

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7586895号
(P7586895)

(45)発行日 令和6年11月19日(2024.11.19)

(24)登録日 令和6年11月11日(2024.11.11)

(51)国際特許分類

F I

G 0 6 F 21/78 (2013.01)

G 0 6 F 21/78

請求項の数 20 (全43頁)

(21)出願番号	特願2022-511208(P2022-511208)	(73)特許権者	314015767
(86)(22)出願日	令和2年6月15日(2020.6.15)		マイクロソフト テクノロジー ライセン
(65)公表番号	特表2022-545012(P2022-545012 A)		シング,エルエルシー
(43)公表日	令和4年10月24日(2022.10.24)		アメリカ合衆国 ワシントン州 9 8 0 5
(86)国際出願番号	PCT/US2020/037681		2 レッドモンド ワン マイクロソフト
(87)国際公開番号	WO2021/034391	(74)代理人	ウェイ
(87)国際公開日	令和3年2月25日(2021.2.25)		100118902
審査請求日	令和5年6月12日(2023.6.12)		弁理士 山本 修
(31)優先権主張番号	16/546,337	(74)代理人	100106208
(32)優先日	令和1年8月21日(2019.8.21)		弁理士 宮前 徹
(33)優先権主張国・地域又は機関	米国(US)	(74)代理人	100196508
			弁理士 松尾 淳一
		(74)代理人	100138759
			弁理士 大房 直樹
		(72)発明者	ブルス,マリク

最終頁に続く

(54)【発明の名称】 メモリー・アパーチャのフラッシュ・オーダーを使用したデーター保存

(57)【特許請求の範囲】

【請求項1】

コンピューティング・システムにおけるデーター保存サブシステムであって、
アパーチャを有するバッテリー・バックアップ・メモリーであって、前記メモリーは物理アドレス空間においてアドレスされるデジタル・データー格納位置を含み、前記アパーチャは前記デジタル・データー格納位置の少なくとも一部分を含み、前記アパーチャは物理アドレスのフラッシュ・オーダーを有し、前記物理アドレスのフラッシュ・オーダーは、データー完全性の脅威に応答して、前記アパーチャからアパーチャ関連不揮発性ストレージにデーターをコピーする順序であり、前記フラッシュ・オーダーは、前記アパーチャのフラッシュ先端物理アドレス(first-to-flush physical address end)および前記アパーチャのフラッシュ後端物理アドレス(last-to-flush physical address end)を定める、バッテリー・バックアップ・メモリーと、
前記バッテリー・バックアップ・メモリーと動作可能に通信するデーター保存回路であって、データー保存ステップを実行するように構成され、前記データー保存ステップは、
(a)データー集合A - データーを前記アパーチャに格納する要求を受けるステップであって、前記A - データーはクリティカル・データーと指定されたデーターを含む、ステップと、
(b)前記アパーチャの未割り当てメモリーの一部A - メモリーを識別するステップであって、前記A - メモリーは前記A - データーを保持できるだけ十分に大きく、前記A - メモリーは、前記アパーチャの他のいずれの未割り当てメモリーの他のいずれの物理

アドレスよりも前記アパーチャのフラッシュ先端物理アドレスに近い物理アドレスを有する、ステップと、

(c) 前記 A - メモリーに割り当て済みの印を付け、前記 A - データのコピーを前記 A - メモリーに入力するステップと、

(d) データ集合 Z - データを前記アパーチャに格納する要求を受けるステップであって、前記 Z - データは、クリティカル・データに指定されたデータを全く含まない、ステップと、

(e) 前記アパーチャの未割り当てメモリーの一部 Z - メモリーを識別するステップであって、前記 Z - メモリーは前記 Z - データを保持できるだけ十分に大きく、前記 Z - メモリーは前記アパーチャの他のいずれの未割り当てメモリーの他いずれの物理アドレスよりも前記アパーチャのフラッシュ後端物理アドレスに近い物理アドレスを有する、ステップと、

10

(f) 前記 Z - メモリーに割り当て済みの印を付け、前記 Z - データのコピーを前記 Z - メモリーに入力するステップとを含む、データ保存回路とを備え、

データ・クリティカル性を考慮せずにデータを前記バッテリー・バックアップ・メモリーに格納することによるよりも、クリティカル・データをフラッシュするのに成功し、これによってそれを保存する可能性を高める、データ保存サブシステム。

20

【請求項 2】

請求項 1 記載のデータ保存サブシステムにおいて、前記バッテリー・バックアップ・メモリー・アパーチャは、ネットワーク・ノード X 内に常駐するか、または、前記ネットワーク・ノード X によって制御され、前記アパーチャに A - データを格納する前記要求は、異なるネットワーク・ノード Y から送られた、データ保存サブシステム。

【請求項 3】

請求項 2 記載のデータ保存サブシステムにおいて、前記ネットワーク・ノード X は、クラウドにおいてストレージ・ノードを含み、前記ネットワーク・ノード Y は前記クラウドにおいて計算ノードを含む、データ保存サブシステム。

【請求項 4】

30

請求項 2 記載のデータ保存サブシステムにおいて、前記ネットワーク・ノード X および前記ネットワーク・ノード Y は、ストレージ関係 R によって関係付けられ、M が N の代わりにデータを格納するとき、2 つのノード M および N が R によって関係付けられるように、R が定められ、R は、前記ネットワーク・ノード X および前記ネットワーク・ノード Y の 1 対 1 の関係ではない、データ保存サブシステム。

【請求項 5】

請求項 1 記載のデータ保存サブシステムにおいて、前記バッテリー・バックアップ・メモリーは N V D I M M メモリーを含む、データ保存サブシステム。

【請求項 6】

請求項 1 記載のデータ保存サブシステムにおいて、前記コンピューティング・システムは仮想機械を備え、前記データ集合 A - データは前記データ保存サブシステムに仮想機械 V M - A から送られ、前記データ集合 Z - データは前記データ保存サブシステムに仮想機械 V M - Z から送られる、データ保存サブシステム。

40

【請求項 7】

コンピューティング・システムにおけるデータ保存方法であって、

データ集合 A - データをバッテリー・バックアップ・メモリーのアパーチャに格納する要求を受けるステップであって、前記メモリーは物理アドレス空間においてアドレスされるデジタル・データ格納位置を含み、前記アパーチャは前記デジタル・データ格納位置の少なくとも一部分を含み、前記アパーチャは物理アドレスのフラッシュ・オーダーを有し、前記物理アドレスのフラッシュ・オーダーは、データ完全性の脅威に応答し

50

て、前記アパーチャからアパーチャ関連不揮発性ストレージにデータをコピーする順序であり、前記フラッシュ・オーダーは、前記アパーチャのフラッシュ先端物理アドレス(first-to-flush physical address end)および前記アパーチャのフラッシュ後端物理アドレス(last-to-flush physical address end)を定め、前記A - データーはクリティカル・データーと指定されたデーターを含む、ステップと、

前記アパーチャの未割り当てメモリーの一部A - メモリーを識別するステップであって、前記A - メモリーは前記A - データーを保持できるだけ十分に大きく、前記A - メモリーは、前記アパーチャの他のいずれの未割り当てメモリーの他のいずれの物理アドレスよりも前記アパーチャのフラッシュ先端物理アドレスに近い物理アドレスを有する、ステップと、

10

前記A - メモリーに割り当て済みの印を付け、前記A - データーのコピーを前記A - メモリーに入力するステップと、

データー集合Z - データーを前記アパーチャに格納する要求を受けるステップであって、前記Z - データーは、クリティカル・データーに指定されたデーターを全く含まない、ステップと、

前記アパーチャの未割り当てメモリーの一部Z - メモリーを識別するステップであって、前記Z - メモリーは前記Z - データーを保持できるだけ十分に大きく、前記Z - メモリーは前記アパーチャの他のいずれの未割り当てメモリーの他いずれの物理アドレスよりも前記アパーチャのフラッシュ後端物理アドレスに近い物理アドレスを有する、ステップと、

