行う。APUは並列的かつ独立にこの処理を実行する。DMACは、共用DRAMに格納されているデータとアプリケーションへのアクセス制御をPUとAPUとによって行う。

40

【0017】

このモジュラー構造によれば、ネットワークのメンバーによって用いられるPEの数は、そのメンバーが必要とする処理パワーに基づく。例えば、1台のサーバーは4つのPEを用いることができ、1台のワークステーションは2つのPEを用いることができ、1つのPDAは1つのPEを用いることができる。特定のソフトウェア・セルの処理に割り当てられるPEのAPUの数は、そのセル内のプログラムとデータの複雑さと大きさによって決まる。

【0018】

好ましい実施形態では、複数のPEが1つの共用DRAMと関連付けられる。好適には、DRAMは複数のセクションに分割され、これらのセクションの各々は複数のメモリ・バ

50

ンクに分割される。特に好ましい実施形態では、DRAMは64個のメモリ・バンクを有し、各バンクは1メガバイトの記憶容量を有する。DRAMの各セクションは、好適には、バンク・コントローラによって制御されることが望ましく、PEの各DMACは、好適には、各バンク・コントローラにアクセスすることが望ましい。したがって、この実施形態の各PEのDMACは、共用DRAMの任意の部分へのアクセスが可能となる。

【0019】

別の態様では、本発明は、共用DRAMからのAPUのデータの読み出しと、共用DRAMへのデータの書き込みのための同期システムと方法とを提供する。このシステムによって、DRAMを共用している複数のAPUと複数のPEとの間のコンフリクトが防止される。このシステムと方法とによれば、DRAM領域が指定され、複数のフル・エンティ・ビットが格納される。これらのフル・エンティ・ビットの各々は、DRAMの指定領域に対応する。この同期システムはDRAMのハードウェアの中に統合化されるので、ソフトウェアの中で実行されるデータ同期方式の計算上のオーバーヘッドはこのシステムによって防止される。

10

【0020】

また本発明によって、DRAM内にサンドボックスが設けられ、1つのAPUのプログラム処理用データから生じる、別のAPUのプログラム処理用データの破損に対するセキュリティが与えられる。各サンドボックスによって、データの読み出しや書き込みが不可能となる共用DRAM領域が画定される。

【0021】

別の態様では、本発明は、PUがAPUへコマンドを出して、アプリケーションとデータのAPUによる処理を開始するためのシステムと方法とを提供するものである。これらのコマンドは、APU遠隔処理命令(ARPC)と呼ばれ、このコマンドによって、APUがコプロセッサの役割を演じることなく、アプリケーションとデータのAPUによる並列処理のPUによる全般的管理と調整が可能となる。

20

【0022】

他の実施形態では、本発明によって、ストリーミング・データ処理用の専用パイプライン構造を設定するシステムと方法とが提供される。このシステムと方法によれば、PUによってこれらのストリーミング・データの処理を行うために、APUの調整グループと、これらのAPUと関連するメモリサンドボックスの調整グループとが設定される。パイプラインの専用APUとメモリサンドボックスとは、データ処理が行われない時間中もパイプライン専用のままである。言い換えれば、専用APU及びこれらの専用APUと関連するサンドボックスとは、この期間中は予約状態となる。

30

【0023】

他の実施形態では、本発明はタスク処理用の絶対タイマーを提供する。この絶対タイマーは、アプリケーションとデータの処理用としてAPUが使用するクロック周波数に依存しない。アプリケーションは、絶対タイマーによって定義される、タスク用の時間に基づいて書かれる。APUが使用しているクロック周波数が、APUの機能の改善などに起因して増加しても、絶対タイマーによって定義される所定のタスク用の時間はそのまま同じである。この方式によれば、古いAPUにおける遅い処理時間を前提として書かれた古いアプリケーションの処理を、これらの新しいAPUでは行わせないこととする必要がなく、かつ、新しいバージョンのAPUによる処理時間の向上を実現することが可能になる。

40

【0024】

また本発明は、より処理速度が高速な新しいAPUを、古いAPUにおける遅い処理速度を前提として書かれた古いアプリケーションの処理に用いることを可能にする、他の方式をも提供するものである。この方式では、速度の改善によって生じるAPUの並列処理の調整における問題の処理の間に、これらの古いアプリケーションの処理時にAPUが使用している命令(マイクロコード)が分析される。APUによる処理の順番がプログラムが予期する順番どおりに維持されるよう、“オペレーションなし”(“NOOP”)命令が、これらのAPUのいくつかによって実行される命令の中へ挿入される。これらの命令の中へ

50

これらの“NOOP”を挿入することにより、APUによるすべての命令を実行するための正しいタイミングが維持される。

【0025】

他の実施形態では、本発明は、光導波路が統合化される集積回路を含むチップ・パッケージを提供するものである。

また、本発明によれば、プログラムと前記プログラムと関連付けられたデータとを格納する第1メモリと、前記プログラムとそれに関連付けられた前記データとを処理する複数の第1処理ユニットと、前記第1処理ユニットによる前記第1メモリへのアクセスを制御するためのメモリ・コントローラと、アクセス・テーブルとキー・テーブルを格納する第2メモリであって、前記アクセス・テーブルは、複数のアクセス・エントリを有し、各アクセス・エントリは、アクセス・キーと、前記アクセス・キーと関連付けられた前記第1メモリ内の1つのメモリ・スペースの識別子を含み、前記キー・テーブルは、複数のキー・エントリを有し、各キー・エントリは、前記第1処理ユニットの識別子と前記第1処理ユニットと関連するリクエスト・キーを含んで成る第2メモリと、前記第1処理ユニットによる前記プログラムとそれに関連付けられた前記データの処理を制御するための第2処理ユニットであって、前記アクセス・テーブルと前記キー・テーブルの作成と更新の操作が可能で、更に、前記第1処理ユニットによる前記プログラムの1つの処理を行う指示を出すことが可能となる前記第2処理ユニットとを有し、前記第1処理ユニットの1つは、前記1つのプログラムの処理において前記メモリ・コントローラにリクエストを出して、前記第1メモリ内の格納位置にアクセスすることが可能であり、前記メモリ・コントローラは、前記リクエストに回答して、前記キー・テーブル内の前記第1処理ユニットの1つと関連付けられたリクエスト・キーと、前記アクセス・テーブル内のアクセス・キーとを比較して、前記アクセス・キーの1つが、前記キー・テーブル内の前記第1処理ユニットの1つと関連付けられたリクエスト・キーと一致して、かつ、前記格納位置が、前記アクセス・テーブル内の前記1つのアクセス・キーと関連付けられたメモリ・スペースと一致する場合は、前記リクエストの実行が可能であることを特徴とするコンピュータ処理システムも提供される。

また、コンピュータで実行される処理方法において、第1メモリに、プログラムと、前記プログラムと関連付けられたデータと、を格納するステップと、複数の第1処理ユニットを用いて、前記プログラムとそれに関連付けられた前記データとを処理するステップと、メモリ・コントローラを用いて、前記第1処理ユニットによる前記第1メモリへのアクセスを制御するステップと、第2処理ユニットを用いて、第2メモリに、アクセス・テーブルとキー・テーブルを作成するステップであって、前記アクセス・テーブルは、複数のアクセス・エントリを有し、各アクセス・エントリが、前記アクセス・キーと、前記アクセス・キーと関連付けられた前記第1メモリのメモリ・スペースの識別子を含み、前記キー・テーブルが、複数のキー・エントリを有し、各キー・エントリが、1つの前記第1処理ユニットの識別子と、前記第1処理ユニットと関連付けられたリクエスト・キーを含むように構成するステップと、前記第2処理ユニットを用いて、前記第1処理ユニットによる前記プログラムとそれに関連付けられた前記データの処理を制御するステップと、前記第2処理ユニットを用いて、前記第1処理ユニットのうちの1つに、前記プログラムのうちの1つの処理を命令するステップと、前記1つの第1処理ユニットから、前記1つのプログラムの処理において、前記メモリ・コントローラにリクエストを出して、前記第1メモリ内の格納位置へアクセスするステップと、前記リクエストに回答して、前記1つの第1処理ユニットと関連付けられたリクエスト・キーと、前記アクセス・テーブル内の前記アクセス・キーとを比較するステップと、1つの前記アクセス・キーが、前記キー・テーブル内の前記1つの処理ユニットと関連付けられたリクエスト・キーと一致し、前記格納位置が、前記アクセス・テーブル内で指定された前記1つのアクセス・キーと関連付けられたメモリ・スペースと一致する場合は、前記リクエストを実行するステップと、を有することを特徴とする方法も提供される。

また、プログラムと前記プログラムと関連付けられたデータを格納する第1メモリであっ

10

20

30

40

50

て、複数のアドレス可能な格納位置を有し、アドレス可能な各格納位置は、前記アドレス可能な格納位置に関連付けられるとともに前記アドレス可能な格納位置のアクセス・キーを含む追加メモリ・セグメント有して成る第1メモリと、前記プログラムとそれに関連付けられた前記データとを処理するための複数の第1処理ユニットと、前記第1メモリへの前記第1処理ユニットによるアクセス制御のためのメモリ・コントローラと、複数のキー・エントリから成るキー・テーブルを格納するための第2メモリであって、各キー・エントリが、前記第1処理ユニットの1つの識別子と、前記第1処理ユニットと関連付けられたリクエスト・キーとを含んでなる第2メモリと、前記第1処理ユニットによる前記プログラムとそれに関連付けられた前記データの処理を制御するための第2処理ユニットであって、前記アクセス・キーの割り当てと更新および前記アクセス・キー・テーブルの作成と更新が可能で、更に、前記第1処理ユニットのうちの1つによる前記プログラムの1つの処理を指示することが可能となる第2処理ユニットとを有し、前記第1処理ユニットの1つは、前記プログラムの1つの処理において前記メモリ・コントローラにリクエストを出して、前記第1メモリ内の格納位置をアクセスすることが可能であり、前記メモリ・コントローラは、前記リクエストにตอบสนองして、前記キー・テーブル内の前記第1処理ユニットの1つと関連付けられたリクエスト・キーと、前記1つのアドレス可能な格納位置と関連付けられている追加メモリ・セグメントに含まれているアクセス・キーとを比較して、前記キー・テーブル内の前記第1処理ユニットの1つと関連付けられたリクエスト・キーと、前記1つのアドレス可能な格納位置と関連付けられている追加メモリ・セグメントに含まれている前記アクセス・キーが一致する場合は、前記リクエストを実行することが可能であることを特徴とするコンピュータ処理システムも提供される。

更に、コンピュータ処理システムにてなされる処理方法において、第1メモリに、プログラムと前記プログラムの関連データを格納するステップであって、複数のアドレス可能な格納位置を有し、前記アドレス可能な各格納位置には、前記アドレス可能な格納位置に関連付けられた追加メモリ・セグメントを含むステップと、各々の前記アドレス可能な格納位置用の各々の前記追加メモリ・セグメントに、前記アドレス可能な格納位置のためのアクセス・キーを格納するステップと、複数の第1処理ユニットを用いて、前記プログラムとそれに関連付けられた前記データを処理するステップと、メモリ・コントローラを用いて、前記第1処理ユニットによる前記第1メモリへのアクセスを制御するステップと、第2メモリに、複数のキー・エントリから成るキー・テーブルを格納するステップであって、各キー・エントリが、前記第1処理ユニットと関連付けられた識別子と、前記第1処理ユニットと関連付けられたリクエスト・キーを含むステップと、前記第2処理ユニットを用いて、前記第1処理ユニットによる前記プログラムとそれに関連付けられた前記データの処理を制御するステップと、前記第2処理ユニットを用いて、各々の前記アクセス・キーを割り当てるステップと、前記第2処理ユニットを用いて、前記キー・テーブルを作成するステップと、前記第2処理ユニットを用いて、前記第1処理ユニットのいずれか1つに、前記プログラムの1つを実行する命令を出すステップと、前記第1処理ユニットの1つから、前記1つのプログラムの処理において、1つの前記アドレス可能な格納位置へアクセスするように、前記メモリ・コントローラへリクエストを出すステップと、前記リクエストにตอบสนองして、前記キー・テーブル内で前記1つの第1処理ユニットと関連付けられたリクエスト・キーと、前記1つのアドレス可能な格納位置に格納されているアクセス・キーとを比較するステップと、前記キー・テーブル内の前記1つの処理ユニットと関連付けられたリクエスト・キーと前記アドレス可能な格納位置と関連付けられた前記追加メモリ・セグメント内に格納されている前記アクセス・キーが一致する場合は、前記リクエストを実行するステップと、を有することを特徴とする方法も提供される。

【0026】

【発明の実施の形態】

図1に、本発明によるコンピュータ・システム101のアーキテクチャ全体を示す。

【0027】

この図に例示されているように、システム101にはネットワーク104が含まれ、複数

10

20

30

40

50

のコンピュータとコンピューティング・デバイスがこのネットワークと接続されている。ネットワーク 104 の例として、LAN、インターネットのようなグローバルネットワーク、あるいは他のコンピュータ・ネットワークが挙げられる。

【0028】

ネットワーク 104 と接続されたコンピュータとコンピューティング・デバイス(ネットワークの“メンバー”)の中には、クライアント側コンピュータ 106、サーバーコンピュータ 108、個人用情報機器(PDA) 110、デジタルテレビ(DTV) 112 およびその他の有線または無線コンピュータとコンピューティング・デバイスなどが含まれる。ネットワーク 104 のメンバーによって用いられるプロセッサは、同じ共通のコンピューティング・モジュールから構成される。またこれらのプロセッサは、好適には、ISA がすべて同じで、好適には同じ命令セットに従って処理を実行する。個々のプロセッサ内に含まれるモジュールの数は、そのプロセッサが必要とする処理パワーによって決められる。

10

【0029】

例えば、システム 101 のサーバー 108 は、クライアント 106 より多いデータ処理およびアプリケーション処理を実行するので、クライアント 106 より多いコンピューティング・モジュールを含むことになる。一方、PDA 110 では最低量の処理しか実行されない。したがって、PDA 110 には最少の数のコンピューティング・モジュールしか含まれない。DTV 112 はクライアント 106 とサーバー 108 の間の処理レベルを実行する。したがって、DTV 112 にはクライアント 106 とサーバー 108 の間のいくつかのコンピューティング・モジュールが含まれる。以下に説明するように、各コンピューティング・モジュールの中には、処理用コントローラと、ネットワーク 104 を介して伝送されるデータおよびアプリケーションの並列処理を実行する複数の同一処理ユニットが含まれる。

20

【0030】

システム 101 がこのように均質な構成を有することから、アダプタビリティ、処理速度および処理効率が改善される。システム 101 の各メンバーが、同じコンピューティング・モジュールのうち 1 つまたはそれ以上(またはコンピューティング・モジュールの一部)を用いて処理を実行するので、データとアプリケーションの実際の処理をどのコンピュータまたはコンピューティング・デバイスで実行するかは重要ではなくなる。さらに、個々のアプリケーションとデータの処理は、ネットワークのメンバーの間で分担することができる。システム全体を通じて、システム 101 が処理したデータとアプリケーションを含むセルを一意的に識別することにより、この処理がどこで行われたかにかかわらず、処理を要求したコンピュータまたはコンピューティング・デバイスへその処理結果を伝送することが可能となる。この処理を実行するモジュールが共通の構造と共通の ISA とを有するので、プロセッサ間の互換性を達成するためのソフトウェアの追加層の計算上の負担が回避される。このアーキテクチャとプログラミング・モデルによって、リアルタイムのマルチメディア・アプリケーションなどの実行に必要な処理速度が改善される。

30

【0031】

システム 101 によって改善される処理速度と効率というさらなる利点を利用するために、このシステムによって処理されるデータとアプリケーションとは、一意的に識別される、それぞれフォーマットが同じであるソフトウェア・セル 102 へとパッケージ化される。各ソフトウェア・セル 102 は、アプリケーションとデータの双方を含むあるいは含み得る。また各ソフトウェア・セルには、ネットワーク 104 とシステム 101 全体の中でセルを識別するためのセル識別子が含まれ、その一例としては、ソフトウェア・セルをグローバルに識別する ID が含まれる。ソフトウェア・セルのこの構造的均一性と、ネットワークの中でのソフトウェア・セルの一意的識別とによって、ネットワークの任意のコンピュータまたはコンピューティング・デバイスでのアプリケーションとデータの処理が改善される。例えば、クライアント 106 は、ソフトウェア・セル 102 の作成を行うこともできるが、クライアント 106 側の処理能力は限られていることから、このソフトウェア・セルをサーバー 108 へ伝送して処理してもらうこともできる。したがって、ソフト

40

50

ウェア・セルは、ネットワーク104全体を移動してネットワーク上での処理用リソースの可用性に基づく処理を行うことが可能となる。

【0032】

また、システム101のプロセッサとソフトウェア・セルが均質な構造を有することで、今日の異質なネットワークの混在という問題の多くを防ぐことができる。例えば任意の命令セットを用いる任意のどのISA上でもアプリケーションの処理を許容しようとする非効率的なプログラミング・モデル(Javaのパーチャル・マシンのような仮想マシンなど)が回避される。したがって、システム101は、今日のネットワークよりもはるかに効率的、かつ、はるかに効果的に広帯域処理の実現が可能となる。

