

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7141908号
(P7141908)

(45)発行日 令和4年9月26日(2022.9.26)

(24)登録日 令和4年9月14日(2022.9.14)

(51)国際特許分類

G 0 6 F	11/14 (2006.01)	F I	G 0 6 F	11/14	6 5 8
G 0 6 F	16/11 (2019.01)		G 0 6 F	16/11	
G 0 6 F	16/14 (2019.01)		G 0 6 F	16/14	

請求項の数 10 (全28頁)

(21)出願番号 特願2018-198161(P2018-198161)
 (22)出願日 平成30年10月22日(2018.10.22)
 (65)公開番号 特開2020-67681(P2020-67681A)
 (43)公開日 令和2年4月30日(2020.4.30)
 審査請求日 令和3年4月16日(2021.4.16)

(73)特許権者	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	110002365 特許業務法人サンネクスト国際特許事務所
(72)発明者	小林 恵美子 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	高重 聰一 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
審査官	漆原 孝治

最終頁に続く

(54)【発明の名称】 データ管理システムおよびデータ管理方法

(57)【特許請求の範囲】**【請求項1】**

複数のストレージ装置を含んで構成されるデータ管理システムであって、
要求対象の時間を示す情報と、要求対象のデータを識別するデータ識別子を示す情報と、
要求対象の精度に関する要求レベルの情報とが含まれる要求を受け付ける要求受付部と、
データを保管するための保管データが取得された時間を示す情報と、前記保管データが
格納されるストレージ装置を示す情報とが対応付けられているデータ管理情報に基づいて
、前記複数のストレージ装置に記憶される保管データの中から、前記要求受付部により受
け付けられた要求に含まれる要求対象の時間に合う保管データを特定するデータ特定部と、
前記データ管理情報に基づいて、前記データ特定部により特定された保管データについて、
前記要求に含まれる要求対象の時間との時間差分を算出し、算出した時間差分が前記
要求に含まれる要求レベルを満たすか否かを評価するデータ評価部と、

前記データ特定部による特定結果を出力する出力部と、
 を備えることを特徴とするデータ管理システム。

【請求項2】

前記データ管理情報には、データを識別する識別子を示す情報が更に対応付けられて含まれ、

前記要求に含まれる要求対象の時間を示す情報として、データを要求する日時を示す情報が指定され、

前記要求には、要求対象のデータを識別する識別子を示す情報が更に含まれ、

前記データ特定部は、前記データ管理情報に基づいて、前記要求に含まれるデータ識別子に対応するデータの保管データがあると判定した場合、前記要求に含まれる要求対象の時間と一致する保管データが存在するか否かを判定し、存在すると判定したときは、前記保管データを特定し、存在しないと判定したときは、前記要求に含まれる要求対象の時間に最も近い日時の保管データを特定する、

ことを特徴とする請求項1に記載のデータ管理システム。

【請求項3】

前記データ管理情報には、データを識別する識別子を示す情報と、保管データ内の最初のタイムスタンプである初回タイムスタンプおよび前記保管データ内の最後のタイムスタンプである最終タイムスタンプの情報とが更に対応付けられ、

前記要求に含まれる要求対象の時間を示す情報として、データを要求する期間を示す開始日時と終了日時との情報が指定され、

前記要求には、要求対象のデータを識別する識別子を示す情報が更に含まれ、

前記データ特定部は、前記データ管理情報に基づいて、前記要求に含まれるデータ識別子に対応するデータの保管データがあると判定した場合、前記データの保管データにタイムスタンプの情報があるか否かを判定し、タイムスタンプの情報があると判定したときは、前記データの保管データの初回タイムスタンプが前記要求に含まれる開始日時以前かつ最新である第1の保管データを特定し、前記データの保管データの最終タイムスタンプが前記要求に含まれる終了日時以降となる第2の保管データを特定し、前記第1の保管データから前記第2の保管データまでの全ての保管データを特定する、

ことを特徴とする請求項1に記載のデータ管理システム。

【請求項4】

前記データ管理情報には、保管データの取得単位として、全てのデータをバックアップするフルバックアップ、フルバックアップからの差分をバックアップする差分バックアップ、または前回のバックアップからの変更分をバックアップする増分バックアップを示す取得単位の情報が更に対応付けられて含まれ、

前記データ特定部は、前記データ管理情報に基づいて、前記要求に含まれる要求対象の時間に合うバックアップデータが存在するか否かを判定し、存在すると判定した場合、前記バックアップデータの取得単位が差分バックアップであるときは、前記バックアップデータの直近のフルバックアップのフルバックアップデータを特定し、前記バックアップデータの取得単位が増分バックアップであるときは、前記バックアップデータの直近のフルバックアップのフルバックアップデータと、前記フルバックアップデータから前記バックアップデータまでの増分バックアップデータと、を特定する、

ことを特徴とする請求項1に記載のデータ管理システム。

【請求項5】

前記要求には、要求対象のデータを識別するデータ識別子を示す情報と、要求対象の可能待ち時間の情報とが更に含まれ、

前記データ特定部により特定された保管データの復元にかかる時間を算出し、算出した復元にかかる時間が前記要求に含まれる可能待ち時間以内であるか否かを評価するデータ評価部を備える、

ことを特徴とする請求項1に記載のデータ管理システム。

【請求項6】

前記データ管理情報には、データを識別する識別子を示す情報が更に対応付けられ、

前記要求には、業務を識別可能な業務識別子を示す情報が更に含まれ、

前記データ特定部は、業務を識別可能な業務識別子を示す情報と、前記業務に係るデータを識別するデータ識別子を示す情報とが対応付けられている業務データ管理情報に基づいて、前記要求に含まれる業務識別子に対応するデータ識別子を特定し、前記データ管理情報に基づいて、特定したデータ識別子に対応するデータの保管データがあるか否かを判定する、

ことを特徴とする請求項1に記載のデータ管理システム。

10

20

30

40

50

【請求項 7】

前記出力部は、ストレージ装置を識別可能な情報と、前記ストレージ装置のプロトコルを示す情報とが対応付けられているストレージ装置管理情報に基づいて、前記データ特定部により特定された保管データが異なるプロトコルのストレージ装置から特定された場合、プロトコルの変換を行い、同じ形式の保管データとして出力する、
ことを特徴とする請求項 1 に記載のデータ管理システム。

【請求項 8】

前記要求には、要求対象のデータを識別するデータ識別子を示す情報と、要求対象の可能待ち時間を示す情報とが含まれ、

前記データ特定部により特定された保管データを識別する保管識別子を示す情報と、前記出力部が前記保管データの復元に要した時間を示す情報と、前記要求に含まれる可能待ち時間を示す情報を対応付けて記憶する記憶部と、

ストレージ装置を識別可能な情報と、前記ストレージ装置に設けられる保管データを格納するディスクを示す情報と前記ディスクの性能を示す情報とが対応付けられているストレージ装置管理情報と、前記記憶部に記憶されている情報に基づいて、復元に要した時間が可能待ち時間よりも長い保管データを格納するディスクよりも性能の高いディスクに前記保管データを移行する処理、および／または、復元に要した時間が可能待ち時間よりも短い保管データを格納するディスクよりも性能の低いディスクに前記保管データを移行する処理を行うデータ移行部と、を備える、

ことを特徴とする請求項 1 に記載のデータ管理システム。

【請求項 9】

前記データ特定部により特定された保管データを識別する保管識別子を示す情報と、前記出力部が前記保管データの復元に要した時間を示す情報と、前記保管データが取得された時間を示す情報を対応付けて記憶する記憶部と、

前記記憶部に記憶されている情報から算出した保管データの復元に要した時間の代表値に基づいて、前記保管データの保管先を変更する処理、および／または、前記記憶部に記憶されている情報から算出した保管データが使用される頻度に基づいて、保管データを取得する間隔を変更する処理を行う処理部を備える、

ことを特徴とする請求項 1 に記載のデータ管理システム。

【請求項 10】

複数のストレージ装置を含んで構成されるデータ管理システムにおけるデータ管理方法であって、

要求受付部が、要求対象の時間を示す情報と、要求対象のデータを識別するデータ識別子を示す情報と、要求対象の精度に関する要求レベルの情報とが含まれる要求を受け付ける第 1 のステップと、

データ特定部が、データを保管するための保管データが取得された時間を示す情報と、前記保管データが格納されるストレージ装置を示す情報とが対応付けられているデータ管理情報に基づいて、前記複数のストレージ装置に記憶される保管データの中から、前記要求受付部により受け付けられた要求に含まれる要求対象の時間に合う保管データを特定する第 2 のステップと、

データ評価部が、前記データ管理情報に基づいて、前記データ特定部により特定された保管データについて、前記要求に含まれる要求対象の時間との時間差分を算出し、算出した時間差分が前記要求に含まれる要求レベルを満たすか否かを評価する第 3 のステップと、

出力部が、前記データ特定部による特定結果を出力する第 4 のステップと、
を備えることを特徴とするデータ管理方法。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、概して、要求対象のデータを提供する技術に関するものである。

【背景技術】

10

20

30

40

50

【0002】

企業内の業務システムにおいて、業務アプリケーションが生成するデータは、記憶装置に保存される。記憶装置のハードウェア障害、ソフトウェア障害、または人的エラーによるデータ破損に備えたデータ保護のため、バックアップを取得する運用が行われる。バックアップを定期的に取得し、複数世代のバックアップデータが保存される場合もある。ここでは、業務アプリケーションが生成するデータが保存される記憶装置を1次ストレージ、バックアップデータが保存される記憶装置を2次ストレージと呼ぶ。1次ストレージと2次ストレージとは、同じ装置内に設けられる場合もあるが、バックアップデータは、保護用であり、使用頻度が低いことから、性能の低いディスク、別の装置などに保存される場合もある。さらにバックアップデータは、長期の保存用にアーカイブ用の記憶装置に保存される場合もある。保存される記憶装置は、ブロックストレージ、ファイルストレージ、オブジェクトストレージなど、仕様およびプロトコルが異なる可能性がある。

