

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6213148号
(P6213148)

(45) 発行日 平成29年10月18日(2017.10.18)

(24) 登録日 平成29年9月29日(2017.9.29)

(51) Int.Cl.

F I

G 0 6 F 3 / 0 6 (2006.01)

G 0 6 F 3 / 0 6 3 0 1 E

請求項の数 9 (全 46 頁)

(21) 出願番号	特願2013-222060 (P2013-222060)	(73) 特許権者	000005223
(22) 出願日	平成25年10月25日(2013.10.25)		富士通株式会社
(65) 公開番号	特開2015-84143 (P2015-84143A)		神奈川県川崎市中原区上小田中4丁目1番1号
(43) 公開日	平成27年4月30日(2015.4.30)	(74) 代理人	100092978
審査請求日	平成28年7月5日(2016.7.5)		弁理士 真田 有
		(74) 代理人	100112678
			弁理士 山本 雅久
		(72) 発明者	尾田 丈一
			石川県かほく市宇野気ヌ98番地の2 株式会社P F U内
		審査官	田中 啓介

最終頁に続く

(54) 【発明の名称】 ストレージ装置、ストレージ装置の制御方法およびストレージ装置制御プログラム

(57) 【特許請求の範囲】

【請求項 1】

複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出する算出部と、

情報処理装置から受信した入出力要求に係る処理量と、前記算出部により算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールするスケジュール部と、

前記スケジュール部によりスケジュールされた前記実行順で前記入出力要求に係る処理を実行する実行部と、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第1所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定する第1判定部と、

当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると前記第1判定部により判定された場合、当該記憶装置についての前記上限値を、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である所定の性能融通単位で変更する変更部と、

を有する、ストレージ装置。

【請求項 2】

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第2所定時間に亘って当該記憶装置の処理性能を下回っているか否かを判定する第2判定部を、さらに有し、

10

20

前記複数の記憶装置のうち第 1 記憶装置に対する入出力要求に係る処理量が前記第 1 所定時間に亘って前記第 1 記憶装置の処理性能を超えていると前記第 1 判定部により判定され、且つ、前記複数の記憶装置のうち第 2 記憶装置に対する入出力要求に係る処理量が前記第 2 所定時間に亘って前記第 2 記憶装置の処理性能を下回っていると前記第 2 判定部により判定された場合、前記変更部は、前記上限値を前記所定の性能融通単位で変更することで、前記第 2 記憶装置から前記第 1 記憶装置への性能融通を前記所定の性能融通単位で行なう、

請求項 1 に記載のストレージ装置。

【請求項 3】

前記変更部は、前記第 1 記憶装置について算出された前記上限値に、前記所定の性能融通単位に対応する値を加算するとともに、前記第 2 記憶装置について算出された前記上限値から、前記所定の性能融通単位に対応する値を減算するように、前記上限値を変更することで、前記第 2 記憶装置から前記第 1 記憶装置への性能融通を前記所定の性能融通単位で行なう、

請求項 2 に記載のストレージ装置。

【請求項 4】

前記記憶装置毎に、前記第 1 所定時間よりも短い時間間隔で、当該記憶装置に対する入出力要求の処理待ち量を取得し、前記第 1 所定時間内に取得された前記処理待ち量のうち最小のものを最小待ち量として取得する第 1 取得部を、さらに有し、

前記第 1 判定部は、前記第 1 記憶装置について前記第 1 取得部により取得された前記最小待ち量が 0 でない場合、前記第 1 記憶装置に対する入出力要求に係る処理量が前記第 1 所定時間に亘って前記第 1 記憶装置の処理性能を超えていると判定する、

請求項 2 または請求項 3 に記載のストレージ装置。

【請求項 5】

前記記憶装置毎に、前記第 2 所定時間よりも短い時間間隔で、当該記憶装置に対する入出力要求に係る処理量の、当該記憶装置の処理性能に対する処理余り量を取得し、前記第 2 所定時間内に取得された前記処理余り量のうち最小のものを最小余り量として取得する第 2 取得部を、さらに有し、

前記第 2 判定部は、前記第 2 取得部により取得された前記最小余り量が 0 でない場合、前記第 2 記憶装置に対する入出力要求に係る処理量が前記第 2 所定時間に亘って前記第 2 記憶装置の処理性能を下回っていると判定する、

請求項 2 ～請求項 4 のいずれか一項に記載のストレージ装置。

【請求項 6】

前記記憶装置毎に、当該記憶装置が、一以上の前記第 1 記憶装置への前記性能融通を行なった前記第 2 記憶装置であり、且つ、当該記憶装置に対する入出力要求の処理待ちの生じた融通復旧対象であるか否かを判定する第 3 判定部と、

前記第 3 判定部により当該記憶装置が前記融通復旧対象であると判定された場合、前記融通復旧対象の記憶装置が他の記憶装置に融通した処理量を、前記一以上の前記第 1 記憶装置から前記融通復旧対象の記憶装置に復帰させる復旧処理部と、

をさらに有する、請求項 2 ～請求項 5 のいずれか一項に記載のストレージ装置。

【請求項 7】

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る、前記第 1 所定時間よりも短い第 3 所定時間間隔における処理量が当該記憶装置の処理性能の所定数倍以上であるか否かを判定する第 4 判定部と、

前記複数の記憶装置のうち第 3 記憶装置に対する入出力要求に係る、前記第 3 所定時間間隔における処理量が前記第 3 記憶装置の処理性能の所定数倍以上であると前記第 4 判定部により判定された場合、前記第 3 記憶装置に対する入出力要求に係る処理量を超える処理余裕をもつ記憶装置から前記第 3 記憶装置への性能融通を行なう融通処理部と、

をさらに有する、請求項 1 ～請求項 6 のいずれか一項に記載のストレージ装置。

【請求項 8】

ストレージ装置が、
複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出し、
情報処理装置から受信した入出力要求に係る処理量と、算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールし、
スケジュールされた前記実行順で前記入出力要求に係る処理を実行し、
前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第1所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定し、
当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると判定された場合、当該記憶装置についての前記上限値を、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である所定の性能融通単位で変更する、
ことを特徴とするストレージ装置の制御方法。

10

【請求項9】

ストレージ装置に含まれるコンピュータに、
複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出し、
情報処理装置から受信した入出力要求に係る処理量と、算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールし、
スケジュールされた前記実行順で前記入出力要求に係る処理を実行し、
前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第1所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定し、
当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると判定された場合、当該記憶装置についての前記上限値を、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である所定の性能融通単位で変更する、
処理を実行させる、ストレージ装置制御プログラム。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ装置、ストレージ装置の制御方法およびストレージ装置制御プログラムに関する。

30

【背景技術】

【0002】

複数のサーバ（以下ホストという）が1台のストレージ装置を共用するストレージシステムでは、ホストの要求した入出力処理がストレージ装置内で競合し、入出力処理（アプリケーション）毎に安定したパフォーマンスを維持できない場合がある。そこで、ストレージ装置は、ホストからの処理要求に対し優先順位（優先度）を付与して処理を行なうことで、入出力処理毎に優先順位に従った適切な性能を維持させるQoS（Quality of Service）機能を有する場合がある。

【0003】

ストレージ装置が有するQoS機能（ストレージQoS）としては、ホストにより指定された優先度に基づきキューなどの内部リソースを割り当て、コマンドを処理する技術が知られている。一例として、ホストAとホストBとの2台が1台のストレージ装置を共有するストレージシステムにおいて、ホストAが優先度を指定してコマンドを送信する場合について説明する。ストレージ装置は、ホストAから優先度が指定された場合、優先度の指定が無い処理に比べて例えば3倍の内部リソースを割り当てる。この結果、ストレージ装置は、ホストAから要求された入出力処理を、ホストBから要求された入出力処理よりも優先して実行する。

40

【0004】

また、ストレージ装置側でホストやLU（Logical Unit；論理ユニット）ごとに目標性能や性能制限を設定することで、入出力処理（アプリケーション）ごとに優先順位に従っ

50

た適切な性能を維持させる技術も知られている。

【 0 0 0 5 】

さらに、特定の入出力処理の負荷状況が変化した場合、特定の入出力処理に対して設定される優先度を変更して上記負荷状況の変化に対応する技術も知られている。例えば、当該技術では、所定時間の間、継続して、実際の入出力処理量がストレージ全体の処理性能を下回り続けた場合、入出力処理量が性能限界値に到達したＬＵ等に設定された優先度を一段階上げることが行なわれている。当該ＬＵ等の優先度を一段階上げることで、当該ＬＵ等の性能限界値が高く変更される。これにより、当該ＬＵ等の性能調整が行なわれ、上記負荷状況の変化に対応することが可能になる。

【 先行技術文献 】

【 特許文献 】

【 0 0 0 6 】

【 特許文献 1 】 特開 2 0 0 6 - 5 3 6 0 1 号公報

【 特許文献 2 】 特開 2 0 0 9 - 9 3 6 3 6 号公報

【 特許文献 3 】 特開 2 0 1 3 - 1 2 7 7 5 8 号公報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 7 】

しかしながら、上述したＱｏＳ機能では、ホストからストレージ装置に要求される性能が変動する場合、その性能変動に対応して、ストレージ装置内のリソース（例えばＬＵ）間での性能調整（性能融通）を臨機応変に行なうことができない。

【 0 0 0 8 】

上述のごとく優先度を変更して負荷状況の変化に対応する技術も知られているが、優先度を変更して性能調整を行なう場合、優先度の変更前後での性能限界値の変更間隔が極めて大きくなる場合がある。性能限界値の変更間隔が大きいと、優先度を上下させるだけでは調整範囲が限定的となり、ホストから要求される性能の変動に細かく追従することができず、ストレージ装置における処理性能が低下する場合もある。

【 0 0 0 9 】

一つの側面で、本発明は、様々な性能変動に対応できるようにすることを目的とする。

なお、前記目的に限らず、後述する発明を実施するための最良の形態に示す各構成により導かれる作用効果であって、従来の技術によっては得られない作用効果を奏することも本件の他の目的の一つとして位置付けることができる。

【 課題を解決するための手段 】

【 0 0 1 0 】

本件のストレージ装置は、複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出する。また、本件のストレージ装置は、情報処理装置から受信した入出力要求に係る処理量と、算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールし、スケジュールされた前記実行順で前記入出力要求に係る処理を実行する。そして、本件のストレージ装置は、前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第 1 所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定する。さらに、本件のストレージ装置は、当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると判定された場合、当該記憶装置についての前記上限値を、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である所定の性能融通単位で変更する。

【 発明の効果 】

【 0 0 1 1 】

一実施形態によれば、様々な性能変動に対応することができる。

【 図面の簡単な説明 】

【 0 0 1 2 】

【 図 1 】 本実施形態に係るストレージシステムのハードウェア構成の一例を示すブロック

10

20

30

40

50

図である。

【図 2】本実施形態のストレージ装置により実行されるファームウェアにより実現される処理部の機能構成を示すブロック図である。

【図 3】本実施形態に係る S C S I ターゲット制御部の機能構成および Q o S 制御域の情報示すブロック図である。

【図 4】優先順位と性能値との対応関係の一例を示す図である。

【図 5】Q o S 管理テーブルの一例を示す図である。

【図 6】本実施形態に係る N ミリ秒タイマスケジュール処理の処理手順を説明するフローチャートである。

【図 7】本実施形態に係る全 L U 限界値加算処理の処理手順を説明するフローチャートである。 10

【図 8】本実施形態に係る Q o S I / O (Input/Output) 起動スケジュール処理の処理手順を説明するフローチャートである。

【図 9】本実施形態に係る Q o S I / O 起動スケジュール処理の処理手順を説明するフローチャートである。

【図 10】本実施形態に係るストレージ装置によるコマンド受信処理の処理手順を説明するフローチャートである。

【図 11】本実施形態に係るミクロ融通処理の処理手順を説明するフローチャートである。

【図 12】本実施形態に係るミクロ融通後処理の処理手順を説明するフローチャートである。 20

【図 13】本実施形態に係るマクロ融通 1 秒チェック処理の処理手順を説明するフローチャートである。

【図 14】本実施形態に係るマクロ融通処理の処理手順を説明するフローチャートである。

【図 15】本実施形態に係るマクロ融通処理の処理手順を説明するフローチャートである。

【図 16】本実施形態に係るマクロ融通復旧処理の処理手順を説明するフローチャートである。

【図 17】本実施形態に係るマクロ融通復旧処理の処理手順を説明するフローチャートである。 30

【図 18】本実施形態に係るマクロ融通復旧処理で作成される融通先ユニット配列 (L U N / ホスト番号配列) の一例を示す図である。

【図 19】本実施形態に係るマクロ融通復旧処理を具体的に説明する図である。

【図 20】本実施形態に係るマクロ融通復旧処理を具体的に説明する図である。

【図 21】本実施形態のストレージ装置制御プログラムを実行するコンピュータを示すブロック図である。

【発明を実施するための形態】

【 0 0 1 3 】

以下に、図面を参照し、本願の開示するストレージ装置、ストレージ装置の制御方法およびストレージ装置制御プログラムの実施形態について、詳細に説明する。ただし、以下に示す実施形態は、あくまでも例示に過ぎず、実施形態で明示しない種々の変形例や技術の適用を排除する意図はない。すなわち、本実施形態を、その趣旨を逸脱しない範囲で種々変形して実施することができる。また、各図は、図中に示す構成要素のみを備えるという趣旨ではなく、他の機能を含むことができる。そして、各実施形態は、処理内容を矛盾させない範囲で適宜組み合わせることが可能である。 40

【 0 0 1 4 】

〔 1 〕本実施形態の構成

まず、図 1 ~ 図 5 を参照しながら、本実施形態の構成について説明する。

〔 1 - 1 〕本実施形態のストレージシステムの構成 50

図1は、本実施形態に係るストレージシステム1の構成の一例を示す図である。図1に示すように、ストレージシステム1は、ホスト10a、ホスト10bおよびストレージ装置100を有する。また、ストレージシステム1において、ストレージ装置100は、ホスト10aおよびホスト10bと接続される。なお、ストレージ装置100と接続するホストの数は図示に限定されるものではなく、変更可能である。

【0015】

〔1-2〕ホストの構成

続いて、図1を参照しながら、ホスト10aおよびホスト10bの構成について説明する。ホスト10aおよびホスト10bは、サーバなどの情報処理装置である。ホスト10aは、FC(Fibre Channel)-A(Adapter)11a、FC-A12a、iSCSI(Internet Small Computer System Interface)-A13a、iSCSI-A14aを有する。FC-A11a、12bは、通常、FC-HBA(Host Bus Adapter)の形態であり、iSCSI-A13a、14aは、通常、NIC(Network Interface Card)を使用し、そこでiSCSIプロトコルを動作させる形態である。

【0016】

FC-A11aは、後述するCM(Controller Module)200a内のFC-CA(Communication Adapter)201aとFCで接続するインターフェースである。また、同様に、FC-A12aは、後述するCM200b内のFC-CA201bとFCで接続するインターフェースである。また、iSCSI-A13aは、後述するCM200a内のiSCSI-CA202aとiSCSIで接続するインターフェースである。同様に、iSCSI-A14aは、後述するCM200b内のiSCSI-CA202bとiSCSIで接続するインターフェースである。

【0017】

ホスト10bは、FC-A11b、FC-A12b、iSCSI-A13b、iSCSI-A14bを有する。FC-A11bは、FC-A11aに対応し、FC-A12bは、FC-A12aに対応する。また、iSCSI-A13bは、iSCSI-A13aに対応し、iSCSI-A14bは、iSCSI-A14aに対応する。なお、以下の説明では、ホスト10aとホスト10bとを区別せず一般化して称する場合には、ホスト10として記載する。

【0018】

〔1-3〕本実施形態に係るストレージ装置の構成

続いて、図1を参照しながら、本実施形態に係るストレージ装置100の構成について説明する。本実施形態に係るストレージ装置100は、HDD(Hard Disk Drive)101a~101dと、CM200a、200bとを有する。

HDD101a~101dは、RAID(Redundant Arrays of Inexpensive Drives)を構成し、ユーザデータを記憶する。CM200a、200bは、ストレージ装置100のシステム全体を制御するストレージ制御モジュールである。CM200a、200bとで冗長化されている。なお、ストレージ装置100が有するHDD、CMの数は、図示したものに限定されるものではなく、変更可能である。また、コントローラエンクロージャのストレージ装置(ストレージ制御装置)として、複数のCMを搭載し、HDD等の物理ディスクは、ディスクエンクロージャのストレージ装置に搭載して、接続することでストレージシステムを構成してもよい。

また、ストレージ装置100は、HDD101a~101d(一以上の物理記憶装置)を論理的に分割した論理ユニット(以下LUという)として使用される。各LU(記憶装置)は、それぞれが論理的な1台のHDDとしてホスト10a、10bにより認識される。

【0019】

CM200aは、FC-CA201a、iSCSI-CA202a、SAS(Serial Attached SCSI)203a、メモリ210aおよびCPU(Central Processing Unit)220aを有する。

【 0 0 2 0 】

F C - C A 2 0 1 a は、ホスト 1 0 a およびホスト 1 0 b と F C で接続するインターフェースである。i S C S I - C A 2 0 2 a は、ホスト 1 0 a およびホスト 1 0 b と i S C S I で接続するインターフェースである。S A S 2 0 3 a は、H D D 1 0 1 a ~ 1 0 1 d と、S A S で接続するインターフェースである。なお、以下の説明において F C - C A 2 0 1 a および i S C S I - C A 2 0 2 a のことを一般化して称する場合には、適宜、ポートと記載する。

【 0 0 2 1 】

メモリ 2 1 0 a は、例えば、R A M (Random Access Memory) などの半導体メモリ素子であり、キャッシュ域 2 1 1 a と制御域 2 1 2 a とを有する。キャッシュ域 2 1 1 a は、ホスト 1 0 a またはホスト 1 0 b と、H D D 1 0 1 a ~ 1 0 1 d のいずれかとの間でやり取りされるデータを一時的に保持する。制御域 2 1 2 a は、ストレージ装置 1 0 0 が実行する各種の処理に要する情報を保持する。また、制御域 2 1 2 a は各種プログラム等も格納している。さらに、制御域 2 1 2 a は、後述する Q o S 制御域 2 1 3 (図 3 参照) を有する。

10

【 0 0 2 2 】

C P U 2 2 0 a は、演算や制御などの各種処理を行う電子回路であり、後述する処理部 3 0 0 (図 2 参照) を実現するために用いられる。

【 0 0 2 3 】

C M 2 0 0 b は、F C - C A 2 0 1 b , i S C S I - C A 2 0 2 b , S A S 2 0 3 b , メモリ 2 1 0 b および C P U 2 2 0 b を有する。F C - C A 2 0 1 b は、F C - C A 2 0 1 a に対応し、i S C S I - C A 2 0 2 b は、i S C S I - C A 2 0 2 a に対応し、S A S 2 0 3 b は、S A S 2 0 3 a に対応する。また、メモリ 2 1 0 b は、メモリ 2 1 0 a に対応し、C P U 2 2 0 b は、C P U 2 2 0 a に対応する。

20

【 0 0 2 4 】

このようなストレージシステム 1 では、ストレージ装置 1 0 0 の C P U 2 2 0 a , 2 2 0 b は、ファームウェアによって以下の処理を実行する。即ち、ストレージ装置 1 0 0 は、複数の L U に設定された優先度に基づいて決定される入出力処理量の上限値を L U 毎に算出する。また、ストレージ装置 1 0 0 は、ホスト 1 0 から受信した入出力要求に係る処理量と、算出された上限値とに基づいて、入出力要求に係る処理の実行順をスケジュールし、スケジュールされた実行順で入出力要求に係る処理を実行する。そして、ストレージ装置 1 0 0 は、L U 毎に、当該 L U に対する入出力要求に係る処理量が第 1 所定時間 (例えば 3 0 秒) に亘って当該 L U の処理性能を超えているか否かを判定する。さらに、ストレージ装置 1 0 0 は、当該 L U に対する入出力要求に係る処理量が当該 L U の処理性能を超えていると判定された場合、当該 L U についての上限値を、所定の性能融通単位 (ユニット) で変更する。このようにして、ストレージ装置 1 0 0 は、ホスト 1 0 からの入出力要求 (I / O 要求, コマンド) の状況やホスト 1 0 の接続環境の変化に応じて、ストレージ性能を最適に制御する。特に、ストレージ装置 1 0 0 は、所定の性能融通単位 (ユニット) で上限値を変更調整することで、ホスト 1 0 から要求される処理性能の変動に細かく追従することができ、様々な性能変動に対応することができる。

30

40

【 0 0 2 5 】

なお、入出力処理量の上限値は、例えば、後述する限界 I / O 数や限界ブロック数であり、上限値, 限界 I / O 数, 限界ブロック数のことを、限界値、または、性能限界値という場合がある。

