

(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2017-138982

(P2017-138982A)

(43) 公開日 平成29年8月10日(2017.8.10)

(51) Int.Cl.		F I		テーマコード (参考)
G06F 3/06	(2006.01)	G06F 3/06	301Z	
G06F 3/08	(2006.01)	G06F 3/08	H	
G06F 13/10	(2006.01)	G06F 13/10	340A	

審査請求 未請求 請求項の数 20 O L (全 20 頁)

(21) 出願番号 特願2017-15503 (P2017-15503)
 (22) 出願日 平成29年1月31日 (2017.1.31)
 (31) 優先権主張番号 62/290, 295
 (32) 優先日 平成28年2月2日 (2016.2.2)
 (33) 優先権主張国 米国 (US)
 (31) 優先権主張番号 15/133, 085
 (32) 優先日 平成28年4月19日 (2016.4.19)
 (33) 優先権主張国 米国 (US)

(特許庁注：以下のものは登録商標)

1. ETHERNET

(71) 出願人 390019839
 三星電子株式会社
 Samsung Electronics
 Co., Ltd.
 大韓民国京畿道水原市靈通区三星路129
 129, Samsung-ro, Yeon
 gtong-gu, Suwon-si, G
 yeonggi-do, Republic
 of Korea

(74) 代理人 110000051
 特許業務法人共生国際特許事務所

(72) 発明者 チェ, インソク スティーヴン
 アメリカ合衆国, カリフォルニア州 94
 063, レッドウッド シティ, メインス
 トリート 333, スイート 301
 最終頁に続く

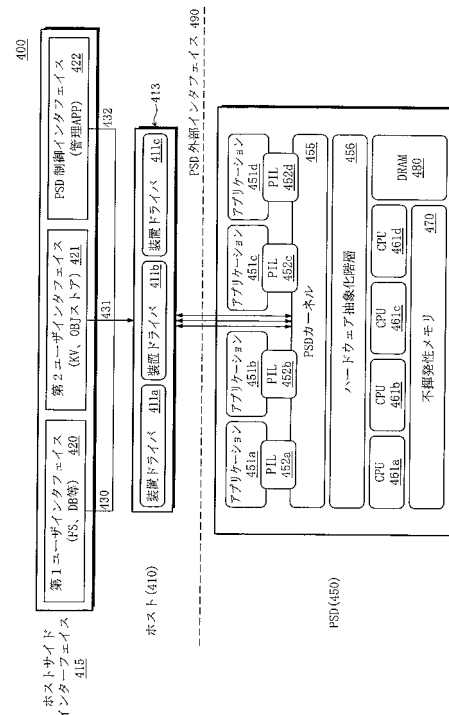
(54) 【発明の名称】 ポリモルフィック格納装置

(57) 【要約】

【課題】ハードウェアリソースの浪費を減らすか、或いはその使用を最適化できるポリモルフィック格納装置 (PSD) を提供する。

【解決手段】1つ以上のアプリケーションを含むアプリケーションコンテナ、ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、複数のパケットに含まれたコマンドに基づいてアプリケーションコンテナ内の1つのアプリケーションへ複数のパケットをルーティングするポリモルフィック格納装置カーネル、並びに1つ以上のアプリケーション及びPSDカーネルの間にインタフェースを提供するPSDインタフェース階層を含む。実施形態において、PSDは、ホストコンピュータが備えるPSD制御インタフェースを通じて受信された制御コマンドに従って自身のハードウェアリソースを再構成する。

【選択図】 図4



【特許請求の範囲】**【請求項 1】**

1つ以上のアプリケーションを含むアプリケーションコンテナと、
ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の1つのアプリケーションへ前記複数のパケットをルーティングするポリモルフィック格納装置 (`polymorphic storage device`、以下、“PSD”)カーネルと、

前記1つ以上のアプリケーション及び前記PSDカーネルの間にインタフェースを提供するPSDインタフェース階層と、を含むことを特徴とする格納装置。

10

【請求項 2】

前記ホストコンピュータからの制御コマンドに従って再構成可能な1つ以上のハードウェアを更に含み、

前記再構成可能なハードウェアは、1つ以上のプロセッサ、不揮発性メモリ、及びDRAM (`dynamic random access memory`) を含むことを特徴とする請求項1に記載の格納装置。

【請求項 3】

PSD制御インタフェース及び1つ以上のアプリケーションインタフェースを有するホストサイドインタフェースが、前記ホストコンピュータに提供されることを特徴とする請求項1に記載の格納装置。

20

【請求項 4】

前記格納装置は、前記PSD制御インタフェースを通じて受信された制御コマンドに従って第1種類の格納装置から第2種類の格納装置に変換されることを特徴とする請求項1に記載の格納装置。

【請求項 5】

前記第1種類の格納装置は、インストレージコンピューティング装置 (`in-storage computing device`)、キーバリューストア (`key-value store`) 装置、及びHadoop分散ファイルシステム (`Hadoop distributed file system`、HDFS) 装置、オブジェクト-ストア装置の中の何れか1つであり、

30

前記第2種類の格納装置は、前記インストレージコンピューティング装置、前記キーバリューストア装置、及び前記HDFS装置の中の他の1つであることを特徴とする請求項4に記載の格納装置。

【請求項 6】

前記PSD制御インタフェースは、前記格納装置に制御コマンド及びメッセージを提供するように構成されることを特徴とする請求項3に記載の格納装置。

【請求項 7】

前記アプリケーションインタフェースは、バックワード交換性のためにエミュレートされたホストサイドインタフェース (`emulated host-side interface`) 及びPSD対応アプリケーションインタフェース (`PSD-aware application interface`) を含むことを特徴とする請求項3に記載の格納装置。

40

【請求項 8】

1つ以上のベンダーコマンド (`vendor commands`) が、前記PSD制御インタフェースを通じて提供されることを特徴とする請求項3に記載の格納装置。

【請求項 9】

前記1つ以上のベンダーコマンドは、クリエートコマンド (`create command`)、スタートコマンド (`start command`)、ストップコマンド (`stop command`)、ポーズコマンド (`pause command`) (別名、サスペンド (`suspend`))、及びレジュームコマンド (`resume`) を含むことを特徴

50

とする請求項 8 に記載の格納装置。

【請求項 10】

前記 1 つ以上のベンダーコマンドは、前記 PSD 制御インタフェースを通じて、前記格納装置において動作するアプリケーションを他の格納装置に移動するための移動コマンド (migration command)、或いは前記格納装置において動作するアプリケーションを他の格納装置において複製するための複製コマンド (duplication command) を含むことを特徴とする請求項 8 に記載の格納装置。

【請求項 11】

前記アプリケーションは、集中化方式 (centralized manner) 或いは、非集中化方式 (decentralized manner) により移動されるか、
或いは複製されることを特徴とする請求項 10 に記載の格納装置。

10

【請求項 12】

ホストコンピュータと、
ポリモルフィック格納装置 (polymorphic storage device、
以下、“ PSD ”) と、

前記ホストコンピュータと前記 PSD との間にインタフェースを提供する PSD 外部インタフェースと、を含み、

前記 PSD は、

1 つ以上のアプリケーションを含むアプリケーションコンテナと、

前記ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の 1 つのアプリケーションへ前記複数のパケットをルーティングする PSD カーネルと、

20

前記 1 つ以上のアプリケーション及び前記 PSD カーネルの間にインタフェースを提供する PSD インタフェース階層と、を含むことを特徴とする格納システム。

【請求項 13】

前記ホストコンピュータから制御コマンドに従って再構成される 1 つ以上のハードウェアを更に含み、

前記 PSD の前記再構成可能なハードウェアは、1 つ以上のプロセッサ、不揮発性メモリ、及び DRAM (dynamic random access memory) を含むことを特徴とする請求項 12 に記載の格納システム。

