

(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



(10) 国際公開番号

WO 2011/121869 A1

(43) 国際公開日

2011 年 10 月 6 日(06.10.2011)

PCT

- (51) 国際特許分類:
G06F 12/00 (2006.01) G06F 13/10 (2006.01)
G06F 3/06 (2006.01)
- (21) 国際出願番号: PCT/JP2011/000311
- (22) 国際出願日: 2011 年 1 月 21 日(21.01.2011)
- (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 2010-075514 2010 年 3 月 29 日(29.03.2010) 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): 菅真樹 (KAN, Masaki) [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). 小林大 (KOBAYASHI, Dai) [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). 板橋康雄 (ITABASHI, Yasuo) [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).

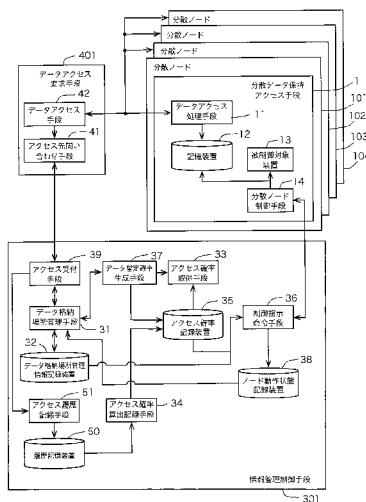
- (74) 代理人: 岩壁冬樹, 外 (IWAKABE, Fuyuki et al.); 〒1040031 東京都中央区京橋一丁目 4 番 1 0 号大野屋京橋ビル 6 階 サンライズ国際特許事務所 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, 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, RO, RS, RU, 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.

[続葉有]

(54) Title: DATA ACCESS LOCATION SELECTION SYSTEM, METHOD AND PROGRAM

(54) 発明の名称: データアクセス場所選択システム、方法およびプログラム

【図2】



(57) Abstract: An access destination determination means determines a read destination and/or a write destination of designated data from among nodes corresponding to data stored in a data store destination storage means. An access probability calculation means calculates an access probability on the basis of a history of access to the data, and stores the calculated access probability in an access probability storage means. If the node corresponding to the designated data is not stored in the data store destination storage means, an access probability estimation means estimates the access probability for the data on the basis of the access probability stored in the access probability storage means, and the access destination determination means determines a write destination of the data based on the estimated access probability.

(57) 要約: アクセス先決定手段は、指定されたデータの読み取り先および/または書き込み先をデータ格納先記憶手段に記憶されているデータに対応するノードの中から決定する。アクセス確率算出手段は、データへのアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させる。指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、アクセス確率推定手段は、アクセス確率記憶手段に記憶されたアクセス確率に基づいてそのデータのアクセス確率を推定し、アクセス先決定手段は、推定されたアクセス確率をもとにそのデータの書き込み先を決定する。

- 1 DISTRIBUTED DATA HOLDING AND ACCESSING MEANS
- 11 DATA ACCESS PROCESSING MEANS
- 12 STORAGE DEVICE
- 13 DEVICE TO BE CONTROLLED
- 14 DISTRIBUTED NODE CONTROL MEANS
- 101, 102, 103, 104 DISTRIBUTED NODE
- 301 INFORMATION MANAGEMENT CONTROL MEANS
- 31 DATA STORE LOCATION MANAGEMENT MEANS
- 32 DATA STORE LOCATION MANAGEMENT INFORMATION RECORDING DEVICE
- 33 ACCESS PROBABILITY ACQUISITION MEANS
- 34 ACCESS PROBABILITY CALCULATION RECORDING MEANS
- 35 ACCESS PROBABILITY RECORDING DEVICE
- 36 CONTROL INSTRUCTION COMMAND MEANS
- 37 DATA PROVISIONAL PROBABILITY GENERATION MEANS
- 38 NODE OPERATIONAL STATE RECORDING DEVICE
- 39 ACCESS RECEIVING MEANS
- 401 DATA ACCESS REQUEST MEANS
- 41 ACCESS DESTINATION INQUIRY MEANS
- 42 DATA ACCESS MEANS
- 50 HISTORY STORAGE DEVICE
- 51 ACCESS HISTORY RECORDING MEANS



(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, 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 (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：

データアクセス場所選択システム、方法およびプログラム

技術分野

[0001] 本発明は、分散ストレージシステムにおいて、データの読み取り先および／または書き込み先を選択するデータアクセス場所選択システム、データアクセス場所選択方法及びデータアクセス場所選択プログラムに関する。

背景技術

[0002] 近年、コンピュータシステムが扱うデータ量が増大している。そのため、大規模なコンピュータシステムにおいては、複数のストレージノードにデータを分散して格納する分散ストレージシステムが利用されている。

[0003] その一方で、コンピュータシステムやストレージシステムが大規模化することにより、コンピュータシステムの消費電力量が問題になりつつある。

[0004] そのため、ディスク装置の回転速度を変更したり、ディスクを停止したりすることによりストレージシステムの省電力化が試みられている。このような仕組みを用いた技術の例が、特許文献 1 に記載されている。特許文献 1 に記載された方法では、ディスク装置にアクセスしない場合にディスク装置を停止させることで、システムの消費電力量を削減している。

[0005] また、分散ストレージシステムにおいて、分散ストレージノードへのアクセス統計情報をもとに算出された無アクセス確率を利用して、データの再配置を行うシステムが特許文献 2 に記載されている。特許文献 2 に記載されたシステムでは、無アクセス時間の期待値が高くなるようにデータ配置を行うことにより、個々ディスクノードが省電力モードになる時間を長くすることができる。

[0006] さらに、特許文献 3 には、I/O（Input/Output）負荷を適切に分散する計算機システムが記載されている。特許文献 3 に記載された計算機システムでは、各統合デバイスに対する I/O アクセス頻度の平均を、

I / O アクセス頻度の予想値として算出する。

先行技術文献

特許文献

- [0007] 特許文献1：特表2005-539303号公報（段落0023）
特許文献2：特開2009-157571号公報（段落0097～0102）
特許文献3：特開2008-59315号公報（段落0194，0198、図12，22）

発明の概要

発明が解決しようとする課題

- [0008] ある計算機ノードのディスクにアクセスが一定期間無い場合にディスクを停止させるなどの省電力化制御を、例えば、特許文献1に記載された方法を用いて行ったとする。しかし、ディスク停止処理の直後に該当ノードへのアクセスがあった場合、ディスクの再回転やノードの再起動などが発生することになる。そのため、ディスクの再回転やノードの再起動により、かえって電力コストが増大してしまう。さらに、クライアント側では、アクセス待ちによる稼働時間も増大するため、電力消費量がさらに増大してしまう。
- [0009] このように、特許文献1に記載された方法では、分散システムの構成ノードの動作状態制御を適切に行うことができず、過剰な電力消費を発生させてしまう恐れがある。
- [0010] また、特許文献2に記載されたシステムを用いた場合、新たなデータの登録後に算出される無アクセス確率に基づいてデータが再配置されることになる。そのため、一旦登録されたデータであっても、算出される確率に基づいて移動させられる蓋然性が高く、その場合には移動コストがかかってしまうという問題がある。ハードディスクドライブ（以下、HDDと記す。）などの記憶装置においては、近年大容量化が進んでいるため、データのコピーには長い時間が必要であり、その分余剰な電力消費が発生してしまうことになる。
- [0011] また、特許文献3に記載された計算機システムでは、新たに追加されるデ

ータについてのアクセス頻度は考慮されておらず、また、既存のストレージのアクセス状態によって新たなデータの記憶先が決定される。このように、新たなデータのアクセス頻度が考慮されずに記憶されてしまうと、そのデータを記憶するノードを適切に制御することが出来ず、結果として、データアクセスに係る消費電力が増大してしまうという結果につながる。そのため、新たなデータが追加されるような場合であっても、データアクセスに係る電力消費を適切に抑制できるようにすることが望ましい。

[0012] そこで、本発明は、分散ストレージシステムにおいて、新たなデータが追加されるような場合であっても、データアクセスに係る電力消費を抑制できるデータアクセス場所選択システム、データアクセス場所選択方法及びデータアクセス場所選択プログラムを提供することを目的とする。

課題を解決するための手段

[0013] 本発明によるデータアクセス場所選択システムは、データを格納する少なくとも1つ以上のノードと、データの読み取り先および／または書き込み先をノードの中から選択する管理サーバとを備え、管理サーバが、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段と、指定されたデータの読み取り先および／または書き込み先を、データ格納先記憶手段に記憶されているそのデータに対応するノードの中から決定するアクセス先決定手段と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、データへのアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出手段と、指定されたデータに対するアクセス確率を、アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、アクセス確率推定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、そのデータのアクセス確率を推定し、アクセス先決定手段は、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先

を決定することを特徴とする。なお、「読み取りおよび／または書き込み」とは、「読み取り」、「書き込み」、または「読み取りと書き込みの双方」を表す。

[0014] 本発明による他のデータアクセス場所選択システムは、データを格納する複数のノードを備え、各ノードが、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段と、指定されたデータの読み取り先および／または書き込み先を、データ格納先記憶手段に記憶されているそのデータに対応するノードの中から決定するアクセス先決定手段と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、データへのアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出手段と、指定されたデータに対するアクセス確率を、アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、アクセス確率推定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、そのデータのアクセス確率を推定し、アクセス先決定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定することを特徴とする。

[0015] 本発明によるデータアクセス場所選択方法は、データを格納する少なくとも1つ以上のノードの中から、データの読み取り先および／または書き込み先を選択する選択装置が、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されているそのデータに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定し、選択装置が、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確率をアクセス確率記憶手段に記憶させ、選択装置が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されてい

い場合、指定されたデータに対するアクセス確率をアクセス確率記憶手段に記憶されたアクセス確率に基づいて推定し、選択装置が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定することを特徴とする。

[0016] 本発明によるデータアクセス場所選択プログラムは、データを格納する少なくとも1つ以上のノードの中から、データの読み取り先および／または書き込み先を選択するコンピュータに適用されるデータアクセス場所選択プログラムであって、コンピュータに、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されているそのデータに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定するアクセス先決定処理、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出処理、および、指定されたデータに対するアクセス確率を、アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定処理を実行させ、アクセス確率推定処理で、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、そのデータのアクセス確率を推定させ、アクセス先決定処理で、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定させることを特徴とする。

発明の効果

[0017] 本発明によれば、分散ストレージシステムにおいて新たなデータが追加されるような場合であっても、データアクセスに係る電力消費を抑制できる。

図面の簡単な説明

[0018] [図1]本発明によるデータアクセス場所選択システムの一実施形態を示す説明図である。

[図2] データアクセス場所選択システムの例を示すブロック図である。

[図3] データの格納先がテーブル形式で記憶された場合の例を示す説明図である。

[図4] アクセス要求命令の履歴の一例を示す説明図である。

[図5] アクセス確率の例を示す説明図である。

[図6] アプリケーションXがデータアクセスする動作の例を示す説明図である。

[図7] ノードの動作状態の例を示す説明図である。

[図8] 本発明におけるデータアクセス処理の例を示すフローチャートである。

[図9] 本発明によるデータアクセス場所選択システムの最小構成の例を示すブロック図である。

[図10] 本発明によるデータアクセス場所選択装置の最小構成の例を示すブロック図である。

発明を実施するための形態

[0019] 以下、本発明の実施形態を図面を参照して説明する。以下の説明では、本発明によるデータアクセス場所選択システムを、複数の計算機ノードから構成される分散ストレージシステムに適用する場合について説明する。

[0020] 図1は、本発明によるデータアクセス場所選択システムの一実施形態を示す説明図である。本発明におけるデータアクセス場所選択システムは、分散ノード101～104と、情報管理制御装置300と、データアクセス要求ノード400とを備えている。分散ノード101～104と、情報管理制御装置300と、データアクセス要求ノード400とは、通信ネットワーク200を介してそれぞれ接続される。

[0021] 分散ノード101～104は、分散ストレージシステムに用いられる計算機であり、それぞれ単数又は複数の記憶装置111～114を備えている。また、分散ノード101～104は、動作状態の制御が可能な要素（以下、被制御要素と記す。）を備えている。この被制御要素とは、分散ノードに用いられる各装置のことであり、例えば、演算装置であるCPU、HDD、計

算機の電源装置などが挙げられる。また、動作状態の制御とは、各被制御要素が行う動作に対して制御指示を行うことであり、例えば、CPUの回転数を下げる、HDDの回転数を下げる、CPUを停止させる、HDDを停止させる、電源装置に電源を落とさせる、などが挙げられる。

[0022] 情報管理制御装置300は、分散ストレージシステムにおいて格納されるデータの配置場所を決定するための情報を保持し、その情報を管理する機能を備える装置である。情報管理制御装置300の役割は、例えば、分散ストレージシステムの一例である分散ファイルシステムにおいて、ファイルのデータブロックをどの分散ノードに格納するかを管理し、そのデータブロックを追跡することに責任を持つメタデータサーバの役割に相当する。

[0023] なお、メタデータサーバを備える分散ファイルシステムについては、以下の参考文献に記載されている。

[参考文献]

