

(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9月13日(13.09.2012)



(10) 国際公開番号

WO 2012/120660 A1

- (51) 国際特許分類:
G06F 12/08 (2006.01)
- (21) 国際出願番号: PCT/JP2011/055488
- (22) 国際出願日: 2011年3月9日(09.03.2011)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人(米国を除く全ての指定国について): 富士通株式会社(FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者; および
- (75) 発明者/出願人(米国についてのみ): 金野 雄次(KONNO, Yuuji) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP). 黒田 康弘(KURODA, Yasuhiro) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).

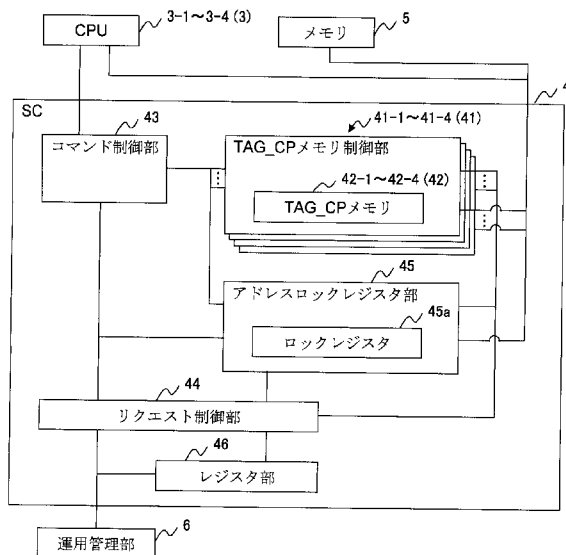
- (74) 代理人: 真田 有, 外(SANADA, Tamotsu et al.); 〒1800004 東京都武蔵野市吉祥寺本町1丁目10番31号 NOF吉祥寺本町ビル5階 Tokyo (JP).
- (81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, 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.
- (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,

[続葉有]

(54) Title: INFORMATION PROCESSING DEVICE AND CACHE CONTROL METHOD

(54) 発明の名称: 情報処理装置、及びキャッシュ制御方法

[図2]



- 5 Memory
6 Operation management unit
41-1 ~ 41-4 (41) TAG_CP memory control unit
42-1 ~ 42-4 (42) TAG_CP memory
43 Command control unit
44 Request control unit
45 Address lock register unit
45a Lock register
46 Register unit

UEに係るリクエストを再発行する。

(57) Abstract: A system controller (4), in the case that there is an uncorrectable error (UE) in data read out from a second tag memory (42) corresponding to a first tag memory (32) of a processing device (3), notifies the processing device (3) of way information of the second tag memory (42) where the UE has occurred. On the basis of the notified way information, the processing device (3) causes the corresponding way in the first tag memory (32) to be degraded and notifies the system controller (4) of the completion of the degrading processing. The system controller (4) causes the way in which the UE has occurred in the second tag memory (42) to be degraded, and after receiving notification from the processing device (3) that the degrading processing of the first tag memory (32) has completed, reissues the request relating to the UE.

(57) 要約: システムコントローラ(4)は、演算処理装置(3)の第1タグメモリ(32)に対応する第2タグメモリ(42)から読み出されたデータが訂正不可能なエラーであるUE(Uncorrectable Error)を起こした場合に、前記UEが発生した前記第2タグメモリ(42)のウェイ情報を前記演算処理装置(3)に通知し、前記演算処理装置(3)は、通知されたウェイ情報に基づいて対応する前記第1タグメモリ(32)のウェイを縮退させ、縮退処理の完了を前記システムコントローラ(4)へ通知して、前記システムコントローラ(4)は、前記第2タグメモリ(42)のUEが発生したウェイを縮退させるとともに、前記演算処理装置(3)から前記第1タグメモリ(32)の縮退処理が完了した旨の通知を受信した後、当該

WO 2012/120660 A1

WO 2012/120660 A1



NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI 添付公開書類:

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, — 国際調査報告 (条約第 21 条(3))
NE, SN, TD, TG).

明 細 書

発明の名称： 情報処理装置、及びキャッシュ制御方法

技術分野

[0001] 本件は、情報処理装置、及びキャッシュ制御方法に関する。

背景技術

[0002] 近年、処理の高速化や耐障害性の向上を図るため、対称型マルチプロセッシング（SMP；Symmetric Multi Processor）方式を用いた、SMPサーバシステムが用いられることがある。

SMPは、複数のCPU（Central Processing Unit）が同等な立場で処理を分担するマルチプロセッサ手法であり、CPUキャッシュを同期させる機能や処理に用いられる各種資源を管理する機能を備える。

[0003] SMPサーバシステムは、複数のCPUやシステムコントローラ（System Controller；以下、SCという）や、RAM（Random Access Memory）等のメモリのほか、システムを制御するファームウェアが搭載された運用管理部などにより構成される。

このようなSMPサーバシステムでは、処理速度の向上のために、CPUのキャッシュタグ（TAG）データのコピー（TAG__CP）をSC内に備えることがある。この場合、各CPUからの問い合わせに対し、対象CPUの前段にあるSCにより、TAG__CPが参照されて応答が返される。これにより、スヌープ方式による高速なキャッシュアクセスを実現し、CPUのキャッシュメモリ（Cache Memory；以下、CMという）の同期処理の高速化を実現する。

[0004] なお、スヌープ方式とは、キャッシュコヒーレンシのアルゴリズムの一つであり、他のキャッシュと更新状態の情報を交換することで、どのキャッシュに最新のデータが格納されているかを把握することができ、最新のデータを取得することができるものである。

また、近年、CPUのCMにおいて、キャッシュライン数の増加に伴い、

複数のWAYによるデータ格納構造であるセットアソシアティブ構成が採用されている。

[0005] セットアソシアティブ構成では、CPUのCMにおいて、キャッシュライン毎に複数のWAYが備えられ、各WAYにデータが格納される。

キャッシュタグデータは、CPU内部のTAGメモリ及びSC内部のTAG__CPメモリにそれぞれ格納され、インデックスと呼ばれるメモリの物理アドレスの一部を使用したアドレスにより管理される。キャッシュタグデータは、CPUからのリクエストに応じて、インデックスにより特定されるCM内の一のキャッシュラインから、一のWAYを絞り込み、CMから所望のデータを取得するのに用いられる。

[0006] なお、CM、TAGメモリ及びTAG__CPメモリとしては、いずれもSRAM (Static RAM) 等のRAMが挙げられる。

上述したSMPサーバシステムでは、CPU、CM、TAGメモリ又はTAG__CPメモリ等の障害が検出されると、運用管理部によって、異常が発生した箇所をシステムから切り離す縮退処理が行なわれる。この縮退処理により、システムの動作を中断することなく運用を継続することができ、耐障害性の向上が実現される。

[0007] 特にミッションクリティカルな分野で使用される大規模SMPサーバシステムで、TAG__CPメモリ等の故障が発生した場合、システムの性能が低下することになるとしても、被疑箇所を切り離して運用を継続することが望ましい。このため、従来のSMPサーバシステムには、CPU内部のTAGメモリやSC内部のTAG__CPメモリの固定的な1ビット (bit) 故障が発生した場合、被疑箇所のWAYを動的に縮退し、運用を停止させずに故障箇所を切り離す仕組みが実装されている。

[0008] なお、1ビット故障では、キャッシュタグデータに含まれるエラー訂正符号 (Error Correcting Code ; 以下、ECCという) により、エラーの訂正が可能である。以下、1ビット故障を訂正可能エラー又はCE (Correctable Error) という。

以下、SC内のTAG_CPメモリでCEが発生した場合の、システムの縮退処理の動作を説明する。

[0009] 図10は、SC400内のTAG_CPメモリ420-2でCEが発生した場合の縮退範囲を示す図であり、図11は、SC400内のTAG_CPメモリ420-2でCEが発生した場合の縮退処理を説明するためのフローチャートである。

図10に例示するように、SMPサーバシステムでは、システムボード（SB；System Board：以下、SBという）200と、運用管理部600とが備えられる。

[0010] また、SB200には、CPU300-1～300-4と、SC400と、メモリ500とが備えられる。なお、以下の説明においてCPU300-1～300-4を区別しない場合には、単にCPU300という。

各CPU300-1～300-4には、それぞれCM310-1～310-4及びTAGメモリ320-1～320-4が備えられる。なお、CM310-1～310-4及びTAGメモリ320-1～320-4の符号におけるハイフン“-”の右側の数字は、CM310-1～310-4及びTAGメモリ320-1～320-4が、それぞれ対応する数字のCPU300-1～300-4に備えられることを示す。

[0011] さらに、SC400には、TAGメモリ320-1～320-4に対応したTAG_CPメモリ420-1～420-4が備えられる。なお、以下の説明においてTAG_CPメモリ420-1～420-4を区別しない場合には、単にTAG_CPメモリ420という。

図10及び図11に示すように、システムの運用中、SC400内のTAG_CPメモリ420-2にCEが発生し、SC400により検出された場合（ステップS101）、SC400からCEが発生したTAG_CPメモリ420に対応するCPU300-2に対して、被疑箇所の情報が通知される（ステップS102）。なお、この情報には、ECCにより訂正された被疑箇所のインデックスと、WAY番号とが含まれる。

[0012] CPU300-2では、通知された被疑箇所に対応したTAGメモリ320-2内のWAYのデータがメモリに書き出されるとともに、当該WAYの縮退処理が行なわれる（ステップS103）。そして、CPU300-2により、SC400へ縮退処理完了の通知が行なわれる（ステップS104）。

縮退処理完了通知を受けたSC400では、被疑箇所のWAYに対して縮退処理が行なわれる（ステップS105）。そして、SC400により、運用管理部600に対して、縮退処理を行なったCPU300-2のWAY番号を含むエラー通知が行なわれ（ステップS106）、運用管理部600の制御情報に故障情報が記録される（ステップS107）。その後、SMPサーバシステムでは運用が継続される（ステップS108）。

[0013] 上述の如く、CPU300-2内部のTAGメモリ320-2の一部（WAY）及びSC400内部のTAG_CPメモリ420-2の一部（WAY）が縮退される（図10中、「縮退範囲」参照）。これにより、システムに多少の性能低下は発生するものの、縮退処理は動的に行なわれるため、運用の停止を回避することができる。

なお、ステップS107において運用管理部600の制御情報に記録された故障情報は、例えばSMPサーバシステムで実行中のOS（Operating System）の再起動等によってCPU及びSC内の縮退状態がリセットされた場合に、被疑箇所のWAYを再度縮退させるために用いられる。

[0014] ところで、SC内部のTAG_CPメモリにエラーの訂正が不可能な故障が発生した場合、SCは、ECCによりエラーの訂正を行なうことができず、キャッシュコヒーレンスを保つことができない。このため、従来のSMPサーバシステムには、SC内のTAG_CPメモリでエラーの訂正が不可能な故障が発生した場合、被疑箇所に対応するCPUを縮退し、運用を一度停止させて故障箇所を切り離す仕組みが実装されている。

[0015] なお、エラーの訂正が不可能な故障とは、キャッシュタグに含まれるECCによってもエラーを訂正することが不可能な故障であり、例えば、2ビット

ト以上の領域の故障である。以下、2ビット以上の領域の故障（多ビット故障）を、訂正不可能エラー又はUE（Uncorrectable Error）という。

以下、SC内のTAG_CPメモリでCEが発生した場合の、システムの縮退処理の動作を説明する。

[0016] 図12は、図10に示すものと同様の構成のSB200及び運用管理部600において、SC400内のTAG_CPメモリ420-2でUEが発生した場合の縮退範囲を示す図である。また、図13は、SC400内のTAG_CPメモリ420-2でUEが発生した場合の縮退処理を説明するためのフローチャートである。

図12及び図13に示すように、システムの運用中、SC400内のTAG_CPメモリ420-2にUEが発生し、SC400により検出された場合（ステップS111）、SC400から運用管理部600に対して、UEが発生したことが割り込みで通知される（ステップS112）。

[0017] 運用管理部600では、割り込み通知に基づいて、被疑箇所に対応したCPU300-2を示す情報とWAY番号とが故障情報として、運用管理部600の制御情報に記録される（ステップS113）。そして、運用管理部600により、SMPサーバシステムで実行中のOSが再起動される（ステップS114）。

OSの再起動後、運用管理部600により、制御情報の故障情報が読み込まれ（ステップS115）、故障情報に記録されているCPU300-2は、立ち上げ処理が行なわれず、他の正常なCPU300-1、300-3及び300-4についてのみ立ち上げ処理が行なわれる。つまり、運用管理部600により、被疑箇所に対応するCPU300-2及び被疑箇所に対応するTAG_CPメモリ420-2が縮退処理された状態で、OSが立ち上がる（ステップS116、図12中、「縮退範囲」参照）。その後、SMPサーバシステムでは、運用が再開される（ステップS117）。

[0018] このように、SC400内部のTAG_CPメモリ420でUEが発生した場合には、SMPサーバシステムの運用が停止し、被疑箇所を含む全コン

ポーネント（例えば１つのCPU300全体）を縮退後、運用を再開するという手法が採られる。

なお、他に、キャッシュメモリを搭載した複数のCPUを備えたマルチプロセッサシステムにおいて、メモリコントロール／コヒーレンシ制御装置が備えるタグメモリから索引されたタグ索引結果に訂正不可能障害が発生した場合であっても、動作を継続できるようにする技術が知られている。

[0019] 具体的には、メモリコントロール／コヒーレンシ制御装置が、タグメモリから索引されたタグ索引結果に訂正不可能障害を検出した場合、各CPUに対して、訂正不可能障害が検出されたタグ索引結果に関連する可能性がある全てのデータを主記憶装置に掃き出すように指示する。これにより、データのコヒーレンシを保障することができる。

なお、訂正不可能障害が検出されたタグ索引結果に関連する可能性がある全てのデータとは、キャッシュメモリに格納されているデータの内の、下位アドレスがタグ索引時に用いられた下位アドレスと一致する全てのデータをいう。

先行技術文献

特許文献

[0020] 特許文献1：特開2008-52550号公報

発明の概要

発明が解決しようとする課題

[0021] 従来、TAG__CPメモリにおける訂正不可能エラー（UE）の発生頻度が低かったため、図12及び図13に例示したように、UEが発生した場合には、被疑箇所に対応したCPU及び被疑箇所のSC内のTAG__CPメモリを縮退処理させる運用が行なわれていた。

しかしながら、UEが発生した場合の上述した手法では、運用が停止する時間が発生し、SMPサーバシステムの可用性が低下するという問題がある。

[0022] また、近年、L S I (Large Scale Integration) 内部の集積度が増すことによりCM容量が増加している。また、SMPサーバシステムに搭載されるCPU数の増加に伴い、SMPサーバシステム内の総CM容量が増加している。このようなSMPサーバシステム内のCM容量の増加によって、以前に比べてUEが発生する確率が高くなっている。

このように、UEの発生確率が高くなっている現状において、SMPサーバシステムの可用性が低下する場面（頻度）が増加するという問題もある。

[0023] さらに、上述したメモリコントロール／コヒーレンシ制御装置を用いた技術では、訂正不可能エラーが検出された場合でもデータのコヒーレンシを保障することができるが、以下の（i）及び（ii）に示す問題がある。

（i）メモリコントロール／コヒーレンシ制御装置内のタグ部で訂正不可能エラーが発生してタグ部の一部が縮退された場合、CPUは、タグ部の一部が縮退されたことを知らないため、CPUから、縮退されたタグ部の一部を再使用するようなリクエストが送信される可能性がある。このようなリクエストが送信された場合、メモリコントロール／コヒーレンシ制御装置は、リクエストに従ったタグ部の使用が不可能であるという応答、或いはキャッシュをタグ部に登録せずに使用するという応答をCPUに返すことになる。