30

【請求項 14】

PSD 制御インタフェース及びアプリケーションインタフェースを有するホストサイドインタフェースが、前記ホストコンピュータに提供されることを特徴とする請求項 12 に記載の格納システム。

【請求項 15】

前記 PSD は、前記 PSD 制御インタフェースを通じて受信された制御コマンドに従って第 1 種類の格納装置から第 2 種類の格納装置に変換されることを特徴とする請求項 14 に記載の格納システム。

【請求項 16】

前記 PSD 制御インタフェースは、前記 PSD に制御コマンド及びメッセージを提供するように構成されることを特徴とする請求項 14 に記載の格納システム。

40

【請求項 17】

格納装置の方法において、

前記格納装置のアプリケーションコンテナに 1 つ以上のアプリケーションを格納する段階と、

ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信する段階と、

前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の 1 つのアプリケーションへ前記複数のパケットをルーティングする段階と、

50

PSD (polymorphic storage device) 制御インタフェースを通じて前記ホストコンピュータから受信された制御コマンドに従って前記格納装置の1つ以上のハードウェアを(再)構成するように前記格納装置でPSDカーネルを実行する段階と、

前記PSDカーネルによって構成されたように前記格納装置を動作する段階と、を含むことを特徴とする方法。

【請求項18】

前記ホストコンピュータから制御コマンドに従って1つ以上のハードウェアを再構成する段階を更に含み、

前記再構成可能なハードウェアは、1つ以上のプロセッサ、不揮発性メモリ、及びDRAM (dynamic random access memory) を含むことを特徴とする請求項17に記載の方法。

10

【請求項19】

前記ホストコンピュータにPSD制御インタフェース及びアプリケーションインタフェースを有するホストサイドインタフェースを提供する段階を更に含むことを特徴とする請求項17に記載の方法。

【請求項20】

前記PSD制御インタフェースを通じて受信された制御コマンドに従って第1種類の格納装置から第2種類の格納装置に前記PSDを変換させる段階を更に含むことを特徴とする請求項19に記載の方法。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は格納装置に係り、更に詳細には装置制御インタフェースを通じて相異なる装置に変換可能なポリモルフィック格納装置及びその方法に係る。

【背景技術】

【0002】

ハードディスクドライバのような伝統的なデータ格納装置と比較して、SSDs (solid state drivers) は、さらなる複雑性 (complexity) と擬似的に汎用なプログラム可能性 (pseudo-general programmability) を提供する。例えば、フラッシュメモリ基盤のSSD (複数) は、益々強力なCPU (複数) と大容量メモリ及びストレージを組み入れ (package)、このようリソースをユーザアプリケーション (user application) に利用可能にする。その上、データ集中化 (data-intensive) アプリケーションの場合、コードをデータに移動することはデータをコードに移動することよりはるかに更に効率的であると考えられている。

30

従って、効率的なデータ処理システムは、早期選択プッシュダウン (early selection pushdown) 及び早期集約 (early aggregation、又は pre-aggregation) のような技術を利用してクエリー (検索要求) 処理パイプライン (query processing pipeline) のなるべく下の方にコードをプッシュしようと試図 (try) する。同様に、並列或いは分散データ処理システムは、データを保持するノードに近いなるべく多くのクエリーを実行できる。

40

【0003】

伝統的に、コード・オフロード (code-offloading) 技術は、臨機応変に (アドホック (ad-hoc)) SSDのファームウェア及びホストコンピュータのOS (operating system) を微調整 (tweak、いじくる) することによって具現された。

永久的なストレージ、揮発性ストレージ、及びプロセシングの間の境界が段々、更に曖

50

味になっているけれども、ホストOS、ライブラリ、固定されたプロトコル、及び格納装置の抽象化層の間に存在する機能的な分離は、冗長（重複的）な具現をよく（commonly）引き起こし、隘路（bottleneck）或いはリソースの浪費屋になり得る。NoSQL（non SQL（Structured Query Language）データベース（DB）（例、キーバリューストア、ドキュメントストア、グラフ、オブジェクトデータベース、等）は、最近のコンピュータシステムで幅広く使用される。関係型（relational）DBと比較して、NoSQL DBは簡単、柔軟、軽量、かつ特定ワークロード（workloads）に対しては優れて拡張性及び大幅な性能向上の利点を期待できる。ビッグデータ（big data）を有するクラウドディングコンピュータ及び大型システムへの速やかな移行は、NoSQL DBsの進展する普及に寄与している。

10

【先行技術文献】

【特許文献】

【0004】

【特許文献1】米国特許第8,583,938号公報

【特許文献2】米国特許第7,676,665号公報

【特許文献3】米国特許第8,812,976号公報

【特許文献4】米国特許公開第2009/0094597号明細書

【発明の概要】

【発明が解決しようとする課題】

20

【0005】

本発明の目的は、ハードウェアリソースの使用を削減し、或いは最適化できるポリモルフィック格納装置及びその方法を提供することにある。

【課題を解決するための手段】

【0006】

本発明の実施形態に係る格納装置は、1つ以上のアプリケーションを含むアプリケーションコンテナ、ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の1つのアプリケーションへ前記複数のパケットをルーティングするポリモルフィック格納装置（polymorphic storage device、以下、“PSD”）カーネル、並びに前記1つ以上のアプリケーション及び前記PSDカーネルの間にインタフェースを提供するPSDインタフェース階層を含む。

30

【0007】

本発明の実施形態に係る格納システムは、ホストコンピュータ、PSD、及び前記ホストコンピュータと前記PSDとの間にインタフェースを提供するPSD制御インタフェースを含み、前記PSDは、1つ以上のアプリケーションを含むアプリケーションコンテナ、前記ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の1つのアプリケーションに前記複数のパケットをルーティングするPSDカーネル、並びに、前記1つ以上のアプリケーション及び前記PSDカーネルの間にインタフェースを提供するPSDインタフェース階層を含む。

40

【0008】

本発明の実施形態に係る格納装置の方法は、前記格納装置のアプリケーションコンテナに1つ以上のアプリケーションを格納する段階、ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信する段階、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナからアプリケーションに前記複数のパケットをルーティングする段階、PSD制御インタフェースを通じて前記ホストコンピュータから受信された制御コマンドに従って前記格納装置の1つ以上のハードウェアを構成するように前記格納装置上でPSDカーネルを動作させる段階、並びに前記PSDカーネルにより構成されたように前記格納装置を動作する段階を含む。

50

【発明の効果】

【0009】

本発明の実施形態に係るポリモルフィック格納装置及びその方法は、ポリモルフィック格納装置制御インタフェースを通じてハードウェアリソースを再構成することによって、リソースの浪費を削減し、或いはリソースの使用を最適化できる。

【図面の簡単な説明】

【0010】

【図1】伝統的なNoSQLデータベースDBシステムを示す図面である。

【図2】他の伝統的なNoSQL DBシステムのためのシステム構造に対する実施形態を示す。

10

【図3】例示的な格納装置のメモリ制御器の階層を示す図面である。

【図4】本発明の実施形態に係るポリモルフィック格納装置PSDを有する例示的な格納システムに対するブロックダイヤグラムである。

【図5】本発明の実施形態に係るポリモルフィック格納装置のファームウェアスタック(firmware stack)を示す図面である。

【図6】(A)は、本発明の実施形態に係る集中化方式に従うアプリケーションの移動及び複製を例示的に示す図面であり、(B)は、本発明の実施形態に係る非集中化方式に従うアプリケーションの移動及び複製を例示的に示す図面である。

