

(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2010-97555
(P2010-97555A)

(43) 公開日 平成22年4月30日 (2010.4.30)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 1/28 (2006.01)	G06F 1/00 333Z	5B011
G06F 3/06 (2006.01)	G06F 3/06	5B065

審査請求 未請求 請求項の数 13 O L (全 60 頁)

(21) 出願番号 特願2008-269921 (P2008-269921)
(22) 出願日 平成20年10月20日 (2008.10.20)

(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
(74) 代理人 100075513
弁理士 後藤 政喜
(74) 代理人 100114236
弁理士 藤井 正弘
(74) 代理人 100120260
弁理士 飯田 雅昭
(72) 発明者 ▲高▼田 正法
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

最終頁に続く

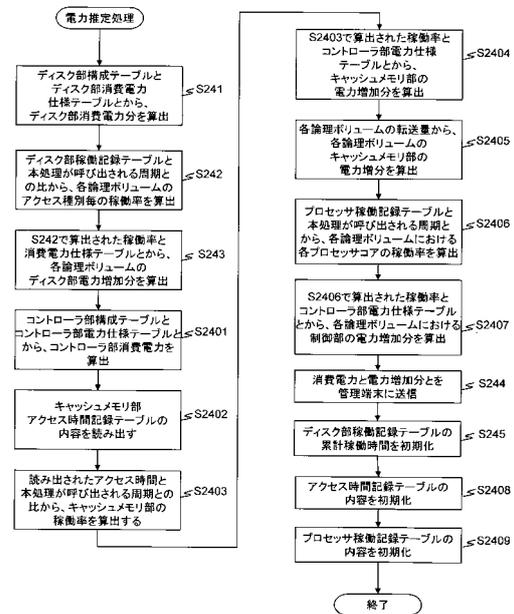
(54) 【発明の名称】 電力推定方法及び計算機システム

(57) 【要約】

【課題】本発明は、論理分割を行うストレージシステムにおいて、電力計を使うことなく、論理分割された領域の電力を求めることを目的とする。

【解決手段】管理計算機と、管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムにおける電力推定方法であって、ストレージシステムは、複数のディスクドライブの実領域上に論理記憶領域を作成し、第3のプロセッサが、論理記憶領域へのアクセスに対するディスクドライブの稼働時間から、論理記憶領域へのアクセスに対するディスクドライブの稼働率を算出するステップと、増加消費電力情報と、算出されたディスクドライブの稼働率とから、各アクセス種別における論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を算出するステップとを含むことを特徴とする。

【選択図】 図 4 2



【特許請求の範囲】

【請求項 1】

管理計算機と、前記管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムにおける電力推定方法であって、

前記管理計算機は、前記ストレージシステムに接続される第 1 のインタフェースと、前記第 1 のインタフェースに接続される第 1 のプロセッサと、前記第 1 のプロセッサに接続される第 1 のメモリとを備え、

前記ホスト計算機は、前記ストレージシステムと接続される第 2 のインタフェースと、前記第 2 のインタフェースに接続される第 2 のプロセッサと、前記第 2 のプロセッサに接続される第 2 のメモリとを備え、

前記ストレージシステムは、複数のディスクドライブへのデータの読み書きを制御する少なくとも一つの第 3 のプロセッサと、前記第 3 のプロセッサと接続される制御メモリと、前記管理計算機と接続される管理インタフェースと、前記ホスト計算機と接続されるホストインタフェースと、前記ディスクドライブと接続されるディスクインタフェースと、データを一時的に格納するキャッシュメモリ部とを備え、

前記第 3 のプロセッサ、前記管理インタフェース、前記ホストインタフェース、前記ディスクインタフェース、及び前記キャッシュメモリ部は、互いに接続され、

前記キャッシュメモリ部は、複数の第 3 のメモリと、前記複数の第 3 のメモリを制御するメモリコントローラとを備え、

前記制御メモリは、前記ストレージシステムを制御するためのストレージ制御プログラムと、前記ストレージシステムを制御するために必要な制御データとを格納し、

前記ストレージ制御プログラムは、電力を推定するための電力推定プログラムを含み、

前記制御データは、前記ストレージシステムが処理を実行しているときにおける前記ディスクドライブ、前記キャッシュメモリ部及び前記第 3 のプロセッサの消費電力の増加分を示す増加消費電力情報を含み、

前記ストレージシステムは、

前記複数のディスクドライブの実領域上に論理記憶領域を作成し、

前記ストレージシステムが備える物理資源を論理的に分割し、前記論理記憶領域と前記論理的に分割された物理資源とから構成される複数の論理区画を生成し、

前記論理区画を前記ホスト計算機に提供し、

前記増加消費電力情報は、前記論理記憶領域へのアクセス種別における前記ディスクドライブの消費電力の増加分に関する情報を含み、

前記方法は、

前記第 3 のプロセッサが、前記論理記憶領域へのアクセスに対する前記ディスクドライブの稼働時間から、前記論理記憶領域へのアクセスに対するディスクドライブの稼働率を算出するステップと、

前記第 3 のプロセッサが、前記増加消費電力情報と、前記算出されたディスクドライブの稼働率とから、各アクセス種別における前記論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を算出するステップと、

前記第 3 のプロセッサが、前記算出された論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を前記管理計算機に送信するステップとを含むことを特徴とする電力推定方法。

【請求項 2】

前記メモリコントローラは、前記第 3 のメモリに対するアクセス時間に関する情報を記憶し、

前記制御データは、前記論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の前記第 3 のメモリへのアクセス量を含み、

前記方法は、さらに、

前記第 3 のプロセッサが、前記ディスクドライブに対するアクセスのときに使用された前記キャッシュメモリ部へのアクセス時間から、前記ディスクドライブに対するアクセス

10

20

30

40

50

のときに使用されたキャッシュメモリ部の稼働率を算出するステップと、

前記第3のプロセッサが、前記算出されたディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の稼働率と前記増加消費電力情報とから、前記ディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を算出するステップと、

前記第3のプロセッサが、前記論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部への全アクセス量と、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部へのアクセス量とから、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の稼働率を算出するステップと、

前記第3のプロセッサが、前記算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の稼働率と、前記増加消費電力情報とから、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を算出するステップと、

前記第3のプロセッサが、前記算出されたディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分、及び前記算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を前記管理計算機に送信するステップとを含むことを特徴とする請求項1に記載の電力推定方法。

【請求項3】

前記制御データは、前記論理記憶領域のアクセスに対する処理に対する前記第3のプロセッサが稼働した稼働時間に関する情報を含み、

前記方法は、さらに、

前記第3のプロセッサが、前記論理記憶領域のアクセスに対する処理に対する前記第3のプロセッサが稼働した稼働時間から、前記論理記憶領域のアクセスに対する処理における第3のプロセッサの稼働率を算出するステップと、

前記第3のプロセッサが、前記増加消費電力情報と、前記算出された論理記憶領域のアクセスに対する処理における第3のプロセッサの稼働率とから、前記論理記憶領域のアクセスに対する第3のプロセッサの消費電力の増加分を算出するステップと、

前記第3のプロセッサが、前記算出された論理記憶領域に対するアクセスにおける第3のプロセッサの消費電力の増加分を前記管理計算機に送信するステップとを含むことを特徴とする請求項1に記載の電力推定方法。

【請求項4】

前記制御データは、前記ディスクドライブの稼働状態における稼働時間及び消費電力に関する情報を含み、

前記方法は、さらに、

前記ストレージシステムが、前記ディスクドライブの各状態における稼働時間から、前記ディスクドライブの各稼働状態の比率を算出するステップと、

前記ストレージシステムが、前記算出されたディスクドライブの各状態の比率と、前記ディスクドライブの各稼働状態における消費電力とから、各状態における前記ディスクドライブの消費電力を算出するステップとを含むことを特徴とする請求項1に記載の電力推定方法。

【請求項5】

前記ディスクドライブは、半導体記憶素子を含み、

前記制御データは、

ストレージ装置が処理を実行している場合における、前記半導体記憶素子の消費エネルギーの増加分を示す増加消費エネルギー情報を含み、

前記増加消費エネルギー情報は、各アクセス種別における単位データ転送量当たりの消費エネルギーの増加分に関する情報を含み、

各アクセス種別における前記論理記憶領域へのデータ転送量を記憶し、

前記方法は、

前記第3のプロセッサが、前記増加消費エネルギー情報と、各アクセス種別における前

10

20

30

40

50

記論理記憶領域へのデータ転送量とから、前記各アクセス種別における半導体記憶素子の消費エネルギーの増加分を算出するステップと、

前記第3のプロセッサが、前記算出された各アクセス種別における半導体記憶素子の消費エネルギーの増加分と、前記電力推定プログラムの実行周期とから、前記各アクセス種別における半導体記憶素子の消費電力の増加分を算出するステップとを含むことを特徴とする請求項1に記載の電力推定方法。

【請求項6】

前記制御データは、稼働するストレージ装置が処理を行っていない場合における、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力に関する情報を含み、

前記第1のメモリは、前記論理記憶領域と前記論理区画との対応関係に関する情報を記憶し、

前記方法は、

前記ストレージシステムが、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力の情報を前記管理端末に送信するステップと、

前記管理計算機が、前記増加消費電力情報から、電力を算出する部位の消費電力が変動するか否かを判定するステップと、

電力を算出する部位の消費電力が変動しないと判定された場合、前記管理計算機が、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサを使用している前記論理区画の数に基づいて、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力を前記各論理区画の定常消費電力として算出するステップと

電力を算出する部位の消費電力が変動しないと判定された場合、前記管理計算機が、前記算出された論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分、前記算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分、及び前記算出された論理記憶領域のアクセスに対する第3のプロセッサの消費電力の増加分と、前記論理記憶領域と前記論理区画との対応関係とから、前記論理区画における前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの消費電力の増加分を算出するステップと、

前記算出された各論理区画におけるディスクドライブ、キャッシュメモリ部及び第3プロセッサの定常消費電力、及び前記算出された各論理区画におけるディスクドライブ、キャッシュメモリ部及び第3プロセッサの消費電力の増加分を表示するステップとを含むことを特徴とする請求項1乃至請求項5に記載の電力推定方法。

【請求項7】

管理計算機と、前記管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムであって、

前記管理計算機は、前記ストレージシステムに接続される第1のインタフェースと、前記第1のインタフェースに接続される第1のプロセッサと、前記第1のプロセッサに接続される第1のメモリとを備え、

前記ホスト計算機は、前記ストレージシステムと接続される第2のインタフェースと、前記第2のインタフェースに接続される第2のプロセッサと、前記第2のプロセッサに接続される第2のメモリとを備え、

前記ストレージシステムは、複数のディスクドライブへのデータの読み書きを制御する少なくとも一つの第3のプロセッサと、前記第3のプロセッサと接続される制御メモリと、前記管理計算機と接続される管理インタフェースと、前記ホスト計算機と接続されるホストインタフェースと、前記ディスクドライブと接続されるディスクインタフェースと、データを一時的に格納するキャッシュメモリ部とを備え、

前記第3のプロセッサ、前記管理インタフェース、前記ホストインタフェース、前記ディスクインタフェース、及び前記キャッシュメモリ部は、互いに接続され、

前記キャッシュメモリ部は、複数の第3のメモリと、前記複数の第3のメモリを制御す

10

20

30

40

50

るメモリコントローラとを備え、

前記制御メモリは、前記ストレージシステムを制御するためのストレージ制御プログラムと、前記ストレージシステムを制御するために必要な制御データとを格納し、

前記ストレージ制御プログラムは、電力を推定するための電力推定プログラムを含み、

前記制御データは、前記ストレージシステムが処理を実行しているときにおける前記ディスクドライブ、前記キャッシュメモリ部及び前記第3のプロセッサの消費電力の増加分を示す増加消費電力情報を含み、

前記ストレージシステムは、

前記複数のディスクドライブの実領域上に論理記憶領域を作成し、

前記ストレージシステムが備える物理資源を論理的に分割し、前記論理記憶領域と前記論理的に分割された物理資源とから構成される複数の論理区画を生成し、

前記論理区画を前記ホスト計算機に提供し、

前記増加消費電力情報は、前記論理記憶領域へのアクセス種別における前記ディスクドライブの消費電力の増加分に関する情報を含み、

前記論理記憶領域へのアクセスに対する前記ディスクドライブの稼働時間から、前記論理記憶領域へのアクセスに対するディスクドライブの稼働率を算出し、

前記増加消費電力情報と、前記算出されたディスクドライブの稼働率とから、各アクセス種別における前記論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を算出し、

前記算出された論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を前記管理計算機に送信することを特徴とする計算機システム。

【請求項8】

前記メモリコントローラは、前記第3のメモリに対するアクセス時間に関する情報を記憶し、

前記制御データは、前記論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の前記第3のメモリへのアクセス量を含み、

前記ストレージシステムは、

前記ディスクドライブに対するアクセスのときに使用された前記キャッシュメモリ部へのアクセス時間から、前記ディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の稼働率を算出し、

前記算出された前記ディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の稼働率と前記増加消費電力情報とから、前記ディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を算出し、

前記論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部への全アクセス量と、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部へのアクセス量とから、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の稼働率を算出し、

前記算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の稼働率と、前記増加消費電力情報とから、前記各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を算出し、

前記算出されたディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分、及び前記算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分を前記管理計算機に送信することを特徴とする請求項7に記載の計算機システム。

【請求項9】

前記制御データは、前記論理記憶領域のアクセスに対する処理に対する前記第3のプロセッサが稼働した稼働時間に関する情報を含み、

前記ストレージシステムは、さらに、

前記論理記憶領域のアクセスに対する処理に対する前記第3のプロセッサが稼働した稼働時間から、前記論理記憶領域のアクセスに対する処理における第3のプロセッサの稼働

10

20

30

40

50

率を算出し、

前記増加消費電力情報と、前記算出された論理記憶領域のアクセスに対する処理における第3のプロセッサの稼働率とから前記論理記憶領域のアクセスに対する第3のプロセッサの消費電力の増加分を算出し、

前記算出された論理記憶領域に対するアクセスにおける第3のプロセッサの消費電力の増加分を前記管理計算機に送信することを特徴とする請求項7に記載の計算機システム。

【請求項10】

前記制御データは、前記ディスクドライブの稼働状態における稼働時間及び消費電力分に関する情報をを含み、

前記ストレージシステムは、さらに、

前記ディスクドライブの各状態における稼働時間から、前記ディスクドライブの各稼働状態の比率を算出し、

前記算出されたディスクドライブの各状態の比率と、前記ディスクドライブの各稼働状態における消費電力とから、各状態における前記ディスクドライブの消費電力を算出することを特徴とする請求項7に記載の計算機システム。

【請求項11】

前記ディスクドライブは、半導体記憶素子を含み、

前記制御データは、

ストレージ装置が処理を実行している場合における、前記半導体記憶素子の消費エネルギーの増加分を示す増加消費エネルギー情報を含み、

前記増加消費エネルギー情報は、各アクセス種別における単位データ転送量当たりの消費エネルギーの増加分に関する情報をを含み、

各アクセス種別における前記論理記憶領域へのデータ転送量を記憶し、

前記ストレージシステムは、

前記増加消費エネルギー情報と、各アクセス種別における前記論理記憶領域へのデータ転送量とから、前記アクセス種別ごとに半導体記憶素子の消費エネルギーの増加分を算出し、

前記算出された各アクセス種別における半導体記憶素子の消費エネルギーの増加分と、前記電力推定プログラムの実行周期とから、前記アクセス種別ごとの半導体記憶素子の消費電力の増加分を算出することを特徴とする請求項7に記載の計算機システム。

【請求項12】

前記制御データは、稼働するストレージ装置が処理を行っていない場合における、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力に関する情報をを含み、

前記第1のメモリは、前記論理記憶領域及び前記論理的に分割された物理資源と、前記論理区画との対応関係に関する情報を記憶し、

前記ストレージシステムは、

前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力の情報を前記管理端末に送信し、

前記管理計算機は、

前記増加消費電力情報を参照し、電力を算出する部位の消費電力が変動するか否かを判定し、

電力を算出する部位の消費電力が変動しないと判定された場合、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサを使用している前記論理区画の数に基づいて、前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの定常消費電力を前記各論理区画の定常消費電力として算出し、

電力を算出する部位の消費電力が変動しないと判定された場合、前記算出された論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分、算出された各論理記憶領域に対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分、及び前記算出された論理記憶領域のアクセスに対する第3のプロセッサの消費電力の増加分

10

20

30

40

50

と、前記論理記憶領域と前記論理区画との対応関係とから、前記論理区画における前記ディスクドライブ、前記キャッシュメモリ部及び前記第3プロセッサの消費電力の増加分を算出し、

前記算出された各論理区画におけるディスクドライブ、キャッシュメモリ部及び第3プロセッサの定常消費電力、及び前記算出された各論理区画におけるディスクドライブ、キャッシュメモリ部及び第3プロセッサの消費電力の増加分を表示することを特徴とする請求項7乃至請求項11に記載の計算機システム。

【請求項13】

管理計算機と、前記管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムであって、

前記管理計算機は、ネットワークに接続される第1のインタフェースと、前記第1のインタフェースに接続される第1のプロセッサと、前記第1のプロセッサに接続される第1のメモリとを備え、

前記ホスト計算機は、前記ストレージシステムと接続される第2のインタフェースと、前記第2のインタフェースに接続される第2のプロセッサと、前記第2のプロセッサに接続される第2のメモリとを備え、

前記ストレージシステムは、

複数のディスクドライブへのデータの読み書きを制御する少なくとも一つの第3のプロセッサと、前記第3のプロセッサと接続される制御メモリと、前記管理計算機と接続される管理インタフェースと、前記ネットワークと接続される少なくとも一つのホストインタフェースと、前記ディスクドライブと接続される少なくとも一つのディスクインタフェースと、データを一時的に格納するキャッシュメモリ部とを備え、

前記第3のプロセッサ、前記管理インタフェース、前記ホストインタフェース、前記ディスクインタフェース、及び前記キャッシュメモリ部は、互いにスイッチ部によって接続され、

前記制御メモリは、前記ストレージシステムを制御するためのストレージ制御プログラムと、前記ストレージシステムを制御するために必要な制御データとを備え、

前記キャッシュメモリ部は、複数の第3のメモリと、前記複数の第3のメモリを制御するメモリコントローラとを備え、

前記メモリコントローラは、前記メモリへのアクセス量を記録するアクセス時間記録テーブルを備え、

前記ストレージシステムは、

前記複数のディスクドライブの実領域上に論理ボリュームを作成し、

前記ストレージシステムが備える物理資源を論理的に分割し、前記論理ボリュームと前記論理的に分割された物理資源とから構成される複数の仮想ストレージを生成し、

前記論理区画を前記ホスト計算機に提供し、

前記制御データは、

ディスクドライブの構成を記憶するディスク部構成テーブルと、

前記論理記ボリュームごとに各アクセス種別における前記ディスクドライブの稼働時間を記憶するディスク部稼働時間記録テーブルと、

各アクセス種別における前記ディスクドライブの消費電力の増加分、及び稼働している前記ストレージシステムが処理を実行していない場合における前記ディスクドライブの消費電力とを記憶するディスク電力仕様テーブルと、

前記制御部及び前記キャッシュメモリ部の構成を記憶するコントロール部構成テーブルと、

稼働している前記ストレージシステムが処理を実行していない場合における前記制御部及び前記キャッシュメモリ部の消費電力を記憶するコントロール部電力仕様テーブルと、

稼働している前記ストレージシステムが処理を実行している場合における前記制御部及び前記キャッシュメモリ部の消費電力増加分を記憶するコントロール部電力変動仕様テーブルと、

10

20

30

40

50

前記論理ボリュームごとに前記プロセッサコアの稼働時間を記憶するプロセッサ稼働記録テーブルと、

前記論理ボリュームに対するアクセスのときに使用されたキャッシュメモリ部へのアクセス量を記憶するキャッシュアクセス量テーブルと、

前記管理計算機は、

前記ストレージシステムを管理するためのストレージ管理プログラムを備え、

前記論理ボリューム及び前記論理的に分割された物理資源と前記仮想ストレージとの対応関係を記憶する、仮想ストレージ - ボリューム対応テーブルを含み、

仮想ストレージの電力を表示するための仮想ストレージ電力表示機能とを含み、

電力を推定するための電力推定機能を備え、

10

前記ストレージシステムは、

ディスク部構成テーブルとディスク電力仕様テーブルとを参照し、前記ディスク部の消費電力を算出し、

前記ディスク部稼働時間記録テーブルと前記電力指定機能が実行される周期とから、前記論理ボリュームへのアクセスに対するディスクドライブの稼働率を算出し、

ディスク電力仕様テーブルと前記前記算出されたディスクドライブの稼働率とから、各々のアクセス種別における前記論理ボリュームへのアクセスに対するディスク部の消費電力増加分を算出し、

コントロール部構成テーブルとコントロール部電力仕様テーブルとから、前記キャッシュメモリ部及び前記制御部の消費電力を算出し、

20

アクセス時間記録テーブルと前記電力指定機能が実行される周期とから、キャッシュメモリ部の稼働率を算出し、

コントロール部電力変動仕様テーブルと前記算出されたキャッシュメモリ部の稼働率とから前記キャッシュメモリ部の電力増加分を算出し、

前記アクセス時間記録テーブル、前記コントロール部電力変動仕様テーブル、及び前記キャッシュアクセス量テーブルから、前記論理ボリュームへのアクセスに対する前記キャッシュメモリ部の消費電力増加分を算出し、