[0024] このような場合、以下の（i-1）～（i-3）に示す状況において、それぞれシステムの性能低下が発生するという問題がある。

（i-1）CPUは、上述の応答を受けることを許容していない場合、つまり上述の応答に対する処理が定義されておらず対応できない場合には、動作不能に陥る可能性がある。

[0025] （i-2）また、CPUは、上述の応答を許容し動作を行なう場合でも、CPUによる処理によっては、縮退されたタグ部の一部を再使用するようなリクエストがメモリコントロール／コヒーレンシ制御装置に繰り返し出力されるおそれがある。このような状況では、CPU及びメモリコントロール／コヒーレンシ制御装置間で、上述したリクエスト及び応答が繰り返し行なわれるため、システムの性能低下を招くことになる。

[0026] (i-3) 或いは、メモリコントロール／コヒーレンシ制御装置は、タグ索引時に用いられた下位アドレスと同じ下位アドレスの別WAYのデータの掃き出し指示を、リクエストに対する応答を返す前に、リクエスト要求元のCPUに対して行なうことが考えられる。この場合、メモリコントロール／コヒーレンシ制御装置は、CPUによる掃き出し完了後、元のリクエストに対する応答を要求元のCPUに返すような動作を行なう。これにより、データのコヒーレンシを保つことはできるが、上述の如く、CPUにより上述の応答に対する処理が行なわれ、システムの性能低下を招くことになる。

[0027] (ii) また、メモリコントロール／コヒーレンシ制御装置は、エントリの縮退状態を示すエントリ使用不可フラグをタグメモリに備える。しかし、タグメモリのアドレスライン系で故障が発生した場合には、エントリ使用不可フラグ自体を正しく読み出せない可能性がある。

すなわち、タグメモリのアドレスライン系で故障が発生した場合、タグメモリのセルに対するアクセスが正常に行なわれず、エントリ使用不可フラグ自体を正しく読み出せないことがある。従って、実際にはエントリ使用不可フラグに縮退を示す情報が設定されている場合であっても、システムからは縮退したように見えず、リトライが行なわれる度にUEの発生が検出されて、処理不能に陥る可能性がある。

[0028] なお、エントリ使用不可フラグを、タグメモリではなく、例えばエントリ毎にラッチで備えることも考えられるが、物量的に困難である。

上述の点に鑑み、本件の目的の1つは、情報処理装置の可用性の向上を実現することである。

なお、前記目的に限らず、後述する発明を実施するための形態に示す各構成により導かれる作用効果であって、従来技術によっては得られない作用効果を奏することも本発明の他の目的の1つとして位置付けることができる。

課題を解決するための手段

[0029] 本件の情報処理装置は、キャッシュメモリ及び第1タグメモリを有する演

算処理装置と、前記演算処理装置と他の処理装置との通信制御を行なうシステムコントローラとを有する情報処理装置であって、前記システムコントローラは、前記演算処理装置から受信したリクエストを保持し、当該リクエストが要求先において処理されなかった場合に当該リクエストを再発行するコマンド制御部と、前記第１タグメモリに記憶されるデータの複製データを保持する第２タグメモリと、前記第２タグメモリから読み出されたデータが訂正不可能なエラーであるＵＥを起こした場合に、前記ＵＥが発生した第２タグメモリのウェイ情報を前記演算処理装置に通知するリクエスト制御部とを有し、前記演算処理装置は、前記リクエスト制御部から前記ＵＥが発生した通知を受けると、当該ＵＥが発生した前記第２タグメモリのウェイに対応する前記第１タグメモリのウェイを縮退させて、前記第１タグメモリのウェイの縮退処理が完了したことを前記リクエスト制御部へ通知し、前記リクエスト制御部は、前記ＵＥが発生した場合に、前記第２タグメモリのＵＥが発生したウェイを縮退させるとともに、前記演算処理装置から、前記第１タグメモリの縮退処理が完了した旨の通知を受信した後、前記コマンド制御部に当該ＵＥに係るリクエストを再発行させる指示を行なうものである。

[0030] また、本件のキャッシュ制御方法は、キャッシュメモリ及び第１タグメモリを有する演算処理装置と、前記演算処理装置と他の処理装置との通信制御を行なうシステムコントローラとを有する情報処理装置のキャッシュ制御方法であって、前記システムコントローラにおいて、前記第１タグメモリに記憶されるデータの複製データを保持する第２タグメモリから読み出されたデータが訂正不可能なエラーであるＵＥを起こした場合に、前記ＵＥが発生した前記第２タグメモリのウェイ情報を前記演算処理装置に通知し、前記演算処理装置において、前記ＵＥが発生した通知を受けると、当該ＵＥが発生した前記第２タグメモリのウェイに対応する前記第１タグメモリのウェイを縮退させて、前記第１タグメモリのウェイの縮退処理が完了したことを前記システムコントローラへ通知して、前記システムコントローラにおいて、前記演算処理装置から、前記第２タグメモリのＵＥが発生したウェイを縮退する

とともに、前記第 1 タグメモリの縮退処理が完了した旨の通知を受信した後、当該 UE に係るリクエストを再発行するものである。

発明の効果

[0031] 開示の技術によれば、情報処理装置の可用性の向上を実現することができる。

図面の簡単な説明

[0032] [図1] 第 1 実施形態の一例としての情報処理システムの構成を示す図である。

[図2] 第 1 実施形態の一例としてのシステムコントローラの構成を示す図である。

[図3] 第 1 実施形態の一例としてのシステムコントローラ内のキャッシュタグメモリで UE が発生した場合の縮退範囲を示す図である。

[図4] 第 1 実施形態の一例としてのシステムコントローラ内のキャッシュタグメモリで UE が発生した場合の縮退処理を説明するためのフローチャートである。

[図5] 第 2 実施形態の一例としての情報処理システムの構成を示す図である。

[図6] 第 2 実施形態の一例としてのメモリのアドレスマップを示す図である。

[図7] 第 2 実施形態の一例としての TAG__CP メモリの構成を示す図である。

[図8] 第 2 実施形態の一例としてのシステムコントローラの構成を示す図である。

[図9] 第 2 実施形態の一例としてのシステムコントローラ内のキャッシュタグメモリで CE 又は UE が発生した場合の縮退処理を説明するためのフローチャートである。

[図10] システムコントローラ内の TAG__CP メモリで CE が発生した場合の縮退範囲を示す図である。

[図11] システムコントローラ内の TAG__CP メモリで CE が発生した場合の縮退処理を説明するためのフローチャートである。

[図12] システムコントローラ内の TAG__CP メモリで UE が発生した場合

の縮退範囲を示す図である。

[図13] システムコントローラ内のTAG__CPメモリでUEが発生した場合の縮退処理を説明するためのフローチャートである。

発明を実施するための形態

[0033] 以下、図面を参照して本発明の実施の形態を説明する。

〔1〕 第1実施形態

〔1-1〕 第1実施形態の構成

図1は、第1実施形態の一例としての情報処理システム1の構成を示す図である。

図1に示すように、情報処理システム（情報処理装置）1は、SB2と、運用管理部6とを備える。

[0034] 情報処理システム1は、例えば、SMPサーバシステムである。

また、第1実施形態における情報処理システム1は、SC4内のTAG__CPメモリ42においてCEが発生した場合に、図11に示す上述した手法により、運用を継続することができる。なお、第1実施形態においては、CEが発生した場合の情報処理システム1の動作については、その詳細な説明を省略する。

[0035] 第1実施形態における情報処理システム1は、SC4内のTAG__CPメモリ42においてUEが発生した場合に、後述の如く、UEが発生したTAG__CPメモリ42のWAY及び当該WAYに対応するCPU3のTAGメモリ32のWAYを動的に縮退させることで、運用を継続させることができる。

SB2は、少なくとも1つ（第1実施形態においては4つ）のCPU3-1～3-4と、SC4と、RAM等のメモリ5とを備える。なお、以下の説明においてCPU3-1～3-4を区別しない場合には、単にCPU3という。

[0036] CPU3-1～3-4は、それぞれSC4に接続され、情報処理システム1における種々の制御や演算を行なう演算処理装置であり、例えば、記憶部

(図示省略)に格納されたプログラムをメモリ5に展開して実行することにより、種々の機能を実現する。

CPU3-1~3-4は、それぞれCM31-1~31-4及びTAGメモリ(第1タグメモリ)32-1~32-4を備える。以下の説明においてCM31-1~31-4を区別しない場合には、単にCM31という。また、TAGメモリ32-1~32-4を区別しない場合には、単にTAGメモリ32という。

[0037] なお、CM31-1~31-4及びTAGメモリ32-1~32-4の符号におけるハイフン“-”の右側の数字は、CM31及びTAGメモリ32が、それぞれ対応する数字のCPU3-1~3-4に備えられることを示す。

CM31は、CPU3とメモリ5との間で転送されるデータを格納する。なお、第1実施形態においては、CM31が、nWAYセットアソシアティブ方式を採用する場合について例示する。

[0038] TAGメモリ32は、CM31で保持されるデータの参照情報であるキャッシュタグデータを格納する。

また、第1実施形態におけるCPU3は、SC4からCEが発生した通知(CE通知リクエスト)又はUEが発生した通知(UE通知リクエスト)を受信すると、当該CE又はUEが発生したTAG_CPメモリ42のWAYに対応するTAGメモリ32のWAYを動的に縮退させる。そして、CPU3は、WAYの縮退後、TAGメモリ32のWAYの縮退処理が完了したことをSC4へ通知する。

[0039] なお、以下、CE又はUEが発生した(検出された)TAG_CPメモリ42のキャッシュタグデータを、被疑箇所ともいう。

SC(システムコントローラ)4は、CPU3及びメモリ5間のアクセスを制御するとともに、CPU3と他のCPU3又はSB2の外部の処理装置との通信制御を行なうLSIである。なお、第1実施形態においては、CPU3及びSC4は、キャッシュコヒーレンシのアルゴリズムとしてスヌープ

方式を採用する場合について例示する。

- [0040] また、第1実施形態におけるSC4は、TAGメモリ32-1～32-4に対応したTAG_CPメモリ（第2タグメモリ）42-1～42-4を備える。

TAG_CPメモリ42-1～42-4は、対応するTAGメモリ32-1～32-4に記憶されるデータの複製データを保持する。以下の説明においてTAG_CPメモリ42-1～42-4を区別しない場合には、単にTAG_CPメモリ42という。

- [0041] なお、CM31、TAGメモリ32及びTAG_CPメモリ42としては、いずれも、例えばSRAM等のRAMが挙げられる。

SC4は、CPU3のキャッシュタグデータのコピーをTAG_CPメモリ42に格納することで、各CPU3からのメモリ5へのアクセス要求等のリクエストに対し、TAG_CPメモリ42を参照してリクエストに応じた所定の処理を行ない、要求元のCPU3に応答を返す。これにより、スヌープ方式による高速なキャッシュアクセスを実現し、CPU3のCM31の同期処理の高速化を実現する。

- [0042] また、SC4は、情報処理システム1の運用中に、TAG_CPメモリ42において訂正不可能エラー（UE）が発生すると、CPU3からのリクエストのうちのUEが検出されたリクエスト（以下、UE検出リクエストという）に係る処理を保留する。

そして、SC4は、UEが検出されたTAG_CPメモリ42に対応するCPU3に対して、エラー情報を含むUE通知リクエストを出力する。なお、エラー情報には、被疑箇所に対応するWAY情報（例えばWAYの番号等）が含まれる。UE通知リクエストを受信したCPU3は、エラー情報に基づいて、TAGメモリ32のWAYを動的に縮退させ、WAYの縮退後、縮退処理が完了した旨の縮退処理完了通知をSC4へ通知する。

- [0043] また、SC4は、TAG_CPメモリ42のUEが発生したWAYを縮退させる。

さらに、S C 4 は、C P U 3 から縮退処理完了通知を受信するとともに、T A G _ C P メモリ 4 2 のW A Y の縮退処理の完了後、U E 発生リクエストに係る処理を再開する。

また、S C 4 は、運用管理部 6 に対してC E 又はU E に関するエラー情報を割り込み通知する。

[0044] S C 4 の詳細な構成については、後述する。

メモリ 5 は、種々のデータやプログラムを一時的に格納する記憶領域であって、C P U 3 がプログラムを実行する際に、データやプログラムを一時的に格納・展開して用いる。なお、第 1 実施形態におけるメモリ 5 は、C P U 3 - 1 ~ 3 - 4 のいずれからもアクセス可能であり、C P U 3 - 1 ~ 3 - 4 により共有して用いられる。

[0045] 運用管理部 6 は、情報処理システム 1 を制御するファームウェアが搭載されており、S C 4 からのC E 又はU E に関するエラー情報の割り込み通知に基づいて、C P U 3 及びS C 4 において縮退されたW A Y に関する情報を故障情報として格納する。なお、縮退されたW A Y に関する情報には、被疑箇所に対応したC P U 3 の情報（例えばC P U の番号等）及びW A Y 情報が含まれる。

[0046] また、運用管理部 6 は、例えば情報処理システム 1 で実行中のO S の再起動等によってC P U 3 及びS C 4 内の縮退状態がリセットされた場合に、故障情報に基づいて、被疑箇所に対応するW A Y を再度縮退させる。なお、運用管理部 6 としては、サービスプロセッサが挙げられる。

なお、情報処理システム 1 は、例えばH D D (Hard Disk Drive) やS S D (Solid State Drive) 等の記憶部（図示省略）を備えることができる。記憶部は、S C 4 を介して各C P U 3 からアクセス可能に構成することができる。

[0047] [1 - 2] 第 1 実施形態のシステムコントローラの構成

図 2 は、第 1 実施形態の一例としてのS C 4 の構成を示す図である。

図 2 に示すように、S C 4 は、複数（第 1 実施形態においては 4 つ）の T

ＡＧ＿ＣＰメモリ制御部４１－１～４１－４と、コマンド制御部４３と、リクエスト制御部４４と、アドレスロックレジスタ部４５と、レジスタ部４６とを備える。

[0048] コマンド制御部４３は、ＣＰＵ３から受信したリクエスト（コマンド）を保持するとともに、リクエストをＴＡＧ＿ＣＰメモリ制御部４１－１～４１－４及びアドレスロックレジスタ部４５へ転送するための制御を行なう。

また、コマンド制御部４３は、ＣＰＵ３から受信したリクエストについてＳＣ４内で処理が完了するまでは、当該リクエストを保持する。つまり、コマンド制御部４３は、リクエストの要求先であるＴＡＧ＿ＣＰメモリ制御部４１において、リクエストが処理中の場合や、リクエストが処理されなかった場合には、当該リクエストを保持する。

[0049] さらに、第１実施形態においては、コマンド制御部４３は、ＣＰＵ３から受信したリクエストが要求先のＴＡＧ＿ＣＰメモリ制御部４１において処理されなかった場合、保持しているリクエストを再発行する。

なお、コマンド制御部４３は、ＴＡＧ＿ＣＰメモリ制御部４１における処理が完了すると、キュー４３ａから該当するリクエストを削除する。

[0050] ＴＡＧ＿ＣＰメモリ制御部４１－１～４１－４は、ＴＡＧ＿ＣＰメモリ４２－１～４２－４にそれぞれ対応して備えられ、コマンド制御部４３から転送されるリクエストに係る処理を実行する。なお、以下の説明においてＴＡＧ＿ＣＰメモリ制御部４１－１～４１－４を区別しない場合には、単にＴＡＧ＿ＣＰメモリ制御部４１という。

具体的には、ＴＡＧ＿ＣＰメモリ制御部４１は、コマンド制御部４３から転送されるリクエストに含まれるメモリ５の実アドレス（ＰＡ；Physical Address）から、キャッシュラインを特定するインデックス及びＷＡＹを特定するエントリアドレス（以下、登録アドレスという）を抽出する。そして、ＴＡＧ＿ＣＰメモリ制御部４１は、対応するＴＡＧ＿ＣＰメモリ４２－１～４２－４内から、抽出したインデックス及び登録アドレスに対応するキャッシュタグデータを検索する。

[0051] なお、検索によりリクエストに係るキャッシュタグデータがTAG_CPメモリ42内でヒット又はミスヒットした場合、その後の処理の内容は、リクエストの内容とキャッシュタグデータのステータスとに応じて決定される。これらの処理の内容の決定は、既知の種々の手法に実現することができるため、ここではその詳細な説明は省略する。

