

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6302049号
(P6302049)

(45) 発行日 平成30年3月28日 (2018. 3. 28)

(24) 登録日 平成30年3月9日 (2018. 3. 9)

| | | | | | |
|-------------------|------------------|------|-------|------|--|
| (51) Int. Cl. | F I | | | | |
| G06F 12/00 | (2006.01) | G06F | 12/00 | 545A | |
| G06F 3/06 | (2006.01) | G06F | 12/00 | 531D | |
| G06F 13/10 | (2006.01) | G06F | 3/06 | 301Z | |
| | | G06F | 13/10 | 340A | |

請求項の数 15 (全 29 頁)

| | | | |
|---------------|-------------------------------|-----------|-----------------------------------|
| (21) 出願番号 | 特願2016-513985 (P2016-513985) | (73) 特許権者 | 506329306 |
| (86) (22) 出願日 | 平成26年5月7日 (2014. 5. 7) | | アマゾン テクノロジーズ インコーポレ イテッド |
| (65) 公表番号 | 特表2016-521426 (P2016-521426A) | | アメリカ合衆国 98108-1226 |
| (43) 公表日 | 平成28年7月21日 (2016. 7. 21) | | ワシントン州 シアトル ビーオー ボッ クス 81226 |
| (86) 国際出願番号 | PCT/US2014/037143 | (74) 代理人 | 110001243 |
| (87) 国際公開番号 | W02014/186191 | | 特許業務法人 谷・阿部特許事務所 |
| (87) 国際公開日 | 平成26年11月20日 (2014. 11. 20) | (72) 発明者 | コリン レアード ラジエール |
| 審査請求日 | 平成27年12月4日 (2015. 12. 4) | | アメリカ合衆国 98109-5210 |
| (31) 優先権主張番号 | 13/895, 237 | | ワシントン州 シアトル テリー アベニ ュー ノース 410 |
| (32) 優先日 | 平成25年5月15日 (2013. 5. 15) | 審査官 | 小林 哲雄 |
| (33) 優先権主張国 | 米国 (US) | | |
| (31) 優先権主張番号 | 13/895, 205 | | |
| (32) 優先日 | 平成25年5月15日 (2013. 5. 15) | | |
| (33) 優先権主張国 | 米国 (US) | | |

最終頁に続く

(54) 【発明の名称】 ハードウェア故障に基づくデータの割り当て

(57) 【特許請求の範囲】

【請求項1】

データを記憶するためのコンピュータ実装方法であって、
実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で、
データ記憶システムに記憶されるデータオブジェクトを受信することであって、前記
データ記憶システムが、複数の物理的データ記憶位置を含み、前記複数の物理的データ記
憶位置の各位置が、複数の物理的特性を有する、前記受信することと、

複数のデータ断片であって、前記データ断片の真部分集合が前記データオブジェクト
を構築するのに使用可能であるように構成された、前記複数のデータ断片を生成するよう
に、冗長符号化スキームを適用することと、

前記複数の物理的データ記憶位置の少なくとも部分集合の物理的特性に適用されたデ
ータ配置規則の集合に少なくとも部分的に基づいて、前記生成されたデータ断片の配置を
決定することであって、前記配置が、前記生成されたデータ断片のうちのそれぞれのデー
タ断片について、前記データ配置規則の集合を満たす前記複数の物理的データ記憶位置か
らの複数の位置を示す、前記決定することと、

前記決定された配置に従って前記データ断片を記憶させることと、を含む、前記コン
ピュータ実装方法。

【請求項2】

前記物理的特性は、多重記憶デバイスの構成内の位置、データ記憶ラック内の多重記憶
デバイスユニットの位置、データ記憶ラックの位置、データセンタ室、データセンタ、及

びデータセンタ地理的位置を含む群から選択される1つ以上の物理的特性を含む、請求項1に記載の前記コンピュータ実装方法。

【請求項3】

前記物理的特性は、プラッタ、オンプラッタ配置、プラッタの面、及び読み取り/書き込みヘッドを含む群から選択される1つ以上の物理的特性を含む、請求項1または2に記載の前記コンピュータ実装方法。

【請求項4】

前記データ配置規則の集合は、前記データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、請求項1～3のいずれか一項に記載の前記コンピュータ実装方法。

10

【請求項5】

前記生成されたデータ断片の前記配置は、前記データオブジェクトを受信する前に前記データ配置規則を満たすと前記データ記憶システムによって予め決定された複数の配置から、前記データ記憶システムが前記配置を選択することによって決定される、請求項1～4のいずれか一項に記載の前記コンピュータ実装方法。

【請求項6】

前記データ記憶システムは、コンピューティングリソースサービスプロバイダによって複数の顧客に提供されたデータ記憶サービスを支援して動作し、

前記データオブジェクトは、前記データ記憶システムに記憶するための顧客のデータを含む、請求項1～5のいずれか一項に記載の前記コンピュータ実装方法。

20

【請求項7】

データを記憶するためのコンピュータ実装方法であって、

実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で、

データ記憶システムに記憶されるデータを受信することであって、前記データ記憶システムが、複数の物理的データ記憶位置を含み、前記複数の物理的データ記憶位置の各位置が、複数の物理的特性を有する、前記受信することと、

前記複数の物理的データ記憶位置の少なくとも部分集合の物理的特性に適用されたデータ配置規則の集合に少なくとも部分的に基づいて、前記データの断片の配置を決定することであって、前記配置が、前記データの断片のうちのそれぞれのデータの断片について、前記データ配置規則の集合を満たす前記複数の物理的データ記憶位置からの複数の位置を示す、前記決定することと、

30

前記決定された配置に従って前記断片を記憶させることと、を含む、前記コンピュータ実装方法。

【請求項8】

冗長符号化スキームを用いて前記断片を生成することをさらに含む、請求項7に記載の前記コンピュータ実装方法。

【請求項9】

前記物理的特性は、多重記憶デバイスの構成を含む多重記憶デバイスユニット内の位置、及び1つ以上の多重記憶デバイスユニットを含むデータ記憶ラック内の位置を含む、請求項7または8に記載の前記コンピュータ実装方法。

40

【請求項10】

前記物理的特性は、記憶デバイスの複数のプラッタからのプラッタ上の位置を示す1つ以上の物理的特性を含む、請求項7～9のいずれか一項に記載の前記コンピュータ実装方法。

【請求項11】

前記データ配置規則の集合は、前記データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、請求項7～10のいずれか一項に記載の前記コンピュータ実装方法。

【請求項12】

前記断片の全体的サイズは、前記受信されたデータのサイズを超え、前記断片の各々は

50

、前記受信されたデータの前記サイズ未満の個別サイズを有する、請求項 7 ~ 11 のいずれか一項に記載の前記コンピュータ実装方法。

【請求項 13】

システムであって、

1つ以上のプロセッサと、

1つ以上のメモリであって、前記システムの前記1つ以上のプロセッサによって実行されるとき、前記システムに、以下のこと、

データ配置規則の集合を、データの断片に、対応する物理的特性を有する複数の物理的データ記憶位置からの複数の位置の中で前記データの断片の配置を決定するように、適用することであって、前記配置が、前記データの断片のうちのそれぞれのデータの断片について、対応する物理的特性を有する物理的データ記憶位置が前記データ配置規則を満たすと決定されたことを示す、前記適用することと、

前記データ配置規則の集合の前記適用に従って、前記データの断片をデータ記憶システムに持続的に記憶させることと、を行わせる命令をその上に記憶させた、前記1つ以上のメモリと、を備える、前記システム。

【請求項 14】

前記物理的特性は、多重記憶デバイスの構成内の位置、データ記憶ラック内の多重記憶デバイスユニットの位置、データ記憶ラック位置、データセンサ室、データセンタ、及びデータセンタの地理的位置を含む群から選択される1つ以上の物理的特性を含む、請求項 13 に記載の前記システム。

【請求項 15】

前記物理的特性は、プラットフォーム、オンプラットフォーム配置、プラットフォームの面、及び読み取り/書き込みヘッドを含む群から選択される1つ以上の物理的特性を含む、請求項 13 または 14 に記載の前記システム。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願の相互参照)

本出願は、2013年5月15日とともに出願された、同時係属中の米国特許出願第 13 / 895 , 237 号及び同第 13 / 895 , 205 号に対する優先権を主張するものであり、これらの内容は、それらの全体が参照により本明細書に組み込まれる。本出願は、「REDUCING INTERFERENCE THROUGH CONTROLLED DATA ACCESS」と題され、2013年5月15日に出願された、米国特許出願第 13 / 895 , 205 号に関連した主題であり、この内容は、その全体が参照により本明細書に組み込まれる。

【背景技術】

【0002】

データ記憶システムは、それらのシステムを使用する組織の要求に対応するために進化してきており、また進化し続けている。例えば、多くの組織は、数ある理由の中でも、大容量のデータのデータアーカイビング、冗長性、及び記憶の目的のために、データ記憶システムを利用する。それらの多くの利点にかかわらず、最新のデータ記憶システムは、データ記憶システムプロバイダ及びそれらのシステムを利用する組織が負担する固有のリスクを抱えている。例えば、それを避ける最良の取り組みにかかわらず、データ記憶システムは、回転磁気媒体等の様々な可動部によって引き起こされるシステム内の振動から生じる過熱及び/または疲労損傷の影響を受けやすい可能性がある構成要素を含むことが多い。その結果として、データ記憶システム内に含まれるデータは、破損または損失の影響を受けやすい場合がある。データの適切な冗長記憶などによるこれらのリスクに適切に対処することは、データ記憶システムに依存する組織、及びサービスプロバイダにさらなるコストを提示する。

【図面の簡単な説明】

10

20

30

40

50

【 0 0 0 3 】

本開示に従う様々な実施形態は、図面を参照することにより記載される。

【 0 0 0 4 】

【図 1】様々な実施形態が実践され得る環境の例示的な実施例を示す。

【図 2】様々な実施形態が実践され得る環境の例示的な実施例を示す。

【図 3】少なくとも一実施形態に従ってデータを記憶するためのプロセスの例示的な実施例を示す。

【図 4】「単純ディスク束」ユニット (J B O D) などの複数の多重記憶デバイスユニットを含むデータ記憶システムラックの例示的な実施例を示す。

【図 5】データ記憶ラック内に含まれ得る J B O D などの多重記憶デバイスユニットの例示的な実施例を示す。

10

【図 6】データ記憶システムをホスティングし得る設備の例示的な実施例を示す。

【図 7】少なくとも一実施形態に従ってデータオブジェクトを記憶するためのプロセスの例示的な実施例を示す。

【図 8】少なくとも一実施形態に従ってデータ記憶システム内のハードウェア故障を追跡するためのプロセスの例示的な実施例を示す。

【図 9】少なくとも一実施形態に従ってデータ記憶にデータオブジェクト断片を配置するためのプロセスの例示的な実施例を示す。

【図 10】少なくとも一実施形態に従ってデータ記憶からデータオブジェクトを検索するためのプロセスの例示的な実施例を示す。

20

【図 11】少なくとも一実施形態に従ってデータ記憶にデータオブジェクト断片を配置するためのプロセスの例示的な実施例を示す。

【図 12】様々な実施形態が実践され得る環境の例示的な実施例を示す。

【 発明を実施するための形態 】

【 0 0 0 5 】

以下の説明では、様々な実施形態が記載される。説明のために、特定の構成及び詳細が実施形態の十分な理解を提供するために記載される。しかしながら、実施形態が特定の詳細を用いることなく実践され得ることも当業者に明らかであろう。さらに、周知の特徴は、記載される実施形態を分かりにくくしないために省略または簡略化され得る。

【 0 0 0 6 】

30

本明細書に記載及び提案された技術は、データ損失または破損のリスクを最小限にしなから、データを記憶し、それにアクセスすることに関する。一実施形態では、エンティティ (例えば、組織) は、サービスへの適切に構成された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) の呼び出しによるなどのサービスに連絡を取り、データオブジェクトのアーカイブまたはそのデータオブジェクトへのアクセスを要求する (例えば、データオブジェクトの検索)。一実施形態では、エンティティは、データ記憶サービスを運営するコンピューティングリソースサービスプロバイダの顧客である。要求を受信すると、サービスプロバイダは、顧客からデータオブジェクトを受信し、冗長符号化アルゴリズムを用いて、システムに記憶するためにオブジェクトを様々な断片に変換することができる。アルゴリズムは、例えば、データオブジェクトが、ごく少数の断片がオブジェクトを再コンパイルするのに必要とされるように、冗長断片に分割されるものであり得る。例となるアルゴリズムとしては、ビットの多重配列の真部分集合がビットの配列を再構築するのに好適であるように、消去コードを利用して、ビットの配列をビットの配列より集団的に大きいビットの多重配列に変換するものが挙げられる。

