(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
2008 年3 月6 日 (06.03.2008)

PCT

(10) 国際公開番号
WO 2008/026497 A1

(51) 国際特許分類:
G06F 3/06 (2006.01)

(21) 国際出願番号: PCT/JP2007/066339

(22) 国際出願日: 2007 年8 月23 日 (23.08.2007)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願2006-230914 2006 年8 月28 日 (28.08.2006) JP

(71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 Tokyo (JP).

(72) 発明者; および

(75) 発明者/出願人 (米国についてのみ): 杉本 欽一 (SUGIMOTO, Kinichi) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP).

(74) 代理人: 高橋 勇 (TAKAHASHI, Isamu); 〒1010031 東京都千代田区東神田１丁目１０番７号 篠田ビル７階 Tokyo (JP).

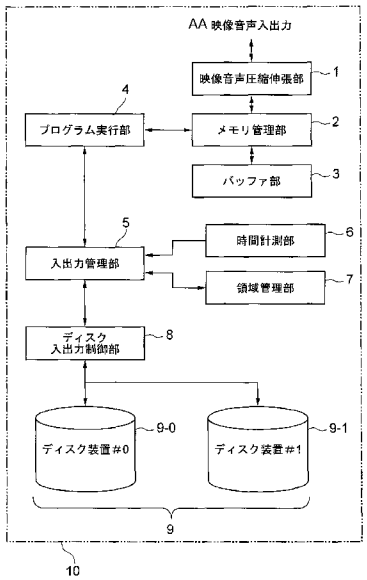
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK,

[続葉有]

(54) **Title:** DISC ARRAY CONTROL DEVICE, DISC ARRAY CONTROL METHOD, AND DISC ARRAY CONTROL PROGRAM

(54) 発明の名称: ディスクアレイ制御装置、ディスクアレイ制御方法、ディスクアレイ制御プログラム



AA VIDEO/AUDIO I/O
1 VIDEO/AUDIO COMPRESSION/DECOMPRESSION UNIT
2 MEMORY MANAGEMENT UNIT
3 BUFFER UNIT
4 PROGRAM EXECUTION UNIT
5 I/O MANAGEMENT UNIT
6 TIME COUNTING UNIT
7 REGION MANAGEMENT UNIT
8 DISC I/O CONTROL UNIT
9-0 DISC DEVICE #0
9-1 DISC DEVICE #1

(57) **Abstract:** [PROBLEMS] To guarantee a stable data I/O speed as a system and realize effective disc device operation management and data arrangement. [MEANS FOR SOLVING PROBLEMS] A disc array control device (10) includes an I/O managing unit (5) which performs input/output by striping to the number of disc devices selected so that the total of I/O speeds at the address as a data I/O object exceeds a predetermined threshold value. Thus, the number of disc devices for which data is subjected to striping is decided in accordance with the necessary I/O speed so as to guarantee a stable data I/O speed and realize an effective disc device operation management and data arrangement.

(57) 要約: 【課題】システムとして安定したデータ入出力速度を保証すると同時に、効率的なディスク装置の動作の管理とデータ配置を実現すること 【解決手段】ディスクアレイ制御装置１０は、データの入出力の対象となるアドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数のディスク装置にストライピングして入出力を行う入出力管理部５を備えている。このようにして、所要の入出力速度に応じてデータをストライピングするディスク装置の台数を決定し、安定したデータ入出力速度を保証すると同時に、効率的なディスク装置の動作の管理とデータ配置を実現する。

WO 2008/026497 A1

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告書

明 細 書

ディスクアレイ制御装置、ディスクアレイ制御方法、ディスクアレイ制御プログラム

技術分野

[0001] 本発明は、磁気ディスクなどの外部記憶装置を利用したディスクアレイ装置の制御装置、制御方法及びプログラムに係り、特に、効率的にデータのストライピングをすることができるディスクアレイ装置の制御装置等に関する。

背景技術

[0002] 関連するストレージ装置においては、ハードディスク装置等の記憶装置の単位時間当たりに入力または出力することができるデータの量(以下、「処理速度」または「速度」という)の差の影響による処理速度のばらつきを最小にすると共に、平均の入出力性能を保証するためにディスク装置の内周側と外周側を組み合わせたディスクアレイを構成するなどにより、性能安定化確保を行っていた。

例えば、特許文献1において、ディスク装置2台を一組として、内周側と外周側を組み合わせたアクセスを行い、平均して性能の良いディスク装置を構成する方式が記載されている。

[0003] また、関連するディスクアレイ装置では、ディスク装置毎の応答速度を安定化させるために、ディスク装置に対するアクセスサイズなどを制御して、常にディスク入出力の応答を安定化させるなどの対策を行っていた。

例えば、特許文献2には、ディスク装置の処理速度を基準にして、各ディスクに分配するデータ量を設定し、各ディスクからの応答時間を一定化する方法が記載されている。

[0004] 特許文献1:特開平10－55597号公報

特許文献2:特開平8－69359号公報

発明の開示

発明が解決しようとする課題

[0005] 関連する技術の問題点は、ハードディスク装置では、データ入出力速度が、内周側

と外周側で大きく変動し、内周側では外周側の約1／2の性能となるため、内周に記録したデータを利用する際のスループットの低下を考慮した設計が必要であった。特に、デジタルAVデータの記録再生では、内周側の性能をその装置の性能とした設計をせざるを得ず、システムのスループットを低く設定せざるを得なかった。

例えば、特許文献1の明細書で示されている通り、ディスクアレイの性能変動の影響を及ばなくするためには、最も性能の悪いディスクを基準にして、最悪性能を見積もるのが一般的であると記載されている。また、この内周側と外周側を組み合わせた2台1組を同時動作させ、性能が安定した入出力を実現するなどの対策を加える方法が記載されている。

[0006] 他の問題点は、ディスク装置の性能変動を回避するために、複数のディスク装置を使用したディスクアレイで構成する場合は、常に複数台のディスクを動作させる必要があった。

例えば、特許文献2で示されている通り、ディスク装置の性能の違いを最小に抑えるため、性能に応じたデータのストライピングサイズを調整し、ディスク装置の応答時間を安定化する方法が記載されている。しかし、この場合も、複数のディスク装置は、常に動作させる必要があった。

[0007] 本発明の目的は、データをディスアレイ装置を構成する磁気ディスク装置や光ディスク装置にストライピングして記録管理する場合に、効率的なディスク装置の動作の管理とデータ配置を実現することにある。

課題を解決するための手段

[0008] 本発明の第1のディスクアレイ制御装置は、複数のディスク装置の少なくとも1つのデータを入出力しようとする記憶領域のアドレスに応じてストライピングに利用するディスク装置の個数を変更して入出力を行う入出力管理部を備えている。

[0009] 上記第1のディスクアレイ制御装置によれば、データを入出力するアドレスに応じてデータをストライピングするディスク装置の台数を決定するので、効率的なディスク装置の動作の管理とデータ配置を実現することができる。すなわち、所要の転送速度や消費電力、静穏性等に応じて適切な数のディスク装置にデータをストライピングすることができる。

発明の効果

[0010] 本発明の、第1の効果は、磁気ディスク装置などの内周側などの性能が低下する領域で効果的にストライピングを行うことが出来るため、ディスク全面で所要性能を保証することができることである。

その理由は、所要性能より性能が低下する領域でのみ、ストライピング動作を行って性能を向上することを可能としているからである。

第2の効果は、必要以上にディスクを動作させることが無く、消費電力の低減を実現することができることにある。

その理由は、本発明ではファイルが要求する所要帯域に応じて領域割り当てを行い、また所要レートに応じてストライピング数を割り当てることになるため、所要レートが低い、あるいはリアルタイム処理の必要が無いデータに対しては、多数のディスクを稼動させない制御が可能となるからである。

第3の効果は、ディスク装置の特性に応じて動作を変更することにより、騒音を抑制化を実現することにある。

理由は、本発明ではアドレス位置によりストライピング数を変更可能としているため、光ディスクのようにアドレス位置により回転数が変わり騒音が発生する装置を使用する場合は、ストライピング数を減らすことにより複数ディスクに対するアクセスを同時に発生させない制御を可能としている。結果として、システム全体の騒音を抑制することが出来る。

発明を実施するための最良の形態

[0011] 本発明の第1の実施形態であるディスクアレイ制御装置10の構成と動作について説明する。

図1、図2にディスクアレイ制御装置10の構成図を示す。

ディスクアレイ制御装置10は、データの入出力を行うプログラム実行部4と、データの管理を行う入出力管理部5と、時間計測を行う時間計測部6と、領域管理を行う領域管理部7と、ディスク入出力制御部8と、2台のディスク装置9－0および9－1からなるディスクアレイ9により構成される。このうちプログラム実行部4は、記録再生を行うデータを生成及び再生する機能を持つが、映像を記録再生する場合は図1に記載され

るような映像音声圧縮伸張部1と、メモリ管理部2と、バッファ部3などの映像処理手段を搭載することが考えられる。

[0012] データの入出力を管理する入出力管理部5は、ディスク入出力制御部8を使用して複数のディスク装置9－0、9－1に、データを適当なサイズに分割して振り分ける動作を行う。この動作をストライピング動作と定義する。ストライピングするデータのサイズは、バイト単位でも構わないし、セクタなどの適当なブロックサイズ単位で実行しても構わない。またストライピング動作を行う際には、対象となる複数のディスク装置9－0等は、稼働中であり、かつ並列動作を行うことが可能である必要がある。

[0013] 本実施形態では、ディスク装置としては磁気ディスク装置を想定して説明しているが、光ディスク装置等のアドレス位置により入出力速度や回転数が変化するディスク装置を複数使用した場合にも同様に適用可能である。

また、本発明の実施形態は、パーソナルコンピュータや、ディジタルデータを利用した映像音声記録再生機器などでの利用を想定している。よって、各部は単一のCPU(Central Processing Unit)上で動作するソフトウェアのモジュールとして実装することも可能であるし、独立したハードウェアでも同様に実現可能である。また、シーケンシャルデータの生成に、映像音声圧縮伸張部1やバッファ部3などを利用して生成することを想定して説明しているが、これらは外部の機器で予めエンコードされたデータを、ネットワークや放送波、他の蓄積装置経由で取得し記録しても良い。

[0014] 図1～8を使用して、ディスクアレイ制御装置10の入出力動作について説明する。

まず、映像や音声などのシーケンシャルデータ及びその他のデータの記録動作について説明する。

外部から入力された映像音声は、映像音声圧縮伸張部1で、圧縮されたディジタルデータに変換され、メモリ管理部2が管理するバッファ部3に格納される。格納されたディジタルデータは、プログラム実行部4がメモリ管理部2を介して生成状況をモニタする。

プログラム実行部4は、前記の圧縮されたディジタルデータが、所定の量バッファ部3に蓄積された時点で、これらのデータをまとめてディスク装置9－0及び9－1に記録するように、入出力管理部5に指示を出すと共に、データを入出力管理部5に送る。

[0015] また、プログラム実行部4は、その他データ、例えばサムネイルデータやファイル管理用のメタデータのように、映像ストリームに直接関連のないデータをディスク装置9－0及び9－1に記録する場合も、同様に入出力管理部5に指示を出すと共に、データを入出力管理部5に送る。

データを論理的なひとまとまりのファイルとして管理する場合は、必要に応じてファイルの種別やファイルの大きさ、所要ビットレートなどのファイルの属性情報を入出力管理部5に通知する。これらの属性はファイルに対するファイル管理用のメタデータとしてディスク装置9－0、9－1に記録して管理される。

