

(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2013-222230

(P2013-222230A)

(43) 公開日 平成25年10月28日(2013.10.28)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 12/00 (2006.01)	G06F 12/00	510B
G06F 3/06 (2006.01)	G06F 12/00	501B
G06F 3/08 (2006.01)	G06F 3/06	301J
	G06F 3/08	F

審査請求 未請求 請求項の数 18 O L (全 49 頁)

(21) 出願番号 特願2012-91514 (P2012-91514)
 (22) 出願日 平成24年4月13日 (2012.4.13)

(特許庁注：以下のものは登録商標)

1. Blu-ray

(71) 出願人 501009849
 株式会社日立エルジーデータストレージ
 東京都港区海岸三丁目2番23号
 (71) 出願人 509189444
 日立コンシューマエレクトロニクス株式会社
 東京都千代田区大手町二丁目2番1号
 (74) 代理人 100100310
 弁理士 井上 学
 (74) 代理人 100098660
 弁理士 戸田 裕二
 (74) 代理人 100091720
 弁理士 岩崎 重美

最終頁に続く

(54) 【発明の名称】 情報処理システム

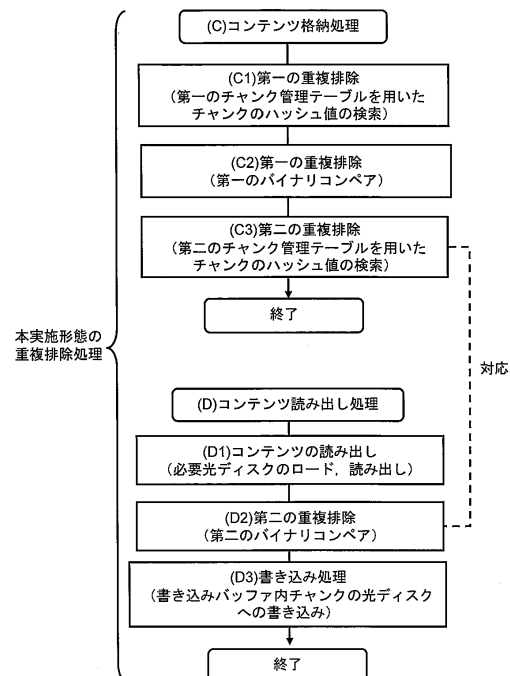
(57) 【要約】

【課題】重複排除においてはバイナリコンペアを行う際に記憶装置に対してランダムアクセスが発生するため、光ディスクライブラリのようなランダムアクセス性能の低い記憶装置を用いて重複排除を行うことは困難であった。

【解決手段】バイナリコンペアを行うために光ディスクドライブに未ロードの光ディスクにアクセスする必要がある場合、バイナリコンペア対象のデータを一時記憶領域に格納してバイナリコンペアを先送りし、光ディスクがロードされたタイミングで先送りされたバイナリコンペアを実行する。また、記憶装置に格納するデータと光ディスクライブラリ内のデータとの重複排除である第二の重複排除を行う前に、記憶装置に格納するデータと一時記憶領域内のデータとの重複排除である第一の重複排除を行う。

【選択図】 図31

Fig.31



【特許請求の範囲】**【請求項 1】**

入力データの格納のときに、重複データを排除する重複排除サーバと、複数の可搬性記録媒体を有し前記入力データを該可搬性記録媒体に格納するライブラリ装置とからなるアーカイブシステムであって、

前記重複排除サーバは、

前記入力データを複数のチャンクに分割し、記憶する第一の記憶領域と、

前記分割されたチャンクに固有のデータを発生させる固有データ発生器と、

前記ライブラリ装置内の可搬性記録媒体に記録されているデータに対応する固有データをテーブル管理する第一の管理部と、を有し

10

前記ライブラリ装置は、

複数の可搬性記録媒体を収納する収納部と、

前記複数の可搬性記録媒体にデータを記録し、および/または、前記複数の可搬性記録媒体からデータを再生するドライブと、

前記収納部から前記可搬性記録媒体を前記ドライブに運搬する運搬部とを有し、

新たに入力データを入力したときは、該入力データを複数のチャンクに分割し、

前記固有データ発生器が該複数の分割されたチャンクごとに固有データを発生し、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致する場合には、該一致したチャンクを第二の記憶領域に記憶しておき、該一致した固有データに対応するデータが格納されている可搬性記録媒体が前記ドライブに挿入されたときに、前記第二の記憶領域に記憶されているチャンクと、前記ドライブに挿入された可搬性記録媒体に格納されているデータとのバイナリコンペアを行う、

20

アーカイブシステム。

【請求項 2】

請求項 1 記載のアーカイブシステムであって、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、記録対象の可搬性記録媒体が前記ドライブに挿入されたときに、前記書き込みバッファ領域に記憶されているチャンクを、前記ドライブに挿入された可搬性記録媒体に書き込む、

30

アーカイブシステム。

【請求項 3】

請求項 1 記載のアーカイブシステムであって、

前記第二の記憶領域に記憶されるチャンクに対応する固有データをテーブル管理する第二の管理部と、を有し

更に新たな入力データを入力したときは、該新たな入力データを複数のチャンクに分割し、

前記固有データ発生器が該複数の分割されたチャンクごとに固有データを発生し、

該分割されたチャンクに対応する固有データと、前記第二の管理部が管理する固有データとの比較を行い、一致する場合には、該一致したチャンクと前記第二の記憶領域に記憶されるチャンクとのバイナリコンペアを行う、

40

アーカイブシステム。

【請求項 4】

請求項 3 記載のアーカイブシステムであって、

該分割されたチャンクに対応する固有データと、前記第二の管理部が管理する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、記録対象の可搬性記録媒体が前記ドライブに挿入されたときに、前記書き込みバッファ領域に記憶されているチャンクを、前記ドライブに挿入された可搬性記録媒体に書き込む、

50

アーカイブシステム。

【請求項 5】

請求項 1 記載のアーカイブシステムであって、
前記可搬性記録媒体は光ディスクである、
アーカイブシステム。

【請求項 6】

請求項 3 記載のアーカイブシステムであって、
前記第二の管理部が管理する固有データとのうち、所定期間内で所定回数以上比較に用いられた固有データに対応するチャンクは、前記第二の記憶領域に留める、
アーカイブシステム。

10

【請求項 7】

請求項 1 記載のアーカイブシステムであって、
該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、前記書き込みバッファ領域に記憶されている関連するチャンクを、前記可搬性記録媒体に書き込む、
アーカイブシステム。

【請求項 8】

請求項 1 記載のアーカイブシステムであって、
該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、該書き込みバッファ領域の空き容量が所定以下の容量になったときに、前記書き込みバッファ領域に記憶されているチャンクを、前記可搬性記録媒体に書き込む、
アーカイブシステム。

20

【請求項 9】

請求項 1 記載のアーカイブシステムであって、
該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致する場合には、該一致したチャンクを第二の記憶領域に記憶しておき、該一致した固有データに対応するデータが格納されている可搬性記録媒体が前記ドライブに挿入されたときに、または、前記第二の記憶領域の空き容量が所定以下の容量となったときに、前記第二の記憶領域に記憶されているチャンクと、前記ドライブに挿入された可搬性記録媒体に格納されているデータとのバイナリコンペアを行う、
アーカイブシステム。

30

【請求項 10】

入力データの格納のときに、重複データを排除する重複排除サーバであって、複数の可搬性記録媒体を有し前記入力データを該可搬性記録媒体に格納するライブラリ装置と接続可能な重複排除サーバであって、

40

前記入力データを複数のチャンクに分割し、記憶する第一の記憶領域と、

前記分割されたチャンクに固有のデータを発生させる固有データ発生器と、

前記ライブラリ装置内の可搬性記録媒体に記録されているデータに対応する固有データをテーブル管理する第一の管理部と、を有し

新たに入力データを入力したときは、該入力データを複数のチャンクに分割し、

前記固有データ発生器が該複数に分割されたチャンクごとに固有データを発生し、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致する場合には、該一致したチャンクを第二の記憶領域に記憶しておき、該一致した固有データに対応するデータが格納されている可搬性記録媒体が前記ドライブに挿入されたときに、前記第二の記憶領域に記憶されているチャンクと、前記ドライブに挿入された可搬性記録媒体に格

50

納されているデータとのバイナリコンペアを行うよう指示する、
重複排除サーバ。

【請求項 1 1】

請求項 1 0 記載の重複排除サーバであって、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、記録対象の可搬性記録媒体が前記ドライブに挿入されたときに、前記書き込みバッファ領域に記憶されているチャンクを、前記ドライブに挿入された可搬性記録媒体に書き込むよう指示する、

重複排除サーバ。

10

【請求項 1 2】

請求項 1 0 記載の重複排除サーバであって、

前記第二の記憶領域に記憶されるチャンクに対応する固有データをテーブル管理する第二の管理部と、を有し

更に新たな入力データを入力したときは、該新たな入力データを複数のチャンクに分割し、

前記固有データ発生器が該複数の分割されたチャンクごとに固有データを発生し、

該分割されたチャンクに対応する固有データと、前記第二の管理部が管理する固有データとの比較を行い、一致する場合には、該一致したチャンクと前記第二の記憶領域に記憶されるチャンクとのバイナリコンペアを行うよう指示する、

20

重複排除サーバ。

【請求項 1 3】

請求項 1 2 記載の重複排除サーバであって、

該分割されたチャンクに対応する固有データと、前記第二の管理部が管理する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、記録対象の可搬性記録媒体が前記ドライブに挿入されたときに、前記書き込みバッファ領域に記憶されているチャンクを、前記ドライブに挿入された可搬性記録媒体に書き込むよう指示する、

重複排除サーバ。

30

【請求項 1 4】

請求項 1 0 記載の重複排除サーバであって、

前記可搬性記録媒体は光ディスクである、

重複排除サーバ。

【請求項 1 5】

請求項 1 2 記載の重複排除サーバであって、

前記第二の管理部が管理する固有データとのうち、所定期間内で所定回数以上比較に用いられた固有データに対応するチャンクは、前記第二の記憶領域に留めるよう指示する、

重複排除サーバ。

【請求項 1 6】

請求項 1 0 記載の重複排除サーバであって、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合には、該一致しないチャンクを書き込みバッファ領域に記憶しておき、前記書き込みバッファ領域に記憶されている関連するチャンクを、前記可搬性記録媒体に書き込むよう指示する、

重複排除サーバ。

40

【請求項 1 7】

請求項 1 0 記載の重複排除サーバであって、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致しない場合に

50

は、該一致しないチャンクを書き込みバッファ領域に記憶しておき、該書き込みバッファ領域の空き容量が所定以下の容量になったときに、前記書き込みバッファ領域に記憶されているチャンクを、前記可搬性記録媒体に書き込むよう指示する、
重複排除サーバ。

【請求項 18】

請求項 10 記載の重複排除サーバであって、

該分割されたチャンクに対応する固有データと、前記第一の管理部が管理する可搬性記録媒体に記録されているデータに対応する固有データとの比較を行い、一致する場合には、該一致したチャンクを第二の記憶領域に記憶しておき、該一致した固有データに対応するデータが格納されている可搬性記録媒体が前記ドライブに挿入されたときに、または、
前記第二の記憶領域の空き容量が所定以下の容量となったときに、前記第二の記憶領域に記憶されているチャンクと、前記ドライブに挿入された可搬性記録媒体に格納されているデータとのバイナリコンペアを行うよう指示する、
重複排除サーバ。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データの保管に必要な容量を低減する重複排除を行う情報処理システムに関する。

【背景技術】

20

【0002】

今日、企業や官公庁、学校などの様々な事業体において、電子ドキュメント（電子ファイル）や電子メールなど、様々な形態の電子データを利用して業務が行われている。インターネットの普及や紙文書の電子データへの移行などを背景として、これら電子データのデータ量は急激に増大しつつある。また、電子データの長期保管に関することを義務づける法律の制定や、各種電子データを長期間蓄積しビジネスに利活用しようとする動きなどに伴い、長期間記憶媒体に保管（アーカイブ）すべき電子データの増大しつつある。事業体は限られた予算で、これら急増する電子データを長期間保管（アーカイブ）する必要があるため、より低コストで記憶媒体に電子データを保管することが求められるようになっている。

30

【0003】

低コストでの電子データ保管に関して、例えば、特許文献 1（米国特許 6704730 号）では、重複排除（Deduplication）と呼ばれる、コンテンツ（ファイルや電子メールメッセージなど）の保管に必要な HDD（ハードディスクドライブ）の容量を削減できる技術が開示されている。具体的には、本技術においては、コンテンツを複数のバイト列（各バイト列はチャンクなどと呼ぶ）に分割し、すでに記憶装置に格納されたチャンクと完全に一致するチャンクである重複チャンクを記憶装置に格納せずに破棄し、重複チャンク以外のチャンクのみを記憶装置に格納する。さらに、破棄したチャンクを含め、複数のチャンクから構成されるコンテンツを管理するデータ構造を保持することで、重複チャンクを破棄した後でも、記憶装置に格納された、重複チャンク以外のチャンクからコンテンツを完全に復元することができる。なお、同技術では、チャンクのハッシュ値を装置に保存しておき、比較対象の 2 つのチャンクのハッシュ値の不一致を判定することにより、2 つのチャンクの不一致の判定を高速に行っている。2 つのチャンクのハッシュ値が一致した場合であっても、2 つのチャンクの内容が完全に一致するとは限らないため、チャンクを構成するバイト列同士の比較（バイナリコンペアとも呼ぶ）を行うことにより、2 つのチャンクの内容が完全に一致するかどうかを確認している。

40

【0004】

一方、電子データを保管する記憶媒体としては、HDD（ハードディスクドライブ。磁気ディスクドライブ）、テープ、光ディスクなどが挙げられる。HDD を記憶媒体として用いる記憶装置としては、アレイ状に構成した HDD を含むディスクアレイ装置がある。

50

テープや光ディスクを記憶媒体として用いる記憶装置として、テープライブラリや光ディスクライブラリがある。テープライブラリや光ディスクライブラリは、ドライブ（テープドライブや光ディスクドライブ）と、記憶媒体を物理的に格納する、ドライブとは物理的に隔離されたスロットと、記憶媒体をスロットとドライブ間で物理的に移動する物理的搬送機構であるロボットと、から構成される。記憶媒体に対する読み書きは記憶媒体をドライブにロードした後に行われる。なお、通常、スロットの数はドライブの数より十分に多い。

【 0 0 0 5 】

電子データを低コストで保管する場合、ビット単価がディスクアレイ装置に比べて低い、テープライブラリや光ディスクライブラリが有効である。しかし、テープライブラリや光ディスクライブラリのランダムアクセス性能（スループット性能、レスポンス性能）は、記憶媒体の物理的な移動を伴う場合があるため、記憶媒体の物理的な移動が一切ないディスクアレイ装置のそれに比べて格段に低い。

10

【先行技術文献】

【特許文献】

【 0 0 0 6 】

【特許文献 1】米国特許 6 7 0 4 7 3 0 号

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 7 】

特許文献 1 における技術は、ランダムアクセス性能の高い HDD を記憶媒体として用いる記憶装置（ディスクアレイ装置など）を前提としていた。特許文献 1 における記憶装置に対して重複排除技術を適用する場合、高頻度でバイナリコンペアを行う必要が生じ、記憶装置に対して多数のランダムアクセスが発生するが、記憶装置のランダムアクセス性能が高いために大きな問題とはなっていなかった。しかし、テープライブラリや光ディスクライブラリに対して同様の技術を適用する場合、そのランダムアクセス性能の低さから、重複排除処理に多大な時間を要し、現実的な時間内に重複排除処理が完了しないという問題があった。

20

【 0 0 0 8 】

また、特許文献 1 における技術においては、チャンク群のハッシュ値を管理するテーブルをシステム内に保管し、重複排除処理の際に参照する必要があるが、コンテンツの量が多い場合、同テーブルのサイズが大きくなるため、同テーブルを容量の限られた一次記憶（DRAM などのメモリ）内に格納することができなくなる。この場合、同テーブルは一次記憶ではなく、一時記憶より格段に容量の大きい、HDD などの二次記憶内に格納することになるが、二次記憶は一次記憶に比べてアクセス性能が低いため、同テーブルに対する参照のオーバーヘッドが大きくなり、重複排除処理の処理時間が長くなるという問題があった。

30

【 0 0 0 9 】

本発明はこれらの背景を鑑みてなされたもので、テープライブラリや光ディスクライブラリのような、記憶媒体内のデータへのアクセス時に記憶媒体の物理的な移動を伴う場合があるような記憶装置において、コンテンツ群に対して重複排除を適用し、コンテンツのデータを記憶装置に格納する際のオーバーヘッドを低減することを可能とする情報処理システムを提供することを一つの目的とする。

40

【課題を解決するための手段】

【 0 0 1 0 】

情報処理システムは、上位装置と、重複排除サーバと、光ディスクライブラリと、を備えて構成され、それらは互いに通信可能なように構成されている。重複排除サーバは、第一の記憶領域と書き込みバッファ領域を有する。光ディスクライブラリは、ロボットと、複数の光ディスクドライブと、複数のスロットと、複数の光ディスクと、を有する。書き込みバッファ領域と光ディスクを合わせて、第二の記憶領域と呼ぶ。重複排除サーバは、第

50

一の記憶領域に格納されたチャンクを管理する第一のチャンク管理テーブルと、第二の記憶領域に格納されたチャンクを管理する第二のチャンク管理テーブルを保持する。第一および第二のチャンク管理テーブルにはそれぞれ、第一および第二の記憶領域に格納されたチャンクのハッシュ値と格納位置の対が記録されている。第一および第二のチャンク管理テーブルはそれぞれ後述する第一および第二の重複排除を行う際に参照される。

【0011】

前述の目的は、重複排除サーバが以下で述べるようなコンテンツ格納処理部とコンテンツ読み出し処理部を有することにより達成される。コンテンツ格納処理部では、重複排除サーバは、上位装置からのコンテンツを受信し、コンテンツをチャンクに分割し、各チャンクを重複排除サーバもしくは光ディスクライブラリに格納する。コンテンツ読み出し処理部では、重複排除サーバは、上位装置からコンテンツの読み出し要求を受信し、必要なチャンクを重複排除サーバもしくは光ディスクライブラリから読み出し、コンテンツを再構築して上位装置に送信する。

10

【0012】

コンテンツ格納処理部は以下のステップ(X1)~(X11)を有することを特徴とする。なお、(X1)~(X6)は第一の重複排除のステップであり、(X7)~(X11)は第二の重複排除の一部のステップである。

【0013】

(X1)：重複排除サーバは、処理対象のチャンクのハッシュ値が第一のチャンク管理テーブルに登録済みかを判定する。

20

【0014】

(X2)：重複排除サーバは、(X1)での判定の結果、登録済みであれば、(X4)に進む。

【0015】

(X3)：重複排除サーバは、(X1)での判定の結果、登録済みでなければ、(X7)に進む。

【0016】

(X4)：重複排除サーバは、(X1)で登録済みであると判定したハッシュ値に対応するチャンクと、処理対象のチャンクとをバイナリコンペアする処理である第一のバイナリコンペアを実行する。

30

【0017】

(X5)：重複排除サーバは、(X4)で第一のバイナリコンペアの結果が一致であれば、処理対象のチャンクを破棄して、次のチャンクの処理に移る。

【0018】

(X6)：(X4)で第一のバイナリコンペアの結果が不一致であれば、(X7)に進む。

【0019】

(X7)：重複排除サーバは、処理対象のチャンクのハッシュ値が第二のチャンク管理テーブルに登録済みかを判定する。

【0020】

(X8)：重複排除サーバは、(X7)での判定の結果、登録済みであれば、(X10)に進む。

40

【0021】

(X9)：重複排除サーバは、(X7)での判定の結果、登録済みでなければ、(X11)に進む。

【0022】

(X10)：重複排除サーバは、処理対象のチャンクを第一の記憶領域に格納し、また、(X7)で登録済みであると判定したハッシュ値に対応するチャンクと、処理対象のチャンクとをバイナリコンペアする処理である第二のバイナリコンペアを実行する予定である旨をバイナリコンペア予定チャンク管理テーブルに登録する。ここでは、説明を簡単化