40

【 0 0 0 7 】

様々な実施形態では、データ記憶システムは、様々なネットワークラックを用いて促進される。これらのラックは、 J B O D 等の様々な多重記憶デバイスユニットを含み、このユニットは同様に、回転磁気媒体 (例えば、プラッタ)、磁気テープ記憶デバイス、またはソリッドステートドライブ (S S D) を利用する記憶デバイスなどの様々な記憶デバイスからなり得る。これらの記憶デバイスは、動作環境による故障の影響を受けやすい場合

50

がある。例えば、回転磁気媒体を利用する記憶デバイスの場合、各記憶デバイスは、極めて近接する他の記憶デバイスに影響を与え得る振動を生じさせる。この種の振動は、ヘッドクラッシュ（例えば、記憶デバイスのアーム上の読み取り/書き込みヘッドが回転プラッタと接触し、1つ以上のプラッタに物理的に損傷を与える）及びデータ損失につながり得る。記憶デバイスはまた、記憶デバイス自体の耐用年数を減少させ得る過熱を起こしやすい。サービスプロバイダは、データ記憶システム内の故障モードに関する情報を有することができ、この情報を使用して、損傷及びデータ損失のリスクを軽減することを試みることができる。これらの故障モードは、必ずしも故障を引き起こすことが既知であるモードではないが、それらが故障を引き起こすものである場合、データ耐久性を損なうであろうモードであり得ることに留意することが重要である。例えば、コンピュータリソースサービスプロバイダは、データが異なる記憶デバイス内の同じプラッタに書き込まれている場合、重大なデータ損失のリスクがあるかを決定する情報を有しないことがあるが、サービスプロバイダは、故障から生じる潜在的損傷のためこの状況を避けることを望むかもしれない。

10

【0008】

様々な実施形態では、サービスプロバイダは、ハードウェア故障によるデータ損失のリスクを減少させる様式でデータ記憶システム内に断片を配置することを決定する。配置は、データ記憶システム内の物理的位置を示し得る。物理的位置は、サービスプロバイダが制御することができるくらい固有であり得る。例えば、配置は、複数の断片の各断片に対して、地理的領域、データセンタ、データセンタ室、ラック、ラックユニット（例えば、ラック内の多重記憶デバイスユニット）、及び/または断片が記憶される記憶デバイスなどの断片の位置の1つ以上の物理的特性を示すか、またはさもなければそれに対応することができる。いくつかの実施形態では、サービスプロバイダは、データ配置に対するさらにより高い粒度の制御を有し、したがって、配置は、複数の断片の各断片に対して、回転磁気媒体（プラッタ）を有する記憶デバイス内のプラッタ、プラッタの側面、プラッタ上の環状領域、記憶デバイスの特定のヘッドによってアクセス可能なプラッタ上の位置、及び/または他の物理的位置などの位置の物理的特性を示すか、またはさもなければそれに対応することができる。サービスプロバイダは、そのような損失を軽減する配置決定をするように構成される配置エンジン内に記述された特定の規則を有し得る。例えば、サービスプロバイダは、データ記憶システム内のハードウェア故障を使用して、故障モード及び他の故障モードを繰り返すことを避けるために配置エンジンに組み込まれた規則を更新する。

20

30

【0009】

いくつかの実施例では、配置エンジンは、データ記憶システムにデータオブジェクトの断片を配置する要求を受信する。配置エンジンが第1の断片の位置を決定した後、エンジンは、データ損失のリスクを軽減するためにサービスプロバイダによって記述された様々な規則に基づいて次の断片のデータ記憶位置を選択することができる。例えば、選択された後続の断片の位置がエンジン内に記述された規則を満たさない場合、これは、適切な位置が識別されるまでその断片の位置を再選択し続ける。その時点で、配置エンジンは、次の断片に進むか、または分析する断片がもう残されていない場合、すべての断片をデータ記憶システムに配置する。

40

【0010】

様々な状況では、データ記憶システムは、多くの顧客からの多くの要求を同時に処理することができるか、または概して、要求処理中に他の理由（例えば、ごみ収集）のためにデータ記憶デバイスにアクセスすることができる。結果として、要求に少なくとも部分的に応答するデータを記憶するデータ記憶システム内の記憶デバイスは、要求が顧客によって行われて、データオブジェクトにアクセスするか、またはそれをアーカイブに入れるかのいずれかであるときに動作している場合がある。上述のように、記憶デバイスは、振動を生じさせることができ、過熱の影響を受けやすい場合がある。したがって、様々な実施形態では、サービスプロバイダは、例えば、配置エンジン内の規則を使用することによ

50

てアーカイブまたは検索プロセス中に特定の記憶デバイスへのアクセスを防ぐことによってデータ損失のリスクを軽減する。例えば、断片が特定の記憶デバイスに記憶されるが、隣接した記憶デバイスが使用中であることを配置エンジンが確認した場合、エンジンは、断片をデータ記憶に配置する前に任意の他の断片の位置を決定することができる。データ検索の例では、データ記憶システムは、上述と同様の動作を実施してもよいが、その代わりにデータオブジェクトを再コンパイルする必要がある必要な断片のみを取得することを選択してもよい。いくつかの実施形態では、読み取り、書き込み、及び/または削除動作は、データ記憶システムの状態が、その読み取り、書き込み、及び/または削除動作が故障のリスクを低下させるために実施されることが許可されるような状態にあるまで遅延され得る。例えば、データが特定の記憶デバイスに書き込まれ、それから読み取られ、またはそれから削除される場合、データを書き込み、読み取り、または削除することは、任意の隣接した記憶デバイスがアイドル状態にあるまで遅延され得る。

10

【 0 0 1 1 】

このように、データ記憶システム内のデータは、システム内のハードウェア故障による損失の影響をあまり受けない。加えて、本明細書に記載及び提案された技術は、さらなる技術的優位性を促進する。例えば、いくつかの実施形態では、ある特定の記憶デバイスへのアクセスが任意の所定の時間での動作環境に基づいて制限され得るため、データ記憶システム内に含まれるハードウェアは、疲労損傷になりにくく、それ故に、サービスプロバイダの維持コストを削減することができる。同様に、これは、顧客のためにコストの減少につながり得る。追加の使用はまた、本明細書に記載された様々な技術によって可能にされる。

20

【 0 0 1 2 】

図 1 は、様々な実施形態が実践され得る環境 1 0 0 の例示的な実施例を示す。環境 1 0 0 では、顧客 1 0 2 は、データ記憶システムに記憶するために、データオブジェクト 1 0 4 (例えば、これは、これらに限定されないが、実行可能なプログラム、図面、及びテキスト文書などの 1 ビット以上からなるファイルを含み得る) をコンピューティングリソースサービスプロバイダに送信する。顧客及びコンピューティングリソースサービスプロバイダは、企業または他の組織などの組織エンティティであり得る。本開示は、コンピューティングリソースサービスプロバイダと通信する顧客を論じているが、特に文脈から明白ではない限り、このような通信は、後述されるものなどの 1 つ以上の通信プロトコルを用いて組織によって、及び/またはその代わりに操作されたデバイス(例えば、コンピュータシステム)によって生じ得ることに留意されるべきである。

30

【 0 0 1 3 】

データオブジェクト 1 0 4 は、冗長符号化エンジン 1 0 6 によって処理される。冗長符号化エンジン 1 0 6 は、1 つ以上のアルゴリズムを利用して、顧客 1 0 2 によって提供されたデータオブジェクトを多くのデータ断片 1 0 8 に変換することができる。データ記憶システムは、親データオブジェクト 1 0 4 の識別子をデータ断片 1 0 8 のデータ記憶位置と関連付けるデータベースを維持することができる。加えて、データ断片 1 0 8 は、システム内で追跡するために符号化され得る。例えば、データ断片 1 0 8 は、親データオブジェクト 1 0 4 の名前、及び後でデータオブジェクトを再コンパイルすることが必要である断片の数で符号化され得る。

40

【 0 0 1 4 】

その後、図 1 に例示されるように、データ断片 1 0 8 は、データ記憶システム内の各データ断片 1 0 8 の位置を決定する配置エンジン 1 1 0 によって処理される。配置エンジン 1 1 0 は、データ断片 1 0 8 をデータ記憶システム内の記憶デバイス(例えば、ドライブ) 1 1 2 に配置するために規則の集合を適用するように構成されたコンピュータシステムまたはコンピュータシステムの構成要素であり得る。配置エンジン 1 1 0 は、サービスプロバイダによって指定された様々な規則を使用して、適切な位置を決定することができる。各データ断片 1 0 8 の位置が指定されると、データ断片 1 0 8 は、記憶するために対応する記憶デバイス 1 1 2 に転送される。

50

【 0 0 1 5 】

記憶デバイス 1 1 2 は、回転磁気媒体、磁気テープドライブ、SSD、または任意の他の形態のデータ記憶デバイスからなり得る。データ断片 1 0 8 は、顧客 1 0 2 がサービスプロバイダにデータオブジェクト 1 0 4 を修正する要求を提供するようときまで、これらの記憶デバイス 1 1 2 に存在する。例えば、顧客 1 0 2 は、データオブジェクト 1 0 4 が削除、置換、または上書きされることを要求することができる。このような場合、サービスプロバイダは、データ断片 1 0 8 を含む記憶デバイス 1 1 2 にアクセスし、顧客の要求に従うであろう。

【 0 0 1 6 】

図 2 は、本開示の様々な実施形態が実践され得る環境 2 0 0 の例示的な実施例を示す。図 2 に例示される環境 2 0 0 は、図 1 に関連して上述されるものなどの構成要素を含み得る。例えば、図 2 の環境 2 0 0 は、顧客 2 0 2 及びコンピューティングリソースサービスプロバイダ 2 0 6 を含む。顧客 2 0 2 及びコンピューティングリソースサービスプロバイダ 2 0 6 は、図 1 に関連して上述されるなどのように構成され得る。図 2 に例示されるように、顧客 2 0 2 は、インターネット 2 0 4 などの 1 つ以上の通信ネットワークを通じてコンピューティングリソースサービスプロバイダ 2 0 6 と通信することができる。顧客 2 0 2 からコンピューティングリソースサービスプロバイダ 2 0 6 への一部の通信は、コンピューティングリソースサービスプロバイダ 2 0 6 を本明細書に記載される様々な技術またはその変形に従って動作させることができる。

【 0 0 1 7 】

上述のように、コンピューティングリソースサービスプロバイダ 2 0 6 は、その顧客 2 0 2 に様々なコンピューティングリソースサービスを提供することができる。例えば、この特定の例示的な実施例では、コンピューティングリソースサービスプロバイダ 2 0 6 は、データ記憶サービス 2 0 8 に加えて、仮想コンピュータシステムサービス、データベースサービス、及び/または 1 つ以上の他の種類のデータ記憶サービスなどの 1 つ以上の他のサービス 2 1 0 を顧客 2 0 2 に提供する。これらの追加のサービス 2 1 0 は、本明細書に明示的に記載されるサービスに加えて、またはその代替として提供されてもよい。

【 0 0 1 8 】

様々な実施形態では、データ記憶サービス 2 0 8 は、顧客のデータを記憶するように集団的に動作するコンピューティングリソースの収集を含む。データ記憶サービス 2 0 8 によって記憶されたデータは、データオブジェクトに整理され得る。データオブジェクトは、恐らく、サイズに対するある特定の制約を除いて任意のサイズを有し得る。したがって、データ記憶サービス 2 0 8 は、様々なサイズの多くのデータオブジェクトを記憶することができる。データ記憶サービス 2 0 8 は、データオブジェクトをデータオブジェクトの識別子と関連付けるキー値ストアとして動作することができる。データ記憶サービス 2 0 8 によって記憶されたデータオブジェクトに関連して検索し、または他の動作を実施するために顧客 2 0 2 によって使用され得るデータオブジェクトの識別子。データ記憶サービス 2 0 8 へのアクセスは、2 0 8 の 1 つ以上のウェブサーバへのアクセスへのウェブサービス呼び出し等の適切に構成された API 呼び出しを通じてであってよい。

【 0 0 1 9 】

図 3 は、本開示が実践され得る様々な実施形態における環境 3 0 0 の例示的な実施例を示す。環境 3 0 0 は、上述の環境 2 0 0 の一部分であり得る。環境 3 0 0 では、顧客 3 0 2 から受信されたデータオブジェクトは、冗長符号化スキームを利用して、データオブジェクトをより小さいデータ断片に分割し得る冗長符号化エンジン 3 0 4 を用いて処理され得る。冗長符号化エンジン 3 0 4 が使用し得るプロセスの一実施例が消去符号化である。消去符号化によって、様々な情報の断片からなるデータオブジェクトは、より小さいデータ断片に縮小される。各データ断片は、すべてのデータ断片内のデータの組み合わせが親データオブジェクトのものより大きくあり得るように、データのその比例配分より多くを含み得る。したがって、すべてのデータ断片が親データオブジェクトを再コンパイルする必要はない。消去符号化が例示の目的のために本開示全体にわたって広範に使用されるが

