

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号  
特許第5932041号  
(P5932041)

(45) 発行日 平成28年6月8日 (2016.6.8)

(24) 登録日 平成28年5月13日 (2016.5.13)

(51) Int.Cl.

G 0 6 F 1 2 / 0 0 ( 2 0 0 6 . 0 1 )

F I

G 0 6 F 1 2 / 0 0 5 0 1 A

請求項の数 11 (全 24 頁)

(21) 出願番号	特願2014-529675 (P2014-529675)	(73) 特許権者	314015767
(86) (22) 出願日	平成23年10月10日 (2011.10.10)		マイクロソフト テクノロジー ライセン
(65) 公表番号	特表2014-526729 (P2014-526729A)		シング, エルエルシー
(43) 公表日	平成26年10月6日 (2014.10.6)		アメリカ合衆国 ワシントン州 9805
(86) 国際出願番号	PCT/US2011/055581		2 レッドモンド ワン マイクロソフト
(87) 国際公開番号	W02013/036254		ウェイ
(87) 国際公開日	平成25年3月14日 (2013.3.14)	(74) 代理人	100107766
審査請求日	平成26年9月10日 (2014.9.10)		弁理士 伊東 忠重
(31) 優先権主張番号	13/229,734	(74) 代理人	100070150
(32) 優先日	平成23年9月11日 (2011.9.11)		弁理士 伊東 忠彦
(33) 優先権主張国	米国 (US)	(74) 代理人	100091214
			弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 プールされたパーティション・レイアウトおよび表現

(57) 【特許請求の範囲】

【請求項 1】

プロセッサを有するコンピュータ上でアクセス可能な少なくとも二つの記憶装置を含む記憶セットを編成する方法であって、当該方法は：

少なくとも二つのプールされた記憶装置の記憶領域内で、プール構成設定を有するプールされたパーティションを生成する段階であって、前記プール構成設定は：

該プールされたパーティションを有する前記プールされた記憶装置；および  
前記プールされたパーティション内に表現された、論理的な記憶容量を提供する少なくとも一つのスペース

を同定する、段階と；

前記スペースの前記論理的な記憶容量を提供するための物理的な記憶容量をリザーブする要求を受領したときに：

前記少なくとも二つのプールされた記憶装置の、他のスペースのために割り当てられていない物理的な記憶領域内で、前記プールされたパーティション内に物理的な記憶容量を提供する広がりを確保し、

該広がり前記プール構成設定において前記スペースと関連付けて、前記スペースの前記論理的な記憶容量を提供するための前記物理的な記憶容量の一部をリザーブする段階とを含む、  
方法。

【請求項 2】

それぞれの記憶装置がパーティション・テーブルを記憶しており、前記パーティション・テーブルは、その記憶装置の記憶領域内のそれぞれのパーティションについて、そのパーティションのパーティション型識別子を指定し；

前記プールされたパーティションは、前記パーティション・テーブル内で、プールされたパーティション型識別子を指定される、  
請求項 1 記載の方法。

【請求項 3】

前記プール構成設定が、広がり、スペースの論理的領域および記憶装置の物理的領域として指定するよう構成されている、請求項 1 記載の方法。

【請求項 4】

請求項 3 記載の方法であって、前記スペースに容量をリザーブする要求を受領したときに：

前記プールされたパーティション内に広がり確保し；

前記スペースに前記広がりをバインドし；

前記広がりおよび前記スペースを特定するよう前記プール構成設定を更新する段階を含む、  
方法。

【請求項 5】

請求項 1 記載の方法であって、

前記プール構成設定は、それぞれのスペースについて、スペース型識別子を指定し；

少なくとも一つのスペースがジャーナルを含み、前記プール構成設定においてジャーナル・スペース型識別子を指定される、  
方法。

【請求項 6】

請求項 1 記載の方法であって、前記スペースが、前記プールされたパーティションのそれぞれのスペースについて、そのスペースの状態を特定するメンテナンス・スペースを有する、方法。

【請求項 7】

請求項 6 記載の方法であって、

前記プール構成設定は、それぞれのスペースについて、スペース型識別子を指定し；

前記メンテナンス・スペースが、前記プール構成設定においてメンテナンス・スペース型識別子を指定される、  
方法。

【請求項 8】

請求項 1 記載の方法であって、

前記記憶セットの少なくとも一つの記憶装置が少なくとも二つのコンピュータによってアクセス可能であり、

当該方法は、前記コンピュータの間で前記プール構成設定への排他的な書き込みアクセスを有するプール構成設定所有者を、前記コンピュータの間で特定する段階を含む、  
方法。

【請求項 9】

当該方法が、前記プール構成設定を更新する要求を受領したとき：

前記コンピュータが前記プール構成設定所有者であるかどうかを判定し；

前記コンピュータが前記プール構成設定所有者でないと判定したとき、前記要求を前記プール構成設定所有者に転送する段階を含む、  
請求項 8 記載の方法。

【請求項 10】

プロセッサを有するコンピュータ上でアクセス可能な少なくとも一つの記憶装置を含む記憶セットを編成するシステムであって、当該システムは：

少なくとも一つの記憶装置の少なくとも二つのパーティションの間に分散されたデータ

10

20

30

40

50

を有するプールされたパーティションを生成するパーティション・マネージャを有し、前記二つのパーティションは、少なくとも二つの記憶装置の物理的な記憶領域の全部よりも少ない部分を占め、前記プールされたパーティションは：

前記記憶セットの第一の記憶装置の物理的な記憶領域の第一の部分範囲内の第一の広がりおよび前記記憶セットの第二の記憶装置の物理的な記憶領域の第二の部分範囲内の第二の広がりを含む、前記少なくとも二つのパーティション内に位置される複数の広がりと；

前記プールされたパーティションを有する前記少なくとも二つの記憶装置を同定する、前記プールされたパーティション内に記憶された構成設定データとを有する、システム。

【請求項 1 1】

10

物理的な記憶容量を少なくとも二つのスペースについての論理的な記憶容量に割り当てるよう構成された記憶装置であって、当該記憶装置は：

物理的な記憶領域のある範囲を含むパーティションと；

前記パーティション内に記憶されるパーティション構成設定とを有しており、前記パーティション構成設定は：

前記パーティション内の物理的な記憶領域の第一の部分範囲を含み、前記パーティションの第一のスペースの論理的な記憶容量のために確保された第一の広がりと；

前記第一の部分範囲と重ならない、前記パーティション内の物理的な記憶領域の第二の部分範囲を含み、前記パーティションの第二のスペースの論理的な記憶容量のために確保された物理的な記憶容量を提供する第二の広がりとを同定する、  
記憶装置。

20

【発明の詳細な説明】

【技術分野】

【0001】

本願はプールされたパーティション・レイアウトおよび表現に関する。

【背景技術】

【0002】

コンピューティングの分野において、多くのシナリオは、さまざまな仕方で割り当てられうる、一つまたは複数の記憶装置（たとえば円盤ベースの磁気および/または光ハードディスク・ドライブ、半導体記憶デバイスおよび不揮発性メモリ回路）のセットに関わる。第一の例として、記憶装置によって提供される記憶スペースは一つまたは複数のパーティションにグループ化されてもよく、各パーティションは動作環境において反映される一つまたは複数の論理ドライブに関連付けられたデータを記憶してもよい。第二の例として、二つ以上の記憶装置の容量が多くの仕方で組み合わせられて、追加的な記憶機能、たとえば改善されたスループット、自動ミラーリングおよび自動パリティ計算といった特徴を提供するさまざまな安価なディスクの冗長アレイ（RAID: Redundant Array of Inexpensive Disks）方式を提供してもよい。第三の例として、記憶スペースは一つまたは複数のコンピュータにとってアクセス可能でありうる（たとえばネットワーク上のいくつかのコンピュータにとって並行してアクセス可能なネットワークアタッチトストレージ装置）。そのような装置によって、ジャーナリングされた（ journaled ）スペースおよびシン・プロビジョニング（ thin provisioning ）といった、より複雑な技法が提供されてもよい。

30

40

【0003】

そのようなシナリオに関わる設計上の選択は、記憶セット（ storage set ）の記憶装置によって提供される数、型、容量、割り当ておよび機能ならびに一つまたは複数の記憶装置のプーリング構成を示すメタデータを与える。このメタデータは多くの仕方で、たとえば記憶コントローラ（たとえば記憶装置に適用されるRAID方式についての情報を記憶するよう構成されたRAIDコントローラ）のメモリ中に、あるいは記憶セットの一つまたは複数の記憶ディスク上の標準化された位置に、記憶されうる。たとえば、記憶装置は、記憶スペースの先頭に、当該記憶装置上に記憶されるパーティションの数、位置、サイズおよび

50

型をリストするマスター・ブート・レコード (MBR: master boot record) を、記憶スペースの末尾に、パーティションによって表出される論理ドライブを示す論理ディスク・モデル (LDM: logical disk model) データベースを有していてもよい。

【発明の概要】

【発明が解決しようとする課題】

【0004】

この概要は、詳細な説明においてさらに後述される概念のセレクションを簡略化された形で紹介するために設けられている。この概要は、特許請求される主題の鍵となる因子または本質的な特徴を特定することは意図されていないし、特許請求される主題の範囲を限定するために使われることも意図されていない。

【0005】

記憶セットを記述する持続するメタデータの態様はさまざまな利点および/または欠点を呈示しうる。第一の例として、記憶コントローラのメモリに記憶されるメタデータは迅速にアクセス可能でありうるが、記憶装置の異なる記憶コントローラまたはコンピュータへの移行を複雑にすることがある(たとえば、記憶装置が異なる機械またはアレイに再配置される場合、その記憶装置上のデータは、パーティション情報が同じ記憶装置上に記憶されていなければアクセス不能になってしまうことがある)。第二の例として、アレイについてのメタデータを一つの記憶装置上にのみ記憶してその記憶装置がその後障害を起こすと、記憶セットの記憶装置のいくつかまたは全部の上のデータが失われることがありうる。第三の例として、記憶装置を二つ以上のコンピュータで並行して共有し、各コンピュータがメタデータを更新しうると、競合状況(race conditions)および/または当該メタデータのミラーリングされたバージョン間でのバージョン間衝突が生じることがある。第四の例では、新たな態様で記憶装置上にメタデータを記憶すると、その記憶装置の他の記憶装置および/またはコンピュータとの互換性および/またはアクセシビリティが低下することがありうる(たとえば、メタデータがしかるべく構成設定されていない装置によって読み出し可能および/または使用可能でなくなることがありうる)。第五の例として、いくつかの表現は、記憶装置が、記憶装置の種々のセットを横断して分散されている複数の記憶プール(たとえば、第一のセットの記憶装置で共有されている第一のプールされたパーティションおよび第二の異なるセットの記憶装置で共有されている第二のプールされたパーティション)に参加することや、あるいは同じ記憶プールに種々の役割において参加すること(たとえば、単一の記憶装置が、ユーザー・データのための第一のパーティションおよび該ユーザー・データについてのパリティ・データのための同じ記憶装置上に割り当てられた第二のパーティションを含むプールされたパーティションを備えること)を許容しないことがありうる。

【課題を解決するための手段】

【0006】

本稿では、記憶セット中の諸記憶装置の、記憶装置および記憶スペースならびにプロビジョンおよびプーリング構成を同定するメタデータを記憶するための技法が呈示される。これらの技法によれば、一つまたは複数の記憶装置が、プール構成設定(pool configuration)および一組の広がり(extent)を有するプールされたパーティションを共有しうる。プール構成設定は、プールされたパーティションを、(プールを同定する)プール・レコード; プールされたパーティションによって表明される(manifested)さまざまなスペース(たとえば、ユーザー・データのためのボリューム、メンテナンス・スペースおよび該ユーザー・データについてのジャーナルを記憶するジャーナル・スペース)を表わす一組のスペース・レコード; プールに参加している記憶装置を表わす一組の記憶装置レコード; および記憶装置上のプールされたパーティション内の物理的な位置の割り当てを、プールされたパーティションのスペース内の論理的な位置にマッピングする一組の広がりレコードとして、表現してもよい。どの記憶装置上でもプールされたパーティションの一貫したメタデータ表現へのアクセスを提供するよう、プール構成設定は、各記憶装置のプールされたパーティションのいちばん上に、ミラーされた仕方でも記憶されてもよい。プー

