

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6313993号
(P6313993)

(45) 発行日 平成30年4月18日 (2018. 4. 18)

(24) 登録日 平成30年3月30日 (2018. 3. 30)

(51) Int. Cl.

F I

G 0 6 F 3 / 0 6 (2006. 01)

G 0 6 F 3 / 0 6 3 0 1 Z

G 0 6 F 1 3 / 1 0 (2006. 01)

G 0 6 F 3 / 0 6 3 0 1 G

G 0 6 F 1 3 / 1 0 3 4 0 A

請求項の数 18 (全 16 頁)

(21) 出願番号 特願2014-33982 (P2014-33982)
 (22) 出願日 平成26年2月25日 (2014. 2. 25)
 (65) 公開番号 特開2014-164769 (P2014-164769A)
 (43) 公開日 平成26年9月8日 (2014. 9. 8)
 審査請求日 平成29年2月24日 (2017. 2. 24)
 (31) 優先権主張番号 13/776, 896
 (32) 優先日 平成25年2月26日 (2013. 2. 26)
 (33) 優先権主張国 米国 (US)

(73) 特許権者 500373758
 シーゲイト テクノロジー エルエルシー
 Seagate Technology
 LLC
 アメリカ合衆国、95014 カリフォル
 ニア州、クパチーノ、サウス・デ・アンザ
 ・ブールバード、10200
 10200 South De Anza
 Blvd Cupertino CA
 95014 United States
 of America

(74) 代理人 110001195
 特許業務法人深見特許事務所

最終頁に続く

(54) 【発明の名称】 ホストのサービス品質インジケータに重みを割り当てるための装置および方法ならびに記憶媒体

(57) 【特許請求の範囲】

【請求項 1】

装置であって、

不揮発性データストレージおよびホストインターフェースに連結することができるコントローラを備え、前記コントローラは、

ホストを介して前記不揮発性データストレージに記憶されたデータに関連する、前記ホストから提供されるサービス品質インジケータを、前記ホストインターフェースを介して受け取ること、

前記サービス品質インジケータに関連するワークロードインジケータを測定すること、

前記サービス品質インジケータと前記測定されたワークロードインジケータとの相関関係に応じて、重みを前記ホストに割り当てること、および

前記ホストからのデータアクセス要求に応じるときに、前記重みを前記サービス品質インジケータに適用すること、を少なくとも行うように構成され、

前記データアクセス要求に応じること、前記重みに基づいて前記ホストからのデータを記憶するための前記不揮発性データストレージの2つ以上のタイプの非階層ストレージから、メモリを選択することを含む、装置。

【請求項 2】

前記2つ以上のタイプの非階層ストレージは、メモリサイズ、待ち時間、スループット、電力消費、保持時間、および信頼性のうちの少なくとも1つに関連する、異なる属性を有することに基づいて選択される、請求項 1 に記載の装置。

10

20

【請求項 3】

前記サービス品質インジケータは、2つを超える値の範囲を含み、前記重みをサービス品質インジケータに適用することは、前記範囲と前記2つ以上のタイプの非階層ストレージとの間のマッピングを調整することを含む、請求項1または請求項2に記載の装置。

【請求項 4】

前記タイプの非階層ストレージの少なくとも1つは、抵抗に基づくメモリまたはスピントルクメモリを備える、請求項1から請求項3のいずれか1項に記載の装置。

【請求項 5】

前記測定されたワークロードインジケータおよびサービス品質インジケータは、関連データが、ランダムであるか順次的であるかを記述する、請求項1から請求項4のいずれか1項に記載の装置。

10

【請求項 6】

前記測定されたワークロードインジケータおよびサービス品質インジケータは、関連データに対して、読み出しアクティビティおよび書き込みアクティビティのうちの少なくとも1つが行なわれると予想されるかどうかを記述する、請求項1から請求項5のいずれか1項に記載の装置。

【請求項 7】

前記測定されたワークロードインジケータおよびサービス品質インジケータは、関連データが、長いデータ保持時間を必要とすると予想されるかどうかを記述する、請求項1から請求項6のいずれか1項に記載の装置。

20

【請求項 8】

前記サービス品質インジケータは、前記ワークロードインジケータに関連しないことが、ホスト提供のメタデータを介して示される、請求項1から請求項7のいずれか1項に記載の装置。

【請求項 9】

前記ホスト提供のメタデータは、ホスト識別子、論理ブロックアドレス範囲、およびファイル識別子のうちの1つ以上を含む、請求項8に記載の装置。

【請求項 10】

コンピュータの動作によって実現される方法であって、

ホストを介して不揮発性データストレージに記憶されたデータに関連する、前記ホストから提供されるサービス品質インジケータを受け取ることと、

30

前記サービス品質インジケータに関連するワークロードインジケータを測定することと、

前記サービス品質インジケータと前記測定されたワークロードインジケータとの相関関係に応じて、重みを前記ホストに割り当てることと、

前記ホストからのデータアクセス要求に応じるときに、前記重みを前記サービス品質インジケータに適用することと、を含む、

前記データアクセス要求に応じて順次データを記憶することは、前記重みに対応する前記ホストからのデータを記憶するための前記不揮発性データストレージの2つ以上のタイプの非階層ストレージから、メモリを選択することを含む、方法。

40

【請求項 11】

前記2つ以上のタイプの非階層ストレージは、メモリサイズ、待ち時間、スループット、電力消費、保持時間、および信頼性のうちの少なくとも1つに関連する、異なる属性を有することに基づいて選択される、請求項10に記載の方法。

【請求項 12】

前記サービス品質インジケータは、2つを超える値の範囲を含み、前記重みを以降のサービス品質インジケータに適用することは、前記範囲と前記2つ以上のタイプの非階層ストレージとの間のマッピングを調整することを含む、請求項10または請求項11に記載の方法。

【請求項 13】

50

前記測定されたワークロードインジケータおよびサービス品質インジケータは、
関連データが、ランダムであるか順次的であるか、
関連データに対して、読み出しアクティビティおよび書き込みアクティビティのうちの
少なくとも1つが行なわれると予想されるかどうか、および
関連データが、長いデータ保持時間を必要とすると予想されるかどうか、のうちの少な
くとも1つを記述する、請求項10から請求項12のいずれか1項に記載の方法。

【請求項14】

前記重みを前記ホストに返信することをさらに含む、請求項10から請求項13のい
ずれか1項に記載の方法。

【請求項15】

請求項10から請求項14のいずれか1項に記載の方法を行うために、プロセッサによ
って操作可能な命令を記憶する、非一時的なコンピュータが読み出し可能な媒体。

【請求項16】

装置であって、
2つ以上の非階層メモリユニットであって、それぞれが異なるタイプの不揮発性メモリ
であり、前記不揮発性メモリの少なくとも1つは、抵抗に基づくメモリを備える、2つ以
上の非階層メモリユニットと、

前記2つ以上のメモリユニットに連結されるコントローラであって、
ホストを介して前記装置に記憶されたデータに関連する、前記ホストから提供されるサ
ービス品質インジケータを記憶すること、