また、TAG_CPメモリ制御部41は、TAG_CPメモリ42でCE又はUEが検出された場合、リクエスト制御部44に対して、被疑箇所の通知（TAG_CPエラー通知）を行なう。

[0052] リクエスト制御部44は、TAG_CPメモリ42から読み出されたデータがCE又はUEを起こした場合に、CE又はUEが発生したTAG_CPメモリ42のWAY情報をCPU3に通知する。

具体的には、リクエスト制御部44は、TAG_CPメモリ42でCE又はUEが検出され、TAG_CPメモリ42からTAG_CPエラー通知を受けると、被疑箇所のインデックス及びWAY情報を含むCE通知リクエスト又はUE通知リクエストをCPU3に発行する。CE又はUE通知リクエストが通知されたCPU3は、当該リクエストに基づいて、CE又はUEが発生したTAG_CPメモリ42のWAYに対応するTAGメモリ32のWAYについて縮退処理を行ない、縮退処理完了通知をリクエスト制御部44へ通知する。

[0053] また、リクエスト制御部44は、TAG_CPメモリ42からUEが検出されたTAG_CPエラー通知を受けると、コマンド制御部43に対して、UE検出リクエストの保留指示を通知する。コマンド制御部43は、リクエスト制御部44から保留指示が通知されると、UE検出リクエストを保留状態として保持する。

さらに、リクエスト制御部44は、CE又はUEが発生した場合に、CE又はUEが発生したTAG_CPメモリ42のWAYを縮退させる。また、リクエスト制御部44は、UEが発生した場合に、CPU3からTAGメモリ32の縮退処理完了通知を受信した後、コマンド制御部43にUE検出リ

クエストを再発行させる指示を行なう。

[0054] 具体的には、リクエスト制御部 44 は、TAG__CPメモリ 42 から TAG__CP エラー通知を受けると、レジスタ部 46 に対して、CE 又は UE が検出された TAG__CPメモリ 42 のWAY の縮退設定を行なう。

また、リクエスト制御部 44 は、UE が発生した場合に、CPU 3 から縮退処理完了通知を受信すると、コマンド制御部 43 に対して、UE 検出リクエストの処理再開（再発行）の指示を通知する。コマンド制御部 43 は、リクエスト制御部 44 から UE 検出リクエストの処理再開の指示が通知されると、UE 検出リクエストの保留状態を解除して当該リクエストの処理を再開（当該リクエストを再発行）する。

[0055] さらに、リクエスト制御部 44 は、CE 又は UE が発生した TAG__CPメモリ 42 のWAY を縮退させた後、運用管理部 6 に対して、CE 又は UE に関するエラー情報を割り込み通知する。割り込み通知を受けた運用管理部 6 は、リクエスト制御部 44 から通知されたエラー情報に基づいて、運用管理部 6 が管理する制御情報に対して、縮退させたWAY に関する情報を故障情報として保持する。なお、故障情報には、被疑箇所に対応した CPU 3 の情報及びWAY 情報が含まれる。

[0056] また、運用管理部 6 は、情報処理システム 1 で実行中の OS が再起動した場合に、保持している故障情報に基づいて、TAGメモリ 32 及び TAG__CPメモリ 42 の当該WAY を縮退させる。

なお、リクエスト制御部 44 は、運用管理部 6 に対して割り込み通知する CE 又は UE に関するエラー情報に、縮退させたWAY に関する情報を含めても良いし、含めなくても良い。リクエスト制御部 44 がエラー情報に縮退させたWAY に関する情報を含めない場合、運用管理部 6 は、リクエスト制御部 44 からのエラー情報の割り込み通知を受信すると、レジスタ部 46 から被疑箇所に対応した CPU 3 の情報及びWAY 情報を取得し、保持するようにしても良い。

[0057] レジスタ部 46 は、使用可能な TAG__CPメモリ 42-1 ~ 42-4 の

WAYを示す構成情報を保持する。

構成情報には、TAG_CPメモリ42-1～42-4のWAY毎に、有効又は無効の状態が含まれ、リクエスト制御部44からの設定変更要求に応じて、レジスタ部46により有効又は無効の状態が設定される。

[0058] 有効又は無効の状態は、例えば有効を示す“0”及び無効を示す“1”のビットを用いた縮退フラグによって表すことができる。

換言すれば、レジスタ部46は、TAG_CPメモリ42-1～42-4のWAYを縮退することを示す縮退フラグを保持するといえる。

すなわち、リクエスト制御部44は、レジスタ部46に対して、CE又はUEが発生したTAG_CPメモリ42のWAYに係る縮退フラグを設定することにより、当該WAYを縮退させる。

[0059] なお、図示を省略しているが、CPU3-1～3-4も、使用可能なTAGメモリ32-1～32-4のWAYを示す構成情報を保持するレジスタ部を備える。

従って、CPU3によるTAGメモリ32のWAYの縮退処理も、リクエスト制御部44による縮退処理と同様に、対応するTAGメモリ32のWAYに係る縮退フラグがCPU3によってレジスタ部に設定されることにより行なわれる。

[0060] アドレスロックレジスタ部45は、ロックレジスタ45aを備え、SC4内で処理中のリクエスト内のアドレス情報をロックレジスタ45aに保持する。

具体的には、アドレスロックレジスタ部45は、コマンド制御部43から転送されるリクエスト内の実アドレスから全てのアドレス（フルアドレス）とインデックスとを抽出し、抽出したフルアドレス、つまりSC4内で処理中のリクエストに係るフルアドレスをロックレジスタ45aに保持する。

[0061] また、アドレスロックレジスタ部45は、後発リクエスト内の実アドレスにおけるフルアドレスが、ロックレジスタ45aに保持されたリクエスト内の実アドレスにおけるフルアドレスと一致した場合、当該後発リクエストに

係るフルアドレスがビジー状態である旨（フルアドレスビジー）をコマンド制御部43に通知する。コマンド制御部43では、フルアドレスビジーが通知されると、後発リクエストを再度TAG__CPメモリ制御部41及びアドレスロックレジスタ部45へそれぞれ転送し、後発リクエストを再発行（リトライ）する。

[0062] このように、アドレスロックレジスタ部45は、コマンド制御部43から転送される後発リクエストに係る処理をキャンセル及びリトライさせて、処理中のリクエストと競合しないようにガード（ロック）するガード（ロック）機能を備える。

また、アドレスロックレジスタ部45は、SC4内で処理中のリクエストに係る処理が完了した場合、ロックレジスタ45aから当該リクエストに係るアドレス情報を削除し、当該リクエストのロックを解除する。

[0063] さらに、第1実施形態におけるアドレスロックレジスタ部45は、上述した処理に加えて、UEが発生した際には、CPU3及びSC4における被疑箇所の縮退処理が完了するまで、UE検出リクエストにおける被疑箇所のフルアドレスをロックレジスタ45aに保持する。そして、後発リクエスト内の実アドレスにおけるフルアドレスが、ロックレジスタ45aに保持された被疑箇所のフルアドレスと一致した場合、アドレスロックレジスタ部45は、後発リクエストのフルアドレスビジーをコマンド制御部43に通知する。

[0064] すなわち、アドレスロックレジスタ部45は、CPU3からTAGメモリ32の縮退処理完了通知を受信するとともに、リクエスト制御部44がTAG__CPメモリ42のUEが発生したWAYを縮退させるまでは、他のリクエストによる当該UEが発生したTAG__CPメモリ42へのアクセスを抑制する。

つまり、第1実施形態におけるアドレスロックレジスタ部45は、UE検出リクエストによりUEが検出された場合、当該UE検出リクエスト内のアドレス情報をロックレジスタ45aに保持してロックし、CPU3からの縮退処理完了通知を受信した場合に、当該リクエストのロックを解除する。

[0065] これにより、他のリクエストによってUE検出リクエストにおけるフルアドレスが参照されないように、被疑箇所をガードすることができる。

なお、アドレスロックレジスタ部45は、UE検出リクエストのアドレスがロックレジスタ45aに保持されている場合、同一インデックスであって且つフルアドレスが不一致のときには、被疑箇所と同一インデックスの領域についてもガードすることができる。これにより、UE発生後の縮退処理中に、後発リクエストによるTAG__CPメモリ42の被疑箇所と同一インデックスの領域への参照を抑止することができる。このため、アドレスロックレジスタ部45は、アドレス比較のビット幅を可変にできるように構成されることが好ましい。

[0066] このように、第1実施形態におけるアドレスロックレジスタ部45は、ガード機能により、処理中のリクエストと競合しないように後発リクエストに係る処理をガードするとともに、UEが発生した場合に被疑箇所をガードすることができる。

〔1-3〕第1実施形態の情報処理システムのUE発生時の動作

次に、上述の如く構成された情報処理システム1におけるSC4のTAG__CPメモリ42でUEが発生した場合の縮退処理について説明する。

[0067] 図3は、第1実施形態の一例としてのSC4内のTAG__CPメモリ42-2でUEが発生した場合の縮退範囲を示す図であり、図4は、第1実施形態の一例としてのSC4内のTAG__CPメモリ42-2でUEが発生した場合の縮退処理を説明するためのフローチャートである。

はじめに、図3及び図4に示すように、システムの運用中、TAG__CPメモリ42-2でUEが検出された場合（ステップS1）、SC4内のTAG__CPメモリ制御部41により、リクエスト制御部44に対してTAG__CPエラー通知が行なわれる。

[0068] TAG__CPエラー通知が入力されると、リクエスト制御部44により、コマンド制御部43に対してUE検出リクエストの保留指示が通知される（ステップS2）。コマンド制御部43では、保留指示の通知を受けて、UE

検出リクエストが保留状態として保持される。

次いで、リクエスト制御部 44 により、UE が検出された TAG__CP メモリ 42-2 に対応する CPU 3-2 に対して、UE が検出された被疑箇所のインデックス及びWAY 情報を含む UE 通知リクエストが通知される（ステップ S3）。

[0069] CPU 3-2 では、受信した UE 通知リクエストに基づいて、被疑箇所に対応する CM 31-2 のWAY の全エントリがメモリ 5 に退避され、当該WAY の縮退処理が行なわれる（ステップ S4、図 3 中、TAG メモリ 32-2 の「縮退範囲」参照）。具体的には、CPU 3 により、CPU 3 内のレジスタ部に対して被疑箇所に対応したWAY を無効にする設定変更要求が出力され、レジスタ部では、構成情報において被疑箇所に対応したWAY に縮退フラグが設定され、当該WAY が無効にされる。その後、CPU 3-2 からリクエスト処理部 44 に対して、縮退処理完了通知が送信される（ステップ S5）。

[0070] また、リクエスト制御部 44 では、SC 4 内で被疑箇所に対応したWAY の縮退処理が行なわれる（ステップ S6、図 3 中、TAG__CP メモリ 42-2 の「縮退範囲」参照）。具体的には、リクエスト処理部 44 により、レジスタ部 46 に対して被疑箇所に対応したWAY を無効にする設定変更要求が出力され、レジスタ部 46 では、構成情報において被疑箇所に対応したWAY に縮退フラグが設定され、当該WAY が無効にされる。

[0071] なお、上述した CPU 3-2 及び TAG__CP メモリ 42-2 の縮退処理が完了するまで、アドレスロックレジスタ部 45 により、他のリクエストによって被疑箇所が参照されないように、被疑箇所がガードされる。これにより、UE の多重発生を防ぐことができる。なお、縮退処理完了通知が入力されると、アドレスロックレジスタ部 45 により、ロックレジスタ 45a に保持された UE 検出リクエストに係るアドレス情報が削除され、当該リクエストのロックが解除される。

[0072] 上述した CPU 3-2 及び TAG__CP メモリ 42-2 の縮退処理の完了

後、リクエスト処理部 4 4 により、コマンド制御部 4 3 に対して、保留状態であった UE 検出リクエストの処理再開の指示が通知される（ステップ S 7）。コマンド制御部 4 3 では、処理再開指示の通知を受けて、UE 検出リクエストに係る処理が再開される。

また、リクエスト制御部 4 4 により、運用管理部 6 に対して、UE に関するエラー情報が割り込み通知される（ステップ S 8）。

[0073] 割り込み通知が入力されると、運用管理部 6 のファームウェアにより、運用管理部 6 が管理する制御情報に対して、縮退させた WAY に関する情報が、故障情報として記録される（ステップ S 9）。

そして、情報処理システム 1 による運用が継続される（ステップ S 10）。

上述の処理により、CPU 3 及び SC 4 において、被疑箇所に対応した WAY の縮退処理が動的に行なわれるため、TAG__CP メモリ 4 2 において UE が発生した場合でも、運用の停止を回避することができる。

[0074] 上述のように、第 1 実施形態の一例としての情報処理システム 1 によれば、SC 4 のリクエスト制御部 4 4 により、TAG__CP メモリ 4 2 で UE が発生した場合に、UE 通知リクエストが CPU 3 に通知される。そして、CPU 3 により、受信した UE 通知リクエストに基づいて UE が発生した WAY に対応した WAY の縮退処理が行なわれる。また、リクエスト制御部 4 4 により、TAG__CP メモリ 4 2 の UE が発生した WAY の縮退処理が行なわれる。

[0075] これにより、情報処理システム 1 は、上述した図 10 及び図 11 等に例示するような TAG__CP メモリ 4 20 に CE が発生した場合と同様に、UE が検出された WAY に対応する、CPU 3 内の TAG メモリ 3 2 の WAY 及び SC 4 内の TAG__CP メモリ 4 2 の WAY を動的に縮退することができる。

従って、システムの継続運用が可能になり、情報処理システム 1 の可用性の向上を実現することができる。

[0076] また、UE検出に伴う縮退処理をWAY単位で行なうことができるため、上述した図12及び図13等に例示するような、UEが発生した場合にCPU300単位及びSC400のTAG__CPメモリ420単位で縮退を行なう従来の手法と比較して、縮退範囲をより限定することができる。

さらに、TAG__CPメモリ42内の被疑箇所に対応するWAYの縮退処理に加えて、CPU3のTAGメモリ32内の被疑箇所に対応するWAYの縮退処理も行なわれる。

[0077] 従って、CPU3側の縮退処理完了後は、当該CPU3から被疑箇所へのリクエストが発行されないため、CPU3の処理負担を抑えることができ、UEが発生した際の情報処理システム1の大幅な性能の低下を抑止することができる。

さらに、TAG__CPメモリ42-1~42-4は、CPU3-1~3-4内のTAGメモリ32-1~32-4と一対一に対応する。従って、TAG__CPメモリ42の一部(WAY)が縮退しても、CPU3は被疑箇所に対応するCPU3以外の他のCPU3へのアクセスを行なうことができるため、CPU3からのリクエスト発行のリトライによる性能低下を防ぐことができる。

[0078] また、第1実施形態の一例としての情報処理システム1によれば、SC4のリクエスト制御部44により、CPU3からTAGメモリ32の縮退処理完了通知を受信した後に、コマンド制御部43にUE検出リクエストを再発行させる指示が行なわれる。

これにより、縮退処理の完了後にコマンド制御部43によってUE検出リクエストが再発行されるため、UE検出リクエストを発行したCPU3は、要求先において処理されなかったUE検出リクエストのリトライを行わずに済む。

[0079] 従って、UE検出リクエストに係る縮退処理に伴うCPU3の処理負担を抑えることができ、UEが発生した際の情報処理システム1の性能の低下を抑止することができる。

さらに、第1実施形態の一例としての情報処理システム1によれば、アドレスロックレジスタ部45により、CPU3及びSC4において被疑箇所に対応するWAYが縮退されるまでは、他のリクエストによる被疑箇所へのアクセスが抑止される。

[0080] これにより、UE検出リクエスト及び被疑箇所への後発リクエストのような、要求元において処理されなかったリクエストが、コマンド制御部43によって再発行される。

