(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2005-293478
(P2005-293478A)
(43)公開日　平成17年10月20日(2005.10.20)

(51)Int.Cl.[7]	F I		テーマコード（参考）
G06F 12/00	G06F 12/00	545A	5B005
G06F 3/06	G06F 12/00	514E	5B014
G06F 12/08	G06F 3/06	301S	5B065
G06F 13/10	G06F 3/06	302A	5B082
	G06F 12/08	501E	

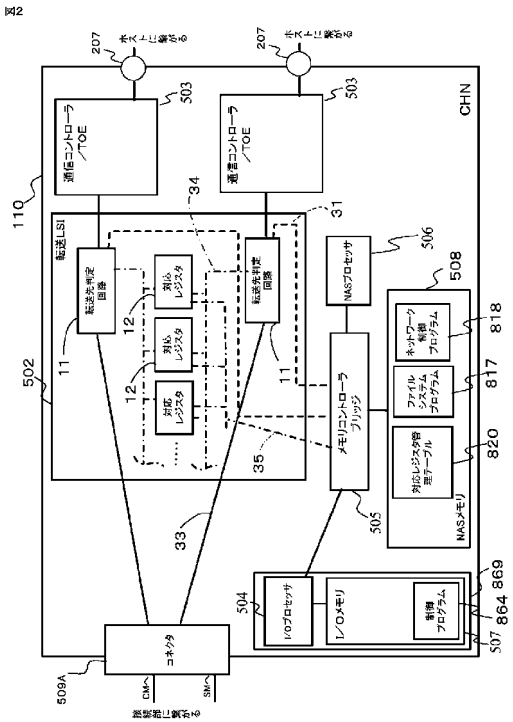
審査請求 未請求　請求項の数 7　ＯＬ　（全 21 頁）　最終頁に続く

(21)出願番号　特願2004-111096 (P2004-111096)
(22)出願日　平成16年4月5日 (2004.4.5)

(71)出願人　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　100095371
弁理士　上村　輝之
(74)代理人　100089277
弁理士　宮川　長夫
(74)代理人　100104891
弁理士　中村　猛
(72)発明者　金井　宏樹
神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
(72)発明者　飯田　純一
神奈川県小田原市中里３２２番２号　株式会社日立製作所ＲＡＩＤシステム事業部内
最終頁に続く

(54)【発明の名称】記憶制御システム、記憶制御システムに備えられるチャネル制御装置、データ転送装置

(57)【要約】
【課題】　上位装置と下位装置との間のデータ転送を高速化する。
【解決手段】　ＣＨＮ１１０は、ホスト端末２００から受信したデータを記憶することができるＮＡＳメモリ５０８と、転送先判定回路１１とを備える。転送先判定回路１１は、ホスト端末２００からアクセス要求を受信した場合、そのアクセス要求が、ホスト端末２００とキャッシュ領域１３１との間で行われるデータ転送を伴うデータアクセス要求であり、データアクセス要求に含まれるファイルシステムアドレスに対応付けられたキャッシュアドレスが識別されたならば、そのキャッシュアドレスが表すキャッシュ領域１３１内のリード対象データを、ＮＡＳメモリ５０８を経由することなくダイレクトにホスト端末２００に転送するか、又は、アクセス要求に含まれるライト対象データを、ＮＡＳメモリ５０８を経由することなくキャッシュ領域１３１に転送する。
【選択図】図２



【特許請求の範囲】
【請求項１】
データを記憶する記憶デバイスと、
外部装置と前記記憶デバイスとの間でやり取りされるデータが格納されるキャッシュ領域を有するキャッシュメモリと、
外部装置からデータを受信して前記キャッシュ領域に格納したり、前記キャッシュ領域に格納されているデータを読み出して前記外部装置に転送したりするチャネル制御部と、
前記キャッシュ領域に格納されているデータを読出して前記記憶デバイスに格納したり、前記記憶デバイスから読出したデータを前記キャッシュ領域に格納したりする記憶デバイス制御部と
を備え、
前記チャネル制御部は、
前記外部装置から受信したデータを記憶することができるデータ記憶メモリと、
データファイル中のデータのファイルレベルのアドレスであるファイルシステムアドレスと、前記ファイルシステムアドレスに対応付けられたキャッシュアドレスとを記憶する対応メモリと、
前記外部装置から受信したアクセス要求が、ファイルシステムアドレスを有するファイルアクセス要求である場合、前記ファイルアクセス要求に含まれるファイルシステムアドレスに対応付けられたキャッシュアドレスが前記対応メモリから識別されたならば、前記識別されたキャッシュアドレスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記外部装置に転送する、又は、前記ファイルアクセス要求が有するアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する転送先判定回路と
を備える、
記憶制御システム。
【請求項２】
前記チャネル制御部は、前記ファイルシステムアドレスを、前記記憶デバイスの記憶領域の管理単位であるブロックレベルのアドレスに変換する第１のプロセッサを更に備え、
前記チャネル制御部及び前記記憶デバイス制御部の少なくとも一方が、前記変換されたブロックレベルのアドレスに基づいて、前記キャッシュメモリ上に前記キャッシュ領域を確保する第２のプロセッサを備え、
前記第１のプロセッサが、前記ファイルアクセス要求が発行されることに先行して前記外部装置から先行コマンドを受信し、前記先行コマンドに応答して、アクセス対象となるファイルシステムアドレスを前記対応メモリに書込み、前記ファイルシステムアドレスを前記ブロックレベルのアドレスに変換し、
前記第２のプロセッサが、前記第１のプロセッサによって変換された前記ブロックレベルのアドレスに基づいて、前記キャッシュ領域を確保し、前記確保したキャッシュ領域を表すキャッシュアドレスを、前記ファイルシステムアドレスが書き込まれた前記対応メモリに書込む、
請求項１記載の記憶制御システム。
【請求項３】
前記チャネル制御部は、前記転送先判定回路よりも前記外部装置側に、前記外部装置から受信するアクセス要求についてインターネットプロトコルの解釈を行うＩＰ回路と、前記外部装置から受信するアクセス要求についてトランスミッションコントロールプロトコルの解釈を行うＴＣＰ回路とを備え、前記外部装置から前記ＩＰ回路及び前記ＴＣＰ回路を経由して前記アクセス要求が前記転送先判定回路に入力されるようになっている、
請求項１記載の記憶制御システム。
【請求項４】
前記対応メモリは、複数のレジスタであり、
前記複数のレジスタの各々には、１つのファイルアクセス要求に含まれる１つのファイ

ルシステムアドレスと、そのファイルシステムアドレスに対応付けられた１又は複数のキャッシュアドレスとが記憶される、
請求項１記載の記憶制御システム。
【請求項５】
データを記憶する記憶デバイスと、外部装置と前記記憶デバイスとの間でやり取りされるデータが格納されるキャッシュ領域を有するキャッシュメモリと、外部装置からデータを受信して前記キャッシュ領域に格納したり、前記キャッシュ領域に格納されているデータを読み出して前記外部装置に転送したりするチャネル制御装置と、前記キャッシュ領域に格納されているデータを読出して前記記憶デバイスに格納したり、前記記憶デバイスから読出したデータを前記キャッシュ領域に格納したりする記憶デバイス制御部とを備える記憶制御システムに搭載される前記チャネル制御装置であって、
前記外部装置から受信したデータを記憶することができるデータ記憶メモリと、
前記外部装置からアクセス要求を受信した場合、前記アクセス要求が、前記外部装置と前記キャッシュ領域との間で行われるデータ転送を伴うデータアクセス要求であり、前記データアクセス要求に含まれる情報要素に対応付けられたキャッシュアドレスが識別されたならば、前記識別されたキャッシュアドレスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記外部装置に転送する、又は、前記アクセス要求に含まれるアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する転送先判定回路と
を備えるチャネル制御装置。
【請求項６】
データファイル中のデータのファイルレベルのアドレスであるファイルシステムアドレスと、前記ファイルシステムアドレスに対応付けられたキャッシュアドレスとを記憶する対応メモリを更に備え、
前記転送先判定回路は、前記外部装置から受信したアクセス要求が、ファイルシステムアドレスを有するファイルアクセス要求である場合、前記ファイルアクセス要求に含まれるファイルシステムアドレスに対応付けられたキャッシュアドレスが前記対応メモリから識別されたならば、前記識別されたキャッシュアドレスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記ホスト端末に転送する、又は、前記ファイルアクセス要求が有するアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する、
請求項５記載のチャネル制御装置。
【請求項７】
通信装置に搭載可能なデータ転送装置において、
前記データ転送装置の上位に存在する外部装置から受信したアクセス要求が、前記外部装置と、前記データ転送装置の下位に存在する下位記憶装置の記憶領域との間で行われるデータ転送を伴うデータアクセス要求であり、前記データアクセス要求に含まれる情報要素に対応付けられた記憶アドレスが識別されたならば、前記識別された記憶アドレスが表す記憶領域であるターゲット記憶領域内のアクセス対象データを、前記通信装置内のデータ記憶メモリを経由することなく前記外部装置に転送する、又は、前記アクセス要求に含まれるアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲット記憶領域に転送する転送先判定回路を備えたデータ転送装置。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、上位装置と下位装置との間でのデータ転送のための技術に関わり、例えば、ＲＡＩＤシステムのような記憶制御システムや、そのような記憶制御システムに搭載可能なチャネル制御装置に関する。

