

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-73039

(P2007-73039A)

(43) 公開日 平成19年3月22日(2007.3.22)

(51) Int. Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00 501B	5B065
G06F 3/06 (2006.01)	G06F 3/06 301Z	5B082
	G06F 12/00 520E	

審査請求 未請求 請求項の数 24 O L 外国語出願 (全 15 頁)

(21) 出願番号	特願2006-215130 (P2006-215130)	(71) 出願人	390009531
(22) 出願日	平成18年8月7日(2006.8.7)		インターナショナル・ビジネス・マシー ズ・コーポレーション
(31) 優先権主張番号	11/199456		INTERNATIONAL BUSIN ESS MACHINES CORPO RATION
(32) 優先日	平成17年8月8日(2005.8.8)		アメリカ合衆国10504 ニューヨーク 州 アーモンク ニュー オーチャード ロード
(33) 優先権主張国	米国 (US)	(74) 代理人	100108501
			弁理士 上野 剛史
		(74) 代理人	100112690
			弁理士 太佐 種一
		(74) 代理人	100091568
			弁理士 市位 嘉宏

最終頁に続く

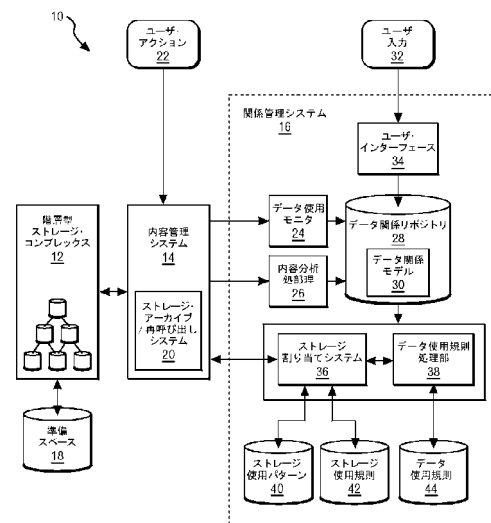
(54) 【発明の名称】 内容ベースの予見的ストレージ管理を提供するシステム、プログラム製品、方法、およびコンピ
ュータ・ソフトウェア

(57) 【要約】

【課題】 階層型ストレージ管理 (HSM) システムおよび方法を提供すること。

【解決手段】 階層型ストレージ管理 (HSM) システムであって、内容管理システムの管理する階層型ストレージ・コンプレックス内のデータ・オブジェクトからデータ・オブジェクト情報を抽出するためのデータ使用モニタと、データ・オブジェクト情報を保存するためのデータ関係リポジトリであって、前記データ・オブジェクト情報は前記階層型ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含む、データ関係リポジトリと、前記関係データを分析し、前記階層型ストレージ・コンプレックス用のデータ管理アクション推奨を作成するシステムとを含む、HSMシステムが提供される。

【選択図】 図 1



【特許請求の範囲】

【請求項 1】

階層型ストレージ管理（H S M）システムであって、
内容管理システムの管理する階層型ストレージ・コンプレックス内のデータ・オブジェクトからデータ・オブジェクト情報を抽出するためのデータ使用モニタと、
データ・オブジェクト情報を保存するためのデータ関係リポジトリであって、前記データ・オブジェクト情報は前記階層型ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含む、データ関係リポジトリと、
前記関係データを分析し、前記階層型ストレージ・コンプレックス用のデータ管理アクション推奨を作成するシステムと
を含む、H S Mシステム。

10

【請求項 2】

前記データ・オブジェクト情報が、データ・オブジェクト記述子、データ・アクセス記述子、およびデータ関係記述子を含む、請求項 1 に記載の H S M システム。

【請求項 3】

データ・オブジェクトに関する内容属性およびメタデータを分析し、前記データ・オブジェクト情報とともに前記データ関係リポジトリ内に保存される内容情報を生成する内容分析処理部をさらに含む、請求項 1 に記載の H S M システム。

【請求項 4】

ユーザが関係データを前記データ関係リポジトリへと入力できるようにするユーザ・インターフェースをさらに含む、請求項 1 に記載の H S M システム。

20

【請求項 5】

前記内容管理システム用のデータ管理アクション推奨の優先度決定を行うストレージ割り当てシステムをさらに含む、請求項 1 に記載の H S M システム。

【請求項 6】

前記ストレージ割り当てシステムが、データ管理アクション推奨の優先度決定をストレージ使用規則およびストレージ使用パターンに基づいて行う、請求項 5 に記載の H S M システム。

【請求項 7】

前記階層型ストレージ・コンプレックスが、階層レベル間で移動中のデータ・オブジェクトに対する準備スペースを含む、請求項 1 に記載の H S M システム。

30

【請求項 8】

前記関係データが、関係するオブジェクトの組、関係するオブジェクトの組ごとの関係タイプ、各組のオブジェクトが関係する頻度、および関係するオブジェクトの組ごとの内容管理アクションを含む、請求項 1 に記載の H S M システム。

【請求項 9】

前記内容管理アクションが、マイグレーション、アーカイブ、および再呼び出しからなる群から選択される、請求項 8 に記載の H S M システム。

【請求項 10】

ストレージ管理システム内のデータ・オブジェクト関係を管理するための、コンピュータ可読媒体上に格納されたプログラム製品であって、

40

内容管理システムの管理するストレージ・コンプレックス内のデータ・オブジェクトからデータ・オブジェクト情報を抽出するために構成されたプログラム・コードと、

データ・オブジェクト情報を保存するためのデータ関係リポジトリであって、前記データ・オブジェクト情報は前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含む、データ関係リポジトリと、