20

前記Z - メモリーに割り当て済みの印を付け、前記Z - データーのコピーを前記Z - メモリーに入力するステップと

を含み、

前記フラッシュ・オーダーにおいて前記クリティカル・データーを非クリティカル・データーよりも前に前記バッテリー・バックアップ・メモリーに入力することによって、クリティカル・データーをフラッシュするのに成功し、これによってそれを保存する可能性を高める、方法。

【請求項 8】

請求項 7 記載の方法であって、

少なくとも部分的に前記バッテリー・バックアップ・メモリーのバッテリー特性に基づいて前記アパーチャのサイズを決定するステップと、

30

少なくとも部分的に前記バッテリー・バックアップ・メモリーのバッテリー特性に基づいて、前記アパーチャの位置を決定するステップと、

少なくとも部分的に前記バッテリー・バックアップ・メモリーのバッテリー特性における変化に基づいて、前記アパーチャの大きさを決定し直すステップと、

少なくとも部分的に前記バッテリー・バックアップ・メモリーのバッテリー特性における変化に基づいて、前記アパーチャを移動させるステップと

のうちの少なくとも 1 つを更に含む方法。

【請求項 9】

請求項 7 記載の方法であって、前記アパーチャの少なくともディフラグ部分をフラッシュ・ディフラグするステップを更に含み、前記ディフラグ部分はクリティカル・データーを含む、方法。

40

【請求項 10】

請求項 7 記載の方法であって、ワークロードのクリティカル性、仮想機械優先度、または、カスタマー・ステータスという判断基準のうちの少なくとも 1 つに基づいて、前記A - データーの少なくとも一部分をクリティカル・データーに指定するステップを更に含む方法。

【請求項 11】

請求項 7 記載の方法であって、

前記データー完全性の脅威を検出するステップと、

50

前記クリティカル・データーの全てを前記アパーチャから前記アパーチャ関連不揮発性ストレージにフラッシュするステップと、

少なくとも一部の非クリティカル・データーを前記アパーチャから前記アパーチャ関連不揮発性ストレージにフラッシュし損なうステップとを更に含む方法。

【請求項 1 2】

請求項 7 記載の方法であって、

前記データー完全性の脅威を検出するステップと、

前記クリティカル・データーの少なくとも全てを前記アパーチャから前記アパーチャ関連不揮発性ストレージにフラッシュするステップと、

その後、フラッシュした前記データーの全てを、前記アパーチャ関連不揮発性ストレージから、揮発性メモリーと、前記データーがフラッシュされた元である前記バッテリー・バックアップ・メモリーと、前記データーがフラッシュされた元である前記バッテリー・バックアップ・メモリーとは異なるバッテリー・バックアップ・メモリーとのうちの少なくとも 1 つに復元するステップと

を更に含む方法。

【請求項 1 3】

請求項 7 記載の方法であって、前記 Z - データーの少なくとも 1 つのコピーを、前記アパーチャの外部および前記アパーチャ関連不揮発性ストレージの外部に保持するステップを更に含む方法。

【請求項 1 4】

請求項 7 記載の方法において、前記バッテリー・バックアップ・メモリーのアパーチャは、ここでは M で示す物理機械内に常駐し、前記アパーチャに前記 A - データーを格納する要求は、ここでは N で示す異なる物理機械から送られた、方法。

【請求項 1 5】

請求項 7 記載の方法であって、前記 A - データーを 1 つのストレージ・ノードのみに格納し、前記 Z - データーのコピーを複数のストレージ・ノードに格納するステップを含む方法。

【請求項 1 6】

データーおよび命令を含むよう構成されたコンピューター読み取り可能記憶媒体であって、前記命令は、実行されると、データー保存サブシステムに、コンピューティング・システムにおけるデーター保存のための方法を実行させ、前記方法は、

複数の要求を受けるステップであって、各要求は、それぞれのデーター集合のバッテリー・バックアップ・メモリーのアパーチャにおける格納を求め、前記メモリーは物理アドレス空間においてアドレスされるデジタル・データー格納位置を含み、前記アパーチャは前記デジタル・データー格納位置の少なくとも一部分を含み、前記アパーチャは物理アドレスのフラッシュ・オーダーを有し、前記物理アドレスのフラッシュ・オーダーは、データー完全性の脅威に応答して、前記アパーチャからアパーチャ関連不揮発性ストレージにデーターをコピーする順序であり、前記物理アドレスのフラッシュ・オーダーは、前記アパーチャのフラッシュ先端物理アドレス(first-to-flush physical address end)および前記アパーチャのフラッシュ後端物理アドレス(last-to-flush physical address end)を定め、それぞれのデーター集合の各々は、クリティカル・データーと指定されたデーターを含む、ステップと、

前記要求のうちの少なくとも 2 つについて、前記アパーチャの未割り当てメモリーのそれぞれの部分を識別するステップであって、未割り当てメモリーの各部分は、前記それぞれのデーター集合を保持できるだけ十分に大きく、未割り当てメモリーの識別された前記それぞれの部分は、前記アパーチャの他のいずれの未割り当てメモリーの他のいずれの物理アドレスよりも前記アパーチャのフラッシュ先端物理アドレスに近い物理アドレスを有する、ステップと、

前記要求のうちの少なくとも 1 つについて、未割り当てメモリーの識別された前記それ

10

20

30

40

50

それぞれの部分に割り当て済みの印を付け、それに前記それぞれのデータ集合のコピーを入力するステップと、

前記データ完全性の脅威を検出するステップと、

前記アパーチャにコピーされた前記クリティカル・データの全てを前記アパーチャから前記アパーチャ関連不揮発性ストレージにフラッシュするステップと

を含み、これによって、前記方法は、前記データ完全性の脅威にもかかわらず、前記アパーチャにコピーされた前記クリティカル・データの全てを保存する、コンピューター読み取り可能記憶媒体。

【請求項 17】

請求項 16 に記載のコンピューター読み取り可能記憶媒体において、前記方法は、アパーチャ・サイズとアパーチャ・アドレスとのうちの少なくとも一方を特定するステップを更に含み、特定する前記ステップは、バッテリー容量と、バッテリーの経年数と、バッテリーの信頼性値とのうちの少なくとも 1 つに基づく、コンピューター読み取り可能記憶媒体。

10

【請求項 18】

請求項 16 に記載のコンピューター読み取り可能記憶媒体において、フラッシュする前記ステップは、統一拡張可能ファームウェア・インターフェースを有するファームウェアを実行するステップを含む、コンピューター読み取り可能記憶媒体。

【請求項 19】

請求項 16 に記載のコンピューター読み取り可能記憶媒体において、前記方法は、クリティカル・データに指定されていないデータを保持するために、前記アパーチャの少なくとも 10 パーセントを予約するステップを更に含む、コンピューター読み取り可能記憶媒体。

20

【請求項 20】

請求項 16 に記載のコンピューター読み取り可能記憶媒体において、前記方法は、

クリティカル・データと指定されたデータのための空間を作るために、クリティカル・データと指定されていないデータを、前記アパーチャの外部の格納位置に移動するステップと、

前記アパーチャの少なくとも一部分をディフラグし、これによって、前記アパーチャの 2 つの割り当て部分の間に位置していた未割り当て部分を削減するか又は消滅させるステップと