[0016] 入出力管理部5は、プログラム実行部4よりファイルの生成を指示された場合は、領域管理部7にディスク装置9上の記録領域の確保を要求し、記録すべき領域の位置、サイズを取得した後、該当する領域に前記のデータを記録するようにディスク入出力制御部8に指示を出す。また、入出力管理部5は、領域を管理するメタデータとして、前記の記録すべき領域の位置、サイズや、ファイルの属性、ファイルの大きさ、所要ビットレートなどのメタデータを生成して、ディスク入出力制御部8に書き込みを指示する。ファイル管理用のメタデータとしては、汎用のファイルシステム、ECMA(European Computer Manufacturers Association)-167やISO(International Organization For Standardization)9293などの汎用のファイルシステムの管理構造や、それを拡張した属性を使用することにより実現可能である。

[0017] また、入出力管理部5は、ディスク装置9－0及び9－1の入出力性能を計測する。具体的に説明すると、入出力管理部5は、ディスク入出力制御部8を制御することで、ディスク装置9－0及び9－1のアドレス範囲を適当な等間隔に区切り、各位置から一定量のデータのシーケンシャルの読み出しを実行して、その処理時間を時間計測部6により計測し、その処理時間を取得する。入出力管理部5は、入出力の際には、ディスク装置9－0及び9－1の制御装置内のバッファなどの容量も考慮して、それらのバッファより十分に大きな一定量のデータの読み出しを実行することにより、より正確な物理的な入出力速度を計測するのが良い。また、ディスク装置のインタフェースのプロトコルオーバーヘッドも考慮し、個々の入出力コマンドが扱うデータ量も、相対的に大きな値を選択するのが良い。

[0018] その結果を元に、入出力管理部5は、アドレス位置とディスク装置の入出力速度の関係を割り出す。このアドレス位置とディスク装置の入出力速度の関係は、予めディスク装置9－0または9－1の特定の位置に記録したり、ディスク装置のコントローラ内の不揮発メモリなどに記録しておき、ディスクアレイ装置を起動する際に、読み出して利用しても良い。あるいは、このアドレス位置とディスク装置の入出力速度の関係は、入出力管理部5内の不揮発性メモリなどに記録し利用しても良い。

ディスク入出力制御部8は入出力管理部5から指示されたアドレス及び長さの領域に対して、要求されたデータの書き込み又は読みだしコマンドをディスク装置9－0及び9－1に出すと共にデータの送受信を実行する。

[0019] 次に、映像や音声などのシーケンシャルデータ及びその他のデータの再生動作について説明する。

まず、プログラム実行部4は、再生したい所望のファイルの読み出しを入出力管理部5に指示する。

入出力管理部5は、プログラム実行部4よりファイルの読み出しを指示され場合、該当ファイルの領域を管理するメタデータ読み出しを入出力制御部8に指示する。ディスク入出力制御部8は前記と同様に、ディスク装置9－0及び9－1からのデータの読み出しを実行する。入出力管理部5は、入出力制御部8から受け取ったメタデータに基づき、ファイルを保持する領域の位置及びサイズを取得した後、再度この位置、サイズのデータの読み出しをディスク入出力制御部8に指示し、受け取ったデータをプログラム実行部4に送る。

[0020] プログラム実行部4は、受け取ったファイルが映像音声データでない場合は、プログラム実行部4に保持して、その他の処理に利用する。プログラム実行部4は、映像音声データの場合、メモリ管理部2を介してバッファ部3にストアする。このデータはメモリ管理部2が再度読み出し、映像音声圧縮伸張部1に送られ、伸張された後、映像音声として外部に出力される。

[0021] 次に、領域管理部7の動作について説明する。領域管理部7は領域割り当ての順を決めるエリアの割り当て、及びストライピング動作の条件を管理する。

図2を使用して領域管理部7の構成を説明する。領域管理部7は複数のディスク装

置を複数のエリアに分けてそれぞれのエリア毎にどのように領域割り当てを実施するかを判断する。

ストライピング管理部71は、エリア管理部73に指示を出して、各エリアのアドレス割り当てを管理する。ストライピング管理部71は、1台あるいは複数のディスク装置に対して、データの振り分けの実施の判断を行うが、これらはストライピング管理メタデータ蓄積部72で管理されるストライピング範囲及びディスク台数管理の設定に基づいて実施される。

エリア管理部73は、各エリアの範囲などを管理するが、各エリアでのディスク装置の性能値を同時に管理し、各データをエリア管理メタデータ蓄積部74に蓄積し保持する。

ストライピング管理メタデータ蓄積部72及びエリア管理メタデータ蓄積部74で保持する各データは、対象のディスク装置9－0及び9－1内の特定アドレスに記録しておき、それをロードして管理を行う。

[0022] ストライピング管理メタデータは、図3のようなリストで、ストライピング領域毎にアクセス順を定めるIndex、ストライピングデータを配置するエリア（部分領域）を特定するエリアIDや、その数を与えるストライピング数などを含んでいる。この図では、エリアIDは、「ディスク装置を特定する情報／そのディスク装置でエリアを特定する情報」という形式で表わしている。例えば、「＃0／0」は、ディスク装置＃0（9－1）の番号0のエリアを示している。

[0023] 一方、エリア管理メタデータは、図4のようなリストにより、各エリアIDに対するディスク装置を特定するための番号、そのディスク装置内でのアドレス範囲、アドレス範囲での入出力性能の最高値、最低値などを保持する。

ディスクアドレス範囲は、そのエリアの開始アドレス－終了アドレスの形式で記録されている。本実施形態では、ディスク装置としてハードディスク装置を想定しているので、アドレスは外周側から内周側に向かって昇順になるように付されている。なお、ディスクアドレス範囲の欄で「0x」はそれに続く数値が16進数で表示されていることを示す記号である。

入出力最高性能、入出力最低性能は、そのエリアにおける最高の入出力速度と最

低の入出力速度をbps等の単位を用いて記録している。本実施形態の場合、「A」はディスク9－0のアドレス0x0での値、「E」はディスク9－0のアドレス0xfffffでの値、「C」はディスク9－0のアドレス0x100000での値、「G」はディスク9－0のアドレス0x17ffffでの値であり、A＞E＞C＞Gの関係にある。ディスク9－1についても同様である。

使用未使用フラグは、そのエリアがストライピング領域として使用されているか否かを示す情報で、「0」は使用されていないことを、「1」は使用されていることを示す。

[0024] このデータの例に基づき2台のディスク装置9－0、9－1にレイアウトした場合の、各エリアとの対応関係を図5に示す。なお、Index=2の2台にストライピングする場合は、ストライピングはディスク装置のセクタやそれより大きな単位を基準にしてストライピングしてよい。ここではストライピングを行うディスク台数であるストライピング数は2台であるが、3台以上でも同様に適用可能である。

[0025] ここで、ディスク装置をエリアに分割する方法について説明する。一般に磁気ディスク装置の入出力速度とアクセス位置の関係は、おおよそ図6のような関係にある。よって、所望の入出力速度（閾値）以上を常に保証する必要がある場合、ディスク装置の内周側にデータを記録した場合、性能不足となることがある。ここでは、点線のような所望入出力速度の保証をする場合、所望速度に対応し境界値となるアドレスXよりも大きい（すなわち内周側になる）アドレス範囲では、性能保証ができないことが判る。よって、このアドレス範囲を基準にして、性能保証可能なエリアと、保証できないエリアを分割し、前記のようなストライピング範囲を設定する。

なお、入出力性能に依存しないデータを保持する場合は、このような速度に基づいたエリアの分割を適用せずに、所要の容量のエリアに任意のストライピング数を割り当てればよい。また、光ディスクのようにアドレス位置により、回転数などが大きく変わり騒音などが発生するディスク装置を利用する場合は、その特性に応じて、予め決まったアドレス範囲ではストライピング数を制限できるようにエリアを分割して、ストライピング範囲を設定すればよい。

[0026] 以上のようにエリアを割り当てることにより、ディスクアレイ9全体としての性能は、図7のような分布となり、所望の入出力速度を常に保証できるようになる。この例では、

ディスク装置9－0とディスク装置9－1が同一の性能を持つものとしている。図7のHの値は図6のCの約2倍、Iは図6のGの約2倍で所望の入出力速度よりは大きな値となっている。

[0027] ファイル管理には各種のファイルシステムで利用されている方式がそのまま適用可能であるが、ファイルシステムが割り付けた仮想的なアドレス及び範囲は、前記の領域管理部のストライピング管理部71に渡され、実際にデータを配置するディスク装置及び物理的なアドレス及び範囲に変換され、入出力管理部5に返される。このデータに基づいて、入出力管理部5からディスク入出力制御部8に入出力指示が出されることになる。

[0028] 次に、入出力管理部5がデータ入出力を行う際の初期化処理の手順を、図8のフローチャートに基づいて説明する。

まず、入出力管理部5は、指示して、ディスク装置9－0及び9－1から、エリア管理メタデータを読み込む。このメタデータは、複数ディスクを一つの論理的ディスクとして扱うための、アドレスの対応関係を指定するものであり、各エリアのディスク番号、ディスク内のアドレス範囲などを規定するものである。入出力管理部5は、同様にストライピング管理メタデータを読み込む。このデータは、各エリアを組み合わせて、論理的なアドレス空間を構成する方法を規定するものである。

[0029] 入出力管理部5は、このエリア管理データ及びストライピング管理メタデータが書き込まれたフォーマット済みの状態か否かを、ディスク装置9－0及び9－1のデータを読み出すことによりチェックする（S11）。

もしフォーマット済みで無い場合は、入出力管理部5は、前述の方法で、ディスク装置9－0及び9－1の入出力速度を計測し（S12）、その結果に基づいて、ディスクアレイ9のエリア分けを決定し、エリア管理メタデータと、ストライピング管理メタデータデータを生成して、共にディスクアレイ9に記録する（S13）。フォーマット済みの場合は、特に何もしない。

[0030] 次に、入出力管理部5は、ディスク装置9－0及び9－1上のエリア管理用メタデータ、ストライピング管理メタデータを読み出し、エリア管理用メタデータをエリア管理メタデータ蓄積部74に、ストライピング管理メタデータを、ストライピング管理メタデータ蓄

積部72に設定する。またその他のファイルシステムが利用するメタデータは、入出力管理部5で保持する(S14)。

以上のような処理シーケンスにより、領域管理部7が管理するアドレス変換機構が初期化処理時に確定し、以降は、入出力管理部5がデータ入出力を実行する。

[0031] 次に、ディスクアレイ制御装置10の効果について説明する。

第1の効果は、磁気ディスク装置などの内周側などの性能が低下する領域で効果的にストライピングを行うことが出来るため、ディスク全面で所要性能を保証することができる。

理由は、内周側の性能の低下を考慮して、予め性能を低く見積もるか、複数のディスクを常に動作させるなどの対策で、性能を保証することを行っていたのに対して、本発明ではそのような考慮無しに、所要性能より性能が低下する領域でのみ、ストライピング動作を行って性能を向上することを可能としているからである。

第2の効果は、必要以上のディスク台数を動作させることが無く、消費電力低減の効果が期待される。

理由は、本発明の実施形態ではファイルが要求する所要帯域に応じて領域割り当てを行い、また所要レートに応じてストライピング数を割り当てることを可能としている。そのため、所要レートが低い場合は、ストライピング数を減らして、稼動するディスク台数を減らす制御が可能である。あるいは所要レートが特に無く非リアルタイム処理でよいデータなどを記録するアドレス範囲では、ストライピングを行わず1台だけディスクを稼動させて、他のディスク装置を停止する制御が可能となる。結果として、システム全体の消費電力の低減や静穏化を実現することが出来る。

