

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2013-532880
(P2013-532880A)

(43) 公表日 平成25年8月19日(2013.8.19)

(51) Int.Cl.

G06F 12/00 (2006.01)
G06F 12/08 (2006.01)

F 1

G06F 12/00
G06F 12/08

テーマコード(参考)

5B005
5B060

審査請求 未請求 予備審査請求 未請求 (全 24 頁)

(21) 出願番号 特願2013-523301 (P2013-523301)
 (86) (22) 出願日 平成23年8月3日 (2011.8.3)
 (85) 翻訳文提出日 平成25年4月1日 (2013.4.1)
 (86) 國際出願番号 PCT/US2011/046412
 (87) 國際公開番号 WO2012/018906
 (87) 國際公開日 平成24年2月9日 (2012.2.9)
 (31) 優先権主張番号 12/849,724
 (32) 優先日 平成22年8月3日 (2010.8.3)
 (33) 優先権主張国 米国(US)

(71) 出願人 591016172
 アドバンスト・マイクロ・ディバイシズ・
 インコーポレイテッド
 ADVANCED MICRO DEVI
 CES INCORPORATED
 アメリカ合衆国、94088-3453
 カリフォルニア州、サンディベイル、ビイ・
 オウ・ボックス・3453、ワン・エイ・
 エム・ディ・プレイス、メイル・ストップ
 ・68(番地なし)
 (74) 代理人 100108833
 弁理士 早川 裕司
 (74) 代理人 100111615
 弁理士 佐野 良太

最終頁に続く

(54) 【発明の名称】メモリ領域を埋めるためのプロセッサ支援

(57) 【要約】

プロセッサおよび/または処理要素の間の処理負荷を分散させることに関する技術が開示される。少なくとも第1処理要素および第2処理要素を有するコンピュータシステムは、1以上のメモリ領域を初期化する要求を前記第2処理要素に処理させることができる。初期化は、前記指定された初期化されるべきメモリ領域を有するメモリへ前記第2処理要素を直接アクセスさせることにより実現されることができる。これにより、前記第2処理要素が前記初期化されるべきメモリ領域を処理する一方、前記第1処理要素は解放されて他の計算タスクを行うことができる。前記第1処理要素に関連付けられたキャッシュは、前記第2処理要素が前記初期化を行う結果、初期化により妨げられることがなく、また、データが前記キャッシュから移動されるのを回避することができる。

【選択図】図4

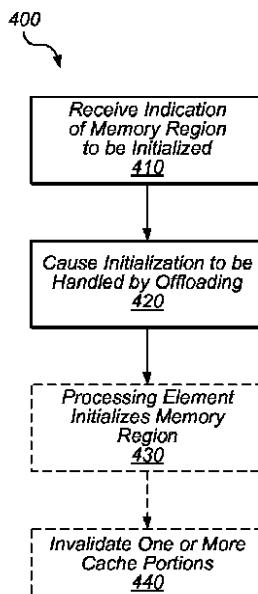


FIG. 4

【特許請求の範囲】**【請求項 1】**

コンピュータデバイスにおける初期化されるメモリ領域を示す指示に応答して、該メモリ領域の初期化を該コンピュータデバイスの第2処理要素に処理させるための、少なくとも該コンピュータデバイスの第1処理要素によって実行可能なプログラム命令を記憶する、

非一時的なコンピュータ読み取り可能な媒体。

【請求項 2】

前記メモリ領域の指示が、第1プログラムから、前記第1処理要素によって実行される制御プログラムにより受けられる、

請求項1の非一時的なコンピュータ読み取り可能な媒体。

【請求項 3】

前記制御プログラムが、前記第1プログラムを実行している、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項 4】

前記指示は、前記制御プログラムで処理可能な1以上のデータオブジェクトに対応したメモリの1以上のメモリ領域を指定しており、

前記1以上のメモリ領域のすべての内容を埋める処理がさらに含まれている、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項 5】

前記制御プログラムは、ガーベジコレクションプロセスの一部として、初期化される複数のメモリ領域を示す複数の指示を生成する処理と、前記複数のメモリ領域に対する初期化を前記第2処理要素に行わせる処理とをさらに含み、

前記初期化には、前記複数のメモリ領域のすべての内容を、プログラミング言語の言語仕様にしたがって指定されたデフォルトの内容で埋める処理が含まれる、

請求項4の非一時的なコンピュータ読み取り可能な媒体。

【請求項 6】

前記制御プログラムは、前記非一時的なコンピュータ読み取り可能な媒体に記憶された1以上のライブラリファイルを有し、かつ、前記指示を受ける前記制御プログラムは、アプリケーション・プログラミング・インターフェース(API)を介して前記指示を受ける前記制御プログラムを有する、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項 7】

前記初期化処理には、前記メモリ領域の内容を変更するために前記第2処理要素によって実行可能な1以上の命令からなる命令群の少なくとも一部を動的に生成する処理が含まれている、

請求項1の非一時的なコンピュータ読み取り可能な媒体。

【請求項 8】

前記メモリの初期化処理を前記第2処理要素に行わせる処理には、前記コンピュータシステムのキャッシュに前記初期化されたメモリ領域の初期化後の内容を記憶させることが含まれず、かつ、

前記キャッシュは、前記第1処理要素が、前記メモリ領域を有する前記コンピュータシステムのメモリにアクセスするのに応答して、前記メモリ領域の内容を記憶するように構成されている、

請求項1の非一時的なコンピュータ読み取り可能な媒体。

【請求項 9】

前記第1処理要素および前記第2処理要素の少なくとも一方を生成するために実行可能なプログラム命令をさらに有する、

請求項1の非一時的なコンピュータ読み取り可能な媒体。

【請求項 10】

10

20

30

40

50

第1処理要素と、第2処理要素と、メモリ領域を含むメモリとを備えたコンピュータデバイスにおける初期化される該メモリ領域を示す指示に応答して、前記第1処理要素で実行されている第1プログラムが、該メモリ領域の初期化を前記第2処理要素に処理させる、方法。

【請求項11】

前記第1処理要素が前記メモリ領域へ直接アクセスすることなく、前記第2処理要素が、ダイレクトメモリアクセス(DMA)を用いて前記メモリ領域を初期化する処理、をさらに含む、

請求項10の方法。

10

【請求項12】

前記指示を生成する前記第1プログラム内においてガーベジコレクションプロセスをさらに含む、

請求項10の方法。

【請求項13】

前記第1プログラムは制御プログラムであって、

前記第2処理要素が、前記制御プログラムの1以上のヒューリスティックな規則にしたがって前記メモリ領域を初期化する処理、をさらに含む、

請求項10の方法。

20

【請求項14】

前記コンピュータデバイスが、前記メモリ領域が初期化されるのに応答して、前記コンピュータデバイスのデータキャッシュの1以上の部分であってメモリ領域初期化前の前記メモリ領域の内容に対応する該部分を無効化する処理、をさらに含む、

請求項10の方法。

【請求項15】

メインメモリを有するメモリサブシステムと、2次記憶装置と、少なくとも第1および第2処理要素と、を有するコンピュータシステムであって、

前記2次記憶装置には、初期化される前記メインメモリのメモリ領域を示す指示に応答して、該メモリ領域の初期化を前記第2処理要素に処理させるように前記コンピュータシステムを動作させるための、前記第1処理要素によって実行可能なプログラム命令、が記憶されている、

30

コンピュータシステム。

【請求項16】

前記第1処理要素と、前記第2処理要素とは異機種である、

請求項15のコンピュータシステム。

【請求項17】

前記第1処理要素と関連付けられたキャッシュをさらに有し、

前記キャッシュは、前記第1処理要素による前記メインメモリへのアクセスに応答して該メインメモリの内容を記憶するよう構成され、

前記メモリ領域の初期化が行われたとき、前記メモリ領域の初期化後の内容が前記キャッシュに記憶されない、

40

請求項15のコンピュータシステム。

【請求項18】

前記第2処理要素に対して前記メインメモリへの直接アクセスを許可するように構成されたメモリアクセスコントローラをさらに有し、

メモリ領域の初期化には、前記第2処理要素が前記メモリアクセスコントローラを用いて前記メモリ領域にアクセスすることが含まれ、前記第1処理要素が前記メモリ領域にアクセスすることが含まれない、

請求項15のコンピュータシステム。

【請求項19】

50

処理要素であって、

メモリデバイスのメモリ領域の初期化を第2処理要素に処理させるように構成されたメモリ初期化回路を有し、

前記メモリ領域を初期化する指示に応答して、前記初期化を第2処理要素に処理させる処理を行う、

処理要素。

【請求項20】

コンピュータシステムで実行可能なプログラムによって利用可能なデータ構造であって、該データ構造によって記述された回路を有する集積回路を製造するプロセスの一部を行うためのデータ構造、を有する非一時的なコンピュータ読み取り可能な記憶媒体であって、

メモリデバイスのメモリ領域の初期化を、コンピュータデバイスの第1処理要素ではなく該コンピュータデバイスの第2処理要素に行わせるように構成されたメモリ初期化ユニットを有し、

前記初期化を前記第2処理要素に行わせる処理が、前記メモリ領域を初期化する指示に応答して行われる、

非一時的なコンピュータ読み取り可能な記憶媒体。

【請求項21】

HDLデータ、VerilogデータおよびGDSIIデータの少なくともいずれかを記憶している、

請求項20の非一時的なコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、コンピュータプロセッサに関し、より具体的には、メモリ領域を埋める要求を受付けるプロセッサに関する。

