

(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
2010 年 7 月 29 日(29.07.2010)

PCT

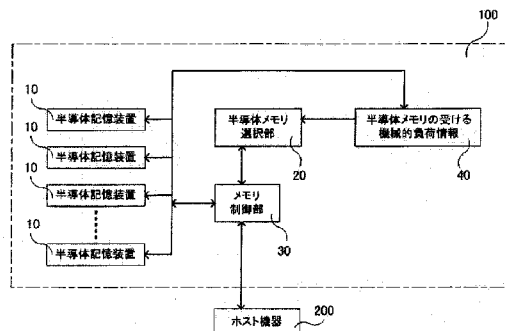
(10) 国際公開番号
WO 2010/084706 A1

- (51) 国際特許分類:
G06F 3/06 (2006.01) G06F 12/16 (2006.01)
- (21) 国際出願番号: PCT/JP2010/000084
- (22) 国際出願日: 2010 年 1 月 8 日(08.01.2010)
- (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 2009-013226 2009 年 1 月 23 日(23.01.2009) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社 東芝(KABUSHIKI KAISHA TOSHIBA) [JP/JP]; 〒1058001 東京都港区芝浦一丁目 1 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): 大森隆広(OMORI, Takahiro). 向井稔(MUKAI, Minoru). 廣畑賢治(HIROHATA, Kenji).
- (74) 代理人: 黒田雄一(KURODA, Yuuichi); 〒1058001 東京都港区芝浦一丁目 1 番 1 号東芝テクノセンター株式会社内 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, 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.
- (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, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, 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))

(54) Title: RAID SYSTEM USING SEMICONDUCTOR MEMORY DEVICE AND CONTROL METHOD THEREFOR

(54) 発明の名称: 半導体記憶装置を用いた R A I D システム及びその制御方法

[図1]



10 SEMICONDUCTOR MEMORY DEVICE
20 SEMICONDUCTOR MEMORY SELECTION UNIT
30 MEMORY CONTROL UNIT
40 INFORMATION ON MECHANICAL LOADS RECEIVED BY SEMICONDUCTOR MEMORIES
200 HOST DEVICE

(57) Abstract: Provided is a RAID system using semiconductor memory devices wherein assumed mechanical loads and electrical loads are leveled among semiconductor memories which store distributed data. The RAID system (100) to perform data transfer with a host device (200), is provided with a plurality of semiconductor memory devices (10) on which a plurality of semiconductor memories (4) are mounted, a semiconductor memory selection unit (20) which selects a combination of the semiconductor memories (4) when the data is distributed and stored among the plurality of semiconductor memory devices (10), and a memory control unit (30) which accesses the semiconductor memories (4) selected by the semiconductor memory selection unit (20) in response to the request from the host device (200), wherein the semiconductor memory selection unit (20) selects the combination of the semiconductor memories so that the mechanical loads received by the semiconductor memories (4) are leveled.

(57) 要約:

[続葉有]

WO 2010/084706 A1

データを分散して記憶させる半導体メモリに対して、想定される機械的負荷及び電氣的負荷を平準化させる半導体記憶装置を用いたRAIDシステムを提供する。ホスト機器200との間でデータ転送を行うRAIDシステム100であって、複数の半導体メモリ4を搭載する複数の半導体記憶装置10と、複数の半導体記憶装置10間でデータを分散して記憶する際に、半導体メモリ4の組み合わせを選択する半導体メモリ選択部20と、半導体メモリ選択部20が選択した半導体メモリ4に対して、ホスト機器200の要求に応じたアクセスを行うメモリ制御部30とを具備し、半導体メモリ選択部20は、半導体メモリ4の受ける機械的負荷が平均化するように、半導体メモリ4の組み合わせを選択する。

明 細 書

発明の名称：

半導体記憶装置を用いたＲＡＩＤシステム及びその制御方法

技術分野

[0001] 本発明は、複数の半導体記憶装置を用いたＲＡＩＤシステム及びその制御方法に関する。

背景技術

[0002] 近年、代表的な大容量記憶装置としてＨＤＤ (Hard Disk Drive) を用いたＲＡＩＤ (Redundant Array of Inexpensive Disks) が広く用いられている。ＲＡＩＤは安価なＨＤＤを複数個組み合わせ、冗長性を持たせた上でデータを分散記録するため、高速で信頼性の高い記憶装置を安価に実現することができる。

[0003] ＨＤＤにおいて、磁気ヘッドによって円盤状のディスク上の磁性体にデータが記録される。ＨＤＤの寿命に影響を与える要因は多数考えられるが、記憶部であるディスクは円盤状であり、ディスク上で不良が発生する場所はランダム性が高い。よって、物理的な情報記録位置によってデータが失われる確率は同じである可能性が高い。このような理由から、故障という観点から見た場合、ＨＤＤ内部でディスクにデータを記録する物理的な場所は大きな問題にはならない。ＲＡＩＤを構成するＨＤＤの場合も、分散記録するデータのＨＤＤ内部での位置は大きな問題とはならない。

[0004] 他方、不揮発性の半導体メモリを用いた半導体記憶装置の大容量化が急速に進んでいる。半導体記憶装置では、データは基板上に実装された半導体メモリに保存される。ＨＤＤのディスクの場合とは異なり、半導体メモリは必ずしも物理的に対称性を持って基板上に実装されるわけではない。よって、半導体メモリの物理的な位置によって、不良が発生する確率に差異が生じる。そのため、データの復元を考慮した冗長性を持たせた分散記録だけでは十分といえない。

[0005] また、半導体メモリの場合、書き込み回数に制限が存在することから、消去回数の平準化について工夫がなされている。例えば、複数のフラッシュメモリ・モジュールにより構成されたRAIDにおいて、複数のモジュールに渡る消去回数平準化法が提案されている（例えば、特許文献1参照。）。これは、半導体メモリに特有の書き込み回数制限を回避するために、RAIDを構成した状態で複数のモジュールにまたがって平準化することにより、RAIDによるストレージ装置の長寿命化を図るものである。

先行技術文献

特許文献

[0006] 特許文献1：特開2007-265265号公報

発明の概要

発明が解決しようとする課題

[0007] 半導体記憶装置でRAIDを構成し、半導体メモリの物理的配置を考慮せずにデータを分散保存させた場合、各半導体記憶装置内の半導体メモリは不良が発生する確率が異なる。各半導体記憶装置にデータを保存する際、その時点で不良が発生する確率が高い半導体メモリに選択的にデータが分散される可能性が生じる。この結果、HDDによるRAIDと比較して、RAIDによる冗長性以上の不良が一度に発生し、データが失われる可能性がより高くなるという問題があった。

