(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
2012 年 10 月 4 日(04.10.2012)

WIPO | PCT

(10) 国際公開番号
WO 2012/132408 A1

(51) 国際特許分類:
G06F 13/10 (2006.01) *G06F 3/06* (2006.01)

(21) 国際出願番号: PCT/JP2012/002116

(22) 国際出願日: 2012 年 3 月 27 日(27.03.2012)

(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 2011-081205 2011 年 3 月 31 日(31.03.2011) JP

(71) 出願人 (米国を除く全ての指定国について): パナソニック株式会社(PANASONIC CORPORATION) [JP/JP]; 〒5718501 大阪府門真市大字門真１００６番地 Osaka (JP).

(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 大坪 紹二(OHTSUBO, Shohji). 阿部 敏久(ABE, Toshihisa). 幸裕 弘(YUKI, Yasuhiro). 寺田 吉希(TERADA, Yoshiki). 廣瀬 勝彦(HIROSE, Katsuhiko).

(74) 代理人: 新居 広守(NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島５丁目３番１０号タナカ・イトーピア新大阪ビル６階新居国際特許事務所内 Osaka (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, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

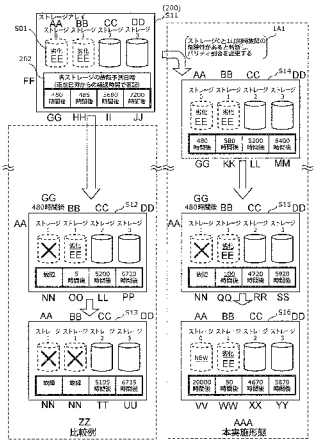
[続葉有]

WO 2012/132408 A1

(54) **Title**: ARRAY MANAGEMENT APPARATUS, ARRAY MANAGEMENT METHOD, COMPUTER PROGRAM, AND INTEGRATED CIRCUIT

(54) 発明の名称 : アレイ管理装置、アレイ管理方法、コンピュータプログラム、集積回路

[図11]



200... TABLEAU DE MÉMOIRES
1A1... LES MÉMOIRES 0 ET 1 SONT DÉTERMINÉES COMME PRÉSENTANT UN RISQUE DE DYSFONCTIONNEMENT EN MÊME TEMPS, ET LES RAPPORTS DE PARITÉ DE CELLES-CI SONT CHANGÉS
AA... MÉMOIRE 0
BB... MÉMOIRE 1
CC... MÉMOIRE 2
DD... MÉMOIRE 3
EE... DÉTÉRIORATION
FF... DATE/HEURE PRÉDITE DE DYSFONCTIONNEMENT DE CHACUNE DES MÉMOIRES (EXPRIMÉE SOUS LA FORME D'UN TEMPS ÉCOULÉ À PARTIR DE LA DATE/HEURE ACTUELLE)
GG... 480 HEURES PLUS TARD
HH... 485 HEURES PLUS TARD
II... 5680 HEURES PLUS TARD
JJ... 7200 HEURES PLUS TARD
KK... 580 HEURES PLUS TARD
LL... 5200 HEURES PLUS TARD
MM... 6400 HEURES PLUS TARD
NN... DYSFONCTIONNEMENT
OO... 5 HEURES PLUS TARD
PP... 6720 HEURES PLUS TARD
QQ... 100 HEURES PLUS TARD
RR... 4720 HEURES PLUS TARD
SS... 5920 HEURES PLUS TARD
TT... 5195 HEURES PLUS TARD
UU... 6715 HEURES PLUS TARD
VV... 20000 HEURES PLUS TARD
WW... 50 HEURES PLUS TARD
XX... 4670 HEURES PLUS TARD
YY... 5870 HEURES PLUS TARD
ZZ... EXEMPLE COMPARATIF
AAA... PRÉSENT MODE DE RÉALISATION

(57) **Abstract**: An array management apparatus (10) that manages an array (200) composed of a plurality of storages is provided with: a determining unit (100x) that determines whether a possibility of data stored in the array (200) being lost, due to at least two storages among the plurality of storages malfunctioning at the same time, exists; and a changing control unit (100y) that changes, when the possibility of data stored in the array (200) being lost is determined as existing, the malfunctioning timing of at least one storage of the aforementioned two storages. This array management apparatus (10) is thereby able to prevent data stored in the array from getting lost.

(57) 要約: 複数のストレージが構成するアレイ（２００）を管理するアレイ管理装置（１０）が、前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、アレイ（２００）が保持するデータが消失する可能性があるか否かを判断する判断部（１００ｘ）と、アレイ（２００）が保持するデータが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、変更制御部（１００ｙ）とを備え、アレイが保持するデータの消失を防ぐことができる。

(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, RW, 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 ：

アレイ管理装置、アレイ管理方法、コンピュータプログラム、集積回路

技術分野

[0001]　本発明は、複数のストレージが構成するアレイを管理するアレイ管理装置に関する。

背景技術

[0002]　従来技術として、複数のストレージが構成するアレイを管理するアレイ管理装置がある（特許文献１などを参照）。

[0003]　アレイでは、例えば、ＲＡＩＤ（Redundant Arrays of Inexpensive Disks）５、ＲＡＩＤ６などの技術により、データが保持される。

先行技術文献

特許文献

[0004] 特許文献1：特開平８－１３７６３３号公報

発明の概要

発明が解決しようとする課題

[0005]　しかしながら、上記従来技術では、アレイの状態が、将来、複数のストレージが同時期に故障することによって、アレイが保持するデータが消失してしまう状態（パリティを使用しても、アレイが保持するデータを復旧することができない状態のこと。ＲＡＩＤ５の場合には、２台以上のストレージが同時故障する状態。ＲＡＩＤ６の場合には、３台以上のストレージが同時故障する状態）になる可能性がある場合が生じる。この場合に、同時故障する可能性がある複数のストレージの故障時期をずらすことができず、アレイが保持するデータが消失する可能性が高くなってしまっていたという課題がある。

[0006]　本発明は、上記従来技術の課題を解決するものであり、アレイが保持するデータの消失を防ぐことを目的とする。

課題を解決するための手段

[0007] 上記目的を達成するために、本発明のアレイ管理は、複数のストレージが構成するアレイを管理するアレイ管理装置であって、前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断部と、前記アレイが保持する前記データが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一方の故障時期を変更する、故障時期変更部とを備えるアレイ管理装置である。

[0008] つまり、例えば、前記故障時期変更手段は、前記各ストレージに対して書き込む、パリティとデータとの割合を変更することによって、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する。

[0009] これにより、将来、複数のストレージが同時期に故障することによって、アレイが保持するデータが消失する可能性があると判断した場合には、同時故障する可能性がある複数のストレージの故障時期をずらされ、データの消失を防ぐことができる。

[0010] なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたは記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラムおよび記録媒体の任意な組み合わせで実現されてもよい。

発明の効果

[0011] 将来、複数のストレージが同時期に故障することによって、アレイが保持するデータが消失する可能性があると判断した場合には、同時故障する可能性がある複数のストレージの故障時期をずらす。このことで、アレイが保持するデータ消失を防ぐことができる。

図面の簡単な説明

[0012] [図1]図１は、本システムを示す図である。

[図2]図２は、本システムを示す図である。

[図3]図３は、ストレージのアドレス空間の概念図である。

[図4]図４は、４つのストレージなどを示す図である。

[図5]図５は、４つのストレージについての表を示す図である。

[図6]図６は、ストレージなどを示す図である

[図7]図７は、アドレス空間などを示す図である。

[図8]図８は、本システムを示す図である。

[図9]図９は、同時故障回避の処理のフローチャートである。

[図10A]図１０Ａは、Ｓ１４０２での動作の詳細を示すフローチャートである。

[図10B]図１０Ｂは、２つの時刻の間の差１Ｆ１、１Ｆ２などを示す図である。

[図10C]図１０Ｃは、２つの時刻の間の差１Ｆ３、１Ｆ４などを示す図である。

[図11]図１１は、故障が生じるまでの時間などを示す図である。

[図12]図１２は、アレイ管理部の構成を示す図である。

[図13]図１３は、本システムの動作のフローチャートである。

[図14]図１４は、本システムの動作のフローチャートである。

[図15]図１５は、処理のそれぞれの段階での表を示す図である。

[図16]図１６は、Ｓ１１０２における動作の詳細を示すフローチャートである。

[図17]図１７は、Ｓ１２０３における動作の詳細を示すフローチャートである。

[図18]図１８は、２つのグラフなどを示す図である。

[図19]図１９は、Ｓ１４０１での動作の詳細を示すフローチャートである。

[図20]図２０は、Ｓ１５０１での動作の詳細を示すフローチャートである。

[図21]図２１は、Ｓ１５０３での動作の詳細を示すフローチャートである。

[図22]図２２は、データ量のグラフを示す図である。

[図23]図２３は、３つのグラフなどを示す図である。

[図24]図２４は、Ｓ１４０４での動作の詳細を示すフローチャートである。

[図25]図２５は、Ｓ１９０１での動作の詳細を示すフローチャートである。

[図26]図２６は、Ｓ２００２での動作の詳細を示すフローチャートである。

[図27]図２７は、Ｓ２００３での動作の詳細を示すフローチャートである。

[図28]図２８は、上欄、中央欄、下欄の３つの欄の表を示す図である。

[図29]図２９は、保存可能なデータ量などを示す表である。

[図30]図３０は、発生頻度などを示す表である。

[図31]図３１は、保存可能なデータ量などを示す図である。

[図32]図３２は、処理効率などを示す図である。

[図33]図３３は、ブロック書き換え量のグラフを示す図である。

[図34]図３４は、ブロック書き換え量などを示す図である。

[図35A]図３５Ａは、パリティ割合などを示す図である。

[図35B]図３５Ｂは、４つのストレージなどを示す図である。

[図35C]図３５Ｃは、書き換え量などを示す表を示す図である。

[図36]図３６は、パリティ割合などを示す表を示す図である。

[図37]図３７は、３つのグラフなどを示す図である。

[図38]図３８は、４つのグラフなどを示す図である。

[図39]図３９は、Ｓ２００４での動作の詳細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

[0013] 以下、本発明の実施形態におけるシステム１について、図面を参照しながら説明する。

[0014] 実施形態のアレイ管理装置１０は、複数のストレージ（ストレージ２１、２２など）が構成するアレイ２００を管理するアレイ管理装置１０であって、前記複数のストレージのうち少なくとも２台のストレージ（α、β）が同時期に故障することによって、前記アレイ２００が保持するデータが消失する可能性があるか否かを判断する判断部１００ｘと（閾値１Ｆ２、差１Ｆ１などを参照）、アレイ２００が保持するデータが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つ（例えばα）の故障時期（日時１ｔを参照）を変更する故障時期変更部１

００ｙとを備えるアレイ管理装置１０である。

[0015] そして、故障時期変更部１００ｙは、各ストレージに対して書き込む、パリティ（領域２０１ｐ参照）とデータ（領域２０１ｐ参照）との割合を変更することによって、前記少なくとも２台のストレージのうちの少なくとも一つ（例えばα）の故障時期を変更する。

[0016] 例えば、前記判断部１００ｘは、前記複数のストレージのうち少なくとも２台のストレージが同時期に故障する可能性があるか否かを判断する同時故障判断部（第１部分１００ｘ１）と、前記少なくとも２台のストレージが同時期に故障すると判断された場合に、前記少なくとも２台のストレージが同時期に故障することによって、前記アレイ２００が保持するデータが消失する可能性があるか否かを判断するアレイ故障判断部（第２部分１００ｘ２）とを含んでもよい。

[0017] 例えば、前記アレイ２００により保持される前記データの消失は、第１の前記ストレージが故障する日時が第１の日時である場合には、発生し難く、第２の日時である場合には発生し易く、第１の前記ストレージは、記憶されるデータとパリティとの全体に占める、前記パリティの割合が、第１の割合である場合と、第２の割合である場合とがあるストレージであり、第１の前記ストレージの前記日時は、前記第１の割合である場合には、消失が発生し難い前記第１の日時であり、前記第２の割合である場合には、発生し易い前記第２の日時であり、前記判断部１００ｘは、２つの割合のうちの一方の前記割合を前記第１の割合と特定し、他方の前記割合を前記第２の割合と特定し、前記故障時期変更部１００ｙは、第１の前記ストレージでの前記割合を、特定された一方の前記割合にさせ、特定された他方の前記割合にはさせなくてもよい。

[0018] 例えば、消失が発生し難い前記第１の日時は、第２の前記ストレージが故障する日時から閾値以上に離れた日時であり、消失が発生し易い前記第２の日時は、第２の前記ストレージが故障する日時から閾値以内の日時でもよい。

[0019] なお、これらの全般的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたは記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラムまたは記録媒体の任意な組み合わせで実現されてもよい。

[0020] 以下、詳しく説明される。

[0021] （同時故障回避）

図１は、システム１の実施イメージ図である。

[0022] 図２は、システム１の構成図である。

[0023] システム１は、アレイ管理装置１０と、アレイ２００（図２参照）とを含む。

[0024] アレイ管理装置１０は、アレイ２００の動作を制御する。

[0025] なお、アレイ管理装置１０は、図１に示されるように、例えば、ＤＶＤ（Digital Video Disc）レコーダ、ＢＤ（Blu-ray Disc）レコーダの全体または一部などでもよい。つまり、例えば、アレイ管理装置１０は、放送波から受信される動画のデータなどのデータを記録するストレージ（図１のストレージ２５、２６などを参照）備えるストレージ装置の全体または一部などでもよい。

[0026] アレイ（ストレージアレイ）２００は、データを保持する。

[0027] なお、例えば、アレイ管理装置１０は、アレイ２００により保持されるデータを、アレイ２００と、アレイ管理装置１０との間で、転送させてもよい。

[0028] なお、アレイ管理装置１０は、例えば、この転送をさせる下位側転送制御部５５（図２）を備えてもよい。

[0029] そして、例えば、アレイ管理装置１０は、アレイ２００により保持されるデータを、アレイ管理装置１０と、保持されるデータの処理をする処理部との間で、転送させてもよい。

[0030] アレイ管理装置１０は、例えば、この転送をさせる上位側転送制御部５４（図２）を備えてもよい。

[0031] そして、例えば、上述の処理部は、アレイ管理装置１０の外部にあるパーソナルコンピュータなどの装置などで動作するオペレーションシステムにおけるファイルシステムなどでもよい。

[0032] なお、アレイ２００により保持されるデータは、例えば、アレイ２００と、このオペレーションシステムの上で動作するアプリケーションとの間を、上述ファイルシステムなどを介して転送されて、このアプリケーションにより利用されてもよい。

[0033] アレイ２００は、ストレージ２０１（図２）を複数個、含んでなり、これら複数のストレージ２０１を用いて、データの保持を行う。

[0034] 含まれる複数のストレージ２０１のうちの少なくとも１つは、例えば、図１のストレージ２５、２６などで示されるような、アレイ管理装置１０に設けられたストレージなどでもよい。

[0035] また、少なくとも１つは、例えば、図１のストレージ２１～２４で示されるような、アレイ管理装置１０の外部にあるストレージでもよい。

[0036] 具体的には、少なくとも１つは、ストレージ２１～２３で示されるような、アレイ管理装置１０が設けられた住宅と同じ住宅に設けられたストレージでもよい。

[0037] また、少なくとも１つは、ストレージ２４で示されるような、他の住宅に設けられたストレージでもよい。なお、例えば、他の住宅は、親夫婦の住む住宅である一方で、アレイ管理装置１０が設けられた住宅は、子供夫婦が住む住宅でもよい、なお、他の住宅にあるストレージ２４は、例えば、インターネットなどであるネットワーク１ＤＬを介して、アレイ管理装置１０からの、動作の制御を受けてもよい。

[0038] なお、図１は、単なる一例である。システム１においては、図１に示される通りでもよいし、図１に示される通りでなくてもよい。

[0039] 図３は、ストレージのアドレス空間の概念図である。

[0040] それぞれのストレージ２０１（例えば、図３のストレージ３）の記憶領域は、複数のブロック１ｘＢ２（ｉ個のブロックＢｋ０３～Ｂｋｉ３などを参

照）に分かれる。

[0041] そして、アレイ２００は、それぞれのストレージ２０１（図３のストレージ０～３）におけるブロック（ｂｋ００、ｂｋ０１、ｂｋ０２、ｂｋ０３参照）が集まってなる記憶領域であるストライプ１ｘＢ１（図３のストライプ０など）を有する。

[0042] 図４は、ストレージ２０１に対するデータ・パリティ格納の概念図である。

[0043] ストライプ１ｘＢ１（例えば、図４のストライプｉ）に含まれるブロック１ｘＢ２（図３）としては、データのブロック１ｘＢ２ｄ（ブロックＢＤｉ０、ＢＤｉ１、ＢＤｉ２などを参照）と、パリティのブロック１ｘＢ２ｐ（ブロックＢＰｉ０などを参照）とがある。

[0044] ブロック１ｘＢ２ｐに記憶されるパリティは、そのブロック１ｘＢ２ｐのあるストレージ２０１以外の他のストレージ２０１（例えば図４のストレージ２）で故障が生じた場合に用いられる。

[0045] つまり、故障が生じた場合に、そのブロック１ｘＢ２ｐに記憶されるパリティと、そのブロック１ｘＢ２ｐのストライプ１ｘＢ１における、故障が生じてないそれぞれのストレージ（ストレージ０、１）でのデータ（ブロックＢｄｉ０、Ｂｄｉ１参照）とが用いられる。

[0046] すなわち、これらから、そのストライプ１ｘＢ１における、故障が生じたストレージ（ストレージ２）におけるデータ（ブロックＢＤｉ２のデータ）が生成される。

[0047] このため、ストライプ１ｘＢ１におけるブロック１ｘＢ２ｄへとデータが書き込まれる際には、そのブロック１ｘＢ２ｄにデータが書き込まれるのと共に、そのストライプ１ｘＢ１のブロック１ｘＢ２ｐに対して、パリティが書き込まれる。

[0048] そして、ストライプ１ｘＢ１における、データが記憶されるブロック１ｘＢ２ｄの個数は、例えば、図４における３個などのように、２個以上である。

[0049] このため、パリティのブロック１ｘＢ２ｐへの書き込みが行われる回数は、その書き込みが、何れの、データのブロック１ｘＢ２ｄへの書き込みがされる際にも行われて、比較的多い回数である。

[0050] つまり、それぞれの、データのブロック１ｘＢ２ｄへの書き込みの回数は、比較的少ない。

[0051] パリティのブロック１ｘＢ２ｐ（例えばＢＰｉ０）への書き込みの回数は、例えば、それぞれの、データのブロック１ｘＢ２ｄ（ＢＤｉ０～ＢＤｉ２）への書き込みの回数が合計された回数などである。つまり、例えば、図４の例では、データのブロック１ｘＢ２ｄの個数が３個であり、データのブロック１ｘＢ２ｄへの書き込みの回数の３倍などである。

[0052] 図５は、それぞれのストレージ２０１でのパリティ割合２０１ｒ（第７行）などを示す表の図である。

[0053] 図５の表におけるそれぞれの列では、その列のストレージ２０１でのことが示される。

[0054] 先述の通り、ストレージ２０１におけるブロック１ｘＢ２としては、データのブロック１ｘＢ２ｄと、パリティのブロック１ｘＢ２ｐとがある。

[0055] 図６は、ストレージ２０１における、データの領域２０１ｄと、パリティの領域２０１ｐなどを示す図である。

[0056] つまり、ストレージ２０１の記憶領域の全体である領域２０１Ａは、データの領域２０１ｄと、パリティの領域２０１ｐとに分かれる。データの領域２０１ｄは、それぞれの、データのブロック１ｘＢ２ｄの領域が集まってなる領域である（図４参照）。そして、パリティの領域２０１ｐは、それぞれの、パリティのブロック１ｘＢ２ｐの領域が集まってなる領域である。

[0057] パリティ割合２０１ｒ（図５）は、全体の領域２０１Ａの大きさに対する、パリティの領域２０１ｐの割合をいう。つまり、パリティ割合２０１ｒは、例えば、ストレージ２０１に含まれる、データのブロック１ｘＢ２ｄの個数と、パリティのブロック１ｘＢ２ｐの個数との合計に占める、パリティのブロック１ｘＢ２ｐの個数の割合などである。

[0058] ストレージ２０１のパリティ割合２０１ｒが、比較的低い割合である場合には、パリティのブロック１ｘＢ２ｐが比較的少ない。パリティブロックは、データブロックと比較して、アクセス頻度が高いため、パリティ割合が少なければ、パリティ割合が多いストレージと比較して、ストレージの劣化が起こりにくく、ストレージ２０１が故障する日時１ｔ（図１０Ｂ）が、比較的遅い（図示される日時１ｔ２を参照）。

[0059] 一方で、パリティ割合２０１ｒが、比較的高い方の割合である場合には、逆に、故障する日時１ｔが、比較的早い（日時１ｔ１を参照）。

[0060] こうして、ストレージ２０１のパリティ割合２０１ｒが一方の割合である場合における、ストレージ２０１の、故障の日時１ｔは、他方の割合である場合における日時１ｔとは異なる。

[0061] 図７は、アレイ２００におけるアドレス空間の概念図である。

[0062] なお、上述の通りであるので、アレイ２００におけるアドレス空間２００ａｓは、それぞれの、ストライプ１ｘＢ１（図３：例えば図４のストライプ０、１など）での部分に分かれる。そして、それぞれのストライプ１ｘＢ１（例えばストライプ１）での部分は、さらに、そのストライプ１ｘＢ１における、それぞれのブロック１ｘＢ２（図７におけるブロックＢＤ１０、ＢＤ１１、ＢＤ１２を参照）での部分に分かれる。

[0063] 図８は、システム１を示す図である。

[0064] アレイ管理装置１０は、アレイ管理部１００（図８）を備える。

[0065] アレイ管理部１００は、パリティ割合管理部１０１と、パリティ割合算出部１０２と、パリティ割合変更部１０３と、故障予測部１０４とを備える。

[0066] なお、アレイ管理装置１０は、図２に示されるように、例えば、ＣＰＵ５２、ＲＡＭ５３、ＲＯＭ５１などを含んでなるコンピュータを含んでもよい。

[0067] そして、アレイ管理部１００（アレイ管理部１００が備える、パリティ割合管理部１０１などのそれぞれの機能ブロック）は、このコンピュータにより、コンピュータプログラム１０１ｐｇ（図２）が実行されることで、アレ

イ管理装置１０に実現される機能の機能ブロックでもよい。

[0068] なお、このコンピュータプログラム１０１ｐｇは、例えば、ＲＯＭ５１により記憶され、記憶されたコンピュータプログラム１０１ｐｇが、上記のコンピュータにより実行されてもよい。

[0069] なお、図２では、説明の便宜上、パリティ割合管理部１０１などが、このＲＯＭ５１の位置に図示される。

[0070] 図９は、同時故障回避の処理のフローチャートである。

[0071] Ｓ１４０１では、それぞれのストレージ２０１で故障が生じる日時である故障予測日時が故障予測部１０４により算出される。Ｓ１４０２では、同時故障が発生する可能性があるか否かが、パリティ割合算出部１０２により算出される。

[0072] 図１０Ａは、同時故障発生の危険性判断の処理のフローチャートである。図９のＳ１４０２では、例えば、この図１０Ａの処理がされてもよい。

[0073] 図１０Ｂは、第１の日時１ｔ１、第２の日時１ｔ２などを示す図である。

[0074] Ｓ１８０１では、最も早い故障予測日時（図１０Ｂの日時１ｔを参照）が算出されたストレージ２０１が、第１のストレージα（図１０Ｂ参照）として特定される。

[0075] Ｓ１８０２（図１０Ａ）では、第２のストレージβが特定される。第２のストレージβは、特定された、複数のストレージ２０１の故障予測日時のうちで、Ｍ＋１番目に早い故障予測日時（図１０Ｂの日時１Ｂ）が算出されたストレージ２０１である。

[0076] このＭは、アレイ２００における故障許容台数（最大故障許容台数）である。故障許容台数は、その台数以下の台数のストレージ２０１で故障が生じるだけであれば、アレイ２００により保持されるデータの消失が生じず、その台数よりも多い台数での故障が生じた場合に、消失が生じる台数である。つまり、故障許容台数は、例えば、アレイ２００が、ＲＡＩＤ５での動作をする際には、１台であり、ＲＡＩＤ６での動作をする際には、２台である。

[0077] つまり、第２のストレージβは、その第２のストレージβが故障するまで

に生じた、それぞれのストレージ２０１での故障が継続するままで、その第２のストレージβも故障すると、データの消失が発生する１以上のストレージ２０１のうちで、故障予測日時が最も早いストレージ２０１である。

[0078] Ｓ１８０３（図１０Ａ）では、第１のストレージαの故障予測日時（日時１ｔ）と、第２のストレージβの故障予測日時（日時１Ｂ）との間の差１Ｆ１が、閾値１Ｆ２以内であるか否かが判定される。