10

20

30

40

50

、本開示の範囲は、必ずしも本明細書に明示的に示されるプロセスに限定されない。

【0020】

データ記憶サービス300の様々な実施形態では、配置エンジン306を含む。ここで例示され、かつ図1のように、配置エンジン306は、データ記憶システム内の様々な位置にデータ断片を配置するために規則の集合を適用するように構成されるコンピュータシステムまたはコンピュータシステムの構成要素であり得る。一般に、データ記憶システム内のデータ記憶位置は、対応する物理的特性を有し得る。例えば、位置は、特定の記憶デバイス308上、特定の多重記憶デバイスユニット(例えば、JBOD)内、特定のデータ記憶システムラック内、特定のデータセンタ室内、特定のデータセンタ内、または特定のデータセンタの地理的位置内として画定され得る。加えて、位置は、記憶デバイス308内に画定され得る。例えば、位置は、記憶デバイス308もしくは多重記憶デバイス内の特定のプラッタ上、プラッタの特定の側面上、プラッタの特定の領域(例えば、内部環状領域もしくは外部環状領域)内、ある特定のヘッドによってアクセス可能な記憶デバイス308内の特定の領域内、または混成記憶デバイスの場合には特定のキャッシュ内として画定され得る。

10

【0021】

規則は、1つ以上の物理的特性の間で断片の異種混在性の条件を強制するように構成され得る。例えば、規則は、同じ記憶デバイスプラッタ上、同じ多重記憶デバイスユニット上、同じデータ記憶ラック上、または同じデータセンタ内に2つ以上のデータ断片の配置を禁止してもよい。したがって、例えば、配置エンジン306の方法に含まれ得る規則は、2つ以上のデータ断片が単一記憶デバイス308内に配置されることを防ぎ得るものである。

20

【0022】

規則はまた、本質的により複雑であってもよい。例えば、規則は、 n 個のデータ断片のうち k 個のみが損失または破損の同じ確率を共有することを可能にすることを含み得、そこで k 及び n は、正の整数である。したがって、 $n - k$ (n マイナス k)個のデータ断片は、損失または破損のより低い確率を有する必要がある、場合により、親データオブジェクトの復元可能性を確実にすることができる。より複雑な規則の別の実施例は、 n 個の断片のうち k 個を、同じプラッタであるが、異なる記憶デバイスに記憶されることから禁止し得るものである。例えば、規則は、2つの断片のみが任意の記憶デバイスの2番目のプラッタに記憶され得ることを宣言してもよい。したがって、記憶デバイスAの2番目のプラッタに1つの断片を、記憶デバイスBの2番目のプラッタに1つの断片を、かつ記憶デバイスCの2番目のプラッタに1つの断片を記憶することは、規則に違反するであろう。したがって、配置エンジン306は、データ断片に対応する記憶デバイス308に配置することができる。加えて、配置エンジン306は、追加の情報をデータベース310に送信することができる。データベース310は、例えば、記憶デバイス308内の親データオブジェクト及び各データ断片の位置に関する情報を含み得る。顧客302がデータ記憶サービス300によってデータオブジェクトへのアクセスを要求する場合、データベース310は、このようなアクセスを取得するのに必要な関連情報のディレクトリとして機能を果たしてもよい。

30

40

【0023】

図4は、データ記憶ラック402が様々な実施形態に従って様々なデータオブジェクトを含むために使用される環境400の例示的な実施例を示す。図4に例示される環境400は、上述されるデータ記憶システムの一部として機能を果たし得る。データ記憶ラック402は、多重記憶デバイスユニット404などの1つ以上の多重記憶デバイスユニットからなり、このユニットは、各々が1つ以上の記憶デバイス406を含む。上述のように、記憶デバイス406は、動作環境による故障の影響を受けやすい場合がある。図4に例示される環境400及びその変形は、例えば、記憶デバイス406に損傷を与え得る過熱及び過度の振動を受ける場合がある。

【0024】

50

図4に例示される環境400では、ファン408は、気流、それ故に、データ記憶ラック402を冷却する方法を提供するために使用される。したがって、ファン408は、多重記憶デバイスユニット404及びその中に含まれる記憶デバイス406内の過熱を防ぐのに役立つ。しかしながら、ファン408がデータ記憶ラック402の下に位置し得るため、多重記憶デバイスユニット404内の温度は、ファン408から離れた距離との関連で増加し得る。したがって、この実施例では、ファン408から最も遠くに離れたデータ記憶ラック402内の記憶デバイス406は、損傷の影響をより受けやすい場合がある。

【0025】

環境400はまた、データ記憶ラック402の動作に必要な他の構成要素を含み得る。例えば、データ記憶ラック402はまた、コンピュータリソースサービスプロバイダの従業員がデータ記憶システムにアクセスする必要がある場合に、電源、データ記憶ラック402の上に位置するスイッチ、及びサーバを含んでもよい。本開示の範囲は、本明細書に明示的に示されたデータ記憶ラック構成に必ずしも限定されないことに留意されるべきである。

【0026】

上述のように、データ記憶ラックは、1つ以上の多重記憶デバイスユニットからなる。図5は、図4に例示されるようなデータ記憶ラックに組み込まれ得る多重記憶デバイスユニット500の例示的な実施例を示す。例示されるように、多重記憶デバイスユニット500は、アレイ状に配列された1つ以上の記憶デバイスを含むエンクロージャ502からなり得る。この実施例では、多重記憶デバイスユニット500内に囲繞された記憶デバイスは、3列で配列され、各々が4つの記憶デバイスを含み、エンクロージャ502内に合計12個の記憶デバイスを含む。しかしながら、多重記憶デバイスユニット500は、任意の数の記憶デバイスからなり得、様々な方法で配列され得る。例えば、多重記憶デバイスユニット500は、記憶デバイスの列（水平かつ/または垂直に配向される）を含んでもよく、この列は、多重記憶デバイスユニット500の中により深くあり、各記憶デバイスに多重記憶デバイスユニット500内の3次元位置を付与する。加えて、いくつかの記憶デバイスは、多重記憶デバイスユニット内に別の方法で配向され得る（例えば、斜めまたは対角線に配置される）。

【0027】

図6は、データ記憶サービスが物理的に存在し得る環境600の例示的な実施例を示す。上述のように、データ記憶ラック602は、1つ以上の多重記憶デバイスユニットからなり得る。これらの多重記憶デバイスユニットは、様々な構成で配列された1つ以上の記憶デバイスを含み得る。環境600では、データ記憶ラック602は、1つ以上のデータセンタ604を含む建物内に維持され得る。これらのデータセンタ604は、コンピューティングリソースサービスプロバイダの従業員のうちの1人以上によって維持され得る。したがって、データセンタは、様々な従業員の任務を支援するために追加の事務所606を有してもよい。

【0028】

ハードウェア故障の場合、その施設の従業員は、故障の根本原因を調査することができ、この情報をカタログに登録することができる。これにより、従業員は上述されるように、1つ以上の故障モードを防ぐために配置エンジン内で実装された規則を更新することができる。

【0029】

図7は、様々な実施形態に従ってデータオブジェクトを記憶するためのプロセス700の例示的な実施例を示す。図7に例示されるプロセス700及びその変形は、冗長符号化エンジンを利用して、データオブジェクトを多くのデータ断片に変換し得るシステムと、データ断片をデータ記憶システムに配置するための規則の集合を含み得る配置エンジンと、データ断片が図3に例示されるように記憶され得る一連の記憶デバイスとを含む、任意の好適なシステムによって実施され得る。一実施形態では、プロセス700は、データ記憶システムに記憶するために顧客からデータオブジェクトを受信する(702)ことを含

10

20

30

40

50

む。データオブジェクトは、任意の好適な様式で受信され得る。例えば、上述のように、顧客は、インターネット等の1つ以上の通信ネットワークを通じてデータ記憶サービスにアクセスすることができる。したがって、データオブジェクトを記憶する要求は、適切に構成されたウェブサービス要求または他のAPI呼び出しとして受信され得る。

【0030】

顧客からデータオブジェクトを受信する(702)と、プロセス700は、記憶するために親データオブジェクトを多くのデータ断片に変換するために冗長符号化スキームを適用する(704)ことを含み得る。符号化スキームは、例えば、すべてのデータ断片が将来、親データオブジェクトを再コンパイルするのに必要ではないように、上述のように消去符号化することを含み得る。さらに、消去符号化が例示の目的のために本開示全体にわたって広範に使用されるが、本開示の範囲は、必ずしも本明細書に明示的に示されるプロセスに限定されない。

10

【0031】

親データオブジェクトがより小さいデータ断片に縮小されると、プロセス700は、様々なデータ断片の配置を決定する(706)ことを含み得る。上述のように、様々な実施形態では、データ断片をデータ記憶システムに配置することは、配置エンジンを用いて行われ得る。したがって、コンピューティングリソースサービスプロバイダは、例えば、データ断片がデータベース内のカタログに登録され、データ記憶システム内のデータ記憶位置が割り当てられるように、上述されるような異種混在性を強制する規則などの様々な規則を配置エンジンに組み込むことができる。加えて、以下に説明されるように、配置エンジンは、データ断片の配置がデータ損失または破損のリスクを最小限にするように、データ記憶システム内の故障の検出に基づいて更新され得る。

20

【0032】

配置エンジンが各データ断片の位置を決定すると、プロセス700は、その断片をデータ記憶システム内のデータ記憶位置に配置する(708)ことができる。データ断片は、記憶するために記憶デバイスに書き込まれ得る。上述のように、記憶デバイスは、回転磁気媒体、磁気テープ記憶デバイス、SSD、または任意の他の形態のデータ記憶デバイスからなり得る。各種の記憶デバイスが様々な故障モードの影響を受けやすい場合があるが、断片の配置(708)は、故障のリスクを軽減するように行われ得る。

【0033】

加えて、様々な追加の動作は、図7に例示されるプロセス700に関連して実施され得る。以下に示されるように、例えば、データ記憶システムは、データ断片が記憶される現在の環境を取得し、そうすることが許可されるまで対応する記憶デバイスへのアクセスを遅延させることができる。加えて、データ記憶システムは、顧客から取得されたデータオブジェクト、または図3に例示されるような冗長符号化エンジンによって生成されたデータ断片を暗号化することができる。したがって、様々な実施形態では、プロセス700は、そのような動作を実施することを含み得る。

30

【0034】

上述のように、配置エンジンは、データ記憶システム内のデータ断片の適切な配置を決定するために使用され得る。データ断片の適切な配置を管理する規則は、例えば、コンピューティングリソースサービスプロバイダの従業員によって、図3に例示されたプロセスの繰り返される反復を通して、またはハードウェア故障等の外部条件によって決定され得る。いくつかの実施形態では、コンピューティングリソースサービスプロバイダの顧客には、コンピューティングリソースサービスプロバイダによって配置エンジンの規則を生成するために使用される配置条件を提供するウェブサービスまたは他のインターフェースによるなどの能力が提供される。データ配置の条件は、上述の条件を含み得る。この条件は、明示的(例えば、自分で規則を指定することによって)または黙示的に、各レベルが配置規則の集合に対応する、いくつかの冗長レベルのうちの1つを選択することによるなど、顧客によって指定され得る。顧客に対する請求金額は、顧客が指定するデータ配置規則がもしあればそれに少なくとも部分的に基づいて異なり得る。

40

50

【 0 0 3 5 】

図 8 は、少なくとも一実施形態に従ってデータ記憶システム内のハードウェア故障を追跡するためのプロセス 8 0 0 の例示的な実施例を示す。図 8 に例示されるプロセス 8 0 0 及びその変形は、配置エンジンを利用して、データ断片がどこに記憶されるべきであるかを決定し得るものを含む、任意の好適なシステムによって実施され得る。プロセス 8 0 0 では、データ記憶システムは、システム内の故障を検出する (8 0 2) ことができる。例えば、多重記憶デバイスユニット内に記憶デバイス故障がある場合、信号は、データ記憶システム、及びコンピュータリソースサービスプロバイダの従業員に送信され、故障を知らせることができる。故障は、多重記憶デバイスユニット内の記憶デバイスに組み込まれたセルフモニタリングアナリシスアンドリポーティングテクノロジー (Self - Mon i t o r i n g A n a l y s i s a n d R e p o r t i n g T e c h n o l o g y) (S M A R T) に基づいて検出され得る。加えて、信号は、電源故障、データ記憶ラック内の通気障害、またはデータセンタ施設内の壊滅的な停電などのデータ記憶システム内のより深刻な障害がある場合に送信され得る。