[0008] そこで、本発明は、上記した問題に鑑みてなされたもので、データを分散して記憶させる半導体メモリに対して、想定される機械的負荷及び電気的負荷を平準化させた半導体記憶装置を用いたRAIDシステム及びその制御方法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

[0009] 本発明の一態様によれば、ホスト機器との間でデータ転送を行う半導体記憶装置を用いたRAIDシステムであって、複数の半導体メモリを搭載する複数の半導体記憶装置と、前記複数の半導体記憶装置間でデータを分散して

記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、前記半導体メモリ選択部が選択した前記半導体メモリに対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、前記半導体メモリ選択部は、前記半導体メモリの受ける機械的負荷が平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いたRAIDシステムが提供される。

[0010] また、本発明の別の一態様によれば、ホスト機器との間でデータ転送を行う半導体記憶装置を用いたRAIDシステムであって、複数の半導体メモリを搭載する複数の半導体記憶装置と、前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、前記半導体メモリ選択部が選択した前記半導体メモリに対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、前記半導体メモリ選択部は、前記半導体メモリが搭載される実装基板と前記半導体メモリのはんだ接合部での損傷値が平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いたRAIDシステムが提供される。

[0011] また、本発明の別の一態様によれば、ホスト機器との間でデータ転送を行う半導体記憶装置を用いたRAIDシステムであって、複数の半導体メモリを搭載する複数の半導体記憶装置と、前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、前記半導体メモリ選択部が選択した前記半導体メモリに対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、前記半導体メモリ選択部は、前記半導体メモリが搭載される実装基板と前記半導体メモリのはんだ接合部での損傷値及び前記半導体メモリへの書き込み回数のいずれもが平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いたRAIDシステムが提供される。

発明の効果

[0012] 本発明によれば、データを分散して記憶させる半導体メモリに対して、想定される機械的負荷及び電氣的負荷を平準化させるように、半導体メモリを組み合わせるので、データ損失のリスクを低減することができる。

図面の簡単な説明

[0013] [図1]本発明の実施の形態に係る半導体記憶装置を用いたRAIDシステムの構成を示すブロック図である。

[図2]本発明の実施の形態に係る半導体記憶装置の概略構成を示す模式図である。

[図3]本発明の実施の形態に係る半導体記憶装置を用いたRAIDシステムの一例を示す模式図である。

[図4]最適化された半導体メモリの組み合わせを説明する図である。

[図5]分散化されたデータの保存例を説明する図である。

[図6]半導体メモリの組み合わせ決定までの処理の流れを示すフローチャートである。

[図7]代表距離の決定の仕方を説明する図である。

[図8]第3の実施形態に係る半導体記憶装置を用いたRAIDシステムの構成を示すブロック図である。

[図9]半導体メモリの組み合わせ決定までの処理の流れを示すフローチャートである。

[図10]第4の実施形態に係る半導体記憶装置を用いたRAIDシステムの構成を示すブロック図である。

[図11]半導体メモリの組み合わせ決定までの処理の流れを示すフローチャートである。

発明を実施するための形態

[0014] 以下、本発明の実施の形態について図面を参照しながら説明する。各図において同一箇所については同一の符号を付すとともに、重複した説明は省略する。

[0015] (第1の実施形態)

図1は、本発明の実施の形態に係る半導体記憶装置を用いたRAIDシステムの構成を示すブロック図である。このRAIDシステム100は、複数の半導体記憶装置10と、メモリ制御部30と、半導体メモリ選択部20から構成され、外部機器としてのホスト機器200と接続している。半導体記憶装置10は後述するように、実装基板上に複数の半導体メモリが搭載されている。半導体メモリ選択部20は、複数の半導体記憶装置10間でデータを分散して記憶する際に、半導体メモリの組み合わせを選択するものである。半導体メモリ選択部20は、それぞれの半導体記憶装置10内の半導体メモリの受ける機械的負荷が平均化するように、半導体メモリの組み合わせを選択する。第1の実施形態では、半導体メモリの受ける機械的負荷情報40が半導体メモリ選択部20に取り込まれている。メモリ制御部30は、半導体メモリ選択部20が選択した半導体メモリに対して、ホスト機器200の要求に応じたアクセスを行い、分散したデータの書き込み等を実行する。ホスト機器200としては、例えばパーソナルコンピュータなどの電子機器が挙げられる。

[0016] 図2は、本発明の実施の形態に係る半導体記憶装置10の概略構成を示す模式図である。図2において、実装基板1上の中央からやや右寄りに、制御用のコントロールIC5が配置されている。コントロールIC5の周囲には、コンデンサ6、センサ7、電源8が配置されている。センサ7は、実装基板1上、またはその周囲環境の物理的変動を計測するもので、モニタリングを行うために配設する。計測する物理量としては、加速度、ひずみ、温度、抵抗、インピーダンスのいずれか一つでも、または複数でもよい。

[0017] さらに、これらを取り囲むように、8個の半導体メモリ4が配置されている。したがって、各半導体メモリ4が配置された位置とコントロールIC5とを結ぶ距離は、異なっている。半導体メモリ4は、例えば2値NAND型フラッシュメモリ、あるいは多値NAND型フラッシュメモリが好適である。図2においては、半導体メモリ4の実装数は8個であるが、一つの半導体記憶装置10に実装される半導体メモリ4の数はこれに限られない。実装基

板 1 の一端部には、コネクタ 2 が配設されている。半導体記憶装置 10 は、樹脂封止などでパッケージングされていてもよい。

[0018] 図 2 に示す半導体記憶装置 10 を n 台用いて、図 3 に示すような RAID システム 100 を構築する。図 3 は、RAID システム 100 の一例を示す模式図である。図 3 に示す RAID システム 100 において、一つのデータを n 個のブロックに分散化して各半導体記憶装置 10 に保存する。データを分散化する方法は、特定のものに限定されることなく、RAID の冗長性のレベルによって様々な周知のものを採用することができる。すなわち、冗長性を持たせずにデータを n 分割して保存する場合 (RAID 0)、データそのもののコピーを n 個保存する場合 (RAID 1)、データを分割して保存し、さらに分割データのパリティも保存する場合 (RAID 5、6) 等である。

[0019] 本発明の第 1 の実施形態においては、コントロール IC 5 の発熱による熱変動 (熱負荷) に対する、半導体メモリ 4 への書き込みの最適化を行う。一般的に、半導体記憶装置 10 に用いられるコントロール IC 5 は副次的に大きな発熱を伴う。実装基板 1 上の実装部品に大きな熱変動が存在する場合、電源の ON、OFF により半導体メモリ 4 と実装基板 1 を接続するはんだ接合部には、大きな熱応力が発生する。