[0079] つまり、第１のストレージαの故障予測日時（日時１ｔ）が、第２のストレージβの故障予測日時（日時１Ｂ）から、閾値１Ｆ２以内で、第２のストレージβの故障予測日時（日時１Ｂ）と同時期であるか否かが判定される。

[0080] 閾値１Ｆ２は、上述の差１Ｆ１が、その閾値１Ｆ２以内である場合には（日時１ｔ２参照）、第１のストレージαの故障予測時刻（日時１ｔ）以後で、かつ、第２のストレージβが故障する故障予測時刻（日時１Ｂ）よりも前に、この第１のストレージαなどの、故障したストレージ２０１の交換、修理などがされず、データの消失が発生する（と推測される）値の閾値である。

[0081] つまり、閾値１Ｆ２は、上述の差１Ｆ１が、その閾値１Ｆ２より大きい場合には、その大きな差１Ｆ１の期間のうちに、故障したストレージ２０１の交換などがされて、データの消失が発生しない（と推測される）値の閾値である。

[0082] Ｓ１８０４（図１０Ａ）では、上述の差１Ｆ１が、閾値１Ｆ２以内と判定される場合に（Ｓ１８０３：ＹＥＳ）、データの消失が発生する（と推測される）ことが判定される。他方、差１Ｆ１が、閾値１Ｆ２以内ではないと判定される場合には（Ｓ１８０３：ＮＯ）、データの消失が発生しない（と推測される）ことが判定される。

[0083] 図１０Ｃは、差１Ｆ３、差１Ｆ４などを示す図である。

[0084] なお、例えば、図１０Ｃで示される通りでもよい。

[0085] つまり、上述の第１のストレージαは、例えば、図１０Ｃで書かれる第１のストレージγ１で、第２のストレージβは、第２のストレージγ２でもよ

い。そして、第３のストレージγは、その第３のストレージγの故障予測日時が、第１のストレージγ１の故障予測日時以後で、かつ、第２のストレージγ２以前であるストレージ２０１である。

[0086] そして、例えば、第１のストレージγ１と、第３のストレージγとの間の差１Ｆ３が、所定の閾値より大きい場合にのみ、消失が発生しないとの判定がされ（Ｓ１８０３：ＮＯ）、閾値より小さい場合には、消失が発生するとの判定がされてもよい（Ｓ１８０３：ＹＥＳ）。

[0087] 同様に、例えば、第３のストレージγと、第２のストレージγ２との間の差１Ｆ４が、所定の閾値より大きい場合にのみ、消失が発生しないとの判定がされ（Ｓ１８０３：ＮＯ）、閾値より小さい場合には、消失が発生するとの判定がされてもよい（Ｓ１８０３：ＹＥＳ）。

[0088] Ｓ１４０３（図９）では、データの消失が発生する可能性があるか否かの判定（Ｓ１４０２）に基づいて、行われる処理が選択される。

[0089] つまり、消失が発生しないと判定される場合には（Ｓ１４０３：ＮＯ）、現在の直前に行われた、アレイ２００への制御と同じ制御が、パリティ割合変更部１０３により行われてもよい。

[0090] 他方、Ｓ１４０４では、消失が発生すると判定される場合に（Ｓ１４０３：ＹＥＳ）、現在の直前に行われた、アレイ２００への制御とは別の制御が、パリティ割合変更部１０３により行われる。別の制御は、その制御がされれば、第１のストレージαが故障する日時１ｔが、第２のストレージβが故障する日時１Ｂから閾値１Ｆ２より離れていて、同時期ではなく、データの消失が発生しない（と推測される）制御である。つまり、別の制御は、例えば、その制御がされれば、アレイ２００に含まれるストレージ２０１におけるパリティ割合２０１ｒ１が、現在の直前に行われた制御がされる際での、そのストレージ２０１におけるパリティ割合２０１ｒ２と違っていて、上述の差１Ｆ１が比較的大きく、故障の時期が同時期ではない制御ではある。

[0091] 図１１は、同時故障回避のイメージ図である。

[0092] 図１１における、本実施形態の欄で示されるように、上述の技術によれば

、データの消失が生じると推測されることが判定される場合には（Ｓ１４０３：ＹＥＳ、図１１の処理１Ａ１）、別の制御がされて（Ｓ１４０４）、行われる制御が変更されることにより、データの消失が発生してしまうことが回避できる（図１１のＳ１４～Ｓ１６）。

[0093] これに対して、図１１に示される比較例は、このような本実施形態とは別の、想定されるその他の技術におけるケースを示す。この比較例では、上述された、同時期か否かの判定（図９：Ｓ１４０２）がされず、行われる制御が変更されなくて、データの消失が発生してしまう（図１１のＳ１２～Ｓ１３）。

[0094] こうして、比較例では、データの消失が発生してしまう一方で、本実施形態では、データの消失の発生が回避できる。

[0095] 図１２は、アレイ管理装置１０を示す図である。

[0096] なお、アレイ管理装置１０は、例えば、判断部１００ｘと、変更制御部（故障時期偏向部）１００ｙとを備えてもよい。

[0097] そして、判断部１００ｘは、上述された、パリティ割合算出部１０２（図８など）と、故障予測部１０４とを含み、変更制御部１００ｙは、パリティ割合管理部１０１と、パリティ割合変更部１０３とを含んでもよい。

[0098] つまり、例えば、パリティ割合算出部１０２により、パリティ割合算出部１０２の上述の動作がされることにより、その動作が、パリティ割合算出部１０２を用いて、判断部１００ｘにより行われてもよい。なお、故障予測部１０４などについても、それぞれ、この例と同様である。

[0099] （受付データ書き込み）

図１３は、受付データ書き込みの処理のフローチャートである。

[0100] 本システム１においては、例えば、図１３で示される処理が行われてもよい。

[0101] Ｓ１００１では、アレイ２００に対して書き込むことが要求されるデータが、アレイ管理装置１０により受け付けられる。Ｓ１００２では、受け付けられたデータが記録されるストライプ１ｘＢ１（図３参照）が特定される。

Ｓ１００３では、そのデータのパリティが、アレイ管理装置１０により生成される。Ｓ１００４では、パリティ割合管理部１０１により記憶される情報１０１Ｉ（図８）に基づく処理が行われる。

[0102] ここで、情報１０１Ｉは、特定されたストライプ１ｘＢ１における、受け付けられたデータが記録される（１以上の）ブロック１ｘＢ２と、生成された、そのデータのパリティが記録されるブロック１ｘＢ２とを、それぞれ特定する情報である。

[0103] なお、情報１０１Ｉは、例えば、図５の表の第２行～第５行におけるデータ構造を有しても良い。情報１０１Ｉは、図１１に示される表などでもよい。

[0104] Ｓ１００５では、受け付けられたデータが、特定される、そのデータが記録されるべき（１以上の）ブロック１ｘＢ２ｄに記録され、生成されたパリティが、特定される、そのパリティが記録されるべきブロック１ｘＢ２ｐに記録される。

[0105] （初期化）

図１４は、パリティ割合の初期化の処理のフローチャートである。

[0106] 図１５は、データ・パリティ対応情報作成作業における対応表（先述の図５参照）の遷移の例を示す図である。

[0107] Ｓ１１０１では、それぞれのストレージ２０１におけるパリティ割合２０１ｒの目標値（図１５のそれぞれの表における第７行）が、アレイ管理装置１０により特定される。なお、各ストレージのパリティ割合の目標値は、各ストレージにおいて非均等であっても良い。Ｓ１１０２では、特定されたそれぞれの目標値に基づいて、情報１０１Ｉが、アレイ管理装置１０により作成される。

[0108] 図１６は、パリティ割合の初期化の処理のフローチャートである。先述された図１４のＳ１１０２では、例えば、この図１６の処理がされてもよい。

[0109] Ｓ１２０１では、情報１０１Ｉの表（図５）における、それぞれのストレージ（各列）のパリティ割合２０１ｒ（第７行）が０に初期化される。Ｓ１

２０２では、最初のストライプ（例えば、図５のストライプ０）が、選択ストライプとして選択される。Ｓ１２０３では、情報１０１Ｉの表における、選択された選択ストライプの行（例えば、ストライプ０の行）の、それぞれのストレージ（各列）の欄の値が特定される。

[0110] Ｓ１２０４では、全てのストライプの処理が終了したか否かが判定される。Ｓ１２０５では、終了していないと判定された場合に（Ｓ１２０４：ＮＯ）、選択ストライプとして、直前のＳ１２０３の処理における選択ストライプの次のストライプが選択される（図１５の処理１Ｃ１～１Ｃ２などを参照）。一方で、全てのストライプの処理が終了したと判定される場合には（Ｓ１２０４：ＹＥＳ）、図１６の処理が終了する。

[0111] 図１７は、ストライプに対するデータ・パリティ対応情報割り当ての処理のフローチャートである。図１６のＳ１２０３では、例えば、この図１７の処理がされてもよい。

[0112] Ｓ１３０１では、所定の台数のストレージ２０１が選択される。

[0113] つまり、先述された、パリティ割合２０１ｒの目標値（Ｓ１１０１）から、現在の割合（図１５の表に第６行）が減じられた値（第８行）の順序が、一番大きい順序から、上述の所定の台数番目の順序までの、それぞれの順序のストレージ２０１が選択される。

[0114] なお、上述の所定の台数は、アレイ２００がＲＡＩＤ５で動作するならば、１台であり、ＲＡＩＤ６で動作するならば、２台である。

[0115] Ｓ１３０２では、この図１７の処理がされる対象のストライプにおける、選択された、所定の台数のストレージ２０１のうちのそれぞれのストレージ２０１のブロック１ｘＢ２が、パリティのブロック１ｘＢ２ｐとして特定される（図３、図４など参照）。

[0116] そして、このＳ１３０２では、他のそれぞれのストレージ２０１のブロック１ｘＢ２が、データのブロック１ｘＢ２ｄとして特定される（図１５の処理１Ｃ１、１Ｃ２などを参照）。

[0117] Ｓ１３０３では、上述のＳ１３０１、Ｓ１３０２での特定の結果が反映さ

れた後における、それぞれのストレージ２０１での、現在の割合（図１５の表の第６行）が算出される。

[0118] （故障予測日時算出の具体例）

図１８は、故障予測日時算出の処理のイメージ図である。

[0119] 図１９は、故障予測日時算出の処理のフローチャートである。先述された図９のＳ１４０１では、例えば、この図１９の処理がされてもよい。

[0120] Ｓ１５０１では、図１８の下欄に示される２つのグラフのうちの、左のグラフで示される関数ｆが特定される。

[0121] 図１８で示されるように、関数ｆは、日時と、その日時における、アレイ２００への書き込み量の累積量との関係の関数である。

[0122] Ｓ１５０２では、最初のストレージが、選択ストレージとして選択される。

[0123] Ｓ１５０３では、選択された選択ストレージについての、図１８の下欄の右のグラフで示される関数g_jが特定される。図１８で示されるように、関数g_jは、アレイ２００に対する書き込みの累積量（下欄のグラフの横軸）と、アレイ２００に対して、その累積量の書き込みがされた日時における、選択ストレージへの書き込み量の累積量との関係を特定する関数である。

[0124] Ｓ１５０４では、選択ストレージにおける保証量が特定される。保証量は、その量のデータ量までの書き込みがされるまでは、その選択ストレージで故障が生じないことが、その選択ストレージのメーカなどにより保証された量である。つまり、保証量は、この量のデータ量だけの書き込みがされた日時（に近い日時）に、故障が生じることが推定される量である。

[0125] Ｓ１５０５では、選択ストレージに対して、保証量のデータ量だけの書き込みがされる際における、アレイ２００に対する書き込みの累積量が特定される。

[0126] Ｓ１５０６では、アレイ２００に対する書き込みの累積量が、特定された上述の累積量になる日時が、その選択ストレージの故障予測日時として特定される。

[0127] Ｓ１５０７では、全てのストレージの故障予測日時が、上述のＳ１５０６で特定されたか否かが判定される。そして、されていると判定された場合には（Ｓ１５０７：ＹＥＳ）、図１９の処理が終了する。他方、Ｓ１５０８では、全てのストレージについての処理が終わってないと判定される場合に（Ｓ１５０７：ＹＥＳ）、直前のＳ１５０３～Ｓ１５０７での処理での選択ストレージの次のストレージ２０１が、次のＳ１５０４～Ｓ１５０７での処理での選択ストレージとして特定される。

[0128] 図２０は、累積書き込みデータ量の増加度合い算出の処理のフローチャートである。上述された図１９のＳ１５０１では、例えば、この図２０の処理がされてもよい。

[0129] Ｓ１６０１では、現在の日時が取得される。なお、例えば、アレイ管理装置１０に設けられた、現在の日時を計時する時計から、計時された、現在の日時が取得されてもよい。

[0130] Ｓ１６０２では、アレイ２００に対する書き込みの、現在における累積量が取得される。なお、例えば、アレイ２００などの、アレイ管理装置１０の外部に設けられた所定の記憶部により、この累積量が記憶されて、記憶された累積量が、アレイ管理装置１０へと取得されてもよい。

[0131] Ｓ１６０３では、アレイ管理装置１０の記憶領域に、取得された、現在の日時と、アレイ２００への書き込みの、現在の累積量とが、互いに対応付けて記憶される。

[0132] 図１８の下欄の２つのグラフのうちの左側のグラフにおける、それぞれの丸印は、アレイ管理装置１０の記憶領域において対応付けて記憶された、日時、および、その日時における、アレイ２００に対する書き込みの累積量を示す。

[0133] Ｓ１６０４では、それぞれの丸印での日時および累積量に基づいて、先述の関数ｆ（図１８の左のグラフを参照）が特定される。なお、この特定の処理としては、従来技術における、この種の処理における技術（回帰分析の処理など）が用いられてもよい。

[0134] 図２１は、ストレージに対する累積書き換え量の増加度合い算出の処理のフローチャートである。先述された図１９のＳ１５０３では、この図２１の処理がされてもよい。

[0135] Ｓ１７０１では、アレイ２００に対する書き込みの、現在における累積量が取得される。Ｓ１７０２では、選択ストレージ（先述）に対する書き込みの、現在の累積量が取得される。なお、例えば、Ｓ１７０２で取得されるこの累積量が、選択ストレージに設けられた所定の記憶部により記憶されて、この記憶部により記憶された累積量が、アレイ管理装置１０へと取得されてもよい。

[0136] Ｓ１７０３では、Ｓ１７０１で取得された累積量と、Ｓ１７０２で取得された累積量とが、それぞれ、アレイ管理装置１０の記憶領域に記憶される。

[0137] Ｓ１７０４では、選択ストレージの関数ｇ$_j$（図１８の右のグラフ参照）が特定される。なお、この特定の処理では、例えば、Ｓ１６０４での、関数ｆの特定の処理と同様に、回帰分析の技術などが用いられてもよい。

[0138] 図２２は、３つの故障予測関数のデータのグラフを示す図である。

[0139] 図示されるように、ストレージｊの故障予測関数は、「関数ｆ×関数ｇ$_j$」（関数ｆと関数ｇ$_j$の合成関数）である。そして、互いに異なるストレージ２０１での故障予測関数は、通常は、互いに異なる関数である。また、図示されるように、例えば、互いに異なるストレージ２０１の保証量（先述）は、互いに異なってもよい。

[0140] 図２３は、同時故障発生のイメージ図である。

[0141] 図２３では、同時故障が発生する第１のケースを示す上段の欄と、第２のケースを示す中断の欄と、第３のケースを示す下段の欄とが示される。

[0142] （変更処理の具体例）

図２４は、同時故障回避の処理のフローチャートである。図９のＳ１４０４では、例えば、この図２４の処理がされてもよい。

[0143] Ｓ１９０１では、それぞれのストレージ２０１における、適切（最適）なパリティ割合２０１ｒが算出される。

[0144] Ｓ１９０２では、情報１０１１Ｉが作成される。作成される情報１０１１Ｉは、その情報での、図５の表における、それぞれのストレージ２０１でのパリティ割合が、算出された、そのストレージ２０１での適切（最適）なパリティ割合２０１ｒである情報である。

[0145] Ｓ１９０３では、それぞれのストライプについて、適宜、スワップの処理がされる。これにより、それぞれのストレージ２０１のそれぞれのブロック１ｘＢ２について、そのブロック１ｘＢ２が、データのブロック１ｘＢ２ｄであるか、パリティのブロック１ｘＢ２ｐであるかが、作成された情報１０１１Ｉ（図５参照）に沿ったものにされて、それぞれのストレージ２０１のデータ構造が、作成された情報１０１１Ｉでのデータ構造にされる。

[0146] Ｓ１９０４では、作成された情報１０１１Ｉが、この図２４の処理がされる以前に、パリティ割合管理部１０１により記憶される情報１０１１Ｉ（図８参照）に対して上書きされて、この上書きがされて以後は、パリティ割合管理部１０１により、作成された新しい情報１０１１Ｉが記憶される。

[0147] 図２５は、最適パリティ割合算出の処理のフローチャートである。図２４のＳ１９０１では、例えば、この図２５の処理がされてもよい。

[0148] Ｓ２００１では、関数ｆ（先述の図１８の左のグラフ参照）が特定される。Ｓ２００２では、それぞれのストレージ２０１の許容残書き換え量が特定される。

[0149] 図２６は、各ストレージの許容残書き換え量算出の処理のフローチャートである。図２５のＳ２００２では、例えば、この図２６の処理がされてもよい。

[0150] Ｓ２１０１では、最初のストレージが、選択ストレージとして選択される。Ｓ２１０２では、選択ストレージの保証量が取得される。Ｓ２１０３では、選択ストレージに対する書き込みの、現在における累積量が取得される。Ｓ２１０４では、取得された保証量から、現在の累積量が減じられた量が、許容残書き換え量として特定される。

[0151] Ｓ２１０５では、全ストレージ２０１の処理が終了したか否かが判定され

る。終了したと判定される場合には（Ｓ２１０５：ＹＥＳ）、図２６の処理が終了する。一方で、Ｓ２１０６では、終了していないと判定される場合に（Ｓ２１０６：ＮＯ）、選択ストレージとして、次のストレージ２０１が選択される。

[0152] Ｓ２００３（図２６）では、関数が特定される。特定される関数は、ストレージ２０１における、パリティ割合２０１ｒと、そのパリティ割合２０１ｒでの、そのストレージ２０１への書き込み量との間の関係を特定する関数である（図３３参照）。

[0153] 図２７は、ストレージのパリティ割合とブロック書き換え量との処理のフローチャートである。図２５のＳ２００３では、例えば、この図２７の処理がされてもよい。

[0154] 図２８は、データブロック同時書き換えの説明図である。

[0155] 図２８の表では、アレイ２００への書き込みに際における第１のパターンが上欄で示され、第２のパターンが中欄で示され、第３のパターンが下欄で示される。第１のパターンでは、１個のストレージ２０１（ストレージ３）に、パリティの書き込みがされると共に、１個のストレージ２０１（ストレージ０）に、データの書き込みがされる。第２のパターンでは、１個のストレージ２０１（ストレージ３）に、パリティの書き込みがされると共に、２個のストレージ２０１（ストレージ０、１）に、データの書き込みがされる。第３のパターンでは、１個のストレージ２０１（ストレージ３）に、パリティの書き込みがされると共に、３個のストレージ２０１（ストレージ０～２）に、データの書き込みがされる。

[0156] 図２９は、同時書き換えデータブロック数と、保存データ量との関係図である。

[0157] それぞれのパターンでの、アレイ２００への書き込みでの、書き込みがされるデータ量は、例えば、図２９で示される通りである。Ｓ２２０１（図２７）では、頻度の情報が取得される。取得される、頻度の情報は、アレイ２００への書き込みとして、第１のパターンの書き込みが生じる第１の頻度（

確率）と、第２のパターンの書き込みが生じる第２の頻度（確率）と、第３のパターンの書き込みが生じる第３の頻度（確率）とをそれぞれ特定する。

[0158] 図３０は、データブロック同時書き換え発生頻度情報管理テーブルの例である。

[0159] 図３０では、取得される情報により、第１の頻度として２５％が特定され、第２の頻度として４０％が特定され、第３の頻度として３５％が特定されるケースが示される。

[0160] 図３１は、データブロック同時書き換え発生頻度に応じたデータの割り当ての図である。

[0161] アレイ２００に対して書き込みがされるのに際して、第１、第２、第３のパターンでの書き込みが、上述の第１、第２、第３の頻度（２５％、４０％、３５％）で行われる。そして、それぞれのパターンでの書き込みでは、１ＭＢ、２ＭＢ、３ＭＢの書き込みがされる。

[0162] このため、第１、第２、第３のパターンで書き込みがされるデータ量の比は、「２５％×１ＭＢ：４０％×２ＭＢ：３５％×３ＭＢ」＝「２５ＭＢ：８０ＭＢ：１０５ＭＢ」である。

[0163] このため、アレイ２００に対して、あるデータ量（２５２０ＭＢ）の書き込みがされる際における、第１、第２、第３のパターンでの書き込みのデータ量は、２５２０ＭＢ×（２５ＭＢ／（２５ＭＢ＋８０ＭＢ＋１０５ＭＢ））＝２５２０ＭＢ×（２５ＭＢ／２１０ＭＢ）＝３００ＭＢ、２５２０ＭＢ×（８０ＭＢ／２１０ＭＢ）＝９６０ＭＢ、２５２０ＭＢ×（１０５ＭＢ／２１０ＭＢ）＝１２６０ＭＢである（図３１の最下部を参照）。

[0164] つまり、第１、第２、第３のパターンでの書き込みがされるストライプの数は、３００ＭＢ／１ＭＢ＝３００個、９６０ＭＢ／２ＭＢ＝４８０個、１２６０ＭＢ／３ＭＢ＝４２０個である。

[0165] こうして、アレイ２００に対する書き込みのデータ量（２５０ＭＢ）から、第１、第２、第３のパターンでの書き込みがされるストライプ数（３００個、４８０個、４２０個）が特定される。

[0166] Ｓ２２０２（図２７）では、アレイ２００に対して、単位データ量（図３１の２５２０ＭＢ参照）の書き込みがされる際における、第１、第２、第３のパターンでの書き込みがされる第１、第２、第３のストライプ数（３００個、４８０個、４２０個を参照）が特定される。

[0167] そして、第１のパターンで、第１のストライプ数（３００個）の書き込みがされるのに際して、ストレージ２０１に対して行われる書き込みの回数は、以下の通りである。

[0168] なお、以下では、説明の便宜上、ストレージ２０１のパリティ割合２０１ｒが、ｘ％と略記される。

[0169] つまり、第１のストライプ数（３００個）のうちの、（１００－ｘ）％の割合の個数（３００個×（１００－ｘ）％）のストライプについては、それらのストライプのうちのそれぞれのストライプにおける、そのストレージ２０１のブロック１ｘＢ２は、データのブロックである。

[0170] そして、第１のパターンでの書き込みでは、図３０の上欄に示されるように、ストライプに含まれる、３つの、データのブロックのうちの１つのブロックにのみ、書き込みがされる。

[0171] このため、それらの、上述された、（３００個×（１００－ｘ）％）の個数のストライプのうちの、１／３の個数｛（３００個×（１００－ｘ）％）／３｝のストライプの書き込みにおいて、そのストレージ２０１への書き込みがされる。

[0172] 図３２は、データブロックにおける書き換え処理発生率の説明図である。

[0173] なお、上述の１／３の点については、適宜、この図３２も参照されたい。

[0174] 一方で、上述された、第１のストライプ数（３００個）のストライプのうちの、ｘ％の割合の個数（３００個××％）のストライプについては、そのストライプにおける、そのストレージ２０１のブロック１ｘＢ２が、パリティのブロック１ｘＢ２ｐである。

[0175] このため、これら、（３００個××％）の個数のストライプの書き込みでは、必ず、そのストレージ２０１への書き込みがされる。

[0176] このため、第１のストライプ数（３００個）のストライプのうちの、「｛（３００個×（１００－ｘ）％）／３｝　＋　（３００個××％）」の個数のストライプの書き込みにおいて、そのストレージ２０１への書き込みがされる。

[0177] こうして、アレイ２００に対して、単位データ量（２５２０ＭＢ参照）の書き込みがされる際に生じる、第１のパターンでの、ストレージ２０１への書き込みの回数Ｗ１（ｘ）が、上述された、「｛（３００個×（１００－ｘ）％）／３｝　＋　（３００個××％）」と特定される。

[0178] 同様にして、第１のパターンでの、ストレージ２０１への書き込みの回数Ｗ２（ｘ）、および、第３のパターンでの、ストレージ２０１への書き込みの回数Ｗ３（ｘ）も特定される。

[0179] これにより、２５２０ＭＢの書き込みがされる際における、ストレージ２０１への書き込みの回数Ｊ（ｘ）が、Ｊ（ｘ）＝Ｗ１（ｘ）＋Ｗ２（ｘ）＋Ｗ３（ｘ）と特定される。