従って、他のリクエストによってUEが検出されたキャッシュタグデータが参照されないように、被疑箇所をガードすることができ、キャッシュコヒーレンシを保つことができる。つまり、上述した図13に例示するようなUEが発生したTAG__CPメモリ42に対応するCPU3自体の縮退のためのOS再起動が不要となるため、システムの継続運用が可能になり、情報処理システム1の可用性の向上を実現することができる。

[0081] また、アドレスロックレジスタ部45によるUE発生被疑箇所のガードは、後発リクエストが、処理中のリクエストと競合しないようにガードするガード機能を利用して実現される。

さらに、第1実施形態の一例としての情報処理システム1によれば、アドレスロックレジスタ部45により、UE検出リクエストのアドレスがロックレジスタ45aに保持されている場合、同一インデックスであって且つフルアドレスが不一致のときには、被疑箇所と同一インデックスの領域についてもガードされる。なお、この場合、アドレスロックレジスタ部45は、アドレス比較のビット幅を可変にできるように構成される。

[0082] これにより、UE発生後の縮退処理中に、後発リクエストによるTAG__CPメモリ42の被疑箇所と同一インデックスの領域への参照を抑止することができる。

従って、上述したアドレスロックレジスタ部45のガード機能による後発リクエストのリトライ条件に、UE検出リクエスト内の実アドレスにおけるフルアドレス又はインデックスと一致した場合を加えることによって、被疑

箇所のガードを実現することができ、新たな回路を設けずに済むため、情報処理システム 1 の製造及び保守コストを減少させることができる。

[0083] さらに、第 1 実施形態の一例としての情報処理システム 1 によれば、レジスタ部 46 により、TAG_CP メモリ 42 の WAY を縮退することを示す縮退フラグが保持される。

これにより、TAG_CP メモリ 42 内に縮退フラグを備える構成と比較して、例えばアドレスライン系で故障が発生した場合でも、リクエスト処理部 44 は、確実に縮退フラグの設定を行なうことができ、UE の縮退処理を確実に行なうことができる。

[0084] [2] 第 2 実施形態

[2-1] 第 2 実施形態の構成

次に、図 5 ～図 9 を参照しながら、第 2 実施形態としての情報処理システム（情報処理装置）1' の構成について説明する。なお、図 5 において既述の符号と同一の符号は同一の部分若しくは略同一の部分を示しているため、重複した説明は省略する。

[0085] 図 5 は、第 2 実施形態の一例としての情報処理システム 1' の構成を示す図である。

図 5 に示すように、第 2 実施形態における情報処理システム 1' は、複数（図 5 に示す例では 16 個）の SB 2 と、運用管理部 6 と、複数（図 5 に示す例では 4 つ）のクロスバー（Cross Bar；以下、XB という）9 とを備える。

そして、情報処理システム 1' は、これら複数の SB 2 の全て若しくは一部を使用した SMP サーバシステムとして機能する。

[0086] XB 9 は、複数の SB 2 間でのデータ転送機能を有する LSI であり、クロスバーユニット（Cross Bar Unit；図 5 中、XBU と表記）8 に実装される。

運用管理部 6 は、第 1 実施形態と同様の構成であり、情報処理システム 1' 全体、つまり各 SB 2 内の CPU 3 及び SC 4 を制御するファームウェア

を搭載する。また、運用管理部 6 は、運用管理部 6 が管理する制御情報において、各 S B 2 における縮退された C P U 3 の情報及び W A Y 情報を故障情報として保持する。

[0087] また、情報処理システム 1' は、第 1 実施形態と同様に、各 S B 2 内の S C 4 からアクセス可能な記憶部（図示省略）を備えても良い。

第 2 実施形態における S C 4 は、第 1 実施形態と同様に、自 S B 2 内の C P U 3 から発行されたリクエストに係る処理を実行することができる。また、第 2 実施形態における S C 4 は、複数の S B 2 間における通信のインターフェース機能を備え、他の S B 2 内の C P U 3 により、X B 9 を経由して自 S B 2 内へのリクエストが発行された場合、当該リクエストに係る処理を実行することができる。

[0088] なお、第 2 実施形態においては、C M 3 1 のキャッシュラインが 2 5 6 バイト (Byte) であり W A Y が 1 2 である場合について例示する。

以下、第 2 実施形態における T A G _ C P メモリ 4 2 の構成について説明する。

図 6 は、第 2 実施形態の一例としてのメモリ 5 のアドレスマップを示す図であり、図 7 は、第 2 実施形態の一例としての T A G _ C P メモリ 4 2 の構成を示す図である。なお、図 7 に示す例においては、各 W A Y のテーブルにおける 1 行が、1 つのキャッシュタグデータに対応する。

[0089] 図 6 に示すように、第 2 実施形態におけるメモリ 5 は、キャッシュライン (2 5 6 バイト) 単位 (ブロック) で S C 4 により管理される。

第 2 実施形態における S C 4 内の T A G _ C P メモリ 4 2 は、キャッシュタグデータを、図 7 に示す形態で管理する。

すなわち、T A G _ C P メモリ 4 2 では、メモリ 5 の実アドレス (P A) のうちの 4 1 : 1 9 ビットがキャッシュタグデータの登録アドレスとして格納される。

[0090] また、T A G _ C P メモリ 4 2 では、キャッシュのステータス (S T S ; Status) の 7 : 0 ビットが、キャッシュタグデータに格納される。

さらに、TAG_CPメモリ42では、キャッシュタグデータのエラー訂正符号（ECC）が、7ビットのデータとしてキャッシュタグデータに付加される。

そして、上述した登録アドレス及びステータスを格納するTAG_CPメモリ42のアドレスには、メモリ5の実アドレスの一部であるインデックスが使用される。

[0091] なお、第2実施形態においては、キャッシュラインを示すインデックスが11ビットであり、キャッシュライン数が2048である場合について例示する。すなわち、図7に示す例においては、メモリ5の実アドレスのうちの18：8ビットがインデックスアドレスに割り当てられる。

従って、図6に示すメモリ5のアドレスマップにおいて、同じインデックスになるブロック（例えばA0及びB0）は、図7に示すように同じインデックスアドレスに割り当てられる。また、これら同じインデックスになるブロックは、順にWAY0、WAY1…WAY11に格納される。

[0092] なお、キャッシュタグデータにおけるステータスは、例えばMOSIプロトコルにおける4状態で表される。MOSIプロトコルとは、M（Modified；変更）、O（Owned；所有）、S（Shared；共有）、及びI（Invalid；無効）の4つのキャッシュステータスを採用するプロトコルである。

また、SC4は、TAG_CPメモリ42においてCE又はUEが検出された場合、第1実施形態と同様の動作を行なうが、図8を参照しながら、より具体的な構成例について説明する。なお、図8において既述の符号と同一の符号は同一の部分若しくは略同一の部分を示しているため、重複した説明は省略する。

[0093] 図8は、第2実施形態の一例としてのSC4の構成を示す図である。

図8に例示するSC4は、第1実施形態におけるSC4に加えて、パイプ部47と、第1I/F（Interface）部48と、第2I/F部49とをさらに備える。

また、図8に例示するSC4は、第1実施形態におけるSC4に加えて、

TAG__CPメモリ制御部41が比較器41aを備え、リクエスト制御部44がアドレス競合検査部45bを備え、レジスタ部46がレジスタ設定変更部46a及び構成制御レジスタ46bを備える。

[0094] さらに、図8に例示するSC4内のコマンド制御部43は、CPU3から受信したリクエストを保持するキュー43aを備え、キュー43a内のリクエストを、パイプ部47へ順に転送するとともに、パイプ部47経由でTAG__CPメモリ制御部41及びアドレスロックレジスタ部45へ順に転送するための制御を行なう。また、コマンド制御部43は、転送したリクエストに係る処理が完了するまでは、当該リクエストを保持する。また、保持するリクエストの再発行（リトライ）の指示を受けると、コマンド制御部43は、再発行の指示に係るリクエストを、再度キュー43aに登録し再発行を行なう。

[0095] 図8に示すように、パイプ部47は、複数のラッチ回路47a-1～47a-n及び47b-1～47b-o（図8中、m、n及びoはそれぞれ備えられるラッチ回路の数、なお、 $m < n$ である）と、結果確定部47cとを備える。

パイプ部47は、コマンド制御部43からのリクエストをラッチ回路47a-1に入力するとともに、当該リクエストを、ラッチ回路47a-1からラッチ回路47a-2並びにTAG__CPメモリ制御部41及びアドレスロックレジスタ部45に対してそれぞれ出力する。

[0096] ラッチ回路47a-2に入力されたリクエストは、TAG__CPメモリ制御部41における検索処理との待ち合わせのため、ラッチ回路47a-2～47a-nを順に通過して、結果確定部47cに出力される。

一方、TAG__CPメモリ制御部41及びアドレスロックレジスタ部45に入力されたリクエストは、TAG__CPメモリ制御部41において検索処理が行なわれ、キャッシュ検索結果としてラッチ回路47b-1に出力される。キャッシュ検索結果は、ラッチ回路47b-1～47b-oを順に通過して、結果確定部47cに出力される。

[0097] 結果確定部 47c は、ラッチ回路 47a-1~47a-n を通過したリクエストと、ラッチ回路 47b-1~47b-o を通過したキャッシュ検索結果とに基づいて、リクエストの転送先を確定し、第 1 I/F 部 48 に出力する。

第 1 I/F 部 48 は、パイプ部 47 の結果確定部 47c から出力されたリクエストを、結果確定部 47c によって確定された転送先、例えば自 SB2 内の CPU3 やメモリ 5、又は XB9 を経由して他の SB2 内の SC4 等へ送信する。

[0098] なお、図 8 に示す例においては、第 1 I/F 部 48 は、CPU3 又はメモリ 5 に対してリクエストを送信するが、例えば、CPU I/F 部及びメモリ I/F 部のように機能を分割しても良い。

ここで、ラッチ回路 47a-1~47a-n 及び 47b-1~47b-o 並びに後述するラッチ回路 40a は、例えばそれぞれがフリップフロップ (Flip-Flop) である。これらのラッチ回路 47a-1~47a-n 及び 47b-1~47b-o により、リクエストとキャッシュ検索結果とが結果確定部 47c に入力されるタイミングの調整が行なわれる。

[0099] TAG_CP メモリ制御部 41-1~41-4 は、コマンド制御部 43 からのリクエストを入力されると、第 1 実施形態と同様に、当該リクエストに含まれるメモリ 5 の実アドレスから、インデックス及び登録アドレスを抽出する。そして、TAG_CP メモリ制御部 41-1~41-4 は、それぞれの TAG_CP メモリ 42-1~42-4 内から、抽出したインデックス及び登録アドレスに対応するキャッシュタグデータを検索する。

[0100] なお、第 2 実施形態においては、上述の如く、インデックスはメモリ 5 の実アドレスのうちの 18 : 8 ビットであり、登録アドレスはメモリ 5 の実アドレスのうちの 41 : 19 ビットである。

具体的には、図 8 に示すように、TAG_CP メモリ制御部 41 は、リクエストから抽出したインデックスに基づいて、TAG_CP メモリ 42 から同じインデックスの登録アドレスを抽出する。そして、TAG_CP メモリ

制御部 4 1 は、リクエストから抽出した登録アドレス（図 8 中、上位 P A [4 1 : 1 9] 参照）と、T A G _ C P メモリ 4 2 から抽出した登録アドレスとを比較器 4 1 a で比較し、一致するか否かを判定する。

- [0101] 一致した場合、つまりリクエストに係るキャッシュタグデータが T A G _ C P メモリ 4 2 内でヒットした場合には、T A G _ C P メモリ制御部 4 1 は、T A G _ C P メモリ 4 2 から抽出した一致した登録アドレスが含まれるキャッシュタグデータを参照する。

なお、検索によりリクエストに係るキャッシュタグデータが T A G _ C P メモリ 4 2 内でヒット又はミスヒットした場合、その後の処理の内容は、リクエストの内容とキャッシュタグデータのステータスとに応じて決定される。これらの処理の内容の決定は、既知の種々の手法に実現することができるため、ここではその詳細な説明は省略する。

- [0102] リクエスト制御部 4 4 は、第 1 実施形態と同様の動作を行なう。

なお、図 8 に示す例においては、リクエスト制御部 4 4 は、T A G _ C P メモリ 4 2 において U E が発生した場合、U E が発生した T A G _ C P メモリ 4 2 の W A Y の縮退処理を行なうため、レジスタ設定変更部 4 6 a に対して設定変更要求を通知する。

レジスタ設定変更部 4 6 a は、構成情報が保持されている構成制御レジスタ 4 6 b に対して、設定変更要求に基づいた縮退フラグの設定を行なう。

- [0103] また、リクエスト制御部 4 4 は、U E が発生した場合、S C 4 と運用管理部 6 との間のインターフェース機能を備える第 2 I / F 部 4 9 に対して、エラー情報の割り込み通知を出力する。第 2 I / F 部 4 9 は、割り込み通知が入力されると運用管理部 6 に対して当該割り込み通知を出力する。

また、S C 4 は、例えば C E 又は U E が発生した場合に、C E 又は U E が発生した T A G _ C P メモリ 4 2 において稼動している W A Y 数が所定の数（例えば 1）以下である場合には、C E 又は U E が発生した T A G _ C P メモリ 4 2 及び対応する C P U 3 自体の縮退を行なっても良い。

- [0104] つまり、T A G _ C P メモリ 4 2 に C E 又は U E が発生し、且つ C E 又は

UEが発生したTAG__CPメモリ42のWAYの稼働数が所定の数以下の場合、CE又はUEが発生したWAYの縮退処理を行なうと、当該TAG__CPメモリ42において稼働するWAYが無くなる。そこで、第2実施形態においては、リクエスト制御部44は、CE又はUEが発生したTAG__CPメモリ42のWAYの稼働数が所定の数以下の場合、CE又はUEが発生したTAG__CPメモリ42全体を縮退するとともに、当該TAG__CPメモリ42に対応するCPU3自体を縮退させる。

[0105] なお、この場合の縮退処理は、上述した図13に例示する手法により行なうことができる。また、リクエスト制御部44は、CE又はUEが発生した場合に、構成制御レジスタ46bに設定された縮退フラグ等の設定情報を読み込むことで、CE又はUEが発生したTAG__CPメモリ42において稼働しているWAY数を把握することができる。

また、リクエスト制御部44は、CEが発生した回数をカウントし、CEの発生回数が所定の閾値よりも大きい場合に、CE検出リクエストをCPU3に通知しても良い。

[0106] 運用管理部6は、例えば情報処理システム1'で実行中のOSの再起動によってCPU3及びSC4内の縮退状態がリセットされた場合に、運用管理部6内に格納された故障情報に基づいて、被疑箇所に対応するWAYを再度縮退させる。その際、運用管理部6は、第2I/F部59を介して、レジスタ設定変更部46aに対し、故障情報に基づいて設定変更要求を通知する。

[0107] レジスタ設定変更部46aは、リクエスト制御部44からの設定変更要求と同様に、構成情報が保持されている構成制御レジスタ46bに対して、設定変更要求に基づいた縮退フラグの設定を行なう。

なお、第2実施形態におけるCPU3においても、構成制御レジスタ（図示省略）が備えられており、運用管理部6は、CPU3が備える構成制御レジスタに対しても、設定変更を行なうことができる。

[0108] アドレスロックレジスタ部45は、コマンド制御部43からのリクエストを入力されると、第1実施形態と同様に、当該リクエスト内のアドレス情報

をロックレジスタ 45 a に保持する。

具体的には、図 8 に示すように、アドレスロックレジスタ部 45 は、コマンド制御部 43 から転送されるリクエスト内の実アドレスからフルアドレス（例えば、第 2 実施形態においては 41 : 3 ビット）とインデックスとを抽出し、抽出したフルアドレスをロックレジスタ 45 a に保持する。

[0109] また、アドレスロックレジスタ部 45 は、処理中のリクエスト内の実アドレスにおけるフルアドレスと、ロックレジスタ 45 a に保持されたリクエスト内の実アドレスにおけるフルアドレスとを比較する、アドレス競合検査部 45 b を備える。