特開2003-522360号公報

[0024] データアクセス要求ノード400は、分散ストレージシステムに対してデータアクセスを行い、アクセスしたデータを用いて任意の演算を行う計算機ノードである。このデータアクセスには、データの読み取り処理、データの書き込み処理、または、データの読み取りと書き込みの双方のいずれかが含まれる。なお、書き込み処理は、新規データの書き込み処理の他、既存データの更新処理及び削除処理を含む。データアクセス要求ノード400は、分散ノード101～104とは別の計算機ノードであってもよく、また、分散ノード101～104自身がデータアクセス要求ノード400として動作しても（すなわち、データアクセス及び演算を行っても）よい。以下の説明では、データアクセス要求ノード400と、分散ノード101～104とが、別の計算機ノードで実現されている場合について説明する。

[0025] 分散ノード101～104、情報管理制御装置300、および、データアクセス要求ノード400は、例えば、CPU、半導体メモリ、通信インタフェースなどを備えるIA（Intel Architecture）サーバ

により実現される。ただし、Intelは、登録商標である。

[0026] 図1に示す例では、データアクセス場所選択システムが分散ノード101から分散ノード104まで4台の分散ノードを備えている場合について示している。ただし、分散ノードの個数は4つに限定されず、2台以上存在すればよい。また、図1に示す例では、各分散ノードが2つの記憶装置を備えている場合について示しているが、記憶装置の個数は、2つに限定されない。記憶装置の個数は、1台であってもよく、3台以上であってもよい。なお、分散ノード101～104までの機能は同様であるため、以下の説明では、分散ノード101を対象として説明する。

[0027] 図2は、データアクセス場所選択システムの例を示すブロック図である。本発明によるデータアクセス場所選択システムは、分散データ保持アクセス手段1、情報管理制御手段301、および、データアクセス要求手段401により実現される。

[0028] 以下の説明では、分散データ保持アクセス手段1が図1における各分散ノード101において、情報管理制御手段301が図1における情報管理制御装置300において、データアクセス要求手段401が図1におけるデータアクセス要求ノード400において、それぞれ動作する場合について説明する。ただし、各手段（より詳しくは、分散データ保持アクセス手段1、情報管理制御手段301、および、データアクセス要求手段401）が動作する装置は、上記装置に限定されない。

[0029] 例えば、図2に示す例では、情報管理制御手段301が分散ノード101と異なる装置（情報管理制御装置300）において動作する場合について示している。ただし、情報管理制御手段301は、分散ノード101～104のうちの1つ又は複数の計算機において動作してもよい。具体的には、情報管理制御手段301として動作させるためのプログラム（ソフトウェア）を分散ノード101～104上に記憶させ、例えば、CPUがそのプログラムに従って動作するようにしてもよい。すなわち、本実施形態によるデータアクセス場所選択システムが、複数の分散ノードを備えるようにしてもよい。

- [0030] 分散データ保持アクセス手段 1 は、データアクセス処理手段 1 1 と、記憶装置 1 2 と、被制御対象装置 1 3 と、分散ノード制御手段 1 4 とを備えている。上述の被制御要素が、図 2 における被制御対象装置 1 3 に相当する。
- [0031] 記憶装置 1 2 は、分散ストレージシステムにおいて対象とされるデータを記憶する。なお、図 1 に例示する記憶装置 1 1 1 が、記憶装置 1 2 に相当する。記憶装置 1 2 は、例えば、HDD や SSD (Solid State Drive)、半導体メモリなどの記憶媒体により実現される。
- [0032] データアクセス処理手段 1 1 は、データアクセス要求手段 4 0 1 からのデータアクセス要求を受け付ける。そして、データアクセス処理手段 1 1 は、記憶装置 1 2 にアクセスして要求対象のデータを読み取り、データアクセス要求手段 4 0 1 に結果を応答する。
- [0033] 被制御対象装置 1 3 は、動作状態の制御が可能な装置（すなわち、被制御要素）である。具体的には、被制御対象装置 1 3 は、分散データ保持アクセス手段 1 において、動作状態の制御が可能なハードウェア装置を示す。被制御対象装置 1 3 は、例えば、CPU などの演算装置、余剰な通信インターフェースなどである。なお、記憶装置 1 2 も、被制御要素 1 2 の一つであるが、図 2 に示す例では、被制御対象装置 1 3 を記憶装置 1 2 と別に示している。
- [0034] 分散ノード制御手段 1 4 は、動作状態の制御が可能な装置（具体的には、被制御要素である記憶装置 1 2 及び被制御対象装置 1 3）に対して制御指示を行う。なお、制御指示の内容は、例えば、被制御要素の説明で例示した内容である。
- [0035] データアクセス要求手段 4 0 1 は、アクセス先問い合わせ手段 4 1 と、データアクセス手段 4 2 とを備えている。アクセス先問い合わせ手段 4 1 は、読み取るデータの読み込み先情報（例えば、データが記憶された分散ノードを表す情報）や、書き込みを行うデータの書き込み先情報（例えば、データを記憶する分散ノードを表す情報）を情報管理制御手段 3 0 1 に問い合わせ、その問合せ結果（以下、アクセス先情報と記す。）を受け取る。以下、データアクセス要求手段 4 0 1 が情報管理制御手段 3 0 1 に対してアクセス先

を問い合わせる要求のことを、データアクセス要求命令と記す。

[0036] データアクセス手段 4 2 は、アクセス先問い合わせ手段 4 1 がアクセス先情報を受け取ると、そのアクセス先である分散ノード 1 0 1 に対してデータアクセスを行う。具体的には、データアクセス手段 4 2 は、分散ノード 1 0 1 の記憶装置 1 2 に対して行いたいアクセス処理（例えば、データの書き込み処理や、読み取り処理）を、データアクセス処理手段 1 1 に要求する。なお、以下の説明では、データの書き込むことを *Write*、データを読み取ることを *Read* と記すこともある。

[0037] 情報管理制御手段 3 0 1 は、データ格納場所管理手段 3 1 と、データ格納場所管理情報記録装置 3 2 と、アクセス確率取得手段 3 3 と、アクセス確率算出記録手段 3 4 と、アクセス確率記録装置 3 5 と、制御指示命令手段 3 6 と、データ暫定確率生成手段 3 7 と、ノード動作状態記録装置 3 8 と、アクセス受付手段 3 9 と、履歴記憶装置 5 0 と、アクセス履歴記録手段 5 1 とを備えている。

[0038] アクセス受付手段 3 9 は、データアクセス要求手段 4 0 1 からのデータアクセス要求命令を受け付ける。そして、アクセス受付手段 3 9 は、データ格納場所管理手段 3 1 にアクセス先の選択を依頼し、選択結果をデータアクセス要求手段 4 0 1 に送信する。

[0039] ここで、データアクセス要求命令には、データを識別する情報、実行命令、及び実行に必要なデータが含まれる。アクセス受付手段 3 9 は、例えば、「ファイル A を *Read* する」、「ファイル B にデータ "XXXX" を *Write* する」のようなデータアクセス要求命令を受け付ける。この場合、「*Read* する」や「*Write* する」が実行命令に相当する。また、実行に必要なデータとして、例えば、*Write* コマンドにおけるデータが挙げられる。なお、このデータは、SCSI (Small Computer System Interface) では、データ・フェーズにおいてイニシエータからターゲットに転送されるデータに相当する。SCSI プロトコルを TCP/IP 上で使用する規格である iSCSI では、実行命令である *Write* コマンドが SCSI C

ommand PDU内で送信され、Writeコマンドのデータは、SCSI DATA Out PDUで送信される。また、データを識別する情報は、Write先のブロックアドレス等が相当し、Writeコマンド内に含まれる。また、例えば、分散Key Value Storeシステムにおいては、put、putの際のKey及びKeyに関連付けられて格納されるValueが、データアクセス要求命令に含まれる。このとき、putが実行命令であり、Keyがデータを識別する情報、Valueが必要なデータである。

[0040] データ格納場所管理情報記録装置32は、分散ストレージシステムで保持管理するデータの格納先を記憶する。具体的には、データ格納場所管理情報記録装置32は、データの格納先を分散ノード101と対応付けて記憶する。データ格納場所管理情報記録装置32が記憶するデータの格納先は、後述のデータ格納場所管理手段31により記憶される。なお、既に分散ノード101に記憶されているデータの格納先については、例えば、ユーザ等により予めデータ格納場所管理情報記録装置32に記憶させておいてもよい。

[0041] また、データ格納場所管理情報記録装置32が記憶する内容は、例えば、「/home/node1」以下のディレクトリおよびファイルを分散ノード101に、「/home/node2」以下のディレクトリおよびファイルを分散ノード102に格納するといったルールを示す情報であってもよい。データ格納場所管理情報記録装置32は、磁気ディスクなど、任意の記録媒体により実現される。

[0042] データ格納場所管理手段31は、分散ストレージシステムで保持管理するデータの実格納先を管理し、アクセス受付手段39のアクセス要求において指定されたデータのアクセス先を決定する。具体的には、指定されたデータに対応する分散ノードの情報がデータ格納場所管理情報記録装置32に記憶されている場合、データ格納場所管理手段31は、指定されたデータに対応する分散ノードの中からデータの読み取り先および/または書き込み先を決定する。

- [0043] 例えば、データ格納場所管理情報記録装置 3 2 がデータベースソフトウェアで定められたテーブル形式でデータの格納先を記憶している場合、データ格納場所管理手段 3 1 は、そのテーブルに対する検索処理を行うことで、データの格納先の検索を行ってもよい。
- [0044] 図 3 は、データの格納先がテーブル形式で記憶された場合の例を示す説明図である。図 3 に示す例では、データ格納場所管理情報記録装置 3 2 がデータを識別するファイル名と、そのファイルを格納しているノードとを対応付けて記憶していることを示す。また、図 3 に示す例では、データをファイル単位で複数のノードに分散して格納していることを示す。なお、分散ノード 1 0 1, 分散ノード 1 0 2, 分散ノード 1 0 3 及び分散ノード 1 0 4 には、それぞれ、`node 1`, `node 2`, `node 3` 及び `node 4` というノード名が付与されているものとする。
- [0045] 例えば、図 3 に例示するファイル「`/home/a.txt`」は、`node 1` 及び `node 2` に格納されていることを示す。ここで、ファイル「`/home/a.txt`」が格納されているノードを知りたい場合、データ格納場所管理手段 3 1 は、「ファイル名」カラムから「`/home/a.txt`」の行を検索し、検索したデータの「ノード」カラムから対応するノードを特定すればよい。
- [0046] なお、データ格納場所管理情報記録装置 3 2 がデータベースソフトウェアで定められたテーブル形式でデータの格納先を記憶している場合、データ格納場所管理手段 3 1 は、そのテーブルに対する追加処理、更新処理を行うことで、データの格納先の更新を行ってもよい。例えば、ファイルの格納場所を変更したい場合、データ格納場所管理手段 3 1 は、実データの格納場所をコピーして移行する処理に併せ、テーブル内の対応するレコードを書き換えればよい。さらに、新たなファイルを格納する場合、データ格納場所管理手段 3 1 は、新たなデータの格納場所を示すレコードをテーブルに追加すればよい。
- [0047] 上記説明では、データの格納先を、ファイル名とノード名との組合せでデ

一タ格納場所管理情報記録装置 3 2 に記憶する場合について説明した。ただし、データ格納場所管理情報記録装置 3 2 が記憶する形式は、上記形式に限定されない。データ格納場所管理情報記録装置 3 2 は、例えば、ノード名リスト情報とハッシュ関数とを記憶していてもよい。この場合、データ格納場所管理手段 3 1 は、コンシステントハッシング法で算出した格納ノード先をデータの読み取り先および／または書き込み先と決定してもよい。

[0048] 一方、指定されたデータに対応する分散ノードの情報がデータ格納場所管理情報記録装置 3 2 に記憶されていない場合、後述のデータ暫定確率生成手段 3 7 の処理結果をもとに、データの読み取り先および／または書き込み先を決定する。なお、本処理については後述する。

[0049] アクセス履歴記録手段 5 1 は、アクセス受付手段 3 9 が受け付けたアクセス要求命令の内容を履歴として履歴記憶装置 5 0 に記憶させる。履歴記憶装置 5 0 は、アクセス履歴記録手段 5 1 の指示に応じ、アクセス要求命令の履歴を記憶する。履歴記憶装置 5 0 は、例えば、磁気ディスク等の記録媒体により実現される。

[0050] アクセス履歴記録手段 5 1 は、アクセス要求命令の履歴を、例えば、ジャーナルログとして履歴記憶装置 5 0 に記憶させてもよく、図 4 に例示するテーブル形式で履歴記憶装置 5 0 に記憶させてもよい。図 4 は、アクセス履歴記録手段 5 1 が記憶するアクセス要求命令の履歴の一例を示す説明図である。図 4 に示す例では、実行命令が読み取り指示か書き込み指示か（図 4 における「Read or Write」）、どのファイルに対して指示が行われたか（図 4 における「dest（ファイル名）」、及び、その指示がいつ行われたか（図 4 における「time」）を対応付けてアクセス要求命令の履歴を記憶していることを示す。

