

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6357587号
(P6357587)

(45) 発行日 平成30年7月11日(2018.7.11)

(24) 登録日 平成30年6月22日(2018.6.22)

(51) Int. Cl.	F I
G06F 12/00 (2006.01)	G06F 12/00 535B
	G06F 12/00 545A

請求項の数 25 (全 31 頁)

(21) 出願番号	特願2017-522597 (P2017-522597)	(73) 特許権者	503433420
(86) (22) 出願日	平成27年12月30日(2015.12.30)		華為技術有限公司
(65) 公表番号	特表2018-503887 (P2018-503887A)		HUAWEI TECHNOLOGIES CO., LTD.
(43) 公表日	平成30年2月8日(2018.2.8)		中華人民共和国 518129 広東省深 ▲チェン▼市龍崗区坂田 華為総部▲ベン ▼公楼
(86) 国際出願番号	PCT/CN2015/100006		Huawei Administrati on Building, Bantia n, Longgang Distric t, Shenzhen, Guangd ong 518129, P. R. Ch ina
(87) 国際公開番号	W02017/113261	(74) 代理人	100107766
(87) 国際公開日	平成29年7月6日(2017.7.6)		弁理士 伊東 忠重
審査請求日	平成29年4月26日(2017.4.26)		最終頁に続く

(54) 【発明の名称】 獲得ロック要求を処理するための方法及びサーバ

(57) 【特許請求の範囲】

【請求項 1】

ロック要求を処理するための方法であり、第1のロックサーバに適用可能であり、前記第1のロックサーバは、第2のロックサーバの引き継ぎロックサーバであり、前記第1のロックサーバは、前記第2のロックサーバのロック管理範囲を記憶する方法であって、

前記第1のロックサーバにより、故障が前記第2のロックサーバで発生したことを習得した後にサイレント状態に入るステップであり、前記サイレント状態は、前記第1のロックサーバがロック権限を再申請するための再請求ロック要求のみを処理することができ、新たなロック権限を申請するための獲得ロック要求を処理することができないことを示し、前記サイレント状態のサイレント範囲は、前記第2のロックサーバが権限を割り当てて いるリソースであるステップと、

前記第1のロックサーバにより、第1の獲得ロック要求を受信するステップであり、前記第1の獲得ロック要求は、第1のリソースをロックすることを要求するために使用され、前記第1の獲得ロック要求は、前記第1のリソースの識別子を搬送するステップと、

前記第1のロックサーバにより、前記第1のリソースが前記第2のロックサーバの前記管理範囲に属すると検出するステップと、

前記第1のロックサーバにより、第1のリソース情報記録リストに問い合わせるステップであり、前記第1のリソース情報記録リストは、前記第2のロックサーバがロック権限を割り当てている前記リソースのリソース識別子を記録し、前記第1のリソース情報記録リストが前記第1のリソースの前記識別子を記録していない場合、前記第1のロックサー

10

20

パにより、前記第 1 の獲得ロック要求に従って前記第 1 のリソースについてロック権限を割り当てるステップと

を含む方法。

【請求項 2】

前記第 1 のロックサーバにより、サイレント状態に入った後に、

前記第 1 のロックサーバにより、第 2 の獲得ロック要求を受信するステップであり、前記第 2 の獲得ロック要求は、第 2 のリソースをロックすることを要求するために使用され、前記第 2 の獲得ロック要求は、前記第 2 のリソースの識別子を搬送するステップと、

前記第 1 のロックサーバにより、前記第 2 のリソースが前記第 1 のロックサーバの管理範囲に属すると検出するステップと、

前記第 1 のロックサーバにより、前記第 2 の獲得ロック要求に従って前記第 2 のリソースについてロック権限を割り当てるステップと

を更に含む、請求項 1 に記載のロック要求を処理するための方法。

【請求項 3】

前記第 1 のロックサーバにより、サイレント状態に入った後に、

前記第 1 のロックサーバにより、第 3 の獲得ロック要求を受信するステップであり、前記第 3 の獲得ロック要求は、第 3 のリソースをロックすることを要求するために使用され、前記第 3 の獲得ロック要求は、前記第 3 のリソースの識別子を搬送するステップと、

前記第 1 のロックサーバにより、前記第 3 のリソースが前記第 2 のロックサーバの前記管理範囲に属すると検出するステップと、

前記第 1 のロックサーバにより、前記第 1 のリソース情報記録リストに問い合わせ、前記第 1 のリソース情報記録リストが前記第 3 の獲得ロック要求で要求された前記リソースの前記リソース識別子を記録している場合、前記第 1 のロックサーバにより、前記第 3 の獲得ロック要求に従って前記第 3 のリソースについてロック権限を割り当てることを拒否するステップと

を更に含む、請求項 1 に記載のロック要求を処理するための方法。

【請求項 4】

前記第 1 のロックサーバにより、前記第 1 のリソースの前記識別子を第 2 のリソース情報記録リストに記録するステップであり、前記第 2 のリソース情報記録リストは、前記第 1 のロックサーバがロック権限を割り当てているリソースのリソース識別子を記録するために使用され、前記第 2 のリソース情報記録リストは、第 3 のロックサーバに記憶されるステップを更に含む、請求項 1 乃至 3 のうちいずれか 1 項に記載のロック要求を処理するための方法。

【請求項 5】

前記第 1 のロックサーバにより、前記第 2 のロックサーバの前記ロック管理範囲を記憶するステップは、

前記第 1 のロックサーバにより、第 1 の通知メッセージを受信するステップであり、前記第 1 の通知メッセージは、前記第 2 のロックサーバの識別情報を搬送するステップと、

前記第 1 のロックサーバにより、前記第 2 のロックサーバの前記識別情報及びロックサーバ引き継ぎ関係に従って、前記第 1 のロックサーバが前記第 2 のロックサーバの前記引き継ぎロックサーバであると決定するステップと、

前記第 1 のロックサーバにより、前記第 2 のロックサーバの前記ロック管理範囲を受信し、前記ロック管理範囲を記憶するステップと

を含む、請求項 1 乃至 3 のうちいずれか 1 項に記載のロック要求を処理するための方法。

【請求項 6】

前記方法の前に、

プロトコルサーバにより、ホストからパケットを受信し、前記パケットから前記第 1 の獲得ロック要求を解析するステップと、

前記プロトコルサーバにより、前記第 1 の獲得ロック要求をロックプロキシに転送する

10

20

30

40

50

ステップと、

前記ロックプロキシにより、前記第 1 の獲得ロック要求で搬送された前記第 1 のリソースの前記識別子に従って決定を実行し、前記第 1 のロックサーバが前記第 1 のリソースを管理すると決定した場合、前記第 1 の獲得ロック要求を前記第 1 のロックサーバに送信するステップと

を更に含む、請求項 1 乃至 3 のうちいずれか 1 項に記載のロック要求を処理するための方法。

【請求項 7】

前記第 1 のロックサーバにより、サイレント状態に入った後に、

再請求ロック要求を受信するステップであり、前記再請求ロック要求は、第 4 のリソースの識別子と、前記第 4 のリソースについて前記第 2 のロックサーバにより割り当てられた権限とを搬送し、前記第 4 のリソースは、前記第 2 のロックサーバが権限を割り当てているリソースであるステップと、

前記第 2 のロックサーバにより割り当てられている前記権限に従って前記第 4 のリソースについて同じ権限を再割り当てするステップと

を更に含む、請求項 1 乃至 3 のうちいずれか 1 項に記載のロック要求を処理するための方法。

【請求項 8】

前記第 2 のロックサーバが権限を割り当てている全てのリソースについて権限が再割り当てされた後に、前記第 1 のロックサーバにより、前記サイレント状態を出るステップ、又は

予め設定された時間の後に、前記第 1 のロックサーバにより、前記サイレント状態を出るステップを更に含む、請求項 7 に記載のロック要求を処理するための方法。

【請求項 9】

前記第 1 のロックサーバにより、前記サイレント状態を出た後に、

前記第 1 のロックサーバにより、前記第 1 のロックサーバの前記管理範囲を更新するステップであり、前記第 1 のロックサーバの更新された管理範囲は、前記更新する前の前記第 1 のロックサーバの前記管理範囲と、前記第 2 のロックサーバの前記管理範囲とを含むステップを更に含む、請求項 8 に記載のロック要求を処理するための方法。

【請求項 10】

他のロック要求管理装置のロック要求を引き継ぐように構成されたロック要求管理装置であって、

第 1 の獲得ロック要求及び通知メッセージを受信するように構成された受信モジュールであり、前記第 1 の獲得ロック要求は、第 1 のリソースの識別子を搬送する受信モジュールと、

前記他のロック要求管理装置のロック管理範囲及び第 1 のリソース情報記録リストを記録するように構成された記憶モジュールであり、前記第 1 のリソース情報記録リストは、前記他のロック要求管理装置がロック権限を割り当てているリソースのリソース識別子を記録する記憶モジュールと、

前記記憶モジュールにより受信された獲得ロック要求を処理するように構成されたロック要求処理モジュールと、

前記通知メッセージに従って、故障が前記他のロック要求管理装置で発生したことが習得された後に、前記ロック要求管理装置をサイレント状態に設定するように構成されたサイレントモジュールであり、前記サイレント状態は、前記ロック要求管理装置がロック権限を再申請するための再請求ロック要求のみを処理することができ、新たなロック権限を申請するための獲得ロック要求を処理することができないことを示し、サイレント範囲は、前記他のロック要求管理装置が権限を割り当てている前記リソースであるサイレントモジュールと

を含み、

前記ロック要求管理装置が前記サイレント状態に入った後に、前記ロック要求処理モジ

10

20

30

40

50

ジュールは、

前記獲得ロック要求が前記サイレント範囲に属する場合、前記第1のリソース情報記録リストに問い合わせ、前記第1のリソース情報記録リストが前記第1のリソースの前記識別子を記録していない場合、前記第1の獲得ロック要求に従って前記第1のリソースについてロック権限を割り当てるように構成される、ロック要求管理装置。

【請求項11】

前記ロック要求管理装置が前記サイレント状態に入った後に、

前記受信モジュールは、第2の獲得ロック要求を受信するように更に構成され、前記第2の獲得ロック要求は、第2のリソースをロックすることを要求するために使用され、前記第2の獲得ロック要求は、前記第2のリソースの識別子を搬送し、

前記ロック要求処理モジュールは、前記第2のリソースが前記ロック要求管理装置の管理範囲に属すると検出した後に、前記第2の獲得ロック要求に従って前記第2のリソースについてロック権限を割り当てるように更に構成される、請求項10に記載のロック要求管理装置。

【請求項12】

前記ロック要求管理装置が前記サイレント状態に入った後に、

前記受信モジュールは、第3の獲得ロック要求を受信するように更に構成され、前記第3の獲得ロック要求は、第3のリソースをロックすることを要求するために使用され、前記第3の獲得ロック要求は、前記第3のリソースの識別子を搬送し、

前記ロック要求処理モジュールは、前記第3のリソースが前記他のロック要求管理装置の前記管理範囲に属すると検出した後に、前記第1のリソース情報記録リストに問い合わせ、前記第1のリソース情報記録リストが前記第3の獲得ロック要求で要求された前記リソースの前記リソース識別子を記録している場合、前記第3の獲得ロック要求に従って前記第3のリソースについてロック権限を割り当てることを拒否するように更に構成される、請求項10に記載のロック要求管理装置。

【請求項13】

前記受信モジュールは、第1の通知メッセージを受信するように更に構成され、前記第1の通知メッセージは、前記他のロック要求管理装置の識別情報を搬送し、

前記受信モジュールは、前記他のロック要求管理装置の前記識別情報及びロック要求管理装置の引き継ぎ関係に従って、前記ロック要求管理装置が前記他のロック要求管理装置の前記引き継ぎロック要求管理装置であると決定された後に、前記他のロック要求管理装置の前記ロック管理範囲を前記記憶モジュールに送信するように更に構成され、

前記記憶モジュールが前記他のロック要求管理装置の前記ロック管理範囲を記憶するように構成されることは、

前記記憶モジュールが前記受信モジュールから前記他のロック要求管理装置の前記ロック管理範囲を受信し、前記ロック管理範囲を記憶するように構成されることを含む、請求項10乃至12のうちいずれか1項に記載のロック要求管理装置。

【請求項14】

前記ロック要求管理装置は、プロトコルサーバモジュール及びロックプロキシモジュールを更に含み、

前記プロトコルサーバモジュールは、ホストからパケットを受信し、前記パケットから前記第1の獲得ロック要求を解析し、前記第1の獲得ロック要求を前記ロックプロキシモジュールに転送するように構成され、

前記ロックプロキシモジュールは、前記第1の獲得ロック要求で搬送された前記第1のリソースの前記識別子に従って決定を実行し、前記ロック要求処理モジュールが前記第1のリソースを管理すると決定した場合、前記受信モジュールを使用することにより、前記第1の獲得ロック要求を前記ロック要求処理モジュールに送信するように構成される、請求項10乃至12のうちいずれか1項に記載のロック要求管理装置。

【請求項15】

前記ロック要求処理モジュールは、前記ロック要求管理装置が前記サイレント状態に入

10

20

30

40

50

った後に、

再請求ロック要求を受信するように更に構成され、前記再請求ロック要求は、第4のリソースの識別子と、前記第4のリソースについて前記他のロック要求管理装置により割り当てられた権限とを搬送し、前記第4のリソースは、前記他のロック要求管理装置が権限を割り当てているリソースであり、

前記他のロック要求管理装置により割り当てられている前記権限に従って前記第4のリソースについて同じ権限を再割り当てするように更に構成される、請求項10乃至12のうちいずれか1項に記載のロック要求管理装置。

【請求項16】

前記サイレントモジュールは、

前記他のロック要求管理装置が権限を割り当てている全てのリソースについて権限が再割り当てされた後に、前記ロック要求管理装置が前記サイレント状態を出ることを可能にするように、或いは

予め設定された時間の後に、前記ロック要求管理装置が前記サイレント状態を出ることを可能にするように更に構成される、請求項15に記載のロック要求管理装置。

【請求項17】

前記記憶モジュールは、前記ロック要求管理装置が前記サイレント状態を出た後に、

前記ロック要求管理装置の前記管理範囲を更新するように更に構成され、前記ロック要求管理装置の更新された管理範囲は、前記ロック要求管理装置の前記管理範囲と、前記他のロック要求管理装置の前記管理範囲とを含む、請求項16に記載のロック要求管理装置

【請求項18】

他のサーバの引き継ぎロック管理サーバであるサーバであって、

獲得ロック要求を受信するように構成されたインタフェースと、

前記他のサーバのロック管理範囲及び第1のリソース情報記録リストを記憶するように構成されたメモリであり、前記第1のリソース情報記録リストは、前記他のサーバがロック権限を割り当てているリソースのリソース識別子を記録するメモリと、

以下のステップ、すなわち、

故障が前記他のサーバで発生したことを習得した後に、前記サーバをサイレント状態に設定するステップであり、前記サイレント状態は、前記サーバがロック権限を再申請するための再請求ロック要求のみを処理することができ、新たなロック権限を申請するための獲得ロック要求を処理することができないことを示し、前記サイレント状態のサイレント範囲は、前記他のサーバが前記権限を割り当てている前記リソースであるステップと、

第1の獲得ロック要求を受信するステップであり、前記第1の獲得ロック要求は、第1のリソースをロックすることを要求するために使用され、前記第1の獲得ロック要求は、前記第1のリソースの識別子を搬送するステップと、

前記第1のリソースが前記他のロックサーバの前記管理範囲に属すると検出するステップと、

前記第1のリソース情報記録リストに問い合わせ、前記第1のリソース情報記録リストが前記第1のリソースの前記識別子を記録していない場合、前記第1の獲得ロック要求に従って前記第1のリソースについてロック権限を割り当てするステップと

を実行するためのプログラムを実行するように構成されたプロセッサと

を含むサーバ。

【請求項19】

前記サーバが前記サイレント状態に入った後に、前記プロセッサは、

第2の獲得ロック要求を受信するように更に構成され、前記第2の獲得ロック要求は、第2のリソースをロックすることを要求するために使用され、前記第2の獲得ロック要求は、前記第2のリソースの識別子を搬送し、

前記第2のリソースが前記サーバの管理範囲に属すると検出し、

前記第2の獲得ロック要求に従って前記第2のリソースについてロック権限を割り当て

10

20

30

40

50

るように更に構成される、請求項 18 に記載のサーバ。

【請求項 20】

前記サーバが前記サイレント状態に入った後に、前記プロセッサは、

第 3 の獲得ロック要求を受信するように更に構成され、前記第 3 の獲得ロック要求は、第 3 のリソースをロックすることを要求するために使用され、前記第 3 の獲得ロック要求は、前記第 3 のリソースの識別子を搬送し、

前記第 3 のリソースが前記他のサーバの前記管理範囲に属すると検出し、

前記第 1 のリソース情報記録リストに問い合わせ、前記第 1 のリソース情報記録リストが前記第 3 の獲得ロック要求で要求された前記リソースの前記リソース識別子を記録している場合、前記第 3 の獲得ロック要求に従って前記第 3 のリソースについてロック権限を割り当てることを拒否するように更に構成される、請求項 18 に記載のサーバ。

10

【請求項 21】

前記メモリが前記他のサーバの前記ロック管理範囲を記憶するように構成されることは、

前記プロセッサが、第 1 の通知メッセージを受信するように構成され、前記第 1 の通知メッセージは、前記他のサーバの識別情報を搬送し、

前記プロセッサが、前記他のサーバの前記識別情報及びサーバ引き継ぎ関係に従って、前記サーバが前記他のサーバの前記引き継ぎサーバであると決定された後に、前記他のサーバの前記ロック管理範囲を前記メモリに送信するように構成され、

前記メモリは、前記他のサーバの前記ロック管理範囲を受信し、前記ロック管理範囲を記憶するように構成されることを含む、請求項 18 乃至 20 のうちいずれか 1 項に記載のサーバ。

20

【請求項 22】

前記サーバは、

ホストからパケットを受信し、前記パケットから前記第 1 の獲得ロック要求を解析するように構成されたプロトコルサーバモジュールであり、前記第 1 の獲得ロック要求をロックプロキシモジュールに転送するように更に構成されるプロトコルサーバモジュールと、

前記第 1 の獲得ロック要求で搬送された前記第 1 のリソースの前記識別子に従って決定を実行し、前記サーバが前記第 1 のリソースを管理すると決定した場合、前記第 1 の獲得ロック要求を前記インタフェースに送信するように構成された前記ロックプロキシモジュールと

30

を更に含む、請求項 18 乃至 20 のうちいずれか 1 項に記載のサーバ。

【請求項 23】

前記サーバが前記サイレント状態に入った後に、前記プロセッサは、

再請求ロック要求を受信するように更に構成され、前記再請求ロック要求は、第 4 のリソースの識別子と、前記第 4 のリソースについて前記他のサーバにより割り当てられた権限とを搬送し、前記第 4 のリソースは、前記他のサーバが権限を割り当てているリソースであり、

前記他のサーバにより割り当てられている前記権限に従って前記第 4 のリソースについて同じ権限を再割り当てするように更に構成される、請求項 18 乃至 20 のうちいずれか 1 項に記載のサーバ。

40

【請求項 24】

前記プロセッサは、

前記他のサーバが権限を割り当てている全てのリソースについて権限が再割り当てされた後に、前記サーバが前記サイレント状態を出ることを可能にするように、或いは

予め設定された時間の後に、前記サーバが前記サイレント状態を出ることを可能にするように更に構成される、請求項 23 に記載のサーバ。

【請求項 25】

前記サーバが前記サイレント状態を出た後に、前記プロセッサは、

前記サーバの前記管理範囲を更新するように更に構成され、前記サーバの更新された管

50

理範囲は、前記サーバの前記管理範囲と、前記他のサーバの前記管理範囲とを含む、請求項 2 4 に記載のサーバ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、コンピュータ技術に関し、特にロック要求を処理するための方法及びサーバに関する。

【背景技術】

【0002】

複数のホストを含む分散システムでは、同時に複数のノードによる同じリソースへの相互排他的アクセスは、ロックサーバを使用することにより実現される。リソース上でいくつかの操作を実行する必要がある場合、ホストは、まず、ロックサーバからのロック権限を要求する必要がある。ホストがロック権限を取得した後にのみ、ホストは、リソース上で対応する操作、例えば、読み取り操作又は書き込み操作を実行することができる。したがって、ロックサーバの性能、可用性及び信頼性は、全体の分散システムの性能、可用性及び信頼性に直接影響を与える。

10

【0003】

分散システムでは、ホストは、NAS (Network Attached Storage、ネットワークアタッチストレージ) ネットワークを使用することによりノードと通信する。ロックサーバは、各ノードに配置され、各ノードは、記憶システムに更に接続され、リソース、例えば、ファイルは、記憶システムに記憶される。

20

【0004】

記憶システム内のリソース上で操作 (例えば、読み取り操作又は書き込み操作) を実行する必要がある場合、ホストは、まず、ホスト上のアプリケーションプログラムを使用することによりロック権限をロックサーバに申請し、次に、リソースについてロックサーバにより割り当てられたロック権限を取得した後に、ファイル上で操作を実行する。リソースについて割り当てられているロック権限とアプリケーションプログラムとの間の対応は、各ノードに記憶されてもよく、或いは各ノードによりアクセスされることができる共有ストレージに記憶されてもよい。例えば、記憶システム内のファイル上で読み取り操作を実行する必要がある場合、ホストは、まず、ファイルについてロック権限をノード内のロックサーバに申請し、ホストがファイルについてロック権限を取得した後にのみ、ファイル上で読み取り操作を実行することができる。ファイルについてのロック権限とロック権限を有するアプリケーションとの間の対応は、ノードに記憶され、ノード又はノード上のアプリケーションは、ロック権限を有する。ノードがロック権限を有する場合であっても、ノードは、更なる分析によって、ノード上のどのアプリケーションが記憶システム内のリソースを使用する必要があるかを依然として習得することができる。

30

【0005】

故障がロックサーバで発生した場合、故障のあるロックサーバ上のサービスは、故障が発生していないロックサーバ (以下では故障のないロックサーバと呼ばれる) に切り替えられる必要がある。NFS (Network File System、ネットワークファイルシステム) 又は Samba のようなプロトコルが使用され、故障のあるロックサーバ上のサービスが故障のないロックサーバに切り替えられた場合、ホストは、ホストのアクセス効率を増加させるために、再請求ロック要求を使用することにより、各アプリケーションによりファイルについて取得されているロック権限を再申請し得る。このように、ロック権限が不適切に制御されたという理由で複数のアプリケーションにより閲覧されるデータが一致せず、複数のアプリケーションが同時にデータを読み書きした場合にデータ破損すらもたらされるという問題を回避するため、分散ロックサーバクラスでは、再請求ロック要求及び獲得ロック要求は、安全に制御される必要がある。

40

【0006】

従来技術では、故障がロックサーバで発生した場合、分散システム内の全ての残りの口

50

ックサーバは、サイレント状態に入る。サイレント中に、分散システム内の全てのロックサーバは、ロック権限を再申請するための再請求ロック要求のみを処理することができ、新たなロック権限を申請するための獲得ロック要求を処理することができない。故障のあるロックサーバにより許可されているロック権限のための再請求ロック要求における処理が完了した後にのみ、分散システム内のロックサーバは、サイレント状態を出て、通常通り獲得ロック要求を処理する。

【0007】

前述の従来技術の解決策では、故障が分散システム内のロックサーバで発生した場合、分散システム内の全ての残りのロックサーバは、サイレント状態に入る必要がある。この場合、ロックサーバは、いずれかのロックサーバにより送信された獲得ロック要求を処理

10

【発明の概要】

【0008】

本発明の第1の態様によれば、ロック要求を処理するための方法が提供され、第1のロックサーバに適用可能であり、第1のロックサーバは、第2のロックサーバの引き継ぎロックサーバであり、第1のロックサーバは、第2のロックサーバのロック管理範囲を記憶し、この方法は、第1のロックサーバにより、故障が第2のロックサーバで発生したことを習得した後にサイレント状態に入るステップであり、サイレント状態のサイレント範囲は、第2のロックサーバが権限を割り当てているリソースであるステップと、第1のロックサーバにより、第1の獲得ロック要求を受信するステップであり、第1の獲得ロック要求は、第1のリソースをロックすることを要求するために使用され、第1の獲得ロック要求は、第1のリソースの識別子を搬送するステップと、第1のロックサーバにより、第1のリソースが第2のロックサーバの管理範囲に属すると検出するステップと、第1のロックサーバにより、第1のリソース情報記録リストに問い合わせるステップであり、第1のリソース情報記録リストは、第2のロックサーバがロック権限を割り当てているリソースのIDを記録し、第1のリソース情報記録リストが第1のリソースの識別子を記録していない場合、第1のロックサーバにより、第1の獲得ロック要求に従って第1のリソースについてロック権限を割り当てるステップとを含む。

20

【0009】

この方法によって、故障が第2のロックサーバで発生した場合、第1のロックサーバは、部分的にのみサイレントである。サイレント中に、第1のロックサーバは、第2のロックサーバがロック権限を割り当てていない獲得ロック要求を処理してもよく、これにより、システム効率を改善する。

30

【0010】

さらに、この方法によって、第1のロックサーバがサイレントである場合、第1のロックサーバの元の管理範囲内のリソースは、サイレント範囲に含まれず、したがって、通常通り処理されてもよい。さらに、第1のロックサーバと第2のロックサーバと他のロックサーバとを含む分散ロック管理システムにおいて、第1のロックサーバがサイレントである場合、第1のロックサーバ及び第2のロックサーバ以外のロックサーバはサイレントでなくてもよく、通常通り動作し続ける。

40

【0011】

第1の態様の第1の可能な実現方式では、この方法は、第1のロックサーバにより、第2の獲得ロック要求を受信するステップであり、第2の獲得ロック要求は、第2のリソースをロックすることを要求するために使用され、第2の獲得ロック要求は、第2のリソースの識別子を搬送するステップと、第1のロックサーバにより、第2のリソースが第1のロックサーバの管理範囲に属すると検出するステップと、第1のロックサーバにより、第2の獲得ロック要求に従って第2のリソースについてロック権限を割り当てるステップとを更に含む。

【0012】

この方法によって、第1のロックサーバがサイレントである場合、第1のロックサーバ

50

の元の管理範囲内のリソースは、サイレント範囲に含まれず、したがって、通常通り処理されてもよい。

【0013】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第2の可能な実現方式では、第1のロックサーバにより、サイレント状態に入った後に、この方法は、第1のロックサーバにより、第3の獲得ロック要求を受信するステップであり、第3の獲得ロック要求は、第3のリソースをロックすることを要求するために使用され、第3の獲得ロック要求は、第3のリソースの識別子を搬送するステップと、第1のロックサーバにより、第3のリソースが第2のロックサーバの管理範囲に属すると検出するステップと、第1のロックサーバにより、第1のリソース情報記録リストに問い合わせ、第1のリソース情報記録リストが第3の獲得ロック要求で要求されたリソースのIDを記録している場合、第1のロックサーバにより、第3の獲得ロック要求に従って第3のリソースについてロック権限を割り当ててを拒否するステップとを更に含んでもよい。

10

【0014】

この方法によって、第1のロックサーバは、第2のロックサーバが権限を割り当てているリソースについて獲得ロック要求を処理することを拒否し、これにより、ロック獲得の衝突を回避する。

【0015】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第3の可能な実現方式では、この方法は、第1のロックサーバにより、第1のリソースの識別子を第2のリソース情報記録リストに記録するステップであり、第2のリソース情報記録リストは、第1のロックサーバがロック権限を割り当てているリソースのIDを記録するために使用され、第2のリソース情報記録リストは、第3のロックサーバに記憶されるステップを更に含んでもよい。

20

【0016】

この方法によって、第1のロックサーバのロック獲得状態が記録されてもよい。将来、故障が第1のロックサーバで発生した後に、第1のロックサーバに対応する引き継ぎロックサーバは、第1のロックサーバから引き継いでよい。引き継ぎ方法は前述のものと同様である。

【0017】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第4の可能な実現方式では、第1のロックサーバにより、第2のロックサーバのロック管理範囲を記憶するステップは、第1のロックサーバにより、第1の通知メッセージを受信するステップであり、第1の通知メッセージは、第2のロックサーバの識別情報を搬送するステップと、第1のロックサーバにより、第2のロックサーバの識別子及びロックサーバ引き継ぎ関係に従って、第1のロックサーバが第2のロックサーバの引き継ぎロックサーバであると決定するステップと、第1のロックサーバにより、第2のロックサーバのロック管理範囲を受信し、ロック管理範囲を記憶するステップとを含む。

30

【0018】

この方法によって、どのように第1のロックサーバが第2のロックサーバのロック管理範囲を取得するかについての解決策が提供される。

40

【0019】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第5の可能な実現方式では、この方法は、プロトコルサーバにより、ホストからパケットを受信し、パケットから第1の獲得ロック要求を解析するステップと、プロトコルサーバにより、第1の獲得ロック要求をロックプロキシに転送するステップと、ロックプロキシにより、第1の獲得ロック要求で搬送された第1のリソースの識別子に従って決定を実行し、第1のロックサーバが第1のリソースを管理すると決定した場合、第1の獲得ロック要求を第1のロックサーバに送信するステップとを更に含んでもよい。

【0020】

50

この方法によって、プロトコルサーバ及びロックプロキシが追加され、ロックサーバとプロトコルサーバとロックプロキシとにより一緒に実行されるロック管理技術が提供される。

【0021】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第6の可能な実現方式では、この方法は、第1のロックサーバにより、サイレント状態に入った後に、再請求ロック要求を受信するステップであり、再請求ロック要求は、第4のリソースの識別子と、第4のリソースについて第2のロックサーバにより割り当てられた権限とを搬送し、第4のリソースは、第2のロックサーバが権限を割り当てているリソースであるステップと、第2のロックサーバにより割り当てられている権限に従って第4のリソースについて同じ権限を再割り当てするステップとを更に含んでもよい。

10

【0022】

この方法によって、サイレント中に、故障が第2のロックサーバで発生する前に割り当てられているロック権限を回復するために、再請求ロック要求が処理される。

【0023】

前述のいずれかの態様又はいずれかの実現方式に基づいて、第1の態様の第7の可能な実現方式では、この方法は、第2のロックサーバが権限を割り当てている全てのリソースについて権限が再割り当てされた後に、第1のロックサーバにより、サイレント状態を出るステップ、又は予め設定された時間の後に、第1のロックサーバにより、サイレント状態を出るステップを更に含んでもよい。

20

【0024】

この方法によって、ロックサーバが長期間サイレント状態にあることを妨げるために、サイレント終了機構が提供される。

【0025】

任意選択で、第1の態様の第8の可能な実現方式では、第1のロックサーバにより、サイレント状態を出た後に、この方法は、第1のロックサーバにより、第1のロックサーバの管理範囲を更新するステップであり、第1のロックサーバの更新された管理範囲は、第1のロックサーバの管理範囲と、第2のロックサーバの管理範囲とを含むステップを更に含む。

【0026】

任意選択で、引き継ぎ関係は、管理ノードにより計算され、次に、各ロックサーバにブロードキャストされてもよく、或いは各ロックサーバ自体により更新されてもよい。

30

【0027】

この方法によって、第1のロックサーバが第2のロックサーバから引き継ぐ全体手順が完了する。

【0028】

第1の態様の第9の可能な実現方式では、第1のリソース情報記録リストは、第1のロックサーバに記憶されてもよく、或いは第1のリソース情報記録リストが第1のロックサーバにより取得されることが出来る限り、他のロックサーバ又は非ロックサーバに記録されてもよい。

40

【0029】

この方法によって、第1のリソース情報記録リストの記憶の柔軟性が改善され、これは、ユーザが実際のニーズに従って製品を設計するのに役立つ。

【0030】

第1の態様の第10の可能な実現方式では、各態様及び実現方式は、仮想マシン環境に実現されてもよく、すなわち、ロックサーバは仮想マシン内で実行する。したがって、ロックサーバは、3つの可能な実現方式、すなわち、ハードウェア、ハードウェア上で実行するソフトウェア、及び仮想マシン内で実行するソフトウェアを有してもよい。

【0031】

第1の態様の第11の可能な実現方式では、引き継ぎが開始した後に、例えば、サイレ

50

ント中に、第1のロックサーバは、問い合わせメッセージを故障のないノードのロックプロキシに更に送信し、問い合わせメッセージを受信した後に、故障のないノードのロックプロキシは、フィードバックメッセージを第1のロックサーバに送信し、フィードバックメッセージは、第2のロックサーバを使用することによりロックプロキシによる申請を通じて取得されたロック権限を搬送し、第1のロックサーバは、ロック権限を詳細リソース情報記録リストに記録する。

【0032】

本発明は、ロック要求管理装置及びサーバの実現方式を更に提供し、ロック要求管理装置及びサーバは、第1の態様及び可能な実現方式の機能を有する。

【0033】

対応して、本発明は、不揮発性コンピュータ読み取り可能記憶媒体及びコンピュータプログラムプロダクトを更に提供する。不揮発性コンピュータ読み取り可能記憶媒体及びコンピュータプログラムプロダクトに含まれるコンピュータ命令が本発明において提供される記憶デバイスのメモリにロードされ、中央処理装置(Central Processing Unit、CPU)がコンピュータ命令を実行した場合、記憶デバイスは、第1の態様及び可能な実現方式の可能な実現解決策を別々に実行することが可能になり、解決策は、実行のために装置又はサーバで実行し得る。

【図面の簡単な説明】

【0034】

【図1】本発明の実施例によるロック管理システムの使用環境のトポロジ図である。

【図2】本発明によるロックサーバ管理範囲及びロックサーバ引き継ぎ関係の実施例の概略図である。

【図3】本発明の実施例によるロック要求を処理するための方法のフローチャートである。

【図4】本発明によるロック要求管理装置の実施例の構成図である。

【図5】本発明によるサーバの実施例の構成図である。

【発明を実施するための形態】

【0035】

以下に、本発明の実施例における添付図面を参照して本発明の実施例における技術的解決策を明確に説明する。明らかに、説明する実施例は、本発明の全ての実施例ではなく、一部のみの実施例である。

【0036】

本発明の実施例では、ロックサーバの間の引き継ぎ関係を確立することが提案される。このように、故障がロックサーバで発生した場合、故障のあるロックサーバの引き継ぎロックサーバは、引き継ぎ関係に従って与えられてもよい。

【0037】

ロックサーバは、ロック要求を処理することができるサーバである。ロック要求(lock request)は、獲得ロック要求(acquire lock request)又は再請求ロック要求(reclaim lock request)でもよい。獲得ロック要求は、読み取りロック要求又は書き込みロック要求でもよい。獲得ロック要求は、リソースをロックすることを申請するために使用される。ロック獲得の後に、権限所有者は、ロック権限を取得する、すなわち、権限所有者のみが、リソースについて対応する操作権限を有する。例えば、読み取りロック要求は、リソースを読み取るための権限を取得することを申請するために使用され、書き込みロック要求は、データをリソースに書き込むための権限を申請するために使用される。再請求ロック要求は、既に取得されたロック権限を再申請するために権限所有者により使用される。例えば、ホストは、ノード1を使用することにより記憶システムに元々アクセスしている。次に、故障がノード1で発生し、ホストがその代わりにノード2を使用することにより記憶システムにアクセスする。ホストは、再請求ロック要求をノード2に送信し、これにより、前に取得されているロック権限を取得する。

【0038】

ロック要求は、解放ロック要求 (release lock request) を更に含んでもよく、解放ロック要求は、ファイルについてロック権限を解放するために使用され、これにより、他のホストは、ファイルについてロック権限を申請してもよい。

【0039】

故障がロックサーバで発生した場合、これは、故障のあるロックサーバと呼ばれ、故障のあるロックサーバのロック管理作業は、故障のあるロックサーバの引き継ぎロックサーバにより引き継がれる。引き継ぎロックサーバのみがサイレント状態に入り、残りのロックサーバは、サイレント状態に入らず、通常通りロック要求を処理してもよい。従来技術に比べて、これは、ロックサーバの故障により全体のシステムにもたらされる影響を低減する。

10

【0040】

さらに、引き継ぎロックサーバがサイレント状態に入ったとしても、引き継ぎロックサーバは、いくつかのリソースのみについてサイレント状態に入る。したがって、引き継ぎロックサーバは、依然としていくつかの獲得ロック要求 (サイレント状態にないリソースについての獲得ロック要求) に通常通り応答することができる。このように、ロックサーバの利用が更に改善され、ロックサーバがサイレント状態に入った後にロックサーバによりシステムにもたらされる影響が低減される。サイレント状態のリソースについて、ロックサーバは、獲得ロック要求を処理しない。サイレント状態のリソースについて、ロックサーバは、再請求ロック要求を処理することができる。非サイレント状態のリソースについてのロック要求について、ロックサーバは、獲得ロック要求を処理することができる。例えば、ロック権限は、リソースについての読み取りロック要求に応じて与えられ、リソースについての書き込みロック要求に応じて、ロック権限は、既に割り当てられた書き込みロックを再利用することにより与えられる。いくつかのリソースについてサイレント状態に入るとは、いくつかのリソースについての獲得ロック要求についてサイレント状態に入ることと考えられてもよい。いくつかのリソースは、故障のあるロックサーバが権限を割り当てており、故障のあるロックサーバの引き継ぎロックサーバがサイレント状態に入るリソースである。

20

【0041】

引き継ぎロックサーバにより元々管理されている獲得ロック要求は、通常状態で維持し、サイレント状態により影響を受けない。受信した獲得ロック要求が故障のあるロックサーバの管理範囲内にあり、この前に、故障のあるロックサーバがロック要求で要求されたリソースについてロック権限を割り当てていない場合、引き継ぎロックサーバは、通常通りロック要求に応答し、ロック要求についてロック権限を割り当ててもよい。受信した獲得ロック要求が故障のあるロックサーバの管理範囲内にあり、この前に、故障のあるロックサーバがロック要求で要求されたリソースについてロック要求を割り当てている場合、引き継ぎロックサーバは、ロック要求についてロック権限を割り当てることを拒否する。

30

【0042】

引き継ぎロックサーバにより元々管理されている獲得ロック要求は、サイレント状態により影響を受けず、引き継ぎロックサーバは、獲得ロック要求についてサイレント状態に入らない。獲得ロック要求についての処理方式は、引き継ぎロックサーバがサイレント状態に入る前に使用されるものと同じであり、したがって、詳細に説明される必要はないかもしれない。以下に特に断りのない限り、「獲得ロック要求」は、故障のあるロックサーバの管理範囲内に元々あり、故障が故障のあるロックサーバで発生した後に引き継ぎロックサーバにより引き継がれた獲得ロック要求を示す。

40

【0043】

本発明のこの実施例は、分散システムに適用されてもよい。分散システムは、複数のノードを含み、各ノードは、いくつかのファイルについてロック権限を管理する。例えば、ノードは、ロックサーバであり、プロセッサと、外部インタフェースと、メモリとを含んでもよい。故障が分散システム内のロックサーバで発生した場合、分散システム内の故障のないロックサーバは、サイレント状態に入り、ロック権限を管理する方法が提供される

50

。ノードは、プロトコルサーバ及びロックプロキシを更に統合してもよく、ロックサーバとプロトコルサーバとロックプロキシとの組み合わせになる。

【0044】

ロック権限を割り当てた後に、ロックサーバは、ロック権限が割り当てられているリソースの識別子を指定のロックサーバにバックアップする。指定のロックサーバは、ロックサーバの引き継ぎロックサーバでもよく、或いはロックサーバのバックアップサーバによりアクセスされることができ他のロックサーバでもよい。獲得ロック要求を受信した後に、故障のあるロックサーバの引き継ぎロックサーバは、リソースのバックアップされた識別子に従って、獲得ロック要求で要求されたロック権限が割り当てられているか否かを決定し、ロック権限が割り当てられている場合、拒否応答メッセージを返信し、或いはロック権限が割り当てられていない場合、獲得ロック要求で要求されたロック権限をホストに割り当てる。ノードは、ロックサーバのみを含んでもよく、或いはプロトコルサーバ及びロックプロキシのような他の機能モジュールを統合してもよい。

10

【0045】

権限を割り当てた後に、ロックサーバは、割り当て記録を生成する。割り当て記録情報は、例えば、ノード1がファイルAについて書き込み権限を割り当てたことを示す{ノード1, ファイルA, 書き込み権限}、又はノード2がファイルBについて読み取り権限を有することを示す{ノード2, ファイルB, 読み取り権限}である。獲得ロック要求を送信したホストに従って、プロトコルサーバは、ノードの割り当て記録をホストの割り当て記録に変更してもよい。例えば、ホスト1が獲得ロック要求を送信した場合、ノード1は、ホスト1に変更され、割り当て記録情報は、ホスト1がファイルAについて書き込み権限を有することを示す{ホスト1, ファイルA, 書き込み権限}になる。ノードは、記憶のために情報を対応するホストに送信してもよい。

20

【0046】

ロック権限が割り当てられたリソースの識別子は、指定のサーバ、例えば、権限を割り当てたロックサーバのバックアップロックサーバ、又は他のロックサーバにバックアップされる。ロック権限の具体的な内容は、バックアップされなくてもよい。すなわち、指定のサーバは、ロック権限が割り当てられたリソースを認識するが、ロック権限が何であるかを認識しない。ロックサーバは、ロック権限が割り当てられたリソースの識別子のみをバックアップしてもよく、ロック権限の具体的な内容をバックアップしなくてもよい。バックアップは、ほとんどシステムリソースを占有せず、分散システムのリソースに大きい影響をもたらさない。

30

【0047】

本発明の実施例において言及される分散システムは、複数のノードを含み、ホストは、NASネットワークを使用することによりノードと通信し、ノードは、記憶システムに接続される。記憶システムに記憶されるリソースは、使用のためにホストに提供される。ホストは、ノードを使用することによりリソースについてロック権限を申請し、ノード内のロックサーバは、ロック権限を管理する。

【0048】

ノード及び記憶デバイスは、分離されてもよく、或いは一緒に結合されてもよい。各ノードは、プロトコルサーバ及びロックプロキシを有する。ホストからのロック要求は、ネットワークファイルシステム(Network File System、NFS)プロトコルに基づいてもよく、或いはサーバメッセージブロック(Server Message Block、SMB)プロトコルに基づいてもよい。プロトコルサーバは、ホストからの1つ以上のプロトコルを処理してもよい。例えば、NFSサーバは、NFSプロトコルをサポートし、SMBサーバは、SMBプロトコルをサポートする。異なるプロトコルサーバとより上位レイヤのホストとの間の通信の動作原理は同様である。プロトコルサーバにより処理されたロック要求は、ロックプロキシにより使用されてもよい。

40

【0049】

図1に示すように、分散システム内の2つのホスト及び4つのノードが説明のための例

50

として使用される。ホスト及びノードの数量は、要件に従って調整されてもよく、その実現原理は同じである。ホストは、ノードを使用することにより記憶システムにアクセスする。ノードにおいて、プロトコルサーバは、ロックプロキシと1対1の対応にある。例えば、ノード1において、プロトコルサーバ1はロックプロキシ1と1対1の対応にあり、ノード2において、プロトコルサーバ2はロックプロキシ2と1対1の対応にあり、以下同様である。信号は、対応に従ってプロトコルサーバとロックプロキシとの間で伝達される。

【0050】

ロックサーバは、プロトコルサーバ及びロックプロキシと共に同じノードに位置してもよく、或いは独立したノードに別々に位置してもよく、或いは他のノードに位置してもよい。ノード内の通信は、コンピュータ内部プロトコル、例えば、バスを使用することにより実行される。ノードの間の通信は、FC又はイーサネットのようなネットワークを使用することにより実行されてもよい。本発明のこの実施例では、サーバ、プロトコルサーバ及びロックプロキシが1つのノードに位置する例が説明のために使用される。例えば、プロトコルサーバ1、ロックプロキシ1及びロックサーバ1は、ノード1に位置する。各ロックサーバは、異なるノード内のロックプロキシに異なるロック権限を許可してもよい。ノード内のロックプロキシは、権限をこのノード内のロックサーバに申請してもよく、或いはロック権限を他のノード内のロックサーバに申請してもよい。

【0051】

管理ノードは、ノードを制御及び管理するために分散システムに別々に配置されてもよく、或いはいずれかのノードが全てのノードを更に制御及び管理してもよい。ノードを管理及び制御するノードは、通常ではプライマリノードであり、管理ノードとも呼ばれてもよい。これは、本発明のこの実施例では限定されず、図面にも別々に示されない。

【0052】

読み取り/書き込み操作が記憶システム内のリソース(ファイル、ディレクトリ、ファイルブロック又はデータブロック等)上で実行される必要がある場合、ホストは、ネットワークを使用することによりロック要求を対応するプロトコルサーバに送信する。ホストは、ロック要求で搬送された情報に従って対応するプロトコルサーバを決定してもよく、或いはIPアドレスセグメントに従って対応するプロトコルサーバを決定してもよい。既存の実現方式は、双方の場合において使用されてもよく、これは、本発明のこの実施例では限定されない。ロック要求を受信した後に、プロトコルサーバは、ロック要求をプロトコルサーバに対応するロックプロキシに送信する。

【0053】

ロックプロキシは、ロックサーバ管理範囲に従って、どのロックサーバがロック要求を処理するかを決定し、次に、処理のためにロック要求を決定されたロックサーバに送信する。ロックサーバ管理範囲は予め設定されてもよく、或いはコンシステントハッシュリングを使用することにより決定されてもよい。ロックサーバ管理範囲は、ロックプロキシが位置するノードのキャッシュに記憶されてもよく、或いは共有ストレージに記憶されてもよく、分散システム内のロックプロキシにより共有される。

【0054】

例えば、獲得ロック要求を受信した後に、ロックプロキシ2は、ローカルに記憶されたロックサーバ管理範囲に従って、獲得ロック要求がロックサーバ3により処理されるべきであると決定し、処理のために獲得ロック要求をロックサーバ3に送信する。或いは、ロックサーバ管理範囲はローカルに記憶されなくてもよいが、その代わりに、ロック要求は、ファイルのIDを搬送し、ロックプロキシは、問い合わせ又は計算によってどのロックサーバがファイルについてのロック権限を管理するかを習得してもよい。或いは、ロックプロキシは、ロックプロキシと共に同じノードに位置するロックサーバにロック要求を直接送信してもよく、次に、同じノードに位置するロックサーバは、ロックサーバ管理範囲に従って、ロック要求を処理する役目をするロックサーバにロック要求を転送する。例えば、ロックプロキシ2は、受信した獲得ロック要求をロックサーバ2に送信し、ロックサーバ2は、ローカルに記憶されたロックサーバ管理範囲に従って、ロックサーバ4が獲得ロック

10

20

30

40

50

要求を処理する役目をすべきと決定し、ロックサーバ2は、処理のために獲得ロック要求をロックサーバ4に転送する。2つの処理方式は、既存の技術を使用することにより実現されてもよく、ここでは更に説明しない。

【0055】

ロックサーバは、ロックサーバにより割り当てられたロック権限を記憶する。ロックプロキシは、ロックプロキシがロックサーバに申請するロック権限を記憶する。分散システム内のロックサーバ管理範囲が変更された場合、分散システム内の管理ノードは、ロックサーバ及びロックプロキシに対して、対応するロックサーバ管理範囲を更新するように命令する。或いは、管理ノードは、ロックサーバ管理範囲を更新し、次に、更新結果を分散システム内のロックプロキシ及びロックサーバにブロードキャストしてもよい。

10

【0056】

ロックサーバがロック要求を受信した後に、ロックサーバが通常動作状態にある（すなわち、サイレント状態にない）場合、ロックサーバによりロック要求を処理する方式は、従来技術のものと同じであり、例えば、ロック要求に従ってロック権限をホストに割り当てる。これは、ここでは更に説明しない。

【0057】

本発明のこの実施例における分散システムは、仮想化された分散システムでもよく、ロックサーバは、仮想マシン内で実行する。ロックプロキシ及びプロトコルサーバも、仮想マシン内で実行してもよい。これらの機能は、仮想化されていない環境のものと同様であるため、説明は別々に提供されない。

20

【0058】

分散システム内のロックサーバ管理範囲及びロックサーバ引き継ぎ関係について、図2を参照する。

【0059】

図2に示すように、ロックサーバは、論理的にリングを形成する。分散システム内のロックサーバ管理範囲は、コンシステントハッシュリングの反時計方向に決定される（他の実現方式では、時計方向が使用されてもよい）。コンシステントハッシュリングは、分散システム内のロックサーバのIDに従った計算によって取得される。例えば、分散システムにおいて、ロックサーバ1のIDは1であり、ロックサーバ2のIDは2であり、ロックサーバ3のIDは3であり、ロックサーバ4のIDは4である。コンシステントハッシュリングを形成するために、ロックサーバは、コンシステントハッシュアルゴリズムを使用することによりIDにおけるハッシュ計算をそれぞれ実行し、計算結果は、昇順に時計方向に配置される。このように、ロックサーバにより取得されたコンシステントハッシュリングは同じである。図2に示すように、コンシステントハッシュリングは0-232であり、ロックサーバのIDにおけるハッシュ計算を実行することにより取得された結果は、順にhash(1)=5000、hash(2)=8000、hash(3)=1024及びhash(4)=512である。時計方向では、0から始めて、ハッシュリングにおけるロックサーバの位置の順序は、ロックサーバ4、ロックサーバ3、ロックサーバ1及びロックサーバ2である。この場合、ロックサーバ4の管理範囲は(8000,232]及び[0,512]であり、ロックサーバ3の管理範囲は(512,1024]であり、ロックサーバ1の管理範囲は(1024,5000]であり、ロックサーバ2の管理範囲は(5000,8000]である。この場合、図2においてロックサーバにより論理的に形成されたリングにおいて、ロックサーバの間の引き継ぎ関係は、コンシステントハッシュリングの時計方向に決定され、すなわち、ロックサーバ1の引き継ぎロックサーバはロックサーバ2であり、ロックサーバ2の引き継ぎロックサーバはロックサーバ4であり、ロックサーバ4の引き継ぎロックサーバは3であり、ロックサーバ3の引き継ぎロックサーバは1である。

30

40

【0060】

明らかに、本発明のこの実施例は、ロックサーバの引き継ぎサーバを決定するための方法を提供する。各ロックサーバが引き継ぎサーバを有することができる限り、引き継ぎ関係は一意ではない点に留意すべきである。例えば、マネージャは、各ロックサーバのための引き継ぎサーバを構成してもよい。例えば、ロックサーバ2は、ロックサーバ1の引き継

50

ぎロックサーバとして構成され、ロックサーバ3は、ロックサーバ2の引き継ぎロックサーバとして構成され、ロックサーバ4は、ロックサーバ3の引き継ぎロックサーバとして構成され、ロックサーバ1は、ロックサーバ4の引き継ぎロックサーバとして構成される。

【0061】

1つのロックサーバは、複数のロックサーバから引き継いでもよい。例えば、故障がロックサーバ3及びロックサーバ1の双方で発生した場合、ロックサーバ3及びロックサーバ1の引き継ぎロックサーバは、ロックサーバ4である。

【0062】

本発明のこの実施例では、ロック要求（例えば、再請求ロック要求又は獲得ロック要求）を受信した後に、ロックプロキシは、記憶されたロックサーバ管理範囲に従って、ロック要求を処理すべきロックサーバを決定する。故障がロック要求を処理すべきロックサーバで発生したと決定した場合（故障がロックサーバで発生した場合、管理ノードは、通知メッセージを分散システム内のロックプロキシにブロードキャストする）、ロックプロキシは、ロックサーバ引き継ぎ関係に従って引き継ぎロックサーバを決定し、処理のためにロック要求を引き継ぎロックサーバに送信する。

10

【0063】

ロックサーバ管理範囲及びロックサーバ引き継ぎ関係は、双方とも管理ノードにより構成され、記憶のために全てのロックプロキシに送信されてもよく、或いは管理ノードが計算によってコンシステントハッシュリングを取得した後にロックプロキシに送信されてもよい。或いは、管理ノードは、ロックプロキシを事前に構成してもよく、ロックプロキシは、同じコンシステントハッシュリングを取得するために計算を別々に実行する。

20

【0064】

ロック要求を受信した後に、ロックプロキシは、計算結果が入る範囲を決定するために、コンシステントハッシュアルゴリズムを使用することにより、ロック要求で搬送されたファイル識別子でハッシュ計算を実行し、次に、対応するロックサーバは、ロック要求を処理する役目をする。例えば、ロック要求は、獲得ロック要求であり、獲得ロック要求で搬送されたファイル識別子（例えば、ファイル名）は、(foo1.txt)である。ロックプロキシは、(foo1.txt)でハッシュ計算を実行し、取得された結果は4500であり、これにより、獲得ロック要求は、ロックサーバ1により管理されるべきであり、ロックプロキシは、獲得ロック要求をロックサーバ1に送信する。他の例では、ロック要求は、再請求ロック要求であり、再請求ロック要求で搬送されたファイル情報は、(foo8.txt)である。ロックプロキシは、(foo8.txt)でハッシュ計算を実行し、取得された結果は9000であり、これにより、再請求ロック要求は、ロックサーバ4により管理されるべきであり、ロックプロキシは、再請求ロック要求をロックサーバ4に送信する。

30

【0065】

引き継ぎロックサーバがサイレントである場合、サイレント範囲内のリソースについて、ホストは、再請求ロック要求を使用することにより、ホストが故障のあるロックサーバに以前に申請している権限を引き継ぎロックサーバから再取得してもよい。引き継ぎロックサーバがサイレント状態を出る前に再請求ロック要求の実行が完了していない場合、実行されるべき再請求ロック要求は、もはや実行されない。再請求ロック要求の部分の詳細については、ステップ309を参照する。

40

【0066】

故障がロックサーバで発生した後に、ロックプロキシは、コンシステントハッシュリング内の故障のあるロックサーバを故障として識別する。ロック要求を受信した後に、ロックプロキシは、ロック要求で搬送されたファイル識別子でハッシュ計算を実行し、ロックサーバ管理範囲に従って、管理範囲が計算結果を含むロックサーバを決定する。決定されたロックサーバが故障状態にある場合、ロックプロキシは、ロックサーバ引き継ぎ関係に従って故障のあるロックサーバの引き継ぎロックサーバを更に決定し、処理のためにロック要求を引き継ぎロックサーバに送信する。ロック要求を受信した後に、引き継ぎロックサーバは、ハッシュ値を取得するために、ファイル識別子に従ってハッシュ計算を実行す

50

る。引き継ぎロックサーバは、ハッシュ値が自分の管理範囲内に入ると見出し、したがって、ロック要求自体を処理する必要がある。他の非ロックサーバがロック要求を受信した場合、ハッシュ値を取得するために、ファイル識別子に従ってハッシュ計算を実行し、ハッシュ値が自分の引き継ぎ範囲内に入らないと見出し、非ロックサーバは処理を実行しない。

【 0 0 6 7 】

例えば、故障が分散システム内のロックサーバ2で発生し、通知メッセージを受信した後に、ロックプロキシは、コンシステントハッシュリング内のロックサーバ2を故障として識別する。ロックプロキシにより受信された再請求ロック要求で搬送されたファイル情報は、(foo5.txt)であり、ロックプロキシは、(foo5.txt)でハッシュ計算を実行する。取得された結果は7000であり、ロック管理範囲に従って、ロックサーバ2が再請求ロック要求を処理する役目をするべきであるが、ロックサーバ2は、現在故障状態にある。ロックサーバ引き継ぎ関係に従って、故障のあるロックサーバ2の引き継ぎロックサーバは、ロックサーバ4である。したがって、ロックプロキシは、処理のために再請求ロック要求を引き継ぎロックサーバ4に送信する。ロックサーバ4は、(foo5.txt)でハッシュ計算を実行し、取得された結果は7000であり、ロックサーバ4の引き継ぎ範囲に属する。したがって、ロックサーバ4は、再請求ロック要求を処理する。

【 0 0 6 8 】

分散システムにおいてホストがファイル上で操作を実行する必要がある場合、ホストは、ホスト上のアプリケーションを使用することにより、ロック要求をプロトコルサーバに送信する。プロトコルサーバは、ロック要求に対応するロックプロキシに送信する。ロックプロキシは、ロック要求で搬送されたファイル識別子(識別子は、例えば、FSID又はFIDである)でハッシュ計算を実行し、計算結果に従って、管理範囲がファイルを含むロックサーバを決定し、対応する処理のためにロック要求をロックサーバに送信する。ファイル識別子でハッシュ計算を実行するために使用されるハッシュアルゴリズムは、コンシステントハッシュリングを生成するために使用されたハッシュアルゴリズムと同じである必要がある。例えば、ロック要求で搬送されたファイル識別子は(foo2.txt)であり、ロックプロキシは、ファイル識別子(foo2.txt)でハッシュ計算を実行する。取得された結果は6500であり、図2に示すように、結果はコンシステントハッシュリング内のロックサーバ1とロックサーバ2との間の範囲内に入り、その範囲は、ロックサーバ2の管理範囲であり、ロック要求は、ロックサーバ2により処理されることが認識され得る。

【 0 0 6 9 】

故障がロックサーバ2で発生した場合、ロックプロキシは、コンシステントハッシュリング内のロックサーバ2を故障として識別する。この場合、ロック要求を受信した後に、ロックプロキシは、ロック要求で搬送されたファイル情報(foo3.txt)でハッシュ計算を実行し、取得された結果は7500であり、コンシステントハッシュリング内のロックサーバ1とロックサーバ2との間の範囲内に入る。しかし、ロックサーバ2は故障状態にある。ロックサーバ2の引き継ぎロックサーバは、コンシステントハッシュリングに従ってロックサーバ4であり、すなわち、その範囲は、ロックサーバ4の管理範囲であり、したがって、ロックプロキシは、処理のためにロック要求をロックサーバ4に送信する。

【 0 0 7 0 】

ノードの名前又はロックサーバのIDに従ってコンシステントハッシュアルゴリズムを使用することにより、コンシステントハッシュリングを取得するための方法において既存の技術が使用されてもよい。詳細はここでは更に説明しない。

【 0 0 7 1 】

図2に示す分散システム内のロックサーバ管理範囲及びロックサーバ引き継ぎ関係に基づいて、本発明の実施例は、分散システム内でロック権限を処理するための方法を提供する。この方法の実施例は、ロックサーバに適用される。この方法の実施例で言及されるプロトコルサーバ及びロックプロキシの方法の実現手順は、前述の方法と同じであり、この方法の実施例では再び別々に説明しない。具体的な手順は図3に示されている。

10

20

30

40

50

【0072】

この方法は、図1に示す分散システムに適用されてもよい。本発明のこの実施例における分散システムにおいて、4つのロックサーバ、すなわち、ロックサーバ1、ロックサーバ2、ロックサーバ3及びロックサーバ4が存在する。この実施例におけるロックサーバの数量は単なる例であり、具体的な数量は、実際のサービス要件に従う。その実現原理は、この実施例のものと同じである。

【0073】

図3を参照して、引き継ぎロックサーバが第1のロックサーバであり、故障が発生したロックサーバが第2のロックサーバである例が、獲得ロック要求の処理を具体的に説明するために以下で使用される。

【0074】

ステップ301：故障が分散システム内のロックサーバで発生した場合、管理ノードは、通知メッセージを分散システム内のロックサーバにブロードキャストする。この実施例では、故障が発生したロックサーバは、第2のロックサーバである。したがって、通知メッセージは、第2のロックサーバのIDを第2のロックサーバの識別情報として搬送する。他の通知メッセージと異ならせるために、故障がロックサーバで発生したという通知メッセージは、第1の通知メッセージと呼ばれる。

【0075】

第1の通知メッセージを受信したロックサーバは、第1の通知メッセージで搬送されたID及びローカルに記憶されたロックサーバ管理範囲に従って、ロックサーバが第2のロックサーバの引き継ぎロックサーバであるか否かを決定する。ロックサーバが第2のロックサーバの引き継ぎロックサーバである場合、ロックサーバは、第2のロックサーバが既に権限を割り当てているリソースについてサイレント状態に入り、ロックサーバが第2のロックサーバの引き継ぎロックサーバでない場合、ロックサーバは、サイレント状態に入らない。

【0076】

さらに、第1の通知メッセージを受信した後に、第1のロックサーバは、タイマを有効にしてもよい。タイマが予め設定された時間の後に満了した場合、第1のロックサーバは、サイレント状態を出て、第1のロックサーバの引き継ぎ関係を更新する。

【0077】

他の故障検出方式は以下の通りである。引き継ぎロックサーバは、引き継ぎ関係についての情報に従って検出メッセージを対応するロックサーバに定期的送信し、故障が対応するロックサーバで発生したと検出した後に、引き継ぎロックサーバは、サイレント状態に入る。

【0078】

ステップ302：第1のロックサーバは、獲得ロック要求を受信し、獲得ロック要求は、ターゲットリソースの識別子を搬送する。ターゲットリソースは、ロックされる必要があるリソースであり、獲得ロック要求の要求対象であり、或いは換言するとロック権限が割り当てられることを待機するリソースである。

【0079】

記憶システム内のリソース上で読み取り又は書き込み操作のような操作を実行する場合、ホストは、プロトコルサーバ及びロックプロキシを使用することにより、獲得ロック要求をロックサーバに送信する必要がある。獲得ロック要求は、ホストがロックすることを要求するリソースのリソース識別子を搬送する。リソース識別子は、操作される必要があるファイルのID又は操作される必要がある論理ユニット番号(Logical Unit Number、LUN)のIDでもよく、獲得ロック要求は、リソースを読み取るため或いはリソースに書き込むための権限を要求する。

【0080】

第1のロックサーバは、リソース識別子に従って、要求されたリソースが第1のロックサーバの管理範囲に属するか否かを決定する。例えば、ハッシュ計算がリソース識別子で

10

20

30

40

50

実行される。取得された値が第1のロックサーバにより予め設定されたハッシュ値の範囲に属する場合、要求されたリソースは、第1のロックサーバの管理範囲に属し、或いは、取得された値が第1のロックサーバにより予め設定されたハッシュ値の範囲に属さない場合、要求されたリソースは、第1のロックサーバの管理範囲に属さない。

【0081】

獲得ロック要求がロックすることを要求するターゲットリソースは、記憶システムに位置し、ロックサーバは、ターゲットリソースについてロック権限を管理する。この処理は既に前述されており、ここでは再び別々に説明しない。

【0082】

処理のために、元々第2のロックサーバに送信されるロック要求を第1のロックサーバに送信するための多くの方法が存在する。例えば、ルータがロックサーバ及びホストのために直接配置され、ルータは、引き継ぎ関係を記録する。故障が第2のロックサーバで発生した場合、ルータは、元々第2のロックサーバに送信されるロック要求を第2のロックサーバの引き継ぎロックサーバに送信する。故障が第2のロックサーバと共に同じノード内のロックプロキシで発生していない場合、前述の解決策が使用されてもよく、ロックプロキシは、元々第2のロックサーバに送信されるロック要求を第2のロックサーバの引き継ぎロックサーバに送信する。

【0083】

図1に示す分散システムにおいて、故障がロックサーバ1で発生したこと、すなわち、ロックサーバ1が第2のロックサーバであることが仮定される。図2に示すロックサーバ管理範囲及びロックサーバ引き継ぎ関係に従って、ロックサーバ2は、ロックサーバ1の引き継ぎロックサーバであり、この場合、ここでの第1のロックサーバは、ロックサーバ2である。

【0084】

ステップ303：第1のロックサーバは、第1のリソース情報記録リストに問い合わせ、第1のリソース情報記録リストは、第2のロックサーバがロック権限を割り当てているリソースのリソース識別子を記録する。

【0085】

第2のロックサーバの引き継ぎロックサーバは、サイレント状態にある。獲得ロック要求を受信した後に、第1のロックサーバを含む各ロックサーバは、まず、ロックサーバが現在サイレント状態にあるか否かを決定する。ロックサーバがサイレント状態にある場合、各ロックサーバは、獲得ロック要求で搬送された識別子に従って決定を更に実行する。獲得ロック要求がロックサーバの引き継ぎ範囲に属する獲得ロック要求であることが習得された場合、ステップ303が実行される。獲得ロック要求を受信したのが第1のロックサーバでなく他のロックサーバである場合、ステップ303をスキップし、全体手順を出る。

【0086】

故障がノードで発生した後に、ノード内のプロトコルサーバは、故障のあるプロトコルサーバになる。故障のあるプロトコルサーバを使用することにより記憶システムに元々アクセスするホストは、その代わりに故障のあるプロトコルサーバの引き継ぎプロトコルサーバを使用することにより記憶システムにアクセスする必要がある。すなわち、引き継ぎプロトコルサーバは、故障のあるプロトコルサーバの作業を引き継ぐ。引き継ぎプロトコルサーバが引き継ぎを完了する前に、引き継ぎプロトコルサーバが位置するノードは、サイレント状態にある（サイレント範囲は、故障のあるロックサーバが権限を割り当てているリソースである）。引き継ぎが完了した後に、サイレントノードは、サイレント状態を出る。さらに、予め設定された時間を超えると、引き継ぎが完了していなくても、サイレントノードは、サイレント状態を出る。

【0087】

サイレントロックサーバが位置するノードもサイレント状態にあり、すなわち、ノードがロックサーバとプロトコルサーバとロックプロキシを含む場合、ノードのプロトコルサーバ及びロックプロキシもサイレント状態に入る。サイレント中に、引き継ぎプロトコ

10

20

30

40

50

ルサーバは、故障のあるプロトコルサーバの作業を引き継ぎ、引き継ぎ処理は以下を含む。ホストは、再請求ロック要求を使用することにより、故障のあるプロトコルサーバを使用することにより以前に処理されている権限を再申請する。

【0088】

従来技術の実現中に、サイレント状態のロックサーバは、いずれかのロック要求について拒否応答メッセージを直接返信する。本発明のこの実施例では、サイレント状態にある場合、第1のロックサーバは、第1のリソース情報記録リストに問い合わせる。第1のリソース情報記録リストは、第2のロックサーバがロック権限を割り当てているリソースのリソース識別子を記憶する。第1のリソース情報記録リストは、ロック権限の具体的な内容、例えば、ロック権限が具体的には読み取り権限であるか書き込み権限であるかを記憶しなくてもよい。したがって、占有される記憶空間はかなり低減される。第1のリソース情報記録リストは、第1のロックサーバにローカルに記憶されてもよく、或いは他のサーバに記憶されてもよく、或いは、第1のロックサーバにローカルに且つ他の指定のサーバに記憶されてもよい。サイレント状態に入った後に、第1のロックサーバは、リソース情報記録リストを他の特定のロックサーバに記憶する。例えば、リソース情報記録リストは、第1のロックサーバにローカルに記憶される。第1のロックサーバがサイレント状態に入った後に、第1のロックサーバは、記憶のためにリソース情報記録リストを第1のロックサーバの引き継ぎロックサーバに送信する。或いは、第1のロックサーバのリソース情報記録が変更された後に、リソース情報記録リストは、同期を保つために、記憶のために第1のロックサーバの引き継ぎロックサーバに遅れずに同期される。

10

20

【0089】

本発明のこの実施例では、ロックサーバが最初にリソースについての獲得ロック要求を受信した場合、ロックサーバは、ロック権限がリソースについて割り当てられているという情報をロックサーバの引き継ぎロックサーバに送信し、引き継ぎロックサーバは、情報を第1のリソース情報記録リストに記憶する。或いは、ロックサーバは、ロック権限がリソースについて割り当てられた後に情報を送信してもよい。その後ロックサーバが同じリソースについてのロック要求を受信した場合、要求された権限が同じであるか否かに拘わらず、情報は、もはやロックサーバの引き継ぎロックサーバに送信されない。

【0090】

具体的な実現方法は以下の通りである。獲得ロック要求を受信する毎に、ロックサーバは、通知メッセージ「ロック権限がリソースについて割り当てられている」がロックサーバの引き継ぎロックサーバに送信されているか否かを決定し、通知メッセージがロックサーバの引き継ぎロックサーバに送信されていない場合、ロックサーバは、通知メッセージを送信し、通知メッセージがロックサーバの引き継ぎロックサーバに送信されている場合、ロックサーバは、通知メッセージを送信しない。

30

【0091】

第2のロックサーバがロック権限を割り当てているリソースのリソース識別子に加えて、第1のリソース情報記録リストは、第1のロックサーバがサイレント状態を出た後にロック要求を処理する場合に第1のロックサーバにより問い合わせるために、第1のロックサーバがロック権限を割り当てているリソースの識別子を更に記憶してもよい。

40

【0092】

ステップ303の後にステップ304又はステップ305が実行される。

【0093】

ステップ304：第1のリソース情報記録リストにターゲットリソースの識別子が存在する場合、第1のロックサーバは、拒否応答メッセージを返信する。

【0094】

リソース情報記録リストがリソース識別子を記憶する場合、これは、第2のロックサーバがリソースについてロック権限を割り当てていることを示す。この場合、第1のロックサーバは、獲得ロック要求を処理せず、これにより、同じリソースについてロック権限の間の衝突を回避する。第1のロックサーバは、ロックプロキシ及びプロトコルサーバを使

50

用することにより拒否応答メッセージをホストに返信する。

【0095】

明らかに、他の実施例では、問い合わせによって第1のロックサーバが他の獲得ロック要求について権限を割り当てていることが見出された場合、第1のロックサーバもまた、拒否応答メッセージを返信する。これは、ここでは詳細に説明しない。

【0096】

ステップ305：リソース情報記録リストがリソースの識別子を有さない場合、第1のロックサーバは、獲得ロック要求で要求された権限に従ってリソースについてロック権限を割り当て、ロックプロキシ及びプロトコルサーバを使用することにより、割り当てられたロック権限をホストに返信する。

10

【0097】

第1のリソース情報記録リストがリソースの識別子を記憶しない場合、これは、ロック権限がリソース識別子に対応するリソースについて割り当てられていないことを示す。この場合、ホストは、リソース識別子に対応するリソース上で操作を実行しない。したがって、第1のロックサーバは、リソース識別子に対応するリソースについてロック権限を割り当ててもよい。第1のロックサーバは、対応するロックプロキシ及びプロトコルサーバを使用することにより、割り当てられたロック権限を、要求を送信したホストに返信し、要求を送信したホストがリソース上で操作を実行することを可能にする。

【0098】

このように、本発明のこの実施例において提供される分散システム内で獲得ロック要求を処理するための方法によって、故障が分散システム内のロックサーバで発生した場合、第2のロックサーバの引き継ぎロックサーバは、いくつかの獲得ロック要求を処理してもよく、ロック権限が要求されたリソースについて割り当てられている場合にのみ、引き継ぎロックサーバは、ロック権限が割り当てられているリソースについて獲得ロック要求を処理しない。したがって、この実施例では、故障が分散システム内のロックサーバで発生したときの影響範囲がより正確に制御されて絞られ、分散システムの性能及び信頼性が改善される。

20

【0099】

ロック権限が割り当てられた後に、第1のロックサーバは、割り当てられた権限をローカルの詳細リソース情報記録リストに記録してもよい。詳細リソース情報記録リストは、リソース識別子、ロック権限、ロック権限の種別及びロック権限の現在の状態のような権限の具体的な内容を記録する。詳細リソース情報記録リスト及び第1のリソース情報記録リストは、分離されてもよく、或いは一緒に統合されてもよい。

30

【0100】

第1のロックサーバが故障のあるサーバから引き継いだ後に、第1のロックサーバは、問い合わせメッセージを故障のないノードのロックプロキシに更に送信する。問い合わせメッセージを受信した後に、それぞれの故障のないノードのロックプロキシは、フィードバックメッセージを第1のロックサーバに送信し、フィードバックメッセージは、第2のロックサーバを使用することによりロックプロキシによる申請を通じて取得されたロック権限を搬送し、第1のロックサーバは、ロック権限をロックサーバの詳細リソース情報記録リストに記録する。したがって、詳細リソース情報記録リストに記録された情報は更新される。第1のロックサーバにより割り当てられた権限の具体的な内容に加えて、詳細リソース情報記録リストは、第2のロックサーバにより割り当てられた権限の具体的な内容を更に記録する。

40

【0101】

前述のように、ステップ304及び305で言及した獲得ロック要求は、双方とも、引き継ぎ範囲に従って元々第2のロックサーバにより処理されるべきである獲得ロック要求である。引き継ぎロックサーバがサイレント状態に入ったとしても、引き継ぎロックサーバは、元々引き継ぎロックサーバの処理範囲内にあるリソースについてサイレント状態に入らず、これらのリソースについての獲得ロック要求は、引き継ぎロックサーバがサイレント状

50

態に入らないのと同じ方法で処理されてもよい。

【0102】

ステップ305の後に、ステップ306がこの方法の実施例において更に実行されてもよい。

【0103】

ステップ306：第1のロックサーバは、ターゲットリソースの識別子を第2のリソース情報記録リストに記憶する。第2のリソース情報記録リストの形式は、第1のリソース情報記録リストの形式と同様であり、第2のリソース情報記録リストは、第1のロックサーバがロック権限を割り当てているリソースのリソース識別子を記録するために使用され、これにより、故障が第1のロックサーバで発生した後に、第1のロックサーバの引き継ぎロックサーバは、第1のロックサーバから引き継いでよい。具体的なステップは、ステップ302からステップ305と同様であり、ここでは詳細に説明しない。

10

【0104】

第1のロックサーバがサイレント状態にない場合、獲得ロック要求内のターゲットリソースについてロック権限を割り当てた後に、第1のロックサーバは、ターゲットリソースの識別子を第2のリソース情報記録リストに記録する。第2のリソース情報記録リストは、第1のロックサーバがロック権限を割り当てているリソースのリソース識別子を記憶する。

【0105】

ステップ306によって、故障が第1のロックサーバで発生した場合、第1のロックサーバの引き継ぎロックサーバは、非サイレント状態からサイレント状態に切り替えられ、サイレント範囲は、第1のロックサーバが権限を割り当てているリソースである。第1のロックサーバの引き継ぎロックサーバにより受信された獲得ロック要求について、第2のリソース情報記録リストがターゲットリソースの識別子を記録していない場合、第1のロックサーバの引き継ぎロックサーバは、ステップ305の操作に従ってターゲットリソースについてロック権限を割り当ててもよく、そうでない場合、ステップ304のように拒否応答メッセージを返信する。

20

【0106】

従来技術の実現中に、リソースについてロック権限を割り当てた後に、ロックサーバは、リソース識別子、ロック権限、ロック権限の種別及びロック権限の現在の状態のような必要な情報をローカルに記憶する。これは、ここで別々に説明しない。本発明のこの実施例では、リソースについてロック権限を割り当てた後に、ロックサーバは、ロック権限が割り当てられたリソースのリソース識別子を更に別々に記憶する。任意選択で、ロックサーバは、リソース識別子を別のリソース情報記録リストに記憶し、リソース情報記録リストをロックサーバの引き継ぎロックサーバに記憶する。

30

【0107】

任意選択で、ステップ304又はステップ306の後に、この方法の実施例は、以下のステップ307を更に含んでもよい。

【0108】

ステップ307：サイレント状態を出す。

【0109】

全てのロックプロキシが第2のロックサーバを使用することによる申請を通じて取得されたロック要求内容を第1のロックサーバに報告した後に、これは、引き継ぎ作業が完了していることを意味し、第1のロックサーバは、事前にサイレント状態を出てもよい。

40

【0110】

さらに、引き継ぎ時間が予め設定されてもよい。予め設定された時間の後に、引き継ぎ作業が完了しているか否かに拘わらず、第1のロックサーバは、サイレント状態を出す。

【0111】

第1のロックサーバがサイレント状態を出た後に、第1のロックサーバの管理範囲及びロックサーバ引き継ぎ関係は更新されてもよい。第1のロックサーバの管理範囲は拡張され、第1のロックサーバの更新された管理範囲は、第1のロックサーバの古い管理範囲と

50

第2のロックサーバの管理範囲との集合である。このステップが実行された後に、第1のロックサーバは、第2のロックサーバから完全に引き継ぐ。対応して、システム内の引き継ぎ範囲も変更される必要がある。例えば、第1のロックサーバの引き継ぎロックサーバ（これは第3のロックサーバと呼ばれる）の引き継ぎ範囲も、第1のロックサーバの管理範囲が変更されたため更新される必要がある。

【0112】

第1の通知メッセージを受信した後に、分散システム内のロックサーバは、タイマを有効にしてもよい。予め設定された時間の後に、サイレント状態のロックサーバは、サイレント状態を出る。前述のように、第1の通知メッセージは、ブロードキャスト方式で管理ノードにより送信され、故障がロックサーバで発生したことを分散システム内のロックサーバに通知するために使用される。第1の通知メッセージを受信した後に、分散システム内の第2ではないロックサーバは、ローカル或いは共有ストレージに記憶されたロックサーバ引き継ぎ関係に従って、第2ではないロックサーバが第2のロックサーバの引き継ぎロックサーバであるか否かを決定する。第2ではないロックサーバが第2のロックサーバの引き継ぎロックサーバである場合、第2ではないロックサーバは、サイレント状態に入り、タイマを有効にする。予め設定された時間の後に、第2ではないロックサーバは、サイレント状態を出て、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する。第2ではないロックサーバが第2のロックサーバの引き継ぎロックサーバでない場合、第2ではないロックサーバは、サイレント状態に入らず、通常通り動作し続ける。

【0113】

さらに、第1の通知メッセージを受信した後に、引き継ぎではないロックサーバは、ローカルに記憶されたロックサーバ管理範囲及びロックサーバ引き継ぎ関係において第2のロックサーバを故障状態として更に識別してもよく、引き継ぎでないロックサーバは、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する。

【0114】

ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新するために分散システム内のロックサーバにより使用されるアルゴリズムは同じである。前述のように、具体的な方式は、ロックサーバのIDでのハッシュ計算を実行することでもよく、詳細はここでは再び説明しない。

【0115】

引き継ぎ関係の更新をトリガーするための複数の方式が存在する。或いは、更新は、管理ノードによりトリガーされてもよい。すなわち、管理ノードの通知メッセージを受信した後に、ロックサーバは管理範囲及びロックサーバ引き継ぎ関係を更新する。このように、管理ノードは、タイマを有効にする必要があり、タイマが予め設定された時間の後に満了した場合、通知メッセージを分散システムにブロードキャストする。管理ノードの通知メッセージを受信した後に、分散システム内で通常通り動作することができる第2ではないロックサーバは、ローカルに記憶されたロックサーバ管理範囲及びロックサーバ引き継ぎ関係を別々に更新する。

【0116】

任意選択で、ステップ307の後に、以下のステップが含まれてもよい。

【0117】

ステップ308：サイレント状態を出た後に、第1のロックサーバは、第1のリソース情報記録リストを削除する。

【0118】

第1のリソース情報記録リストは、第1のロックサーバにローカルに記憶されてもよく、或いは他のサーバに記憶されてもよい。第1のリソース情報記録リストが他のサーバに記憶される場合、第1のロックサーバは、他のサーバに対して第1のリソース情報記録リストを削除するように命令してもよい。

【0119】

第1のリソース情報記録リストは、第2のロックサーバがロック権限を割り当てている

10

20

30

40

50

リソースのリソース識別子を記録し、リソース識別子の内容は、例えば、「リソースID：権限割り当て済み」である。第1のロックサーバが第2のロックサーバのロックサービスを引き継いだ後、すなわち、第1のロックサーバがサイレント状態に入った後に、獲得ロック要求を受信した場合、第1のロックサーバは、第1のリソース情報記録リストに従って、獲得ロック要求で搬送されたリソース識別子に対応するリソースについてロック権限を割り当てるか否かを決定する。サイレント状態を出た後に、第1のロックサーバは、通常の手順に従って獲得ロック要求を処理し、この場合、第1のリソース情報記録リストに記録された情報は、もはやロック要求を処理するための基礎として使用されない。したがって、この場合、第1のリソース情報記録リストは削除されてもよい。このように、記憶される情報は最小化されることができ、分散システムの占有されるシステムリソースはできる限り低減されることができ、分散システムの性能への影響は最小化されることができ

10

【0120】

前述のように、サイレント中に、第1のロックサーバは、サイレント範囲内のリソースについて再請求ロック要求を処理し続けてもよい。したがって、第1のロックサーバは、ステップ301とステップ307との間にステップ309を更に実行してもよい。

【0121】

ステップ309：第1のロックサーバは、再請求ロック要求を受信し、再請求ロック要求は、他のターゲットリソースの識別子及び他のターゲットリソースについて第2のロックサーバにより割り当てられたロック権限を搬送し、第2のロックサーバにより割り当てられたロック権限は、故障が第2のロックサーバで発生する前に他のターゲットリソースについて第2のロックサーバにより割り当てられている。次に、第1のロックサーバは、第2のロックサーバにより既に割り当てられたロック権限に従って他のターゲットリソースについてロック権限を再割り当てし、再割り当てされたロック権限は、故障が第2のロックサーバで発生する前に第2のロックサーバにより他のターゲットリソースに割り当てられたロック権限と同じである。明らかに、再割り当てされたロック権限の権限所有者も、前の権限所有者と同じである。再請求ロック要求は、ホストにより開始され、サイレント状態を出る前に、第1のロックサーバは、複数の再請求ロック要求を処理してもよい。サイレント状態を出た後に、第1のロックサーバは、もはや再請求ロック要求を処理しない

20

30

【0122】

例えば、第2のロックサーバがリソースについて書き込み権限を権限所有者に割り当てた後に、故障が第2のロックサーバで発生する。再請求ロック要求を受信した後に、第1のロックサーバは、リソースについて書き込み権限を権限所有者に再割り当てする。本発明のこの実施例では、リソースについてロック権限を最初に割り当てた後に、ロックサーバは、ロック権限が割り当てられたリソースのリソース識別子を引き継ぎロックサーバに記憶する。故障がロックサーバで発生した場合、サイレント状態の第2のロックサーバの引き継ぎロックサーバは、記憶されたリソース識別子に従って、受信した獲得ロック要求が処理され得るか否かを決定する。このように、故障がロックサーバで発生した場合に影響を受ける獲得ロック要求の範囲が最小化され得る。さらに、リソース識別子のみが記憶されるため、ほとんどシステムリソースが占有されず、システムの性能への影響も弱く、全体の分散システムの安定性及び信頼性が改善される。

40

【0123】

前述の方法において、バックアップを簡略化し、占有されるシステムリソースをできる限り低減するために、ロック権限が割り当てられたリソースのリソース識別子のみがバックアップされる。十分なシステムリソースが存在する場合、ロックサーバについての情報は完全にバックアップされてもよく、すなわち、ロックサーバの詳細リソース情報記録リストは、完全にバックアップされ、例えば、第1のロックサーバの引き継ぎロックサーバにバックアップされる。この場合、ここでの処理方法の原理は、前述の方法のものと同様であるが、より多くの情報がバックアップされ、より多くのシステムリソースを占有する

50

。しかし、完全なロック権限は、第1のロックサーバにバックアップされるため、第1のロックサーバが第2のロックサーバから引き継ぐ場合、全てのノードのロックプロキシが申請を通じて取得されたロック権限を引き継ぎロックサーバに再び報告することは必要とされない。すなわち、ステップ305において言及した、第1のロックサーバにより、問い合わせメッセージを故障のないノードのロックプロキシに更に送信するステップは省略されてもよい。したがって、サイレント時間が最小化され得る。

【0124】

前述の方法は、仮想化された分散システムに適用されてもよい。更に、仮想化された分散システムにおいて、ロックサーバは、仮想マシン内で実行する。

【0125】

仮想化されたシナリオでは、ロックサーバの引き継ぎロックサーバが物理ノードに配置される場合、故障が現在のロックサーバで発生したとき、データは同じ物理ノード内で速く送信されるため、引き継ぎ時間が短縮され得る。

【0126】

仮想化されたシナリオでは、第2のロックサーバが位置するノードが通常に戻った場合、新たなロックサーバがそのノードに配置されてもよい。他の第1のロックサーバは、そのノードに直接マイグレーションされてもよく、すなわち、他の第1のロックサーバのアドレスマッピング関係が変更される。或いは、新たなロックサーバがそのノードに生成されてもよく、他の第1のロックサーバ上のロックサービスは、新たに生成されたロックサーバにマイグレーションされる。仮想化されたシステムでは、操作を簡略化するために、他の第1のロックサーバは、そのノードに直接マイグレーションされる。

【0127】

特定の条件を満たす場合、分散システム内のロックサーバのロックサーバ管理範囲及びロックサーバ引き継ぎ関係は更新されてもよい。前述のように、故障が分散システム内のロックサーバで発生した場合、故障のないロックサーバは、予め設定されたルールに従ってロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する。さらに、管理ノードは、分散システム内の第2ではないロックサーバに対して、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新するように命令してもよい。或いは、管理ノードは、ロックサーバ引き継ぎ関係を更新し、次に、更新されたロックサーバ引き継ぎ関係を分散システム内のロックサーバにブロードキャストしてもよい。例えば、新たなロックサーバが参加した場合、管理ノードは、分散システム内のロックサーバに対して、ロックサーバ引き継ぎ関係を別々に更新するように命令する。

【0128】

引き継ぎ関係は、2つの可能な場合に基づいて更新される。一方は、ロックサーバが故障又は他の理由のため、もはや使用されないことであり、他方は、新たなロックサーバが参加することである。2つの場合について以下に別々に説明する。

【0129】

前述の方法の実施例を参照して、分散システム内の第2ではないロックサーバが管理ノードの第1の通知メッセージ（第1の通知メッセージは、故障が分散システム内のロックサーバで発生したことを分散システム内の第2ではないロックサーバに通知するために使用される）を受信した後に、分散システム内のそれぞれの故障のないロックサーバは、故障のないロックサーバのロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する。第2ではないロックサーバは、予め設定された方法（例えば、コンシステントハッシュアルゴリズム）に従って第2ではないロックサーバの管理範囲及びロックサーバ引き継ぎ関係を更新してもよい。或いは、管理ノードは、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新し、次に、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を分散システム内のロックサーバにブロードキャストする。第2ではないロックサーバのロックサーバ管理範囲及びロックサーバ引き継ぎ関係は、第2ではないロックサーバにローカルに記憶されてもよく、共有ストレージに記憶されてもよい。これは、この方法の実施例では限定されない。

10

20

30

40

50

【0130】

新たなロックサーバが分散システムに参加した場合、分散システム内のロックサーバも、ロックサーバのロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する必要がある。分散システム内のロックサーバは、第2の通知メッセージを受信し、第2の通知メッセージは、新たなロックサーバの識別子を搬送する。分散システム内のロックサーバは、ロックサーバのロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新する。

【0131】

新たなロックサーバが分散システムに参加した場合、管理ノードは、第2の通知メッセージを分散システム内のロックサーバに送信し、第2の通知メッセージは、新たなロックサーバのIDを搬送する。第2の通知メッセージを受信した後に、ロックサーバ（新たなロックサーバを含む）は、計算によって、予め設定されたルール（例えば、コンシステントハッシュアルゴリズム）に従って新たなロックサーバ管理範囲及び新たなロックサーバ引き継ぎ関係を取得する。同様に、或いは、管理ノードは、ロックサーバ管理範囲及びロックサーバ引き継ぎ関係を更新し、次に、更新されたロックサーバ管理範囲及び更新されたロックサーバ引き継ぎ関係を分散システム内のロックサーバに送信してもよい。

10

【0132】

ロックサーバ管理範囲及びロックサーバ引き継ぎ関係が更新された後に、各ロックサーバは、更新の後に取得されたロックサーバ管理範囲及びロックサーバ引き継ぎ関係に従ってロックサーバの新たなバックアップロックサーバを決定し、新たなバックアップロックサーバに記憶される必要があるリソース情報記録リスト又はリソース識別子を新たなバックアップロックサーバに送信する。

20

【0133】

本発明の実施例は、ロック要求を処理するためのロック要求管理装置4を更に提供する。ロック要求管理装置4は、例えば、ロックサーバであり、ロック要求管理装置4の構成は、図4に示されている。ロック要求管理装置4は、図3及び前述の方法の実施例に適用されてもよい。詳細な説明は、方法の実施例及び対応する図3において既に行われているため、ロック要求管理装置4のモジュールの機能のみについて、以下に簡単に説明する。全ての詳細な機能については、前述の方法の実施例を参照する。複数のロック要求管理装置4は、分散ロック管理システムを形成してもよい。

【0134】

ロックサーバ4は、受信モジュール41と、記憶モジュール42と、ロック要求処理モジュール43と、サイレントモジュール44とを含む。

30

【0135】

受信モジュール41は、第1の獲得ロック要求及び通知メッセージを受信するように構成され、第1の獲得ロック要求は、第1のリソースの識別子を搬送する。記憶モジュール42は、他のロック要求管理装置のロック管理範囲及び第1のリソース情報記録リストを記録するように構成され、第1のリソース情報記録リストは、他のロック要求管理装置がロック権限を割り当てているリソースのリソース識別子を記録する。ロック要求処理モジュール43は、記憶モジュール42により受信された獲得ロック要求を処理するように構成される。サイレントモジュール44は、通知メッセージに従って、故障が他のロック要求管理装置で発生したことが習得された後に、ロック要求管理装置4をサイレント状態に設定するように構成され、サイレント範囲は、他のロック要求管理装置が権限を割り当てているリソースであり、ロック要求管理装置4がサイレント状態に入った後に、ロック要求処理モジュール43は、獲得ロック要求がサイレント範囲に属する場合、第1のリソース情報記録リストに問い合わせ、第1のリソース情報記録リストが第1のリソースの識別子を記録していない場合、第1の獲得ロック要求に従って第1のリソースについてロック権限を割り当てるように具体的に構成される。

40

【0136】

任意選択で、ロック要求管理装置4がサイレント状態に入った後に、受信モジュール41は、第2の獲得ロック要求を受信するように更に構成され、第2の獲得ロック要求は、第

50

2のリソースをロックすることを要求するために使用され、第2の獲得ロック要求は、第2のリソースの識別子を搬送する。ロック要求処理モジュール43は、第2のリソースがロック要求管理装置4の管理範囲に属すると検出した後に、第2の獲得ロック要求に従って第2のリソースについてロック権限を割り当てるように更に構成される。

【0137】

任意選択で、ロック要求管理装置4がサイレント状態に入った後に、受信モジュール41は、第3の獲得ロック要求を受信するように更に構成され、第3の獲得ロック要求は、第3のリソースをロックすることを要求するために使用され、第3の獲得ロック要求は、第3のリソースの識別子を搬送する。ロック要求処理モジュール43は、第3のリソースが他のロック要求管理装置の管理範囲に属すると検出した後に、第1のリソース情報記録リストに問い合わせ、第1のリソース情報記録リストが第3の獲得ロック要求で要求されたリソースのリソース識別子を記録している場合、第3の獲得ロック要求に従って第3のリソースについてロック権限を割り当てることを拒否するように更に構成される。

10

【0138】

任意選択で、ロック要求管理装置4がサイレント状態に入った後に、ロック要求処理モジュール43は、再請求ロック要求を受信するように更に構成され、再請求ロック要求は、第4のリソースの識別子と、第4のリソースについて他のロック要求管理装置により割り当てられた権限とを搬送し、第4のリソースは、他のロック要求管理装置が権限を割り当てているリソースであり、他のロック要求管理装置により割り当てられている権限に従って第4のリソースについて同じ権限を再割り当てするように更に構成される。

20

【0139】

任意選択で、受信モジュール41は、第1の通知メッセージを受信するように更に構成され、第1の通知メッセージは、他のロック要求管理装置の識別情報を搬送する。受信モジュール41は、他のロック要求管理装置の識別子及びロック要求管理装置4の引き継ぎ関係に従って、ロック要求管理装置が他のロック要求管理装置の引き継ぎロック要求管理装置4であると決定された後に、他のロック要求管理装置のロック管理範囲を記憶モジュール42に送信するように更に構成される。前述の記憶モジュール42が他のロック要求管理装置のロック管理範囲を記憶するように構成されることは、記憶モジュール42が受信モジュール41から他のロック要求管理装置のロック管理範囲を受信し、ロック管理範囲を記憶するように構成されることを具体的に含む。

30

【0140】

任意選択で、ロック要求管理装置4は、プロトコルサーバモジュール45及びロックプロキシモジュール46を更に含んでもよい。プロトコルサーバモジュール45は、ホストからパケットを受信し、パケットから第1の獲得ロック要求を解析し、第1の獲得ロック要求をロックプロキシモジュール46に転送するように構成される。ロックプロキシモジュール46は、第1の獲得ロック要求で搬送された第1のリソースの識別子に従って決定を実行し、ロック要求処理モジュール43が第1のリソースを管理すると決定した場合、受信モジュール41を使用することにより、第1の獲得ロック要求をロック要求処理モジュール43に送信するように構成される。

【0141】

任意選択で、サイレントモジュール44は、他のロック要求管理装置が権限を割り当てている全てのリソースについて権限が再割り当てされた後に、ロック要求管理装置がサイレント状態を出ることを可能にするように、或いは予め設定された時間の後に、ロック要求管理装置4がサイレント状態を出ることを可能にするように更に構成される。

40

【0142】

任意選択で、記憶モジュール42は、ロック要求管理装置4がサイレント状態を出た後に、ロック要求管理装置の管理範囲を更新するように更に構成され、ロック要求管理装置の更新された管理範囲は、ロック要求管理装置の管理範囲と、他のロック要求管理装置の管理範囲とを含む。

【0143】

50

本発明の実施例は、分散システム内の他のサーバ5を更に提供し、サーバ5の構成は、図5に示されている。サーバ5は、インタフェース51と、メモリ52と、プロセッサ53とを含む。サーバ5は、方法の実施例における方法を実行してもよい。具体的には、サーバ5のプロセッサ53は、方法のステップを実行する。インタフェース51は、外部データインタフェースを提供し、メモリ52は、データ記憶空間を提供する。サーバ5について以下に簡単にのみ説明する。具体的な内容については、前述の説明を参照する。

【0144】

インタフェース51は、例えば、獲得ロック要求又は再請求ロック要求を受信するための外部インタフェースを提供する。

【0145】

メモリ52は、第1のリソース情報記録リストを記憶するように構成され、第1のリソース情報記録リストは、他のサーバがロック権限を割り当てているリソースのリソース識別子を記録する。方法の実施例から、メモリ52は、他の情報、例えば、第2のリソース情報記録リスト又は詳細リソース情報記録リストを記憶するように更に構成されてもよく、メモリ52は、他のサーバのロック管理範囲を記憶するように更に構成されてもよいことが習得され得る。

【0146】

プロセッサ53は、方法の実施例におけるステップ、例えば、故障が他のサーバ5で発生したことを習得した後に、サーバをサイレント状態に設定するステップであり、サイレント状態のサイレント範囲は、他のサーバが権限を割り当てているリソースであるステップと、第1の獲得ロック要求を受信するステップであり、第1の獲得ロック要求は、第1のリソースをロックするために使用され、第1の獲得ロック要求は、第1のリソースの識別子を搬送するステップと、第1のリソースが他のロックサーバの管理範囲に属すると検出するステップと、第1のロックサーバにより、第1のリソース情報記録リストに問い合わせ、第1のリソース情報記録リストが第1のリソースの識別子を記録していない場合、第1のロックサーバにより、第1の獲得ロック要求に従って第1のリソースについてロック権限を割り当てるステップとを実行するためのプログラムを実行するように構成される。

【0147】

方法の実施例における動作、例えば、サイレント状態に入る動作、サイレント状態を出る動作、問い合わせる動作、決定する動作及び権限を割り当てる動作は、全てプロセッサ53により実行されてもよい。

【0148】

任意選択で、サーバ5は、プロトコルサーバモジュール54及びロックプロキシモジュール55を更に含んでもよい。

【0149】

プロトコルサーバモジュール54は、ホストからパケットを受信し、パケットから第1の獲得ロック要求を解析するように構成され、第1の獲得ロック要求をロックプロキシモジュール55に転送するように更に構成される。ロックプロキシモジュール55は、第1の獲得ロック要求で搬送された第1のリソースの識別子に従って決定を実行し、サーバ5が第1のリソースを管理すると決定した場合、第1の獲得ロック要求をインタフェース51に転送するように構成される。

【0150】

本発明の各実施例で言及される「リスト」、例えば、第1のリソース情報記録リストの形式は、シート又はリストに限定されず、「リスト」に記憶される内容は、「リスト」の定義として使用される。

【0151】

本発明の各態様又は各態様の可能な実現方式は、システム、方法又はコンピュータプログラムプロダクトとして具体的に実現されてもよい。したがって、本発明の各態様又は各態様の可能な実現方式は、完全なハードウェアの実施例、完全なソフトウェアの実施例（ファームウェア、常駐ソフトウェア等を含む）又はソフトウェアとハードウェアとを組み

10

20

30

40

50

合わせた実施例の形式を使用してもよく、これらは、ここでは併せて「回路」、「モジュール」又は「システム」と呼ばれる。さらに、本発明の各態様又は各態様の可能な実現方式は、コンピュータプログラムプロダクトの形式を使用してもよく、コンピュータプログラムプロダクトは、コンピュータ読み取り可能媒体に記憶されたコンピュータ読み取り可能プログラムコードを示す。

【0152】

コンピュータ読み取り可能媒体は、コンピュータ読み取り可能信号媒体又はコンピュータ読み取り可能記憶媒体でもよい。コンピュータ読み取り可能記憶媒体は、ランダムアクセスメモリ(RAM)、読み取り専用メモリ(ROM)、消去可能プログラム可能読み取り専用メモリ(EPROM又はフラッシュメモリ)、ファイバ、及びコンパクトディスク読み取り専用メモリ(CD-ROM)のような、電子、磁気、光、電磁気、赤外線若しくは半導体のシステム、デバイス若しくは装置、又は前述の項目のいずれか適切な組み合わせを含むが、これらに限定されない。

10

【0153】

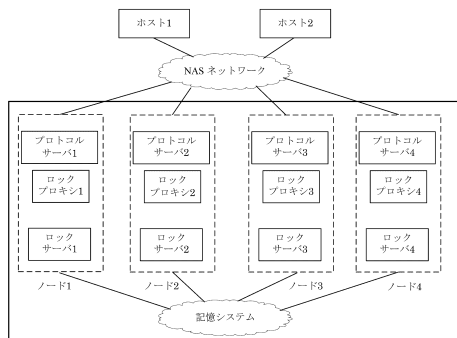
コンピュータ内のプロセッサは、コンピュータ読み取り可能媒体に記憶されたコンピュータ読み取り可能プログラムコードを読み取り、これにより、プロセッサは、フローチャート内の各ステップ又はステップの組み合わせで指定された機能及び動作を実行することができ、ブロック図内の各ブロック又はブロックの組み合わせで指定された機能及び動作を実現する装置が生成される。

【0154】

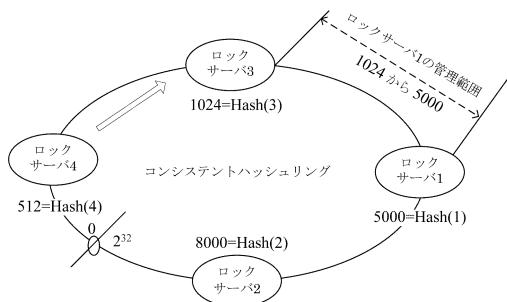
明らかに、当業者は、本発明の精神及び範囲を逸脱することなく、本発明に対して様々な修正及び変更を行い得る。このように、本発明は、以下の特許請求の範囲及びこれらの等価な技術により定義される保護範囲内に入るという条件で、これらの修正及び変更をカバーすることを意図する。

20

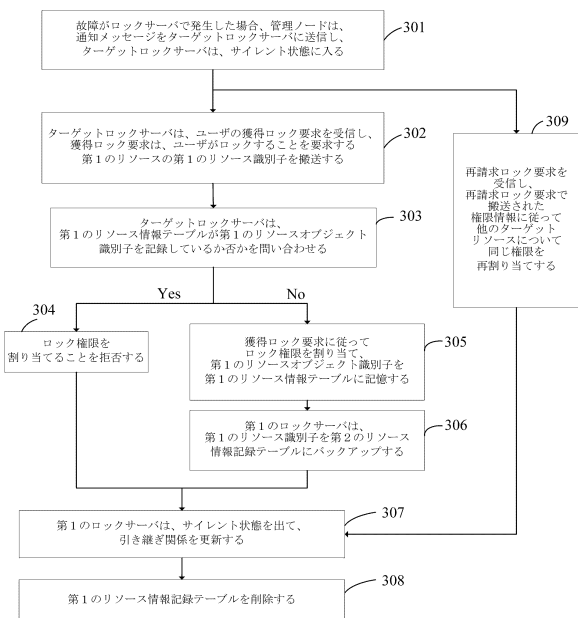
【図1】



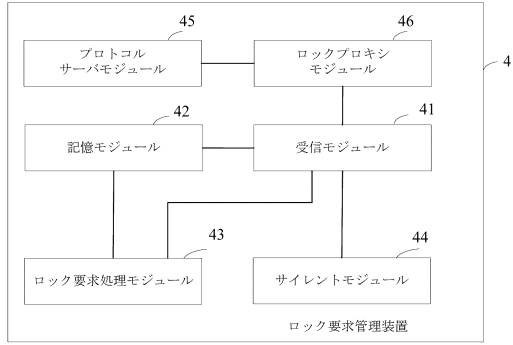
【図2】



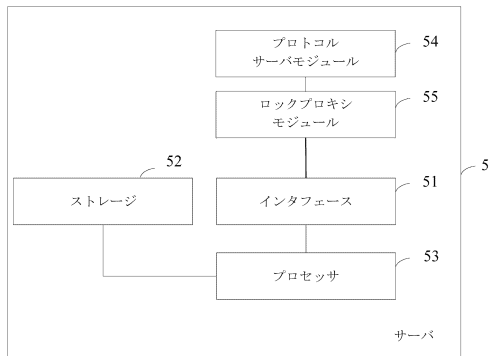
【図3】



【図4】



【図5】



フロントページの続き

(74)代理人 100070150

弁理士 伊東 忠彦

(74)代理人 100091214

弁理士 大貫 進介

(72)発明者 馮 銳

中国518129 広 東 省深 チェン 市 龍 崗 区坂田 華 為 総 部 辦
公楼

(72)発明者 陳 光

中国518129 広 東 省深 チェン 市 龍 崗 区坂田 華 為 総 部 辦
公楼

(72)発明者 劉 軍

中国518129 広 東 省深 チェン 市 龍 崗 区坂田 華 為 総 部 辦
公楼

審査官 月野 洋一郎

(56)参考文献 米国特許第07356531(US, B1)

米国特許第08296599(US, B1)

特開2011-242949(JP, A)

特開2014-048969(JP, A)

特開平11-338843(JP, A)

特開2011-210106(JP, A)

特表2006-525573(JP, A)

特表2005-528684(JP, A)

米国特許出願公開第2006/0212453(US, A1)

米国特許出願公開第2013/0174165(US, A1)

米国特許出願公開第2015/0347239(US, A1)

新田 淳 外1名, ディスク共用DBクラスタに向けた分散ロックマネージャ, 情報処理学会研究
報告 2012(平成24)年度 3 [CD-ROM], 日本, 一般社団法人情報処理学会
, 2012年11月 5日, p.1-p.10, 研究報告 情報技術とアクセス技術(IFIT) No.107

(58)調査した分野(Int.Cl., DB名)

G06F 12/00