【0003】

業務アプリケーションのデータに対しては、企業内での分析、監査、アプリケーションの開発テストといった業務アプリケーション以外の用途での利用要求がある。その場合、データを管理するデータ管理者は、利用要求に合致するデータを抽出して提供する。要求されたデータは、1次ストレージにある場合、2次ストレージにある場合、またはアーカイブ用ストレージにしかない場合もあり、データ管理者は、データが格納されているストレージを特定してデータを抽出しなければならないため、作業負荷がかかる。

【0004】

ここで、業務アプリケーションのデータを別用途に提供する場合、業務アプリケーションには影響を与えないことが望ましい。よって、1次ストレージのデータについて、スナップショットを取得して提供する方法がある。この点、データの提供のために、データオブジェクトのスナップショットを取得し、ライブクローンボリュームを作成し、アプリケーションにマウントさせる一連のワークフローを定義し、自動実行する方法が示されている（特許文献1参照）。

【先行技術文献】**【特許文献】****【0005】****【文献】米国特許第9665437号明細書**

30

【発明の概要】**【発明が解決しようとする課題】****【0006】**

特許文献1に記載の技術では、利用にあわせてワークフローを定義し、取得したスナップショットからデータを提供する。しかしながら、特許文献1に記載の技術は、バックアップが既に取得されているバックアップデータには適用することができない。このため、利用したい単位が過去1年の顧客毎の消費分析といった場合、必要なデータが複数のアプリケーション（顧客名簿管理、売り上げ管理など）のデータであって、長期期間分であるため、同一のストレージには保存されない可能性が高く、さらに、データの利用要求で指定される時間とバックアップの取得時間とは一致しない可能性が高いため、要求されたデータを提供することが困難となってしまう。

【0007】

本発明は以上の点を考慮してなされたもので、要求対象の時間に合ったデータを提供し得るデータ管理システムを提案しようとするものである。

【課題を解決するための手段】**【0008】**

かかる課題を解決するため本発明においては、複数のストレージ装置を含んで構成されるデータ管理システムであって、要求対象の時間を示す情報が含まれる要求を受け付ける要求受付部と、データを保管するための保管データが取得された時間を示す情報と、前記保管データが格納されるストレージ装置を示す情報とが対応付けられているデータ管理情

10

20

40

50

報に基づいて、前記複数のストレージ装置に記憶される保管データの中から、前記要求受付部により受け付けられた要求に含まれる要求対象の時間に合う保管データを特定するデータ特定部と、前記データ特定部による特定結果を出力する出力部と、を設けるようにした。

【0009】

上記構成では、例えば、要求対象の時間と、データの保管データが取得された日時との関連付けが把握されているので、要求されたデータの時間要件に合った保管データを特定でき、当該保管データに係る情報から当該保管データを得ることができるようになる。

【発明の効果】

【0010】

本発明によれば、要求対象のデータを提供することができる。

【図面の簡単な説明】

【0011】

【図1】第1の実施の形態によるデータ管理システムに係る構成の一例を示す図である。

【図2】第1の実施の形態による管理サーバのハードウェア構成の一例を示す図である。

【図3】第1の実施の形態によるデータ管理プログラムの機能モジュールに係る構成の一例を示す図である。

【図4】第1の実施の形態によるデータ管理情報の一例を示す図である。

【図5】第1の実施の形態によるA P管理情報の一例を示す図である。

【図6】第1の実施の形態によるデータ属性管理情報の一例を示す図である。

【図7】第1の実施の形態によるデータ提供処理に係るフローチャートの一例を示す図である。

【図8】第1の実施の形態による日時指定時のデータ特定処理に係るフローチャートの一例を示す図である。

【図9】第1の実施の形態による期間指定時のデータ特定処理に係るフローチャートの一例を示す図である。

【図10】第1の実施の形態による期間指定時のバックアップデータの選択の概念図である。

【図11】第1の実施の形態によるストレージ装置管理情報の一例を示す図である。

【図12】第1の実施の形態による要求レベル確認処理に係るフローチャートの一例を示す図である。

【図13】第1の実施の形態による復元時間確認処理に係るフローチャートの一例を示す図である。

【図14】第1の実施の形態による提供形態決定処理に係るフローチャートの一例を示す図である。

【図15】第1の実施の形態による入出力インターフェースの一例を示す図である。

【図16】第2の実施の形態による業務データ管理情報の一例を示す図である。

【図17】第3の実施の形態によるストレージ階層管理情報の一例を示す図である。

【図18】第3の実施の形態による使用履歴管理情報の一例を示す図である。

【発明を実施するための形態】

【0012】

以下図面について、本発明の一実施の形態を詳述する。

【0013】

(1) 第1の実施の形態

図1において、100は全体として第1の実施の形態によるデータ管理システムを示す。

【0014】

図1は、データ管理システム100に係る構成の一例を示す図である。データ管理システム100は、管理サーバ110、ホスト計算機120、SDS (Software Defined Storage) システム130、およびクライアント端末140を含んで構成される。

【0015】

10

20

30

40

50

管理サーバ110とホスト計算機120とは、管理用NW(Network)150で接続される。管理サーバ110とSDSシステム130とは、管理用NW150で接続される。管理サーバ110とクライアント端末140とは、アクセス用NW160で接続される。ホスト計算機120とSDSシステム130とは、NW170(IP(Internet Protocol)ネットワークまたはSAN(Storage Area Network))で接続される。なお、ネットワークの構成は、上述の構成に限られるものではなく、適宜の構成を採用可能である。

【0016】

管理サーバ110は、データ管理プログラム111を実行し、SDSシステム130に格納(記憶)されているデータを管理する。

【0017】

ホスト計算機120は、一または複数設けられ、1つ以上のAP(Application Program)を含んで構成される業務AP121を提供し、SDSシステム130にデータを格納する。

【0018】

SDSシステム130は、コントロールノード131を備える。コントロールノード131は、制御ソフトウェアを実行し、SDSシステム130を構成するストレージ装置132を仮想化し、ホスト計算機120に対するインターフェースを提供する。SDSシステム130を構成するストレージ装置132として、業務AP121のデータを格納する1次ストレージ、当該データの保管データ(当該データのバックアップを取得したバックアップデータ、当該データのアーカイブを取得したアーカイブデータ、スナップショットにより取得したスナップショットデータ)を格納する2次ストレージがある。以降では、バックアップデータとアーカイブデータとスナップショットデータとを特に区別せず、バックアップデータの場合を例に挙げて説明するが、アーカイブデータ、スナップショットデータにも適用可能である。なお、ストレージ装置132は、複数の汎用サーバのディスクを仮想化したストレージであってもよいし、一般的なストレージサブシステムであってもよい。

10

20

【0019】

クライアント端末140は、データを利用するユーザが使用(操作)する端末であって、データ管理プログラム111にアクセス(例えば、ユーザがデータを要求するための利用要求を送信)する。

30

【0020】

図2は、データ管理プログラム111を実行する管理サーバ110のハードウェア構成の一例を示す図である。管理サーバ110は、1つ以上の、CPU201(中央処理装置)、メモリ202、ハードディスク等の2次記憶装置203、キーボード、マウスからの入力とディスプレイへの出力とを制御する入出力IF(interface)204、ネットワーク(管理用NW150、アクセス用NW160等)に接続するためのネットワークIF205を含んで構成される。

【0021】

管理サーバ110のメモリ202上には、データ管理プログラム111がロードされ、CPU201により実行される。なお、管理サーバ110の機能(例えば、図3に示すデータ管理プログラム111の機能モジュール)は、上述のようにソフトウェアにより実現されてもよいし、専用の回路などのハードウェアにより実現されてもよいし、ソフトウェアとハードウェアとが組み合わされて実現されてもよい。また、管理サーバ110の機能の一部は、管理サーバ110と通信可能な他のコンピュータにより実現されてもよい。

40

【0022】

また、2次記憶装置203には、データ管理プログラム111が使用するデータ管理DB(database)206が保存される。なお、管理サーバ110は、物理マシンでなく仮想マシンとして実装されてもよい。

【0023】

付言するならば、SDSシステム130として構成されるストレージ装置132のうち

50

、汎用サーバのハードウェア構成については、図2に示す構成と同様であり、その説明を省略する。

【0024】

図3は、データ管理プログラム111の機能モジュールに係る構成の一例を示す図である。データ管理プログラム111は、セルフサービスポータルIF部301、データ管理DB部302、要求受付部303、データ特定部304、データ情報収集部305、バックアップ情報収集部306、データ評価部307、および出力部308を含んで構成される。

【0025】

セルフサービスポータルIF部301は、クライアント端末140と通信を行い、データを利用するユーザとのインターフェースを提供する。データ管理DB部302は、データ管理DB206（データに関する属性、バックアップ情報、復元情報など）を管理する。要求受付部303は、例えば、要求対象の時間（例えば、後述の対象時間）を示す情報が含まれる要求（例えば、ユーザからのデータの利用要求）を受け付ける。

10

【0026】

データ特定部304は、例えば、データを保管するための保管データが取得された時間（日付、時刻など）を示す情報と、保管データが格納されるストレージ装置（保管場所、ストレージ識別子、ボリューム識別子など）を示す情報とが対応付けられているデータ管理情報（例えば、後述の図4参照）に基づいて、複数のストレージ装置132に記憶される保管データの中から、要求受付部303により受け付けられた要求に含まれる要求対象の時間に合う保管データを特定する。

20

【0027】

データ情報収集部305は、データに関する情報を業務APP121から収集したり、データの管理者からの入力により収集したりする。バックアップ情報収集部306は、データのバックアップ情報をSDSシステム130から収集したり、ホスト計算機120のバックアップ実行ツールから収集したりする。

