

(19) 日本国特許庁 (JP)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2008-527555

(P2008-527555A)

(43) 公表日 平成20年7月24日 (2008.7.24)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00 501A	5B014
G06F 3/06 (2006.01)	G06F 3/06 301J	5B065
G06F 13/10 (2006.01)	G06F 12/00 514E	5B082
	G06F 12/00 545A	
	G06F 13/10 340A	
審査請求 未請求 予備審査請求 未請求 (全 23 頁)		

(21) 出願番号 特願2007-550798 (P2007-550798)
 (86) (22) 出願日 平成18年1月17日 (2006.1.17)
 (85) 翻訳文提出日 平成19年7月17日 (2007.7.17)
 (86) 国際出願番号 PCT/EP2006/050260
 (87) 国際公開番号 W02006/077215
 (87) 国際公開日 平成18年7月27日 (2006.7.27)
 (31) 優先権主張番号 11/037, 421
 (32) 優先日 平成17年1月18日 (2005.1.18)
 (33) 優先権主張国 米国 (US)

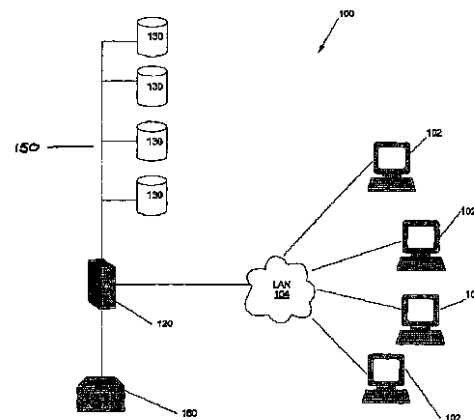
(71) 出願人 390009531
 インターナショナル・ビジネス・マシーンズ・コーポレーション
 INTERNATIONAL BUSINESS MACHINES CORPORATION
 アメリカ合衆国10504 ニューヨーク州 アーモンク ニュー オーチャードロード
 (74) 代理人 100108501
 弁理士 上野 剛史
 (74) 代理人 100112690
 弁理士 太佐 種一
 (74) 代理人 100091568
 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】 仮想化ストレージ・サブシステム内の仮想化ストレージ割り当ての自動パフォーマンス最適化を提供するための方法、装置およびプログラム・ストレージ・デバイス

(57) 【要約】

複数の仮想ストレージ・サブシステムにおけるストレージ割り当てを管理するためのポリシー・プロビジョニング・アドバイザーが開示される。データ・ストレージおよび作業負荷要求についての要求はユーザから受信される。仮想化エンジン内のストレージ・コントローラおよび管理ディスクのためのパフォーマンス情報が受信される。受信されたパフォーマンス情報および利用可能な管理ディスク、ならびに追跡されたそれらの利用度に基づいて、リソース・グループが定義される。仮想化エンジン内のリソース・グループ・パフォーマンスは、ユーザ作業負荷要求および現在または履歴リソース・グループ利用度に基づいて予測される。ストレージは、予測されたリソース・グループ・パフォーマンスに基づいて、仮想化エンジン内の1つ以上の管理ディスクに割り当てられる。



【特許請求の範囲】**【請求項 1】**

プログラム・ストレージ・デバイスであって、
仮想ストレージ・システムにおけるストレージ割り当てを管理するための操作を実行する処理デバイスにより実行可能なプログラム命令を含み、前記操作は、
作業負荷プロファイルを評価するステップと、
管理ディスクのパフォーマンス特性を決定するステップと、
前記管理ディスクとリソース・グループとの間の関係を決定するステップと、
前記管理ディスクが割り当てられるリソース・グループに基づいて、前記管理ディスクのセットを含む少なくとも 1 つの仮想ディスクの割り当てを推奨するステップとを含む、
プログラム・ストレージ・デバイス。

10

【請求項 2】

前記操作は、
前記作業負荷プロファイルを定義するステップをさらに含む、
請求項 1 に記載のプログラム・ストレージ・デバイス。

【請求項 3】

前記操作は、
前記管理ディスクが割り当てられるリソース・グループに基づいて、管理ディスクのセットを含む仮想ディスクを作成するステップをさらに含む、
請求項 1 または 2 に記載のプログラム・ストレージ・デバイス。

20

【請求項 4】

関係を決定する前記ステップは、ユーザ入力に基づく、請求項 1、2 または 3 に記載のプログラム・ストレージ・デバイス。

【請求項 5】

関係を決定する前記ステップは、自動的に提供される入力に基づく、請求項 1 ~ 3 のいずれか 1 項に記載のプログラム・ストレージ・デバイス。

【請求項 6】

仮想ディスクを作成する前記ステップは、前記管理ディスクが割り当てられる前記リソース・グループに基づいて、管理ディスクの前記セット内に複数の仮想ディスクを作成するステップをさらに含む、請求項 3 に記載のプログラム・ストレージ・デバイス。

30

【請求項 7】

管理ディスクのセットおよび特定の作業負荷についてのスループット・レイバビリティを含むリソース・グループと関連付けられたパフォーマンス特性のセットを識別することにより仮想ディスクのための位置を選択するステップと、

管理ディスクのセットおよびリソース・グループと関連付けられたパフォーマンス特性のセットに基づいて最適なりソース利用度を推定するステップとをさらに含む、

先行請求項のいずれか 1 項に記載のプログラム・ストレージ・デバイス。

【請求項 8】

作業負荷プロファイルを定義する前記ステップは、所望ストレージ量、前記作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成るグループから選ばれる少なくとも 1 つのものを選択するステップをさらに含む、請求項 2 に記載のプログラム・ストレージ・デバイス。

40

【請求項 9】

管理ディスクのパフォーマンス・レイバビリティを得るステップをさらに含む、先行請求項のいずれか 1 項に記載のプログラム・ストレージ・デバイス。

【請求項 10】

管理ディスクのパフォーマンス・レイバビリティを得る前記ステップは、管理ディスクのパフォーマンス・レイバビリティを示す手動入力を提供するステップをさらに含む、請求項 9 に記載のプログラム・ストレージ・デバイス。

【請求項 11】

50

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、管理ディスク・コンフィギュレータからパフォーマンス・ケイパビリティを受信するステップをさらに含む、請求項 9 に記載のプログラム・ストレージ・デバイス。

【請求項 12】

管理ディスク・コンフィギュレータからパフォーマンス・ケイパビリティを受信する前記ステップは、管理ディスクの前記パフォーマンス特性およびそれらの関連付けられたリソース・グループに関する情報を得るステップをさらに含み、前記リソース・グループは、管理ディスク間の潜在的な関係を識別する、請求項 11 に記載のプログラム・ストレージ・デバイス。

【請求項 13】

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、制御された較正を実行するステップおよび前記制御された較正からの前記結果を用いるステップをさらに含む、請求項 9 に記載のプログラム・ストレージ・デバイス。

【請求項 14】

制御された較正を実行する前記ステップは、前記管理ディスクの挙動を識別するために、前記管理ディスクに対して指定された I/O ロードを動作させるステップを含む、請求項 13 に記載のプログラム・ストレージ・デバイス。

【請求項 15】

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、制御されない較正を実行するステップおよび前記制御されない較正からの前記結果を用いるステップをさらに含む、請求項 9 に記載のプログラム・ストレージ・デバイス。

【請求項 16】

制御されない較正を実行する前記ステップは、前記管理ディスクの挙動を識別するために、アプリケーションが制御されない作業負荷として作動する際に前記パフォーマンス挙動を分析するステップを含む、請求項 15 に記載のプログラム・ストレージ・デバイス。

【請求項 17】

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、ユーザからインタフェースを介して入力を受信するステップをさらに含む、請求項 9 に記載のプログラム・ストレージ・デバイス。

【請求項 18】

管理ディスクのパフォーマンス特性を決定する前記ステップは、前記管理ディスクが属するリソース・グループを識別するステップをさらに含む、先行請求項のいずれか 1 項に記載のプログラム・ストレージ・デバイス。

【請求項 19】

前記ユーザ入力は、管理ディスクのパフォーマンス・プロファイルからの選択を含む、請求項 17 に記載のプログラム・ストレージ・デバイス。

【請求項 20】

前記ユーザ入力は、パフォーマンス・プロファイル・リソース・グループからの選択を含む、請求項 17 に記載のプログラム・ストレージ・デバイス。

【請求項 21】

作業負荷プロファイルを評価するステップと、
管理ディスクのパフォーマンス特性を決定するステップと、
前記管理ディスクとリソース・グループとの間の関係を決定するステップと、
前記管理ディスクが割り当てられるリソース・グループに基づいて、前記管理ディスクのセットを含む少なくとも 1 つの仮想ディスクの割り当てを推奨するステップとを含む、方法。

