

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2012年7月26日(26.07.2012)



(10) 国際公開番号

WO 2012/098591 A1

- (51) 国際特許分類:
G06F 12/00 (2006.01)
- (21) 国際出願番号: PCT/JP2011/001680
- (22) 国際出願日: 2011年3月22日(22.03.2011)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2011-007260 2011年1月17日(17.01.2011) JP
- (71) 出願人(米国を除く全ての指定国について): 株式会社日立製作所 (HITACHI, LTD.) [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目6番6号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 金子聡 (KANEKO, Satoshi) [JP/JP]; 〒2440817 神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 システム開発研究所内 Kanagawa (JP). 山本彰(YAMAMOTO, Akira) [JP/JP]; 〒1008220 東京都

千代田区丸の内一丁目6番1号 株式会社日立製作所 研究開発本部内 Tokyo (JP). 柴山 司 (SHIBAYAMA, Tsukasa) [JP/JP]; 〒2440817 神奈川県横浜市戸塚区吉田町292番地 株式会社日立製作所 システム開発研究所内 Kanagawa (JP).

(74) 代理人: 大賀真司, 外(OHGA, Shinji et al.); 〒1050022 東京都港区海岸1-2-20 汐留ビルディング3階サンネクスト国際特許事務所 Tokyo (JP).

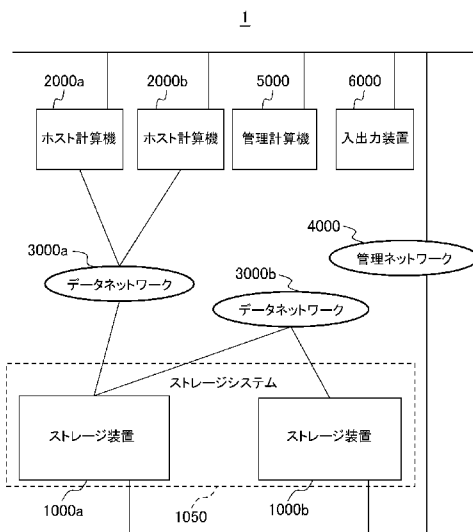
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[続葉有]

(54) Title: COMPUTER SYSTEM, ADMINISTRATION COMPUTER, AND STORAGE ADMINISTRATION METHOD

(54) 発明の名称: 計算機システム、管理計算機およびストレージ管理方法

図1



1000a, 1000b Storage apparatus
1050 Storage system
2000a, 2000b Host computer
3000a, 3000b Data network
4000 Administration network
5000 Administration computer
6000 I/O device

(57) Abstract: A storage apparatus (1000) comprises: a storage device (1100) which stores data which is read and written by a host computer; and a control device (1220) which controls the writing of data to the storage device (1100). The control device (1220) provides prescribed storage regions of the storage device (1100) as one or more volumes to a host computer (2000), and provides statistical information relating to the storage regions to an administration computer (5000) according to a request from the administration computer (5000). The administration computer (5000) comprises: a storage device (5100) which stores a storage region administration table which administers the storage regions of a plurality of the storage apparatuses (1000); and a control device (5200) which administers the configuration of the storage regions of the plurality of storage apparatuses. The control device (5200) administers the data configuration of the plurality of storage apparatuses (1000) based on the statistical information relating to the storage regions of the plurality of storage apparatuses (1000) which is provided from the storage apparatuses (1000).

(57) 要約: ストレージ装置1000は、ホスト計算機により読み書きされるデータを格納する記憶デバイス1100と、記憶デバイス1100へのデータの書き込みを制御する制御デバイス1220と、を備え、制御デバイス1220は、記憶デバイス1100の所定の記憶領域を1つ以上のボリュームとしてホスト計算機2000に提供し、管理計算機5000の要求に応じて、記憶領域に関する統計情報を管理計算機5000に提供し、管理計算機5000は、複数のストレージ装置1000の記憶領域を管理する記憶領域管理テーブルを格納する記憶デバイス5100と、複数のストレージ装置の記憶領域の構成を管理する制御デバイス5200と、を備え、制御デバイス5200は、複数のストレージ装置1000から提供されたストレージ装置1000の記憶領域に関する統計情報を元に、複数のストレージ装置1000のデータ

構成を管理する。



(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

計算機システム、管理計算機およびストレージ管理方法

技術分野

[0001] 本発明は、計算機システム、管理計算機およびストレージ管理方法に関し、ストレージ装置のデータ移行を管理する計算機システム、管理計算機およびストレージ管理方法に適用して好適なるものである。

背景技術

[0002] 近年、計算機システムが扱うデータ量は年々増加する傾向にあり、かかるデータ量の増加に対し、ストレージ装置を新規に追加したり、既存のストレージシステム間でデータを移行したりすることによって対応している。

[0003] ストレージシステム間で移行されるデータの単位は、ストレージ装置にデータを格納するホスト計算機に利用される記憶領域の単位であり、例えば、単一のハードディスク装置（HDD：Hard Disk Drive）から構成される論理ボリュームであったり、複数のハードディスク装置から構成されるRAID（Redundant Arrays of Independent Disks）グループ上の論理ボリュームであったり、シンプロビジョニング（Thin Provisioning）を利用した仮想的な論理ボリューム（以下、これを仮想ボリュームと呼ぶ）などを例示できる。

[0004] ここで、シンプロビジョニングとは、ホスト計算機に対して仮想ボリュームを提供し、ホスト計算機から仮想ボリュームに対するデータの書き込み要求があったときに、データを格納するための物理的な記憶領域を動的に仮想ボリュームに割り当てる機能である。このようなシンプロビジョニング機能によれば、現実提供可能な記憶領域よりも大きな容量の仮想ボリュームをホスト計算機に提供することができる。また、最近では、性能の異なる複数種類のハードディスク装置を用いて効率的にストレージ管理を行うことが行われている。例えば、階層ストレージ管理（HSM：Hierarchical Storage Management）と呼ばれる管理手法では、アクセス頻度の高いデータが格納された

領域に対しては高速および高性能な記憶階層から記憶領域を割り当て、アクセス頻度の低いデータが格納された領域に対しては低速および低性能な記憶階層から記憶領域を割り当てる。

- [0005] また、特許文献 1 には、移行されるデータが格納されている移行元の記憶領域の情報とデータの移行先の記憶領域の情報とを取得し、データを移行するために必要な所要時間を算出して、最適なデータ移行プランを作成する技術が開示されている。例えば、高性能な記憶階層に格納されたデータを、低性能な記憶階層に格納されたデータより優先して移行することにより、アクセス頻度の高いデータを優先的に移行することを可能としている。

先行技術文献

特許文献

- [0006] 特許文献1：特開 2 0 0 8 - 2 0 3 9 3 7 号公報

発明の概要

発明が解決しようとする課題

- [0007] しかし、特許文献 1 では、データ移行の所要時間を算出するために、複数のストレージ装置を管理する管理計算機は、データの移行元となるストレージ装置およびデータの移行先となるストレージ装置から任意の記憶領域に関するすべての情報を取得する必要があった。このため、管理計算機とストレージ装置との間で転送されるデータ量が膨大となり、ストレージシステム全体の性能を劣化させる可能性があるという課題があった。

- [0008] 本発明は以上の点を考慮してなされたもので、ストレージ装置の記憶領域に関する情報を効率よく取得することが可能なストレージシステム、管理計算機およびストレージ管理方法を提案しようとするものである。

課題を解決するための手段

- [0009] かかる課題を解決するために本発明においては、複数のストレージ装置と、前期複数のストレージ装置にデータの書き込みを要求する複数のホスト装置と、前記複数のストレージ装置および複数のホスト装置を管理する管理計

算機とがネットワークを介してそれぞれ相互に接続された計算機システムが提供される。前記ストレージ装置は、前記ホスト計算機により読み書きされるデータを格納する記憶デバイスと、前記記憶デバイスへのデータの書き込みを制御する制御デバイスと、を備え、前記制御デバイスは、前記記憶デバイスの所定の記憶領域を1つ以上のボリュームとして前記ホスト計算機に提供し、前記管理計算機の要求に応じて、前記記憶領域に関する統計情報を前記管理計算機に提供し、前記管理計算機は、前記複数のストレージ装置の記憶領域を管理する記憶領域管理テーブルを格納する記憶デバイスと、前記複数のストレージ装置の前記記憶領域の構成を管理する制御デバイスと、を備え、前記制御デバイスは、前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する。

[0010] かかる構成によれば、ストレージ装置は、前記ストレージ装置の所定の記憶領域を1つ以上のボリュームとして前記ホスト計算機に提供し、前記管理計算機が、前記記憶領域に関する統計情報を前記ストレージ装置に要求し、前記ストレージ装置が、前記管理計算機の要求に応じて、前記記憶領域に関する統計情報を前記管理計算機に提供し、前記管理計算機が、前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する。これにより、ストレージ装置に格納されている記憶領域に関するすべての情報を取得しなくとも、該記憶領域の統計情報を取得してストレージ装置に格納されているデータの移行時間やデータの移行先におけるデータ I/O 性能などを算出することができるため、転送するデータ量や転送時間を大幅に短縮することが可能となる。

発明の効果

[0011] 本発明によれば、ストレージ装置の記憶領域に関する情報を効率よく取得して、ストレージシステム全体の性能劣化を防止することができる。

図面の簡単な説明

[0012] [図1]本発明の第1の実施形態にかかる計算機システムの全体構成を示すブロック図である。

[図2]同実施形態にかかるストレージ装置の構成を示すブロック図である。

[図3]同実施形態にかかるディスクデバイスの構成を示すブロック図である。

[図4]同実施形態にかかるストレージ側物理論理記憶領域対応テーブルの一例を示す図表である。

[図5]同実施形態にかかるストレージ側物理論理記憶領域対応テーブルに格納されている情報の例を示す概念図である。

[図6]同実施形態にかかるリソース構成テーブルの一例を示す図表である。

[図7]同実施形態にかかるストレージ側外部ストレージ構成テーブルの一例を示す図表である。

[図8]同実施形態にかかるホスト計算機の構成を示すブロック図である。

[図9]同実施形態にかかる管理計算機の構成を示すブロック図である。

[図10]同実施形態にかかるリソース管理テーブルの一例を示す図表である。

[図11]同実施形態にかかる管理側外部ストレージ構成テーブルの一例を示す図表である。

[図12]同実施形態にかかる管理側物理論理記憶領域対応テーブルの一例を示す図表である。

[図13]同実施形態にかかるディスクタイプ性能情報テーブルの一例を示す図表である。

[図14]同実施形態にかかるポート性能情報テーブルの一例を示す図表である。

[図15]同実施形態にかかるシステム管理プログラムの処理の処理手順を示すフローチャートである。

[図16]同実施形態にかかるデータ移行所要時間計算プログラムの処理の処理手順を示すフローチャートである。

[図17]同実施形態にかかるデータ移行所要時間算出処理の具体例を示すフローチャートである。

[図18] 同実施形態にかかる統計情報取得プログラムの処理の処理手順を示すフローチャートである。

[図19] 本発明の第2の実施形態にかかる管理計算機の構成を示すブロック図である。

[図20] 同実施形態にかかる管理側物理論理記憶領域対応テーブルの一例を示す図表である。

[図21] 同実施形態にかかる管理側プール物理論理記憶領域対応テーブルの一例を示す図表である。

[図22] 同実施形態にかかる移行先ボリューム性能予測プログラムの処理の処理手順を示すフローチャートである。

[図23] 同実施形態にかかる移行先ボリューム性能予測プログラムの具体例を示すフローチャートである。

発明を実施するための形態

[0013] 以下図面について、本発明の一実施の形態を詳述する。

[0014] なお、以降の説明では、「aaaテーブル」等の表現にて本発明の情報を説明するが、これらの情報は必ずしもテーブルのデータ構造以外で表現されていてもよい。そのため、データ構造に依存しないことを示すために「aaaテーブル」、「aaaリスト」、「aaaDB」、「aaaキュー」等について「aaa情報」と呼ぶことがある。さらに、各情報御の内容を説明する際に、「識別情報」、「識別子」、「名」、「名前」、「ID」という表現を用いるが、これらについてはお互いに置換が可能である。

[0015] (1) 第1の実施の形態

(1-1) 計算機システムの構成

まず、図1を参照して、本実施の形態にかかる計算機システム1の構成について説明する。図1に示すように、本実施の形態にかかる計算機システム1は、ストレージシステム1050と、ホスト計算機2000a、2000b（以降、ホスト計算機2000a、2000bを単にホスト計算機2000と称する場合もある。）と、データネットワーク3000a、3000b

(以降、データネットワーク 3000a、3000bを単にデータネットワーク 3000と称して説明する場合もある。)と、管理ネットワーク 4000と、管理計算機 5000と、入出力装置 6000と、を備えている。

[0016] ストレージシステム 1050は、複数のストレージ装置、例えば、ストレージ装置 1000a、ストレージ装置 1000bから構成される。ストレージ装置 1000a、1000bは、以降の説明では、単にストレージ装置 1000と称する場合もある。また、あるストレージ装置に接続されるストレージ装置と、接続先のストレージ装置とを明示的に区別するために、前者を外部接続元ストレージ装置、後者を外部接続先ストレージ装置と称して説明する場合もある。

[0017] 複数のストレージ装置 1000とホスト計算機 2000とは、データネットワーク 3000aを介して接続されている。また、ストレージシステム 1050を構成する複数のストレージ装置 1000間は、データネットワーク 3000bを介して接続されている。データネットワーク 3000aおよび 3000bは、SAN (Storage Area Network) であっても、IP (Internet Protocol) ネットワークであってもよく、これら以外のデータ通信用ネットワークであってもよい。

[0018] また、ホスト計算機 2000およびストレージ装置 1000は、管理ネットワーク 4000を介して管理計算機 5000と接続されている。管理ネットワーク 4000は、例えば、IPネットワークを例示できるが、SANであっても、これら以外のデータ通信用ネットワークであってもよい。また、データネットワーク 3000a、3000bと管理ネットワーク 4000とは、異なるネットワークでもよいし、同一のネットワークでもよい。また、本実施形態では、ホスト計算機 2000と管理計算機 5000とは、別体の計算機として説明しているが、かかる例に限定されず、ホスト計算機 2000と管理計算機 5000とを一体の計算機として構成してもよい。また、入出力装置 6000は、管理計算機 5000と管理ネットワーク 4000を介して接続されている。

- [0019] ストレージ装置 1000 は、データを格納する記憶デバイスを備えた情報処理装置であって、ホスト計算機 2000 から送信されたコマンドを解釈してストレージ装置 1000 の記憶領域内へのリード／ライトを実行する。ホスト計算機 2000 は、ユーザ業務に応じた各種演算処理を実行するコンピュータ装置であり、例えばパーソナルコンピュータ、ワークステーション又はメインフレーム等から構成される。
- [0020] 管理計算機 5000 は、システム管理者が計算機システム 1 を管理するために利用する情報処理装置であって、例えば、パーソナルコンピュータ、ワークステーション又はメインフレーム等から構成される。入出力装置 6000 は、システム管理者からの入力を受け付けたり、管理計算機 5000 から送信される処理結果等を受信して出力したりする端末装置であって、例えば、キーボード、スイッチ、ポインティングデバイスなどの入力装置と、ディスプレイやスピーカなどの出力装置とから構成される。本実施形態では、管理計算機 5000 と入出力装置 6000 とは別体の装置としたが、かかる例に限定されず、管理計算機 5000 と入出力装置 6000 とを一体の装置として構成してもよい。
- [0021] また、図 1 では 2 つのストレージ装置 1000 から構成される 1 つのストレージシステムと、2 つのホスト計算機 2000 と、1 つの管理計算機 5000 とから計算機システム 1 が構成されているが、各装置の数はこれらに限定されない。
- [0022] ここで、本実施形態の概要について説明する。通常、計算機システム 1 が扱うデータ量が増加した場合に、データ量の増加に対し、ストレージ装置 1000 を新規に追加したり、既存のストレージシステム 1050 間でデータを移行したりすることによって対応している。
- [0023] ストレージシステム 1050 間で移行されるデータの単位は、ストレージ装置 1000 にデータを格納するホスト計算機 2000 に利用される記憶領域の単位であり、例えば、単一のハードディスク装置（HDD : Hard Disk Drive）から構成される論理ボリュームであったり、複数のハードディスク装