10

【 0 0 3 6 】

検出された故障の深刻度及び影響を決定する際に、データ記憶システムまたはコンピュータリソースサービスプロバイダの従業員は、特定の故障モード、例えば、データ記憶システム内のすべての以前の障害の履歴の記録を含むログで記録することができる。したがって、このログは、データ記憶システムの故障統計を更新する (8 0 4) ために使用され得る。故障統計を更新する (8 0 4) ための目的の 1 つは、例えば、データ記憶システム内の将来のハードウェア障害の可能性を取得することである。これは、1 つ以上の機能を果たすことができる。例えば、サーバのファンが特定の記憶デバイスの耐用年数を減少させ得るか、またはこれらの記憶デバイス内に間欠故障を引き起こし得る振動を発生させる場合、この情報は、コンピュータリソースサービスプロバイダがこれらの特定の記憶デバイスの使用を制限することを求め得るように、このプロバイダに対して価値のあるものであり得る。加えて、データサービスラック内の電源が配線の束または内部記憶デバイスなどのある特定の多重記憶デバイスユニット構成要素の耐性を超え得る熱を生じており、異なる故障モードをもたらす場合、コンピュータリソースサービスプロバイダは、その特定の多重記憶デバイスユニットへのアクセスを制限すること、またはデータサービスラックに記憶されるデータの量を制限することを求めることができる。

20

30

【 0 0 3 7 】

故障統計を更新する (8 0 4) ことによって取得された十分な情報を用いて、コンピュータリソースサービスプロバイダの従業員またはデータ記憶システムは、1 つ以上のアルゴリズムによって、配置エンジンを更新する (8 0 6) ことができる。配置エンジンの更新 (8 0 6) は、図 1 及び 3 に例示されるように、データ記憶システム内の 1 つ以上のデータ断片の配置を決定するために使用され得る新しい規則の修正または実装を含み得る。例えば、同じ親データオブジェクトを有する多重データ断片が、例えば、同じ記憶デバイスプラットフォームまたは記憶デバイスに記憶されることを防ぐ新しい規則が実装されてもよい。加えて、この規則は、例えば、2 つ以上のデータ断片が同じデータ記憶ラック内、データセンタの同じ室内、同じデータセンタ内、または同じ地理的指定内に記憶されないような系を含み得る。

40

【 0 0 3 8 】

上述のように実装され得る規則は、静的でなくてもよい。例えば、データ記憶システムがさらなるハードウェア故障を検出する (8 0 2) 場合、プロセス 8 0 0 は繰り返してもよく、故障統計のさらなる更新 (8 0 4) をもたらし得る。したがって、配置エンジンは、データ記憶システム内のハードウェア故障の頻度に基づいていかなる時でも更新され得る (8 0 6) 。

【 0 0 3 9 】

図 9 は、少なくとも一実施形態に従ってデータオブジェクトを記憶するためのプロセス 9 0 0 の例示的な実施例である。上述のように、データオブジェクトは、図 3 に例示され

50

るものなどの冗長符号化エンジンを用いて、多くのデータ断片に変換され得る。この時点では、プロセス900は、親データオブジェクトのデータ断片をデータ記憶システムに配置する要求を受信する(902)ことができるものである。この時点では、図3に例示されるものなどの配置エンジンは、第1のセグメントの位置を選択する(904)ことができる。第1のセグメントの位置の選択(904)は、データ記憶システム内の任意の既知のハードウェア故障モードから独立してもよく、同じ親データオブジェクトを有する他のデータ断片の配置から独立してもよい。しかしながら、規則は、配置エンジン内に実装され、第1のセグメントが特定の位置に記憶されることを必要にし得る。

【0040】

第1のデータ断片位置が選択される(904)と、プロセス900は、次のデータ断片の位置を選択する(906)ことを含み得る。第1のデータセグメントの位置を考慮すると、配置エンジンは、その中に含まれる上述のものなどの実装された規則に基づいて次のデータ断片の位置を調査する(908)ことができる。選択された位置が配置エンジン内に実装された規則を満たさない場合、データ記憶システムは、現在のデータ断片の新しいデータ記憶位置を選択する(910)ことができる。例えば、第1のデータ断片が多重記憶デバイスユニットXに記憶されると配置エンジンが決定し、配置エンジン状態内に実装された規則が、2つ以上のデータ断片が同じ多重記憶デバイスユニットに記憶されないと述べる場合、配置エンジンは、任意の後続のデータ断片に割り当てられた位置を調査することができ、それが多重記憶デバイスユニットXに記憶されるように現在割り当てられる場合にその位置を再選択することができる。データ断片の位置が配置エンジンに記述された規則を満たす場合、データ記憶システムは、これ以上データ断片が残っているかを評価する(912)ことができる。位置を必要とするより多くのデータ断片が存在する場合、プロセス900は、該データ断片の位置を決定する(906)ことを含み得る。このように、各データ断片の位置は、配置エンジンに含まれる様々な規則を満たすことができる。

【0041】

各データ断片に配置エンジンに記述されたすべての規則を満たす位置が割り当てられると、プロセス900は、すべてのデータ断片をデータ記憶システムに書き込む(914)ことができる。規則の実装が例示の目的のために本開示全体にわたって広範に使用されるが、本開示の範囲は、本明細書に明示的に示されるプロセスに必ずしも限定されない。例えば、配置エンジンは、後続のデータ断片の可能な位置を制限する規則に加えて、プロセス900の前の繰り返しに基づいて現在のデータ断片がデータシステムに記憶される位置の前の集合を選択するために実施され得る一連のアルゴリズムを含んでもよい。前の集合のリストは、データ損失または破損のリスクの減少をもたらしたこれまでに使用された位置を含み得る。

【0042】

上述のように、記憶デバイスは、例えば、振動及び過熱から生じる損傷の影響を受けやすい場合がある。回転磁気媒体を利用する記憶デバイスの場合、このような記憶デバイスの通常動作は、周囲のデータ環境に悪影響を与え得る振動をもたらす。したがって、図10は、少なくとも一実施形態に従って周囲の悪条件によるデータ損失のリスクを減少させ得る様式でデータ記憶からデータオブジェクトを検索するためのプロセス1000の例示的な実施例である。

【0043】

上記に詳述されたプロセス900と同様に、プロセス1000は、様々な記憶デバイスへのアクセスを必要とし得る動作を実施するように、例えば、顧客またはコンピューティングリソースサービスプロバイダからの要求を受信する(1002)ことを含み得る。これらの記憶デバイスは、親データオブジェクトと関連付けられたデータ断片を含み得る。要求を受信する(1002)と、データ記憶システムは、上述のようなデータベースを参照して、データ記憶システムに記憶されたデータ断片の位置を決定する(1004)ことができる。位置が決定される(1004)と、データ記憶システムは、現在のデータ断片の位置を取り囲む環境の現在の活動状態を取得する(1006)ことができる。例えば、

現在のデータ断片が特定の記憶デバイスに位置する場合、その環境は、現在のデータ断片を含む記憶デバイスに空間的に近接した（例えば、振動伝達媒体または共通の取り付け具を共有する）任意の記憶デバイス、及び目標とした記憶デバイスを有する多重記憶デバイスユニットに空間的に近接した任意の多重記憶デバイスユニットを含み得る。加えて、その環境は、目標とした記憶デバイスを取り囲む現在の温度及び含水量、ならびに気圧または空気中の汚染物質を含み得る。現在の活動状態は、データ記憶システム内の任意の構成要素の動作状態を参照することができる。例えば、記憶デバイスについて、現在の活動状態は、記憶デバイスが回転している（例えば、記憶デバイス内の記憶媒体のアクセスを読み取り/書き込み/削除する）か、かつどの速度で、またはアイドル状態にあるかを指すことができる。上述のように、動作中の回転磁気媒体を利用する記憶デバイスは、振動及び熱を生じ得る。したがって、記憶デバイスの現在の活動状態は、周囲環境に対応する影響を与え得る。

10

【 0 0 4 4 】

周囲環境の評価は、既知の故障モードに基づいて1つ以上の分析を含み得る。これらの故障モードの深刻度は、上述のプロセス800などのプロセスによって、または既知の製作公差によってコンピューティングリソースサービスプロバイダに既知であり得る。例えば、任意のハードウェア構成要素（例えば、記憶デバイス、多重記憶デバイスユニット、データ記憶ラック）の供給前に、製造者は、これらの構成要素の耐性及び/または耐用年数を決定するために試験を実施することができる。これは、周囲の悪条件（例えば、過度の熱、振動、湿気）を決定する器具が埋め込まれた記憶デバイスなどの試験デバイスの使用を含み得る。したがって、製造者は、周囲環境に基づいてあらゆるハードウェア構成要素への損傷のリスクを詳述する情報を取得することができる。この情報は、例えば、コンピューティングリソースサービスプロバイダへのハードウェア構成要素の販売中、または製造者によって最初に発見されたとき、コンピューティングリソースサービスプロバイダに伝達され得る。この場合、製造者は、企業または他の組織などの組織エンティティであり得る。

20

【 0 0 4 5 】

コンピューティングリソースサービスプロバイダは、上述のように、プロセス800を用いて取得された統計、または製造者から取得された情報を使用して、周囲環境がデータ損失または破損のリスクを増加させるならば、記憶デバイスへのアクセスを防ぐであろう条件の集合を生じさせることができる。例えば、コンピューティングリソースサービスプロバイダは、アクセスされる記憶デバイスへの空間的近接に基づいた条件の集合を提供することができる。例えば、親データオブジェクトへのアクセスに必要なデータ断片を有するものに隣接した1つ以上の記憶デバイスが現在アクティブである（例えば、記憶デバイスがその媒体に1つ以上の動作を実施している）場合、データ記憶システムは、すべての隣接した記憶デバイスがアイドル状態にあるまで、データ断片を含む記憶デバイスへのアクセスを拒否することができる。空間的近接はまた、隣接した記憶デバイスに限定されない。例えば、統計的に、目標とした記憶デバイスに隣接しないが、振動伝達媒体を目標とした記憶デバイスと、または共通の取り付け具を目標とした記憶デバイスと共有する現在アクティブな記憶デバイスから生じる振動が目標とした記憶デバイスにより大きい潜在的損傷をもたらす場合、データ記憶システムはまた、これらのアクティブな記憶デバイスがアイドル状態にあるまで、データ断片を含む記憶デバイスへのアクセスを拒否することができる。したがって、コンピューティングリソースサービスプロバイダによって実装され得る条件を用いて、プロセス1000は、現在のデータ断片を含む記憶デバイスがアクセスされ得るかの決定1008を含み得る。現在の記憶デバイスへのアクセスが周囲の悪条件により許可されない場合、データ記憶システムは、現在のデータ断片を迂回し、次のデータ断片の位置を決定する（1004）ことができる。さもなければ、データ記憶システムは、すべての必要なデータ断片が位置特定され、かつ環境がデータ断片の検索に有利であるかを決定することを求める（1010）ことができる。

30

40

【 0 0 4 6 】

50

上述の条件は、様々な方法で実装され得ることに留意することが重要である。例えば、データ記憶システムは、記憶デバイスへのアクセスが必要であるときに実行される小さいスクリプトを含み得る。このスクリプトは、必要なデータ断片を含む記憶デバイスに隣接した記憶デバイスを調査し、すべての隣接した記憶デバイスの状態を含むファイルを生成することができる。同様に、このファイルは、プロセス1000と一致するデータ記憶システムによって処理され得る。

【0047】

図3に上述されるように、冗長符号化スキームを利用する冗長符号化エンジンは、データオブジェクトをより小さいデータ断片に縮小するために使用され得る。消去符号化によって、冗長符号化エンジンを用いて生成されたすべてのデータ断片が親データオブジェクトを復元するのに必要でなくてもよい。したがって、プロセス1000は、データ記憶システムが親データオブジェクトを復元するのに必要なデータ断片を有するかを評価する(1010)ことを含むことができる。この評価は、データ断片を含む記憶デバイスが上述のようにアクセスされ得るとデータ記憶システムが決定する(1008)たびに生じ得る。必要なデータ断片を含む記憶デバイスがアクセスされ得るとデータ記憶システムが決定した場合、データ記憶システムは、必要な記憶デバイスへのアクセスを許可することができる。しかしながら、より多くのデータ断片が親データオブジェクトを再生するのに必要である場合、データ記憶システムは、次のデータ断片を選択し、データ記憶システム内の該データ断片の位置を決定する(1004)ことができる。

【0048】

親データオブジェクトを再生するのに必要なすべてのデータ断片が識別されたと決定され(1010)、必要なデータ断片を含む記憶デバイスにアクセスする許可があるとデータ記憶システムが決定する(1008)と、データ記憶システムは、データ断片にアクセスし(1012)、親データオブジェクトを再コンパイルすることができる。

