

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5299869号
(P5299869)

(45) 発行日 平成25年9月25日(2013.9.25)

(24) 登録日 平成25年6月28日(2013.6.28)

(51) Int.Cl.

G06F 9/48 (2006.01)

F 1

G06F 9/46 452 E

請求項の数 4 (全 16 頁)

(21) 出願番号 特願2009-516504 (P2009-516504)
 (86) (22) 出願日 平成19年6月6日 (2007.6.6)
 (65) 公表番号 特表2009-541852 (P2009-541852A)
 (43) 公表日 平成21年11月26日 (2009.11.26)
 (86) 國際出願番号 PCT/US2007/013452
 (87) 國際公開番号 WO2007/149228
 (87) 國際公開日 平成19年12月27日 (2007.12.27)
 審査請求日 平成22年5月20日 (2010.5.20)
 (31) 優先権主張番号 11/471,466
 (32) 優先日 平成18年6月19日 (2006.6.19)
 (33) 優先権主張国 米国(US)
 (31) 優先権主張番号 11/546,072
 (32) 優先日 平成18年10月10日 (2006.10.10)
 (33) 優先権主張国 米国(US)

(73) 特許権者 508371792
 コンデューシブ・テクノロジーズ・コーポ
 レイション
 CONDUSIV TECHNOLOGI
 ES CORPORATION
 アメリカ合衆国、91504 カリフォル
 ニア州、バーバンク、ノース・グレノーク
 ス・ブルバード、7590
 (74) 代理人 100064746
 弁理士 深見 久郎
 (74) 代理人 100085132
 弁理士 森田 俊雄
 (74) 代理人 100083703
 弁理士 仲村 義平

前置審査

最終頁に続く

(54) 【発明の名称】コンピュータマイクロジョブ

(57) 【特許請求の範囲】

【請求項 1】

コンピュータジョブを所有し、実行のために前記コンピュータジョブが提供されるオペレーティングシステムとは別個のプログラムによって、前記コンピュータジョブを複数のマイクロジョブに分割するステップを備え、

前記プログラムは前記複数のマイクロジョブのうちの第1のマイクロジョブのサイズを、前記第1のマイクロジョブを実行するのに必要なリソースの使用のために前記オペレーティングシステムによって前記コンピュータジョブに割当てられたタイムスライスの持続期間に基づいて選択し、さらに、

前記複数のマイクロジョブの各マイクロジョブの実行を要求することによって、前記コンピュータジョブを実行させるステップを備え、

プロセッサを備えるシステムによって実行される、方法。

【請求項 2】

前記複数のマイクロジョブの各マイクロジョブは、前記リソースの所有権が当該マイクロジョブに与えられる実際の割当時間以下の時間で実行される、請求項1に記載の方法。

【請求項 3】

前記各マイクロジョブの実行は、当該マイクロジョブを動作させるのに必要な1つ以上のリソースが1つ以上の遊休基準に準拠しているのに応答して要求される、請求項1に記載の方法。

【請求項 4】

10

20

1つ以上のプロセッサと、
1つ以上のコンピュータプロセッサに通信結合されたコンピュータ読取可能記憶媒体と
を備え、

前記コンピュータ読取可能媒体は、前記1つ以上のプロセッサによって実行されると、
前記1つ以上のプロセッサに請求項1～3の1つ以上に記載の方法を実行させる命令の1
つ以上のシーケンスを格納している、システム。

【発明の詳細な説明】

【技術分野】

【0001】

発明の分野

10

この発明は、コンピュータ環境でソフトウェアアプリケーションを実行することに関する。
特に、この発明の実施例は、アプリケーションのコンピュータまたは入出力ジョブを
マイクロジョブに分割すること、およびそのマイクロジョブを実行することに関する。

【背景技術】

【0002】

背景

多くのマルチタスクオペレーティングシステムにおいて、プロセスはいくつかのスレッド
に細かくされる。スレッドは、オペレーティングシステム(O/S)によって実行され
る1つのコードである。マルチスレッディングの概念は、1つのプロセス内のいくつかの
コード(またはスレッド)を「同時に」動作可能にすることである。たとえば、ワードブ
ロセッサが動作している場合、ユーザは「メニュー項目を探す」をクリックしてポップア
ップボックスを表示することができる。このポップアップは、メインのワードプロセッサ
ウインドウから独立して移動および操作され得る。したがって、ポップアップのためにメ
インのワードプロセッサウインドウが非アクティブになることはない。これが、ワードブ
ロセッサプロセス内で動作する2つの異なるスレッドの例である。

20

【発明の概要】

【発明が解決しようとする課題】

【0003】

マルチタスキングの概念は、単一のコンピュータプロセッサ上で同時に実行されている
複数のコードの見かけを与えるという点でマルチスレッディングと同様である。相違点は
、マルチタスキングはコンピュータ上で動作する1つよりも多いプロセスを指し、マルチ
スレッディングは上記の例のように同一のプロセス内で動作する1つよりも多いスレッド
を指すことである。

30

【0004】

同時に動作する1つよりも多いプロセスまたはスレッドの見かけは、マルチタスクスケ
ジューラが、スレッドを「量子」と称され得る非常に小さな時間増分で動作するようにス
ケジュールする結果である。量子は、スレッドに与えられ、かつその間はそのスレッドが
CPUリソースを所有するタイムスライスである。量子の長さは、最近のオペレーティン
グシステムでは約20ミリ秒から約120ミリ秒の範囲内にある。正確な時間は、O/S
が動作しているハードウェアに依存して異なり得る。さらに、O/Sは、特定のスレッド
に与えられる量子の値を変更することができる。たとえば、スレッドがその第1の量子中
に完了しない場合、O/Sは、そのスレッドの実行がスケジュールされている次回に量子
のサイズを増大または減少させ得る。

40

【0005】

人間の時間感覚と比べて量子の長さが短いため、およびラウンドロビン方法でスレッド
を実行することによって、スレッドは同時に動作するように見える。最近のマルチタスク
O/Sスケジューラはスレッドに優先順位を加え、優先順位の低いスレッドの前に優先順
位の高いスレッドを最適に動作させるためのさまざまなアルゴリズムが存在する。しかし
、すべてのスレッドは即時実行のためにO/Sスケジューラに提示され、O/Sスケジュ
ーラはそれらの優先順位に基づいてできる限り速くすべてのスレッドに実行を完了させる