【発明を実施するための形態】

【0011】

20

ここに開示された特徴及び教示の各々は、装置制御インタフェースを通じて他の装置に変換できるポリモルフィック格納装置(polymorphic storage device)を提供するために、個別的に又は他の特徴及び教示と組み合わせて利用できる。このような追加的特徴及び教示の多くを、個別的に、及び組み合わせて利用する実施形態は、添付された図面を参照して詳細に説明される。このような詳細な説明は単に、当業者が本教示の態様を実行できるよう教示することを意図し、請求項の範囲を制限しない。従って、詳細な説明に上述された特徴の組み合わせは、最も広い意味で教示を実施するためには必要とは限らず、代わりに、単に本教示の特に代表的な実施形態を記述するために教示される。

30

【0012】

以下の説明で、説明の目的として、特定名称が本発明の完全な理解を提供するために説明される。しかし、このような特定細部事項は本発明の思想を実施しなくともよいことが当業者に明確である。

【0013】

本明細書の一部ではコンピュータメモリ内にデータビットに対する演算のシンボル表現とアルゴリズムの用語で表現される。このようなアルゴリズムの説明及び表現は効果的に他の当業者に作業の実体を伝達するためにデータ処理分野の当業者によって使用される。ここで、アルゴリズムは、伝統的に、最終結果に至る段階の一貫性あるシーケンス(self-consistent sequence)が考慮される。段階は物理的量の物理的操作を要求するものである。ほとんど、必須的ではないが、このような量は、便利のためにしばしば電気或いは磁気信号の形態に取ってもよい。このような信号はビット、値、要素、シンボル、キャラクター、用語、数字、或いはそのようなことと言及される。

40

【0014】

しかし、これらの全て及び類似な用語は、適切な物理量と関連され単なるこのような量に適用される便利なラベルであることを有意しなければならぬ。具体的には下の説明から明になるように、説明を通じて、このような“プロセッシング”、“コンピューティング”、“計算”、“決定”、“表示”、或いはそのような用語を利用する論議は、コンピュータシステム或いは類似な電子コンピューティング装置のアクション(actions)及び処理であると言及される。コンピュータシステム或いは類似な電子コンピューティン

50

グ装置は、コンピュータシステムのレジスタ内部の物理的（電子的）量として表現されたデータを、コンピュータシステムメモリ或いはレジスタ或いは他の情報格納、伝送或いはディスプレイ装置内の物理的量として類似に表現された他のデータ操作及び変換する。

【0015】

ここに提供されたアルゴリズムは本質的に任意の特定コンピュータ又は他の装置に関連されたことではない。多様な汎用システム、コンピュータサーバー又は個人用コンピュータは本明細書の教示に従うプログラムと共に使用され、或いは要求された方法段階を遂行するにより特殊化された装置を構成することが便利であることができる。多様なこれらのシステムのための要求される構造は、下の説明から開示される。多様なプログラミング言語が本明細書に記載されたように発明の教示を具現するのに使用されることもできると理解されるべきである。

10

【0016】

また、代表的な実施形態の多様な特徴と従属請求項は本教示の追加的な有用な実施形態を提供するために列挙されない方式に結合されることができ。また、請求された発明を制限する目的のみならず、オリジナル発明の目的のためにエンティティグループのすべての値の範囲或いは指示は、可能なすべての中間値又は中間エンティティを開示することが明確である。また、明示的な基準及び図面に図示された構成要素の形状は、本開示で実施される方式を理解できるように設計するが、寸法及び実施形態に示した形状に限定されないことを有意しなければならない。

20

【0017】

図1は伝統的なNoSQL_DB（データベース）システムを示す図面である。NoSQL_DBシステム100は、オーラップされた構成要素及び冗長なマッピングを含む。NoSQL_DBシステム100は、ホストコンピュータ110及び格納装置150を含む。ここで、格納装置150はホストツースSDインタフェース（Host-to-SSD interface）190を通じてホストコンピュータ110と通信する。ホストコンピュータ110は、キーをLBA（logical block address）変換するためのキーLBAマッピングテーブル121を有するNoSQL_DB（120）、オペレーティングシステム（OS）130、及びSSDインタフェース140を含む。

30

ホストツースSDインタフェース190の実例は、下記に限定されないが、PCIe（peripheral component interconnect express）、SAS（serial attached SCSI）、イーサネット（ethernet）、及びSATA（serial AT attachment）を含み得る。

格納装置150はメモリ制御器160及び不揮発性メモリ170を含む。メモリ制御器160はホストインタフェース161、LBA-PBAマッピングテーブル162、エラー訂正コード（ECC）メモリ163、及び不揮発性メモリ170とのインタフェーシングのためのフラッシュインタフェース164を含む。ホストコンピュータ110と格納装置150とは、ホストツースSDインタフェース190を通じてコマンド、アドレス（LBAs）、及びデータを通信する。

40

【0018】

ここで、2つの冗長なマッピング階層が存在する。キー（key）LBAマッピングのための第1マッピング階層と、LBA-PBAマッピングのための第2マッピング階層である。伝統的に、ホストコンピュータ110及び格納装置150は、格納装置150における格納位置のアドレスを特定するためにLBAに基づいて互いに通信する。互換性のために、格納装置150はLBAを用いて動作し、メモリ制御器160のフラッシュ変換階層FTL（162）はLBA-PBA変換を維持する。

FTLはLBA-PBAマッピングテーブル162を利用して論理ブロックアドレス（

50

LBA)を物理ブロックアドレス(PBA)に変換し、不揮発性メモリ170の物理ブロックに対する要請された動作(例、読出し、書込み、読出し/変更(modify)/書込み(RMW)、及び削除)を遂行するためにマッピングテーブル162によって提示されたように不揮発性メモリ170の物理ブロックにアクセスする。

【0019】

ホストコンピュータ110のキーLBAマッピングは、ホストコンピュータと格納装置との間に適切な約束が形成されておれば、省略できる。

図2は、他の伝統的なNoSQLDBシステムのためのシステムを示す。NoSQLDBシステム200はホストコンピュータ210及びホストツールSSDインタフェース290を通じて通信する格納装置250を含む。ホストコンピュータ210はNoSQLユーザアプリケーションプログラムインタフェース(API)220、オペレーティングシステム(OS)230、及びSSDインタフェース240を含む。図1に示されたホストツールSSDインタフェース190と同様に、ホストツールSSDインタフェース290はPCIe、SAS、イーサネット、及びSATA或いはそのようなものの中で任意のものである。

格納装置250は不揮発性メモリ270とメモリ制御器260を含む。メモリ制御器260はホストインタフェース261、NoSQL__DB統合(integrated)FTL(262)、エラー訂正コード(ECC)メモリ263、及び不揮発性メモリ270とのインタフェーシングのためのフラッシュインタフェース264を含む。NoSQL__DB統合(integrated)FTL(262)はホストコンピュータ210内のキーLBAマッピングテーブルを排除する。従って、NoSQL__DBシステム200は、2つの重複マッピング階層を要求する図1のNoSQL__DBシステム100と比較して更に集約(aggregate)された性能を有する。

【0020】

ホストコンピュータ210及び格納装置250は、ホストツールSSDインタフェース290を通じてクエリーコマンド、キー、及びデータを通信する。不揮発性メモリ270は、下記に限定されないが、フラッシュメモリ、PRAM(phase change random access memory)、STT-MRAM(spin-torque magnetic random access memory)、及びReRAM(resistive RAM)等を含む多様な種類の不揮発性メモリである。