置から構成されるRAIDグループ上の論理ボリュームであったり、シンプロビジョニングを利用した仮想ボリュームなどを例示できる。

[0024] ここで、シンプロビジョニングとは、ホスト計算機2000に対して仮想ボリュームを提供し、ホスト計算機2000から仮想ボリュームに対するデータの書き込み要求があったときに、データを格納するための物理的な記憶領域を動的に仮想ボリュームに割り当てる機能である。このようなシンプロビジョニング機能によれば、現実提供可能な記憶領域よりも大きな容量の仮想ボリュームをホスト計算機2000に提供することができる。また、最近では、性能の異なる複数種類のハードディスク装置を用いて効率的にストレージ管理を行うことが行われている。例えば、階層ストレージ管理と呼ばれる管理手法では、アクセス頻度の高いデータが格納された領域に対しては高速および高性能な記憶階層から記憶領域を割り当て、アクセス頻度の低いデータが格納された領域に対しては低速および低性能な記憶階層から記憶領域を割り当てる。

[0025] また、移行されるデータが格納されている移行元の記憶領域の情報と、データの移行先の記憶領域の情報とを取得し、データを移行するために必要な所要時間を算出して、最適なデータ移行プランを作成する技術が開示されている。例えば、高性能な記憶階層に格納されたデータを、低性能な記憶階層に格納されたデータより優先して移行することにより、アクセス頻度の高いデータを優先的に移行することを可能としている。

[0026] しかし、上記技術では、データ移行の所要時間を算出するために、複数のストレージ装置1000を管理する管理計算機5000は、データの移行元となるストレージ装置1000aおよびデータの移行先となるストレージ装置1000bから任意の記憶領域に関するすべての情報を取得する必要があった。このため、管理計算機5000とストレージ装置1000との間で転送されるデータ量が膨大となり、ストレージシステム全体の性能を劣化させる可能性があるという課題があった。本実施形態にかかるストレージシステム1では、ストレージ装置1000の記憶領域に関する情報を大幅に削減し

て、記憶領域に関するすべての情報を得た場合と同等の結果を得ることにより、データの移行処理等を効率化させることを可能としている。

[0027] (1-2) 各装置の構成

(1-2-1) ストレージ装置の構成

次に、ストレージ装置 1000 のハードウェア構成およびソフトウェア構成について説明する。図 2 に示すように、ストレージ装置 1000 は、ストレージ装置 1000 全体を制御するディスクコントローラ 1200 と、データを格納するディスクデバイス 1100 とから構成される。

[0028] ディスクコントローラ 1200 は、メインメモリ 1210 と、制御デバイス 1220 と、ホスト I/F 1230 と、管理 I/F 1240 と、ディスク I/F 1250 と、外部接続 I/F 1260 とを備える。なお、ストレージ装置 1000 が、外部接続先として利用される場合には、外部接続 I/F を備えていなくともよい。

[0029] メインメモリ 1210 には、構成管理プログラム 1211 と、統計情報作成プログラム 1215 と、ストレージ側物理論理記憶領域対応テーブル 1212 と、リソース構成テーブル 1213 と、ストレージ側外部ストレージ構成テーブル 1214 と、が記憶されている。なお、ストレージ装置 1000 が、外部接続先として利用される場合には、ストレージ側外部ストレージ構成テーブル 1214 がメインメモリ 1210 に記憶されていなくともよい。また、ストレージ装置 1000 が、シンプロビジョニング機能に対応していない場合には、ストレージ側物理論理記憶領域対応テーブル 1212 がメインメモリ 1210 に記憶されていなくともよい。各テーブルの詳細については、後述する。

[0030] 構成管理プログラム 1211 は、ストレージ装置 1000 の構成情報や性能情報を管理するプログラムである。構成管理プログラム 1211 は、例えば、ストレージ側物理論理記憶領域対応テーブル 1212、リソース構成テーブル 1213 またはストレージ側外部ストレージ構成テーブル 1214 に格納されている情報を参照したり更新したりする。また、構成管理プログラ

ム 1 2 1 1 は、後述する管理計算機 5 0 0 0 のシステム管理プログラム 5 1 1 0 と通信して、各種情報を送受信する。

[0031] 統計情報作成プログラム 1 2 1 5 は、管理計算機 5 0 0 0 の要求に応じて、ストレージ装置 1 0 0 0 の構成に関する統計情報を作成する機能を有する。ここで、統計情報とは、所定の記憶領域の単位と、所定の統計情報の粒度とに対応する情報であって、例えば、記憶領域の単位として論理ボリュームまたは仮想ボリュームが例示でき、統計情報の粒度としてリソース毎のセグメント割当数や、論理ボリューム内のデータの占有率などを例示できる。統計情報作成プログラム 1 2 1 5 は、管理計算機 5 0 0 0 から与えられた記憶領域の単位および統計情報の粒度に対応する情報を各テーブルから取得する。メインメモリ 1 2 1 0 に格納されている各テーブルについては、後で詳細に説明する。

[0032] 制御デバイス 1 2 2 0 は、演算処理装置として機能し、メインメモリ 1 2 1 0 に記憶されている各種プログラムや演算パラメータ等に当たって、ストレージ装置 1 0 0 0 全体の動作を制御する。

[0033] ホスト I / F 1 2 3 0 は、データネットワーク 3 0 0 0 を介してホスト計算機 2 0 0 0 と接続するためのインタフェースであって、データや制御命令などの各種情報をホスト計算機 2 0 0 0 に送受信する。管理 I / F 1 2 4 0 は、管理ネットワーク 4 0 0 0 を介してホスト計算機 2 0 0 0 や管理計算機 5 0 0 0 に接続するためのインタフェースであって、データや制御命令などの各種情報をホスト計算機 2 0 0 0 や管理計算機 5 0 0 0 に送受信する。

[0034] ディスク I / F 1 2 5 0 は、ディスクコントローラ 1 2 0 0 とディスクデバイス 1 1 0 0 とを接続するインタフェースであって、データや制御命令などの各種情報をディスクデバイス 1 1 0 0 に送受信する。外部接続 I / F 1 2 6 0 は、データネットワーク 3 0 0 0 と外部接続されたストレージ装置 1 0 0 0 とを接続するためのインタフェースであって、データや制御命令などの各種情報を外部接続されたストレージ装置 1 0 0 0 に送受信する。

[0035] ディスクデバイス 1 1 0 0 は、複数の物理デバイス (PDEV) 1 1 5 0

a、1150b、1150c、1150d（以降、物理デバイス（PDEV）1150と称して説明する場合もある。）から構成される。ここで、ディスクデバイス1100の論理的な構成について説明する。

[0036] 図3に示すように、ディスクデバイス1100は、一つ以上の論理ボリューム1110（1110a、1110b、1110c、1110d）とプール1120とを備えている。論理ボリューム1110とは、データの記憶単位となる論理的な記憶領域であって、一つ以上の物理リソース1121または仮想リソース1122から生成される。論理ボリューム1110には、ホスト計算機2000によってリード／ライトされる各種データが格納されている。なお、本実施形態では、ディスクデバイス1100には、4つの論理ボリューム1110と1つのプール1120が備えられているが、論理ボリューム等の数はこれらに限定されない。

[0037] プール1120は、一つ以上の物理リソース1121または一つ以上の仮想リソースから構成されている。プール1120は、ストレージ装置1000がホスト計算機2000に提供する記憶領域の集合であって、例えば、RAIDグループにより提供されるプールでもよいし、シンプロビジョニングにより提供されるプールであってもよい。例えば、RAIDグループにより提供されるプールの場合には、複数のディスクにより1つのRAIDグループが構成され、1つのRAIDグループが提供する記憶領域上に1または複数の論理ボリュームが定義される。そして、複数のRAIDグループが提供する論理ボリュームが1つのプールとして管理される。RAIDグループとしては、例えば、複数のハードディスクデバイスを1つにまとめて巨大な記憶領域として提供する「RAID0」や、ハードディスクデバイスの冗長度を高めるために、ハードディスク間でミラーリングを行う「RAID1」などを例示できる。

[0038] 図3では、プール1120は、物理リソース1121と仮想リソース1122から構成されているが、かかる例に限定されず、プール1120は、物理リソース1121のみから構成されていてもよいし、仮想リソース112

2のみから構成されていてもよい。また本実施形態では、複数の物理デバイス（PDEV）1150から構成されるRAID上の論理デバイスを、物理リソースとして定義している。また、外部接続技術によって、外部接続元のストレージシステム1050内に作成される論理デバイスを、仮想リソースとして定義している。また、物理リソースと仮想リソースとを特に区別する必要がない場合には、単にリソースと称して説明する。

[0039] さらに、図3では、論理ボリューム1110の各種種類を示している。論理ボリューム1110の種類としては、例えば、通常論理ボリューム1110a、外部接続ボリューム1110b、シンプロビジョニングボリューム1110c、ダイナミックシンプロビジョニング1110dなどが挙げられる。

[0040] 通常論理ボリューム1110aは、ハードディスクデバイスから構成される物理リソース1121（RAIDグループ等）から構成される論理的な記憶領域である。また、外部接続ボリューム1110bは、記憶領域の実体が外部接続先（例えば、ストレージ装置1000b）に存在し、仮想リソース1122から構成される論理的な記憶領域である。

[0041] また、シンプロビジョニングボリューム1110cは、上記したように、動的に容量を拡張することが可能な論理ボリュームである。シンプロビジョニングボリューム1110cは、ホスト計算機2000からのデータの書き込み要求を受けた時点で、プール1120に含まれる物理リソース1123または仮想リソース1124からセグメント（所定の領域）を割り当てることにより、動的にボリュームの容量を拡張することができる。

[0042] また、ダイナミックシンプロビジョニングボリューム1110dは、シンプロビジョニングボリューム1110cと同様に、動的に容量を拡張することが可能な論理ボリュームである。ダイナミックシンプロビジョニングボリューム1110dは、一度割り当てられたセグメントを、論理ボリュームアクセス状況に応じて、性能（応答性、信頼性など）の異なるセグメントへ動的に割り当てることができる。

- [0043] なお、図3では、シンプロビジョニングボリューム1110dおよびダイナミックシンプロビジョニングボリューム1110cは、物理リソース1123および仮想リソース1124の両方からセグメントを割り当てることとなっているが、かかる例に限定されず、物理リソース1123または仮想リソース1124のいずれか一方からセグメントを割り当てるようにしてもよい。
- [0044] 次に、図4～図7を参照して、メインメモリ1210に格納されている各テーブルの詳細について説明する。
- [0045] ストレージ側物理論理記憶領域対応テーブル1212は、シンプロビジョニング機能における、論理ボリューム1110の論理アドレスと、物理リソース1121または仮想リソース1122とのアドレスの対応関係を管理するテーブルである。ストレージ側物理論理記憶領域対応テーブル1212は、図4に示すように、ボリューム識別子欄12120、論理開始アドレス欄12121、論理終了アドレス欄12122、プール識別子欄12123、リソース識別子欄12124、物理開始アドレス欄12125および物理終了アドレス欄12126から構成される。
- [0046] ボリューム識別子欄12120には、各論理ボリューム1110を識別するボリューム識別子が格納される。また、論理開始アドレス欄12121には、論理ボリューム1110のセグメントの開始アドレスが格納され、論理終了アドレス欄12122には、論理開始アドレス欄12121に格納された論理開始アドレスに対応する論理ボリューム1110の終了アドレスが格納される。
- [0047] そして、プール識別子欄12123には、セグメントの割当て元であるリソースを有するプールを識別するプール識別子が格納される。リソース識別子欄12124には、プール1120の物理リソース1121または仮想リソース1122の識別子であるリソース識別子が格納される。物理開始アドレス欄12125には、プール1120の物理リソース1120または仮想リソース1122のリソースの開始アドレスが格納され、物理終了アドレス

欄 1 2 1 2 6 には、プール 1 1 2 0 の物理リソース 1 1 2 0 または仮想リソース 1 1 2 2 のリソースの終了アドレスが格納される。

[0048] ここで、ストレージ側物理論理記憶領域対応テーブル 1 2 1 2 に格納されている情報について説明する。

[0049] 図 5 は、図 4 に示すストレージ側物理論理記憶領域対応テーブル 1 2 1 2 に格納されている情報の例を示している。図 5 に示すように、プール識別子「1」のプール 1 1 2 0 の物理リソース 0 の 1 0 0 0 番地から 1 9 9 9 番地までが、論理ボリューム（ボリューム識別子 0）1 1 1 0 の 0 番地から 9 9 9 番地に対応している。同様に、物理リソース「1」の 0 番地から 9 9 9 番地が、論理ボリューム（ボリューム識別子が「0」）1 1 1 0 の 1 0 0 0 番地から 1 9 9 9 番地に対応している。そして、物理リソース「2」の 0 番地から 2 9 9 9 番地が、論理ボリューム（ボリューム識別子が「0」）1 1 1 0 の 2 0 0 0 番地から 3 9 9 9 番地に対応している。また、論理ボリューム 1 1 1 0 の 4 0 0 0 番地から 1 9 9 9 9 番地までは、物理リソースが割り当てられていない状態である。なお、ダイナミックシンプロビジョニングボリューム 1 1 1 0 d では、図 5 に示す対応関係が、動的に変更される。

[0050] 次に、リソース構成テーブル 1 2 1 3 について説明する。リソース構成テーブル 1 2 1 3 は、ストレージ装置 1 0 0 0 のディスクデバイス 1 1 1 0 に備えられているリソースに関する情報を管理するテーブルである。図 6 に示すように、リソース構成テーブル 1 2 1 3 は、プール識別子欄 1 2 1 3 0、リソース識別子欄 1 2 1 3 1、リソース種別欄 1 2 1 3 2 およびリソース構成欄 1 2 1 3 3 から構成される。

[0051] プール識別子欄 1 2 1 3 0 には、各リソースが所属するプールを識別するプール識別子が格納される。リソース識別子欄 1 2 1 3 1 には、各リソースを識別するリソース識別子が格納される。リソース種別欄 1 2 1 3 2 には、各リソースの種別を示す情報が格納される。リソース種別とは、記憶媒体の種別を表し、例えば、F C（Fibre channel）、S S D（Solid State Drive）または S A T A（Serial AT Attachment）などを例示できる。リソース構

成欄 1 2 1 3 3 には、リソースの構成を示す情報が格納される。リソースの構成を示す情報とは、ストレージ装置 1 0 0 0 内部の物理リソースか、ストレージ装置 1 0 0 0 に接続された外部のストレージ装置の仮想リソースかを示す情報である。図 6 では、ストレージ装置 1 0 0 0 内部の物理リソースを「内部」とし、ストレージ装置外部の仮想リソースを「外部」としている。

