

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6616827号
(P6616827)

(45) 発行日 令和1年12月4日 (2019. 12. 4)

(24) 登録日 令和1年11月15日 (2019. 11. 15)

(51) Int. Cl.

F I

G O 6 F 16/182 (2019. 01)

G O 6 F 16/182

G O 6 F 11/16 (2006. 01)

G O 6 F 11/16 6 6 6

請求項の数 19 (全 20 頁)

(21) 出願番号 特願2017-511880 (P2017-511880)
 (86) (22) 出願日 平成27年9月11日 (2015. 9. 11)
 (65) 公表番号 特表2017-527911 (P2017-527911A)
 (43) 公表日 平成29年9月21日 (2017. 9. 21)
 (86) 国際出願番号 PCT/US2015/049548
 (87) 国際公開番号 W02016/040722
 (87) 国際公開日 平成28年3月17日 (2016. 3. 17)
 審査請求日 平成30年8月1日 (2018. 8. 1)
 (31) 優先権主張番号 14/485, 497
 (32) 優先日 平成26年9月12日 (2014. 9. 12)
 (33) 優先権主張国・地域又は機関
 米国 (US)

(73) 特許権者 314015767
 マイクロソフト テクノロジー ライセン
 シング, エルエルシー
 アメリカ合衆国 ワシントン州 9805
 2 レッドモンド ワン マイクロソフト
 ウェイ
 (74) 代理人 100079108
 弁理士 稲葉 良幸
 (74) 代理人 100109346
 弁理士 大貫 敏史
 (74) 代理人 100117189
 弁理士 江口 昭彦
 (74) 代理人 100134120
 弁理士 内藤 和彦

最終頁に続く

(54) 【発明の名称】 スケーラブルなデータストレージプール

(57) 【特許請求の範囲】

【請求項 1】

方法のためのコンピュータ実行可能命令を実行する際に、1つまたは複数のプロセッサによって実行される方法であって、

ストレージ階層における故障領域およびデータストレージプールにおける利用可能なストレージリソースを記述するデータを1つまたは複数のコンピューティングデバイスによって取得するステップと、

前記ストレージ階層の1つまたは複数のレベル内の前記利用可能なストレージリソースに関連するデバイスの動作特性を、1つまたは複数のコンピューティングデバイスによって確認するステップと、

記述された前記故障領域と、前記ストレージ階層の1つまたは複数のレベル内のデバイスの確認された前記動作特性とに基づいて、前記データストレージプール内の1つまたは複数の特定のデータストレージデバイスに、前記1つまたは複数のコンピューティングデバイスによるメタデータの分配を割り当てるステップとを含み、

前記割り当ては、前記メタデータの少なくとも一部を受信するように構成されたデータストレージプール内の特定数のデータストレージデバイスに対する、故障領域の使用の数を最大化するように実行される方法。

【請求項 2】

前記メタデータは、データがデータストレージプールに記憶される、対応する位置を見

つけるために使用可能である、請求項 1 に記載の方法。

【請求項 3】

前記動作特性は、正常性、データ接続タイプ、メディアタイプ、または現在のメタデータステータスを含む、請求項 1 に記載の方法。

【請求項 4】

割り当てられたメタデータを受信するために使用される、前記データストレージプール内の 1 つまたは複数の特定のデータストレージデバイスの数は、前記データストレージプールについて指定された 1 つまたは複数の弾力的制約条件に少なくとも部分的に基づいている、請求項 1 に記載の方法。

【請求項 5】

前記弾力的制約条件は、動作可能のまま前記データストレージプールによって容認されるデバイス障害の数を特定する、請求項 4 に記載の方法。

【請求項 6】

前記故障領域がデータセンター、ラック、筐体、データストレージデバイス、およびストレージ階層におけるストレージユニットレベルを含む、請求項 1 に記載の方法。

【請求項 7】

指定された数のデータストレージデバイスは、データストレージプールについて指定された 1 つまたは複数の弾力的制約条件に少なくとも部分的に基づいている、請求項 1 に記載の方法。

【請求項 8】

少なくとも部分的にハードウェアに実装された 1 つまたは複数のモジュールを備えるシステムであって、前記 1 つまたは複数のモジュールは、

複数の並列スレッドを使用して複数のデータストレージデバイスを同時にフォーマットするステップと、

同時フォーマットされた前記複数のデータストレージデバイスを単一のトランザクションを介してデータストレージプールに追加するステップと、

前記追加に応じて、前記データストレージプール内の各データストレージデバイスの動作特性に基づいて、前記データストレージプールにおけるメタデータの分配を割り当てるステップと、

前記複数の並列スレッドを使用して前記データストレージプールから少なくとも 1 つまたは複数の、他のデータストレージデバイスを除去するステップと、

前記データストレージプール内に残っている各データストレージデバイスの動作特性に基づいて、前記データストレージプール内のメタデータの分配の再均衡化するステップとを含む動作を実行するように構成されている、システム。

【請求項 9】

前記動作特性は、正常性、データ接続タイプ、メディアタイプ、または現在のメタデータステータスを含む、請求項 8 に記載のシステム。

【請求項 10】

前記割り当ては、前記データストレージプール内の利用可能なストレージリソースのストレージ階層における故障領域に少なくとも部分的に基づく、請求項 8 に記載のシステム。

【請求項 11】

前記割り当ては、前記メタデータの少なくとも一部を受信するように構成されたデータストレージプール内の特定数のデータストレージデバイスに対する、故障領域の使用の数を最大化するように実行される、請求項 10 に記載のシステム。

【請求項 12】

複数のデータストレージデバイスのデータストレージプールへの追加にตอบสนองして、データストレージプール単一データストレージデバイス内のメタデータを更新するステップをさらに含む、請求項 8 に記載のシステム。

【請求項 13】

10

20

30

40

50

少なくとも部分的にハードウェアに実装された１つまたは複数のモジュールを備えるシステムであって、前記１つまたは複数のモジュールは、

データストレージプール内のメタデータの分配を再均衡化するための指示を受信するステップと、

前記データストレージプールについて指定された弾力的制約条件を記述するデータを取得するステップと、

記述された前記弾力的制約条件に基づいて、前記データストレージプール内の１つまたは複数の特定のデータストレージデバイスへのメタデータの再分配を割り当てるステップと

を含む動作を実行するように構成されており、

10

前記割り当ては、前記データストレージプール内の利用可能なストレージリソースのストレージ階層における故障領域に少なくとも部分的に基づき、

前記割り当ては、メタデータの少なくとも一部を受信するように構成されたデータストレージプール内の特定数のデータストレージデバイスのための故障領域の使用の数を最大にするように実行される、システム。

【請求項１４】

前記弾力的制約条件は、動作可能のまま前記データストレージプールによって容認されるデバイス障害の数を特定する、請求項１３に記載のシステム。

【請求項１５】

データストレージプール内のメタデータの割り当て再分配は、データストレージプール内のデバイスの動作特性に少なくとも部分的に基づく、請求項１３に記載のシステム。

20

【請求項１６】

前記動作特性は、正常性、データ接続タイプ、メディアタイプ、または現在のメタデータステータスを含む、請求項１５に記載のシステム。

【請求項１７】

前記複数の並列スレッドを使用して前記データストレージプールから少なくとも１つまたは複数の、他のデータストレージデバイスを除去するステップと、

前記データストレージプール内に残っている各データストレージデバイスの動作特性に基づいて、前記データストレージプール内のメタデータの分配の再均衡化するステップとをさらに含む、請求項１３に記載のシステム。

30

【請求項１８】

メタデータは、データがデータストレージプールに記憶される対応する位置を見つけるために使用可能である、請求項１３に記載のシステム。

【請求項１９】

前記故障領域がデータセンター、ラック、筐体、データストレージデバイス、およびストレージ階層におけるストレージユニットレベルを含む、請求項１３に記載のシステム。

【発明の詳細な説明】

【背景技術】

【０００１】

[0001] データセンタによる「クラウドにおける」データストレージの普及は、増加し続けている。たとえば、ウェブサービスは、ウェブサービスの動作をサポートするため、そしてさらには、ウェブサービスのための単一の目的として機能するため、このデータが様々な異なるデバイスによって容易にアクセスされ得るように、ネットワークの一部としてデータを遠隔的に記憶するように構成され得る。

40

【０００２】

[0002] データストレージプールの使用を通じてデータセンタにおけるデータを管理するための技術が、開発されてきている。データストレージプールは、典型的には、たとえば、データセンタの特定の消費者のための、特定のデータストレージタスクのために構成されているデータストレージデバイスの集合体として、形成される。しかし、これらのデータストレージプールを形成し管理するのに用いられる従来型の技術は、データストレージ

