

(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
2016年5月19日(19.05.2016)



(10) 国際公開番号

WO 2016/075779 A1

- (51) 国際特許分類:
G06F 3/06 (2006.01)
- (21) 国際出願番号:
PCT/JP2014/079986
- (22) 国際出願日:
2014年11月12日(12.11.2014)
- (25) 国際出願の言語:
日本語
- (26) 国際公開の言語:
日本語
- (71) 出願人: 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: 坂庭 秀紀(SAKANIWA Hidenori); 〒
1008280 東京都千代田区丸の内一丁目6番6号
株式会社日立製作所内 Tokyo (JP). 岡田 渡
(OKADA Wataru); 〒1008280 東京都千代田区丸の
内一丁目6番6号 株式会社日立製作所内
Tokyo (JP). 大平 良徳(OHIRA Yoshinori); 〒

1008280 東京都千代田区丸の内一丁目6番6号
株式会社日立製作所内 Tokyo (JP). 赤川 悅太郎
(AKAGAWA Etsutarou); 〒1008280 東京都千代田区
丸の内一丁目6番6号 株式会社日立製作所内
Tokyo (JP). 牧 晋広(MAKI Nobuhiro); 〒1008280 東
京都千代田区丸の内一丁目6番6号 株式会社
日立製作所内 Tokyo (JP). 森口 美緒子
(MORIGUCHI Mioko); 〒1008280 東京都千代田区
丸の内一丁目6番6号 株式会社日立製作所内
Tokyo (JP).

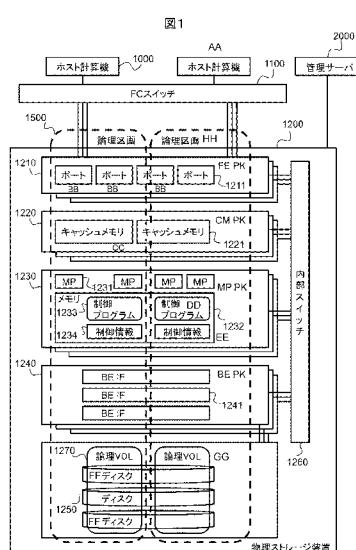
(74) 代理人: 青稜特許業務法人(SEIRYO I.P.C.); 〒
1040032 東京都中央区八丁堀二丁目24番2号
Tokyo (JP).

(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, BN, 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,

[続葉有]

(54) Title: COMPUTER SYSTEM AND STORAGE DEVICE

(54) 発明の名称: 計算機システムおよびストレージ装置



- 1000 AA Host computer
- 1100 FC switch
- 1200 Physical storage device
- 1210 FE PK
- 1211,BB Port
- 1220 CM PK
- 1221,CC Cache memory
- 1230 MP PK
- 1231 Memory
- 1232 DD Control program
- 1234 EE Control information
- 1240 BE PK
- 1241 BE IF
- 1250 FF Disk
- 1260 Internal switch
- 1270 GG Logic VOL
- 1500,HH Logic partition
- 2000 Management server

(57) Abstract: A computer system that is composed of a host computer, a storage device and a management computer. The storage device comprises a port for connecting with the host computer, a cache memory, a processor, and a plurality of logic volumes which are logical memory regions. For each logic volume, the port, the cache memory and the processor are divided into logic partitions, as resources that are used for reading and writing in the logic volume. The host computer reads and writes with respect to the logic volumes. If a failure occurs in the storage device, the management computer issues a command to the storage device to allocate the resources of a logic partition for which reading/writing performance is not ensured to a logic partition for which reading/writing performance is ensured.

(57) 要約: ホスト計算機とストレージ装置と管理計算機から構成される計算機システムであって、前記ストレージ装置は、前記ホスト計算機と接続するためのポートと、キャッシュメモリと、プロセッサと、論理的な記憶領域である複数の論理ボリュームと、を有し、前記論理ボリュームごとに、前記論理ボリュームの読み書きに使用される資源として、前記ポートと前記キャッシュメモリと前記プロセッサが論理区分に分割され、前記ホスト計算機は、前記論理ボリュームに対して読み書きを行い、前記管理計算機は、前記ストレージ装置に障害が発生した場合、読み書きの性能が保証されない前記論理区画の資源を、読み書きの性能が保証される前記論理区画へ割当てるよう、前記ストレージ装置へ指示を出す。



IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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.

- (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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユー

ラシア (AM, AZ, BY, KG, KZ, 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, KM, 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：計算機システムおよびストレージ装置

技術分野

[0001] 本発明は計算機システムおよびストレージ装置に関するものである。

背景技術

[0002] ストレージ装置のコンソリデーションが進み、データセンタなどでは複数の会社や複数の部署などが単一のストレージ装置を共有して使用するマルチテナント型の使用形態が増えてきている。同時に、ストレージ装置の大規模、複雑化が進み、限られた人数でストレージ装置全体を管理することが困難になった。このような状況に対して、一つのストレージ装置を複数の論理区画に分割し、論理区画毎に個別に管理可能にする技術が出てきた。これにより、ストレージ装置全体の管理者が論理区画を作成して会社毎又は部署毎に用意したストレージ管理者に割り当てることで、ストレージ装置の管理業務を委譲でき、管理の負荷を分散させることができた。

[0003] このようなストレージ装置を複数の論理区画に分割する技術に関して、例えば特許文献1には「単純にクラスタ型ストレージシステムに論理分割技術を適用すると、クラスタ間をまたがって論理区画を形成し、割り当てた資源量に応じた性能の論理区画を保証できない。・・・1つの論理区画に対して、第1のクラスタ内の資源を割り当てる。・・・更に、第1のクラスタに障害が起きた際、第2のクラスタが、第1のクラスタの処理を継続できるように構成をとってもよい」と記載されている。

先行技術文献

特許文献

[0004] 特許文献1：米国特許出願公開第2009／0307419号明細書

発明の概要

発明が解決しようとする課題

[0005] 特許文献1の記載によれば、割り当てた資源量（リソース量）に応じた性

能が保証される。しかしながら、第1のクラスタに障害が起きた際、第2のクラスタは性能を保証するに十分なリソース量を有しているとは限らない。

- [0006] エンタープライズ環境では、一般的に論理区画の性能を保証する技術が採用され、クラウド環境での一つのストレージ装置に複数の論理区画が存在する環境においては、障害時にも性能を保証しなければならない論理区画と、縮退運用を迫られる論理区画が混在する状況が存在する。
- [0007] そこで、本発明は、障害時に限られたリソースを論理区画に再配置し、必要な性能を保証する論理区画を提供することを目的とする。

課題を解決するための手段

- [0008] 本発明に係る代表的な計算機システムは、ホスト計算機とストレージ装置と管理計算機から構成される計算機システムであって、前記ストレージ装置は、前記ホスト計算機と接続するためのポートと、キャッシュメモリと、プロセッサと、論理的な記憶領域である複数の論理ボリュームと、を有し、前記論理ボリュームごとに、前記論理ボリュームの読み書きに使用される資源として、前記ポートと前記キャッシュメモリと前記プロセッサが論理区分に分割され、前記ホスト計算機は、前記論理ボリュームに対して読み書きを行い、前記管理計算機は、前記ストレージ装置に障害が発生した場合、読み書きの性能が保証されない前記論理区画の資源を、読み書きの性能が保証される前記論理区画へ割当るように、前記ストレージ装置へ指示を出すことを特徴とする。

発明の効果

- [0009] 本発明によれば、障害時に限られたリソースを論理区画に再配置し、必要な性能を保証する論理区画を提供できる。

図面の簡単な説明

- [0010] [図1]計算機システムの構成の例を示す図である。
[図2]管理サーバの構成の例を示す図である。
[図3]リソース管理テーブルの例を示す図である。
[図4]論理区画管理テーブルの例を示す図である。

[図5]リソース確保上限管理テーブルの例を示す図である。

[図6]リソース利用管理テーブルの例を示す図である。

[図7]リソース再配置設定の処理フローの例を示す図である。

[図8]一般的なリソース選択の処理フローの例を示す図である。

[図9A]障害発生時のリソース割当変更の例を示す図である。

[図9B]障害発生時のリソース割当変更なしの例を示す図である。

[図10]障害発生時のリソース確保の上限設定変更の例を示す図である。

[図11]FEポートのリソース情報テーブルの例を示す図である。

[図12]FEポートのリソース選択の処理フローの例を示す図である。

[図13]FEポートのチェックの処理フローの例を示す図である。

[図14]MPのリソース情報テーブルの例を示す図である。

[図15]MPのリソース選択の処理フローの例を示す図である。

[図16]キャッシュメモリのリソース情報テーブルの例を示す図である。

[図17]キャッシュメモリのリソース選択の処理フローの例を示す図である。

[図18]ディスクドライブのリソース情報テーブルの例を示す図である。

[図19]ディスクドライブのリソース選択の処理フローの例を示す図である。

[図20]I/O利用状況を監視する管理サーバの構成の例を示す図である。

[図21]テーブル管理情報の例を示す図である。

[図22]I/O性能に基づくリソース再配置設定の処理フローの例を示す図である。

[図23]I/O性能に基づくリソース選択の処理フローの例を示す図である。

発明を実施するための形態

[0011] 以下で、本発明を実施する形態の例を、実施例を用いて説明する。なお各実施例は、本発明の特徴を説明するものであって、本発明を限定するものではない。実施例を用いて説明される実施形態では、当業者が実施するのに十分なほど詳細にその説明がなされているが、他の実装・形態も可能であり、本発明の技術的思想の範囲と精神を逸脱することなく構成・構造の変更や多様な要素の置き換えの可能であることが理解される必要がある。

- [0012] 従って、以降の記述をその記述に限定して解釈してはならない。ある実施例の構成要素は、本発明の技術的思想の範囲を逸脱しない範囲で、他の実施例に追加又は他の実施例の構成要素と代替可能である。本実施形態は、後述されるように、汎用コンピュータ上で稼動するソフトウェアにより実装されてもよいし、専用ハードウェア又はソフトウェアとハードウェアの組み合わせにより実装されてもよい。
- [0013] なお、以後の説明では、主に「テーブル」形式によって本実施形態で利用される情報について説明するが、情報は必ずしもテーブルによるデータ構造で表現されていなくても良く、リスト、DB、キューなどのデータ構造やそれ以外で表現されていてもよい。
- [0014] 以下では「プログラム」を主語（動作主体）として本実施形態における各処理について説明を行う場合、プログラムはプロセッサによって実行されることで定められた処理をメモリ及び通信ポート（通信制御装置）を用いながら行うものである。このため、プロセッサを主語とした説明としてもよい。
- [0015] また、プログラムを主語として開示された処理は、管理サーバなどの計算機又はストレージシステムが行う処理としてもよい。プログラムの一部又は全ては専用ハードウェアで実現してもよく、また、モジュール化されていてもよい。
- [0016] 各機能を実現するプログラム、テーブル、ファイルなどの情報は、不揮発性半導体メモリ、ハードディスクドライブ、SSD (Solid State Drive) などの記憶デバイス、または、ICカード、SDカード、DVDなどの計算機読み取り可能な非一時的データ記憶媒体に格納されてもよく、プログラム配布サーバや非一時的記憶媒体によって計算機や計算システムにインストールされてもよい。

実施例 1

- [0017] 図1は、計算機システムの構成の例を示す図である。計算機システムは、ホスト計算機1000、スイッチ1100、物理ストレージ装置1200、管理サーバ2000で構成される。これらそれぞれの装置は1つまたは複数

で構成される。

- [0018] ホスト計算機1000は、一般的なサーバであってもよいし、仮想化機能を有するサーバであってもよい。一般的なサーバである場合は、ホスト計算機1000上で稼働するOSやアプリケーション（DBやファイルシステムなど）が物理ストレージ1200の提供する記憶領域に対してデータの入出力をを行うことになる。また、仮想化機能を有するサーバであれば、この仮想化機能あるいは仮想化機能により提供されるVM（Virtual Machine）上のアプリケーションが物理ストレージ1200の提供する記憶領域に対してデータの入出力をを行うことになる。
- [0019] ホスト計算機1000と物理ストレージ装置1200はFC(Fibre Channel)ケーブルにより接続されている。この接続を用いて、ホスト計算機1000あるいはホスト計算機1000上で稼働するVMは、物理ストレージ装置1200が提供する記憶領域に対してデータの入出力をを行う。ホスト計算機1000と物理ストレージ装置1200は直接接続されてもよいが、例えばFCスイッチであるスイッチ1100を介すことで、複数のホスト計算機1000や複数の物理ストレージ装置1200と接続する事もできる。複数のスイッチ1100が有る場合、スイッチ1100同士を接続する事で、さらに多くのホスト計算機1000と物理ストレージ装置1200が接続可能となる。
- [0020] 本実施例では、ホスト計算機1000と物理ストレージ装置1200はFCケーブルにより接続されているが、iSCSI (internet SCSI) などのプロトコルを使用する場合はEthernet (登録商標) ケーブルにより接続されてもよいし、その他のデータ入出力用に利用可能な接続方式で接続されてもよい。その場合のスイッチ1100はIP (Internet Protocol) スイッチでもよいし、その他の接続方式に適したスイッチング機能を持つ機器が導入されてもよい。
- [0021] 管理サーバ2000は、物理ストレージ装置1200を管理するためのサーバである。物理ストレージ装置1200を管理するために物理ストレージ装置1200とEthernetケーブルにより接続される。管理サーバ2000と

物理ストレージ装置1200は直接接続されてもよいが、IPスイッチを介することで複数の管理サーバや複数の物理ストレージ装置1200と接続する事もできる。本実施例では、管理サーバ2000と物理ストレージ装置1200はEthernetケーブルにより接続されているが、その他の管理用のデータ送受信が可能な接続方式で接続されてもよい。

- [0022] 物理ストレージ装置1200は前述のとおり、ホスト計算機1000とFCケーブルで接続されているが、これに加えて、複数の物理ストレージ装置1200の有る場合、物理ストレージ装置1200同士で接続されてもよい。ホスト計算機1000、スイッチ1100、物理ストレージ装置1200、管理計算機2000の数は、図1に記載された数に依らず、1つ以上であればいくつであってもよい。また、管理サーバ2000は物理ストレージ装置1200内に格納されていてもよい。
- [0023] 物理ストレージ装置1200は、管理サーバ2000により、複数の論理区画 (LPAR : Logical Partition) 1500に分割されて管理される。物理ストレージ装置1200は、FEPK (Front End Package) 1210、CMPK (Cache Memory Package) 1220、MPPK (Micro Processor Package) 1230、BEPK (Back End Package) 1240、ディスクドライブ1250、内部スイッチ1260で構成される。FEPK 1210、CMPK 1220、MPPK 1230、BEPK 1240は高速な内部バスなどでお互い接続されている。この接続は内部スイッチ1260を介して行われてもよい。
- [0024] FEPK 1210はデータ入出力用のインターフェース (Front End Interface) であるポート1211を一つ以上備えており、これを介してホスト計算機1000や他の物理ストレージ装置1200やスイッチ1100と接続される。データ入出力がFCケーブルを介した通信により行われる場合はFCポートとなるが、それ以外通信形態で行われる場合はその形態に適したIF (Interface) を備える。
- [0025] CMPK 1220はRAM(Random Access Memory)やSSD(Solid State Drive)などの高速アクセスが可能な記憶領域であるキャッシュメモリ1221を一つ以

