

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2012-505440
(P2012-505440A)

(43) 公表日 平成24年3月1日(2012.3.1)

(51) Int.Cl.		F I			テーマコード (参考)
G06F 3/06	(2006.01)	G06F 3/06	301J		5B014
G06F 13/10	(2006.01)	G06F 13/10	340A		5B065
		G06F 3/06	302A		

審査請求 有 予備審査請求 未請求 (全 26 頁)

(21) 出願番号	特願2011-514963 (P2011-514963)	(71) 出願人	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(86) (22) 出願日	平成21年2月25日 (2009.2.25)	(74) 代理人	100093861 弁理士 大賀 真司
(85) 翻訳文提出日	平成23年4月7日 (2011.4.7)	(74) 代理人	100129218 弁理士 百本 宏之
(86) 国際出願番号	PCT/JP2009/054004	(72) 発明者	平山 洋志 神奈川県横浜市戸塚区吉田町292番地株 式会社日立製作所 横浜研究所内
(87) 国際公開番号	W02010/097960	Fターム(参考)	5B014 EB04 GC15 5B065 CC08 CE01 CS04
(87) 国際公開日	平成22年9月2日 (2010.9.2)		

最終頁に続く

(54) 【発明の名称】 ストレージ装置、ストレージ装置でのデータ処理方法

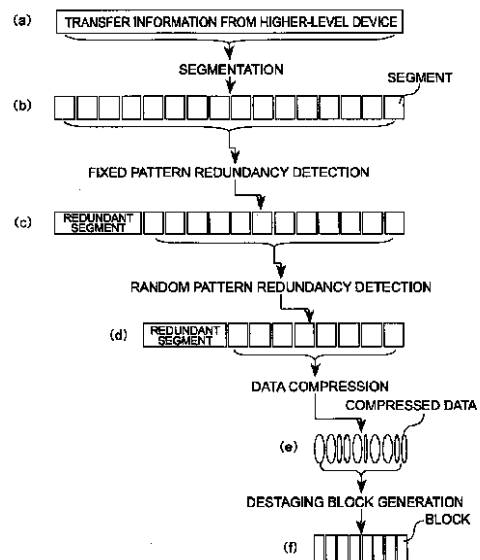
(57) 【要約】

【課題】本発明は、転送データに対する重複排除と、データ圧縮まで含めた処理時間の短縮、ストレージ装置容量の効率的利用を図る。

【解決手段】転送情報に含まれるデータパターンに注目し、最適なパターン検出、重複排除処理を段階的に実施する。更に重複排除後のデータを対象にデータ圧縮処理を実施する。更に、段階的に行われる重複排除、圧縮処理の実施判定を、各処理段階の残データ量や、装置負荷により判断する。

【選択図】図2

FIG.2



【特許請求の範囲】

【請求項 1】

上位装置から転送される情報を受信する第 1 のインターフェースと、
 前記情報を一時記憶するキャッシュメモリと、
 前記情報を格納するための記憶領域を有する記憶デバイスと、
 前記情報を前記記憶デバイスに送信する第 2 のインターフェースと、
 前記情報に対する制御処理を実行する制御装置と、
 を備えるストレージ装置であって、
 前記制御装置は、前記情報から、既に前記記憶領域に存在する重複部分を除き、残りの
 部分を圧縮して当該記憶領域に格納するストレージ装置。

10

【請求項 2】

請求項 1 記載のストレージ装置において、
 前記制御装置は、
 前記情報を複数のセグメントに分けて処理する第 1 の処理と、
 前記複数のセグメントのそれぞれが有する複数のデータの組み合わせが、規則性を持った
 第 1 のパターンであるか否かを判定する第 2 の処理と、
 前記第 1 のパターンからなるセグメントの検出のみで前記重複を判定するか、或いは、
 前記第 1 のパターンからなるセグメントと前記記憶領域に記憶されているセグメントとの
 間で前記重複があるか否かを判定する第 3 の処理と、
 前記第 3 の処理が肯定されたセグメントが、前記キャッシュメモリに占める領域を開放
 する第 4 の処理と、
 前記複数のセグメントのうち、前記第 3 の処理が肯定されたもの以外を、複数のデータ
 の組み合わせがランダムである第 2 のパターンを持ったセグメントであると判定する第 5
 の処理と、
 前記第 2 のパターンを持ったセグメントを前記キャッシュメモリから読み出す第 6 の処
 理と、
 前記第 2 のパターンを持ったセグメントについて、前記第 2 のパターンの異同を区別す
 るためのハッシュ値を演算する第 7 の処理と、
 前記記憶領域に格納されている複数のセグメントそれぞれの前記ハッシュ値を登録した
 ハッシュテーブルを参照し、前記第 7 の処理で得られたハッシュ値と同等のハッシュ値が
 前記ハッシュテーブルに存在するか否かを判定する第 8 の処理と、
 前記第 8 の処理が肯定されたセグメントが、前記キャッシュメモリに占める領域を開放
 する第 9 の処理と、
 前記第 8 処理が肯定されたセグメント以外のセグメントが前記キャッシュメモリに占め
 る領域に記憶された情報を圧縮する第 10 の処理と、
 前記圧縮された情報を備え、前記キャッシュメモリから前記記憶領域へ情報を格納する
 際の単位であるページブロックを一つ又は複数生成する第 11 の処理と、
 前記ページブロックを前記記憶領域に格納する第 12 の処理と、を
 実行するものである請求項 1 記載のストレージ装置。

20

30

【請求項 3】

前記制御装置は、前記情報を構成する複数のデータの組み合わせのパターンに着目し、
 当該パターンが特定のものである部分を前記情報から除き、残りの部分に対して、前記重
 複の判定のための演算を実行する、請求項 1 記載のストレージ装置。

40

【請求項 4】

前記制御装置は、前記情報から、前記データの組み合わせが規則的なパターンで構成さ
 れる部分を除き、前記複数のデータの組み合わせがランダムなパターンで構成される部分
 に、前記重複の判定のための演算を実行する請求項 3 記載のストレージ装置。

【請求項 5】

前記制御装置は、前記演算を、前記複数のデータの組み合わせがランダムなパターンの
 代表値であるハッシュ値を求めるハッシュ演算に基づいて、実行する、請求項 4 記載のス

50

トレージ装置。

【請求項 6】

前記制御装置は、

前記上位装置から転送される情報と、前記記憶領域に記憶されている情報との間での前記重複の判定、及び、前記転送された情報から前記重複した部分を前記圧縮の対象から排除する処理を、第 1 の領域を単位として行い、

前記圧縮の対象に、第 2 の領域を単位とした圧縮を行って圧縮データを生成し、

前記生成された圧縮データを、第 3 の領域に割り付け、当該第 3 の領域を前記記憶デバイスに格納する、請求項 1 記載のストレージ装置。

【請求項 7】

前記制御装置は、

前記転送される情報が格納される前記第 1 の領域には、前記上位装置からのアクセスのための空間に対して定められた第 1 のブロックアドレスを付加し、

前記圧縮データが格納される前記第 3 の領域には、前記記憶デバイスへのアクセスのための空間に対して定められた第 2 のブロックアドレスを付加する、

請求項 1 記載のストレージ装置。

【請求項 8】

前記制御装置は、

前記第 1 のブロックアドレスと、前記第 2 のブロックアドレスとの対応を管理する制御テーブルを備える、ローカルメモリを備える、請求項 7 記載のストレージ装置。

【請求項 9】

前記制御装置は、

前記圧縮データを前記第 3 の領域に格納する際に、圧縮データ量と当該第 3 の領域に格納するデータ量から、前記第 3 の領域の数を求め、

(圧縮前のデータ量に対する領域数) M (M は正の整数で、 M 第 3 の領域数) の条件が成立した際前記圧縮データを破棄し、当該条件が不成立である際、前記第 3 の領域に前記圧縮データを割り付け、前記記憶デバイスへ格納する請求項 6 記載のストレージ装置。

【請求項 10】

前記制御装置は、前記重複部分の排除の処理、又は、前記圧縮の処理の対象となる前記情報の量が閾値を越えると判定した場合、当該判定の対象となった処理をスキップし、

前記圧縮がされた、又は、当該圧縮がされなかった情報を前記第 3 の領域に割り付ける、請求項 6 記載のストレージ装置。

【請求項 11】

前記制御装置は、前記圧縮の処理を前記記憶デバイスに格納された前記情報を対象に行い、

前記上位装置からのアクセス頻度によって、前記圧縮処理の対象とする前記記憶デバイスの領域を選択する請求項 1 記載のストレージ装置。

【請求項 12】

前記制御装置は、複数のデータが同一であるパターン、及び、複数のデータの組み合わせが周期をもって繰り返すパターンを、前記規則性を持っているパターンと判定する、請求項 4 記載のストレージ装置。

【請求項 13】

上位装置から転送される情報を受信する第 1 のインターフェースと、

前記情報を一時記憶するキャッシュメモリ、

前記情報を格納するための記憶領域を有する記憶デバイスと、

前記情報を前記記憶デバイスに送信する第 2 のインターフェースと、

前記情報に対する制御処理を実行する制御装置と、

を備えるストレージ装置でのデータ処理方法であって、

前記制御装置は、前記情報から、既に前記記憶領域に存在する重複部分を除き、残りの

10

20

30

40

50

情報を圧縮して当該記憶領域に格納する前記データ処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明はストレージ装置、及び、ストレージ装置でのデータ処理方法に関するものである。

【背景技術】

【0002】

従来、ストレージ装置は上位装置（ホスト）から転送された転送情報を、キャッシュメモリへ一時的に保持した後、ハードディスクなどの記憶デバイスに格納していた。近年、ストレージ装置が、転送情報とストレージ装置に既に格納されている情報との間で重複を検出し、重複していると判定した転送情報を記憶デバイスに格納しないようにする技術が提案されている。ストレージ装置が、この種の“重複情報”を記憶デバイスに格納しないことによって、物理的な記録容量を効率的に活用することができる。これにより、ストレージ装置がホストに対して見せる、見かけ上の記憶容量を増加することができる。

10

【0003】

ここでの重複を検出する一つの例が、米国特許公報第6,928,526に記載されている。ストレージ装置が、転送情報をセグメント単位に処理し、セグメントに含まれるデータに対して代表値・特徴値（ハッシュ値）を計算する。次いで、ストレージ装置は、転送情報のハッシュ値と、記憶デバイスに格納されている格納情報に対するハッシュ値とを比較することによって、記憶デバイスに同一の内容の情報を書き込まないようにしている。