【 0 0 2 6 】

〔 1 - 4 〕ファームウェアにより実現される処理部の機能構成

次に、図 2 を参照しながら、ストレージ装置 1 0 0 により実行されるファームウェアにより実現される処理部 3 0 0 の機能構成について説明する。図 2 は、ストレージ装置 1 0 0 により実行されるファームウェアにより実現される処理部 3 0 0 の機能構成を示すブロック図である。なお、処理部 3 0 0 は、メモリ 2 1 0 a およびメモリ 2 1 0 b と、C P U

50

220aおよびCPU220bとが協働して実現される。

【0027】

図2に示すように、処理部300は、FC I/F (InterFace) 制御部301, iSCSI I/F 制御部302, SCSIターゲット制御部303, SCSIコマンド制御部304, リソース制御部305, キャッシュ制御部306, RAID制御部307を有する。また、処理部300は、ディスク制御部308, SAS I/F 制御部309, メモリ制御部310, カーネル部311, システム構成制御部312, 保守インタフェース制御部313を有する。

【0028】

FC I/F 制御部301は、ホスト10とFC接続する場合のFCプロトコルチップの制御を行なう。iSCSI I/F 制御部302は、ホスト10とiSCSI接続する場合のiSCSIプロトコルチップの制御を行なう。

【0029】

SCSIターゲット制御部303は、SCSIターゲットの制御を行なう。SCSIターゲット制御部303は、例えば、ホスト管理, LU管理, コマンド実行管理, コマンド以外の異常系イベント管理などを行なう。なお、SCSIターゲット制御部303の詳細な機能については、図3を参照しながら後述する。

【0030】

SCSIコマンド制御部304は、SCSIコマンド種ごとの制御を行なう。リソース制御部305は、全CMをまたいで論理ボリューム管理, ホストパス管理, 負荷管理を行なう。キャッシュ制御部306は、キャッシュメモリの管理を行なう。RAID制御部307は、RAIDグループの管理を行なう。ディスク制御部308は、ディスクコマンド制御を行なう。SAS I/F 制御部309は、HDD 101a ~ 101dとSAS接続するプロトコルチップの制御を行なう。

【0031】

メモリ制御部310は、各制御部が使用する制御メモリを管理する。カーネル部311は、基本ソフト部であり、各制御部の下位に位置する。システム構成制御部312は、ストレージ装置100の構成情報を管理する。保守インターフェース制御部313は、保守用のユーザインターフェースを制御する。

【0032】

〔1-5〕本実施形態に係るSCSIターゲット制御部の機能構成およびQoS制御域の情報

次に、図3を参照しながら、SCSIターゲット制御部303の機能構成およびQoS制御域213の情報について説明する。図3は、本実施形態に係るSCSIターゲット制御部303の機能構成およびQoS制御域213の情報を示すブロック図である。

【0033】

図3に示すように、SCSIターゲット制御部303は、後述するQoS管理テーブル501や、後述する各LU (LUN # 1 ~ # m; mは2以上の整数) の性能融通情報や、後述するミクロ融通有フラグなどに基づき、制御を行なう。QoS管理テーブル501や、各LUの性能融通情報や、ミクロ融通有フラグは、例えば、メモリ210a, 210bの制御域212a, 212bにおけるQoS制御域213に保存される。

【0034】

また、図3に示すように、SCSIターゲット制御部303は、待ちキュー格納部502, 性能設定部503, コマンド受信部504, 算出部505, スケジュール部506, 実行部507を有する。さらに、SCSIターゲット制御部303は、マクロ融通チェック部510, マクロ融通処理部520, マクロ融通復旧処理部530, ミクロ融通処理部540を有する。なお、以下に説明するSCSIターゲット制御部303の各機能は、メモリ210aおよびメモリ210bと、CPU220aおよびCPU220bとが協働して実現される。

【0035】

10

20

30

40

50

〔 1 - 5 - 1 〕 優先順位と性能値との対応関係、および、QoS 管理テーブル

QoS 管理テーブル 501 は、対象毎に、「優先順位」と「性能値」と「実行可能な入出力処理数」と「実行した入出力処理数」とを対応付けた情報を記憶する。ここで、QoS 管理テーブル 501 の詳細な説明をする前に、まず、図 4 を参照しながら「優先順位」と「性能値」との対応関係について説明する。図 4 は、優先順位と性能値との対応関係の一例を示す図である。

【 0036 】

図 4 に示すように、「優先順位」は、優先度を示し、例えば、1 から 15 までの 15 段階のレベルとして設定される。なお、優先順位のレベルは、例示に過ぎず、1 から 15 の 15 段階に限定されるものではない。

【 0037 】

また、「優先順位」の段階毎に「性能値」が対応付けられる。言い換えると、「優先順位」は、性能制限を示す。「性能値」としては、単位時間当たりの通信量である MB/s (megabytes per second) と、単位時間当たりに実行可能な入出力処理数である IOPS (Input Output Per Second) とが用いられる。「優先順位」(性能設定レベル)のそれぞれに対する具体的な「性能値 (MB/s)」および「性能値 (IOPS)」は、例えば、ディスクドライブなどの現実的な性能を考慮して、図 4 に示すように設定される。

【 0038 】

図 4 に示す例では、「優先順位」が「1」に設定される場合、単位時間当たりの通信量が「800 (MB/s)」であり、単位時間当たりに「15000」の入出力処理が実行されることを示す。同様に、「優先順位」が「3」に設定される場合、単位時間当たりの通信量が「600 (MB/s)」であり、単位時間当たりに「10000」の入出力処理が実行されることを示す。

【 0039 】

また、図 4 に示す例では、優先順位 1 ~ 8 での性能値 (MB/s) の間隔は 100 であり、優先順位 8 ~ 10 での性能値 (MB/s) の間隔は 30 であり、優先順位 10 ~ 11 での性能値 (MB/s) の間隔は 15 であり、優先順位 11 ~ 15 での性能値 (MB/s) の間隔は 5 である。つまり、性能値 (MB/s) は、「優先順位」の低いレベル (15) から高いレベル (1) に向かって大きくなるように設定されている。

【 0040 】

同様に、優先順位 1 ~ 5 での性能値 (IOPS) の間隔は 2500 であり、優先順位 5 ~ 7 での性能値 (IOPS) の間隔は 2000 であり、優先順位 7 ~ 10 での性能値 (IOPS) の間隔は 200 であり、優先順位 10 ~ 11 での性能値 (IOPS) の間隔は 100 であり、優先順位 11 ~ 15 での性能値 (IOPS) の間隔は 60 である。つまり、性能値 (IOPS) も、「優先順位」の低いレベル (15) から高いレベル (1) に向かって大きくなるように設定されている。

【 0041 】

次に、図 5 を参照しながら、QoS 管理テーブル 501 の一例について説明する。図 5 は、QoS 管理テーブル 501 の一例を示す図である。図 5 に示すように、QoS 管理テーブル 501 は、「対象」、「優先順位」、「N ミリ秒 I/O 数」、「N ミリ秒ブロック数」、「限界 I/O 数」、「限界ブロック数」、「現 I/O 数」および「現ブロック数」を対応付けた情報を記憶する。

【 0042 】

ここで、QoS 管理テーブル 501 が記憶する「対象」は、優先順位を設定する対象を示す。ここで、「ポート」、「ホスト」、「ホストと LUN との組合せ」、「ポートと LUN との組合せ」のそれぞれが優先順位を設定する対象となる。例えば、「対象」には、「FC - CA201a」、「ホスト 10a」、「ホスト 10a - LUN - A」などが格納される。なお、以下の説明では、「対象」として「ホストと LUN との組合せ」または「ポートと LUN との組合せ」を指す場合については、単に、「LU」として適宜記載する。

10

20

30

40

50

【 0 0 4 3 】

「優先順位」は、対象に設定された優先度を示す。ここでは、優先順位のレベルが 1 から 15 までの 15 段階のいずれかのレベルが設定されるものとする。「優先順位」には、例えば「3」、「5」、「7」などのレベルが格納される。

【 0 0 4 4 】

「Nミリ秒 I / O 数」は、所定の時間 (Nミリ秒) の間に実行可能な I / O 数 (入出力要求数, 入出力コマンド数) を示す。ここで、所定の時間 (Nミリ秒) として、例えば 10 ミリ秒、つまり $N = 10$ が設定される。また、「Nミリ秒 I / O 数」には、例えば「100」、「50」、「10」などの値が格納される。

【 0 0 4 5 】

「Nミリ秒ブロック数」は、所定の時間 (Nミリ秒) の間に書込み可能なブロック数を示す。ブロック数は、1 ブロックを 512 バイトとして、単位時間当たりの通信量 (MB / s) から算出される。なお、ここでは、所定の時間 (Nミリ秒) として、例えば 10 ミリ秒が設定される。すなわち、「Nミリ秒ブロック数」には、10 ミリ秒間に書込み可能なブロック数として、例えば「12288」、「8192」、「4096」などの値が格納される。

【 0 0 4 6 】

「限界 I / O 数」は、現時点での実行可能な I / O 数 (入出力処理量) の上限値を示す。限界 I / O 数は、「Nミリ秒 I / O 数」に「Nミリ秒が経過した回数 (n)」を乗じた値から、後述する「現 I / O 数」を減算することによって算出される。例えば、「限界 I / O 数」には、「 $100 \times n - a$ 」、「 $50 \times n - b$ 」、「 $10 \times n - c$ 」などの値が格納される。なお、「n」は、「Nミリ秒が経過した回数」であり、「a」から「c」は、後述するように、対応する「現 I / O 数」の値である。

【 0 0 4 7 】

「限界ブロック数」は、現時点での書込み可能なブロック数 (入出力処理量) の上限値を示す。限界ブロック数は、「Nミリ秒ブロック数」に「Nミリ秒が経過した回数 (n)」を乗じた値から、後述する「現ブロック数」を減算することによって算出される。例えば、「限界ブロック数」には、「 $12288 \times n - A$ 」、「 $8192 \times n - B$ 」、「 $4096 \times n - C$ 」などの値が格納される。なお、「n」は、「Nミリ秒が経過した回数」であり、「A」から「C」は、後述するように、対応する「現ブロック数」の値である。

【 0 0 4 8 】

「現 I / O 数」は、所定の時間までに要求された I / O 数の累積値を示す。例えば、「現 I / O 数」には、「a」、「b」、「c」などが格納される。「現ブロック数」は、所定の時間までに要求されたブロック数の累積値を示す。例えば、「現ブロック数」には、「A」、「B」、「C」などが格納される。

【 0 0 4 9 】

図 5 に示す例では、QoS 管理テーブル 501 は、「ポート」の優先順位が「3」であり、「Nミリ秒 I / O 数」が「100」であり、「Nミリ秒ブロック数」が「12288」であることを示す。また、QoS 管理テーブル 501 は、「限界 I / O 数」が「 $100 \times n - a$ 」であり、「限界ブロック数」が「 $12288 \times n - A$ 」であり、「現 I / O 数」が「a」であり、「現ブロック数」が「A」であることを示す。

【 0 0 5 0 】

なお、QoS 管理テーブル 501 が記憶する「対象」、「優先順位」、「Nミリ秒 I / O 数」、「Nミリ秒ブロック数」は、後述する性能設定部 503 により初期設定される値である。また、QoS 管理テーブル 501 が記憶する「限界 I / O 数」、「限界ブロック数」、「現 I / O 数」および「現ブロック数」は N ミリ秒ごとに更新される値である。

【 0 0 5 1 】

〔 1 - 5 - 2 〕 待ちキュー格納部

待ちキュー格納部 502 は、キューであり、コマンド受信部 504 が受信したコマンドのうち、所定時間経過後にスケジュールされるコマンドを格納する。また、待ちキュー格

10

20

30

40

50

納部 502 は、ホストと LUN との組合せ毎またはポートと LUN との組合せ毎 (LU 毎) に設けられる。

【0052】

〔1-5-3〕性能設定部

性能設定部 503 は、QoS 管理テーブル 501 が記憶する情報のうち「対象」, 「優先順位」, 「N ミリ秒 I/O 数」, 「N ミリ秒ブロック数」のそれぞれに、管理者から受け付けた値を設定する。また、性能設定部 503 は、表 1 を参照しながら後述するミクロ性能融通およびマクロ性能融通の各種条件や、当該各種条件に係る値などを、管理者からの指示に応じて設定する機能も有する。なお、管理者は、各ホスト 10 の性能や各ホスト 10 が実行するアプリケーションなどを加味して、優先順位や各種条件に係る値などを設定する。さらに、性能設定部 503 は、ストレージ装置 100 に接続するホスト 10 や、ホスト 10 と LUN との組合せ (LU) が変化した場合に、QoS 管理テーブル 501 が記憶する情報を更新する。

10

【0053】

〔1-5-4〕コマンド受信部

コマンド受信部 504 は、ホスト 10 からコマンド (入出力要求) を受信して受け付けるとともに、実行待ちのコマンドがあるか否かを判定する。例えば、コマンド受信部 504 は、実行待ちのコマンドがあると判定した場合、受信したコマンドを待ちキュー格納部 502 に格納させる。また、コマンド受信部 504 は、実行待ちのコマンドがないと判定した場合、後述するスケジュール部 506 に受信したコマンドの実行順をスケジュールさせる。なお、コマンド受信部 504 の詳細な動作 (コマンド受信処理の処理手順) については、図 10 を参照しながら後述する。

20

【0054】

〔1-5-5〕算出部

算出部 505 は、ホスト 10, ポートおよび LU のそれぞれに対して設定された優先順位に基づいて決定される入出力処理量の上限値 (即ち限界 I/O 数と限界ブロック数) を所定の時間 (N ミリ秒) 毎に算出する。ここで、算出部 505 は、10 ミリ秒毎に限界 I/O 数と限界ブロック数とを算出するものとして説明する。例えば、算出部 505 は、10 ミリ秒毎に、限界 I/O 数に N ミリ秒 I/O 数を加算する。また、算出部 505 は、10 ミリ秒毎に、限界ブロック数に N ミリ秒ブロック数を加算する。算出部 505 は、算出した限界 I/O 数と限界ブロック数とを QoS 管理テーブル 501 に格納させる。

30

【0055】

さらに、算出部 505 は、後述するマクロ性能融通に対応すべく、以下のような機能を有する。つまり、算出部 505 は、10 ミリ秒毎に、融通先 LU (後述する融通フラグ F_{IN} がオンの LU) の限界 I/O 数に、上述のごとく N ミリ秒 I/O 数を加算するとともに、後述する融通ユニット数 N_B に対応する I/O 数をさらに加算する (図 7 参照)。同様に、算出部 505 は、10 ミリ秒毎に、当該融通先 LU の限界ブロック数に、上述のごとく N ミリ秒ブロック数を加算するとともに、上記融通ユニット数 N_B に対応するブロック数をさらに加算する (図 7 参照)。

40

【0056】

一方、算出部 505 は、10 ミリ秒毎に、融通元 LU (後述する融通フラグ F_{OUT} がオンの LU) の限界 I/O 数に、上述のごとく N ミリ秒 I/O 数を加算した値から、上記融通ユニット数 N_B に対応する I/O 数を減算する (図 7 参照)。同様に、算出部 505 は、10 ミリ秒毎に、当該融通元 LU の限界ブロック数に、上述のごとく N ミリ秒ブロック数を加算した値から、上記融通ユニット数 N_B に対応するブロック数を減算する (図 7 参照)。

【0057】

〔1-5-6〕スケジュール部

スケジュール部 506 は、ホスト 10 から受け付けたコマンド数と、算出部 505 により算出された限界 I/O 数および限界ブロック数とに基づいて、コマンドの実行順をスケ

50

ジュールする。例えば、スケジュール部 506 は、所定の時間が経過する毎に、コマンドの実行順をスケジュールする。なお、スケジュール部 506 の詳細な動作（QoS I/O 起動スケジュール処理の処理手順）については、図 8 および図 9 を参照しながら後述する。

【0058】

また、スケジュール部 506 は、マークしたホストとマークした LUN とを復元する。そして、スケジュール部 506 は、復元したホストと LUN との組合せから順に、以下の処理を実行する。なお、ホストのマークおよび LUN のマークについては後述する。

【0059】

まず、スケジュール部 506 は、ホスト 10 から受け付けたコマンド数が、算出部 505 により算出されたホスト、ポート、LUN それぞれの限界 I/O 数および限界ブロック数以内で処理可能であるか否かを判定する。ここで、スケジュール部 506 は、受け付けたコマンド数が限界 I/O 数および限界ブロック数以内であり、処理可能であると判定した場合、実行部 507 にコマンドを起動させる。

【0060】

また、スケジュール部 506 は、処理可能であると判定した場合、QoS 管理テーブル 501 が記憶する「現 I/O 数」に、起動したコマンド分の I/O 数を加算した値を格納させる。同様に、スケジュール部 506 は、処理可能であると判定した場合、QoS 管理テーブル 501 が記憶する「現ブロック数」に、起動したコマンド分のブロック数を加算した値を格納させる。

【0061】

また、スケジュール部 506 は、処理可能であると判定した場合、QoS 管理テーブル 501 が記憶する「限界 I/O 数」から、起動したコマンド分の I/O 数を減算した値を格納させる。同様に、スケジュール部 506 は、処理可能であると判定した場合、QoS 管理テーブル 501 が記憶する「限界ブロック数」から、起動したコマンド分のブロック数を減算した値を格納させる。

【0062】

一方、スケジュール部 506 は、受け付けたコマンド数が限界 I/O 数および限界ブロック数以内ではないと判定した場合、所定の時間が経過した後に、ホスト 10 から受け付けたコマンド数が、新たに算出された限界 I/O 数および限界ブロック数以内であるか否かを判定する。

【0063】

例えば、スケジュール部 506 は、受け付けたコマンド数がポートの限界 I/O 数および限界ブロック数以内ではないと判定した場合、ホストと LUN との組合せをマークし、所定の時間が経過後に、マークしたホストと LUN との組合せからスケジュールを実行する。また、例えば、スケジュール部 506 は、ポートの限界 I/O 数および限界ブロック数以内ではあるが、ホストの限界 I/O 数および限界ブロック数以内ではないと判定した場合、LUN をマークする。そして、スケジュール部 506 は、次の LUN をセットし、ホスト 10 から受け付けたコマンド数が、新たに算出された限界 I/O 数および限界ブロック数以内であるか否かを判定する。ここで、次の LUN が無い場合、スケジュール部 506 は、次のホストをセットし、受け付けたコマンド数が、新たに算出された限界 I/O 数および限界ブロック数以内であるか否かを判定する。

【0064】

また、スケジュール部 506 は、コマンド受信部 504 により実行待ちのコマンドがないと判定された場合、所定の時間が経過するのを待つことなくコマンドの実行順をスケジュールする。例えば、スケジュール部 506 は、ホスト 10 から受け付けたコマンド数が、算出部 505 により算出されたホスト、ポート、LUN それぞれの限界 I/O 数および限界ブロック数以内で処理可能であるか否かを判定する。ここで、スケジュール部 506 は、ホスト 10 から受け付けたコマンド数がポート、ホスト、LUN のいずれかの限界 I/O 数および限界ブロック数以内で処理可能ではないと判定した場合、コマンドを待ちキ

キュー格納部 502 に格納する。そして、スケジュール部 506 は、所定の時間が経過した後に、ホスト 10 から受け付けたコマンド数が、新たに算出された限界 I/O 数および限界ブロック数以内であるか否かを判定する。

【0065】

〔1-5-7〕実行部

実行部 507 は、スケジュール部 506 によりスケジュールされた実行順でコマンドを待ちキュー格納部 502 からデキューし、コマンドを起動して入出力処理を実行する。

【0066】

〔1-5-8〕ミクロ性能融通、およびマクロ性能融通

次に、マクロ融通チェック部 510、マクロ融通処理部 520、マクロ融通復旧処理部 530、ミクロ融通処理部 540 について説明するのに先立ち、本実施形態で採用される 2 種類の性能融通方式について説明する。2 種類の性能融通方式は、ミクロ性能融通方式とマクロ性能融通方式とである。

【0067】

なお、以下では、ストレージ装置 100 の処理性能を改善すべく、各 LU（または各ホスト）における入出力処理量の上限值つまり性能限界値を調整することを「性能融通」という。「性能融通」では、処理性能不足を解消するために性能限界値を高く変更する LU を、「融通先 LU」という。また、「性能融通」では、「融通先 LU」の性能限界値を高くする場合、「融通先 LU」の性能限界値を高くする分だけ、処理性能に余裕のある LU の性能限界値を低く変更する。これにより、処理性能に余裕のある LU から「融通先 LU」への性能融通が行なわれる。上述のごとく「融通先 LU」に対して処理性能を融通した「処理性能に余裕のある LU」を、「融通元 LU」という。

【0068】

「ミクロ性能融通」は、一時的なスパイク性能などを吸収する際に用いて有効な方式である。「ミクロ性能融通」では、例えば 1 秒以下のサイクル（ミクロサイクル）で、処理性能が不足する LU に、処理性能に余裕のある LU から処理性能が融通される。このときの融通量（後述するミクロ融通ユニット数 N_H ）は、記憶されず、次のミクロサイクルでは性能限界値の調整に反映されない。従って、ストレージ装置 100 は、次のミクロサイクルでは、元の処理性能で動作する。このように、「ミクロ性能融通」は一時的な融通であるため、「ミクロ性能融通」を用いることで一度に大量の性能融通が可能である。