前記記憶されたサービス品質インジケータに関連するワークロードインジケータを測定
すること、

前記記憶されたサービス品質インジケータと前記測定されたワークロードインジケータ
との相関関係に応じて、重みを前記ホストに割り当てること、および

以降のホストのデータアクセス要求に応じて、前記重みを、以降のデータを記憶する
ときに前記2つ以上のメモリユニットの間でデータをどこに記憶するかを選択するための以
降のサービス品質インジケータに適用すること、を少なくとも行うように構成される、コ
ントローラとを備え、

前記データアクセス要求に応じることは、前記重みに基づいて前記ホストからのデー
タを記憶するための前記装置の2つ以上のタイプの非階層ストレージから、メモリを選択す
ることを含む、装置。

【請求項17】

前記2つ以上のメモリユニットは、データアクセス速度、電力消費、保持時間、および
信頼性のうちの少なくとも1つに関連する、異なる属性を有する、請求項16に記載の装
置。

【請求項18】

前記測定されたワークロードインジケータおよびサービス品質インジケータは、
関連データが、ランダムであるか順次的であるか、
関連データに対して、読み出しアクティビティおよび書き込みアクティビティのうちの
少なくとも1つが行なわれると予想されるかどうか、および

関連データが、長いデータ保持時間を必要とすると予想されるかどうか、のうちの少な
くとも1つを記述する、請求項16または請求項17に記載の装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ホストのサービス品質インジケータに重みを割り当てるための装置および方
法ならびに記憶媒体に関する。

【発明の概要】

【課題を解決するための手段】

【0002】

10

20

30

40

50

本開示は、重みをホストのサービス品質インジケータに割り当てることに関する。一実施例において、方法および装置は、ホストを介して不揮発性データストレージに記憶されたデータに関連する、ホストから提供されるサービス品質インジケータを、ホストインターフェースを介して受け取るのを容易にする。サービス品質インジケータに関連するワークロードインジケータが測定され、サービス品質インジケータと測定されたワークロードインジケータとの相関関係に応じて、重みがホストに割り当てられる。重みは、ホストからのデータアクセス要求に応じるときに、サービス品質インジケータに適用される。

【0003】

別の実施例において、装置は、2つ以上の非階層メモリユニットを含む。各ユニットは、異なるタイプの不揮発性メモリであり、不揮発メモリのタイプの少なくとも1つは、抵抗に基づくメモリを含む。この2つ以上のメモリユニットに、コントローラが連結される。このコントローラは、ホストを介して装置に記憶されたデータに関連する、ホストから提供される品質サービスインジケータを記憶することを少なくとも行うように構成される。記憶されたサービス品質インジケータに関連するワークロードインジケータが測定され、記憶されたサービス品質インジケータと測定されたワークロードインジケータとの相関関係に応じて、重みがホストに割り当てられる。以降のホストのデータアクセス要求に応じて、重みが、以降のデータを記憶するときに2つ以上のメモリユニット間で選択するための以降のサービス品質インジケータに適用される。

【0004】

種々の実施形態のこれらのおよび他の特徴および態様は、以下の詳細な議論および添付図面を考慮して理解され得る。

【0005】

以下の図表では、複数の図面の中の類似する/同じ構成要素を識別するために、同じ参照番号が使用され得る。

【図面の簡単な説明】

【0006】

【図1】例示的な実施形態による、装置のブロック図である。

【図2】例示的な実施形態による、ホストおよびワークロードインジケータのデータを示すブロック図である。

【図3】例示的な実施形態に従って、ワークロードインジケータとホスト提供のインジケータとを比較したときに、重みがどのように調整され得るのかを示す表である。

【図4】例示的な実施形態に従って、複数のワークロードインジケータを考慮して、重みがどのように調整され得るのかを示す表である。

【図5】例示的な実施形態による、方法のフローチャートである。

【図6】例示的な実施形態による、方法のフローチャートである。

【図7】例示的な実施形態による、方法のフローチャートである。

【発明を実施するための形態】

【0007】

以下の種々の例示的な実施形態の説明では、本明細書の一部を形成し、種々の例示的な実施形態が実例として示される、添付図面を参照する。他の実施形態が用いられ得ること、また、本明細書に添付される特許請求の範囲から逸脱することなく、構造上および操作上の変更が行われ得ることを理解されたい。

【0008】

本開示は、全体として、固体メモリストレージを使用するデバイス等の、持続的データストレージデバイスに関する。固体メモリストレージとしては、磁気ディスク、フラッシュメモリ（例えば、NANDまたはNORタイプメモリ）、抵抗に基づくメモリ（例えば、抵抗性ランダムアクセスメモリ、相変化メモリ）、およびスピントランスファトルクのランダムアクセスメモリが挙げられるが、それらに限定されない。これらのメモリタイプのそれぞれは、異なる特徴および利点を有し得るが、異なるメモリタイプを使用するメモリデバイスの効果的な使用は、デバイスに記憶されたホストデータと関連付けられるワー

10

20

30

40

50

クロード属性 / インジケータを効果的に特徴付けることに関わり得る。

【 0 0 0 9 】

「ワークロード属性」または「ワークロードインジケータ」という用語は、通常操作中に測定され得る特定のデータユニットと関連付けられる、時間および場所に基づく特徴を指す。例えば、インジケータは、頻度データが読み出されること、頻度データが書き込まれる / 変更されること、保持時間、データへのアクセスが順次的であるかランダムであるか、等の任意の組み合わせに基づいて、データユニットに割り当てられ得る。インジケータは、データユニットに向けられたアクションによってだけでなく、近傍のデータユニットに向けられたアクションによっても決定され得る。

【 0 0 1 0 】

属性が関連するデータユニットは、コンピューティングアーキテクチャのどのような層がデータストレージを管理しているかに基づいて変動し得る。例えば、ホストアプリケーションは、ファイルまたはファイルシステムメタデータに関するデータユニットを確認し得る。下位レベルのホストドライバは、論理ブロックアドレス (L B N A) の範囲に関して、データユニットを確認し得る。デバイスの内部プロセッサは、論理アドレスおよび物理アドレスの双方に関して、データユニットを確認し得る。

【 0 0 1 1 】

本開示は、持続的データストレージデバイスのワークロード属性の特徴付けおよび管理に関する。類似したインジケータは、ホストデバイスから通信され得、ワークロード属性 / インジケータと比較され得る。これは、図 1 で示されるデバイス 1 0 0 等のデータストレージデバイスで使用され得る。

【 0 0 1 2 】

図 1 において、ブロック図は、例示的な実施形態による、データストレージデバイス 1 0 0 を図示する。このデバイス 1 0 0 は、固体メモリの任意の組み合わせを利用する、ソリッドステートドライブ (S S D) (またはその副構成要素) として構成され得る。デバイス 1 0 0 の特徴は、固体メモリおよび磁気ディスクの組み合わせを使用する混合ドライブ等の、他のタイプのハードドライブデバイスに適用可能であり得る。デバイス 1 0 0 の特徴はまた、標準的なハードドライブデータインターフェースを利用しない、特殊用途の固体および / またはディスクデータストレージデバイス (またはその副構成要素) にも適用可能であり得る。例えば、デバイス 1 0 0 は、例えばホスト 1 1 4 の中央処理ユニット (C P U) およびメモリコントローラと緊密に統合される、ホストのメインボード基盤の一部として構成され得る。