【請求項 22】

前記作業負荷プロファイルを定義するステップをさらに含む、
請求項 21 に記載の方法。

【請求項 23】

10

20

30

40

50

前記管理ディスクが割り当てられるリソース・グループに基づいて、前記管理ディスクのセットを含む仮想ディスクを作成するステップをさらに含む、

請求項 2 1 または 2 2 に記載の方法。

【請求項 2 4】

関係を決定する前記ステップは、ユーザ入力に基づく、請求項 2 1、2 2 または 2 3 に記載の方法。

【請求項 2 5】

関係を決定する前記ステップは、自動的に提供される入力に基づく、請求項 2 1 ~ 2 3 のいずれか 1 項に記載の方法。

【請求項 2 6】

仮想ディスクを作成する前記ステップは、前記管理ディスクが割り当てられる前記リソース・グループに基づいて、管理ディスクの前記セット内に複数の仮想ディスクを作成するステップをさらに含む、請求項 2 3 に記載の方法。

【請求項 2 7】

管理ディスクのセットおよび特定の作業負荷についてのスループット・レイバビリティを含むリソース・グループと関連付けられたパフォーマンス特性のセットを識別することにより仮想ディスクのための位置を選択するステップと、

管理ディスクのセットおよびリソース・グループと関連付けられたパフォーマンス特性のセットに基づいて最適なリソース利用度を推定するステップとをさらに含む、

請求項 2 1 ~ 2 6 のいずれか 1 項に記載の方法。

【請求項 2 8】

作業負荷プロファイルを定義する前記ステップは、所望ストレージ量、前記作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成る前記グループから選ばれる少なくとも 1 つのものをを選択するステップをさらに含む、請求項 2 2 に記載の方法。

【請求項 2 9】

管理ディスクのパフォーマンス・レイバビリティを得るステップをさらに含む、請求項 2 1 ~ 2 8 のいずれか 1 項に記載の方法。

【請求項 3 0】

管理ディスクのパフォーマンス・レイバビリティを得る前記ステップは、管理ディスクのパフォーマンス・レイバビリティを示す手動入力を提供するステップをさらに含む、請求項 2 9 に記載の方法。

【請求項 3 1】

管理ディスクのパフォーマンス・レイバビリティを得る前記ステップは、管理ディスク・コンフィギュレータからパフォーマンス・レイバビリティを受信するステップをさらに含む、請求項 2 9 に記載の方法。

【請求項 3 2】

管理ディスク・コンフィギュレータからパフォーマンス・レイバビリティを受信する前記ステップは、管理ディスクの前記パフォーマンス特性およびそれらの関連付けられたリソース・グループに関する情報を得るステップをさらに含み、前記リソース・グループは、管理ディスク間の潜在的な関係を識別する、請求項 3 1 に記載の方法。

【請求項 3 3】

管理ディスクのパフォーマンス・レイバビリティを得る前記ステップは、制御された較正を実行するステップおよび前記制御された較正からの前記結果を用いるステップをさらに含む、請求項 2 9 に記載の方法。

【請求項 3 4】

制御された較正を実行する前記ステップは、前記管理ディスクの挙動を識別するために、前記管理ディスクに対して指定された I / O ロードを動作させるステップを含む、請求項 3 3 に記載の方法。

【請求項 3 5】

10

20

30

40

50

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、制御されない較正を実行するステップおよび前記制御されない較正からの前記結果を用いるステップをさらに含む、請求項 29 に記載の方法。

【請求項 36】

制御されない較正を実行する前記ステップは、前記管理ディスクの挙動を識別するために、アプリケーションが制御されない作業負荷として作動する際に前記パフォーマンス挙動を分析するステップを含む、請求項 35 に記載の方法。

【請求項 37】

管理ディスクのパフォーマンス・ケイパビリティを得る前記ステップは、ユーザからインタフェースを介して入力を受信するステップをさらに含む、請求項 29 に記載の方法。

【請求項 38】

管理ディスクのパフォーマンス特性を決定する前記ステップは、前記管理ディスクが属するリソース・グループを識別するステップをさらに含む、請求項 21 ~ 37 のいずれか 1 項に記載の方法。

【請求項 39】

前記ユーザ入力は、管理ディスクのパフォーマンス・プロファイルからの選択を含む、請求項 37 に記載の方法。

【請求項 40】

前記ユーザ入力は、パフォーマンス・プロファイルリソース・グループからの選択を含む、請求項 37 に記載の方法。

【請求項 41】

仮想ストレージ・システムにおけるストレージ割り当てを管理するためのボリューム・プロビジョニング・アドバイザーであって、該ボリューム・プロビジョニング・アドバイザーは、請求項 21 ~ 40 のいずれか 1 項に記載の方法を実行するために動作可能である、ボリューム・プロビジョニング・アドバイザー。

【請求項 42】

リソース・グループは、名称、パフォーマンス属性および前記リソース・グループに属する管理ディスクのセットを含む、請求項 41 に記載のボリューム・プロビジョニング・アドバイザー。

【請求項 43】

管理ディスクは、1 つ以上のリソース・グループに属し、リソース・グループは、管理ディスク間の潜在的な関係を識別する、請求項 41 または 42 に記載のボリューム・プロビジョニング・アドバイザー。

【請求項 44】

コンピュータ・プログラムであって、前記プログラムがコンピュータ上で実行される場合、請求項 21 ~ 40 のいずれか 1 項に記載の方法を実行するようにされたプログラム・コード手段を含む、コンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ネットワーク・ストレージ・システムに関し、より具体的には、仮想化ストレージ・サブシステム内の仮想化ストレージ割り当ての自動パフォーマンス最適化を提供するための方法、装置およびプログラム・ストレージ・デバイスに関する。

【背景技術】

【0002】

会社、政府機関または他の事業体において用いられるような企業データ処理機構において、情報は、サーバに上に格納され、例えば、ネットワーク越しにユーザによりアクセスされることがよくある。情報は、処理されるべきどのようなタイプのプログラムまたはデータあるいはその両方の情報も含み得る。自分のパーソナル・コンピュータ、ワークステーション等（一般に、「コンピュータ」）を使用するユーザは、自分のコンピュータが、

10

20

30

40

50

処理されるべき情報を取得し、さらに、情報を、例えば、遠隔サーバ上に格納できるようにする。

【 0 0 0 3 】

一般に、サーバは、多くのディスク・ストレージ・ユニットを通常含む大容量ストレージ・サブシステムにデータを格納する。データは、ファイルのような単位で格納される。1つのサーバ中で、1つのファイルは、1つのディスク・ストレージ・ユニット上に格納されてもよく、あるいは、1つのファイルのいくつかの部分がいくつかのディスク・ストレージ・ユニット上に格納されてもよい。サーバは、多くのユーザからのアクセス要求に同時にサービスでき、同時にサービスされるアクセス操作が、それらのアクセス操作が同時にサービスされるように複数のディスク・ストレージ・ユニット全体にわたって分布された情報と関連していることが好ましいことが理解されるであろう。別の言い方をすれば、他のディスク・ドライブ・ユニットが軽負荷またはアイドル状態になっている一方で、1つのディスク・ドライブ・ユニットに負荷がかかりすぎる、すなわちビジー・サービング・アクセスにならないやり方で情報をディスク・ストレージ・ユニット中に格納することが一般に望ましい。

10

【 0 0 0 4 】

ビジネスのコンピュータ・ネットワークは、互いにそしてビジネス・ユーザから離れて位置する複数のストレージ・ネットワークを有することがある。ストレージ・ネットワークは、異なるタイプのシステム上でホストされることもある。作業を正しく実行するため、ビジネス・ユーザは、ストレージ・ネットワークのすべてに含まれているデータへの迅速かつ信頼性の高いアクセスを要求することがある。情報科学（IT）要員は、高速で信頼性の高いアクセスをビジネス・ユーザに提供できなければならない。

20

【 0 0 0 5 】

ストレージ・エリア・ネットワーク（SAN）は、データ・ストレージ・デバイスをサーバに論理的に接続する高速な高帯域幅ストレージ・ネットワークである。そして今度は、ビジネス・ユーザが通常、サーバを介してデータ・ストレージ・デバイスに接続される。SANは、伝統的なサーバ/ストレージ接続により示される概念を拡張し、より高い柔軟性、可用性、統合管理およびパフォーマンスを提供する。SANは、企業内の情報ヘユーザがいつでもアクセスできるようにするための第1のITソリューションである。一般に、SANは、ホスト、相互接続デバイス、ストレージ・デバイス、およびネットワーク・アタッチ・サーバ（NAS）デバイスのようなネットワーク・デバイスを定義するための管理ソフトウェアを含む。SAN管理ソフトウェアは、リンクがデバイス間で定義されることも可能にする。SAN内では、ソフトウェアは、単一の仮想ディスクとして特徴付けられると同時に、多数のストレージ・ディスクにわたってデータが格納される仮想ストレージを定義することができる。