[0180] 図３３は、パリティ割合２０１ｒ（上述のｘ）と、回数Ｊ（ｘ）だけの書き込みで書き込まれるデータ量との関係を特定する関数Ｋ（ｘ）のグラフを示す図である。先述のＪ（Ｘ）が特定されることにより、このＫ（ｘ）が特定される。

[0181] 図３４は、同時書き換えデータブロック数を考慮したブロック書き換え量を示す図である。

[0182] すなわち、上述された、第１のパターンでの回数であるＷ１（ｘ）は、図３４の上欄の左、中央、右の３つのグラフのうちの左のグラフでの関数である。そして、Ｗ２（ｘ）は、中央のグラフでの関数であり、Ｗ３（ｘ）は、右のグラフでの関数である。

[0183] そして、下欄のグラフの関数は、上述された、Ｊ（ｘ）＝Ｗ１（ｘ）＋Ｗ２（ｘ）＋Ｗ３（ｘ）から特定されるＫ（Ｘ）である。Ｓ２２０４～Ｓ２２０６（図２７）では、上述のＷ１（ｘ）、Ｗ２（ｘ）、Ｗ３（ｘ）が特定される。Ｓ２２０７では、特定された、これらＷ１（ｘ）、Ｗ２（ｘ）、Ｗ３

（ｘ）から、上述のＫ（ｘ）、つまり、図３４の下欄のグラフの関数が特定される。

[0184] 図３５Ａは、複数のストレージ２０１（ストレージ０～３）でのパリティ割合２０１ｒの組合わせを示す図である。

[0185] 図３５Ｂは、図３５Ａの組合わせの状態でのアレイ２００を示す図である。

[0186] 図３５Ｃは、パリティ割合２０１ｒの組合わせが、図３５Ａの組合わせである際における、それぞれのストレージ２０１での、書き込みがされるデータ量を示す図である。

[0187] 図３５Ｂの組合わせでは、図３５Ｃで示されるように、ストレージ０～４での、書き込みのデータ量（上述のＫ（ｘ））の比が、「１：１：１：３」である。

[0188] 図３６は、図３５Ａでの組合わせとは異なる組合わせを示す図である。

[0189] アレイ管理装置１０により特定される適切な組合わせは、通常は、図３５Ａで示されるような組合わせでなく、図３６で示されるような組合わせである。

[0190] 図３７は、３つのグラフを示す図である。

[0191] 上段のグラフは、パリティ割合２０１ｒと、書き込みのデータ量との間の関係を示すグラフであり、上述のＫ（ｘ）のグラフである（図３４の下欄を参照）。中段のグラフは、パリティ割合２０１ｒと、アレイ２００に対する書き込みのデータ量との間の関係を示すグラフ（Ｌ（ｘ））である。上段のグラフ（Ｊ（ｘ））から、この、中段のグラフ（Ｌ（ｘ））が特定される。

[0192] 下段のグラフは、パリティ割合２０１ｒと、ストレージ２０１の故障予測日時との関係を示すグラフ（Ｍ（ｘ））である。中段のグラフ（Ｌ（ｘ））から、この、下段のグラフ（Ｍ（ｘ））が特定される。

[0193] Ｓ２００４（図２５）では、特定された上述のＪ（ｘ）から特定される、このＭ（Ｘ）に基づいて故障予測日時が特定される。つまり、例えば、それぞれのストレージ２０１での、そのストレージ２０１が、それぞれのパリテ

ィ割合２０１ｒのときにおける故障予測日時が特定される。

[0194] 図３８は、４つのグラフを示す図である。すなわち、Ｓ２００４では、図３８の４つのグラフが特定される。

[0195] 図３９は、各パリティ割合における各ストレージの故障予測日時算出の処理のフローチャートである。図２５のＳ２００４では、例えば、この図３９の処理がされてもよい。

[0196] 図２６の処理では、それぞれのストレージ２０１を対象ストレージとして、Ｓ３３０２の処理が行われる。

[0197] Ｓ３３０２（図３９）では、それぞれのストレージ２０１（対象ストレージ）についての、それぞれのパリティ割合２０１ｒでの故障予測日時が特定される。

[0198] Ｓ２００５（図２５）では、アレイ２００に含まれる複数のストレージ２０１のパリティ割合２０１ｒの組合わせのそれぞれについて、その組合わせでの、それぞれのストレージ２０１の故障予測日時が特定される。つまり、Ｓ２００５では、それぞれの組合わせについて、その組合わせにおける、複数の故障予測日時が特定される。

[0199] そして、Ｓ２００５では、複数の組合わせ（上述）のうちから、特定される、複数の故障予測日時が、最も適切である組合わせを特定する（図３８の処理１Ｇ、先述の図１０Ｂなどを参照）。つまり、複数の組合わせから、最も適切な組合わせが特定される。なお、最も適切である組合わせでの、特定される複数の故障予測日時においては、例えば、先述された第１のストレージαの故障予測日時（図１０Ｂの日時１ｔ）と、第２のストレージβの故障予測日時１Ｂとの間の差１Ｆ１が、最大でもよい。

[0200] そして、図２４のＳ１９０３～Ｓ１９０４では、特定された、最適である、パリティ割合の組合わせに基づいて、先述された処理が行われる。つまり、例えば、このＳ１９０３～Ｓ１９０４では、アレイ２００の動作が制御されてもよい。つまり、制御がされた動作では、アレイ２００における、それぞれのストレージ２０１でのパリティ割合が、特定された、最適な組合わせ

での、そのストレージ２０１のパリティ割合と同じ割合でもよい（先述された、Ｓ１９０３、Ｓ９０４の説明などを参照）。

[0201] （数式等による説明）

以下では、より詳細な説明がされる。

[0202] まず、記号の説明がされる。以下の説明で、記号「N_{stripe}」は、アレイ２００におけるストライプ１ｘＢ１の数（ストライプ数）を示す。つまり、ストライプ番号ｉは、０～N_{stripe}－１の値と取りうる。また、記号「$N_{storage}$」は、アレイ２００におけるストレージ２０１の数（ストレージ数）を示す。ただし、スペアストレージは含まない。つまり、ストレージ番号ｊは、０～$N_{storage}$－１の値を取りうる。また、記号「N_{data}」は、各ストライプにおける、データブロック（ブロック１ｘＢ２ｄ）のストレージ数（図４等では３個）を示す。また、記号「N_{parity}」は、各ストライプにおける、パリティブロック（ブロック１ｘＢ２ｐ）のストレージ数を示す。つまり、ＲＡＩＤ５では、N_{parity}＝１であり、また、ＲＡＩＤ６ではN_{parity}＝２である。ここで、N_{stripe}、$N_{storage}$、N_{data}、N_{parity}は、一般に、アレイの新規構築時に決定し、運用中は変化しない。例として、$N_{storage}$＝４、N_{data}＝３、N_{parity}＝１とする。

[0203] 図３などにおいて、それぞれのブロック１ｘＢ２は、Ｂｋ○△という形で表現される。ここで、○はストライプ番号であり、△はストレージ番号である。

[0204] 図４、図７、図５などにおいて、データに対応するブロック（データのブロック１ｘＢ２ｄ）は、ＢＤ○＃という形で表現される。なお、ここで、「＃」は、「各ストライプにおけるデータブロックに対応するストレージ番号」である。一方、パリティに対応するブロック（ブロック１ｘＢ２ｐ）は、ＢＰ○＄という形で表現される。ここで、「＄」は、「各ストライプにおけるパリティブロックに対応するストレージ番号」である。

[0205] 次に、図１８の２つのグラフのうちの、左のグラフについて述べる。

[0206] まず、記号の説明がされる。記号「T_{now}」は、現在時刻を示す。記号「D_{now}」は、アレイに対する、現在の累積データ書き込み量を示す。記号「T_j」は

、アレイにおいて、ストレージｊが破損すると予測される日時を示す。記号「D_j」は、日時T_jにおける、アレイに対する累積データ書き込み量を示す。記号「t_j」は、現在時刻からストレージｊの故障予測日時までの所要時間（$T_j = T_{now} + t_j$）を示す。記号「d_j」は、日時T_jにおける、アレイに対する累積データ書き込み量（$D_j = D_{now} + d_j$）を示す。

[0207] ここで、T_{now}は取得可能である。また、D_{now}は、予め、過去のデータ書き込み量を記録しておく仕組みがアレイ管理装置１０に設けられ、それを利用し、取得可能とする。そして、以下の説明では、t_j、d_jを求めることが述べられる。これらt_j、d_jより、T_j、D_jを決定することが可能である。

[0208] 故障予測関数ｆが、線形とすると、以下の式での表現がされる。ここで、ａは係数である。

[0209] [数1]

$$a \cdot t_j = d_j \quad \cdots (1)$$

[0210] 係数ａについては、「過去のデータ書き込み量」と、「その時点の日時」の記録を用いた回帰分析等により決定される。

[0211] 次に、図１８の右のグラフについて説明される。

[0212] まず、記号の説明がされる。記号「W_{jnow}」は、ストレージｊに対する現在の累積ブロック書き換え量を示す。記号「W_{jmax}」は、ストレージｊに対するブロック書き換え量の保証値を示す。記号「w_j」は、ストレージｊに対する許容残書き換え量（$w_j = W_{jmax} - W_{jnow}$）を示す。ここで、$W_{jnow}$は、予め、過去のブロック書き換え量を記録しておく仕組みをアレイ管理装置１０に設けておき、それを利用し、取得可能とする。W_{jmax}は、メーカ保証値、あるいは、加速度試験などにより、予め決定するとする（S.M.A.R.T.（Self-Monitoring, Analysis and Reporting Technology）を用いてもよい）。W_{jnow}、W_{jmax}が決定可能なので、w_jも決定可能である。

[0213] 故障予測関数g_jが線形とすると、以下の式で表現される（ここで、b_jは係

数）。

[0214] [数2]

$$b_j \cdot d_j = w_j \quad \cdots (2)$$

[0215] 係数b_jは、ストレージごとに異なる。「過去のブロック書き換え量」と、「その時点のアレイに対するデータ書き込み量」の記録を用いた回帰分析等により決定される。

[0216] W_{jnow}、W_{jmax}、ａ、b_jが決定されることで、上述の式（１）（2）を用いて、t_jを算出できる。

[0217] 例として、以下の場合を考える（データ量の単位をＭＢ、時間の単位をｈとする）。アレイに対する書き込みデータ量は、平均して１時間あたり、４００ＭＢ（ａ＝４００）である。アレイに対して、１０ＭＢのデータ書き込みが発生する際に、ストレージｊに対して、４ＭＢのブロック書き換えが発生するとする（b_j＝０．４）。ストレージｊにおける許容残書き換え量は４００００ＭＢ（w_j＝４００００）である。式（１）、（２）より、d_j＝１０００００、t_j＝２５０であり、つまりストレージｊは、２５０時間後に破損するという予測が算出できる。全ストレージに対して、同様に、t_jを算出する。なお、故障予測関数ｆ、g_jは非線形でも良い。

[0218] 次に、図１０Ａ～図１０Ｃについて述べられる。

[0219] まず、記号の説明がされる。記号「Ｍ」は、冗長化アレイにおける最大許容故障台数を示す。つまり、ＲＡＩＤ５ではＭ＝１であり、ＲＡＩＤ６ではＭ＝２である。ＲＡＩＤ５、６の場合には、Ｍは、N_{parity}と一致する。記号「T_α」は、アレイにおいて、最初にストレージが破損すると予測される日時を示す。ここで、最初に破損すると予測されているストレージをαとする。記号「T_β」は、アレイにおいて、Ｍ＋１番目にストレージが破損する（つまり、アレイが破損する）と予測される日時を示す。ここで、Ｍ＋１番目に破損すると予測されているストレージをβとする。記号「$\Delta_{\alpha\beta}$」は、ストレージ

αと、ストレージβの破損の間隔（$\Delta_{\alpha\beta}=T_{\beta}-T_{\alpha}$）を示す。

[0220] ここで、Mは、一般に、アレイの新規構築時に決定し、運用中は変化しない。T_{α}、T_{β}は、全ストレージについてT_{j}を算出後、並び替えることで、決定可能である。そして、T_{α}、T_{β}が決定可能なので、$\Delta_{\alpha\beta}$も決定可能である。そして、$\Delta_{\alpha\beta}$が、所定値以内である場合、同時故障発生の危険性があると判断する。

[0221] 続けて、図１０Ｃについて述べられる。

[0222] まず、記号の説明がされる。記号「T_{γ}」は、アレイにおいて、ｘ番目にストレージが破損すると予測される日時を示す。ここで、ｘ番目に破損すると予測されているストレージをγとする。ｘは、2～Mの値を取りうる。記号「$T_{\gamma 1}$」は、アレイにおいて、ｘ－１番目にストレージが破損すると予測される日時を示す。ここで、ｘ－１番目に破損すると予測されているストレージをγ１とする。記号「$T_{\gamma 2}$」は、アレイにおいて、ｘ＋１番目にストレージが破損すると予測される日時を示す。ここで、ｘ－２番目に破損すると予測されているストレージをγ２とする。記号「$\Delta_{\gamma\gamma 1}$」は、ストレージγと、ストレージγ１の破損の間隔の時間（$\Delta_{\gamma\gamma 1}=T_{\gamma}-T_{\gamma 1}$）を示す。記号「$\Delta_{\gamma\gamma 2}$」は、ストレージγと、ストレージγ２の破損の間隔の時間（$\Delta_{\gamma\gamma 2}=T_{\gamma 2}-T_{\gamma}$）を示す。

[0223] ここで、ｘの値は２～Mの整数値を取りうる。ｘが決定すれば、T_{γ}、$T_{\gamma 1}$、$T_{\gamma 2}$、$\Delta_{\gamma\gamma 1}$、$\Delta_{\gamma\gamma 2}$も決定可能である。任意のｘについて、それぞれ$\Delta_{\gamma\gamma 1}$、$\Delta_{\gamma\gamma 2}$を算出する。そして、$\Delta_{\gamma\gamma 1}$または$\Delta_{\gamma\gamma 2}$が、所定値以内である場合、同時故障発生の危険性があると判断する。

[0224] 次に、図２８、図３０について述べられる。

[0225] 記号の説明がされる。記号「ｖ」は、同時書き換えデータブロック数を示す。ｖは、１～N_{data}の値を取りうる。

[0226] 以下の説明では、まず、ｖが、任意の値を取る場合について説明する。その後に、ｖの発生頻度を用いて、全てのｖについて算出した結果を統合して考える。

[0227] 次に、図２９について述べられる。

[0228] まず、記号の説明がされる。記号「ＢＫ」は、ブロックサイズ（ここでは固定とするが、可変でもよい）を示す。記号「C_v」は、同時書き換えデータブロック数ｖにおいて、１回のストライプ書き込みで保存可能なデータ量を示す。

[0229] C_vは、以下の式により算出される。

[0230] [数3]

$$C_v = v \cdot BK \quad \cdots (3)$$

[0231] 次に、図３１について述べられる。

[0232] まず、記号の説明がされる。記号「R_v」は、同時書き換えデータブロック数ｖにおける、書き込みストライプ数を示す。記号「P_v」は、同時書き換えデータブロック数ｖにおける、同時書き換え発生頻度。（P_vは０～１の値を取りうる。また、全同時書き換えデータブロック数における合計値はＲＡＩＤ５なら１、ＲＡＩＤ６なら２となる）を示す。

[0233] ここで、媒介変数ｃを用いて、以下の式での表現ができる。

[0234] [数4]

$$R_v = c \cdot P_v \quad \cdots (4)$$

[0235] 受付データのデータ量をｄとすると、このｄは、以下の式で算出される。

[0236] [数5]

$$d = \sum_{v=1}^{N_{data}} (R_v \cdot C_v) \quad \cdots (5)$$

つまり、(３)、(４)より、下記の通りである。

[0237]

[数6]

$$d = c \cdot BK \cdot \sum_{v=1}^{N_{data}} (v \cdot P_v)$$

・・・（６）

例として、以下の場合を考える（データ量の単位をＭＢとする）。つまり、N_{data}＝３、P_1＝０．２５、P_2＝０．４、P_3＝０．３５、ＢＫ＝１、ｄ＝２５２０の場合が考えられる。

[0238] 式（６）を解くと、ｃ＝１２００となる。つまり、R_1＝３００、R_2＝４８０、R_3＝４２０と算出することができる。

[0239] 次に、図３２について述べられる。まず、記号の説明がされる。記号「Q_v」は、同時書き換えデータブロック数ｖにおける、データブロックにおける書き換え処理発生率を示す。

[0240] Q_vは、以下の式により算出可能である。

[0241] [数7]

$$Q_v = \frac{v}{N_{data}}$$

・・・（７）

例として、N_{data}＝３、ｖ＝１なら、Q_1＝１／３である。

[0242] 次に、図３５Ａ～図３５Ｃについて述べられる。

[0243] 同時書き換えデータブロック数ｖにおいて、R_v個のストライプに対して書き込みを行うことを考える。ここで、各ストレージに対してブロック書き換えが発生する。

[0244] まず、各ストレージにおけるパリティ割合が０％または１００％である場合を考える。

[0245] 最初に、記号の説明がされる。記号「w_{vdata}」は、同時書き換えデータブロック数ｖにおいて、R_v個のストライプに対して書き込みを行う際に、「パリティ割合０％」のストレージに対して発生するブロック書き換え量を示す。記号「$w_{vparity}$」は、同時書き換えデータブロック数ｖにおいて、R_v個のスト

ライプに対して書き込みを行う際に、「パリティ割合１００％」のストレージに対して発生するブロック書き換え量を示す。

[0246] w_{vdata}、$w_{vparity}$は、以下の式により算出される。

[0247] [数8]

$$w_{vdata} = Q_v \cdot R_v \cdot BK \quad \cdots (8)$$

[0248] [数9]

$$w_{vparity} = R_v \cdot BK \quad \cdots (9)$$

[0249] 例として、以下の場合を考える（データ量の単位をＭＢ）。つまり、N_{data}＝３、ｖ＝１、ＢＫ＝１、R_v＝３００である場合が考えられる。

[0250] 式(８)、(９)より、w_{vdata}＝１００、$w_{vparity}$＝３００、つまりアレイに対して３００ＭＢのデータ書き込みを行う際に、パリティ割合０％のストレージに対するブロック書き換え量は１００ＭＢであるといえる。

[0251] 次に、図６、３３、３６について述べられる。

[0252] パリティ割合が０％および１００％以外の場合について考える。

[0253] まず、記号の説明がされる。記号「w_{vr}」は、同時書き換えデータブロック数ｖにおいて、R_v個のストライプに対して書き込みを行う際に、「パリティ割合ｒ」のストレージに対して発生するブロック書き換え量。以下の式により算出される（ただしｒは０～１の範囲を示す）。

[0254] w_{vr}は、以下の式により算出される。

[0255] [数10]

$$w_{vr} = (1 - r) \cdot w_{vdata} + r \cdot w_{vparity} \quad \cdots (10)$$

[0256] 例として、w_{vdata}＝１００、$w_{vparity}$＝３００、Ｒ＝０．３とすると、w_{vr}＝

１６０と算出できる。

[0257] 次に、図３４について述べられる。ここでは、前記で各ｖについて算出した結果を統合して考える。まず、記号の説明がされる。記号「w_r」は、「パリティ割合ｒ」のストレージに対して発生するブロック書き換え量を示す。

[0258] ここで、w_rは、各同時書き換えデータブロック数における合計値となる。つまり、w_rは、以下の式により算出される。

[0259] [数11]

$$w_r = \sum_{v=1}^{N_{data}} (w_{vr})$$

・・・（１１）

（１０）より、

[0260] [数12]

$$w_r = \sum_{v=1}^{N_{data}} ((1-r)\cdot w_{vdata} + r\cdot w_{vparity}) = \sum_{v=1}^{N_{data}} (w_{vdata} + (w_{vparity} - w_{vdata})\cdot r)$$

・・・（１２）

である。

[0261] （８）、（９）より、

[0262] [数13]

$$w_r = BK\cdot \sum_{v=1}^{N_{data}} (Q_v\cdot R_v + (1-Q_v)\cdot R_v\cdot r)$$

・・・（１３）

である。

[0263] （７）より、

[0264] [数14]

$$w_r = BK\cdot \sum_{v=1}^{N_{data}} (\frac{v}{N_{data}}\cdot R_v + (1-\frac{v}{N_{data}})\cdot R_v\cdot r)$$

・・・（１４）

と表せる。

[0265] 例として、以下の場合を考える（データ量の単位をＭＢ）。つまり、N_{data}＝３、R_1＝３００、R_2＝４８０、R_3＝４２０、ＢＫ＝１である場合が考えられる。

[0266] 式(１４)への代入がされると、以下の式となる。

[0267] [数15]

$$w_r = 360r + 840 \quad \cdots (15)$$

[0268] 上記の式に、例えばｒ＝０．３を代入すると、「アレイに対して２５２０ＭＢの書き込みを行う際の、パリティ割合３０％のストレージに対するブロック書き換え量」が、w_r＝９４８として算出できる。

[0269] さらに、(１４)に対して(４)を適用すると

[0270] [数16]

$$w_r = c \cdot BK \cdot \sum_{v=1}^{N_{data}} \left(\frac{v}{N_{data}} \cdot P_v + \left(1 - \frac{v}{N_{data}}\right) \cdot P_v \cdot r\right) \quad \cdots (16)$$

である。(６)より、

[0271] [数17]

$$c \cdot BK = \frac{d}{\sum_{v=1}^{N_{data}} (v \cdot P_v)} \quad \cdots (17)$$

なので、

[0272] [数18]

$$w_r = \frac{d}{\sum_{v=1}^{N_{data}} (v \cdot P_v)} \cdot \sum_{v=1}^{N_{data}} \left(\frac{v}{N_{data}} \cdot P_v + \left(1 - \frac{v}{N_{data}}\right) \cdot P_v \cdot r\right)$$

・・・（１８）

である。整理すると

[0273] [数19]

$$w_r = d \cdot (\frac{1}{N_{data}} + (\frac{\sum_{v=1}^{N_{data}} (P_v)}{\sum_{v=1}^{N_{data}} (v \cdot P_v)} - \frac{1}{N_{data}}) \cdot r)$$

・・・（１９）

である。ここで、

[0274] [数20]

$$\sum_{v=1}^{N_{data}} P_v = 1$$

・・・（２０）

なので、

[0275] [数21]

$$w_r = d \cdot (\frac{1}{N_{data}} + (\frac{1}{\sum_{v=1}^{N_{data}} (v \cdot P_v)} - \frac{1}{N_{data}}) \cdot r)$$

・・・（２１）

である。

[0276] つまり、式(２１)より、ｒ、ｄ、N_{data}、およびそれぞれの同時書き換えデータブロック数ｖにおける、同時書き換え発生頻度P_vが決定すれば、w_rを決定することができる。

[0277] ここで、

[数22]

$$\sum_{v=1}^{N_{data}} (v \cdot P_v)$$

・・・（２２）

をＹとおく。Ｙの値は、同時書き換え発生頻度P_vとN_{data}により決定する。N_{data}は、アレイ新規構成時に決定し、P_vは、ストレージに対して書き込みを行うアプリケーションやシステムなど外部の特性により決定する（シーケンシャルライト、ランダムライトの割合、その時のデータ量など）。つまり、Ｙは、アレイを構成する各ストレージのパリティ割合の影響を受けない。

[0278] 次に、図３７について述べられる。

[0279] ストレージｊのパリティ割合をｒに再設定することを考える。まず、記号の説明がされる。記号「T_{jr}」は、ストレージｊのパリティ割合をｒに再設定した時に、アレイにおいて、ストレージｊが破損すると予測される日時を示す。記号「D_{jr}」は、日時T_{jr}における、アレイに対する累積データ書き込み量を示す。記号「t_{jr}」は、現在時刻から、ストレージｊのパリティ割合をｒに再設定した時のストレージｊの故障予測日時までの所要時間（$T_{jr}=T_{now}+t_{jr}$）を示す。記号「d_{jr}」は、日時T_{jr}における、アレイに対する累積データ書き込み量を示す（$D_{jr}=D_{now}+d_{jr}$）。

[0280] ストレージｊのパリティ割合をｒに変更することで、故障予測日時を変化させることができる。なお、故障予測関数ｆについては影響を受けない。ここで、式(20)において$w_r=w_j$、$d=d_{jr}$とすることで、d_{jr}を算出できる。

[0281] [数23]

$$w_j = d_{jr} \cdot \left(\frac{1}{N_{data}} + \left(\frac{1}{Y} - \frac{1}{N_{data}}\right) \cdot r\right)$$

・・・（２３）

変形すると

[0282] [数24]

$$d_{jr} = \frac{N_{data} \cdot Y}{Y + (N_{data} - Y) \cdot r} \cdot w_j$$

・・・（２４）

である。さらに、故障予測関数ｆを用いて故障予測日時を算出する。(１)より、

[0283] [数25]

$$t_{jr} = \frac{N_{data} \cdot Y}{Y + (N_{data} - Y) \cdot r \cdot a} \cdot w_j$$

