

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2022-63148

(P2022-63148A)

(43)公開日 令和4年4月21日(2022.4.21)

(51)国際特許分類		F I		テーマコード(参考)	
G 0 6 F	11/07 (2006.01)	G 0 6 F	11/07	1 9 3	5 B 0 4 2
G 0 6 F	13/10 (2006.01)	G 0 6 F	13/10	3 4 0 A	
G 0 6 F	13/14 (2006.01)	G 0 6 F	13/14	3 3 0 A	
G 0 6 F	3/06 (2006.01)	G 0 6 F	13/14	3 3 0 F	
G 0 6 F	9/50 (2006.01)	G 0 6 F	3/06	3 0 1 Z	
		審査請求	未請求	請求項の数	15 O L (全40頁) 最終頁に続く

(21)出願番号	特願2020-171539(P2020-171539)	(71)出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22)出願日	令和2年10月9日(2020.10.9)	(74)代理人	110002365 特許業務法人サンネクスト国際特許事務所
		(72)発明者	柴山 司 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	弘中 和衛 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		(72)発明者	佐藤 賢太 東京都千代田区丸の内一丁目6番6号 株式会社日立製作所内
		最終頁に続く	

(54)【発明の名称】 計算機システム及び計算機システム運用管理方法

(57)【要約】

【課題】

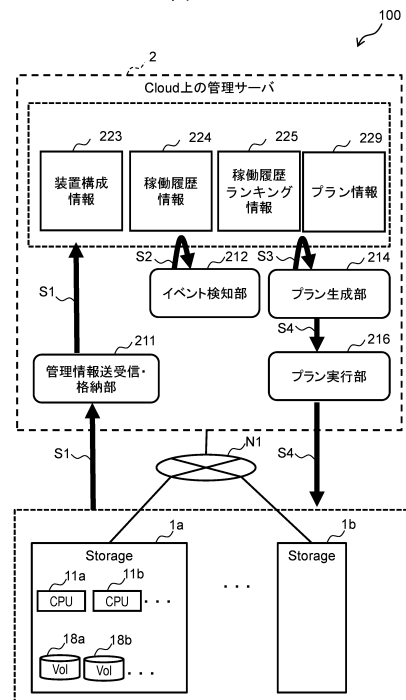
ストレージ装置の複数のプロセッサで負荷に偏りが生じた場合に、オーナー権の変更処理による影響を考慮した上で、計算機システムの運用に与える影響を抑制しながらオーナー権の変更処理を実行する。

【解決手段】

計算機システム100は、1以上のストレージ装置1と管理計算機(管理サーバ2)を含み、管理計算機は、情報収集部(管理情報送受信・格納部211)、イベント検知部212、プラン生成部214、及びプラン実行部216を有する。そしてプラン生成部214は、プランにおける制御権の変更処理の対象ボリューム、制御権の移動元のプロセッサ、及び制御権の移動先のプロセッサを決定し、当該プランにおける制御権の変更処理による影響を見積もり、プラン実行部216は、影響の見積もり及びストレージ装置の稼働情報に基づいて、プランの実行時刻を決定する。

【選択図】図1

図 1



【特許請求の範囲】

【請求項 1】

プロセッサとメモリとキャッシュメモリとを有して、上位システムがデータを入出力する、1以上の記憶デバイスにより構成される1以上のボリュームを提供する1以上のストレージ装置と、

前記ストレージ装置を管理する管理計算機と、

を含む計算機システムであって、

前記管理計算機は、

前記ストレージ装置における構成情報及び稼働情報を含む運用情報を収集する情報収集部と、

前記収集された運用情報から、所定の検知条件に基づいて、前記プロセッサの負荷の不均衡を示すイベントを検知するイベント検知部と、

前記検知されたイベントを解決する、前記ボリュームに対する前記プロセッサの制御権を変更するプランを生成するプラン生成部と、

前記生成されたプランを実行するプラン実行部と、

を備え、

前記稼働情報には、前記ストレージ装置における前記プロセッサ、前記ボリューム、及び前記キャッシュメモリの稼働履歴を示す情報が含まれ、

前記プラン生成部は、前記プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定し、当該プランにおける前記制御権の変更処理による影響を見積もり、

前記プラン実行部は、前記影響の見積もり及び前記ストレージ装置の前記稼働情報に基づいて、前記プランの実行時刻を決定する

ことを特徴とする計算機システム。

【請求項 2】

前記検知条件は、基準値の超過を検知する時間間隔である検知間隔と、前記基準値の超過を検知する検知期間と、前記検知期間の間に前記基準値の超過を検知した検知回数と、を含んで構成され、

前記イベント検知部は、前記イベントを検知する際に、前記検知間隔を用いて、検知処理の時間方向のダウンサンプリングを実施する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 3】

前記イベント検知部は、前記イベントを検知する際に、前記検知間隔による時間方向のダウンサンプリングと前記検知期間と前記検知回数とを用いて、前記プロセッサの負荷の不均衡の発生傾向を検知する

ことを特徴とする請求項 2 に記載の計算機システム。

【請求項 4】

前記稼働情報には、前記ボリュームの稼働履歴において負荷の高いボリュームの順位及びその負荷量を保持するボリューム稼働履歴ランキング情報がさらに含まれ、

前記プラン生成部は、前記ボリューム稼働履歴ランキング情報から、前記プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 5】

前記プラン生成部は、前記制御権の変更処理による影響として、キャッシュメモリ容量、プロセッサ負荷、キャッシュメモリの書き込み待ち割合、及びストレージ装置のハードウェアスペックに基づいて、当該変更処理の処理時間を予測する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 6】

前記プラン生成部は、前記制御権の変更処理による影響として、当該変更処理の実行中の

10

20

30

40

50

I/O 負荷を予測する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 7】

前記プラン生成部は、前記制御権の変更処理による影響として、当該変更処理の実行中のプロセッサ負荷を予測する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 8】

前記プラン生成部は、前記プランを生成する際、前記ボリュームに直接紐付かない負荷を他のリソースの稼働情報から見積もる処理を踏まえて、当該プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 9】

前記プラン生成部は、負荷の見積もりとして、過去の検知点における最大値を採用することを特徴とする請求項 1 に記載の計算機システム。

【請求項 10】

前記管理計算機は、前記プラン生成部が生成した前記プランについて、実行の希望時刻または禁止時刻のユーザによる指定を受け付けるインターフェースをさらに備え、前記インターフェースが前記ユーザによる指定を受け付けた場合、前記プラン実行部は、前記影響の見積もり及び前記ストレージ装置の稼働情報と、前記ユーザによって指定された前記希望時刻または前記禁止時刻とを考慮して、前記プランの実行時刻を決定することを特徴とする請求項 1 に記載の計算機システム。

【請求項 11】

前記管理計算機は、前記プラン実行部による前記プランの実行後に、前記ストレージ装置における稼働情報に基づいて、前記イベントが解決したか否かを確認するプラン実行完了確認部をさらに備える

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 12】

前記検知条件は、基準値の超過を検知する時間間隔である検知間隔と、前記基準値の超過を検知する検知期間と、前記検知期間の間に前記基準値の超過を検知した検知回数と、を含んで構成され、

前記プラン実行完了確認部は、前記検知条件から前記検知期間をを基にした確認条件に基づいて、前記イベントが解決したか否かを確認する

ことを特徴とする請求項 11 に記載の計算機システム。

【請求項 13】

前記プラン実行完了確認部は、前記プランの実行後に前記イベントを検知した時点から一定時間が経過した場合には、前記イベントが解決したと見なす

ことを特徴とする請求項 11 に記載の計算機システム。

【請求項 14】

複数のストレージノードで 1 台のクラスタが構成され、当該複数のストレージノードの間でメモリ及びキャッシュメモリのデータを共有しない構成である場合に、

前記プラン生成部は、前記イベント検知部で検知されたイベントを解決するための前記プランの生成として、前記複数のストレージノードに跨って前記ボリュームに対する前記プロセッサの制御権を変更するプランを生成し、当該プランにおける前記制御権の変更処理による影響として、キャッシュメモリ容量、プロセッサ負荷、キャッシュメモリの書き込み待ち割合、ボリューム容量、及びストレージ装置のハードウェアスペックに基づいて、当該変更処理の処理時間を予測する

ことを特徴とする請求項 1 に記載の計算機システム。

【請求項 15】

プロセッサとメモリとキャッシュメモリとを有して、上位システムがデータを入出力する

10

20

30

40

50

、 1 以上の記憶デバイスにより構成される 1 以上のボリュームを提供する 1 以上のストレージ装置と、前記ストレージ装置を管理する管理計算機と、を含む計算機システムによる計算機システム運用管理方法であって、
 前記管理計算機が、前記ストレージ装置における構成情報及び稼働情報を含む運用情報を収集する情報収集ステップと、
 前記管理計算機が、前記情報収集ステップで収集された運用情報から、所定の検知条件に基づいて、前記プロセッサの負荷の不均衡を示すイベントを検知するイベント検知ステップと、
 前記管理計算機が、前記イベント検知ステップで検知されたイベントを解決する、前記ボリュームに対する前記プロセッサの制御権を変更するプランを生成するプラン生成ステップと、
 前記管理計算機が、前記プラン生成ステップで生成されたプランを実行するプラン実行ステップと、
 を備え、
 前記稼働情報には、前記ストレージ装置における前記プロセッサ、前記ボリューム、及び前記キャッシュメモリの稼働履歴を示す情報が含まれ、
 前記プラン生成ステップにおいて、前記プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定し、当該プランにおける前記制御権の変更処理による影響を見積もり、
 前記プラン実行ステップにおいて、前記影響の見積もり及び前記ストレージ装置の前記稼働情報に基づいて、前記プランの実行時刻を決定することを特徴とする計算機システム運用管理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、計算機システム及び計算機システム運用管理方法に関し、ボリュームを提供する 1 以上のストレージ装置における複数のプロセッサの間でボリュームの制御権（オーナー権）を変更する計算機システム及び計算機システム運用管理方法に適用して好適なものである。

【背景技術】

【0002】

近年、計算機システムなどの運用コストを削減するために、運用管理操作の自動化が進んでおり、ストレージ装置の構成情報や稼働情報を収集・分析し対処するような運用管理を、高度化または容易化する技術が普及している。特に、ストレージ装置の運用管理では、性能負荷を監視して適切な対処を実施するためには高度な運用スキルが必要となるため、運用管理を容易化する様々な技術が開示されている。

【0003】

例えば特許文献 1 には、ストレージ装置から稼働情報を収集して負荷を監視し、ストレージ装置においてプロセッサ間の負荷が不均衡になった場合に、負荷の高いプロセッサから負荷の低いプロセッサへプロセッサ毎のボリュームの制御権（オーナー権）を移動する構成変更案を作成し、実行する技術が開示されている。また例えば、特許文献 2 には、ストレージ装置の追加や閉塞等の構成変更を検知した場合に、プロセッサが持つボリュームのオーナー権を適当なプロセッサへ変更する技術が開示されている。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特許第 6 2 3 5 1 5 6 号公報

【特許文献 2】特表 2 0 1 6 - 5 1 0 4 4 0 号公報

【発明の概要】

【発明が解決しようとする課題】

10

20

30

40

50

【 0 0 0 5 】

ところで、オーナー権の変更処理はそれ自体で負荷が発生するが、上述した従来技術ではオーナー権の変更処理による影響が考慮されていないため、変更中に問題が発生するおそれがあった。また、上述した従来技術では、オーナー権の変更処理を実行する時刻について、アプリケーションの稼働やユーザ操作による計算機システムの運用への影響を考慮しておらず、例えば、ユーザが意図しない時刻に、意図しないオーナー権の変更処理が実行されてしまう可能性もあった。

【 0 0 0 6 】

本発明は以上の点を考慮してなされたもので、計算機システムにおいてストレージ装置の複数のプロセッサで負荷に偏りが生じた場合に、オーナー権の変更処理による影響を考慮した上で、計算機システムの運用に与える影響を抑制しながらオーナー権の変更処理を実行することができる計算機システム及び計算機システム運用管理方法を提案しようとするものである。

10

【課題を解決するための手段】

【 0 0 0 7 】

かかる課題を解決するため本発明においては、プロセッサとメモリとキャッシュメモリとを有して、上位システムがデータを入出力する、1以上の記憶デバイスにより構成される1以上のボリュームを提供する1以上のストレージ装置と、前記ストレージ装置を管理する管理計算機と、を含む計算機システムであって、前記管理計算機は、前記ストレージ装置における構成情報及び稼働情報を含む運用情報を収集する情報収集部と、前記収集された運用情報から、所定の検知条件に基づいて、前記プロセッサの負荷の不均衡を示すイベントを検知するイベント検知部と、前記検知されたイベントを解決する、前記ボリュームに対する前記プロセッサの制御権を変更するプランを生成するプラン生成部と、前記生成されたプランを実行するプラン実行部と、を備え、前記稼働情報には、前記ストレージ装置における前記プロセッサ、前記ボリューム、及び前記キャッシュメモリの稼働履歴を示す情報が含まれ、前記プラン生成部は、前記プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定し、当該プランにおける前記制御権の変更処理による影響を見積もり、前記プラン実行部は、前記影響の見積もり及び前記ストレージ装置の前記稼働情報に基づいて、前記プランの実行時刻を決定する計算機システムが提供される。

20

30

【 0 0 0 8 】

また、かかる課題を解決するため本発明においては、プロセッサとメモリとキャッシュメモリとを有して、上位システムがデータを入出力する、1以上の記憶デバイスにより構成される1以上のボリュームを提供する1以上のストレージ装置と、前記ストレージ装置を管理する管理計算機と、を含む計算機システムによる以下の計算機システム運用管理方法が提供される。この計算機システム運用管理方法は、前記管理計算機が、前記ストレージ装置における構成情報及び稼働情報を含む運用情報を収集する情報収集ステップと、前記管理計算機が、前記情報収集ステップで収集された運用情報から、所定の検知条件に基づいて、前記プロセッサの負荷の不均衡を示すイベントを検知するイベント検知ステップと、前記管理計算機が、前記イベント検知ステップで検知されたイベントを解決する、前記ボリュームに対する前記プロセッサの制御権を変更するプランを生成するプラン生成ステップと、前記管理計算機が、前記プラン生成ステップで生成されたプランを実行するプラン実行ステップと、を備え、前記稼働情報には、前記ストレージ装置における前記プロセッサ、前記ボリューム、及び前記キャッシュメモリの稼働履歴を示す情報が含まれ、前記プラン生成ステップにおいて、前記プランにおける前記制御権の変更処理の対象ボリューム、前記制御権の移動元のプロセッサ、及び前記制御権の移動先のプロセッサを決定し、当該プランにおける前記制御権の変更処理による影響を見積もり、前記プラン実行ステップにおいて、前記影響の見積もり及び前記ストレージ装置の前記稼働情報に基づいて、前記プランの実行時刻を決定する。

40

【発明の効果】

50

【 0 0 0 9 】

本発明によれば、計算機システムにおいてストレージ装置の複数のプロセッサで負荷に偏りが生じた場合に、オーナー権の変更処理による影響を考慮した上で、計算機システムの運用に与える影響を抑制しながらオーナー権の変更処理を実行することができる。

【 図面の簡単な説明 】

【 0 0 1 0 】

【 図 1 】 本発明の第 1 の実施形態における構成及び処理の概要を説明する図である。

【 図 2 】 計算機システム 1 0 0 の全体構成例を示す図である。

【 図 3 】 ストレージ装置 1 の構成例を示す図である。

【 図 4 】 ストレージ装置 1 のメモリ 1 2 上のプログラム及び情報を示す図である。 10

【 図 5 】 クラウド上の管理サーバ 2 に保持されるプログラム及び情報を示す図である。

【 図 6 】 稼働情報 1 2 5 の一例を示す図である。

【 図 7 】 構成情報 1 2 6 の一例を示す図である。

【 図 8 】 オーナー権管理情報 1 2 7 の一例を示す図である。

【 図 9 】 キャッシュ管理情報 1 2 8 の一例を示す図である。

【 図 1 0 】 組織情報 2 2 1 の一例を示す図である。

【 図 1 1 】 装置情報 2 2 2 の一例を示す図である。

【 図 1 2 】 装置構成情報 2 2 3 の一例を示す図である。

【 図 1 3 】 稼働履歴情報 2 2 4 の一例を示す図である。

【 図 1 4 】 稼働履歴ランキング情報 2 2 5 の一例を示す図である。 20

【 図 1 5 】 検知情報 2 2 6 の一例を示す図である。

【 図 1 6 】 性能異常履歴情報 2 2 7 の一例を示す図である。

【 図 1 7 】 イベント情報 2 2 8 の一例を示す図である。

【 図 1 8 】 プラン情報 2 2 9 の一例を示す図である。

【 図 1 9 】 ジョブ情報 2 3 0 の一例を示す図である。

【 図 2 0 】 第 1 の実施形態における全体的な処理の処理手順例を示すフローチャートである。

【 図 2 1 】 情報収集処理の処理手順例を示すフローチャートである。

【 図 2 2 】 検知処理の処理手順例を示すフローチャートである。

【 図 2 3 】 プラン生成処理の処理手順例を示すフローチャート（その 1 ）である。 30

【 図 2 4 】 プラン生成処理の処理手順例を示すフローチャート（その 2 ）である。

【 図 2 5 】 プラン実行処理の処理手順例を示すフローチャート（その 1 ）である。

【 図 2 6 】 プラン実行処理の処理手順例を示すフローチャート（その 2 ）である。

【 図 2 7 】 本発明の第 2 の実施形態に係る計算機システム 2 0 0 の全体構成例を示す図である。

【 図 2 8 】 第 2 の実施形態における構成情報 1 2 6 ' の一例を示す図である。

【 図 2 9 】 第 2 の実施形態における装置構成情報 2 2 3 ' の一例を示す図である。

【 図 3 0 】 第 2 の実施形態におけるプラン生成処理の処理手順例を示すフローチャートである。

【 発明を実施するための形態 】 40

【 0 0 1 1 】

以下、図面を参照して、本発明の好適な実施形態を説明する。以下の説明において、同一または類似の要素及び処理に対しては、同一の符号を付して差分を説明し、重複する説明を省略する。また、後出の実施形態では、既出の実施形態との差分を説明し、重複する説明を省略する。