【0033】

ネットワーク104のすべてのメンバーのための基本となる処理用モジュールはプロセッサ・エレメント(PE)である。図2にPEの構造が例示されている。この図に示すように、PE201は、処理ユニット(PU)203、DMAC205、複数の付加処理ユニット(APU)、すなわち、APU207、APU209、APU211、APU213、APU215、APU217、APU219、APU221を具備する。ローカルPEバス223は、APUと、DMAC205と、PU203との間でデータとアプリケーションとを伝送する。ローカルPEバス223は、従来型のアーキテクチャなどを備えていてもよいし、あるいは、パケット交換式ネットワークとして実現されてもよい。パケット交換式ネットワークとして実現される場合、より多くのハードウェアが必要となり、その一方で、利用可能な帯域幅が増加する。

【0034】

PE201は、デジタル論理回路を実現する様々な方法を用いて構成可能である。しかし、PE201は、好適には、シリコン基板上の単一の集積回路として構成されることが望ましい。基板用代替材料の中には、ガリウム砒素、ガリウム・アルミニウム砒素、砒素および多種多様のドーパントを用いるその他のいわゆるIII-B化合物が含まれる。またPE201は、超伝導材料(高速単一磁束量子(RSFQ)論理処理など)を用いて実現することもできる。

【0035】

PE201は、高帯域メモリ接続部227を介してダイナミック・ランダム・アクセス・メモリ(DRAM)225と密接に関連する。DRAM225はPE201用メイン・メモリとして機能する。DRAM225は好適には、ダイナミック・ランダム・アクセス・メモリであることが望ましいとはいえ、他の手段、例えばスタティック・ランダム・アクセス・メモリ(SRAM)として、磁気ランダム・アクセス・メモリ(MRAM)、光メモリまたはホログラフィ・メモリなどを用いてDRAM225を実現することもできる。DMAC205によって、DRAM225と、PE201のAPUとPUとの間のデータ転送が改善される。以下さらに説明するように、DMAC205によって、各APUに対するDRAM225内の排他的領域が指定されるが、この排他的領域の中へはAPUだけしかデータの書き込みができず、また、APUだけしかこの排他的領域からのデータ読み出しを行うことができない。この排他的領域は“サンドボックス”と呼ばれる。

【0036】

PU203は、データとアプリケーションのスタンド・アローン型処理が可能な標準的プロセッサなどであってもよい。作動時に、PUは、APUによって、データとアプリケーションの処理のスケジュール管理と全般的管理とを行う。APUは好適には、単一命令、複数データ(SIMD)プロセッサであることが望ましい。PU203の制御によって、APUは、並列的かつ独立にこれらのデータとアプリケーションの処理を実行する。DMAC205は、共用DRAM225に格納されているデータとアプリケーションへのPU203とAPUによるアクセス制御を行う。PE201は、好適には8個のAPUを含むことが望ましいとはいえ、必要とする処理パワーに応じて、PE内でこの数より多少上下する個数のAPUを用いてもよい。また、PE201のようないくつかのPEを結合(まとめてパッケージ化)して処理パワーの改善を図ることもできる。

10

20

30

40

50

【 0 0 3 7 】

例えば、図 3 に示すように、1 以上のチップ・パッケージなどの中に 4 つの P E をパッケージ化(まとめて結合)してネットワーク 1 0 4 のメンバー用の単一プロセッサを形成してもよい。この構成は広帯域エンジン(B E)と呼ばれる。図 3 に示すように、B E 3 0 1 には 4 つの P E (P E 3 0 3、P E 3 0 5、P E 3 0 7、P E 3 0 9)が含まれる。これらの P E 間の通信は B E バス 3 1 1 を介して行われる。広帯域メモリ接続部 3 1 3 によって共用 D R A M 3 1 5 とこれらの P E 間の通信が行われる。B E バス 3 1 1 の代わりに、B E 3 0 1 の P E 間の通信は、D R A M 3 1 5 とこのメモリ接続部とを介して行うことができる。

【 0 0 3 8 】

入力/出力(I/O)インターフェース 3 1 7 と外部バス 3 1 9 とは、広帯域エンジン 3 0 1 とネットワーク 1 0 4 のその他のメンバー間で通信を行う。B E 3 0 1 の各 P E は、P E の A P U によって行われるアプリケーションとデータの並列的かつ独立した処理と同様の並列的かつ独立した方法で、データとアプリケーションの処理を実行する。

【 0 0 3 9 】

図 4 は A P U の構造を例示する図である。A P U 4 0 2 には、ローカル・メモリ 4 0 6、レジスタ 4 1 0、4 つの浮動小数点演算ユニット 4 1 2 および 4 つの整数演算ユニット 4 1 4 が含まれる。しかし、ここでもまた、必要とする処理パワーに応じて、4 個より多少上下する個数の浮動小数点演算ユニット 4 1 2 と整数演算ユニット 4 1 4 を用いてもよい。1 つの好ましい実施形態では、ローカル・メモリ 4 0 6 には 1 2 8 キロバイトの記憶容量が含まれ、レジスタ 4 1 0 の容量は 1 2 8 × 1 2 8 ビットである。浮動小数点演算ユニット 4 1 2 は、毎秒 3 2 0 億浮動小数点演算(3 2 G L O P S)の速度で好適に作動し、整数演算ユニット 4 1 4 は、毎秒 3 2 0 億回の演算速度(3 2 G O P)で好適に作動する。

【 0 0 4 0 】

ローカル・メモリ 4 0 6 はキャッシュ・メモリではない。ローカル・メモリ 4 0 6 は、好適には S R A M として構成されることが望ましい。A P U に対するキャッシュ・コヒーレンシー、つまりキャッシュの整合性のサポートは不要である。P U では、当該 P U で開始されるダイレクト・メモリー・アクセス(D M A)をサポートするためにキャッシュの整合性が要求される場合もある。しかし、A P U によって開始される D M A に対する、あるいは、外部装置からのおよび外部装置へのアクセスに対するキャッシュの整合性のサポートは不要である。

【 0 0 4 1 】

A P U 4 0 2 にはさらに、A P U へおよび A P U からアプリケーションとデータとを伝送するためのバス 4 0 4 が含まれる。1 つの好ましい実施形態ではこのバスは 1 0 2 4 ビットの幅を持つ。A P U 4 0 2 にはさらに内部バス 4 0 8、4 2 0、4 1 8 が含まれる。1 つの好ましい実施形態では、バス 4 0 8 は 2 5 6 ビットの幅を持ち、ローカル・メモリ 4 0 6 とレジスタ 4 1 0 間で通信を行う。バス 4 2 0 と 4 1 8 とは、それぞれ、レジスタ 4 1 0 と浮動小数点演算ユニット 4 1 2 との間、および、レジスタ 4 1 0 と整数演算ユニット 4 1 4 間で通信を行う。ある好ましい実施形態では、レジスタ 4 1 0 から浮動小数点演算ユニット 4 1 2 または整数演算ユニット 4 1 4 へのバス 4 1 8 と 4 2 0 の幅は、3 8 4 ビットであり、浮動小数点演算ユニット 4 1 2 または整数演算ユニット 4 1 4 からレジスタ 4 1 0 へのバス 4 1 8 と 4 2 0 の幅は 1 2 8 ビットである。浮動小数点演算ユニット 4 1 2 または整数演算ユニット 4 1 4 からレジスタ 4 1 0 への幅より広い、レジスタ 4 1 0 から浮動小数点演算ユニットまたは整数演算ユニットへの上記バスの広い幅によって、レジスタ 4 1 0 からのより広いデータ・フローが処理中に許容される。最大 3 ワードが各計算には必要となる。しかし、各計算の結果は、一般に、ただ 1 ワードだけである。

【 0 0 4 2 】

図 5 ~ 1 0 は、ネットワーク 1 0 4 のメンバーのプロセッサのモジュラー構造をさらに例示する図である。例えば、図 5 に示すように、1 つのプロセッサには単一の P E 5 0 2 を含むことができる。上述のように、この P E には、一般に、P U、D M A C および 8 個の

10

20

30

40

50

A P Uが含まれる。各A P Uにはローカル・ストレージ(L S)が含まれる。一方、プロセッサは、ビジュアライザ(V S)5 0 5の構造を有する場合もある。図5に示すように、V S 5 0 5はP U 5 1 2、D M A C 5 1 4および4つのA P U(A P U 5 1 6、A P U 5 1 8、A P U 5 2 0、A P U 5 2 2)を有する。P Eのその他の4つのA P Uによって通常占有されるチップ・パッケージ内のスペースは、この場合、ピクセル・エンジン5 0 8、画像用キャッシュ5 1 0およびブラウン管コントローラ(C R T C)5 0 4によって占有される。P E 5 0 2またはV S 5 0 5に求められる通信速度に応じて、チップ・パッケージの中に光インターフェース5 0 6が含まれる場合もある。

【0043】

この標準化されたモジュラー構造を用いて、多数の他のプロセッサの変更例を容易にかつ効率的に構成することが可能となる。例えば、図6に示すプロセッサは、2つのチップ・パッケージ(B Eを備えるチップ・パッケージ6 0 2と、4つのV Sを含むチップ・パッケージ6 0 4)を有する。入出力部(I / O)6 0 6によって、チップ・パッケージ6 0 2のB Eとネットワーク1 0 4との間にインターフェースが設けられる。バス6 0 8はチップ・パッケージ6 0 2とチップ・パッケージ6 0 4との間の通信を行う。入出力プロセッサ(I O P)6 1 0によってデータ・フローが制御され、I / O 6 0 6へのまたはI / O 6 0 6からの入出力が行われる。I / O 6 0 6はA S I C (Application Specific Integrated Circuit)として製造が可能である。V Sからの出力はビデオ信号6 1 2である。

10

【0044】

図7は、ネットワーク1 0 4のその他のメンバーへ超高速通信を行う2つの光インターフェース7 0 4と7 0 6とを備えたB E 7 0 2用のチップ・パッケージ(またはローカルに接続された他のチップ・パッケージ)を例示する。B E 7 0 2は、ネットワーク1 0 4上でサーバーなどとして機能することができる。

20

【0045】

図8のチップ・パッケージは、2つのP E 8 0 2及び8 0 4および2つのV S 8 0 6及び8 0 8を有する。I / O 8 1 0は、チップ・パッケージとネットワーク1 0 4との間にインターフェースを与える。チップ・パッケージからの出力はビデオ信号である。この構成は画像処理用ワークステーションなどとして機能することができる。

【0046】

図9はさらに別の構成を例示する。この構成は、図8に例示されている構成の処理パワーの1 / 2を含む。2つのP Eの代わりに1つのP E 9 0 2が設けられ、2つのV Sの代わりに1つのV S 9 0 4が設けられる。I / O 9 0 6は、図8に例示されているI / Oの帯域幅の1 / 2の帯域幅を有する。またこのようなプロセッサは、画像処理用ワークステーションとして機能することができる。

30

【0047】

最後の構成が図10に図示されている。このプロセッサは、単一のV S 1 0 0 2とI / O 1 0 0 4だけから構成される。この構成はP D Aなどとして機能することができる。

【0048】

図11は、ネットワーク1 0 4のプロセッサのチップ・パッケージの中への光インターフェースの統合を例示する図である。これらの光インターフェースによって、光信号は電気信号に変換され、電気信号は光信号に変換される。また、これらの光インターフェースは、ガリウム砒素、アルミニウム・ガリウム砒素、ゲルマニウムその他の元素や化合物などを含む様々な材料から構成することができる。この図に示すように、光インターフェース1 1 0 4と1 1 0 6とはB E 1 1 0 2のチップ・パッケージの上に組み立てられる。B Eバス1 1 0 8はB E 1 1 0 2のP E、すなわち、P E 1 1 1 0、P E 1 1 1 2、P E 1 1 1 4、P E 1 1 1 6およびこれらの光インターフェースとの間での通信を行う。光インターフェース1 1 0 4には2つのポート(ポート1 1 1 8とポート1 1 2 0)が含まれ、また光インターフェース1 1 0 6には2つのポート(ポート1 1 2 2とポート1 1 2 4)が含まれる。ポート1 1 1 8、1 1 2 0、1 1 2 2、1 1 2 4は、光導波路1 1 2 6、1 1 2 8、1 1 3 0、1 1 3 2とそれぞれ接続される。光信号は、光インターフェース1 1 0 4と

40

50

1106のポートを介して、これらの光導波路の中を通り、BE1102へおよびBE1102から伝送される。

【0049】

このような光導波路と各BEの4つの光ポートとを用いて様々な構成において複数のBEをまとめて接続してもよい。例えば、図12に示すように、このような光ポートを介して2つまたはそれ以上のBE(BE1152、BE1154、BE1156など)を直列に接続することができる。この例では、BE1152の光インターフェース1166は、その光ポートを介しBE1154の光インターフェース1160の光ポートと接続される。同様に、BE1154の光インターフェース1162の光ポートは、BE1156の光インターフェース1164の光ポートと接続される。

10

【0050】

図13にマトリクス構成が例示される。この構成では、各BEの光インターフェースは2つの他のBEと接続される。この図に示すように、BE1172の光インターフェース1188の光ポートの中の1つが、BE1176の光インターフェース1182の光ポートと接続される。光インターフェース1188のもう一方の光ポートは、BE1178の光インターフェース1184の光ポートと接続される。同様に、BE1174の光インターフェース1190の1つの光ポートはBE1178の光インターフェース1184のもう一方の光ポートと接続される。光インターフェース1190のもう一方の光ポートは、BE1180の光インターフェース1186の光ポートと接続される。このマトリクス構成は他のBEに対しても同様に拡張することができる。

20

【0051】

シリアル構成かマトリクス構成のいずれかを用いて、任意の所望のサイズとパワーから成るネットワーク104用プロセッサの構成が可能となる。言うまでもなく、BEの光インターフェースに対して、または、BEよりPE数の少ないプロセッサに対して追加ポートを加えて、他の構成を形成してもよい。

【0052】

図14はBEのDRAMに対する制御システムと構造を例示する図である。同様の制御システムと構造が、別のサイズを持ち、多少異なる数のPEを含むプロセッサの中で用いられる。この図に示すように、クロスバ交換機によって、BE1201を備える4つのPEからなる各DMAC1210が8つのバンク・コントロール1206と接続される。各バンク・コントロール1206によって、DRAM1204の8つのバンク1208(4つだけしか図示されていない)が制御される。したがって、DRAM1204は、合計64のバンクを具備することになる。好ましい実施形態では、DRAM1204は64メガバイトの容量を持ち、各バンクは1メガバイトの容量を持っている。各バンク内の最小のアドレス指定可能単位は、この好ましい実施形態では1024ビットのブロックである。

30

【0053】

BE1201にはスイッチ・ユニット1212も含まれる。スイッチ・ユニット1212によって、BE1201と密接に接続されているBEの他のAPUのDRAM1204へのアクセスが可能となる。したがって、第2のBEを第1のBEと密接に接続することが可能となり、さらに、各BEの各APUは、APUが通常アクセス可能なメモリ・ロケーションの数の2倍のアドレス指定を行うことが可能となる。スイッチ・ユニット1212のようなスイッチ・ユニットを介して、第1のBEのDRAMから第2のBEのDRAMへのデータの直接読み出し、または、第2のBEのDRAMから第1のBEのDRAMへのデータの直接書き込みを行うことが可能となる。

40

【0054】

例えば、図15に示すように、このような書き込みを行うために、第1のBEのAPU(BE1222のAPU1220など)によって、第2のBEのDRAM(通常の場合のようなBE1222のDRAM1224ではなく、BE1226のDRAM1228など)のメモリ・ロケーションへの書き込みコマンドが出される。BE1222のDMAC1230は、クロスバ交換機1221を介して、バンク・コントロール1234へ書き込みコマ

50

ンドを送り、バンク・コントロール1234は、バンク・コントロール1234と接続された外部ポート1232へコマンドを伝送する。BE1226のDMAC1238は書き込みコマンドを受け取り、BE1226のスイッチ・ユニット1240へこのコマンドを転送する。スイッチ・ユニット1240は書き込みコマンドの中に含まれるDRAMアドレスを識別し、BE1226のバンク・コントロール1242を介して、DRAM1228のバンク1244へ、DRAMアドレス内に格納するデータを送る。したがって、スイッチ・ユニット1240によって、DRAM1224とDRAM1228の双方は、BE1222のAPU用の単一メモリ空間として機能することが可能になる。

【0055】

図16はDRAMの64個のバンク構成を図示する。これらのバンクは、8つの行(1302、1304、1306、1308、1310、1312、1314、1316)と8つの列(1320、1322、1324、1326、1328、1330、1332、1334)とで構成されている。各行は1つのバンク・コントローラによって制御される。したがって、各バンク・コントローラは8メガバイトのメモリを制御する。

10

【0056】

図17と18は、最小のアドレス指定可能な格納単位(1024ビットのブロックなど)でのDRAMの格納とアクセスを行うための異なる構成を例示する。図17で、DMAC1402は単一のバンク1404の中に8つの1024ビット・ブロック1406を格納する。図18では、DMAC1412によって、1024ビットを含むデータ・ブロックの読み出しと書き込みが行われるものの、これらのブロックは、2つのバンク(バンク1414とバンク1416)の間で分配される。したがって、これらのバンクの各々には16個のデータ・ブロックが含まれ、データの各ブロックには512ビットが含まれる。この分配によって、DRAMのアクセスをさらに高速なものに改善することが可能となり、ある種のアプリケーションの処理に役立つ。