30

【 0 0 0 6 】

高速で信頼性の高いアクセスをビジネス・ユーザに提供するというこの目標に到達する際の1つの重要な構成要素は、SANを設計および保守する人々にSANが完全に理解されるようにすることである。SANは、その複雑さため、迅速に理解することが困難なことがよくある。SANおよびSAN内の仮想システムの構成を迅速に理解および変更されるようにするツールが有益である。

40

【 0 0 0 7 】

SANの利点の1つは、多数のクライアントについてストレージ・アクセスを管理するサーバにおいて生じ得る隘路の排除である。ストレージへの共用アクセスを可能にすることにより、SANは、より低いデータ・アクセス待ち時間および向上したパフォーマンスに備えることができる。しかしながら、SANアタッチ・ストレージのような大きいストレージ・ネットワークにおいては、新たに割り当てられたスペースが考え得る最良のパフォーマンスを達成するように、ストレージの増加分をどこに割り当てるかをストレージ・アドミニストレータが知ることは困難である。これは、多数の仮想化ストレージ・サブシステムを含み得るネットワークが複雑であること、作業負荷を分析することが複雑である

50

こと、および物理ストレージ属性がアプリケーションから隠されることがあるためである。

【 0 0 0 8 】

過去において、大ストレージ環境のためのストレージ割り当ては手動で実行されてきた。アルゴリズムに基づいてストレージを割り当て、またはどこに割り当てるかを推奨できるストレージ管理ソフトウェアが利用可能である。それにもかかわらず、これらのアルゴリズムは実際に、S A N内の仮想ストレージ・システムを含む利用可能なストレージの制約内で生産性能要件を満たすことを試みない。

【 発明の開示 】

【 発明が解決しようとする課題 】

10

【 0 0 0 9 】

本発明は、仮想化ストレージ・サブシステム内の仮想化ストレージ割り当ての自動パフォーマンス最適化を提供するための方法、装置およびプログラム・ストレージ・デバイスを開示する。

【 課題を解決するための手段 】

【 0 0 1 0 】

第1の態様によれば、プログラム・ストレージ・デバイスが提供され、プログラム・ストレージ・デバイスは、仮想ストレージ・システムにおいてストレージ割り当てを管理するための操作を実行する処理デバイスにより実行可能なプログラム命令を含み、操作は、作業負荷プロファイルを評価するステップと、管理ディスクのパフォーマンス特性を決定するステップと、管理ディスクとリソース・グループとの間の関係を決定するステップと、管理ディスクが割り当てられるリソース・グループに基づいて、管理ディスクのセットを含む少なくとも1つの仮想ディスクの割り当てを推奨するステップとを含む。

20

【 0 0 1 1 】

1つの実施形態において、操作は、作業負荷プロファイルを定義するステップをさらに含む。

【 0 0 1 2 】

1つの実施形態において、操作は、管理ディスクが割り当てられるリソース・グループに基づいて、管理ディスクのセットを含む仮想ディスクを作成するステップをさらに含む。

30

【 0 0 1 3 】

1つの実施形態において、関係を決定するステップは、ユーザ入力に基づく。

【 0 0 1 4 】

1つの実施形態において、関係を決定するステップは、自動的に提供される入力に基づく。

【 0 0 1 5 】

1つの実施形態において、仮想ディスクを作成するステップは、管理ディスクが割り当てられるリソース・グループに基づいて管理ディスクのセット内に複数の仮想ディスクを作成するステップをさらに含む。

【 0 0 1 6 】

40

1つの実施形態において、仮想ディスクのための位置は、管理ディスクのセットおよび特定の作業負荷についてのスループット・レイバリティを含むリソース・グループと関連付けられたパフォーマンス特性のセットを識別することにより選択され、最適なりソース利用度は、管理ディスクおよびリソース・グループと関連付けられたパフォーマンス特性のセットに基づいて推定される。

【 0 0 1 7 】

1つの実施形態において、作業負荷プロファイルを定義するステップは、所望ストレージ量、作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成るグループから選ばれる少なくとも1つのものを選択するステップをさらに含む。

50

【 0 0 1 8 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティが得られる。

【 0 0 1 9 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティを得るステップは、管理ディスクのパフォーマンス・ケイパビリティを示す手動入力を提供することをさらに含む。

【 0 0 2 0 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティを得るステップは、管理ディスク・コンフィギュレータからパフォーマンス・ケイパビリティを受信するステップをさらに含む。

10

【 0 0 2 1 】

1つの実施形態において、管理ディスク・コンフィギュレータからパフォーマンス・ケイパビリティを受信するステップは、管理ディスクのパフォーマンス特性および管理ディスクの関連付けられたリソース・グループに関する情報を得るステップをさらに含み、リソース・グループは、管理ディスク間の潜在的な関係を識別する。

【 0 0 2 2 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティを得るステップは、制御された較正を実行するステップおよび制御された較正からの結果を用いるステップをさらに含む。

20

【 0 0 2 3 】

1つの実施形態において、制御された較正を実行するステップは、管理ディスクの挙動を識別するために、管理ディスクに対して指定されたI/Oロードを動作させるステップを含む。

【 0 0 2 4 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティを得るステップは、制御されない較正を実行するステップおよび制御されない較正からの結果を用いるステップをさらに含む。

【 0 0 2 5 】

1つの実施形態において、制御されない較正を実行するステップは、管理ディスクの挙動を識別するために、アプリケーションが制御されない作業負荷として作動する際にパフォーマンス挙動を分析するステップを含む。

30

【 0 0 2 6 】

1つの実施形態において、管理ディスクのパフォーマンス・ケイパビリティを得るステップは、ユーザからインタフェースを介して入力を受信するステップをさらに含む。

【 0 0 2 7 】

1つの実施形態において、管理ディスクのパフォーマンス特性を決定するステップは、管理ディスクが属するリソース・グループを識別するステップをさらに含む。

【 0 0 2 8 】

1つの実施形態によれば、プログラム・ストレージ・デバイスが提供され、プログラム・ストレージ・デバイスは、仮想ストレージ・システムにおけるストレージ割り当てを管理するための操作を実行する処理デバイスにより実行可能なプログラム命令を含み、操作は、作業負荷プロファイルを定義するステップと、管理ディスクのパフォーマンス特性を決定するステップと、管理ディスクとリソース・グループとの間の関係を決定するステップと、管理ディスクが割り当てられるリソース・グループに基づいて管理ディスクのセットを含む仮想ディスクを作成するステップとを含む。

40

【 0 0 2 9 】

1つの実施形態によれば、仮想ストレージ・システムにおけるストレージ割り当てを管理するデバイスが提供される。このデバイスは、ストレージ・システム情報を格納するためのメモリおよびこのメモリに結合されたプロセッサを含む。このプロセッサは、作業負

50

荷プロファイルを定義する際に使用するユーザ・インタフェースならびに管理ディスクのパフォーマンス特性を決定し、管理ディスクとリソース・グループとの間の関係をユーザ定義されたまたは自動化された入力に基づいて決定し、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクを作成するための仮想ディスク・アロケータを提供するように構成される。

【0030】

1つの実施形態において、プロセッサは、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセット内に1つ以上の仮想ディスクを作成する。

【0031】

1つの実施形態において、プロセッサ・アロケータは、管理ディスクのセットおよび特定の作業負荷についてのスループット・レイパビリティを含むリソース・グループに関連付けられたパフォーマンス特性のセットを識別し、管理ディスクのセットおよびリソース・グループに関連付けられたパフォーマンス特性に基づいて最適リソース利用を推定することにより、仮想ディスクのための位置を推奨する。

10

【0032】

1つの実施形態において、作業負荷プロファイルは、所望ストレージの量、作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成るグループから選ばれる少なくとも1つのものを含む。

【0033】

1つの実施形態において、プロセッサは、グラフィック・インタフェースを提供する。

20

【0034】

1つの実施形態において、プロセッサは、コマンド・ライン・インタフェースを用いるスクリプト化されたアプリケーションを提供する。

【0035】

1つの実施形態において、プロセッサは、アプリケーション・プログラミング・インタフェース（API）を用いるスクリプト化されたアプリケーションを提供する。

【0036】

1つの実施形態において、プロセッサは、手動入力、管理ディスク・コンフィギュレータ、制御された較正結果、制御されない較正結果および標準インタフェース入力を含むグループのうちの少なくとも1つを用いて管理ディスクのパフォーマンス・レイパビリティを決定する。