前記プロセッサ稼働記録テーブルと前記電力指定機能が実行される周期とから、前記論理ボリュームへのアクセスに対する前記プロセッサコアの稼働率を算出し、

前記コントロール部電力変動仕様テーブルと前記算出されたプロセッサコアの稼働率とから、前記論理ボリュームへのアクセスに対する制御部の消費電力増加分を算出し、

30

前記算出されえた各消費電力及び各消費電力増加分を前記管理計算機に送信し、

前記管理計算機は、

仮想ストレージ ボリューム対応テーブルを参照し、前記仮想ボリュームごとにディスク部の電力増加分、キャッシュメモリ部の電力増加分、制御部の電力増加分を算出し、算出結果、前記ディスク部の消費電力分及びコントロール部消費電力を表示することを特徴とする計算機システム。

【発明の詳細な説明】

【技術分野】

【0001】

40

本発明は、データを複数の磁気ディスク装置に格納するストレージシステムに関する。

【背景技術】

【0002】

複数の磁気ディスク装置から構成されるディスクアレイでサブシステムを構成し、データを前記ディスクアレイに格納する、いわゆるストレージシステムが知られている。

【0003】

このようなストレージシステムは、多数の計算機を接続して用いるため、データセンタ等の、多くの計算機が集中する環境で用いられることが多い。

【0004】

一方、近年、データセンタ、及びデータセンタにおける機器の消費電力は増大する傾向

50

にある。その中でもストレージシステムの消費電力は大きく、ストレージシステム内の装置の稼働状況に応じて、ディスクの電源を落とす等の、データの移行を行うなどの省電力制御が求められている。前述の省電力制御を実現するためには、装置の稼働状況に応じた消費電力を、細かい構成部位単位で、精確に求める必要がある。

【 0 0 0 5 】

また、ストレージシステムは、一般的に多数のホスト計算機と接続されるため、複数のホスト計算機によってストレージシステム内の同じ物理資源が共有される。一般的には、ストレージシステム内の物理資源は論理的に分割され、ストレージシステムは、当該論理資源を各ホスト計算機に提供している。

【 0 0 0 6 】

このとき、ホスト計算機及びストレージシステムを含む全体のシステムで省電力制御を行うためには、ストレージシステムの消費電力を、対応するホスト計算機毎、即ち論理資源毎に求められることが望ましい。

【 0 0 0 7 】

前述した課題を解決する方法としては、各々の稼働部位の電力を測定し、測定された電力と、定格電力、I/O数、及び転送量などとを関連付けて、管理者に提示する方法がある（例えば、特許文献1参照）。

【 0 0 0 8 】

他の方法としては、装置の消費電力に閾値を設け、電力が閾値以下になるようディスクの回転数、及び電源のON/OFFの制御を行う方法がある（例えば、特許文献2参照）。特許文献2に記載されている発明では、各部品が消費する電力の最大値、即ち、定格電力を用いた制御を行っている。

【 0 0 0 9 】

また、他の方法として、ノートPC等に用いられる小型ハードドライブドライブについて、詳細な状態シミュレーション（I/Oトレースから、シーク、頭出し待ち、及びアクセスなどの状態遷移のシミュレーション）を行い、消費電力を推定する方法がある（例えば、非特許文献1参照）。また、非特許文献1には、簡略な方法として、稼働時間の割合を用いて消費電力を推定する方法も開示されている。

【特許文献1】特開2007-079754号公報

【特許文献2】特開2008-003719号公報

【非特許文献1】"Modeling Hard-Disk Power Consumption", Proceedings of 2nd USENIX Conference on File and Storage Technologies, 2003

【発明の開示】

【発明が解決しようとする課題】

【 0 0 1 0 】

しかし、特許文献1に記載された発明のように電力を実測する方法では、例えば、論理ボリュームに対応するディスクなど、細かい単位で消費電力を知りたい場合、消費電力を求める構成単位毎に電力計を用いる必要がある。また、任意の複数のディスクを組み合わせると論理ボリュームが構成される場合、消費電力を求める構成単位が変化するため、更に細かい単位、即ち、各ディスクに電力計を用いて消費電力を求める必要がある。また、複数の論理ボリュームが同一のディスクを共有している場合、前述した方法では各論理ボリューム毎の電力を求めることができない。

【 0 0 1 1 】

また、特許文献2に記載されている発明のように定格電力による消費電力積み上げでは、実際の消費電力と大きな乖離があるため、精度のよい制御ができない。また、複数の論理区画が同一のハードウェア資源を共有する場合における各論理区画毎の消費電力を求める方法が開示されていない。

【 0 0 1 2 】

また、非特許文献1に記載された発明のようにディスクの状態シミュレーションによる推定では、ディスクの詳細な状態をシミュレーションする必要があり、計算量が大きくな

10

20

30

40

50

る。また、稼働時間の割合を用いた推定方法では、ディスクから稼働時間の情報を取り出せるようにディスクのファームウェアを改変し、その情報をストレージのコントローラで取り出す仕組みを作る必要がある。即ち、全ディスクの種別毎に、ディスクのファームウェアを改変する必要がある。また、非特許文献 1 に記載された発明では、ディスクアクセスの種類（空間、時間局所性）に特徴（偏り）がある場合、十分な精度が出ない。また、ストレージシステムが複数の仮想ストレージをサポートし、それらが同一のハードウェア資源を共有する場合、仮想ストレージ毎の消費電力を求める方法が開示されていない。また、非特許文献 1 には、コントローラ部について、電力を推定する方法も開示されていない。

【 0 0 1 3 】

以上のように、従来の方法では、ストレージシステムについて、コントローラ部、ディスクアレイ部分の電力を求めることができない。また、仮想ストレージをサポートするストレージシステムにおいて、各仮想ストレージの消費電力を求めることができない。

【 0 0 1 4 】

本発明は、前述の問題点を鑑みてなされたものであり、ストレージシステムにおいて、各部位、及び各論理区画毎の消費電力を求めることである。

【 課題を解決するための手段 】

【 0 0 1 5 】

本発明の代表的な一例を示せば以下の通りである。即ち、管理計算機と、前記管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムにおける電力推定方法であって、前記管理計算機は、前記ストレージシステムに接続される第 1 のインタフェースと、前記第 1 のインタフェースに接続される第 1 のプロセッサと、前記第 1 のプロセッサに接続される第 1 のメモリとを備え、前記ホスト計算機は、前記ストレージシステムと接続される第 2 のインタフェースと、前記第 2 のインタフェースに接続される第 2 のプロセッサと、前記第 2 のプロセッサに接続される第 2 のメモリとを備え、前記ストレージシステムは、複数のディスクドライブへのデータの読み書きを制御する少なくとも一つの第 3 のプロセッサと、前記第 3 のプロセッサと接続される制御メモリと、前記管理計算機と接続される管理インタフェースと、前記ホスト計算機と接続されるホストインタフェースと、前記ディスクドライブと接続されるディスクインタフェースと、データを一時的に格納するキャッシュメモリ部とを備え、前記第 3 のプロセッサ、前記管理インタフェース、前記ホストインタフェース、前記ディスクインタフェース、及び前記キャッシュメモリ部は、互いに接続され、前記キャッシュメモリ部は、複数の第 3 のメモリと、前記複数の第 3 のメモリを制御するメモリコントローラとを備え、前記制御メモリは、前記ストレージシステムを制御するためのストレージ制御プログラムと、前記ストレージシステムを制御するために必要な制御データとを格納し、前記ストレージ制御プログラムは、電力を推定するための電力推定プログラムを含み、前記制御データは、前記ストレージシステムが処理を実行しているときにおける前記ディスクドライブ、前記キャッシュメモリ部及び前記第 3 のプロセッサの消費電力の増加分を示す増加消費電力情報を含み、前記ストレージシステムは、前記複数のディスクドライブの実領域上に論理記憶領域を作成し、前記ストレージシステムが備える物理資源を論理的に分割し、前記論理記憶領域と前記論理的に分割された物理資源とから構成される複数の論理区画を生成し、前記論理区画を前記ホスト計算機に提供し、前記増加消費電力情報は、前記論理記憶領域へのアクセス種別における前記ディスクドライブの消費電力の増加分に関する情報を含み、前記方法は、前記第 3 のプロセッサが、前記論理記憶領域へのアクセスに対する前記ディスクドライブの稼働時間から、前記論理記憶領域へのアクセスに対するディスクドライブの稼働率を算出するステップと、前記第 3 のプロセッサが、前記増加消費電力情報と、前記算出されたディスクドライブの稼働率とから、各アクセス種別における前記論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を算出するステップと、前記第 3 のプロセッサが、前記算出された論理記憶領域へのアクセスに対するディスクドライブの消費電力の増加分を前記管理計算機に送信するステップとを含むことを特徴とする。

10

20

30

40

50

【発明の効果】

【0016】

本発明によれば、ストレージシステムのディスク部分、及びその他の構成部分の消費電力を求めることができ、これによって論理区画毎の消費電力を求めることが可能になる。その結果、より精度の良い電力制御が可能となる。

【発明を実施するための最良の形態】

【0017】

以下、本発明を実施するための形態について、図を用いて説明する。

【0018】

[第1の実施形態]

本発明の第1の実施形態は、ストレージシステムにおいて、ストレージシステムに含まれディスク、及びストレージシステムのコントローラ部分の稼働状況に応じた消費電力を推定する発明である。

【0019】

まず、第1の実施形態の概要について説明する。

【0020】

第1の実施形態では、稼働状況に応じた消費電力を、以下のようにして求める。ストレージシステムが、各々のディスクの種別における、Idle時の消費電力、及び各I/O種別（ランダムまたはシーケンシャル、及びリードまたはライト）のアクセスを常に処理している状態での消費電力（以下、消費電力基礎データ）に関する情報を保持している場合、I/O種別毎に、ストレージシステムの制御部がディスクからの応答を待っている時間を集計する。ストレージシステムは、保持される消費電力基礎データと集計された稼働情報とから、稼働状況に応じた消費電力を求める。また、ストレージシステムは、コントローラ部分について、プロセッサの稼働時間、及びキャッシュメモリへのアクセス時間を記録し、前述のプロセッサの稼働時間及びキャッシュメモリへのアクセス時間から消費電力を求める。なお、以下、任意のI/O種別のアクセスを常に処理している状態を、限界まで稼働している状態ともいう。

【0021】

図1は、本発明の第1の実施形態の計算機システムの構成を説明するブロック図である。

【0022】

計算機システムは、ストレージシステム1、管理用端末4、及びホストコンピュータ71を備える。ストレージシステム1は、管理用端末4とホストコンピュータ71とに接続されている。

【0023】

ストレージシステム1は、ディスクドライブ6、ホストインターフェース（I/F）部72、ディスクインターフェース（I/F）部74、キャッシュメモリ部73、制御部75、管理端末用インターフェース（I/F）部77、及びスイッチ部76を備える。

【0024】

ホストI/F部72は、ホストコンピュータ71と接続するためのインターフェースであり、ストレージシステム1はホストI/F部72を介してホストコンピュータ71とデータをやり取りする。

【0025】

ホストI/F部72は、プロトコル変換LSI721及びデータ転送LSI722を備える。

【0026】

プロトコル変換LSI721は、ホストコンピュータ71との接続に用いられるFibre Channelのプロトコル、及びストレージ内部で用いられるPCIのプロトコルを変換する。

【0027】

データ転送 L S I 7 2 2 は、制御部 7 5 からの指示に従って、プロトコル変換 L S I 7 2 1 とキャッシュメモリ部 7 3 との間のデータを転送する。

【 0 0 2 8 】

なお、図 1 に示す例において、ホスト I / F 部 7 2 は二つであるが、ストレージシステム 1 は、二つ以上ホスト I / F 部 7 2 を備えていてもよい。また、プロトコル変換 L S I 7 2 1 は、F i b r e C h a n n e l のプロトコル、及び P C I のプロトコル以外のプロトコルを変換するものであってもよい。

【 0 0 2 9 】

キャッシュメモリ部 7 3 は、ディスクドライブ 6 に格納されているデータの一部に高速にアクセスするためのデータを格納している。

【 0 0 3 0 】

ディスク I / F 部 7 4 は、ストレージシステム 1 が備えるディスクドライブ 6 と接続するためのインタフェースであり、ストレージシステム 1 はディスク I / F 部 7 4 を介してディスクドライブ 6 とデータのやり取りしている。

【 0 0 3 1 】

ディスク I / F 部 7 4 は、プロトコル変換 L S I 7 4 1 及びデータ転送 L S I 7 4 2 を備える。

【 0 0 3 2 】

プロトコル変換 L S I 7 4 1 は、ディスクドライブ 6 との接続に用いられる F i b r e C h a n n e l または S A S (S e r i a l A t t a c h e d S C S I) のプロトコル、及びストレージ内部で用いられるプロトコルを変換する。

【 0 0 3 3 】

データ転送 L S I 7 4 2 は、制御部 7 5 からの指示に従って、ディスクドライブ 6 とキャッシュメモリ部 7 3 との間のデータを転送する。

【 0 0 3 4 】

なお、プロトコル変換 L S I 7 4 1 は、F i b r e C h a n n e l のプロトコル、S A S のプロトコル、及びストレージ内部で用いられるプロトコル以外のプロトコルを変換するものであってもよい。

【 0 0 3 5 】

制御部 7 5 は、ストレージシステム 1 を制御している。制御部 7 5 は、プロセッサ 7 8 と制御メモリ 7 9 とを備える。また、プロセッサ 7 8 は、複数のプロセッサコア 7 8 1 を備える。

【 0 0 3 6 】

制御メモリ 7 9 は、ストレージ制御プログラム 2 と、ストレージ制御に必要な情報である制御データ 3 とを格納する。プロセッサ 7 8 は、ストレージシステム 1 を制御するストレージ制御プログラム 2 を実行し、ストレージシステム 1 を制御する。例えば、制御部 7 5 は、ストレージシステム 1 内のデータ転送を制御する。

【 0 0 3 7 】

管理端末用 I / F 部 7 7 は、管理用端末 4 と接続するためのインタフェースであり、ストレージシステム 1 は管理端末用 I / F 部 7 7 を介して管理用端末 4 とデータのやり取りをする。

【 0 0 3 8 】

スイッチ部 7 6 は、ホスト I / F 部 7 2 、キャッシュメモリ部 7 3 、ディスク I / F 部 7 4 、制御部 7 5 及び管理端末用 I / F 7 7 を接続する。

【 0 0 3 9 】

管理用端末 4 は、プロセッサ (図示省略) 、不揮発性記憶媒体 (図示省略) 、及びメモリ (図示省略) を備え、不揮発性記憶媒体はストレージシステム 1 の管理を行うストレージ管理プログラム 5 が格納されており、当該ストレージ管理プログラム 5 によって、ストレージシステム 1 を管理している。

【 0 0 4 0 】

10

20

30

40

50

ホストコンピュータ 71 は、プロセッサ（図示省略）、不揮発性記憶媒体（図示省略）、及びメモリ（図示省略）を備え、ストレージシステム 1 が備えるディスクドライブ 6 からデータを読み出したり、また、ディスクドライブ 6 にデータを書き込んだりする。

【0041】

図 2 は、本発明の第 1 の実施形態のキャッシュメモリ部 73 の構成を説明するブロック図である。

【0042】

キャッシュメモリ部 73 は、メモリコントローラ L S I 731 とメモリモジュール（D I M M）733 とを備える。

【0043】

メモリコントローラ L S I 731 は、ストレージシステム 1 の内部で用いられる P C I のプロトコル、及びメモリモジュール 733 へのアクセスに用いられる S D R A M アクセスプロトコルを変換する。また、メモリコントローラ L S I 731 は、メモリモジュール 733 へアクセスされた時間を記録するアクセス時間記録テーブル 732 を備える。なお、メモリコントローラ L S I 731 は、ストレージシステム 1 の内部で用いられる P C I のプロトコル、及びメモリモジュール 733 へのアクセスに用いられる S D R A M アクセスプロトコル以外のプロトコルを変換するものであってもよい。

【0044】

図 3 は、本発明の第 1 の実施形態のアクセス時間記録テーブル 732 の一例を示す図である。

【0045】

アクセス時間記録テーブル 732 は、チャンネル番号 7321 及びアクセス時間 7322 を含む。チャンネル番号 7321 は、メモリモジュール 733 とメモリコントローラ L S I 731 との間のデータ転送路であるチャンネルの識別子を格納する。アクセス時間 7322 は、それぞれのチャンネルにおいて過去にアクセスされた累積時間を示す値を格納する。

【0046】

図 4 は、本発明の第 1 の実施形態のストレージ制御プログラム 2 が備える機能を説明するブロック図である。

【0047】

ストレージ制御プログラム 2 は、ホスト I / O 受領機能 21、ディスク I / O 発行機能 22、ディスク I / O 完了処理機能 23、電力推定機能 24、ジョブ起動処理機能 25、及びジョブ終了処理機能 26 を備える。各々の機能についての詳細は、図 16 ~ 図 20 を用いて後述する。

【0048】

図 5 は、本発明の第 1 の実施形態の制御データ 3 の構成を説明するブロック図である。

【0049】

制御データ 3 は、ディスク部構成テーブル 31、ディスク電力仕様テーブル 32、ディスク稼働記録テーブル 33、T a g 管理テーブル 34、I / O 発行状況テーブル 35、コントローラ部構成テーブル 36、コントローラ部電力仕様テーブル 37、コントローラ部電力変動仕様テーブル 38、プロセッサ稼働記録テーブル 39、及びプロセッサ稼働状況テーブル 30 を備える。以下、各々のテーブルについて説明する。

【0050】

図 6 は、本発明の第 1 の実施形態のディスク部構成テーブル 31 の一例を示す図である。

【0051】

ディスク部構成テーブル 31 は、ストレージシステム 1 に搭載されているディスクドライブ 6 の種別、及びディスクドライブ 6 が搭載されている物理的な位置を管理する。ディスク部構成テーブル 31 は、ディスク 6 を増設または減設するとき更新される。また、電力を推定するときに、電力推定機能 24 がディスク部構成テーブル 31 を参照する。

【0052】

10

20

30

40

50

ディスク部構成テーブル31は、ディスク番号311及びディスク型番312を含む。

【0053】

ディスク番号311は、ディスクドライブ6が搭載される物理的な位置示す情報を格納する。例えば、ストレージシステム1のスロット番号を識別する識別子が格納される。

【0054】

ディスク型番312は、ディスク番号311に対応する位置に、どの種別のディスクドライブ6が搭載されているかを示す情報を格納する。例えば、ディスクドライブ6の型番が格納される。電力に関する仕様が異なるディスクドライブ6を識別できればどのような情報であってもよい。

【0055】

図7は、本発明の第1の実施形態のディスク電力仕様テーブル32の一例を示す図である。

【0056】

ディスク電力仕様テーブル32は、ストレージシステム1に搭載されているディスクドライブ6の種別毎に、各々の状態における消費電力を管理する。ディスク電力仕様テーブル32は、新規種別のディスクドライブ6が追加されたときに更新される。また、電力を推定するときに、電力推定機能24がディスク電力仕様テーブル32を参照する。

【0057】

ディスク電力仕様テーブル32は、ディスク型番321、Idle時電力322、Random Read時電力増分323、Random Write時電力増分324、Sequential Read時電力増分325、及びSequential Write時電力増分326を含む。

【0058】

ディスク型番321は、ディスクの種別を識別する識別子を格納する。例えば、ディスクドライブ6の型番などが格納される。図6のディスク型番312と同一のものである。

【0059】

Idle時電力322は、対応するディスクドライブ6がIdle状態のときの消費電力を示す値を格納する。ここで、Idle状態とは、電源が入っておりアクセスを受け付け可能であるが、アクセスが来ていない状態を示す。

【0060】

Random Read時電力増分323は、対応するディスクドライブ6がRandom Readのみを受け付けており、かつ、限界まで稼働している状態における消費電力とIdle時電力322との差分を示す値を格納する。ここで、Random Readとは、ホストコンピュータ71がディスクドライブ6の全体に対し、局所性が無いリードアクセスを発行している状態を示す。

【0061】

Random Write時電力増分324は、対応するディスクドライブ6がRandom Writeのみを受け付けており、かつ、限界まで稼働している状態における消費電力とIdle時電力322との差分を示す値を格納する。ここで、Random Writeとは、ホストコンピュータ71がディスクの全体に対し、局所性が無いライトアクセスを発行している状態を示す。

【0062】

Sequential Read時電力増分325は、対応するディスクドライブ6がSequential Readのみを受け付けており、かつ、限界まで稼働している状態における消費電力とIdle時電力との差分を示す値を格納する。ここで、Sequential Readとは、ホストコンピュータ71がディスクの連続した領域に対し、順番にリードアクセスを発行している状態を示す。

【0063】

Sequential Write時電力増分326は、対応するディスクドライブ6がSequential Writeのみを受け付けており、かつ、限界まで稼働してい

10

20

30

40

50

る状態における消費電力とIdle時電力との差分を示す値を格納する。ここで、Sequential Writeとは、ホストコンピュータ71がディスクの連続した領域に対し、順番にライトアクセスを発行している状態を示す。