20

【0004】

一方、ストレージの記録容量を効率的に利用できる別の方法として、データ圧縮技術が挙げられる。ストレージ装置には可逆圧縮が可能なアルゴリズムが適用される。代表的なアルゴリズムの一例として、LZ符号化が挙げられる。

【先行技術文献】

【特許文献】

【0005】

【特許文献1】米国特許公報第6,928,526号

【発明の開示】

30

【発明が解決しようとする課題】

【0006】

ストレージ装置が情報の重複を判定する際には、転送情報に含まれるデータをセグメント毎に分けて、重複を判定する。セグメントは、ストレージ装置が記憶デバイスへアクセスする際の単位であるブロック（例512バイト）の複数から構成される。1セグメントのデータ量が、数キロバイトから、数十メガバイトと小単位とされることによって、重複の検出の確率が高くなるようにすることが望ましい。

しかしながら、セグメントが増えれば、ハッシュ値の管理テーブルの容量は膨大となって、転送情報と格納情報との重複を検証するために要する時間が嵩む。また、セグメント数が多くなると、ハッシュ値を得るための演算には、より一意性の高いアルゴリズム（例えばSHA: Secure Hash Algorithm）を適用し、ハッシュ値の衝突を回避しなければならない。そのようなアルゴリズムでは演算量が多く、処理レーテンシーが増加する。

40

【0007】

また、さらに、上位装置からストレージ装置への転送データは、近年特に増加傾向が著しい。これに対処するために、ストレージ装置の記憶容量も、テラバイト、ペタバイトレベルになっている。したがって、重複を検証するための演算量も増加の一途をたどっている。

【0008】

以上のことから、ストレージ装置において、重複を高精度に判定しようとする、そのための処理時間が増大し、ストレージ装置の上位計算機に対するI/O性能が低下すると

50

いう課題があった。

また、ストレージ装置が膨大な転送情報の全てを圧縮して記憶デバイスに格納しようとする
ると結局のところ記憶容量を効率よく利用することができないばかりか、情報の圧縮、圧縮した情報の伸長に対する処理に時間を費やすことにもなる。

【0009】

そこで、本発明は、この課題を解決するために、ストレージ装置が、記憶デバイスの記憶容量を効率的に利用しながら、転送情報の圧縮、伸長に要する処理時間が増えないようにすることを目的とするものである。

【0010】

さらに、本発明のその他の目的は、ストレージ装置が上位装置から転送された転送情報と記憶デバイスに格納された格納情報との間の重複の判定を高精度に行いながらも、そのための演算効率を向上することによって、重複判定のための処理時間が増えないようすることを目的とするものである。

10

【課題を解決するための手段】

【0011】

前記目的を達成するために、本発明は、ストレージ装置またはそのデータ処理方法において、転送情報から、既に記憶デバイスに存在する重複部分を除き、残りの情報を圧縮して記憶デバイスに格納することによって、記憶デバイスの記憶容量を有効に活用することを特徴とするものである。

さらに、前記目的を達成する他の発明は、ストレージ装置が、転送情報を構成するデータの組み合わせのパターンに着目し、転送情報のうち、パターンが特定である情報部分を除き、残りの情報部分に対して、重複判定のための演算を実行することを特徴とするものである。

20

【0012】

この発明の有用な形態は、ストレージ装置が、転送情報から、データの組み合わせが規則的なパターンで構成される固定情報を除き、データの組み合わせがランダムなパターンで構成されるランダム情報に対して、重複判定のための演算を実行することにより、重複判定をすべき対象となる情報を低減させることによって、重複判定のための処理を効率良く実行することを特徴とするものである。

【発明の効果】

30

【0013】

本発明によれば、ストレージ装置が、記憶デバイスの記憶容量を効率的に利用しながら、転送情報の圧縮、伸長に要する処理時間が増えないようにすることができる。さらに、ストレージ装置が、上位装置から転送された転送情報と記憶デバイスに格納された格納情報との間の重複の判定を高精度に行いながらも、そのための演算効率を向上することによって、重複判定のための処理時間が増えないようにすることができる。

【図面の簡単な説明】

【0014】

【図1】本発明が適用されるストレージ装置の構成図。

【図2】転送データに対する重複排除、データ圧縮処理の説明図。

40

【図4】ストレージ装置に適用される重複排除、データ圧縮処理の説明図。

【図3】固定データの構成例の説明図。

【図5】圧縮ブロックの生成方法を示す説明図。

【図6】圧縮ブロックの別の生成方法を示す説明図。

【図7】圧縮ブロックの生成方法を説明するフローチャート。

【図8】圧縮ブロックの別の生成方法を説明するフローチャート。

【図9】転送ブロック、圧縮ブロックの構成例を示す説明図。

【図10】ハッシュテーブルの構成例の説明図。

【図11】セグメントテーブルの構成例の説明図。

【図12】重複排除、データ圧縮処理の連携方法を説明するフローチャート。

50

【図13】重複排除、データ圧縮処理実施の判定方法を説明するフローチャート。

【図14】デステージデータに対する処理の実施方法を説明するフローチャート。

【図15】上位装置からの転送要求に対する装置動作説明するフローチャート。

【発明を実施するための最良の形態】

【0015】

次に、本発明の実施形態を図面に基づいて説明する。なお、本発明は、以下に説明する実施形態に限定されるものではない。

【0016】

図1は、本発明に係わるストレージ装置100のブロック構成を示したものである。図1において、ストレージ装置100の構成例を示すブロック図である。ストレージ装置は、10二重化された、第1のコントローラ（CTL0：1a）と第2のコントローラ（CTL1：1b）とを備えている。二つのコントローラの間は、Dual I/Oによって接続されている。

【0017】

各コントローラにはキャッシュメモリ（Cache0：4a、または、Cache1：4b）、ホストインターフェイス（Host I/F0：2a、または、Host I/F1：2b）、ドライブインターフェイス（Drv I/F0：3a、または、Drv I/F1：3b）、そして、マイクロプロセッサ（MP0：5a、または、MP1：5b）などのデバイスが接続されている。各コントローラは、20複数のデバイス間で、I/O（情報・データ転送経路）を介した情報・データの転送を制御している。

【0018】

ホストインターフェイス2a、2bは、上位装置（ホスト）と間で、通信プロトコルに従い、20パケット単位での情報の転送を制御するインターフェースである。ドライブインターフェース3a、3bは、記憶デバイス（HDD、SSD等）との間で、通信プロトコルに従い、ブロック単位（記憶デバイスのアクセス単位）で情報の転送を制御するインターフェースである。

【0019】

キャッシュメモリ4a、4bは、ホストインターフェイス2a、2b及びドライブインターフェース3a、3bからの転送データを一時的に蓄積するデバイスである。ストレージ装置は、30キャッシュメモリ4a、4bの存在により、ホストからのアクセスと記憶デバイスへの書き込み（デステージ）を分けることができ、ホストからのアクセスに対する装置応答を高速化し、キャッシュヒットによるリード性能を向上することができる。

【0020】

ドライブインターフェース3a、3bには、記憶デバイスが接続され、ホストから見たアクセス空間である論理ボリューム7a、7b（LU0、LU1、LU：Logical Unit）が割り当てられ、30論理ボリュームはマイクロプロセッサ（5a、5b）（MP0、MP1）により管理される。

【0021】

マイクロプロセッサ5a、5bは、各コントローラ1a、1b、ホストインターフェイス2a、2b、40ドライブインターフェース3a、3bの制御部あるいは制御手段を成す。各マイクロプロセッサが実行する制御に必要な情報は、ローカルメモリ（LM0：7a、または、LM1：7b）に格納されている。なお、キャッシュメモリ4をコントローラ1に接続されているとして説明したが、キャッシュメモリをマイクロプロセッサ5に接続されているローカルメモリ6と統合させるものであってもよい。

【0022】

次に、ストレージ装置100が実現する、既述の“情報の重複”を排除・除外等する動作、40“情報の圧縮”を実行する動作、それぞれの原理について説明する。図2は、これを説明するブロック図である。“情報の重複”を排除する動作に続いて“情報の圧縮”の動作が実行される。

図2の（a）は、上位装置からストレージ装置に転送された情報（転送情報）を示す。（50a）は、上位装置から見えるストレージ装置の記憶領域のアクセス空間（ボリューム空間

)の一部でもある。

図2の(b)は、転送情報に重複が有るか無いかの検査・判定、そして、重複した情報を記憶領域に格納しないという、情報の排除が、セグメントに基づいて行われることを示している。ストレージ装置は、重複情報に関する処理を、転送情報を複数のセグメント毎に分割、あるいはセグメント毎に処理することによって達成する。ストレージ装置は、一つのセグメントに含めるデータ量を、例えばブロック(例512バイト)の整数倍として、記憶デバイスへの格納単位(デステージブロック)と整合させるようにしている。

【0023】

転送情報には、転送情報を構成する複数のデータの組み合わせの形態が規則的なものである“固定パターン”と、これがランダムである“ランダムパターン”とがある。前者の例としては、複数のデータが同一であるもの、あるいは、複数のデータが規則的あるいは周期的に同じように繰り返すものがある。これを図3に基づいて説明する。

10

【0024】

図3の(a)は一つのブロック(BLK)に含まれる全てのデータが“ゼロ”であるパターンであることを示している。セグメント(SEG)はn個のブロックから構成されている。ストレージ装置は、(a)に該当するセグメントを“固定パターン”であると判定する。

これに対して(b)は、一つのブロックが同一の値のデータの複数から構成され、複数のブロックが“繰り返しサイズ”を構成し、複数の繰り返しサイズより繰り返し周期を構成することを示している。この繰り返し周期の間で、繰り返しサイズ毎にデータの値が連続的に変化する。ストレージ装置は、(b)についても、固定パターンと判定する。

20

【0025】

転送情報に含まれる複数のデータが全て同一あるいは複数のデータの組み合わせ・変化に規則性がある場合には、ストレージ装置は、転送情報を固定パターンと判定する。何ストレージ装置のマイクロプロセッサはローカルメモリの管理プログラムに基づいて、転送情報・セグメントのパターンを固定パターンと判定する。