50

プールの一部としてのメタデータストレージの従来型の制限のために、スケーラブルでない。

【発明の概要】

【課題を解決するための手段】

【0003】

[0003] スケーラブルなデータストレージ技術を説明する。1つまたは複数の実現において、ストレージ階層における故障領域とデータストレージプールにおける利用可能なストレージリソースとを記述するデータが、1つまたは複数のコンピューティングデバイスによって、取得される。ストレージ階層の1つまたは複数のレベル内で利用可能なストレージリソースと関連するデバイスの動作特性が、1つまたは複数のコンピューティングデバイスによって、確認される。記述された故障領域とストレージ階層の1つまたは複数のレベル内のデバイスの確認された動作特性とに基づき、1つまたは複数のコンピューティングデバイスによって、メタデータの分配が、データストレージプール内の1つまたは複数の特定のデータストレージデバイスに割り当てられる。

10

【0004】

[0004] 1つまたは複数の実現において、システムは、少なくとも部分的にはハードウェアとして実現された1つまたは複数のモジュールを含む。この1つまたは複数のモジュールは、複数のデータストレージデバイスを、単一のトランザクションを経由して、データストレージプールに追加することを含む動作を実行するように構成されている。追加したことに応答して、データストレージプール内の各データストレージデバイスの動作特性に基づき、データストレージプール内のメタデータの分配が、割り当てられる。

20

【0005】

[0005] 1つまたは複数の実現において、システムは、少なくとも部分的にはハードウェアとして実現された1つまたは複数のモジュールを含む。1つまたは複数のモジュールは、データストレージプールにおけるメタデータの分配を再均衡化するための指示を受け取ることと、データストレージプールに対して特定された弾力的制約条件を記述するデータを取得することとを含む動作を実行するように構成されている。メタデータの再分配は、記述された弾力的制約条件に基づき、データストレージプール内の1つまたは複数の特定のデータストレージデバイスに、割り当てられる。

【0006】

30

[0006] この発明の概要は、下記の発明を実施するための形態でさらに述べられる概念の抜粋を単純化された形式で紹介するために、提供されている。この発明の概要は、特許請求される主題の鍵となる特徴または必要不可欠な特徴を識別するものとして意図されず、また特許請求される主題の範囲を決定する際に役立つものとして用いられることも意図されていない。

【0007】

[0007] 添付の図面を参照しながら、詳細な説明を記述する。図面では、参照番号の最も左の桁は、その参照番号が最初に現れる図面を表す。説明および図面の異なる箇所では同じ参照番号が用いられる場合には、同様または同一の項目を示すことがある。図面に表されているものは、1つまたは複数のものを示すことがあり得るので、論じられている単数または複数の形態のものを、相互交換可能な態様で参照することがある。

40

【図面の簡単な説明】

【0008】

【図1】[0008]スケーラブルなデータストレージプール技術を用いるように動作可能な例示的な実現における環境の図解である。

【図2】[0009]ある例示的な実現における、図1のデータストレージプールが実現され得るストレージ階層を示すシステムの図である。

【図3】[0010]データストレージプール内の故障領域に関する知識を活用することによって、メタデータをストレージ階層内のストレージデバイスに割り当てるための技術が説明されている、例示的な実現における手順を示す流れ図である。

50

【図 4】[0011]データストレージプール内のデータストレージデバイスの管理においてプールマネージャモジュールによって並列処理が用いられる、例示的な実現におけるシステムの図である。

【図 5】[0012]メタデータが割り当てられデータストレージプールにデータストレージデバイスが追加される例におけるシステムの図である。

【図 6】[0013]メタデータが割り当てられデータストレージプールにデータストレージデバイスが追加される例示的な実現における手順を示す流れ図である。

【図 7】[0014]指示に応答してメタデータがデータストレージプールにおいて再均衡化される例におけるシステムの図である。

【図 8】[0015]メタデータがデータストレージプールにおいて再均衡化される例示的な実現における手順の流れ図である。

10

【図 9】[0016]本明細書で説明される技術の実施形態を実現するために、図 1 ~ 図 8 を参照して説明されるいずれかのタイプのコンピューティングデバイスとして実現され得る例示的なデバイスの様々なコンポーネントを含む例示的なシステムの図解である。

【発明を実施するための形態】

【 0 0 0 9 】

概観

[0017] スケーラブルなデータストレージ技術が、説明される。1つまたは複数の実現において、データストレージプールにおけるデータストレージデバイス（たとえば、ハードドライブ）の最大の個数が、それらのデータストレージデバイスの部分集合におけるデータストレージプールにおいて「どこで」データが見出され得るかを記述するメタデータを配置することによって、増加させる技術が説明される。

20

【 0 0 1 0 】

[0018] どのデータストレージデバイスが、メタデータを受け取るように割り当てられるのかは、様々な考察に基づく。たとえば、この割り当ては、データストレージデバイスの正常性、データ接続タイプ、媒体タイプ、データストレージプール内のどこで用いられるか、または現在のメタデータ状態に基づく。さらに、データストレージデバイスによってデータがいかんして記憶されるのか、データセンタ、ラック、およびエンクロージャの全体におけるデータストレージデバイスの分配など、データストレージ階層内の故障領域も、考察される場合がある。このように、データストレージプールは、データストレージプールのストレージ階層内の異なるレベルにおけるデバイスの故障に対する保護を与える。さらに、データストレージデバイスの個数も、用いられるべき故障領域の個数などの考察も、同じように、そのデータストレージプールに対して特定される弾力的制約条件によって、特定される。

30

【 0 0 1 1 】

[0019] さらにまた、データストレージプールの管理は、データストレージデバイスをデータストレージプールに追加する、または、データストレージデバイスをデータストレージプールから削除することの一部として並列処理が用いられる技術を、利用する場合がある。たとえば、複数のスレッドを用いて、複数のデータストレージデバイスが、並列的にフォーマットされ、次に、単一のトランザクションとしてデータストレージプールに追加されることがある。同様の技術が、データストレージデバイスをデータストレージプールから削除するために、用いられることもある。

40

【 0 0 1 2 】

[0020] メタデータの割り当てのための再均衡化技術も、データストレージプールによって用いられることがある。たとえば、メタデータの分配における変更がなされるべきだという指示が、ユーザなどから、ストレージ階層におけるデータストレージデバイスまたは他のデバイスの追加、除去、故障などに起因して、受け取られることがある。それに応答して、メタデータが、データストレージプールの弾力的制約条件と同様に、動作特性や故障領域などの他の上述した考察に基づいて、再均衡化されることがある。これらおよび他の技術に関するさらなる説明は、以下のセクションとの関係で見出される。

50

【 0 0 1 3 】

[0021] 以下の説明では、本明細書で説明されるスケーラブルなデータストレージプール技術を用い得る例示的な環境が、先に説明される。次に、例示的な環境でも他の環境でも同じように実行され得る例示的な手順が、説明される。その結果として、例示的な手順のパフォーマンスは例示的な環境に限定されることはなく、例示的な環境は例示的な手順のパフォーマンスに限定されることはない。

【 0 0 1 4 】

[0022] 以下の説明は、先に述べたシステムおよびデバイスを用いて実現されるスケーラブルなデータストレージプール技術について、説明するものである。各手順の諸態様は、ハードウェア、ファームウェア、ソフトウェア、またはそれらの組合せとして実現される。これらの手順は、1つまたは複数のデバイスによって実行される動作を特定するブロックの組として示されており、それぞれのブロックによる動作を実行するためには、示されている順序に必ずしも限定されない。以下の説明のいくつかの部分では、上述した図面が参照される。

【 0 0 1 5 】

[0023] 図1～図8の例と関係して説明される機能、特徴、および概念は、本明細書において説明される手順の文脈において、用いられ得る。さらに、下記の異なる手順と関係して説明される機能、特徴、および概念は、それらの異なる手順の間で相互に交換され得るのであって、ある個別の手順の文脈における実現に限定されることはない。さらに、本明細書における異なる代表的な手順および対応する図面と関連する複数のブロックは、全体として適用される、および/または、異なる方法で組み合わせられることがあり得る。このように、本明細書における異なる例示的な環境、デバイス、コンポーネント、および手順と関係して説明される個別的な機能、特徴、および概念は、いずれかの適切な組合せで用いられ得るのであって、列挙されている例によって表されている特定の組合せには限定されない。

【 0 0 1 6 】

例示的な環境

[0024] 図1は、本明細書で説明される技術を用いるように動作可能な例示的な実現における環境100の図解である。図示された環境100は、ネットワーク106を経由して通信可能に結合されるストレージサービスプロバイダ102と消費者104を含む。ストレージサービスプロバイダ102を実現し得るコンピューティングデバイス108、110と、消費者104とは、様々な方法で構成される。

【 0 0 1 7 】

[0025] たとえば、コンピューティングデバイス110は、デスクトップコンピュータ、モバイルステーション、娯楽機器、表示デバイスと通信可能に結合されたセットトップボックス、ワイヤレス電話、ゲームコンソールなどのような、ネットワーク106を経由して通信することが可能なコンピュータとして、構成される。このように、コンピューティングデバイス110は、実質的なメモリとプロセッサリソースとを有するフルリソースのデバイス（たとえば、パーソナルコンピュータ、ゲームコンソール）から、限定されたメモリおよび/または処理リソースを有する低リソースのデバイス（たとえば、伝統的なセットトップボックス、ハンドヘルド型のゲームコンソール）までの範囲を有し得る。さらに、コンピューティングデバイスは、ストレージサービスプロバイダなどのコンピューティングデバイス108による場合などのように業務によって動作を実行するのに用いられる複数のサーバなど、複数の異なるデバイスを表す場合もある。

