

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4677443号
(P4677443)

(45) 発行日 平成23年4月27日(2011.4.27)

(24) 登録日 平成23年2月4日(2011.2.4)

(51) Int.Cl.	F 1
G06F 12/00	(2006.01) G06F 12/00 531M
G06F 11/20	(2006.01) G06F 11/20 310C
G06F 3/06	(2006.01) G06F 3/06 304F
	G06F 12/00 533J

請求項の数 9 (全 16 頁)

(21) 出願番号	特願2007-512162 (P2007-512162)	(73) 特許権者 390009531 インターナショナル・ビジネス・マシーンズ・コーポレーション INTERNATIONAL BUSINESS MACHINES CORPORATION アメリカ合衆国 10504 ニューヨーク 州 アーモンク ニュー オーチャード ロード
(86) (22) 出願日	平成17年4月25日 (2005.4.25)	
(65) 公表番号	特表2007-537519 (P2007-537519A)	
(43) 公表日	平成19年12月20日 (2007.12.20)	
(86) 國際出願番号	PCT/EP2005/051845	
(87) 國際公開番号	W02005/109201	
(87) 國際公開日	平成17年11月17日 (2005.11.17)	
審査請求日	平成20年3月26日 (2008.3.26)	
(31) 優先権主張番号	0410540.9	
(32) 優先日	平成16年5月12日 (2004.5.12)	(74) 代理人 100108501 弁理士 上野 剛史
(33) 優先権主張国	英國 (GB)	(74) 代理人 100112690 弁理士 太佐 種一
早期審査対象出願		(74) 代理人 100091568 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】データ・ストレージ・システムおよび複数の異機種ストレージ・コントローラに渡って書き込みセットおよび書き込みセットの境界を保持するための方法

(57) 【特許請求の範囲】

【請求項 1】

主要局(100)における第1のデータ記憶部(106a)に前記主要局の現在の書き込みセットの一部としてデータを記憶するように動作する第1のストレージ・コントローラ(104a)と、

前記主要局における第2のデータ記憶部(106b)に前記現在の書き込みセットの一部としてデータを記憶するように動作する第2のストレージ・コントローラ(104c)と、

書き込みセットの境界に渡って從属書き込みの整合性を維持しつつ、各前記コントローラ(104a、104c)に新規の書き込みセットの識別子を提供するように動作する制御ノード(110)とを備え、

前記第1のストレージ・コントローラ(104a)は、フリーズ/解凍(F/T)ストレージ・コントローラであり、前記第2のストレージ・コントローラ(104c)は、ポーリング・ストレージ・コントローラであり、

前記制御ノード(110)は、前記現在の書き込みセットを保持し、かつ、前記ポーリング・ストレージ・コントローラ(104c)によるポーリング要求に応じて前記現在の書き込みセットを前記ポーリング・ストレージ・コントローラへ報告し、またフリーズ・コマンドを各前記F/Tストレージ・コントローラ(104a)へ送付することにより、前記F/Tストレージ・コントローラ(104a)に対して新規の書き込みセットへの変更を行わせるように動作する、データ・ストレージ・システム。

【請求項 2】

前記 F / T ストレージ・コントローラ (104a) は、フリーズ・コマンドを受信したことにより、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作する、請求項 1 に記載のシステム。

【請求項 3】

前記 F / T ストレージ・コントローラ (104a) は、解凍コマンドを受信したことにより、書き込み動作の実行を継続するように動作する、請求項 2 に記載のシステム。

【請求項 4】

前記制御ノード (110) は、フリーズ・コマンドを各前記 F / T ストレージ・コントローラ (104a) へ送付した後に、各前記 F / T ストレージ・コントローラから前記フリーズ・コマンド受信の受領確認を受信したことにより、前記現在の書き込みセットを前記新規の書き込みセットに変更するように動作する、請求項 1 から 3 のいずれか 1 つに記載のシステム。

10

【請求項 5】

前記制御ノード (110) は、前記新規の書き込みセットへ変更した後に、解凍コマンドを前記 F / T ストレージ・コントローラ (104a) へ送付するように動作する、請求項 1 から 4 のいずれか 1 つに記載のシステム。

【請求項 6】

前記 F / T ストレージ・コントローラ (104a) は、解凍コマンドを受信したことにより、書き込み動作の要求者に対して当該書き込み動作を受領確認するように動作し、解凍コマンドが前記 F / T ストレージ・コントローラによって受信された後に前記書き込み受領確認が送付された前記書き込み動作はいずれも、前記新規の書き込みセットの一部として書き込まれる、請求項 1 から 5 のいずれか 1 つに記載のシステム。

20

【請求項 7】

前記 F / T ストレージ・コントローラ (104a) および前記ポーリング・ストレージ・コントローラ (104c) は、前記現在の書き込みセットまたは前記新規の書き込みセットの一部として前記主要局 (100) における前記第 1 のデータ記憶部 (106a) または前記第 2 のデータ記憶部 (106b) に書き込まれたデータを、2 次局 (102) における対応する遠隔ストレージ・コントローラに対して非同期的にコピーするように動作する、請求項 1 から 6 のいずれか 1 つに記載のシステム。

30

【請求項 8】

前記ポーリング・ストレージ・コントローラ (104c) は、前記主要局 (100) における前記第 2 のデータ記憶部 (106b) に前記現在の書き込みセットの一部としてデータを記憶する前に、前記制御ノード (110) をポーリングして、前記現在の書き込みセットの識別子を受信するように動作する、請求項 1 から 7 のいずれか 1 つに記載のシステム。

【請求項 9】

複数の異機種ストレージ・コントローラに渡って書き込みセットおよび書き込みセットの境界を保持するための方法であって、

主要局における制御ノードによって、当該主要局における少なくとも 1 つの F / T ストレージ・コントローラへフリーズ・コマンドを送付することにより、前記 F / T ストレージ・コントローラを、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作させるステップと、

40

各 F / T ストレージ・コントローラからの前記フリーズ・コマンドの受領確認を受信すると、前記制御ノードによって、前記主要局の現在の書き込みセットを新規の書き込みセットへ変更するステップであって、前記現在の書き込みセットは、前記主要局における少なくとも 1 つのポーリング・ストレージ・コントローラに対して、当該ポーリング・ストレージ・コントローラによるポーリング要求に応じて提供可能であるステップと、

前記制御ノードによって、解凍コマンドを前記 F / T ストレージ・コントローラへ送付することにより、前記 F / T ストレージ・コントローラを、前記要求者に対して前記書き

50

込み動作を受領確認するように動作させるステップであって、前記解凍コマンドが前記 F/Tストレージ・コントローラによって受信された後に受領確認が送られた前記書き込み動作はいずれも、前記新規の書き込みセットの一部として書き込まれるステップとを含む、方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的に、コンピュータ・システムのための災害復旧戦略に関し、より特定的には、2次局におけるストレージの非同期更新をサポートする書き込みセット境界管理に関する。 10