【 0 0 1 2 】

また、以下の説明及び各図によって示される構成及び処理は、本発明の理解及び実施に必要な程度で実施形態の概要を例示するものであり、本発明に係る実施の態様を限定することを意図する趣旨ではない。また、各実施形態及び各変形例は、本発明の趣旨を逸脱せず、整合する範囲内で、一部又は全部を組み合わせることができる。

【 0 0 1 3 】

以下の説明においては、数字に付加された添字や枝番号によって区別される符号が付与された類似の要素を、数字のみの符号によって、添え字や枝番号に関係なく総称する。例えば「100a」「100b」や「200-1」「200-2」といった符号が付与された要素を、「100」や「200」といった符号を付与して総称する。また、「XXインターフェース14a」「YYインターフェース14b」といった、数字に添え字や枝番が付加された符号が付与された類似の要素を、「インターフェース14」のように、要素の名称の共通部分と数字のみの符号部分を以って総称する。

【 0 0 1 4 】

また、以下の説明においては、各種情報を表（テーブル）形式にて説明するが、情報は表形式に限らず、ドキュメント形式やその他の形式であってもよい。また、表の構成は一例であり、表は適宜、統合または分散できる。また、以下の説明において、各表の項目（カラム）として挙げられているIDや名前は、レコードを区別可能であれば、番号及び文字列の何れでもよい。

10

【 0 0 1 5 】

また、以下では、「プログラム」を主語として処理を説明する場合がある。プログラムは、プロセッサ（例えばCPU（Central Processing Unit））によって実行されることで、定められた処理を、適宜に記憶デバイス（例えばメモリ）及び/又は通信インターフェイスデバイス（例えば通信ポート）等を用いながら行うため、処理の主語がプロセッサとされてもよい。プログラムを主語として説明された処理は、プロセッサあるいはその

20

【 0 0 1 6 】

また、プログラムを実行するプロセッサを、目的の処理機能を実現する装置として「XX部」と呼ぶこともできる。また、プロセッサは、処理の一部又は全部を行うハードウェア回路を含んでもよい。プログラムは、プログラムソースから各コントローラにインストールされてもよい。プログラムソースは、例えば、プログラム配布計算機または計算機が読み取り可能な記憶メディアであってもよい。

【 0 0 1 7 】

(1) 第1の実施形態

(1-1) 概要

図1は、本発明の第1の実施形態における構成及び処理の概要を説明する図である。図1に示した計算機システム100は、本実施形態に係る計算機システムの一例であって、1以上のストレージ装置1（個別にはストレージ装置1a, 1b）と、クラウド上の管理サーバ2と、を備える。

30

【 0 0 1 8 】

ストレージ装置1はそれぞれ、複数のCPU11（個別にはCPU11a, 11b）とボリューム18（個別にはボリューム18a, 18b）とを有する。各ストレージ装置1a, 1bとクラウド上の管理サーバ2は、ネットワークN1によって通信可能に接続される。

【 0 0 1 9 】

管理サーバ2は、例えばクラウド上に構成されており、管理情報送受信・格納部211、イベント検知部212、プラン生成部214、プラン実行部216、装置構成情報223、稼働履歴情報224、稼働履歴ランキング情報225、及びプラン情報229を有する。なお、管理サーバ2は、上記以外の構成要素も有しており、これら全体については、図5を参照しながら後述する。

40

【 0 0 2 0 】

図1には、本実施形態における全体処理の概要的な処理手順として、ステップS1～S4が示されている。

【 0 0 2 1 】

ステップS1は、構成情報や稼働情報等の運用情報を収集する処理（情報収集処理）であ

50

る。ステップ S 1 では、管理情報送受信・格納部 2 1 1 が、ストレージ装置 1 における構成情報（CPU 1 1 とボリュームとの制御関係を示すオーナー権に関する情報を含む）及び稼働情報を、ストレージ装置 1 から収集し、クラウド上の管理サーバ 2 に保持する。

【0022】

ステップ S 2 は、ステップ S 1 で収集した運用情報に基づいて、イベントを検知する処理（検知処理）である。ステップ S 2 では、イベント検知部 2 1 2 が、装置構成情報 2 2 3、稼働履歴情報 2 2 4、及び稼働履歴ランキング情報 2 2 5 を参照し、イベントを検知する。ここで「イベント」とは、構成情報や稼働情報が予め定められた所定の検知条件に合致したことを意味する。この検知条件は、具体的には例えば、ストレージ装置 1 における「性能異常」の発生を検知するための条件であって、詳細については図 1 5 の説明等において後述する。

10

【0023】

ステップ S 3 は、ステップ S 2 によるイベントの検知を受けて、イベントを対処するプランを生成する処理（プラン生成処理）である。ステップ S 3 では、プラン生成部 2 1 4 がイベントを解決するプランを立案し、プラン情報 2 2 9 に保存する。ここで「プラン」とは、イベントを解決するための処理内容や時刻情報等を含む情報であって、イベントを解決するための処理として、ボリュームの制御権（オーナー権）を変更する処理が含まれる。

【0024】

ステップ S 4 は、ステップ S 3 で生成されたプランを実行する処理（プラン実行処理）である。ステップ S 4 では、プラン実行部 2 1 6 が、生成されたプランを取得し、ストレージ装置 1 に対して適当な時刻で、当該プランに記載された処理を実行する。

20

【0025】

上記ステップ S 1 ~ S 4 の処理の実行を通じて、本実施形態に係る計算機システム 1 0 0 は、オーナー権変更処理の影響を考慮した上で、その実行時刻を決定することができ、運用管理者の管理コストを低減しつつ適切にオーナー権変更の処理を実行することができる。

【0026】

(1-2) 構成

図 2 は、計算機システム 1 0 0 の全体構成例を示す図である。図 2 では、図 1 よりも詳細に計算機システム 1 0 0 の構成を示しており、計算機システム 1 0 0 は、クラウド上の管理サーバ 2 と複数のデータセンタ 3（個別にはデータセンタ 3 a , 3 b）とを備える。管理サーバ 2 とデータセンタ 3 a , 3 b は、データセンタ外のネットワーク N 1 で接続される。

30

【0027】

また、図 2 において、データセンタ 3 はそれぞれ、1 以上のストレージ装置 1（個別には例えばストレージ装置 1 a , 1 b , 1 c）を有し、1 つのデータセンタ 3 内において複数のストレージ装置の間は、データセンタ内のネットワーク N 2 で接続される。また、各データセンタ 3 は、データセンタ内のネットワーク N 2 をデータセンタ外のネットワーク N 1 に接続するためのスイッチ 4 を有する。なお、スイッチ 4 は必ずしも必要ではなく、データセンタ内のネットワーク N 2 とデータセンタ外のネットワーク N 1 とが同一ネットワークであってもよい。また、これらのネットワーク N 1 , N 2 は冗長化された構成であってもよい。また、これらのネットワーク N 1 , N 2 は、イーサネット（Ethernet（登録商標））、インフィニバンド（InfiniBand）、または無線等の何れの通信手段であってもよい。また、図 2 では、ストレージ装置 1 a , 1 b , 1 c やデータセンタ 3 a , 3 b の数が固定数となっているが、本実施形態においてストレージ装置 1 及びデータセンタ 3 の数は上記固定数に限定されるものではない。

40

【0028】

クラウド上の管理サーバ 2 は、ストレージ装置 1 の管理情報を保持したり、ストレージ装置 1 を操作するためのプログラムを保持したりするための管理サーバである。例えば、管理サーバ 2 は、プログラムを介してストレージ装置 1 a , 1 b , 1 c から情報を収集したり、処理を実行したり、GUI（Graphical User Interface）を介してユーザ（管理

50

者)へ管理情報を表示し、管理者から入力された操作リクエストをストレージ装置1 a, 1 b, 1 cに送信したりする。

【0029】

なお、本実施形態では、クラウド上の1つの管理サーバ2が、全てのデータセンタ3が有する複数のストレージ装置1をまとめて管理する構成として説明するが、例えばクラウド上に複数の管理サーバ2が存在し、それぞれの管理サーバ2が各データセンタを管理するといった構成であってもよい。

【0030】

図3は、ストレージ装置1の構成例を示す図である。図3に示したように、ストレージ装置1は、プロセッサの一例であるCPU11(個別にはCPU11a, 11b)と、記憶デバイスの一例であるメモリ12、ドライブ13(個別にはドライブ13a, 13b, 13c)、及びキャッシュメモリ16と、通信インターフェイスデバイスの一例であるネットワークI/F14(個別にはネットワークI/F14a, 14b)とを有し、それぞれの構成がバス15で相互に接続されている。また、ストレージ装置1は、ドライブ13の物理領域を仮想的に束ねる論理領域としてプール17(個別にはプール17a, 17b)を有し、プール17にデータをRead/Writeする場合に論理的にアクセスする領域としてボリューム18(個別にはボリューム18a, 18b, 18c)を有する。

10

【0031】

なお、本実施形態に係るストレージ装置1において、CPU11、メモリ12、キャッシュメモリ16の数は、図3に例示した数に限定されない。また、ドライブ13の数、及び論理的な領域であるプール17やボリューム18の数も、図3に例示した数に限定されない。また、メモリ12とキャッシュメモリ16は物理的に分かれていなくてもよく、共通の物理領域を論理的に分割して、それぞれをメモリ領域またはキャッシュメモリ領域として利用してもよい。また、CPU11内には、複数のコアが含まれていてもよい。

20

【0032】

ドライブ13は、物理的な記憶デバイスであり、具体的には、HDD(Hard Disk Drive)やSSD(Solid State Drive)、またはその他の不揮発メモリ(例えばSCM(Storage Class Memory))等の何れであってもよい。

【0033】

ネットワークI/F14aは、データセンタ内のネットワークN2に接続するインターフェースであり、ネットワークI/F14bは、データにアクセスするホスト(図示省略)と接続する入出力(I/O)データ用ネットワーク(図示省略)に接続するインターフェースである。なお、ネットワークI/F14a, 14bは、ファイバチャネル(Fibre Channel)やイーサネット等の何れのインターフェースでもよい。また、データセンタ内のネットワークN2とI/Oデータ用ネットワークとが同じ場合は、ネットワークI/F14a, 14bが同じであってもよい。図3では、ネットワークI/F14a, 14bをそれぞれ1つ示したが、ネットワークI/F14a, 14bの数はこれに限定されない。

30

【0034】

バス15は、CPU11、メモリ12、ドライブ13、ネットワークI/F14、及びキャッシュメモリ16を接続する内部バスである。

40

【0035】

キャッシュメモリ16は、ドライブ13にアクセスするデータに高速にアクセスするために、一時的にデータをキャッシュする領域であって、以降は省略してキャッシュと称することがある。

【0036】

図4は、ストレージ装置1のメモリ12上のプログラム及び情報を示す図である。図4に示したように、メモリ12には、管理情報送受信部121、稼働情報収集・格納部122、管理操作実行部123、キャッシュ管理部124、稼働情報125、構成情報126、オーナー権管理情報127、及びキャッシュ管理情報128が記憶されている。

50

【0037】

管理情報送受信部121は、ストレージ装置1a, 1b, 1cが保持する稼働情報125、構成情報126、及びオーナー権管理情報127を定期的に管理サーバ2に送信したり、管理サーバ2から送信される操作のリクエストを受信したりするプログラムである。なお、管理情報送受信部121と管理サーバ2との間のやり取りは、Push/Pollingの何れでもよい。

【0038】

稼働情報収集・格納部122は、ストレージ装置1a, 1b, 1cの内部で発生するI/O (Input / Output) 処理等によるI/Oの量、データ転送量、及び、ストレージ装置1内の各リソースの使用量を定期的に収集し、収集した情報を稼働情報125に格納するプログラムである。

10

【0039】

管理操作実行部123は、管理サーバ2から送信されるストレージ装置1の管理操作を受け付けて実行するプログラムである。また、管理操作実行部123は、上記管理操作の内容に従って、構成情報126やオーナー権管理情報127の情報を更新する。

【0040】

キャッシュ管理部124は、ストレージ装置1a, 1b, 1c内のキャッシュ16を管理するプログラムである。また、キャッシュ管理部124は、管理操作やI/Oの内容に従って、キャッシュ管理情報128の情報を更新する。

【0041】

稼働情報125は、ストレージ装置1a, 1b, 1cの稼働情報を保持する表である。稼働情報125の詳細は、具体例を示した図6を参照しながら後述する。

20

【0042】

構成情報126は、ストレージ装置1a, 1b, 1cの構成情報を保持する表である。構成情報126の詳細は、具体例を示した図7を参照しながら後述する。

【0043】

オーナー権管理情報127は、ストレージ装置1a, 1b, 1cのCPU11ごとのボリュームの制御権であるオーナー権の管理情報を保持する表である。オーナー権管理情報127の詳細は、具体例を示した図8を参照しながら後述する。

【0044】

キャッシュ管理情報128は、ストレージ装置1a, 1b, 1c内のキャッシュ16のポインタ(キャッシュ管理ポインタ)とその状態を保持する表である。キャッシュ管理情報128の詳細は、具体例を示した図9を参照しながら後述する。

30

【0045】

図5は、クラウド上の管理サーバ2に保持されるプログラム及び情報を示す図である。図5に示したように、クラウド上の管理サーバ2には、管理情報送受信・格納部211、イベント検知部212、イベント登録部213、プラン生成部214、プラン提示・修正部215、プラン実行部216、プラン実行完了確認部217、組織情報221、装置情報222、装置構成情報223、稼働履歴情報224、稼働履歴ランキング情報225、検知情報226、性能異常履歴情報227、イベント情報228、プラン情報229、及びジョブ情報230が記憶されている。

40

【0046】

管理情報送受信・格納部211は、ストレージ装置1a, 1b, 1cの管理情報送受信部121と通信し、ストレージ装置1a, 1b, 1cからの情報を受信したり、管理サーバ2からの管理操作をストレージ装置1a, 1b, 1cに送信したり、受信した情報を装置情報222、装置構成情報223、稼働履歴情報224、稼働履歴ランキング情報225、または性能異常履歴情報227に格納したりするプログラムである。

【0047】

イベント検知部212は、稼働履歴情報224、稼働履歴ランキング情報225、及び性能異常履歴情報227の情報をを用いて、所定の検知条件に合致したかをチェックし、条件

50

に合致した場合はイベントとして検知するプログラムである。

【0048】

イベント登録部213は、イベント検知部212によってイベントが検知された場合に呼び出されるプログラムであり、イベント検知部212が検知したイベントをイベント情報228に格納するプログラムである。

【0049】

プラン生成部214は、イベント登録部213によってイベント情報228に登録されたイベントを解決するプランを生成するプログラムである。また、プラン生成部214は、生成したプランをプラン情報229及びジョブ情報230に格納する。

【0050】

プラン提示・修正部215は、プラン生成部214によって生成されたプラン情報をユーザ（管理者）に提示し、承認を得るプログラムである。また、プラン提示・修正部215は、ユーザ（管理者）から、プランの実行を希望する時刻（実行希望時刻）やプランの実行を許可しない時刻（実行禁止時刻）を受け付け、プランを修正するプログラムでもある。なお、プラン提示・修正部215がユーザ（管理者）に提示するインターフェースは、GUI、CLI（Command Line Interface）、またはAPI（Application Programming Interface）等、任意のI/Fであってよい。

【0051】

プラン実行部216は、プラン生成部214によって生成されたプログラム、またはプラン提示・修正部215によって修正されたプランを実行するプログラムである。プラン実行部216は、プラン情報229及びジョブ情報230の情報を管理情報送受信・格納部211に送信することにより、プランを実行するためのストレージ操作の情報をストレージ装置1a, 1b, 1cに送信する。