前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション推奨を作成するために構成されたプログラム・コードと

を含む、プログラム製品。

【請求項 11】

50

前記データ・オブジェクト情報が、データ・オブジェクト記述子、データ・アクセス記述子、およびデータ関係記述子を含む、請求項 10 に記載のプログラム製品。

【請求項 12】

データ・オブジェクトに関する内容属性およびメタデータを分析し、前記データ・オブジェクト情報とともに前記データ関係リポジトリ内に保存される内容情報を生成するために構成されたプログラム・コードをさらに含む、請求項 10 に記載のプログラム製品。

【請求項 13】

ユーザが関係データを前記データ関係リポジトリへと入力できるようにするユーザ・インターフェースをさらに含む、請求項 10 に記載のプログラム製品。

【請求項 14】

前記内容管理システム用のデータ管理アクション推奨の優先度決定を行うために構成されたプログラム・コードをさらに含む、請求項 10 に記載のプログラム製品。

【請求項 15】

前記データ管理アクション推奨の優先度決定が、ストレージ使用規則およびストレージ使用パターンに基づいて行われる、請求項 14 に記載のプログラム製品。

【請求項 16】

前記ストレージ・コンプレックスが、ストレージ・レベル間で移動中のデータ・オブジェクトに対する準備スペースを含む、請求項 10 に記載のプログラム製品。

【請求項 17】

前記関係データが、関係するオブジェクトの組、関係するオブジェクトの組ごとの関係タイプ、各組のオブジェクトが関係する頻度、および関係するオブジェクトの組ごとの内容管理アクションを含む、請求項 10 に記載のプログラム製品。

【請求項 18】

前記内容管理アクションが、マイグレーション、アーカイブ、および再呼び出しからなる群から選択される、請求項 10 に記載のプログラム製品。

【請求項 19】

ストレージ・コンプレックス内のデータ・オブジェクト用の予見的ストレージ管理を提供する方法であって、

データ・オブジェクトを分析して内容情報を生成すること、

前記内容情報を、類似の内容情報をもつオブジェクトを識別する関係リポジトリへと格納すること、および

推奨されるストレージ管理アクションを、前記関係リポジトリ内で識別された関係に基づいて作成すること

を含む方法。

【請求項 20】

前記推奨されるストレージ管理アクションの優先度決定を行うさらに別のステップを含む、請求項 19 に記載の方法。

【請求項 21】

前記内容情報が、内容特徴およびメタデータを含む、請求項 19 に記載の方法。

【請求項 22】

前記推奨されるストレージ管理アクションが、マイグレーション、アーカイブ、および再呼び出しからなる群から選択されるアクションを含む、請求項 19 に記載の方法。

【請求項 23】

ストレージ・コンプレックス内のデータ・オブジェクトを管理するための予見的ストレージ管理アプリケーションを展開する方法であって、

データ・オブジェクト情報を前記ストレージ・コンプレックス内のデータ・オブジェクトから抽出すること、

前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含む、データ・オブジェクト情報を格納すること、および

前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション

10

20

30

40

50

ン推奨を作成すること

を行うように動作可能であるコンピュータ・インフラストラクチャを提供することを含む方法。

【請求項 24】

ストレージ・コンプレックス内のデータ・オブジェクトを管理するための予見的ストレージ管理アプリケーションを実装する伝播信号の形で実施されるコンピュータ・ソフトウェアであって、前記コンピュータ・ソフトウェアは、コンピュータに以下の機能、すなわち、

データ・オブジェクト情報を前記ストレージ・コンプレックス内のデータ・オブジェクトから抽出すること、

前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含む、データ・オブジェクト情報を格納すること、および

前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション推奨を作成すること

を行わせる命令を含む、コンピュータ・ソフトウェア。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ストレージ管理に関し、より詳細には、内容ベースの予見的ストレージ管理を提供するシステムおよび方法に関する。

【背景技術】

【0002】

階層型ストレージ管理（HSM、Hierarchical Storage Management）は、今では、事実上すべてのアーカイバル・システムにとって必須の最小限の機能と考えられている。現在のHSMシステムは、経過時間（age）、サイズ、使用頻度、または何らかの、ユーザによって割り当てられたファイル優先度のような低レベルの属性から導出される、明示的に述べられた規則ベースのポリシーによって駆動されている。

【0003】

ディスクなどの高価なメディアから、テープなどのより安価で低速なメディアへのデータの移動を、通常、律している（govern）単純な規則では、データ属性、内容、および相互関係の包括的な考慮に基づいた、知的で、事前対策となる（proactive）データ・マイグレーションはサポートされない。したがって、階層型ストレージ空間では、準最適な管理しか行われず、システム・パフォーマンスが低下する。

【0004】

現在のHSMシステムでは、アーカイバルおよび空間の管理の問題を解決するために、低レベルのファイル属性に基づいて明示的なポリシーを設定する機能を提供している。TSM HSM（商標）およびLEGATO Disk Extender（商標）は、こうした機能を提供するシステムのいくつかである。これらのシステムでは、アーカイバル用の候補であるデータを判断する非常に単純な手段を使用している。同様に、データ・再呼び出し動作は、通常、特定のユーザ要求によってトリガされる。

