

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2008-59478
(P2008-59478A)

(43) 公開日 平成20年3月13日(2008.3.13)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 3/06 (2006.01)	G06F 3/06 301E	5B014
G06F 13/10 (2006.01)	G06F 13/10 340A	5B065

審査請求 未請求 請求項の数 10 O L (全 18 頁)

(21) 出願番号	特願2006-238150 (P2006-238150)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成18年9月1日(2006.9.1)	(74) 代理人	100079108 弁理士 稲葉 良幸
		(74) 代理人	100093861 弁理士 大賀 眞司
		(72) 発明者	山口 寧彦 神奈川県小田原市中里322番2号 株式会社日立製作所RAIDシステム事業部内
		(72) 発明者	杉本 定広 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内
		Fターム(参考)	5B014 EB04 GD31 5B065 BA01 CE01

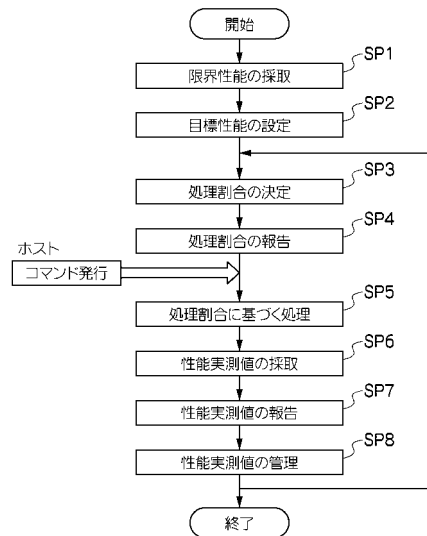
(54) 【発明の名称】 記憶システム及びデータ入出力制御方法

(57) 【要約】

【解決課題】優先したいデータの入出力性能を保障し得る信頼性の高い記憶システム及びデータ入出力制御方法を提案する。

【解決手段】ストレージ装置におけるデータ入出力処理の限界性能を検出し、検出した限界性能と、予め指定されたデータ入出力処理を優先すべきターゲットとなる優先ターゲットごとの予め設定されたデータ入出力処理の目標性能とに基づいて、優先ターゲット及び他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定し、ストレージ装置が、優先ターゲット及び非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する処理割合に応じた頻度で実行するようにした。

【選択図】 図5



【特許請求の範囲】**【請求項 1】**

上位装置から与えられるデータ入出力要求に応じたデータ入出力処理を行うストレージ装置を有する記憶システムにおいて、

前記ストレージ装置における単位時間当たりの限界データ入出力処理数である限界性能を検出する限界性能検出部と、

前記限界性能検出部により検出された前記限界性能と、I/O処理要求の種別であるターゲットのうちデータ入出力処理を優先すべきターゲットとなる優先ターゲット毎の予め設定された単位時間当たりの目標データ入出力処理数である目標性能と、に基づいて、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する処理割合決定部と、

を備え、

前記ストレージ装置は、

前記優先ターゲット及び前記非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行する

ことを特徴とする記憶システム。

【請求項 2】

各前記ターゲットは、

該I/O処理要求を発行したアプリケーション毎又は、データの読み書きの対象となるデバイス毎である

ことを特徴とする請求項 1 に記載の記憶システム。

【請求項 3】

前記優先ターゲット及び前記非優先ターゲット群ごとの単位時間あたりのデータ入出力処理回数をそれぞれ計測する計測部

を備え、

前記処理割合決定部は、

前記計測部により計測された前記優先ターゲット及び前記非優先ターゲット群の単位時間あたりのデータ入出力処理回数の実測値である前記優先ターゲット及び前記非優先ターゲットごとの性能実測値と、前記限界性能検出部により検出された前記限界性能と、前記優先ターゲットごとの前記目標性能とに基づいて、周期的に、前記優先ターゲット及び前記非優先ターゲット群ごとの新たな前記処理割合をそれぞれ決定する

ことを特徴とする請求項 1 に記載の記憶システム。

【請求項 4】

前記処理割合決定部は、

前記優先ターゲット又は前記非優先ターゲット群の前記性能実測値が予め定められた下限値以下であったときには、前記限界性能検出部により検出された前記限界性能と、前記優先ターゲットごとの前記目標性能とに基づいて、前記優先ターゲット及び前記非優先ターゲット群ごとの新たな前記処理割合を決定する

ことを特徴とする請求項 3 に記載の記憶システム。

【請求項 5】

前記処理割合決定部は、

前記優先ターゲット及び前記非優先ターゲット群ごとの前記処理割合を、時間帯ごとに当該時間帯について予め定められた前記処理割合にそれぞれ変更する

ことを特徴とする請求項 1 に記載の記憶システム。

【請求項 6】

上位装置から与えられるデータ入出力要求に応じたデータ入出力処理を行うストレージ装置のデータ入出力を制御するデータ入出力制御方法において、

前記ストレージ装置における単位時間当たりの限界データ入出力処理数である限界性能を検出する第 1 のステップと、

検出された前記限界性能と、I/O処理要求の種別であるターゲットのうちデータ入出

10

20

30

40

50

力処理を優先すべきターゲットでなる優先ターゲット毎の予め設定された単位時間当たりの目標データ入出力処理数である目標性能と、に基づいて、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する第2のステップと、

前記ストレージ装置が、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行する第3のステップと、

を備えることを特徴とするデータ入出力制御方法。

【請求項7】

各前記ターゲットは、

該I/O処理要求を発行したアプリケーション毎又は、データの読み書きの対象となるデバイス毎である

ことを特徴とする請求項6に記載のデータ入出力制御方法。

【請求項8】

前記第3のステップでは、

前記優先ターゲット及び前記非優先ターゲット群ごとの単位時間あたりのデータ入出力処理回数をそれぞれ計測し、

計測した前記優先ターゲット及び前記非優先ターゲット群の単位時間あたりのデータ入出力処理回数の実測値でなる前記優先ターゲット及び前記非優先ターゲットごとの性能実測値と、前記限界性能検出部により検出された前記限界性能と、前記優先ターゲットごとの前記目標性能とに基づいて、周期的に、前記優先ターゲット及び前記非優先ターゲット群ごとの新たな前記処理割合をそれぞれ決定する第4のステップを備える

ことを特徴とする請求項6に記載のデータ入出力制御方法。

【請求項9】

前記第4のステップでは、

前記優先ターゲット又は前記非優先ターゲット群の前記性能実測値が予め定められた下限値以下であったときには、前記限界性能検出部により検出された前記限界性能と、前記優先ターゲットごとの前記目標性能とに基づいて、前記優先ターゲット及び前記非優先ターゲット群ごとの新たな前記処理割合を決定する

ことを特徴とする請求項8に記載のデータ入出力制御方法。

【請求項10】

前記第2のステップでは、

前記優先ターゲット及び前記非優先ターゲット群ごとの前記処理割合を、時間帯ごとに当該時間帯について予め定められた前記処理割合にそれぞれ変更する

ことを特徴とする請求項6に記載のデータ入出力制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、記憶システム及びデータ入出力制御方法に関し、例えばホスト装置がアプリケーションプログラムに基づき各種業務の処理を実行する記憶システムに適用して好適なものである。