【0021】

格納装置の行動を定義する本発明のポリモルフィック格納装置(present polymorphic storage device)のファームウェアはファームウェアアップデートを通じて再構成される。ファームウェアアップデートを通じて、本発明のポリモルフィック格納装置はオリジナル構成から異なる型の装置に変換できる。例えば、本発明のポリモルフィック格納装置は、最初、汎用格納装置として構成され、ファームウェアアップデートによって、特殊目的装置に変換できる。このような特殊目的装置は、インストレージコンピューティング装置(in-storage computing device、格納装置内コンピューティング可能な格納装置)、キーバリューストア装置(key-value store device)、Hadoop分散ファイルシステム(Hadoop(又は、Apache__Hadoop) distributed file system、HDFS)装置、オブジェクトストア装置(object-store device)、その他である。

【0022】

図3は、従来技術に係る例示的な格納装置のメモリ制御器の階層を示す図面である。メモリ制御器360は、HIL(host interface layer、ホストインタフェース階層)310、FTL(flash translation layer、フラッシュ翻訳階層)320、VFL(virtual flash layer、仮想フラッシュ階層)330、及びFIL(flash interface layer、

10

20

30

40

50

フラッシュインタフェース階層) 340の4つの階層を含む。このような4つの階層は格納装置のファームウェア内に具現される。ここで、該ファームウェアは格納装置に埋め込まれた(embodied)1つ以上のプロセッサ上で作動する。HIL(310)はSATA、SAS、或いはPCIeと同一バスインタフェースプロトコルを具現する。FTL(320)はホストOSの論理ブロックアドレス(LBA)を格納装置の不揮発性メモリの物理ブロックアドレス(PBA)にマッピングする役割を遂行する。VFL(330)及びFIL(340)は格納装置の不揮発性メモリと格納装置のDRAMとの間のデータを伝送する役割を遂行する。格納装置のファームウェアは、ECCロジック及びDMA(direct memory access)を含む他の特徴を提供できる。不揮発性メモリの入出力性能を向上するために、メモリ制御器360はチップレベル及び/或いはチャンネルレベルでインタリーピング技術を利用できる。

10

【0023】

本発明は装置制御インタフェース(device control interface)を通じて1つの形態の装置から他の形態の装置に変換可能であるポリモルフィック格納装置(polymorphic storage device、以下、“PSD”)を提供する。

図4は本発明の実施形態に係るポリモルフィック格納装置を有する例示的な格納システムに関するブロックダイヤグラムである。格納システム400はホストコンピュータ410、及びPSD外部インタフェース490を通じてホストコンピュータ410と通信するPSD(450)を含む。ホストコンピュータ410は、PSD(450)を再構成するのに使用される複数の装置ドライバ411a~411cを格納するための装置ドライバプール413を含む。

20

PSD(450)はPSDカーネル455、PSDカーネル455とアプリケーション451とをインタフェーシングするためのPSDインタフェース階層(PIL)452、及び再構成及び拡張可能なハードウェアを含む。ここで、再構成及び拡張可能なハードウェアは複数のCPU(461a~461d)、不揮発性メモリ470、及びDRAM480を含む。

PSD(450)は、再構成及び拡張可能なハードウェアとの通信をアプリケーション451に対して可能にする装置ドライバインタフェースを提供するためのハードウェア抽象化階層(hardware abstraction layer)456を更に含む。

30

本実施形態では、PSD(450)が再構成及び拡張可能なハードウェアを含むように開示されているが、本発明のPSD(450)は再構成及び拡張可能でないハードウェアを含み得ると理解されるべきである。再構成及び拡張可能なハードウェアはPSD(450)の柔軟性(flexibility)及びプログラマビリティ(programmability)を提供できる。

【0024】

本発明の実施形態において、ホストサイドインタフェース415は第1ユーザインタフェース420、第2ユーザインタフェース421、PSD(450)のためのPSD制御インタフェース422を含む。

40

第1ユーザインタフェース420は、例えば、ファイルシステム(FS)及び/或いはデータベース(DB)、或いはそのようなものとのインタフェースを提供する。第1ユーザインタフェース420を通じて、ファイルシステム及びデータベースに要求される入出力(I/O)制御コマンド(例、装置制御コマンド及びメタデータ)430はホストコンピュータ410へ/から、伝送される。

第2ユーザインタフェース421は、例えばキーバリューストア(KV)ストア及び/或いはオブジェクト(OBJ)ストアにインタフェースを提供し、キーバリューストア及び/或いはオブジェクトストアのためのクエリー431(例、put、get)を提供する。

PSD制御インタフェース422は、PSD(450)に対する装置制御インタフェースであって、PSD外部インタフェース490を通じてホストコンピュータ410及びP

50

PSD(450)にコマンド及びメッセージ432を提供する。ホストコンピュータ410の装置ドライバプール413内の複数の装置ドライバ411a~411cの中で何れか1つを利用して、PSD(450)はPSD制御インタフェース422から受信されたコマンド(コマンド及びメッセージ432のコマンド部分)に従って構成(或いは再構成)される。

PSD制御インタフェース422を通じた該コマンドはPSD(450)で動作するアプリケーション451a~451dを制御する。ここで、PSD(450)で動作するアプリケーション451a~451dは、PSD対応(PSD-aware、正確には、「PSDを意識した」である)アプリケーションと言及される。何故ならば、アプリケーション451a~451dは、PSD制御インタフェース422を通じて受信された該コマンドを利用してPSD(450)を(再)構成し、PSD(450)を第1種類の格納装置から第2種類格納装置に変換するからである。

また、PSD制御インタフェース422はホストサイドの装置ドライバプール413に格納された装置ドライバ及びそれらの装置ドライバマネージャに対する制御も提供する。

【0025】

本発明の実施形態において、PSD外部インタフェース490はバックワード互換性のためのホストサイドユーザインタフェース420、421、或いはPSD(450)上で動作するPSD対応アプリケーション451a~451dを管理及び制御するためのPSD制御インタフェース422をエミュレート(emulate)する。

ベンダー固有(vendor-specific)コマンドはPSD外部インタフェース490を通じてPSD対応アプリケーション451a~451dに提供される。このようなベンダー固有コマンドの実施形態は、下記に限定されないが、“生成(create)”、“開始/停止(start/stop)”、“休止/再開(pause/resume)”、等を含み得る。PSD(450)はホストコンピュータ410と通信してファームウェアアップデートをせずに機能とアプリケーションを追加、削除、及び変更できる。

【0026】

本発明の実施形態において、ホストコンピュータ410とPSD(450)との間のPSDインタフェース490は、PSD(450)内のオフロードされた(offloaded)或いは集積化された(integrated)機能を提供する。

例えば、PSD対応アプリケーション451或いはホストサイドアプリケーションは、格納システムインタフェースとよりは、むしろPSD(450)のユーザインタフェースと相互作用する。ここで、格納システムインタフェースはホストOS及びホストコンピュータ410の装置ドライバ411a~411cによって伝統的に提供される。

伝統的に、ホストコンピュータ410上で動作するホストサイドアプリケーションは、アプリケーションロジックを有するホストコンピュータでその機能性を具現する。本発明の実施形態において、ホストコンピュータ410で動作するホストサイドアプリケーションのためのアプリケーションロジックはPSD(450)内に分散される。

例えば、PSD(450)の集積化されたキーバリューフTLはホストサイドアプリケーションで実質的に(virtually)全てのロジックを除去でき、キーバリュースQL動作のためのユーザAPIのみがPSD(450)のPSDカーネル455に提供される。この動作の残りはPSD(450)の内部で処理され、PSD対応アプリケーション451はPSD外部インタフェース490を通じて受信されたコマンド及びメッセージを利用してキーバリュースQL動作を実行する。