50

。

【0006】

しかし、この様のスケジューリングの問題点は、コンピュータ性能が予想され得るより悪いということである。しばしば、プロセスは一瞬止まったり、動かなくなったりさえする。たとえば、ユーザ入力に基づいて表示画面に現われるプロセスはしばしば、別のプロセスが消費しているプロセッサ時間が多過ぎるために、ユーザがデータを入力するよう現れることができない。

【0007】

この節で説明された方策は、追求され得る方策であるが、以前に着想または追求された方策であるとは限らない。したがって、特に明記しない限り、この節で説明された方策のいずれも、単にこの節に含まれているからといって先行技術と見なされると仮定されるべきではない。

10

【0008】

この発明は、同様の参照番号が同様の要素を指す添付の図面において、限定的ではなく例示的に図示される。

【図面の簡単な説明】

【0009】

【図1】この発明の実施例に従った、マイクロジョブを実行するためのアーキテクチャの図である。

【図2】この発明の実施例に従った、コンピュータジョブをマイクロジョブに分割するマイクロスケジューラの図である。

20

【図3】この発明の実施例に従った、コンピュータジョブをマイクロジョブに分割するプロセスのステップを図示するフローチャートである。

【図4】本発明の実施例が実現され得るコンピュータシステムを図示するブロック図である。

【発明を実施するための形態】

【0010】

詳細な説明

以下の説明では、説明のために、多数の特定的な詳細が記載されてこの発明の完全な理解を与える。しかし、この発明はこれらの特定的な詳細がなくても実施され得ることが明らかになるであろう。他の場合では、この発明を不要に不明瞭にするのを避けるため、周知の構造および装置がブロック図形態で示される。

30

【0011】

概要

大多数のコンピュータは、常に自身のリソース容量のすべてを利用しているとは限らない。これは典型的に、1日24時間、週7日よく使用されているように見えるコンピュータ、および毎日わずかな間しか立ち上げられないコンピュータについても当てはまる。したがって、コンピュータ時間およびリソースが無駄である。たとえば、24時間にわたって、かなり多用され、かつ活動時に短時間のスパイクを有し得るコンピュータシステムは、平均して約5から20パーセントのリソースしか使用していない可能性がある。

40

【0012】

コンピュータジョブをマイクロジョブに分割することによってこれらの使われていないコンピュータリソースを利用するための方法、システム、および装置が本明細書中に開示される。マイクロジョブは、非常に小さなコンピュータコード（たとえば比較的少数の命令）であり得る。1つの実施例では、マイクロジョブのサイズは、そのマイクロジョブが処理リソースを割当てられた時間内にそのマイクロジョブが実行を終了するようなものである。たとえば、1つの実施例では、マイクロジョブ内の演算の数は、マイクロジョブが量子内に実行を完了するようなものである。マイクロジョブの長さは量子よりも短いことがある。

【0013】

50

1つの実施例では、コンピュータジョブ全体がマイクロジョブに分割され、コンピュータジョブはその後マイクロジョブごとに実行されて、コンピュータジョブ全体が完了する。現在のリソース制約を考慮してコンピュータジョブをできる限り迅速に動作させようと試みること、または他のジョブおよびアプリケーションに重大な影響を与えるのを避けるためにジョブを「時間外」に動作するようにスケジュールすることとは異なり、コンピュータジョブは継続的にコンピュータ上で動作し得るが、あまりに微小であるため、ユーザまたは他のコンピュータジョブには感知不可能であり得る。したがって、ジョブは、ユーザならびに他のジョブおよびアプリケーションにとって完全に透明であり得る。ユーザはジョブをスケジュールする必要はなく、この方法を用いれば、ジョブは性能臨界時間中を含むいつでも動作可能である。

10

【0014】

1つの実施例では、マイクロジョブは、選択された基準に基づいて時々挿入されて実行される。1つの実施例では、マイクロのジョブの実行は、マイクロジョブどうしの間の間隔をあけるための何らかの時間間隔の判断に基づいて、ある時間にわたって分散される。時間間隔は、マイクロジョブの各々の間で同一である必要はない。1つの実施例では、マイクロジョブ実行の基準は、リソース可用性に基づく。たとえば、マイクロジョブを実行するために、そのマイクロジョブが使用する1つ以上のリソースが1つ以上の遊休基準に準拠しているかどうかについての判断が下される。遊休基準が満たされている場合、そのマイクロジョブが実行される。

【0015】

20

マイクロジョブ

1つの実施例では、マイクロジョブのサイズは、特定のマイクロジョブが処理ジョブを実行するために使用されるリソースを所有する割当時間内にその特定のマイクロジョブが完了可能なサイズである。1つの実施例では、各マイクロジョブは、自身の割当時間内に完了するようなサイズである。しかし、マイクロジョブの中には、大きすぎて自身の割当時間内に実行を完了不可能なものもあり得る。

【0016】

30

1つの実施例では、割当時間は量子である。上述のように、量子は、コンピュータコードの一部（たとえばスレッド）に与えられ、かつその間はそのコード部分がCPUリソースを所有するタイムスライスである。これも上述のように、異なるオペレーティングシステムは異なる量子を使用する。さらに、特定のコード部分に割振られる量子は、実行時間中に状況に基づいて変化し得る。たとえば、O/Sは、スレッドに割当てられる量子のサイズを増大または減少させ得る。1つの実施例では、コンピュータジョブは、そのコンピュータジョブに割当てられると予想される量子のサイズに基づいてマイクロジョブに分割される。別の実施例では、コンピュータジョブは、そのコンピュータジョブに割当てられた量子のサイズに基づいてマイクロジョブに分割される。コンピュータジョブのどの部分をマイクロジョブとして分けるべきかについての判断は、実行時間前または実行時間中のいずれかに下され得る。

【0017】

40

マイクロジョブは、1つの実施例に従って、次のマイクロジョブが実行されるまでの実行の一時休止を安全に見込んでおきつつ单一の単位として完了され得る、実質的により小さい（たとえば最小の）作業単位である。実行の一時休止を安全に見込んでおくというのは、特定のマイクロジョブの実行を、すべてのマイクロジョブの実行から生じる結果に影響を及ぼすことなく遅らせることができるという意味である。

【0018】

マイクロジョブはスレッドの一部であり得る。たとえば、スレッドが複数のマイクロジョブに分割され得る。これらのマイクロジョブは、スレッドがスケジュールされる方法と同様にスケジュールされ得る。しかし、上述のように、1つの実施例では、マイクロジョブは、自身が処理リソースを所有する量子または他の期間にわたって実行可能であれば実行を完了することになる。