[0051] なお、アクセス履歴記録手段 5 1 は、アクセス受付手段 3 9 が受け付けたアクセス要求命令の全ての履歴を履歴記憶装置 5 0 に記憶させなくてもよい。アクセス履歴記録手段 5 1 は、アクセス要求命令の一部をサンプリングして履歴記憶装置 5 0 に記憶させてもよい。

- [0052] アクセス確率記録装置 35 は、各分散ノードに記憶されたデータの読み取りおよび／または書き込みが行われる確率を表すアクセス確率を記憶する。アクセス確率記録装置 35 には、後述のアクセス確率算出記録手段 34 又はデータ暫定確率生成手段 37 が算出したアクセス確率が記憶される。アクセス確率記録装置 35 は、例えば、磁気ディスク等の任意の記録媒体により実現される。
- [0053] アクセス確率算出記録手段 34 は、データに対するアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記録装置 35 に記憶させる。具体的には、アクセス確率算出記録手段 34 は、履歴記憶装置 50 が記憶するアクセス要求命令の履歴を参照し、同一ファイル名に対するアクセス履歴の情報をもとにアクセス確率を算出する。アクセス確率は、アクセス履歴に基づいて算出されることから、各ファイルがアクセスされる度合いを表す情報であるとも言える。
- [0054] アクセス確率算出記録手段 34 は、例えば、データが記憶されてからの期間でそのデータのアクセス回数を割ることにより、アクセス確率を算出してもよい。例えば、アクセス確率算出記録手段 34 が、ファイル「／home／a. txt」が生成された時点からの経過日数を分母とし、0時から12時までの間に発生したREADアクセス回数を分子として計算することにより、0時から12時までの間にREADアクセスが発生するアクセス確率を算出することができる。
- [0055] 図5は、アクセス確率の例を示す説明図である。図5に示す例では、アクセス確率記録装置 35 が、各ファイル名によって特定されるデータに対し、時間帯別のアクセス確率と、日付別のアクセス確率とを対応付けて記憶していることを示す。図5に示す例では、0時から12時の間に、ファイル「／home／a. txt」に対して、READアクセスが発生する確率が10%であり、Writeアクセスが発生する確率が0%であることを示す。例えば、10日に1度の割合で午前3時にREADアクセスが起きるようなファイルの場合、アクセス確率は、上記確率になる。

- [0056] また、図5に示す例では、ファイル「／home／a．txt」に対して、月の月上旬にREADアクセスが発生する確率が100%であり、月の中旬及び下旬にREADアクセスが発生する確率は0%である。このことから、ファイル「／home／a．txt」は、月の月上旬（例えば、1日～10日）に、3．3日に1度程度の割合でREADアクセスが起きるファイルであることがわかる。
- [0057] なお、図5に示す例では、アクセス確率記録装置35が月の月上旬、中旬及び下旬の期間でアクセス確率を記憶する場合について説明した。ただし、アクセス確率を記憶させる期間は、上記期間に限定されない。例えば、アクセス確率算出記録手段34が、1年を4つの期間（例えば、1Q、2Q、3Q、4Qなどの期間）に分けてアクセス確率を算出し、アクセス確率記録装置35にそれぞれの期間ごとのアクセス確率を記憶させてもよい。また、アクセス確率算出記録手段34が、月ごと（すなわち、1月～12月の期間）にアクセス確率を算出し、月ごとのアクセス確率をアクセス確率記録装置35に記憶させてもよい。
- [0058] また、アクセス確率算出記録手段34は、所定の期間内にデータアクセスが行われた回数（以下、アクセス回数と記す。）を示す情報を、アクセス確率記録装置35に記憶させてもよい。アクセス確率算出記録手段34は、例えば、単位時間以内のアクセス回数が既定値以上になった場合に、そのアクセス回数をアクセス確率記録装置35に記憶させてもよい。
- [0059] また、アクセス確率算出記録手段34は、アクセス確率とアクセス回数とにより算出される情報（以下、評価値と記す。）をアクセス確率記録装置35に記憶させてもよい。例えば、アクセス確率算出記録手段34は、アクセス確率にアクセス回数を乗じた値を評価値として算出し、算出した評価値をアクセス確率記録装置35に記憶させてもよい。
- [0060] 以下、具体例を用いて、評価値の算出方法について説明する。例えば、ある規定時間帯のアクセス確率が10%であるとする。このとき、平均アクセス回数100回のファイルと、1回のファイルではアクセス特性が異なる。

そこで、アクセス確率算出記録手段 34 は、平均アクセス回数 100 回のファイルの評価値を、 $1 / 10 \times 100 = 10$ と算出し、平均アクセス回数 1 回のファイルの評価値を、 $1 / 10 \times 1 = 0.1$ と算出する。そして、アクセス確率算出記録手段 34 は、算出した評価値を対象のファイルと対応付けてアクセス確率記録装置 35 に記憶させる。

[0061] このように評価値を算出することにより、アクセス確率が同一であっても、アクセス特性が異なるファイルを区別することが可能になる。なお、評価値の算出方法は、上記方法に限定されない。以上のことから、評価値とは、データへのアクセス特性を示す情報であると言える。

[0062] また、データ格納場所管理情報記録装置 32 の場合と同様、アクセス確率記録装置 35 は、データベースソフトウェアで定められたテーブル形式でデータの格納先を記憶していてもよい。この場合、アクセス確率算出記録手段 34 は、そのテーブルに対する検索処理や追加処理、更新処理を行うことで、アクセス確率の検索や更新を行ってもよい。

[0063] アクセス確率取得手段 33 は、各データにおけるアクセス確率の取得要求をデータ暫定確率生成手段 37 から受け付ける。そして、アクセス確率取得手段 33 は、指定されたデータに対応するアクセス確率をアクセス確率記録装置 35 から検索及び取得し、取得したアクセス確率をデータ暫定確率生成手段 37 に送信する。

[0064] データ暫定確率生成手段 37 は、指定されたデータに対応する分散ノードがデータ格納場所管理情報記録装置 32 に記憶されていない場合、分散ノードの記憶装置 12 に対して新規に格納されるデータの暫定的なアクセス確率を算出する。以下、暫定的なアクセス確率を算出することを、アクセス確率を推定すると記すこともある。

[0065] 以下、暫定的なアクセス確率の算出方法について具体的に説明する。データ a がアクセスされてから s 秒後にデータ b がアクセスされる共起確率を $P_s(b : a)$ としたとき、共起確率 $P_s(b : a)$ は、以下の式 1 を用いて算出される。なお、データ b がアクセスされた回数は、履歴記憶装置 50 に

記憶されたアクセス履歴から抽出可能である。

[0066] $P_s(b:a)$ = データ a がアクセスされた s 秒後にデータ b がアクセスされた回数 / データ a がアクセスされた総回数 (式 1)

[0067] よって、データ a がアクセスされてから t 秒後までにデータ b がアクセスされる確率 $P_t(b:a)$ は、 $\sum P_s(b:a)$ になる。

[0068] さらに、データ暫定確率生成手段 37 は、集合 W の中から選択されたデータ集合 $A\{a, b, \dots, n\}$ が t 秒後までにアクセスされる確率を $P_t(A:W)$ としたとき、確率 $P_t(A:W)$ を、以下の式 2 を用いて算出してもよい。

[0069] $P_t(A:W) = P_t(a:W) + P_t(b:W) + \dots + P_t(n:W)$ (式 2)

[0070] ここで、集合 W を、システム全体において今までのアクセスデータ全てを対象としてもよく、システム全体において直近の x 秒以内のアクセスデータ全てを対象としてもよい。さらに、あるノードに対するアクセスデータを集合 W としてもよく、相関マイニングで共起関係にあるデータのみを集合 W (それ以外のデータの確率は 0 にする、など) としてもよい。なお、集合 W は、データ量や、確率の計算に要する時間などを考慮し、予め定めておけばよい。

[0071] 以上のように、データ暫定確率生成手段 37 が、データへのアクセス履歴及びアクセス確率に基づいて共起確率を算出することにより、指定されたデータに対するアクセス確率を推定できる。データ暫定確率生成手段 37 は、このように算出した共起確率を暫定的なアクセス確率としてもよい。ただし、アクセス確率の算出方法は、上記方法に限定されない。

[0072] 例えば、データ暫定確率生成手段 37 は、新規に作成するファイルの属性に着目して暫定的なアクセス確率を算出してもよい。以下、具体例を用いて、暫定的なアクセス確率を算出する方法について説明する。ここでは、「 $\text{home}/x.\text{log}$ 」という名前のファイルを新規に作成し、ファイル「 $\text{home}/x.\text{log}$ 」の将来的なアクセス確率として、暫定的なアクセ

ス確率を算出する場合について説明する。

[0073] まず、ファイル「／home／x．log」の拡張子「log」に着目し、データ暫定確率生成手段37は、アクセス確率取得手段33に対し、拡張子が「log」であるファイルのアクセス確率をアクセス確率記録装置35から全て取得するように依頼する。データ暫定確率生成手段37は、得られたlogファイル全てのアクセス確率を平均し、平均した結果をファイル「／home／x．log」の暫定的なアクセス確率とする。そして、データ暫定確率生成手段37は、暫定的なアクセス確率をアクセス確率記録装置35に記憶させる。

[0074] なお、上記説明では、拡張子に基づいて暫定的なアクセス確率を算出する方法について説明した。さらに、データ暫定確率生成手段37は、拡張子だけでなく、ディレクトリパス（例えば、「／home」）を考慮に入れてもよい。この場合、データ暫定確率生成手段37は、アクセス確率取得手段33に対し、拡張子が「log」であり、かつ、ディレクトリパス「／home」に格納されるファイルのアクセス確率をアクセス確率記録装置35から全て取得するように依頼してもよい。この場合、データ暫定確率生成手段37は、得られたアクセス確率の平均をファイル「／home／x．log」の暫定的なアクセス確率としてもよい。

[0075] また、データ暫定確率生成手段37は、暫定的なアクセス確率を算出する際、ファイル名やファイルの格納場所ではなく、そのファイルに対するデータアクセスを行った（すなわち、ファイルの呼び出し元の）アプリケーションを考慮に入れてもよい。この場合、データ暫定確率生成手段37は、アクセス確率取得手段33に対し、データアクセスを行ったアプリケーション名が同一のアクセス確率をアクセス確率記録装置35から全て取得するように依頼し、得られたアクセス確率の平均を暫定的なアクセス確率としてもよい。

[0076] 以下、ファイルにアクセスするアプリケーションを考慮して暫定的なアクセス確率を算出する場合について詳細に説明する。図6は、アプリケーショ

ンXがデータアクセスする動作の例を示す説明図である。図6において丸で囲まれたA、B及びCは、それぞれアプリケーションXからアクセスされるデータを表す。また、図の左から右方向に向かって、アプリケーションXがデータにアクセスする時間が経過していることを示す。

[0077] 図6に示すように、あるアプリケーションからアクセスされるデータは、時間的な相関性があると想定される。図6に示す例では、データAがアクセスされた後にデータBがアクセスされ、さらに、データA及びデータBがアクセスされた後にデータCがアクセスされるという相関性である。

[0078] このような相関性から、同一のアプリケーションからアクセスされるデータは、同時にアクセスされる確率が高いと判断できる。そこで、データ暫定確率生成手段37は、新たなデータに対し、アクセスしたアプリケーション名が同一のアクセス確率をもとに暫定的なアクセス確率を算出する。なお、暫定的なアクセス確率の算出方法は、例えば、共起確率を用いてアクセス確率を算出する方法を用いてもよい。

[0079] なお、ファイルに対するデータアクセスを行ったアプリケーションに基づいたアクセス確率をデータ暫定確率生成手段37が取得するために、アクセス確率記録装置35は、ファイル名とともに、そのファイルに対するデータアクセスを行ったアプリケーション名を記憶しておけばよい。

[0080] 具体的には、データアクセス要求手段401からのデータアクセス要求にアプリケーション名を付与しておく。そして、アクセス受付手段39が、そのデータアクセス要求を受け取ると、アクセス履歴記録手段51が、アクセス受付手段39が受け付けたアクセス要求命令の内容にアプリケーション名を付加した履歴を履歴記憶装置50に記憶させる。そして、アクセス確率算出記録手段34が、アクセス要求命令の履歴を用いてファイルのアクセス確率をアプリケーションごとに算出し、アクセス確率記録装置35に記憶させる。なお、履歴記憶装置50及びアクセス確率記録装置35は、ファイル名とアプリケーション名とを対応付けて記憶すればよい。

[0081] また、データ暫定確率生成手段37は、アクセス確率の代わりに評価値を

用いて、指定されたデータのアクセス確率を推定してもよい。なお、アクセス確率を評価値を用いて推定する方法は、アクセス確率記録装置 35 に記憶されたアクセス確率を用いて推定する方法と同様の方法を用いればよい。

[0082] このように、データ暫定確率生成手段 37 は、指定されたデータに対するアクセス確率を、アクセス確率記録装置 35 に記憶されたアクセス確率に基づいて推定する。そして、データ暫定確率生成手段 37 は、推定したアクセス確率をアクセス確率記録装置 35 に記憶させる。