【0064】

図8は、本発明の第1の実施形態のディスク稼働記録テーブル33の一例を示す図である。

【0065】

ディスク稼働記録テーブル33は、ディスクドライブ6にアクセスされた時間を管理する。ディスク稼働記録テーブル33は、ディスクドライブ6にアクセスされる毎に更新される。また、電力を推定するときに、電力推定機能24がディスク稼働記録テーブル33を参照する。

10

【0066】

ディスク稼働記録テーブル33は、ディスク番号331、Random Read累計稼働時間332、Random Write累計稼働時間333、Sequential Read累計稼働時間334、及びSequential Write累計稼働時間335を含む。

【0067】

ディスク番号331は、ディスクドライブ6が搭載される物理的な位置示す情報を格納する。図6のディスク番号311と同一のものである。

【0068】

Random Read累計稼働時間332は、ディスク番号331に対応するディスクドライブ6に対してRandom Readアクセスが発行されていた時間の累計時間を格納する。

20

【0069】

Random Write累計稼働時間333は、ディスク番号331に対応するディスクドライブ6に対してRandom Writeアクセスが発行された時間の累積時間を格納する。

【0070】

Sequential Read累計稼働時間334は、ディスク番号331に対応するディスクドライブ6に対してSequential Readアクセスが発行された時間の累積時間を格納する。

30

【0071】

Sequential Write累計稼働時間335は、ディスク番号331に対応するディスクドライブ6に対してSequential Writeアクセスが発行された時間の累積時間を格納する。

【0072】

図9は、本発明の第1の実施形態のTag管理テーブルの34の一例を示す図である。

【0073】

Tag管理テーブル34は、各ディスクドライブ6に対し発行されているアクセスを識別するための識別子であるTagを管理する。これによって、一つのディスクドライブ6に対して多重にアクセスが発行されている場合でも、応答がどのアクセス要求に対するものであるかを識別できる。

40

【0074】

Tag管理テーブル34は、ディスク番号341、Tag番号342、及びアクセス種別343を含む。

【0075】

ディスク番号341は、ディスクドライブ6が搭載される物理的な位置示す情報を格納する。図6のディスク番号311と同一のものである。

【0076】

Tag番号342は、ディスクドライブ6に対し同時に発行されている複数のアクセス

50

を一意に識別するための識別子を格納する。T a g 番号 3 4 2 は、制御部 7 5 によってアクセス発行時に指定される。

【 0 0 7 7 】

アクセス種別 3 4 3 は、ディスクドライブ 6 にアクセスすることになった契機を示す情報を格納する。具体的には、ディスクドライブ 6 に発行されたアクセスの種別が格納される。

【 0 0 7 8 】

なお、T a g 管理テーブル 3 4 は、アクセス種別 3 4 3 のみを格納し、どのホストコンピュータ 7 1 からのアクセスかは管理していないが、制御部 7 5 は、各ホストコンピュータ 7 1 の I / O を識別するための識別子を付与し、当該識別子と T a g 番号 3 4 2 とを対応付けたテーブル（図示省略）を別に備え、当該対応テーブルによって、ホストコンピュータ 7 1 及びホストコンピュータ 7 1 から発行されたアクセスの種別を管理している。

10

【 0 0 7 9 】

図 1 0 は、本発明の第 1 の実施形態の I / O 発行状況テーブル 3 5 の一例を示す図である。

【 0 0 8 0 】

I / O 発行状況テーブル 3 5 は、各ディスクドライブ 6 に対し、発行されたアクセスの種別、及びアクセスが発行された時刻を管理する。I / O 発行状況テーブル 3 5 は、ディスク I / O 発行機能 2 2 がディスクドライブ 6 へのアクセスを新規に発行したとき、または、ディスク I / O 完了処理機能 2 3 がディスクドライブ 6 からの応答を処理したときに参照され、また、更新される。

20

【 0 0 8 1 】

I / O 発行状況テーブル 3 5 は、ディスク番号 3 5 1、アクセス種別 3 5 2、現在発行中コマンド数 3 5 3、及びディスク稼働開始時刻 3 5 4 を含む。

【 0 0 8 2 】

ディスク番号 3 5 1 は、ディスクドライブ 6 が搭載される物理的な位置示す情報を格納する。図 6 のディスク番号 3 1 1 と同一のものである。

【 0 0 8 3 】

アクセス種別 3 5 2 は、ディスクドライブ 6 にアクセスすることになった契機を示す情報を格納する。図 9 のアクセス種別 3 4 3 と同一のものである。

30

【 0 0 8 4 】

現在発行中コマンド数 3 5 3 は、ディスク番号 3 5 1 に対応するディスクドライブ 6 に、現在、アクセス種別 3 5 2 に対応するアクセスが現在いくつ発行されているかを示す値を格納する。

【 0 0 8 5 】

ディスク稼働開始時刻 3 5 4 は、ディスク番号 3 5 1 に対応するディスクドライブ 6 に、アクセス種別 3 5 2 に対応するアクセスがいつから発行されているかを示す値を格納する。具体的には、ストレージシステム 1 が稼働してからカウントされる、ストレージシステム 1 の内部時間が格納される。

【 0 0 8 6 】

図 1 1 は、本発明の第 1 の実施形態のコントローラ部構成テーブル 3 6 の一例を示す図である。

40

【 0 0 8 7 】

コントローラ部構成テーブル 3 6 は、ストレージシステム 1 に搭載されている、キャッシュメモリ部 7 3 または制御部 7 5 などの構成部の具体的な種類、及び構成部の物理的な搭載位置を管理する。コントローラ部構成テーブル 3 6 は、部品を増設または減設するときに更新される。また、また、電力を推定するときに、電力推定機能 2 4 がコントローラ部構成テーブル 3 6 を参照する。

【 0 0 8 8 】

コントローラ部構成テーブル 3 6 は、種別 3 6 1、スロット番号 3 6 2、及び部品名 3

50

63を含む。

【0089】

種別361は、制御部75やキャッシュメモリ部73などのストレージシステム1の構成部の種別を格納する。スロット番号362は、各構成部の部品を増設する場合に使用されるスロットの番号を格納する。スロットの数は、ストレージシステム1に予め備えられた数だけ存在する。

【0090】

部品名363は、スロットに搭載されている各構成部の部品の具体的な種別を示す情報を格納する。なお、任意のスロット番号362に対応するスロットに構成部の部品が搭載されていない場合は、「無し」が格納される。ただし、搭載されていないことを示す情報であれば、他の情報を格納してもよい。

10

【0091】

コントローラ部構成テーブル36を備えることによって、ストレージシステム1の具体的な装置構成が分かる。

【0092】

図12は、本発明の第1の実施形態のコントローラ部電力仕様テーブル37の一例を示す図である。

【0093】

コントローラ部電力仕様テーブル37は、ストレージシステム1に搭載されている各種部品の定常状態（Idle状態）における消費電力の情報を管理する。コントローラ部電力仕様テーブル37は、新規種別の部品が追加されたときに更新される。また、電力を推定するときに、電力推定機能24がコントローラ部電力仕様テーブル37を参照する。

20

【0094】

コントローラ部電力仕様テーブル37は、種別371、部品名372、及びIdle時電力373を含む。

【0095】

種別371は、制御部75またはキャッシュメモリ部73などのストレージシステム1の構成部の種別を格納する。種別371は、図11の種別361と同一のものである。

【0096】

部品名372は、性能、または容量が異なる部品を区別するため部品の型番など、電力に関する仕様が異なるもの識別できる情報を格納する。部品名372は、図11の部品名363と同一のものである。

30

【0097】

Idle時電力373は、通電されている部品が特に稼働していないとき（Idle時）の消費電力を格納する。

【0098】

図13は、本発明の第1の実施形態のコントローラ部電力変動仕様テーブル38の一例を示す図である。

【0099】

コントローラ部電力変動仕様テーブル38は、ストレージシステム1に搭載されている各構成部の稼働状況における電力の変動を管理する。

40

【0100】

コントローラ部電力変動仕様テーブル38は、種別381、部品名382、及び電力増分383を含む。

【0101】

種別381は、制御部75やキャッシュメモリ部73などのストレージシステム1の構成部の種別を格納する。種別381は、図11の種別361と同一のものである。

【0102】

部品名382は、性能、または容量が異なる部品を区別するための部品の型番など、電力に関する仕様が異なるもの識別できる情報を格納する。部品名382は、図11の部品

50

名 3 6 3 と同一のものである。

【 0 1 0 3 】

電力増分 3 8 3 は、該当する部品の最大稼働時における消費電力の増加分を格納する。

【 0 1 0 4 】

コントローラ部電力変動仕様テーブル 3 8 は、消費電力が変動するストレージシステム 1 の構成部に関する情報のみを格納してもよいし、ストレージシステム 1 の全ての構成部に関する情報を格納してもよい。ストレージシステム 1 の全ての構成部に関する情報が格納される場合、消費電力が変動しない部品の電力増分 3 8 3 は、「 0 」が格納される。

【 0 1 0 5 】

図 1 4 は、本発明の第 1 の実施形態のプロセッサ稼働記録テーブル 3 9 の一例を示す図である。

【 0 1 0 6 】

プロセッサ稼働記録テーブル 3 9 は、制御部 7 5 が備えるプロセッサ 7 8 の稼働時間を管理する。プロセッサ稼働記録テーブル 3 9 は、プロセッサ 7 8 上で動作するスケジューラによってジョブ、即ち、ストレージ制御プログラム 2 が備える、ホスト I / O 処理機能、またはディスク I / O 処理機能などの I / O 処理機能が実行されているときに更新される。また、電力を推定するときに、電力推定機能 2 4 がプロセッサ稼働記録テーブル 3 9 を参照する。

【 0 1 0 7 】

プロセッサ稼働記録テーブル 3 9 は、制御部番号 3 9 1、プロセッサコア番号 3 9 2、及び累計稼働時間 3 9 3 を含む。

【 0 1 0 8 】

制御部番号 3 9 1 は、ストレージシステム 1 内が備える複数の制御部 7 5 を識別するための識別子を格納する。

【 0 1 0 9 】

プロセッサコア番号 3 9 2 は、制御部 7 5 が備えるプロセッサ 7 8 の内部に存在する複数のプロセッサコア 7 8 1 を識別する識別子を格納する。

【 0 1 1 0 】

累計稼働時間 3 9 3 は、プロセッサコア番号 3 9 2 に対応するプロセッサコア 7 8 1 上でジョブが実行されていた時間の累計を格納する。

【 0 1 1 1 】

図 1 5 は、本発明の第 1 の実施形態のプロセッサ稼働状況テーブル 3 0 の一例を示す図である。

【 0 1 1 2 】

プロセッサ稼働状況テーブル 3 0 は、ストレージ制御プログラム 2 のジョブが実行されているか否か、及びストレージ制御プログラム 2 のジョブが開始された時間を管理する。

【 0 1 1 3 】

プロセッサ稼働状況テーブル 3 0 は、制御部番号 3 0 1、プロセッサコア番号 3 0 2、及びジョブ開始時刻 3 0 3 を含む。

【 0 1 1 4 】

制御部番号 3 0 1 は、ストレージシステム 1 内が備える複数の制御部 7 5 を識別するための識別子を格納する。制御部番号 3 0 1 は、図 1 4 の制御部番号 3 9 1 と同一のものである。

【 0 1 1 5 】

プロセッサコア番号 3 0 2 は、制御部 7 5 が備えるプロセッサ 7 8 の内部に存在する複数のプロセッサコア 7 8 1 を識別するための識別子を格納する図 1 4 のプロセッサコア番号 3 9 2 と同一のものである。

【 0 1 1 6 】

ジョブ開始時刻 3 0 3 は、現在ストレージ制御プログラム 2 のジョブが実行されている

10

20

30

40

50

か否か、及びストレージ制御プログラム2のジョブが開始された時間を格納する。具体的には、ストレージ制御プログラム2のジョブが実行されていない場合、ジョブ開始時刻303には何も格納されない。また、ストレージ制御プログラム2のジョブが実行されている場合、ジョブ開始時刻303にはストレージ制御プログラム2のジョブが開始された時間を格納する。なお、格納される時間は、ストレージシステム1の内部時間である。

【0117】

以上が制御データ3が備える各テーブルの説明である。以下、ストレージ制御プログラム2が備える各機能について説明する。

【0118】

図16は、第1の実施形態のホストI/O受領機能21の一例を説明するフローチャートである。

10

【0119】

ホストI/O受領機能21は、ホストコンピュータ71からリード要求のコマンド、またはライト要求のコマンドを受信したときに呼び出され、受信したコマンドの種別を判定し、判別されたコマンドの種別に応じた処理を実行する。

【0120】

まず、ホストI/O受領機能21は、ホストコンピュータ71からのアクセス要求のアクセス種別を解析する(S2101)。具体的には、アクセス種別は、Random Read、Random Write、Sequential Read、またはSequential Writeのいずれかである。

20

【0121】

ホストI/O受領機能21は、解析されたコマンドの種別が前述のいずれに該当するかを判定する(S2102)。

【0122】

アクセス種別がRandom Readであると判定された場合、ホストI/O受領機能21は、キャッシュメモリ部73に、要求されたデータが格納されているか否かを判定する(S2111)。

【0123】

要求されたデータが格納されていないと判定された場合、ホストI/O受領機能21は、ディスクI/O発行機能22を起動する(S2112)。このとき、ホストI/O受領機能21は、アクセス時間を集計するため、ディスクアクセスを発行する契機をパラメータとしてディスクI/O発行機能22に与える。具体的には、ホストI/O受領機能21は、契機をRandom Read、アクセスの転送元をホストから要求されたデータが存在するディスク上のアドレス、転送先をキャッシュメモリ部73上の空き領域をパラメータとして、ホストI/O受領機能21はディスクI/O発行機能22に与えて、ディスクI/O発行機能22を起動し、S2113に進む。

30

【0124】

S2111において、要求されたデータが格納されていないと判定された場合、ホストI/O受領機能21は、S2113に進む。

【0125】

ホストI/O受領機能21は、要求されたデータをホストコンピュータ71への応答として送信され(S2113)、処理を終了する。

40

【0126】

S2102において、アクセス種別がRandom Writeであると判定された場合、ホストI/O受領機能21は、ホストコンピュータ71から送られたデータをキャッシュメモリ部73へ書き込む(S2121)。

【0127】

次に、ホストI/O受領機能21は、ホストコンピュータ71に対し、ライト完了の応答を送信する(S2122)。

【0128】

50

次に、ホストI/O受領機能21は、キャッシュメモリ部73の空き領域が十分であるか否かを判定する(S2123)。

【0129】

キャッシュメモリ部73の空き領域が十分であると判定された場合、ホストI/O受領機能21は、処理を終了する。

【0130】

キャッシュメモリ部73の空き領域が十分でないとして判定された場合、ホストI/O受領機能21は、ホストコンピュータ71からのライトを受けられるようにするため、キャッシュメモリ部73に格納されているデータをディスク6への書き込みを実行する。具体的には、アクセス契機をRandom Writeとし、転送元をキャッシュメモリ部73上の最近アクセスされていない領域とし、転送先をアクセスされていない領域に対応するディスクドライブ6の対応するアドレスをパラメータとして、ホストI/O受領機能21はディスクI/O発行機能22に与え、ディスクI/O発行機能22を起動し(S2124)、処理を終了する。

10

【0131】

S2102において、アクセス種別がSequential Readであると判定された場合、ホストI/O受領機能21は、ホストコンピュータ71からアクセスがあることを予測して要求されたデータをキャッシュメモリ部73に読み出しているため、読み出されているデータをホストコンピュータ71への応答として送信する(S2131)。

【0132】

ホストI/O受領機能21は、ホストコンピュータ71が次の領域へアクセスすることが予測されるため、次の領域のデータをキャッシュメモリ部73に読み出す。具体的には、アクセス契機をSequential Readとし、転送元をホストコンピュータ71が要求した領域の次の領域、転送先をキャッシュメモリ部73上の空き領域をパラメータとして、ホストI/O受領機能21はディスクI/O発行機能22に与え、ディスクI/O発行機能22を起動し(S2132)、処理を終了する。

20

【0133】

S2102において、アクセス種別がSequential Writeであると判定された場合、ホストI/O受領機能21は、ホストコンピュータ71から送られたデータをキャッシュメモリ部73へ書き込む(S2141)。

30

【0134】

次に、ホストI/O受領機能21は、ホストコンピュータ71に対し、ライト完了の応答を送信する(S2142)。

【0135】

Sequential Writeの場合、同じ領域に対して再度書き込みが行われる可能性は低いため、書き込まれたデータをそのままキャッシュメモリ部73に保持しても性能面での効果は期待できない。そのため、ホストI/O受領機能21は、ホストコンピュータ71から書き込まれたデータをディスクドライブ6に退避する。具体的には、アクセス契機をSequential Write、転送元をホストが書いた領域、転送先を対応するディスクドライブ6の対応するアドレスをパラメータとして、ホストI/O受領機能21はディスクI/O発行機能22に与え、ディスクI/O発行機能22を起動し(S2143)、処理を終了する。

40

【0136】

図17は、本発明の第1の実施形態のディスクI/O発行処理機能22の一例を説明するフローチャートである。

【0137】

ディスクI/O発行処理機能22は、ディスクドライブ6へのアクセスを発行する機能であり、ホストI/O受領機能21からの要求に応じて起動する。

【0138】

まず、ディスクI/O発行処理機能22は、呼び出し元であるホストI/O受領機能2

50

1 から与えられたパラメータを解析し、アクセス先のディスクドライブ 6、及びアクセス種別を抽出する (S 2 2 1)。

【0139】

ディスク I / O 発行処理機能 2 2 は、T a g 管理テーブル 3 4 を参照し、使用されていない T a g 番号 3 4 2、即ち、T a g 番号 3 4 2 に対応するアクセス種別 3 4 3 が記入されていない T a g 番号 3 4 2 を選択する (S 2 2 2)。

【0140】

ディスク I / O 発行処理機能 2 2 は、選択された T a g 番号 3 4 2 に対応するエントリのアクセス種別 3 4 3 に、抽出されたアクセス種別を格納する (S 2 2 3)。

【0141】

ディスク I / O 発行処理機能 2 2 は、ディスク I / F 部 7 4 のプロトコル変換 L S I 7 4 1 へ、ディスクアクセスコマンドを送信する (S 2 2 4)。

【0142】

次に、ディスク I / O 発行処理機能 2 2 は、I / O 発行状況テーブル 3 5 を参照し、抽出されたアクセス先のディスクドライブ 6、及びアクセス種別に対応するエントリの現在発行中コマンド数 3 5 3 が「0」であるか否かを判定する (S 2 2 5)。

【0143】

対応するエントリの現在発行中コマンド数 3 5 3 が「0」であると判定された場合、アクセス先のディスクドライブ 6 は稼働していなかったことを示しており、ディスク I / O 発行処理機能 2 2 は、アクセス先のディスクドライブ 6 はこれから稼働するため、対応するエントリのディスク稼働開始時刻 3 5 4 に現在の時刻を格納し (S 2 2 6)、S 2 2 7 に進む。

【0144】

S 2 2 5 において、対応するエントリの現在発行中コマンド数 3 5 3 が「0」でないと判定された場合、既に同種のアクセスが同じディスクドライブ 6 へ発行されている状態、即ち、アクセス先のディスクドライブ 6 は既に稼働中のため、ディスク I / O 発行処理機能 2 2 は、S 2 2 7 に進む。

【0145】

ディスク I / O 発行処理機能 2 2 は、I / O 発行状況テーブル 3 5 の対応するエントリの現在発行中コマンド数 3 5 3 を「1」加算し (S 2 2 7)、処理を終了する。

【0146】

図 1 8 は、本発明の第 1 の実施形態のディスク I / O 完了処理機能 2 3 の一例を説明するフローチャートである。

【0147】

ディスク I / O 完了処理機能 2 3 は、ディスク I / O 発行機能 2 2 が発行したディスクアクセスコマンドに対する応答をディスクドライブ 6 からディスク I / O 発行機能 2 2 が受け取ったときに呼び出される機能であり、アクセスに要した時間を管理する。

【0148】

まず、ディスク I / O 完了処理機能 2 3 は、ディスクドライブ 6 から受信した応答を解析し、T a g 番号 3 4 2 を抽出する (S 2 3 1)。

【0149】

ディスク I / O 完了処理機能 2 3 は、T a g 管理テーブル 3 4 を参照し、抽出された T a g 番号 3 4 2 に対応するエントリのアクセス種別 3 4 3 を読み出す (S 2 3 2)。

【0150】

ディスク I / O 完了処理機能 2 3 は、T a g 管理テーブル 3 4 の前記エントリのアクセス種別 3 4 3 を消去する (S 2 3 3)。

【0151】

次に、ディスク I / O 完了処理機能 2 3 は、I / O 発行状況テーブル 3 5 の対応するエントリの現在発行中コマンド数 3 5 3 を「1」減算する (S 2 3 4)。