[0052] 次に、ストレージ側外部ストレージ構成テーブル 1 2 1 4 について説明する。ストレージ側外部ストレージ構成テーブル 1 2 1 4 は、外部接続の接続元となるストレージ装置 1 0 0 0 a（以降、外部接続元ストレージ装置 1 0 0 0 a と称する。）内の仮想リソース 1 1 2 2 と、外部接続の接続先となるストレージ装置 1 0 0 0 b（以降、外部接続先ストレージ装置 1 0 0 0 b と称する。）の論理ボリュームとの対応関係を管理するテーブルである。図 7 に示すように、ストレージ側外部ストレージ構成テーブル 1 2 1 4 は、リソース識別子欄 1 2 1 4 0、ポート識別子欄 1 2 1 4 1、外部ストレージ識別子欄 1 2 1 4 2、外部ストレージ側リソース識別子欄 1 2 1 4 3 および外部ストレージポート識別子欄 1 2 1 4 4 から構成される。

[0053] リソース識別子欄 1 2 1 4 0 には、外部接続元ストレージ装置 1 0 0 0 a の仮想リソース 1 1 2 2 を識別するリソース識別子が格納される。ポート識別子欄 1 2 1 4 1 には、外部接続元ストレージ装置 1 0 0 0 の仮想リソース 1 1 2 2 と、外部接続先ストレージ装置 1 0 0 0 b の論理ボリューム 1 1 1 0 を接続している、外部接続元ストレージ装置 1 0 0 0 a 側のポートを識別するポート識別子が格納される。

[0054] 外部ストレージ識別子欄 1 2 1 4 2 には、外部接続元ストレージ装置 1 0 0 0 a の仮想リソース 1 1 2 2 に対応する外部接続先ストレージ装置 1 0 0 0 b を識別する識別子が格納される。当該識別子は、外部接続先ストレージ装置 1 0 0 0 b を特定可能な情報であれば、どのような形式の情報であってもよい。

[0055] 外部ストレージ側リソース識別子欄 1 2 1 4 3 には、外部接続元ストレージ装置 1 0 0 0 a の仮想リソース 1 1 2 2 と対応する外部接続先ストレージ

装置 1000b の論理ボリューム 1110 を構成するリソースを識別する識別子が格納される。また、外部ストレージポート識別子欄 12144 には、外部接続元ストレージ装置 1000a の仮想リソース 1122 と外部接続先ストレージ装置 1000b の論理ボリューム 1110 を接続している、外部接続元ストレージ装置 1000a のポートを識別するポート識別子が格納される。

[0056] (1-2-2) ホスト計算機の構成

次に、図 8 を参照して、ホスト計算機 2000 のハードウェア構成およびソフトウェア構成について説明する。図 8 に示すように、ホスト計算機 2000 は、メインメモリ 2100 と、制御デバイス 2200 と、ホスト I/F 2300 と、管理 I/F 2400 とを備える。なお、ホスト計算機 2000 には、図示していない、キーボードや表示デバイス等の入出力デバイスが備えられていてもよい。

[0057] メインメモリ 2100 には、各種プログラムや演算パラメータが格納されている。制御デバイス 2200 は、演算処理装置として機能し、メインメモリ 2100 に記憶されている各種プログラムや演算パラメータ等にしたがって、ホスト計算機 2000 全体の動作を制御する。

[0058] ホスト I/F 2300 は、データネットワーク 3000 を介してストレージ装置 1000 と接続するためのインタフェースであって、データや制御命令などの各種情報をストレージ装置 1000 に送受信する。管理 I/F 2400 は、管理ネットワーク 4000 を介してストレージ装置 1000 または管理計算機 5000 と接続するためのインタフェースであって、データや制御命令などの各種情報をストレージ装置 1000 または管理計算機 5000 に送受信する。

[0059] (1-2-3) 管理計算機の構成

次に、図 9 を参照して、管理計算機 5000 のハードウェア構成およびソフトウェア構成について説明する。図 9 に示すように、管理計算機 5000 は、メインメモリ 5100 と、制御デバイス 2200 と、管理 I/F 530

０と、ホストＩ／Ｆ５４００とを備える。

- [0060] メインメモリ５１００には、システム管理プログラム５１１０と、データ移行所要時間計算プログラム５１１２と、統計情報取得プログラム５１１３と、リソース管理テーブル５１５０と、管理側外部ストレージ構成テーブル５１５５と、管理側物理論理記憶領域対応テーブル５１６０と、ディスクタイプ性能情報テーブル５１７０と、ポート性能情報テーブル５１８０と、が記憶されている。
- [0061] システム管理プログラム５１１０は、ストレージシステム１０５０の構成を管理するプログラムであって、具体的には、ストレージ装置１０００の構成情報をストレージ装置１０００の構成管理プログラム１２１１から取得して、取得した情報をもとにリソース管理テーブル５１５０や管理側外部ストレージ構成テーブル５１５５を更新するプログラムである。
- [0062] データ移行所要時間計算プログラム５１１２は、ユーザ操作に応じて、移行対象のボリュームと移行先のボリュームの情報の入力を受け付けて、リソース管理テーブル５１５０、管理側外部ストレージ構成テーブル５１５５、管理側物理論理記憶領域対応テーブル５１６０、ディスクタイプ性能情報テーブル５１７０およびポート性能情報テーブル５１８０を参照して、入力されたボリューム間のデータ移行の所要時間を算出するためのプログラムである。
- [0063] 統計情報取得プログラム５１１３は、ユーザ操作に応じて、統計情報を取得する対象として、任意の記憶領域の単位と、統計情報の粒度の入力を受け付けて、ストレージ装置１０００の統計情報作成プログラム１２１５に入力された情報を提供する。そして、統計情報取得プログラム５１１３は、ストレージ装置１０００の統計情報作成プログラム１２１５から、入力に応じた統計情報を取得する。
- [0064] 次に、リソース管理テーブル５１５０について説明する。リソース管理テーブル５１５０は、管理計算機５０００に管理されている１つ以上のストレージ装置１０００が有するリソースの情報を管理するテーブルである。リソ

ース管理テーブル5150は、図10に示すように、装置識別子欄51500、プール識別子欄51501、リソース識別子欄51502、リソース種別欄51503およびリソース構成欄51504から構成される。

[0065] 装置識別子欄51500には、ストレージ装置1000を識別するストレージ装置の識別子が格納される。プール識別子欄51501には、ストレージ装置1000のプール1120を識別するプール識別子が格納される。リソース識別子欄51502には、各プール内のリソースを識別するリソース識別子が格納される。リソース種別欄51503には、各リソースの種別を示す情報が格納される。リソース種別とは、記憶媒体の種別を表し、例えば、FC、SSD、SATAなどを例示できる。リソース構成欄51504には、リソースの構成を示す情報が格納される。リソースの構成を示す情報とは、ストレージ装置1000内部の物理リソースか、ストレージ装置1000に接続された外部のストレージ装置の仮想リソースかを示す情報である。図10では、ストレージ装置1000内部の物理リソースを「内部」とし、ストレージ装置外部の仮想リソースを「外部」としている。

[0066] 次に、管理側外部ストレージ構成テーブル5155について説明する。管理側外部ストレージ構成テーブル5155は、管理計算機5000に管理され、1つ以上のストレージ装置1000aと、外部接続された外部ストレージ装置1000bとの外部接続の構成を示すテーブルである。管理側外部ストレージ構成テーブル5155は、図11に示すように、ストレージ識別子欄51550、リソース識別子欄51551、ポート識別子欄51552、外部ストレージ識別子欄51553、外部ストレージ側リソース識別子51554および外部ストレージポート識別子欄51555から構成される。

[0067] ストレージ識別子欄51550には、外部接続元のストレージ装置1000aを識別するストレージ識別子が格納される。リソース識別子欄51551には、外部接続元のストレージ装置1000a内の仮想リソース1122を識別するリソース識別子が格納される。ポート識別子欄51552には、

外部接続元のストレージ装置 1000 a 内の仮想リソース 1122 と外部接続先ストレージ装置 1000 b 内の論理ボリューム 1110 を接続している、外部接続元ストレージ装置 1000 a 側のポートを識別するポート識別子が格納される。

[0068] 外部ストレージ識別子欄 51553 には、外部接続元ストレージ装置 1000 a 内の仮想リソース 1122 に対応する外部接続先ストレージ装置 1000 b を識別する外部ストレージ識別子が格納される。外部ストレージ側リソース識別子 51554 には、外部接続元ストレージ装置 1000 a 内の仮想リソース 1122 に対応する外部接続先ストレージ装置 1000 b 内の論理ボリュームを構成するリソースを識別する外部ストレージ側リソース識別子が格納される。外部ストレージポート識別子欄 51555 には、外部接続元ストレージ装置 1000 a 内の仮想リソース 1122 と外部接続先ストレージ装置 1000 b 内の論理ボリューム 1110 を接続している、外部接続元ストレージ装置 1000 b 側のポートを識別する外部ストレージポート識別子が格納される。

[0069] 次に、管理側物理論理記憶領域対応テーブル 5160 について説明する。管理側物理論理記憶領域対応テーブル 5160 は、ストレージ装置 1000 の構成に関する任意の統計情報を管理するテーブルである。具体的には、統計情報取得プログラム 5113 が、ストレージ装置 1000 のストレージ側物理論理記憶領域対応テーブル 1212 の情報を元にした任意の統計情報をストレージ装置 1000 から取得して管理側物理論理記憶領域対応テーブル 5160 に格納する。統計情報とは、例えば、所定の記憶領域の単位と所定の統計情報の粒度に対応する情報であって、例えば、記憶領域の単位として論理ボリュームや仮想ボリュームが挙げられ、統計情報の粒度としてリソース毎のセグメント割当て数や論理ボリューム内のデータの占有率などが挙げられる。図 12 では、仮想ボリュームに割り当てられているリソースのセグメント数を統計情報として格納する例を挙げて説明する。

[0070] 管理側物理論理記憶領域対応テーブル 5160 は、図 12 に示すように、

装置識別子欄 5 1 6 0 0、ボリューム識別子欄 5 1 6 0 1、プール識別子欄 5 1 6 0 2、リソース識別子欄 5 1 6 0 3 および割当済セグメント数 5 1 6 0 4 から構成される。装置識別子欄 5 1 6 0 0 には、ストレージ装置 1 0 0 0 を識別する装置識別子が格納される。ボリューム識別子欄 5 1 6 0 1 には、論理ボリューム 1 1 1 0 を識別するボリューム識別子が格納される。プール識別子欄 5 1 6 0 2 には、セグメントの割当て元であるリソースを有するプールを識別するプール識別子が格納される。リソース識別子欄 5 1 6 0 3 には、プール 1 1 2 0 の物理リソース 1 1 2 1 または仮想リソース 1 1 2 2 の識別子であるリソース識別子が格納される。割当済セグメント数 5 1 6 0 4 には、各ボリュームに割り当てられた任意リソースのセグメント数が格納される。例えば、ストレージ側物理論理記憶領域対応テーブル 1 2 1 2 に格納されている情報のうち、同一種別のリソースから割り当てられているセグメント数を合計した値が格納される。

[0071] 次に、ディスクタイプ性能情報テーブル 5 1 7 0 について説明する。ディスクタイプ性能情報テーブル 5 1 7 0 は、物理リソース 1 1 2 1 または仮想リソース 1 1 2 2 の各リソースのリード（R e a d）／ライト（W r i t e）性能を管理するテーブルである。ディスクタイプ性能情報テーブル 5 1 7 0 は、図 1 3 に示すように、装置識別子欄 5 1 7 0 0、リソース種別欄 5 1 7 0 1、R e a d 性能欄 5 1 7 0 2 および W r i t e 性能欄 5 1 7 0 3 から構成される。

[0072] 装置識別子欄 5 1 7 0 0 には、ストレージ装置 1 0 0 0 を識別する装置識別子が格納される。リソース種別欄 5 1 7 0 1 には、各リソースの種別を示す情報が格納される。リソース種別とは、記憶媒体の種別を表し、例えば、F C、S S D、S A T A などを例示できる。R e a d 性能欄 5 1 7 0 2 には、リソース種別に対応するリード（R e a d）性能が格納される。リード性能は、例えば、「G／S e c」の単位で示されるが、かかる例に限定されず、他の単位で表されてもよい。W r i t e 性能欄 5 1 7 0 3 には、リソース種別に対応するライト（W r i t e）性能が格納される。例えば、「G／S

e c」の単位で示されるが、かかる例に限定されず、他の単位で表されてもよい。

[0073] 次に、ポート性能情報テーブル5180について説明する。ポート性能情報テーブル5180は、ストレージ装置1000に含まれる各ポートの性能を示すIOPS（Input Output Per Second）を管理するテーブルである。ポート性能情報テーブル5180は、図14に示すように、装置識別子欄51800、ポート識別子欄51801および最大IOPS欄51802から構成される。

[0074] 装置識別子欄51800には、ストレージ装置1000を識別する装置識別子が格納される。ポート識別子欄51801には、ストレージ装置1000に含まれる各ポートを識別するポート識別子が格納される。最大IOPS欄51802には、各ポートのIOPSの値が格納される。当該IOPSの値は、カタログやマニュアル等に占められるポートのIOPSの値が格納されてもよい。また、システム管理プログラム5110によりストレージ装置1000から取得されたポートのIOPSの測定値が格納されてもよいし、過去に測定された複数の測定値から傾向を割りだした予測値を格納してもよい。

[0075] 図9に戻り、制御デバイス5200は、演算処理装置として機能し、メインメモリ5100に記憶されている各種プログラムや演算パラメータ等にしたがって、管理計算機5000全体の動作を制御する。

[0076] 管理I/F5300は、管理ネットワーク4000を介してストレージ装置1000およびホスト計算機2000と接続するためのインタフェースであって、データや制御命令などの各種情報をストレージ装置1000およびホスト計算機2000に送受信する。

[0077] ホストI/F5400は、管理ネットワーク4000を介してホスト計算機2000と接続するためのインタフェースであって、データや制御命令などの各種情報をホスト計算機2000に送受信する。

[0078] なお、上記したストレージ装置1000、ホスト計算機2000および管

理計算機 5000 に備えられている各機能は、各装置の制御デバイスが、各装置のメインメモリから読み出して実行する各プログラムによって実現される。各プログラムは、予め各装置のメインメモリに格納されていてもよいし、必要に応じて、各装置が利用可能な媒体を介して、他の装置からメインメモリに格納されてもよい。各装置が利用可能な媒体とは、例えば、読取装置等に着脱可能な記憶媒体や、各装置に接続されるネットワークや、ネットワークを伝搬する搬送波やデジタル信号などの通信媒体などを例示できる。

[0079] (1-3) 各装置の動作の詳細

次に、計算機システム 1 における各装置の動作の詳細について説明する。以下では、特に、本実施形態の特徴的な処理を実行する管理計算機 5000 の動作の詳細について説明する。

[0080] 以降の説明では「プログラム」を主語として説明を行う場合があるが、プログラムは制御デバイスによって実行されることで定められた処理をメインメモリ及びポート（通信制御デバイス）を用いながら行うため、制御デバイスを主語とした説明としてもよい。また、プログラムを主語として開示された処理は管理計算機等の計算機、情報処理装置が行う処理としてもよい。また、プログラムの一部または全ては専用のハードウェアによって実現されてもよい。また、各種プログラムはプログラム配布サーバや、計算機が読み取り可能な記憶メディア（computer-readable memory media）によって各計算機にインストールされてもよい。

[0081] まず、システム管理プログラム 5110 の処理の詳細について説明する。システム管理プログラム 5110 は、以下の処理を定期的に行ってもよいし、ユーザ操作に応じて実行してもよい。システム管理プログラム 5110 は、ストレージ装置 1000 から構成情報を取得して、取得した構成情報を所定のテーブルに格納する処理を行う。