ルされたパーティション内のスペースを表明する (manifest) 要求は、そのスペースを表わす、可能性としてはそのスペースのプロビジョンされた容量を含むスペース・レコードを生成することによって満たされてもよい。そのスペースに物理的な容量を割り当てるため、広がりも、そのスペースに割り当てられ、そのスペースにバインドされてもよい。このバインドは、そのスペースの生成時に即座に実行されてもよく、あるいはそのスペースの容量が利用されるまで遅延されてもよい。追加的に、プールされたパーティションにアクセスする諸コンピュータの間で、プール構成設定を更新することを排他的に許容されるプール構成設定所有者が同定されてもよく、プール構成設定を変更する (たとえばスペースを追加するまたはスペースに広がり割り当てる) いかなる要求もそのプール構成設定所有者に転送されてもよい。

10

#### 【0007】

プールされたパーティションのこの表現は、プールされたパーティションの他の表現に対していくつかの利点を可能にしうる。第一の例として、記憶装置のプールされたパーティション内のデータは、記憶装置が (たとえば異なる機械または異なる記憶アレイに) 再配置される場合に、あるいはプールされたパーティションを共有している別の記憶装置がクラッシュするまたは利用不能になる場合に、アクセス可能のままでありうる。プール構成設定がプールされたパーティションを共有する各記憶装置上にミラーされているからである。第二の例として、この表現は、スペースのプロビジョニングを広がり割り当てから区別する。この後者は遅延された仕方、たとえばジャストインタイム式に実行されてもよい。この区別は、記憶セットのさまざまな型の柔軟性、たとえばスペースをサイズ変更するために広がり割り当て直す簡単な仕方、記憶装置上の容量の割り当てを含まない迅速なプロビジョニング技法およびシン・プロビジョニング (thin provisioning) 機能 (諸記憶装置の利用可能な物理的容量を超える容量をもつスペースをプロビジョニングしておいて、そのスペースの容量が尽きる際に追加的な記憶装置の形で物理的な記憶容量を追加することによってプロビジョニングされた容量を充足するオプションがある) を可能にしうる。第三の例として、このようにして表現されるプールされたパーティションは同じ記憶装置上で二つ以上のパーティションにまたがることがありうる。第四の例として、第一のプールされたパーティションがプールされていない諸パーティション (これはプールされたパーティションを認識しない記憶システム上でアクセス可能であってもよい) と共存しうる。第五の例として、記憶装置は、それぞれ異なるセットの他の記憶装置と共有される二つ以上のプールされたパーティションを記憶しうる。第六の例として、記憶装置にアクセスしうるコンピュータまたは装置の間でのプール構成設定所有者の選択は、二つ以上の装置がプール構成設定を並行して更新する場合に普通なら起こりうるプール構成設定の非同期につながりうる競合状態を軽減しうる。これらおよび他の利点は、本稿に呈示される技法に基づくプールされるパーティションの表現を通じて達成可能であってもよい。

20

30

#### 【0008】

上記および関係する目的の達成に向け、以下の記述および付属の図面はある種の例示的な側面および実装を記述する。これらは、一つまたは複数の側面が用いられうるさまざまな仕方のほんのいくつかを示す。本開示の他の側面、利点および新規な特徴は、付属の図面との関連で考慮されるとき、以下の詳細な記述から明白となるであろう。

40

#### 【図面の簡単な説明】

#### 【0009】

【図1】記憶装置の容量の一組のパーティションへの割り当ておよび該記憶装置上のそれらのパーティションの構成設定の表現を表わす例示的なシナリオの図解である。

【図2】安価なディスクの冗長アレイ (RAID) レベル5記憶方式に従っていくつかの記憶装置にまたがるボリュームへの記憶セットの容量の割り当てを表わす例示的なシナリオの図解である。

【図3】二つの記憶装置の容量の、それらの記憶装置にまたがる動的パーティションへの割り当ておよびそれらの記憶装置上の該動的ボリュームの表現を表わす例示的なシナリオ

50

の図解である。

【図４】本稿に提示される技法に基づく、プールされたパーティションへの割り当てであって、前記プールされたパーティションを共有する諸記憶装置上に割り当てられる一組の広がりによってバインドされた一組のスペースを含む割り当てと、それらの記憶装置上の前記プールされたパーティションの表現とを表わす例示的なシナリオの図解である。

【図５】本稿に提示される技法に基づく、それぞれ異なる他の記憶装置と共有される二つのプールを有する記憶装置と、それらの記憶装置上の前記プールされたパーティションの表現とを表わす例示的なシナリオの図解である。

【図６】本稿に提示される技法に基づく、少なくとも二つの記憶装置を含む記憶セットを配分する（apportioning）例示的な方法を示すフローチャートの図解である。

10

【図７】本稿に記載される仕組みの一つまたは複数を具現するように構成されたプロセッサ実行可能命令を有する例示的なコンピュータ可読媒体の図解である。

【図８】記憶装置のプールされたパーティション内のスペースをプロビジョニングし、該スペースに広がり割り当てのための第一の例示的な技法の図解である。

【図９】記憶装置のプールされたパーティション内のスペースをプロビジョニングし、該スペースに広がり割り当てのための第二の例示的な技法の図解である。

【図１０】記憶装置のプールされたパーティション内のスペースをプロビジョニングし、該スペースに広がり割り当てのための第三の例示的な技法の図解である。

【図１１】本稿に記載される仕組みの一つまたは複数が実装されうる例示的なコンピューティング環境を示す図である。

20

【発明を実施するための形態】

【００１０】

特許請求される主題についてここで図面を参照して述べる。図面を通じて、同様の要素を指すために同様の参照符号が使われる。以下の記述では、説明の目的で、特許請求される主題の十全な理解を与えるために、数多くの個別的詳細が記述される。しかしながら、特許請求される主題がそうした個別的詳細なしでも実施されうことは明白であろう。他の場合には、特許請求される主題の記述を容易にするために、構造および装置がブロック図の形で示される。

【００１１】

#### A．序

30

コンピューティングの分野において、多くのシナリオは、一つまたは複数のコンピュータまたは他の装置に代わってデータを記憶するよう構成された記憶装置（たとえばハードディスク・ドライブ、半導体記憶デバイスまたは揮発性もしくは不揮発性メモリ回路）に関わる。記憶されるデータは多くの型のオブジェクト（たとえば、ファイル、メディア・オブジェクト、メディア・オブジェクト、データ・セットまたはデータベース・レコード）を含んでいてもよく、さまざまな仕方で（たとえば名前を付けられたオブジェクトの集合として、階層構造として、またはリレーショナルな仕方で）編成されてもよい。しかしながら、より低いレベルでは、記憶装置はしばしば該装置によって使用可能なあるレイアウトに従って構成設定される。たとえば、記憶されるべきデータの型および編成には関わりなく、記憶装置はしばしば、コンピュータおよび他の装置が読むことができる特定の仕方で編成される（たとえばオーガナイズされる）。一つのそのような例として、記憶装置の利用可能な容量はしばしば一組のパーティションとして編成され、記憶装置の先頭において生成されるパーティション・テーブルがそれらのパーティションの位置、サイズおよび型を示す。記憶装置の容量を諸ボリュームに分離することは、さまざまな利点を提供しうる（たとえば、異なるパーティションが異なるグループのデータを隔離しうる；異なるボリュームを表明するために使用されうる；あるいは異なる型のデータを記憶するようおよび／または異なる型のパーティションを利用する異なるコンピュータまたは装置上でのアクセスのためにデータを記憶するよう異なる編成にされうる）。

40

【００１２】

図１は、一連のパーティション１０４として編成されたデータ記憶容量をもつ記憶装置

50

102を表わす例示的なシナリオ100の図解である。この例示的なシナリオ100では、それぞれのパーティション104は、さまざまな型のユーザー・データ114を記憶するように構成されており、コンピュータまたは装置上で表明されうる一つまたは複数の論理ボリューム106を含んでおり、ファイル・システムをもってさらにフォーマットされ、一組のファイルを記憶するために使われる。たとえば、記憶装置102は、基本的な仕方

で編成され、一つの論理ボリューム106を記憶することができる第一のパーティション104を有していてもよく、該論理ボリュームは、該論理ボリューム106のユーザー・データ112を含むファイルのカタログを表わすマスター・ファイル・テーブル108を有する。記憶装置102はまた、拡張された仕方

で編成され、よって二つ以上の論理ボリューム106を記憶することができる第二のパーティション106を有していてもよい。たとえば、第二のパーティション106は、それぞれが(拡張された第二のパーティション106内の)第三のパーティション106に含まれる二つの論理ボリューム106を含んでいてもよい。第三のパーティションは論理ボリューム106を含む。トップ・レベルのパーティション104についての情報は、記憶装置102の利用可能な容量の先頭に記憶されており、各パーティション104についてのメタデータを含むパーティション・テーブル110に含まれていてもよい。該メタデータは、パーティション104が始まる物理位置、パーティション104がコンピュータがオペレーティング環境にブートすることができるようにする情報を含んでいるか否かおよび/またはパーティション104の型を識別するパーティション型指標(たとえば、第一のパーティション104については基本パーティション型、第二のパーティション104については拡張パーティション型)を含んでいる。拡張された第二のパーティション104は、各含まれているパーティション108に、該含まれているパーティション108についてのメタデータを含むマスター・ファイル・テーブル(MFT: master file table)を先行させていてもよい。このようにして、記憶装置102の容量は、それぞれ一つまたは複数の論理ボリューム106を含む一組のパーティション104として編成されることができ、よって記憶装置102の記憶容量をいくつかの論理ボリューム106に区画化して表出することができる。

#### 【0013】

図1の例示的なシナリオ100におけるパーティション104の例示的なフォーマットはいくつかのシナリオを満足させるが、いくつかの追加的な特徴を可能にするような仕方

でフォーマットを設計することが望ましいことがありうる。第一の例として、二つ以上のパーティション104の間の、ミラーされた関係(各パーティション104に同一のデータ・セットが記憶され、一つのパーティション104内のデータへの変更が自動的に他のパーティション104に伝搬させられる)のような関係を確立することが望ましいことがありうる。そのようなミラーリングは、データ破損の場合のデータ・セットの耐久性を増進しうる。たとえば、パーティション104の破損またはパーティション104を含む記憶装置102の障害の場合、データ・セットは、(可能性としては異なる記憶装置102上の)第二のパーティション104にミラーされていれば、無傷かつアクセス可能なままでありうる。ミラーリングは、パフォーマンスをも改善しうる。たとえば、第一の記憶装置102は一組のパフォーマンス特性に限定されてもよく(たとえば、アクセス要求を開始する際に関わる最小のレイテンシーまたはデータにアクセスする際の最大の持続可能なスループット)、二つ以上の記憶装置102に横断してデータ・セットをミラーすることにより、すべての記憶装置102の総和を表わすパフォーマンス特性を可能にしうる。記憶装置102のいくつかのフォーマットによって可能にされうる第二の特徴は、データ・セットを二つ以上の記憶装置102のパーティション104にわたって横断的にまたがらせることである。たとえば、図1の例示的なシナリオ100では、論理ボリューム106のサイズは論理パーティション106を有するパーティション104のサイズに限定され、該パーティション104のサイズは記憶装置102の利用可能な容量に限定される。しかしながら、他のフォーマットは二つ以上の記憶装置102の二つ以上のパーティションをプールし、それによりいずれか一つのパーティション104または記憶装置102のサイズ制限を超えることを可能にすることがある。記憶装置102のフォーマット形式は

10

20

30

40

50

、誤り検出および誤り訂正（たとえば、それぞれのデータ・セットについて、記憶装置 102 は自動的にチェックサムのような検証子のセットを生成し、記憶しうる。これが、可能性としては誤って書き込まれたデータを検出するためにデータ・セットの内容と比較されてもよい）およびデータ・セットへの共有されるアクセス（たとえば、記憶セット 102 は、データ・セットを、データ・セットへの一貫しないまたは確定しない更新につながりうる書き込みベースの競合条件にさらすことなく複数のコンピュータまたは他の装置を可能にしてもよい）を増進するために工夫されてもよい。