【0049】

本明細書に記載されるすべてのプロセスと同様に、プロセス1000の変形は、本開示の範囲内であると見なされる。例えば、上述のように、プロセス1000は、記憶デバイスへのアクセス(1008)が存在する悪条件により禁止される場合、データ断片の位置を決定する(1004)ことを含む得る。プロセス1000の変形の一実施例は、記憶デバイスへのアクセスが次のデータ断片の位置を決定する(1004)前に許可されるまでデータ記憶システムが待つことができるものであり得る。プロセス1000の変形の別の実施例は、上述されるように、データ記憶システムがデータオブジェクトを構築するのに必要な単に最小の数の断片とは対照的に、すべてのデータ断片(またはより多数のデータ断片)を取得することを選択するものであり得る。

【0050】

記憶デバイスへのアクセスは、読み取り動作のみに限定されなくてもよい。したがって、図11は、少なくとも一実施形態に従って周囲の悪条件によりデータ損失のリスクを減少させ得る様式でデータオブジェクト断片をデータ記憶に配置するためのプロセス1100の例示的な実施例である。上述のプロセス900とかなり類似して、プロセス1100は、親データオブジェクトの1つ以上のデータ断片をデータ記憶システムに配置する要求を受信する(1102)ことを含む得る。

【0051】

上述のプロセス900と同様に、プロセス1100は、第1のデータ断片を記憶するように位置を決定する(1104)ことを含む得る。この決定1104は、図3に例示されるなどのように配置エンジンを用いて行われ得る。上述のように、第1のセグメントの位置の選択は、データ記憶システム内の任意の既知のハードウェア故障モードから独立してもよく、同じ親データオブジェクトを有する他のデータ断片の配置から独立してもよい。しかしながら、プロセス1100はその後、データ断片の所望の位置を取り囲む環境の現在の状態を取得する(1106)ことを含む得る。周囲環境の評価は、上記に記載された技術を利用する様式で行われ得る。加えて、コンピューティングリソースサービスプロバ

10

20

30

40

50

イダは、上記に記載されたプロセス800に基づいて、またはデータ記憶システム構成要素の製造者から受信された情報に基づいて、配置エンジンに組み込まれた規則を更新することができる。

【0052】

プロセス1000と同様に、プロセス1100は、データ断片の記憶のために選択された記憶デバイスへのアクセスが許可されるかどうかの決定1108を含み得る。この決定は、上記に記載されるような方法を用いて行われ得る。選択された記憶デバイスへのアクセスが許可されない場合、データ記憶システムは、次のデータ断片の位置を決定する(1104)ために進むことができる。さもなければ、データ記憶システムは、親データオブジェクトと関連付けられたすべてのデータ断片が位置特定され、かつその環境がデータ断片の記憶に有利であるかを決定する(1110)ことを求めることができる。データ記憶システムがすべてのデータ断片を記憶するように識別された記憶デバイスにアクセスする許可を有すると、データ記憶システムは、これらのデータ断片をそれぞれの記憶デバイスに配置する(1112)ことができる。

【0053】

本明細書に記載されるすべてのプロセスと同様に、プロセス1100の変形は、本開示の範囲内であると見なされる。例えば、プロセス1100は代わりに、選択された位置が環境の悪条件により許可されない場合、データ断片の新しい位置を決定することを含み得る。したがって、データ記憶システムは、1つのデータ断片に焦点を合わせ、好適な位置が選択されるまで次のデータ断片に移らないようにすることができる。別の実施例は、n個のデータ断片のうちk個が同じ損失または破損の確率を共有することを可能にするものなど、配置エンジン内のより複雑な規則に基づいてもよい。したがって、データ記憶システムは、k個のデータ断片の選択された位置を取り囲む環境を無視することを含み得る。n - k個の断片が残ると、データ記憶システムは、これらの断片の周囲環境を評価し始め、プロセス1100に従って適切な位置を選択するであろう。

【0054】

図12は、様々な実施形態に従った態様を実装するための環境1200の態様を例示する。理解され得るように、ウェブベースの環境が説明の目的のために使用されるが、異なる環境は必要に応じて、様々な実施形態を実装するために使用され得る。この環境は、電子クライアントデバイス1202を含み、これは、適切なネットワーク1204上で要求、メッセージ、または情報を送受信し、かつ情報をデバイスのユーザに送り返すように動作可能な任意の適切なデバイスを含み得る。このようなクライアントデバイスの実施例としては、パーソナルコンピュータ、携帯電話、ハンドヘルドメッセージングデバイス、ラップトップコンピュータ、タブレットコンピュータ、セットトップボックス、携帯情報端末、埋込み型コンピュータシステム、電子書籍リーダーなどが挙げられる。ネットワークとしては、イントラネット、インターネット、セルラーネットワーク、ローカルエリアネットワーク、もしくは任意の他のこのようなネットワーク、またはこれらの組み合わせを含む任意の適切なネットワークが挙げられ得る。このようなシステムに使用される構成要素は、選択されたネットワーク及び/または環境の種類に少なくとも部分的に依存し得る。このようなネットワークを介して通信するためのプロトコル及び構成要素は周知であり、本明細書で詳細に説明されない。ネットワーク上の通信は、有線または無線接続及びこれらの組み合わせによって有効にされ得る。この実施例では、ネットワークは、その環境が要求を受信し、かつそれに応答してコンテンツを供給するためのウェブサーバ1206を含むときにインターネットを含むが、他のネットワークでは、当業者に明らかであるように、同様の目的を果たす代替のデバイスが使用され得る。

【0055】

例示的な環境は、少なくとも1つのアプリケーションサーバ1208と、データストア1210とを含む。つながれるか、または別の方法で構成され得る、いくつかのアプリケーションサーバ、層、または他の要素、プロセス、もしくは構成要素があり得、これらは、適切なデータストアからデータを取得することなどのタスクを実施するようにやり取り

10

20

30

40

50

することができる。本明細書で使用されるとき、サーバは、ハードウェアデバイスまたは仮想コンテンツシステムなどの様々な方法で実装され得る。場合により、サーバは、コンピュータシステムで実行されるプログラミングモジュールを指すことができる。本明細書で使用されるとき、用語「データストア」は、データを記憶し、それにアクセスし、かつそれを検索することができる任意のデバイス、またはデバイスの組み合わせを指し、これは、あらゆる基準の分散型またはクラスタ化環境で任意の組み合わせ及び任意の数のデータサーバ、データベース、データ記憶デバイス、及びデータ記憶媒体を含み得る。アプリケーションサーバは、クライアントデバイスのための1つ以上のアプリケーションの態様を実行する必要に応じて、データストアと一体化し、アプリケーションのためのデータアクセス及びビジネスロジックのうちいくつか（さらに大部分）を処理するための任意の適切なハードウェア及びソフトウェアを含み得る。アプリケーションサーバは、データストアと協働でアクセス制御サービスを提供することができ、ユーザに転送されるテキスト、グラフィックス、音声、及び/または映像などのコンテンツを生成することができ、ハイパーテキストマークアップ言語（「HTML」）、拡張マークアップ言語（「XML」）、またはこの実施例における別の適切な構造化言語の形態でウェブサーバによってユーザに供給され得る。すべての要求及び応答の処理、ならびにクライアントデバイス1202とアプリケーションサーバ1208との間のコンテンツの配信は、ウェブサーバによって処理され得る。ウェブサーバ及びアプリケーションサーバは必須ではなく、本明細書に説明される構造化コードが本明細書のどこかで考察されるような任意の適切なデバイスまたはホストマシンで実行され得るように、単に例示的な構成要素であることが理解されるべきである。さらに、単一デバイスによって実施されると本明細書に記載される動作は、特に文脈から明らかでない限り、分散型システムを形成し得る多重デバイスによって集団的に実施され得る。

【0056】

データストア1210は、本開示の特定の態様に関するデータを記憶するためのいくつかの別個のデータ表、データベース、または他のデータ記憶機構及び媒体を含み得る。例えば、例示されるデータストアは、生産側のコンテンツを供給するために使用され得る生産データ1212及びユーザ情報1216を記憶するための機構を含み得る。データストアはまた、報告、分析、または他のこのような目的のために使用され得るログデータ1214を記憶するための機構を含むように示される。ページイメージ情報のためなどのデータストアに記憶され、かつ正確な情報にアクセスする必要があり得る多くの他の態様があり得、これらは、必要に応じて上記に列記された機構のうちいずれかに、またはデータストア1210における追加の機構において記憶され得ることが理解されるべきである。データストア1210は、それと関連付けられた論理によって、アプリケーションサーバ1208から命令を受信し、かつそれに応答してデータを取得、更新、またはさもなければ処理するように動作可能である。一実施例では、ユーザは、ユーザによって操作されたデバイスによって、ある特定の種類の項目の検索要求を提出してもよい。この場合、データストアは、ユーザの身元を検証するユーザ情報にアクセスしてもよく、カタログ詳細情報にアクセスして、その種類の項目に関する情報を取得することができる。次に、この情報は、ユーザがユーザデバイス1202上のブラウザを介して見ることができるウェブページ上に列挙する結果などでユーザに返され得る。特定の対象項目の情報は、ブラウザの専用ページまたはウィンドウで見られ得る。しかしながら、本開示の実施形態がウェブページの文脈に必ずしも限定されないが、要求が必ずしもコンテンツの要求ではない場合、概して処理要求により一般に適用可能であり得ることに留意されるべきである。

【0057】

各サーバは典型的に、そのサーバの一般的な管理及び動作のために実行可能なプログラム命令を提供するオペレーティングシステムを含み、典型的に、サーバのプロセッサによって実行されるとき、サーバがその意図された機能を実施することを可能にする命令を記憶するコンピュータ可読記憶媒体（例えば、ハードディスク、ランダムアクセスメモリ、読み取り専用メモリなど）を含む。オペレーティングシステムの好適な実装形態及びサー

10

20

30

40

50

バの一般的な機能性は既知であるか、または市販のものであり、特に本明細書の本開示に照らして当業者に容易に実装される。

【 0 0 5 8 】

一実施形態における環境は、1つ以上のコンピュータネットワークまたは直接接続を用いて、通信リンクを介して相互接続されるいくつかのコンピュータシステム及び構成要素を利用する分散型コンピューティング環境である。しかしながら、このようなシステムは、図12に例示されるより少ないか、またはより多い数の構成要素を有するシステムで等しく良好に動作し得ることが当業者に理解されるであろう。したがって、図12におけるシステム1200の描写は、本質的に例示的であり、本開示の範囲に限定しないと見なされるべきである。

10

【 0 0 5 9 】

様々な実施形態はさらに、多種多様の動作環境で実装され得、これらは、場合により、いくつかのアプリケーションのうちのいずれかを動作するために使用され得る1つ以上のユーザコンピュータ、コンピューティングデバイス、または処理デバイスを含み得る。ユーザデバイスまたはクライアントデバイスは、標準的オペレーティングシステムを実行するデスクトップ、ラップトップ、またはタブレットコンピュータなどのいくつかの汎用パーソナルコンピュータ、ならびにモバイルソフトウェアを実行し、かついくつかのネットワーク及びメッセージングプロトコルを支援することができるセルラー、無線、及びハンドヘルドデバイスのうちのいずれかを含み得る。このようなシステムはまた、様々な市販のオペレーティングシステム、ならびに開発及びデータベース管理などの目的のため
20
の他の既知のアプリケーションのうちのいずれかを実行するいくつかのワークステーションを含み得る。これらのデバイスはまた、ダミー端子、シンクライアント、ゲーミングシステム、及びネットワークを介して通信することができる他のデバイスなどの他の電子デバイスを含み得る。

20

【 0 0 6 0 】

本開示の様々な実施形態は、伝送制御プロトコル/インターネットプロトコル(「TCP/IP」)、開放型システム間相互接続(「OSI」)モデルの様々な層で動作するプロトコル、ファイル転送プロトコル(「FTP」)、ユニバーサルプラグアンドプレイ(「UpnP」)、ネットワークファイルシステム(「NFS」)、共通インターネットファイルシステム(「CIFS」)、及びAppleTalkなどの様々な市販のプロトコル
30
のうちのいずれかを介して通信を支援するために当業者によく知られている少なくとも1つのネットワークを利用する。ネットワークは、例えば、ローカルエリアネットワーク、広域ネットワーク、仮想プライベートネットワーク、インターネット、イントラネット、エクストラネット、公衆交換電話網、赤外線ネットワーク、無線ネットワーク、及びこれらの任意の組み合わせであり得る。