50

【0019】

マイクロジョブは、どの時点においても非常に少量のリソース（たとえばCPU時間、メモリ割当）しか必要とし得ない。どの時点においてもそのように最小限のリソースしか使用しないため、見えないプロセスになり得る。マイクロジョブを小さく保つことによって、コンピュータジョブが1度に使用するコンピュータリソースが少量で済むようになる。したがって、この発明の1つの実施例に従って、マイクロジョブの実行は、コンピュータシステム内の他のアプリケーションの性能に重大な影響を与えないように十分少量のリソースを消費する。

【0020】

コンピュータジョブのマイクロジョブへの分割

10

I) アプリケーションが自身のコンピュータジョブをマイクロジョブに分割する

1つの実施例では、アプリケーションプログラムが自身のコンピュータジョブを複数のマイクロジョブに分割する。本明細書中で用いられているように、複数という用語は、1よりも大きいいずれかの数を意味する。図1は、この発明の実施例に従った、マイクロジョブを実行するためのアーキテクチャ100の図である。各MJS対応のアプリケーション115(1)～115(n)が、自身のコンピュータジョブをマイクロジョブ125に分割して実行する。たとえば、アプリケーションプログラムがアプリケーションコード内の適切な場所で、マイクロジョブ125を実行するMJS110の許可を要求する呼出しを行うことができ、これが事実上コンピュータジョブをマイクロジョブ125に分割する。例として、コンピュータジョブは、バックアップ、索引付け、ソフトウェア更新、ウイルスおよび有害ソフトスキャン、ならびにフラグメンテーション解消などのメンテナンスを行い得る。しかし、MJS対応のアプリケーション115(1)～115(n)はメンテナンス以外のソフトウェアであってもよい。

20

【0021】

引続き図1を参照して、マイクロジョブスケジューラ(MJS)110は、いつマイクロジョブ125が実行可能であるかを判断する。MJSとともに作動可能なアプリケーションを、本明細書中ではMJS対応のアプリケーション115と称する。この実施例では、MJS110は、特定のMJS対応のアプリケーション115(たとえば115(1))が、1つ以上のマイクロジョブ125を実行可能にするよう要求できるようにするためのアプリケーションプログラムインターフェイス(API)130を有する。API130はまた、以下により十分に説明されるように、MJS対応のアプリケーション115が、マイクロジョブ125がどれだけ分散され得るかを指定できるようにする。例示的なAPIが以下の本明細書中に含まれる。しかし、アーキテクチャ100はこの例示的なAPIに限定されない。

30

【0022】

1つの実施例では、MJS110は、次にどのマイクロジョブ125を実行可能にすべきかを判断できるように、マイクロジョブ待ち行列を維持する。マイクロジョブ125の実行は、他のアプリケーションに対する影響が無視できるものであるようにMJS110によってタイミングが取られる。

40

【0023】

マイクロジョブ125の実行は、アプリケーションとMJSとの間のAPIの呼出しまして他の通信方法において、MJS対応のアプリケーション115によって指定され得る。MJS110が、次のマイクロジョブ125が他のジョブに影響を与えることなく実行可能であると判断すると、MJS110は、MJS対応のアプリケーション115(1)にマイクロジョブ125を実行するよう命令することによって、MJS対応のアプリケーション115に応答する。

【0024】

1つの実施例では、コンピュータリソース利用が監視および分析されて、リソース利用が1つ以上の遊休基準に準拠しているかどうかを判断する。MJSは、1つ以上の遊休基準が満たされるとマイクロジョブ125を実行させる。1つの実施例では、特定のマイク

50

ロジョブ125を実行させるのに必要な時間は量子より小さいかそれに等しいため、そのマイクロジョブ125が使用するいずれのリソースも、別のジョブがそのリソースを必要とする前に放棄される。したがって、マイクロジョブ125によるリソース利用は気づかれないことがある。マイクロジョブ125はそのアプリケーションの環境には見えないことがある。1つの実施例では、MJSは、時間間隔に基づいてマイクロジョブ125をスケジュールする。時間間隔に基づいたスケジューリングが以下に説明される。

【0025】

MJS110はまた、1つの実施例では、メモリマネージャ140も有する。MJS110は初期化されると、オペレーティングシステムによってメモリが割当てられ、そのうちのいくらかは自身の目的のために使用し、そのうちのいくらかはMJS対応のアプリケーション115に割当てる。MJS対応のアプリケーション115(1)は起動すると、MJS110からのメモリ割当てを要求する。MJS110は、すべてのプロセスによる現在のコンピュータシステムメモリ利用およびMJS対応のアプリケーション115(1)のニーズなどの要因に基づいて、どれくらいのメモリをMJS対応のアプリケーション115(1)に割当てるべきかを判断し得る。メモリ要件は、各MJS対応のアプリケーション115に特定的であってもよいし、コンピュータソフトウェアプログラマによってMJS対応のアプリケーション115にプログラムされてもよい。

【0026】

I I) スケジューラがコンピュータジョブをマイクロジョブに分割する
1つの実施例では、MJS110がコンピュータジョブをマイクロジョブ125に分割する。図2を参照して、MJS110は、コンピュータジョブ205をマイクロジョブ125に分割するマイクロジョブ分割ロジック210を有する。MJS110はまた、マイクロジョブ125を実行するためにスケジュールするマイクロジョブスケジューリングロジック220も有する。たとえば、MJS110は、MJS対応でないアプリケーションプログラムに巻付けられたシェルとして作動し得る。したがって、この例では、シェルMJS110は、それを通じて如何なる実行形式のファイルも動作させることができ完全なソフトウェアアプリケーションである。したがって、1つの実施例では、シェルMJS110は、コンピュータジョブを実行形式のファイルからマイクロジョブ125に自動的に分割する。換言すれば、この実施例では、アプリケーションプログラムはアプリケーションをマイクロジョブ125に分割する必要がない。

【0027】

シェルMJS110は、1つの実施例では、リソース利用に基づいてコンピュータジョブを実行形式のファイルからマイクロジョブ125に分割する。シェルMJS110は、アプリケーションおよびそのアプリケーションがどのように動作するかを分析して、そのアプリケーションがどのリソースを使用するかを調べ得る。たとえば、シェルMJS110は、そのアプリケーションがどのリソースを使用するか、およびそのアプリケーションがどの程度までそのリソースを使用するかを分析する。たとえば、ディスクデフラグ用プログラムが動作すると、シェルMJS110は、アプリケーションがどのリソースを使用するか(たとえばCPU、ネットワーク、ディスクI/O)を判断することができる。シェルMJS110は、1つの実施例では、この分析に基づいてどのようにアプリケーションをマイクロジョブ125に分割するかを自動的に判断する。シェルMJS110はまた、この分析に基づいて、どのようにマイクロジョブ125をスケジュールするかも判断し得る。