【背景技術】

【0002】

従来、ストレージ装置においては、ホスト装置からのデータ入出力要求に対するデータ入出力処理は、そのホスト装置に接続されたポートに対応したプロセッサが行なう。このため、優先度の高い業務及び優先度の低い業務のデータ入出力処理が同一のプロセッサにより行われる場合がある。このような状況では、優先度の低い業務の負荷が大きい場合に、その影響を受けて優先度の高い業務の帯域が確保されないことがあった。

【0003】

かかる問題を解決するための方法として、従来、単一のストレージ装置において、ポー

10

20

30

40

50

トや、W W N (World Wide Name) 又はデバイスを単位とするデータ入出力処理の上限値又は閾値をユーザが設定し、非優先のホスト装置のデータの入出力を抑止し、優先したいホスト装置のデータの入出力の帯域を確保する方法が提案されている(特許文献1参照)

。

【0004】

【特許文献1】特開2002-108567号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

ところが、かかる特開2002-108567号公報に開示された方法は、非優先のデータの入出力を抑止するのみで、優先したいデータの入出力性能が保障されない問題があった。

10

【0006】

本発明は以上の点を考慮してなされたもので、優先したいデータの入出力性能を保障し得る信頼性の高い記憶システム及びデータ入出力制御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】

【0007】

かかる課題を解決するため本発明においては、上位装置から与えられるデータ入出力要求に応じたデータ入出力処理を行うストレージ装置を有する記憶システムにおいて、前記ストレージ装置における単位時間当たりの限界データ入出力処理数である限界性能を検出する限界性能検出部と、前記限界性能検出部により検出された前記限界性能と、I/O処理要求の種別であるターゲットのうちデータ入出力処理を優先すべきターゲットでなる優先ターゲット毎の予め設定された単位時間当たりの目標データ入出力処理数である目標性能と、に基づいて、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する処理割合決定部と、を備え、前記ストレージ装置は、前記優先ターゲット及び前記非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行することを特徴とする。

20

【0008】

この結果、この記憶システムでは、優先ターゲット及び前記非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行することができる。

30

【0009】

また本発明においては、上位装置から与えられるデータ入出力要求に応じたデータ入出力処理を行うストレージ装置のデータ入出力を制御するデータ入出力制御方法において、前記ストレージ装置における単位時間当たりの限界データ入出力処理数である限界性能を検出する第1のステップと、検出された前記限界性能と、I/O処理要求の種別であるターゲットのうちデータ入出力処理を優先すべきターゲットでなる優先ターゲット毎の予め設定された単位時間当たりの目標データ入出力処理数である目標性能と、に基づいて、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する第2のステップと、前記ストレージ装置が、前記データ入出力処理を優先ターゲット及びその他のターゲットからなる非優先ターゲット群ごとの処理割合をそれぞれ決定する処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行する第3のステップと、を備えることを特徴とする。

40

【0010】

この結果、このデータ入出力制御方法では、優先ターゲット及び前記非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行することができる。

【0011】

50

この結果、このデータ入出力制御方法によれば、優先ターゲット及び前記非優先ターゲット群についてのデータ入出力処理を、それぞれ処理割合決定部により決定された対応する前記処理割合に応じた頻度で実行することができる。

【発明の効果】

【0012】

本発明によれば、優先したいデータの入出力性能を保障し得る信頼性の高い記憶システム及びデータ入出力制御方法を実現できる。

【発明を実施するための最良の形態】

【0013】

以下、各図を参照しながら本発明の一実施形態について説明する。

10

【0014】

(1) 記憶システムの全体構成

図1において、1は全体として本実施の形態による記憶システムを示す。この記憶システム1は、ホスト装置2がスイッチ3を介してストレージ装置4に接続されると共に、当該ストレージ装置4にIP(Internet Protocol)ネットワーク5を介して管理サーバ6が接続されることにより構成されている。

【0015】

上位装置としてのホスト装置2は、CPU(Central Processing Unit)10、メモリ11、バス制御部13等の情報処理資源を備えたコンピュータ装置であり、パーソナルコンピュータや、ワークステーション、メインフレームなどから構成される。ホスト装置2は、例えばキーボード、スイッチやポインティングデバイス、マイクロフォン等の情報入力装置(図示せず)と、モニタディスプレイやスピーカ等の情報出力装置(図示せず)とを備える。

20

【0016】

CPU10は、ホスト装置2全体の動作制御を司るプロセッサである。またメモリ11は、ユーザ業務で使用されるアプリケーションプログラムを保持するために使用されるほか、CPU10のワークメモリとしても用いられる。このメモリ11に保持されたアプリケーションプログラムをCPU10が実行することにより、ホスト装置2全体として対応する業務に関する各種処理が行われる。バス制御部13は、バスを複数使用する場合に、全てのバスに均等に処理コマンドを発行する制御を行う。

30

【0017】

スイッチ3は、ファイバーチャネルネットワークを構成するもので、複数のファイバーチャネルスイッチから構成される。このスイッチ3を介したホスト装置2及びストレージ装置4間の通信は、ファイバーチャネルプロトコルに従って行われる。

【0018】

ストレージ装置4は、図1に示すように、データを記憶するための記憶領域を提供するディスクデバイス部30と、ディスクデバイス部30に対するデータの入出力を制御する複数のコントロール部(DKC)20とから構成される。

【0019】

ディスクデバイス部30は、複数のディスクデバイス(図示せず)及び複数のポート31を備える。ディスクデバイスは、例えばSCSI(Small Computer System Interface)ディスク等の高価なディスクドライブ、及びSATA(Serial AT Attachment)ディスクや光ディスク等の安価なディスクドライブから構成される。

40

【0020】

これらディスクデバイスは、コントロール部20によりRAID方式で運用される。1又は複数のディスクデバイスにより提供される物理的な記憶領域上に、1又は複数の論理デバイス(LDEV)32が構成される。そしてホスト装置2からのデータは、この論理デバイス32内に所定の大きさのブロック(以下、これを論理ブロックと呼ぶ)を単位として記憶される。

【0021】

50

各論理デバイス 3 2 には、それぞれ固有の識別番号が付与される。本実施の形態の場合、データの入出力は、この識別番号と、各論理ブロックにそれぞれ付与されるその論理ブロックに固有の番号（以下、これを L B A (Logical Block Address) と呼ぶ）との組み合わせたものをアドレスとして、当該アドレスを指定して行われる。

【 0 0 2 2 】

一方、コントロール部 2 0 は、複数のチャネルプロセッサ (CHP) 2 1、複数のディスク側ポート 2 2 及び管理端末 (SVP) 2 4 を備えて構成される。