[0020] 図 2 に示す半導体記憶装置 10 では、コントロール IC 5 から距離が離れるに従って実装基板 1 上の温度が下降していく。そこで、RAID システム 100 を構成する各半導体記憶装置 10 において、分散してデータを保存する各半導体メモリ 4 の、コントロール IC 5 からの代表距離を極力平均化するように組み合わせを決定する。これにより、RAID システム 100 を構成する半導体メモリ 4 全体に渡って、熱変動の影響を平均化することができる。代表距離としては、コントロール IC 5 の中心から各半導体メモリ 4 の中心までの距離あるいはコントロール IC 5 の中心から最も近接する各半導体メモリのはんだ接合部までの距離などが好適である。

[0021] 代表距離を平均化するための組み合わせは、半導体記憶装置 10 の設計時に決定することができる。しかし、半導体メモリ 4 の劣化等の情報、例えば

既書き込み回数等をも合わせて考慮して、組み合わせの優先順位付けは随時変更することも可能である。

- [0022] 具体的な半導体メモリ 4 の組み合わせの決定方法として、各半導体メモリ 4 とコントロール IC 5 の代表距離を用いて、組み合わせ最適化の観点から決定することができる。k 番目の半導体記憶装置 10 において、コントロール IC 5 から各半導体メモリ 4 までの代表距離を d_{ki} ($k=1 \cdots n$, $i=1 \cdots 8$)、半導体記憶装置 A の 1～8 の半導体メモリ 4 と組み合わせでデータを保存する、他の半導体記憶装置の半導体メモリの組を $B(i)$, $C(i) \cdots$, $N(i)$ (ここで、 $B(i)$, $C(i) \cdots$, $N(i)$ は 1 から 8 までの自然数のいずれかで、かつ $i=1 \cdots 8$ である) とすると、組み合わせを決定する方法は、以下の最適化問題を解くことに帰着される。 $\min. f$ は関数 f を最小化することを表す。

[数1]

$$\min. f = \sum_{i=1}^8 ((d_{1i} + d_{2B(i)} + d_{3C(i)} + \cdots + d_{nN(i)}) / (n - m))^2 \quad (1)$$

$$m = \sum_{i=1}^8 (d_{1i} + d_{2B(i)} + d_{3C(i)} + \cdots + d_{nN(i)}) / (8 \times n)$$

- [0023] 式 (1) は遺伝的アルゴリズム等の一般的な最適化アルゴリズムによって解を得ることができる。分散してデータを記録する半導体メモリ 4 の組み合わせを考慮して、半導体メモリ 4 の順序を並べ替えた様子を図 4 に示す。図 4 に示すように、RAID システム 100 は、N 個の半導体記憶装置 10 から成り、各半導体記憶装置 10 はそれぞれ 8 個の半導体メモリ 4 を搭載している。半導体記憶装置 B 乃至半導体記憶装置 N においては、最適化アルゴリズムによって各半導体メモリ 4 について並べ替えが行われた状態で配置されていることを表わしている。図 4 の最適化された半導体メモリ 4 の組み合わせに対して、データを図 5 のように分散して記録する。図 5 に示すように、分散データ a 1 を記録する半導体メモリ A 1 に対して、組み合わせとして選定された半導体メモリ B (1) に分散データ b 1 を記録し、組み合わせとして選定された半導体メモリ C (1) に分散データ c 1 を記録している。分散データ n 1 についても同様に記録する。

[0024] 本実施形態によれば、分散データを記録する半導体メモリの組み合わせに、コントロール IC 等から受ける熱等の機械的負荷を考慮するので、データ損失のリスクを低減することができる

(第 2 の実施形態)

本発明に係る第 2 の実施形態として、外部からの荷重や振動に対する書き込み最適化を行う場合を説明する。図 2 に示すように、支持部（ボス穴）3 が形成された実装基板 1 においては、外部から荷重や振動が加えられた際に、支持部 3 を介して実装基板 1 に荷重が伝達される。支持部 3 を経由して荷重が加えられた場合、一般的に、支持部 3 に近接するに従って実装基板の曲率が大きくなる傾向がある。半導体メモリ 4 と実装基板 1 との間のはんだ接合部には、はんだ接合部における実装基板 1 の曲率が大きいほど大きな応力が発生する。

[0025] そこで、第 2 の実施形態においては、各半導体メモリ 4 と支持部 3 までの代表距離が平均化するようにデータを分散保存する。代表距離として、支持部 3 の中心から各半導体メモリ 4 の中心までの距離あるいは支持部 3 の中心から最も近接する各半導体メモリ 4 のはんだ接合部までの距離などを選定することが好適である。上記式 (1) により、代表距離を設定して最適化問題に帰着させることで、組み合わせの解を得ることができる。

[0026] 代表距離を平均化するための組み合わせは、半導体記憶装置 10 の設計時に決定することができる。しかし、半導体メモリ 4 の劣化等の情報、例えば既書き込み回数等をも合わせて考慮して、組み合わせの優先順位付けは随時変更することも可能である。

[0027] ここで、上記した第 1 の実施形態及び第 2 の実施形態における半導体メモリ 4 の組み合わせ決定までの処理の流れについて、図 6 により説明する。

[0028] まず、基準位置となる熱源やボス穴の位置を決定する（ステップ S 6 1）。ここで、熱源はコントロール IC 5 あるいは電源 8 であり、ボス穴は支持部 3 である。各半導体メモリ 4 への代表距離 d_{ki} を決定する（ステップ S 6 2）。代表距離 d_{ki} の決定は、一つの決め方に限られない。そこで、例えば

、図7に示すよう決定することが好適である。図7は、代表距離の決定の仕方を説明する図である。図7に示すように、熱源となるコントロールIC5から半導体メモリ4のはんだ接合部9への最短距離 d_{ts} 、熱源となるコントロールIC5から半導体メモリ4の中心部への最短距離 d_{tc} 、ボス穴3から半導体メモリ4のはんだ接合部9への最短距離 d_{ms} 、ボス穴3から半導体メモリ4の中心部への最短距離 d_{mc} のいずれでもよい。

[0029] 代表距離 d_{ki} の決定後は、最適化問題を解いて半導体メモリ4の組み合わせを決定する（ステップS63）。RAIDシステム100に設けたメモリ制御部30へ、決定した半導体メモリ4の組み合わせを格納する（ステップS64）。

[0030] （第3の実施形態）

本発明に係る第3の実施の形態として、物理的状态を測定しながら各半導体メモリ4のはんだ接合部に蓄積される損傷値を推定し、推定された損傷値を元に書き込み最適化を行う場合について説明する。図8は、本発明の第3の実施形態に係る半導体記憶装置を用いたRAIDシステム100の構成を示すブロック図である。このRAIDシステム100は、複数の半導体記憶装置10と、メモリ制御部30と、半導体メモリ選択部20から構成され、外部機器としてのホスト機器200と接続している。さらに、第3の実施形態においては、疲労特性データベース50を予め準備する。この疲労特性データベース50には、例えば、はんだ接合部9のひずみ範囲と寿命の関係を表すデータが蓄積されている。このデータベース50は、実装基板1の設計情報や使用している材料の情報を利用して、半導体メモリ4が搭載された半導体記憶装置10に対して現象解析を実行し、はんだ接合部9に生じる負荷を予測することにより、はんだ接合部9の破損に関する関係を算出して構築することができる。

