

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4756553号
(P4756553)

(45) 発行日 平成23年8月24日(2011.8.24)

(24) 登録日 平成23年6月10日(2011.6.10)

(51) Int.Cl.

G06F 9/50 (2006.01)

F 1

G06F 9/46 465D

請求項の数 8 (全 17 頁)

(21) 出願番号 特願2006-335130 (P2006-335130)
 (22) 出願日 平成18年12月12日 (2006.12.12)
 (65) 公開番号 特開2008-146503 (P2008-146503A)
 (43) 公開日 平成20年6月26日 (2008.6.26)
 審査請求日 平成21年11月27日 (2009.11.27)

(73) 特許権者 310021766
 株式会社ソニー・コンピュータエンタテインメント
 東京都港区港南1丁目7番1号
 (74) 代理人 100105924
 弁理士 森下 賢樹
 (74) 代理人 100109047
 弁理士 村田 雄祐
 (74) 代理人 100109081
 弁理士 三木 友由
 (74) 代理人 100134256
 弁理士 青木 武司

最終頁に続く

(54) 【発明の名称】分散処理方法、オペレーティングシステムおよびマルチプロセッサシステム

(57) 【特許請求の範囲】

【請求項 1】

制御用のメインプロセッサ、複数の演算用のサブプロセッサ、およびメモリインタフェースがバスで接続されたマルチコアプロセッサと、共有メモリとを含み、前記メインプロセッサと複数の前記サブプロセッサが前記メモリインタフェースを介して前記共有メモリにアクセス可能なマルチプロセッサシステムにおける分散処理方法であって、

各サブプロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させ、前記共有メモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないサブプロセッサに割り当てる実行させることにより、複数のパイプライン処理系で実行される前記特定処理の複数のタスクの内、処理時間が所定の閾値よりも長い各パイプライン処理系の高負荷タスクが異なるサブプロセッサに割り当てられて実行され、その結果、各サブプロセッサの負荷が均一化されることを特徴とする分散処理方法。

【請求項 2】

前記サブプロセッサの数を前記特定処理を構成する前記高負荷タスクの数で除算して得られる値を超えない整数値の数だけ前記パイプライン処理系を動作させることを特徴とする請求項1に記載の分散処理方法。

【請求項 3】

10

20

前記パイプライン処理系で実行される前記特定処理のタスク間でやりとりされるデータの入出力関係を記述した設定ファイルをもとにタスク間の入出力チャネルを構築し、前記入出力チャネルを介したタスク間のストリーム通信を実行することを特徴とする請求項1または2に記載の分散処理方法。

【請求項4】

前記設定ファイルに記述されたタスクの入出力経路に新たなタスクを直列または並列に挿入することにより、前記設定ファイルを動的に変更する手順をさらに含むことを特徴とする請求項3に記載の分散処理方法。

【請求項5】

前記パイプライン処理系で実行される前記特定処理の各タスクは、当該タスクを割り当てられたサブプロセッサによって互いに実行され、各サブプロセッサは、割り当てられたタスクの入力チャネルから入力を受け取って当該タスクを処理し、そのタスクの出力チャネルに実行結果を出力することを特徴とする請求項3に記載の分散処理方法。 10

【請求項6】

制御用のメインプロセッサ、複数の演算用のサブプロセッサ、およびメモリインタフェースがバスで接続されたマルチコアプロセッサと、共有メモリとを含み、前記メインプロセッサと複数の前記サブプロセッサが前記メモリインタフェースを介して前記共有メモリにアクセス可能なマルチプロセッサシステム上で動作するオペレーティングシステムであって、

各サブプロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させる機能と、 20

前記共有メモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないサブプロセッサに割り当てる実行させることにより、複数のパイプライン処理系で実行される前記特定処理の複数のタスクの内、処理時間が所定の閾値よりも長い各パイプライン処理系の高負荷タスクを異なるサブプロセッサに割り当てる実行させ、その結果、各サブプロセッサの負荷を均一化させる機能とを前記マルチプロセッサシステムに実現させることを特徴とするオペレーティングシステム。

【請求項7】

制御用のメインプロセッサ、それぞれがローカルメモリをもつ複数の演算用のサブプロセッサ、およびメモリインタフェースがバスで接続されたマルチコアプロセッサと、共有メモリとを含み、前記メインプロセッサと複数の前記サブプロセッサが前記メモリインタフェースを介して前記共有メモリにアクセス可能なマルチプロセッサシステムであって、

前記複数の演算用のサブプロセッサ上で動作するオペレーティングシステムは、

各サブプロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させる機能と、

前記共有メモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないサブプロセッサの前記ローカルメモリにロードして実行させることにより、複数のパイプライン処理系で実行される前記特定処理の複数のタスクの内、処理時間が所定の閾値よりも長い各パイプライン処理系の高負荷タスクを異なるサブプロセッサに割り当てる実行させ、その結果、各サブプロセッサの負荷を均一化させる機能とを含むことを特徴とするマルチプロセッサシステム。 40

【請求項8】

前記サブプロセッサに割り当たられた各タスクは、前記メインプロセッサを介在せることなく、通信チャネルを介して互いにデータをやりとりしながら実行されることを特徴とする請求項7に記載のマルチプロセッサシステム。

【発明の詳細な説明】

10

20

30

40

50

【技術分野】**【0001】**

この発明は、複数のプロセッサを含むマルチプロセッサシステムにおける分散処理技術に関する。

【背景技術】**【0002】**

CPU (Central Processing Unit) の処理速度の高速化のために、CPUの動作周波数を上げる工夫がなされてきた。動作周波数を上げるためにRISC (Reduced Instruction Set Computer) アーキテクチャが採用され、CPUの各処理ステージの処理速度を均一化して並列度を高めるために深いパイプラインが形成されてきた。また、基板上の配線幅や配線間隔を小さくしてより多くの論理をチップ内に組み込んで高密度化を図るとともに、チップの消費電力を抑える工夫もなされてきた。

【0003】

しかしながら、配線が細かくなつたことでリーク電流が増加し、消費電力が下がらなくなってきた。また、消費電力を下げることができないため、CPUの動作周波数を上げることができず、従来のように動作周波数を上げることでCPUの高速化を図るのは難しくなってきている。

【0004】

また、CPU単体の高速化とは別に、複数のCPUをもつマルチプロセッサシステムによって処理の高速化が図られている。特にゲームのように複数のプロセスが動作するアプリケーションの高速化にはマルチプロセッサシステムが適している。

【0005】

図1(a)、(b)はCPUの高速化の一手法であるパイプライン処理の原理を説明する図である。図1(a)に示すように、CPUでは、一つの命令(Instruction)が、フェッチ(IF; Instruction Fetch)、デコード(ID; Instruction Decode)、レジスタフェッチ(RF; Register Fetch)、実行(EXE; Execution)、メモリアクセス(MEM; Memory Access)、書き戻し(WB; Write Back)といった6つのステージを経て処理される。通常は、一つ前の命令の処理ステージがすべて完了しなければ次の命令の処理を開始することができないが、パイプライン処理では、各ステージの処理機構を独立して動作させることにより、流れ作業のように、一つ前の命令について全処理ステージの実行が完了する前に次の命令を処理し始めることができる。これにより、複数の命令からなるプログラム全体の処理時間の短縮を図ることができる。