[0083] ここで、データ格納場所管理手段 31 は、データ暫定確率生成手段 37 が指定されたデータに対する暫定的なアクセス確率を算出すると、推定されたアクセス確率をもとにデータの書き込み先を決定する。データ格納場所管理手段 31 は、アクセス確率記録装置 35 に記録されたアクセス確率が、暫定的なアクセス確率と所定の値の範囲内にあるデータが格納されている分散ノードを、データの書き込み先と決定してもよい。

[0084] 具体的には、データ格納場所管理手段 31 は、暫定的なアクセス確率に類似したデータ（例えば、アクセス確率の差が所定の範囲内のデータ）をアクセス確率記録装置 35 から抽出する。次に、データ格納場所管理手段 31 は、特定したデータが格納されている分散ノードをデータ格納場所管理情報記録装置 32 から読み取る。そして、データ格納場所管理手段 31 は、読み取った分散ノードをデータの格納場所（以下、配置場所と記すこともある。）として決定する。

[0085] また、データ格納場所管理手段 31 は、推定されたアクセス確率が最も高い、もしくは、所定のアクセス確率を超える時間帯に稼働している分散ノードをデータの書き込み先として決定してもよい。なお、稼働しているか否かについては、後述するスケジュール情報に基づいて判断される。

[0086] なお、以下の説明では、ファイルアクセス要求命令が、ファイル名と実行命令とを含む場合について説明する。ただし、データアクセス要求手段 401 は、データの暫定的なアクセス確率を算出するアルゴリズムや、アクセス確率として管理保持したい形式に応じて、アクセス要求命令に必要な情報を

付与すればよく、また、付与した情報に応じて命令形式を定義すればよい。
なお、読み取りおよび／または書き込みを行う対象（以下、アクセス対象と記す）のデータと、そのアクセス内容を認識できる情報を含んでいれば、データアクセス要求手段４０１は、通常のファイルアクセスプロトコルを用いてもよい。

[0087] ノード動作状態記録装置３８は、分散データ保持アクセス手段１が動作する計算機（すなわち分散ノード１０１～１０４）の動作状態を記憶する。図７は、ノードの動作状態の例を示す説明図である。ここで、ノードが正常稼働している状態を「active」と記す。また、ノードの低速状態（例えば、ディスクの回転数を落としている状態）を「low speed」と記す。また、ノードが停止している状態を「standby」と記す。図７に示す例では、node 1, 4が正常稼働、node 2が低速状態、node 3が停止状態であることを示す。

[0088] ノード動作状態記録装置３８が記憶する動作状態は、後述の制御指示命令手段３６がノードの状態を変更する処理を行った際に更新される。なお、システム稼働開始の時点では、データアクセス場所選択システムに接続される全分散ノードを、例えば、停止状態として予めノード動作状態記録装置３８に記憶させておけばよい。なお、ノード動作状態記録装置３８はデータベースソフトウェアで定められたテーブル形式でノードの動作状態を記憶してもよい。

[0089] 制御指示命令手段３６は、データ格納場所管理情報記録装置３２と、アクセス確率記録装置３５に格納されている情報とを用いて、分散データ保持アクセス手段１に制御命令を発行する。ここで、制御命令とは、例えば、ディスクなどの計算機内の構成要素の動作状態を変更すること（具体的には、ディスクの回転数を下げたり、CPUのクロック数を下げたり、停止させたりすること）や、分散ノードの計算機全体を停止することなどが挙げられる。

[0090] 例えば、制御指示命令手段３６は、ノード動作状態記録装置３８の該当する動作状態を書き換えた後、各分散ノードに対して動作状態を変更する制御

指示を行う。なお、各分散ノードがノード動作状態記録装置 38 を備えていてもよい。この場合、制御指示命令手段 36 が、分散ノードの状態を変更する前に、情報管理制御手段 301 が備えるノード動作状態記録装置 38 の該当する動作状態を書き換え、書き換えた動作状態を他の分散ノード 101 ~ 104 に通知する。各分散ノード 101 ~ 104 は、動作状態の更新後、制御指示命令手段 36 に対して応答通知を発行する。制御指示命令手段 36 は、全ての分散ノードから応答通知を受けたあと、動作状態を変更する制御指示を行う。

[0091] このように、制御指示命令手段 36 は、分散ノード制御手段 14 に対して、分散データ保持アクセス手段 1 における被制御要素の動作状態を制御する指示を行う。なお、制御指示の内容は、被制御要素に応じて定義しておけばよい。

[0092] ここで、制御指示命令手段 36 は、以下に例示する方法を用いてノードの制御を行う。なお、以下の制御方法は例示であり、他の方法を用いてノードの制御を行ってもよい。

[0093] (1) アクセスの無いことが予想されるノードを停止する。

これは、分散ノードが保持しているデータ全体に対するアクセス確率を一定期間ごとに算出し、アクセス確率が閾値以下のノードを停止させる方法である。

[0094] この方法では、まず、アクセス確率算出記録手段 34 が、定期的に一定期間のアクセス履歴を履歴記憶装置 50 から読み取ってアクセス確率を算出し、算出結果をアクセス確率記録装置 35 に記憶させる。制御指示命令手段 36 は、データ格納場所管理情報記録装置 32 及びアクセス確率記録装置 35 を参照し、各分散ノードが保持するデータのアクセス確率の和を算出する。そして、制御指示命令手段 36 は、算出したアクセス確率の和が閾値以下の分散ノードに対して停止指示を行う。

[0095] なお、制御指示命令手段 36 は、対象の分散ノードに対し、すぐに停止指示を行ってもよく、所定期間経過後に停止指示を行ってもよい。所定期間経

過後に停止指示を行う場合、例えば、制御指示命令手段 36 は、停止スケジュールを決定し、そのスケジュールに従って分散ノードに対する停止指示を行ってもよい。このとき、制御指示命令手段 36 は、停止スケジュールを情報管理制御装置 300 が備える任意の記憶装置（図示せず。以下、スケジュール記憶手段と記す。）に記憶させておいてもよい。なお、スケジュール記憶手段に記憶させる情報は、停止スケジュールに限定されず、分散ノードを起動させるスケジュールであってもよい。

[0096] 上記説明では、一定期間のアクセス確率をもとにデータアクセスが無いことが予想される分散ノードを特定し、その分散ノードを停止させる場合について説明した。なお、分散ノードを停止可能な状態として、以下のようなケースが挙げられる。

[0097] 例えば、READ アクセスされるデータが、複数の分散ノードに保持されている場合、制御指示命令手段 36 は、これらの分散ノードのうち、少なくとも 1 つの分散ノードを起動するように制御すればよい。このように、起動させる分散ノードを減少させることで、更なる省電力化が期待できる。

[0098] また、アクセス先の分散ノードに Write アクセスを行わずに一旦別の分散ノードに Write アクセスを行い、その後、アクセス先の分散ノードに所定のタイミング（例えば、アクセス先の分散ノードが起動した後）でデータを移動させるとする。このような制御が行われる場合、制御指示命令手段 36 は、Write アクセスを行う分散ノードを停止する制御を行ってもよい。このように、起動させる分散ノードを減少させることで、更なる省電力化が期待できる。

[0099] （２）あらかじめスケジュールリングしておく。

これは、あらかじめ各分散ノードの起動及び停止の時間帯を決めておく方法である。例えば、月曜日から金曜日には全てのノードを稼働させ、週末には半数ノードを停止させる、といった制御である。

[0100] このような制御を行うことにより、総電力利用量を抑制し、例えば、週末の総電力利用量が一定量を超えないようにすることが可能になる。この方法

を用いる場合、制御指示命令手段 3 6 は、決められたスケジュールに従って、分散ノードへの起動命令及び停止命令を発行する。

[0101] 具体的には、制御指示命令手段 3 6 は、期間ごとに算出したアクセス確率の和が閾値以下の分散ノードを特定した場合、スケジュール記憶手段（図示せず）に対して特定した分散ノードの停止スケジュールを設定する。そして、制御指示命令手段 3 6 は、期間ごとに設定したスケジュールに従って、分散ノードの停止及び起動を行う。なお、設定した停止期間内にアクセス確率の高いデータが存在する場合、制御指示命令手段 3 6 は、停止していない分散ノードにそのデータを転送する制御を行ってもよい。このように、アクセス確率をもとに決定されたスケジュールをスケジュール情報と記す。

[0102] さらに、制御指示命令手段 3 6 は、このスケジュール情報（すなわち、期間毎に起動する、停止するなどの情報）に基づいて、データの配置状態及びアクセス確率に応じて定期的にデータの再配置を行ってもよい。制御指示命令手段 3 6 は、例えば、スケジュール情報をもとに、各分散ノードで停止可能な時間帯の候補を抽出する。そして、制御指示命令手段 3 6 は、抽出した分散ノードにおいて、停止すると判断された時間帯にアクセス確率が所定の値を超えるデータを、別の分散ノードに再配置するよう、分散ノード制御手段 1 4 に対して制御指示を行う。このようにすることで、更なる省電力化を図ることが出来る。

[0103] また、制御指示命令手段 3 6 は、高いアクセス確率が集中する分散ノードのデータを別の分散ノードに移動するよう制御指示を行い、移動先の分散ノードをその時間帯に稼働するように制御してもよい。このようにした場合には、性能向上を図ることが可能になる。

[0104] データ格納場所管理手段 3 1 と、アクセス確率取得手段 3 3 と、アクセス確率算出記録手段 3 4 と、制御指示命令手段 3 6 と、データ暫定確率生成手段 3 7 と、アクセス受付手段 3 9 と、アクセス履歴記録手段 5 1 とは、プログラム（データアクセス場所選択プログラム）に従って動作するコンピュータの CPU によって実現される。例えば、プログラムは、情報管理制御装置

300の記憶部（図示せず）に記憶され、CPUは、そのプログラムを読み込み、プログラムに従って、データ格納場所管理手段31、アクセス確率取得手段33、アクセス確率算出記録手段34、制御指示命令手段36、データ暫定確率生成手段37、アクセス受付手段39、及び、アクセス履歴記録手段51として動作してもよい。

[0105] また、データ格納場所管理手段31と、アクセス確率取得手段33と、アクセス確率算出記録手段34と、制御指示命令手段36と、データ暫定確率生成手段37と、アクセス受付手段39と、アクセス履歴記録手段51とは、それぞれが専用のハードウェアで実現されていてもよい。

[0106] 次に、動作について説明する。図8は、本実施形態におけるデータアクセス処理の例を示すフローチャートである。データアクセス処理は、データアクセス要求手段401のアクセス先問い合わせ手段41が情報管理制御手段301のアクセス受付手段39に対して、アクセス要求命令を発行することで開始される。すなわち、まず、アクセス先問い合わせ手段41が、アクセス受付手段39に対して、アクセス要求命令を発行する。

[0107] アクセス受付手段39は、アクセス要求命令を受け付けると、データ格納場所管理手段31にアクセス対象データが格納されている分散ノードを問い合わせる（ステップS101）。例えば、データ格納場所管理情報記録装置32が図3に例示する内容を記憶している場合に、ファイル「／home／a．txt」のアクセス先（すなわち、格納ノードの情報）は、node1，node2になる。

[0108] 次に、データ格納場所管理手段31は、アクセス要求命令におけるアクセスの種類について判断する（ステップS102）。アクセスの種類がWrite処理の場合（ステップS102における「write」）、ステップS103へ進み、READ処理である場合（ステップS102における「Read」）、ステップS106に進む。なお、READ処理とは、データの読み取り処理のことを表し、Write処理とは、データに対する書き込み処理の他、データの削除処理、データの更新処理、ファイルのクリエイト処理な

ど、データの読み取り処理以外の処理のことを表す。

[0109] アクセスの種類がWrite処理の場合、データ格納場所管理手段31は、アクセス対象のデータが新規データか否かを判断する（ステップS103）。具体的には、データ格納場所管理手段31は、アクセス要求命令が新規データのWrite処理か、アクセス対象ファイルに対する更新処理かを判断する。以下の説明では、データ格納場所管理情報記録装置32にアクセス対象のファイル名が存在しない場合に新規データのWrite処理と判断され、ファイル名が存在する場合にアクセス対象ファイルに対する更新処理と判断されるものとする。

[0110] 新規データと判断された場合（ステップS103における「新規データ」）には、ステップS104に進み、更新処理と判断された場合（ステップS103における「更新」）にはステップS106へ進む。

[0111] 新規データであった場合、データ暫定確率生成手段37は、新規データのアクセス処理要求命令に含まれる情報を用いて、その新規データの将来的なアクセス確率を推定する（ステップS104）。例えば、アクセス処理要求命令が、「／home／e. txtというファイルを新規に作成する」という内容であった場合、データ暫定確率生成手段37は、ファイル「／home／e. txt」のアクセス確率を推定する。すなわち、データ暫定確率生成手段37は、暫定的なアクセス確率を算出する。なお、上記例では、「／home／e. txt」が、アクセス処理要求命令に含まれる情報に相当する。アクセス処理要求命令には、他にも、命令の発行元プログラム名やプロセス名、プログラムの実行ユーザ名などを含んでもよい。

[0112] 以下、アクセス処理要求命令が、「／home／e. txtというファイルを新規に作成する」という内容であった場合の動作について説明する。まず、アクセス確率取得手段33は、データ暫定確率生成手段37の指示に応じ、アクセス確率記録装置35から拡張子がtxtであるファイルのアクセス確率を全て取得する。次に、データ暫定確率生成手段37は、各アクセス確率の値をもとに平均値を算出し、その平均値をファイル「／home／e