【0026】

各セグメントは、固定パターン、ランダムパターン、又は両者が混在するパターンから構成される。なお、ストレージ装置は両者が混在するパターンを、通常、ランダムパターンとして扱う。

30

【0027】

図2に戻り説明を続ける。セグメントを単位としての重複の排除は、まず、固定パターンを対象に行う(図2の(b)(c))。ストレージ装置は、各セグメントに対し、例えば、図3の(a)、(b)のような規則性を持ったパターンを有しているか、そして、情報が格納されている記憶領域に同じ規則性から成るパターンを持つセグメントが存在するかを検証する。ストレージ装置は、この両方を肯定する、或いは規則性を持ったパターンを有することのみが肯定された場合に転送情報のうち固定パターンの判定の対象となったセグメントを“重複セグメント”(図2の(c))として、これを、記憶デバイスへ格納する対象から除外する。すなわち、重複セグメントは、記憶デバイスに格納される対象とはならない。互いに比較される複数のセグメント同士が規則的なパターンを有するか否かの判定については、パターン値によって直接判定すればよく既述のハッシュ値を求める演算までは必須ではない。ブロック内、セグメント内に含まれるデータが同一、或いは周期性がある場合、同一であるデータ値、規則性があるデータ値がパターン値に相当する。

40

【0028】

次に、ストレージ装置は、転送情報を構成する複数のセグメントから、固定パターンを有する重複セグメントを除いた残りのセグメントに対し、記憶領域に格納されたセグメントと、ランダムパターンが一致するか否かの判定を行い、同じパターンを有するセグメントを“重複セグメント”と判定する。ランダムパターンが一致するか否かは、互いに比較すべき複数のセグメントに対して、それぞれのハッシュ値が一致するか否かの判定結果によって決せられる。ハッシュ値が一致したセグメントは重複セグメント(図2の(d))

50

として以後の処理対象から除外される。

【0029】

次に、ストレージ装置は、図2(c)および(d)で特定された重複セグメントが除かれた転送情報に対して、データ圧縮処理を行う(図2(e))。データ圧縮の単位は、セグメントに拘束されず、複数セグメント、或いは、セグメントを構成する複数ブロック毎に行うことができる。

データ圧縮処理によって転送情報の情報量が減少する。各圧縮データの単位において、データ量は基本的にそれぞれ異なる。最後に、ストレージ装置は、圧縮データに基づいて、記憶デバイスへ格納される情報の単位であるデステージブロックを生成し、これを記憶デバイスへ格納する(図2(f))。

10

【0030】

図2に示す一連の処理において、固定パターンに着目して重複セグメントを判定することをランダムパターンに着目して重複セグメントを判定することに先行させる理由は、固定パターンに係わる重複セグメントを転送情報から予め除いて、ハッシュ値の演算対象となるセグメント数を減少させるためである。ランダムパターンの一致・不一致を高精度に行おうとすると、ハッシュ値の演算が重くなるためである。固定パターンを持つセグメントを極力ハッシュ値の演算対象を減少させることによって、転送情報を記憶デバイスに格納させるための処理のレーテンシーを最小に抑えることができる。

【0031】

また、重複排除のための処理では、セグメントそのものを排除しているため、データ圧縮に比べ、重複して情報を記憶デバイスに格納すること削減する効果が高い。そこで、重複セグメントを除外する処理を、情報の圧縮処理に先行して行う。特に、データ圧縮対象を、重複セグメントを排除した残りのセグメントに限定するので、圧縮処理におけるレーテンシーを最小に抑えながら、データ圧縮により記憶容量の効率的な利用を図ることができる。

20

【0032】

次に、図2の処理を実現する、ストレージ装置のより詳しい構成について説明する。図4はストレージ装置の機能ブロック図である。ストレージ装置100は、コントローラ、キャッシュメモリ、そして、マイクロプロセッサを連携させて重複排除処理(図2(a)-(d))、データ圧縮処理(図2(e)、(f))を実現する。なお、図4は、二重化されたコントローラのうち、代表して一方のコントローラに関する動作及び構成を示している。コントローラ1aは、ホストインターフェイス2aとキャッシュ4a間及び、Dual I/Oを介したもう一方のコントローラ1bとの間でのデータ転送を制御するHDMA8を有する。さらに、コントローラ1aは、固定パターンを持った情報(セグメント)に対して、パターン値を検出する固定データ検出手段(検出部あるいは検出回路)9と、キャッシュ4aに格納された情報、セグメントに対してハッシュ値を演算するランダムデータ検出手段10とを備える。

30

【0033】

コントローラ1aはさらに、キャッシュ4aに格納されたデータに対して圧縮処理を行うデータ圧縮手段11と、キャッシュ4a上のデータとドライブインターフェイス間でのデータ転送を制御するDDMA12とを備えている。

40

【0034】

次に、マイクロプロセッサ5aの動作について説明する。マイクロプロセッサ5aは図4に示すそれぞれの処理を行う処理部または処理手段を備える。すなわち、マイクロプロセッサは、パターン値、及び、ハッシュ値に基づくセグメントの重複判定、重複を判定した際に、その対象となったセグメントがキャッシュ領域に有する領域を開放する、すなわち、開放領域に他の転送情報が書き込まれることを可とするため排除指示、圧縮の対象となるキャッシュ領域を指定し、データ圧縮手段11の起動を制御する圧縮指示、DDMA12を起動し、キャッシュに存在する圧縮データをその格納領域からドライブインターフェイス3aへ転送することを制御するデステージ指示と、を実現する。

50

【0035】

マイクロプロセッサ5aは、転送情報をキャッシュメモリに格納する時に、ブロック単位でアドレス(Logical Address LA1)を付加する。上位装置は論理アドレスを頼りに、記憶デバイスに格納された情報にアクセスすることができる。しかしながら、既述の重複排除や、データ圧縮によって、上位装置から見えるアクセス空間におけるブロックアドレスと、それに相当するデステージブロックのアドレスとの対応が取れなくなってしまう。そこで、マイクロプロセッサはその対応を管理する目的でアクセス空間の領域管理を行う。

【0036】

マイクロプロセッサがデータ圧縮を行うとデータ量が減少するため、圧縮の対象となったブロックと、その圧縮データより生成したデステージブロックとの対応が一致しなくなる。従って、マイクロプロセッサは圧縮データから生成されたデステージブロックに新たにアドレス(LA2)を付加し、ホストからストレージ装置に情報を転送した時のブロックとデステージブロックと間の対応を図っている。

10

【0037】

マイクロプロセッサによるローカルメモリの利用について説明する。ローカルメモリには、ランダムパターンに対する重複検査に用いるものであり、記憶デバイスに格納済みのセグメント情報(例えば、アドレス)と、そのセグメントのハッシュ値との対応を管理するハッシュテーブルを備えている。転送情報から求めたハッシュ値と、ハッシュテーブルに登録されたハッシュ値との一致・不一致の検証を行うことによって、ランダムパターンについて、重複セグメントの有無を判定する。さらに、ローカルメモリ6aは、上位装置から見える、アクセス空間(論理ボリューム)のアドレスと、デステージブロックのアドレスとの対応を管理するセグメントテーブルを備える。これらテーブルの詳細については、後述する。

20

【0038】

次に、圧縮処理(圧縮データを含むデステージブロック(圧縮ブロック)生成例、ブロック構成例、圧縮ブロックの生成方法)について具体的に説明する。図5、図6に圧縮ブロックの生成例を示す。データ圧縮単位は、転送情報を成す複数のブロックから構成される。図5、図6では4ブロックでデータ圧縮単位を構成する例を示している。データ圧縮はこのデータ圧縮単位に対して行われる。データ圧縮単位には、転送ブロックに付加され、転送時のアドレス(LA1)等のヘッダも含まれる。ヘッダを除く転送情報のみを圧縮対象としても構わない。

30

【0039】

図5は、データ圧縮処理を行った結果、データ圧縮単位に含まれるデータ量(100%)が70%のデータ量に圧縮された、圧縮ブロックの生成を説明している。転送情報のデータ量に対して、70%となった圧縮データを、デステージブロック単位(1ブロックデータ量の割合は、 $100 \div 4$ ブロック=25%)に分割する。25%、25%、20%の割合で3つのデステージブロックに分割されることになる。圧縮後のデータの割り当て量が20%のブロックは5%分の無効データを含む圧縮ブロックとなる。

【0040】

コントローラあるいはマイクロプロセッサが、この圧縮ブロックを生成する際には、キャッシュメモリ上の圧縮データを格納する領域に対し、ブロックデータ量分のキャッシュアドレス範囲を指定し、その単位でキャッシュからデータを読み出すことによって、圧縮ブロックを生成する。ブロックデータ量を満たさない場合であっても、ブロックデータ量をキャッシュメモリから読み出すことで、無効データを含んだ圧縮ブロックの転送が実現される。或いは、無効データとしてマイクロプロセッサは特別なパターンをキャッシュメモリの該当領域に書込み、ブロックデータ量を満たすことも考えられる。マイクロプロセッサは、さらに、圧縮ブロックヘッダを付加して圧縮ブロックを生成する。

40

【0041】

図5の場合、圧縮単位(4ブロック)に対するデータを圧縮した結果、圧縮ブロックが3

50

ブロックになったことから、圧縮効果がある。圧縮ブロックを生成した結果、圧縮前のブロック数と圧縮後のブロック数と同数の場合もありうる。この場、マイクロプロセッサは圧縮の効果が無いものとして、圧縮データを破棄する。

【0042】

図5に示す圧縮ブロックの生成を、マイクロプロセッサ5a又はコントローラ1aの動作を示すフローチャートを用いて説明する。図7は、データ圧縮単位が一つである場合における、圧縮ブロックの生成方法を説明するものである。

【0043】

マイクロプロセッサは、データ圧縮単位から得られる圧縮データ量を把握する(S141)。圧縮データ量の把握は、キャッシュメモリにある圧縮データの格納範囲から求めることができる。S142において、圧縮データ量と、圧縮ブロックに割り当てるデータ量との間の除算を行うことにより、圧縮データ量に対する分割数と、剰余バイト数を計算する。

【0044】

S143にて、マイクロプロセッサは、剰余分を含む圧縮データ格納に必要なブロック数(分割数+1)と閾値との比較を行う。マイクロプロセッサは、例えばデータ圧縮単位に含まれるブロック数4に対し、圧縮効果が得られるブロック数3以下を閾値として与える。