【0006】

図1(b)は、6個の命令1~6がパイプライン処理される様子を示している。各ステージが1クロックサイクルで処理されるとすると、命令1は第1サイクルでIFステージに投入され、以降、1サイクル毎に後続のステージに結果が渡されて各ステージの処理が実行される。命令2は、命令1についてIFステージの処理が終わった後、第2サイクルでIFステージに投入され、以降、同様に1サイクル毎に後続の各ステージに送られ、処理される。同様に命令3~6は、それぞれ第3~第6サイクルでIFステージに投入されて各ステージの処理がなされる。第6サイクルにおいては、6つのステージが並列に動作して各命令を処理することになるから、CPUのスループットが最大化される。

【0007】

特許文献1には、プログラムの処理時間を短縮することのできる命令パイプライン処理方法が開示されている。

【特許文献1】特開2000-172502号公報**【発明の開示】****【発明が解決しようとする課題】****【0008】**

マルチプロセッサシステムにおいてプログラムを並列に実行させるためには、プログラムを並列に動作する複数のタスクに分解し、タスクをプロセッサに割り当てるににより

10

20

30

40

50

、並列に実行させる必要がある。プログラムを並列に動作するタスクに分解する作業は自動化が難しく、プログラマの手作業によるしかなく、並列処理のプログラミングはたいへん手間と時間がかかる。また、パイプライン処理は各ステージの処理時間が均一であることが前提となるため、処理時間の異なる複数のタスクからなるプログラムの処理にパイプライン処理の原理をそのまま適用することはできない。

【0009】

本発明はこうした課題に鑑みてなされたものであり、その目的は、マルチプロセッサシステムにおいて分散処理を効率的に実行するための技術を提供することにある。

【課題を解決するための手段】

【0010】

上記課題を解決するために、本発明のある態様は、複数のプロセッサを含むマルチプロセッサシステムにおける分散処理方法である。この方法では、各プロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させ、メインメモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないプロセッサに割り当てて実行させることにより、複数のパイプライン処理系で実行される前記特定処理の複数のタスクの内、処理時間が所定の閾値よりも長い高負荷タスクが異なるプロセッサに割り当てられて実行される。

10

【0011】

本発明の別の態様は、オペレーティングシステムである。このオペレーティングシステムは、複数のプロセッサを含むマルチプロセッサシステム上で動作するオペレーティングシステムであって、各プロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させる機能と、メインメモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないプロセッサに割り当てて実行させる機能とを前記マルチプロセッサシステムに実現させる。

20

【0012】

本発明のさらに別の態様は、マルチプロセッサシステムである。このマルチプロセッサシステムは、制御用のメインプロセッサと、それぞれがローカルメモリをもつ複数の演算用のサブプロセッサと、共有メモリとを含む。前記複数の演算用のサブプロセッサ上で動作するオペレーティングシステムは、各サブプロセッサの計算資源を時分割して複数のタスクに割り当てるにより、複数のタスクが並列に実行されるマルチタスク環境において、タスクの実行結果を別のタスクに与えることにより、負荷の異なる複数のタスクからなる特定処理を実行するためのパイプライン処理系を構築し、当該パイプライン処理系を複数動作させる機能と、前記共有メモリにコンテキストが退避されて実行可能状態にあるタスクをいずれのタスクも実行していないサブプロセッサの前記ローカルメモリにロードして実行させる機能とを含む。

30

【0013】

なお、以上の構成要素の任意の組合せ、本発明の表現を方法、装置、システム、コンピュータプログラム、データ構造、記録媒体などの間で変換したものもまた、本発明の態様として有効である。

【発明の効果】

【0014】

本発明によれば、マルチプロセッサシステムにおいて分散処理を効率的に実行することができる。

【発明を実施するための最良の形態】

40

50

【0015】

実施の形態1

図2は、実施の形態1に係るマルチプロセッサシステム100の構成図である。マルチコアプロセッサ110は、複数のプロセッサを一つのパッケージに集積したものであり、一つのプロセッサユニット(PU)10、複数(ここでは8個)のサブプロセッサユニット(SPU)20a～20h、メモリインタフェース40、I/Oインタフェース50がリングバス30で接続された構成である。PU10およびSPU20a～20hは、メモリインタフェース40を介してメインメモリ120にアクセスすることができる。I/Oインタフェース50は、外部デバイスとのインタフェースである。

【0016】

10

PU10はマルチプロセッサシステム100全体を制御するためのメインプロセッサであり、キャッシュメモリ12をもつ。SPU20a～20hは演算用のプロセッサであり、内部にローカルメモリ22a～22hをもつ。SPU20a～20hは、DMA(Direct Memory Access)コントローラ(DMAC)24a～24hをもち、DMAC24a～24hは、メインメモリ120とローカルメモリ22a～22hの間でデータをDMA転送する。以下、SPU20a～20hを総称するときは添え字a～hを省略して単にSPU20と表記する。また、8個のSPU20a～20hをSPU1～SPU8と表記することもある。

【0017】

20

各SPU20上では分散オペレーティングシステムが動作する。分散オペレーティングシステムにおけるカーネルプログラムが複数のSPU20上で協調して動作し、各SPU20上で動作するユーザプログラムを時分割で切り替えながら実行する。プログラムの時分割切り替えにはPU10が介在しないため、非常に小さいオーバーヘッドでプログラムの切り替えが可能である。

【0018】

SPU20上で動作するカーネルプログラムは、SPU20の計算資源、具体的にはCPUタイムを時分割して複数のタスクに割り当てることにより、SPU20上で仮想的に複数のタスクが並列に実行されるマルチタスク環境を実現する。一つのユーザプログラムを複数のSPU20で実行することも可能であり、複数のSPU20は、PU10を介さずに互いに直接通信することができる。

30

【0019】

ユーザプログラムは複数のタスクに分割され、メインメモリ120に保持される。各SPU20は、メインメモリ120に保持された実行可能状態にあるタスクをローカルメモリ22にDMA転送し、そのタスクを実行する。各SPU20は、時分割されたCPUタイムをタスクの実行に割り当て、タスクを実行し、割り当てられたCPUタイムを消費すると、タスクをローカルメモリ22からメインメモリ120にDMA転送して退避する。タスクのコンテキストの退避処理はカーネルプログラムが自動的に実行するため、プログラマは意識する必要はない。

【0020】

40

タスク間でデータのやりとりをするために、タスクを実行するSPU20は互いに通信する。SPU20間の通信帯域は非常に広帯域であるため、一つのプログラムをタスクに分割して複数のSPU20で分散処理してもレーテンシが問題となることはなく、むしろ複数のSPU20で並列にタスクを処理することによりスループットが向上して、プログラムの処理時間が短くなる。また、タスクはメインメモリ120とローカルメモリ22の間をDMA転送され、プロセッサは転送に関わらないため、プロセッサに負荷がかからない。

【0021】

複数のSPU20上で実現されるマルチタスク環境において、タスクの実行結果を別のタスクの入力として与えることにより、複数のタスクからなる特定処理を実行するためのパイプライン処理系が構築される。

50

【0022】

パイプライン処理系で実行される特定処理のタスク間でデータストリームをやりとりするための入出力チャネルが構築され、入出力チャネルを介したタスク間のストリーム通信が実現される。各タスクは、当該タスクを割り当てられたプロセッサによって互いに非同期に実行され、各プロセッサは、割り当てられたタスクの入力チャネルから入力を受け取って当該タスクを処理し、そのタスクの出力チャネルに実行結果を出力する。

【0023】

図3は、タスクがスケジュールされてプロセッサに割り当てられる様子を説明する図である。ここでは4つのSPU1～SPU4にタスクが割り当てられる場合を示す。特定の処理が複数のタスクに分割され、これらのタスクはカーネルに実装されたタスクスケジューラによりスケジュールされ、いずれかのSPU20に割り当てられて実行される。