【背景技術】

【0002】

コンピュータシステムのための災害復旧戦略は、典型的には、主要局から幾分離れた場所にある2次局に記憶されたデータをコピーすることを伴う。主要コピーおよび2次コピー間でのコピーは、同期的または非同期的に行われる。コピーが同期的に行われる場合、更新が主要コピーに書き込まれるたびに、当該更新は、2次コピーが書き込まれるべき2次局にも送られる。2次局が主要局に対して2次コピーが更新されたことを通知した後のみ、主要局は主要コピーに対する更新を受領確認し、次の更新を書き込む準備が整うことになる。よって、更新は、主要コピーおよび2次コピーへ同じ順序で書き込まれる。コピーが非同期的に行われる場合、複数の更新が、主要コピーに対して書き込まれて、任意の更新が2次局へ送られる前に、受領確認される場合がある。これは、主要コピーが2次コピーから独立して保持された状態で行われる。更新は、典型的には本明細書において「カラー」と称する書き込みのセットとして、周期的に2次局へ送られて、2次コピーに書き込まれるが、これは、主要コピーに書き込まれたのと必ずしも同じ順序で行われるわけではない。 20

【0003】

単一のカラーが主要局における複数のストレージ・コントローラに渡って保持される場合、次のカラーに切り換える際に、すべてのストレージ・コントローラが次のカラーへ同等に切り換えて、カラー境界を渡る「従属書き込み」の整合性を維持することが必要である。例えば、データベース更新トランザクションのための以下の典型的な従属書き込みのシーケンスを考える。 30

【0004】

1. データベース更新が生じようとする事を示す、データベース・ログを更新するための第1の書き込みを実行し、その後、

【0005】

2. データベースを更新するための第2の書き込みを実行し、最後に

【0006】

3. データベース更新がうまく完了したことを示す、データベース・ログを更新するための第3の書き込みを実行する。

【0007】

この場合、これらの従属書き込みは、すべて同じカラーに属するか、または、カラー境界を越える場合は、前の(1つ以上の)書き込みは古いカラーに属し、後の(1つ以上の)書き込みは新しいカラーに属することが必須である。この例において書き込み1、2、および3は、それぞれ異なるストレージ・コントローラによって書き込まれると仮定する。もし、書き込み1および3がカラー・グループである「赤」の一部として書き込まれ、そして書き込み2が次のカラー・グループである「青」の一部として書き込まれるならば、「赤」グループが2次コピーに書き込まれた後で且つ「青」グループが2次コピーに書き込まれる前に主要コピーが失われる場合には、2次コピーのデータベース・ログは、実際にはデータベースは更新されていないのに、2次書き込みが生じたと誤って示すことになる。 40

【0008】

主要局における複数のストレージ・コントローラに渡ってカラーとカラー境界とを保持するための手法の一つでは、書き込みをカラーと関連付ける前に、各ストレージ・コントローラは、現在のカラーを保持するカラー制御ノードに対してポーリングを行う。カラー制御ノードは、ストレージ・コントローラに対して現在のカラーを通告し、ストレージ・コントローラは、当該カラーの一部として、書き込みを行う。

【0009】

主要局における複数のストレージ・コントローラに渡ってカラーとカラー境界とを保持するための他の手法では、カラー制御ノードが現在のカラーを保持するが、ストレージ・コントローラは、カラー制御ノードに対して現在のカラーをポーリングしない。むしろ、カラー制御ノードが新規のカラーを形成したい場合には、「フリーズ」コマンドを、すべてのストレージ・コントローラへ送付して、新規のカラーを示す。ストレージ・コントローラが「フリーズ」コマンドを受信すると、要求者からの書き込み動作の受領確認を保留する。その後、ストレージ・コントローラは、「フリーズ」コマンドの受領確認をカラー制御ノードへ送付する。カラー制御ノードがすべてのストレージ・コントローラから受領確認を受信すると、「解凍」コマンドをストレージ・コントローラへ送付する。ストレージ・コントローラが「解凍」コマンドを受信すると、要求者に対して書き込み動作の受領確認を行い、「フリーズ」コマンドを受信後に受領確認が送られたすべての書き込み動作は、新規のカラーの一部として書き込まれる。

【0010】

10

ストレージ・システムが、システム・スケーリングまたは移行をサポートするため、ポーリング・ストレージ・コントローラと、フリーズ／解凍ストレージ・コントローラとの両方を使用すれば有利になるが、主要局において複数の異機種ストレージ・コントローラに渡ってカラーとカラー境界とを保持するための方法が必要である。

【発明の開示】

【課題を解決するための手段】

【0011】

本発明は、2次局におけるストレージの非同期更新をサポートする異機種ストレージ・コントローラのための書き込みセット境界管理を開示する。

【0012】

30

本発明の一局面において、データ・ストレージ・システムが提供され、当該データ・ストレージ・システムは、第1のカラー・ポリシーを有し、主要局における第1のデータ記憶部に主要局の現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのストレージ・コントローラと、第2のカラー・ポリシーを有し、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのストレージ・コントローラと、カラー境界に渡って従属書き込みの整合性を維持しつつ、各コントローラに新規のカラー情報を提供するように動作するカラー制御ノードとを含む。

【0013】

本発明の他の局面において、第1のカラー・ポリシーは、フリーズ／解凍（F／T）ストレージ・コントローラのものであり、第2のカラー・ポリシーは、ポーリング・ストレージ・コントローラのものである。

【0014】

本発明の他の局面において、主要局における第1のデータ記憶部に主要局における現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのフリーズ／解凍（F／T）ストレージ・コントローラと、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのポーリング・ストレージ・コントローラと、現在のカラーを保持し、且つ、ポーリング・ストレージ・コントローラによるポーリング要求に応じて現在のカラーをポーリング・ストレージ・コントローラへ報告し、F／Tストレージ・コントローラに対して新規のカラーへの変更をおこな

40

50

わせるように動作するカラー制御ノードとを含む、データ・ストレージ・システムが提供される。

【0015】

本発明の他の局面において、カラー制御ノードは、フリーズ・コマンドを各F/Tストレージ・コントローラへ送付するように動作する。

【0016】

本発明の他の局面において、F/Tストレージ・コントローラは、フリーズ・コマンドを受信したことに応じて、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作する。

【0017】

本発明の他の局面において、F/Tストレージ・コントローラは、解凍コマンドを受信したことに応じて、書き込み動作の実行を継続するように動作する。

10

【0018】

本発明の他の局面において、カラー制御ノードは、フリーズ・コマンドを各F/Tストレージ・コントローラへ送付した後に、各F/Tストレージ・コントローラからフリーズ・コマンド受信の受領確認を受信したことに応じて、現在のカラーを新規のカラーに変更するように動作する。