30

【0037】

1つの実施形態において、手動入力は、管理ディスクのパフォーマンス・プロファイルからの選択を含む。

【0038】

1つの実施形態において、手動入力は、パフォーマンス・プロファイル・リソース・グループからの選択を含む。

【0039】

本発明の1つの実施形態において、仮想ストレージ・システムにおいてストレージ割り当てを管理するための方法が提供される。この方法は、ユーザ定義された作業負荷プロファイルを得るステップと、管理ディスクのパフォーマンス特性を決定するステップと、管理ディスクとリソース・グループとの間の関係をユーザ定義されたまたは自動化された入力に基づいて決定するステップと、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクを作成するステップとを含む。

40

【0040】

別の態様によれば、作業負荷プロファイルを評価するステップと、管理ディスクのパフォーマンス特性を決定するステップと、管理ディスクとリソース・グループとの間の関係を決定するステップと、管理ディスクが割り当てられるリソース・グループに基づいて管理ディスクのセットを含む1つ以上の仮想ディスクの割り当てを推奨するステップとを含む方法が提供される。

50

【 0 0 4 1 】

1つの実施形態によれば、仮想ストレージ・システムにおいてストレージ割り当てを管理するボリューム・プロビジョニング・アドバイザーが提供される。ボリューム・プロビジョニング・アドバイザーは、作業負荷プロファイルを定義する際に使用するユーザ・インタフェースおよびこのユーザ・インタフェースに作動的に結合された仮想ディスク・アロケータを含み、仮想ディスク・アロケータは、管理ディスクのパフォーマンス特性を決定し、管理ディスクとリソース・グループとの間の関係をユーザ定義されたまたは自動化された入力に基づいて決定し、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクを作成する。

【 0 0 4 2 】

1つの実施形態において、作業負荷プロファイルは、所望ストレージの量、作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成るグループから選ばれる少なくとも1つのものを含む。

【 0 0 4 3 】

1つの実施形態において、リソース・グループは、名称、パフォーマンス属性、およびリソース・グループに属する管理ディスクのセットを含む。

【 0 0 4 4 】

1つの実施形態において、管理ディスクは、1つ以上のリソース・グループに属し、リソース・グループは、管理ディスク間の潜在的な関係を識別する。

【 0 0 4 5 】

別の態様によれば、仮想ストレージ・システムにおいてストレージ割り当てを管理するためのボリューム・プロビジョニング・アドバイザーが提供され、このボリューム・プロビジョニング・アドバイザーは、作業負荷プロファイルを評価する作業負荷評価器および仮想ディスク・アロケータを含み、仮想ディスク・アロケータは、管理ディスクのパフォーマンス特性を決定し、管理ディスクとリソース・グループとの間の関係を決定し、管理ディスクが割り当てられるリソース・グループに基づいて管理ディスクのセットを含む1つ以上の仮想ディスクの割り当てを推奨する。

【 0 0 4 6 】

別の実施形態によれば、仮想ストレージ・システムにおいてストレージ割り当てを管理するためのボリューム・プロビジョニング・アドバイザーが提供され、このボリューム・プロビジョニング・アドバイザーは、ユーザ定義された作業負荷プロファイルを得るための手段と、管理ディスクのパフォーマンス特性を決定するための手段と、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクの割り当てを推奨するための手段とを含む。

【 0 0 4 7 】

本発明の別の実施形態において、別のボリューム・プロビジョニング・アドバイザーが提供される。ボリューム・プロビジョニング・アドバイザーのこの実施形態は、ユーザ定義された作業負荷プロファイルを得るための手段と、管理ディスクのパフォーマンス特性を決定するための手段と、管理ディスクとリソース・グループとの間の関係をユーザ定義されたまたは自動化された入力に基づいて決定するための手段と、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクを作成するための手段とを含む。

【 0 0 4 8 】

別の実施形態によれば、別のボリューム・プロビジョニング・アドバイザーが提供される。このボリューム・プロビジョニング・アドバイザーは、ユーザ定義された作業負荷プロファイルを得るための手段と、管理ディスクのパフォーマンス特性を決定するための手段と、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクを作成するための手段とを含む。

【 0 0 4 9 】

別の実施形態によれば、仮想ストレージ・システムにおいてストレージ割り当てを管理

10

20

30

40

50

するための方法が提供され、この方法は、ユーザ定義された作業負荷プロファイルを得るステップと、管理ディスクのパフォーマンス特性を決定するステップと、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクの割り当てを推奨するステップとを含む。

【 0 0 5 0 】

1つの実施形態において、管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクが作成される。

【 0 0 5 1 】

1つの実施形態において、仮想ディスクを作成するステップは、管理ディスクのセットおよび特定の作業負荷についてのスループット・レイバビリティを含むリソース・グループと関連付けられたパフォーマンス特性のセットを識別するステップと、管理ディスクのセットおよびリソース・グループと関連付けられた識別されたパフォーマンス特性のセットから導き出されたリソース利用を推定するステップとをさらに含む。

【 0 0 5 2 】

1つの実施形態において、ユーザ定義された作業負荷プロファイルを得るステップは、所望ストレージの量、作業負荷の特性の識別および所望のスループットのレベル、ならびに割り当てヒントから成るグループから少なくとも1つを選択するステップをさらに含む。

【 0 0 5 3 】

1つの実施形態において、管理ディスクのパフォーマンス・レイバビリティを決定するステップは、手動入力、得られた構成データ、制御された較正結果、制御されない較正結果および標準インタフェース入力をから成るグループのうちの少なくとも1つを用いるステップを含む。

【 発明を実施するための最良の形態 】

【 0 0 5 4 】

本発明の実施形態は、単に例として、そして以下の図面（同じ参照符号は、全体を通して対応する部材を表す）に関連して説明される。

【 0 0 5 5 】

実施形態の以下の説明において、本明細書の一部を成す添付図面が参照され、図面には、本発明を実施できる特定の実施形態が例示として示される。他の実施形態が利用され得ることが理解されるべきである。なぜならば、構造変化は、本発明の範囲を逸脱することなくなされ得るからである。

【 0 0 5 6 】

本発明は、仮想化ストレージ・サブシステム内の仮想化ストレージ割り当ての自動パフォーマンス最適化を提供するための方法、装置およびプログラム・ストレージ・デバイスを提供する。

【 0 0 5 7 】

図1は、ローカル・エリア・ネットワーク（LAN）の形態のコンピュータ・ネットワーク100を例示する。図1において、ワークステーション・ノード102は、LAN相互接続104を介してサーバ120に結合されている。データ・ストレージ130は、データ・バス150を介してサーバ120に結合されている。LAN相互接続100は、イーサネット（登録商標）のような、任意数のネットワーク・トポロジーであってよい。

【 0 0 5 8 】

図1に示されるネットワークは、ネットワークのクライアント・サーバ・モデルとして知られている。クライアントは、サービスまたは他のリソースを共用するネットワークに接続されたデバイスである。サーバ120が、これらのサービスまたはリソースを管理する。サーバ120は、コンピュータまたはソフトウェアプログラムであって、クライアント102にサービスを提供する。サーバにより管理され得るサービスとしては、データ・ストレージ130へのアクセス、サーバ120または他の接続されたノード（図示せず）から提供されるアプリケーション、あるいはプリンタ160の共用が含まれる。

10

20

30

40

50

【 0 0 5 9 】

図 1 において、ワークステーション 1 0 2 は、サーバ 1 2 0 のクライアントであり、サーバ 1 2 0 により管理されるデータ・ストレージ 1 3 0 へのアクセスを共用する。ワークステーション 1 0 2 の 1 つがデータ・ストレージ 1 3 0 へのアクセスを必要とする場合、ワークステーション 1 0 2 は、LAN 相互接続 1 0 0 を介してサーバ 1 2 0 に要求を提出する。サーバ 1 2 0 は、ワークステーション 1 0 2 からデータ・ストレージ 1 3 0 へのアクセスについての要求に応える。サーバとストレージとの間での考え得る相互接続技術は、ファイバ・チャンネル・プロトコル (F C P) またはスモール・コンピュータ・システム・インタフェース (S C S I) である。

【 0 0 6 0 】

図 1 に示されるようなネットワークが成長すると、新しいクライアント 1 0 2 が付加され、より多くのストレージ 1 3 0 が付加され、サービス要求が増大し得る。上述のように、サーバ 1 2 0 は、ストレージ 1 3 0 へのアクセスについてのすべての要求に応える。その結果、ストレージ 1 3 0 上での I O 負荷が劇的に増大することがあり、サーバ 1 2 0 上での作業負荷が劇的に増大し、パフォーマンス低下の可能性という結果になることがある。従来のクライアント・サーバ・モデルの帯域幅制限の低減を助けるため、ストレージ・エリア・ネットワーク (S A N) が近年いよいよ普及してきている。ストレージ・エリア・ネットワークは、サーバおよびストレージと高速で相互接続する。LAN のような既存のネットワークング・モデルをストレージ・エリア・ネットワークと組み合わせることにより、コンピュータ・ネットワーク全体のパフォーマンスが改善され得る。