[0082] 具体的には、システム管理プログラム 5110 は、図 15 に示すように、ストレージ装置 1000 の構成管理プログラム 1212 に対して、ストレージ装置 1000 内のリソースおよび外部ストレージの構成情報を要求する（

S 1 0 1 0)。

[0083] そして、システム管理プログラム 5 1 1 0 は、構成管理プログラム 1 2 1 2 により取得されたストレージ装置 1 0 0 0 内のリソース構成テーブル 1 2 1 3 およびストレージ側外部ストレージ構成テーブル 1 2 1 4 の情報を取得する (S 1 0 2 0)。

[0084] そして、システム管理プログラム 5 1 1 0 は、ステップ S 1 0 2 0 で取得して構成情報を、リソース管理テーブル 5 1 5 0 および管理側外部ストレージ構成テーブル 5 1 5 5 に格納する (S 1 0 3 0)。具体的に、システム管理プログラム 5 1 1 0 は、各ストレージ装置 1 0 0 0 の装置識別子とリソース構成情報とを対応付けてリソース管理テーブル 5 1 5 0 に格納する。また、システム管理プログラム 5 1 1 0 は、各ストレージ装置 1 0 0 0 の装置識別子 (ストレージ識別子) と外部ストレージの構成情報とを対応付けて管理側外部ストレージ構成テーブル 5 1 5 5 に格納する。

[0085] 次に、データ移行所要時間計算プログラム 5 1 1 2 の処理の詳細について説明する。データ移行所要時間計算プログラム 5 1 1 2 は、システム管理者からの入力に応じて以下の処理を実行してもよいし、ストレージ装置 1 0 0 0 の論理ボリュームに対するリソースの割り当て状況等を監視する監視プログラム (図示せず) の要求に応じて以下の処理を実行してもよい。

[0086] 図 1 6 に示すように、データ移行所要時間計算プログラム 5 1 1 2 は、システム管理者からの入力等に応じてデータ移行にかかる所要時間を算出する処理を行う。具体的には、まず、データ移行所要時間計算プログラム 5 1 1 2 は、入力装置 6 0 0 0 を介して管理者により入力されたデータ移行のための情報を取得する (S 2 0 0 5)。データ移行のための情報とは、例えば、データ移行の対象となるボリューム (論理ボリューム) の情報やデータの移行先のボリュームの情報などである。ボリュームの情報としては、例えば、図 1 0 または図 1 1 に示したリソースを識別する情報などである。

[0087] そして、データ移行所要時間計算プログラム 5 1 1 2 は、統計情報取得プログラム 5 1 1 3 に対して、所定の入力に応じた統計情報を要求する (S 2

010)。具体的に、データ移行所要時間計算プログラム5112は、統計情報取得プログラム5113にステップS2005で取得した移行元ボリュームの情報および移行先のボリュームの情報と、統計情報の単位と、統計情報の粒度および種類とを統計情報取得プログラム5113に提供して、統計情報取得プログラム5113から入力に応じた統計情報を取得する。ここで、統計情報の単位とは、例えば仮想ボリューム単位を例示でき、統計情報の粒度または種類とは、リソース毎のセグメント割当数を例示できる。

[0088] そして、データ移行所要時間計算プログラム5112は、ステップS2010で取得した指定したボリュームのリソース割り当てに関する統計情報を取得する（S2020）。なお、本実施形態では、取得する統計情報をボリュームのリソース割り当てに関する情報としたが、かかる例に限定されず、ステップS2010での入力に応じて変化する情報である。統計情報取得プログラム5113による統計情報の取得処理については後で詳細に説明する。

[0089] 次に、データ移行所要時間計算プログラム5112は、ステップS2020で取得した統計情報を、管理側物理論理記憶領域対応テーブル5160に格納する（S2030）。管理側物理論理記憶領域対応テーブル5160には、上記したように、ストレージ側物理論理記憶領域対応テーブル1212に格納されている情報のうち、同一種別のリソースから割り当てられているセグメント数を合計した値が格納される。これにより、ストレージ側物理論理記憶領域対応テーブル1212に格納されているすべての情報よりも少ない情報をまとめて取得することができるため、転送するデータ量や転送時間を大幅に短縮することが可能となる。

[0090] そして、データ移行所要時間計算プログラム5112は、リソース管理テーブル5150、管理側外部ストレージ構成テーブル5155、管理側物理論理記憶領域対応テーブル5160、ディスクタイプ性能情報テーブル5170、ポート性能情報テーブル5180に格納されている情報を元に、データ移行の所要時間を算出する（S2040）。

- [0091] ここで、データ移行所要時間算出処理の具体例について説明する。図 17 に示すように、データ移行所要時間計算プログラム 5112 は、管理側物理論理記憶領域対応テーブル 5160 を参照して、ステップ S2005 において入力されたデータ移行の対象となるボリュームを構成するリソースを特定する (S2050)。
- [0092] そして、データ移行所要時間計算プログラム 5112 は、リソース管理テーブル 5150 を参照して、ステップ S2050 で特定したリソースの種別と構成を特定する (S2052)。
- [0093] そして、データ移行所要時間計算プログラム 5112 は、ステップ S2052 において特定した構成情報が外部か否かを判定する (S2054)。ステップ S2054 において、構成情報が外部であると判定された場合には、データ移行所要時間計算プログラム 5112 は、管理側外部ストレージ構成テーブル 5155 を参照して、特定したリソースに対応する、外部接続に用いられるポートを特定する (S2056)。さらに、データ移行所要時間計算プログラム 5112 は、ポート性能情報テーブル 5180 を参照して、ステップ S2056 において特定したポートの性能情報を特定する (S2058)。
- [0094] ステップ S2054 において、構成情報が外部ではないと判定された場合には、ステップ S2060 の処理を実行する。そして、データ移行所要時間計算プログラム 5112 は、ディスクタイプ性能情報テーブル 5170 を参照して、ステップ S2050 において特定したリソースの種別に対応するリード／ライト性能を特定する (S2060)。
- [0095] 次に、データ移行所要時間計算プログラム 5112 は、ステップ S2050 において特定したリソースに格納されたデータのデータ移行所要時間を計算する (S2062)。例えば、移行元ボリュームが 200GB の SSD および 800GB の FC である場合には、ステップ S2060 において SSD のリード性能が 10 (G/Sec)、FC のリード性能が 8 (G/Sec) と特定される。したがって、移行元ボリュームのデータ移行所要時間は、次

式により算出される。

[0096] [数1]

$$200/10 + 800/8 = 120(\text{Sec}) \quad \dots(1)$$

[0097] 移行先ボリュームが1000GBのFCでの仮想リソースで構成されている場合には、ステップS2058において、ポート性能が特定される。ポート性能の最大IOPSが10 (G/sec)であった場合には、10/5 (G/sec)をポート性能としてもよい。また、ステップS2060において、FCのライト性能が4 (G/Sec)と特定される。したがって、移行先ボリュームへのデータ移行所要時間は、 $1000/4 + 1000/5 = 250 + 200 = 450$ (sec)となる。よって、データ移行所要時間計算プログラム5112は、移行元ボリュームのデータ移行所要時間と移行先ボリュームのデータ移行所要時間とを加算した値 ($120 + 450 = 570$ (sec))をデータ移行所要時間とする。

[0098] そして、データ移行所要時間計算プログラム5112は、すべてのリソースについてデータ移行所要時間を算出したか否かを判定する (S2064)。ステップS2064において、すべてのリソースについてデータ移行所要時間を算出したと判定された場合には、データ移行所要時間計算プログラム5112は、処理を終了する。一方、ステップS2064において、すべてのリソースについてデータ移行所要時間を算出していないと判定された場合には、データ移行所要時間計算プログラム5112は、ステップS2050以降の処理を繰り返す。

[0099] 次に、統計情報取得プログラム5113の処理の詳細について説明する。統計情報取得プログラム5113は、ストレージ装置1000の統計情報作成プログラム1215に対して所定の統計情報を要求する。統計情報取得プログラム5113は、上記したデータ移行所要時間計算プログラム5112の指示のもと、以下の統計情報取得処理を実行する。

[0100] 図18に示すように、まず、統計情報取得プログラム5113は、データ

移行所要時間計算プログラム 5 1 1 2 から、統計情報取得対象の情報の入力を受け付ける（S 3 0 0 5）。統計情報取得対象の情報とは、例えば、統計情報取得対象の単位と、統計情報取得対象を特定する情報と、統計情報の粒度または種類などである。具体的に、統計情報取得対象の単位として仮想ボリュームを例示でき、統計情報取得対象を特定する情報として移行元または移行先の論理ボリュームを例示でき、統計情報の粒度または種類の情報としてリソース毎のセグメント割当数を例示できる。

[0101] そして、統計情報取得プログラム 5 1 1 3 は、ステップ S 3 0 0 5 において受け付けた統計情報取得対象の情報を、ストレージ装置 1 0 0 0 の統計情報作成プログラム 1 2 1 5 に提供し、統計情報作成プログラム 1 2 1 5 に対して、統計情報を要求する（S 3 0 1 0）。そして、統計情報取得プログラム 5 1 1 3 は、統計情報作成プログラム 1 2 1 5 から、統計情報取得対象の情報に対応する統計情報を取得する（S 3 0 2 0）。具体的に、統計情報作成プログラム 1 2 1 5 は、例えば、移行元または移行先の論理ボリュームのうち、対象となる仮想ボリュームについて、仮想ボリュームに割り当てられているセグメント数を、リソース毎に加算した値を統計情報として作成して、統計情報取得プログラム 5 1 1 3 に提供する。

[0102] なお、ステップ S 3 0 1 0 において、統計情報取得プログラム 5 1 1 3 は、前記取得した時刻を統計情報作成プログラムへ送信してもよい。この場合、統計情報取得プログラム 5 1 1 3 は、ステップ S 3 0 2 0 で統計情報作成プログラムから取得した統計情報の範囲内で、前回取得した時刻から構成が変更された箇所についてのみ統計情報を取得するようにしてもよい。かかる方法によれば、管理計算機 5 0 0 0 とストレージ装置 1 0 0 0 との間のデータ転送量をさらに削減することができる。

[0103] （２）第２の実施の形態

次に、図 1 9 ～図 2 2 を参照して、本発明の第２の実施の形態について説明する。第１の実施の形態では、ストレージ装置 1 0 0 0 から管理計算機 5 0 0 0 に転送される統計情報として、仮想ボリュームを構成するセグメント

割当情報に関するデータを例示したが、第2の実施の形態では、セグメント割当情報に加えて、セグメントのIOPSの情報、プール内のリソースのセグメント割当情報と、セグメントのIOPSの情報を転送する点で第1の実施の形態と異なっている。以下では、第1の実施の形態と異なる構成について特に詳細に説明し、第1の実施の形態と同様の構成の詳細な説明について省略する。

[0104] 図19に示すように、本実施形態にかかる管理計算機5000は、メインメモリ5100と、制御デバイス2200と、管理I/F5300と、ホストI/F5400とを備える。

[0105] 制御デバイス2200と、管理I/F5300と、ホストI/F5400は、第1の実施形態と同様の機能を有するため、詳細な説明は省略する。

[0106] メインメモリ5100には、システム管理プログラム5110と、移行先ボリューム性能予測プログラム5114と、統計情報取得プログラム5113と、リソース管理テーブル5150と、管理側外部ストレージ構成テーブル5155と、管理側物理論理記憶領域対応テーブル5160Bと、ディスクタイプ性能情報テーブル5170と、ポート性能情報テーブル5180と、管理側プール物理論理記憶領域対応テーブル5190と、が記憶されている。本実施形態では、第1の実施の形態において説明したデータ移行所要時間計算プログラム5112がメインメモリ5100に格納されていないが、データ移行所要時間計算プログラム5112がメインメモリ5100に格納されていてもよい。

[0107] 以下では、第1の実施の形態と異なるプログラムおよびテーブルについて特に詳細に説明する。移行先ボリューム性能予測プログラム5114は、任意のダイナミックシンプロビジョニングボリューム（以降、動的仮想ボリュームと称する）のデータを異なるプールに属する動的仮想ボリュームに移行した場合の性能（例えばレスポンスタイム）を予測するプログラムである。なお、本実施の形態では、移行先ボリューム性能予測プログラム5114の性能予測の対象となるボリュームを動的仮想ボリュームとしたが、かかる例

に限定されず、第1の実施の形態で説明した論理ボリューム1110は、例えば、通常論理ボリューム1110a、外部接続ボリューム1110bであってもよい。また、本実施の形態では、動的仮想ボリュームのデータ移行のルールは、IOPSが高いデータほど性能の高いリソースのセグメントに移行するルールとしているが、かかる例に限定されない。ここで、性能とは、マニュアルやカタログ等で示される値（G/sec）を例示できる。

[0108] 次に、管理側物理論理記憶領域対応テーブル5160Bについて説明する。図20に示すように、管理側物理論理記憶領域対応テーブル5160Bは、ストレージ装置1000の構成に関する任意の統計情報を管理するテーブルである。管理側物理論理記憶領域対応テーブル5160Bは、第1の実施の形態にかかる管理側物理論理記憶領域対応テーブル5160に格納されている情報に加えて、任意の割当済セグメントへの平均IOPSの情報を保持する。

[0109] 具体的に、平均IOPS欄51605には、ボリュームに割り当てられた任意のリソースのセグメントに対するIOPSの値の平均値が格納される。なお、本実施の形態では、管理側物理論理記憶領域対応テーブル5160Bに保持されるIOPSの値を割当セグメントの平均IOPSの値としたが、例えば、50パーセントタイルのようなIO分布を示す統計情報を格納してもよい。

[0110] 次に、管理側プール物理論理記憶領域対応テーブル5190について説明する。管理側プール物理論理記憶領域対応テーブル5190は、ストレージ装置1000のプール1120内のリソースに関する統計情報を管理するテーブルであって、図21に示すように、管理側プール物理論理記憶領域対応テーブル5190は、装置識別子欄51900、プール識別子欄51901、リソース識別子欄51902、総セグメント数欄51903、割当済セグメント数欄51904、割当済セグメントへの平均IOPS欄51905から構成される。

[0111] 装置識別子欄51900には、ストレージ装置を識別する装置識別子が格

納される。プール識別子欄 5 1 9 0 1 には、プールと識別するプール識別子が格納される。リソース識別子欄 5 1 9 0 2 には、リソースを識別するリソース識別子欄が格納される。総セグメント数欄 5 1 9 0 3 には、リソースが有する総セグメント数が格納される。割当済セグメント数欄 5 1 9 0 4 には、リソースが有するセグメントのうち、割当済のセグメント数が格納される。割当済セグメントへの平均 I O P S 欄 5 1 9 0 5 には、任意のリソースの割当済みセグメントへの平均 I O P S の値が格納される。

[0112] 次に、本実施の形態にかかる管理計算機 5 0 0 の動作の詳細について説明する。以下では、特に、第 1 の実施形態とは異なる本実施形態の特徴的な処理について説明する。以下では特に、移行先ボリューム性能予測プログラム 5 1 1 4 の処理の詳細について説明する。

[0113] 図 2 2 に示すように、移行先ボリューム性能予測プログラム 5 1 1 4 は、まず、入力装置 6 0 0 0 を介して管理者により入力されたデータ移行のための情報を取得する (S 4 0 1 0)。データ移行のための情報とは、例えば、データ移行の対象となるボリューム (論理ボリューム) の情報やデータの移行先のボリュームの情報などである。