20

【0057】

図19はPE内のDMAC1506のアーキテクチャを例示する。この図に例示されているように、各APU1502がDMAC1506の構造上のノード1504へ直接アクセスを行うように、DMAC1506を含む構造上のハードウェアは全てのPEを通じて配設される。各ノードは、ノードが直接アクセスを行う対象のAPUによるメモリ・アクセスに適した論理処理を実行する。

30

【0058】

図20はDMACの他の実施形態、すなわち、非分配型アーキテクチャを図示する。この場合、DMAC1606の構造上のハードウェアは集中型である。APU1602とPU1604は、ローカルPEバス1607を介してDMAC1606を用いて通信を行う。DMAC1606はクロスバー・スイッチを介してバス1608と接続される。バス1608はDRAM1610と接続されている。

【0059】

上述のように1つのPEの複数のAPUのすべては、独立に、共用DRAM内のデータへのアクセスが可能である。その結果、第1のAPUがあるデータとそのローカル・ストレージで処理しているときに、第2のAPUがこれらのデータを要求する場合もある。その時点で共用DRAMから第2のAPUへ当該データが出力された場合、データの値を変化させ得る第1のAPUの進行中の処理に起因して、そのデータが無効になる場合がある。したがって、その時点で第2のプロセッサが共用DRAMからデータを受け取った場合、第2のプロセッサでエラー結果が生じるおそれがある。例えば、このようなデータとしては、グローバル変数用の具体的な値が挙げられる。第1のプロセッサがその処理中その値を変えた場合、第2のプロセッサはもう使用されていない値を受け取ることになる。したがって、共用DRAMの範囲内でメモリ・ロケーションからのおよびメモリ・ロケーションへのAPUによるデータの読み出しと書き込みを同期させる何らかの方式が必要となる。この方式では、別のAPUがそのローカル・ストレージで現在働きかけている対象データであって、したがって最新のものではないデータのメモリ・ロケーションからの読み出

40

50

しと、最新のデータを格納するメモリ・ロケーションの中へのデータの書き込みと、を行わないようにする必要がある。

【 0 0 6 0 】

これらの問題を解決するために、D R A Mの各アドレス指定が可能なメモリ・ロケーションに対して、そのメモリ・ロケーションの中に格納されているデータに関連する状態情報を格納するために、D R A Mの中でメモリの追加セグメントの割り振りが行われる。この状態情報の中には、フル/エンプティ(F / E)ビットと、メモリ・ロケーションからデータを要求するA P Uの識別子(A P U I D)と、要求されたデータを読み出す読み出し先となるA P Uのローカル・ストレージのアドレス(L Sアドレス)とが含まれる。D R A Mのアドレス指定が可能なメモリ・ロケーションは任意のサイズとすることができる。ある好ましい実施形態ではこのサイズは1 0 2 4ビットである。

10

【 0 0 6 1 】

F / Eビットの1への設定は、メモリ・ロケーションに格納されているデータが最新のものであることを示す。一方、F / Eビットの0への設定は、関連するメモリ・ロケーションに格納されたデータが最新のものではないことを示す。このビットが0に設定されているとき、A P Uがそのデータを要求しても、A P Uによってそのデータの即時読み出しは妨げられる。この場合、そのデータを要求しているA P Uを識別するA P U I Dと、データが最新のものになっているとき、そのデータを読み出す読み出し先となるこのA P Uのローカル・ストレージ内のメモリ・ロケーションを識別するL Sアドレスとが、追加メモリ・セグメントの中へ入力される。

20

【 0 0 6 2 】

また追加メモリ・セグメントは、A P Uのローカル・ストレージ内の各メモリ・ロケーションに対して割り振られる。この追加メモリ・セグメントは、“ ビジー・ビット ” と呼ばれる1ビットを格納する。このビジー・ビットは、D R A Mから検索される固有データの格納用として関連するL Sメモリ・ロケーションの予約を行うために使用される。ローカル・ストレージ内の特定のメモリ・ロケーションに対してビジー・ビットが1に設定されている場合、これらの固有データの書き込み用としてのみA P Uはこのメモリ・ロケーションを使用することができる。一方、ビジー・ビットが、ローカル・ストレージ内の特定のメモリ・ロケーションに対して0に設定されている場合、A P Uは、任意のデータの書き込み用としてこのメモリ・ロケーションを使用することができる。

30

【 0 0 6 3 】

F / Eビット、A P U I D、L Sアドレスおよびビジー・ビットが、P Eの共用D R A Mからの、および、P Eの共用D R A Mへのデータの読み出しと書き込みを同期させるために使用される方法を示す例が図 2 1 - 3 5 に例示されている。

【 0 0 6 4 】

図 2 1 に示すように、1以上のP E (P E 1 7 2 0 など) が D R A M 1 7 0 2 を使用する。P E 1 7 2 0 にはA P U 1 7 2 2 とA P U 1 7 4 0 とが含まれる。A P U 1 7 2 2 には制御論理回路 1 7 2 4 が含まれ、A P U 1 7 4 0 には制御論理回路 1 7 4 2 が含まれる。A P U 1 7 2 2 にはローカル・ストレージ 1 7 2 6 も含まれる。このローカル・ストレージには複数のアドレス可能なメモリ・ロケーション 1 7 2 8 が含まれる。A P U 1 7 4 0 にはローカル・ストレージ 1 7 4 4 が含まれ、このローカル・ストレージにも複数のアドレス可能なメモリ・ロケーション 1 7 4 6 が含まれる。これらのアドレス可能なメモリ・ロケーションのすべては好適にはサイズが1 0 2 4ビットであることが望ましい。

40

【 0 0 6 5 】

メモリの追加セグメントは各L Sのアドレス可能なメモリ・ロケーションと関連付けられる。例えば、メモリ・セグメント 1 7 2 9 と 1 7 3 4 とはそれぞれ、ローカルなメモリ・ロケーション 1 7 3 1 と 1 7 3 2 とに関連付けられ、メモリ・セグメント 1 7 5 2 はローカルなメモリ・ロケーション 1 7 5 0 と関連付けられる。上述のような“ ビジー・ビット ” はこれらの追加メモリ・セグメントの各々の中に格納される。ローカルなメモリ・ロケーション 1 7 3 2 は、このメモリ・ロケーションがデータを含むことを示すいくつかの

50

×印を用いて示されている。

【0066】

DRAM 1702には、メモリ・ロケーション1706と1708とを含む複数のアドレス可能なメモリ・ロケーション1704が含まれる。これらのメモリ・ロケーションは、好適にはサイズが1024ビットであることが望ましい。メモリの追加セグメントはまたこれらのメモリ・ロケーションの各々とも関連付けられる。例えば、追加メモリ・セグメント1760はメモリ・ロケーション1706と関連し、追加メモリ・セグメント1762はメモリ・ロケーション1708と関連付けられる。各メモリ・ロケーションに格納されるデータに関連する状態情報は、メモリ・ロケーションと関連付けられたメモリ・セグメントに格納される。この状態情報の中には、上述のように、F/Eビット、APU IDおよびLSアドレスが含まれる。例えば、メモリ・ロケーション1708については、この状態情報にはF/Eビット1712、APU ID 1714およびLSアドレス1716が含まれる。

10

【0067】

この状態情報とビジー・ビットとを用いて、PEのAPU、または1グループのPE間での、共用DRAMからの、および、同期した共用DRAMからの読み出しと、同期した共用DRAMへのデータの書き込みを行うことができる。

【0068】

図22はAPU 1722のLSメモリ・ロケーション1732から、DRAM 1702のメモリ・ロケーション1708へのデータの同期書き込みの開始を例示する図である。APU 1722の制御論理回路1724によってこれらのデータの同期書き込みが開始される。メモリ・ロケーション1708がエンプティであるため、F/Eビット1712は0に設定される。その結果、メモリ・ロケーション1708の中へLSメモリ・ロケーション1732内のデータを書き込むことが可能となる。一方、このビットが1に設定されていて、メモリ・ロケーション1708がフル状態であり、最新の有効データを含むことが示されている場合、制御回路1722はエラー・メッセージを受け取ることになり、このメモリ・ロケーションへのデータの書き込みは禁止される。

20

【0069】

メモリ・ロケーション1708への成功したデータの同期書き込みの結果が図23に示されている。この書き込まれたデータはメモリ・ロケーション1708の中に格納され、F/Eビット1712は1に設定される。この設定によって、メモリ・ロケーション1708がフル状態であること、および、このメモリ・ロケーションの中のデータが最新の有効データであることが示される。

30

【0070】

図24は、DRAM 1702のメモリ・ロケーション1708からローカル・ストレージ1744のLSメモリ・ロケーション1750へのデータの同期読み出しの開始を例示する図である。この読み出しを開始するために、LSメモリ・ロケーション1750のメモリ・セグメント1752の中のビジー・ビットが1に設定されて、上記データ用としてこのメモリ・ロケーションが予約される。このビジー・ビットを1に設定することによって、APU 1740がこのメモリ・ロケーションに他のデータを格納することはなくなっている。

40

【0071】

図25に示すように、制御論理回路1742は次にDRAM 1702のメモリ・ロケーション1708に対して同期読取りコマンドを出す。このメモリ・ロケーションと関連付けられるF/Eビット1712は1に設定されているので、メモリ・ロケーション1708の中に格納されたデータは最新の、有効データであると見なされる。その結果、メモリ・ロケーション1708からLSメモリ・ロケーション1750へのデータ転送の準備の際に、F/Eビット1712は0に設定される。この設定が図26に示されている。このビットを0に設定されているということは、これらのデータの読み出しの後に、メモリ・ロケーション1708のデータは無効になることを示す。

50

【 0 0 7 2 】

図 2 7 に示すように、メモリ・ロケーション 1 7 0 8 内のデータは、次に、メモリ・ロケーション 1 7 0 8 から L S メモリ・ロケーション 1 7 5 0 へ読み出される。図 2 8 は最終状態を示す図である。メモリ・ロケーション 1 7 0 8 のデータのコピーは L S メモリ・ロケーション 1 7 5 0 に格納される。F / E ビット 1 7 1 2 は 0 に設定され、メモリ・ロケーション 1 7 0 8 のデータが無効であることが示される。この無効は、A P U 1 7 4 0 によって行われた上記データの変更の結果である。メモリ・セグメント 1 7 5 2 内のビジー・ビットもまた 0 に設定される。この設定によって、A P U 1 7 4 0 が L S メモリ・ロケーション 1 7 5 0 を任意の目的に利用できること、すなわち、この L S メモリ・ロケーションがもはや固有データの受信を待機している予約状態ではないことが示される。したがって、任意の目的のために A P U 1 7 4 0 による L S メモリ・ロケーション 1 7 5 0 へのアクセスが可能となる。

10

【 0 0 7 3 】

図 2 9 ~ 図 3 5 には、D R A M 1 7 0 2 のメモリ・ロケーション用の F / E ビットが、0 に設定されていて、このメモリ・ロケーションのデータが最新のものでもなく有効なものでもないことが示されている場合の、D R A M 1 7 0 2 (メモリ・ロケーション 1 7 0 8 など)のメモリ・ロケーションから、A P U のローカル・ストレージ(ローカル・ストレージ 1 7 4 4 の L S メモリ・ロケーション 1 7 5 2 など)の L S メモリ・ロケーションへのデータの同期読み出しが例示されている。図 2 9 に示すように、この転送を開始するために、L S メモリ・ロケーション 1 7 5 0 のメモリ・セグメント 1 7 5 2 内のビジー・ビットは 1 に設定され、このデータ転送用としてこの L S メモリ・ロケーションが予約される。図 3 0 に示すように、制御論理回路 1 7 4 2 は、次に、D R A M 1 7 0 2 のメモリ・ロケーション 1 7 0 8 に対して同期読取りコマンドを出す。このメモリ・ロケーションと関連付けられた F / E ビット(F / E ビット 1 7 1 2)は 0 に設定されているので、メモリ・ロケーション 1 7 0 8 に格納されているデータは無効である。その結果、信号は制御論理回路 1 7 4 2 へ伝送され、このメモリ・ロケーションからのデータの即時読み出しが阻止される。

20

【 0 0 7 4 】

図 3 1 に示すように、A P U I D 1 7 1 4 とこの読取りコマンド用の L S アドレス 1 7 1 6 とはメモリ・セグメント 1 7 6 2 の中へ書き込まれる。この場合、A P U 1 7 4 0 用の A P U I D と、L S メモリ・ロケーション 1 7 5 0 用の L S メモリ・ロケーションとはメモリ・セグメント 1 7 6 2 の中へ書き込まれる。したがって、メモリ・ロケーション 1 7 0 8 の範囲内のデータが最新のものになっているとき、この A P U I D と L S メモリ・ロケーションは、最新のデータを伝送する伝送先のメモリ・ロケーションを決定するために使用される。

30

【 0 0 7 5 】

メモリ・ロケーション 1 7 0 8 内のデータは、A P U がこのメモリ・ロケーションの中へデータを書き込むと、有効で最新のデータとなる。A P U 1 7 2 2 のメモリ・ロケーション 1 7 3 2 などからメモリ・ロケーション 1 7 0 8 の中へのデータの同期書き込みが図 2 9 に例示されている。このメモリ・ロケーション用の F / E ビット 1 7 1 2 が 0 に設定されているため、これらのデータのこの同期書き込みは許される。

40

【 0 0 7 6 】

図 3 3 に示すように、この書き込み後、メモリ・ロケーション 1 7 0 8 の中のデータは最新の有効データになる。したがって、メモリ・セグメント 1 7 6 2 から得られる A P U I D 1 7 1 4 と L S アドレス 1 7 1 6 とは、メモリ・セグメント 1 7 6 2 から即座に読み出され、次いでこの情報はこのセグメントから削除される。メモリ・ロケーション 1 7 0 8 の中のデータの即時読み出しを予期して、F / E ビット 1 7 1 2 もまた 0 に設定される。図 3 4 に示すように、A P U I D 1 7 1 4 と L S アドレス 1 7 1 6 とを読み出すと、A P U 1 7 4 0 の L S メモリ・ロケーション 1 7 5 0 へメモリ・ロケーション 1 7 0 8 内の有効データを読み出すためにこの情報は直ちに使用される。最終状態が図 3 5 に図示され

50

ている。この図は、メモリ・ロケーション1708からメモリ・ロケーション1750へコピーされた有効データと、0に設定されたメモリ・セグメント1752内のビジー・ビットと、0に設定されたメモリ・セグメント1762内のF/Eビット1712とを図示する。このビジー・ビットの0への設定によって、任意の目的のためにAPU1740がLSメモリ・ロケーション1750のアクセスを行うことが可能になる。このF/Eビットの0への設定によって、メモリ・ロケーション1708内のデータがもはや最新のものでもなく、有効なものでもないことが示される。

【0077】

図36は、上述のオペレーションと、DRAMのメモリ・ロケーションの様々な状態とを要約する図であり、この状態は、F/Eビットの状態と、APUIDと、メモリ・ロケーションに対応するメモリ・セグメントの中に格納されたLSアドレスとに基づく。このメモリ・ロケーションは、3つの状態を持つことが可能である。これらの3つの状態として、F/Eビットが0に設定され、APUIDまたはLSアドレスに対して情報が提供されないエンpty状態1880と、F/Eビットが1に設定され、APUIDまたはLSアドレスに対して情報が提供されないフル状態1882と、F/Eビットが0に設定され、APUIDとLSアドレスに対して情報が提供されるブロッキング状態1884とがある。

【0078】

この図に示すように、エンpty状態1880では、同期書き込みオペレーションが許され、フル状態1882への遷移という結果が得られる。しかし、メモリ・ロケーションがエンpty状態であるときはメモリ・ロケーション内のデータが最新のものではないので、同期読み出しオペレーションに対しては、ブロッキング状態1884へ遷移するという結果となる。

【0079】

フル状態1882では、同期読み出しオペレーションが許され、エンpty状態1880への遷移という結果が得られる。一方、有効データの上書きを避けるために、フル状態1882の同期書き込みオペレーションは禁止される。このような書き込みオペレーションがこの状態で試みられる場合、状態の変化は生じず、エラー・メッセージがAPUの対応する制御論理回路へ伝送される。

【0080】

ブロッキング状態1884では、メモリ・ロケーションの中へのデータの同期書き込みが許され、エンpty状態1880への遷移という結果が得られる。一方、ブロッキング状態1884での同期読み出しオペレーションは禁止される。このブロッキング状態を生じさせることとなった前同期読み出しオペレーションとのコンフリクトを阻止するためである。同期読み出しオペレーションが、ブロッキング状態1884で試みられた場合、状態変化は生じないでAPUの対応する制御論理回路へエラー・メッセージが伝送される。

【0081】

共用DRAMからのデータの同期読み出しと、共用DRAMへのデータの同期書き込みを行う上述の方式は、外部装置からのデータ読み出しと外部装置へのデータ書き込み用プロセッサとして通常専用の計算用リソースを取り除くためにも利用が可能である。この入出力(I/O)機能はPUによって行うこともできる。しかし、この同期方式の変更を利用して、適切なプログラムを実行するAPUがこの機能を実行してもよい。例えば、この方式を利用して、外部装置によって開始された、I/Oインターフェースからのデータ伝送を求める割込み要求を受け取るPUは、このAPUにこの要求の処理を委任してもよい。次いで、APUはI/Oインターフェースに対して同期書き込みコマンドを出す。今度はこのインターフェースによって、現在DRAMの中へデータを書き込むことができる旨の信号が外部装置へ送られる。次にAPUはDRAMに対して同期読取りコマンドを出し、DRAMの関連するメモリ空間をブロッキング状態に設定する。APUはまた、データを受け取る必要があるAPUのローカル・ストレージのメモリ・ロケーションに対してビジー・ビットを1に設定する。ブロッキング状態では、DRAMの関連するメモリ空間と関連