【0045】

S143において、マイクロプロセッサは、閾値以下の場合、剰余バイトに続くキャッシュ領域に無効データを付加し(S144)、更に、圧縮ブロックヘッダを付加する。S144では、マイクロプロセッサはブロックデータ量をキャッシュメモリから読み出すことによって無効データの付加を行わない場合もある。最後に、マイクロプロセッサは、圧縮ブロックヘッダと、割り当てるデータ量分の圧縮データをキャッシュより読出し、圧縮ブロックとして、記憶デバイスへ格納し、セグメントテーブルを更新する。セグメントテーブルの更新については、図10で説明する。S143にて、閾値を超えた場合、キャッシュ上の圧縮データ格納領域は開放され、そして、圧縮データは破棄され(S146)、圧縮前の転送ブロックのまま記憶デバイスへ格納される。

【0046】

図6は、複数の圧縮データに対する圧縮ブロックの生成を示している。複数の圧縮データを対象とするため、圧縮効果が得られなかった圧縮データは除外される。3つの圧縮データの総データ量に対してブロック分割が行われる。従って、圧縮データ同士の境界を跨ぐブロック分割が行われ、ブロックデータ量に満たない分は、無効データを含むブロックに分割される。ブロック分割後、同様に圧縮ブロックヘッダの付加で圧縮ブロックを構成する。ブロック分割、圧縮ブロックの生成は、図5と同様、キャッシュ上の圧縮データ格納領域に対する指定、アクセスによって行われる。

【0047】

図6に示される圧縮ブロックの生成方法を説明する。図8は、複数のデータ圧縮単位を対象とした圧縮ブロックの生成方法を説明するフローチャートである。マイクロプロセッサは、各データ圧縮単位から得られる圧縮データ量を把握し、圧縮効果のない圧縮データを破棄する(該当するキャッシュ領域開放、S151)。圧縮データ量の把握は、キャッシュメモリ上の圧縮データ格納範囲から求めることができる。

S152において、マイクロプロセッサは、キャッシュメモリ上に残った圧縮データに対する総データ量を算出する。S153からS156の処理は、総データ量を基に、図7のS142からS145の処理に相当する。

S154において、図6の場合、マイクロプロセッサは、複数の圧縮データに対する転送ブロック数12に対し、圧縮効果が得られるブロック数11以下を閾値として与える。閾値は常に一定ではなく、処理の対象となる複数の圧縮データ数に対する転送ブロック数により可変としてもよい。

【0048】

S154において、マイクロプロセッサは、剰余分を含むブロック数(分割数+1)が、閾値を超えた場合、ブロック分割の対象とする圧縮データ数を判定する(S157)。S157に

10

20

30

40

50

において、マイクロプロセッサが単独の圧縮データであると判定すると、圧縮データを破棄し、それに対する転送ブロックを、非圧縮のまま記憶デバイスに格納する。もしくは、マイクロプロセッサは、図5で説明の単一の圧縮データに対する処理を行う。

【0049】

マイクロプロセッサがS157を不成立と判定すると、複数の圧縮データに対し、例えばそのデータ量が最大の圧縮データを破棄し（S158）、S152以降で、残りの圧縮データに対する処理を行う。S158において、マイクロプロセッサは、破棄された圧縮データに対する転送ブロックを非圧縮のまま記憶デバイスに格納する。以上説明した処理は、S151で排除される圧縮データの間に含まれる複数の圧縮データ単位を対象とするが、常に一定の圧縮データ数に対する処理が行われるわけではない。

10

【0050】

以上、圧縮ブロック生成の際にブロック分割を複数の圧縮データに対して行うことで、分割後のブロック内に含まれる無効データ量を抑制することが可能となり、単独の圧縮データを処理対象とする図5、図7の方法に対し、圧縮ブロック数を抑えることができる。

【0051】

図9は転送データブロックの構成例(a)、圧縮データを格納する圧縮ブロックの構成例(b)である。転送データブロックは、上位装置からのアクセスに用いられるブロックアドレス(LA1)と、補償コードとリザーブビットにより構成されるブロックヘッダと、転送データより構成される。圧縮データブロックは、圧縮データとそれに対するブロックヘッダより構成される。

20

【0052】

圧縮データは、ブロックヘッダを含む転送データブロック複数から成る圧縮単位に対するものである。ブロックヘッダは、圧縮データを対象とした記憶デバイスへの格納、アクセスを行うため、新たに付加される圧縮ブロックアドレス(LA2)と、補償コードからなる。更に、構成例(a)のリザーブビットに、例えば、無効データの有無を識別するパディングフラグ、無効データを除く圧縮データ量を示す有効バイト数、複数の圧縮データに対するブロック分割の結果、異なる圧縮データが含まれるブロックであることを識別するリンクフラグ、データ圧縮処理に適用された符号化アルゴリズム識別情報(符号化方法)を含める。

30

【0053】

リンクフラグは、記憶デバイスに格納の圧縮ブロックをキャッシュへ格納する場合、このフラグによって、ブロック分割前の圧縮データすべてが揃ったことを識別するものである。符号化方法は、圧縮データに対する復号時に、適用する復号アルゴリズムを識別するものである。更に圧縮データを含む圧縮ブロックを識別するものである。パディングフラグにより、圧縮単位に対する複数の圧縮ブロックの中で最後の圧縮ブロックであることが識別できる。

【0054】

圧縮ブロックの構成例(b)に含まれる補償コードの生成は、例えばキャッシュ上の圧縮データを読み出し、記憶デバイスへ格納する際に、圧縮データ読み出しと同時に補償コードを計算、圧縮ブロックヘッダとして付加する。

40

【0055】

図4に示す、ストレージ装置の機能ブロックに基づいて、図2で説明した、重複排除、データ圧縮、そして、デステージ動作のそれぞれを説明する。図2(b)のセグメント化は、コントローラ1aが、転送情報を格納する、キャッシュメモリの領域を、セグメント単位で管理することで実現される。キャッシュメモリの領域に対する開放の指示などのアクセスは、マイクロプロセッサの管理によって、コントローラ1aがセグメント単位で行う。HDMA8は、転送情報をキャッシュメモリへ転送すると同時に、Dual I/Oを経由して、他系のコントローラ1bと、キャッシュメモリ4bに転送し(マルチキャスト)、二重に転送情報を保持するようにしている。

【0056】

50

図2(c)に示す、固定パターンに対する重複排除は、HDMAが転送情報をキャッシュメモリ4aに格納すると同時に、転送情報を受領した固定データ検出手段9が、例えばブロック、或いはセグメント単位のデータ量毎にデータパターンを検証し、パターン値をマイクロプロセッサ5aに転送することにより行われる。マイクロプロセッサ5aは、得られたパターン値からセグメント単位で固定パターン(例えば、図3)があるか否かを検出する。

【0057】

パターン値については、固定データ検出手段は、HDMAからキャッシュメモリへ逐次格納される転送データに対し、バイト単位で検証を行った結果得られるものである。例えばブロック単位の転送データ量範囲で、全てのバイトデータが同一であることを判定し、更に連続する複数のブロック(セグメント)単位で、パターン値の同一或いは、周期性の判定結果に相当する。

10

【0058】

マイクロプロセッサ5aは、固定パターンを検出した際に、重複を判定し、マイクロプロセッサ5aは、“排除指示”を生成し、コントローラ1aを介してこの指示をキャッシュメモリ4aに送る。キャッシュメモリ4aは、固定パターンを持ったセグメントが格納されている領域を開放する。開放された領域は、新たなキャッシュアクセスに対する書込みが許可され、管理上は空き領域として管理される。更にもう他系のキャッシュメモリに二重に保持されている転送データに対しても、開放が適用される。すなわち、第1の系のキャッシュメモリ4aと、第2の系のキャッシュメモリ4bとの利用状況を同等とする。

20

【0059】

図2(d)で説明する、ランダムパターンを持つセグメントに対する重複排除のための処理は、固定パターンのセグメントが除かれた後のセグメントを対象に行われる。マイクロプロセッサ5aは、残セグメントを格納している、キャッシュメモリ4aの領域へアクセスする。ランダムパターン検出手段10の起動を制御する。同手段10はランダムパターンからなるセグメントを検出、当該セグメント毎にハッシュ値を算出し、マイクロプロセッサ5aにハッシュ値を転送する。

【0060】

マイクロプロセッサ5aは求めたハッシュ値とハッシュテーブルに格納済みセグメントに対するハッシュ値との一致・不一致を検証する(重複判定)。一致を判定した場合、固定パターンを持ったセグメントと同様に、ランダムパターンを持ったセグメントが占める、キャッシュメモリの領域を開放する。

30

図2(e)のデータ圧縮は、重複セグメントを除いた後の、キャッシュメモリ上のデータに対して行われる。マイクロプロセッサ5aは、データ圧縮手段11がキャッシュメモリ上の残セグメントにアクセスする領域を指定する。データ圧縮手段11は、データ圧縮単位(各セグメントの構成単位であるブロック或いは、複数ブロック単位)で圧縮処理を行う。マイクロプロセッサ5aは、圧縮されたデータを、キャッシュメモリ上に確保した別領域に格納する。圧縮データの格納に利用されるキャッシュ領域は、例えばデータ圧縮単位に含まれるデータ量と同等のサイズ分確保される。確保した領域へ圧縮データの格納が行われ、確保領域を超えた書き込みが行われた、或いは、データ圧縮単位に対し、圧縮データ量がほぼ同等で、圧縮効果が得られない場合がある。この場合、マイクロプロセッサ5aは、格納した圧縮データを破棄する(領域開放)。

40

【0061】

図2(f)に説明した、デステージブロック生成は、圧縮データに対して行われる。データ圧縮手段11は、キャッシュ上に格納された圧縮データを、ブロック単位のデータ量毎にオリジナルのアドレス(LA1)に新たにアドレス(LA2)、圧縮データに対する補償コード(CRC)を含むブロックヘッダを付加して成るデステージブロックを生成する。図7、図8によってデータ圧縮効果が得られなかったブロックについては、圧縮前のブロックがデステージブロックとなる。