上備えている。キャッシュメモリ1221には、ホスト計算機1000との入出力を行う際の一時データや、物理ストレージ装置1200が各種機能を動作させるための設定情報やストレージの構成情報などが格納される。

- [0026] MPPK1230はMP (Micro Processor) 1231、メモリ1232で構成される。MP1231は、メモリ1232に格納されたホスト計算機1000との入出力を行うためのプログラムや、物理ストレージ装置1200の各種機能のためのプログラムを実行するプロセッサである。ホスト計算機1000との入出力を行うためのプログラムや、物理ストレージ装置1200の各種機能のためのプログラムを実行するプロセッサが複数のコアからなる場合は、図1に示されたMP1231のそれぞれがコアであってもよい。
- [0027] メモリ1232は、RAMなどの高速アクセス可能な記憶領域であり、ホスト計算機1000との入出力を行うためのプログラムや、物理ストレージ装置1200の各種機能のプログラムである制御プログラム1233と、それらのプログラムで使用される制御情報1234を格納する。特に本実施例では、設定された論理区画に応じて入出力処理やストレージの各種機能を制御するための論理区画情報が格納される。
- [0028] MP1231、メモリ1232の数は、図1に記載された数に依らず、1つ以上であればいくつであってもよい。また、MPPK1230は管理用のインターフェースを持ち、これを介して管理サーバ2000と接続される。物理ストレージ装置1200の管理がEthernetケーブルを介した通信により行われる場合はEthernetポートとなるが、それ以外の通信形態で行われる場合はその形態に適したIFを備える。
- [0029] BEPK1240は、ディスクドライブ1250と接続するためのインターフェースであるBEIF (Back End Interface) 1241を備える。この接続形態は、SCSI(Small Computer System Interface)、SATA (Serial AT Attachment)、SAS(Serial Attached SCSI)などが一般的であるが、その他の接続形態でもよい。ディスクドライブ1250は、HDD(Hard Disk Drive)や、SSD(Solid State Drive)や、CDドライブ、DVDドライブなどの記憶装置である。

FEPK 1210、CMPK 1220、MPPK 1230、BEPK 1240、ディスクドライブ 1250、内部スイッチ 1260 の数は、図 1 に記載された数に依らず、1つ以上であればいくつであってもよい。

[0030] ここで、本実施例の制御プログラム 1233 について説明する。制御プログラム 1233 は、一般的なストレージ装置が保有するデータ入出力の処理プログラムを含んでいる。制御プログラム 1233 は、複数のディスクドライブ 1250 を用いて RAID(Redundant Arrays of Inexpensive Disks) グループを構成し、これを一つ以上の論理的な記憶領域に分割した論理ボリューム(論理VOL) 1270 をホスト計算機 1000 に提供することができる。この場合、データ入出力の処理は、この論理ボリューム 1270 に対する入出力を物理的なディスクドライブ 1250 への入出力に変換する処理を含む。本実施例ではこの論理ボリューム 1270 へのデータ入出力を行うことを前提とする。

[0031] また、このデータ入出力の処理は、論理区画 1500 間の性能影響を回避するために、各論理区画 1500 は割り当てられたリソースだけを使って処理を行うように制御を行う。例えば、入出力を行う際に MP 1231 の処理能力を使うが、MP 1231 の使用率が 50% 分割り当てられている場合は、この使用率がモニタリングされている。そして、使用率が 50% を越える場合には、その論理区画 1500 の処理をスリープして、他の論理区画 1500 の処理に MP 1231 を明け渡す制御が行なわれる。

[0032] あるいは、データ入出力の処理は、キャッシュメモリ 1221 の使用量が 50% 分割り当てられている場合、この使用率がモニタリングされており、使用率が 50% を越える場合には、この論理区画で使用しているキャッシュメモリ 1221 の一部をデステージするなどで解放し、空き領域を作った後に処理を進めるなどの制御を行う。

[0033] このような割り当てられたリソースのみを使って処理が行なわれればよく、どのように行われるかは特定される必要がない。すなわち、各論理区画 1500 の処理が、他の論理区画 1500 からの影響を受けずに、割り当てら

れた分のリソースを使用して処理を進められることがある物理ストレージ装置1200であればよい。

- [0034] また、制御プログラム1233は、二つの物理ストレージ装置1200間でデータをコピーするリモートコピー機能を有してもよい。リモートコピーでは、MP1231がコピー元の論理ボリューム1270のデータを読み出し、ポート1211を介してコピー先の論理ボリューム1270を有する物理ストレージ装置1200へ送信する。この送信を、コピー先の論理ボリューム1270を有する物理ストレージ装置1200のMP1231は、ポート1211を介して受信して、コピー先の論理ボリューム1270に書き込む。このようにしてコピー元論理ボリューム1270のデータは全てコピー先論理ボリューム1270にコピーされる。
- [0035] また、コピー中におけるコピー済み領域への書き込みは、コピー元論理ボリューム1270とコピー先論理ボリューム1270の両方に実施される必要がある。このため、コピー元の物理ストレージ装置1200への書き込み命令をコピー先の物理ストレージ装置1200に転送する。これらの物理ストレージ装置1200の機能は、さまざまな高機能化や簡略化が可能であるが、本実施例はその本質を変えることなくそれらの機能へ適用可能であるため、本実施例では、上述の機能を前提として説明を行う。
- [0036] 図2は管理サーバ2000の構成の例を示す図である。管理サーバ2000は、CPU (Central Processing Unit) であるプロセッサ2010、入出力IF2020、メモリ2030で構成される。プロセッサ2010は、メモリ2030に格納された各種プログラムを実行するための装置である。入出力IF2020はキーボード、マウス、タブレット、タッチペンなどからの入力を受け付け、ディスプレいやプリンタ、スピーカなどに出力するためのインターフェースである。メモリ2030は、RAMなどのデータ格納領域であり、各種プログラムやデータ、一時データなどを格納する。特に本実施例では、論理区画設定管理情報2040とリソース使用状況情報2050と論理区画設定プログラム2060が格納される。

- [0037] 図3は論理区画設定管理情報2040を構成するリソース管理テーブルの例を示す図である。ストレージ装置ID3000は本計算機システムにおける物理ストレージ装置1200のIDが格納される。この格納されたIDが指す物理ストレージ装置1200に属するリソースの種別がリソース種別3010に格納され、各リソースの実体を指示するIDがリソースID3020に格納される。性能・容量3030は、各リソースの最大性能、最大容量が格納される。
- [0038] 本実施例では、リソース種別3010にはMP1231のコアを示す「MP_Core」、キャッシュメモリ1221を示す「キャッシュメモリ」、ポート1211を示す「FEポート」、BE IF1241を示す「BE IF」、ディスクドライブ1250を示す「HDD」が格納される。性能・容量3030には、MP1231のコアの処理速度(MIPS)、キャッシュメモリ1221やディスクドライブ1250の容量(GB)、FE ポート1211やBE IF1241の性能(Gbps)が格納される。
- [0039] 障害時制約3040には、各リソースの障害発生時の制約情報が格納される。キャッシュメモリは、障害時にはデータが消失する可能性があるので、ライトスルー動作になり、書き込み性能が劣化するなどの制約情報が格納される。HDDに関しては、RAID構成であれば障害が発生したディスクドライブのデータ回復処理が動作し、RAIDグループ内のアクセス性能が劣化するなどの制約情報が格納される。これらの値を、あらかじめ論理区画設定プログラム2060がユーザ入力や物理ストレージ装置1200から収集した情報をもとに設定しておく。
- [0040] 図4は論理区画設定管理情報2040を構成する論理区画管理テーブルの例を示す図である。論理区画ID4000は論理区画1500のIDである。障害発生時に性能を保証しなければいけない論理区画か、縮退動作をする論理区画かの情報が、障害時の性能保証フラグ4010に格納される。論理区画IDにあらかじめ設定された性能要件が、性能要件4020に格納される。これらの値は、論理区画設定プログラム2060により、ユーザが論理区画15

OOを作成するときに設定される。

- [0041] 図5は論理区画設定管理情報2040を構成するリソース確保上限管理テーブルの例を示す図である。論理区画IDに設定された性能要件が性能要件4020に設定された場合に、その論理区画に割当てられるリソース確保量の上限の情報が格納される。例えば、ホスト1000と論理区画1500とのI/O性能を示すIOPS (Input/Output Operations Per Second) が200IOPSを満たす論理区画は、ポート1211を0.3個分、MP1231を0.5個分、キャッシュメモリ1221を200MB、ディスクドライブ1250を160GBの枠を上限としてリソースを確保する。
- [0042] ここでIOPSを満たすリソース上限は、ストレージ装置に所定の負荷をかけた場合の統計情報を元に作成してもよい。この4つのリソース確保量のパターンは環境によって大きく異なる可能性があるので、管理サーバで測定したIOPS、各リソースの使用状況に応じて、所定のIOPSを満たすためのリソース割当は変更されてもよい。性能要件のIOPSに近い状態のリソース使用状況を記憶しておいて、その値でリソース確保上限管理テーブルは更新されてもよい。または、現状のIOPSとその時使用されているリソース使用量の関係を利用して、性能要件のIOPS時のリソース確保上限は、前記の関係に比例した値で更新されてもよい。このリソースを確保すれば、想定範囲内の負荷時でも、性能要件を満たせるリソース量が設定される。
- [0043] 各論理区画1500は特定のリソースを上限の量だけ最初から割当てられて、その割当ては論理区画1500毎のリソースの所有権であってもよい。この場合、リソースであるポートやキャッシュメモリ、MP、ディスクドライブ毎に、どの論理区画1500が所有しているかを示すフラグが設けられてもよい。これにより、どの論理区画1500のリソースが、どの論理区画1500に貸し出しされているかなどが明確になり、性能保証フラグと連動したリソースの貸し借り管理が容易になるメリットがある。
- [0044] また、この上限は、リソースを確保できる権利の上限を意味してもよい。リソースの特定の所有権は設けられず、物理ストレージ装置1200のリソ

ース全体を管理サーバ2000が管理し、各論理区画1500は、必要なりソースを借りる（確保）できる権利を管理する。これにより、管理サーバ2000は、リソース全体の使用量と未使用量を管理して、論理区画1500が解放する量を指定することで、その解放した分のリソースの量を他の論理区画1500が利用できる。このように、リソースを共有化して、各論理区画1500に設定された上限分のリソースを確保できる権利に基づいて、各論理区画1500は共有リソースからリソースを確保する。リソース管理に関しては、この他のどのような管理の構成でもよい。

[0045] 図6はリソース使用状況管理情報2050を構成するリソース利用管理テーブルの例を示す図である。論理区画ID6000には論理区画1500のIDが格納される。ストレージ装置ID6010には、論理区画ID6000を構成する本計算機システムにおける物理ストレージ装置1200のIDが格納される。論理区画1500に割り当てられているリソースを示す情報が、リソース種別6020、リソースID6030、割り当て率・アドレス6040、使用率・使用状況・障害6050である。リソース種別6020には、割り当てられているリソースの種別が格納される。本実施例では、MP1231のコアを示す「MP_Core」、キャッシュメモリ1221を示す「キャッシュメモリ」、ポート1211を示す「FE ポート」、BE IF1241を示す「BE IF」、ディスクドライブ1250を示す「HDD」が格納される。

[0046] リソースID6030には割り当てられている具体的なリソースのIDが格納される。割り当て率・アドレス6040は、リソース種別により、格納される値の意味が変わる。リソース種別6020がMP_Core、FE ポート、BE IFであれば、各リソースの最大性能に対して論理区画1500が使用してよい割合が格納される。リソース種別6020がキャッシュメモリである場合は、使用可能なブロックのアドレスが格納される。本実施例では4KB (4096Byte) 単位でブロックが作成されるとし、各ブロックの先頭アドレスがここに格納されている。ディスクドライブ1250の場合は使用可能な容量がここに格納される。

- [0047] 使用率・使用状況・障害 6050 も、リソース種別により、格納される値の意味が変わる。リソース種別 6020 が MP_Core、FE ポート、BE IF、HDD であれば、各リソースの最大性能・容量に対して、論理区画 1500 が使用している割合が格納される。リソース種別 6020 がキャッシュメモリである場合は、キャッシュメモリ 1221 の使用状況が格納される。
- [0048] この使用状況は、キャッシュメモリ 1221 がどのようなデータを格納しているかを示す。例えば、この使用状況は、ホスト計算機 1000 から Write / Read 命令を受け付けて、ディスクドライブ 1250 に書き込む前のデータを保持ししているキャッシュ、ディスクドライブ 1250 から読み出されたデータを保持しているキャッシュに使われる Write / Read キャッシュである。また、この使用状況は、リモートコピー中に発生した書き込みデータが一時的に格納されたリモートコピーバッファ (R. C. Buffer) であってもよいし、一時的にリモートコピーバッファとなり、その後にコピーの完了したデータが格納された (R. C. Buffer (転送済み)) などであってもよい。
- [0049] 未使用の場合は「— (ハイフン)」などが格納される使用率・使用状況・障害 6050 の値は、リソースを他の論理区画 1500 に貸し出している場合、貸し出している量も加算された値である。たとえば、MP_Core を貸出元の論理区画 1500 が 10% 使っており、同一の MP_Core が他の論理区画 1500 に 10% 分貸し出されている場合、使用率・使用状況・障害 6050 の値は 20% となる。FE ポート、BE IF、HDD の場合も同じように、使用率・使用状況・障害 6050 は貸し出している量を加算された値である。
- [0050] キャッシュメモリの場合、使用率・使用状況・障害 6050 は貸出先での使用状況が格納される。また、障害発生時には、障害情報が格納される。さらに、リモートコピーバッファの使用率が高くなってきた場合には、ホスト計算機 1000 から論理区画 1500 へのデータの流入制限をすることによって、リモートコピーバッファが一杯にならないようにする制御が実行されてもよいが、性能保証フラグが設定されている論理区間の場合には、ホスト計算機 1000 と論理区画 1500 との間の IOPS 低下を防ぐために、リモー