10

20

30

40

50

付けられた追加メモリ・セグメントの中に、A P UのIDとA P Uのローカル・ストレージの関連するメモリ・ロケーションのアドレスが含まれる。次に外部装置は同期書き込みコマンドを出し、D R A Mの関連するメモリ空間へデータが直接書き込まれる。このメモリ空間はブロッキング状態にあるので、データは、このスペースの中から、追加メモリ・セグメントの中で識別されたA P Uのローカル・ストレージのメモリ・ロケーションの中へ直ちに読み出される。次いで、これらのメモリ・ロケーション用のビジー・ビットは0に設定される。外部装置がデータの書き込みを完了したとき、A P Uは、伝送が完了した旨を示す信号をP Uへ出す。

【0082】

したがって、この方式を用いて、P Uに対する最小の計算上の負荷で、外部装置からのデータ転送処理を行うことができる。しかし、この機能を委任されたA P UはP Uに対して割込み要求を出せることが望ましく、外部装置がD R A Mに対して直接アクセスを行うことが望ましい。

【0083】

各P EのD R A Mには複数の“サンドボックス”が含まれる。サンドボックスによって共用D R A M領域が画定され、この領域を越えて、特定のA P Uまたは1組のA P Uがデータの読み出しや書き込みを行うことはできない。これらのサンドボックスによって、1つのA P Uが処理するデータに起因する、別のA P Uによって処理されるデータの破損に対するセキュリティが与えられる。またこれらのサンドボックスによって、ソフトウェア・セルが全D R A Mの中でデータの破損を生じる可能性なく、ネットワーク104から特定のサンドボックスの中へソフトウェア・セルのダウンロードを行うことが許される。本発明では、サンドボックスは、D R A MとD M A Cとから成るハードウェアの中に設けられる。ソフトウェアの代わりに、このハードウェア内にこれらのサンドボックスを設けることにより、速度とセキュリティという利点が得られる。

【0084】

P EのP UはA P Uへ割り当てられるサンドボックスの制御を行う。P Uは、オペレーティング・システムのような信頼のおけるプログラムだけしか通常作動させないので、本方式によってセキュリティが危険にさらされることはない。本方式に従って、P Uはキー管理テーブルの構築と維持とを行う。図37にこのキー管理テーブルが例示されている。この図に示すように、キー管理テーブル1902内の各エントリには、A P U用の識別子(I D)1904と、そのA P U用のA P Uキー1906と、キー・マスク1908とが含まれる。このキー・マスクの用途について以下説明する。キー管理テーブル1902は、スタティック・ランダム・アクセス・メモリ(S R A M)のような比較的高速のメモリに好適に格納され、D M A Cと関連付けられる。キー管理テーブル1902へのエントリはP Uによって制御される。A P Uが、D R A Mの特定の格納位置(ストレージロケーション)へのデータの書き込みあるいはD R A Mの特定の格納位置からのデータの読み出しを要求すると、D M A Cは、その格納位置と関連付けられたメモリ・アクセス・キーに対して、キー管理テーブル1902内のそのA P Uへ割り当てられたA P Uキー1906の評価を行う。

【0085】

図38に示すように、D R A M 2002の各アドレス可能な格納位置2006に対して専用メモリ・セグメント2010が割り当てられる。この格納位置用のメモリ・アクセス・キー2012はこの専用メモリ・セグメントの中に格納される。上述のように、やはり各アドレス可能な格納位置2006と関連付けられたさらなる追加専用メモリ・セグメント2008によって、格納位置へのデータ書き込みと、格納位置からのデータの読み出しを行うための同期情報が格納される。

【0086】

作動時に、A P UはD M A CへD M Aコマンドを出す。このコマンドには、D R A M 2002の格納位置2006のアドレスが含まれる。このコマンドを実行する前に、D M A Cは、キー管理テーブル1902におけるA P UのID1904を用いて要求を行っている

10

20

30

40

50

A P Uのキー1906を調べる。次いで、D M A Cは、A P Uがアクセスを求める対象先であるD R A Mの格納位置と関連付けられた専用メモリ・セグメント2010内に格納されているメモリ・アクセス・キー2012と、要求を行っているA P UのA P Uキー1906との比較を行う。2つのキーが一致しない場合、D M Aコマンドは実行されない。一方、2つのキーが一致した場合、D M Aコマンドは進行し、要求されたメモリ・アクセスが実行される。

【0087】

図39に他の実施形態の一例を示す。この例では、P Uはメモリ・アクセス管理テーブル2102の維持も行う。メモリ・アクセス管理テーブル2102にはD R A M内にある各サンドボックス用のエントリが含まれる。図39の特定の例では、D R A Mには64個のサンドボックスが含まれる。メモリ・アクセス管理テーブル2102内の各エントリには、サンドボックス用識別子(I D)2104と、ベース・メモリ・アドレス2106と、サンドボックス・サイズ2108と、メモリ・アクセス・キー2110と、アクセス・キーマスク2110とが含まれる。ベース・メモリ・アドレス2106によって、D R A M内にアドレスが設けられ、このアドレスによって特定のメモリ・サンドボックスの最初の部分が示される。サンドボックス・サイズ2108によってサンドボックスのサイズが与えられ、したがって、このサイズによって特定のサンドボックスのエンドポイントが与えられる。

【0088】

図40は、キー管理テーブル1902とメモリ・アクセス管理テーブル2102とを用いてD M Aコマンドを実行するためのステップを示すフロー・チャートである。ステップ2202では、A P Uによって、サンドボックス内の特定の一定あるいは複数のメモリ・ロケーションに対するアクセス用D M AコマンドがD M A Cへ出される。このコマンドには、アクセス要求を行う対象先である特定のサンドボックスの識別を行うサンドボックスI D2104が含まれる。ステップ2204では、D M A Cは、A P UのI D1904を利用して、キー管理テーブル1902内の要求を行っているA P Uのキー1906を調べる。ステップ2206で、D M A Cは、メモリ・アクセス管理テーブル2102で、サンドボックスと関連付けられたメモリ・アクセス・キー2110を調べるコマンドで、サンドボックスI D2104を利用する。ステップ2208で、D M A Cは、要求を行っているA P Uへ割り当てられているA P Uキー1906をサンドボックスと関連付けられたアクセス・キー2110と比較する。ステップ2210で、この2つのキーが一致するかどうかの決定が行われる。この2つのキーが一致しない場合、処理はステップ2212へ移行し、そこでD M Aコマンドは先へ進まず、要求を行っているA P UとP Uのいずれかまたはその双方へエラー・メッセージが送信される。一方、ステップ2210で、2つのキーの一致が得られた場合、処理はステップ2214へ進み、そこでD M A CはD M Aコマンドを実行する。

【0089】

A P Uキー用およびメモリ・アクセス・キー用のキー・マスクによってこのシステムに大きな柔軟性が与えられる。キー用のキー・マスクによって、マスクされたビットはワイルド・カードに変換される。例えば、A P Uキー1906と関連付けられたキー・マスク1908が、キー・マスク1908内のこれらのビットを1に設定することなどにより、その最後の2ビットが“マスク”に設定されている場合、A P Uキーは1または0のいずれかになることができ、そのままメモリ・アクセス・キーに一致することになる。例えば、A P Uキーが1010であるとする。通常、このA P Uキーによって1010のアクセス・キーを持つサンドボックスへのアクセスだけが可能になる。しかし、このA P Uキー用のA P Uキー・マスクが0001に設定されている場合、このA P Uキーを用いて1010または1011のいずれかのアクセス・キーを持つサンドボックスへのアクセスを行うことが可能となる。同様に、1010または1011のいずれかのA P Uキーを持つA P Uによって、0001に設定されたマスクを持つアクセス・キー1010のアクセスを行うことが可能である。A P Uキー・マスクとメモリ・キー・マスクの双方を同時に使用す

10

20

30

40

50

ることができるので、多数のバリエーションのサンドボックスに対するA P Uによるアクセシビリティの設定が可能となる。

【 0 0 9 0 】

また本発明はシステム 1 0 1 のプロセッサ用の新しいプログラミング・モデルも提供するものである。このプログラミング・モデルではソフトウェア・セル 1 0 2 が用いられる。ネットワーク 1 0 4 上の任意のプロセッサへ処理用としてこれらのセルの伝送を行うことが可能である。またこの新しいプログラミング・モデルでは、システム 1 0 1 のユニークなモジュラー形アーキテクチャと、システム 1 0 1 のプロセッサとが利用される。

【 0 0 9 1 】

ソフトウェア・セルはA P Uのローカル・ストレージからA P Uによって直接処理される。A P Uは、D R A M内のいずれのデータまたはプログラムに対しても直接働きかけることは行わない。D R A M内のデータとプログラムは、A P Uがこれらのデータとプログラムの処理を行う前に、A P Uのローカル・ストレージの中に読み込まれる。したがって、A P Uのローカル・ストレージには、プログラム・カウンタと、スタックと、これらのプログラムを実行するための他のソフトウェア・エレメントとが含まれることになる。P Uは、D M A Cに対してD M Aコマンドを出すことによりA P Uの制御を行う。

10

【 0 0 9 2 】

ソフトウェア・セル 1 0 2 の構造が図 4 1 に例示されている。この図に示すように、ソフトウェア・セル 2 3 0 2 などのソフトウェア・セルの中には、ルート選定情報セクション 2 3 0 4 と本体部分 2 3 0 6 とが含まれる。ルート選定情報セクション 2 3 0 4 に含まれる情報は、ネットワーク 1 0 4 のプロトコルに依って決められる。ルート選定情報セクション 2 3 0 4 の中には、ヘッダ 2 3 0 8、宛先 I D 2 3 1 0、ソース I D 2 3 1 2 および応答 I D 2 3 1 4 が含まれる。宛先 I D にはネットワーク・アドレスが含まれる。T C P / I P プロトコルの下で、例えば、ネットワーク・アドレスはインターネット・プロトコル (I P) アドレスである。さらに宛先 I D 2 3 1 0 には、処理のためにセルを送送すべき伝送先の P E 及び A P U の識別子が含まれる。ソース I D 2 3 1 4 にはネットワーク・アドレスが含まれ、このソース I D によって P E と A P U とが識別され、この P E と A P U とからセルが起動し、必要な場合に、宛先 P E と A P U とがセルに関する追加情報を得ることが可能となる。応答 I D 2 3 1 4 にはネットワーク・アドレスが含まれ、この応答 I D 2 3 1 4 によって、セルに関するクエリとセルの処理の結果とを送る送り先の P E と A P U とが識別される。

20

30

【 0 0 9 3 】

セルの本体部分 2 3 0 6 にはネットワークのプロトコルとは無関係の情報が含まれる。図 4 1 の分解部分はセルの本体部分 2 3 0 6 の細部を図示する。セルの本体部分 2 3 0 6 のヘッダ 2 3 2 0 によってセル本体の開始部が識別される。セル・インターフェース 2 3 2 2 にはセルの利用に必要な情報が含まれる。この情報の中には、グローバルな一意的 I D 2 3 2 4 と、要求される A P U 2 3 2 6 と、サンドボックス・サイズ 2 3 2 8 と、前回のセルの I D 2 3 3 0 とが含まれる。

【 0 0 9 4 】

グローバルな一意的 I D 2 3 2 4 によって、ネットワーク 1 0 4 全体を通じてソフトウェア・セル 2 3 0 2 が一意的に識別される。グローバルな一意的 I D 2 3 2 4 が、ソース I D 2 3 1 2 (ソース I D 2 3 1 2 内の P E または A P U の一意的識別子など) と、ソフトウェア・セル 2 3 0 2 の作成または伝送の時刻と日付とに基づいて作成される。必要な A P U 2 3 2 6 によってセルの実行に必要な最低数の A P U が与えられる。サンドボックス・サイズ 2 3 2 8 によって、セルの実行に必要な D R A M と関連する必要な A P U 内に、保護されたメモリ量が与えられる。前回のセル I D 2 3 3 0 によって、シーケンシャルな実行を要求する 1 グループのセル (ストリーミング・データなど) 内の前回のセルの識別子が提供される。

40

【 0 0 9 5 】

実行セクション 2 3 3 2 の中にはセルのコア情報が含まれる。この情報の中には D M A コ

50

マンド・リスト 2334 と、プログラム 2336 と、データ 2338 とが含まれる。プログラム 2336 には、A P U プログラム 2360 と 2362 などの A P U によって実行されるプログラム(“アプレット”と呼ばれる)が含まれ、データ 2338 にはこれらのプログラムを用いて処理されるデータが含まれる。DMA コマンド・リスト 2334 には、プログラムの起動に必要な一連の DMA コマンドが含まれる。これらの DMA コマンドには DMA コマンド 2340、2350、2355、2358 が含まれる。P U は DMA C へこれらの DMA コマンドを出す。

【0096】

DMA コマンド 2340 には V I D 2342 が含まれる。V I D 2342 は、DMA コマンドが出されたとき物理 I D に対して対応づけられる A P U のバーチャル I D である。DMA コマンド 2340 にはロード・コマンド 2344 とアドレス 2346 も含まれる。ロード・コマンド 2344 は、A P U に D R A M から特定の情報を読み出しローカル・ストレージの中へ入れるように命令する。アドレス 2346 によってこの特定情報を含む D R A M 内のバーチャル・アドレスが与えられる。この特定情報は、プログラム・セクション 2336 からのプログラムや、データ・セクション 2338 からのデータや、あるいはその他のデータなどであってもよい。最終的に、DMA コマンド 2340 にはローカル・ストレージのアドレス 2348 が含まれる。このアドレスによって、情報をロードできそうなローカル・ストレージのアドレスが識別される。DMA コマンド 2350 には類似の情報が含まれる。その他の DMA コマンドも使用可能である。

【0097】

DMA コマンド・リスト 2334 には一連のキック・コマンド(キック・コマンド 2355 と 2358 など)も含まれる。キック・コマンドとは、P U によって A P U へ出されるセルの処理を開始するコマンドである。DMA キック・コマンド 2355 には、バーチャル A P U I D 2352 と、キック・コマンド 2354 と、プログラム・カウンタ 2356 とが含まれる。バーチャル A P U I D 2352 はキックすべき対象 A P U を識別し、キック・コマンド 2354 は関連するキック・コマンドを与え、プログラム・カウンタ 2356 は、プログラムの実行用プログラム・カウンタのためのアドレスを与える。DMA キック・コマンド 2358 は、同じ A P U または別の A P U に対して同様の情報を与える。

【0098】

上述したように、P U は独立したプロセッサとして A P U を扱い、コプロセッサとして扱うものではない。したがって、A P U による処理を制御するために、P U は、遠隔手順呼出しに類似したコマンドを使用する。これらのコマンドは“ A P U 遠隔手順呼出し(A R P C)”と呼ばれる。P U は、一連の DMA コマンドを DMA C へ出すことにより A R P C を実行する。DMA C は、A P U プログラムとそれと関連するスタック・フレームとを A P U のローカル・ストレージの中へロードする。次いで、P U は A P U へ最初のキックを出し、A P U プログラムを実行する。

【0099】

図 4 2 は、アプレットを実行するための A R P C のステップを例示する。指定 A P U によるアプレットの処理の開始時に P U が実行するこれらのステップが、図 4 2 の第 1 の部分 2402 に示され、指定 A P U が実行するステップが、図 4 2 の第 2 の部分 2404 に示されている。

【0100】

ステップ 2410 で、P U はアプレットを評価し、次いで、アプレットの処理用 A P U を指定する。ステップ 2412 で、P U は、必要な単複のサンドボックス用のメモリ・アクセス・キーの設定を行う DMA コマンドを DMA C へ出すことにより、アプレットの実行用スペースを D R A M 内に割り振る。ステップ 2414 で、P U は、指定 A P U への割り込み要求による、アプレットの完了信号の伝送を可能にする。ステップ 2418 で、P U は、D R A M から A P U のローカル・ストレージへアプレットをロードする DMA コマンドを DMA C へ出す。ステップ 2420 で、DMA コマンドが実行され、アプレットが D R

10

20

30

40

50

AMからローカル・ストレージへ読み出される。ステップ2422で、PUは、アプレットと関連付けられたスタック・フレームをDRAMからAPUのローカル・ストレージへロードするDMAコマンドをDMACへ出す。ステップ2423で、DMAコマンドが実行され、スタック・フレームがDRAMからAPUのローカル・ストレージへ読み出される。ステップ2424で、PUは、DMACがAPUへキーを割り当てて、ステップ2412で指定された、一又は複数のハードウェア・サンドボックスからのデータ読み出しと、その一又は複数のハードウェア・サンドボックスへのデータ書き込みを行うことをAPUに許可するDMAコマンドを出す。ステップ2426で、DMACは、APUへ割り当てられたキーを用いてキー管理テーブル(KTAB)の更新を行う。ステップ2428で、PUは、プログラムの処理を開始するDMAコマンド“キック”をAPUへ出す。特定の