【0062】

50

マイクロプロセッサ 5 a は、D D M A を起動し、キャッシュメモリ上に格納された圧縮データの読出しと同時に、上記説明に従い圧縮ブロックに対するヘッダを付加によってデステージブロックを生成、ドライブインターフェイス 3 a を介して、記憶デバイスにデステージブロックを格納する。格納後、キャッシュメモリのデステージブロック領域を開放する。もう一方のキャッシュに対しても同様に該当領域を開放する。

【 0 0 6 3 】

なお、図 3 では、パターン値、ハッシュ値の検出、データ圧縮、そして、デステージブロックの生成はコントローラ 1 a 内で行い、キャッシュメモリ 4 a がコントローラ 1 a と接続される構成を一例として説明したが、キャッシュメモリ 4 a がマイクロプロセッサ 5 a に接続される構成であってもよい。

更に、コントローラ 1 a の固定パターン検出、ランダムパターン検出、そして、データ圧縮の各処理はマイクロプロセッサ 5 a が行うことでもよい。この場合、マイクロプロセッサが転送情報をキャッシュメモリに格納すると同時には固定パターンの検出ができないことも考えられる。その場合は、ランダムパターンの検出と同様に、キャッシュメモリに転送情報を格納し後、キャッシュメモリの転送情報に対して固定パターンの検出を行うことでもよい。

【 0 0 6 4 】

本発明は、ハッシュ値によるパターン検出を、固定パターンの検出にも適用することを必ずしも妨げるものではない。固定パターンの種類は限定されたものであることから、複雑な計算アルゴリズムを用いなくとも、ハッシュ値の衝突を回避できる。従って、固定パターン検出時には、ランダムパターンに適用のアルゴリズムの短縮（繰返し演算回数の制限）や、処理ステップ数が少ないアルゴリズムを適用して、処理レーテンシーを抑制することができる。ハッシュテーブルもランダムデータとは別のテーブルを備えればよい。マイクロプロセッサで、固定パターンを検出しようとする場合には、ランダムパターンを検出するための実行アルゴリズムとは変更したものを使用すればよい。

【 0 0 6 5 】

なお、キャッシュメモリと各デバイスとの間のデータ転送は、コントローラに備える D M A (H D M A 、 D D M A) を介したものに限定されず、マイクロプロセッサに備える D M A 或いはホスト、ドライブ I / F 内に備える D M A を介することもできる。この場合コントローラは、H D M A 、 D D M A を含まず、ホストインターフェイス、ドライブインターフェイスとマイクロプロセッサを接続し、転送データの分配を制御する機能のみを備えた手段となることも考えられる。

【 0 0 6 6 】

更に固定データ検出手段、ランダムデータ検出手段、データ圧縮手段も、コントローラ内に備えることに限定されない。コントローラとは別的手段、例えば専用ハードウェアをマイクロプロセッサに直結する場合や、固定パターンの検査、ハッシュ値の算出、データ圧縮処理をマイクロプロセッサによる演算処理で実現することも考えられる。

【 0 0 6 7 】

次に、ローカルメモリ (6 a , 6 b) に存在するハッシュテーブルについて説明する。図 1 0 はハッシュテーブルの一例を示すものであり、テーブルの構成は、上位装置から見えるボリューム空間に与えられる転送ブロックアドレス (B l o c k L A) と、重複排除を行う処理の単位である Segment の番号 (アドレス) と、デステージフラグ (N o 又は Y e s) と、ハッシュ値 (H a s h m e a n) と、デステージブロックアドレス (D e - s t a g e L A) とが対応されて登録されたもの構成される。マイクロプロセッサは、ハッシュテーブルを図 2 に示す処理を行う過程で作成する。ボリューム空間が複数存在する場合には、このテーブルもボリューム毎に存在する。

【 0 0 6 8 】

図 1 0 は、セグメントが全て同一のブロック数で構成される場合を示している。ストレージ装置が、セグメント長を変更できる場合や、ボリューム空間毎にセグメント長を変更する場合、そのセグメントに含まれるブロックアドレス範囲を変更する。図 1 0 は 1 セグ

10

20

30

40

50

メント = 8 ブロックであることを説明している。

【 0 0 6 9 】

デステージフラグは、セグメント毎の重複排除の実施 (No)、または記憶デバイスへの格納 (Yes) をマイクロプロセッサが識別できるようにするためのものである。重複排除が実施された場合、デステージブロックアドレスに、重複セグメントと同一パターンデータが格納されたセグメントに対するアドレスが登録される。これによって、マイクロプロセッサおよびコントローラは、重複セグメントに同一データパターンが格納された記憶領域へアクセスすることができる。排除の対象となったセグメントに対しては、デステージブロックアドレスの代わりに、記憶領域に格納されている同一のデータパターンを有するセグメントの番号を登録するものであってもよい。

10

【 0 0 7 0 】

マイクロプロセッサは、転送情報のセグメントのデータを記憶デバイスへ格納する場合、当該セグメントに対するハッシュ値をハッシュテーブルに登録する。マイクロプロセッサは、重複セグメントがあるか否かの判定に際し、転送情報のセグメントについてハッシュ値を計算し、計算したハッシュ値に同等なハッシュ値がハッシュテーブルに存在するか否かを検証する。

【 0 0 7 1 】

マイクロプロセッサは、ハッシュテーブルにより、上位装置から見えるボリューム空間に対するブロックアドレスと、デステージブロックアドレスとの対応、つまりボリューム空間と、記憶デバイスのアクセス空間との対応を管理する。したがって、転送情報を送信する要求の対象となったアクセスブロックから、対応するデステージブロックへのアクセスが容易になる。

20

【 0 0 7 2 】

次にセグメントテーブル (図 1 1) について説明する。テーブルの構成は、図 1 0 と同様に、転送ブロックアドレス、重複排除の単位であるセグメント番号 (アドレス)、デステージブロックアドレス、更に重複排除、データ圧縮の識別フラグから構成される。ボリューム空間が複数存在する場合には、このテーブルもボリューム毎に存在する。

【 0 0 7 3 】

セグメント番号は図 1 0 と同様であり説明は省略するが、セグメントに含まれるデータブロック数は一定としている。固定パターン、ランダムパターンを持ったセグメントの重複を排除する毎にセグメントの大きさを変更することも可能である。重複排除に対する識別フラグ (重複フラグ) は、行われた重複排除の種類 (固定データ : Fixing、ランダムデータ : Random) と、排除の発生が無いことを識別する (No) ことができるものである。

30

【 0 0 7 4 】

データ圧縮に対する識別フラグ (圧縮フラグ) は、圧縮処理の有無を識別できるものである。重複排除されたセグメント (重複フラグ = Fixing or Random)、データ圧縮を行わないブロック (転送アドレス 0x00 1 0 ~ 1 3) については、圧縮フラグとして「No」が登録されている。更に圧縮効果が得られないブロック (転送アドレス 0x00 1 C ~ 1 F) については圧縮フラグとして「オーバーフロー」が登録されている。データ圧縮単位は、重複排除単位 (セグメント) に一致させなくてもよい。図 9 では一例として 2 ブロックをデータ圧縮のための単位としている。

40

【 0 0 7 5 】

ここで、セグメントテーブル内のデステージブロックアドレスに注目して欲しい。セグメント 0x0000、0x0001 については固定パターン、又は、ランダムパターンを持ったセグメントに対する重複排除の処理が行われ、重複するデステージブロックアドレス、或いは、これに代えて、同一パターンを持ち、記憶デバイスに記憶されているセグメントの番号が登録される。

【 0 0 7 6 】

固定パターンには上位装置で利用されないデータパターン (例えば記憶デバイスに対するフォーマットパターン) も存在する。このようなパターンを含むセグメントが排除され

50

た場合には、ストレージ装置は、デステージブロックアドレスを管理する必要がない。或いはそれらのパターン全てを排除の対象とし、すなわち、重複セグメントとして記憶デバイスへ格納しない場合もある。

【0077】

図11において、ブロックアドレス0x0010~13、0x001C~1Fの範囲は圧縮が行われなかったか、或いは圧縮効果の無かったデータブロックであり、このブロックそのもののサイズを単位としてデステージアドレスが登録される。マイクロプロセッサは、ブロックアドレス0x0014~17、0x0018~1Bで示されるデータ圧縮単位に対し、圧縮データに対するデステージブロックアドレスをブロックアドレスに割り当てている。図11の例では、2ブロックをデータ圧縮単位としており、効果の得られる1ブロックに圧縮された場合のデステージブロックアドレスを割り付けている。データ圧縮によって、データ圧縮単位のデータ量を減少できるため、マイクロプロセッサ及びコントローラは、図5或いは、図6で説明した方法を用いて、圧縮ブロック(デステージブロック)を生成する。

10

【0078】

以上によって、マイクロプロセッサは、セグメントテーブルに基づいて、上位装置から見えるボリューム空間に対するブロックアドレスと、デステージブロックアドレスとの対応を確認することができる。マイクロプロセッサ及びコントローラは、上位装置から転送要求のあったアクセスブロックに対応するデステージブロックへのアクセス、圧縮データに対する伸長処理の実施を行うことができる。

なお、セグメントテーブルは、既述の構成に限定されるものではなく、上位装置から見えるボリューム空間に対するブロックアドレスと、デステージブロックアドレスとの対応が取れるものであればいかなる形態を持つこともできる。圧縮データの判定は、圧縮ブロックにおける圧縮ブロックヘッダの分析(例えば符号化方法を示すフラグ)でも行うことができる。

20

【0079】

また、セグメントテーブルを重複セグメントの排除、データ圧縮、未処理ブロックの各テーブルに分けて管理しても構わない。上位装置から要求された転送アドレスを、例えば重複排除テーブルまたは、未処理ブロックテーブル、圧縮テーブルの順にテーブルを選択する。テーブルの参照順について、特にデータ圧縮は、長時間アクセスの無かったボリュームや、その中の領域を対象に行われ、上位装置からのアクセス確率は低いと考えられる。従って、テーブルのアクセス順も最後とする。更に、参照したテーブルで管理されるデステージブロックと、テーブルの種類に従って、上位装置への転送データを処理(圧縮データに対する伸長処理等)する。以上、テーブルは様々な構成が考えられるが、その内容は、図10及び図11において説明したものと同様である。