トバッファ割当て量が増やされてもよい。

- [0051] 所定の時点のリモートコピーバッファ使用率と、所定の時点から一定期間の使用増加率に基づき、ある所定期間に80%以上の使用率になると予測された場合には、その所定期間に60%になるようにリモートコピーバッファの量を増加させる処理が実行されてもよい。これにより、性能要件のIOPSが維持可能となる。リソース利用管理テーブルの値は、論理区画設定プログラム2060により、ユーザが論理区画を作成するときに設定される。また、使用率・使用状況・障害6050は論理区画設定プログラム2060による定期的なモニタリングで更新される。
- [0052] 次に、論理区画設定プログラム2060の処理の流れを説明する。図7は、論理区画設定プログラム2060の障害発生時のリソースの再配置設定の処理フローの例を示す図である。管理サーバ2000のスケジューラにより定期的に起動されることにより、図7に示された処理フローは開始する。
- [0053] 起動されると、プロセッサ2010は、物理ストレージ装置1200から障害検出情報を取得し(S7000)、障害リソースがある場合には、そのリソースを論理区画に割当てないように割当て禁止処理を行う(S7010)。さらに各論理区画1500の各リソースの使用状況を取得し、図6に示されたリソース利用管理テーブルを更新する(S7020)。障害が発生したことによって、リソース利用が論理区画確保の上限となってしまっている仮想ストレージが存在するかを確認する(S7030)。
- [0054] S7030においてNOの場合は、障害が発生していても現在割当てられているリソースを使い切らずに処理が行えている状況であるので、プロセッサ2010はリソースの再配置は実行せずに処理を終了する。S7030でYESの場合、プロセッサ2010は、図4に示された論理区画管理テーブルを参照して、その論理区画が障害時に性能保証をするかしないかを性能保証フラグ4010により確認する(S7040)。
- [0055] 性能保証フラグ4010が設定されていない場合には、その論理区画は障害により確保できるリソースが制限され、性能が保証できない状況になる。

この際に、その論理区画に設定された性能要件を満たすためのリソース確保の上限設定を減少させる（S7050）。すなわち、障害が要因で性能保証できない論理区画は利用できるリソース量は減少しているので、この減少を他のリソースで補わないように、上限設定を減少させる必要がある。

- [0056] S7040で性能保証フラグが設定されている（YES）の場合、プロセッサ2010は、他の論理区画に貸し出している未使用リソースの有無をチェックする（S7060）。貸し出しているリソースがある場合には、貸し出している先の論理区画に返却処理を要求してリソースを回収する（S7070）。この回収により、性能を満たすリソースが確保できれば（S7080でN0）、処理は終了となる。
- [0057] 貸し出しリソースがない場合（S7060のN0）や、リソースが不足している場合（S7080のYES）には、プロセッサ2010は、性能を保証するために必要なリソース量を算出する（S7090）。これは、図5に示された性能要件（IOPS）に対するリソース確保量を参照し、算出してもよいし、発生した障害のリソース量に基づいて算出してもよい。発生した障害のリソース量と等価なリソース量が必要としてもよい。
- [0058] プロセッサ2010は、リソース選択処理を行う（S7100）。リソース選択処理においては、性能保証フラグを設定した論理区画において性能が保証できるかできないかの判定を行い、性能を保証できない場合には、警告フラグをONにする（図8を用いて説明する）。警告フラグがONの際は、管理者へのIF2020を介して性能が保証できない旨の警告を通知する（S7120）。
- [0059] 図8は、論理区画設定プログラム2060の障害発生時のリソース選択の処理フローの例を示す図である。リソース選択は、図7を用いて説明したS7100の処理である。プロセッサ2010は、性能保証ができないことを管理者に通知するかしないかを後半の処理で判定するため、初期設定として警告フラグをOFFに設定する（S8000）。まずは、性能保証フラグの設定がない論理区画の未使用リソースを利用して、性能保証フラグが設定さ

れている論理区画のリソース追加が可能かを確認する（S8010）。

- [0060] これは、図6に示したリソース利用管理テーブルを参照するなどして未使用リソースの量を算出することで確認できる。必要なリソースが確保できる場合（S8010のYES）、プロセッサ2010は、性能保証が設定されていない論理区画の未使用リソースの借用処理を行う（S8020）。まずは未使用リソースから借用することで、性能保証フラグが設定されていない論理区間においても現状の性能を可能な限り低下させないことが可能である。
- [0061] 未使用リソースのみではリソースが確保できない場合（S8010のNO）、プロセッサ2010は、性能保証フラグが設定されていない論理区画が使用しているリソースを減らすことで、リソースを確保し、確保したリソースを貸し出す（S8030）。図6に示したリソース利用管理テーブルを参照して、リソース使用量が少ない論理区画から順にリソースを解放していく。
- [0062] 例えばキャッシュの場合は、リソース解放時にデステージ処理が必要になり、デステージ処理の対象領域が広いと、デステージ処理時間が多く必要となるので、デステージ処理の影響する時間が長くなってしまう。このため、使用領域の少ない論理区画から解放処理をした方が、性能へ影響する時間を短縮できる可能性がある。そして、デステージ処理の済んだ領域は未使用領域として利用する。
- [0063] S8030を実行しても障害が要因の性能劣化を解消するためのリソースが確保できない場合（S8040のYES）に、プロセッサ2010は、性能保証フラグ設定ありの論理区間の未使用リソースから借用できるかを確認する（S8050、S8060）。この借用は、性能保証フラグの設定がある論理区間の間でのリソースの貸し借りになるが、リソースを貸す論理区画の動作が優先されるようにする。
- [0064] すなわち、リソースを貸す論理区画が一時的にリソースを貸した後でも、リソースを貸した論理区画が、リソースを必要とした場合は、借りた論理区画の状況に依らず直ちにリソースの返却を求められるようになる。このようになると、リソースを借りた論理区画は性能を保証されなくなるが、貸した

論理区画の都合に合わせた処理が実行される。

- [0065] この実施例では、リソースを確保可能であるかの確認（S 8050）と、その確保可能なリソースを一時的に借用可能であるかの確認（S 8060）とが分かれているが、これらの確認は一つの判断にまとめられてもよい。借用可能な場合（S 8060のYES）、プロセッサ2010は、性能保証フラグ設定ありの論理区画の未使用リソースの借用処理を行う（S 8070）。S 8070を実行しても障害が要因の性能劣化を解消するためのリソースが確保できない場合（S 8080のYES）には、性能保証フラグが設定されている論理区画で性能が保証できない旨を管理者に通知するための警告フラグをONにする（S 8090）。
- [0066] 図9A、図9B、図10は、障害発生時の論理区画のリソース確保の上限設定変更の例を示す図である。図7と図8を用いて説明した論理区画設定プログラム2060による処理の結果の一例である。
- [0067] 図9Aは、性能保証フラグ設定が有効な論理区画に割当てられているリソースで障害が発生した例を示す図である。この場合、性能を保証するために、性能保証フラグの設定されていない論理区画のリソースは、性能保証フラグのある論理区画に割当てられている（図9Aに示した右向き矢印）。その分だけ性能保証をしない論理区画で使用可能なリソースは縮小され、リソースが制限された中でベストエフォートな性能を出す。この際に、利用できるリソースの全体量は減少しているので、図7に示されたS7050の様に、この論理区画に設定しているリソース上限は縮小する必要がある。
- [0068] 図9Bは、性能保証フラグ設定が無効な論理区画に割当てられているリソースで障害が発生した例を示す図である。性能保証フラグが有効な論理区画は性能に障害の直接の影響を受けないので再配置せず、性能保証フラグが無効な論理区画で使用可能なリソースが縮小される。図9Aを用いた説明と同様に、この論理区画に設定しているリソース上限も縮小する必要がある。
- [0069] 図10は、正常時と障害時の各論理区画のリソースの上限の例を示す図である。正常時は論理区画のリソース上限があらかじめ決められ、その上限の

範囲内で必要なだけリソースが利用される形態である。障害時には、利用可能なリソースの全体量が減少するので、性能保証フラグが無効な論理区画2と論理区画3のリソース上限は減らされ、その枠の中で必要なリソースが割当てられて使用される。図10は、動作している処理への影響が少ない様に、未使用リソースが大きい論理区画ほど、リソース上限の縮小幅は大きくなる場合の例を示す図となっている。逆に、性能保証フラグが有効な論理区画1のリソース上限は大きくなっている。

- [0070] 基本的には、性能保証フラグが有効な論理区画では、障害前と同じリソース上限であれば性能への影響はないが、障害が発生した箇所によって性能保証のための安全係数があらかじめ用意されてもよい。これは、障害が発生した箇所によって、他への影響を考慮した係数であり、この係数に応じて論理区画の上限が増やされる。例えば、性能保証フラグありの論理区画が利用しているMPに障害が発生した場合、そのMPでの処理が発生しないようにスケジューリングを変更するために、本来の上限よりも多くのMPリソースが割当てられることによって、障害時でも性能が保証されうる。
- [0071] また、例えば、RAD5やRAD6で構成されているHDDに障害が発生した場合には、障害の発生したHDDの周辺に保存されている情報から、障害の発生したHDDのデータの回復処理が動作する。データの回復処理は複数の物理HDDへのアクセスが発生し、BEIF1241での切替処理などのために、直接影響のないリソース(HDD)での障害であっても論理区画に影響がある場合がある。そのような場合に、処理を高速化するために、図5を用いて説明したリソース上限よりもキャッシュリソースの割当てが多く設定されてもよい。
- [0072] また、図10は、障害時に論理区画2、3から論理区画1へリソースを貸し出す例を示す図であるが、リソース未使用率の高い論理区画2からまずリソースを借用し、さらに足りない分を、次に未使用率の高い論理区画3から借用している例である。
- [0073] 障害の発生したリソースの上限の増減量に比例して、障害の発生している

いリソースの上限も合わせて増減してもよい。障害の発生したリソースの上限が減る場合には、他の障害の発生していないリソースも使われる量を減られ、他の論理区画がリソースを必要としている時に、融通できるリソースは増やされる。また、障害が発生したリソースの上限を増やした場合には、障害の発生していないリソースも現状確保している上限よりも多く使用される可能性が高いので、比例して上限が増やされることで、性能保証に必要なリソースは確保される。

[0074] 図11は、各論理区画に割当てられたポート1211のリソース管理情報テーブルの例を示す図である。ポート1211のリソース管理情報テーブルは論理区画設定プログラム2060の中で参照される。このテーブルはリソースの貸し出し可能量を示しているが、これは、図6を用いて説明したリソースの使用量に加えて、未使用のリソースをX%（Xはあらかじめ設定された値）の余裕をもって残し、この余裕以外の未使用のリソースを貸し出せる分のリソース量として示している。このテーブルを元にどのポート1211の未使用リソースを借用できるかをチェックする。このテーブルは、図7を用いて説明した処理中のS7100のリソース選択処理で利用される。

[0075] 論理区画でのポート1211の借用は、障害に対応するため、通常時にはパスとして使われないが、すぐにパスが利用可能なように事前にマルチパスの設定がなされている場合には、その利用可能なパスを有効にしてその論理区画のポート番号を使用するように変更するのみで完了するので、性能劣化もなく借用、貸し出しが可能である。しかし、事前に論理区画用の利用可能なパスが設定されていない場合には、新たにパスが生成される必要がある。そのため、パス生成による時間でIOPS性能劣化をさせないために、事前にマルチパス設定がなされているポートを優先して割当てる処理が行われる。

[0076] 物理ポートにはキャッシュが存在するため、キャッシュにデータが残っている間は、以前に設定してあるポートへI/Oデータが送られてしまうので、ポートキャッシュがクリアされるまでの時間を待たされる場合がある。この際には、一時的にポートのキャッシュをOFFにして、ポートの切り替え処理が

行なわれる。

- [0077] このリソース情報管理テーブルによると、リソースの貸出可能量 11040 のみでは、リソースを選択できないので、図 12 を用いて説明する処理フローでこのテーブルを利用して、障害時のリソース不足を補うためにどこからリソースを借用するかの判断が行われる。例えば、性能保証フラグ 11010 が有効（「1」）な VPS2 の FE ポートのリソースが不足である場合には、まず性能保証フラグ 11010 のない Port # A-4、5、6、Port # B-1、2（貸出リソース 11030）が選ばれる。
- [0078] VPS5（貸出元論理区画 11000）に割当てられている Port # B-1、2 は、ストレージ装置 ID 11020 が別ストレージ装置であることを示しているので、ストレージ装置構成の変更に時間のかかる可能性がある。そこで、まずは、ストレージ装置 ID 11020 が同じストレージ装置内を示す Port # A-4、5、6 を選択し、その内で貸出可能量 11040 の値が最も大きい Port # A-6 が選択されることになる。障害時使用制約 11050 の設定があるポートがあれば、そのポートを選択するとリスクがあるので、そのポートは選択されないようにする。
- [0079] 以上ではポート単位での貸し借りを説明したが、1つのポートが時分割で利用されて、その割当て時間でリソースが配分されてもよい。
- [0080] 図 12 は、図 7 の S7100 で行われる FE ポートのリソース選択の処理フローの例を示す図である。この処理フローは論理区画設定プログラム 2060 の一部である。FE ポートの場合、論理区画での FE ポートの借用の処理フローは、マルチパスが張られていれば、論理区画のポート番号を変更するのみで完了するので、基本的には図 8 を用いて説明した処理フローと同じである。このため、異なる処理の部分のみの説明をする。
- [0081] S12030 のポートチェックは、FE ポートをチェックして、リソースを借用できる場合（S12030 の YES）に、プロセッサ 2010 は、図 8 を用いて既に説明した処理を行う。S12030 のポートチェックで NO となつた場合には、リソースが確保できないということで、管理者にその旨を通知

する処理S12100を実行する。ポートチェックの処理内容を、図13を用いてさらに説明する。

- [0082] 図13は、図12のS12030で行われるFEポートのリソースが割当て可能かの事前チェックの処理フローの例を示す図である。この処理フローは論理区画設定プログラム2060の一部である。まず、プロセッサ2010は、論理区画と接続されたホスト計算機1000とマルチパスが構成されているかをチェックする(S13000)。マルチパス構成である場合(S13000のYES)には、論理区画のポート番号を変更するのみでリソース割り当てが可能である。この場合は“YES”が設定されて処理が終了となる。
- [0083] マルチパス構成でない場合(S13000のNO)に、プロセッサ2010は、マルチパス構成が構築可能かをチェックする(S13010)。例えば、ホスト計算機1000と物理ストレージ装置1200とが実際に結線されていないとマルチパスは実現できなく、また、物理ストレージ装置1200の構成管理情報を大幅に変更する必要である場合は、マルチパスの構築処理に多くの時間が必要となるため、構築可能でないと判定される。
- [0084] マルチパス構成が可能な場合(S13010のYES)に、プロセッサ2010は、マルチパス構築処理を実行し(S13020)、論理区画でのリソース貸し借りが自由にできるようになるため“YES”が設定されて処理が終了となる。結線がなされていない、または物理ストレージ装置1200の構成上マルチパスが構築困難な場合(S13010のNO)には、“NO”が設定されて処理が終了となる。
- [0085] 図14は、各論理区画に割当てられたMP1231のリソース管理情報テーブルである。MP1231のリソース管理情報テーブルは論理区画設定プログラム2060の中で参照される。決められたMP1231のみが任意の論理ボリューム1270の情報を取得できる権利(オーナー権)が設定されている場合がある。これにより、任意の論理ボリューム1270に対するデータの入出力処理を行えるため、任意の論理ボリューム1270の構成情報を

設定情報をキャッシュメモリ1221からローカルなメモリ1232に一度取り込んでしまえば、MP1231は構成情報と設定情報を取得するためにキャッシュメモリ1221へアクセスする必要がなくなる。