【背景技術】
【０００２】
例えば、データセンタ等のような大規模なデータを取り扱うデータベースシステムでは、ホストコンピュータとは別に構成された記憶制御システムを用いてデータを管理する。この記憶制御システムは、例えば、多数の記憶デバイスをアレイ状に配設して構成されたRAID（Redundant Array of Independent Inexpensive Disks）のようなディスクアレイシステムである。
【０００３】
このような記憶制御システムには、例えば、特開２００３－３１６７１３号公報に開示されているように、ファイル単位のＩ／Ｏ要求を処理するネットワークチャネルアダプタ（以下、ＣＨＮ）を備えることにより、ＮＡＳ（Network Area Storage）となることができるものがある。また、この記憶制御システムは、ディスクと、ディスクアダプタ（ＤＫＡ）と、ＣＨＮとＤＫＡとに共有されるデータであって、ディスクに保存されるデータが格納されるキャッシュメモリとが備えられる。
【０００４】
【特許文献１】特開２００３－３１６７１３号公報。
【発明の開示】
【発明が解決しようとする課題】
【０００５】
上述した公報によれば、ＣＨＮにはメモリが搭載されており、そのメモリには、ファイルデータをキャッシュするため等に用いられるデータバッファが存在する。従って、ＮＡＳクライアントとキャッシュメモリとの間のデータ転送は、そのメモリ内のデータバッファを経由して行われることになる。このデータ転送を高速化することができれば、より利便性が高まると考えられる。
【０００６】
従って、本発明の目的は、上位装置と下位装置との間のデータ転送を高速化することにある。
【０００７】
本発明の更なる目的は、後の記載から明らかになるであろう。
【課題を解決するための手段】
【０００８】
本発明の第１の観点に従う記憶制御システムは、データを記憶する記憶デバイスと、外部装置と前記記憶デバイスとの間でやり取りされるデータが格納されるキャッシュ領域を有するキャッシュメモリと、外部装置からデータを受信して前記キャッシュ領域に格納したり、前記キャッシュ領域に格納されているデータを読み出して前記外部装置に転送したりするチャネル制御部と、前記キャッシュ領域に格納されているデータを読出して前記記憶デバイスに格納したり、前記記憶デバイスから読出したデータを前記キャッシュ領域に格納したりする記憶デバイス制御部とを備える。前記チャネル制御部は、前記外部装置から受信したデータを記憶することができるデータ記憶メモリと、データファイル中のデータのファイルレベルのアドレスであるファイルシステムアドレスと、前記ファイルシステムアドレスに対応付けられたキャッシュアドレスとを記憶する対応メモリと、転送先判定回路を備える。転送先判定回路は、前記外部装置から受信したアクセス要求が、ファイルシステムアドレスを有するファイルアクセス要求である場合、前記ファイルアクセス要求に含まれるファイルシステムアドレスに対応付けられたキャッシュアドレスが前記対応メモリから識別されたならば、前記識別されたキャッシュアドレスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記外部装置に転送する、又は、前記ファイルアクセス要求が有するアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する。
【０００９】

本発明の第１の観点に従う記憶制御システムの第１の実施態様では、前記チャネル制御部は、前記ファイルシステムアドレスを、前記記憶デバイスの記憶領域の管理単位であるブロックレベルのアドレスに変換する第１のプロセッサを更に備える。前記チャネル制御部及び前記記憶デバイス制御部の少なくとも一方が、前記変換されたブロックレベルのアドレスに基づいて、前記キャッシュメモリ上に前記キャッシュ領域を確保する第２のプロセッサを備える。前記第１のプロセッサが、前記ファイルアクセス要求が発行されることに先行して前記外部装置から先行コマンドを受信し、前記先行コマンドに応答して、アクセス対象となるファイルシステムアドレスを前記対応メモリに書込み、前記ファイルシステムアドレスを前記ブロックレベルのアドレスに変換する。前記第２のプロセッサが、前記第１のプロセッサによって変換された前記ブロックレベルのアドレスに基づいて、前記キャッシュ領域を確保し、前記確保したキャッシュ領域を表すキャッシュアドレスを、前記ファイルシステムアドレスが書き込まれた前記対応メモリに書込む。

【００１０】

本発明の第２の観点に従う記憶制御システムの第２の実施態様では、前記チャネル制御部は、前記転送先判定回路よりも前記外部装置側に、前記外部装置から受信するアクセス要求についてインターネットプロトコルの解釈を行うＩＰ回路と、前記外部装置から受信するアクセス要求についてトランスミッションコントロールプロトコルの解釈を行うＴＣＰ回路とを備える。それにより、前記外部装置から前記ＩＰ回路及び前記ＴＣＰ回路を経由して前記アクセス要求が前記転送先判定回路に入力されるようになっている。

【００１１】

本発明の第１の観点に従う記憶制御システムの第３の実施態様では、前記対応メモリは、複数のレジスタである。前記複数のレジスタの各々には、１つのファイルアクセス要求に含まれる１つのファイルシステムアドレスと、そのファイルシステムアドレスに対応付けられた１又は複数のキャッシュアドレスとが記憶される。

【００１２】

本発明の第２の観点に従うチャネル制御装置は、以下の記憶制御システム、すなわち、データを記憶する記憶デバイスと、外部装置と前記記憶デバイスとの間でやり取りされるデータが格納されるキャッシュ領域を有するキャッシュメモリと、外部装置からデータを受信して前記キャッシュ領域に格納したり、前記キャッシュ領域に格納されているデータを読み出して前記外部装置に転送したりするチャネル制御装置と、前記キャッシュ領域に格納されているデータを読出して前記記憶デバイスに格納したり、前記記憶デバイスから読出したデータを前記キャッシュ領域に格納したりする記憶デバイス制御部とを備える記憶制御システムに搭載される前記チャネル制御装置である。チャネル制御装置は、前記外部装置から受信したデータを記憶することができるデータ記憶メモリと、転送先判定回路とを備える。転送先判定回路は、前記外部装置からアクセス要求を受信した場合、前記アクセス要求が、前記外部装置と前記キャッシュ領域との間で行われるデータ転送を伴うデータアクセス要求であり、前記データアクセス要求に含まれる情報要素に対応付けられたキャッシュアドレスが識別されたならば、前記識別されたキャッシュアドレスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記外部装置に転送する、又は、前記アクセス要求に含まれるアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する。

【００１３】

本発明の第２の観点に従うチャネル制御装置の第１の実施態様では、チャネル制御装置は、データファイル中のデータのファイルレベルのアドレスであるファイルシステムアドレスと、前記ファイルシステムアドレスに対応付けられたキャッシュアドレスとを記憶する対応メモリを更に備える。その場合、前記転送先判定回路は、前記外部装置から受信したアクセス要求が、ファイルシステムアドレスを有するファイルアクセス要求である場合、前記ファイルアクセス要求に含まれるファイルシステムアドレスに対応付けられたキャッシュアドレスが前記対応メモリから識別されたならば、前記識別されたキャッシュアド

レスが表すキャッシュ領域であるターゲットキャッシュ領域内のアクセス対象データを、前記データ記憶メモリを経由することなく前記ホスト端末に転送する、又は、前記ファイルアクセス要求が有するアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲットキャッシュ領域に転送する。

【００１４】

本発明の第３の観点に従うデータ転送装置は、通信装置に搭載可能なデータ転送装置であり、転送先判定回路を備える。転送先判定回路は、前記データ転送装置の上位に存在する外部装置から受信したアクセス要求が、前記外部装置と、前記データ転送装置の下位に存在する下位記憶装置の記憶領域との間で行われるデータ転送を伴うデータアクセス要求であり、前記データアクセス要求に含まれる情報要素に対応付けられた記憶アドレスが識別されたならば、前記識別された記憶アドレスが表す記憶領域であるターゲット記憶領域内のアクセス対象データを、前記通信装置内のデータ記憶メモリを経由することなく前記外部装置に転送する、又は、前記アクセス要求に含まれるアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲット記憶領域に転送する。

【００１５】

本発明の第４の観点に従うデータ転送方法は、第１と第２のステップを有する。前記データ転送方法は、第１のステップでは、上位装置からアクセス要求を受信する。また、前記データ転送方法は、前記第２のステップでは、前記受信したアクセス要求が、前記上位装置と、下位装置の記憶領域との間で行われるデータ転送を伴うデータアクセス要求であり、前記データアクセス要求に含まれる情報要素に対応付けられた記憶アドレスが識別されたならば、前記識別された記憶アドレスが表す記憶領域であるターゲット記憶領域内のアクセス対象データを、通信装置内のデータ記憶メモリを経由することなく前記外部装置に転送するか、又は、前記アクセス要求に含まれるアクセス対象データを、前記データ記憶メモリを経由することなく前記ターゲット記憶領域に転送する。

【発明の効果】

【００１６】

本発明によれば、上位装置（例えば外部装置）と下位装置（例えばキャッシュメモリ）との間のデータ転送を高速化することができる。

【発明を実施するための最良の形態】

【００１７】

以下、図面を参照して、本発明の一実施形態について説明する。

【００１８】

図１は、本実施形態に係る記憶制御システムを備えるコンピュータシステムの構成を示すブロック図である。

【００１９】

このコンピュータシステム１では、通信ネットワーク８２０に、１以上のホスト端末２００と、記憶制御システム６００とが接続されている。通信ネットワーク８２０は、ファイルレベルのデータがやり取りされる通信ネットワークであり、例えば、ＬＡＮ、インターネット、専用回線、公衆回線等を場合に応じて適宜用いることができる。