なお、アドレス競合検査部 45 b は、図 8 に示す例においては、リクエスト制御部 44 内に備えられているが、アドレスロックレジスタ部 45 と結線されており、アドレスロックレジスタ部 45 の一機能として動作する。なお、アドレス競合検査部 45 b は、アドレスロックレジスタ部 45 内に備えられても良い。

[0110] アドレス競合検査部 45 b は、比較器 45 b a を備え、後発リクエスト内の実アドレスにおけるフルアドレスが、ラッチ回路 40 a に入力されたタイミングで、ラッチ回路 40 a から比較器 45 b a の PA [41 : 3] に入力される。また、ロックレジスタ 45 a に保持されたリクエスト内の実アドレスにおけるフルアドレスが、ロックレジスタ 45 a（図 8 中、アドレスロックレジスタ部 45 内の REG_ADRS [41 : 3]）から比較器 45 b a の REG_ADRS [41 : 3] に入力される。

[0111] そして、アドレス競合検査部 45 b において、比較器 45 b a により、入力された二つのフルアドレスが一致したと判定された場合には、比較器 45 b a は、当該後発リクエストに係るフルアドレスがビジー状態である旨（フルアドレスビジー）をコマンド制御部 43 に通知する。

このように、第 2 実施形態におけるアドレスロックレジスタ部 45 は、第 1 実施形態と同様に、コマンド制御部 43 から転送される後発リクエストに係る処理をキャンセル及びリトライさせて、処理中のリクエストと競合しな

いようにガードするガード機能を備える。

- [0112] また、アドレスロックレジスタ部 4 5 は、第 1 実施形態と同様に、上述した処理に加えて、UE が発生した際には、CPU 3 及び SC 4 による被疑箇所の縮退処理が完了するまで、UE 検出リクエストにおける被疑箇所のフルアドレスをロックレジスタ 4 5 a に保持する。

これにより、アドレス競合検査部 4 5 b は、後発リクエストに係るフルアドレス (PA [4 1 : 3]) と UE 検出リクエストにおける被疑箇所のフルアドレス (REG_ADRS [4 1 : 3]) とが一致した場合、第 1 実施形態と同様に、後発リクエストのフルアドレスビジーをコマンド制御部 4 3 に通知する。

- [0113] すなわち、アドレスロックレジスタ部 4 5 は、CPU 3 から TAG メモリ 3 2 の縮退処理完了通知を受信するとともに、リクエスト制御部 4 4 が TAG_CP メモリ 4 2 の UE が発生した WAY を縮退させるまでは、他のリクエストによる当該 UE が発生した TAG_CP メモリ 4 2 へのアクセスを抑制する。

このように、第 2 実施形態におけるアドレスロックレジスタ部 4 5 は、第 1 実施形態と同様に、ガード機能により、処理中のリクエストと競合しないように後発リクエストに係る処理をガードするとともに、UE が発生した場合に被疑箇所をガードすることができる。

- [0114] なお、図 8 に示す例においては、パイプ部 4 7 の結果確定部 4 7 c と、第 1 I/F 部 4 8 との間から、各 TAG_CP メモリ制御部 4 1 及びアドレスロックレジスタ部 4 5 ヘステータス更新用のラインが設けられている。これにより、リクエストが結果確定部 4 7 c から出力された段階で、各 TAG_CP メモリ制御部 4 1 におけるステータスの更新及びアドレスロックレジスタ部 4 5 におけるロックの制御が行なわれる。

- [0115] つまり、アドレスロックレジスタ部 4 5 は、UE が発生した場合に、結果確定部 4 7 c からステータス更新用のラインを介してロックの維持を示す情報を入力され、ロックレジスタ 4 5 a において UE 検出リクエストに係るロ

ックを維持することができる。

〔２－２〕第２実施形態の情報処理システムのＣＥ又はＵＥ発生時の動作次に、上述の如く構成された情報処理システム１′におけるＳＣ４のＴＡＧ＿ＣＰメモリ４２でＣＥ又はＵＥが発生した場合の縮退処理について説明する。

[0116] 図９は、第２実施形態の一例としてのＳＣ４内のＴＡＧ＿ＣＰメモリ４２－２でＣＥ又はＵＥが発生した場合の縮退処理を説明するためのフローチャートである。

はじめに、図９に示すように、システムの運用中、ＴＡＧ＿ＣＰメモリ４２－２でエラーが発生し、ＴＡＧ＿ＣＰメモリ制御部４１－２によって検出された場合（ステップＳ１１）、ＳＣ４により、検出されたエラーがＣＥであるか否かが判断される（ステップＳ１２）。

[0117] 検出されたエラーがＣＥであると判断された場合（ステップＳ１２のＹｅｓルート）、リクエスト制御部４４により、ＣＥの発生数が所定の閾値より大きいかが判断される（ステップＳ１３）。

ＣＥ発生数が所定の閾値以下であると判断された場合（ステップＳ１３のＮｏルート）、リクエスト制御部４４により、ＣＥ発生数をカウントするカウンタの値がインクリメントされ、情報処理システム１′の運用に戻る。

[0118] 一方、ＣＥ発生数が所定の閾値より大きいと判断された場合（ステップＳ１３のＹｅｓルート）、リクエスト制御部４４により、ＣＥが発生したＴＡＧ＿ＣＰメモリ４２－２において稼動しているＷＡＹ数が所定の数（ここでは１）以下か否かが判断される（ステップＳ１４）。

稼動しているＷＡＹ数が１よりも大きいと判断された場合（ステップＳ１４のＮｏルート）、リクエスト制御部４４からＣＥが発生したＴＡＧ＿ＣＰメモリ４２－２に対応するＣＰＵ３－２に対して、ＣＥ通知リクエストが通知される（ステップＳ１５）。なお、このリクエストには、ＥＣＣにより訂正された被疑箇所のインデックスと、ＷＡＹ情報とが含まれる。

[0119] ＣＰＵ３－２では、被疑箇所の情報の通知により、ＣＥが発生したＴＡＧ

メモリ 32-2 内のWAYのキャッシュデータがメモリに掃き出されるとともに、CEが発生したTAGメモリ 32-2のWAYに対して縮退処理が行なわれる（ステップS16）。そして、CPU3-2により、SC4へ縮退処理完了の通知が行なわれる（ステップS17）。

縮退処理完了通知を受けたリクエスト制御部44では、CEが発生したTAG__CPメモリ42-2のWAYに対して縮退処理が行なわれる（ステップS18）。そして、リクエスト制御部44により、運用管理部6に対して、CEに関するエラー情報が通知され（ステップS19）、運用管理部6の制御情報に故障情報が記録される（ステップS20）。その後、情報処理システム1'では、運用が継続される（ステップS21）。

[0120] 一方、ステップS14において、リクエスト制御部44により、稼動しているWAY数が1以下と判断された場合（ステップS14のYesルート）、リクエスト制御部44から運用管理部6に対して、CEが発生し、CPU3-2の縮退処理を行なうことが割り込みで通知される（ステップS22）。

割り込み通知後、運用管理部6では、自身で管理する制御情報に、被疑対象のTAG__CPメモリ42-2に対応したTAGメモリ32を有するCPU3-2を示す情報とWAY情報とが故障情報として記録される（ステップS23）。そして、運用管理部6により、情報処理システム1'で実行中のOSが再起動される（ステップS24）。

[0121] OSの再起動後、運用管理部6により、制御情報の故障情報が読み込まれ（ステップS25）、故障情報に記録されているCPU3-2は、立ち上げ処理が行なわれず、他の正常なCPU3-1、3-3及び3-4についてのみ立ち上げ処理が行なわれる。つまり、運用管理部6により、被疑箇所に対応するCPU3-2が縮退処理される（ステップS26）。その後、情報処理システム1'では、運用が再開される（ステップS27）。

[0122] 一方、ステップS12において、検出されたエラーがCEではない、つまりUEであると判断された場合（ステップS12のNoルート）、リクエス

ト制御部44により、UEが発生したTAG__CPメモリ42-2において稼動しているWAY数が所定の数（ここでは1）以下か否かが判断される（ステップS28）。

稼動しているWAY数が1よりも大きいと判断された場合（ステップS28のNルート）、図4を用いて上述したステップS2～S10の処理が行なわれる（ステップS29）。

[0123] すなわち、TAG__CPメモリ制御部41からリクエスト制御部44にTAG__CPエラー通知が通知され、リクエスト制御部44により、コマンド制御部43に対してUE検出リクエストの保留指示が通知される（ステップS2）。そして、コマンド制御部43内でUE検出リクエストが保留状態として保持される。

次いで、リクエスト制御部44により、UEが検出されたTAG__CPメモリ42-2に対応するCPU3-2に対して、UE通知リクエストが通知される（ステップS3）。

[0124] CPU3-2では、被疑箇所に対応するCM31-2のWAYの縮退処理が行なわれ（ステップS4）、リクエスト処理部44に対して、縮退処理完了通知が送信される（ステップS5）。

また、リクエスト制御部44により、SC4内で被疑箇所に対応したWAYの縮退処理が行なわれる（ステップS6）。

[0125] CPU3-2から縮退処理完了通知が通知され、UEが検出されたTAG__CPメモリ42-2のWAYの縮退処理の完了後、リクエスト処理部44により、コマンド制御部43に対してUE検出リクエストの処理再開の指示が通知される（ステップS7）。コマンド制御部43内では、UE検出リクエストに係る処理が再開（再発行）される。

そして、リクエスト制御部44により、運用管理部6に対して、UEに関するエラー情報が割り込み通知される（ステップS8）。

[0126] 割り込み通知が入力されると、運用管理部6のファームウェアにより、運用管理部6が管理する制御情報に対して、縮退されたWAYに関する情報が

故障情報として格納され（ステップS 9）、情報処理システム1' による運用が継続される（ステップS 10）。

一方、ステップS 28において、リクエスト制御部44により、稼動しているWAY数が1以下と判断された場合（ステップS 28のYesルート）、リクエスト制御部44から運用管理部6に対して、UEが発生し、CPU3-2の縮退処理を行なうことが割り込み通知される（ステップS 22）。以降、UEが発生したTAG__CPメモリ42に対応するCPU3-2について、上述したステップS 23以降の処理が行なわれる。

[0127] このように、第2実施形態としての情報処理システム1'（特にSC4）によれば、上述した第1実施形態と同様の効果を得ることができる。

また、第2実施形態の一例としての情報処理システム1'によれば、リクエスト制御部44により、CEの発生回数がカウントされ、CEの発生回数が所定の閾値よりも大きい場合に、CE検出リクエストがCPU3に通知される。

[0128] これにより、CEの発生回数が所定の閾値以下の場合には、WAYの縮退処理が行なわれないため、WAYの縮退処理に伴う情報処理システム1'の性能低下を抑えることができる。

〔3〕その他

以上、本発明の好ましい実施形態及び変形例について詳述したが、本発明は、かかる特定の実施形態及び変形例に限定されるものではなく、本発明の趣旨を逸脱しない範囲内において、種々の変形、変更して実施することができる。

[0129] 例えば、上述した第1及び第2実施形態においては、SB2内のCPU3が4つの場合について説明したが、CPU3の個数は、これに限定されるものではなく、1つでも良いし、他の数でも良い。CPU3がいずれの数の場合でも、SC4内に、各TAGメモリ32（CPU3）に対応したTAG__CPメモリ42が備えられれば良い。

また、上述した第1及び第2実施形態においては、UEが検出されたリク

エストについては、CPU 3 及び SC 4 における対象のWAYの縮退処理が完了するまでは処理を保留するが、これに限定されるものではない。例えば、SC 4 は、UEが検出された際に、キャッシュタグデータから、他のCPU 3、つまりUEが検出されたTAG__CPメモリ 4 2に対応するCPU 3とは異なるCPU 3に最新のデータがあることが確認できた場合、UE検出リクエストに対しては通常通り処理を行なっても良い。なお、他のCPUに最新のデータがある場合は、UEが検出されたリクエストに対する他WAYのTAG__CPメモリ 4 2のキャッシュタグデータのステータスは、MOSSI プロトコルにおいては、例えば“M”又は“O”である。

[0130] 例えば、UE検出リクエストについて、通常処理の過程でアドレスロックを外す処理、つまりUE検出リクエストについて通常通り処理を行ない、当該処理の完了後にロックレジスタ 4 5 aからアドレスを削除する処理を抑止する。そして、CPU 3の縮退処理に伴うCM 3 1のデータの掃き出し処理が完了した後に、当該アドレスロックを外す処理を行なう。このために、ロックレジスタ 4 5 aで保持されるアドレスや、パイプ部 4 7を通過するデータに、リクエストに係るステータスが“M”や“O”であることを示す情報が付加されることが好ましい。

[0131] あるいは、CPU 3の縮退処理に伴うCM 3 1のデータの掃き出し処理が完了するまでは、UEが複数回発生してしまうのを許容することでも良い。

上述のように、UEが発生したTAG__CPメモリ 4 2に対応するCPU 3と、UE検出リクエストの要求元のCPU 3及び最新データを保持しているCPU 3とが異なるCPUの場合には、UE検出リクエストの保留による処理遅延を防ぐことができる。

[0132] さらに、上述した第2実施形態における図9に示すフローチャートにおいて、ステップS 1 4のYesルートからのステップS 2 2～ステップS 2 7の処理、つまり、TAG__CPメモリ 4 2にCEが発生し、且つCEが発生したTAG__CPメモリ 4 2で稼働しているWAYが所定の数以下の場合の処理は、これに限定されるものではない。例えば、情報処理システム 1' は

、ステップS 2 2～ステップS 2 7の処理を実行せずに、そのまま運用を継続してもよい。

[0133] また、上述した第2実施形態における図9に示すフローチャートにおいて、ステップS 2 8のY e sルートからのステップS 2 2～ステップS 2 7の処理、つまり、T A G_ C Pメモリ4 2にU Eが発生し、且つU Eが発生したT A G_ C Pメモリ4 2で稼働しているW A Yが所定の数以下の場合の処理は、これに限定されるものではない。例えば、情報処理システム1' は、ステップS 2 2～ステップS 2 7の処理を実行せずに、S C 4がO Sに対して最後のW A YでU Eが発生した旨を通知し、O Sでは処理中のプロセスを全て終了させた後に終了（Shutdown）処理を行なっても良い。

[0134] さらに、第1実施形態及び第2実施形態における図4に示すフローチャートにおいて、ステップS 6の処理は、ステップS 3～ステップS 5の処理の前、或いはステップS 3～ステップS 5の間に行なわれても良い。つまり、リクエスト処理部4 4によるU Eが発生したT A G_ C Pメモリ4 2のW A Yの縮退処理は、当該W A Yに対応するC P U 3のW A Yの縮退処理よりも前に行なわれても良いし、並行して行なわれても良い。

符号の説明

[0135] 1, 1' 情報処理システム（情報処理装置）
 2 S B
 3, 3-1, 3-2, 3-3, 3-4 C P U（演算処理装置）
 3 0 0, 3 0 0-1, 3 0 0-2, 3 0 0-3, 3 0 0-4 C P U
 3 1, 3 1-1, 3 1-2, 3 1-3, 3 1-4 C M（キャッシュメモリ）
 3 1 0-1, 3 1 0-2, 3 1 0-3, 3 1 0-4 C M（キャッシュメモリ）
 3 2, 3 2-1, 3 2-2, 3 2-3, 3 2-4 T A Gメモリ（第1タグメモリ）
 3 2 0-1, 3 2 0-2, 3 2 0-3, 3 2 0-4 T A Gメモリ

4, 400 SC (システムコントローラ)

40a ラッチ回路

41, 41-1, 41-2, 41-3, 41-4 TAG_CPメモリ
制御部

41a 比較器

42, 42-1, 42-2, 42-3, 42-4 TAG_CPメモリ
(第2タグメモリ)