【0069】

「マクロ性能融通」は、ホスト 10 から要求される処理性能が様々な状況に変動する場合に、性能融通量（性能限界値の調整量）を学習して調整し、各 LU の処理性能を実際の入出力性能に追従させる方式である。「マクロ性能融通」では、例えば 2 秒以上の、より大きなサイクル（マクロサイクル）で処理性能が監視される。そして、所定時間（例えば 1 マクロサイクル）に亘り連続して性能不足となっている LU に対し、前記所定時間に亘り連続して処理性能に余裕のある LU から、処理性能が融通される。このように、「マクロ性能融通」は、性能融通量を学習により調整するため、ホスト 10 から要求される処理性能の継続的な変動に有効に対応することができる。

【0070】

なお、本実施形態の「マクロ性能融通」では、各 LU について、例えば、1 秒毎に処理性能の不足/余裕をチェックすることで、2 秒以上のマクロサイクル毎にマクロ性能融通を実行するか否かの判定が行なわれる。例えば、1 マクロサイクル内において、1 秒毎に、処理性能が不足しているか、あるいは、処理性能に余裕があるかがチェックされる。このとき、1 マクロサイクルの間、常に処理性能が不足している融通先 LU が存在し、且つ、1 マクロサイクルの間、常に処理性能に余裕のある融通元 LU が存在していることが、マクロ性能融通の実行条件として設定される。そして、マクロ性能融通の実行条件を満たす場合、融通元 LU から融通先 LU へのマクロ性能融通が、所定の性能融通単位で実行される。

【0071】

また、本実施形態の「マクロ性能融通」では、1マクロサイクル当たりの融通量の変更は、所定の性能融通単位（ユニット単位）で行なわれる。本実施形態において、所定の性能融通単位（ユニット単位）の入出力処理量は、上記優先度に基づいて決定される入出力処理量（性能値MB/s，IOPS）の最小間隔以下に設定される。以下、所定の性能融通単位の入出力処理量を「ユニット」という。例えば、図4に示す優先順位と性能値との対応関係では、上記優先度についての性能値の最小間隔は、5MB/sおよび60IOPSであるので、「1ユニット」は、例えば、5MB/sおよび60IOPSに設定される。

【0072】

ここで、上述のごとく優先度を変更して負荷状況の変化に対応する従来技術を用いた場合を考える。この場合、図4に示す例では、優先順位1～7の範囲における性能限界値の変更間隔（融通量）は、性能値MB/sについて100、性能値IOPSについて2500または2000であり、優先順位8～15の範囲における変更間隔（融通量）に比べ、かなり大きい。性能限界値の変更間隔が大きいと、優先順位を上下させるだけでは調整範囲が限定的となり、ホスト10から要求される性能の変動に細かく追従することができない。これに対して、本実施形態の「マクロ性能融通」では、上記優先度についての性能値の最小間隔5MB/s，60IOPSを「1ユニット」とする単位で、性能融通が行なわれる。したがって、優先度を変更して性能限界値を変更する場合に比べ、性能限界値の変更間隔（融通量）を大幅に小さくすることができ、ホスト10から要求される処理性能の変動に細かく追従することができ、様々な性能変動に対応することができる。

【0073】

本実施形態では、上述したミクロ性能融通およびマクロ性能融通の特徴を踏まえ、ミクロ性能融通およびマクロ性能融通のいずれか一方が使用されてもよいし、両方が同時に使用されてもよい。一方のみを使用する場合、表1におけるパラメータ等を適切に設定することにより、より有効な融通が行なわれる。また、両方を使用する場合、一時的なスパイク性能などを吸収する際にはミクロ性能融通を使用する一方、継続的な性能変動を吸収する際にはマクロ性能融通を使用するように制御を行なうことで、両方の特徴を有効に利用することができる。

【0074】

ついで、下記表1を参照しながら、ミクロ性能融通およびマクロ性能融通の各種条件や、当該各種条件に係る値などについて説明する。当該各種条件や、当該各種条件に係る値などは、前述した通り、性能設定部503により、管理者からの指示に応じて設定される。下記表1のごとく設定された各種条件や値に基づいて、ミクロ性能融通およびマクロ性能融通がそれぞれ実行される。

【0075】

下記表1に示すように、本実施形態の「ミクロ性能融通」および「マクロ性能融通」では、それぞれ、以下のように各種条件や値が設定され使用される。各種条件や値としては、「サイクル」，「1サイクルでの融通LU数」，「融通回数制限」，「融通量」，「融通を受ける条件」，「融通する条件」，「復帰条件」が設定される。

【0076】

【表 1】

	ミクロ性能融通	マクロ性能融通
サイクル	例えば、0.5秒	例えば、30秒
1サイクルでの融通LU数	最大4LU。16LUまでチューニング可能。	最大4LU。16LUまでチューニング可能。
融通回数制限	マクロ融通1サイクル当たり4回まで。 チューニング可能。	なし。 ただし、性能値の上限および下限はあり。
融通量	1回の融通量は設定性能の2倍。 チューニング可能。	1回の融通はユニット単位。 1ユニットは例えば60 IOPSと5MB/sで 両方同時に融通。
融通を受ける条件	1サイクルで設定性能の2～4倍の待ちI/O (I/O数あるいはブロック数)を検出。	30秒間続けて処理性能不足を検出。
融通する条件	上記融通量以上の余裕がある。	30秒間続けて処理性能に余裕がある。
復帰条件	—	他に融通しているLU (融通元LU)で処理性能不足を検出。 QoS関連設定変更時。

10

【0077】

「サイクル」は、性能融通を実行する周期（秒）である。ミクロ性能融通を実行するサイクル（ミクロサイクル）は、例えば、0.5秒とし、マクロ性能融通を実行するサイクル（マクロサイクル）は、例えば、30秒とする。

【0078】

「1サイクルでの融通LU数」は、1サイクルの間に性能融通を行なうLUの数の上限値である。ミクロ性能融通の「1サイクルでの融通LU数」は、例えば、最大4LU（デフォルト）とし、16LUまでチューニング可能とする。また、マクロ性能融通の「1サイ

20

【0079】

「融通回数制限」は、性能融通の実行回数の制限に係る情報である。ミクロ性能融通の「融通回数制限」は、例えば、一つのLUについて、マクロ性能融通1サイクル（マクロサイクル；30秒）当たり4回までとし、チューニング可能とする。また、マクロ性能融通の「融通回数制限」は、「なし」に設定される。ただし、上限および下限は設定される。

【0080】

「融通量」は、性能融通1回当たりに融通される量である。ミクロ性能融通の「融通量」は、例えば、1回当たり、設定性能（例えば図4に示す優先度に応じて設定された性能値）の2倍とし、チューニング可能とする。また、マクロ性能融通の「融通量」は、例えば、1回当たり、上述した1ユニット（所定の性能融通単位）、つまり60 IOPSおよび5MB/sの両方で同時に融通する。

30

【0081】

「融通を受ける条件」は、LUが他のLUから性能融通を受ける条件、つまりLUが融通先LUになる条件である。ミクロ性能融通の「融通を受ける条件」は、1ミクロサイクルで設定性能（性能値）の2～4倍の待ちI/O（I/O数またはブロック数）を検出したLUであることとする。また、マクロ性能融通の「融通を受ける条件」は、30秒（第1所定時間）の間、継続して処理性能不足を検出したLUであることとする。

40

【0082】

「融通する条件」は、LUが他のLUへ性能融通を行なう条件、つまりLUが融通元LUになる条件である。ミクロ性能融通の「融通する条件」は、ミクロ性能融通を行なう際の上記「融通量」以上の余裕があることとする。また、マクロ性能融通の「融通する条件」は、30秒（第2所定時間）の間、継続して処理性能に余裕のあるLUであることとする。

【0083】

「復帰条件」は、融通元LUから融通先LUへ融通された入出力処理量を、融通先LUから融通元LUに復帰（復旧）させる条件である。ミクロ性能融通では、上述したように、融通量は記憶されず次のミクロサイクルでストレージ装置100は元の処理性能で動作

50

するため、ミクロ性能融通の「復帰条件」は設定されない。また、マクロ性能融通の「復帰条件」は、他のLUに性能融通を行なっている融通元LUで処理性能不足を検出することとする。また、マクロ性能融通の復帰（復旧）は、QoS関連設定変更時にも実行される。

【0084】

〔1-5-9〕LUの性能融通情報

次に、上述したミクロ性能融通およびマクロ性能融通を実行すべく、図3に示すように、QoS制御域213に、LU（LUN#1～#m；mは2以上の整数）毎に保持される、各LUの性能融通情報について説明する。つまり、QoS制御域213には、LU毎に、例えば、6種類の融通フラグ F_{IN} 、 F_{OUT} 、 F_{INMAX} 、 F_{OUTMAX} 、 F_{MICIN} 、 F_{MICOUT} と8種類の数値 N_B 、 N_C 、 N_D 、 N_E 、 N_F 、 N_G 、 N_H 、 N_I とが性能融通情報として保存される。これらの融通フラグや数値は、管理者からの指示に応じて性能設定部503により設定されるか、もしくは、ミクロ性能融通およびマクロ性能融通の実行中にSCSインターゲット制御部303の機能によって設定・更新される。

【0085】

融通フラグ F_{IN} は、対象LUが他のLUからマクロ性能融通を受けている場合つまり対象LUが融通先LUである場合にオンに設定され、それ以外の場合にオフに設定される。

融通フラグ F_{OUT} は、対象LUが他のLUに対しマクロ性能融通を行なっている場合つまり対象LUが融通元LUである場合にオンに設定され、それ以外の場合にオフに設定される。

【0086】

融通フラグ F_{INMAX} は、対象LUが他のLUから受けているマクロ性能融通の融通量が予め設定された上限に到達している場合（上限到達時）にオンに設定され、それ以外の場合にオフに設定される。

融通フラグ F_{OUTMAX} は、対象LUから他のLUに対するマクロ性能融通の融通量が予め設定された下限に到達している場合（下限到達時）にオンに設定され、それ以外の場合にオフに設定される。

【0087】

融通フラグ F_{MICIN} は、対象LUが他のLUからミクロ性能融通を受けている場合にオンに設定され、それ以外の場合にオフに設定される。

融通フラグ F_{MICOUT} は、対象LUが他のLUに対しミクロ性能融通を行なっている場合にオンに設定され、それ以外の場合にオフに設定される。

【0088】

融通ユニット数 N_B は、対象LUが他のLUからマクロ性能融通を受けている場合あるいは対象LUが他のLUに対しマクロ性能融通を行なっている場合の融通量に対応するユニット数である。ここで、融通ユニット数 N_B は、当該融通量を、上述した所定の性能融通単位（ユニット）の数に換算した値である。

【0089】

待ちI/O数 N_C は、対象LUにおける、QoS制御で待ち状態のI/O（ホスト10からの入出力要求またはコマンド）の数である。

待ちブロック数 N_D は、対象LUにおける、QoS制御で待ち状態のI/O（ホスト10からの入出力要求またはコマンド）の総ブロック数である。

これらの待ちI/O数 N_C および待ちブロック数 N_D から、QoS制御で待ち状態のI/Oに対応するユニット数（待ちユニット数）が換算される。

【0090】

最小余りユニット数 N_E は、対象LUにおける1秒毎の余りユニット数 N_I の最小値（最小余り量）である。ここで、余りユニット数 N_I は、対象LUに対するI/Oに係る処理量の、対象LUの処理性能に対する処理余り量（余裕値）に対応するもので、当該処理余り量を、上述した所定の性能融通単位（ユニット）の数に換算した値である。

【0091】

10

20

30

40

50

最小待ちユニット数 N_F は、対象 LU における 1 秒毎の待ちユニット数の最小値（最小待ち量）である。ここで、待ちユニット数は、上述したように、待ち I/O 数 N_C および待ちブロック数 N_D から換算される、 QoS 制御で待ち状態の I/O に対応するユニット数である。

【0092】

ミクロ融通数 N_G は、対象 LU がミクロ性能融通を行なう際の「融通回数制限」（表 1 参照）に対応する値で、対象 LU における、1 マクロサイクル（例えば 30 秒）内のミクロ性能融通の実行回数の制限値（例えば 4 回）である。

ミクロ融通ユニット数 N_H は、今回のミクロ性能融通によって、対象 LU が他の LU から受ける一時的な融通量、もしくは、対象 LU が他の LU に対し融通する一時的な融通量に対応するユニット数である。つまり、ミクロ融通ユニット数 N_H は、当該一時的な融通量を、上述した所定の性能融通単位（ユニット）の数に換算した値である。

10

【0093】

余りユニット数 N_I は、最小余りユニット数 N_E の説明において上述した通り、対象 LU に対する I/O に係る処理量の、対象 LU の処理性能に対する処理余り量（余裕値）に対応する。そして、余りユニット数 N_I は、当該処理余り量を、上述した所定の性能融通単位（ユニット）の数に換算した値である。

【0094】

〔1 - 5 - 10〕マクロ融通チェック部

マクロ融通チェック部 510 は、第 1 判定部 521 および第 2 判定部 522 によりマクロ性能融通を実行するか否かを判定するために必要な情報（最小余りユニット数 N_E および最小待ちユニット数 N_F ）をチェックして取得する。このため、マクロ融通チェック部 510 は、第 1 取得部 511 および第 2 取得部 512 を有する。なお、マクロ融通チェック部 510 の詳細な動作（マクロ融通 1 秒チェック処理の処理手順）については、図 13 を参照しながら後述する。

20

【0095】

第 1 取得部 511 は、後述する第 1 判定部 521 がマクロ性能融通の上記「融通を受ける条件」の判定を行なうために必要な情報である最小待ちユニット数 N_F を取得する。つまり、第 1 取得部 511 は、 LU 毎に、1 マクロサイクル（例えば 30 秒；第 1 所定時間）よりも短い時間間隔（例えば 1 秒間隔）で、当該 LU に対する待ちユニット数（ I/O の処理待ち量）を取得する。このとき、1 秒毎の待ちユニット数は、上述したように、待ち I/O 数 N_C および待ちブロック数 N_D から換算される。

30

【0096】

そして、第 1 取得部 511 は、 LU 毎に、1 マクロサイクル中に取得される待ちユニット数のうち最小のものを最小待ちユニット数（最小待ち量） N_F として取得する。このとき、第 1 取得部 511 は、1 マクロサイクル中の 1 秒毎に待ちユニット数を取得する都度、今回取得した待ちユニット数と、 QoS 制御域 213 に保存されている値（現時点までの最小待ちユニット数） N_F とを比較する。今回取得した待ちユニット数が値 N_F よりも小さい場合、第 1 取得部 511 は、値 N_F を、今回取得した待ちユニット数に置き換える。第 1 取得部 511 が以上の処理を繰り返すことにより、1 マクロサイクルが経過した時点で QoS 制御域 213 に保存されている値 N_F が、最小待ちユニット数 N_F となる。

40

【0097】

第 2 取得部 512 は、後述する第 2 判定部 522 がマクロ性能融通の上記「融通する条件」の判定を行なうために必要な情報である最小余りユニット数 N_E を取得する。つまり、第 2 取得部 512 は、 LU 毎に、1 マクロサイクル（例えば 30 秒；第 2 所定時間）よりも短い時間間隔（例えば 1 秒間隔）で、当該 LU についての余りユニット数（処理余り量）を取得する。このとき、1 秒毎の余りユニット数は、 QoS 制御域 213 に保存される余りユニット数 N_I を 1 秒毎に参照することで取得される。

【0098】

そして、第 2 取得部 512 は、 LU 毎に、1 マクロサイクル中に取得される余りユニッ

50

ト数 N_I のうち最小のものを最小余りユニット数 (最小余り量) N_E として取得する。このとき、第2取得部512は、1マクロサイクル中の1秒毎に余りユニット数 N_I を取得する都度、今回取得した余りユニット数 N_I と、QoS制御域213に保存されている値 (現時点までの最小余りユニット数) N_E とを比較する。今回取得した余りユニット数が値 N_E よりも小さい場合、第2取得部512は、値 N_E を、今回取得した余りユニット数 N_I に置き換える。第2取得部512が以上の処理を繰り返すことにより、1マクロサイクルが経過した時点でQoS制御域213に保存されている値 N_E が、最小余りユニット数 N_E となる。

【0099】

〔1-5-11〕マクロ融通処理部

マクロ融通処理部520は、マクロ融通チェック部510の第1取得部511および第2取得部512によりそれぞれ取得された最小待ちユニット数 N_F および最小余りユニット数 N_E に基づき、マクロ性能融通を実行する。マクロ融通処理部520は、第1判定部521、第2判定部522および変更部523を有する。なお、マクロ融通処理部520の詳細な動作 (マクロ融通処理の処理手順) については、図14および図15を参照しながら後述する。

【0100】

第1判定部521は、各LUがマクロ性能融通の上記「融通を受ける条件」を満たすか否かを判定する。つまり、第1判定部521は、LU毎に、当該LUに対するI/Oに係る処理量が1マクロサイクル (第1所定時間; 30秒) に亘って当該LUの処理性能を超えているか否かを判定する。より具体的に、第1判定部521は、当該LUについて第1取得部511により取得された最小待ちユニット数 N_F が0でない場合、当該LUに対するI/Oに係る処理量が1マクロサイクルに亘って当該LUの処理性能を超えていると判定する。つまり、第1判定部521は、当該LUがマクロ性能融通の上記「融通を受ける条件」を満たす融通先LUであると判定する。

【0101】

第2判定部522は、各LUがマクロ性能融通の上記「融通する条件」を満たすか否かを判定する。つまり、第2判定部522は、LU毎に、当該LUに対するI/Oに係る処理量が1マクロサイクル (第2所定時間; 30秒) に亘って当該LUの処理性能を下回っているか否かを判定する。より具体的に、第2判定部522は、当該LUについて第2取得部512により取得された最小余りユニット数 N_E が0でない場合、当該LUに対するI/Oに係る処理量が1マクロサイクルに亘って当該LUの処理性能を下回っていると判定する。つまり、第2判定部522は、当該LUがマクロ性能融通の上記「融通する条件」を満たす融通元LUであると判定する。

【0102】

変更部523は、第1判定部521および第2判定部522による判定結果に従って、対象LUの性能限界値をユニット単位で変更することで、融通元LUから融通先LUへのマクロ性能融通を、1ユニット単位で行なう。つまり、第1判定部521による判定結果が下記項目(a1)に示す場合であり、且つ、第2判定部522による判定結果が下記項目(a2)に示す場合である時、変更部523は、下記項目(a3)に示す変更処理を行なう。

【0103】

(a1) 第1判定部521による判定結果: 一のLU (第1記憶装置) がマクロ性能融通の上記「融通を受ける条件」を満たす融通先LUである場合。

(a2) 第2判定部522による判定結果: 他のLU (第2記憶装置) がマクロ性能融通の上記「融通する条件」を満たす融通元LUである場合。

(a3) 変更処理: 融通先LUについて算出された性能制限値に、上記所定の性能融通単位で1ユニット分に対応する値を加算するとともに、融通元LUについて算出された性能制限値から、上記所定の性能融通単位で1ユニット分に対応する値を減算する。

【0104】

10

20

30

40

50

〔 1 - 5 - 1 2 〕 マクロ融通復旧処理部

マクロ融通復旧処理部 5 3 0 は、各 L U が上記「復旧条件」を満たした場合、融通先 L U に融通した処理量を、融通先 L U から融通元 L U に復旧させる復旧処理を実行する。マクロ融通復旧処理部 5 3 0 は、第 3 判定部 5 3 1 および復旧処理部 5 3 2 を有する。なお、マクロ融通復旧処理部 5 3 0（特に復旧処理部 5 3 2）の詳細な動作（マクロ融通復旧処理の処理手順）については、図 1 6 ~ 図 2 0 を参照しながら後述する。

【 0 1 0 5 】

第 3 判定部 5 3 1 は、各 L U がマクロ性能融通の上記「復旧条件」を満たすか否かを判定する。つまり、第 3 判定部 5 3 1 は、L U 毎に、当該 L U が、一以上の融通先 L U（第 1 記憶装置）へのマクロ性能融通を行なった融通元 L U（第 2 記憶装置）であり、且つ、当該 L U に対する I / O の処理待ちの生じた融通復旧対象であるか否かを判定する。

10

【 0 1 0 6 】

復旧処理部 5 3 2 は、第 3 判定部 5 3 1 により当該 L U が融通復旧対象（以下、融通復旧 L U という）であると判定された場合、融通復旧 L U が他の L U に融通した処理量を一以上の融通先 L U から融通復旧 L U に復旧させる復旧処理を実行する。その際、復旧処理部 5 3 2 は、一以上の融通先 L U に対し融通された処理量の残量が均等もしくはほぼ均等になるように、当該処理量を一以上の融通先 L U から融通復旧 L U に復旧させる。