【0028】

シェルMJS110はさまざまなパラメータを使用して、どのようにコンピュータジョブをマイクロジョブ125に分割するかを判断し得、および/またはどのようにマイクロジョブ125を実行するためにスケジュールするかを判断し得る。これらのパラメータは、ユーザ入力に基づき得るか、またはシェルMJS110によって確立され得る。たとえば、ユーザは、特定のアプリケーションの優先順位が高いように指定し得る。

【0029】

10

20

30

40

50

シェルMJS110は、1つの実施例では、時間間隔に基づいてマイクロジョブ125を実行するためにスケジュールする。

【0030】

1つの実施例では、MJS110はオペレーティングシステムの一部である。この実施例では、オペレーティングシステム内のMJS110がコンピュータジョブをマイクロジョブ125に分割し得る。

【0031】

時間間隔に基づいたマイクロジョブのスケジューリング

1つの実施例では、マイクロジョブ125は時間間隔に基づいて実行される。たとえば、MJS110は時間間隔に基づいてマイクロジョブ125をスケジュールする。たとえば、MJS110はマイクロジョブ125の実行をある時間にわたって分散させる。たとえば、コンピュータジョブは、マイクロジョブ125に分割されていない場合、完了するのに12分かかり得る。しかし、マイクロジョブ125に分割された場合、コンピュータジョブ全体の実行を24時間にわたって分散させることができ、各特定のマイクロジョブ125は2,3秒に1度実行される。

【0032】

特定の例として、量子が20ミリ秒である場合、コンピュータジョブ全体は約36,000量子で完了し得る。したがって、コンピュータジョブは約36,000個のマイクロジョブ125に分割される。マイクロジョブ125のいくつかが量子よりも小さい場合、マイクロジョブ125の数は若干多くなり得る。20ms量子であるとすると、24時間にわたって4,320,000量子が存在する。したがって、MJS110は、120量子ごと、または2.4秒ごとに1つのマイクロジョブ125を実行し得る。この特定の例では、マイクロジョブ対応のアプリケーションは完了するのに24時間かかることになるが、極端に繁忙のシステムでも検出不可能であり得る。

【0033】

事象に基づいたマイクロジョブのスケジューリング

1つの実施例では、マイクロジョブ125は、事象に基づいて実行がスケジュールされる。たとえば、MJS110は、発生するある数の演算に基づいてマイクロジョブ125をスケジュールする。別の例として、MJS110は、発生するある数の量子に基づいてマイクロジョブ125をスケジュールする。異なる量子は異なるサイズであり得る。したがって、1つの実施例では、MJS110は事象に基づいてマイクロジョブ125の実行を分散させる。

【0034】

リソースに基づいたマイクロジョブのスケジューリング

1つの実施例では、マイクロジョブ125はリソース利用に基づいてスケジュールされる。たとえば、1つの実施例では、MJS110はリソースに基づいたスケジューラである。たとえば、MJS110は、マイクロジョブ125が遊休リソースしか利用しないようにマイクロジョブ125をスケジュールする。MJS110は、リソース利用が1つ以上の遊休基準に準拠しているかどうかを判断してスケジューリング決定を下す。例として、MJS110はディスク使用率を分析し得る。マイクロジョブ125を有するアプリケーション以外のアプリケーションがそのディスクを使用している場合、MJS110は、この他のアプリケーションが終わるのを待ってからマイクロジョブ125をスケジュールする。MJS110はディスクI/O利用を監視し続け、ディスクI/Oへのアクセスを求めて他のアプリケーションがなければ、別のマイクロジョブ125をスケジュールできるようにする。しかし、別のアプリケーションがディスクI/Oの利用を求めている場合は、MJS110は別のマイクロジョブ125をスケジュールできないようにして、この他のアプリケーションがディスクI/Oを利用することができる。

【0035】

別の例として、MJS110はネットワーク使用率を分析し得る。ネットワークトラフィックが高すぎる場合、MJS110はトラフィックが遅くなるまでマイクロジョブ12

10

20

30

40

50

5を1つもスケジュールしないことになる。ネットワークトラフィックが十分低い場合は、MJS110はマイクロジョブを実行するためにスケジュールする。MJS110は、ネットワークトラフィックが十分低くあり続けていることを確認し続ける。ネットワークトラフィックが十分低くあり続けている場合、別のマイクロジョブ125がスケジュールされ得る。しかし、トラフィックが高くなりすぎると、さらに他のマイクロジョブ125は実行がスケジュールされない。

【0036】

MJS110は、いずれかの種類のコンピュータリソースおよびリソースのいずれかの組合せに基づいて、リソースに基づいたスケジューリング決定を下し得る。1つの実施例では、MJS110は、実行される許可を待っているマイクロジョブ125の複数の待ち行列を有する。各待ち行列は特定のリソースに対応し得る。たとえば、ディスクI/Oを利用する必要があるマイクロジョブ125の待ち行列、ネットワークを利用する必要があるマイクロジョブ125の待ち行列などがあり得る。リソースの組合せを利用するマイクロジョブ125の1つ以上の待ち行列もあり得る。MJS110は、特定のリソースまたはリソースの組合せが利用可能であればマイクロジョブ125を配布する。特定のマイクロジョブ125は、2つのリソースの使用を必要とし得る。たとえば、特定のマイクロジョブ125は、ネットワークリソースおよびディスクリソースの使用を必要とし得る。しかし、この特定のマイクロジョブ125はCPUリソースを必要としない。CPUリソース利用が現在高いとしても、この特定のマイクロジョブ125をスケジュールおよび実行することができる。

【0037】

MJS対応のアプリケーション115は、この発明の実施例に従って、MJS110にパラメータを送ってリソース利用を制御する。リソース利用の制御は、ディスクI/O、CPUおよびネットワークを含むが、これらに限定されない。たとえば、MJS対応のアプリケーション115は、上記の3つのリソースの閾値レベルのいずれかの組合せが得られるまでは、マイクロジョブ125の実行を要求することができる。さらに、MJS対応のアプリケーション115は、異なるマイクロジョブ125に異なるリソース閾値レベルを指定することができる。たとえば、MJS対応のアプリケーション115は、1つの実施例に従って、各マイクロジョブ125に異なるリソース閾値レベルを指定する。したがって、細分化されたリソース管理が可能である。MJS110がリソース利用を計算する場合、この発明の1つの実施例に従うと、測定されるのは、MJS対応のアプリケーション115以外のプロセスのリソース利用である。CPU利用閾値が20パーセントに設定される以下の例を用いて説明する。MJS対応のアプリケーション115を実行可能にする前にCPU利用が20パーセント未満である場合、CPU利用は、マイクロジョブが実行されると20パーセントより高く増加し得る。この20パーセントを超える増加は、この例ではCPUリソース利用違反とはみなされない。同様の原理がネットワークおよびディスクI/Oリソースに当てはまる。