【0019】

本発明の他の局面において、カラー制御ノードは、新規のカラーへ変更した後に、解凍コマンドをF/Tストレージ・コントローラへ送付するように動作する。

20

【0020】

本発明の他の局面において、F/Tストレージ・コントローラは、解凍コマンドを受信したことに応じて、書き込み動作の要求者に対して当該書き込み動作を受領確認するように動作し、解凍コマンドがF/Tストレージ・コントローラによって受信された後に書き込み受領確認が送付された書き込み動作はいずれも、新規のカラーの一部として書き込まれる。

【0021】

本発明の他の局面において、F/Tストレージ・コントローラおよびポーリング・ストレージ・コントローラは、2次局における対応する遠隔ストレージ・コントローラに対して、いずれかのカラーの一部として書き込まれたデータを非同期的にコピーするように動作する。

30

【0022】

本発明の他の局面において、ポーリング・ストレージ・コントローラは、カラー制御ノードをポーリングして、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶する前に、現在のカラーを受信するように動作する。

【0023】

本発明の他の局面において、複数の異機種ストレージ・コントローラに渡ってカラーおよびカラー境界を保持するための方法が提供され、当該方法は、主要局における少なくとも1つのF/Tストレージ・コントローラへフリーズ・コマンドを送付することによって、F/Tストレージ・コントローラを、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作させるステップと、各F/Tストレージ・コントローラからのフリーズ・コマンドの受領確認を受信すると、主要局の現在のカラーを新規のカラーへ変更するステップであって、現在のカラーは、主要局における少なくとも1つのポーリング・ストレージ・コントローラに対して、ポーリング・ストレージ・コントローラによるポーリング要求に応じて提供可能であるステップと、解凍コマンドをF/Tストレージ・コントローラへ送付することによって、F/Tストレージ・コントローラを、書き込み動作の要求者に対して当該書き込み動作を受領確認するように動作させるステップであって、フリーズ・コマンドがF/Tストレージ・コントローラによって受信された後に受領確認が送られた書き込み動作はいずれも、新規のカラーの一部として書き込まれるステップとを含む。

40

50

【0024】

本発明の他の局面において、データ・ストレージ・システムであって、主要局における第1のデータ記憶部に主要局の現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのフリーズ／解凍(F/T)ストレージ・コントローラと、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのポーリング・ストレージ・コントローラと、現在のカラーを保持し且つポーリング・ストレージ・コントローラによるポーリング要求に応じて、ポーリング・ストレージ・コントローラに対して現在のカラーを報告するように動作するポーリング・カラー制御ノードと、F/Tストレージ・コントローラおよびポーリング・カラー制御ノードの何れかに対して新規のカラーへの変更をさせるように動作するF/Tカラー制御ノードとを含む、データ・ストレージ・システムが提供される。 10

【0025】

本発明の他の局面において、F/Tカラー制御ノードは、フリーズ・コマンドをF/Tストレージ・コントローラへ送付するように動作する。

【0026】

本発明の他の局面において、F/Tストレージ・コントローラは、フリーズ・コマンドを受信したことに応じて、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作する。

【0027】

本発明の他の局面において、F/Tカラー制御ノードは、各F/Tストレージ・コントローラからのフリーズ・コマンドの受信の受領確認を受信したことに応じて、現在のカラーを新規のカラーへ変更するためのコマンドをポーリング・カラー制御ノードへ送付するように動作する。 20

【0028】

本発明の他の局面において、カラーを変更するコマンドは、フリーズ・コマンドである。

【0029】

本発明の他の局面において、カラーを変更するコマンドは、解凍コマンドである。

【0030】

本発明の他の局面において、ポーリング・カラー制御ノードは、カラー変更コマンドの受信に応じて、現在のカラーを新規のカラーへ変更するように動作する。 30

【0031】

本発明の他の局面において、ポーリング・カラー制御ノードは、新規のカラーへの変更の後、カラー変更受領確認をF/Tカラー制御ノードへ送付するように動作する。

【0032】

本発明の他の局面において、F/Tカラー制御ノードは、カラー変更受領確認を受信したことに応じて、解凍コマンドをF/Tストレージ・コントローラへ送付するように動作する。

【0033】

本発明の他の局面において、F/Tストレージ・コントローラは、解凍コマンドを受信したことに応じて、書き込み動作の要求者に対して書き込み動作を受領確認するように動作し、解凍コマンドがF/Tストレージ・コントローラによって受信された後に書き込み受領確認が送付された任意の書き込み動作は、新規のカラーの一部として書き込まれたとみなされる。 40

【0034】

本発明の他の局面において、F/Tストレージ・コントローラおよびポーリング・ストレージ・コントローラは、2次局における対応する遠隔ストレージ・コントローラに対して、いずれかのカラーの一部として書き込まれたデータを非同期的にコピーするように動作する。

【0035】

本発明の他の局面において、ポーリング・ストレージ・コントローラは、ポーリング・カラー制御ノードをポーリングして、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶する前に、現在のカラーを受信するように動作する。

【0036】

本発明の他の局面において、複数の異機種ストレージ・コントローラに渡ってカラーおよびカラー境界を保持するための方法が提供され、当該方法は、主要局における少なくとも1つのF/Tストレージ・コントローラへフリーズ・コマンドを送付することによって、F/Tストレージ・コントローラを、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作させるステップと、各F/Tストレージ・コントローラからのフリーズ・コマンドの受領確認を受信すると、主要局におけるポーリング・カラー制御ノードの現在のカラーを新規のカラーへ変更するステップであって、現在のカラーは、ポーリング・ストレージ・コントローラによるポーリング要求に応じて、主要局における少なくとも1つのポーリング・ストレージ・コントローラに対して提供可能であるステップと、カラー変更の受信の受領確認を受信した後に、解凍コマンドをF/Tストレージ・コントローラへ送付することによって、F/Tストレージ・コントローラが書き込み動作を要求者に対して受領確認するように動作させるステップであって、解凍コマンドがF/Tストレージ・コントローラによって受信された後に受領確認が送られた書き込み動作はいずれも、新規のカラーの一部として書き込まれるステップとを含む。

【0037】

本発明の他の局面において、変更するステップは、解凍コマンドをポーリング・カラー制御ノードへ送付するステップを含む。

【0038】

本発明の他の局面において、変更するステップは、解凍コマンドをポーリング・カラー制御ノードへ送付するステップを含む。

【0039】

本発明の他の局面において、F/Tカラー制御ノードは、解凍コマンドをF/Tストレージ・コントローラおよびポーリング・カラー制御ノードへ送付するように動作する。

【0040】

本発明の他の局面において、F/Tカラー制御ノードは、各F/Tストレージ・コントローラおよびポーリング・カラー制御ノードからのフリーズ・コマンドの受信の受領確認を受信したことに応じて、現在のカラーを新規のカラーへ変更するための解凍コマンドをポーリング・カラー制御ノードへ送付するように動作する。