10

【0024】

いずれかのSPU20によって実行されているタスクは、実行中状態（Running状態）にあるという。SPU20が実行中状態にあるタスクを解放すると、そのタスクのコンテキストはメインメモリ120に退避される。メインメモリ120にコンテキストが退避されたタスクは、実行可能状態（Ready状態）または待ち状態（Waiting）にある。

【0025】

実行可能状態とは、タスクの入力チャネルに入力値が与えられており、タスクがいつでもSPU20によって実行可能な状態にあることである。実行可能状態にあるタスクを割り当て可能なSPU20が確保されると、そのタスクは実行中状態に遷移する。

20

【0026】

待ち状態とは、タスクの実行に必要な条件が満たされていないため、タスクはまだ実行可能ではなく、待機している状態である。たとえば、タスクの入力チャネルにまだ入力値が与えられていない場合は、入力値が与えられるまではタスクは待ち状態にある。待ち状態にあるタスクの実行に必要な条件が満たされると、そのタスクは実行可能状態に遷移する。

【0027】

図3では、SPU1、SPU2およびSPU4のローカルメモリ22a、22b、22dにはタスクがロードされて実行されている。SPU3にはタスクが割り当てられていない。メインメモリ120の待ち行列には、待ち状態のタスクが1つ、実行可能なタスクが2つがキューイングされている。タスクスケジューラは、メインメモリ120に退避されている最初の実行可能なタスクをSPU3に割り当てる。割り当てられたタスクはSPU3のDMA24cによりローカルメモリ22cにDMA転送される。

30

【0028】

メインメモリ120に退避されたタスクは待ち行列にキューイングされ、ラウンドロビン方式などによりスケジュールされ、タスクが割り当てられていないアイドル状態のSPU20に割り当てられる。

【0029】

図4(a)、(b)は、特定処理を複数のタスクに分割してパイプライン処理する様子を説明する図である。マルチコアプロセッサ110内のSPUが4個(SPU1～SPU4)であるとする。特定処理は4つのタスクA、B、C、Dに分割され、それぞれのタスクの処理時間Tは共通であるとする。

40

【0030】

図4(a)には、6つの特定処理(P1～P6)がパイプライン処理される様子が示されている。各処理P1～P6を構成する4つのタスクA～Dには、ジョブ番号の添え字をつけています。処理P1について、時刻T0においてタスクA1が実行され、タスクA1の処理結果はタスクB1に渡され、時刻T1においてタスクB1が実行される。以降、同様にして処理結果を引き継ぎながら、時刻T2においてタスクC1が実行され、時刻T3においてタスクD1が実行される。

50

【0031】

処理P2については、時刻T0においてジョブ1のタスクA1が実行された後、時刻T1においてタスクA2が実行され、以降、時刻T2、T3、T4において、タスクB2、C2、D2が順次実行される。

【0032】

以下、同様に、処理P3～P6について、時刻T2～T5においてタスクA3～A6の実行が開始され、以降、処理時間T毎に後続のタスクが順次実行される。

【0033】

図4(b)は、時刻T3、T4、T5における各SPUのタスク割り当て状況を示す。タスクAはSPU1、タスクBはSPU2、タスクCはSPU3、タスクDはSPU4にそれぞれ割り当てられたとする。その割り当てのもとで処理P1～P6が図4(a)に示すようにパイプライン処理されたとする。

【0034】

時刻T3では、SPU1にはタスクA4、SPU2にはタスクB3、SPU3にはタスクC2、SPU4にはタスクD1がそれぞれ割り当てられる。

時刻T4では、SPU1にはタスクA5、SPU2にはタスクB4、SPU3にはタスクC3、SPU4にはタスクD2がそれぞれ割り当てられる。

時刻T5では、SPU1にはタスクA6、SPU2にはタスクB5、SPU3にはタスクC4、SPU4にはタスクD3がそれぞれ割り当てられる。

このように、時刻T3～T5の間、4つのSPU1～SPU4には一つずつタスクが割り当てられているから、負荷の比率は1:1:1:1である。

【0035】

図4(a)、(b)の例では、各タスクの処理時間が同じであったが、これは特殊な場合であり、一般には特定処理は負荷の異なる複数のタスクに分割される。

【0036】

図5(a)、(b)は、負荷の異なる複数のタスクから構成される特定処理がパイプライン処理される様子を説明する図である。特定処理は4つのタスクA、B、C、Dに分割される。タスクCの処理時間は、タスクA、B、Dの処理時間Tの3倍であるとする。

【0037】

図5(a)には、8つの特定処理P1～P8がパイプライン処理される様子が示されている。各処理は処理時間Tだけずらして投入される。タスクAはSPU1、タスクBはSPU2、タスクCはSPU3、タスクDはSPU4にそれぞれ割り当てられたとする。タスクCの処理時間が他のタスクよりも長いため、スループットが最大化される時刻T5～T7において、特定のSPUにタスクCが複数個割り当てられることになる。

【0038】

図5(b)は、時刻T5、T6、T7における各SPUのタスク割り当て状況を示す。

時刻T5では、SPU1にはタスクA6、SPU2にはタスクB5、SPU3にはタスクC4、C3、C2、SPU4にはタスクD1がそれぞれ割り当てられる。

時刻T6では、SPU1にはタスクA7、SPU2にはタスクB6、SPU3にはタスクC5、C4、C3、SPU4にはタスクD2がそれぞれ割り当てられる。

時刻T7では、SPU1にはタスクA8、SPU2にはタスクB7、SPU3にはタスクC6、C5、C4、SPU4にはタスクD3がそれぞれ割り当てられる。

このように、時刻T5～T7の間、SPU1、SPU2、SPU4には一つずつタスクが割り当てられているが、SPU3には3つのタスクが割り当てられている。したがって、SPU1～SPU4の負荷の比率は1:1:3:1になる。

【0039】

このように、特定処理を構成する各タスクの処理時間が異なる場合に、パイプライン処理を実行すると、高負荷のタスクが割り当てられたSPUに負荷が集中してしまい、マルチコアプロセッサ110の演算処理のスループットを上げることができなくなる。そこで本実施の形態では、特定処理を実行するパイプライン処理系を複数動作させ、各パイプラ

10

20

30

40

50

イン処理で実行される特定処理の高負荷タスクを異なる SPU に割り当てることで負荷分散を図る。

【0040】

図 6 (a) は、1つのパイプライン処理系をマルチプロセッサシステム 100 で動作させた場合のタスクのプロセッサへの割り当てを説明する図である。タスク A、B、C、D を順次実行するパイプライン処理系 150 を 4 つの SPU 1 ~ SPU 4 で実行する場合、たとえば、タスク A は SPU 1 に、タスク B は SPU 2 に、タスク C は SPU 3 に、タスク D は SPU 4 にそれぞれ割り当てられる。したがって、タスク C の処理時間が長い場合、SPU 3 に負荷が集中する。

【0041】

図 6 (b) は、複数のパイプライン処理系をマルチプロセッサシステム 100 で動作させた場合のタスクのプロセッサへの割り当てを説明する図である。4 つのパイプライン処理系 #1 ~ #4 (符号 151 ~ 154) を並列に実行させるとする。

【0042】

4 つのパイプライン処理系 #1 ~ #4 のそれぞれのタスク A ~ D のすべてがタスクスケジューリングの対象となる。パイプライン処理系 #1 の高負荷タスク C が SPU 3 に割り当てられたとする。このとき、SPU 3 は高負荷タスク C の実行に時間がかかり、他のタスクのためにアイドル状態になるまでに時間がかかる。したがって、パイプライン処理系 #2 の高負荷タスク C は同じ SPU 3 に割り当てられることはなく、別の SPU、たとえば SPU 4 に割り当てられる。