#### 【0014】

よって、図 1 の例示的なシナリオ 100 において示される基本フォーマットは、ミラーリング、並行アクセス、障害耐性および誤り検出および訂正といった機能を可能にするために拡張されてもよい。たとえば、一組の記憶装置 102 が、安価なディスクの冗長アレイ（RAID）方式に従って編成されてもよい。RIAD 方式はそのような機能を実装するためにハードウェアおよび / またはソフトウェアで実装されてもよい。第一の例として、RAID 0 方式では、いくつかの記憶装置 102 のいくつかのパーティション 104 に記憶されているデータが単一の論理ボリューム 104 として呈示される。第二の例として、RAID 1 方式では、記憶されるデータは一つまたは複数の記憶装置 102 上のいくつかのパーティション 104 を横断して自動的にミラーリングされ、それによりデータ・セットへのより高速のアクセスおよび / または記憶装置 102 またはパーティション 104 が信頼できなくなったまたはアクセスできなくなった場合の障害耐性を可能にする。第三の例として、RAID 4 方式では、データ・セットのそれぞれの部分についてチェックサムが計算され、該チェックサムがデータ・セットの完全性（たとえば、データ・セットに書き込まれたデータが該データ・セットから整合する形で読まれ、記憶の間に意図しない変更を受けていないこと）を検証するために、および / またはパーティション 104 および / または記憶装置 102 の障害から回復するために（たとえば、記憶装置 102 のパーティション 104 が破損した場合、あるいは記憶装置 102 が除去された場合、パーティション 104 内に含まれるデータが、欠けているデータのチェックサムおよびチェックサムによって表わされるデータ・セットの他の部分を使って回復可能でありうる）、使用されてもよい。それにより、RAID 方式や他のフォーマット仕様は、記憶装置 102 上での特定の編成およびデータ・アクセスを通じてこれらおよびその他の機能を可能にする。

#### 【0015】

図 2 は、特定の記憶機能を有する容量のプール 202 を確立するために RAID 4 方式に従って四つの記憶装置 102 のセットをフォーマットすることを表わす例示的なシナリオ 200 の図解である。この例示的なシナリオ 200 では、それぞれの記憶装置 102 は、ユーザー・データ 114 をいくつかのパーティション 104 を横断して延在する論理ボリューム 106 内に記憶し、ユーザー・データ 114 の完全性を検証し、記憶装置 102 の障害の場合にデータ・セットを復元するためにパリティ・データ 204 を自動的に生成し、記憶するよう構成される。よって、各記憶装置 102 は、該記憶装置 102 の先頭に、該記憶装置 102 のパーティション 104 を記述するメタデータを含むパーティション・テーブル 110 を記憶していてもよい。各記憶装置 102 はまた、パーティション 104 を記憶する。最初の三つの記憶装置 102 のパーティション 104 は、単一の論理ボリューム 106 を呈するプール 202 として構成設定される（たとえば、マスター・ファイル・テーブル 112 およびユーザー・データ 114）。さらに、第四の記憶装置 102 の容量はプール 202 に含まれるが、パリティ・データ 204 のためにリザーブされている。パリティ・データは、最初の三つの記憶装置 102 に記憶される対応するデータ・セットについて自動的に計算される。たとえば、最初の三つの記憶装置 102 のパーティション 104 内のある特定の物理アドレスに記憶されるデータの語について、語長のチェックサムが計算され、第四の記憶装置 102 のパーティション 104 内の同じ物理アドレスに記憶されてもよい。したがって、記憶装置のこのフォーマット形式 RAID 4 レイアウトは、記憶装置 102 の容量を総合して、論理ボリューム 106 および記憶セット 102 の障害耐性を改善しうる検証子データを呈するプール 202 を生成することを可能にする。



## 【 0 0 1 6 】

しかしながら、図 2 の例示的なシナリオ 2 0 0 におけるレイアウトはいくつかの制限を呈しうる。第一の例として、RAID 4 方式は、マッチする特性をもつ記憶装置 1 0 2 をプールすることができるだけであることがある。たとえば、この方式を、さまざまな容量の記憶装置 1 0 2 (たとえば、1 テラバイト記憶装置 1 0 2 と 2 テラバイト記憶装置) で適用することは可能でないことがある。この制限は、最初の三つの記憶装置 1 0 2 上のデータ・セットの対応する物理アドレスについての、第四の記憶装置 1 0 2 上でのチェックサムの計算から生じることがありうる。たとえば、アルゴリズムは、これらの計算を、(パフォーマンスの低減なしには) 互いに異なるサイズの記憶装置 1 0 2 に適応させることができないことがある。第二の例として、RAID 4 方式は、互いに異なるパフォーマンス特性をもつ記憶装置 1 0 2 を含む記憶セットについては、低減した一貫性、信頼性および/またはパフォーマンスを示すことがある(たとえば、たとえ記憶装置 1 0 2 が同一のサイズであっても、異なる構成および/またはモデルの記憶装置 1 0 2 のレイテンシーおよびスループット特性は、問題、あるいは記憶セットの低減したパフォーマンスにつながる) ことがある。第三の例として、諸記憶装置 1 0 2 の相互依存性が、いずれか一つの記憶装置 1 0 2 のポータビリティを低減させることがありうる。たとえば、異なるアレイに移したら、記憶装置 1 0 2 の内容にアクセスすることが可能でなくなることがある。記憶装置 1 0 2 の内容を記述するメタデータが、他の記憶装置 1 0 2 および/またはそのコントローラと相互に関係していることがありうるからである(たとえば、プール 2 0 2 を記述するメタデータは、RAID コントローラの別個のメモリ・コンポーネント上に記憶されていてもよく、一つまたは複数の記憶装置 1 0 2 を異なる RAID コントローラに移すことは、メタデータを移し損なうことがあり、その結果、データ・セットにアクセスできなくなる)。第四の例として、これらの記憶装置 1 0 2 のフォーマット形式は、RAID 4 方式をサポートするよう構成されているシステム(たとえば記憶コントローラおよびソフトウェア)によってのみ使用可能であることがある。第五の例として、この方式は、記憶装置 1 0 2 すべての容量全部を利用することがある。したがって、プール 2 0 2 の外で、たとえばプール 2 0 2 の外の第二のパーティション 1 0 4 として、あるいは同じまたは他の記憶装置 1 0 2 と一緒に別のプール 2 0 2 に参加するパーティション 1 0 4 として、使うために、これらの記憶装置 1 0 2 上の何らかの容量を割り当てることが可能でないことがある。さらに、記憶装置 1 0 2 のセット上で異なる複数の RAID 方式を実装することが難しい。たとえば、RAID コントローラは、それが使える記憶装置 1 0 2 すべての容量全体に対して一つの RAID 方式を適用するよう構成されていることがある。

## 【 0 0 1 7 】

図 3 は、図 2 の例示的なシナリオ 2 0 0 の制限のいくつかを克服しうる、論理ディスク・マネージャ (LDM: Logical Disk Manager) に基づく記憶装置 1 0 2 の例示的なフォーマット形式を表わす例示的なシナリオ 3 0 0 の図解である。この例示的なシナリオ 3 0 0 では、二つの記憶装置 1 0 2 は、パーティション・テーブル 1 1 0 で始まり、連続的な物理アドレスの大きな割り当てられたブロックを表わす一組のサブディスク 3 0 4 を含む動的パーティション 3 0 2 が続くよう編成される。該連続的な物理アドレスは表明される論理ボリューム 1 0 6 に総合され、表明される論理ボリューム 1 0 6 は、両方の記憶装置 1 0 2 にまたがり、第一の記憶装置 1 0 2 の動的パーティション 3 0 2 および第二の記憶装置 1 0 2 の動的パーティション 3 0 2 の容量のプール 2 0 2 を表わす。動的パーティションは、第一の記憶装置 1 0 2 の動的パーティション 3 0 2 に記憶される第一のサブディスク 3 0 4 および第二の記憶装置 1 0 2 の動的パーティション 3 0 2 に記憶される第二のサブディスク 3 0 4 から総合される論理アドレスのシーケンスとして、論理的に表現される。異なる記憶装置 1 0 2 上の物理アドレスの別個のブロックを含んでいるが、これらのサブディスク 3 0 4 は、論理アドレスの連続的なブロックとして総合され、表明される。さらに、プール 2 0 2 についてのメタデータは、各記憶装置 1 0 2 の末尾に位置され、記憶装置 1 0 2 のサブディスク 3 0 4 の割り当てを論理ボリューム 1 0 6 として表現する論理ディスク・マネージャ (LDM) データベース 3 0 6 に記憶される。たとえば、この例

10

20

30

40

50

示的なシナリオ 300 では、論理ディスク・マネージャ・データベース 306 は、第一の記憶装置 102 が二つのサブディスク 304 を有し、第二の記憶装置 102 が一つのサブディスク 304 を有するが、第一の記憶装置 102 の第一のサブディスク 304 は第一の論理ボリューム 106 として表明される一方、第一の記憶装置 102 の第二のサブディスク 304 および第二の記憶装置 102 の唯一のサブディスク 304 が一緒に第二の論理ボリューム 106 として表明される。さらに、論理ディスク・マネージャ・データベース 306 は、両方の記憶装置 102 上でミラーされ、よって各記憶装置 102 と一緒にポータブルであり、それに含まれる動的パーティション 302 の論理的な表明を指示する。

#### 【0018】

論理ディスク・マネージャ・データベース 306 をもつ記憶装置 102 のフォーマット付けは、いくつかの利点を呈しうる。第一の例として、図 1 に示した例示的なシナリオ 100 に比べ、論理ディスク・マネージャ・データベース 306 は、複数の記憶装置 102 上のサブディスク 304 からなるプール 202 の表現を可能にする。また、図 2 の例示的なシナリオ 200 に比べ、図 3 の例示的なシナリオ 300 に示したフォーマット付けは、異なるサイズおよび/または異なるパフォーマンス特性の記憶装置 102 上で実装されてもよい（たとえば、図 3 における第一および第二の記憶装置 102 は、異なる記憶コントローラをもつ、異なる製造業者によって生成された、異なる総容量を有していてもよい）。さらに、図 2 に示した例示的なシナリオ 200 とさらに比べ、論理ディスク・マネージャ・データベース 306 は、プール 202 の、該プール 202 に含まれない、論理ディスク・マネージャ・データベース 306 を使う機能のないシステム（たとえば、記憶コントローラおよびソフトウェア）にとってアクセス可能かつ使用可能でありうる他のパーティション 104（たとえば、第一の記憶装置 102 上の第一のパーティション 104）との共存を可能にしうる。

#### 【0019】

しかしながら、サブディスク 304 のプール 202 を表わすために論理ディスク・マネージャ・データベース 306 を使うことも、いくつかの制限を呈することがある。第一の例として、多くの論理ディスク・マネージャ・データベース 306 は一つのプール 202 を表わすことができるのみであり、論理ディスク・マネージャ・データベース 306 をサブディスク 304 の複数のプール 202 を表わすよう拡張することは難しいことがある。第二の例として、記憶装置 104 の末尾にある論理ディスク・マネージャ・データベース 306 の位置は、論理ディスク・マネージャ・データベース 306 のサイズを制限することがあり、論理ディスク・マネージャ・データベース 306 を拡大することは論理ディスク・マネージャ・データベース 306 全体を各記憶装置 102 上の異なる位置に再書き込みすることに関わることがある。第三の例として、パーティション 104 の一部のみを含む論理ボリューム 106 の表明を表現することが難しいことがある。たとえば、図 1 の例示的なシナリオ 100 において同様に示されているように、各パーティション 104 は完全に一つの論理ボリューム 106 に割り当てられる。この対応のため、パーティション 104 を柔軟な仕方では割り当てる（たとえば、パーティション 104 の容量を二つ以上の論理ボリューム 106 を横断して分散させる）ことが難しいことがあり、論理ボリューム 106（それに含まれるサブディスク 304 を含む）のサイズ変更は、記憶装置 102 の諸パーティション 104 のかなりの変更を伴うことがある（たとえば、論理ボリューム 106 のサイズ変更は、パーティション 104 の再割り当てを伴うことがあり、可能性としてはパーティション 104 に記憶されるデータの完全な再書き込みを伴うことがある）。