【００２０】

１以上のホスト端末２００の各々は、例えば、ＣＰＵ（Central Processing Unit）やメモリ等の情報処理資源を備えたコンピュータ装置であり、例えば、パーソナルコンピュータ、ワークステーション、メインフレーム等として構成される。ホスト端末２００は、例えば、キーボードスイッチやポインティングデバイス、マイクロフォン等の情報入力装置（図示せず）と、例えば、モニタディスプレイやスピーカー等の情報出力装置（図示せず）とを備えている。さらに、ホスト端末２００は、例えば、ＮＡＳを利用するためのＮＡＳ利用ソフトウェア２００Ｂと、Windows（登録商標）又はＵＮＩＸ（登録商標）等のＯＳ（オペレーティングシステム）２００Ｃと、ネットワークドライバ２００Ｆとを備える。ホスト端末２００は、例えば、ファイル名を指定してファイル単位でのデータ入出力を記憶制御システム６００に要求する。ＮＡＳ用ソフトウェア２００Ｂとしては、例えば

、ＯＳ２００ＣがＵＮＩＸ（登録商標）の場合、ＮＦＳ（Network File System）であり、ＯＳ２００ＣがＷｉｎｄｏｗｓ（登録商標）の場合、ＣＩＦＳ（Common Interface File System）である。また、ネットワークドライバ２００Ｆには、例えば、ＴＣＰ（Transmission Control Protocol）に基づくデータ処理を行うＴＣＰドライバ２００Ｄと、ＩＰ（Internet Protocol）に基づくデータ処理を行うＩＰドライバ２００Ｅとが含まれる。

【００２１】

記憶制御システム６００は、例えば、アレイ状に配列された多数の物理記憶デバイス３００を備えるＲＡＩＤシステムである。記憶制御システム６００は、記憶制御装置１００と記憶装置ユニット１０１とに大別することができる。記憶制御装置１００は、例えば、複数のチャネルアダプタＮＡＳ（以下、「ＣＨＮ」と略記）１１０と、複数のディスクアダプタ（以下、ＤＫＡ）１４０と、キャッシュメモリ１３０と、共有メモリ１２０と、接続部１５０とを備えている。

【００２２】

ＣＨＮ１１０は、ホスト端末２００との間のデータ通信を行うものである。ＣＨＮ１１０は、例えば、ＣＰＵやメモリ等を備えたマイクロコンピュータシステムとして構成されており、ホスト端末２００から受信した各種コマンドを解釈して実行する。ＣＨＮ１１０には、自分を識別するためのネットワークアドレス（例えば、ＩＰアドレスやＷＷＮ）が割り当てられている。ＣＨＮ１１０は、ホスト端末２００から通信ネットワーク８２０を介してファイル単位でのＩ／Ｏ要求（例えば、ファイル名と、そのファイル名を持つファイルをリード又はライトする命令とを含んだコマンド、以下、「ファイルＩ／Ｏ要求」と言う）を受けて、そのファイルＩ／Ｏ要求を処理するＮＡＳ（Network Attached Storage）として振る舞うことができるようになっている。ＣＨＮ１１０は、ホスト端末２００から受信したデータをキャッシュメモリ１３０に格納したり、ＤＫＡ１４０によってキャッシュメモリ１３０に格納されたデータを取得してホスト端末２００に送信したりする。

【００２３】

ＤＫＡ１４０は、記憶装置ユニット１０１内の論理的な記憶ユニット（以下、ＬＵ）３１０との間のデータ授受を行うものである。ＤＫＡ１４０は、ＬＵ３１０を備える物理記憶デバイス３００に接続するための図示しない通信ポートを備えている。また、ＤＫＡ１４０は、ＣＰＵやメモリ等を備えたマイクロコンピュータシステムとして構成されている。ＤＫＡ１４０は、ＣＨＮ１１０からキャッシュメモリ１３０に書き込まれたデータを取得してＬＵ３１０に書込んだり、また、ＬＵ３１０から読み出したデータをキャッシュメモリ１３０に格納したりする。ＤＫＡ１４０は、ＬＵ３１０との間でデータ入出力を行う場合、論理的なアドレスを物理的なアドレスに変換する。

【００２４】

キャッシュメモリ（以下、「ＣＭ」と略記する場合有り）１３０は、例えば揮発性又は不揮発性のメモリであり、ホスト端末２００から受信してＬＭ３１０へ書き込まれるデータや、ＬＵ３１０から読出されてホスト端末２００へ転送されるデータを一時的に記憶するものである。

【００２５】

共有メモリ（以下、「ＳＭ」と略記する場合有り）１２０は、例えば不揮発性のメモリであり、ホスト端末との間でやり取りされるデータに関する制御情報（例えば、ＣＭ１３０上に確保されたどのキャッシュ領域にどのデータが格納されるべきかを示す情報）等が格納される。また、共有メモリ１２０には、例えば、ワーク領域（例えば、ＣＨＮ１１０及びＤＫＡ１４０のＣＰＵ間でやり取りされるメッセージを一時的に記憶する領域）が設定される。なお、図示の例では、ＣＭ１３０とＳＭ１２０は、物理的に分離しているが、一つのメモリであっても良い。その場合、そのメモリ上のメモリ空間が、論理的に、ＣＭ用の空間とＳＭ用の空間とに分けられても良い。

【００２６】

接続部１５０は、各ＣＨＮ１１０、各ＤＫＡ１４０、キャッシュメモリ１３０及び共有

メモリ１２０を相互に接続させる。接続部１５０は、例えば、高速スイッチング動作によってデータ伝送を行う超高速クロスバスイッチ等のような高速バスとして構成することができる。

【００２７】

記憶装置ユニット１０１には、アレイ状に配列された複数の物理記憶デバイス３００が含まれている。物理記憶デバイス３００としては、例えば、ハードディスク、フレキシブルディスク、磁気テープ、半導体メモリ、光ディスク等のようなデバイスを用いることができる。物理記憶デバイス３００の記憶領域上には、論理的な記憶デバイスである複数の論理ユニット（以下、「ＬＵ」と略記）３１０が備えられている。

【００２８】

以下、本実施形態についてより詳細に説明する。なお、以下の説明では、ＮＡＳ用ソフトウェア２００Ｂは、ＮＦＳであるとする。

【００２９】

図２は、ＣＨＮ１１０の構成例を示すブロック図である。

【００３０】

ＣＨＮ１１０は、複数（又は１つ）の通信ポート２０７と、２以上（又は１つ）の通信コントローラ／ＴＯＥ５０３と、Ｉ／Ｏプロセッサ５０４及びＩ／Ｏメモリ５０７を備えた１つ（又は複数）の入出力制御部８６９と、メモリコントローラブリッジ５０５と、ＮＡＳプロセッサ５０６と、ＮＡＳメモリ５０８と、コネクタ５０９と、転送ＬＳＩ５０２とを備えている。

【００３１】

通信コントローラ／ＴＯＥ５０３は、通信コントローラとＴＯＥ（TCP/IPオフロードエンジン）とが一体になったハードウェア回路であり、通信ポート２０７及び転送ＬＳＩ５０２に接続される。通信コントローラは、例えばＬＡＮコントローラであり、ＩＰに基づくデータ処理を行う。ＴＯＥは、ＴＣＰに基づくデータ処理を行う。なお、通信コントローラとＴＯＥは、物理的に分離していても良い。

【００３２】

メモリコントローラブリッジ５０５は、ＮＡＳプロセッサ５０６、ＮＡＳメモリ５０８、Ｉ／Ｏプロセッサ５０４及び転送ＬＳＩ５０２に接続されており、それらの間の通信を中継するＬＳＩ（Large-Scale Integrated circuit）である。メモリコントローラブリッジ５０５には、後述するメモリブリッジ転送バス３１及びプロセッサレジスタバス３５が接続される。

【００３３】

ＮＡＳメモリ５０８は、ＮＡＳプロセッサ５０６の制御を司るプログラムを記憶することができる。ＮＡＳメモリ５０８は、例えば、ファイルシステムプログラム８１７、ネットワーク制御プログラム８１８、対応レジスタ管理テーブル８２０等を記憶することができる。ファイルシステムプログラム８１７は、例えば、ファイルＩ／Ｏ要求に含まれているファイル名と、そのファイル名を有するファイルが格納されている場所のアドレス情報（例えばＬＵＮ及び先頭論理ブロックアドレス）との対応づけを管理し、その対応付けに基づいて、ファイルＩ／Ｏ要求をブロックＩ／Ｏ要求に変換する。ネットワーク制御プログラム８１８は、例えば、ＮＦＳ（Network File System）とＳａｍｂａの２つのファイルシステムプロトコルを含んで構成される。ＮＦＳは、ＮＦＳが動作するＵＮＩＸ（登録商標）オペレーティングシステムを搭載したホスト端末からのファイルＩ／Ｏ要求を受け付ける。一方、Ｓａｍｂａは、ＣＩＦＳ（Common Interface File System）が動作するＷｉｎｄｏｗｓ（登録商標）オペレーティングシステムを搭載したホスト端末らのファイルＩ／Ｏ要求を受け付ける。対応レジスタ管理テーブル８２０には、複数の対応レジスタ１２にそれぞれ対応した複数のレジスタＩＤ及び複数の使用状態データが登録されている。

