(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014 年 2 月 27 日(27.02.2014)

WIPO | PCT

(10) 国際公開番号
WO 2014/030252 A1

(51) 国際特許分類:
G06F 3/06 (2006.01) *G06F 12/00* (2006.01)

(21) 国際出願番号: PCT/JP2012/071424

(22) 国際出願日: 2012 年 8 月 24 日(24.08.2012)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人 (米国を除く全ての指定国について): 株式会社日立製作所 **(HITACHI, LTD.)** [JP/JP]; 〒1008280 東京都千代田区丸の内一丁目６番６号 Tokyo (JP). 株式会社日立情報通信エンジニアリング **(HITACHI INFORMATION & TELECOMMUNICATION ENGINEERING, LTD.)** [JP/JP]; 〒2206122 神奈川県横浜市西区みなとみらい二丁目３番３号 Kanagawa (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 岸 雅之 **(KISHI, Masayuki)** [JP/JP]; 〒2590180 神奈川県足柄上郡中井町境７８１番地 株式会社日立情報通信エンジニアリング内 Kanagawa (JP).

(74) 代理人: 大賀 眞司, 外 **(OHGA, Shinji** et al.); 〒1400002 東京都品川区東品川二丁目３番１２号 シーフォートスクエア センタービルディング １６階 サンネクスト国際特許事務所 Tokyo (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

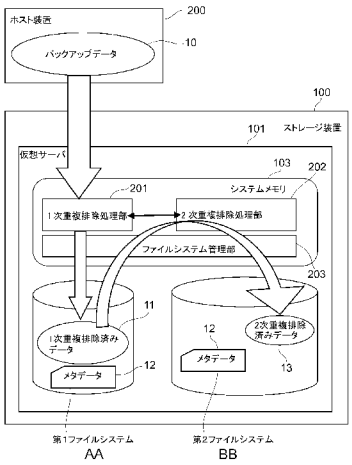
(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW,

[続葉有]

(54) **Title**: STORAGE DEVICE AND DATA MANAGEMENT METHOD

(54) 発明の名称 : ストレージ装置及びデータ管理方法

[図3]



10	Backup data
11	Primary deduplicated data
12	Metadata
13	Secondary deduplicated data
100	Storage device
101	Virtual server
103	System memory
200	Host device
201	Primary deduplication processing unit
202	Secondary deduplication processing unit
203	File system management unit
AA	First file system
BB	Second file system

(57) **Abstract**: [Problem] To efficiently perform deduplication processing by taking into consideration the advantages of two or more deduplication mechanisms. [Solution] A control unit for a storage device partitions received data into one or more chunks and compresses the partitioned chunks. The control unit subjects chunks with a compression ratio of less than or equal to a threshold to first deduplication processing by calculating hash values for the compressed chunks without storing the chunks in a first storage area and comparing the hash values with hash values of other data already stored in a second storage area. The control unit subjects chunks with a compression ratio greater than the threshold to second deduplication processing by reading out the compressed chunks from the first storage area after the compressed chunks have been stored in the first storage area, calculating hash values of the compressed chunks, and comparing the hash values with hash values of other data already stored in the second storage area.

(57) 要約: 【課題】２つ以上の重複排除機構の利点を考慮して効率的に重複排除処理を実行する。 【解決手段】ストレージ装置の制御部は、受信したデータを１または２以上のチャンクに分割し、分割した前記チャンクを圧縮し、圧縮率が閾値以下のチャンクに対し、前記第１記憶領域に格納せずに、圧縮された前記チャンクのハッシュ値を算出し、前記ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第１の重複排除処理を実行し、圧縮率が閾値より大きいチャンクに対し、前記第１記憶領域に圧縮された前記チャンクを格納した後に、前記圧縮された前記チャンクを前記第１記憶領域から読み出し、圧縮された前記チャンクのハッシュ値を算出し、該ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第２の重複排除処理を実行する。

WO 2014/030252 A1

MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:

— 国際調査報告（条約第 21 条(3)）

明　細　書

発明の名称 ： ストレージ装置及びデータ管理方法

技術分野

[0001] 本発明は、ストレージ装置及びデータ管理方法に関し、２つ以上の重複排除機構を利用して重複排除処理を行うストレージ装置及びデータ管理方法に適用して好適なるものである。

背景技術

[0002] ストレージ装置は、ホスト装置からの大規模データを記憶するために、大容量な記憶領域を保持している。ホスト装置からのデータは、年々増加の一途をたどっており、ストレージ装置のサイズやコストの問題から、大規模データを効率的に記憶する必要がある。そこで、記憶領域に格納するデータ量の増大を抑制し、データ容量効率を高めるため、データの重複を検出して排除するデータの重複排除処理が注目されている。

[0003] データの重複排除処理は、新たに記憶デバイスに書き込むデータ、いわゆるライトデータが、既に磁気ディスクに格納されているデータと同一内容の場合、重複するデータを磁気ディスクに書き込まない技術である。ライトデータが磁気ディスクに格納済みのデータと同一内容であるか否かは、一般的にデータのハッシュ値を用いて検証されている。

[0004] 従来、ホスト装置からのデータのすべてをディスクに記憶した後に重複排除処理を行う方式（以降、ポストプロセス方式とも称する）が採用されていた。しかし、ポストプロセス方式では、ホスト装置からのデータのすべてをディスクに書き込む必要があるため、大容量の記憶領域が必要となってしまう。そこで、ポストプロセス方式だけでなく、ディスクに書き込む前に重複排除処理を行う方式（以降、インライン方式とも称する）も併用して、重複排除処理を実行する技術が開示されている（例えば、特許文献１）。

先行技術文献

特許文献

[0005] 特許文献1：米国特許出願公開第２０１１／０２８９２８１号明細書

発明の概要

発明が解決しようとする課題

[0006] 特許文献１では、重複排除処理において、単にポストプロセス方式とインライン方式を併用することのみ開示されている。しかし、ポストプロセス方式ではすべてのデータを一旦ディスクに書き込むため、全体の処理性能がディスクの書き込み性能に依存してしまう。また、インライン方式では、データをディスクに書き込む際に重複排除処理を行うため、全体の処理性能が重複排除処理の性能に依存してしまう。そこで、両方式の利点を考慮して重複排除処理を実行する必要があった。また、ポストプロセス方式とインライン方式を併用した場合、両方式で同様の重複排除処理を実行してしまい、無駄な重複排除処理が発生してしまう可能性があるという問題があった。

[0007] そこで、２つ以上の重複排除機構の利点を考慮して効率的に重複排除処理を実行することが可能なストレージ装置及びデータ管理方法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に本発明においては、第１記憶領域と第２記憶領域とを提供する記憶装置と、前記記憶装置へのデータの入出力を制御する制御部と、を備え、前記制御部は、受信したデータを１または２以上のチャンクに分割し、分割した前記チャンクを圧縮し、圧縮率が閾値以下のチャンクに対し、前記第１記憶領域に格納せずに、圧縮された前記チャンクのハッシュ値を算出し、前記ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第１の重複排除処理を実行し、圧縮率が閾値より大きいチャンクに対し、前記第１記憶領域に圧縮された前記チャンクを格納した後に、前記圧縮された前記チャンクを前記第１記憶領域から読み出し、圧縮された前記チャンクのハッシュ値を算出し、該ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第２の重複排除処理を実行することを特徴とする、ストレージ装置が提供

される。

[0009] かかる構成によれば、受信したデータを１または２以上のチャンクに分割し、分割したチャンクを圧縮し、チャンクの圧縮率が所定の閾値以下の場合に、圧縮された該チャンクのハッシュ値を算出し、該ハッシュ値と既に格納されているデータのハッシュ値とを比較して第１の重複排除処理を実行し、チャンクの圧縮率が所定の閾値より大きい場合に、圧縮された該チャンクを第１のファイルシステムに格納した後に、圧縮された該チャンクのハッシュ値を算出し、該ハッシュ値と既に格納されているデータのハッシュ値とを比較して第２の重複排除処理を実行する。

[0010] これにより、重複排除処理のうち、処理負荷の小さいデータの分割処理を１次重複排除処理時に行うことができ、チャンクの圧縮率に基づいて、該チャンクを１次重複排除処理で重複排除を行うか、２次重複排除処理で重複排除処理を行うかを決定し、１次重複排除処理と２次重複排除処理のそれぞれの利点を考慮して効率的に重複排除処理を実行することが可能となる。

発明の効果

[0011] 本発明によれば、２つ以上の重複排除機構の利点を考慮して効率的に重複排除処理を実行することにより重複排除処理の負荷を分散することができる。

図面の簡単な説明

[0012] [図1]本発明の第１の実施形態に係る概要を説明する概念図である。

[図2]同実施形態にかかる計算機システムのハードウェア構成を示すブロック図である。

[図3]同実施形態にかかるストレージ装置のソフトウェア構成を示すブロック図である。

[図4]同実施形態にかかるメタデータについて説明する図表である。

[図5]同実施形態にかかるチャンクの管理情報を説明する概念図である。

[図6]同実施形態にかかる１次重複排除済みデータを示す概念図である。

[図7]同実施形態にかかるチャンクの圧縮ヘッダを説明する図表である。

[図8]同実施形態にかかるバックアップ処理を示すフローチャートである。

[図9]同実施形態にかかるデータの書き込み処理を示すフローチャートである。

[図10]同実施形態にかかる１次重複排除処理を示すフローチャートである。

[図11]同実施形態にかかる２次重複排除処理を示すフローチャートである。

[図12]同実施形態にかかるデータのRead処理を示すフローチャートである。

[図13]同実施形態にかかるデータのRead処理を示すフローチャートである。

[図14]本発明の第２の実施形態にかかるストレージ装置のソフトウェア構成を示すブロック図である。

発明を実施するための形態

[0013] 以下図面について、本発明の一実施の形態を詳述する。

[0014] （１）第１の実施の形態

（１－１）本実施の形態の概要

まず、図１を参照して、本実施の形態の概要について説明する。本実施形態では、ストレージ装置１００は、ホスト装置２００からのバックアップデータを記憶領域に格納する。なお、ホスト装置は、バックアップサーバ等のサーバ、他のストレージ装置であってもよい。ストレージ装置１００のバックアップデータの記憶領域として、バックアップデータを一時的に格納する記憶領域（第１ファイルシステム）と、重複排除処理実施後のバックアップデータの記憶領域（第２ファイルシステム）とが備えられている。

[0015] ストレージ装置１００は、第１ファイルシステムにバックアップデータを格納する際に、最初の重複排除処理（以降、１次重複排除処理と称して説明する。）を実行する。このように、ホスト装置２００からのバックアップデータを格納する前に重複排除処理を行う方式を、インライン方式と称する。

[0016] そして、ストレージ装置１００は、第１ファイルシステムに格納されたバックアップデータに対して、さらに重複排除処理（以降、２次重複排除処理と称して説明する。）を実行して、第２ファイルシステムにバックアップデータを格納する。このように、一旦バックアップデータを格納した後に重複

排除処理を行う方式をポストプロセス方式と称する。

[0017] ポストプロセス方式では、すべてのデータを一旦ディスクに書き込むため、全体の処理性能がディスクの書き込み性能に依存してしまう。さらに、ポストプロセス方式では、全てのデータを一旦ディスクに書き込むため、データ格納のため大きな記憶容量が消費されてしまう。また、インライン方式では、データをディスクに書き込む際に重複排除処理を行うため、全体の処理性能が重複排除処理の性能に依存してしまう。そこで、両方式の利点を考慮して重複排除処理を実行する必要がある。また、ポストプロセス方式とインライン方式を併用した場合、両方式で同様の重複排除処理を実行してしまい、無駄な重複排除処理が発生してしまう可能性があるという問題がある。

[0018] そこで、本実施の形態では、データの圧縮率に基づいて、該データを１次重複排除処理で重複排除を行うか、２次重複排除処理で重複排除処理を行うかを決定する。また、重複排除処理のうち、処理負荷の小さいデータの分割処理を１次重複排除処理時に行う。これにより、１次重複排除処理と２次重複排除処理のそれぞれの利点を考慮して効率的に重複排除処理を実行することが可能となる。また、圧縮率が閾値より低いデータに対してのみ１次重複排除処理を行うため、インライン方式での処理負荷を小さくしつつ、データの一時格納のための記憶領域の消費量を小さくできる。

[0019] （１－２）計算機システムの構成

次に、本実施の形態にかかる計算機システムのハードウェア構成について説明する。図２に示すように、計算機システムは、ストレージ装置１００とホスト装置２００とから構成されている。ホスト装置２００は、ＳＡＮ（Ｓｔｏｒａｇｅ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）などのネットワークを介してストレージ装置１００と接続されている。なお、図中には表記していないが、ストレージ装置１００をコントロールする管理端末を含んでもよい。

[0020] ストレージ装置１００は、ホスト装置２００から送信されたコマンドを解釈して、ディスクアレイ装置１１０の記憶領域内へのリード／ライトを実行する。ストレージ装置１００は、複数の仮想サーバ１０１ａ、１０１ｂ、１

０１ｃ・・・１０１ｎ（以降、仮想サーバ１０１と総称して説明する場合もある。）と、ファイバチャネルケーブル（図中ＦＣケーブルと表記）１０６と、ディスクアレイ装置１１０とから構成される。仮想サーバ１０１とディスクアレイ装置１１０とは、ファイバチャネルポート１０５、１０７に接続されたファイバチャネルケーブル１０６を介して接続されている。なお、本実施形態では仮想サーバを用いているが、物理サーバであってもよい。

[0021] 仮想サーバ１０１は、ストレージ装置１００内に仮想的に再現された計算機環境である。仮想サーバ１０１は、ＣＰＵ１０２、システムメモリ１０３、ＨＤＤ（Hard Disk Drive）１０４及びファイバチャネルポート（図中ＦＣポートと表記）１０５などを含む。