【 0 0 2 3 】

各チャネルプロセッサ 2 1 は、それぞれマイクロプロセッサ、メモリ及び通信インタフェース等を備えたマイクロコンピュータシステムとして構成されており、それぞれスイッチ 3 に接続するためのホスト側ポート 2 2 A を備える。チャネルプロセッサ 2 1 は、ホスト装置 2 から送信される各種コマンドを解釈して、必要な処理を実行する。各チャネルアダプタ 2 1 のホスト側ポート 2 2 A には、それぞれを識別するためのネットワークアドレス（例えば、IP アドレスや WWN）が割り当てられており、これにより各チャネルプロセッサ 2 1 がそれぞれ個別に N A S (Network Attached Storage) 装置として振る舞うことができるようになっている。

10

【 0 0 2 4 】

ディスク側ポート 2 2 B は、例えばファイバチャネル規格に準拠したポートであり、ディスクデバイス部 3 0 の対応するポート 3 1 とケーブルを介して接続される。コントロール部 2 0 におけるディスクデバイス部 3 0 との間でのデータ授受は、このディスク側ポート 2 2 B を介して行なわれる。

20

【 0 0 2 5 】

また管理端末 2 4 は、ストレージ装置 4 を保守又は管理するための端末装置である。記憶システム 1 の保守員は、この管理端末 2 4 を操作することにより、例えば、ディスクデバイス部 2 0 におけるディスクデバイスの増設等、R A I D 構成や E C C (Error Correcting Code) グループ構成の設定変更等を行うことができる。

【 0 0 2 6 】

管理端末 2 4 は、メモリ 2 5 や C P U 2 6 等の情報処理資源を備えたマイクロコンピュータシステムとして構成される。メモリ 2 5 は、各種制御プログラムや各種テーブルを記憶保持するために使用されるほか、C P U 2 6 のワークメモリとしても用いられる。後述の目標性能設定テーブル 5 0 もこのメモリ 2 5 に記憶保持される。メモリ 2 5 に格納された制御プログラムを C P U 2 6 が実行することにより、管理端末 2 4 全体として種々の処理を実行する。

30

【 0 0 2 7 】

管理サーバ 6 は、ストレージ装置 4 の監視及び管理を行うためのサーバであり、稼働情報監視部 7 及び負荷制御部 8 から構成される。これら稼働情報監視部 7 及び負荷制御部 8 は、IP ネットワーク 5 を介してホスト装置 2 及びストレージ装置 4 と接続されている。

【 0 0 2 8 】

稼働情報監視部 7 は、ストレージ装置 4 の稼働情報を収集及び管理する機能を有し、C P U 4 0 及びメモリ 4 1 等の情報処理資源を備えて構成される。メモリ 4 1 には各種制御プログラムや、後述の性能実測値テーブル 5 2 及び業務別性能実測値テーブル 5 3 が格納される。このメモリ 4 1 に格納された制御プログラムを C P U 4 0 が実行することにより、各種の処理が実行される。

40

【 0 0 2 9 】

また負荷制御部 8 は、各チャネルアダプタ 2 1 におけるデータ入出力処理の負荷を制御する機能を有し、C P U 4 2 及びメモリ 4 3 等の情報処理資源を備えて構成される。メモリ 4 3 には各種制御プログラムや、後述の目標性能設定テーブル 5 0 が格納される。メモリ 4 3 に格納された制御プログラムを C P U 4 2 が実行することにより、各種の処理が実行される。

【 0 0 3 0 】

50

(2) 本実施の形態によるデータ入出力制御機能

(2-1) 各種テーブルの構成

次に、この記憶システム1に採用されたデータ入出力制御機能について説明する。

【0031】

本実施の形態による記憶システム1は、ホスト装置2がそれぞれ対応するアプリケーションプログラムに基づいて各種処理を行う業務のうち、優先したい業務のデータ入出力処理をユーザが設定した性能に応じた頻度で行うように制御するデータ入出力制御機能が採用されている点を特徴とする。

【0032】

このようなデータ入出力制御機能に関する各種処理を行うための手段として、コントローラ部20内の管理端末24のメモリ25及び負荷制御部8のメモリ43に目標性能設定テーブル50が格納され、稼動情報制御部7のメモリ40に性能実測値テーブル52、業務別性能実測値テーブル53が格納されている。

10

【0033】

目標性能設定テーブル50は、ホスト装置2が行う各業務のうち、特にデータ入出処理を優先すべき業務(以下、これを優先指定業務と呼ぶ)及びこれ以外の業務群(以下、これを非優先指定業務群と呼ぶ)についてユーザがそれぞれ設定した目標性能を管理するために用いるテーブルである。ここで「目標性能」とは、その優先指定業務又は非優先指定業務群について行う1秒あたりのデータ入出力処理の処理回数(IOPS)をいう。そしてこの目標性能設定テーブル50は、図2に示すように、ターゲットフィールド50A、

20

【0034】

このうちターゲットフィールド50Aには、対応する優先指定業務又は非優先指定業務群に付与されたIDが格納される。また、各優先指定業務に対応する目標性能フィールド50Bには、ユーザが設定したその優先指定業務について期待する目標性能が格納される。さらに、処理割合フィールド50Cには、予め測定した記憶システム1全体の1秒あたりのデータ入出力処理回数(以下、これを限界性能と呼ぶ)に対するその優先指定業務又は非優先指定業務群の目標性能の割合(以下、これを処理割合と呼ぶ)が格納される。さらに目標性能設定テーブル50の最下段の行50Dには、予め測定した限界性能が格納される。

30

【0035】

従って、図2の例では、優先指定業務として「業務1」～「業務3」があり、このうち例えば「業務1」という優先指定業務については、ユーザにより設定された目標性能が「10000」、その処理割合は「50%」であり、限界性能は「20000」回であることが分かる。

【0036】

なお、後述のように、各優先指定業務の処理割合は、ユーザが設定したその優先指定業務の目標性能及び予め測定した限界性能に基づいて算出され、非優先指定業務群の目標性能及び処理割合は、かかる各優先指定業務の目標性能及びかかる限界性能に基づいて算出される。このため、目標性能設定テーブル50における非優先指定業務群に対応する目標性能フィールド50B(図2において「その他」の行の目標性能フィールド50B)内には目標性能が格納されない。

40

【0037】

また性能実測値テーブル52は、優先指定業務及び非優先指定業務群ごとの1秒あたりのデータ入出力処理の処理回数の実測値(以下、これを性能実測値と呼ぶ)を管理するためのテーブルであり、図3に示すように、ターゲットフィールド52A、目標性能フィールド52B、処理割合フィールド52C及び性能実測値フィールド52Dから構成される。

【0038】

このうちターゲットフィールド52A、目標性能フィールド52B及び処理割合フィールド52Cには、それぞれ図2について上述した目標性能設定テーブル50のターゲット

50

フィールド 5 0 A、目標性能フィールド 5 0 B 及び処理割合フィールド 5 0 C と同様の情報が格納される。また性能実測値フィールド 5 2 D には、その優先指定業務又は非優先指定業務群について取得された上述の性能実測値が格納される。