第3の効果は、ディスク装置の特性に応じて動作を変更することにより、騒音の抑制効果が期待される。

理由は、光ディスク装置などではディスクのアドレス位置に応じてディスクの回転数を変えるなどして性能向上を図っているが、そのためアクセス位置に依存して騒音が発生することになる。ストライピングを行うことにより性能向上を図った場合、ディスク台数分、更に騒音が発生することとなり問題となる。これに対して、本発明の実施形態ではアドレス位置によりストライピング数を変更可能としているため、騒音が問題となる

アドレス範囲を使用する場合は、ストライピング数を減らすことにより、複数ディスクに対するアクセスを同時に発生させない制御を可能としている。結果として、システムの騒音の抑制を実現することが出来る。

[0032] 次に、本発明の第2の実施形態であるディスクアレイ制御装置11の構成と動作について説明する。

図9にディスクアレイ制御装置11の構成図を示す。

ディスクアレイ制御装置11は、データの入出力を指示するプログラム実行部4と、データの入出力管理を行う入出力管理部5と、時間計測を行う時間計測部6と、領域管理を行う領域管理部7と、ディスク入出力制御部8と、ディスク装置9－0及び、9－1、9－2、9－3から成るディスクアレイ9から構成される。このうちプログラム実行部4は、記録再生を行うデータを生成及び再生する手段を持つが、映像を記録再生する場合は図1に記載されるような映像音声圧縮伸張部1と、メモリ管理部2と、バッファ部3などの映像処理手段を搭載することが考えられる。

[0033] 本実施形態では、ディスク装置としては磁気ディスク装置4台を想定して説明しているが、光ディスク装置や5台以上のディスク装置を使用した場合にも同様に適用可能である。また、ディスク装置を4個のエリアに分けて構成した場合の例を説明しているが、5個以上のエリアに分けて管理することも、同様に可能である。

[0034] 図9～12使用して、ディスクアレイ制御装置11の入出力動作について説明する。なお、基本的な動作は、ディスクアレイ制御装置10と同様であるため、動作や構成が異なる点に絞り説明する。

ディスクアレイ制御装置11の構成について説明する。図9のように、本実施形態では、ディスクアレイ9はディスク装置9－0及び、9－1、9－2、9－3のように3台以上の複数のディスク装置から構成されている。各ディスク装置は、ディスク入出力制御部8からアクセスが可能であり、入出力管理部5からの指示に基づいて入出力動作が実施される。各ディスク装置のエリアの管理は、第1の実施の形態同様に領域管理部7で実施される。

[0035] 領域管理部7でのエリアの管理及びストライピングの管理は、基本的に第1の実施形態と同様に適用されるが、入出力管理部5からの指示に基づくストライピングの管

理において、必要に応じてストライピング管理メタデータを更新する点が異なる。よって、以下、入出力管理部5で行われる入出力動作のアルゴリズムを、図10を使用して説明する。

[0036] データの入出力動作は、まず入出力管理部5内で、エリア管理用メタデータ及びストライピング管理用メタデータが読み込まれているかをチェックする(S21)。エリア管理用メタデータは例えば、図4のような管理情報を保持しているデータ構造であり、ストライピング管理用メタデータは、例えば図3のような管理情報を保持しているデータ構造である。もしこれらのメタデータが読み込まれていない場合は、このメタデータの読み出しを実行する(S22)。

[0037] 次に、入出力動作が書き込み動作かを判定して(S23)、読み出し動作の場合は、データエリアの読み出しを、領域管理部7のアドレス変換に基づいて実行する(S24)。

[0038] 入出力動作が書き込みの場合は、入出力管理部5は、所要の入出力帯域をファイル属性より入手し、その所要帯域が現在のディスク装置1台に対するアクセスで不足するかを判定して(S25)、もし領域が不足する場合は入出力管理部5は、領域管理部7に所要帯域量を要求し(S26)、領域管理部7では所要帯域を保証できるエリア及びストライピング数を、エリア管理部に問い合わせ未使用エリアから選択して割り当てる(S27)。ストライピング数は、所要帯域を保証できる最小のストライピング数を選択するように制御することが可能である。また、この割り当てられたストライピング数だけディスク装置を動作させ、それ以外のディスク装置を停止しておくように、入出力管理部5から指示を出すことも可能である。

[0039] 入出力管理部5は、その結果に基づいて、ストライピング管理用メタデータを更新する(S28)。入出力管理部5は、これらの確保されたエリア又はストライピングエリアに対して書き込み動作を実行する(S29)。さらに、入出力管理部5は、エリア管理用メタデータ及びストライピング管理用メタデータを更新して(S30)処理を終了する。

[0040] これらの処理シーケンスに汎用ファイルシステムを実装する場合は、領域管理部7で管理している帯域保証しているストライピングエリア毎の領域管理、及び所要レートなどの属性を管理できるファイルシステムを構築することが必要である。ただし、領域

割り当てアルゴリズムをレートに応じて多段にすることで、所要帯域属性の異なった複数の領域の割り当てアルゴリズムが実装できるため、所要帯域の判定(S25)と割り当て動作(S29)を汎用ファイルシステムの管理構造に実装することは容易である。

[0041] また、入出力管理部5において、ファイル入出力に対して所要帯域が指定されていない場合は、ストライピングエリアの割り当ては行わず、ストライピングされていないエリアから領域割り当てを実行する。また、所要帯域が既に割り当たられたストライピングエリアで十分な場合は、そのエリアから領域割り当てを実行する。

[0042] 次に、第2の実施の形態にもとづき、複数のストライピングエリアを割り当てた例を、図11及び図12で説明する。図11では、ストライピングエリアとして、＃1、＃2、＃3それぞれが、ストライピングの数が3、2、4の場合である。また、エリアを昇順に割り当てることにより、実際の割り当てられた領域の帯域保証レートは図12に示すような領域により可変な帯域割り当てが成されることになる。

[0043] ディスクアレイ制御装置11によれば、ディスクアレイ制御装置10と同様の効果に加え、ディスクアレイ9に入力するファイルの特性に応じて、適切な帯域のストライピングエリアを作成し、割り当てることができる。

[0044] 次に、本発明の他の実施形態について説明する。

[0045] ディスクアレイ制御装置において、ディスク装置はアドレスにより入出力速度が異なるものとし、入出力管理部は、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数の前記ディスク装置にストライピングして前記入出力を行うようにしてもよい。

このようにすれば、所要の入出力速度に応じてデータをストライピングするディスク装置の台数を決定するので、安定したデータ入出力速度を保証することができる。

上記第1のディスクアレイ制御装置において、ディスク装置はディスクの内周側ほど処理速度が小さいものであり、アドレスは外周側から内周側に向かって昇順に付されており、入出力管理部は、入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて台数を決定するようにしてもよい。

このようにすれば、CAV制御を行う磁気ディスク装置や光ディスク装置でディスクアレイを構成した場合に、アドレス位置を基準としてストライピングに使用するディスク装

置の台数を決定することができる。

[0046] 上記第1のディスクアレイ制御装置において、アドレスと入出力速度とを対応付けて予め記憶したエリア管理メタデータを記憶するエリア管理メタデータ記憶部を備え、入出力管理部は、エリア管理メタデータを参照して入出力速度を取得するようにしてもよい。エリア管理メタデータ記憶部は、入出力管理部に設けてもよいしディスク装置に設けてもよい。

このようにすれば、データの入出力を行う都度入出力の対象となるアドレスにおける入出力速度を算出する必要がなくなり、処理を高速化することができる。

[0047] 上記第1のディスクアレイ制御装置において、ディスク装置の複数のアドレスにおいて入出力速度を計測し、エリア管理メタデータを生成する時間計測部を備えるようにしてもよい。

このようにすれば、実際に計測された各アドレス位置における入出力速度に基づいて、より正確にディスク装置の台数を決定することができる。

[0048] 上記第1のディスクアレイ制御装置において、入出力管理部は、ストライピングの対象となっていないディスク装置のスピンドルを停止させるようにしてもよい。

このようにすれば、ディスクアレイの消費電力の削減と静穏化を実現することができる。

[0049] 各ディスク装置の記憶領域を処理速度に基づいて論理的に複数の部分領域に分割して管理する領域管理部と、所定の閾値以上の入出力速度が得られるように、1台のディスク装置の一つの部分領域を選択してその部分領域に、または、複数のディスク装置からそれぞれ一つの部分領域を選択してそれらの部分領域にストライピングしてデータの入出力を行う入出力管理部を備えを備えるようにしてもよいものである。

[0050] これにより、所要の入出力帯域に応じて必要な数の部分領域を選択してデータの入出力を行うことできる。そのため、安定したデータ入出力速度を保証すると同時に、効率的なディスク装置の動作の管理とデータ配置を実現することができる。

[0051] 上記第2のディスクアレイ制御装置において、各部分領域における処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けて記憶するエリア管理メタデータ記憶部と、複数のディスク装置からそれぞれ一つの部分領域を選

択したストライピング領域を特定するストライピング領域特定情報とそのストライピング領域に含まれる各部分領域のエリア特定情報とを対応付けて記憶するストライピング管理メタデータ記憶部と、を備え、入出力管理部は、エリア管理メタデータ記憶部とストライピング管理メタデータ記憶部を参照して閾値以上の入出力速度が得られるストライピング領域を選択するようにしてもよい。

このようにすれば、ストライピング領域が予め作成されると共に各ストライピング領域の入出力速度が記憶されているので、入出力要求のあったときに迅速に適切なストライピング領域を選択することができる。

[0052] 上記第2のディスクアレイ制御装置において、入出力管理部は、ディスク装置に対して書き込みを行う際に、閾値の入力を受け、既存のストライピング領域に入力された閾値を上回る入出力速度を得られるものがない場合には、入力された閾値を上回る入力速度が得られるように既存のストライピング領域に組み込まれていない部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して書き込みを行い、新たなストライピング領域に関する情報をストライピング管理メタデータ記憶部に書き込むようにしてもよい。

このようにすれば、ディスク装置に入力するデータの特性に応じて適切なストライピング領域を選択することができるとともに、既存のストライピング領域では入出力速度が不足する場合にも対処することができる。

[0053] 本発明の実施形態のディスクアレイ制御方法は、複数のディスク装置の少なくとも1つのデータを入出力しようとする記憶領域のアドレスに応じてストライピングに利用するディスク装置の個数を変更して入出力を行うようにしてもよいものである。

[0054] これにより、データを入出力するアドレスに応じてデータをストライピングするディスク装置の台数を決定するので、効率的なディスク装置の動作の管理とデータ配置を実現することができる。すなわち、所要の転送速度や消費電力、静穏性等に応じて適切な数のディスク装置にデータをストライピングすることができる。

[0055] 上記第1のディスクアレイ制御方法において、ディスク装置はアドレスにより入出力速度が異なるものとし、入出力管理部工程では、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数の前

記ディスク装置にストライピングして前記入出力を行うようにしてもよい。

このようにすれば、所要の入出力速度に応じてデータをストライピングするディスク装置の台数を決定するので、安定したデータ入出力速度を保証することができる。

上記第1のディスクアレイ制御方法において、ディスク装置はディスクの内周側ほど処理速度が小さいものであり、アドレスは外周側から内周側に向かって昇順に付されており、入出力管理工程では、入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて台数を決定するようにしてもよい。

このようにすれば、CAV制御を行う磁気ディスク装置や光ディスク装置でディスクアレイを構成した場合に、アドレス位置を基準としてストライピングに使用するディスク装置の台数を決定することができる。

上記第1のディスクアレイ制御方法において、入出力管理工程では、アドレスと入出力速度とを対応付けて予め記憶したエリア管理メタデータを参照して入出力速度を取得するようにしてもよい。エリア管理メタデータは、入出力管理部に記憶してもよいし、ディスク装置に記憶してもよい。