【0052】

プラン実行完了確認部217は、プラン実行部216によるプラン実行後に、ストレージ装置1における稼働情報を収集して（実際には稼働履歴情報224を参照して）、イベントが解決されたかどうかを確認するプログラムである。プラン実行完了確認部217は、イベントが解決されたことを確認したか、またはプラン実行から一定時間が経過した場合に、イベント情報228を更新する。

【0053】

組織情報221は、データセンタ3及びデータセンタ3内のストレージ装置1a, 1b, 1cを利用する組織の情報を保持する表である。組織情報221の詳細は、具体例を示した図10を参照しながら後述する。

【0054】

装置情報222は、組織情報221に保持される組織ごとに、各組織の管理対象のストレージ装置1a, 1b, 1cの識別子、名前、及び状態を示す情報を保持する表である。装置情報222の詳細は、具体例を示した図11を参照しながら後述する。

【0055】

装置構成情報223は、装置情報222で管理対象として保持されたストレージ装置1a, 1b, 1cそれぞれの構成情報を保持する表である。装置構成情報223の詳細は、具体例を示した図12を参照しながら後述する。

【0056】

稼働履歴情報224は、装置情報222で管理対象として保持されたストレージ装置1a, 1b, 1cからそれぞれ収集した稼働情報の時系列の履歴情報を保持する表である。稼働履歴情報224の詳細は、具体例を示した図13を参照しながら後述する。

【0057】

稼働履歴ランキング情報225は、稼働履歴情報224に保持された、管理対象のストレージ装置1a, 1b, 1cから収集した稼働情報の時系列の履歴情報のうちから、負荷の高いリソースについてのランキング及びその負荷量を保持する表である。稼働履歴ランキング情報225の詳細は、具体例を示した図14を参照しながら後述する。

10

20

30

40

50

【 0 0 5 8 】

検知情報 2 2 6 は、上記管理対象のストレージ装置 1 a , 1 b , 1 c の稼働情報からイベントを検知するための情報を保持する表であって、イベントを検知する所定の検知条件の構成要件に関する情報を保持する。検知情報 2 2 6 の詳細は、具体例を示した図 1 5 を参照しながら後述する。

【 0 0 5 9 】

性能異常履歴情報 2 2 7 は、上記管理対象のストレージ装置 1 a , 1 b , 1 c の稼働情報が、性能異常を示す所定の基準を超えた場合に、そのリソース識別子、閾値、時刻、及び稼働率等を保持する表である。性能異常履歴情報 2 2 7 の詳細は、具体例を示した図 1 6 を参照しながら後述する。

10

【 0 0 6 0 】

イベント情報 2 2 8 は、イベントの検知条件に合致したときのイベントの情報を保持する表である。イベント情報 2 2 8 の詳細は、具体例を示した図 1 7 を参照しながら後述する。

【 0 0 6 1 】

プラン情報 2 2 9 は、イベント情報 2 2 8 に登録されたイベントを解決するためのプランに関する諸情報（例えば、操作名、登録時刻、予測処理時間、プラン実行に伴う副作用に関する情報等）、及びプラン実行に必要な情報（例えば、実行予定時刻または実行禁止時刻等）を保持する表である。プラン情報 2 2 9 の詳細は、具体例を示した図 1 8 を参照しながら後述する。

20

【 0 0 6 2 】

ジョブ情報 2 3 0 は、プラン情報 2 2 9 に登録されたプランを実行するときのストレージ装置 1 a , 1 b , 1 c への具体的な操作の内容を保持する表である。ジョブ情報 2 3 0 の詳細は、具体例を示した図 1 9 を参照しながら後述する。

【 0 0 6 3 】

(1 - 3) 各種データ

図 6 は、稼働情報 1 2 5 の一例を示す図である。図 6 に示したように、稼働情報 1 2 5 は、ストレージ装置 1 が有するリソースの種別情報を示すリソース種別 1 2 5 0 1、リソースの識別子を示すリソース ID 1 2 5 0 2、対象リソースにおけるメトリックの名称を示すメトリック 1 2 5 0 3、及びメトリックの値を示す値 1 2 5 0 4 を管理する。

30

【 0 0 6 4 】

このうち、リソース種別 1 2 5 0 1 は、オーナー権変更の計算に必要な負荷情報である、CPU 1 1、ボリューム 1 8、キャッシュ 1 6 の情報を少なくとも保持するが、これらのリソースのみに限定されるものではない。

【 0 0 6 5 】

また、メトリック 1 2 5 0 3 は、CPU 1 1 に関するメトリックとして、CPU 1 1 の稼働率を計算する元の情報となる各種の「BusyTime」（プロセッサが非アイドル状態にある時間を指す）や、CPU 1 1 の利用率が高いボリューム 1 8 のランキングを示す「UsageRanking」の情報を保持する。メトリック「UsageRanking」の値 1 2 5 0 4 には、ボリューム 1 8 の識別子だけでなく、ボリュームごとのBusyTimeも設定される。なお、CPU 1 1 に関するメトリックは、上記例に限定される必要はなく、他のメトリックを保持してもよい。また、1 の CPU 1 1 内のコアごとに稼働情報を管理してもよく、その場合は、1 の CPU 1 1 全体として各コアの稼働情報の平均値等を用いて、CPU 1 1 全体の値を保持するようにしてもよい。

40

【 0 0 6 6 】

また、メトリック 1 2 5 0 3 は、ボリューム 1 8 に関するメトリックとして、シーケンシャル I / O またはランダム I / O それぞれの Read / Write 情報である「SequentialReadIoCount」、「SequentialWriteIoCount」、「RandomReadIoCount」、「RandomWriteIoCount」、「SequentialReadHitIoCount」、「RandomReadHitIoCount」や、I / O のレスポンスタイムの計算に必要な「ReadAccumulatedR

50

responseTime」、「WriteAccumulatedResponseTime」や、データ転送量の計算に必要な「SequentialReadBlockCount」、「SequentialWriteBlockCount」、「RandomReadBlockCount」、「RandomWriteBlockCount」といった値を保持してよい。なお、ボリューム18に関するメトリックは、上記例に限定される必要はない。

【0067】

また、メトリック12503は、キャッシュ16に関するメトリックとして、キャッシュメモリ16へのWriteが滞留する割合を示す「WritePendingRate」を保持してもよい。なお、キャッシュ16に関するメトリックは、上記例に限定される必要はない。

【0068】

なお、本実施形態では、ストレージ装置1側で稼働情報の生情報（累積値）を保持し、累積値を平均値に変換してからクラウド上の管理サーバ2へ送信する例を示すが、それに限定される必要はなく、ストレージ装置1で平均値を計算しておいて、その情報をそのままクラウド上の管理サーバ2へ送信する構成としてもよい。

【0069】

図7は、構成情報126の一例を示す図である。図7に示したように、構成情報126は、ストレージ装置1を識別するための製品番号を示す装置製番12601、ストレージ装置1のモデル（型式）を示す装置モデル12602、ストレージ装置1が有するリソースの種別情報を示すリソース種別12603、リソースの識別子を示すリソースID12604、各リソース間の関連を示す関連リソース12605、各リソースのスペック（サイズ、速度等）を示すスペック12606、及び各ストレージ装置1におけるリソースの属性情報を示す属性12607を管理する。

【0070】

このうち、リソース種別12603は、それぞれのストレージ装置1a, 1b, 1cが有する、CPU11、キャッシュ16、ポート（ネットワークI/F14）、ドライブ13、プール17、及びボリューム18の情報を保持する。

【0071】

また、関連リソース12605は、ドライブ13とプール17との間の物理/論理の関係や、プール17とボリューム18との間の論理リソース間の関係等を保持する。

【0072】

また、スペック12606は、物理的または論理的な構成の数字上の性能値（スペック）を保持する。

【0073】

また、属性12607は、例えば対象のリソースがボリューム18の場合、当該ボリュームの圧縮機能が有効であることを「圧縮」という属性情報で表したり、当該ボリュームにコピー設定がされているときには「コピー」という属性情報で表したりすることができる。なお、上記「コピー」の属性情報には、当該ボリュームとのペア状態が設定されている別のボリュームの識別子も記載されるとしてもよい。また、属性12607には、1つのリソースに複数の属性情報が設定されてもよい。

【0074】

図8は、オーナー権管理情報127の一例を示す図である。オーナー権管理情報127は、ストレージ装置1においてどのボリューム18がどのCPU11に制御されるか（どのCPU11がどのボリューム18を制御するか）、というストレージ装置1が持つオーナー権（制御権）に関する情報を保持する。具体的には図8に示したように、オーナー権管理情報127は、上記オーナー権による制御関係を有するボリューム18及びCPU11の組み合わせについて、ボリューム18の識別子を示すVolumeID12701と、CPU11の識別子を示すCPU_ID12702とを保持する。

【0075】

図9は、キャッシュ管理情報128の一例を示す図である。図9に示したように、キャッシュ管理情報128は、CPU11の識別子を示すCPU_ID12801、ボリューム

10

20

30

40

50

18の識別子を示すVolume ID 12802、ボリューム18のどの領域がキャッシュ16に保持されているかを示すキャッシュ上のポインタであるキャッシュ管理ポインタ12803、及びキャッシュ管理ポインタの状態を示す状態12804の情報を保持する。

【0076】

ここで、特許文献2にも開示されるように、オーナー権の管理においては、ボリュームのどの領域がキャッシュのどの領域に保存されるかを管理する必要があり、ボリュームを管理するCPUのオーナー権があるCPUから別のCPUに移動すると、移動先のCPUでの管理情報（キャッシュ管理ポインタ）の情報もコピーする必要がある。図9に示したキャッシュ管理情報128の状態12804では、このコピー中の状態を「移行中」とし、移行が完了している通常状態を「通常」としている。なお、上記の「移行中」の状態でI/Oが発生すると、キャッシュが有効でないため、I/O性能が劣化する。

10

【0077】

図10は、組織情報221の一例を示す図である。複数のストレージを管理する場合、複数台のストレージを1つの組織として管理することがある。本実施形態ではそれに従い、複数のストレージを組織という単位で管理する。他にはテナントや部署単位、もしくは単に集合体に意味を持たせずグループにIDを持たせてグループ毎に複数のストレージを管理してもよい。組織情報221は、この組織に関する情報を保持するものであり、具体的には図10に示したように、組織の識別子を示すID 22101、組織に割り当てられたコード名を示す組織コード22102、組織の種別を示すタイプ22103、及び組織の名前を示す名前22104、を保持する。このうち、タイプ22103は、例えば顧客、パートナーといった値を取り得る。また、組織情報221が保持する情報は、例えば運用を開始するときに最初に設定される。

20

【0078】

図11は、装置情報222の一例を示す図である。図11に示したように、装置情報222は、ストレージ装置1を保有または管理する組織の識別子を示す組織ID 22201、ストレージ装置1の製品番号を示す装置製番22202、ストレージ装置1のモデル（型式）を示す装置モデル22203、ストレージ装置1の名前を示す名前22204、及びストレージ装置1のクラウドとの接続状態を示す状態22205を保持する。

【0079】

装置情報222は、例えば管理情報送受信・格納部211がストレージ装置1と通信して定期的に情報を更新するか、またはイベントドリブンで通信して情報を更新する。そして、組織ID 22201は、図10に示した組織情報221のID 22101に対応する。また、装置製番22202及び装置モデル22203はそれぞれ、図7に示した構成情報126の装置製番12601及び装置モデル12602に対応する。

30

【0080】

図12は、装置構成情報223の一例を示す図である。図12に示したように、装置構成情報223は、ストレージ装置1を保有または管理する組織の識別子を示す組織ID 22301、ストレージ装置1の製品番号を示す装置製番22302、ストレージ装置1のモデル（型式）を表す装置モデル22303、ストレージ装置1が有するリソースの種別情報を示すリソース種別22304、リソースの識別子を示すリソースID 22305、各リソース間の関連を示す関連リソース22306、リソースのスペック（サイズ、速度等）を示すスペック22307、及びストレージ装置1におけるリソースの属性情報を示す属性22308を保持する。

40

【0081】

装置構成情報223は、例えば管理情報送受信・格納部211がストレージ装置1と通信して定期的に情報を更新するか、またはイベントドリブンで通信して情報を更新する。そして、組織ID 22301は、図10に示した組織情報221のID 22101に対応する。また、装置製番22302、装置モデル22303、リソース種別22304、リソースID 22305、スペック22307、及び属性22308はそれぞれ、図7に示し

50

た構成情報 1 2 6 の装置製番 1 2 6 0 1、装置モデル 1 2 6 0 2、リソース種別 1 2 6 0 3、リソース ID 1 2 6 0 4、スペック 1 2 6 0 6、及び属性 1 2 6 0 7 に対応する。なお、装置構成情報 2 2 3 の関連リソース 2 2 3 0 6 には、構成情報 1 2 6 (図 7 参照) の関連リソース 1 2 6 0 5 に記載される情報に加え、オーナー権管理情報 1 2 7 (図 8 参照) に示されるボリューム 1 8 と CPU 1 1 との関連情報も含めて保持される。例えば図 1 2 では、関連リソース 2 2 3 0 6 に、ボリューム 1 8 の関連リソースとして CPU 1 1 の ID を記載しているが、CPU 1 1 の関連リソースとしてボリューム 1 8 の ID を設定するようにしてもよい。またその場合、複数のボリューム 1 8 の ID を設定してもよい。

【0082】

図 1 3 は、稼働履歴情報 2 2 4 の一例を示す図である。稼働履歴情報 2 2 4 は、ストレージ装置 1 に保持される稼働情報 1 2 5 を基に算出される情報であって、ストレージ装置 1 の管理情報送受信部 1 2 1 とクラウド上の管理サーバ 2 の管理情報送受信・格納部 2 1 1 との間で行われる定期的に通信を経て更新される。詳細には、稼働履歴情報 2 2 4 に記録される情報は、管理情報送受信部 1 2 1 が稼働情報 1 2 5 の情報をもとに稼働履歴情報 2 2 4 に登録する値を算出した上で、管理情報送受信・格納部 2 1 1 がこれらを受信して更新する。なお、別例としては、管理情報送受信・格納部 2 1 1 が管理情報送受信部 1 2 1 から稼働情報 1 2 5 の情報を受信し、受信後に稼働情報 1 2 5 の情報をもとに値を算出して稼働履歴情報 2 2 4 を更新するとしてもよい。

10

【0083】

図 1 3 に示したように、稼働履歴情報 2 2 4 は、ストレージ装置 1 を保有または管理する組織の識別子を示す組織 ID 2 2 4 0 1、ストレージ装置 1 の製品番号を示す装置製番 2 2 4 0 2、ストレージ装置 1 のモデル(型式)を示す装置モデル 2 2 4 0 3、ストレージ装置 1 が有するリソースの種別情報を示すリソース種別 2 2 4 0 4、リソースの識別子を示すリソース ID 2 2 4 0 5、当該レコードの情報を算出(または受信)した時刻を示す時刻 2 2 4 0 6、対象リソースにおけるメトリックの名称を示すメトリック 2 2 4 0 7、及びメトリックの値を示す値 2 2 4 0 8 を管理する。

20

【0084】

このうち、組織 ID 2 2 4 0 1 は、図 1 0 に示した組織情報 2 2 1 の ID 2 2 1 0 1 に対応する。また、装置製番 2 2 2 0 2 及び装置モデル 2 2 2 0 3 はそれぞれ、図 7 に示した構成情報 1 2 6 の装置製番 1 2 6 0 1 及び装置モデル 1 2 6 0 2 に対応する。また、リソース種別 2 2 4 0 4 及びリソース ID 2 2 4 0 5 はそれぞれ、図 6 に示した稼働情報 1 2 5 のリソース種別 1 2 5 0 1 及びリソース ID 1 2 5 0 2 に対応する。

30

【0085】

なお、本実施形態では、図 6 の稼働情報 1 2 5 で説明したようにストレージ装置 1 側で稼働情報の生情報(累積値)を保持しており、クラウド上の管理サーバ 2 側では、生の稼働情報(累積値)を平均値とした値を稼働履歴情報 2 2 4 に保持する例を示す。このとき、時刻 2 2 4 0 6 には平均を計算した時刻が示され、値 2 2 4 0 8 にはメトリックの平均値が示される。時刻の間隔はどのような値を取ってもよい。

【0086】

また、メトリック 2 2 4 0 7 は、CPU 1 1 に関するメトリックとして、CPU 1 1 の稼働率を示す各種の「BusyTime」や、CPU 1 1 の利用率が高いボリューム 1 8 のランキングを示す「UsageRanking」の情報を保持する。稼働率は例えば、「BusyTime」の 2 点間の差を、取得時刻の 2 点間の差の時刻で割ることによって算出できる。なお、メトリック 2 2 4 0 7 に保持される CPU 1 1 に関するメトリックは上記例に限定される必要はなく、他のメトリックを保持してもよい。

40

【0087】

また、メトリック 2 2 4 0 7 は、ボリューム 1 8 に関するメトリックとして、シーケンシャル I/O またはランダム I/O それぞれの Read/Write 情報である「SequentialReadIOPS」、「SequentialWriteIOPS」、「RandomReadIOPS」、「RandomWriteIOPS」、「SequentialReadHitIOPS」、「RandomReadHitIOPS」や、I