- [0086] MPリソースの再配置は、MPを使用するオーナー権を切替るのみであり、オーナー権の切替により別の論理区画でのMPの利用が可能となる。基本的にMPのリソース選択の処理フローは、図8を用いて既に説明した処理フローと同じである。図8を用いて説明した処理フローと、MPのリソース選択の処理フローとでは、どの未使用リソースを選択するかの基準として、図14に示したMP1231のリソース管理情報テーブルが利用されることが異なる。
- [0087] MP1231のスリープ期間は未使用として識別されてもよい。MP1231のスリープ期間はMP1231が使われていない期間であるので、この期間を他の論理区画が利用するようにスケジューリング処理を行うことによって、MPリソースの割当は調整される。MPリソースの貸し借りは、MP1231のコア単位ではなく、MP1231の単位で貸し借りされてもよい。
- [0088] コア単位でリソースが貸し借りされてしまうと、他の論理区画の処理とMP1231の内部のL2キャッシュが共有されてしまい、他の論理区画との間で性能の影響がでてしまう可能性もある。このような可能性のない場合は、MP1231の単位で貸し借りされてもよい。さらに、MPPK1230の内部のメモリ1232や図示を省略したバスが共有されると影響の出る場合、メモリ1232やバスも論理区画毎に割当ることが望ましい。
- [0089] 例えば、性能保証フラグが有効なVPS2のMPリソースが不足の場合は、まず性能保証フラグのないMP2_Core#a、b、c、MP3_Core#a、bが選ばれる。VPS5に割当てられているMP3_Core#a、bは、別の物理ストレージ装置であるので、まずは、同じ物理ストレージ装置内のMP2_Core#a、b、cが選択される。
- [0090] この選択された中で貸し出し可能量は、すべて同じ35%になっているが

、VPS3にはMP2_Core#a、bの2つが割り当たっているため、貸し出し可能量はVPS4よりもVPS3の方が多い。このため、MP2_Core#aを選択してMPリソースをVPS2に貸し出す処理が行なわれる。障害時にオーナー権を固定して使用するなどの障害時制約があるMPは、選択優先度が低くなる。

[0091] 図15は、図7のS7100で行われるMPリソース選択の処理フローの例を示す図である。この処理フローは論理区画設定プログラム2060の一部である。既に説明したように、MPリソースの再配置は、MPを使用するオーナー権を切替るのみにより、別の論理区画での利用が可能である。これ以外は図8を用いて説明したリソース選択の処理フローと同じであるので、説明を省略する。

[0092] 図16は、各論理区画に割当てられたキャッシュメモリ1221のリソース管理情報テーブルの例を示す図である。キャッシュメモリ1221のリソース管理情報テーブルは論理区画設定プログラム2060の中で参照される。キャッシュメモリ1221に障害があり、貯まっているデータが破壊されると回復が不可能となるので、基本的にキャッシュメモリ1221は2重化される。さらに、キャッシュメモリ1221に障害が発生した場合、キャッシュメモリ1221の1部の領域のみが使えなくなる状況は少なく、キャッシュメモリ1221の1面全部が使用できなくなることが多い。そのため、障害により、2重化して動作しているキャッシュメモリ1221の1面全部が使えない状態において、性能の保証される必要がある。

[0093] また、キャッシュメモリ1221が2重化されていない場合、キャッシュメモリ1221に格納されたデータが障害時に破壊されてもよいように、キャッシュメモリ1221はライトスルー設定され、データはキャッシュメモリ1221に書き込まれると同時に、論理ボリューム1270に書き込まれる場合がある。この場合には、正常に動作している1面のキャッシュが、仮想的に2面化してライトスルー用の領域とReadキャッシュ用の領域で分離されてもよい。サーバからシーケンシャルな連続データの読み出しがある場合には、Readキャッシュにデータがプリフェッチされることで、ReadのI/O性

能が向上する。

- [0094] また、別の物理ストレージ装置1200が管理するキャッシュメモリ1221の貸出可能な領域がある場合には、別の物理ストレージ装置1200のキャッシュリソースを借用して割当てる。これにより、ライトスルーで使用される領域以外にReadキャッシュやリモートコピー用のバッファに利用でき、ReadやリモートコピーのI/O性能の改善が期待できる。
- [0095] 図16に示したリソース管理情報テーブルでは、キャッシュリソースの算出された貸出可能量が格納されているが、キャッシュメモリ1221にデータが残っている場合、デステージが発生し、このデステージの領域が大きければ、それだけデステージに時間が長くなるので、性能劣化になってしまう可能性がある。障害時にライトスルー設定にならないシステムの場合、キャッシュリソースの未使用領域の選択時は、デステージ時間による性能劣化を最小限とするため、リソース管理情報テーブルの貸出可能量が多い、すなわち、割り当てキャッシュ量に対して使用率が低いキャッシュリソースを選択することにより、デステージされるデータ量が少なくなる。
- [0096] 図17は、図7のS7100で行われるキャッシュリソース選択の処理フローの例を示す図である。この処理フローは論理区画設定プログラム2060の一部である。キャッシュメモリ1221は、障害時の影響が大きい部分であり、ライトスルー設定になってしまふと、一気に性能が低下する可能性もあるので、図17に示すリソース選択の処理フローは、図8に示す処理フローから大きく変更されている。
- [0097] まず、プロセッサ2010は、警告フラグをOFFにし(S17000)、障害時にキャッシュメモリ1221がライトスルー動作になるかどうかを判定する(S17010)。ライトスルー動作になつまう場合には、性能劣化は避けられないが、それでも性能保証フラグが有効な論理区画の性能が確保される場合(S17020のN0)には、そのままの装置構成で問題ないため、そこで処理が終了となる。
- [0098] ライトスルー動作による性能劣化を軽減するため、別の物理ストレージ装

置 1200 のキャッシュメモリ 1221 が利用されてもよい。HAクラスタ (High Availability Cluster) 構成によって複数の物理ストレージ装置 1200 が接続されている場合には、別の物理ストレージ装置 1200 のキャッシュメモリ 1221 を利用できる可能性がある (S17030)。HAクラスタ構成以外の構成であっても、別の物理ストレージ装置 1200 のリソースを共有することができる場合には、障害の発生していない物理ストレージ装置 1200 のキャッシュメモリ 1221 を共有することで、性能劣化の軽減される可能性がある。

[0099] しかし、このようなシステム構成でない場合 (S17040 のN0) には、プロセッサ 2010 は、警告フラグをONにし (S17130)、性能保証フラグが有効な論理区画の性能が保証されないことを管理者に通知する。このようなシステム構成であった場合 (S17040 のYES) でも、プロセッサ 2010 は、キャッシュリソースの借用処理を行い (S17050)、性能保証フラグが有効な論理区画の性能が確保されない場合 (S17060 のN0) には、警告フラグをONにする (S17130)。

[0100] 障害時にライトスルー動作にならないが (S17010 のN0)、キャッシュリソースが不足する場合 (S17070 のYES) に、プロセッサ 2010 は I/O パタンを確認する (S17080)。I/O パタンがシーケンシャルの場合 (S17080 のYES) には、リードキャッシュのリソース量を増やすことにより (S17090)、リード性能の向上を試みる。それでも性能不足の場合 (S17100 のYES) には、物理ストレージ装置 1200 によっては、キャッシュメモリ 1221 の性能が高く、キャッシュリソースを増やすと性能が上がる場合もあるので、プロセッサ 2010 は、図 16 に示したリソース管理情報テーブルを参照し、性能保証フラグが無効な論理区画から未使用リソースの多い順にキャッシュのリソースを借用する (S17110)。

[0101] しかし、キャッシュリソースを増やしても確実に I/O 性能が向上するかは不明であるので、S17070～S17110 は省略されてもよい。性能を保証するためのキャッシュリソースが不足の場合 (S170120 のYES) に

は、プロセッサ2010は警告フラグをONにする（S17130）。

[0102] 図18は、各論理区画に割当てられたディスクドライブ1250のリソース管理情報テーブルである。このリソース管理情報テーブルは、各論理区画に設定された、性能保証フラグの有無、ストレージ装置ID、貸出リソース（HDD／SSDなど）や、その貸出可能量、障害時制約情報などを含む。物理ストレージ装置1200では、複数のディスクドライブ1250によりRAIDを構成するため、そのRAID構成によって障害時にデータ回復が可能か、回復までの時間などが決まる。

[0103] このディスクドライブ1250のリソース管理情報テーブルが参照され、リソース選択の処理が行なわれる。まずは性能保証フラグが無効な論理区画からディスクリソースは借用され、同じ物理ストレージ装置1200内か、ディスクドライブ1250の種類がHDDかSSDかなどの性能と、貸出可能量に基づきリソース選択の処理が行なわれる。

[0104] 図19は、図7のS7100で行われるディスクドライブ1250のリソース選択の処理フローの例を示す図である。ディスクドライブ1250の障害は、キャッシュメモリ1221の障害と同様にハードウェア的な制約が大きく、図19に示す処理フローは、図8を用いて説明した処理フローから大きく変更した処理フローとなる。まず、プロセッサ2010は、警告フラグをOFFにし（S19000）、障害時にデータ回復処理が動作するかを確認する（S19010）。データ回復処理中でも性能が保証できる場合（S19020のNO）には、そこでリソース選択の処理は終了になる。

[0105] 性能を保証するためのリソースが不足する場合（S19020のYES）には、データ回復による性能劣化を補うために、ディスクアクセスの高速化処理が実施される（S19030）。この高速化処理は、Dynamic ProvisioningやDynamic Tieringなどと呼ばれる処理であって、データの再配置により高速なディスクドライブ1250へデータを移動するなど、障害の発生したデータの回復が高速化されてもよい。

[0106] データ回復処理がない場合（S19010のNO）には、データは壊れてし

まう状況であるので、プロセッサ2010は、障害の発生したディスクドライブ1250へのアクセスを禁止する処理を行う（S19050）。リソースが不足する場合（S19060のYES）には、性能保証フラグが無効な論理区画から未使用リソースが多い順にリソースを借用する処理を行う（S19070）。性能保証フラグが有効な論理区画に性能を保証するリソースが割り当てられない場合（S19080のYES）には、プロセッサ2010は、警告フラグをONにして（S19090）、管理者にその旨を警告する。

[0107] 以上で説明したように、本実施例によれば、障害発生時に、性能保証をしなければならない論理区画が、性能保証をしない論理区画からリソースを借用し、性能保証をしなければならない論理区画の性能を保証することができる。また、性能保証をしなければならない論理区画間でもリソースの借用が可能になる。

[0108] なお、本実施例では、論理区画設定プログラム2060での障害検出を契機にリソースの貸し借りが実施される例を示したが、この処理が物理ストレージ装置1200内で行われてもよい。また、障害検出の契機ではなく、ユーザの指示により実施されてもよいし、ウィルス検出によるデータ障害やデータベース異常の検出を契機としてもよい。

[0109] また、最初から未割当のリソースが存在する場合、リソース不足に陥った論理区画は、未割当のリソースを優先的に借用するようにして、借用できる未割当リソースが無くなった時点で、論理区画間の借用が行われるようしてもよい。

実施例 2

[0110] 実施例1では、図5を用いて説明したように、IO性能（IOPS）に必要なリソースの上限があらかじめ設定されていて、障害時にリソースを貸し借りする処理であった。これに対して、実施例2では、管理サーバ2000が実際のIOの量を監視して、IOPSが性能要件を満たなくなる状況を検出し、監視したIOの量に基づいてリソースを貸し借りすることにより性能を保証する。なお、実施例2は、多くの部分で実施例1と同じ構成であるた

め、以下では異なる構成のみを説明する。

- [0111] 図20は管理サーバ20000の構成の例を示す図である。管理サーバ20000は、図2に示した管理サーバ2000に対して、IOの利用状況を監視して、その情報を管理するIO利用状況管理情報20010をさらに有する。
- [0112] 図21はIO利用状況管理情報20010のテーブル管理情報の例を示す図である。各論理区画のIOPSが測定され、IO利用状況管理情報20010のテーブル管理情報は、その測定結果の平均IOPS21020とMax IOPS21030をテーブルで管理している。この管理のために、テーブル管理情報は性能保証フラグ21000とストレージ装置ID21010を含んでもよい。平均IOPS21020は、通常運用時にどのくらいIOPS性能が確保されていたかを表す。また、Max IOPS21030は、IOアクセス負荷が上がってきた際に、どのくらいまで性能を保証すればよいかを表す。さらに、IOPSの平均と分散値21040または標準偏差の値を算出して管理しておけば、どのようなばらつきでIOアクセスが行われているか、その時のリソース使用率の傾向を表せる。
- [0113] これにより、性能が劣化し始めてきたタイミングを検出することも可能となり、事前に論理区画へリソースを割り当てて、性能を保証することも可能になる。IOPSとリソース使用量の関係の傾向としては、例えば、分散が少ない場合には、現在割当ているリソースの平均の量で性能が確保できていることを示す。この場合、その際のリソース使用量を図5のリソース確保上限管理テーブルの上限として採用してもよい。
- [0114] また、分散が多い場合には、平均量を確保しつつ、その際にそのリソースの使用量の変化量を監視することで、確保しなければいけないリソースが確保できる。確保しなければいけないリソースの特定ができれば、あるタイミングで未使用率が高いリソースであっても、そのリソースは解放せず、その論理区画が確保維持してもよい。
- [0115] 図22は、実施例1の図7に相当する、障害時のリソースの再配置設定の

処理フローの例を示す図である。障害が発生すると、プロセッサ2010は、その障害を検出し（S22000）、その障害の発生したリソースの割当てを禁止する（S22010）。そして、I/O利用状況を監視し（S22020）、I/O性能が性能要件を満たしているかを確認する（S22030）。I/O性能が不足している際の、リソース使用状況を取得する（S22040）。リソース選択（S22090）を除くその他の処理（S22050～S22080、S22100～S22110）は、図7を用いて既に説明した処理フローと同じであるため説明を省略する。リソース選択（S22090）については、図23を用いて後で説明する。

[0116] 実施例1とは、プロセッサ2010が、図5に示したテーブルを参照してリソース確保の上限値を超えているかどうかで判断するのではなく、I/O性能を監視し、そのI/O性能が性能要件を満たしているかどうかを契機に、リソース確保を行う点で異なる。実際のI/O性能を監視することから、I/Oの性能を保証する目的は、直接に達成される。図21に示した平均IOPS21020、Max IOPS21030、IOPS分散値21040を使用して、保証すべきI/O性能の傾向を算出し、事前にI/O性能を確保するためのリソースの再配置が行われてもよい。

[0117] また、リソースを貸し出す論理区画のI/O性能も監視し、リソースを再配置する前の性能と、リソースを再配置後に劣化した性能を取得できる。これにより、性能保証フラグが有効な論理区画のみでなく、性能保証フラグが無効な論理区画において、性能の劣化量が制限されてもよい。

[0118] 図23は、リソース選択の処理フローの例を示す図である。このリソース選択は、図22のS22090の処理である。この処理フローは、基本的には、実施例1の図8を用いて説明した処理フローと同じであるが、未使用リソースの量に基づいてリソースを借用する先を選択するのではなく、I/O利用状況に基づいてI/O利用率の低い論理区画からリソースを借用する点（S23010）で異なる。これは、I/O利用が低いということは、割当てられたリソースをあまり使っていない、つまり未使用リソースが多いことを利用してい