【 0 0 1 8 】

[0026] ネットワーク106はインターネットとして図示されているが、このネットワークは、広い範囲の構成を想定し得る。たとえば、ネットワーク106は、ワイドエリアネットワーク（WAN）、ローカルエリアネットワーク（LAN）、ワイヤレスネットワーク、公衆電話ネットワーク、イントラネットなどを含み得る。さらに、単一のネットワーク106が示されているが、ネットワーク106は、複数のネットワークを含むように構成され

る。

【 0 0 1 9 】

【0027】 ストレージサービスプロバイダ 1 0 2 は、たとえば、ソーシャルネットワークサービスの写真を記憶するなどのウェブサービスの一部として実現されるときなど、消費者 1 0 4 および / またはストレージサービスプロバイダ 1 0 2 自体のデータを記憶するのに用いられ得る複数のデータストレージプール 1 1 2 を含むものとして、図示されている。データストレージプールの管理を含む機能は、プールマネージャモジュール 1 1 4 によって表される。先に述べたように、この管理は、データストレージプール 1 1 2 におけるどのデータストレージデバイスがメタデータを含むべきかという選択を含み、これに関するさらなる説明は、図 2 および図 3 との関係で、見出される。データストレージプール 1 1 2 の管理は、また、データストレージデバイスを同時にフォーマットおよび再フォーマットするなどの並列処理を用い、これに関するさらなる説明は、図 4 ~ 図 6 との関係で見出される。さらに、プールマネージャモジュール 1 1 4 は、ユーザ入力やデータストレージ階層におけるデバイスの故障にตอบสนองしてデータストレージデバイスの追加または削除など、メタデータの再均衡化技術をサポートでき、これについては、図 7 および図 8 との関係でさらに説明される。

10

【 0 0 2 0 】

【0028】 図 2 は、図 1 のデータストレージプール 1 1 2 が実現され得るストレージ階層を示す例示的な実現におけるシステムを示す。この例におけるプールマネージャモジュール 1 1 4 は、メタデータがどのようにデータストレージプールに割り当てられるかを選択するための機能を表すメタデータマネージャモジュール 2 0 2 を含むものとして、図示されている。

20

【 0 0 2 1 】

【0029】 システム 2 0 0 は、異なる地理的位置に配置される 1 つまたは複数のデータセンタ 2 0 4 を含むストレージ階層の例を含んでおり、よって、各データセンタ 2 0 4 は、ある地理的位置におけるデータセンタ 2 0 4 のデバイスの故障（たとえば、設備全体への電力の損失）が別の地理的位置におけるデータセンタの利用可能性によって克服される故障領域を提供する。

【 0 0 2 2 】

【0030】 データセンタ 2 0 4 の各々は、複数のラック 2 0 6 を含む。ラック 2 0 6 は、たとえば、エンクロージャ 2 0 8 を包囲する電力およびスイッチングを含むように構成され、エンクロージャ 2 0 8 はまた、データストレージデバイス 2 1 0 を包囲する電力およびスイッチングを含む。よって、ラック 2 0 6 は、また、ラック 2 0 6 の 1 つにおける故障（たとえば、そのラックにおけるスイッチ）が別のラックによって克服される故障領域を提供し得る。同様に、ラック内のエンクロージャ 2 0 8 も、データが複数のエンクロージャを経由して提供される場合にはこれらのコピーが故障を容認するように用いられるように、あるエンクロージャの故障は別のエンクロージャを用いて克服される故障領域を提供する。またさらに、データストレージデバイス 2 1 0 も、複数のデータストレージデバイス 2 1 0 におけるデータのコピーを通じて同様に故障領域を提供する。ストレージ階層における他のレベルは、たとえばノードなどのストレージユニットのように故障領域を提供することも考えられる。

30

40

【 0 0 2 3 】

【0031】 このように、ストレージ階層内のデバイスは、それらの領域内でデバイスの故障に対する保護を与えるために使用可能な故障領域を提供する。さらに、故障領域は、また、デバイスの故障に対してさらなる保護を与えるために、階層的に用いられることもある。したがって、メタデータマネージャモジュール 2 0 2 は、データストレージプール 1 1 2 のどこにデータが記憶されるのかを記述するメタデータをどのデバイスが受け取るべきかを選択するために、図 2 のデータストレージプール 1 1 2 のストレージ階層内の故障領域に関する知識を利用する。たとえば、メタデータマネージャモジュール 2 0 2 は、特定の弾力的制約条件（たとえば、容認され得る故障の個数）を用いて、この弾力的制約条件を満

50

たしストレージ階層内で利用可能なリソースの故障領域の使用を最大化するメタデータを、いくつかのデバイスに拡大する。この技術に関するさらなる説明は、以下の手順との関係で見出される。

【 0 0 2 4 】

[0032] 図 3 は、データストレージプール内の故障領域に関する知識を活用することによって、メタデータを、ストレージ階層内のストレージデバイスに割り当てるための技術が説明されている、例示的な実現における手順を示している。ストレージ階層における故障領域とデータストレージプールにおける利用可能なストレージリソースとを記述するデータが、プールマネージャモジュール 1 1 4 によって、獲得される（ブロック 3 0 2）。プールマネージャモジュール 1 1 4 は、たとえば、図 1 のデータストレージプール 1 1 2 の管理の一部として、データストレージプール 1 1 2 にデータを記憶するために、どのデータストレージデバイス 2 1 0 が用いられるかを決定する。プールマネージャモジュール 1 1 4 は、また、データストレージプール 1 1 2 の一部として用いられるストレージ階層における他のデバイスを決定することがあり、この対応する故障領域から、データストレージプール 1 1 2 を実現するのに用いられるエンクロージャ 2 0 8、ラック 2 0 6 および異なるデータセンタ 2 0 4 などさえも、決定できる。

10

【 0 0 2 5 】

[0033] プールマネージャモジュール 1 1 4 は、また、ストレージ階層の 1 つまたは複数のレベル内で利用可能なストレージリソースと関連するデバイスの動作特性を確認する（ブロック 3 0 4）。これは、ストレージ階層の他のレベルと関連するデバイスに加えて、データストレージデバイス 2 1 0 の動作特性を確認することを含むのであって、すなわち、エンクロージャ 2 0 8、ラック 2 0 6、およびデータセンタ 2 0 4 のレベルにおける場合など、異なる故障境界を超える場合もある。

20

【 0 0 2 6 】

[0034] 様々な異なる動作特性が、プールマネージャモジュール 1 1 4 によって、確認される。たとえば、デバイスの正常性（ブロック 3 0 6）が確認される場合があり、それによって、非正常なデバイスではなく正常なデバイスを選択するのに、この知識が活用される。特定のタイプのデータストレージデバイス 2 1 0 のバスにより大きな重み（たとえば、他のバスタイプではなく NVMe）、エンクロージャ 2 0 8 とラック 2 0 6 との間やラック 2 0 6 とデータセンタ 2 0 4 との間のデータ接続のタイプなど、を与えるなどのために、データ接続タイプ（ブロック 3 0 8）が確認されることもある。たとえば、より大きな重みを、よって、HDDではなくSSDを選択する蓋然性を与えるなどのために、データストレージデバイス 2 1 0 のための媒体タイプ（ブロック 3 1 0）が確認される場合もある。

30

【 0 0 2 7 】

[0035] プールマネージャモジュール 1 1 4 も、ストレージ階層におけるデバイスがどのように用いられているかを確認する場合があります（ブロック 3 1 2）。たとえば、データストレージデバイス 2 1 0 については、使用のタイプは、マニュアルタイプ、ジャーナルタイプ、ホットスベアタイプなどとは異なり、一般に、データストレージプール 1 1 2 の一部としてデータを受け取り記憶するのにデータストレージデバイス 2 1 0 が利用可能である自動配分を含み得る。1 つまたは複数の実現においては、他の各タイプには、もはや用いられないタイプに対する優先性が与えられるのであって、たとえば、データストレージデバイス 2 1 0 は、データストレージプール 1 1 2 からの除去のための印が付される。現在のメタデータ状態（ブロック 3 1 4）も確認され、それによって、プールマネージャモジュール 1 1 4 が、より大きな重みを（およびしたがって、優先性を）、データストレージデバイス 2 1 0 と、メタデータを記憶するための現在用いられているストレージ階層における関連のデバイスとに与えることがある。

40

【 0 0 2 8 】

[0036] メタデータの分配は、記述された故障領域と、ストレージ階層の 1 つまたは複数のレベル内のデバイスの確認された動作特性とに基づいて、データストレージプール内の 1 つまたは複数の特定のデータストレージデバイスに割り当てられる（ブロック 3 1 6）