【0041】

本発明の他の局面において、複数の異機種ストレージ・コントローラに渡ってカラーおよびカラー境界を保持するための方法が提供され、当該方法は、主要局における少なくとも1つのF/Tストレージ・コントローラおよびポーリング・カラー制御ノードへフリーズ・コマンドを送付することによって、F/Tストレージ・コントローラを、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作させるステップと、各F/Tストレージ・コントローラおよびポーリング・カラー制御ノードからのフリーズ・コマンドの受領確認を受信すると、主要局におけるポーリング・カラー制御ノードの現在のカラーを新規のカラーへ変更するステップであって、現在のカラーは、ポーリング・ストレージ・コントローラによるポーリング要求に応じて、主要局における少なくとも1つのポーリング・ストレージ・コントローラに対して提供可能であるステップと、カラー変更の受信の受領確認を受信した後に、解凍コマンドをF/Tストレージ・コントローラへ送付することによって、F/Tストレージ・コントローラが書き込み動作を要求者に対して受領確認するように動作させるステップであって、解凍コマンドがF/Tストレージ・コントローラによって受信された後に受領確認が送られた書き込み動作はいずれも、新規のカラーの一部として書き込まれるステップとを含む。

【0042】

本発明の他の局面において、変更するステップは、解凍コマンドをポーリング・カラー

10

20

30

40

50

制御ノードへ送付するステップを含む。

【0043】

本発明の他の局面において、F/Tカラー制御ノードは、ポーリング・カラー制御ノードからのフリーズ・コマンドの受信の受領確認を受信したことに応じて、解凍コマンドをポーリング・カラー制御ノードへ送付するように動作する。

【0044】

本発明の他の局面において、F/Tカラー制御ノードは、ポーリング・カラー制御ノードからの解凍コマンドの受信の受領確認を受信したことに応じて、解凍コマンドをF/Tストレージ・コントローラへ送付するように動作する。

【0045】

本発明の他の局面において、ポーリング・カラー制御ノードは、F/Tカラー制御ノードからフリーズ・コマンドおよび解凍コマンドの何れかを受信したことに応じて、現在のカラーを変更するように動作する。

【0046】

本発明の他の局面において、複数の異機種ストレージ・コントローラに渡ってカラーおよびカラー境界を保持するための方法が提供され、当該方法は、主要局における少なくとも1つのF/Tストレージ・コントローラへ第1のフリーズ・コマンドを送付することによって、F/Tストレージ・コントローラを、書き込み動作の要求者からの当該書き込み動作の受領確認を保留するように動作させるステップと、F/Tカラー制御ノードが各F/Tストレージ・コントローラから第1のフリーズ・コマンドの受信の受領確認を受信した後に、ポーリング・カラー制御ノードへ第2のフリーズ・コマンドを送付するステップと、F/Tカラー制御ノードがポーリング・カラー制御ノードから第2のフリーズ・コマンドの受信の受領確認を受信した後に、ポーリング・カラー制御ノードへ第1の解凍コマンドを送付するステップと、第2のフリーズ・コマンドおよび第1の解凍コマンドの何れかに応じて、主要局におけるポーリング・カラー制御ノードの現在のカラーを新規のカラーへ変更するステップと、第1の解凍コマンドの受信の受領確認を受信した後に、F/Tストレージ・コントローラへ第2の解凍コマンドを送付することによって、F/Tストレージ・コントローラが書き込み動作を要求者に対して受領確認するように動作させるステップであって、第2の解凍コマンドがF/Tストレージ・コントローラによって受信された後に受領確認が送られた書き込み動作はいずれも、新規のカラーの一部として書き込まれるステップとを含む。

【0047】

本発明の他の局面において、データ・ストレージ・システムを配備するための方法が提供され、当該方法は、第1のカラー・ポリシーを有し、主要局における第1のデータ記憶部に主要局の現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのストレージ・コントローラを設置するステップと、第2のカラー・ポリシーを有し、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶するように動作する少なくとも1つのストレージ・コントローラを設置するステップと、カラー境界に渡って従属書き込みの整合性を維持しつつ、各コントローラに新規のカラー情報を提供するように動作するカラー制御ノードを設置するステップとを含む。

【0048】

本発明の他の局面において、コンピュータ読み取り可能な媒体上に実施されたコンピュータ・プログラムが提供され、当該コンピュータ・プログラムは、主要局における第1のデータ記憶部に主要局の現在のカラーの一部としてデータを記憶するための第1のカラー・ポリシーを有する少なくとも1つのストレージ・コントローラを制御するように動作する第1のコード・セグメントと、主要局における第2のデータ記憶部に現在のカラーの一部としてデータを記憶するための第2のカラー・ポリシーを有する少なくとも1つのストレージ・コントローラを制御するように動作する第2のコード・セグメントと、カラー境界に渡って従属書き込みの整合性を維持しつつ、各コントローラに新規のカラー情報を提供するためのカラー制御ノードを制御するように動作する第3のコード・セグメントとを

10

20

30

40

50

含む。

【0049】

本明細書および請求項を全体を通じて、現在のカラーを受信するためのポーリングについての言及は、任意の既知の情報表現手法を使用してカラー情報を受信することをいい、カラーが、カラーに関連付けられるシリアル番号のような数値形式で表現されることを含むものである。

【0050】

本発明は、添付の図面と共に以下の詳細な説明から、さらに十分に理解および評価されるだろう。

【発明を実施するための最良の形態】

10

【0051】

今から図1を参照する。図1は、本発明の好ましい実施形態に従って構築され且つ動作する、主要および2次データ・ストレージ・システムの簡略ブロック図である。図1のシステムにおいて、主要局100は、ストレージ・コントローラ104a, 104b, および104cを含み、それぞれ、記憶部106a, 106b, および106cに結合されている。ストレージ・コントローラ104a, 104b, および104cは、ホスト・システム(図示せず)からの記憶部106a, 106b, および106c内のボリュームに対する入力/出力(I/O)アクセスを管理する。ストレージ・コントローラ104a, 104b, および104cは、2次局102における対応する遠隔ストレージ・コントローラ108a, 108b, および108cに対して、接続112a, 112b, および112cを介してデータを非同期にコピーするように構成されるのが好ましい。ストレージ・コントローラ104a, 104b, および104cは、本明細書において「カラー」と称されるデータ書き込みのセットを、遠隔ストレージ・コントローラ108a, 108b, および108cに対して、従来の手法に従って周期的に送信する。遠隔ストレージ・コントローラ108a, 108b, および108cは、それぞれに付属した記憶部114a, 114b, および114cにデータを書き込む。図1は、主要局100および2次局102における一定数の構成要素、例えばそれぞれ3つのストレージ・コントローラおよびストレージを示したが、任意の数のストレージ・コントローラおよび記憶部を使用してもよいことが理解される。