10

【0101】

上記のように、図42の第2の部分2404は、アプレットの実行時にAPUによって行われるステップを例示するものである。ステップ2430で、APUは、ステップ2428で出されるキック・コマンドに応じてアプレットの実行を開始する。ステップ2432で、アプレットの指示で、APUは、アプレットの関連スタック・フレームの評価を行う。ステップ2434で、APUは、DMACへ複数のDMAコマンドを出し、スタック・フレームが必要に応じてDRAMからAPUのローカル・ストレージへ指定するデータのロードを行う。ステップ2436で、これらのDMAコマンドが実行され、データは、D

20

【0102】

PUの指示の下で独立にタスクを実行するAPUの能力によって、1グループのAPUと、1グループのAPUと関連付けられたメモリ・リソースとを拡張タスクの実行専用にすることが可能になる。例えば、1つのPUは、1以上のAPUと、これらの1以上のAPUと関連付けられた1グループのメモリサンドボックスとを、拡張された時間中ネットワーク104を介して伝送されてくるデータの受信専用とし、また、1以上の他のAPUとそれらと関連付けられたメモリ・サンドボックスへ、この時間中受信したデータのさらなる処理を行うための送信専用とすることができる。この能力は、ネットワーク104を介して伝送されるストリーミング・データ(ストリーミングMPEGまたはストリーミングATRACオーディオまたはビデオ・データなど)の処理にとって特に好適である。PUは、1以上のAPUおよびそれらと関連付けられたメモリ・サンドボックスをこれらのデータの受信専用とし、1以上の他のAPUおよびそれらと関連付けられたメモリ・サンドボックスをこれらのデータの解凍と処理専用とすることができる。言い換えれば、PUは、APUのグループとそれらと関連付けられたメモリ・サンドボックスとの間でこのよう

30

40

【0103】

しかし、このような処理を効率的に実行するためには、パイプ・ラインの専用APUとメモリサンドボックスとが、データ・ストリームを含むアプレットの処理が行われない時間中もパイプ・ライン専用のものであることが望ましい。言い換えれば、専用APUおよびそれらと関連するサンドボックスが、これらの時間中予約状態のままに置かれることが望ましい。アプレットの処理の完了時における、APUとその関連付けられた一又は複数のメモリ・サンドボックスを予約、即ちリザーブ状態としておくことは、“常駐終了”と呼ばれる。常駐終了はPUからの命令に応じて行われる。

【0104】

50

図43、44、45は、1グループのAPUおよびそれらと関連するサンドボックスを含む、ストリーミング・データ(ストリーミングMPEGデータなど)を処理するための専用パイプライン構造の設定を例示する。図43に示すように、このパイプライン構造の構成要素にはPE2502とDRAM2518とが含まれる。PE2502の中には、PU2504、DMAC2506およびAPU2508、APU2510、APU2512を含む複数のAPUが含まれる。PU2504、DMAC2506およびこれらのAPU間の通信はPEバス2514を介して行われる。広帯域幅のバス2516によってDMAC2506はDRAM2518と接続される。DRAM2518の中には、複数のサンドボックス(サンドボックス2520、サンドボックス2522、サンドボックス2524、サンドボックス2526など)が含まれる。

10

【0105】

図44に、専用パイプラインを設定するためのステップを例示する。ステップ2610で、PU2504は、ネットワーク・アプレットを処理するようにAPU2508を割り当てる。ネットワーク・アプレットは、ネットワーク104のネットワーク・プロトコルの処理用プログラムを有する。この場合、このプロトコルは 伝送制御プロトコル/インターネット用プロトコル(TCP/IP)である。このプロトコルに従うTCP/IPデータ・パケットはネットワーク104を介して伝送される。受信時に、APU2508はこれらのパケットを処理し、パケット内のデータを組み立て、ソフトウェア・セル102の中へ入れる。ステップ2612で、PU2504は、ネットワーク・アプレットの処理の完了時に常駐終了を実行するようにAPU2508に指示する。ステップ2614で、PU2504は、APU2510及び2512がMPEGアプレットの処理を行うように割り当てる。ステップ2615で、PU2504は、MPEGアプレットの処理の完了時に常駐終了を実行するようにAPU2510及び2512に指示する。ステップ2616で、PU2504は、APU2510によるアクセス用ソース・サンドボックス及びAPU2508によるアクセス用宛先サンドボックスとしてサンドボックス2520を指定する。ステップ2618で、PU2504は、APU2510によるアクセス用宛先サンドボックス及びAPU2512によるアクセス用ソースサンドボックスとしてサンドボックス2522を指定する。ステップ2620で、PU2504は、APU2512によるアクセス用宛先サンドボックス及びパイプライン内の後段のAPUによるアクセス用ソースサンドボックスとしてサンドボックス2524を指定する。ステップ2622で、PU2504は、パイプライン内の後段のAPUによるアクセス用宛先サンドボックス及びアクセス用ソースサンドボックスとしてサンドボックス2526を指定する。ステップ2624で、APU2510とAPU2512とは、それぞれ、ソース・サンドボックス2520とソース・サンドボックス2522の範囲内のメモリ・ブロックへ同期読取りコマンドを送り、これらのメモリ・ブロックをブロック状態に設定する。最後に、処理はステップ2628へ移り、そこで、専用パイプラインの設定が完了し、パイプ・ライン専用のリソースが予約される。このようにして、APU2508、2510、2512等およびそれらと関連するサンドボックス2520、2522、2524および2526は予約状態に入る。

20

30

【0106】

図45に、この専用パイプラインによるストリーミングMPEGデータの処理ステップを例示する。ステップ2630で、APU2508は、ネットワーク・アプレットを処理し、そのローカル・ストレージの中で、TCP/IPデータ・パケットをネットワーク104から受信する。ステップ2632で、APU2508は、これらのTCP/IPデータ・パケットを処理し、これらのパケット内のデータをアSEMBルし、ソフトウェア・セル102の中へ入れる。ステップ2634で、APU2508はソフトウェア・セルのヘッダ2320(図23)をチェックし、セルがMPEGデータを含むかどうかの判定を行う。セルがMPEGデータを含まない場合、ステップ2636で、APU2508は、専用パイプライン内に含まれない他のAPUによって他のデータを処理するために、DRAM2518内に指定される汎用サンドボックスへそのセルを伝送する。またAPU2508は

40

50

この伝送について P U 2 5 0 4 に通知する。

【 0 1 0 7 】

一方、ソフトウェア・セルが M P E G データを含む場合、ステップ 2 6 3 8 で、A P U 2 5 0 8 はそのセルの前のセルの I D 2 3 3 0 (図 4 1) をチェックし、そのセルが属する M P E G データ・ストリームを識別する。ステップ 2 6 4 0 で、A P U 2 5 0 8 はセルの処理用の専用パイプラインの A P U を選択する。この場合、A P U 2 5 0 8 は、これらのデータを処理する A P U 2 5 1 0 を選択する。この選択は前回のセル I D 2 3 3 0 とロード・バランシング・ファクタ (負荷平衡係数) とに基づく。例えば、そのソフトウェア・セルが属する M P E G データ・ストリームの前回のソフトウェア・セルが処理用として A P U 2 5 1 0 へ送られたことが前のセル I D 2 3 3 0 によって示されている場合、現在のソフトウェア・セルも通常の処理用として A P U 2 5 1 0 へ送られる。ステップ 2 6 4 2 で、A P U 2 5 0 8 は、サンドボックス 2 5 2 0 へ M P E G データを書き込む同期書き込みコマンドを出す。このサンドボックスは予めブロッキング状態に設定されているので、ステップ 2 6 4 4 で、M P E G データは、サンドボックス 2 5 2 0 から A P U 2 5 1 0 のローカル・ストレージへ自動的に読み出される。ステップ 2 6 4 6 で、A P U 2 5 1 0 はそのローカル・ストレージで M P E G データを処理してビデオ・データを生成する。ステップ 2 6 4 8 で、A P U 2 5 1 0 はサンドボックス 2 5 2 2 へビデオ・データを書き込む。ステップ 2 6 5 0 で、A P U 2 5 1 0 はサンドボックス 2 5 2 0 へ同期読取りコマンドを出し、このサンドボックスに追加 M P E G データの受信を準備させる。ステップ 2 6 5 2 で、A P U 2 5 1 0 は常駐終了処理を行う。この処理によってこの A P U は予約状態に入り、この予約状態の間 A P U は、M P E G データ・ストリームの中で追加 M P E G データの処理を行うべく待機する。

10

20

【 0 1 0 8 】

他のタイプのデータ処理用として 1 グループの A P U およびそれらと関連するサンドボックス間でその他の専用構造の設定が可能である。例えば、図 4 6 に示すように、A P U の専用グループ (A P U 2 7 0 2 、 2 7 0 8 、 2 7 1 4 など) を設定し、3 次元オブジェクトに対して幾何学的変換を実行して 2 次元ディスプレイ・リストの生成を行うことが可能となる。これらの 2 次元ディスプレイ・リストを他の A P U によってさらに処理 (レンダ) し画素データの生成を行うようにすることが可能である。この処理を実行するために、3 次元オブジェクトと、これらのオブジェクト処理から結果として生じるディスプレイ・リストの格納用として、サンドボックスが、A P U 2 7 0 2 、 2 7 0 8 、 2 4 1 4 の専用となる。例えば、ソース・サンドボックス 2 7 0 4 、 2 7 1 0 、 2 7 1 6 は、それぞれ、A P U 2 7 0 2 、 A P U 2 7 0 8 、 A P U 2 7 1 4 によって処理された 3 次元オブジェクトの格納専用となる。同様に、宛先サンドボックス 2 7 0 6 、 2 7 1 2 、 2 7 1 8 は、それぞれ、A P U 2 7 0 2 、 A P U 2 7 0 8 、 A P U 2 7 1 4 によるこれらの 3 次元オブジェクトの処理から結果として生じるディスプレイ・リストの格納専用となる。

30

【 0 1 0 9 】

調整用 A P U 2 7 2 0 は、そのローカル・ストレージにおける、宛先サンドボックス 2 7 0 6 、 2 7 1 2 、 2 7 1 8 からのディスプレイ・リストの受信専用である。A P U 2 7 2 0 は、これらのディスプレイ・リスト間での調整を行い、画素データのレンダリングのためにこれらのディスプレイ・リストを他の A P U へ送る。

40

【 0 1 1 0 】

システム 1 0 1 のプロセッサは絶対タイマーも使用する。この絶対タイマーは A P U と P E の他のエレメントへクロック信号を出力する。このクロック信号はこれらのエレメントを駆動するクロック信号に依存せず、かつ、このクロック信号より高速である。この絶対タイマーの利用が図 4 7 に例示されている。

【 0 1 1 1 】

この図に示すように、この絶対タイマーによって A P U によるタスク・パフォーマンスのためのタイム・バジェット (割り当て時間) が決定される。このタイム・バジェットによって、これらのタスクの完了時間が設定されるが、この時間は A P U によるタスク処理に

50

必要な時間より長い時間になる。その結果、各タスクについて、タイム・バジェットの範囲内に、ビジーな時間とスタンバイ時間とが存在することになる。すべてのアプレットは、A P U の実際の処理時間にかかわらず、このタイム・バジェットに基づいて処理を行うように書かれる。

【 0 1 1 2 】

例えば、P E の特定のA P U 用として、タイム・バジェット 2 8 0 4 のビジー時間 2 8 0 2 中に特定のタスクを行うことができる。ビジー時間 2 8 0 2 がタイム・バジェット 2 8 0 4 未満であるため、スタンバイ時間 2 8 0 6 がタイム・バジェット中に生じる。このスタンバイ時間中、A P U は、A P U が消費するパワーが少なくなるスリープモードに入る。

10

【 0 1 1 3 】

タイム・バジェット 2 8 0 4 が満了するまで、他のA P U またはP E の他のエレメントがタスク処理の結果を予想することはない。したがって、A P U の実際の処理速度にかかわらず、絶対タイマーによって決定されるタイム・バジェットを用いてA P U の処理結果が常時調整される。

【 0 1 1 4 】

将来、A P U による処理速度はさらに高速になる。しかし、絶対タイマーによって設定されるタイム・バジェットは同じままである。例えば、図 4 7 に示すように、将来のA P U は、さらに短時間でタスクを実行することになり、したがって、スタンバイ時間はさらに長くなるであろう。したがって、ビジー時間 2 8 0 8 はビジー時間 2 8 0 2 より短くなり、スタンバイ時間 2 8 1 0 はスタンバイ時間 2 8 0 6 より長くなる。しかし、絶対タイマーによって設定された同じタイム・バジェットに基づいて処理を行うようにプログラムが書かれているので、A P U 間での処理結果の調整が維持される。その結果、さらに高速のA P U が、その処理の結果が予想される時点でコンフリクトを生じることなく、低速のA P U 用として書かれたプログラムの処理を行うことが可能となる。

20

【 0 1 1 5 】

動作速度の向上や動作速度が異なることに起因するA P U の並列処理の調整問題に対しては、A P U 間での調整を決定する絶対タイマーに代えて、P U または1以上の指定A P U において、A P U が実行している特定の命令(マイクロコード)の分析をアプレットの処理時に行うようにすることもできる。“オペレーションなし”(“NOOP”)命令を命令の中へ挿入し、A P U のいくつかによってこの命令を実行して、アプレットによって予想されるA P U による処理を1ステップずつ適切に行うことが可能となる。命令の中へこれらのNOOPを挿入することにより、すべての命令のA P U による実行を行うための正しいタイミングの維持が可能となる。

30

【 0 1 1 6 】

以上特定の実施形態に関して本明細書で本発明について説明したが、これらの実施形態は本発明の原理と適用を示す単に例示的なものであると理解すべきである。したがって、添付の請求項によって画定されているような本発明の精神と範囲から逸脱することなく、以上の例示の実施形態に対して多数の改変を行うことが可能であり、また、他の構成を考案することが可能である。

40

【 図面の簡単な説明 】

【 図 1 】 本発明によるコンピュータ・ネットワークのアーキテクチャ全体を例示する。

【 図 2 】 本発明によるプロセッサ・エレメント(P E)の構造を例示する図である。

【 図 3 】 本発明による広帯域エンジン(B E)の構造を例示する図である。

【 図 4 】 本発明による付加処理ユニット(A P U)の構造を例示する図である。

【 図 5 】 本発明によるプロセッサ・エレメントと、ビジュアライザ(V S)と、光インターフェースとの構造を例示する図である。

【 図 6 】 本発明によるプロセッサ・エレメントの1つの組合せを例示する図である。

【 図 7 】 本発明によるプロセッサ・エレメントの別の組合せを例示する図である。

【 図 8 】 本発明によるプロセッサ・エレメントのさらに別の組合せを例示する図である

50

- 。
- 【図 9】 本発明によるプロセッサ・エレメントのさらに別の組合せを例示する図である。
- 。
- 【図 10】 本発明によるプロセッサ・エレメントのさらに別の組合せを例示する図である。
- 【図 11】 本発明によるチップ・パッケージ内での光インターフェースの統合化を例示する図である。
- 【図 12】 図 11 の光インターフェースを用いるプロセッサの 1 つの構成を示す図である。
- 【図 13】 図 11 の光インターフェースを用いるプロセッサの別の構成を示す図である 10
- 。
- 【図 14】 本発明によるメモリ・システムの構造を例示する図である。
- 【図 15】 本発明による第 1 の広帯域エンジンから第 2 の広帯域エンジンへのデータの書き込みを例示する図である。
- 【図 16】 本発明によるプロセッサ・エレメントのための共用メモリの構造を示す図である。
- 【図 17】 図 16 に示すメモリ・バンク用の 1 つの構成を例示する図である。
- 【図 18】 図 16 に示すメモリ・バンク用の別の構成を例示する図である。
- 【図 19】 本発明による D M A C のための構造を例示する図である。
- 【図 20】 本発明による D M A C のための代替の構造を例示する図である。 20
- 【図 21】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 22】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 23】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 24】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 25】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 26】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 27】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 28】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 29】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 30】 本発明によるデータ同期オペレーションを例示する図である。 30
- 【図 31】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 32】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 33】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 34】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 35】 本発明によるデータ同期オペレーションを例示する図である。
- 【図 36】 本発明のデータ同期方式によるメモリ・ロケーションの様々な状態を例示する 3 つの状態のメモリ図である。
- 【図 37】 本発明によるハードウェア・サンドボックス用のキー管理テーブルの構造を例示する図である。
- 【図 38】 本発明によるハードウェア・サンドボックス用メモリ・アクセス・キーの格納方式を例示する図である。 40
- 【図 39】 本発明によるハードウェア・サンドボックス用メモリ・アクセス管理テーブルの構造を例示する図である。
- 【図 40】 図 37 のキー管理テーブルと図 39 のメモリ・アクセス管理テーブルとを用いてメモリ・サンドボックスにアクセスするステップを示すフロー・チャートである。
- 【図 41】 本発明によるソフトウェア・セルの構造を例示する図である。
- 【図 42】 本発明による、 A P U へ遠隔処理命令を出すステップを示すフロー・チャートである。
- 【図 43】 本発明による、ストリーミング・データ処理用専用パイプラインの構造を例示する図である。 50