【 0 0 3 9 】

従って、図 3 では、例えば「業務 1」という優先指定業務は、目標性能が「10000」であったのに対し、性能実測値が「8000」であったことが示されている。

【 0 0 4 0 】

さらに業務別性能実測値テーブル 5 3 は、優先指定業務及び非優先指定業務群ごとの複数回分の性能実測値を管理するために用いられるテーブルであり、図 4 に示すように、ターゲットフィールド 5 3 A、目標性能フィールド 5 3 B、第 1 及び第 2 の処理割合フィールド 5 3 C、5 3 E 並びに第 1 及び第 2 の性能実測値フィールド 5 3 D、5 3 F から構成されている。

【 0 0 4 1 】

そしてターゲットフィールド 5 3 A、目標性能フィールド 5 3 B 及び処理割合フィールド 5 3 C には、それぞれ図 2 について上述した目標性能設定テーブル 5 0 のターゲットフィールド 5 0 A、目標性能フィールド 5 0 B 及び処理割合フィールド 5 0 C と同様の情報が格納される。

【 0 0 4 2 】

また第 1 の性能実測値フィールド 5 3 D には、最後から 2 番目に取得されたその優先指定業務又は非優先指定業務群に関する上述の性能実測値が格納され、第 2 の処理割合フィールド 5 3 E には、当該性能実測値に基づき算出されたその優先指定業務及び非優先指定業務群に対する上述の処理割合が格納される。

【 0 0 4 3 】

さらに第 2 の性能実測値フィールド 5 3 F には、その優先指定業務又は非優先指定業務の処理割合を第 2 の処理割合フィールド 5 3 E に設定した場合に得られたその優先指定業務又は非優先指定業務群に関する上述の性能実測値（つまり最後に得られた性能実測値）が格納される。

【 0 0 4 4 】

従って、図 4 では、例えば、「業務 1」という優先指定業務は、目標性能が「10000」で処理割合が「50%」のときの性能実測値が「8000」であったのに対し、処理割合を「63%」に変更したところ性能実測値が「9800」となったことが示されている。

【 0 0 4 5 】

(2 - 2) データの入出力優先処理

次に、本実施の形態の記憶システム 1 におけるデータの入出力優先処理について説明する。図 5 は、記憶システム 1 におけるデータ入出力制御機能に関する一連の処理の流れを示すフローチャートである。

【 0 0 4 6 】

この記憶システム 1 の場合、かかるデータ入出力制御機能に関し、まず、記憶システム 1 が稼働前であればシステムテストの段階で、また記憶システム 1 が稼働中であれば運用時におけるストレージ装置 4 の限界性能を管理端末 2 4 が採取する (S P 1)。このようにして採取された限界性能は、管理端末 2 4 内のメモリ 2 5 に保持された上述の目標性能設定テーブル 5 0 (図 2) の最下段に格納されて保持される。

【 0 0 4 7 】

続いて、ユーザが管理端末 2 4 を操作して、所望する 1 又は複数の優先指定業務を指定すると共に、ステップ S P 1 で採取した限界性能を参照しながら、その優先指定業務ごとの目標性能をそれぞれ設定する (S P 2)。

【 0 0 4 8 】

このとき管理端末 2 4 の C P U 2 6 は、指定された各優先指定業務に対してそれぞれ固有の I D を付与し、これら I D と、その I D に対応する優先指定業務について設定された目標性能とを、それぞれ目標性能設定テーブル 5 0 の対応するターゲットフィールド 5 0

10

20

30

40

50

A及び目標性能フィールド50Bに格納する。また管理端末24AのCPU26は、これら各優先指定業務のIDと、そのIDに対応する優先指定業務について設定された目標性能と、ステップSP1において採取したストレージ装置4の性能限界とを稼動情報監視部7及び負荷制御部8に送信する。

【0049】

稼動情報監視部7のCPU40は、管理端末24から送信されるこれらの情報のうち、各優先指定業務のID及びその目標性能を、性能実測値テーブル52の対応するターゲットフィールド52A及び目標値性能フィールド52Bと、業務別性能実測値テーブル53の対応するターゲットフィールド53A及び目標性能フィールド53Bとにそれぞれ格納する。

10

【0050】

また負荷制御部8のCPU42は、管理端末24から送信されるこれらの情報に基づいて、管理端末24のメモリ25に保持された目標性能設定テーブル50と同じ内容の目標性能設定テーブル50を作成し、これをメモリ43に格納する。

【0051】

この後、負荷制御部8のCPU42は、上述のようにしてメモリ43に格納された目標性能設定テーブル50に登録されている各優先指定業務の目標性能及びストレージ装置4の限界性能に基づいて、優先指定業務及び非優先指定業務群ごとの処理割合を算出及び決定する(S P 3)。

【0052】

具体的にCPU42は、図4に示すように、優先指定業務として「業務1」～「業務3」が存在し、これらの目標性能がそれぞれ「10000」、「5000」及び「2000」であり、ストレージ装置4の限界性能が「20000」であった場合、「業務1」については、次式(数1)

20

$$10000 / 20000 \times 100 = 50 [\%] \quad \dots \dots (1)$$

のように処理割合を算出及び決定し、「業務2」については、次式

(数2)

$$5000 / 20000 \times 100 = 25 [\%] \quad \dots \dots (2)$$

のように処理割合を算出する。

【0053】

またCPUは、「業務3」については、次式

30

(数3)

$$2000 / 20000 \times 100 = 10 [\%] \quad \dots \dots (3)$$

のように処理割合を算出して決定し、非優先指定業務群(図4の「その他」)については、「業務1」～「業務3」について算出した各処理割合を利用して、次式

(数4)

$$100 - (50 + 25 + 10) = 15 [\%] \quad \dots \dots (4)$$

のように処理割合を算出して決定する。

【0054】

そしてCPU42は、このようにして各優先指定業務及び非優先指定業務群の処理割合を決定すると、これら各優先指定業務及び非優先指定業務群の処理割合を、目標性能設定テーブル50の処理割合フィールド50Cに格納すると共に稼動情報監視部7に報告する(S P 4)。

40

【0055】

また稼動情報監視部7のCPU40は、かかる報告を受けると、各優先指定業務及び非優先指定業務群の処理割合をそれぞれ性能実測値テーブル52の処理割合フィールド52Cと、業務別性能実測値テーブル53の第1の処理割合フィールド53Cとにそれぞれ格納すると共に、これら各処理割合を各管理端末24に報告する(S P 4)。

【0056】

さらに各管理端末24は、かかる報告を受けると、各優先指定業務及び非優先指定業務

50

群の処理割合をそれぞれ目標性能設定テーブル50の目標性能フィールド50Bに格納すると共に、これら各処理割合を自己の管理下にある各チャンネルプロセッサ21に報告する(S P 4)。

【0057】

一方、かかる報告を受けた各チャンネルプロセッサ21は、各優先指定業務及び非優先指定業務群に対する処理割合を、指定された処理割合として内部設定する。そして各チャンネルプロセッサ21は、この後、ホスト装置2のバス制御部13から送信されるデータ入出力要求コマンドに応じて、上述のように内部設定した処理割合を満たすように、例えば図6の「I/O処理順序」欄51Cに示すような順番で各優先指定業務及び非優先指定業務群の各業務についてのデータ入出力処理を実行する(S P 5)。