[0031] 第1の実施形態、第2の実施形態において述べたように、はんだ接合部9は機械的な負荷により損傷を受ける。ある一定の損傷が蓄積すると、はんだ接合部9は疲労により電氣的に断線する。本実施形態では、センサ7で測定

された物理量を、別途保存してある疲労特性データベース50と比較することにより、各半導体メモリ4のはんだ接合部9の損傷値を算出する。この損傷値についての情報51を元に、半導体メモリ選択部20において、分散したデータを書き込む半導体メモリ4の組み合わせを決定する。半導体メモリ4の組み合わせを更新する頻度は、ファイル書き込み毎、一定期間使用后、半導体メモリ内のガベージコレクションのタイミング等、の場合が考えられる。また、第1の実施形態、第2の実施形態で算出される代表距離を損傷値に置き換えることで、データを分散させる半導体メモリ4の組み合わせを、最適化問題を解くことにより得ることができる。

[0032] ここで、第3の実施形態における半導体メモリ4の組み合わせ決定までの処理の流れについて、図9により説明する。

[0033] まず、センサ7からの情報を取得する（ステップS91）。取得したセンサ情報は一定時間ごとにセンサ情報データベース60に記憶する（ステップS92）。半導体メモリ4の組み合わせを更新する必要があるか否かを判定するために、組み合わせ更新イベントを検出する（ステップS93）。組み合わせ更新イベントを検出しなかった場合には、ステップS91に戻り、組み合わせ更新イベントを検出した場合には、センサ情報が格納されたセンサ情報データベース60と疲労特性が格納された疲労特性データベース50を用いて、各半導体メモリ4のはんだ接合部9の損傷値を計算する（ステップS94）。次いで、各半導体メモリ4内の最も損傷値が大きいのはんだ接合部9の損傷値を代表損傷値 d_{ki} とする（ステップS95）。代表損傷値 d_{ki} から、最適化問題を解いて半導体メモリ4の組み合わせ $B(i) \cdots N(i)$ を決定する（ステップS96）。RAIDシステムに設けたメモリ制御部30へ、決定した半導体メモリ4の組み合わせを格納する（ステップS97）。その後、各半導体メモリ4内に記憶されたデータの組み合わせを更新（ステップS98）し、ステップS91に戻る。

[0034] （第4の実施形態）

次に、本発明に係る第4の実施形態について説明する。図10は、第4の

実施形態に係る半導体記憶装置を用いたRAIDシステム100の構成を示すブロック図である。このRAIDシステム100は、複数の半導体記憶装置10と、メモリ制御部30と、半導体メモリ選択部20から構成され、外部機器としてのホスト機器200と接続している。さらに、第4の実施形態においては、疲労特性データベース50と、各半導体メモリ4への書き込み回数情報71を格納した書き込み回数データベース70を予め準備している。一般的に、半導体メモリ4では、記憶素子の記録回数は有限であり、数百万回程度で記憶素子が劣化して書き込みが不可能になる。一方、情報は、記憶素子の集まりであるブロック単位で書き込みあるいは消去が行われる。そこで、或るブロックに書き込みを行った回数を別途記録し、次に書き込みを行うブロックを決める際の参考とし、書き込みの回数が分散化され、特定のブロックに書き込みが集中することを防いでいる。

[0035] そこで、本実施形態では、各半導体メモリ4の記憶素子の劣化の情報を考慮するもので、それぞれの半導体メモリ4の総書き込み回数と、機械的な損傷の可能性を示す指標の両方を利用して、書き込みを行う半導体メモリ4の組み合わせを決定する。たとえば、 k 番目の半導体記憶装置10において、 i 番目の半導体メモリ4の総書き込み回数を D_{ki} とすると、以下の式(2)のような最適化問題によって組合せを決定することができる。

[数2]

$$\begin{aligned} \min. f &= \sum_{i=1}^8 \left((d_{1i} + d_{2B(i)} + d_{3C(i)} + \dots + d_{nN(i)}) / (n-m) \right)^2 + \alpha g(D_{ki}) \\ m &= \sum_{i=1}^8 (d_{1i} + d_{2B(i)} + d_{3C(i)} + \dots + d_{nN(i)}) / (8 \times n) \end{aligned} \quad (2)$$

[0036] α は正の定数を表す。関数 $g(D_{ki})$ と定数 α の決め方によって、機械的な損傷と記憶素子の書き込みによる劣化とのバランスをとることができる。一例として、以下の式(3)のように、機械的な損傷と同様、平均化を目的とする関数を選択することもできる。

[数3]

$$g = \sum_{i=1}^8 \left((D_{1i} + D_{2B(i)} + D_{3C(i)} + \dots + D_{nN(i)}) / (n - M) \right)^2 \quad (3)$$

$$M = \sum_{i=1}^8 (D_{1i} + D_{2B(i)} + D_{3C(i)} + \dots + D_{nN(i)}) / (8 \times n)$$

[0037] ここで、第4の実施形態における半導体メモリ4の組み合わせ決定までの処理の流れについて、図11により説明する。

[0038] まず、センサ7からの情報を取得する（ステップS111）。取得したセンサ情報は一定時間ごとにセンサ情報データベース60に記憶する（ステップS112）。半導体メモリ4の組み合わせを更新する必要があるか否かを判定するために、組み合わせ更新イベントを検出する（ステップS113）。組み合わせ更新イベントを検出しなかった場合には、ステップS111に戻り、組み合わせ更新イベントを検出した場合には、センサ情報が格納されたセンサ情報データベース60と疲労特性が格納された疲労特性データベース50を用いて、各半導体メモリ4のはんだ接合部9の損傷値を計算する（ステップS114）。次いで、各半導体メモリ4内の最も損傷値が大きいのはんだ接合部9の損傷値を代表損傷値 d_{ki} とする（ステップS115）。代表損傷値 d_{ki} を決定した後、半導体メモリ書き込み回数データベース70に格納された半導体メモリ書き込み回数 D_{ki} から、最適化問題を解いて半導体メモリ4の組み合わせ $B(i) \dots N(i)$ を決定する（ステップS116）。RAIDシステム100に設けたメモリ制御部30へ、決定した半導体メモリの組み合わせを格納する（ステップS117）。その後、各半導体メモリ4内に記憶されたデータの組み合わせを更新（ステップS118）し、ステップS111に戻る。