420, 420-1, 420-2, 420-3, 420-4 TAG_
CPメモリ

43 コマンド制御部

43a キュー

44 リクエスト制御部

45 アドレスロックレジスタ部

45a ロックレジスタ

45b アドレス競合検査部

45ba 比較器

46 レジスタ部

46a レジスタ設定変更部

46b 構成制御レジスタ

47 パイプ部

47a-1~47a-n, 47b-1~47b-o ラッチ回路

47c 結果確定部

48 第1I/F部

49 第2I/F部

5, 500 メモリ

6, 600 運用管理部

7 記憶部

8 XBU

9 X B

請求の範囲

- [請求項1] キャッシュメモリ及び第1タグメモリを有する演算処理装置と、前記演算処理装置と他の処理装置との通信制御を行なうシステムコントローラとを有する情報処理装置であって、
- 前記システムコントローラは、
- 前記演算処理装置から受信したリクエストを保持し、当該リクエストが要求先において処理されなかった場合に当該リクエストを再発行するコマンド制御部と、
- 前記第1タグメモリに記憶されるデータの複製データを保持する第2タグメモリと、
- 前記第2タグメモリから読み出されたデータが訂正不可能なエラーであるUE（Uncorrectable Error）を起こした場合に、前記UEが発生した第2タグメモリのウェイ情報を前記演算処理装置に通知するリクエスト制御部とを有し、
- 前記演算処理装置は、前記リクエスト制御部から前記UEが発生した通知を受けると、当該UEが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリのウェイを縮退させて、前記第1タグメモリのウェイの縮退処理が完了したことを前記リクエスト制御部へ通知し、
- 前記リクエスト制御部は、前記UEが発生した場合に、前記第2タグメモリのUEが発生したウェイを縮退させるとともに、前記演算処理装置から、前記第1タグメモリの縮退処理が完了した旨の通知を受信した後、前記コマンド制御部に当該UEに係るリクエストを再発行させる指示を行なうことを特徴とする、情報処理装置。
- [請求項2] 前記リクエスト制御部は、前記UEが発生した場合に、前記コマンド制御部に対して前記UEに係るリクエストを保留させる指示を行なうことを特徴とする、請求項1記載の情報処理装置。
- [請求項3] 前記システムコントローラは、

前記UEが発生した場合に、前記演算処理装置から前記第1タグメモリの縮退処理が完了した旨の通知を受信するとともに、前記第2タグメモリのUEが発生したウェイを縮退させるまでは、他のリクエストによる前記第2タグメモリの当該UEが発生したデータへのアクセスを抑止するアドレスロックレジスタ部を有することを特徴とする、請求項1又は請求項2記載の情報処理装置。

[請求項4]

前記アドレスロックレジスタ部は、

リクエスト内のアドレス情報を保持するロックレジスタを備えるとともに、

前記UEが発生した場合に、当該UEに係るリクエスト内のアドレス情報を抽出し、抽出したアドレス情報を前記ロックレジスタに保持させ、当該UEに係るリクエストの後発リクエスト内のアドレス情報が、前記ロックレジスタに保持された前記UEに係るリクエスト内の前記アドレス情報と一致した場合、前記コマンド制御部に対して当該後発リクエストを再発行させる指示を行なうことを特徴とする、請求項3記載の情報処理装置。

[請求項5]

前記システムコントローラは、

前記第2タグメモリのウェイを縮退することを示す縮退フラグを保持するレジスタ部を有し、

前記リクエスト制御部は、前記レジスタ部に対して、UEが発生した前記第2タグメモリのウェイに係る縮退フラグを設定することにより、前記第2タグメモリのUEが発生したウェイを縮退させることを特徴とする、請求項1～4のいずれか1項記載の情報処理装置。

[請求項6]

前記情報処理装置は、前記キャッシュメモリ及び前記第1タグメモリを有する前記演算処理装置を複数備え、

前記システムコントローラは、前記複数の演算処理装置にそれぞれ備えられた前記複数の第1タグメモリに対応して、前記第2タグメモリを複数備えることを特徴とする、請求項1～5のいずれか1項記載

の情報処理装置。

[請求項7]

前記情報処理装置は、

前記情報処理装置に係る制御を行なう運用管理部を有し、

前記リクエスト制御部は、前記第2タグメモリのUEが発生したウェイを縮退させた後、前記運用管理部に対して、前記UEに関するエラー情報を通知し、

前記運用管理部は、前記リクエスト制御部からの通知に基づいて、縮退されたウェイに関する情報を保持するとともに、前記情報処理装置において実行されるOS（Operating System）が再起動した場合に、保持している縮退させたウェイに関する情報に基づいて、前記第1及び第2タグメモリの当該ウェイを縮退させることを特徴とする、請求項1～6のいずれか1項記載の情報処理装置。

[請求項8]

前記リクエスト制御部は、前記第2タグメモリにおいてUEが発生した際に稼働中のウェイが所定の数以下である場合、前記運用管理部に対して、当該UEが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリを有する演算処理装置を示す情報を通知し、

前記運用管理部は、前記リクエスト制御部から通知された前記演算処理装置を示す情報を保持するとともに、前記情報処理装置において実行されるOSを再起動させ、保持している前記演算処理装置を示す情報に基づいて、当該演算処理装置を縮退させることを特徴とする、請求項7記載の情報処理装置。

[請求項9]

前記リクエスト制御部は、前記第2タグメモリから読み出されたデータが訂正可能なエラーであるCE（Correctable Error）を起こした場合に、前記CEが発生した第2タグメモリのウェイ情報を前記演算処理装置に通知し、

前記演算処理装置は、前記リクエスト制御部から前記CEが発生した通知を受けると、当該CEが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリのウェイを縮退させて、前記第1タグ

メモリのウェイの縮退処理が完了したことを前記リクエスト制御部へ通知し、

前記リクエスト制御部は、前記ＣＥが発生した場合に、前記第２タグメモリのＣＥが発生したウェイを縮退させることを特徴とする、請求項７記載の情報処理装置。

[請求項10]

前記リクエスト制御部は、前記第２タグメモリにおいてＣＥが発生した際に稼働中のウェイが所定の数以下である場合、前記運用管理部に対して、当該ＣＥが発生した前記第２タグメモリのウェイに対応する前記第１タグメモリを有する演算処理装置を示す情報を通知し、

前記運用管理部は、前記リクエスト制御部から通知された前記演算処理装置を示す情報を保持するとともに、前記情報処理装置において実行されるＯＳを再起動させ、保持している前記演算処理装置を示す情報に基づいて、当該演算処理装置を縮退させることを特徴とする、請求項９記載の情報処理装置。

[請求項11]

キャッシュメモリ及び第１タグメモリを有する演算処理装置と、前記演算処理装置と他の処理装置との通信制御を行なうシステムコントローラとを有する情報処理装置のキャッシュ制御方法であって、

前記システムコントローラにおいて、

前記第１タグメモリに記憶されるデータの複製データを保持する第２タグメモリから読み出されたデータが訂正不可能なエラーであるＵＥ（Uncorrectable Error）を起こした場合に、前記ＵＥが発生した前記第２タグメモリのウェイ情報を前記演算処理装置に通知し、

前記演算処理装置において、

前記ＵＥが発生した通知を受けると、当該ＵＥが発生した前記第２タグメモリのウェイに対応する前記第１タグメモリのウェイを縮退させて、前記第１タグメモリのウェイの縮退処理が完了したことを前記システムコントローラへ通知して、

前記システムコントローラにおいて、

前記演算処理装置から、前記第2タグメモリのUEが発生したウェイを縮退するとともに、前記第1タグメモリの縮退処理が完了した旨の通知を受信した後、当該UEに係るリクエストを再発行させることを特徴とする、情報処理装置のキャッシュ制御方法。

[請求項12]

前記システムコントローラにおいて、

前記第2タグメモリから読み出されたデータが前記UEを起こした場合に、前記UEに係るリクエストを保留させることを特徴とする、請求項11記載の情報処理装置のキャッシュ制御方法。

[請求項13]

前記システムコントローラにおいて、

前記演算処理装置から前記第1タグメモリの縮退処理が完了した旨の通知を受信するとともに、前記第2タグメモリのUEが発生したウェイを縮退させるまでは、他のリクエストによる前記第2タグメモリの当該UEが発生したデータへのアクセスを抑止することを特徴とする、請求項11又は請求項12記載の情報処理装置のキャッシュ制御方法。

[請求項14]

前記システムコントローラにおいて、

前記UEが発生した場合に、当該UEに係るリクエスト内のアドレス情報を抽出し、抽出したアドレス情報を保持するとともに、当該UEに係るリクエストの後発リクエスト内のアドレス情報が、前記保持している前記UEに係るリクエスト内の前記アドレス情報と一致した場合、当該後発リクエストを再発行させることを特徴とする、請求項13記載の情報処理装置のキャッシュ制御方法。

[請求項15]

前記システムコントローラにおいて、

前記UEが発生した場合に、前記システムコントローラが備えるレジスタ部に対して、UEが発生した前記第2タグメモリのウェイを縮退することを示す縮退フラグを設定することにより、前記第2タグメモリのUEが発生したウェイを縮退することを特徴とする、請求項11～14のいずれか1項記載の情報処理装置のキャッシュ制御方法。

[請求項16] 前記情報処理装置は、前記キャッシュメモリ及び前記第1タグメモリを有する前記演算処理装置を複数備え、

前記システムコントローラは、前記複数の演算処理装置にそれぞれ備えられた前記複数の第1タグメモリに対応して、前記第2タグメモリを複数備えることを特徴とする、請求項11～15のいずれか1項記載の情報処理装置のキャッシュ制御方法。

[請求項17] 前記情報処理装置は、

前記情報処理装置に係る制御を行なう運用管理部を有し、

前記システムコントローラにおいて、

前記第2タグメモリのUEが発生したウェイを縮退させた後、前記運用管理部に対して、前記UEに関するエラー情報を通知し、

前記運用管理部において、

前記システムコントローラからの通知に基づいて、縮退されたウェイに関する情報を保持するとともに、前記情報処理装置において実行されるOS（Operating System）が再起動した場合に、保持している縮退させたウェイに関する情報に基づいて、前記第1及び第2タグメモリの当該ウェイを縮退させることを特徴とする、請求項11～16のいずれか1項記載の情報処理装置のキャッシュ制御方法。

[請求項18] 前記システムコントローラにおいて、

前記第2タグメモリにおいてUEが発生した際に稼働中のウェイが所定の数以下である場合、前記運用管理部に対して、当該UEが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリを有する演算処理装置を示す情報を通知し、

前記運用管理部において、

前記システムコントローラから通知された前記演算処理装置を示す情報を保持するとともに、前記情報処理装置において実行されるOSを再起動させ、保持している前記演算処理装置を示す情報に基づいて、当該演算処理装置を縮退させることを特徴とする、請求項17記載

の情報処理装置のキャッシュ制御方法。

[請求項19]

前記システムコントローラにおいて、

前記第2タグメモリから読み出されたデータが訂正可能なエラーであるC E (Correctable Error) を起こした場合に、前記C Eが発生した第2タグメモリのウェイ情報を前記演算処理装置に通知し、

前記演算処理装置において、

前記システムコントローラから前記C Eが発生した通知を受けると、当該C Eが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリのウェイを縮退させて、前記第1タグメモリのウェイの縮退処理が完了したことを前記システムコントローラへ通知し、

前記システムコントローラにおいて、

前記C Eが発生した場合に、前記第2タグメモリのC Eが発生したウェイを縮退させることを特徴とする、請求項17記載の情報処理装置のキャッシュ制御方法。

[請求項20]

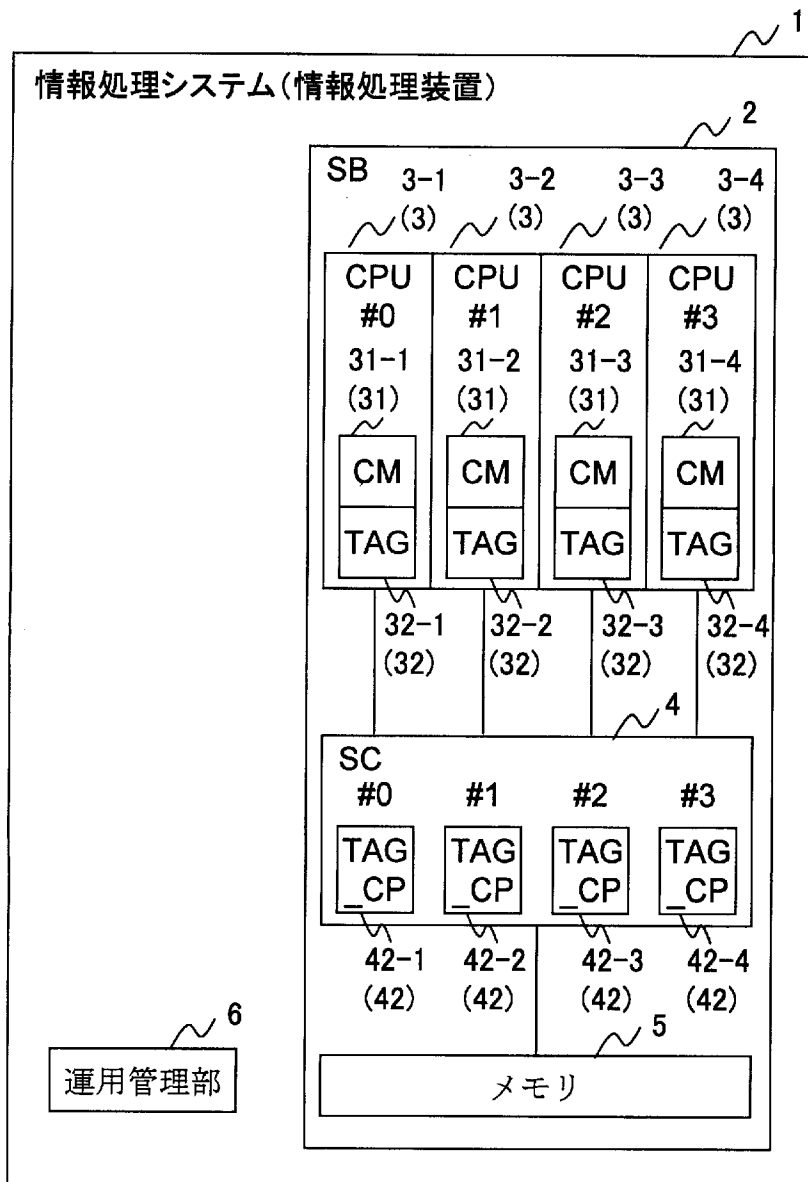
前記システムコントローラにおいて、

前記第2タグメモリにおいてC Eが発生した際に稼働中のウェイが所定の数以下である場合、前記運用管理部に対して、当該C Eが発生した前記第2タグメモリのウェイに対応する前記第1タグメモリを有する演算処理装置を示す情報を通知し、