【0027】

PSD(450)が1回に1つ以上のアプリケーション(例、415a~451d)を処理できるので、対応するホストサイド装置ドライバ411は装置ドライバプール413に存在する。もし、PSD(450)が装置ドライバの唯一の活性インスタンスを有すれば、装置ドライバプール413は伝統的な格納システムと同様の方法で1つの装置ドライバ411として動作する。

10

20

30

40

50

図4に示したように、PSD(450)はPSD(450)の内部アプリケーションのインスタンス451a~451dを制御するPSD制御インタフェース422を利用。この意味で、PSD制御インタフェース422がアプリケーション451と直接相互作用するのではなく、アプリケーション451の制御メカニズムを提供する、という点で、PSD制御インタフェース422の提供(provisioning)は、仮想マシン(VM)マネージャ或いはハイパーバイザー(hypervisor)に類似する。。

【0028】

多重アプリケーション451a~451dはPSD(450)上で同時に実行される。しかし、ホストコンピュータ410とPSD(450)との間の物理的な通信チャンネルの個数は物理的装置の個数により制限される。このような理由で、各アプリケーション451a~451dとPSD(450)の内部アプリケーションコンテナとの間の通信はPSD(450)の内部でマルチプレクシングされ、デマルチプレクシングされ得る。PSD外部インタフェース490は任意の物理インタフェースを含み得る。該物理インタフェースは、下記に限定されないが、SATA、SAS、NVMe、イーサネット、FC(fiber channel)、IB(InfiniBand)、RDMA(remote direct memory access)、等である。

また、ホストサイドの装置ドライバブル413は、物理インタフェースを提供する現存の通信プロトコル内に全てのメッセージ/コマンド/データを埋め込み/カプセル化(embed/encapsulate)するマルチプレクシングに係る処理を行う。幾つかの実施形態において、PSD(450)の自己のプロトコルは現存するプロトコルにデータをオーバーディングする際のオーバーヘッドを低減するために使用される。

【0029】

図5は本発明の実施形態に係るポリモルフィック格納装置PSDのファームウェアスタック(firmware stack)を示す図面である。PSDファームウェアスタック500はフラッシュ制御器階層535、PSDカーネル555、及びPSDインタフェース階層552、及びアプリケーションコンテナ510を含む。アプリケーションコンテナ510は1つ以上のアプリケーション511a、511b、511cを含む。

本発明の実施形態において、アプリケーション511cは残りのアプリケーション511a、511bから自分自身を差別化してPSD500で動作する現在のテナント(tenant)としてユーザアプリケーション511cと言及される。以後、アプリケーション511a、511b、511cはアプリケーション511として集合的に言及される。アプリケーション511は、ホストプロトコル及びメッセージプロセッサ513、ATL(application-specific address translation logic、特定アプリケーション用アドレス変換ロジック)514、アプリケーション固有ウェアレベリング(WL)515、及びアプリケーション固有ガーベッジコレクションロジック(GC)516を含む。

【0030】

例えば、アプリケーション511は、キーバリュキャッシングアプリケーション(key-value caching application)、HDFS(Hadoop分散ファイルシステム)アプリケーション、及びMongoDB(ドキュメント指向DB)アプリケーションの中の何れか1つとする。

キーバリュアプリケーションの場合、アプリケーション固有アドレス変換ロジックATL(514)はインデックス集積化(index-integrated)フラッシュ変換階層FTLであり得、アプリケーション固有ガーベッジコレクションロジックGC(516)はLRU(least-recently used)リスト(list)を利用することによって更に効率的になる。

HDFSアプリケーションの場合、ATL(514)はメモリ-フットプリントの観点(memory-footprint perspective)からは、例えば細分性(granularity)のために従来一律に必要とされる64MBよりも更に小さく構成され得る。従って、該アプリケーション固有アドレス変換ロジックATL514が必

10

20

30

40

50

須になる。

その上に、幾つかのHDFSアプリケーションの場合、GS516及びWL515は、それが含む、全部又は大部分のデータがコールドデータ（一旦生成された後に全くアップデートされないデータ）である場合、最小化（或いは最適化）され得る。

MongoDBアプリケーションの場合、ホスト動作の一部はATL（514）及びGC（516）にオフロードされる。各アプリケーション511は該アプリケーション511の特性及び要求に適するように相異なるATL（514）、GC（516）、WL（515）を含む。

【0031】

フラッシュ制御器階層535は仮想フラッシュ階層（virtual flash layer、VFL）530及びフラッシュインタフェース階層（FIL）540を含む。フラッシュ制御器階層535はPSDの効率性を向上するハードウェア支援を提供する。即ち、フラッシュ制御器階層535は、PSD上で動作するアプリケーション511に依存しながら、フラッシュメモリに効率的なインタフェースを提供するように再構成及び拡張可能なハードウェアの使用を最適化する。結局、フラッシュ制御器階層535は、アプリケーション511の特徴を支援するようにファームウェアをアップデートすることによってアップデートされる。

【0032】

ホストコンピュータから受信されたパケット（例、データ、メッセージ、及びコマンド）は先ず、PSDカーネル555に伝送される。コンテキスト（CTX）マネージャ556及びPSDカーネル555のメッセージ（MSG）ルータ557は、ホストコンピュータから受信されたパケットをルーティングする役割を遂行する。

CTXマネージャ556は、現在PSDのテナントとなっているアプリケーションをその特性と共に認識（aware）している。そのため、CTXマネージャ556は、現在動作しているアプリケーションと他の現存するアプリケーションの双方をトラッキングするためのアプリケーションID、及びATL（514）、WL（515）、及びGC（516）に関連する構成のテーブルを含む。

また、CTXマネージャ556は、スタート/ストップ（start/stop）、レジューム（resume）、及びポーズ（pause）（別名、サスペンド（suspend））、等の、現在テナントのアプリケーションの施行及びライブネス（liveness）（例、コンテキスト制御）を制御する役割を遂行する。

これらのコマンドは、ベンダーコマンドに由来する場合、PSD制御インタフェース422を通じてPSDカーネル455内のCTXマネージャ556に伝えられる。

PSDカーネル555はルーティングされたパケットのカプセル化を解除し、従ってルーティングされたパケットを処理する。例えば、アプリケーションを“クリエート（生成）”するためのコマンドを受信する場合、CTXマネージャ556はホストから提供されたデータと共にアプリケーション511のアプリケーションコンテキストを発生する。ホストから提供されるデータは、ATL（514）、WL（515）、及びGC（516）のための構成プロファイル或いはロジック、PSD上で実行可能なアプリケーション511のバイナリコード、そしてアプリケーション511がPSDカーネル555と通信可能にするPSDライブラリを含むメタデータを含む。

【0033】

もし、ホストコンピュータから受信されたメッセージがPSDに登録されたアプリケーション511に関連され、登録されたアプリケーション511のコンテキストが活性であれば、PSDカーネル555はメッセージ及び対応するデータ及び/或いはコマンドを登録されたアプリケーション511にルーティングする。もし、登録されたアプリケーションのコンテキストが非活性であれば、PSDカーネル555はホストコンピュータにエラーメッセージをリターンする。もし、登録されないか、或いは認識されないメッセージがホストコンピュータから受信されれば、PSDカーネル555はホストコンピュータに他のエラーメッセージを伝送する。

10

20

30

40

50

【0034】

CTXマネージャ556はアプリケーションコンテナ510を制御する役割を遂行する。CTXマネージャ556はアプリケーション511のためのアプリケーションコンテキストを生成し、PSDのハードウェアリソースをATL(514)、WL(515)、及びGC(516)に分配及び管理する。CTXマネージャ556は仮想マシン(VM)と類似な完全(フル)仮想化方式(fully virtualized manner)、或いは近似的(パラ)仮想化方式(para-virtualized manner)により具現される。