【図44】 本発明によるストリーミング・データの処理時の図43の専用パイプラインによって実行されるステップを示すフロー・チャートである。

【図45】 本発明によるストリーミング・データの処理時の図43の専用パイプラインによって実行されるステップを示すフロー・チャートである。

【図46】 本発明によるストリーミング・データ処理用の専用パイプラインの他の構造を例示する図である。

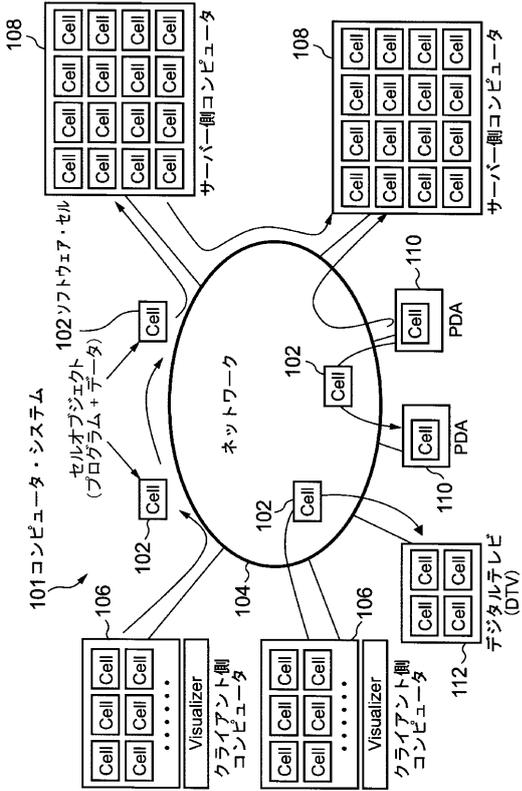
【図47】 本発明によるAPUによるアプリケーションとデータの並列処理を調整するための絶対タイマー方式を例示する図である。

【符号の説明】

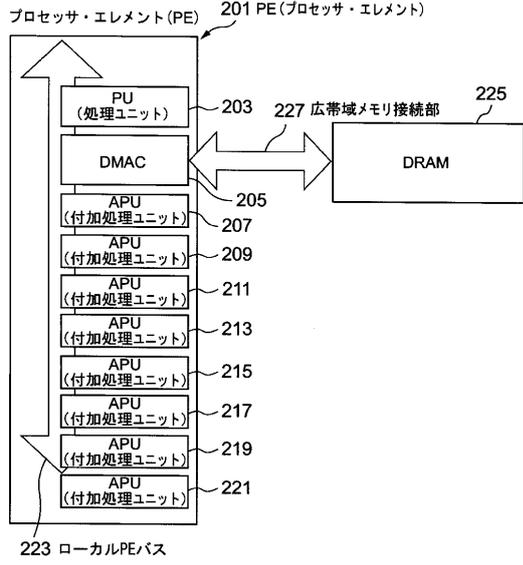
101	システム	10
1010	キー	
102	セル	
104	ネットワーク	
106	クライアント	
108	サーバーコンピュータ	
1104	光インターフェース	
1108	バス	
1118	ポート	
1122	ポート	
1126	光導波路	20
1160	光インターフェース	
1162	光インターフェース	
1164	光インターフェース	
1166	光インターフェース	
1182	光インターフェース	
1184	光インターフェース	
1186	光インターフェース	
1188	光インターフェース	
1188	光インターフェース	
1190	光インターフェース	30
1190	光インターフェース	
1206	コントロール	
1212	ユニット	
1221	クロスバ交換機	
1232	外部ポート	
1234	コントロール	
1240	ユニット	
1242	コントロール	
1244	バンク	
1406	ブロック	40
1414	バンク	
1416	バンク	
1504	ノード	
1607	バス	
1608	バス	
1722	制御回路	
1724	制御論理回路	
1726	ストレージ	
1728	ロケーション	
1729	セグメント	50

1 7 3 1	ロケーション	
1 7 3 2	ロケーション	
1 7 4 2	制御論理回路	
1 7 4 6	ロケーション	
1 7 5 0	ロケーション	
1 7 5 2	セグメント	
1 7 6 0	セグメント	
1 7 6 2	セグメント	
1 8 8 0	エンプティ状態	
1 8 8 2	フル状態	10
1 8 8 4	ブロッキング状態	
1 9 0 2	キー管理テーブル	
1 9 0 6	キー	
1 9 0 8	マスク	
2 0 0 6	格納位置	
2 0 0 8	セグメント	
2 0 1 0	セグメント	
2 0 1 2	キー	
2 1 0 2	アクセス管理テーブル	
2 1 0 6	アドレス	20
2 1 1 0	キー	
2 1 1 0	キーマスク	
2 2 3	バス	
2 2 7	高帯域メモリ接続部	
2 3 0 2	セル	
2 3 0 8	ヘッダ	
2 3 2 0	ヘッダ	
2 3 2 2	インターフェース	
2 3 3 2	実行セクション	
2 3 3 4	リスト	30
2 5 2 0	サンドボックス	
2 5 2 2	サンドボックス	
2 5 2 4	サンドボックス	
2 5 2 6	サンドボックス	
2 7 0 4	サンドボックス	
2 7 0 6	宛先サンドボックス	
3 0 1	広帯域エンジン	
3 1 1	バス	
3 1 3	広帯域メモリ接続部	
3 1 7	インターフェース	40
3 1 9	外部バス	
4 0 6	メモリ	
4 0 8	内部バス	
4 1 0	レジスタ	
4 1 2	浮動小数点演算ユニット	
4 1 4	整数演算ユニット	
4 2 0	バス	
5 0 6	パッケージの中に光インターフェース	
5 0 8	エンジン	
5 1 0	画像用キャッシュ	50