「t x t」のアクセス確率として採用する。

[0113] そして、データ暫定確率生成手段 37 は、採用したアクセス確率をアクセス確率記録装置 35 に記憶させる。また、データ暫定確率生成手段 37 は、推定したアクセス確率をデータ格納場所管理手段 31 に通知する。

[0114] 次に、データ格納場所管理手段 31 は、通知された暫定的なアクセス確率からデータ配置場所を決定する（図 8 のステップ S105）。データ格納場所管理手段 31 は、例えば、暫定的なアクセス確率の値と類似したデータ（例えば、アクセス確率の差が一定値以内）と同一の分散ノードを選択し、その分散ノードをデータ配置場所と決定する。このような分散ノードを選択することにより、その分散ノードにデータアクセスを集中させることができる。このように、ある時間帯のアクセスを限られた分散ノードに集中させることによって、他の分散ノードを停止（または、動作状態を変更して省電力化）させることが可能になる。

[0115] さらに、データ格納場所管理手段 31 は、ノード動作状態記録装置 38 の情報を参照し、正常稼働している分散ノードの中から書き込み先の分散ノードを選択するようにしてもよい。また、例えば、データ配置場所として決定した分散ノードが正常稼働していない場合、データ格納場所管理手段 31 は、正常稼働している分散ノードを、一時的に書き込みを行う分散ノードと決定してもよい。なお、この場合、正常稼働していなかった分散ノードが正常稼働状態に戻ったときに、制御指示命令手段 36 が、一時的に書き込みが行われた分散ノードに対し、正常稼働状態に戻った分散ノードへデータを移動するように制御指示を行う。このような制御を行うことで、可用性を確保しつつ、不要な分散ノードの稼働を抑制することが出来る。

[0116] また、データ格納場所管理手段 31 は、スケジュール記憶手段（図示せず）に記憶されたスケジュール情報に基づいてデータの配置場所を決定してもよい。具体的には、算出したデータのアクセス確率のうち、そのデータへのアクセスがある時間帯に正常稼働（すなわち、active）のノードをスケジュール記憶手段から読み取り、そのノードをデータの配置場所と決定し

てもよい。なお、起動中（正常稼働中）の分散ノードが複数ある場合、データ格納場所管理手段 31 は、予め定めたロジックに基づいて、1 つの分散ノードを決定すればよい。

[0117] 以下、スケジュール情報に基づいてデータの配置場所を決定する方法について、具体例を用いて説明する。例えば、3 つの分散ノード `node 1`, `node 2`, `node 3` の稼働スケジュールが、`node 1` については月曜、水曜、金曜に稼働、`node 2` については火曜、木曜、土曜に稼働、`node 3` については日曜に稼働、と設定されているとする。ここで、データ A が新規に作成され、その暫定的なアクセス確率によると、月曜日に集中してアクセスされるデータ（すなわち、月曜日のアクセス確率が突出して高いデータ）であることが分かったとする。この場合、データ格納場所管理手段 31 は、`node 1` をデータ A の配置場所と決定する。

[0118] スケジュール情報に基づいてデータの配置場所を決定する方法について、さらに、他の具体例を用いて説明する。新規に作成されたデータ A の暫定的なアクセス確率を算出した結果、平日の昼間のアクセス確率が高く、日曜日にまれにアクセスされるデータであることが分かったとする。また、このときの稼働スケジュールが、平日の昼間には `node 1` ~ 100 の分散ノードが、平日の夜間には `node 1` ~ 10 の分散ノードが、土日休日には `node 11` ~ 15 の分散ノードが稼働する、と設定されていたとする。この場合、データ格納場所管理手段 31 は、`node 11` ~ `node 15` をデータ A の配置場所と決定する。なお、土日休日にアクセスされるデータが急増された場合、例えば、制御指示命令手段 36 が、別途スケジュール情報を「土日休日には `node 11` ~ 20 の分散ノードが稼働する」と変更してもよい。

[0119] 上記説明では、アクセス確率が類似するファイルが格納された分散ノードの中から、データ配置場所を決定する場合について説明した。ただし、データの配置場所を決定する方法は、実装するシステムやファイルへのアクセス状況に応じて、種々の変更が可能である。

[0120] 例えば、データの書き込みを複数の分散ノードに行うシステムにおいて、

最初に書き込みがされた後、読み取りばかりされるファイルが存在するとする。この場合、データ格納場所管理手段 31 は、書き込み先の分散ノードのうち、少なくとも 1 つの分散ノードにアクセスを集中させるように分散ノードを決定してもよい。アクセスを集中させた分散ノードさえ稼働させておけば、他の書き込み先ノードを停止させた場合であっても、対象のデータに対して読み取り処理を行うことが可能だからである。

[0121] なお、データ格納場所管理手段 31 は、同一のアクセスパターンのファイルを別のノードに格納するように決定し、分散ストレージシステムの性能を確保するようにしてもよい。

[0122] ステップ S105 の後はステップ S107 へ進む。

[0123] ステップ S102 においてアクセスの種類が READ 処理であった場合、又は、ステップ S103 においてアクセス要求命令が更新処理であった場合、データ格納場所管理手段 31 は、アクセス先の分散ノードを選択する（ステップ S106）。

[0124] 具体的には、ステップ S102 においてアクセスの種類が READ 処理であった場合、データ格納場所管理手段 31 は、ステップ S101 で得られた（すなわち、データ格納場所管理情報記録装置 32 から読み取った）分散ノードのうちのいずれか 1 つを選択する。例えば、図 3 に例示するファイル「／home／a. txt」の場合、READ 先の候補は node1 又は node2 になる。データ格納場所管理手段 31 は、例えば、ラウンドロビン法を用いて分散ノードを選択してもよい。他にも、データ格納場所管理手段 31 は、ノード動作状態記録装置 38 に記憶された分散ノードの稼働状態をもとに、正常動作している分散ノードの中から 1 つを任意の方法で選択してもよい。ただし、分散ノードを選択する方法は、上記方法に限定されない。

[0125] 一方、ステップ S103 においてアクセス要求命令が更新処理であった場合、データ格納場所管理手段 31 は、書き込み対象の分散ノードを全て選択する。なお、データ格納場所管理手段 31 は、ノード動作状態記録装置 38 の情報を参照し、対象の分散ノードが停止している（もしくは、省電力動作

状態である) 場合、正常稼働している分散ノードを、一時的に書き込みを行う分散ノードとして選択してもよい。この場合、正常稼働していなかった分散ノードが正常稼働状態に戻ったときに、制御指示命令手段 36 が、一時的に書き込みが行われた分散ノードに対し、正常稼働状態に戻った分散ノードへデータを移動するように制御指示を行う。

[0126] また、ステップ S 106 において、書き込み先として選択された分散ノードが停止している場合、制御指示命令手段 36 が、その分散ノードに対して起動命令を発行してもよい。

[0127] ステップ S 105 又はステップ S 106 の後、アクセス受付手段 39 は、データアクセスを行う対象の分散ノード情報をデータアクセス要求手段 401 に応答し、アクセス履歴記録手段 51 にアクセス履歴記録処理を依頼する (ステップ S 107)。最後に、データアクセス要求手段 401 がアクセス先の分散ノード情報を受け取ると、データアクセス手段 42 は、指定された分散ノードの分散データ保持アクセス手段 1 にアクセス要求命令を発行する (ステップ S 108)。データアクセス処理手段 11 は、受け取ったアクセス要求命令に基づいて処理を行い、処理結果をデータアクセス要求手段 401 に送信する。

[0128] また、上記説明では、新規データに対するアクセス要求のタイミングで、データ暫定確率生成手段 37 が暫定的なアクセス確率を算出し、算出したアクセス確率をアクセス確率記録装置 35 に記憶させる場合について説明した。一方、アクセス確率算出記録手段 34 は、任意の周期でアクセス確率を計算し、計算したアクセス確率をアクセス確率記録装置 35 に記憶させてもよい。

[0129] 例えば、アクセス確率算出記録手段 34 は、データを再配置する処理が行われるタイミングや、アクセス確率が読み取られるタイミングで、その都度アクセス確率を計算してもよい。さらに、アクセス確率算出記録手段 34 は、暫定的なアクセス確率についても、一定期間経過後、実際のアクセス履歴に基づいてアクセス確率を算出し、算出した値でアクセス確率記録装置 35

の内容を更新してもよい。

[0130] なお、複数の分散ノードにデータを格納する分散ストレージシステムでは、データの書き込み処理において、複製先として指定された複数の分散ノードに対して同時に書き込み処理をしなくてもよい場合がある。例えば、アクセス集中時に備え、6つの分散ノードに複製データを格納しなければならない場合であっても、システム可用性を確保するため、最初の時点では、2つの分散ノードにデータを格納すればよい場合がある。

[0131] この場合、最初にデータアクセス要求命令を受け取ったときには、データ格納場所管理手段31は、最小限の2つの分散ノードを、書き込み処理を行う分散ノードとしてデータアクセス要求手段401に応答してもよい。そして、データ格納場所管理手段31は、アクセス確率記録装置35に記録されている情報に基づいてデータアクセスが集中する時間帯を判断し、その時間までに、残りの4つの分散ノードを、書き込み処理を行う分散ノードとしてデータアクセス要求手段401に応答してもよい。

[0132] 以上の説明では、情報管理制御装置300が複数の分散ノードを101～104を管理する場合について説明した。すなわち、情報管理制御装置300が情報管理制御手段301を備え、分散ノード101が分散データ保持アクセス手段1を備える場合について説明した。このような構成ではなく、分散ノード101が、分散データ保持アクセス手段1と情報管理制御手段301の両方備えていてもよい。この場合、複数の分散ノード101～104が協調して動作することにより、上述の実施形態のように動作してもよい。例えば、それぞれの分散ノード内に備える記憶装置内のデータについてアクセス確率を算出し、そのアクセス確率をもとに被制御要素の制御を行えばよい。

[0133] 以上のように、本実施形態によれば、データ格納場所管理手段31が、データ格納場所管理情報記録装置32に記憶されているデータに対応する分散ノードの中から指定されたデータの読み取り先および／または書き込み先を決定する。そして、アクセス確率算出記録手段34が、履歴記憶装置50に

記憶されたアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記録装置 35 に記憶させる。一方、指定されたデータに対応するノードがデータ格納場所管理情報記録装置 32 に記憶されていない場合、データ暫定確率生成手段 37 が、指定されたデータに対するアクセス確率をアクセス確率記録装置 35 に記憶されたアクセス確率に基づいて推定する。この場合、データ格納場所管理手段 31 が、推定されたアクセス確率をもとにそのデータの書き込み先を決定する。

[0134] そのため、本発明による分散ストレージシステムでは、新たなデータが追加されるような場合であっても、データアクセスに係る電力消費を抑制できる。

[0135] 具体的には、分散ストレージシステムにおいて、データ格納場所管理手段 31 が、データアクセス先を将来のアクセス確率（すなわち、暫定的なアクセス確率）に基づいて決定する。よって、新規に作成されたデータや、そのレプリカの物理的な割り当て場所を、初期割り当ての時点において決定できるため、再配置などのデータ移動などの処理量を削減でき、省電力化が可能になる。

[0136] また、制御指示命令手段 36 が、暫定的なアクセス確率に基づいて分散システムを構成する分散ノードの動作状態をデータの登録前に事前に変更出来るため、効率的な省電力制御が可能になる。

[0137] さらに、制御指示命令手段 36 が、アクセス確率に応じて分散ノードのデータを移動させるため、システムの省電力稼働状態を維持するように分散ノードを選択させることが可能になる。

[0138] また、アクセス履歴記録手段 51 が、データごとのアクセス履歴として、データの識別名や利用内容を記録する。そのため、同一サービス内もしくは同一ユーザ内でアクセスパターンが異なるデータに対しても省電力効果を提供できる。

[0139] 例えば、特許文献 1 に記載された方法や、特許文献 2 に記載されたシステムでは、新規データ及びそのレプリカの物理的な割り当て場所を、データ格

納の最初の時点で決定することができない。過去のアクセス履歴データをもとにアクセス確率を算出する方法では、そのデータが将来的にどのように使われるかを予想できず、データの初期配置場所を決めるための情報が不足しているからである。そのため、結果的に、登録後にアクセス確率を算出後、データコピーを行うことになる。このような状況でデータコピーが大量に発生すると、ネットワークやバスの I/O 帯域を大量に利用してしまうため、分散システム上で動作するアプリケーションの性能に悪影響を与えてしまう。しかし、本実施形態によれば、新規データに対する暫定的なアクセス確率を算出して適切な位置にデータを配置する。よって、データの移動コストが抑制されるため、結果として、消費電力を抑制することが可能になる。

[0140] また、例えば、特許文献 2 に記載されたシステムでは、データのアクセス履歴やアクセス確率は、ボリュームを一定単位ごとに区切った物理配置場所のスライス単位で扱っている。このような場合、同一スライスに格納されたデータについては、再配置処理では分離できなくなる。すなわち、サービスやユーザごとにボリュームを割り当てた場合には、これらのサービス及びユーザごとにアクセスパターンが異なるデータごとに対応することは困難であった。しかし、本実施形態では、データ単位でアクセス確率を算出するため、データごとに柔軟な対応が可能になる。