【0043】

同様に、高負荷タスク C を割り当てられた SPU 4 はアイドル状態になるまでに時間を要するため、パイプライン処理系 #3 の高負荷タスク C は SPU 3、SPU 4 には割り当てることができず、それ以外の SPU、たとえば SPU 1 に割り当てられる。さらに、パイプライン処理系 #4 の高負荷タスク C は、既に高負荷タスク C を割り当てられた SPU 3、SPU 4、SPU 1 を避けて、SPU 2 に割り当てられる。このように、4 つのパイプライン処理系 #1 ~ #4 を動作させれば、各パイプライン処理系の高負荷タスク C を異なる SPU に割り当てるができるようになり、特定の SPU に負荷が集中するのを避けることができる。

【0044】

なお、マルチプロセッサシステム 100 において、タスクの SPU 20 への割り当ては、PU 10 が行うのではなく、各 SPU 20 が自律的にメインメモリ 120 の待ち行列から実行可能なタスクを取り出すことで行われる。したがって、図 6 (b) の 4 つのパイプライン処理系 #1 ~ #4 のそれぞれの高負荷タスク C の SPU 1 ~ SPU 4 への割り当ては事前に決まっているのではない。SPU 1 ~ SPU 4 がそれぞれ、待ち行列から実行可能なタスクを取り出して実行することで、結果的に、4 つの高負荷タスク C が 4 つの SPU 1 ~ SPU 4 に分散されて割り当てられることになる。

【0045】

図 7 (a) ~ (e) は、複数のパイプライン処理系を実行するマルチプロセッサシステム 100 において、各プロセッサのタスク割り当て状況と全プロセッサの負荷比率を説明する図である。

【0046】

図 7 (a) は、図 6 (b) のパイプライン処理系 #1 について、時刻 T5、T6、T7 における各 SPU のタスク割り当て状況を示す。これは図 4 (b) のタスク割り当て状況と同じである。図 7 (b) は、パイプライン処理系 #2 について、時刻 T5、T6、T7 における各 SPU のタスク割り当て状況を示す。パイプライン処理系 #2 ではタスク C は SPU 4 に割り当てられるから、SPU 4 にタスク C が 3 つ割り当てられることになる。同様に、図 7 (c) に示すように、パイプライン処理系 #3 については SPU 1 にタスク C が 3 つ割り当てられ、図 7 (d) に示すように、パイプライン処理系 #4 については SPU 2 にタスク C が 3 つ割り当てられる。

10

20

30

40

50

【0047】

図7(e)は、SPU1～SPU4の負荷比率を説明する図である。図7(a)～図7(d)に示したパイプライン処理系#1～#4のタスク割り当てにより、高負荷タスクCはSPU1～SPU4に分散され、負荷が均一化されるから、SPU1～SPU4の負荷比率は1:1:1:1になる。

【0048】

以上述べたように、特定処理を実行するためのパイプライン処理系を複数動作させれば、各パイプライン処理系の高負荷タスクが異なるSPUに分散されて割り当てられることになり、マルチプロセッサシステム100のスループットを向上させることができる。

【0049】

図1(a)、(b)で説明したように、CPUは、命令のフェッチ、デコードなどの専用の回路をもち、これらの専用回路を並列に動作させてパイプライン処理を実行する。パイプライン処理の各ステージの処理時間を等しくするために、これらの専用回路における処理時間は等しくなるように(たとえば、1クロックサイクルで実行が完了するように)設計されている。このように、一般的なCPUにおけるパイプライン処理では、命令を実行時間が同じサイクル数で実行できるステージに細分化し、並列処理を行うことでスループットの向上を図っている。パイプライン処理を実現するためには、各ステージの処理サイクル数が揃うように機能を分割する必要がある。

【0050】

それに対して、マルチプロセッサシステム100の複数のSPU20は、汎用のプロセッサであるため、どんな処理でも実行可能である。したがって、特定処理を複数の異なるタスクに分割してパイプライン処理するとき、各SPU20にはどのタスクを割り当ててもかまわない。したがって、特定処理を実行するためのパイプライン処理系を複数動作させて、各パイプライン処理系の高負荷タスクをアイドル状態にある異なる複数のSPU20に分散させて割り当てることが可能である。それにより、特定のSPU20に負荷が集中する事態を避けることができ、全体のスループットを上げることができる。

【0051】

SPU20に割り当てるタスクの負荷を均一にする必要はないから、特定処理が負荷の異なるタスクに分割されるという一般的な状況に対応することができ、並列プログラミングが容易になる。また、タスクの数とタスクを割り当てるプロセッサの数は独立しているため、プログラマは、物理的なプロセッサの構成を気にする必要もない。

【0052】

特定処理を構成する高負荷タスクが2つあり、この特定処理を実行するためパイプライン処理系を4つ動作させるとすると、高負荷タスクが全体で8個あることになる。高負荷タスクを割り当てるプロセッサ数も8個あると、各プロセッサに高負荷タスクを1つずつ割り当てるため、好都合である。一般にはプロセッサの数はシステムによって決まっているから、特定処理を構成する高負荷タスクの数が与えられると、プロセッサの数を特定処理の高負荷タスクの数で除算して得られる値を超えない整数値の数だけパイプライン処理系を動作させればよい。

【0053】

実施の形態2

実施の形態2は、実施の形態1のマルチプロセッサシステム100におけるパイプライン処理をビデオやオーディオからメタ情報を抽出する処理に応用したものである。

【0054】

図8は、実施の形態2に係るメタ情報処理装置200の機能構成図である。同図は機能に着目したブロック図を描いており、これらの機能ブロックはハードウェアのみ、ソフトウェアのみ、またはそれらの組合せによっていろいろな形で実現することができる。すなわち、これらの機能構成ブロックの少なくとも一部は、図2で説明したマルチプロセッサシステム100のハードウェア構成により実現され、ハードウェア構成で実現されない機能ブロックは、メインメモリ120にロードされたプログラムをPU10またはSPU2

10

20

30

40

50

0が実行することにより実現される。

【0055】

メタ情報処理装置200は、ビデオやオーディオなどを含むコンテンツからメタ情報を抽出する装置であり、メタ情報抽出ブロック210と、ジョブデータベース220と、メタ情報データベース230とを有する。

【0056】

メタ情報抽出ブロック210は、ジョブデータベース220からジョブを取り出し、コンテンツに含まれるビデオ、オーディオ、テキストからそれぞれメタ情報を抽出し、抽出したメタ情報をメタ情報データベース230に登録する。メタ情報抽出ブロック210によるメタ情報抽出処理は、複数のタスクに分解される。メタ情報抽出処理を構成するタスクをマルチプロセッサシステム100の複数のSPU20に割り当て、パイプライン処理によって実行する。

【0057】

メタ情報抽出ブロック210の各機能構成を説明する。リクエスト処理部240は、ジョブデータベース220からジョブを取り出し、ファイル転送部242に与える。ファイル転送部242はジョブによって指定されたファイルをデータベースから読み出し、ファイル解析・分割部244に与える。ファイル解析・分割部244は、ファイルを解析し、オーディオデータをオーディオデコーダ250に、テキストデータをテキスト変換部264に、画像データを画像デコーダ252に与える。