30

【 0 0 6 1 】

ウェブサーバを利用する実施形態では、ウェブサーバは、ハイパーテキスト転送プロトコル(「HTTP」)サーバ、FTPサーバ、共通ゲートウェイインターフェース(「CGI」)サーバ、データサーバ、Javaサーバ、及び業務アプリケーションサーバを含む様々なサーバまたは中間階層アプリケーションのうちのいずれかを実行することが
40
できる。サーバ(複数可)はまた、Java(登録商標)、C、C#、もしくはC++等の任意のプログラミング言語、またはPerl、Python、もしくはTCL等のスクリプト言語、ならびにこれらの組み合わせで書き込まれた1つ以上のスクリプトまたはプログラムとして実装され得る1つ以上のウェブアプリケーションを実行することによるなど、ユーザデバイスからの要求に回答してプログラムまたはスクリプトを実行することができる。サーバ(複数可)はまた、これらに限定するものではないが、Oracle(登録商標)、Microsoft(登録商標)、Sybase(登録商標)、及びIBM(登録商標)から市販のものを含むデータベースサーバを含み得る。

40

【 0 0 6 2 】

この環境は、上記に説明されるような様々なデータストア、他のメモリ、及び記憶媒体

50

を含み得る。これらは、コンピュータのうちの1つ以上のローカル（及び/またはそれに固有）か、またはネットワークにわたってコンピュータのうちのいずれかもしくはすべてからリモートの記憶媒体上などの様々な位置に存在することができる。特定の実施形態の集合では、情報は、当業者によく知られている記憶領域ネットワーク（「SAN」）に存在し得る。同様に、コンピュータ、サーバ、またはネットワークデバイスに起因する機能を実施するための任意の必要なファイルは、必要に応じてローカル及び/またはリモートに記憶され得る。システムがコンピュータ化されたデバイスを含む場合、各々のこのようなデバイスは、バスを介して電氣的に連結され得るハードウェア要素を含み得、この要素は、例えば、少なくとも1つの中央処理装置（「CPU」または「プロセッサ」）、少なくとも1つの入力デバイス（例えば、マウス、キーボード、コントローラ、タッチスクリーン、またはキーパッド）、及び少なくとも1つの出力デバイス（例えば、表示デバイス、プリンタ、またはスピーカ）を含む。このようなシステムはまた、ランダムアクセスメモリ（「RAM」）または読み取り専用メモリ（「ROM」）、ならびに取り外し可能なメディアデバイス、メモリカード、フラッシュカードなどのディスクドライブ、光学記憶デバイス及びソリッドステート記憶デバイスなどの1つ以上の記憶デバイスを含み得る。

【0063】

このようなデバイスはまた、コンピュータ可読記憶媒体リーダー、通信デバイス（例えば、モデム、ネットワークカード（無線または有線）、赤外線通信デバイスなど）、及び上記に記載されるようなワーキングメモリを含み得る。コンピュータ可読記憶媒体リーダーは、リモート、ローカル、固定、及び/または取り外し可能な記憶デバイスを表すコンピュータ可読記憶媒体、ならびにコンピュータ可読情報を一時的及び/またはより永久に含み、記憶し、送信し、かつ検索するための記憶媒体と接続されるか、またはそれらを受信するように構成され得る。システム及び様々なデバイスはまた、典型的に、いくつかのソフトウェアアプリケーション、モジュール、サービス、または少なくとも1つのワーキングメモリデバイス内に位置する他の要素を含み、これらは、クライアントアプリケーションまたはウェブブラウザ等のオペレーティングシステム及びアプリケーションプログラムを含む。代替の実施形態が上記に記載されるものから多くの変形を有し得ることが理解されるべきである。例えば、カスタマイズされたハードウェアも使用されてもよく、かつ/または特定の要素がハードウェア、ソフトウェア（アプレット等の携帯用ソフトウェアを含む）、もしくはその両方で実装され得る。さらに、ネットワーク入力/出力デバイス等の他のコンピューティングデバイスへの接続が利用されてもよい。

【0064】

コードまたはコードの部分を含むための記憶媒体及びコンピュータ可読媒体は、限定されないが、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータなどの情報の記憶及び/または送信のための任意の方法または技術で実装された揮発性及び不揮発性の取り外し可能、及び取り外し可能ではない媒体などの記憶媒体及び通信媒体を含む、当該技術分野において既知であるか、または使用される任意の適切な媒体を含み得、これらの媒体としては、RAM、ROM、電氣的消去可能なプログラマブル読み取り専用メモリ（「EEPROM」）、フラッシュメモリもしくは他のメモリ技術、コンパクトディスク読み取り専用メモリ（「CD-ROM」）、デジタル多用途ディスク（DVD）もしくは他の光学記憶、磁気カセット、磁気テープ、磁気ディスク記憶、もしくは他の磁気記憶デバイス、または所望の情報を記憶するために使用され得、かつシステムデバイスによってアクセスされ得る任意の他の媒体が挙げられる。本明細書に提供される本開示及び教示に基づいて、当業者であれば、様々な実施形態を実装する他の手法及び/または方法を理解するであろう。

【0065】

したがって、本明細書及び図面は、限定的意味ではなく例示で見なされる。しかしながら、特許請求の範囲に記載される本発明のより広範な趣旨及び範囲から逸脱することなく、様々な修正及び変更がそれに行われ得ることが明白であろう。

【0066】

10

20

30

40

50

前述は、以下の付記を考慮して理解され得る。

【 0 0 6 7 】

1. データを記憶するためのコンピュータ実装方法であって、
実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で、
データ記憶システムに記憶されるデータオブジェクトを受信することであって、データ記憶システムが、複数の物理的データ記憶位置を含み、複数の物理的データ記憶位置の各位置が、複数の物理的特性を有する、受信することと、

データ断片の真部分集合がデータオブジェクトを構築するのに使用可能であるように構成された複数のデータ断片を生成するように、冗長符号化スキームを適用することと、

複数の物理的データ記憶位置の少なくとも部分集合の物理的特性に適用されたデータ配置規則の集合に少なくとも部分的に基づいて、生成されたデータ断片の配置を決定することであって、配置が、データ配置規則の集合を満たす複数の物理的データ記憶位置からの複数の位置を示す、決定することと、

決定された配置に従ってデータ断片を記憶させることと、を含む、コンピュータ実装方法。

10

【 0 0 6 8 】

2. 物理的特性は、多重記憶デバイスの構成内の位置、データ記憶ラック内の多重記憶デバイスユニットの位置、データ記憶ラックの位置、データセンタ室、データセンタ、及びデータセンタ地理的位置を含む群から選択される1つ以上の物理的特性を含む、付記1に記載のコンピュータ実装方法。

20

【 0 0 6 9 】

3. 物理的特性は、プラッタ、オンプラッタ配置、プラッタ側、及び読み取り/書き込みヘッドを含む群から選択される1つ以上の物理的特性を含む、付記1または2に記載のコンピュータ実装方法。

【 0 0 7 0 】

4. データ配置規則の集合は、データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、付記1～3のいずれか一付記に記載のコンピュータ実装方法。

【 0 0 7 1 】

5. 生成されたデータ断片の配置は、データオブジェクトを受信する前にデータ配置規則を満たすと予め決定された複数の配置からその配置を選択することによって決定される、付記1～4のいずれか一付記に記載のコンピュータ実装方法。

30

【 0 0 7 2 】

6. データ記憶システムは、コンピューティングリソースサービスプロバイダによって複数の顧客に提供されたデータ記憶サービスを支援して動作し、

データオブジェクトは、データ記憶システムに記憶するための顧客のデータを含む、付記1～5のいずれか一付記に記載のコンピュータ実装方法。

【 0 0 7 3 】

7. データを記憶するためのコンピュータ実装方法であって、
実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で、
データ記憶システムに記憶されるデータを受信することであって、データ記憶システムが、複数の物理的データ記憶位置を含み、複数の物理的データ記憶位置の各位置が、複数の物理的特性を有する、受信することと、

40

複数の物理的データ記憶位置の少なくとも部分集合の物理的特性に適用されたデータ配置規則の集合に少なくとも部分的に基づいて、データの断片の配置を決定することであって、配置が、データ配置規則の集合を満たす複数の物理的データ記憶位置からの複数の位置を示す、決定することと、

決定された配置に従って断片を記憶させることと、を含む、コンピュータ実装方法。

【 0 0 7 4 】

8. 冗長符号化スキームを用いて断片を生成することをさらに含む、付記7に記載の

50

コンピュータ実装方法。

【0075】

9．物理的特性は、多重記憶デバイスの構成を含む多重記憶デバイスユニット内の位置、及び1つ以上の多重記憶デバイスユニットを含むデータ記憶ラック内の位置を含む、付記7または8に記載のコンピュータ実装方法。

【0076】

10．物理的特性は、記憶デバイスの複数のプラットフォームからのプラットフォーム上の位置を示す1つ以上の物理的特性を含む、付記7～9のいずれか一付記に記載のコンピュータ実装方法。

【0077】

11．データ配置規則の集合は、データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、付記7～10のいずれか一付記に記載のコンピュータ実装方法。

【0078】

12．断片の集団的サイズは、受信されたデータのサイズを超え、断片の各々は、受信されたデータのサイズ未満の個別サイズを有する、付記7～11のいずれか一付記に記載のコンピュータ実装方法。

【0079】

13．システムであって、
 複数のデータ記憶デバイスであって、各々が、対応する複数の物理的特性を有する1つ以上の物理的データ記憶位置を含む、複数のデータ記憶デバイスと、
 1つ以上のプロセッサと、
 メモリであって、1つ以上のプロセッサによって実行されるとき、システムに、以下のこと、

データの断片の配置を決定することであって、配置が、データ配置規則の集合を満たす対応する物理的特性を有する複数の物理的データ記憶位置から複数の物理的データ記憶位置を示す、決定することと、

配置に従って複数の物理的データ記憶デバイスの少なくとも部分集合の中で断片を記憶することと、を行わせる命令を含むメモリと、を備える、システム。

【0080】

14．物理的特性は、多重記憶デバイスユニット内の位置、及びデータ記憶ラック内の位置を含む群から選択される1つ以上の物理的特性を含む、付記13に記載のシステム。

【0081】

15．物理的特性は、記憶デバイス内の物理的配置を示す1つ以上の物理的特性を含む、付記13または14に記載のシステム。

【0082】

16．データ配置規則の集合は、データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、付記13～15のいずれか一付記に記載のシステム。

【0083】

17．データの断片は、冗長符号化エンジンをデータに適用することによって生成され、

断片は、データのサイズ未満のデータ量を個別に含むがデータのサイズを超えるデータ量を集団的に含む、付記13～16のいずれか一付記に記載のシステム。

【0084】

18．システムは、コンピューティングリソースサービスプロバイダによって複数の顧客に提供されたデータ記憶サービスを支援して動作し、

データの断片は、システム内の記憶のためにコンピューティングリソースサービスプロバイダの顧客のデータを含む、付記13～17のいずれか一付記に記載のシステム。

10

20

30

40

50

【 0 0 8 5 】

19．1つ以上のコンピュータ可読記憶媒体であって、システムの1つ以上のプロセッサによって実行されるとき、システムに、以下のこと、

データ配置規則の集合を、データの断片に、対応する物理的特性を有する複数の物理的データ記憶位置からの複数の位置の中でデータの断片の配置を決定するように、適用することであって、配置が、対応する物理的特性を有する物理的データ記憶位置がデータ配置規則を満たすと決定されたことを示す、適用することと、

データ配置規則の集合の適用に従って、データの断片をデータ記憶システムに持続的に記憶させることと、を行わせる命令をその中に集団的に記憶させる、1つ以上のコンピュータ可読記憶媒体。

10

【 0 0 8 6 】

20．物理的特性は、多重記憶デバイスの構成内の位置、データ記憶ラック内の多重記憶デバイスユニットの位置、データ記憶ラック位置、データセンサ室、データセンタ、及びデータセンタの地理的位置を含む群から選択される1つ以上の物理的特性を含む、付記19に記載の1つ以上のコンピュータ可読記憶媒体。

【 0 0 8 7 】

21．物理的特性は、プラッタ、オンプラッタ配置、プラッタ側、及び読み取り/書き込みヘッドを含む群から選択される1つ以上の物理的特性を含む、付記19または20に記載の1つ以上のコンピュータ可読記憶媒体。

【 0 0 8 8 】

22．データ配置規則の集合は、データ配置規則の集合が満たされるとき、1つ以上の異種混在性の条件を満たすように構成される、付記19～21のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

20

【 0 0 8 9 】

23．データ配置規則の集合を適用することは、データの断片を受信する前にデータ配置規則を満たすと決定された複数の配置から配置を選択することを含む、付記19～22のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

【 0 0 9 0 】

24．コンピュータ可読記憶媒体は、コンピューティングリソースサービスプロバイダによって複数の顧客に提供されたデータ記憶サービスを支援して動作し、

データの断片は、データ記憶システムによって記憶するためのコンピューティングリソースサービスプロバイダの顧客のデータを含む、付記19～23のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

30

【 0 0 9 1 】

25．実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で、

データ記憶システム内の動作を実施する要求を受信することであって、データ記憶システムが、複数の記憶デバイスを含み、複数の記憶デバイスの部分集合の各記憶デバイスが、動作を実施するのに使用可能な記憶デバイスと同時にアクティブであるときに、記憶デバイスの障害を潜在的に引き起こすと決定される、受信することと、