実施形態によっては、アプリケーション511はPSDカーネル555を具現せずに、一般的な(generic)格納装置上で通常の(normal)ファームウェアとして動作する。他の実施形態においては、アプリケーション511は、アプリケーションの特徴の認識能力を備えたPSDに合わせて特別に設計された新しいアプリケーションとして動作する。

【0035】

一部のアプリケーション511は、アプリケーション固有ウェアレベリングWL及び/或いはガーベッジコレクションGCロジックを必要とする。このようなロジックはPSDの性能のために非常に重要である。PSDカーネル555はPSDライブラリ558を提供する。ここで、PSDライブラリ558は、アプリケーション固有WL/GCインタフェース、及びPSDカーネル555とインタフェースするための一般的なAPIを含む。

後者の場合、アプリケーション511は自分自身のWL/GCロジック515、516を具現する代わりに提供された(一般的な)ライブラリを登録する。前者の場合、PSDカーネル555はPSDファームウェアスタック500の階層における全ての権限をアプリケーション511に提供する代わりに、WL/GCに関連する作業を一切遂行しない。

【0036】

アプリケーション511はPSDインタフェース階層552を通じて基底のPSDカーネル555と通信する。PSDインタフェース階層552は通信API及びパケット(例、メッセージ、コマンド、及びデータ)を提供する。ここで、パケットはPSDインタフェース階層552を通じてPSDにパスされる。

【0037】

複数のアプリケーション511は、アプリケーションコンテキストの形態を取って単一のPSDに共存し得る。アプリケーションコンテナ510は活性(active)或いは遊休(idle)状態にあるアプリケーション511のアプリケーションコンテキストを保持する。アプリケーション511はアプリケーションロジック、ホストプロトコルとメッセージプロセッサの階層513、アプリケーション固有のアドレス変換ロジック(ATL)514、ウェアレベリング(WL)ロジック515、ガーベッジコレクション(GC)ロジック516、等を含む。

幾つかの実施形態において、WL及びGCロジック514、515はオプションである。例えば、キーバリューストアアプリケーションは、コアロジックのためのそれらのキーバリューストア構造とアルゴリズムを有し得る。ホストプロトコルとメッセージプロセッサの階層513はコアロジックの機能のラッパー(wrapper)と同様に簡単である。キーバリューストアアプリケーションの場合に、アプリケーション固有のATL(514)は統合(integrated)FTL(例、図2のNoSQL__DB統合(integrated)FTL(262))に合併(merge)される。

【0038】

追加的なロジックは、最近のSSDハードウェアの観点からは強制的(mandatorily)に要求されるわけではないが、次に述べる因子はPSDの実行を加速できる。PSDの実行の加速因子の例としては、以下に限定されないが、(a)多重TLB(multiple translation lookaside buffers、多重の変換ルックアサイドバッファ)或いは更に大きいTLB、(b)アプリケーションコンテナIDを有するキャッシュ、(c)衝突を低減する更に多いバンクを有するメモリ及び/

10

20

30

40

50

或いはDIMM (dual inline memory module) 当たり、更に多いレジスタ、(d) DMA (direct memory access) に対す多い柔軟した提供、そして(e)更に多い個数の現存するDMA、(f)埋め込み型プロセッサから仮想化のための支援、(g)アプリケーションコンテキスト認識フラッシュ制御器、(h)ポリモルフィズム (polymorphism) を敏感に認識するファイングレイン電源制御 (fine-grain power control)、(i)ハードウェア支援メッセージルーティング (hardware-assisted message routing) 等を含む。

他の実施形態において、PSDはアプリケーション固有の機能性を加速して提供するFPGA (field programmable gate array) を含む。機械学習 (machine-learning) カーネル、ビッグデータ処理における科学的計算などのアプリケーションは、伝統的な装置のCPU上でコードを走らせるよりFPGAを用いて加速できる。ここで、ビッグデータ処理は、ニューラルネットワーク (Neural network)、マトリクス操作、FFT (fast fourier transform、高速フーリエ変換)、パターンマッチング、等を含む。

【0039】

一実施形態において、格納システム400はシングルPSD (450) 又は複数のPSD毎に複数のインスタンス (アプリケーション) を提供する。追加に、格納システム400は複数のPSD (450) の間における移動 (migration) 及び複製 (replication) を支援する。

【0040】

図6(A)及び図6(B)は複数のPSDの間におけるアプリケーションの移動及び複製を例示的に示す図面である。複数のPSD (650a~650c) は集中化方式 (centralized manner) 或いは非集中化方式 (decentralized manner) の内の何れか1つの方式によりアプリケーションの移行及び複製を支援する。

【0041】

図6(A)は本発明の実施形態に係る集中化方式によるアプリケーションの移動及び複製を例示する図面である。PSD制御インタフェイス622は複数のPSD (650a~650c) の間における移動及び複製を遂行するためのコマンド及びメッセージ432を提供する。集中化方式の具現の際、移動及び複製動作はPSD (650a~650c) のCTXマネージャ (例、図5のCTXマネージャ556) 及びPSD制御インタフェイス622によって具現される。PSD制御インタフェイス622は集中化された移動及び複製動作を実行するためのPSD外部インタフェイス653を生成 (spawn) する。

【0042】

図6(B)は本発明の実施形態に係る非集中化方式によるアプリケーションの移動及び複製を例示する図面である。非集中化方式による具現の場合、各PSDは隣接する装置の間でピアツーピア (peer-to-peer) 通信に基づいてアプリケーション移動/複製動作を遂行する。

【0043】

本発明の実施形態に係る格納装置は、1つ以上のアプリケーションを含むアプリケーションコンテナと、ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナ内の1つのアプリケーションへ前記複数のパケットをルーティングするポリモルフィック格納装置 (polymorphic storage device、以下、“PSD”) カーネルと、前記1つ以上のアプリケーション及び前記PSDカーネルの間にインタフェイスを提供するPSDインタフェイス階層と、を含む。

【0044】

一実施形態において、前記ホストコンピュータからの制御コマンドに従って再構成可能な1つ以上のハードウェアを更に含み、前記再構成可能なハードウェアは1つ以上のプロ

10

20

30

40

50

セッサ、不揮発性メモリ、及びDRAM (dynamic random access memory) を含む。

【0045】

一実施形態において、PSD制御インタフェイス及び1つ以上のアプリケーションインタフェイスを有するホストサイドインタフェイスが、前記ホストコンピュータに提供される。

【0046】

一実施形態において、前記格納装置は、前記PSD制御インタフェイスを通じて受信された制御コマンドに従って第1種類の格納装置から第2種類の格納装置に変換される。

【0047】

一実施形態において、前記第1種類の格納装置はインストレージコンピューティング装置 (in-storage computing device)、キーバリューストア (key-value store) 装置、及びHadoop分散ファイルシステム (Hadoop distributed file system; HDFS) 装置、オブジェクト-ストア装置の中の何れか1つであり、前記第2種類の格納装置は前記インストレージコンピューティング装置、前記キーバリューストア装置、及び前記HDFS装置の中の他の1つである。

【0048】

一実施形態において、前記PSD制御インタフェイスは前記格納装置に制御コマンド及びメッセージを提供するように構成される。

【0049】

一実施形態において、前記アプリケーションインタフェイスは、バックワード交換性のためにエミュレートされたホストサイドインタフェイス (emulated host-side interface) 及びPSD対応アプリケーションインタフェイス (PSD-aware application interface) を含む。

【0050】

一実施形態において、1つ以上のベンダーコマンド (vendor commands) が、前記PSD制御インタフェイスを通じて提供されることができる。

【0051】

一実施形態において、前記1つ以上のベンダーコマンドは、クリエートコマンド (create command)、スタートコマンド (start command)、ストップコマンド (stop command)、ポーズコマンド (pause command)、及びレジュームコマンド (resume) を含む。