【0058】

オーディオデコーダ250は、オーディオデータを復号し、12音解析部260に与え得る。画像デコーダ252は、画像データを復号し、画像解析部262に与える。

【0059】

12音解析部260は、復号されたオーディオデータを12音解析技術により周波数分析して特徴量を抽出する。12音解析部260は、抽出した特徴量にもとづいてオーディオのメタデータを生成し、データベース登録部270に与える。

【0060】

テキスト変換部264は、テキストデータからキーワードなどを抽出することでテキストのメタデータを生成し、データベース登録部270に与える。

【0061】

画像解析部262は、復号された画像データを画像処理して特徴量を抽出し、抽出した特徴量にもとづいてメタデータを生成する。生成された画像のメタデータはデータベース登録部270に与えられる。

【0062】

データベース登録部270は、オーディオのメタデータ、テキストのメタデータ、画像のメタデータをコンテンツのメタ情報としてメタ情報データベース230に登録する。データベース登録部270による登録が終わると、結果通知部272は、ジョブデータベース220にジョブの完了を通知する。

【0063】

メタ情報抽出ブロック210の各機能構成がタスクに対応し、各機能構成の入出力がそのままタスクの入出力チャネルに対応する。

【0064】

図9は、メタ情報抽出パイプライン処理系300を1つだけ動作させてメタ情報抽出処理を実行した場合の各タスクのプロセッサへの割り当てを説明する図である。このパイプライン処理系300で実行されるメタ情報抽出処理は、リクエスト処理部240、ファイル転送部242、ファイル解析・分割部244、オーディオデコーダ250、12音解析部260、データベース登録部270、および結果通知部272の7つのタスクからなり、それぞれのタスクは、SPU1～SPU7に割り当てられ、SPU8はアイドル状態である。12音解析部260がもっとも負荷の高いタスクであるため、このタスクを割り当てられたSPU5は稼働率が100%になる。オーディオデコーダ250は次に負荷の高

10

20

30

40

50

いタスクであり、このタスクを割り当てられた SPU 4 は稼働率が 50 % となる。それ以外の処理時間が短いタスクを割り当てられた SPU の稼働率は極めて低いものとなっている。

【 0 0 6 5 】

図 10 (a) ~ (d) は、4 つのメタ情報抽出パイプライン処理系 #1 ~ #4 (符号 300a ~ 300d) を動作させてメタ情報抽出処理を実行した場合の高負荷タスクのプロセッサへの割り当てを説明する図である。各パイプライン処理系 #1 ~ #4 において、12 音解析部 260a ~ 260d と画像解析部 262a ~ 画像解析部 262d が高負荷タスクであり、合計すると 8 個の高負荷タスクがある。

【 0 0 6 6 】

10

8 つの SPU 1 ~ SPU 8 は、処理時間の短いタスクを実行していることもあるが、そのタスクの実行が終わると、アイドル状態になる。8 個の高負荷タスクは、アイドル状態の SPU に割り当てられていくことになり、いったん高負荷タスクが割り当てられた SPU は再びアイドル状態になるまで時間がかかる。結果的に、8 個の SPU に均等に 8 個の高負荷タスクが割り当てられることになる。

【 0 0 6 7 】

同図に示すように、パイプライン処理系 #1 の 12 音解析部 260a は SPU 1 に、画像解析部 262a は SPU 2 にそれぞれ割り当てられる。パイプライン処理系 #2 の 12 音解析部 260b は SPU 3 に、画像解析部 262b は SPU 4 にそれぞれ割り当てられる。パイプライン処理系 #3 の 12 音解析部 260c は SPU 5 に、画像解析部 262c は SPU 6 にそれぞれ割り当てられる。パイプライン処理系 #4 の 12 音解析部 260d は SPU 7 に、画像解析部 262d は SPU 8 にそれぞれ割り当てられる。これにより、SPU 1 ~ SPU 8 はすべて稼働率 100 % になり、マルチプロセッサシステム 100 の処理スループットが最大化する。

20

【 0 0 6 8 】

図 11 (a) ~ (c) は、メタ情報抽出パイプライン処理系のバリエーションを説明する図である。図 11 (a) のパイプライン処理系 A (符号 300) は、オーディオのメタデータを抽出するタスクで構成される。これに対して、図 11 (b) のパイプライン処理系 B (符号 310) は、オーディオ解析と画像解析を行い、オーディオのメタデータと画像のメタデータを抽出するタスクで構成される。図 11 (a) と比べると、画像デコーダ 252 と画像解析部 262 のタスクが、オーディオデコーダ 250 と 12 音解析部 260 のタスクに対して並列に追加されている。

30

【 0 0 6 9 】

このようなタスクの追加は、実行ファイルを変更することなく、タスクの接続関係を記述した設定ファイルを変更するだけで行うことができ、設定ファイルの変更により動的にパイプライン処理系のタスクを自由自在に変更できる。この設定ファイルはたとえば、XML (Extensible Markup Language) により記述される。設定ファイルは、タスク間でやりとりされるデータの入出力関係を記述したものであり、これをもとにタスク間の入出力チャネルが構築され、入出力チャネルを介してタスク間のストリーム通信が実行される。

40

【 0 0 7 0 】

たとえば、タスク A の出力をタスク B の入力に接続するには、「タスク A の出力チャネルをタスク B の入力チャネルに接続する」という記述を設定ファイルに設ければよい。図 11 (a) のパイプライン処理系 A の例で言えば、オーディオデコーダ 250 の出力チャネルを 12 音解析部 260 の入力チャネルに接続するという記述を設けることになる。

【 0 0 7 1 】

図 11 (b) のパイプライン処理系 B を構築するには、図 11 (a) のパイプライン処理系 A の設定ファイルを一部変更するだけでよい。ファイル解析・分割部 244 の出力チャネルをオーディオデコーダ 250 の入力チャネルだけでなく、画像デコーダ 252 の入力チャネルにもつなげる。そして、画像デコーダ 252 の出力チャネルを画像解析部 26

50

2の入力チャネルにつなげる。さらに、画像解析部262の出力チャネルをデータベース登録部270の入力チャネルにつなげる。以上の設定ファイルの変更により、図11(a)の処理系Aが図11(b)の処理系Bに変更される。

【0072】

図11(c)は、図11(a)の12音解析部260を12音低レベル解析部260aと12音高レベル解析部260bにタスク分解したパイプライン処理系Cを示す。12音低レベル解析部260aは、人間の聴覚能力に応じた低レベルの特徴を解析するものであり、一例としてビート検出、コード進行検出、楽曲構造解析、キー(調)検出などを行い、ビート、コード、曲構造などの低レベルの特徴量を抽出する。12音高レベル解析部260bは、人間の知的な理解能力に応じた高レベルの特徴を解析するものであり、一例として統計解析、ジャンル判別、楽器音検出、ムード検出、音質検出などを行い、ジャンル、楽器音、ムード、音質などの高レベルの特徴量を抽出する。

10

【0073】

12音低レベル解析部260aと12音高レベル解析部260bは両方とも高負荷タスクであるが、このように2つのタスクに分解すれば、別々のSPUに割り当てられることになるから、負荷の分散を図ることができる。図11(c)のパイプライン処理系Cを構築するには、図11(a)のパイプライン処理系Aの設定ファイルを一部変更する。図11(a)のオーディオデコーダ250から12音解析部260を経てデータベース登録部270に至るデータストリームの経路において、12音解析部260を削除し、12音低レベル解析部260aと12音高レベル解析部260bを直列につなぎ直す。