50

するため、(X7)で登録済みであると判定したハッシュ値に対応するチャンクは光ディスクドライブに未ロードの光ディスク内に格納されているものとする。

【0023】

(X11)：重複排除サーバは、処理対象のチャンクを書き込みバッファ領域に格納し、処理対象のチャンクの光ディスクへの書き込みを実行する予定である旨を、書き込み予定チャンク管理テーブルに登録する。ここでは、説明を簡単化するため、処理対象のチャンクの書き込み先の光ディスクは光ディスクドライブに未ロードの光ディスクであるものとする。

【0024】

コンテンツ格納処理部は、第二のバイナリコンペアと書き込みを先送りすることを特徴とする。 10

【0025】

コンテンツ読み出し処理部は、以下のステップ(Y1)～(Y7)を有することを特徴とする。なお、(Y4)～(Y7)は第二の重複排除の一部のステップである。

【0026】

(Y1)：重複排除サーバは、上位装置よりコンテンツ読み出し要求を受信する。

【0027】

(Y2)：重複排除サーバは、コンテンツをチャンクに分割し、必要なチャンクを特定する。

【0028】

(Y3)：重複排除サーバは、必要なチャンクの読み出しに必要な光ディスクを光ディスクドライブにロードし、必要なチャンクを第一の記憶領域もしくは第二の記憶領域から読み出して、コンテンツを再構築して、再構築したコンテンツを上位装置に送信する。 20

【0029】

(Y4)：重複排除サーバは、(Y3)でロードされた光ディスクに関連する、バイナリコンペア予定チャンク管理テーブルに登録された第二のバイナリコンペアを実行する。

【0030】

(Y5)：(Y4)の第二のバイナリコンペアの結果が一致であれば、第二のバイナリコンペアの対象の第一の記憶領域内のチャンクを破棄する。

【0031】

(Y6)：(Y4)の第二のバイナリコンペアの結果が不一致であれば、第二のバイナリコンペアの対象の第一の記憶領域内のチャンクを書き込みバッファ領域に移動し、移動したチャンクの光ディスクへの書き込みを実行する予定である旨を、書き込み予定チャンク管理テーブルに登録する。ここでは、説明を簡単化するため、処理対象のチャンクの書き込み先の光ディスクは光ディスクドライブに未ロードの光ディスクであるものとする。 30

【0032】

(Y7)：重複排除サーバは、(Y3)でロードされた光ディスクに関連する、書き込み予定チャンク管理テーブルに登録された書き込みを実行する。

【0033】

コンテンツ読み出し処理部は、コンテンツ格納処理部にて先送りされた第二のバイナリコンペアおよび書き込みを実行することを特徴とする。 40

【0034】

その他、本願が開示する課題、及びその解決方法は、発明を実施するための形態の欄の記載、及び図面の記載等により明らかにされる。

【発明の効果】

【0035】

本発明によれば、第二の重複排除に先立って、第一の重複排除を行うことにより、第二の重複排除の対象となるチャンクを削減し、重複排除を効率的に行うことが可能になる。また、第二のバイナリコンペアおよび光ディスクへの書き込みを、コンテンツ読み出し処理部の実行時まで先送りすることにより、光ディスクの光ディスクドライブへのロード回 50

数を削減することが可能になり、ランダムアクセス性能が低い光ディスクライブラリにおいて効率的に第二のバイナリコンペアおよび光ディスクへの書き込みを行うことが可能となる。また、本発明によれば、光ディスクの光ディスクドライブへのロード回数削減により、光ディスクライブラリ内のロボットの動作回数を低減させ、ロボットを長寿命化することが可能になる。

【図面の簡単な説明】

【0036】

【図1】本発明の第一の実施形態における情報処理システムの全体構成を示す図である。

【図2】本発明におけるコンテンツとチャンクの関係を示す図である。

【図3】本発明の実施形態のアーカイブサーバのメモリの構成例を示す図である。

10

【図4】本発明の実施形態の重複排除サーバのメモリおよびHDDの構成例を示す図である。

【図5】本発明におけるチャンクファミリーの概念図を示す図である。

【図6】本発明の実施形態におけるファイル管理テーブルの構成例を示す図である。

【図7】本発明の実施形態における光ディスク管理テーブルの構成例を示す図である。

【図8】本発明の実施形態における光ディスクドライブ管理テーブルの構成例を示す図である。

【図9】本発明の実施形態におけるコンテンツ管理テーブルの構成例を示す図である。

【図10】本発明の実施形態における第一のチャンク管理テーブルの構成例を示す図である。

20

【図11】本発明の実施形態における第二のチャンク管理テーブルの構成例を示す図である。

【図12】本発明の実施形態におけるチャンクファミリー管理テーブルの構成例を示す図である。

【図13】本発明の実施形態におけるバイナリコンペア予定チャンク管理テーブルの構成例を示す図である。

【図14】本発明の実施形態における書き込み予定チャンク管理テーブルの構成例を示す図である。

【図15】本発明の実施形態における統計情報管理テーブルの構成例を示す図である。

【図16】本発明の実施形態における管理サーバのディスプレイの表示例を示す図である。

30

【図17】本発明の実施形態におけるコンテンツ格納処理のフローチャートを示す図である。

【図18】本発明の実施形態におけるコンテンツ格納処理の部分Bのフローチャートを示す図である。

【図19】本発明の実施形態におけるコンテンツ格納処理の部分Cのフローチャートを示す図である。

【図20】本発明の実施形態におけるコンテンツ読み出し処理のフローチャートを示す図である。

【図21】本発明の実施形態におけるバイナリコンペア処理のフローチャートを示す図である。

40

【図22】本発明の実施形態におけるバイナリコンペア処理の部分Dのフローチャートを示す図である。

【図23】本発明の実施形態における書き込み処理のフローチャートを示す図である。

【図24】本発明の実施形態における強制バイナリコンペア処理のフローチャートを示す図である。

【図25】本発明の実施形態における強制書き込み処理のフローチャートを示す図である。

【図26】本発明の第二の実施形態における情報処理システムの全体構成を示す図である。

50

【図 27】本発明の第二の実施形態における、ファイルストレージのコントローラのメモリ、および、ファイルストレージの HDD の構成例を示す図である。

【図 28】本発明の第三の実施形態における情報処理システムの全体構成を示す図である。

【図 29】本発明の第三の実施形態におけるアーカイブサーバのメモリの構成例を示す図である。

【図 30】従来の重複排除処理の概念図を示す図である。

【図 31】本発明の実施形態における重複排除処理の概念図を示す図である。

【発明を実施するための形態】

【0037】

10

以下、本発明の実施の形態について、図面を参照しながら説明する。なお実施の形態を説明するための図面において、同一部には原則として同一符号を付し、その繰り返しの説明は省略する。また説明上プログラムまたは機能が主語となる場合、実際にはプログラムを実行するプロセッサや回路によって処理が実行される。

【0038】

以下、第一の実施形態を説明する。図 1 ~ 図 25 を用いて第一の実施形態の情報処理システムについて説明する。また、本実施形態で用いるテーブルやプログラムの処理フローなどの詳細説明の前に、図 1 を用いて情報処理システム 1 の全体的な概略動作について説明すると共に、図 30、図 31 を用いて従来の重複排除処理と本実施形態における重複排除処理の概念処理について説明する。

20

【0039】

図 1 は、第一の実施形態における情報処理システム 1 を示している。情報処理システム 1 は、ファイル（コンテンツとも記す）を管理する機能などを有する。

【0040】

まず、情報処理システム 1 の構成要素について述べる。

【0041】

情報処理システム 1 は、重複排除サーバ 100、光ディスクライブラリ 200、ファイルストレージ 300、アーカイブサーバ 400、クライアント装置 500、管理サーバ 600 が、相互結合網 700a、700b を介して通信可能に接続されて構成される。相互結合網 700 は、例えば各種の LAN (Local Area Network) や WAN (Wide Area Network) 等により構成される通信網である。例えば、相互結合網 700 は、イーサネット（登録商標）を使い、通信プロトコルとして TCP/IP (Transmission Control Protocol / Internet Protocol) を利用することができる。

30

【0042】

次に、情報処理システム 1 の各構成要素について説明する。

【0043】

まず、アーカイブサーバ 400 について説明する。

【0044】

アーカイブサーバ 400 は、情報処理システム 1 が管理するコンテンツを統合的に制御するコンピュータである。アーカイブサーバ 400 は、相互結合網 700a を介して、重複排除サーバ 100、光ディスクライブラリ 200、ファイルストレージ 300、クライアント装置 500、管理サーバ 600 と通信可能に接続されている。

40

【0045】

情報処理システム 1 が管理するファイル（コンテンツ）の実体は、アーカイブサーバ 400 と通信可能に接続されたファイルストレージ 300 や光ディスクライブラリ 200 に含まれる記憶媒体に格納されているが、アーカイブサーバ 400 はそれらファイル（コンテンツ）の読み書きを中継する機能を担っている。また、アーカイブサーバ 400 は、コンテンツへのアクセス頻度や無アクセス期間、あるいは、クライアント装置 500 からのコンテンツへの読み出し要求などに応じて、ファイルストレージ 300 と光ディスクライ

50

ブラリ 200 の間でコンテンツを相互に移動あるいはコピーする機能も担っている。

【0046】

アーカイブサーバ 400 がファイルストレージ 300 内のコンテンツを読み書きする際に用いるネットワーク・プロトコルとしては、NFS (Network File System) や CIFS (Common Internet File System) を利用することができる。アーカイブサーバ 400 が重複排除サーバ 100 に対してコンテンツを読み書きする際に用いるネットワーク・プロトコルとしては、NFS や CIFS といった一般的なプロトコルの他に独自プロトコルを利用することができる。本実施形態では、アーカイブサーバ 400 が重複排除サーバ 100 に対してコンテンツを読み書きする際に、独自プロトコルを用いる例 (後述) を示すが、これに限られるものではない。

10

【0047】

アーカイブサーバ 400 は、メモリ 410 の他に、図示しない、CPU (Central Processing Unit)、HDD (Hard Disk Drive)、ネットワーク・インターフェースといった内部構成要素を有し、それら内部構成要素は内部バスを介して相互に接続されている。

【0048】

図 3 は、アーカイブサーバ 400 のメモリ 410 の内容を示している。メモリ 410 には、後述するファイル管理テーブル 1000 としたテーブル、および、マイグレーション処理 411 のプログラム、後述する読み出し要求処理 412 のプログラム、書き込み要求処理 413 のプログラムといった各種プログラムが記憶されている。メモリ 410 に格納された各種プログラムの動作については後述する。メモリ 410 に格納された各種プログラムは、図示しないアーカイブサーバ 400 の CPU に読み込まれて実行される。

20

【0049】

次に、図 1 に戻って、クライアント装置 500 について説明する。

【0050】

クライアント装置 500 は、アーカイブサーバ 400 により提供されるファイル読み書き機能を利用しながら業務を行う従業員等 (以下、ユーザとも記す) が使用するコンピュータである。クライアント装置 500 はアーカイブサーバ 400 と相互結合網 700b を介して通信可能に接続されている。

【0051】

クライアント装置 500 は、アーカイブサーバ 400 を介して、ファイルストレージ 300 が有する記憶媒体に格納されたコンテンツに対して読み書きを行ったり、光ディスクライブラリ 200 が有する記憶媒体に格納されたチャンク (後述) に対して読み込みを行ったりする。クライアント装置 500 がアーカイブサーバ 400 に対してコンテンツを読み書きする際に用いるネットワーク・プロトコルとしては、NFS や CIFS を利用することができる。

30

【0052】

クライアント装置 500 は、図示しない、CPU、メモリ、HDD、ネットワーク・インターフェースを内部構成要素として有し、それら内部構成要素は内部バスで接続されている。さらに、クライアント装置 500 には、ユーザ・インターフェースとして、図示しない、ディスプレイ、キーボード及びマウスが接続されている。クライアント装置 500 のメモリには、ネットワーク・プロトコルを介してアーカイブサーバ 400 にコンテンツを書き込むためのコンテンツ書き込みプログラム (図示しない)、ネットワーク・プロトコルを介してアーカイブサーバ 400 からコンテンツを読み出すためのコンテンツ読み出しプログラム (図示しない)、コンテンツの移動・複製・削除・ファイル名変更を行うプログラムなどが記憶されている。これら各種プログラムはクライアント装置 500 の CPU に読み出されて実行される。ユーザはユーザ・インターフェースを用いてクライアント装置 500 を操作し、アーカイブサーバ 400 に対してコンテンツの読み書きなどを行うことにより、各種業務を遂行する。

40

【0053】

50

次に、ファイルストレージ 300 について説明する。

【0054】

ファイルストレージ 300 は、ユーザが利用するコンテンツを主に格納するためのコンピュータである。なお、ファイルストレージ 300 は、重複排除サーバ 100 のための一時記憶領域としても用いることができる。重複排除サーバ 100 がどのようにファイルストレージ 300 を用いるかについては後述する。ファイルストレージ 300 は、相互結合網 700 a を介して、重複排除サーバ 100、アーカイブサーバ 400、管理サーバ 600 と通信可能に接続されている。重複排除サーバ 100、アーカイブサーバ 400 がファイルストレージ 300 内のコンテンツを読み書きする際に用いるネットワーク・プロトコルとしては、NFS や CIFS を利用することができる。

10

【0055】

ファイルストレージ 300 は、コントローラ 310 と HDD 320 を内部構成要素として有し、それら内部構成要素は内部バスで接続されている。コントローラ 310 は、図示しない、CPU、メモリ、HDD、ネットワーク・インターフェースを内部構成要素として有し、それら内部構成要素は内部バスで接続されている。ファイルストレージ 300 はファイルストレージ 300 のコントローラ 310 のネットワーク・インターフェースを介して相互結合網 700 a に接続されている。

【0056】

ファイルストレージ 300 のコントローラ 310 のメモリ内には、RAID (Redundant Array of Independent Disks) プログラム、スナップショット・プログラム、ファイルシステム、ネットワーク・プロトコル・プログラムなどといった各種プログラムが記憶されている。ファイルストレージ 300 のコントローラ 310 のメモリ内の各種プログラムはコントローラ 310 の図示しない CPU に読み込まれて実行される。

20

【0057】

RAID プログラムは、複数の HDD 320 の記憶領域から構成される論理的な記憶領域であるボリュームを提供し、ボリュームに対する読み書きをできるようにするプログラムである。RAID の技術は公知であるため詳細については説明しないが、本技術を利用することにより、高可用性かつ高性能な記憶領域を提供することが可能になる。スナップショット・プログラムはボリュームの物理的あるいは論理的な複製を作成する機能である。スナップショット・プログラムにより、ボリュームが物理的あるいは論理的に破損した場合でも、破損前の状態に復元することが可能になる。ファイルシステムは、オペレーティングシステムが提供する機能であり、ボリューム上に階層構造を持った記憶領域を構築することができる。ファイル(コンテンツ)はファイルシステム内で管理され、ファイルシステムに含まれるファイル(コンテンツ)は、ディレクトリ名とファイル名から構成されるファイルパスによって一意に識別することができる。ネットワーク・プロトコル・プログラムは、ファイル(コンテンツ)に対する相互結合網 700 a を介した読み書き機能を提供するプログラムである。これらファイルストレージ 300 内の各種プログラムにより、アーカイブサーバ 400 や重複排除サーバ 100 はファイルストレージ 300 内のファイル(コンテンツ)を、相互結合網 700 a を介して読み書きすることができる。

30

40

【0058】

次に、光ディスクライブラリ 200 について説明する。

【0059】

光ディスクライブラリ 200 は、内蔵する光ディスクにデータを保管する装置である。

【0060】

光ディスクライブラリ 200 は、コントローラ 210、光ディスクドライブ 220、スロット 230、光ディスク 240、ロボット 250 を内部構成要素として有する。コントローラ 210、光ディスクドライブ 220、ロボット 250 は内部バスで接続されている。光ディスクドライブ 220、スロット 230、光ディスク 240 の数は 1 または複数である。また、コントローラ 210 は性能向上や可用性向上のため、冗長化されていてもよ

50

い。

【0061】

光ディスクライブラリ200のコントローラ210は、図示しない、メモリ、CPUおよびHDDを内部構成要素として有し、それら内部構成要素は内部バスで接続されている。コントローラ210のメモリには、後述するロードの指示を受け付け、ロードを行うプログラムや、後述するアンロードの指示を受け付け、アンロードを行うプログラムや、データの書き込みや読み出しの要求を受け付け、データの書き込みや読み出しを行うプログラムなどが記憶される。これら各種プログラムはコントローラ210のCPUに読み出されて実行される。

【0062】

光ディスク240は記憶媒体の一種であり、可搬媒体（リムーバブル・メディア）であることを特徴とする。スロット230は、光ディスク240を格納する物理的な棚である。光ディスクドライブ220は光ディスク240を内部にロードし、ロードした光ディスク240に対してデータを読み書きするデバイスである。光ディスクドライブ220とスロット230は物理的に隔離されている。ロボット250は光ディスク240をスロット230と光ディスクドライブ220の間で物理的に移動する搬送機構である。

【0063】

光ディスクライブラリ200は、重複排除サーバ100に対して、光ディスク240の光ディスクドライブ220へのロード/アンロード機能や、相互結合網700aを介した、光ディスク240に対するデータの読み書き機能を提供する。

【0064】

光ディスクライブラリ200は、重複排除サーバ100からの光ディスク240のロード/アンロード要求に応じて、指定された光ディスク240を光ディスクドライブ220にロード（Insert）したり、光ディスク240を光ディスクドライブ220からアンロード（Eject）したりする。

【0065】

また、光ディスクライブラリ200は、重複排除サーバ100からのデータの書き込み要求に応じて、重複排除サーバ100から送信されたデータを光ディスクに書き込んだり、重複排除サーバ100からのデータの読み出し要求に応じて、光ディスクからデータを読み出して重複排除サーバ100に送信したりする。

【0066】

重複排除サーバ100が、相互結合網700aを介して光ディスクライブラリ200に対してデータを読み書きする際、NFSやCIFSのような一般的なネットワーク・プロトコルを用いることもできるが、本実施形態では独自のネットワーク・プロトコルを用いる例を示している。本実施形態では、相互結合網700aを介して光ディスクライブラリ200に対してデータを書き込む際、書き込み先の光ディスクドライブ220を明示的に指定して、データを書き込む。また、相互結合網700aを介して光ディスクライブラリ200に対してデータを読み込む際、読み込み元の光ディスクドライブ220を明示的に指定して、データを読み込む。

【0067】

重複排除サーバ100が、光ディスクライブラリ200に対して光ディスク240の光ディスクドライブ220へのロードを指示する場合、重複排除サーバ100は、ロード対象の光ディスク240を光ディスクライブラリ200内で一意に識別する識別子と、ロード先の光ディスクドライブ220を光ディスクライブラリ200内で一意に識別する識別子を指定する。これにより、指定された光ディスク240はロボット250により指定された光ディスクドライブ220に搬送され、その光ディスクドライブ220にロードされる。空き状態の光ディスクドライブが存在しない場合、使用中でない（後述する「Unlocked」状態の）光ディスクドライブ240内の光ディスク240をアンロードすることで、光ディスクドライブ220を空き状態にした後に、ロードを行う。

【0068】

10

20

30

40

50

重複排除サーバ100が、光ディスクライブラリ200に対して光ディスク240の光ディスクドライブ220からのアンロードを指示する場合、重複排除サーバ100は、アンロード対象の光ディスク240を光ディスクライブラリ200内で一意に識別する識別子を指定する。指定された光ディスク240は光ディスクドライブ220から排出され、ロボット250により元のスロット230に移動される。

【0069】

ここで、光ディスクおよびそれに格納されるデータについて述べる。

【0070】

光ディスクの規格としては、ライト・ワンス（書き換え不可）を特徴とするDVD-R、BD-Rなどといった規格や、リライタブル（書き換え可能）を特徴とするDVD-RW、BD-REなどといった規格が知られている。なお、DVD、BDはそれぞれDigital Versatile Disk、Blu-ray Diskの略である。ライト・ワンスの光ディスクであっても追記書き込み機能をサポートしている。本実施形態においては、いずれの規格を用いるかは規定しない。