【背景技術】

【0002】

コンピュータの動作の間、メモリの領域がある値を用いて初期化される（埋められる）必要が生じることがある。メモリ領域の初期化には、コンピュータ上のリソースが用いられる。例えば、初期化を行うプロセッサは、一連のメモリロケーションに値を書き込まなければならないことがあり、これには時間がかかり得る。このような初期化の間、プロセッサは他の演算タスクを行うことができないことがある。

【発明の概要】

【発明が解決しようとする課題】

【0003】

さらに、メモリの初期化処理は、プロセッサと関連付けられたキャッシュに混乱を生じさせることがある。キャッシュ性能は、キャッシュの内容がメモリの初期化の間にキャッシュから移動されるので、プロセッサによって負の影響を受けることがある。例えば、キャッシュの（メモリ領域の初期化が開始される以前からの）既存の内容の幾つかまたはすべてが、初期化されるメモリ領域の内容によって置き換えられることもありうる。このような置換では、それまでキャッシュ内に存在していたデータを取得するために他のメモリが続けてアクセスされることがあるため、プログラムの実行が遅くなることがある。

【課題を解決するための手段】

【0004】

コンピュータシステムまたは計算装置が、あるメモリ処理を第1処理要素から第2処理要素へ分散することを可能とする方法および構造の様々な実施形態がここに開示される。

【0005】

記載された一実施形態において、コンピュータデバイスの初期化されるメモリ領域を示す指示を受付けることと、この受付に対応してメモリ領域の初期化をコンピュータデバイスの第2処理要素に処理せることとを含む処理を、コンピュータデバイスの少なくとも

10

20

30

40

50

第1処理要素によって実行可能なプログラム命令を有する、コンピュータ読み取り可能な媒体が開示される（ここで、プログラム命令はコンピュータ読み取り可能な媒体に記憶されている）。さらなる実施形態においては、当該指示は、第1処理要素によって実行される制御プログラムから受付ける。

【0006】

別の実施形態は、コンピュータデバイスの初期化されるメモリ領域を示す指示を受ける第1プログラムを含み、当該第1プログラムをコンピュータデバイスの第1処理要素で実行し、第1プログラムが、当該受付に対応して、メモリ領域の初期化をコンピュータデバイスの第2処理要素に処理させる方法を含む。さらなる実施形態においては、第2処理要素は、第1処理要素によるメモリ領域への直接アクセスを伴わずにメモリ領域を初期化するために、ダイレクトメモリアクセス（DMA）を利用する。

10

【0007】

さらに別の実施形態は、メインメモリを有するメモリサブシステムと、2次記憶装置と、少なくとも第1および第2処理要素と、を有するコンピュータシステムであって、前記2次記憶装置には、初期化されるメモリ領域を示す指示をコンピュータシステムに受付けさせるための、第1処理要素によって実行可能なプログラム命令が記憶され、メモリ領域は前記メインメモリ内に存在し、かつ、前記受付けに応答して、メモリ領域の初期化をコンピュータデバイスの前記第2処理要素に処理させるコンピュータシステムである。さらなる実施形態においては、コンピュータシステムは、前記第1処理要素と関連付けられたキャッシュを有し、前記キャッシュは、前記第1処理要素による前記メインメモリへのアクセスに応答して前記メインメモリの内容を記憶するように構成され、かつ、前記メモリ領域の初期化を前記第2処理要素に処理させることは、前記キャッシュが前記メモリ領域の初期化後の内容を記憶することをもたらさない。

20

【図面の簡単な説明】

【0008】

【図1】図1は、メモリの初期化を第1処理要素から第2処理要素へ分散させるように構成されたコンピュータシステムの一実施形態を示すブロック図である。

【図2A】図2Aは、初期化前後における例示的なメモリ領域を示すブロック図である。

【図2B】図2Bは、初期化前後における例示的なメモリ領域を示すブロック図である。

【図3A】図3Aは、メモリの初期化を行うように構成された制御プログラムを有するメモリサブシステムの実施形態を説明するブロック図である。

30

【図3B】図3Bは、メモリの初期化を行うように構成されたオペレーティングシステムを有するメモリサブシステムの実施形態を説明するブロック図である。

【図3C】図3Cは、メモリの初期化を行うように構成されたJAV A（登録商標）仮想マシンプログラムを含む実施形態を説明するブロック図である。

【図4】図4は、メモリの初期化を第1処理要素から第2処理要素へ分散する方法の一実施形態を説明するフロー図である。

【図5】図5は、メモリの初期化を第1処理要素から第2処理要素へ分散するコンピュータシステムの別の実施形態を説明するブロック図である。

40

【発明を実施するための形態】

【0009】

この明細書は、「一実施形態（one embodiment）」または「実施形態（an embodiment）」なる記載を含んでいる。「一実施形態において（in one embodiment）」または「実施形態において（in an embodiment）」というフレーズが出現しても、同一の実施形態を意味するとはかぎらない。ある特徴、構造または特性は、この開示と整合性のある任意の適切な態様において組み合わせられてもよい。

【0010】

（用語）

この開示（添付の請求項を含む）にみられる用語の定義および／または文脈について、

50

以下の数段落で述べておく。

【0011】

「含む・有する (Comprising)」または「含む・備える (Including)」について。これらの用語はオープンエンドな用語である。添付した請求項において用いられているとおり、これらの用語は追加的な構造またはステップを排除するものではない。1以上の処理要素を有する装置に言及している請求項を考えてみると、このような請求項は追加的な構成要素（例えば、ネットワークインターフェースユニット、グラフィックス用回路等）を有するということから当該装置を排除するものではない。

【0012】

「ように構成された (Configured To)」について。様々なユニット、回路またはその他の構成要素が、1つまたは複数のタスクを行う「ように構成されている」というように記載されることがある。このような文脈において、「ように構成された」は、当該ユニット／回路／構成要素が、動作中に当該1つまたは複数のタスクを行う構造（例えば回路）を備えている、ということを示すことによってその構造を特定するために用いられている。そのようなわけで、当該ユニット／回路／構成要素は当該具体的なユニット／回路／構成要素が現在動作中でない（例えば、オンにされていない）場合であっても当該タスクを行うように構成されているということもできる。「ように構成された」表現と合わせて用いられるユニット／回路／構成要素は、ハードウェア（例えば回路や、処理を実施するために実行可能なメモリ記憶プログラム命令等）を含む。ユニット／回路／構成要素が1以上のタスクを行う「ように構成されている」という表現は、当該ユニット／回路／構成要素について米国特許法第112条第6パラグラフの適用の問題が生じないようにすることを明示的に意図したものである。追加的に、「ように構成された」は、ソフトウェアおよび／またはファームウェア（例えば、フィールド・プログラマブル・ゲート・アレイ（FPGA）または汎用プロセッサが実行するソフトウェア）によって操作される総称的な構造（例えば総称的回路）を含むことができる。さらに、「ように構成された」は、問題となるタスク（または複数のタスク）を行うことができるような態様にて動作するために1以上のタスクを実装したまたは行うように構成されたデバイス（例えば、集積回路）を製造するための製造工程（例えば、半導体製造施設）を適合させることを含んでもよい。

【0013】

「処理要素 (Processing Element)」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、コンピュータ命令を実行することができるデバイス（例えば回路）またはデバイスの組合せを含む。処理要素は、様々な実施形態において、シングルコアプロセッサ、マルチコアプロセッサのコアの1つ、または、マルチコアプロセッサの2以上のコアからなるグループをさす。

【0014】

「プロセッサ (Processor)」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、1以上の処理要素を備えたデバイスを含む。プロセッサは、以下のものに限定する趣旨ではないが、中央演算処理装置（CPU）、コプロセッサ、演算処理装置、グラフィックス処理装置、デジタルシグナルプロセッサ（DSP）等をさすことがある。

【0015】

「第1 (first)」、「第2 (second)」等について。ここに用いられているように、これらの用語は当該用語に先行された名詞に付したラベルとして用いられており、どのような形式であれ順序（例えば、空間的、経時的、論理的順序等）を示唆するものではない。例えば、8つの処理要素またはコアを有するプロセッサにおいては、「第1」処理要素および「第2」処理要素という用語は当該8つの処理要素のうちの任意の2つの処理要素を指示するために用いられることができる。換言すれば、「第1」処理要素および「第2」処理要素は論理的な処理要素0および1に限定されるものではない。

【0016】

10

20

30

40

50

「コンピュータ（Computer）」または「コンピュータシステム（Computer System）」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、一緒に動作する1以上のコンピュータデバイスおよびそれらが有する任意のソフトウェアを含む。コンピュータデバイスは、1以上の処理要素およびメモリサブシステムを含む。メモリサブシステムは、様々なタスクを行うために1以上の処理要素により実行可能なプログラム命令を記憶してもよい。

【0017】

「コンピュータ読み取り可能な媒体（Computer-readable Medium）」について。ここに用いられているように、この用語は、コンピュータまたはコンピュータシステムによって読み取り可能な（非一時的かつ有形的な）媒体をさし、ハードドライブ、光ディスク、DVD、揮発性または不揮発性RAMデバイス、ホログラフィックストレージ、プログラマブルメモリ等の磁気、光学式および半導体記憶媒体が含まれる。コンピュータ読み取り可能な媒体との関連で用いられる「非一時的な（non-transitory）」という用語は、もっぱら、一時的な（無形な）媒体（例えば、搬送波）のような、米国特許法第101条の規定により適格性を有しないとされる任意の主題を請求の範囲から除外することを意図したものであって、法定の要件を具備する他の任意の主題を除外することを意図するものではない。