【0038】

MJS110がオペレーティングシステムの外部で実行される場合、MJS110は、1つの実施例では、自身のリソース利用において自己制限する。たとえば、MJS110は自身のリソース利用を監視し、自身のリソース利用が高くなりすぎると、MJS110は、ある期間の間だけMJS110のスケジューリングを停止するようオペレーティングシステムに要求を出す。

【0039】

プロセスフロー

図3は、この発明の実施例に従った、コンピュータジョブをマイクロジョブ125に分割することによってコンピュータジョブを実行するためのプロセス300のステップを図示するフローチャートである。ステップ302において、コンピュータジョブが開始される。コンピュータジョブは、MJS対応のアプリケーション115からのものであり得る。しかし、コンピュータジョブは、MJS対応のアプリケーション115に関連付けられ

10

20

30

40

50

なくてもよい。

【0040】

例として、ステップ302において、1つの実施例では、コンピュータシステムが立ち上るとMJS対応のアプリケーション115プログラムが起動される。MJS対応のアプリケーション115が行うべきコンピュータジョブを有していない場合、MJS対応のアプリケーション115は、動作すべきコンピュータジョブを有するまで遊休状態にあり続ける。この遊休状態では、MJS対応のアプリケーション115は、時折の監視などのいくつかの機能を果たし得る。ある点で、MJS対応のアプリケーション115は、記憶媒体のデフラグ、またはウイルススキャンなどの行うべきコンピュータジョブを有する。コンピュータジョブは、単一のディスクまたはそれに記憶されるファイルをデフラグすることであり得、MJS対応のアプリケーション115は継続的にそのディスクをデフラグする。10

【0041】

MJS対応のアプリケーション115は、起動されると少量のメモリが割当てられ得る。MJS対応のアプリケーション115は典型的に、1度に単一のマイクロジョブ125しか実行しようとしないため、割当てられる量は非常に少量であり得る。しかし、場合によっては、MJS対応のアプリケーション115は、実行すべき他のプロセスを待たずに複数のマイクロジョブ125を実行しようし得る。たとえば、必要なコンピュータシステムリソースが遊休状態にあるとMJS110が判断した場合、MJS110は、複数のマイクロジョブ125が利用するリソースを利用している別のプロセスがなければ、MJS対応のアプリケーション115がその複数のマイクロジョブ125を続けて実行できるようにし得る。20

【0042】

ステップ304において、コンピュータジョブはマイクロジョブ125に分割される。特定のマイクロジョブ125のサイズは、1つの実施例では、その特定のマイクロジョブ125が処理ジョブを実行するために使用されるリソースを所有する割当時間内にその特定のマイクロジョブ125が完了可能なサイズである。マイクロジョブ125は、マイクロジョブ125の実行が、コンピュータシステム内の他のジョブの性能に重大な影響を与えないよう十分少量のリソースを利用するようなサイズであり得る。1つの実施例では、コンピュータジョブをマイクロジョブ125に分割することは、各マイクロジョブ125が、そのコンピュータジョブにそのマイクロジョブ125を実行するために使用されるリソースの所有権が与えられる割当時間内に実行を完了できるように、マイクロジョブ125のサイズを選択することを備える。30

【0043】

1つの実施例では、コンピュータジョブは、そのコンピュータジョブを所有するアプリケーションによってマイクロジョブ125に分割される。コンピュータジョブのマイクロジョブ125への分割は、MJS対応のアプリケーション115内の命令によって達成され得る。一般に、これらの命令は、MJS対応のアプリケーション115内の決定点である。たとえば、命令は、マイクロジョブ125を実行する許可を要求する、MJS110へのAPIの呼出しであり得る。しかし、MJS110はMJS対応のアプリケーション115に統合されてもよく、その場合命令は、MJS対応のアプリケーション115内のスケジューリング機能への呼出しであり得る。他の技術を用いてコンピュータジョブをマイクロジョブ125に分割してもよい。たとえば、1つの実施例では、コンピュータジョブは、シェルMJS110内のマイクロジョブ分割ロジックによってマイクロジョブ125に分割される。40

【0044】

ステップ306はマイクロジョブ125の実行である。1つの実施例では、コンピュータジョブ全体がマイクロジョブ125に分割され、コンピュータジョブはその後マイクロジョブ125ごとに実行されて、コンピュータジョブ全体が完了する。たとえば、デフラグジョブ全体がマイクロジョブ125に分割され、それらが1度に1つずつ実行されて、50

デフラグジョブ全体が完了する。特定のマイクロジョブ125は、1つの実施例では、割当時間より短い時間で、または割当時間と等しい時間で実行を完了する。

【0045】

マイクロジョブ125ごとの実行は、順次の1つずつの実行に限定されないが、順次の1つずつの実行を含む。複数のマイクロジョブ125が同時に実行され得る。たとえば、複数のCPUがある場合、異なるマイクロジョブ125が異なるCPUで同時に実行され得る。

【0046】

1つの実施例では、マイクロジョブ125のスケジューリングはリソース利用に基づく。この実施例では、マイクロジョブ125の各々について、特定のマイクロジョブ125が使用するコンピュータシステムの1つ以上のリソースの利用が1つ以上の遊休基準を満たすかどうかについての判断が下される。特定のリソースの遊休基準は1つ以上の要因に基づき得る。たとえば、1つの実施例では、CPUリソース利用の遊休基準としてCPU利用が使用される。したがって、マイクロジョブ125が実行されるのは、マイクロジョブ125が必要とするコンピュータシステムのリソースが十分遊休状態にあるときのみである。遊休基準は、1つの実施例では、リソース閾値に基づく。たとえば、MJS対応のアプリケーション115のマイクロジョブ125が、他のプロセスによるリソース利用がそのMJS対応のアプリケーション115が指定する閾値未満である場合にのみ実行されるようなリソース閾値が使用され得る。以下に説明される例示的なAPIは、いくつかのリソース閾値パラメータの一例を含む。しかし、プロセス300はこれらのリソース閾値パラメータに限定されない。