・・・（２５）

である。

[0284] 例として、以下の場合を考える（データ量の単位をＭＢ）。つまり、N_{data}＝３、P_1＝０．２５、P_2＝０．４、P_3＝０．３５、w_j＝４００００である場合が考えられる。

[0285] 式(２４)への代入がされると、以下の式が成り立つ。

[0286] [数26]

$$d_{jr} = \frac{840000}{7 + 3 \cdot r}$$

・・・（２６）

さらに、ａ＝４００とすると、

[0287] [数27]

$$t_{jr} = \frac{2100}{7 + 3 \cdot r}$$

・・・（２７）

という式が成り立つ。

[0288] 上記の式に、例えばｒ＝０．３を代入すると、「ストレージｊのパリティ割合を３０％とした際の、故障予測日時（現在日時からの経過時間）がt_{jr}＝２６６、つまり２６６時間後として算出できる。そして、例えば、パリティ割合を０％に変更すればt_{jr}＝３００、つまり故障予測日時を３４時間変化させることができる。

[0289] 最後に、図３８について述べられる。図３８では、全てのストレージに対して、故障予測日時までの所要時間t_{jr}と、パリティ割合ｒの関係関数を導出することが示される。

[0290] （その他）

なお、例えば、次のようでもよい。

[0291] すなわち、アレイ２００により保持されるデータの消失は、第１のストレージαが故障する日時１ｔが第１の日時１ｔ１である場合には、発生し難く、第２の日時１ｔ２である場合には発生し易く、第１のストレージαは、第１のストレージαにより記憶されるデータ（領域２０１ｄ参照）とパリティ（領域２０１ｐ参照）との全体（領域２０１Ａ参照）に占める、パリティの割合であるパリティ割合２０１ｒが、第１の割合２０１ｒ１である場合と、第２の割合２０１ｒ２である場合とがあるストレージであり、第１のストレージαの故障の日時１ｔは、第１の割合２０１ｒ１である場合には、消失が発生し難い第１の日時１ｔ１であり、第２の割合２０１ｒ２である場合には、発生し易い第２の日時１ｔ２であり、判断部１００ｘは、予め定められた２つの割合のうちの一方の割合を第１の割合２０１ｒ１と特定し、他方の割合を第２の割合２０１ｒ２と特定し、変更制御部１００ｙは、第１のストレージαでの割合２０１ｒを、特定された一方の割合（第１の割合２０１ｒ１）にさせ、特定された他方の割合（第２の割合２０１ｒ２）にはさせない制御をしてもよい。

[0292] そして、消失が発生し難い第１の日時１ｔ１は、アレイ２００に含まれる、第１のストレージαとは別の第２のストレージβが故障する日時１Ｂから閾値１Ｆ２以上に離れた日時であり、消失が発生し易い第２の日時１ｔ２は、第２のストレージβが故障する日時１Ｂから閾値１Ｆ２以内の日時でもよい。

[0293] これにより、データの消失の発生が防げる。

[0294] しかも、単なる、パリティの割合に応じた処理がされるだけで済み、処理が簡単にできる。

[0295] なお、本実施例のアレイ管理装置は、典型的には半導体集積回路であるＬＳＩとして実現される。これらは個別に１チップ化されてもよいし、一部またはすべてを含むように１チップ化されても良い。ここではＬＳＩとしたが、集積度の違いにより、ＩＣ、システムＬＳＩ、スーパーＬＳＩ、ウルトラ

ＬＳＩと呼称されることもある。

[0296] また、集積回路化の手法はＬＳＩに限るものではなく、専用回路または汎用プロセッサで実現しても良い。ＬＳＩ製造後に、プログラムすることが可能なＦＰＧＡ（Ｆｉｅｌｄ Ｐｒｏｇｒａｍｍａｂｌｅ Ｇａｔｅ Ａｒｒａｙ）や、ＬＳＩ内部の回路セルの接続や設定を再構成可能なリコンフィギュラブル・プロセッサを利用しても良い。

[0297] さらには、半導体技術の進歩または派生する別技術によりＬＳＩに置き換わる集積回路化の技術が登場すれば、当然、その技術を用いて機能ブロックの集積化を行っても良い。バイオ技術の適応などが可能性として有り得る。

[0298] さらに加えて、本実施例の～装置を集積化した半導体チップと、画像を描画するためのディスプレイとを組み合せて、様々な用途に応じた描画機器を構成することができる。携帯電話やテレビ、デジタルビデオレコーダ、デジタルビデオカメラ、カーナビゲーション等における情報描画手段として、本発明を利用することが可能である。ディスプレイとしては、ブラウン管（ＣＲＴ）の他、液晶やＰＤＰ（プラズマディスプレイパネル）、有機ＥＬなどのフラットディスプレイ、プロジェクターを代表とする投射型ディスプレイなどと組み合わせることが可能である。

[0299] なお、上記各実施の形態において、各構成要素は、専用のハードウェアで構成されるか、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、ＣＰＵまたはプロセッサなどのプログラム実行部が、ハードディスクまたは半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。ここで、上記各実施の形態のアレイ管理装置などを実現するソフトウェアは、次のようなプログラムである。

[0300] すなわち、このプログラムは、複数のストレージが構成するアレイを管理するコンピュータに、前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断ステップと、前記アレイが保持する

前記データが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、故障時期変更ステップとを実行させるためのコンピュータプログラムなどである。

産業上の利用可能性

[0301] 本発明のアレイ管理装置は、様々な用途に利用可能である。例えば、携帯電話や携帯音楽プレーヤー、デジタルカメラ、デジタルビデオカメラ等の電池駆動の携帯表示端末や、テレビ、デジタルビデオレコーダ、カーナビゲーション等の高解像度の情報表示機器におけるメニュー表示やＷｅｂブラウザ、エディタ、ＥＰＧ、地図表示等における情報表示手段として利用価値が高い。

[0302] 将来、複数のストレージが同時期に故障することによって、アレイが保持するデータが消失する可能性があると判断した場合には、同時故障する可能性がある複数のストレージの故障時期をずらす。このことで、アレイが保持するデータ消失を防ぐことができる。

符号の説明

[0303] １ｘＢ２、１ｘＢ２ｄ、１ｘＢ２ｐ　ブロック

１０　アレイ管理装置

１００ｘ　判断部

１００ｙ　変更制御部

１０１　パリティ割合管理部

１０２　パリティ割合算出部

１０３　パリティ割合変更部

１０４　故障予測部

２００　アレイ

２０１　ストレージ

２０１ｄ　領域

２０１ｐ　領域

請求の範囲

[請求項1] 複数のストレージが構成するアレイを管理するアレイ管理装置であって、

前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断部と、

前記アレイが保持する前記データが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、故障時期変更部と、

を備えるアレイ管理装置。

[請求項2] 前記故障時期変更部は、前記各ストレージに対して書き込む、パリティとデータとの割合を変更することによって、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、

請求項１記載のアレイ管理装置。

[請求項3] 前記判断部は、

前記複数のストレージのうち少なくとも２台のストレージが同時期に故障する可能性があるか否かを判断する同時故障判断部と、

前記少なくとも２台のストレージが同時期に故障すると判断された場合に、前記少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断するアレイ故障判断部とを含む

請求項１記載のアレイ管理装置。

[請求項4] 前記アレイにより保持される前記データの消失は、第１の前記ストレージが故障する日時が第１の日時である場合には、発生し難く、第２の日時である場合には発生し易く、

第１の前記ストレージは、記憶されるデータとパリティとの全体に占める、前記パリティの割合が、第１の割合である場合と、第２の割合である場合とがあるストレージであり、

第１の前記ストレージの前記日時は、前記第１の割合である場合には、消失が発生し難い前記第１の日時であり、前記第２の割合である場合には、発生し易い前記第２の日時であり、

前記判断部は、２つの割合のうちの一方の前記割合を前記第１の割合と特定し、他方の前記割合を前記第２の割合と特定し、

前記故障時期変更部は、第１の前記ストレージでの前記割合を、特定された一方の前記割合にさせ、特定された他方の前記割合にはさせない、

請求項１記載のアレイ管理装置。

[請求項5] 消失が発生し難い前記第１の日時は、第２の前記ストレージが故障する日時から閾値以上に離れた日時であり、

消失が発生し易い前記第２の日時は、第２の前記ストレージが故障する日時から閾値以内の日時である、

請求項４記載のアレイ管理装置。

[請求項6] 複数のストレージが構成するアレイを管理する集積回路であって、

前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断部と、

前記アレイが保持する前記データが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、故障時期変更部と、

を備える集積回路。

[請求項7] 複数のストレージが構成するアレイを管理するアレイ管理方法であって、

前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断ステップと、

前記アレイが保持する前記データが消失する可能性があると判断さ

れた場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、故障時期変更ステップと、

を含むアレイ管理方法。

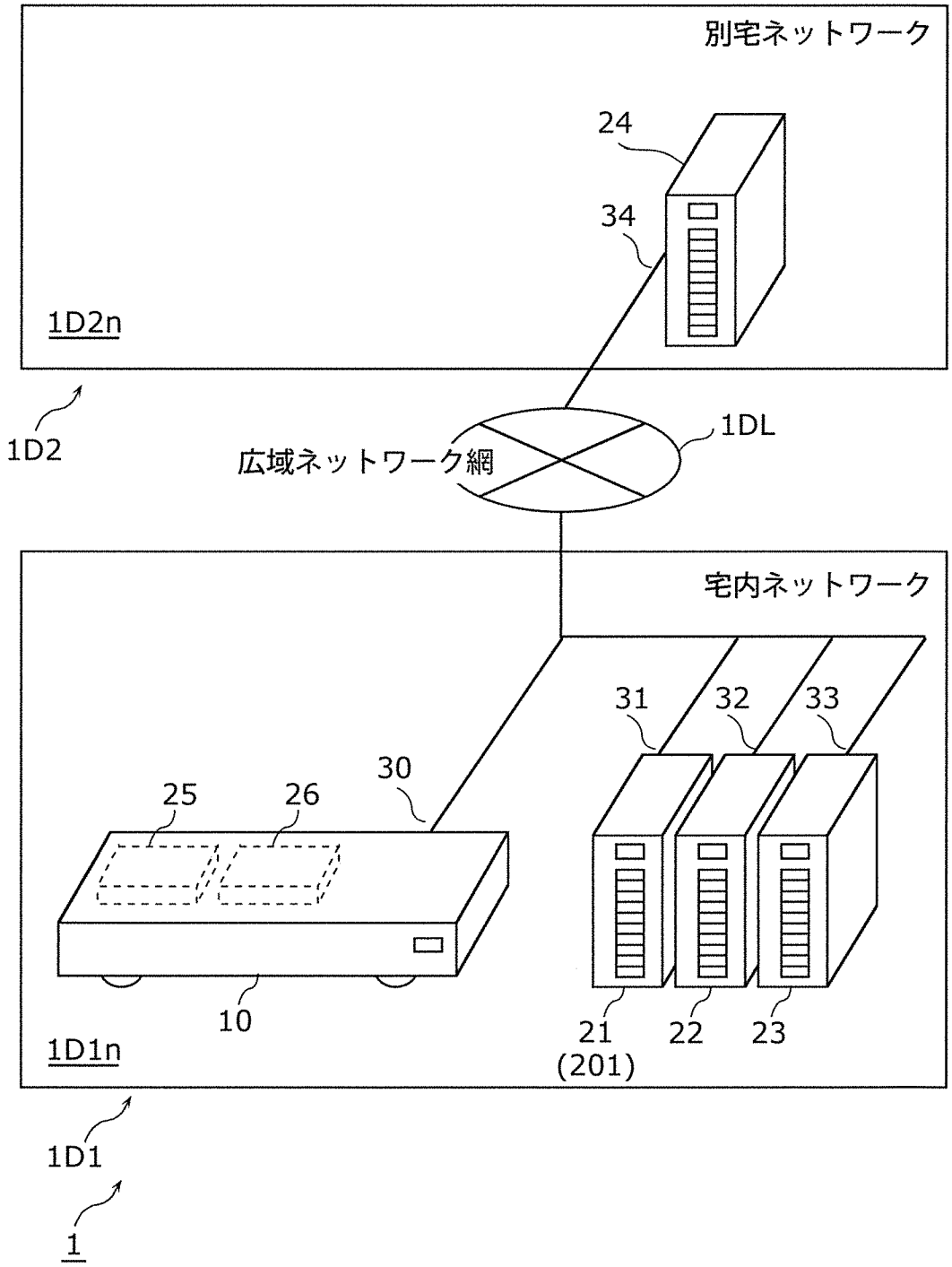
［請求項8］ 複数のストレージが構成するアレイを管理するコンピュータに、

前記複数のストレージのうち少なくとも２台のストレージが同時期に故障することによって、前記アレイが保持するデータが消失する可能性があるか否かを判断する判断ステップと、

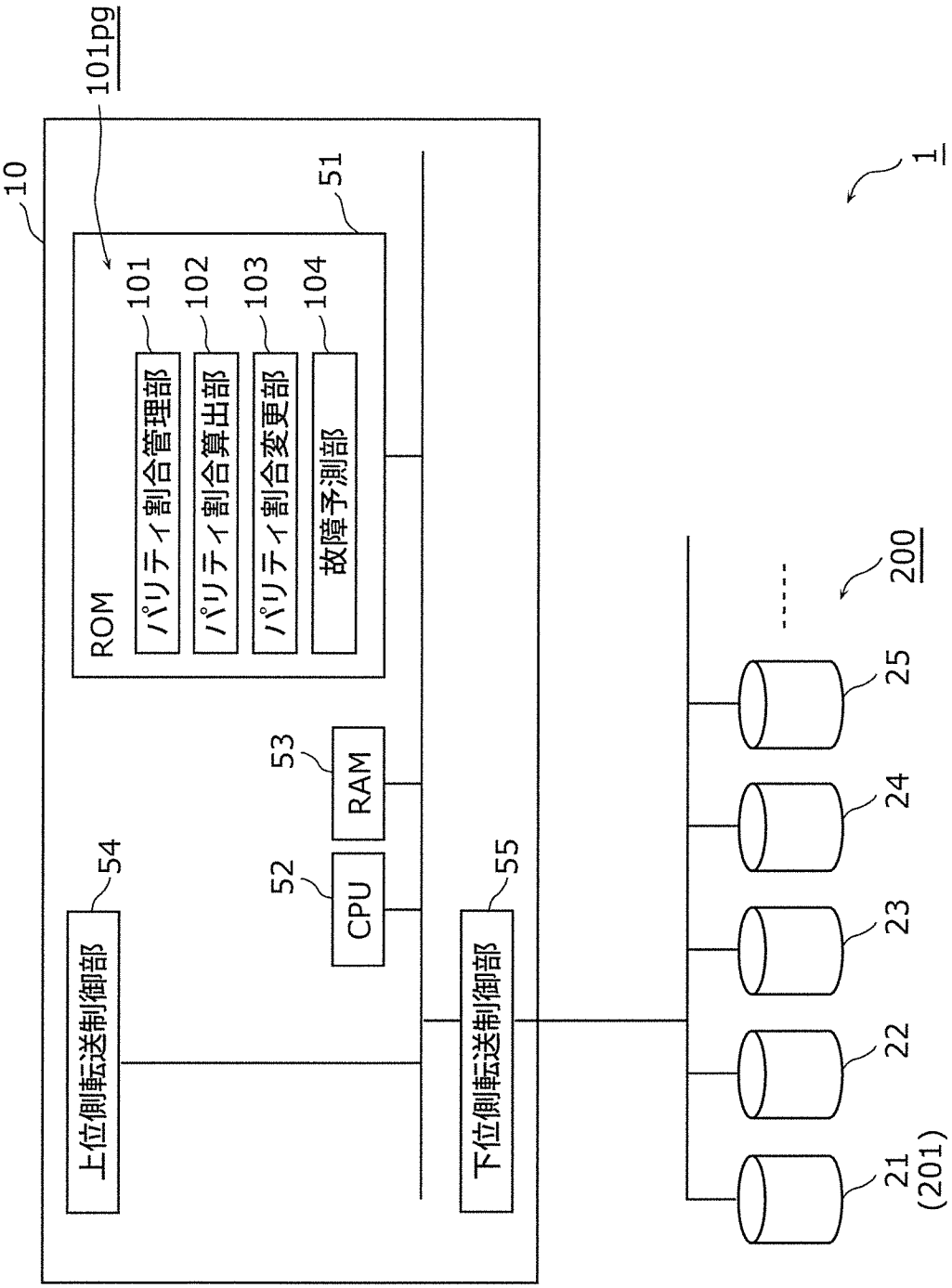
前記アレイが保持する前記データが消失する可能性があると判断された場合には、前記少なくとも２台のストレージのうちの少なくとも一つの故障時期を変更する、故障時期変更ステップと、

を実行させるためのコンピュータプログラム。

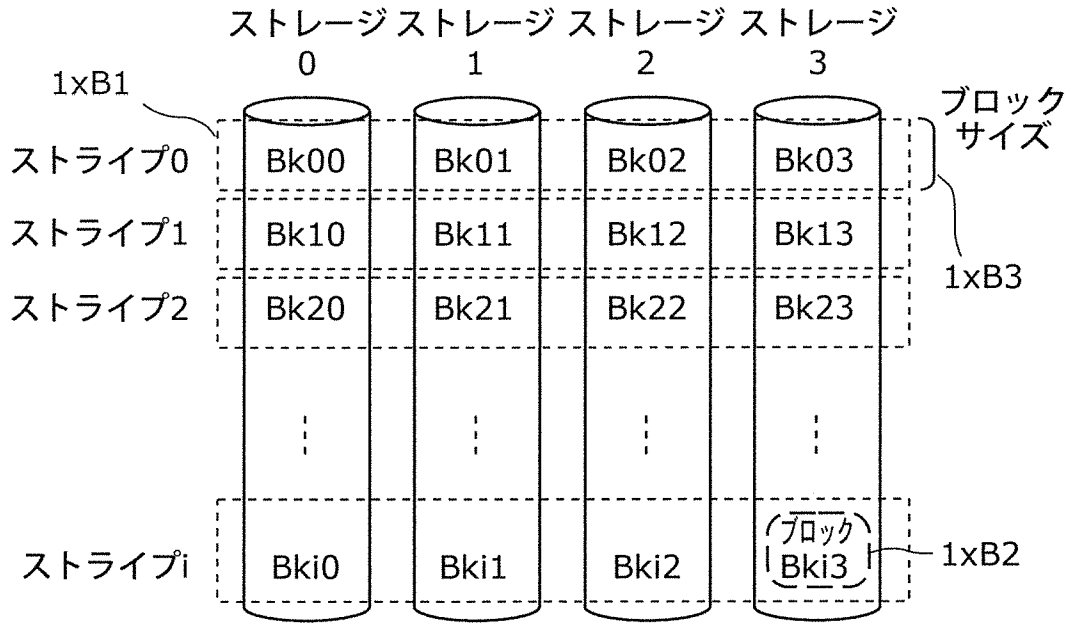
[図1]


別宅ネットワーク
24
34
1D2n
1D2
広域ネットワーク網
1DL
宅内ネットワーク
31
32
33
30
25
26
10
1D1n
21
(201)
22
23
1D1
1

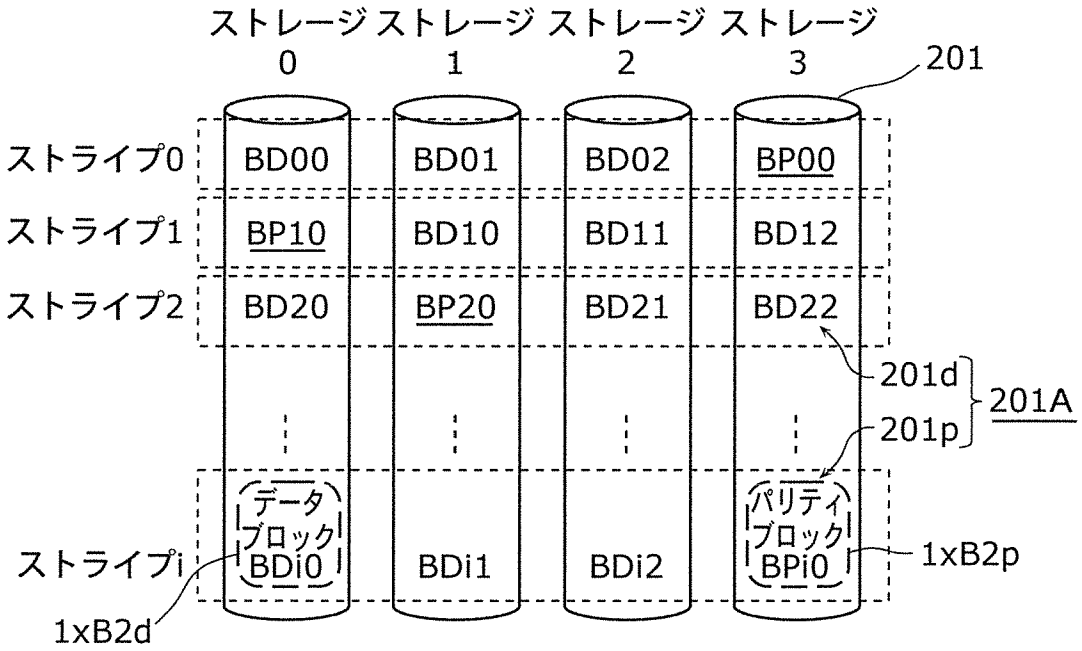
[図2]


101pg
10
51
ROM
101
パリティ割合管理部
102
パリティ割合算出部
103
パリティ割合変更部
104
故障予測部
53
RAM
52
CPU
54
上位側転送制御部
55
下位側転送制御部
1
200
25
24
23
22
21
(201)

[図3]


ストレージ ストレージ ストレージ ストレージ
0
1
2
3
1xB1
ブロック
サイズ
ストライプ0
Bk00
Bk01
Bk02
Bk03
ストライプ1
Bk10
Bk11
Bk12
Bk13
1xB3
ストライプ2
Bk20
Bk21
Bk22
Bk23
ストライプi
Bki0
Bki1
Bki2
ブロック
Bki3
1xB2

[図4]


ストレージ ストレージ ストレージ ストレージ
0
1
2
3
201
ストライプ0
BD00
BD01
BD02
BP00
ストライプ1
BP10
BD10
BD11
BD12
ストライプ2
BD20
BP20
BD21
BD22
201d
201A
201p
データ
ブロック
BDi0
BDi1
BDi2
パリティ
ブロック
BPi0
ストライプi
1xB2p
1xB2d

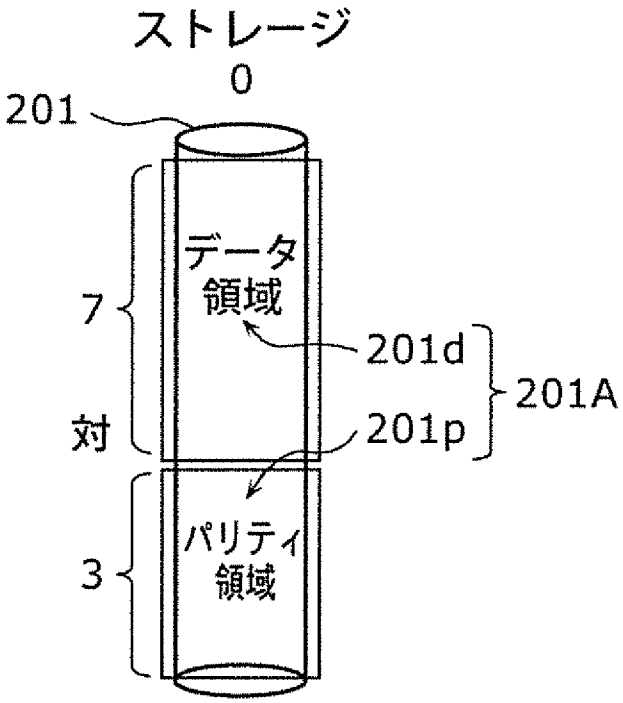
[図5]

	ストレージ0	ストレージ1	ストレージ2	ストレージ3
ストライプ0	BD00	BD01	BD02	BP00
ストライプ1	BP10	BD10	BD11	BD12
ストライプ2	BD20	BP20	BD21	BD22