る。

[0119] I/Oを多く利用している論理区画からリソースを借用すると、性能保証フラグが無効とは言え、性能が急激に低下する可能性がある。クラウド環境を前提とした、すべてのユーザが使い易いシステムである場合、急激な性能劣化は避けた方が、ユーザからのクレームが少なくて済む可能性があるので、I/O利用の少ない論理区画からリソースを借用する。

[0120] また、I/O利用状況を監視しているので、I/O利用傾向から事前にI/O利用率が予測されてもよく、性能保証フラグが有効な論理区画のI/O性能が不足し始めたら、事前に性能保証フラグが無効な論理区画を利用しているホスト計算機1000に対して、I/O利用を抑えるように指示を出す（S23030）。これによって、性能保証フラグが無効な論理区画の未使用リソースが多く確保されることとなり、性能保証フラグが有効な論理区画に多くのリソースが割り当てられてもよい。図23に示す処理フローのその他の処理は、図8を用いて説明した処理フローと同じであるので、説明を省略する。

[0121] 以上で説明したように、実施例2によれば、障害発生時に性能保証をしなければならない論理区画が、性能保証をしない論理区画からリソースを借用し、性能保証をしなければならない論理区画の性能を保証することができる。特に、性能を測定して保証するため、正確な性能の保証が可能になる。

符号の説明

- [0122] 1000：ホスト計算機
1200：ストレージ装置
1210：FE PK
1220：CM PK
1230：MP PK
1240：BE PK
1250：ディスクドライブ
1270：論理ボリューム
1500：論理区画

2000：管理サーバ

請求の範囲

- [請求項1] ホスト計算機とストレージ装置と管理計算機から構成される計算機システムであって、
前記ストレージ装置は、
前記ホスト計算機と接続するためのポートと、
キャッシュメモリと、
プロセッサと、
論理的な記憶領域である複数の論理ボリュームと、
を有し、
前記論理ボリュームごとに、前記論理ボリュームの読み書きに使用
される資源として、前記ポートと前記キャッシュメモリと前記プロセ
ッサが論理区分に分割され、
前記ホスト計算機は、前記論理ボリュームに対して読み書きを行い
、
前記管理計算機は、
前記ストレージ装置に障害が発生した場合、読み書きの性能が保証
されない前記論理区画の資源を、読み書きの性能が保証される前記論
理区画へ割当てるように、前記ストレージ装置へ指示を出すこと
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、
前記読み書きの性能が保証されない論理区画と前記読み書きの性能
が保証される論理区画を識別する第1の情報を有し、
前記第1の情報に基づき、読み書きの性能が保証されない前記論理
区画の資源を、読み書きの性能が保証される前記論理区画へ割当てる
ように、前記ストレージ装置へ指示を出すこと
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] 前記管理計算機は、
前記読み書きの性能が保証されない論理区画の資源に障害が発生し

た場合、前記指示を出さず、

前記読み書きの性能が保証される論理区画の資源に障害が発生した場合、前記指示を出すこと

を特徴とする請求項2に記載の計算機システム。

[請求項4]

前記管理計算機は、

前記論理区画単位の資源を使用する量と、前記論理区画単位の前記読み書きの性能との関係の第2の情報を有し、

前記第2の情報と前記保証される読み書きの性能とに基づき、前記読み書きの保証される論理区画の初期資源が割当てられ、

前記発生した障害により使用できない資源の不足量を、前記読み書きの性能が保証されない論理区画から、前記読み書きの性能が保証される論理区画へ割当てるよう、前記ストレージ装置へ指示を出すこと

を特徴とする請求項3に記載の計算機システム。

[請求項5]

前記管理計算機は、

前記読み書きの性能が保証されない論理区画の未使用的資源を、前記読み書きの性能が保証される論理区画へ割当てるよう、前記ストレージ装置へ指示を出すこと

を特徴とする請求項4に記載の計算機システム。

[請求項6]

前記管理計算機は、

前記読み書きの性能が保証されない論理区画の未使用的資源の量が、前記資源の不足量より少ない場合、

前記読み書きの性能が保証されない論理区画の資源の使用量を減らして未使用的資源の量を増やし、前記増やした未使用的資源を、前記読み書きの性能が保証される論理区画へ割当てるよう、前記ストレージ装置へ指示を出すこと

を特徴とする請求項5に記載の計算機システム。

[請求項7]

前記管理計算機は、

前記読み書きの性能が保証される論理区画のポートに障害が発生した場合、

前記ホスト計算機と前記ストレージ装置との間はマルチパス構成であるかを判定し、前記判定がマルチパス構成でないと、マルチパス構成を構築すること

を特徴とする請求項 6 に記載の計算機システム。

[請求項8] 前記ストレージ装置は、2重化された第1のキャッシュメモリと第2のキャッシュメモリとを有し、

前記管理計算機は、

前記読み書きの性能が保証される論理区画のキャッシュデータが含まれる前記第1のキャッシュメモリに障害が発生した場合、

前記第2のキャッシュメモリがライトスルー動作であるか否かに応じて、異なるキャッシュメモリを資源としての割当てるように、前記ストレージ装置へ指示を出すこと

を特徴とする請求項 6 に記載の計算機システム。

[請求項9] 前記管理計算機は、

前記読み書きの性能が保証される論理区画のプロセッサに障害が発生した場合、

前記読み書きの性能が保証されない論理区画のプロセッサを、プロセッサ単位で、前記読み書きの性能が保証される論理区画へ割当てるよう、前記ストレージ装置へ指示を出すこと

を特徴とする請求項 6 に記載の計算機システム。

[請求項10] 前記管理計算機は、

前記ストレージ装置に障害が発生した場合、前記読み書きの性能が保証される論理区画の性能を取得し、前記取得した性能が保証される性能より低いと、前記読み書きの性能が保証されない論理区画の資源を、前記読み書きの性能が保証される論理区画へ割当てるよう、前記ストレージ装置へ指示を出すこと

を特徴とする請求項 1 に記載の計算機システム。

- [請求項11] 前記管理計算機は、
前記読み書きの性能が保証されない論理区画を複数有し、
前記読み書きの性能が保証されない複数の論理区画の中で、前記論
理ボリュームの読み書きの最も少ない論理区画の資源を、前記読み書
きの性能が保証される論理区画へ割当るように、前記ストレージ裝
置へ指示を出すこと
を特徴とする請求項 10 に記載の計算機システム。
- [請求項12] 前記管理計算機は、
前記読み書きの性能が保証されない論理区画に属する論理ボリュー
ムを読み書きする前記ホスト計算機へ読み書きの制限を指示すること
を特徴とする請求項 11 に記載の計算機システム。
- [請求項13] ホスト計算機と接続されるストレージ装置であつて、
前記ホスト計算機と接続するためのポートと、
キャッシュメモリと、
プロセッサと、
論理的な記憶領域である複数の論理ボリュームと、
を有し、
前記論理ボリュームごとに、前記論理ボリュームの読み書きに使用
される資源として、前記ポートと前記キャッシュメモリと前記プロセ
ッサが論理区分に分割され、
前記ストレージ装置に障害が発生した場合、読み書きの性能が保証
されない前記論理区画の資源を、読み書きの性能が保証される前記論
理区画へ割当てを行うこと
を特徴とするストレージ装置。
- [請求項14] 前記読み書きの性能が保証されない論理区画と前記読み書きの性能
が保証される論理区画を識別する情報を有し、
前記情報に基づき、読み書きの性能が保証されない前記論理区画の

資源を、読み書きの性能が保証される前記論理区画へ割当てを行うこと

を特徴とする請求項13に記載のストレージ装置。

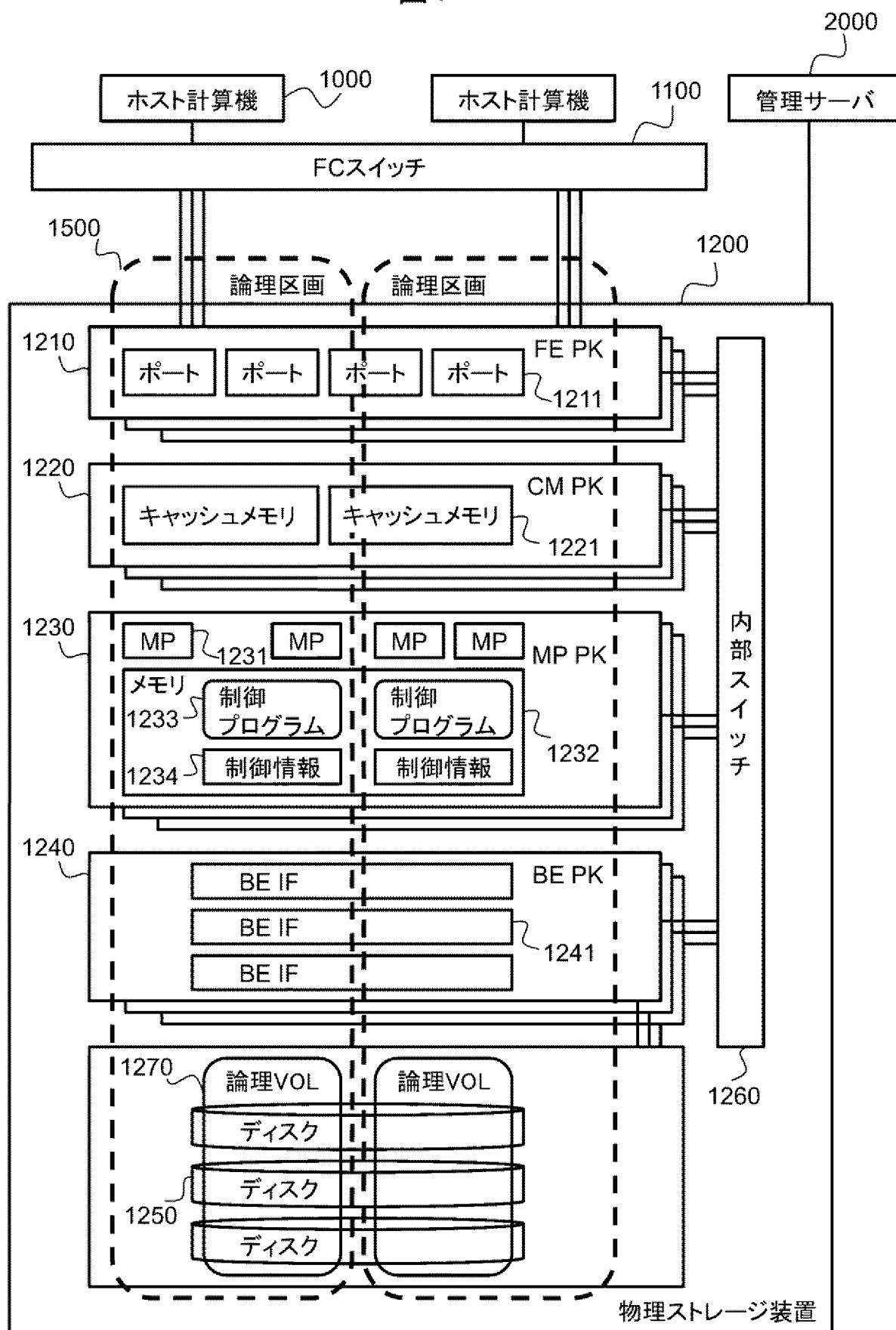
[請求項15] 前記読み書きの性能が保証されない論理区画の資源に障害が発生した場合、前記割当てを行い、

前記読み書きの性能が保証される論理区画の資源に障害が発生した場合、前記割当てを行わないこと

を特徴とする請求項14に記載のストレージ装置。

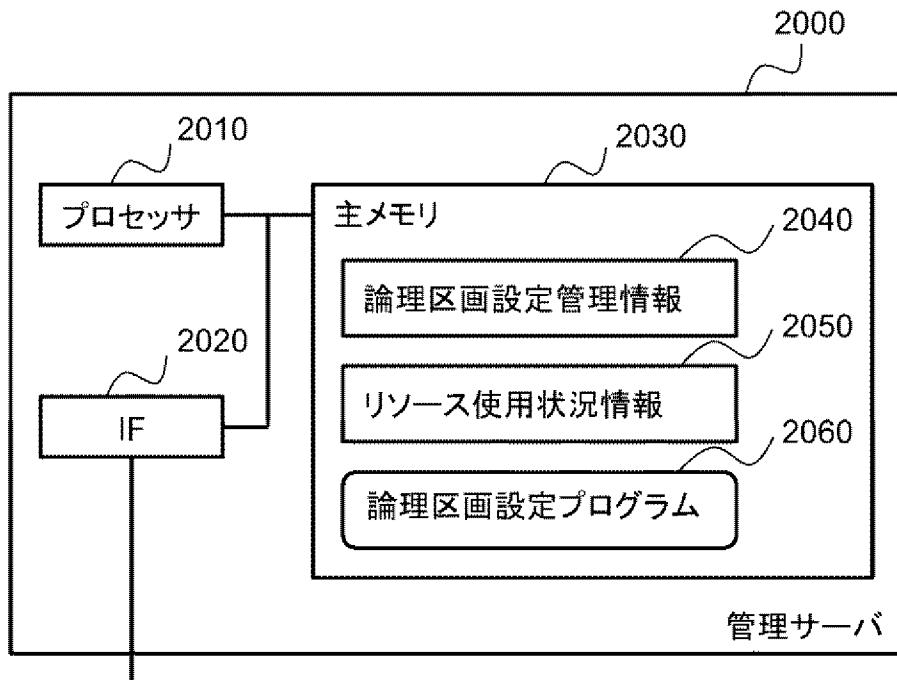
[図1]

図1



[図2]

図2



[図3]

図3

Figure 3 is a table showing resource configuration. The columns are: Storage Device ID, Resource Type, Resource ID, Performance-Capacity, and Failure-time Constraint. The rows represent different resources grouped by their type.

ストレージ装置ID	リソース種別	リソースID	性能・容量	障害時制約
100	MP_Core	MP1_Core#a	5220MIPS	なし
		MP1_Core#b	5220MIPS	
		...		
	キャッシュメモリ	RAM1	8GB	ライトスルーアクション
		RAM2	8GB	
		...		
	FE ポート	FE Port1	8Gbps	なし
		FE Port2	8Gbps	
		...		
	BE IF	BE Port1	8Gbps	なし
BE Port2		8Gbps		
...				
HDD	HDD1	160GB	データ回復処理	
	HDD2	160GB		
	...			
101	...			

[図4]

図4

Figure 4 is a table showing performance guarantees. The columns are: Logical Area ID, Failure-time Performance Guarantee Flag, and Performance Requirements.

論理区画ID	障害時性能保証フラグ	性能要件
VPS1	1	1000IOPS
VPS2	1	800IOPS
VPS3	0	—
...		