[0141] 次に、本発明の最小構成について説明する。図 9 は、本発明によるデータアクセス場所選択システムの最小構成の例を示すブロック図である。また、図 10 は、本発明によるデータアクセス場所選択システムの他の最小構成の例を示すブロック図である。

[0142] 図 9 に例示するデータアクセス場所選択システムは、データを格納する少なくとも 1 つ以上のノード 80（例えば、分散ノード 101）と、データの読み取り先および／または書き込み先をノードの中から選択する管理サーバ 90（例えば、情報管理制御装置 300）とを備えている。

[0143] 管理サーバ 90 は、データの格納先をノード 80 と対応付けて記憶するデータ格納先記憶手段 91（例えば、データ格納場所管理情報記録装置 32）

と、指定されたデータの読み取り先および／または書き込み先を、データ格納先記憶手段 9 1 に記憶されているそのデータに対応するノードの中から決定するアクセス先決定手段 9 2（例えば、データ格納場所管理手段 3 1）と、各ノード 8 0 に対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段 9 3（アクセス確率記録装置 3 5）と、データへのアクセス履歴（例えば、履歴記憶装置 5 0 に記憶されたアクセス履歴）に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出手段 9 4（例えば、アクセス確率算出記録手段 3 4）と、指定されたデータに対するアクセス確率を、アクセス確率記憶手段 9 3 に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段 9 5（例えば、データ暫定確率生成手段 3 7）とを備えている。

[0144] アクセス確率推定手段 9 5 は、指定されたデータに対応するノードがデータ格納先記憶手段 9 1 に記憶されていない場合、そのデータのアクセス確率を推定する（例えば、拡張子が同一のデータのアクセス確率の平均値を算出する、ディレクトリが同一のデータのアクセス確率の平均値を算出する、アクセス先のアプリケーションが同一のデータのアクセス確率の平均値を算出する、もしくは、共起確率を算出するなど）。

[0145] アクセス先決定手段 9 2 は、指定されたデータに対応するノードがデータ格納先記憶手段 9 1 に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定する（例えば、アクセス確率が類似するデータを格納するノードをアクセス先と決定する、もしくは、所定のアクセス確率を超える期間内に稼働中のノードを選択する）。

[0146] また、図 1 0 に例示するデータアクセス場所選択システムは、複数のノード 8 0（例えば、分散ノード 1 0 1 ～ 1 0 4）を備えている。各ノード 8 0 は、データ格納先記憶手段 9 1 と、アクセス先決定手段 9 2 と、アクセス確率記憶手段 9 3 と、アクセス確率算出手段 9 4 と、アクセス確率推定手段 9 5 とを備えている。なお、データ格納先記憶手段 9 1、アクセス先決定手段

９２、アクセス確率記憶手段９３、アクセス確率算出手段９４及びアクセス確率推定手段９５の内容は、図９に示す内容と同様である。

[0147] そのような構成により、分散ストレージシステムにおいて、新たなデータが追加されるような場合であっても、データアクセスに係る電力消費を抑制できる。

[0148] なお、少なくとも以下に示すようなデータアクセス場所選択システムも、上記に示すいずれかの実施形態に開示されている。

[0149] (１) データを格納する少なくとも１つ以上のノード（例えば、分散ノード１０１）と、データの読み取り先および／または書き込み先をノードの中から選択する管理サーバ（例えば、情報管理制御装置３００）とを備え、管理サーバが、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段（例えば、データ格納場所管理情報記録装置３２）と、指定されたデータの読み取り先および／または書き込み先を、データ格納先記憶手段に記憶されているそのデータに対応するノードの中から決定するアクセス先決定手段（例えば、データ格納場所管理手段３１）と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段（アクセス確率記録装置３５）と、データへのアクセス履歴（例えば、履歴記憶装置５０に記憶されたアクセス履歴）に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出手段（例えば、アクセス確率算出記録手段３４）と、指定されたデータに対するアクセス確率を、アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段（例えば、データ暫定確率生成手段３７）とを備え、アクセス確率推定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、そのデータのアクセス確率を推定し、アクセス先決定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定する（例えば、アクセス確率が類似するデータを格納するノードをアクセス先と

決定する、もしくは、所定のアクセス確率を超える期間内に稼働中のノードを選択する) データアクセス場所選択システム。

[0150] (2) アクセス先決定手段は、アクセス確率記録手段に記録されたアクセス確率のうち、推定されたアクセス確率と所定の値の範囲内にあるアクセス確率のデータ(例えば、暫定的なアクセス確率に類似したデータ)を格納するノードを、データの書き込み先と決定するデータアクセス場所選択システム。

[0151] (3) 推定されたアクセス確率に基づいて、時間帯ごとの各ノードの動作スケジュール(例えば、停止スケジュール)を決定するノード動作スケジュール決定手段(例えば、制御指示命令手段36)を備え、アクセス先決定手段が、推定されたアクセス確率が所定の値を超える時間帯における動作スケジュールが稼働中であるノードをデータの読み取り先および/または書き込み先と決定するデータアクセス場所選択システム。

[0152] (4) ノードに対して制御指示命令を発行する制御指示命令手段(例えば、制御指示命令手段36)を備え、制御指示命令手段が、動作スケジュールに基づいてノードに対する制御指示命令(例えば、停止指示)を行うデータアクセス場所選択システム。

[0153] (5) アクセス確率推定手段が、指定されたデータと同一の拡張子(例えば、拡張子「.log」)を有するデータのアクセス確率をアクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定するデータアクセス場所選択システム。

[0154] (6) アクセス確率推定手段が、指定されたデータと同一のディレクトリ(例えば、ディレクトリパス「/home」)に格納されるデータのアクセス確率をアクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定するデータアクセス場所選択システム。

[0155] (7) アクセス確率推定手段が、データへのアクセス履歴及びアクセス確率記憶手段に記憶されたアクセス確率に基づいて共起確率を算出する(例えば

、（式１）を用いて算出する）ことにより、指定されたデータのアクセス確率を推定するデータアクセス場所選択システム。

[0156] （８）データを格納する複数のノード（例えば、分散ノード１０１～１０４）を備え、各ノードが、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段（例えば、データ格納場所管理情報記録装置３２）と、指定されたデータの読み取り先および／または書き込み先を、データ格納先記憶手段に記憶されているそのデータに対応するノードの中から決定するアクセス先決定手段（例えば、データ格納場所管理手段３１）と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段（アクセス確率記録装置３５）と、データへのアクセス履歴（例えば、履歴記憶装置５０に記憶されたアクセス履歴）に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出手段（例えば、アクセス確率算出記録手段３４）と、指定されたデータに対するアクセス確率を、アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段（例えば、データ暫定確率生成手段３７）とを備え、アクセス確率推定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、そのデータのアクセス確率を推定し、アクセス先決定手段が、指定されたデータに対応するノードがデータ格納先記憶手段に記憶されていない場合、推定されたアクセス確率をもとにそのデータの書き込み先を決定する（例えば、アクセス確率が類似するデータを格納するノードをアクセス先と決定する、もしくは、所定のアクセス確率を超える期間内に稼働中のノードを選択する）データアクセス場所選択システム。

[0157] 上記の実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。

[0158] （付記１）データを格納する少なくとも１つ以上のノードと、データの読み取り先および／または書き込み先を前記ノードの中から選択する管理サーバとを備え、前記管理サーバは、データの格納先をノードと対応付けて記憶す

るデータ格納先記憶手段と、指定されたデータの読み取り先および／または書き込み先を、前記データ格納先記憶手段に記憶されている当該データに対応するノードの中から決定するアクセス先決定手段と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、データへのアクセス履歴に基づいて前記アクセス確率を算出し、算出したアクセス確率を前記アクセス確率記憶手段に記憶させるアクセス確率算出手段と、指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、前記アクセス確率推定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定し、前記アクセス先決定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定することを特徴とするデータアクセス場所選択システム。

- [0159] （付記 2） アクセス先決定手段は、アクセス確率記録手段に記録されたアクセス確率のうち、推定されたアクセス確率と所定の値の範囲内にあるアクセス確率のデータを格納するノードを、データの書き込み先と決定する付記 1 記載のデータアクセス場所選択システム。
- [0160] （付記 3） 推定されたアクセス確率に基づいて、時間帯ごとの各ノードの動作スケジュールを決定するノード動作スケジュール決定手段を備え、アクセス先決定手段は、推定されたアクセス確率が所定の値を超える時間帯における前記動作スケジュールが稼働中であるノードをデータの読み取り先および／または書き込み先と決定する付記 1 または付記 2 記載のデータアクセス場所選択システム。
- [0161] （付記 4） ノードに対して制御指示命令を発行する制御指示命令手段を備え、前記制御指示命令手段は、動作スケジュールに基づいてノードに対する制御指示命令を行う付記 3 記載のデータアクセス場所選択システム。
- [0162] （付記 5） アクセス確率推定手段は、指定されたデータと同一の拡張子を有

するデータのアクセス確率を前記アクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定する付記 1 から付記 4 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0163] (付記 6) アクセス確率推定手段は、指定されたデータと同一のディレクトリに格納されるデータのアクセス確率を前記アクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定する付記 1 から付記 4 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0164] (付記 7) アクセス確率推定手段は、データへのアクセス履歴及び前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて共起確率を算出することにより、指定されたデータのアクセス確率を推定する付記 1 から付記 4 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0165] (付記 8) アクセス確率算出手段は、データへのアクセス元を示す情報を含むアクセス履歴に基づいてアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させ、アクセス確率推定手段は、指定されたデータのアクセス元が同一のデータのアクセス確率を前記アクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値をアクセス確率と推定する付記 1 から付記 7 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0166] (付記 9) ノードの動作状態を記憶するノード状態記憶手段を備え、アクセス先決定手段は、前記ノード状態記憶手段に記憶されたノードの動作状態が稼働中のノードの中からデータの読み取り先および／または書き込み先を決定する付記 1 から付記 8 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0167] (付記 10) データへのアクセスログを収集するアクセスログ収集手段を備え、アクセス確率算出手段は、前記アクセス収集手段が収集したアクセスログに基づいてアクセス確率を算出する付記 1 から付記 9 のうちのいずれか 1

つに記載のデータアクセス場所選択システム。

[0168] (付記 1 1) アクセスログ収集手段は、データへのアクセス内容を示す情報である付与情報を含めてアクセスログを収集し、アクセス確率算出手段は、アクセスログに基づいて前記付与情報ごとにアクセス確率を算出し、算出したアクセス確率をアクセス確率記憶手段に記憶させる付記 1 0 記載のデータアクセス場所選択システム。

[0169] (付記 1 2) アクセス確率算出手段は、データへのアクセス履歴に基づいてアクセス確率及び当該データへのアクセスの頻度を示す情報であるアクセス頻度を算出し、前記アクセス確率及びアクセス頻度を基に、データへのアクセス特性を示す情報である評価値を算出し、当該評価値をアクセス確率記憶手段に記憶させ、アクセス確率推定手段は、指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶された評価値に基づいて推定する付記 1 から付記 1 1 のうちのいずれか 1 つに記載のデータアクセス場所選択システム。

[0170] (付記 1 3) データを格納する複数のノードを備え、前記各ノードは、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段と、指定されたデータの読み取り先および／または書き込み先を、前記データ格納先記憶手段に記憶されている当該データに対応するノードの中から決定するアクセス先決定手段と、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、データへのアクセス履歴に基づいて前記アクセス確率を算出し、算出したアクセス確率を前記アクセス確率記憶手段に記憶させるアクセス確率算出手段と、指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、前記アクセス確率推定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定し、前記アクセス先決定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記

アクセス確率をもとに当該データの書き込み先を決定することを特徴とするデータアクセス場所選択システム。

[0171] (付記 1 4) データを格納する少なくとも 1 つ以上のノードの中から、データの読み取り先および／または書き込み先を選択する選択装置が、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されている当該データに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定し、前記選択装置が、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確率をアクセス確率記憶手段に記憶させ、前記選択装置が、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、指定されたデータに対するアクセス確率を前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定し、前記選択装置が、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定することを特徴とするデータアクセス場所選択方法。

[0172] (付記 1 5) アクセス確率記録手段に記録されたアクセス確率のうち、推定されたアクセス確率と所定の値の範囲内にあるアクセス確率のデータを格納するノードを、データの書き込み先と決定する付記 1 4 記載のデータアクセス場所選択方法。

[0173] (付記 1 6) データを格納する少なくとも 1 つ以上のノードの中から、データの読み取り先および／または書き込み先を選択するコンピュータに適用されるデータアクセス場所選択プログラムであって、前記コンピュータに、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されている当該データに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定するアクセス先決定処理、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確

率をアクセス確率記憶手段に記憶させるアクセス確率算出処理、および、指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定処理を実行させ、前記アクセス確率推定処理で、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定させ、前記アクセス先決定処理で、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定させるためのデータアクセス場所選択プログラム。