10

【0018】

「オペレーティングシステム（Operating System）」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、コンピュータシステムのリソースへのアクセスを（例えばアプリケーションからの要求に応答して）制御するプログラムまたはプログラム群を含む。幾つかの実施形態においては、オペレーティングシステムは通信デバイス、ストレージデバイス等のI/Oデバイスへのアクセスを制御する。ここに記載したように、オペレーティングシステムは、ある実施形態において、第2処理要素にメモリの初期化を行わせるために実行可能な命令を含む。

20

【0019】

「キャッシュ（Cache）」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、メモリまたはデータを保存するその他のストレージを含み、幾つかの他のメモリまたはストレージと比較してより高速なアクセスを提供することでこのようなデータのための将来の要求を改善することができるものである。

30

【0020】

「コンピュータシステムに動作ないし演算を行わせる（Causing a Computer System to Perform Operations）」について。プログラム命令の実行は、「コンピュータシステムに動作を行わせる」として記載または主張されてもよい。この語句は広く解釈されるべきであって、実行されると対象動作を行う命令や、実行されると当該動作が行われるコードをインストールまたは作成する命令が含まれる。例えば、コンピュータ読み取り可能な媒体は、コンピュータシステムにメモリ領域に対するメモリ初期化をコンピュータシステムの第1処理要素からコンピュータシステムの第2処理要素へ分散させるために実行可能な命令を含んでもよい。

40

【0021】

「実行可能な（Executable）」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、1以上のある処理要素（すなわちあるインストラクション・セット・アーキテクチャ（ISA））と関連付けされたフォーマットにおける命令を含むが、また、処理要素のISAのための命令を生成するために制御プログラム（例えばJAVA（登録商標）仮想マシン）によって解釈されることができる中間フォーマット（例えば、JAVA（登録商標）バイトコード）での命令を含む。この定義にしたがい、第1処理要素で「実行される」プログラムは、当該第1処理要素によって実行される命令の少なくとも幾つかを含む（当該プログラムの他の命令は別の処理要素によって実行されてもよいが）。プログラムの実行はまた、プログラムの解釈を含む。

【0022】

50

「アプリケーション・プログラミング・インターフェース（A P I）」について。この用語は、技術分野において通常用いられかつ受容された語義を有し、ソフトウェアが他のソフトウェアとやりとりをすることを可能にするインターフェースを含む。プログラムは、アプリケーション、ライブラリルーチン、オペレーティングシステム等の機能を利用するためのA P I呼び出しを作成してもよい。

【0023】

ここに記載したように、コンピュータプログラムは、コンピュータメモリをあるデータで初期化する（埋める）ことで当該メモリにそれまで記憶されていたデータを消去するという必要性を有することがある。幾つかの実施形態においては、メモリを初期化するという必要性は、（新たな）メモリの割り当てを受付ける要求にしたがって生じことがある。一実施形態においては、（他のJ A V A（登録商標）プログラムを動作させるために用いられる）J A V A（登録商標）仮想マシン（J V M）プログラムは、J A V A（登録商標）プログラムがメモリ領域をブランクの（デフォルト）データで利用を開始することができるよう、これらのメモリ領域をゼロで埋める（z e r o o u t）。別の実施形態においては、オペレーティングシステムは、例えば、ユーザプログラムが当該メモリにアクセスすることを許可する前にメモリ領域をすべてゼロで上書きすることもできる（幾つかの実施形態においては、消去されたデータは、パスワード、クレジットカード番号またはオペレーティングシステムがユーザプログラムにアクセスできることを望まないような他のデータを保持していたと考えることもできる）。他のタイプのプログラムによる多くの他の種類のメモリ初期化も想定しうるので、この開示はJ V Mまたはオペレーティングシステムソフトウェアに限定されるものではない。初期化中にメモリ領域を埋めるデータは、すべてゼロであってもよいが必須ではない点については後述するとおりである。

10

20

30

40

【0024】

一実施形態においては、コンピュータシステムは、中央演算処理装置（C P U）といった、例えば汎用命令を実行するように構成された第1プロセッサを有している。コンピュータシステムはまた、グラフィックス処理装置（G P U）といった、グラフィックス命令等の専用命令を実行するように構成された第2プロセッサを有している。他の実施形態においては、第1プロセッサ（または処理要素）は、C P UおよびG P U双方の機能性を单一デバイス、パッケージまたは集積回路内に有してもよい。コンピュータシステムはまた、メモリサブシステムを有している。一実施形態において、コンピュータシステムは、所定の命令シーケンスが前記第2のプロセッサによって行われるように構成されている（すなわち、プログラムされている）。これらの命令シーケンスは、第1プロセッサによって実行される命令によって生成されてもよく、メモリの初期化ルーチンを含むことができる。したがって、第1プロセッサは、第2プロセッサが初期化を行っている間は他のタスクを自由に行うことができる（例えば、初期化されるメモリ領域が直ちに必要とされなければ、第1プロセッサは、第2プロセッサがメモリの初期化を行っている間、プログラムの実行を継続することができる）。第1プロセッサのパフォーマンスを改善させることに加えて、ここに開示した技術は、例えばキャッシュからのデータの移動を避けることにより、第1プロセッサと関連付けられたデータキャッシュのパフォーマンスを改善させることもできる。

【0025】

次に図1を参照すると、メモリの初期化を第1処理要素から第2処理要素へ分散させるように構成されたコンピュータシステム10の一実施形態が図示されている。コンピュータシステム10は、バス20を介して連結された第1処理要素100Aと第2処理要素100Bとを有している。一実施形態においては、バス20により、処理要素100Aおよび処理要素100Bはメモリサブシステム60内の1以上のメモリ領域64にアクセスすることができる。メモリサブシステム60は、様々なプログラム62を格納することができ、そのようなプログラムの幾つかは、処理要素100Bを用いてメモリを初期化することを要求しまたは初期化をさせるために実行可能なプログラムである。追加的に、図1においては視覚的に区別できる構成要素として図示されているが、メモリサブシステム60

50

の一部またはすべては、処理要素 100A、処理要素 100B の回路の一部、または、処理要素 100A および処理要素 100B の双方を備えた単一のデバイスの一部をなしてもよい。一実施形態においては、キャッシュ 30 は、処理要素 100A からアクセス可能であって、メモリサブシステム 60 に保存されたデータに対応するデータを保存するよう構成されている。一実施形態においては、メモリアクセスコントローラ 75 は、処理要素 100A、処理要素 100B、メモリサブシステム 60 の任意の組合せに接続される（またはその内部に実装され）ことができ、また、バス 20 に接続されてもよい。コンピュータシステム 10 は、様々な実施形態において異なる様様にて構成され得る。

【0026】

処理要素 100A および処理要素 100B は、任意のタイプのプロセッサ（例えば、中央演算処理装置、演算処理装置、グラフィックス処理装置、デジタル信号処理装置等）に相当して（またはその内部に配置されて）もよい。一実施形態においては、処理要素 100A は中央演算処理装置（または 1 以上のコアからなるグループ）であり、処理要素 100B は異なるタイプの処理装置、例えば、グラフィックス処理装置（1 以上のコアを有してもよい）である。幾つかの実施形態においては、処理要素 100A および処理要素 100B のいずれかまたは双方は、複数のコアを有してもよい。他の実施形態においては、処理要素 100A および処理要素 100B は、同一チップ上に配置された 1 以上のプロセッサ・コアからなる異なるグループとされてもよい。処理要素 100A および処理要素 100B は、幾つかの実施形態においては、様々な処理要素からなるクラスタまたはグループで構成されてもよい（例えば、処理要素 100A は 2 つのクアッドコア・プロセッサからなるグループとされることができる）。

10

20

30

40

【0027】

一実施形態においては、処理要素をメモリサブシステム 60 に接続するバス 20 は、ノースブリッジバス、任意の他のプロセッサバスまたは当業者に公知のプロセッサ相互接続部であってもよい。バス 20 は、一実施形態においては、同一チップ上に配置されてもよい（1 以上のグループの）プロセッサ・コア間の相互接続部である。バス 20 は、单一のバスまたは相互接続部に限定される必要はなく、1 以上のバス、（2 地点間）相互接続部、またはここに記載した構造体にデータを伝達するのに適した他の通信経路および通信装置の任意の組合せであってもよい。

【0028】

メモリサブシステム 60 は、1 以上のメモリデバイスを備えている。様々な実施形態において、これらのメモリデバイスは、RAM モジュール、組込みメモリ（例えば eDRAM）、ソリッドステート記憶デバイス、ハードドライブといった 2 次記憶装置または上記定義のとおりその他の任意のコンピュータ読み取り可能な媒体で構成されてもよい。一実施形態においては、メモリサブシステム 60 は、当該メモリサブシステム 60 における 1 以上のメモリデバイス内に 1 以上のメモリ領域 64 を備えている。メモリ領域 64 は固定サイズまたは固定ロケーションである必要はなく、これに代えて任意の開始位置（または開始アドレス）および終了位置（または終了アドレス）を含むメモリの 1 以上の部分をさすものであってもよい。このように、具体的な一実施形態においては、第 1 メモリ領域はサイズが 4000KB の一連のメモリロケーションであってもよく、他方、第 2 メモリ領域はサイズが 32KB の一連のメモリロケーションであってもよい。一実施形態においては、メモリ領域 64 は、複数のメモリデバイスにわたって（または複数タイプのメモリデバイスにわたって、例えば、単一のメモリ領域が RAM モジュールとハードドライブとの記憶スペースで構成されるように）構成されてもよい。メモリ領域は、物理的または論理的に連続的であってもよく、連続的でなくてもよい。