[0114] 次に、移行先ボリューム性能予測プログラム 5 1 1 4 は、統計情報取得プログラム 5 1 1 3 に対して、所定の入力に応じた統計情報を要求する (S 4 0 2 0)。具体的に、移行先ボリューム性能予測プログラム 5 1 1 4 は、ステップ S 4 0 1 0 で取得した移行元ボリュームの情報および移行先のボリュームの情報と、統計情報の単位と粒度および種類を統計情報取得プログラム 5 1 1 3 に提供して、統計情報取得プログラム 5 1 1 3 から入力に応じた統計情報を取得する。ここで統計情報の単位とは、仮想ボリュームおよびプールであり、統計情報の粒度および種類とは、リソース毎のセグメント割当数および平均 I O P S である。なお、本実施の形態では、リソース毎のセグメント割当数および平均 I O P S を統計情報の粒度および種類として例示したが、5 0 パーセントタイルのような粒度および種類であってもよい。

[0115] そして、移行先ボリューム性能予測プログラム 5 1 1 4 は、ステップ S 2

010で取得した指定したボリュームのリソース割り当てに関する統計情報を取得する（S4030）。なお、本実施形態では、取得する統計情報をボリュームのリソース割り当ておよび平均IOPSに関する情報としたが、かかる例に限定されない。

[0116] 次に移行先ボリューム性能予測プログラム5114は、ステップS4030で取得した統計情報を、管理側物理論理記憶領域対応テーブル5160Bおよび管理側プール物理論理記憶領域対応テーブル5190に格納する（S4040）。管理側物理論理記憶領域対応テーブル5160Bには、上記したように、ストレージ側物理論理記憶領域対応テーブル1212に格納されている情報のうち、同一種別のリソースから割り当てられているセグメント数を合計した値と各リソースの平均IOPSが格納される。これにより、ストレージ側物理論理記憶領域対応テーブル1212に格納されているすべての情報よりも少ない情報をまとめて取得することができるため、転送するデータ量や転送時間を大幅に短縮することが可能となる。

[0117] そして、移行先ボリューム性能予測プログラム5114は、リソース管理テーブル5150、管理側外部ストレージ構成テーブル5155、管理側物理論理記憶領域対応テーブル5160B、ディスクタイプ性能情報テーブル5170、ポート性能情報テーブル5180、管理側プール物理論理記憶領域対応テーブル5190に格納されている情報を元に、移行先の動的仮想ボリュームの性能を算出する（S4050）。

[0118] ここで、移行先ボリューム性能予測処理の具体例について説明する。図23に示すように、移行先ボリューム性能予測プログラム5114は、管理側物理論理記憶領域対応テーブル5160を参照して、ステップS4010において入力された移行元ボリュームを構成するリソースおよび移行先ボリュームに対応するプールを構成するリソースを特定する（S4060）。

[0119] そして、移行先ボリューム性能予測プログラム5114は、リソース管理テーブル5150を参照して、ステップS4060で特定したリソースの種類と構成を特定する（S4062）。

- [0120] そして、移行先ボリューム性能予測プログラム5114は、ステップS4062において特定した構成情報が外部か否かを判定する（S4064）。ステップS4064において、構成情報が外部であると判定された場合には、移行先ボリューム性能予測プログラム5114は、管理側外部ストレージ構成テーブル5155を参照して、特定したリソースに対応する、外部接続に用いられるポートを特定する（S4066）。さらに、移行先ボリューム性能予測プログラム5114は、ポート性能情報テーブル5180を参照して、ステップS4066において特定したポートの性能情報を特定する（S4068）。
- [0121] ステップS4064において、構成情報が外部ではないと判定された場合には、ステップS4070の処理を実行する。そして、移行先ボリューム性能予測プログラム5114は、ディスクタイプ性能情報テーブル5170を参照して、ステップS4060において特定したリソースの種別に対応するリード／ライト性能を特定する（S4070）。
- [0122] 次に、移行先ボリューム性能予測プログラム5114は、ステップS4060において特定したリソースの移行における移行先ボリュームの性能を計算する（S4072）。例えば、移行元ボリュームがリソース1の1000セグメント（平均IOPS=5000）と、リソース2の1000セグメント（平均IOPS=6000）と、リソース3の2000セグメント（平均IOPS3000）とから構成されているとする。そして、移行先ボリュームに対応するプールがリソース10の全セグメント（1000セグメント）が割当済み（平均IOPS=4000）で、リソース11の全セグメント（1000セグメント）が割当済み（平均IOPS=3000）で、リソース2の全8000セグメントのうち2000セグメントが割当済（平均IOPS=2500）と特定できたとする。
- [0123] また、上記したように、動的仮想ボリュームのデータ移行のルールは、IOPSが高いデータほど性能の高いリソースのセグメントに移行されるため、移行元ボリュームのリソース2上のデータ、移行元ボリュームのリソース

1上のデータ、移行先のリソース10上のデータ・・・、の順で高性能のリソースが割り当てられると推測できる。したがって、この推測結果に基づいて、移行先において割り当てられるリソースの割合とメディアの性能とから移行先におけるボリュームの性能を算出することができる。

[0124] 例えば、移行先のストレージ装置1000bにおける性能の異なる複数のリソースにデータが格納される割合に、各リソースの平均性能を乗じることにより、レスポンス性能を算出することができる。具体的に、移行先のストレージ装置bにおいて、1TB（1024GB）のデータ量のうち、SSDに20%、FCに60%、SATAに20%のデータが格納されると推測されたとする。そして、各リソースの平均レスポンス性能が、SSDが1（Sec）、FCが2（Sec）、SATAが3（Sec）であったとする。この場合、移行先のストレージ装置1000bの全体の平均レスポンス性能は、次式により算出される。

[0125] [数2]

$$1 \times (20/100) + 2 \times (60/100) + 3 \times (20/100) = 2.0(\text{Sec}) \quad \dots(2)$$

[0126] また、本実施の形態では、全セグメント毎のIOPSの情報をストレージ装置1000から取得した方が、ボリュームの性能をより高い精度で見積もることができるが、移行先ボリューム性能予測プログラム5114は、ステップS4020において統計情報の粒度をより細かく設定することにより、より詳細度の高いIOPSの統計情報を取得することができる。統計情報の粒度をより細かく設定するとは、例えば、IOPSが0～1000のセグメント数、1001～2000までのセグメント数・・・などを例示できる。これにより、移行先ボリューム性能予測プログラム5114は、データ転送量と統計情報の精度を調整することが可能となる。

[0127] なお、ステップS4020で前回取得した時刻を統計情報作成プログラム1215へ送信し、ステップS4030で統計情報作成プログラム1215から取得した統計情報の範囲内で、前回取得した時刻から構成が変更された

箇所についてのみ統計情報を取得するようにしてもよい。かかる方法によれば、管理計算機５０００とストレージ装置１０００との間のデータ転送量をさらに削減することができる。

[0128] （４）他の実施の形態

なお、上述の実施形態においては、管理計算機５０００に格納されている各種プログラムに基づいて、管理計算機５０００の制御デバイス５２００が本発明のバックアップ容量算出部、移行元データ決定部、移行先ストレージ装置決定部およびデータ移行指示部などの各種機能を実現しているが、かかる例に限定されない。例えば、制御デバイス５２００を管理計算機５０００とは別体の他の装置に設けて、当該ＣＰＵと協同して各種機能を実現するようにしてもよい。また、管理計算機５０００に格納されている各種プログラムを管理計算機５０００とは別体の他の装置に設けて、当該プログラムが制御デバイス５２００に呼び出されることにより各種機能を実現するようにしてもよい。

[0129] また、例えば、本明細書の管理計算機５０００等の処理における各ステップは、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はない。すなわち、管理計算機５０００等の処理における各ステップは、異なる処理であっても並列的に実行されてもよい。

[0130] また、管理計算機５０００等に内蔵されるＣＰＵ、ＲＯＭおよびＲＡＭなどのハードウェアを、上述した管理計算機５０００等の各構成と同等の機能を発揮させるためのコンピュータプログラムも作成可能である。また、該コンピュータプログラムを記憶させた記憶媒体も提供される。

産業上の利用可能性

[0131] 本発明は、ストレージ装置のデータ移行を管理する計算機システム、管理計算機およびストレージ管理方法に適用することができる。

符号の説明

[0132] １０００ ストレージ装置
 １１００ ディスクデバイス

1 2 1 0	メインメモリ
1 2 1 1	構成管理プログラム
1 2 1 2	ストレージ側物理論理記憶領域対応テーブル
1 2 1 3	リソース構成テーブル
1 2 1 4	ストレージ側外部ストレージ構成テーブル
1 2 1 5	統計情報作成プログラム
1 2 2 0	制御デバイス
1 2 3 0	ホスト I / F
1 2 4 0	管理 I / F
1 2 5 0	ディスク I / F
1 2 6 0	外部接続 I / F
2 0 0 0	ホスト計算機
2 1 0 0	メインメモリ
2 2 0 0	制御デバイス
2 3 0 0	ホスト I / F
2 4 0 0	管理 I / F
3 0 0 0	データネットワーク
4 0 0 0	管理ネットワーク
5 0 0 0	管理計算機
5 1 0 0	メインメモリ
5 1 1 0	システム管理プログラム
5 1 1 2	データ移行所要時間計算プログラム
5 1 1 3	統計情報取得プログラム
5 1 5 0	リソース管理テーブル
5 1 5 5	管理側外部ストレージ構成テーブル
5 1 6 0	管理側物理論理記憶領域対応テーブル
5 1 7 0	ディスクタイプ性能情報テーブル
5 1 8 0	ポート性能情報テーブル

5 2 0 0 制御デバイス

5 3 0 0 管理 I / F

6 0 0 0 入出力装置

請求の範囲

[請求項1]

複数のストレージ装置と、前記複数のストレージ装置にデータの書き込みを要求する複数のホスト装置と、前記複数のストレージ装置および複数のホスト装置を管理する管理計算機とがネットワークを介してそれぞれ相互に接続された計算機システムであって、

前記ストレージ装置は、

前記ホスト計算機により読み書きされるデータを格納する記憶デバイスと、

前記記憶デバイスへのデータの書き込みを制御する制御デバイスと、を備え、

前記制御デバイスは、

前記記憶デバイスの所定の記憶領域を1つ以上のボリュームとして前記ホスト計算機に提供し、前記管理計算機の要求に応じて、前記記憶領域に関する統計情報を前記管理計算機に提供し、

前記管理計算機は、

前記複数のストレージ装置の記憶領域を管理する記憶領域管理テーブルを格納する記憶デバイスと、

前記複数のストレージ装置の前記記憶領域の構成を管理する制御デバイスと、を備え、

前記制御デバイスは、

前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する

ことを特徴とする、計算機システム。

[請求項2]

前記管理計算機は、前記ストレージ装置に、前記記憶領域に関する統計情報の単位、種類および／または粒度を提供し、

前記ストレージ装置は、前記管理計算機に提供された前記記憶領域に関する統計情報の単位、種類、粒度に基づいて、前記記憶領域に関

する統計情報を生成する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項3] 前記ストレージ装置の制御デバイスは、前記ホスト計算機による書き込み要求に応じて、前記ボリュームのデータ格納領域に前記記憶領域のうち未割当領域を割り当て、

前記管理計算機の制御デバイスは、前記ストレージ装置の前記ボリュームに格納されているデータ構成を管理する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項4] 前記ストレージ装置の制御デバイスは、既に前記記憶領域が割り当てられ散る前記ボリュームに格納されている前記データを、当該データへのアクセス状況に応じて、性能の異なる前記記憶領域へ割り当てを動的に変更し、

前記管理計算機の制御デバイスは、前記ストレージ装置の前記ボリュームに格納されているデータ構成を管理する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項5] 前記管理計算機の制御デバイスは、
前記ストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記ストレージ装置の前記ボリュームに格納されたデータを別のボリュームに移行する際の、データ移行所要時間を算出する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項6] 前記管理計算機の制御デバイスは、
前記ストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記ストレージ装置の前記ボリュームに格納されたデータを別のボリュームに移行する際の、移行先の前記別のボリュームのデータアクセス性能を算出する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項7] ストレージ装置の制御デバイスは、

前記記憶デバイスの前記記憶領域の構成情報や性能情報の更新履歴を管理し、

前記管理計算機の要求に応じて、前記更新履歴に基づく前記記憶領域に関する統計情報を前記管理計算機に提供し、

前記管理計算機の制御デバイスは、

前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する

ことを特徴とする、請求項 1 に記載の計算機システム。

[請求項 8]

複数のストレージ装置、前記複数のストレージ装置にデータの書き込みを要求する複数のホスト装置と、ネットワークを介してそれぞれ相互に接続された管理計算機であって、

前記ストレージ装置は、前記ストレージ装置の所定の記憶領域を 1 つ以上のボリュームとして前記ホスト計算機に提供し、前記管理計算機の要求に応じて、前記記憶領域に関する統計情報を前記管理計算機に提供し、

前記管理計算機は、

前記複数のストレージ装置の記憶領域を管理する記憶領域管理テーブルを格納する記憶デバイスと、

前記複数のストレージ装置の前記記憶領域の構成を管理する制御デバイスと、を備え、

前記制御デバイスは、

前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する

ことを特徴とする、管理計算機。

[請求項 9]

複数のストレージ装置と、前記複数のストレージ装置にデータの書き込みを要求する複数のホスト装置と、前記複数のストレージ装置お

よび複数のホスト装置を管理する管理計算機とがネットワークを介してそれぞれ相互に接続された計算機システムにおけるストレージ管理方法であって、

前記ストレージ装置が、前記ストレージ装置の所定の記憶領域を 1 つ以上のボリュームとして前記ホスト計算機に提供する第 1 のステップと、

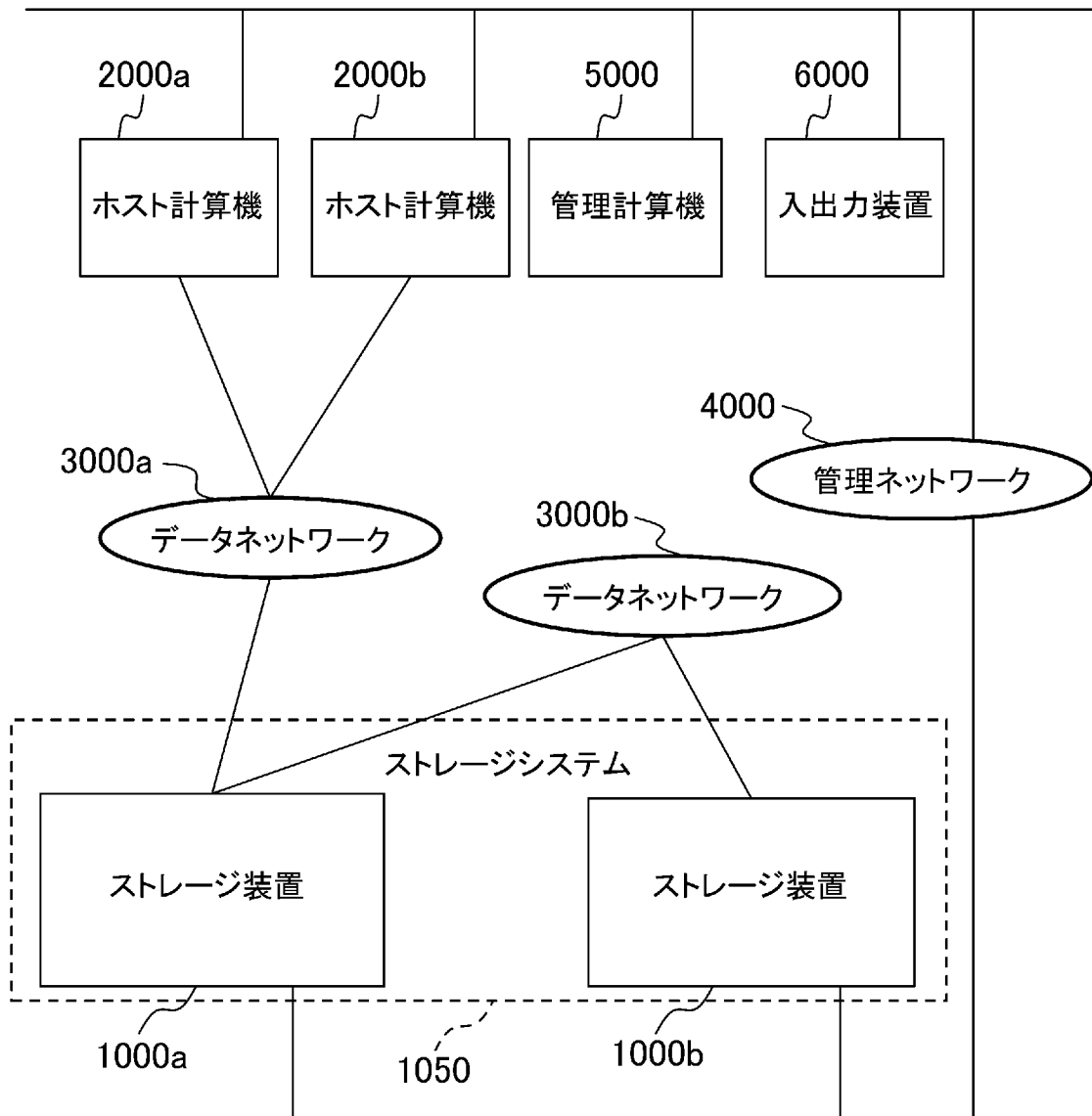
前記管理計算機が、前記記憶領域に関する統計情報を前記ストレージ装置に要求する第 2 のステップと、

