

(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2016-224962

(P2016-224962A)

(43) 公開日 平成28年12月28日 (2016. 12. 28)

(51) Int. Cl.		F I		テーマコード (参考)
G06F 12/00	(2006.01)	G06F 12/00	533A	5B034
G06F 11/20	(2006.01)	G06F 11/20	682	

審査請求 有 請求項の数 20 O L (全 23 頁)

(21) 出願番号 特願2016-139400 (P2016-139400)
 (22) 出願日 平成28年7月14日 (2016. 7. 14)
 (62) 分割の表示 特願2014-518599 (P2014-518599) の分割
 原出願日 平成24年6月8日 (2012. 6. 8)
 (31) 優先権主張番号 13/174, 271
 (32) 優先日 平成23年6月30日 (2011. 6. 30)
 (33) 優先権主張国 米国 (US)

(71) 出願人 314015767
 マイクロソフト テクノロジー ライセンシング, エルエルシー
 アメリカ合衆国 ワシントン州 98052 レッドモンド ワン マイクロソフト ウェイ
 (74) 代理人 100140109
 弁理士 小野 新次郎
 (74) 代理人 100075270
 弁理士 小林 泰
 (74) 代理人 100101373
 弁理士 竹内 茂雄
 (74) 代理人 100118902
 弁理士 山本 修

最終頁に続く

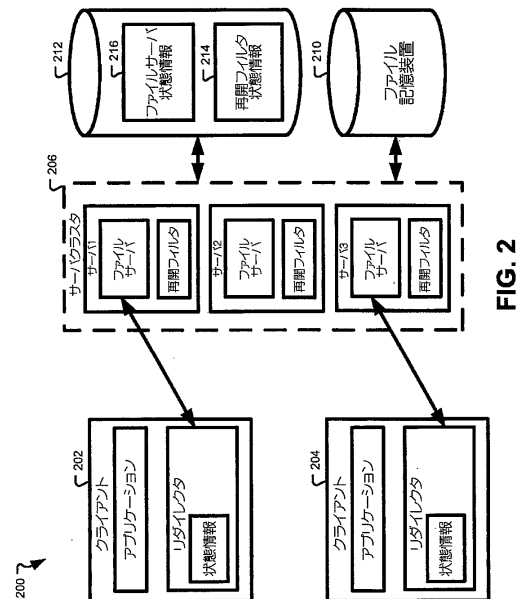
(54) 【発明の名称】 トランスペアレントなフェールオーバー

(57) 【要約】 (修正有)

【課題】 ネットワーク障害およびサーバフェールオーバーをまたいで状態を維持する技術を提供する。

【解決手段】 ソフトウェア環境200において、永続的ハンドルは、ファイルサーバとのセッションが確立された後でクライアント202, 204によって要求される。永続的ハンドルを求める要求は、クライアントによって生成されたハンドル識別子を含む。サーバは、ハンドル識別子を使用して状態情報と関連付ける。ネットワーク障害またはサーバフェールオーバーがありクライアントへの再接続があったとき、ハンドル識別子を使用して、もし再実行されればサーバ上で一貫性のない状態を生み出すことになるような再実行された要求が識別される。次いで、再実行された要求は適切に処理される。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

サーバクラスタ上の共有ファイルシステムにアクセスするクライアントに一貫した可用性を提供するコンピュータ実施方法であって、

共有ファイルシステム中のファイル情報にアクセスするために、ファイルサーバに接続する要求を、前記ファイルサーバにおいて受け取るステップであって、接続する前記要求はファイルアクセスプロトコルに従ってフォーマットされており、前記ファイルサーバはサーバクラスタ中の複数のサーバのうちの 1 つである、ステップと、

前記ファイルサーバから第 1 の応答を送るステップであって、前記第 1 の応答は、前記共有ファイルシステム中のファイル情報へのアクセスを可能にするためにクライアントとのセッションを確立し、前記第 1 の応答は前記ファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記クライアントによって前記共有ファイルシステム中のファイルにアクセスするために前記ファイルサーバ上で永続的ハンドルを開く要求を、前記ファイルサーバにおいて受け取るステップであって、永続的ハンドルを開く前記要求が、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

永続的ハンドルを開く前記要求を受け取るのに応答して、前記ファイルサーバが、

前記セッションに関する第 1 の状態情報を前記ハンドル識別子と関連付けて記憶し、

前記ファイルへのアクセスを許可する第 2 の応答を前記クライアントに送る

ステップと、

前記ファイルサーバにおいてファイルアクセス要求を受け取るステップと、

クライアント切断後に、前記永続的ハンドルを使用して前記セッションを再確立する要求を受け取るステップと、

前記第 1 の状態情報を使用して前記セッションを再確立するステップと、

前記セッションの再確立の後で、前記ハンドル識別子を含む新しいファイルアクセス要求を前記クライアントから受け取るステップと、

前記ハンドル識別子を使用して、前記新しいファイルアクセス要求が以前に前記ファイルサーバに送られた前記ファイルアクセス要求の再実行であると判定するステップとを含む方法。

【請求項 2】

請求項 1 に記載の方法であって、前記第 1 の状態情報が、前記クライアントによって再送されれば前記ファイルサーバが一貫性のない状態になるような操作の状態を含む、方法。

【請求項 3】

請求項 2 に記載の方法であって、前記ファイルアクセスプロトコルは、サーバメッセージブロック (SMB) プロトコルのあるバージョンである、方法。

【請求項 4】

請求項 3 に記載の方法であって、前記第 1 の状態情報は、前記クライアントによってなされるバイト範囲ロックと、前記クライアントによって送られる作成の操作とのうちの 1 以上の状態を含む、方法。

【請求項 5】

請求項 3 に記載の方法であって、前記クライアントの切断の前に、前記セッションに関する第 2 の状態情報を記憶するために再開フィルタを登録するステップを更に含む方法。

【請求項 6】

請求項 5 に記載の方法であって、前記第 2 の状態情報は、前記クライアントによってなされるリースと、前記クライアントによってなされるオープンとのうちの 1 以上の状態を含む、方法。

【請求項 7】

請求項 5 に記載の方法であって、前記セッションを再確立する前記ステップは、前記第 2 の状態情報を使用するステップを更に含む、方法。

10

20

30

40

50

【請求項 8】

請求項 1 に記載の方法であって、前記クライアントの切断は前記ファイルサーバの障害のせいで発生し、前記セッションの前記再確立は前記サーバクラスタ中の第 2 のファイルサーバによって実施される、方法。

【請求項 9】

コンピュータ実行可能命令を含むコンピュータ可読記憶媒体であって、前記コンピュータ実行可能命令は、コンピュータに、一貫した状態を維持する方法を実施させ、前記方法は、

ファイル情報にアクセスするためにサーバに接続する要求をクライアントによって送るステップであって、接続する前記要求がファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記サーバから第 1 の応答を受け取るステップであって、前記第 1 の応答が、前記サーバ上のファイル情報へのアクセスを可能にするために前記クライアントとのセッションを確立するステップであって、前記第 1 の応答は前記ファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記クライアントによって前記サーバ上のファイルにアクセスするために前記サーバ上で永続的ハンドルを開く要求を送るステップであって、永続的ハンドルを開く前記要求が、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

前記ファイルへのアクセスを許可する第 2 の応答を前記クライアントにおいて受け取るステップと、

ファイルアクセス要求を送るステップと、

前記クライアントが前記サーバから切断されたことを検出するステップと、

前記永続的ハンドルを使用して前記セッションを再確立する要求を送るステップであって、前記セッションを再確立する前記要求は前記ハンドル識別子を含む、ステップと、

前記セッションが再確立されたと判定するステップと、

以前に送った前記ファイルアクセス要求の再実行である新しいファイルアクセス要求を送るステップと

を含む、コンピュータ可読記憶媒体。

【請求項 10】

請求項 9 に記載のコンピュータ可読記憶媒体であって、前記サーバはサーバクラスタ中の複数のサーバのうちの 1 つであり、前記切断は前記サーバの障害のせいで発生し、前記セッションの前記再確立は前記サーバクラスタ中の前記複数のサーバのうちの第 2 のサーバによって実施される、コンピュータ可読記憶媒体。

【請求項 11】

請求項 10 に記載のコンピュータ可読記憶媒体であって、前記ファイルアクセスプロトコルは、サーバメッセージブロック (SMB) プロトコルのあるバージョンである、コンピュータ可読記憶媒体。

【請求項 12】

請求項 9 に記載のコンピュータ可読記憶媒体であって、前記新しい要求は、前記ハンドル識別子を含み、前記切断の前の前記クライアントによってなされるバイト範囲ロックと、前記切断の前の前記クライアントによって送られる作成の要求とのうちの 1 以上の要求の再実行である、コンピュータ可読記憶媒体。

【請求項 13】

サーバクラスタ上の共有ファイルシステムにアクセスするクライアントに一貫した可用性を提供するためのシステムであって、該システムは、

少なくとも 1 つのプロセッサと、

前記少なくとも 1 つのプロセッサに動作上結合され、前記少なくとも 1 つのプロセッサにより実行されると前記少なくとも 1 つのプロセッサにある方法を実施させる命令を含むメモリと

を備え、前記方法は、

10

20

30

40

50

共有ファイルシステム中のファイル情報にアクセスするために、ファイルサーバに接続する要求を、前記ファイルサーバにおいて受け取るステップであって、接続する前記要求はファイルアクセスプロトコルに従ってフォーマットされており、前記ファイルサーバはサーバクラスタ中の複数のサーバのうちの1つである、ステップと、

前記ファイルサーバから第1の応答を送るステップであって、前記第1の応答は、前記共有ファイルシステム中のファイル情報へのアクセスを可能にするためにクライアントとのセッションを確立し、前記第1の応答は前記ファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記クライアントによって前記共有ファイルシステム中のファイルにアクセスするために前記ファイルサーバ上で永続的ハンドルを開く要求を、前記ファイルサーバにおいて受け取るステップであって、永続的ハンドルを開く前記要求が、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

永続的ハンドルを開く前記要求を受け取るのに応答して、前記ファイルサーバが、前記セッションに関する第1の状態情報を前記ハンドル識別子と関連付けて記憶し、前記ファイルへのアクセスを許可する第2の応答を前記クライアントに送るステップと、

前記ファイルサーバにおいてファイルアクセス要求を受け取るステップと、

クライアント切断後に、前記永続的ハンドルを使用して前記セッションを再確立する要求を受け取るステップと、

前記第1の状態情報を使用して前記セッションを再確立するステップと、

前記セッションの再確立の後で、前記ハンドル識別子を含む新しいファイルアクセス要求を前記クライアントから受け取るステップと、

前記ハンドル識別子を使用して、前記新しいファイルアクセス要求が以前に前記ファイルサーバに送られた前記ファイルアクセス要求の再実行であると判定するステップとを含む、システム。

【請求項14】

請求項13に記載のシステムであって、前記第1の状態情報が、前記クライアントによって再送されれば前記ファイルサーバが一貫性のない状態になるような操作の状態を含む、システム。

【請求項15】

請求項14に記載のシステムであって、前記ファイルアクセスプロトコルは、サーバメッセージブロック(SMB)プロトコルのあるバージョンである、システム。

【請求項16】

請求項15に記載のシステムであって、前記第1の状態情報は、前記クライアントによってなされるバイト範囲ロックと、前記クライアントによって送られる作成の操作とのうちの1以上の状態を含む、システム。

【請求項17】

請求項15に記載のシステムであって、前記クライアントの切断の前に、前記セッションに関する第2の状態情報を記憶するために再開フィルタを登録するステップを更に含むシステム。

【請求項18】

請求項17に記載のシステムであって、前記第2の状態情報は、前記クライアントによってなされるリースと、前記クライアントによってなされるオープンとのうちの1以上の状態を含む、システム。

【請求項19】

請求項17に記載のシステムであって、前記セッションを再確立する前記ステップは、前記第2の状態情報を使用するステップを更に含む、システム。

【請求項20】

請求項13に記載のシステムであって、前記クライアントの切断は前記ファイルサーバの障害のせいで発生し、前記セッションの前記再確立は前記サーバクラスタ中の第2のフ

10

20

30

40

50

ファイルサーバによって実施される、システム。

【請求項 2 1】

請求項 9 から 1 2 のうちの何れか一項に記載のコンピュータ実行可能命令を含むプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

[0001]サーバクラスタは、フェールオーバーおよび高い情報可用性をクライアントに提供するために一般に使用される。

10

【背景技術】

【0002】

サーバクラスタの使用により、クライアント上でサーバ動作を要求するアプリケーションに対してどんなサーバ障害もトランスペアレントであるように、トランスペアレントなフェールオーバーがクライアントにもたらされる。サーバクラスタは、ネットワーク中の複数のクライアントにファイル情報へのアクセスを提供するための共有ファイルシステムにおいて、有用である可能性がある。

【発明の概要】

【発明が解決しようとする課題】

【0003】

20

しかし、共有ファイルシステムがサーバメッセージブロック (SMB) プロトコルなどのステートフルなプロトコルを利用するとき、問題が生じることがある。サーバクラスタ中のサーバに障害が発生したとき、いくつかのステートフルなプロトコルは、クライアント状態を障害サーバから代替サーバに移行させる方法を提供しない。また、いくらかの状態情報を記憶することを可能にするファイルアクセスプロトコルであっても、異なるコンポーネントが異なる状態情報を記憶することは可能にしない。

【0004】

[0002]実施形態は、これらおよび他の考慮事項に鑑みてなされたものである。また、比較的特定の問題を論じたが、この背景において識別された特定の問題の解決に実施形態が限定されるべきでないことを理解されたい。

30

【課題を解決するための手段】

【0005】

[0003]この概要は、詳細な説明のセクションでさらに後述する概念の精選を単純な形で紹介するために提供する。この概要は、特許請求される主題の重要な特徴または本質的な特徴を識別するものとはせず、また、特許請求される主題の範囲を決定する際の補助として使用されるものともしない。

【0006】

[0004]共有ファイルシステム中で永続的ハンドルを利用する実施形態について述べる。永続的ハンドルを使用して、ネットワーク障害およびサーバフェールオーバーをまたいで状態が維持される。永続的ハンドルは、ファイルサーバとのセッションが確立された後でクライアントによって要求される。永続的ハンドルを求める要求は、クライアントによって生成されたハンドル識別子を含む。サーバは、ハンドル識別子を使用して状態情報と関連付ける。ネットワーク障害またはサーバフェールオーバーがありクライアントへの再接続があったとき、ハンドル識別子を使用して、もし再実行されればサーバ上で一貫性のない状態を生み出すこととなるような再実行された要求が識別される。次いで、再実行された要求は適切に処理される。

40

【0007】

[0005]実施形態は、コンピュータプロセス、コンピューティングシステム、または、コンピュータプログラム製品やコンピュータ可読媒体などの製造品として、実現することができる。コンピュータプログラム製品は、コンピュータシステムによって読取可能な、か

50

つコンピュータプロセスを実行するための命令のコンピュータプログラムをエンコードする、コンピュータ記憶媒体とすることができる。コンピュータプログラム製品はまた、コンピューティングシステムによって読取可能な、かつコンピュータプロセスを実行するための命令のコンピュータプログラムをエンコードする、搬送波上の伝搬信号とすることもできる。

【0008】

[0006]以下の図に関して、非制限的かつ非網羅的な実施形態について述べる。

【図面の簡単な説明】

【0009】

【図1】[0007]実施形態を実現するのに使用できるシステムを示す図である。

10

【図2】[0008]いくつかの実施形態による、ファイルアクセスプロトコルを使用して通信するクライアントおよびファイルサーバクラスタのブロック図である。

【図3】[0009]いくつかの実施形態による、サーバフェールオーバー時に再実行防御を提供するための動作フローを示す図である。

【図4】[0010]いくつかの実施形態による、ファイルの一貫した可用性を維持するための動作フローを示す図である。

【図5】[0011]実施形態を実現するのに適したコンピューティング環境のブロック図である。

【発明を実施するための形態】

【0010】

20

[0012]以下、本明細書の一部をなし特定の例示的な実施形態を示す添付の図面を参照しながら、様々な実施形態についてより完全に述べる。しかし、実施形態は多くの異なる形で実現することができ、本明細書に示す実施形態に限定されるものと解釈すべきではない。そうではなく、これらの実施形態は、本開示が十分に完全なものとなるように、かつ実施形態の範囲を当業者に完全に伝えることとなるように提供する。実施形態は、方法、システム、またはデバイスとして実践することができる。したがって、実施形態は、ハードウェア実装形態の形、完全にソフトウェア実装形態の形、またはソフトウェア面とハードウェア面を組み合わせた実装形態の形をとることができる。よって、以下の詳細な説明は、限定的な意味にとるべきではない。

【0011】

30

[0013]図1に、いくつかの実施形態を実現するのに使用できるシステム100を示す。システム100は、クライアント102および104、ならびにサーバクラスタ106を含む。クライアント102および104は、ネットワーク108を介してサーバクラスタ106と通信する。サーバクラスタ106は、クライアント102および104上のアプリケーションからアクセスされる情報を記憶する。クライアント102および104は、クラスタ106とのセッションを確立して、クラスタ106上の情報にアクセスする。図1ではクライアント102および104のみがクラスタ106と通信するのが示されているが、他の実施形態では、サーバクラスタ106からの情報にアクセスするクライアントが3つ以上あってもよい。

【0012】

40

[0014]図1に示すように、サーバクラスタ106は、サーバ106A、106B、および106Cを含み、これらのサーバは、クラスタ106に記憶された情報について、高い可用性と冗長性の両方を提供する。実施形態では、クラスタ106は、クライアント102および104からアクセスされるファイルシステムを有する。図1には3つのサーバが示されているが、他の実施形態では、クラスタ106は4つ以上のサーバまたは2つ以下のサーバを含んでもよい。実施形態では、クライアント102および104上のアプリケーションがファイルシステムにファイル情報を要求し、アプリケーションに対してトランスペアレントに、サーバクラスタ106上の共有ファイルシステムからファイル情報が取り出される。

【0013】

50

[0015]一実施形態によれば、サーバ106A、106B、および106Cを利用して、クラスタ106に記憶されたファイルシステムの一貫した可用性が提供される。これは、クライアント102および104上ならびにサーバ106A、106B、および106C上のコンポーネントを利用して状態情報を記憶することによって行われ、この状態情報を使用して、ネットワーク108の障害またはサーバ106A、106B、106Cのうちの1つの障害があった場合にクライアント102および104とクラスタ106との間でセッションを再確立することができる。後でより詳細に述べるように、状態情報を記憶することで、クライアント102および104は、クライアント102および104上で実行されるアプリケーションに対してトランスペアレントな、一貫したファイルアクセスおよびフェールオーバーを有することができる。

10

【0014】

[0016]クラスタ106のサーバ、例えば106A、106B、および106Cは、実施形態では、それぞれファイル情報へのアクセスをクライアントに提供するものであり、ファイル情報の一貫した可用性をクライアントに提供するように構成される。一実施形態を例示すると、クライアント102は、クラスタ106のサーバとのセッションを確立する要求を送ることができる。例えば、クライアント102は、サーバ106Aとのセッションを確立して、サーバクラスタ106に記憶された共有ファイルシステムにアクセスすることができる。セッションを確立するプロセスの一部として、クライアント102は、ファイルアクセスプロトコルを利用することができる。実施形態では、ファイルアクセスプロトコルは、ネットワークファイルシステム(NFS)またはサーバメッセージブロック(SMB)プロトコルのバージョンである。

20

【0015】

[0017]セッションの確立は、クライアント102とサーバ106Aとの間で伝送されるいくつかの交渉要求および応答の交換を含んでよい。SMBプロトコルのバージョンでは、セッション中に使用されることになるプロトコルの正確なバージョンを交渉するのに使用される、かつクライアント(例えば102)とサーバ(例えば106A)の両方の能力を相互に対して公開するのに使用される、特に定義された交渉パケットがある。一実施形態では、交渉パケットは、サーバ106Aがクラスタ(例えばクラスタ106)の一部であることを示す指示を含んでよい。これによりクライアントは、一貫した可用性、言い換えればトランスペアレントなフェールオーバー能力を、サーバ106Aが提供できることを

30

【0016】

[0018]上の例を続けるが、セッションが確立された後、クライアント102は、ファイルシステム中のファイルにアクセスするための永続的ハンドルを求めて、ファイルアクセスプロトコルに従ってフォーマットされたメッセージをサーバ106Aに送ることができる。永続的ハンドルを要求することは、実施形態では、サーバ106Aがクラスタ106の一部であることの結果として利用可能であるトランスペアレントなフェールオーバー能力を、クライアントが利用したいことを示す。実施形態では、要求は、グローバル意識別子であるハンドル識別子を含む。

【0017】

[0019]サーバ106Aは、永続的ハンドルを求める要求を受け取り、クライアント102とのセッションに関する状態情報と共にハンドル識別子を記憶することになる。状態情報を記憶することは単に、ファイルサーバが記憶域へのハンドル識別子を保持すること、および状態情報をハンドル識別子と関連付けて記憶することを含めばよい。後でより詳細に述べるように、いくつかの実施形態では、フィルタなどの別々のコンポーネントを使用して、種々のタイプの状態情報を記憶することができる。さらに他の実施形態では、永続的ハンドルに関する情報は、ノード間で複製され、ファイルシステム上の永続記憶域には記憶されない。さらに他の実施形態では、永続的ハンドルに関する情報は、ノード間で複製され、かつ、ファイルシステム上の永続記憶域に記憶される。

40

【0018】

50

[0020]サーバ106Aは、応答をクライアント102に送り、永続的ハンドルとファイル情報へのアクセスとを許可する。次いでクライアント102は、ファイルに対して様々な操作を実施するための他の要求を送ることに進むことができる。例えば、クライアント102は、ファイル情報を読み取ること、ファイルに書き込むこと、ファイルの属性を列挙すること、ファイルを閉じること、およびファイルに対する様々なロックを求めることの要求を送ることができる。クライアントによって要求された各操作の結果、状態情報が更新されてよく、それにより、クライアントが切断された場合にクライアントの状態を確実に元通りにできるようになる。この更新は、追加の状態情報をハンドル識別子と関連付けて保存することを含んでよい。

【0019】

[0021]何らかの時点で、クライアント102はサーバから切断される場合がある。切断は、例えばネットワーク障害または混乱のせいであることがある。あるいは、切断は、サーバ106Aの障害のせいであることがある。ネットワーク障害を伴う実施形態では、クライアント102は、切断が生じたことを検出し、ネットワークが利用可能になるのを待機して、サーバ106Aに再接続することができる。他の実施形態では、クライアント102は、障害を検出すると、再接続する要求をクラスタ106に送り、クラスタ106は、フェールオーバーサーバを提供して再接続要求を処理することになる。

【0020】

[0022]いずれの場合でも、クライアント102は、再接続する要求を送る。要求は、ハンドル識別子を含むことになる。サーバ106Aまたは代替サーバ(106Bもしくは106C)は、ハンドル識別子に基づいて状態情報を取り出し、状態情報を使用して前の状態を再確立し、再接続が成功であることを示す応答をクライアントに送ることになる。いくつかの実施形態では、前の状態情報が失われたかさもなければ利用不可能である場合は、再接続が可能でないことがある。これらの状況では、サーバは、再接続要求をセッション確立要求として扱い、相応に応答することができる。

【0021】

[0023]セッションが再確立された後、クライアント102は、新しいファイルアクセス要求を送る。いくつかの実施形態では、新しいファイルアクセス要求のうちの1つは、前の要求の再実行である場合がある。再実行された要求は、もしサーバによってそれが再実行であると認識されずに処理されればサーバ上で一貫性のない状態を生み出すことになるようなタイプの要求である場合がある。正確な要求タイプは、使用されているファイルアクセスプロトコルによってどのように要求が処理されるかに依存する。例えば、SMBプロトコルのバージョンでは、バイト範囲ロックがファイルの各部分に対して要求され許可される場合がある。したがって、ファイルの各部分をロックする要求をクライアントが送り、要求が完了したが、切断前にクライアントに通知されない場合、クライアントは、前の要求を再実行する可能性がある。サーバは、要求が再実行であることを識別できることが必要となる。したがって、実施形態では、永続的ハンドルを求める元の要求と共に送られたハンドル識別子を使用して、再実行された要求が識別される。識別されると、再実行された要求を、サーバ上の一貫性のない状態を回避するために処理することができる。

【0022】

[0024]いくつかの実施形態では、トランスペアレントなフェールオーバーをクライアント102上のアプリケーションに提供するために、クライアント102に記憶される状態情報があってよい。すなわち、サーバ106A(またはフェールオーバーサーバ)は、再接続後に状態を復元するのに必要な情報の全てを記憶することは担わなくてよい。いくつかの実施形態では、クライアントが、いくらかの状態を再確立することを担うことができる。例えば、ファイル情報を読み取る要求が切断前に送られた場合、サーバは、読取要求に関する状態情報を保存することを担わなくてよい。再接続が行われたとき、クライアントは、読取要求を再送することを担うことができる。異なるコンポーネントによって状態情報が復元される実施形態に関する追加の説明については、図2に関してより詳細に後述する。

10

20

30

40

50

【 0 0 2 3 】

[0025]以上の記述は、図1に示す実施形態がどのように動作できるかに関する一例に過ぎない。より詳細に後述するように、実施形態は、種々のステップまたは動作を含み得る。これらは、任意の適切なソフトウェアまたはハードウェアコンポーネントまたはモジュールを使用して実施することができる。

【 0 0 2 4 】

[0026]次に図2に移るが、図2には、クライアント202と、クライアント204と、3つのサーバ(サーバ1、サーバ2、およびサーバ3)を有するサーバクラスタ206とを含むソフトウェア環境200のブロック図を示す。また、ファイルシステムがファイル情報を記憶するファイル記憶装置210、および、サーバ1、サーバ2、サーバ3のうちの1つまたは複数によって状態情報を記憶できる記憶装置212も示す。

10

【 0 0 2 5 】

[0027]図2に示すように、クライアント202およびクライアント204はそれぞれ、ファイル情報を要求することのできるアプリケーションを備える。アプリケーションは、例えば、ワードプロセッシングアプリケーション、表計算アプリケーション、ブラウザアプリケーション、または、ファイルへのアクセスを要求する他の任意のアプリケーションとすることができる。図2に示す実施形態では、ファイルは、ファイル記憶装置210に記憶された共有ファイルシステム中に位置する。クライアント202およびクライアント204はそれぞれ、ファイルを求める要求をアプリケーションからファイルサーバにリダイレクトするリダイレクタをさらに備え、ファイルサーバは、共有ファイルシステムへのアクセスを提供する。リダイレクタは、ファイルアクセスプロトコルを使用してファイルサーバと通信する。いくつかの実施形態では、ファイルアクセスプロトコルは、NFSのまたはSMBプロトコルのバージョンとすることができる。例示の目的で、図2については、クライアント202およびクライアント204中のリダイレクタがSMB2.0などのSMBプロトコルのバージョンを使用してファイルサーバと通信すると仮定して述べる。しかし、実施形態は、SMBプロトコルの使用に限定されない。

20

【 0 0 2 6 】

[0028]図2では、サーバ1、サーバ2、およびサーバ3がそれぞれファイルサーバを備えるのが示されている。上に言及したように、ファイルサーバは、SMBプロトコルのバージョンを使用して、クライアント202およびクライアント204上のリダイレクタと通信することができる。サーバ1、サーバ2、およびサーバ3のそれぞれは再開フィルタも備え、再開フィルタは、いくつかの実施形態では、クライアントリダイレクタとファイルサーバとの間で確立されたセッションの状態情報を記憶するのに使用される。

30

【 0 0 2 7 】

[0029]SMBプロトコルを使用してクライアントとサーバとの間におけるセッションを確立することは、リダイレクタ(クライアント202上のリダイレクタなど)がファイルサーバ(サーバクラスタ206中のサーバ1など)に交渉要求を送ることで開始する。リダイレクタとファイルサーバは、交渉パケットを交換して、セッションに使用されることになるSMBのバージョンを交渉する。加えて、交渉中に能力を交換することもできる。一実施形態では、ファイルサーバは、ファイルサーバからクライアントに送られる交渉応答パケットの中に能力フラグを含めて、ファイルサーバが永続的ハンドルの使用をサポートすることをクライアントに示すことができる。いくつかの実施形態では、これは、ファイルサーバが、クラスタ中の別のサーバへのフェールオーバーによって一貫した可用性をクライアントに提供できるクラスタの一部であるような状況で行われる。他の実施形態では、スタンドアロンサーバもまた、ネットワーク障害がある場合にクライアントに再接続できるように、この能力を有することができる。

40

【 0 0 2 8 】

[0030]交渉が完了すると、クライアント上のリダイレクタとファイルサーバとは、セッションを確立する。次いで、クライアントリダイレクタは、ファイルアクセス要求をファイルサーバに送ることができる。一実施形態では、リダイレクタは、永続的ハンドルを要

50

求する。SMBプロトコルのバージョンは、切断されたクライアントに再接続するのに使用できる持続的ハンドルを可能にする。しかし、持続的ハンドルは、クライアントが再接続した後で状態を記憶および再確立することを必ずしも可能にするとは限らない。したがって、実施形態では、リダイレクタは、クライアントリダイレクタが永続的ハンドルを要求していることを示すための何らかの追加フラグおよび/またはインジケータと共に、持続的ハンドルの要求を送ることができる。加えて、クライアントは、再実行された要求を再接続後に識別するのに使用できるハンドル識別子を含んでよい。以下は、永続的ハンドルを要求するためにSMBプロトコルのバージョンにおいて使用できる持続的ハンドル要求構造の一実施形態である。

```
struct SMB2__DURABLE__HANDLE__REQUEST__V2 {
    ULONG Flags;
    GUID HandleId; //クライアントがこのハンドルに対する固有IDを提供した
    // (再実行を検出するのに使用される)
    ULONG Timeout; //タイムアウト(秒)
    ULONG Reserved; //ZEROに設定されなければならない}

```

【0031】要求に回答して、サーバ1上のファイルサーバは、実施形態では、永続的ハンドルを許可しファイル識別子をクライアント202上のクライアントリダイレクタに提供することによって応じる。次いで、クライアントリダイレクタは、永続的ハンドルとファイル識別子とに関連するファイルからの情報にアクセスすることができる。いくつかの実施形態では、クライアントリダイレクタは、ディレクトリに対する永続的ハンドルを要求することができる。すなわち、永続的ハンドルは個別ファイルに関連するのではなく、ハンドルはディレクトリに関連してよい。

【0029】

【0032】サーバ1上のファイルサーバが永続的ハンドルを許可するのに加えて、ファイルサーバはまた、状態情報を記憶装置212に記憶することになる。状態情報は、クライアントリダイレクタによって生成されたハンドル識別子と関連付けて記憶することができ、また、クライアント202上のクライアントリダイレクタに提供されたファイル識別子と関連付けて記憶することができる。後でより詳細に述べるように、ファイルサーバは、状態情報をファイルサーバ状態情報216として直接に記憶することができる。他の実施形態では、ファイルサーバは、再開フィルタを利用して状態情報を記憶することができる。さらに他の実施形態では、ファイルサーバは、状態情報を直接に記憶し、かつ、他の状態情報を記憶するために再開フィルタを使用することができる。

【0030】

【0033】交渉が完了した後、クライアントリダイレクタは、例えばSMBプロトコルのバージョンを使用して、ファイルアクセス要求を送る。いくつかの実施形態では、ファイルサーバは、クライアントリダイレクタから受け取った各要求につき、状態情報を記憶することになる。何らかの時点で、例えばネットワーク障害またはサーバ1の障害の結果として、クライアント202とサーバ1との間で切断がある場合がある。クライアント202は、障害がネットワーク障害に基づくものであった場合にはサーバ1と、またはフェールオーバーサーバ(サーバ2とサーバ3のうちの一方)と、接続を再確立することができる。再接続の一部として、クライアント202は、前に提供したハンドル識別子と、元のセッションの交渉時にファイルサーバによって提供されたファイル識別子とを含む再接続要求を送ることができる。状態情報は、サーバクラスタ206中の全てのサーバからアクセス可能な記憶装置212中で利用可能なので、フェールオーバーサーバは、再接続要求の中でクライアントによって提供されたハンドル識別子および/またはファイル識別子に基づいて、前の状態情報を識別することができる。クライアントがサーバ1との接続を再確立しようと試みている実施形態では、サーバ1上のファイルサーバもまた、記憶装置212上の状態情報にアクセスして、クライアントとのセッションの前の状態を再確立することができる。

10

20

30

40

50

【 0 0 3 1 】

[0034]上に言及したように、いくつかの実施形態では、切断されたクライアントに状態の再確立を提供するために、環境 2 0 0 中の異なるコンポーネントが、異なるタイプの状態情報を記憶することを担う。図 2 に示すように、各ファイルサーバは、再開フィルタを備える。実施形態では、再開フィルタを使用して、クライアントが再接続されたときに状態を再確立するための状態情報が記憶される。再開フィルタは、ファイルサーバによって使用される特定のファイルアクセスプロトコルに依存しない。実施形態では、ファイルサーバは最初に、特定の状態情報を記憶するために再開フィルタに登録することになる。登録されると、ファイルサーバは、状態情報を再開フィルタに渡すことができ、再開フィルタは、状態情報を再開フィルタ状態情報 2 1 4 として記憶装置 2 1 2 に記憶する。再開フィルタ状態情報 2 1 4 に加えて、サーバは、ファイルサーバ状態情報 2 1 6 として示される別個の状態情報を、記憶装置に 2 1 2 に記憶することができる。実施形態では、この異なる状態情報は、再開フィルタ状態情報 2 1 4 とは別の記憶位置に記憶することができる。ファイルサーバ状態情報 2 1 6 および再開フィルタ状態情報 2 1 4 は、ログファイルなど任意の適切な方式で記憶することができる。後でより詳細に述べるように、再開フィルタによって記憶される状態情報のタイプは、実施形態では、一般情報であり、サーバ情報は、より特定の状態情報である。

10

【 0 0 3 2 】

[0035]いくつかの実施形態では、クライアントもまた、いくらかの状態情報を記憶することを担う。図 2 に示すように、クライアント 2 0 2 および 2 0 4 は、切断後にクライアントが再接続されたときに状態を再確立するのに使用される状態情報を記憶する。これらの実施形態では、切断後にクライアントが再接続されたときにクライアントの状態を再確立するためにファイルサーバが全ての状態情報を記憶することを必要とする代わりにクライアントが状態を再確立するようにすることには、いくらかのコスト節約があるであろう。例えば、ファイルサーバが全ての状態情報を記憶する必要がある場合は、ファイルサーバは、ファイルに対して実施すべき何らかの操作を伴う何らかの要求をクライアントリダイレクタから受け取る度に、この要求または操作に関する何らかの情報を記憶することが必要とされる。クライアントリダイレクタが状態情報のうちのいくらかを記憶する必要があるようにすることにより、クライアントから受け取った要求または操作ごとにファイルサーバが状態情報を記憶しなければならないことのコストが削減される。

20

30

【 0 0 3 3 】

[0036]理解できるように、環境 2 0 0 中の異なるコンポーネントに記憶される状態情報は、異なる設計上の考慮事項に依存する。例えば、状態情報が整合性を有し一貫して利用可能であることをファイルサーバが保証する必要がある、十分に重要な何らかの情報がある場合がある。この場合、この情報は、ファイルサーバおよび/または再開フィルタによって記憶されるべきである。例えば、サーバが共有モードを施行し、アクセスを要求する新しいクライアントが既存のクライアントのアクセスと干渉しないように保証するために、実施形態によれば、状態情報はサーバに記憶されなければならない。他の状態情報は、それほどクリティカルでない場合があり、情報中でいくらかの不整合は許容され得る。例として、クライアントが、ローカルにキャッシュされたファイルプロパティを有する場合がある。クライアントが切断に続いてファイルサーバに再接続した後、キャッシュされたファイルプロパティを新たに要求することができる。

40

【 0 0 3 4 】

[0037]一実施形態では、クライアントリダイレクタとファイルサーバとの間の通信に S M B プロトコルのバージョンが使用される場合、S M B プロトコルは、環境 2 0 0 中に示す様々なコンポーネントによって記憶されることになる特定の状態を可能にすることができる。一実施形態では、S M B プロトコルを使用して利用可能な操作は、3つのグループに分けられる。各グループに関連する状態情報は、異なるコンポーネントによって記憶される。

【 0 0 3 5 】

50

[0038]第1のグループは、非べき等 (non-idempotent) 操作と一般に呼ぶことができ、これは、これらの操作が再実行された場合、例えばクライアント切断の前にすでに一度ファイルに適用された後で再適用された場合に、ファイルサーバ上で一貫性のない状態を生み出すであろうことを意味する。SMBプロトコルのバージョンでは、バイト範囲ロックが、再実行検出を必要とする操作の一例である。というのは、これらのロックはスタックされてアンスタックされるからである。他の例は、付加としての書込みおよびオープン/作成を含み、これらは、例えば新しいファイルの作成または既存のファイルの上書きによって、ディスク状態を修正する可能性がある。実施形態では、ファイルサーバはこれらの操作が再実行されていることを認識しなければならないので、これらのタイプの操作に関連する状態はファイルサーバによって記憶される。図2に示す実施形態では、これらの操作に関連する状態は、サーバ1、サーバ2、およびサーバ3のそれぞれの上のファイルサーバによって、ファイルサーバ状態情報216の一部として記憶装置212に記憶されることになる。いくつかの実施形態では、前述のような、セッションの交渉中にクライアントによって提供されたハンドル識別子を使用して、要求が前の要求の再実行であることが識別される。

10

20

30

40

50

【0036】

[0039]操作の第2のグループは、データオープン操作に係る。これらの操作は、ファイル中の情報の読取り、書込み、実行、または削除を要求する場合がある。共有モードを施行し、他のクライアントが既存のクライアントに影響を及ぼさないようにすることができるために、実施形態によれば、これらのオープン操作に関する状態は、サーバ側で記憶されなければならない。オープン操作に関する状態はまた、ローカル操作が永続的ハンドルと干渉するのを阻止するために、サーバ側で記憶される。例えば、クラスタノード上で実行されるプログラムが、クライアントのために予約されているハンドルを修正するかさもなければそれに影響を及ぼすことが防止される。実施形態では、これらのタイプの操作に関する情報は、再開フィルタによって記憶される。上に言及したように、再開フィルタは、実施形態では、SMBプロトコルに特有ではなく、NFSなどの異なるファイルアクセスプロトコルをファイルサーバが使用しているときでも使用することができる。図2に示す実施形態では、サーバ1、サーバ2、およびサーバ3のそれぞれの上の再開フィルタは、オープン操作の状態情報を、再開フィルタ状態情報214の一部として記憶装置212に記憶する。

【0037】

[0040]操作の第3のグループは、サーバにおいて再適用されてもサーバの最終状態を変更しないような操作を含む。これらは、べき等 (idempotent) 操作と呼ぶことができる。このグループ中のいくつかの操作は、読取り、付加でない書込み、削除、名前変更、メタデータ設定操作、およびメタデータ照会操作を含むが、これらに限定されない。リース状態もまた、クライアントによって記憶することができ、サーバによって保持される必要はない。実施形態では、リースは、性能を増大させネットワーク使用を低減させるためにクライアントが一貫した方式でそのバッファリング戦略を動的に改変するのを可能にするように設計されたメカニズムである。リモートファイル操作に関するネットワーク性能は、クライアントがファイルデータをローカルにバッファリングできる場合に増大させることができ、ローカルにバッファリングすることで、ネットワークパケットを送受信する必要性が低減するかまたはなくなる。クライアントは、他のどんなクライアントもデータにアクセスしていないことを確認した場合、リモートサーバ上のファイルに情報を書き込まなくてよいものとするができる。同様に、クライアントは、他のどんなクライアントもリモートファイルにデータを書き込んでいないことを確認した場合、リモートファイルからの事前に読み取ったデータをバッファリングすることができる。

【0038】

[0041]実施形態によれば、リース状態は、サーバ上で保持される必要はない。というのは、クライアントがフェールオーバー後にこれらのハンドルを再開している間、再開フィルタは所与のファイルに対する全ての作成をブロックするからである。このことは暗黙的

に、クライアントが猶予期間中に再接続する / それらのハンドルを再開する場合は、フェールオーバープロセス中にハンドルリースが失われることは決してないという保証をもたらす。言い換えれば、クライアントは、再開段階中はそれらのハンドルリースを常に取り戻すことになる。さらに、読取り / 書込み、読取り / 書込み / ハンドルリースなどの、排他的リースは、どんな所与の時点でも、単一のクライアントのみに許可される。このことは、他のどんなクライアントからの、ファイルに対する他のどんなデータオープンもないことを含意する。したがって、フェールオーバー中は、再開フィルタは、排他的リースを保持するクライアントがその全てのハンドルを再開し終えるまで、ファイルに対する新しい作成を可能にしないことになるので、クライアントがその排他的リースを取り戻すことになるという保証がある。読取りリースなど、肯定応答を必要としない共有リースは、サーバと再開フィルタとのいずれにも知られることなく、いつでも失われる可能性がある。というのは、基礎をなすファイルシステムが、中断を引き起こした操作の続行を可能にするからである。このようなリースの場合、クライアントは、実施形態では、リースがフェールオーバー全体にわたって中断されていると見なし、そのキャッシュを消去して古い読取りを防止する。したがって、第3のグループ中の操作の状態は、サーバからのどんな追加サポートもなしにクライアントによって再現することができる。図2に示す実施形態では、クライアント202および204上のリダイレクタが、第3の操作グループの状態情報を記憶する。

10

【0039】

[0042] 動作において、環境200は、クライアント202および204上のアプリケーションが、共有ファイルシステム中のファイル記憶装置210に記憶されたファイルへのアクセスを要求するのを可能にする。アプリケーションは、ファイル情報をトランスペアレントに要求することができる。クライアント上のリダイレクタは、前述のようにクラスタ206中のサーバのうちの1つとのセッションを確立し、切断があった場合にリダイレクタが再接続してセッションを再確立できるように永続的ハンドルを要求することになる。ファイルサーバは、状態情報を、ファイルサーバ状態情報216として直接に、または再開フィルタを使用して再開フィルタ状態情報214として、記憶装置212に記憶することになる。いくつかの実施形態では、クライアントもまた、いくらかの状態情報を記憶することになる。切断が発生した場合、リダイレクタは、ファイルサーバまたはフェールオーバーサーバに再接続するのを要求することができる。次いで、サーバ側（例えば記憶装置212中）およびクライアント側で記憶された状態情報を使用して、クライアントの前の状態を再確立することができる。これは全て、クライアント202および204上のアプリケーションに対してトランスペアレントに行われる。

20

30

【0040】

[0043] 理解できるように、環境200に関する上の記述は、本明細書に述べる実施形態を限定する意図はない。図2およびその記述は、単に、いくつかの実施形態の実装形態を例示するものとする。他の実施形態では、環境200中の異なるコンポーネントに、異なるタイプの状態情報を記憶することができる。また、上に示したように、記憶される状態情報のタイプ、ならびにどんなコンポーネントが状態情報を記憶するかを決定し得る、種々のファイルアクセスプロトコルを使用することができる。したがって、実施形態は、図2に示し記述するものに限定されない。

40

【0041】

[0044] 以下の図3および4に関する記述は、サーバメッセージブロック(SMB)プロトコルをファイルアクセスプロトコルとして使用して行う。しかし、実施形態はこれに限定されない。実施形態では、SMBまたはネットワークファイルシステム(NFS)の種々のバージョンを含めた、任意のファイルアクセスプロトコルを、ファイルアクセスプロトコルとして使用することができる。SMBは、単に例示を簡便かつ容易にするために記述の中で使用されているに過ぎない。

【0042】

[0045] 図3および4に、実施形態による動作フロー300および400を示す。動作フ

50

ロー 300 および 400 は、任意の適切なコンピューティング環境で実施することができる。例えば、動作フローは、図 1 および 2 に示すようなシステムおよび環境によって実行することができる。したがって、動作フロー 300 および 400 に関する記述は、図 1 および 2 のコンポーネントのうち少なくとも 1 つに言及することができる。しかし、図 1 および 2 のコンポーネントへのそのような言及はどれも説明の目的に過ぎず、図 1 および 2 の実装形態は、動作フロー 300 および 400 のための非限定的な環境であることを理解されたい。

【0043】

[0046] さらに、動作フロー 300 および 400 を特定の順序で連続的に図示および記述するが、他の実施形態では、異なる順序で、複数回にわたって、および / または平行して動作が実施されてもよい。さらに、いくつかの実施形態では、1 つまたは複数の動作が省略または結合されてもよい。

10

【0044】

[0047] 動作フロー 300 に、サーバフェールオーバー時に再実行防御を提供するためのステップを示す。実施形態では、図 3 に示すフロー 300 は、サーバクラスタの一部であるサーバ（例えばクラスタ 206 のサーバ 1、サーバ 2、およびサーバ 3（図 2））上で実行されているファイルサーバによって実施することができる。フロー 300 は動作 302 で開始し、ファイルサーバに接続する要求が受け取られる。動作 302 で受け取られる要求は、ファイルサーバを介してアクセス可能な共有ファイルシステムに記憶されたファイル情報にアクセスするために、ファイルサーバとのセッションを確立する要求である。要求は、クライアント、例えばクライアント 202 および 204（図 2）によって送ることができる。動作 302 の後、フロー 300 は動作 304 に進み、セッションが確立されたことを示す応答が送られる。いくつかの実施形態では、動作 302 および 304 で送られる要求および応答は、セッションを交渉するためにクライアントとサーバとの間で交換されるいくつかのメッセージの一部とすることができる。メッセージの交換は、ファイルサーバの永続的ハンドル提供能力を含めた能力の交換を含んでよい。

20

【0045】

[0048] 動作フローは動作 304 から動作 306 に進み、永続的ハンドルを求める第 2 の要求が受け取られる。要求はクライアントによって送られ、要求はまた、クライアントによって生成されたハンドル識別子を含む。ハンドル識別子は、実施形態では、クライアントとファイルサーバとの間で確立されたセッションに関する状態情報を記憶するために、サーバによって使用される。状態情報を記憶することの一部として、フロー 300 は、実施形態では、ファイルサーバがいくらかの状態情報を記憶するために再開フィルタに登録する動作 308 を含んでよい。実施形態では、再開フィルタは、プロトコル層と、基礎をなすストレージシステムとの間に位置し、実施形態では、ファイルサーバとクライアントとの間で確立されたセッションに関する状態情報を記憶するのに使用することができる。

30

【0046】

[0049] 動作 310 で、状態情報は、ハンドル識別子と関連付けて記憶される。状態情報は、テーブル、データベース、またはログファイルなど、任意の適切な形で記憶することができる。記憶域は、永続的であり、必要時に状態を再確立するためにファイルサーバによって利用可能である。状態情報は、ファイルサーバによって直接に記憶することができる。他の実施形態では、フロー 300 は動作 312 を含み、再開フィルタを使用して状態情報が記憶される。上に示したように、ファイルサーバは、いくつかの実施形態では、再開フィルタに登録して状態情報を記憶することができる。

40

【0047】

[0050] フロー 300 は動作 312 から動作 314 に進み、永続的ハンドルを使用したファイルへのアクセスを許可する応答がクライアントに送られる。応答は、ファイルサーバによって応答の中で提供される、また動作 310（および任意選択で動作 312）で記憶された状態情報と関連付けて記憶された、ファイル識別子を含む。

【0048】

50

[0051]次いで、フロー 300 は動作 316 に進み、動作 316 では、任意選択でいくつかのファイルアクセス要求が受信される。ファイルアクセス要求は、永続的ハンドルに関連するファイルに対して実施するいくつかのファイル操作を含んでよい。操作は、例えば、データ読取り/書込みのためのオープン、属性の列挙、データをローカルにキャッシュするのを可能にするためのリース要求、または他のファイルアクセス操作とすることができる。動作 316 におけるファイルアクセス要求の受信に関連する様々な状態を、動作 318 で更新することができる。すなわち、これらの要求がクライアントに許可されると、前の動作(310 および 312)で記憶された状態情報は、追加の状態情報を反映するように更新される。

【0049】

[0052]動作 318 の後に、枠 319 内で識別されるいくつかの追加動作がある。これらの動作は、クライアントがファイルサーバから切断された結果として実施されてよい。理解できるように、動作 302 ~ 318 を元々実施したファイルサーバが障害のせいで利用不可能である状況では、枠 319 内の追加動作は、フェールオーバーサーバによって実施される。障害がネットワーク問題の結果であるような他の実施形態では、枠 319 内の動作は、同じファイルサーバによって実施される。

【0050】

[0053]動作 320 で、再接続する要求が受け取られる。要求は、ファイルサーバによって前に提供されたファイルハンドル、ならびに、クライアントが永続的ハンドルの要求時に使用したハンドル識別子を含む。動作 320 で要求を受け取ったファイルサーバは、ハンドル識別子およびファイル識別子を使用して、状態情報を検索することができる。上に示したように、この動作は、前に再開フィルタを使用して保存された状態情報を取り出すために、再開フィルタを使用することを含んでよい。

【0051】

[0054]フロー 300 は動作 320 から動作 322 に進み、状態情報を使用して、クライアントとの接続および前の状態が再確立される。動作 322 の後、フローは動作 324 に進み、新しいファイルアクセス要求を受け取られる。したがって、動作 324 はいくつかの操作を含む可能性があり、各操作は、ファイルアクセス要求をクライアントから受け取ることを含む。

【0052】

[0055]動作 324 で受け取られる要求のいくつかは、ファイルサーバとクライアントとの間の切断の前に送られた前の要求の再実行である場合がある。この結果、これらの操作のいくつかは、ファイルサーバにおいて再適用されれば、一貫性のない状態を生み出す可能性がある。動作 326 で、再実行である新しいファイルアクセス要求が検出される。実施形態では、この動作は、前にクライアントによって提供されたハンドル識別子を使用してファイルアクセス要求を識別することを含んでよい。動作 326 で再実行が検出されると、動作 328 で、要求は適切に処理される。すなわち、再実行された操作がファイルサーバ上で一貫性のない状態を生み出すことになる場合は、前の操作がうまく実施されたならばこれらの操作を無視することができる。あるいは、前の操作がうまく実施されなかったならば、再実行された操作を適用することができる。次いで、フロー 300 は 330 で終了する。

【0053】

[0056]動作フロー 400 に、一貫した可用性を維持するためのステップを示す。実施形態では、フロー 400 は、共有ファイルシステム中のファイルにアクセスするためにファイルサーバと通信しているクライアント(クライアント 202 および 204 (図 2)など)上のリダイレクタによって実施することができる。クライアントは、実施形態では、SMB プロトコルのバージョンまたは NFS のバージョンなどのファイルアクセスプロトコルを使用してファイルサーバと通信する。

【0054】

[0057]フロー 400 は動作 402 で開始し、ファイルサーバに接続する要求が送られる

10

20

30

40

50

。動作 4 0 2 で送られる要求は、ファイルサーバを介してアクセス可能な共有ファイルシステムに記憶されたファイル情報にアクセスするために、ファイルサーバとのセッションを確立する要求である。要求は、サーバクラスタの一部であるサーバ（例えばサーバ 1、サーバ 2、およびサーバ 3）上のファイルサーバに送ることができる（図 2）。要求は、SMB または FNS のバージョンなどのファイルアクセスプロトコルに従ってフォーマットされる。

【 0 0 5 5 】

[0058]動作 4 0 2 の後、フロー 4 0 0 は動作 4 0 4 に進み、セッションが確立されたことを示す応答が受け取られる。いくつかの実施形態では、動作 4 0 2 および 4 0 4 は、セッションを交渉するためにクライアントとサーバとの間で交換されるいくつかのメッセージの一部とすることができる。メッセージの交換は、ファイルサーバの永続的ハンドル提供能力を含めた能力の交換を含んでよい。

10

【 0 0 5 6 】

[0059]動作フローは動作 4 0 4 から動作 4 0 6 に進み、永続的ハンドルを求める要求を受け取られる。交渉プロセス（動作 4 0 2 および 4 0 4）の結果として、クライアントには、ファイルサーバが永続的ハンドルを提供できることが通知されたものとしてすることができる。クライアント上のアプリケーションが切断および再接続の後にそれらの状態を確実に再確立させることができるようにするために、動作 4 0 6 で、クライアントは永続的ハンドルを要求することができる。要求は、クライアントによって生成されたハンドル識別子を含む。

20

【 0 0 5 7 】

[0060]フロー 4 0 0 は動作 4 0 6 から動作 4 0 8 に進み、永続的ハンドルを使用したファイルへのアクセスを許可する応答を受け取られる。応答は、ファイルサーバによって応答の中で提供されるファイル識別子を含む。

【 0 0 5 8 】

[0061]動作 4 1 0 で、いくつかの実施形態では、状態情報をクライアントによって記憶することができる。状態情報は、ハンドル識別子と関連付けて、かつ、永続的ハンドルを許可する受信済み応答の中で提供されたファイル識別子と関連付けて記憶される。状態情報は、テーブル、データベース、またはログファイルなど、任意の適切な形で記憶することができる。記憶域は、永続的であり、必要時に状態を再確立するためにクライアントによって利用可能である。理解できるように、クライアントによって記憶される状態情報は、実施形態では、ファイルサーバ上で一貫性のない状態を生み出すことなく安全にファイルサーバに対して再実行して返すことのできる操作の状態情報である。再実行される操作は、例えば、データをローカルにキャッシュするためのリース、読取り、書込み、削除、およびメタデータ列挙とすることができる。

30

【 0 0 5 9 】

[0062]フロー 4 0 0 は動作 4 1 0 から動作 4 1 2 に進み、クライアントは、いくつかのファイルアクセス要求を送る。したがって、動作 4 1 2 は、実施形態によれば、ファイル操作を実施する複数の要求を送ることを含んでよい。動作 4 1 2 の後に動作 4 1 4 が続き、クライアント上の状態情報が更新される。理解できるように、動作 4 1 4 は、何度も、すなわち動作 4 1 2 でファイルアクセス要求がクライアントによって送られる度に、行われてよい。

40

【 0 0 6 0 】

[0063]動作 4 1 4 から、フローは動作 4 1 6 に進み、切断が検出される。検出は、タイムアウト、イベント通知、または他の何らかの手段によって行われてよい。動作 4 1 6 に続き、動作 4 1 8 で、接続して前にファイルサーバと確立されたセッションを再確立する要求が送られる。要求は、前にファイルサーバによって提供されたファイルハンドル、ならびに、クライアントが永続的ハンドルの要求時に使用したハンドル識別子を含む。

【 0 0 6 1 】

[0064]フロー 4 0 0 は動作 4 1 8 から動作 4 2 0 に進み、再接続が成功であることが判

50

定される。動作 4 2 0 の後、フローは動作 4 2 2 に進み、クライアントに記憶された状態情報を使用して前の状態が再確立される。動作 4 2 2 は、読取り、書込み、列挙、ロックの要求、または、前の状態を再確立するための他の操作を含めた、いくつかの異なる要求を送ることを含んでよい。フローは動作 4 2 2 から動作 4 2 4 に進み、クライアントは、新しいファイルアクセス要求を送る。フローは 4 2 6 で終了する。

【 0 0 6 2 】

[0065] 図 5 に、本明細書に述べる実施形態を実現するのに使用できる一般的なコンピュータシステムシステム 5 0 0 を示す。コンピュータシステム 5 0 0 は、コンピューティング環境の一例に過ぎず、コンピュータおよびネットワークアーキテクチャーの使用または機能の範囲に関するどんな限定を示唆するものともしない。また、コンピュータシステム 5 0 0 は、この例示的なコンピュータシステム 5 0 0 中に示されるコンポーネントの任意の 1 つまたは組合せに関係するどんな依存または要件を有すると解釈されるべきではない。実施形態では、システム 5 0 0 は、図 1 に関して上述したクライアントおよび/またはサーバとして使用することができる。

10

【 0 0 6 3 】

[0066] システム 5 0 0 は、その最も基本的な構成では通常、少なくとも 1 つの処理ユニット 5 0 2 と、メモリ 5 0 4 とを備える。コンピューティングデバイスの正確な構成およびタイプに応じて、メモリ 5 0 4 は、揮発性 (RAM など)、不揮発性 (ROM、フラッシュメモリ、その他など)、またはこの 2 つの何らかの組合せとすることができる。この最も基本的な構成を、図 5 では破線 5 0 6 によって示す。実施形態では、システムメモリ 5 0 4 は、ファイル情報へのアクセスを要求するアプリケーション 5 2 3 などのアプリケーションを記憶する。システムメモリ 5 0 4 はまた、実施形態により要求をインターセプトしてこれらの要求をファイルサーバに通信するリダイレクタ 5 2 2 も備える。

20

【 0 0 6 4 】

[0067] 本明細書においてコンピュータ可読媒体という用語は、コンピュータ記憶媒体を含み得る。コンピュータ記憶媒体は、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータなどの情報を記憶するための任意の方法または技術で実現された、揮発性および不揮発性、リムーバブルおよびノンリムーバブル媒体を含み得る。システムメモリ 5 0 4、リムーバブル記憶装置、およびノンリムーバブル記憶装置 5 0 8 は全て、コンピュータ記憶媒体の例 (すなわちメモリ記憶装置) である。実施形態では、データ、例えば状態情報 5 2 0 などが記憶される。コンピュータ記憶媒体は、RAM、ROM、電氣的に消去可能な読取可能メモリ (EEPROM)、フラッシュメモリ、もしくは他のメモリ技術、CD-ROM、デジタル多用途ディスク (DVD)、もしくは他の光学記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、もしくは他の磁気記憶デバイス、または、情報を記憶するのに使用できコンピュータシステム 5 0 0 からアクセスできる他の任意の媒体を含み得るが、これらに限定されない。このようなコンピュータ記憶媒体はどれも、デバイス 5 0 0 の一部となり得る。コンピューティングデバイス 5 0 0 はまた、キーボード、マウス、ペン、音声入力デバイス、タッチ入力デバイスなどの、入力デバイス (複数可) 5 1 4 を有することができる。表示装置、スピーカー、プリンターなどの、出力デバイス (複数可) 5 1 6 も備わってよい。前述のデバイスは例であり、他のデバイスが使用されてもよい。

30

40

【 0 0 6 5 】

[0068] 本明細書においてコンピュータ可読媒体という用語はまた、通信媒体を含み得る。通信媒体は、搬送波や他のトランスポート機構などの被変調データ信号中の、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータによって具体化することができる、任意の情報送達媒体を含む。用語「被変調データ信号」は、信号中の情報を符号化するようにして 1 つまたは複数の特性が設定または変更される信号を記述することができる。限定ではなく例として、通信媒体は、有線ネットワークや直接有線接続などの有線媒体と、音響、無線周波数 (RF)、赤外線および他のワイヤレス媒体などのワイヤレス媒体とを含み得る。

50

【 0 0 6 6 】

[0069]本明細書の全体を通して、「一実施形態」または「実施形態」に言及したが、これらは、述べた特定の特徴、構造、または特性が、少なくとも1つの実施形態に含まれることを意味する。したがって、このような句の使用は、1つだけでなく複数の実施形態に言及することができる。さらに、述べた特徴、構造、または特性は、1つまたは複数の実施形態で、任意の適切な方式で組み合わせることができる。

【 0 0 6 7 】

[0070]しかし、具体的詳細の1つもしくは複数がなくとも、または、他の方法、リソース、材料などと共に、実施形態を実践できることを、当業者なら認識するであろう。他の場合には、周知の構造、リソース、または動作については詳細に図示または記述していないが、これは単に、実施形態の態様を曖昧にするのを避けるためである。

【 0 0 6 8 】

[0071]例示的な実施形態および適用例を図示および記述したが、実施形態は前述の正確な構成およびリソースに限定されないことを理解されたい。特許請求される実施形態の範囲を逸脱することなく、当業者に明らかな様々な修正、変更、および変形を、本明細書に開示する方法およびシステムの構成、動作、および詳細に加えることができる。

【 図 1 】

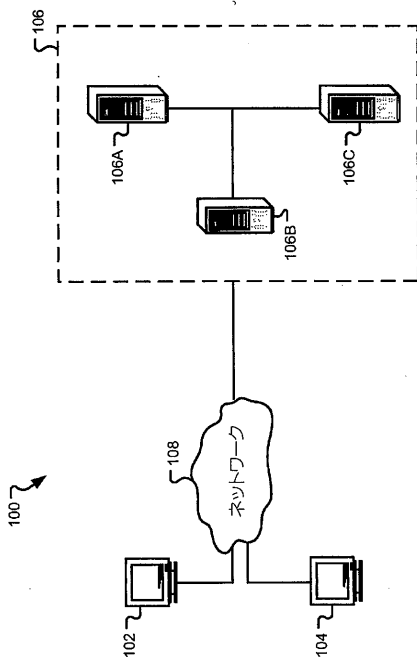


FIG.1

【 図 2 】

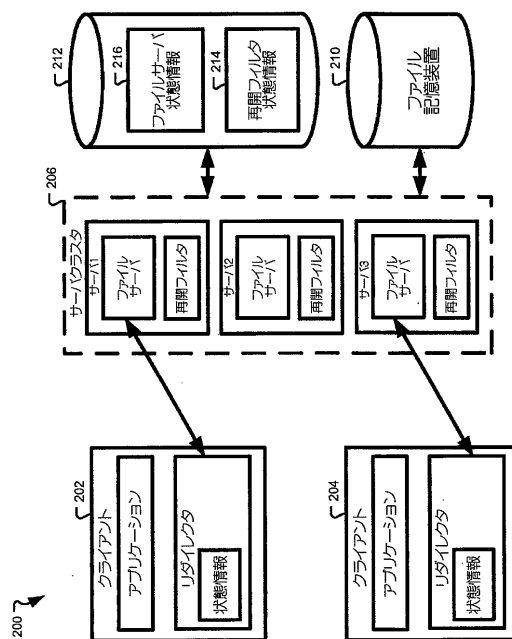


FIG. 2

【 図 3 】

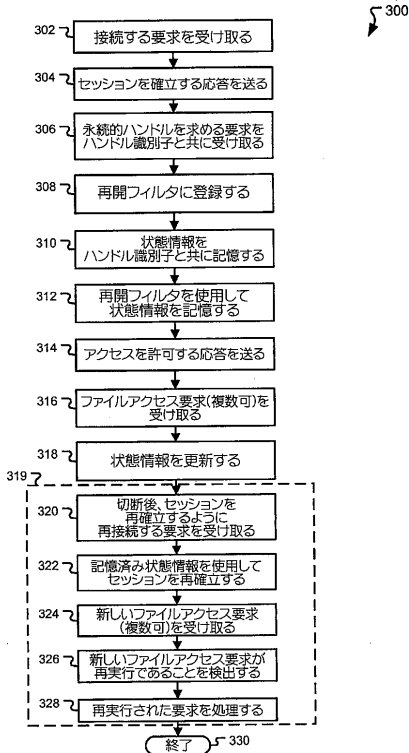


FIG.3

【 図 4 】

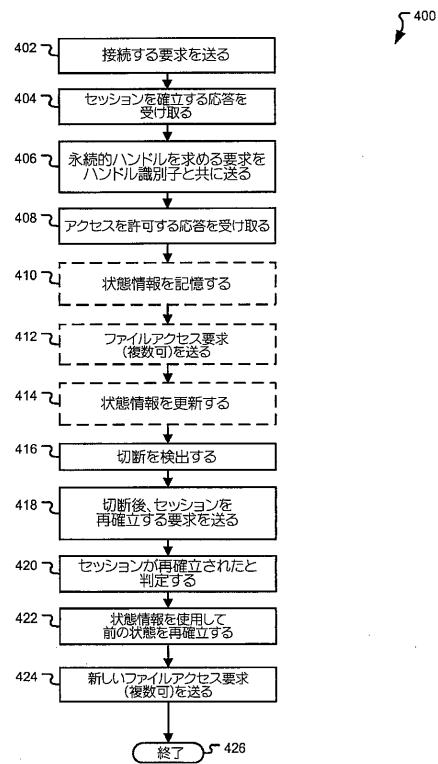


FIG.4

【 図 5 】

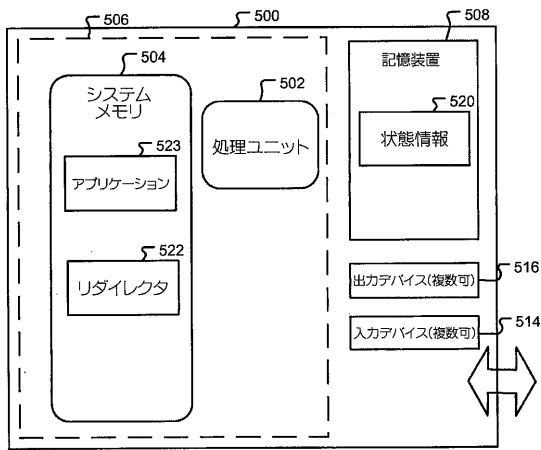


FIG.5

【手続補正書】**【提出日】**平成28年8月9日(2016.8.9)**【手続補正1】****【補正対象書類名】**特許請求の範囲**【補正対象項目名】**全文**【補正方法】**変更**【補正の内容】****【特許請求の範囲】****【請求項1】**

サーバクラス上の共有ファイルシステムにアクセスするクライアントに一貫した可用性を提供するコンピュータ実施方法であって、

クライアントとのセッションを確立するステップと、

ファイルサーバにおいて、前記共有ファイルシステム中のファイルに前記クライアントがアクセスするために前記ファイルサーバ上で永続的ハンドルを開く要求を受けるステップであって、前記要求は、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

前記要求を受けるのに応答して、前記ファイルサーバが、

前記セッションに関する第1の状態情報を前記ハンドル識別子と関連付けて記憶し、
前記ファイルへのアクセスを許可する応答を前記クライアントに送る

ステップと、

クライアント切断後に、前記永続的ハンドルを使用して前記セッションを再確立する第2の要求を受けるステップと、

前記第1の状態情報を使用して前記セッションを再確立するステップと、

前記セッションの再確立の後に、前記ハンドル識別子を含む新しい要求を前記クライアントから受けるステップと、

前記ハンドル識別子を使用して、前記新しい要求が以前に送られた要求の再実行であると判定するステップと

を含む方法。

【請求項2】

請求項1に記載の方法であって、前記第1の状態情報が、前記クライアントによって再送されれば前記ファイルサーバが一貫性のない状態になるような操作の状態を含む、方法。

【請求項3】

請求項2に記載の方法であって、前記セッションはファイルアクセスプロトコルを使用して確立され、前記ファイルアクセスプロトコルは、サーバメッセージブロック(SMB)プロトコルのあるバージョンである、方法。

【請求項4】

請求項3に記載の方法であって、前記第1の状態情報は、前記クライアントによってなされるバイト範囲ロックと、前記クライアントによって送られる作成の操作とのうちの1以上の状態を含む、方法。

【請求項5】

請求項3に記載の方法であって、前記クライアントの切断の前に、前記セッションに関する第2の状態情報を記憶するために再開フィルタを登録するステップを更に含む方法。

【請求項6】

請求項5に記載の方法であって、前記第2の状態情報は、前記クライアントによってなされるリースと、前記クライアントによってなされるオープンとのうちの1以上の状態を含む、方法。

【請求項7】

請求項5に記載の方法であって、前記セッションを再確立する前記ステップは、前記第2の状態情報を使用するステップを更に含む、方法。

【請求項 8】

請求項 1 に記載の方法であって、前記クライアントの切断は前記ファイルサーバの障害のせいで発生し、前記セッションの前記再確立は、前記サーバクラスタ中の第 2 のファイルサーバによって実施される、方法。

【請求項 9】

コンピュータ実行可能命令を含むコンピュータ可読記憶媒体であって、前記コンピュータ実行可能命令は、プロセッサにより実行されると、一貫した状態を維持する方法を実施させ、前記方法は、

ファイル情報にアクセスするためにサーバに接続する第 1 の要求をクライアントが送るステップであって、前記第 1 の要求はファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記サーバから第 1 の応答を受けるステップであって、前記第 1 の応答は、前記サーバ上のファイル情報へのアクセスを可能にするために前記クライアントとのセッションを確立するステップであって、前記第 1 の応答は前記ファイルアクセスプロトコルに従ってフォーマットされた、ステップと、

前記クライアントが前記サーバ上のファイルにアクセスするために前記サーバ上で永続的ハンドルを開く第 2 の要求を送るステップであって、前記第 2 の要求は、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

前記ファイルへのアクセスを許可する第 2 の応答を前記クライアントにおいて受けるステップと、

前記ファイルに関するべき等操作のための第 3 の要求を前記サーバに送るステップと、前記クライアントにおいて、前記べき等操作に関係した状態情報を記憶するステップと、

前記クライアントが前記サーバから切断されたことを検出するステップと、

前記永続的ハンドルを使用して前記セッションを再確立する第 4 の要求を送るステップであって、前記セッションを再確立する前記第 4 の要求は、前記ハンドル識別子を含む、ステップと、

前記セッションが再確立されたと判定するステップと、

前記ハンドル識別子及び前記状態情報を含む、前記べき等操作のための新しい要求を送るステップと

を含む、コンピュータ可読記憶媒体。

【請求項 10】

請求項 9 に記載のコンピュータ可読記憶媒体であって、前記サーバはサーバクラスタ中の複数のサーバのうちの 1 つであり、前記切断は前記サーバの障害のせいで発生し、前記セッションの前記再確立は、前記サーバクラスタ中の前記複数のサーバのうちの第 2 のサーバによって実施される、コンピュータ可読記憶媒体。

【請求項 11】

請求項 10 に記載のコンピュータ可読記憶媒体であって、前記ファイルアクセスプロトコルは、サーバメッセージブロック (SMB) プロトコルのあるバージョンである、コンピュータ可読記憶媒体。

【請求項 12】

請求項 9 に記載のコンピュータ可読記憶媒体であって、前記セッションが再確立した後に、前記新しい要求は前記第 3 の要求の再実行であると判定される、コンピュータ可読記憶媒体。

【請求項 13】

請求項 12 に記載のコンピュータ可読記憶媒体であって、前記新しい要求は、前記切断の前に前記クライアントによってなされるバイト範囲ロックと、前記切断の前に前記クライアントによって送られる作成の要求とのうちの 1 以上の要求の再実行である、コンピュータ可読記憶媒体。

【請求項 14】

サーバクラスタ上の共有ファイルシステムにアクセスするクライアントに一貫した可用性を提供するためのシステムであって、該システムは、

少なくとも1つのプロセッサと、

前記少なくとも1つのプロセッサに動作上結合され、前記少なくとも1つのプロセッサにより実行されると前記少なくとも1つのプロセッサにある方法を実施させる命令を記憶したメモリと

を備え、前記方法は、

クライアントとのセッションを確立するステップと、

ファイルサーバにおいて、前記共有ファイルシステム中のファイルに前記クライアントがアクセスするために前記ファイルサーバ上で永続的ハンドルを開く要求を受けるステップであって、前記要求は、前記クライアントによって提供されたハンドル識別子を含む、ステップと、

前記要求を受けるのに応答して、前記ファイルサーバが、

前記セッションに関する第1の状態情報を前記ハンドル識別子と関連付けて記憶し、

前記ファイルへのアクセスを許可する応答を前記クライアントに送る

ステップと、

クライアント切断後に、前記永続的ハンドルを使用して前記セッションを再確立する第2の要求を受けるステップと、

前記第1の状態情報を使用して前記セッションを再確立するステップと、

前記セッションの再確立の後に、前記ハンドル識別子を含む新しい要求を前記クライアントから受けるステップと、

前記ハンドル識別子を使用して、前記新しい要求が以前に送られた要求の再実行であると判定するステップと

を含む、システム。

【請求項15】

請求項14に記載のシステムであって、前記第1の状態情報は、前記クライアントによって再送されれば前記ファイルサーバが一貫性のない状態になるような操作の状態を含む、システム。

【請求項16】

請求項15に記載のシステムであって、前記セッションはファイルアクセスプロトコルを使用して確立され、前記ファイルアクセスプロトコルは、サーバメッセージブロック(SMB)プロトコルのあるバージョンである、システム。

【請求項17】

請求項16に記載のシステムであって、前記第1の状態情報は、前記クライアントによってなされるバイト範囲ロックと、前記クライアントによって送られる作成の操作との中の1以上の状態を含む、システム。

【請求項18】

請求項16に記載のシステムであって、前記方法は、

前記クライアントの切断の前に、前記セッションに関する第2の状態情報を記憶するために再開フィルタを登録するステップ

を更に含むシステム。

【請求項19】

請求項18に記載のシステムであって、前記第2の状態情報は、前記クライアントによってなされるリースと、前記クライアントによってなされるオープンとの中の1以上の状態を含む、システム。

【請求項20】

請求項18に記載のシステムであって、前記セッションを再確立する前記ステップは、前記第2の状態情報を使用するステップを更に含む、システム。

フロントページの続き

- (74)代理人 100153028
弁理士 上田 忠
- (72)発明者 ジョージ, マシュー
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 クルーズ, デーヴィッド・エム
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 ピンカートン, ジェームズ・ティー
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 バッテパティ, ルーペシュ・シー
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 ジョリー, トム
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 スワン, ポール・アール
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 シャーン, ミンドン
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント
- (72)発明者 ロヴィンガー, ダニエル・エドワード
アメリカ合衆国ワシントン州 9 8 0 5 2 - 6 3 9 9 , レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイ - インターナショナル・パテント

Fターム(参考) 5B034 CC04 DD06