【0029】

メモリサブシステム 60 とそのメモリ領域は、処理要素 100A によってアクセス可能とされている。例えば、処理要素 100A は、バス 20 を介してメモリサブシステム 60 からデータを取得（またはデータをメモリサブシステム 60 に保存）してもよい。様々な実施形態において、以下に説明するように、メモリサブシステム 60 はまた、処理要素 1

50

00B によってもアクセス可能とされている。様々な実施形態において、メモリサブシステム 60 は、1 以上のプログラム 62 を記憶している。プログラム 62（または複数のプログラム）は、コンピュータシステム 10 上で実行可能な任意のプログラム（または複数のプログラム）とされてもよい。このように、様々な実施形態において、プログラム 62 は、Java（登録商標）仮想マシン（JVM）、オペレーティングシステム、API ライブラリ、JVM またはオペレーティングシステム上で動作するユーザプログラム等であってもよい。様々な実施形態において、プログラム 62 は、以下にさらに議論するように、メモリの初期化を処理要素 100A から処理要素 100B へ分散させる機能を有することができる。

【0030】

10

メモリアクセスコントローラ 75 は、一実施形態においてはメモリサブシステム 60 に接続されており、様々な実施形態においてメモリサブシステム 60 の制御、管理、調整および／または処理要素 100 からメモリサブシステム 60 へのメモリアクセスの許可を行うように構成されている。メモリアクセスコントローラ 75 は、一実施形態においてはダイレクトメモリアクセス（DMA）コントローラであり、処理要素 100A および／または処理要素 100B と同一チップ上に配置されてもよい。様々な実施形態において、メモリアクセスコントローラ 75 は、処理要素 100A によって警告、通知を受けまたは権限を付与された場合を除いて処理要素 100B がメモリ領域 64 にアクセスすることを制限してもよい（このような場合には、アクセスコントローラ 75 はメモリサブシステム 60 の領域の幾つかまたはすべてへのアクセスを許可してもよい）。メモリアクセスコントローラ 75 は、一実施形態においてバス 20 を使用し（および／またはバス 20 に接続され）てもよい。

20

【0031】

キャッシュ 30 は、処理要素 100A によりアクセス可能とされており、メモリサブシステム 60 に対応するデータを保持するように構成されたキャッシュを有している。キャッシュ 30 は、このようにして、当該データへの処理要素 100A の高速アクセスを提供するために、メモリサブシステム 60 に保存されたデータのサブセットを保持するように構成されてもよい。様々な実施形態において、キャッシュ 30 は、L1 キャッシュ、L2 キャッシュ、L3 キャッシュまたはその他のキャッシュを含む階層型キャッシュシステムを含んでもよい。キャッシュ 30 は、様々な実施形態において処理要素 100A 内に部分的または完全に配置されてもよく、あるいは、処理要素 100A の外部に部分的または完全に配置されてもよい（例えば、一実施形態においては、キャッシュ 30 は処理要素 100A 内の L1 キャッシュおよび処理要素 100A の外部の L2 キャッシュとで構成されている）。所定の処理要素と「関連付け」られたキャッシュは、当該処理要素によってアクセスされるように構成されている。

30

【0032】

幾つかの例においては、キャッシュ処理によって、キャッシュ 30 に既に記憶されていたデータが他のデータと置換される（または他のデータによって置換される）ことがある。幾つかの実施形態においては、処理要素 100A がメモリサブシステム 60 のメモリ領域に直接アクセスする際には、アクセスされたデータを記憶するためにキャッシュ 30 の一部が用いられることになる。例えば、もし処理要素 100A がメモリサブシステム 60 に直接アクセスしてメモリ領域 64 を初期化するのであれば、キャッシュ 30 内の既存データが当該メモリ領域のための新たに初期化されたデータによって移動されることが考えられる。キャッシュから移動されたデータはアクセスに時間がかかることがある、これにより実行時間が長くなりうる。例えば、以下の C コードを考えてみるとよい：

40

```
int C = A + B;
int *Freespace = malloc(8192);
```

```
E = C;
```

このコードでは（コンパイルおよび実行されると）、キャッシュされる変数「C」のためのデータ値が最初にもたらされる。次いで、malloc（）関数の呼び出しによりメモ

50

リ 8 1 9 2 バイト分の初期化が生じると考えられ、キャッシュから「C」のための値が移動される。次の命令（「C」の値を変数Eに割り当てる命令）が実行されると、キャッシュは「ミス（外れ）」となり、より低レベルのキャッシュまたはより離隔したメモリから変数Cの値を取得しなければならなくなり、遅延が生じる。もしCの値が最初の段階でキャッシュから移動されていなかったとすれば、この遅延は回避し得たのであって、パフォーマンスがスピードアップしていたかもしれない。キャッシュ30についてのデータの移動／置換は、様々な実施形態において、LRU (least recently used) 方式の置換を含む、当業者が考えうる任意の数のハードウェアまたはソフトウェアスキームを含む置換ポリシーによって規律されてもよい。

【0033】

次に図2Aを参照すると、初期化前のメモリ領域64の一例が図示されている。図示されているように、メモリ領域64は、複数のメモリロケーション（ロケーション212-216を含む）を備えており、それぞれが個別にアドレス可能とされており様々な実施形態において所定量のデータを記憶するように構成されている。図示されているように、メモリロケーション212は、データ205を記憶している。メモリロケーション212内のデータ205は、幾つかの実施形態においては、コンピュータシステム10によって実行されているプログラムによってあらかじめ書き込まれたものであってもよく、あるいは、単純に任意的（ランダム）であってもよい。

【0034】

図2Bにおいて、初期化後のメモリ領域64の一例が図示されている。この実施形態においては、メモリロケーション212内のデータ205は、ゼロの値を有する複数ビットからなるビット列に初期化することにより「ゼロで埋められ（zeroed out）」ている。ここにさらに議論されるように、この初期化は、は幾つかの実施形態においては処理要素100Bによって行われる。「ゼロで埋める処理（Zeroing out）」は初期化の一形態にすぎない；その他の初期化には、テストパターンでデータを書き込む（例えば、すべて負の値に対応する値、16進値0xDEADBEEF等）が含まれる。初期化は、幾つかの実施形態においては、JAVA（登録商標）プログラミング言語仕様といった外部仕様にしたがって行われてもよい。初期化は上述したデータ形式および値に限定されるものではなく、様々な実施形態において、1以上のメモリ領域を埋める任意のデータを含んでもよい。

【0035】

幾つかの実施形態においては、メモリの初期化は、ある最小サイズのメモリ領域を初期化することに限定されてもよい（最小サイズについては初期化要求に応じる制御プログラムが任意に判断することが考えられる）。例えば、メモリの初期化は、（例えば8KBのページといった、コンピュータシステム10のオペレーティングシステムによって定義されるような）ページ1単位、キャッシュライン幅または（1024バイトといった）所定の固定サイズ等を下回らない領域を初期化することに限定されてもよい。これらの実施形態においては、メモリの初期化のための最小サイズ閾値は、小さいメモリ領域を初期化するために第2処理要素を使用することに伴い生じうる性能ペナルティを回避するために設定されることも考えられるが、これは、第1処理要素ではなく第2処理要素を使用して初期化を行うことで様々な実施形態において不可避的なオーバーヘッド・コストが生じることがあるからである。

【0036】

次に図3Aを参照すると、メモリサブシステム60内にユーザプログラム304と制御プログラム310とを備えた実施形態を示すブロック図が図示されている。一実施形態においては、プログラム304およびプログラム310はともに図1に関して前述した各別のプログラム62である。様々な実施形態において、ユーザプログラム304は、メモリに直接アクセスしてメモリ領域（複数のメモリ領域）を初期化する権限を欠いている（またはそのようにプログラムおよび／または設計されていない）こともあるが、他方、制御プログラム310は、メモリ領域を初期化するように実行可能とされている（例えば初期

10

20

30

40

50

化ルーチン 313 を使用して)。例えばプログラム 304 は J A V A (登録商標) プロセスおよび / またはユーザアプリケーションであってもよく、他方、プログラム 310 は J V M またはオペレーティングシステムであってもよい(下記の図 3B ないし図 3C による説明を参照)。様々な実施形態において、ユーザプログラム 304 および制御プログラム 310 は、サブシステム 60 内の 1 以上のメモリデバイス内に記憶されている(例えば、制御プログラム 310 はハードドライブに記憶されたうえで実行中に R A M モジュールに(完全または部分的に)ロードされてもよい)。

【0037】

制御プログラム 310 は、様々な実施形態において、処理要素 100A および / または処理要素 100B によって実行可能な命令を有している。すなわち、所定の制御プログラム 310 は、処理要素 100A、処理要素 100B、あるいは処理要素 100A および処理要素 100B の組合せによって実行可能な命令を有してもよい。例えば、一実施形態においては、制御プログラム 310 は、処理要素 100A および処理要素 100B の双方によって実行可能な単一のインストラクション・セット・アーキテクチャ(i n s t r u c t i o n _ s e t _ a r c h i t e c t u r e ; I S A) に属する命令を有している。他方、別の実施形態においては、制御プログラム 310 は、処理要素 100A により実行可能な第 1 の I S A に属する命令を有しており、また、処理要素 100B により実行可能な第 2 の異なる I S A に属する命令を有している。メモリの初期化ルーチン 313 は、このようにして、幾つかの実施形態において制御プログラム 310 の他の部分とは異なる I S A に属する命令を有してもよい。