[0174] (付記 17) アクセス先決定処理で、アクセス確率記録手段に記録されたアクセス確率のうち、推定されたアクセス確率と所定の値の範囲内にあるアクセス確率のデータを格納するノードを、データの書き込み先と決定させる付記 16 記載のデータアクセス場所選択プログラム。

[0175] 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およ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。

[0176] この出願は、2010年3月29日に提出された日本特許出願2010-75514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。

産業上の利用可能性

[0177] 本発明は、分散ストレージシステムにおいて、データの読み取り先および／または書き込み先を選択するデータアクセス場所選択システムに好適に適用される。すなわち、本発明を、複数の計算機にデータを分散して格納する分散ストレージシステムに適用が可能であり、また、それらを利用できるITシステム全般にも適用が可能である。

符号の説明

[0178] 1 分散データ保持アクセス手段

- 1 1 データアクセス処理手段
- 1 2 記憶装置
- 1 3 被制御対象装置
- 1 4 分散ノード制御手段
- 3 1 データ格納場所管理手段
- 3 2 データ格納場所管理情報記録装置
- 3 3 アクセス確率取得手段
- 3 4 アクセス確率算出記録手段
- 3 5 アクセス確率記録装置
- 3 6 制御指示命令手段
- 3 7 データ暫定確率生成手段
- 3 8 ノード動作状態記録装置
- 3 9 アクセス受付手段
- 4 1 アクセス先問い合わせ手段
- 4 2 データアクセス手段
- 5 0 履歴記憶装置
- 5 1 アクセス履歴記録手段
- 1 0 1 ~ 1 0 4 分散ノード
- 1 1 1 ~ 1 1 4 記憶装置
- 2 0 0 通信ネットワーク
- 3 0 0 情報管理制御装置
- 3 0 1 情報管理制御手段
- 4 0 0 データアクセス要求ノード
- 4 0 1 データアクセス要求手段

請求の範囲

[請求項1]

データを格納する少なくとも1つ以上のノードと、
データの読み取り先および／または書き込み先を前記ノードの中から選択する管理サーバとを備え、
前記管理サーバは、
データの格納先をノードと対応付けて記憶するデータ格納先記憶手段と、
指定されたデータの読み取り先および／または書き込み先を、前記データ格納先記憶手段に記憶されている当該データに対応するノードの中から決定するアクセス先決定手段と、
各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、
データへのアクセス履歴に基づいて前記アクセス確率を算出し、算出したアクセス確率を前記アクセス確率記憶手段に記憶させるアクセス確率算出手段と、
指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、
前記アクセス確率推定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定し、
前記アクセス先決定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定することを特徴とするデータアクセス場所選択システム。

[請求項2]

アクセス先決定手段は、アクセス確率記録手段に記録されたアクセス確率のうち、推定されたアクセス確率と所定の値の範囲内にあるアクセス確率のデータを格納するノードを、データの書き込み先と決定

する

請求項 1 記載のデータアクセス場所選択システム。

[請求項3] 推定されたアクセス確率に基づいて、時間帯ごとの各ノードの動作スケジュールを決定するノード動作スケジュール決定手段を備え、

アクセス先決定手段は、推定されたアクセス確率が所定の値を超える時間帯における前記動作スケジュールが稼働中であるノードをデータの読み取り先および／または書き込み先と決定する

請求項 1 または請求項 2 記載のデータアクセス場所選択システム。

[請求項4] ノードに対して制御指示命令を発行する制御指示命令手段を備え、前記制御指示命令手段は、動作スケジュールに基づいてノードに対する制御指示命令を行う

請求項 3 記載のデータアクセス場所選択システム。

[請求項5] アクセス確率推定手段は、指定されたデータと同一の拡張子を有するデータのアクセス確率を前記アクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定する

請求項 1 から請求項 4 のうちのいずれか 1 項に記載のデータアクセス場所選択システム。

[請求項6] アクセス確率推定手段は、指定されたデータと同一のディレクトリに格納されるデータのアクセス確率を前記アクセス確率記憶手段から抽出し、抽出したアクセス確率の平均値を、指定されたデータのアクセス確率と推定する

請求項 1 から請求項 4 のうちのいずれか 1 項に記載のデータアクセス場所選択システム。

[請求項7] アクセス確率推定手段は、データへのアクセス履歴及び前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて共起確率を算出することにより、指定されたデータのアクセス確率を推定する

請求項 1 から請求項 4 のうちのいずれか 1 項に記載のデータアクセ

ス場所選択システム。

[請求項8]

データを格納する複数のノードを備え、

前記各ノードは、

データの格納先をノードと対応付けて記憶するデータ格納先記憶手段と、

指定されたデータの読み取り先および／または書き込み先を、前記データ格納先記憶手段に記憶されている当該データに対応するノードの中から決定するアクセス先決定手段と、

各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率を記憶するアクセス確率記憶手段と、

データへのアクセス履歴に基づいて前記アクセス確率を算出し、算出したアクセス確率を前記アクセス確率記憶手段に記憶させるアクセス確率算出手段と、

指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定手段とを備え、

前記アクセス確率推定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定し、

前記アクセス先決定手段は、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定する

ことを特徴とするデータアクセス場所選択システム。

[請求項9]

データを格納する少なくとも1つ以上のノードの中から、データの読み取り先および／または書き込み先を選択する選択装置が、データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されている当該データに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定し、

前記選択装置が、各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確率をアクセス確率記憶手段に記憶させ、

前記選択装置が、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、指定されたデータに対するアクセス確率を前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定し、

前記選択装置が、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定する

ことを特徴とするデータアクセス場所選択方法。

[請求項10]

データを格納する少なくとも1つ以上のノードの中から、データの読み取り先および／または書き込み先を選択するコンピュータに適用されるデータアクセス場所選択プログラムであって、

前記コンピュータに、

データの格納先をノードと対応付けて記憶するデータ格納先記憶手段に記憶されている当該データに対応するノードの中から指定されたデータの読み取り先および／または書き込み先を決定するアクセス先決定処理、

各ノードに対してデータの読み取りおよび／または書き込みが行われる確率を示すアクセス確率をデータへのアクセス履歴に基づいて算出し、算出されたアクセス確率をアクセス確率記憶手段に記憶させるアクセス確率算出処理、および、

指定されたデータに対するアクセス確率を、前記アクセス確率記憶手段に記憶されたアクセス確率に基づいて推定するアクセス確率推定処理を実行させ、

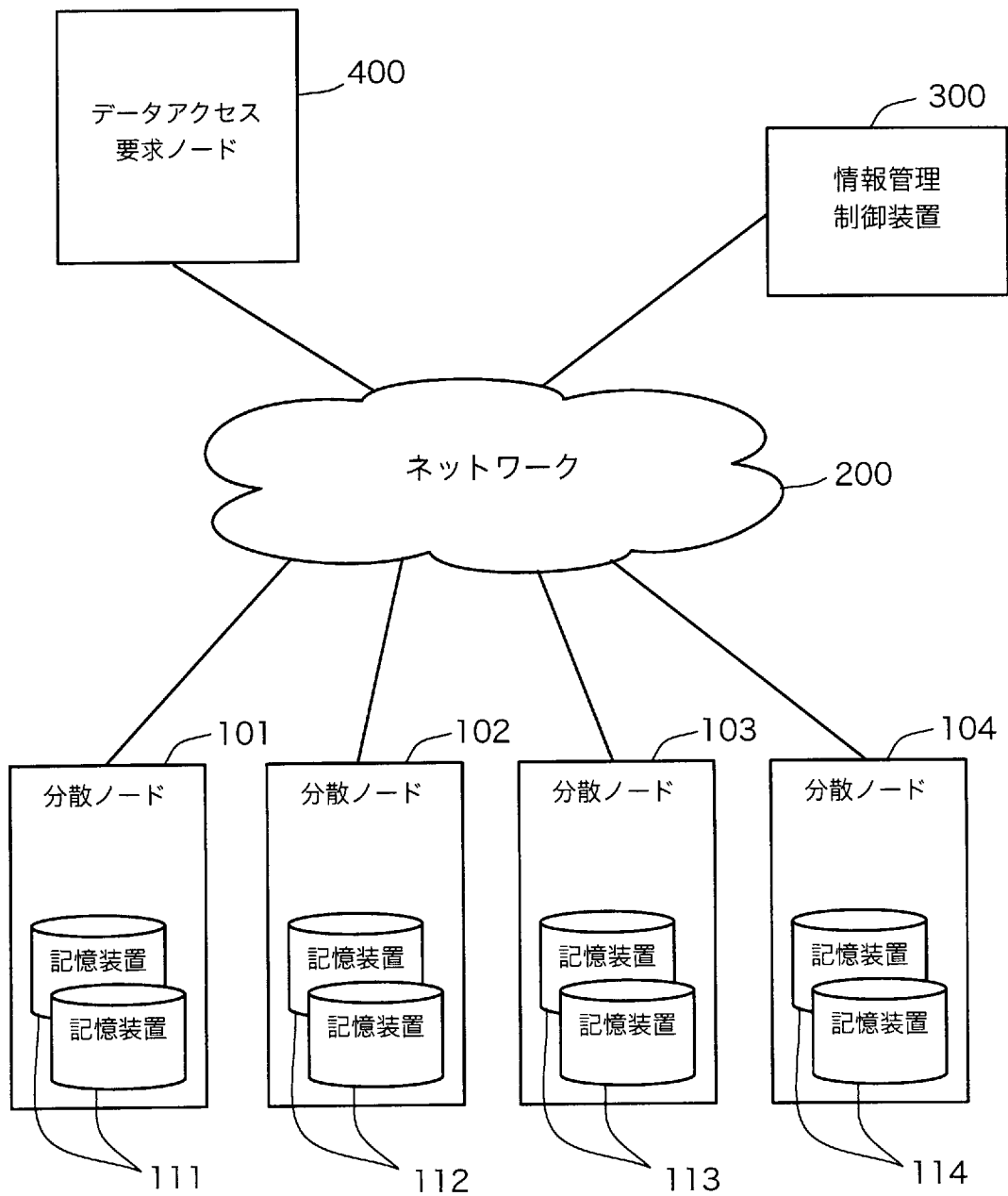
前記アクセス確率推定処理で、指定されたデータに対応するノード

が前記データ格納先記憶手段に記憶されていない場合、当該データのアクセス確率を推定させ、

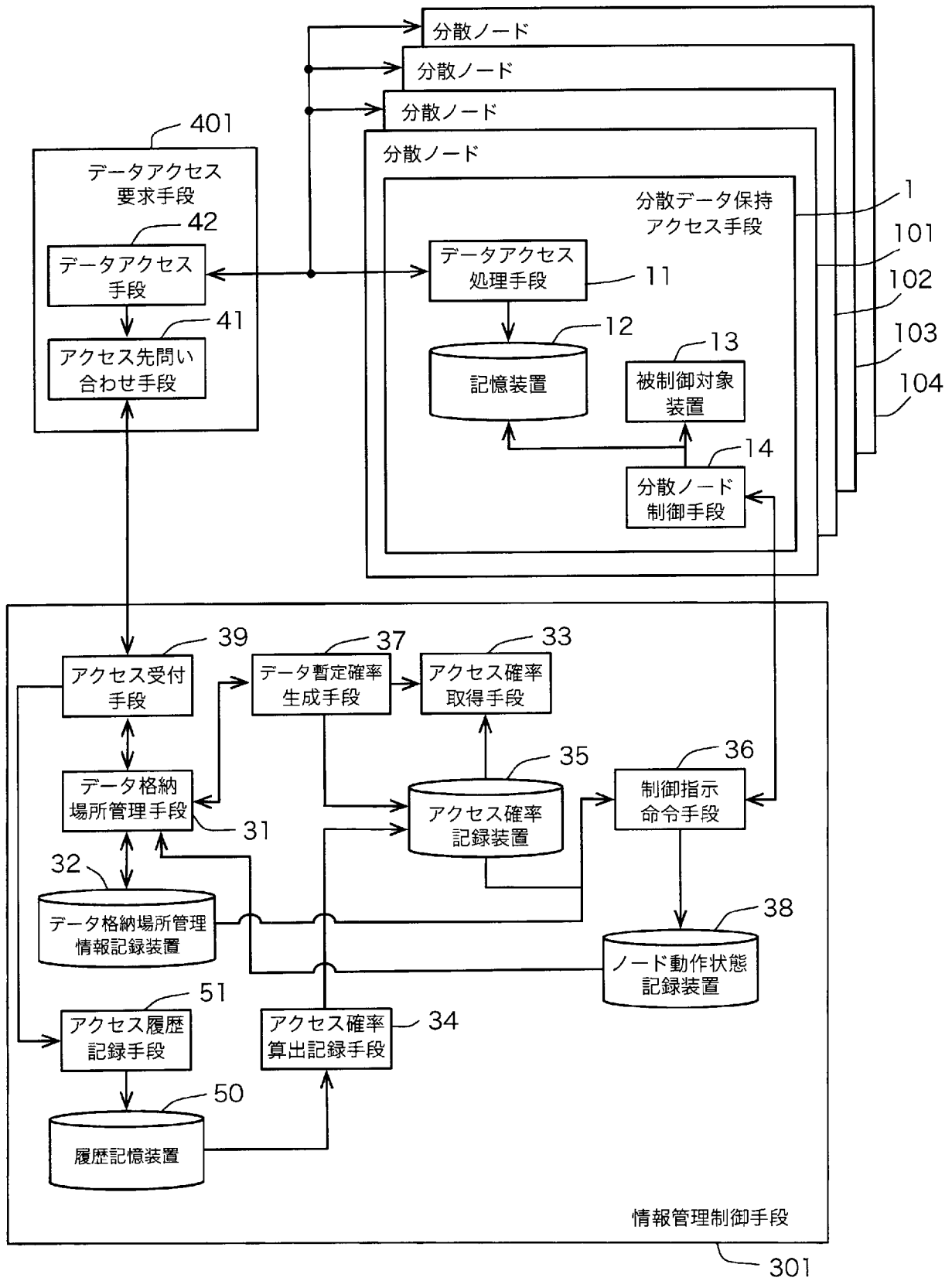
前記アクセス先決定処理で、指定されたデータに対応するノードが前記データ格納先記憶手段に記憶されていない場合、推定された前記アクセス確率をもとに当該データの書き込み先を決定させる

ためのデータアクセス場所選択プログラム。

[図1]



[図2]



[図3]

ファイル名	ノード
/home/a.txt	node1 ,node2
/home/b.txt	node1 ,node2
/home/c.txt	node1 ,node3
/tmp/a.txt	node4
...	...