[0039] 本実施形態によれば、半導体記憶装置列によって構成されたRAIDシステム100において、特定の半導体メモリの組み合わせにデータを保存したことによる、機械的な破壊を原因としたデータ損失のリスクを低減することができる。

[0040] なお、本発明は上記の実施形態のそのままに限定されるものではなく、実

施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。
また、上記の実施形態に開示されている複数の構成要素の適宜な組み合わせにより、種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態にわたる構成要素を適宜組み合わせてもよい。

符号の説明

[0041] 1・・・実装基板、2・・・コネクタ、3・・・ボス穴、4・・・半導体メモリ、5・・・コントロールIC、6・・・コンデンサ、7・・・センサ、8・・・電源、9・・・はんだ接合部、10・・・半導体記憶装置、20・・・半導体メモリ選択部、30・・・メモリ制御部、50・・・疲労特性データベース、60・・・センサ情報データベース、70・・・半導体メモリ書き込み回数データベース、100・・・半導体記憶装置を用いたRAIDシステム、200・・・ホスト機器。

請求の範囲

- [請求項1] ホスト機器との間でデータ転送を行う半導体記憶装置を用いた R A I D システムであって、
 複数の半導体メモリを搭載する複数の半導体記憶装置と、
 前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、
 前記半導体メモリ選択部が選択した前記半導体メモリの組み合わせに対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、
 前記半導体メモリ選択部は、前記半導体メモリの受ける機械的負荷が平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いた R A I D システム。
- [請求項2] 前記機械的負荷の大きさは、実装基板上の前記半導体メモリと前記実装基板上の熱源との距離に基づいて判定することを特徴とする請求項 1 記載の半導体記憶装置を用いた R A I D システム。
- [請求項3] 前記機械的負荷の大きさは、実装基板上の前記半導体メモリと実装基板を支持するための支持部との距離に基づいて判定することを特徴とする請求項 1 記載の半導体記憶装置を用いた R A I D システム。
- [請求項4] 前記半導体メモリの組み合わせは、ファイル書き込み毎、一定期間使用後、前記半導体メモリ内のガベージコレクションのタイミング等で更新を実行することを特徴とする請求項 1 に記載の半導体記憶装置を用いた R A I D システム。
- [請求項5] ホスト機器との間でデータ転送を行う半導体記憶装置を用いた R A I D システムであって、
 複数の半導体メモリを搭載する複数の半導体記憶装置と、
 前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、
 前記半導体メモリ選択部が選択した前記半導体メモリの組み合わせ

に対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、

前記半導体メモリ選択部は、前記半導体メモリが搭載される実装基板と前記半導体メモリのはんだ接合部での損傷値が平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いたRAIDシステム。

[請求項6] 前記実装基板上、または前記実装基板の周囲環境の物理的変動を計測するセンサと、前記はんだ接合部の疲労特性データベースを備え、

前記はんだ接合部での損傷値は、前記センサからの情報と前記はんだ接合部の疲労特性に基づいて計算することを特徴とする請求項5記載の半導体記憶装置を用いたRAIDシステム。

[請求項7] ホスト機器との間でデータ転送を行う半導体記憶装置を用いたRAIDシステムであって、

複数の半導体メモリを搭載する複数の半導体記憶装置と、

前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの組み合わせを選択する半導体メモリ選択部と、

前記半導体メモリ選択部が選択した前記半導体メモリの組み合わせに対して、前記ホスト機器の要求に応じたアクセスを行うメモリ制御部とを具備し、

前記半導体メモリ選択部は、前記半導体メモリが搭載される実装基板と前記半導体メモリのはんだ接合部での損傷値及び前記半導体メモリへの書き込み回数のいずれもが平均化するように、前記半導体メモリの組み合わせを選択することを特徴とする半導体記憶装置を用いたRAIDシステム。

[請求項8] 前記実装基板上、または前記実装基板の周囲環境の物理的変動を計測するセンサと、前記はんだ接合部の疲労特性データベースを備え、

前記はんだ接合部での損傷値は、前記センサからの情報と前記はんだ接合部の疲労特性に基づいて計算することを特徴とする請求項7記

載の半導体記憶装置を用いたRAIDシステム。

[請求項9]