30

のうちの少なくとも 1 つを更に含む、コンピューター読み取り可能記憶媒体。

【発明の詳細な説明】

【背景技術】

【0001】

【0001】 システム・アーキテクト、およびコンピューティング・システムを設計、実装、修正、または最適化する他の人々は、データ・ストレージの選択が多くの場合トレードオフを伴うことを認識している。例えば、一部のダイナミック・ランダム・アクセス・メモリー・デバイスのような、デジタル・データを格納することができるデバイスの中には、比較的速いデータ格納動作をサポートするが、停電の場合データを信頼性高く保持できないものがある。これらのデバイスを一般に「揮発性」記憶デバイスと呼ぶ。容量、コスト、期待動作寿命(working life)、動作速度、交換の難度、電氣的要件、およびハードウェアまたはソフトウェア規格もしくはベスト・プラクティスとの互換性というような、異なる技術的特性を有する揮発性記憶デバイスが多数存在する。電気機械式ハード・ディスク・ドライブおよびソリッド・ステート・ドライブというような他のデータ記憶デバイスは、電源を遮断した後でも、データ値を信頼性高く保持する。これらのデバイスを一般に「不揮発性」記憶デバイスと呼ぶ。この場合もまた、異なる技術的特性を有する不揮発性記憶デバイスが多く存在する。

40

【0002】

【0002】 したがって、システム・アーキテクトおよび入手可能な記憶メカニズムと手順との間で選択しなければならない他の人々は、過度に多い数の記憶デバイスの選択肢に直

50

面し、技術の選択肢とシステム性能、コスト、利便性、信頼性、および他の特性との間には、多くの相互依存するトレードオフがある。したがって、複雑さを減らし、予測可能性を高めるためには、特定の性能仮定(performance assumption)、目標、または知見に的を絞り、ストレージ・アーキテクチャの選択肢を狭めて優先順位を付けることが有用であろうが、真剣な検討が必要となる。

【発明の概要】

【0003】

[0003] 本文書において説明する実施形態の中には、特に、仮想機械を実行するクラウドのような、ネットワーク型マルチモード環境において、改良されたデータ保存ツールおよび技法を提供するものがある。ある実施形態では、バッテリー・バックアップされたメモリ・アパーチャのフラッシュ・オーダーに基づいて、更にバッテリー特性についての関連情報に基づいて、メモリ割り当てを行うことによって、クリティカル・データを失う危険性を減らす。このようなストレージ割り当ても、データの完全性を損なうことなく、データ冗長性要件を低下させることができる。

10

【0004】

[0004] ある実施形態では、コンピューティング・システムは、アパーチャを有しバッテリー・バックアップされたメモリを含むデータ保存サブシステムを有する。アパーチャはフラッシュ・オーダーを有する。フラッシュ・オーダーとは、停電またはリブートのようなデータ完全性の脅威にตอบสนองして、アパーチャからアパーチャ関連不揮発性ストレージ(aperture-associated non-volatile storage)にデータをコピーする順序である。フラッシュ・オーダーは、アパーチャのフラッシュ先端(first-to-flush end)およびアパーチャのフラッシュ後端(last-to-flush end)を定める。

20

【0005】

[0005] この例では、データ保存回路が、動作可能にバッテリー・バックアップ・メモリと通信する。データ保存回路は、プロセッサとファームウェアとを含み、データ保存ステップを実行するように構成することができる。この例では、これらのデータ保存ステップは、以下を含んでもよい。

(a) データ集合A - データをアパーチャに格納する要求を受けるステップ。A - データはクリティカル・データと指定されたデータを含む。(b) アパーチャの未割り当てメモリの一部分A - メモリを識別するステップ。A - メモリはA - データを保持できるだけ十分に大きく、A - メモリは、アパーチャの他のいずれの未割り当てメモリの他のいずれのアドレスよりもアパーチャのフラッシュ先端に近いアドレスを有する。(c) A - メモリに割り当て済みの印を付け、A - データのコピーをA - メモリに入力するステップ。これらのステップ(a) ~ (c)は、クリティカル・データを収容するそれぞれの異なるデータ集合の分だけ、複数回繰り返すことができる。

30

【0006】

[0006] この例では、データ保存ステップは、以下も含んでもよい。

(d) データ集合Z - データをアパーチャに格納する要求を受けるステップ。Z - データは、クリティカル・データに指定されたデータを全く含まない。(e) アパーチャの未割り当てメモリの一部分Z - メモリを識別するステップ。Z - メモリはZ - データを保持できるだけ十分に大きく、Z - メモリは、アパーチャの他のいずれの未割り当てメモリの他のいずれのアドレスよりもアパーチャのフラッシュ後端に近いアドレスを有する。(f) Z - メモリに割り当て済みの印を付け、Z - データのコピーをZ - メモリに入力するステップ。これらのステップ(d) ~ (f)は、クリティカル・データを収容しないそれぞれの異なるデータ集合の分だけ繰り返すことができ、クリティカル・データだけが格納されているときは省略してもよい。

40

【0007】

[0007] ラベル(a) ~ (f)は、ここでは単なる識別子として役割を果たすに過ぎず、必ずしも動作の順序を指定する訳ではない。実例をあげると、いずれのA - データを格納する前でも、Z - データをアパーチャに格納してもよい。

50

【 0 0 0 8 】

[0008] このコンテキストでは、データ保存サブシステムの一例は、データ・クリティカル性を考慮せずにデータをバッテリー・バックアップ・メモリーに格納するのではなく、フラッシュ・オーダーにおいて、クリティカル・データを非クリティカル・データに先立ってバッテリー・バックアップ・メモリーに入力することによって、クリティカル・データをフラッシュするのに成功し、これによってそれを保存する可能性(likelihood)を高める。脅威を受けたコンピューティング・システムの外部にクリティカル・データの複製または他の冗長コピーを維持するというような、データ保存の代替形態は、本明細書において説明するデータ保存サブシステムの能力(capabilities)を考慮すれば、削減するまたは不要にすることができる。

10

【 0 0 0 9 】

[0009] 動作において、本明細書において説明するデータ保存のある実施形態は、複数の要求を受け、各要求が、バッテリー・バックアップ・メモリーのアパーチャ内へのそれぞれのデータ集合の格納を求める。アパーチャはフラッシュ・オーダーを有する。フラッシュ・オーダーとは、データ完全性の脅威に応答して、アパーチャからアパーチャ関連不揮発性ストレージにデータをコピーする順序である。フラッシュ・オーダーは、アパーチャのフラッシュ先端とアパーチャのフラッシュ後端とを定める。この例では、それぞれのデータ集合は、各々、クリティカル・データに指定されたデータを含む。要求の内少なくとも2つについて、この実施形態は、アパーチャの未割り当てメモリーのそれぞれの部分を識別し、未割り当てメモリーの各部分は、それぞれのデータ集合を保持できるだけ十分に大きい。識別された未割り当てメモリーのそれぞれの部分は、アパーチャの他のいずれの未割り当てメモリーの他いずれのアドレスよりもアパーチャのフラッシュ先端に近いアドレスを有する。要求の内少なくとも1つについて、この実施形態は、識別した未割り当てメモリーのそれぞれの部分に割り当て済みの印を付け、その中にそれぞれのデータ集合のコピーを入力する。

20

【 0 0 1 0 】

[0010] この例を続けると、少なくとも1つのクリティカル・データを格納した後のある時点において、この実施形態はデータ完全性の脅威を検出する。例えば、近づきつつあるリブート、または差し迫った外部電源の停電と言った、結果的にバッテリー電力への切り替えが起こる脅威を検出する。この脅威の検出に応答して、この実施形態は、アパーチャにコピーされたクリティカル・データの全てをフラッシュすることによって、そのクリティカル・データをアパーチャから不揮発性ストレージにコピーする。このフラッシュによって、アパーチャにコピーされたクリティカル・データの全てを保存し、したがって、アパーチャの範囲外にデータの冗長コピーを有することにも、その不揮発性バックアップ・メモリーにも全く頼ることなく、クリティカル・データを保存する(そしてできる限り非クリティカル・データの一部または全部も保存する)。