40

複数の記憶デバイスの部分集合の現在の活動状態に少なくとも部分的に基づいて、動作を実施する様式を決定することであって、現在の活動状態が、部分集合の1つ以上の記憶デバイスが現在アクティブであることを示す、決定することと、

決定された様式に従って動作を実施させることと、を含む、コンピュータ実装方法。

【 0 0 9 2 】

26．複数の記憶デバイスの部分集合は、記憶デバイスに物理的に隣接した1つ以上の記憶デバイスを含む、付記25に記載のコンピュータ実装方法。

【 0 0 9 3 】

27．部分集合の記憶デバイスのうちの1つ以上が現在アクティブであると現在の活動状態が示すとき、動作を実施する様式を決定することは、現在の活動状態が、動作の実

50

施が許可される状態に変更するまで動作を遅延させることを決定することを含む、付記 25 または 26 に記載のコンピュータ実装方法。

【0094】

28. 部分集合の記憶デバイスのうちの 1 つ以上が現在アクティブであると現在の活動状態が示すとき、動作を実施する様式を決定することは、動作を実施するための異なる記憶デバイスを選択することを含む、付記 25 ~ 27 のいずれか一付記に記載のコンピュータ実装方法。

【0095】

29. 動作を実施する要求は、データ記憶システムから検索されることを要求されたデータオブジェクトのためのデータを取得する要求である、付記 25 ~ 28 のいずれか一付記に記載のコンピュータ実装方法。

10

【0096】

30. データ記憶システムは、コンピューティングリソースサービスプロバイダによって複数の顧客に提供されたデータ記憶サービスを支援して動作し、

動作を実施する要求は、複数の顧客のうちの 1 人の顧客のデータに関連している、付記 25 ~ 29 のいずれか一付記に記載のコンピュータ実装方法。

【0097】

31. 実行可能な命令によって構成された 1 つ以上のコンピュータシステムの制御下で、

データ記憶デバイスの記憶媒体に少なくとも部分的にアクセスすることによって実施可能な動作を実施する要求を受信することであって、データ記憶デバイスが、データ記憶デバイスの動作を潜在的に妨げるように決定されたデバイスの集合に対応する、受信することと、

20

デバイスの集合の現在の活動状態に少なくとも部分的に基づいて、動作が実施され得る様式を決定することと、

決定された様式で動作を実施することと、を含む、コンピュータ実装方法。

【0098】

32. データ記憶デバイスは、1 つ以上の機械的可動部を利用してデータにアクセスする、付記 31 に記載のコンピュータ実装方法。

【0099】

33. データ記憶デバイスの記憶媒体へのアクセスは、データ記憶デバイス上の読み取り、書き込み、または削除動作を実施する能力を含む、付記 31 または 32 に記載のコンピュータ実装方法。

30

【0100】

34. 現在の活動状態は、デバイスの集合の 1 つ以上のデバイス状態がデータ記憶デバイスの動作を潜在的に妨げるために 1 つ以上の条件を満たすかを示す、付記 31 ~ 33 のいずれか一付記に記載のコンピュータ実装方法。

【0101】

35. エラーを潜在的に引き起こすと各々決定されたデバイスの集合は、要求されたデータ記憶デバイスへの空間的近接のための 1 つ以上の条件を満たす、付記 31 ~ 34 のいずれか一付記に記載のコンピュータ実装方法。

40

【0102】

36. 動作を実施することは、データ記憶デバイスの記憶媒体へのアクセスを遅延させることを含む、付記 31 ~ 35 のいずれか一付記に記載のコンピュータ実装方法。

【0103】

37. システムであって、
複数のデータ記憶デバイスと、

1 つ以上のプロセッサと、

メモリであって、1 つ以上のプロセッサによって実行されるとき、システムに、以下のこと、

50

特定のデータ記憶デバイスの動作を潜在的に妨げると決定された複数のデータ記憶デバイスからのデータ記憶デバイスの集合の現在の活動状態を決定することと、

決定された現在の活動状態に少なくとも部分的に基づいて、動作が特定のデータ記憶デバイスを用いて実施可能である様式を決定することと、

決定された様式に従って動作を実施させることと、を行わせる命令を含むメモリと、を備える、システム。

【0104】

38. 複数のデータ記憶デバイスのうちのデータ記憶デバイスは、1つ以上の機械的可動部を利用して1つ以上の動作を実施する、付記37に記載のシステム。

【0105】

39. 動作は、特定のデータ記憶デバイス上の読み取り、書き込み、または削除動作を含む、付記37または38に記載のシステム。

【0106】

40. 現在の活動状態は、その集合内のデータ記憶デバイスの少なくとも部分集合がアクティブであるかに少なくとも部分的に基づいている、付記37~39のいずれか一付記に記載のシステム。

【0107】

41. 特定のデータ記憶デバイスの動作を潜在的に妨げると決定された記憶デバイスは、特定のデータ記憶デバイスに空間的に近接であると決定される、付記37~40のいずれか一付記に記載のシステム。

【0108】

42. 特定のデータ記憶デバイスに空間的に近接であると決定された記憶デバイスは、共通の取り付け具を共有する、付記41に記載のシステム。

【0109】

43. 特定のデータ記憶デバイスに空間的に近接であると決定された記憶デバイスは、振動伝達媒体を共有する、付記41または42に記載のシステム。

【0110】

44. 動作を実施することは、現在の活動状態が変化するまで動作を遅延させることを含む、付記37~43のいずれか一付記に記載のシステム。

【0111】

45. 1つ以上のコンピュータ可読記憶媒体であって、システムの1つ以上のプロセッサによって実行されるとき、システムに、以下のこと、

動作を実施するデータ記憶システム内のデータ記憶デバイスを選択することと、

選択されたデータ記憶デバイスの動作を潜在的に妨げると決定されたデバイスの現在の活動状態に少なくとも部分的に基づいて、動作を実行する様式を決定することと、

決定された様式に従って動作を実施することと、を行わせる命令をその中に記憶させる、1つ以上のコンピュータ可読記憶媒体。

【0112】

46. デバイスは、データ記憶デバイスを含むデータ記憶デバイスのアレイからのものである、付記45に記載の1つ以上のコンピュータ可読記憶媒体。

【0113】

47. 動作は、選択されたデータ記憶デバイス上の読み取り、書き込み、または削除動作を含む、付記45または46に記載の1つ以上のコンピュータ可読記憶媒体。

【0114】

48. 現在の活動状態は、選択されたデータ記憶デバイスの動作を潜在的に妨げると決定されたデバイスのうちの1つ以上のデバイスがアクティブであるかに少なくとも部分的に基づいている、付記45~47のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

【0115】

49. 動作を実施する様式を決定することは、動作を実施するための異なるデータ記

10

20

30

40

50

憶デバイスを選択することを含む、付記45～48のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

【0116】

50. 選択されたデータ記憶デバイスを用いて実施される動作は、選択されたデータ記憶デバイスの動作を潜在的に妨げると決定された記憶デバイスがアイドル状態にあるまで遅延する、付記45～49のいずれか一付記に記載の1つ以上のコンピュータ可読記憶媒体。

【0117】

他の変形は、本開示の趣旨の範囲内である。したがって、開示される技術は、様々な修正及び代替の構築物の影響を受けやすいが、ある特定の例示されるこれらの実施形態は、
10 図面に示され、詳細に上記に記載されている。しかしながら、本発明を開示された特定の1つの形態または複数の形態に限定する意図はないが、それとは逆に、意図は、添付の特許請求の範囲に規定されるように、本発明の趣旨及び範囲内に入るすべての修正、代替の構築物、及び均等物を網羅することであることが理解されるべきである。

【0118】

開示される実施形態を説明することにおいて（特に以下の特許請求の範囲において）、用語「1つの(a)」、「1つの(an)」、「その(the)」、及び同様の指示語の使用は、特に本明細書に示されない限り、または文脈により明らかに矛盾しない限り、単数形及び複数形の両方を網羅するように解釈されるべきである。用語「備える(comprising)」、「有する(having)」、「含む(including)」、及び「含む(containing)」は、特に記載されない限り、非限定的用語（すなわち、「含むが限定されない」を意味する）と解釈されるべきである。用語「接続された(connecting)」は、修正されず、物理的接続を指すとき、何か介在するものがある場合であっても、部分的または完全に中に含まれ、取り付けられ、またはともに結合されると解釈されるべきである。本明細書の値の範囲の列挙は、特に本明細書に示されない限り、範囲内に入る各別個の値を個別に参照する簡単な方法として機能を果たすことが単に意図され、各別個の値は、本明細書に個別に列挙されるかのように本明細書に組み込まれる。用語「集合」（例えば、「項目の集合」）または「部分集合」は、特に記載されない限り、または文脈により矛盾しない限り、1つ以上の元を含む非空集合と解釈されるべきである。さらに、特に記載されない限り、または文脈により矛盾しない限り、対応する
20 30 集合の用語「部分集合」は、対応する集合の真部分集合を必ずしも示さないが、部分集合及び対応する集合は等しくてもよい。

【0119】

「A、B、及びCのうちの少なくとも1つ」または「A、B及びCのうちの少なくとも1つ」という形態の語句などの接続語は、具体的に特に記載されない限り、または特に文脈により明らかに矛盾しない限り、そうでなければ、項目、用語などがA、B、もしくはC、またはA、B、及びCの集合の任意の非空部分集合のいずれかであり得ることを提示するために一般に使用される文脈で理解される。例えば、上記の接続語で使用される3つの元を有する集合の例示的な実施例では、「A、B、及びCのうちの少なくとも1つ」及び「A、B及びCのうちの少なくとも1つ」は、以下の集合{A}、{B}、{C}、{A、B}、{A、C}、{B、C}、{A、B、C}のうちのいずれかを指す。したがって、このような接続語は、ある特定の実施形態がAのうちの少なくとも1つ、Bのうちの少なくとも1つ、及びCのうちの少なくとも1つが各々存在する必要があることを暗示することが概して意図されない。

【0120】

本明細書に記載されるプロセスの動作は、特に本明細書に示さない限り、または特に文脈により明らかに矛盾しない限り、任意の好適な順序で実施され得る。本明細書に記載されるプロセス（またはこれらの変形及び/または組み合わせ）は、実行可能な命令によって構成された1つ以上のコンピュータシステムの制御下で実施されてもよく、これらのハードウェアまたは組み合わせによって、1つ以上のプロセッサ上で集団的に実行するコー
40 50

ド（例えば、実行可能な命令、1つ以上のコンピュータプログラム、または1つ以上のアプリケーション）として実装されてもよい。このコードは、例えば、1つ以上のプロセッサによって実行可能な複数の命令を含むコンピュータプログラムの形態でコンピュータ可読記憶媒体に記憶されてもよい。コンピュータ可読記憶媒体は、非一時的であってもよい。

【0121】

本明細書に提供された任意及びすべての実施例、または例示的な用語（例えば、「などの（such as）」）の使用は、本発明の実施形態を単により良く説明することが意図され、特に特許請求されない限り、本発明の範囲に制限を課すものではない。本明細書における用語は、本発明の実践に不可欠なものとしてあらゆる特許請求されない要素も示すと解釈されるべきではない。

10

【0122】

本開示の好ましい実施形態が本明細書に記載され、本発明を実施するために本発明者に既知である最良のモードを含む。これらの好ましい実施形態の変形は、前述の説明を読むことで当業者に明らかになり得る。本発明者は、当業者が必要に応じてこのような変形を利用することを予期し、本発明者は、本開示の実施形態に対して、本明細書に具体的に記載されるのとは別の方法で実践されることを意図する。したがって、本開示の範囲は、適用法によって許可されるとき、ここに添付される特許請求の範囲に列挙される主題のすべての修正及び均等物を含む。さらに、すべての可能な変形における上述の要素の任意の組み合わせは、特に本明細書に示されない限り、または特に文脈により明らかに矛盾しない限り、本開示の範囲によって包含される。