10

【0071】

DVDやBDといった光ディスクはUDF（ユニバーサルディスクフォーマット）と呼ばれる光ディスク用のファイルシステムを採用している。リライタブルな光ディスクを用いる場合、光ディスク内のデータをブロック単位で更新し、ファイルシステムを変更することができ、ファイルの追加、更新、削除を行うことが可能である。ライト・ワンスの光ディスクを用いる場合であっても、一旦書き込んだメタデータおよびデータを物理的に消去することはできないものの、追記書き込みを行い、ファイルシステムを論理的に変更することにより、ファイルの追加、論理的な更新、論理的な削除を行うことが可能である。

20

【0072】

光ディスク内に含まれるファイルは、ディレクトリ名とファイル名から構成されるファイルパスで一意に指定することができる。光ディスク240を複数含む光ディスクライブラリ200においては、光ディスクライブラリ200の光ディスク240を一意に識別する識別子と、ファイルパスに用いることにより、光ディスクライブラリ200に含まれるファイルを一意に指定することができる。

【0073】

次に、重複排除サーバ100について説明する。

30

重複排除サーバ100は、アーカイブサーバ400に対して相互結合網700aを介してのコンテンツの読み書き機能を提供し、アーカイブサーバ400から書込まれたコンテンツを光ディスクライブラリ200に書き込んだり、アーカイブサーバ400から読み込み要求のあったコンテンツを光ディスクライブラリ200から読み込んだりするコンピュータである。重複排除サーバ100は、コンテンツを光ディスクライブラリ200に書き込む際に、前述した重複排除を用いて、コンテンツを複数のバイト列（チャンクと呼ぶ）に分割し、重複チャンク、即ち、すでに光ディスクライブラリ200に格納されたチャンクと完全に一致するチャンクを光ディスクライブラリ200に格納せずに破棄し、重複チャンク以外のチャンクのみを光ディスクライブラリ200に格納する。図2には、コンテンツ2が1または複数のチャンク3に分割される様子を図示している。なお、チャンクのサイズは固定長とすることもできるし、可変長とすることもできる。二つのチャンクが完全に一致する可能性があるか判定するために、チャンクのハッシュ値を用いることができる。例えば、MD5（Message Digest 5）やSHA-1（Secure Hash Algorithm 1）といった公知のハッシュ関数を用いて、チャンクのデータからハッシュ値を算出することができる。例えば、MD5、SHA-1であれば、ハッシュ値はそれぞれ128ビット、160ビットのビット列である。ハッシュ値が一致する二つのチャンクは、そのデータのバイト列が一致する可能性がある。二つのチャンクのハッシュ値が一致する場合であっても、それらのバイト列は一致しない場合もある。算出したチャンクのハッシュ値は、例えば、重複排除サーバ100のメモリ110に記録しておき、後でハッシュ値の比較に用いることができる。

40

50

【 0 0 7 4 】

重複排除技術を用いることによりコンテンツの保管に必要な記憶領域の容量を削減することができる。重複排除の概略動作の説明は後述する。

【 0 0 7 5 】

前述したように、本実施形態では、アーカイブサーバ400が重複排除サーバ100に対してコンテンツを格納したり、重複排除サーバ100からコンテンツを読み出したりする際に、独自プロトコルを用いる例を示すが、これに限定されるものではない。アーカイブサーバ400が重複排除サーバ100に対してコンテンツを格納する際の処理は、後述するコンテンツ格納処理2000で詳細に示す。アーカイブサーバ400が重複排除サーバ100からコンテンツを読み出す際の処理は、後述するコンテンツ読み出し処理2100で詳細に示す。

10

【 0 0 7 6 】

重複排除サーバ100は、チャンクを光ディスクライブラリ200に書き込む際、光ディスクライブラリ200に対して、書き込み先の光ディスク240の識別子を指定して光ディスク240のロードを指示し、その後、ロードされた光ディスク240に対して、当該チャンクを光ディスクライブラリ200内で一意に識別できる識別子をファイルパスとして指定して当該チャンクを書き込む。

【 0 0 7 7 】

重複排除サーバ100は、チャンクを光ディスクライブラリ200から読み出す際、光ディスクライブラリ200に対して、読み込み元の光ディスク240の識別子を指定して光ディスク240のロードを指示し、その後、ロードされた光ディスク240に対して、当該チャンクを光ディスクライブラリ200内で一意に識別できる識別子をファイルパスとして指定して当該チャンクを読み出す。

20

【 0 0 7 8 】

重複排除サーバ100は、メモリ110、HDD120の他に、図示しない、CPU、ネットワーク・インターフェースを内部構成要素として有し、それら内部構成要素は内部バスで接続されている。重複排除サーバ100は重複排除サーバ100のネットワーク・インターフェースを介して相互結合網700aに接続されている。

【 0 0 7 9 】

図4は、重複排除サーバ100のメモリ110、HDD120の内容を示している。メモリ110には、後述する、光ディスク管理テーブル1100、光ディスクドライブ管理テーブル1200、コンテンツ管理テーブル1300、第一のチャンク管理テーブル1400、チャンクファミリー管理テーブル1600、バイナリコンペア予定チャンク管理テーブル1700、書き込み予定チャンク管理テーブル1800、統計情報管理テーブル1900といったテーブル、および、後述する、コンテンツ格納処理2000、コンテンツ読み出し処理2100、バイナリコンペア処理2200、書き込み処理2300、強制バイナリコンペア処理2400、強制書き込み処理2500といったプログラムが記憶されている。メモリ110に格納された各種プログラムは、図示しない重複排除サーバ100のCPUに読み込まれて実行される。これら各種プログラムの動作については後述する。メモリ110に格納された各種テーブルは、各種プログラムの実行時に重複排除サーバ100のCPUにより参照・更新される。これら各種テーブルの詳細については後述する。なお、本実施形態では、これら各種テーブルをメモリ110に記憶する例を示したが、これら各種テーブルは重複排除サーバ100のHDD120、ファイルストレージ300のコントローラ310のメモリ、ファイルストレージ300のHDD320、光ディスクライブラリ200のコントローラ210のメモリあるいはHDD、光ディスクライブラリ200の光ディスク240、のいずれかに記憶するように構成してもよい。

30

40

【 0 0 8 0 】

HDD120には、後述する、第二のチャンク管理テーブル1500というテーブルが記憶されている。本実施形態では、このテーブルをHDD120に格納する例を示したが、このテーブルは重複排除サーバ100のメモリ110、ファイルストレージ300のコ

50

ントローラ 3 1 0 のメモリ，ファイルストレージ 3 0 0 の HDD 3 2 0 ，光ディスクライブラリ 2 0 0 のコントローラ 2 1 0 のメモリあるいは HDD ，光ディスクライブラリ 2 0 0 の光ディスク 2 4 0 ，のいずれかに記憶するように構成してもよい。

【 0 0 8 1 】

次に，図 1 に戻って，管理サーバ 6 0 0 について説明する。

【 0 0 8 2 】

管理サーバ 6 0 0 は，重複排除サーバ 1 0 0 ，光ディスクライブラリ 2 0 0 ，ファイルストレージ 3 0 0 ，アーカイブサーバ 4 0 0 を管理するためのコンピュータである。情報処理システム 1 の管理者は，管理サーバ 6 0 0 を用いることにより重複排除サーバ 1 0 0 ，光ディスクライブラリ 2 0 0 ，ファイルストレージ 3 0 0 ，アーカイブサーバ 4 0 0 に関する各種情報を参照したり，設定を変更したりすることができる。

10

【 0 0 8 3 】

管理サーバ 6 0 0 は，図示しない，CPU ，メモリ，HDD ，ネットワーク・インターフェースを内部構成要素として有し，それら内部構成要素は内部バスで接続されている。さらに，管理サーバ 6 0 0 には，ユーザ・インターフェースとして，図示しない，ディスプレイ，キーボード及びマウスが接続されている。管理者はそれらを用いて，重複排除サーバ 1 0 0 ，光ディスクライブラリ 2 0 0 ，ファイルストレージ 3 0 0 ，アーカイブサーバ 4 0 0 の管理を行う。

【 0 0 8 4 】

管理サーバ 6 0 0 に用いることで，どのような情報を参照できるかの例については後述する。

20

【 0 0 8 5 】

次に，情報処理システム 1 内の記憶領域について述べる。

【 0 0 8 6 】

ファイルストレージ 3 0 0 には，プライマリ記憶領域 1 0 ，第一の記憶領域 2 0 ，書き込みバッファ領域 4 0 といった，複数のファイルシステム領域が含まれる。プライマリ記憶領域 1 0 はアーカイブサーバ 4 0 0 が利用するファイルシステム領域である。第一の記憶領域 2 0 および書き込みバッファ領域 4 0 は重複排除サーバ 1 0 0 が利用するファイルシステム領域である。ファイルストレージ 3 0 0 の書き込みバッファ領域 4 0 と光ディスク 2 4 0 全ての記憶領域を合わせた記憶領域を第二の記憶領域と呼ぶ。書き込みバッファ領域 4 0 は，光ディスク 2 4 0 に書き込む前に一時的にデータ（チャンク）を保管するために用いられる。第一の記憶領域 2 0 ，書き込みバッファ領域 4 0 は HDD 3 2 0 上に格納される。

30

【 0 0 8 7 】

なお，本実施形態では，第一の記憶領域 2 0 および書き込みバッファ領域 4 0 を HDD 3 2 0 上に格納する例を示したが，これらの記憶領域は光ディスク 2 4 0 上に格納するように構成してもよい。ただし，この場合，ランダムアクセス性能を考慮して，光ディスク 2 4 0 を光ディスクドライブ 2 2 0 に常駐させる対応が必要となる。

【 0 0 8 8 】

あるいはまた，第一の記憶領域 2 0 および書き込みバッファ領域 4 0 を重複排除サーバ 1 0 0 の HDD 1 2 0 に格納するように構成してもよい。

40

【 0 0 8 9 】

以上が第一の実施形態における情報処理システム 1 の説明である。

【 0 0 9 0 】

次に，再び図 1 を参照しつつ，本実施形態の情報処理システム 1 の全体的な概略動作について述べる。なお，ここでは，概略動作の説明を簡単化するため，基本的な動作についてのみ述べる。

【 0 0 9 1 】

まず，クライアント装置 5 0 0 がコンテンツをアーカイブサーバ 4 0 0 に書き込んだ後に，そのコンテンツが最終的に第二の記憶領域 3 0 もしくは第一の記憶領域 2 0 に格納さ

50

れるまでの概略動作について述べる。

【0092】

クライアント装置500はまず、ユーザの指示に従い、クライアント装置500のメモリに記憶されたコンテンツ書き込みプログラムを用いて、コンテンツをアーカイブサーバ400に書き込む。アーカイブサーバ400はクライアント装置500から書き込まれたコンテンツを、書き込み要求処理413のプログラムを用いて、ファイルストレージ300のプライマリ記憶領域10に書き込む。その後、クライアント装置500は、ユーザの指示に従い、アーカイブサーバ400を介して、当該コンテンツに対して読み書きを行う。

【0093】

次に、当該コンテンツはアーカイブサーバ400のメモリ410に記憶されたマイグレーション処理411のプログラムにより、重複排除サーバ100に書き込まれる。

【0094】

ここで、マイグレーション、および、マイグレーション処理411のプログラムについて簡単に説明する。マイグレーションとは、性質の異なる記憶階層間でコンテンツを移動することである。記憶階層は、例えば、ビット単価が高いが、高性能である記憶媒体（例えば、ファイルストレージ300のHDD320）から構成される記憶領域である上位階層と、ビット単価の低いが、低性能である記憶媒体（例えば、光ディスクライブラリ200の光ディスク240）から構成される記憶領域である下位階層を含む。一般に、時間経過に伴い書き込まれたコンテンツが古くなっていくと、次第にそのコンテンツに対するアクセス頻度は低下していく。そのようなアクセス頻度の低いコンテンツは上位階層から下位階層に移動（マイグレーション）することにより、コンテンツの保管コストを低減することができる。なお、下位階層に格納されたコンテンツに対するアクセス頻度が高くなった場合、そのコンテンツは下位階層から上位階層に移動されることもある。また、コンテンツへのアクセス性能を向上するために、コンテンツの本体を下位階層に保管し、コンテンツの複製を上位階層に保管することもできる。さらにまた、記憶階層は2つ以上存在する場合もある。

【0095】

マイグレーション処理411のプログラムは、定期的（例えば一日毎など）に起動され、予め管理者により管理サーバ600を用いて設定されたポリシーに従ってマイグレーション対象のコンテンツを選択し、マイグレーションを行う。ポリシーには、マイグレーション対象とするコンテンツの選択方法が指定され、例えば、アクセス頻度がある設定された閾値以下であるコンテンツや、無アクセス期間がある設定された閾値以上のコンテンツを上位階層から下位階層へのマイグレーション対象として選択したり、逆に、アクセス頻度がある設定された閾値以上であるコンテンツを下位階層から上位階層へのマイグレーション対象として選択したりする、コンテンツの選択方法が指定される。

【0096】

アーカイブサーバ400は、マイグレーションされたコンテンツがファイルストレージ300と重複排除サーバ100のどちらにアクセスすれば読み出すことができるかを、後述するファイル管理テーブル1000に記録している。

【0097】

以上がマイグレーション、および、マイグレーション処理411のプログラムについての説明であり、再び概略動作の説明に戻る。

【0098】

マイグレーション処理411のプログラムにより重複排除サーバ100に書き込まれたコンテンツに含まれるチャンクは、重複排除サーバ100のメモリ110に記憶されたコンテンツ格納処理2000などにより、重複排除処理が行われ、最終的に第一の記憶領域20あるいは第二の記憶領域30に格納される。重複排除処理の概念動作については、後ほど図31を参照しながら説明する。記憶されたコンテンツ格納処理2000などの重複排除サーバ100のメモリ110に格納された、コンテンツの格納に関連する各種プログ

10

20

30

40

50

ラムの動作に関しては後述する。第一の記憶領域 20 に格納されたチャンクは後述する第一のチャンク管理テーブル 1400 で管理され、また、第二の記憶領域 30 に格納されたチャンクは後述する第二のチャンク管理テーブル 1500 で管理される。

【0099】

以上が、クライアント装置 500 がコンテンツをアーカイブサーバ 400 に書き込んだ後に、そのコンテンツが最終的に第二の記憶領域 30 もしくは第一の記憶領域 20 に格納されるまでの概略動作についての説明である。

【0100】

ところで、第二の記憶領域 30 には複数の光ディスク 240 が含まれるが、第二の記憶領域 30 の光ディスク 240 にコンテンツを格納する際、コンテンツを構成するチャンクをどの光ディスク 240 に格納するかを選択方法について、図 5 を参照しながら補足説明する。

10

【0101】

図 5 はこれから定義するチャンクファミリーの概念図を示している。図 2 で示したようにコンテンツ 2 は 1 または複数のチャンク 3 から構成される。しかし、チャンク 3 は重複排除されるため、重複排除後にはコンテンツ 2 は、チャンク 3 に重複排除を適用した後のチャンクである重複排除済みチャンク 4 から構成されることになる。図 5 の例では、コンテンツ 2 a は重複排除済みチャンク 4 p、4 q から構成されている。また、コンテンツ 2 b は重複排除済みチャンク 4 q、4 r から構成されている。この例では、重複排除済みチャンク 4 q は、コンテンツ 2 a、2 b に共通して含まれるチャンクである。重複排除済みチャンク 4 q に対応する、コンテンツ 2 a のチャンク 3 とコンテンツ 2 b のチャンク 3 のうち一方は、データのバイト列が一致するため、重複排除により削除されている。なお、以降の図 5 以外の説明においては、チャンク 3 および重複排除済みチャンク 4 を単にチャンクと呼ぶことがある。

20

【0102】

ここで以下のような定義をする。即ち、(1) コンテンツ 2 を親とし、そのコンテンツ 2 に含まれる重複排除済みチャンク 4 をその親の子とし、(2) 共通の親 (コンテンツ 2) を持つ複数の重複排除済みチャンク 4 は兄弟であるとし、(3) 兄弟の兄弟もまた兄弟であるとし、(4) 兄弟関係にある重複排除済みチャンク 4 群をチャンクファミリー 5 であるとする。つまり、チャンクファミリー 5 はコンテンツ 2 を介して互いに関連する重複排除済みチャンク 4 の集合体である。

30

【0103】

コンテンツを構成する重複排除済みチャンク 4 を光ディスク 240 に格納する際には、このチャンクファミリーの考え方をを用いる。即ち、チャンクファミリー 5 に含まれる重複排除済みチャンク 4 群は 1 枚の光ディスク 240 に格納する。チャンクファミリー 5 に含まれる重複排除済みチャンク 4 群の総容量が大きく、1 枚の光ディスク 240 に収まらない場合は、チャンクファミリー 5 を分割して、複数の光ディスク 240 に格納する。1 枚の光ディスク 240 に複数のチャンクファミリー 5 が含まれていてもよいものとする。

【0104】

このようなチャンクファミリーの考え方をを用いてチャンクを光ディスク 240 に格納することにより、コンテンツ 2 にアクセスする際にアクセスする必要のある光ディスク 240 の枚数を低減することができる。例えば、チャンクファミリーがある 1 枚の光ディスク 240 に格納されていることを想定すると、そのチャンクファミリーに関連するいずれのコンテンツにアクセスする場合であっても、当該の 1 枚の光ディスク 240 にのみアクセスすればよい。そのため、光ディスク 240 をスロット 230 から光ディスクドライブ 220 に移動してロードする回数を最小限に抑えることができる。

40

【0105】

以上がチャンクをどの光ディスク 240 に格納するかを選択方法についての補足説明である。

【0106】

50

次に、クライアント装置 500 が、アーカイブサーバ 400 に書き込んだコンテンツを読み出す際の概略動作について説明する。

【0107】

クライアント装置 500 はまず、ユーザの指示に従い、クライアント装置 500 のメモリに記憶されたコンテンツ読み出しプログラムを用いて、コンテンツをアーカイブサーバ 400 から読み出すための要求をアーカイブサーバ 400 に対して発行する。その要求を受信したアーカイブサーバ 400 の読み出し要求処理 412 のプログラムは、ファイル管理テーブル 1000 を参照し、読み出し対象のコンテンツがファイルストレージ 300 のプライマリ記憶領域 10 と重複排除サーバ 100 のいずれから読み出すことができるかを特定すると共に、ファイルストレージ 300 内あるいは重複排除サーバ 100 内における、読み出し対象のコンテンツの識別子を特定する。

10

【0108】

読み出し対象のコンテンツがファイルストレージ 300 のプライマリ記憶領域 10 から読み出すことができると判定した場合は、アーカイブサーバ 400 はファイルストレージ 300 から読み出し対象のコンテンツを読み出し、読み出したコンテンツをクライアント装置 500 に送信する。

【0109】

読み出し対象のコンテンツが重複排除サーバ 100 から読み出すことができると判定した場合は、アーカイブサーバ 400 は重複排除サーバ 100 に対して、読み出し対象のコンテンツを重複排除サーバ 100 から読み出すための要求を重複排除サーバ 100 に対して発行する。コンテンツ読み出し処理 2100 のプログラムにより、コンテンツの読み出しの要求を受信した重複排除サーバ 100 は、コンテンツを構成する 1 または複数のチャンクそれぞれが第一の記憶領域 20 と第二の記憶領域 30 のどちらに格納されているかを判定し、必要なチャンクのデータを第一の記憶領域 20 あるいは第二の記憶領域 30 から読み出した後、読み出したチャンクからコンテンツを再構築し、アーカイブサーバ 400 に再構築したコンテンツを送信する。重複排除サーバ 100 から、読み出し対象のコンテンツを受信したアーカイブサーバ 400 は、受信したコンテンツをクライアント装置 500 に送信する。