10

20

30

【0038】

一実施形態においては、制御プログラム 310 は、ユーザプログラム 304 からメモリ要求 305 を受付けるために実行可能な初期化ルーチン 313 を含むプログラム命令群を備えている(別の実施形態においては、制御プログラム 310 は内部でメモリ要求 305 を生成する)。メモリの初期化ルーチン 313 は、処理要素 100B(処理要素 100A ではなく)に初期化要求 305 で指定される 1 以上のメモリ領域 64 を初期化させるために実行可能である。メモリの初期化ルーチン 313 は、様々な実施形態において、O P E N C L、J A V A(登録商標)、C++ 等のプログラミング言語で書かれたコードに対応する命令を含む。ルーチン 313 に対応するコードは初期化ルーチン 313 を様々な実施形態にて行うために解釈および / またはコンパイルされてもよい。

【0039】

O P E N C L コードがどのようにして処理要素 100B により実行可能な命令を生成するためには用いることができるかの一例は、2010 年 5 月 21 日出願の米国特許出願第 12/785,052 号、発明の名称:「コンピューティング・プラットフォームにおける処理負荷の分散(D I S T R I B U T I N G W O R K L O A D S I N A C O M P U T I N G P L A T F O R M)」に見出すことができるところ、その内容は引用により本明細書に組み込まれる。

30

【0040】

メモリの初期化ルーチン 313 は、様々な実施形態において、処理要素 100B にメモリ領域 64 を初期化させるために実行されてもよい。一実施形態においては、初期化ルーチン 313 の実行は、初期化要求 305 に応答して開始され、この初期化要求 305 は、ユーザプログラム 304 によって生成されてもよい。初期化要求 305 は、様々な実施形態において様々な形式をとってもよく、初期化すべき 1 以上のメモリ領域 64 を識別または判定するのに利用できる情報を含んでいる。一実施形態においては、要求 305 は、データオブジェクトの名称を指定する。一実施形態においては、初期化要求 305 は、初期化されるべきメモリ空間のメモリベースアドレスおよびオフセット値(長さ)を含んでいる。他の実施形態においては、初期化要求 305 は、初期化されるべきメモリの基底の「開始」アドレスおよびメモリ上限の「停止」アドレスを含んでいる。しかしながら、メモリ要求 305 は、このように限定されたものではなく、初期化されるべき 1 以上のメモリ領域 64 を判定するのに利用可能な任意の情報を含んでもよい。

40

50

【0041】

実行中、制御プログラム310は、処理要素100Aおよび/または処理要素100Bによって実行されるが、少なくとも一実施形態においては、初期化ルーチン313の実行は、初期化要求307を用いてもっぱら処理要素100Bによって行われる。処理要素100Bによるルーチン313の実行は、様々な実施形態において異なる様にして進行してもよい。一実施形態においては、制御プログラム310の一部は、処理要素100Bによるルーチン313の実行を「設定」するために処理要素100Aによって実行可能とされてもよい。処理要素100Aは、ルーチン313への参照を含む処理要素100Bへの制御メッセージ、通知または命令を送信してもよい。このような制御メッセージを受付けると、処理要素100Bはさらにルーチン313を実行することが考えられる（例えば、ルーチン313の命令が記憶されているメモリおよび/またはキャッシュに直接アクセスすることによって）。別の実施形態においては、初期化ルーチン313に対する命令は、単純にバス（バス20等）上に出力されてもよく、その時点で処理要素100Bは当該命令を認識して実行するであろう。一実施形態においては、処理要素100Aは、処理要素100Bがメモリ領域64に直接アクセスすることをメモリアクセスコントローラ75に許可させるコンフィギュレーション動作を含む、処理要素100Bのための1以上のコンフィギュレーション動作を行うために、（処理要素100AのISAに属する）命令を実行してもよい。処理要素100Bに初期化ルーチン313を実行させるために、当業者が想到しうる様々な他の技術もまた利用することができる。

10

【0042】

20

初期化ルーチン313の命令は、一実施形態においては、初期化すべき1以上のメモリ領域64を指定する1以上の参照と1以上のメモリ領域を初期化するために処理要素100Bにより実行可能な命令とを含む。初期化されたメモリ領域を埋めるデータはすべてゼロ、すべて負の値、パターン化されたデータまたは前述したようなその他のデータとすることができます。幾つかの実施形態においては、初期化ルーチン313の一部分（またはその全体）は、制御プログラム310によって動的に生成されてもよい。動的生成は、一実施形態においてはメモリ要求305に含まれる情報に応答して行われてもよい。例えば、もしメモリ要求305がRAMのうち8MB相当の部分が初期化すべきであると指定しているならば、少なくとも初期化ルーチン313の一部分はこの8MB値を反映するように動的に変更されてもよい。

30

【0043】

初期化ルーチン313は、様々なソフトウェアプログラムの一部として行われてもよい。例えば、一実施形態においては、ルーチン313は、アプリケーション・プログラミング・インターフェース（API）の仕様にしたがってメモリ要求305が作成されてライブラリルーチンの一部として行われてもよい。別の実施形態においては、ルーチン313は、JAVAFX（登録商標）ガベジコレクションプロセスの一部として行われてもよい（図3Cを参照して後述するとおり）。しかしながら、初期化ルーチン313は上述したタイプのプログラムに限定されるものではない。

【0044】

40

次に図3Bを参照すると、コンピュータシステム10のオペレーティングシステム320がメモリの初期化を第1処理要素から第2処理要素へ分散させるように構成されている実施形態を示したブロック図が図示されている。一実施形態においては、オペレーティングシステム320は、制御プログラム310について前述した動作のいずれかまたはすべてを完全にまたは部分的に行うようにしててもよい。様々な実施形態において、オペレーティングシステム320は、1以上のメモリ要求305を受付、生成および/または処理して1以上のメモリ領域64を初期化してもよい。一実施形態においては、メモリ要求305は、オペレーティングシステム320内のライブラリ（またはモジュール）により受けられてもよく、プログラム62、プログラム304といったプログラムまたはオペレーティングシステム320それ自体によって呼び出し可能とされてもよい。これらのライブラリは、様々な実施形態において、メモリサブシステム60内の1以上のファイルとして

50

記憶されてもよい、要素 322, 324 といったモジュールのための API インタフェースを有してもよい、これは、C プログラミング言語関数 `malloc()` および `init()` に対応するものである。例えば、コンピュータシステム 10 上で実行されるプログラム 62 は、`malloc()` ルーチンを呼び出すことにより（より多くの）メモリの割り当てを受けることを要求してもよい。オペレーティングシステム 320 はしたがって、一実施形態においては、適切な命令（初期化ルーチン 313 等）をロードおよび／または動的に生成したうえでこれらのロードまたは生成された命令を第 2 処理要素 100B に実行させることにより、当該要求に応じる。このような初期化は、例えば初めて割り当てられたメモリブロックがデータを 1 つのプログラムから別のプログラムへと漏洩することを回避するなどのセキュリティ上の理由で望ましいであろう。`Init` モジュール 324 は、一実施形態においてはメモリに別のプロセスをロードするために用いられてもよく、そのためメモリ要求 305 を内部で生成してもよいだろう（これに応じて初期化要求 307 が処理要素 100B へ送信されてもよい）。

【0045】

次に図 3C を参照すると、JAVA（登録商標）仮想マシン（JVM）330 がメモリの初期化を第 1 処理要素から第 2 処理要素へ分散させるように構成されている実施形態を示したブロック図が図示されている。JVM330 は、制御プログラム 310 について前述した動作のいずれかまたはすべてを完全にまたは部分的に行うようにしてよく、メモリサブシステム 60（図示せず）に記憶されてもよい。一実施形態においては、JVM330 は、メモリサブシステム 60 に記憶された 1 以上の JAVA（登録商標）プログラムの JAVA（登録商標）バイトコードを実行するように構成されている（したがって、制御プログラム 310 はこのようにして他のプログラムを実行してもよく、さらに、この観点では JAVA（登録商標）プログラムに限定されるものではない）。JAVA（登録商標）バイトコードの実行は、任意の数の JAVA（登録商標）オブジェクト 331 が作成および／または破棄されることを生じさせるかもしれない。JAVA（登録商標）オブジェクトのためのデフォルトの初期値は、JVM330 の様々な実施形態においてすべてゼロに設定されてもよい。このような初期化は、様々な実施形態において、ガーベジコレクションプロセス 332 および／またはコンストラクタルーチン 334（これは、幾つかの実施形態においては、全体としてまたは部分的に、初期化ルーチン 313 に対応してもよい）によって行われてもよい。ガーベジコレクションプロセス 332 の最後のステップとして、一実施形態においては、1 以上のメモリ領域のすべてが、メモリ領域をゼロで埋めることにより将来のオブジェクト割り当てのために利用可能とされてもよい（次のガーベジコレクションが追加的に初期化されたメモリをもたらすまで、これにより既に初期化されたメモリを確保する）。あるいは、ゼロで埋める処理は、様々な実施形態において、新たなオブジェクトが JAVA（登録商標）ユーザプログラムによって割り当てを受けるたびにその都ベースで行われることができる。

【0046】

一実施形態においては、ガーベジコレクションプロセス 332 は、どの JAVA（登録商標）オブジェクトがもはや使用されていないかを判定してそれらの使用されていないオブジェクトに対するメモリの割り当てを解除する。このメモリの割り当て解除のプロセスにおいて、JVM330 は、1 以上の対応するメモリ領域を初期化してゼロ値を含ませるようにしてよい。JVM330 はまた、1 以上のコンストラクタルーチン 334（または複数のルーチン）を実行させるようにしてよい。コンストラクタルーチン 334（または複数のルーチン）は、デフォルトのルーチンであってもよく、解放されているメモリの割り当て、JVM330 で実行されている 1 以上の JAVA（登録商標）プログラムに対して行うことを必要とするかもしれない、同様にそれらの JAVA（登録商標）プログラム（これは様々な実施形態においてユーザプログラム 304 に対応してもよい）の実行中に 1 以上のメモリ領域 64 の初期化を行わせてよい。JVM330 によるメモリ領域の初期化を最適化するための様々な技術、変形が当業者において想起されると思われる。例えば、JVM330 は、より大きな容量のメモリ（例えば 1MB）を「ゼロ」にして新た