【 0 0 1 3 】

デバイス 1 0 0 は、デバイス 1 0 0 の不揮発性メモリのいくつかまたは全部を含む、2 つ以上のメモリユニット 1 0 2、1 0 3 を含む。メモリユニット 1 0 2、1 0 3 は、1 つ以上のそれぞれ別個の物理ユニット 1 0 4、1 0 5、例えばメモリチップを含み得る。この実施例において、メモリユニット 1 0 2、1 0 3 は非階層ユニットであり、それぞれの物理ユニット 1 0 4、1 0 5 は、それぞれが、互いに異なるタイプの不揮発性メモリストレージ媒体を含む。各物理ユニット 1 0 4、1 0 5 内で、メモリは、より小さいブロック 1 0 6、1 0 7 にグループ化され得る。物理ユニット 1 0 4、1 0 5 の下層の媒体が異なるので、ブロック 1 0 6、1 0 7 のメモリサイズは異なり得る。デバイス 1 0 0 の特徴のいくつかは、非階層混合媒体ストレージに適用可能である一方で、この概念はまた、単一のストレージ媒体タイプを使用するデバイスでも利用され得るが、ユニット 1 0 2、1 0 3 間で異なる構成を伴い、メモリサイズ、読み出し待ち時間、書き込み待ち時間、電力消費、電力消費、保持時間、信頼性等のうちの 1 つ以上に関連する、異なる性能属性をもたらす。

【 0 0 1 4 】

デバイス 1 0 0 は、ホストインターフェース 1 1 2 を介してホスト 1 1 4 から受け取られる要求をサービスすることを容易にする、1 つ以上のコントローラ 1 1 0 を含み得る。ホストインターフェース 1 1 2 は、ハードドライブインターフェース、または他のタイプ

10

20

30

40

50

のコンピュータインタフェース（例えば、メモリコントローラインターフェース、周辺バスインタフェース）を含み得る。コントローラ 110 は、一般に、論理アドレスを参照するホスト 114 から、読み出し要求または書き込み要求を受け取り得る。コントローラ 110 は、論理アドレスを物理アドレスに翻訳し、メモリユニット 102、103 の適切な物理アドレスに対してそれぞれの読み出し操作または書き込み操作を行う。

【0015】

メモリユニット 102、103 は、特定のメモリタイプに適切な読み出し / 書き込み操作を行うために、媒体への信号の符号化、復号化、および他のアプリケーションを少なくとも行う、別々のコントローラ（図示せず）を含み得る。別々のコントローラはまた、特定のメモリアーキテクチャに適切な、それら自体の論理 - 物理マッピングも行い得る。そのような場合、主コントローラ 110 は、論理ホストアドレスを、メモリユニット 102、103 によって使用可能な内部論理アドレスに翻訳し得る。

【0016】

デバイス 100 は、数ある中でも、不揮発性メモリユニット 102、103 のための揮発性キャッシュ 118 に使用され得る、揮発性ランダムアクセスメモリ（RAM）116 を含み得る。一般に、揮発性キャッシュ 118 は、不揮発性メモリ 102、103 の一部分をミラーリングするが、不揮発性メモリ 102、103 よりも迅速にそこからの読み出しおよび / またはそこへの書き込みを行うことができる、階層メモリ構造である。いくつかの状況、例えば、短時間に読み取り / 書き込みアクティビティが繰り返されると分かるデータについて、揮発性キャッシュ 118 は、性能を高める。

【0017】

上で述べられるように、メモリユニット 102、103 は、異なるタイプのメモリストレージ媒体を含む、非階層ユニットである。例えば、メモリユニット 102、103 は、それぞれが、フラッシュメモリ、抵抗性 RAM（ReRAM）、スピントルク RAM（STRAM）、または相変化メモリ（PSM）ユニットのうちの異なるものであり得る。コントローラ 110 は、データと関連付けられるワークロード属性に基づいて、ユニット 102、103 のうちの特定の 1 つに記憶するためのデータを選択し得る。例えば、メモリユニット 102、103 は、メモリサイズ、読み出し / 書き込み待ち時間、読み出し / 書き込みスループット、電力消費、保持時間、信頼性等に関して、異なる性能を有し得る。ホスト 114 はまた、記憶されたデータと関連する所望の性能属性に基づいて、ユニット 102、103 に記憶する特定のデータも目標とし得る。所望のデータ属性としては、高速アクセスに対する必要性、書き込み / 更新の頻度、信頼性に対する必要性、長い保持時間、順次的またはランダムなアクセス等が挙げられる。

【0018】

一般に、所望のデータ属性を、データが記憶されるメモリユニット 102、103 の特徴と一致させることが有用である。例えば、頻繁に書き込まれるデータは、高耐久性のメモリタイプに最良に記憶され得る。高速アクセスを必要とするデータは、低い読み出し待ち時間のメモリタイプに最良に記憶され得る。書き込まれる頻度は低い、高い信頼性を必要とするデータは、良好な信頼性および保持特徴を有する、比較的高い待ち時間のメモリタイプに最良に記憶することができる。データは、より良好にまたはより容易に特定のメモリサイズと一致され得る。例えば、ある要求サイズが、メモリページサイズに適合され得るか、または一致していない要求が、読み出し - 修正 - 書き込み操作（例えば、低い待ち時間、高い耐久性）に、より良好に適するメモリに適合され得る。

【0019】

メモリユニット 102、103 は、非階層であるので、階層キャッシュ配置のようにユニット 102、103 間で冗長データを同期させる必要はない。一般に、データを記憶するために、ユニット 102、103 が選択される（例えば、データの論理アドレスが、選択されたメモリユニット 102、103 の物理アドレスにマップされる）と、データは、データが必要とされている限り、そのユニットの中に存在し続け得る。データ属性が再評価され、再評価に基づいてデータが一方のユニット 102、103 からもう一方のユニッ

10

20

30

40

50

トに移動されるときに事例があり得る。再評価は、一定の間隔でスケジュールされたプロセスである場合もあれば、そうでない場合もあり、例えば、ユーザ要求に基づいて、デバイスがかなりのアクティビティを起こす、満杯になる、増加したビットエラーレートを経験する、等のときに起こり得る。再評価中に、データに影響を及ぼす現在のアクティビティが、初期のデータの分類に基づいて予想されるアクティビティと異なる場合、データは、移動され得る。

【0020】

図示される実施形態において、ホスト114は、例えばホストインターフェース114を介してシグナリングすることによって、所望のストレージ属性をデバイス100に報告し得る。このシグナリングは、データが転送されているとき、またはいつかの他のときに生じ得る。本開示の目的で、この所望の属性のシグナリングで使用されるデータは、サービス品質(QoS)インジケータと称される。

