

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4278917号
(P4278917)

(45) 発行日 平成21年6月17日(2009.6.17)

(24) 登録日 平成21年3月19日(2009.3.19)

(51) Int. Cl.	F I		
G06F 3/06	(2006.01)	G06F 3/06	301X
G06F 12/00	(2006.01)	G06F 12/00	514E
G06F 13/10	(2006.01)	G06F 12/00	52OP
G06F 13/14	(2006.01)	G06F 13/10	34OB
		G06F 13/14	31OH

請求項の数 10 (全 19 頁)

(21) 出願番号	特願2002-113571 (P2002-113571)	(73) 特許権者	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目7番3号
(22) 出願日	平成14年4月16日(2002.4.16)	(74) 代理人	100123434 弁理士 田澤 英昭
(65) 公開番号	特開2003-308174 (P2003-308174A)	(74) 代理人	100101133 弁理士 濱田 初音
(43) 公開日	平成15年10月31日(2003.10.31)	(72) 発明者	清水 英弘 東京都千代田区丸の内二丁目2番3号 三 菱電機株式会社内
審査請求日	平成17年3月29日(2005.3.29)	(72) 発明者	郡 光則 東京都千代田区丸の内二丁目2番3号 三 菱電機株式会社内
		審査官	木村 雅也

最終頁に続く

(54) 【発明の名称】 デバイス間データ転送装置、デバイス間データ転送方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、

上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、

オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備え、上記転送経路を介して上記記憶デバイスと上記データ処理デバイスとの間でデータ転送するデバイス間データ転送装置において、

上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、

複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送をすべき優先順位を指定す

る優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を読み出して、データ転送を優先すべきボリュームを特定する情報として記録するとともに、

上記冗長性管理テーブルから、上記記録した冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、

上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、

上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理手段と、

上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率とから、上記ボリューム冗長性管理手段による判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御手段とを備えたことを特徴とするデバイス間データ転送装置。

【請求項 2】

ボリューム冗長性管理手段は、論理ボリュームにデータが書き込まれると、この論理ボリュームに対応する冗長ボリュームの複数の論理ボリュームに対して上記データの複製を自動的に書き込むことを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 3】

ボリューム冗長性管理手段は、新たな論理ボリュームを冗長ボリュームに追加した際、この冗長ボリュームに既存の論理ボリュームのデータの複製を新たな論理ボリュームに書き込むことを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 4】

論理ボリュームの優先度は、転送経路のデータ転送効率に応じて設定した値であることを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 5】

論理ボリュームの優先度は、転送経路の障害率に応じて設定した値であることを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 6】

論理ボリュームの優先度は、記憶デバイスを構成する記憶装置の故障率、論理ボリュームを構成する記憶装置の故障率、及びデータ処理デバイスの故障率に応じて設定した値であることを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 7】

論理ボリュームの優先度は、転送経路のデータ転送効率、記憶デバイスを構成する記憶装置の故障率、論理ボリュームを構成する記憶装置の故障率、データ処理デバイスの故障率、及び上記転送経路の障害率に応じて設定した値であることを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 8】

データ転送制御手段は、転送時のエラー又は転送時の実際の転送効率に合わせて論理ボリュームの優先度を動的に変更することを特徴とする請求項 1 記載のデバイス間データ転送装置。

【請求項 9】

10

20

30

40

50

複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、

上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、

オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備えたデバイス間データ転送装置のデバイス間データ転送方法において、

上記デバイス間データ転送装置のボリューム冗長性管理手段が、

上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、

複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送すべき優先順位を指定する優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を読み出して、データ転送を優先すべきボリュームを特定する情報として記録するとともに、

上記冗長性管理テーブルから、上記記録した冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、

上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、

上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理ステップと、

上記デバイス間データ転送装置のデータ転送制御手段が、

上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率から、上記ボリューム冗長性管理ステップにおける判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御ステップとを備えたことを特徴とするデバイス間データ転送方法。

【請求項10】

複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、

上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、

オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備えたコンピュータを、上記転送経路を介して上記記憶デ

10

20

30

40

50

バスと上記データ処理デバイスとのデバイス間でデータ転送するデバイス間データ転送装置として機能させるためのプログラムであって、

上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、

複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送すべき優先順位を指定する優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を、データ転送を優先すべきボリュームを特定する情報として記録するとともに、

上記冗長性管理テーブルから、上記記録した冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、

上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、

上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理手段、

上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率から、上記ボリューム冗長性管理手段による判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御手段として上記コンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明はバス又はネットワークに接続された複数のデバイス間のデータ転送を制御するデバイス間データ転送装置、デバイス間データ転送方法及びプログラムに関するものである。

【0002】

【従来の技術】

中央処理装置の負荷やデータ転送の負荷を軽減するために、主記憶装置を介さない peer-to-peer データ転送（以下、デバイス間データ転送と称する）を行うことがある。デバイス間データ転送において、デバイスの物理的位置やこれらの中に介在するバス、ネットワークの種類によって、データ転送の効率が悪くなったり、データ転送ができない場合がある。

【0003】

図5は従来のデバイス間データ転送装置を含むシステムの構成を示す図である。図において、100は中央処理装置113と主記憶装置109と各デバイスとを接続するバスであ

10

20

30

40

50

って、複数のバスを接続するためのバス結合装置102a, 102bを介してバス101a, 101bに接続している。また、バス100はデバイス間データ転送できず、バス101a, 101bはpeer-to-peer転送可能であるものとする。102a, 102bはバス結合装置で、例えば異なるバスやネットワーク間を中継するブリッジである。磁気ディスク装置106a, 107aは、ディスクコントローラ103aによってバス101aに接続して、1つの記憶デバイス105aを構成している。また、磁気ディスク装置106b, 107bも同様に、ディスクコントローラ103bによってバス101bに接続して、1つの記憶デバイス105bを構成している。以下、磁気ディスク装置106a, 107a, 106b, 107bは、物理ディスクともいう。

【0004】

104はバス101aに接続するI/Oプロセッサであって、内部に記憶装置を有するデータ処理デバイスである。このI/Oプロセッサ104が、記憶デバイス105aに関するデバイス間データ転送の際のデータ転送先となる。108は論理ボリュームで、磁気ディスク装置106a, 107a, 106b, 107b上の複数のパーティションから構成される。また、論理ボリューム108は、物理ディスク106a, 107a, 106b, 107bから構成されるミラーディスクであり、物理ディスク106aと物理ディスク106bに格納される内容がまったく同一である。同様に、物理ディスク107aと物理ディスク107bに格納される内容がまったく同一である。