50

。たとえば、プールマネージャモジュール304は、たとえば、正常性の次にデータ接続タイプ、その次に媒体タイプ、その次にどのように用いられているか、その次に現在のメタデータ状態などのリスト化された順序で、これらの動作特性により大きな重みを与えるように構成され得る。これらの動作特性の各々がデータストレージデバイス210に対して等しい場合には、メタデータのためにどのデータストレージデバイス210が用いられるべきかという割り当ては、その集合から、ランダムに選択される。さらに、故障領域が、また、たとえば、利用可能であるときの、異なるエンクロージャ208、ラック206、データセンタ204などにおけるデータストレージデバイス210の利用など、データストレージプール122の利用可能なリソースに基づいて、故障境界の個数が最大化されるように、考慮される。

10

【0029】

[0037] メタデータを記憶するのに用いられるべきデータストレージデバイス210の個数は、様々な方法で決定され得る。たとえば、容認されるべき故障の個数など、弾力的制約条件が特定され、上述したストレージ階層にも基づき得るこれらの制約条件に基づいて、データストレージデバイス210の個数が選択される。たとえば、1つのエンクロージャの場合には、そのエンクロージャに対して6つのメタデータドライブが特定され得、2つから4つのエンクロージャの場合には、1エンクロージャ当たり3つのメタデータドライブが特定され、5つから16のエンクロージャの場合には、1エンクロージャ当たり2つのエンクロージャが特定され、16よりも多いエンクロージャの場合には、1エンクロージャ当たり1つのエンクロージャが特定される、などである。各エンクロージャ（または、より高い故障領域）では、最も適格なデバイスから開始するメタデータを有するために、適切なデータストレージデバイス210が、選択される。希望する個数の適格なデバイスが見つからない場合には、適格なデバイスを可能な限り多く含むような選択が行われる。さらに、これらの技術は、データストレージデバイス210が既にメタデータを有する場合には、それは引き続きメタデータを有し続けるという点で、適切である。

20

【0030】

[0038] このようにして、ストレージ階層に表現されたストレージトポロジとその階層におけるデバイスの動作特性とに関する知識が、データストレージプール112の弾力的制約条件によって特定される信頼性の保証に影響を与えないような方法で、データストレージデバイス210の特定の部分集合に、メタデータストレージの範囲を定めるのに用いられる。これらの特徴は、さらに後述されるメタデータの再均衡化を含めて、データストレージデバイス210の追加と削除との一部としても用いられる。

30

【0031】

[0039] 図4は、データストレージプール内のデータストレージデバイスの管理において、プールマネージャモジュール114による並列処理が用いられる、例示的な実現におけるシステム400を示す。従来型の技術を用いて、データストレージプール112を作成する、または、データストレージデバイス210を既存のデータストレージプールに追加するときには、時間の大半は、データストレージデバイス210をフォーマットすること、そして、データストレージデバイス210の各々をデータストレージプール112に追加するためのトランザクションを実行することに費やされる。しかし、この例では、複数のデータストレージデバイスのためのフォーマットを並列的に実行し、次に、単一のトランザクションを経由してこれらのデバイスをデータストレージプールに追加するために、並列処理を用いられて、それにより、データストレージデバイス210を作成してデータストレージプール112に追加するのに要する時間を短縮できる。

40

【0032】

[0040] たとえば、図4のシステム400は、第1の段階402、第2の段階404、および第3の段階406を用いて、図示されている。第1の段階402では、リクエスト408が、ストレージサービスプロバイダ102のプールマネージャモジュール114によって、受け取られる。リクエスト406は、たとえば、ネットワーク106を経由してアクセスを求めるために通信モジュール116を用いている消費者104、ストレージサービ

50

スプロバイダ 1 0 2 と関連する技術者などにより生じる。さらに、リクエスト 4 0 6 は、データストレージプール 1 1 2 の作成を生じさせ、既存のデータストレージプール 1 1 2 にデータストレージデバイス 2 1 0 を追加するなどのように構成される。

【 0 0 3 3 】

[0041] 第 2 の段階 4 0 4 では、プールマネージャモジュール 1 1 4 が、複数のデータストレージデバイス 2 1 0 (1) ~ 2 1 0 (N) を並列にフォーマットするために、並列処理を用いる。たとえば、プールマネージャ 1 1 4 は、対応するデータストレージデバイス 2 1 0 (1) ~ 2 1 0 (N) を同時にフォーマットするために各スレッドが用いられるように、複数のスレッド 4 1 0 (1) ~ 4 1 0 (N) を用いる。これは、デバイスの各々に対する安全な分割などのフォーマットを含む。フォーマットされるべきデータストレージデバイスの個数が利用可能なスレッドの個数よりも少ない場合には、これは、1 回のパスで実行される。

10

【 0 0 3 4 】

[0042] そうでない場合には、いったんスレッドによるフォーマットが完了されると、別のデータストレージデバイスがフォーマットのためにキューから選択されるように、ワークキューが用いられる。たとえば、3 2 のスレッドが利用可能であるが、6 4 のデータストレージデバイスがフォーマットされなければならない場合には、3 2 のスレッドの各々が、フォーマットを実行するために、同時に、そして、連続的に、用いられることがある。しかし、スレッドの方が先にフォーマットを終了する場合には、そのスレッドは、フォーマットのタスクが実行されなければならない場合にはどのスレッドもアイドル状態ではないように、直ちにキューからデータストレージデバイスを選択することもある。このようにして、スレッドのグループは、希望通りに、フォーマットを反復する。

20

【 0 0 3 5 】

[0043] 第 3 の段階 4 0 6 では、第 2 の段階 4 0 4 からの、フォーマットされたデータストレージデバイス 2 1 0 (1) ~ 2 1 0 (N) が、従来型の技術を用いて実行される場合のように各個別的なデータストレージデバイス 2 1 0 (1) から 2 1 0 (N) のために別個のトランザクションを用いるというよりむしろ単一のトランザクションで、データストレージプール 1 1 2 に追加される。こうして、この例示的な並列処理では、単一のトランザクションの使用は、データストレージデバイス 2 1 0 を作成またはデータストレージプール 1 1 2 に追加するのに要する時間量を、短縮するのに用いられる。この並列処理は、また、メタデータの割り当てと共に実行され、それについてのさらなる説明は、以下に見出され、対応する図面において示されている。

30

【 0 0 3 6 】

[0044] 図 5 は、メタデータが割り当てられ、データストレージプールにデータストレージデバイスが追加される例におけるシステム 5 0 0 を示している。図 6 は、メタデータが割り当てられ、データストレージプールにデータストレージデバイスが追加される例示的な実現における手順 6 0 0 を示している。したがって、以下の説明は、図 5 および図 6 の両方を交換可能に参照する。複数のデータストレージデバイスが、複数のスレッドを同時に用いて、1 つまたは複数のコンピューティングデバイスにより、フォーマットされる (ブロック 6 0 2) 。プールマネージャモジュール 1 1 4 を実現するコンピューティングデバイスは、たとえば、図 3 との関係で上述されたように、複数のデータストレージデバイス 2 1 0 (1) ~ 2 1 0 (N) をフォーマットするために、複数のスレッドを用いる。

40

【 0 0 3 7 】

[0045] 第 1 の段階 5 0 2、第 2 の段階 5 0 4、第 3 の段階 5 0 6、および第 4 の段階 5 0 8 が、図 5 のシステム 5 0 0 において図示されている。第 1 の段階 5 0 2 では、新たなストレージプールを作成するときに、そのストレージプールの 1 つまたは複数のデータストレージデバイスが、データストレージプール 1 1 2 のためのメタデータを記憶するのに用いられるために、選択される。たとえば、1 0 0 のデータストレージデバイスをデータストレージプール 1 1 2 に追加するためのリクエストが行われる。したがって、これらのデバイスの 1 つまたは複数は、これらの追加的なデバイスのためのメタデータを記憶する

50

ように構成され、データストレージプール 1 1 2 に追加される。

【 0 0 3 8 】

[0046] 次に、第 2 の段階 5 0 4 では、複数のデータストレージデバイス 2 1 0 (2) ~ 2 1 0 (1 0 0) が、単一のトランザクションを経由して、1 つまたは複数のコンピューティングデバイスにより、データストレージプールに追加される (ブロック 6 0 6) 。先の例を継続すると、単一のトランザクションは、プールを作成し、メタデータを選択されたドライブに書き込むために用いられ得る。

【 0 0 3 9 】

[0047] データストレージデバイスが既存のデータストレージプールに追加される別の例では、単一のトランザクションは、これらのデバイスをデータストレージプールに追加するためのメタデータを記憶するのに用いられる既存のデータストレージデバイスに向かう。したがって、第 3 の段階 5 0 6 におけるそのような例では、メタデータの更新 5 1 2 は、これらの追加的なデータストレージデバイス 2 1 0 (2) ~ 2 1 0 (1 0 0) もメタデータストレージデバイス 5 0 8 自体も同様に、データストレージプール 1 1 2 の一部として含まれるように、単一のトランザクションでデータストレージプール 1 1 2 のメタデータストレージデバイス 5 0 8 に提供される。