ホスト機器との間でデータ転送を行い、複数の半導体メモリを搭載する複数の半導体記憶装置を用いたRAIDシステムにおける制御方法であって、

前記複数の半導体記憶装置間でデータを分散して記憶する際に、前記半導体メモリの受ける機械的負荷が平均化するように、前記半導体メモリの組み合わせを選択し、

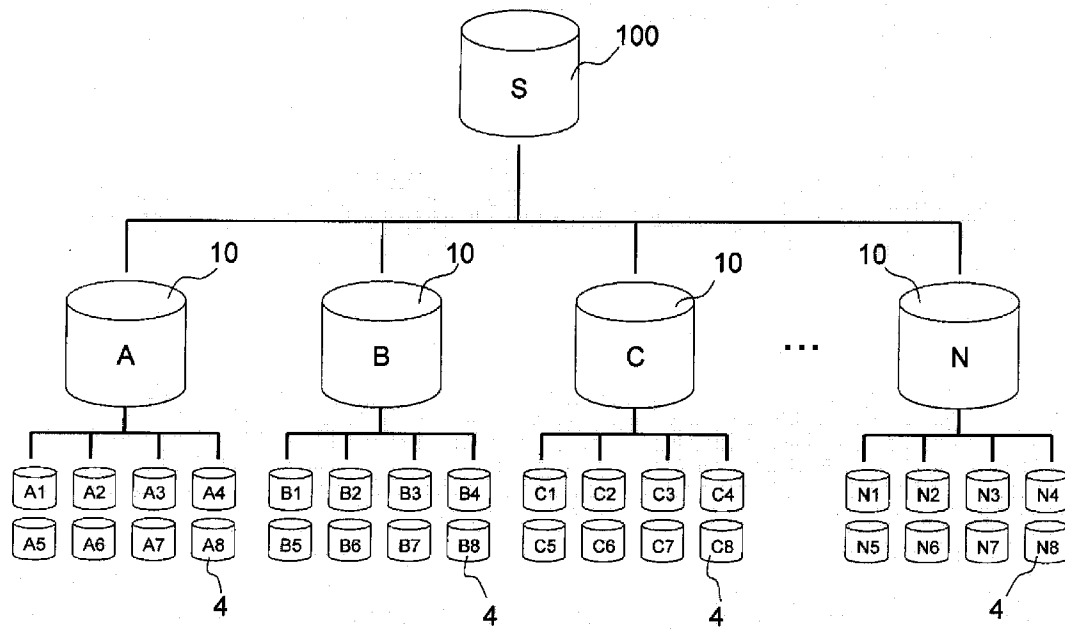
選択した前記半導体メモリの組み合わせに対して、前記ホスト機器の要求に応じたアクセスを行う、ことを特徴とする半導体記憶装置を用いたRAIDシステムにおける制御方法。

```

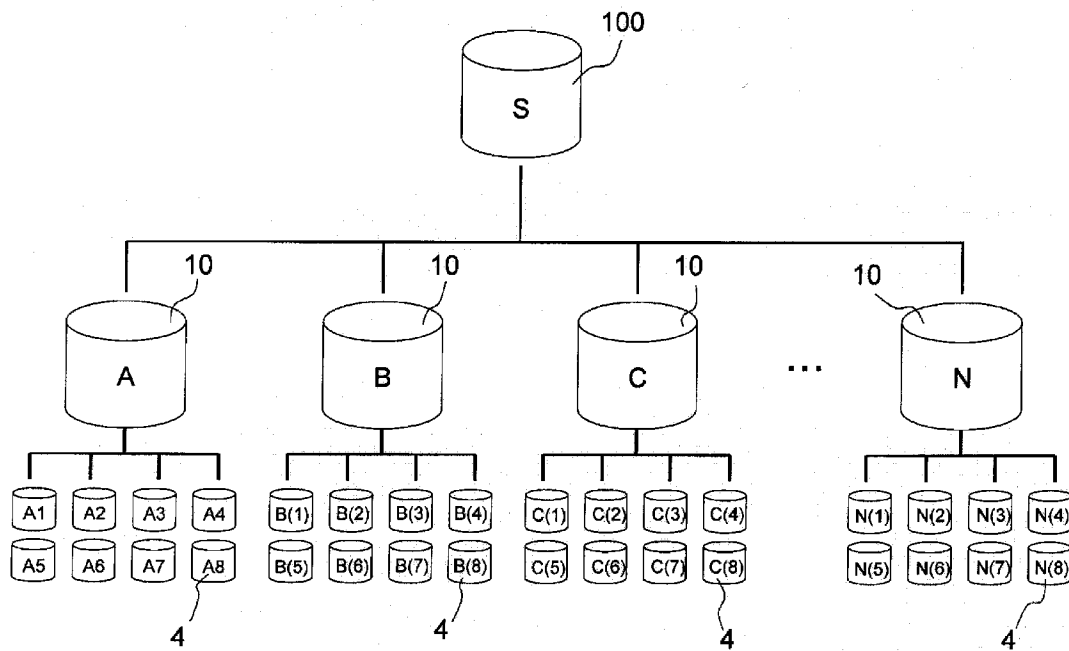
graph TD
    200[ホスト機器 200] <--> 30[メモリ制御部 30]
    30 <--> 20[半導体メモリ選択部 20]
    30 <--> 10[半導体記憶装置 10]
    20 <--> 40[半導体メモリの受ける機械的負荷情報 40]
    40 --> 20
    40 --> 10
  
```

Figure 1 is a schematic diagram of a rectangular frame assembly 1. The frame 1 is defined by a rectangular border 3 and four corner fasteners 3. It contains a grid of rectangular panels 4. A central panel 5 is surrounded by a cluster of small rectangular components 6 and 7. A vertical strip 8 is located on the right side of the frame.

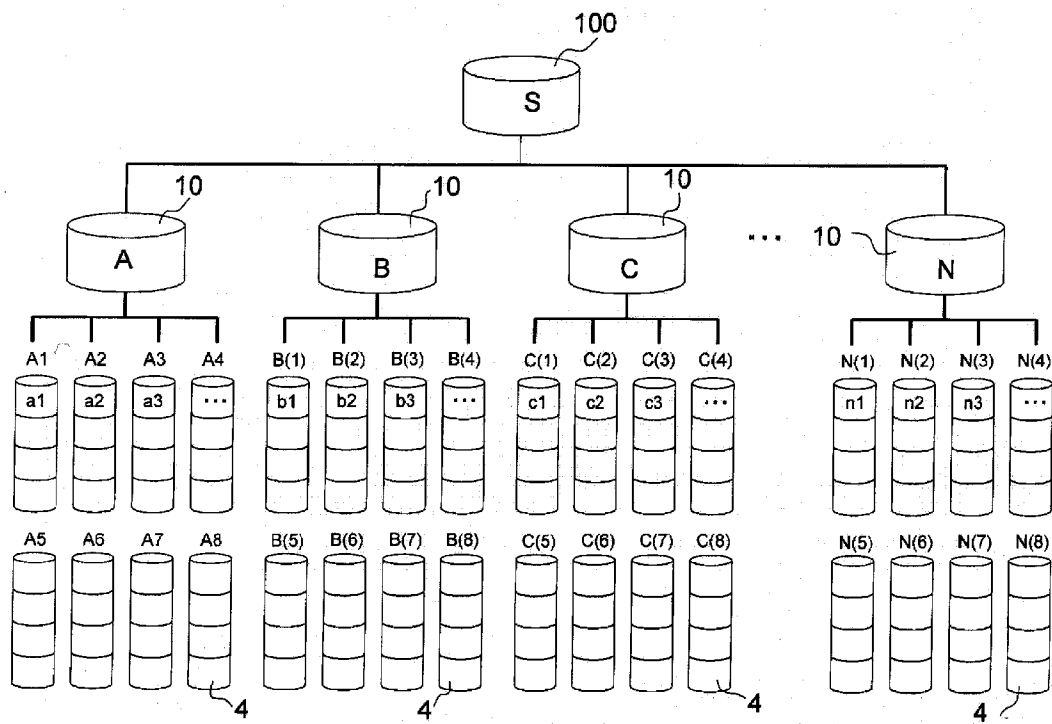
[図3]