【 0 1 0 7 】

〔 1 - 5 - 1 3 〕 ミクロ融通処理部

ミクロ融通処理部 5 4 0 は、ミクロ性能融通を実行するもので、第 4 判定部 5 4 1，融通処理部 5 4 2 および融通後処理部 5 4 3 を有する。なお、ミクロ融通処理部 5 4 0 の詳細な動作（ミクロ融通処理およびミクロ融通後処理の処理手順）については、図 1 1 および図 1 2 を参照しながら後述する。ミクロ融通処理部 5 4 0 によるミクロ性能融通の実行回数は、1 マクロサイクル当たり、Q o S 制御域 2 1 3 においてミクロ融通数 N_G として設定される値（例えば 4）に制限される。

20

【 0 1 0 8 】

第 4 判定部 5 4 1 は、各 L U がミクロ性能融通の「融通を受ける条件」を満たすか否かを判定する。つまり、第 4 判定部 5 4 1 は、L U 毎に、当該 L U に対する I / O に係る処理量（待ちユニット数）が当該 L U の処理性能（性能値 / 設定値）の所定数倍以上（例えば 4 倍以上）であるか否かを判定する。

30

【 0 1 0 9 】

融通処理部 5 4 2 および融通後処理部 5 4 3 は、協働して、第 4 判定部 5 4 1 により「融通を受ける条件」を満たすと判定された L U に対し、当該 L U に対する I / O に係る処理量を超える処理余裕をもつ L U からミクロ性能融通を行なう。なお、1 回のミクロ性能融通で融通される「融通量」は、設定性能の 2 倍に制限される（上記表 1 参照）。また、ミクロ性能融通の「融通量」は、Q o S 制御域 2 1 3 においてミクロ融通ユニット数 N_H として保持される。Q o S 制御域 2 1 3 におけるミクロ融通ユニット数 N_H の値は、図 1 1 を参照しながら後述するように、1 ミクロサイクルの開始時にクリアされる。このため、ミクロ融通ユニット数 N_H の値に応じた融通量でのミクロ性能融通は、1 回（1 ミクロサイクル）だけ実行される。

40

【 0 1 1 0 】

融通処理部 5 4 2 は、スケジュール部 5 0 6 による Q o S I / O 起動スケジュール処理前（図 6 参照）に動作する。融通処理部 5 4 2 は、1 ミクロサイクル毎に、大きく性能不足となる（例えば設定値の 4 倍以上の性能不足となる）融通先 L U（待ち L U）を検索し、もし該当する融通先 L U があれば、処理性能に余裕のある融通元 L U（余り L U）を検索する。そして、融通処理部 5 4 2 は、検索された待ち L U の待ちユニット数と検索された余り L U の余りユニット数とを、それぞれ大きい順に並べた待ちユニット配列および余りユニット配列として記憶する。融通処理部 5 4 2 は、待ちユニット配列および余りユニット配列に基づき、ミクロ性能融通を行なう待ち L U と余り L U との組合せを決定する。また、融通処理部 5 4 2 は、決定された組合せの L U の Q o S 制御域 2 1 3 にミクロ融

50

通ユニット数 N_H および融通フラグ F_{MICIN} 、 F_{MICOUT} を設定する。さらに、融通処理部542は、マイクロ性能融通を行なう待ちLUと余りLUとの組合せを一組でも決定した場合、QoS制御域213のマイクロ融通有フラグ(図3参照)をオンに設定し、マイクロ性能融通を行なう旨を融通後処理部543に通知する。融通処理部542は、以上のようなマイクロ融通処理を、図11を参照しながら詳述する処理手順で実行する。

【0111】

融通後処理部543は、スケジュール部506によるQoS I/O起動スケジュール処理後(図6参照)、QoS制御域213のマイクロ融通有フラグ(図3参照)がオンに設定されている場合に動作する。融通後処理部543は、1マイクロサイクル毎に、マイクロ融通ユニット数 N_H および融通フラグ F_{MICIN} 、 F_{MICOUT} に基づいて、待ちLUおよび余りLUの限界値に対する加減算処理を行ない、次のNミリ秒サイクルでマイクロ性能融通を実行する。ただし、本実施形態では、QoS制御用のメモリで保持可能な数値範囲を超えるのを抑止すべく、図6に示すように、1秒毎に、QoS制御に係るQoSカウンタがクリアされる。当該QoSカウンタのクリアに対応すべく、融通後処理部543は、1秒毎のQoSカウンタクリアタイミングとそれ以外のタイミングとで異なる処理を行なっている。各タイミングでの、融通後処理部543による処理については、図12を参照しながら後述する。融通後処理部543は、以上のようなマイクロ融通後処理を、図12を参照しながら詳述する処理手順で実行する。

10

【0112】

〔2〕本実施形態の動作

20

次に、図6～図20を参照しながら、本実施形態に係るストレージ装置100の動作、特に、本実施形態に係るストレージ装置100による処理の処理手順について説明する。

【0113】

〔2-1〕Nミリ秒タイマスケジュール処理

まず、図6に示すフローチャート(ステップS101～S111)に従って、本実施形態のSCSIターゲット制御部303によって実行されるNミリ秒タイマスケジュール処理の処理手順について説明する。Nは、例えば、10であり、SCSIターゲット制御部303は、各ポートについて、10ミリ秒毎に、図6に示すNミリ秒タイマスケジュール処理を実行する。これにより、図7～図9および図11～図20を参照しながら詳述するとく、QoS制御(QoS I/O起動スケジュール)、マクロ性能融通およびマイクロ性能融通が実行される。

30

【0114】

Nミリ秒タイマスケジュール処理が起動されると、まず、算出部505は、各ポートの限界値である限界I/O数および限界ブロック数に、それぞれNミリ秒I/O数およびNミリ秒ブロック数を加算する。加算結果のポート限界値は、QoS管理テーブル501に上書き保存される(ステップS101)。ポート限界値に加算される限界I/O数および限界ブロック数は、QoS設定(図4の性能値)の $N/1000$ ($N=10$ の場合、 $1/100$)の値となる(図5参照)。

【0115】

また、算出部505は、全てのホスト10の限界値である限界I/O数および限界ブロック数に、それぞれNミリ秒I/O数およびNミリ秒ブロック数を加算する。加算結果の全ホスト限界値は、QoS管理テーブル501に上書き保存される(ステップS102)。各ホスト限界値に加算される限界I/O数および限界ブロック数は、QoS設定(図4の性能値)の $N/1000$ ($N=10$ の場合、 $1/100$)の値となる(図5参照)。

40

【0116】

さらに、本実施形態では、LUを性能融通対象としているため、算出部505は、全てのLUのそれぞれに対し、図7に示す全LU限界値加算処理を実行する(ステップS103)。ここで、図7に示すフローチャート(ステップS121～S125)に従って、全LU限界値加算処理の処理手順について説明する。

【0117】

50

図 7 に示すように、算出部 505 は、全ての LU の限界値である限界 I / O 数および限界ブロック数に、それぞれ N ミリ秒 I / O 数および N ミリ秒ブロック数を加算する。加算結果の全ホスト限界値は、QoS 管理テーブル 501 に上書き保存される（ステップ S121）。各ホスト限界値に加算される限界 I / O 数および限界ブロック数は、QoS 設定（図 4 の性能値）の N / 1000（N = 10 の場合、1 / 100）の値となる（図 5 参照）。

【0118】

そして、算出部 505 は、QoS 制御域 213 を参照し、各 LU の融通フラグ F_{IN} がオンであるか否かを判定する（ステップ S122）。融通フラグ F_{IN} がオンである場合（ステップ S122 の YES ルート）、つまり対象 LU がマクロ性能融通を受けている融通先 LU である場合、算出部 505 は、さらに、以下の加算処理を行なう（ステップ S123）。

10

【0119】

即ち、算出部 505 は、対象 LU の限界値である限界 I / O 数に、QoS 制御域 213 の融通ユニット数 N_B に対応する I / O 数をさらに加算する。具体的には、[融通ユニット数 N_B] × [ユニット当たりの I / O 数] の N / 1000 の値が限界 I / O 数に加算される。同様に、算出部 505 は、対象 LU の限界値である限界ブロック数に、QoS 制御域 213 の融通ユニット数 N_B に対応する I / O 数をさらに加算する。具体的には、[融通ユニット数 N_B] × [ユニット当たりのブロック数] の N / 1000 の値が限界ブロック数に加算される。加算結果は、QoS 管理テーブル 501 に上書き保存される。

20

【0120】

ステップ S123 の処理後、もしくは、融通フラグ F_{IN} がオフである場合（ステップ S122 の NO ルート）、算出部 505 は、QoS 制御域 213 を参照し、各 LU の融通フラグ F_{OUT} がオンであるか否かを判定する（ステップ S124）。融通フラグ F_{OUT} がオンである場合（ステップ S124 の YES ルート）、つまり対象 LU がマクロ性能融通を行っている融通元 LU である場合、算出部 505 は、さらに、以下の減算処理を行なう（ステップ S125）。

【0121】

即ち、算出部 505 は、対象 LU の限界値である限界 I / O 数から、QoS 制御域 213 の融通ユニット数 N_B に対応する I / O 数を減算する。具体的には、[融通ユニット数 N_B] × [ユニット当たりの I / O 数] の N / 1000 の値が限界 I / O 数から減算される。同様に、算出部 505 は、対象 LU の限界値である限界ブロック数から、QoS 制御域 213 の融通ユニット数 N_B に対応する ブロック数 を減算する。具体的には、[融通ユニット数 N_B] × [ユニット当たりのブロック数] の N / 1000 の値が限界ブロック数から減算される。減算結果は、QoS 管理テーブル 501 に上書き保存される。

30

【0122】

ステップ S125 の処理後、もしくは、融通フラグ F_{OUT} がオフである場合（ステップ S124 の NO ルート）、算出部 505 は、全 LU 限界値加算処理を完了する。

上述のような全 LU 限界値加算処理（ステップ S103, S121 ~ S125）を N ミリ秒毎に実行することで、マクロ性能融通が N ミリ秒毎に実行される。なお、ホスト 10 を性能融通対象とする場合には、算出部 505 は、ステップ S102 において、全てのホストのそれぞれに対し、図 7 に示す全 LU 限界値加算処理と同様の全ホスト限界値加算処理を実行する。

40

【0123】

さて、限界値加算処理（ステップ S101 ~ S103）を完了すると、図 6 に示すように、第 4 判定部 541 および融通処理部 542 によるミクロ融通処理が、例えば 0.5 秒（1 ミクロサイクル）毎に実行される（ステップ S104）。第 4 判定部 541 および融通処理部 542 によるミクロ融通処理については、図 11 を参照しながら後述する。

【0124】

そして、マクロ融通チェック部 510 によるマクロ融通 1 秒チェック処理が、1 秒毎に

50

実行されてから（ステップ S 1 0 5）、第 3 判定部 5 3 1 により融通元 L U で処理性能不足が発生したか否かが、N ミリ秒毎に判定される（ステップ S 1 0 6）。融通元 L U で処理性能不足が発生した場合（ステップ S 1 0 6 の Y E S ルート）、つまり融通復旧 L U が存在する場合、復旧処理部 5 3 2 によるマクロ融通復旧処理が実行される（ステップ S 1 0 7）。上述したステップ S 1 0 5 ~ S 1 0 7 の処理は、マクロ融通前処理に相当する。なお、マクロ融通チェック部 5 1 0 によるマクロ融通 1 秒チェック処理については、図 1 3 を参照しながら後述し、復旧処理部 5 3 2 によるマクロ融通復旧処理については、図 1 6 および図 1 7 を参照しながら後述する。

【 0 1 2 5 】

ステップ S 1 0 7 の処理後、もしくは、融通復旧 L U が存在しない場合（ステップ S 1 0 6 の N O ルート）、スケジュール部 5 0 6 は、Q o S I / O 起動スケジュール処理を実行する（ステップ S 1 0 8）。Q o S I / O 起動スケジュール処理については、図 8 および図 9 を参照しながら後述する。

【 0 1 2 6 】

ついで、マクロ融通処理部 5 2 0 によるマクロ融通処理が、例えば 3 0 秒（1 マクロサイクル）毎に実行され（ステップ S 1 0 9）、1 秒毎に、Q o S 制御に係る Q o S カウントがクリアされる（ステップ S 1 1 0）。Q o S カウントのクリアにより、Q o S 制御用のメモリで保持可能な数値範囲を超えることが抑止される。この後、融通後処理部 5 4 3 によるミクロ融通後処理が、例えば 0 . 5 秒（1 ミクロサイクル）毎に実行され（ステップ S 1 1 1）、N ミリ秒タイマスケジュールの 1 サイクル分（N ミリ秒毎）の処理が完了する。マクロ融通処理部 5 2 0 によるマクロ融通処理については、図 1 4 および図 1 5 を参照しながら後述する。また、融通後処理部 5 4 3 によるミクロ融通後処理については、図 1 2 を参照しながら後述する。

【 0 1 2 7 】

〔 2 - 2 〕 Q o S I / O 起動スケジュール処理

次に、図 8 および図 9 に示すフローチャート（ステップ S 2 0 1 ~ S 2 1 7）に従って、本実施形態のスケジュール部 5 0 6 によって実行される Q o S I / O 起動スケジュール処理（図 6 のステップ S 1 0 8）の処理手順について詳細に説明する。Q o S I / O 起動スケジュール処理では、L U 毎に設けた待ちキュー内にコマンドが待機している場合、以下のように、当該 L U の限界 I / O 数および限界ブロック数に従って、待ちキュー内のコマンドが実行される。

【 0 1 2 8 】

スケジュール部 5 0 6 は、マークしたホストを復元する（ステップ S 2 0 1）。続いて、スケジュール部 5 0 6 は、マークした L U を復元する（ステップ S 2 0 2）。そして、スケジュール部 5 0 6 は、待ちキュー格納部 5 0 2 内に次のコマンドがあるか否かを判定する（ステップ S 2 0 3）。

【 0 1 2 9 】

ここで、スケジュール部 5 0 6 は、待ちキュー格納部 5 0 2 内に次のコマンドがあると判定した場合（ステップ S 2 0 3 の Y E S ルート）、コマンド数がポートの限界 I / O 数および限界ブロック数以内であるか否かを判定する（ステップ S 2 0 4）。そして、スケジュール部 5 0 6 は、コマンド数がポートの限界 I / O 数および限界ブロック数以内であると判定した場合（ステップ S 2 0 4 の Y E S ルート）、コマンド数がホストの限界 I / O 数および限界ブロック数以内であるか否かを判定する（ステップ S 2 0 5）。

【 0 1 3 0 】

ここで、スケジュール部 5 0 6 は、コマンド数がホストの限界 I / O 数および限界ブロック数以内であると判定した場合（ステップ S 2 0 5 の Y E S ルート）、コマンド数が L U の限界 I / O 数および限界ブロック数以内であるか否かを判定する（ステップ S 2 0 6）。そして、スケジュール部 5 0 6 は、コマンド数が L U の限界 I / O 数および限界ブロック数以内であると判定した場合（ステップ S 2 0 6 の Y E S ルート）、ポートの限界値（限界 I / O 数および限界ブロック数）から、コマンド分の I / O 数およびブロック数を

減算する（ステップS207）。

【0131】

また、スケジュール部506は、ホストの限界値（限界I/O数および限界ブロック数）から、コマンド分のI/O数およびブロック数を減算する（ステップS208）。続いて、スケジュール部506は、LUの限界値（限界I/O数および限界ブロック数）から、コマンド分のI/O数およびブロック数を減算する（ステップS209）。そして、実行部507は、コマンドを起動する（ステップS210）。また、コマンドを起動した後、実行部507は、コマンドを待ちキュー格納部502からデキューする（ステップS211）。この処理の後、ストレージ装置100は、ステップS203に移行する。

【0132】

また、スケジュール部506は、コマンド数がポートの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS204のNORルート）、ホストとLUとをマークし（ステップS212）、処理を終了する。また、スケジュール部506は、コマンド数がホストの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS205のNORルート）、LUをマークし（ステップS213）、ステップS214に移行する。

【0133】

また、スケジュール部506は、待ちキュー格納部502内に次のコマンドがないと判定した場合（ステップS203のNORルート）、ステップS214に移行する。また、スケジュール部506は、コマンド数がLUの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS206のNORルート）、ステップS214に移行する。

【0134】

ステップS214において、スケジュール部506は、次のLUをセットする。そして、スケジュール部506は、次のLUがマークしたLUと同じLUであるか否かを判定する（ステップS215）。ここで、スケジュール部506は、次のLUがマークしたLUと同じLUではないと判定した場合（ステップS215のNORルート）、ステップS203に移行する。

【0135】

一方、スケジュール部506は、次のLUがマークしたLUと同じLUであると判定した場合（ステップS215のYESルート）、次のホストをセットする（ステップS216）。続いて、スケジュール部506は、次のホストがマークしたホストと同じホストであるか否かを判定する（ステップS217）。ここで、スケジュール部506は、次のホストがマークしたホストと同じホストではないと判定した場合（ステップS217のNORルート）、ステップS202に移行する。一方、スケジュール部506は、次のホストがマークしたホストと同じホストであると判定した場合（ステップS217のYESルート）、処理を終了する。

【0136】

〔2-3〕コマンド受信時の処理

ここで、図10に示すフローチャート（ステップS301～S309）に従って、本実施形態のコマンド受信部504によって実行されるコマンド受信処理の処理手順について詳細に説明する。ストレージ装置100は、例えば、コマンドを受信したことを契機にこのコマンド受信処理を実行する。コマンド受信処理では、ホストからコマンドを受信した際、以下のように、受信したコマンドは、限界I/O数および限界ブロック数に応じて、即座に実行されるか、もしくは、待ちキューにエンキューされる。

【0137】

図10に示すように、コマンド受信部504は、受信したコマンドの発行元ホストとLUとの間で待ちI/Oがあるか否かを判定する（ステップS301）。ここで、コマンド受信部504により発行元ホストとLUとの間で待ちI/Oがないと判定された場合（ステップS301のNORルート）、スケジュール部506は、以下の処理を実行する。すなわち、スケジュール部506は、コマンド数がポートの限界I/O数および限界ブロッ

10

20

30

40

50

ク数以内であるか否かを判定する（ステップS302）。そして、スケジュール部506は、コマンド数がポートの限界I/O数および限界ブロック数以内であると判定した場合（ステップS302のYESルート）、コマンド数がホストの限界I/O数および限界ブロック数以内であるか否かを判定する（ステップS303）。

【0138】

ここで、スケジュール部506は、コマンド数がホストの限界I/O数および限界ブロック数以内であると判定した場合（ステップS303のYESルート）、コマンド数がLUの限界I/O数および限界ブロック数以内であるか否かを判定する（ステップS304）。そして、スケジュール部506は、コマンド数がLUの限界I/O数および限界ブロック数以内であると判定した場合（ステップS304のYESルート）、ポートの限界値（限界I/O数および限界ブロック数）から、コマンド分のI/O数とブロック数とを減算する（ステップS305）。

10

【0139】

また、スケジュール部506は、ホストの限界値（限界I/O数および限界ブロック数）から、コマンド分のI/O数およびブロック数を減算する（ステップS306）。続いて、スケジュール部506は、LUの限界値（限界I/O数および限界ブロック数）から、コマンド分のI/O数およびブロック数を減算する（ステップS307）。また、実行部507は、コマンドを起動する（ステップS308）。ストレージ装置100は、ステップS308の終了後、処理を終了する。

【0140】

20

コマンド受信部504は、受信したコマンドの発行元ホストとLUとの間で待ちI/Oがあると判定した場合（ステップS301のYESルート）、ステップS309に移行する。また、スケジュール部506は、コマンド数がポートの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS302のNORルート）、ステップS309に移行する。また、スケジュール部506は、コマンド数がホストの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS303のNORルート）、ステップS309に移行する。また、スケジュール部506は、コマンド数がLUの限界I/O数および限界ブロック数以内ではないと判定した場合（ステップS304のNORルート）、ステップS309に移行する。そして、コマンド受信部504は、ステップS309において、コマンドを待ちキュー格納部502にエンキューし、処理を終了する。

30

【0141】

〔2-4〕ミクロ融通処理

次に、図11に示すフローチャート（ステップS401～S411）に従って、本実施形態のミクロ融通処理部540（第4判定部541および融通処理部542）によって実行されるミクロ融通処理（図6のステップS104）の処理手順について説明する。ミクロ融通処理は、図6に示すNミリ秒タイマスケジュール処理がNミリ秒毎に実行される中で、1ミクロサイクル（例えば0.5秒＝500ミリ秒）毎に、スケジュール部506によるQoS I/O起動スケジュール処理前に実行される。