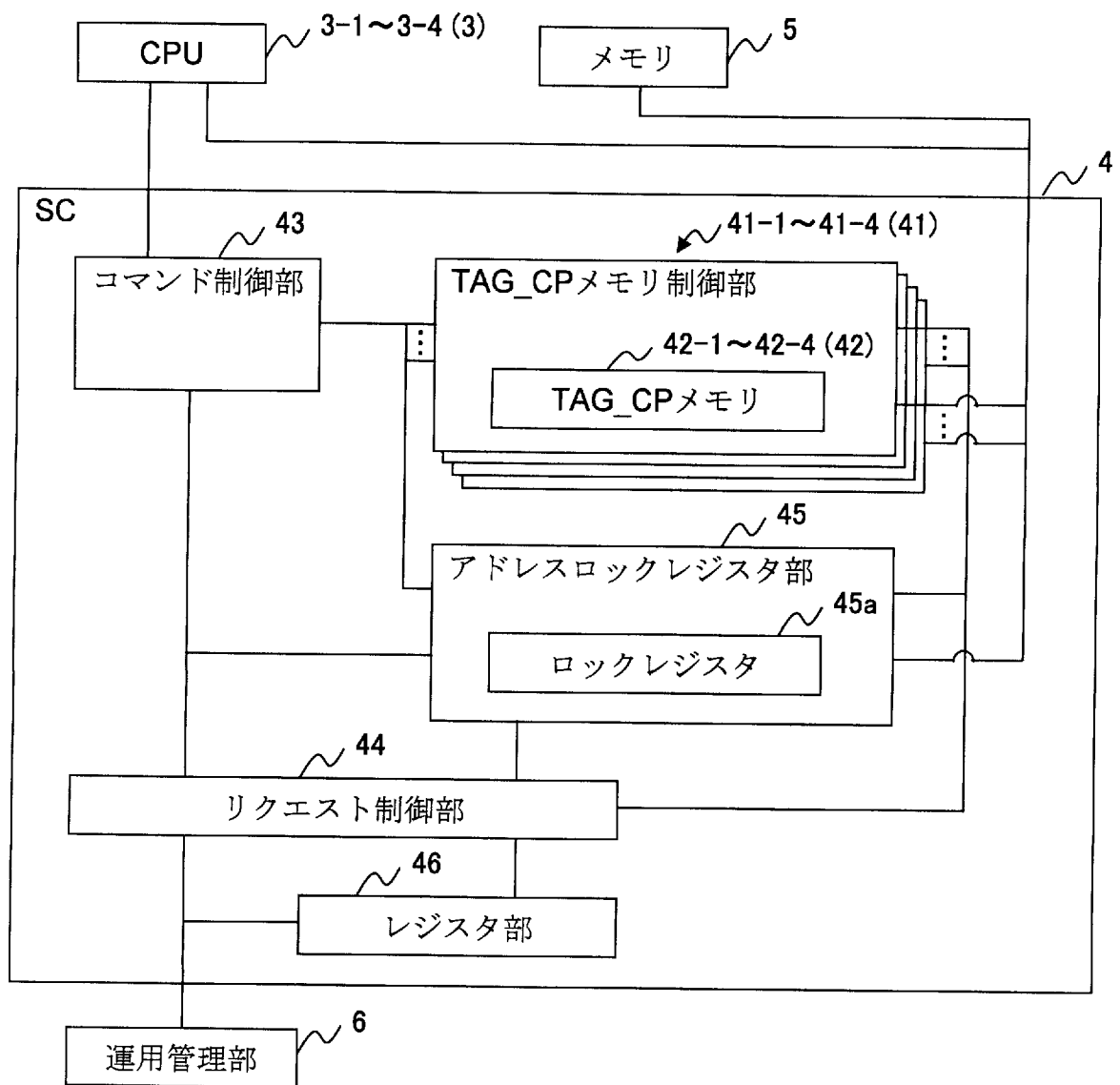
前記運用管理部において、

前記システムコントローラから通知された前記演算処理装置を示す情報を保持するとともに、前記情報処理装置において実行されるOSを再起動させ、保持している前記演算処理装置を示す情報に基づいて、当該演算処理装置を縮退させることを特徴とする、請求項19記載の情報処理装置のキャッシュ制御方法。

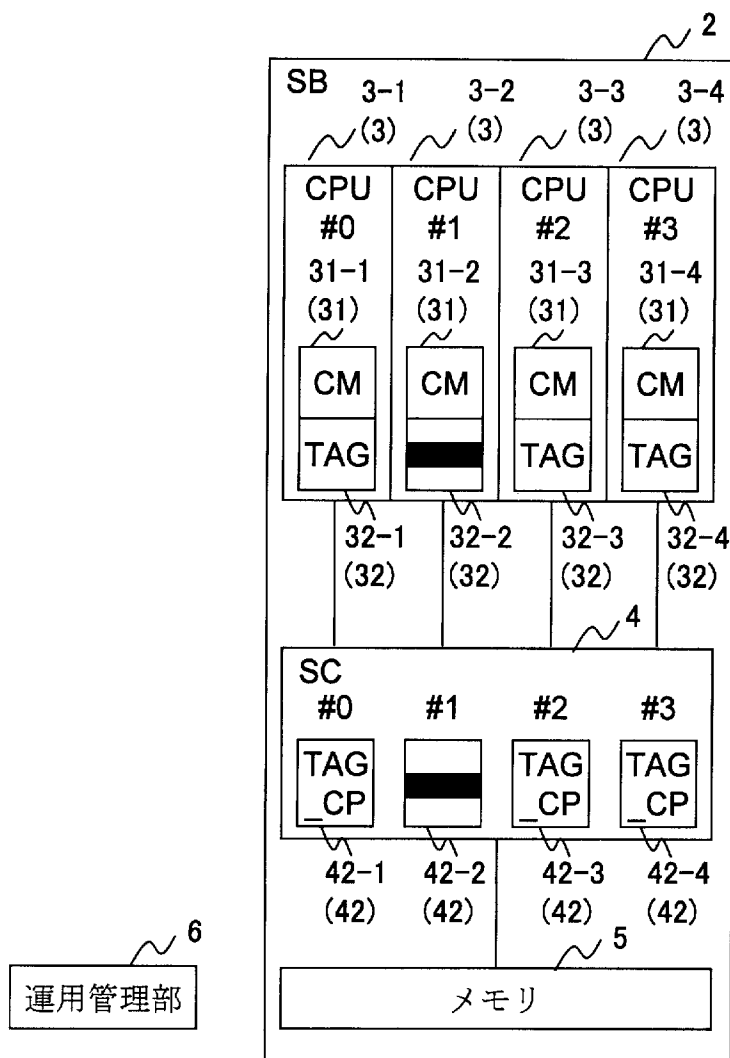
[図1]



[図2]

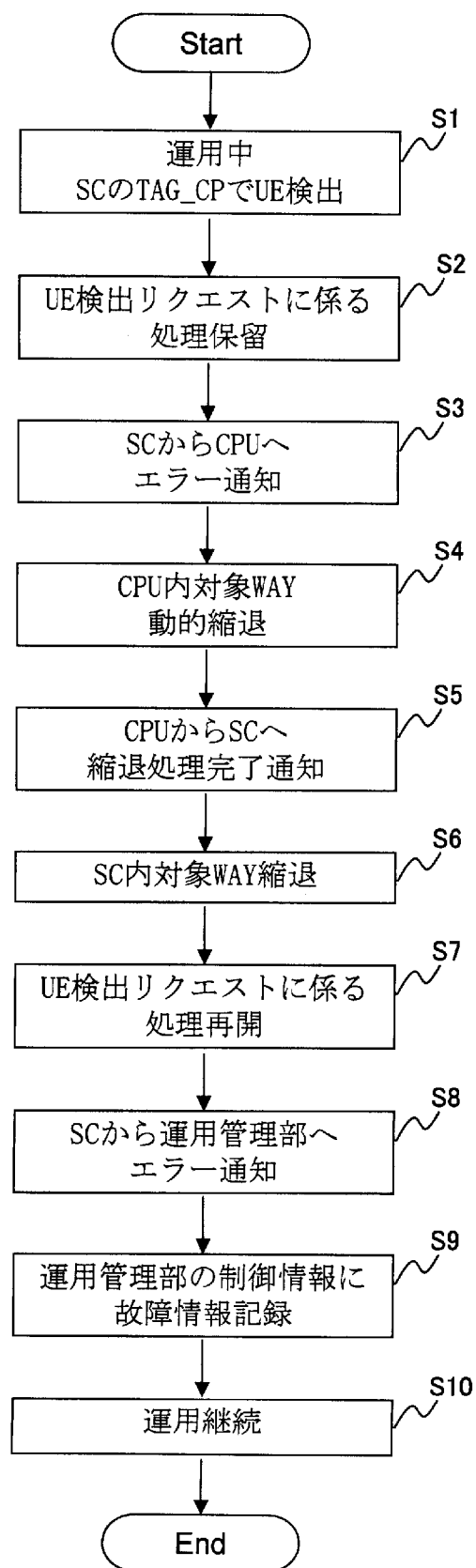


[図3]

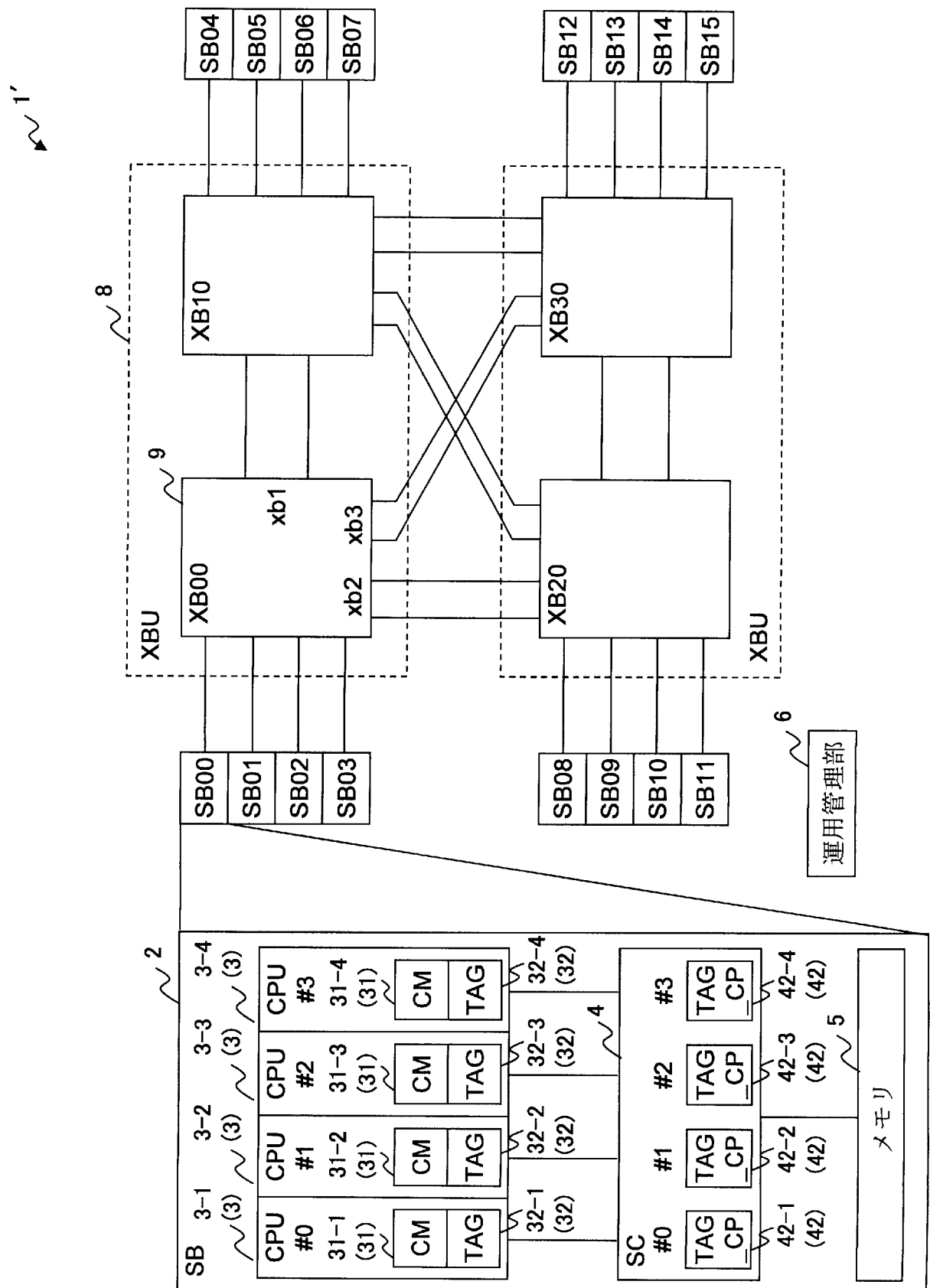


■ : 縮退範囲

[図4]



[図5]



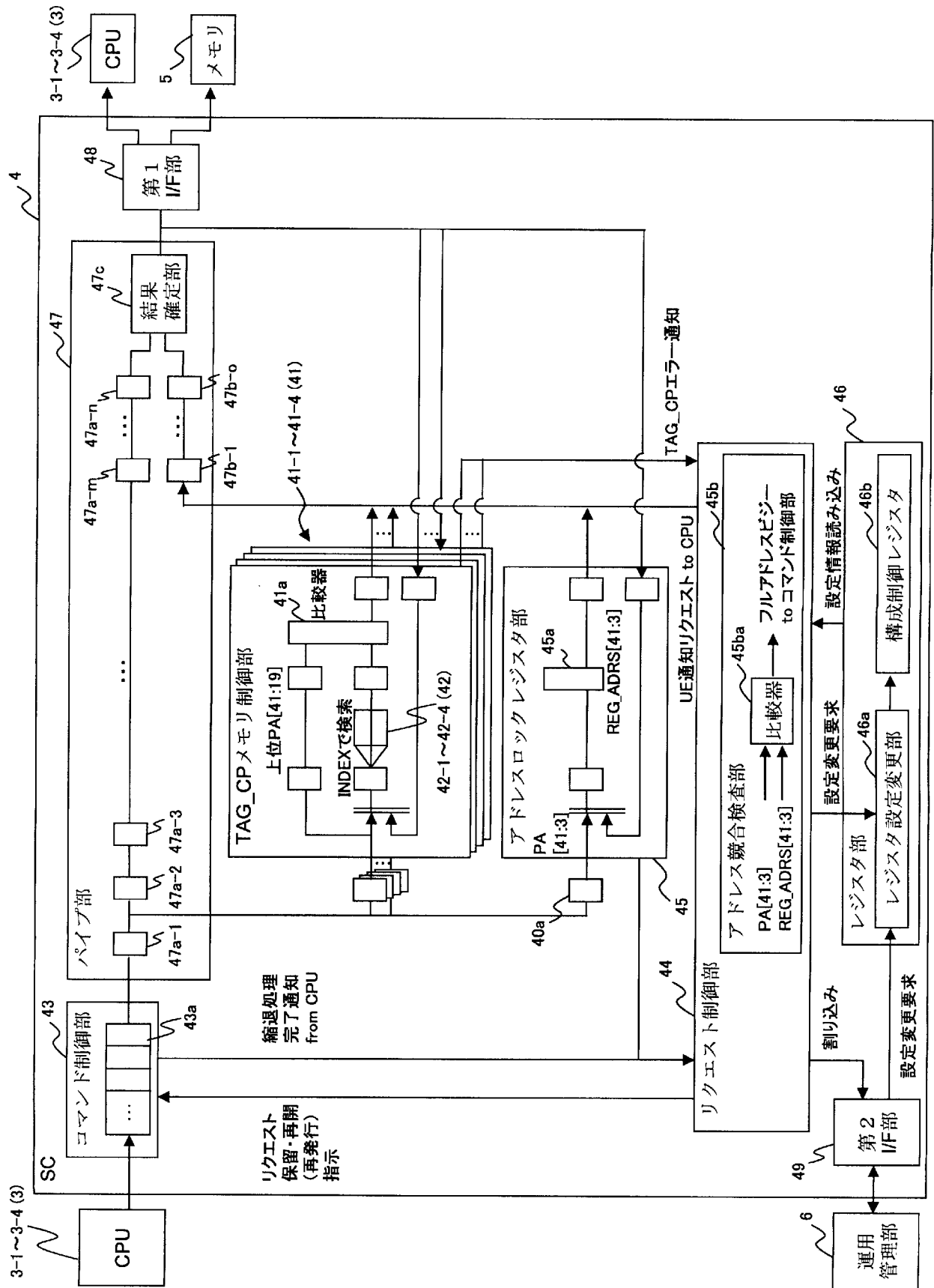
[図6]

A0 (256B) A1 (256B) A2 (256B) A2047 (256B)
B0 (256B) B1 (256B) B2 (256B) B2047 (256B)