10

20

30

40

50

に作成された J A V A (登録商標) オブジェクトの要求を満たすのに必要なだけメモリ分配するように構成されてもよい(クラスが作成されるたびにその都度メモリを初期化するのではなく)。

【0047】

様々な実施形態において、オペレーティングシステム 320 および JVM 330 を除く多くのプログラムが、コンピュータシステム 10 にメモリ領域を初期化するタスクの処理要素 100A から処理要素 100B への分散を、コンピュータシステム 10 行わせてもよい。コンピュータシステム 10 の処理要素によってコンパイルおよび実行(または解釈)されるように設計された異なるプログラミング言語は、メモリの初期化の分散(またはオフロード)能力の利点を活用するように設計された API ルーチンを含むライブラリを有してもよい。さらにコンパイラは、高水準ソースコードから実行可能なコードを生成する際に、ここに記載した技術を用いてメモリの初期化の分散を生じさせるように設計されることも考えうる。コンパイラは、一実施形態においては、プログラムが 1 以上のメモリ埋め処理を第 1 処理要素から第 2 要素へ分散することでどの段階で利益を得ることができるかを判定するために、ヒューリスティックスを利用することができる(例えば、このようなヒューリスティックスの基礎をなしうる要素としては、メモリ領域のサイズ(当該領域が十分に大きい場合にはおそらくオフロード / 分散)、初期化に続いてどの程度頻繁にかつ直ちにメモリ領域がアクセスされるか、初期化が行われた後所定の期間においてアクセスされた初期化された領域のバイト数、所定のメモリの初期化をオフロード化しないことによるキャッシュ移動の結果として予想されるキャッシュミスの量等が考えられる)。幾つかの実施形態においては、ここに記載したメモリの初期化技術は、ソースコードプログラマに対して幾つかの場合において透過的である。例えば、ソースコードプログラマは、C プログラミング言語による `malloc()` に対する呼び出しを、当該プログラミング言語の言語仕様にしたがって、当該呼び出しを扱うライブラリルーチンがメモリの初期化を第 1 要素から第 2 要素へと分散させることを認識することなくプログラムすることができる。

10

20

30

【0048】

次に図 4 を参照すると、1 以上のメモリ領域の初期化の第 1 処理要素による第 2 処理要素へのオフロードに係る方法 400 の一実施形態のフロー図が図示されている。方法 400 は、全体としてまたは部分的に、コンピュータシステム 10 または後述するシステム 500 といった他の任意の適切なコンピュータシステムまたはコンピュータデバイスによつて行われてもよい。ステップ 410 において、初期化される 1 以上のメモリ領域の指示が受け付けられる。このステップは、一実施形態においては、例えばプログラム 304 からのメモリ要求を受け付けるために制御プログラム 310 を実行する処理要素 100A によつて行われてもよい。一実施形態においては、ステップ 410 は、JVM 330 のプロセス 332 といったガベージコレクションプロセスによって生成された要求を受け付けることが含まれる。

40

【0049】

ステップ 420 において、ステップ 410 の指示の受付に応答して、コンピュータシステム 10 は、要求されたメモリ領域の初期化を処理要素 100A から処理要素 100B へオフロードする。一実施形態においては、ステップ 420 は、処理要素 100A によつて行われ、要求されたメモリ領域の初期化を処理要素 100B へオフロードさせる。ステップ 420 はまた、様々な実施形態において、コンフィギュレーション動作を行うかそうでなければ処理要素 100B にメモリ領域 64 を初期化させるような態様にて処理要素 100B とやり取りを行う処理要素 100A を含んでもよい(例えば初期化ルーチン 313 を実行するように処理要素 100B を設定するなど)。

50

【0050】

ステップ 430 において、ステップ 410 の初期化要求がオフロードされた(すなわち分散された)処理要素は、指示された 1 以上のメモリ領域を初期化する。一実施形態においては、このステップは要求されたメモリ領域を初期化するために(コントローラ 75 を

50

介して) ダイレクトメモリアクセスを利用して処理要素 100B によって行われる。このように、ステップ 430 の様々な実施形態において、初期化は、処理要素 100A が初期化されるべきメモリ領域のための値を直接変更することなしに行われる。ある実施形態において、ステップ 430 は、制御プログラム 310 の 1 以上の所定のルール、ルーチン等にしたがって行われる。これらのルールは、ヒューリスティクス(例えば前述したようなヒューリスティクス)を含むことが考えられる。

【0051】

ステップ 440 において、コンピュータシステム 10 のキャッシュの 1 以上の部分を無効化してもよい。複数の処理要素(コンピュータシステム 10 等)を有するシステムにおいては、メモリ領域 64 内のデータのコピーは、幾つかの実施形態においてメモリ階層(キャッシュ 30 を含む)に記憶されてもよい。もしメモリ領域 64 が方法 400 にしたがって初期化されるならば、幾つかの事例においておよび幾つかの実施形態においてはキャッシュ無効化手順を行って、コンピュータシステム 10 のキャッシュ(例えば、キャッシュ 30)内に残っている初期化されたメモリ領域 64 に対応するデータの古いコピーが存在しないことを確保することが必要となることがある。ステップ 440 は、様々な実施形態において、処理要素 100A、処理要素 100B および/またはメモリアクセスコントローラ 75 により様々に開始されてもよく、当業者に知られた様々な技術を用いて行ってもよい。

10

【0052】

次に図 5 を参照すると、前述した様々な実施形態を実装することができる例示的なコンピュータシステム 500 を示すブロック図が図示されている。コンピュータシステム 500 のコンポーネントは、全体的または部分的にコンピュータシステム 10 のコンポーネントと同一または類似であってよい。例えば、コンピュータシステム 500 は図示したとおり、メモリサブシステム 60、処理要素 100A、処理要素 100B、キャッシュ 30、メモリアクセスコントローラ 75 を備えている。コンピュータシステム 500 は、サーバシステム、パーソナルコンピュータシステム、デスクトップコンピュータ、ラップトップノートブックコンピュータ、メインフレームコンピュータシステム、携帯用コンピュータ、ワープステーション、ネットワークコンピュータ、携帯電話、ポケットベルまたは携帯情報端末(PDA)といった消費者装置等の任意の様々なタイプのデバイスを含んでもよいが、これらに限定されるものではない。コンピュータシステム 500 はまた、ストレージデバイス、スイッチ、モデム、ルータ等の任意のタイプのネットワーク接続周辺デバイスであってもよい。図 5 には便宜上单一のコンピュータシステム 500 が図示されているが、システム 500 はまた、一緒に動作する 2 以上のコンピュータシステムとして実装されてもよい。

20

30

【0053】

コンピュータシステム 500 の一実施形態において、メモリサブシステム 60 は、2 次記憶装置 455 と RAM モジュール 444, 446 を備えている。一実施形態においては、2 次記憶装置 455 には、プログラム命令が記憶されており、これらのプログラム命令は、コンピュータシステムに初期化されるべきメモリ領域の指示を受付けさせるために、第 1 処理要素 100A により実行可能とされている。ここで、メモリ領域はコンピュータシステムのメモリ内にあり、前記指示の受付けに応答して、メモリ領域の初期化をコンピュータデバイスの処理要素 100B に処理させる。処理要素 100A および処理要素 100B は、ある実施形態において、異種(すなわち異なるタイプに属する)のものであってよい。例えば、処理要素 100A が中央演算処理装置(CPU)であり、処理要素 100B がグラフィックス処理装置(GPU)であるような場合であってもよい。さらに、一実施形態においては、キャッシュ 30 は、処理要素 100A がメモリにアクセスするのに応じてメモリサブシステム 60 内の 1 以上のメモリデバイスの内容を記憶するように構成されてもよく、ここで、メモリ領域の初期化を行うことには、当該メモリ領域の初期化後の内容をキャッシュに記憶せることを含まない。すなわちキャッシュ 30 は、初期化されたメモリ領域に対応する新たに初期化されたデータによって他のデータがキャッシュ 3

40

50

0 内から移動されることを、回避することができる。メモリアクセスコントローラ 75 は、様々な実施形態において、処理要素 100B に対してメモリサブシステム 60 内の 1 以上のメモリデバイスへのアクセスを提供するように構成されてもよく、ここで、メモリ領域の初期化を行うことには、処理要素 100B がメモリアクセスコントローラ 75 を用いてメモリ領域にアクセスすることが含まれ、かつ、初期化を行うことには処理要素 100A がメモリ領域にアクセス（すなわちメモリ領域を変更）することは含まれない。

【 0054 】

追加的に、一実施形態においては、I/O 装置 444 は、バス 20 を介してメモリサブシステム 60 に接続されている。様々な実施形態において、I/O 装置は、他のストレージデバイス（ハードドライブ、光学ドライブ、リムーバブルフラッシュドライブ、ストレージアレイ、S AN またはそれらに関連するコントローラ）、ネットワークインターフェースデバイス（例えばローカルエリアまたはワイドエリアネットワーク）またはその他のデバイス（例えばグラフィックス、ユーザインターフェースデバイス等）を含んでもよい。一実施形態においては、コンピュータシステム 500 は、ネットワークインターフェースデバイスを介してネットワークに接続されている。I/O 装置は、様々なタイプのインターフェースを含んでもよく、これは様々な実施形態にしたがって、他のデバイスおよびそれらのインターフェースに接続され通信を行うように構成されてもよい。一実施形態においては、I/O インタフェースはフロントサイドバスから 1 以上のバックサイドバスへのブリッジチップ（例えばサウスブリッジ）である。