【００３４】

ＮＡＳプロセッサ５０６は、例えばＣＰＵ又はマイクロプロセッサであり、メモリコン

トローラブリッジ５０５に接続されている。ＮＡＳプロセッサ５０６は、ＮＡＳメモリ５０８に格納されているファイルシステムプログラム８１７及びネットワーク制御プログラム８１８等を読出し、読み出したコンピュータプログラムに従う処理を実行することができる。例えば、ＮＡＳプロセッサ５０６は、ネットワーク制御プログラム８１８により、ホスト端末２００からのファイルＩ／Ｏ要求を受け付ける。また、ＮＡＳプロセッサ５０６は、ファイルシステムプログラム８１７により、ホスト端末２００から受信しＮＡＳメモリ５０８に格納されたファイルＩ／Ｏ要求をブロックＩ／Ｏ要求に変換してＩ／Ｏプロセッサ５０４に出力することができる（なお、ブロックとは、ＬＵ３１０上の記憶領域におけるデータの管理単位である）。また、ＮＡＳプロセッサ５０６は、データ転送を伴うアクセス要求がホスト端末２００から出力される場合、それに先行して、所定の先行コマンドをホスト端末２００から受ける。その場合、ＮＡＳプロセッサ５０６は、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、複数の対応レジスタ１２の中から選択された対応レジスタ１２に、ファイルシステムアドレスを登録する。また、ＮＡＳプロセッサ５０６は、上記選択された対応レジスタ１２のレジスタＩＤを、Ｉ／Ｏプロセッサ５０４に通知する。

【００３５】

Ｉ／Ｏプロセッサ５０４は、例えばＣＰＵ又はマイクロプロセッサであり、Ｉ／Ｏメモリ５０７から読み出した制御プログラム８６４により、接続部１５０との間のデータの授受や、ＮＡＳプロセッサ５０６と接続部１５０との間のデータ通信の中継や、キャッシュメモリ１３０上のキャッシュアドレスの管理等を実行することができる。また、Ｉ／Ｏプロセッサ５０４は、ＮＡＳプロセッサ５０６からのブロックＩ／Ｏ要求に応答して、キャッシュメモリ１３０上にキャッシュ領域を確保すると共に、確保されたキャッシュ領域を表すキャッシュアドレスを、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、ＮＡＳプロセッサ５０６から通知されたレジスタＩＤに対応した対応レジスタ１２に登録することができる。

【００３６】

Ｉ／Ｏメモリ５０７は、Ｉ／Ｏプロセッサ５０４の制御を司るコンピュータプログラム等を格納する。

【００３７】

コネクタ５０９は、接続部１５０に接続される。具体的には、例えば、コネクタ５０９は、接続部１５０に含まれるＣＭ転送経路（ＣＭ１３０に接続された転送経路）及びＳＭ転送経路（ＳＭ１２０に接続されて転送経路）に接続される。なお、必ずしも、一つのコネクタ５０９に、ＣＭ転送経路とＳＭ転送経路とが混在する必要はない。例えば、コネクタ５０９の代わりに、ＣＭ転送経路に接続される第１のコネクタと、ＳＭ転送経路に接続される第２のコネクタとが備えられても良い。

【００３８】

転送ＬＳＩ５０２は、例えば、通信コントローラ／ＴＯＥ５０３、メモリコントローラブリッジ５０５、キャッシュメモリ１３０及び共有メモリ１２０の相互の通信を可能とするためのＬＳＩ（Large-Scale Integrated circuit）である。転送ＬＳＩ５０２は、複数（又は１つ）の対応レジスタ１２と、１以上（例えば通信コントローラ／ＴＯＥ５０３と同数）の転送先判定回路１１とを備えている。

【００３９】

複数の対応先レジスタ１２の各々には、その対応先レジスタ１２とプロセッサ５０４又は５０６との間でやり取りされるデータが経由するプロセッサレジスタバス３５と、後述する判定レジスタバス３４とが接続される。複数の対応レジスタ１２の各々には、後に詳述するように、ファイルシステムアドレスと、そのファイルシステムアドレスに対応付けられたキャッシュアドレスとが登録される。なお、転送先ＬＳＩ５０２に備えられる対応先レジスタ１２の数は、例えば、ＣＨＮ１１０が同時に受けることができるホスト端末２００からのアクセス数と同数である。

【００４０】

転送先判定回路１１は、例えば純粋なハードウェア回路である。転送先判定回路１１には、複数の対応先レジスタ１２に対するアクセス経路となる判定レジスタバス３４と、コネクタ５９０を介してキャッシュメモリ１３０や共有メモリ１２０との間でやり取りされるデータが経由するコネクタ転送バス３３と、メモリコントローラブリッジ５０５との間でやり取りされるデータが経由するメモリブリッジ転送バス３５とが接続される。転送先判定回路１１は、例えば、ホスト端末２００から通信コントローラ／ＴＯＥ５０３を介して受信したアクセス要求が、データ転送を伴うアクセス要求であるか否かを判別し、その判別の結果、データ転送を伴うアクセス要求であり、且つ、複数の対応レジスタ１２の少なくとも１つからアクセス先キャッシュアドレスを認識できる場合には、認識されたアクセス先キャッシュアドレスとの間で、ＮＡＳメモリ５０８を介すること無くダイレクトにデータ転送を行う。具体的には、例えば、転送先判定回路１１は、ホスト端末２００から通信コントローラ／ＴＯＥ５０３を介してリード要求（例えば、「NFSPROC_READ」）を受信した場合、そのリード要求に含まれているファイルシステムアドレスを記憶した対応レジスタ１２が存在するならば、その対応レジスタ１２に記憶されているキャッシュアドレスが示すキャッシュ領域からコネクタ転送バス３３を介してリード対象データを読出し、そのリード対象データを、ＮＡＳメモリ５０８を介することなくダイレクトに通信コントローラ／ＴＯＥ５０３を介してホスト装置２００に転送する。また、例えば、転送先判定回路１１は、ホスト端末２００から通信コントローラ／ＴＯＥ５０３を介してライト要求（例えば、「NFSPROC_WRITE」）を受信した場合、そのライト要求に含まれているファイルシステムアドレスを記憶した対応レジスタ１２が存在するならば、そのライト要求が有するライト対象データを、その対応レジスタ１２に記憶されているキャッシュアドレスが示すキャッシュ領域に、ＮＡＳメモリ５０８を介することなくコネクタ転送バス３３を介してダイレクトに転送する。

【００４１】

以上が、ＣＨＮ１１０の構成例である。なお、上述したプロセッサレジスタバス３５、判定レジスタバス３４及びメモリブリッジ転送バス３１は、論理的なバスを示す。転送ＬＳＩ５０２内の各構成要素とメモリコントローラブリッジ５０５とは、ＰＣＩバス等の物理的な単一バス（図示せず）で互いに接続される（この点は、後述の図７についても同様である）。

【００４２】

図３は、対応先レジスタ１２の構成例を示す。

【００４３】

図３に示すように、対応先レジスタ１２には、ホスト端末２００のＮＦＳ２００Ｂから出力されるファイルシステムアドレスと、そのファイルシステムアドレスに対応付けられた１又は複数のキャッシュアドレスとが登録される。

【００４４】

ファイルシステムアドレスとは、アクセス対象のデータファイル（例えば、ホスト端末２００がＬＵ３１０に対して入出力するユーザファイル）のアドレスに関する情報、例えば、アクセス対象のデータファイル中のどの場所からのどれぐらいの長さのデータのことを指すかを示す情報である。具体的には、例えば、ファイルシステムアドレスには、ファイルハンドルと、オフセットと、データ長とが含まれている。ファイルハンドルは、例えば、アクセス対象のデータファイルの先頭アドレスを表す。オフセットとは、ファイルハンドルからのオフセットのことであり、具体的には、アクセス対象のデータファイルにおける実際のアクセス対象先頭アドレスを表す。データ長とは、アクセス対象のデータファイルにおける上記オフセットからのデータ長のことであり、例えば、図示のように、データファイル３０がＭバイト単位でアクセス（書込み又は読込み）される場合には、データ長はＭを表す。

【００４５】

１又は複数のキャッシュアドレスとは、キャッシュメモリ１３０上の記憶領域を示すアドレスのことであり、例えば、１つのファイルシステムアドレスに基づいてキャッシュメ

モリに書込まれる又は読み出されるデータサイズ分の記憶領域を表す。具体的には、例えば、キャッシュメモリ１３０がキャッシュブロック１３０Ｓ単位で管理されている場合、１つのキャッシュアドレスは１つのキャッシュブロック１３０Ｓを表し、キャッシュアドレスは、１つのファイルシステムアドレスに対してＫ（Ｋ≧１）個登録される。具体的には、例えば、データファイル３０がＭバイト（例えば４０９６バイト）単位でアクセスされ、キャッシュメモリ１３０上の各キャッシュブロック１３０Ｓの領域サイズがＮバイト（例えば５１２バイト）の場合、キャッシュアドレスの数Ｋは、Ｍ÷Ｎの商の値（例えば４０９６／５１２＝８）となる。なお、余りが生じた場合は、商の値に１が加えられてもよい。

【００４６】

図４は、対応レジスタ管理テーブル８２０の構成例を示す。

【００４７】

対応レジスタ管理テーブル８２０には、複数の対応レジスタ１２にそれぞれ対応した複数のレジスタＩＤ及び複数の使用状態データが登録される。

【００４８】

レジスタＩＤとは、対応する対応レジスタの識別情報（例えば番号）である。

【００４９】

使用状態データとは、対応する対応レジスタ１２の使用状態を示すデータであり、例えば、使用中か否かを示す情報（例えば、使用中であれば「１」であり、未使用であれば「０」）である。

【００５０】

以下、本実施形態で行われるデータ転送処理流れの例を説明する。

【００５１】

図５は、ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルリードを要求した場合に行なわれる処理流れを示す。

【００５２】

ホスト端末２００のＮＦＳ２００Ｂは、ユーザからファイルリード要求を受けた場合（ステップＳ１）、要求されたファイル名を有するデータファイルを取得するためにどこにアクセスすれば良いかを要求する先行コマンドを送信する。具体的には、例えば、ホスト端末２００のＮＦＳ２００Ｂは、ファイルハンドラの要求（例えば「NFSPROC_LOOKUP」）を送信する（Ｓ２）。