【 0 0 4 0 】

[0048] 第 4 の段階 5 0 8 では、メタデータの更新された割り当て 5 1 4 が、これらの追加的なデータストレージデバイスと、関連するデバイスとをストレージ階層においてアドレス指定するために、データストレージプール 1 1 2 に対して実行される。たとえば、データストレージプール 1 1 2 内の 1 つまたは複数の特定のデータストレージデバイスへのメタデータの分配は、データストレージプール内のデータストレージデバイスの確認された動作特性に基づいて、割り当てられる (ブロック 6 1 0) 。データストレージデバイスが既存のプールに追加されるシナリオについては、これは、データストレージプール 1 1 2 に先に含まれていたデータストレージデバイスと同様にデータストレージプール 1 1 2 に追加された追加的なデータストレージデバイスの動作特性を含み得る。この割り当ては、図 2 および図 3 との関係で先に述べた技術を用いるなど、様々な方法で、実行される。

【 0 0 4 1 】

[0049] メタデータは、次に、割り当てられた分配に従って、シーケンシャルに、一度に 1 つずつ、1 つまたは複数の特定のデータストレージデバイスに分配される (ブロック 6 1 2) 。上述のシナリオに従い、データストレージデバイスが既存のデータストレージプールに追加される場合には、この分配は、故障に対する保護を与えるために、割り当てられた特定のデータストレージデバイスに対して、シーケンシャルに、実行される。

【 0 0 4 2 】

[0050] たとえば、メタデータの分配の間に故障に遭遇する場合の「過半数が勝利する」アプローチにおいては、データストレージプールにおけるデータストレージデバイスを、たとえば「ロールバック」、すなわち、新たな割り当てに向かう移動、に設定する状態を決定するために、投票技術が実行される。このようにして、デバイスをシーケンシャルに追加することによって、過半数が維持される。しかし、データストレージプール 1 1 4 が作成されているシナリオでは、そのような関心は、どれも妥当ではなく、したがって、割り当てられた特定のデータストレージデバイスが、たとえば、上述した並列処理を通じて、メタデータの分配を同時に受け取る。

【 0 0 4 3 】

[0051] データストレージプール 1 1 2 へのデータストレージデバイス 2 1 0 の追加について説明してきたが、並列処理および単一トランザクションに関するこれらの技術は、複数のデータストレージデバイスのデータストレージプールからの除去にも等しく適用可能であることは、容易に明らかなはずである。先の例を継続すると、追加されたデータストレージデバイスは、データストレージプール 1 1 2 において、既存のデータストレージデバイスを置き換えるために用いられ得る。その結果として、新たなデバイスが最初に追加され、メタデータが再分配される。除去されるべきデバイスが、次に、排除されたものと

10

20

30

40

50

して印が付され、その次には、これらの印が付されたデバイスからメタデータが除去されるように、メタデータの割り当てが実行され、スレッドの使用を通じた並列処理が、排除されるデバイスをデータストレージプール 112 から除去するのに用いられることになり、それらのデバイスの再フォーマットなどがなされる。このように、並列処理と単一トランザクションとの使用が、データストレージプール 112 とプールマネージャモジュール 114 とにより、従来型の技術と比較して効率とリソース消費とを向上させるために、用いられる。

【0044】

[0052] 図 7 は、指示に応答してメタデータがデータストレージプールにおいて再均衡化される例におけるシステム 700 を、示している。図 8 は、メタデータがデータストレージプールにおいて再均衡化される例示的な実現における手順 800 を示している。したがって、以下の説明は、図 7 および図 8 の両方を、交換可能に、参照する。

10

【0045】

[0053] 指示が、1つまたは複数のコンピューティングデバイスによって、データストレージプールにおけるメタデータの分配を再均衡化するために、受け取られる（ブロック 802）。第 1 の段階 702 では、たとえば、再均衡化の指示 708 が、ストレージサービスプロバイダ 112 のプールマネージャモジュール 114 によって受け取られる。再均衡化の指示 708 は、先に述べたように、消費者 104 または技術者から生じ得、データストレージプール 112 のストレージ階層内のデバイスの故障、追加、または削除の検出に応答して、プールマネージャモジュール 114 によって自動的に生成され得る、などである。

20

【0046】

[0054] 容認される故障の個数やデフォルトで特定されるものなど、データストレージプールに対して特定された弾力的制約条件を記述するデータが、取得される（ブロック 804）。したがって、メタデータの再分配は、記述された弾力的制約条件に基づいて、データストレージプール内の 1つまたは複数の特定のデータストレージデバイスに割り当てられ得る（ブロック 806）。たとえば、図 2 ~ 図 6 との関係で上述したように、メタデータ選択 710 技術は、動作特性、故障境界などに基づいて、データストレージプール 112 内の特定のデータストレージデバイス 210 (1) ~ 210 (N) 上のストレージにメタデータを割り当てするのに用いられ得る。さらに、弾力的制約条件 712 は、また、たとえば、いくつかの故障をサポートするいくつかのデータストレージデバイスを特定するためなど、この割り当ての一部としても、用いられ得る。

30

【0047】

[0055] 第 3 の段階 706 では、再均衡化動作 714 が実行され、これは、複数のトランザクションの使用や並列処理などを含み得るのであって、上述したように、それは、データストレージプール 112 におけるメタデータを、再び割り当てる。このようにして、データストレージプール 112 のプールマネージャモジュール 114 による管理は、スケーラビリティとパフォーマンスとのゴールをサポートするために、複数の異なる考察のうちの 1つまたは複数扱うことになる。

【0048】

例示的なシステムおよびデバイス

40

[0056] 図 9 は、本明細書で説明される様々な技術を実現し得る 1つもしくは複数のコンピューティングシステムおよび/またはデバイスを表す例示的なコンピューティングデバイス 902 を含む例示的なシステムを、全体として 900 において図示している。これは、プラットフォームによって抽象化されたリソースの一部としてプールマネージャモジュール 114 を含むことを通じて、図示されている。たとえば、コンピューティングデバイス 902 は、サービスプロバイダのサーバ、クライアントと関連するデバイス（たとえば、クライアントデバイス）、オンチップシステム、および/または任意の他の適切なコンピューティングデバイスもしくはコンピューティングシステムであり得る。

【0049】

[0057] 図示されている例示的なコンピューティングデバイス 902 は、処理システム 9

50

04、1つまたは複数のコンピュータ可読媒体906、および相互通信可能に結合されている1つまたは複数のI/Oインタフェース908を含む。図示されていないが、コンピューティングデバイス902は、様々なコンポーネントを相互に結合するシステムバスまたは他のデータおよびコマンド転送システムを、さらに含み得る。システムバスは、メモリバスもしくはメモリコントローラ、周辺バス、ユニバーサルシリアルバス、および/または、様々なバスアーキテクチャのうちのいずれかを用いるプロセッサバスもしくはローカルバスなど、異なるバス構造のいずれか1つまたはそれらの組合せを含み得る。制御線およびデータ線など、多様な他の例も、考察される。

【0050】

[0058] 処理システム904は、ハードウェアを用いて1つまたは複数の動作を実行する機能を表す。したがって、処理システム904は、プロセッサ、機能ブロックなどとして構成され得るハードウェア要素910を含むものとして、図示されている。これは、特定用途向け集積回路、または、1つまたは複数の半導体を用いて形成される他のロジックデバイスとしての、ハードウェアによる実現を含む。ハードウェア要素910は、ハードウェア要素910がそれによって形成される材料によって、または、ハードウェア要素910において用いられる処理機構によって、限定されることはない。たとえば、プロセッサは、1つまたは複数の半導体および/またはトランジスタ（たとえば、電子集積回路（IC））から構成され得る。そのような文脈では、プロセッサ実行可能な命令とは、電子的に実行可能な命令である。

【0051】

[0059] コンピュータ可読媒体906は、メモリ/ストレージ912を含むものとして、図示されている。メモリ/ストレージ912は、1つまたは複数のコンピュータ可読媒体と関連するメモリ/ストレージ能力を表す。メモリ/ストレージ912は、揮発性媒体（ランダムアクセスメモリ（RAM）など）および/または不揮発性媒体（リードオンリメモリ（ROM）、フラッシュメモリ、光ディスク、磁気ディスクなど）を含む。メモリ/ストレージコンポーネント912は、固定された媒体（たとえば、RAM、ROM、固定ハードドライブなど）と同様に着脱可能な媒体（たとえば、フラッシュメモリ、着脱可能なハードドライブ、光ディスクなど）も含む。コンピュータ可読媒体906は、さらに後述されるように、様々な他の方法で構成されることもある。