10

【0058】

この場合において、チャンネルプロセッサ21は、当該チャンネルプロセッサ21においてデータ入出力要求コマンド処理を行った各優先指定業務及び非優先指定業務群全体のデータ入出力処理の処理回数と、各優先指定業務及び非優先指定業務群ごとのデータ入出力処理の処理回数とを、それぞれ個別にカウンタ(図示せず)によりカウントするようにして計測する。

【0059】

そしてチャンネルプロセッサ21は、各優先指定業務及び非優先指定業務群に対してそれぞれその処理割合に基づいて予め設定された処理回数、例えば図6の場合であれば、「業務1」については50回、「業務2」については25回、「業務3」については10回及び「その他」については15回のデータ入出力処理を行った段階で、その優先指定業務又は非優先指定業務群についてのデータ入出力処理を中止する。さらにチャンネルプロセッサ21は、各優先指定業務及び非優先指定業務群全体のデータ入出力処理の処理回数の総数が各優先指定業務及び非優先指定業務群に対してそれぞれ設定された処理回数の合計値(例えば「100回」)に達した段階で各カウンタのカウント値をクリアする(「0」に戻す)。このようにしてチャンネルプロセッサ21は、内部設定した処理割合に応じた頻度で、各優先指定業務及び非優先指定業務群についてのデータ入出力処理を実行する。

20

【0060】

この際、チャンネルプロセッサ21は、いずれかの優先指定業務又は非優先指定業務群についてのデータ入出力要求コマンドを受信しなかった場合には、その分のデータ入出力処理を他の優先指定業務又は非優先指定業務群に割り当てる。ただし、このままかかる優先指定業務又は非優先指定業務群についてのデータ入出力要求コマンドを全く受信しなかった場合、他の優先指定業務及び非優先指定業務群についてのデータ入出力処理の処理回数がそれぞれ対応する既定数に達した段階でその後のデータ入出力処理が行われなくなる問題がある。そこで本実施の形態の場合、チャンネルプロセッサ21は、限界性能を基準に各カウンタのカウント値をクリアする。具体的には、限界性能が「20000」である場合、10000回のデータ入出力処理を行うために要する時間は、次式

30

(数5)

$$100 / 20000 \times 100 = 0.005 \text{ [s]} = 5 \text{ [ms]} \quad \dots \dots (5)$$

のように5[ms]であるため、チャンネルプロセッサ21は、5[ms]周期で各カウンタのカウント値をクリアすることとなる。

40

【0061】

なお、チャンネルアダプタ21がこのような処理を行い得るようにするため、ホスト装置2のバス制御部13は、データ入出力要求コマンドがいずれの業務についてのものかをチャンネルプロセッサ21が認識できるように、業務ごとに異なるIDを付加したデータ入出力要求コマンドを発行する。

【0062】

他方、チャンネルプロセッサ21は、上述のようなデータ入出力処理の実行中に、既存の技術を利用して優先指定業務及び非優先指定業務群ごとの性能実測値を所定周期(例えば1秒周期)で採取する(S P 6)。そしてチャンネルプロセッサ21は、採取した優先指定

50

業務及び非優先指定業務群ごとの性能実測値を、管理端末 2 4 を経由して稼働情報制御部 7 に報告する (S P 7) 。

【 0 0 6 3 】

そして稼働情報制御部 7 の C P U 4 0 は、この報告を受信すると、取得した各優先指定業務及び非優先指定業務群の性能実測値を、性能実測値テーブル 5 2 の対応する性能実測値フィールド 5 2 D と、業務別性能実測値テーブル 5 3 の対応する第 1 又は第 2 の性能実測値フィールド 5 3 D , 5 3 F とにそれぞれ格納して管理する (S P 8) 。

【 0 0 6 4 】

この後、記憶システム 1 では、処理割合の決定 (S P 3) 以降の処理が一定周期 (以下、これを処理割合決定周期と呼ぶ) を繰り返す (S P 3 ~ S P 8) 。

10

【 0 0 6 5 】

この場合において、かかる処理割合決定周期は一定時間であるものとし、かかる一定時間における平均性能実測値を元に、次の処理割合決定周期における各優先指定業務及び非優先指定業務群の性能実測値の処理割合をそれぞれ決定する。ただし、性能が安定しているときなど、各優先指定業務及び非優先指定業務群の処理割合を頻繁に変更する必要がない場合には、処理割合決定周期を管理端末 2 4 から所定時間単位 (例えば 1 0 分単位) で変更できるようにしても良い。

【 0 0 6 6 】

以上のように、チャンネルアダプタ 2 1 におけるデータの入出力処理はステップ S P 3 において決定された処理割合に基づいて行われるが、実際にはチャンネルアダプタ 2 1 以外の部位も各優先指定業務及び非優先指定業務群のデータ入出力処理の処理割合の決定に影響する。このため、かかる処理割合に基づいて上述のようなデータ入出力処理を行った場合においても、性能実測値が目標性能に達しないときや、目標性能以上の性能がでていることがある。なお、チャンネルアダプタ 2 1 の性能が目標性能以下となるときの例としては、データ入出力処理に時間がかかり、記憶システム 1 全体のデータ入出力処理回数が既定回数 (例えば「100回」) に達する前に、次の処理割合決定周期が到来した場合などが挙げられる。またチャンネルアダプタ 2 1 の性能が目標性能以上となるときの例としては、次の処理割合決定周期が到来する前に記憶システム 1 全体のデータ入出力処理の処理回数が既定回数に達した場合などが挙げられる。

20

【 0 0 6 7 】

そこで本実施の形態の場合、負荷制御部 8 は、1 つ前の処理割合決定周期において決定した優先指定業務及び非優先指定業務群ごとの処理割合と、そのとき得られた優先指定業務及び非優先指定業務群ごとの性能実測値と、各優先指定業務及び非優先指定業務群についてそれぞれ設定されている目標性能とに基づいて、次式

30

(数 6)

$$\text{目標性能} : \text{次の処理割合} = \text{性能実測値} : \text{前の処理割合} \quad \dots \dots (6)$$

で与えられる関係を満たすように、次の処理割合決定周期における各優先指定業務及び非優先指定業務群の処理割合を決定する。

【 0 0 6 8 】

例えば、図 3 における性能実測値テーブル 5 2 を例にとると、「業務 1 」については、目標性能が「10000」、処理割合が「50%」とされているところ、そのときの性能実測値が「8000」であったため、次の処理割合決定周期における「業務 1 」の処理割合 a は、(6) 式を用いて次式

40

(数 7)

$$a = 10000 \times 50 / 8000 = 63 [\%]$$

のように求めることができる。同様にして、次の処理割合決定周期における「業務 2 」の処理割合は「25%」、「業務 3 」の処理割合は「7%」、非優先指定業務群の処理割合は「5%」と求められる。