20

【0074】

このようにパイプライン処理系の新たにタスクを追加したり、タスクの構成を変更する際、一般にはタスクの負荷を考慮してパイプライン処理系を適切に設計する必要がある。しかし、本実施の形態では、複数のパイプライン処理系を動作させてタスクをプロセッサに割り当てる構成であり、負荷が自動的に均一化されて複数のプロセッサに分散される。そのため、プログラマはタスクの負荷に注意を払うことなく、単に設定ファイルを変更するだけで簡単にパイプライン処理系を設計し直すことができ、プログラミングに柔軟性をもたらすことができる。

【0075】

図12は、4つのメタ情報抽出パイプライン処理系#1～#4が互いに独立にデータベースにアクセスしながら並列に動作する構成を説明する図である。マルチプロセッサシステム100の複数のSPU20は、PU10の介在なしに、互いに独立に自律的に動作し、必要に応じてタスク間通信により協調動作する。PU10が集中管理すると、PU10に負荷が集中してボトルネックになるからである。

30

【0076】

図12に示すように、各パイプライン処理系#1～#4のリクエスト処理部240a～240dは互いに独立にジョブデータベース220をポーリングし、未処理のジョブがジョブデータベース220に登録されると、SPU20はその未処理ジョブを開始する。

【0077】

各パイプライン処理系#1～#4のデータベース登録部270a～270dは互いに独立にメタ情報データベース230にメタデータを登録する。また、結果通知部272a～272dも、互いに独立にジョブデータベース220にジョブの終了を通知する。

40

【0078】

ジョブの投入とジョブの実行結果の登録は同期させる必要がないため、制御用のPU10が介在して同期を取ったり、排他制御をすることはない。そのため、メタ情報抽出処理を複数のタスクに分割し、複数のSPU20がタスクを自律的に実行するという複雑な処理系を構成しても、制御用のPU10に負荷が集中してボトルネックとなる事態を避けることができる。

【0079】

以上、本発明を実施の形態をもとに説明した。実施の形態は例示であり、それらの各構

50

成要素や各処理プロセスの組合せにいろいろな変形例が可能であること、またそうした変形例も本発明の範囲にあることは当業者に理解されるところである。

【図面の簡単な説明】

【0080】

【図1】図1(a)、(b)はCPUの高速化の一手法であるパイプライン処理の原理を説明する図である。

【図2】実施の形態1に係るマルチプロセッサシステムの構成図である。

【図3】タスクがスケジュールされてプロセッサに割り当てる仕組みを説明する図である。

【図4】図4(a)、(b)は、特定処理を複数のタスクに分割してパイプライン処理する様子を説明する図である。 10

【図5】図5(a)、(b)は、負荷の異なる複数のタスクから構成される特定処理がパイプライン処理される様子を説明する図である。

【図6】図6(a)は、1つのパイプライン処理系をマルチプロセッサシステムで動作させた場合のタスクのプロセッサへの割り当てを説明する図であり、図6(b)は、複数のパイプライン処理系をマルチプロセッサシステムで動作させた場合のタスクのプロセッサへの割り当てを説明する図である。

【図7】図7(a)～(e)は、複数のパイプライン処理系を実行するマルチプロセッサシステムにおいて、各プロセッサのタスク割り当て状況と全プロセッサの負荷比率を説明する図である。 20

【図8】実施の形態2に係るメタ情報処理装置の機能構成図である。

【図9】メタ情報抽出パイプライン処理系を1つだけ動作させてメタ情報抽出処理を実行した場合の各タスクのプロセッサへの割り当てを説明する図である。

【図10】図10(a)～(d)は、4つのメタ情報抽出パイプライン処理系#1～#4(符号300a～300d)を動作させてメタ情報抽出処理を実行した場合の高負荷タスクのプロセッサへの割り当てを説明する図である。

【図11】図11(a)～(c)は、メタ情報抽出パイプライン処理系のバリエーションを説明する図である。

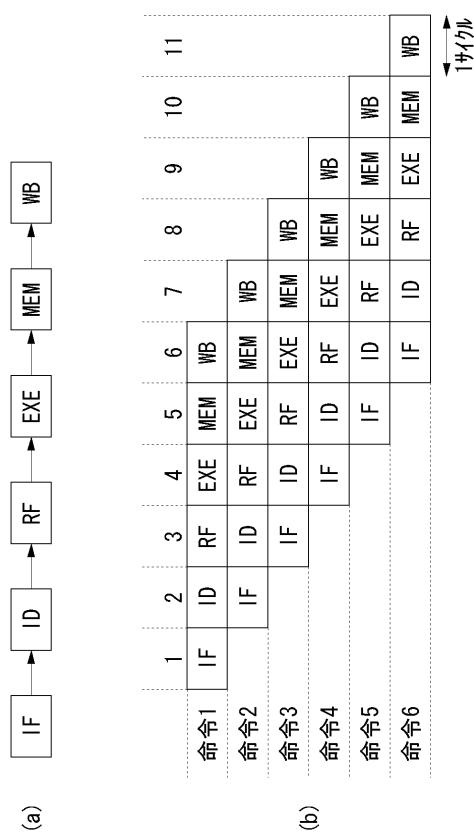
【図12】4つのメタ情報抽出パイプライン処理系が互いに独立にデータベースにアクセスしながら並列に動作する構成を説明する図である。 30

【符号の説明】

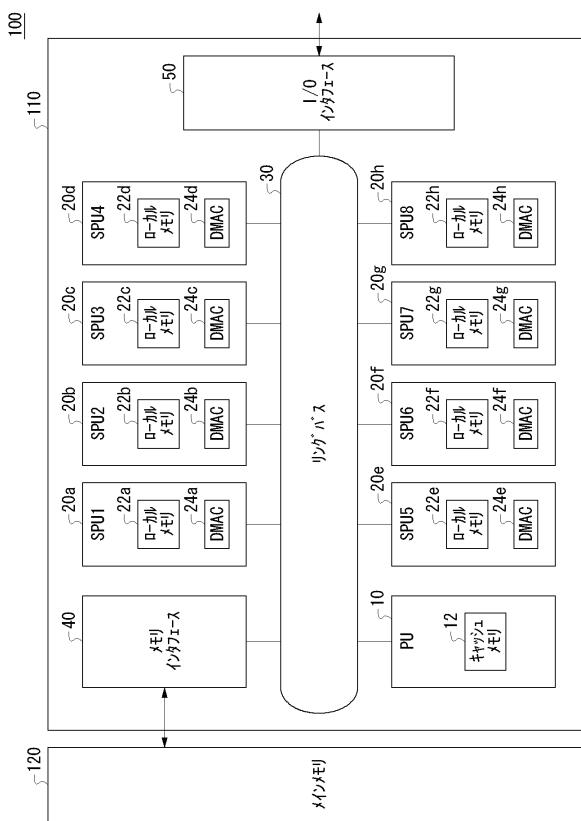
【0081】

1 プロセッサユニット、 20 サブプロセッサユニット、 22 ローカルメモリ、 24 DMAコントローラ、 30 リングバス、 40 メモリインターフェース、 50 I/Oインターフェース、 100 マルチプロセッサシステム、 110 マルチコアプロセッサ、 120 メインメモリ、 150 パイプライン処理系、 200 メタ情報処理装置、 210 メタ情報抽出ブロック、 220 ジョブデータベース、 230 メタ情報データベース、 240 リクエスト処理部、 242 ファイル転送部、 244 ファイル解析・分割部、 250 オーディオデコーダ、 252 画像デコーダ、 260 12音解析部、 262 画像解析部、 264 テキスト変換部、 270 データベース登録部、 272 結果通知部、 300 メタ情報抽出パイプライン処理系。 40