10

【 0055 】

メモリサブシステム 60 は、処理要素 100A および / または処理要素 100B により様々な実施形態において利用可能なメモリを有している。サブシステム 60 内のメモリは、ハードディスク・ストレージ、フロッピー（登録商標）ディスク・ストレージ、リムーバブル・ディスク・ストレージ、フラッシュメモリ、ランダムアクセスメモリ（RAM S RAM、EDO RAM、SDRAM、DDR SDRAM、RAMBUS RAM 等）、読み出し専用メモリ（P ROM、E E P ROM 等）といった異なる物理的記憶媒体を用いて実装されてもよい。コンピュータシステム 500 内のメモリは RAM 444, 446 および 2 次記憶装置 455 のようなストレージに限定されるものではない。むしろ、コンピュータシステム 500 はまた、ここに記載しなかったキャッシュメモリや I/O 装置 444（例えば、ハードドライブ、ストレージアレイ等）上のストレージといった他の形態のストレージを有してもよい。幾つかの実施形態においては、これらの他の形態のストレージはまた、処理要素 100A および / または処理要素 100B により実行可能なプログラム命令を記憶してもよい。

20

30

【 0056 】

前述した技術および方法は、任意の適切なコンピュータ読み取り可能な媒体に記憶されたコンピュータ読み取り可能な命令として実装されてもよい。これらの命令は、コンピュータシステムおよび / またはコンピュータデバイスが前述した態様にて動作することを可能にするソフトウェアであってもよく、メモリサブシステム 60 内のコンピュータ読み取り可能な媒体（またはメモリサブシステム 60 内にはない別のコンピュータ読み取り可能な媒体）に記憶されてもよい。このように、ライブラリルーチン、ガーベジコレクションプロセス、その他のソフトウェアルーチンおよびオブジェクトならびに任意のもしくはすべてのソフトウェア 62, 304, 310, 313, 320, 322, 324, 330, 331, 332, 334 はこのようなコンピュータ読み取り可能な媒体上に記憶されてもよい（本明細書の段落 [0017] に注記したとおり、このような媒体は非一時的なものとされてもよい）。

40

【 0057 】

さらに、前述した技術および方法は、幾つかの実施形態においてはハードウェアとして実装されてもよい。例えば、一実施形態は、メモリデバイスのメモリ領域の初期化を第 2 処理要素に処理させるように構成されたメモリ初期化回路を備えた処理要素であり、ここで、初期化を行うことは、メモリ領域が初期化されるべきであるとの指示に応答して行わ

50

れる。ハードウェア実施形態は、前述したアルゴリズムおよび技術を実装するための回路論理を用いてもよい（例えば、方法400）。

【0058】

ハードウェアの実施形態は、ハードウェア生成命令を用いて生成されてもよい。例えば、ハードウェア生成命令は、VerilogまたはVHDLといった高水準設計言語（HDL）によるビヘイビア・レベルまたはレジスタ転送レベル（RTL）の記述レベルでハードウェア機能を記述した1以上のデータ構造の概要を示すものであってもよい。記述内容は合成ツールにより読み取られてもよく、合成ツールは記述を合成してネットリストを生成するものであってもよい。ネットリストは、ゲート群（例えば合成ライブラリに定義された）を含んでもよく、メモリの初期化の分散化／オフロード化を実現するように構成された処理要素（100Aおよび／または100B等）の機能性を表現するものである。このようなネットリストを配置・経路処理したうえでマスクに対して適用すべき幾何学形状を記述したデータセットが生成されてもよい。このようなマスクは1以上の処理要素（100Aおよび／または100B等）に対応する半導体回路または回路系を製造するために様々な半導体製造工程において利用されてもよい。代替的に、データベースは所望によりネットリスト（合成ライブラリ付きまたは合成ライブラリなし）またはデータセットであってもよい。このように、ハードウェア生成命令は、前述した方法および技術を実装するプロセッサおよび／または処理要素を、製造分野における知識を有する当業者に知られた技術にしたがって生成または作成するために実行されてもよい。追加的に、このようなハードウェア生成命令は、任意の適切なコンピュータ読み取り可能な媒体（これは要素60のようなメモリサブシステム内に存在してもよく、あるいは他のコンピュータ読み取り可能な媒体に存在していてもよい）に記憶されてもよい。

前述したコンピュータ読み取り可能な記憶媒体は、幾つかの実施形態において、プログラムにより読み出される命令を記憶するために用いられることができ、また、処理要素100Aおよび／または100Bを有するハードウェアを製造するために直接的または間接的に用いられることができる。例えば、命令は、VerilogまたはVHDLといった高水準設計言語（HDL）によるビヘイビア・レベルまたはレジスタ転送レベル（RTL）の記述レベルでハードウェア機能を記述した1以上のデータ構造の概要を示すものであってもよい。記述内容は合成ツールにより読み取られてもよく、合成ツールは記述を合成してネットリストを生成するものであってもよい。ネットリストは、ゲート群（例えば合成ライブラリに定義された）を含んでもよく、処理要素100、メモリ初期化ユニットおよび／またはメモリ初期化回路の機能性を表現するものである。このようなネットリストを配置・経路処理して、マスクに対して適用すべき幾何学形状を記述したデータセットが生成されてもよい。このようなマスクはハードウェア実施形態に対応する半導体回路または回路系を製造するために様々な半導体製造工程において利用されてもよい。代替的に、データベースは所望によりネットリスト（合成ライブラリ付きまたは合成ライブラリなし）またはデータセットであってもよい。一実施形態はこのようにして（非一時的な）コンピュータシステムで実行可能なプログラムによって利用可能なデータ構造であって該データ構造によって記述された回路を有する集積回路を製造するプロセスの一部を行うためのデータ構造を有する（非一時的な）コンピュータ読み取り可能な記憶媒体であって、前記データ構造に記述された前記回路が、メモリデバイスのメモリ領域の初期化をコンピュータデバイスの第1処理要素ではなく該コンピュータデバイスの第2処理要素に行わせるように構成されたメモリ初期化ユニットを有し、かつ、前記初期化を前記第2処理要素に行わせる処理が、前記メモリ領域が初期化されるべきとの指示に応答して行われることを特徴とする非一時的なコンピュータ読み取り可能な記憶媒体、となる。

【0059】

以上のとおり具体的な実施形態について説明したが、これらの実施形態は、本開示の範囲を限定することを意図するものではなく、ある特徴に関して单一の実施形態のみが記載されている場合でも同様である。開示に含まれる特徴の例は説明のためのものであって特に断らないかぎり限定のためのものではない。前述した説明はこの開示の利益を有する当

10

20

30

40

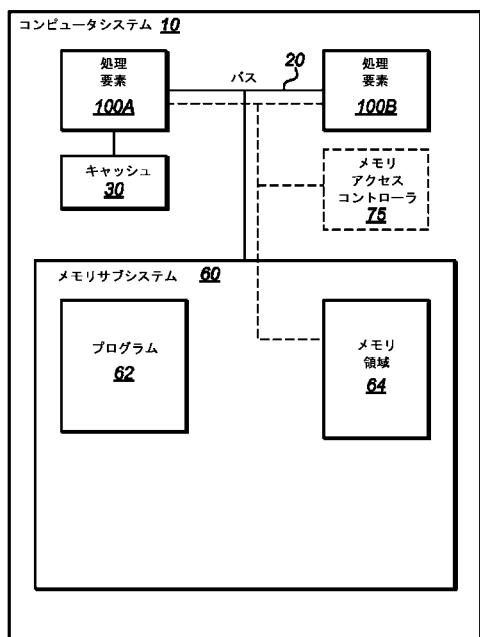
50

業者に自明と考えられるように代替、変更、均等物を包含することを意図したものである。

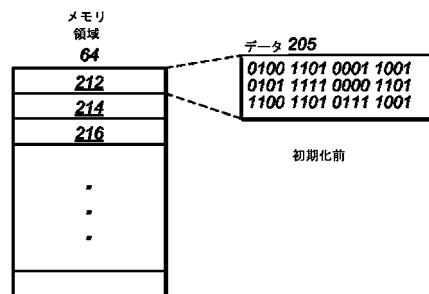
【0060】

本開示の範囲は、ここに検討した問題点のいずれかまたはそのすべてを軽減するかどうかにかかわらず、ここに（明示的もしくは暗示的に）開示したあらゆる特徴、それら特徴の組合せおよびそれらの任意の一般化を含むものである。したがって、このような特徴の組合せを対象としてこの出願（またはこの出願が優先権を主張するところの出願）の審査の間に新たな請求項が構成されてもよい。特に、添付の請求項については、従属項に係る特徴は独立項の特徴と組み合わせることができ、各従属項からの特徴は任意の適切な態様にて組み合わせられてもよく、添付の請求項において具体的に適示された具体的な組合せに限定されるものではない。