30

【 0 0 1 1 】

[0011] 本明細書における教示に関連する他の技術的活動および特性も、当業者には明白になるであろう。示した例は単なる例示に過ぎない。この摘要は、特許請求する主題の主要な特徴または必須の特徴を特定することを意図するのではなく、特許請求する主題の範囲を限定するために使用することを意図するのでもない。逆に、この摘要は、詳細な説明において以下で更に説明するいくつかの技術的概念を、簡略化した形態で紹介するために設けられている。技術革新(innovation)は、請求項によって適正に理解される通りに定められ、この摘要が特許請求の範囲と矛盾する場合、特許請求の範囲を優先する。

40

【 図面の簡単な説明 】

【 0 0 1 2 】

[0012] 添付図面を参照しながら、更に特定のな説明を行う。これらの図面は、選択した態様を図示するに過ぎず、したがって適用範囲即ち特許請求の範囲を余すところなく断定するのではない。

【 図 1 】 コンピューター・システム全体を示し、更に構成された記憶媒体全体も示すプロ

50

ック図である。

【図 2】計算ノードと少なくとも 1 つのストレージ・ノードとを含む環境を示すブロック図である。

【図 3】あるコンピューティング環境のある態様を示すブロック図である。

【図 4】データ保存機能が構成されたシステムの態様を示すブロック図である。

【図 5】バッテリー特性のいくつかの例を示すブロック図である。

【図 6】データ・クリティカル性の判断基準のいくつかの例を示すブロック図である。

【図 7】あるコンピューティング・システムのある例を示すブロック図である。

【図 8】様々な仮想機械を有し(bear)、バッテリー・バックアップ・メモリー・アパーチャを有する様々なストレージ・ノードと共にネットワークを組む計算ノードを示すアーキテクチャ図である。

10

【図 9】様々な仮想機械を実行し、ストレージ・バックアップ・メモリー・アパーチャ(storage-backed memory aperture)を使用するオペレーティング・システムを示すアーキテクチャ図である。

【図 10】あるデータ保存方法におけるステップを示すフローチャートである。

【図 11】あるデータ保存方法におけるステップを更に示すフローチャートである。

【発明を実施するための形態】

【0013】

[0024] 全体像

[0025] 技術革新はそれらの原点を超えて拡大することができるが、技術革新の原点を理解すると、一層深くその技術革新の真意を認めるのに役立つことができる。本願の場合、本明細書において説明する教示は、機械学習用のビッグ・データを格納するという技術的課題が動機となった。もののインターネット・デバイスは、データを大量にそして高速(rapid rate)で産出し(pump out)、機械学習ツールおよび技法は、このデータの洪水を有益な方法で処理できる場合もある。ある機械学習データベースは、何百万ものレコードを実際のデータベース内に含み、したがって高い読み取り速度が重要となる。大量のデータが関与するにも拘わらず、個々のデータに対する不正な意図的変更は望ましくないという意味、そしてリブートまたは予期されない停電によって起こるデータ変化は望ましくないという意味の双方で、データの完全性も重要である。

20

【0014】

[0026] また、この動機付けになる状況(motivational situation)では、他の要因も役割を果たす。具体的には、コストについての懸念が、全てのデータを N V D I M M または同様のバッテリー・バックアップされた高速ランダム・アクセス・メモリーに単純に入力することを諦めさせるか、または躊躇させる。全てのデータを N V D I M M に入力することが可能であっても、バッテリーの故障によって生ずるデータ完全性に対する危険性には対応しない(address)。バッテリー・バックアップは、例えば、ハード・ディスクまたは D V D またはテープ、若しくは何らかの同様の不揮発性媒体上における格納信頼性は高くない。バッテリーの寿命は、時間の経過と共に変化する。安全性向上のためにデータを複数の N V D I M M デバイスに跨がって複製することも考えられるが、この場合もコストが重要な要因となる。更にまた、データ完全性に対する容認可能な危険性のレベルは、常に一定であるという訳ではない。例えば、カスタマーが異なれば、異なるサービス品質保証(S L A : Service Level Agreement)または他のサービス目標もしくは保証を求める(have)のはもっともである。

30

【0015】

[0027] 端的に言うと、本明細書において説明するいくつかのデータ保存ツールおよび技法は、大量の機械学習データの集合を保存する取り組みにおいて表出した技術的課題がある程度動機付けになった。しかしながら、本明細書において提供する教示は、多くの他の技術的シナリオにも同様に、有益な応用可能性を有することは、当業者には認められよう。

40

【0016】

50

【0028】 本明細書において説明するある実施形態は、完全性の脅威に対抗してデータを保存するために、動作ステップとデバイス特性との組み合わせを実装する。カスタマー要件、ワークロード・クリティカル性、または仮想機械のクリティカル性というような判断基準に基づいて、データをクリティカル・データおよび非クリティカル・データに分割する。クリティカル・データは、バッテリー・バックアップ・メモリーの物理アドレスに格納され、非クリティカル・データに先立ってフラッシュされる。ある実施形態では、バッテリーの電力が、クリティカル・データおよび非クリティカル・データの双方を安全にフラッシュするのに十分でない危険性を認識し、例えば、非クリティカル・データが時間内にフラッシュされない場合に備えて、非クリティカル・データの冗長コピーを保持することによって、この危険性に取り組む。バッテリー・バックアップ・メモリーのオーバーチャージは、バッテリーの特性にしたがって大きさが決められ、条件が変化するに連れて、大きさを変更することもできる。オーバーチャージの使用を最適化するために、特に、クリティカル・データを保持する部分内において、ディフラグging(defragging)を実行する。革新的なデータ保存の他の態様についても、本明細書において論ずる。

【0017】

【0029】 本明細書において説明する実施形態は、人によっては、より広いコンテキストで見ることにもできる。実例をあげると、割り当て、データ、順序、電力、保存、および要求というような概念は、特定の実施形態に関連すると見なすこともできる。しかしながら、広いコンテキストの利用可能性から、本明細書において抽象的な理念に対して排他的権利を求める訳ではなく、これらを求めてはいない。むしろ、本開示は、NVDIMMのコストをサービス品質保証データ可用性要件に対してどのように釣り合わせるかというような、特定の技術的問題を完全にまたは部分的に解決する技術的效果を有するように、しかるべく具体化した実施形態を提供することに絞る。他の構成された記憶媒体、システム、および割り当て、データ、オーダー、電力、保存、または要求を伴うプロセスは、本願の範囲には入らない。したがって、本開示を正しく理解することによって、曖昧さ、単なる抽象性、技術的特徴(character)の欠如、および付随する証明問題も回避される。

【0018】

【0030】 更に一般的には、本開示のあらゆる部分、または本明細書におけるあらゆる特定の詳細が、必ずしも、実施可能性、書面による説明、または最良の態様というような法的判断基準を満たさなければならないのではないことは、当業者には認められよう。また、実施形態は、特定の動機付けになる例、要求、応答、オペレーティング・システムまたは他のカーネル、ソフトウェア開発環境、インターフェース規格、ソフトウェア・プロセス、開発ツール、識別子、ファイル、データ構造、表記、制御フロー、擬似コード、命名規則、ノード・アーキテクチャ、または本明細書において説明する他の実施態様選択肢には限定されない。他のいずれの特許開示との明らかな矛盾は、本技術革新の所有者からであっても、本開示において提示する特許請求の範囲を解釈するには全く役に立たない。

【0019】

【0031】 技術的特徴

【0032】 本明細書において説明する実施形態の技術的特長は、当業者には明白であろう。そして、広い範囲の注意深い読者にも幾通りも方法で明白になるであろう。ある実施形態は、メモリー割り当て、揮発性ストレージから不揮発性ストレージへのデータのフラッシュ、バッテリー特性に基づくメモリー・オーバーチャージのサイズ決定またはサイズ決定し直し、およびメモリーのディフラグgingというような、各々、コンピューティング技術に深く根付いた技術的活動に取り組む。論ずる技術的メカニズムの一部には、例えば、NVDIMMまたは他のバッテリー・バックアップ・メモリー、統一拡張可能ファームウェア・インターフェース(Unified Extensible Firmware Interface)、計算ノード、ストレージ・ノード、メモリー・マップ、フラッシュ・オーダー、および仮想機械が含まれる。論ずる技術的效果の一部には、例えば、クリティカル・データが完全性の脅威に対抗して保存される可能性を高めること、データ保存を行うためのデータ冗長性依拠を減らすこと、バッテリー特性および非クリティカル・データのために確保するメモリー量