【0142】

図11に示すように、ミクロ融通処理部540は、0.5秒毎に起動されると、まず、各LUについて、QoS制御域213の融通フラグ F_{MICIN} および F_{MICOUT} をクリアしてオフ状態にする（ステップS401）。また、ミクロ融通処理部540は、各LUについて、QoS制御域213のミクロ融通ユニット数 N_H をクリアして0にする（ステップS402）。これにより、前回のミクロ性能融通に係る情報が全てクリアされる。従って、ミクロ融通ユニット数 N_H の値に応じた融通量でのミクロ性能融通は、1回（1ミクロサイクル）だけ実行される。

40

【0143】

この後、第4判定部541は、LU毎に、当該LUに対するI/Oに係る処理量に対応する待ちユニット数が当該LUの処理性能（性能値/設定値）の所定数倍以上（例えば4倍以上）であるか否かを判定する（ステップS403）。待ちユニット数は、QoS制御

50

域 2 1 3 の待ち I / O 数 N_C および待ちブロック数 N_D から換算される。そして、待ちユニット数が設定値の所定数倍以上である場合（ステップ S 4 0 3 の Y E S ルート）、融通処理部 5 4 2 は、今回の対象 L U の待ちユニット数を、待ちユニット数の大きい順に並べた待ちユニット配列（リスト）の適切な箇所に挿入する（ステップ S 4 0 4）。マイクロ性能融通用の待ちユニット配列（リスト）は、例えば Q o S 制御域 2 1 3 内に作成され、L U を特定する L U N と当該 L U における待ちユニット数とを対応付けて保存している。なお、本実施形態において、マイクロ性能融通での「1 サイクルでの融通 L U 数」は、例えば、4 L U としているので、待ちユニット配列における要素数は 4 で十分である。

【 0 1 4 4 】

ステップ S 4 0 4 の処理後、もしくは、待ちユニット数が設定値の所定数倍未満である場合（ステップ S 4 0 3 の N O ルート）、融通処理部 5 4 2 は、全ての L U について待ちユニット数の判定を行なったか否かを判定する（ステップ S 4 0 5）。全ての L U について待ちユニット数の判定を行っていない場合（ステップ S 4 0 5 の N O ルート）、マイクロ融通処理部 5 4 0 は、ステップ S 4 0 1 の処理に戻る。一方、全ての L U について待ちユニット数の判定を行なった場合（ステップ S 4 0 5 の Y E S ルート）、融通処理部 5 4 2 は、マイクロ性能融通用の待ちユニット配列に待ちユニットが存在するか否かを判定する（ステップ S 4 0 6）。

【 0 1 4 5 】

待ちユニット配列に待ちユニットが存在する場合（ステップ S 4 0 6 の Y E S ルート）、融通処理部 5 4 2 は、今回の対象 L U の余りユニット数を、余りユニット数の大きい順に並べた余りユニット配列（リスト）の適切な箇所に挿入する（ステップ S 4 0 7）。マイクロ性能融通用の余りユニット配列（リスト）は、例えば Q o S 制御域 2 1 3 内に作成され、L U を特定する L U N と当該 L U における余りユニット数とを対応付けて保存している。対象 L U の余りユニット数としては、Q o S 制御域 2 1 3 に保存される余りユニット数 N_I が読み出されて用いられる。

【 0 1 4 6 】

この後、融通処理部 5 4 2 は、待ちユニット配列における待ちユニット数と、余りユニット配列における余りユニット数とを、大きい方から順に比較し、待ちユニット数が余りユニット数以下であるか否かを判定する（ステップ S 4 0 8）。待ちユニット数が余りユニット数以下である場合（ステップ S 4 0 8 の Y E S ルート）、融通処理部 5 4 2 は、このときの待ち L U と余り L U とを、マイクロ性能融通を行なう待ち L U と余り L U との組合せとして決定する。なお、組合せが決定された場合、当該組合せの待ち L U と余り L U とは、待ちユニット配列および余りユニット配列から削除される。

【 0 1 4 7 】

そして、融通処理部 5 4 2 は、決定された組合せの待ち L U の Q o S 制御域 2 1 3 において、余り L U から融通されるマイクロ融通ユニット数 N_H を設定するとともに、当該待ち L U がマイクロ性能融通を受けることを示す融通フラグ F_{MICIN} をオンに設定する。また、融通処理部 5 4 2 は、決定された組合せの余り L U の Q o S 制御域 2 1 3 において、待ち L U に融通するマイクロ融通ユニット数 N_H を設定するとともに、当該余り L U がマイクロ性能融通を行なっていることを示す融通フラグ F_{MICOUT} をオンに設定する（ステップ S 4 0 9）。

【 0 1 4 8 】

さらに、融通処理部 5 4 2 は、マイクロ性能融通を行なう待ち L U と余り L U との組合せを一組でも決定すると、Q o S 制御域 2 1 3 のマイクロ融通有フラグをオンに設定し（ステップ S 4 1 0）、マイクロ性能融通を行なう旨を融通後処理部 5 4 3 に通知する。そして、融通処理部 5 4 2 は、待ち配列にまだ待ちユニットが有るか否かを判定する（ステップ S 4 1 1）。待ちユニットが有る場合（ステップ S 4 1 1 の Y E S ルート）、融通処理部 5 4 2 は、ステップ S 4 0 8 の処理に戻る。

【 0 1 4 9 】

待ちユニットが無い場合（ステップ S 4 1 1 の N O ルート）、マイクロ融通処理部 5 4 0

10

20

30

40

50

は、ミクロ融通処理を完了する。また、待ちユニット数が余りユニット数よりも大きい場合（ステップS 4 0 8のN O ルート）、ミクロ融通処理部5 4 0は、ミクロ融通処理を完了する。同様に、待ちユニット配列に待ちユニットが存在しない場合（ステップS 4 0 6のN O ルート）、ミクロ融通処理部5 4 0は、ミクロ融通処理を完了する。

【0 1 5 0】

〔2 - 5〕ミクロ融通後処理

次に、図1 2に示すフローチャート（ステップS 5 0 1～S 5 0 7）に従って、本実施形態の融通後処理部5 4 3によって実行されるミクロ融通後処理（図6のステップS 1 1 1）の処理手順について説明する。ミクロ融通後処理は、図6に示すNミリ秒タイムスケジュール処理がNミリ秒毎に実行される中で、1ミクロサイクル（0.5秒）毎に、スケジュール部5 0 6によるQ o S I / O 起動スケジュール処理後に実行される。特に、ミクロ融通後処理は、図6に示すように、Q o S I / O 起動スケジュール処理およびマクロ融通処理が実行され、1秒毎のQ o S カウントがクリアされた後の、Nミリ秒タイムスケジュール処理の最終ステップS 1 1 1で実行される。

【0 1 5 1】

図1 2に示すように、融通後処理部5 4 3は、まず、Q o S 制御域2 1 3のミクロ融通有フラグがオンであるか否かを判定する（ステップS 5 0 1）。ミクロ融通有フラグがオフである場合（ステップS 5 0 1のN O ルート）、融通後処理部5 4 3は、何ら処理を行なうことなく、ミクロ融通後処理を完了する。

【0 1 5 2】

ミクロ融通有フラグがオンである場合（ステップS 5 0 1のY E S ルート）、融通後処理部5 4 3はQ o S 制御域2 1 3のミクロ融通有フラグをオフに設定する（ミクロ融通有設定クリア後；ステップS 5 0 2）。この後、融通後処理部5 4 3は、各L Uについて、融通フラグF_{MICIN}がオンであるか否かを判定する（ステップS 5 0 3）。

【0 1 5 3】

融通フラグF_{MICIN}がオンである場合（ステップS 5 0 3のY E S ルート）、即ち対象L Uが待ちL Uである場合、融通後処理部5 4 3は、当該L Uの限界I / O数にミクロ融通ユニット数N_H分のI / O数を加算する。また、融通後処理部5 4 3は、当該L Uの限界ブロック数にミクロ融通ユニット数N_H分のブロック数を加算する（ステップS 5 0 4）。

【0 1 5 4】

ステップS 5 0 4の処理後、もしくは、融通フラグF_{MICIN}がオフである場合（ステップS 5 0 3のN O ルート）、融通後処理部5 4 3は、各L Uについて、下記所定の判定条件を満たしているか否かを判定する（ステップS 5 0 5）。ここで、所定の判定条件は、対象L Uの融通フラグF_{MICOUT}がオンで、且つ、今回のタイミングが1秒毎のQ o S カウントクリア後ではないという条件である。

【0 1 5 5】

所定の判定条件を満たす場合（ステップS 5 0 5のY E S ルート）、即ち今回のタイミングが1秒毎のQ o S カウントクリア後ではなく且つ対象L Uが余りL Uである場合、融通後処理部5 4 3は、以下の減算処理を行なう（ステップS 5 0 6）。つまり、融通後処理部5 4 3は、当該L Uの限界I / O数からミクロ融通ユニット数N_H分のI / O数を減算するとともに、当該L Uの限界ブロック数からミクロ融通ユニット数N_H分のブロック数を減算する。

【0 1 5 6】

所定の判定条件を満たさない場合（ステップS 5 0 5のN O ルート）、つまり対象L Uの融通フラグF_{MICOUT}がオフ、または、今回のタイミングが1秒毎のQ o S カウントクリア後である場合、融通後処理部5 4 3は、ステップS 5 0 6の減算処理を行なうことなくステップS 5 0 7の処理に移行する。今回のタイミングが1秒毎のQ o S カウントクリア後である場合、Q o S カウントがクリアされているため、減算処理を行なう必要が無いからである。

【 0 1 5 7 】

ステップ S 5 0 6 の処理後、もしくは、ステップ S 5 0 5 で N O 判定後、融通後処理部 5 4 3 は、判定対象の L U がまだ有るか否かを判定する（ステップ S 5 0 7）。判定対象の L U がまだ有る場合（ステップ S 5 0 7 の Y E S ルート）、融通後処理部 5 4 3 は、ステップ S 5 0 3 の処理に戻る。判定対象の L U が無い場合（ステップ S 5 0 7 の N O ルート）、融通後処理部 5 4 3 は、ミクロ融通後処理を完了する。

【 0 1 5 8 】

〔 2 - 6 〕マクロ融通 1 秒チェック処理

次に、図 1 3 に示すフローチャート（ステップ S 6 0 1 ~ S 6 0 8）に従って、本実施形態のマクロ融通チェック部 5 1 0 によって実行されるマクロ融通 1 秒チェック処理（図 6 のステップ S 1 0 5）の処理手順について説明する。マクロ融通 1 秒チェック処理では、第 1 判定部 5 2 1 および第 2 判定部 5 2 2 によりマクロ性能融通を実行するか否かを判定するために必要な情報である、最小余りユニット数 N_E および最小待ちユニット数 N_F が取得される。マクロ融通 1 秒チェック処理は、図 6 に示す N ミリ秒タイムスケジュール処理が N ミリ秒毎に実行される中で、例えば 1 秒毎に実行される。

【 0 1 5 9 】

図 1 3 に示すように、マクロ融通チェック部 5 1 0 は、まず、復帰値のデフォルト値として“FALSE”を Q o S 制御域 2 1 3 等に記憶させる（ステップ S 6 0 1）。ここで、復帰値“FALSE”は、マクロ性能融通による融通量を復帰させる融通復旧 L U が存在しないことを示す情報である。この後、マクロ融通チェック部 5 1 0 は、各 L U について、以下の処理（ステップ S 6 0 2 ~ S 6 0 8）を実行する。

【 0 1 6 0 】

マクロ融通チェック部 5 1 0 は、対象 L U の融通フラグ F_{OUT} がオンで、且つ、対象 L U に待ち I / O があるか否か、つまり、対象 L U が、融通元 L U でありながら処理性能不足の生じた融通復旧 L U であるか否かを判定する（ステップ S 6 0 2）。

【 0 1 6 1 】

対象 L U が融通復旧 L U である場合（ステップ S 6 0 2 の Y E S ルート）、マクロ融通チェック部 5 1 0 は、復帰値として“TRUE”を Q o S 制御域 2 1 3 等に記憶させる（ステップ S 6 0 8）。ここで、復帰値“TRUE”は、融通復旧 L U が存在することを示す情報である。この後、マクロ融通チェック部 5 1 0 は、ステップ S 6 0 7 の処理に移行する。

【 0 1 6 2 】

対象 L U が融通復旧 L U でない場合（ステップ S 6 0 2 の N O ルート）、第 2 取得部 5 1 2 は、当該 L U の現在の余りユニット数 N_I を取得する。各 L U の余りユニット数 N_I は、Q o S 制御域 2 1 3 に保存される各 L U の余りユニット数 N_I を参照することで取得される。そして、第 2 取得部 5 1 2 は、今回取得した余りユニット数 N_I と、Q o S 制御域 2 1 3 に保存されている値（現時点までの最小余りユニット数） N_E とを比較する（ステップ S 6 0 3）。

【 0 1 6 3 】

今回取得した余りユニット数 N_I が値 N_E よりも小さい場合（ステップ S 6 0 3 の Y E S ルート）、第 2 取得部 5 1 2 は、値 N_E を今回取得した余りユニット数 N_I に置き換える（ステップ S 6 0 4）。

【 0 1 6 4 】

ステップ S 6 0 4 の処理後、もしくは、今回取得した余りユニット数 N_I が値 N_E 以上である場合（ステップ S 6 0 3 の N O ルート）、第 1 取得部 5 1 1 は、当該 L U の現在の待ちユニット数を取得する。各 L U の待ちユニット数は、上述したように対象 L U の待ち I / O 数 N_C および待ちブロック数 N_D から換算される。そして、第 1 取得部 5 1 1 は、今回取得した待ちユニット数と、Q o S 制御域 2 1 3 に保存されている値（現時点までの最小待ちユニット数） N_F とを比較する（ステップ S 6 0 5）。

【 0 1 6 5 】

今回取得した待ちユニット数が値 N_F よりも小さい場合（ステップ S 6 0 5 の Y E S ル

ート)、第1取得部511は、値 N_F を、今回取得した待ちユニット数に置き換える(ステップS606)。

【0166】

ステップS606やS608の処理後、または、今回取得した待ちユニット数が値 N_F 以上である場合(ステップS605のNORルート)、マクロ融通チェック部510は、全てのLUについてマクロ融通1秒チェック処理を終了したか否か判定する(ステップS607)。全てのLUについてチェック処理を終了していない場合(ステップS607のNORルート)、マクロ融通チェック部510は、ステップS602の処理に戻る。全てのLUについてチェック処理を終了した場合(ステップS607のYESルート)、マクロ融通チェック部510は、マクロ融通1秒チェック処理を完了する。

10

【0167】

以上のようなマクロ融通1秒チェック処理を、1マクロサイクル(30秒)の間、1秒毎に実行することで、最終的にQoS制御域213に保存されている値 N_E が、1マクロサイクル中の最小余りユニット数 N_E となる。また、同様に、最終的にQoS制御域213に保存されている値 N_F が、1マクロサイクル中の最小待ちユニット数 N_F となる。なお、ステップS603およびS604の処理と、ステップS605およびS606の処理とは、逆の順序で実行されてもよい。

【0168】

〔2-7〕マクロ融通処理

次に、図14および図15に示すフローチャート(ステップS701~S713)に従って、本実施形態のマクロ融通処理部520によって実行されるマクロ融通処理(図6のステップS109)の処理手順について説明する。マクロ融通処理は、図6に示すNミリ秒タイマスケジュール処理がNミリ秒毎に実行される中で、1マクロサイクル(例えば30秒)毎に、スケジュール部506によるQoS I/O起動スケジュール処理後に起動される。マクロ融通処理では、QoS I/O起動スケジュール後に、1マクロサイクル毎のタイミングで、連続的に処理性能に余裕のあるLUから連続的に処理性能が不足しているLUにマクロ性能融通が行なわれる。

20

【0169】

まず、第1判定部521は、各LUについて、当該LUの最小待ちユニット数 N_F が0以外の値で、且つ、当該LUの融通フラグ F_{INMAX} がオフであるかを判定する(ステップS701)。ここで、当該LUの最小待ちユニット数 N_F が0以外の値である場合、当該LUは、1マクロサイクルにおける一秒刻みのタイミングでは、常に待ち状態である。したがって、当該LUは、1マクロサイクルに亘って処理性能不足の状態であるとみなせ、本実施形態では、上記「融通を受ける条件」を満たしているものと判断される。ただし、当該LUの融通フラグ F_{INMAX} がオンである場合、当該LUにおいては、他のLUから受けているマクロ性能融通の融通量が上限に到達しており、マクロ性能融通を行なうことができない。そこで、本実施形態では、各LUについて、当該LUの最小待ちユニット数 N_F が0以外の値で、且つ、当該LUの融通フラグ F_{INMAX} がオフであるかが判定されている。

30

【0170】

当該LUの最小待ちユニット数 N_F が0以外の値で、且つ、当該LUの融通フラグ F_{INMAX} がオフである場合(ステップS701のYESルート)、第1判定部521は、今回の対象LUの最小待ちユニット数 N_F を、待ちユニット数の大きい順に並べた待ちユニット配列の適切な箇所に挿入する(ステップS702)。マクロ性能融通用の待ちユニット配列は、例えばQoS制御域213内に作成され、LUを特定するLUNと当該LUにおける最小待ちユニット数 N_F とを対応付けて保存している。なお、本実施形態において、マクロ性能融通での「1サイクルでの融通LU数」は、例えば、4LUとしているので、待ちユニット配列における要素数は4で十分である。

40

【0171】

ステップS702の処理後、もしくは、当該LUの最小待ちユニット数 N_F が0又は当

50

該LUの融通フラグ F_{INMAX} がオンである場合（ステップS701のNORルート）、第1判定部521は、全てのLUについて最小待ちユニット数 N_F の判定を行なったか否かを判定する（ステップS703）。全てのLUについて待ちユニット数の判定を行っていない場合（ステップS703のNORルート）、第1判定部521は、ステップS701の処理に戻る。全てのLUについて待ちユニット数の判定を行なった場合（ステップS703のYESルート）、マクロ融通処理部520は、マクロ性能融通用の待ちユニット配列に待ちユニットが存在するか否かを判定する（ステップS704）。マクロ性能融通用の待ちユニット配列に待ちユニットが存在しない場合（ステップS704のNORルート）、マクロ融通処理部520は、マクロ融通処理を完了する。

【0172】

マクロ性能融通用待ちユニット配列に待ちユニットが存在する場合（ステップS704のYESルート）、第2判定部522は、各LUについて、当該LUの最小余りユニット数 N_E が0以外の値で、且つ、当該LUの融通フラグ F_{OUTMAX} がオフであるかを判定する（ステップS705）。ここで、当該LUの最小余りユニット数 N_E が0以外の値である場合、当該LUは、1マクロサイクルにおける一秒刻みのタイミングでは、常に余り状態である。したがって、当該LUは、1マクロサイクルに亘って処理性能に余裕のある状態であるとみなせ、本実施形態では、上記「融通する条件」を満たしているものと判断される。ただし、当該LUの融通フラグ F_{OUTMAX} がオンである場合、当該LUにおいては、他のLUに対するマクロ性能融通の融通量が下限に到達しており、マクロ性能融通を行なうことができない。そこで、本実施形態では、各LUについて、当該LUの最小余りユニット数 N_E が0以外の値で、且つ、当該LUの融通フラグ F_{OUTMAX} がオフであるかが判定されている。

【0173】

当該LUの最小余りユニット数 N_E が0以外の値で、且つ、当該LUの融通フラグ F_{OUTMAX} がオフである場合（ステップS705のYESルート）、第2判定部522は、今回の対象LUの最小余りユニット数 N_E を、余りユニット数の大きい順に並べた余りユニット配列の適切な箇所に挿入する（ステップS706）。マクロ性能融通用の余りユニット配列は、例えばQoS制御域213内に作成され、LUを特定するLUNと当該LUにおける最小余りユニット数 N_E とを対応付けて保存している。なお、本実施形態において、マクロ性能融通での「1サイクルでの融通LU数」は、例えば、4LUとしているので、余りユニット配列における要素数は4で十分である。

【0174】

ステップS706の処理後、もしくは、当該LUの最小余りユニット数 N_E が0又は当該LUの融通フラグ F_{OUTMAX} がオンである場合（ステップS705のNORルート）、第2判定部522は、全てのLUについて最小余りユニット数 N_E の判定を行なったか否かを判定する（ステップS707）。全てのLUについて待ちユニット数の判定を行っていない場合（ステップS707のNORルート）、第2判定部522は、ステップS705の処理に戻る。

【0175】

全てのLUについて待ちユニット数の判定を行なった場合（ステップS707のYESルート）、マクロ融通処理部520は、待ちユニット配列の要素[i]が0であるか、あるいは、余りユニット配列の要素[i]が0であるかを判定する（ステップS708）。待ちユニット配列の要素[i]が0、または、余りユニット配列の要素[i]が0である場合（ステップS708のYESルート）、マクロ融通処理部520は、マクロ融通処理を完了する。