【 0 0 6 1 】

図 2 は、本発明の実施形態による S A N 2 0 0 の 1 つの実施形態を示す。図 2 において、サーバ 2 0 2 は、S A N 相互接続 2 0 4 を介してデータ・ストレージ・デバイス 2 3 0 に結合されている。各サーバ 2 0 2 および各ストレージ・デバイス 2 3 0 は、S A N 相互接続 2 0 4 に結合されている。サーバ 2 0 2 は、S A N 相互接続に接続されたストレージ・デバイス 2 3 0 のいずれにも直接アクセスできる。S A N 相互接続 2 0 4 は、ファイバ・チャンネルまたは S C S I のような、高速相互接続であり得る。図 2 が示すように、サーバ 2 0 2 およびストレージ・デバイス 2 3 0 は、それら自体の内外にネットワークを含む。

【 0 0 6 2 】

図 2 の S A N 2 0 0 において、どのサーバ 2 0 2 も、LAN におけるように特定のストレージ・デバイス 2 3 0 の専用ではない。どのサーバ 2 0 2 も、図 2 における S A N 2 0 0 上のどのストレージ・デバイス 2 3 0 にもアクセスできる。S A N 2 0 0 の典型的な特性としては、高帯域幅、サーバからストレージ・ノードへの多数の経路、大きい接続距離、および非常に大きいストレージ容量が含まれ得る。その結果、S A N の複雑さに伴い、ファイバ・チャンネル・ベースの S A N 2 0 0 のパフォーマンス、柔軟性およびスケラビリティは、典型的な S C S I ベースのシステムのそれよりもかなり大きくなり得る。

【 0 0 6 3 】

図 2 は、S A N 相互接続 2 0 4 に結合されたネットワーク・アドミニストレータ 2 7 0 も示す。ディスクからデータ経路まで、すべての利用可能な S A N リソースを有効に利用し、十分なデータ保護および回復性に備えるやり方で S A N 2 0 0 中のストレージ 2 3 0 を割り当てることが特に重要である。アドミニストレータ 2 7 0 は、ストレージ要素の大きいネットワーク内のストレージ位置の選択を支援するように構成され得る。アドミニストレータ 2 7 0 は、仮想ディスク・アロケータ (V D A) 2 7 2 を含み、これは、本発明の実施形態によれば、所望のパフォーマンスの水準、ユーザの作業負荷の属性、ストレージの変化するパフォーマンス属性および異なるタイプの作業負荷に対するその応答、ならびにネットワーク内の競合作業負荷の存在が与えられれば、顧客の指定パフォーマンスおよびスペース要件に従って入力 / 出力ストレージ割り当てを処理する。

【 0 0 6 4 】

仮想ディスク・アロケータ 2 7 2 は、ストレージ要素のネットワーク内でのストレージ

についての要求を、要求によって、またはストレージ方針機構を介して指定されたパフォーマンス要求を適合するようなやり方で満足させる。仮想ディスク・アロケータ 272 は、仮想化ストレージ・サブシステムである IBM（登録商標）2145 SAN ボリューム・コントローラ（SAN VC）のような環境において動作し得る。（IBM は、インターナショナル・ビジネス・マシーニズ・コーポレーションの登録商標である。）仮想ディスク・アロケータ 272 は、管理ディスクのパフォーマンス特性を決定する。仮想ディスク・アロケータ 272 は、管理ディスクとリソース・グループとの間の関係を、ユーザ定義されたまたは自動化された入力に基づいて決定し、リソース・グループ、ならびに管理ディスクが割り当てられるキャッシュおよびデータ経路のようなリソース・グループ・ストレージリソースを考慮して、管理ディスクのセットを含む仮想ディスクを作成する。

10

【0065】

仮想ディスク・アロケータ 272 は、ポリシー・ベースの考え方をオープン・システム環境に拡張し、パフォーマンス要件を満たすために仮想化ストレージ・サブシステム内のストレージ要素の選択を自動化する。仮想化ストレージ・システム内の記憶素子を選択して提示する際には、OS またはボリューム・マネージャ・ソフトウェア、あるいは DB2（登録商標）およびその他のデータベース製品のようなストライプ化ボリュームの概念をサポートする（データベース・アプリケーションのような）アプリケーションが最適に使用されることを考慮する。（DB2 は、インターナショナル・ビジネス・マシーニズ・コーポレーションの登録商標である。）仮想ディスク・アロケータ 272 は、長期データ使用パターンを考慮してストレージを割り当てるという考えも拡張する。仮想ディスク・アロケータ 272 は、データ配置のインテリジェント選択を行うために必要とされる様々なアルゴリズムを実装している。

20

【0066】

仮想ディスク・アロケータ 272 が、どのノード、すなわち、仮想化エンジン 274 のようなエンジンがデータにアクセスしてよいか、およびどの管理ディスク・グループ（MDG）、すなわち仮想ディスクをサポートするディスクのグループが、選択されるべき LUN を構成するかを決定を行うようにしてもよい。MDG 内には少なくとも 1 つの管理ディスクがあり、この管理ディスクは、仮想ディスク内のデータをストライプ化するために仮想化エンジン 274 およびボリューム・マネージャ 276 により用いられ、これは、エンタプライズ・ストレージ・システム（ESS）の論理ディスクに相当する。従って、仮想ディスク・アロケータ 272 は、顧客の所望のパフォーマンス・レベルを満たすために、1 つの LUN または複数のストレージ要素全体にわたる複数のリソース・グループ中の複数の LUN を選択することができる。

30

【0067】

アドミニストレータは、システムに接続されているディスクのパフォーマンス・ケイパビリティを見出すために、較正プロセス 278 を実行できる。これには、リソース・グループにおいて特定の構成に編成された個々のディスクのパフォーマンス・ケイパビリティについての特定の知識を単に用いるのではなく、それらのディスクのグループのパフォーマンス・パラメータを見出すために、特定のテストを実施することも当然含まれよう。

【0068】

図 3 は、本発明の実施形態による仮想ディスク・アロケータ 272 に実装される属性表 300 を例示する。これらの属性としては、ユーザ作業負荷属性および所望のパフォーマンス・レベルの情報（310）、管理ディスクまたは任意のリソース・グループのパフォーマンス特性の決定（312）、ユーザ定義されたまたは自動化された入力に基づいて決定される管理ディスクとリソース・グループとの間の関係（314）、および管理ディスクが割り当てられるリソース・グループを考慮に入れて生成される管理ディスクのセットを有する仮想ディスク（316）が含まれる。

40

【0069】

アプリケーションが要求する作業負荷、あるいは少なくともそれらの作業負荷について妥当な見積もりをすることなく、インテリジェントにデータ配置の決定を行うことはほと

50

んど不可能である。例えば、ユーザが 100GB のストレージを要求すれば、軽い性能が要求されるのならば、単一の 100GB 論理ディスクの割り当てでもよいかもしれないが、高性能アプリケーションの場合は、10 台のディスク・アレイ全体にわたり 10 個の 10GB 論理的ディスクの割り当て、およびそれらの論理ディスク全体にわたるデータのストライピングを要求するかも知れない。あいにく、ほとんどの顧客は、自らの作業負荷がどのように見えるかについて知識を持っていない。

【0070】

図 4 は、本発明の実施形態による、管理ディスクをリソース・グループに割り当てる際に、仮想ディスク・アロケータによって使用される作業負荷プロファイルを得るためのユーザ・インタフェースのための機構 400 を例示する。作業負荷プロファイルは、作業負荷および所望のスループットのレベルの特性を記述し、所望のストレージの量および割り当てヒントのような情報を含んでもよい。

【0071】

最初に、予め用意された作業負荷プロファイル (canned workload profile) が提供されてもよい (410)。図 2 を参照すると、仮想ディスク・アロケータ 272 が、メモリ 292 中に予め用意された作業負荷プロファイルを提供してもよい。予め用意された作業負荷プロファイル 410 は、様々な産業および用途における顧客環境の特性に基づいて定めることができる。例として、有名な所与のバンド作業負荷のセット、例えば、SAP__OLTP、DB2 ビジネス・インテリジェンス等が、提供され得る。アプリケーション専門家からの助言を得て、顧客は最初に、これらの所与の作業負荷 410 のうちの 1 つを選択する。