【0028】

データ評価部307は、例えば、データ管理情報に基づいて、データ特定部304により特定された保管データについて、要求に含まれる要求対象の時間との時間差分を算出し、算出した時間差分が要求に含まれる要求レベルを満たすか否かを評価する。また、例えば、データ評価部307は、データ特定部304により特定された保管データの復元にかかる時間を算出し、算出した復元にかかる時間が要求に含まれる可能待ち時間以内であるか否かを評価する。出力部308は、データ特定部304による特定結果を出力する（特定した保管データを提供するための形式変換、転送等を行う）。例えば、出力部308は、データ特定部304により特定された保管データに係る情報を出力する。出力部308は、当該保管データの保管先を示す情報を出力するものであってもよいし、当該保管データを出力するものであってもよいし、当該保管データを識別する識別子および当該データを保管する保管先を識別する識別子を出力するものであってもよい。

30

【0029】

図4は、データ管理DB206に含まれるデータ管理情報の一例（データ管理テーブル400）を示す図である。データ管理テーブル400では、バックアップが取得されているデータについて、データ毎に、時間軸でバックアップ情報を管理する。データ管理テーブル400は、データ識別子フィールド401、日付フィールド402、識別子フィールド403、バックアップ取得単位フィールド404、保管場所フィールド405、サイズフィールド406、および取得時刻フィールド407を含んで構成される。

40

【0030】

データ識別子フィールド401は、データを識別するデータ識別子を示す情報を格納する。日付フィールド402は、バックアップデータの取得の日付（バックアップの日付）を示す情報を格納する。識別子フィールド403は、バックアップデータ（バックアップファイル）を識別する識別子を格納する。バックアップ取得単位フィールド404は、バ

50

ツクアップの取得単位を示す情報を格納する。保管場所フィールド405は、バックアップ先（保管先）を示すストレージ情報を格納する。サイズフィールド406は、バックアップデータのサイズを示す情報を格納する。取得時刻フィールド407は、バックアップデータの取得の時刻を示す情報を格納する。

【0031】

ここで、データ識別子は、DBのデータの場合は、DBのテーブル名またはカラム名、オブジェクトの場合は、オブジェクト識別子、ファイルの場合は、ファイル識別子など、APのデータを識別するデータの単位である。保管場所は、ストレージ識別子、ストレージ内のボリューム識別子、ディレクトリ等であり、バックアップデータの保管先を識別するためのものである。図4の例では、ストレージ識別子（ストレージ1）とボリューム識別子（LU1）とで保管場所を示している

10

。取得単位としては、フルバックアップ、差分バックアップ、増分バックアップなど、バックアップの方式を管理する。

【0032】

フルバックアップは、全てのデータを保持するバックアップの方式である。差分バックアップは、フルバックアップからの差分を保持するバックアップの方式である。増分バックアップは、前回のバックアップからの増分（変更分）を保持するバックアップの方式である。例えば、下記のようにデータが更新され、バックアップが取られた場合において、差分バックアップと増分バックアップとの違いを説明する。

時刻A：データWにデータXが追加

20

時刻B：フルバックアップ

時刻C：データYの追加

時刻D：差分バックアップ（または増分バックアップ）

時刻E：データZの追加

時刻F：差分バックアップ（または増分バックアップ）

【0033】

時刻Bのフルバックアップでは、全てのデータ（データWおよびデータX）がバックアップデータとして保存される。

【0034】

時刻Dの差分バックアップでは、前回のフルバックアップのバックアップデータ（データWおよびデータX）との差分のデータ（データY）がバックアップデータとして保存される。時刻Dの増分バックアップでは、前回のバックアップ（時刻B）からの変更分のデータ（データY）がバックアップデータとして保存される。

30

【0035】

時刻Fの差分バックアップでは、前回のフルバックアップのバックアップデータ（データWおよびデータX）との差分のデータ（データYおよびデータZ）がバックアップデータとして保存される。時刻Fの増分バックアップでは、前回のバックアップ（時刻D）からの変更分のデータ（データZ）がバックアップデータとして保存される。

【0036】

なお、本実施の形態では、バックアップの方式として、フルバックアップ、差分バックアップ、および増分バックアップを例に挙げて説明するが、マルチレベル増分バックアップ、逆増分バックアップなど、他のバックアップの方式を採用することもできる。

40

【0037】

バックアップ情報については、データ管理プログラム111が、バックアップ実行ツールから取得して登録してもよいし、SDSシステム130から2次ストレージに保管されたデータの情報を取得して登録してもよい。データ識別子については、データ管理プログラム111が、業務AP121から取得してもよいし、バックアップ情報の取得時にバックアップデータを復元し、データの内容を読み出して登録してもよい。登録したデータ識別子については、バックアップデータの復元、読み出しをバックアップ情報の取得の都度、または定期的に実行することで、追加があれば更新する。

50

【 0 0 3 8 】

また、図4には示さないが、例えば、図10に示すように、データ管理テーブル400では、データ内にタイムスタンプがある場合には、データ内の最初のタイムスタンプである初回タイムスタンプとデータ内の最後のタイムスタンプである最終タイムスタンプとを追加で管理する。タイムスタンプの情報については、例えば、データ管理プログラム111がバックアップデータを一旦復元し、データの内容を読み出す際に併せて登録する。

【 0 0 3 9 】

図5は、AP管理情報の一例(AP管理テーブル500)を示す図である。AP管理テーブル500は、データ管理DB206に含まれる。AP管理テーブル500は、バックアップデータが取得されたAPの情報を管理するテーブルである。AP管理テーブル500は、AP識別子フィールド501、APタイプフィールド502、およびバックアップ手段フィールド503を含んで構成される。

10

【 0 0 4 0 】

AP識別子フィールド501は、APを識別するAP識別子を示す情報を格納する。APタイプフィールド502は、APのタイプ(例えば、ソフトウェア名)を示す情報を格納する。バックアップ手段フィールド503は、バックアップの実行手段を示す情報(データの提供時にデータを復元する際に用いる情報であって、どのようにバックアップが行われたかを識別可能な情報)を格納する。図5では、APをDBシステムとして、AP識別子には、DB名、APタイプには、DB管理システムの名称を登録する例を示す。AP識別子、APタイプ、およびバックアップ手段の情報は、例えば、管理者により登録される。

20

【 0 0 4 1 】

図6は、データ属性管理情報の一例(データ属性管理テーブル600)を示す図である。データ属性管理テーブル600は、データ管理DB206に含まれる。データ属性管理テーブル600は、データの属性を管理するテーブルである。データ属性管理テーブル600は、データ識別子フィールド601、タイプフィールド602、およびタイムスタンプカラムフィールド603を含んで構成される。

【 0 0 4 2 】

データ識別子フィールド601は、データを識別するデータ識別子を示す情報を格納する。タイプフィールド602は、データのタイプを示す情報を格納する。タイムスタンプカラムフィールド603は、データ内のタイムスタンプを示す情報を格納する。データのタイプとしては、データの追加、更新、削除に対する仕様が追記型であるのか更新型であるのかが管理される。

30

【 0 0 4 3 】

図7は、データ管理プログラム111がデータの利用要求に対するデータを提供するデータ提供処理に係るフローチャートの一例を示す図である。

【 0 0 4 4 】

ステップS701では、データ管理プログラム111は、クライアント端末140からデータの利用要求を受け付ける。利用要求には、ユーザから要求されたデータのデータ名(データ識別子)、データを取得する時間範囲(対象時間)、データの精度に関する要求レベル、データの提供まで待つことを許容する時間(可能待ち時間)などの情報が含まれる。付言するならば、利用要求には、上述した全ての情報が含まれるものではない。例えば、要求対象のデータの種別が1種類である場合は、データ識別子が含まれていなくてもよい。

40

【 0 0 4 5 】

なお、ユーザインターフェースの一例(画面)については、図15を用いて後述する。要求レベルは、要求された対象時間に対して提供可能なデータの時間との差分に対する精度を要求するものであり、要求レベルとしては、例えば、2段階(低、高)から選択される。可能待ち時間としては、ユーザがデータの提供を希望する時間(例えば、1時間以内)が指定される。

50

【0046】

ステップS702では、データ管理プログラム111は、ユーザから要求されたデータ名（データ識別子）でデータ管理テーブル400を検索する。

【0047】

ステップS703では、データ管理プログラム111は、要求されたデータ名（データ識別子）のデータが存在するか否か（一致するデータ名があるか否か）を判定し、存在すると判定した場合、ステップS704に処理を移し、存在しないと判定した場合、ステップS711に処理を移す。

【0048】

ステップS704では、データ管理プログラム111は、要求されたデータの条件に合うデータの候補を抽出するために、要求されるデータの対象時間が日時指定か期間指定か（日時指定であるか否か）を判定し、日時指定であると判定した場合、ステップS705に処理を移し、期間指定であると判定した場合、ステップS706に処理を移す。 10

【0049】

ステップS705では、データ管理プログラム111は、日時指定時のデータ特定処理を実行し、ステップS707に処理を移す。なお、日時指定時のデータ特定処理の詳細については、図8を用いて後述する。

【0050】

ステップS706では、データ管理プログラム111は、期間指定時のデータ特定処理を実行し、ステップS707に処理を移す。なお、期間指定時のデータ特定処理の詳細については、図9を用いて後述する。 20

【0051】

ステップS707では、データ管理プログラム111は、利用要求で要求レベルが指定されている場合、要求レベルを満足するデータを提供できるかを確認する処理（要求レベル確認処理）を行う。なお、要求レベル確認処理の詳細については、図12を用いて後述する。

【0052】

ステップS708では、データ管理プログラム111は、利用要求で可能待ち時間が指定されている場合、特定したデータの候補について、復元に要する時間を算出し、要求された可能待ち時間を満足するかを確認する処理（復元時間確認処理）を行う。なお、復元時間確認処理の詳細については、図13を用いて後述する。 30

【0053】

ステップS709では、データ管理プログラム111は、提供形態決定処理を行う。なお、提供形態決定処理では、提供するデータの提供形態が決定されるが、詳細については、図14を用いて後述する。

【0054】

ステップS710では、データ管理プログラム111は、データを復元して提供形態のデータを作成する。データの復元については、例えば、データ管理プログラム111は、A P 管理テーブル500に基づいて、バックアップ手段（例えば、D B のインスタンス）を特定し、特定したバックアップ手段を用いてデータを復元する。 40