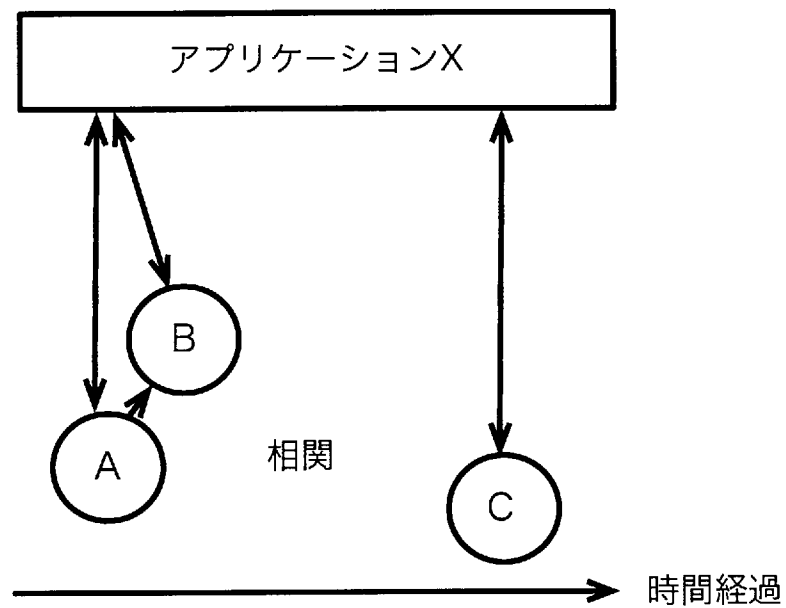
[図4]

Read or Write	dest (ファイル名)	time
Write	/ home / a. t x t	2009.01.01 01:00:01
Write	/ home / b. t x t	2009.01.01 01:00:01
Write	/ home / c. log	2009.01.01 01:00:03
Write	/ home / c. log	2009.01.01 01:00:04
Read	/ home / a. t x t	2009.01.01 01:00:06
Write	/ home / a. t x t	2009.01.01 01:00:07
Write	/ home / b. t x t	2009.01.01 01:00:09
Read	/ home / b. t x t	2009.01.01 01:00:10
Read	/ home / c. log	2009.01.01 01:00:10
Read	/ home / a. t x t	2009.01.01 01:00:11
...	...	...

[図5]

ファイル名	一日内の時間帯別確率				月別確率					
	0 - 12		12 - 24		上旬		中旬		下旬	
	Read	Write	Read	Write	Read	Write	Read	Write	Read	Write
/home/a.txt	10%	0%	0%	5%	100%	33%	0%	33%	0%	33%
/home/b.txt	0%	0%	3%	3%	0%	0%	0%	100%	100%	0%
/home/c.log	0%	30%	100%	100%	100%	100%	100%	100%	100%	100%
/home/a.txt	100%	100%	0%	0%	100%	100%	100%	100%	100%	100%

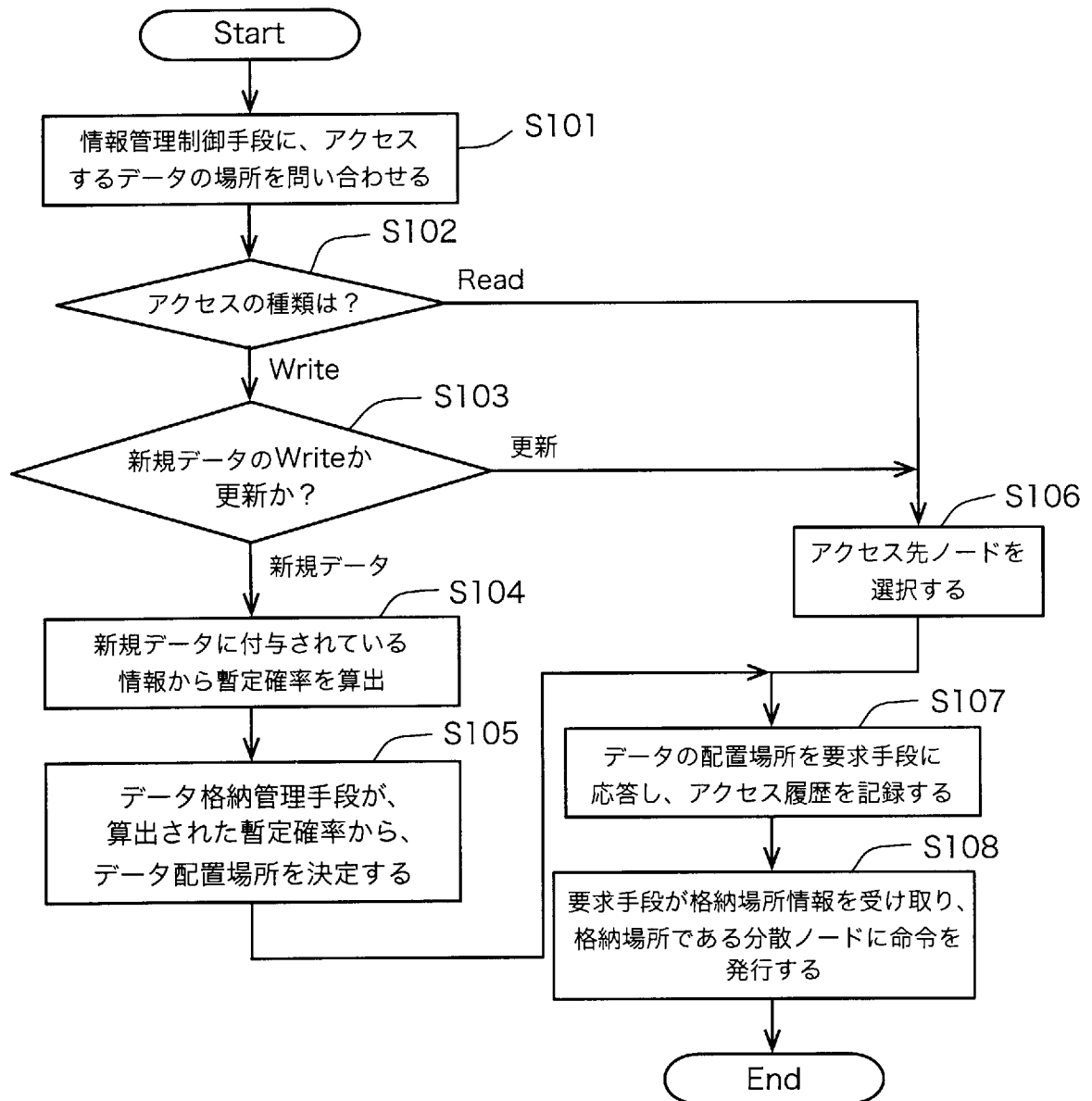
[図6]



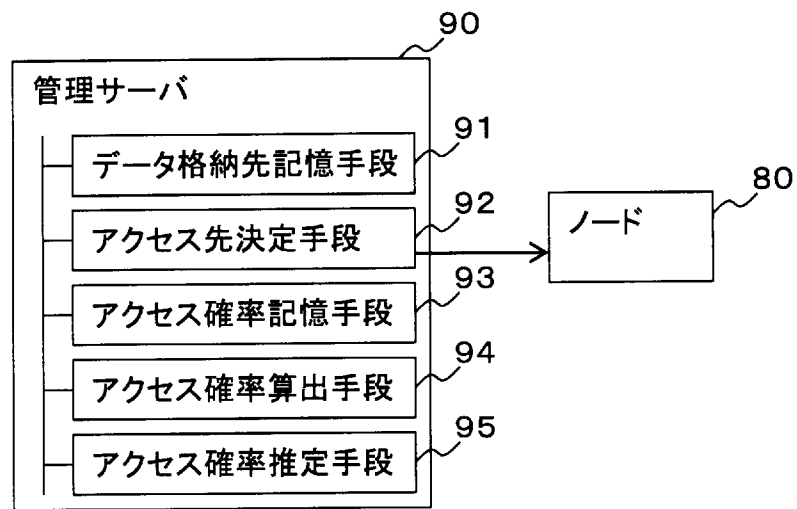
[図7]

ノード名	状態
node1	active
node2	low speed
node3	standby
node4	active
...	...

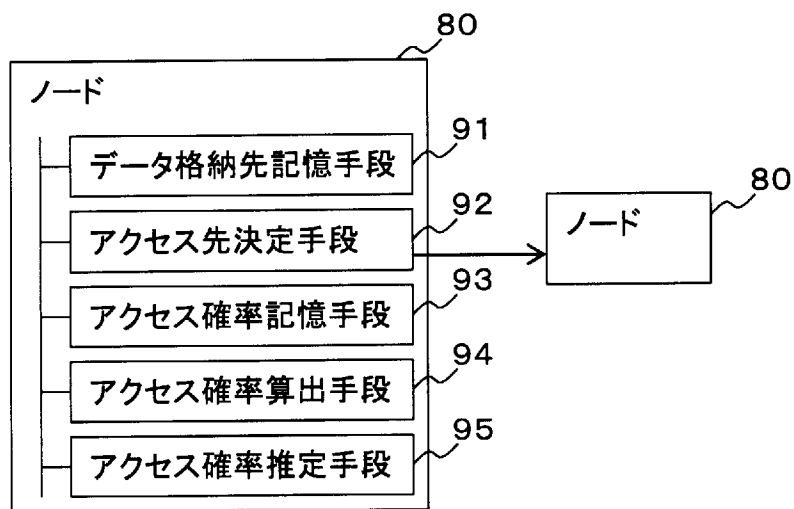
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/000311

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/00(2006.01)i, G06F3/06(2006.01)i, G06F13/10(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)

G06F12/00, G06F3/06, G06F13/10

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 1996-2011

Kokai Jitsuyo Shinan Koho 1971-2011 Toroku Jitsuyo Shinan Koho 1994-2011

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

JSTPlus (JDreamII)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2009-157571 A (Fujitsu Ltd.), 16 July 2009 (16.07.2009), entire text; all drawings & US 2009/0254702 A1	1-10
A	JP 2004-240803 A (Hitachi, Ltd.), 26 August 2004 (26.08.2004), entire text; all drawings & US 2004/0205143 A1	1-10
A	JP 2003-280950 A (Fujitsu Ltd.), 03 October 2003 (03.10.2003), entire text; all drawings (Family: none)	1-10



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

02 March, 2011 (02.03.11)

Date of mailing of the international search report

15 March, 2011 (15.03.11)

Name and mailing address of the ISA/

Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類（国際特許分類（I P C）） Int.Cl. G06F12/00(2006.01)i, G06F3/06(2006.01)i, G06F13/10(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（I P C）） Int.Cl. G06F12/00, G06F3/06, G06F13/10			
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 日本国公開実用新案公報 日本国実用新案登録公報 日本国登録実用新案公報 1 9 2 2－1 9 9 6年 1 9 7 1－2 0 1 1年 1 9 9 6－2 0 1 1年 1 9 9 4－2 0 1 1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） JSTPlus(JDreamII)			
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	JP 2009-157571 A（富士通株式会社）2009.07.16, 全文, 全図 & US 2009/0254702 A1		1-10
A	JP 2004-240803 A（株式会社日立製作所）2004.08.26, 全文, 全図 & US 2004/0205143 A1		1-10
A	JP 2003-280950 A（富士通株式会社）2003.10.03, 全文, 全図 （ファミリーなし）		1-10
☐ 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 0 2 . 0 3 . 2 0 1 1		国際調査報告の発送日 1 5 . 0 3 . 2 0 1 1	
国際調査機関の名称及びあて先 日本国特許庁（I S A / J P） 郵便番号1 0 0－8 9 1 5 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） 北村 学 電話番号 0 3－3 5 8 1－1 1 0 1 内線 3 5 9 9	5 M 4 5 3 5