【0176】

待ちユニット配列の要素[i]も余りユニット配列の要素[i]も0でない場合（ステップS708のNORルート）、マクロ融通処理部520（変更部523）は、以下の処理（ステップS709～S713）を実行する。ここでは、待ちユニット配列のi番目の要素[i]に係る融通先LUと、余りユニット配列のi番目の要素[i]に係る融通元LUとを、マク

10

20

30

40

50

ロ性能融通を行なう一つの組合せとする。本実施形態において、各配列の要素数は4としているので、 i は、例えば0, 1, 2, 3とする。また、ステップS708～S713の処理は、各配列の待ちユニット数または余りユニット数の大きい順に実行される。

【0177】

なお、ステップS709～S713の処理のうち、ステップS709の処理は、融通先LUのQoS制御域213を対象としている。また、ステップS709～S713の処理のうち、ステップS710～S712の処理は、融通元LUのQoS制御域213を対象としている。

【0178】

変更部523は、待ちユニット配列の i 番目の融通先LUに係る、QoS制御域213における性能融通情報を、以下のように設定・変更する(ステップS709)。つまり、変更部523は、 i 番目の融通先LUの融通フラグ F_{IN} をオンに設定する。また、変更部523は、 i 番目の融通先LUの融通ユニット数 N_B を1インクリメントする。これにより、融通先LUについて算出される性能制限値に、1ユニット分に対応する値(例えば5MB/s, 60IOPS)が加算される(図7のステップS123参照)。さらに、変更部523は、融通ユニット数 N_B を1インクリメントすることで他のLUからの融通量が上限に到達した場合、融通フラグ F_{INMAX} をオンに設定する。

【0179】

そして、変更部523は、余りユニット配列の i 番目の融通元LUの融通フラグ F_{IN} がオンであるか否かを判定する(ステップS710)。融通フラグ F_{IN} がオンである場合(ステップS710のYESルート)、 i 番目の融通元LUは、処理性能に余裕があるにもかかわらず他のLUから性能融通を受けているものと考えられる。そこで、変更部523は、余りユニット配列の i 番目の融通元LUに係る、QoS制御域213における性能融通情報を、以下のように設定・変更する(ステップS711)。

【0180】

つまり、変更部523は、 i 番目の融通元LUの融通ユニット数 N_B を1デクリメントする。これにより、融通元LUについて算出される性能制限値から、1ユニット分に対応する値(例えば5MB/s, 60IOPS)が減算されることになる(図7のステップS123参照)。また、 i 番目の融通元LUの融通フラグ F_{INMAX} がオンである場合、即ち他のLUからの融通量が上限に到達していた場合、変更部523は、 i 番目の融通元LUの融通フラグ F_{INMAX} をオフに設定する。さらに、融通ユニット数 N_B を1デクリメントすることで融通ユニット数 N_B が0になった場合、 i 番目の融通元LUが他のLUから受ける融通量は0となるので、変更部523は、 i 番目の融通元LUの融通フラグ F_{IN} をオフに設定する。

【0181】

i 番目の融通元LUの融通フラグ F_{IN} がオフである場合(ステップS710のNOルート)、変更部523は、 i 番目の融通元LUに係る、QoS制御域213における性能融通情報を、以下のように設定・変更する(ステップS712)。つまり、変更部523は、 i 番目の融通元LUの融通フラグ F_{OUT} をオンに設定する。また、変更部523は、 i 番目の融通元LUの融通ユニット数 N_B を1インクリメントする。これにより、融通元LUについて算出される性能制限値から、1ユニット分に対応する値(例えば5MB/s, 60IOPS)が減算される(図7のステップS125参照)。さらに、変更部523は、融通ユニット数 N_B を1インクリメントすることで、他のLUへの融通量が上限に到達した場合、融通フラグ F_{OUTMAX} をオンに設定する。

【0182】

ステップS711またはS712の処理後、マクロ融通処理部520は、マクロ性能融通の待ちユニット配列に、次の要素 $[i]$ があるか否かを判定する(ステップS713)。次の要素 $[i]$ がある場合(ステップS713のYESルート)、マクロ融通処理部520は、ステップS708の処理に戻る。次の要素 $[i]$ がない場合(ステップS713のNOルート)、マクロ融通処理部520は、マクロ融通処理を完了する。

【 0 1 8 3 】

〔 2 - 8 〕 マクロ融通復旧処理

次に、図 1 6 および図 1 7 に示すフローチャート（ステップ S 8 0 1 ~ S 8 1 8 ）に従って、本実施形態の復旧処理部 5 3 2 によって実行されるマクロ融通復旧処理（図 6 のステップ S 1 0 7 ）の処理手順について説明する。マクロ融通復旧処理は、図 6 に示す N ミリ秒タイマスケジュール処理が N ミリ秒毎に実行される中で、他の L U へ性能を融通している L U で性能不足を検出した場合（図 6 のステップ S 1 0 6 の Y E S ルート）、当該 L U の性能を元の設定性能に戻すために実行される。つまり、マクロ融通復旧処理では、融通復旧 L U が他の L U に融通した処理量が、融通先 L U から融通復旧 L U に復帰される。その際、復旧処理部 5 3 2 は、融通先 L U に対し融通された処理量の残量が均等もしくは

10

【 0 1 8 4 】

復旧処理部 5 3 2 は、まず、融通復旧 L U が存在するか否かを確認する（ステップ S 8 0 1 ）。融通復旧 L U が存在しない場合（ステップ S 8 0 1 の N O ルート）、復旧処理部 5 3 2 は、マクロ融通復旧処理を完了する。

【 0 1 8 5 】

融通復旧 L U が存在する場合（ステップ S 8 0 1 の Y E S ルート）、復旧処理部 5 3 2 は、各 L U について、以下の処理（ステップ S 8 0 2 ~ S 8 0 5 ）を実行する。復旧処理部 5 3 2 は、対象 L U の融通フラグ F_{IN} がオンであるか否か、つまり対象 L U が融通先 L U であるか否かを判定する（ステップ S 8 0 2 ）。

20

【 0 1 8 6 】

対象 L U が融通先 L U である場合（ステップ S 8 0 2 の Y E S ルート）、復旧処理部 5 3 2 は、今回の対象 L U の融通ユニット数 N_B を、融通ユニット数 N_B の大きい順に並べた融通先ユニット配列の適切な箇所に挿入する（ステップ S 8 0 3 ）。併せて、復旧処理部 5 3 2 は、融通先 L U を特定する L U N と、融通先 L U に係るホストを特定する番号とを対応付ける配列も作成する（ステップ S 8 0 4 ）。L U / ホスト番号配列を含む融通先ユニット配列は、例えば Q o S 制御域 2 1 3 内に作成され、融通先 L U を特定する L U N と当該 L U における融通ユニット数 N_B とを対応付けて保存している。なお、融通先ユニット配列（L U N / ホスト番号配列）の一例については、図 1 8 を参照しながら後述する。また、融通先ユニット配列における要素数は、例えば 1 5 である。対象が「ポートと L U N との組合せ」である場合、復旧処理部 5 3 2 は、ステップ S 8 0 4 において、融通先 L U を特定する L U N と、融通先 L U に係るポートを特定する情報とを対応付ける配列も作成する。

30

【 0 1 8 7 】

ステップ S 8 0 4 の処理後、もしくは、対象 L U が融通先 L U でない場合（ステップ S 8 0 2 の N O ルート）、復旧処理部 5 3 2 は、全ての L U について融通フラグ F_{IN} の判定を行なったか否かを判定する（ステップ S 8 0 5 ）。全ての L U について融通フラグ F_{IN} の判定を行っていない場合（ステップ S 8 0 5 の N O ルート）、復旧処理部 5 3 2 は、ステップ S 8 0 2 の処理に戻る。全ての L U について融通フラグ F_{IN} の判定を行なった場合（ステップ S 8 0 5 の Y E S ルート）、復旧処理部 5 3 2 は、ステップ S 8 0 3 , S 8 0 4 で作成された融通先ユニット配列を参照しながら、以下の処理（ステップ S 8 0 6 ~ S 8 1 8 ）を実行する。これにより、復旧処理部 5 3 2 は、融通先 L U に対し融通された処理量の残量が均等もしくはほぼ均等になるように、当該処理量を融通先 L U から融通復旧 L U に復帰させる。

40

【 0 1 8 8 】

復旧処理部 5 3 2 による以下の処理を実行するに際し、融通復旧 L U に復旧させるべきユニット数（必要ユニット数）を A、復旧残りユニット数を R とする。復旧残りユニット数 R の初期値は、必要ユニット数 A である。また、融通先ユニット配列数（融通先ユニット配列の要素数）を M、配列内ユニット数を $X_0, X_1, X_2, \dots, X_{M-1}$ （ただし $X_0, X_1, X_2, \dots, X_{M-1}$ ）、配列内 L U を $L_0, L_1, L_2, \dots, L_{M-1}$ とする（ステップ S 8 0 6

50

）。

【 0 1 8 9 】

そして、復旧処理部 5 3 2 は、 $i = 0$ としてから（ステップ S 8 0 7）、復旧残りユニット数 R が $(X_i - X_{i+1}) \times (i + 1)$ 以下であるか否かを判定する（ステップ S 8 0 8）。復旧残りユニット数 R が $(X_i - X_{i+1}) \times (i + 1)$ よりも大きい場合（ステップ S 8 0 8 の N O ルート）、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、配列内 L U の L_0 から L_i までの融通ユニット数 N_B から、 $(X_i - X_{i+1})$ を減算する。このとき、 $L_0 \sim L_i$ のそれぞれの融通フラグ F_{INMAX} がオンであれば、復旧処理部 5 3 2 は、融通フラグ F_{INMAX} をオフに設定する。そして、復旧処理部 5 3 2 は、復旧残りユニット数 R を $R - (X_i - X_{i+1}) \times (i + 1)$ に置き換える（ステップ S 8 0 9）。この後、復旧処理部 5 3 2 は、 i を $i + 1$ にインクリメントしてから（ステップ S 8 1 0）、 i が $M - 1$ と一致するか否かを判定する（ステップ S 8 1 1）。 i が $M - 1$ と一致しない場合（ステップ S 8 1 1 の N O ルート）、復旧処理部 5 3 2 は、ステップ S 8 0 8 の処理に戻る。

10

【 0 1 9 0 】

一方、復旧残りユニット数 R が $(X_i - X_{i+1}) \times (i + 1)$ 以下である場合（ステップ S 8 0 8 の Y E S ルート）、復旧処理部 5 3 2 は、以下の処理（ステップ S 8 1 2）を実行する。つまり、復旧処理部 5 3 2 は、復旧残りユニット数 R を $(i + 1)$ 個以内の整数値に分割する。そして、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、配列内 L U の L_0 から L_i までの融通ユニット数 N_B から、分割して得られた整数値を減算する。このとき、 $L_0 \sim L_i$ のそれぞれの融通フラグ F_{INMAX} がオンであれば、復旧処理部 5 3 2 は、融通フラグ F_{INMAX} をオフに設定する。そして、復旧処理部 5 3 2 は、融通ユニット数 N_B の減算処理に伴い融通ユニット数 N_B が 0 になった場合、融通フラグ F_{IN} をオフに設定する。そして、復旧処理部 5 3 2 は、復旧残りユニット数 R を 0 に置き換える。

20

【 0 1 9 1 】

ステップ S 8 1 2 の処理後、もしくは、 i が $M - 1$ と一致した場合（ステップ S 8 1 1 の Y E S ルート）、復旧処理部 5 3 2 は、復旧残りユニット数 R が 0 と一致しているか否かを判定する（ステップ S 8 1 3）。復旧残りユニット数 R が 0 でない場合（ステップ S 8 1 3 の N O ルート）、復旧処理部 5 3 2 は、融通フラグ F_{IN} がオンの全 L Uの中から、 LUN の最も若い L U を選択する（ステップ S 8 1 4）。この後、復旧処理部 5 3 2 は、復旧残りユニット数 R が、選択した L U の融通ユニット数 N_B 以下であるか否かを判定する（ステップ S 8 1 5）。

30

【 0 1 9 2 】

復旧残りユニット数 R が融通ユニット数 N_B よりも大きい場合（ステップ S 8 1 5 の N O ルート）、復旧処理部 5 3 2 は、復旧残りユニット数 R を、 R から融通ユニット数 N_B を減算した値に置き換える。また、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、選択した L U の融通ユニット数 N_B を 0 に設定する。さらに、選択した L U の融通フラグ F_{INMAX} がオンであれば、復旧処理部 5 3 2 は、融通フラグ F_{INMAX} をオフに設定する。そして、復旧処理部 5 3 2 は、融通フラグ F_{IN} をオフに設定する（ステップ S 8 1 6）。

【 0 1 9 3 】

ステップ S 8 1 6 の処理後、復旧処理部 5 3 2 は、ステップ S 8 1 4 に戻り、融通フラグ F_{IN} がオンの全 L Uの中から、 LUN が次に若い L U を選択し、同様の処理（ステップ S 8 1 5、S 8 1 6）を実行する。

40

【 0 1 9 4 】

復旧残りユニット数 R が融通ユニット数 N_B 以下である場合（ステップ S 8 1 5 の Y E S ルート）、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、選択した L U の融通ユニット数 N_B から、復旧残りユニット数 R を減算する。また、選択した L U の融通フラグ F_{INMAX} がオンであれば、復旧処理部 5 3 2 は、融通フラグ F_{INMAX} をオフに設定する。さらに、復旧処理部 5 3 2 は、融通ユニット数 N_B の減算処理に伴い融通ユニット数 N_B が 0 になった場合、融通フラグ F_{IN} をオフに設定する。そして、復旧処理部 5 3 2 は、復旧残りユニット数 R を 0 に置き換える（ステップ S 8 1 7）。

50

【 0 1 9 5 】

ステップ S 8 1 7 の処理後、または、復旧残りユニット数 R が 0 である場合（ステップ S 8 1 3 の Y E S ルート）、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、融通復旧 L U の融通フラグ F_{OUT} および F_{OUTMAX} をいずれもオフに設定する。また、復旧処理部 5 3 2 は、Q o S 制御域 2 1 3 において、融通復旧 L U の融通ユニット数 N_B を 0 に設定する（ステップ S 8 1 8）。この後、復旧処理部 5 3 2 は、マクロ融通復旧処理を完了する。

【 0 1 9 6 】

次に、図 1 7 のステップ S 8 0 7 ~ S 8 1 3 の処理について、図 1 8 ~ 図 2 0 に示す例を参照しながら、具体的に説明する。なお、図 1 8 は、本実施形態に係るマクロ融通復旧処理（図 1 6 のステップ S 8 0 3 , S 8 0 4）で作成される融通先ユニット配列（L U N / ホスト番号配列）の一例を示す図である。また、図 1 9 および図 2 0 は、本実施形態に係るマクロ融通復旧処理（特に図 1 7 のステップ S 8 0 7 ~ S 8 1 3）を具体的に説明する図である。

【 0 1 9 7 】

図 1 8 に示す融通先ユニット配列では、「ホストと L U N との組合せ」を特定するホスト番号および L U N と、当該 L U N に対応する融通先 L U の融通ユニット数 N_B とが、対応付けられ、融通ユニット数 N_B の大きい順にソートされている。図 1 8 および図 1 9 に示す例において、配列内 L U の L_0 は、ホスト # 1 および L U N # 1 に対応し、 L_0 の融通ユニット数 $N_B (= X_0)$ は、5 0 である。また、配列内 L U の L_1 は、ホスト # 2 および L U N # 3 に対応し、 L_1 の融通ユニット数 $N_B (= X_1)$ は、2 5 であり、配列内 L U の L_2 は、ホスト # 1 および L U N # 4 に対応し、 L_2 の融通ユニット数 $N_B (= X_2)$ は、1 5 である。さらに、配列内 L U の L_3 は、ホスト # 2 および L U N # 2 に対応し、 L_3 の融通ユニット数 $N_B (= X_3)$ は、1 0 であり、配列内 L U の L_4 は、ホスト # 2 および L U N # 5 に対応し、 L_4 の融通ユニット数 $N_B (= X_4)$ は、8 である。なお、融通先ユニット配列数（融通先ユニット配列の要素数）M は 1 0 とする。

【 0 1 9 8 】

以下では、図 1 8 および図 1 9 に示す融通先ユニット配列に基づき、復旧処理部 5 3 2 が、必要ユニット数 $A = 64$ のユニットを融通復旧 L U に復旧させる場合の処理手順について具体的に説明する。このとき、復旧残りユニット数 R の初期値は、必要ユニット数 A の 6 4 となる。復旧処理部 5 3 2 は、 $i = 0$ を設定してから（ステップ S 8 0 7）、まず、 $i = 0$ のときの値 $(X_i - X_{i+1}) \times (i + 1) = (X_0 - X_1) \times (0 + 1) = 25 \times 1 = 25$ を算出する。R の値は 6 4 であるので、 $R > 25$ であり（ステップ S 8 0 8 の N O ルート）、 L_0 から融通復旧ユニットに 2 5 ユニット分の性能を復旧させるべく、復旧処理部 5 3 2 は、 L_0 の融通ユニット数 N_B から 2 5 を減算する。そして、復旧処理部 5 3 2 は、復旧残りユニット数 R を、6 4 から、 $64 - 25 = 39$ に置き換える（ステップ S 8 0 9；図 2 0 の最上段の式参照）。

【 0 1 9 9 】

ついで、復旧処理部 5 3 2 は、 i を 1 インクリメントして $i = 1$ にする（ステップ S 8 1 0）。このとき、 $i = 1$ $M - 1 = 9$ であるため（ステップ S 8 1 1 の N O ルート）、復旧処理部 5 3 2 は、ステップ S 8 0 8 に戻り、 $i = 1$ のときの値 $(X_i - X_{i+1}) \times (i + 1) = (X_1 - X_2) \times (1 + 1) = 10 \times 2 = 20$ を算出する。R の値は 3 9 であるので、 $R > 20$ であり（ステップ S 8 0 8 の N O ルート）、 L_0 および L_1 から融通復旧ユニットに 2 0 ユニット分の性能を復旧させるべく、復旧処理部 5 3 2 は、 L_0 および L_1 の融通ユニット数 N_B からそれぞれ 1 0 を減算する。そして、復旧処理部 5 3 2 は、復旧残りユニット数 R を、3 9 から、 $39 - 20 = 19$ に置き換える（ステップ S 8 0 9；図 2 0 の上から 2 段目の式参照）。

【 0 2 0 0 】

ついで、復旧処理部 5 3 2 は、 i を 1 インクリメントして $i = 2$ にする（ステップ S 8 1 0）。このとき、 $i = 2$ $M - 1 = 9$ であるため（ステップ S 8 1 1 の N O ルート）、

復旧処理部 532 は、ステップ S808 に戻り、 $i = 2$ のときの値 $(X_i - X_{i+1}) \times (i + 1) = (X_2 - X_3) \times (2 + 1) = 5 \times 3 = 15$ を算出する。R の値は 19 であるので、 $R > 15$ であり（ステップ S808 の NO ルート）、 L_0, L_1, L_2 から融通復旧ユニットに 15 ユニット分の性能を復旧させるべく、復旧処理部 532 は、 L_0, L_1, L_2 の融通ユニット数 N_B からそれぞれ 5 を減算する。そして、復旧処理部 532 は、復旧残りユニット数 R を、19 から、 $19 - 15 = 4$ に置き換える（ステップ S809；図 20 の上から 3 段目の式参照）。

【0201】

ついで、復旧処理部 532 は、 i を 1 インクリメントして $i = 3$ にする（ステップ S810）。このとき、 $i = 3$ 、 $M - 1 = 9$ であるため（ステップ S811 の NO ルート）、復旧処理部 532 は、ステップ S808 に戻り、 $i = 3$ のときの値 $(X_i - X_{i+1}) \times (i + 1) = (X_3 - X_4) \times (3 + 1) = 2 \times 4 = 8$ を算出する。R の値は 4 であるので、R 8 であり（ステップ S808 の YES ルート；図 20 の上から 4 段目の式参照）、復旧処理部 532 は、復旧残りユニット数 $R = 4$ を $(i + 1) = 4$ 個以内の整数値に分割する。ここでは、 $R = 4$ は、4 つの LU である $L_0 \sim L_3$ に分割され、復旧処理部 532 は、4 つの $L_0 \sim L_3$ の融通ユニット数 N_B から、それぞれ 1 を減算し、復旧残りユニット数 R を 4 から 0 に置き換える（ステップ S812）。このとき $R = 0$ であるため（ステップ S813 の YES ルート）、復旧処理部 532 は、ステップ S818 の処理に移行する。

【0202】

以上の処理によって、融通復旧 LU が他の LU に融通した性能（処理量）が、融通先 LU から融通復旧 LU に復帰される。その際、融通先 LU に対し融通された性能（処理量）の残量が均等もしくはほぼ均等になるように、当該性能を融通先 LU から融通復旧 LU に復帰させることができる。

【0203】

〔3〕本実施形態の効果

本実施形態のストレージ装置 100 によれば、使用されるホストと LU との組合せの性能レベル（優先度）に従って、ストレージ装置 100 の性能を配分し、ストレージの QoS 機能を容易に実現することができる。