【0072】

作業負荷プロファイルはまた、顧客の作業負荷 412 の観測に基づいて自動的に作成されてもよい。それぞれの顧客の作業負荷は固有の属性を有するので、顧客の環境におけるストレージ・アクセス・パターンを観測することにより、より良好な作業負荷見積もりが得られる。図 2 を参照すると、仮想ディスク・アロケータ 272 は、その判断の多くを、観測されたディスク・アクセス挙動に基づかせることができ、仮想ディスク・アロケータ 272 は、この観測されたディスク・アクセス挙動をデータベースの形でメモリ 292 中に保持する。仮想ディスク・アロケータ 272 と関連して動作するグラフィック・インタフェースのようなユーザ・インタフェースにより、ユーザはボリュームをグループ化し特定の時間枠を指定し、次にそれらのボリュームの観測された挙動に基づいて作業負荷プロファイルを作成できる。このように、仮想ディスク・アロケータ 272 は、顧客の作業負荷について学習し、時間の経過につれてその意志決定を向上させる。なお、ユーザ・インタフェースは、コマンド・ライン・インタフェース (CLI) を用いるスクリプト化アプリケーションであってもよい。

【0073】

作業負荷プロファイルは、インテリジェント・ソフトウェア・コンポーネント 414 によっても提供され得る。図 2 を参照すると、仮想ディスク・アロケータ 272 は、作業負荷プロファイルを提供するインテリジェント・ソフトウェア・コンポーネントも含み得る。これらの作業負荷プロファイルは、アプリケーション固有の特別な知識に基づくこともある。

【0074】

図 5 は、管理ディスクまたは任意のリソース・グループ (以下、管理ディスク) のパフォーマンス・ケイパビリティを決定するための機構 500 を例示する。管理ディスク・パフォーマンス・ケイパビリティは、管理ディスクをリソース・グループに割り当てるために、作業負荷プロファイルと共に仮想ディスク・アロケータ 272 により用いられる。そのような機構は、アドミニストレータからからの手動入力から、パフォーマンス・ケイパビリティを導き出すための自動化された発見的方法に及ぶあらゆる方法を含む。また、制御されたまたは制御されないストレージ環境のいずれかにおいて、較正作業負荷の使用によるケイパビリティの較正を用いることによりパフォーマンスを決定するための機構も含

まれる。

【 0 0 7 5 】

管理ディスクのパフォーマンス・ケイパビリティを決定するための1つの機構として、手動入力アプローチ510でもよい。このアプローチは簡単であるが、管理ディスク・パフォーマンス・ケイパビリティの理解へのアプローチとしてはあまり望ましくない。例えば、少数の管理ディスク・パフォーマンス・プロファイルが定義できるにすぎないであろう（例えば、「ミラード・ディスク」、「8つのディスクを備えるRAID-5アレイ」）。これらのプロファイルの各々は、特定のデフォルト性能属性を有する。アドミニストレータはその場合、管理されたディスク構成に一致する適切なプロファイルを選択するであろう。このアプローチでは、管理ディスク・ケイパビの非常に多様性が生かされないであろう。

10

【 0 0 7 6 】

管理ディスクのパフォーマンス・ケイパビリティの把握への別のアプローチは、管理ディスク・コンフィギュレータ512を用いることである。デバイス特有のコンフィギュレータは、「把握された」ディスク・コントローラの有限集合に対するパフォーマンス・特性および管理ディスクとストレージ・コントローラ間の関係に関する入力を提供できるであろう。例えば、FAST900コンフィギュレータは、RAIDアレイ、LUN、および管理ディスクをサブシステム上に作成する統一的な技術を提供し、それらの管理ディスクについて管理ディスク・パフォーマンス見積もりを提供するであろう。コンフィギュレータは、管理ディスクとストレージ・コントローラとの間の関係を動的に認識することもできるであろう。

20

【 0 0 7 7 】

別のアプローチにおいて、管理ディスクの挙動を観測する間、管理ディスクに対し指定されたI/O負荷を走らせるために、制御された較正（キャリブレーション）514が用いられる。通常、制御された較正は、管理ディスク上にスペースが割り当てられる前に実行されるが、スペースが予約されていれば、後に実行されてもよい。

【 0 0 7 8 】

制御されない較正を行う手法516では、管理ディスクのパフォーマンス挙動は、アプリケーションが管理ディスクに対して正常作業負荷を走らせる際に分析される。管理ディスクがすでに割り当てられる場合には（そして、較正用に予約された空きスペースがまったくなければ）、このアプローチが用いられるであろう。制御された較正ステップ以来条件が変わった（例えば、管理ディスクが、劣化した条件で動作している）かどうかを決定することも有用である。例として、制御されない較正によれば、特定の負荷点において応答時間が一貫して高いことを観測し、この負荷点が最大スループット・ケイパビリティをより正確に反映すると見積もることもあるだろう。

30

【 0 0 7 9 】

もう1つのアプローチは、LUNおよびLUNと物理リソースとの関連付け、ならびにそれらの要素の静的パフォーマンス・ケイパビリティを識別する概念モデルをSMIS仕様が提供できるようにし得る標準化インタフェース518を用いる。

【 0 0 8 0 】

上述のアプローチに類似した他のアプローチを用いて割り当て決定を行うことができる。同様に、上述のアプローチの組み合わせを選ぶことができる。例えば、アドミニストレータは、SANVCの背後に構成された特定のタイプの第1のストレージ・コントローラに対して、制御された較正ステップを走らせ、次に、その後のコントローラについてその同じタイプの属性、例えば、「これらの管理ディスクは、それらの管理ディスクと同様に見える」、を手動で割り当てても知れない。

40

【 0 0 8 1 】

仮想ディスク・アロケータ272により用いられる作業負荷パラメータは、ディスク・ストレージ・パフォーマンスを正確に予測するそれらの能力に基づいて、またデータ収集ツールを介してのそれらの一般的な入手しやすさに基づいて選択される。用いられる作業

50

負荷パラメータとしては、管理ディスクの各々に最大スループット、毎秒あたりの最大ランダム読み取りおよび書き込み、毎秒あたりの最大混合読み取りおよび書き込み、毎秒あたりの最大順次読み取りおよび書き込み、毎秒あたりの最大混合読み取りおよび書き込み、ならびに低負荷および最大負荷における待ち時間が含まれる。

【 0 0 8 2 】

混合読み取りおよび書き込みは、コンポーネントが双方向機能を有しているかどうかを決定するのに重要である。短ブロック・ランダム操作は、処理能力およびコストを決定するために用いられ、大ブロック順次計量は、データ転送能力およびコストを決定するために用いられる。

【 0 0 8 3 】

これらの最大スループットは次に、読み取りおよび書き込み操作あたり、ならびに読み取りおよび書き込み転送メガバイトあたりのオーバーヘッドを決定するために用いられる。オーバーヘッドは次に、管理ディスクの予測される利用を決定するために、予測される作業負荷特性に適用される。

【 0 0 8 4 】

管理ディスクは、個別のパフォーマンス・ケイパビリティを有しており、それらの管理ディスクと関連付けられた他のリソースのパフォーマンス・ケイパビリティを考慮する場合に異なるパフォーマンス・ケイパビリティを有するので、仮想ディスク・アロケータ 272 は、リソース・グループのセットを定義し、各リソース・グループは、そのリソース・グループに属する管理ディスクのセットについて関連するパフォーマンス属性を有する。管理ディスク・パフォーマンス・ケイパビリティおよびそれらの関連付けられたリソースのパフォーマンス・ケイパビリティを考慮することにより、管理ディスクの全体パフォーマンスを決定することができる。これらの他のリソースは、管理ディスクがある物理ディスクまたはコントローラを含むかも知れない。しかしながら、管理ディスクと関連付けられた考え得るリソースすべてを識別する必要はない。代わりに、仮想ディスク・アロケータ 272 は、リソース・グループを定義する時に、リソースとの関連付けを一般化する。

【 0 0 8 5 】

各リソース・グループは、(管理ディスクについてと同じ)パフォーマンス属性を与えられる。リソース・グループのパフォーマンス・ケイパビリティは、手動で指定、または制御された較正と呼ばれる制御されたベンチマーク手順を介して決定され得る。較正は、識別された管理ディスクに対して、アドミニストレータにより識別された時間で実行され得る。管理ディスクが仮想ディスクに割り当てられる前に、較正は管理ディスクに対して実行され得る。これは、指定されたストレージ・コントローラ上での管理ディスクの作成を自動化する操作手順内に含まれることもあり得る。較正は、グループ中の管理ディスクすべてが割り当てられない限り、個別の管理ディスクおよび関連付けられたリソース・グループすべてに対して実行されるであろう。較正において用いられる技術は、競合するディスク・サブシステムに対してエンジニアリング測定を実施する際に用いられる技術と同様である。

【 0 0 8 6 】

管理ディスクが 1 つ以上のリソース・グループに属し得るので、リソース・グループの較正により、データ割り当て決定の最適化が可能になる。しかしながら、管理ディスクと関連リソースとの間の潜在的な関係すべてを見出す必要はなく、疑われる関係が存在しない場合、影響はまったく生じない。仮想ディスク・アロケータ 272 を様々な実施することにより、関係を見出す際に、可変レベルの高度化を組み込むことができるかも知れない。