[図5]

図5

性能要件 (IOPS)	リソース確保量			
	Port	MP	Cache	Disk
...				
200	0.3	0.5	200MB	160GB
...				
8000	4	4	2GB	2TB
8050	5	4	2GB	2TB
...				
...				

[図6]

図6

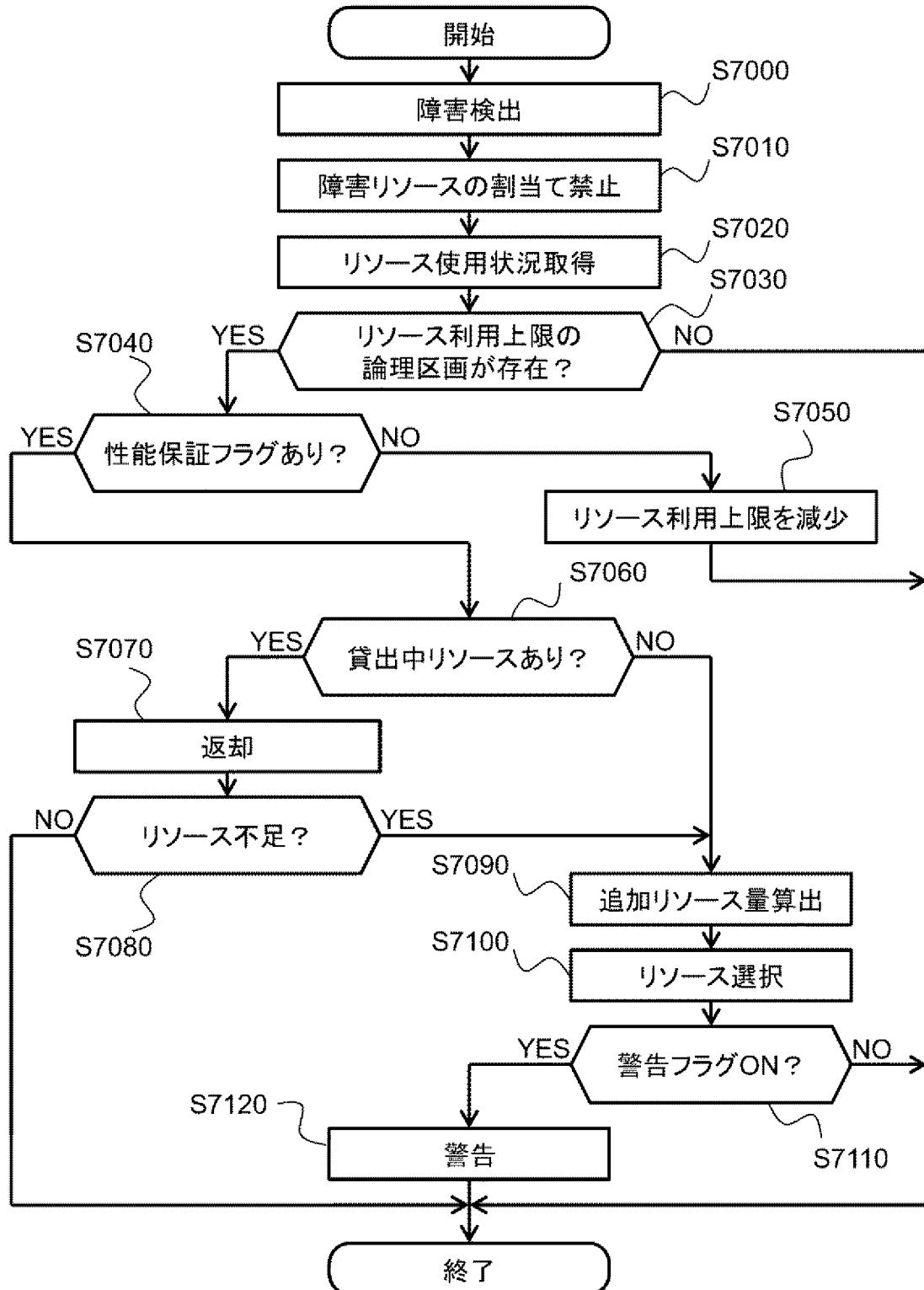
The diagram illustrates the mapping of table columns to specific identifiers:

- Column 1 (論理区画ID) is connected to 6000.
- Column 2 (ストレージ装置ID) is connected to 6010.
- Column 3 (リソース種別) is connected to 6020.
- Column 4 (リソースID) is connected to 6030.
- Column 5 (割り当て率・アドレス) is connected to 6040.
- Column 6 (使用率・使用状況・障害) is connected to 6050.

論理区画ID	ストレージ装置ID	リソース種別	リソースID	割り当て率・アドレス	使用率・使用状況・障害
VPS1	100	MP_Core	MP1_Core#a	30%	20%
			MP1_Core#b	30%	20%
			...		
		キャッシュメモリ	RAM1	0	W/R Cache
			RAM1	4096	W/R Cache
			RAM2	8192	R. C. Buffer
			...		
		FE ポート	FE Port1	30%	20%
			FE Port2	30%	障害
		
		BE IF	BE Port1	30%	20%
			BE Port1	30%	20%
			...		
		HDD	HDD1	160GB	20%
			HDD2	160GB	20%
			...		
VPS2	100	...			

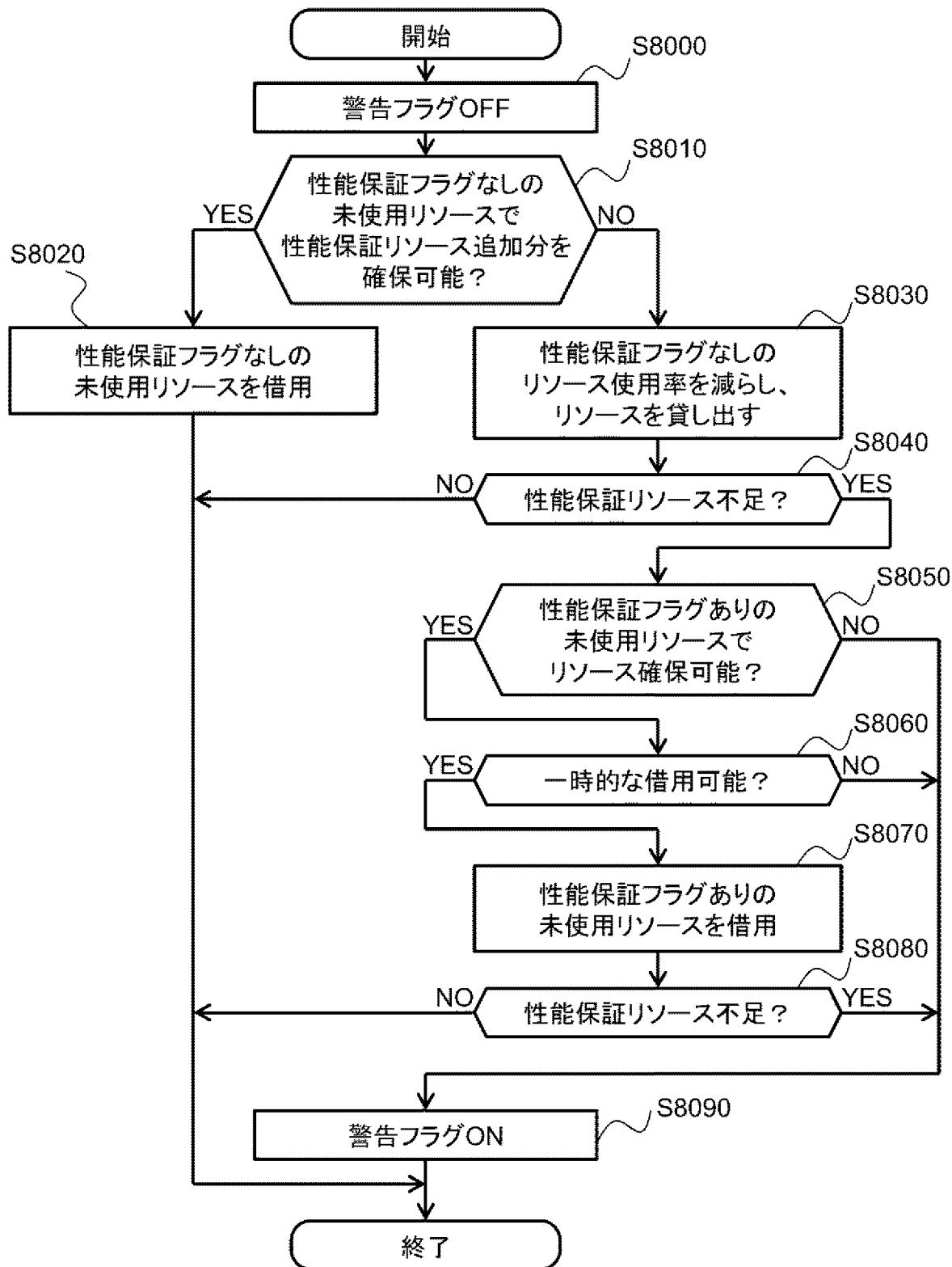
[図7]

図7



[図8]

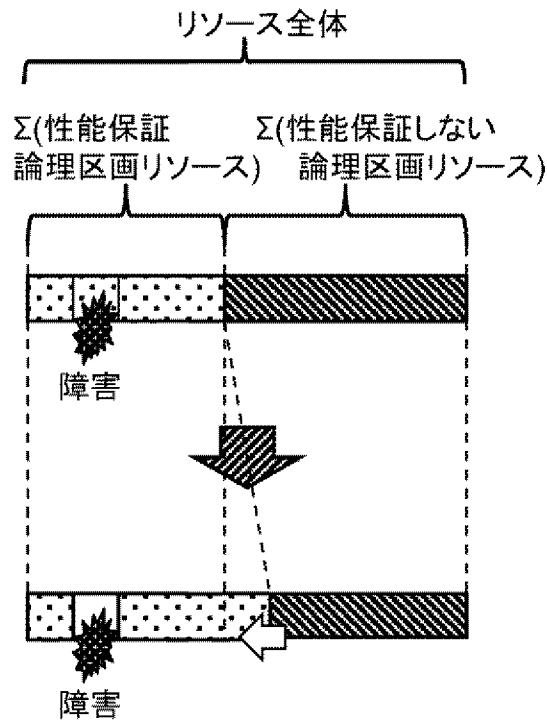
図8



[図9A]

図9A

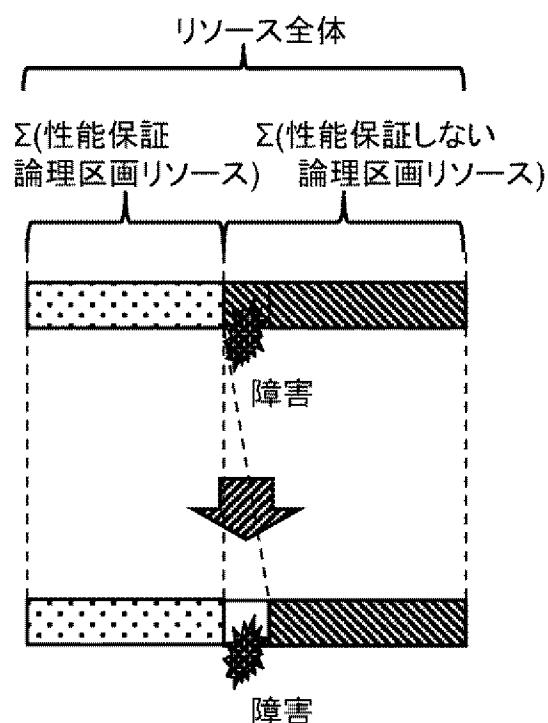
性能保証フラグ設定ありに
割当てられているリソースの障害



[図9B]

図9B

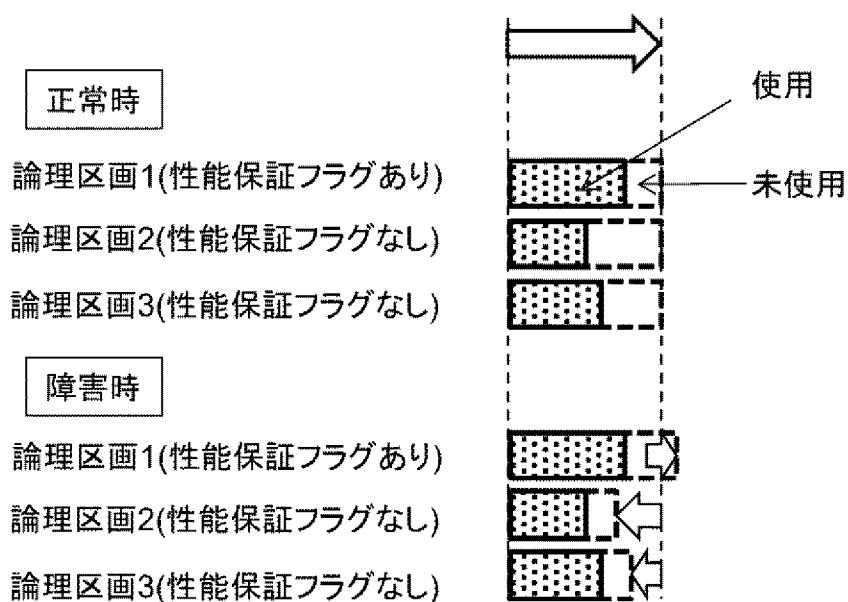
性能保証フラグなしに
割当てられるリソースの障害



[図10]