[0022] ＣＰＵ１０２は、演算処理装置として機能し、システムメモリ１０３に記憶されている各種プログラムや演算パラメータ等にしたがって、ストレージ装置１００全体の動作を制御する。システムメモリ１０３には、主に、１次重複排除処理を実行するプログラム及び２次重複排除処理を実行するプログラムが記憶されている。

[0023] ＨＤＤ１０４は、複数の記憶媒体から構成されている。例えば、ＳＳＤ（Solid State Disk）、ＳＣＳＩ（Small Computer System Interface）ディスク等の高価なハードディスクドライブ、または、ＳＡＴＡ（Serial AT Attachment）ディスク等の安価なハードディスクドライブでなる複数のハードディスクドライブから構成されてもよい。なお、本実施形態では、記憶媒体としてＨＤＤを用いているが、ＳＳＤ等の他の記憶媒体であってもよい。

[0024] 複数のＨＤＤ１０４により１つのＲＡＩＤ（Redundant Array of Inexpensive Disks）グループが構成され、１又は複数のＲＡＩＤグループが提供する物理的な記憶領域上に、１又は複数の論理ユニット（ＬＵ）が設定される。そしてホスト装置２００からのデータは、この論理ユニット（ＬＵ）内に所定大きさのブロックを単位として格納される。本実施の形態では、ディスクアレイ装置１１０の複数のＨＤＤ１０４から構成されるＬＵ０を第１のファイルシステムにマウントし、ＬＵ１を第２のファイルシステムにマウントし

て利用する。

[0025] ホスト装置２００は、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）などの演算装置や、メモリ、ディスクなどの記憶領域等の情報処理資源と、必要に応じて、キーボード、マウス、モニタディスプレイ、スピーカー、通信Ｉ／Ｆカード等の情報入出力装置を備えた計算機装置であり、例えばパーソナルコンピュータやワークステーション、メインフレーム等で構成される。

[0026] （１－３）ストレージ装置のソフトウェア構成

次に、図３を参照して、ストレージ装置１００のソフトウェア構成について説明する。図３に示すように、ストレージ装置１００のシステムメモリ１０３には、１次重複排除処理部２０１、２次重複排除処理部２０２及びファイルシステム管理部２０３などのプログラムが格納されている。なお、これらのプログラムは、ＣＰＵにより実行される。従って以下の説明において、これらのプログラムを主語として処理を説明している場合には、実際にはＣＰＵによりそのプログラムを実行することにより処理を実現することを意味する。

[0027] １次重複排除処理部２０１は、ホスト装置２００からのバックアップデータ１０を１次重複排除して第１ファイルシステムに格納する。２次重複排除処理部２０２は、第１ファイルシステムに格納された１次重複排除済みデータ１１を２次重複排除して第２ファイルシステムに格納する。

[0028] 本実施の形態では、１次重複排除処理部２０１により実行される１次重複排除処理と、２次重複排除処理部２０２により実行される２次重複排除処理とで、異なる重複排除処理を実行している。１次重複排除処理では、重複排除処理において負荷の小さいデータの分割処理と圧縮処理を行う。また、圧縮処理後のデータの圧縮率に基づいてデータのハッシュ値の計算と重複排除処理とを１次重複排除処理で実行するか２次重複排除処理で実行するか判定している。そして、２次重複排除処理では、１次重複排除処理においてハッシュ値の計算が行われなかったデータに対して重複排除処理を実行する。

[0029] 上記したように、バックアップデータのすべてをインライン方式である１次重複排除処理で行うと、重複排除処理に時間がかかり、ストレージ装置１００全体の処理性能が重複排除処理の性能に依存してしまう。また、バックアップデータのすべてをポストプロセス方式で重複排除した場合、すなわち、一旦第１ファイルシステムに格納した後、２次重複排除処理で重複排除処理を行う場合には、全体の処理性能がディスクの書き込み性能に依存してしまう。さらに、ポストプロセス方式では、全てのデータを一旦ディスクに書き込むため、データ格納のため大きな記憶容量が消費されてしまう。また、単純に、１次重複排除処理と２次重複排除処理を併用するだけでは、両処理で同様の重複排除処理を実行してしまい、無駄な重複排除処理が発生してしまう。

[0030] そこで、本実施の形態では、１次重複排除処理で、重複排除処理のうち負荷の小さいデータの分割処理と圧縮処理を行い、さらに、圧縮率の低い分割データ（一時的なデータ記憶領域の容量を大きく消費するデータ）に対して重複判定処理を実行する。ここで、１次重複排除処理において分割されたデータをチャンクと称して以下説明する。データの分割処理については、後で詳細に説明する。

[0031] 重複排除処理における重複判定処理は、分割されたデータ（チャンク）の圧縮率に関わらず、ほぼ同等の時間がかかる。したがって、１次重複排除処理では、圧縮率の低いチャンクに対して重複判定処理を実行することにより、重複判定処理の負荷を軽減しつつ、データの書き込み処理を高速化させることが可能となる。さらに、圧縮率の低いチャンクをインライン方式にて重複排除処理することにより、データの一時格納のための記憶領域の消費量を小さくできる。

[0032] 一方、２次重複排除処理では、１次重複排除処理で既に重複判定処理を実行したチャンク以外のチャンクに対して重複判定処理を実行することにより、１次重複排除処理と２次重複排除処理とで同様の重複排除処理を実行することを防止している。具体的に、１次重複排除処理において重複判定処理が

実行されたチャンクについては、各チャンクのデータヘッダに、既に重複判定処理を実行したことを示すフラグを設定する。そして、２次重複排除処理において、設定されたフラグを参照して、１次重複排除処理で重複判定処理が実行されていないチャンクに対して重複判定処理を実行する。

[0033] 次に、図４を参照して、第１ファイルシステム及び第２ファイルシステムに格納されているメタデータ１２について説明する。メタデータ１２は、第１ファイルシステムに格納される１次重複排除済みデータまたは第２ファイルシステムに格納される２次重複排除済みデータの管理情報を示すデータである。

[0034] 図４に示すように、メタデータ１２には各種テーブルが含まれる。具体的に、スタブファイル（Stub file）１２１、チャンクデータセット（Chunk Data Set）１２２、チャンクデータセットインデックス（Chunk Data Set Index）１２３、コンテンツ管理テーブル１２４及びチャンクインデックス１２５などのテーブルがメタデータ１２に含まれる。

[0035] スタブファイル１２１は、バックアップデータとコンテンツＩＤとを関連付けるためのテーブルである。バックアップデータは、複数のファイルデータから構成される。当該ファイルデータを、記憶領域に格納される単位である論理的にまとまったコンテンツ（content）と称する。各コンテンツは複数のチャンクに分割され、各コンテンツは、コンテンツＩＤによって識別される。このコンテンツＩＤがスタブファイル１２１に格納される。ストレージ装置１００がディスクアレイ装置１１０に格納されたデータのリード／ライトを行う場合には、まず、スタブファイル１２１のコンテンツＩＤが呼び出される。

[0036] チャンクデータセット１２２は、複数のチャンクから構成されるユーザデータであって、ストレージ装置１００に格納されるバックアップデータである。チャンクデータセットインデックス１２３は、チャンクデータセット１２２に含まれる各チャンクの情報が格納されている。具体的に、チャンクデータセットインデックス１２３には、各チャンクの長さ情報とチャンクデー

タとが対応付けて格納されている。

[0037] コンテンツ管理テーブル１２４は、コンテンツ内のチャンク情報を管理するテーブルである。ここで、コンテンツとは、上記したコンテンツＩＤによって識別されるファイルデータである。また、チャンクインデックス１２５は、各チャンクがどのチャンクデータセット１２２に存在するかを示す情報である。また、チャンクインデックス１２５には、各チャンクを識別するチャンクのフィンガープリントと、チャンクが存在するチャンクデータセット１２２を識別するチャンクデータセットＩＤとが対応付けられている。

[0038] 次に、図５を参照して、チャンクの管理情報について詳細に説明する。図５に示すように、スタブファイル（図中Stub fileと表記）１２１には、オリジナルのデータファイルを識別するコンテンツＩＤ（図中Content IDと表記）が格納されている。そして、１つのスタブファイル１２１に対して１つのコンテンツファイルが対応し、各コンテンツファイルはコンテンツ管理テーブル（図中Content Mng Tblと表記）１２４で管理されている。

[0039] コンテンツ管理テーブル１２４で管理されている各コンテンツファイルは、コンテンツＩＤ（図中Content IDと表記）によって識別される。コンテンツファイルには、各チャンクのオフセット（Content Offset）、チャンク長（Chunk Length）、チャンクが存在するコンテナの識別情報（Chunk Data Set ID）、各チャンクのハッシュ値（Fingerprint）が格納されている。

[0040] また、チャンクデータセットインデックス（図中Chunk Data Set Indexと表記）１２３には、チャンクの管理情報として、チャンクデータセット（図中Chunk Data Setと表記）１２２に格納されているチャンクのハッシュ値（Fingerprint）と、チャンクのオフセット及びデータ長とが対応付けられて格納されている。各チャンクデータセット１２２は、チャンクデータセットＩＤ（図中Chunk Data Set IDと表記）で識別されている。チャンクデータセットインデックス１２３では、チャンクの管理情報がチャンクデータセットごとにまとめられて管理されている。

[0041] チャンクデータセット１２２は、所定数のチャンクを１つのコンテナとし

て管理している。各コンテナは、チャンクデータセットＩＤにより識別され、各コンテナには、チャンク長が付された複数のチャンクデータが含まれる。チャンクデータセット１２２のコンテナを識別するチャンクデータセットＩＤと、上記したチャンクデータセットインデックス１２３のチャンクデータセットＩＤとが対応づけられている。

[0042] チャンクインデックス１２５は、各チャンクのハッシュ値（Fingerprint）とチャンクが存在するコンテナの識別情報（Chunk Data Set ID）とが対応づけられて格納されている。チャンクインデックス１２５は、重複排除処理を実行する際に、各チャンクから計算されたハッシュ値をもとに、いずれのコンテナに格納されているかを判定するためのテーブルである。

[0043] 上記したように、バックアップデータであるコンテンツは、１次重複排除処理において、複数のチャンクに分割される。コンテンツは、通常のファイルの他、例えば、アーカイブファイル、バックアップファイルまたは仮想ボリュームファイルなどの通常のファイルを集約したファイルなどを例示することができる。

[0044] 重複排除処理は、コンテンツからチャンクを順次切り出す処理と、切り出したチャンクの重複の有無を判定する処理と、チャンクの格納保存処理とからなる。重複排除処理を効率よく実行するためには、チャンクの切り出し処理において、内容が同一のデータセグメントをより多く切り出すことが重要となる。

[0045] チャンクの切り出し方法としては、固定長チャンク切り出し方式、可変長チャンク切り出し方式などがある。固定長チャンク切り出し方式は、例えば、４キロバイト（ＫＢ）や１メガバイト（ＭＢ）といった一定の長さのチャンクを順次切り出す方法である。また、可変長チャンク方式は、コンテンツデータの局所的な条件をもとにチャンクの切り出しの境界を決定してコンテンツを切り出す方法である。

[0046] しかし、固定長チャンク切り出し方式は、チャンクを切り出すためのオーバーヘッドが小さいが、コンテンツデータの変更がデータの挿入などの変更

の場合、データが挿入された後のチャンクがずれて切り出されるため、重複排除効率が低下してしまう。一方、可変長チャンク切り出し方式は、データが挿入されてチャンクがずれてもチャンクを切り出すための境界の位置は変わらないため重複排除効率を上げることができるが、チャンクの境界を探索するための処理のオーバーヘッドが大きくなってしまう。また、基本データ切り出し方式では、基本データを切り出すために伸長処理を繰り返す必要があり、重複排除処理のオーバーヘッドが大きくなるという問題があった。

[0047] したがって、重複排除効率と重複排除処理のオーバーヘッドのトレードオフを考慮すると、上記したチャンク切り出し方式のうち、いずれか一つのチャンク切り出し方式を用いて重複排除処理を行っても、重複排除処理全体の最適化を図ることができないという問題があった。

[0048] そこで、本実施の形態では、各コンテンツ、あるいはコンテンツの各部分の特性に基づいて、チャンクの切り出し処理において適用するチャンク切り出し方式を切り替えることにより、各コンテンツの種別に応じて最適なチャンク切り出し方式を選択する。コンテンツの種別は、各コンテンツに付加されている種別を識別する情報を検出することにより判定することが可能となる。コンテンツの種別に対応するコンテンツの特性や構造を予め知っておくことにより、コンテンツの種別に応じて最適なチャンク切り出し方式を選択することが可能となる。

[0049] 例えば、あるコンテンツについて、変更があまりない種別であれば、当該コンテンツについては固定長チャンク方式を適用してチャンクを切り出すことが好適である。また、サイズの大きいコンテンツの場合には、チャンクサイズを大きく取ったほうが、処理オーバーヘッドが小さくなり、サイズの小さいコンテンツの場合には、チャンクサイズを小さく取ることが好ましい。また、コンテンツへの挿入がある場合には、可変長チャンク方式を適用してチャンクを切り出すことが好適である。コンテンツへの挿入があるが、変更が少ない場合には、チャンクのサイズを大きめに取ることにより、重複排除効率を低下させずに、処理効率を向上させて管理オーバーヘッドを低減させ

ることが可能となる。

[0050] また、所定の構造を有するコンテンツは、ヘッダ部、ボディ部、トレイラ部などの各部に分けることができ、部分毎に適用すべきチャンク切り出し方式が異なる。各部分に好適なチャンク切り出し方式を適用することにより、重複排除効率と処理効率とを最適化することが可能となる。