【0005】

現在の最新の階層型ストレージ管理ツールの使用する単純な属性は、階層的に配置された1組のストレージ・システム全体にわたるデータ・マイグレーション（アーカイバルおよび再呼び出し）の現在の精緻な必要性に取り組むものではない。皮肉にも、アーカイバル動作では、アーカイブへと提出されたデータ・オブジェクト間の関係を記述する、容易に利用可能で重要な情報が無視されている。同様に、再呼び出し動作では、ユーザの必要としそうなデータ・オブジェクトの再呼び出しを予見していない。したがって、データ・マイグレーション（再呼び出しおよび2次的アーカイバル）は、より精緻な規則、ポリシー、および情報とともに、機能しておらず、またそれは可能でもない。

【0006】

10

20

30

40

50

現在の H S M システムは、ファイル・サイズ、経過時間、使用頻度などの低レベルの属性に基づいている。このため、H S M ポリシは、限られた属性の組に対して働くように制約されている。これは、単にファイル・サイズなどではなく、内容に基づいてファイルをマイグレーションすることにユーザの興味がある複雑なストレージの想定状況では、限られた役にしか立たない。ユーザには、ストレージ階層全体にわたる移行のためにより高レベルのポリシを、「1998 年より前に行われた薬品の裁判に関し、化合物 X、Y、および Z に触れているファイルをすべて、テープ・ストレージへとマイグレーションせよ」、または「プロジェクト番号 1 S 2 3 に言及しているファイルをすべて、より安価な S A T A ディスクへとマイグレーションせよ」、または「退院している患者のカルテおよび関係する文書をすべてマイグレーションせよ」のように設定する柔軟性はない。同様に、現在の H S M システムでは、X を標準の S Q L 述語または条件とすることができ、Y を、場合によってはストレージ・プールの一部として、コストおよびパフォーマンスの決められたあるタイプのストレージ・メディアとする、「X を満たすファイルをすべて、ストレージ・メディア Y へとマイグレーションせよ」などのより複雑なポリシを扱うことができない。

10

【0007】

こうしたタイプのポリシは、既存の H S M システムによってはサポートすることができないが、これは、そのようなシステムは、ポリシ属性の数および性質によって厳しく制限されているためである。さらに、ストレージ属性は、特に、ストレージ・メディアのパフォーマンスおよびコストの特性、したがって、問題のメディアによってもたらされる潜在的なサービス・レベルに関して、比較的制限されている。

20

【0008】

現在の H S M システムは、ポリシ属性（データ・オブジェクトおよびストレージ・メディア）に依存する、「X ならば Y」（If X, then Y）という明示的に定義された規則によって駆動される。しかし、現在のシステム・ストレージの想定状況はより複雑である。ユーザは、常に可能な規則をすべて定義できるわけではない。理想的には、H S M システムも、1 組の暗黙の規則によって律されるべきである。たとえば、ユーザは、常に、病院請求書およびカルテからなる特定の組を同時にマイグレーションすることができる。病院請求書およびカルテに関係するオブジェクト・クラスの組の間には暗黙の関係があること、たとえば、それらが退院させた特定の患者に属するかもしれないということは、極めてあり得ることである。このため、退院させた患者の病院請求書をすべて転送する明示的な H S M ポリシがある場合は、その患者のカルテもマイグレーションすることには意味がある。

30

【0009】

同様に、ある患者のカルテがテープ・ストレージから引き出されている場合には、病院請求書も引き出すのが賢明であり得る。カルテを見ている保険代理人（insurance agent）が、病院請求書も調べたいと思うことがあり得る。明示的な別々のデータ・マイグレーション・コマンドを 2 つ発行する代わりに、記録が自動的に単一のコマンド 1 つで引き出され、それにより、テープ待ち時間（tape latency）が減らされるならば、それは好ましいはずである。こうしたタイプの規則は、現在の H S M システムでは実装することができない。

40

【発明の開示】

【発明が解決しようとする課題】

【0010】

本発明は、上述の、ならびに他の諸問題に取り組むために、データのアーカイバルおよび再呼び出しに関する情報を識別および追跡し、この情報をデータ・アーカイバル・システムのパフォーマンスを改善するためのしくみの内部で規則、ポリシ、およびデータ関係として成文化するためのシステムおよび方法を提供している。本発明は、明示的または暗黙のデータ関係を、システム監視およびユーザ入力を通して捉えるための機構を含む。また、本発明は、データ関係を、ストレージ空間の割り当ておよび使用の改善のための推奨

50

(recommendations) を出力するアルゴリズムによって処理するための機構を含む。

【課題を解決するための手段】

【 0 0 1 1 】

第 1 の態様では、本発明により、階層型ストレージ管理 (H S M) システムであって、内容管理システムの管理する階層型ストレージ・コンプレックス内のデータ・オブジェクトからデータ・オブジェクト情報を抽出するためのデータ使用モニタと、データ・オブジェクト情報を保存するためのデータ関係リポジトリであって、前記データ・オブジェクト情報は前記階層型ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含み、データ関係リポジトリと、前記関係データを分析し、前記階層型ストレージ・コンプレックス用のデータ管理アクション推奨を作成するシステムとを含む、 H S M システムが提供される。 10

【 0 0 1 2 】

第 2 の態様では、本発明により、ストレージ管理アクション内のデータ・オブジェクト関係を管理するための、コンピュータ可読媒体上に格納されたプログラム製品であって、内容管理システムの管理するストレージ・コンプレックス内のデータ・オブジェクトからデータ・オブジェクト情報を抽出するために構成されたプログラム・コードと、データ・オブジェクト情報を保存するためのデータ関係リポジトリであって、前記データ・オブジェクト情報は前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含み、データ関係リポジトリと、前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション推奨を作成するために構成されたプログラム・コードとを含む、プログラム製品が提供される。 20