【0005】

ここで、用語「パーティション」は、物理ディスク全体であっても、その一部であってもよい。用語「論理ボリューム」は、複数のパーティションを組み合わせて論理的な1つの記憶装置として扱えるようにしたものである。例えば、用語「論理ボリューム」は、いわゆるソフトウェアRAID (Redundant Arrays of Inexpensive Disk) の全てのレベル(0~5)及び「spanned volume」も含んでいる。これらの用語の定義は、後述する本発明においても同様である。

【0006】

109は主記憶装置であって、システム全体を制御するオペレーティングシステム110を格納する。111はオペレーティングシステム110に組み込まれたソフトウェアとして実装されるボリューム管理部で、記憶デバイス105a, 105bの物理的位置やI/Oプロセッサ104の物理的位置などのデバイスに関する情報、及びバス101a, 101bのデータ転送効率などの転送経路に関する情報に基づいてデータ転送可能な記憶デバイス105a, 105bの組を特定する。112はオペレーティングシステム110に組み込まれたソフトウェアとして実装されるデータ転送制御部であって、記憶デバイス105a, 105bの物理的位置やI/Oプロセッサ104の物理的位置などのデバイスに関する情報、及びバス101a, 101bのデータ転送効率などの転送経路に関する情報に基づいて記憶デバイス105a, 105bとI/Oプロセッサ104との間でデータ転送の可否を識別する。113は中央処理装置(CPU)で、オペレーティングシステム110を実行してシステム全体の各装置を制御する。

【0007】

次に動作について説明する。

まず、転送の対象となるデータがどのファイルに存在するかは、ユーザの知るところである。また、対象となるファイルがどの論理ボリューム上に存在するかは、ユーザの知るところである。論理ボリュームがどの物理ディスクから構成されているかは、システム管理者の知るところである。しかしながら、ファイルがどの物理ディスク上に存在するかは、オペレーティングシステムの中に隠蔽されているので、通常は、ユーザは知ることができない。

【0008】

そこで、上述した従来のデバイス間データ転送装置では、ファイル及びデータブロックの物理ディスク上の配置を識別するために、ボリューム管理部111及びデータ転送制御部112をオペレーティングシステム110の中に組み込んでいる。ここで、用語「ファイ

10

20

30

40

50

ル」とは、記憶デバイスに格納されたデータ転送の対象となる数値列であり、用語「データブロック」とは、データ転送の単位となる数値列の固まりを意味する。1つのファイルは、複数のデータブロックから構成される。

【0009】

ボリューム管理部111及びデータ転送制御部112をオペレーティングシステム110の中に組み込むことにより、ボリューム管理部111は、オペレーティングシステム110中に隠蔽されているファイル及びその内部のデータブロックがどの物理ディスク上に存在するかという情報を得ることができる。具体的には、ボリューム管理部111が、データごとに、オペレーティングシステムレベル又はハードウェアレベルでのデバイスに関する情報と転送経路に関する情報を取得し、デバイス及び経路を特定する。

10

【0010】

ここで、バス100はデバイス間データ転送不可能なバスで、バス102a, 102bはデバイス間データ転送可能なバスである。バス102bには、データ転送先となるI/Oプロセッサが接続されていない。論理ボリューム108は、物理ディスク106a, 107a, 106b, 107bから構成されるミラーディスクである。このようなハードウェア構成である場合、以下に述べるように、データ転送制御部112は、物理ディスク106a, 107a又は物理ディスク106b, 107bのどちらか一方の転送可能性から、I/Oプロセッサを選択することができる。

【0011】

論理ボリューム108に格納されるデータをデバイス間データ転送する場合、オペレーティングシステム110の内部で転送の対象となるデータがどの物理ディスクに存在するかを識別する。このとき、上述したように、当該物理ディスク及びデータを転送する転送経路は、物理ディスクなどのデバイスに関する情報と転送経路に関する情報のみから特定される。そして、物理ディスク106a, 107aがデータ転送の対象となり、データのミラーである物理ディスク106b, 107bに蓄えられているデータが転送の対象とならないことが明確である場合、論理ボリューム108とI/Oプロセッサ104との間でデータ転送が可能となる。

20

【0012】

このとき、物理ディスク106b, 107bに蓄えられているデータが転送の対象となる場合には、バス100がデバイス間データ転送不可能なバスなので、論理ボリューム108とI/Oプロセッサ104との間でデータ転送が不可能となる。

30

【0013】

【発明が解決しようとする課題】

従来のデバイス間データ転送装置は以上のように構成されているので、異なる転送経路を介する複数の論理ボリュームからなる冗長なボリュームを転送先又は転送元とした場合、転送効率の良い転送経路にある物理ディスク上のデータを適宜選択することができないという課題があった。

【0014】

例えば、転送可能な経路を介するデバイスと転送不可能な経路を介するデバイスを組み合わせた論理ボリューム上のデータについては、オペレーティングシステムレベル又はハードウェアレベルでのデバイスに関する情報と転送経路に関する情報（これらをデバイス識別情報と称する）のみからデバイス及び経路を特定していたため、同じ内容のデータを重複して格納する冗長なデータを選択することができなかった。

40

【0015】

また、データごとに、オペレーティングシステムレベル又はハードウェアレベルでのデバイス識別情報を取得した場合、必要なデータが存在するデバイスを特定するのは、オペレーティングシステム110に依存していたため、効率の良いデバイス間データ転送が行える経路にあるデバイスの選択をすることができなかった。例えば、論理ボリューム108がRAIDやミラーディスクなどの冗長な論理ボリュームである場合、ディスク装置106a, 107aに記憶されるデータの内容と、ディスク装置106b, 107bに記憶さ

50

れるデータの内容とが同一であったとしても、転送の対象となるデータがディスク装置 106a, 107aにあるかディスク装置 106b, 107bにあるかは、オペレーティングシステム 110任せである。

【0016】

このため、オペレーティングシステム 110によっては、例えば物理ディスク 106b, 107bに蓄えられているデータが転送の対象となった場合においても、データ転送効率の悪い経路にあるディスク装置 106b, 107bと主記憶装置 109との間でのデータ転送を行わずに、より効率の良い転送経路にあり同一内容を有するディスク装置 106a, 107aとデバイス 104との間でデータ転送を行うことを選択できなかった。