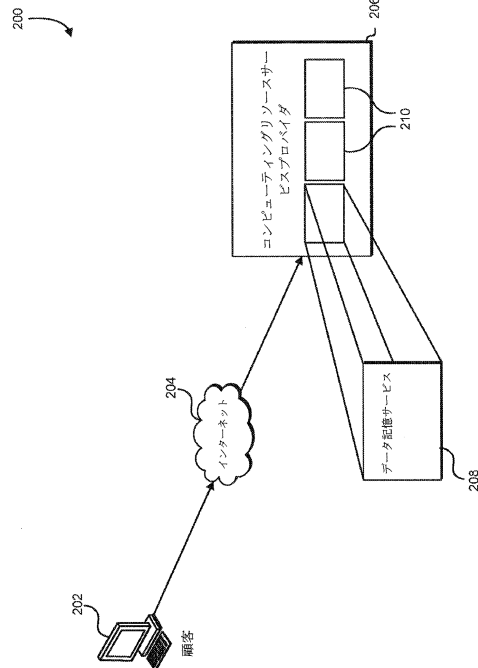
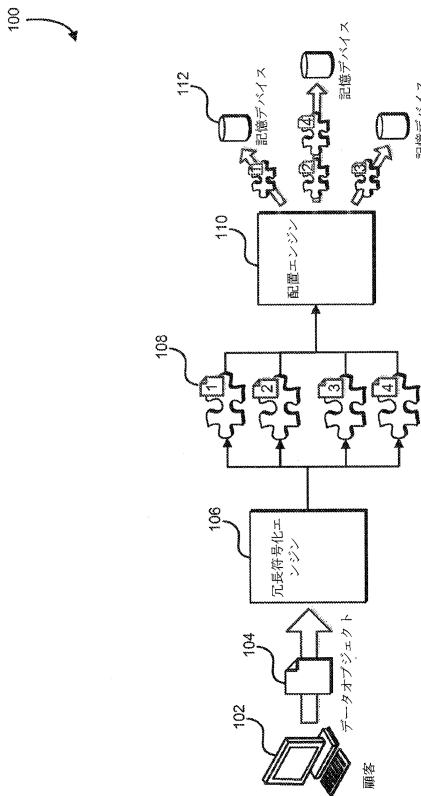
20

【0123】

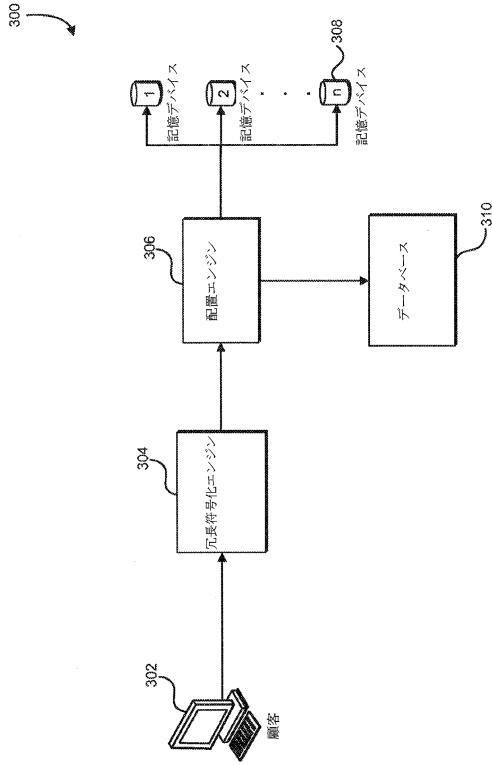
本明細書に引用される公表文献、特許出願、及び特許を含むすべての参考文献は、各参考文献が参照によって組み込まれるように個別かつ具体的に示され、その全体が本明細書に記載されたかのように、同じ程度まで参照により本明細書に組み込まれる。

【図1】

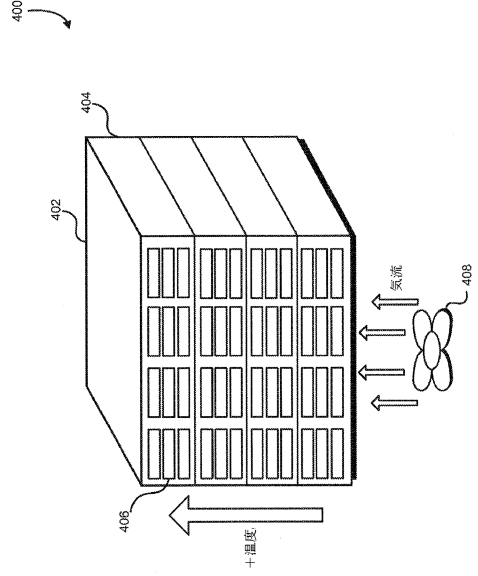
【図2】



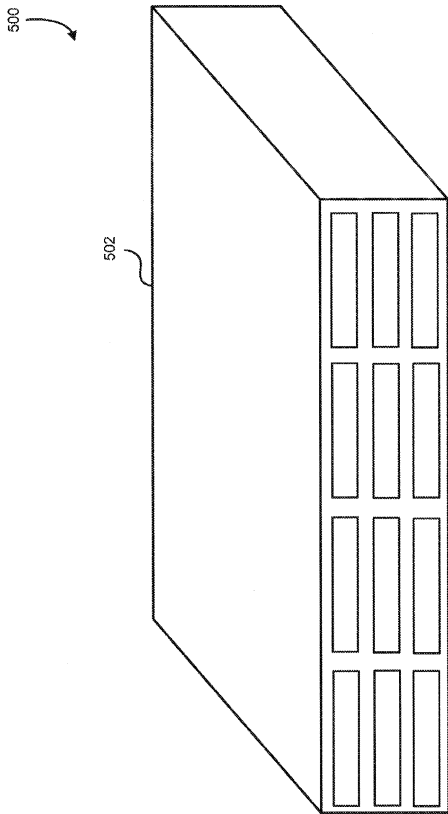
【図3】



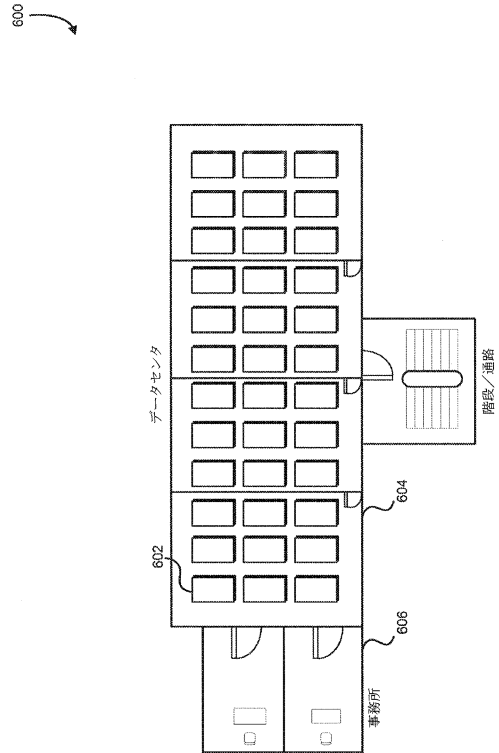
【図4】



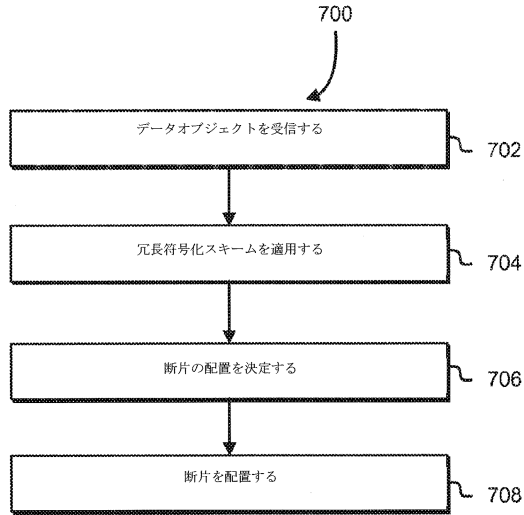
【図5】



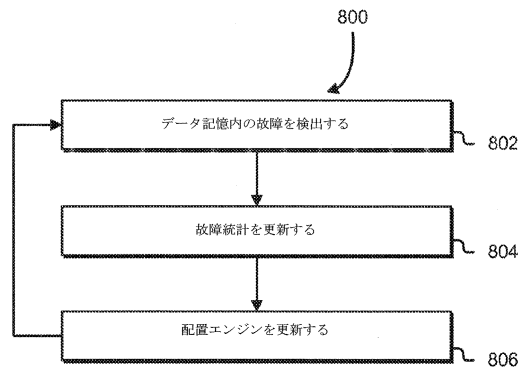
【図6】



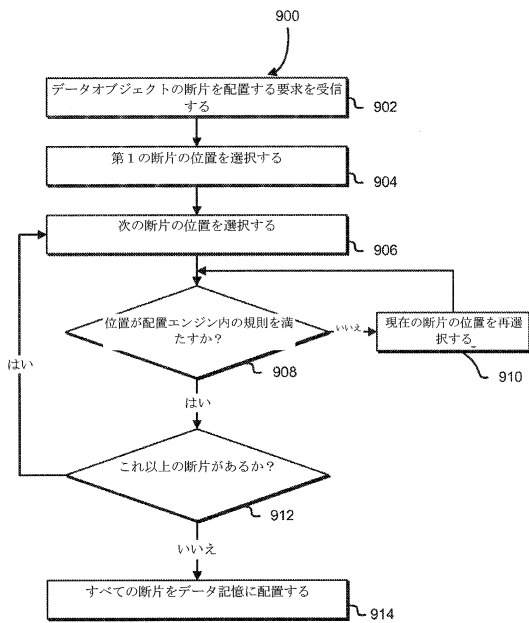
【図7】



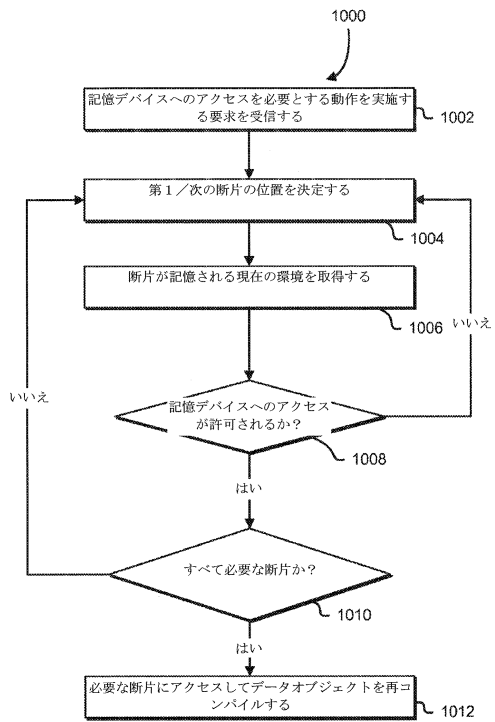
【図8】



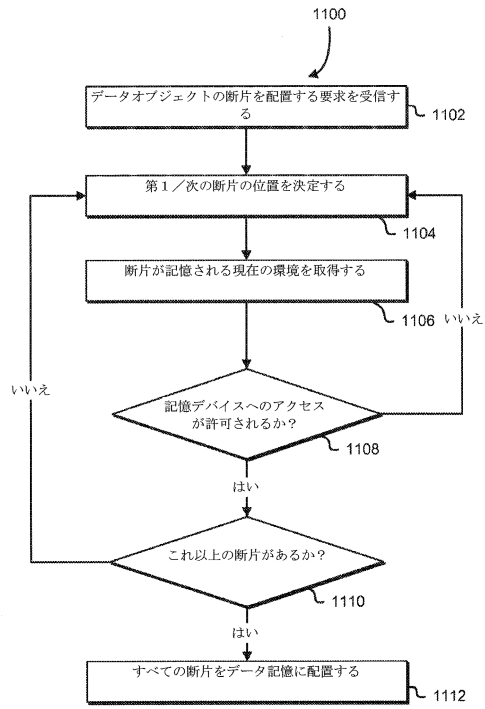
【図9】



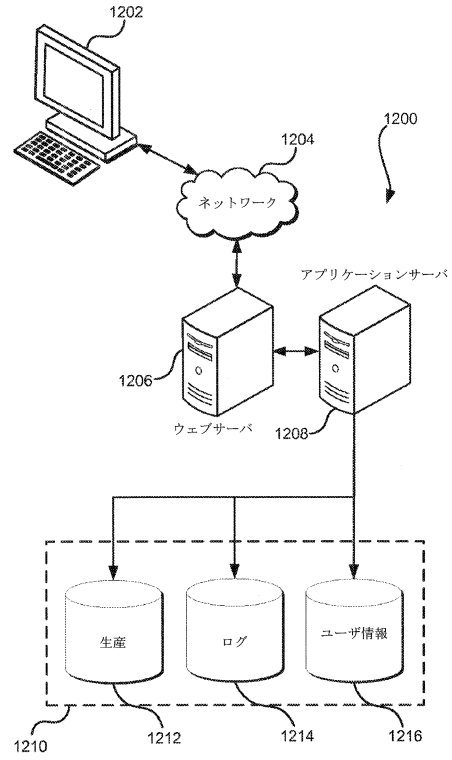
【図10】



【図11】



【図12】



フロントページの続き

- (56)参考文献 特開2007-206993(JP,A)
米国特許出願公開第2010/0023722(US,A1)
特開2007-012061(JP,A)
特開2008-191915(JP,A)
特開2012-018624(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00
G06F 3/06
G06F 13/10