【0152】

10

20

30

40

50

次に、ディスクI/O完了処理機能23は、I/O発行状況テーブル35の対応するエントリの現在発行中コマンド数353が「0」であるか否かを判定する(S235)。

【0153】

対応するエントリの現在発行中コマンド数353が「0」と判定された場合、当該アクセス種別の該当するアクセスは1つも発行されていないこと、つまり、ディスクドライブ6へのアクセスが終了したことを示しており、ディスクI/O完了処理機能23は、これまでのアクセス時間を稼働情報として記録する。即ち、ディスクI/O完了処理機能23は、I/O発行状況テーブル35の対応するエントリのディスク稼働開始時刻354と現在時刻との差分を算出し、当該アクセス種別に対応する累計稼働時間(332~335)に、算出された時間を加算し(S236)、処理を終了する。

10

【0154】

S235において、対応するエントリの現在発行中コマンド数353が「0」でないと判定された場合、まだ他に同一アクセス種別のアクセスが実行中であるため、処理を終了する。

【0155】

図19は、本発明の第1の実施形態のジョブ起動処理機能25の一例を説明する図である。

【0156】

ジョブ起動処理機能25は、ホストコンピュータ71からのコマンドまたはディスクドライブ6からの応答を受け取った場合、受信したコマンドまたは応答から要求された処理を実行するプロセッサコア781によって呼び出され、当該処理に応じた機能(ジョブ)を呼び出す。

20

【0157】

まず、ジョブ起動処理機能25は、プロセッサ稼働状況テーブル30を参照し、要求された処理を実行するプロセッサコア781を含む制御部75を識別するための制御部番号301、及び当該プロセッサコア781を識別するプロセッサコア番号302に対応したエントリを確認する。ジョブ起動処理機能25は、該当するエントリのジョブ開始時刻303に、現在の時刻を記入する(S251)。その後、要求された処理に応じたジョブを起動(S252)し、処理を終了する。

【0158】

図20は、本発明の第1の実施形態のジョブ終了処理機能26の一例を説明する図である。

30

【0159】

ジョブ終了処理機能26は、前述のジョブが終了したプロセッサコア781によって呼び出され、処理に要した時間の集計を実行する。

【0160】

まず、ジョブ終了処理機能26は、プロセッサ稼働状況テーブル30を参照し、要求された処理を実行しているプロセッサコア781を含む制御部75を識別するための制御部番号301、及び当該プロセッサコア781を識別するプロセッサコア番号302に対応したエントリを読み出す。次に、ジョブ終了処理機能26は、読み出された情報と一致するエントリのジョブ開始時刻303を参照し、参照されたジョブ開始時刻303と現在時刻との差分を算出する(S261)。

40

【0161】

ジョブ終了処理機能26は、算出された差分を、プロセッサ稼働記録テーブル39の対応するエントリの累計稼働時間393に加算する(S262)。

【0162】

ジョブ終了処理機能26は、プロセッサ稼働状況テーブル30の対応するエントリのジョブ開始時刻303の値を消去し(S263)、処理を終了する。

【0163】

図21は、本発明の第1の実施形態の電力推定機能24の一例を説明するフローチャー

50

トである。

【0164】

電力推定機能24は、ディスクドライブ6、制御部75、及びキャッシュメモリ部73の稼働率を算出し、各テーブルの情報と算出された稼働率から消費電力を推定し、推定された消費電力の情報を管理用端末4に送信する。電力推定機能24は、一定の周期で呼び出される。

【0165】

まず、電力推定機能24は、ディスク部構成テーブル31とディスク電力仕様テーブル32とからディスク部消費電力分を算出する(S241)。具体的には、電力推定機能24は、ディスク部構成テーブル31のエントリのディスク型番312を参照し、ディスク電力仕様テーブル32において、参照されたディスク型番312に対応するエントリのIdle時電力322をディスク部消費電力分とする。

10

【0166】

次に、電力推定機能24は、ディスク稼働記録テーブル33の情報と本処理が呼び出される周期との比を用いて、各ディスクドライブ6についてアクセス種別毎に稼働率を算出する(S242)。具体的には、稼働率は、アクセス種別の累計稼働時間(332~335)と本処理が呼び出される周期とを用いて算出される。例えば、本処理の呼び出し間隔が「1秒」で、ディスク稼働記録テーブル33のRandom Read累計稼働時間332が「1ミリ秒」である場合、Random Readにおける稼働率は、「1/1000」となる。

20

【0167】

次に、電力推定機能24は、算出された稼働率とディスク電力仕様テーブル32の情報とから、各ディスクドライブ6の電力増加分を算出する(S243)。例えば、Random Readにおける稼働率が「1/10」で、Random Read時電力増分323が「4.9W」である場合、電力増加は「0.49W」となる。算出式については、図22を用いて後述する。

【0168】

次に、電力推定機能24は、コントローラ部構成テーブル36とコントローラ部電力仕様テーブル37とから、コントローラ部の消費電力分を算出する(S2401)。具体的には、種別361毎の消費電力分が算出される。例えば、図11及び図12に示す例において、制御部75のスロット0及びスロット1には、各々「C-1」と「C-2」とが搭載されており、当該「C-1」のIdle時電力373が「85W」であり、また、当該「C-2」のIdle時電力3073が「65W」であり、他のスロットには何にも搭載されていない場合、ストレージシステム1における制御部75の消費電力は、「85W+65W=150W」と算出される。算出された値がコントローラ部における制御部75の消費電力分となる。他の構成部についても同様の方法を用いて算出される。

30

【0169】

次に、電力推定機能24は、キャッシュメモリ部73のアクセス時間記録テーブル732の内容を読み出す(S2402)。

【0170】

電力推定機能24は、読み出されたアクセス時間7322と本処理が呼び出される周期との比を用いて、キャッシュメモリ部73の稼働率を算出する(S2403)。例えば、本処理の呼び出し周期が「1秒」で、読みだされたアクセス時間7322が「10ミリ秒」である場合、キャッシュメモリ部73の稼働率は、「1/100」となる。

40

【0171】

電力推定機能24は、算出されたキャッシュメモリ部73の稼働率とコントローラ部電力変動仕様テーブル38の内容とから、キャッシュメモリ部73の電力増加分を算出する(S2404)。例えば、図12に示す例において、キャッシュメモリ部73の稼働率が「1/10」であり、部品名382が「M-2」の電力増分383が「5.3W」であり、当該メモリモジュール733の電力増分は、「1/10×5.3W=0.53W」と算

50

出される。

【0172】

次に、電力推定機能24は、プロセッサ稼働記録テーブル39と本処理が呼び出される周期との比を用いて、各プロセッサコア781の稼働率を算出する(S2405)。算出方法は、S242と同様である。

【0173】

次に、電力推定機能24は、算出されたプロセッサコア781の稼働率から、制御部75の稼働率を算出する(S2406)。具体的には、制御部75が備える各プロセッサコア781の稼働率の平均を算出し、算出されたプロセッサコア781の稼働率の平均を制御部75の稼働率とする。例えば、制御部75が4つのプロセッサコア781を備え、それぞれの稼働率が「20%」、「40%」、「60%」、「80%」であった場合、制御部75の稼働率は「50%」と算出される。

10

【0174】

次に、電力推定機能24は、算出された制御部75の稼働率とコントローラ部電力変動仕様テーブル38の内容とから、制御部75の電力増加分を算出する(S2407)。例えば、コントローラ部電力変動仕様テーブル38の電力増分383が「42W」であり、稼働率が「50%」である場合、制御部75の電力増加分は、「42W×0.5=21W」と算出される。

【0175】

電力推定機能24は、以上の処理によって算出された各構成部の消費電力及び電力増加分を管理用端末4に送信する(S244)。

20

【0176】

電力推定機能24は、ディスク稼働記録テーブル33の累計稼働時間(332~335)を初期化する(S245)。

【0177】

次に、キャッシュメモリ部73のアクセス時間記録テーブル732のアクセス時間7322を初期化する(S2408)。

【0178】

電力推定機能24は、プロセッサ稼働記録テーブル39の累計稼働時間393を初期化し(S2409)、処理を終了する。

30

【0179】

なお、前述の処理では、消費電力を算出したが、消費エネルギーを算出してよい。具体的には、算出された消費電力に稼働時間を乗算することによって消費エネルギーを算出することができる。また、算出された消費エネルギーは、管理用端末4に送信されてもよい。

【0180】

図22は、本発明の第1の実施形態の電力推定機能24がS241及びS243におけるディスク部電力増加分を算出方法の一例を示す図である。

【0181】

算出値1001は、図21における各ステップで算出される電力を示す情報である。具体的には、P__Static、P__RndRd、P__RadWr、P__SeqRd、P__SeqWr、及びP__Dynamicを格納する。

40

【0182】

P__Staticは、任意のディスク型番312における定常時の消費電力である。P__RndRdは、任意のディスク型番312におけるRandom Readのアクセス時の電力増加分である。P__RadWrは、任意のディスク型番312におけるRandom Writeのアクセス時の電力増加分である。P__SeqRdは、任意のディスク型番312におけるSequential Readのアクセス時の電力増加分である。P__SeqWrは、任意のディスク型番312におけるSequential Writeのアクセス時の電力増加分である。P__Dynamicは、任意のディスク型番301

50

2における全てのアクセスを含めた時の電力増加分である。

【0183】

算出式1002は、算出値1001の具体的な算出式を示す。以下各算出値1001における算出式1002について説明する。

【0184】

P_Static は、以下の式で算出される。

(式1) $P_Static = P_Idle$

ここで、 P_Idle は、 $Idle$ 時の消費電力である。

【0185】

P_RndRd は、以下の式で算出される。

(式2) $P_RndRd = AR_RndRd \times p_RndRd$

ここで、 AR_RndRd は $Random\ Read$ アクセスの稼働率であり、 p_RndRd は $Random\ Read$ 時電力増分323である。

【0186】

P_RadWr は、以下の式で算出される。

(式3) $P_RadWr = AR_RadWr \times p_RadWr$

ここで、 AR_RadWr は $Random\ Write$ アクセスの稼働率であり、 p_RndRd は $Random\ Write$ 時電力増分324である。

【0187】

P_SeqRd は、以下の式で算出される。

(式4) $P_SeqRd = AR_SeqRd \times p_SeqRd$

ここで、 AR_SeqRd は $Sequential\ Read$ アクセスの稼働率であり、 p_SeqRd は $Sequential\ Read$ 時電力増分325である。

【0188】

P_SeqWr は、以下の式で算出される。

(式5) $P_SeqWr = AR_SeqWr \times p_SeqWr$

ここで、 AR_SeqWr は $Sequential\ Write$ アクセスの稼働率であり、 p_SeqWr は $Sequential\ Write$ 時電力増分326である。

【0189】

$P_Dynamic$ は、以下の式で算出される。

(式6) $P_Dynamic = P_RndRd + P_RadWr + P_SeqRd + P_SeqWr$

以上で説明した式を用いて、各ディスク部の電力増加分が算出される。

【0190】

図23は、本発明の第1の実施形態の電力推定機能24が実行する処理のS244における管理用端末4に送信される表示データ990の一例を示す図である。

【0191】

表示データ990は、部位991、平均消費電力992、電力増加993、及び $Idle$ 時電力994を含む。

【0192】

部位991は、ストレージシステム1の構成部、または、部品を示す情報を格納する。例えば、種別361、及び部品名363の情報を格納する。なお、構成部、または部品を区別できる情報であれば、どのような情報であってもよい。

【0193】

平均消費電力992は、部位991の平均消費電力を示す情報を格納する。ただし、平均消費電力は、算出された電力増分と $Idle$ 時電力994との合計値を示す値である。

【0194】

電力増加993は、ディスクドライブ6またはコントローラ部の電力増加量を示す情報を格納する。部位991がディスクドライブ6を示す情報の場合、各アクセス種別の電力増加に情報が格納される。また、部位991がコントローラ部の場合、コントローラ部に

10

20

30

40

50

電力増加に情報が格納される。

【0195】

Idle時電力994は、ストレージシステム1の各構成部または部品が処理をしていないときの電力を示す情報を格納する。

【0196】

図24は、本発明の第1の実施形態のストレージ管理プログラム5が備える機能及びテーブルを説明するブロック図である。

【0197】

ストレージ管理プログラム5は、消費電力履歴テーブル51、及び消費電力更新機能52を備える。

【0198】

消費電力履歴テーブル51は、ストレージシステム1から送られた電力データを管理する。消費電力更新機能52は、ストレージシステム1から送られた電力データに基づいて、消費電力履歴テーブル51を更新する。

【0199】

図25は、本発明の第1の実施形態の消費電力履歴テーブル51の一例を示す図である。

【0200】

消費電力履歴テーブル51は、管理用端末4のストレージ管理プログラム5に含まれ、ストレージシステム1から送られた電力データを管理し、管理用端末利用者からの要求に応じてその内容を表示するために用いられる。

【0201】

消費電力履歴テーブル51は、日時511、部位512、平均消費電力513、電力増加514、及びIdle時電力515を含む。

【0202】

日時511は、格納されているデータが、いつの情報であることを示す情報を格納する。部位512、平均消費電力513、電力増加514、及びIdle時電力515は、図22の部位991、平均消費電力992、電力増加993、及びIdle時電力994と同一のものである。

【0203】

以上説明したように、第1の実施形態によれば、ストレージシステム1内のすべての構成部、即ちディスク部とコントローラ部との稼働情報教に応じた消費電力について、より正確な値を算出することができるようになる。また、管理者は、算出された消費電力の情報を知ることが可能になる。その結果、より精度の良い電力制御ができる。

【0204】

[第2の実施形態]

本発明の第2の実施形態は、ストレージシステム1が複数の仮想ストレージ(LPAR)として扱うことができる場合、それぞれの仮想ストレージの消費電力を推定する発明である。

【0205】

第2の実施形態の計算機システムの構成は、第1の実施形態と同様のものであり、説明は省略する。

【0206】

第2の実施形態では、ストレージシステム1が備える物理資源を分割し、分割された物理資源から仮想ストレージが構成される。仮想ストレージは、ホストコンピュータ71からは一つのストレージシステム1として認識される。ディスクドライブ6は、論理ボリュームに分割されている。各テーブルは論理ボリューム毎の情報が格納される。電力推定機能24は、論理ボリューム毎に推定消費電力を算出し、算出された推定消費電力に基づいて、仮想ストレージの消費電力を推定することができる。

【0207】

10

20

30

40

50

図 26 は、本発明の第 2 の実施形態のストレージ管理プログラム 5 が備える機能及びテーブルを説明するブロック図である。

【0208】

第 2 の実施形態において、ストレージ管理プログラム 5 は、新たに、仮想ストレージ - ボリューム対応テーブル 53 と仮想ストレージ電力表示機能 54 とを備える。

【0209】

仮想ストレージ - ボリューム対応テーブル 53 は、仮想ストレージと論理ボリュームとの対応を管理する。仮想ストレージ電力表示機能 54 は、仮想ストレージ毎の電力推定の結果を表示する。

【0210】

図 27 は、本発明の第 2 の実施形態の仮想ストレージ - ボリューム対応テーブル 53 の一例を示す図である。

【0211】

仮想ストレージ - ボリューム対応テーブル 53 は、仮想ストレージ (L P A R) 番号 531、及び論理ボリューム番号 532 を管理する。

【0212】

仮想ストレージ番号 531 は、ストレージシステム 1 内に構成されている仮想ストレージを識別するための識別子を格納する。

【0213】

論理ボリューム番号 532 は、ディスクドライブ 6 上に作成されている論理ボリュームを識別するための識別子を格納する。

【0214】

ストレージシステム 1 は、論理ボリューム番号 532 から、どの仮想ストレージにおける処理であるかが分かる。

【0215】

図 28 は、本発明の第 2 の実施形態のストレージ制御プログラム 2 が備える機能を説明するブロック図である。

【0216】

第 1 の実施形態と比べ、第 2 の実施形態のストレージ制御プログラム 2 は、新たにデータ転送起動機能 27 を備える。データ転送起動機能 27 は、論理ボリュームの操作のためにキャッシュメモリ部 73 にアクセスされたデータ量を管理するための機能である。詳細については、図 41 を用いて後述する。

【0217】

図 29 は、本発明の第 2 の実施形態の制御データ 3 の構成を説明するブロック図である。

【0218】

第 1 の実施形態と比べ、第 2 の実施形態の制御データ 3 は、新たにキャッシュアクセス量テーブル 91、及び L D E V - 物理ボリューム対応テーブル 92 を備える。

【0219】

また、第 2 の実施形態におけるディスク稼働記録テーブル 33、T a g 管理テーブル 34、I / O 発行状況テーブル 35、プロセッサ稼働記録テーブル 39、及びプロセッサ稼働状況テーブル 30 に格納される情報も第 1 の実施形態のものとは異なる。

【0220】

まず、L D E V - 物理ボリューム対応テーブル 92 について説明する。

【0221】

図 30 は、本発明の第 2 の実施形態の L D E V - 物理ボリューム対応テーブル 92 の一例を示す図である。

【0222】

L D E V - 物理ボリューム対応テーブル 92 は、ディスクドライブ 6 と、一以上のディスクドライブ 6 から構成される R A I D グループ上に作成される論理ボリューム (L D E

10

20

30

40

50

V (L o g i c a l D e v i c e)) との対応を管理する。

【 0 2 2 3 】

L D E V - 物理ボリューム対応テーブル 9 2 は、論理ボリューム番号 9 2 1、R A I D レベル 9 2 2、ディスク番号 9 2 3、及び L B A 範囲 9 2 4 を含む。

【 0 2 2 4 】

論理ボリューム番号 9 2 1 は、論理ボリュームを識別するための識別子を格納する。

【 0 2 2 5 】

R A I D レベル 9 2 2 は、構成される R A I D の種別を示す情報を格納する。

【 0 2 2 6 】

ディスク番号 9 2 3 は、R A I D を構成するディスクドライブ 6 の識別子を格納する。

10

【 0 2 2 7 】

L B A 範囲 9 2 4 は、ディスクドライブ 6 に割り当てられているアドレスを格納する。

【 0 2 2 8 】

以下、ディスク稼働記録テーブル 3 3、T a g 管理テーブル 3 4、I / O 発行状況テーブル 3 5、プロセッサ稼働記録テーブル 3 9、及びプロセッサ稼働状況テーブル 3 0 の差異について説明する。

【 0 2 2 9 】

図 3 1 は、本発明の第 2 の実施形態のディスク稼働記録テーブル 3 3 の一例を示す図である。

【 0 2 3 0 】

20

ディスク稼働記録テーブル 3 3 は、第 1 の実施形態と比べ、新たに論理ボリューム番号 3 3 6 を含む。

【 0 2 3 1 】

論理ボリューム番号 3 3 6 は、ディスクドライブ 6 上に作成された論理ボリュームを識別するための識別子を格納する。これによって、ディスク稼働記録テーブル 3 3 は、各論理ボリュームのアクセス種別毎の稼働累計時間 (3 3 2 ~ 3 3 5) を管理できる。

【 0 2 3 2 】

なお、論理ボリュームを識別するための識別子は、論理ボリュームが作成されるときに、管理用端末 4 から指定された識別子である。

【 0 2 3 3 】

30

図 3 2 は、本発明の第 2 の実施形態の T a g 管理テーブル 3 4 の一例を示す図である。

【 0 2 3 4 】

T a g 管理テーブル 3 4 は、第 1 の実施形態と比べ、新たに論理ボリューム番号 3 4 4 を含む。論理ボリューム番号 3 4 4 は、図 3 1 の論理ボリューム番号 3 3 6 と同一のものである。これによって、T a g 管理テーブル 3 4 は、T a g 番号 3 4 2 毎に論理ボリューム番号 3 4 4 が対応付けられており、T a g 番号 3 4 2 に対応するアクセスが、どの論理ボリュームからのアクセスであるかを管理できる。

【 0 2 3 5 】

図 3 3 は、本発明の第 2 の実施形態の I / O 発行状況テーブル 3 5 の一例を示す図である。

40

【 0 2 3 6 】

I / O 発行状況テーブル 3 5 は、第 1 の実施形態と比べ、新たに論理ボリューム番号 3 5 5 を含む。論理ボリューム番号 3 5 5 は、図 3 1 の論理ボリューム番号 3 3 6 と同一のものである。これによって、I / O 発行状況テーブル 3 5 は、各論理ボリュームのアクセス種別毎に現在発行中コマンド数 3 5 3、及びディスク稼働開始時刻 3 5 4 を管理できる。

【 0 2 3 7 】

図 3 4 は、本発明の第 2 の実施形態のプロセッサ稼働記録テーブル 3 9 の一例を示す図である。

【 0 2 3 8 】

50

プロセッサ稼働記録テーブル39は、第1の実施形態と比べ、新たに論理ボリューム番号394を含む。論理ボリューム番号394は、図31の論理ボリューム番号336と同一のものである。これによって、プロセッサ稼働記録テーブル39は、プロセッサコア781がどの論理ボリュームのジョブを実行するために稼働していた時間を管理できる。

【0239】

図35は、本発明の第2の実施形態のプロセッサ稼働状況テーブル30の一例を示す図である。

【0240】

プロセッサ稼働状況テーブル30は、第1の実施形態と比べ、新たに論理ボリューム番号304を含む。論理ボリューム番号304は、図31の論理ボリューム番号336と同一のものである。これによって、プロセッサ稼働状況テーブル30は、論理ボリューム毎にプロセッサコア781のジョブ開始時刻303を管理できる。