20

【0052】

30

2次局102は、主要局100から数百から数千マイル(1マイルは約1.6キロ)離れていてもよい。ストレージ・コントローラの対の間の接続112a, 112b, および112cは、遠距離電気通信搬送波によって維持される専用光ファイバ・ネットワークまたは接続、チャンネル・エクステンダを有する長距離光ファイバ接続のような専用回線を備えてもよく、またはストレージ・エリア・ネットワーク(SAN)、主要エリア・ネットワーク(LAN)、またはインターネットのようなネットワークを介して延長されてもよい。

【0053】

主要局100におけるホスト・システム(図示せず)は、記憶部106a, 106b, および106c内のボリュームに対して入力/出力(I/O)動作を行ってもよい。ホスト・システムは、ストレージ・コントローラ104a, 104b, および104cと、当該技術において周知の任意のネットワークまたは接続を介して通信を行ってもよい。

40

【0054】

ストレージ・コントローラ104a, 104b, 104c, 108a, 108b, および108cは、好ましくは、プロセッサ集合体を含み、IBM社のEnterprise Storage Server(ESS)(IBM社の登録商標)、3990ストレージ・コントローラのような、当該技術において周知の任意のストレージ・コントローラまたはサーバを備えてもよい。記憶部106a, 106b, 106c, 114a, 114b, および114cは、単なるディスクの集合体(JBOD)、または独立ディスクの冗長アレイ(RAID)アレイなどの記憶装置のアレイを備えてよい。

50

【0055】

主要局100におけるストレージ・コントローラ104a, 104b, および104cは、ネットワーク118を介して、カラー制御ノード110と通信するのが好ましい。カラー制御ノード110は、同一のデータ・ストレージ局内で動作する複数のストレージ・コントローラと共に機能するのが好ましいが、同一のデータ・ストレージ局内において、すべてのストレージ・コントローラが同一のカラー・ポリシーを有するわけではない。例えば、カラー制御ノード110は、ポーリング・ストレージ・コントローラがポーリングするための、本明細書において「カラー」と称する現在の書き込みセットを識別する識別子を保持する一方で、「フリーズ」コマンドをフリーズ／解凍(F/T)ストレージ・コントローラへ発行することによって、F/Tストレージ・コントローラに対して次の書き込みセット／カラーへの変更を行わせてもよい。主要局100において、任意の数および組み合わせのポーリング・ストレージ・コントローラおよびF/Tストレージ・コントローラを使用してもよく、例えば、ストレージ・コントローラ104aおよび104bが、カラー制御ノード110が「フリーズ」コマンドを発行するまで現在のカラーを保持するF/Tストレージ・コントローラであってもよいし、ストレージ・コントローラ104cが、カラー制御ノード110をポーリングして、書き込みをカラーに関連付ける前に、現在のカラーを要求するポーリング・ストレージ・コントローラであってもよい。ストレージ・コントローラ104a, 104b, および104cは、従来の手法に従って、当該カラーの一部としてストレージに対する書き込みを行う。

【0056】

今度は図2を参照する。図2は、本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。図2の方法は、図1の主要および2次データ・ストレージ・システムを参照して理解されてもよい。図2の方法は、好ましくは、ストレージ・コントローラ104a, 104b, および104cと、カラー制御ノード110とによって実施される。図2の方法において、ポーリング・ストレージ・コントローラ104cが、現在のカラーについて、当該カラーの一部として書き込みを実行する前にカラー制御ノード110に対して現在のカラーについてポーリングするのに対し、F/Tストレージ・コントローラ104aおよび104bは、現在のカラーを保持し、すべての書き込みを当該カラーの一部として実行する。例えば、予め定義されたスケジュールに従って新規のカラーに切り換える直前に、カラー制御ノード110は、新規のカラーを示す「フリーズ」コマンドをF/Tストレージ・コントローラ104aおよび104bへ発行する。「フリーズ」コマンドを受信すると、F/Tストレージ・コントローラ104aおよび104bは、書き込み動作の実行を継続してもよいが、要求者からの書き込み動作の受領確認を保留するか、または、書き込み動作をまとめて中断してもよい。その後、F/Tストレージ・コントローラ104aおよび104bは、「フリーズ」コマンドの受領確認をカラー制御ノード110へ送付する。一旦カラー制御ノード110がすべてのF/Tストレージ・コントローラから受領確認を受信すると、カラー制御ノード110は、そのカラーを変更する。よって、この時点以降、カラー制御ノード110は、新規のカラーを示すことによって、ポーリング・ストレージ・コントローラ104cから受信するポーリングに応答する。カラー制御ノード110は、その後、「解凍」コマンドをF/Tストレージ・コントローラ104aおよび104bへ送付し、「解凍」コマンドを受信すると、F/Tストレージ・コントローラ104aおよび104bは、中断された書き込み動作がもしあればそれを実行して、「フリーズ」コマンドを受信後に実行された書き込み動作の受領確認を、書き込み動作の要求者に対して送付してもよい。ここで、「フリーズ」コマンドを受信後に受領確認が送付されるすべての書き込み動作は、新規のカラーの一部として書き込まれる。

【0057】

今度は図3を参照する。図3は、本発明の好ましい実施形態に従って構築され且つ動作する、主要および2次データ・ストレージ・システムの簡略ブロック図である。図3のシステムは、図1のシステムと実質的に同様であり、重要な例外は、主要局100における

10

20

30

40

50

F / Tストレージ・コントローラ104aおよび104bは、「フリーズ」コマンドをF / Tストレージ・コントローラ104aおよび104bへ発行することによってフリーズ / 解凍(F / T)ストレージ・コントローラに対して次のカラーへの変更を行わせるF / Tカラー制御ノード300と通信するのに対して、ポーリング・ストレージ・コントローラ104cは、ポーリング・カラー制御ノードとして機能してポーリング・ストレージ・コントローラのために現在のカラーを保持するポーリング・カラー制御ノード302と通信するという点である。