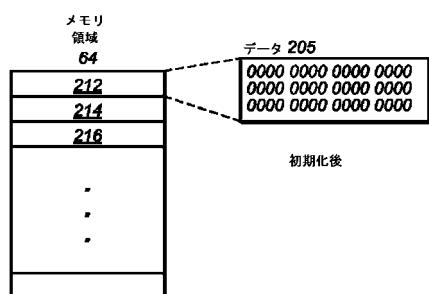
【図1】



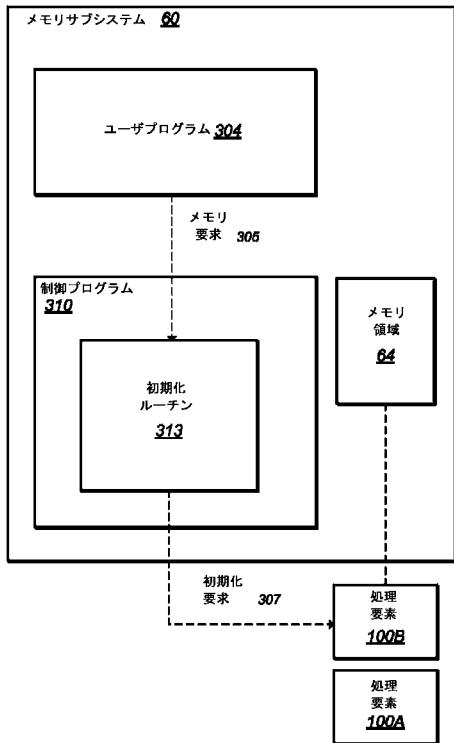
【図2 A】



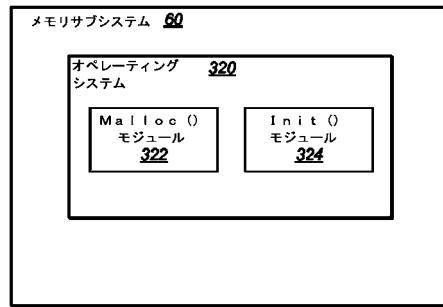
【図2 B】



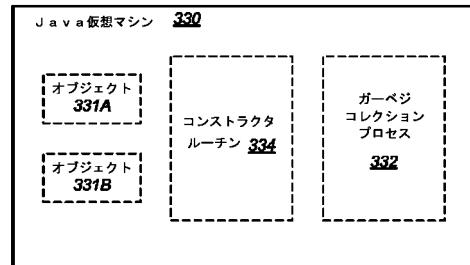
【図3A】



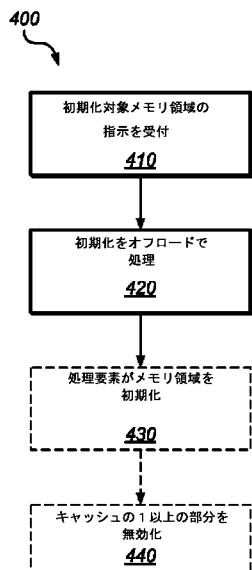
【図3B】



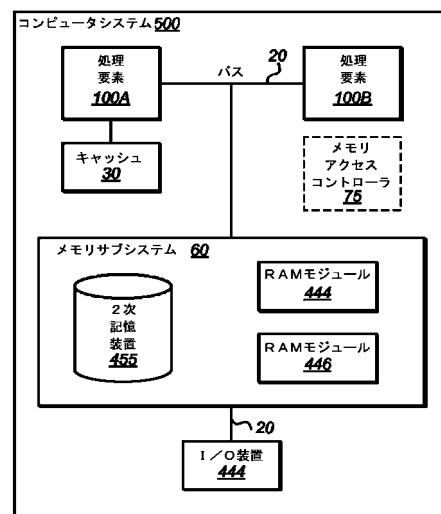
【図3C】



【図4】



【図5】



【手続補正書】

【提出日】平成25年4月2日(2013.4.2)

【手続補正1】

【補正対象書類名】特許請求の範囲

【補正対象項目名】全文

【補正方法】変更

【補正の内容】

【特許請求の範囲】

【請求項1】

コンピュータデバイスにおける初期化されるメモリ領域を示す指示に応答して、該メモリ領域の初期化を該コンピュータデバイスの第2処理要素に処理させるための、少なくとも該コンピュータデバイスの第1処理要素によって実行可能なプログラム命令を記憶する、

非一時的なコンピュータ読み取り可能な媒体。

【請求項2】

前記メモリ領域の指示が、第1プログラムから、前記第1処理要素によって実行される制御プログラムにより受付けられる、

請求項1の非一時的なコンピュータ読み取り可能な媒体。

【請求項3】

前記制御プログラムが、前記第1プログラムを実行している、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項4】

前記指示は、前記制御プログラムで処理可能な1以上のデータオブジェクトに対応したメモリの1以上のメモリ領域を指定しており、

前記1以上のメモリ領域のすべての内容を埋める処理がさらに含まれている、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項5】

前記制御プログラムは、ガーベジコレクションプロセスの一部として、初期化される複数のメモリ領域を示す複数の指示を生成する処理と、前記複数のメモリ領域に対する初期化を前記第2処理要素に行わせる処理とをさらに含み、

前記初期化には、前記複数のメモリ領域のすべての内容を、プログラミング言語の言語仕様にしたがって指定されたデフォルトの内容で埋める処理が含まれる、

請求項4の非一時的なコンピュータ読み取り可能な媒体。

【請求項6】

前記制御プログラムは、前記非一時的なコンピュータ読み取り可能な媒体に記憶された1以上のライブラリファイルを有し、かつ、前記指示を受付ける前記制御プログラムは、アプリケーション・プログラミング・インターフェース(API)を介して前記指示を受付ける前記制御プログラムを有する、

請求項2の非一時的なコンピュータ読み取り可能な媒体。

【請求項7】

第1処理要素と、第2処理要素と、メモリ領域を含むメモリとを備えたコンピュータデバイスにおける初期化される該メモリ領域を示す指示に応答して、前記第1処理要素で実行されている第1プログラムが、該メモリ領域の初期化を前記第2処理要素に処理させる、

方法。

【請求項8】

前記第1処理要素が前記メモリ領域へ直接アクセスすることなく、前記第2処理要素が、ダイレクトメモリアクセス(DMA)を用いて前記メモリ領域を初期化する処理、をさらに含む、

請求項7の方法。

【請求項 9】

前記指示を生成する前記第1プログラム内においてガーベジコレクションプロセスをさらに含む、

請求項7の方法。

【請求項 10】

前記第1プログラムは制御プログラムであって、

前記第2処理要素が、前記制御プログラムの1以上のヒューリスティックな規則にしたがって前記メモリ領域を初期化する処理、をさらに含む、

請求項7の方法。

【請求項 11】

前記コンピュータデバイスが、前記メモリ領域が初期化されるのに応答して、前記コンピュータデバイスのデータキャッシュの1以上の部分であってメモリ領域初期化前の前記メモリ領域の内容に対応する該部分を無効化する処理、をさらに含む、

請求項7の方法。

【請求項 12】

メインメモリを有するメモリサブシステムと、2次記憶装置と、少なくとも第1および第2処理要素と、を有するコンピュータシステムであって、

前記2次記憶装置には、初期化される前記メインメモリのメモリ領域を示す指示に応答して、該メモリ領域の初期化を前記第2処理要素に処理させるように前記コンピュータシステムを動作させるための、前記第1処理要素によって実行可能なプログラム命令、が記憶されている、

コンピュータシステム。

【請求項 13】

前記第1処理要素と、前記第2処理要素とは異機種である、

請求項12のコンピュータシステム。

【請求項 14】

前記第1処理要素と関連付けられたキャッシュをさらに有し、

前記キャッシュは、前記第1処理要素による前記メインメモリへのアクセスに応答して該メインメモリの内容を記憶するように構成され、

前記メモリ領域の初期化が行われたとき、前記メモリ領域の初期化後の内容が前記キャッシュに記憶されない、

請求項12のコンピュータシステム。

【請求項 15】

前記第2処理要素に対して前記メインメモリへの直接アクセスを許可するように構成されたメモリアクセスコントローラをさらに有し、

メモリ領域の初期化には、前記第2処理要素が前記メモリアクセスコントローラを用いて前記メモリ領域にアクセスすることが含まれ、前記第1処理要素が前記メモリ領域にアクセスすることが含まれない、

請求項12のコンピュータシステム。

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No PCT/US2011/046412

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F9/50 ADD. G06F12/02		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, COMPENDEX, INSPEC, IBM-TDB		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	"NVIDIA CUDA Compute Unified Device Architecture, programming guide", , 27 November 2007 (2007-11-27), pages I-XIII, I-128, XP008139068. Retrieved from the Internet: URL: http://developer.download.nvidia.com/compute/cuda/1_1/NVIDIA_CUDA_Programming_Guide_1.1.pdf [retrieved on 2007-11-29] in particular appendices D.5 and E.8; the whole document -----	1-21
<input type="checkbox"/> Further documents are listed in the continuation of Box C.		<input type="checkbox"/> See patent family annex.
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubt on priority, claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search	Date of mailing of the international search report	
12 October 2011	25/10/2011	
Name and mailing address of the ISA/ European Patent Office, P.B. 5618 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax (+31-70) 340-3016	Authorized officer Steinmetz, Christof	

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,R,S,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IS,JP,KE,KG,KM,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LT,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PE,PG,PH,PL,PT,QA,RO,RS,RU,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,ZA,ZM

(74)代理人 100162156

弁理士 村雨 圭介

(72)発明者 エリック アール・カスپール

アメリカ合衆国 94025 カリフォルニア州、メンロー パーク、ウェイブレイ ストリート
311 #2

(72)発明者 ローラン モリケッティ

アメリカ合衆国 95112 カリフォルニア州、サンノゼ、サウス 14番 ストリート 37
7

F ターム(参考) 5B005 KK14 MM01 PP21

5B060 AB30