【 0 0 1 3 】

第 3 の態様では、本発明により、ストレージ・コンプレックス内のデータ・オブジェクト用の予見的ストレージ管理を提供する方法であって、データ・オブジェクトを分析して内容情報を生成すること、前記内容情報を、類似の内容情報をもつオブジェクトを識別する関係リポジトリへと格納すること、および推奨されるストレージ管理アクションを、前記関係リポジトリ内で識別された関係に基づいて作成することを含む方法が提供される。

【 0 0 1 4 】

第 4 の態様では、本発明により、ストレージ・コンプレックス内のデータ・オブジェクトを管理するための予見的ストレージ管理アプリケーションを展開する方法であって、データ・オブジェクト情報を前記ストレージ・コンプレックス内のデータ・オブジェクトから抽出すること、前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含み、データ・オブジェクト情報を格納すること、および前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション推奨を作成することを行うように動作可能であるコンピュータ・インフラストラクチャを提供することを含む方法が提供される。 30

【 0 0 1 5 】

第 5 の態様では、本発明により、ストレージ・コンプレックス内のデータ・オブジェクトを管理するための予見的ストレージ管理アプリケーションを実装する伝播信号の形で実施されるコンピュータ・ソフトウェアであって、前記コンピュータ・ソフトウェアは、コンピュータに以下の機能、すなわち、データ・オブジェクト情報を前記ストレージ・コンプレックス内のデータ・オブジェクトから抽出すること、前記ストレージ・コンプレックス内のデータ・オブジェクトに関する関係データを含み、データ・オブジェクト情報を格納すること、および前記関係データを分析し、前記ストレージ・コンプレックス用のデータ管理アクション推奨を作成することを行わせる命令を含む、コンピュータ・ソフトウェアが提供される。 40

【 0 0 1 6 】

上記によって、本発明により、予測的なアーカイバルおよび再呼び出しの動作を可能にすること、予測的なマイグレーション準備スペースを規則 / ポリシ・ベースで律することを可能にすること、およびユーザに対してアーカイバルおよび再呼び出しの動作のための 50

反応時間の改善を可能にすることを含む様々なストレージ管理機能が提供される。

【0017】

この発明のこうしたおよび他の機能は、本発明の様々な態様の以下の詳細な説明を添付の図面とともに解釈することからより容易に理解されよう。

【発明を実施するための最良の形態】

【0018】

ここで図面を参照すると、図1に示すのは、階層型ストレージ・コンプレックス12、内容管理システム14、およびデータ関係管理システム16を全体として含む、階層型ストレージ管理(HSM)システム10である。階層型ストレージ・コンプレックス12は、様々なタイプのストレージ機構、たとえば、ハードディスク・ドライブ(hard drive)、テープ、ディスク、光ディスク(optical)などを、データ・オブジェクトの格納のために含んでもよい。この発明の目的で、HSMシステム10は、データ・オブジェクトのあるストレージ機構から別のストレージ機構へと移動させることができる、どのようなタイプのストレージ管理システムを含んでもよい。

10

【0019】

1つのデータ・オブジェクトは、HSMシステム10が単一の要素として管理するデータのどのような部分を含んでもよい。通常、各ファイルは別々のオブジェクトと考えるが、1組のファイルも1つのデータ・オブジェクトと考えてよい。HSMシステム10の仕様により、個々のファイル、ファイルの組、またはそのどちらもがそのストレージ階層全体にわたって管理される(すなわち、アーカイブされ、再呼び出しされ、マイグレーションされる)かどうかは決定されることになる。したがって、本発明は、個々のファイル、ファイルの組、またはそのどちらともすることのできるデータ・オブジェクトに対して動作する。

20

【0020】

内容管理システム14は、一企業内の内容を管理するための典型的なシステムを含むが、これは、ファイルがすべて、単にファイル・システム内に格納される非常に低レベルのシステムから、データが、それぞれの特定のデータ・オブジェクトを記述するメタデータのあるデータベース内に格納されるより精緻なシステムに及ぶ可能性がある。内容管理システム14は、階層型ストレージ・コンプレックス12内のデータ・オブジェクトに関係するどのようなアクション22にも応答する。こうした動作には、たとえば、ファイルを開く/編集すること、アーカイブされたデータ・オブジェクトを再呼び出しする要求などがある。

30

【0021】

言及した通り、内容はメタデータとともに格納してもよい。内容管理システム14のタイプに応じて、メタデータは、ファイル・サイズ、経過時間、使用頻度などの低レベルの属性を含むのでも、またはファイル内容の記述、たとえば、電子カルテ用の、患者番号、疾病名、処方の詳細などのような高レベルの属性を含むのでもよい。内容管理システム14には、データ・オブジェクトに対して関連付けられているメタデータとともにアクセスする機能がある。

【0022】

内容管理システム14とともに含まれているのが、ストレージ・アーカイブ/再呼び出し(SAR、storage archive/recall)システム20である。SARシステム20は、内容管理システム14とともに相互運用される独立のシステムとすることもでき、または内容管理システム14と完全に一体化させることもできる。SARシステム20では、ストレージ・コンポーネントすべてを常に把握している。そこにある情報は、ストレージのタイプ(たとえば、テープ、ディスク、光ディスク、など)、各ストレージ・タイプに関連付けられているコスト(たとえば、そのストレージ・メディアの価格およびデータをそのストレージから引き出すことに関連する取り出し時間)、各ストレージ・サブタイプの容量、各サブタイプで使用されるストレージの量、データを「より安価な」メディアへとマイグレーションすることが必要になる前に許されるストレージ使用の最大量を指定するユ