【 0 0 8 7 】

図 6 は、本発明の実施形態による、仮想化ストレージ・サブシステム中の重要なパフォーマンス要素を抽出するために仮想ディスク・アロケータ 272 により使用されるデータ構造 600 を例示する。データ構造は、クラスタ、デバイス・アダプタ、個別ディスクま

10

20

30

40

50

たはディスク・アレイおよび任意の関連リソースのようなストレージ素子を表すノードのツリーとして表すことができる。しかしながら、当業者は、本発明が図6に示される構造に限定されることを意図するものでないことを理解するであろう。むしろ、ツリー構造ではなく、より一般的なノードのネットワークが用いられることがある。

【0088】

データ構造600は、ストレージ割り当ての目標志向アプローチを提供するために使用される。例として、SAN VCは、伝統的なストレージ・コントローラと多少異なるストレージのモデルを提示する。アプリケーション・サーバは、IBMエンタープライズ・ストレージ・サーバ（登録商標）（ESS）における論理ディスクに相当する仮想ディスクを目指して提示される。仮想ディスク・アロケータ272は、特定の作業負荷について10
のスループット・レイバビリティを含む、管理ディスクのセットおよびリソース・グループと関連付けられたパフォーマンス特性を識別し、管理ディスクのセットおよびリソース・グループと関連付けられたパフォーマンス特性のセットから導き出されたリソース利用を推定することにより、仮想ディスクを作成する。仮想ディスクは、例えば、SAN VC 1次ノードを介してアクセスされ、仮想ディスクのためのデータはそのノード中にキャッシュされ、これはESSクラスタに相当する機能である。図6を参照すると、仮想ディスク605のためのストレージは、1つ以上の管理ディスク610、620、630および640にわたって割り当てられる。例示される管理ディスクは、管理ディスク・グループ（MDG）650にグループ化される。仮想ディスク605は、単一のMDG650に20
関連付けられている。一般に、最高パフォーマンスのため、仮想ディスク605は、MDG650中の管理ディスクすべてにわたってストライプされるであろう。

【0089】

管理ディスク610、620、630および640の各々は、他のアプリケーションと共用される単一の物理ディスクの一部として割り当てられたり、いくつかのディスクから成るRAIDアレイ全体として割り当てられるかも知れない。加えて、ストレージ・アドミニストレータは、管理ディスク構成の選択において非常に大きい柔軟性を与えられる。他のアプリケーションとの関連付けの結果として、1つの管理ディスク上でアクティビティは、別の管理ディスクのパフォーマンスに対して、それらが同じストレージ・デバイス上または同じストレージ・コントローラの背後にある場合のように、影響を与え得る。図6に関連して、管理ディスク610は、3つのリソース611、612、および613に30
関連付けられている一方で、管理ディスク620は、リソース613、621および622に関連付けられている。リソース613が両方の管理ディスク610、620に関連付けられているので、管理ディスクのパフォーマンス特性の各々は、未知かつ非常に可変であり得る。従って、インテリジェントなデータ配置決定の鍵は、管理ディスクのレイバビリティおよび挙動に関する妥当なパフォーマンス見積もりを取得し、維持することである。

【0090】

本発明の実施形態によれば、各管理ディスクは、1以上のリソース・グループ（RG）に関連付けられる。RG1（660）、RG2（670）、およびRG3（680）が仮想ディスク・アロケータ272により定義される。RG1（660）は、管理ディスク610ならびにリソース611、612および613を含む。RG2（670）は、管理ディスク620、630および管理ディスク640の一部を含み、かつリソース613、621、622、631、632および641を有する。RG3（680）は、管理ディスク640、ならびにリソース632および641の一部を含む。見て取れるように、リソース613は、管理ディスク610および620により共用され、リソース632および641は、管理ディスク630および640により共用される。管理ディスクおよび関連付けられたリソースがグループ化されて、リソース・グループ内およびリソース・グループ間での相互作用の観測が可能になるので、管理ディスクのレイバビリティおよび挙動に関し、強化されたパフォーマンス見積もりが仮想ディスク・アロケータ272によりなされ得る。40
50

【0091】

例えば、ボリューム選択プロセスの間、仮想ディスク・アロケータ272は、管理ディスクおよびリソース・グループの利用度を、ESS用VPAにおける構成ツリーの構成要素についてそれらの利用度を維持するやり方と同様に、常時監視する。当てのための潜在的な管理ディスクを選択する場合、1つの管理ディスクが属するリソース・グループのすべてについて影響が評価される。管理ディスクの利用度は次に、すべての関連RGのうちで最も高いと想定される。これは、VPAの現行バージョンにより用いられる構成ツリーを進む手法と同様である。仮想ディスク・アロケータ272は、指定された容量およびパフォーマンス要件を満たすストレージ割り当てについて推奨する。

【0092】

10

図2を再度参照すると、仮想ディスク・アロケータ272は、ストレージ素子が実際にどのように動作しているかを知ることにより改善されるが、極めて正確な情報には依存せず、これは、仮想ディスク・アロケータ272が、異なるベンダの不均質タイプのストレージのために動作できるからである。しかしながら、正確なリアルタイムまたは履歴パフォーマンス・データは、1つのベンダ売主の製品を他と区別するため、ならびにストレージ割り当てにバイアスをかけて、対象となる期間の間競合する可能性がある作業負荷から離すために用いられ得る。

【0093】

20

仮想ディスク・アロケータ272の重要な側面は、利用可能なリソース全体にわたってストレージ割り当てを均衡させるために容量およびパフォーマンス構造の使用を伴う。仮想ディスク・アロケータ272において複数の選択が可能な場合、容量およびパフォーマンス構造は、疑似乱数を使用して割り当てを1つのリソースのセットにバイアスするために用いることができる。いくつかのサンプル割り当てをこのようにして選択することができ、サンプル中で最良のものが答えとして選ばれる。この手法により、同様な状況でアルゴリズムが同じ推奨をすることを防止でき、利用可能なリソース全体にわたって作業負荷の広がりを増大させる。このように、ストレージ割り当ては、指定された作業負荷を扱う能力が最も高い、ネットワーク中の要素に向かってバイアスをかけられる。

【0094】

30

図7は、本発明の実施形態による、仮想ストレージ・システムにおけるストレージ割り当てを管理するための方法の流れ図700を例示する。この方法は、ユーザ定義された作業負荷プロファイルを得るステップ710と、管理ディスクのパフォーマンス特性を決定するステップ720とを含む。管理ディスクとリソース・グループとの間の関係が、ユーザ定義されたまたは自動化された入力に基づいて決定される(730)。管理ディスクが割り当てられるリソース・グループを考慮して管理ディスクのセットを含む仮想ディスクが生成される(740)。

【0095】

40

図3図7に関連して例示されるプロセスは、コンピュータ可読媒体またはキャリア、例えば、図2に例示される固定式または取り外し可能あるいはその両方のデータ・ストレージ・デバイス288のうちの1つ以上、あるいは他のデータ・ストレージ・デバイスまたはデータ通信デバイスにおいて具体化され得る。データ・ストレージ・デバイス288またはコンピュータ・プログラム290は、実行のためにアドミニストレータ270または仮想ディスク・アロケータ272を構成するために、メモリ292にロードされ得る。そのようなコンピュータ・プログラムは命令を含んでおり、これらの命令は、図2のプロセッサ294のようなプロセッサによって読み取られると、アドミニストレータ270または仮想ディスク・アロケータ272に、本発明のステップまたは要素を実行するのに必要なステップを実行させる。

【0096】

本発明の代表的な実施形態の上記の説明は、例示および説明の目的で提示された。網羅的であることや、本発明を、開示された正確な形態に限定することは意図されていない。上記の教示に照らし、多くの変更および変形が可能である。本発明の範囲は、この詳細な

50

説明によってではなく、むしろ添付の特許請求の範囲により限定されるものである。

【図面の簡単な説明】

【0097】

【図1】ローカル・エリア・ネットワークの形態のコンピュータ・ネットワークを例示する。

【図2】本発明の実施形態によるストレージ・エリア・ネットワークの1つの実施形態を例示する。

【図3】本発明の実施形態による仮想ディスク・アロケータに組み込まれる属性の表を例示する。

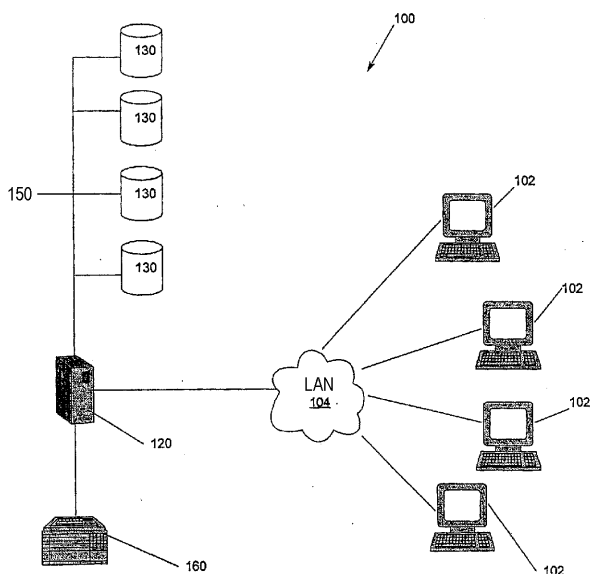
【図4】本発明の実施形態による、管理ディスクをリソース・グループに割り当てる際に仮想ディスク・アロケータによって使用される作業負荷プロファイルを得るユーザ・インタフェースのための機構を例示する。