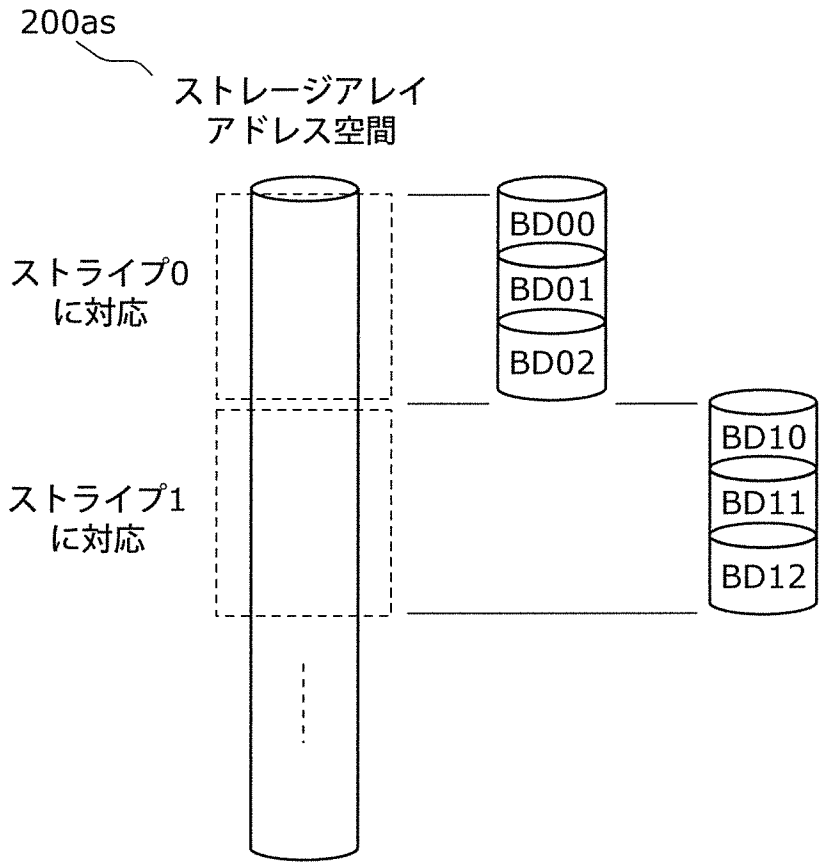
ストライプi	BDi0	BDi1	BDi2	BPi0
各ストレージにおけるパリティ割合	30%	20%	15%	35%

201r
(201r1,201r2)

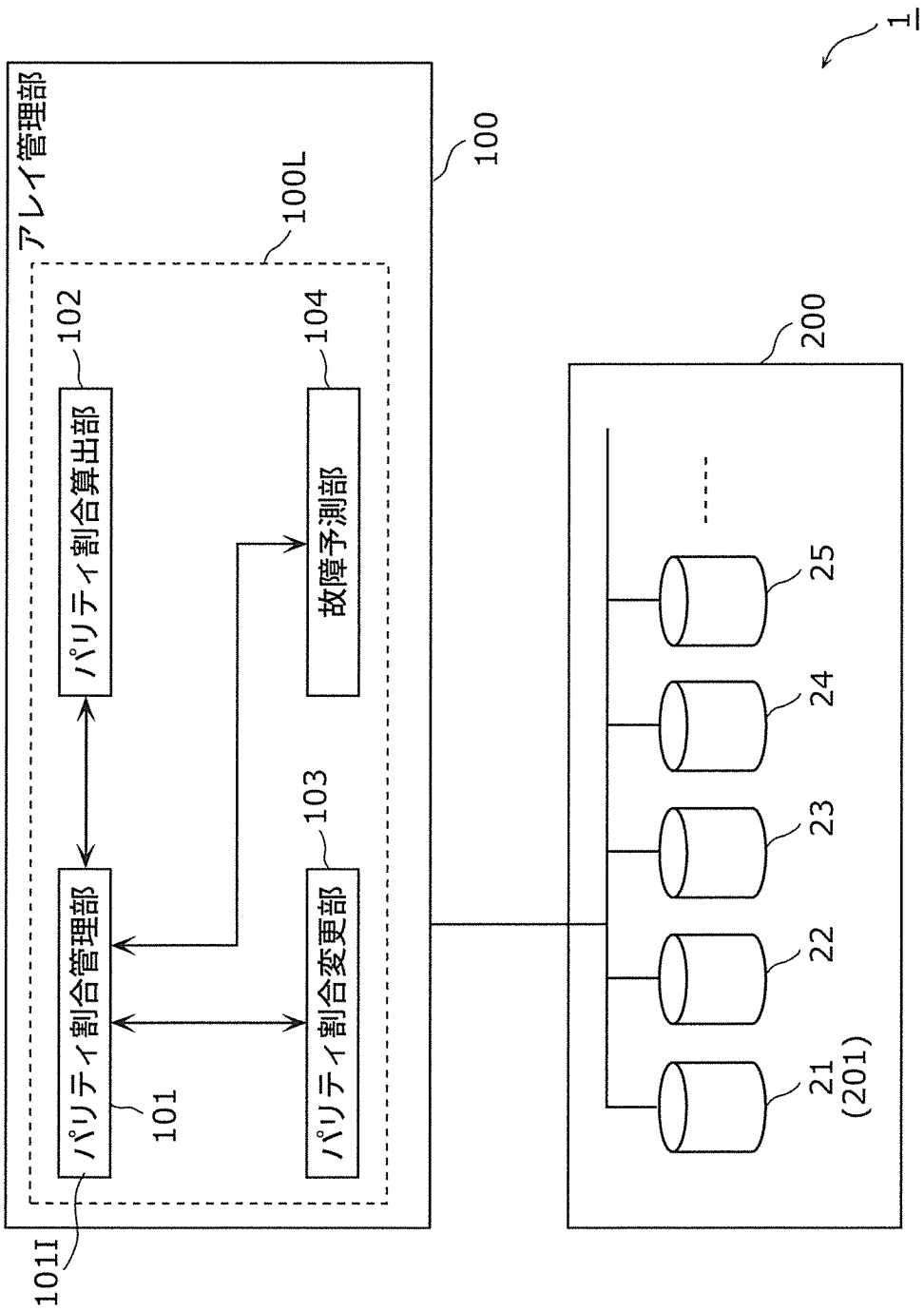
[図6]



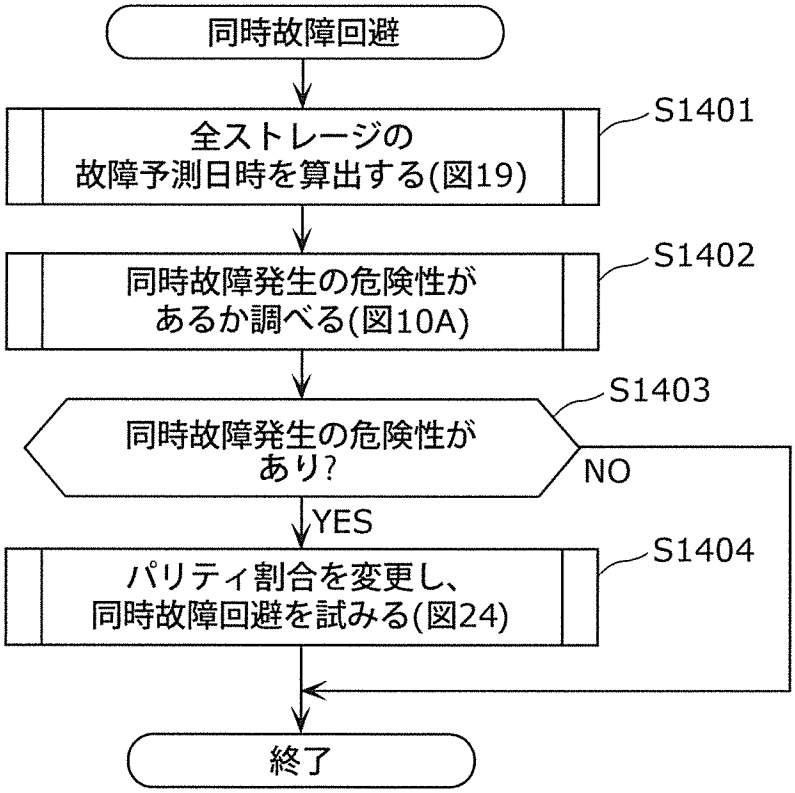
[図7]


200as
ストレージアレイ
アドレス空間
ストライプ0
に対応
ストライプ1
に対応
BD00
BD01
BD02
BD10
BD11
BD12

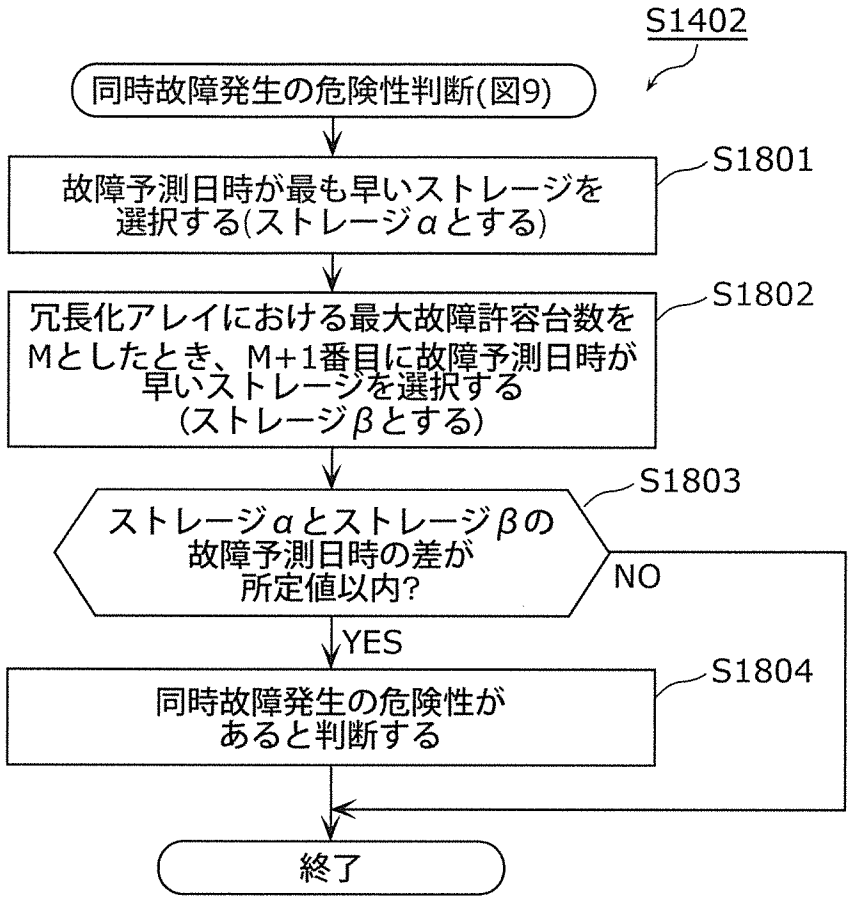
[図8]


1
アレイ管理部
100
100L
102
パリティ割合算出部
104
故障予測部
103
パリティ割合変更部
パリティ割合管理部
101
101I
200
25
24
23
22
21
(201)

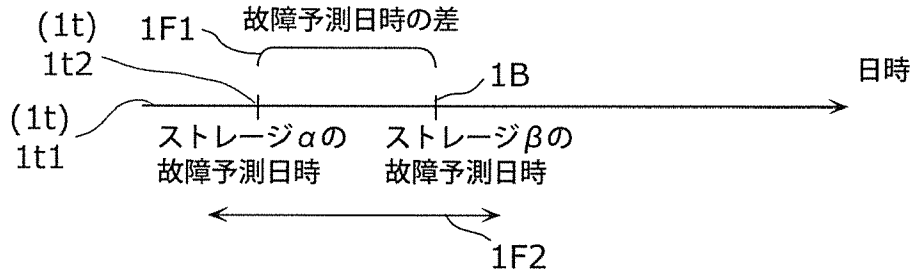
[図9]


同時故障回避
全ストレージの
故障予測日時を算出する(図19)
S1401
同時故障発生の危険性が
あるか調べる(図10A)
S1402
同時故障発生の危険性が
あり?
S1403
NO
YES
パリティ割合を変更し、
同時故障回避を試みる(図24)
S1404
終了

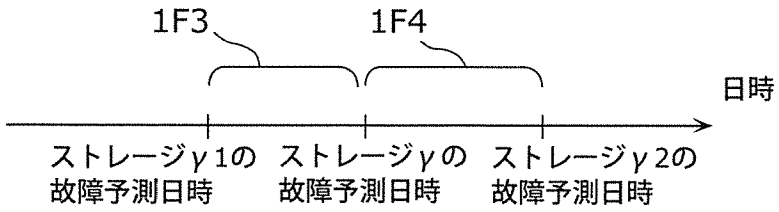
[図10A]


S1402
同時故障発生の危険性判断(図9)
故障予測日時が最も早いストレージを
選択する(ストレージαとする)
S1801
冗長化アレイにおける最大故障許容台数を
Mとしたとき、M+1番目に故障予測日時が
早いストレージを選択する
(ストレージβとする)
S1802
ストレージαとストレージβの
故障予測日時の差が
所定値以内?
S1803
NO
YES
同時故障発生の危険性が
あると判断する
S1804
終了

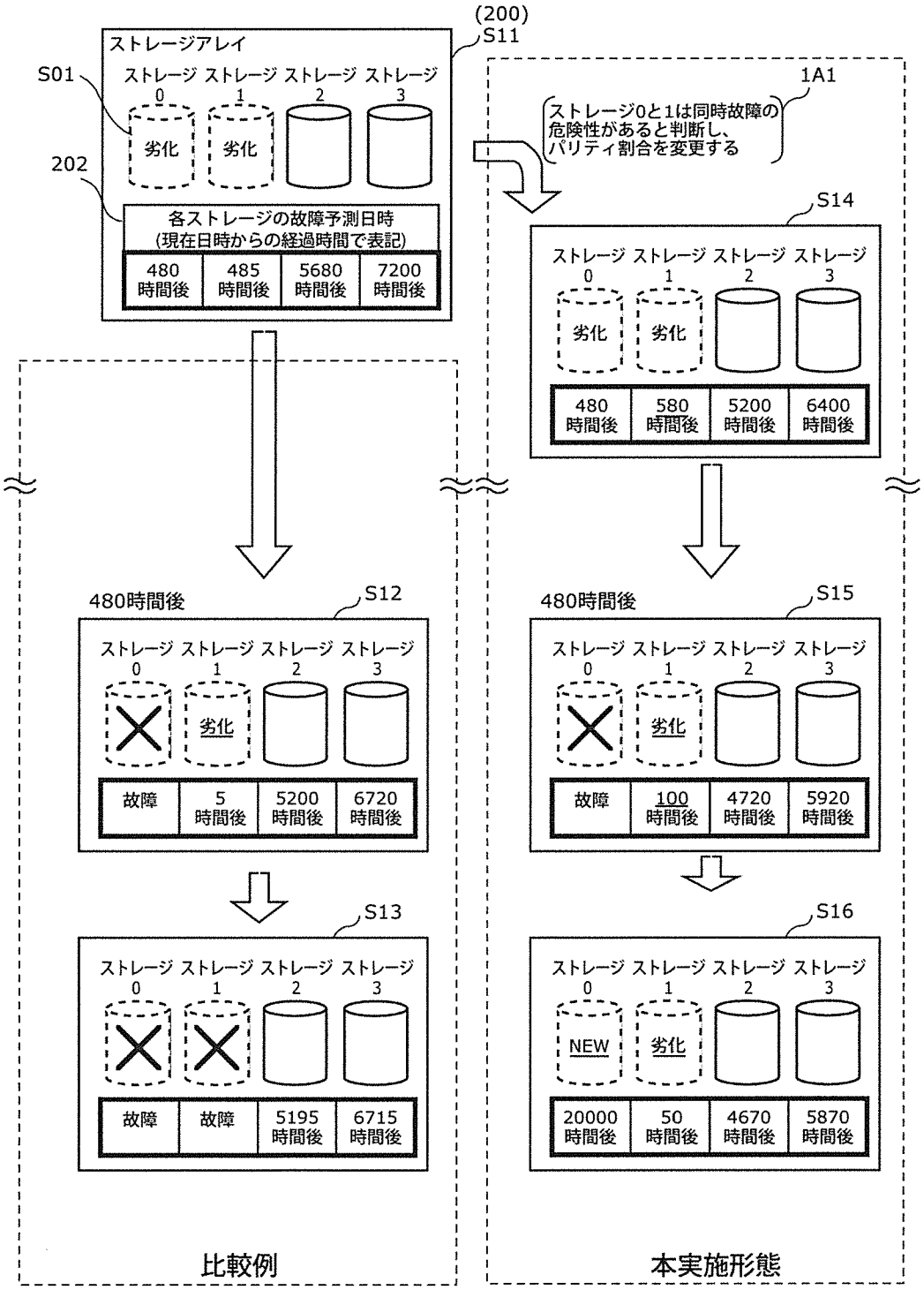
[図10B]


(1t)
1t2
1F1
故障予測日時の差
1B
日時
(1t)
1t1
ストレージαの
故障予測日時
ストレージβの
故障予測日時
1F2

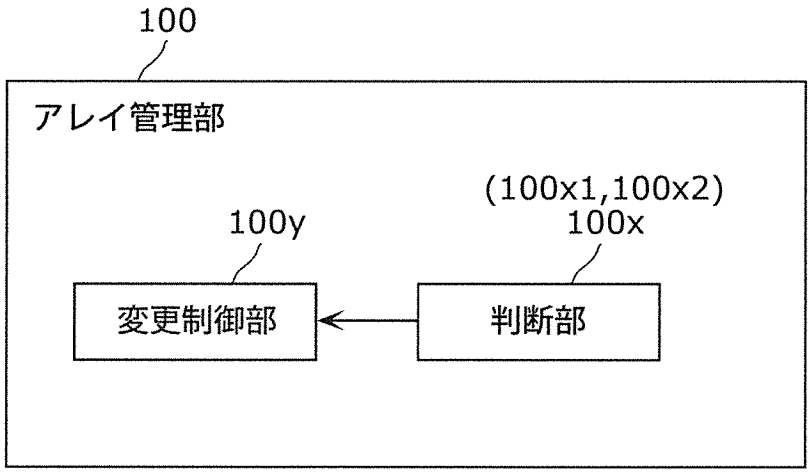
[図10C]


1F3
1F4
日時
ストレージγ1の
故障予測日時
ストレージγの
故障予測日時
ストレージγ2の
故障予測日時

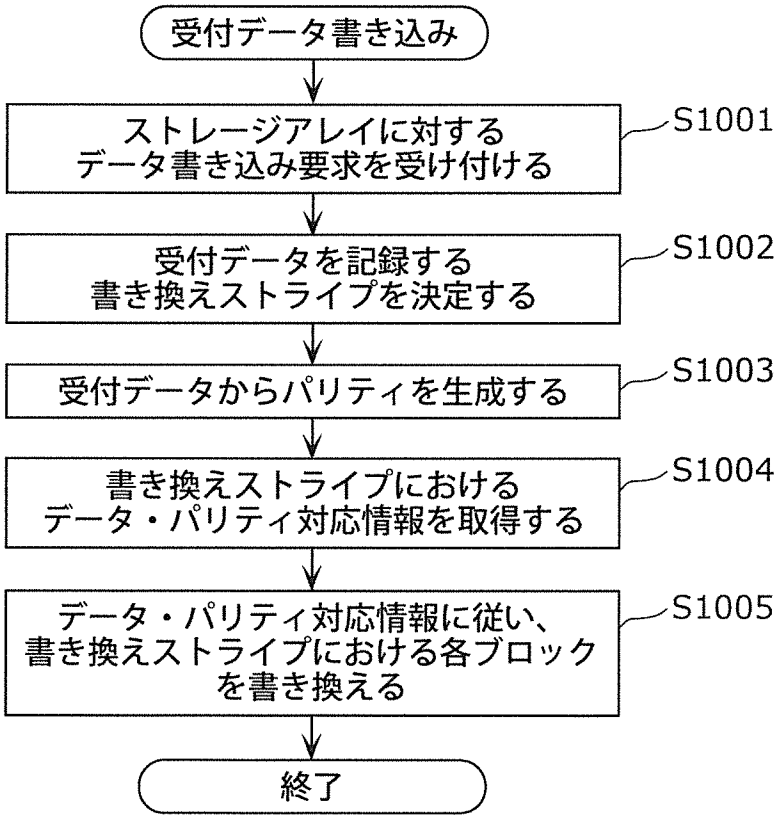
[図11]


(200)
S11
ストレージアレイ
S01
ストレージ ストレージ ストレージ ストレージ
0 1 2 3
劣化 劣化
202
各ストレージの故障予測日時
(現在日時からの経過時間で表記)
480 時間後
485 時間後
5680 時間後
7200 時間後
1A1
ストレージ0と1は同時故障の危険性があると判断し、パリティ割合を変更する
S14
480 580 5200 6400 時間後
480時間後
S12
故障 5 5200 6720 時間後
S13
故障 故障 5195 6715 時間後
S15
故障 100 4720 5920 時間後
S16
NEW
20000 50 4670 5870 時間後
比較例
本実施形態

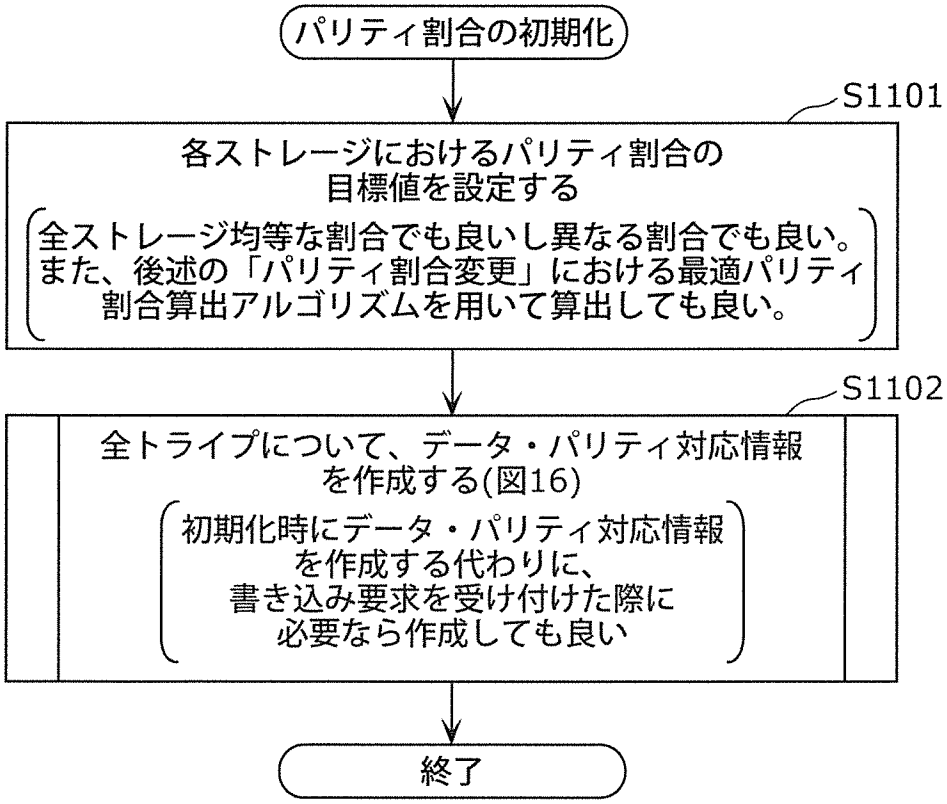
[図12]


100
アレイ管理部
(100x1,100x2)
100x
100y
変更制御部
判断部

[図13]


受付データ書き込み
ストレージアレイに対する
データ書き込み要求を受け付ける
S1001
受付データを記録する
書き換えストライプを決定する
S1002
受付データからパリティを生成する
S1003
書き換えストライプにおける
データ・パリティ対応情報を取得する
S1004
データ・パリティ対応情報に従い、
書き換えストライプにおける各ブロック
を書き換える
S1005
終了

[図14]


パリティ割合の初期化
S1101
各ストレージにおけるパリティ割合の
目標値を設定する
全ストレージ均等な割合でも良いし異なる割合でも良い。
また、後述の「パリティ割合変更」における最適パリティ
割合算出アルゴリズムを用いて算出しても良い。
S1102
全トライプについて、データ・パリティ対応情報
を作成する(図16)
初期化時にデータ・パリティ対応情報
を作成する代わりに、
書き込み要求を受け付けた際に
必要なら作成しても良い
終了

[図15]

	ストレージ0	ストレージ1	ストレージ2	ストレージ3
ストライプ0	—	—	—	—
ストライプ1	—	—	—	—
ストライプ2	—	—	—	—

各ストレージにおけるパリティ割合	0%	0%	0%	0%
目標値	30%	20%	15%	35%
目標値から現在値を減じた値	30%	20%	15%	(35%)

⇩ (ストライプ0において、ストレージ3に対応するブロックにパリティを割り当て
それ以外のブロックについてはデータを割り当てる) 1C1

	ストレージ0	ストレージ1	ストレージ2	ストレージ3
ストライプ0	BD00	BD01	BD02	BP00
ストライプ1	—	—	—	—
ストライプ2	—	—	—	—