【0055】

ステップS711では、データ管理プログラム111は、利用要求に対する結果を出力し、データ提供処理を終了する。例えば、データ管理プログラム111は、結果として、提供形態のデータへのアクセス方法を出力する。なお、出力の例については、図15を用いて後述する。

【0056】

ここで、要求されたデータの時間に一致するデータが最新のバックアップデータ以降である場合、バックアップデータには要求に合うデータがないことになる。その場合、1次ストレージのデータボリュームのスナップショットを取得して保存してもよい。取得したデータについては、バックアップデータと同様に取得時刻、保存先などを管理してもよい。 50

【 0 0 5 7 】

図 8 は、日時指定時のデータ特定処理（ステップ S 7 0 5）に係るフローチャートの一例を示す図である。

【 0 0 5 8 】

ステップ S 8 0 1 では、データ管理プログラム 1 1 1 は、指定日時（指定された日時）に取得日時（日付および取得時刻）が一致するバックアップデータがあるかをデータ管理テーブル 4 0 0 から検索する。

【 0 0 5 9 】

ステップ S 8 0 2 では、データ管理プログラム 1 1 1 は、検索の結果、一致するものがあるか否かを判定し、あると判定した場合、ステップ S 8 0 3 に処理を移し、ないと判定した場合、ステップ S 8 0 4 に処理を移す。
10

【 0 0 6 0 】

ステップ S 8 0 3 では、データ管理プログラム 1 1 1 は、該当のバックアップデータを特定し、ステップ S 8 0 5 に処理を移す。

【 0 0 6 1 】

ステップ S 8 0 4 では、データ管理プログラム 1 1 1 は、バックアップの取得日時が指定日時に最も近いバックアップのバックアップデータを特定し、ステップ S 8 0 5 に処理を移す。

【 0 0 6 2 】

ステップ S 8 0 5 では、データ管理プログラム 1 1 1 は、選択したバックアップデータ（該データ）が差分データ（差分バックアップで取得されたバックアップデータ）であるか否かを判定し、差分データであると判定した場合、ステップ S 8 0 6 に処理を移し、差分データでないと判定した場合、ステップ S 8 0 7 に処理を移す。
20

【 0 0 6 3 】

ステップ S 8 0 6 では、データ管理プログラム 1 1 1 は、元となるフルバックアップデータを特定し、日時指定時のデータ特定処理を終了する。

【 0 0 6 4 】

ステップ S 8 0 7 では、データ管理プログラム 1 1 1 は、選択したバックアップデータ（該データ）が増分データ（増分バックアップで取得されたバックアップデータ）であるか否かを判定し、増分データであると判定した場合、ステップ S 8 0 8 に処理を移し、増分データでないと判定した場合、該データがフルバックアップのバックアップデータであるので、日時指定時のデータ特定処理を終了する。
30

【 0 0 6 5 】

ステップ S 8 0 8 では、データ管理プログラム 1 1 1 は、フルバックアップデータと、フルバックアップデータから該データまでの増分データとを特定し、日時指定時のデータ特定処理を終了する。

【 0 0 6 6 】

このように、データ特定部 3 0 4 は、例えば、データ管理情報に基づいて、要求に含まれるデータ識別子に対応するデータの保管データがある（例えば、ステップ S 7 0 3 で「Y e s」）と判定した場合、要求に含まれる要求対象の時間と一致する保管データが存在するか否かを判定し（例えば、ステップ S 8 0 2）、存在すると判定したときは、保管データを特定し（例えば、ステップ S 8 0 3）、存在しないと判定したときは、要求に含まれる要求対象の時間に最も近い日時の保管データを特定する（例えば、ステップ S 8 0 4）。
40

【 0 0 6 7 】

図 9 は、期間指定時のデータ特定処理（ステップ S 7 0 6）に係るフローチャートの一例を示す図である。期間指定時のデータ特定処理については、適時、図 1 0 を用いて説明する。

【 0 0 6 8 】

図 1 0 は、期間指定時のバックアップデータの選択の概念図である。図 1 0 (A) は、バックアップの取得時間に対して、要求された対象期間のバックアップデータを取り出す

10

20

30

40

50

場合の例を時間軸で表す図である。図10(B)は、(A)のバックアップの特定の例に併せて、データ管理テーブル400の一部として管理し得る初回タイムスタンプ408と最終タイムスタンプ409との例を示す。

【0069】

ステップS901では、データ管理プログラム111は、要求されたデータ(対象データ)が追記型であるか否かを判定し、追記型であると判定した場合、ステップS902に処理を移し、追記型でないと判定した場合、ステップS908に処理を移す。

【0070】

ステップS902では、データ管理プログラム111は、対象データにタイムスタンプカラムがあるか否かを判定し、あると判定した場合、ステップS903に処理を移し、ないと判定した場合、ステップS908に処理を移す。10

【0071】

ステップS903では、データ管理プログラム111は、データ管理テーブル400の初回タイムスタンプが要求された開始日時以前、かつ、最新のバックアップデータ(ここではバックアップデータAとする)を特定する。図10の例では、対象期間が4/25~5/10であるので、初回タイムスタンプが開始日時4/24以前で最新の4/20であるレコード1001のバックアップデータが特定される。

【0072】

ステップS904では、データ管理プログラム111は、特定したバックアップデータについて、最終タイムスタンプが要求された終了日時以降であるか否かを判定し、以降であると判定した場合、ステップS906に処理を移し、以降でないと判定した場合、ステップS905に処理を移す。図10の例では、レコード1001の最終タイムスタンプ(2017/4/29)は、要求された終了日時(2017/5/10)以降ではないと判定される。20

【0073】

ステップS905では、データ管理プログラム111は、該バックアップデータ以降にバックアップが取得されたバックアップデータについて、最終タイムスタンプが要求された終了日時以降になるまで、確認を繰り返す。

【0074】

ステップS906では、データ管理プログラム111は、最終タイムスタンプが終了日時以降のバックアップデータ(ここではバックアップデータB)を特定する。図10の例では、レコード1003が、要求された終了日時(2017/5/10)以降で直近の終了タイムスタンプ(2017/5/13)を有するバックアップデータの情報である。30

【0075】

ステップS907では、データ管理プログラム111は、最初に選択したバックアップデータAから終了日時が含まれる該バックアップデータBまでの全バックアップデータを対象として特定し、ステップS911に処理を移す。図10の例では、レコード1001とレコード1003との間にあるレコード1002のバックアップデータが特定される。なお、データ管理プログラム111は、最初に特定したバックアップデータAの最終タイムスタンプが終了日時以降の場合、該バックアップデータAのみを特定する。40

【0076】

ステップS908では(データが更新型である場合、または、データが追記型であり、タイムスタンプカラムがない場合)、データ管理プログラム111は、バックアップの取得日時が、要求された開始日時以前で、かつ開始日時に最も近い日時のバックアップデータ(バックアップデータC)を特定する。

【0077】

ステップS909では、データ管理プログラム111は、バックアップの取得日時が、要求された終了日時以降で、かつ終了日時に最も近い日時のバックアップデータ(バックアップデータD)を特定する。

【0078】

10

20

30

40

50

ステップ S 910 では、データ管理プログラム 111 は、特定したバックアップデータ C からバックアップデータ D までの全バックアップデータを対象として特定する。

【0079】

ステップ S 911 では、データ管理プログラム 111 は、特定したバックアップデータ（特定データ）に差分データがあるか否かを判定し、あると判定した場合、ステップ S 912 に処理を移し、ないと判定した場合、ステップ S 913 に処理を移す。

【0080】

ステップ S 912 では、データ管理プログラム 111 は、元となるフルバックデータを特定し、期間指定時のデータ特定処理を終了する。

【0081】

ステップ S 913 では、データ管理プログラム 111 は、特定したバックアップデータ（特定データ）に増分データがあるか否かを判定し、あると判定した場合、ステップ S 914 に処理を移し、ないと判定した場合、期間指定時のデータ特定処理を終了する。

【0082】

ステップ S 914 では、データ管理プログラム 111 は、フルバックアップデータと、フルバックアップデータから該データまでの増分データとを特定し、期間指定時のデータ特定処理を終了する。

【0083】

このように、データ特定部 304 は、例えば、データ管理情報に基づいて、要求に含まれるデータ識別子に対応するデータの保管データがある（例えば、ステップ S 703 で「Yes」）と判定した場合、データの保管データにタイムスタンプの情報があるか否かを判定し（例えば、ステップ S 902）、タイムスタンプの情報があると判定したときは、データの保管データの初回タイムスタンプが要求に含まれる開始日時以前かつ最新である第 1 の保管データを特定し（例えば、ステップ S 903）、データの保管データの最終タイムスタンプが前記要求に含まれる終了日時以降となる第 2 の保管データを特定し（例えば、ステップ S 906）、第 1 の保管データから第 2 の保管データまでの全ての保管データを特定する（例えば、ステップ S 907）。

【0084】

また、例えば、データ特定部 304 は、データ管理情報に基づいて、要求に含まれる要求対象の時間に合うバックアップデータが存在するか否かを判定し（例えば、ステップ S 703、ステップ S 801～ステップ S 804、ステップ S 901～ステップ S 910）、存在すると判定した場合、バックアップデータの取得単位が差分バックアップであるときは、バックアップデータの直近のフルバックアップのフルバックアップデータを特定し（例えば、ステップ S 806、ステップ S 912）、バックアップデータの取得単位が増分バックアップであるときは、バックアップデータの直近のフルバックアップのフルバックアップデータと、フルバックアップデータからバックアップデータまでの増分バックアップデータと、を特定する（例えば、ステップ S 808、ステップ S 914）。

【0085】

図 11 は、ストレージ装置管理情報の一例（ストレージ装置管理テーブル 1100）を示す図である。ストレージ装置管理テーブル 1100 は、データ管理 DB206 に含まれる。ストレージ装置管理テーブル 1100 は、ストレージ識別子フィールド 1101、タイプフィールド 1102、プロトコルフィールド 1103、ディスク種別フィールド 1104、および性能フィールド 1105 を含んで構成される。