#### 【0020】

##### B. 呈示される技法

本願で呈示されるのは、図 3 の例示的なシナリオにおける論理ディスク・マネージャ（LDM）・データベース 306 が示すものも含め他のフォーマット付け技法の制限のいくつかに対処しうる、記憶装置 102 をフォーマットする技法である。これらの技法によれば、記憶装置 102 の利用可能な容量を、さまざまな記憶機能（たとえば、複数の記憶装置 102 によって共有される記憶容量の一つまたは複数のプール 202 の生成；RAID 1A

10

20

30

40

50

レイにおけるミラーリングまたはRAID 4アレイにおけるチェックサム利用といった、さまざまな協調された記憶機能の規定および自動化された適用；複数のコンピュータまたは装置による並行アクセス）の実装を容易にする仕方で編成するフォーマットが工夫されてもよい。そのような技法は、記憶装置102上に記憶される記憶セットの堅牢さ（たとえば、異なる記憶コントローラに再配置された場合の、記憶装置102についてのメタデータの保存）；編成の柔軟さ（たとえば、パーティション104のサイズ変更を容易にする、記憶装置102の利用可能な容量の割り当ておよび再割り当てならびに当該編成の他の編成技法との両立性（たとえば、記憶装置が、他の記憶装置102と同期した記憶のプール202のようなより複雑な編成構造および基本パーティション104のような標準化された構造の両方を含むことができるようにする）を増進してもよい。

10

#### 【0021】

したがって、これらおよびその他の考察に基づき、本願で呈示される技法は、それぞれが他の記憶装置102と共有されてもよく、さまざまな記憶機能（たとえば種々のRAID方式または他の機能）を実装しうる一つまたは複数のプールされたパーティションを含むよう、記憶装置102を編成することに関わる。記憶装置102をおおまかに見ると、プールされるパーティションは、他のパーティション104と同様の仕方で、たとえばパーティション・テーブル110内でインデックス付けされた、取っておかれた利用可能な容量のブロックとして、割り当てられる。このように、プールされるパーティションは、異なる型の他のパーティション104と共存しうる。しかしながら、プールされるパーティションの容量は異なる仕方で利用される。第一の例として、プールされるパーティションを記述するメタデータは別個の構造（たとえば論理ディスク・マネージャ・データベース306）内ではなく、プールされるパーティション内に記憶される。それにより、同じ記憶装置102上に、任意の異なる型の（たとえば他の記憶装置102の異なるセットと共有される、および/または異なるRAID方式または他の記憶機能を実装する）複数のプールされるパーティションが記憶されることができる。第二の例として、プールされるパーティションは、異なる構造体を表わしうるスペースのセットを定義しうる（たとえば、ユーザー・データを記憶する論理ボリューム106；再同期を提案する陳腐性（staleness）または記憶装置102の交換を提案する障害など、それらのスペースについてのメンテナンス・メタデータ；または記憶セットへの書き込みの一時的ストアとして構成されたジャーナル）。さらに、パーティション102の容量を完全に一つの論理ボリューム106に割り当てたり、あるいは少数の大きな連続ブロックの形で少数の論理ボリューム106に割り当てたりするのではなく、プールされるパーティションの容量は、それぞれのスペースにマッピングされうる小さなブロック（本稿では「広がり（extent）」と称する）において割り当てられてもよい。これらの関連付けは、プールされるパーティションの構成設定において記憶されてもよい。さらに、前記関連付けは、（たとえば広がり割り当てたり、あるいは割り当て解除したりすることによる）スペースのサイズ変更や、プロビジョニングの遅延および/またはそれらのスペースについての広がりバインドといった他の機能を容易にしてもよく、それによりシン・プロビジョニングのような機能を可能にしてもよい。これらおよびその他の特徴が、本願で呈示される技法に基づく記憶装置102のフォーマット付けおよび編成を通じて達成可能でありうる。

20

30

40

#### 【0022】

図4は、これらおよびその他の機能を可能にしうる記憶装置102の例示的な表現を表わす例示的なシナリオ400の図解である。対照のため、図3の例示的なシナリオ300では、論理ディスク・マネージャ・データベース306は記憶装置102を、一組の論理ボリューム106および一組の記憶装置102としてプロビジョニングすることを表わしており、一組の記憶装置102はそれぞれの動的パーティション302に記憶される一組のサブディスク304によって関連付けられる。この例示的なシナリオ400では、プロビジョニングはプール404として（たとえば、プール404の名前のようなプール404についてのメタデータを記憶し、プール404の所有者を同定するプール・レコードによって表現される）、およびプール404から表明されるスペース406のセットとして

50

(たとえば、スペース 4 0 6 の名前、スペース 4 0 6 の型、スペース 4 0 6 のプロビジョンされた容量およびRAID方式のようなスペース 4 0 6 によって実装される記憶機能のようなそれぞれのスペース 4 0 6 についてのメタデータを記憶するスペース・レコードによって表現される) 表わされる。多くの型のスペース 4 0 6 がプール 4 0 4 によって表明されてもよく、ユーザー・データのための容量を提供する論理ボリューム 1 0 6 ; それぞれのスペース 4 0 6 の健全性(たとえば陳腐性、障害または非可用性)を表わす健全性指標 4 1 2 のような、プール 4 0 4 の他のスペース 4 0 6 についてのメタデータを記憶するメンテナンス・スペース 4 1 0 ; RAID 4 方式でのような他のスペース 4 0 6 についてのチェックサムを記憶するチェックサム・スペース; および(たとえばパッチ書き込みを増進するおよび/またはRAID書き込みホールを軽減するために)別のスペース 4 0 6 に書き込まれるべきデータ・セットが一時的に記憶されうるジャーナルに構成設定されたジャーナル・スペースを含む。物理的には、プロビジョニングは一組の記憶装置 1 0 2 として(たとえば、名前および総容量を含め、記憶装置 1 0 2 についてのメタデータを記憶する記憶装置レコードによって表わされる)、およびそれぞれのスペース 4 0 6 に関連付けられた記憶装置 1 0 2 の割り当てられた部分を表わす一組の広がり 4 0 8 として(たとえば、記憶装置 1 0 2 のある範囲の物理アドレスをスペース 4 0 6 のある範囲の論理アドレスにマッピングする広がりレコードによって表わされる)、表現される。このように、プール構成設定 4 0 2 は、プール 4 0 4 内で表明されるスペース 4 0 6 に割り当てられる広がり 4 0 8 の関連付けを通じて、記憶装置 1 0 2 の物理的な容量のプール 4 0 4 へのプロビジョニングを表わす。

#### 【 0 0 2 3 】

図 5 は、図 4 の例示的なシナリオ 4 0 0 において示されている表現を使って記憶装置 1 0 2 の編成を示す例示的なシナリオ 5 0 0 の図解である。この例示的なシナリオ 5 0 0 では、第一の記憶装置 1 0 2 は二つのプール 4 0 4 を記憶するよう構成されていてもよい。第一のプール 4 0 4 は第二の記憶装置 1 0 2 と共有され、第二のプール 4 0 4 は第三の記憶装置 1 0 2 と共有される。第一の記憶装置 1 0 2 はパーティション・テーブル 1 1 0 で始まり、第一のプールされるパーティション 4 0 4 が続く。第一のプールされるパーティション 4 0 4 は、第二の記憶装置 1 0 2 上に記憶される第二のプールされるパーティション 4 0 4 との協働において表明されるスペースを含む。両方のプールされるパーティション 4 0 4 は、図 4 に示した表現の編成に従って記憶装置のプロビジョニングを記述するレコードを含むプール構成設定 4 0 2 で始まる。プール構成設定 4 0 2 は、プール 1 0 4 から表明される一組のスペース 4 0 6 を同定し、それぞれスペース 4 0 6 の一組の論理アドレスにマッピングされる第一および第二の記憶装置 1 0 2 のプールされるパーティション 4 0 4 内において割り当てられる(つまり確保される)広がり 4 0 8 (たとえば、物理的な割り当て)を指示する。たとえば、第一の記憶装置 1 0 2 の第一のプールされるパーティション 4 0 4 は、メンテナンス・スペース(たとえば図 4 に示される例示的なメンテナンス・スペース 4 1 0)としてそれぞれ利用される四つの広がり 4 0 8 と; ユーザー・データを記憶する二つの広がり 4 0 8 と; プールされるパーティション 4 0 4 の一つまたは複数の他の広がり 4 0 8 および/またはスペース 4 0 6 についてのジャーナルを表わす広がり 4 0 8 とを記憶する。割り当ては、スペースの論理的な容量についての物理的な記憶スペース(たとえば広がり 4 0 8)のプロビジョニングにおけるいくつかの柔軟性を提供しうる(たとえば、第二のスペース 4 0 6 は第一の記憶装置 1 0 2 の第二の広がり 4 0 8 で始まり、第一の記憶装置 1 0 2 の第三の広がり 4 0 6 を通じて続き、論理的に第二の記憶装置 1 0 2 上の第二の広がり 4 0 8 がそれに続く)。プール構成設定 4 0 2 は、スペース 4 0 8 の間のさまざまな型の関係、たとえば第一の記憶装置 1 0 2 のメンテナンス・スペースおよび第二の記憶装置 1 0 2 の第一のプールされるパーティション 4 0 4 内の対応するメンテナンス・スペースのミラーリング 5 0 4 をも表わしうる。さらに、第二の記憶装置 1 0 2 と共有される第一の記憶装置 1 0 2 上の第一のプールされるパーティション 4 0 4 は、第一の記憶装置 1 0 2 上に記憶され、第三の記憶装置 1 0 2 と共有される第二のプールされるパーティション 4 0 4 および基本的なパーティション 1 0 4 と共存してもよ

い。

#### 【 0 0 2 4 】

本願で呈示される技法に基づく図 4 ~ 図 5 の例示的なシナリオ 4 0 0、5 0 0 における記憶装置 1 0 2 の編成は、他の編成 ( 図 1 ~ 図 3 で呈示したものを含む ) に対していくつかの利点を可能にしうる。第一の例として、プール構成設定 4 0 2 は、異なる型のスペース 4 0 6 ( たとえば、メンテナンス・スペース、ユーザー・データ・スペースおよびジャーナル・スペース ) の指定ならびに各スペース 4 0 6 に適用される記憶機能の指定 ( これは、異なるスペース 4 0 6 について異なる RAID 方式など異なる記憶機能を使うことを可能にする ) を可能にする。第二の例として、プール構成設定 4 0 2 は、プール 4 0 4 によって表明される一組のスペース 4 0 6 を定義しうるが、プール 4 0 4 のスペース 4 0 6 に、記憶装置 1 0 2 の物理的な容量を含む広がり 4 0 8 を割り当てることにおいて、いくらかの柔軟性を呈しうる。たとえば、サブディスク 3 0 4 のような大きなブロックではなく一連の小さな広がり 4 0 8 を割り当てることによって、この編成は、広がり 4 0 8 の再割り当てを通じて、スペース 4 0 6 の粒度の細かいサイズ変更を可能にする。さらに、スペース 4 0 6 は特定のサイズをもってプロビジョニングされることがあるが、ある実施形態は、広がり 4 0 8 を割り当て、該広がり 4 0 8 の物理アドレスを、それらのスペース 4 0 6 の論理アドレスにあとで、たとえばまだ広がり 4 0 4 にバインドされていないスペース 4 0 6 内の位置への書き込みが受領されるときに、バインドしてもよい。遅延された割り当ておよび遅延されたバインドというこれらの概念は、記憶装置 1 0 2 のフォーマット付けの迅速な完了 ( たとえば、スペース 4 0 6 を生成する要求の受領時に広がり 4 0 8 すべてを割り当てて初期化する必要がない ) およびオーバプロビジョニング ( たとえば、記憶装置 1 0 2 の利用可能な容量 5 0 8 を超える定義される容量をもつスペース 4 0 6 を生成し、スペース 4 0 6 の使用される容量が利用可能な容量 5 0 8 を使い尽くすときに、記憶セットに容量を追加するようユーザーに促す ) を可能にしうる。これらおよびその他の特徴は、図 4 ~ 図 5 の例示的なシナリオ 4 0 0、5 0 0 において示される記憶装置 1 0 2 の編成を通じて、本願で呈示される技法を通じて達成可能であってもよい。