10

【0241】

次に、キャッシュアクセス量テーブル91について説明する。

【0242】

図36は、本発明の第2の実施形態のキャッシュアクセス量テーブル91の一例を示す図である。

【0243】

キャッシュアクセス量テーブル91は、各論理ボリュームのキャッシュメモリアクセスマンを管理する。また、論理ボリューム毎の消費電力を推定するときに、電力推定機能24がキャッシュアクセス量テーブル91を参照する。

20

【0244】

キャッシュアクセス量テーブル91は、論理ボリューム番号911、及び累計アクセス量912を含む。

【0245】

論理ボリューム番号911は、図31の論理ボリューム番号336と同一のものである。

【0246】

累計アクセス量912は、論理ボリュームの操作のためにキャッシュメモリ部73にアクセスされたデータ量を格納する。

30

【0247】

図37は、本発明の第2の実施形態のディスクI/O発行処理機能22の一例を示す図である。以下、第1の実施形態と異なるステップを中心に説明する。

【0248】

S221～S223までの処理は、第1の実施形態と同様である。なお、ホストI/O受領機能21から渡されたパラメータには論理ボリュームの識別子が含まれる。S223後、新たにS228の処理が実行される。

【0249】

S228において、ディスクI/O発行処理機能22は、Tag管理テーブル34を参照し、S222において選択されたTag番号342に対応するエントリの、論理ボリューム番号344に、抽出された論理ボリュームの識別子を格納する。

40

【0250】

S224の処理は、第1の実施形態と同様である。S225における処理が第1の実施形態と異なる。

【0251】

S225において、ディスクI/O発行処理機能22は、I/O発行状況テーブル35を参照し、抽出されたアクセス先のディスクドライブ6、論理ボリューム番号355、及びアクセス種別に対応するエントリの現在発行中コマンド数353が「0」であるか否かを判定する。

【0252】

50

S 2 2 6 ~ S 2 2 7 までの処理は、第 1 の実施形態と同様である。

【 0 2 5 3 】

図 3 8 は、本発明の第 2 の実施形態のディスク I / O 完了処理機能 2 3 の一例を示す図である。以下、第 1 の実施形態と異なるステップを中心に説明する。

【 0 2 5 4 】

S 2 3 1 ~ S 2 3 3 までの処理は、第 1 の実施形態と同様である。S 2 3 3 後、新たに S 2 3 7 及び S 2 3 8 の処理が実行される。

【 0 2 5 5 】

S 2 3 7 において、ディスク I / O 完了処理機能 2 3 は、T a g 管理テーブル 3 4 を参照し、抽出された T a g 番号 3 4 2 に対応するエントリの論理ボリューム番号 3 4 4 を読み出す。また、S 2 2 8 において、ディスク I / O 完了処理機能 2 3 は、前記エントリの論理ボリューム番号 3 4 4 を消去する。

10

【 0 2 5 6 】

S 2 3 4 ~ S 2 3 6 までの処理は、第 1 の実施形態と同様である。ただし、抽出された論理ボリューム番号 3 4 4 に基づいて処理が実行される。

【 0 2 5 7 】

図 3 9 は、本発明の第 2 の実施形態のジョブ起動処理機能 2 5 の一例を示す図である。以下、第 1 の実施形態と異なるステップを中心に説明する。

【 0 2 5 8 】

S 2 5 1 及び S 2 5 2 の処理は、第 1 の実施形態と同様である。第 1 の実施形態と比べ、新たに、S 2 5 3 及び S 2 5 4 の処理が実行される。

20

【 0 2 5 9 】

S 2 5 3 において、ジョブ起動処理機能 2 5 は、起動されるジョブがどの論理ボリュームからの要求であるかを判定する。なお、該要求には、論理ボリュームの識別子が含まれており、ジョブ起動処理機能 2 5 は、当該論理ボリュームの識別子を参照して、起動されるジョブがどの論理ボリュームからの要求であるかを判定する。

【 0 2 6 0 】

S 2 5 4 において、ジョブ起動処理機能 2 5 は、S 2 5 1 でジョブ開始時刻 3 0 3 が格納されたエントリの論理ボリューム番号 3 0 4 に、S 2 5 3 で判定された論理ボリュームの識別子を格納する。

30

【 0 2 6 1 】

図 4 0 は、本発明の第 2 の実施形態のジョブ終了処理機能 2 6 の一例を示す図である。以下、第 1 の実施形態と異なるステップを中心に説明する。

【 0 2 6 2 】

S 2 6 1 及び S 2 6 3 の処理は、第 1 の実施形態と同様である。第 1 の実施形態と比べ、S 2 6 2 の処理が異なり、また、新たに、S 2 6 4 及び S 2 6 5 の処理が実行される。

【 0 2 6 3 】

S 2 6 4 において、ジョブ終了処理機能 2 6 は、プロセッサ稼働状況テーブル 3 0 を参照し、要求された処理を実行しているプロセッサコア 7 8 1 を含む制御部番号 3 0 1、プロセッサコア 7 8 1 を識別するプロセッサコア番号 3 0 2、及び論理ボリューム番号 3 0 4 に対応したエントリを読み出す。

40

【 0 2 6 4 】

S 2 6 2 において、ジョブ終了処理機能 2 6 は、読み出された情報（制御部番号 3 0 1、プロセッサコア番号 3 0 2、及び論理ボリューム番号 3 0 4）に一致するエントリの累計稼働時間 3 9 3 に、S 2 6 1 で算出された差分を加算する。

【 0 2 6 5 】

S 2 6 5 において、ジョブ終了処理機能 2 6 は、プロセッサ稼働状況テーブル 3 0 の対応するエントリの論理ボリューム番号 3 0 4 を消去し、処理を終了する。

【 0 2 6 6 】

図 4 1 は、本発明の第 2 の実施形態のデータ転送起動機能 2 7 の一例を示す図である。

50

【0267】

データ転送起動機能27は、キャッシュメモリ部73とホストインターフェース部72との間でデータ転送を行うとき、またはキャッシュメモリ部73とディスクインターフェース部74との間でデータ転送を行うとき起動される。

【0268】

まず、データ転送起動機能27は、どの論理ボリュームにおける転送処理であるかを判定する(S271)。通常、データ転送起動機能27を呼び出すジョブがどの論理ボリュームの処理を行っているかを管理しており、データ転送起動機能27は、当該ジョブが管理する情報を参照し、どの論理ボリュームの操作を行っているかを判定する。

【0269】

次に、データ転送起動機能27は、転送されるデータ転送量を、S271で判定された論理ボリュームと一致する論理ボリューム番号911のエントリに、転送されるデータの転送量を加算する(S272)。

【0270】

データ転送起動機能27は、データ転送を起動し(S273)、処理を終了する。具体的には、データ転送起動機能27は、ホストインターフェース部72のデータ転送LSI722、または、ディスクインターフェース部74のデータ転送LSI742に対し、データ転送コマンドを送信する。

【0271】

図42は、本発明の第2の実施形態の電力推定機能24の一例を説明するフローチャートである。以下、第1の実施形態と異なるステップを中心に説明する。

【0272】

S241、S2401~S2403、S244~245、及びS2408~S2409は、第1の実施形態と同様である。

【0273】

S242において、電力推定機能24は、ディスク稼働記録テーブル33の情報と本処理が呼び出される周期との比を用いて、各論理ボリュームについてアクセス種別毎に稼働率を算出する。

【0274】

S243において、電力推定機能24は、算出された稼働率とディスク電力仕様テーブル32の情報とから、各論理ボリュームの電力増加分を算出する。

【0275】

S2405において、電力推定機能24は、各論理ボリュームのデータ転送量から、論理ボリューム毎のキャッシュメモリ部73の電力増加分を算出する。

【0276】

具体的には、電力推定機能24は、キャッシュアクセス量テーブル91の累計アクセス量912の総和に対する任意の論理ボリュームの累計アクセス量912の比を算出し、算出された比と図3のアクセス時間7322とを乗算し、任意の論理ボリュームに対するキャッシュメモリ部73のアクセス時間を算出する。

【0277】

次に、電力推定機能24は、本処理の呼び出し周期と、算出された任意の論理ボリュームに対するキャッシュメモリ部73のアクセス時間とから論理ボリュームにおけるキャッシュメモリ部73の稼働率を算出し、算出された論理ボリュームにおけるキャッシュメモリ部73の稼働率を用いて、キャッシュメモリ部73の電力増加分を算出する。なお、キャッシュメモリ部73の電力増加分の算出方法は、第1の実施形態と同様である。

【0278】

S2406において、電力推定機能24は、プロセッサ稼働記録テーブル39の累計稼働時間393と本処理が呼び出される周期との比を用いて、論理ボリューム毎にプロセッサコア781の稼働率を算出する。算出方法は、第1の実施形態と同様である。

【0279】

10

20

30

40

50

S 2 4 0 7において、電力推定機能 2 4 は、算出されたプロセッサの稼働率から、論理ボリューム毎の制御部 7 5 の稼働率を算出する。算出方法は、第 1 の実施形態と同様である。

【 0 2 8 0 】

図 4 3 は、本発明の第 2 の実施形態の電力推定機能 2 4 が実行する処理の S 2 4 4 における管理用端末 4 に送信される表示データ 9 9 0 の一例を示す図である。

【 0 2 8 1 】

表示データ 9 9 0 は、第 1 の実施形態と比べ、新たに論理ボリューム番号 9 9 5 を含む。論理ボリューム番号 9 9 5 は、図 3 1 の論理ボリューム番号 3 3 6 と同一のものである。

10

【 0 2 8 2 】

図 4 4 は、本発明の第 2 の実施形態の消費電力履歴テーブル 5 1 の一例を示す図である。

【 0 2 8 3 】

消費電力履歴テーブル 5 1 は、第 1 の実施形態と比べ、新たに論理ボリューム番号 5 1 6 を含む。論理ボリューム番号 5 1 6 は、図 3 1 の論理ボリューム番号 3 3 6 と同一のものである。これによって、管理用端末 4 は、論理ボリューム毎に電力増加分が分かる。

【 0 2 8 4 】

図 4 5 は、本発明の第 2 の実施形態の仮想ストレージ電力表示機能 5 4 の一例を説明するフローチャートである。

20

【 0 2 8 5 】

仮想ストレージ電力表示機能 5 4 は、管理用端末 4 を操作する管理者からの要求によって呼び出され、仮想ストレージの消費電力を表示する。

【 0 2 8 6 】

まず、仮想ストレージ電力表示機能 5 4 は、消費電力履歴テーブル 5 1 を参照し、I d 1 e 時電力 5 1 5 を画面に表示する (S 5 4 1) 。

【 0 2 8 7 】

仮想ストレージ電力表示機能 5 4 は、仮想ストレージ - ボリューム対応テーブル 5 3 を参照し、任意の仮想ストレージ番号に対応する仮想ストレージのエントリを選択する (S 5 4 2) 。

30

【 0 2 8 8 】

次に、仮想ストレージ電力表示機能 5 4 は、仮想ストレージ - ボリューム対応テーブル 5 3 を参照し、選択された仮想ストレージに含まれる論理ボリュームを列挙する (S 5 4 3) 。

【 0 2 8 9 】

次に、仮想ストレージ電力表示機能 5 4 は、消費電力履歴テーブル 5 1 を参照し、列挙された論理ボリュームの電力増加分を合算する (S 5 4 4) 。

【 0 2 9 0 】

仮想ストレージ電力表示機能 5 4 は、合算された計算結果を画面に表示する (S 5 4 5) 。

40

【 0 2 9 1 】

仮想ストレージ電力表示機能 5 4 は、全ての仮想ストレージのエントリに対して前述の処理が実行されたか否かを判定する (S 5 4 6) 。

【 0 2 9 2 】

全ての仮想ストレージのエントリに対して前述の処理が実行されていないと判定された場合、仮想ストレージ電力表示機能 5 4 は、S 5 4 2 に戻り、同様の処理を実行する。

【 0 2 9 3 】

全ての仮想ストレージのエントリに対して前述の処理が実行されたと判定された場合、仮想ストレージ電力表示機能 5 4 は、処理を終了する。

【 0 2 9 4 】

50

図46は、本発明の第2の実施形態の仮想ストレージ電力表示機能54が実行する仮想ストレージ消費電力の出力の一例を示す図である。

【0295】

図46に示す例では、仮想ストレージ電力表示機能54は、定常電力591、及び電力増加分592を表示している。

【0296】

定常電力591は、通電され、処理を何も実行していないときにストレージシステム1が消費する電力、即ち、Idle時の電力である。電力増加分592は、各仮想ストレージに関する仕事による電力増加である。

【0297】

更新部593は、当該更新部593を操作することによって各データを更新する。

【0298】

終了部594は、当該終了部594を操作することによって表示を終了させる。

【0299】

以上説明したように、第2の実施形態によれば、複数の仮想ストレージをサポートするストレージシステム1において、仮想ストレージ毎の電力増加量を求めることができる。

【0300】

その結果、仮想ストレージ機能を使って一台のストレージシステム1において、複数のストレージシステム1について個別に電力を求める場合と同様に、仮想ストレージ毎の消費電力を求めることができる。

【0301】

なお、第2の実施形態では、ディスクの稼働時間を論理ボリューム毎に測定することによって求めていたが、キャッシュメモリ部73の電力推定方法と同様、ストレージシステム1は、稼働時間をI/O数または転送量などを論理ボリューム毎に格納し、電力増加量を論理ボリューム毎のI/O数または転送量に応じて比例配分する形態であってもよい。

【0302】

また、第2の実施形態では、仮想ストレージ毎の消費電力を求めたが、同様の方法で、特定のホストコンピュータ71毎に消費電力を求めることもできる。

【0303】

また、第2の実施形態では、論理ボリュームの識別子を用いて、論理ボリューム毎の推定電力を算出したが、各テーブルが仮想ストレージを識別する識別子を含み、仮想ストレージ毎に推定電力を算出する形態であってもよい。

【0304】

[第3の実施形態]

本発明の第3の実施形態は、ディスクドライブ6の状態制御(アクティブ、スタンバイ、または回転OFF)を実行するストレージシステム1において、状態変更を考慮した電力推定を行う発明である。

【0305】

第3の実施形態のストレージシステムの構成は、第2の実施形態と同様のものであり、説明は省略する。

【0306】

図47は、本発明の第3の実施形態のストレージ制御プログラム2が備える機能を説明するブロック図である。

【0307】

第2の実施形態と比べ、第3の実施形態のストレージ制御プログラム2は、新たにディスク状態変更機能28を備える。ディスク状態変更機能28は、ディスクドライブ6の状態を管理する。

【0308】

図48は、本発明の第3の実施形態の制御データ3の構成を説明するブロック図である。

。

10

20

30

40

50

【0309】

第2の実施形態と比べ、第3の実施形態の制御データ3は、新たにディスク状態テーブル93を備える。ディスク状態テーブル93については、図51を用いて後述する。

【0310】

図49は、第3の実施形態のディスク電力仕様テーブル32の一例を示す図である。

【0311】

ディスク電力仕様テーブル32は、第2の実施形態と比べ、新たに、スタンバイ時電力327、及びスピンドアウン時電力328を含む。

【0312】

スタンバイ時電力327は、ディスクドライブ6が、スタンバイ状態にあるときの消費電力を格納する。

10

【0313】

スピンドアウン時電力328は、ディスクドライブ6が、スピンドアウン状態にあるときの消費電力を格納する。

【0314】

図50は、本発明の第3の実施形態のディスク稼働記録テーブル33の一例を示す図である。

【0315】

ディスク稼働記録テーブル33は、第2の実施形態と比べ、新たに、スタンバイ時間337、及びスピンドアウン時間338を含む。スタンバイ時間337、及びスピンドアウン時間338は、ディスク番号331毎に格納される。

20

【0316】

スタンバイ時間337は、ディスクドライブ6がスタンバイ状態になっている累計時間を格納する。

【0317】

スピンドアウン時間338は、ディスクドライブ6がスピンドアウン状態になっている累計時間を格納する。

【0318】

図51は、本発明の第3の実施形態のディスク状態テーブル93の一例を示す図である。

30

【0319】

ディスク状態テーブル93は、ストレージシステム1内のディスクドライブ6の現在の稼働状態、及び状態の変更開始時間を管理する。

【0320】

ディスク状態テーブル93は、ディスク番号931、現在の状態932、及び開始時刻933を含む。

【0321】

ディスク番号931は、図6のディスク番号311と同一のものである。

【0322】

現在の状態932は、現在のディスクドライブ6の稼働状態を示す情報を格納する。例えば、通常の稼働状態の場合「アクティブ」が格納され、スタンバイ状態の場合「スタンバイ」が格納され、スピンドアウン状態の場合「スピンドアウン」が格納される。

40

【0323】

開始時刻933は、現在状態932に示されている稼働状態になった時刻、即ち、状態変更の開始時間を格納する。

【0324】

図52は、本発明の第3の実施形態のディスク状態変更機能28の一例を説明するフローチャートである。

【0325】

ディスク状態変更機能28は、ホストコンピュータ71または管理用端末4からの指示

50

によって、ディスクドライブ 6 の状態を変更するときに呼び出され、ディスクの状態を変更し、状態の変更時間を管理する。

【0326】

まず、ディスク状態変更機能 28 は、対象ディスクドライブ 6 が現在スタンバイ状態、またはスピンドウン状態にあるか否かを判定する (S281)。

【0327】

対象ディスクドライブ 6 がスタンバイ状態、またはスピンドウン状態であると判定された場合、即ち、スタンバイ状態またはスピンドウン状態が終了すると判定された場合、ディスク状態変更機能 28 は、ディスク状態テーブル 93 を参照し、対象ディスクドライブ 6 の開始時刻 933 と現在時刻との差分を算出する (S282)。

10

【0328】

ディスク状態変更機能 28 は、算出された差分をディスク稼働記録テーブル 33 の対応するエントリのスタンバイ時間 337、またはスピンドウン時間 338 へ加算する (S283)。

【0329】

ディスク状態変更機能 28 は、ディスク状態テーブル 93 の対象ディスクドライブ 6 のエントリを初期化し (S284)、S285 に進む。

【0330】

S281 において、対象ディスクドライブ 6 がスタンバイ状態、またはスピンドウン状態でないと判定された場合、ディスク状態変更機能 28 は、新しく変更される状態がスタンバイまたはスピンドウンであるか否かを判定する (S285)。

20

【0331】

新しく変更される状態がスタンバイまたはスピンドウンでないと判定された場合、ディスク状態変更機能 28 は、S288 へ進む。

【0332】

S285 において、新しく変更される状態がスタンバイまたはスピンドウンであると判定された場合、ディスク状態変更機能 28 は、ディスク状態テーブル 93 の対象ディスクドライブ 6 の開始時刻 933 に現在時刻を格納する (S286)。

【0333】

ディスク状態変更機能 28 は、ディスク状態テーブル 93 の現在状態 932 に、新しい状態を格納し (S287)、S288 に進む。

30

【0334】

ディスク状態変更機能 28 は、ホストコンピュータ 71 または管理用端末 4 からの指示された状態に、ディスクの状態を変更し (S288)、処理を終了する。

【0335】

図 53 は、本発明の第 3 の実施形態の電力推定機能 24 の一例を説明するフローチャートである。以下、第 1 の実施形態と異なるステップを中心に説明する。

【0336】

S240 において、電力推定機能 24 は、ディスク稼働記録テーブル 33 のスタンバイ時間 337、及びスピンドウン時間 338 と、本処理が呼び出される周期との比を用いて、ディスクドライブ 6 の各状態における比率を算出する。算出方法は、第 1 の実施形態と同様である。

40

【0337】

S241 において、電力推定機能 24 は、算出された比率とディスク電力仕様テーブル 32 の情報とからディスク部消費電力を算出する。具体的には、ディスクドライブ 6 の状態における消費電力も算出される。なお、ディスク部消費電力の算出方法は、第 1 の実施形態と同様である。

【0338】

S242 ~ S245、及び S2401 ~ S2409 の処理は第 1 の実施形態と同一のものである。

50

【0339】

以上説明したように、第3の実施形態によれば、ディスクドライブ6の状態制御（アクティブ、スタンバイ、または回転OFF）を実行するストレージシステム1において、その状態変更を考慮した電力推定を行うことができる。

【0340】

[第4の実施形態]

本発明の第4の実施形態は、ストレージシステム1が、ディスクドライブ6と半導体ディスク（SSD（Solid State Drive））とを備える場合に、両者の特性の違いを考慮した電力推定を行う発明である。

【0341】

第4の実施形態では、ディスクドライブ6に対してはアクセス時間を用い電力推定を行い、また、SSDに対しては転送量を用いて電力推定を行うことによって、精度の高い電力推定を行う。

【0342】

第4の実施形態のストレージシステムの構成は、第2の実施形態と同様のものであり、説明は省略する。ただし、ディスクドライブ6には、SSDが含まれる。以下、SSD6とも記載する。

【0343】

図54は、本発明の第4の実施形態の制御データ3の構成を説明するブロック図である。

【0344】

第3の実施形態と比べ、第4の実施形態の制御データ3は、新たにSSD電力仕様テーブル94、及びSSD稼働記録テーブル95を備える。SSD電力仕様テーブル94、及びSSD稼働記録テーブル95については、図55及び図56を用いて後述する。