[0051] 上記したように、１次重複排除処理部２０１は、コンテンツを複数のチャンクに切り出し、各チャンクを圧縮する。１次重複排除処理部２０１は、図６に示すように、まず、コンテンツをヘッダ部(図中Metaと表記)とボディ部（図中FileXと表記）に分割する。そして、１次重複排除処理部２０１は、さらに、ボディ部を固定長または可変長に分割する。コンテンツを固定長で分割する場合には、例えば、４キロバイト（ＫＢ）や１メガバイト（ＭＢ）といった一定の長さのチャンクを順次切り出す。また、コンテンツを可変長で分割する場合には、コンテンツの局所的な条件をもとにチャンクの切り出しの境界を決定してチャンクを切り出す。また、例えば、vmdkファイル、vdiファイル、vhdファイル、zipファイルまたはgzipファイルなどコンテンツの構造に変更があまりないファイルを固定長に分割し、これらのファイル以外のファイルを可変長に分割する。

[0052] そして、１次重複排除処理部２０１は、分割したチャンクを圧縮し、圧縮率の低いチャンク（圧縮率が閾値よりも低いチャンク）に対して１次重複排除処理を行う。１次重複排除処理部２０１は、１次重複判定処理の対象となるチャンクのハッシュ値を算出して、該ハッシュ値をもとに同一チャンクがＨＤＤ１０４に既に格納されているかを判定する。１次重複排除処理部２０１は、１次重複排除処理を行った結果、既にＨＤＤ１０４に格納されているチャンクを排除して、第１ファイルシステムに格納するための１次重複排除済みデータを生成する。１次重複排除処理部２０１は、圧縮した各チャンクに圧縮後のデータ情報を示す圧縮ヘッダを付して管理する。なお、一次重複排除処理（インライン方式）において、圧縮率が閾値よりも高いチャンクのハッシュ値の算出及び重複排除処理を実行しない。

[0053] 次に、チャンクの圧縮ヘッダについて説明する。図７は、圧縮された各チャンクに付される圧縮ヘッダを説明する概念図である。図７に示すように、圧縮ヘッダは、マジックナンバー３０１、ステータス３０２、フィンガープリント３０３、チャンクデータセットＩＤ３０４、圧縮前length３０５及び圧縮後length３０６を含む。

[0054] マジックナンバー３０１には、１次重複排除処理済みのチャンクであることを示す情報が格納される。ステータス３０２には、チャンクが重複判定処理を実行されたかを示す情報が格納される。例えば、ステータス３０２にステータス１が格納されている場合には、重複判定未実施であることを示す。ステータス３０２にステータス２が格納されている場合には、重複判定実施済みであり、未だＨＤＤ１０４に格納されていない新規チャンクであることを示す。また、ステータス３０２にステータス３が格納されている場合には、重複判定実施済みであり、既にＨＤＤ１０４に格納されている既存チャンクであることを示す。

[0055] フィンガープリント３０３には、チャンクから算出されたハッシュ値が格納される。なお、１次重複排除処理において、重複判定処理が行われなかったチャンクについては、フィンガープリント３０３には無効な値が格納される。すなわち、ステータス１のチャンクについては、未だ重複判定処理が実行されていないため、フィンガープリント３０３には無効値が格納される。

[0056] チャンクデータセットＩＤ３０４には、チャンク格納先のチャンクデータセットＩＤが格納される。チャンクデータセットＩＤ３０４は、チャンクを格納するコンテナ（Chunk Data Set１２２）を識別する情報である。なお、１次重複排除処理が実行されていないチャンクや未だＨＤＤ１０４に格納されていない新規チャンクについては、チャンクデータセットＩＤ３０４に無効な値が格納される。すなわち、ステータス１やステータス２のチャンクのチャンクデータセットＩＤ３０４には、無効値が格納される。

[0057] 圧縮前length３０５には、圧縮前のチャンク長が格納される。圧縮後length３０６には、圧縮後のチャンク長が格納される。

[0058] ２次重複排除処理部２０２は、１次重複排除処理部２０１により生成された１次重複排除データに含まれるチャンクの圧縮ヘッダを参照して、各チャンクの重複判定処理を実行するかを判定する。具体的に、２次重複排除処理部２０２は、チャンクの圧縮ヘッダのステータスを参照し、重複判定処理を行うか否か判断する。

[0059] 例えば、チャンクの圧縮ヘッダのステータス３０２がスタータス１の場合には、１次重複排除処理において重複判定処理が実行されていないため、２次重複排除処理において重複判定処理を実行する。また、チャンクの圧縮ヘッダのステータス３０２がステータス２の場合には、１次重複判定処理において重複判定処理は実行されているが、チャンクデータセット１２２には格納されていないチャンクであるため、チャンクの格納先を決定して該チャンクを書き込む。また、チャンクの圧縮ヘッダのステータス３０２がステータス３の場合には、１次重複判定処理において重複判定処理が実行され、既にチャンクデータセット１２２に格納されているチャンクであるため、重複判定処理は実行せずに、チャンクの格納先を取得する。

[0060] 上記したように、１次重複排除処理部２０１は、重複排除処理のうち負荷のかからない分割処理や圧縮処理を行い、圧縮率の低いチャンクに対してハッシュ値の計算及び重複判定処理を行う。そして、２次重複排除処理部２０２は、各チャンクの圧縮ヘッダを参照して、１次重複排除処理部２０２により重複判定処理が行われていないチャンクに対して重複判定処理を実行する。これにより、重複判定処理の負荷を軽減しつつ、データの書き込み処理を高速化させることが可能となる。さらに、圧縮率の低い（データサイズの大きい）チャンクをインライン方式にて重複排除処理することにより、データの一時格納のための記憶領域の消費量を小さくできる。

[0061] （１－４）重複排除処理

本実施の形態にかかる重複排除処理は、ホスト装置２００からの要求に応じてデータのバックアップを開始する。ストレージ装置１００におけるデータのバックアップ処理は、図８に示すように、まず、データの書き込み先を

オープンして（Ｓ１０１）、バックアップデータのサイズ分データの書き込み処理（Ｓ１０３）を繰り返す（Ｓ１０２～Ｓ１０４）。ストレージ装置１００は、データの書き込み処理終了後、書き込み先をクローズして（Ｓ１０５）バックアップ処理を終了する。

[0062] 上記したステップＳ１０３におけるデータの書き込み処理において、図９に示すように、ストレージ装置１００は、ホスト装置２００からのバックアップデータをメモリ上のバッファに滞留させる（Ｓ１１１）。

[0063] そして、ストレージ装置１００は、バッファに規定量のデータが溜まったかを判定する（Ｓ１１２）。ステップＳ１１２において、バッファに規定量のデータが溜まったと判定された場合には、１次重複排除処理部２０１に１次重複排除処理を実行させる。一方、ステップ１１２において、バッファに規定量のデータが溜まっていないと判定された場合には、さらにバックアップデータを受領する（Ｓ１０２）。

[0064] （１－４－１）１次重複排除処理の詳細

次に、図１０を参照して、１次重複排除処理部２０１による１次重複排除処理の詳細について説明する。図１０に示すように、１次重複排除処理部２０１は、バッファに滞留したデータについて、バッファサイズ分ステップＳ１２１～ステップＳ１３７までの処理を繰り返す。

[0065] １次重複排除処理部２０１は、上記した分割処理により、バッファから固定長または可変長で１チャンクを切り出す（Ｓ１２２）。そして、１次重複排除処理部２０１は、ステップＳ１２２において切り出したチャンクを圧縮して（Ｓ１２３）、チャンクの圧縮率を算出する（Ｓ１２４）。

[0066] そして、１次重複排除処理部２０１は、変数FingerPrintにnull値を代入し（Ｓ１２５）、変数ChunkDataSetIDにnull値を代入する（Ｓ１２６）。

[0067] 続いて、１次重複排除処理部２０１は、ステップＳ１２４において算出したチャンクの圧縮率が所定の閾値より低いか否かを判定する（Ｓ１２７）。ステップＳ１２７において、チャンクの圧縮率が所定の閾値より低い場合とは、圧縮前後でチャンク長があまり変わらない場合である。

[0068] ステップＳ１２７において、チャンクの圧縮率が所定の閾値より低いと判定された場合には、ステップＳ１２８以降の処理を実行する。一方、ステップＳ１２７において、チャンクの圧縮率が所定の閾値より高いと判定された場合には、ステップＳ１３１以降の処理を実行する。

[0069] ステップＳ１２８において、１次重複排除処理部２０１は、チャンクのデータからハッシュ値を算出して、算出結果を変数FingerPrintに代入する（Ｓ１２８）。

[0070] そして、１次重複排除処理部２０１は、算出したハッシュ値を用いて、チャンクがチャンクデータセットに格納されているか、格納されている場合にはチャンクデータセットのチャンクデータセットＩＤ（ChankDataSetID）を確認する（Ｓ１２９）。

[0071] そして、１次重複排除処理部２０１は、重複判定処理の対象となるチャンクと同一のチャンクがチャンクデータセットに格納されているかを判定する（Ｓ１３０）。ステップＳ１３０において、同一のチャンクがあると判定された場合には、１次重複排除処理部２０１は、ステップＳ１３５以降の処理を実行する。一方、ステップＳ１３０において同一のチャンクがないと判定された場合には、ステップＳ１３３以降の処理を実行する。

[0072] ステップＳ１２７において、圧縮率が閾値よりも高いと判定された場合には、１次重複排除処理部２０１は、重複判定処理を実行せずに、ステータス１のチャンクヘッダを生成する（Ｓ１３１）。ステータス１のチャンクヘッダとは、上記したように、重複判定未実施のチャンクに付される圧縮ヘッダである。図７に示すように、チャンクヘッダがステータス１の場合、チャンクとチャンクヘッダとが第１ファイルシステムに書き込まれる。なお、重複判定処理が実施されていないため、チャンクヘッダのフィンガープリント３０３とチャンクデータセットＩＤ３０４はｎｕｌｌ値のままである。

[0073] また、ステップＳ１２７において、圧縮率が閾値よりも低いと判定され、重複判定処理が実行された結果、同一チャンクがチャンクデータセット１２２に存在しないと判定された場合には、ステータス２のチャンクヘッダを生

成する（Ｓ１３３）。ステータス２のチャンクヘッダとは、上記したように、重複判定が実施済みであり、チャンクデータセット１２２に同一チャンクがない場合にチャンクに付される圧縮ヘッダである。図７に示すように、チャンクヘッダがステータス２の場合、チャンクとチャンクヘッダとが第１ファイルシステムに書き込まれる（Ｓ１３４）。なお、チャンクヘッダのフィンガープリント３０３には、チャンクから算出したハッシュ値が格納される。また、チャンクデータセットＩＤ３０４は、チャンクが未だ見つかっていないため、ｎｕｌｌ値のままである。

[0074] また、ステップＳ１２７において、圧縮率が閾値よりも低いと判定され、重複判定処理が実行された結果、同一チャンクがチャンクデータセット１２２に存在すると判定された場合には、ステータス３のチャンクヘッダを生成する（Ｓ１３５）。ステータス３のチャンクヘッダとは、上記したように、重複判定が実施済みであり、チャンクデータセット１２２に同一チャンクがある場合にチャンクに付される圧縮ヘッダである。図７に示すように、チャンクヘッダがステータス３の場合、チャンクヘッダのみ第１ファイルシステムに書き込まれる（Ｓ１３６）。つまり、チャンクのデータ自体は第１ファイルシステムに書き込まれず、記憶容量を削減することができる。

[0075] （１－４－２）２次重複排除処理の詳細

以上、１次重複排除処理の詳細について説明した。次に、図１１を参照して、２次重複排除処理部２０２による２次重複排除処理の詳細について説明する。２次重複排除処理は、所定時間ごとに定期的に実行するようにしてもよいし、予め決められたタイミングで実行するようにしてもよいし、管理者の入力に応じて実行するようにしてもよい。さらに、第１ファイルシステムの容量が一定量を超えた場合に、実行を開始してもよい。

[0076] 図１１に示すように、２次重複排除処理部２０２は、まず、変数offsetに０を代入する（Ｓ２０１）。続いて、１次重複排除済みファイル（第１ファイルシステム）をオープンして、１次重複済みファイル分、２次重複排除処理を繰り返す（Ｓ２０３～Ｓ２２２）。

[0077] ステップＳ２０２において、１次重複排除済みファイルをオープンした２次重複排除処理部２０２は、変数offsetに代入された値からチャンクヘッダサイズ分のデータを読み出す（Ｓ２０４）。そして、２次重複排除処理部２０２は、チャンクヘッダの変数Lengthの値から、圧縮後のチャンク長を取得する（Ｓ２０５）。さらに、２次重複排除処理部は、チャンクヘッダの変数FingerPrintから、チャンクのハッシュ値（フィンガープリント）を取得する（Ｓ２０６）。なお、１次重複排除処理において未だ１次重複判定処理が未実施の場合には、チャンクヘッダのFingerPrintに無効な値（null）が格納されている。

[0078] 続いて、２次重複排除処理部２０２は、チャンクのチャンクヘッダに含まれるステータス（Status）を確認する（Ｓ２０７）。ステップＳ２０７において、ステータスがステータス１の場合、すなわち、対象となるチャンクが重複判定未実施である場合、２次重複排除処理部２０２は、ステップＳ２０８以降の処理を実行する。また、ステップＳ２０７において、ステータスがステータス２の場合、すなわち、対象となるチャンクが１次重複排除処理により重複判定済であるが、チャンクデータセット１２２にチャンクが存在しない場合、２次重複排除処理部２０２は、重複排除処理を実行せずにステップＳ２１６以降の処理を実行する。また、ステップＳ２０７において、ステータスがステータス３の場合、すなわち、対象となるチャンクが１次重複排除処理により重複判定済であり、チャンクデータセット１２２にチャンクが存在する場合、２次重複排除処理部２０２は、重複排除処理を実行せずにステップＳ２２４の処理を実行する。