各ストレージにおけるパリティ割合	0%	0%	0%	0%
目標値	30%	20%	15%	35%
目標値から現在値を減じた値	(30%)	20%	15%	-65%

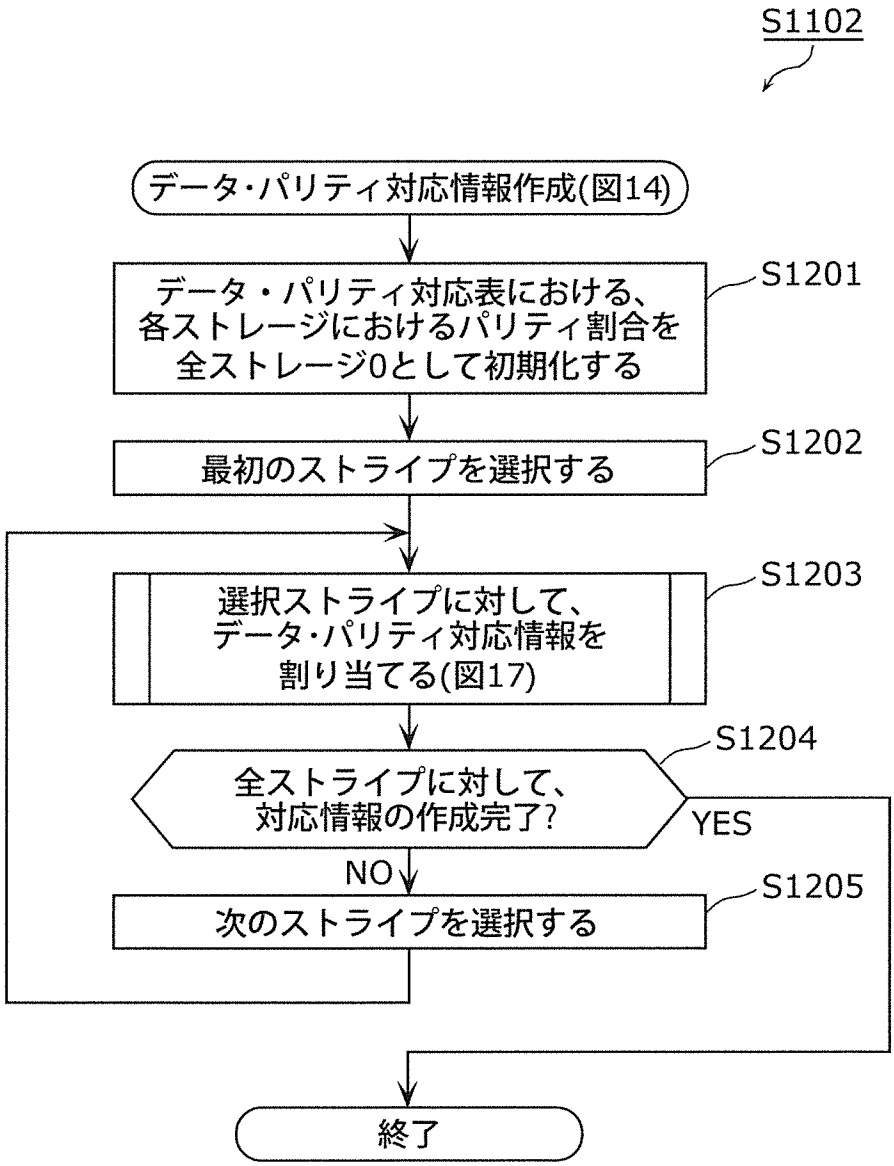
⇩ (ストライプ1において、ストレージ0に対応するブロックにパリティを割り当て
それ以外のブロックについてはデータを割り当てる) 1C2

	ストレージ0	ストレージ1	ストレージ2	ストレージ3
ストライプ0	BD00	BD01	BD02	BP00
ストライプ1	BP10	BD10	BD11	BD12
ストライプ2	—	—	—	—

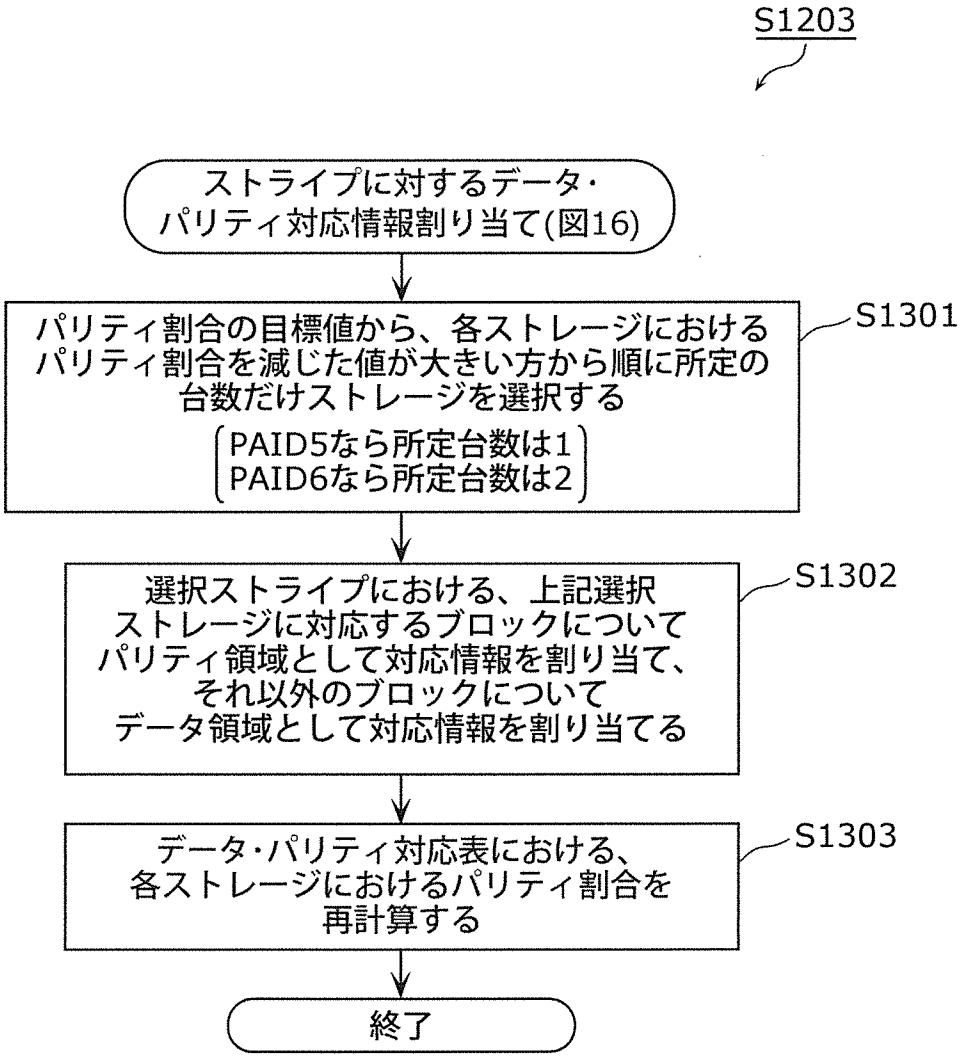
各ストレージにおけるパリティ割合	50%	0%	0%	0%
目標値	30%	20%	15%	35%
目標値から現在値を減じた値	-20%	(20%)	15%	-15%

以下、同様に繰り返す

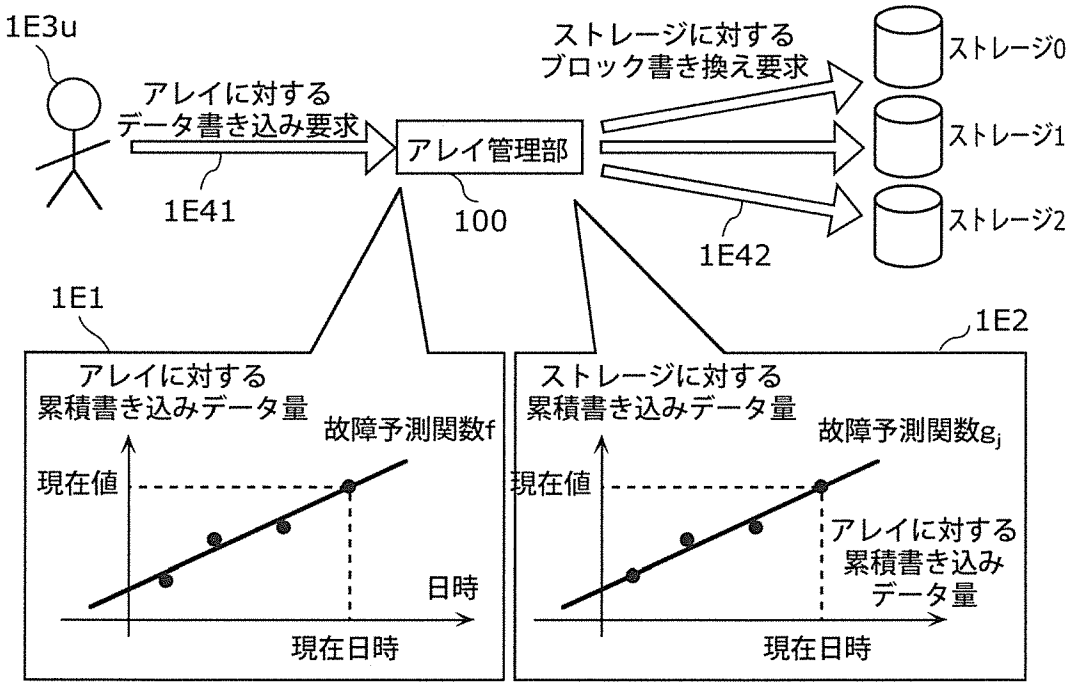
[図16]


S1102
データ・パリティ対応情報作成(図14)
データ・パリティ対応表における、
各ストレージにおけるパリティ割合を
全ストレージ0として初期化する
S1201
最初のストライプを選択する
S1202
選択ストライプに対して、
データ・パリティ対応情報を
割り当てる(図17)
S1203
全ストライプに対して、
対応情報の作成完了?
S1204
YES
NO
次のストライプを選択する
S1205
終了

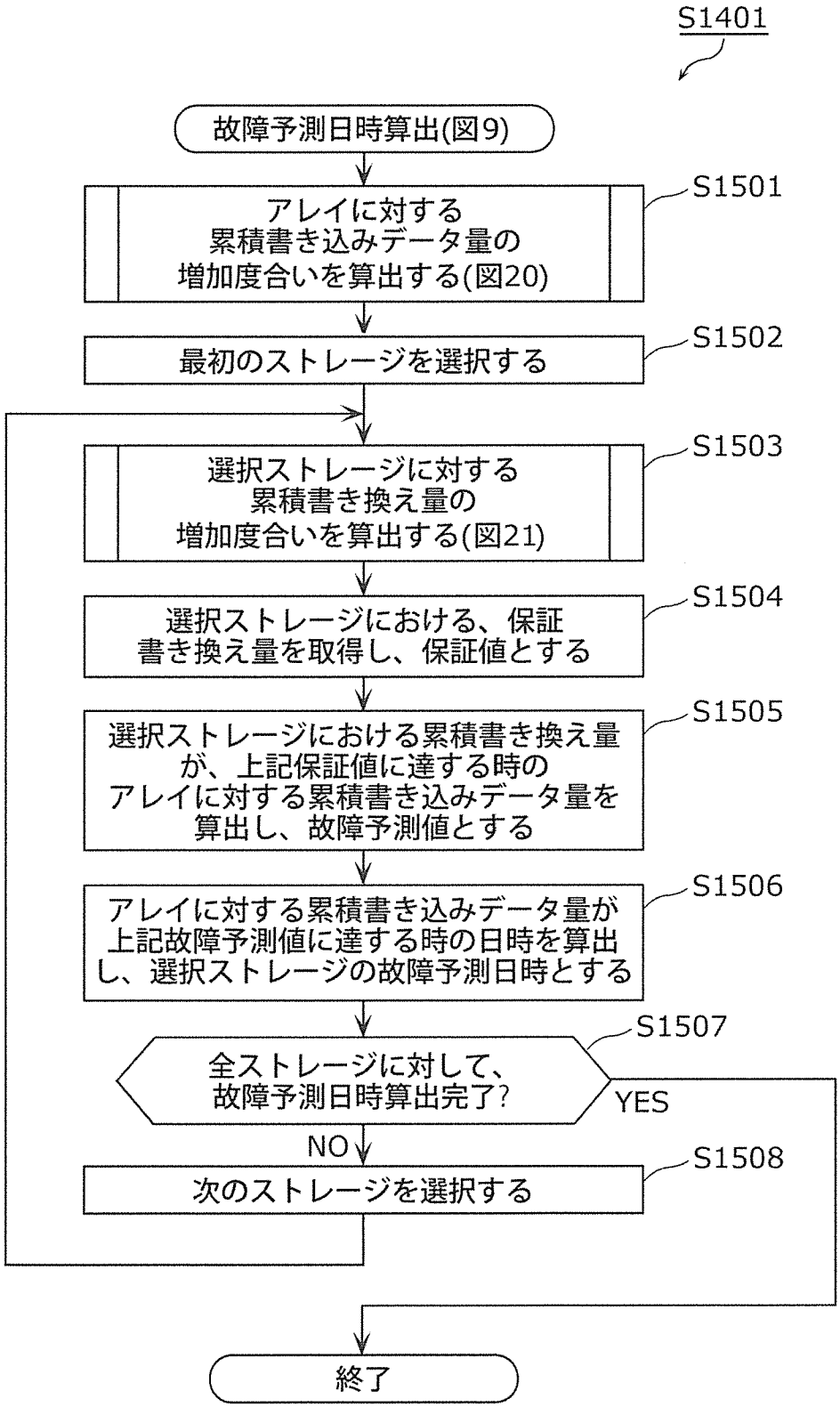
[図17]


S1203
ストライプに対するデータ・
パリティ対応情報割り当て(図16)
パリティ割合の目標値から、各ストレージにおける
パリティ割合を減じた値が大きい方から順に所定の
台数だけストレージを選択する
PAID5なら所定台数は1
PAID6なら所定台数は2
S1301
選択ストライプにおける、上記選択
ストレージに対応するブロックについて
パリティ領域として対応情報を割り当て、
それ以外のブロックについて
データ領域として対応情報を割り当てる
S1302
データ・パリティ対応表における、
各ストレージにおけるパリティ割合を
再計算する
S1303
終了

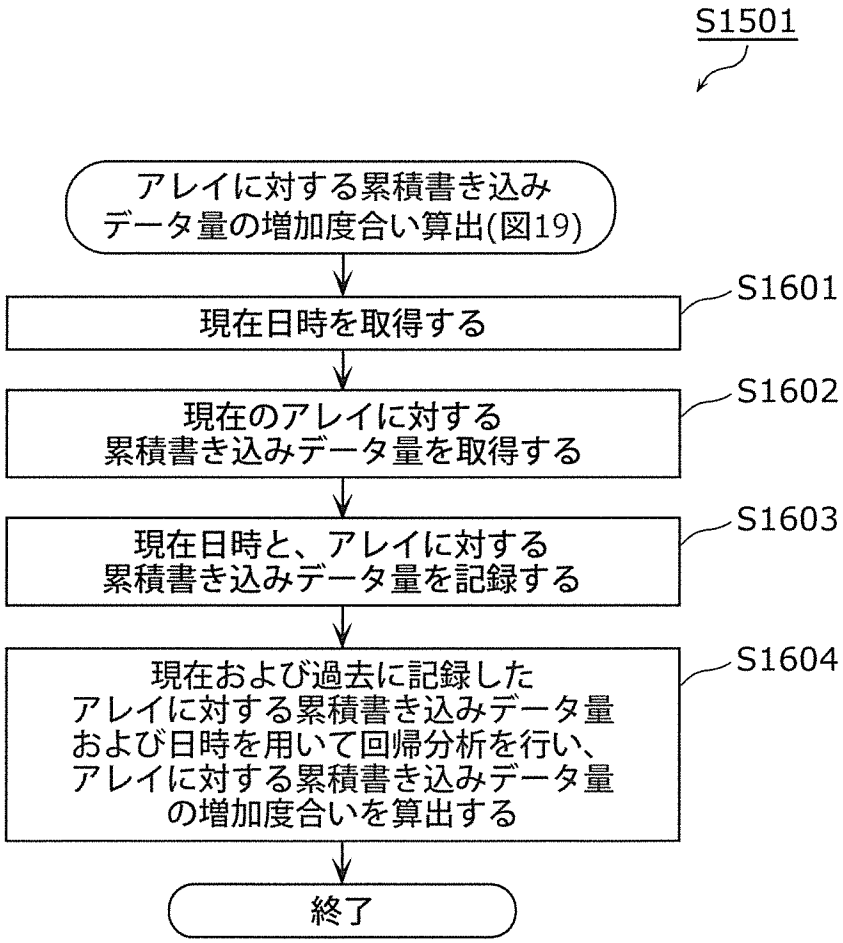
[図18]


1E3u
アレイに対する
データ書き込み要求
1E41
アレイ管理部
100
ストレージに対する
ブロック書き換え要求
1E42
ストレージ0
ストレージ1
ストレージ2
1E1
アレイに対する
累積書き込みデータ量
故障予測関数f
現在値
日時
現在日時
1E2
ストレージに対する
累積書き込みデータ量
故障予測関数g_j
現在値
アレイに対する
累積書き込み
データ量
現在日時

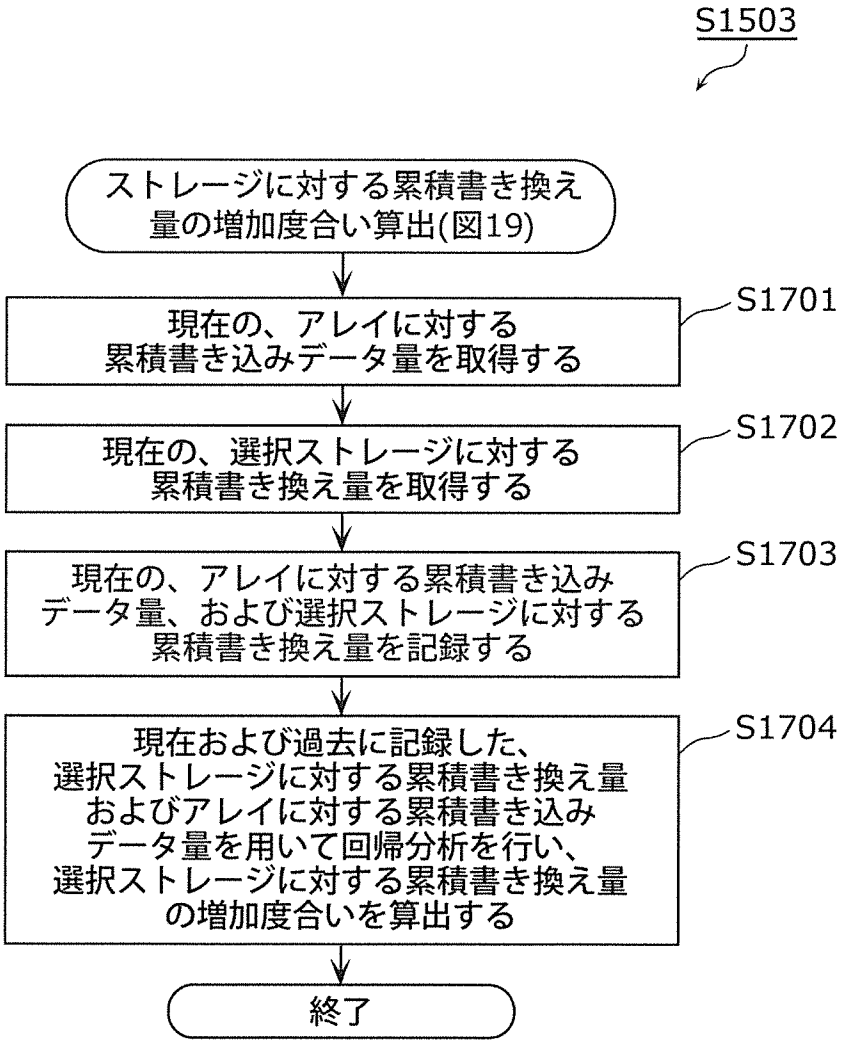
[図19]


S1401
故障予測日時算出(図9)
アレイに対する
累積書き込みデータ量の
増加度合いを算出する(図20)
S1501
最初のストレージを選択する
S1502
選択ストレージに対する
累積書き換え量の
増加度合いを算出する(図21)
S1503
選択ストレージにおける、保証
書き換え量を取得し、保証値とする
S1504
選択ストレージにおける累積書き換え量
が、上記保証値に達する時の
アレイに対する累積書き込みデータ量を
算出し、故障予測値とする
S1505
アレイに対する累積書き込みデータ量が
上記故障予測値に達する時の日時を算出
し、選択ストレージの故障予測日時とする
S1506
全ストレージに対して、
故障予測日時算出完了?
S1507
YES
NO
次のストレージを選択する
S1508
終了

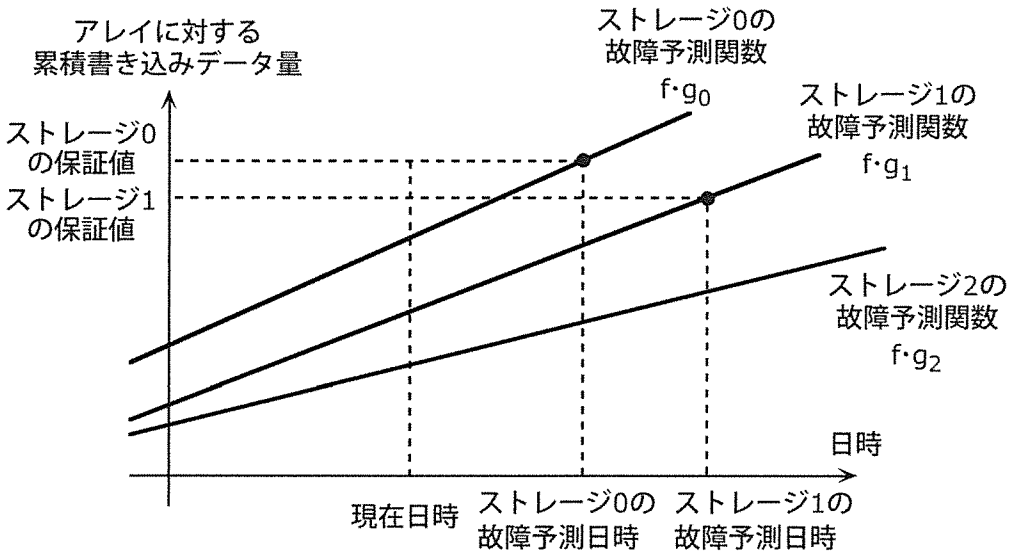
[図20]


S1501
アレイに対する累積書き込み
データ量の増加度合い算出(図19)
S1601
現在日時を取得する
S1602
現在のアレイに対する
累積書き込みデータ量を取得する
S1603
現在日時と、アレイに対する
累積書き込みデータ量を記録する
S1604
現在および過去に記録した
アレイに対する累積書き込みデータ量
および日時を用いて回帰分析を行い、
アレイに対する累積書き込みデータ量
の増加度合いを算出する
終了