図10

論理区画のリソース上限



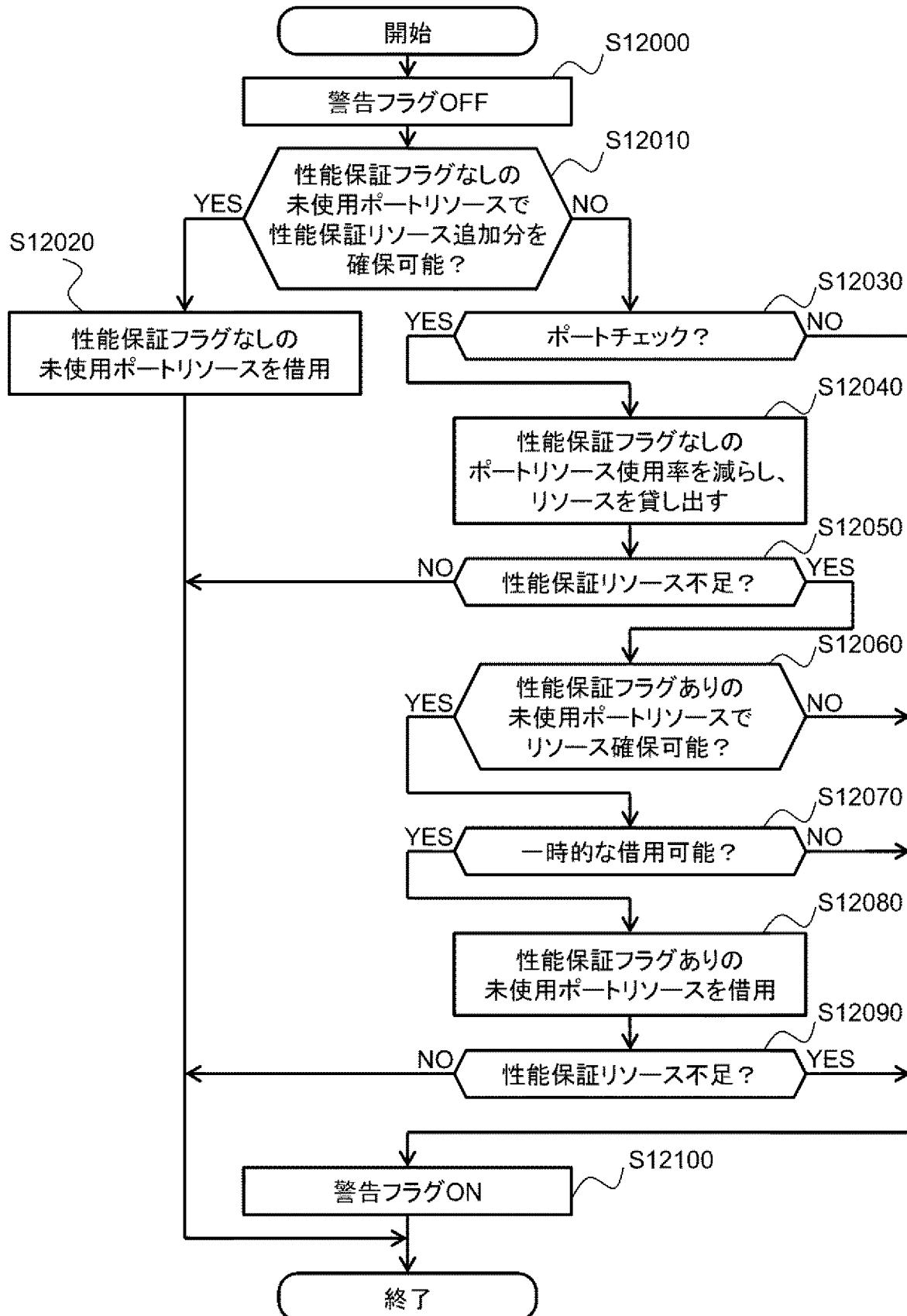
[図11]

図11

貸出元 論理区画	性能保証 フラグ	ストレージ 装置ID	貸出 リソース	貸出 可能量	障害時 使用制約
VPS1	1	100	Port #A-1	0%	なし
VPS2	1		Port #A-2	5%	なし
VPS3	0		Port #A-3	6%	なし
VPS4	0		Port #A-4	10%	なし
VPS5	0		Port #A-5	15%	なし
...			Port #A-6	40%	なし
		101	Port #B-1	20%	なし
			Port #B-2	50%	なし

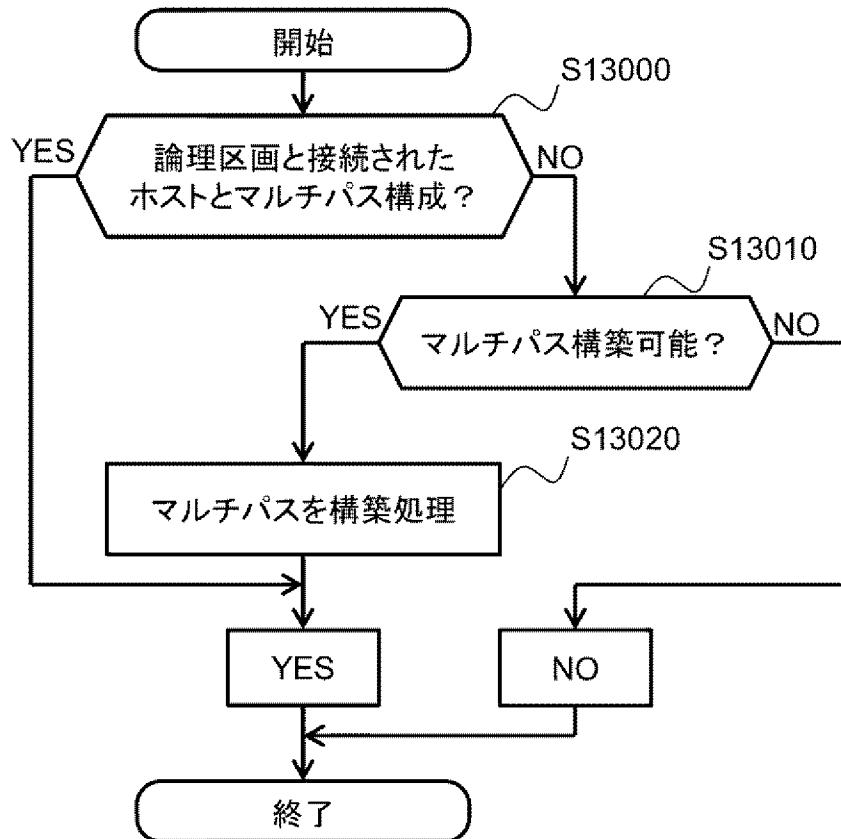
[図12]

図12



[図13]

図13



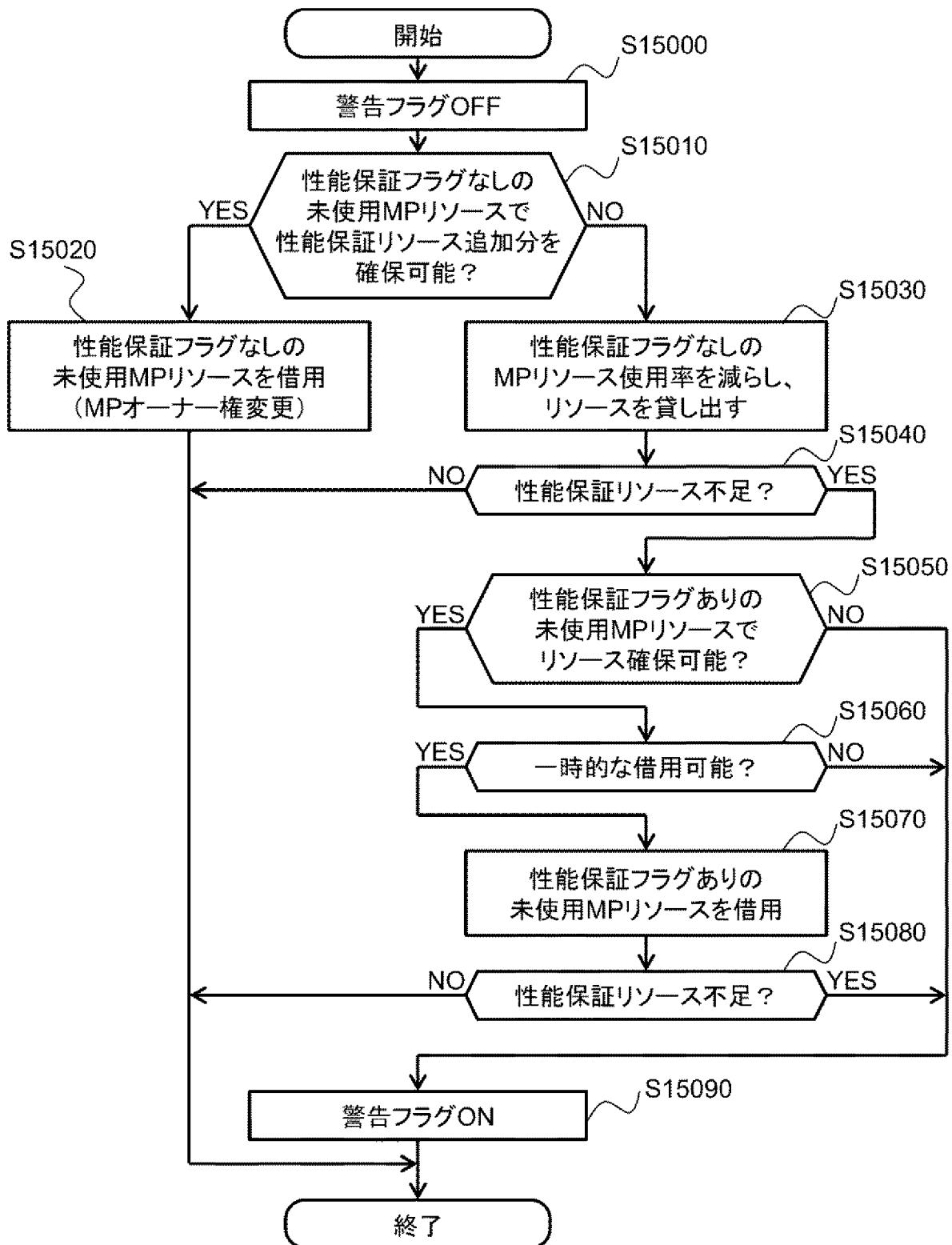
[図14]

図14

貸出元 論理区画	性能保証 フラグ	ストレージ 装置ID	貸出 リソース	貸出 可能量	障害時制約
VPS1	1	100	MP1_Core#a	0%	オーナー権固定
VPS2	1		MP1_Core#b	5%	なし
VPS3	0		MP1_Core#c	3%	なし
VPS4	0	101	MP2_Core#a	35%	なし
VPS5	0		MP2_Core#b	35%	なし
...			MP2_Core#c	35%	なし
			MP3_Core#a	35%	なし
			MP3_Core#b	45%	なし

[図15]

図15



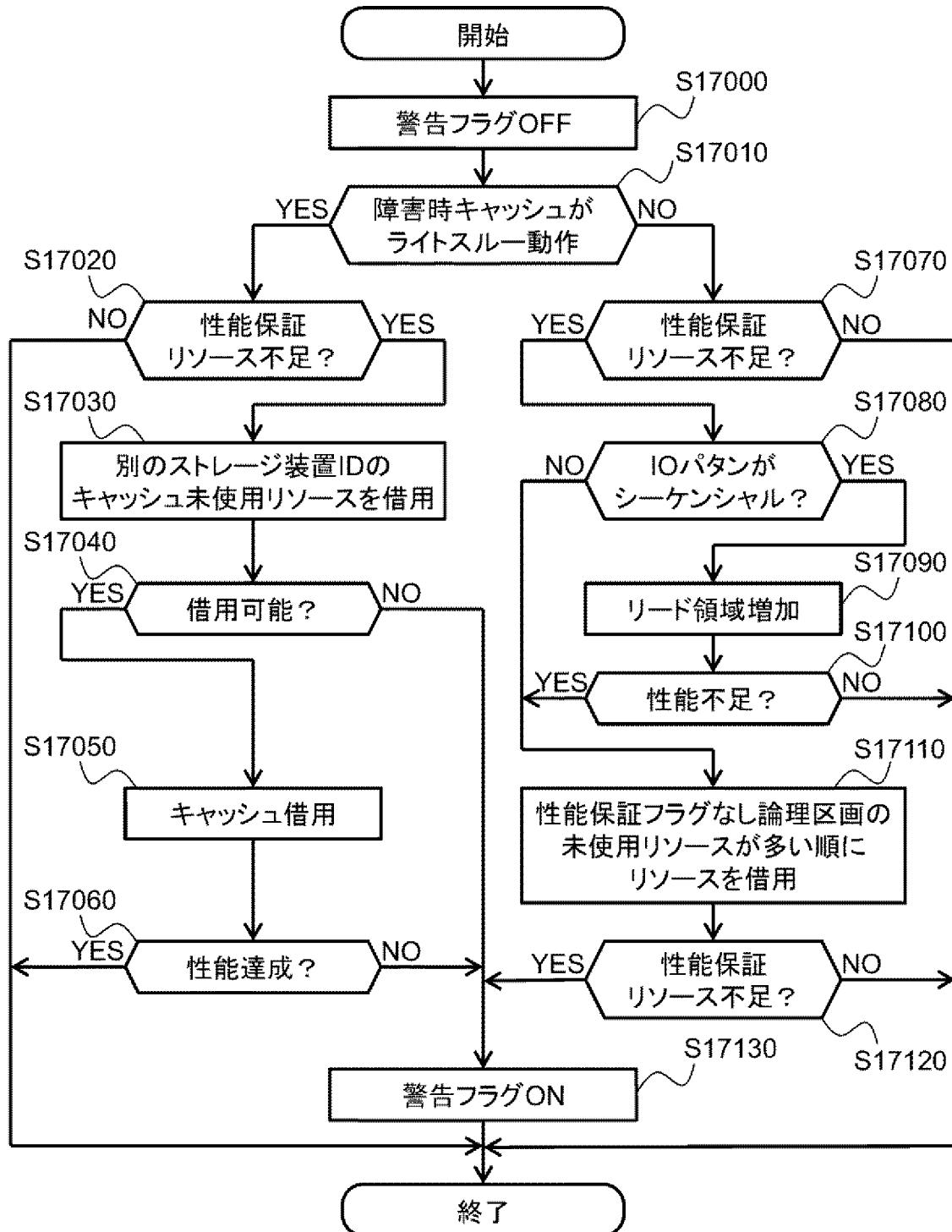
[図16]

図16

貸出元 論理区画	性能保証 フラグ	ストレージ 装置ID	貸出 リソース	貸出 可能量	障害時制約
VPS1	1	100	Cache #A 0-1023	0%	障害時は ライトスルー
VPS2	1		Cache #A 1024-2047	0%	
			Cache #A 2048-4095	0%	
VPS3	0		Cache #A	0%	
VPS4	0		Cache #A	0%	
VPS5	0		Cache #B	20%	
...		101	Cache #B	50%	

[図17]