【0058】

今度は図4を参照する。図4は、本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。図4の方法は、図3の主要および2次データ・ストレージ・システムを参照して理解されてもよい。図4の方法は、好ましくは、ストレージ・コントローラ104a, 104b, および104cと、ポーリング・カラー制御ノード302と、F / Tカラー制御ノード300とによって実施される。図4の方法において、ポーリング・ストレージ・コントローラ104cが、現在のカラーについて、当該カラーの一部として書き込みを実行する前にポーリング・カラー制御ノード302に対してポーリングするのに対し、F / Tストレージ・コントローラ104aおよび104bは、現在のカラーを保持する。例えば、予め定義されたスケジュールに従って新規のカラーに切り換える直前に、F / Tカラー制御ノード300は、新規のカラーを示す「フリーズ」コマンドをF / Tストレージ・コントローラ104aおよび104bへ発行する。「フリーズ」コマンドを受信すると、F / Tストレージ・コントローラ104aおよび104bは、書き込み動作の実行を継続してもよいが、要求者からの書き込み動作の受領確認を保留するか、または、書き込み動作をまとめて中断してもよい。その後、F / Tストレージ・コントローラ104aおよび104bは、「フリーズ」コマンドの受領確認をF / Tカラー制御ノード300へ送付する。一旦F / Tカラー制御ノード300がすべてのF / Tストレージ・コントローラから「フリーズ」コマンドの受領確認を受信すると、F / Tカラー制御ノード300は、コマンドをポーリング・カラー制御ノード302へ送付して、カラーを変更する。当該コマンドは、標準的な「フリーズ」または「解凍」コマンドの形式であってもよい。一旦ポーリング・カラー制御ノード302がF / Tカラー制御ノード300からコマンドを受信してそのカラーを変更すると、ポーリング・カラー制御ノード302はそのカラーを変更して、受領確認をF / Tカラー制御ノード300へ送付する。よって、この時点以降、ポーリング・カラー制御ノード302は、新規のカラーを示すことによって、ポーリング・ストレージ・コントローラ104cから受信するポーリングに応答する。一旦F / Tカラー制御ノード300がポーリング・カラー制御ノード302からカラーを変更した旨を示す受領確認を受信すると、F / Tカラー制御ノード300は、「解凍」コマンドをF / Tストレージ・コントローラ104aおよび104bへ送付し、「解凍」コマンドを受信すると、F / Tストレージ・コントローラ104aおよび104bは、中断された書き込み動作がもしあればそれを実行して、「フリーズ」コマンドを受信後に実行された書き込み動作の受領確認を、書き込み動作の要求者に対して送付してもよい。ここで、「フリーズ」コマンドを受信後に受領確認が送付されるすべての書き込み動作は、新規のカラーの一部として書き込まれる。また、F / Tカラー制御ノード300は、「解凍」コマンドをポーリング・カラー制御ノード302へ送付してもよく、それは無視されてもよい。

【0059】

今度は図5を参照する。図5は、本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。図5の方法は、図3の主要および2次データ・ストレージ・システムを参照して理解されてもよい。図5の方法は、好ましくは、ストレージ・コントローラ104a, 104b, および104cと、ポーリング・カラー制御ノード302と、F / Tカラー制御ノード300とによって実施される。図5の方法において、ポーリン

10

20

30

40

50

グ・ストレージ・コントローラ 104c が、現在のカラーについて、当該カラーの一部として書き込みを実行する前にポーリング・カラー制御ノード 302 に対してポーリングするのに対し、F/Tストレージ・コントローラ 104a および 104b は、現在のカラーを保持し、すべての書き込みを当該カラーの一部として実行する。例えば、予め定義されたスケジュールに従って新規のカラーに切り換える直前に、F/Tカラー制御ノード 300 は、新規のカラーを示す「フリーズ」コマンドを F/Tストレージ・コントローラ 104a および 104b と共にポーリング・カラー制御ノード 302 へ発行する。「フリーズ」コマンドを受信すると、F/Tストレージ・コントローラ 104a および 104b は、書き込み動作の実行を継続してもよいが、要求者からの書き込み動作の受領確認を保留するか、または、書き込み動作をまとめて中断してもよい。その後、F/Tストレージ・コントローラ 104a および 104b ならびにポーリング・カラー制御ノード 302 は、「フリーズ」コマンドの受領確認を F/Tカラー制御ノード 300 へ送付する。一旦 F/Tカラー制御ノード 300 がすべての F/Tストレージ・コントローラ および ポーリング・カラー制御ノード 302 から受領確認を受信すると、F/Tカラー制御ノード 300 は、コマンドをポーリング・カラー制御ノード 302 へ送付して、カラーを変更する。当該コマンドは、標準的な「解凍」コマンドの形式であってもよい。一旦 ポーリング・カラー制御ノード 302 が F/Tカラー制御ノード 300 からコマンドを受信してそのカラーを変更すると、ポーリング・カラー制御ノード 302 はそのカラーを変更して、受領確認を F/Tカラー制御ノード 300 へ送付する。よって、この時点以降、ポーリング・カラー制御ノード 302 は、新規のカラーを示すことによって、ポーリング・ストレージ・コントローラ 104c から受信するポーリングに応答する。一旦 F/Tカラー制御ノード 300 がポーリング・カラー制御ノード 302 からカラーを変更した旨を示す受領確認を受信すると、F/Tカラー制御ノード 300 は、「解凍」コマンドを F/Tストレージ・コントローラ 104a および 104b へ送付し、「解凍」コマンドを受信すると、F/Tストレージ・コントローラ 104a および 104b は、中断された書き込み動作がもしあればそれを実行して、「フリーズ」コマンドを受信後に実行された書き込み動作の受領確認を、書き込み動作の要求者に対して送付してもよい。ここで、「フリーズ」コマンドを受信後に受領確認が送付されるすべての書き込み動作は、新規のカラーの一部として書き込まれる。

【0060】

今度は図 6 を参照する。図 6 は、本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。図 6 の方法は、図 3 の主要および 2 次データ・ストレージ・システムを参照して理解されてもよい。図 6 の方法は、好ましくは、ストレージ・コントローラ 104a, 104b, および 104c と、ポーリング・カラー制御ノード 302 と、F/Tカラー制御ノード 300 とによって実施される。図 6 の方法において、ポーリング・ストレージ・コントローラ 104c が、現在のカラーについて、当該カラーの一部として書き込みを実行する前にポーリング・カラー制御ノード 302 に対してポーリングするのに対し、F/Tストレージ・コントローラ 104a および 104b は、現在のカラーを保持し、すべての書き込みを当該カラーの一部として実行する。例えば、予め定義されたスケジュールに従って新規のカラーに切り換える直前に、F/Tカラー制御ノード 300 は、新規のカラーを示す「フリーズ」コマンドを F/Tストレージ・コントローラ 104a および 104b へ発行する。「フリーズ」コマンドを受信すると、F/Tストレージ・コントローラ 104a および 104b は、書き込み動作の実行を継続してもよいが、要求者からの書き込み動作の受領確認を保留するか、または、書き込み動作をまとめて中断してもよい。その後、F/Tストレージ・コントローラ 104a および 104b は、「フリーズ」コマンドの受領確認を F/Tカラー制御ノード 300 へ送付する。一旦 F/Tカラー制御ノード 300 がすべての F/Tストレージ・コントローラ から受領確認を受信すると、F/Tカラー制御ノード 300 は、「フリーズ」コマンドをポーリング・カラー制御ノード 302 へ送付する。一旦 ポーリング・カラー制御ノード 302 が F/Tカラー制