【0017】

また、RAIDやミラーディスクなどの冗長性を持たせた論理ボリュームは、オペレーティングシステム若しくはハードウェアの機能を使い、その上位層でデバイス間転送を行っている。このため、オペレーティングシステムやハードウェアに異なるディスク装置の複数のパーティションからなる論理ボリュームを生成する機能がない場合や、ネットワークドライブやリムーバブルドライブなどのように、オペレーティングシステムの機能範囲では論理ボリュームとして組み合わせ不可能なパーティションを使った冗長な論理ボリュームを生成したい場合は、冗長性を持たせた論理ボリュームをデバイス間データ転送の対象とすることができない。つまり、オペレーティングシステムの機能範囲を超える冗長な論理ボリュームを、データ転送先又は転送元とすることができないという課題があった。

【0018】

この発明は上記のような課題を解決するためになされたもので、複数のパーティションからなる論理ボリュームを構成することができないオペレーティングシステムにおいても、冗長性を有する論理ボリュームとデータ転送先又は転送元デバイスとのデバイス間データ転送を実現すると共に、転送効率の良い経路にあるデータを適宜選択することができ、また、データ転送時に故障したディスク装置のデータや転送不可能な経路にデータを流さないデバイス間データ転送装置、デバイス間データ転送方法及びプログラムを得ることを目的とする。

【0019】

【課題を解決するための手段】

この発明に係るデバイス間データ転送装置は、複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備え、上記転送経路を介して上記記憶デバイスと上記データ処理デバイスとの間でデータ転送するデバイス間データ転送装置において、上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送をすべき優先順位を指定する優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を読み出して、データ転送を優先すべきボリュームを特定する情報として記録するとともに、上記冗長性管理テーブルから、上記記録し

10

20

30

40

50

た冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理手段と、上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率とから、上記ボリューム冗長性管理手段による判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御手段とを備えるものである。

10

【0021】

この発明に係るデバイス間データ転送装置は、ボリューム冗長性管理手段が、論理ボリュームにデータが書き込まれると、この論理ボリュームに対応する冗長ボリュームの複数の論理ボリュームに対して上記データの複製を自動的に書き込むものである。

【0022】

この発明に係るデバイス間データ転送装置は、ボリューム冗長性管理手段が、新たな論理ボリュームを冗長ボリュームに追加した際、この冗長ボリュームに既存の論理ボリュームのデータの複製を新たな論理ボリュームに書き込むものである。

20

【0023】

この発明に係るデバイス間データ転送装置は、論理ボリュームの優先度が転送経路のデータ転送効率に応じて設定した値であるものである。

【0024】

この発明に係るデバイス間データ転送装置は、論理ボリュームの優先度が転送経路の障害率に応じて設定した値であることを特徴とするものである。

【0025】

この発明に係るデバイス間データ転送装置は、論理ボリュームの優先度が、記憶デバイスを構成する記憶装置の故障率、論理ボリュームを構成する記憶装置の故障率、及びデータ処理デバイスの故障率に応じて設定した値であることを特徴とするものである。

30

【0026】

この発明に係るデバイス間データ転送装置は、論理ボリュームの優先度が、転送経路のデータ転送効率、記憶デバイスを構成する記憶装置の故障率、論理ボリュームを構成する記憶装置の故障率、及びデータ処理デバイスの故障率及び転送経路の障害率に応じて設定した値であることを特徴とするものである。

【0027】

この発明に係るデバイス間データ転送装置は、データ転送制御手段が、転送時のエラー又は転送時の実際の転送効率に合わせて論理ボリュームの優先度を動的に変更するものである。

40

【0028】

この発明に係るデバイス間データ転送方法は、複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備えたデバイス間データ転送装置のデバイス間データ転送方法において、上記デバイス間データ転送装置の

50

ボリューム冗長性管理手段が、上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送をすべき優先順位を指定する優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を読み出して、データ転送を優先すべきボリュームを特定する情報として記録するとともに、上記冗長性管理テーブルから、上記記録した冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率

が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理ステップと、上記デバイス間データ転送装置のデータ転送制御手段が、上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率から、上記ボリューム冗長性管理ステップにおける判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率

が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御ステップとを備えるものである。

10

20

30

【 0 0 3 0 】

この発明に係るプログラムは、複数の記憶装置と、上記記憶装置と転送経路の間のデータ転送を中継するディスクコントローラとを有する複数の記憶デバイスと、上記転送経路に接続し、上記転送経路を介して上記記憶デバイスから転送されたデータを処理する複数のデータ処理デバイスと、オペレーティングシステムを実行し、この実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報に基づいて、上記記憶デバイスから読み込んだデータを処理する中央処理装置とを備えたコンピュータを、上記転送経路を介して上記記憶デバイスと上記データ処理デバイスとのデバイス間でデータ転送するデバイス間データ転送装置として機能させるためのプログラムであって、上記記憶装置のデータを上記記憶デバイスから上記データ処理デバイスへデータ転送するにあたり、上記オペレーティングシステムの実行に伴って利用される転送対象のデータが存在する記憶装置を特定する識別情報を用いず、複数の上記記憶装置の組からなる論理ボリュームを特定する論理ボリューム識別情報、この論理ボリューム識別情報で特定される複数の論理ボリュームからなり、複数の上記論理ボリューム間で同一のデータがミラーリングされた冗長ボリュームを特定する冗長ボリューム識別情報、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送の対象となる上記データ処理デバイスを特定するデバイス特定情報、及び上記記憶デバイスと上記データ処理デバイスとの間のデータ転送における各転送経路のデータ転送効率に応じて上記冗長ボリュームの論理ボリュームに付与したデータを転送をすべき優先順位を指定する優先度からなる冗長性情報を格納する冗長性管理テーブルから、上記データの転送元の論理ボリュームを特定する論理ボリューム識別情報及び

40

50

転送先の上記データ処理デバイスを特定するデバイス特定情報に合致する上記冗長性情報を、データ転送を優先すべきボリュームを特定する情報として記録するとともに、上記冗長性管理テーブルから、上記記録した冗長性情報の冗長ボリューム識別情報及びデバイス特定情報に対応する他の冗長性情報を探索し、上記他の冗長性情報がなければ、上記記録した冗長性情報で特定される冗長ボリュームの論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定し、上記他の冗長性情報が存在すると、上記記録した冗長性情報の論理ボリュームに付与された優先度と、上記他の冗長性情報の論理ボリュームに付与された上記優先度を比較して優先度の高い論理ボリュームを、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームと判定するボリューム冗長性管理手段、上記記憶装置の物理的位置、複数の上記データ処理デバイスの物理的位置及び複数の上記転送経路のデータ転送効率を記憶するボリューム管理テーブルから読み出した上記記憶装置及び上記データ処理デバイスの物理的位置と上記複数の転送経路のデータ転送効率から、上記ボリューム冗長性管理手段による判定で得られた上記記憶デバイスと上記データ処理デバイスとの間のデータ転送効率が他より良い転送経路に位置する論理ボリュームを構成する記憶装置のデータ転送の可否を識別し、上記記憶デバイスと上記データ処理デバイスとの間のデータ転送を制御するデータ転送制御手段として上記コンピュータを機能させるものである。

10

【0032】

【発明の実施の形態】

20

以下、この発明の実施の一形態を説明する。

実施の形態 1 .