【0086】

ストレージ識別子フィールド 1101 は、ストレージ装置 132 を識別するストレージ識別子を示す情報を格納する。タイプフィールド 1102 は、アクセスタイプを示す情報を格納する。プロトコルフィールド 1103 は、対応するプロトコルを示す情報を格納する。ディスク種別フィールド 1104 は、構成するディスクの種別を示す情報を格納する。性能フィールド 1105 は、ディスクの性能（例えば、転送速度）を示す情報を格納する。これらは、管理者により設定されてもよいし、ストレージ管理ツールから取得して登

10

20

30

40

50

録されてもよい。

【0087】

さらに、ボリューム識別子フィールド1106を設け、ボリュームの作成時に、対応するディスクの種別に併せてボリューム識別子を示す情報が登録されてもよい。

【0088】

図12は、要求レベルを満足するデータを提供できるかを確認する要求レベル確認処理(ステップS707)に係るフローチャートの一例を示す図である。

【0089】

ステップS1201では、データ管理プログラム111は、特定したバックアップデータについて、対象時間との時間差分を算出する。例えば、データ管理プログラム111は、要求されたデータの対象時間が日時指定である場合、特定したバックアップデータの取得日時と指定された日時との差分を算出する。また、例えば、データ管理プログラム111は、対象時間が期間指定である場合、指定された開始日時と特定したバックアップデータの最初の日時(タイムスタンプカラムがない場合はバックアップの取得日時)との差分と、指定された終了日時と特定したバックアップデータの最後の日時との差分とを算出し、差分の合計とする。

10

【0090】

ステップS1202では、データ管理プログラム111は、利用要求(入力された利用条件)から要求レベルを取得する。

【0091】

ステップS1203では、データ管理プログラム111は、時間差分が要求レベルを満たすか否かを判定し、満たすと判定した場合、ステップS1204に処理を移し、満たさないと判定した場合、ステップS1205に処理を移す。例えば、要求レベルは、低、高の2段階で設定されるものとし、判定のためのしきい値をレベル別に設ける。要求レベルが「低」の場合は、時間差分が「高」の場合より大きい第1のしきい値となり、「高」の場合は、時間差分が「低」の場合より小さい第2のしきい値となるように設定される。設定については、データ管理プログラム111の情報として予め定義されてもよいし、管理者により定義されてもよいし、ユーザにより定義されてもよい。この場合、データ管理プログラム111は、要求レベルを満たすか否かの判定は、時間差分が、要求レベルに対応するしきい値以内であるか否かを判定する。

20

【0092】

ステップS1204では、データ管理プログラム111は、要求レベルの判定結果を「満足」とし、要求レベル確認処理を終了する。

【0093】

ステップS1205では、データ管理プログラム111は、要求レベルの判定結果を「不満足」とし、要求レベル確認処理を終了する。

【0094】

ここで要求レベルを満たさない場合は、追加の処理として、データ管理プログラム111は、データの更新記録(例えば、ジャーナルファイル)からの復元を実施してもよい。追加の処理として、データの過去の更新記録のバックアップを利用する方法がある。データ管理プログラム111は、データの過去の更新記録のバックアップが保存されていて利用可能であるか否かを判定し、利用不可である場合、要求レベルの判定結果を「不満足」とする。データ管理プログラム111は、データの更新記録が利用可能である場合、特定したバックアップデータの日時と対象日時(期間指定の場合は終了日時)との差分となっている時間帯を指定して更新記録があるかを検索する。検索の結果、差分の時間帯の更新記録があった場合は、更新記録を対象データとして、復元するように追加する。その場合、データ管理プログラム111は、適用する更新記録の最新日時と対象日時との差分を再算出し、要求レベルに対する判定を実施する。

40

【0095】

図13は、要求された可能待ち時間を満足するかを確認する復元時間確認処理(ステッ

50

PS 708)に係るフローチャートの一例を示す図である。

【0096】

ステップS1301では、データ管理プログラム111は、特定したバックアップデータについて復元時間の算出を行う。算出処理では、例えば、データ管理プログラム111は、データ特定処理で対象と特定したバックアップデータのうち、フルバックアップデータについて、データ管理テーブル400で管理するバックアップ情報から保管先とサイズとを読み出す。データ管理プログラム111は、ストレージ装置132の構成情報(例えば、ボリューム識別子フィールド1106がある場合は、ストレージ装置管理テーブル1100)からボリュームの使用するディスクを判別し、ストレージ装置管理テーブル1100からバックアップデータの保管されているストレージ装置132の仕様のうち、ディスクの性能の情報を読み出す。データ管理プログラム111は、該当のサイズのデータを該当のディスクから読み出す時間(読出時間)を算出する(データサイズ/IOPS(Input/Output Per Second))。例えば、データ管理プログラム111は、当該復元データが格納されているストレージ装置132から管理サーバ110へのバックアップデータの転送時間を算出して読出時間に加算して復元時間とする。また、例えば、データ管理プログラム111は、復元データを所定のサーバに設ける場合、当該復元データが格納されているストレージ装置132から管理サーバ110へのバックアップデータの転送時間と、管理サーバ110から所定のサーバへの復元データの転送時間を算出して読出時間に加算して復元時間とする。

10

【0097】

なお、算出には、ストレージ装置132のプロセッサ使用率を重み付けしてもよい。また、特定されたバックアップデータが差分バックアップまたは増分バックアップについても同様に算出する。

20

【0098】

ここで、復元時間の算出については、過去の履歴を使用する方法でもよい。バックアップデータを復元して提供する場合に、復元にかかった実行時間を計測し、履歴として管理(2次記憶装置203に記憶)しておく。対象として特定されたバックアップデータが履歴にある場合、計測された実行時間を読み出して、復元時間とする。複数回、復元が実行されている場合は、全ての実行時間を読み出して代表値(例えば、平均値)を算出してもよい。

30

【0099】

ステップS1302では、データ管理プログラム111は、利用要求(入力されたデータの利用条件)から可能待ち時間を取得する。

【0100】

ステップS1303では、データ管理プログラム111は、算出した復元時間が可能待ち時間以内であるか否かを判定し、以内であると判定した場合、ステップS1304に処理を移し、以内でないと判定した場合、ステップS1305に処理を移す。

【0101】

ステップS1304では、データ管理プログラム111は、復元時間の判定結果を「満足」とし、復元時間確認処理を終了する。

40

【0102】

ステップS1305では、データ管理プログラム111は、復元時間の判定結果を「不満足」とし、復元時間確認処理を終了する。

【0103】

ここで、要求レベルを満たさない場合は、追加の処理として、データ管理プログラム111は、別のバックアップデータを代替として使用可能であるか否かを判定する処理を実施してもよい。追加処理では、データ管理プログラム111は、最初に対象として特定したバックアップデータに差分バックアップデータまたは増分バックアップデータが含まれているか否かを判定する。データ管理プログラム111は、差分バックアップデータまたは増分バックアップデータが含まれていると判定した場合、フルバックアップデータに対

50

する時間差分および復元時間を算出し直す。データ管理プログラム 111 は、算出し直した復元時間が要件レベルを満たすか否かを判定し、満たすと判定した場合、差分バックアップデータまたは増分バックアップデータを対象データから除き、判定結果を「満足」とする。

【0104】

図 14 は、データの提供の形態（提供形態）を決定する提供形態決定処理（ステップ S709）に係るフロー・チャートの一例を示す図である。提供形態決定処理では、提供データのサイズ（容量）とネットワーク性能とに基づいて提供形態が決定される。

【0105】

ステップ S1401 では、データ管理プログラム 111 は、提供するバックアップデータのサイズから復元した場合のデータサイズ（復元データのサイズ）を見積もる。

10

【0106】

ステップ S1402 では、データ管理プログラム 111 は、見積もったデータサイズとネットワークの性能とから転送時間（例えば、所定のサーバからクライアント端末 140 への転送時間）を見積もる。

【0107】

ステップ S1403 では、データ管理プログラム 111 は、転送時間の見積もりがしきい値以下であるか否かを判定し、しきい値以下であると判定した場合、ステップ S1404 に処理を移し、しきい値を超えると判定した場合、ステップ S1405 に処理を移す。

【0108】

ステップ S1404 では、データ管理プログラム 111 は、ファイルでの提供形態（ファイル形式で提供）と決定し、提供形態決定処理を終了する。この場合、データ管理プログラム 111 は、データをファイルに書き出し、所定のサーバ（ファイルを利用者がアクセス可能なサーバ）に転送することになる。

20

【0109】

ステップ S1405 では、データ管理プログラム 111 は、仮想マシンでの提供形態（仮想マシン形式で提供）と決定し、提供形態決定処理を終了する。この場合、データ管理プログラム 111 は、仮想マシンに復元したデータを含む A P 環境を構築し、仮想マシンの A P にアクセスしてデータを利用可能にし、結果の出力において、仮想マシンへのアクセス情報を出力することになる。

30

【0110】

また、データ管理プログラム 111 は、提供データのセキュリティポリシーに従って提供形態を決定するようにしてもよい。セキュリティポリシーは、管理者により設定される。セキュリティポリシーとして、アクセスコントロールを行う場合、提供データを管理者側環境に配置して制御する形態、仮想マシン形態で A P に対するアクセスを制御する形態等を管理者が設定可能とする。

【0111】

また、利用者が提供形態を指定するインターフェースが設けられてもよく、その場合、データ管理プログラム 111 は、指定された形態でデータを提供する。

【0112】

図 15 は、出入力インターフェースの一例（入出力画面 1500）を示す図である。入出力画面 1500 の上段エリア 1501 は、利用者が利用要求（利用条件）を入力するエリアであり、下段エリア 1502 は、利用要求に対する結果を出力するエリアである。

40

【0113】

入力情報として、データ名 1503、要求するデータの対象時間（期間指定 1504 または日時指定 1505）、要求レベル 1506、可能待ち時間 1507 がある。要求するデータの対象時間としては、ユーザは、期間指定 1504 または日時指定 1505 を選択して設定する。ここで、ユーザが日時指定 1505 を設定するケースとしては、例えば、当該時点のデータを再現してテストを行うケースが挙げられる。また、ユーザが期間指定 1504 を設定するケースとしては、例えば、期間内に更新されたデータを用いた分析を