【0345】

図55は、本発明の第4の実施形態のSSD電力仕様テーブル94の一例を示す図である。

【0346】

SSD電力仕様テーブル94は、ストレージシステム1に搭載されているSSD6の種別毎に、各々の状態における消費エネルギーを管理する。SSD電力仕様テーブル94は、新規種別のディスクドライブ6が追加されたときに更新される。また、電力を推定するときに、電力推定機能24がSSD電力仕様テーブル94を参照する。

【0347】

SSD電力仕様テーブル94は、ディスク型番941、Idle時電力942、Random Read時消費エネルギー943、Random Write時消費エネルギー944、Sequential Read時消費エネルギー945、及びSequential Write時消費エネルギー946を含む。

【0348】

ディスク型番941は、SSD6の種別を識別する識別子を格納する。

【0349】

Idle時電力942は、対応するSSD6がIdle状態のときの消費電力を示す値を格納する。

【0350】

Random Read時消費エネルギー943は、ディスク型番941に対応するSSD6がRandom Readのみを受け付けており、かつ、限界まで稼働している状態におけるデータ転送量当たりの消費エネルギーを示す値を格納する。

【0351】

Random Write時消費エネルギー944は、ディスク型番941に対応するSSD6がRandom Writeのみを受け付けており、かつ、限界まで稼働している状態におけるデータ転送量当たりの消費エネルギーを示す値を格納する。

10

20

30

40

50

【0352】

Sequential Read時消費エネルギー945は、ディスク型番941に対応するSSD6がSequential Readのみを受け付けており、かつ、限界まで稼働している状態におけるデータ転送量当たりの消費エネルギーを示す値を格納する。

【0353】

Sequential Write時消費エネルギー946は、ディスク型番941に対応するSSD6がSequential Writeのみを受け付けており、かつ、限界まで稼働している状態におけるデータ転送量当たりの消費エネルギーを示す値を格納する。

【0354】

図56は、本発明の第4の実施形態のSSD稼働記録テーブル95の一例を示す図である。

10

【0355】

SSD稼働記録テーブル95は、SSD6にアクセスされたときのデータ転送量を管理する。SSD稼働記録テーブル95はSSD6にアクセスされる毎に更新される。また、電力を推定するときに、電力推定機能24がSSD稼働記録テーブル95を参照する。

【0356】

SSD稼働記録テーブル95は、ディスク番号951、論理ボリューム番号952、Random Read累計転送量953、Random Write累計転送量954、Sequential Read累計転送量955、及びSequential Write累計転送量956を含む。

20

【0357】

ディスク番号951及び論理ボリューム番号952は、図31のディスク番号331及び論理ボリューム番号336と同一のものである。

【0358】

Random Read累計転送量953は、対応するSSD6に対してRandom Readアクセスにおける転送量の累積転送量を格納する。

【0359】

Random Write累計転送量954は、対応するSSD6に対してRandom Writeアクセスにおける転送量の累積転送量を格納する。

30

【0360】

Sequential Read累計転送量955は、対応するSSD6に対してSequential Readアクセスにおける転送量の累積転送量を格納する。

【0361】

Sequential Write累計転送量956は、対応するSSD6に対してSequential Writeアクセスにおける転送量の累積転送量を格納する。

【0362】

図57は、本発明の第4の実施形態のディスクI/O発行処理機能22の一例を説明するフローチャートである。

【0363】

S221～S224までの処理は、第3の実施形態と同様である。

40

【0364】

S229において、ディスクI/O発行処理機能22は、アクセス先がSSD6であるか否かを判定する。具体的には、ディスクI/O発行処理機能22は、S221で抽出されたアクセス先のディスクドライブ6に関する情報を参照して判定する。

【0365】

アクセス先がSSD6でないと判定された場合、ディスクI/O発行処理機能22は、S225に進む。以下、S225～S227の処理は第3の実施形態と同様である。

【0366】

アクセス先がSSD6であると判定された場合、ディスクI/O発行処理機能22は、

50

処理を終了する。アクセス先がSSD6の場合、稼働時間の代わりに転送量を用いて電力推定がされるため、S225～S227の処理は省略される。

【0367】

図58は、本発明の第4の実施形態のディスクI/O完了処理機能23の一例を説明するフローチャートである。

【0368】

S231～S238までの処理は、第3の実施形態と同様である。

【0369】

S239において、ディスクI/O完了処理機能23は、アクセス先がSSD6であるか否かを判定する。

10

【0370】

アクセス先がSSD6でないと判定された場合、ディスクI/O完了処理機能23は、S234に進む、以下、SSD234～S236の処理は第3の実施形態と同様である。

【0371】

アクセス先がSSD6であると判定された場合、ディスクI/O完了処理機能23は、SSD稼働記録テーブル95の対応するエントリの各アクセス種別の累積転送量(943～946)毎に転送量を加算し(S230)、処理を終了する。

【0372】

図59A及び図59Bは、本発明の第4の実施形態の電力推定機能24の一例を説明するフローチャートである。

20

【0373】

電力推定機能24は、ディスク部構成テーブル31を参照し、電力推定の対象を選択する(S2411)。

【0374】

次に、電力推定機能24は、選択された電力推定の対象がSSDであるか否かを判定する(S2412)。

【0375】

選択された電力推定の対象がSSDでないと判定された場合、電力推定機能24は、S241に進む。以下、S241～S243までの処理は第3の実施形態と同様である。電力推定機能24は、S243の後に、S2416に進む。

30

【0376】

選択された電力推定の対象がSSDであると判定された場合、電力推定機能24は、ディスク部構成テーブル31及びSSD電力仕様テーブル94を参照し、SSD定常電力分を算出する(S2413)。算出方法は、S241と同様である。

【0377】

次に、電力推定機能24は、SSD電力仕様テーブル94及びSSD稼働記録テーブル95を参照し、SSD6が消費したエネルギーを算出する(S2414)。具体的には、各アクセス種別の、消費エネルギー(933～936)と累積転送量(943～946)とを乗算することによって算出される。

40

【0378】

次に、電力推定機能24は、S2414で算出された消費エネルギーと本処理が呼び出される周期とを用いてSSDの電力増加分を算出する(S2415)。具体的には、S2414で算出された消費エネルギーを本処理が呼び出される周期で除算して算出される。電力推定機能24は、SSDの電力増加分を算出したエネルギーを算出した後、S2416に進む。

【0379】

電力推定機能24は、全ての電力推定の対象について処理を実行したか否かを判定する(S2416)。

【0380】

全ての電力推定の対象について処理を実行していないと判定された場合、電力推定機能

50

24は、S2411に戻り、同様の処理を実行する。

【0381】

全ての電力推定の対象について処理を実行していないと判定された場合、電力推定機能24は、S2401に進む。以下、S2401～S245の処理は、第3の実施形態と同様である。

【0382】

S2417において、電力推定機能24は、SSD稼働記録テーブル95の各アクセス種別の累計転送量(943～946)を初期化する。

【0383】

以上説明したように、第4の実施形態によれば、ディスクドライブ6とSSD6とを備えるストレージシステム1においても、両者の特性の違いを考慮した精度の高い電力推定を行うことができる。

【0384】

[第5の実施形態]

本発明の第5の実施形態は、論理分割を行うストレージシステム1において、各仮想ストレージの電力に定常電力を含めて管理者に提示する発明である。

【0385】

第5の実施形態では、ストレージシステム1の各物理資源がどの仮想ストレージに利用されているかを判定し、仮想ストレージにおける物理資源の利用の程度に応じて、定常電力を分配し、当該分配された電力が表示される。

【0386】

図60は、本発明の第5の実施形態のストレージ管理プログラム5が備える機能及びテーブルを説明するブロック図である。

【0387】

第3の実施形態と比べ、ストレージ管理プログラム5は、新たに、仮想ストレージ-ディスク資源対応テーブル55、及び仮想ストレージ-コントローラ資源対応テーブル56を備える。仮想ストレージ-ディスク資源対応テーブル55、及び仮想ストレージ-コントローラ資源対応テーブル56については、図60及び図61を用いて後述する。

【0388】

図61は、本発明の第5の実施形態の仮想ストレージ-ディスク資源対応テーブル55の一例を示す図である。

【0389】

仮想ストレージ-ディスク資源対応テーブル55は、各ディスクドライブ6がどの仮想ストレージに用いられているかを管理する。仮想ストレージ-ディスク資源対応テーブル55は、仮想ストレージ毎の電力表示を行うときに参照される。

【0390】

仮想ストレージ-ディスク資源対応テーブル55は、ディスク番号551、及び利用仮想ストレージ番号552を含む。

【0391】

ディスク番号551は、図6のディスク番号311と同一のものである。

【0392】

利用仮想ストレージ番号552は、対応する各ディスクドライブ6がどの仮想ストレージに利用されているかを示す情報を格納する。

【0393】

図62は、本発明の第5の実施形態の仮想ストレージ-コントローラ資源対応テーブル56の一例を示す図である。

【0394】

仮想ストレージ-コントローラ資源対応テーブル56は、ストレージシステム1の各物理資源が、どの仮想ストレージに用いられているかを管理する。仮想ストレージ-コントローラ資源対応テーブル56は、仮想ストレージ毎の電力表示を行うときに参照される。

10

20

30

40

50

【0395】

仮想ストレージ - コントローラ資源対応テーブル56は、種別561、スロット番号、及び利用仮想ストレージ番号563を含む。

【0396】

種別561及びスロット番号562は、図11の種別361及びスロット番号362と同一のものである。

【0397】

利用仮想ストレージ番号は、種別561及びスロット番号562に対応する物理資源がどの仮想ストレージに利用されているかを示す情報を格納する。具体的には、仮想ストレージ番号531が格納される。

10

【0398】

図63は、本発明の第5の実施形態の仮想ストレージ電力表示機能54の一例を示す図である。

【0399】

仮想ストレージ電力表示機能54は、管理者の電力表示要求等によって呼び出され、仮想ストレージ毎の消費電力を算出し、算出された仮想ストレージ毎の消費電力を表示する。

【0400】

まず、仮想ストレージ電力表示機能54は、ディスクドライブまたはコントローラ部位から消費電力の算出対象を選択する(S5411)。

20

【0401】

仮想ストレージ電力表示機能54は、仮想ストレージ - ディスク資源対応テーブル55、及び仮想ストレージ - コントローラ資源対応テーブル56を参照し、該当する部位を用いている仮想ストレージを選択する(S5412)。

【0402】

次に、仮想ストレージ電力表示機能54は、該当部位が電力変動するか否かを判定する(S5413)。例えば、ディスクドライブ6、または制御部75は電力変動すると判定され、ホストI/F72、またはスイッチ部76は電力変動しないと判定される。

【0403】

該当部位が電力変動すると判定された場合、仮想ストレージ電力表示機能54は、仮想ストレージ - ボリューム対応テーブル53を参照し、選択された仮想ストレージに含まれる論理ボリュームを列挙する(S5417)。

30

【0404】

次に、仮想ストレージ電力表示機能54は、消費電力履歴テーブル51を参照し、選択された部位について、各論理ボリュームの電力増加分を合算し、仮想ストレージの電力増加分を算出する(S5418)。

【0405】

次に、仮想ストレージ電力表示機能54は、選択された部位について、全ての論理ボリュームの電力増加分を合算し、電力増加合計を算出する(S5419)。

【0406】

仮想ストレージ電力表示機能54は、電力増加合計に対する、仮想ストレージ電力増加の割合を算出し、算出された割合と前託された部位の定常電力とを乗じ、仮想ストレージの該当部位の定常電力を算出し(S5420)、S5421に進む。

40

【0407】

S5414において、該当部位が電力変動しないと判定された場合、仮想ストレージ電力表示機能54は、仮想ストレージ - ディスク資源対応テーブル55、または仮想ストレージ - コントローラ資源対応テーブル56を参照し、当該資源を利用している仮想ストレージの数を算出する(S5414)。

【0408】

仮想ストレージ電力表示機能54は、消費電力履歴テーブル51を参照して、該当部位

50

の定常電力を求める (S 5 4 1 5) 。

【 0 4 0 9 】

算出された該当部位の定常電力を該当する部位を利用している仮想ストレージの数で除算し、仮想ストレージの該当部位の消費電力を算出する (S 5 4 1 6) 。

【 0 4 1 0 】

仮想ストレージ電力表示機能 5 4 は、以上の処理のよって算出された結果を画面に表示する (S 5 4 2 1) 。

【 0 4 1 1 】

仮想ストレージ電力表示機能 5 4 は、全ての仮想ストレージに対し処理を実行したか否かを判定する (S 5 4 2 2) 。

10

【 0 4 1 2 】

全ての仮想ストレージに対し処理を実行していないと判定された場合、仮想ストレージ電力表示機能 5 4 は、S 5 4 1 2 に戻り同様の処理を実行する。

【 0 4 1 3 】

全ての仮想ストレージに対し処理を実行したと判定された場合、仮想ストレージ電力表示機能 5 4 は、全ての部位に対して処理を実行したか否かを判定する (S 5 4 2 3) 。

【 0 4 1 4 】

全ての部位に対して処理を実行していないと判定された場合、仮想ストレージ電力表示機能 5 4 は、S 5 4 1 1 に戻り同様の処理を実行する。

【 0 4 1 5 】

全ての部位に対して処理を実行したと判定された場合、仮想ストレージ電力表示機能 5 4 は、処理を終了する。

20

【 0 4 1 6 】

以上の処理によって、電力変動する部位については、各仮想ストレージが該当部位を利用率に応じて定常電力を配分する。また、電力変動しない部位については、仮想ストレージ毎に定常電力を均等に割り振る。これによって、各仮想ストレージの利用状況に応じて、各仮想ストレージの電力の合計と実際の消費電力の合計とが対応するように消費電力を割り当てることができる。

【 0 4 1 7 】

図 6 4 は、本発明の第 5 の実施形態の仮想ストレージ消費電力表示画面の一例を示す図である。

30

【 0 4 1 8 】

各仮想ストレージ (5 9 2 A 、 5 9 2 B) に、各部位の消費電力を表示する。

【 0 4 1 9 】

図 6 4 に示す例では、仮想ストレージ電力表示機能 5 4 は、各仮想ストレージの消費電力を、ストレージシステム 1 の物理資源毎に表示している。具体的には、ストレージ 0 (5 9 2 A) 及びストレージシステム 1 (5 9 2 B) について、各仮想ストレージが利用している物理資源毎に消費電力が表示されている。は、通電され、処理を何も実行していないときにストレージシステム 1 が消費する電力である。

【 0 4 2 0 】

以上説明したように、第 5 の実施形態によれば、論理分割を行うストレージシステム 1 において、各仮想ストレージの電力に定常電力を含めて管理者に提示することができる。

40

【 0 4 2 1 】

特許請求の範囲に記載した以外の発明の観点の代表的なものとして、次のものがあげられる。

【 0 4 2 2 】

(1) 管理計算機と、前記管理計算機及びホスト計算機に接続されたストレージシステムとを備える計算機システムにおける電力推定方法であって、

前記管理計算機は、前記ストレージシステムに接続される第 1 のインタフェースと、前記第 1 のインタフェースに接続される第 1 のプロセッサと、前記第 1 のプロセッサに接続

50

される第 1 のメモリとを備え、

前記ホスト計算機は、前記ストレージシステムと接続される第 2 のインタフェースと、前記第 2 のインタフェースに接続される第 2 のプロセッサと、前記第 2 のプロセッサに接続される第 2 のメモリとを備え、

前記ストレージシステムは、複数のディスクドライブへのデータの読み書きを制御する少なくとも一つの第 3 のプロセッサと、前記第 3 のプロセッサと接続される制御メモリと、前記管理計算機と接続される管理インタフェースと、前記ホスト計算機と接続される少なくとも一つのホストインタフェースと、前記ディスクドライブと接続される少なくとも一つのディスクインタフェースと、データを一時的に格納するキャッシュメモリ部とを備え、

前記第 3 のプロセッサ、前記管理インタフェース、前記ホストインタフェース、前記ディスクインタフェース、及び前記キャッシュメモリ部は、互いに接続され、

前記制御メモリは、前記ストレージシステムを制御するためのストレージ制御プログラムと、前記ストレージシステムを制御するために必要な制御データとを格納し、

前記ストレージ制御プログラムは、電力を推定するための電力推定プログラムを含み、

前記キャッシュメモリ部は、複数のメモリと、前記複数のメモリを制御するメモリコントローラとを備え、

前記制御メモリは、

ストレージ装置が処理を実行している場合における、前記ディスクドライブ、前記キャッシュメモリ部及び前記第 3 のプロセッサの消費電力の増加分を示す増加消費電力情報を含み、

前記方法は、

前記第 3 のプロセッサが、前記ディスクドライブの稼働時間から、前記ディスクドライブへの各アクセス種別に対するディスクドライブの稼働率を算出し、

前記第 3 のプロセッサが、前記増加消費電力情報と前記算出された各アクセス種別に対するディスクドライブの稼働率とから、各アクセス種別における前記ディスクドライブの消費電力の増加分を算出し、

前記第 3 のプロセッサが、前記ディスクドライブに対するアクセスのときに使用された前記キャッシュメモリ部へのアクセス時間から前記キャッシュメモリ部の稼働率を算出し、

前記第 3 のプロセッサが、前記増加消費電力情報と前記算出されたキャッシュメモリ部の稼働率とから、前記ディスクドライブに対するアクセスのときに使用された前記キャッシュメモリ部の消費電力の増加分を算出し、

前記第 3 のプロセッサが、前記ディスクドライブのアクセスに対する処理によって前記第 3 のプロセッサが稼働した稼働時間から、前記ディスクドライブのアクセスに対する処理における第 3 のプロセッサの稼働率を算出するステップと、

前記第 3 のプロセッサが、前記増加消費電力情報と前記算出されたディスクドライブのアクセスに対する処理における第 3 のプロセッサの稼働率とから、前記ディスクドライブのアクセスに対する処理における第 3 のプロセッサの消費電力の増加分を算出し、

前記第 3 のプロセッサが、前記算出された各アクセス種別におけるディスクドライブの消費電力の増加分、前記算出されたディスクドライブに対するアクセスのときに使用されたキャッシュメモリ部の消費電力の増加分、及び前記算出されたディスクドライブのアクセスに対する処理における第 3 のプロセッサの消費電力の増加分を前記管理計算機に送信することを特徴とする計算機システム。