#### 【 0 0 2 5 】

##### C . 例示的な実施形態

図 6 は、これらの技法のある例示的な実施形態の図解であり、少なくとも二つの記憶装置 1 0 2 を含む記憶セットを編成する例示的な方法 6 0 0 として示されている。例示的な方法 6 0 0 は、たとえば、装置のメモリ・コンポーネント ( たとえば、メモリ回路、ハードディスク・ドライブの円盤、半導体メモリ・コンポーネントまたは磁気もしくは光学式ディスク ) に記憶される、装置のプロセッサによって実行されたときに該装置に本願で呈示される技法を実行させる命令の組として実装されてもよい。例示的な方法 6 0 0 は、6 0 2 で始まり、プロセッサ上で命令を実行 6 0 4 することに関わる。具体的には、前記命令は、少なくとも二つのプールされた記憶装置 1 0 2 の記憶領域内で、プール構成設定 4 0 2 を含むプールされたパーティション 5 0 2 を生成 6 0 6 するよう構成されている。プール構成設定 4 0 2 は、プールされたパーティション 4 0 2 を記憶するプールされた記憶装置 1 0 2 と、プールされたパーティション 4 0 2 内で表現される少なくとも一つのスペース 4 0 6 を指定 6 0 8、6 1 0 する。前記命令はまた、スペース 4 0 6 について広がり 4 0 8 を割り当てる要求を受領 6 1 2 したときに、プールされたパーティション 4 0 2 内に該スペース 4 0 6 についての広がり 4 0 8 を割り当て 6 1 4、該広がり 4 0 8 をプール構成設定 4 0 2 においてスペース 4 0 6 と関連付けるようにも構成されている。このようにして、前記命令は、本稿に呈示される技法に基づく記憶装置 1 0 2 の編成を達成し、例示的な方法 6 0 0 は 6 1 8 で終了する。

#### 【 0 0 2 6 】

もう一つの実施形態は、本稿に提示される技法を適用するよう構成されたプロセッサ実行可能命令を有する例示的なコンピュータ可読媒体に関わる。そのようなコンピュータ可読媒体はたとえば、メモリ半導体 ( たとえば、静的ランダム・アクセス・メモリ ( SRAM )、動的ランダム・アクセス・メモリ ( DRAM ) および / または同期動的ランダム・アクセス

・メモリ（SDRAM）技術を利用した半導体）、ハードディスク・ドライブの円盤、フラッシュメモリ・デバイスまたは磁気もしくは光ディスク（たとえばCD-R、DVD-Rまたはフロッピー（登録商標）ディスク）のような有体なデバイスに関わるコンピュータ可読記憶媒体を含んでいてもよく、装置のプロセッサによって実行されたときに該装置に本稿に提示される技法を実装させるコンピュータ可読命令のセットをエンコードしていてもよい。そのようなコンピュータ可読媒体は、さまざまな物理現象を通じて伝搬させられうる信号（たとえば電磁信号、音波信号または光信号）のようなさまざまな型の通信媒体をも（コンピュータ可読記憶媒体とは異なる技術のクラスとして）含みうる。そのような伝搬は、さまざまな有線シナリオ（たとえばイーサネット（登録商標）または光ファイバーケーブルを介する）および／または無線シナリオ（たとえばWiFiのような無線ローカル・エリア・ネットワーク（WLAN）、ブルートゥースのようなパーソナル・エリア・ネットワーク（PAN）またはセルラーもしくは電波ネットワーク）においてでありうる。そのような媒体は、装置のプロセッサによって実行されたときに該装置に本稿に提示される技法を実装させるコンピュータ可読命令のセットをエンコードしていてもよい。

【0027】

これらの仕方で工夫されうる例示的なコンピュータ可読媒体が図7に示されている。図7では、実装700は、コンピュータ可読データ704がエンコードされているコンピュータ可読媒体702（たとえば、CD-R、DVD-Rまたはハードディスク・ドライブの円盤）を有する。このコンピュータ可読データ704は、本稿に記載される原理に基づいて動作するよう構成されるコンピュータ命令706の組を有する。あるそのような実施形態では、プロセッサ実行可能命令706は、装置710のプロセッサ712によって実行されたときに、該装置710に、図6の例示的な方法600のような記憶装置102の容量を編成する方法を実行させてもよい。このコンピュータ可読媒体のいくつかの実施形態は、このように構成されたプロセッサ実行可能命令を記憶するよう構成された非一時的なコンピュータ可読記憶媒体（たとえば、ハードディスク・ドライブ、光ディスクまたはフラッシュメモリ・デバイス）を含んでいてもよい。本稿に呈示される技法に従って機能するよう構成されている多くのそのようなコンピュータ可読媒体が当業者によって考案されうる。

【0028】

#### D．変形

本稿で論じられる技法は、多くの側面における変形をもって考案されてもよい。いくつかの変形は、追加的な利点を提示し、および／またはこれらおよび他の技法の他の変形に関する不都合な点を軽減しうる。さらに、いくつかの変形は、組み合わせて実装されてもよく、いくつかの組み合わせは、相乗作用を通じて追加的な利点および／または軽減された不都合な点を特徴としうる。そうした変形は、さまざまな実施形態に組み込まれて、そのような実施形態に対して個別のおよび／または相乗的な利点を付与しうる。

【0029】

#### D1．シナリオ

これらの技法の実施形態の間で変わりうる第一の側面は、そうした技法が利用されうるシナリオに関する。この第一の側面の第一の変形として、これらの技法は、ハードディスク・ドライブ、半導体記憶デバイス、不揮発性メモリ回路、テープ・ベースの記憶デバイスおよび磁気および光ディスクを含む多くの型の記憶装置102とともに使用されうる。そのような記憶装置102は、これらの技法を実装する装置710（コンピュータなど）に直接接続されてもよいし；有線または無線のローカル・エリア・ネットワーク（たとえば、802.11WiFiネットワークまたはアドホック接続または赤外線接続）を通じてアクセス可能であってもよいし；および／または有線または無線の広域ネットワーク（たとえばセルラー・ネットワークまたはインターネット）を通じてアクセス可能であってもよい。さらに、これらの技法が使われる二つ以上の記憶装置102は、独立して動作するのでも（たとえば、ソフトウェア・プロセスを通じて独立してアクセスされる記憶装置106）；ゆるやかな協働をもって動作するのでも（たとえば、独立して動作するが記憶セットを共有する他の記憶装置102のことを知らされており、該他の記憶装置と通信してもよい記

憶装置 102) ; あるいは緊密な協同をもって動作するのでも (たとえば、安価なディスクの冗長アレイ (RAID) コントローラがいくつかの記憶装置 106 を記憶システムの構成要素として管理する) よい。この第一の側面の第二の変形として、これらの技法は、さまざまな型のデータ・セットを含む多くの型の記憶セットとの関連で使用されてもよい。それには、さまざまな型のバイナリー・オブジェクトを記憶するバイナリー記憶システム ; ファイルを記憶するファイル・システム ; メディア・オブジェクトを記憶するメディア・ライブラリ ; 多くの型のオブジェクトを記憶するオブジェクト・システム ; レコードを記憶するデータベース ; 電子メール・メッセージを記憶する電子メール・システムが含まれる。この第一の側面の第三の変形として、これらの技法の一部または全部がコンピューティング環境内の一つまたは複数のコンポーネント内で、記憶装置 102 へのアクセスをもつコンピュータまたは装置の揮発性または不揮発性に記憶されているソフトウェア命令のセット (たとえば、オペレーティング・システム・プロセスまたはハードウェア・ドライバ) のように ; 記憶装置 102 とインターフェースをもつよう構成された記憶システム (たとえば RAID コントローラ) によって ; あるいは記憶セットのそれぞれの記憶装置 102 において、実装されてもよい。

10

#### 【0030】

この第一の側面の第四の変形として、記憶装置 104 の容量を編成するこれらの技法は、他の技法 (図 1 ~ 図 3 に示したようなもの) では達成するのが難しいことがあるいくつかの結果を達成しうる。第一の例として、これらの技法は、第一の記憶装置 102 上の少なくとも二つのパーティション 104 にまたがり、それらのパーティションを総合するプールされたパーティション 502 の生成を可能にしうる。第二の例として、これらの技法は、第一の記憶装置 102 上のプールされたパーティション 502 に記憶された少なくとも二つの広がり 408 ならびに第二の記憶装置 102 上のプールされたパーティション 502 に記憶された少なくとも一つの広がり 408 を含むスペース 406 の生成を可能にしうる。第三の例として、三つの記憶装置 102 のセットの間で、これらの技法は、第一の記憶装置 102 が、第一のプール 404 を第二の記憶装置 102 とは共有するが第三の記憶装置 102 とは共有せず、第二の記憶装置 102 が、第二のプール 404 を第三の記憶装置 102 とは共有するが第二の記憶装置 102 とは共有しないことを可能にしうる。この型の共有は、論理ディスク・マネージャ・データベース 306 のような他の技法の使用を通じては達成可能でないことがあり、特に、第一のプール 404 を共有する諸記憶装置 102 が、記憶装置 102 の異なるセットと共有される第二のプール 404 のプール構成設定 402 を更新することからは隔離された仕方ではプール構成設定 402 を更新することを可能にできなかったことがありうる。そのような多くの記憶シナリオが、本稿に呈示される技法に基づく記憶装置 102 の編成を通じて達成されうる。

20

30

#### 【0031】

##### D2 . プール構成設定

本技法の実施形態の間で変わりうる第二の側面は、記憶装置 102 によって共有されるプール 404 を表わすプール構成設定 (pool configuration) 402 の性質および使用に関する。この第二の側面の第一の変形として、プール構成設定 402 は一つの記憶装置 102 上に記憶されてもよく、あるいはプール 404 を記憶する諸記憶装置 102 上でミラーされていてもよい。この第二の側面の第二の変形として、プールされたパーティション 502 は多くの仕方では記憶装置 102 上で同定されうる。たとえば、記憶装置 102 は、それぞれのパーティション 104 の位置、サイズおよびパーティション型識別子を指定するパーティション・テーブル 110 を有していてもよく、プールされたパーティション 502 は、プールされたパーティション型識別子をもってパーティション・テーブル中で同定されうる。該識別子は、プールされたパーティション 502 を利用するよう構成された装置 710 によって理解可能であり、使用可能であってもよく、他の型の装置 710 によって無視されてもよい (よって誤使用されない)。

40

#### 【0032】

この第二の側面の第三の変形として、プール構成設定 402 は、プールされるパーティ

50

ション 5 0 2 の多くの領域に記憶されうる。それには、プールされる領域 5 0 2 の先頭、プールされる領域 5 0 2 の末尾、あるいはプールされる領域 5 0 2 内の定義されたまたは識別可能な位置が含まれる。簡単なアクセスのためおよび/またはプール構成設定 4 0 2 の増大を可能にするため、プール構成設定 4 0 2 をプールされる領域 5 0 2 の先頭に位置させることが有利であることがある（対照的に、プール構成設定 4 0 2 をプールされる領域 5 0 2 の末尾に位置させる場合には、上記増大は、プール構成設定 4 0 2 をプールされる領域 5 0 2 内の異なる位置に再書き込みすることなくしては達成が難しいことがある）。この第二の側面の第四の変形として、プール構成設定 5 0 2 は、リレーショナル・データベース、階層構造にされた文書（たとえば拡張可能マークアップ言語（XML）文書）またはテーブルなど、多くの仕方で構成されていてもよい。この第二の側面の第五の変形として、プール構成設定 5 0 2 は、名前、製造業者、モデル、容量、パフォーマンス特性および使用を含む、それぞれのエンティティ（たとえば、プール 4 0 2、スペース 4 0 6、広がり 4 0 8 および記憶装置 1 0 2）を記述するさまざまな型のメタデータを含んでいてもよい。特に、プール構成設定 5 0 2 は、ユーザー・データを記憶するスペース 4 0 6 のためのユーザー・スペース型識別子；他のスペース 4 0 6 のチェックサム（または他の型の検証子）を記憶するスペース 4 0 6 のためのチェックサム・スペース識別子；および他のスペース 4 0 6 に適用される書き込みを記憶するためのジャーナルとして使われるスペース 4 0 6 のためのジャーナル・スペース型識別子のような、それぞれのスペース 4 0 6 の型を同定するスペース型識別子を指定してもよい。さらに、スペース 4 0 6 は、シーケンシャルなまたはネストする関係を同定するよう親スペース 4 0 6 を同定してもよい（たとえば、第一のユーザー・データ・スペース 4 0 6 と連続している第二のユーザー・データ・スペース 4 0 6、あるいはジャーナル・スペースが親スペースを同定し、該親スペースについて、当該ジャーナル・スペースが該親スペースに適用されるべきジャーナル更新を記憶する）。