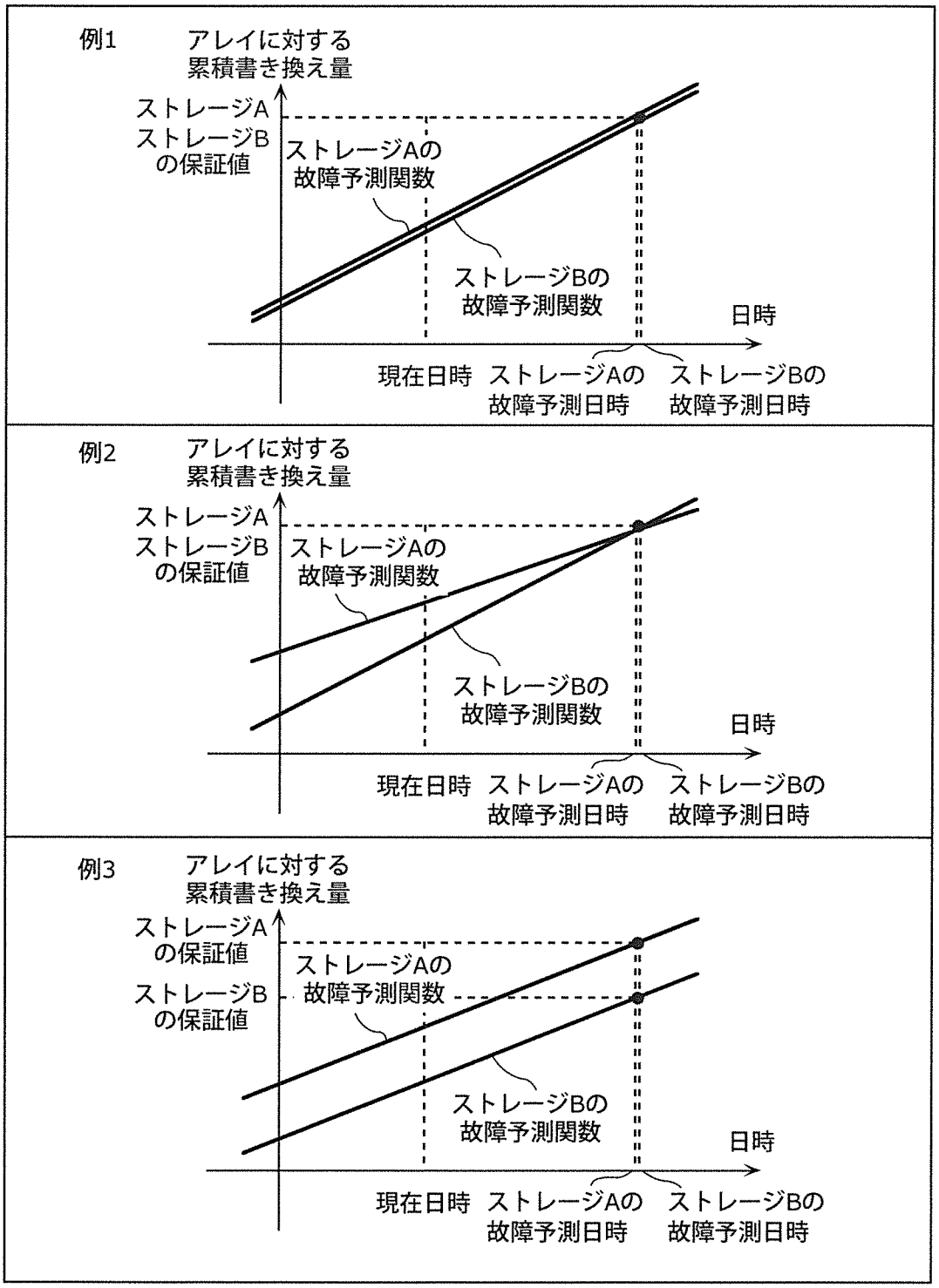
[図21]


S1503
ストレージに対する累積書き換え量の増加度合い算出(図19)
現在の、アレイに対する累積書き込みデータ量を取得する
S1701
現在の、選択ストレージに対する累積書き換え量を取得する
S1702
現在の、アレイに対する累積書き込みデータ量、および選択ストレージに対する累積書き換え量を記録する
S1703
現在および過去に記録した、選択ストレージに対する累積書き換え量およびアレイに対する累積書き込みデータ量を用いて回帰分析を行い、選択ストレージに対する累積書き換え量の増加度合いを算出する
S1704
終了

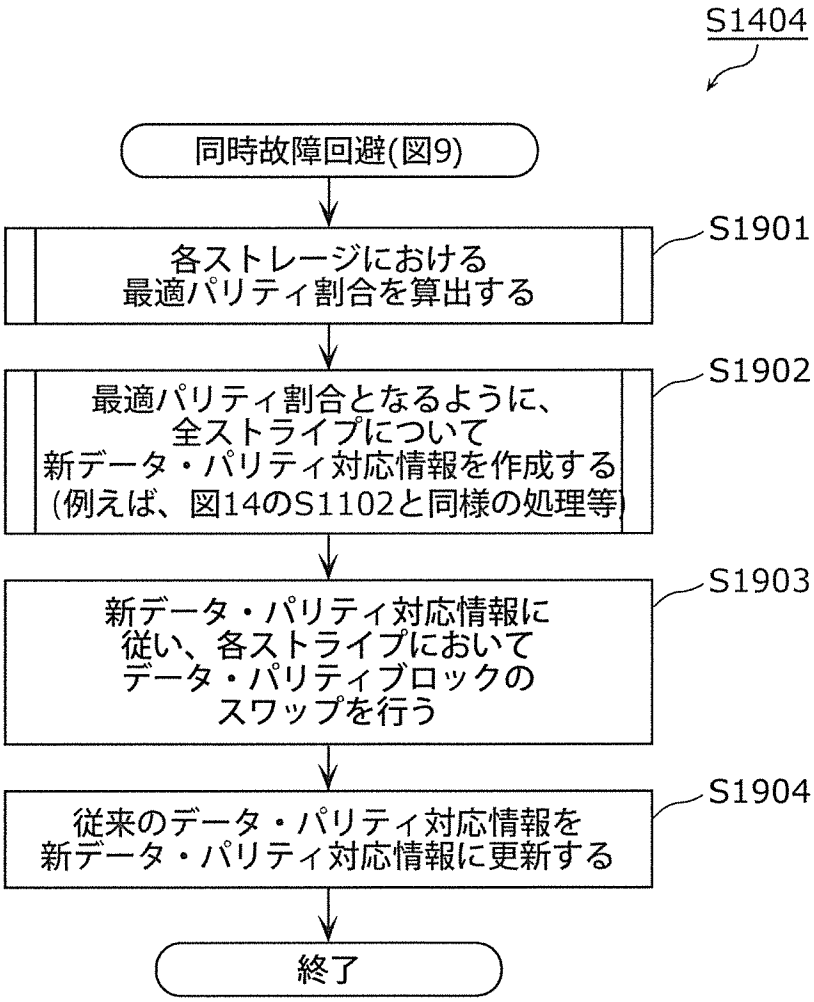
[図22]


アレイに対する累積書き込みデータ量
ストレージ0の故障予測関数
f·g0
ストレージ1の故障予測関数
f·g1
ストレージ2の故障予測関数
f·g2
ストレージ0の保証値
ストレージ1の保証値
日時
現在日時
ストレージ0の故障予測日時
ストレージ1の故障予測日時

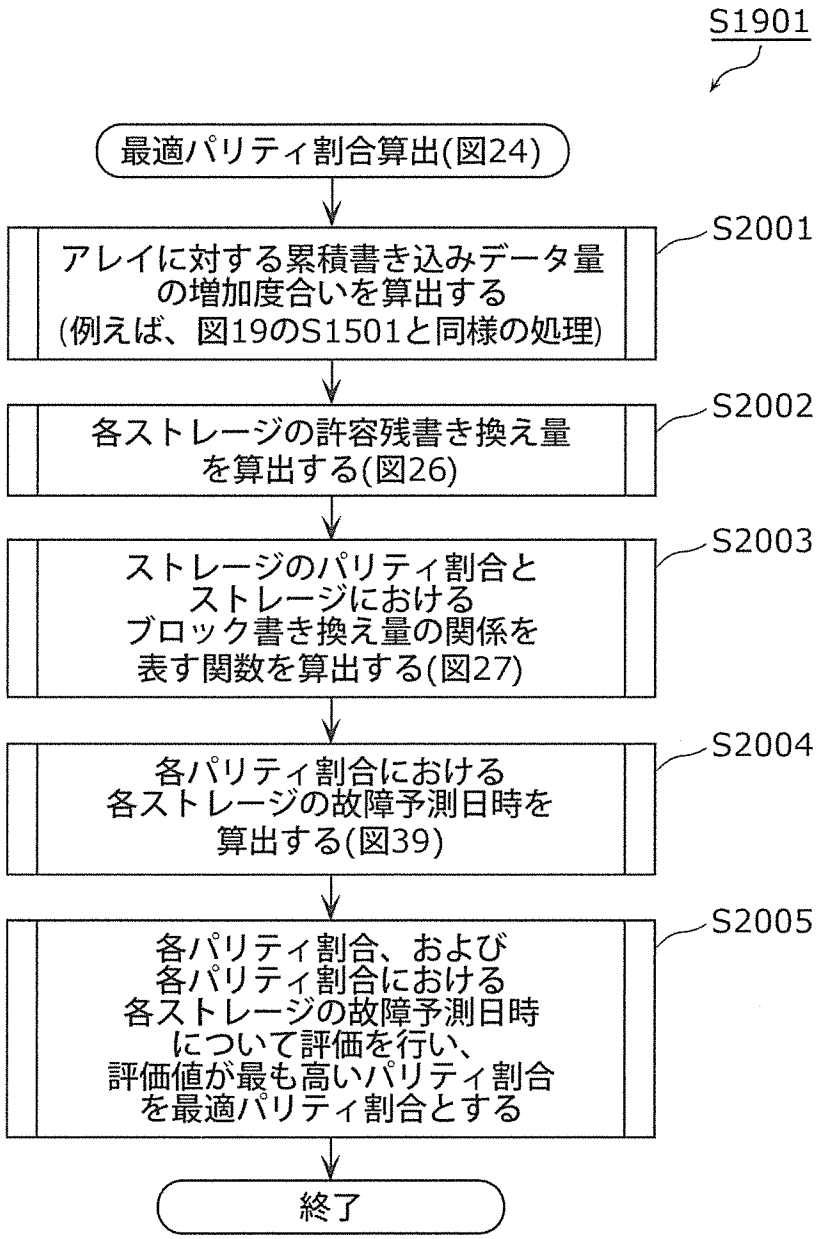
[図23]


例1
アレイに対する
累積書き換え量
ストレージA
ストレージB
の保証値
ストレージAの
故障予測関数
ストレージBの
故障予測関数
日時
現在日時
ストレージAの
故障予測日時
ストレージBの
故障予測日時
例2
アレイに対する
累積書き換え量
ストレージA
ストレージB
の保証値
ストレージAの
故障予測関数
ストレージBの
故障予測関数
日時
現在日時
ストレージAの
故障予測日時
ストレージBの
故障予測日時
例3
アレイに対する
累積書き換え量
ストレージA
の保証値
ストレージB
の保証値
ストレージAの
故障予測関数
ストレージBの
故障予測関数
日時
現在日時
ストレージAの
故障予測日時
ストレージBの
故障予測日時

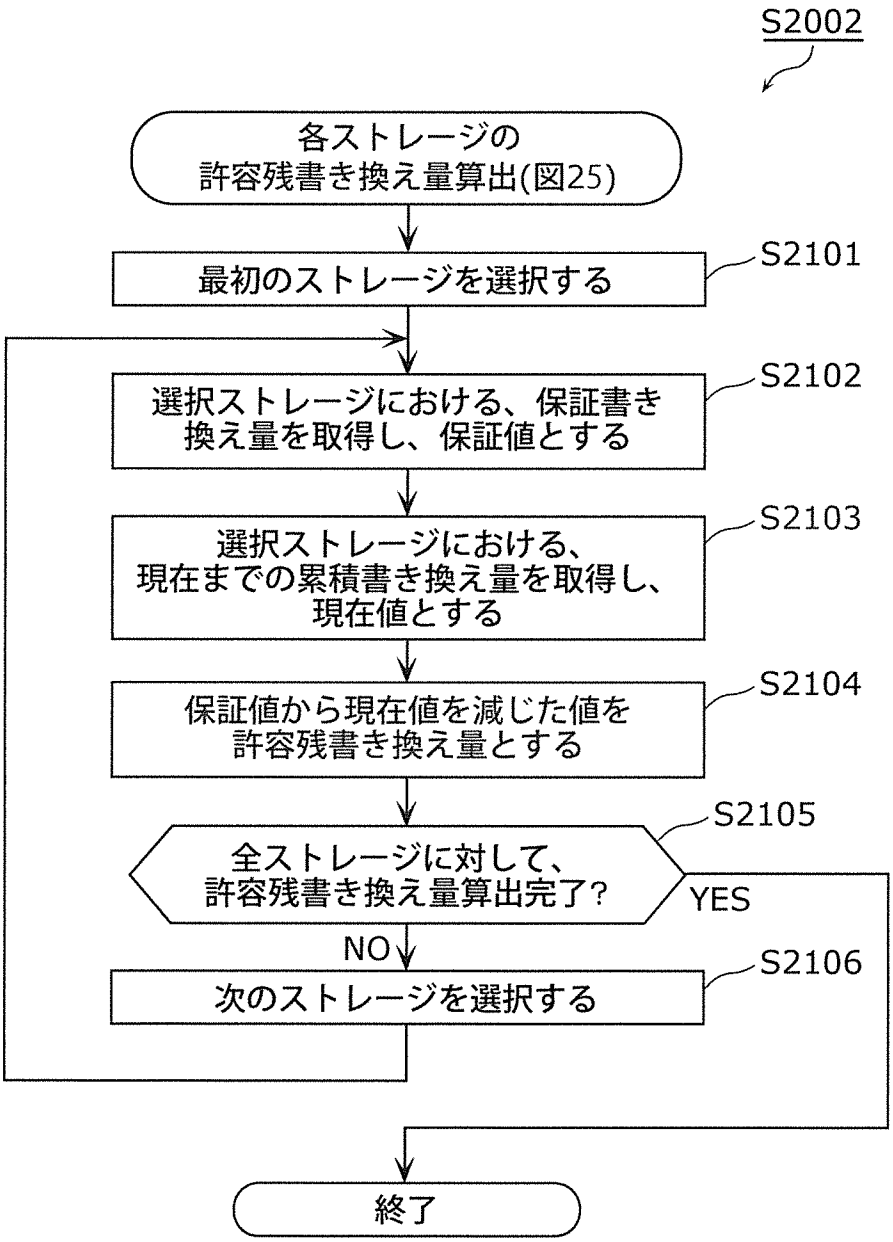
[図24]


S1404
同時故障回避(図9)
各ストレージにおける
最適パリティ割合を算出する
S1901
最適パリティ割合となるように、
全ストライプについて
新データ・パリティ対応情報を作成する
(例えば、図14のS1102と同様の処理等)
S1902
新データ・パリティ対応情報に
従い、各ストライプにおいて
データ・パリティブロックの
スワップを行う
S1903
従来のデータ・パリティ対応情報を
新データ・パリティ対応情報に更新する
S1904
終了

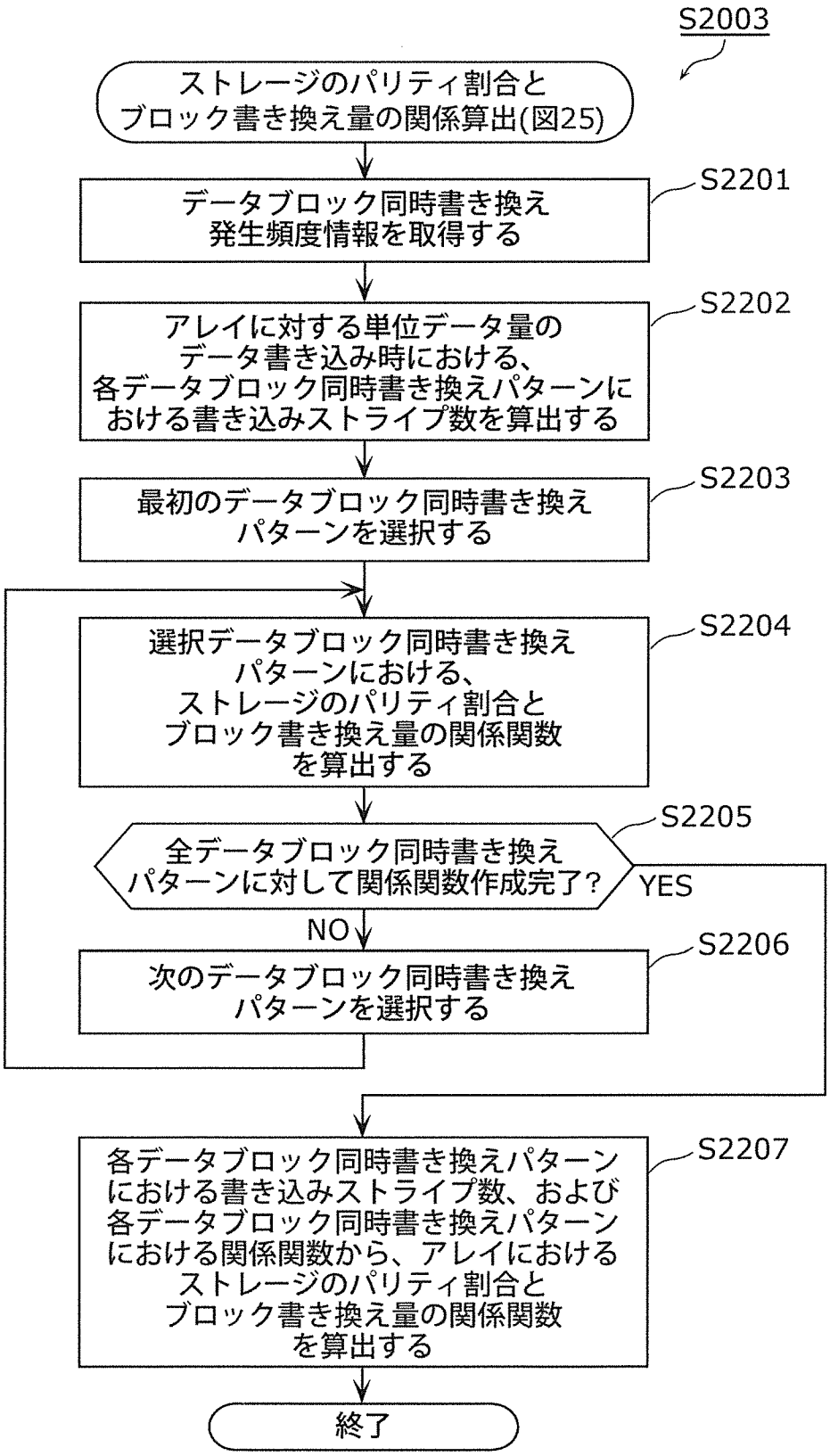
[図25]


S1901
最適パリティ割合算出(図24)
アレイに対する累積書き込みデータ量
の増加度合いを算出する
(例えば、図19のS1501と同様の処理)
S2001
各ストレージの許容残書き換え量
を算出する(図26)
S2002
ストレージのパリティ割合と
ストレージにおける
ブロック書き換え量の関係を
表す関数を算出する(図27)
S2003
各パリティ割合における
各ストレージの故障予測日時を
算出する(図39)
S2004
各パリティ割合、および
各パリティ割合における
各ストレージの故障予測日時
について評価を行い、
評価値が最も高いパリティ割合
を最適パリティ割合とする
S2005
終了

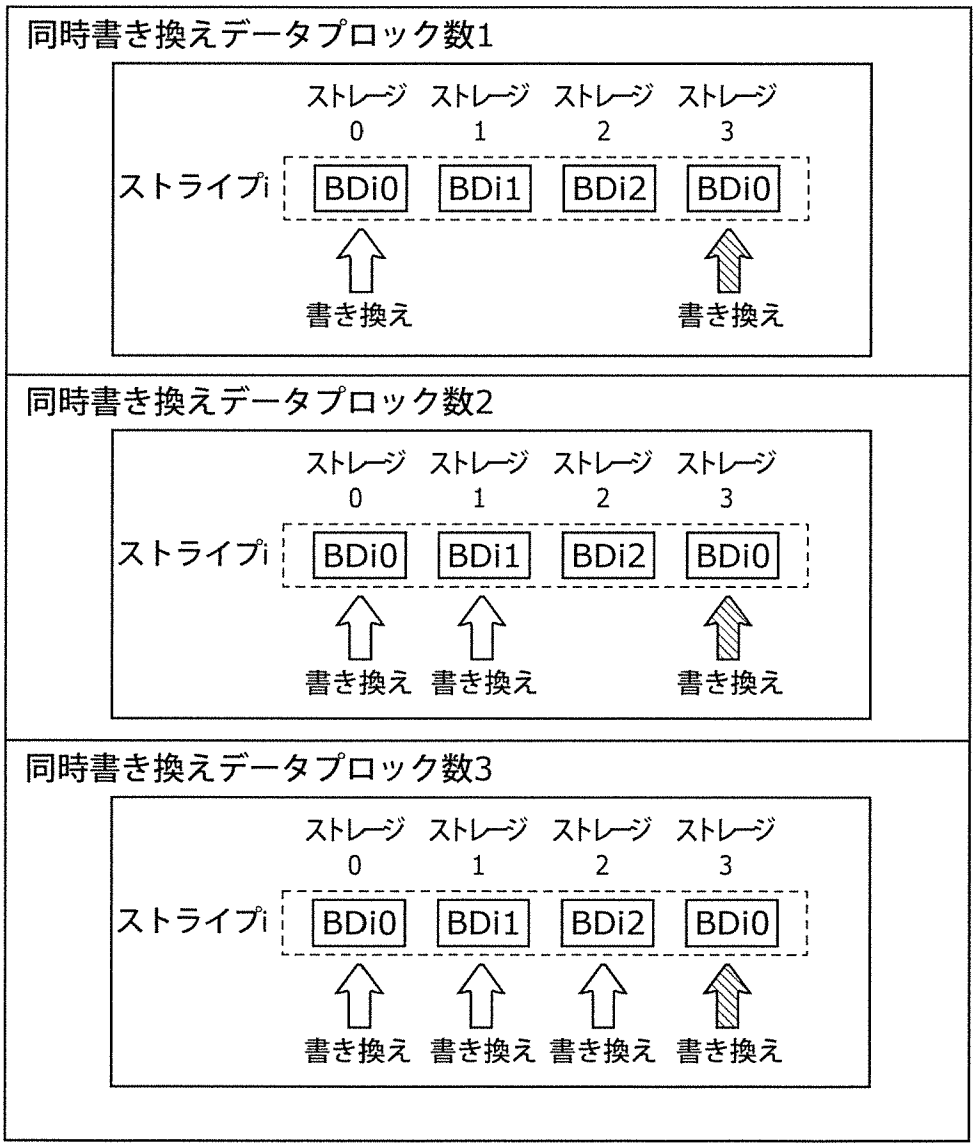
[図26]


S2002
各ストレージの
許容残書き換え量算出(図25)
最初のストレージを選択する
S2101
選択ストレージにおける、保証書き
換え量を取得し、保証値とする
S2102
選択ストレージにおける、
現在までの累積書き換え量を取得し、
現在値とする
S2103
保証値から現在値を減じた値を
許容残書き換え量とする
S2104
全ストレージに対して、
許容残書き換え量算出完了?
S2105
YES
NO
次のストレージを選択する
S2106
終了

[図27]


S2003
ストレージのパリティ割合と
ブロック書き換え量の関係算出(図25)
S2201
データブロック同時書き換え
発生頻度情報を取得する
S2202
アレイに対する単位データ量の
データ書き込み時における、
各データブロック同時書き換えパターンに
おける書き込みストライプ数を算出する
S2203
最初のデータブロック同時書き換え
パターンを選択する
S2204
選択データブロック同時書き換え
パターンにおける、
ストレージのパリティ割合と
ブロック書き換え量の関係関数
を算出する
S2205
全データブロック同時書き換え
パターンに対して関係関数作成完了?
YES
NO
S2206
次のデータブロック同時書き換え
パターンを選択する
S2207
各データブロック同時書き換えパターン
における書き込みストライプ数、および
各データブロック同時書き換えパターン
における関係関数から、アレイにおける
ストレージのパリティ割合と
ブロック書き換え量の関係関数
を算出する
終了

[図28]



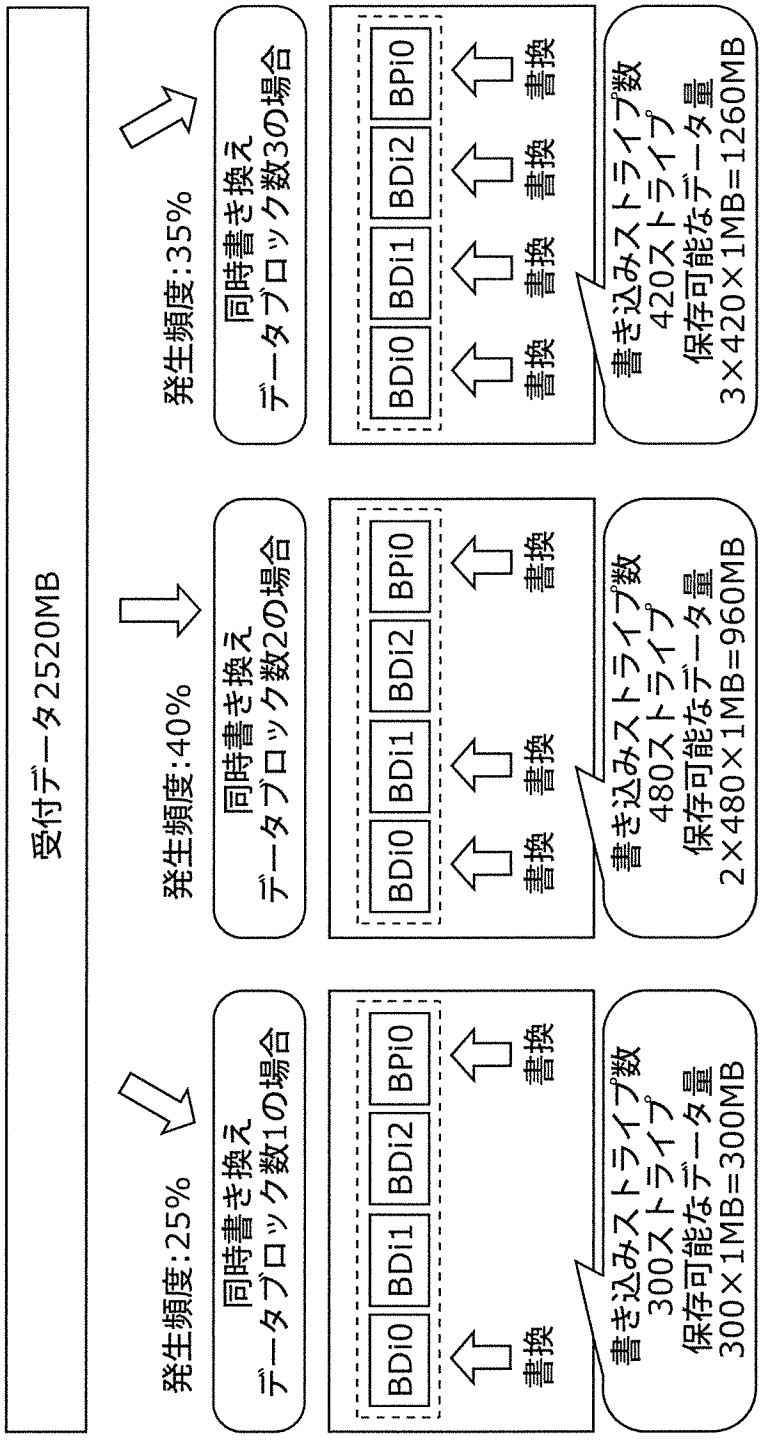
[図29]