【図面の簡単な説明】

【0423】

【図 1】本発明の第 1 の実施形態の計算機システムの構成を説明するブロック図である。

【図 2】本発明の第 1 の実施形態のキャッシュメモリ部の構成を説明するブロック図である。

【図 3】本発明の第 1 の実施形態のアクセス時間記録テーブルの一例を示す図である。

10

20

30

40

50

【図 4】本発明の第 1 の実施形態のストレージ制御プログラムが備える機能を説明するブロック図である。

【図 5】本発明の第 1 の実施形態の制御データの構成を説明するブロック図である。

【図 6】本発明の第 1 の実施形態のディスク部構成テーブルの一例を示す図である。

【図 7】本発明の第 1 の実施形態のディスク電力仕様テーブルの一例を示す図である。

【図 8】本発明の第 1 の実施形態のディスク稼働記録テーブルの一例を示す図である。

【図 9】本発明の第 1 の実施形態の T a g 管理テーブルの一例を示す図である。

【図 10】本発明の第 1 の実施形態の I / O 発行状況テーブルの一例を示す図である。

【図 11】本発明の第 1 の実施形態のコントローラ部構成テーブルの一例を示す図である。

【図 12】本発明の第 1 の実施形態のコントローラ部電力仕様テーブルの一例を示す図である。

【図 13】本発明の第 1 の実施形態のコントローラ部電力変動仕様テーブルの一例を示す図である。

【図 14】本発明の第 1 の実施形態のプロセッサ稼働状況テーブルの一例を示す図である。

【図 15】本発明の第 1 の実施形態のプロセッサ稼働状況テーブルの一例を示す図である。

【図 16】第 1 の実施形態のホスト I / O 受領機能の一例を説明するフローチャートである。

【図 17】本発明の第 1 の実施形態のディスク I / O 発行機能の一例を説明するフローチャートである。

【図 18】本発明の第 1 の実施形態のディスク I / O 完了処理機能の一例を説明するフローチャートである。

【図 19】本発明の第 1 の実施形態のジョブ起動処理機能の一例を説明する図である。

【図 20】本発明の第 1 の実施形態のジョブ終了処理機能の一例を説明する図である。

【図 21】本発明の第 1 の実施形態の電力推定機能の一例を説明するフローチャートである。

【図 22】本発明の第 1 の実施形態の電力推定機能が S 2 4 1 及び S 2 4 3 におけるディスク部電力増加分を算出方法の一例を示す図である。

【図 23】本発明の第 1 の実施形態の電力推定機能が実行する処理の S 2 4 4 における管理用端末に送信される表示データの一例を示す図である。

【図 24】本発明の第 1 の実施形態のストレージ管理プログラムが備える機能及びテーブルを説明するブロック図である。

【図 25】本発明の第 1 の実施形態の消費電力履歴テーブルの一例を示す図である。

【図 26】本発明の第 2 の実施形態のストレージ管理プログラムが備える機能及びテーブルを説明するブロック図である。

【図 27】本発明の第 2 の実施形態の仮想ストレージ ボリューム対応テーブルの一例を示す図である。

【図 28】本発明の第 2 の実施形態のストレージ制御プログラムが備える機能を説明するブロック図である。

【図 29】本発明の第 2 の実施形態の制御データの構成を説明するブロック図である。

【図 30】本発明の第 2 の実施形態の L D E V 物理ボリューム対応テーブルの一例を示す図である。

【図 31】本発明の第 2 の実施形態のディスク稼働記録テーブルの一例を示す図である。

【図 32】本発明の第 2 の実施形態の T a g 管理テーブルの一例を示す図である。

【図 33】本発明の第 2 の実施形態の I / O 発行状況テーブルの一例を示す図である。

【図 34】本発明の第 2 の実施形態のプロセッサ稼働状況テーブルの一例を示す図である。

【図 35】本発明の第 2 の実施形態のプロセッサ稼働状況テーブルの一例を示す図である

10

20

30

40

50

。

【図 3 6】本発明の第 2 の実施形態のキャッシュアクセス量テーブルの一例を示す図である。

【図 3 7】本発明の第 2 の実施形態のディスク I / O 発行機能の一例を示す図である。

【図 3 8】本発明の第 2 の実施形態のディスク I / O 完了処理機能の一例を示す図である。

。

【図 3 9】本発明の第 2 の実施形態のジョブ起動処理機能の一例を示す図である。

【図 4 0】本発明の第 2 の実施形態のジョブ終了処理機能の一例を示す図である。

【図 4 1】本発明の第 2 の実施形態のデータ転送起動機能の一例を示す図である。

【図 4 2】本発明の第 2 の実施形態の電力推定機能の一例を説明するフローチャートである。 10

【図 4 3】本発明の第 2 の実施形態の電力推定機能が実行する処理の S 2 4 4 における管理用端末に送信される表示データの一例を示す図である。

【図 4 4】本発明の第 2 の実施形態の消費電力履歴テーブルの一例を示す図である。

【図 4 5】本発明の第 2 の実施形態の仮想ストレージ電力表示機能の一例を説明するフローチャートである。

【図 4 6】本発明の第 2 の実施形態の仮想ストレージ電力表示機能が実行する仮想ストレージ消費電力の出力の一例を示す図である。

【図 4 7】本発明の第 3 の実施形態のストレージ制御プログラムが備える機能を説明するブロック図である。 20

【図 4 8】本発明の第 3 の実施形態の制御データの構成を説明するブロック図である。

【図 4 9】第 3 の実施形態のディスク電力仕様テーブルの一例を示す図である。

【図 5 0】本発明の第 3 の実施形態のディスク稼働記録テーブルの一例を示す図である。

【図 5 1】本発明の第 3 の実施形態のディスク状態テーブルの一例を示す図である。

【図 5 2】本発明の第 3 の実施形態のディスク状態変更機能の一例を説明するフローチャートである。

【図 5 3】本発明の第 3 の実施形態の電力推定機能の一例を説明するフローチャートである。

【図 5 4】本発明の第 4 の実施形態の制御データの構成を説明するブロック図である。

【図 5 5】本発明の第 4 の実施形態の SSD 電力仕様テーブルの一例を示す図である。 30

【図 5 6】本発明の第 4 の実施形態の SSD 稼働記録テーブルの一例を示す図である。

【図 5 7】本発明の第 4 の実施形態のディスク I / O 発行機能の一例を説明するフローチャートである。

【図 5 8】本発明の第 4 の実施形態のディスク I / O 完了処理機能の一例を説明するフローチャートである。

【図 5 9 A】本発明の第 4 の実施形態の電力推定機能の一例を説明するフローチャートである。

【図 5 9 B】本発明の第 4 の実施形態の電力推定機能の一例を説明するフローチャートである。

【図 6 0】本発明の第 5 の実施形態のストレージ管理プログラムが備える機能及びテーブルを説明するブロック図である。 40

【図 6 1】本発明の第 5 の実施形態の仮想ストレージ - ディスク資源対応テーブルの一例を示す図である。

【図 6 2】本発明の第 5 の実施形態の仮想ストレージ - コントローラ資源対応テーブルの一例を示す図である。

【図 6 3】本発明の第 5 の実施形態の仮想ストレージ電力表示機能の一例を示す図である。

。

【図 6 4】本発明の第 5 の実施形態の仮想ストレージ消費電力表示画面の一例を示す図である。

【符号の説明】

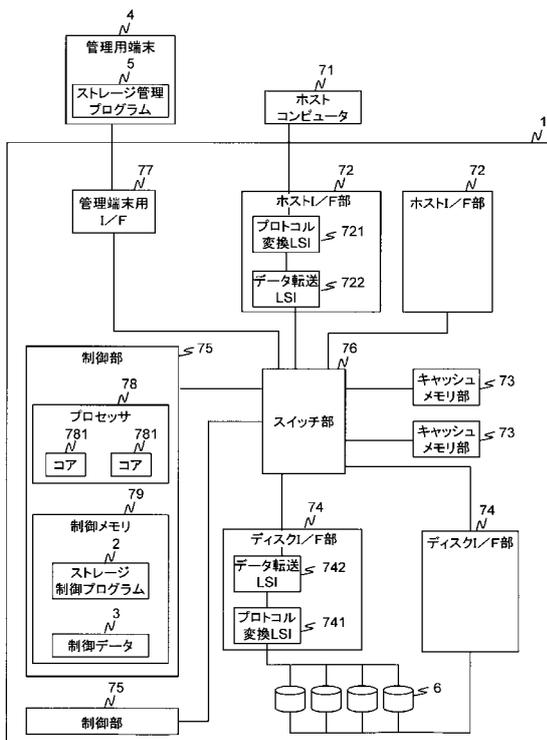
【 0 4 2 4 】

- 1 ストレージシステム
- 4 管理用端末
- 6 ディスクドライブ
- 7 1 ホストコンピュータ
- 7 2 ホスト I / F 部
- 7 3 キャッシュメモリ部
- 7 4 ディスク I / F 部
- 7 5 制御部
- 7 6 スイッチ部
- 7 7 管理端末用 I / F 部
- 7 8 プロセッサ
- 7 9 制御メモリ
- 7 2 1 プロトコル変換 L S I
- 7 2 2 データ転送 L S I
- 7 3 1 メモリコントローラ L S I
- 7 3 3 メモリモジュール (D I M M)
- 7 4 1 プロトコル変換 L S I
- 7 4 2 データ転送 L S I
- 7 8 1 プロセッサコア

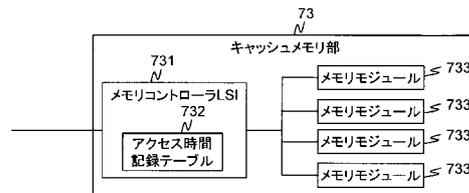
10

20

【 図 1 】



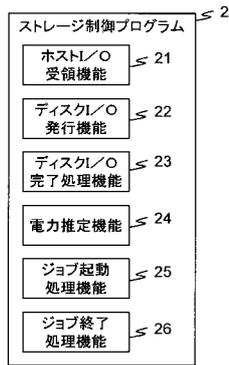
【 図 2 】



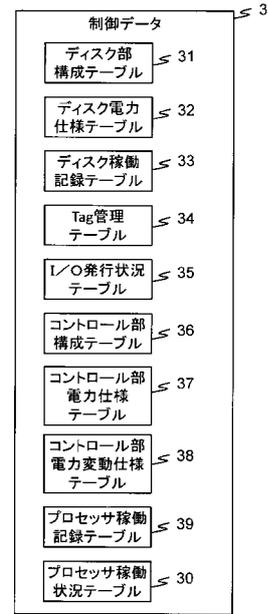
【 図 3 】

7321 チャンネル番号	7322 アクセス時間[μs]
0	64820
1	473,404
...	...

【 図 4 】



【 図 5 】



【 図 6 】

311 N ディスク番号	312 N ディスク型番
#0	A
#1	A
#2	B
#3	B
#4	C
#5	C
#6	なし
...	...

【 図 8 】

331 N ディスク 番号	332 N Random Read 累計稼働 時間[μs]	333 N Random Write 累計稼働 時間[μs]	334 N Sequential Read 累計稼働 時間[μs]	335 N Sequential Write 累計稼働 時間[μs]
#0	82,148	8,651	32,823	6,647
#1	9,384	46,095	5,020	107
...

【 図 7 】

321 N ディスク 型番	322 N Idle時 電力[W]	323 N Random Read時 電力増分[W]	324 N Random Write時 電力増分[W]	325 N Sequential Read時 電力増分[W]	326 N Sequential Write時 電力増分[W]
A	14.1	4.9	1.9	1.6	1.9
B	8.9	5.1	2.3	2.0	2.3
C	11.4	5.6	2.4	1.8	2.3
...

【 図 9 】

ディスク番号	Tag番号	アクセス種別
#0	0	Random Read
	1	Random Read
	2	なし
...
#1	0	Sequential Read
	1	Sequential Read
...

【 図 1 0 】

ディスク番号	アクセス種別	現在発行中 コマンド数	ディスク稼働 開始時刻
#0	Random Read	5	8,421652
	Random Write	0	-
	Sequential Read	10	8,421120
#1	Sequential Write	0	-
	Random Read	0	-
#1	Random Write	3	8,421573

【 図 1 1 】

種別	スロット番号	部品名
制御部	0	C-A
	1	C-B

キャッシュメモリ部	0	あり
	1	あり

ホストI/F	0	H-A
	1	H-B

ディスクI/F	0	D-A
	1	D-B

キャッシュメモリ メモリモジュール	0	M-A
	1	M-B

...

【 図 1 2 】

種別	部品名	Idle時電力[W]
制御部	C-A	85
	C-B	65

キャッシュメモリ部	-	40
ホストI/F	H-A	52
	H-B	56

ディスクI/F	-	48
キャッシュメモリ メモリモジュール	M-A	2.5
	M-B	4.2

...

【 図 1 3 】

381 N	382 N	383 N	38 N
種別	部品名	電力増分[W]	
制御部	C-A	42	
	C-B	36	
	
キャッシュメモリ メモリモジュール	M-A	3.7	
	M-B	5.3	
	
...	

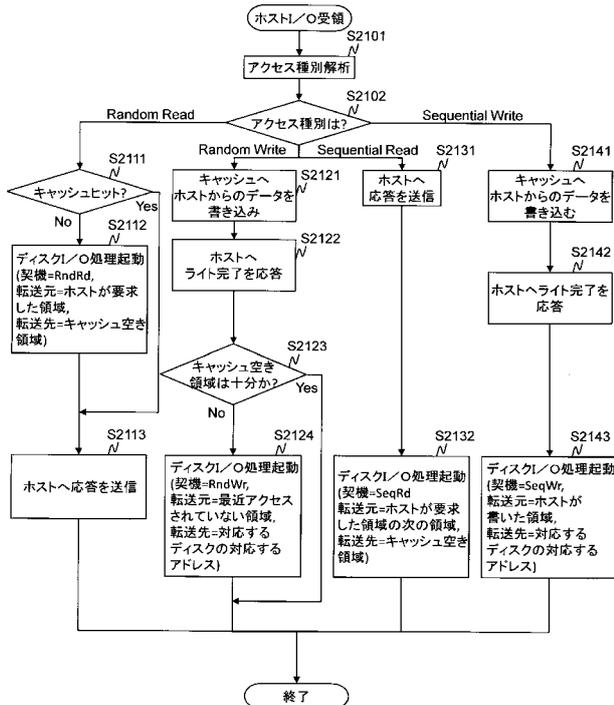
【 図 1 5 】

301 N	302 N	303 N	30 N
制御部番号	プロセッサコア番号	ジョブ開始時刻	
0	0	8,421,652	
	1	-	
	2	8,421,210	
	3	-	
1	0	-	
	1	8,421,573	
...	

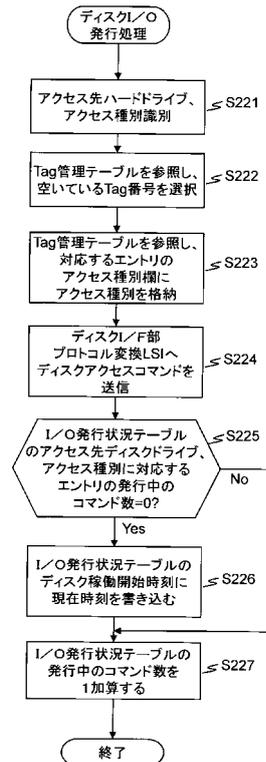
【 図 1 4 】

391 N	392 N	393 N	39 N
制御部番号	プロセッサコア番号	累計稼働時間[μs]	
0	0	758,243	
	1	426,197	
	2	548,164	
	3	137,654	
1	0	84,792	
	1	46,879	
...	

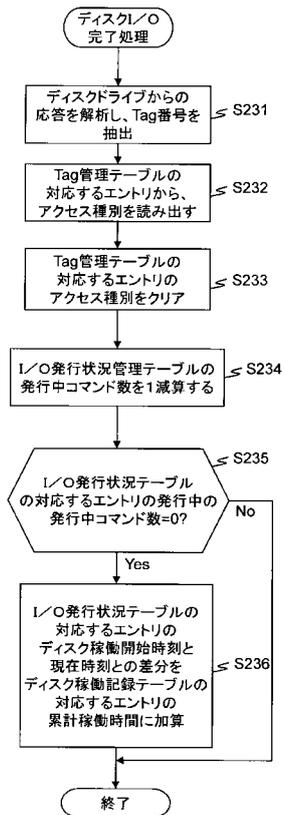
【 図 1 6 】



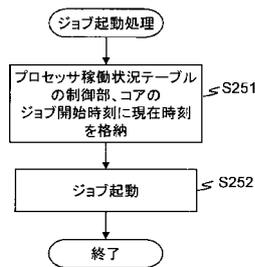
【 図 1 7 】



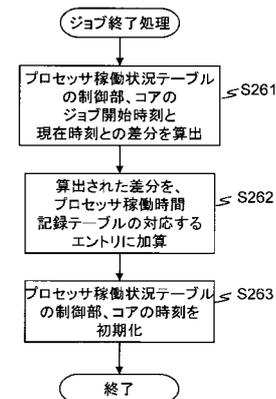
【 図 1 8 】



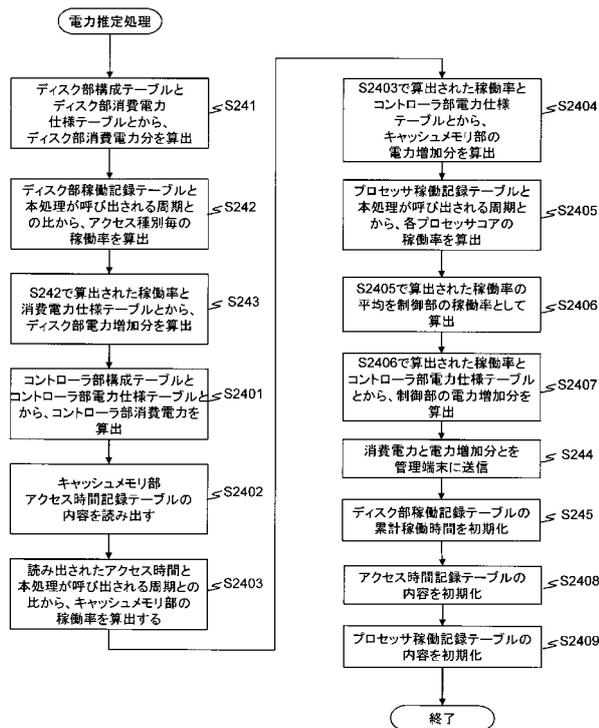
【 図 1 9 】



【 図 2 0 】



【 図 2 1 】



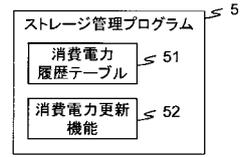
【 図 2 2 】

1001 N	算出値	1002 N	算出式
	P_static		P_idle
	P_RndRd		AR_RndRd × p_RndRd
	P_RndWr		AR_RndWr × p_RndWr
	P_SeqRd		AR_SeqRd × p_SeqRd
	P_SeqWr		AR_SeqWr × p_SeqWr
	P_Dynamic		P_RndRd + P_RndWr + P_SeqRd + P_SeqWr

【 図 2 3 】

部位	平均消費電力 [W]	電力増加[W]					Idle時電力 [W]
		Random Read 電力増加 [W]	Random Write 電力増加 [W]	Sequential Read 電力増加 [W]	Sequential Write 電力増加 [W]	コントローラ部電力増加 [W]	
ディスク0	18.0	3.7	0.2	0.0	0.0	-	14.1
ディスク1	16.3	0.0	0.0	2.0	0.2	-	14.1
キャッシュメモリ部0	40.0	-	-	-	-	0.0	40.0
キャッシュメモリ部1	40.0	-	-	-	-	0.0	40.0
DIMM 0	3.7	-	-	-	-	1.2	2.5
DIMM 1	7.2	-	-	-	-	3.0	4.2
制御部0	122.0	-	-	-	-	37.0	85.0
制御部1	75.0	-	-	-	-	10.0	65.0
ホスト I/F部0	52.0	-	-	-	-	0.0	52.0
ホスト I/F部1	56.0	-	-	-	-	0.0	56.0
ディスク I/F部0	48.0	-	-	-	-	0.0	48.0
...
合計	1,325	30	18	0	12		1,260

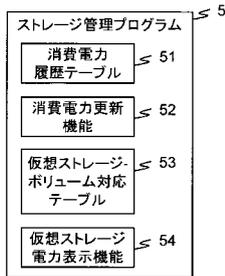
【 図 2 4 】



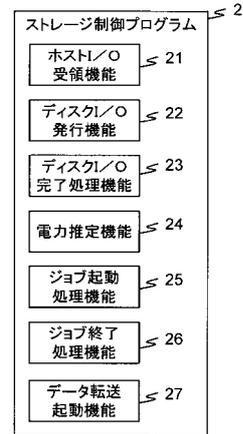
【 図 2 5 】

日時	部位	平均消費電力 [W]	電力増加[W]					Idle時電力 [W]
			Random Read 電力増加 [W]	Random Write 電力増加 [W]	Sequential Read 電力増加 [W]	Sequential Write 電力増加 [W]	コントローラ部電力増加 [W]	
05/11 10:00	ディスク0	18	3.7	0.2	0	0	-	14.1
11:00	制御部0	122	-	-	-	-	37	85
...
...	合計	1,325	30	8	15	12	152	1,260
05/11 12:00
...	合計	1,570	257	12	23	18	215	1,260
...

【 図 2 6 】



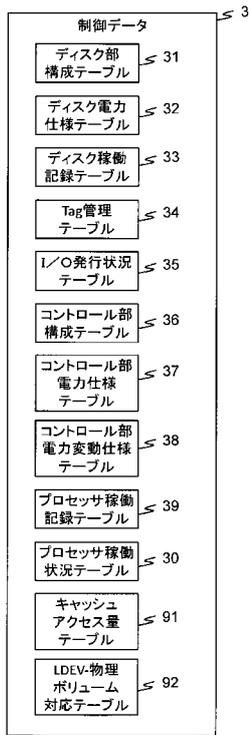
【 図 2 8 】



【 図 2 7 】

仮想ストレージ (LPAR)番号	論理ボリューム番号
0	0,1,2,3,5,7,8,12
1	4,6,9,10,11,13,14
...	...

【 図 2 9 】



【 図 3 0 】

921 N	922 N	923 N	924 N
論理ボリューム番号	RAIDレベル	ディスク番号	LBA範囲
0	1	0	0-20971519
		1	0-20971519
	
1	5	4	0-16777215
		5	0-16777215
		6	0-16777215
		7	0-16777215
2	1	0	20971520-37748735
		1	20971520-37748735
...

【 図 3 1 】

331 N	336 N	332 N	333 N	334 N	335 N
ディスク番号	論理ボリューム番号	Random Read 累計稼働時間[μs]	Random Write 累計稼働時間[μs]	Sequential Read 累計稼働時間[μs]	Sequential Write 累計稼働時間[μs]
#0	0	43,705	6,385	13,980	3,376
	1	40,273	2,013	18,534	3,482

#1	0	9,384	46,095	5,020	107

...

【 図 3 2 】

341 N	342 N	344 N	343 N
ディスク番号	Tag番号	論理ボリューム番号	アクセス種別
#0	0	0	Random Read
	1	1	Random Read
	2	-	なし

#1	0	0	Sequential Read
	1	0	Sequential Read
...

【 図 3 3 】

351 N	355 N	352 N	353 N	354 N
ディスク番号	論理ボリューム番号	アクセス種別	現在発行中コマンド数	ディスク稼働開始時刻
#0	0	Random Read	5	8,421652
		Random Write	0	-
		Sequential Read	10	8,421120
#1	0	Sequential Write	0	-
		Random Read	0	-
#1	0	Random Write	3	8,421573
	

【図 3 4】

391 N	392 N	394 N	393 N
制御部番号	プロセッサコア番号	論理ボリューム番号	累計稼働時間[μs]
0	0	0	374,105
		1	374,120
	1	0	398,752
		1	25,182
	2	0	268,675
		1	240,357
3	0	42,970	
	1	60,723	
1	0	0	75,804
		1	9,750
	1	0	40,795
1	1	5,972	5,972
...