30

【0080】

次に、改めて、図2及び、図3で説明した、転送情報の重複の排除、データ圧縮処理の流れをフローチャート(図12)を用いて説明する。ストレージ装置は、通常、転送情報を受信した後、これらの処理を順次処理(オンライン処理)で進める。オンラインとは、上位装置からのアクセス(データ書込みまたは、読出し)命令がコントローラに対し連続的に発行され、コントローラが記憶デバイスへ連続的にアクセス(格納、読出し)している状態である。

40

【0081】

一方、コントローラが上位装置からのアクセス命令に対するレスポンスを重視する場合や、ストレージ装置の装置構成の如何によっては、ストレージ装置は、複数の処理をオフライン或いはニアランで進める。オフライン、或いはニアラインとは、上位装置からのアクセス要求が無い或いは、頻度が少ない状態である。

【0082】

コントローラ1aのHDMA8は、上位装置から転送情報を受信し、さらに、キャッシュメモリ4aに転送情報を格納する(S701)。マイクロプロセッサ5aは、重複排除1(固定パターンの排除)の要否を判定する(S702)。この判定の際、マイクロプロセッサ5

50

a は、固定データ検出手段より得られる転送情報のパターン値に基づいて、固定パターンが含まれるセグメントを判定する。マイクロプロセッサがこのセグメントの排除が要と判定すると、セグメント単位で固定データの重複を排除する (S703)。

【0083】

S702の判定が、重複するセグメントの排除 (S703) をスキップする場合、マイクロプロセッサ 5 a 又はコントローラ 1 a は、S704、又はS706への移行や、重複排除、データ圧縮全てをスキップしてS709へ移行する。特に、S709へスキップすることは、ストレージ装置がホストからのアクセス命令に対するレスポンスを重視する場合に選択され、重複排除、圧縮処理はデステージ後のオフライン (ニアライン) 状態で行われる。

【0084】

処理がS704へスキップする場合、固定パターンをランダムパターンの一つとして扱い、固定パターンを持ったデータについての重複排除を、ランダムパターンを持ったデータの重複排除で纏めて扱う場合が考えられる。

【0085】

S704で、マイクロプロセッサ及びコントローラは重複排除 2 (重複したランダムパターンの除外) の要否を判定する。重複したセグメントを除外することを肯定すると判定する場合、S705において、キャッシュメモリに格納されたデータに対しセグメント単位での重複を除外する処理を行う。S704の判定が、この除外を否定する場合は、マイクロプロセッサ及びコントローラは、重複排除 2 の処理 (S705) をスキップし、S706へ移行するか、あるいはデータ圧縮をスキップしてS709へ移行する。S709へスキップするのは、たとえば、前処理 (固定パターン重複排除) の結果、未排除のセグメント数がゼロとなった場合である。

【0086】

以上説明した中で、特に、圧縮処理がスキップされて、処理がS709へ移行する場合、データ圧縮の対象とするデータの条件が限定されることがある。非圧縮のデータのままデステージされた後、上位装置からのアクセスが長時間発生していないボリューム空間であること或いは、ボリューム空間の一部でそのような領域であることをマイクロプロセッサが判定する。マイクロプロセッサは、該当するデータ領域に対してデータ圧縮を行う。一方、頻繁にアクセスが発生する領域は非圧縮とする。これによって、マイクロプロセッサは伸長時間を費やすことを避け、転送命令に対する装置レスポンスを最優先とする。データ圧縮は、ストレージ装置がオフライン (ニアライン) 状態で実施される。

【0087】

S706では、マイクロプロセッサは、データ圧縮処理の要否を判定する。肯定判定する場合、図 5、6 で説明した内容に従い、圧縮ブロックを生成し (S707、S708)、デステージ処理を行う (S709)。

マイクロプロセッサが、S702、S704、S706において種々の判定を行う場合、キャッシュメモリ内の残セグメント数に注目することが考えられる。図 13 に示すように、マイクロプロセッサはキャッシュメモリ内のセグメント数を把握する (S801)。セグメント数は、マイクロプロセッサがコントローラを介してキャッシュメモリからデータ量を求め、セグメント当たりのデータ量で除算することにより得られる。マイクロプロセッサは、セグメント数と閾値 M (M は正の整数で $M \geq 1$) を比較する (S802)。(セグメント) $\geq M$ の条件で、マイクロプロセッサはS709を実行し、ブロックデータのデステージを行う。(残セグメント) $< M$ の条件で、マイクロプロセッサは、更にセグメントの有無を判定する (S803)。(残セグメント数) = 0 の場合、処理の対象となるセグメント或いは、デステージデータが存在しないので、処理は終了となる。S803で (残セグメント数) = 0 の場合、各処理が引き続き実行される。図 13 の処理を、S702の判定に適用する場合、S703或いはS704、S706、S704の判定に適用する場合、S705或いはS706、S706判定に適用する場合、S707から処理が行われる。

【0088】

S802の判定からS709への分岐は、処理対象のデータ量が閾値 M 以上の場合、つまり、ス

10

20

30

40

50

ストレージ装置が、上位装置に対してオンラインで状態であって、ストレージ装置から上位装置へのレスポンスを優先させる場合に実現される。

【0089】

マイクロプロセッサは装置負荷（ニアライン、オフライン）により、Mを変化させてもよい。マイクロプロセッサは、装置負荷が軽い状態ではMを増加し、重複排除の処理、及び、圧縮処理が行われるデータ量を増やす。

【0090】

装置負荷の判定は、マイクロプロセッサが行い、1セグメントに対する重複排除に要する時間、或いは、データ圧縮単位に対するデータ圧縮に要する処理時間と、単位時間あたりにストレージ装置に対するアクセス命令が発生する頻度から決められる。

10

【0091】

既述の実施形態にかかわる、情報処理方法は、上装置からストレージ装置への転送情報に含まれるデータのパターンの相違に注目して、ストレージ装置が、重複する情報を段階的に、記憶する対象から除外をするようにし、さらに、除外された情報以外の残りの情報に対して圧縮処理を行うようにしたので、情報の受信から記憶デバイスへの格納までに要する時間を短縮し、そして、ストレージ装置が上位装置に記録容量を効率的に利用させることができる。

【0092】

なお、既述の実施形態では、重複したセグメントのデータをキャッシュメモリから開放されたが、開放しないか、あるいは、転送情報が2重にキャッシュメモリに記録されているかぎり、第1の系のキャッシュメモリについては開放を行い、第2の系のキャッシュメモリについては開放しないようにしてもよい。このようにすることによって、キャッシュメモリやコントローラ内部に発生した障害によって、転送データが消失されることを回避できる。

20

【0093】

次に、記憶デバイスに格納された転送情報に対して、重複を排除する処理、データ圧縮の処理について説明する。これらの処理は、例えば、図12の各判定（S702、S704、S706）において、処理がスキップされる場合で、ストレージ装置がオフライン又はニアラインの状態で行われる。各処理が実施されたか否かの判定は、マイクロプロセッサがセグメントテーブルを分析することで可能となる。図14はこのことを示すフローチャートである。

30

【0094】

マイクロプロセッサは、まずローカルメモリ上のセグメントテーブルを分析し、重複排除処理、圧縮処理が実施されていないデステージブロックを特定する（S901）。図11のセグメントテーブルにおいて、圧縮フラグ、重複フラグより、非圧縮ブロック（オーバーフローフラグ無し）と、重複排除の対象とならなかったセグメントを特定する（S902、S903）。更に、装置状態がオフライン（ニアライン）であることを判定する（S904）。

【0095】

上位装置からのアクセス命令発生の場合はその処理を優先して実施する（S905）。マイクロプロセッサが、オフライン（ニアライン）を認定した場合は、S902、S903で特定したセグメント、非圧縮ブロックに対する処理を順に行う（S906、S907）。S906では図2と同様、固定パターンに対する重複排除を先に実施する。

40

【0096】

図11のセグメントテーブルの場合、重複排除の対象となるセグメントとして0x0002、0x0003、0x0004がある。圧縮フラグより、セグメントに含まれるブロックが全て非圧縮であるセグメントは0x0004のみである。従って、S902の判定によって重複排除の対象はセグメント0x0004となる。

非圧縮ブロックは、圧縮フラグにより特定される。圧縮効果の得られないオーバーフローブロック、圧縮ブロック以外のブロックが相当する。マイクロプロセッサは、非圧縮ブロックに対し圧縮後、図5、6で説明した方法によって、圧縮ブロックを生成する。

50

【 0 0 9 7 】

次に、ストレージ装置が、既述のデータ圧縮、重複排除処理を経て記憶デバイスに格納されたブロックデータを上位装置へ転送（再生）する方法について説明する。

【 0 0 9 8 】

図 15 は、その方法を実現するための、ストレージ装置の動作を説明するフローチャートである。ストレージ装置は、上位装置からのアクセス命令に対し、ボリューム空間におけるアクセス領域（ブロックアドレス）を検出する（S101）。マイクロプロセッサは、セグメントテーブルを参照し、アクセスブロックアドレスに対するデステージブロックアドレスを求め、記憶デバイスからブロックアドレスに対応するブロックデータをリードし、ブロックデータをキャッシュメモリに格納する（S102）。

10

【 0 0 9 9 】

S103において、マイクロプロセッサは、キャッシュメモリに格納されたブロックのヘッダからアドレス情報を得、さらに、セグメントテーブルを参照して求めたデステージブロックアドレスとの一致・不一致を判定する。マイクロプロセッサは、不一致を判定した場合は、フローチャートをエラー終了する。この判定によって、ストレージ装置は所望のデステージブロックがキャッシュメモリに格納されたことを補償する。

【 0 1 0 0 】

S104で、マイクロプロセッサはキャッシュメモリに格納された情報を判別する。判別はキャッシュメモリに格納したブロックに対する圧縮フラグを用いて行われる。S104において、マイクロプロセッサが、成立（格納された情報は圧縮されたもの）を肯定判定すると、圧縮情報に対する伸張を行い、キャッシュメモリの別領域に伸張後の情報（転送ブロックデータ）を格納する（S105）。更に、マイクロプロセッサは、伸長により得られた転送ブロックを構成し、上位装置からのデータ転送時に付加されたブロックアドレス（LA1）と、アクセス命令に対するブロックアドレスとの一致・不一致（S106）を判定する。不一致を判定すると、一連の処理をエラー終了させる。一致を判定すると、キャッシュメモリのブロック情報を上位装置に転送する（S107）。以上のように、ストレージ装置は、デステージデータの格納時、圧縮データを伸長した後、上位装置への転送前に、ブロックアドレスの検証を段階的に行う。