同時書き換え データブロック数	1回のストライプ書き込みで 保存可能なデータ量
1	1MB
2	2MB
3	3MB

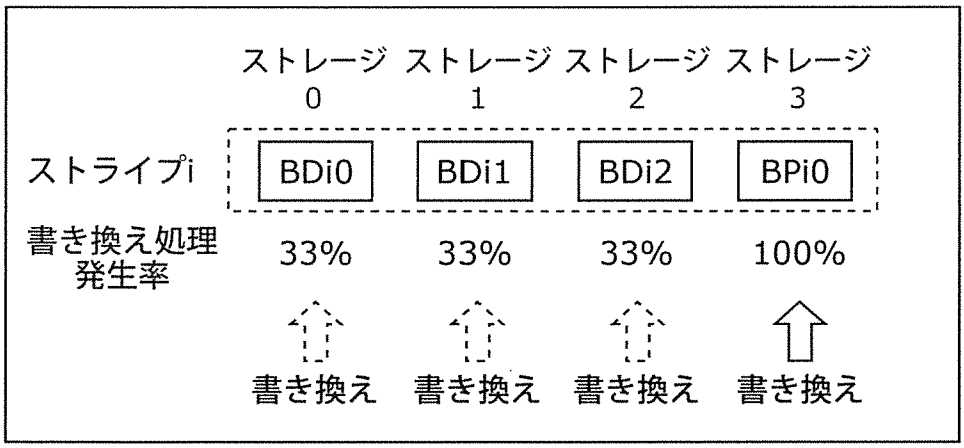
[図30]

同時書き換えデータブロック数	累積発生ストライプ数	発生頻度
1	60000	25%
2	96000	40%
3	84000	35%
(合計)	240000	

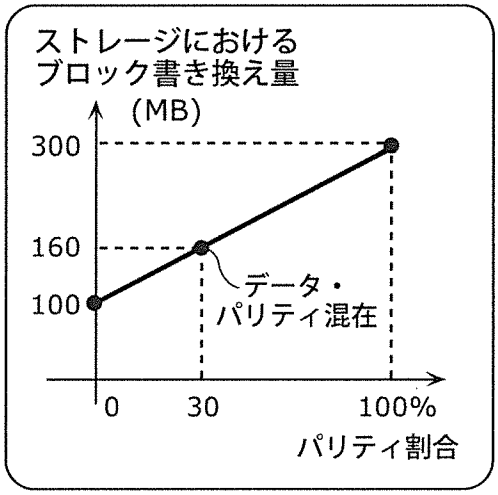
[図31]



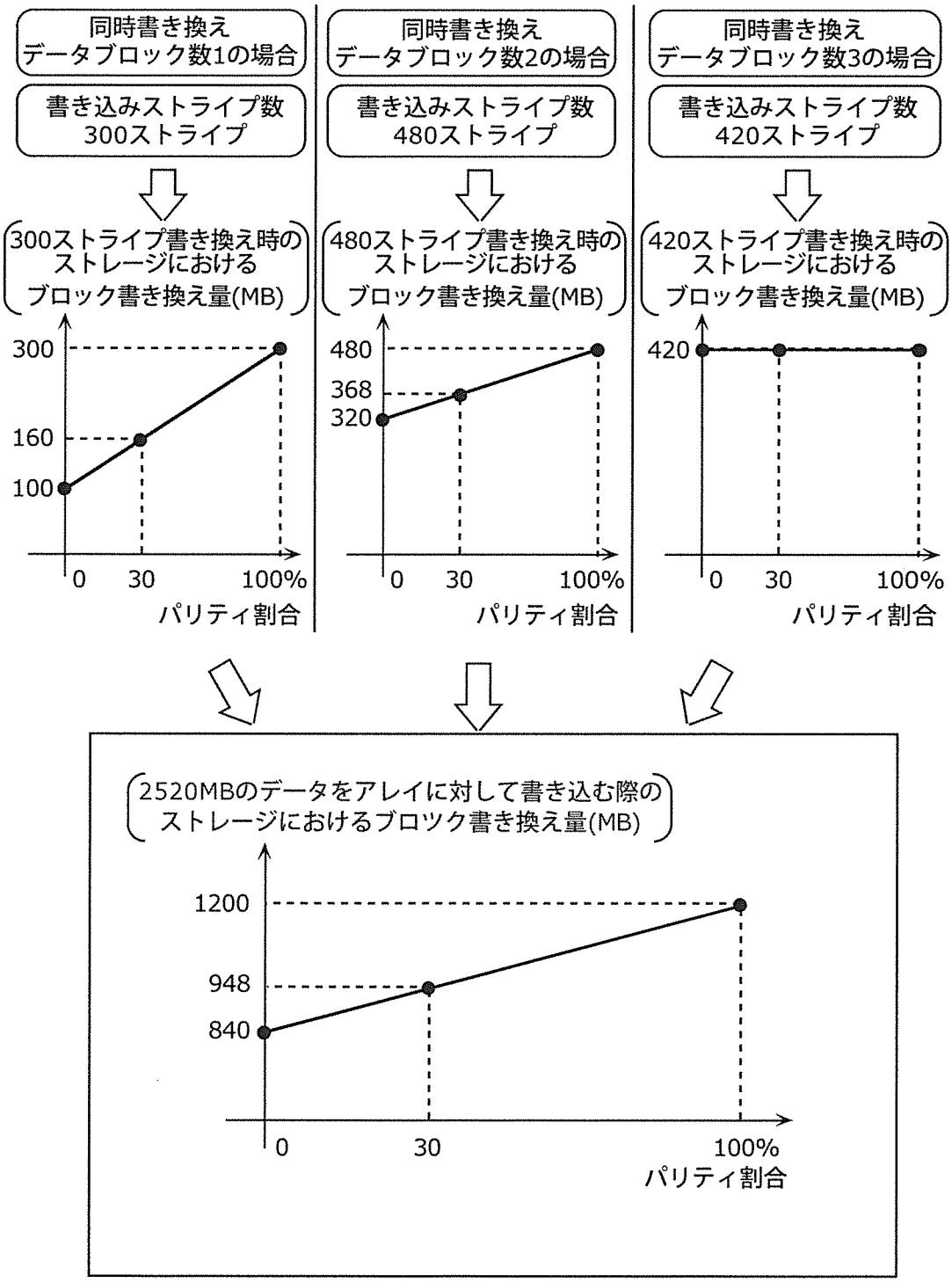
[図32]


ストレージ 0
ストレージ 1
ストレージ 2
ストレージ 3
ストライプi
BDi0
BDi1
BDi2
BPi0
書き換え処理発生率
33%
33%
33%
100%
書き換え
書き換え
書き換え
書き換え

[図33]


ストレージにおける
ブロック書き換え量
(MB)
300
160
100
データ・
パリティ混在
0
30
100%
パリティ割合

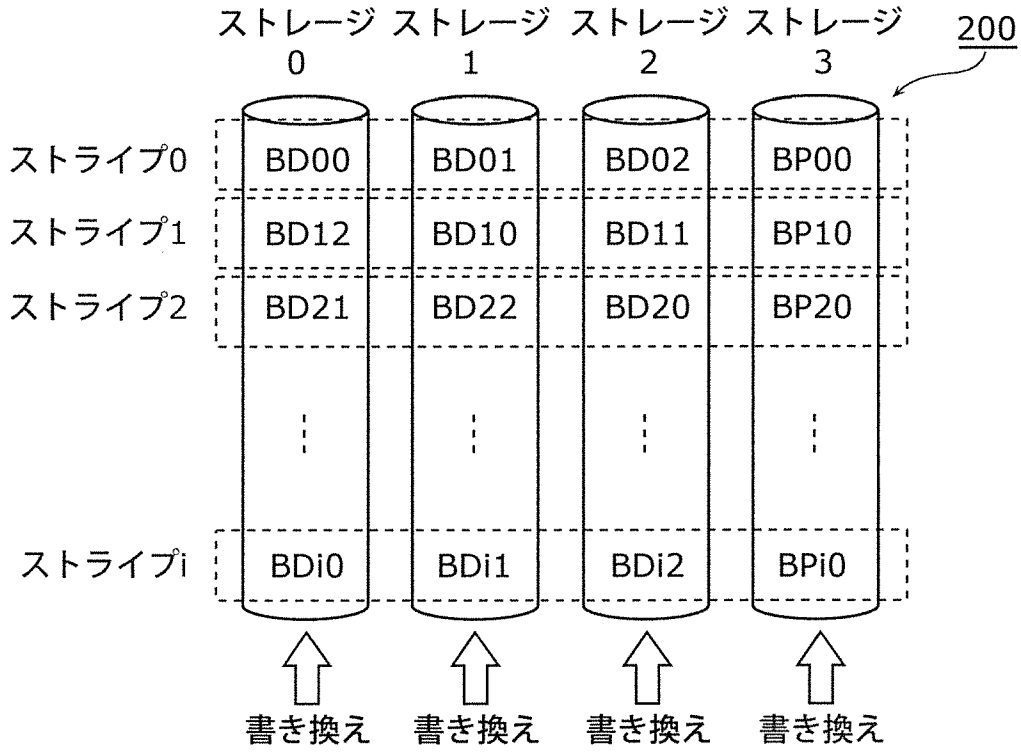
[図34]



[図35A]

	ストレージ				合計
	0	1	2	3	
パリティ割合	0%	0%	0%	100%	100%

[図35B]



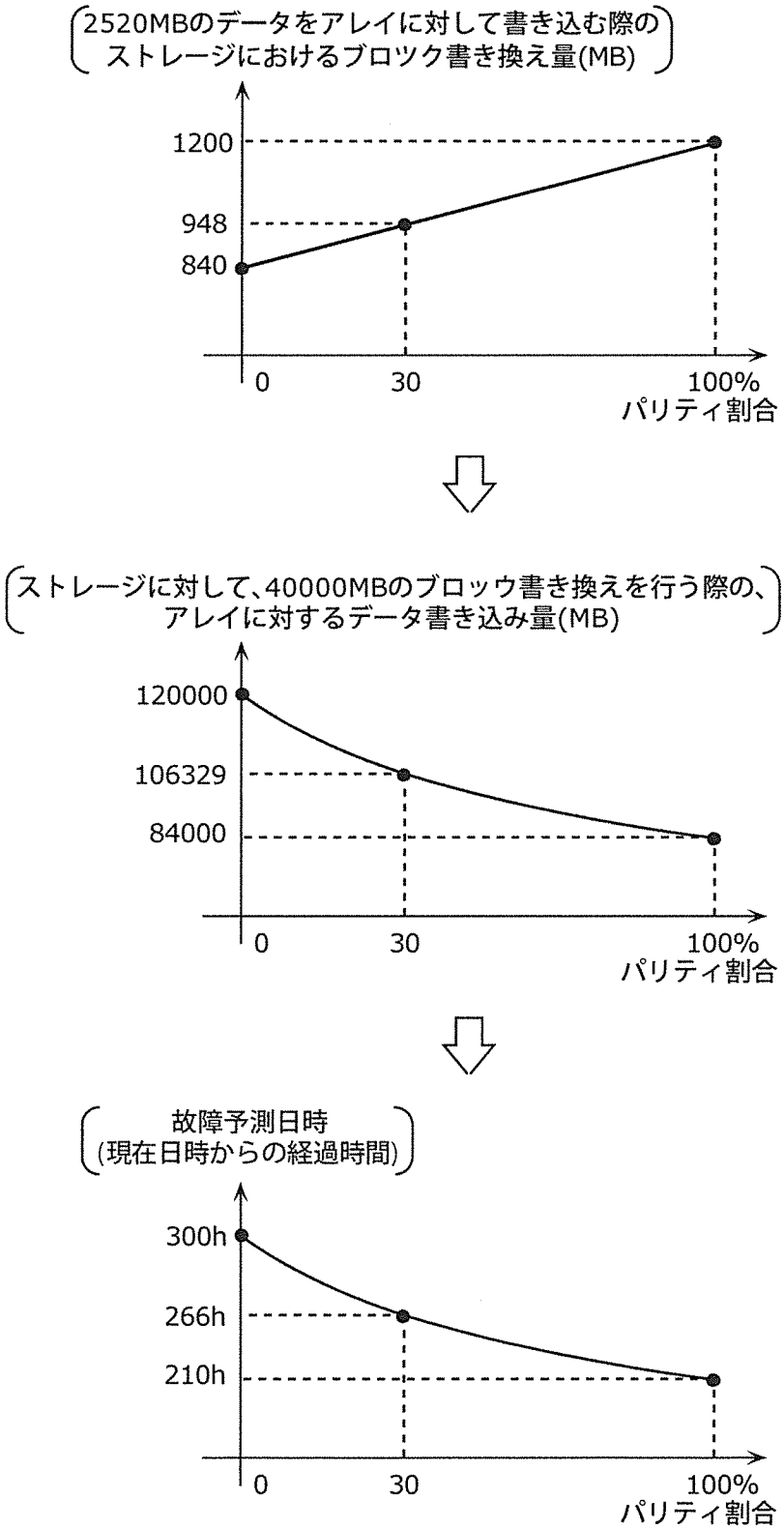
[図35C]

	ストレージ				合計
	0	1	2	3	
書き換え量	100MB	100MB	100MB	300MB	600MB

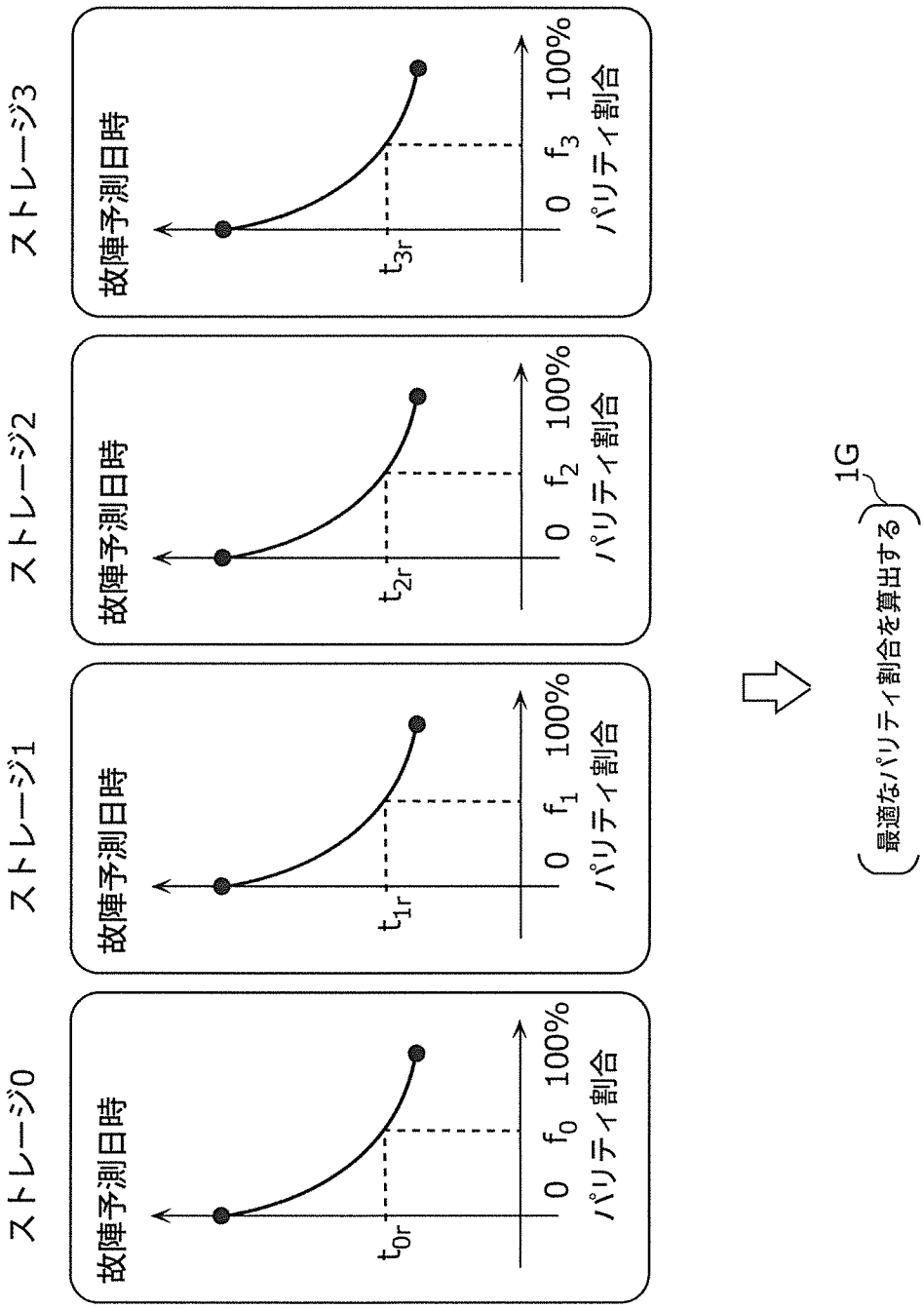
[図36]

	ストレージ				合計
	0	1	2	3	
パリティ割合	30%	20%	15%	35%	100%

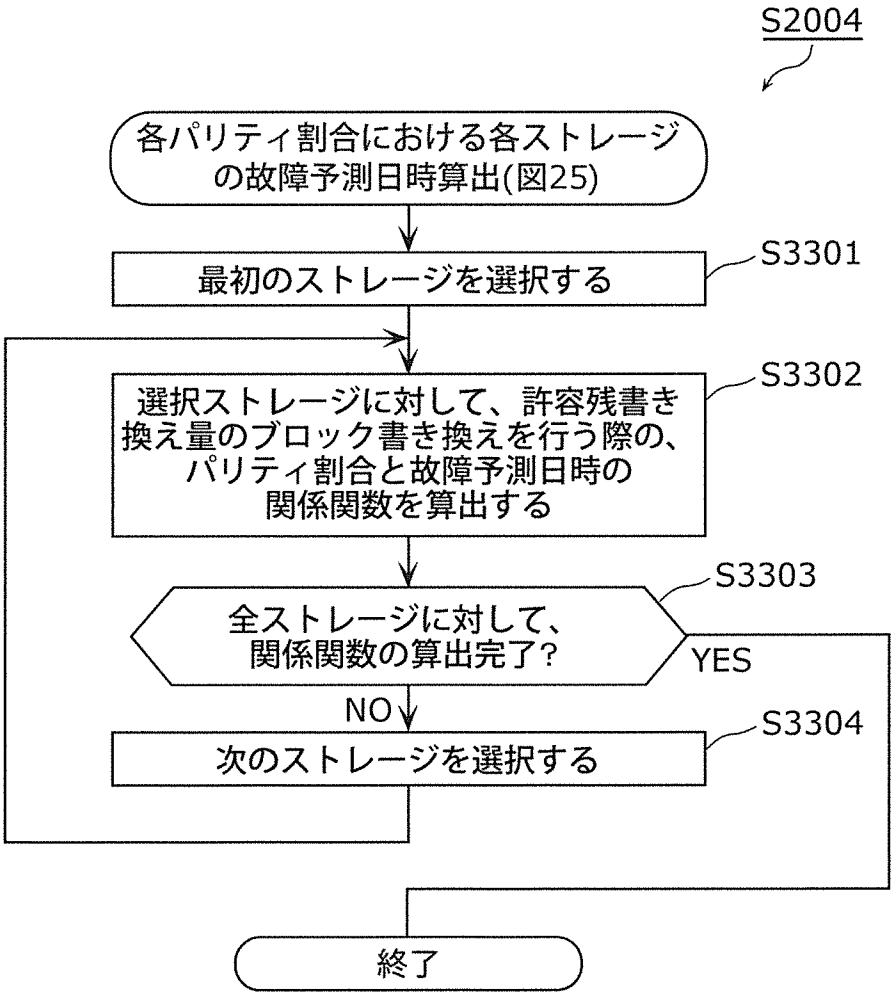
[図37]


2520MBのデータをアレイに対して書き込む際の
ストレージにおけるブロック書き換え量(MB)
1200
948
840
0
30
100%
パリティ割合
ストレージに対して、40000MBのブロッウ書き換えを行う際の、
アレイに対するデータ書き込み量(MB)
120000
106329
84000
0
30
100%
パリティ割合
故障予測日時
(現在日時からの経過時間)
300h
266h
210h
0
30
100%
パリティ割合

[図38]


ストレージ0
故障予測日時
t_{0r}
0 f_0 100%
パリティ割合
ストレージ1
故障予測日時
t_{1r}
0 f_1 100%
パリティ割合
ストレージ2
故障予測日時
t_{2r}
0 f_2 100%
パリティ割合
ストレージ3
故障予測日時
t_{3r}
0 f_3 100%
パリティ割合
1G
（最適なパリティ割合を算出する）

[図39]


S2004
各パリティ割合における各ストレージの故障予測日時算出(図25)
最初のストレージを選択する
S3301
選択ストレージに対して、許容残書き換え量のブロック書き換えを行う際の、パリティ割合と故障予測日時の関係関数を算出する
S3302
全ストレージに対して、関係関数の算出完了?
S3303
YES
NO
次のストレージを選択する
S3304
終了

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2012/002116

A. CLASSIFICATION OF SUBJECT MATTER
G06F13/10(2006.01)i, *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)
G06F13/10, 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	1996-2012
Kokai Jitsuyo Shinan Koho	1971-2012	Toroku Jitsuyo Shinan Koho	1994-2012

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2010-015516 A (Toshiba Corp.), 21 January 2010 (21.01.2010), paragraphs [0006] to [0016], [0045], [0059] to [0070] & US 2010/0005228 A1	1-8
Y	JP 2009-037304 A (Hitachi, Ltd.), 19 February 2009 (19.02.2009), paragraphs [0064] to [0066] & US 2009/0037656 A1	1-3,6-8
Y	JP 2007-206993 A (Fujitsu Ltd.), 16 August 2007 (16.08.2007), paragraphs [0027], [0039] to [0042], [0063]; fig. 14, 15 & US 2007/0180294 A1	1-3,6-8

☐ 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 21 June, 2012 (21.06.12)	Date of mailing of the international search report 03 July, 2012 (03.07.12)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１２／００２１１６

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F13/10(2006.01)i, 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.　G06F13/10, 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	１９７１－２０１２年
日本国実用新案登録公報	１９９６－２０１２年
日本国登録実用新案公報	１９９４－２０１２年

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

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
X	JP 2010-015516 A（株式会社東芝）2010.01.21, 段落【０００６】－【００１６】,【００４５】, 【００５９】－【００７０】 & US 2010/0005228 A1	1-8
Y	JP 2009-037304 A（株式会社日立製作所）2009.02.19, 段落【００６４】－【００６６】 & US 2009/0037656 A1	1-3, 6-8
Y	JP 2007-206993 A（富士通株式会社）2007.08.16, 段落【００２７】,【００３９】－【００４２】,【００６３】, 図１４，１５ & US 2007/0180294 A1	1-3, 6-8

☐　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	国際調査報告の発送日
２１．０６．２０１２	０３．０７．２０１２

国際調査機関の名称及びあて先

日本国特許庁（ＩＳＡ／ＪＰ）

郵便番号１００－８９１５

東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｔ　３５６５

横山　佳弘

電話番号　０３－３５８１－１１０１　内線　３５６８