【0204】

また、本実施形態のストレージ装置 100 によれば、ホスト 10 からの入出力要求（I/O 要求、コマンド）の状況やホスト 10 の接続環境の変化に応じて、ストレージ性能を最適に制御することができる。特に、本実施形態のマクロ性能融通では、上記優先度についての性能値の最小間隔 5 MB/s、60 IOPS を「1 ユニット」とする単位で、性能融通が行なわれる。したがって、優先度を変更して性能限界値を変更する場合に比べ、性能限界値の変更間隔（融通量）を大幅に小さくすることができ、ホスト 10 から要求される処理性能の変動に細かく追従することができ、様々な性能変動に対応することができる。その際、複数ボリューム（LU）間で、処理性能に余裕のある LU から、性能不足の LU に対し、性能が融通されるため、性能融通を確実にこなうことができる。

【0205】

さらに、本実施形態のストレージ装置 100 では、ミクロ性能融通およびマクロ性能融通の特徴を踏まえ、ミクロ性能融通およびマクロ性能融通のいずれか一方が使用される。ミクロ性能融通およびマクロ性能融通の一方を使用する場合、ストレージ装置 100 のストレージ QoS を使用する動作環境に応じて、適切な性能融通を選択することができる。また、両方を使用する場合、一時的なスパイク性能などを吸収する際にはミクロ性能融通を使用する一方、継続的な性能変動を吸収する際にはマクロ性能融通を使用するように制御を行なうことで、両方の特徴を有効に利用することができる。

【0206】

また、本実施形態のミクロ性能融通では、例えば 1 秒以下ミクロサイクルで、処理性能が不足する LU に、処理性能に余裕のある LU から処理性能が融通される。このときの融通量は、記憶されず、次のミクロサイクルでは性能限界値の調整に反映されない。従って

10

20

30

40

50

、ストレージ装置 100 は、次のマイクロサイクルでは、元の処理性能で動作する。このように、マイクロ性能融通は一時的な融通であるため、マイクロ性能融通を用いることで一度に大量の性能融通が可能である。

【0207】

さらに、本実施形態のマクロ性能融通では、例えば 2 秒以上の、より大きなマクロサイクルで処理性能が監視される。そして、所定時間（例えば 1 マクロサイクル）に亘り連続して性能不足となっている LU に対し、前記所定時間に亘り連続して処理性能に余裕のある LU から、処理性能が融通される。このように、マクロ性能融通は、性能融通量を学習により調整するため、ホスト 10 から要求される処理性能の継続的な変動に有効に対応することができる。

10

【0208】

またさらに、本実施形態のストレージ装置 100 では、融通元 LU で処理性能不足が発生した場合、つまり融通復旧 LU が存在する場合、融通復旧 LU が他の LU に融通した性能（処理量）が、融通先 LU から融通復旧 LU に復帰される。その際、融通先 LU に対し融通された性能（処理量）の残量が均等もしくはほぼ均等になるように、当該性能を融通先 LU から融通復旧 LU に復帰させることができる。つまり、融通復旧 LU の性能復旧後の複数の融通先 LU において融通量を均等もしくはほぼ均等に残すことができ、複数の融通先 LU において融通量の偏りがなくなる。したがって、当該複数の融通先 LU を含むストレージ装置 100 において、以降の処理を効率よく行なうことができる。

【0209】

20

〔4〕その他

以上、本発明の好ましい実施形態について詳述したが、本発明は、係る特定の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲内において、種々の変形、変更して実施することができる。

【0210】

また、図示した各構成部は、機能概念的なものであり、必ずしも物理的に図示のごとく構成されていることを要しない。例えば、SCSI ターゲット制御部 303 において、コマンド受信部 504 とスケジュール部 506 とは統合されてもよい。さらに、各装置にて行なわれる各処理機能は、その全部または任意の一部が、CPU および当該 CPU にて解析実行されるプログラムにて実現され、あるいは、ワイヤードロジックによるハードウェアとして実現され得る。

30

【0211】

また、本実施形態において説明した各種処理は、予め用意されたプログラムをストレージ装置が有するコンピュータで実行することによって実現することができる。そこで、以下では、図 21 を参照しながら、本実施形態と同様の機能を有するストレージ装置制御プログラムを実行するコンピュータの一例を説明する。図 21 は、本実施形態のストレージ装置制御プログラムを実行するコンピュータを示すブロック図である。

【0212】

図 21 に示すように、コンピュータ 1000 は、ホストとのインターフェースである FC - CA 1010 と、ホストとのインターフェースである iSCSI - CA 1020 と、ディスク装置とのインターフェースである SAS 1030 とを有する。また、コンピュータ 1000 は、各種情報を一時記憶する RAM 1040 と、書き換え可能であり、電源を切ってもデータが消えない不揮発性のフラッシュメモリ 1050 とを有する。さらに、コンピュータ 1000 は、読み出し専用メモリである ROM (Read Only Memory) 1060 と、各種演算処理を実行する CPU 1070 とを有する。また、コンピュータ 1000 が有する各部は、バス 1080 を介して接続される。

40

【0213】

そして、フラッシュメモリ 1050 には、図 3 および図 5 に示した QoS 管理テーブル 501 に対応する QoS 管理テーブル 1051 や、図 3 に示した各 LU の性能融通情報などが記憶される。また、ROM 1060 には、図 3 に示した算出部 505、スケジュール

50

部 5 0 6 , 実行部 5 0 7 , マクロ融通チェック部 5 1 0 , マクロ融通処理部 5 2 0 , マクロ融通復旧処理部 5 3 0 , ミクロ融通処理部 5 4 0 と同様の機能を有するストレージ装置制御プログラム 1 0 6 1 が記憶される。

【 0 2 1 4 】

そして、CPU 1 0 7 0 は、フラッシュメモリ 1 0 5 0 から読み出した Q o S 管理テーブル 1 0 5 1 や各 L U の性能融通情報を参照し、ROM 1 0 6 0 から読み出したストレージ装置制御プログラム 1 0 6 1 をストレージ装置制御プロセス 1 0 7 1 として実行する。

【 0 2 1 5 】

なお、上記ストレージ装置制御プログラム 1 0 6 1 は、必ずしも ROM 1 0 6 0 に格納されている必要はない。例えば、当該プログラム 1 0 6 1 は、フレキシブルディスク、C 10 D (C D - R O M , C D - R , C D - R W など) , D V D (D V D - R O M , D V D - R A M , D V D - R , D V D - R W , D V D + R , D V D + R W など) , ブルーレイディスク等のコンピュータ読取可能な記録媒体に記録された形態で提供されてもよい。この場合、コンピュータ 1 0 0 0 はその記録媒体からプログラム 1 0 6 1 を読み取って内部記憶装置または外部記憶装置に転送し格納して用いる。また、公衆回線、インターネット、L A N (Local Area Network) , W A N (Wide Area Network) 等を介してコンピュータ 1 0 0 0 に接続される他のコンピュータ (またはサーバ) 等にこのプログラム 1 0 6 1 を記憶させておき、コンピュータ 1 0 0 0 がこれらからプログラム 1 0 6 1 を読み出して実行するようにしてもよい。

【 0 2 1 6 】

〔 5 〕 付記

以上の各実施例を含む実施形態に関し、さらに以下の付記を開示する。

(付記 1)

複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出する算出部と、

情報処理装置から受信した入出力要求に係る処理量と、前記算出部により算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールするスケジュール部と、

前記スケジュール部によりスケジュールされた前記実行順で前記入出力要求に係る処理を実行する実行部と、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第 1 所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定する第 1 判定部と、

当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると前記第 1 判定部により判定された場合、当該記憶装置についての前記上限値を、所定の性能融通単位で変更する変更部と、

を有する、ストレージ装置。

【 0 2 1 7 】

(付記 2)

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第 2 所定時間に亘って当該記憶装置の処理性能を下回っているか否かを判定する第 2 判定部を、さらに有し

、
前記複数の記憶装置のうち第 1 記憶装置に対する入出力要求に係る処理量が前記第 1 所定時間に亘って前記第 1 記憶装置の処理性能を超えていると前記第 1 判定部により判定され、且つ、前記複数の記憶装置のうち第 2 記憶装置に対する入出力要求に係る処理量が前記第 2 所定時間に亘って前記第 2 記憶装置の処理性能を下回っていると前記第 2 判定部により判定された場合、前記変更部は、前記上限値を前記所定の性能融通単位で変更することで、前記第 2 記憶装置から前記第 1 記憶装置への性能融通を前記所定の性能融通単位で行なう、

付記 1 に記載のストレージ装置。

【 0 2 1 8 】

(付記 3)

前記変更部は、前記第 1 記憶装置について算出された前記上限値に、前記所定の性能融通単位に対応する値を加算するとともに、前記第 2 記憶装置について算出された前記上限値から、前記所定の性能融通単位に対応する値を減算するように、前記上限値を変更することで、前記第 2 記憶装置から前記第 1 記憶装置への性能融通を前記所定の性能融通単位で行なう、

付記 2 に記載のストレージ装置。

【0219】

(付記 4)

前記所定の性能融通単位に対応する値は、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である、

付記 3 に記載のストレージ装置。

【0220】

(付記 5)

前記記憶装置毎に、前記第 1 所定時間よりも短い時間間隔で、当該記憶装置に対する入出力要求の処理待ち量を取得し、前記第 1 所定時間内に取得された前記処理待ち量のうち最小のものを最小待ち量として取得する第 1 取得部を、さらに有し、

前記第 1 判定部は、前記第 1 記憶装置について前記第 1 取得部により取得された前記最小待ち量が 0 でない場合、前記第 1 記憶装置に対する入出力要求に係る処理量が前記第 1 所定時間に亘って前記第 1 記憶装置の処理性能を超えていると判定する、

付記 2 ～付記 4 のいずれか一項に記載のストレージ装置。

【0221】

(付記 6)

前記第 1 取得部は、前記最小待ち量として、前記最小待ち量を前記所定の性能融通単位に換算した最小待ちユニット数を取得する、

付記 5 に記載のストレージ装置。

【0222】

(付記 7)

前記記憶装置毎に、前記第 2 所定時間よりも短い時間間隔で、当該記憶装置に対する入出力要求に係る処理量の、当該記憶装置の処理性能に対する処理余り量を取得し、前記第 2 所定時間内に取得された前記処理余り量のうち最小のものを最小余り量として取得する第 2 取得部を、さらに有し、

前記第 2 判定部は、前記第 2 取得部により取得された前記最小余り量が 0 でない場合、前記第 2 記憶装置に対する入出力要求に係る処理量が前記第 2 所定時間に亘って前記第 2 記憶装置の処理性能を下回っていると判定する、

付記 2 ～付記 6 のいずれか一項に記載のストレージ装置。

【0223】

(付記 8)

前記第 2 取得部は、前記最小余り量として、前記最小余り量を前記所定の性能融通単位に換算した最小余りユニット数を取得する、

付記 7 に記載のストレージ装置。

【0224】

(付記 9)

前記記憶装置毎に、当該記憶装置が、一以上の前記第 1 記憶装置への前記性能融通を行なった前記第 2 記憶装置であり、且つ、当該記憶装置に対する入出力要求の処理待ちの生じた融通復旧対象であるか否かを判定する第 3 判定部と、

前記第 3 判定部により当該記憶装置が前記融通復旧対象であると判定された場合、前記融通復旧対象の記憶装置が他の記憶装置に融通した処理量を、前記一以上の前記第 1 記憶装置から前記融通復旧対象の記憶装置に復帰させる復旧処理部と、

をさらに有する、付記 2 ～付記 8 のいずれか一項に記載のストレージ装置。

10

20

30

40

50

【 0 2 2 5 】

(付 記 1 0)

前記復旧処理部は、前記一以上の前記第 1 記憶装置に対し融通された処理量の残量が均等もしくはほぼ均等になるように、当該処理量を前記一以上の前記第 1 記憶装置から前記融通復旧対象の記憶装置に復帰させる、

付記 9 に記載のストレージ装置。

【 0 2 2 6 】

(付 記 1 1)

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能の所定数倍以上であるか否かを判定する第 4 判定部と、

当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能の所定数倍以上であると前記第 4 判定部により判定された場合、当該記憶装置に対する入出力要求に係る処理量を超える処理余裕をもつ記憶装置から当該記憶装置への性能融通を行なう融通処理部と、

をさらに有する、付記 1 ～付記 1 0 のいずれか一項に記載のストレージ装置。

【 0 2 2 7 】

(付 記 1 2)

前記複数の記憶装置のそれぞれは、一以上の物理記憶装置を論理的に分割した複数の論理ユニットの一つである、

付記 1 ～付記 1 1 のいずれか一項に記載のストレージ装置。

【 0 2 2 8 】

(付 記 1 3)

ストレージ装置が、

複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出し、

情報処理装置から受信した入出力要求に係る処理量と、算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールし、

スケジュールされた前記実行順で前記入出力要求に係る処理を実行し、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第 1 所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定し、

当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると判定された場合、当該記憶装置についての前記上限値を、所定の性能融通単位で変更する、

ことを特徴とするストレージ装置の制御方法。

【 0 2 2 9 】

(付 記 1 4)

前記ストレージ装置が、さらに、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第 2 所定時間に亘って当該記憶装置の処理性能を下回っているか否かを判定し、

前記複数の記憶装置のうち第 1 記憶装置に対する入出力要求に係る処理量が前記第 1 所定時間に亘って前記第 1 記憶装置の処理性能を超えていると前記第 1 判定部により判定され、且つ、前記複数の記憶装置のうち第 2 記憶装置に対する入出力要求に係る処理量が前記第 2 所定時間に亘って前記第 2 記憶装置の処理性能を下回っていると前記第 2 判定部により判定された場合、前記上限値を前記所定の性能融通単位で変更することで、前記第 2 記憶装置から前記第 1 記憶装置への性能融通を前記所定の性能融通単位で行なう、

付記 1 3 に記載のストレージ装置の制御方法。

【 0 2 3 0 】

(付 記 1 5)

前記ストレージ装置が、さらに、

前記第 1 記憶装置について算出された前記上限値に、前記所定の性能融通単位に対応す

10

20

30

40

50

る値を加算するとともに、前記第2記憶装置について算出された前記上限値から、前記所定の性能融通単位に対応する値を減算するように、前記上限値を変更することで、前記第2記憶装置から前記第1記憶装置への性能融通を前記所定の性能融通単位で行なう、
付記14に記載のストレージ装置の制御方法。

【0231】

(付記16)

前記所定の性能融通単位に対応する値は、前記優先度に基づいて決定される前記入出力処理量(性能値)の最小間隔以下である、

付記15に記載のストレージ装置の制御方法。

【0232】

(付記17)

ストレージ装置に含まれるコンピュータに、

複数の記憶装置に設定された優先度に基づいて決定される入出力処理量の上限値を前記記憶装置毎に算出し、

情報処理装置から受信した入出力要求に係る処理量と、算出された前記上限値とに基づいて、前記入出力要求に係る処理の実行順をスケジュールし、

スケジュールされた前記実行順で前記入出力要求に係る処理を実行し、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第1所定時間に亘って当該記憶装置の処理性能を超えているか否かを判定し、

当該記憶装置に対する入出力要求に係る処理量が当該記憶装置の処理性能を超えていると判定された場合、当該記憶装置についての前記上限値を、所定の性能融通単位で変更する、

処理を実行させる、ストレージ装置制御プログラム。

【0233】

(付記18)

前記コンピュータに、さらに、

前記記憶装置毎に、当該記憶装置に対する入出力要求に係る処理量が第2所定時間に亘って当該記憶装置の処理性能を下回っているか否かを判定し、

前記複数の記憶装置のうち第1記憶装置に対する入出力要求に係る処理量が前記第1所定時間に亘って前記第1記憶装置の処理性能を超えていると前記第1判定部により判定され、且つ、前記複数の記憶装置のうち第2記憶装置に対する入出力要求に係る処理量が前記第2所定時間に亘って前記第2記憶装置の処理性能を下回っていると前記第2判定部により判定された場合、前記上限値を前記所定の性能融通単位で変更することで、前記第2記憶装置から前記第1記憶装置への性能融通を前記所定の性能融通単位で行なう、

処理を実行させる、付記17に記載のストレージ装置制御プログラム。

【0234】

(付記19)

前記コンピュータに、さらに、

前記第1記憶装置について算出された前記上限値に、前記所定の性能融通単位に対応する値を加算するとともに、前記第2記憶装置について算出された前記上限値から、前記所定の性能融通単位に対応する値を減算するように、前記上限値を変更することで、前記第2記憶装置から前記第1記憶装置への性能融通を前記所定の性能融通単位で行なう、

付記18に記載のストレージ装置制御プログラム。

【0235】

(付記20)

前記所定の性能融通単位に対応する値は、前記優先度に基づいて決定される前記入出力処理量の最小間隔以下である、

付記19に記載のストレージ装置制御プログラム。

【符号の説明】

【0236】

10

20

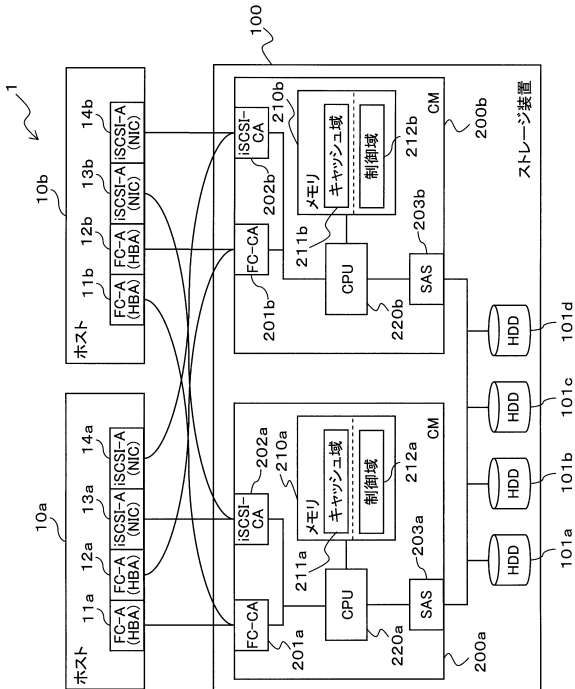
30

40

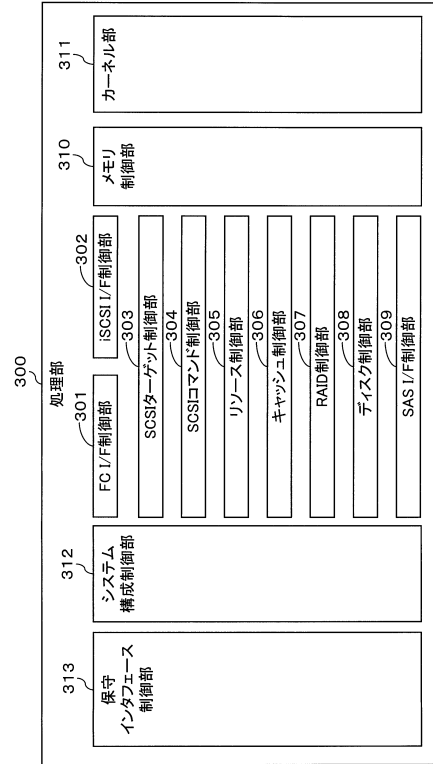
50

1	ストレージシステム	
1 0 a , 1 0 b	ホスト (情報処理装置)	
1 0 0	ストレージ装置	
1 0 1 a ~ 1 0 1 d	H D D (一以上の物理記憶装置)	
2 0 0 a , 2 0 0 b	C M	
2 0 1 a , 2 0 1 b , 1 0 1 0	F C - C A (ポート)	
2 0 2 a , 2 0 2 b , 1 0 2 0	i S C S I - C A (ポート)	
2 0 3 a , 2 0 3 b , 1 0 3 0	S A S	
2 1 0 a , 2 1 0 b	メモリ	
2 1 1 a , 2 1 1 b	キャッシュ域	10
2 1 2 a , 2 1 2 b	制御域	
2 1 3	Q o S 制御域	
2 2 0 a , 2 2 0 b , 1 0 7 0	C P U (コンピュータ)	
3 0 3	S C S I ターゲット制御部	
3 0 4	S C S I コマンド制御部	
5 0 1	Q o S 管理テーブル	
5 0 2	待ちキュー格納部	
5 0 3	性能設定部	
5 0 4	コマンド受信部	
5 0 5	算出部	20
5 0 6	スケジュール部	
5 0 7	実行部	
5 1 0	マクロ融通チェック部	
5 1 1	第 1 取得部	
5 1 2	第 2 取得部	
5 2 0	マクロ融通処理部	
5 2 1	第 1 判定部	
5 2 2	第 2 判定部	
5 2 3	変更部	
5 3 0	マクロ融通復旧処理部	30
5 3 1	第 3 判定部	
5 3 2	復旧処理部	
5 4 0	ミクロ融通処理部	
5 4 1	第 4 判定部	
5 4 2	融通処理部	
5 4 3	融通後処理部	
1 0 0 0	コンピュータ	
1 0 4 0	R A M	
1 0 5 0	フラッシュメモリ	
1 0 6 0	R O M	40
1 0 6 1	ストレージ装置制御プログラム	
1 0 7 1	ストレージ装置制御プロセス	