というような判断基準を考慮してバッテリー・バックアップ・メモリーのアパーチャのサイズを効率的に決定することが含まれる。つまり、純粋に心理的なプロセスは、明確に除外される。ある実施形態は、N V D I M Mコスト、カスタマー・サービス品質保証要件、およびデータ冗長性のためのノード・ストレージの可用性というような判断基準の均衡を取りながら、完全性の脅威に対抗してデータを保存することによって、コンピューティング・システムおよびサービスの機能(functioning)を改善する。本教示の技術特性に基づく他の利点も、提示する説明から、当業者には明白になるであろう。

【 0 0 2 0 】

[0033] 頭字語、省略、名称、および記号

[0034] 一部の頭字語、省略、名称、および記号について、以下に定義する。他のものは、本明細書の他の場所で定義されるか、または当業者が理解するために本明細書における定義が必要ないものである。

【 0 0 2 1 】

[0035] A C P I : 高度構成および電力インターフェース

[0036] A L U : 算術演算および論理ユニット

[0037] A P I : アプリケーション・プログラム・インターフェース

[0038] B I O S : 基本入力/出力システム

[0039] B M C : ベースボード管理コントローラ

[0040] C D : コンパクト・ディスク

[0041] C P U : 中央処理ユニット

[0042] E F I : 拡張可能ファームウェア・インターフェース

[0043] D R A M : ダイナミック・ランダム・アクセス・メモリー

[0044] D V D : デジタル・バーサタイル・ディスクまたはデジタル・ビデオ・ディスク

[0045] F P G A : フィールド・プログラマブル・ゲート・アレイ

[0046] F P U : 浮動小数点処理ユニット

[0047] G P U : グラフィカル処理ユニット

[0048] G U I : グラフィカル・ユーザー・インターフェース

[0049] H D D : ハード・ディスク・ドライブ (例えば、ソリッド・ステート、電気機械、光学)

[0050] I a a S または I A A S : サービスとしてのインフラストラクチャ

[0051] I D : 識別または識別情報(identity)

[0052] I o T : もののインターネット

[0053] L A N : ローカル・エリア・ネットワーク

[0054] N V D I M M : 不揮発性デュアル・インライン・メモリー・モジュール

[0055] N V M e : 不揮発性メモリー・エクスプレス

[0056] O S : オペレーティング・システム

[0057] P a a S または P A A S : サービスとしてのプラットフォーム

[0058] R A M : ランダム・アクセス・メモリー

[0059] R O M : リード・オンリー・メモリー

[0060] S A T A : シリアル A T A (コンピューター・バス・インターフェース)

[0061] S L A : サービス品質保証

[0062] S M M : システム管理モード

[0063] T C P / I P : 送信制御プロトコル/インターネット・プロトコル

[0064] U E F I : 統一拡張可能ファームウェア・インターフェース

[0065] V M : 仮想機械

[0066] W A N : ワイド・エリア・ネットワーク

[0067] いくつかの追加の用語

【 0 0 2 2 】

[0068] 本明細書では、図面に示すような例示的实施形態(exemplary embodiment)

を参照し、これらを説明するために、本明細書では、具体的な文言を使用する。しかし、本明細書において例示される特徴の変更(alteration)および更に他の修正(modification)、ならびに本明細書における特定実施形態によって例示される抽象的原理の他の技術的応用も、関連技術(1つまたは複数)に習熟し本開示を入手できるものには想起され、特許請求の範囲内に該当すると見なされて当然である。

【0023】

【0069】本開示において、用語の意味が明確化されるので、特許請求の範囲は、これらの明確化に対して細心の注意を払って、読解されてしかるべきである。具体的な例をあげるが、関連技術(1つまたは複数)に習熟する者には、他の例も、使用される用語の意味に当てはまり、1つ以上の請求項の範囲内に該当する場合もあることが理解されよう。用語は、本明細書では、これらが一般的な使用(特に、技術以外での使用)において有する意味、特定業界の使用における意味、または特定の辞書もしくは1組の辞書において有する意味とは、必ずしも同じ意味を有するとは限らない。用語の深さを示すのに役立つように、種々の言葉使いと共に参照番号を使用する場合もある。所与のテキストにおいて参照番号が抜けていても、必ずしも、図の内容がそのテキストによって論じられていないことを、意味するのではない。本発明者は、特定のそして選択された辞書編纂に対する権利を主張し行使する。引用符合が付いた用語は、明示的に定義されているが、引用符合を使用せずに、用語を暗示的に定義する場合もある。本明細書の詳細な説明においておよび/または出願申請書内のどこかで、明示的または暗示的に用語を定義することもできる。

【0024】

【0070】本明細書において使用する場合、「コンピューター・システム」(computer system)(別名「コンピューティング・システム」(computing system))は、例えば、1つ以上のサーバー、マザーボード、処理ノード、ラップトップ、タブレット、パーソナル・コンピューター(携帯用またはそれ以外)、パーソナル・デジタル・アシスタント、スマートフォン、スマートウォッチ、スマートバンド、携帯電話または移動体電話、少なくともプロセッサとメモリーとを有する他の移動体デバイス、ビデオ・ゲーム・システム、拡張現実システム、ホログラフィ投影システム、テレビジョン、ウェアラブル・コンピューティング・システム、および/または少なくとも部分的に命令によって制御される1つ以上のプロセッサを備えた(provide)他のデバイス(1つまたは複数)を含むことができる。命令は、ファームウェアの形態、あるいはメモリーおよび/または特殊回路における他のソフトウェアの形態であってもよい。

【0025】

【0071】「マルチスレッド型」(multithreaded)コンピューター・システムとは、複数の実行スレッドをサポートするコンピューター・システムのことである。「スレッド」(thread)という用語は、スケジューリングが可能なコード、またはスケジューリングおよびおそらく同期を受けるコードを含むように理解されてしかるべきである。スレッドは、本開示以外では、例えば、「タスク」(task)、「プロセス」(process)、または「ルーチン」(coroutine)のような、他の名称で知られている場合もある。しかしながら、本明細書では、スレッドはプロセス内部において実行パスを定義するとして、スレッドとプロセスとの間で区別することとする。また、プロセスのスレッドは、所与のアドレス空間を共有するが、異なるプロセスは異なるそれぞれのアドレス空間を有する。プロセスのスレッドは、並列に、順次、または並列実行および順次実行の組み合わせで(例えば、タイムスライス状)実行することができる。

【0026】

【0072】「プロセッサ」(processor)とは、同時マルチスレッディング実施態様におけるコアのような、スレッド処理ユニットである。プロセッサは、ハードウェアを含む。所与のチップは、1つ以上のプロセッサを保持することができる。プロセッサは汎用でもよく、またはベクトル処理、グラフィクス処理、信号処理、浮動小数点算術処理、暗号化、I/O処理、機械学習等のような、特定の使用に合わせて個別に製造することもできる。

10

20

30

40

50

【 0 0 2 7 】

【0073】 「カーネル」(kernel)は、オペレーティング・システム、ハイパーバイザー、仮想機械、BIOSまたはUEFIコード、および同様のハードウェア・インターフェース・ソフトウェアを含む。

【 0 0 2 8 】

【0074】 「コード」(code)とは、プロセッサ命令、データ（定数、変数、およびデータ構造を含む）、または命令およびデータの双方を意味する。「コード」および「ソフトウェア」は、本明細書では相互交換可能に使用される。実行可能コード、インタプリタ・コード、およびファームウェアはコードのいくつかの例である。