このようにすれば、データの入出力を行う都度入出力の対象となるアドレスにおける入出力速度を算出する必要がなくなり、処理を高速化することができる。

[0056] 上記第1のディスクアレイ制御方法において、ディスク装置の複数のアドレスにおいて入出力速度を計測し、エリア管理メタデータを生成する時間計測工程を備えるようにしてもよい。

このようにすれば、実際に計測された各アドレス位置における入出力速度に基づいて、より正確にディスク装置の台数を決定することができる。

[0057] 上記第1のディスクアレイ制御方法において、ストライピングの対象となっていないディスク装置のスピンドルを停止させるようにしてもよい。

このようにすれば、ディスクアレイの消費電力の削減と静穏化を実現することができる。

[0058] 本発明の実施形態のディスクアレイ制御方法は、各ディスク装置の記憶領域を処理速度に基づいて論理的に複数の部分領域に分割して管理する領域管理工程と、所定の閾値以上の入出力速度が得られるように、1台のディスク装置の一つの部分

領域を選択してその部分領域に、または、複数のディスク装置からそれぞれ一つの部分領域を選択してそれらの部分領域にストライピングしてデータの入出力を行う入出力管理工程を備えていてもよい。

[0059] 上記第2のディスクアレイ制御方法によれば、所要の入出力帯域に応じて必要な数の部分領域を選択してデータの入出力を行うことできる。そのため、安定したデータ入出力速度を保証すると同時に、効率的なディスク装置の動作の管理とデータ配置を実現することができる。

[0060] 上記第2のディスクアレイ制御方法において、入出力管理工程では、各部分領域における処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けたエリア管理メタデータと、複数のディスク装置からそれぞれ一つの部分領域を選択したストライピング領域を特定するストライピング領域特定情報とそのストライピング領域に含まれる各部分領域のエリア特定情報とを対応付けたストライピング管理メタデータ、を参照して閾値以上の入出力速度が得られるストライピング領域を選択するようにしてもよい。

このようにすれば、ストライピング領域が予め作成されると共に各ストライピング領域の入出力速度が記憶されているので、入出力要求のあったときに迅速に適切なストライピング領域を選択することができる。

[0061] 入出力管理工程では、ディスク装置に対して書き込みを行う際に、閾値の入力を受け、既存のストライピング領域に入力された閾値を上回る入出力速度を得られるものがない場合には、入力された閾値を上回る入力速度が得られるように既存のストライピング領域に組み込まれていない部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して書き込みを行い、新たなストライピング領域に関する情報をストライピング管理メタデータ記憶部に書き込むようにしてもよい。

このようにすれば、ディスク装置に入力するデータの特性に応じて適切なストライピング領域を選択することができるとともに、既存のストライピング領域では入出力速度が不足する場合にも対処することができる。

[0062] ディスク装置はアドレスにより入出力速度が異なるものとし、入出力管理処理では、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値

を上回るように選択した台数の前記ディスク装置にストライピングして前記入出力を行うようにしてもよい。

このようにすれば、所要の入出力速度に応じてデータをストライピングするディスク装置の台数を決定するので、安定したデータ入出力速度を保証することができる。

[0063] 上記第1のディスクアレイ制御プログラムにおいて、ディスク装置はディスクの内周側ほど処理速度が小さいものであり、アドレスは外周側から内周側に向かって昇順に付されており、入出力管理処理では、入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて台数を決定するようにしてもよい。

このようにすれば、CAV制御を行う磁気ディスク装置や光ディスク装置でディスクアレイを構成した場合に、アドレス位置を基準としてストライピングに使用するディスク装置の台数を決定することができる。

[0064] 上記第1のディスクアレイ制御プログラムにおいて、ディスク装置の複数のアドレスにおいて入出力速度を計測し、エリア管理メタデータを生成する時間計測処理をコンピュータに実行させるようにしてもよい。エリア管理メタデータは、入出力管理部に記憶してもよいし、ディスク装置に記憶してもよい。

このようにすれば、データの入出力を行う都度入出力の対象となるアドレスにおける入出力速度を算出する必要がなくなり、処理を高速化することができる。

[0065] ディスク装置の複数のアドレスにおいて入出力速度を計測し、エリア管理メタデータを生成する時間計測処理をコンピュータに実行させるようにしてもよい。

このようにすれば、実際に計測された各アドレス位置における入出力速度に基づいて、より正確にディスク装置の台数を決定することができる。

[0066] 上記第1のディスクアレイ制御プログラムにおいて、入出力管理処理では、ストライピングの対象となっていないディスク装置のスピンドルを停止させるようにしてもよい。

このようにすれば、ディスクアレイの消費電力の削減と静穏化を実現することができる。

[0067] 入出力管理処理では、各部分領域における処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けたエリア管理メタデータと、複数のディスク装置からそれぞれ一つの部分領域を選択したストライピング領域を特定するスト

ライピング領域特定情報とそのストライピング領域に含まれる各部分領域のエリア特定情報とを対応付けたストライピング管理メタデータ、を参照して閾値以上の入出力速度が得られるストライピング領域を選択するようにしてもよい。

このようにすれば、ストライピング領域が予め作成されると共に各ストライピング領域の入出力速度が記憶されているので、入出力要求のあったときに迅速に適切なストライピング領域を選択することができる。

[0068] 上記第2のディスクアレイ制御プログラムにおいて、入出力管理処理では、ディスク装置に対して書き込みを行う際に、閾値の入力を受け、既存のストライピング領域に入力された閾値を上回る入出力速度を得られるものがない場合には、入力された閾値を上回る入力速度が得られるように既存のストライピング領域に組み込まれていない部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して書き込みを行い、新たなストライピング領域に関する情報をストライピング管理メタデータ記憶部に書き込むようにしてもよい。

このようにすれば、ディスク装置に入力するデータの特性に応じて適切なストライピング領域を選択することができるとともに、既存のストライピング領域では入出力速度が不足する場合にも対処することができる。

[0069] 以上の実施形態は例示したモノであって、本発明は図示した実施形態の範囲に限定されるものではなく、請求の範囲内において種々変更できるものである。

[0070] 以上、実施形態(及び実施例)を参照して本願発明を説明したが、本願発明は上記実施形態(及び実施例)に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

[0071] この出願は2006年8月28日に出願された日本出願特願2006－230914を基礎とする優先権を主張し、その開示の全てをここに取り込む。

図面の簡単な説明

[0072] [図1]本発明の第1の実施の形態であるディスクアレイ制御装置の構成図である。

[図2]図1の、領域管理部の詳細な構成図である。

[図3]ストライピング管理メタデータの例である。

[図4]エリア管理メタデータの例である。

[図5]磁気ディスク装置上のエリア配置の説明図である。

[図6]磁気ディスク装置の典型的な読み出し速度特性の説明図である。

[図7]エリアレイアウトに基づいてストライピング制御した場合の入出力特性の説明図である。

[図8]図1の入出力管理部における、初期化処理のフローチャートである。

[図9]本発明の第2の実施の形態であるディスクアレイ制御装置の構成図である。

[図10]図9の入出力管理部における、ファイルの入出力処理のフローチャートである。

[図11]磁気ディスク装置のストライピング制御の適用例の説明図である。

[図12]磁気ディスク装置のストライピング制御の適用例の入出力特性の説明図である。

符号の説明

[0073] 1 映像音声圧縮伸張部

2 メモリ管理部

3 バッファ部

4 プログラム実行部

5 入出力管理部

6 時間計測部

7 領域管理部

8 ディスク入出力制御部

9 ディスクアレイ

9－0、9－1、9－2、9－3 ディスク装置

10、11 ディスクアレイ制御装置

71 ストライピング管理部

72 ストライピング管理部メタデータ蓄積部

73 エリア管理部

74 エリア管理部メタデータ蓄積部

請求の範囲

[1] 複数のディスク装置からなるディスクアレイを制御する装置において、

前記複数のディスク装置の少なくとも1つの、データを入出力しようとする記憶領域のアドレスに応じてストライピングに利用する前記ディスク装置の個数を変更して前記入出力を行う入出力管理部を備えたことを特徴とする、ディスクアレイ制御装置。

[2] 前記ディスク装置は前記アドレスにより入出力速度が異なるものであり、

前記入出力管理部は、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数の前記ディスク装置にストライピングして前記入出力を行うことを特徴とした請求項1に記載のディスクアレイ制御装置。

[3] 前記ディスク装置はディスクの内周側ほど前記処理速度が小さいものであり、前記アドレスは外周側から内周側に向かって昇順に付されており、

前記入出力管理部は、前記入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて前記台数を決定することを特徴とした請求項2に記載のディスクアレイ制御装置。

[4] 前記アドレスと前記入出力速度とを対応付けて予め記憶したエリア管理メタデータを記憶するエリア管理メタデータ記憶部を備え、

前記入出力管理部は、前記エリア管理メタデータを参照して前記入出力速度を取得することを特徴とした請求項2に記載のディスクアレイ制御装置。

[5] 前記エリア管理メタデータ記憶部は、前記入出力管理部に設けられていることを特徴とした請求項4に記載のディスクアレイ制御装置。

[6] 前記エリア管理メタデータ記憶部は、前記ディスク装置に設けられていることを特徴とした請求項4に記載のディスクアレイ制御装置。

[7] 前記ディスク装置の複数のアドレスにおいて前記入出力速度を計測し、前記エリア管理メタデータを生成する時間計測部を備えたことを特徴とした請求項4ないし請求項6のいずれか一つに記載のディスクアレイ制御装置。

[8] 前記入出力管理部は、ストライピングの対象となっていない前記ディスク装置のスピンドルを停止させることを特徴とする、請求項1ないし請求項7のいずれか一つに記

載のディスクアレイ制御装置。

[9] アドレスにより入出力速度が異なる複数のディスク装置からなるディスクアレイを制御する装置において、

前記各ディスク装置の記憶領域を前記処理速度に基づいて論理的に複数の部分領域に分割して管理する領域管理部と、

所定の閾値以上の入出力速度が得られるように、1台の前記ディスク装置の一つの前記部分領域を選択してその部分領域に、または、複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択してそれらの部分領域にストライピングしてデータの入出力を行う入出力管理部を備えたことを特徴とするディスクアレイ制御装置。

[10] 前記各部分領域における前記処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けて記憶するエリア管理メタデータ記憶部と、

複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択したストライピング領域を特定するストライピング領域特定情報とそのストライピング領域に含まれる各部分領域の前記エリア特定情報とを対応付けて記憶するストライピング管理メタデータ記憶部と、を備え、

前記入出力管理部は、エリア管理メタデータ記憶部とストライピング管理メタデータ記憶部を参照して前記閾値以上の入出力速度が得られる前記ストライピング領域を選択することを特徴とした請求項9に記載のディスクアレイ制御装置。

[11] 前記入出力管理部は、前記ディスク装置に対して書き込みを行う際に、前記閾値の入力を受け、

既存の前記ストライピング領域に前記入力された閾値を上回る前記入出力速度を得られるものがない場合には、前記入力された閾値を上回る前記入力速度が得られるように既存の前記ストライピング領域に組み込まれていない前記部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して前記書き込みを行い、前記新たなストライピング領域に関する情報を前記ストライピング管理メタデータ記憶部に書き込むことを特徴とした請求項9に記載のディスクアレイ制御装置。

[12] 複数のディスク装置からなるディスクアレイを制御する方法において、

前記複数のディスク装置の少なくとも1つの、データを入出力しようとする記憶領域

のアドレスに応じてストライピングに利用する前記ディスク装置の個数を変更して前記入出力を行う入出力管理工程を備えたことを特徴とする、ディスクアレイ制御方法。

[13] 前記ディスク装置は前記アドレスにより入出力速度が異なるものであり、

前記入出力管理工程では、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数の前記ディスク装置にストライピングして前記入出力を行うことを特徴とした請求項12に記載のディスクアレイ制御方法。