図 1 はこの発明の実施の形態 1 によるデバイス間データ転送装置を含むシステムの構成を示す図である。図において、1 はネットワーク結合装置 3 a , 3 b によってバス（転送経路）2 , 2 a に接続されるネットワーク（転送経路）である。2 は中央処理装置 1 5 と主記憶装置 1 0 を接続しているバス（転送経路）であり、複数のバスを接続するためのバス結合装置 3 c によってバス（転送経路）2 b に接続されている。バス結合装置 3 c は、例えばブリッジである。4 a , 4 b はディスクコントローラであって、ディスクコントローラ 4 a によって磁気ディスク装置 6 a , 7 a をバス 2 a に接続して 1 つの記憶デバイス 9 a を構成する。同様に、ディスクコントローラ 4 b によって磁気ディスク装置 6 b , 7 b をバス 2 b に接続して 1 つの記憶デバイス 9 b を構成している。5 a , 5 b はバス 2 a , 2 b にそれぞれ接続するデータ処理デバイスであり、例えば内部に記憶装置を有する I / O プロセッサを含むものとする。バス 2 a , 2 b は、それ自体が内部的に複数のバスが結合されたバスの集合であってもよい。

30

【0033】

8 は冗長ボリュームであって、デバイス 9 a を構成する磁気ディスク装置 6 a , 7 a からなる論理ボリューム 8 a とデバイス 9 b を構成する磁気ディスク装置 6 b , 7 b からなる論理ボリューム 8 b との 2 つの論理ボリュームにより構成される。この実施の形態 1 において、冗長ボリュームとは、多重度が任意のミラーディスクを意味する。つまり、冗長ボリューム 8 は 2 重の冗長性を持っており、論理ボリューム 8 a 上に記憶されるデータと論理ボリューム 8 b 上に記憶されるデータは同一とする。

40

【0034】

また、デバイス 9 a , 9 b 及びデバイス 5 a , 5 b は、1 システムの中に複数接続されており、相互に直接データ転送が行われる。通常、デバイス 9 a , 9 b の磁気ディスク装置 7 a , 7 b に蓄えられているデータは、主記憶装置 1 0 に転送され中央処理装置 1 5 で処理される。しかしながら、この実施の形態に係るデバイス間データ転送装置では、主記憶装置 1 0 を経由することなく、デバイス 5 a , 5 b の内部にある記憶装置に転送され、デバイス 5 a , 5 b 内の I / O プロセッサによって処理されるものとする。

【0035】

1 0 は主記憶装置であって、ボリューム管理部 1 1、データ転送制御部 1 2 及びボリュー

50

ム冗長性管理部 1 4 や、システムの制御を行うオペレーティングシステム 1 3 などのソフトウェアを格納する。1 1 はボリューム管理部で、ボリューム管理に要する構成要素を特定するボリューム管理情報を格納し、他の構成要素に適宜提供する。また、ボリューム管理部 1 1 は、オペレーティングシステム 1 3 の初期化後に、ボリューム構成テーブル 1 6 を直接参照するか、若しくは、オペレーティングシステム 1 3 に対して、ボリューム構成情報、バス情報、物理ディスク・パーティション情報、バス接続情報、ネットワーク接続情報や I / O プロセッサ情報などの情報の採取を要求して、ボリューム管理テーブル 1 7 を生成する。

【 0 0 3 6 】

1 2 はデータ転送制御部（データ転送制御手段）であって、ボリューム管理部 1 1 が管理する情報に基づいて磁気ディスクとデータ処理デバイスとの間でのデータ転送の可否を識別する。1 3 はオペレーションシステムで、システム全体のソフトウェアを制御する。また、オペレーティングシステム 1 3 は、初期時若しくはユーザによるツールの使用時に、デバイスドライバ 2 0 に情報採取を要求し、ボリューム構成テーブル 1 6 及びバス・ネットワーク情報 1 8 を生成する。1 4 はボリューム冗長性管理部（ボリューム冗長性管理手段）であって、ユーザの指定により冗長性がある論理ボリュームの組（つまり、上述した冗長ボリュームを構成する論理ボリューム群）を特定する情報やデバイスを特定する情報を格納する。

【 0 0 3 7 】

また、ボリューム管理部 1 1、データ転送制御部 1 2 及びボリューム冗長性管理部 1 4 は、オペレーティングシステム 1 3 とは分離してプログラムされたソフトウェアとして主記憶装置 1 0 上に実装されて、オペレーティングシステム 1 3 に対してシステムの情報を問い合わせることができる。さらに、ボリューム管理部 1 1、データ転送制御部 1 2 及びボリューム冗長性管理部 1 4 は、ファームウェアとして実現されてもよいし、ハードウェアとして実現されてもよいし、ソフトウェアとファームウェアとハードウェアの組み合わせで実現されてもよい。1 5 は中央処理装置（CPU）で、オペレーティングシステム 1 3 などを実行してシステム内の各構成要素を制御する。

【 0 0 3 8 】

図 2 は実施の形態 1 によるデバイス間データ転送装置の構成要素の相互関係を示す図である。図において、1 6 はボリューム構成テーブルであって、オペレーティングシステム 1 3 が管理するボリュームの構成情報が格納される。1 7 はボリューム管理部 1 1 に生成・管理されるボリューム管理テーブルで、主にボリューム管理に要する構成要素を特定するボリューム管理情報を格納する。このボリューム管理情報としては、システム内の構成要素の処理性能を特定したり、相互関係を示す指標となる、例えばバス転送情報（転送効率や障害率など）、バス結合装置情報、物理ディスク情報、論理ボリューム情報及び I / O プロセッサ情報（例えば、故障率）などがある。1 8 はバス・ネットワーク情報であって、オペレーティングシステム 1 3 が管理するネットワーク 1 やバス 2 , 2 a , 2 b の構成情報が格納される。