20

【0110】

以上が、クライアント装置 500 がアーカイブサーバ 400 からコンテンツを読み出す際の概略動作についての説明である。

30

【0111】

次に、図 30、図 31 を用いて従来の重複排除処理と本実施形態における重複排除処理の概念処理について説明する。なお、図 30、図 31 では詳細な処理の記載は省略してある。また、以下の説明で主体が省略されている場合、当該主体は前述の CPU、コントローラ等である。

【0112】

図 30 は、従来の重複排除処理の概念図である。なお、図 30 の説明においては、重複排除サーバ 100 が管理する、光ディスクライブラリ 200 もしくはファイルストレージ 300 内の記憶領域を、単に重複排除サーバ 100 の記憶領域と呼ぶことがある。

40

【0113】

従来の重複排除処理においては、コンテンツを管理するためにチャンク管理テーブルとコンテンツ管理テーブルの 2 つのテーブルを用いる。

【0114】

チャンク管理テーブルには、重複排除サーバ 100 の記憶領域に格納されたチャンクに関する情報が記録されており、各チャンクにつき、チャンク識別子（チャンクを重複排除サーバ 100 内で一意に識別するための識別子）、チャンクのハッシュ値、チャンクの格納位置が記録されている。なお、チャンク管理テーブルには同一のハッシュ値が複数登録されている場合がある。これは、複数のチャンクのデータが完全に一致しないにも関わらず、それらチャンクが同一のハッシュ値を持つ場合に起きる。

50

【 0 1 1 5 】

コンテンツ管理テーブルには、重複排除サーバ100で管理されているコンテンツに関する情報が記録されており、各コンテンツにつき、コンテンツ識別子と、コンテンツに含まれるチャンクのチャンク識別子のリストが記録されている。

【 0 1 1 6 】

図30の上側のフローチャートは、コンテンツを重複排除サーバ100の記憶領域に格納する際の処理である、(A)コンテンツ格納処理の概念図を示している。

【 0 1 1 7 】

(A)においては、まず、格納対象のコンテンツは1または複数のチャンクに分割される。その後、各チャンクに対して、(A1)および(必要であれば)(A2)の処理が適用される。

【 0 1 1 8 】

(A1)、(A2)の処理においては、重複排除サーバ100の記憶領域内に当該チャンクと完全に一致するチャンクが存在すれば、当該チャンクを破棄し、そうでなければ、当該チャンクを重複排除サーバの記憶領域に格納する。

【 0 1 1 9 】

具体的には、(A1)では、チャンク管理テーブルを参照し、処理対象のチャンクのハッシュ値と一致するハッシュ値を検索する。

【 0 1 2 0 】

処理対象のチャンクのハッシュ値と一致するハッシュ値が見つかった場合、処理対象のチャンクと完全に一致するチャンクが重複排除サーバ100の記憶領域に存在する可能性があるため、(A2)に進む。

【 0 1 2 1 】

処理対象のチャンクのハッシュ値と一致するハッシュ値が見つからなかった場合、処理対象のチャンクと完全に一致するチャンクは重複排除サーバ100の記憶領域に存在しないため、処理対象のチャンクを重複排除サーバ100の記憶領域に格納すると共に、チャンク管理テーブルに、処理対象のチャンクのチャンク識別子、ハッシュ値、格納位置を登録し、次のチャンクの処理に移る。

【 0 1 2 2 】

また、具体的には、(A2)では、処理対象のチャンクと、処理対象のチャンクと完全に一致する可能性のあるチャンクが完全に一致するかを、それらをバイナリコンペア(2つのチャンクのバイト列が一致するかの比較)することにより確認する。

【 0 1 2 3 】

バイナリコンペアの結果が一致である場合、処理対象のチャンクは重複排除サーバ100の記憶領域内のチャンクと重複するため、処理対象のチャンクを破棄し、次のチャンクの処理に移る。

【 0 1 2 4 】

バイナリコンペアの結果が不一致である場合、処理対象のチャンクと完全に一致するチャンクは重複排除サーバ100の記憶領域に存在しないため、処理対象のチャンクを重複排除サーバ100の記憶領域に格納すると共に、チャンク管理テーブルに処理対象のチャンクのチャンク識別子、ハッシュ値、格納位置を登録し、次のチャンクの処理に移る。

【 0 1 2 5 】

当該コンテンツに含まれるすべてのチャンクの処理が完了した後、コンテンツ管理テーブルにコンテンツ識別子と、当該コンテンツに含まれるチャンクのチャンク識別子のリストを登録し、(A)の処理を終了する。

【 0 1 2 6 】

図30の下側のフローチャートは、コンテンツを重複排除サーバ100の記憶領域から読み出す際の処理である、(B)コンテンツ読み出し処理の概念図を示している。(B)は、読み出し対象のコンテンツ識別子を指定して呼び出される。

【 0 1 2 7 】

10

20

30

40

50

(B)においては、まず、コンテンツ管理テーブルを参照して、指定されたコンテンツ識別子に対するコンテンツに含まれるチャンク識別子のリストを特定する。そして、(B1)にて、各チャンク識別子の対応するチャンクを重複排除サーバの記憶領域から読み出し、コンテンツを再構築し、(B)の呼び出し元に再構築したコンテンツを送信する。(B1)においては、必要チャンクが光ディスク240に格納されているのであれば、必要に応じて光ディスク240をロードして、必要チャンクを読み出す。

【0128】

以上が従来の上記重複排除処理の概念図の説明である。

【0129】

図31は、本発明の実施形態における重複排除処理の概念図である。

10

【0130】

本実施形態の重複排除処理においては、コンテンツを管理するために、第一のチャンク管理テーブル1400、第二のチャンク管理テーブル1500、コンテンツ管理テーブル1300の3つのテーブルを用いる。各テーブルの詳細な説明は図6以降の説明にて行う。

【0131】

重複排除サーバが管理する記憶領域には、第一の記憶領域20と第二の記憶領域30の2つが存在する。第一の記憶領域20には、後述する第一の重複排除と後述する第二の検索処理とが完了したチャンクであって、後述する第二のバイナリコンペアの処理を待つチャンクが一時的に格納される。第一の記憶領域20に格納されたチャンクは前記第一のチャンク管理テーブル1400で管理される。第二の記憶領域30は、第二の重複排除が完了したチャンクが格納される記憶領域であり、書き込みバッファ領域40と光ディスク240から構成される。書き込みバッファ領域40は、光ディスク240に書き込まれるのを待つチャンクが一時的に格納される記憶領域である。光ディスク240は第二の重複排除が完了したチャンクが最終的に格納される記憶領域である。第二の記憶領域30に格納されたチャンクは前記第二のチャンク管理テーブル1500で管理される。

20

【0132】

第一のチャンク管理テーブル1400には、第一の記憶領域20に格納されたチャンクに関する情報が記録されており、各チャンクにつき、チャンクを第一の記憶領域20内で一意に識別するための識別子であるチャンク識別子、チャンクのハッシュ値、チャンクの格納位置などが記録されている。

30

【0133】

第二のチャンク管理テーブル1500には、第二の記憶領域30に格納されたチャンクに関する情報が記録されており、各チャンクにつき、チャンクを第二の記憶領域30内で一意に識別するための識別子であるチャンク識別子、チャンクのハッシュ値、チャンクの格納位置などが記録されている。

【0134】

コンテンツ管理テーブル1300には、重複排除サーバ100で管理されているコンテンツに関する情報が記録されており、各コンテンツにつき、コンテンツ識別子と、コンテンツに含まれるチャンクのチャンク識別子(第一のチャンク管理テーブル1400で管理されるチャンクか第二のチャンク管理テーブル1500で管理されるチャンクかを識別できる情報も含む)のリスト、などが記録されている。

40

【0135】

図31の上側のフローチャートは、コンテンツを第一の記憶領域20もしくは第二の記憶領域30に格納する際の処理である、(C)コンテンツ格納処理の概念図を示している。(C1)および(C2)は第一の重複排除の処理であり、(C3)は第二の重複排除の処理の一部である。

【0136】

(C)においては、まず、格納対象のコンテンツは1または複数のチャンクに分割される。その後、各チャンクに対して、(C1)、(必要であれば)(C2)、(必要であれ

50

ば) (C3) の処理が適用される。

【0137】

(C1), (C2) の処理においては, 第一の記憶領域20に処理対象のチャンクと完全に一致するチャンクが存在すれば, 処理対象のチャンクを破棄し, そうでなければ, 処理対象のチャンクに対して(C3)の処理を適用する。

【0138】

具体的には, (C1) では, 第一のチャンク管理テーブル1400を参照し, 処理対象のチャンクのハッシュ値と一致するハッシュ値を検索する処理である第一の検索処理を行う。

【0139】

(C1) にて, 処理対象のチャンクのハッシュ値と一致するハッシュ値が見つかった場合, 処理対象のチャンクと完全に一致するチャンクが第一の記憶領域20に存在する可能性があるため, (C2) に進む。(C1) にて, 処理対象のチャンクのハッシュ値と一致するハッシュ値が見つからなかった場合, 処理対象のチャンクと完全に一致するチャンクは第一の記憶領域20に存在しないため, (C3) に進む。

【0140】

また, (C2) では, 処理対象のチャンクと, 処理対象のチャンクと完全に一致する可能性のあるチャンクが完全に一致するかを, それらのバイト列が一致するか比較する処理である第一のバイナリコンペアを実行することにより確認する。

【0141】

第一のバイナリコンペアの比較結果が一致である場合, 処理対象のチャンクは, 第一の記憶領域20内のチャンクと重複するため, 破棄される。第一のバイナリコンペアの比較結果が不一致である場合, 処理対象のチャンクと完全に一致するチャンクは第一の記憶領域20に存在しないため, (C3) に進む。

【0142】

(C3) では, 第二のチャンク管理テーブル1500を参照し, 処理対象のチャンクのハッシュ値と一致するハッシュ値を検索する処理である第二の検索処理を行う。

【0143】

(C3) にて, 処理対象のチャンクのハッシュ値と一致するハッシュ値が見つかった場合, 処理対象のチャンクと完全に一致するチャンクが第二の記憶領域30に存在する可能性があるため, 当該チャンクを第一の記憶領域20に格納し, 第一のチャンク管理テーブル1400に処理対象のチャンクのチャンク識別子, ハッシュ値, 格納位置を登録し, (D2) における第二のバイナリコンペアを(C)の終了後のある時点で実行する予定である旨の情報(本情報はバイナリコンペア予定チャンク管理テーブル1700にて管理される)を登録し, コンテンツ管理テーブル1300にコンテンツ識別子と, コンテンツに含まれるチャンクのチャンク識別子のリストを登録し, 次のチャンクの処理に移る。(C3) では, 要するに, 第二のバイナリコンペアの実行は先送りされるということである。なお, (D2) における第二のバイナリコンペアでは, 当該チャンクと, 当該チャンクと完全に一致する可能性のある, 第二の記憶領域30内のチャンクをバイナリコンペアする。

【0144】

(C3) にて, 処理対象のチャンクのハッシュ値と一致するハッシュ値が見つからなかった場合, 処理対象のチャンクと完全に一致するチャンクは第二の記憶領域30に存在しないため, 第二の記憶領域30内の書き込みバッファ領域40に格納し, 第二のチャンク管理テーブル1500に処理対象のチャンクのチャンク識別子, ハッシュ値, 格納位置を登録し, (D3) における書き込み処理を(C)の終了後のある時点で実行する予定である旨の情報(本情報は書き込み予定チャンク管理テーブル1800にて管理される)を登録し, コンテンツ管理テーブル1300にコンテンツ識別子と, コンテンツに含まれるチャンクのチャンク識別子のリストを登録し, 次のチャンクの処理に移る。(C3) では, 要するに, 光ディスクへの書き込みの実行は先送りされるということである。

【0145】

10

20

30

40

50

当該コンテンツに含まれるすべてのチャンクの処理が完了した後，コンテンツ管理テーブル1300にコンテンツ識別子と，当該コンテンツに含まれるチャンクのチャンク識別子のリストを登録し，(C)の処理を終了する。

【0146】

図31の下側のフローチャートは，コンテンツを重複排除サーバ100の記憶領域から読み出す際の処理である，(D)コンテンツ読み出し処理の概念図を示している。(D)は，読み出し対象のコンテンツの識別子を指定して呼び出される。(D)は前述の図30の(B)と異なり，指定されたコンテンツの読み出しを行う処理(B1)と同様の処理(D1)を実行する他に，(D2)にて第二の重複排除における第二のバイナリコンペアを実行し，また，(D3)にて書き込みバッファ領域40内のチャンクを光ディスクへ書き込む書き込み処理を実行する。

10

【0147】

具体的には，(D2)では，(C3)にて実行予定として登録した第二のバイナリコンペアのうち，指定されたコンテンツの読み出し時にロードされた光ディスク240に含まれるチャンクに関連する第二のバイナリコンペアを実行する。(D2)では，要するに，(C3)にて先送りされた第二のバイナリコンペアを実行するということである。

【0148】

第二のバイナリコンペアの結果が一致である場合，処理対象の第一の記憶領域20内のチャンクは処理対象の光ディスク240内のチャンクと重複するため，処理対象の第一の記憶領域20内のチャンクを破棄する。

20

【0149】

第二のバイナリコンペアの結果が不一致である場合，処理対象の第一の記憶領域20内のチャンクと完全に一致するチャンクは光ディスク240内に存在しないため，処理対象の第一の記憶領域20内のチャンクを第二の記憶領域30内の書き込みバッファ領域40に格納し，第二のチャンク管理テーブル1500に処理対象のチャンクのチャンク識別子，ハッシュ値，格納位置を登録し，(D3)における書き込み処理を(D2)の終了後のある時点で実行する予定である旨の情報(本情報は書き込み予定チャンク管理テーブル1800にて管理される)を登録し，処理対象の第一の記憶領域20内のチャンクを破棄する(併せて，対応する第一のチャンク管理テーブル1400のエントリを削除する)。

【0150】

また，(D3)では，(C3)あるいは(D2)にて実行予定として登録された書き込み処理のうち，指定されたコンテンツの読み出し時にロードされた光ディスク240を書き込み先とする書き込み処理を実行する。光ディスク240への書き込み処理が完了した書き込みバッファ領域40内のチャンクは破棄し，それに対応して第二のチャンク管理テーブルを更新する。(D3)では，要するに，(C3)にて先送りされた光ディスクへの書き込みを実行するということである。

30

【0151】

以下では，本実施形態のメリットについて述べる。

【0152】

通常，第二の記憶領域30には膨大な数のチャンクが格納されるため，それらチャンクを管理するための第二のチャンク管理テーブル1500のサイズは，重複排除サーバ100のメモリ110のサイズより大きくなる場合が多い。そのため，第二のチャンク管理テーブル1500は，メモリ110ではなくHDD120に記憶するのが適当である。しかし，この場合，HDD120はメモリ110に比べてアクセスコストが大きいため，第二のチャンク管理テーブル1500へのアクセスのオーバーヘッドが問題となる。本実施形態では，第二のチャンク管理テーブル1500へのアクセスが発生する第二の重複排除の前に第一の重複排除を行い，第一の重複排除で破棄されなかったチャンクに対してのみ第二の重複排除を適用し，第一の重複排除で破棄されたチャンクに対しては第二の重複排除の適用をスキップできるため，第二のチャンク管理テーブル1500へのアクセス回数を低減することができ，また，第二の重複排除における第二のバイナリコンペアの回数を低

40

50

減することができる。第二のチャンク管理テーブル1500へのアクセス回数を低減することにより、本実施形態の重複排除処理に要する時間を短縮することができる。

【0153】

また、本実施形態においては、第二のバイナリコンペアの実行を、コンテンツ読み出し処理にて当該第二のバイナリコンペアに関連する光ディスク240がロードされる時まで先送りすることにより、第二のバイナリコンペアの実行のためだけの光ディスク240のロードを削減し、全体としての光ディスク240のロード回数を低減することができる。

【0154】

さらにまた、本実施形態においては、チャンクの光ディスク240への書き込み処理の実行を、コンテンツ読み出し処理にて当該書き込み処理に関連する光ディスク240がロードされる時まで先送りすることにより、書き込み処理のためだけの光ディスク240のロードを削減し、全体としての光ディスク240のロード回数を低減することができる。

10

【0155】

光ディスク240のロード回数を低減することにより、光ディスクライブラリ200のロボット250の動作回数を低減し、ロボット250の寿命を延ばすことができる。

【0156】

次に、図6～図15を参照しながら、本実施形態における各種テーブルについて説明する。

【0157】

図6は、ファイル管理テーブル1000の構成例を示している。

20

【0158】

ファイル管理テーブル1000は、アーカイブサーバ400が管理するコンテンツ（ファイル）に関する情報を管理するためのテーブルである。

【0159】

ファイル管理テーブル1000は、ファイルパス1001、ファイルストレージ内コンテンツ識別子1002、重複排除サーバ内コンテンツ識別子1003、および、図示しないメタデータから構成される。

【0160】

ファイルパス1001は、アーカイブサーバ400内でコンテンツ（ファイル）を一意に識別するための識別子である。例えば、ディレクトリ名とファイル名を連結した文字列として表現される。ファイルパス1001で指定されるコンテンツを当該コンテンツと呼ぶ。

30

【0161】

ファイルストレージ内コンテンツ識別子1002は、当該コンテンツのファイルストレージ300内での識別子を示す情報である。当該コンテンツがファイルストレージ300内に格納されている場合は、ファイルストレージ内コンテンツ識別子1002のエントリには、当該コンテンツのファイルストレージ300内での識別子が記録される。当該コンテンツがファイルストレージ300内に格納されていない場合は、ファイルストレージ内コンテンツ識別子1002のエントリには、無効値の「NULL」が記録される。

40

【0162】

重複排除サーバ内コンテンツ識別子1003は、当該コンテンツの重複排除サーバ100での識別子を示す情報である。当該コンテンツが重複排除サーバ100により管理されている場合は、重複排除サーバ内コンテンツ識別子1003のエントリには、当該コンテンツの重複排除サーバ100内での識別子が記録される。当該コンテンツが重複排除サーバ100内により管理されていない場合は、重複排除サーバ内コンテンツ識別子1003のエントリには、無効値の「NULL」が記録される。

【0163】

ファイル管理テーブル1000のメタデータは、当該コンテンツに関する情報であり、サイズ、作成時刻、最終更新時刻、最終アクセス時刻、所有者、アクセス制御情報などから構成される。

50

【 0 1 6 4 】

図6の例では、アーカイブサーバ400がファイルパス「/foo/bar1.txt」、
「/foo/bar2.txt」、
「/foo/bar3.txt」を持つコンテンツ（ファイル）を管理していることを示している。ファイルパス「/foo/bar1.txt」で指定されるコンテンツは、ファイルストレージ300内に「/foo/bar1.txt」という識別子を持つコンテンツとして保管されていることを示している。ファイルパス「/foo/bar2.txt」で指定されるコンテンツは、重複排除サーバ100内で「1」という識別子を持つコンテンツとして管理されていることを示している。ファイルパス「/foo/bar3.txt」で指定されるコンテンツは、ファイルストレージ300内に「/foo/bar3.txt」という識別子を持つコンテンツとして保管されており、また、重複排除サーバ100内で「10」という識別子を持つコンテンツとして管理されていることを示している（この場合、コンテンツは2カ所に存在することを示している）。

10

【 0 1 6 5 】

図7は、光ディスク管理テーブル1100の構成例を示している。

【 0 1 6 6 】

光ディスク管理テーブル1100は、光ディスクライブラリ200に含まれる光ディスク240の物理的な格納位置を管理するためのテーブルである。

【 0 1 6 7 】

光ディスク管理テーブル1100は、光ディスク#1101、格納位置1102、空き領域サイズ1103から構成される。

20

【 0 1 6 8 】

光ディスク#1101は、光ディスクライブラリ200内で光ディスク240を一意に識別するための識別子である。光ディスク#1101で識別される光ディスク240を当該光ディスクと呼ぶ。