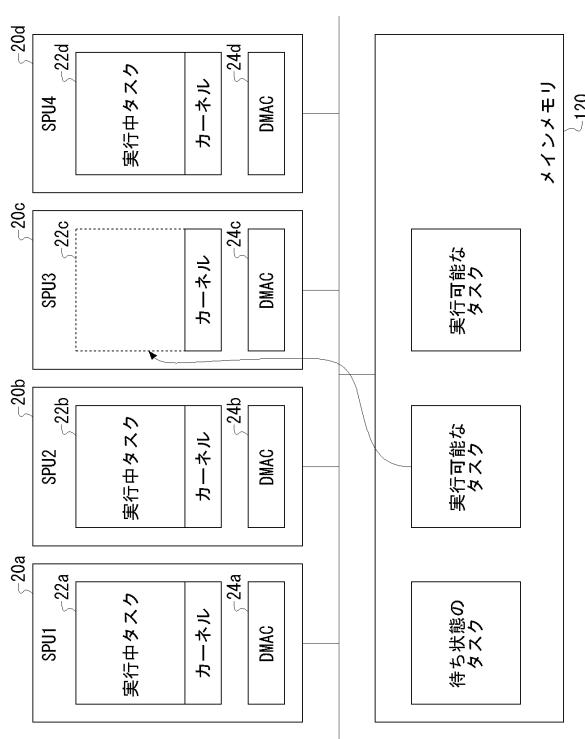
【図1】



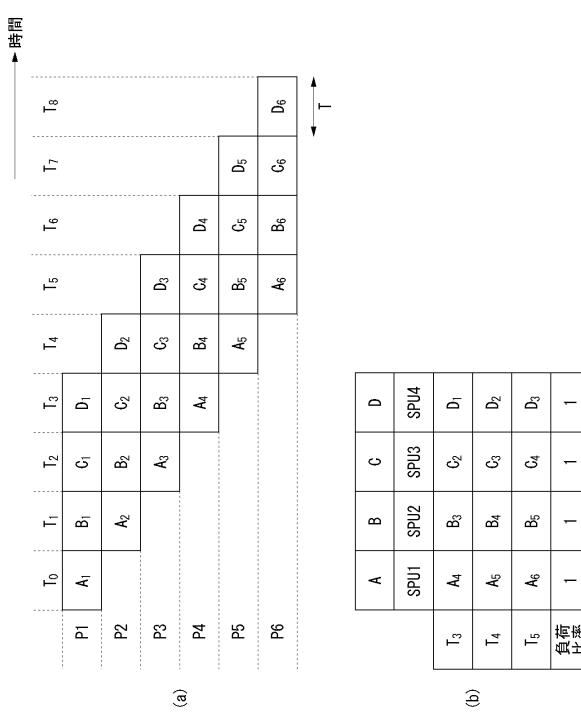
【図2】



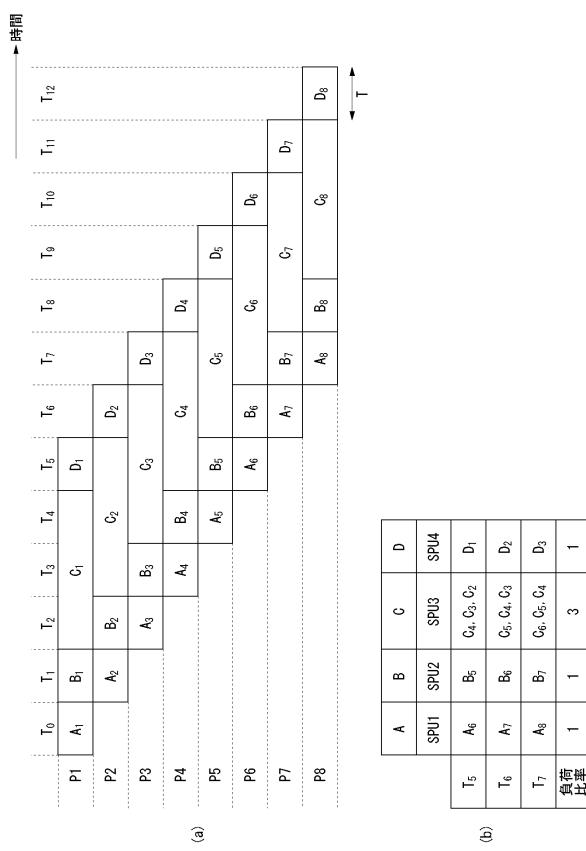
【図3】



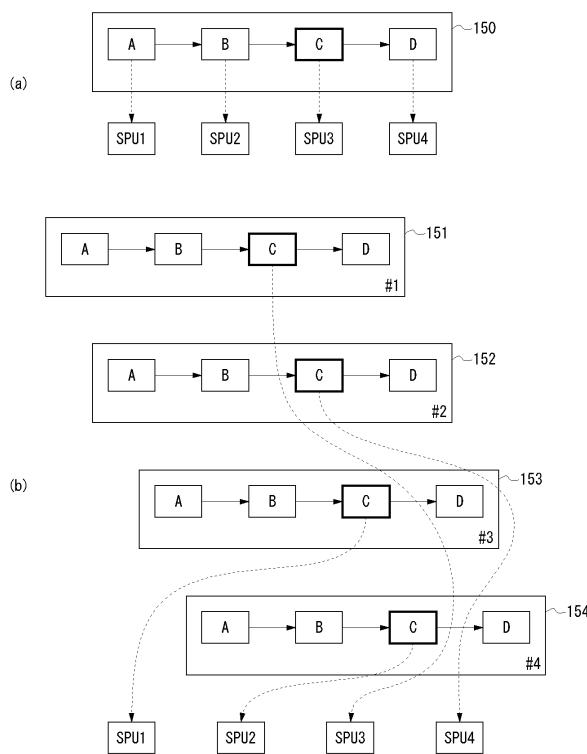
【図4】



【図5】



【図6】



【図7】

(a) Task mapping: T₅ to SPU1-SPU2-SPU3-SPU4. T₆ to SPU2-SPU3-SPU4. T₇ to SPU3-SPU4.

(b) Task mapping: T₅ to SPU1-SPU2-SPU4. T₆ to SPU2-SPU3-SPU4. T₇ to SPU3-SPU4.

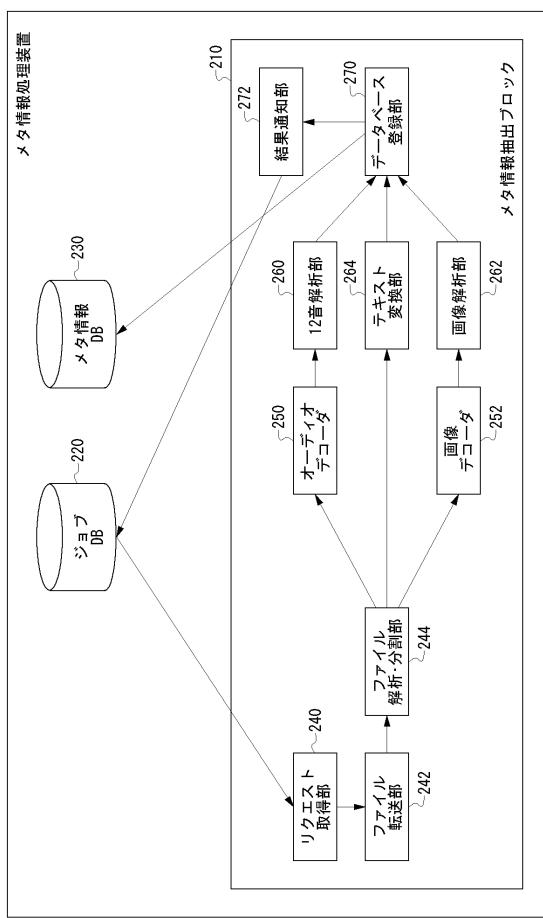
(c) Task mapping: T₅ to SPU3-SPU2-SPU1-SPU4. T₆ to SPU2-SPU1-SPU4. T₇ to SPU1-SPU4.