50

ノのレスポンスタイムである「ReadResponseTime」、「WriteResponseTime」や、データ転送量である「SequentialReadTransferRate」、「SequentialWriteTransferRate」、「RandomReadTransferRate」、「RandomWriteTransferRate」といった値を保持してよい。例えば、各種の「IOPS」は、各種「IoCount」（図6参照）の2点間の差を、取得時刻の2点間の差の時刻で割ることによって算出できる。各種「ResponseTime」は、各「AccumulatedResponseTime」（図6参照）の2点間の差を、取得時刻の2点間の差の時刻で割ることによって算出できる。各「TransferRate」は、各種「BlockCount」（図6参照）の2点間の差を、取得時刻の2点間の差の時刻で割ることによって算出できる。なお、メトリック22407に保持されるボリューム18に関するメトリックは上記例に限定される必要はない。

10

【0088】

また、メトリック22407は、キャッシュ16に関するメトリックとして、キャッシュメモリ16へのWriteが滞留する割合を示す「WritePendingRate」を保持してもよいが、キャッシュ16に関するメトリックは、上記例に限定される必要はない。

【0089】

図14は、稼働履歴ランキング情報225の一例を示す図である。稼働履歴ランキング情報225は、図13に示した稼働履歴情報224におけるメトリック「UsageRanking」の値をもとに作成される。図14に示したように、稼働履歴ランキング情報225は、ストレージ装置1を保有または管理する組織の識別子を示す組織ID22501、ストレージ装置1の製品番号を示す装置製番22502、ストレージ装置1のモデル（型式）を示す装置モデル22503、CPU11の識別子を示すCPU_ID22504、時刻22505、ランキングを示すRank22506、ボリューム18の識別子を示すVolumeID22507、及びボリューム18ごとのCPU稼働率を示すBusyRateInPercent22508を保持する。

20

【0090】

図15は、検知情報226の一例を示す図である。検知情報226は、イベントを検知する所定の検知条件の構成要件に関する情報を保持するものであって、図15に示したように、検知情報の識別子を示すID22601、ストレージ装置1を保有または管理する組織の識別子を示す組織ID22602、ストレージ装置1の製品番号を示す装置製番22603、ストレージ装置1のモデル（型式）を示す装置モデル22604、性能異常の判定に用いられる閾値を示す閾値22605、閾値超過を検知する時間間隔として設定される最小検知間隔（MinimumDetectionInterval）を示す最小検知間隔22606、閾値超過の検知を行う検知期間（DetectionPeriod）を示す検知期間22607、及び検知期間の間に閾値超過を検知した回数である検知数（DetectionCount）を示す検知数22608を保持する。

30

【0091】

このうち、閾値22605は、固定値でも非固定値であってもよい。固定値とする場合は、システムが事前に定義した値を用いてもよい。非固定値とする場合は、ユーザの運用に合わせて適当な値を設定してもよく、運用前に値を設定して運用中に変更するようにしてもよい。さらに、ユーザ運用だけでなく、学習等によって適当な値を判断するよ

40

【0092】

次に、最小検知間隔22606、検知期間22607、及び検知数22608についての例を説明する。本例では、検知期間22607及び検知数22608を異常発生条件に用いる。

【0093】

基本的な考え方として、CPU11とボリューム18の制御権における性能不均衡のような問題は、ある一瞬のずれを見るのではなく、長期的な傾向を見る必要がある。長期的な傾向を見るためには、短時間の性能変動はノイズとなる。そこでまずノイズ除去を行い、次に長期的な傾向を判断することが好ましい。ノイズの除去の一例として、最小検知間隔

50

を利用し、長期的な傾向判断の一例として、検知期間と検知数を利用する。

【0094】

最小検知間隔 22606 は、閾値超過を検知する時間間隔（解像度）である。例えばストレージ装置 1 からクラウド上の管理サーバ 2 に 5 分間隔で稼働情報を送信する場合、閾値超過を検知すると、性能異常履歴登録や他 CPU で発生した閾値超過履歴の確認を行うが、その処理を短い時間間隔で行うと、管理サーバ 2 側の処理負荷が高くなる。これを避けるために、最小検知間隔（例えば 1 時間とする）を設け、ある CPU で平均値よりもずれが大きいうという不均衡が最小検知間隔後も発生するかどうかを見ることで、管理サーバ 2 側の処理負荷を少なくすることができる。つまり、最小検知間隔は、時間方向で不均衡の検知をダウンサンプリングするために利用することができる。

10

【0095】

ここで、最小検知間隔のみを利用するとした場合には、次のような欠点が想定される。性能情報は傾向があるといっても変動の大きなメトリックであるため、ある時点（5 分間隔のある時点）でたまたま不均衡が発生していない場合があり得る。このような場合に、たまたま不均衡が発生していないタイミングの時刻（から最小検知間隔までの間）において異常ではないと判断してしまうことがある。本例では、以上のような欠点を補うために、検知期間 22607 及び検知数 22608 による 2 つの判断基準を設ける。

【0096】

検知期間 22607 は、現在時刻から過去のどこまでの期間を検知対象とするか、を示す期間である。検知数 22608 は、検知期間の間に基準の閾値を超過した（不均衡が発生した）回数を示す値である。例として、検知期間を 5 日間、検知数を 30 回とする場合を考える。検知期間（5 日間）のうちに、最小検知間隔（1 時間）おきのチェックで性能異常を検知し、その回数が検知数（30 回）を超える場合に、定常的に性能異常（不均衡）が発生すると判断する。このように、最小検知間隔を設けることで、時間方向でダウンサンプリングを実施し、管理サーバ 2 側の処理負荷を少なくすることができる。また、検知期間と検知数を設けることで、長期的な傾向として、性能不均衡度合が次第に大きくなる場合だけでなく、性能メトリック自体は増減の変動が発生するが長期の傾向として不均衡が繰り返される場合（1 時間 * 30 回 / (24 時間 * 5 日) = 25%。すなわち、検知期間のうち 25% の時間で不均衡が発生していると考えられる場合）も、性能異常（不均衡）としてイベントを検知することができるようになる。

20

30

【0097】

なお、図 15 の例ではストレージ装置 1 ごとに閾値、最小検知間隔、検知期間、検知数が設定されているが、システムで共通の値を持ってよい。また、これらの値は固定値でも非固定値でもよい。固定値の場合はシステムが事前に定義した値を用いてもよく、非固定値の場合は、ユーザの運用に合わせて適当な値を設定してもよく、運用前に設定して運用中に値を変更してもよい。さらに、ユーザ運用だけでなく、学習等によって適当な値を判断して設定するようにしてもよい。

【0098】

また、ノイズ除去方法は、処理の軽い最小検知間隔を利用した上記のダウンサンプリング方法だけでなく、例えば、ローパスフィルタやウィンドウ集計（タンプリングウィンドウや、スライディングウィンドウ）等の方法を利用してもよい。また、長期的な傾向の判断にも、例えば高速フーリエ変換を使って周期性を確認したり、回帰計算をすることで予測を実施してもよい。

40

【0099】

図 16 は、性能異常履歴情報 227 の一例を示す図である。性能異常履歴情報 227 はイベント登録部 213 によって更新される。図 16 に示したように、性能異常履歴情報 227 は、性能異常の識別子を示す ID 22701、ストレージ装置 1 を保有または管理する組織の識別子を示す組織 ID 22702、ストレージ装置 1 の製品番号を示す装置製番 22703、ストレージ装置 1 のモデル（型式）を示す装置モデル 22704、CPU 11 の識別子を示す CPU_ID 22705、性能異常の判定に用いられた閾値を示す閾値 2

50

2706、性能異常が発生した時刻を示す異常発生時刻22707、及び異常発生時刻のCPU11の稼働率を示す稼働率22708を保持する。なお、閾値22705は、固定値でも非固定値でもよい。固定値の場合はシステムが事前に定義した値を用いてもよく、非固定値の場合は、ユーザの運用に合わせて適当な値を設定してもよく、運用前に設定して運用中に値を変更してもよい。さらに、ユーザ運用だけでなく、学習等によって適当な値を判断し設定してもよい。

【0100】

図17は、イベント情報228の一例を示す図である。イベント情報228はイベント登録部213によって更新される。図17に示したように、イベント情報228は、イベントの識別子を示すID22801、ストレージ装置1を保有または管理する組織の識別子を示す組織ID22802、ストレージ装置1の製品番号を示す装置製番22803、ストレージ装置1のモデル(型式)を示す装置モデル22804、イベントの発生時刻を示す発生時刻22805、イベントへの対処が実行されて解決した時刻を示す更新時刻22806、及びイベントが解決したか否かのステータスを示す解決済22807を保持する。イベント情報228では、イベントが発生し、それに対する対処が実行されて解決すると、その解決時刻で更新時刻22806が更新され、解決済22807が「Yes」とされる。

10

【0101】

図18は、プラン情報229の一例を示す図である。プラン情報229は、イベント情報228に記載されたイベントを解決するためのプランに関する情報を保持し、プラン生成部214やプラン提示・修正部215によって更新される。具体的には図18に示したように、プラン情報229は、プランの識別子を示すID22901、イベントの識別子(イベントID)を示すイベントID22902、ストレージ装置1の製品番号を示す装置製番22903、ストレージ装置1のモデル(型式)を示す装置モデル22904、プランの状態を示す状態22905、プラン情報229にプランが登録された時刻を示すプラン登録時刻22906、プランの変更があった場合にその時刻を示すプラン変更時刻22907、プランの実行時に予測される処理時間を示す予測処理時間22908、プランを実行した場合の(実行中の)I/Oへの影響の予測値(I/O影響負荷)を示すI/O影響予測22909、プランを実行した場合の(実行中の)CPU11への影響の予測値(CPU負荷上昇量)を示すCPU影響予測22910、プランの実行時刻に対するユーザからの指定条件を示すユーザ指定条件22911、プランの実行予定時刻を示す実行予定時刻22912、及びユーザから指定されたプランの実行を禁止する時刻を示すプランの実行禁止時刻22913を保持する。

20

30

【0102】

このうち、イベントID22902は、図17に示したイベント情報228のID22801に対応する。また、状態22905が取り得る値は、プランが登録され未実行である「Scheduled」やプランが実行中である「Executing」等がある。なお、これらの値に限定されず他の値を設定してもよい。

【0103】

また、予測処理時間22908は、後述するジョブ情報230に記載されるジョブ処理の実行に要する時間の見積値が設定される。

40

【0104】

また、I/O影響予測22909は、プランを実行したときにI/O性能が劣化する場合に、その劣化度合いを示す。図18のI/O影響予測22909では割合を記載しているが、これに限らず、応答時間(ResponseTime)の絶対値や、他のメトリックの割合や絶対値でもよい。

【0105】

また、CPU影響予測22910は、プランを実行するプラン実行処理によりCPU11に負荷が掛かるため、その影響予測情報を設定する。図18のCPU影響予測22910では、現在の状態から変化する割合を記載しているが、他にも例えばCPU稼働率(Bus

50

yRate) の劣化の値を設定する等してもよい。

【0106】

また、ユーザ指定条件 22911 は、プランをユーザに提示したときにユーザから指定条件（実行希望時刻や実行禁止時刻）が設定された場合に、指定条件の設定があった旨を示すものであり、例えば実行希望時刻が設定された場合には「希望時刻」と記載され、実行禁止時刻が設定された場合には「禁止時刻」と記載される。また、実行予定時刻 22912 は、プランの実行予定時刻が記載されるものであるが、ユーザによって実行希望時刻が設定された場合には、その時刻が設定される。一方、ユーザによる実行希望時刻の設定がない場合には、システムによって決定されるプランの実行推奨時刻が設定される。また、実行禁止時刻 22913 には、ユーザによって実行禁止時刻が設定された場合に、その値が設定される。

10

【0107】

図 19 は、ジョブ情報 230 の一例を示す図である。ジョブ情報 230 は、プランを実行するための具体的なストレージ装置 1 に対する具体的な処理の内容を示す情報を保持するものであって、プラン生成部 214 によって更新される。本実施形態では、1つのプランに複数のジョブが対応する場合がある。図 19 に示したように、ジョブ情報 230 は、ジョブの識別子を示す ID 23001、プランの識別子を示すプラン ID 23002、ストレージ装置 1 の製品番号を示す装置製番 23003、ストレージ装置 1 のモデル（型式）を示す装置モデル 23004、ジョブで実行される操作を示す操作名 23005、オーナー権の移動元 CPU を示す移動元 CPU 23006、オーナー権が変更されるボリューム 18 である移動対象のボリュームを示す移動対象 Volume 23007、及びオーナー権の移動先 CPU を示す移動先 CPU 23008 を保持する。このうち、プラン ID 23002 は、図 18 に示したプラン情報 229 の ID 22901 に対応する。

20

【0108】

(1-4) 処理

本実施形態に係る計算機システム 100 において実行される処理について説明する。

【0109】

本実施形態における処理フローは、主にストレージ装置 1 から構成情報や稼働情報等の運用情報をクラウド上の管理サーバ 2 へ収集する処理（情報収集処理）、収集した情報からイベントを検知する処理（検知処理）、検知したイベントを対処するプランを生成する処理（プラン生成処理）、及びプランを実行する処理（プラン実行処理）の 4 つのフローに分かれる。また、本実施形態に係る計算機システム 100 では、上記 4 つの処理以外に 2 つの定期的な処理が実行される。1 つは、ストレージ装置 1 の稼働情報収集・格納部 122 がストレージ装置 1 の稼働情報（I/O や CPU の負荷等）を定期的に収集し、稼働情報 125 に保存する処理である。もう 1 つは、特許文献 2 にも示されるように、ストレージ装置 1 のキャッシュ管理部 124 がキャッシュ管理情報 128 を更新する処理である。この処理は、キャッシュ管理情報 128 における CPU_ID 12801 と Volume ID 12802 の関係を定期的にバックグラウンドでチェックしキャッシュ管理ポイント 12803 を定期的に更新する（キャッシュのダーティ領域の定期スキャンを実施する）ものである。これらの処理は本特許の特徴でなく公知であるため、図示を省略し、以下では最初に列挙した 4 つの処理フローについて、図示とともにその詳細を説明する。

30

40

【0110】

図 20 は、第 1 の実施形態における全体的な処理の処理手順例を示すフローチャートである。

【0111】

図 20 に示した全体的な処理 S1000 によれば、まず、ストレージ装置 1 からクラウド上の管理サーバ 2 に情報を送信する情報収集処理が実行される（ステップ S1100）。本例では、ステップ S1100 の処理は定期的に行われるとするが、本実施形態においてステップ S1100 の処理の実行タイミングはこれに限定されるものではなく、例えばストレージ装置 1 からの Push 通知が上がったときに実行する等としてもよい。なお、

50

ステップ S 1 1 0 0 は、図 1 に示したステップ S 1 に相当する。

【 0 1 1 2 】

次に、ステップ S 1 1 0 0 が実行されて各種情報がクラウド上の管理サーバ 2 に送信されて保存された後は、イベントを検知する検知処理が実行される（ステップ S 1 2 0 0）。ステップ S 1 2 0 0 は、図 1 に示したステップ S 2 に相当する。

【 0 1 1 3 】

次に、ステップ S 1 2 0 0 が実行されてイベントがイベント情報 2 2 8 に登録されると、イベントを解決するためのプランを生成するプラン生成処理が実行される（ステップ S 1 3 0 0）。ステップ S 1 3 0 0 は、図 1 に示したステップ S 3 に相当する。なお、ステップ S 1 2 0 0 からステップ S 1 3 0 0 への遷移は、イベントドリブンで実行してもよいし、一度 GUI 等でユーザ（管理者）にイベントの発生を通知してユーザ（管理者）から当該イベントに対応する旨の指示があった場合に実行するとしてもよい。

10

【 0 1 1 4 】

そして、ステップ S 1 3 0 0 でプランが生成された後、プランを実行するプラン実行処理が実行される（ステップ S 1 4 0 0）。ステップ S 1 4 0 0 は、図 1 に示したステップ S 4 に相当する。なお、ステップ S 1 3 0 0 からステップ S 1 4 0 0 への遷移は、イベントドリブンで実行してもよいし、ユーザ（管理者）からプランの実行指示があった場合に実行するとしてもよいし、プランの有無をシステム内で定期的にチェックしてプランが存在することが確認された場合に実行する等としてもよい。

【 0 1 1 5 】

以下では、上述したステップ S 1 1 0 0 ~ S 1 4 0 0 の各処理を詳しく説明する。

20

【 0 1 1 6 】

（ 1 - 4 - 1 ） 情報収集処理

図 2 1 は、情報収集処理の処理手順例を示すフローチャートである。図 2 1 に示すフローチャートは、図 2 0 に示したステップ S 1 1 0 0 の情報収集処理の詳細な処理手順例を示すものである。