10

20

30

40

50

御ノード300から「フリーズ」コマンドを受信すると、ポーリング・カラー制御ノード302は、そのカラーを変更してもよい。その後、ポーリング・カラー制御ノード302は、受領確認をF/Tカラー制御ノード300へ送付する。一旦F/Tカラー制御ノード300がポーリング・カラー制御ノード302から受領確認を受信すると、F/Tカラー制御ノード300は、「解凍」コマンドをポーリング・カラー制御ノード302へ送付する。一旦ポーリング・カラー制御ノード302がF/Tカラー制御ノード300から「解凍」コマンドを受信すると、ポーリング・カラー制御ノード302は、そのカラーを変更してもよい。よって、ポーリング・カラー制御ノード302は、「フリーズ」コマンドまたは「解凍」コマンドの何れかを受信後にそのカラーを変更してもよい。その後、ポーリング・カラー制御ノード302は、「解凍」コマンドの受領確認をF/Tカラー制御ノード300へ送付する。一旦F/Tカラー制御ノード300がポーリング・カラー制御ノード302から「解凍」コマンドの受領確認を受信すると、F/Tカラー制御ノード300は、「解凍」コマンドをF/Tストレージ・コントローラ104aおよび104bへ送付し、「解凍」コマンドを受信すると、F/Tストレージ・コントローラ104aおよび104bは、中断された書き込み動作がもしあればそれを実行して、「フリーズ」コマンドを受信後に実行された書き込み動作の受領確認を、書き込み動作の要求者に対して送付してもよい。ここで、「フリーズ」コマンドを受信後に受領確認が送付されるすべての書き込み動作は、新規のカラーの一部として書き込まれる。

【0061】

本明細書において説明された任意の方法の1つ以上のステップは、本発明の真の精神および範囲を逸脱することなく、省略されてもよく、または、図示の順序とは異なる順序で実行されてもよいことが理解される。

【0062】

本明細書において開示された方法および装置を、特定のコンピュータ・ハードウェアまたはソフトウェアを参照して説明してきたが、本明細書において説明した方法および装置は、従来の手法を使用して、コンピュータ・ハードウェアまたはソフトウェアにおいて容易に実施されてもよいことが理解される。

【0063】

本発明を1つ以上の特定の実施形態を参照して説明してきたが、本説明は、本発明を全体として例示することを意図したものであり、本発明を図示の実施形態に限定するよう解釈されることを意図したものではない。本明細書において特に図示されていないが、様々な変更が当業者に想起されてもよく、それでもやはり、これらの様々な変更は、本発明の真の精神および範囲内のものであることが理解される。