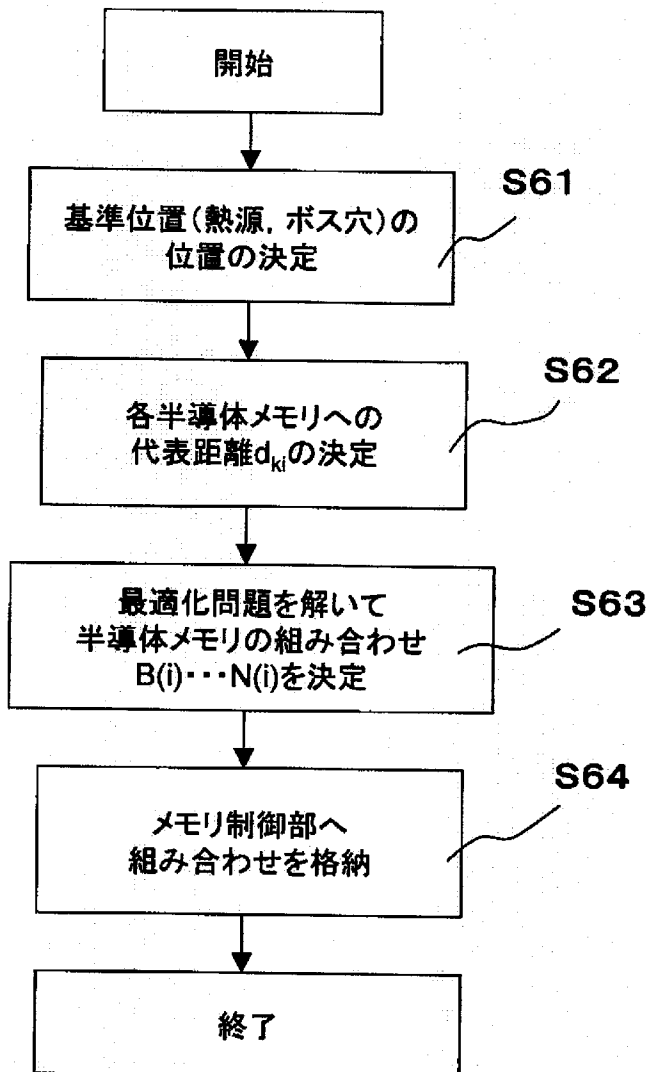
[図4]



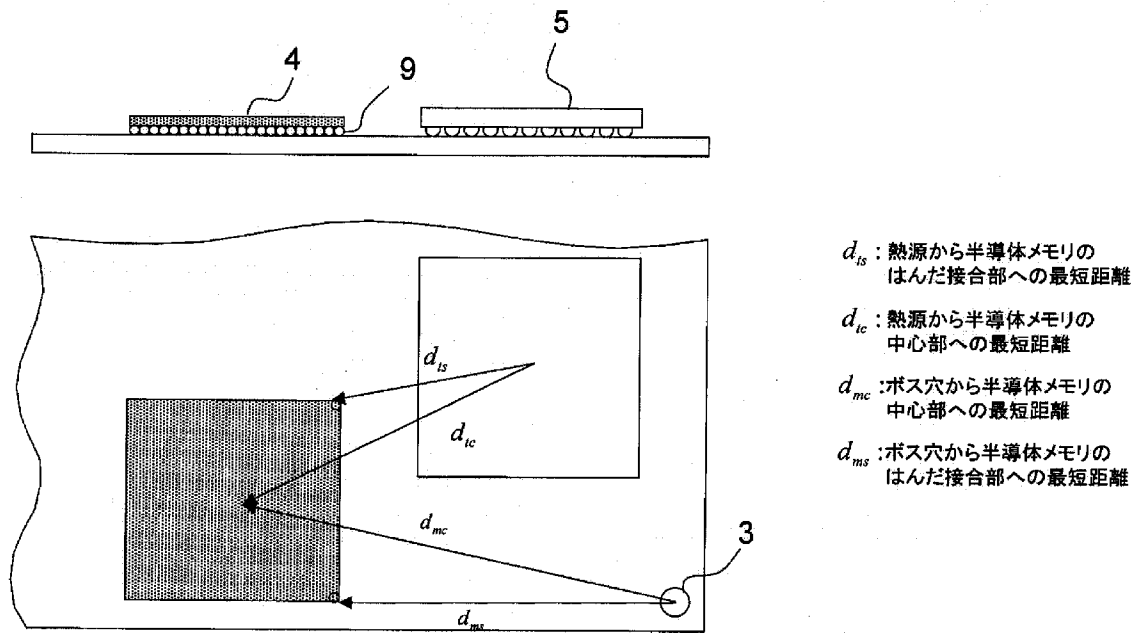
[図5]



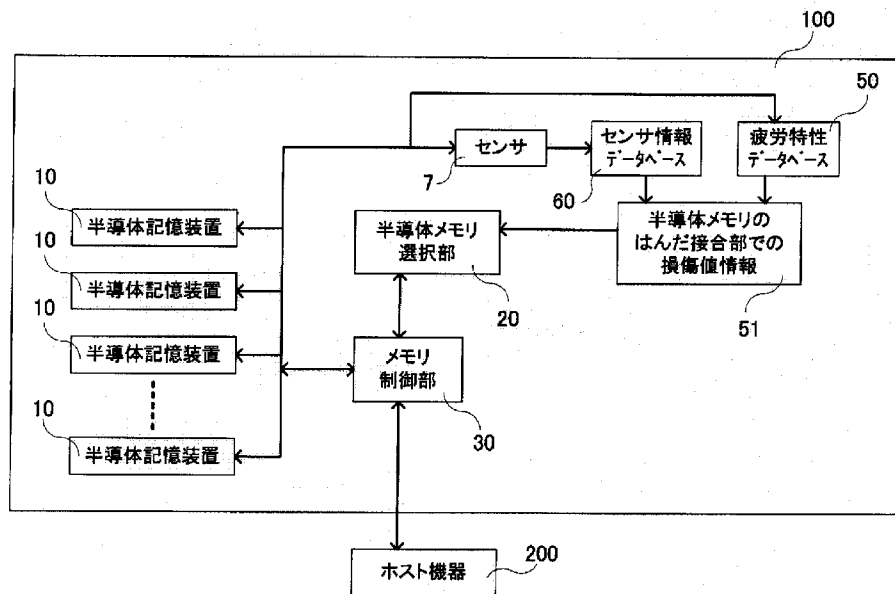
[図6]



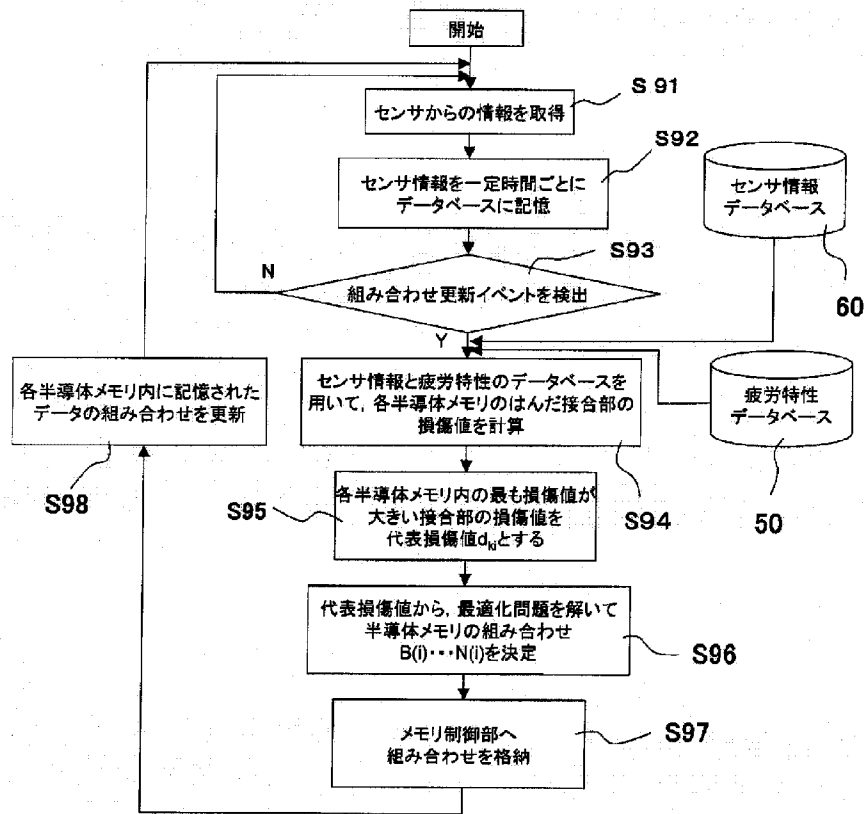
[図7]