[0079] 次に、チャンクヘッダのステータスがステータス１の場合、すなわち、重複判定未実施の場合の処理について説明する。２次重複排除処理部２０２は、offsetの値にチャンクヘッダサイズを加算した長さ分のデータを読み出す（Ｓ２０８）。そして、ステップＳ２０８において読み出したチャンクのデータからハッシュ値（FingerPrint）を算出する（Ｓ２０９）。

[0080] 次に、２次重複排除処理部２０２は、ステップＳ２０９において算出したF

ingerPrintをもとに、チャンクデータセット１２２のチャンクの有無を確認して（Ｓ２１０）、チャンクデータセット１２２に対象となるチャンクと同一のチャンクが存在するか判定する（Ｓ２１１）。

[0081] ステップＳ２１１において、チャンクデータセット１２２に同一のチャンクが存在すると判定された場合には、２次重複排除処理部２０２は、変数ChunkDataSetIDに既に格納されている同一のチャンクの格納先のチャンクデータセットＩＤ（ChunkDataSetID）と同じＩＤを代入して（Ｓ２１２）、ステップＳ２２０以降の処理を実行する。

[0082] 一方、ステップＳ２１１において、チャンクデータセット１２２に同一のチャンクが存在しないと判定された場合には、２次重複排除処理部２０２は、チャンクを格納する格納先のチャンクデータセット（ChunkDataSet）１２２を決定して、決定した該チャンクデータセット１２２のチャンクデータセットＩＤを変数ChunkDataSetIDに代入する（Ｓ２１３）。

[0083] そして、２次重複排除処理部２０２は、チャンクデータセット（ChunkDataSet）１２２にチャンクヘッダとチャンクデータを書き込む（Ｓ２１４）。さらに、２次重複排除処理部２０２は、ステップＳ２０９において変数FingerPrintに代入した値とステップＳ２１３において変数ChunkDataSetIDに代入した値をチャンクインデックス１２５に登録して（Ｓ２１５）、ステップＳ２２０以降の処理を実行する。

[0084] 次に、チャンクヘッダのステータスがステータス２の場合、すなわち、重複判定実施済みであるが、チャンクデータセット１２２にチャンクが存在しない場合の処理について説明する。２次重複排除処理部２０２は、offsetの値にチャンクヘッダサイズを加算した長さ分のデータを読み出す（Ｓ２１６）。

[0085] そして、２次重複排除処理部２０２は、チャンクを格納する格納先のチャンクデータセット（ChunkDataSet）１２２を決定して、決定した該チャンクデータセット１２２のチャンクデータセットＩＤを変数ChunkDataSetIDに代入する（Ｓ２１７）。

[0086] そして、２次重複排除処理部２０２は、チャンクデータセット（ChunkDataSet）１２２にチャンクヘッダとチャンクデータを書き込む（Ｓ２１８）。さらに、２次重複排除処理部２０２は、ステップＳ２０６においてFingerPrintに代入した値と、ステップＳ２１７において変数ChunkDataSetIDに代入した値をチャンクインデックス１２５に登録して（Ｓ２１９）、ステップＳ２２０以降の処理を実行する。

[0087] 次に、チャンクヘッダのステータスがステータス３の場合、すなわち、重複判定実施済みであり、チャンクデータセット１２２にチャンクが存在する場合の処理について説明する。２次重複排除処理部２０２は、チャンクヘッダからチャンクデータセットＩＤ（ChunkDataSetID）を取得して、変数ChunkDataSetIDに代入する（Ｓ２２４）。そして、２次重複排除処理部２０２は、ステップＳ２２０以降の処理を実行する。なお、チャンクヘッダに格納されているチャンクデータセットＩＤ（ChunkDataSetID）は、一次重複排除処理において重複排除されたデータと同一のデータであって、既に格納されているデータの格納先を示すＩＤである。

[0088] そして、２次重複排除処理部２０２は、コンテンツ管理テーブル１２４に、チャンク長（Length）、オフセット（Offset）、フィンガープリント（FingerPrint）、チャンクデータセットＩＤ（ChunkDataSetID）を設定する（Ｓ２２０）。

[0089] そして、変数Offsetの値にチャンクヘッダのサイズとチャンク長（Length）とを加算して、変数Offsetに代入する（Ｓ２２１）。

[0090] ステップＳ２０３～ステップＳ２２の処理を１次重複排除済みファイルのサイズ分繰り返した後、１次重複排除済みファイルをクローズして（Ｓ２２３）、２次重複排除処理を終了する。

[0091] （１－５）Ｒｅａｄ処理の詳細

次に、図１２を参照して、１次重複排除処理及び２次重複排除処理が行われたデータのＲｅａｄ処理について説明する。重複排除済みデータのＲｅａｄ処理は、１次重複排除処理部２０１及び２次重複排除処理部２０２によっ

て行われる。

[0092] 図１２に示すように、１次重複排除処理部２０２は、まず、Ｒｅａｄ対象が２次重複排除済みのデータであるかを判定する（Ｓ３０１）。例えば、１次重複排除処理部２０２は、当該データがスタブ化されている場合に、当該データが２次重複排除済みのデータであると判定する。

[0093] ステップＳ３０１において、Ｒｅａｄ対象のデータが２次重複排除済みであると判定された場合には、２次重複排除済みデータのＲｅａｄ処理を実行する（Ｓ３０２）。一方、ステップＳ３０１において、Ｒｅａｄ対象のデータが２次重複排除済みではないと判定された場合には、ステップＳ３０３以降の処理を実行する。

[0094] 図１３に、２次重複排除済みデータのＲｅａｄ処理の詳細を示す。図１３に示すように、２次重複排除処理部２０２は、コンテンツデータのコンテンツＩＤ（content ID）に対応するコンテンツ管理テーブル１２４を読み出す（Ｓ３１１）。

[0095] そして、２次重複排除処理部２０２は、コンテンツのチャンクの数分ステップＳ３１２～ステップＳ３１８の処理を繰り返す。

[0096] まず、２次重複排除処理部２０２は、コンテンツ管理テーブル１２４からフィンガープリント（FingerPrint）を取得する（Ｓ３１３）。さらに、２次重複排除処理部２０２は、コンテンツ管理テーブル１２４からチャンクデータセットＩＤ（ChunkDataSetID）を取得する（Ｓ３１４）。

[0097] そして、２次重複排除処理部２０２は、ステップＳ３１３において取得したフィンガープリント（FingerPrint）をキーにして、チャンクデータセットインデックス（ChunkDataSetIndex）１２３からチャンクのチャンク長（Length）及びオフセット（Offset）を取得する（Ｓ３１５）。

[0098] そして、２次重複排除処理部２０２は、ステップＳ３１５において取得したチャンクデータセットのオフセット（Offset）からチャンク長（Length）分のデータを読み出す（Ｓ３１６）。そして、２次重複排除処理部２０２は、ステップＳ３１６において読み出したチャンクデータを第１ファイルシス

テムに書き込む（Ｓ３１７）。

[0099] 図１２に戻り、ステップＳ３０２において２次重複排除済みのデータのRead処理が実行された後、１次重複排除処理部２０１は、１次重複排除済みファイルをReadする（Ｓ３０３）。

[0100] そして、ステップＳ３０３においてReadしたデータを伸長する（Ｓ３０４）。そして、データを要求したホスト装置２００等のデータ要求元に圧縮前のオリジナルデータを返却する（Ｓ３０５）。以上、重複排除済みデータのRead処理について説明した。

[0101] （１－６）本実施形態の効果

以上のように、本実施の形態によれば、１次重複排除処理部２０１は、ホスト装置２００からのデータを１または２以上のチャンクに分割し、分割したチャンクを圧縮し、チャンクの圧縮率が所定の閾値より低い場合に、圧縮された該チャンクのハッシュ値を算出し、該ハッシュ値とＨＤＤ１０４に既に格納されているデータのハッシュ値とを比較して第１の重複排除処理を実行し、チャンクの圧縮率が所定の閾値より大きい場合に、圧縮された該チャンクを第１のファイルシステムに格納した後に、２次重複排除処理部２０２が、圧縮された該チャンクのハッシュ値を算出し、該ハッシュ値と既にＨＤＤ１０４に格納されているデータのハッシュ値とを比較して２次重複排除処理を実行する。

[0102] これにより、重複排除処理のうち、処理負荷の小さいデータの分割処理を１次重複排除処理時に行うことができ、チャンクの圧縮率に基づいて、該チャンクを１次重複排除処理で重複排除を行うか、２次重複排除処理で重複排除処理を行うかを決定し、１次重複排除処理と２次重複排除処理のそれぞれの利点を考慮して効率的に重複排除処理を実行することが可能となる。

[0103] （２）第２の実施形態

次に、図１４を参照して、第２の実施形態について説明する。以下では、上記した第１の実施形態と同様の構成については詳細な説明は省略し、第１の実施形態と異なる構成について特に詳細に説明する。計算機システムのハ

ードウェア構成は、第１の実施形態と同様であるため、詳細な説明は省略する。

[0104] （２－１）ホスト装置及びストレージ装置のソフトウェア構成

本実施形態では、図１４に示すように、ホスト装置２００’に１次重複排除処理部２０１が備えられ、ストレージ装置１００’には、２次重複排除処理部２０２が備えられた構成となっている。ホスト装置２００’は、バックアップサーバ等のサーバ、他のストレージ装置であってもよい。

[0105] このように、ホスト装置２００’において１次重複排除処理を実行することにより、データのバックアップ時に、ホスト装置２００’からストレージ装置１００’へのデータ量を削減することができる。例えば、ホスト装置２００’の処理能力が高く、ホスト装置２００’とストレージ装置１００’との間の転送能力が低い場合には、本実施形態の如く構成することが好ましい。

符号の説明

[0106]

１００ ストレージ装置
１０１ 仮想サーバ
１０３ システムメモリ
１０５ ファイバチャネルポート
１０６ ファイバチャネルケーブル
１１０ ディスクアレイ装置
１２１ スタブファイル
１２２ チャンクデータセット
１２３ チャンクデータセットインデックス
１２４ コンテンツ管理テーブル
１２５ チャンクインデックス
２００ ホスト装置
２０１ １次重複排除処理部
２０２ ２次重複排除処理部

２０３　　ファイルシステム管理部

請求の範囲

[請求項1] 第１記憶領域と第２記憶領域とを提供する記憶装置と、

前記記憶装置へのデータの入出力を制御する制御部と、

を備え、

前記制御部は、

受信したデータを１または２以上のチャンクに分割し、

分割した前記チャンクを圧縮し、

圧縮率が閾値以下のチャンクに対し、前記第１記憶領域に格納せずに、圧縮された前記チャンクのハッシュ値を算出し、前記ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第１の重複排除処理を実行し、

圧縮率が閾値より大きいチャンクに対し、前記第１記憶領域に圧縮された前記チャンクを格納した後に、前記圧縮された前記チャンクを前記第１記憶領域から読み出し、圧縮された前記チャンクのハッシュ値を算出し、該ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第２の重複排除処理を実行する

ことを特徴とする、ストレージ装置。

[請求項2] 前記制御部は、

前記第１記憶領域と第１のファイルシステムとを対応付け、前記第２記憶領域と第２のファイルシステムとを対応付け、

前記第１の重複排除処理により重複排除できないチャンクと、圧縮率が前記閾値より大きいチャンクと、を第１のファイルシステムに格納し、

前記第１のファイルシステムに格納したチャンクに対して前記第２の重複排除処理を実行した前記チャンクを第２のファイルシステムに格納する

ことを特徴とする、請求項１に記載のストレージ装置。

[請求項3] 前記制御部は、

圧縮した前記チャンクに前記第１の重複排除処理を実行したかを示す情報を含む圧縮ヘッダを付して前記第１のファイルシステムに格納し、

前記圧縮ヘッダを参照して、前記第１の重複排除処理を実行していない場合に、前記チャンクに前記第２の重複排除処理を実行する

ことを特徴とする、請求項２に記載のストレージ装置。

[請求項4] 前記制御部は、

前記チャンクに前記第１の重複排除処理を実行していない場合に、前記圧縮ヘッダに第１のフラグを設定し、

前記チャンクに前記第１の重複排除処理を実行し、当該チャンクのハッシュ値と同一のハッシュ値である他のデータが前記第２記憶領域に格納されていない場合に、前記圧縮ヘッダに第２のフラグを設定し、

前記チャンクに前記第１の重複排除処理を実行し、当該チャンクのハッシュ値と同一のハッシュ値である他のデータが前記第２記憶領域に記憶されている場合に、前記圧縮ヘッダに第３のフラグを設定する

ことを特徴とする、請求項３に記載のストレージ装置。

[請求項5] 前記制御部は、

前記圧縮ヘッダに前記第１のフラグを設定した場合に、前記チャンク及び該チャンクの圧縮ヘッダを前記第１のファイルシステムに格納し、

前記圧縮ヘッダに前記第２のフラグを設定した場合に、前記チャンク及び該チャンクの圧縮ヘッダを前記第１のファイルシステムに格納し、

前記圧縮ヘッダに前記第３のフラグを設定した場合に、前記チャンクの圧縮ヘッダのみ前記第１のファイルシステムに格納する

ことを特徴とする、請求項４に記載のストレージ装置。

[請求項6] 前記制御部は、

前記圧縮ヘッダに前記第１のフラグが設定されている場合に、前記チャンクに前記第２の重複排除処理を実行し、

前記圧縮ヘッダに前記第２のフラグが設定されている場合に、前記チャンクを前記第２記憶領域に格納し、

前記圧縮ヘッダに前記第３のフラグが設定されている場合に、前記チャンクの前記第２記憶領域の格納先を取得する

ことを特徴とする、請求項４に記載のストレージ装置。

［請求項7］ 第１記憶領域と第２記憶領域とを提供する記憶装置と、前記記憶装置へのデータの入出力を制御する制御部と、を備えたストレージ装置におけるデータ管理方法であって、