【0021】

QoSインジケータは、高速アクセスに対する必要性、書き込み/更新の頻度、信頼性に対する必要性、長い保持時間、順次的またはランダムなアクセス等の、本明細書で説明される所望のストレージ属性の任意の組み合わせを示すために使用され得る。QoSインジケータは、アクセス速度、信頼性、ランダム性等の目標属性を明示的に記述し得る。QoSインジケータは、暗示的であり得、例えば、ワークロードインジケータに関連しない、他のホスト提供のメタデータから導き出され得るが、そこからは、QoS要件が導き出され得る。そのような暗示的QoSインジケータとしては、ファイル識別子、ホスト識別子、論理ブラックアドレス範囲、アプリケーション名等が挙げられるが、それらに限定されない。これらの暗示的インジケータは、履歴的解析、帯域外シグナリング等によって決定され得る。ホスト114は、QoSインジケータを、論理ブロックアドレス(LBA)範囲、ファイル、ディレクトリ、アプリケーション等の、任意の記憶されたデータユニットと関連付け得る。ホスト114は、ホスト要求が論理アドレスに影響を及ぼす前、その間、またはその後、QoSインジケータを論理アドレスと関連付けることによって、インジケータをデバイス100に通信し得る。

【0022】

図1で示されるように、デバイス100は、コントローラ110を介して操作可能な機能モジュールのためのメモリの予約部分120と、デバイスによって内部的に使用される持続的データを記憶するためのデータベース122とを含む。例えば、ホスト属性モジュール124は、デバイス100の操作中に、ホスト114から受け取られるQoSインジケータを監視し得る。ホスト属性モジュール124はまた、QoSデータをデータベース122に記憶し得る。QoSデータは、インデックスとして個々の論理アドレスまたはアドレスの範囲を使用して、データベースに記憶され得る。

【0023】

ホスト供給のQoSインジケータが、データのワークロード要件に関して常時正しい場合、QoSインジケータを記憶する必要はない。そのような場合、ホスト供給のQoSインジケータを使用して、データをどこに記憶するのかを決定することができ、次いで、QoSインジケータを安全に破棄することができる。しかしながら、ホスト供給のインジケータが正しくない場合があり得る。例えば、ホストプログラムまたはオペレーティングシステムは、正しくない推測を行い得ること、正しくないまたは次善のデフォルト値に依存し得ること、等があり得る。

【0024】

データが最初に書き込まれるときに、ホスト供給のQoSデータが正しい場合であっても、ホスト通信プロトコルは、例えば記憶されたデータの以降の読み出しまたは更新中に、QoSを更新するためのプロビジョニングを有しないことがあり得る。その結果、データが最初に記憶されたときにメモリユニット102、103のうちの1つに最適にプロビジョニングされたある量のデータが、経時的なデータ変更の使用につれて、後で準最適にプロビジョニングされる。

10

20

30

40

50

【 0 0 2 5 】

ホスト供給のQoSデータを有効に管理するために、デバイス100は、経時的にデータワークロードアクティビティを追跡する、ワークロード属性モジュール126を含み、アクティビティのインジケータをデータベース122に記憶する。ワークロードは、到来する書き込みデータを監視し、以降の読み出しを測定することによって決定することができる。ワークロードはまた、空間的局所性のための関連付けられるコマンドを時間的に監視することによっても決定され得る。

【 0 0 2 6 】

時間的局所性は、ある時点で参照またはアクセスされ、近い将来に再度参照される可能性がある情報を含む。空間的局所性は、近隣のデータが最近アクセスされていれば、データがアクセスされる可能性がより高いという概念を連想させる。順次データは、複数のコマンドが、特定の空間的局所性に当てはまる時間的局所性を有するときに決定される。これは、2つ以上の異なる組のLBA範囲（またはファイル）が、1組として（例えば、全てのiノード、ファイルおよびiノード、キーおよび値等）まとめてアクセスされるときに生じ得る。ワークロード属性は、時間的に関連付けられたコマンドと、共通の時間的アフィニティを有するグループ化コマンドとを比較することによって決定される、ワークロード関連データセットとして構成することができる。

10

【 0 0 2 7 】

ワークロードデータは、ワークロード属性モジュール126およびホスト属性モジュール124からのデータを比較する、解決/重みモジュール128に利用できるようにされる。解決/重みモジュール128は、特定のデータユニットについてホスト114によって提供されるQoSインジケータと、該データユニットについて実際に測定されたワークロードアクティビティとの間の相違を解決する。これは、ホストQoSインジケータの重みを作成するために使用することができる。

20

【 0 0 2 8 】

重みは、データをメモリユニット102、103のどこに記憶するかを選択するとき、データ割り当てモジュール130によって適用することができる。重みは、測定されたワークロード属性によって補足される「ヒント」とみなされる、QoSインジケータをもたらし得る。別の実施例において、QoSインジケータは、ホスト114からの以前のQoSインジケータの精度に基づいてスケールリングすることができる。単一の重みが、特定のホスト114からの全てのインジケータに適用され得るか、またはいくつかの重みが、異なる細かさのレベル、例えば、QoSインジケータのタイプ、ホストコマンドのタイプ、メモリ範囲、ホストコマンドまたはデータのコンテキスト等で使用され得る。

30

【 0 0 2 9 】

重みは、異なるメモリタイプ間でのより良好なデータ移送/割り当ての際にホスト114を補助するために、ホストインターフェース112を介してホスト114に返信することができる。重みは、データに対する予想された変更のため、データのマッピング状態に対する変更が行われない、または以前の内部追跡データがデータセットのトリミング時にクリアされる、データのトリミング等の他のヒントに適用され得る。重み付けは、メモリユニット102、103間でのデータのバックグラウンドの再割り当てに応じて行われ得る。バックグラウンドのプロセスは、時として、既存のデータについてQoS値を再確認し、リソースのより良好な使用をもたらす場合に、データを再割り当てすることができる。例えば、ホスト供給のQoSは、それが頻繁にアクセスされることを示唆しているが、ワークロードインジケータは、それが長期間にわたってアクセスされていないことを示す場合、そのようなプロセスは、データの格下げを生じさせ得る。

40

【 0 0 3 0 】

本明細書で説明される実施形態の1つの態様は、異なる抽象化層（ホストファイル、デバイス、LBA、コントローラ、物理アドレス）が必要とされ得ないことである。例えば、デバイス100およびコントローラ110は、ファイルまたはオブジェクトのより抽象的な概念によって作動することができ、したがって、システムの中の層の数を低減させる

50

ことができる。ホスト 1 1 4 とストレージデバイス 1 0 0 との間のこの種の通信の概念は、デバイス 1 0 0 が、実際にはホスト 1 1 4 の一部とみなされること、例えば、オペレーティングシステムおよび / またはカーネルレベルのアクセスが (例えば、インテリジェント D R A M コントローラのような) デバイスコントローラ 1 1 0 によって直接理解されることであり得る。そのような場合、サービス品質インジケータの通信は、要求の性質によって暗示され得る。例えば、オペレーティングシステムのページングが高速ストレージを獲得し得ること、アプリケーションのページングが低速ストレージを獲得し得ること、動的リンクライブラリ (D L L) のローディングが中速を獲得すること等である。