40

50

ーザ定義の閾値、そのストレージ内での各データ・ファイルのアクセス・パターン（たとえば、同時に移動されるオブジェクトを把握したもの）についてのものである。

【 0 0 2 3 】

データ関係管理システム 1 6 では、データ・オブジェクトが利用される様子に関する関係および他の情報を分析および識別することによって、内容管理システム 1 4 のとるストレージのアーカイブおよび取り出し（retrieval）のアクションを最適化している。データ関係管理システム 1 6 の中央には、データ関係リポジトリ 2 8 があり、ここでは、データ関係モデル 3 0 に従ってデータの編成が行われる。データ関係リポジトリ 2 8 は、一意に識別されるデータ・オブジェクトそれぞれについての情報を含んでいる。また、これは、データ・オブジェクト同士の関係に関するすべての情報を含んでいる。データ関係モデル 3 0 により、データ・オブジェクト情報を編成する構造が提供される。1 組のデータ使用規則 4 4 が、この構造化された情報に適用されて、階層型ストレージ・コンプレックス 1 2 内部でのデータ・オブジェクトの格納および取り出しが管理される。このモデル 3 0 は、3 つのタイプのデータを含む。すなわち、データ・オブジェクト記述子、データ・アクセス記述子、およびデータ関係記述子である。

10

【 0 0 2 4 】

データ・オブジェクト記述子は、各データ・オブジェクトに対する一意の識別子、（たとえば、バイトまたは何か他の尺度での）サイズを含めて、各データ・オブジェクトに対する記述情報、およびデータの型を含む。また、データ・オブジェクト記述子は、特定のデータ・オブジェクトに関係する内容情報（たとえば、メタデータまたは実際のデータ内容）を含んでもよい。データ・アクセス記述子は、データ・アクセス・イベントを記述し、各イベントに対してユーザ識別子、時刻および日付、アクセスされたデータに対する一意の識別子、およびアクセスのタイプ（読み取り、読み取り／書き込み、など）を含む。データ関係記述子は、データ・オブジェクト同士の関係を記述する。データ関係は、一般に、関係する 2 つ以上のデータ・オブジェクトのリスト、関係タイプ、各オブジェクトがその組（set）の一部として現れる頻度、およびその関係の「使用」（usage）を含む。関係タイプとは、データ・オブジェクトがどのように関係するか、たとえば、どちらも類似の内容情報を含む、どちらもユーザによって通常同時に開かれる、などを指す。「使用」により、その関係があてはまる内容管理アクションのタイプ（マイグレーション、アーカイブ、または再呼び出し）が識別される。

20

30

【 0 0 2 5 】

情報のデータ関係リポジトリ 2 8 への流入は、データ使用モニタ 2 4、内容分析処理部 2 6、またはユーザ・インターフェース 3 4 を介して行われる。データ使用モニタ 2 4 では、ストレージ管理システム 1 0 内部で生じるデータへのアクセスに関する情報を受け取る。通常、この情報は、内容管理システム 1 4 内部のあるコンポーネントから送られるはずであり、一意のユーザ識別子、アクセスの時刻および日付、アクセスされたデータに対する一意の識別子、データの（バイトまたは何か他の尺度での）サイズおよび型、およびアクセスのタイプ（たとえば、読み取り、読み取り／書き込み、など）を含むはずである。自由選択で、この内容管理コンポーネントは、一意のアクセス・イベント識別子および 1 組の関係するそれまでのアクセス・イベント識別子を提供してもよい。データ使用モニタ 2 4 は、この情報を処理して、アーカイバルおよび再呼び出しのプロセスならびにストレージ階層内部でのマイグレーションに関するどのようなデータ関係も含めて、データ・オブジェクト記述子情報を作成する。

40

【 0 0 2 6 】

内容分析処理部（CAP、content analysis processor）2 6 は、データ・オブジェクトおよび関連するメタデータ情報を内容管理システム 1 4 から取り込んで、データ・オブジェクトごとの内容情報を作成する。この内容情報は、そのオブジェクトに関連付けられているメタデータと、そのオブジェクトから抽出された内容属性の集合和（union）である。したがって、CAP 2 6 は、データ・オブジェクト記述子を内容情報で詳細にする。CAP 2 6 の実装は、処理すべきデータ・オブジェクトの型および特定のドメインに依存

50

することに注意されたい。たとえば、画像データ用のCAP26は、JPEG、GIF、TIFF、PNGなどを含む幅広い範囲の画像フォーマットを処理するように実装できるはずである。CAP26は、たとえば、色およびテキストに基づく特徴を抽出する画像処理サブルーチンをもつことができるはずである。同様に、薬品試験（drugtrial）データ用のCAP26では、様々な試験管理アプリケーションによって作成された様々な薬品試験フォーマットのサポートが可能であるべきであり、患者名、疾病タイプなど、ヘッダおよび内容情報の抽出が可能であるべきである。

【0027】

内容属性は、データ・オブジェクト内容から抽出でき、かつ適切なデータ管理アクションを決定するのに有用な分類情報を提供する特徴である。内容管理システム14内のメタデータ情報は、内容属性を含むことができることに注意されたい。例としては、スプレッドシート中の日付および財務データ型を示す列見出し（column headers）、画像の色・テキスト・レイアウトの情報、DICOM（医用におけるデジタル画像と通信）画像中にあるDICOMヘッダ情報の内容などがある。