前記ストレージ装置が、前記管理計算機の要求に応じて、前記記憶領域に関する統計情報を前記管理計算機に提供する第 3 のステップと、

前記管理計算機が、前記複数のストレージ装置から提供された前記ストレージ装置の前記記憶領域に関する統計情報を元に、前記複数のストレージ装置のデータ構成を管理する第 4 のステップと
を備えることを特徴とする、ストレージ管理方法。

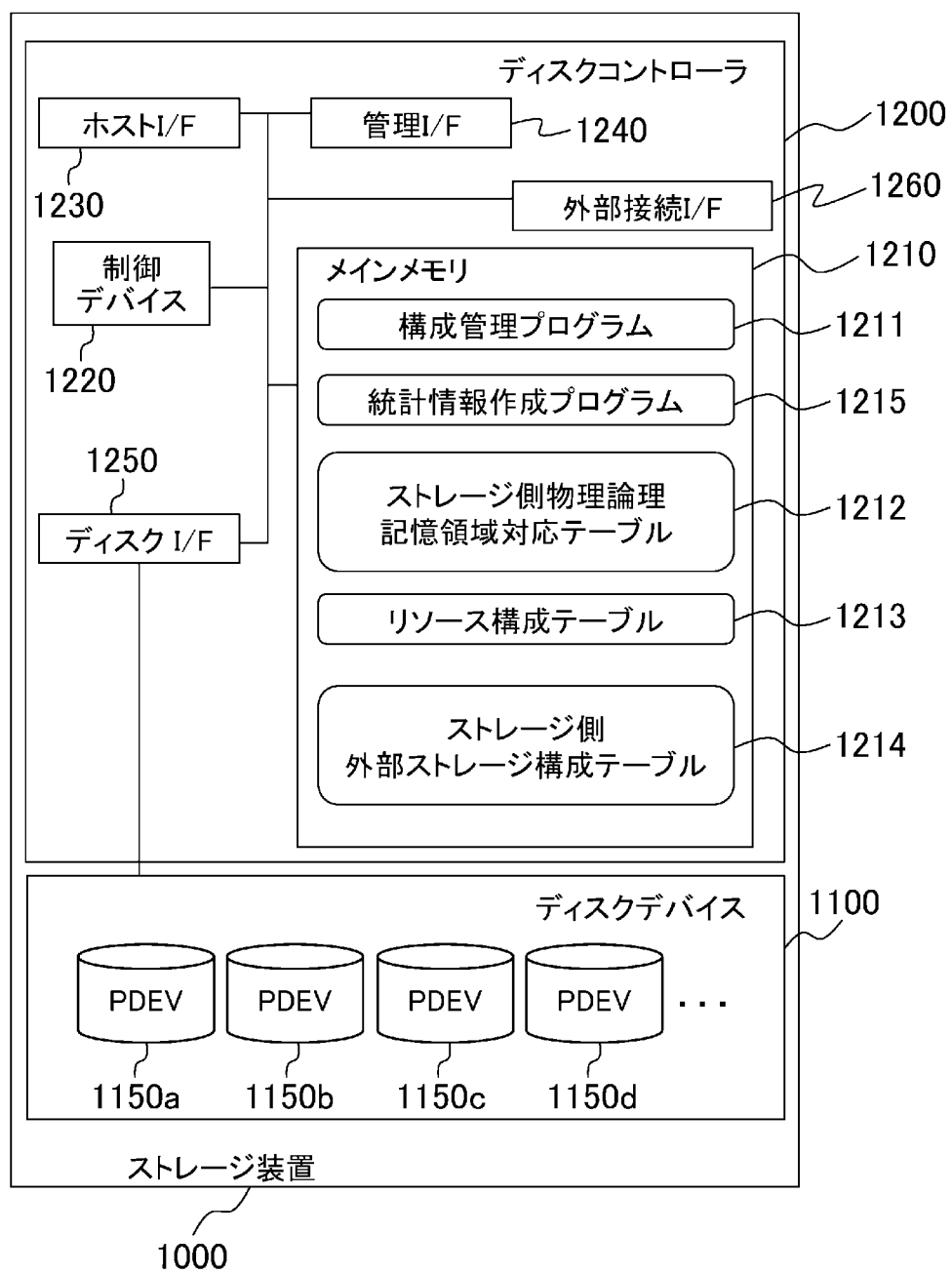
[図1]

図 1

1

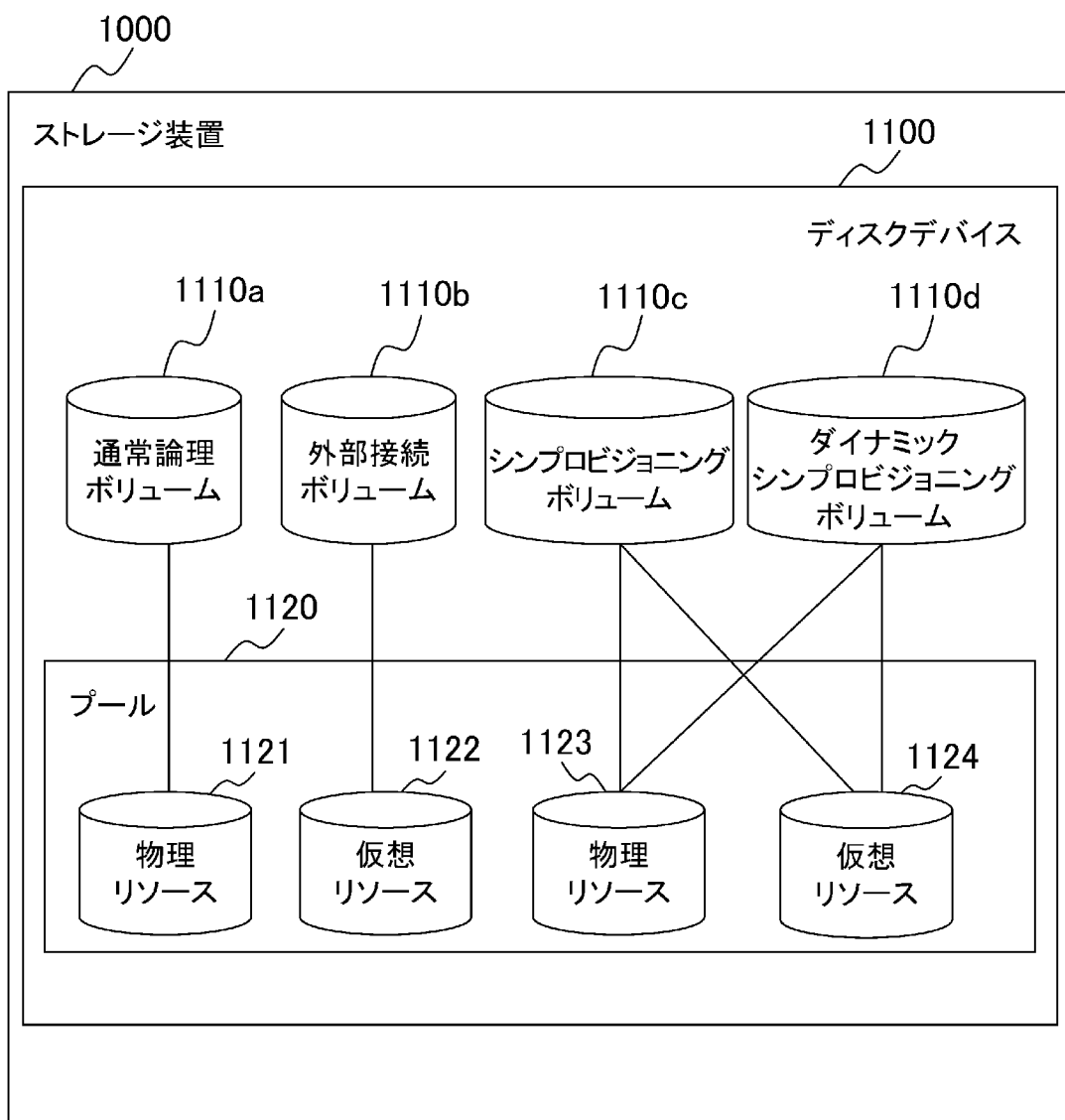
[図2]

図2



[図3]

図3



[図4]

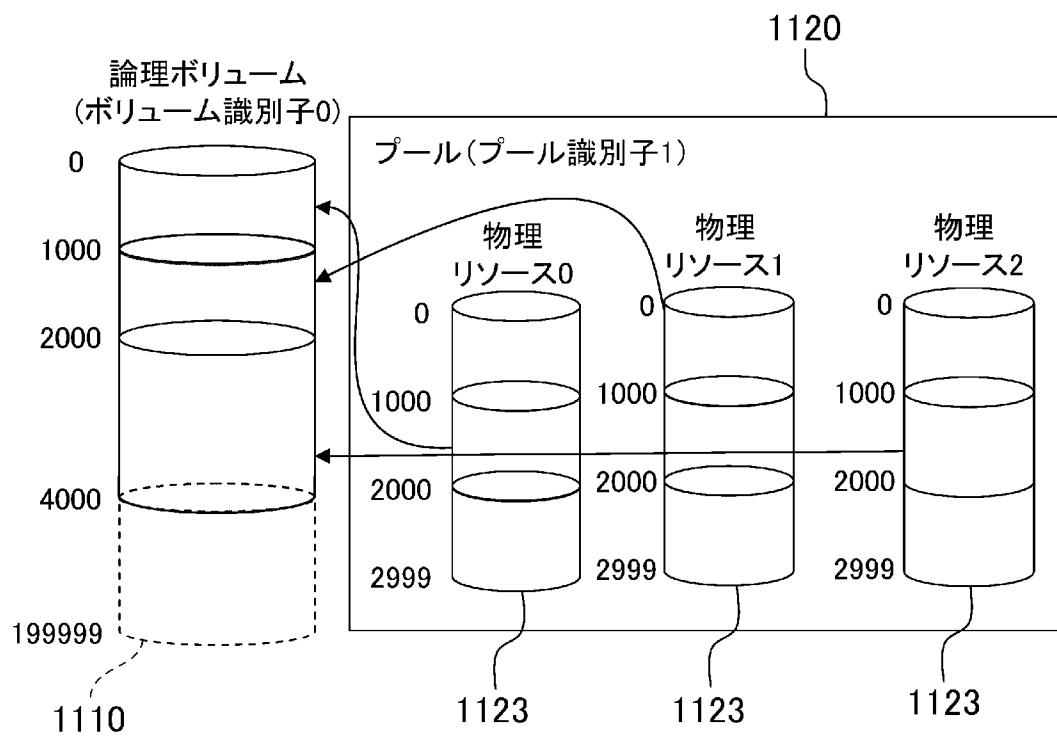
図4

ボリューム 識別子	論理開始 アドレス	論理終了 アドレス	プール 識別子	リソース 識別子	物理開始 アドレス	物理終了 アドレス
0	0	999	1	0	1000	1999
	1000	1999		1	0	999
	2000	2999		2	0	999
	3000	3999		2	1000	2999
	4000	199999		null	null	null
⋮	⋮	⋮	⋮	⋮	⋮	⋮
12120	12121	12122	12123	12124	12125	12126

1212

[図5]

図5



[図6]

図6

プール 識別子	リソース 識別子	リソース 種別	リソース 構成
0	0	FC	内部
	1	FC	内部
	2	SSD	内部
1	3	FC	外部

12130

12131

12132

12133

1213

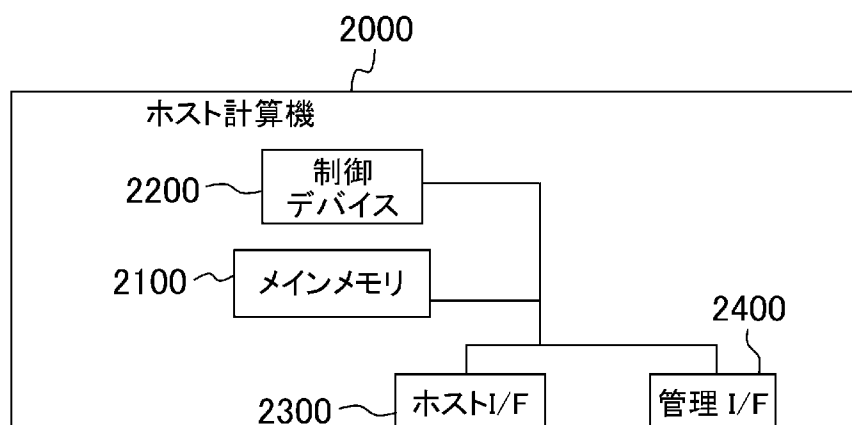
[図7]

図7

リソース 識別子	ポート 識別子	外部ストレージ 識別子	外部ストレージ側 リソース識別子	外部ストレージ ポート識別子	1214
3	0	Storage2	0	1	
⋮	⋮	⋮	⋮	⋮	
12140 12141		12142	12143	12144	