(d) Task mapping: T₅ to SPU1-SPU3-SPU2-SPU4. T₆ to SPU2-SPU3-SPU4. T₇ to SPU3-SPU4.

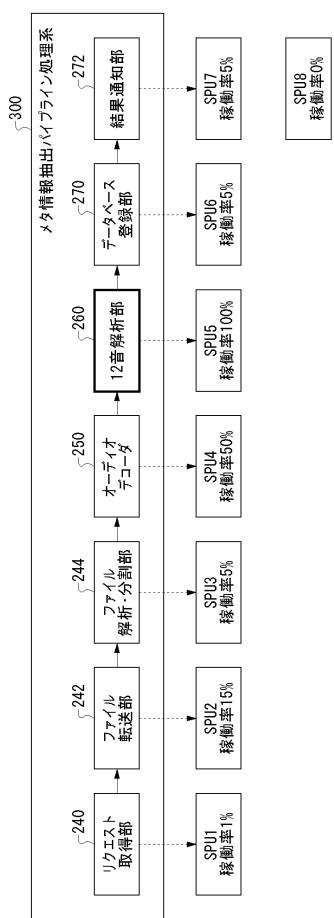
(e) Overall load distribution table:

	SPU1	SPU2	SPU3	SPU4
負荷比率	1	1	1	1

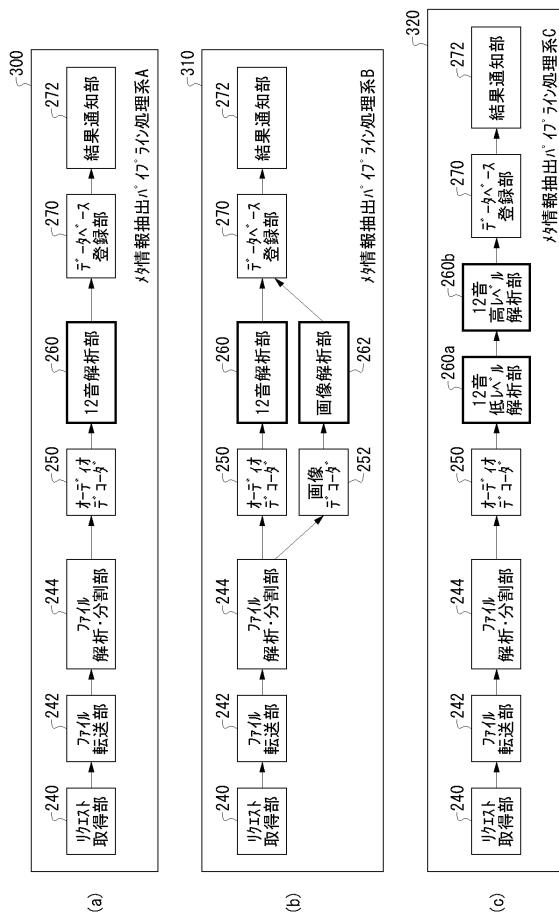
【図8】



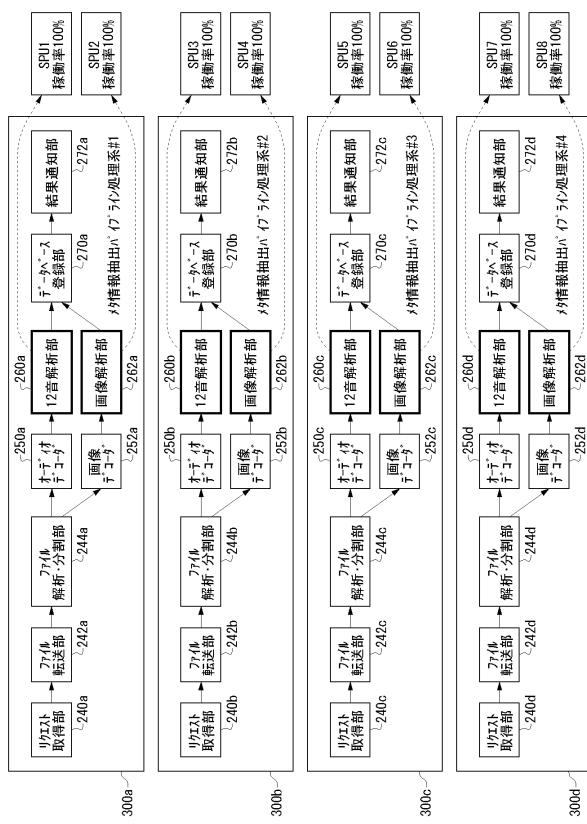
【 図 9 】



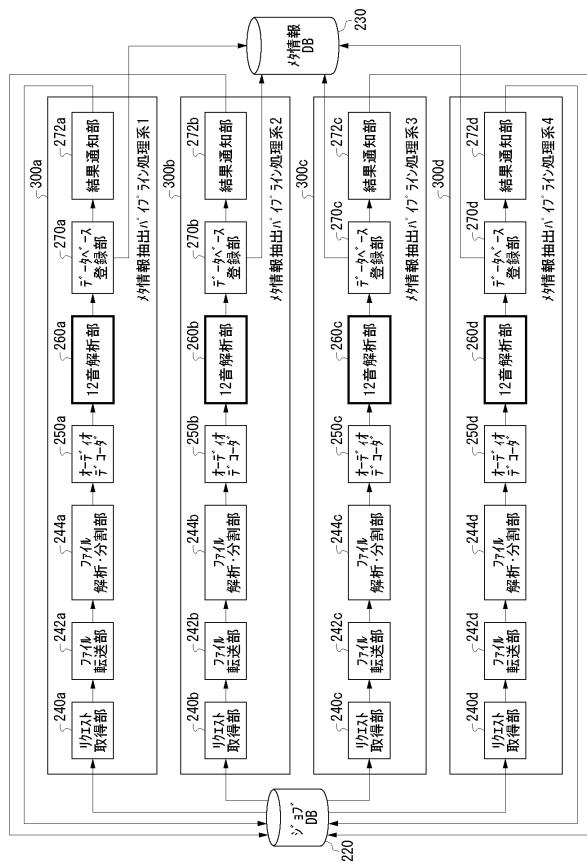
【図11】



【 図 1 0 】



【図12】



フロントページの続き

(72)発明者 斎藤 勝

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

(72)発明者 赤羽 誠

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

(72)発明者 鈴木 章

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

(72)発明者 池田 望

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

(72)発明者 高橋 良和

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

(72)発明者 森 晴

東京都港区南青山2丁目6番21号 株式会社ソニー・コンピュータエンタテインメント内

審査官 北元 健太

(56)参考文献 特開平7-84967(JP,A)

特開2004-199674(JP,A)

特開平10-116338(JP,A)

特開2006-106859(JP,A)

特開2000-353099(JP,A)

特開平6-12392(JP,A)

特表2004-509386(JP,A)

特開2006-99579(JP,A)

(58)調査した分野(Int.Cl., DB名)

G 06 F 9 / 46 - 9 / 54