図17



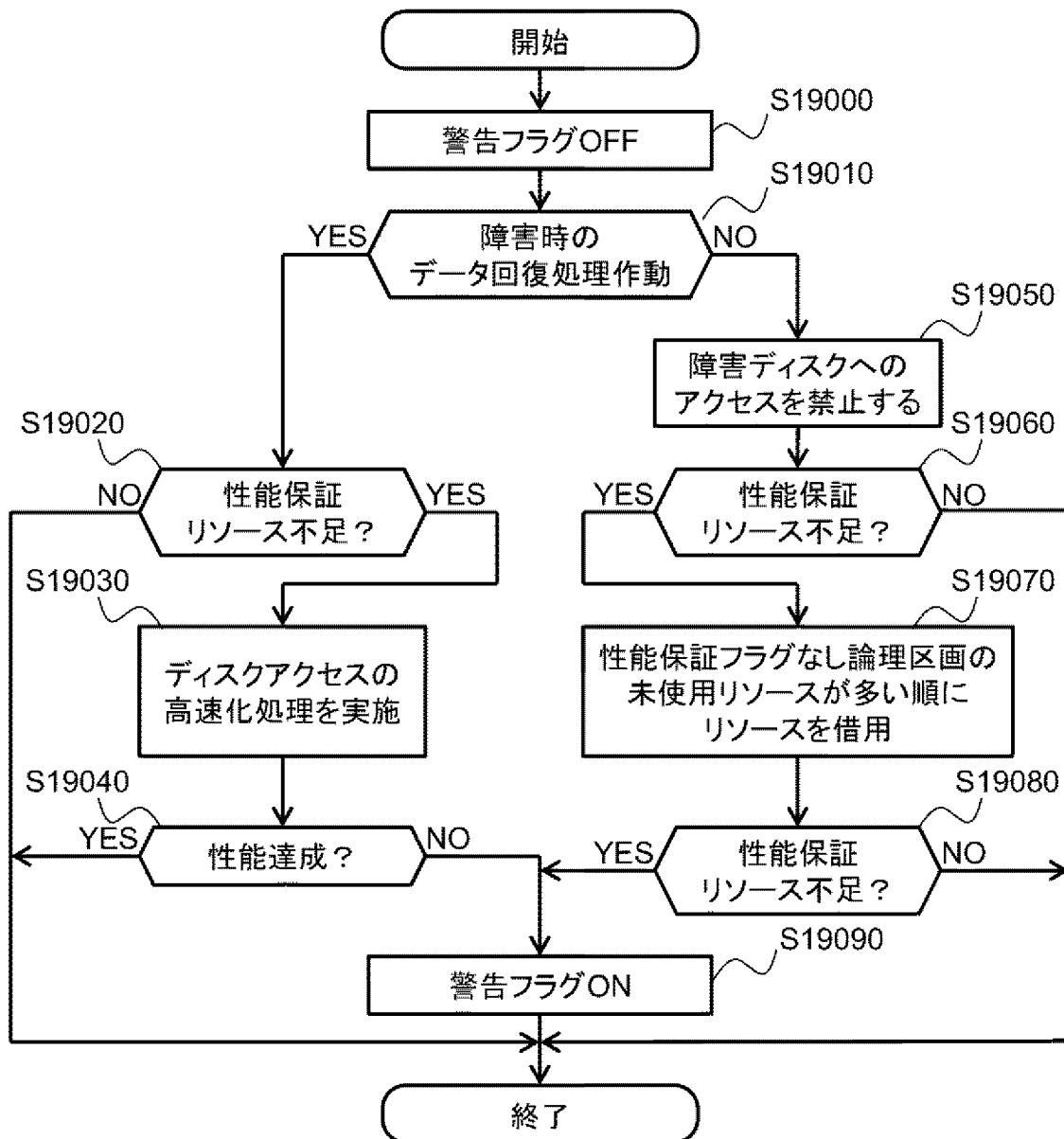
[図18]

図 18

貸出元 論理区画	性能保証 フラグ	ストレージ 装置ID	貸出 リソース	貸出 可能量	障害時制約
VPS2	1	100	HDD#1	25%	データ回復 処理
VPS3	0		SSD#1	10%	なし
...					
...		101			

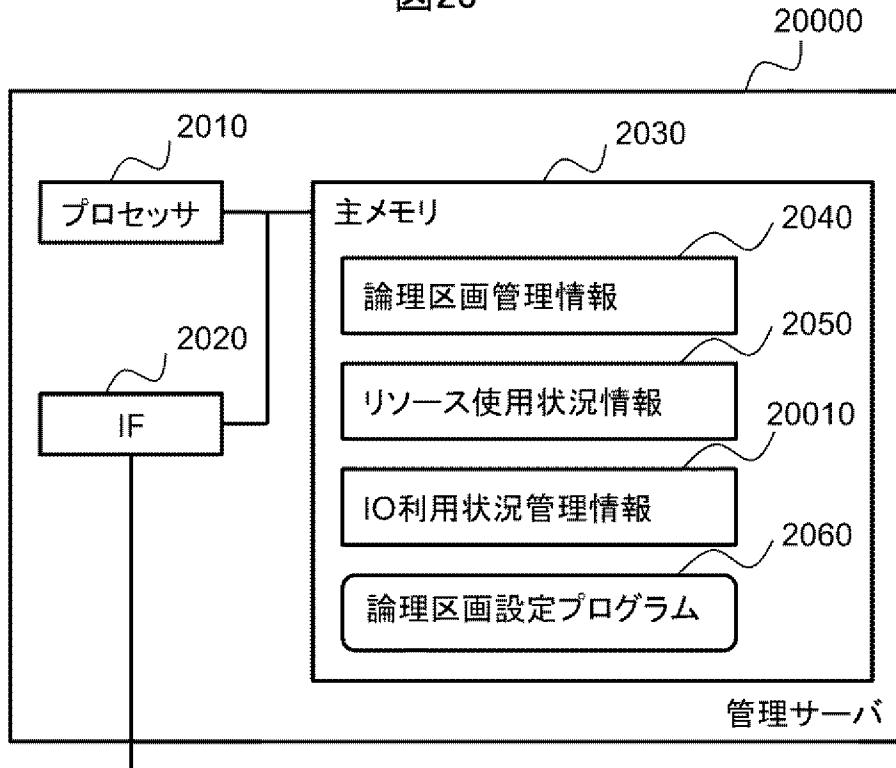
[図19]

図19



[図20]

図20



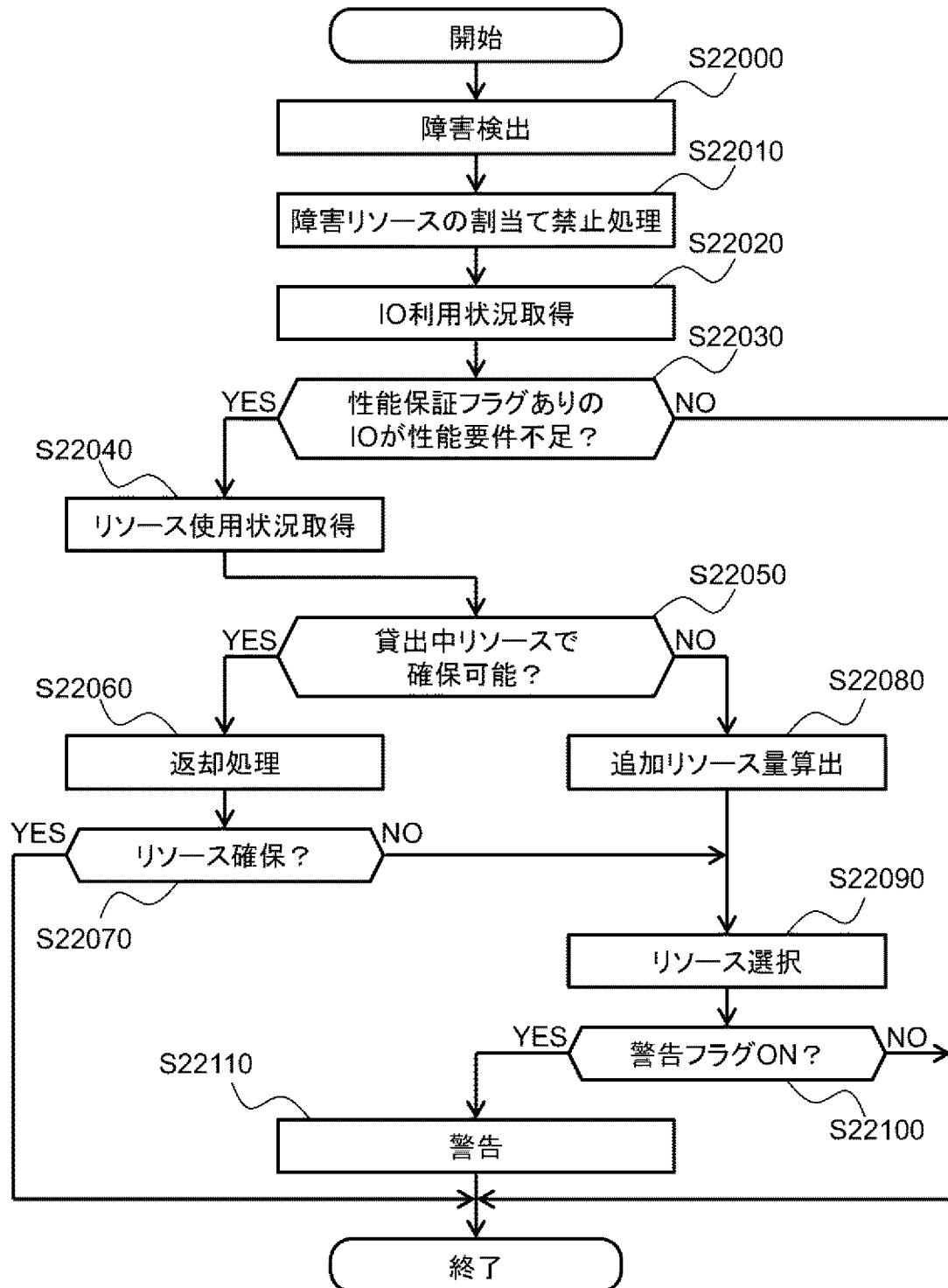
[図21]

図21

論理区画ID	障害時性能保証フラグ	ストレージ装置ID	平均IOPS(測定値)	Max IOPS(測定値)	IOPS分散値
VPS 1	1	100	550	800	1.5
VPS 2	1	100	120	150	10.2
VPS 3	0	100	200	300	0.5
VPS 4	0	101	15	30	2.7
...					

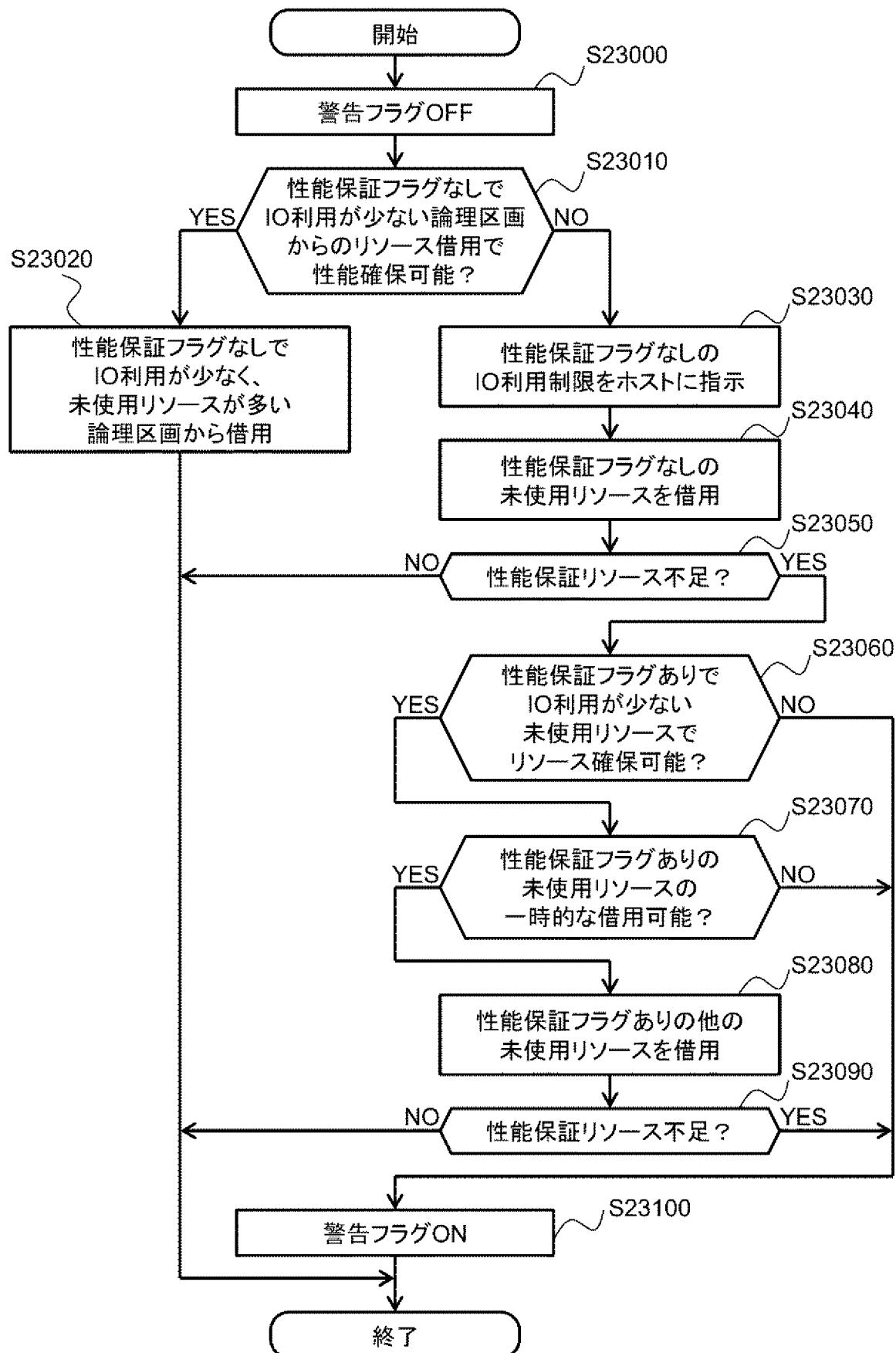
[図22]

図22



[図23]

図23



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/079986

A. CLASSIFICATION OF SUBJECT MATTER
G06F3/06(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)
G06F3/06

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 1996-2015
 Kokai Jitsuyo Shinan Koho 1971-2015 Toroku Jitsuyo Shinan Koho 1994-2015

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.
A	WO 2011/108027 A1 (Hitachi, Ltd.), 09 September 2011 (09.09.2011), paragraphs [0047] to [0054] & US 2011/0219271 A1	1-15
A	JP 2004-246852 A (Hitachi, Ltd.), 02 September 2004 (02.09.2004), paragraph [0012] & US 2004/0123180 A1	1-15
A	JP 2012-221340 A (Fujitsu Ltd.), 12 November 2012 (12.11.2012), paragraphs [0011] to [0028] & US 2012/0265907 A1	1-15

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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"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	"&" document member of the same patent family

Date of the actual completion of the international search
02 February 2015 (02.02.15)

Date of mailing of the international search report
10 February 2015 (10.02.15)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/079986

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2006-285808 A (Hitachi, Ltd.), 19 October 2006 (19.10.2006), paragraphs [0006] to [0010] & US 2006/0224854 A1	1-15

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F3/06(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（IPC））

Int.Cl. G06F3/06

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	1922-1996年
日本国公開実用新案公報	1971-2015年
日本国実用新案登録公報	1996-2015年
日本国登録実用新案公報	1994-2015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
A	WO 2011/108027 A1 (株式会社日立製作所) 2011.09.09, 段落[0047]-[0054] & US 2011/0219271 A1	1-15
A	JP 2004-246852 A (株式会社日立製作所) 2004.09.02, 段落【0012】 & US 2004/0123180 A1	1-15
A	JP 2012-221340 A (富士通株式会社) 2012.11.12, 段落【0011】 - 【0028】 & US 2012/0265907 A1	1-15

 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」国際出願目前で、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

02.02.2015

国際調査報告の発送日

10.02.2015

国際調査機関の名称及びあて先

日本国特許庁（ISA/JP）

郵便番号100-8915

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）

木村 雅也

5 T 3980

電話番号 03-3581-1101 内線 3568

C (続き) . 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
A	JP 2006-285808 A (株式会社日立製作所) 2006.10.19, 段落【0006】 - 【0010】 & US 2006/0224854 A1	1-15