10

【0047】

1つの実施例では、マイクロジョブ125のスケジューリングは時間間隔に基づく。1つの実施例では、コンピュータジョブの実行はある期間にわたって分散される。たとえば、コンピュータジョブは数時間にわたって分散され得る。コンピュータジョブがいくつのマイクロジョブ125に分割されるかに基づいて、どのようにマイクロジョブ125をある時間にわたって分散させるかについての判断が下され得る。連続するマイクロジョブ125どうしの間の時間は均一である必要はない。

20

【0048】

1つの実施例では、コンピュータジョブを開始したアプリケーションプログラムは継続的に動作し、コンピュータシステムが立ち上がっているかぎり動作し続けるが、行わなければならぬ作業がない間は遊休状態にあり続ける。たとえば、ディスクデフラグ用プログラムからウイルス検出アプリケーションプログラムへと動作し続ける。したがって、コンピュータジョブが完了したとしても、アプリケーションプログラムは作業の次のセグメントを待ちながら遊休モードで動作し続ける。したがって、アプリケーションプログラムは、行うべき別のコンピュータジョブがあるときに再起動される必要がない。この結果、アプリケーションプログラムは、アプリケーションプログラムの起動に典型的な付加的なリソースを消費しない。アプリケーションプログラムが、行うべき別のコンピュータジョブを有すると判断した場合、コンピュータジョブはマイクロジョブ125に分割され、マイクロジョブ125はある時間にわたって実行される。たとえば、ディスクデフラグ用プログラムのアプリケーションプログラムは、コンピュータの記憶媒体に対する変更に基づいて付加的な記憶媒体デフラグがなされるべきであると判断し得る。

30

【0049】

例示的なAPI

この発明の実施例は、MJS対応のアプリケーション115がMJS110とインターフェイス接続できるようにするためのAPIである。例示的なAPIは、CPU、ディスク、およびネットワークについて以下のリソース閾値パラメータを有する。

40

【0050】

- ・CPU利用閾値

- ・保留ディスクI/Oカウント閾値

50

・ネットワーク利用閾値

各マイクロジョブ125について上記のパラメータが指定され得る。換言すれば、異なるマイクロジョブ125には異なるリソース閾値パラメータが割振られ得る。たとえば、ネットワークを使用するマイクロジョブ125については、ネットワーク閾値が使用され得る。しかし、ネットワークを使用しないマイクロジョブ125については、ネットワーク閾値はゼロであり得る。したがって、この発明の実施例に従って、細分化されたリソース管理が提供される。

【0051】

特定の例として、MJS対応のアプリケーション115は、CPU利用が50%未満であり、I/Oディスク利用が40%未満であり、ネットワークトラフィックが60%未満である場合にのみ特定のマイクロジョブ125の実行を要求し得る。全く何もなしも含めて、リソース閾値因子のいずれかの組合せが用いられ得る。CPU利用閾値は、この発明の実施例に従って、いずれかの他のジョブの利用閾値とは対照的に、MJSのCPUの使用ごとに異なる。

10

【0052】

以下の2つのパラメータを用いて、リソース利用をどれほど頻繁に測定すべきかを指定する。

【0053】

- ・CPU利用ウィンドウ
- ・ネットワーク利用ウィンドウ

20

CPU利用ウィンドウパラメータは、それにわたるCPU利用が計算されるタイムウィンドウを規定する。たとえば、最新のnミリ秒にわたるCPU利用が平均化される。ネットワーク利用ウィンドウは、それにわたるネットワーク利用が計算されるタイムウィンドウを規定する。これらのパラメータはMJS110に内在し得る。しかし、MJS対応のアプリケーション115はこれらのパラメータをオーバーライドし得る。保留ディスクI/Oはどの時点においても絶対であり、したがって計算する必要はない。

【0054】

必須の遊休時間パラメータがMJS対応のアプリケーション115からMJSに渡されて、どのようにマイクロジョブ125がある時間にわたって分散されるかを制御し得る。必須の遊休時間パラメータは任意である。さらに、使用時には、必須の遊休パラメータの値はゼロであり得る。

30

【0055】

- ・必須の遊休時間

MJS110は、すべてのマイクロジョブ125が実行された後のシステム遊休時間であると規定される「遊休時間」を追跡する。上述のように、MJS対応のアプリケーション115は、MJS110を用いてマイクロジョブ125を待ち行列に入れることができる。MJS待ち行列にマイクロジョブ125がない場合、MJS110は指定された必須の遊休時間だけ待ってから起きて、MJS対応のアプリケーション115が付加的な作業を行なうことを認可する。たとえば、MJS対応のデフラグ用プログラムは、まずある数のマイクロジョブ125を実行してディスクドライブをデフラグし、その後MJS110によって休止され得る。指定された必須の遊休時間の後、MJS110は、MJS対応のデフラグ用プログラムを呼出して付加的な作業を認可する。たとえば、MJS対応のデフラグ用プログラムは、メモリの解放などのクリーンアップジョブを実行し得る。必須の遊休時間は、MJS対応のアプリケーション115によって調整され得るデフォルトパラメータであり得る。

40

【0056】

以下のパラメータは、リソース利用が閾値を上回るときにマイクロジョブ125の実行を待つことに関する。

【0057】

- ・待ち時間

50

・最大待ち時間

リソース利用が現在高すぎてマイクロジョブを実行できないとMJS110が判断した場合、MJS110は、指定された待ち時間だけ待ってから、リソース利用を再確認する。待ち時間パラメータは、MJS110がリソース利用が高すぎると判断するたびに増加され得る。たとえば、MJS110は、最大待ち時間に達するまで待ち時間パラメータを増加させることができる。これらのパラメータは、MJS対応のアプリケーション115が最初に始動されるときにアプリケーション115によって指定され得る。MJS対応のアプリケーション115は、実行時間中にこれらのパラメータを調整することができる。