【図5】本発明の1つの実施形態による、管理ディスクのパフォーマンス・ケイパビリティを決定するための機構を例示する。

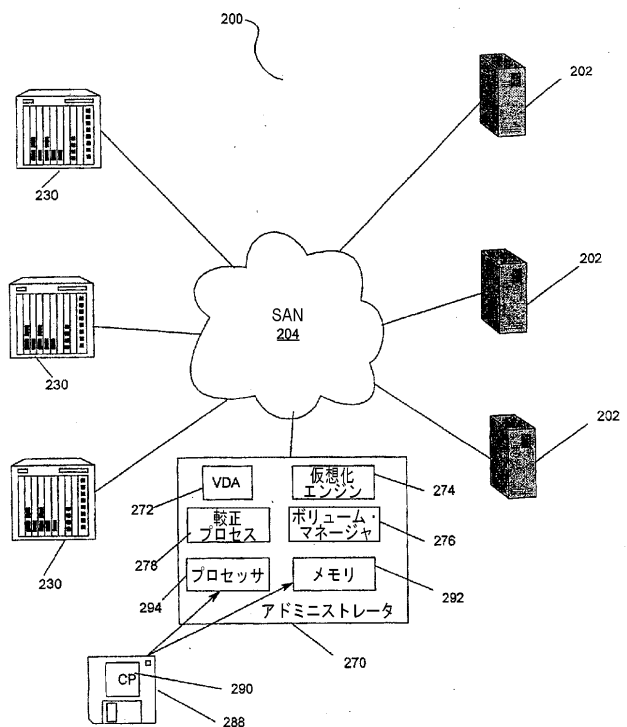
【図6】本発明の実施形態による、仮想化ストレージ・サブシステム中の重要なパフォーマンス要素を抽出するために仮想ディスク・アロケータにより使用されるデータ構造を例示する。

【図7】本発明の実施形態による、仮想ストレージ・システムにおけるストレージ割り当てを管理するための方法の流れ図を例示する。

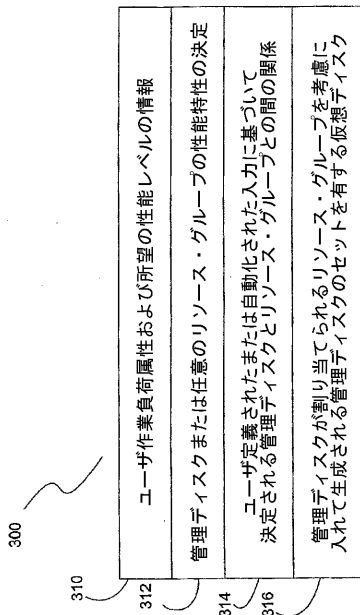
【図1】



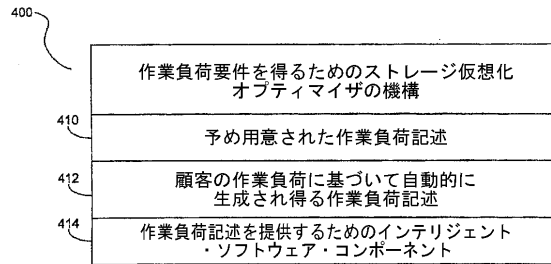
【図2】



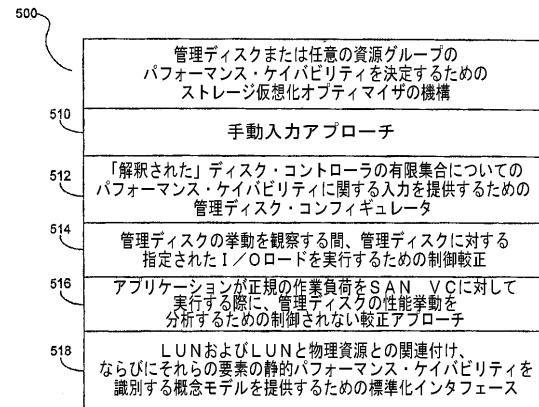
【 図 3 】



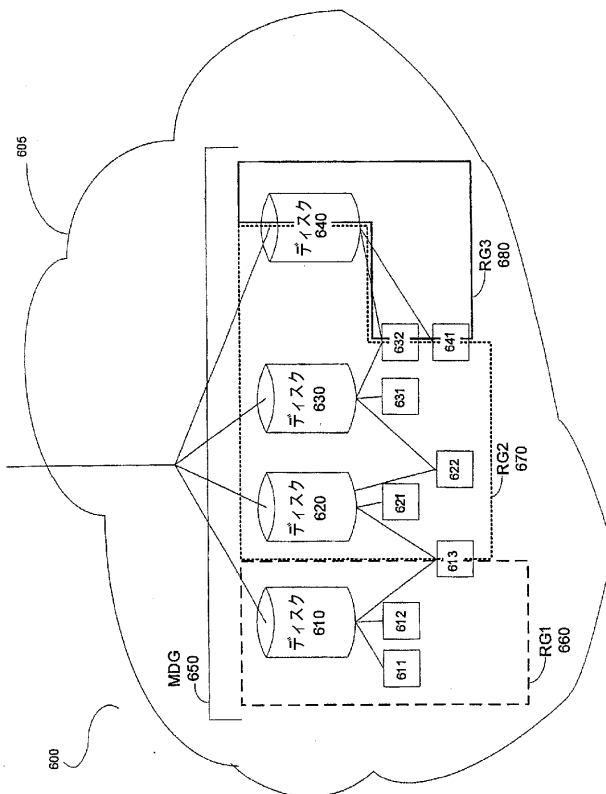
【圖 4】



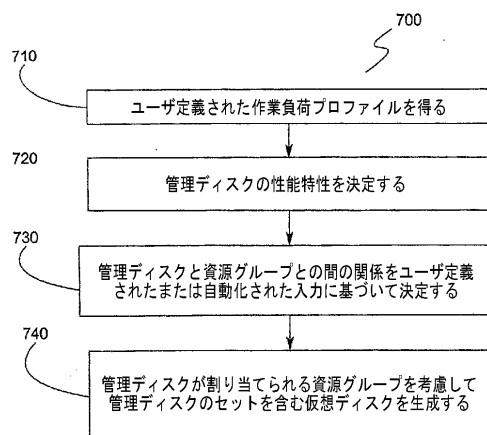
【 図 5 】



【 図 6 】



【圖 7】



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2006/050260

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F9/46

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 820 035 B1 (ZAHAVI WILLIAM) 16 November 2004 (2004-11-16) column 1, line 24 - line 29 column 2, line 5 - line 34 column 7, line 49 - line 53 column 11, line 32 - column 13, line 51	1-44
A	US 2003/233382 A1 (GEMBA RYU ET AL) 18 December 2003 (2003-12-18) the whole document	1-44
A	WO 02/25870 A (STORAGEAPPS INC) 28 March 2002 (2002-03-28) the whole document	1-44

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *I* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *S* document member of the same patent family

Date of the actual completion of the international search

6 April 2006

Date of mailing of the international search report

19/04/2006

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+31-70) 340-3016

Authorized officer

Dewyn, T

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2006/050260

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 6820035	B1	16-11-2004	NONE	
US 2003233382	A1	18-12-2003	JP 2004021557 A	22-01-2004
WO 0225870	A	28-03-2002	AU 3271901 A	02-04-2002

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 アショフ、ジョン

アメリカ合衆国 9 5 0 7 6 カリフォルニア州、ワトソンビル、カイル・キャニオン 3 6 8

(72)発明者 シェリアン、ニーナ

アメリカ合衆国 9 5 1 2 0 カリフォルニア州、サン・ホセ、カジェ・エスペランサ 6 0 9 0

(72)発明者 スミス、ブライアン

アメリカ合衆国 9 4 1 0 7 カリフォルニア州、サンフランシスコ、ブランナン # 1 7 ジー 2
2 9

F ターム(参考) 5B014 EB04 GC31

5B065 BA01 BA06 CC03 EK02 ZA01

5B082 CA19