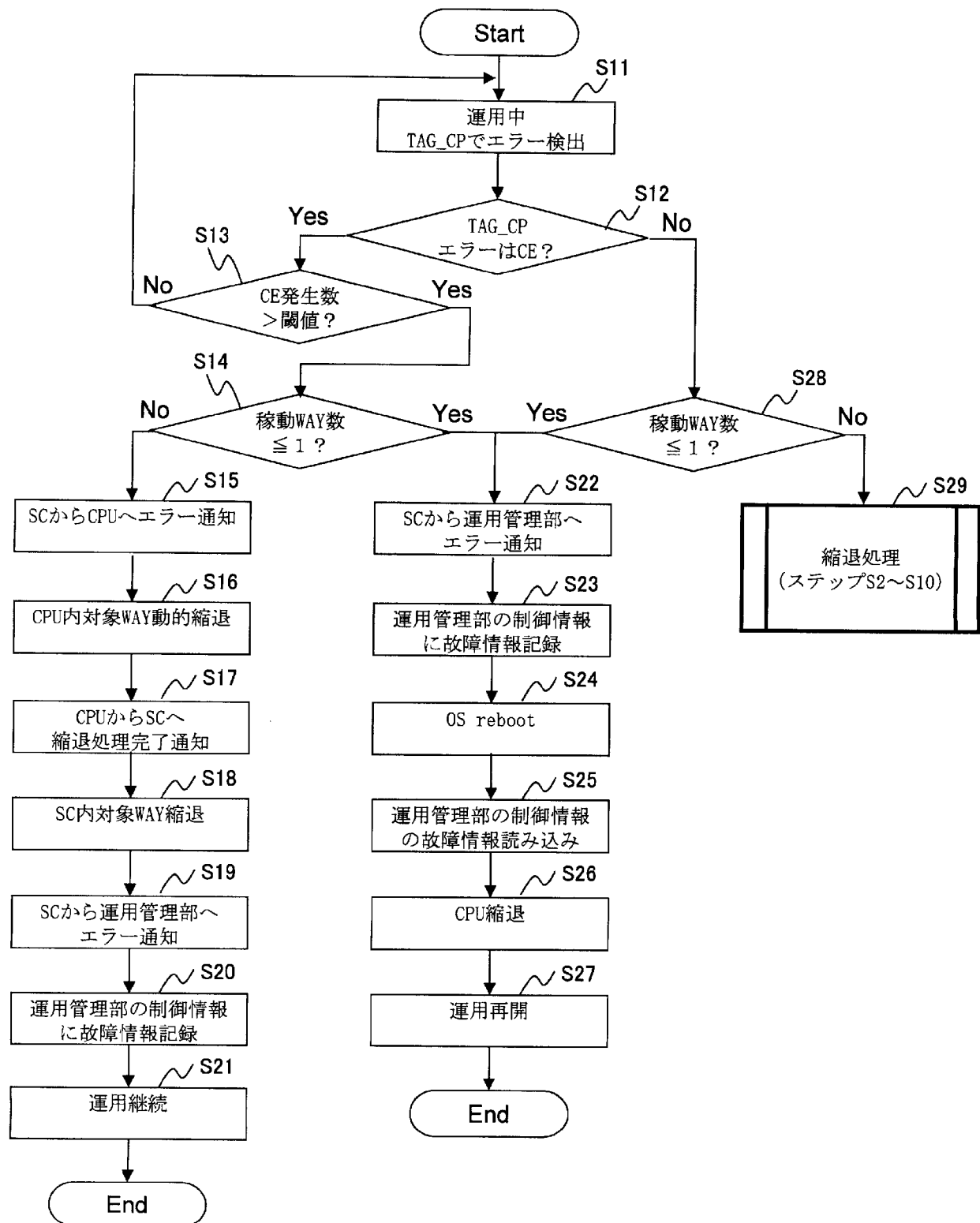
[図7]

LINE	WAY 0	WAY 1		WAY 11		
INDEX	登録アドレス	ステータス	登録アドレス	ステータス		
PA[18:8]	PA[41:19]	STS[7:0]	ECC[6:0]	PA[41:19]	STS[7:0]	ECC[6:0]
* 000000000000	A0			B0		
* 000000000001	A1			B1		
* 000000000010						
.....						
.....						
* 111111111111	A2047			B2047		

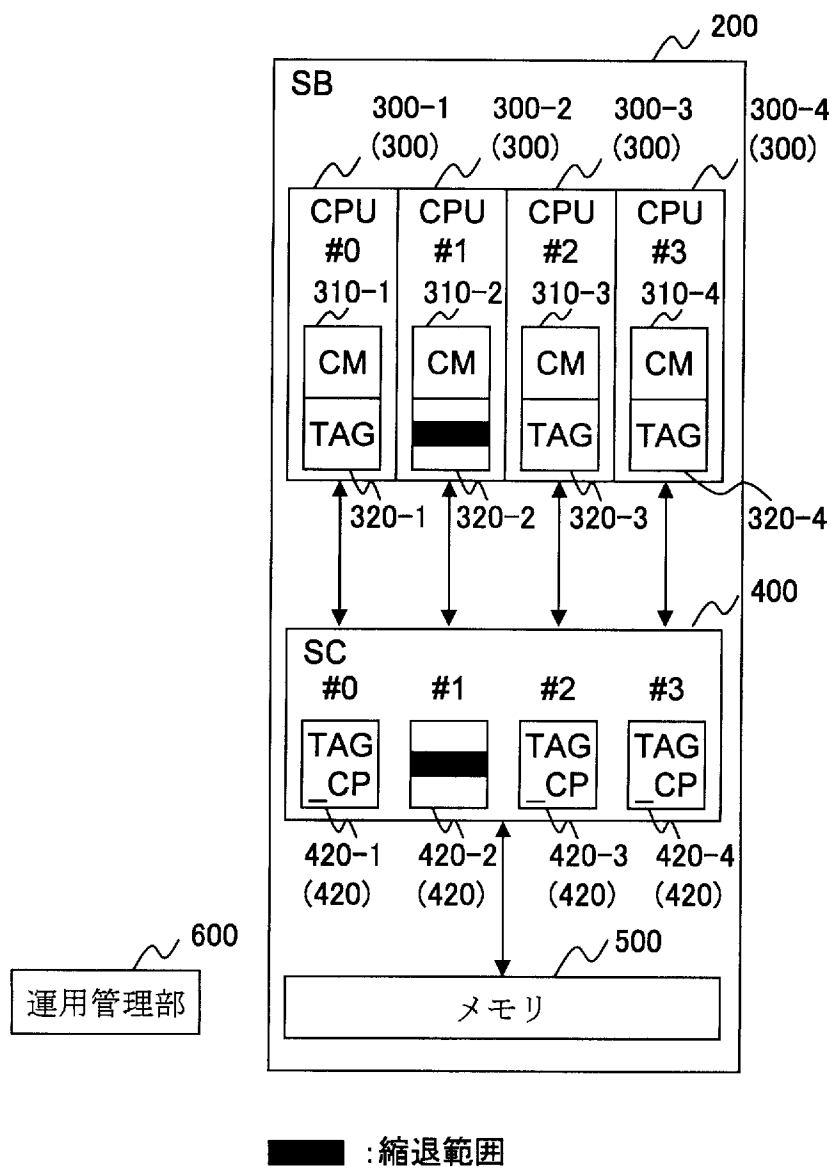
[図8]



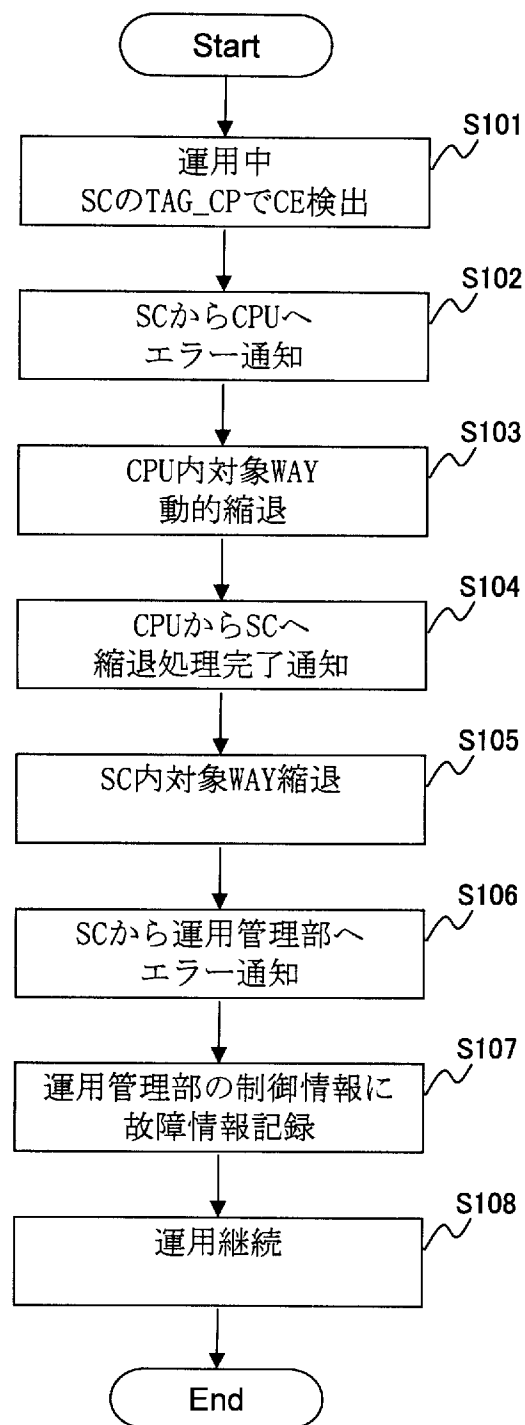
[図9]



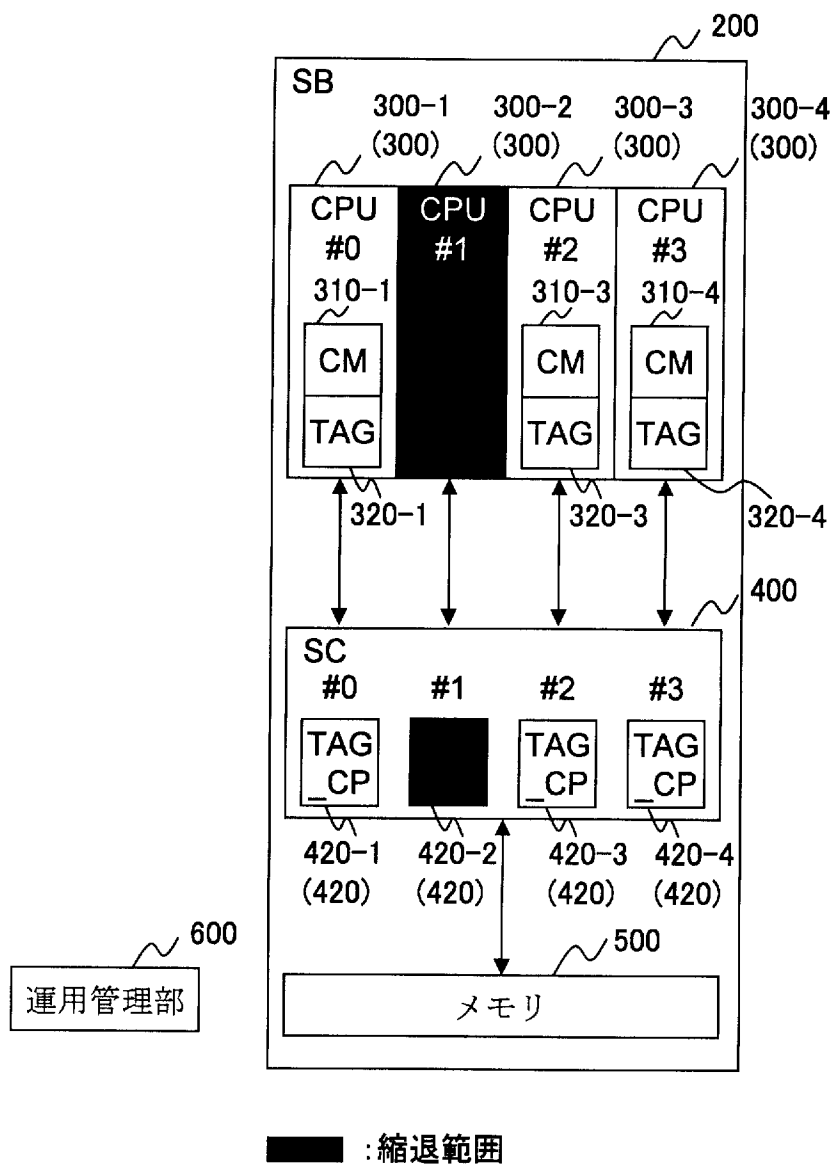
[図10]



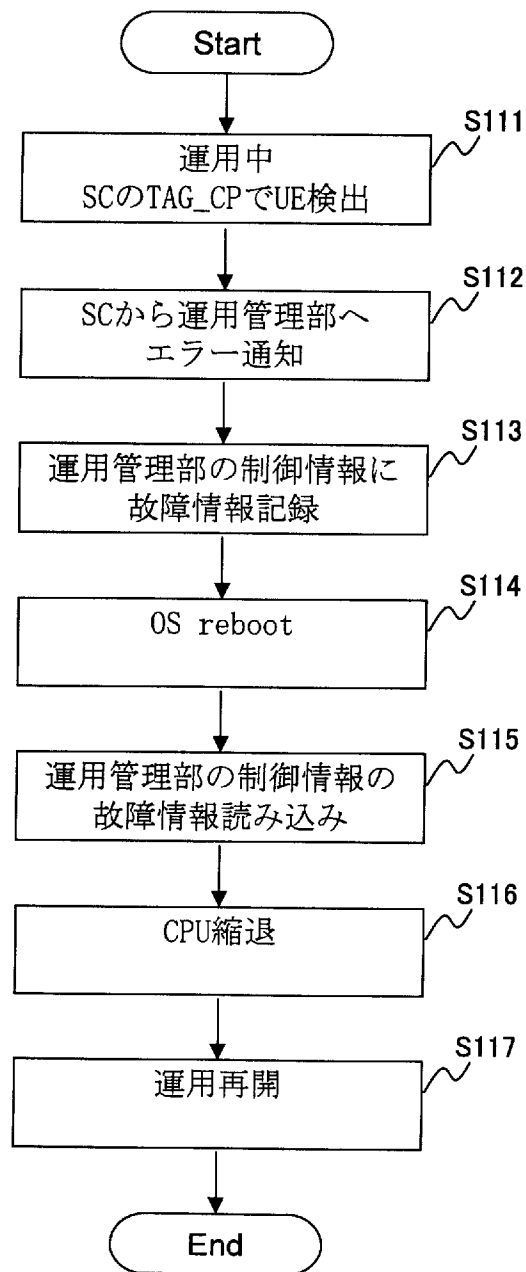
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2011/055488

A. CLASSIFICATION OF SUBJECT MATTER

G06F12/08 (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)

G06F12/08

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)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	WO 2007/097019 A1 (Fujitsu Ltd.), 30 August 2007 (30.08.2007), paragraphs [0033] to [0048] & US 2008/0282037 A1	1-7, 9, 11-17, 19 8, 10, 18, 20
Y A	JP 2008-046701 A (NEC Computertechno, Ltd.), 28 February 2008 (28.02.2008), paragraphs [0098] to [0111] (Family: none)	1-7, 9, 11-17, 19 8, 10, 18, 20

☐ 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 May, 2011 (12.05.11)

Date of mailing of the international search report
24 May, 2011 (24.05.11)

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. G06F12/08 (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. G06F12/08

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 1 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 1 年
日本国登録実用新案公報	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）

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
Y A	WO 2007/097019 A1（富士通株式会社）2007.08.30, 段落 0033-0048 & US 2008/0282037 A1	1-7, 9, 11-17, 19 8, 10, 18, 20
Y A	JP 2008-046701 A（エヌイーシーコンピュータテクノ株式会社） 2008.02.28, 段落 0098-0111（ファミリーなし）	1-7, 9, 11-17, 19 8, 10, 18, 20

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 5 . 2 0 1 1

国際調査報告の発送日

2 4 . 0 5 . 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）

浜岸 広明

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9

5 M

9 8 4 5