【0052】

一実施形態において、前記1つ以上のベンダーコマンドは、前記PSD制御インタフェイスを通じて前記格納装置で動作するアプリケーションを他の格納装置に移動するための移動コマンド (migration command)、或いは前記格納装置において動作するアプリケーションを他の格納装置において複製するための複製コマンド (duplication command) を含む。

【0053】

一実施形態において、前記アプリケーションは集中化方式 (centralized manner) 或いは非集中化方式 (decentralized manner) により移動されるか、或いは複製される。

【0054】

本発明の実施形態に係る格納システムは、ホストコンピュータと、ポリモルフィック格納装置 (polymorphic storage device、以下、“PSD”) と、前記ホストコンピュータと前記PSDとの間にインタフェイスを提供するPSD外部インタフェイスと、を含み、前記PSDは1つ以上のアプリケーションを含むアプリケーションコンテナと、前記ホストコンピュータからデータ、メッセージ、及びコマンドを有する複数のパケットを受信し、前記複数のパケットに含まれたコマンドに基づいて前記アプ

10

20

30

40

50

リケーションコンテナ内の1つのアプリケーションへ前記複数のパケットをルーティングするPSDカーネルと、前記1つ以上のアプリケーション及び前記PSDカーネルの間にインタフェースを提供するPSDインタフェース階層と、を含む。

【0055】

一実施形態において、前記ホストコンピュータから制御コマンドに従って再構成される1つ以上のハードウェアを更に含み、前記PSDの前記再構成可能なハードウェアは1つ以上のプロセッサ、不揮発性メモリ、及びDRAM(dynamic random access memory)を含む。

【0056】

一実施形態において、PSD制御インタフェース及びアプリケーションインタフェースを有するホストサイドインタフェースが、前記ホストコンピュータに提供される。

10

【0057】

一実施形態において、前記PSDは前記PSD制御インタフェースを通じて受信された制御コマンドに従って第1種類の格納装置から格納装置の第2種類の格納装置に変換される。

【0058】

一実施形態において、前記PSD制御インタフェースは前記PSDに制御コマンド及びメッセージを提供するように具現される。

【0059】

本発明の実施形態に係る格納装置の方法は、前記格納装置のアプリケーションコンテナに1つ以上のアプリケーションを格納する段階と、ホストコンピュータ内の1つのからデータ、メッセージ、及びコマンドを有する複数のパケットを受信する段階と、前記複数のパケットに含まれたコマンドに基づいて前記アプリケーションコンテナからアプリケーションに前記複数のパケットをルーティングする段階と、PSD(polymorphic storage device)制御インタフェースを通じて前記ホストコンピュータから受信された制御コマンドに従って前記格納装置の1つ以上のハードウェアを(再)構成するように前記格納装置でPSDカーネルを駆動する段階と、前記PSDカーネルによって構成されたように前記格納装置を動作する段階と、を含む。

20

【0060】

一実施形態において、前記ホストコンピュータから制御コマンドに従って1つ以上のハードウェアを再構成する段階を更に含み、前記再構成可能なハードウェアは1つ以上のプロセッサ、不揮発性メモリ、及びDRAM(dynamic random access memory)を含む。

30

【0061】

一実施形態において、前記ホストコンピュータにPSD制御インタフェース及びアプリケーションインタフェースを有するホストサイドインタフェースを提供する段階を更に含む。

【0062】

一実施形態において、前記PSD制御インタフェースを通じて受信された制御コマンドに従って第1種類の格納装置から第2種類の格納装置に前記PSDを変換させる段階を更に含む。

40

【0063】

一方、詳述された本発明の内容は発明を実施するための具体的な実施形態に過ぎない。本発明は具体的であり実際に利用できる手段のみならず、将来の技術で活用できる、抽象的であり且つ概念的なアイデアである技術的思想を含む。

【符号の説明】

【0064】

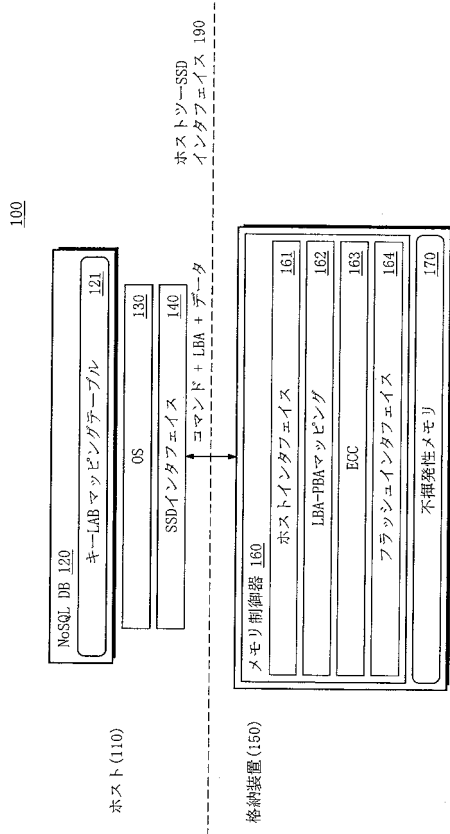
- 100 伝統的な、NoSQL__DBシステム
- 200 他の伝統的な、NoSQL__DBシステム
- 110、210、410 ホスト、ホストコンピュータ

50

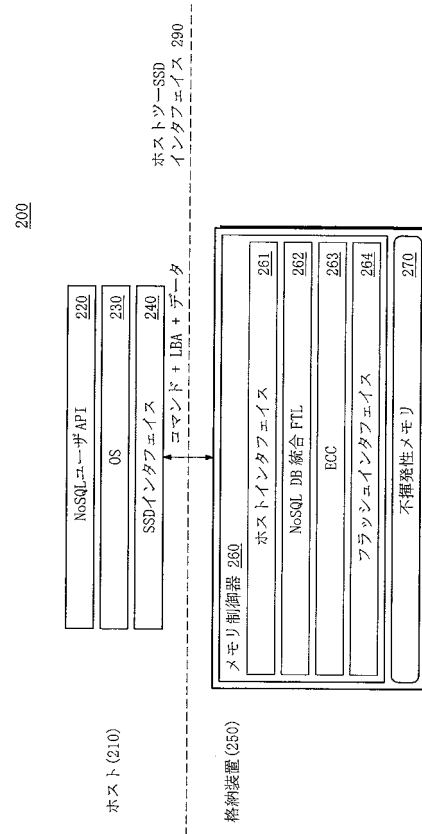
1 2 0	N o S Q L _ _ D B	
2 2 0	N o S Q L _ _ A P I	
1 2 1	キー L B A マッピングテーブル	
1 3 0 , 2 3 0	O S	
1 4 0、2 4 0	S S D インタフェイス	
1 5 0、2 5 0	格納装置	
1 6 0、2 6 0	メモリ制御器	
1 6 1、2 6 1	ホストインタフェイス	
1 6 2	L B A - P B A マッピングテーブル、F T L (フラッシュ変換 (翻訳) 階層)	10
2 6 2	N o S Q L _ _ D B 統合 (i n t e g r a t e d) F T L	
1 6 3、2 6 3	エラー訂正コード (E C C) メモリ	
1 6 4、2 6 4	フラッシュインタフェイス	
1 7 0、2 7 0	不揮発性メモリ	
1 9 0、2 9 0	ホストツース S S D インタフェイス	
3 6 0	メモリ制御器	
3 1 0	H I L (ホストインタフェイス階層)	
3 2 0	F T L (フラッシュ翻訳階層)	
3 3 0	V F L (仮想フラッシュ階層)	
3 4 0	F I L (フラッシュインタフェイス階層)	20
4 0 0	格納システム	
4 1 0	ホストコンピュータ、ホスト	
4 1 1 a ~ 4 1 1 c	装置ドライバ	
4 1 3	装置ドライバプール	
4 1 5	ホストサイドインタフェイス	
4 2 0、4 2 1	第 1、第 2 ユーザインタフェイス	
4 2 2	P S D 制御インタフェイス	
4 3 0	入出力 (I / O) 制御コマンド (例、装置制御コマンド及びメタデータ)	
4 3 1	クエリー (p u t、g e t、等)	
4 3 2	コマンド (再構成、移動 / 複製) 及びメッセージ	30
4 5 0	P S D、ポリモルフィック格納装置	
4 5 1、4 5 1 a ~ 4 5 1 d	アプリケーション	
4 5 2、4 5 2 a ~ d	P I L (P S D インタフェイス階層)	
4 5 5	P S D カーネル	
4 5 6	ハードウェア抽象化階層	
4 6 1、4 6 1 a ~ 4 6 1 d	C P U	
4 8 0	D R A M	
4 9 0	P S D 外部インタフェイス	
5 0 0	P S D ファームウェアスタック	
5 1 0	アプリケーションコンテナ	40
5 1 1	アプリケーション (集合的名称)	
5 1 1 a、5 1 1 b	アプリケーション	
5 1 1 c	アプリケーション、ユーザアプリケーション	
5 1 3	ホストプロトコル及びメッセージプロセッサ	
5 1 4	A T L (アプリケーション固有アドレス変換ロジック)	
5 1 5	W L (アプリケーション固有ウェアレベリング)	
5 1 6	G C (アプリケーション固有ガーベッジコレクションロジック)	
5 3 0	V F L (仮想フラッシュ階層)	
5 3 5	フラッシュ制御器階層	
5 4 0	F I L (フラッシュインタフェイス階層)	50