【 0 0 3 1 】

以下、図 2 を参照すると、ブロック図は、例示的な実施形態による、属性データセット 2 0 0 の比較を図示する。ブロック 2 0 2 ~ 2 0 5 は、4 つの異なるアドレス (例えば、L B A) について収集されたメタデータを表す。ブロック 2 0 2 ~ 2 0 5 のメタデータは、個々のアドレスおよび / またはアドレスのブロックと関連付けられるアクティビティを記述し得る。各ブロック 2 0 2 ~ 2 0 5 のデータは、2 つのカテゴリ、例えば、ブロック 2 0 2 で示されるように、ホストインジケータ 2 0 2 A およびワークロードインジケータ 2 0 2 B に分割される。

【 0 0 3 2 】

ホストインジケータ 2 0 2 A は、この実施例において速度 (例えば、待ち時間、スループット)、保持、およびデータがランダムであるか否かに関する、ホストによって提供される Q o S インジケータを含む。これらのインジケータの最初の 2 つは、2 ビットの数 (0 ~ 3) によって指定され、ランダム性は、単一のビット、すなわち、ランダムの場合は「 1」、順次的な場合「 0」である。この実施例では、ホストとストレージデバイスとの間の通信オーバーヘッドを低減させるために、ホストインジケータが最小のワードサイズを使用すると予想される。しかしながら、本実施形態は、ホストまたはワークロードインジケータの任意の特定の形態またはワード長に限定されない。

【 0 0 3 3 】

ワークロードインジケータ 2 0 2 B は、ストレージデバイスによって測定される。ワークロードインジケータ 2 0 2 B も、ホストインジケータ 2 0 2 A との比較を簡単にするために、2 ビットの数として示される。実際には、ワークロードインジケータは、比較的大きなデータセットから数学的に導き出され得るので、ワークロードインジケータ 2 0 2 B は、より高い精度 (例えば、8 ビット以上の数) を使用して表され得る。そのような場合、ワークロードインジケータ 2 0 2 B は、ホストインジケータ 2 0 2 A の利用可能な値に適合する範囲に分割することができる。

【 0 0 3 4 】

この実施例においてデバイスによって測定されるワークロードインジケータは、一般に、読み出し / 書き込みの回数および / または読み出し / 書き込みの時間的近さの尺度である、「読み出しホットネス」および「書き込みホットネス」を含む。高い値の読み出しホットネスおよび書き込みホットネスは、データブロックがかなりの量のアクティビティを受けることを示し、これは、どのくらい前にアクティビティが発生したのか従って重み付けされる場合もあれば、そうでない場合もある。キャッシュは、一般に、最近の読み出しまたは書き込みアクティビティを追跡することだけしか必要としないので、これは、読み出しキャッシュまたは書き込みキャッシュがアクティビティをどのように追跡し得るのかということとは異なり得る。しかしながら、アクティビティが比較的低い頻度で (例えば、システムのブート中にだけ) 起こる場合であっても、性能的な目的で、繰り返しの集中的なアクティビティを経験する、あるデータユニットへの高速アクセスを提供することが望ましくなり得る。

【 0 0 3 5 】

図 2 の実施例のデバイスはまた、ランダム性ワークロードインジケータによって示されるように、アドレス範囲に対する読み出しがどのくらい頻繁に同じ順序で生じるのかも測定する。順次データユニット (例えば、メディアファイル) は、通常、予測可能な順序で

10

20

30

40

50

読み直される。ハードドライブのメタデータ（例えば、ジャーナル、ファイルアロケーションテーブル）等の他のデータは、ランダムな順序で書き込まれ、読み出され、および／または更新され得る。あるデータストレージタイプおよび関連アクセスアルゴリズムは、ランダムなデータに関して比較的より良好に機能し得、他のものは、順次データに関して比較的より良好に機能し得る。

【 0 0 3 6 】

図 1 の解決 / 重みモジュール 1 2 8 等の解析モジュールは、ホストインジケータ 2 0 2 A をワークロードインジケータ 2 0 2 B と関連させるように構成される。例えば、システム性能は、一般に、高頻度または低頻度で集中的なアクティビティを経験するデータが、より高速なストレージ媒体に配置される場合に向上し得る。このように、速度 Q o S インジケータは、読み出し / 書き込みホットネスのワークロードインジケータに関連させることができる。インジケータ 2 0 2 A、2 0 2 B の間には、負の相関関係があり得る。例えば、長期の保持は、頻繁に書き直されるデータにとって優先事項となり得ないので、Q o S インジケータ 2 0 2 A における保持は、ワークロードインジケータ 2 0 2 B の書き込みホットネスに逆相関し得る。この例は、仮想メモリであり、それは、短期間だけ依存され得る。

【 0 0 3 7 】

ブロック 2 0 2 において、ホストインジケータ 2 0 2 A のいくつかは、いくつかの測定されたワークロードインジケータ 2 0 2 B と良好に相関しないと思われる。例えば、ホストは、最も高い速度のインジケータを示したが、読み出しホットネスおよび書き込みホットネスは、データが最初に記憶されてから、かなりの読み出し / 書き込みアクセスが見られなかったことを示す。しかしながら、データのランダム性は、ホストによって示されるものと相関すると思われるが、この場合はデータがほとんど使用されていないと思われるので、このパラメータは、性能にあまり影響を及ぼし得ない。

【 0 0 3 8 】

ブロック 2 0 3 および 2 0 4 における関連するホストインジケータは、測定されたワークロードインジケータ、例えば高い / 低い読み出し / 書き込みホットネスのワークロードインジケータと適合する高い / 低い速度の Q o S インジケータと密に相関すると思われる。ブロック 2 0 5 における読み出しおよび書き込みホットネスのインジケータは、速度および保持のためのホストインジケータに矛盾すると思われる。その結果、この少数のサンプルについて、ホストは、ランダム / 順次データを良好に分類していると思われるが、アクセス速度に対する必要性を過大評価し、データ保持に対する必要性を過小評価している可能性がある」と結論され得る。このように、ストレージシステムは、将来、データを異なるメモリユニットに割り当てるときに、より少ない重みをホスト提供の速度および保持インジケータに与える傾向があり得る。

【 0 0 3 9 】

以下、図 3 を参照すると、表 3 0 0 は、例示的な実施形態に従って、重みがシステムにどのように適用され得るのかを図示する。表 3 0 0 は、ストレージデバイスによって独立して評価することができる、ホストによってストレージデバイスに提要される任意のメトリックまたはインジケータに適用可能であり得る。縦軸は、ホスト提供のインジケータを示し、横軸は、測定されたワークロードインジケータを示す。この実施例において、インジケータは、横軸および縦軸上の「低」および「高」という表記から分かるように、ある値の範囲を有し得る。

【 0 0 4 0 】

図中の 4 つの領域 3 0 2 ~ 3 0 5 は、ホスト提供のインジケータを、測定されたワークロードインジケータとどのように比較するのかに応じて、ホストインジケータが、どのように重み付けされ得るのかを図示する。領域 3 0 3 および 3 0 4 は、測定されたワークロードインジケータがホスト提供のインジケータと一致する場所を示す。領域 3 0 2 は、測定されたワークロードインジケータと比較して、ホストインジケータが高い場所を示し、領域 3 0 4 は、測定されたワークロードインジケータと比較して、ホストインジケータが