【0058】

ハードウェア概要

10

図4は、本発明の実施例が実現され得るコンピュータシステム400を図示するブロック図である。プロセス300のステップは、システム400の1つ以上のコンピュータ読取可能媒体の命令として記憶され、コンピュータシステム400のプロセッサ上で実行される。コンピュータシステム400は、情報を通信するためのバス402または他の通信メカニズムと、バス402に結合される、情報を処理するためのプロセッサ404とを含む。コンピュータシステム400はまた、バス402に結合される、プロセッサ404によって実行されるべき情報および命令を記憶するための、ランダムアクセスメモリ(RAM)または他の動的記憶装置などのメインメモリ406も含む。メインメモリ406はまた、プロセッサ404によって実行されるべき命令の実行中に、一時変数または他の中間情報を記憶するために用いられ得る。コンピュータシステム400はさらに、バス402に結合される、プロセッサ404のための静的情報および命令を記憶するための読み出し専用メモリ(ROM)408または他の静的記憶装置を含む。情報および命令を記憶するための、磁気ディスクまたは光学ディスクなどの記憶装置410が設けられ、バス402に結合される。コンピュータシステム400は、プロセッサ400をいくつでも有してよい。たとえば、1つの実施例では、コンピュータシステム400はマルチプロセッサシステムである。プロセッサ404は、コアをいくつでも有してよい。1つの実施例では、プロセッサ404はマルチコアプロセッサ404である。コンピュータシステム400は、ハイパースレッドマシン内で用いられ得る。

20

【0059】

コンピュータシステム400は、コンピュータユーザに情報を表示するための、陰極線管(CRT)などのディスプレイ412にバス402を介して結合され得る。プロセッサ404に情報および指令選択を通信するための、英数字および他のキーを含む入力装置414がバス402に結合される。別の種類のユーザ入力装置は、プロセッサ404に方向情報および指令選択を通信するため、ならびにディスプレイ412上のカーソルの動きを制御するためのマウス、トラックボール、またはカーソル方向キーなどのカーソル制御部416である。この入力装置は典型的に、第1の軸(たとえばx)および第2の軸(たとえばy)の2本の軸における2自由度を有し、これにより、この装置が平面内の位置を指定することができる。

30

【0060】

本発明は、本明細書において説明される技術を実現するためのコンピュータシステム400の使用に関する。本発明の1つの実施例によると、それらの技術は、プロセッサ404がメインメモリ406に含まれる1つ以上の命令の1つ以上のシーケンスを実行するのに応答して、コンピュータシステム400によって行われる。そのような命令は、記憶装置410などの別の機械読取可能媒体からメインメモリ406に読み込まれ得る。メインメモリ406に含まれる命令のシーケンスの実行により、プロセッサ400が、本明細書において説明されるプロセスステップを実行することになる。代替的な実施例では、ソフトウェア命令の代わりに、またはソフトウェア命令と組合せて、配線による回路構成を用いて本発明を実現してもよい。したがって、本発明の実施例は、ハードウェア回路構成およびソフトウェアのいずれの特定的な組合せにも限定されない。

40

【0061】

50

本明細書中で用いられるような「機械読取可能媒体」という用語は、機械に特定的な様で操作させるデータを与えるのに関与するいずれかの媒体を指す。コンピュータシステム400を用いて実現される実施例では、たとえばプロセッサ404に実行用の命令を与える際に、さまざまな機械読取可能媒体が含まれる。そのような媒体は、不揮発性媒体、揮発性媒体、および送信媒体を含む多くの形態を取り得るがこれらに限定されない。不揮発性媒体は、たとえば、記憶装置410などの光学または磁気ディスクを含む。揮発性媒体は、メインメモリ406などの動的メモリを含む。送信媒体は、バス402を備える配線を含む、同軸ケーブル、銅線および光ファイバを含む。送信媒体はまた、電波および赤外線データ通信中に発生されるような音波または光波の形態を取り得る。すべてのそのような媒体は、媒体によって搬送される命令が、機械に命令を読込む物理メカニズムによって検出され得るように具体的である必要がある。10

【0062】

機械読取可能媒体の一般的な形態は、たとえば、フロッピー（登録商標）ディスク、フレキシブルディスク、ハードディスク、磁気テープ、もしくはいずれかの他の磁気媒体、CD-ROM、いずれかの他の光学媒体、パンチカード、紙テープ、孔のパターンを有するいずれかの他の物理媒体、RAM、PROM、EPROM、FLASH-E PROM、いずれかの他のメモリチップもしくはカートリッジ、後で説明されるような搬送波、またはそこからコンピュータが読込可能ないずれかの他の媒体を含む。

【0063】

実行用の1つ以上の命令の1つ以上のシーケンスをプロセッサ404に搬送する際、さまざまな形態の機械読取可能媒体が含まれ得る。たとえば、命令は当初、リモートコンピュータの磁気ディスクで搬送され得る。リモートコンピュータは、自身の動的メモリに命令をロードし、モデムを用いて電話線で命令を送ることができる。コンピュータシステム400にローカルなモデムが電話線上的データを受け、赤外線送信機を用いてそのデータを赤外線信号に変換することができる。赤外線検出器が、赤外線信号内に搬送されるデータを受信することができ、適切な回路構成がそのデータをバス402に乗せることができる。バス402は、データをメインメモリ406に搬送し、そこからプロセッサ404が命令を取出して実行する。メインメモリ406が受けた命令は、プロセッサ404による実行の前または後のいずれかに、記憶装置410に任意で記憶され得る。20

【0064】

コンピュータシステム400はまた、バス402に結合される通信インターフェイス418も含む。通信インターフェイス418は、ローカルネットワーク422に接続されるネットワークリンク420に、双方向データ通信結合を与える。たとえば、通信インターフェイス418は、対応する種類の電話線にデータ通信接続を与えるための統合サービスデジタル網（ISDN）カードまたはモデムであり得る。別の例として、通信インターフェイス418は、互換性のあるLANにデータ通信接続を与えるためのローカルエリアネットワーク（LAN）カードであり得る。無線リンクも実現され得る。いずれのそのような実現例においても、通信インターフェイス418は、さまざまな種類の情報を表わすデジタルデータストリームを搬送する電気、電磁または光学信号を送受信する。30

【0065】

ネットワークリンク420は典型的に、1つ以上のネットワークを通じて他のデータ装置にデータ通信を与える。たとえば、ネットワークリンク420は、ローカルネットワーク422を通じてホストコンピュータ424に、またはインターネットサービスプロバイダ（ISP）426によって操作されるデータ機器に接続を与える。ISP426は次に、現在一般的に「インターネット」428と称される広域パケットデータ通信網を通じてデータ通信サービスを与える。ローカルネットワーク422およびインターネット428は両方とも、デジタルデータストリームを搬送する電気、電磁または光学信号を用いる。さまざまなネットワークを通る信号、ならびにコンピュータシステム400との間でデジタルデータを搬送し合うネットワークリンク420上のおよび通信インターフェイス418を通る信号は、情報を移送する搬送波の例示的な形態である。40