【 0 0 6 9 】

また、非優先指定業務群の処理割合は、優先指定業務の処理割合の決定後に残りを割り

50

当てるため、優先指定業務の処理割合の合計が「100」に達してしまった場合、非優先指定業務の処理割合が「0」又はマイナスの値となり、非優先指定業務群についてのデータ入出力処理が実施されないことになる。同様に、いずれかの優先指定業務の処理割合が「100」に決定されたときも、非優先指定業務の処理が実施されないことになる。

【0070】

例えば、図8を例にとると、「業務1」については、目標性能が「10000」、処理割合が「50%」、性能実測値が「5000」であるため、次の処理割合決定周期における処理割合が、次式

(数8)

$$10000 \times 50 / 5000 = 100 \text{ [\%]} \quad \dots\dots (8)$$

のように算出される。

【0071】

同様に、「業務2」については、次式

(数9)

$$5000 \times 25 / 5000 = 25 \text{ [\%]} \quad \dots\dots (9)$$

「業務3」については、次式

(数10)

$$2000 \times 10 / 5000 = 4 \text{ [\%]} \quad \dots\dots (10)$$

非優先指定業務群については、次式

(数11)

$$100 - (100 + 25 + 4) = -29 \text{ [\%]} \quad \dots\dots (11)$$

のように処理割合がそれぞれ算出及び決定される。

【0072】

そして、このように非優先指定業務群の処理割合がマイナスになると、非優先指定業務群が処理されなくなる不都合が生じる。そこで、本実施の形態においては、上述した各優先指定業務及び非優先指定業務群の処理割合の下限値がいずれも「1」に設定されており、これら優先指定業務及び非優先指定業務群について、最低「1」の処理分を確保し得るようになされている。そして負荷制御部8は、優先指定業務及び非優先指定業務群のいずれかの処理割合が「0」又はマイナスになった場合、全体の平均をとって処理割合を決定する。

【0073】

例えば、図8に示す例の場合、「業務1」については、次式

(数12)

$$100 / (100 + 25 + 4 + 1) \times 100 = 77 \text{ [\%]} \quad \dots\dots (12)$$

のように処理割合を算出及び決定し、「業務2」については、次式

(数13)

$$25 / (100 + 25 + 4 + 1) \times 100 = 19 \text{ [\%]} \quad \dots\dots (13)$$

のように処理割合を算出及び決定し、「業務3」については、次式

(数14)

$$4 / (100 + 25 + 4 + 1) \times 100 = 3 \text{ [\%]} \quad \dots\dots (14)$$

のように処理割合を算出及び決定し、非優先指定業務については、次式

(数15)

$$100 - (77 + 19 + 3) = 1 \text{ [\%]} \quad \dots\dots (15)$$

のように処理割合を算出及び決定する。

【0074】

このようにすることによって、各優先指定業務及び非優先指定業務の処理割合の総和が「100」となり、各業務の処理割合が「0」又はマイナスとなることはない。

【0075】

続いて、図5のステップSP5において説明したチャンネルプロセッサ21の処理について、図7を参照して説明する。図7は、ホスト装置2からのデータ入出力要求コマンドを

10

20

30

40

50

受信したチャンネルプロセッサ 2 1 が行う処理の一連の手順を示したフローチャートである。

【 0 0 7 6 】

チャンネルプロセッサ 2 1 は、この処理割合に基づくデータ入出力処理を開始すると、まず、負荷制御部 8 により算出された各優先指定業務及び非優先指定業務群に対する処理割合の報告を受けているかを確認する (S P 1 0) 。

【 0 0 7 7 】

そしてチャンネルプロセッサ 2 1 は、かかる処理割合の報告を受けているときには (S P 1 0 : Y E S) 、処理割合の受信内容を設定する (S P 1 1) 。これに対してチャンネルプロセッサ 2 1 は、処理割合の受信内容が設定されるか、処理割合の報告を受けていないとき
10

【 0 0 7 8 】

この後、チャンネルプロセッサ 2 1 は、ホスト装置 2 からのコマンド又はコマンドキューを受領すると (S P 1 2) 、コマンド毎に付与された I D があるときは、かかる I D に基づいて、かかるコマンドがどの業務に対するコマンドであるかを判別する。またチャンネルプロセッサ 2 1 は、当該業務における処理回数の割合が、記憶システム 1 全体としての目標の処理割合以下であるかを確認する (S P 1 3) 。

【 0 0 7 9 】

そしてチャンネルプロセッサ 2 1 は、コマンドが目標の処理割合以下であるときには (S P 1 3 : Y E S) コマンド処理を実行する (S P 1 4) 。これに対してチャンネルプロセッサ 2 1 は、コマンドが目標処理割合を超えているときには (S P 1 3 : N O) 、コマンド処理を行うことなく、コマンドを溜めるコマンドキューイングを行う (S P 2 1) 。

10

20

【 0 0 8 0 】

チャンネルプロセッサ 2 1 は、コマンド処理を実行した後、当該業務における限界性能に加算を行うとともに (S P 1 5) 、当該業務のコマンド処理回数及び記憶システム 1 全体の処理回数を確認する (S P 1 6) 。チャンネルプロセッサ 2 1 は、処理回数の確認が終わるとき (S P 1 6) 、又はコマンドキューイングが終了したときには (S P 2 1) 、記憶システム 1 全体としての処理回数が 1 0 0 になったか、又はカウンタクリア周期を経過したかを確認する (S P 1 7) 。

30

【 0 0 8 1 】

チャンネルプロセッサ 2 1 は、記憶システム 1 全体の処理回数が 1 0 0 になるか、カウンタクリア周期が経過しているときには (S P 1 7 : Y E S) 、記憶システム 1 全体の処理回数又はカウンタクリア周期をクリアして (S P 1 8) 、さらに 1 秒間経過するのを待ち受け (S P 1 9 : Y E S) 、管理端末 2 4 に対して限界性能を報告し (S P 2 0) 、処理を終了する。さらに続けて処理を行うときには、ステップ S P 1 0 に戻り、対象とする処理を続けて行う処理に順次切り替えながら同様の処理を繰り返す (ステップ S P 1 0 ~ ステップ S P 2 0) 。 1 秒間経過するのを待ち受けないときには (S P 1 9 : N O) 、管理端末 2 4 に対して限界性能を報告することなく、ステップ S P 1 0 に戻り、対象とする処理を、続けて行う処理に順次切り替えながら同様の処理を繰り返す (ステップ S P 1 0 ~ ステップ S P 2 0) 。

40

【 0 0 8 2 】

次に、図 5 に示したフローチャートにおいて、ステップ S P 8 終了後に、ステップ S P 3 に戻る場合の処理割合の決定方法について説明する。

【 0 0 8 3 】

図 9 は、第 1 回目の処理を終えた後に、第 2 回目の処理割合を決定するための一連の手順を示したフローチャートである。負荷制御部 8 の C P U 4 2 は、優先指定業務群について、第 2 回目の処理割合を算出する (S P 3 0) 。 C P U 4 2 はさらに、非優先指定業務群の第 2 回目の処理割合を算出する (S P 3 1) 。 C P U 4 2 は、非優先指定業務の処理割合が 0 以下であるかを確認する (S P 3 2) 。 C P U 4 2 は、非優先指定業務の処理割
50