- 5 5 2 P S D インタフェイス階層
- 5 5 5 P S D カーネル 5 5 6 C T X (コンテキスト) マネージャ
- 5 5 7 M S G (メッセージ) ルータ
- 5 5 8 P S D ライブラリ
- 6 2 2 P S D 制御インタフェイス
- 6 5 3 P S D 外部インタフェイス
- 6 5 0 a ~ 6 5 0 c P S D

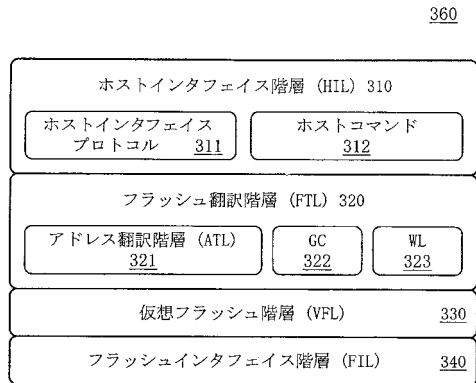
【 図 1 】



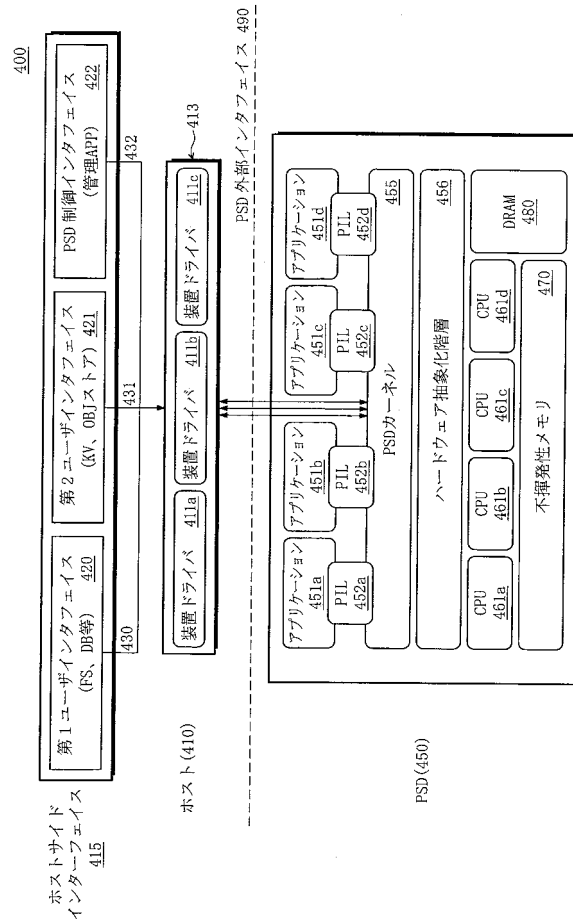
【 図 2 】



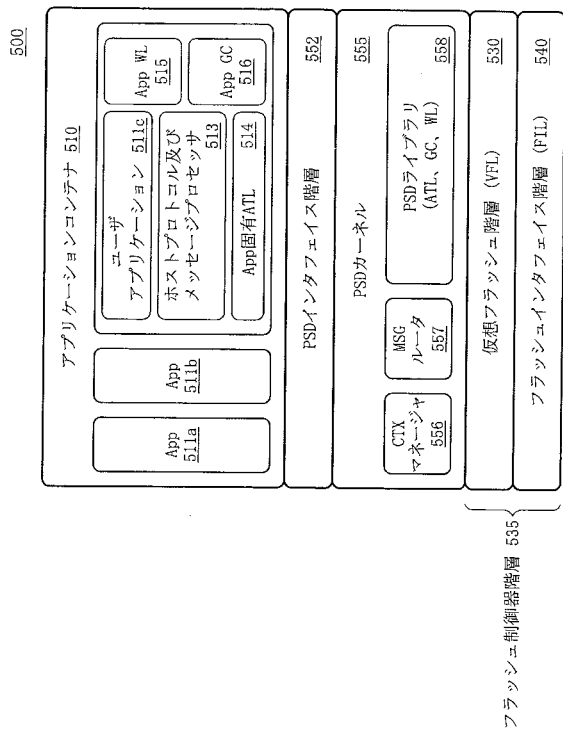
【図3】



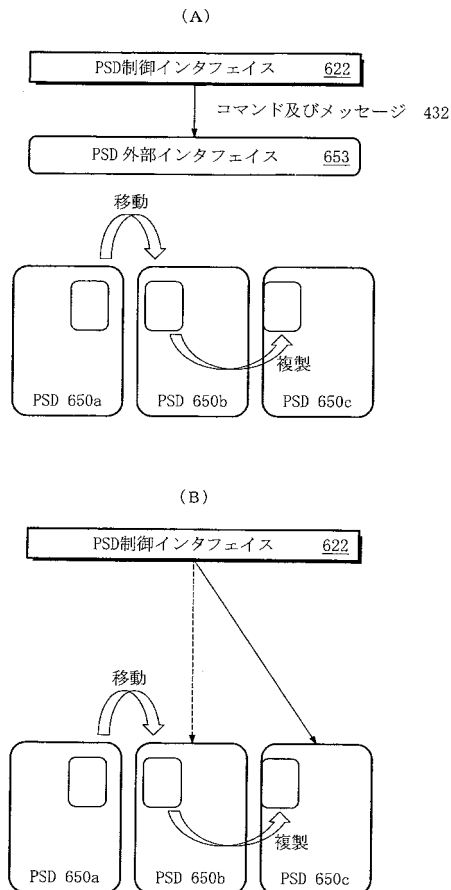
【図4】



【図5】



【図6】



フロントページの続き

(72)発明者 キ, ヤンソク

アメリカ合衆国, カリフォルニア州 94303, パロアルト, アルテア ウォーク 873