50

行うケースが挙げられる。

【0114】

出力情報として、提供データ名1508、提供形態1509、アクセス方法を示すアクセス先1510、要件を満たすか否かを示す要件満足レベル1511がある。要件満足レベルの付加情報として、時間の差分と提供までの時間とが出力される。

【0115】

要件満足レベルは、対象期間に対する提供データの時間差分から要求レベルを満たすか否かの判定を行う処理(図12)と、算出した復元時間が可能待ち時間を満たすか否かの判定を行う処理(図13)の結果をもとに決定される。ここでの要件満足レベルは、例えば、3つのレベルを設け、両方の判定結果が満足である場合、レベル「高」、どちらか一方の判定結果が満足である場合、レベル「中」、両方の判定結果とも不満足である場合、レベル「低」とする。

10

【0116】

なお、利用要求に対して、データを復元する前に特定したバックアップデータについて要件満足レベル(時間の差分、提供までの時間)を出力し、ユーザが確認後に復元する構成としてもよい。

【0117】

以上、上述した構成によれば、利用者からの要求に対して、保存するバックアップデータから最適なバックアップデータを特定し、利用者にデータを提供することを可能とする。例えば、データの利用要求で指定される時間断面が、バックアップの取得日時とは異なる場合でも、既存のバックアップデータの中から利用要求に合うデータを抽出することができる。また、例えば、利用要求を満たすかを示す補助情報として、指定日時との差分情報(例えば、時間の差分)と、準備目安時間(例えば、提供までの時間)とを附加して、結果を出力することを可能とする。かかる時間の差分によれば、ユーザは、提供されるデータの精度を把握できるようになる。かかる提供までの時間によれば、ユーザは、データが提供されるまでの時間を把握できるようになる。

20

【0118】

本実施の形態によれば、様々な種類の記憶装置に様々なタイミングで保存された保管データについて、時間情報、保管場所情報などを管理することで、利用単位にあった適切な保管データを適切な単位で提供することを可能にする。また、業務APが生成するデータへの利用要求に対して、別のストレージ装置に保管された保管データを使用して、利用要件にあった保管データを特定して提供することで、業務APの稼動への影響を回避できるようになる。

30

【0119】

(2) 第2の実施の形態

本実施の形態では、データを要求する単位に対して、複数のデータを含むケースについて説明する。本実施の形態では、第1の実施の形態と異なる構成について主に説明する。

【0120】

図16は、第1の実施の形態の管理情報に追加して本実施の形態で管理する業務データ管理情報の一例(業務データ管理テーブル1600)を示す。業務データ管理テーブル1600は、データ管理DB206に含まれる。業務データ管理テーブル1600は、業務識別子フィールド1601、データ項目名フィールド1602、APフィールド1603、およびAP内データ識別子フィールド1604を含んで構成される。

40

【0121】

業務識別子フィールド1601は、業務を識別可能な業務識別子を示す情報を格納する。データ項目名フィールド1602は、データの項目名を示す情報を格納する。APフィールド1603は、APとそのタイプを識別する情報を格納する。AP内データ識別子フィールド1604は、AP内でのデータを識別する識別子(業務に係るデータを識別するデータ識別子)を示す情報を格納する。

【0122】

50

本実施の形態では、利用者は、利用条件の入力時に業務名（業務識別子）を指定する。データ管理プログラム111は、業務データ管理テーブル1600を検索して、一致する業務識別子を特定する。データ管理プログラム111は、業務識別子に対応するデータ項目名と、APと、AP内データ識別子とを特定する。データ管理プログラム111は、特定したAP内データ識別子をデータ名（データ識別子）として、データ管理テーブル400を検索し（ステップS702）、以降の処理を同様に実施する。なお、時間要件の指定については第1の実施の形態と同様である。

【0123】

このように、データ特定部304は、業務データ管理情報に基づいて、要求に含まれる業務識別子に対応するデータ識別子を特定し、データ管理情報に基づいて、特定したデータ識別子に対応するデータの保管データがあるか否かを判定する。

10

【0124】

以上により、データを要求する単位を業務として、複数のデータを含む場合においても、利用者が業務名を指定することで、複数のデータについて、それぞれ必要な時間のデータを提供することが可能になる。

【0125】

次に、利用単位に対する複数のデータのバックアップが仕様の異なるストレージ装置132に保管されている場合の処理を説明する。

【0126】

図16の例では、同じ業務識別子に対応するデータが異なるAPであり、バックアップデータは異なるストレージ装置132に保管されているものとする。ここで、第1の実施の形態の図11に示したストレージ装置管理テーブル1100に登録されているストレージ装置132の仕様を参照する。例えば、顧客DBのバックアップデータは、ストレージAに保管され、在庫DBのバックアップデータは、ストレージBに保管されているものとする。この場合、SDSシステム130では、バックアップデータを取り出すとき、異なるプロトコルに対して、プロトコル変換を行い、共通のプロトコル（同じ形式）で取り出す。例えば、ロックストレージに保管された顧客DBのバックアップデータをファイルとして取り出し、ファイルストレージに保管された在庫DBのバックアップデータは、そのままファイルとして取り出す。共通のプロトコルで取り出すことで、業務単位で必要となる複数のデータをまとめて提供することができる。

20

【0127】

(3) 第3の実施の形態

本実施の形態では、バックアップデータのストレージ装置132への配置を制御する方法について説明する。本実施の形態では、上述した実施の形態と異なる構成について主に説明する。

【0128】

SDSシステム130では、多種のストレージ装置132を仮想化して管理し、ストレージ装置132のディスクの性能別に階層を設け、階層間のデータ移行を行う階層制御機能（階層制御部）を提供する。管理サーバ110（データ管理プログラム111の図示は省略するデータ移行部）は、バックアップデータの使用履歴に基づいて、バックアップデータの保管先に対するデータ移行をSDSシステム130の階層制御機能に要求する。

40

【0129】

図17は、ストレージ階層管理情報の一例（ストレージ階層管理テーブル1700）を示す図である。ストレージ階層管理テーブル1700は、データ管理DB206に含まれる。ストレージ階層管理テーブル1700は、レベルフィールド1701、ストレージ識別子フィールド1702、ディスクフィールド1703、性能フィールド1704、および空き容量フィールド1705を含んで構成される。

【0130】

レベルフィールド1701は、階層のレベルを示す情報を格納する。ストレージ識別子フィールド1702は、管理対象のストレージ装置132を識別するストレージ識別子を

50

示す情報を格納する。ディスクフィールド 1703 は、ディスクの種別を示す情報を格納する。性能フィールド 1704 は、ディスクの性能（例えば、転送速度（MB/s））を示す情報を格納する。空き容量フィールド 1705 は、ディスクの空き容量を示す情報を格納する。

【0131】

レベルに対するストレージ装置 132、ディスクの割当については、管理者またはシステムにより事前に設定される。性能、空き容量については、管理サーバ 110 が定期的に SDS システム 130 または各ストレージ装置 132 から情報を取得して登録し、例えば、平均転送速度、取得時点での空き容量が管理される。階層は、例えば、性能のよい順に高、中、低の 3 階層とし、それぞれレベル 1、2、3 としてレベルフィールド 1701 で管理される。

10

【0132】

管理サーバ 110（データ管理プログラム 111 の図示は省略する記憶部）は、バックアップデータを使用したデータの提供時に復元までの実行時間を計測し、履歴として管理（2 次記憶装置に記憶）する。本実施の形態では、管理サーバ 110 は、バックアップデータを使用する際の利用条件である可能待ち時間についても履歴として記録（記憶）する。管理サーバ 110 は、バックアップデータ毎に提供のために使用した履歴を管理する。管理サーバ 110 は、期間要求に併せて提供データのサイズが異なる場合であっても、バックアップからの復元時間はバックアップデータ全体に対して行うため、その際にかかった時間を記録する。なお、記録については、本実施の形態では、使用履歴管理情報として登録する。

20

【0133】

図 18 は、使用履歴管理情報の一例（使用履歴管理テーブル 1800）を示す図である。使用履歴管理テーブル 1800 は、データ管理 DB206 に含まれる。使用履歴管理テーブル 1800 は、バックアップ識別子フィールド 1801、使用日時フィールド 1802、復元時間フィールド 1803、可能待ち時間フィールド 1804 を含んで構成される。なお、記録は、テーブルへの出力（登録）に限られるものではなく、ログファイル等への出力であってもよい。

30

【0134】

バックアップ識別子フィールド 1801 は、バックアップデータを識別するバックアップ識別子（保管識別子）を示す情報を格納する。使用日時フィールド 1802 は、復元に使用した日時を示す情報を格納する。復元時間フィールド 1803 は、バックアップデータの復元に要した時間を示す情報を格納する。可能待ち時間フィールド 1804 は、要求時に指定された可能待ち時間を示す情報を格納する。

40

【0135】

管理サーバ 110 では、バックアップデータ毎の使用頻度が予め設定した一定レベル（例えば、1 回 / 週）以上になった場合、要求された可能待ち時間の平均と復元時間の平均とを算出し、可能待ち時間に対する復元時間の時間差分を算出する。管理サーバ 110 は、差分が予め設定した上限基準値以上（復元時間が可能待ち時間より長い）の場合、該バックアップデータが保管される階層を、現在の階層より高い階層がある場合、ストレージ階層管理テーブル 1700 を参照し、空き容量を確認する。管理サーバ 110 は、バックアップデータのサイズ（データ管理テーブル 400 参照）以上の空き容量がある場合、該階層にデータを移行するように SDS システム 130 に指示を出す。