【0066】

コンピュータシステム400は、ネットワーク、ネットワークリンク420および通信インターフェイス418を通じて、メッセージを送り、プログラムコードを含むデータを受信することができる。インターネットの例では、サーバ430が、インターネット428、ISP426、ローカルネットワーク422および通信インターフェイス418を通じてアプリケーションプログラムについての要求コードを送信し得る。

【0067】

受けたコードは受信時にプロセッサ404によって実行され得、および/または後で実行するために記憶装置410もしくは他の不揮発性記憶装置に記憶される。この様態で、コンピュータシステム400は、搬送波の形態のアプリケーションコードを得ることができる。

10

【0068】

上記の明細書では、本発明の実施例が、実現例ごとに異なり得る多数の特定的な詳細を参照して説明された。したがって、発明であるもの、および出願人によって発明であると意図されるものを唯一および独占的に示すものは、いずれの後の訂正も含む、この出願から生じる一組の請求項であり、そのような請求項が生じる特定的な形態におけるものである。そのような請求項に含まれる用語について本明細書中において明示的に記載されるいずれの定義も、請求項において用いられるようなそのような用語の意味を決定するものとする。したがって、請求項において明示的に列挙されていない限定、要素、性質、特徴、利点または属性は、そのような請求項の範囲を如何なる意味でも限定するものではない。

20

明細書および図面はしたがって、限定的ではなく例示的に認識されるべきである。

【図1】

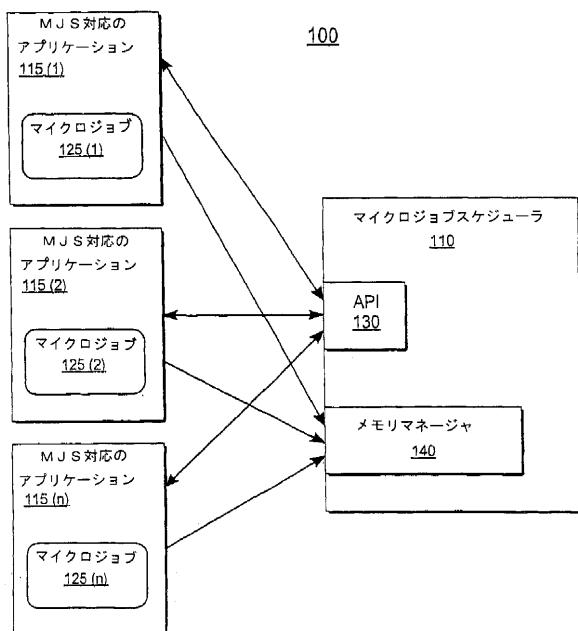


Fig. 1

【図2】

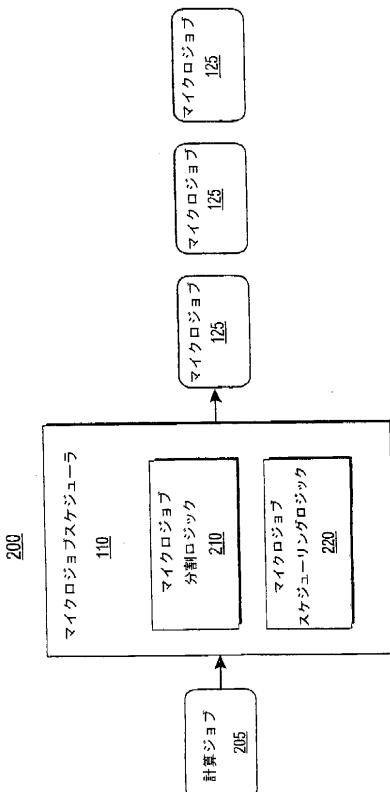


Fig. 2

【図3】

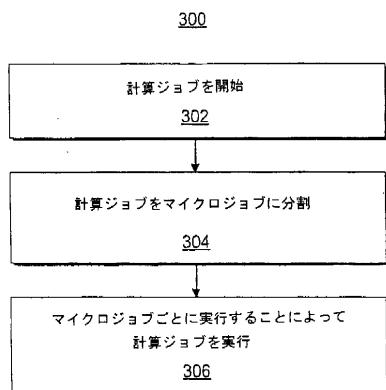
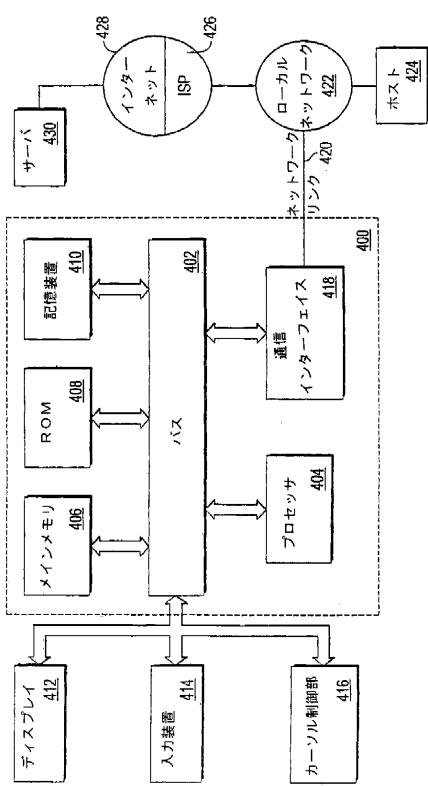


Fig. 3

【図4】



フロントページの続き

(74)代理人 100096781
弁理士 堀井 豊

(74)代理人 100098316
弁理士 野田 久登

(74)代理人 100111246
弁理士 荒川 伸夫

(72)発明者 ジェンセン, クレイグ
アメリカ合衆国、33756 フロリダ州、クリアウォーター、エス・フォート・ハリソン・アベニュ、611、ナンバー・357

(72)発明者 スタファー, アンドリュー
アメリカ合衆国、91342 カリフォルニア州、シルマー、アルタ・ビスタ・ウェイ、13270

(72)発明者 トマス, バジル
アメリカ合衆国、91342 カリフォルニア州、シルマー、ウィロウグリーン・レーン、14537

(72)発明者 キャドルビ, リチャード
アメリカ合衆国、93065 カリフォルニア州、シミ・バレー、ソナタ・ウェイ、530、ナンバー・ディ

審査官 田中 幸雄

(56)参考文献 特開平03-085642 (JP, A)

Tiago Ferreto et al, Scheduling Divisible Workloads Using the Adaptive Time Factoring Algorithm, Lecture Notes in Computer Science, ドイツ, Springer-Verlag, 2005年, Vol. 3719, pages232-239, Distributed and Parallel Computing

(58)調査した分野(Int.Cl., DB名)

G 06 F 9 / 48