20

【 0 1 0 1 】

S104において、不成立が判定されると、S108にて、キャッシュメモリに格納された情報が、重複排除処理の対象となったセグメントに含まれているか否かを判定する。S108の判定が不成立であった場合、キャッシュメモリに格納された情報は非圧縮のもので、排除の対象とならなかった情報であると判断され、S106に移行する。S108の判定が成立であると、排除の対象となった情報と判断される。この場合アクセス要求アドレスと、キャッシュメモリに格納されたブロックアドレスは一致せず、そのままS107を実行する。以上の条件判定、処理の起動制御などは、マイクロプロセッサにより行われる。

30

【 産業上の利用可能性 】

【 0 1 0 2 】

本発明は、転送情報に含まれるデータのパターンに注目し、最適なパターン検出、重複排除処理を段階的に実施する。更に、重複排除後のデータを対象にデータ圧縮処理を実施する。更に、段階的に行われる重複排除、圧縮処理の実施判定を、各処理段階の残データ量や、装置負荷により判断する。したがって、本発明は、キャッシュメモリやホスト、記憶デバイス、及び、コントローラやプロセッサを備えて構成されるストレージ装置に広く適用される。

40

【 符号の説明 】

【 0 1 0 3 】

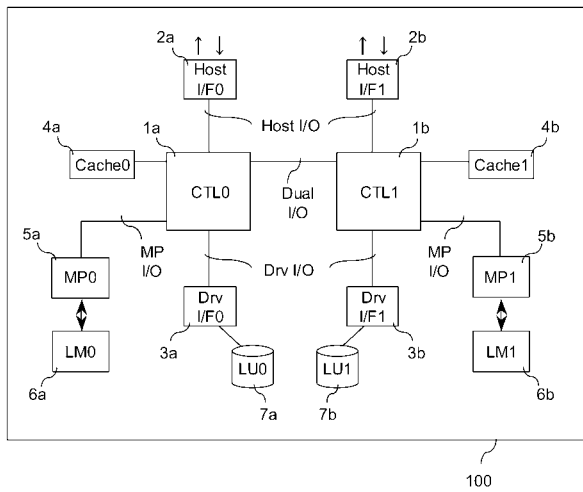
- 1 コントローラ
- 2 ホストインターフェイス
- 3 ドライブインターフェイス
- 4 キャッシュメモリ

50

- 5 マイクロプロセッサ
- 6 ローカルメモリ
- 7 記憶デバイス
- 8 HDMA
- 9 固定データ検出手段
- 10 ランダムデータ検出手段
- 11 データ圧縮手段
- 12 DDMA