【 0 1 6 9 】

格納位置1102は、当該光ディスクの物理的な格納位置を示す情報である。光ディスク#1101には、スロット230あるいは光ディスクドライブ220のいずれかの識別子が記憶される。

【 0 1 7 0 】

空き領域サイズ1103は、当該光ディスクの空き領域のサイズを示す情報である。

30

【 0 1 7 1 】

図7の例では、識別子「1」で識別される光ディスク240は「1」で識別されるスロット230に格納されている。また、識別子「2」で識別される光ディスク240は「1」で識別される光ディスクドライブ220に格納されている。

【 0 1 7 2 】

図8は、光ディスクドライブ管理テーブル1200の構成例を示している。

【 0 1 7 3 】

光ディスクドライブ管理テーブル1200は、光ディスクドライブ220の状態を管理するためのテーブルである。

40

【 0 1 7 4 】

光ディスクドライブ管理テーブル1200は、光ディスクドライブ#1201、ロード済み光ディスク#1202、ロック状態1203、最終使用時刻1204から構成される。

【 0 1 7 5 】

光ディスクドライブ#1201は、光ディスクライブラリ200内で光ディスクドライブ220を一意に識別するための識別子である。光ディスクドライブ#1201で識別される光ディスクドライブ220を当該光ディスクドライブと呼ぶ。

【 0 1 7 6 】

ロード済み光ディスク#1202は、当該光ディスクドライブがロード済みである光デ

50

ディスク 240 の識別子を示す情報である。光ディスクドライブ 220 に光ディスク 240 がロードされていない場合、無効値の「NULL」が記録される。

【0177】

ロック状態 1203 は、当該光ディスクドライブが重複排除サーバ 100 内のプログラムのスレッドにより占有（ロック）されているかを示す情報である。なお、重複排除サーバ 100 内のプログラムはマルチスレッドで動作することができるため、プログラムのスレッドは複数存在しうる。ロック状態 1203 の値「Locked」はロックされていることを示し、また、値「Unlocked」はアンロックされている（ロックされていない）ことを示す。当該光ディスクドライブ 220 がロックされている場合、ロックしたプログラムのスレッド以外のスレッドは当該光ディスクドライブ 220 を使用（読み書き、ロード、アンロード）することができない。

10

【0178】

最終使用時刻 1204 は、重複排除サーバ 100 内のプログラムが光ディスクドライブ #1201 で識別される光ディスクドライブ 220 を最後に使用（読み書き、ロード、アンロード）した時刻を示す情報である。

【0179】

図 8 の例では、「1」で識別される光ディスクドライブ 220 には、「5」で識別される光ディスク 240 がロード済みであり、重複排除サーバ 100 内のプログラムのスレッドにより占有（ロック）されていることを示している。また、「2」で識別される光ディスクドライブ 220 には、「10」で識別される光ディスク 240 がロード済みであり、重複排除サーバ 100 内のプログラムのスレッドにより占有（ロック）されていないことを示している。さらにまた、「3」で識別される光ディスクドライブ 220 には光ディスク 240 がロードされておらず、重複排除サーバ 100 内のプログラムのスレッドにより占有（ロック）されていないことを示している。

20

【0180】

ここで、重複排除サーバ 100 内のプログラムのスレッドによる光ディスクドライブ 220 の使い方のポリシーの例について述べる。

【0181】

まず、光ディスクドライブ 220 のロック/アンロックおよびマウント/アンマウントのポリシーについて述べる。重複排除サーバ 100 内のプログラムのスレッドが、ある光ディスクドライブ 220 をロックし、光ディスク 240 を当該光ディスクドライブ 220 にロードし、ロードした光ディスク 240 に対して読み書きを行った後、当該光ディスクドライブ 220 をアンロックするが、当該光ディスクドライブ 220 からロード済みの光ディスク 240 をアンロードしないような、光ディスク 240 を積極的にアンロードしないポリシーを採用することができる。このようなポリシーを採用する場合、前記スレッドが前記光ディスク 240 をアンロックした後、短い時間経過後、再び前記光ディスク 240 を使用する場合、すでに前記光ディスクはロード済みであるため、再度前記光ディスクをロードするオーバーヘッドを削減することができる。

30

【0182】

次に、使用する光ディスクドライブ 220 の選択のポリシーについて述べる。使用する光ディスクドライブ 220 の選択のポリシーとしては、例えば、ロック状態 1203 が「Unlocked」である光ディスクドライブ 220 のうち、最終更新時刻 1204 が最も古い光ディスクドライブ 220 を選択するというポリシーが挙げられる。このようなポリシーを使用することにより、再度利用される可能性の低い光ディスクドライブ 220 を選択することができる。

40

【0183】

図 9 は、コンテンツ管理テーブル 1300 の構成例を示している。

【0184】

コンテンツ管理テーブル 1300 は、重複排除サーバ 100 内でコンテンツ 2 を管理するための情報である。

50

【 0 1 8 5 】

コンテンツ管理テーブル 1 3 0 0 は、コンテンツ # 1 3 0 1、構成チャンクリスト 1 3 0 2、アクセス必要光ディスク枚数 1 3 0 3 の他、図示しないサイズから構成される。

【 0 1 8 6 】

コンテンツ # 1 3 0 1 は、重複排除サーバ 1 0 0 内でコンテンツを一意に識別するための識別子である。コンテンツ # 1 3 0 1 で識別されるコンテンツを当該コンテンツと呼ぶ。

【 0 1 8 7 】

構成チャンクリスト 1 3 0 2 は、当該コンテンツを構成するチャンクのリストを示す情報である。図 9 の例では、例えば、識別子「1」で識別されるコンテンツは、（後述する）第一のチャンク管理テーブル 1 4 0 0 内の識別子「1」および「2」で識別されるチャンクから構成されることを示している。また、識別子「2」で識別されるコンテンツは、（後述する）第一のチャンク管理テーブル 1 4 0 0 内の識別子「1」で識別されるチャンク、および、（後述する）第二のチャンク管理テーブル 1 5 0 0 内の識別子「2」で識別されるチャンクから構成されることを示している。

10

【 0 1 8 8 】

アクセス必要光ディスク枚数 1 3 0 3 は、当該コンテンツを読み出す際に必要となる光ディスクの枚数を示す情報である。例えば、コンテンツを構成するチャンクが全て 1 枚の光ディスク 2 4 0 に格納されている場合、アクセス必要光ディスク枚数 1 3 0 3 は 1 枚となる。また、例えば、コンテンツを構成するチャンクが全て、第一の記憶領域 2 0、もしくは、書き込みバッファ領域 4 0 に格納されていれば、アクセス必要光ディスク枚数 1 3 0 3 は 0 枚となる。

20

【 0 1 8 9 】

なお、チャンクファミリーを分割し、当該チャンクファミリーに含まれるチャンクを複数の光ディスク 2 4 0 に格納する場合、当該チャンクファミリーに含まれるコンテンツ 2 のアクセス必要光ディスク枚数 1 3 0 3 の合計がなるべく小さくなるような、チャンクファミリーの分割を行う。

【 0 1 9 0 】

コンテンツ管理テーブル 1 3 0 0 のサイズは、当該コンテンツのサイズを示す情報である。

30

【 0 1 9 1 】

図 1 0 は、第一のチャンク管理テーブル 1 4 0 0 の構成例を示している。

【 0 1 9 2 】

第一のチャンク管理テーブル 1 4 0 0 は、第一の記憶領域 2 0 に格納されたチャンクを管理するための情報である。第一の記憶領域 2 0 にチャンクが格納された際には、同テーブルに新規エントリが追加され、また、第一の記憶領域 2 0 からチャンクが削除された際には、同テーブルから、削除されたチャンクに対応するエントリが削除される。

【 0 1 9 3 】

第一のチャンク管理テーブル 1 4 0 0 の各エントリに高速にアクセスするためにインデックスを用いることができる。インデックスのデータ構造としては、例えば、キーと値の組（エントリと呼ぶ）を複数個格納し、キーに対応する値をすばやく参照するための公知のデータ構造であるハッシュテーブルを用いることができる。その場合、ハッシュテーブルのキーは後述するハッシュ値 1 4 0 2 である。インデックスとしてハッシュテーブルを用いる場合、第一のチャンク管理テーブル 1 4 0 0 には、後述するハッシュ値 1 4 0 2 を格納しない。

40

【 0 1 9 4 】

第一のチャンク管理テーブル 1 4 0 0 は、チャンク # 1 4 0 1、ハッシュ値 1 4 0 2、格納位置情報 1 4 0 3 の他、図示しないサイズから構成される。

【 0 1 9 5 】

チャンク # 1 4 0 1 は、第一の記憶領域 2 0 内でチャンクを一意に識別するための識別

50

子である。チャンク # 1 4 0 1 で識別される第一の記憶領域 2 0 内のチャンクを当該チャンクと呼ぶ。

【 0 1 9 6 】

ハッシュ値 1 4 0 2 は、当該チャンクのデータに対するハッシュ値である。ハッシュ値の算出には、前述のように、MD 5 や S H A - 1 などの公知のハッシュ関数を用いることができる。

【 0 1 9 7 】

格納位置情報 1 4 0 3 は、当該チャンクの格納位置を示す情報である。

【 0 1 9 8 】

第一のチャンク管理テーブル 1 4 0 0 のサイズは、当該チャンクのサイズを示す情報である。

10

【 0 1 9 9 】

図 1 0 の例では、識別子「 1 」で指定されるチャンクのハッシュ値は「ハッシュ値 1 」(例えば、1 2 8 ビットや 1 6 0 ビットのビット列)であり、格納位置は第一の記憶領域 2 0 内の、ファイルパス「 / p r i m a r y / c h u n k 1 」で指定される位置である。

【 0 2 0 0 】

図 1 1 は、第二のチャンク管理テーブル 1 5 0 0 の構成例を示している。

【 0 2 0 1 】

第二のチャンク管理テーブル 1 5 0 0 は、第二の記憶領域 3 0 に格納されたチャンクを管理するための情報である。第二の記憶領域 3 0 にチャンクが格納された際には、同テーブルに新規エントリが追加され、また、第二の記憶領域 3 0 からチャンクが削除された際には、同テーブルから、削除されたチャンクに対応するエントリが削除される。

20

【 0 2 0 2 】

第二のチャンク管理テーブル 1 5 0 0 の各エントリに高速にアクセスするためにインデックスを用いることができる。インデックスのデータ構造としては、前述のハッシュテーブルを用いることができる。その場合、ハッシュテーブルのキーは後述するハッシュ値 1 5 0 2 である。

【 0 2 0 3 】

第二のチャンク管理テーブル 1 5 0 0 は、チャンク # 1 5 0 1 , ハッシュ値 1 5 0 2 , 格納位置情報 1 5 0 3 , 重複頻度情報 1 5 0 4 の他、図示しないサイズから構成される。

30

【 0 2 0 4 】

チャンク # 1 5 0 1 は、第二の記憶領域 3 0 内で重複排除済みチャンク 4 を一意に識別するための識別子である。チャンク # 1 5 0 1 で識別される第二の記憶領域 3 0 内のチャンクを当該チャンクと呼ぶ。

【 0 2 0 5 】

ハッシュ値 1 5 0 2 は、当該チャンクのデータに対するハッシュ値である。ハッシュ値の算出には、前述のように、MD 5 や S H A - 1 などの公知のハッシュ関数を用いることができる。

【 0 2 0 6 】

格納位置情報 1 5 0 3 は、当該チャンクの格納位置を示す情報である。

40

【 0 2 0 7 】

重複頻度情報 1 5 0 4 は、当該チャンクに関しての、所定期間内に重複が検出された頻度(回数)を示す情報である。

【 0 2 0 8 】

第二のチャンク管理テーブル 1 5 0 0 のサイズは、当該チャンクのサイズを示す情報である。

【 0 2 0 9 】

図 1 1 の例では、識別子「 1 」で指定される重複排除済みチャンク 4 のハッシュ値は「ハッシュ値 3 」(例えば、1 2 8 ビットや 1 6 0 ビットのビット列)であり、格納位置は書き込みバッファ領域 4 0 内の、ファイルパス「 / s e c o n d a r y / c h u n k 1 」

50

で指定される位置である。また、識別子「2」で指定される重複排除済みチャンク4のハッシュ値は「ハッシュ値4」（例えば、128ビットや160ビットのビット列）であり、格納位置は、識別子「1」で指定される光ディスク240内の、ファイルパス「/secondary/chunk2」で指定される位置である。

【0210】

図12は、チャンクファミリー管理テーブル1600の構成例を示している。

【0211】

チャンクファミリー管理テーブル1600は、前述したチャンクファミリー5の構成を管理するための情報である。

【0212】

チャンクファミリー管理テーブル1600は、チャンクファミリー#1601、所属チャンクリスト1602、格納光ディスクリスト1603、合計サイズ1604から構成される。

【0213】

チャンクファミリー#1601は、重複排除サーバ100内でチャンクファミリー5を一意に識別するための識別子である。チャンクファミリー#1601で指定されるチャンクファミリー5を当該チャンクファミリーと呼ぶ。

【0214】

所属チャンクリスト1602は、当該チャンクファミリーに所属する重複排除済みチャンク4のリストである。

【0215】

格納光ディスクリスト1603は、当該チャンクファミリーに所属する重複排除済みチャンク4を格納している光ディスク240のリストである。当該チャンクファミリーに所属する重複排除済みチャンク4群は、1枚の光ディスク240に格納される場合もあるし、複数の光ディスク240に格納される場合もある（チャンクファミリー5の合計サイズが大きい場合など）。

【0216】

合計サイズ1604は、当該チャンクファミリーに所属する重複排除済みチャンク4群の合計サイズである。

【0217】

図12の例では、識別子「1」で指定されるチャンクファミリー5は、第二のチャンク管理テーブル1500内の識別子「1」で指定される重複排除済みチャンク4と第二のチャンク管理テーブル1500内の識別子「2」で指定される重複排除済みチャンク4を構成要素として含み、それら重複排除済みチャンク4は識別子「1」で指定される光ディスク240に格納されていることを示している。

【0218】

図13は、バイナリコンペア予定チャンク管理テーブル1700の構成例を示している。

【0219】

バイナリコンペア予定チャンク管理テーブル1700は、後述する第二のバイナリコンペアを適用する予定であるが、まだ第二のバイナリコンペアを適用していない、第一の記憶領域20内のチャンクと第二の記憶領域30（より正確には光ディスク240）内のチャンクの対を管理する情報である。第二のバイナリコンペアとは、あるチャンクと第二の記憶領域30（より正確には光ディスク240）内のチャンクとのバイナリコンペアのことを指す。

【0220】

バイナリコンペア予定チャンク管理テーブル1700は、ID1701、必要光ディスク#1702、コンペア元チャンク1703、コンペア先チャンク1704から構成される。

【0221】

10

20

30

40

50

ID 1701は、実行予定である第二のバイナリコンペアを一意に識別するための識別子である。ID 1701で指定される第二のバイナリコンペアを当該第二のバイナリコンペアと呼ぶ。

【0222】

必要光ディスク# 1702は、当該第二のバイナリコンペアを行うために必要な光ディスク240を指定する識別子である。

【0223】

コンペア元チャンク1703は、当該第二のバイナリコンペアのコンペア元のチャンクである、第一のチャンク管理テーブル1400で管理される第一の記憶領域20内のチャンクを示す情報である。

10

【0224】

コンペア先チャンク1704は、当該第二のバイナリコンペアのコンペア先のチャンクである、第二のチャンク管理テーブル1500で管理される第二の記憶領域30内のチャンクを示す情報である。

【0225】

図13の例では、コンペア元チャンクが第一のチャンク管理テーブル1400内の識別子「1」で指定されるチャンクであり、コンペア先チャンクが第二のチャンク管理テーブル1500内の識別子「10」で指定されるチャンクであり、コンペア元チャンクとコンペア先チャンクに対して第二のバイナリコンペアを行う際に識別子「1」で指定される光ディスク240が必要であること（つまり、コンペア先チャンクが識別子「1」で指定される光ディスク240に格納されていること）を示している。

20

【0226】

図14は、書き込み予定チャンク管理テーブル1800の構成例を示している。

【0227】

書き込み予定チャンク管理テーブル1800は、光ディスク240に書き込む予定であるが、まだ書き込みを行っていない、書き込みバッファ領域40内のチャンクを管理するテーブルである。

【0228】

書き込み予定チャンク管理テーブル1800は、ID 1801、書き込み先光ディスク# 1802、書き込みチャンク1803から構成される。

30

【0229】

ID 1801は、実行予定である書き込みを一意に識別するための識別子である。ID 1801で指定される書き込みを当該書き込みと呼ぶ。

【0230】

書き込み先光ディスク# 1802は、当該書き込みの書き込み先の光ディスク240を指定する識別子である。

【0231】

書き込みチャンク1803は、当該書き込みの対象となる、第二のチャンク管理テーブル1500で管理される書き込みバッファ領域40内のチャンクを示す情報である。

40

【0232】

図14の例では、識別子「1」で指定される書き込みは、第二のチャンク管理テーブル1500内の識別子「10」で指定されるチャンクが対象であり、その書き込み先の光ディスク240は識別子「1」で指定される光ディスク240であることを示している。

【0233】

図15は、統計情報管理テーブル1900の構成例を示している。

【0234】

統計情報管理テーブル1900は、重複排除サーバ100で管理している各種統計情報を管理するテーブルである。

【0235】

統計情報管理テーブル1900は、統計情報種別1901、統計値1902から構成さ

50

れる。

【0236】

統計情報種別1901は統計情報の種別を示す情報である。統計情報の種別としては、「第一のチャンク管理テーブルの使用容量[MB]」、「第一のチャンク管理テーブルの空き容量[MB]」、「第二のチャンク管理テーブルの使用容量[MB]」、「第二のチャンク管理テーブルの空き容量[MB]」、「重複排除サーバ管理下のコンテンツの合計サイズ[MB]」、「第一の記憶領域の使用容量[MB]」、「第一の記憶領域の空き容量[MB]」、「第二の記憶領域の使用容量[MB]」、「第二の記憶領域の空き容量[MB]」、「書き込みバッファ領域の使用容量[MB]」、「書き込みバッファ領域の空き容量[MB]」が存在する。以下では、これらの統計情報の意味と更新契機について述べる。

10

【0237】

「第一のチャンク管理テーブルの使用エントリ数」は第一のチャンク管理テーブル1400が使用済みのエントリの合計数を示している。「第一のチャンク管理テーブルの空きエントリ数」は第一のチャンク管理テーブル1400の空きエントリの合計数を示している。「第一のチャンク管理テーブルの使用エントリ数」、「第一のチャンク管理テーブルの空きエントリ数」の統計値1902は、第一のチャンク管理テーブル1400にエントリを追加したり、第一のチャンク管理テーブル1400からエントリを削除したりした際に更新される。

20

【0238】

「第二のチャンク管理テーブルの使用エントリ数」は第二のチャンク管理テーブル1500が使用済みのエントリの合計数を示している。「第二のチャンク管理テーブルの空きエントリ数」は第二のチャンク管理テーブル1500の空きエントリの合計数を示している。「第二のチャンク管理テーブルの使用エントリ数」、「第二のチャンク管理テーブルの空きエントリ数」の統計値1902は、第二のチャンク管理テーブル1500にエントリを追加、あるいは、第二のチャンク管理テーブル1500からエントリを削除した際に更新される。

30

【0239】

「重複排除サーバ管理下のコンテンツの合計サイズ[MB]」はコンテンツ管理テーブル1300に含まれるコンテンツ2のサイズの合計を示している。「重複排除サーバ管理下のコンテンツの合計サイズ[MB]」の統計値1902は重複排除サーバ100にコンテンツを格納、あるいは、重複排除サーバ100からコンテンツを削除した際に更新される。

40

【0240】

「第一の記憶領域の使用容量[MB]」は第一のチャンク管理テーブル1400に含まれるチャンクのサイズの合計を示している。「第一の記憶領域の空き容量[MB]」は第一の記憶領域20の空き領域の容量を示している。「第一の記憶領域の使用容量[MB]」、「第一の記憶領域の空き容量[MB]」の統計値1902は、第一の記憶領域20にチャンクを格納、あるいは、第一の記憶領域20からチャンクを削除した際に更新される。