【００５３】

ＮＡＳプロセッサ５０６は、ホスト端末２００のＮＦＳ２００Ｂから通信コントローラ／ＴＯＥ５０３を介して受けたファイルハンドラ要求に従って、ファイルシステムプログラム８１７に基づくファイル処理を行う（Ｓ３）。そして、ＮＡＳプロセッサ５０６は、そのファイル処理によって取得されたファイルハンドラ（受信したファイルハンドラ要求に対応したファイルハンドラ）を、ホスト端末２００のＮＦＳ２００Ｂに送信する（Ｓ４）。

【００５４】

また、ＮＡＳプロセッサ５０６は、複数の対応レジスタ１２の中から未使用の対応レジスタ１２を探す（Ｓ５）。具体的には、ＮＡＳプロセッサ５０６は、メモリコントローラブリッジ５０５を介して対応レジスタ管理テーブル８２０を参照し、対応レジスタ管理テーブル８２０に記録されている複数の使用状態データの中から未使用を示す使用状態データを探す。

【００５５】

Ｓ５において、未使用の対応レジスタ１２が探し出された場合（Ｓ５でＹ）、ＮＡＳプロセッサ５０６は、探し出された未使用の対応レジスタ１２を確保する（Ｓ６）。具体的には、ＮＡＳプロセッサ５０６は、対応レジスタ管理テーブル８２０において、探し出された対応レジスタ１２に対応する使用状態データを「未使用」から「使用中」に変更する。

【００５６】
また、ＮＡＳプロセッサ５０６は、Ｓ５でＹの場合、上記取得されたファイルハンドラに基づくファイルシステムアドレスを、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、Ｓ６で確保した対応レジスタ１２に書込む（Ｓ７）。
【００５７】
また、ＮＡＳプロセッサ５０６は、Ｓ５でＹの場合、上記取得されたファイルハンドラに基づくファイルシステムアドレスを、ＬＵ３１０の記憶領域の管理単位であるブロックレベルのアドレスに変換し、そのブロックレベルのアドレスを有するリード要求（以下、ブロックリード要求）と、Ｓ６で確保された対応レジスタ１２のレジスタＩＤとを、Ｉ／Ｏプロセッサ５０４に送信する（Ｓ８）。
【００５８】
Ｉ／Ｏプロセッサ５０４は、レジスタＩＤ及びブロックリード要求を受信した場合、ブロックリード要求に基づいて、キャッシュメモリ１３０上に、そのブロックリード要求に従って読み出すデータ（以下、リード対象データ）を格納するためのキャッシュ領域１３１を確保する（Ｓ９）。そして、Ｉ／Ｏプロセッサ５０４は、ＤＫＡ１４０上の図示しないＩ／Ｏプロセッサとプロセッサ間通信を行うことにより、受信したブロックリード要求に含まれるアドレスが表す場所（ＬＵ３１０内の場所）からリード対象データを読み出し、そのリード対象データを、上記確保したキャッシュ領域１３１に格納する（Ｓ１０）。
【００５９】
また、Ｉ／Ｏプロセッサ５０４は、確保したキャッシュ領域１３１を表すＫ個のキャッシュアドレスを、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、上記受信したレジスタＩＤに対応する対応レジスタ１２に書込む（Ｓ１１）。
【００６０】
さて、ホスト端末２００のＮＦＳ２００Ｂは、Ｓ４の処理によってファイルハンドラを受信した場合、受信したファイルハンドラに基づくファイルシステムアドレスを含んだファイルリード要求を、ＣＨＮ１１０に送信する（Ｓ１２）。
【００６１】
ＣＨＮ１１０の転送先判定回路１１は、ファイルシステムアドレスを含んだファイルリード要求を通信コントローラ／ＴＯＥ５０３を介して受信する。受信したファイルリード要求は、通信コントローラ／ＴＯＥ５０３において、既に、ＩＰ及びＴＣＰに基づくデータ処理が行われている。
【００６２】
転送先判定回路１１は、受信したアクセス要求はキャッシュメモリ１３０へのデータ転送を伴うアクセス要求であり、且つ、そのアクセス要求に対応したキャッシュアドレスが複数の対応レジスタ１２のいずれかに記録されているか否かを判断する（Ｓ１３）。具体的には、例えば、転送先判定回路１１は、受信したアクセス要求が「NFSPROC_READ」又は「NFSPROC_WRITE」であり、且つ、そのアクセス要求に含まれるファイルシステムアドレスを記憶する対応レジスタ１２を複数の対応レジスタ１２の中に存在するか否かを判断する。
【００６３】
Ｓ１３において、否定的な判断結果が得られた場合、すなわち、以下の（１）又は（２）の場合、
（１）受信したアクセス要求がキャッシュメモリ１３０へのデータ転送を伴うアクセス要求では無い（例えば「NFSPROC_READ」及び「NFSPROC_WRITE」のいずれでもない）と判断された場合、
（２）受信したアクセス要求がキャッシュメモリ１３０へのデータ転送を伴うアクセス要求（例えば「NFSPROC_READ」）であっても、受信したアクセス要求に含まれるファイルシステムアドレスを記憶する対応レジスタ１２が複数の対応レジスタ１２の中から見つからなかった場合、
転送先判定回路１１は、ＮＡＳプロセッサ５０６にデータ読出し指令を出力する。それに

より、ＮＡＳプロセッサ５０６からＩ／Ｏプロセッサ５０４に、ブロックリード要求が出力される。そして、そのブロックリード要求を受信したＩ／Ｏプロセッサ５０４によって、キャッシュメモリ１３０にステージングされているリード対象データが読み出されてＮＡＳメモリ５０８に格納され（Ｓ１５及びＳ１６）、ＮＡＳプロセッサ５０６によって、ＮＡＳメモリ５０８に格納されたリード対象データが読み出されて、ホスト端末２００に転送される（Ｓ１７及びＳ１８）。

【００６４】

一方、Ｓ１３において、肯定的な判断結果が得られた場合、すなわち、受信したアクセス要求がキャッシュメモリ１３０へのデータ転送を伴うアクセス要求（例えば「NFSPROC_READ」）であり、且つ、受信したアクセス要求に含まれるファイルシステムアドレスを記憶する対応レジスタ１２が複数の対応レジスタ１２の中から見つかった場合、転送先判定回路１１は、ダイレクトデータ転送処理を行う。具体的には、転送先判定回路１１は、その見つかった対応レジスタ１２に記憶されているＫ個のキャッシュアドレスを判定レジスタバス３４を介して取得し、取得されたＫ個のキャッシュアドレスが示すキャッシュ領域１３１（例えばＫ個のキャッシュブロック１３０Ｓ）から、コネクタ転送バス３３を経由してリード対象データを読み出し、読出したリード対象データを、ＮＡＳメモリ５０８を介することなく通信コントローラ／ＴＯＥ５０３を介してホスト端末２００に転送する（Ｓ１９）。

【００６５】

以上が、ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルリードを要求した場合に行なわれる処理流れである。なお、この処理流れにおいて、Ｓ４の処理は、Ｓ１１の後に行われても良い。また、Ｓ１２のファイルリード要求に含まれるファイルシステムアドレスは、例えば、転送先判定回路１１によって、ＮＡＳメモリ５０８に格納されても良いし、Ｉ／Ｏプロセッサ５０４によって、Ｉ／Ｏメモリ５０７に格納されても良い。

【００６６】

図６は、ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルライトを要求した場合に行なわれる処理流れを示す。なお、以下の説明において、図５を参照して説明した部分と重複する部分については説明を省略或いは簡略する。

【００６７】

ホスト端末２００のＮＦＳ２００Ｂは、ユーザからファイルライト要求を受けた場合（ステップＳ２１）、要求されたファイル名を有するデータファイルを書込むためにどこにアクセスすれば良いかを要求するためのファイルハンドラ要求を送信する（Ｓ２２）。

【００６８】

ＮＡＳプロセッサ５０６は、受信したファイルハンドラ要求に従ってファイル処理を行い（Ｓ２３）、そのファイル処理によって取得されたファイルハンドラを、ホスト端末２００のＮＦＳ２００Ｂに送信する（Ｓ２４）。

【００６９】

また、ＮＡＳプロセッサ５０６は、複数の対応レジスタ１２の中から未使用の対応レジスタ１２を探す（Ｓ２５）。

【００７０】

Ｓ２５において、未使用の対応レジスタ１２が探し出された場合（Ｓ２５でＹ）、ＮＡＳプロセッサ５０６は、探し出された未使用の対応レジスタ１２を確保する（Ｓ２６）。

【００７１】

また、ＮＡＳプロセッサ５０６は、Ｓ２５でＹの場合、上記取得されたファイルハンドラに基づくファイルシステムアドレスを、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、Ｓ２６で確保した対応レジスタ１２に書込む（Ｓ２７）。

【００７２】

また、ＮＡＳプロセッサ５０６は、Ｓ２５でＹの場合、上記取得されたファイルハンドラに基づくファイルシステムアドレスをブロックレベルのアドレスに変換し、そのブロッ

クレベルのアドレスを有するライト要求（以下、ブロックライト要求）と、Ｓ２６で確保された対応レジスタ１２のレジスタＩＤとを、Ｉ／Ｏプロセッサ５０４に送信する（Ｓ２８）。