【0028】

ユーザ・インターフェース34により、ユーザが、自分の使用するデータ同士の関係を記述するユーザ入力32を与えることができるようにする機構が提供される。ユーザ・インターフェース34では、ユーザの記述した関係を、データ関係モデル30の規定するフォーマットへと変換し、関係情報を、データ関係リポジトリ28内のデータ関係記述子の一部として保存する。

【0029】

データ関係リポジトリ28に、データ関係モデル30によるデータ・オブジェクト情報がロードされた後、データ使用規則処理部38は、データ使用規則44をこの情報に適用して、推奨されるデータ管理アクションを作成することができる。推奨されるデータ管理アクションは、たとえば、データ・オブジェクトのリスト、アクション（たとえば、マイグレーション、再呼び出し、アーカイブ）、その推奨が有効である期間、確率、または優先度（prioritization）、およびそれらの組み合わせを含むことができる。データ使用規則44は、内容管理システム14内で生じる特定のアクションに応答して適用することができ、たとえば、ファイルをテープ・バックアップから取り出すユーザ要求により、データ関係リポジトリ28内のデータ関係すべてを探す規則をトリガさせることができる。推奨されるデータ管理アクションは、それに付随する確率を含んでいてもよいことに注意されたい。たとえば、そのアクションに、ファイルXが取り出されるとき、ファイルYも取り出される確率がN%あることが述べられていてもよい。

【0030】

ストレージ割り当てシステム36では、データ使用規則処理部38からの推奨されるデータ管理アクションの優先度決定を行い、内容管理システム14に対して1組のストレージ管理アクションを作成する。この1組のストレージ管理アクションは、1組の構成可能なストレージ使用規則42およびそれまでのストレージ使用パターン40に基づく。ストレージ使用規則42には、ストレージ管理アクションが不許可、許可、および推奨になる状況が記述される。1つのストレージ使用規則は、タイプ（不許可、許可、推奨）、1組のストレージ管理アクション（マイグレーション、再呼び出し、アーカイブ）、および条件（ストレージ占有率閾値/上限値、実行可能期間）を含む。ストレージ使用パターン40には、階層型ストレージ・コンプレックス12の様々なレベルにわたるデータの経時的な移動が記述される。

【0031】

言及した通り、階層型ストレージ・コンプレックス12は、物理ストレージ・メディアからなる。これは、異なるストレージ・タイプに関連付けられている異なるストレージ・レベルからなるとすることができる。各ストレージ・レベルは、価格コストおよびアクセス・コストが類似した一群のストレージ・メディアである。たとえば、テープ・メディアはすべてあるストレージ・レベルに属し、高速磁気ディスクはすべて別のストレージ・レ

10

20

30

40

50

ベルに属する。価格コスト (P C、price cost) は、ストレージ・メディアの原価 (cost price) であり、すなわち、テープ・ストレージの価格コストは、ディスク・ストレージよりもずっと低い。アクセス・コスト (A C、accesscost) は、そのストレージ・メディア上の一単位のストレージへのアクセスに必要な時間である。たとえば、テープ・ストレージのアクセス・コストは、ディスク・ストレージよりもずっと高い。

【0032】

ある例示的な実施形態では、階層型ストレージ・コンプレックスの設定を、ストレージ・レベルを1からMax Levelまで番号付けして行うことができ、ここで、Max Levelは存在するストレージ・レベルの総数である。Storage Level (1) は、P Cが最高でA Cが最小のストレージ・メディアからなる。Storage Level (Max Level) は、P Cが最小でA Cが最高のストレージ・メディアからなる。

【0033】

準備スペース (staging space) 18は、ストレージ割り当てシステム36からの確率ベースのストレージ管理推奨を取り扱うように実装することの可能な、階層型ストレージ・コンプレックス12の一部である。各ストレージ・レベル n ($1 \leq n < \text{Max Level}$) に対して、Staging (n)とマークされた、関連付けられたストレージ・コンポーネントがある。Storage Level (n)のある特定の部分が、Staging (n)とマークされる。データを確率的規則 ($0 < \text{確率} < 1$) に基づいて移動させるべきときは、まず準備エリア18に移動させる。したがって、データを予見的な (anticipated) 再呼び出し規則に基づいてStorage Level (n)からStorage Level (n-1)へと移動すべき場合は、まずStaging (n-1)へと移動させる。これにより、実際の再呼び出しコマンドが与えられたときに、データがより高いストレージ・レベルにあるようにする。これに対応して、データを確率的規則に基づいてStorage Level (n)からStorage Level (n+1)へと移動すべき場合は、まずStaging (n+1)へと移動させる。

【0034】

ここで図2を参照すると、ストレージ管理システム10の動作の例示的な実装形態を、流れ図の形で示している。まず、ステップS1で、ユーザが内容管理システム14の制御下にあるデータにアクセスすると、内容管理システム14は、データ・アクセス記述をデータ使用モニタへと送る。この記述は、一意のユーザ識別子、時刻および日付、アクセスされたデータに対する一意の識別子、データの (バイトまたは何か他の尺度での) サイズおよび型、およびアクセスのタイプ (読み取り、読み取り/書き込み、など) を含む。自由選択で、内容管理コンポーネント14は、一意のアクセス・イベント識別子および1組の関係するそれまでのアクセス・イベント識別子を提供してもよい。データ使用モニタ24は、この情報を処理して、アーカイバル、再呼び出し、およびストレージ階層内部でのマイグレーションに関する1組のデータ関係を作成する。このデータ関係は、同じユーザが同時にアクセスしたデータ・オブジェクトを、同時アクセスの頻度に関する指標によって示す。