40

【0136】

管理サーバ 110 は、算出した差分が予め設定した下限基準値以下（復元時間が可能待ち時間より短い）の場合、該バックアップデータが保管される階層を、現在の階層より低い階層がある場合、ストレージ階層管理テーブル 1700 を参照し、当該階層のストレージ装置 132（ディスク）の空き容量を確認する。また、管理サーバ 110 は、当該階層のストレージ装置 132 の性能（転送速度）を参照して、バックアップデータのサイズ（データ管理テーブル 400 参照）から、復元した場合の復元時間を算出し、平均可能待ち

50

時間内に収まるかを確認する。管理サーバ110は、空き容量があり、復元時間が収まる場合、該階層のストレージ装置132にデータを移行する。

【0137】

このように、データ移行部は、例えば、ストレージ装置管理情報と、記憶部に記憶されている情報（使用履歴管理情報）とに基づいて、復元に要した時間が可能待ち時間よりも長い保管データを格納するディスクよりも性能の高いディスクに保管データを移行する処理、および／または、復元に要した時間が可能待ち時間よりも短い保管データを格納するディスクよりも性能の低いディスクに保管データを移行する処理を行う。

【0138】

本実施の形態では、利用者の要求を満たすように保管データの配置を適切に実行し、ストレージ装置132の効率利用を可能とする。

10

【0139】

(4) 第4の実施の形態

本実施の形態では、バックアップデータの利用履歴に基づいて、バックアップの取得を制御する方法について説明する。本実施の形態では、上述した実施の形態と異なる構成について主に説明する。

【0140】

SDSシステム130においては、ボリュームのスナップショット機能、ローカル装置または別装置へボリュームをコピーするコピー機能を提供する。バックアップの実行頻度、保管先等は、管理者がバックアップ管理手段（ソフトウェア等）に設定し、バックアップ管理手段が設定に基づいてSDSシステム130のスナップショット機能またはコピー機能を実行する。管理サーバ110（データ管理プログラム111の図示は省略する処理部）は、定期的に使用履歴から、バックアップデータの保管先、フルバックアップの取得間隔を判断し、バックアップ実行手段に対して、要求を送信することで、バックアップの取得に係る制御を行う。

20

【0141】

管理サーバ110は、保管先の判断では、第3の実施の形態で示したバックアップデータの使用履歴から、使用頻度と平均復元時間とを算出する。管理サーバ110は、使用頻度が基準値以上の高頻度であり、かつ、平均復元時間が基準値以上である場合、保管先として最も高い階層（レベル1）を選択する。また、管理サーバ110は、使用頻度が基準値以上であり、平均復元時間が基準値未満である場合、または、使用頻度が基準値未満であり、平均復元時間が基準値以上である場合、中間の階層（レベル2）を選択する。また、管理サーバ110は、使用頻度が基準値未満であり、かつ、平均復元時間が基準値未満である場合、最も低い階層（レベル3）を選択する。管理サーバ110は、選択した階層（レベル）に複数のストレージ装置132がある場合、例えば、空き容量がバックアップデータのサイズより大きく、最も空き容量が大きいストレージ装置132を保管先として選択する。管理サーバ110は、空き容量がバックアップデータのサイズより大きなストレージ装置132がない場合、1段低い階層のストレージ装置132から選択してもよい。

30

【0142】

なお、管理サーバ110は、バックアップデータの使用履歴から平均復元時間を算出して、使用頻度を用いることなく、保管先を判断してもよい。

40

【0143】

さらに、管理サーバ110では、使用履歴をもとに、バックアップデータの取得間隔を判断する。管理サーバ110は、バックアップデータ別に使用頻度を算出する。ここでは、バックアップデータの取得単位を元に、フルバックアップデータに関連する差分バックアップデータ、増分バックアップデータを識別する。

【0144】

管理サーバ110は、フルバックアップデータの使用頻度が基準値以下である場合、フルバックアップデータ、差分バックアップデータまたは増分バックアップデータの間隔を長くする。長くする割合としては、予め決めた値を設定してもよいし、使用頻度に併せて

50

割合を決定してもよい。

【 0 1 4 5 】

管理サーバ110は、差分バックアップデータまたは増分バックアップデータの使用頻度が基準値以上であり、差分バックアップデータまたは増分バックアップからの平均復元時間が基準値以上である場合、フルバックアップの取得間隔を短くする。短くする割合については、予め決めた値を設定してもよいし、差分バックアップまたは増分バックアップの間隔として、差分バックアップまたは増分バックアップ間隔を併せて短くしてもよい。使用頻度の比較では、フルバックアップデータのみの使用頻度と差分バックアップデータまたは増分バックアップデータの使用頻度とを比較してもよい。差分バックアップデータまたは増分バックアップデータの使用頻度の方がフルバックアップのみより多い場合、上記を設定する。

10

【 0 1 4 6 】

管理サーバ110は、差分バックアップデータまたは増分バックアップデータの使用頻度が基準値以上で、差分バックアップまたは増分バックアップからの平均復元時間が基準値未満である場合、差分バックアップまたは増分バックアップの取得間隔を短くする。この場合、フルバックアップの間隔は変更しない。

【 0 1 4 7 】

それ以外の場合は、フルバックアップ、差分バックアップまたは増分バックアップとも取得間隔を変更しない。

【 0 1 4 8 】

ここで、基準値については、予め管理者またはシステムにより設定された値を使用する。平均復元時間に対する基準値については、使用履歴から要求された可能待ち時間の平均値を使用してもよい。

20

【 0 1 4 9 】

このように、処理部は、例えば、記憶部に記憶されている情報から算出した保管データの復元に要した時間の代表値（例えば、平均復元時間）に基づいて、保管データの保管先を変更する処理、および／または、記憶部に記憶されている情報（使用履歴管理情報）から算出した保管データが使用される頻度に基づいて、保管データを取得する間隔を変更する処理を行う。

【 0 1 5 0 】

本実施の形態によれば、保管データの使用状況に併せて、提供にかかる復元時間を考慮してデータの精度を向上できるよう、保管データの取得に係る制御を行うことを可能にする。

30

【 0 1 5 1 】

なお、第3の実施の形態、第4の実施の形態については、バックアップデータが本来の使用目的である障害時のデータ回復時に使用についても、要求される復元時間を満たす必要がある。よって、管理サーバ110がバックアップデータの配置、取得頻度の制御を行ってよいかどうかをAP毎に管理者が設定してもよい。その場合、制御が許可された場合のみ上記の制御を実行する。

【 0 1 5 2 】

40

(5) 他の実施の形態

なお上述の実施の形態においては、本発明をデータ管理システムに適用するようにした場合について述べたが、本発明はこれに限らず、この他種々のシステム、装置、方法、プログラムに広く適用することができる。

【 0 1 5 3 】

また上述の実施の形態においては、説明の便宜上、XXテーブル、XXファイルを用いて各種のデータを説明したが、データ構造は限定されるものではなく、XX情報などと表現してもよい。

【 0 1 5 4 】

また、上記の説明において各機能を実現するプログラム、テーブル、ファイル等の情報

50

は、メモリや、ハードディスク、S S D (Solid State Drive) 等の記憶装置、または、I C カード、S D カード、D V D 等の記録媒体に置くことができる。

【0155】

また上述した構成については、本発明の要旨を超えない範囲において、適宜に、変更したり、組み替えたり、組み合わせたり、省略したりしてもよい。

【符号の説明】

【0156】

1 0 0 データ管理システム、1 1 0 管理サーバ、1 2 0 ホスト計算機、1 3 0 S D S システム、1 4 0 クライアント端末。

10

20

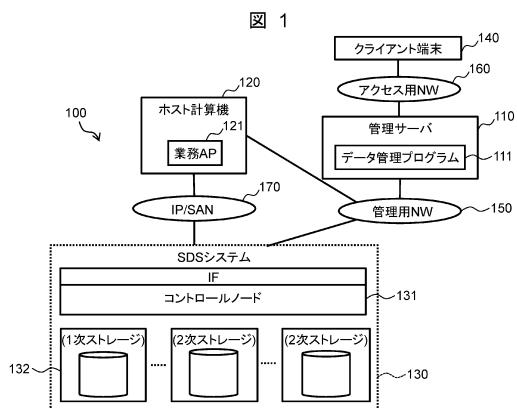
30

40

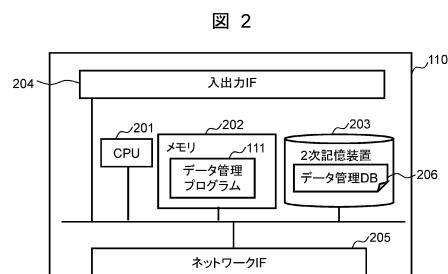
50

【図面】

【図 1】



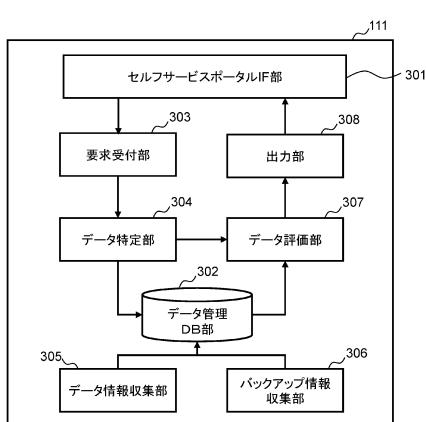
【図 2】



10

【図 3】

図 3



【図 4】

図 4

データ識別子				バックアップ情報			
日付	識別子	取得単位	保管場所	サイズ	取扱時刻	ストレージ1/LU1	ストレージ2/LU2
DB1/ Table-A	2017/5/1 FB01	フルバックアップ	FB01	2GB	00:00:00		
	2017/5/7 FB01-A1	FB01に対する差分	FB01-LU2	100MB	00:00:00		
	2017/5/14 FB01-A2	FB01に対する差分	FB01-LU2	150MB	00:00:00		
	...						