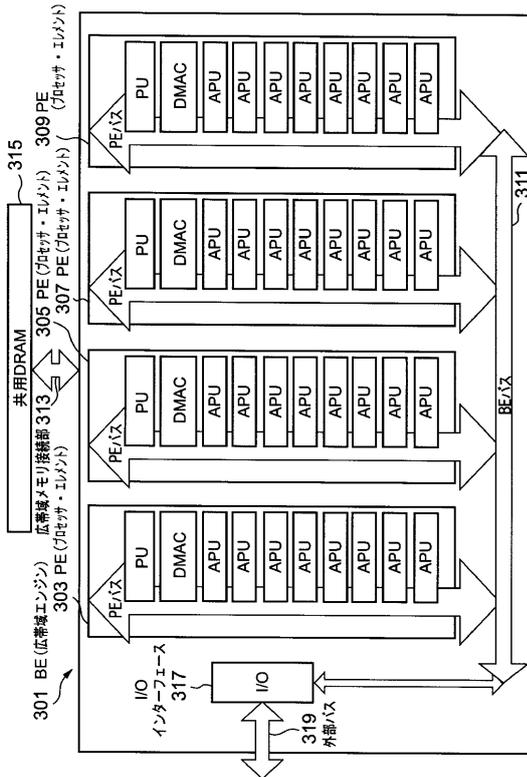
【図1】



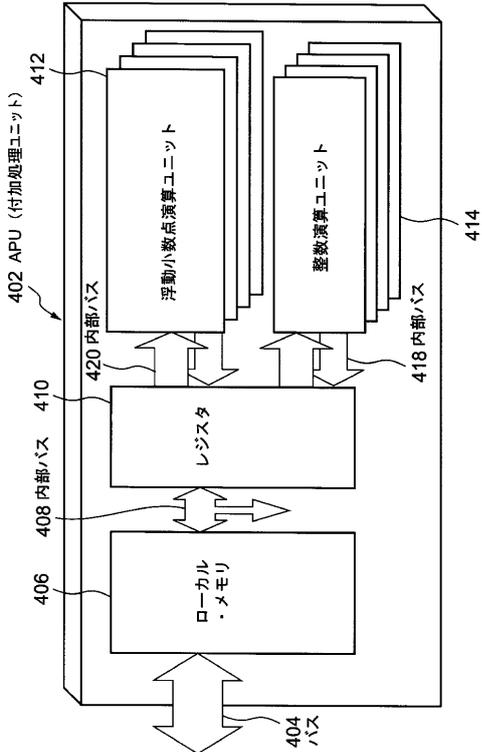
【図2】



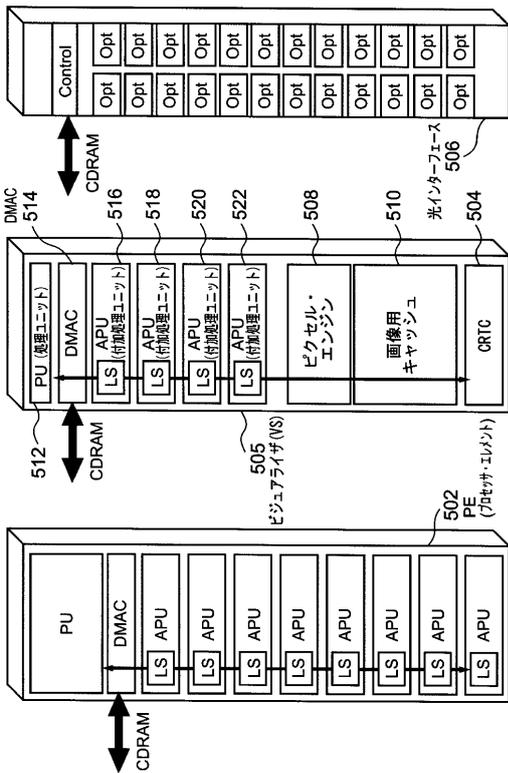
【図3】



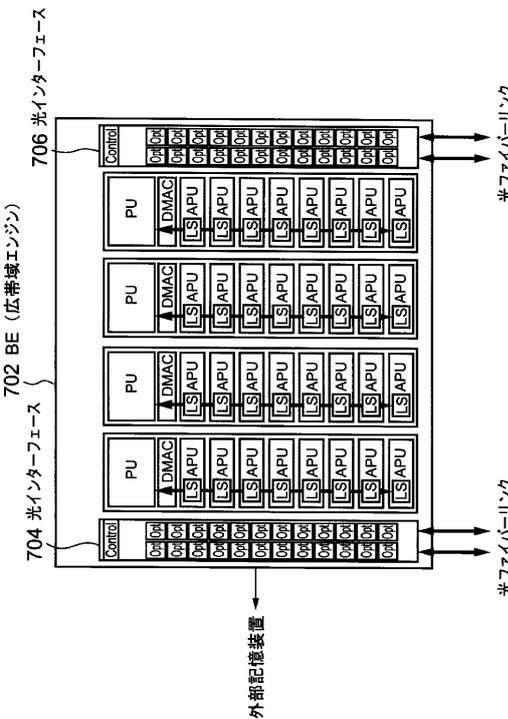
【図4】



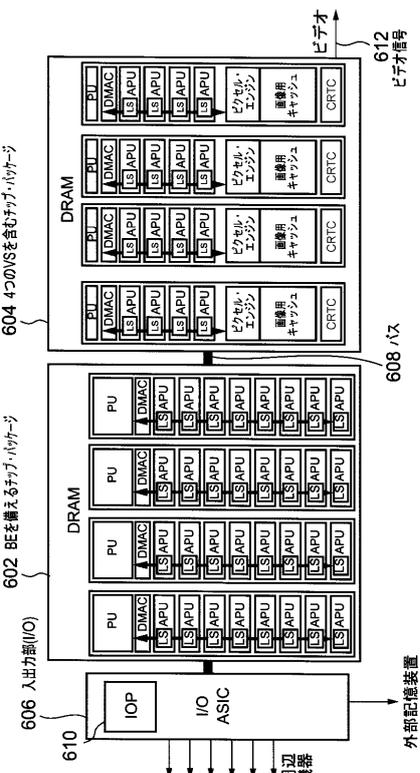
【 図 5 】



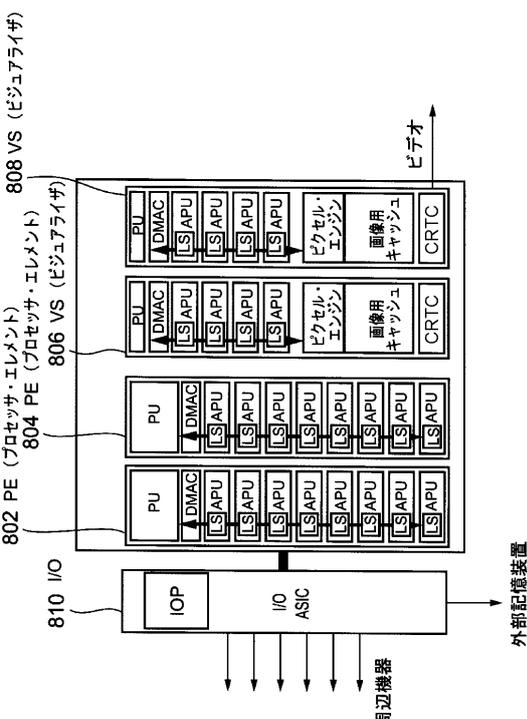
【 図 7 】



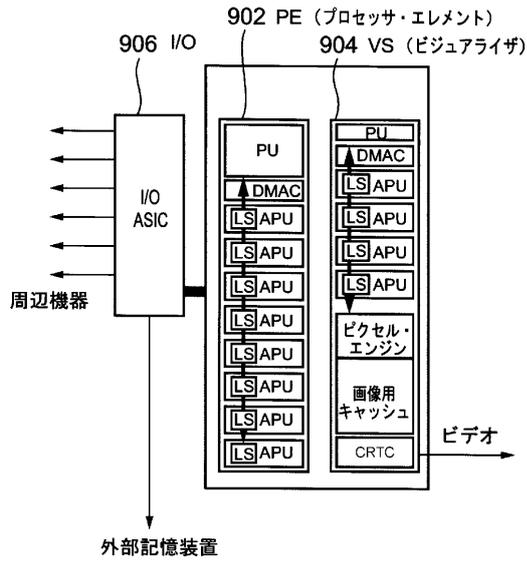
【 図 6 】



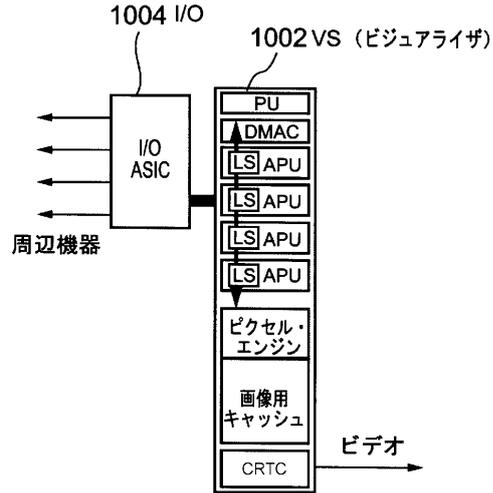
【 図 8 】



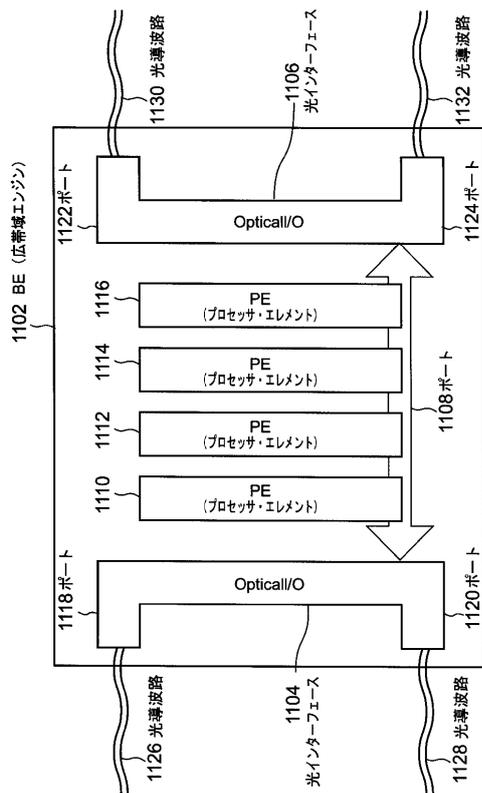
【図9】



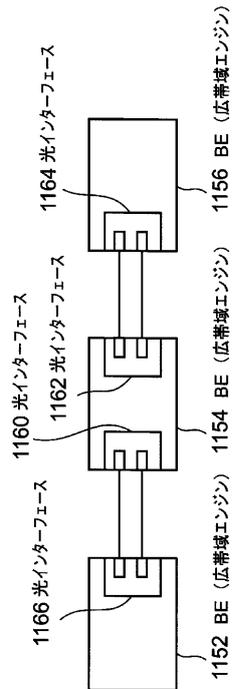
【図10】



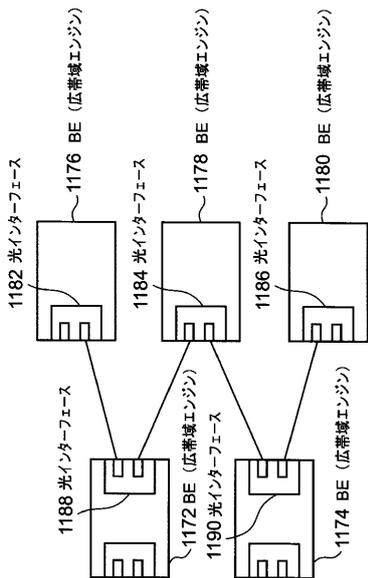
【図11】



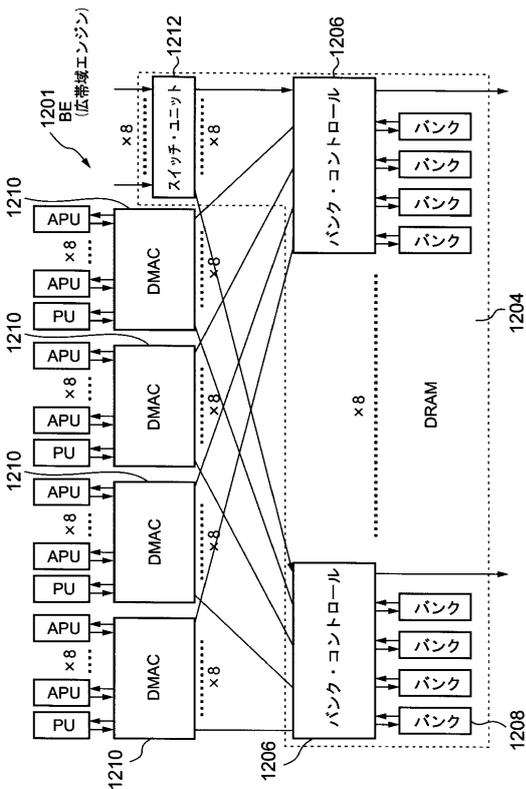
【図12】



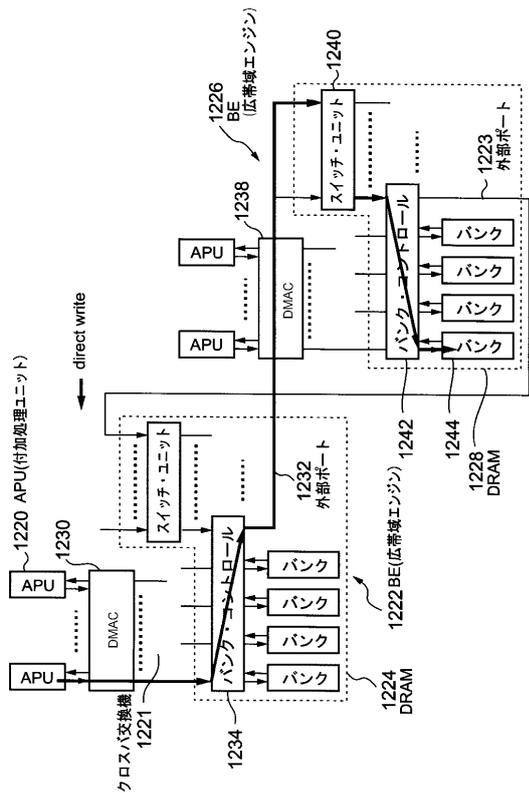
【図 13】



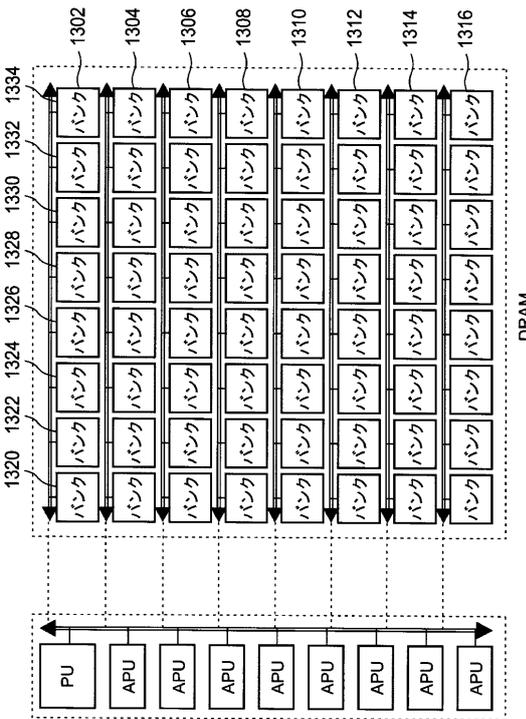
【図 14】



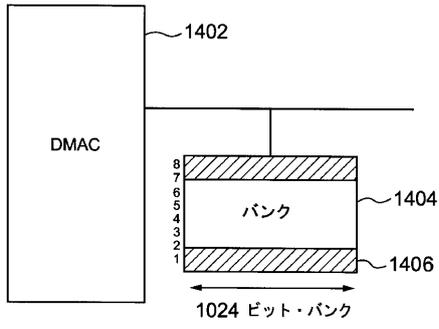
【図 15】



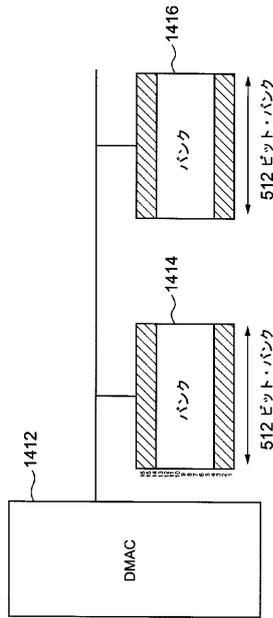
【図 16】



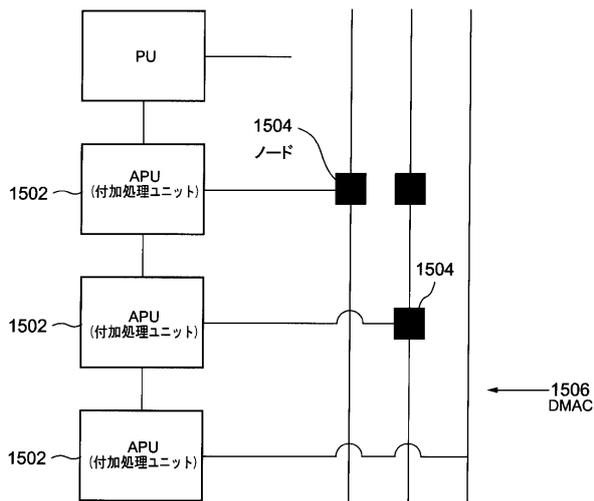
【図 17】



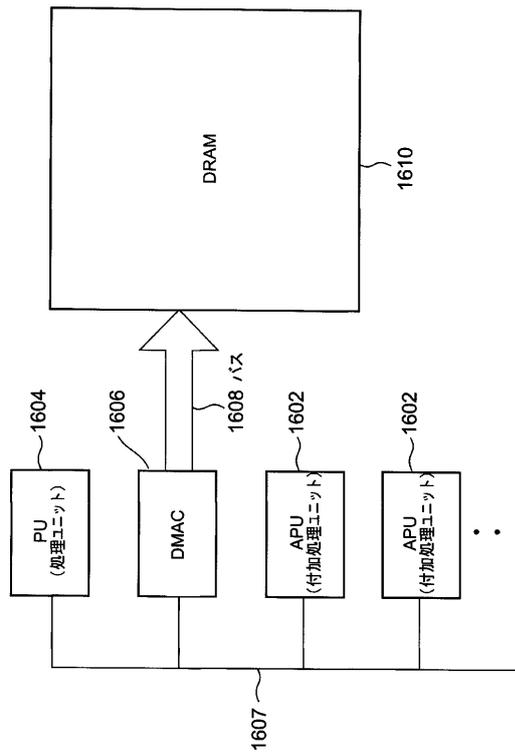
【図 18】



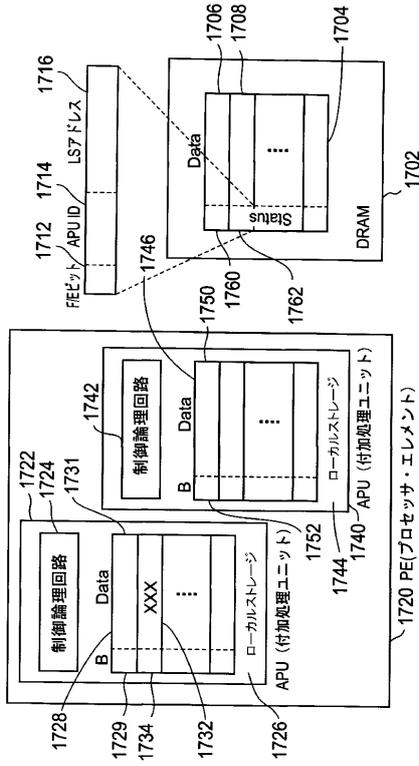
【図 19】



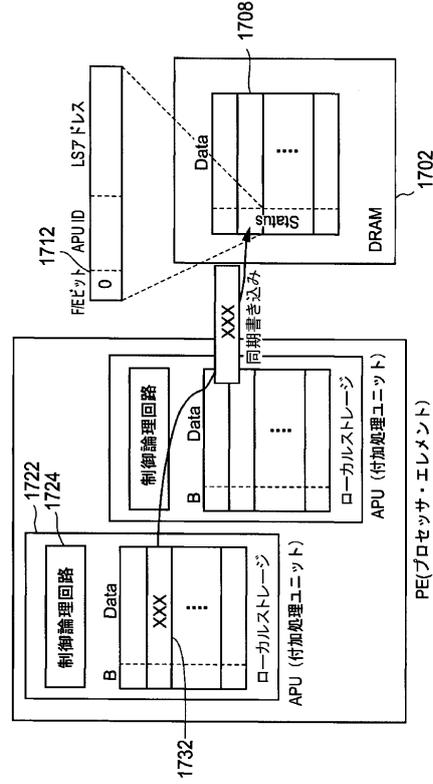
【図 20】



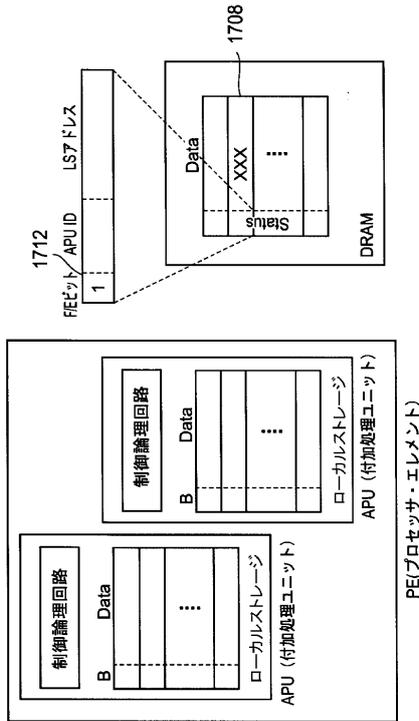
【図 2 1】



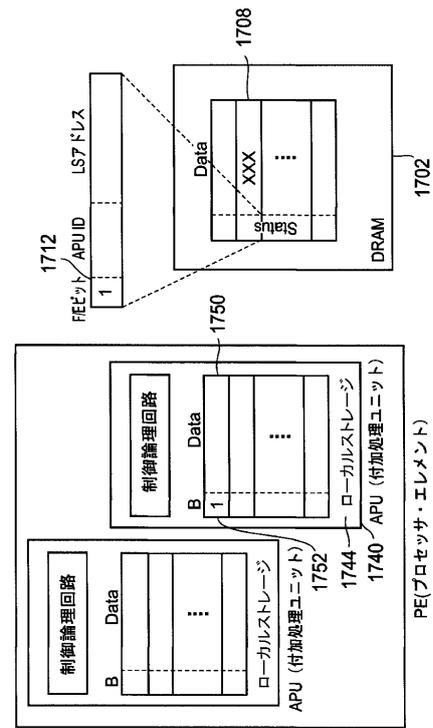
【図 2 2】



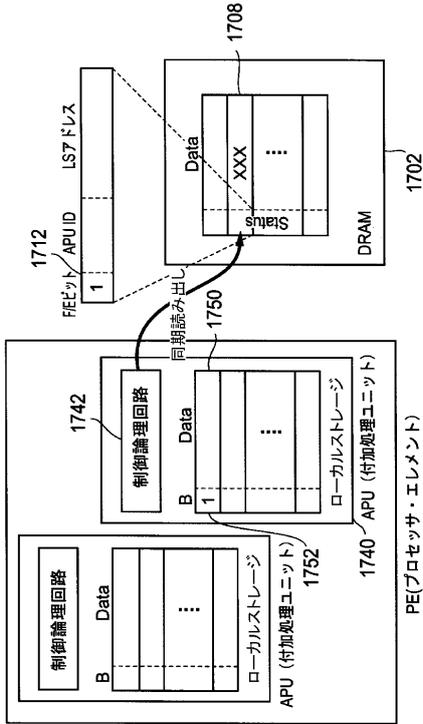
【図 2 3】



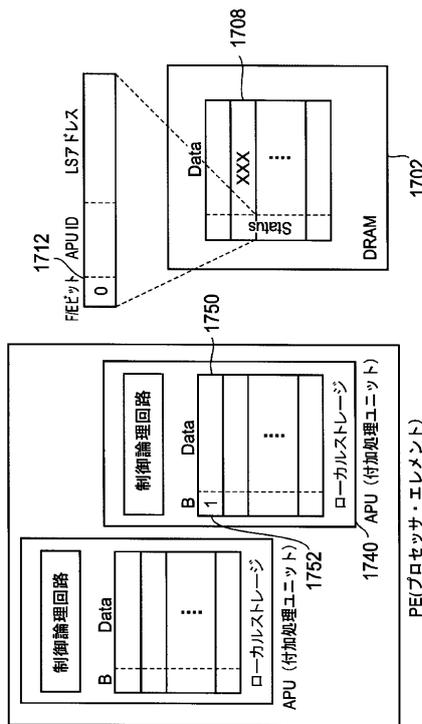
【図 2 4】



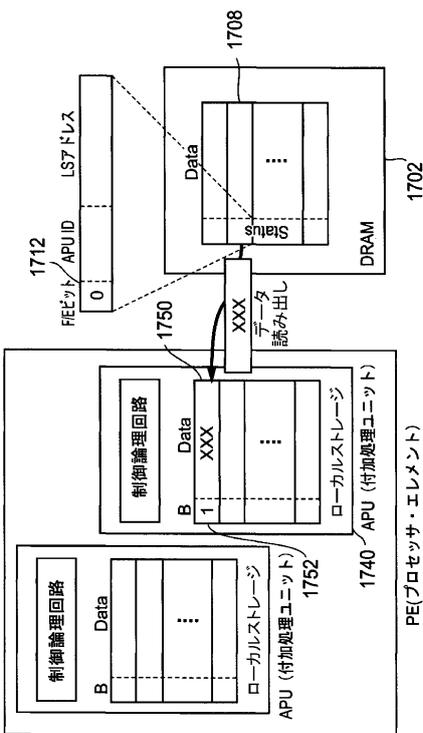
【図 25】



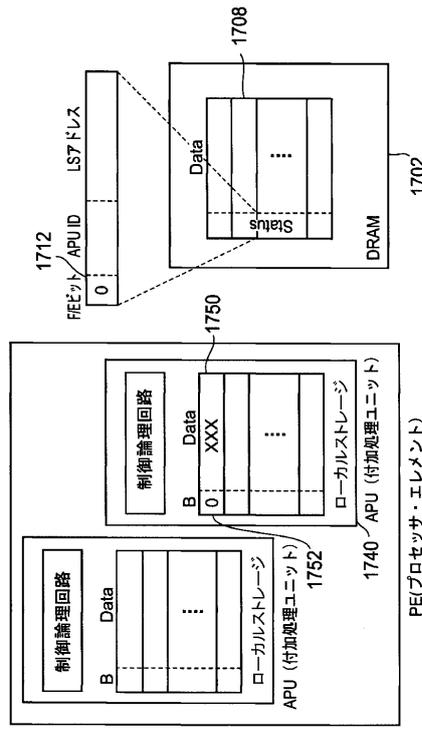
【図 26】



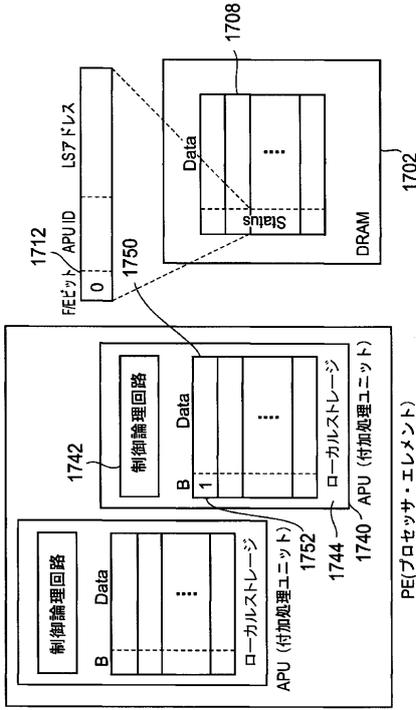
【図 27】



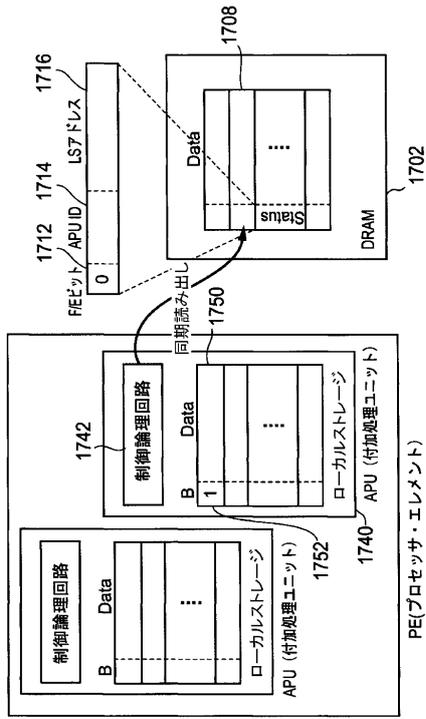
【図 28】



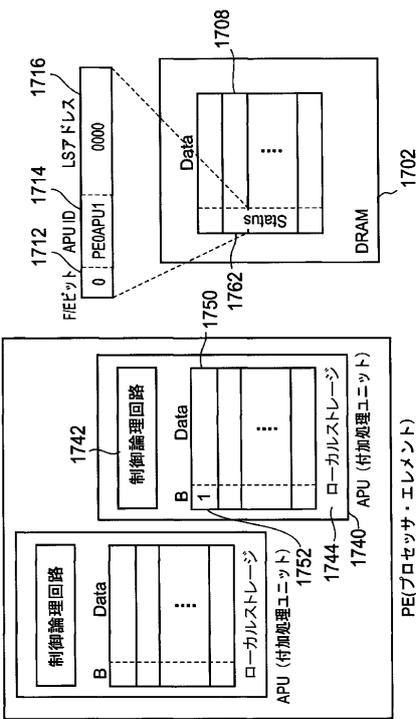
【図 29】