【 0 1 1 7 】

図 2 1 によればまず、ストレージ装置 1 の管理情報送受信部 1 2 1 が、稼働情報 1 2 5 に保持されている稼働情報の値を平均値に変換する（ステップ S 1 1 0 1）。平均値の計算方法は、例えば図 1 3 の稼働履歴情報 2 2 4 で説明した方法を利用する。

30

【 0 1 1 8 】

次に、管理情報送受信部 1 2 1 は、クラウド上の管理サーバ 2 の管理情報送受信・格納部 2 1 1 に、ステップ S 1 1 0 1 で平均値に変換した稼働情報と、構成情報 1 2 6 及びオーナー権管理情報 1 2 7 を送付する（ステップ S 1 1 0 2）。

【 0 1 1 9 】

次に、管理サーバ 2 の管理情報送受信・格納部 2 1 1 が、ステップ S 1 1 0 2 で受信した情報を、装置構成情報 2 2 3、稼働履歴情報 2 2 4、及び稼働履歴ランキング情報 2 2 5 に適宜格納する（ステップ S 1 1 0 3）。なお、図 1 4 の稼働履歴ランキング情報 2 2 5 で説明したように、稼働履歴ランキング情報 2 2 5 は、稼働履歴情報 2 2 4 におけるメトリック「UsageRanking」の値から作成される。

40

【 0 1 2 0 】

そしてステップ S 1 1 0 3 の処理が完了すると、情報収集処理を終了し、ステップ S 1 2 0 0 の検知処理に遷移する。

【 0 1 2 1 】

（ 1 - 4 - 2 ） 検知処理

図 2 2 は、検知処理の処理手順例を示すフローチャートである。図 2 2 に示すフローチャートは、図 2 0 に示したステップ S 1 2 0 0 の検知処理の詳細な処理手順例を示すものである。

【 0 1 2 2 】

図 2 2 によればまず、イベント検知部 2 1 2 が、稼働履歴情報 2 2 4 を参照し、現在時刻

50

におけるシステム全体（1台のストレージ装置1が保持する全CPU11）の平均CPU稼働率を計算する（ステップS1201）。

【0123】

次に、イベント検知部212は、システム全体の平均CPU稼働率が基準より低く、かつ、1つのCPU11の稼働率が基準を超えているものがあるか否かをチェックする（ステップS1202）。ここで基準とする値は、検知情報226に設定されている閾値22605を利用する。

【0124】

次に、イベント検知部212は、ステップS1202のチェック結果を確認する（ステップS1203）。ステップS1203で肯定結果が得られた場合、すなわち、システム全体の平均CPU稼働率が基準よりも低く、かつ、1つのCPU11の稼働率が基準を超えているものがある場合には（ステップS1203のYES）、ステップS1204に進む。一方、ステップS1203で否定結果が得られた場合、すなわち、システム全体の平均CPU稼働率が基準以上であるか、または、システム全体の平均CPU稼働率が基準よりも低いものの、稼働率が基準を超えているCPU11が1つも存在しない場合には（ステップS1203のNO）、ステップS1211に進む。

10

【0125】

ステップS1204では、イベント検知部212が、ステップS1203で基準を超えている1つのCPU11について、性能異常履歴情報227から同一CPUの異常発生時刻22707の値を取得する。

20

【0126】

次いで、イベント検知部212は、ステップS1204で扱った同一のCPU11について、検知情報226の最小検知間隔22606の時間が経過しているか否かを判定する（ステップS1205）。ステップS1205で肯定結果が得られた場合には（ステップS1205のYES）、ステップS1206に進み、ステップS1205で否定結果が得られた場合には（ステップS1205のNO）、ステップS1211に進む。

【0127】

ステップS1206では、イベント登録部213が、性能異常履歴情報227に新たなレコードを作成し、作成したレコードの各情報を登録する。具体的には例えば、イベント登録部213は、ステップS1204で扱った1つのCPU11の識別子をCPU_ID22705に、ステップS1203で基準を超えたことが確認された閾値を閾値22706に、現在時刻を異常発生時刻22707に、現時点での当該CPU11の稼働率を稼働率22708に登録する。

30

【0128】

次いで、イベント登録部213は、検知情報226を参照し、一定期間における異常発生条件（本例では検知期間22607及び検知数22608で構成されるとしている）に合致する状況であるか否かをチェックする（ステップS1207）。

【0129】

そしてイベント登録部213は、ステップS1207のチェック結果を確認する（ステップS1208）。ステップS1208で肯定結果が得られた場合、すなわち異常発生条件に合致する状況である場合には（ステップS1208のYES）、ステップS1209に進み、ステップS1208で否定結果が得られた場合、すなわち異常発生条件に合致しない状況である場合には（ステップS1208のNO）、ステップS1211に進む。

40

【0130】

上記のステップS1205～S1208の処理が行われることにより、最小検知間隔を利用して、時間方向で不均衡の検知をダウンサンプリングすることができる。

【0131】

ステップS1209では、イベント登録部213がイベント情報228を確認し、ステップS1206で性能異常履歴情報227に新たに登録した性能異常と同一のイベント（同伴イベント）が未解決の状態で登録済みであるか否かを判定する。具体的には、イベント

50

登録部 2 1 3 は、イベント情報 2 2 8 に、解決済 2 2 8 0 7 が「No」となっている同件イベントが存在しているか否かを確認する。ステップ S 1 2 0 9 において肯定結果が得られた場合、すなわち未解決の同件イベントが登録済みである場合は（ステップ S 1 2 0 9 の YES）、ステップ S 1 2 1 1 に進み、ステップ S 1 2 0 9 において否定結果が得られた場合、すなわち未解決の同件イベントが未登録である場合は（ステップ S 1 2 0 9 の NO）、ステップ S 1 2 1 0 に進む。

【 0 1 3 2 】

ステップ S 1 2 1 0 では、イベント登録部 2 1 3 が、イベント情報 2 2 8 に新たなレコードを作成し、当該レコードに、ステップ S 1 2 0 8 で異常発生条件に合致していると判定したイベントに関する情報（装置情報や発生時刻）を登録する。その後ステップ S 1 2 1 1 に進む。

10

【 0 1 3 3 】

ステップ S 1 2 1 1 では、全てのストレージ装置 1 の全ての CPU 1 1 について、ステップ S 1 2 0 2 以降の処理を行って異常発生の判定等の確認を行ったかをチェックする。ステップ S 1 2 1 1 で未確認の CPU 1 1 が存在した場合には（ステップ S 1 2 1 1 の NO）、ステップ S 1 2 0 2（ステップ S 1 2 0 3 でもよい）に戻り、未確認の CPU 1 1 に対してステップ S 1 2 0 2（またはステップ S 1 2 0 3）以降の処理を行う。一方、ステップ S 1 2 1 1 で全ての CPU 1 1 について確認済みであった場合には（ステップ S 1 2 1 1 の YES）、ステップ S 1 2 1 2 に進む。

【 0 1 3 4 】

そしてステップ S 1 2 1 2 では、例えばイベント登録部 2 1 3 が、今回の検知処理におけるここまでの処理でイベント情報 2 2 8 に新たに登録したイベントの有無を確認する。ステップ S 1 2 1 2 において新たな登録イベントが存在した場合には（ステップ S 1 2 1 2 の YES）、検知処理を終了し、ステップ S 1 3 0 0 のプラン生成処理に遷移する。一方、ステップ S 1 2 1 2 において新たな登録イベントが存在しなかった場合には（ステップ S 1 2 1 2 の NO）、検知処理を終了し、ステップ S 1 1 0 0 の情報収集処理に遷移する。

20

【 0 1 3 5 】

（ 1 - 4 - 3 ）プラン生成処理

図 2 3 及び図 2 4 は、プラン生成処理の処理手順例を示すフローチャート（その 1，その 2）である。図 2 3，図 2 4 に示すフローチャートは、図 2 0 に示したステップ S 1 3 0 0 のプラン生成処理の詳細な処理手順例を示すものである。

30

【 0 1 3 6 】

図 2 3，図 2 4 によればまず、プラン生成部 2 1 4 が、性能異常履歴情報 2 2 7 から異常発生時刻 2 2 7 0 7 及び稼働率 2 2 7 0 8 の情報を取得する（ステップ S 1 3 0 1）。ステップ S 1 3 0 1 の処理は、ステップ S 1 2 0 0 の検知処理で検知されたイベントの発生時刻、及びイベント発生時の CPU 稼働率を取得するものである。したがって、ステップ S 1 3 0 1 で情報の取得対象とする期間は、イベント検知後にイベントドリブンでプラン生成処理を開始する場合は、現在の時刻から検知情報 2 2 6 の検知期間 2 2 6 0 7 に設定された検知期間だけ遡った時刻までを対象期間とすればよい。また、生成したプランをユーザに提示した後にユーザの承認後にプランを実行する場合は、前回のユーザの承認時刻から現在時刻までを対象期間とするか、または、ユーザの承認時刻から、イベント検知時刻から検知期間を遡った時刻までを対象期間としてもよい。例えば、イベント検知後にイベントドリブンでプラン生成処理を開始する場合は、現在の時刻から検知情報 2 2 6 の検知期間に遡った時刻までの異常履歴情報を利用してよい1月10日8時から5日前(1月5日8時)の異常履歴情報を元にプラン生成する。この期間には必ず30回異常情報が含まれる。プラン生成をユーザに提示した後に、ユーザの承認後に実行する場合は、ユーザの承認時刻から検知期間を遡った時刻までの異常履歴情報を利用してよい。1月11日8時から5日前(1月6日8時)の異常履歴情報を元にプラン生成する。この期間には異常情報が30回でない可能性がある。ユーザの承認時刻がプラン生成時間だと考え、最新の稼働

40

50

情報を見るべきだという考えの元に、ユーザ承認時刻から5日間前の稼働情報を元にプラン生成します。ユーザの承認時刻から、イベント検知時刻から検知期間を遡った時刻、までの異常履歴情報を利用してよい。1月11日8時から1月5日8時までの異常履歴情報を元にプラン生成する。両者のORの時間となる(ユーザは通知が出てから5日以内には普通は判断・承認すると仮定)。これは異常検知した期間の稼働情報+ユーザが承認した時刻までの最新の稼働情報も加味してプラン生成する。

【0137】

次に、プラン生成部214は、ステップS1301で取得した1以上の異常発生時刻におけるCPU稼働率の時間平均を計算する(ステップS1302)。

【0138】

次に、プラン生成部214は、性能異常履歴情報227から時刻情報(異常発生時刻)を、稼働履歴ランキング情報225から異常発生時刻における上位ランクのボリュームごとの稼働情報を取得する(ステップS1303)。

【0139】

次に、プラン生成部214は、装置構成情報223の属性22308を参照し、オーナー権変更不可のボリュームの有無をチェックし、オーナー権変更不可のボリュームがあれば、プランの候補から除外する(ステップS1304)。具体的には例えば、ストレージ装置1内の特別な用途に用いられ、ユーザデータの保管用には利用されないボリューム等については、オーナー権変更不可の属性が付されており、ステップS1304ではこのようなオーナー権変更不可な属性等をチェックすることにより、オーナー権の変更処理を行うことができないボリューム18を、プランの候補から除外する。

【0140】

次に、プラン生成部214は、各ボリュームの上位ランクに入っている時刻における負荷の値を合計してスコアリングする(ステップS1305)。コピー構成の場合は、同じコピーペアボリュームは同じオーナー権で扱うと性能が良い場合があるため、コピーペアのボリュームはグループで1つのボリュームとみなしてよい。なお、時刻によって一部ボリュームがランク外(欠損)になっている場合は、その時刻における閾値超過への寄与が小さいと考えられるため、負荷の値をゼロとみなしてよい。

【0141】

次に、プラン生成部214は、性能異常履歴情報227を参照し、プラン対象のイベントが検知された検知期間(例えば5日間)の間に異常が発生したCPUを、オーナー権の移動対象候補から除外する(ステップS1306)。

【0142】

次に、プラン生成部214は、上位スコアのボリュームについて、上位のボリュームから順に、稼働率の低いCPUをオーナー権の移動先と仮定し、当該ボリュームのオーナー権を移動した際に、移動元CPU及び移動先CPUにおける稼働率がどのように加減するかを予測する処理を、予測結果のCPU稼働率が以下になるまで、繰り返し実行する(ステップS1307)。すなわち、ステップS1307では、上位スコアのボリュームから順に、オーナー権を移動した際のCPU稼働率の変化予測を繰り返し、CPU稼働率が基準値以下を達成することが見込まれた時点で、この繰り返しを終了する。また、全てのボリュームについて繰り返しを終了しても、CPU稼働率が基準値以下を達成できないこともあり得る。

【0143】

なお、ステップS1307の処理において、オーナー権の移動時に加減するボリュームの負荷は、過去の検知点における最大値を採用する。最大値を採用する理由は、性能値は時刻とともに変動する可能性があり、検知した最大の値が別の時刻でも発生している可能性があるためである。なお、CPUごとの利用率ランキング(UsageRanking)において記録されるボリュームごとの稼働率は、コピー系のバックグラウンド処理が紐付かない場合がある。このように、直接ボリュームに紐付かないがオーナー権変更によって影響がある負荷は、稼働履歴情報224に記載されたI/O量(I/Oのメトリックの値)やボリューム

10

20

30

40

50

ムごとの稼働率から見積もるようにしてもよい。

【0144】

また、移動先CPUと移動対象のボリュームの考え方は、上記した方法に限定されるものではない。例えば、性能予測が外れるリスクが高いコピー構成のボリュームは、移動させる優先度を下げて、上位スコアの通常ボリュームから移動プランを生成し、通常ボリュームの移動だけでは基準値以下にならない場合だけ、コピー構成のボリュームも移動させるとする方法でもよい。また、単なる稼働率の加減に基づいて稼働率を計算するのではなく、I/O量と稼働率の関係のモデルを内包し、オーナー権の変更による移動後のボリュームのI/O性能を予測し、その予測結果と前述のI/O量と稼働率の関係のモデルに基づいて、移動後の稼働率を算出する方法でもよい。

10

【0145】

ステップS1307の処理が完了した後、プラン生成部214は、ステップS1307においてCPU稼働率を基準値以下にする移動先CPU及びボリュームを決定できたか否かを判定する(ステップS1308)。言い換えると、ステップS1307における繰り返し処理によって、CPU稼働率の予測結果を基準値以下とすることができた場合にはイベントに対処する適切なプランが生成可能であることを意味することから、ステップS1308ではプランの生成可否を判定する。ステップS1308において基準値以下を達成するCPU及びボリュームを決定できなかった場合、すなわちプランが生成不可である場合には(ステップS1308のNO)、ステップS1309に進む。

【0146】

ステップS1309では、ステップS1301～ステップS1308の処理を経てイベントに対処するプランが生成不可であるという結果が得られたため、プラン生成部214(またはプラン提示・修正部215等でもよい)が、今回発生した性能異常(イベント)はシステム内のオーナー権変更では解決できないことをGUI等でユーザ(管理者)に通知し、プラン生成処理を終了する。

20

【0147】

一方、ステップS1308において基準値以下を達成するCPU及びボリュームを決定できた場合、すなわちプランが生成可能である場合には(ステップS1308のYES)、図24のステップS1310に進む。

【0148】

ステップS1310では、プラン生成部214が、キャッシュ容量、CPU稼働率、キャッシュのWritePendingRate、及び装置種別(装置種別から紐づく各装置のスペック情報)に基づいて、オーナー権変更実行時間を予測する。以下に詳しく説明する。

30

【0149】

オーナー権変更実行時間は、構成変更リクエスト受付時間と、キャッシュ管理情報128を定期的に更新する処理時間(キャッシュメモリ16の管理ポインタ情報の移行時間)によって決まる。

【0150】

上記のキャッシュ管理情報128を定期的に更新する処理時間は、キャッシュ容量とWritePendingRateに比例し、CPU稼働率の空き(動作時間割合)、装置のハードウェアスペックに反比例する。また、常にバックグラウンドでキャッシュ管理(移行処理含む)のためにダーティ領域のスキャンを実施している。スキャン処理とオーナー権移動処理は非同期で実行されるため、現在進行中のスキャン処理において、移行対象のキャッシュ領域が既に通り返している場合がある。このため、確実にスキャンされるように、スキャン2週分の時間で見積もる。構成変更リクエストの時間は移動するボリューム数によって決まる。以上のことを踏まえると、オーナー権変更実行時間は、以下の式1に従って予測することができる。

40

【0151】

【数1】

50

実行時間

$$= X * \text{ボリューム数} + \text{キャッシュ容量} * \text{WritePendingRate} / (Y * \text{動作時間割合}) \dots (式1)$$

ここで、Xはリクエスト実行時間であり、Yは装置種別固有のパーティ容量当たりの処理速度であり、それぞれストレージ装置種別によって決定される。また、動作時間割合は「(100 - CPU稼働率) [%]」で決定される。

【0152】

次に、プラン生成部214は、プランに沿ってオーナー権変更を実行した場合の(実行中の)副作用の1つとして、I/Oへの影響の予測値(I/O影響負荷)を予測する(ステップS1311)。I/O影響負荷の見積もり方法は、例えば事前に計測した値を参考に固定値としてもよいし、計測した値から作成したモデルを利用して導出してもよい。また、過去のオーナー権変更時のI/O影響負荷の実績を元に学習して予測をしてもよい。