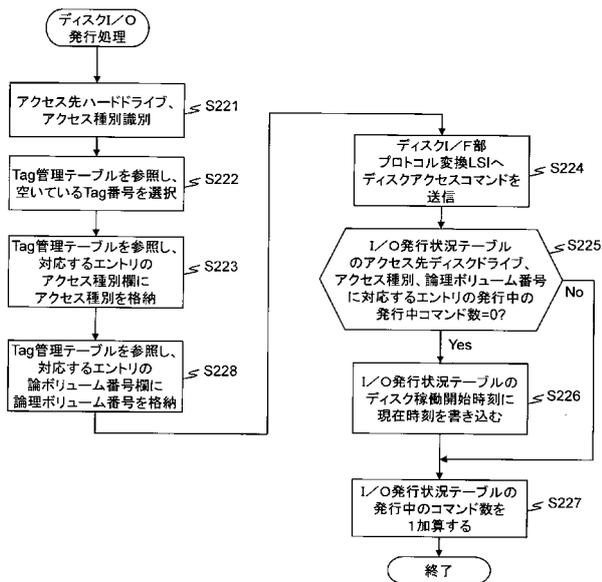
【図 3 5】

301 N	302 N	304 N	303 N
制御部番号	プロセッサコア番号	論理ボリューム番号	ジョブ開始時刻
0	0	0	8,421,652
	1	-	-
	2	1	8,421,210
	3	-	-
1	0	-	-
	1	0	8,421,573
...

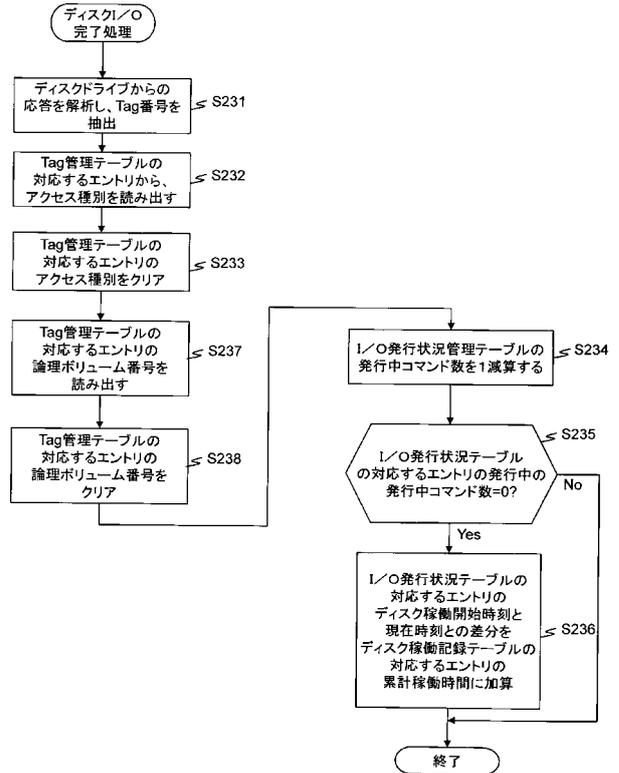
【図 3 6】

911 N	912 N
論理ボリューム番号	累積アクセス量[B]
0	1,653,688,320
1	833,028,096
...	...

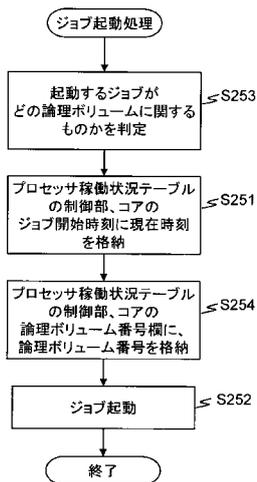
【図 3 7】



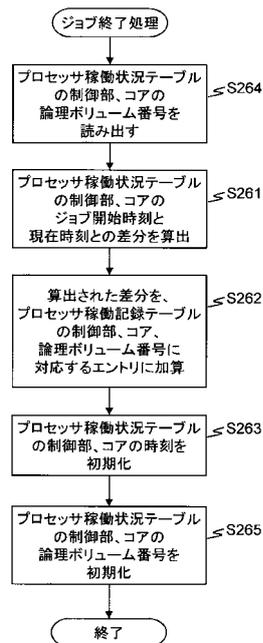
【図 3 8】



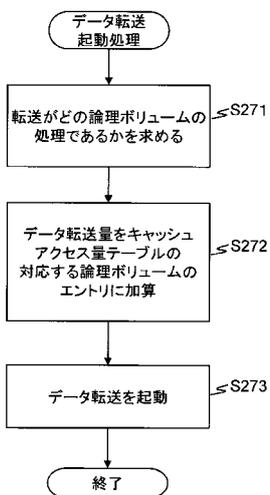
【 図 3 9 】



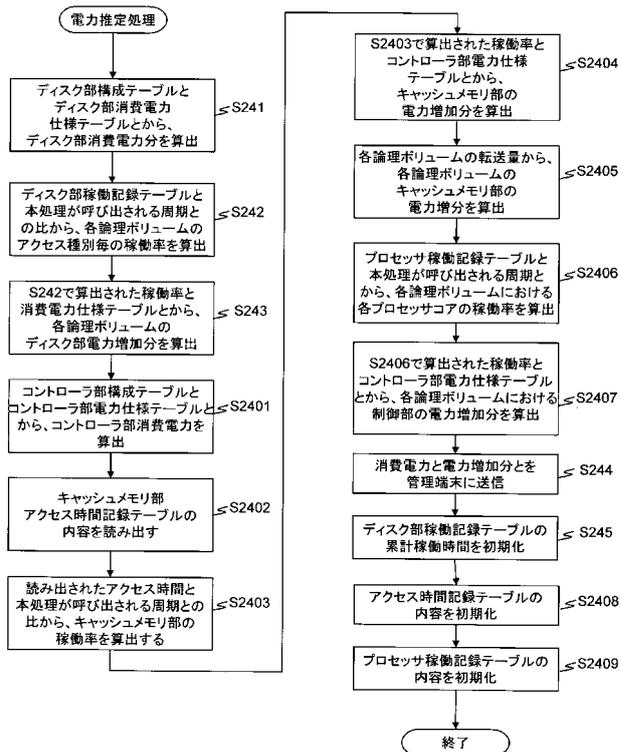
【 図 4 0 】



【 図 4 1 】



【 図 4 2 】



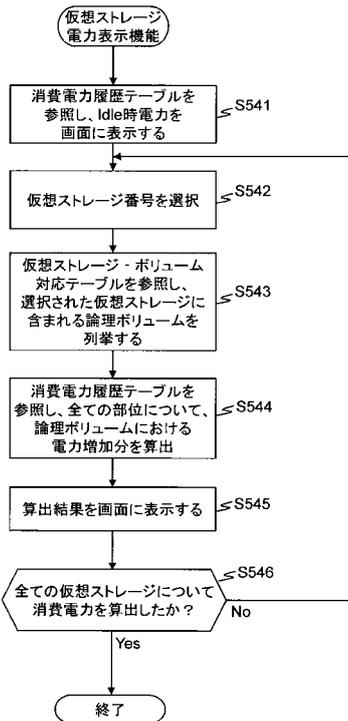
【 図 4 3 】

部位	991 N	992 N	994 N	995 N	993 N				コントローラ部電力増加[W]
	平均消費電力[W]	Idle時電力[W]	論理ボリューム番号	Random Read電力増加[W]	Random Write電力増加[W]	Sequential Read電力増加[W]	Sequential Write電力増加[W]	電力増加[W]	
ディスク0	18.0	14.1	0	3.7	0.2	0.0	0.0	-	
ディスク1	
DIMM 0	3.7	2.5	0	-	-	-	-	0.5	
...	
合計	1,325	1,260	-	30	18	0	12	15	

【 図 4 4 】

日時	部位	平均消費電力[W]	論理ボリューム番号	電力増加[W]				コントローラ部電力増加[W]	Idle時電力[W]
				Random Read電力増加[W]	Random Write電力増加[W]	Sequential Read電力増加[W]	Sequential Write電力増加[W]		
05/11 10:00	ディスク0	18	0	3.7	0.2	0	0	-	14.1
	制御部0	122	0	-	-	-	-	3.7	...
-11:00	85
	合計	1,325	...	30	8	15	12	152	1,260
05/11 -12:00
	合計	1,570	...	257	12	23	18	215	1,260

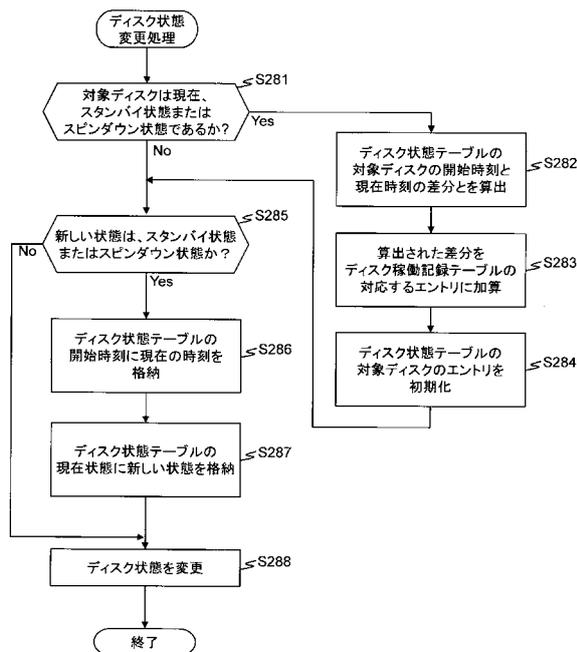
【 図 4 5 】



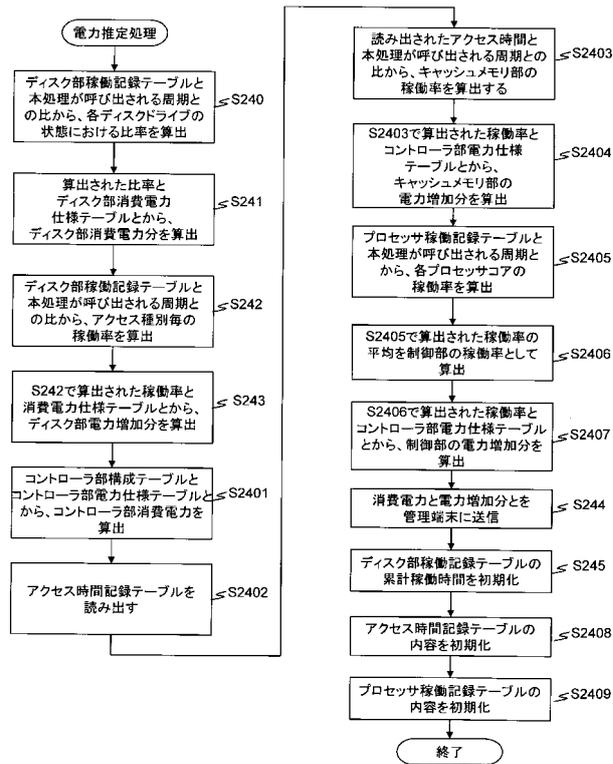
【 図 4 6 】

	591 N	592 N	...
	定常電力	仮想ストレージ0	
合計	1500[W]	120[W]	...
ディスク0	12.1[W]	1.8[W]	...
制御部0	65[W]	27[W]	...
...
		593 N	594 N
		更新	終了

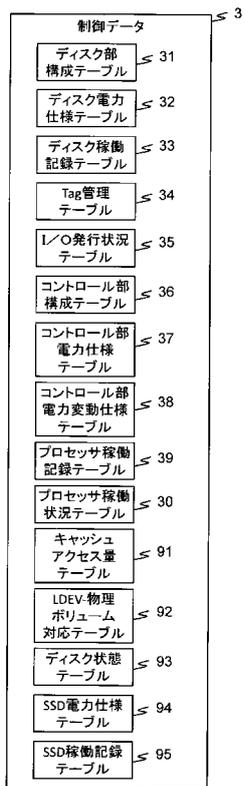
【図 5 2】



【図 5 3】



【図 5 4】



【図 5 5】

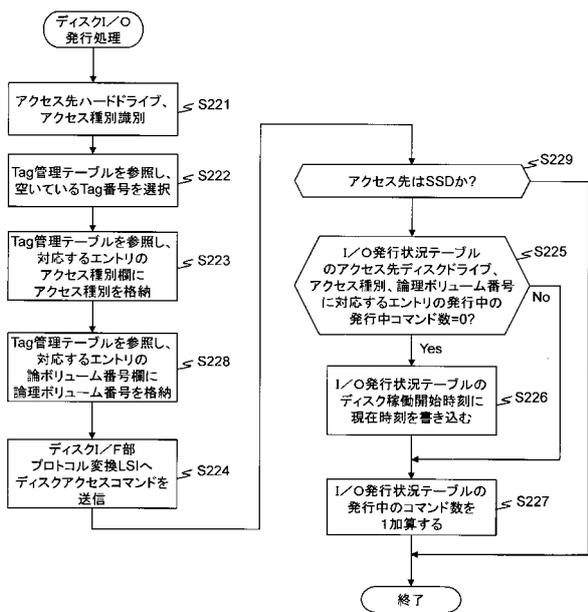
941 N	942 N	943 N	944 N	945 N	946 N	94 N
ディスク型番	Idle時電力[W]	Random Read時消費エネルギー[J/kB]	Random Write時消費エネルギー[J/kB]	Sequential Read時消費エネルギー[J/kB]	Sequential Write時消費エネルギー[J/kB]	
S-A	14.1	4.9	1.9	1.6	1.9	
S-B	8.9	5.1	2.3	2.0	2.3	
...

【図 5 6】

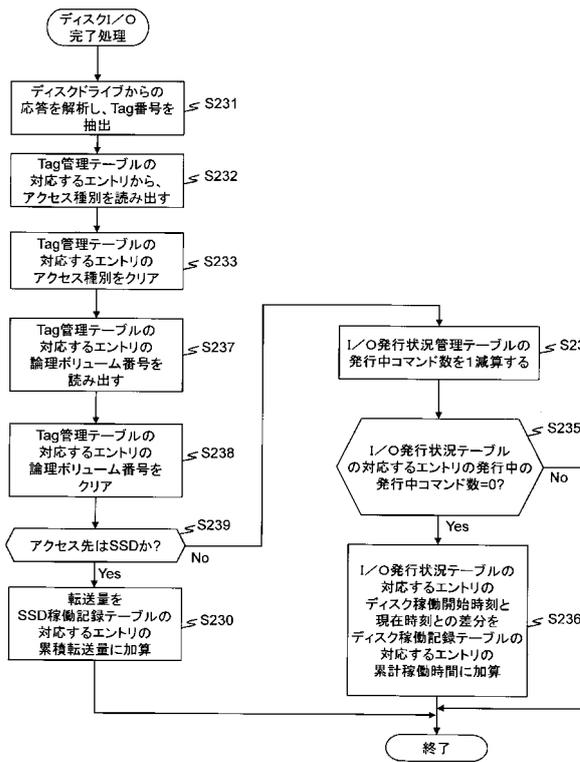
951 N	952 N	953 N	954 N	955 N	956 N	95 N
ディスク番号	論理ボリューム番号	Random Read 累計転送量 [kB]	Random Write 累計転送量 [kB]	Sequential Read 累計転送量 [kB]	Sequential Write 累計転送量 [kB]	
#0	0	43,705	6,385	13,980	3,376	
	1	40,273	2,013	18,534	3,482	

#1	0	9,384	46,095	5,020	107	

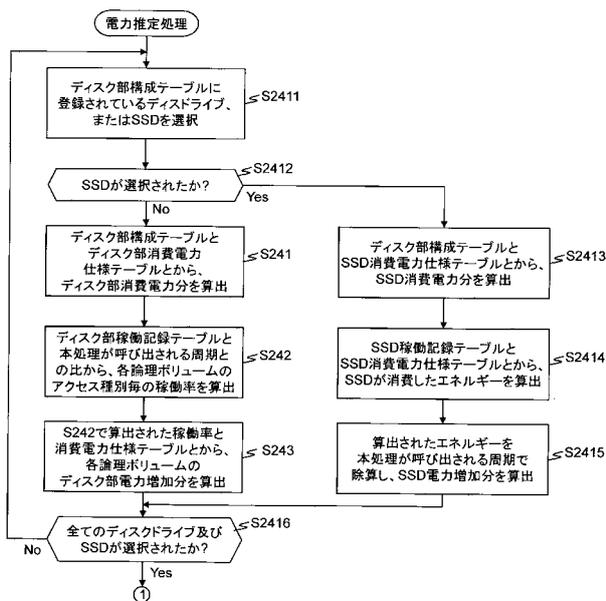
【図57】



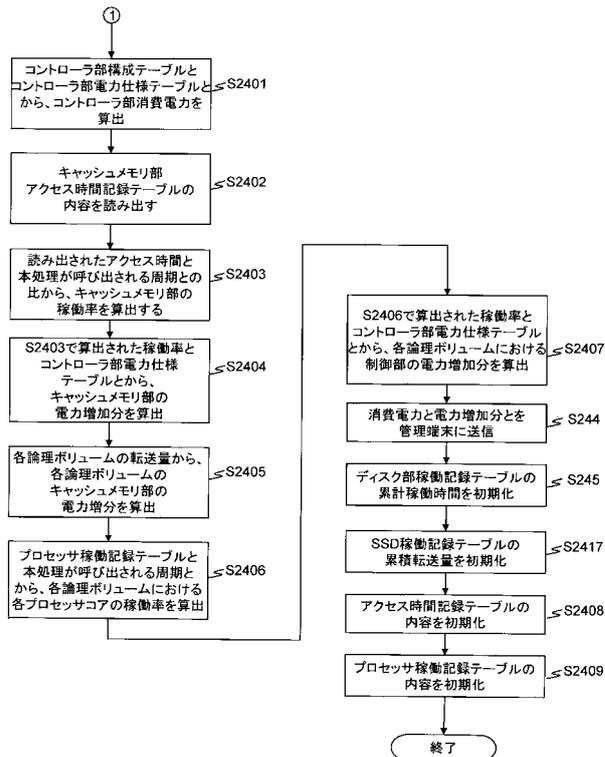
【図58】



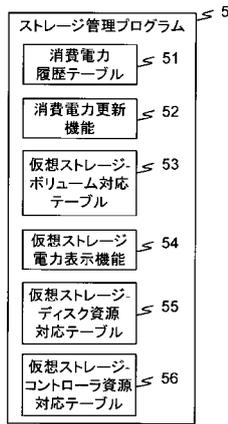
【図59A】



【図59B】



【図60】



【図61】

ディスク番号	利用仮想ストレージ番号
#0	0,1,2
#1	0,1
#2	0
#3	0
#4	2
#5	1
#6	1
...	...

【図62】

種別	スロット番号	利用仮想ストレージ番号
制御部	0	0,1,2
	1	0,1

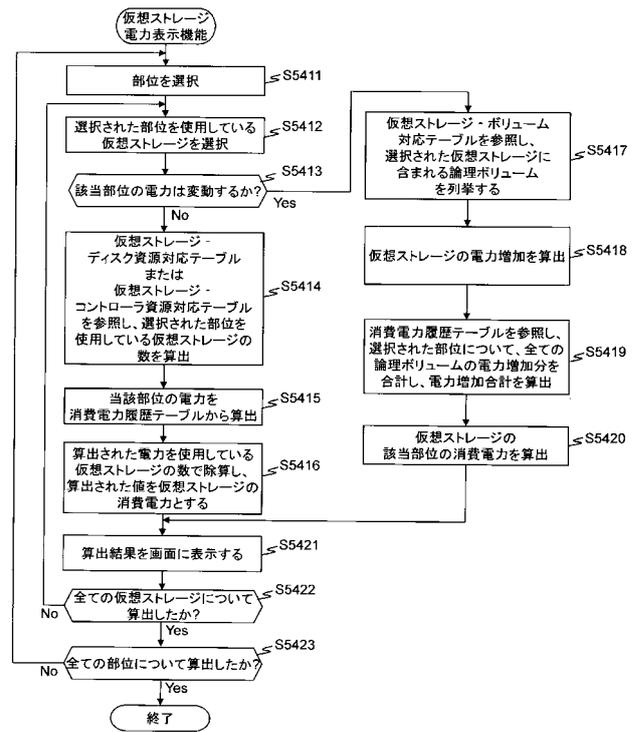
キャッシュメモリ部	0	0,1,2
	1	0

ホストI/F	0	0
	1	1,2

ディスクI/F	0	0,1
	1	1,2

キャッシュメモリメモリモジュール	0	0,1,2
	1	0,1

【図63】



【 図 6 4 】

仮想ストレージ消費電力			
	592A N	592B N	
	仮想ストレージ0	仮想ストレージ1	...
合計	273[W]	263[W]	...
ディスク0	3[W]	0[W]
制御部0	33[W]	15[W]
....
		593 N	594 N
	<input type="button" value="更新"/>		<input type="button" value="終了"/>

フロントページの続き

(72)発明者 中村 崇仁

神奈川県川崎市麻生区王禅寺 1 0 9 9 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) 5B011 EB07 GG00 HH01

5B065 BA01 CA30 ZA14