【0052】

[0060] 1つまたは複数の入力/出力インタフェース908は、ユーザが様々な入力デバイスを用いて、コマンドおよび情報をコンピューティングデバイス902に入力することを可能にすると共に、様々な出力デバイスを用いて、情報をユーザおよび/または他のコンポーネントもしくはデバイスに提示することを可能にする機能を表す。入力デバイスの例は、キーボード、カーソル制御デバイス（たとえば、マウス）、マイクロフォン、スキャナ、タッチ機能（たとえば、物理的タッチを検出するように構成される静電容量式センサまたは他のセンサ）、（たとえば、タッチを含まないジェスチャとしての動きを認識するために、赤外線周波数などの可視波長または非可視波長を用いる）カメラなどを含む。出力デバイスの例は、ディスプレイデバイス（たとえば、モニターやプロジェクタ）、スピーカ、プリンタ、ネットワークカード、触覚応答デバイスなどを含む。したがって、コンピューティングデバイス902は、ユーザとの対話をサポートするために、さらに後述されるように、多様な方法で構成される。

【0053】

[0061] 様々な技術が、ソフトウェア、ハードウェア要素、またはプログラムモジュールの一般的な文脈において、本明細書で説明される。一般に、そのようなモジュールは、特定のタスクを実行するまたは特定の抽象的なデータタイプを実現するルーチン、プログラム、オブジェクト、要素、コンポーネント、データ構造などを含む。本明細書で用いられる「モジュール」、「機能」、および「コンポーネント」という用語は、一般に、ソフトウェア、ファームウェア、ハードウェア、またはこれらの組合せを表す。本明細書で説明される技術の特徴は、プラットフォームには依存することなく、つまり、その技術が、

10

20

30

40

50

多様なプロセッサを有する多様な商用コンピューティングプラットフォームにおいて実現され得る、ということを意味している。

【 0 0 5 4 】

[0062] 説明されているモジュールおよび技術の実現は、何らかの形態のコンピュータ可読媒体に記憶され得る、または、何らかの形態のコンピュータ可読媒体を経由して伝送され得る。コンピュータ可読媒体は、コンピューティングデバイス 902 によってアクセスされ得る多様な媒体を含み得る。限定ではなく例として、コンピュータ可読媒体は、「コンピュータ可読記憶媒体」と「コンピュータ可読信号媒体」とを含む。

【 0 0 5 5 】

[0063] 「コンピュータ可読記憶媒体」とは、単なる信号伝送、搬送波、または信号自体とは対照的に、情報の永続的および／もしくは非一時的な記憶を可能にする媒体ならびに／またはデバイスを指し得る。したがって、コンピュータ可読記憶媒体とは、非信号伝送媒体を指す。コンピュータ可読記憶媒体は、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、ロジック要素／回路、または他のデータなどの情報を記憶するのに適した方法または技術により実現された、揮発性および不揮発性の着脱可能および着脱不可能な媒体および／またはストレージデバイスなどのハードウェアを含む。コンピュータ可読記憶媒体の例は、RAM、ROM、EEPROM、フラッシュメモリ、もしくは他のメモリ技術、CD-ROM、デジタルバーサタイルディスク(DVD)、もしくは他の光ストレージ、ハードディスク、磁気カセット、磁気テープ、磁気ディスクストレージ、もしくは他の磁気記憶デバイス、もしくは、所望の情報を記憶するのに適した、コンピュータがアクセスできる他の記憶デバイス、有体物である媒体、または製品を含み得るが、これらには限定されない。

【 0 0 5 6 】

[0064] 「コンピュータ可読信号媒体」とは、ネットワークを経由してなど、コンピューティングデバイス 902 のハードウェアに命令を伝送するよう構成された信号伝送媒体を指してもよい。信号媒体は、典型的には、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、または他のデータを、搬送波、データ信号、または他の伝送機構などの変調されたデータ信号において具現化できる。信号媒体は、任意の情報配信媒体も含む。「変調されたデータ信号」という用語は、信号における情報をエンコードするように設定もしくは変更されたその特性セットのうちの1つまたは複数の特性を有する信号を意味する。限定ではなく例として、通信媒体は、ワイヤードネットワークまたは直接配線接続などの有線媒体と、音響、無線周波数(RF)、赤外線、および他の無線媒体などの無線媒体と、を含む。

【 0 0 5 7 】

[0065] 前述したように、ハードウェア要素 910 およびコンピュータ可読媒体 906 は、1つまたは複数の命令を実行するなどの、本明細書で説明される技術の少なくともいくつかの態様を実現するためにいくつかの実施形態において用いられ得る、ハードウェアとして実現された、モジュール、プログラマブルデバイスロジック、および／または固定デバイスロジックを表す。ハードウェアは、集積回路もしくはオンチップシステムのコンポーネント、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、コンプレックスプログラマブルロジックデバイス(CPLD)、および、シリコンによる他の実現物または他のハードウェアを含み得る。この文脈では、ハードウェアは、そのハードウェアにより具現化される命令および／またはロジックにより定義されるプログラムタスクを実行する処理デバイスとして、また同様に、たとえば前述したコンピュータ可読記憶媒体なども、実行のための命令を記憶するために利用されるハードウェアとして動作できる。

【 0 0 5 8 】

[0066] 前述したものの組合せは、本明細書で説明される様々な技術を実現するためにも用いられる。したがって、ソフトウェア、ハードウェア、または実行可能なモジュールは、何らかの形態のコンピュータ可読記憶媒体上で具現化されるおよび／または1つまたは

複数のハードウェア要素 910 により具現化される 1 つまたは複数の命令および / またはロジックとして実現される。コンピューティングデバイス 902 は、ソフトウェアモジュールおよび / またはハードウェアモジュールに対応する特定の命令および / または機能を実現するように構成され得る。したがって、ソフトウェアとしてコンピューティングデバイス 902 により実行可能なモジュールの実現は、たとえば、処理システム 904 のコンピュータ可読記憶媒体および / またはハードウェア要素 910 の使用を通じて、ハードウェアとして少なくとも部分的に実現され得る。命令および / または機能は、本明細書で説明される技術、モジュール、および例を実現する 1 つまたは複数の製品 (たとえば、1 つまたは複数のコンピューティングデバイス 902 および / または処理システム 904) により、実行可能 / 動作可能である。

10

【0059】

[0067] 図 9 にさらに図示されているように、例示的なシステム 900 は、パーソナルコンピュータ (PC)、テレビジョンデバイス、および / またはモバイルデバイス上でアプリケーションを実行する際に、シームレスなユーザエクスペリエンスのためのユビキタス環境を可能にする。サービスおよびアプリケーションは、アプリケーションを利用している間、ビデオゲームをプレイしている間、ビデオを見ている間などに、あるデバイスから次のデバイスに移る際の共通のユーザエクスペリエンスのために、3 つすべての環境において実質的に同様に動作する。

【0060】

[0068] 例示的なシステム 900 では、複数のデバイスが、中央コンピューティングデバイスを通じて、相互接続される。中央コンピューティングデバイスは、複数のデバイスに対してローカルな場合があり得、または、複数のデバイスからリモートに位置する場合もあり得る。ある実施形態では、中央コンピューティングデバイスは、ネットワーク、インターネット、または他のデータ通信リンクを通じて複数のデバイスに接続されている 1 つまたは複数のサーバコンピュータのクラウドである。

20

【0061】

[0069] ある実施形態では、この相互接続アーキテクチャは、複数のデバイスのユーザに共通でシームレスなエクスペリエンスを提供するように、複数のデバイスにわたって機能を配信することを可能にする。複数のデバイスの各々は、異なる物理的要件および能力を有し得、中央コンピューティングデバイスは、各デバイス向けにテイラーメイドに作成されながらもすべてのデバイスに共通なエクスペリエンスを各デバイスに提供することを可能にするプラットフォームを用いる。ある実施形態において、ターゲットデバイスのクラスが作成され、エクスペリエンスが、デバイスの一般的なクラスに向けてテイラーメイドされる。デバイスのクラスは、デバイスの物理的特徴、使用のタイプ、または他の共通な特性により定義される。

30

【0062】

[0070] 様々な実現において、コンピューティングデバイス 902 は、コンピュータ 914、モバイル 916、およびテレビジョン 918 の使用のためなど、様々な異なる構成を想定し得る。これらの構成の各々は、一般に異なる構成および能力を有するデバイスを含んでおり、したがって、コンピューティングデバイス 902 は、異なるデバイスクラスのうちの 1 つまたは複数に従って構成される。たとえば、コンピューティングデバイス 902 は、パーソナルコンピュータ、デスクトップコンピュータ、マルチスクリーンコンピュータ、ラップトップコンピュータ、ネットブックなどを含むデバイスのコンピュータ 914 クラスとして実現される。

40

【0063】

[0071] コンピューティングデバイス 902 は、また、携帯電話機、ポータブル音楽プレーヤ、ポータブルゲームデバイス、タブレットコンピュータ、マルチスクリーンコンピュータなどのモバイルデバイスを含む、デバイスのモバイル 916 クラスとしても実現され得る。コンピューティングデバイス 902 は、また、一般に、くつろいだ視聴環境において、より大型のスクリーンを有するデバイスまたはより大型のスクリーンに接続されるデ