[14] 前記ディスク装置はディスクの内周側ほど前記処理速度が小さいものであり、前記アドレスは外周側から内周側に向かって昇順に付されており、

前記入出力管理工程では、前記入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて前記台数を決定することを特徴とした請求項13に記載のディスクアレイ制御方法。

[15] 前記入出力管理工程では、前記アドレスと前記入出力速度とを対応付けて予め記憶したエリア管理メタデータを参照して前記入出力速度を取得することを特徴とした請求項13に記載のディスクアレイ制御方法。

[16] 前記エリア管理メタデータは、前記入出力管理工程を実行する装置が備える記憶装置に記憶されていることを特徴とした請求項15に記載のディスクアレイ制御方法。

[17] 前記エリア管理メタデータは、前記ディスク装置に記憶されていることを特徴とした請求項15に記載のディスクアレイ制御方法。

[18] 前記ディスク装置の複数のアドレスにおいて前記入出力速度を計測し、前記エリア管理メタデータを生成する時間計測工程を備えたことを特徴とした請求項15ないし請求項17のいずれか一つに記載のディスクアレイ制御方法。

[19] ストライピングの対象となっていない前記ディスク装置のスピンドルを停止させることを特徴とする、請求項12ないし請求項18のいずれか一つに記載のディスクアレイ制御方法。

[20] アドレスにより入出力速度が異なる複数のディスク装置からなるディスクアレイを制御する方法において、

前記各ディスク装置の記憶領域を前記処理速度に基づいて論理的に複数の部分

領域に分割して管理する領域管理工程と、

所定の閾値以上の入出力速度が得られるように、1台の前記ディスク装置の一つの前記部分領域を選択してその部分領域に、または、複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択してそれらの部分領域にストライピングしてデータの入出力を行う入出力管理工程を備えたことを特徴としたディスクアレイ制御方法。

[21] 前記入出力管理工程では、

前記各部分領域における前記処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けたエリア管理メタデータと、

複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択したストライピング領域を特定するストライピング領域特定情報とそのストライピング領域に含まれる各部分領域の前記エリア特定情報とを対応付けたストライピング管理メタデータ、

を参照して前記閾値以上の入出力速度が得られる前記ストライピング領域を選択することを特徴とした請求項20に記載のディスクアレイ制御方法。

[22] 前記入出力管理工程では、前記ディスク装置に対して書き込みを行う際に、前記閾値の入力を受け、

既存の前記ストライピング領域に前記入力された閾値を上回る前記入出力速度を得られるものがない場合には、前記入力された閾値を上回る前記入力速度が得られるように既存の前記ストライピング領域に組み込まれていない前記部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して前記書き込みを行い、前記新たなストライピング領域に関する情報を前記ストライピング管理メタデータ記憶部に書き込むことを特徴とした請求項20に記載のディスクアレイ制御方法。

[23] 複数のディスク装置からなるディスクアレイを制御するプログラムにおいて、

コンピュータに、

前記複数のディスク装置の少なくとも1つの、データを入出力しようとする記憶領域のアドレスに応じてストライピングに利用する前記ディスク装置の個数を変更して前記入出力を行う入出力管理処理を実行させることを特徴とする、ディスクアレイ制御プログラム。

[24] 前記ディスク装置は前記アドレスにより入出力速度が異なるものであり、

前記入出力管理処理では、データの入出力の対象となる前記アドレスにおける入出力速度の合計が所定の閾値を上回るように選択した台数の前記ディスク装置にストライピングして前記入出力を行うことを特徴とした請求項23に記載のディスクアレイ制御プログラム。

[25] 前記ディスク装置はディスクの内周側ほど前記処理速度が小さいものであり、前記アドレスは外周側から内周側に向かって昇順に付されており、

前記入出力管理処理では、前記入出力の対象となるアドレスと所定の境界値との間の大小関係に基づいて前記台数を決定することを特徴とした請求項24に記載のディスクアレイ制御プログラム。

[26] 前記入出力管理処理では、前記アドレスと前記入出力速度とを対応付けて予め記憶したエリア管理メタデータを参照して前記入出力速度を取得することを特徴とした請求項24に記載のディスクアレイ制御プログラム。

[27] 前記エリア管理メタデータは、前記入出力管理処理を実行する装置が備える記憶装置に記憶されていることを特徴とした請求項26に記載のディスクアレイ制御プログラム。

[28] 前記エリア管理メタデータは、前記ディスク装置に記憶されていることを特徴とした請求項26に記載のディスクアレイ制御プログラム。

[29] 前記ディスク装置の複数のアドレスにおいて前記入出力速度を計測し、前記エリア管理メタデータを生成する時間計測処理を前記コンピュータに実行させることを特徴とした請求項26ないし請求項28のいずれか一つに記載のディスクアレイ制御プログラム。

[30] 前記入出力管理処理では、ストライピングの対象となっていない前記ディスク装置のスピンドルを停止させることを特徴とする、請求項23ないし請求項29のいずれか一つに記載のディスクアレイ制御プログラム。

[31] アドレスにより入出力速度が異なる複数のディスク装置からなるディスクアレイを制御するプログラムにおいて、

コンピュータに、

前記各ディスク装置の記憶領域を前記処理速度に基づいて論理的に複数の部分

領域に分割して管理する領域管理処理と、

所定の閾値以上の入出力速度が得られるように、1台の前記ディスク装置の一つの前記部分領域を選択してその部分領域に、または、複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択してそれらの部分領域にストライピングしてデータの入出力を行う入出力管理処理を実行させることを特徴としたディスクアレイ制御プログラム。

[32] 前記入出力管理処理では、

前記各部分領域における前記処理速度の最高値と最低値をその部分領域を特定するエリア特定情報と対応付けたエリア管理メタデータと、

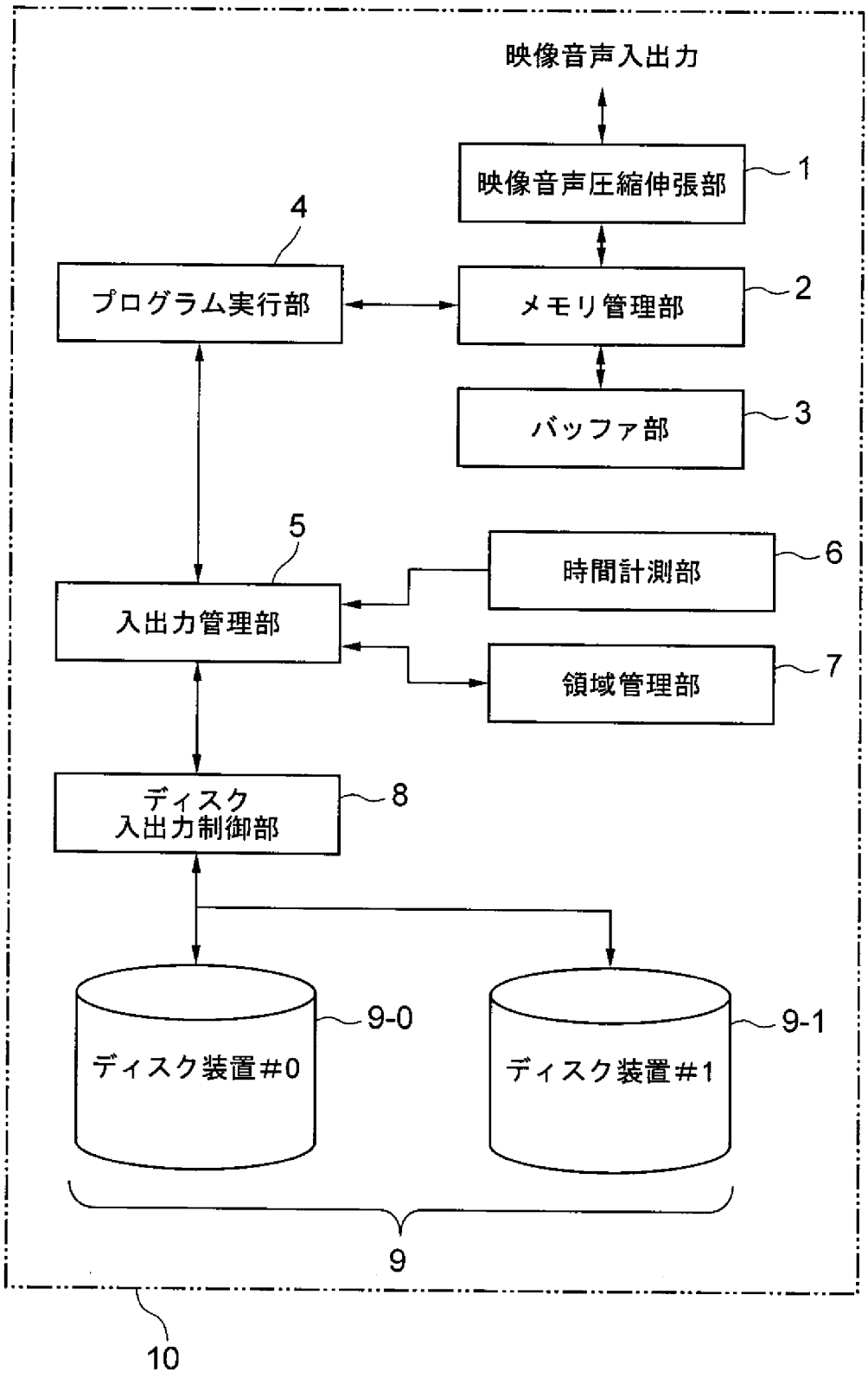
複数の前記ディスク装置からそれぞれ一つの前記部分領域を選択したストライピング領域を特定するストライピング領域特定情報とそのストライピング領域に含まれる各部分領域の前記エリア特定情報とを対応付けたストライピング管理メタデータ、

を参照して前記閾値以上の入出力速度が得られる前記ストライピング領域を選択することを特徴とした請求項31に記載のディスクアレイ制御プログラム。

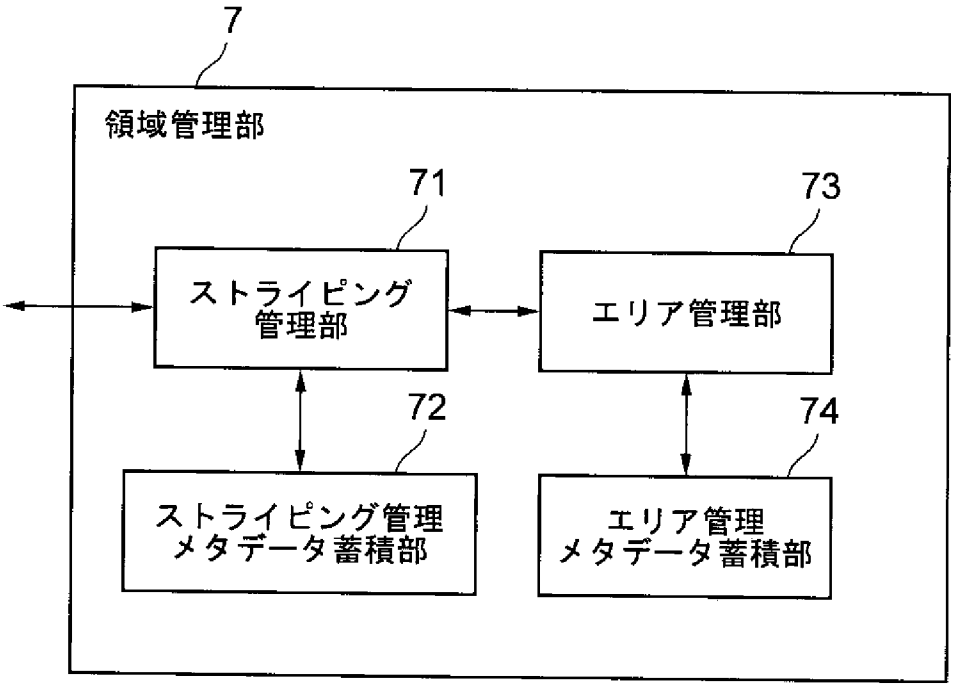
[33] 前記入出力管理処理では、前記ディスク装置に対して書き込みを行う際に、前記閾値の入力を受け、