【００７３】

Ｉ／Ｏプロセッサ５０４は、レジスタＩＤ及びブロックライト要求を受信した場合、ブロックライト要求に基づいて、キャッシュメモリ１３０上に、そのブロックライト要求に従って書込むデータ（以下、ライト対象データ）を格納するためのキャッシュ領域１３１を確保する（Ｓ２９）。そして、Ｉ／Ｏプロセッサ５０４は、ＤＫＡ１４０上の図示しないＩ／Ｏプロセッサとプロセッサ間通信を行うことにより、受信したブロックライト要求に含まれるアドレスが表す場所（ＬＵ３１０内の場所）から空データ（例えば全て「０」のビットから成るデータ）を読み出し、その空データを、上記確保したキャッシュ領域１３１に格納する（Ｓ３０）。なお、このＳ３０の処理は行わなくても良い。

【００７４】

また、Ｉ／Ｏプロセッサ５０４は、確保したキャッシュ領域１３１を表すＫ個のキャッシュアドレスを、メモリコントローラブリッジ５０５及びプロセッサレジスタバス３５を介して、上記受信したレジスタＩＤに対応する対応レジスタ１２に書込む（Ｓ３１）。

【００７５】

さて、ホスト端末２００のＮＦＳ２００Ｂは、Ｓ２４の処理によってファイルハンドラを受信した場合、受信したファイルハンドラに基づくファイルシステムアドレスを含んだファイルライト要求を、ＣＨＮ１１０に送信する（Ｓ３２）。

【００７６】

ＣＨＮ１１０の転送先判定回路１１は、ファイルシステムアドレス及びライト対象データを含んだファイルライト要求を通信コントローラ／ＴＯＥ５０３を介して受信する。転送先判定回路１１は、受信したアクセス要求はキャッシュメモリ１３０へのデータ転送を伴うアクセス要求であり、且つ、そのアクセス要求に対応したキャッシュアドレスが複数の対応レジスタ１２のいずれかに記録されているか否かを判断する（Ｓ３３）。

【００７７】

Ｓ３３において、否定的な判断結果が得られた場合、転送先判定回路１１は、ＮＡＳプロセッサ５０６にデータ書込み指令を出力する（Ｓ３４）。それにより、ＮＡＳプロセッサ５０６によってＮＡＳメモリ５０８にライト対象データが書込まれ、且つ、ＮＡＳプロセッサ５０６からブロックライト要求がＩ／Ｏプロセッサ５０４に出力される。そして、ブロックライト要求を受けたＩ／Ｏプロセッサ５０４によって、ＮＡＳメモリ５０８に格納されているライト対象データが読み出されてキャッシュ領域１３１に格納される（Ｓ３５及びＳ３６）。

【００７８】

一方、Ｓ３３において、肯定的な判断結果が得られた場合、すなわち、受信したアクセス要求がキャッシュメモリ１３０へのデータ転送を伴うアクセス要求（例えば「NFSPROC_WRITE」）であり、且つ、受信したアクセス要求に含まれるファイルシステムアドレスを記憶する対応レジスタ１２が複数の対応レジスタ１２の中から見つかった場合、転送先判定回路１１は、ダイレクトデータ転送処理を行う。具体的には、転送先判定回路１１は、その見つかった対応レジスタ１２に記憶されているＫ個のキャッシュアドレスを判定レジスタバス３４を介して取得し、取得されたＫ個のキャッシュアドレスが示すキャッシュ領域１３１（例えばＫ個のキャッシュブロック１３０Ｓ）に、ＮＡＳメモリ５０８を介することなく、コネクタ転送バス３３を経由してライト対象データを転送する（Ｓ３７）。

【００７９】

以上が、ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルライトを要求した場合に行なわれる処理流れである。なお、この処理流れにおいて、Ｓ２４の処理は、Ｓ３１の後に行われても良い。また、Ｓ３２のファイルライト要求に含まれるファイルシステムアドレスは、例えば、転送先判定回路１１によって、ＮＡＳメモリ５０８に格納されても良いし、Ｉ／Ｏプロセッサ５０４によって、Ｉ／Ｏメモリ５０７に格納されても良い

。

【００８０】

以上、上述した実施形態によれば、データ転送を伴うアクセス要求がホスト端末２００から発行されることに先行して、所定の先行コマンドがホスト端末２００から発行される。ＣＨＮ１１０において所定の先行コマンドが検出された場合、そのアクセス要求でやり取りされるデータのキャッシュ領域が確保されると共に、そのキャッシュ領域を表すキャッシュアドレスが対応レジスタ１２上でファイルシステムに対応付けられる。また、ホスト端末２００から出力されたアクセス要求は、ホスト端末２００のＴＣＰドライバ２００Ｄ及びＩＰドライバ２００ＥによってＴＣＰ及びＩＰに基づく処理が施されるが、そのアクセス要求のＩＰ層及びＴＣＰ層の部分は、通信コントローラ／ＴＯＥ５０３によって解釈されるので、転送先判定回路１１が受信するアクセス要求は、ＮＡＳ用ソフトウェア（例えばＮＦＳ）２００Ｂから出力されたアクセス要求と同じものである。これらのことにより、ホスト端末２００から、実際に、データ転送を伴うアクセス要求が発行された場合は、そのアクセス要求に従うデータは、そのアクセス要求に含まれるファイルシステムアドレスとそれに対応付けられたキャッシュアドレスとに基づいて、ＣＨＮ１１０上のメモリを介することなく、ホスト端末２００とキャッシュメモリ１３０との間で直接やり取りされる。これにより、ホスト端末２００とキャッシュメモリ１３０との間のデータのやり取りが高速化される。これは、ホスト端末２００から出力されたライト対象データをキャッシュメモリ１３０に転送する場合には特に効果的であると考えられる。また、これは、ホスト端末２００とキャッシュメモリ１３０との間でやり取りされるデータが長くなればなるほど効果的であると考えられる。

【００８１】

また、上述した実施形態によれば、ＮＡＳプロセッサ５０６が、ファイルレベルのＩ／Ｏ要求をブロックレベルのＩ／Ｏ要求に変換する仕様になっており、且つ、Ｉ／Ｏプロセッサ５０４が、キャッシュメモリ１３０を管理する（例えばキャッシュメモリ１３０上にキャッシュ領域を確保する）仕様になっている場合、それらの仕様を変更することなく、上述した高速化を実現することができる。

【００８２】

ところで、上述した実施形態について、例えば幾つかの変形例が考えられる。以下、各変形例について述べる。

【００８３】

（１）第１の変形例。

【００８４】

図７は、本実施形態の第１の変形例におけるＣＨＮ１１０の構成例を示す。なお、以下の説明では、上述した実施形態の説明と重複する部分については説明を省略或いは簡略する。

【００８５】

第１の変形例では、転送先判定回路１１は、通信コントローラ／ＴＯＥ３に搭載されている。また、この場合、通信コントローラ／ＴＯＥ３において、ＩＰやＴＣＰの解釈或いは処理を行うＩＰ／ＴＣＰ部４は、転送先判定回路１１の場所よりも上位の場所（つまりホスト端末側の場所）に搭載されている。通信コントローラ／ＴＯＥ３がアクセス要求を受信した場合に、ＩＰ及びＴＣＰの解釈をした後のアクセス要求が転送先判定回路１１に受信されるようにするためである。

【００８６】

この第１の変形例では、上述した実施形態と同様に、転送先判定回路１１に、メモリブリッジ転送バス３１や、判定レジスタバス３４を接続することができる。また、コネクタ転送バス３３を介して、転送ＬＳＩ５と転送先判定回路１１とを接続することができる。この場合、例えば、転送先判定回路１１からダイレクトデータ転送処理により出力されたライト対象データは、コネクタ転送バス３３及びデータ転送ＬＳＩ５を介して、キャッシュメモリ１３０に転送される。

【００８７】
（２）第２の変形例
図８は、本実施形態の第２の変形例における、ＣＨＮ１１０とＤＫＡ１４０との間のデータ通信例を示す。
【００８８】
第２の変形例では、ＣＨＮ１１０には、Ｉ／Ｏプロセッサは搭載されず、ＣＨＮ１１０内のＩ／Ｏプロセッサ５０４が行う処理を、ＤＫＡ１４０のＩ／Ｏプロセッサ６０３に行わせる。
【００８９】
例えば、図８（Ａ）に示すように、ＣＨＮ１１０とＤＫＡ１４０との間に、専用の割込み線５１０が備えられる。この場合、ＮＡＳプロセッサ５０６から出力されたブロックレベルのＩ／Ｏ要求は、専用の割込み線５１０を介して、ＤＫＡ１４０のＩ／Ｏプロセッサ６０３に送信される。これにより、ＤＫＡ１４０のＩ／Ｏプロセッサ６０３によって、例えば、図５のＳ９～Ｓ１１の処理や、図６のＳ２９～Ｓ３１の処理が行われる。
【００９０】
また、例えば、図８（Ｂ）に示すように、ＣＨＮ１１０上のＮＡＳメモリ５０８に、コマンドキュー５１１が設けられる。ＮＡＳプロセッサ５０６から出力されたブロックレベルのＩ／Ｏ要求は、そのコマンドキュー５１１に格納される（Ｓ５１）。ＤＫＡ１４０のＩ／Ｏプロセッサ６０３は、コマンドキューをポーリングし（Ｓ５２）、Ｉ／Ｏ要求が存在することを検出した場合には、そのＩ／Ｏ要求をコマンドキューから取得する（Ｓ５３）。
【００９１】
（３）第３の変形例。
【００９２】
図９は、本実施形態の第３の変形例に係る対応レジスタ管理テーブル１８２０の構成例を示す。
【００９３】
対応レジスタ管理テーブル８２０には、各対応レジスタ毎に、更に、ファイルシステムアドレスのアドレス登録領域が存在する。各対応レジスタに対応したアドレス登録領域には、その対応レジスタに読み出されるデータのファイルシステムアドレスが登録される。
【００９４】
例えば、Ｉ／Ｏプロセッサ５０４は、図５のＳ１１の後に、キャッシュアドレスの書込みが済んだことをＮＡＳプロセッサ５０６に通知する（Ｓ１１Ａ）。ＮＡＳプロセッサ５０６は、その通知を受けた場合、Ｓ６で確保された対応レジスタ１２に対応する対応レジスタ管理テーブル８２０上のアドレス登録領域に、Ｓ７で書き込んだファイルシステムアドレスを書込む（１１Ｂ）。これにより、そのファイルシステムアドレスに対応するデータを記憶するキャッシュ領域１３０のキャッシュアドレスが登録済みの対応レジスタ１２がどれであるかがわかるようになる。
【００９５】
このような処理を行っておくことで、以後、ファイルハンドラ要求に応答してファイルハンドラが取得された場合には、そのファイルハンドラに基づくファイルシステムアドレスに対応したキャッシュアドレスが既に登録されている対応レジスタ１２が存在するか否かの判定（以下、「登録判定処理」）を実行することができる。なお、Ｓ１１Ｂの処理は、例えば、Ｓ７の際に行なわれても良い。
【００９６】
図１０は、この第３の変形例に係る登録判定処理が行われるタイミングの例を示す。
【００９７】
図１０（Ａ）は、リードの場合の例である。すなわち、例えば、図５のＳ３又はＳ４の後に、ＮＡＳプロセッサ５０６は、登録判定処理を実行する（Ｓ４Ａ）。具体的には、例えば、ＮＡＳプロセッサ５０６は、取得されたファイルハンドラに基づくファイルシステ