20

30

40

50

【図 5】

図 5

AP識別子	APタイプ	バックアップ手段
DB1	RDB_A	APツール
DB2	RDB_A	APツール
DB3	RDB_B	ストレージコピー

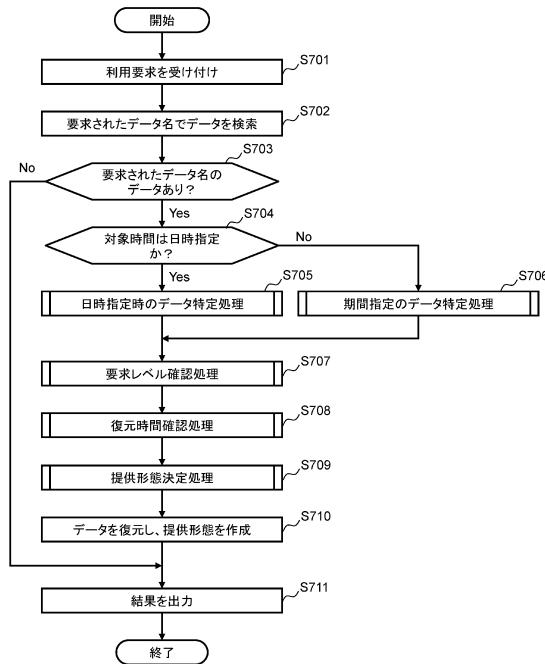
【図 6】

図 6

データ識別子	タイプ	タイムスタンプカラム
Table-A	更新型	更新時刻
Table-B	追記型	追加時刻
Table-C	更新型	—

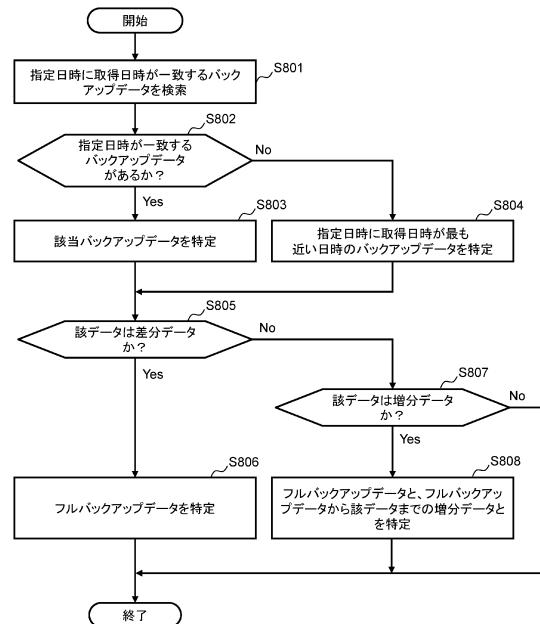
【図 7】

図 7



【図 8】

図 8



10

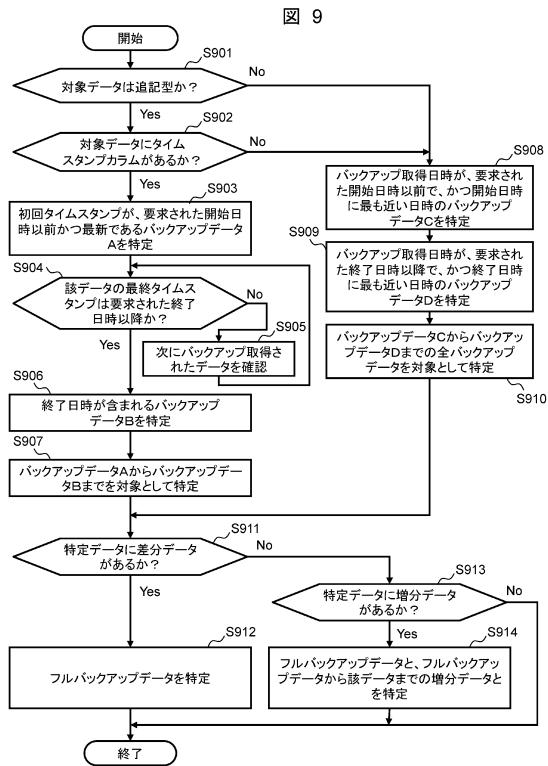
20

30

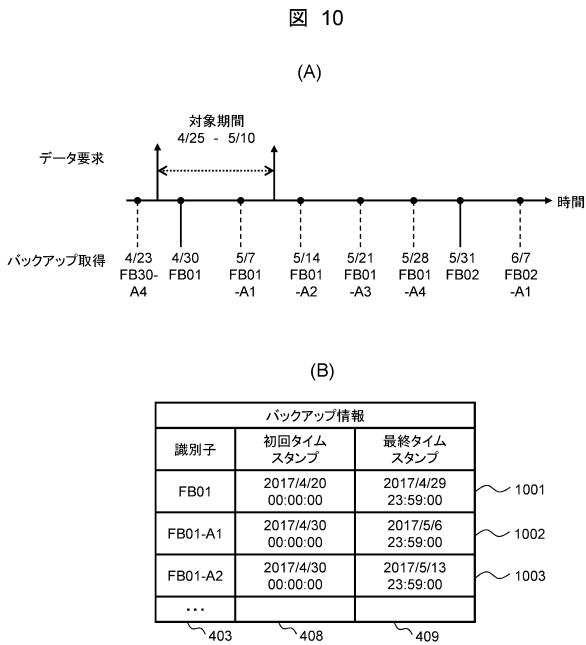
40

50

【図 9】



【図 10】



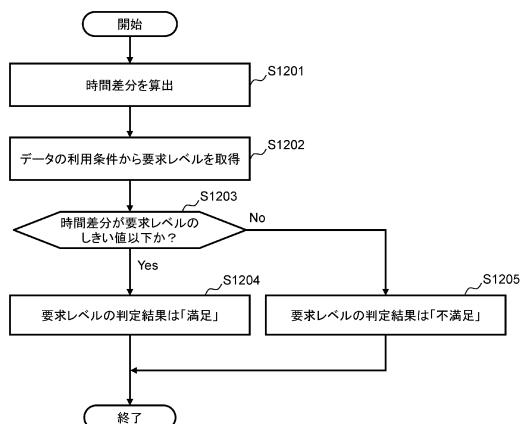
【図 11】

図 11

ストレージ識別子	タイプ	プロトコル	ディスク種別	性能	ボリューム識別子
ストレージA	ブロック	FC, iSCSI, ...	SSD	3,000	LU1, LU3
			SAS	200	LU2
ストレージB	ファイル	NFS, CIFS, ...	SAS	250	—
ストレージC	オブジェクト	HTTP	SATA	100	—
...					

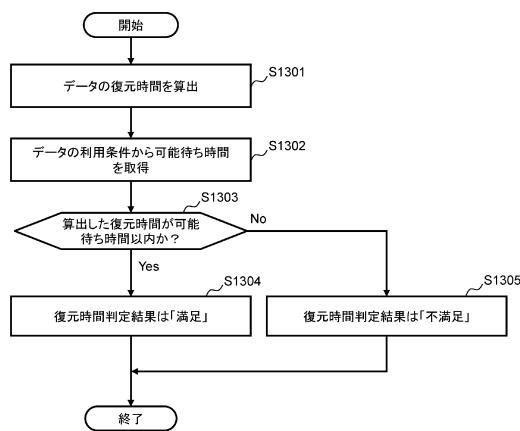
【図 12】

図 12



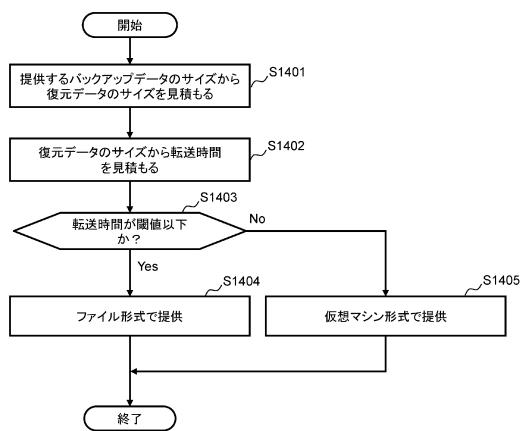
【図 13】

図 13



【図 14】

図 14



【図 15】

図 15

1500

[要件入力]		
1503 データ名	DB1:Table A	
対象時間		
1504 固定期間指定	開始日時 2017/04/20 00:00:00	終了日時 2017/05/10 00:00:00
対象日時		
1505 口日時指定	2017/04/20 00:00:00	
要求レベル		
1506 高		
可能待ち時間		
1507 1 [hour]		
[結果]		
1508 提供データ名	DB1:Table A	
1509 提供形態	CSVファイル	
1510 アクセス先	¥¥result¥result.csv	
要件満足レベル		
- 時間の差分	1日	
- 提供までの時間	30分	

1501 1502

This form captures user requirements for data recovery, including source table, time range, and delivery preferences. It also displays the resulting output parameters like file format and access location.

【図 16】

図 16

1600

業務識別子	データ項目名	AP	AP内データ識別子
顧客別売上分析	顧客氏名	顧客管理・RDB	顧客DB/顧客テーブル
	品名	品名管理・NoSQL	在庫DB/品名テーブル
	売上記録	伝票画像管理	伝票オブジェクト
...			

1601 1602 1603 1604

This table maps business identifiers to specific data items and their corresponding Application Programming Interface (API) and internal data identifier names.

10

20

30

40

50

【図17】

図17

レベル	ストレージ識別子	ディスク	性能	空き容量	1700
1	ストレージ1	SSD	3,000	1TB	1701
2	ストレージ1	SAS	300	1TB	1702
	ストレージ2	SAS	300	1TB	1703
3	ストレージ3	SATA	100	1TB	1704
	...				1705

【図18】

図18

バックアップ識別子	使用日時	復元時間	可能待ち時間	1800
FB01	2018/04/01	15m	10m	1801
	2018/04/02	20m	10m	1802
	2018/04/10	20m	10m	1803
	2018/04/15	25m	10m	1804
...				

10

20

30

40

50

フロントページの続き

(56)参考文献 特開2010-231394(JP,A)
 特開2000-347925(JP,A)
 特開2010-152707(JP,A)
 特開2005-122611(JP,A)

(58)調査した分野 (Int.Cl., DB名)
 G06F 11/14
 G06F 16/11
 G06F 16/14