【0153】

次に、プラン生成部214は、プランに沿ってオーナー権変更を実行した場合の(実行中の)副作用の1つとして、CPU11への影響の予測値を、CPU負荷の上昇量(CPU負荷上昇量)によって予測する(ステップS1312)。CPU負荷の見積もり方法は、例えば事前に計測した値を参考に固定値としてもよいし、計測した値から作成したモデルを利用して導出してもよい。また、過去のオーナー権変更時のCPU負荷の実績を元に学習して予測をしてもよい。また、ステップS1311とステップS1312の実行順序は入れ替えてもよい。

【0154】

次に、プラン生成部214は、ステップS1310~S1312の予測結果をプラン情報229に格納するとともに、そのプランに対応する操作に関する情報をジョブ情報230に格納する(ステップS1313)。具体的にはプラン生成部214は、プラン情報229に対して、ステップS1310で予測したオーナー権変更実行時間を予測処理時間22908に、ステップS1311で予測したI/O影響負荷をI/O影響予測22909に、ステップS1312で予測したCPU負荷上昇量をCPU影響予測22910に、それぞれ格納する。また、プラン生成部214は、ジョブ情報230に対して、当該プランにおける操作内容(例えば「オーナー権変更」)を操作名23005に、当該プランにおけるオーナー権の移動元のCPU11の識別子を移動元CPU23006に、当該プランにおけるオーナー権変更の対象となるボリューム18の識別子を移動対象Volume23007に、当該プランにおけるオーナー権の移動先のCPU11の識別子を移動先CPU23008に、それぞれ格納する。なお、具体的な説明は省略するが、ステップS1313において、プラン生成部214は、プラン情報229及びジョブ情報230に対して、上述した項目以外の項目(装置製番や装置モデル等)についても適宜情報を登録するようにしてよい。

【0155】

次に、プラン提示・修正部215が、プラン情報229及びジョブ情報230の内容をGUI等でユーザ(管理者)へ提示する(ステップS1314)。ステップS1314では、イベントを対処するためのプランに関してステップS1314までにプラン情報229及びジョブ情報230に登録された任意の情報を提示するようにしてよい。

【0156】

そして、ステップS1314による提示を受けて、ユーザ(管理者)がプランの内容を承認すると(ステップS1315)、ステップS1316に進む。

【0157】

なお、プラン提示・修正部215は、ステップS1315においてユーザから承認を得る際(またはステップS1314の提示タイミングでもよい)、プランの実行時刻に対する

10

20

30

40

50

ユーザの希望を入力可能な I / F を提示するようにする。この提示された I / F に対してユーザは必要に応じて適宜入力操作を行うことにより、プランの実行を希望する時刻（実行希望時刻）やプランの実行を禁止する時刻（実行禁止時刻）を指定することができる。

【 0 1 5 8 】

ステップ S 1 3 1 6 では、プラン提示・修正部 2 1 5 が、ステップ S 1 3 1 5 においてユーザ（管理者）から実行希望時刻または実行禁止時刻が入力（指定）されたか否かを判定する。ステップ S 1 3 1 6 で肯定結果が得られた場合、すなわち実行希望時刻または実行禁止時刻が入力された場合は（ステップ S 1 3 1 6 の YES）、ステップ S 1 3 1 7 に進む。一方、ステップ S 1 3 1 6 で否定結果が得られた場合、すなわち実行希望時刻または実行禁止時刻の何れも入力されなかった場合は（ステップ S 1 3 1 6 の NO）、プラン生成処理を終了し、ステップ S 1 4 0 0 のプラン実行処理に遷移する。

10

【 0 1 5 9 】

ステップ S 1 3 1 7 では、プラン提示・修正部 2 1 5 が、ユーザからの指定内容に応じた情報を、プラン情報 2 2 9 のユーザ指定条件 2 2 9 1 1、実行予定時刻 2 2 9 1 2、または実行禁止時刻 2 2 9 1 3 に情報を格納する。補足すると、ユーザから実行希望時刻が指定された場合、指定された時刻は実行予定時刻 2 2 9 1 2 に登録される。そしてステップ S 1 3 1 7 が完了すると、プラン生成処理を終了し、ステップ S 1 4 0 0 のプラン実行処理に遷移する。

【 0 1 6 0 】

なお、図 2 3 及び図 2 4 に示したプラン生成処理の処理手順の変形例として、ステップ S 1 3 1 5 ~ S 1 3 1 7 を省略し、計算機システム 1 0 0（主に管理サーバ 2）が自動で処理を進めるようにしてもよい。

20

【 0 1 6 1 】

（ 1 - 4 - 4 ）プラン実行処理

図 2 5 及び図 2 6 は、プラン実行処理の処理手順例を示すフローチャート（その 1，その 2）である。図 2 5，図 2 6 に示すフローチャートは、図 2 0 に示したステップ S 1 4 0 0 のプラン実行処理の詳細な処理手順例を示すものである。

【 0 1 6 2 】

図 2 5，図 2 6 によればまず、プラン実行部 2 1 6 が、CPU 負荷の情報、実行予測時間（オーナー権変更実行時間）、副作用の情報、及びユーザに指定された実行希望時刻または実行禁止時刻に基づいて、プランの実行時刻を決定する（ステップ S 1 4 0 1）。

30

【 0 1 6 3 】

ステップ S 1 4 0 1 におけるプランの実行時刻の決定方法について詳しく説明する。ユーザ（管理者）によって実行希望時刻が指定されている場合（プラン情報 2 2 9 のユーザ指定条件 2 2 9 1 1 が「希望時刻」で、実行予定時刻 2 2 9 1 2 が登録されている場合）は、実行予定時刻 2 2 9 1 2 に登録された実行希望時刻をプランの実行時刻に決定すればよい。この場合、プラン実行部 2 1 6 は、プラン情報 2 2 9 の実行予定時刻 2 2 9 1 2 を更新する必要はない。

【 0 1 6 4 】

一方、実行禁止時刻が設定されている場合（プラン情報 2 2 9 のユーザ指定条件 2 2 9 1 1 が「禁止時刻」で、実行禁止時刻 2 2 9 1 3 が登録されている場合）は、プラン実行部 2 1 6 は、実行禁止時刻 2 2 9 1 3 に登録された実行禁止時刻（時間帯でもよい）を避けてプランの実行時刻を決定し、プラン情報 2 2 9 の実行予定時刻 2 2 9 1 2 に登録する。具体的には次のような決定方法で決定すればよい。例えば CPU 負荷が一定値よりも低い時刻になると実行してもよい。また、過去一定期間の CPU 負荷の履歴を見て、負荷が低くなる時刻を予測し、実行予測時間とその間の副作用（I / O 影響、CPU 負荷上昇）から実行時刻を決定してもよい。また、過去の履歴情報を学習して将来の CPU 負荷を予測して、実行予測時間とその間の副作用から実行時刻を決定してもよい。

40

【 0 1 6 5 】

ステップ S 1 4 0 1 の終了後、プラン実行部 2 1 6 は、プランの実行時刻になったときに

50

、管理情報送受信・格納部 2 1 1 を介してストレージ装置 1 に、オーナー権変更処理のリクエスト（ジョブ情報 2 3 0 の内容）を送付する（ステップ S 1 4 0 2 ）。

【 0 1 6 6 】

そしてストレージ装置 1 の管理情報送受信部 1 2 1 が、ステップ S 1 4 0 2 で送付されたリクエストを受け付けて、管理操作実行部 1 2 3 に送付する（ステップ S 1 4 0 3 ）。

【 0 1 6 7 】

次に、管理操作実行部 1 2 3 がリクエストに応じたオーナー権変更処理を実行し、実行後に、オーナー権管理情報 1 2 7 及びキャッシュ管理情報 1 2 8 における CPU 1 1 とボリューム 1 8 との関係を更新する（ステップ S 1 4 0 4 ）。

【 0 1 6 8 】

そして管理操作実行部 1 2 3 は、管理情報送受信部 1 2 1 を介してオーナー権変更処理の実行結果をクラウド上の管理サーバ 2 に返戻する（ステップ S 1 4 0 5 ）。

【 0 1 6 9 】

そして、管理サーバ 2 側では、プラン実行完了確認部 2 1 7 によって、オーナー権変更処理のプラン実行によりイベントが解消したか否かを確認する処理として、ステップ S 1 4 0 6 ~ S 1 4 0 9 の処理が行われる。

【 0 1 7 0 】

まずステップ S 1 4 0 6 では、プラン実行完了確認部 2 1 7 が、稼働履歴情報 2 2 4 を確認する。

【 0 1 7 1 】

次に、プラン実行完了確認部 2 1 7 は、稼働履歴情報 2 2 4 に基づいて、イベントが解決したかを確認する（ステップ S 1 4 0 7 ）。イベント解決の確認方法としては、例えば、実行対象のイベントについて、ステップ S 1 2 0 0 の検知処理でイベント検知に用いられた検知条件（検知情報 2 2 6 参照）を元に、「検知期間」を短縮化した（例えば検知情報 2 2 6 の検知期間 2 2 6 0 7 が「5日」となっている場合に、短縮して1日とする）確認条件で、異常が検知されない（言い換えればイベントが解決した）状態になるかを確認する。また例えば、実行対象のイベントについて、ステップ S 1 2 0 0 の検知処理でイベント検知に用いられた検知条件を元に、「検知期間」を「最寄りの前回閾値超過を検知した曜日・時刻から 2 4 時間経過後」に置き換えた確認条件で、異常が検知されない（言い換えればイベントが解決した）状態になるかを確認する。また例えば、単純にプラン実行の完了結果を受領したことによってイベントが解決したと判断してもよい。

【 0 1 7 2 】

次に、プラン実行完了確認部 2 1 7 は、ステップ S 1 4 0 7 の確認によってイベントの解決を確認できたか否かを判定する（ステップ S 1 4 0 8 ）。ステップ S 1 4 0 8 で肯定結果が得られた場合、すなわちイベントの解決が確認できた場合は（ステップ S 1 4 0 8 の YES ）、ステップ S 1 4 1 0 に進み、ステップ S 1 4 0 8 で否定結果が得られた場合、すなわちイベントの解決が確認できなかった場合は（ステップ S 1 4 0 8 の NO ）、ステップ S 1 4 0 9 に進む。

【 0 1 7 3 】

ステップ S 1 4 0 9 では、プラン実行完了確認部 2 1 7 は、プラン実行後に一定時間が経過したかを確認する。ここでの一定時間とは、ステップ S 1 4 0 7 におけるイベント解決の確認方法で説明した確認条件を構成する「検知期間」と捉えてよい。すなわち、元の検知条件から変更された「検知期間」である「1日」や「最寄りの前回閾値超過を検知した曜日・時刻から 2 4 時間経過後」を「一定期間」としてよく、この一定期間を経過していない場合は（ステップ S 1 4 0 9 の NO ）、ステップ S 1 4 0 6 に戻り、最新の稼働履歴情報 2 2 4 に基づいてイベント解決の確認を繰り返す。一方、一定期間が経過した場合は（ステップ S 1 4 0 9 の YES ）、プラン実行完了確認部 2 1 7 はイベントが終了した（解決した）と見なし、ステップ S 1 4 1 0 に進む。

【 0 1 7 4 】

そしてステップ S 1 4 1 0 では、プラン実行完了確認部 2 1 7 がイベント情報 2 2 8 の解

10

20

30

40

50

決済み 2 2 8 0 7 を「 Y e s 」に変更し、プラン実行処理を終了する。

【 0 1 7 5 】

以上に説明したように、本実施形態に係る計算機システム 1 0 0 によれば、あるデータセンタ 3 または組織によって管理される複数のストレージ装置 1 において、CPU稼働率の不均衡が発生した場合に、その情報を収集し、不均衡の発生を検知し、検知した不均衡を解決するプランを生成し、プランを実行することができる。さらに、プラン生成時には、オーナー権変更処理の影響として処理時間の予測や、I/Oへの影響、CPU負荷自体の影響を考慮しつつ、処理の実行時刻を決定することができ、またユーザからの実行希望時刻や実行禁止時刻を受け付けることで、アプリケーションやユーザの運用への影響を考慮して処理の実行時刻を決定することができる。かくして、本実施形態に係る計算機システム 1 0 0 は、ストレージ装置 1 の CPU 1 1 で負荷に偏りが生じた場合に、オーナー権変更の処理を運用に与える影響を少なくして実行できるようになり、ストレージ装置 1 で将来発生し得る性能不安定リスクを早期に検知、対処することができるようになり、運用管理者の管理コストを低減することができる。

【 0 1 7 6 】

なお、本実施形態では、計算機システム 1 0 0 において 1 台のストレージ装置 1 内でボリューム 1 8 に対する CPU 1 1 のオーナー権を変更する場合について説明したが、本発明は 1 台のストレージ装置 1 内におけるオーナー権変更だけに限定されるものではない。例えば、本実施形態の変形例として、計算機システムが複数のストレージ装置（ストレージノード）を備える構成であっても、メモリ及びキャッシュメモリのデータを共有し、1 台のストレージ装置としてユーザに見せる場合には、本実施形態と同様にして本発明を適用可能である。

【 0 1 7 7 】

(2) 第 2 の実施形態

第 2 の実施形態では、第 1 の実施形態との相違点として、複数のストレージノードで 1 台のクラスタを構成し、さらにそのストレージノード間でメモリおよびキャッシュメモリのデータを共有しない構成におけるオーナー権変更について説明する。以下、第 2 の実施形態では、第 1 の実施形態との差分のみを説明し、重複する構成や処理等については、同一の符号を付してそれらの説明を省略する。

【 0 1 7 8 】

(2 - 1) 構成

図 2 7 は、本発明の第 2 の実施形態に係る計算機システム 2 0 0 の全体構成例を示す図である。計算機システム 2 0 0 では、図 2 に示した第 1 の実施形態に係る計算機システム 1 0 0 の全体構成との差分として、それぞれのデータセンタ 3 内において、ストレージ装置 1 a , 1 b , 1 c が、ストレージノード 1 a ' , 1 b ' , 1 c ' になり、これら複数のストレージノード 1 ' を束ねるストレージクラスタ 5 が存在する。

【 0 1 7 9 】

計算機システム 2 0 0 において、ストレージノード 1 ' の構成は、図 3 に示したストレージ装置 1 の構成と同様である。ストレージクラスタ 5 の仕様次第では、プール 1 7 a , 1 7 b 及びボリューム 1 8 a , 1 8 b , 1 8 c がストレージノード内に閉じる構成であったり、ストレージノードを跨がる構成であったりする場合がある。本実施形態は上記の何れの構成にも適用できるため、ストレージノード 1 ' の構成の図示を省略する。

【 0 1 8 0 】

なお、ストレージノード 1 ' (個別にはストレージノード 1 a ' , 1 b ' , 1 c ') の各メモリ上で動作するプログラムやテーブルは、各ストレージノードで個別の情報を保持するようにしてもよいし、代表ノードが存在し、その代表ノードのメモリ上ですべてのプログラムやテーブルを保持するようにしてもよい。本実施形態では簡単のために、代表ノードが存在する形態で説明する。第 2 の実施形態では、ストレージノード 1 ' のメモリやクラウド上の管理サーバ 2 は、第 1 の実施形態において図 4 や図 5 に示したのと同様のプログラム及び情報 (テーブル) を保持する。このうち第 2 の実施形態において第 1 の実施形態と

の差分を有する情報（テーブル）については、第 1 の実施形態で説明した符号にダッシュ「'」を添える。具体的には、構成情報 1 2 6 を構成情報 1 2 6 ' と称し、装置構成情報 2 2 3 を装置構成情報 2 2 3 ' と称する。

【0181】

図 2 8 は、第 2 の実施形態における構成情報 1 2 6 ' の一例を示す図である。構成情報 1 2 6 ' は、ストレージノード 1 a ' , 1 b ' , 1 c ' の構成情報を保持する表であって、ストレージノード 1 ' を識別するための製品番号を示す装置製番 1 2 6 0 1 ' 、ストレージノード 1 ' のモデル（型式）を示す装置モデル 1 2 6 0 2 ' 、ストレージノード 1 ' が有するリソースの種別情報を示すリソース種別 1 2 6 0 3 ' 、リソースの識別子を示すリソース ID 1 2 6 0 4 ' 、各リソース間の関連を示す関連リソース 1 2 6 0 5 ' 、各リソースのスペック（サイズ、速度等）を示すスペック 1 2 6 0 6 ' 、及び各ストレージノード 1 ' におけるリソースの属性情報を示す属性 1 2 6 0 7 ' を管理する。

【0182】

構成情報 1 2 6 ' における第 1 の実施形態の構成情報 1 2 6 との差分は、管理対象のリソース種別 1 2 6 0 3 ' に「Node」が追加されること、及びリソース種別 1 2 6 0 3 ' が CPU , Cache , Port , Drive の場合は関連リソース 1 2 6 0 5 ' にノードの情報が設定されることである。また、ストレージクラス 5 の仕様によっては、プールやボリュームがストレージノードをまたがない場合があり、その場合は、プールやボリュームの関連リソース 1 2 6 0 5 ' にノードの情報が設定されてもよい。