【0035】

次に、ステップS2で、データ使用モニタ24により、このデータ関係は、データ関係リポジトリ28へと保存される。ステップS3で、ユーザは既存のデータ管理アクション規則を閲覧および更新でき、新しい規則をユーザ・インターフェースを通して追加することができる。規則は、直接に入力することもでき、またはユーザが、ユーザ・インターフェース内で提供される1組のプロンプトを通して、規則を作成するための情報を提供することもできる。データ使用規則は、ユーザ・インターフェース34によって作成される。

【0036】

ステップS4で、内容管理システム14は、内容およびこの内容に関連付けられているメタデータを、内容分析処理部 (CAP) 26へと送る。このメタデータは、MIMEの定めるものなどの明示的なデータ型を含んでもよく、または内容フォーマットを (ファイル・ヘッダによるなどして) それ自体で定義してもよい。一部の場合には、データ・フォ

ーマットを未定義としてもよい。内容分析処理部 26 では、内容が C A P の処理できるフォーマットになっている場合、内容情報をメタデータおよび内容に基づいて生成する。そうでない場合は、メタデータだけを使用して内容情報を作り出す。ステップ S 5 で、内容情報は、データ関係リポジトリ 28 内に、処理中のデータ・オブジェクトに関するデータ記述子の一部としてファイルされる。

【0037】

ステップ S 6 で、データ使用規則処理部 38 およびストレージ割り当てシステム 36 により、推奨されるストレージ管理アクションのリストが作成される。このプロセスの呼び出しは、スケジュールに基づいて、あるいは、新しい空間が利用可能であること、または利用可能な空間の使用率 (usage) があらかじめ設定された閾値に達することに反応するデータ使用モニタ 24 からの促し (prompt) などの外部信号に基づいて行われる。データ使用規則処理部 38 は、データ・オブジェクト管理アクションが期待される尤度が高いデータ・オブジェクトを識別しながら、データ関係リポジトリ 28 へのアクセスおよびスキャンを行う。たとえば、第 2 のデータ・オブジェクトを所与のユーザがアーカイブから再呼び出したばかりであることを、データ使用モニタ 24 が示し、別のデータ・オブジェクトへのアクセスが普通同時に行われることが、そのデータ・オブジェクトに関するデータ関係によって示されている場合、データ使用規則処理部 38 では、内容管理システム 14 がその第 2 の関係するデータ・オブジェクトも再呼び出しする旨の推奨を生成することになる。この推奨されるデータ管理アクションは、ストレージ割り当てシステム 36 へと渡されて優先度決定を行い、内容管理システム 14 へと転送されることになる。

10

20

【0038】

ステップ S 7 で、ストレージ割り当てシステムは、優先度決定の行われたアクションの組をストレージ・アーカイブ / 再呼び出しシステム 20 へと送り、これがそのアクション、たとえば、指定されたデータの異なる Storage Level (ストレージ・レベル) および Staging Space (準備スペース) 間での必要なマイグレーションを実行する。

【0039】

ステップ S 8 で、ストレージ・アーカイブ / 再呼び出しシステム 20 は、階層型ストレージ・コンプレックス 12 の属性すべてのリストをストレージ割り当てシステム 36 へと送る。この属性のリストが含むのは、ストレージのタイプ (テープ、ディスク、光ディスク、など)、存在する様々なストレージ・レベル、各ストレージ・レベルに関連付けられているコスト (P C と A C のどちらも)、各ストレージ・レベルの総容量、各レベルで使用中のストレージの量、データをマイグレーションすることが必要になる前に許されるストレージ使用の最大量を指定するユーザ定義の閾値、そのストレージ内での各データ・ファイルのアクセス・パターン (たとえば、どのくらい頻繁に使用されているか、同時に移動されるオブジェクトを把握したものなど) などである。

30

【0040】

一般に、上で説明した様々なコンポーネントは、コンピュータ・システム、たとえば、デスクトップ、ラップトップ、ワークステーションなどの上で実装することができ、それらは、クライアントまたはサーバあるいはその両方の一部としても実装することができる。このコンピュータ・システムへのアクセスは、インターネット、L A N (local area network、構内通信網)、W A N (wide area network、広域通信網)、V P N (virtual private network) などのネットワークを介して提供することができる。通信は、直接ハードワイヤード接続 (たとえば、シリアル・ポート) を介して、または有線または無線あるいはその両方の伝送方法のどのような組み合わせを利用してもよいアドレス可能な接続を介して行うことができよう。しかも、トークン・リング、イーサネット (R)、W i F i、または他の従来の通信標準などの従来のネットワーク接続を使用できよう。さらに、接続は、従来の T C P / I P ソケット・ベースのプロトコルによって提供できよう。この場合、I S P (Internet service provider) を使用して、相互接続を確立できよう。接続は、J M S などのメッセージ・ベースのプロトコル、あるいは W e b サービスなど、自己発

40

50

見型の (self-discovered) または他のサービス・ベースのプロトコルによって提供できよう。さらに、上で示したように、通信はクライアント/サーバまたはサーバ/サーバの環境で行うことができよう。

【0041】

本発明の教示は、ビジネス・モデル (business method) として予約支払 (subscription) 制または料金 (fee) 制で提供できるはずであることを理解されたい。たとえば、関係管理システム 16 を含むコンピュータ・システムの作成、維持、または展開あるいはそれらの組み合わせは、本明細書で説明した機能を顧客に提案するサービス・プロバイダが行うことができよう。すなわち、サービス・プロバイダは、上で説明した推奨されるデータ管理アクションを提供する内容管理システムへのインターフェースの提供を提案すること