50

合が 0 以下であるときには (S P 3 2 : Y E S)、非優先指定業務の処理割合を「 1 」にし (S P 3 3)、優先指定業務群の処理割合を補正する (S P 3 4)。さらに C P U 4 2 は、再度、非優先指定業務の処理割合を算出する (S P 3 5)。C P U 4 2 は、すべての業務の処理割合を決定すると、その処理割合を、各チャネルプロセッサ 2 1 A , 2 1 B , 2 1 C , 2 1 D 等に報告し (S P 3 6)、処理を終了する。

【 0 0 8 4 】

(2 - 3) 本実施形態の効果

本実施形態によれば、優先指定業務について、ユーザが所望する目標性能に応じた頻度でデータ入出力処理を行うことができる。さらに、処理割合について、前の処理割合決定周期において取得した各優先指定業務及び非優先指定業務群の性能実測値に基づいた変更を行うことにより、優先指定業務についてのデータ入出力処理の頻度をより精度良く目標性能に近づけるような制御を行なうことができる。

10

【 0 0 8 5 】

(3) 他の実施の形態

なお、本実施形態においては、同一のストレージ装置 4 内において、複数のコントローラ部 2 0 を有する場合について述べた。しかし、本発明はこれに限らず、図 1 1 に示すように、単数のコントローラ部 2 0 を有してもよい。

【 0 0 8 6 】

また、本実施形態においては、ストレージ装置 4 が一つのホスト装置 2 に接続される場合について述べた。しかし、本発明はこれに限らず、図 1 2 に示すように、複数のホスト装置 2 に接続されてもよい。

20

【 0 0 8 7 】

さらに、本実施形態においては特に言及していないが、時間帯により業務の使用頻度が変わり、時間帯により業務の処理割合を変更したいときには、時間帯ごとに目標性能を管理端末 2 4 にてあらかじめ設定しておき、負荷制御部 8 の C P U 4 2 がその設定に従って当該時刻毎になったところで目標性能を再設定することにより、記憶システム 1 の運用開始時に一通りの設定を行うだけで連続稼動が可能となるように設定してもよい。図 1 0 に示すように、「業務 1」という I D の優先指定業務について、6時から7時は目標性能値を「10000」とし、7時から8時には「5000」、8時から9時には「2000」のように時間帯に応じて目標性能値を変更するようにしてもよい。

30

【 0 0 8 8 】

また、本実施形態では、データ入出力優先処理の制御対象となるデータ入出力処理要求の種別であるターゲットが、データ入出力要求を発行した業務 (アプリケーション) 毎に区別される例を述べた。しかし、本発明はこれに限らず、図 1 3 に示すように、データの読書きをする対象となるデバイス (L D E V) でもよい。また、ターゲットとは、データ処理を行うための処理の単位としてもよく、データ処理をするプログラムごとに区別してもよいし、データを格納する領域ごとに区別してもよいし、その他でもよい。

【 産業上の利用可能性 】

【 0 0 8 9 】

本発明は、種々の構成の記憶システムに広く適用することができる。

40

【 図面の簡単な説明 】

【 0 0 9 0 】

【 図 1 】 本実施の形態による記憶システム 1 を示すブロック図である。

【 図 2 】 本実施の形態による目標性能設定テーブル 5 0 の構成を示す図である。

【 図 3 】 本実施の形態による性能実測値テーブル 5 2 の構成を示す図である。

【 図 4 】 本実施の形態による性能実測値テーブル 5 3 の構成を示す図である。

【 図 5 】 本実施の形態による処理フローを示すフローチャートである。

【 図 6 】 本実施の形態によるデータ入出力処理順序表を示す図である。

【 図 7 】 チャネルプロセッサ 2 1 による処理フローを示すフローチャートである。

【 図 8 】 本実施の形態による性能実測値テーブルの構成を示す図である。

50

【図9】処理割合を再設定する処理フローを示すフローチャートである。

【図10】時間帯ごとに業務の処理割合を変更する場合のテーブルの構成を示す図である。

【図11】単数のコントロール部20を備えた記憶システム1を示すブロック図である。

【図12】複数のホスト装置2を備えた記憶システム1を示すブロック図である。

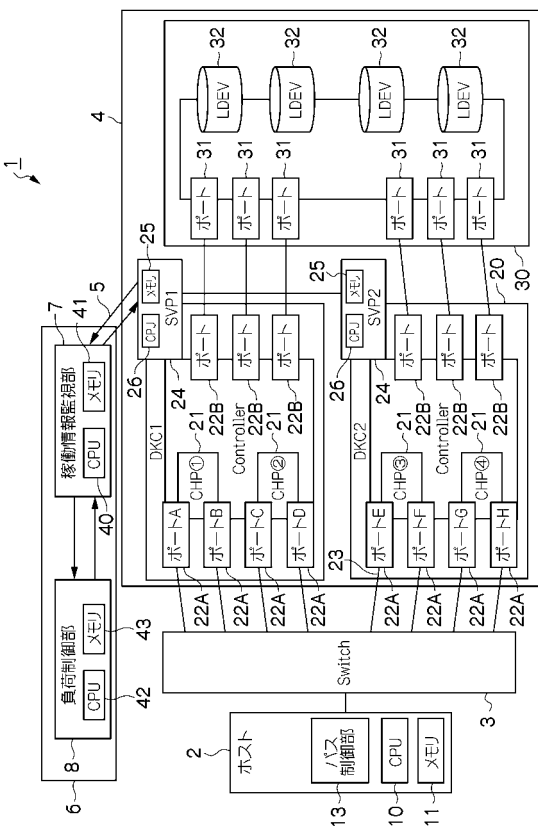
【図13】本実施の形態によるデータ入出力処理順序表を示す図である。

【符号の説明】

【0091】

- 1 …… 記憶システム、 2 …… ホスト装置、 4 …… ストレージ装置、 5 …… IPネットワーク、 6 …… 管理サーバ、 7 …… 稼働情報監視部、 8 …… 負荷制御部、 10 …… CPU、
- 11 …… メモリ、 20 …… コントロール部、 21 …… チャンnelアダプタ、 22 …… ポート、 24 …… 管理端末、 30 …… ディスクデバイス部。

【図1】



【図2】

50	50A	50B	50C
	ターゲット	目標性能[IOPS]	処理割合[%]
	業務1	10000	50
	業務2	5000	25
	業務3	2000	10
	その他	-	15
	限界性能	20000[IOPS]	

【図3】

52	52A	52B	52C	52D
	ターゲット	目標性能[IOPS]	処理割合[%]	性能実測値[IOPS]
	業務1	10000	50	8000
	業務2	5000	25	5000
	業務3	2000	10	3000
	その他	-	15	4000