【0183】

図 2 9 は、第 2 の実施形態における装置構成情報 2 2 3 ' の一例を示す図である。装置構成情報 2 2 3 ' は、ストレージノード 1 ' を保有または管理する組織の識別子を示す組織 ID 2 2 3 0 1 ' 、ストレージノード 1 ' の製品番号を示す装置製番 2 2 3 0 2 ' 、ストレージノード 1 ' のモデル（型式）を表す装置モデル 2 2 3 0 3 ' 、ストレージノード 1 ' が有するリソースの種別情報を示すリソース種別 2 2 3 0 4 ' 、リソースの識別子を示すリソース ID 2 2 3 0 5 ' 、各リソース間の関連を示す関連リソース 2 2 3 0 6 ' 、リソースのスペック（サイズ、速度等）を示すスペック 2 2 3 0 7 ' 、及びストレージノード 1 ' におけるリソースの属性情報を示す属性 2 2 3 0 8 ' を保持する。

【0184】

装置構成情報 2 2 3 ' における第 1 の実施形態の装置構成情報 2 2 3 との差分は、図 2 8 に示した構成情報 1 2 6 ' の場合と同様、ストレージノード 1 ' に関する情報が追加されていることである。差分の具体的な内容は図 2 8 の構成情報 1 2 6 ' と同様であるため、説明を省略する。

【0185】

（2-2）処理

第 2 の実施形態に係る計算機システム 2 0 0 において実行される処理について説明する。第 2 の実施形態における処理フローは、第 1 の実施形態で図 2 0 に示した処理フローと同様、情報収集処理、検知処理、プラン生成処理、そしてプラン実行処理の流れとなる。第 1 の実施形態との差分としては、各ストレージノード 1 ' ではメモリ及びキャッシュメモリの内容を共有しないため、オーナー権変更時はノード間でキャッシュデータをコピーする必要があり、及びノード跨りでキャッシュが効かなくなることからボリュームも移行する必要があり、である。そのため、第 2 の実施形態では、プラン生成時に影響を予測する処理が第 1 の実施形態とは異なる。

【0186】

図 3 0 は、第 2 の実施形態におけるプラン生成処理の処理手順例を示すフローチャートである。図 3 0 に示したフローチャートは、プラン生成処理の後半部にあたり、図 2 3 及び図 2 4 に示した第 1 の実施形態におけるプラン生成処理の処理手順のうち、図 2 4 の処理フローに置き換えられるものである。さらに言えば、図 3 0 に示す処理フローは、図 2 4 に示されたステップ S 1 3 1 0 がステップ S 1 3 1 0 ' に置き換えられる点で異なる。

【0187】

20

30

40

50

図30に示したように、ステップS1310'では、プラン生成部214が、キャッシュ容量、移動先ストレージノードのCPU稼働率、移動先ストレージノードのキャッシュのWritePendingRate、ボリューム容量、及び装置種別（装置種別から紐付く各装置のスペック情報）に基づいて、オーナー権変更実行時間を予測する。以下に詳しく説明する。

【0188】

オーナー権変更実行時間は、構成変更リクエスト受付時間と、キャッシュ管理情報128を定期的に更新する処理時間（キャッシュメモリの管理ポインタ情報の移行時間）、キャッシュデータをコピーする時間、及びボリュームを移行する時間によって決まる。

【0189】

上記のキャッシュ管理情報128を定期的に更新する処理時間は、キャッシュ容量と移動先ストレージノードのWritePendingRateとボリューム容量とに比例し、移動先ストレージノードのCPU稼働率の空き（動作時間割合）、装置のハードウェアスペックに反比例する。また、常にバックグラウンドでキャッシュ管理（移行処理含む）のためにデータ領域のスキャンを実施している。スキャン処理とオーナー権移動処理は非同期で実行されるため、現在進行中のスキャン処理において、移行対象のキャッシュ領域が既に通り過ぎている場合がある。このため、確実にスキャンされるように、スキャン2週分の時間で見積もる。構成変更リクエストの時間は移動するボリューム数によって決まる。以上のことを踏まえると、オーナー権変更実行時間は、以下の式2に従って予測することができる。

【0190】

【数2】

実行時間

$$= X' * \text{ボリューム数} \\ + \text{キャッシュ容量} * \text{WritePendingRate} / (Y' * \text{動作時間割合}) \\ + (\text{移動するボリュームの合計容量} + \text{移動するボリュームが利用するキャッシュの容量}) / Z \quad \dots \text{(式2)}$$

ここで、X'はリクエスト実行時間であり、Y'は装置種別固有のデータ容量当たりの処理速度であり、Zは装置種別固有のデータコピー容量当たりの処理速度であり、それぞれストレージノードの装置種別によって決定される。また、動作時間割合は「(100 - CPU稼働率) [%]」で決定される。また、移動するボリュームが利用するキャッシュの容量は、キャッシュ管理情報128の管理ポインタ数（キャッシュ管理ポインタ12803）から求められる（図9参照）。

【0191】

以上に説明した第2の実施形態に係る計算機システム200によれば、第1の実施形態に係る計算機システム100のようにメモリ及びキャッシュメモリのデータを共有する構成でなくても、複数のストレージノード1'間でメモリやキャッシュメモリを共有しない構成でのストレージクラス5にて、CPU稼働率の不均衡が発生した場合に、その情報を収集し、不均衡の発生を検知し、検知した不均衡を解決するプランを生成し、プランを実行することができる。さらに、プラン生成時には、オーナー権変更処理の影響として処理時間の予測や、I/Oへの影響、CPU負荷自体の影響を考慮しつつ、処理の実行時刻を決定することができ、またユーザからの実行希望時刻や実行禁止時刻を受け付けることで、アプリケーションやユーザの運用への影響を考慮して処理の実行時刻を決定することができる。かくして、本実施形態に係る計算機システム200は、ストレージノード1'のCPUで負荷に偏りが生じた場合に、オーナー権変更の処理を運用に与える影響を少なくして実行することができるようになり、ストレージノード1'で将来発生し得る性能不安定リスクを早期に検知、対処することができるようになり、運用管理者の管理コストを低減することができる。

【0192】

10

20

30

40

50

なお、本発明は上述した各実施形態に限定されるものではなく、様々な変形例を含む。例えば、第1の実施形態では管理プログラムはすべてクラウド上に存在するとしたが、データセンタ内またはデータセンタ外のクラウドでない管理サーバ上のメモリに存在するとしてもよい。

【0193】

その他、上述した実施形態は本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。また、矛盾しない限りにおいて、ある実施形態の構成の一部を他の実施形態の構成で置き換え、ある実施形態の構成に他の実施形態の構成を加えることも可能である。また、各実施形態の構成の一部について、構成の追加、削除、置換、統合、又は分散をすることが可能である。また実施形態で示した構成及び処理は、処理効率又は実装効率に基づいて適宜分散、統合、又は入れ替えることが可能である。

10

【0194】

また、図面において制御線や情報線は説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。実際には殆ど全ての構成が相互に接続されていると考えてもよい。

【符号の説明】

【0195】

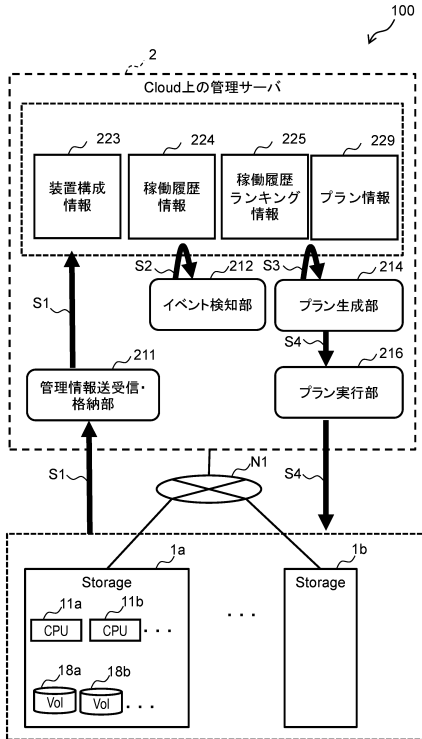
1	ストレージ装置	
2	管理サーバ	20
3	データセンタ	
4	スイッチ	
5	ストレージクラスタ	
11	CPU	
12	メモリ	
13	ドライブ	
14	ネットワークI/F	
15	バス	
16	キャッシュメモリ(キャッシュ)	
17	プール	30
18	ボリューム	
100, 200	計算機システム	
121	管理情報送受信部	
122	稼働情報収集・格納部	
123	管理操作実行部	
124	キャッシュ管理部	
125	稼働情報	
126	構成情報	
127	オーナー権管理情報	
128	キャッシュ管理情報	40
211	管理情報送受信・格納部	
212	イベント検知部	
213	イベント登録部	
214	プラン生成部	
215	プラン提示・修正部	
216	プラン実行部	
217	プラン実行完了確認部	
221	組織情報	
222	装置情報	
223	装置構成情報	50

- 2 2 4 稼働履歴情報
- 2 2 5 稼働履歴ランキング情報
- 2 2 6 検知情報
- 2 2 7 性能異常履歴情報
- 2 2 8 イベント情報
- 2 2 9 プラン情報
- 2 3 0 ジョブ情報

【 図 面 】

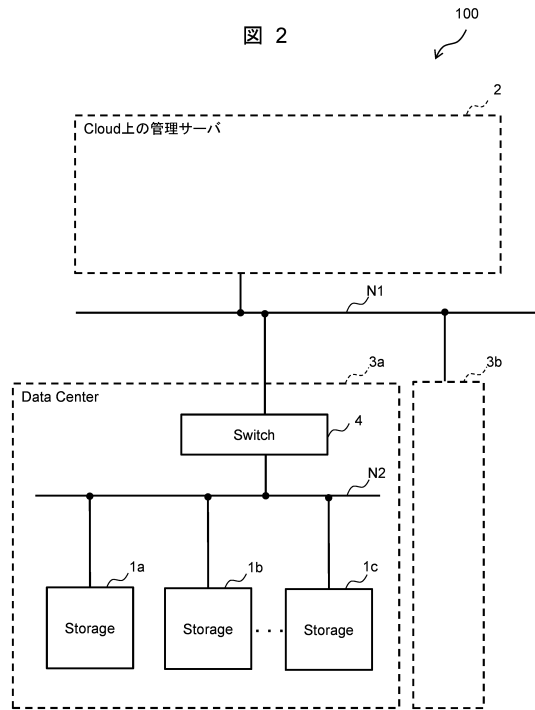
【 図 1 】

図 1



【 図 2 】

図 2



10

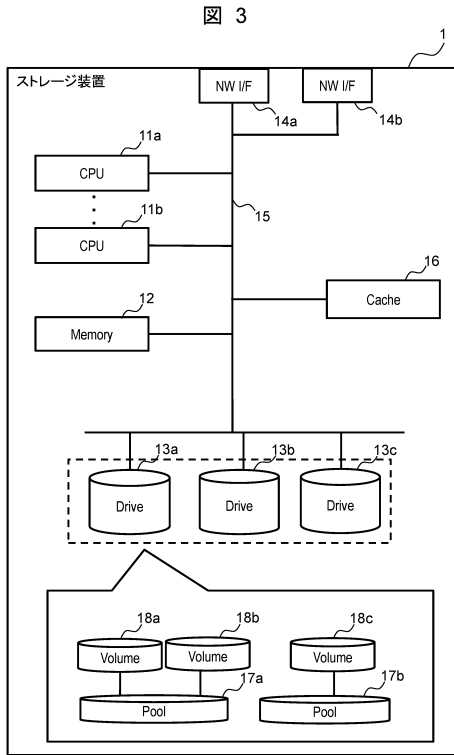
20

30

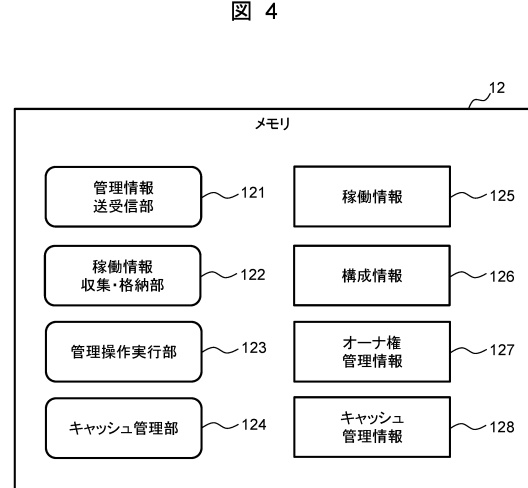
40

50

【 図 3 】



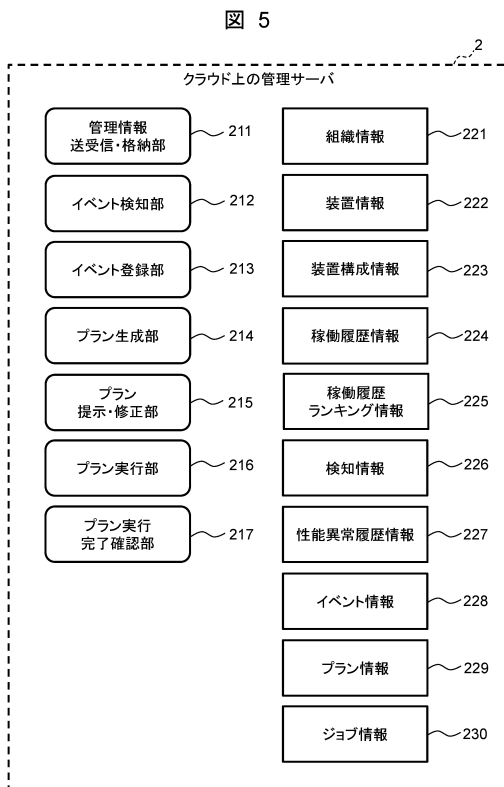
【 図 4 】



10

20

【 図 5 】



【 図 6 】

図 6

リソース種別	リソースID	メトリック	値
CPU	CPU1	TotalBusyTime	13500
		TargetBusyTime	8500
		UsageRanking[]	Vol1:1000, Vol3:500,...
	...		
	CPU2		
	...		
Volume	Vol1	SequentialReadIoCount	100
		SequentialWriteIoCount	10
		RandomReadIoCount	50
		RandomWriteIoCount	40
		SequentialReadHitIoCount	10
		RandomReadHitIoCount	5
		ReadAccumulatedResponseTime	12000
		WriteAccumulatedResponseTime	6000
		SequentialReadBlockCount	10
		SequentialWriteBlockCount	100
	RandomReadBlockCount	50	
RandomWriteBlockCount	50		
...		40	
	Vol2		
	...		
Cache	Cache1	WritePendingRate	20%
	...		
...			

30

40

50

【 図 7 】

図 7

装置 製番	装置 モデル	リソース 種別	リソース ID	関連 リソース	スペック	属性
001	Model A	CPU	CPU1	-	4.0GHz	
			CPU2	-	4.0GHz	
			...			
		Cache	Cache1	-	128GB	
			Cache2	-	128GB	
			...			
		Port	Port1	-	100GB/sec	
			Port2	-	100GB/sec	
			...			
		Drive	Drive1	Pool1	1TB MLC	
			Drive2	Pool1	1TB MLC	
			...			
		Pool	Pool1	-	1.5TB	
			Pool2	-	15TB	
			...			
		Volume	Vol1	Pool1	1TB	
Vol2	Pool2		5TB	圧縮		
Vol3	Pool2		5TB	コピー (Vol4)		
...						
...						

【 図 8 】

図 8

VolumeID	CPU_ID
Vol1	CPU1
Vol2	CPU1
Vol3	CPU2
...	

10

20

【 図 9 】

図 9

CPU_ID	VolumeID	キャッシュ管理ポインタ	状態
CPU1			移行中
CPU2			通常
...			

【 図 10 】

図 10

ID	組織コード	タイプ	名前
1	Org1	Customer	ABC
2	Org2	Customer	DEF
3	Org3	Partner	XYZ
...			

30

40

50

【 図 1 1 】

図 11

組織ID	装置製番	装置モデル	名前	状態
1	001	Model A	Storage A	Online
1	002	Model B	Storage B	Offline
2				
...				

【 図 1 2 】

図 12

組織 ID	装置 製番	装置 モデル	リソース 種別	リソース ID	関連 リソース	スペック	属性
1	001	Model A	CPU	CPU1	-	4.0GHz	
				CPU2	-	4.0GHz	
				...			
			Cache	Cache1	-	128GB	
				Cache2	-	128GB	
				...			
			Port	Port1	-	100GB/ sec	
				Port2	-	100GB/ sec	
				...			
			Drive	Drive1	Pool1	1TB MLC	
				Drive2	Pool1	1TB MLC	
				...			
			Pool	Pool1	-	1.5TB	
				Pool2	-	15TB	
				...			
			Volume	Volume 1	Pool1, CPU1	1TB	
				Volume 2	Pool2, CPU2	5TB	圧縮
				Volume 3	Pool2, CPU1	5TB	コピー (Vol4)
				...			
			...				