低い場所を示す。

【 0 0 4 1 】

領域 3 0 2 ~ 3 0 5 は、ホスト提供のインジケータの精度の一般的なインジケータを提供し得るが、重みのそれぞれの増加または減少は、各領域 3 0 2 ~ 3 0 5 に対して均一に適用される必要はない。例えば、ホスト提供のインジケータの大部分が領域 3 0 4 と 3 0 5 との間で等しく分割される場合、これは、ホストが圧倒的に低く評価していることを示し得る。領域 3 0 4 および 3 0 5 によって示される重みの変化が等しく適用された場合、ホストは、中間の重みを有することになる。しかしながら、領域 3 0 4 における正しい指示は、偶然に起因し得、それは、領域 3 0 4 の中にあるインジケータが重みに対して及ぼす影響がより少ない（または全く影響を及ぼさない）ことが望ましくなり得る。類似する解析を、領域 3 0 3 および領域 3 0 2 の中のインジケータに適用し得る。

10

【 0 0 4 2 】

ホストが一貫して高いまたは低いと評価する場合、これはまた、Q o S インジケータを異なるストレージユニットにマップするときに、ホストインジケータがどのように翻訳され、適用され得るかという指示も提供し得る。図 2 による例示的な速度メトリックを使用すると、ホストが適度に正確であり、異なる速度属性を伴う 2 つのメモリユニットタイプがあった場合、0 または 1 のホスト提供の速度メトリックを有するデータは、より遅い方のメモリユニットにマップされる。2 または 3 のホスト提供のメトリックを有するデータは、より速い方のメモリユニットにマップされる。

20

【 0 0 4 3 】

上で説明されるマッピングは、ホストの重みに基づいて調整され得る。例えば、ホストのインジケータが一貫して領域 3 0 2 および 3 0 3 の中にある（ホストが、速度に対する必要性を過大評価している）場合、0 ~ 2 のホスト提供のメトリックを有するデータが、より遅い方のメモリユニットに適用され、3 のホスト提供のメトリックを有するデータが、より速い方のメモリユニットに適用される。一方で、ホストのインジケータが一貫して領域 3 0 4 および 3 0 5 の中にある（ホストが、速度に対する必要性を過小評価している）場合、0 のホスト提供のメトリックを有するデータが、より遅い方のメモリユニットに適用され、1 ~ 3 のホスト提供のメトリックを有するデータが、より速い方のメモリユニットに適用される。

30

【 0 0 4 4 】

ホストが 1 つを超える Q o S メトリックを提供する場合、ホスト提供のインジケータに対する重みおよび修正は、各メトリックを単独で調べる必要がないことが認識されるであろう。一実施例として図 2 によるメトリックを使用すると、ホストは、速度インジケータが低いときはいつも正確な保持インジケータを提供し得るが、速度インジケータが高いときにはいつも不正確な保持インジケータを提供し得る。その結果、データを記憶する場所に関する決定は、図 4 にあるような、統合決定表を使用し得、該図は、例示的な実施形態による、統合決定表 4 0 0 を示す。

【 0 0 4 5 】

図 4 において、2 つ以上のメモリユニットタイプのうちの 1 つに 1 つのユニットを記憶するという決定は、2 つのホスト提供のインジケータ I 1 および I 2 に依存し得る。Q o S 値 Q は、 $Q = W \times (C_1 \times I_1 + C_2 \times I_2)$ 使用して見出され得、式中、C 1 および C 2 は、スケーリング定数である。Q_T を閾値とすると、Q > Q_T である場合、データは、一方のメモリユニットに配置され得、Q < Q_T である場合、データは、もう一方のメモリユニットに配置され得る。これは、複数の閾値 Q_T を使用することによって、任意の数のメモリユニットに拡張され得る。重み W は、ホストインジケータが表 4 0 0 の四半部 4 0 2 ~ 4 0 5 のどこに入るかに依存して調整され得る。これは、 $Q = W_1 \times C_1 \times I_1 + W_2 \times C_2 \times I_2$ に設定することによって、さらに精緻化され得る。そのような場合、W₁ および W₂ のそれぞれは、ホストインジケータが入る表 4 0 0 の四半部 4 0 2 ~ 4 0 5 に基づいて、独立に決定 / 調整される。

40

【 0 0 4 6 】

50

表 3 0 0 および表 4 0 0 は、さらなる仕切りを含むように修正され得、そのような仕切りは、互いに対して直線的または非直線的にサイズ決定され得ることが認識されるであろう。表 3 0 0 および表 4 0 0 はまた、「低」/「高」の表記が「0」/「1」、「真」/「偽」等と置き換えられる、2 進数のホスト提供のインジケータまたはワークロードインジケータにも使用され得る。そのような 2 進数のインジケータは、一方または双方の軸に使用され得、前者の場合、多値軸とともに使用され得る。

【 0 0 4 7 】

以下、図 5 を参照すると、フローチャートは、例示的な実施形態による手順を図示する。フローチャートは、5 0 0 で、記憶場所に影響を及ぼす要求、例えば、読み出し、書き込み、更新をホストから受け取ることによって引き起こされる。記憶場所は、1 つ以上のアドレス（例えば、隣接する範囲）を含み得る。この要求は、決定ブロック 5 0 2 で示されるように、Q o S インジケータを含む場合もあれば、含まない場合もある。読み出し等のいくつかの要求は、Q o S インジケータを含み得ないが、それでも、内部メトリックを追跡する目的で重要であり得る。要求が Q o S インジケータを有する場合、5 0 4 で、それがデータベースに加えられ、5 0 6 で、メトリックカウンタが 1 つ増加される。

【 0 0 4 8 】

メトリックカウンタは、時として、ホスト供給の Q o S インジケータの重みの更新を引き起こすことが意図されるが、経過時間等の他のトリガーも使用され得る。決定ブロック 5 0 8 で試験されたときに、メトリックカウンタが閾値を超えた場合、5 1 0 で、図 6 でさらに詳細に示される機能を使用して調整される。5 1 0 で重みが調整された後に、カウンタは、所望の数の操作が再度生じるまで 5 1 0 での調整が繰り返されないように、5 1 2 でリセットされる。

【 0 0 4 9 】

ブロック 5 0 2 で Q o S インジケータが見つかったか否かに関わらず、要求は、それでも、5 1 4 で、記憶場所に関する内部ワークロードインジケータを更新するために使用され得る。例えば、Q o S インジケータが、最初にデータが書き込まれるときにだけ使用される場合、以降の読み出しまたは更新操作は、Q o S インジケータを有しないが、それでも、5 1 4 でワークロードインジケータを更新するのに有用である。更新されたワークロードインジケータは、データが書き込まれたときに提供されたアドレスと関連付けられる Q o S インジケータと比較することができる。ワークロードインジケータは、キャッシングおよびガーベージコレクション等の他の操作にも使用され得る。