ムアドレスと同一のファイルシステムアドレスが対応レジスタ管理テーブル１８２０に存在するか否かを判断する。その結果、その同一のファイルシステムアドレスが対応レジスタ管理テーブル１８２０から見つからない場合（Ｓ４ＡでＮ）、上述したＳ５が行われる。一方、Ｓ４Ａで、同一のファイルシステムアドレスが対応レジスタ管理テーブル１８２０から見つかった場合（Ｓ４ＡでＹ）、Ｓ５～１１が行われることなくＳ１２が実行される。

【００９８】

図１０（Ｂ）は、ライトの場合の例である。すなわち、例えば、図６のＳ２３又はＳ２４の後に、ＮＡＳプロセッサ５０６は、登録判定処理を実行する（Ｓ２４Ａ）。その結果、取得されたファイルハンドラに基づくファイルシステムアドレスと同一のファイルシステムアドレスが対応レジスタ管理テーブル１８２０から見つからない場合（Ｓ２４ＡでＮ）、上述したＳ２５が行われる。一方、Ｓ２４Ａで、上記同一のファイルシステムアドレスが対応レジスタ管理テーブル１８２０から見つかった場合（Ｓ２４ＡでＹ）、Ｓ２５～３１が行われることなくＳ３２が実行される。

【００９９】

ホスト端末２００から、同一のファイルリード要求が連続して発行される場合がある。また、ホスト端末２００から発行されるファイルライト要求に含まれるライト対象データは、過去にファイルリード要求を発行することによって読み出されたデータである場合がある。それらの場合、既に、図５に示した処理流れによって、ファイルリード要求或いはファイルライト要求に含まれるファイルシステムアドレスに対応したキャッシュアドレスが既に登録された対応レジスタ１２が存在することになる。このような場合に、わざわざ、図５のＳ６～Ｓ１１の処理や、図６のＳ２６～Ｓ３１の処理が行われるのは無駄である。上述した第３の変形例によれば、その無駄を防ぐことができる。

【０１００】

以上、本発明の実施形態及び変形例を説明したが、これらは本発明の説明のための例示であって、本発明の範囲をこの実施形態及び変形例にのみ限定する趣旨ではない。本発明は、他の種々の形態でも実施することが可能である。例えば、ＣＨＮ１１０とＤＫＡ１４０は一体に作られていても良い。また、例えば、複数の対応レジスタ１２に代えて、同一のメモリ上で、複数のファイルシステムアドレスと、複数のファイルシステムアドレスの各々に対応するＫ個のキャッシュアドレスとが記録されても良い。

【図面の簡単な説明】

【０１０１】

【図１】本発明の一実施形態に係る記憶制御システムを備えるコンピュータシステムの構成を示すブロック図である。

【図２】ＣＨＮ１１０の構成例を示すブロック図である。

【図３】対応レジスタ１２の構成例を示す。

【図４】対応レジスタ管理テーブル８２０の構成例を示す。

【図５】ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルリードを要求した場合に行なわれる処理流れを示す。

【図６】ホスト端末２００のユーザがＮＦＳ２００Ｂに対してファイルライトを要求した場合に行なわれる処理流れを示す。

【図７】本実施形態の第１の変形例におけるＣＨＮ１１０の構成例を示す。

【図８】本実施形態の第２の変形例における、ＣＨＮ１１０とＤＫＡ１４０との間のデータ通信例を示す。

【図９】本実施形態の第３の変形例に係る対応レジスタ管理テーブル１８２０の構成例を示す。

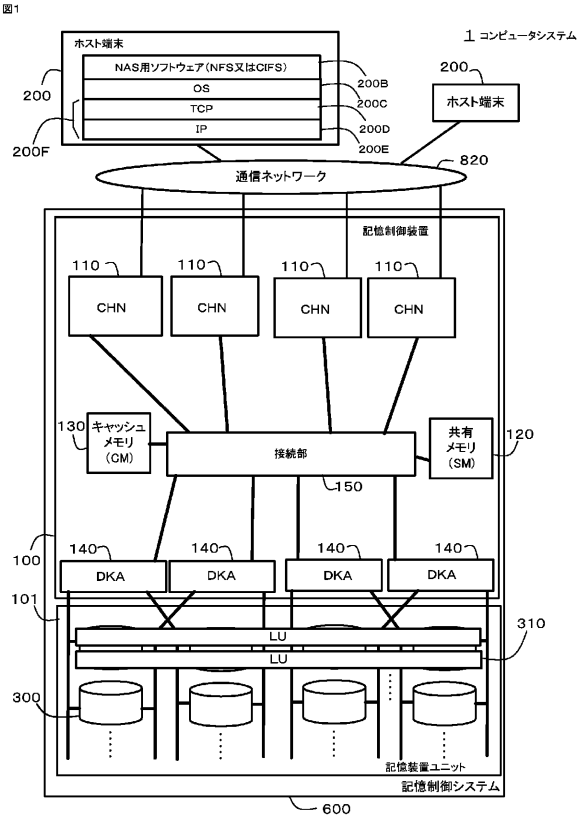
【図１０】本実施形態の第３の変形例に係る登録判定処理が行われるタイミングの例を示す。

【符号の説明】

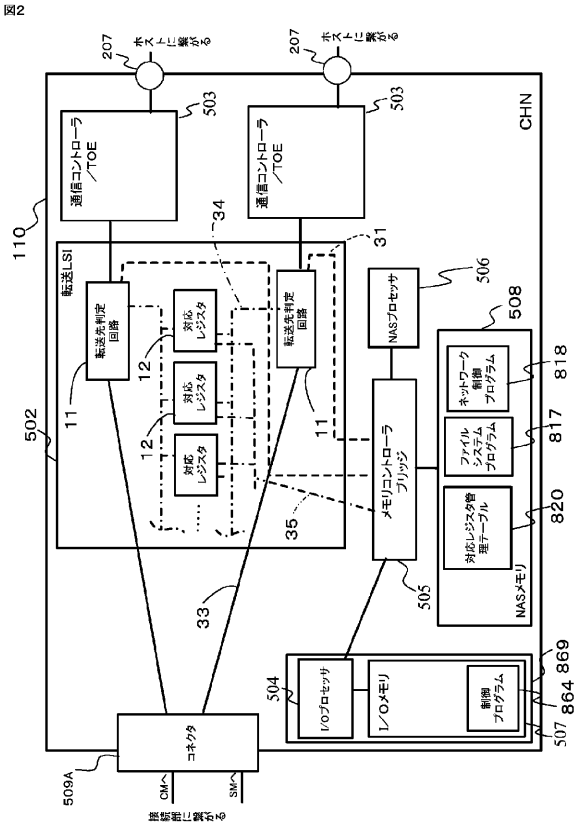
【０１０２】

１００…記憶制御装置　１０１…記憶装置ユニット　１０３…データベース　１０４…データファイル　１１０…チャネルアダプタＮＡＳ（ＣＨＮ）　１２０…共有メモリ　１３０…キャッシュメモリ　１４０…ディスクアダプタ　１５０…接続部　２００…ホスト端末　２００Ｂ…ＮＡＳ用ソフトウェア　２００Ｃ…ＯＳ　２００Ｄ…ＴＣＰドライバ　２００Ｅ…ＩＰドライバ　２００Ｆ…ネットワークドライバ　３１０…論理ユニット　６００…記憶制御システム