【 0 0 3 9 】

1 9 はボリューム冗長性管理部 1 4 に生成・管理される冗長性管理テーブルで、装置内に構成されている冗長性がある論理ボリュームを特定する情報やこれらに対するデータ転送時における優先度などを冗長性情報として格納する。この冗長性管理テーブル 1 9 に記憶される内容は、後述する図 3 において詳細に説明する。また、ボリューム冗長性管理部 1 4 は、冗長性管理テーブル 1 9 に格納される情報に基づいて、データの書き込み時に関係する全ての論理ボリュームに同一のデータを書き込む要求をオペレーティングシステム 1 3 に依頼する。2 0 はデバイスドライバであって、オペレーティングシステム 1 3 と、バス結合装置、物理ディスクや I / O プロセッサなどのハードウェアとのデータ送受を中継するソフトウェアである。また、図中の矢印は、問い合わせ、参照、処理要求、生成を行う側から受ける側を指している。なお、図 1 と同一構成要素には同一符号を付して重複する説明を省略する。

10

20

30

40

50

【 0 0 4 0 】

図 3 は図 2 中の冗長性管理テーブルの構成を示す図である。図において、21 は冗長性情報であって、論理ボリューム名 22、デバイス番号 23、優先度 24 及び冗長セット番号 25 から構成される。冗長性情報 21 は、冗長性管理テーブル 19 内に、冗長ボリュームを構成する論理ボリュームの数だけ格納される。22 は論理ボリューム名（ボリューム特定情報）であって、デバイスドライバ 20 に対して要求を出す際に各論理ボリュームを識別する識別子に相当する。23 はデバイス番号（デバイス特定情報）で、デバイス間データ転送の対象となるデバイスを特定する番号である。

【 0 0 4 1 】

24 は論理ボリュームのデータを転送すべき優先順位を指定する優先度であって、データの転送効率に相当する情報である。この優先度 24 は、論理ボリュームと I/O プロセッサとの間の転送経路の最大の転送効率、論理ボリュームと I/O プロセッサ間の転送経路の障害率、論理ボリュームの処理性能（例えば、これを構成する記憶デバイスの故障率など）、論理ボリュームの故障率や、I/O プロセッサの処理性能（故障率など）のいずれから算出するか、これらの組み合わせによる計算で算出しても良い。

10

【 0 0 4 2 】

例えば、優先度 24 を論理ボリュームと I/O プロセッサとの最大の転送効率と論理ボリュームの故障率から算出する場合、転送効率を 5 段階評価とし、故障率を 5 段階評価として、優先度 24 を表す式として（転送効率）× 5 + 故障率を用いる。この場合、転送効率を重視した論理ボリュームの選択を行うことができる。また、論理ボリュームと I/O プロセッサとの最大の転送効率や、論理ボリュームと I/O プロセッサ間の転送経路の障害率は、ボリューム管理部 11 に問い合わせることにより取得される。このとき、ボリューム管理部 11 は、ボリューム管理テーブル 17 を参照する。

20

【 0 0 4 3 】

また、デバイス間データ転送におけるデータ転送効率のよい転送経路に位置する論理ボリュームに対して高い優先度 24 が付与される。ここで、ボリューム冗長性管理部 14 は、ユーザが優先度 24 を指定したとき、若しくは、コンピュータの初期化時、若しくは、データ転送後にデータ転送速度の統計情報をデータ転送制御部 12 から受け取って優先度 24 の更新を依頼された時、冗長性管理テーブル 19 の優先度 24 を更新する。

30

【 0 0 4 4 】

25 は冗長ボリュームを特定する冗長セット番号（ボリューム特定情報）であって、同一の冗長ボリュームに属する論理ボリュームには同一値が与えられる。つまり、同一の冗長セット番号 25 で特定されるディスク装置には、互いに同一のデータが格納される。なお、図 2 と同一構成要素には同一符号を付して重複する説明を省略する。

【 0 0 4 5 】

次に動作について説明する。

データ転送制御部 12 は、データ転送前に、転送対象のデータを格納する論理ボリュームを特定する情報である論理ボリューム名 22 及びデバイス番号 23 をボリューム冗長性管理部 14 に送信する。ボリューム冗長性管理部 14 では、データ転送制御部 12 から受信した論理ボリューム名 22 及びデバイス番号 23 について冗長性管理テーブル 19 の優先度 24 を参照して、転送対象のデータを格納しデータ転送がより優先される論理ボリュームを特定する論理ボリューム名 22 を返信する。続いて、データ転送制御部 12 は、ボリューム管理部 11 に情報採取を要求して転送経路の転送可能性若しくは転送効率を取得する。このとき、ボリューム管理部 11 は、ボリューム管理テーブル 17 を参照して転送経路の転送可能性などをデータ転送制御部 12 に提供する。

40

【 0 0 4 6 】

ここで、転送経路の転送可能性は、ボリューム管理テーブル 17 においてバス転送情報として格納される。この転送可能性とは、デバイス間データ転送の可能不可能を含む転送効率である。転送可能性の最も単純な形式は、デバイス間データ転送の可能不可能を ON / OFF で記憶することである。或いは、転送効率を 5 段階で表すとすれば、1 ~ 5 の値を

50

記憶する。例えば、1は転送効率最大であり、 $1 > 2 > 3 > 4 > 5$ の順に転送効率を示す。5は転送不可能を示す。バス転送情報には、転送経路となるバスを特定するバス番号とバスの転送可能性の組がバスの数だけ格納される。

【0047】

データ転送制御部12は、上記転送経路の情報に基づいてデバイス間データ転送すべきか、主記憶装置10に転送して中央処理装置15にて処理すべきかを判断する(データ転送制御ステップ)。このあと、データ転送制御部12は、オペレーティングシステム13に対して判断結果の転送経路に沿ったデータ転送の要求を出す。オペレーティングシステム13では、データ転送制御部12の要求に従ってデバイスドライバ20にデータ転送の要求を渡す。これにより、データ転送が開始される。

10

【0048】

次に転送対象のデータを格納し、且つデータ転送が最も優先される論理ボリュームを求める動作について詳細に説明する。

図4は図2中のボリューム冗長性管理部の動作を示すフロー図であり、この図及び図3を参照して説明する。

まず、データ転送制御部12は、データの転送元の論理ボリュームを特定する論理ボリューム名22及び転送先のデバイスを特定するデバイス番号23をボリューム冗長性管理部14に送信する(ステップST1)。ボリューム冗長性管理部14は、受信した論理ボリューム名22とデバイス番号23との組に対応する冗長性情報21を冗長性管理テーブル19から読み出す。このあと、ボリューム冗長性管理部14は、データ転送を最優先すべきボリュームを特定する情報として該冗長性情報21を記録する(ステップST2)。これらステップST1及びステップST2が選別ステップに相当する。