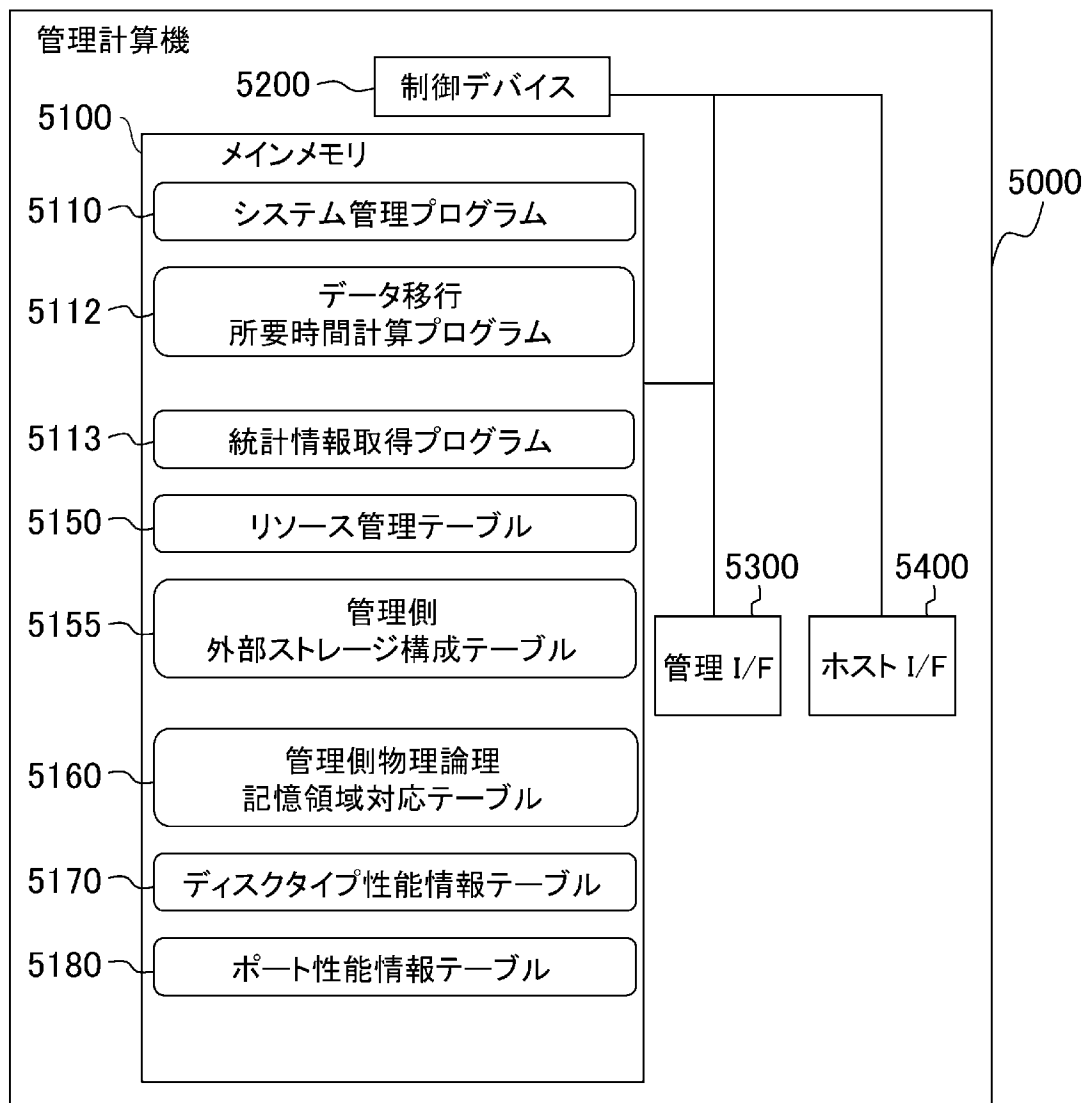
[図8]

図8



[図9]

図9



[図10]

図10

装置 識別子	プール 識別子	リソース 識別子	リソース 種別	リソース 構成
Storage1	0	0	FC	内部
	1	1	FC	内部
		2	SSD	内部
		3	FC	外部
	2	4	SATA	外部
⋮	⋮	⋮	⋮	⋮

51500

51501

51502

51503

51504

5150

[図11]

図 11

ストレージ 識別子	リソース 識別子	ポート 識別子	外部 ストレージ 識別子	外部 ストレージ側 リソース 識別子	外部 ストレージ ポート 識別子	
Storage1	3	0	Storage2	0	1	5155
	4	1	Storage2	1	2	
⋮	⋮	⋮	⋮	⋮	⋮	
51550		51551	51552	51553	51554	51555

[図12]

図12

装置 識別子	ボリューム 識別子	プール 識別子	リソース 識別子	割当済 セグメント数	
Storage1	0	1	1	1000	
			2	1000	
			3	2000	
⋮	⋮	⋮	⋮	⋮	5160
51600		51601	51602	51603	51604

[図13]

図13

装置 識別子	リソース種別	Read性能 (G/sec)	Write性能 (G/sec)
Storage1	SSD	10	5
	FC	8	4
	SATA	3	1
⋮	⋮	⋮	⋮

5170

51700 51701 51702 51703

[図14]

図 14

装置 識別子	ポート 識別子	最大IOPS (G/sec)
Storage1	0	10
	1	8
	2	3
⋮	⋮	⋮

51800

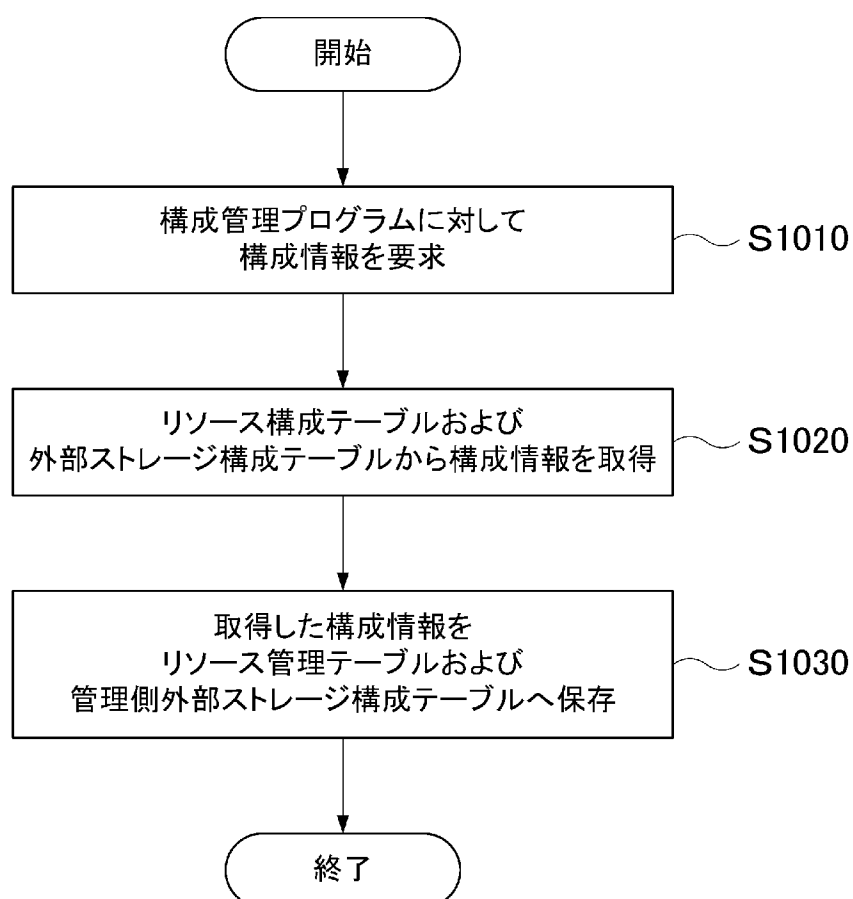
51801

51802

5180

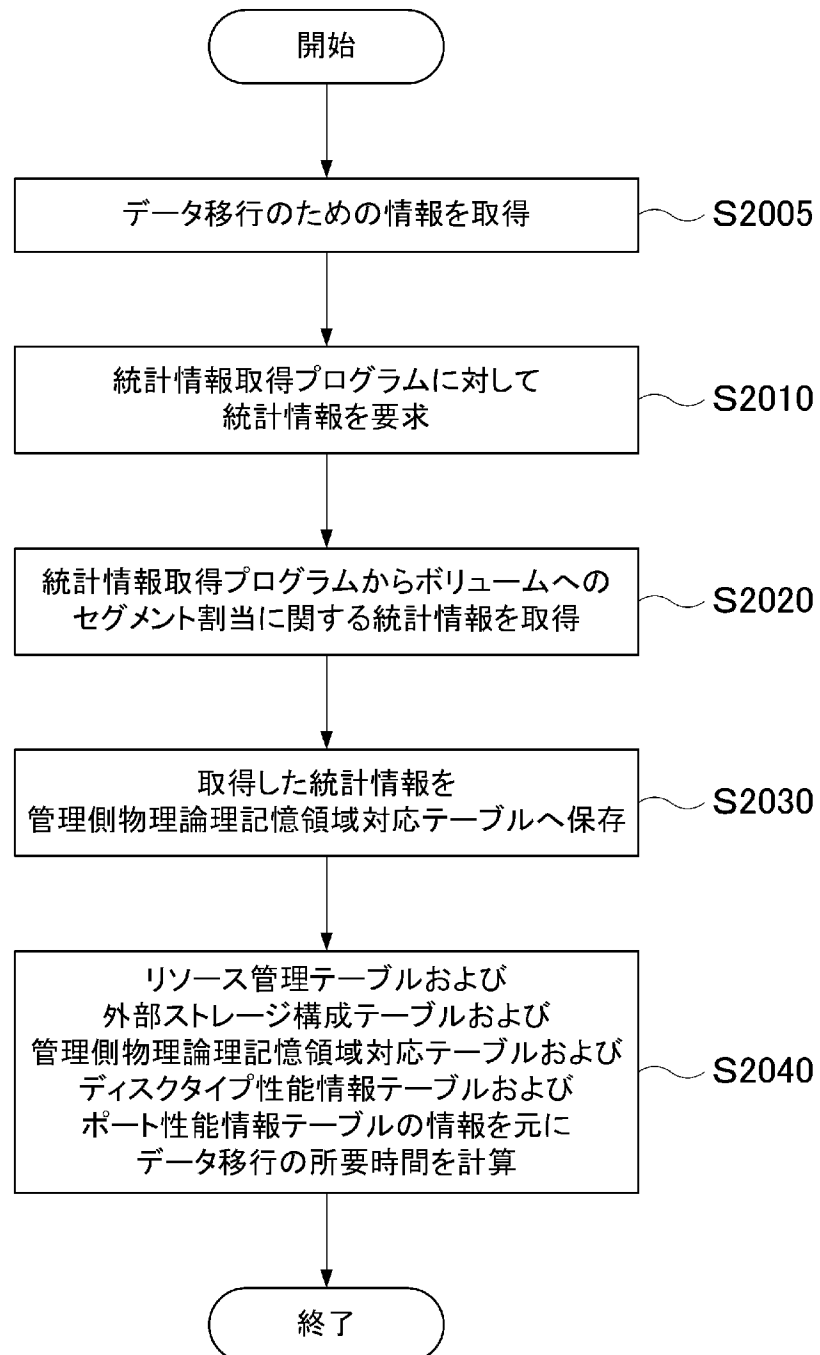
[図15]

図15



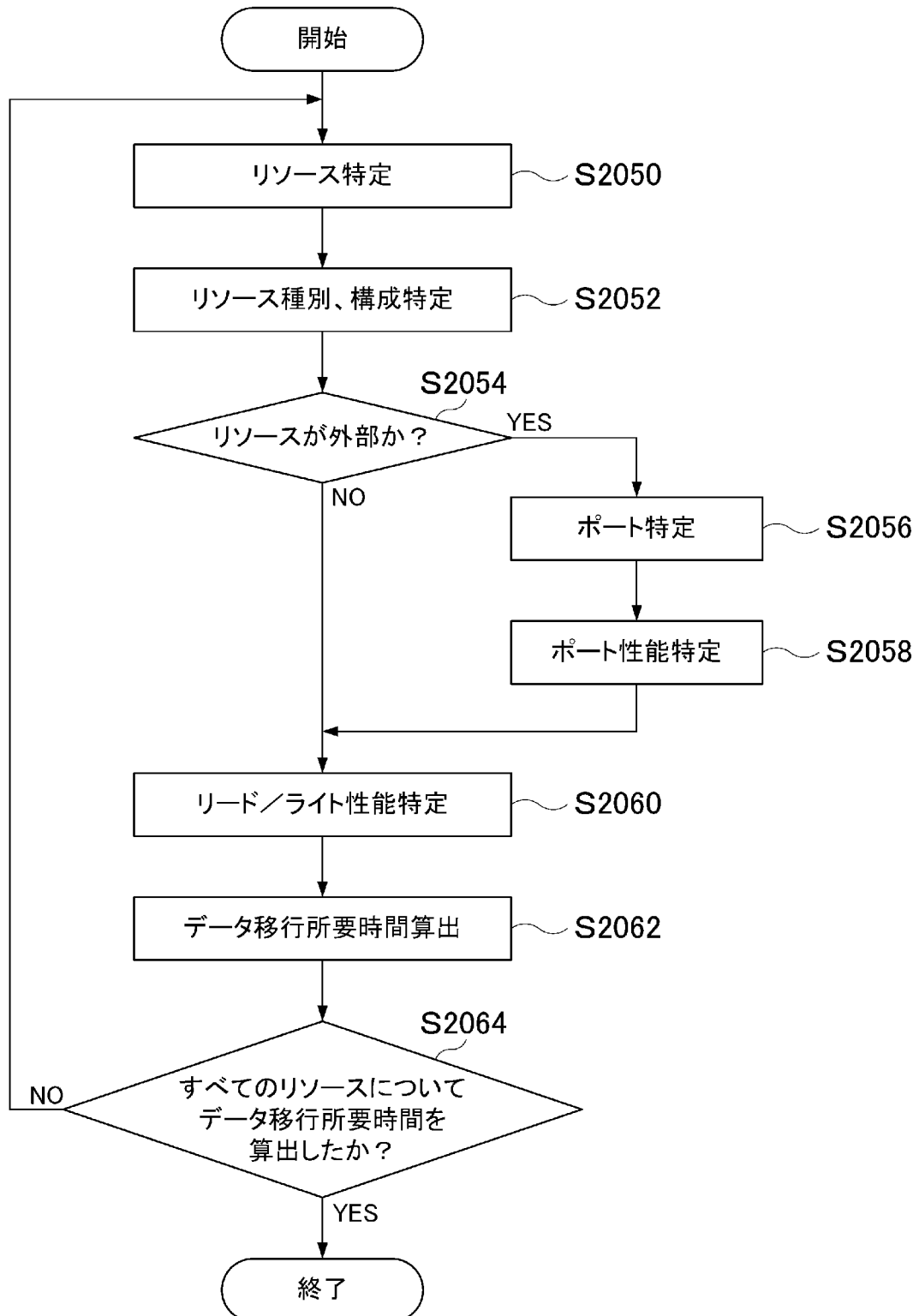
[図16]