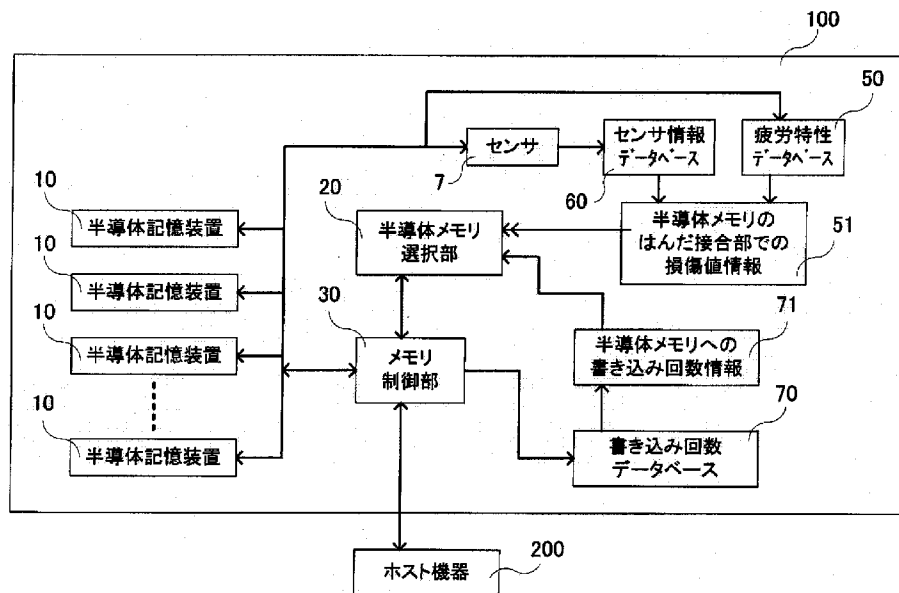
[図8]



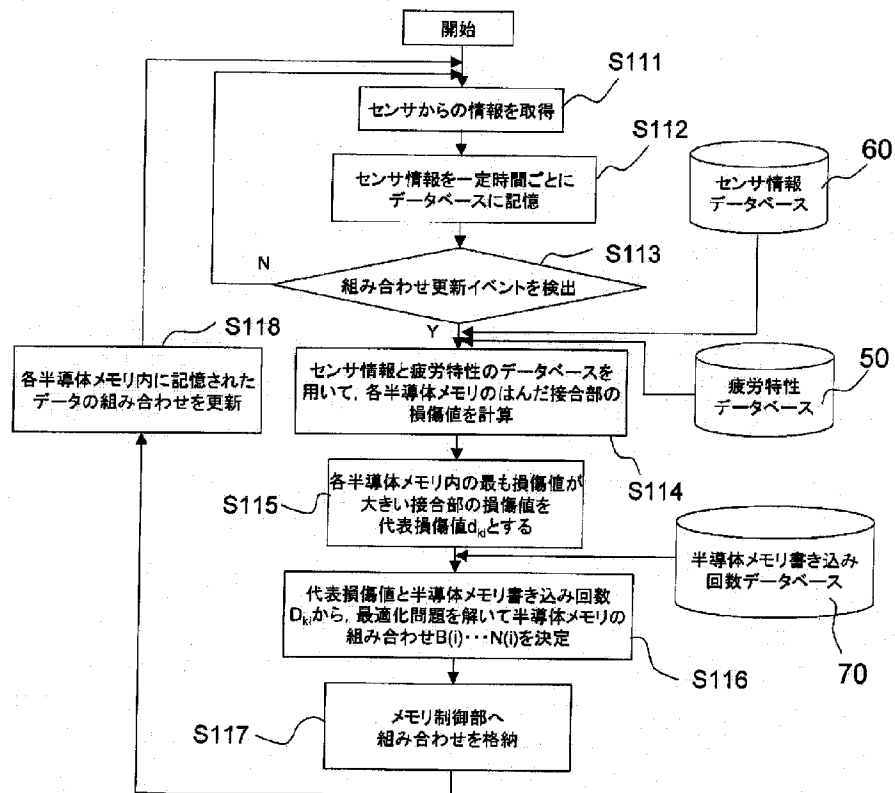
[図9]



[図10]



[図11]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/000084

A. CLASSIFICATION OF SUBJECT MATTER

G06F3/06(2006.01) i, G06F12/16(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, G06F3/08, G06F12/16

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	1996-2010
Kokai Jitsuyo Shinan Koho	1971-2010	Toroku Jitsuyo Shinan Koho	1994-2010

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2003-330815 A (NEC System Technologies, Ltd.), 21 November 2003 (21.11.2003), paragraphs [0021] to [0022]; fig. 1 (Family: none)	1-9
A	JP 2008-204041 A (Hitachi, Ltd.), 04 September 2008 (04.09.2008), paragraphs [0030] to [0040]; fig. 1A & US 2008/0201392 A1	1-9

☐ 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
12 April, 2010 (12.04.10)

Date of mailing of the international search report
20 April, 2010 (20.04.10)

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. G06F3/06(2006.01)i, G06F12/16(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. G06F3/06, G06F3/08, G06F12/16

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 0 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 0 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 0 年

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

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2003-330815 A (エヌイーシーシステムテクノロジー株式会社) 2003.11.21, 段落【0021】～【0022】、図1 (ファミリーなし)	1 - 9
A	JP 2008-204041 A (株式会社日立製作所) 2008.09.04, 段落【0030】～【0040】、図1 A & US 2008/0201392 A1	1 - 9

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

1 2 . 0 4 . 2 0 1 0

国際調査報告の発送日

2 0 . 0 4 . 2 0 1 0

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)

堀江 義隆

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 8 6

5 N

9 1 7 2