【 0 0 2 9 】

【0075】 「プログラム」(program)は、本明細書では広く使用され、アプリケーション、カーネル、ドライバ、割り込みハンドラ、ファームウェア、状態機械、ライブラリ、ならびにプログラマー（開発者とも呼ばれる）によって書かれたたおよび/または自動的に生成された他のコードを含む。

【 0 0 3 0 】

【0076】 「サービス」(service)とは、クラウド・コンピューティング環境、あるいは他のネットワークまたはコンピューティング・システム環境における、消費可能なプログラムの提供を意味し、リソースまたはリソースへのアクセスを複数のプログラムに提供する。

【 0 0 3 1 】

【0077】 「クラウド」(cloud)とは、従量制オンデマンド・サービスに柔軟に利用可能な、コンピューティング、ストレージ、およびネットワーキングのために蓄えられた(pooled)リソースを意味する。クラウドは、プライベート、パブリック、コミュニティ、またはハイブリッドでもよく、クラウド・サービスは、サービスとしてのインフラストラクチャ(IaaS)、サービスとしてのプラットフォーム(PaaS)、サービスとしてのソフトウェア(SaaS)、または他のサービスの形態で提供することができる。特に明記しない限り、ファイルからの読み取りまたはファイルへの書き込みについての論述はいずれも、ローカル・ファイルの読み取り/書き込み、またはクラウド・ネットワークもしくは他のネットワークでもよいネットワークを通じた読み出し/書き込み、あるいは双方(ローカルおよびネットワークを通じた読み取り/書き込み)を行うことも含んでもよい。

【 0 0 3 2 】

【0078】 「IoT」即ち「もののインターネット」とは、アドレス可能な埋め込みコンピューティング・ノードのあらゆるネットワーク状集合体を意味する。このようなノードは、本明細書において定義するようなコンピューター・システムの例であるが、これらは以下の特性の内少なくとも2つも有する。(a)ローカルでない、人が読み取ることができるディスプレイ、(b)ローカルでないキーボード、(c)入力の主要なソースが、非言語データのソースを追跡するセンサーである。(d)ローカルでない回転ディスク・ストレージ-RAMチップまたはROMチップが、ローカルのためのメモリーを提供する。(e)CDまたはDVD以外のドライブ、(f)家庭用器具または家庭用備品における埋め込み、(g)植え込み型医療デバイスまたはウェアラブル医療デバイスにおける埋め込み、(h)車両における埋め込み、(i)プロセス自動制御システムにおける埋め込み、または(j)以下の内1つに的を絞った設計：環境監視、都市インフラストラクチャ監視、工業用機器監視、エネルギー使用監視、人間もしくは動物の健康監視、物理的安全保護、もしくは物理輸送システム監視。IoTストレージは、クラウドを通じた不正アクセス、または他のネットワークを通じた不正アクセス、または直接的な局所アクセス試行による不正アクセスの標的となるおそれがある。

【 0 0 3 3 】

【0079】 状況によっては、例えば、秘匿性-完全性-可用性(confidentiality-integrity-availability)の3要素(triad)から成るサイバーセキュリティについての議論の中では、「データ可用性」の意味と「データ完全性」の意味との間で、区別するのはもっと

10

20

30

40

50

ものである。しかし、本明細書において使用する場合、「データー完全性」とは、データーにアクセスできること、および正しい値を有するデーターにアクセスできることの双方を包含することを意味する。つまり、記憶デバイスの電力停止、値の上書き、不可逆的な暗号化または符合化または圧縮、改竄、見当違い(misplacement)、スクランプリング、アクセス不可、およびデーター値不確定性は全て、データー完全性に対する脅威の例である。
【 0 0 3 4 】

【0080】 特に明示的に述べない限り、本明細書において使用する場合、「ディフラグ」(defrag)とはフラッシュ・ディフラグのことを指す(「ディフラグ」は「ディフラグメント」(defragment)の省略形である)。フラッシュ・ディフラグとは、アパーチャのフラッシュ先端近くにおいて未割り当てメモリーの量を削減する、または未割り当てメモリーのアパーチャのフラッシュ先端からの距離を延長させる、または双方を行う処理である。フラッシュ・ディフラグは、例えば、データーをコピーする、またはデーターを移動させる、または未だ割り当てられていないメモリーにデーターを挿入することを伴う場合もある。

10

【 0 0 3 5 】

【0081】 「フラッシュ・ディフラグ」が「統合ディフラグ」(consolidation defrag)と異なるのは、これらが異なる目標を有することである。フラッシュ・ディフラグは、クリティカル・データーのフラッシングを非クリティカル・データーのフラッシング、または未割り当てメモリー内にあるガベージもしくは無関係データーのフラッシングよりも優先するという目標を有する。統合ディフラグは、対照的に、使用するためにストレージから、特に、回転ディスク・プлатターからまたはテープから、データーを引き出すために要する時間を短縮するという目標を有する。統合ディフラグは、特定の所有者に属するデーター全て、例えば、特定のファイル内にあるデーター全てを纏めようとする。これらは同等の処理ではない。

20

【 0 0 3 6 】

【0082】 例えば、AまたはBは、それぞれ、AまたはBが所有するクリティカル・データーを表すとし、下線はメモリーにおける未割り当て空間を表すとし、メモリー・アパーチャの境界(bounds)は中括弧(curly braces)によって表すとする。すると、{ A B A _ } から { A A B _ } への変化は、統合ディフラグであり、フラッシュ・ディフラグではない。また、{ A _ B _ } から { A B _ _ } への変化は、フラッシュ・ディフラグであり(フラッシュ先端が左端である場合)、統合ディフラグではない。

30

【 0 0 3 7 】

【0083】 本明細書において使用する場合、「含む」(include)は、特に明記しない限り、追加の要素を許容する(即ち、含む(include)は備える(comprise)を意味する)。

【 0 0 3 8 】

【0084】 「最適化する」(optimize)は、必ずしも完璧にではないが、改良することを意味する。例えば、最適化されているプログラムまたはアルゴリズムにおいて更に改良することも可能な場合もある。

【 0 0 3 9 】

40

【0085】 「プロセス」(process)は、コンピューティング・サイエンス技術の用語として、本明細書においてときどき使用され、その技術的意味では、計算リソース・ユーザー(computational resource user)を包含する。また、計算リソース・ユーザーは、例えば、コルーチン、スレッド、タスク、割り込みハンドラー、アプリケーション・プロセス、カーネル・プロセス、手順、またはオブジェクト・メソッドも含んでもよく、またはこのように呼んでもよい。実際には、「プロセス」とは、Windows(登録商標) Task Manager、Linux(登録商標) ps、または他のオペレーティング・システム環境における同様のユーティリティ(それぞれ、Microsoft Corporation、Linus、Torvaldsの商標)のようなシステム・ユーティリティによって識別される計算エンティティである。また、「プロセス」は、本明細書におい

50

て、例えば、システムの請求項または製品（構成された記憶媒体）の請求項の対語として、プロセスの請求項を記載するときに、特許法の技術用語としても使用される。同様に、「方法」(method)は、本明細書では、コンピューティング・サイエンス技術における技術用語として（「ルーチン」の一種）、また特許法の技術用語（「プロセス」）としても使用されることがある。「プロセス」および「方法」は、特許法の意味では、本明細書においては相互交換可能に使用される。尚、当業者であれば、どちらの意味が特定の場において意図されているか理解でき、更に所与の特許請求されるプロセスまたは方法（特許法の意味）は、1つ以上のプロセスまたはメソッド（コンピューティング・サイエンスの意味）を使用して実装されてもよいことも理解されよう。