20

【0049】

続いて、ボリューム冗長性管理部14は、上述した最優先すべきボリュームと冗長セット番号25が同じで且つデバイス番号23が同じ論理ボリュームが存在するか否かを判定する(ステップST3)。具体的には、ボリューム冗長性管理部14が、ステップST2にて読み出した冗長性情報21内の冗長セット番号25及びデバイス番号23をキーとして、冗長性管理テーブル19の内容を検索し、同一の冗長セット番号25及びデバイス番号23を有する冗長性情報21を探索する。このとき、同一の冗長セット番号25及びデバイス番号23が存在しなければ、ステップST8の処理に進む。

30

【0050】

ステップST8において、ボリューム冗長性管理部14は、ステップST2にて読み出した冗長性情報21で特定される論理ボリュームが、転送対象のデータを格納し、且つデータ転送が最も優先される論理ボリュームであるものと判断して、該論理ボリュームを特定する論理ボリューム名22をデータ転送制御部12に返信する。

【0051】

一方、ステップST3において、同一の冗長セット番号25及びデバイス番号23を有する論理ボリュームを特定する冗長性情報21が検出されると、ボリューム冗長性管理部14は、該冗長性情報21を冗長性管理テーブル19から取り出す(ステップST4)。このあと、ボリューム冗長性管理部14は、取り出した冗長性情報21から優先度24を抽出・保持する(ステップST5)。

40

【0052】

さらに、ボリューム冗長性管理部14は、取得した優先度24の値が最優先のボリュームの優先度24より上か否かを判定する(ステップST6、優先ボリューム選択ステップ)。ここでは、ステップST1にてデータ転送制御部12から指定された論理ボリュームの冗長性情報21内の優先度24と、ステップST5にて抽出・保持された冗長性情報21内の優先度24とが大小比較される。このとき、前者の優先度24の値が後者の優先度24より高ければ、ステップST3の処理に戻って、同一の冗長セット番号25及びデバイス番号23を有する他の論理ボリュームの優先度24との比較処理が行われる。

【0053】

50

一方、後者の優先度 24 が前者の優先度 24 より高ければ、ボリューム冗長性管理部 14 は、ステップ S T 5 にて抽出・保持した冗長性情報 21 で特定される論理ボリュームを、データ転送を最優先すべきボリュームとして記録する（ステップ S T 7）。

【 0 0 5 4 】

このあと、ステップ S T 3 の処理に戻って、ステップ S T 7 にてデータ転送を最優先すべきボリュームとして記録された論理ボリュームの優先度 24 と、これと同一の冗長セット番号 25 及びデバイス番号 23 を有する他の論理ボリュームの優先度 24 とが比較処理される。

【 0 0 5 5 】

以上の処理によって、ボリューム冗長性管理部 14 は、ステップ S T 1 にて受信した論理ボリューム名 22 及びデバイス番号 23 によって特定される論理ボリュームが格納する転送対象のデータと同一のデータを格納する論理ボリュームのうち、データ転送をより優先すべきボリュームが選択される。そして、該論理ボリュームを特定する論理ボリューム名 22 が、データ転送制御部 12 に返信される。

【 0 0 5 6 】

図 1 で、例えば I / O プロセッサ 5 b が接続されておらず、論理ボリューム 8 b からのデータは、バス 2 b、バス結合装置 3 c 及びバス 2 を経由して主記憶装置 10 に転送され、中央処理装置 15 で処理される場合を考える。この場合、冗長性がある論理ボリューム（冗長ボリューム 8）として論理ボリューム 8 a, 8 b を冗長性管理テーブル 19 に指定するにあたり、実システムにあった転送効率を反映させた優先度 24 によって論理ボリューム 8 a の方が論理ボリューム 8 b より優先される。これにより、データ転送効率の悪い経路にあるディスク装置 6 b, 7 b からなる論理ボリューム 8 b と主記憶装置 10 との間でのデータ転送が行われず、より効率の良い転送経路にあるディスク装置 6 a, 7 a からなる論理ボリューム 8 a とデバイス 5 a との間でのデータ転送を選択することができる。

【 0 0 5 7 】

このように、論理ボリュームを構成するディスクの物理的な位置や転送経路を考慮して優先度 24 を設定することで、より効率の良い転送経路にあるディスク装置からなる論理ボリュームとデータ処理デバイスとの間でのデータ転送を選択することができる。

【 0 0 5 8 】

また、データ転送の際、デバイスドライバ 20 がデバイスのエラーを検出した場合、そのエラー情報は、データ転送制御部 12 に伝えられる。このとき、発生したエラーがハードウェアの致命的なエラーであり、以降エラーが検出されたデバイスが使用できないことが明白な場合、データ転送制御部 12 は、ボリューム冗長性管理部 14 に対してエラーとなったデバイスが関与する論理ボリュームの冗長性情報 21 を冗長性管理テーブル 19 より削除することを要求する。

【 0 0 5 9 】

この要求に従い、ボリューム冗長性管理部 14 は、対象のボリュームに関する冗長性情報 21 を冗長性管理テーブル 19 より削除する。このとき、ユーザが当該ボリュームに関する冗長性情報 21 を新たに冗長性管理テーブル 19 に追加しなければ、削除されたボリューム上のデータがデバイス間データ転送の対象となることはない。

【 0 0 6 0 】

また、データ転送の際、データ転送制御部 12 によって転送サイズと転送時間より、実際の転送速度が求められる。データ転送制御部 12 は、ボリューム冗長性管理部 14 に対して転送速度の統計情報を通知する。ボリューム冗長性管理部 14 は、対象の論理ボリューム名 22 及びデバイス番号 23 の優先度 24 の値を更新する。

【 0 0 6 1 】

さらに、ユーザが、既にデータが格納されている論理ボリュームと同一の冗長セット番号 25 を有する論理ボリュームを追加する旨をボリューム冗長性管理部 14 に対して指定した場合、ボリューム冗長性管理部 14 は、冗長性管理テーブル 19 に新たな論理ボリュームに係る冗長性情報 21 を追加する。続けて、ボリューム冗長性管理部 14 は、既にデー

10

20

30

40

50

タが格納されている論理ボリュームのデータを新たに追加された論理ボリュームにコピーする。

【 0 0 6 2 】

さらに、冗長ボリュームを構成する任意の論理ボリュームにデータを追加した場合、ボリューム冗長性管理部 1 4 は自動的に他の構成要素の論理ボリューム（冗長ボリュームを構成する他の論理ボリューム）についても、同じ内容を追加する。