10

【0042】

本明細書で説明したシステム、機能、機構、方法、エンジン、およびモジュールは、ハードウェア、ソフトウェア、またはハードウェアとソフトウェアの組み合わせの形で実装できることが了解されている。これらは、本明細書で説明した諸方法の実現用に適合されたどのようなタイプのコンピュータ・システムまたは他の機器によっても実装することができる。ハードウェアとソフトウェアの典型的な組み合わせは、ロードされ実行されると、本明細書で説明した諸方法を実現するようなコンピュータ・システムを制御する、コンピュータ・プログラム付きの汎用コンピュータ・システムとすることができよう。あるいは、本発明の機能タスクのうちの1つまたは複数を実現するための特化したハードウェア

20

【0043】

また、本発明は、本明細書で説明した方法および機能の実装を可能にする特徴すべてを含み、コンピュータ・システムへとロードされるとこれらの方法および機能を実現することのできる、コンピュータ・プログラム製品中に埋め込むことができる。コンピュータ・プログラム、ソフトウェア・プログラム、プログラム、プログラム製品、ソフトウェアなどの用語は、この文脈では、情報処理能力のあるシステムに、直接に、あるいは、以下のこと、すなわち、(a) 別の言語、コード、または記法への変換と、(b) 異なる有形形式での複製とのどちらかまたは両方の後で、特定の機能を実行させるものである1組の命令の、任意の言語、コード、または記法における任意の表現を意味する。

30

【0044】

本発明の上述の説明は、例示および説明を目的として提示してきた。網羅的であること、または本発明を開示したまさにその形式に限定する、明らかに、多くの修正形態および変形形態が可能である。当業者に明らかである可能性のある修正形態および変形形態は、添付の特許請求の範囲の定めるこの発明の範囲内に含まれるものとする。

【図面の簡単な説明】

【0045】

【図1】本発明の一実施形態による階層型ストレージ管理システムを示す図である。

40

【図2】本発明の一実施形態による例示的な動作を示す流れ図である。

【符号の説明】

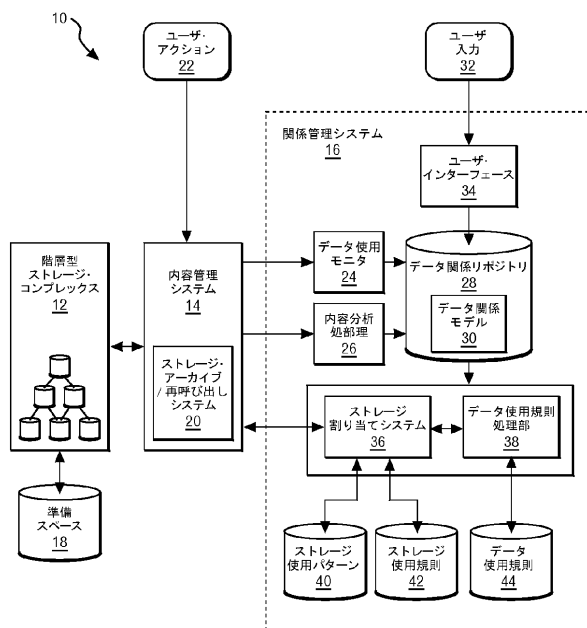
【0046】

- 10 階層型ストレージ管理システム、HSMシステム、ストレージ管理システム
- 12 階層型ストレージ・コンプレックス
- 14 内容管理システム
- 16 データ関係管理システム
- 18 準備スペース、準備エリア
- 20 ストレージ・アーカイブ/再呼び出しシステム、SARシステム
- 22 アクション

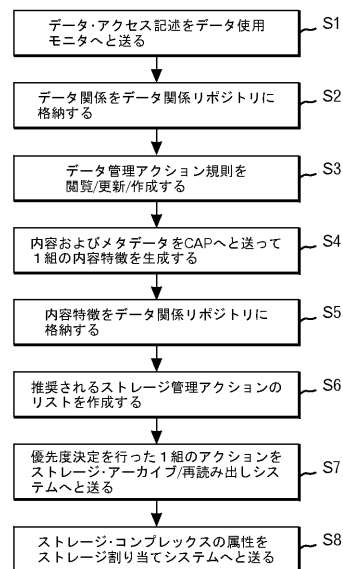
50

- 2 4 データ使用モニタ
- 2 6 内容分析処理部、CAP
- 2 8 データ関係リポジトリ
- 3 0 データ関係モデル、モデル
- 3 4 ユーザ・インターフェース
- 3 6 ストレージ割り当てシステム
- 3 8 データ使用規則処理部
- 4 0 ストレージ使用パターン
- 4 2 ストレージ使用規則
- 4 4 データ使用規則

【図 1】



【図 2】



フロントページの続き

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 ウィリアム・エイ・タルスキー

アメリカ合衆国 1 0 5 8 8、ニューヨーク州 スクラブオーク デインストリート 3 6 2 0

(72)発明者 バムシ・ケイ・ブツクル

アメリカ合衆国 1 0 5 4 9、ニューヨーク州 マウントキスコ グローブストリート 2 8

F ターム(参考) 5B065 BA01 BA03 BA07 CE23 ZA01

5B082 CA11 CA13 CA18 EA07

【外国語明細書】

[2007073039000001.pdf](#)

[2007073039000002.pdf](#)

[2007073039000003.pdf](#)

[2007073039000004.pdf](#)