【0040】

10

【0086】「自動的に」(automatically)は、自動化を用いないことの対語として、自動化の使用によることを意味する（例えば、本明細書において論ずる特定の動作および技術的效果を求めるためにソフトウェアによって構成された汎用コンピューティング・ハードウェア）。具体的には、「自動的に」実行されるステップは、手作業で紙面上で実行されるのでも、人間の思考において実行されるのでもないが、人間によって開始されても、または人間によって対話処理で誘導されてもよい。自動ステップは、このように提供される技術的相互作用がなくては実現されない1つ以上の技術的效果を得るために、機械によって実行される。

【0041】

【0087】当業者には、技術的效果は技術的实施形態の推定上の目的であることが理解されよう。例えば、ある実施形態は計算を伴うが、いくつかの計算は技術的コンポーネント（例えば、紙と鉛筆によって、または精神的処理過程としてであっても）がなくても実行できるという単なる事実は、技術的效果の存在を取り除くことはなく、その実施形態の具体的および技術的な性質を変化させることもない。未割り当てメモリーを特定する、データをメモリーにコピーする、メモリーに割り当て済みの印を付ける、揮発性ストレージから不揮発性ストレージにデータをフラッシュする、および本明細書において論ずる多くの他の動作というような、データ保存動作は本質的にデジタルであると考えられている。人間の精神は、必要なデータを読み取りおよび書き込み、本明細書において教示されるデータ保存ステップを実行するために、直接CPUまたは他のプロセッサとも、あるいはRAMまたは他のデジタル・ストレージともインターフェースすることはできない。このことは、当業者であれば、本開示を考慮して全て十分に理解されようが、他の人々にはこれを知らせなければならない、または気付かせなければならない場合もある。

20

30

【0042】

【0088】「計算的に」(computationally)も同様に、コンピューティング・デバイス（少なくともプロセッサ＋メモリー）が使用されていることを意味し、単に人間の思考または単なる人間の行動だけによって結果を得ることを除外する。例えば、紙と鉛筆で算術演算を行うことは、本明細書においては理解されるように、計算的に算術演算を行っていることにはならない。計算結果は、より速く、より広く、より深く、より精度高く、より一貫性があり、より総合的であり、および/またはそれ以外で人間の処理能力(performance)のみの範囲を超える技術的效果を提供する。「計算ステップ」(computational step)とは、計算的に実行されるステップである。「自動的に」も「計算的に」も、必ずしも「直ちに」を意味しない。「計算的に」および「自動的に」は、本明細書では相互交換可能に使用される。

40

【0043】

【0089】「積極的に」(proactively)は、ユーザーからの直接的な要求を伴わないことを意味する。実際に、実施形態による積極的なステップが、そのステップの結果がユーザーに提示されるまで、可能であることに、ユーザーが気が付かないことさえもある。特に明記しない限り、本明細書において説明する計算および/または自動ステップはいずれも、積極的に行うことができる。

【0044】

50

[0090] この文書全体を通じて、随意的複数「(1つまたは複数)」(s),(es)または (ies)の使用は、示された特徴(feature)が1つ以上存在することを意味する。例えば、「プロセッサ(1つまたは複数)」は、「1つ以上のプロセッサ」または同等に「少なくとも1つのプロセッサ」を意味する。

【0045】

[0091] 米国の法律および慣行に限って言えば、本明細書、特許請求の範囲、または他の場所における「ステップ」(step)という単語の使用は、ミーンズ・プラス・ファンクション、ステップ・プラス・ファンクション、または35米国特許法第112条第6項/第112(f)条の請求項の解釈を引き合いに出すことを意図するのではない。その主旨の仮定はいずれも、これによって明示的に反証される。

10

【0046】

[0092] 米国の法律および慣行に限って言えば、請求項は、これらが「ための手段」(means for)という語句を使用していなければ、ミーンズ・プラス・ファンクションの解釈を引き合いに出すことを意図していない。ミーンズ・プラス・ファンクションの文言として解釈されるように意図された請求項の文言がある場合はいずれも、「ための手段」(means for)という語句を使用することによって、その意図を明示的に記載する。ミーンズ・プラス・ファンクションの解釈が適用されるとき、「ための手段」の使用によるか、および/または請求項の文言の裁判所の法的解釈によるかには関係なく、所与の名词または所与の動詞について明細書において記載される手段は、その請求項の文言に結び付けられ、更に、図面におけるブロック図の同じブロック内に現れること、同じまたは同様の名称による指定、同じ参照番号による指定(denotation)、図のいずれかにおいて表される機能的関係、本開示の本文において指摘される機能的関係の内任意のものによって、本明細書では一緒に結び付けられると理解されてしかるべきである。例えば、請求項の限定が「ザック・ウィジェット」(zac widget)を記載し、この請求項の限定がミーンズ・プラス・ファンクションの解釈の対象となる場合、あらゆる図のブロック、段落、あるいは「ザック・ウィジェット」について述べる例、またはザック・ウィジェットに割り当てられたあらゆる参照番号によって一緒に結び付けられる例、またはザック・ウィジェットの構造または動作と機能的関係を有するものとして開示される例において、本明細書のいずれかのところで識別される全ての構造は、少なくとも、ザック・ウィジェットに対する適用において識別される構造の一部であると見なされ、ザック・ウィジェット構造について1組の均等物を定める助けになる。

20

30

【0047】

[0093] 尚、当業者には、この技術革新の開示では、種々のデータ値およびデータ構造について論ずることが認められ、更にこのような項目はメモリー(RAM、ディスク等)内に常駐することによって、メモリーを構成することが認められよう。また、この技術革新の開示では、種々のアルゴリズム・ステップについても論じ、これらのアルゴリズム・ステップは所与の実施態様において実行可能コードに具体化されること、そしてこのようなコードもメモリー内に常駐すること、更にこのコードは、それを実行する任意の汎用プロセッサを事実上構成し、それによって汎用プロセッサから、機能的に特殊目的ハードウェアである特殊目的プロセッサに変形することが、当業者には認められよう。

40

【0048】

[0094] したがって、当業者であれば、(a)請求項に記載されるメモリー、および(b)その請求項に記載されるデータ構造またはデータ値またはコードを、非重複品目として扱うという間違いを犯さないであろう。請求項が、あらゆるデータ構造またはデータ値もしくはコード(piece of code)の常駐性(residency)を明示的に記載しなくても、言及されるデータ構造およびデータ値ならびにコードがメモリー内に常駐することは、暗黙裏に了解されよう(understood)。したがって、このような常駐性の明示的な記載は不要である。しかしながら、これらは禁止されているのではなく、強調のために1つまたは2つの選択の記載があってもよく、それによって他のデータ値およびデータ構造ならびにコード全てが常駐性から除外されるのではない。同様に、請求項において記載

50

されるコード機能は、その構成品質が請求項において明示的に記載されているか否かには関係なく、プロセッサを構成することは、暗黙裏に了解されよう。

【 0 0 4 9 】

[0095] 本文書全体にわたって、特に明記しない限り、プロセスにおけるステップについて言及するときはいつでも、そのステップが直接対象の要素(party of interest)によって実行されても、および/または仲介メカニズムおよび/または仲介エンティティを通じて、その対象の要素によって間接的に実行されてもよく、なおもそのステップの範囲内に該当することを仮定する。即ち、対象の要素によるステップの直接的な実行は、直接的な実行が明示的に述べられた要件でなければ、必須ではない。例えば、割り当てる、コピーする、ディフラグする、指定する、検出する、決定する、消滅させる、実行する、失敗する、フラッシュする、識別する、保持する、印を付ける、移動させる、置く、保存する、提供する、受け取る、削減する、予約する、常駐する(residing)、大きさを決め直す、再現する(restoring)、保存する、送る、指定する、格納する(および割り当てる(allocates)、割り当てられた(allocated)、コピーする(copies)、コピーされた(copied)等)のような、宛先または他の主題(subject)に対する、対象の要素によるアクションを伴うステップは、転送する、コピーする、アップロードする、ダウンロードする、エンコードする、デコードする、圧縮する、解凍する、暗号化する、解読する、認証する、呼び出す、および本文書において記載されるあらゆるアクションを含むその他というような、何らかの他の要素(party)による、仲介アクションを伴ってもよいが、それでもなおも対象の要素によって直接実行されたものと理解されてよい。