#### 【 0 0 3 3 】

この第二の側面の第六の変形として、広がり 4 0 8 は、記憶装置 1 0 2 上の固定した物理位置（たとえばトラックおよびセクタ）、記憶装置 1 0 2 上の物理アドレスのシーケンス内のある物理アドレス（たとえば、ある物理アドレス範囲）またはプールされるパーティション 5 0 2 内のオフセット（たとえばプールされるパーティション 5 0 2 の開始物理アドレスからのオフセット）を使って、プールされたパーティション 5 0 2 内の物理アドレスのブロックとして指定されてもよく、そのアドレス範囲を、スペース 4 0 6 内の論理アドレス範囲と関連付けてもよい（たとえば、開始論理アドレスおよび長さもしくは終了論理アドレスを指定）。

#### 【 0 0 3 4 】

この第二の側面の第七の変形として、スペース 4 0 6 への広がり 4 0 8 の割り当ておよびバインドはさまざまな仕方で達成されうる。図 8 ~ 図 1 0 は、このバインドを達成するためのいくつかの代替的なシナリオを呈示している。（これらの各図は、それぞれ図の左部分および右部分として示される、二つの時点における記憶装置 1 0 2 の状態を呈示する。）図 8 の例示的なシナリオ 8 0 0 では、第一の時点 8 0 2 において、記憶装置 1 0 2 は、プール 4 0 4 のためのプール構成設定 4 0 2 およびプール 4 0 4 を記憶する記憶装置 1 0 2 のレコードを記憶しているプールされたパーティション 5 0 2 を、大量の利用可能な容量 5 0 8 のほかに含む。第二の時点 8 0 4 では、ユーザー・データを記憶するスペース 4 0 6 を生成するスペース割り当て要求 8 0 6 が受容されるとき、プール構成設定 4 0 2 は、スペース 4 0 6 についての、その型およびプロビジョンされた容量を含むレコードを追加するよう更新されてもよい。さらに、この第二の時点 8 0 4 において、二つの広がり 4 0 8 が、プールされたパーティション 5 0 2 の利用可能な容量 5 0 8 から生成されてもよく、プール構成設定 4 0 2 は、該広がり 4 0 8 を、スペース 4 0 6 の論理位置にバインドする、該広がり 4 0 8 についてのレコードを追加するよう更新されてもよい。このようにして、スペース 4 0 6 についての広がり 4 0 8 はすぐに割り当てられ、使用のために利用可能となりうる。



## 【 0 0 3 5 】

あるいはまた、スペース 4 0 6 についての広がり 4 0 8 のバインドは遅延されてもよく、ジャストインタイム式に実行されてもよい。たとえば、図 9 の例示的なシナリオ 9 0 0 に示されるように、第一の時点 9 0 2 において、スペース 4 0 6 を生成するようスペース割り当て要求が受領されてもよく、スペース 4 0 6 についてのレコードがプール構成設定 4 0 2 内に生成されてもよい。しかしながら、スペース 4 0 6 の容量は第一の時点 9 0 2 ではフルには割り当てられなくてもよい。たとえば、第一の時点 9 0 2 では、一つの広がり 4 0 8 だけが割り当てられて、スペース 4 0 6 の他の部分（たとえば他のアドレス・ブロック）は割り当てられなくてもよい。しかしながら、第二の時点 9 0 4 において、広がり 4 0 8 がまだ割り当てられていないスペース 4 0 6 の特定のアドレスに書き込みするよう要求が受領されるとき、プールされたパーティション 5 0 2 の利用可能な容量 5 0 8 から第二の広がり 4 0 8 が割り当てられて、プール構成設定 4 0 4 内の広がりレコード 4 0 8 においてスペース 4 0 6 にバインドされてもよい。この遅延された割り当ては、スペース 4 0 6 の生成要求をより素早く充足することを可能にすることができ（第一の時点 9 0 2 で割り当てる必要があるのが全部より少ない広がり 4 0 8 なので）、および／またはシン・プロビジョニングを可能にすることができる。シン・プロビジョニングでは、スペース 4 0 6 は、スペース 4 0 6 にバインドされている広がり 4 0 8 の広がりサイズと、プール 4 0 4 を共有する記憶装置 1 0 2 の利用可能な容量 5 0 8 の和より大きいプロビジョニング・サイズをもってプロビジョンされてもよい。

## 【 0 0 3 6 】

図 1 0 は、スペース 4 0 6 の割り当てにおけるさらなる変形を含む例示的なシナリオ 1 0 0 0 の図解である。ここでは、スペース 4 0 6 についての広がり 4 0 8 の割り当ては、スペース 4 0 6 の特定の論理位置への広がり 4 0 8 のバインドからは分離されてもよい。この例示的なシナリオ 1 0 0 0 では、第一の時点 1 0 0 2 において、記憶装置 1 0 2 は、二つの広がり 4 0 8 が割り当てられているスペース 4 0 6 を有するプールされたパーティション 5 0 2 を有している。しかしながら、広がり 4 0 8 は、スペース 4 0 6 内の特定の論理位置にはマッピングされておらず、単にスペース 4 0 6 のための利用可能な容量として予約されている。この予約は、たとえばスペース 4 0 6 を生成する要求を受領した際に、実行されてもよく、それにより、スペース 4 0 6 のプロビジョニングされた容量の全部より少ない部分が予約され、スペース 4 0 6 にデータが書き込まれるべきアドレスがまだわからないとはいえ、何らかの容量が予約され、使用のために入手可能となる。よって、第二の時点 1 0 0 4 において、まだ広がり 4 0 8 がバインドされていないスペース 4 0 6 内の論理アドレスに書き込みをする書き込み要求が受領されるとき、バインドされていない広がり 4 0 8 の一つが選択され、書き込み要求において指定されている論理アドレスを含むスペース 4 0 6 の位置にバインドされうる。このようにして、スペース 4 0 6 の広がり 4 0 8 へのバインドは、スペース 4 0 6 のための容量の利用可能性を損なうことなく、スペース 4 0 6 の広がり 4 0 8 へのバインドを遅らせることができる。これらおよびその他の仕方において、改善されたパフォーマンスおよびシン・プロビジョニングのようなさまざまな特徴を可能にするよう広がり 4 0 8 がスペース 4 0 6 に割り当てられることができる。当業者は、本稿に呈示される技法を実装する際にプール構成設定 4 0 8 の使用における多くの変形を工夫しうる。

## 【 0 0 3 7 】

D 3 . プール構成設定所有者

本技法の実施形態の間で変わりうる第三の側面は、プール 4 0 4 のプール所有者の同定に関する。たとえば、プールをなす記憶装置 1 0 2 へのアクセスを有するコンピュータまたは他の装置 7 1 0 の間で、プール構成設定 4 0 2 に対する、コンピュータの間での書き込みアクセスを排他的に有するプール構成設定所有者が選定されてもよい。プール構成設定 4 0 2 を更新する要求（たとえば、スペース 4 0 6 を生成するまたは広がり 4 0 8 をスペース 4 0 6 にバインドする要求）を受領したとき、コンピュータは、自分がプール構成設定所有者であるかどうかを判定してもよく、そうでなければ、その要求をプール構成設

定所有者に転送してもよい。さらに、プール構成設定所有者が反応しなくなるまたは利用不能になる場合、他のコンピュータは、プール構成設定所有者の障害の検出に際して、それらのコンピュータの間で代替プール構成設定所有者を特定してもよい（たとえば、新たなプール構成設定所有者を選定する）。この選出は、プール構成設定 402 への更新が、競合状態から生じる衝突する更新を回避する仕方で行われることを可能にする。衝突する更新は、プール構成設定 402 を整合しない状態にすることがありうる。当業者は、プール構成設定 402 およびプール 404 の完全性をそのような状況から保護するようプール構成設定 402 を更新する仕方において多くの変形を構想し、利用しうる。

#### 【0038】

##### E. コンピューティング環境

図 11 は、本稿に提示される技法が実装されうるコンピューティング装置 1102 内の例示的なコンピューティング環境の図解を示している。例示的なコンピューティング装置は、これに限られないが、パーソナル・コンピュータ、サーバー・コンピュータ、ハンドヘルドもしくはラップトップ装置、モバイル装置（携帯電話、携帯情報端末（PDA）、メディア・プレーヤーなど）、マルチプロセッサ・システム、消費者電子装置、ミニコンピュータ、メインフレーム・コンピュータおよび上記のシステムもしくは装置のいずれかを含む分散式のコンピューティング環境を含む。

#### 【0039】

図 11 は、本稿に提示される一つまたは複数の実施形態を実装するよう構成されているコンピューティング装置 1102 を有するシステム 1100 の例を示している。ある構成では、コンピューティング装置 1102 は少なくとも一つのプロセッサ 1106 および少なくとも一つのメモリ・コンポーネント 1108 を含む。コンピューティング装置の厳密な構成および型に依存して、メモリ・コンポーネント 1108 は揮発性（たとえば RAM）、不揮発性（たとえば ROM、フラッシュメモリなど）または中間的もしくはハイブリッド型のメモリ・コンポーネントでありうる。この構成は、図 11 では破線 1104 によって示されている。

#### 【0040】

いくつかの実施形態では、装置 1102 は追加的な特徴／機能を含んでいてもよい。たとえば、装置 1102 は、これに限られないがハードディスク・ドライブ、半導体記憶デバイスおよび／または他のリムーバブルもしくは非リムーバブルな磁気もしくは光学式媒体を含む一つまたは複数の追加的な記憶コンポーネント 1110 を含んでいてもよい。ある実施形態では、本稿で与えられる一つまたは複数の実施形態を実装するコンピュータ可読かつプロセッサ実行可能な命令が記憶コンポーネント 1110 に記憶される。記憶コンポーネント 1110 は、オペレーティング・システムのコンポーネント、一つまたは複数のアプリケーションをなす実行可能なバイナリー、プログラミング・ライブラリ（たとえば、アプリケーション・プログラミング・インターフェース（API））、メディア・オブジェクトおよび文書類のような他のデータ・オブジェクトをも記憶していてもよい。コンピュータ可読命令は、プロセッサ 1106 による実行のためにメモリ・コンポーネント 1108 にロードされてもよい。

#### 【0041】

コンピューティング装置 1102 はまた、コンピューティング装置 1102 に他の装置と通信することを許容する一つまたは複数の通信コンポーネント 1116 をも含んでいてもよい。一つまたは複数の通信コンポーネント 1116 は（たとえば）モデム、ネットワーク・インターフェース・カード（NIC）、電波周波数送信機／受信機、赤外線ポートおよびユニバーサル・シリアル・バス（USB）USB 接続を有していてもよい。そのような通信コンポーネント 1116 は有線接続（物理的なコード、ケーブルまたはワイヤを通じてネットワークに接続する）または無線接続（可視光、赤外線または一つもしくは複数の電波周波数などを通じてネットワーキング装置と無線で通信する）を有していてもよい。

#### 【0042】

コンピューティング装置 1102 は、キーボード、マウス、ペン、音声入力装置、タッ

10

20

30

40

50

チ入力装置、赤外線カメラまたはビデオ入力装置のような一つまたは複数の入力コンポーネント 1 1 1 4 および / または一つまたは複数のディスプレイ、スピーカーおよびプリンターのような一つまたは複数の出力コンポーネント 1 1 1 2 を含んでいてもよい。入力コンポーネント 1 1 1 4 および / または出力コンポーネント 1 1 1 2 は有線接続、無線接続またはそれらの任意の組み合わせを介してコンピューティング装置 1 1 0 2 に接続されてもよい。ある実施形態では、別のコンピューティング装置からの入力コンポーネント 1 1 1 4 または出力コンポーネント 1 1 1 2 がコンピューティング装置 1 1 0 2 のための入力コンポーネント 1 1 1 4 および / または出力コンポーネント 1 1 1 2 として使用されてもよい。