前記制御部が、受信したデータを１または２以上のチャンクに分割し、分割した前記チャンクを圧縮する第１のステップと、

前記制御部が、圧縮率が閾値以下のチャンクに対して、前記第１記憶領域に格納せずに、圧縮された前記チャンクのハッシュ値を算出し、前記ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第１の重複排除処理を実行する第２のステップと、

前記制御部が、圧縮率が閾値より大きいチャンクに対し、前記第１記憶領域に圧縮された前記チャンクを格納した後に、前記圧縮された前記チャンクを前記第１記憶領域から読み出し、圧縮された前記チャンクのハッシュ値を算出し、該ハッシュ値と既に前記第２記憶領域に格納されている他のデータのハッシュ値とを比較して第２の重複排除処理を実行する第３のステップと

を含むことを特徴とする、データ管理方法。

［請求項8］ 前記第１記憶領域と第１のファイルシステムとが対応付けられ、前記第２記憶領域と第２のファイルシステムとが対応付けられており、

前記第２のステップにおいて、前記制御部が前記第１の重複排除処理により重複排除できないチャンクと、圧縮率が前記閾値より大きい

チャンクとを第１のファイルシステムに格納する第４のステップと、

前記第３のステップにおいて、前記制御部が前記第１のファイルシステムに格納したチャンクに対して前記第２の重複排除処理を実行した前記チャンクを第２のファイルシステムに格納する第５のステップと

を含むことを特徴とする、請求項７に記載のデータ管理方法。

[請求項9] 前記第４のステップにおいて、前記制御部が圧縮した前記チャンクに前記第１の重複排除処理を実行したかを示す情報を含む圧縮ヘッダを付して前記第１のファイルシステムに格納する第６のステップと、

前記圧縮ヘッダを参照して、前記第１の重複排除処理を実行していない場合に、前記チャンクに前記第２の重複排除処理を実行する第７のステップと

を含むことを特徴とする、請求項８に記載のデータ管理方法。

[請求項10] 前記制御部が

前記チャンクに前記第１の重複排除処理を実行していない場合に、前記圧縮ヘッダに第１のフラグを設定し、

前記チャンクに前記第１の重複排除処理を実行し、当該チャンクのハッシュ値と同一のハッシュ値である他のデータが前記第２記憶領域に格納されていない場合に、前記圧縮ヘッダに第２のフラグを設定し、

前記チャンクに前記第１の重複排除処理を実行し、当該チャンクのハッシュ値と同一のハッシュ値である他のデータが前記第２記憶領域に記憶されている場合に、前記圧縮ヘッダに第３のフラグを設定する

第８のステップを含むことを特徴とする、請求項９に記載のデータ管理方法。

[請求項11] 前記制御部が、

前記圧縮ヘッダに前記第１のフラグを設定した場合に、前記チャンク及び該チャンクの圧縮ヘッダを前記第１のファイルシステムに格納

し、

前記圧縮ヘッダに前記第２のフラグを設定した場合に、前記チャンク及び該チャンクの圧縮ヘッダを前記第１のファイルシステムに格納し、

前記圧縮ヘッダに前記第３のフラグを設定した場合に、前記チャンクの圧縮ヘッダのみ前記第１のファイルシステムに格納する

第９のステップを含むことを特徴とする、請求項１０に記載のデータ管理方法。

［請求項12］ 前記制御部は、

前記圧縮ヘッダに前記第１のフラグが設定されている場合に、前記チャンクに前記第２の重複排除処理を実行し、

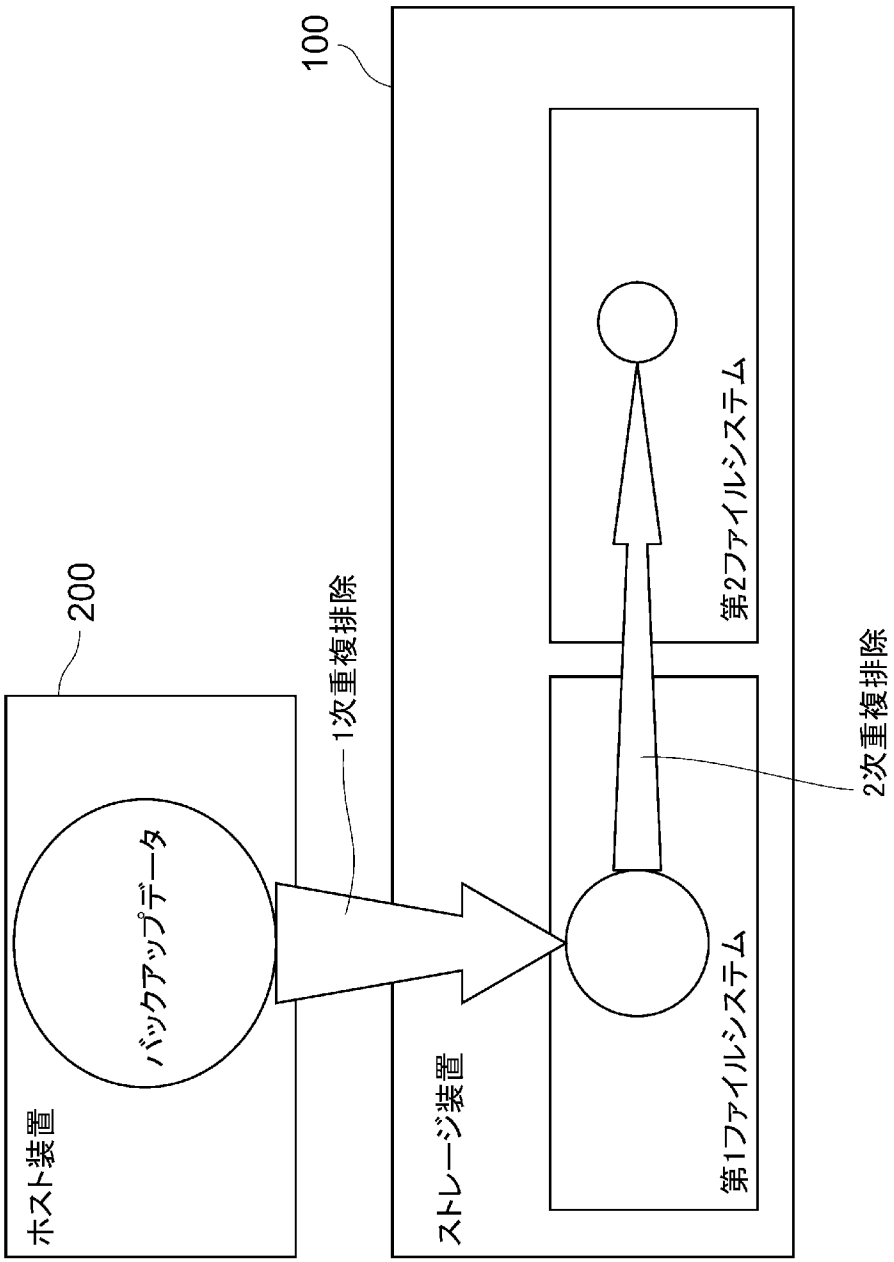
前記圧縮ヘッダに前記第２のフラグが設定されている場合に、前記チャンクを前記第２記憶領域に格納し、

前記圧縮ヘッダに前記第３のフラグが設定されている場合に、前記チャンクの前記第２記憶領域の格納先を取得する

第１０のステップを含むことを特徴とする、請求項１０に記載のデータ管理方法。

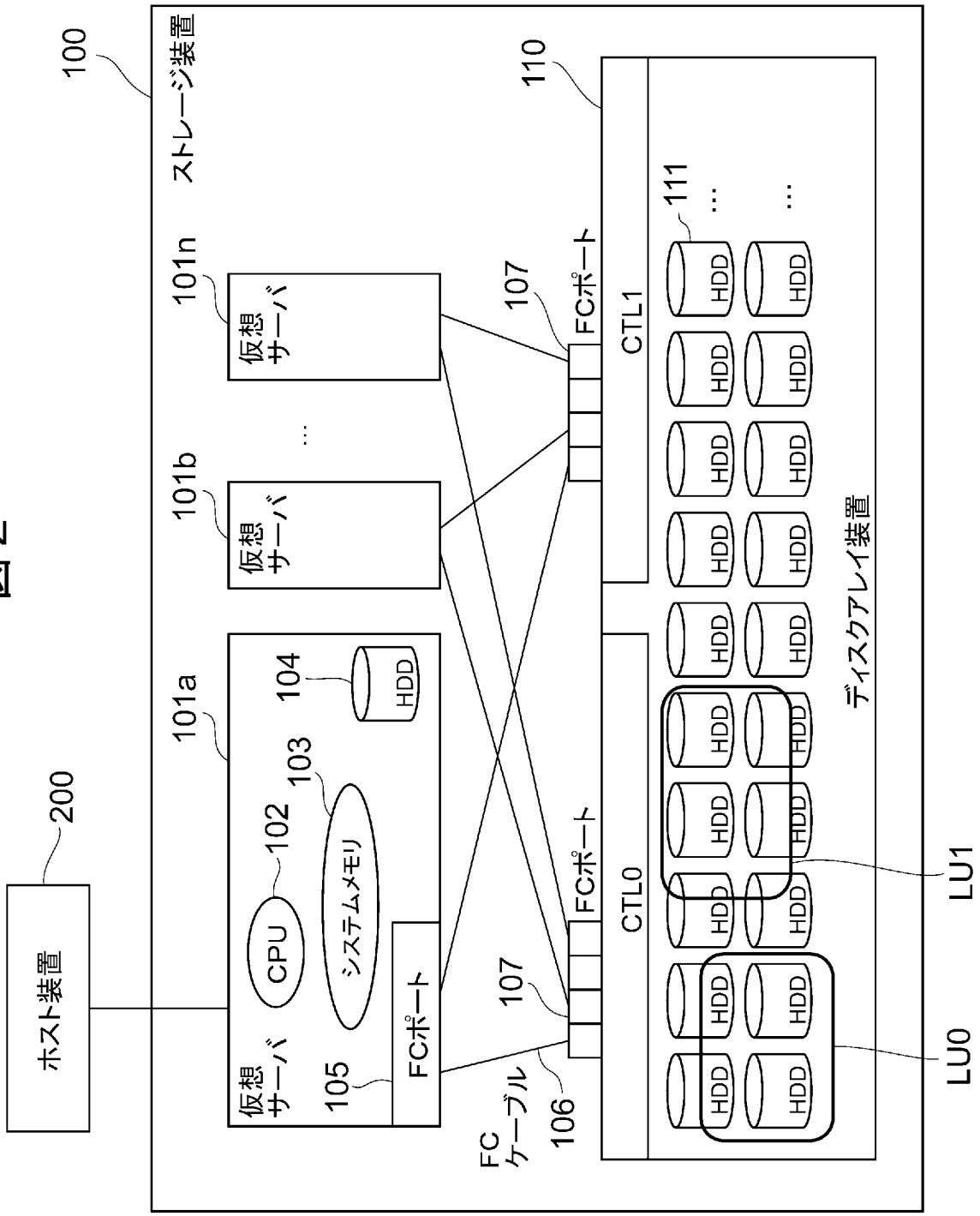
[図1]

図１


100
200
ホスト装置
バックアップデータ
1次重複排除
ストレージ装置
第1ファイルシステム
第2ファイルシステム
2次重複排除

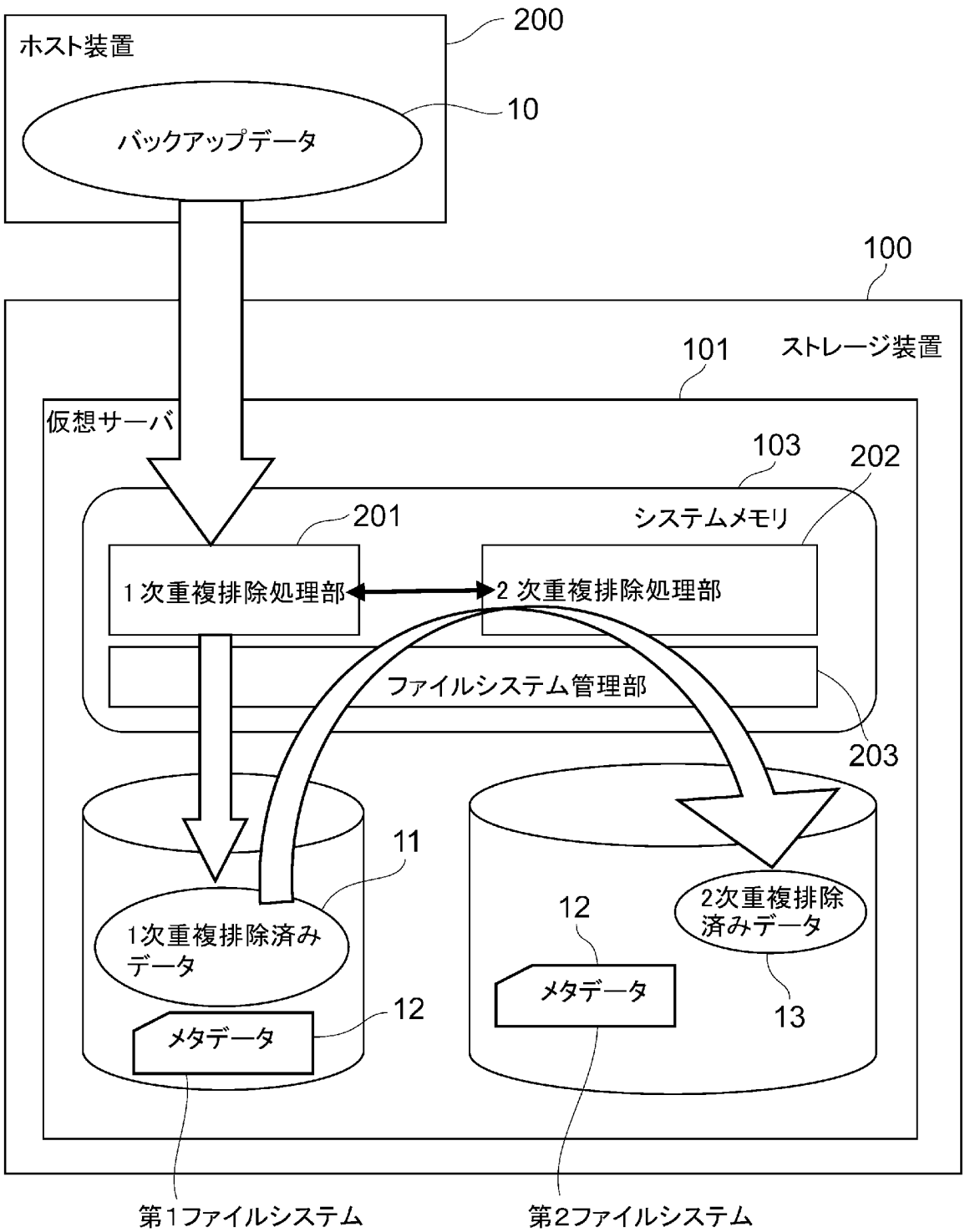
[図2]