【 0 0 5 0 】

決定ブロック 5 1 6 で、イベントが新しいストレージ場所を画定する（例えば、新しいデータを書き込む）ことを要求すると決定された場合、サービス品質インジケータの重み付けした値に基づいて、5 1 8 で、データを記憶する適切なメモリユニットが決定される。Q o S インジケータが、新しいデータを書き込むことに対する要求にも随意である配設では、5 1 8 で、重み付けした Q o S の代わりに、ワークロードインジケータ（例えば、時間またはアドレスが現行データと関連する他のデータと関連付けられるインジケータ）、またはデフォルト値等の、他のデータが使用され得る。Q o S が提供される場合であっても、ホストの重みが十分に低い場合、デバイスは、Q o S インジケータの代わりに、ワークロードインジケータまたはいくつかの他の属性を使用し得る。

【 0 0 5 1 】

要求は、5 2 0 で、ルーチンを出る前に満たされる。5 2 0 で要求を満たすことは、1 つまたは複数の目標アドレスの読み出し、書き込み、更新を含み得る。これらの操作の順序が変更され得ること、およびいくつかの操作が並列に行われ得ることが認識されるであろう。例えば、要求は、5 2 0 で、ルーチンの開始時に満たされ得、他の操作は、並列に行われ得るか、または後で実行するようにキューに入れられ得る。

【 0 0 5 2 】

以下、図 6 を参照すると、フローチャートは、例示的な実施形態による、ホストの重みを調整するために使用される、図 5 による調整機能 5 1 0 の実施例を図示する。手順は、

10

20

30

40

50

ループ端ブロック 600 で示されるように、ホスト提供の属性を有する全てのアドレスを通してループすることを含む。例えば、ホスト提供の QoS 属性が、LBA によってインデックスが付けられたデータベース（例えば、図 1 のデータベース 122）に記憶されている場合、ループ 600 を通して繰り返すために、LBA によって記憶された全ての記録に対するクエリーが使用され得る。

【0053】

602 で、ループ 600 の各アドレスについて、ホスト提供の QoS インジケータが見つけれられる。604 で、同じアドレスについて、測定されたワークロードインジケータが見つけれられる。606 で、ワークロードインジケータとの比較に基づいて、ホスト提供の QoS インジケータの精度が決定される。この比較は、ルックアップテーブルに基づいて、または当技術分野で知られる他の相関関数を使用して、数字で示される。608 で、この比較に基づいて、ホストの重みが調整される。重みは、総じてホストに適用可能であり得、または異なるホスト提供の QoS インジケータおよび / または QoS インジケータの組み合わせに対する、1 つを超える別々の重みを含み得る。ループ 600 を出た後に、610 で、1 つ以上の調整されたまたは新しい重みが記憶される。

10

【0054】

608 で重みを調整することは、既存の重みを修正すること、または新しい重みを作成することを含み得る。例えば、ループ 600 を通した各繰り返しは、累算器に対して加算または減算することができる。610 で、ループは出たが、重みが記憶される前に、累算器は、適切な値、例えば 0 と 1 との間に平均化し、スケーリングすることができる。このスケーリングされた値は、存在していれば以前の重みと組み合わせる（例えば、平均化する）ことができ、または任意の以前の重みを置き換えることができる。

20

【0055】

以下、図 7 を参照すると、フローチャートは、例示的な実施形態による方法を図示する。この方法は、702 で、ホストから提供されるサービス品質インジケータを、（例えば、ホストインターフェースを介して）受け取ることを含む。サービス品質インジケータは、ホストを介して不揮発性データストレージに記憶されるデータに関連する。

【0056】

702 で、サービス品質インジケータに関連するワークロード属性が測定され、704 で、属性と監視されたワークロード属性との相関関係に応じて、重みがホストに割り当てられる。測定されたワークロードインジケータおよびサービス品質インジケータは、a) 関連データがランダムであるか、順次的であるか、b) 関連データが、かなりの読み出しアクティビティおよびかなりの書き込みアクティビティのうちの少なくとも 1 つを経験すると予想されるかどうか、c) 関連データが、長いデータ保持時間を必要とすると予想されるかどうか、のうちの少なくとも 1 つを記述し得る。

30

【0057】

706 で、以降のホストのデータアクセス要求に応じるときに、重みがサービス品質インジケータに適用される。例えば、重み付けした以降のサービス品質インジケータに応じて、以降のデータを記憶するメモリが、不揮発性データストレージの 2 つ以上のタイプの非階層ストレージから選択され得る。2 つ以上のタイプの非階層ストレージは、例えば、メモリサイズ、読み出し待ち時間、書き込み待ち時間、電力消費、保持時間、および信頼性のうちの少なくとも 1 つに関連する、異なる属性を有することに基づいて選択され得る。重みは、ホストに返信して、例えば、以降のデータの分類、および QoS インジケータの使用において、ホストを支援し得る。

40

【0058】

別の配設において、サービス品質の測定値は、2 つを超える値の範囲を含み得る。そのような場合、重みを以降のサービス品質インジケータに適用することは、範囲と 2 つ以上のタイプの非階層ストレージとの間のマッピングを調整することを含む。2 つ以上のタイプの非階層ストレージの少なくとも 1 つは、抵抗に基づくメモリ（例えば、ReRAM、PCM）を含み得る。

50

【 0 0 5 9 】

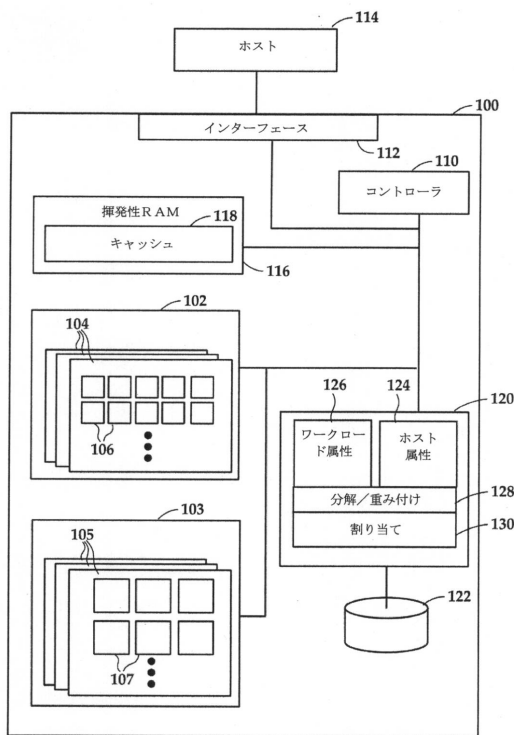
上で説明される種々の実施形態は、特定の結果を提供するように相互作用する回路および/またはソフトウェアモジュールを使用して実現され得る。コンピューティング技術の当業者は、モジュールレベルで、または総じて当技術分野で一般に知られている知識を使用して、そのような説明される機能を容易に実現することができる。例えば、本明細書で図示されるフローチャートは、プロセッサによって実行するためのコンピュータが読み出し可能な命令/コードを作成するために使用され得る。そのような命令は、コンピュータが読み出し可能な媒体に記憶され、当技術分野で知られているように実行するためのプロセッサへ転送され得る。上で示される構造および手順は、上で説明されるデータストレージデバイスにおけるキャッシングの管理を容易にするために使用することができる実施形態の代表例に過ぎない。