40

【0241】

「第二の記憶領域の使用容量[MB]」は第二のチャンク管理テーブル1500に含まれるチャンクのサイズの合計を示している。「第二の記憶領域の空き容量[MB]」は第二の記憶領域30内の空き領域の容量を示している。「第二の記憶領域の使用容量[MB]」、「第二の記憶領域の空き容量[MB]」の統計値1902は、第二の記憶領域30にチャンクを格納、あるいは、第二の記憶領域30からチャンクを削除した際に更新される。

40

【0242】

「書き込みバッファ領域の使用容量[MB]」は第二のチャンク管理テーブル1500に含まれるチャンクのうち書き込みバッファ領域40に格納されたチャンクのサイズの合

50

計を示している。「書き込みバッファ領域の空き容量 [MB]」は書き込みバッファ領域 40 内の空き領域の容量を示している。「書き込みバッファ領域の使用容量 [MB]」, 「書き込みバッファ領域の空き容量 [MB]」の統計値 1902 は, 書き込みバッファ領域 40 にチャンクを格納, あるいは, 書き込みバッファ領域 40 からチャンクを削除した際に更新される。

【0243】

以上が本実施形態における各種テーブルについての説明である。

【0244】

図16は, 管理サーバ600のディスプレイ610の表示例を示している。

【0245】

管理者によるマウス/キーボード操作に応じて, 管理サーバ600はディスプレイ610に統計情報管理テーブル1900に格納された各種統計情報(統計情報の種別および統計値)を表示する。

【0246】

次に, 図17~図25を参照しながら, 本実施形態における情報処理システム1が実行する各種処理の流れについて説明する。以下の説明で主体が省略されている場合, 当該主体は前述のCPU, コントローラ等である。Sはステップを表す。「Y」, 「N」はそれぞれ「Yes」, 「No」を表す。

【0247】

図17~図19はコンテンツ格納処理2000の処理フローの一例を示している。

【0248】

コンテンツ格納処理2000は, アーカイブサーバ400が, マイグレーション処理411のプログラムにおいて重複排除サーバ100にコンテンツを格納する際に, 重複排除サーバ100が実行する処理である。

【0249】

重複排除サーバ100は, S2001では, アーカイブサーバ400から格納対象のコンテンツを受信する。以下では格納対象のコンテンツを当該コンテンツと呼ぶ。重複排除サーバ100は, 当該コンテンツを一時的にメモリ110に格納する。なお, 重複排除サーバ100は, 当該コンテンツを一時的に, メモリ110ではなく, HDD120, あるいは, HDD320に格納してもよい。

【0250】

重複排除サーバ100は, S2002では, コンテンツ管理テーブル1300内に新規エントリを追加する。追加した新規エントリを当該新規エントリと呼ぶ。当該新規エントリのコンテンツ#1301には登録済みエントリの識別子と重複しない識別子を登録する。当該新規エントリの構成チャンクリスト1302には空のリストを登録する。当該新規エントリのアクセス必要光ディスク枚数1303には無効値の「NULL」を登録する。

【0251】

重複排除サーバ100は, S2003では, S2002にて登録したコンテンツの識別子をアーカイブサーバ400に送信する。アーカイブサーバ400は, 送信された識別子を受信し, 当該コンテンツに対応する, ファイル管理テーブル1000のエントリの重複排除サーバ内コンテンツ識別子1003に受信した識別子を登録する。

【0252】

重複排除サーバ100は, S2004では, 当該コンテンツを1または複数のチャンクに分割する。チャンクのサイズとしては予め指定された固定長のサイズを用いることができるし, 当該コンテンツの内容によって個々のチャンクのサイズを調整する可変長のサイズを用いることもできる。

【0253】

重複排除サーバ100は, 分割により得られた1または複数のチャンクに対して, S2005~S2012の処理を実行する。

【0254】

10

20

30

40

50

重複排除サーバ100は、S2005では、未処理のチャンクを先頭から順番にひとつ選択する。以下のコンテンツ格納処理2000の説明においては、選択されたチャンクを当該チャンクと呼ぶ。

【0255】

重複排除サーバ100は、S2006では、当該チャンクのハッシュ値を算出する。

【0256】

重複排除サーバ100は、S2007では、算出したハッシュ値が第一のチャンク管理テーブル1400に登録済みであるかを判定する。判定結果が「N」であれば、図18の「B」に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第一の記憶領域20内には存在しないことを意味している。判定結果が「Y」であれば、S2008に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第一の記憶領域20内に存在する可能性があることを示している。

10

【0257】

重複排除サーバ100は、S2008では、当該チャンクと、S2007で登録済みであると判定したハッシュ値に対応する、第一の記憶領域20内のチャンクとをバイナリコンペアする処理である第一のバイナリコンペアを実行する。

【0258】

重複排除サーバ100は、S2009では、第一のバイナリコンペアの結果が一致であるかどうかを判定する。判定結果が「N」であれば、図18の「B」に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第一の記憶領域20内には存在しないことを意味している。判定結果が「Y」であれば、S2010に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第一の記憶領域20内は存在したことを意味している。

20

【0259】

重複排除サーバ100は、S2010で、当該チャンクを破棄する。

【0260】

重複排除サーバ100は、S2011で、コンテンツ管理テーブル1300を更新する。具体的には、重複排除サーバ100は、当該コンテンツに対応するコンテンツ管理テーブル1300の構成チャンクリスト1302に、S2007で登録済みであると判定したハッシュ値に対応する、第一の記憶領域20内のチャンクの識別子を追加する。

30

【0261】

重複排除サーバ100は、S2012で、未処理のチャンクがあるかを判定する。判定結果が「N」であれば、コンテンツ格納処理2000を終了する。判定結果が「Y」であれば、S2005に戻る。

【0262】

重複排除サーバ100は、S2013では、S2006で算出したハッシュ値が第二のチャンク管理テーブル1500に登録済みであるかを判定する。判定結果が「N」であれば、S2020に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第二の記憶領域30内には存在しないことを意味している。判定結果が「Y」であれば、S2014に進む。この判定結果は、当該チャンクと完全に一致するチャンクが第二の記憶領域30内に存在する可能性があることを示している。

40

【0263】

重複排除サーバ100は、S2014では、後述する第二のバイナリコンペアに必要なチャンクが即時アクセス可能かを判定する。具体的には、(1)S2013で登録済みであると判定したハッシュ値に対応する、第二のチャンク管理テーブル1500のエントリの格納位置情報1503が書き込みバッファ領域40内のチャンクを指し示している場合、あるいは、(2)同エントリの格納位置情報1503が光ディスク240内のチャンクを指し示しており、かつ、同光ディスク240に対応する、光ディスク管理テーブルの格納位置1102が光ディスクドライブ220を指し示している場合に、即時アクセス可能である(「Y」)と判定する。それ以外の場合は、同光ディスク240がスロット230

50

に格納されていることを意味するため、即時アクセス可能でない(「N」)と判定する。判定結果が「Y」の場合は、S 2 0 1 7に進む。判定結果が「N」の場合はS 2 0 1 5に進む。

【0264】

重複排除サーバ100は、S 2 0 1 5では、当該チャンクを第一の記憶領域20に格納する。

【0265】

重複排除サーバ100は、S 2 0 1 6では、当該チャンクを第一のチャンク管理テーブル1400に登録し、バイナリコンペア予定チャンク管理テーブル1700とコンテンツ管理テーブル1300を更新する。

10

【0266】

具体的には、重複排除サーバ100は、第一のチャンク管理テーブル1400に新規エントリを追加し、新規エントリのチャンク#1401、ハッシュ値1402、格納位置情報1403にはそれぞれ、(1)他のエントリの識別子と重複しない識別子、(2)S 2 0 0 6で算出したハッシュ値、(3)S 2 0 1 5で当該チャンクを格納した格納位置に対する識別子を設定する。

【0267】

さらに、重複排除サーバ100は、バイナリコンペア予定チャンク管理テーブル1700に新規エントリを追加し、(1)新規エントリのID1701に他のエントリと重複しない識別子を設定し、(2)新規エントリの必要光ディスク1702に、S 2 0 1 3で登録済みであると判定したハッシュ値に対応する、第二の記憶領域30内のチャンクが格納されている光ディスク240の識別子を設定し、(3)コンペア元チャンク1703に、S 2 0 1 5で第一の記憶領域20に格納したチャンクを重複排除サーバ100内で一意に識別する識別子を設定し、(4)コンペア先チャンク1704に、S 2 0 1 3で登録済みであると判定したハッシュ値に対応する、第二の記憶領域30内のチャンクを重複排除サーバ100内で一意に識別する識別子を設定する。

20

【0268】

さらにまた、重複排除サーバ100は、当該コンテンツに対応するコンテンツ管理テーブル1300の構成チャンクリスト1302に、S 2 0 1 5で第一の記憶領域20に格納したチャンクを重複排除サーバ100内で一意に識別する識別子を追加する。その後、重複排除サーバ100は、図17の「A」に進む。

30

【0269】

重複排除サーバ100は、S 2 0 1 7では、当該チャンクと、S 2 0 1 3で登録済みであると判定したハッシュ値に対応する、第二の記憶領域30内のチャンクとをバイナリコンペアする処理である第二のバイナリコンペアを実行する。

【0270】

重複排除サーバ100は、S 2 0 1 8では、S 2 0 1 7における第二のバイナリコンペアの結果が一致であるかを判定する。判定結果が「Y」(一致である)の場合はS 2 0 1 9に進む。判定結果が「N」(一致でない)の場合はS 2 0 2 0に進む。

40

【0271】

重複排除サーバ100は、S 2 0 1 9では、図19の「C」の処理を実行する。

【0272】

重複排除サーバ100は、S 2 0 2 0では、当該チャンクを書き込みバッファ領域40に格納する。

【0273】

重複排除サーバ100は、S 2 0 2 1では、第二のチャンク管理テーブル1500を更新する。具体的には、第二のチャンク管理テーブル1500に新規エントリを追加し、新規エントリのチャンク#1501、ハッシュ値1502、格納位置情報1503、重複頻度情報1504にそれぞれ、(1)第二のチャンク管理テーブル1500内の他のエントリと重複しない識別子、(2)S 2 0 0 6で算出したハッシュ値、(3)S 2 0 2 0で当

50

該チャンクを格納した書き込みバッファ領域 40 内の位置を一意に示す識別子，(4) 初期値の「0」，を設定する。

【0274】

なお，以下のコンテンツ格納処理 2000 のステップにおいて，S2021 で追加した新規エントリに対応するチャンクを新規チャンクと呼ぶ。

【0275】

重複排除サーバ 100 は，S2022 では，コンテンツ管理テーブル 1300 を更新する。具体的には，重複排除サーバ 100 は，当該コンテンツに対応するコンテンツ管理テーブル 1300 の構成チャンクリスト 1302 に，S2013 で登録済みであると判定したハッシュ値に対応する，第二の記憶領域 30 内のチャンクの識別子を追加する。

10

【0276】

重複排除サーバ 100 は，S2023 では，チャンクファミリー管理テーブル 1600 を更新する。具体的には，まず，チャンクファミリー管理テーブル 1600，コンテンツ管理テーブル 1300，第二のチャンク管理テーブル 1500 を参照して，新規チャンクが所属するチャンクファミリー 5 を特定する。次に，特定したチャンクファミリー 5 がチャンクファミリー管理テーブル 1600 に登録されていないければ，チャンクファミリー管理テーブル 1600 に新規エントリを追加し，(1) 新規エントリのチャンクファミリー #1601 に特定したチャンクファミリー 5 を一意に識別する識別子を設定し，(2) 新規エントリの所属チャンクリスト 1602 に，前記新規チャンクを重複排除サーバ 100 内で一意に識別する識別子を追加し，(3) 新規エントリの格納光ディスクリスト 1603 に，光ディスク管理テーブル 1100 の空き領域サイズ 1103 が予め設定された閾値（例えば，50 [MB]）以上である光ディスク 240 のうち，光ディスク #1101 が最も小さい光ディスク 240 の光ディスク #1101 を設定し，(4) 新規エントリの合計サイズ 1604 に前記新規チャンクのサイズを設定する。特定したチャンクファミリー 5 がチャンクファミリー管理テーブル 1600 に登録されていれば，当該チャンクファミリー 5 に対応するエントリの所属チャンクリスト 1602 に，前記新規チャンクを重複排除サーバ 100 内で一意に識別する識別子を追加し，合計サイズ 1604 に前記新規チャンクのサイズを足し込む。また，格納光ディスクリスト 1603 には，所属チャンクリスト 1602 に含まれるチャンクが格納された光ディスク 240 のリスト（第一のリストと呼ぶ），もしくは，第一のリストに含まれる全ての光ディスク 240 の空き領域サイズが

20

30

【0277】

重複排除サーバ 100 は，S2024 では，書き込み予定チャンク管理テーブル 1800 を更新する。具体的には，書き込み予定チャンク管理テーブル 1800 に新規エントリを追加し，新規エントリの ID 1801，書き込み先光ディスク #1802，書き込みチャンク 1803 にそれぞれ，(1) 書き込み予定チャンク管理テーブル 1800 内で重複しない識別子，(2) 前記新規チャンクを重複排除サーバ 100 内で一意に識別する識別子，(3) 前記新規チャンクが所属するチャンクファミリー 5 に対応する，チャンクファミリー管理テーブル 16 の格納光ディスクリスト 1603 に含まれる光ディスク 240 のうち，空き領域サイズ 1103 が最も小さい光ディスク 240 の光ディスク #1101，を設定する。その後，重複排除サーバ 100 は図 17 の「A」に進む。

40

【0278】

重複排除サーバ 100 は，S2025 では，第二のチャンク管理テーブル 1500 の重複頻度情報 1504 を更新する。具体的には，S2013 で登録済みであると判定したハッシュ値に対応する，第二のチャンク管理テーブル 1500 のエントリの重複頻度情報 1504 を 1 だけインクリメントする。

50

【0279】

重複排除サーバ100は、S2026では、第二のチャンク管理テーブル1500に登録されたチャンクが頻出チャンクであるかを判定する。具体的には、S2013で登録済みであると判定したハッシュ値に対応する、第二のチャンク管理テーブル1500のエントリの重複頻度情報1504が、予め管理者により設定された閾値（例えば「10」回）より大きければ、頻出チャンクである（「Y」）と判定し、そうでなければ、頻出チャンクでない（「N」）と判定する。判定結果が「Y」であれば、S2029に進み、判定結果が「N」であれば、S2027に進む。

【0280】

重複排除サーバ100は、S2027では、当該チャンクを破棄する。

10

【0281】

重複排除サーバ100は、S2028では、コンテンツ管理テーブル1300を更新する。具体的には、重複排除サーバ100は、当該コンテンツに対応するコンテンツ管理テーブル1300の構成チャンクリスト1302に、S2013で登録済みであると判定したハッシュ値に対応する、第二の記憶領域30内のチャンクの識別子を追加し、図19の「C」の処理を終了する。

【0282】

重複排除サーバ100は、S2029、S2030ではそれぞれ、S2015、S2016と同様の処理を行う。その後、図19の「C」の処理を終了する。

20

【0283】

S2026～S2029において、頻出チャンクを第一の記憶領域20に格納することにより、以後、頻出チャンクと完全に一致するチャンク（一致チャンクと呼ぶ）は第一の重複排除で破棄されるため、一致チャンクに対しては第二の重複排除を行う必要がなくなり、本実施形態の重複排除の処理のオーバーヘッドを低減することができる。

【0284】

図20はコンテンツ読み出し処理2100の処理フローの一例を示している。

【0285】

コンテンツ読み出し処理2100では、重複排除サーバ100は、読み出し要求のあったコンテンツを構成するチャンクを、第一の記憶領域20、および、第二の記憶領域30から読み出し、コンテンツを再構築して読み出し要求元へ送信する。また、読み出し要求のあったコンテンツを構成するチャンクを第二の記憶領域30の光ディスク240のから読み出した場合、当該光ディスク240に関連する、(1)バイナリコンペア予定チャンク管理テーブル1700のエントリにより指定される第二のバイナリコンペア、および、(2)書き込み予定チャンク管理テーブル18のエントリにより指定される書き込み処理、をそれぞれ実行する。

30

【0286】

重複排除サーバ100は、S2101では、アーカイブサーバ400からコンテンツの読み出し要求を受信する。読み出し要求には、読み出し対象のコンテンツのコンテンツ#1301が含まれる。以下では、読み出し対象のコンテンツを当該コンテンツと呼ぶ。

【0287】

重複排除サーバ100は、S2102では、当該コンテンツを送信する際に用いる送信バッファを、重複排除サーバ100のメモリ110内に確保する。なお、送信バッファは、HDD120あるいはHDD320内に確保してもよい。送信バッファのサイズとしては読み出し対象のコンテンツのサイズを用いる。

40

【0288】

重複排除サーバ100は、S2103では、第一の記憶領域20もしくは書き込みバッファ40に格納された必要チャンクを読み出し、確保した送信バッファに書き込む。具体的には、まず、当該コンテンツに対応する、コンテンツ管理テーブル1300の構成チャンクリスト1302を参照して、当該コンテンツの再構築に必要なチャンク（必要チャンクと呼ぶ）を特定する。次に、第一のチャンク管理テーブル1400を参照して、特定し

50

た必要チャンクのうち、第一の記憶領域 20 に格納されたチャンクを特定し、特定したチャンクを第一の記憶領域 20 から読み出し、読み出したチャンクを前記送信バッファの適切な位置に書き込む。次に、第二のチャンク管理テーブル 1500 を参照して、特定した必要チャンクのうち、書き込みバッファ領域 40 に格納されたチャンクを特定し、特定したチャンクを書き込みバッファ領域 40 から読み出し、読み出したチャンクを前記送信バッファの適切な位置に書き込む。

【0289】

重複排除サーバ 100 は、S2104 では、当該コンテンツを再構築するために必要な光ディスク（必要光ディスクと呼ぶ）を特定する。具体的には、第二のチャンク管理テーブル 1500 を参照して、前記必要チャンクが格納された 1 または複数の光ディスク 240 を特定する。

10

【0290】

次に、重複排除サーバ 100 は、必要光ディスクに対して、S2105 ~ S2111 の処理を実行する。

【0291】

重複排除サーバ 100 は、S2105 では、必要光ディスクのうち、未処理の光ディスクをひとつ選択する。選択した光ディスクを当該光ディスクと呼ぶ。

【0292】

重複排除サーバ 100 は、S2106 では、当該光ディスクがロード済みかを判定する。具体的には、当該光ディスクに対応する、光ディスク管理テーブル 1100 の格納位置 1102 を参照し、格納位置 1102 が光ディスクドライブ 220 であれば、ロード済みである（「Y」）と判定し、格納位置 1102 がスロット 230 であれば、ロード済みでない（「N」）と判定する。判定結果が「Y」であれば、S2108 に進み、「N」であれば、S2107 に進む。

20

【0293】

重複排除サーバ 100 は、S2107 では、当該光ディスクをスロット 230 から光ディスクドライブ 220 にロードする。

【0294】

重複排除サーバ 100 は、S2108 では、必要チャンクのうち、当該光ディスクに含まれるチャンクを読み出し、前記送信バッファに書き込む。具体的には、第二のチャンク管理テーブル 1500 の格納位置情報 1503 を参照し、前記必要チャンクのうち、当該光ディスクに含まれるチャンクを特定し、特定したチャンクを当該光ディスクから読み出し、読み出したチャンクを前記送信バッファの適切な位置に書き込む。

30

【0295】

重複排除サーバ 100 は、S2109 では、図 21、図 22 のバイナリコンペア処理 2200 を実行する。

【0296】

重複排除サーバ 100 は、S2110 では、図 23 の書き込み処理 2300 を実行する。

【0297】

重複排除サーバ 100 は、S2111 では、未処理の光ディスクがあるかを判定する。未処理の光ディスクがあれば（「Y」）、S2105 に進み、未処理の光ディスクがなければ（「N」）、S2112 に進む。