図2


100
ストレージ装置
101n
仮想サーバ
101b
仮想サーバ
101a
仮想サーバ
CPU
102
103
システムメモリ
104
HDD
105
FCポート
107
106
FCケーブル
200
ホスト装置
110
CTL1
CTL0
FCポート
111
HDD
ディスクアレイ装置
LU1
LU0

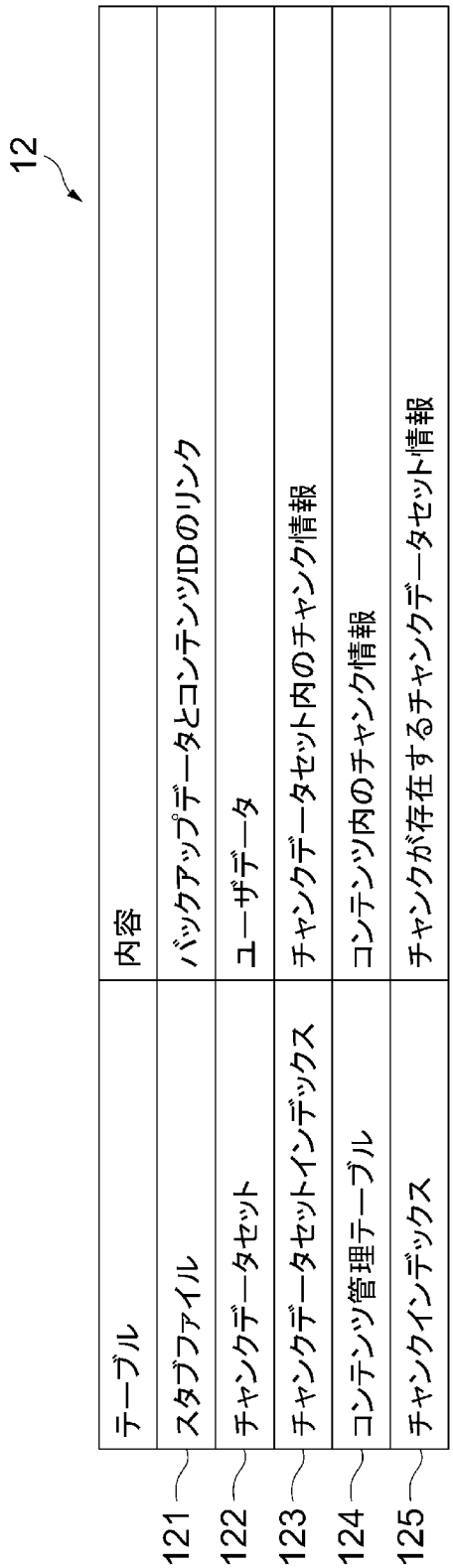
[図3]

図３


ホスト装置
200
バックアップデータ
10
100
101
ストレージ装置
仮想サーバ
103
202
201
システムメモリ
1次重複排除処理部
2次重複排除処理部
ファイルシステム管理部
203
11
1次重複排除済みデータ
12
メタデータ
12
メタデータ
2次重複排除済みデータ
13
第1ファイルシステム
第2ファイルシステム

[図4]

図 4


12
テーブル
内容
スタブファイル
バックアップデータとコンテンツIDのリンク
チャンクデータセット
ユーザデータ
チャンクデータセットインデックス
チャンクデータセット内のチャンク情報
コンテンツ管理テーブル
コンテンツ内のチャンク情報
チャンクインデックス
チャンクが存在するチャンクデータセット情報
121
122
123
124
125

[図5]

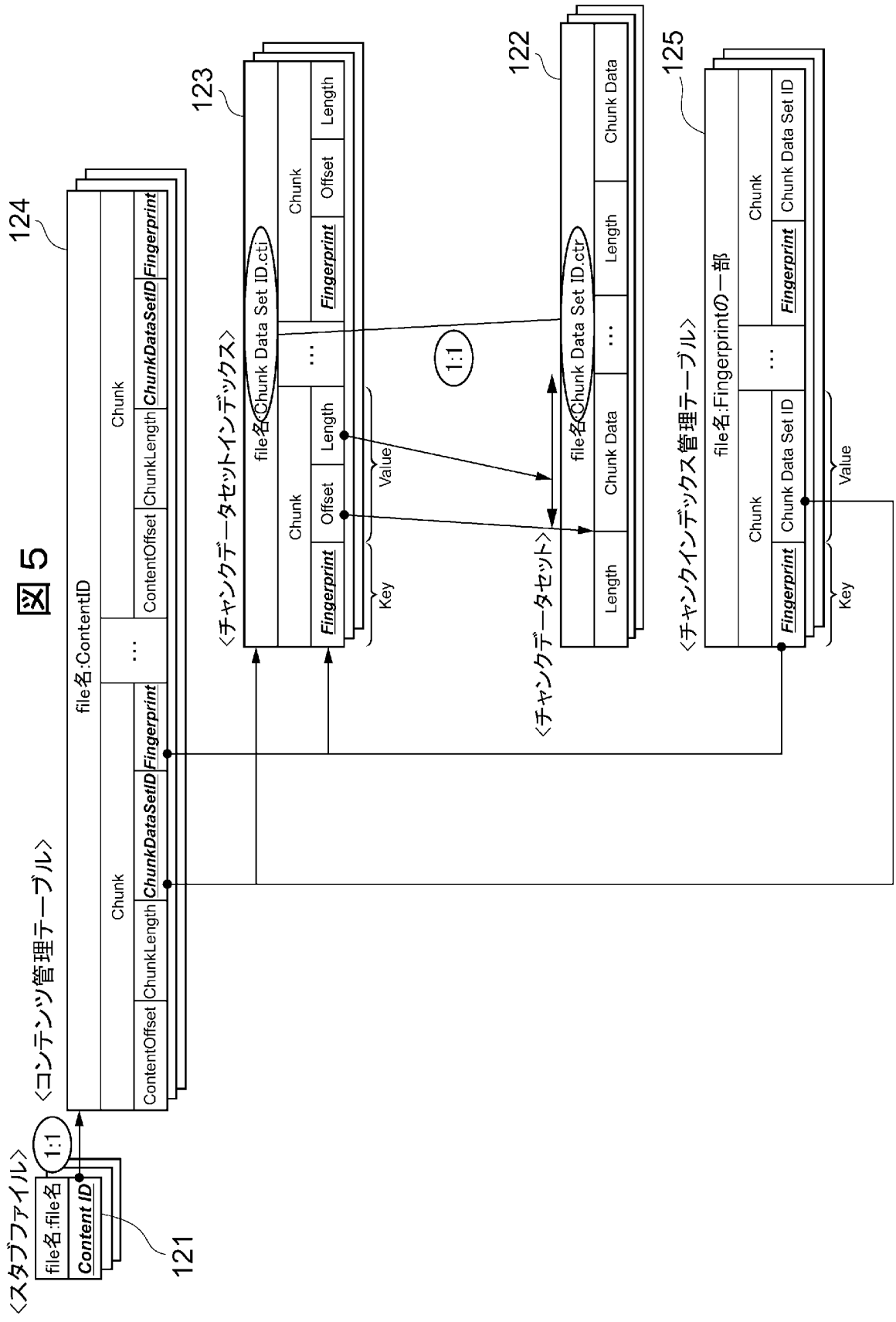
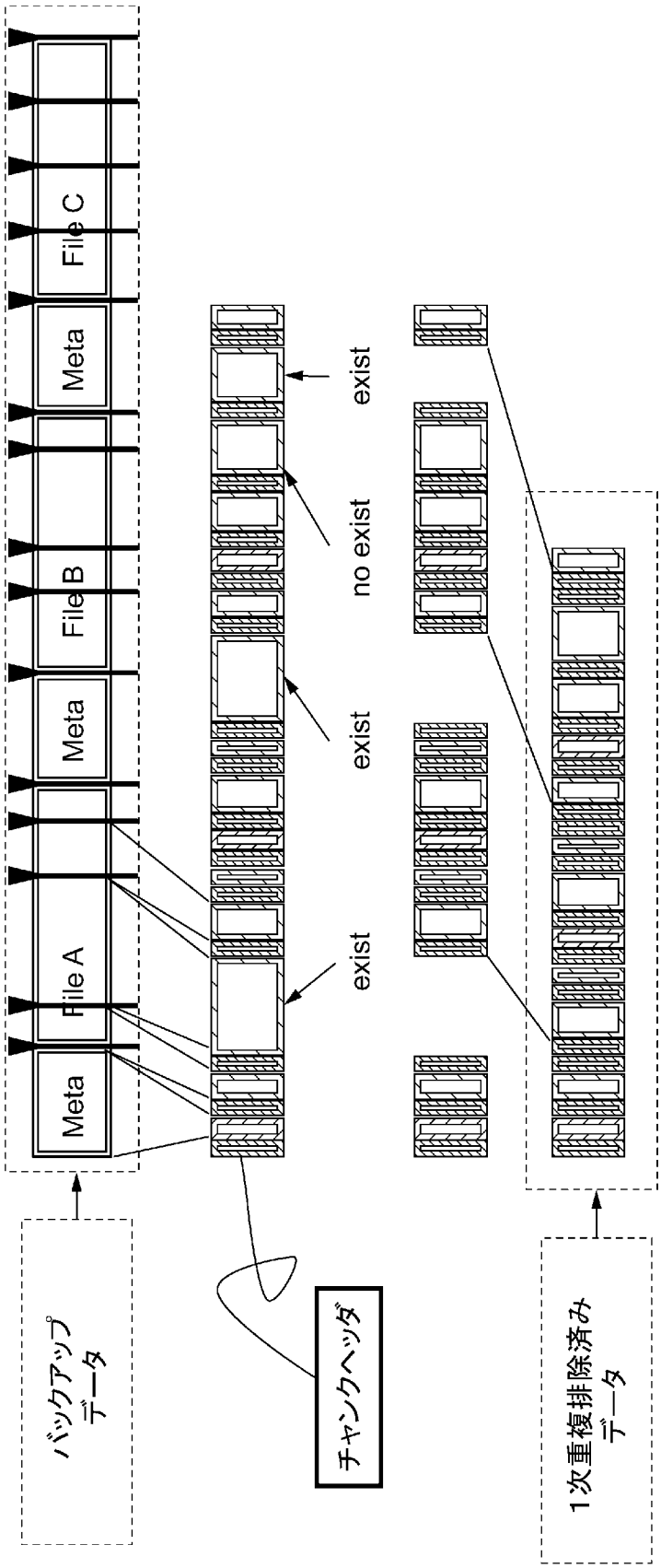


図5
〈スタブファイル〉
file名:file名
Content ID
1:1
121
〈コンテンツ管理テーブル〉
124
file名:ContentID
Chunk
ContentOffset
ChunkLength
ChunkDataSetID
Fingerprint
〈チャンクデータセットインデックス〉
123
file名:Chunk Data Set ID.cti
Chunk
Offset
Length
Fingerprint
Key
Value
1:1
〈チャンクデータセット〉
122
file名:Chunk Data Set ID.ctr
Chunk Data
Length
〈チャンクインデックス管理テーブル〉
125
file名:Fingerprintの一部
Chunk Data Set ID