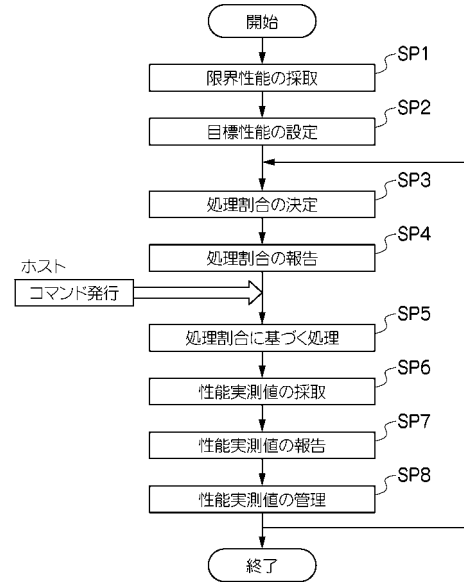
【 図 4 】

53

ターゲット	Step1		Step2		Step3
	目標性能 [IOPS]	処理割合[%]	性能実測値 [IOPS]	処理割合[%]	性能実測値 [IOPS]
業務1	10000	50	8000	63	9800
業務2	5000	25	5000	25	5000
業務3	2000	10	3000	7	2200
その他	-	15	4000	5	1500

53A
53B
53C
53D
53E
53F

【 図 5 】

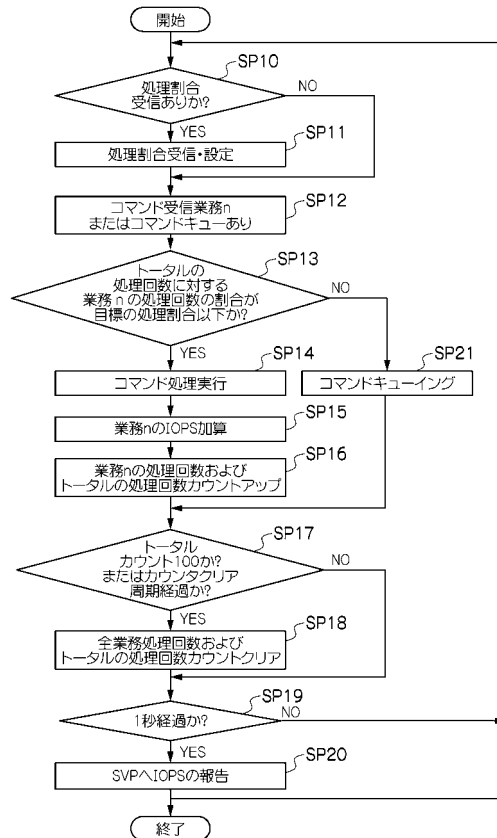


【 図 6 】

51

ターゲット	処理割合	IO処理順序									
		1	5	9	12	14	16	17	18	19	20
業務1	50%	1	5	9	12	14	16	17	18	19	20
業務2	25%	2	6	10	13	15	-	-	-	-	-
業務3	10%	3	7	-	-	-	-	-	-	-	-
その他	15%	4	8	11	-	-	-	-	-	-	-

【 図 7 】



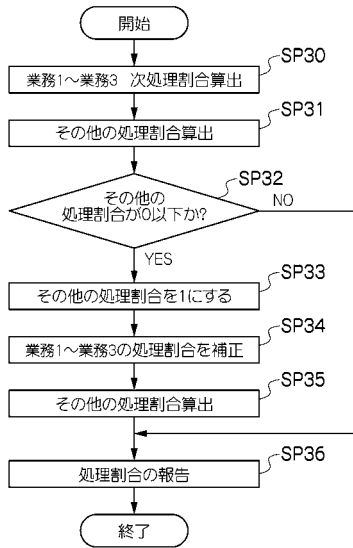
【 図 8 】

54			
52A	54B	54C	54D
ターゲット	目標性能[IOPS]	コマンド処理割合[%]	性能実測値
業務1	10000	50	5000
業務2	5000	25	5000
業務3	2000	10	5000
その他	-	15	5000

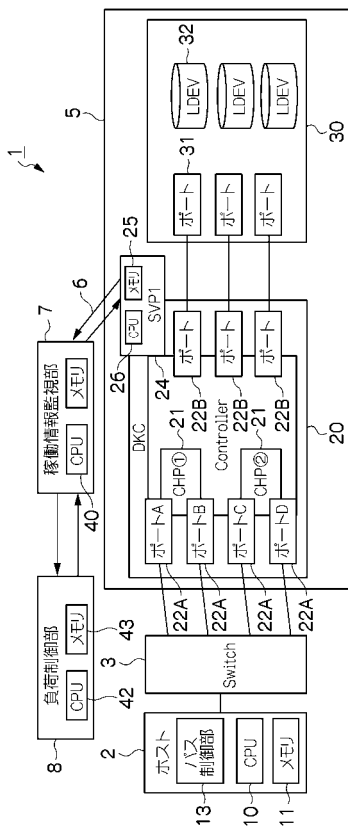
【 図 1 0 】

55				
55A	55B	55C	55D	55E
ターゲット	6:00~7:00	7:00~8:00	8:00~9:00	...
業務1	10000	5000	2000	...
業務2	5000	5000	8000	...
業務3	2000	3000	5000	...

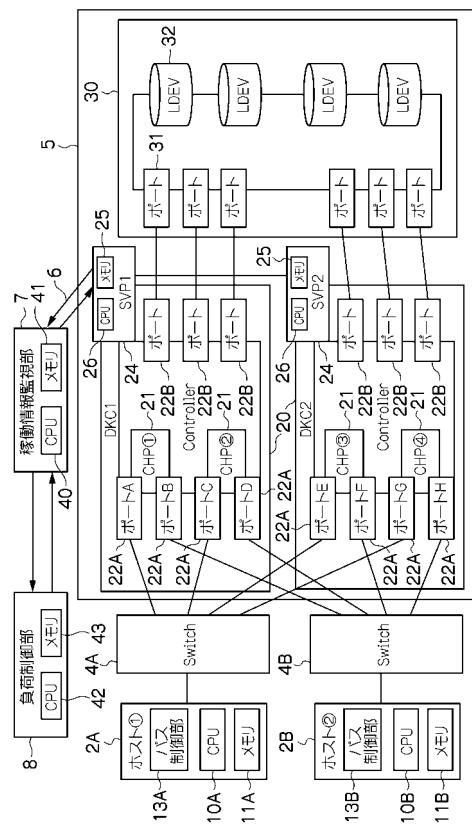
【 図 9 】



【 図 1 1 】



【 図 1 2 】



【 図 1 3 】

56

56A		56B		56C							
ターゲット	処理割合	コマンド処理順序									
LDEV1	50%	1	5	9	12	14	16	17	18	19	20
LDEV2	25%	2	6	10	13	15	-	-	-	-	-
LDEV3	10%	3	7	-	-	-	-	-	-	-	-
その他	15%	4	8	11	-	-	-	-	-	-	-