40

【0298】

重複排除サーバ 100 は、S2112 では、前記送信バッファ内の当該コンテンツのデータを、コンテンツの読み出し要求元に送信し、コンテンツ読み出し処理 2100 を終了する。

【0299】

図 21 及び図 22 はバイナリコンペア処理 2200 の処理フローの一例を示している。

【0300】

50

バイナリコンペア処理 2200 では、バイナリコンペア予定チャンク管理テーブル 1700 で指定される第二のバイナリコンペアのうち、当該光ディスクに関連する第二のバイナリコンペアを実行する。バイナリコンペア処理 2200 により、コンテンツ格納処理 2000 の S2015 ~ S2016 で実行を先送りされた第二のバイナリコンペアが実行される。

【0301】

重複排除サーバ 100 は、S2201 では、バイナリコンペア予定チャンク管理テーブル 1700 の必要光ディスク # 1702 を参照し、必要光ディスク # 1702 が当該光ディスクの識別子と一致する、バイナリコンペア予定チャンク管理テーブル 1700 のエントリを特定する。

10

【0302】

重複排除サーバ 100 は、S2201 で特定したエントリに対して、S2202 ~ S2211 の処理を実行する。

【0303】

重複排除サーバ 100 は、S2202 では、S2201 で特定したエントリの中から未処理のエントリをひとつ選択する。

【0304】

以下のステップでは、S2202 で選択したエントリのコンペア元チャンク 1703、コンペア先チャンク 1704 で指定されるチャンクをそれぞれ、単にコンペア元チャンク、コンペア先チャンクと呼ぶ。

20

【0305】

重複排除サーバ 100 は、S2203 では、選択したエントリで指定された第二のバイナリコンペアを実行する。具体的には、コンペア元チャンクおよびコンペア先チャンクに対してバイナリコンペアを実行する。

【0306】

重複排除サーバ 100 は、S2204 では、S2203 でのバイナリコンペアの結果が一致であるかどうかを判定する。判定結果が一致である（「Y」）場合、S2205 に進み、判定結果が一致でない（「N」）場合、S2207 に進む。

【0307】

重複排除サーバ 100 は、S2205 では、図 22 の「D」の処理を実行する。

30

【0308】

重複排除サーバ 100 は、S2206 では、未処理のエントリがあるかどうかを判定する。未処理のエントリがある（「Y」）ならば、S2202 に進み、未処理のエントリがあとに（「N」）ならば、バイナリコンペア処理 2200 を終了する。

【0309】

重複排除サーバ 100 は、S2207、S2208 では、それぞれ、S2020、S2021 と同様の処理を行う。ただし、S2020、S2021 の「当該チャンク」を「コンペア元チャンク」に読み替える必要がある。

【0310】

重複排除サーバ 100 は、S2209 では、コンテンツ管理テーブル 1300 を更新する。具体的には、重複排除サーバ 100 は、コンテンツ管理テーブル 1300 の構成チャンクリスト 1302 のエントリのうち、構成チャンクリスト 1302 内にコンペア元チャンクの識別子を含むエントリ内の、コンペア元チャンクの識別子を、S2207 で書き込みバッファ領域 40 に格納したチャンクに対する識別子に置き換える。S2209 の処理は、コンペア元チャンクの格納位置を第一の記憶領域 20 から書き込みバッファ領域 40 に移動したことに伴う、コンテンツ管理テーブル 1300 の更新を意味する。

40

【0311】

重複排除サーバ 100 は、S2210、S2211 では、それぞれ、S2023、S2024 と同様の処理を行う。ただし、S2210、S2211 における「新規チャンク」を、「S2207 で書き込みバッファ領域 40 に格納したチャンク」に読み替える必要が

50

ある。その後、重複排除サーバ100は、S2205に進む。

【0312】

重複排除サーバ100は、S2212では、第二のチャンク管理テーブル1500の重複頻度情報1504を更新する。具体的には、S2207で書き込みバッファ領域40に格納したチャンクに対応する、第二のチャンク管理テーブル1500の重複頻度情報1504の値を1だけインクリメントする。

【0313】

重複排除サーバ100は、S2213では、コンペア元チャンクが頻出チャンクであるかどうかを判定する。具体的には、S2207で書き込みバッファ領域40に格納したチャンクに対応する、第二のチャンク管理テーブル1500の重複頻度情報1504の値が、予め設定された閾値（例えば「10」回）より大きければ、頻出チャンクである（「Y」と判定し、そうでなければ、頻出チャンクでない（「N」と判定する。判定結果が「Y」であれば、図22の「D」の処理を終了し、判定結果が「N」であれば、S2214に進む。

10

【0314】

重複排除サーバ100は、S2214では、コンペア元チャンクを第一の記憶領域20から削除する。

【0315】

重複排除サーバ100は、S2215では、コンペア元チャンクに対応する、第一のチャンク管理テーブル1400のエントリを削除し、図22の「D」の処理を終了する。

20

【0316】

S2212～S2215において、頻出チャンクを第一の記憶領域20から削除しないことにより、以後、頻出チャンクと完全に一致するチャンク（一致チャンクと呼ぶ）は第一の重複排除で破棄されるため、一致チャンクに対しては第二の重複排除を行う必要がなくなり、本実施形態の重複排除の処理のオーバーヘッドを低減することができる。

【0317】

図23は書き込み処理2300の処理フローの一例を示している。

【0318】

書き込み処理2300では、コンテンツ格納処理2000やバイナリコンペア処理2200で実行を先送りされた、書き込みバッファ領域40内のチャンクの光ディスク240への書き込みが実行される。

30

【0319】

重複排除サーバ100は、S2301では、書き込み予定チャンク管理テーブル1800内のエントリのうち、S2105で選択された光ディスク240に関連するエントリを特定する。

【0320】

重複排除サーバ100は、S2302では、特定したエントリに対応する、光ディスクへの書き込みを行う。具体的には、特定したエントリの書き込みチャンク1803で指定される、書き込みバッファ領域40内のチャンクを、S2105で選択された光ディスク240に書き込む。そして、書き込み予定チャンク管理テーブル1800内のうち特定したエントリを削除し、また、書き込んだチャンクに対応する、第二のチャンク管理テーブル1500のエントリを削除する。

40

【0321】

図24は強制バイナリコンペア処理2400の処理フローの一例を示している。

【0322】

強制バイナリコンペア処理2400は、強制的にバイナリコンペア予定チャンク管理テーブル1700のエントリで指定される第二のバイナリコンペアを実行し、第一の記憶領域20に溜まったチャンクの量を削減する処理である。重複排除サーバ100は、後述する、「第一の閾値[MB]<第二の閾値[MB]」を満たすような二つの閾値を用いて、削減するチャンクの量を制御する。強制バイナリコンペア処理2400は定期的（例えば

50

、1日毎)に起動される。

【0323】

重複排除サーバ100は、S2401では、統計情報管理テーブル1900の「第一の記憶領域の使用容量[MB]」の統計値1902を参照し、「第一の記憶領域の使用容量[MB]」の統計値1902が、予め設定された第二の閾値[MB]を超えるかを判定する。超える場合(「Y」)はS2402に進み、超えない場合(「N」)は、強制バイナリコンペア処理2400を終了する。

【0324】

重複排除サーバ100は、S2402では、強制バイナリコンペア処理2400に必要な光ディスクを特定する。具体的には、バイナリコンペア予定チャンク管理テーブル1700の必要光ディスク#1702に含まれる全光ディスク240をリストアップする。リストアップした光ディスク240を対象光ディスクと呼ぶ。

10

【0325】

次に、重複排除サーバ100は、対象光ディスクに対して、S2403～S2406の処理を実行する。

【0326】

重複排除サーバ100は、S2403では、対象光ディスクのうち未処理の光ディスク240をひとつ選択する。選択した光ディスク240を当該光ディスクと呼ぶ。

【0327】

重複排除サーバ100は、S2404では、当該光ディスクに対して、すでに説明したバイナリコンペア処理2200を実行する。

20

【0328】

重複排除サーバ100は、S2405では、統計情報管理テーブル1900の「第一の記憶領域の使用容量[MB]」の統計値1902を参照し、「第一の記憶領域の使用容量[MB]」の統計値1902が、予め設定された第一の閾値[MB]を下回るかを判定する。下回る場合(「Y」)は、強制バイナリコンペア処理2400を終了し、下回らない場合は、S2406に進む。

【0329】

重複排除サーバ100は、S2406では、未処理の光ディスクがあるかを判定する。未処理の光ディスクがある(「Y」)場合、S2403に進み、未処理の光ディスクがない(「N」)場合、強制バイナリコンペア処理2400を終了する。

30

【0330】

図25は強制書き込み処理2500の処理フローの一例を示している。

【0331】

強制書き込み処理2500は、強制的に書き込み予定チャンク管理テーブル1800のエントリで指定される書き込みを実行し、書き込みバッファ領域40に溜まったチャンクの量を削減する処理である。重複排除サーバ100は、後述する第三の閾値[MB]を用いて、チャンクの量を削減する必要があるかを判定する。強制書き込み処理2500は定期的(例えば、1日毎)に起動される。

【0332】

重複排除サーバ100は、S2501では、統計情報管理テーブル1900の「書き込みバッファ領域の使用容量[MB]」の統計値1902を参照し、「書き込みバッファ領域の使用容量[MB]」の統計値1902が、予め設定された第三の閾値[MB]を超えるかを判定する。超える場合(「Y」)は、S2502に進み、超えない場合(「N」)は、強制書き込み処理2500を終了する。

40

【0333】

重複排除サーバ100は、S2502では、書き込み予定チャンク管理テーブル1800に含まれる書き込み先光ディスク#1802をリストアップする。リストアップした書き込み先光ディスク#1802に対応する光ディスク240を対象光ディスクと呼ぶ。

【0334】

50

次に、重複排除サーバ100は、対象光ディスクに対して、S2503～S2505の処理を実行する。

【0335】

重複排除サーバ100は、S2503では、対象光ディスクのうち未処理の光ディスク240をひとつ選択する。選択した光ディスクを当該光ディスクと呼ぶ。

【0336】

重複排除サーバ100は、S2504では、当該光ディスクに対して、図23で説明した書き込み処理2300を実行する。ただし、図23の説明における「S2105で選択された光ディスク240」を「当該光ディスク」に読み替える必要がある。

【0337】

重複排除サーバ100は、S2505では、対象光ディスクの中に未処理の光ディスクがあるかを判定する。未処理の光ディスクがある（「Y」）場合は、S2503に進み、未処理の光ディスクがない（「N」）場合は、強制書き込み処理2500を終了する。

【0338】

以下、第二の実施形態を説明する。図27、図27は第二の実施形態の情報処理システムを説明する図である。

【0339】

図27は、第二の実施形態における情報処理システム1を示している。第二の実施形態の情報処理システム1には、第一の実施形態と異なり、重複排除サーバ100およびアーカイブサーバ400が存在しない。第一の実施形態において重複排除サーバ100およびアーカイブサーバ400が実行していた処理は、第二の実施形態においては、すべてファイルストレージ300により実行される。

【0340】

図28は、第二の実施形態におけるメモリ311、HDD320上の各種テーブル、および、メモリ311上の各種プログラム、を示している。第二の実施形態におけるメモリ311上には、第一の実施形態のアーカイブサーバ400のメモリ410上に記憶されたテーブルおよび各種プログラム、および、第一の実施形態の重複排除サーバ100のメモリ110上に記憶されたテーブルおよび各種プログラム、が記憶されている。

【0341】

なお、第二の実施形態におけるメモリ311、HDD320上の各種テーブルは、光ディスク240、HDD320、コントローラ210内のメモリ、あるいは、コントローラ210内のHDDに記憶するように構成してもよい。

【0342】

以下、第三の実施形態を説明する。図28、図29は第三の実施形態の情報処理システムを説明する図である。

【0343】

図28は、第三の実施形態における情報処理システム1を示している。第三の実施形態の情報処理システム1は、第一の実施形態の情報処理システム1とほぼ同様であるが、第一の実施形態のアーカイブサーバ400の代わりに、電子メールサーバ800を有する。第三の実施形態における情報処理システム1は、電子メールサーバ800により、クライアント装置500に対して、電子メールメッセージを電子メールサーバ800からクライアント装置500に読み込んだり、電子メールメッセージを電子メールサーバ800から送信したりする機能などを提供する。電子メールメッセージはファイルストレージ300のプライマリ記憶領域10に記憶されている。

【0344】

図29は、電子メールサーバ800のメモリ810上のテーブルや各種プログラムを示す図である。メモリ810上には、電子メール管理テーブル811といったテーブルや、マイグレーション処理812、電子メール読み出し要求処理813といったプログラムが記録されている。電子メール管理テーブル811は情報処理システム1内の電子メールメッセージを管理するテーブルである。マイグレーション処理812のプログラムは、ブラ

10

20

30

40

50

イマリ記憶領域 10 に記憶された電子メールメッセージのうち、古いメッセージを第一の記憶領域 20 あるいは第二の記憶領域 30 に移動（マイグレーション）するプログラムである。マイグレーション処理 812 の動作はマイグレーション処理 411 とほぼ同様であるが、マイグレーション処理 812 のプログラムが取り扱うデータがファイルではなく電子メールメッセージである。電子メール読み出し要求処理 813 のプログラムは、クライアント装置 500 からの電子メールメッセージ読み出し要求に応じて、クライアント装置 500 に電子メールメッセージを送信するプログラムである。電子メール読み出し要求処理 813 の動作は読み出し要求処理 412 とほぼ同様であるが、電子メール読み出し要求処理 813 が取り扱うデータはファイルではなく電子メールメッセージである。

【0345】

以上、本発明の好適な実施形態を説明したが、これらは本発明の説明のための例示であって、本発明の範囲を実施例にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。

【0346】

例えば、本発明の光ディスクライブラリはテープライブラリに置き換えることが可能である。また、図 31 で説明した本発明の実施形態の重複排除処理における第一の重複排除である（C1）（C2）の処理を省略することも可能である（図 17 の S2007 ~ S2010 の処理に代えて、図 18 の「B」の処理を行うことを意味する）。ただし、この場合、第二の重複排除の対象となるチャンクの量が增大する可能性がある。一方で、この場合でも、第二のバイナリコンペアは先送りされて、コンテンツ読み出し処理内で実行されるため、光ディスクのロード回数を低減することは可能である。

【0347】

また、例えば、異なる二つのチャンクのハッシュ値が同一になる確率は非常に低く、現実的には無視できるくらいの確率であることを想定し、図 31 の（C2）の第一のバイナリコンペア、および、（D2）の第二のバイナリコンペアを省略することも可能である。この場合、図 31 の処理（C2）（D2）は実施しない。また、図 31 の（C3）の処理では、（1）処理対象のチャンクのハッシュ値と一致するハッシュ値が見つかった場合、処理対象のチャンクを破棄して次のチャンクの処理に移り、（2）処理対象のチャンクのハッシュ値と一致するハッシュ値が見つからなかった場合、処理対象のチャンクを書き込みバッファに格納して次のチャンクの処理に移る。つまり、処理対象のチャンクの光ディスクへの書き込みの先送りのみが行われることを意味している。先送りされた、光ディスクへの書き込みは図 31 の（D3）にて実行される。光ディスクへの書き込みを先送りすることにより、光ディスクの光ディスクドライブへのロード回数を低減することができる。

【符号の説明】

【0348】

- 1 情報処理システム
- 20 第一の記憶領域
- 30 第二の記憶領域
- 40 書き込みバッファ領域
- 100 重複排除サーバ
- 200 光ディスクライブラリ
- 240 光ディスク
- 300 ファイルストレージ
- 400 アーカイブサーバ
- 500 クライアント装置
- 600 管理サーバ
- 700 相互結合網
- 1400 第一のチャンク管理テーブル
- 1500 第二のチャンク管理テーブル
- 1700 バイナリコンペア予定チャンク管理テーブル

10

20

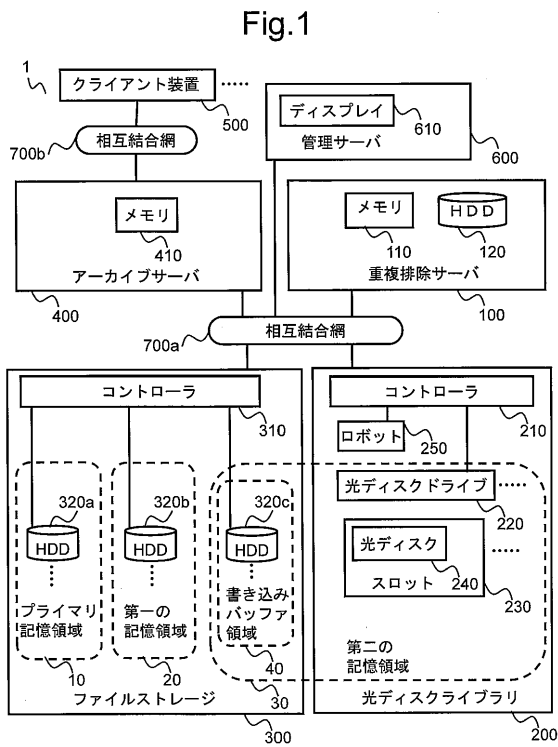
30

40

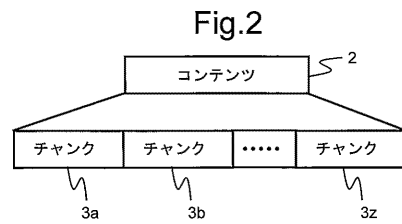
50

1800 書き込み予定チャンク管理テーブル

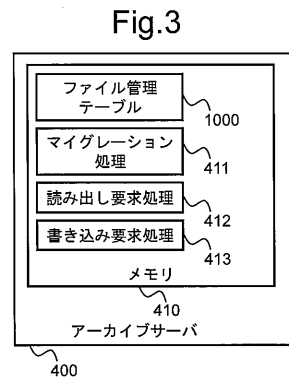
【図1】



【図2】

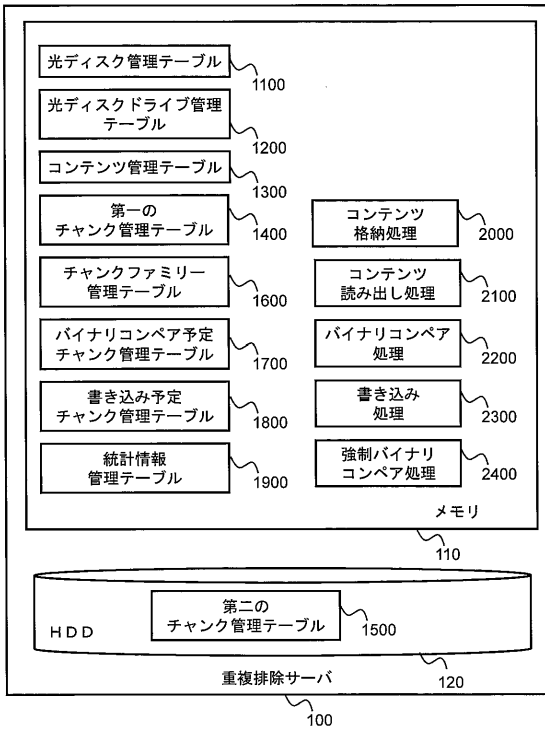


【図3】



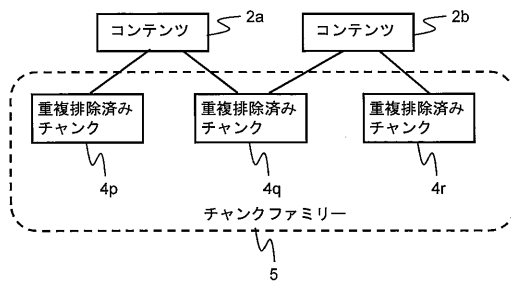
【 図 4 】

Fig.4



【 図 5 】

Fig.5



【 図 6 】

Fig.6

1001	1002	1003
ファイルパス	ファイルストレージ内コンテンツ識別子	重複排除サーバ内コンテンツ識別子
/foo/bar1.txt	/foo/bar1.txt	NULL
/foo/bar2.txt	NULL	1
/foo/bar3.txt	/foo/bar3.txt	10
:	:	:

【 図 7 】

Fig.7

1101	1102	1103
光ディスク#	格納位置	空き領域サイズ
1	スロット#1	
2	光ディスクドライブ#1	
:	:	:

【 図 8 】

Fig.8

1201	1202	1203	1204
光ディスクドライブ#	ロード済み光ディスク#	ロック状態	最終使用時刻
1	5	Locked	
2	10	Unlocked	
3	NULL	Unlocked	
:	:	:	:

【 図 9 】

Fig.9

1301	1302	1303
コンテンツ#	構成チャンクリスト	アクセス必要光ディスク枚数
1	「第一のチャンク管理テーブル」の1, 「第一のチャンク管理テーブル」の2	
2	「第一のチャンク管理テーブル」の1, 「第二のチャンク管理テーブル」の2	
:	:	:

【 図 10 】

Fig.10

1401	1402	1403
チャンク#	ハッシュ値	格納位置情報
1	ハッシュ値1	第一の記憶領域、/primary/chunk1
2	ハッシュ値2	第一の記憶領域、/primary/chunk2
:	:	:

【図11】

Fig.11

1500

第二のチャンク管理テーブル			
チャンク#	ハッシュ値	格納位置情報	重複頻度情報
1	ハッシュ値3	書き込みバッファ領域、 /secondary/chunk1	
2	ハッシュ値4	光ディスク#1、 /secondary/chunk2	
:	:	:	:

【図12】

Fig.12

1600

チャンクファミリー管理テーブル			
チャンクファミリー#	所属チャンクリスト	格納光ディスクリスト	合計サイズ
1	「第二のチャンク管理テーブル」の1、 「第二のチャンク管理テーブル」の2	光ディスク#1	
:	:	:	:

【図13】

Fig.13

1700

バイナリコンペア予定チャンク管理テーブル			
ID	必要光ディスク#	コンペア元チャンク	コンペア先チャンク
1	1	「第一のチャンク管理テーブル」の1	「第二のチャンク管理テーブル」の10
2	1	「第一のチャンク管理テーブル」の2	「第二のチャンク管理テーブル」の11
:	:	:	:

【図14】

Fig.14

1800

書き込み予定チャンク管理テーブル		
ID	書き込み先光ディスク#	書き込みチャンク
1	1	「第二のチャンク管理テーブル」の10
2	1	「第二のチャンク管理テーブル」の11
:	:	:

【図15】

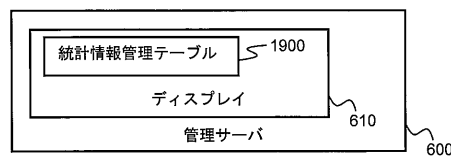
Fig.15

1900

統計情報管理テーブル	
統計情報種別	統計値
第一のチャンク管理テーブルの使用エントリ数	X1
第一のチャンク管理テーブルの空きエントリ数	X2
第二のチャンク管理テーブルの使用エントリ数	X3
第二のチャンク管理テーブルの空きエントリ数	X4
重複排除サーバ管理下のコンテンツの合計サイズ[MB]	X5
第一の記憶領域の使用容量[MB]	X6
第一の記憶領域の空き容量[MB]	X7
第二の記憶領域の使用容量[MB]	X8
第二の記憶領域の空き容量[MB]	X9
書き込みバッファ領域の使用容量[MB]	X10
書き込みバッファ領域の空き容量[MB]	X11

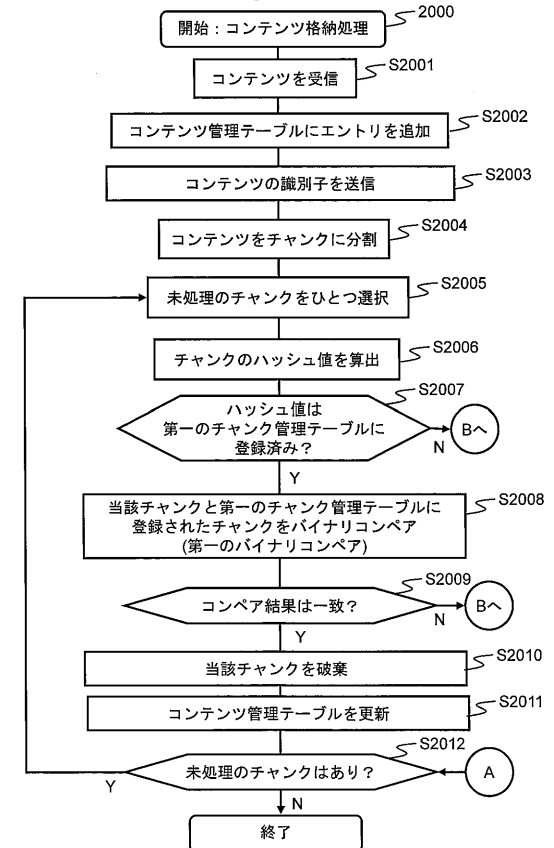
【図16】

Fig.16



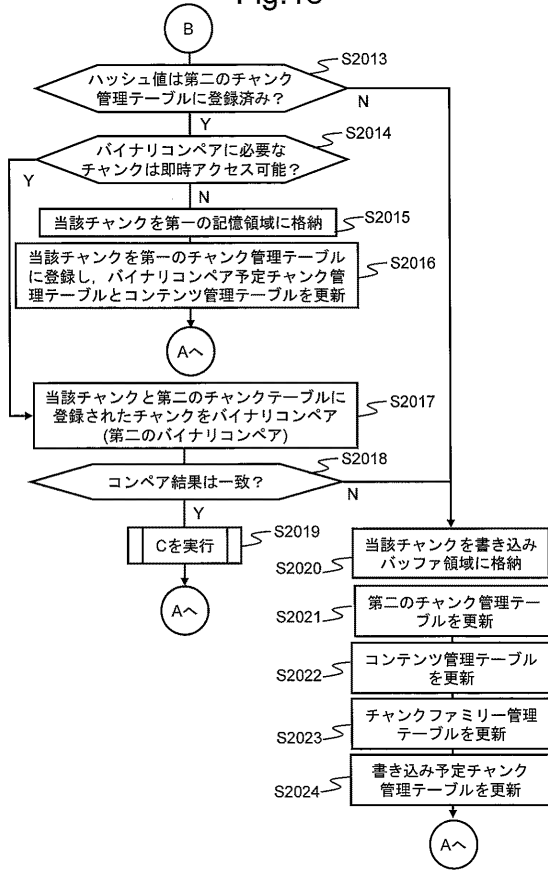
【図17】

Fig.17



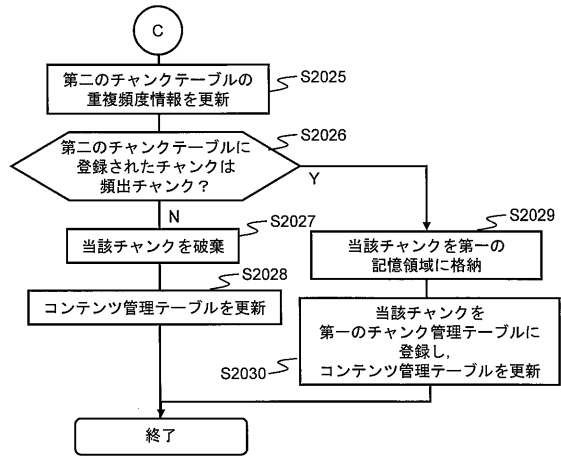
【 図 1 8 】

Fig.18



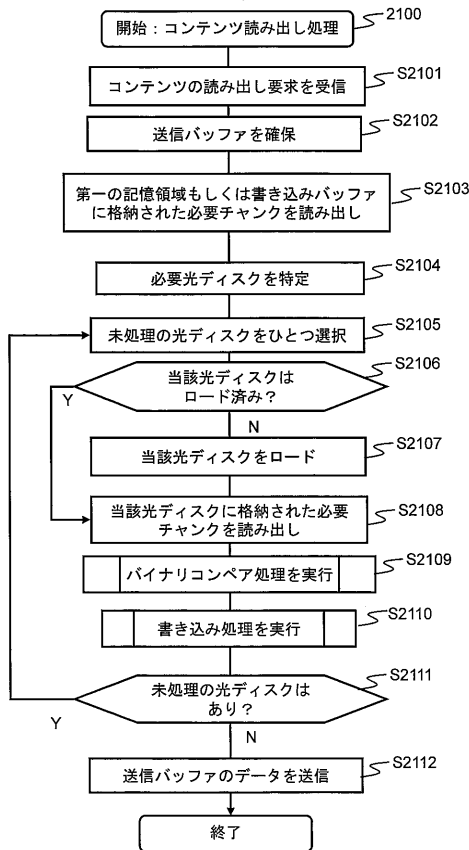
【 図 1 9 】

Fig.19



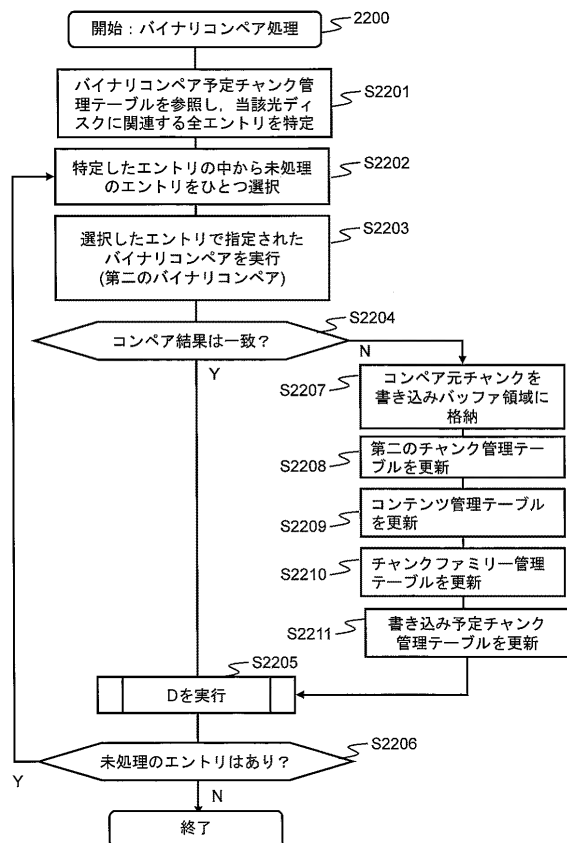
【 図 2 0 】

Fig.20

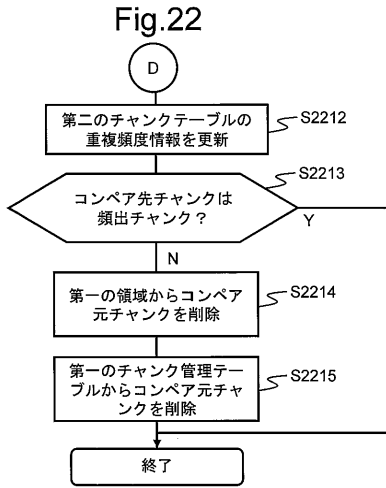


【 図 2 1 】

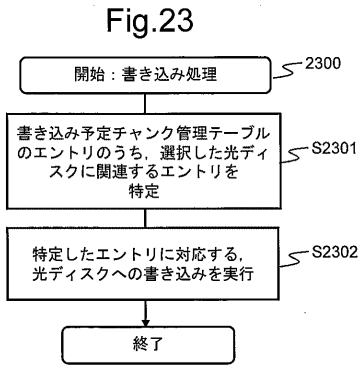
Fig.21



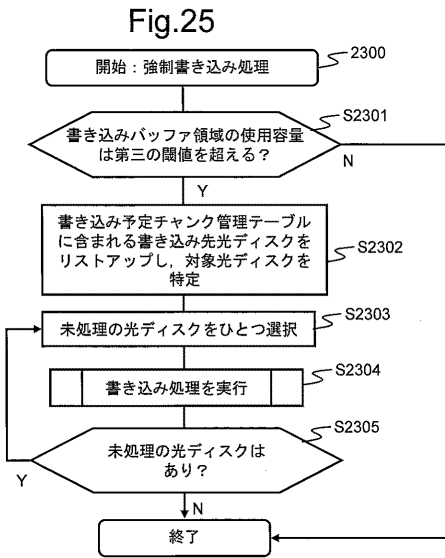
【 図 2 2 】



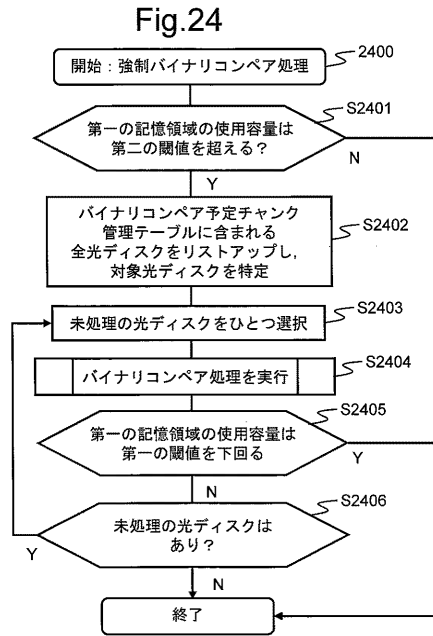
【 図 2 3 】



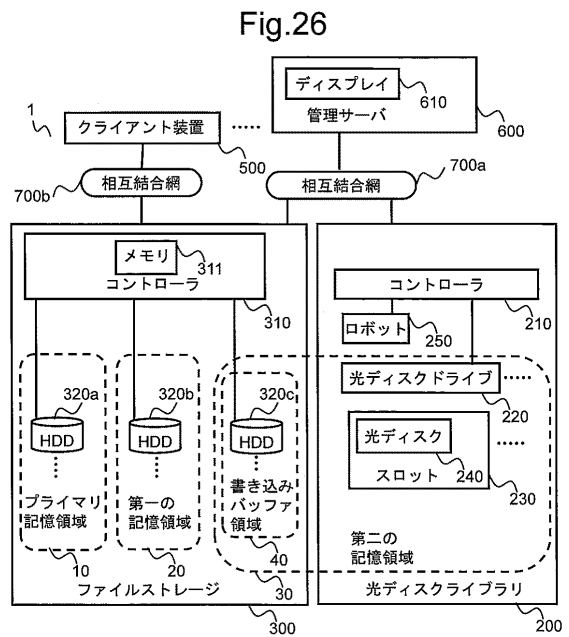
【 図 2 5 】



【 図 2 4 】

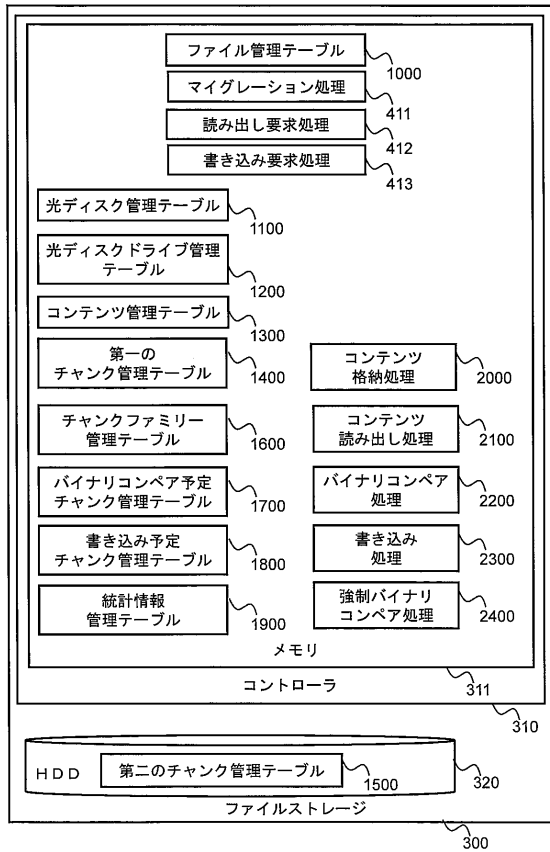


【 図 2 6 】



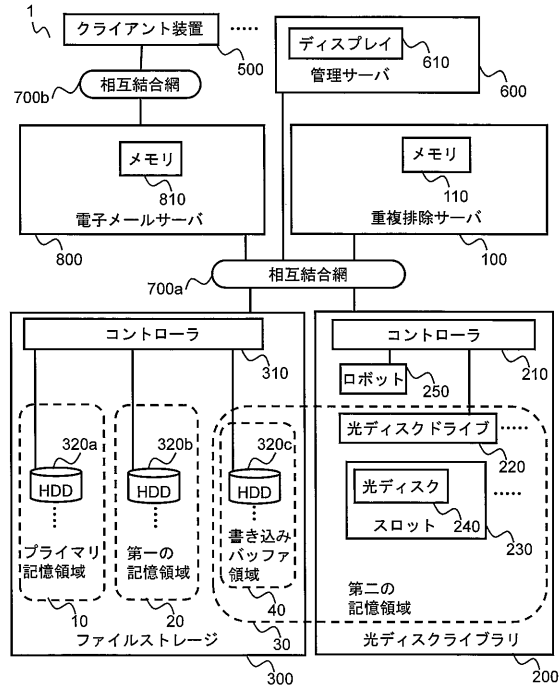
【 図 2 7 】

Fig.27



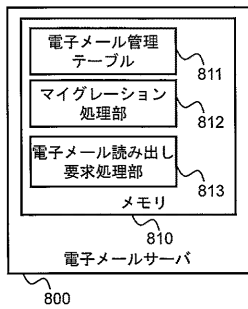
【 図 2 8 】

Fig.28



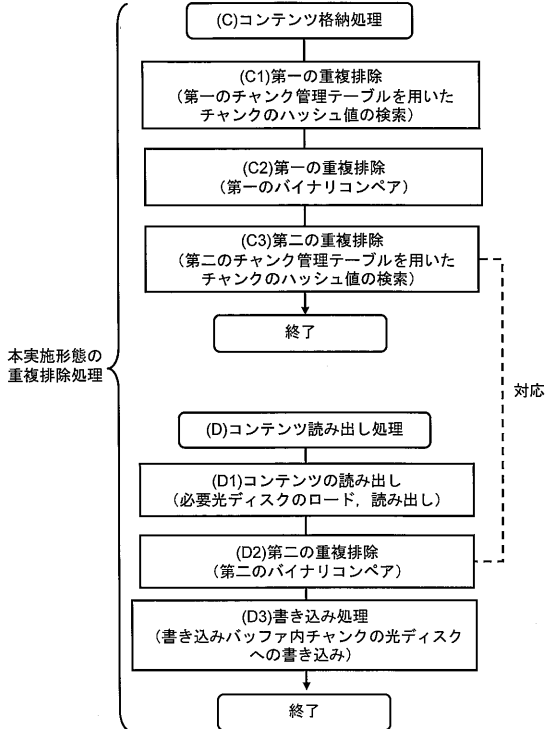
【 図 2 9 】

Fig.29



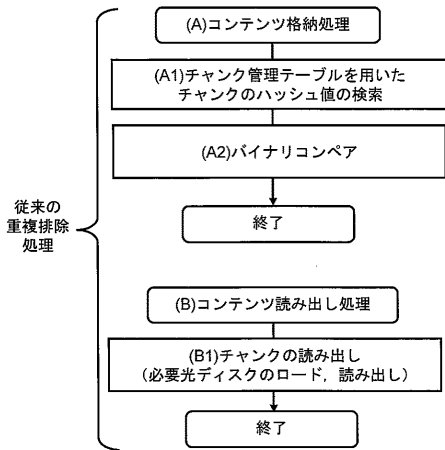
【 図 3 1 】

Fig.31



【 図 3 0 】

Fig.30



フロントページの続き

(72)発明者 渡辺 恭男

神奈川県横浜市戸塚区吉田町2-9-2番地 株式会社日立製作所横浜研究所内

(72)発明者 芹沢 一

神奈川県横浜市戸塚区吉田町2-9-2番地 株式会社日立製作所横浜研究所内