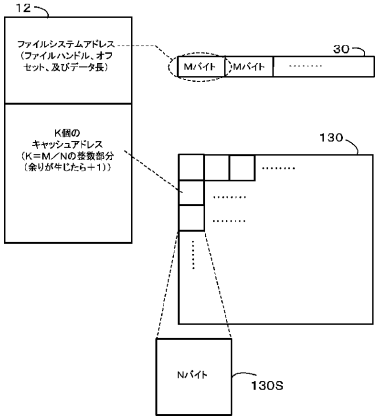
【図１】



【図２】



【図３】

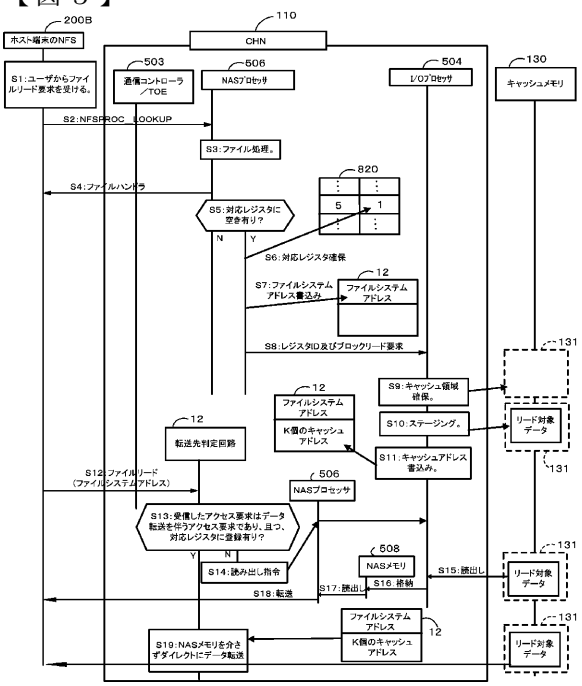


【図４】

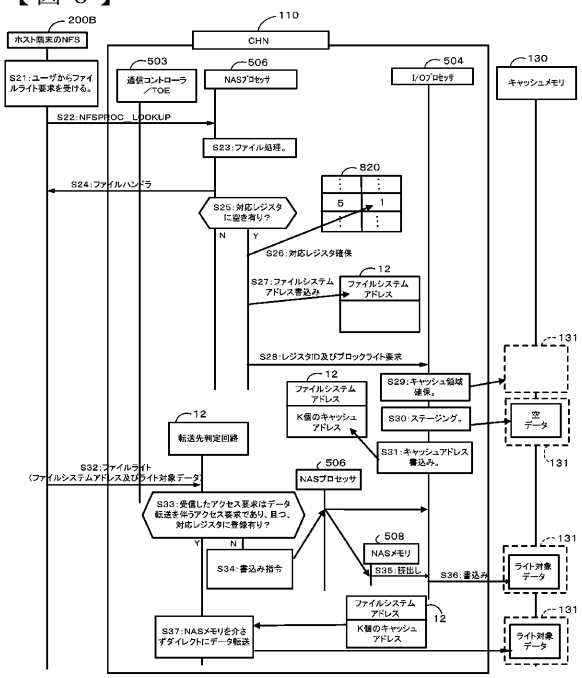
820

対応レジスタID	使用状態データ
1	1
2	0
3	1
⋮	⋮

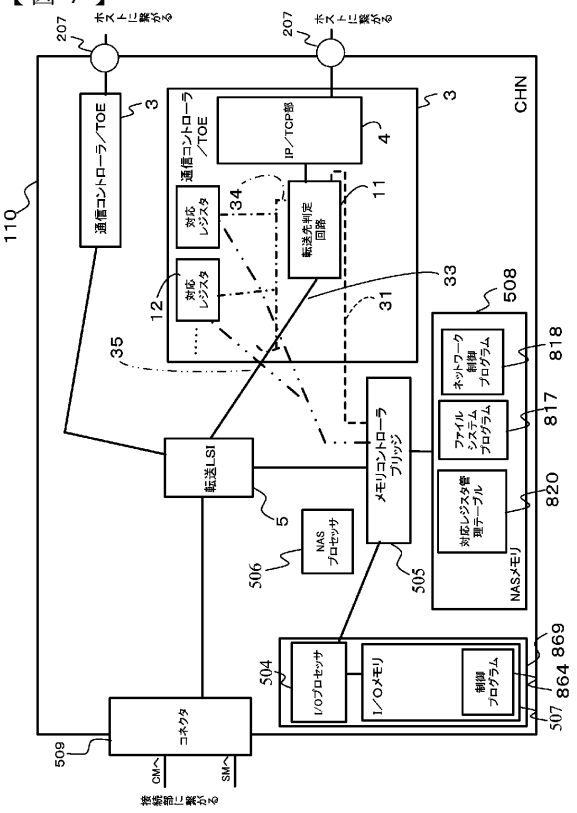
【図５】



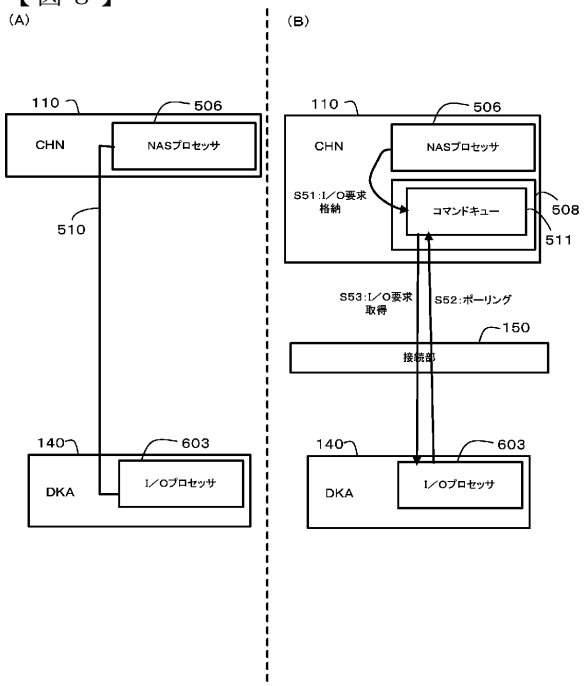
【図６】



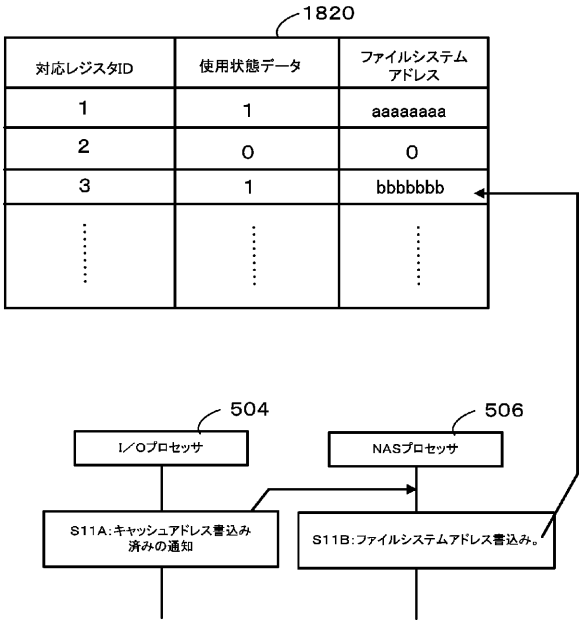
【図７】



【図８】

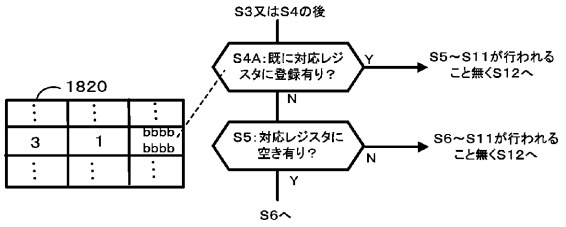

(A)
(B)
110
506
CHN
NASプロセッサ
510
S51:I／O要求格納
コマンドキュー
508
511
S53:I／O要求取得
S52:ポーリング
150
接続部
140
603
DKA
I／Oプロセッサ

【図９】

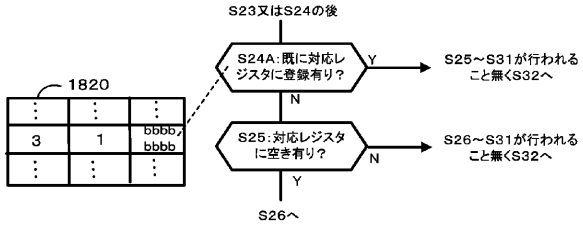

1820
対応レジスタID
使用状態データ
ファイルシステムアドレス
1
1
aaaaaaaa
2
0
0
3
1
bbbbbbb
504
I／Oプロセッサ
506
NASプロセッサ
S11A:キャッシュアドレス書込み済みの通知
S11B:ファイルシステムアドレス書込み。

【図１０】

(A)


S3又はS4の後
S4A:既に対応レジスタに登録有り?
Y
S5～S11が行われること無くS12へ
N
S5:対応レジスタに空き有り?
N
S6～S11が行われること無くS12へ
Y
S6へ
1820
3
1
bbbb
bbbb

(B)


S23又はS24の後
S24A:既に対応レジスタに登録有り?
Y
S25～S31が行われること無くS32へ
N
S25:対応レジスタに空き有り?
N
S26～S31が行われること無くS32へ
Y
S26へ
1820
3
1
bbbb
bbbb

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
	Ｇ０６Ｆ 12/08	５０３Ｚ	
	Ｇ０６Ｆ 12/08	５１７Ｂ	
	Ｇ０６Ｆ 12/08	５５１Ｚ	
	Ｇ０６Ｆ 12/08	５５７	
	Ｇ０６Ｆ 13/10	３４０Ｂ	

Ｆターム(参考) 5B005 JJ12 MM11 NN12
5B014 EB05 GC07
5B065 BA01 BA06 CA12 CA30 CC08 CH01 CH20
5B082 FA04 FA12