図16



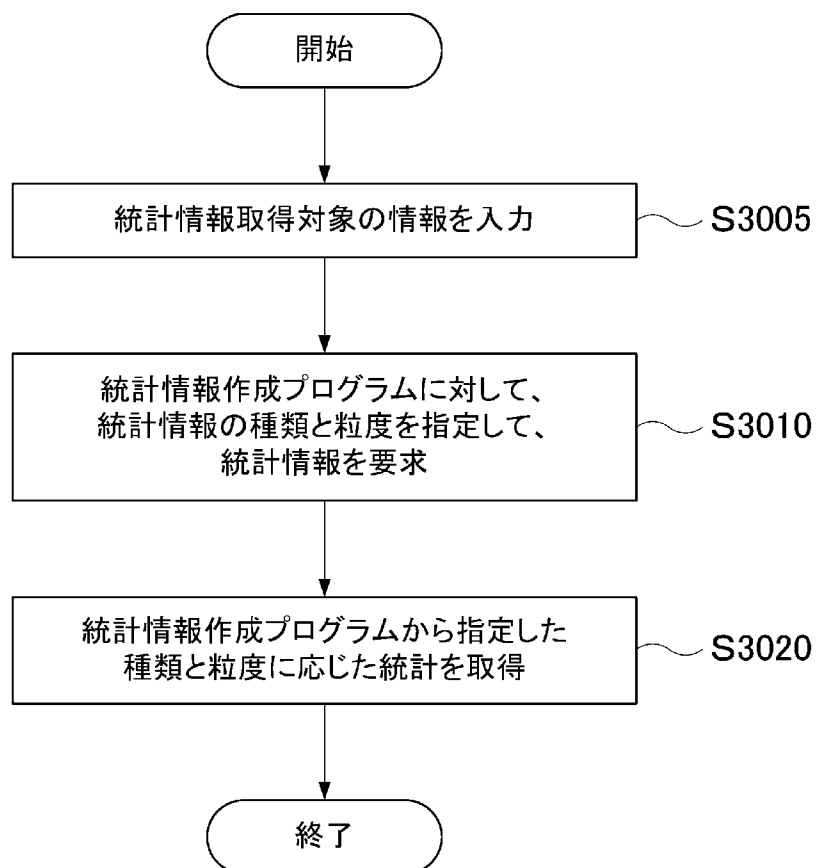
[図17]

図17



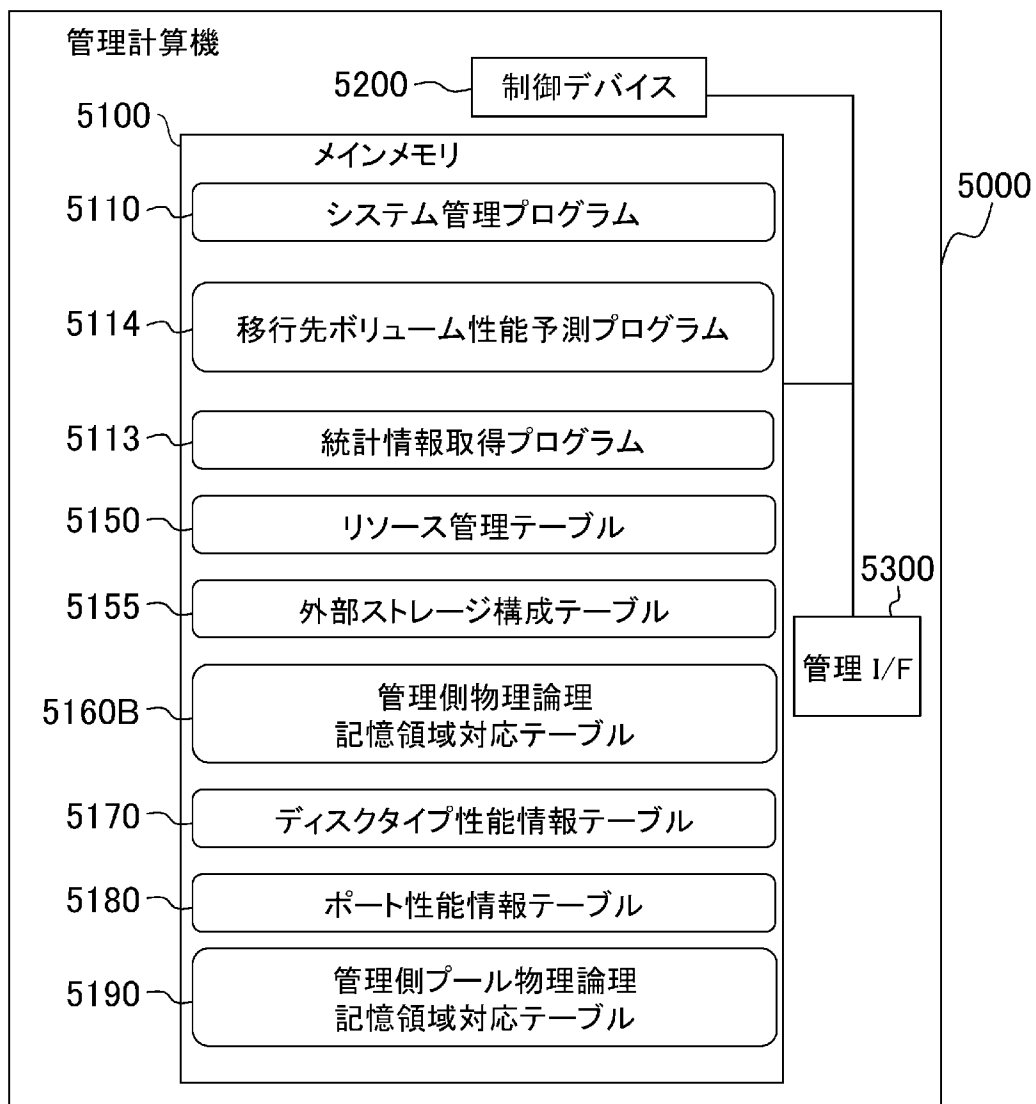
[図18]

図 18



[図19]

図 19



[図20]

図20

装置 識別子	ボリューム 識別子	プール 識別子	リソース 識別子	割当済 セグメント数	平均 IOPS
Storage1	0	1	1	1000	6000
			2	1000	5000
			3	2000	3000
⋮	⋮	⋮	⋮	⋮	⋮
51600	51601	51602	51603	51604	51605

5160B

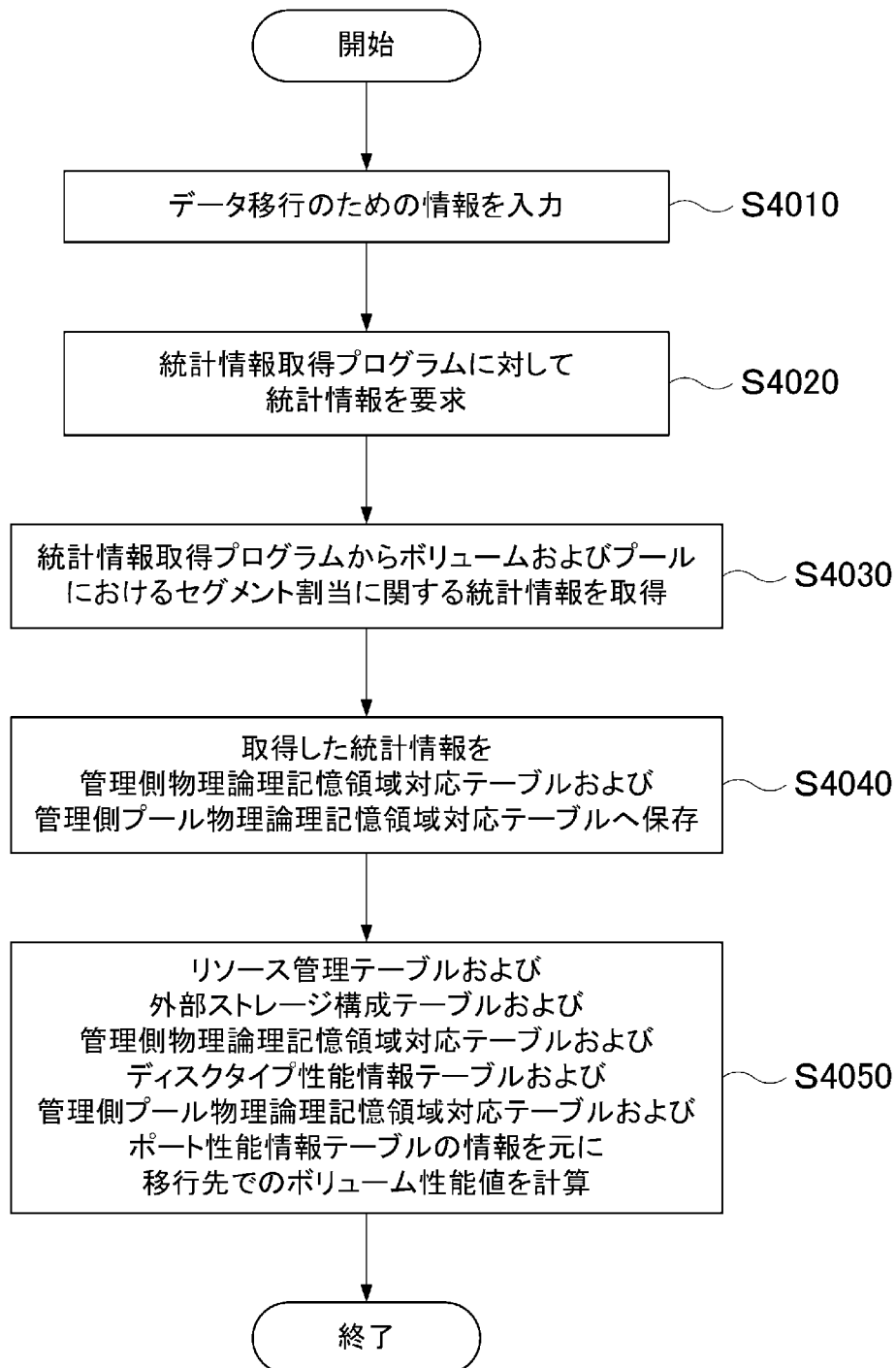
[図21]

図21

装置 識別子	プール 識別子	リソース 識別子	総 セグメント数	割当済 セグメント数	割当済 セグメントへ の平均IOPS
Storage1	1	10	1000	1000	4000
		11	1000	1000	3000
		12	8000	2000	2500
⋮	⋮	⋮	⋮	⋮	⋮
5190					
51900	51901	51902	51903	51904	51905

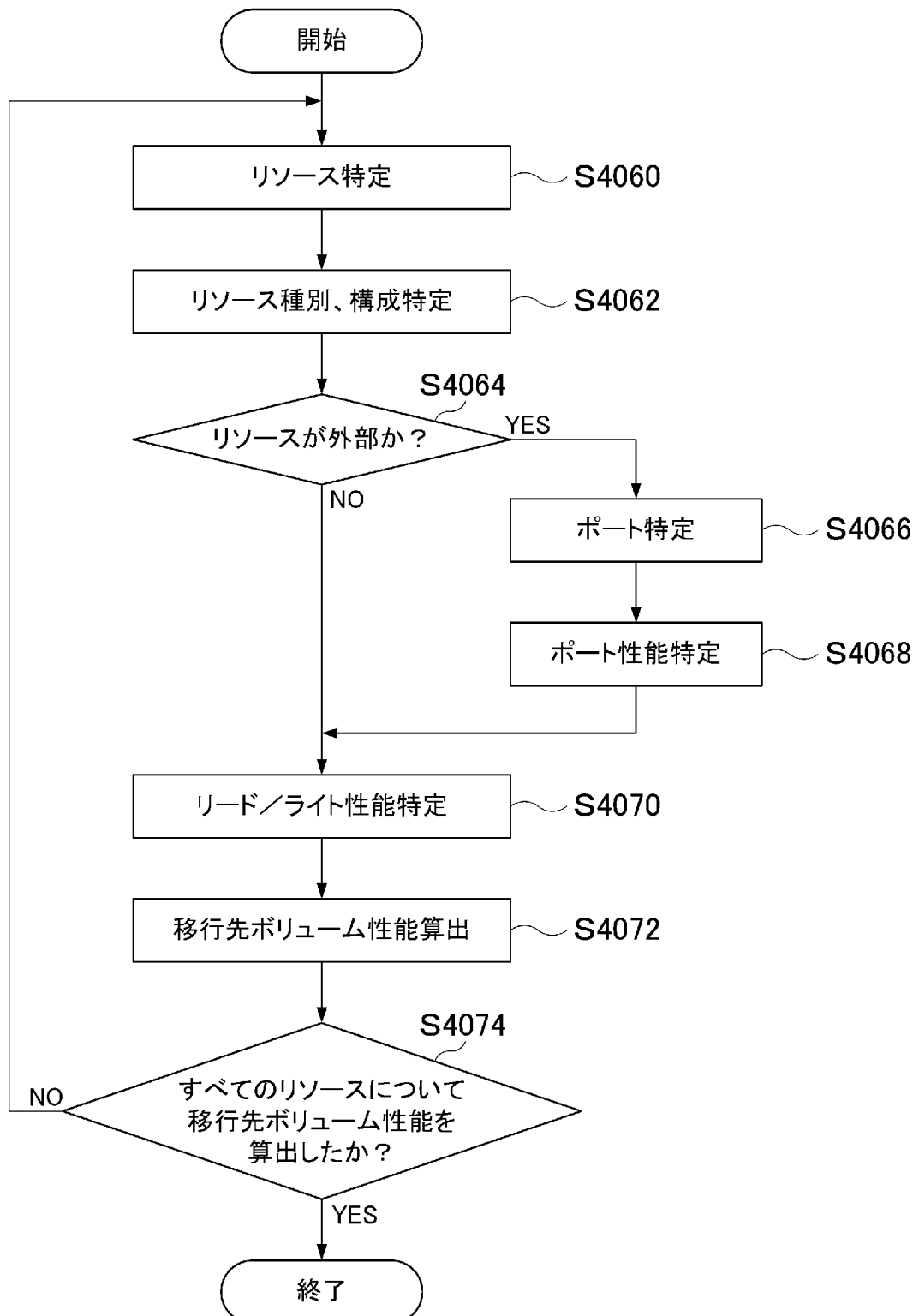
[図22]

図22



[図23]

図23



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/001680

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/00 (2006.01) i

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)

G06F12/00

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2011
Kokai Jitsuyo Shinan Koho	1971-2011	Toroku Jitsuyo Shinan Koho	1994-2011

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 2008-203937 A (Hitachi, Ltd.), 04 September 2008 (04.09.2008), paragraphs [0011] to [0020], [0043] to [0058], [0080] to [0090], [0099] to [0103]; fig. 1, 4, 9, 11 & US 2008/0201542 A1	1-5, 7-9 6
Y	JP 2008-47156 A (Hitachi, Ltd.), 28 February 2008 (28.02.2008), paragraphs [0039] to [0050], [0084] to [0090]; fig. 2, 8, 9 & JP 2006-99748 A & JP 2007-109262 A & US 2006/0047909 A1 & US 2006/0047930 A1 & US 2006/0143418 A1 & US 2008/0091898 A1 & EP 1630658 A1 & EP 1770493 A2	6

☐ 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 application or patent but published on or after the international filing date

“L” 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

“&” document member of the same patent family

Date of the actual completion of the international search
28 April, 2011 (28.04.11)

Date of mailing of the international search report
17 May, 2011 (17.05.11)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06F12/00(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06F12/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 1 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 1 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 1 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X Y	JP 2008-203937 A（株式会社日立製作所）2008.09.04, [0011]-[0020], [0043]-[0058], [0080]-[0090], [0099]-[0103]段落、 第 1, 4, 9, 11 図 & US 2008/0201542 A1	1-5, 7-9 6
Y	JP 2008-47156 A（株式会社日立製作所）2008.02.28, [0039]-[0050], [0084]-[0090]段落、第 2, 8, 9 図 & JP 2006-99748 A & JP 2007-109262 A & US 2006/0047909 A1 & US 2006/0047930 A1 & US 2006/0143418 A1 & US 2008/0091898 A1 & EP 1630658 A1 & EP 1770493 A2	6

C 欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の 1 以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

2 8 . 0 4 . 2 0 1 1

国際調査報告の発送日

1 7 . 0 5 . 2 0 1 1

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目 4 番 3 号

特許庁審査官（権限のある職員）

桜井 茂行

5 M

2 9 4 5

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9