10

20

【 図 1 3 】

図 13

組織 ID	装置 製番	装置 モデル	リソース 種別	リソース ID	時刻	メトリック	値
1			CPU	CPU1	10:00 :00	TotalBusyRate	30%
						TargetBusyRate	20%
						UsageRanking[]	Vol1:10%, Vol3:5%,...
						...	
						...	
						...	
						...	
						...	
						...	
						...	
			Volume	Vol1	10:00 :00	SequentialReadIOPS	100
						SequentialWriteIOPS	10
						RandomReadIOPS	50
						RandomWriteIOPS	40
						SequentialReadHitIOPS	10
						RandomReadHitIOPS	5
						ReadResponseTime	200
						WriteResponseTime	100
						SequentialReadTransferRate	300
						SequentialWriteTransferRate	100
RandomReadTransferRate	50						
RandomWriteTransferRate	50						
...	40						
...							
...							
...							
...							

30

40

【 図 1 4 】

図 14

組織 ID	装置 製番	装置 モデル	CPU_ID	時刻	Rank	VolumeID	BusyRate InPercent
			CPU1	10:00:00	1	Vol1	15
					2	Vol3	10
					...		
					...		
					...		
			CPU2	...			
				...			
				...			
				...			
				...			

50

【 図 1 5 】

図 15

ID	組織ID	装置製番	装置モデル	閾値	最小検知間隔	検知期間	検知数
1	1	001	Model A	45%	1hour	5days	30
2	2						
...							

【 図 1 6 】

図 16

ID	組織ID	装置製番	装置モデル	CPU_ID	閾値	異常発生時刻	稼働率
1	1	001	Model A	CPU1	40%	10:00:00	50%
2	2						
...							

10

20

【 図 1 7 】

図 17

ID	組織ID	装置製番	装置モデル	発生時刻	更新時刻	解決済
1	1	001	Model A	10:00:00	-	No
2	1	002	Model B	10:00:05	11:00:00	Yes
3	2					
...						

【 図 1 8 】

図 18

イベント ID	装置製番	装置モデル	状態	プラン登録時刻	プラン更新時刻	予測処理 I/O 影響時間	CPU 影響率	ユーザ指定条件	実行予定時刻	実行禁止時刻
									希望時刻	禁止時刻
1	001	Model A	Scheduled	10:05:00	-	15min	10%	禁止時刻	-	4:00-7:00
2	002	Model B	Scheduled	0:00:00:00	-	20min	10%	希望時刻	4:00:00	-
...										

30

40

50

【 図 19 】

図 19

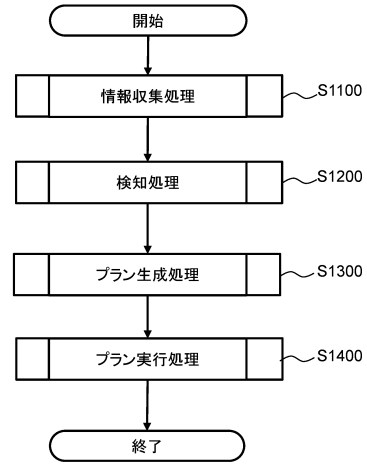
230

ID	プランID	装置製番	装置モデル	操作名	移動元CPU	移動対象Volume	移動先CPU
1	1	001	Model A	オーナー権変更	CPU1	Vo1	CPU2
2	1	001	Model A	オーナー権変更	CPU1	Vo13	CPU2
...							

【 図 20 】

図 20

S1000



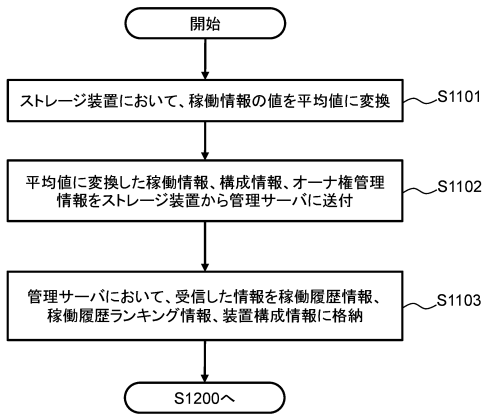
10

20

【 図 21 】

図 21

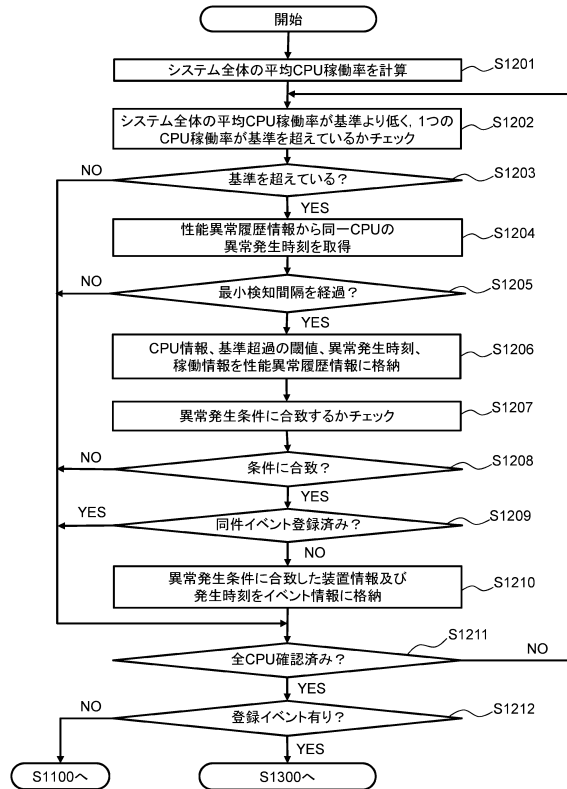
S1100



【 図 22 】

図 22

S1200

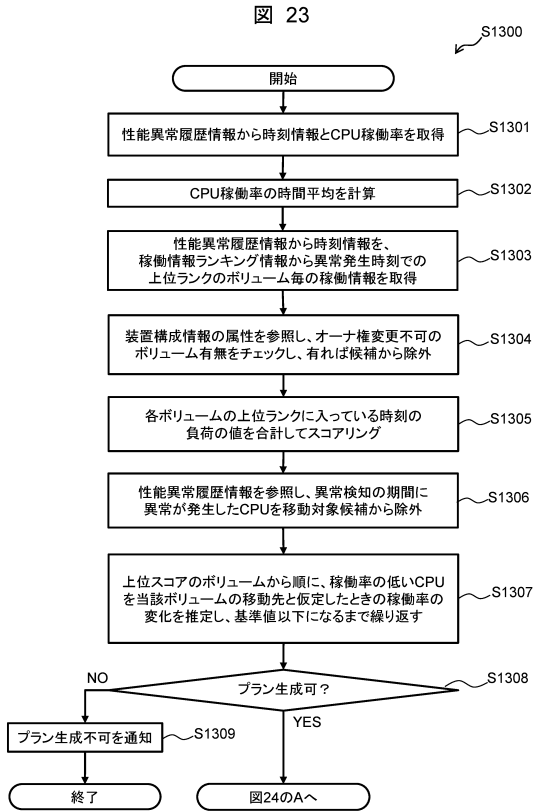


30

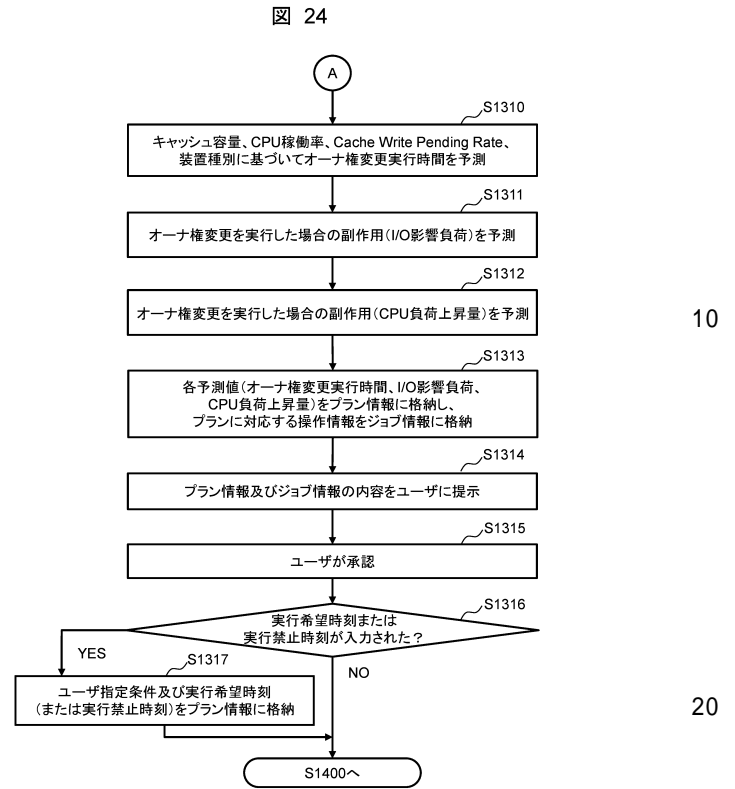
40

50

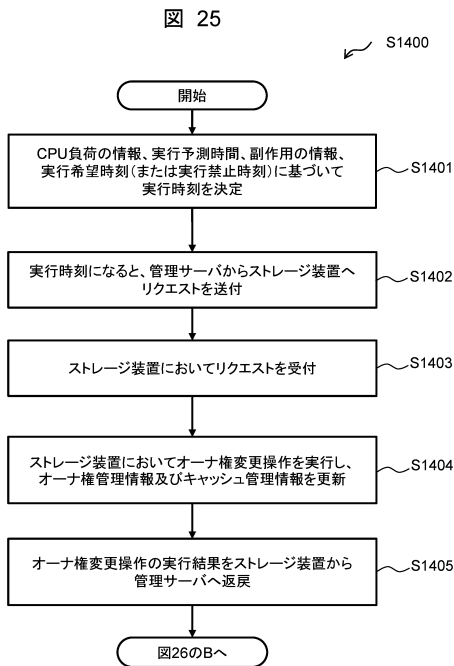
【 図 2 3 】



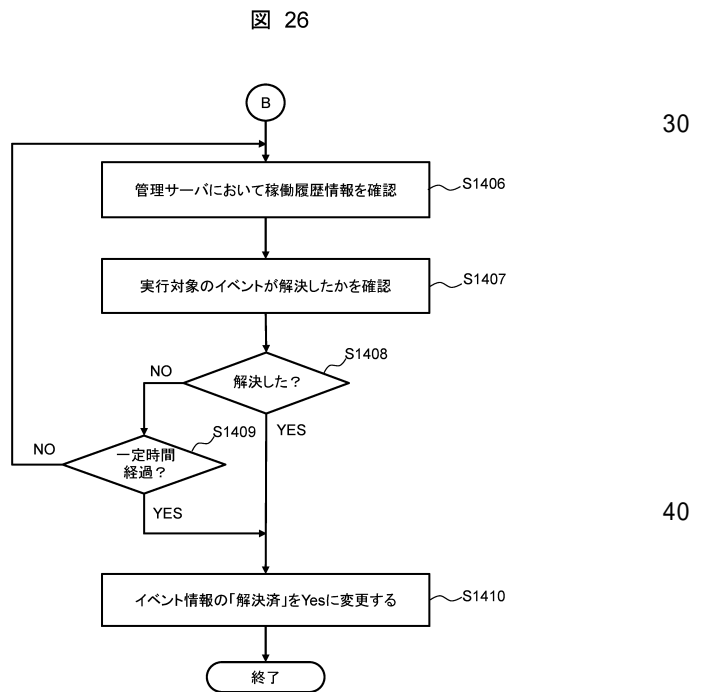
【 図 2 4 】



【 図 2 5 】



【 図 2 6 】



10

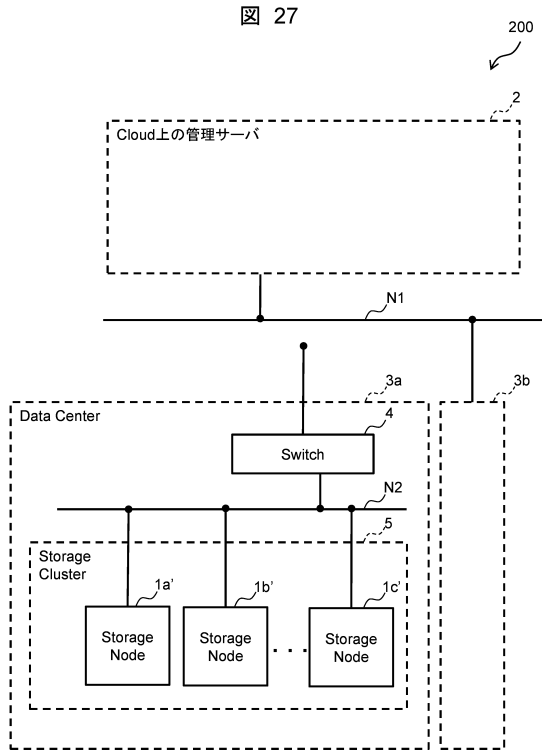
20

30

40

50

【 図 2 7 】



【 図 2 8 】

図 28

装置製番	装置モデル	リソース種別	リソースID	関連リソース	スペック	属性	
12601'	Model A	Node	Node1	-	-		
			...	-	-		
		CPU	CPU1	Node1	4.0GHz		
			CPU2	Node2	4.0GHz		
			...				
		Cache	Cache1	Node1	128GB		
			Cache2	Node1	128GB		
			...				
		Port	Port1	Node1	100GB/sec		
			Port2	Node1	100GB/sec		
			...				
		Drive	Drive1	Node1, Pool1	1TB MLC		
			Drive2	Node1, Pool1	1TB MLC		
			...				
		Pool	Pool1	-	1.5TB		
			Pool2	-	15TB		
			...				
		Volume	Vol1	Pool1	1TB		
			Vol2	Pool2	5TB	圧縮	
			Vol3	Pool2	5TB	コピー (Vol4)	
...							
...							

10

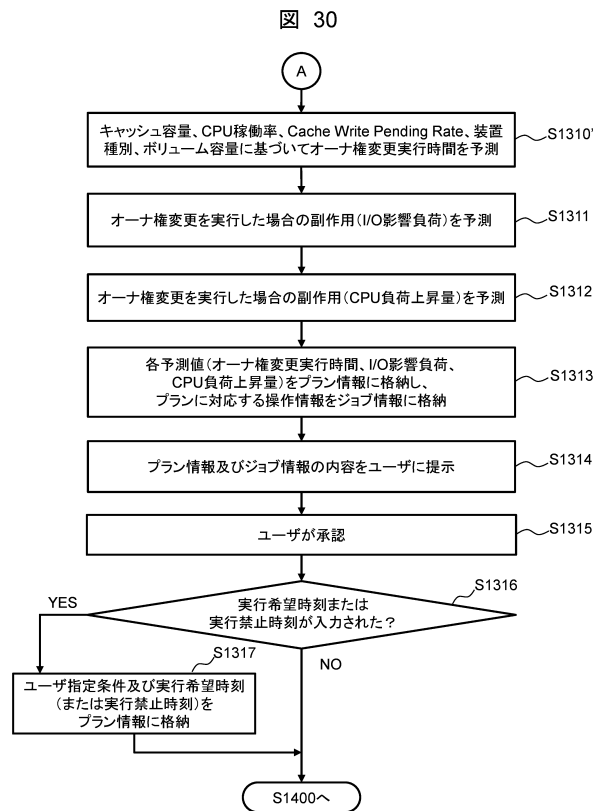
20

【 図 2 9 】

図 29

組織ID	装置製番	装置モデル	リソース種別	リソースID	関連リソース	スペック	属性	
1	001	Model A	Node	Node1				
				...				
			CPU	CPU1	Node1	4.0GHz		
				CPU2	Node2	4.0GHz		
				...				
			Cache	Cache1	Node1	128GB		
				Cache2	Node1	128GB		
				...				
			Port	Port1	Node1	100GB/sec		
				Port2	Node1	100GB/sec		
				...				
			Drive	Drive1	Node1, Pool1	1TB MLC		
				Drive2	Node1, Pool1	1TB MLC		
				...				
			Pool	Pool1	-	1.5TB		
				Pool2	-	15TB		
				...				
			Volume	Volume1	Pool1, CPU1	1TB		
				Volume2	Pool2, CPU2	5TB	圧縮	
				Volume3	Pool2, CPU1	5TB	コピー (Vol4)	
...								
...								

【 図 3 0 】



30

40

フロントページの続き

(51)国際特許分類

G 0 6 F 11/34 (2006.01)

F I

G 0 6 F	9/50	1 5 0 D
G 0 6 F	11/34	1 4 2
G 0 6 F	11/07	1 6 0

テーマコード (参考)

Fターム (参考) 5B042 GA34 JJ29 KK04 KK17 MA08 MC21 MC24 MC35