#### 【 0 0 4 3 】

コンピューティング装置 1 1 0 2 のコンポーネントは、バスのようなさまざまな相互接続によって接続されてもよい。そのような相互接続は、PCI エクスプレスのような周辺コンポーネント相互接続 (PCI)、ユニバーサル・シリアル・バス (USB)、ファイアワイヤ (IEEE1394)、光学式バス構造などを含みうる。別の実施形態では、コンピューティング装置 1 1 0 2 のコンポーネントはネットワークによって相互接続されてもよい。たとえば、メモリ・コンポーネント 1 1 0 8 は、ネットワークによって相互接続された種々の物理的位置に位置されている複数の物理的メモリ・ユニットから構成されてもよい。

#### 【 0 0 4 4 】

当業者は、コンピュータ可読命令を記憶するために利用される記憶装置がネットワークを横断して分散されてもよいことを認識するであろう。たとえば、ネットワーク 1 1 1 8 を介してアクセス可能なコンピューティング装置 1 1 2 0 が本稿で与えられる一つまたは複数の実施形態を実装するためのコンピュータ可読命令を記憶してもよい。コンピューティング装置 1 1 0 2 はコンピューティング装置 1 1 2 0 にアクセスし、コンピュータ可読命令の一部または全部を実行のためにダウンロードしてもよい。あるいはまた、コンピューティング装置 1 1 0 2 は必要に応じてコンピュータ可読命令の諸片をダウンロードしてもよく、あるいはいくつかの命令はコンピューティング装置 1 1 0 2 において実行され、いくつかはコンピューティング装置 1 1 2 0 において実行されてもよい。

#### 【 0 0 4 5 】

##### F. 用語の用法

本願での用法では、用語「コンポーネント」、「モジュール」、「システム」、「インターフェース」などは、ハードウェア、ハードウェアとソフトウェアの組み合わせ、ソフトウェアまたは実行中のソフトウェアのいずれであれコンピュータ関係のエンティティを指すことが一般に意図されている。たとえば、コンポーネントは、これに限られないが、プロセッサ上で走っているプロセス、プロセッサ、オブジェクト、実行可能物、実行のスレッド、プログラムおよび / またはコンピュータであってもよい。例として、コントローラ上で走るアプリケーションおよび前記コントローラの両方がコンポーネントであることができる。一つまたは複数のコンポーネントがプロセスおよび / または実行のスレッド内に存在することができ、コンポーネントは一つのコンピュータ上に局在することもできるし、および / または二つ以上のコンピュータの間で分散させられることもできる。

#### 【 0 0 4 6 】

さらに、特許請求される主題は、開示される主題を実装するようコンピュータを制御するためのソフトウェア、ファームウェア、ハードウェアまたはそれらの任意の組み合わせを製造するための標準的なプログラミングおよび / またはエンジニアリング技法を使って、方法、装置または製造物として実装されてもよい。本稿で用いるところの用語「製造物」は、任意のコンピュータ可読デバイス、担体または媒体からアクセス可能なコンピュータ・プログラムを包含することが意図されている。むろん、当業者は、特許請求される主題の範囲および精神から外れることなく、この構成に多くの修正がなしうることを認識するであろう。

#### 【 0 0 4 7 】

本稿では諸実施形態のさまざまな動作が与えられている。ある実施形態では、記載され

10

20

30

40

50

る動作の一つまたは複数は、一つまたは複数のコンピュータ可読媒体上に記憶されているコンピュータ可読命令であって、コンピューティング装置によって実行されたら、記載される動作を該コンピューティング装置に実行させるものをなしてもよい。動作の一部または全部が記載される順序は、それらの動作が必ず順序依存であることを含意するものと解釈すべきではない。代替的な順序付けが、本稿の恩恵を受ける当業者によって理解されるであろう。さらに、すべての動作が本稿で与えられる各実施形態において必ず存在するわけではないことは理解されるであろう。

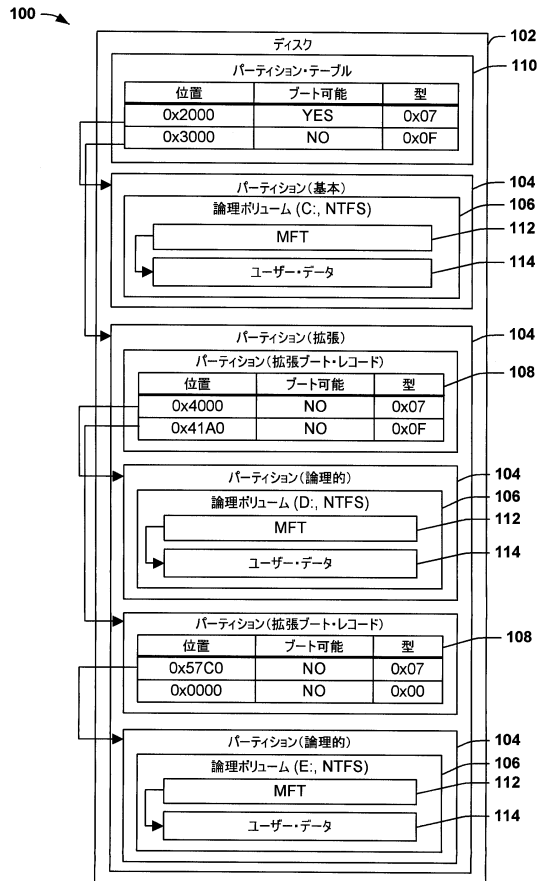
【0048】

さらに、「例示的」という語は、本稿では、例、事例または例解のはたらきをすることを意味するために使われる。本稿に「例示的」として記載される任意の側面または設計は、必ずしも、他の側面または設計に対して有利であると解釈されるものではない。むしろ、例示的という語の使用は、概念を具体的な仕方で提示することを意図したものである。本願での用法では、用語「または」は排他的な「または」ではなく包含的な「または」を意味することが意図されている。すなわち、特に断わりのない限りまたは文脈から明らかではない限り、「XはAまたはBを用いる」は自然な包含的な置換の任意のものを意味することが意図されている。すなわち、XがAを用いる；XがBを用いる；またはXがAおよびBの両方を用いるのであれば、「XはAまたはBを用いる」は上記の事例のいずれのもとでも満たされる。さらに、本願および付属の請求項で使われるところの冠詞「a」および「an」は、そうでないことが記載されているまたは単数形に向けられていることが文脈から明らかなのでない限り、「一つまたは複数」を意味する。

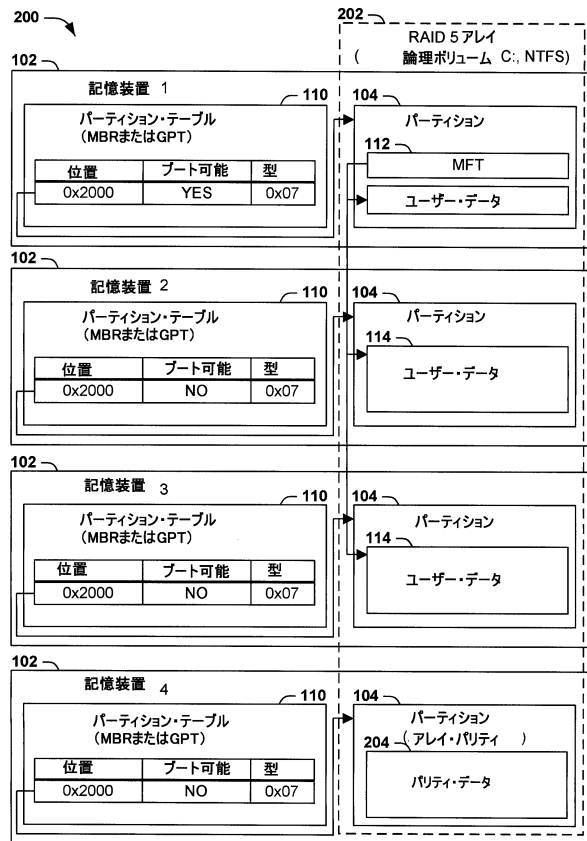
【0049】

また、本開示は、一つまたは複数の実装に関して図示され記述されたが、本明細書および付属の図面を読み、理解することに基づいて、他の当業者にも等価な変更および修正が思いつくであろう。本開示はそのようなすべての修正および変更を含み、付属の請求項の範囲によってのみ限定される。特に上記のコンポーネント（たとえば要素、資源など）によって実行されるさまざまな機能に関し、そのようなコンポーネントを記述するために使われる用語は、そうでないことが示されていない限り、本開示の本稿で例解された例示的な実装における当該機能を実行する開示された構造と構造的に等価でないとしても、記述されるコンポーネントの指定される機能を実行する（たとえば、機能的に等価な）任意のコンポーネントに対応することが意図されている。さらに、本開示の特定の特徴はいくつかの実装の一つのみに関して開示されたことがありうるが、そのような特徴は、任意の所与のまたは特定の応用のために所望されるまたは有利でありうるように、他の実装の一つまたは複数の他の特徴と組み合わせられてもよい。さらに、用語「含む」「有する」「もつ」またはそれらの変形が詳細な説明または請求項において使用される限りにおいて、そのような用語は用語「comprising」と同様の仕方で包含的であることが意図されている。