【 0 0 6 3 】

以上のように、この実施の形態 1 によれば、冗長性管理テーブル 1 9 に記憶した優先度 2 4 に基づいて、転送対象のデータを格納する複数の論理ボリュームからデータ処理デバイスとの間のデータ転送効率がより良い転送経路に位置する論理ボリュームを識別するので、複数の論理ボリュームからなる冗長ボリューム上のデータをデバイス間データ転送の対象とした場合、データ転送効率の良い経路にあるデバイス上のデータが格納される論理ボリュームを選択してデバイス間データ転送を行うことができる。

10

【 0 0 6 4 】

また、データ転送時に転送効率の悪い経路にデータを流すことを防ぐことができ、データ転送装置全体のパフォーマンスの劣化を回避することができる。

【 0 0 6 5 】

さらに、転送可能な経路を含むデバイスと転送不可能な経路を含むデバイスとを混在させて冗長ボリュームを構成した場合であっても、オペレーティングシステムに組み込むこと無しに効率の良い転送経路を選択してデバイス間データ転送を行うことができる。

20

【 0 0 6 6 】

つまり、従来の技術で述べたように、転送対象のデータがどの物理ディスク上に存在するかは、オペレーティングシステム 1 3 の中に隠蔽されているので、通常は、ユーザは知ることができない。一方、転送の対象となるデータがどの論理ボリューム上に存在するかは、ユーザの知るところである。また、論理ボリュームがどの物理ディスクから構成されているかは、システム管理者の知るところである。

【 0 0 6 7 】

これにより、冗長ボリューム、論理ボリュームや論理ボリュームを構成する物理ディスク（記憶デバイス）の構造は、ユーザにとって既知である。そこで、上述したボリュームの構造に基づいて、データ処理デバイスとの間のデータ転送効率などに応じた論理ボリュームの優先度 2 4 を設けることで、オペレーティングシステムに隠蔽された情報を要することなく、上記データ転送効率が最も良い転送経路に位置する論理ボリュームを選択することができる。

30

【 0 0 6 8 】

また、従来では R A I D やミラーディスクのようにデータに冗長性を持たせるために、オペレーティングシステムやハードウェアの論理ボリューム生成機能を用いる必要があったが、ボリューム冗長性管理部 1 4 及び冗長性管理テーブル 1 9 によってデータに冗長性を持たせることができる。これにより、オペレーティングシステムに異なるディスク装置の複数のパーティションからなる論理ボリュームを生成する機能がない場合や、組み合わせ不可能なパーティションを使った冗長性のある論理ボリュームを生成したい場合であっても、自由に冗長ボリュームを構成し、デバイス間データ転送の対象とすることができる。

40

【 0 0 6 9 】

さらに、この実施の形態 1 によれば、ハードウェア障害などで不良となった論理ボリュームを、エラーの発生と同時に自動的にデバイス間データ転送の対象から除外するので、データ転送時に転送不可能な経路にデータを流すことを防ぐことができ、致命的な障害を回避することができる。

【 0 0 7 0 】

さらに、実際の転送速度を優先度に反映させることにより、実システムにあった転送効率で論理ボリュームを選択する。

【 0 0 7 1 】

50

さらに、上記実施の形態では、論理ボリュームを構成する記憶デバイスとしてネットワーク1を介したりリモートドライブを例にして説明したが、1つのコンピュータ装置に搭載される「物理ディスク」として機能する磁気テープや半導体ディスクなどの記憶媒体であってもよい。

【0072】

なお、本実施形態では、デバイス間データ転送はネットワークを介したデバイス間データ転送であってもよい。

【0073】

【発明の効果】

以上のように、この発明によれば、複数の記憶デバイスと複数のデータ処理デバイスとが複数の転送経路を介して接続し、デバイス間で転送経路を介してデータを転送するデバイス間データ転送装置において、複数の記憶デバイスの組からなる論理ボリュームと複数の論理ボリュームの組からなる冗長ボリュームとをそれぞれ特定するボリューム特定情報、デバイス間データ転送の対象となるデバイスを特定するデバイス特定情報、及び各論理ボリュームの優先度を記憶しておき、該冗長性管理テーブルに記憶した優先度に基づいて、ボリューム特定情報及びデバイス特定情報にて選別された論理ボリュームから、デバイス間のデータ転送効率が他よりも良い転送経路に位置する論理ボリュームを識別するので、複数の論理ボリュームからなる冗長ボリューム上のデータをデバイス間データ転送の対象とした場合、データ転送効率の良い経路にあるデバイス上のデータが格納される論理ボリュームを選択することができるという効果がある。

10

20

【0074】

この発明によれば、記憶デバイスの物理的位置、複数のデータ処理デバイスの物理的位置及び複数の転送経路のデータ転送効率を記憶するボリューム管理テーブルに記憶された各デバイスの物理的位置と複数の転送経路のデータ転送効率とに基づいて、優先度が最も高い論理ボリュームを構成する記憶デバイスのデータ転送の可否を識別するので、複数の論理ボリュームからなる冗長ボリューム上のデータをデバイス間データ転送の対象とした場合、データ転送効率の良い経路にあるデバイスのデータが格納される論理ボリュームを選択してデバイス間データ転送を行うことができるという効果がある。

【0075】

この発明によれば、論理ボリュームにデータが書き込まれると、該論理ボリュームに関連する複数の論理ボリュームに対してデータの複製を自動的に書き込むので、論理ボリュームや冗長ボリュームを構成する際の負担を軽減することができるという効果がある。

30

【0076】

この発明によれば、新たな論理ボリュームを冗長なボリュームに追加した際に、自動的に既存の論理ボリュームのデータの複製を新たな論理ボリュームに書き込むので、冗長ボリュームを変更する際の負担を軽減することができるという効果がある。

【0077】

この発明によれば、論理ボリュームの優先度が、転送経路の障害率に応じて設定した値、論理ボリュームの優先度が、デバイスの故障率に応じて設定した値、若しくは、転送経路のデータ転送効率、デバイスの故障率及び転送経路の障害率に応じて設定した値であるので、実システムにあった転送効率で論理ボリュームを選択することができるという効果がある。

40

【0078】

この発明によれば、転送時のエラー又は転送時の実際の転送効率に合わせて論理ボリュームの優先度を動的に変更するので、データ転送時に転送不可能な経路にデータを流すことを防ぐことができ、致命的な障害を回避することができるという効果がある。また、実システムにあった転送効率で論理ボリュームを選択することができるという効果がある。