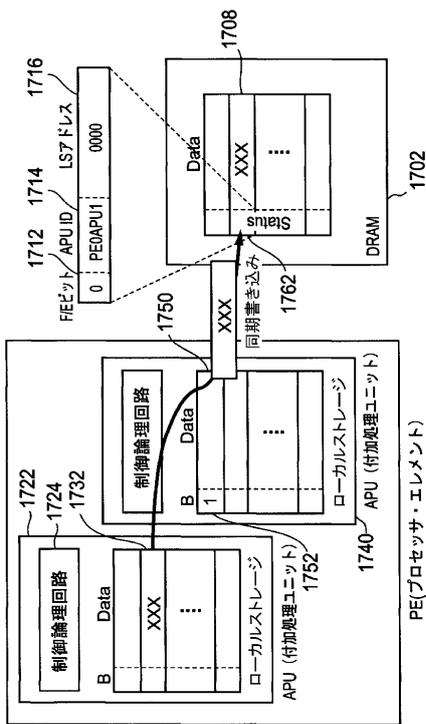
【図 30】



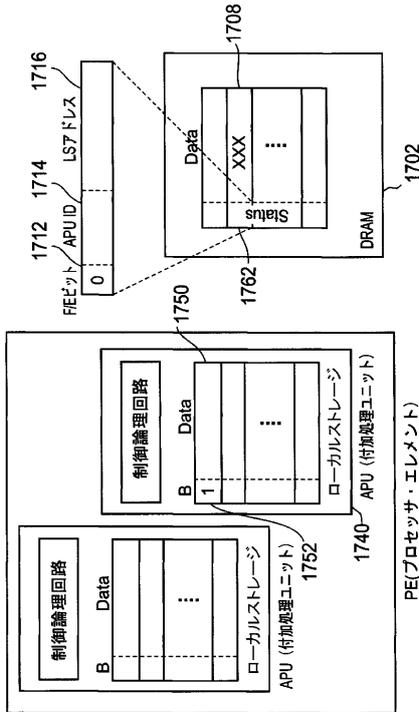
【図 31】



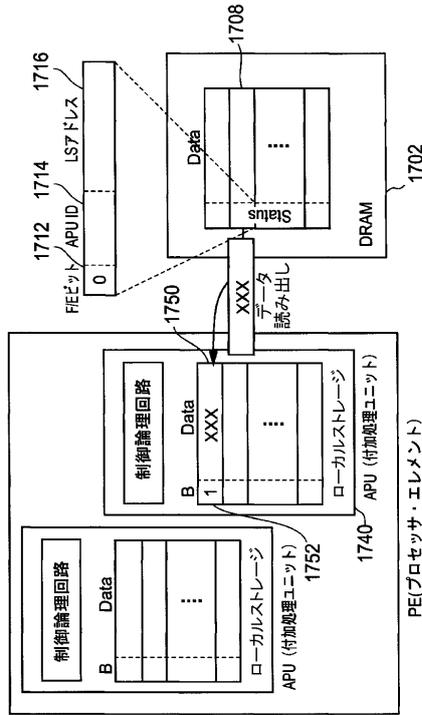
【図 32】



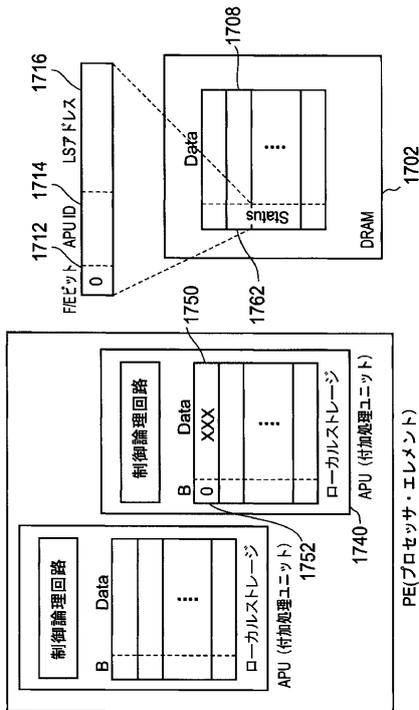
【図 3 3】



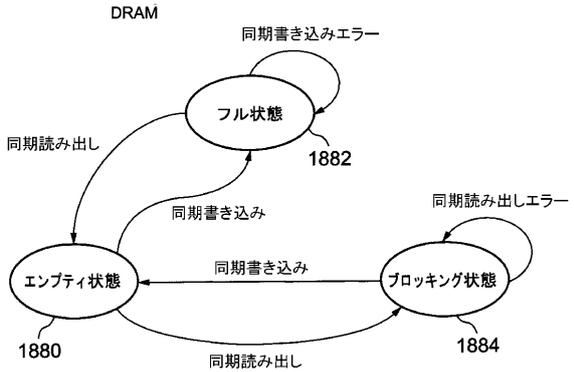
【図 3 4】



【図 3 5】



【図 3 6】

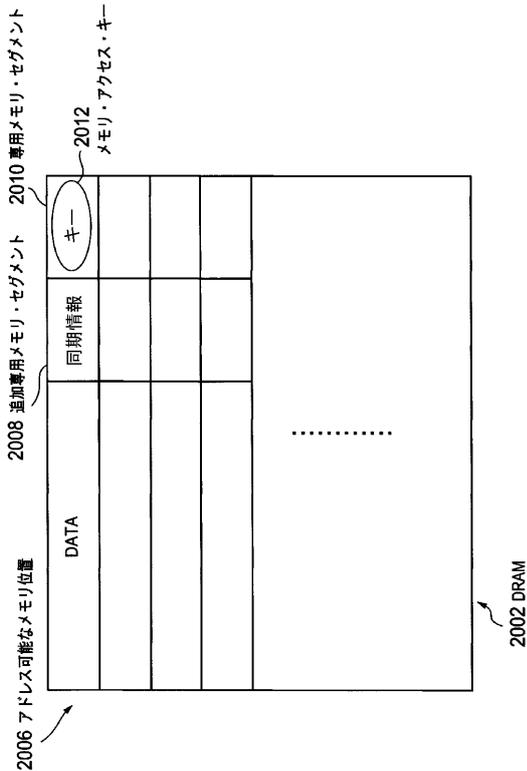


【図 3 7】

1902 キー管理テーブル

ID	APU用の識別子(ID)	キー・マスク
0	APUキー	キー・マスク
1	APUキー	キー・マスク
2	APUキー	キー・マスク
...
7	APUキー	キー・マスク

【図38】



【図39】

2102 メモリ・アクセス管理テーブル

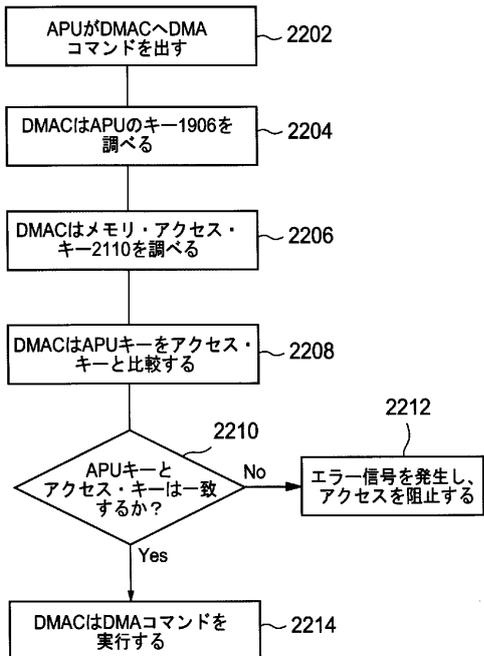
ベース・メモリ・アドレス サンドボックス・サイズ アクセス・キー

2106 2108 2110 2112 アクセス・キーマスク

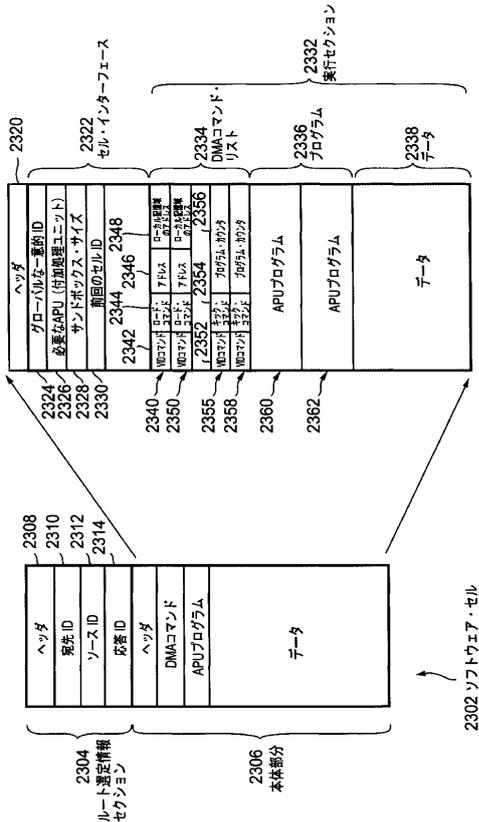
ID	Base	Size	Access Key	Access Key Mask
0	Base	Size	Access Key	Access Key Mask
1	Base	Size	Access Key	Access Key Mask
2	Base	Size	Access Key	Access Key Mask
.....
63	Base	Size	Access Key	Access Key Mask

2104 サンドボックス用識別子(ID)

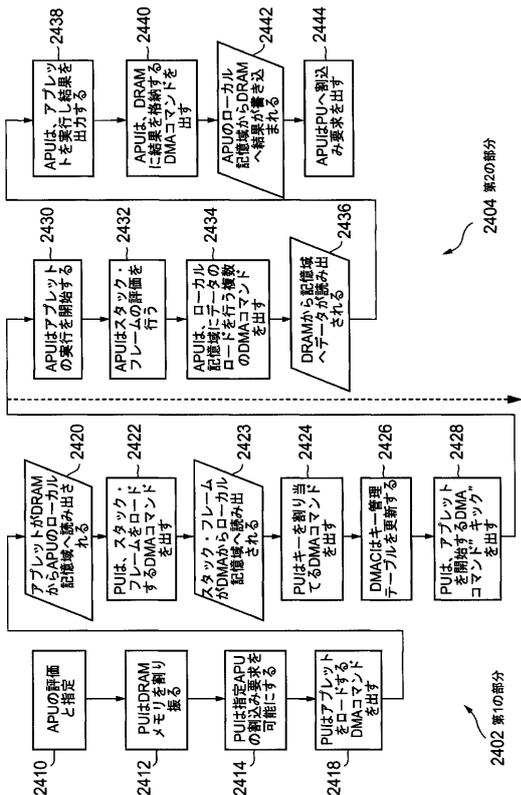
【図40】



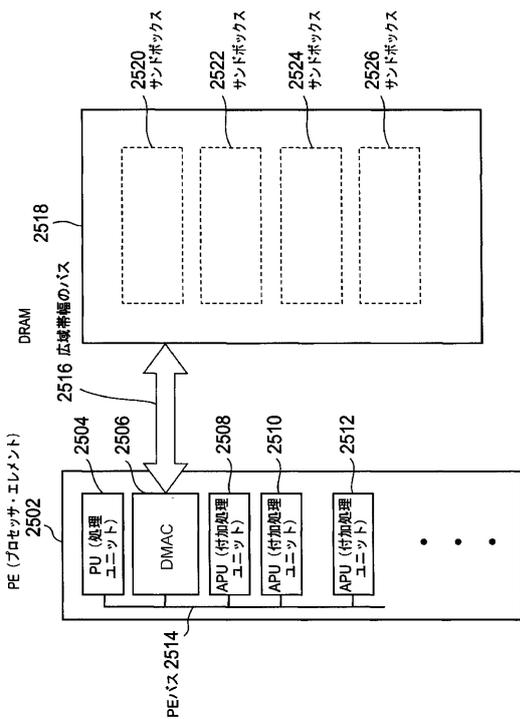
【図41】



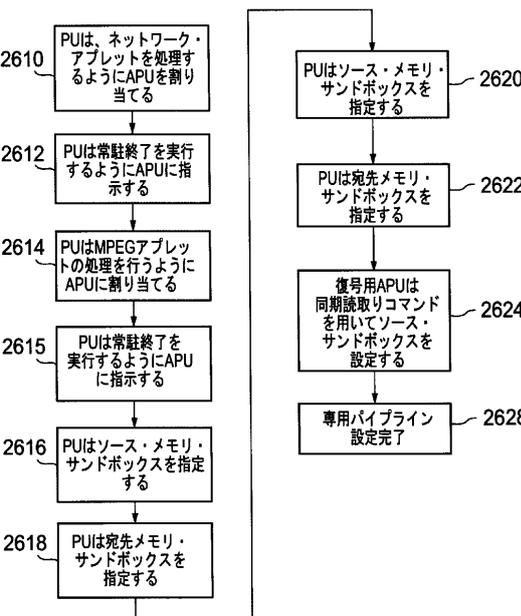
【図42】



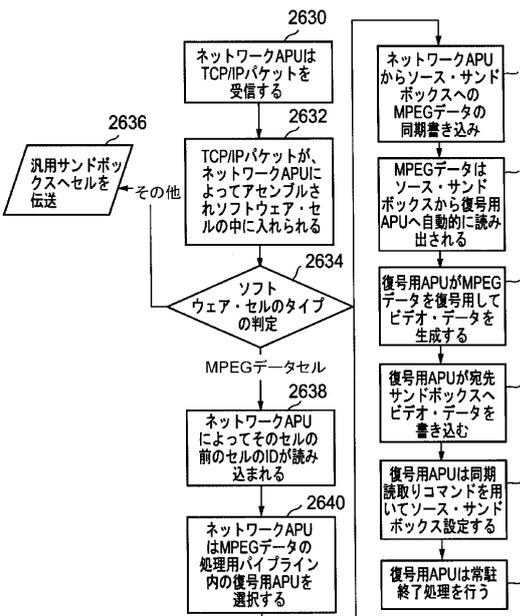
【図43】



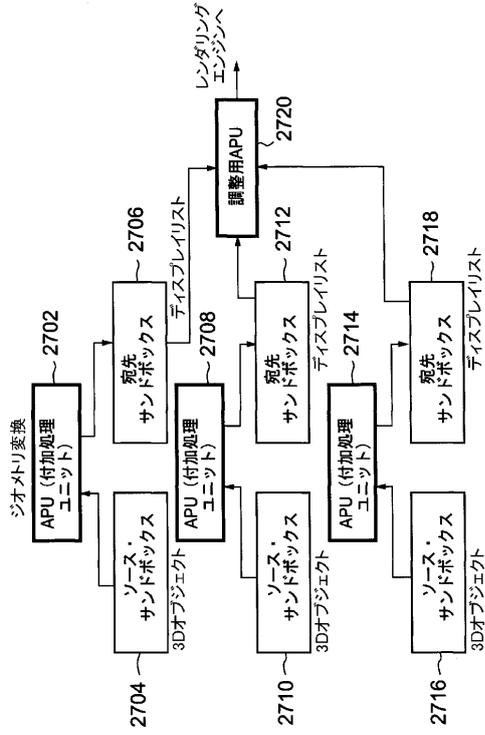
【図44】



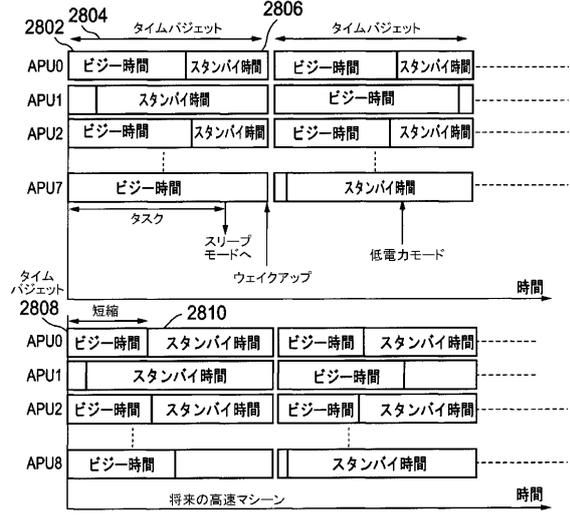
【図45】



【図46】



【図47】



将来の高速マシン

フロントページの続き

(72)発明者 山崎 剛

アメリカ合衆国、カリフォルニア州 94404 - 2175、フォスター シティ、セカンド
フロア、イースト ヒルスデイル ブルバード 919 ソニー コンピュータエンタテインメン
ト アメリカ、インク。内

審査官 鈴木 修治

(56)参考文献 特開2000 - 132411 (JP, A)

特開平02 - 211575 (JP, A)

特開2000 - 284974 (JP, A)

特開平07 - 325725 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 9/46-9/54