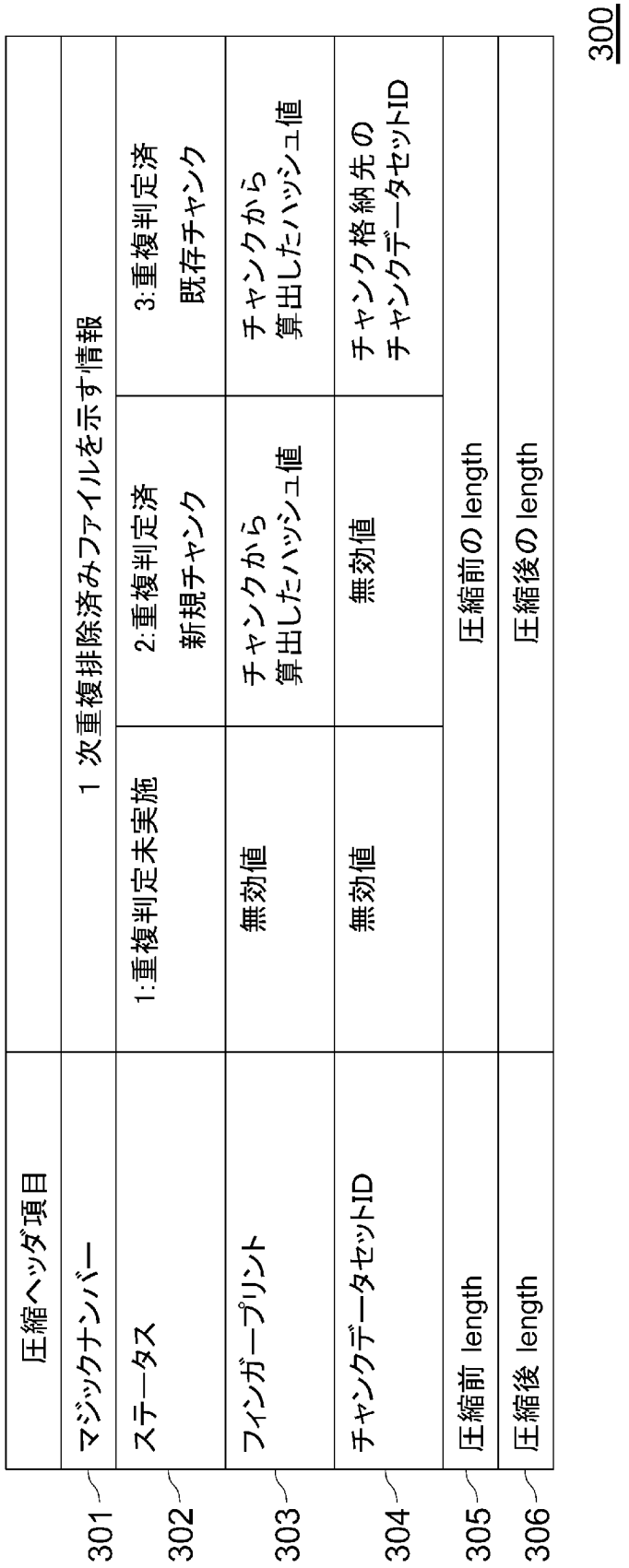
[図6]

図6


Meta
File A
Meta
File B
Meta
File C
exist
no exist
exist
exist
バックアップデータ
チャンクヘッダ
1次重複排除済みデータ

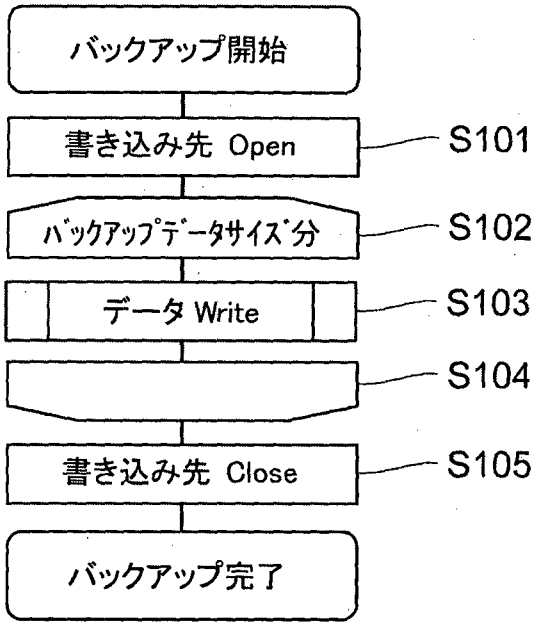
[図7]

図 7


圧縮ヘッダ項目
マジックナンバー
1 次重複排除済みファイルを示す情報
ステータス
1:重複判定未実施
2:重複判定済
新規チャンク
3:重複判定済
既存チャンク
フィンガープリント
無効値
チャンクから
算出したハッシュ値
チャンクから
算出したハッシュ値
チャンクデータセットID
無効値
無効値
チャンク格納先の
チャンクデータセットID
圧縮前 length
圧縮前の length
圧縮後 length
圧縮後の length
301
302
303
304
305
306
300

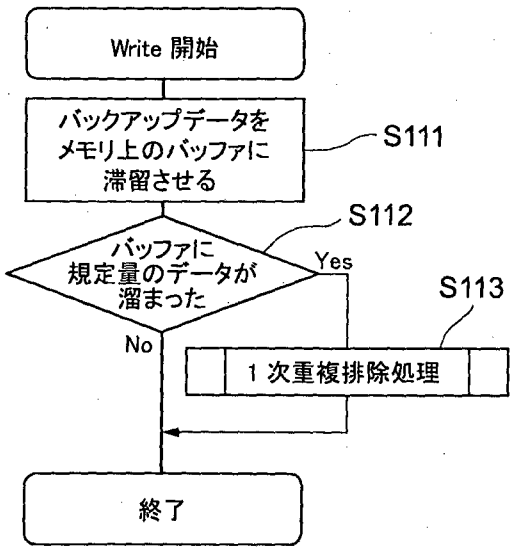
[図 8]

図 8


バックアップ開始
書き込み先 Open
S101
ﾊﾞｯｸｱｯﾌﾟﾃﾞｰﾀｻｲｽﾞ分
S102
データ Write
S103
S104
書き込み先 Close
S105
バックアップ完了

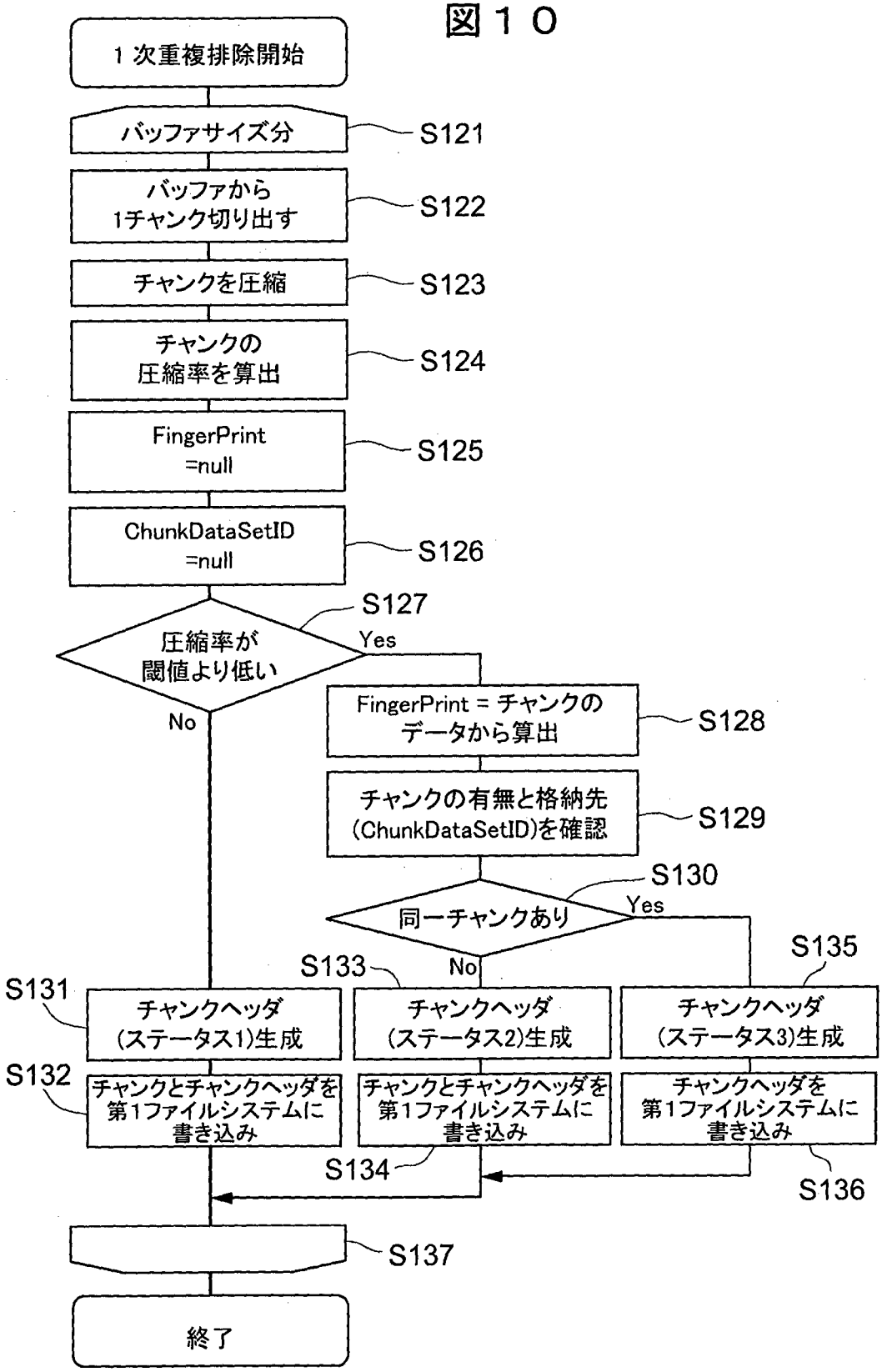
[図 9]

図 9


Write 開始
バックアップデータを
メモリ上のバッファに
滞留させる
S111
S112
バッファに
規定量のデータが
溜まった
Yes
No
S113
1 次重複排除処理
終了

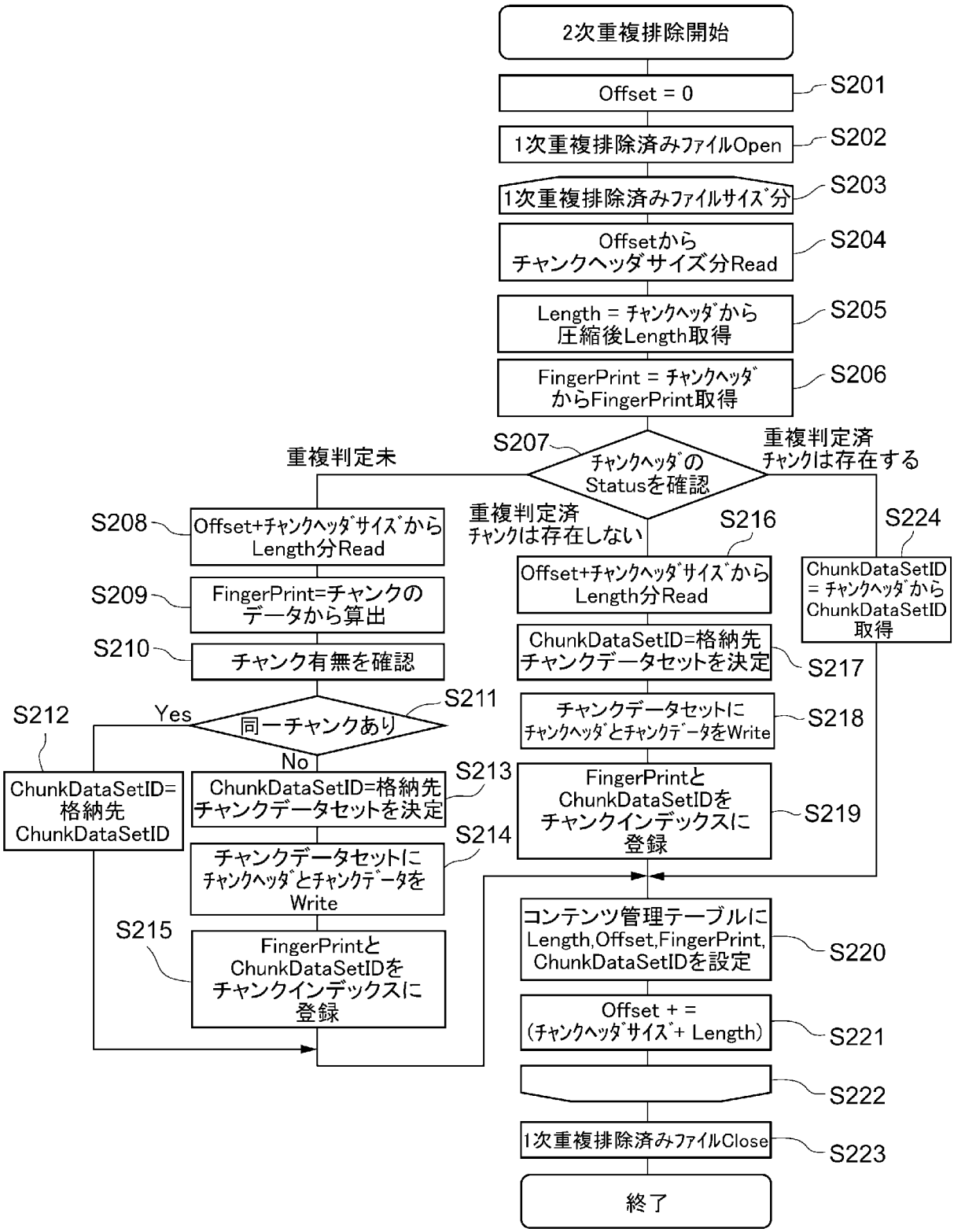
[図 10]

図１０


1 次重複排除開始
バッファサイズ分
S121
バッファから
1チャンク切り出す
S122
チャンクを圧縮
S123
チャンクの
圧縮率を算出
S124
FingerPrint
=null
S125
ChunkDataSetID
=null
S126
S127
圧縮率が
閾値より低い
Yes
No
FingerPrint = チャンクの
データから算出
S128
チャンクの有無と格納先
(ChunkDataSetID)を確認
S129
S130
同一チャンクあり
Yes
No
S131
S132
S133
S135
チャンクヘッダ
(ステータス1)生成
チャンクとチャンクヘッダを
第1ファイルシステムに
書き込み
チャンクヘッダ
(ステータス2)生成
チャンクとチャンクヘッダを
第1ファイルシステムに
書き込み
チャンクヘッダ
(ステータス3)生成
チャンクヘッダを
第1ファイルシステムに
書き込み
S134
S136
S137
終了