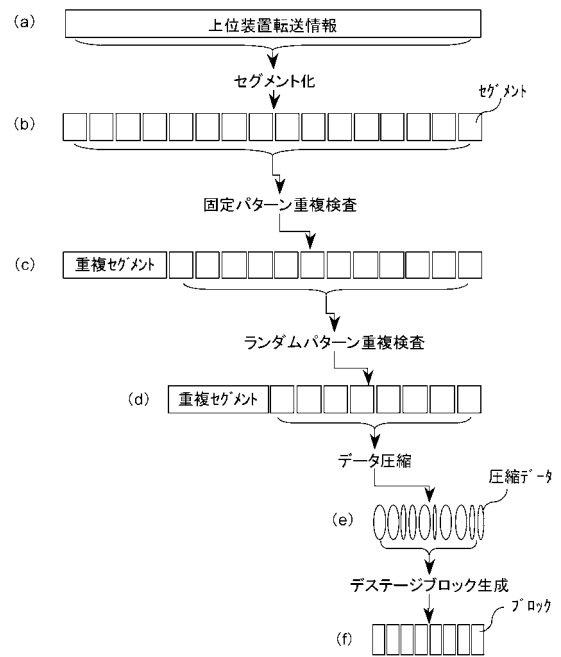
【 図 1 】

図 1

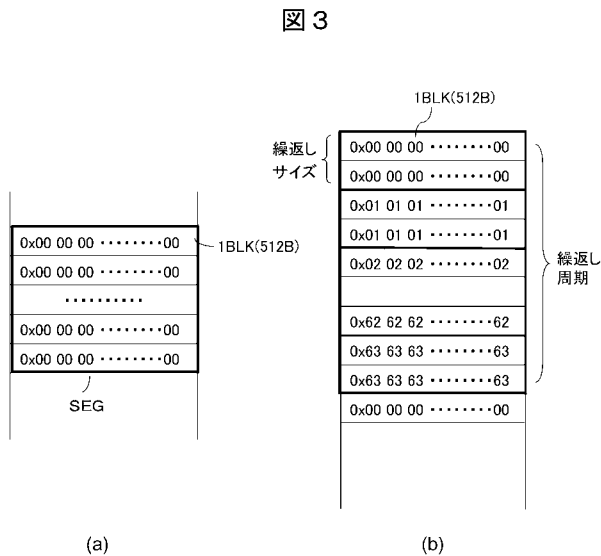


【 図 2 】

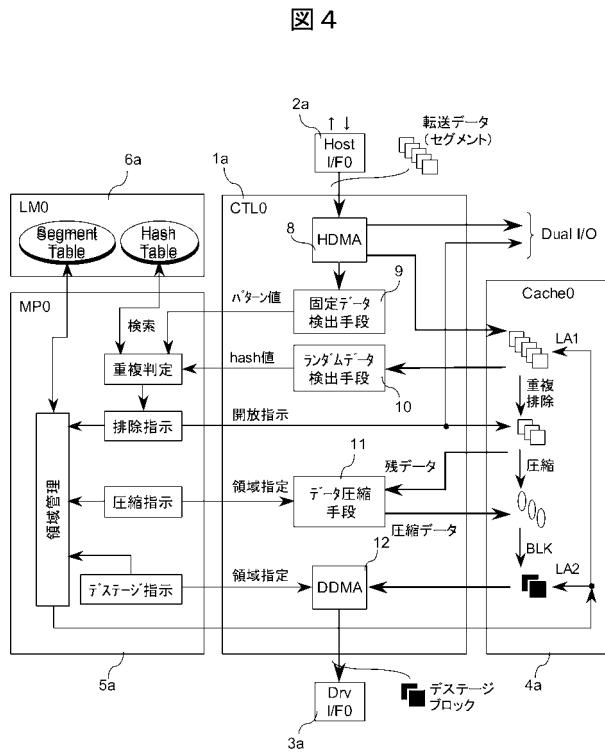
図 2



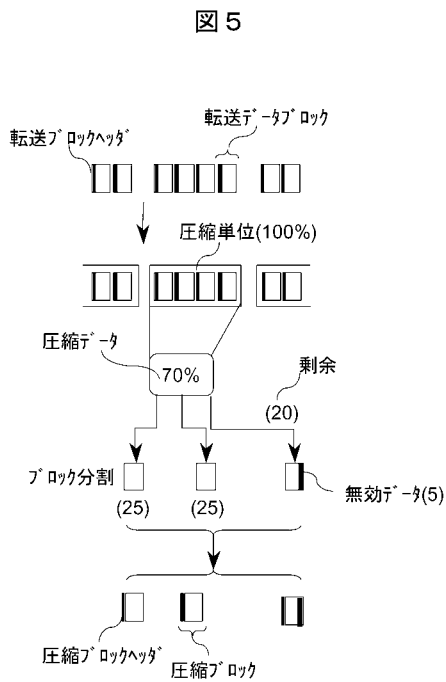
【 図 3 】



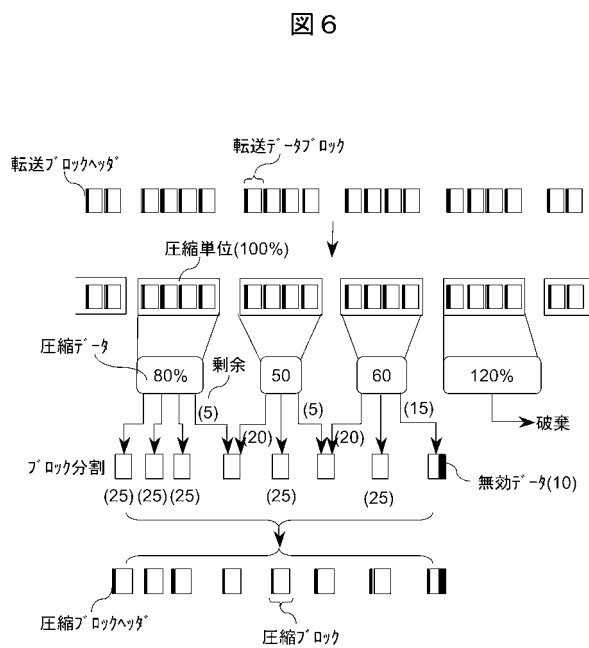
【 図 4 】



【 図 5 】

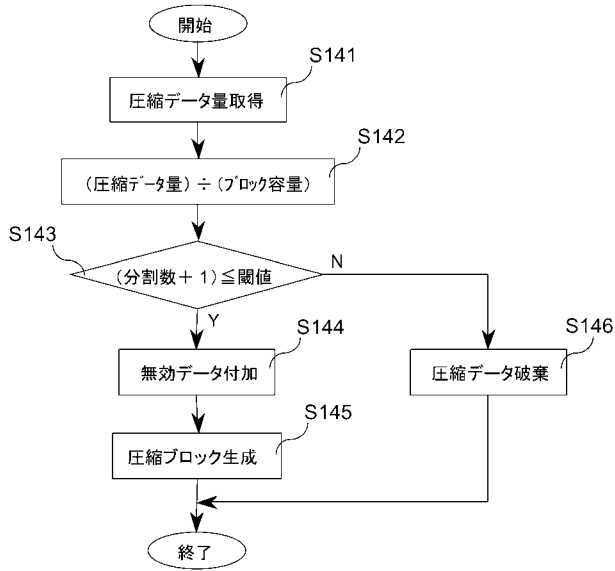


【 図 6 】



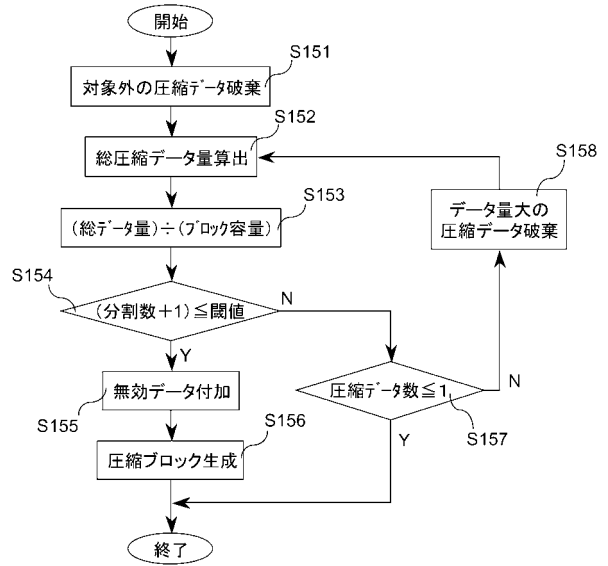
【 図 7 】

図 7



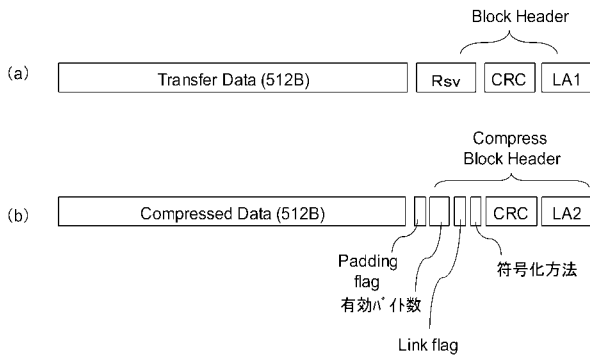
【 図 8 】

図 8



【 図 9 】

図 9



【 図 10 】

図 10

Block LA	Segment*	De-stage	Hash mean	De-stage LA			
0x0000	0x0000	No	-	0xFF00			
0x0001				0xFF01			
...				...			
0x0007				0xFF07			
0x0008				0x0001	No	-	0xFF08
0x0009							0xFF09
...							...
0x000F							0xFF0F
...	...						
0xFF00	0x1F00	Yes	0x000012				0xFF00
0xFF01							0xFF01
...							...
0xFF07				0xFF07			
0xFF08				0x1F01	Yes	0x000034	0xFF08
0xFF09							0xFF09
...							...
0xFF0F							0xFF0F

* 1segment = 8 Block (4KB)

【 図 1 1 】

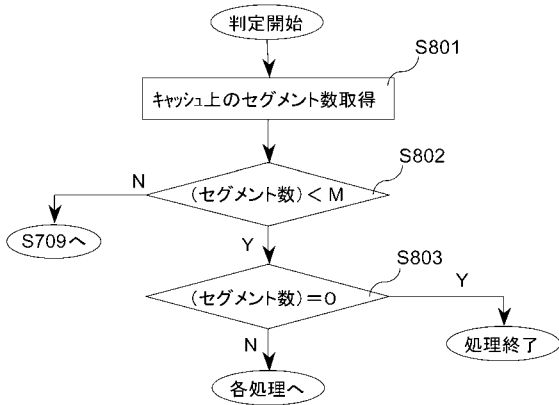
図 1 1

Block LA	Segment*	De-duplication	Compression	De-stage LA
0x0000	0x0000	Fixing	No	0xFF00
0x0001			No	
....				
0x0007	0x0001	Random	No	0xFF08
0x0008			No	0xFF09
0x0009			No
....	0x0002	No	No	0xFF0F
0x0010			No	0x0000
0x0011			No	0x0001
0x0012	0x0003	No	No	0x0002
0x0013			Yes	0x0003
0x0014			Yes	0x0004
0x0015	0x0004	No	Yes	0x0005
0x0016			Yes	0x0006
0x0017			Yes	0x0007
0x0018	0x0004	No	Over Flow	0x0008
0x0019			Over Flow	0x0009
0x001A			Over Flow	0x000A
0x001B	0x0004	No	No	0x000B
0x001C			No	0x000C
0x001D			No	0x000D
0x001E	0x0004	No	No	0x000E
0x001F			No	0x000F
0x0020			No	0x0010
0x0021	0x0004	No	No	0x0011
0x0022			No	0x0012
0x0023			No	0x0013
0x0024	0x0004	No	No	0x0012
0x0025			No	0x0013
0x0026			No	0x0013
0x0027

* 1segment = 8 Block (4KB)

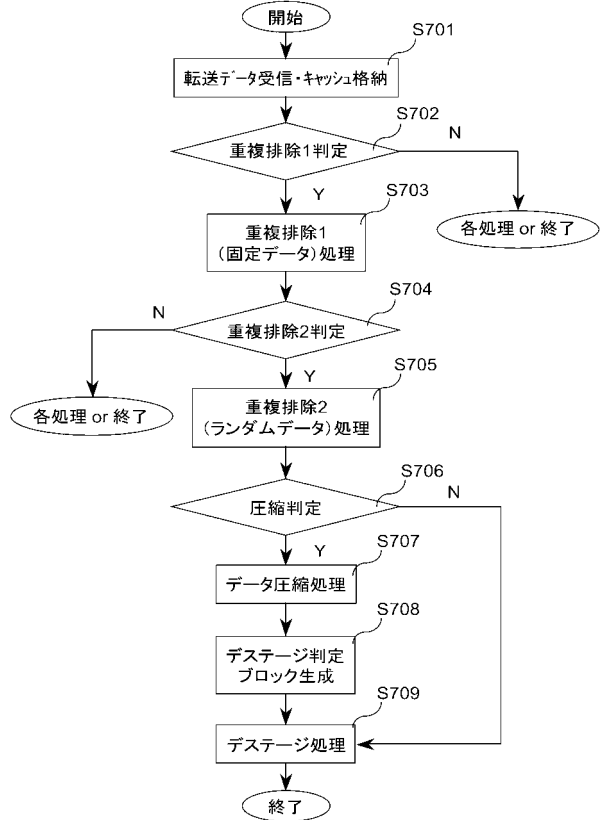
【 図 1 3 】

図 1 3



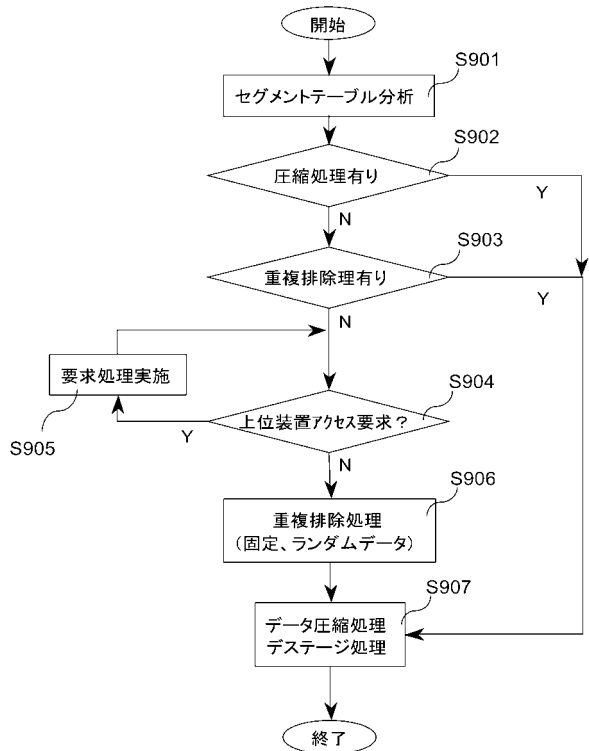
【 図 1 2 】

図 1 2



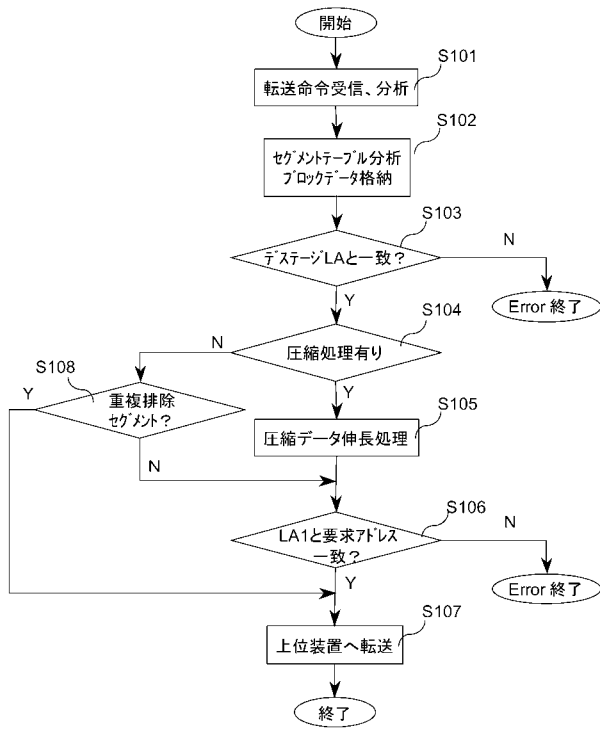
【 図 1 4 】

図 1 4



【 図 1 5 】

図 1 5



【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2009/054004

A. CLASSIFICATION OF SUBJECT MATTER INV. G06F11/14 H03M7/30		
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) G06F H03M		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 928 526 B1 (ZHU MING BENJAMIN [US] ET AL) 9 August 2005 (2005-08-09) cited in the application	1, 13
Y	paragraph [0050] paragraph [0026] paragraph [0020] paragraph [0018] paragraph [0022] paragraph [0017] paragraph [0052]	2-12
Y	WO 03/081433 A1 (NOKIA CORP [FI]; HAETOENEN KIMMO [FI]; MIETTINEN MARKUS [FI]) 2 October 2003 (2003-10-02) page 4, line 20 - page 5, line 3	2-12
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 document 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 2 July 2009		Date of mailing of the international search report 10/07/2009
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 LV Rijswijk Tel. (+31-70) 340-2040 Fax (+31-70) 340-3016		Authorized officer Filip, Liviu

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/JP2009/054004

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
US 6928526	B1	09-08-2005	US 2008183767 A1	31-07-2008
			US 2005216669 A1	29-09-2005
WO 03081433	A1	02-10-2003	AU 2002244892 A1	08-10-2003
			EP 1490769 A1	29-12-2004
			US 2005138483 A1	23-06-2005

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(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, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW