

【公報種別】特許法第17条の2の規定による補正の掲載
【部門区分】第6部門第3区分
【発行日】令和4年1月26日(2022.1.26)

【公開番号】特開2019-128959(P2019-128959A)
【公開日】令和1年8月1日(2019.8.1)
【年通号数】公開・登録公報2019-031
【出願番号】特願2019-6709(P2019-6709)
【国際特許分類】

G 0 6 F 12/00(2006.01)

G 0 6 F 3/06(2006.01)

G 0 6 F 11/10(2006.01)

【F I】

G 0 6 F 12/00 5 4 5 A

G 0 6 F 3/06 3 0 5 C

G 0 6 F 11/10 6 1 2

G 0 6 F 12/00 5 3 1 Z

【手続補正書】

【提出日】令和4年1月18日(2022.1.18)

【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】

単一の論理ユニットとして1つの信頼性グループにグループ化されて仮想装置管理レイヤによって管理される1つ以上のストレージ装置を含むキーバリューストレージシステムにおけるデータ格納方法であって、

前記データが前記データを格納するために客體複製を含む信頼性メカニズムに対応する閾値を満足することを決定する段階と、

メタデータに含まれて前記信頼性メカニズムを識別する信頼性メカニズム識別子によって識別される前記信頼性メカニズムを使用する段階と、

前記信頼性メカニズムに応じて前記データを格納する段階と、を有し、

前記データを格納する段階は、

K Vバリューストレージ装置を選択する段階と、

前記K Vバリューストレージ装置に対応するキーをハッシングするためのハッシュを計算する段階と、

前記K Vバリューストレージ装置に対応するキー客體のレプリカを格納するために前記1つ以上のストレージ装置の中のストレージ装置の部分集合を決定する段階と、

同一の使用者キー名の下で前記ストレージ装置の部分集合に前記K Vバリューストレージ装置に対応するアップデートされたバリューストレージ装置を書込む段階と、を含むことを特徴とする方法。

【請求項2】

前記閾値は、前記データの客體サイズ、前記データの処理能力考慮 (throughput consideration)、前記データの読出し／書込み温度 (read/write temperature)、及び前記1つ以上のストレージ装置の基本 (underlying) 消去コーディング能力の中の1つ以上に基づくことを特徴とする請求項1に記載の方法。

【請求項3】

ブルームフィルタ又はキャッシュを使用して前記信頼性メカニズムに対する前記データ

10

20

30

40

50

をテストする段階を更に含むことを特徴とする請求項 1 に記載の方法。

【請求項 4】

前記信頼性メカニズム、前記データを格納する前記 1 つ以上のストレージ装置に対するチェックサムを含む前記メタデータ、前記データを格納する前記 1 つ以上のストレージ装置に格納された前記データのバリュー客体のサイズ、及び前記 1 つ以上のストレージ装置の中のどこに前記データが格納されているのかを示す前記 1 つ以上のストレージ装置のパリティグループのメンバーの位置を記録するための前記メタデータを前記データに対応するキーと共に挿入する段階を更に含むことを特徴とする請求項 1 に記載の方法。

【請求項 5】

前記信頼性メカニズムは、パッキング (p a c k i n g) を含み、
前記データを格納する段階は、
前記信頼性グループの前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置にそれぞれ格納された 1 つ以上のキー客体を選択する段階と、
前記 1 つ以上のキー客体に対応する 1 つ以上のバリュー客体を取り出す段階と、
前記 1 つ以上のバリュー客体の仮想バリューサイズが同一になるように前記 1 つ以上のバリュー客体の中の最大のバリューサイズを有しないバリュー客体の終端に仮想ゼロをパディングする段階と、
前記 1 つ以上のキー客体から 1 つ以上のパリティ客体を生成する段階と、
前記 1 つ以上のキー客体を前記第 1 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、
前記 1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、を含み、
前記第 2 のそれぞれの 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別 (d i s t i n c t) され、
前記 1 つ以上のキー客体の数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 1 に記載の方法。

【請求項 6】

前記信頼性メカニズムは、消去コーディングを使用するパッキングを含み、
前記 1 つ以上のストレージ装置は、M D S (m a x i m u m d i s t a n c e s e p a r a b l e) 消去コーディングで構成されることを特徴とする請求項 5 に記載の方法。

【請求項 7】

前記信頼性メカニズムは、再生成消去コーディングを使用するパッキングを含み、
前記 1 つ以上のストレージ装置は、再生成消去コーディングで構成されることを特徴とする請求項 5 に記載の方法。

【請求項 8】

前記信頼性メカニズムは、スプリットिंगを含み、
前記データを格納する段階は、
前記 K V バリューを 1 つ以上の同一サイズの客体に分割する段階と、
前記 1 つ以上の同一サイズの客体から 1 つ以上のパリティ客体を生成する段階と、
前記ハッシュに基づいて前記 1 つ以上のストレージ装置の中の前記 K V バリューが位置する主要ストレージ装置を決定する段階と、
前記主要ストレージ装置から始めて連続する順序で、1 つ以上のバリュー客体を前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置に書き込み、1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、を含み、
前記 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別 (d i s t i n c t) され、
前記 1 つ以上のバリュー客体の数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 1 に記載の方法。

【請求項 9】

10

20

30

40

50

前記信頼性メカニズムは、消去コーディングを使用するスプリッティングを含み、
前記 1 つ以上のストレージ装置は、MDS (maximum distance separable) 消去コーディングで構成されることを特徴とする請求項 8 に記載の方法。
 【請求項 10】

前記信頼性メカニズムは、再生成消去コーディングを使用するスプリッティングを含み、
前記 1 つ以上のストレージ装置は、再生成消去コーディングで構成され、
前記データを格納する段階は、
前記再生成消去コーディングを使用して前記 1 つ以上の同一サイズの客体を 1 つ以上のサブパケットに分割する段階と、
前記 1 つ以上のパリティ客体を 1 つ以上のパリティサブパケットに分割する段階と、を更に含むことを特徴とする請求項 8 に記載の方法。

10

【請求項 11】
信頼性メカニズムに基づいてデータを格納するためのデータ信頼性システムであって、
ステートレスデータ保護を使用して仮想装置として構成される 1 つ以上のストレージ装置と、

客体複製を含む前記信頼性メカニズムに応じて前記 1 つ以上のストレージ装置を前記仮想装置として管理して前記 1 つ以上のストレージ装置にデータを格納するように構成された仮想装置管理レイヤと、を備え、
前記仮想装置管理レイヤは、

20

前記データを格納するための前記信頼性メカニズムに対応する閾値を決定し、
メタデータに含まれて前記信頼性メカニズムを識別する信頼性メカニズム識別子によって識別される前記信頼性メカニズムを使用し、
前記信頼性メカニズムに応じて前記データを格納するように構成され、

前記仮想装置管理レイヤは、
KV バリユーを選択し、
前記 KV バリユーに対応するキーをハッシングするためのハッシュを計算し、
前記 KV バリユーに対応するキー客体のレプリカを格納するために前記 1 つ以上のストレージ装置の中のストレージ装置の部分集合を決定し、

同一の使用者キー名の下で前記ストレージ装置の部分集合に前記 KV バリユーに対応するアップデートされたバリユーを書込むことによって、前記データを格納するように構成されることを特徴とするデータ信頼性システム。

30

【請求項 12】
 前記信頼性メカニズムは、パッキングを含み、
 前記仮想装置管理レイヤは、

前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置に格納された 1 つ以上のキー客体を選択し、
前記 1 つ以上のキー客体に対応する 1 つ以上のバリユー客体を取り出し、
前記 1 つ以上のバリユー客体の仮想バリユーサイズが同一になるように前記 1 つ以上のバリユー客体の中の最大のバリユーサイズを有しないバリユー客体の終端に仮想ゼロをパディングし、

40

前記 1 つ以上のキー客体から 1 つ以上のパリティ客体を生成し、
前記 1 つ以上のキー客体を前記第 1 のそれぞれの 1 つ以上のストレージ装置に書込み、
前記 1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込むことによって、前記データを格納するように構成され、

前記第 2 のそれぞれの 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別され、

前記 1 つ以上のキー客体の数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 11 に記載のデータ信頼性システム。

50

【請求項 1 3】

前記信頼性メカニズムは、スプリッティングを含み、

前記仮想装置管理レイヤは、

前記 K V バリユーを 1 つ以上の同一サイズの客体に分割し、

前記 1 つ以上の同一サイズの客体から 1 つ以上のパリティ客体を生成し、

前記ハッシュに基づいて前記 1 つ以上のストレージ装置の中の前記 K V バリユーが位置する主要ストレージ装置を決定し、

前記主要ストレージ装置から始めて連続する順序で、1 つ以上のバリユー客体を前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置に書き込み、1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込むことによって、前記データを格納するように構成され、