【図面の簡単な説明】

【図1】 この発明の実施の形態1によるデバイス間データ転送装置を含むシステムの構成を示す図である。

50

【図2】 実施の形態1によるデバイス間データ転送装置の構成要素の相互関係を示す図である。

【図3】 図2中の冗長性管理テーブルの構成を示す図である。

【図4】 図2中のボリューム冗長性管理部の動作を示すフロー図である。

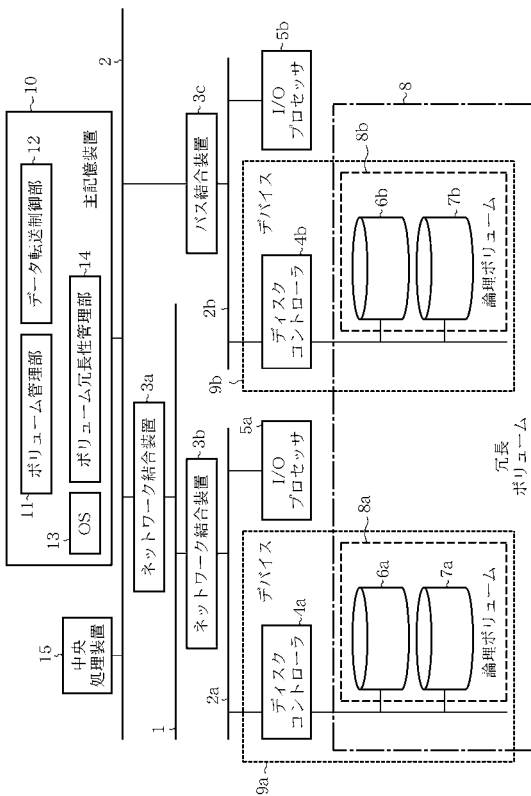
【図5】 従来のデバイス間データ転送装置を含むシステムの構成を示す図である。

【符号の説明】

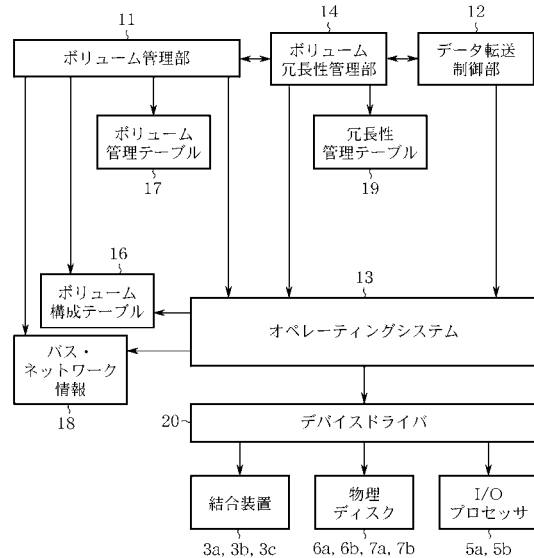
1 ネットワーク（転送経路）、2, 2a, 2b バス（転送経路）、3a, 3b ネットワーク結合装置、3c バス結合装置、4a, 4b ディスクコントローラ、5a, 5b I/Oプロセッサ（データ処理デバイス）、6a, 6b, 7a, 7b 磁気ディスク装置、8 冗長ボリューム、8a, 8b 論理ボリューム、9a, 9b 記憶デバイス、10 主記憶装置、11 ボリューム管理部、12 データ転送制御部（データ転送制御手段）、13 オペレーティングシステム、14 ボリューム冗長性管理部（ボリューム冗長性管理手段）、15 中央処理装置（CPU）、16 ボリューム構成テーブル、17 ボリューム管理テーブル、18 バス・ネットワーク情報、19 冗長性管理テーブル、20 デバイスドライバ、21 冗長性情報、22 論理ボリューム名（ボリューム特定情報）、23 デバイス番号（デバイス特定情報）、24 優先度、25 冗長セット番号（ボリューム特定情報）。

10

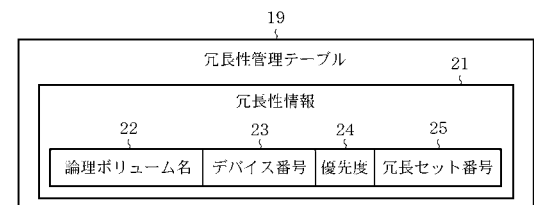
【図1】



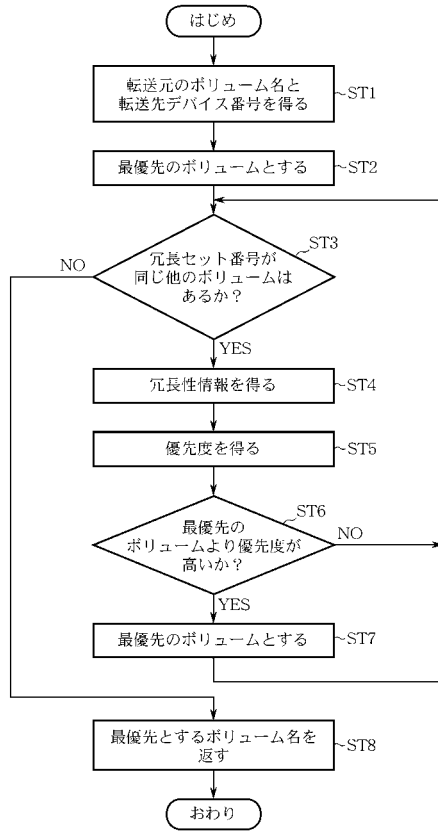
【図2】



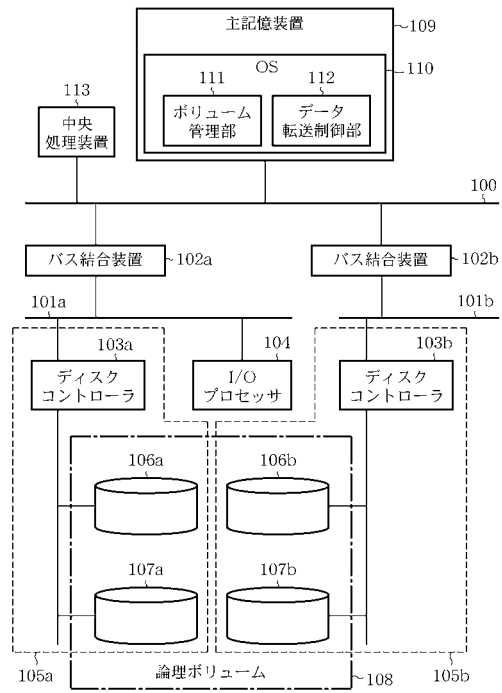
【図3】



【図4】



【図5】



フロントページの続き

(56)参考文献 特開平 1 1 - 3 4 5 1 9 2 (J P , A)
特開平 0 6 - 3 3 2 6 2 5 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G06F 3/06

G06F 12/00

G06F 13/10

G06F 13/14