50

バイスを含むデバイスのテレビジョン 9 1 8 クラスとしても、実現され得る。これらのデバイスは、テレビジョン、セットトップボックス、ゲームコンソールなどを含む。

【 0 0 6 4 】

[0072] 本明細書で説明される技術は、コンピューティングデバイス 9 0 2 のこれら様々な構成によりサポートされ、本明細書で説明される技術の特定の例に限定されることはない。この機能は、後述されるように、プラットフォーム 9 2 2 を経由して「クラウド」 9 2 0 における分散システムの使用により、全部または一部が実現される。

【 0 0 6 5 】

[0073] クラウド 9 2 0 は、リソース 9 2 4 のためのプラットフォーム 9 2 2 を含む、および / または、そのようなプラットフォーム 9 2 2 を表す。プラットフォーム 9 2 2 は、クラウド 9 2 0 のハードウェアリソース（たとえば、サーバ）とソフトウェアリソースの基礎的な機能を、抽象化している。リソース 9 2 4 は、コンピューティングデバイス 9 0 2 からリモートにあるサーバ上でコンピュータ処理が実行されている間に用いることが可能なアプリケーションおよび / またはデータを含み得る。リソース 9 2 4 は、また、インターネット、および / または、セルラネットワークもしくは W i - F i ネットワークなどの加入者ネットワークを通じて提供されるサービスも含む。

【 0 0 6 6 】

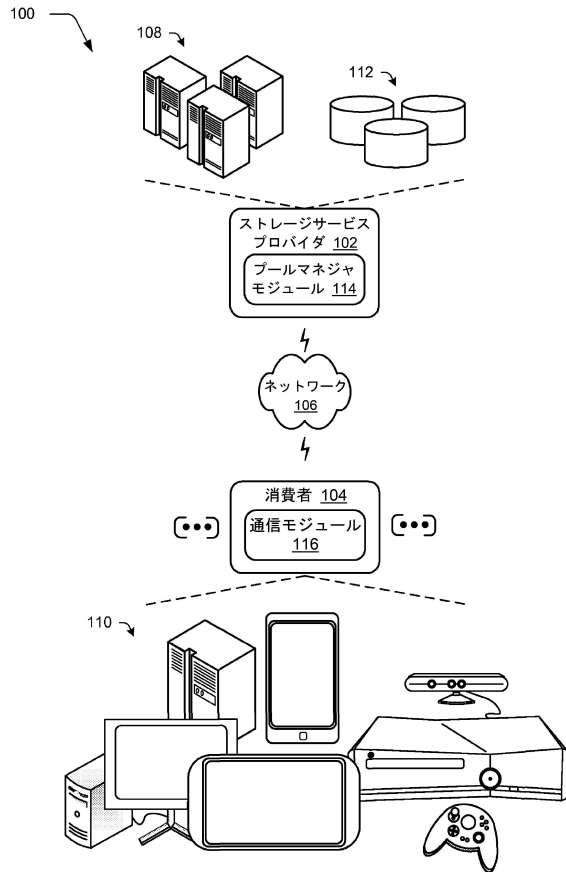
[0074] プラットフォーム 9 2 2 は、コンピューティングデバイス 9 0 2 と他のコンピューティングデバイスとを接続するためのリソースおよび機能を抽象化し得る。プラットフォーム 9 2 2 は、また、リソースのスケーリングを抽象化し、プラットフォーム 9 2 2 を経由して実現されるリソース 9 2 4 を求める直面されている要求に、対応するレベルのスケールを提供するように、機能する。したがって、相互接続されたデバイスの実施形態では、本明細書で説明される機能の実現は、システム 9 0 0 の全体に、分散される。たとえば、本明細書で説明される機能は、コンピューティングデバイス 9 0 2 において部分的に実現されると共に、クラウド 9 2 0 の機能を抽象化するプラットフォーム 9 2 2 を経由しても実現される。

【 0 0 6 7 】

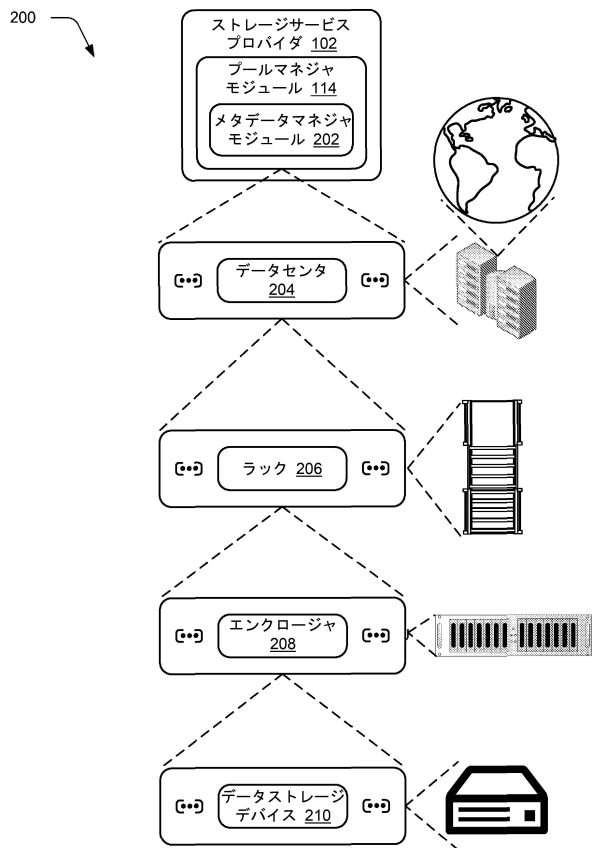
結論

[0075] 以上では、例示的な実現を、構造上の特徴および / または方法論的な動作に特有の言葉を用いて説明してきたが、添付の特許請求の範囲において定義される実現は、説明された特定の特徵または動作には必ずしも限定されない、ということが理解されるべきである。むしろ、特定の特徵および動作は、特許請求されている特徴を実現する例示的な形態として、開示されている。