既存の前記ストライピング領域に前記入力された閾値を上回る前記入出力速度を得られるものがない場合には、前記入力された閾値を上回る前記入力速度が得られるように既存の前記ストライピング領域に組み込まれていない前記部分領域から新たなストライピング領域を生成し、そのストライピング領域に対して前記書き込みを行い、前記新たなストライピング領域に関する情報を前記ストライピング管理メタデータ記憶部に書き込むことを特徴とした請求項31に記載のディスクアレイ制御プログラム。

[図1]


映像音声入出力
映像音声圧縮伸張部
1
4
プログラム実行部
メモリ管理部
2
バッファ部
3
5
時間計測部
6
入出力管理部
領域管理部
7
ディスク
入出力制御部
8
9-0
ディスク装置#0
9-1
ディスク装置#1
9
10

[図2]


7
領域管理部
71
ストライピング
管理部
73
エリア管理部
72
ストライピング管理
メタデータ蓄積部
74
エリア管理
メタデータ蓄積部

[図3]

Index	ストライピング数	エリアID
0	1	#0/0
1	1	#1/0
2	2	#0/1, #1/1

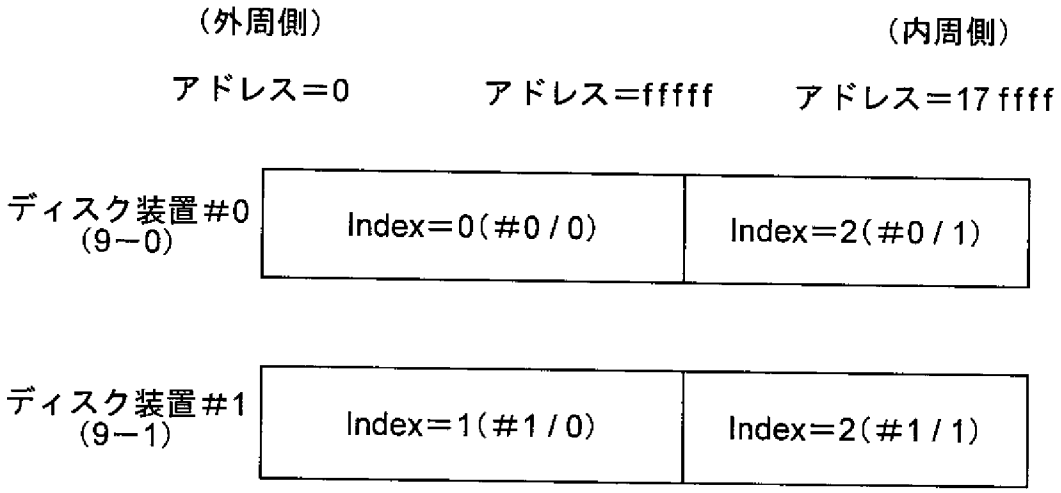
72

[図4]

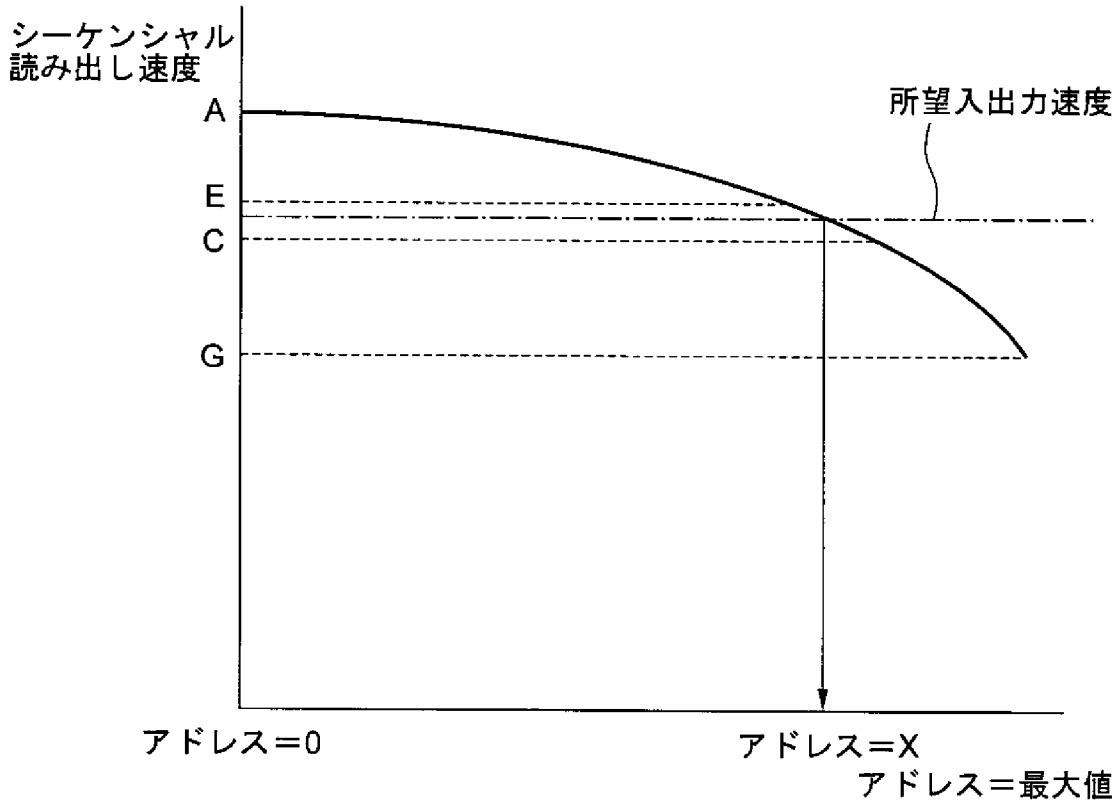
エリアID	ディスク番号	ディスクアドレス範囲	入出力最高性能	入出力最低性能	使用未使用フラグ
#0/0	0	0x0 – 0xfffff	A	E	0
#1/0	1	0x0 – 0xfffff	B	F	0
#0/1	0	0x100000 – 0x17ffff	C	G	0
#1/1	1	0x100000 – 0x17ffff	D	H	0

74

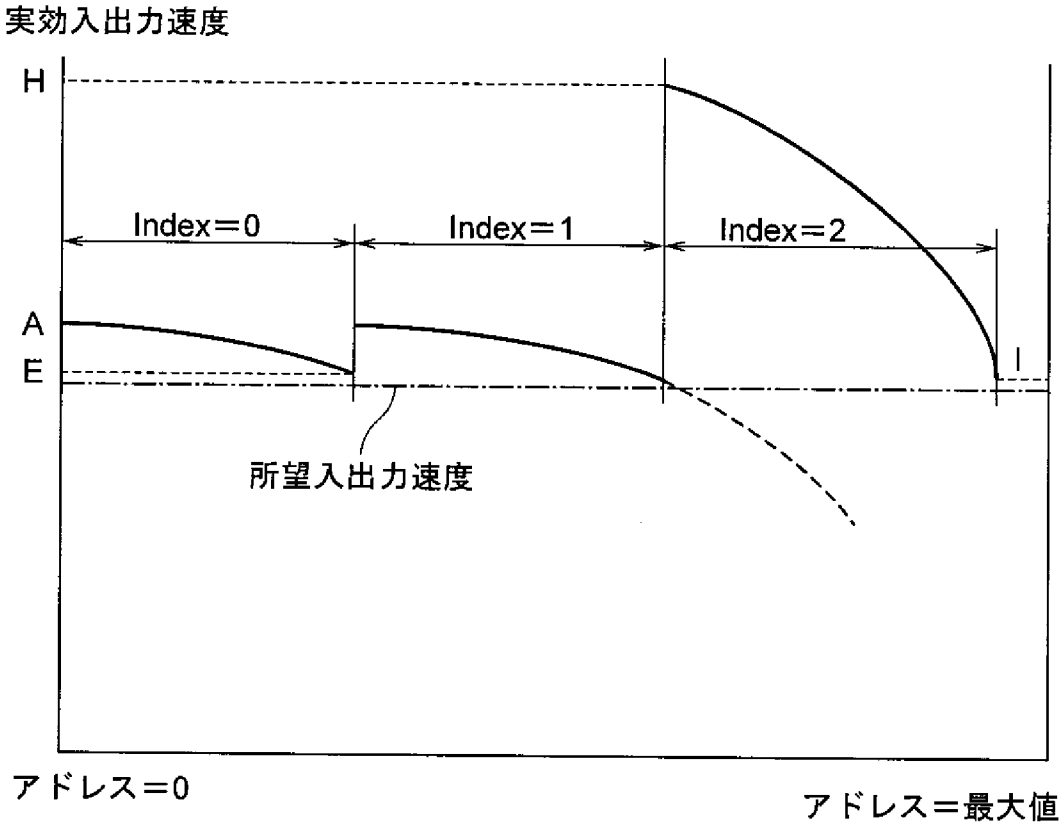
[図5]


（外周側）
（内周側）
アドレス＝0
アドレス＝fffff
アドレス＝17 ffff
ディスク装置#0
(9−0)
Index＝0(#0 / 0)
Index＝2(#0 / 1)
ディスク装置#1
(9−1)
Index＝1(#1 / 0)
Index＝2(#1 / 1)

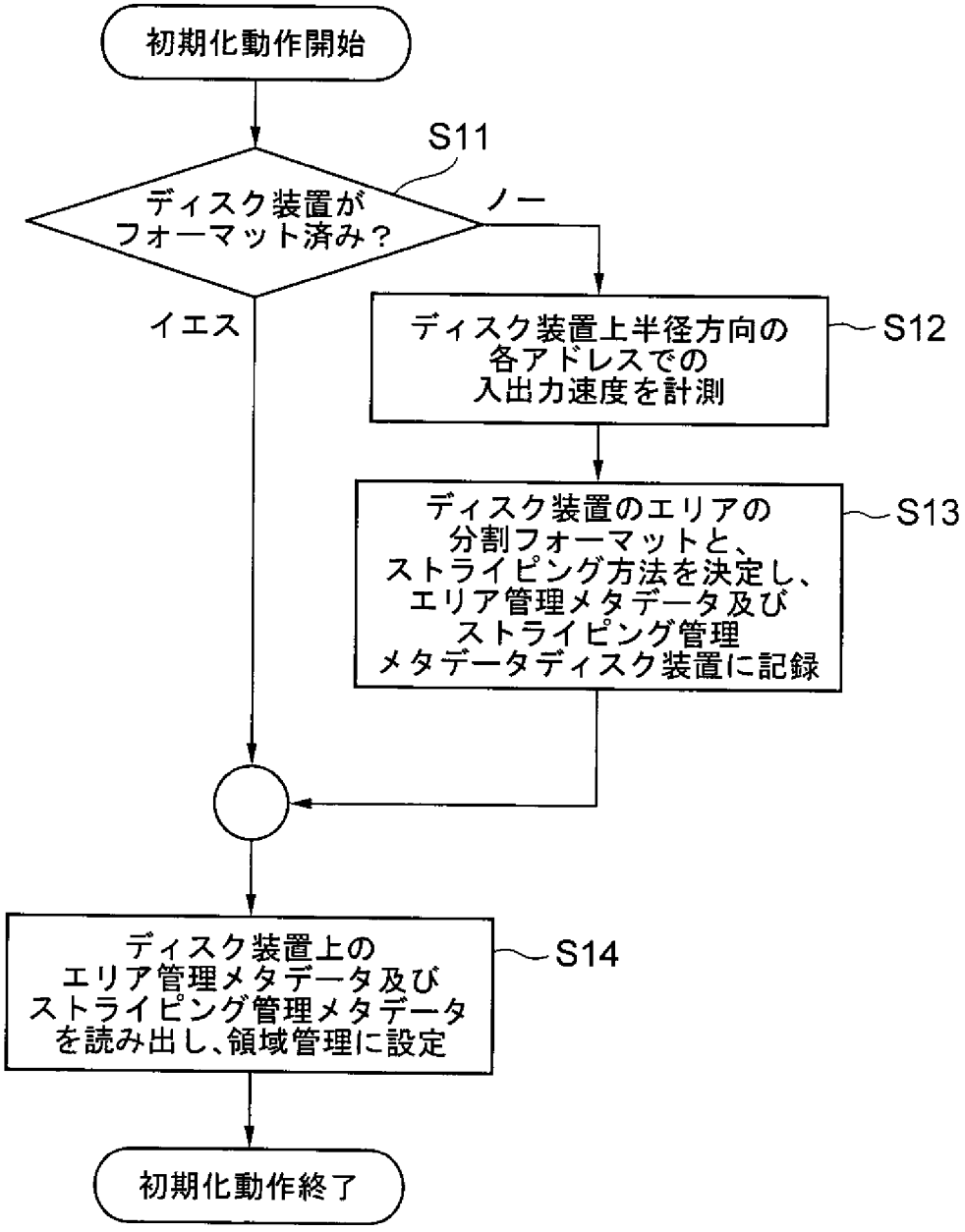
[図6]