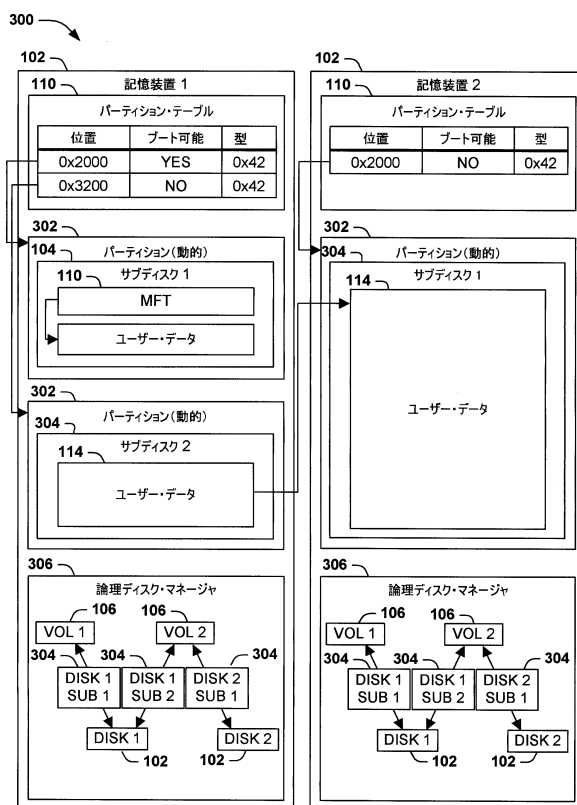
【図 1】



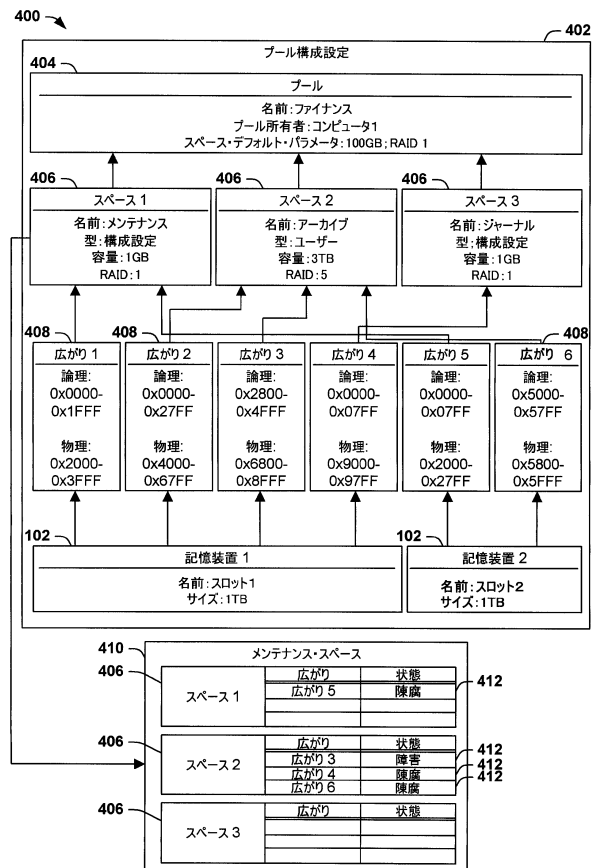
【図 2】



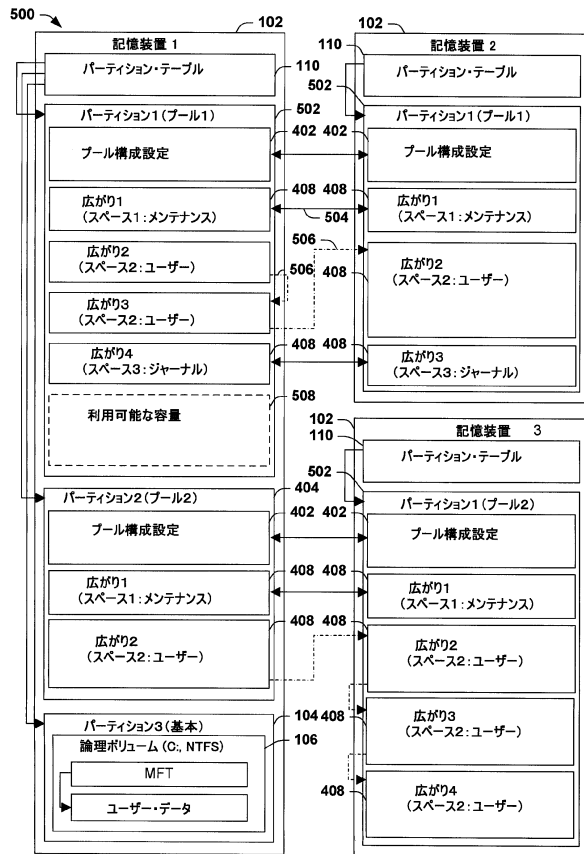
【図 3】



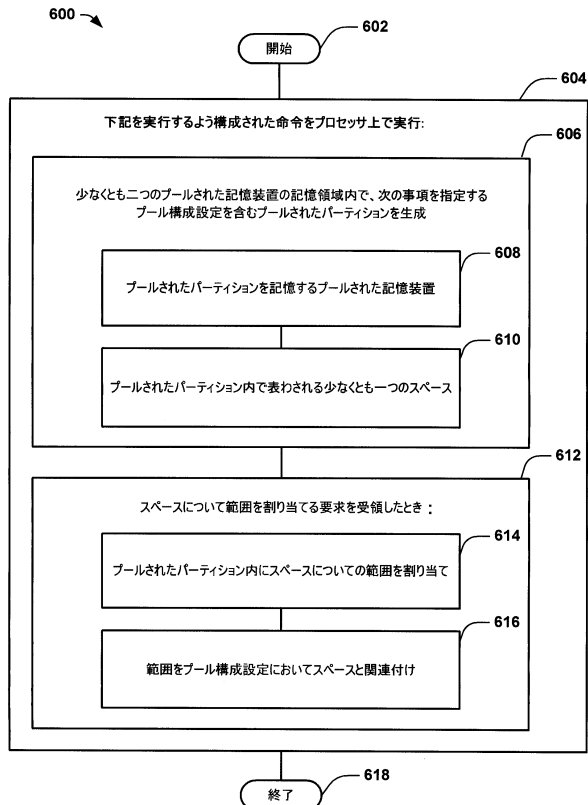
【図 4】



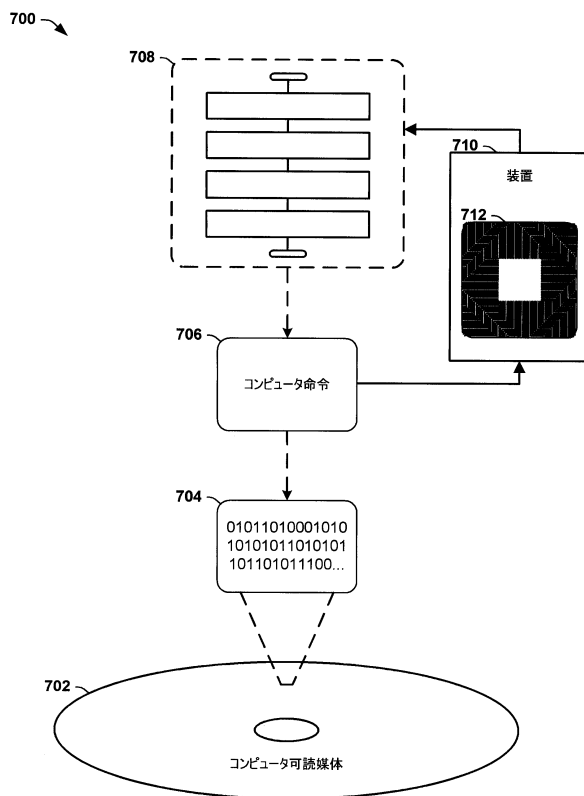
【図 5】



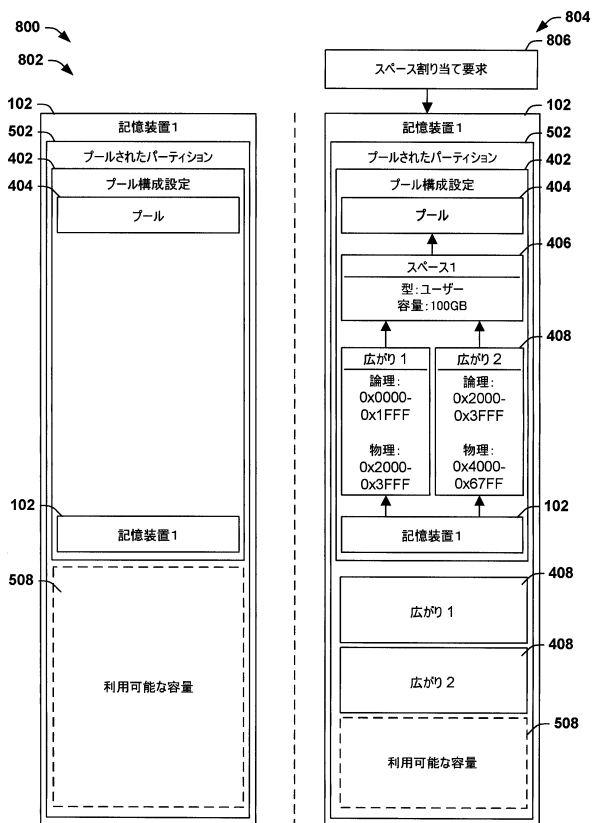
【図 6】



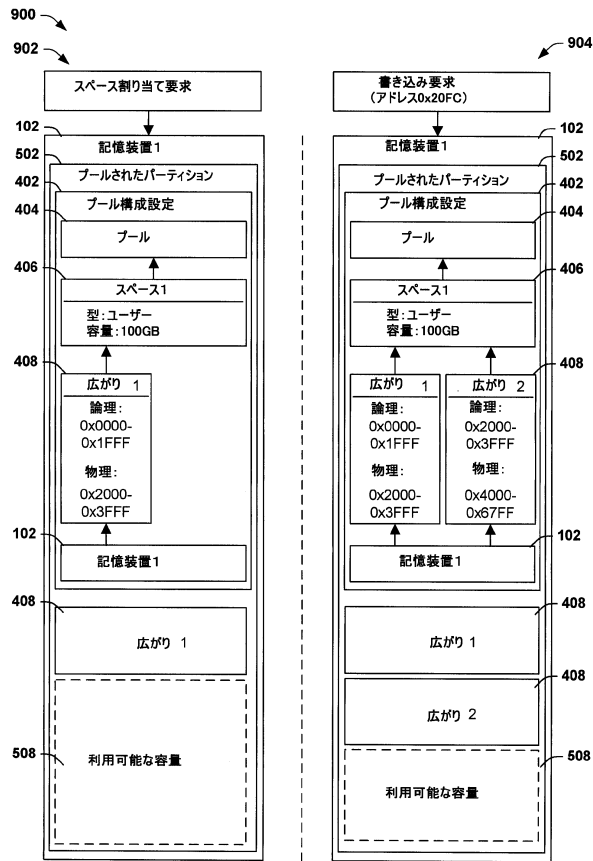
【図 7】



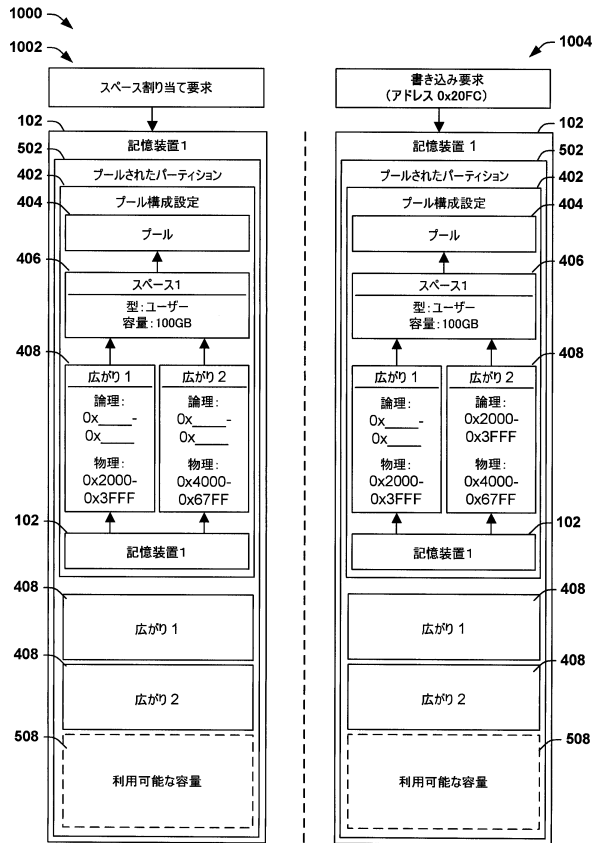
【図 8】



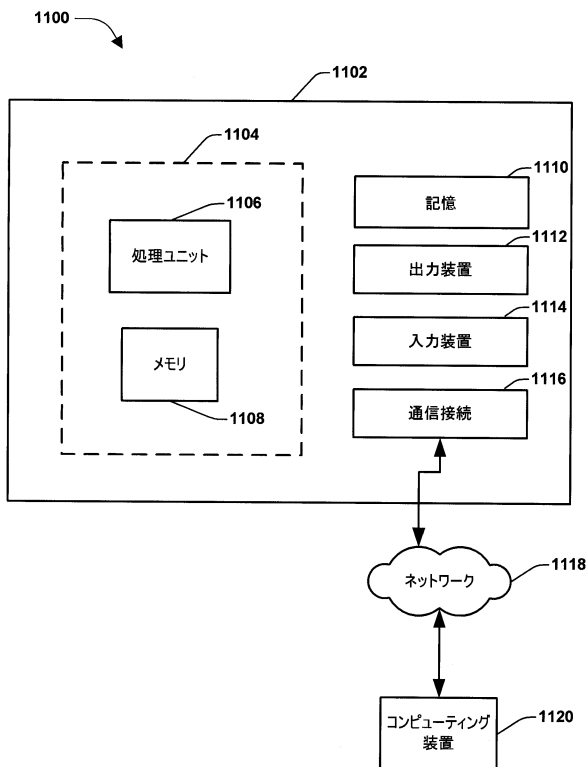
【図 9】



【図 10】



【図 11】



---

フロントページの続き

(72)発明者 メーラ, カラン

アメリカ合衆国 98052-6399 ワシントン州 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 コーポレーション エルシーエー - インターナショナル パテンツ 内

(72)発明者 パレオログ, エマニュエル

アメリカ合衆国 98052-6399 ワシントン州 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 コーポレーション エルシーエー - インターナショナル パテンツ 内

審査官 金沢 史明

(56)参考文献 国際公開第2011/108027(WO, A1)

米国特許出願公開第2008/0229048(US, A1)

米国特許出願公開第2004/0153606(US, A1)

米国特許出願公開第2005/0235337(US, A1)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00