10

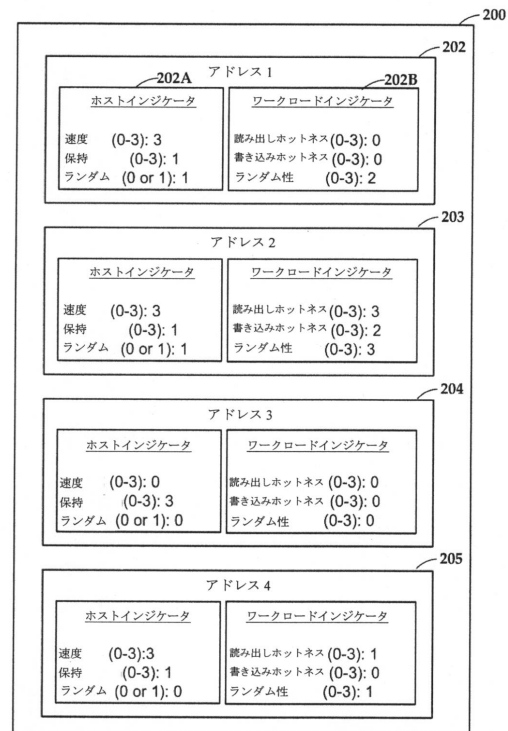
【 0 0 6 0 】

例示的な実施形態の前述の説明は、例示および説明の目的で提示したものである。本記述は、網羅的であること、または本発明の概念を開示された形態に限定することを意図したものではない。上述の教示に照らして、多数の修正および変形が可能である。開示される実施形態の特徴のいずれかまたは全ては、個々にまたは任意の組み合わせで適用することができ、限定することを意味するものではなく、単なる事例である。範囲は、この詳細な説明によって限定されるのではなく、むしろ本明細書に添付される特許請求の範囲によって限定されることを意図する。

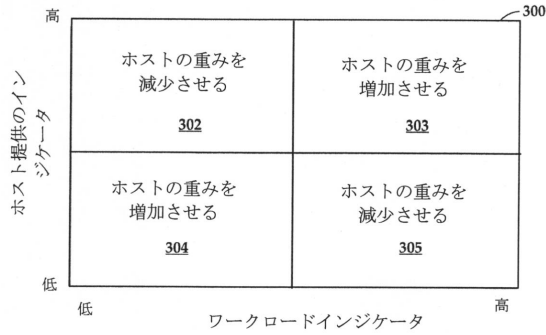
【 図 1 】



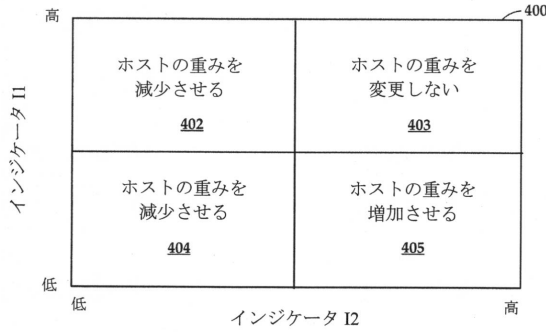
【 図 2 】



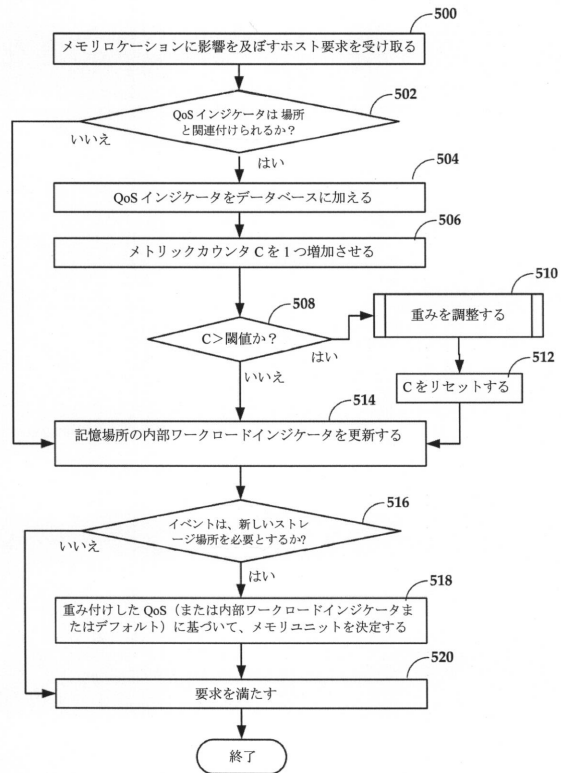
【図 3】



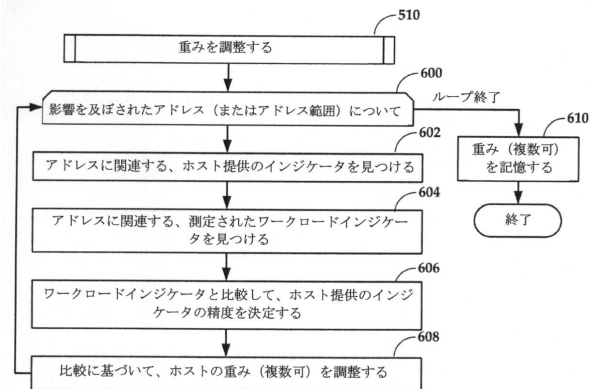
【図 4】



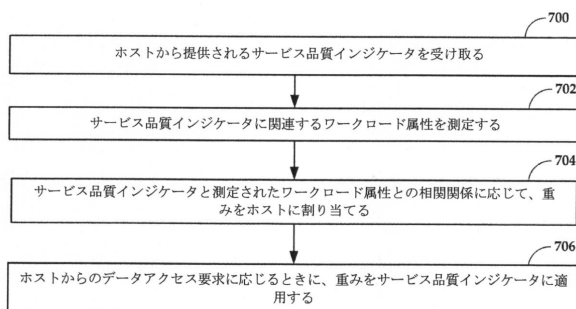
【図 5】



【図 6】



【図 7】



フロントページの続き

- (72)発明者 リャン・ジェイムズ・ゴス
アメリカ合衆国、55372 ミネソタ州、プライヤー・レイク、ポンドセッジ・レーン、508
8
- (72)発明者 マイケル・ジョセフ・スタイナー
アメリカ合衆国、55104 ミネソタ州、セント・ポール、キャロル・アベニュー、2104
- (72)発明者 マーク・アレン・ギャートナー
アメリカ合衆国、55127 ミネソタ州、バッドネ・ハイツ、コスタ・ドライブ、3273
- (72)発明者 デビッド・スコット・エブセン
アメリカ合衆国、55305 ミネソタ州、ミネトンカ、リッジマウント・アベニュー・ウェスト、
11905

審査官 田名網 忠雄

- (56)参考文献 米国特許出願公開第2013/0031298 (US, A1)
特開2005-228278 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06 - 3/08
G06F 12/00
G06F 13/10 - 13/14