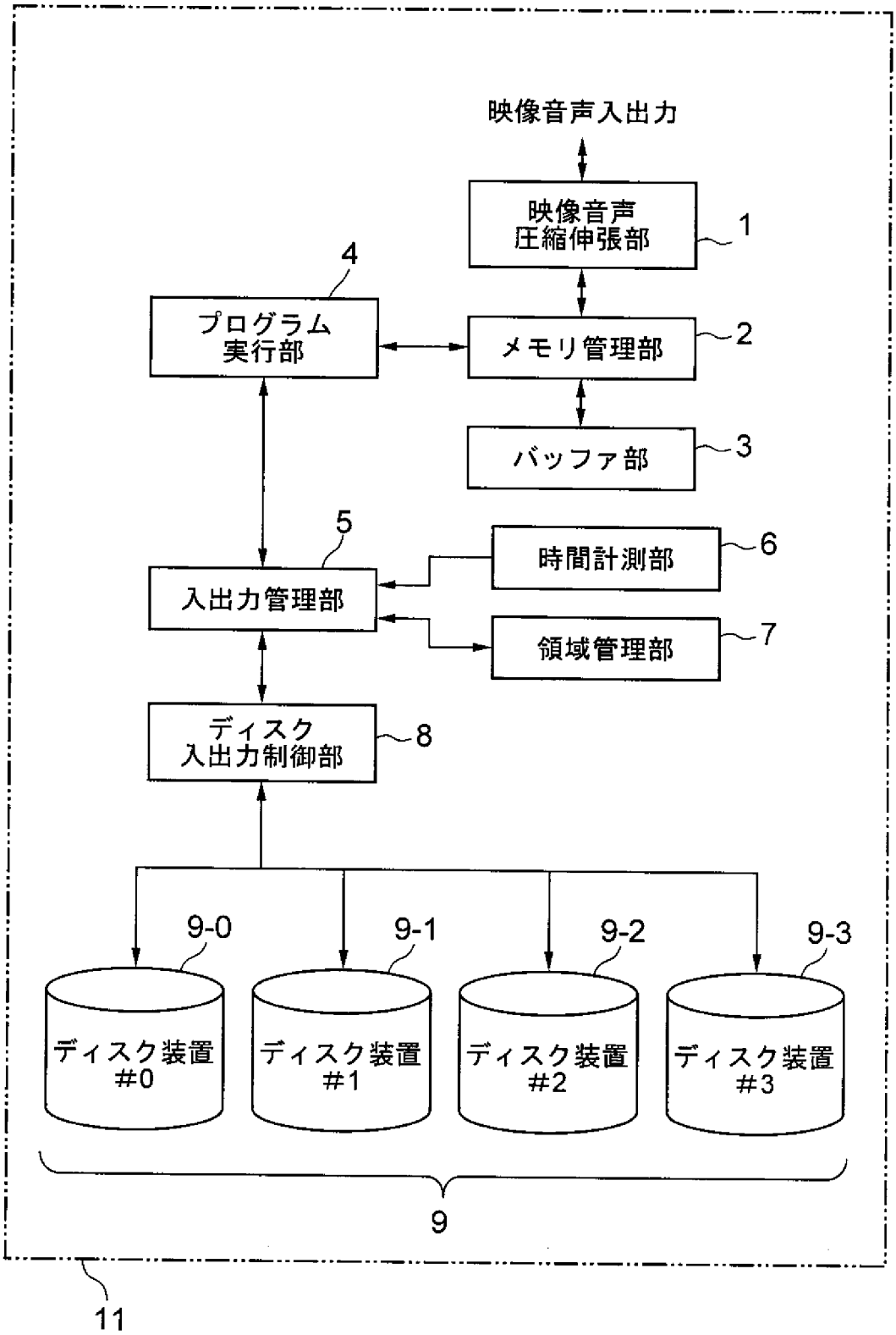
[図7]



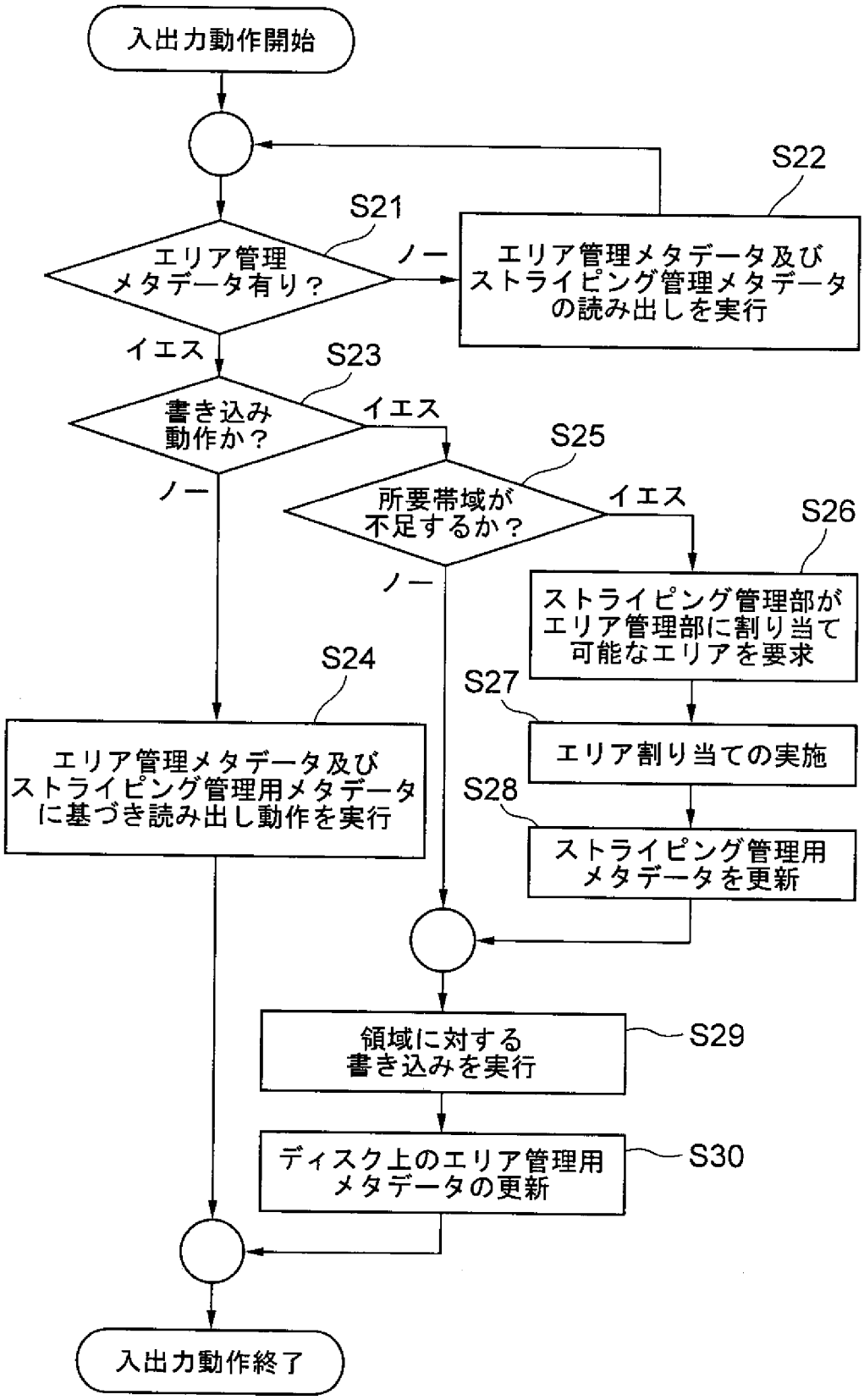
[図8]


初期化動作開始
S11
ディスク装置が
フォーマット済み？
ノー
イエス
ディスク装置上半径方向の
各アドレスでの
入出力速度を計測
S12
ディスク装置のエリアの
分割フォーマットと、
ストライピング方法を決定し、
エリア管理メタデータ及び
ストライピング管理
メタデータディスク装置に記録
S13
ディスク装置上の
エリア管理メタデータ及び
ストライピング管理メタデータ
を読み出し、領域管理に設定
S14
初期化動作終了