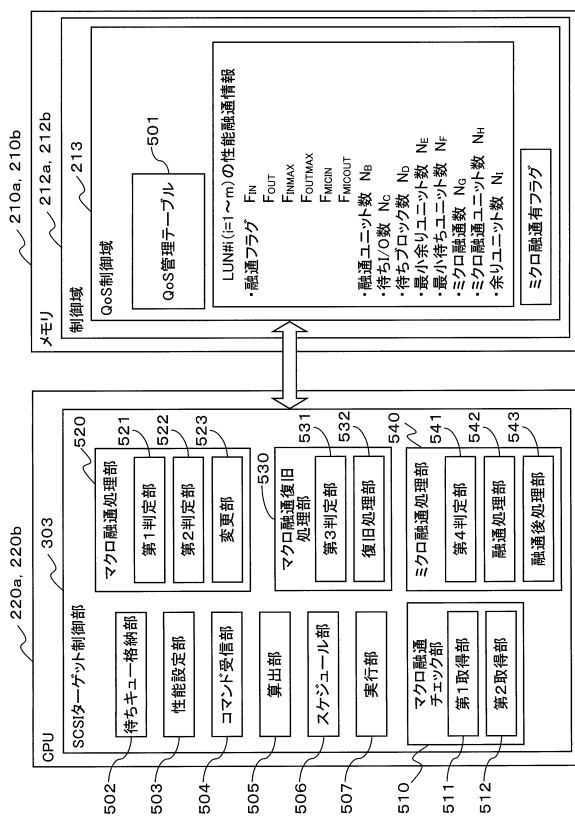
【 図 1 】



【 図 2 】



【 図 3 】



【 図 4 】

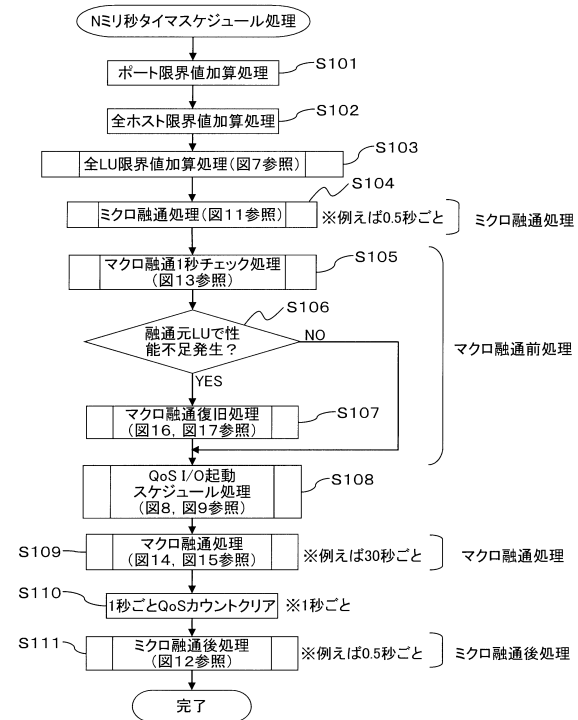
優先順位	性能値 (MB/s)	性能値 (IOPS)
1	800	15000
2	700	12500
3	600	10000
4	500	7500
5	400	5000
6	300	3000
7	200	1000
8	100	800
9	70	600
10	40	400
11	25	300
12	20	240
13	15	180
14	10	120
15	5	60

【図 5】

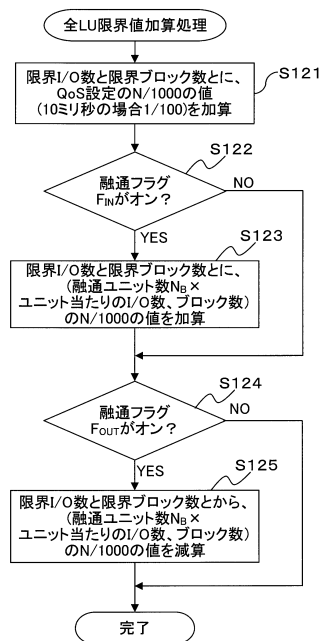
501

対象	優先順位	Nミリ秒I/O数	Nミリ秒ブロック数	限界I/O数	限界ブロック数	現I/O数	現ブロック数
FC-CA 201a	3	100	12288	$100 \times n-a$	$12288 \times n-A$	a	A
ホスト 10a	5	50	8192	$50 \times n-b$	$8192 \times n-B$	b	B
ホスト 10b	7	10	4096	$10 \times n-c$	$4096 \times n-C$	c	C
ホスト 10a-LUN-A	10	4	819.2	$4 \times n-d$	$819.2 \times n-D$	d	D
ホスト 10a-LUN-B	13	1.8	307.2	$1.8 \times n-e$	$307.2 \times n-E$	e	E
ホスト 10b-LUN-C	15	0.6	102.4	$0.6 \times n-f$	$102.4 \times n-F$	f	F

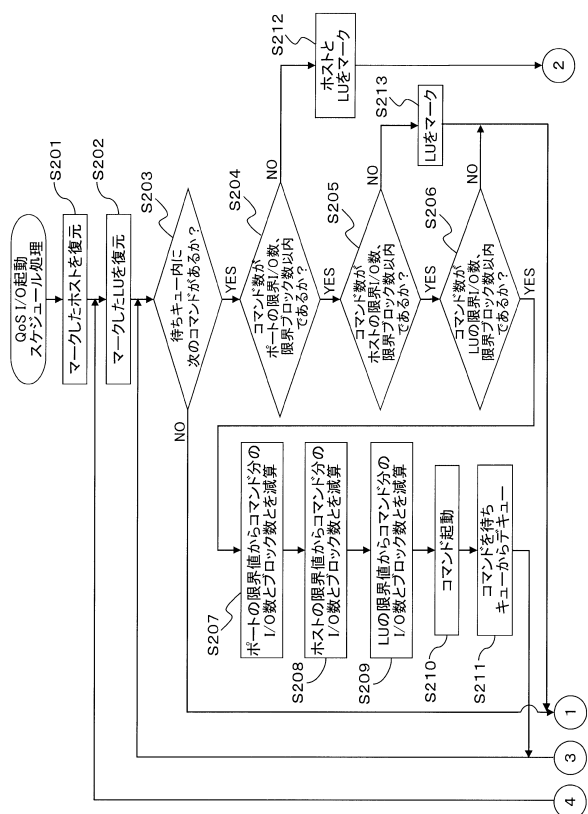
【図 6】



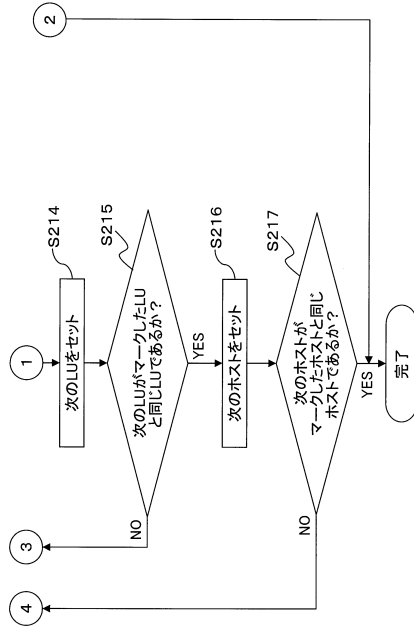
【図 7】



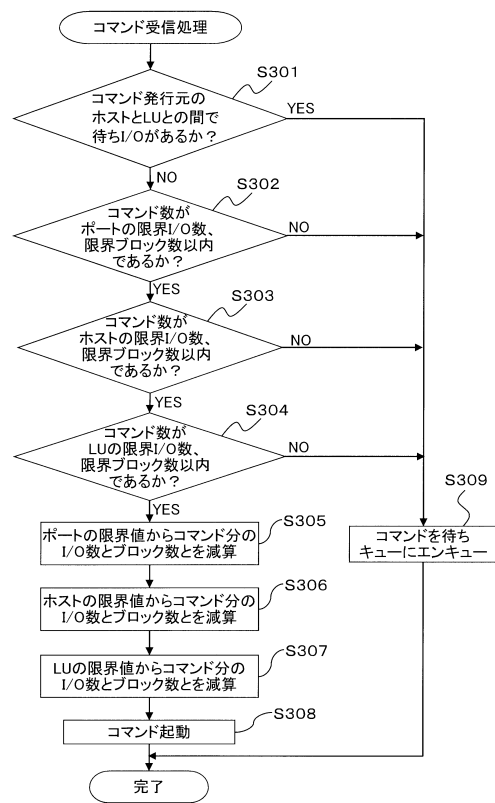
【図 8】



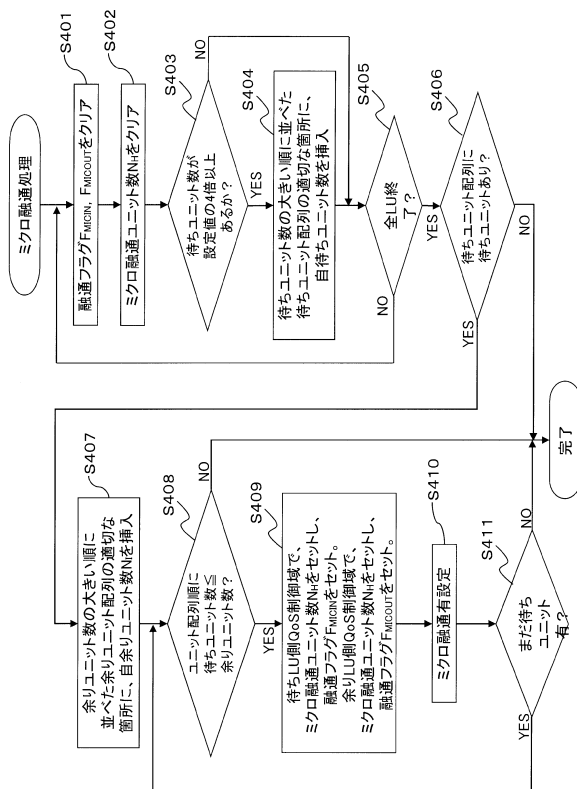
【図 9】



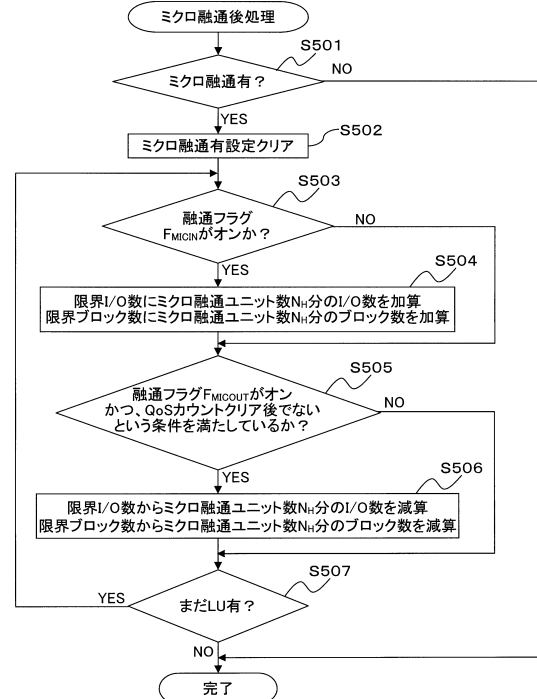
【図 10】



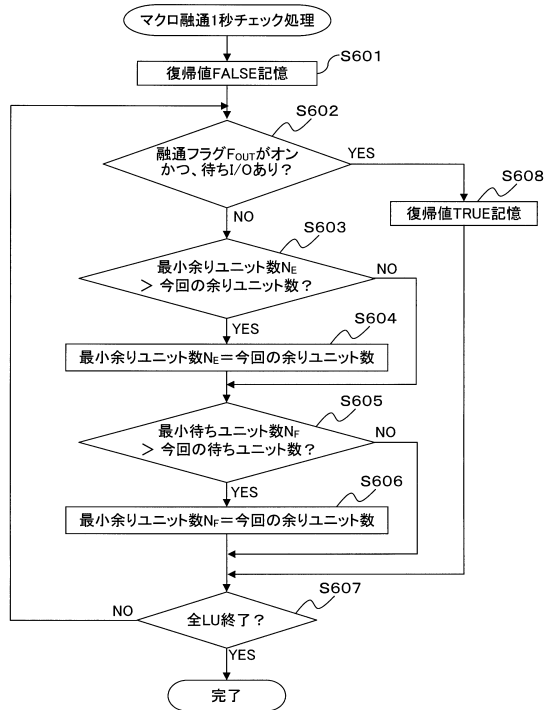
【図 11】



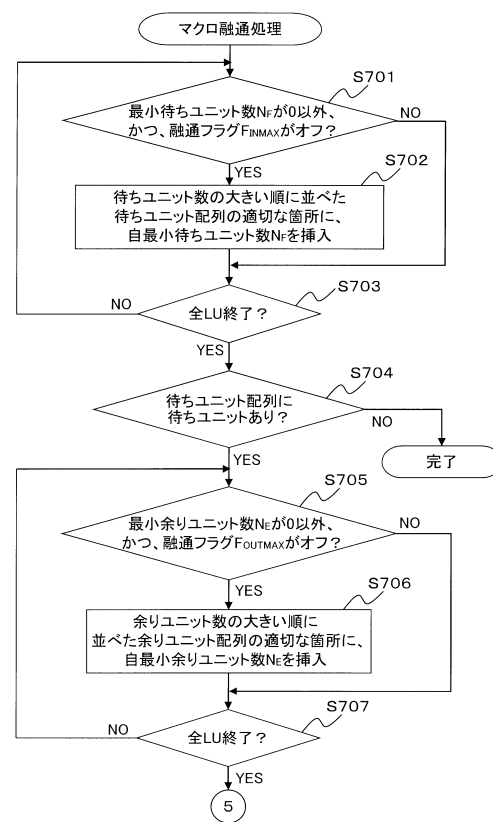
【図 12】



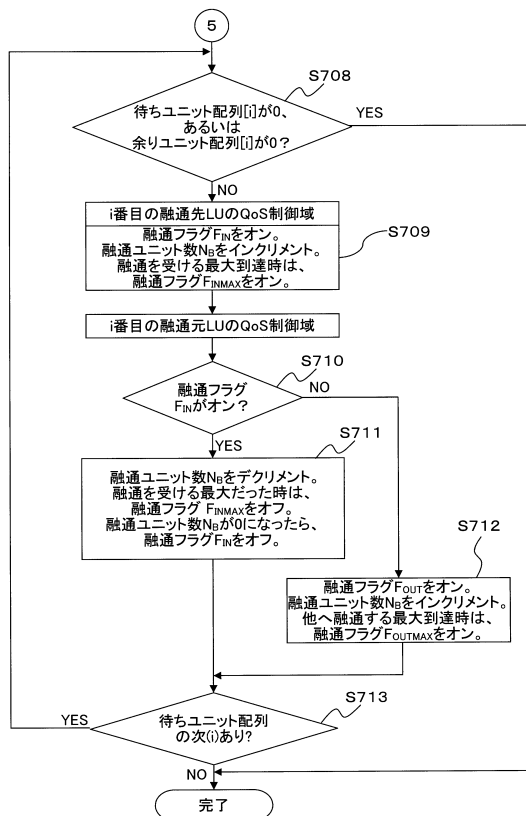
【図 13】



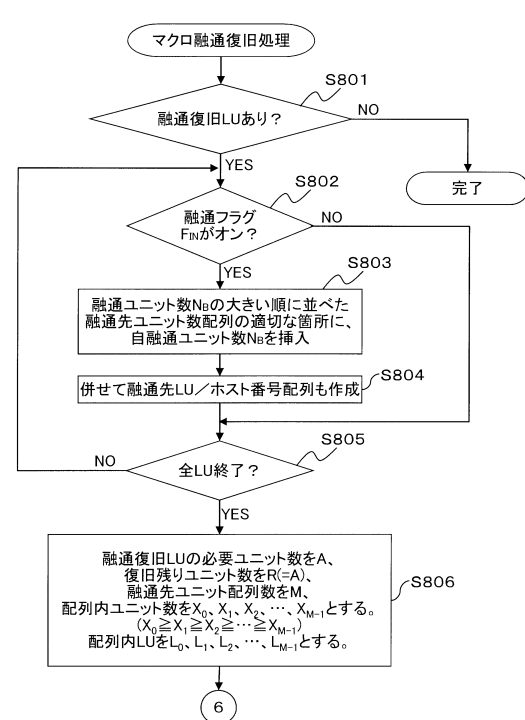
【図 14】



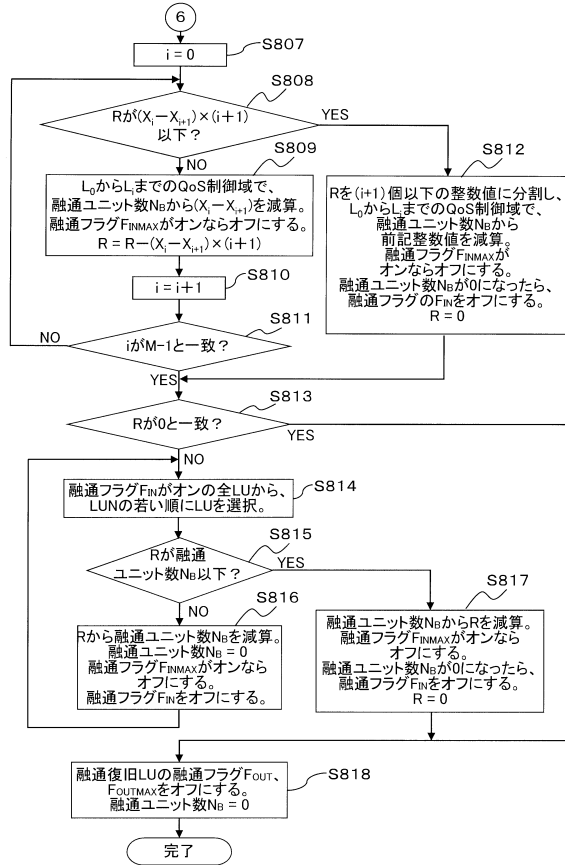
【図 15】



【図 16】



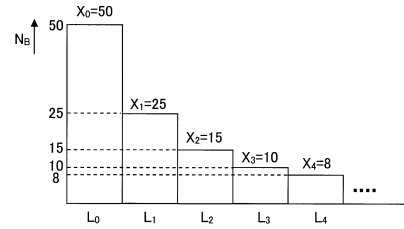
【図 17】



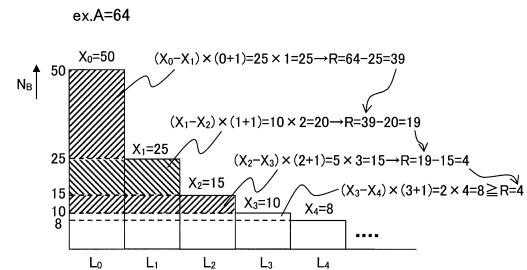
【図 18】

	ホスト	LUN	融通ユニット数Nb
L ₀	ホスト#1	LUN#1	50
L ₁	ホスト#2	LUN#3	25
L ₂	ホスト#1	LUN#4	15
L ₃	ホスト#2	LUN#2	10
L ₄	ホスト#2	LUN#5	8
⋮	⋮	⋮	⋮

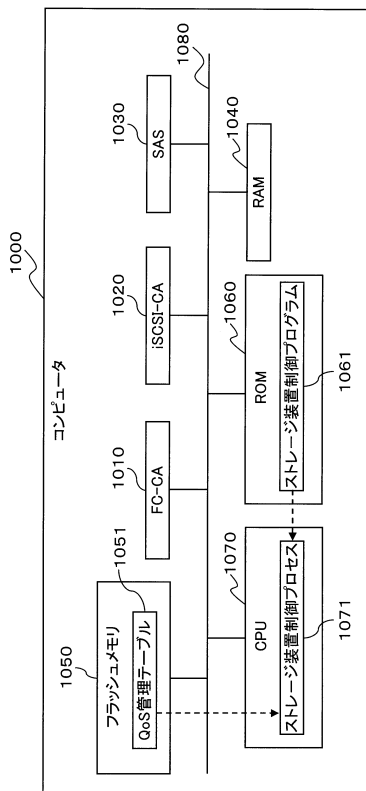
【図 19】



【図 20】



【図 21】



フロントページの続き

- (56)参考文献 特開2013-127758(JP,A)
特開2008-059478(JP,A)
国際公開第2013/101947(WO,A1)
特開2009-087175(JP,A)
特開2013-084177(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F3/06-3/08
G06F11/07、11/28-11/34、11/36
G06F13/10-13/14