前記 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別 (distinct) され、

前記 1 つ以上のバリユー客体の数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 1 1 に記載のデータ信頼性システム。

【請求項 1 4】

前記信頼性メカニズムは、再生成消去コーディングを使用するスプリッティングを含み、

前記 1 つ以上のストレージ装置は、再生成消去コーディングで構成され、

前記仮想装置管理レイヤは、前記再生成消去コーディングを使用して前記 1 つ以上の同一サイズの客体を 1 つ以上のサブパケットに分割し、前記 1 つ以上のパリティ客体を 1 つ以上のパリティサブパケットに分割することによって、前記データを格納するように更に構成されることを特徴とする請求項 1 3 に記載のデータ信頼性システム。

【請求項 1 5】

プロセッサ上で実行される場合、単一の論理ユニットとして 1 つの信頼性グループにグループ化されて仮想装置管理レイヤによって管理される 1 つ以上のストレージ装置を含むキーバリユー信頼性システムにおけるデータ格納方法を具現するコンピュータコードを有する非一時的なコンピュータ読み取り可能な記録媒体であって、

前記方法は、

前記データが前記データを格納するために客体複製を含む信頼性メカニズムに対応する閾値を満足することを決定する段階と、

メタデータに含まれて前記信頼性メカニズムを識別する信頼性メカニズム識別子によって識別される前記信頼性メカニズムを使用する段階と、

前記信頼性メカニズムに応じて前記データを格納する段階と、を有し、

前記データを格納する段階は、

K V バリユーを選択する段階と、

前記 K V バリユーに対応するキーをハッシングするためのハッシュを計算する段階と、

前記 K V バリユーに対応するキー客体のレプリカを格納するために前記 1 つ以上のストレージ装置の中のストレージ装置の部分集合を決定する段階と、

同一の使用者キー名の下で前記ストレージ装置の部分集合に前記 K V バリユーに対応するアップデートされたバリユーを書込む段階と、を含むことを特徴とする非一時的なコンピュータ読み取り可能な記録媒体。

【請求項 1 6】

前記信頼性メカニズムは、パッキングを含み、

前記データを格納する段階は、

前記信頼性グループの前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置に格納された 1 つ以上のキー客体を選択する段階と、

前記 1 つ以上のキー客体に対応する 1 つ以上のバリユー客体を取り出す段階と、

前記 1 つ以上のバリユー客体の仮想バリユーサイズが同一になるように前記 1 つ以上のバ

10

20

30

40

50

リユー客体の中の最大のバリュースイズを有しないバリュースイズの終端に仮想ゼロをパディングする段階と、

前記 1 つ以上のキー客体から 1 つ以上のパリティ客体を生成する段階と、

前記 1 つ以上のキー客体を前記第 1 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、

前記 1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、を含み、

前記 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別 (distinct) され、

前記 1 つ以上のキー客体の数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 15 に記載の非一時的なコンピュータ読み取り可能な記録媒体。

10

【請求項 17】

前記信頼性メカニズムは、スプリッティングを含み、

前記データを格納する段階は、

前記 K V バリュースイズを 1 つ以上の同一サイズの客体に分割する段階と、

前記 1 つ以上の同一サイズの客体から 1 つ以上のパリティ客体を生成する段階と、

前記ハッシュに基づいて前記 1 つ以上のストレージ装置の中の前記 K V バリュースイズが位置する主要ストレージ装置を決定する段階と、

前記主要ストレージ装置から始めて連続した順序で、1 つ以上のバリュースイズを前記 1 つ以上のストレージ装置の中の同一数の第 1 のそれぞれの 1 つ以上のストレージ装置に書き込み、1 つ以上のパリティ客体を前記 1 つ以上のストレージ装置の中の同一数の第 2 のそれぞれの 1 つ以上のストレージ装置に書込む段階と、を含み、

20

前記第 2 のそれぞれの 1 つ以上のストレージ装置は、前記第 1 のそれぞれの 1 つ以上のストレージ装置から峻別 (distinct) され、

前記 1 つ以上のバリュースイズの数と前記 1 つ以上のパリティ客体の数との和は、前記 1 つ以上のストレージ装置の数と同一であることを特徴とする請求項 15 に記載の非一時的なコンピュータ読み取り可能な記録媒体。

30

40

50