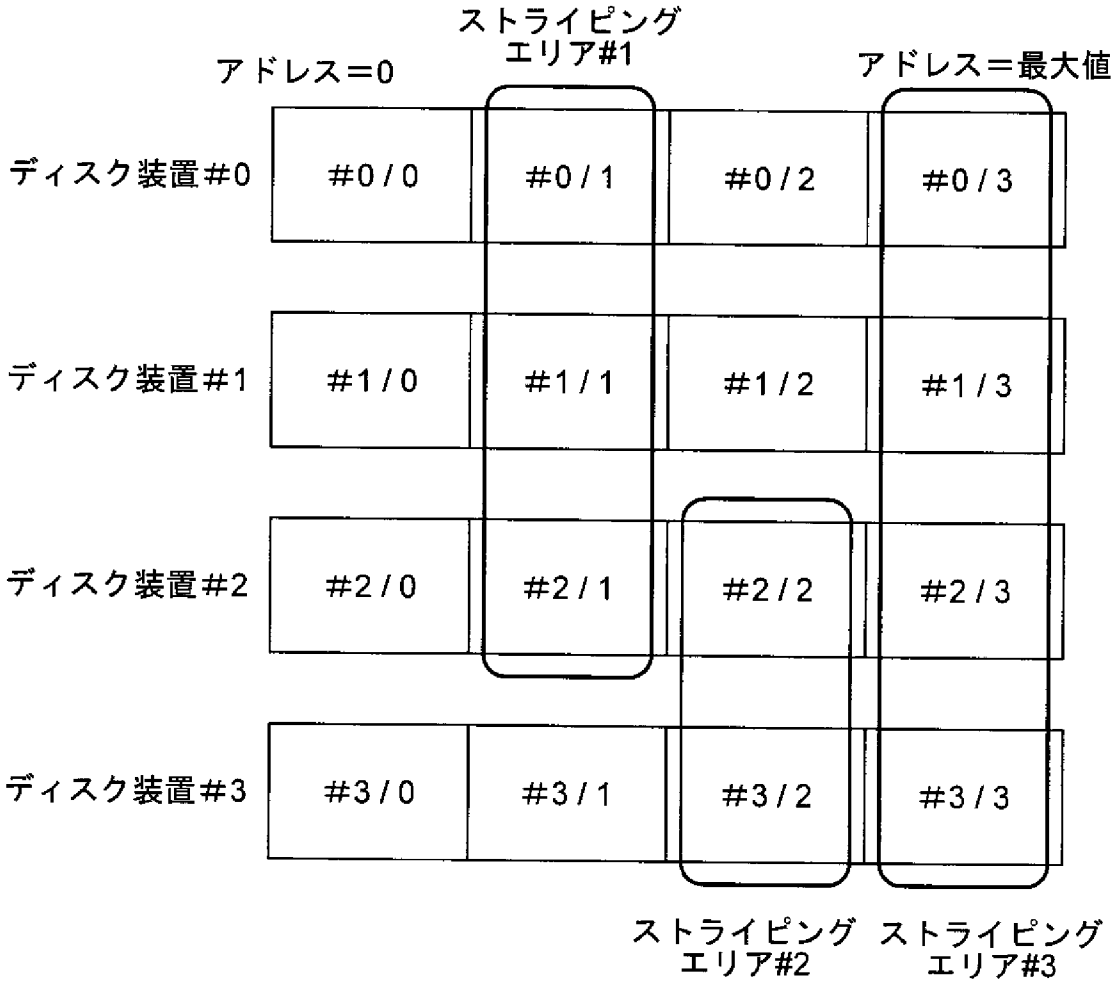
[図9]


映像音声入出力
映像音声
圧縮伸張部
1
4
プログラム
実行部
メモリ管理部
2
バッファ部
3
5
時間計測部
6
入出力管理部
領域管理部
7
ディスク
入出力制御部
8
9-0
9-1
9-2
9-3
ディスク装置
#0
ディスク装置
#1
ディスク装置
#2
ディスク装置
#3
9
11

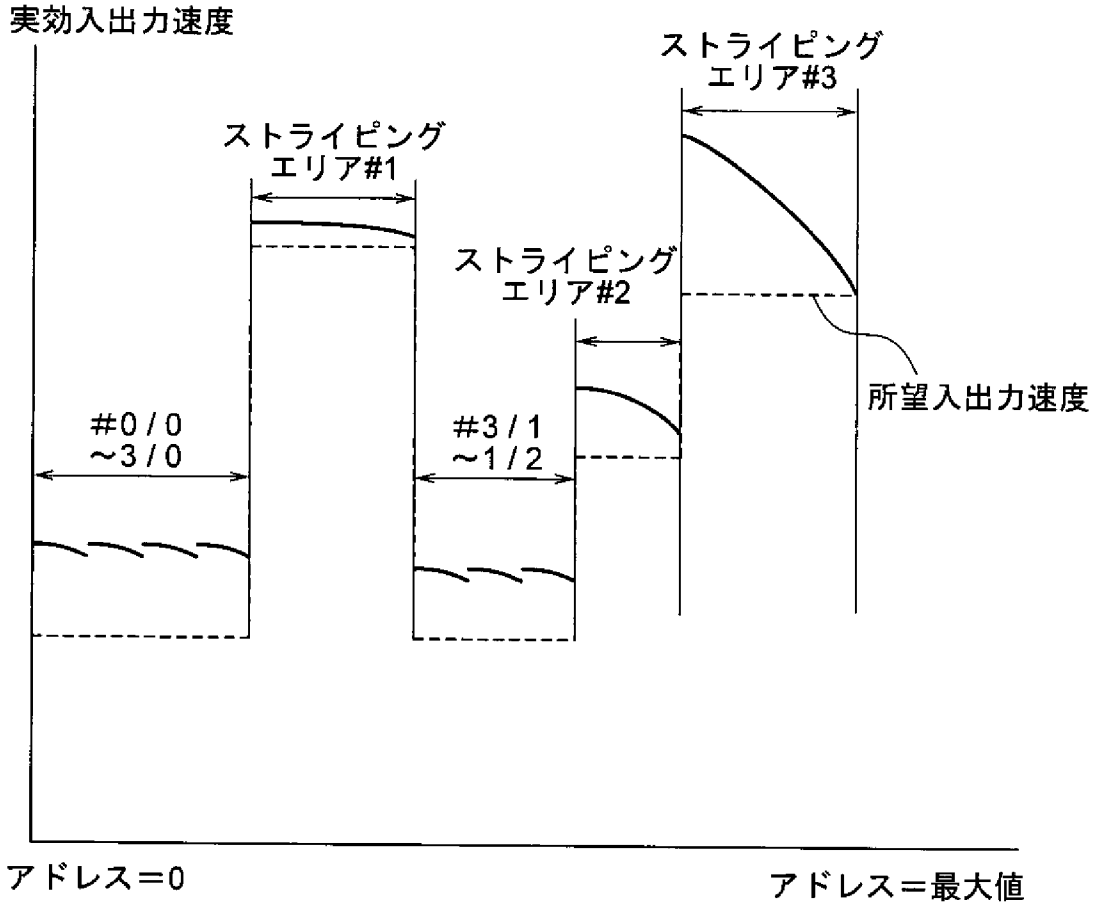
[図10]


入出力動作開始
S21
エリア管理
メタデータ有り？
ノー
S22
エリア管理メタデータ及び
ストライピング管理メタデータ
の読み出しを実行
イエス
S23
書き込み
動作か？
イエス
S25
所要帯域が
不足するか？
イエス
S26
ストライピング管理部が
エリア管理部に割り当て
可能なエリアを要求
ノー
ノー
S24
エリア管理メタデータ及び
ストライピング管理用メタデータ
に基づき読み出し動作を実行
S27
エリア割り当ての実施
S28
ストライピング管理用
メタデータを更新
領域に対する
書き込みを実行
S29
ディスク上のエリア管理用
メタデータの更新
S30
入出力動作終了

[図11]


ストライピング
エリア#1
アドレス＝0
アドレス＝最大値
ディスク装置#0
ディスク装置#1
ディスク装置#2
ディスク装置#3
#0/0
#0/1
#0/2
#0/3
#1/0
#1/1
#1/2
#1/3
#2/0
#2/1
#2/2
#2/3
#3/0
#3/1
#3/2
#3/3
ストライピング
エリア#2
ストライピング
エリア#3

[図12]


実効入出力速度
ストライピング
エリア#3
ストライピング
エリア#1
ストライピング
エリア#2
所望入出力速度
#0/0
～3/0
#3/1
～1/2
アドレス＝0
アドレス＝最大値

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2007/066339

A. CLASSIFICATION OF SUBJECT MATTER
G06F3/06(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

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 1996-2007
Kokai Jitsuyo Shinan Koho 1971-2007 Toroku Jitsuyo Shinan Koho 1994-2007

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.
X	JP 08-249132 A (Hitachi, Ltd.), 27 September, 1996 (27.09.96), Par. Nos. [0021] to [0070]; Figs. 1, 4, 6 & US 5754756 A & EP 0732653 A2 & DE 69632219 T	1,8,12,19, 23,30
X	JP 08-115171 A (Hitachi, Ltd.), 07 May, 1996 (07.05.96),	9-11,20-22, 31-33
Y	Par. Nos. [0094] to [0122]; Figs. 6, 10 (Family: none)	1-8,12-19, 23-30
Y	JP 08-328759 A (Mitsubishi Electric Corp.), 13 December, 1996 (13.12.96), Par. Nos. [0056] to [0061] & US 5761526 A	1-8,12-19, 23-30

☒ 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
16 November, 2007 (16.11.07)	27 November, 2007 (27.11.07)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2007/066339

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 2006-106980 A (Hitachi, Ltd.), 20 April, 2006 (20.04.06), Par. No. [0035] & US 2006/0075190 A1 & EP 1643353 A2	1-33
A	JP 2003-521759 A (3Ware Co.), 15 July, 2003 (15.07.03), Full text; all drawings & US 6484235 B1 & US 6487633 B1 & US 6591338 B1 & US 6591339 B1 & EP 1198793 A & WO 2000/067250 A2 & AU 4702700 A	1-33
A	JP 09-198194 A (Symbios Logic Inc.), 31 July, 1997 (31.07.97), Par. Nos. [0022] to [0024] & US 5719983 A & EP 0780758 A2 & DE 69624667 T	1-33
A	JP 10-172246 A (Fuji Xerox Co., Ltd.), 26 June, 1998 (26.06.98), Fig. 2 (Family: none)	1-33
A	JP 10-055597 A (Sony Corp.), 24 February, 1998 (24.02.98), Full text; all drawings (Family: none)	1-33
A	JP 07-084732 A (Hewlett-Packard Co.), 31 March, 1995 (31.03.95), Par. Nos. [0024] to [0039] & US 5392244 A & EP 0639811 A2 & DE 69432545 T	1-33
A	JP 08-063298 A (Hitachi, Ltd.), 08 March, 1996 (08.03.96), Par. Nos. [0024] to [0068] (Family: none)	1-33

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２００７／０６６３３９

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

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

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	１９７１－２００７年
日本国実用新案登録公報	１９９６－２００７年
日本国登録実用新案公報	１９９４－２００７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	関連する請求の範囲の番号
X	JP 08-249132 A（株式会社日立製作所）1996.09.27, 段落【００２１】－【００７０】、図１、図４、図６ & US 5754756 A & EP 0732653 A2 & DE 69632219 T	１，８，１２， １９，２３， ３０
X	JP 08-115171 A（株式会社日立製作所）1996.05.07, 段落【００９４】－【０１２２】、図６、図１０（ファミリーなし）	９－１１， ２０－２２， ３１－３３
Y		１－８， １２－１９，

☑　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 １６．１１．２００７	国際調査報告の発送日 ２７．１１．２００７
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）　５Ｉ　３８０１ 梅景　篤 電話番号　０３－３５８１－１１０１　内線　３５６５

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	関連する請求の範囲の番号
		２３－３０
Ｙ	JP 08-328759 A（三菱電機株式会社）1996.12.13, 段落【００５６】－【００６１】 & US 5761526 A	１－８， １２－１９， ２３－３０
Ａ	JP 2006-106980 A（株式会社日立製作所）2006.04.20, 段落【００３５】 & US 2006/0075190 A1 & EP 1643353 A2	１－３３
Ａ	JP 2003-521759 A（３ウェア　カンパニー）2003.07.15, 全文、全図 & US 6484235 B1 & US 6487633 B1 & US 6591338 B1 & US 6591339 B1 & EP 1198793 A & WO 2000/067250 A2 & AU 4702700 A	１－３３
Ａ	JP 09-198194 A（シンバイオス・ロジック・インコーポレイテッド） 1997.07.31, 段落【００２２】－【００２４】 & US 5719983 A & EP 0780758 A2 & DE 69624667 T	１－３３
Ａ	JP 10-172246 A（富士ゼロックス株式会社）1998.06.26, 図２ （ファミリーなし）	１－３３
Ａ	JP 10-055597 A（ソニー株式会社）1998.02.24, 全文、全図 （ファミリーなし）	１－３３
Ａ	JP 07-084732 A（ヒューレット・パッカード・カンパニー） 1995.03.31, 段落【００２４】－【００３９】 & US 5392244 A & EP 0639811 A2 & DE 69432545 T	１－３３
Ａ	JP 08-063298 A（株式会社日立製作所）1996.03.08, 段落【００２４】－【００６８】（ファミリーなし）	１－３３