【図面の簡単な説明】

【0064】

【図1】本発明の好ましい実施形態に従って構築され且つ動作する、主要および2次データ・ストレージ・システムの簡略ブロック図である。

【図2】本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。

【図3】本発明の好ましい実施形態に従って構築され且つ動作する、主要および2次データ・ストレージ・システムの簡略ブロック図である。

【図4】本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。

【図5】本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである。

【図6】本発明の好ましい実施形態に従って動作する、複数の異機種ストレージ・コントローラに渡るカラーおよびカラー境界を保持するための方法の簡略フローチャートである

10

20

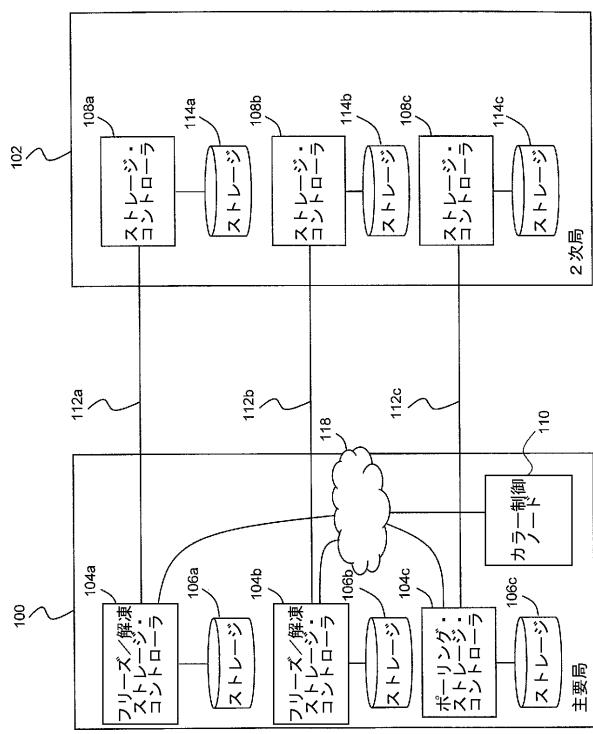
30

40

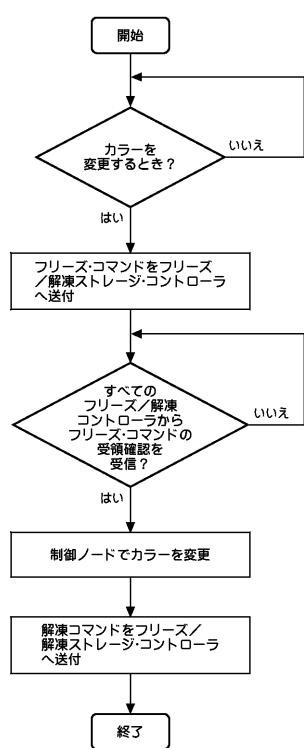
50

。

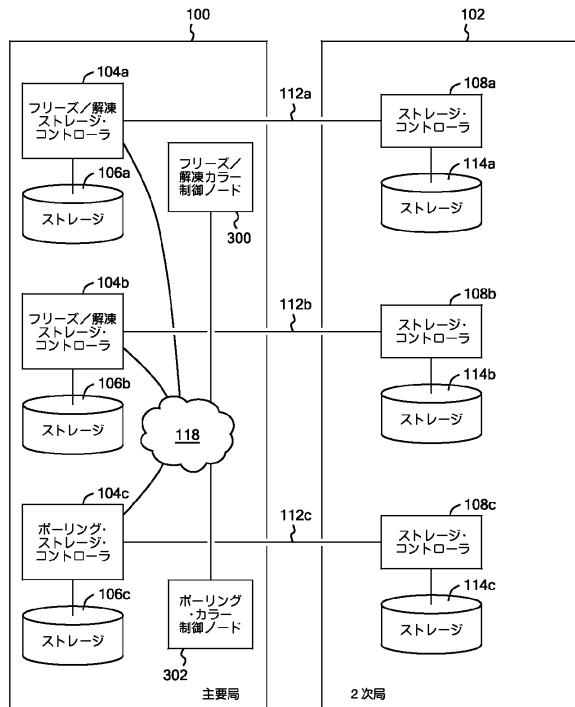
【図1】



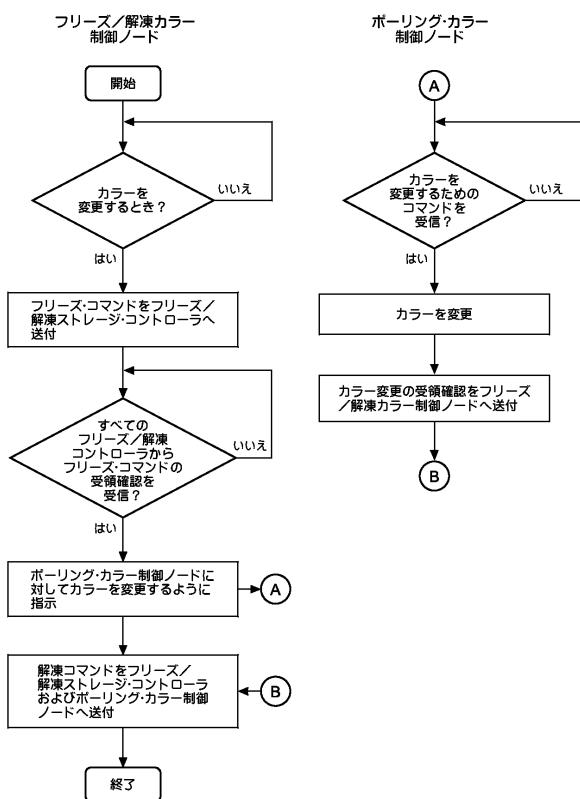
【図2】



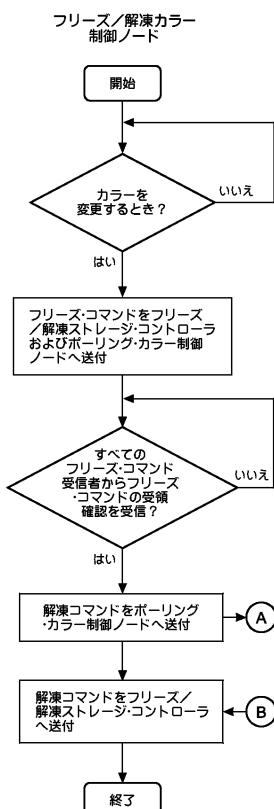
【図3】



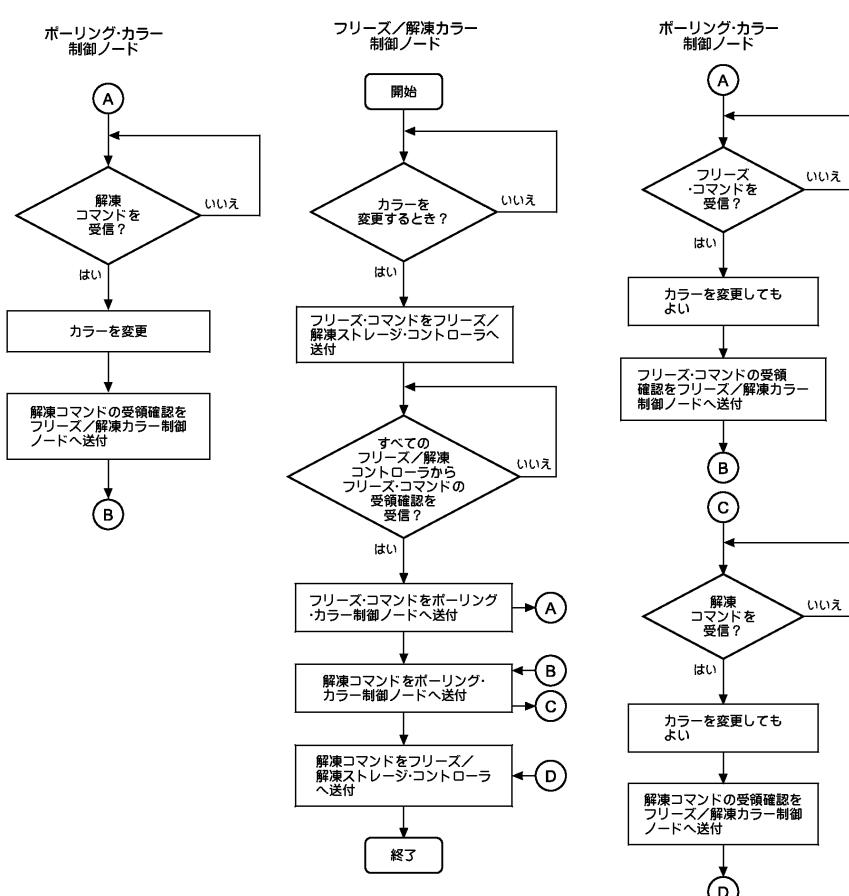
【 図 4 】



【 図 5 】



【図6】



フロントページの続き

(74)代理人 100086243
弁理士 坂口 博

(72)発明者 ベンドール、シラ
イスラエル国 ギバ・シュメール ヨニ・ネタンヤフ・ストリート 30

(72)発明者 バターワース、ヘンリー
イギリス国 S 0 5 3 5 R P イーストレイ・ハンプシャー チャンドラーズ・フォード ハー
トフィールド・ロード 17

(72)発明者 クレディ、アミル
イスラエル国 テルアビブ・ヤツファ 6 4 1 6 1 ブローク・ストリート 17

(72)発明者 ニッサンメシング、オリト
イスラエル国 ホド・ハシャロン ザルマン・シャザール・ストリート 5

(72)発明者 ウォルマン、アダム
イスラエル国 ホド・ハシャロン ハネシーム・ストリート 6 4 / 1 1

(72)発明者 ズロトニク、アビアド
イスラエル国 D . N . G . T 1 5 2 9 5 ミゼフ・ネトファ

審査官 北村 学

(56)参考文献 特開平07-244597 (JP, A)
特開平08-305500 (JP, A)
特開平11-085408 (JP, A)
米国特許第06148383 (US, A)
米国特許出願公開第2003/0028736 (US, A1)
国際公開第2003/017022 (WO, A1)

(58)調査した分野(Int.Cl., DB名)

G06F 3/06
G06F 11/20
G06F 12/00
JSTPlus(JDreamII)