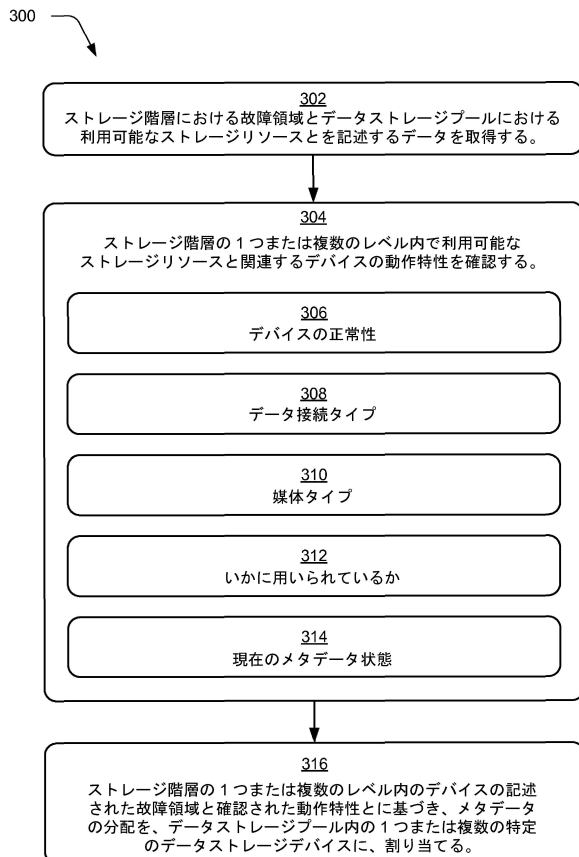
【図 1】



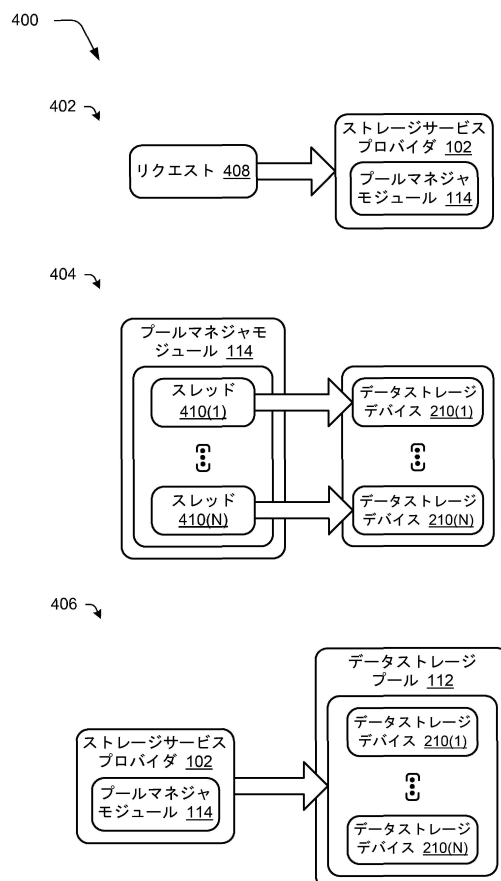
【図 2】



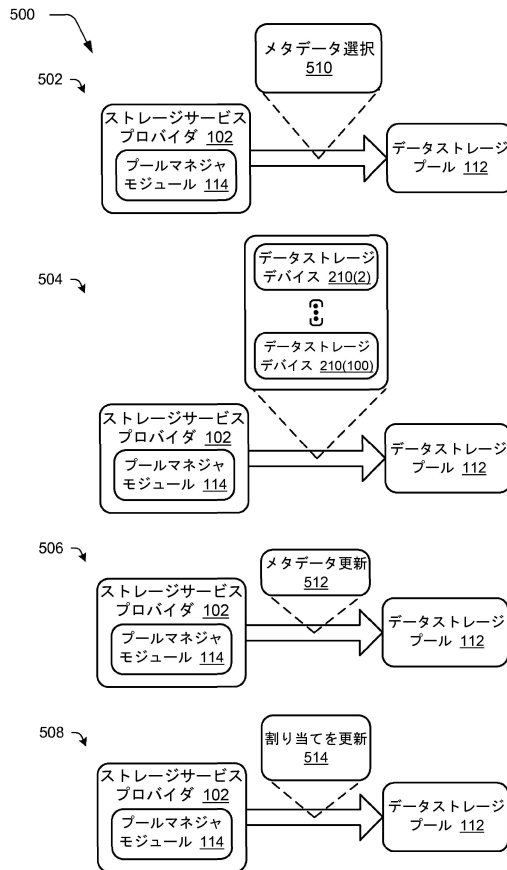
【図 3】



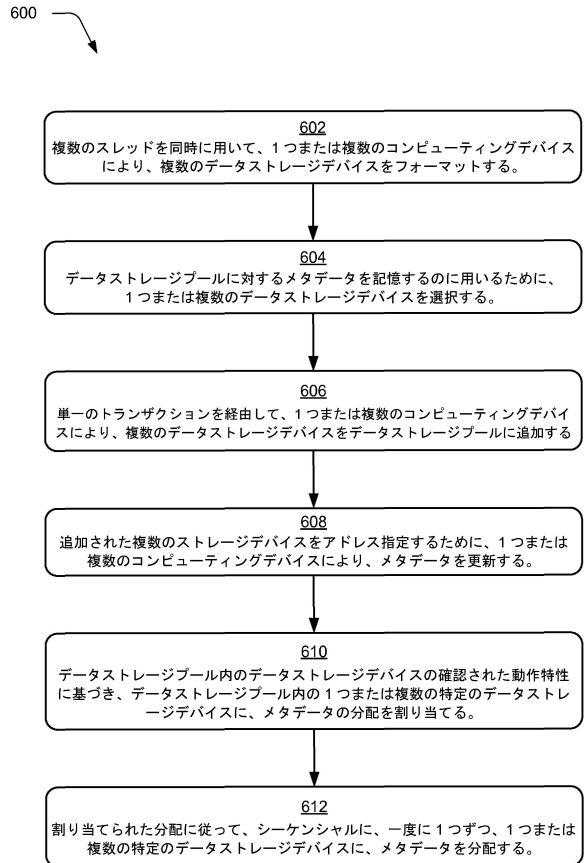
【図 4】



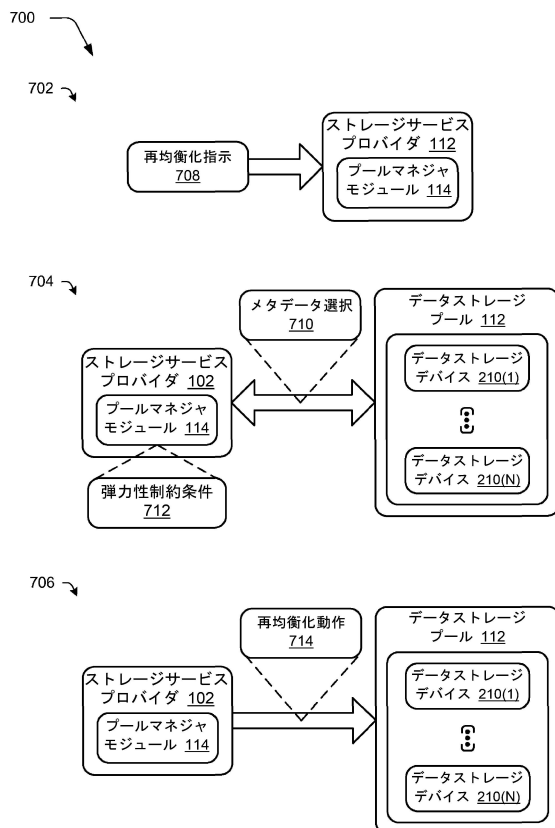
【図 5】



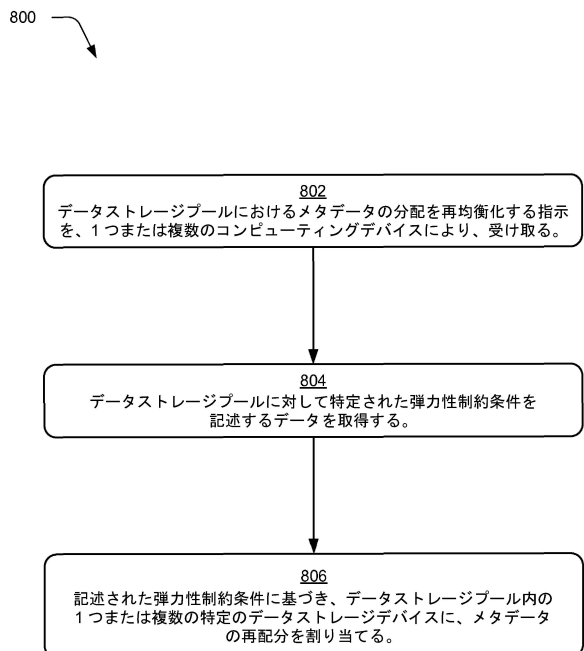
【図 6】



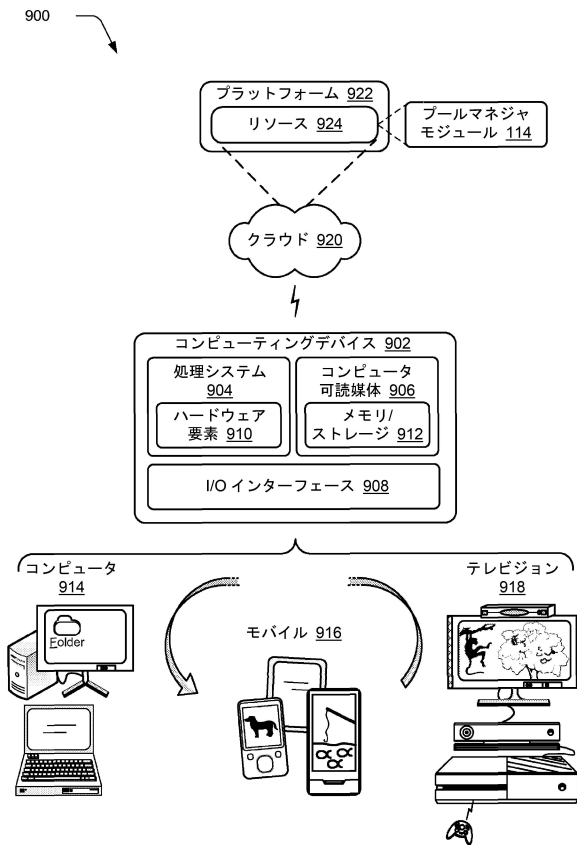
【図 7】



【図 8】



【図 9】



フロントページの続き

(74)代理人 100108213

弁理士 阿部 豊隆

(74)代理人 100162950

弁理士 久下 範子

(72)発明者 ベンチュラ, アシュリー ピー.

アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ
: パテント グループ ドケティング (ビルディング 8 / 1000)

(72)発明者 ラマニ, タルン

アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ
: パテント グループ ドケティング (ビルディング 8 / 1000)

(72)発明者 メーラ, カラン

アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ
: パテント グループ ドケティング (ビルディング 8 / 1000)

(72)発明者 ラジバル, シブ クマー

アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ
: パテント グループ ドケティング (ビルディング 8 / 1000)

(72)発明者 ハベワラ, サロシュ シー.

アメリカ合衆国, ワシントン州 98052-6399, レッドモンド, ワン マイクロソフト
ウェイ, マイクロソフト テクノロジー ライセンシング, エルエルシー, エーティーティーエヌ
: パテント グループ ドケティング (ビルディング 8 / 1000)

審査官 後藤 彰

(56)参考文献 特表2008-518284(JP, A)

特表2009-529295(JP, A)

特開2009-163664(JP, A)

米国特許出願公開第2006/0106898(US, A1)

米国特許第08495111(US, B1)

米国特許出願公開第2008/0005191(US, A1)

米国特許出願公開第2014/0181039(US, A1)

(58)調査した分野(Int.Cl., DB名)

G06F 16/182

G06F 11/16