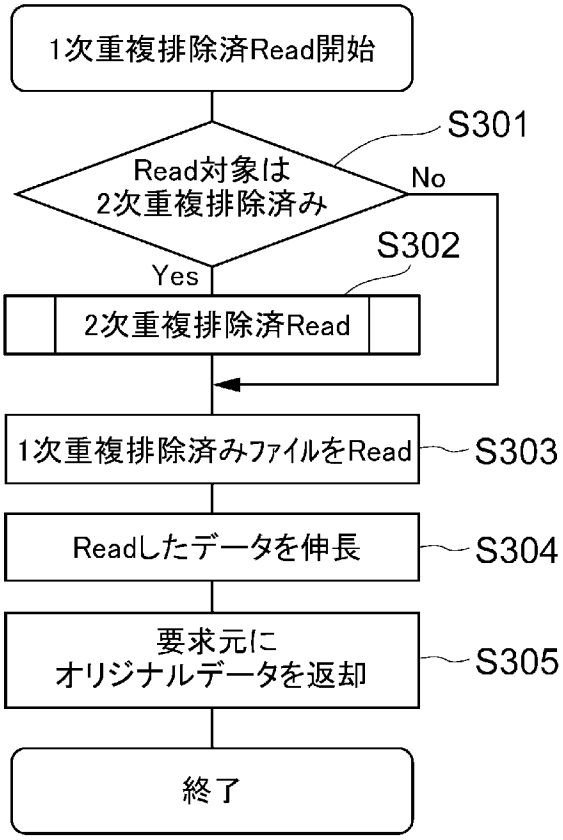
[図11]

図１１


2次重複排除開始
Offset = 0
S201
1次重複排除済みﾌｧｲﾙOpen
S202
1次重複排除済みﾌｧｲﾙｻｲｽﾞ分
S203
Offsetから
ﾁｬﾝｸﾍｯﾀﾞｻｲｽﾞ分Read
S204
Length = ﾁｬﾝｸﾍｯﾀﾞから
圧縮後Length取得
S205
FingerPrint = ﾁｬﾝｸﾍｯﾀﾞ
からFingerPrint取得
S206
S207
ﾁｬﾝｸﾍｯﾀﾞの
Statusを確認
重複判定未
重複判定済
ﾁｬﾝｸは存在する
重複判定済
ﾁｬﾝｸは存在しない
S208
Offset+ﾁｬﾝｸﾍｯﾀﾞｻｲｽﾞから
Length分Read
S209
FingerPrint=チャンクの
データから算出
S210
チャンク有無を確認
S211
同一チャンクあり
Yes
No
S212
ChunkDataSetID=
格納先
ChunkDataSetID
S213
ChunkDataSetID=格納先
チャンクデータセットを決定
S214
チャンクデータセットに
ﾁｬﾝｸﾍｯﾀﾞとﾁｬﾝｸﾃﾞｰﾀを
Write
S215
FingerPrintと
ChunkDataSetIDを
チャンクインデックスに
登録
S216
Offset+ﾁｬﾝｸﾍｯﾀﾞｻｲｽﾞから
Length分Read
S217
ChunkDataSetID=格納先
チャンクデータセットを決定
S218
チャンクデータセットに
ﾁｬﾝｸﾍｯﾀﾞとﾁｬﾝｸﾃﾞｰﾀをWrite
S219
FingerPrintと
ChunkDataSetIDを
チャンクインデックスに
登録
S224
ChunkDataSetID
= ﾁｬﾝｸﾍｯﾀﾞから
ChunkDataSetID
取得
S220
コンテンツ管理テーブルに
Length,Offset,FingerPrint,
ChunkDataSetIDを設定
S221
Offset + =
(ﾁｬﾝｸﾍｯﾀﾞｻｲｽﾞ+ Length)
S222
S223
1次重複排除済みﾌｧｲﾙClose
終了

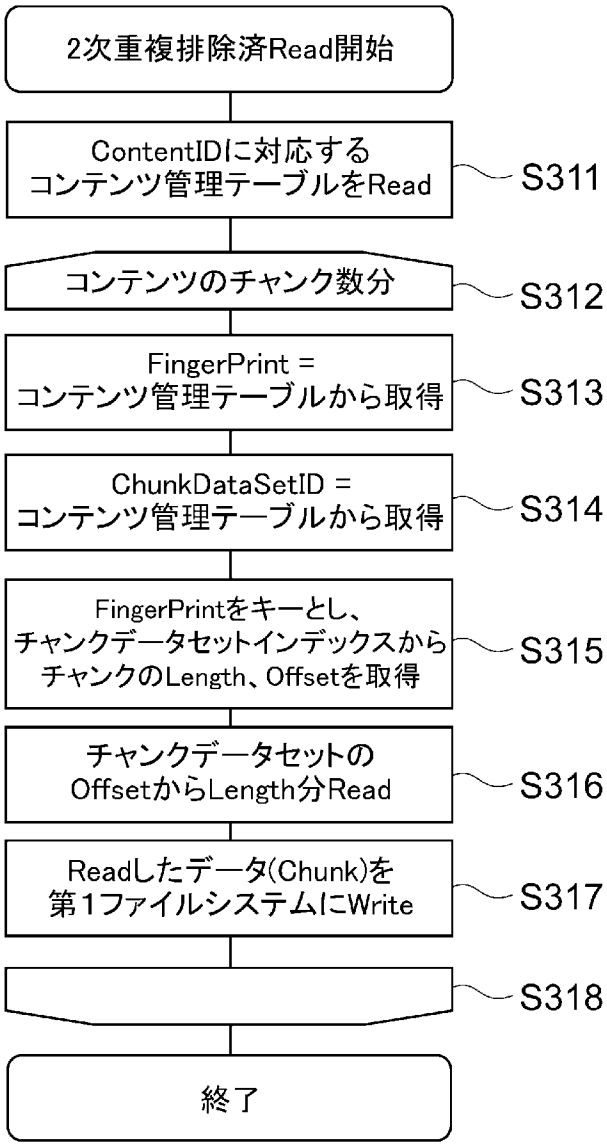
[図12]

図１２


1次重複排除済Read開始
S301
Read対象は
2次重複排除済み
No
Yes
S302
2次重複排除済Read
1次重複排除済みファイルをRead
S303
Readしたデータを伸長
S304
要求元に
オリジナルデータを返却
S305
終了

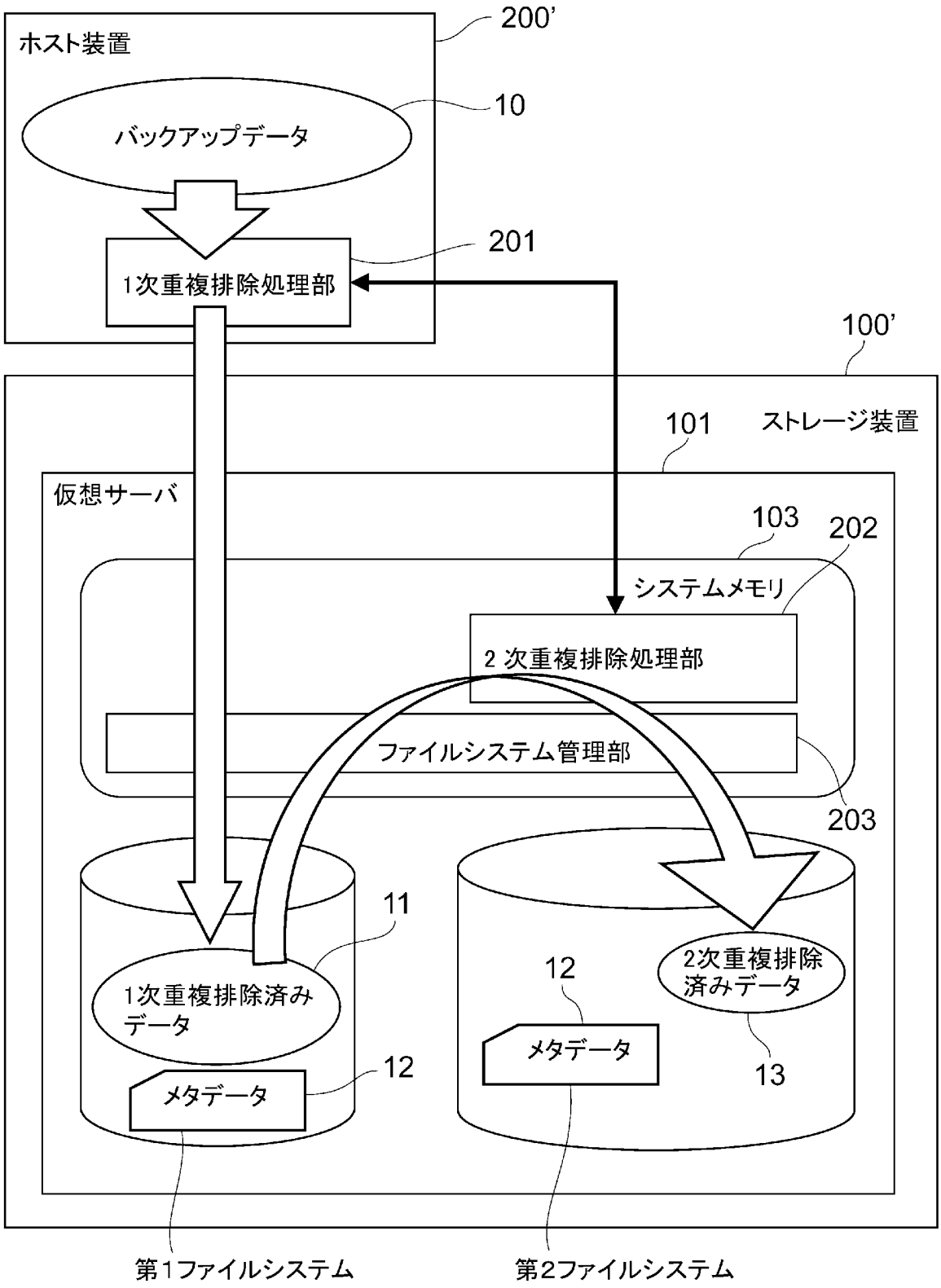
[図13]

図１３


2次重複排除済Read開始
ContentIDに対応する
コンテンツ管理テーブルをRead
S311
コンテンツのチャンク数分
S312
FingerPrint =
コンテンツ管理テーブルから取得
S313
ChunkDataSetID =
コンテンツ管理テーブルから取得
S314
FingerPrintをキーとし、
チャンクデータセットインデックスから
チャンクのLength、Offsetを取得
S315
チャンクデータセットの
OffsetからLength分Read
S316
Readしたデータ(Chunk)を
第1ファイルシステムにWrite
S317
S318
終了

[図14]

図１４


ホスト装置
200'
バックアップデータ
10
1次重複排除処理部
201
100'
101
ストレージ装置
仮想サーバ
103
202
システムメモリ
2 次重複排除処理部
ファイルシステム管理部
203
11
1次重複排除済み
データ
12
メタデータ
12
2次重複排除
済みデータ
メタデータ
13
第1ファイルシステム
第2ファイルシステム

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2012/071424

A. CLASSIFICATION OF SUBJECT MATTER
G06F3/06(2006.01)i, *G06F12/00*(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06F3/06, G06F12/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2011/0125722 A1 (Ocarina Networks), 26 May 2011 (26.05.2011), entire text; all drawings (particularly, the parts relating to fig. 8) (Family: none)	1-12
A	Wataru KATSURASHIMA, Storage ni Okina Henka o Motarasu Chofuku Haijo Gijutsu ga Kakushin suru Storage no Sekai, Gekkan Ascii Dot Technologies 2011 Nen 1 Gatsu Go, 25 November 2010 (25.11.2010), vol.16, no.1, pages 108 to 115	1-12
A	Wataru KATSURASHIMA, Storage Bun'ya no Yottsu no Chumoku Gijutsu, Gekkan Ascii Dot Technologies 2011 Nen 2 Gatsu Go, 24 December 2010 (24.12.2010), vol.16, no.2, pages 56 to 59	1-12

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 November, 2012 (30.11.12)	11 December, 2012 (11.12.12)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2012/071424

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-304307 A (Sanyo Electric Co., Ltd.), 28 October 2004 (28.10.2004), entire text; all drawings (particularly, the parts relating to fig. 3 to 5) (Family: none)	1-12

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１２／０７１４２４

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F3/06(2006.01)i, G06F12/00(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F3/06, G06F12/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１２年
日本国実用新案登録公報	１９９６－２０１２年
日本国登録実用新案公報	１９９４－２０１２年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	US 2011/0125722 A1 (Ocarina Networks) 2011.05.26, 全文、全図（特に、Fig.8に係る記載）（ファミリーなし）	1-12
A	桂島航，ストレージに大きな変化をもたらす 重複排除技術が革新するストレージの世界，月刊アスキードットテクノロジーズ2011年1月号，2010.11.25，第16巻 第1号，pp.108-115	1-12
A	桂島航，ストレージ分野の４つの注目技術，月刊アスキードットテクノロジーズ2011年2月号，2010.12.24，第16巻第2号，pp.56-59	1-12

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日 ３０．１１．２０１２	国際調査報告の発送日 １１．１２．２０１２
国際調査機関の名称及びあて先 日本国特許庁（ＩＳＡ／ＪＰ） 郵便番号１００－８９１５ 東京都千代田区霞が関三丁目４番３号	特許庁審査官（権限のある職員）　５Ｔ　９８５７ 木村　貴俊 電話番号　０３－３５８１－１１０１　内線　３５６８

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Ａ	JP 2004-304307 A（三洋電機株式会社）2004.10.28，全文、全図（特に、図３－５に係る記載参照）（ファミリーなし）	1-12