【 0 0 5 0 】

[0096] データまたは命令に言及するときはいつでも、これらの品目がコンピューター読み取り可能メモリーおよび/またはコンピューター読み取り可能記憶媒体を構成し、これによって、例えば、単純な紙面上の存在、人間の心における存在、またはワイヤ上を伝搬される単なる信号に対立するものとして、それを特定の物品(article)に変換することは、暗黙裏に了解されよう。米国における特許保護に限って言えば、メモリーまたは他のコンピューター読み取り可能記憶媒体は、米国特許および商標庁(USPTO)のIn re Nuijten事件の解釈の下における特許可能な主題の範囲から外れる、伝搬信号でも搬送波でも単なるエネルギーでもない。米国では信号自体または単なるエネルギーをその適用範囲とする請求項はなく、本開示を考慮して、そうでないことを主張するいずれの請求項の解釈も、文面上不合理である。米国以外において付与された請求項において特に明示的に述べられていないならば、請求項は信号自体も単なるエネルギーもその適用範囲としない。

【 0 0 5 1 】

[0097] 更に、本明細書において明らかに異なる定めがあっても、(a)一方で、コンピューター読み取り可能記憶媒体およびコンピューター読み取り可能メモリーと、(b)他方で、信号媒体とも呼ばれる伝送媒体との間に、明確な区別があることは理解されよう。伝送媒体は、伝搬信号または搬送波コンピューター読み取り可能媒体である。対照的に、コンピューター読み取り可能記憶媒体およびコンピューター読み取り可能メモリーは、伝搬信号でも、搬送波のコンピューター読み取り可能媒体でもない。請求項において特に明記しない限り、「コンピューター読み取り可能媒体」とは、コンピューター読み取り可能記憶媒体を意味し、伝搬信号自体でも単なるエネルギーでもない。

【 0 0 5 2 】

[0098] 「実施形態」(embodiment)は、本明細書では例である。「実施形態」という用語は、「本発明」と相互交換可能でない。実施形態は、結果的に得られる態様の組み合わせが、それ自体本明細書において明示的に記載されていなくても、他の実施形態を創作するために自由に態様を共有または借用してもよい(結果が動作可能であることを前提とする)。あらゆる許容される組み合わせを明示的にそして個別に記載することの要求は、当業者には不要であり、特許明細書は当業者である読者のために書かれることを認める方針に反する。少数の組み合わせ可能な特徴からでさえも得られる可能な組み合わせの数に関する正式な組み合わせの計算および非公式な共通の直観も、多数の態様の組み合わせ

が、本明細書において記載する態様に対して存在することを示す。したがって、あらゆる組み合わせの明示的な記載を要求することは、特許明細書が簡潔であり、該当する技術分野において読者に知識があることを求める方針に反することである。

【 0 0 5 3 】

[0099] 参照番号のリスト

[00100] 以下のリストは、便利であるように、そして図面の支援のために提示され、複数の項目を参照して技術革新を説明する明細書の本文の一部として提示されるものである。このリストに掲示されない項目も、しかしながら、所与の実施形態の一部であってもよい。本文を読みやすくするために、所与の参照番号は、本文において参照される項目の、全てではないが一部の記載の近くに記載される。同じ参照番号が、所与の項目の異なる例または異なる実例を参照するために使用される場合もある。参照番号のリストは次の通りである。

【 0 0 5 4 】

[00101] 1 0 0 コンピューティング環境とも呼ばれる動作環境

[00102] 1 0 2 計算システムまたはコンピューティング・システムとも呼ばれるコンピューター・システム

[00103] 1 0 4 ユーザー

[00104] 1 0 6 周辺機器

[00105] 1 0 8 例えば、L A N、W A N、ソフトウェア定義ネットワーク、クラウド、および他の有線またはワイヤレス・ネットワークを含むネットワーク全体

[00106] 1 1 0 プロセッサ

[00107] 1 1 2 コンピューター読み取り可能記憶媒体、例えば、R A M、ハード・ディスク。

【 0 0 5 5 】

[00108] 1 1 4 リムーバブル構成コンピューター読み取り可能記憶媒体

[00109] 1 1 6 プロセッサによって実行可能な命令であって、リムーバブル記憶媒体上または他のメモリー（揮発性または不揮発性または双方）内に常駐することができる。

【 0 0 5 6 】

[00110] 1 1 8 データー

[00111] 1 2 0 カーネル（1つまたは複数）、例えば、オペレーティング・システム（1つまたは複数）、B I O S、U E F I、デバイス・ドライバ

[00112] 1 2 2 ツール、例えば、ウィルス予防ソフトウェア、ファイアウォール、パケット・スニファァー・ソフトウェア、侵入検出システム、侵入防止システム、デバッガー、プロファイラー、コンパイラー、インタプリター、デコンパイラー、アセンブラー、ディスアセンブラー、ソース・コード・エディター、自動完了ソフトウェア、シミュレーター、ファザー(fuzzer)、レポジトリ・アクセス・ツール、バージョン制御ツール、オプティマイザー、協働ツール、ソフトウェア開発ツールおよびツール・スイート、ハードウェア開発ツールおよびツール・スイート、診断等。

【 0 0 5 7 】

[00113] 1 2 4 アプリケーション、例えば、ワード・プロセッサ、ウェブ・ブラウザ、スプレッドシート、ゲーム、電子メール・ツール。

[00114] 1 2 6 表示画面。「ディスプレイ」とも呼ぶ。

【 0 0 5 8 】

[00115] 1 2 8 他の参照番号 1 0 6、1 0 8、1 1 0、1 1 2、1 1 4 と関連付けられない(not otherwise)コンピューティング・ハードウェア。図 8 では、番号 1 2 8 はプロセッサ 1 1 0 およびメモリー 1 1 2 ハードウェアも示す。

【 0 0 5 9 】

[00116] 2 0 0 例えば、計算ノードにおける 計算機能。

[00117] 2 0 2 計算ノード。

10

20

30

40

50

[00118] 204 データー・ソース、例えば、データーを出力するまたそうでなければ生成するプログラム。

【0060】

[00119] 206 仮想機械、例えば、ハードウェア仮想化を行い、オペレーティング・システムを含むコンピューティング構造(computing construct)。例えば、作業用メモリー・リソース、CPUリソース、I/Oリソース、および不揮発性メモリーを含むことができる。

【0061】

[00120] 208 システム・コード、例えば、オペレーティング・システム、ハイパーバイザー、BIOS、UEFI、ファームウェア。

[00121] 210 例えば、ストレージ・ノードにおける格納機能。

【0062】

[00122] 212 ストレージ・ノード

[00123] 214 不揮発性ストレージ、例えば、NVDIMM、ディスク、フラッシュ・メモリー。

【0063】

[00124] 216 バッテリー

[00125] 218 バッテリー・バックアップ・メモリー

[00126] 220 メモリー・アパーチャ。物理デバイスにおけるデジタル・データ格納位置と関連付けられた物理アドレス空間の一部分を指すことができ、またはデジタル・データ格納位置自体を指すこともできる。

【0064】

[00127] 222 データー格納要求

[00128] 224 データー格納要求に対する応答

[00129] 226 データーの複製

[00130] 300 コンピューティング環境の態様

[00131] 302 クラウド。「クラウド・コンピューティング環境」と呼んでもよい。

【0065】

[00132] 304 データー保存回路。電子回路と、これらの電子回路またはその使用データー保存回路を制御する任意のファームウェアとを含み、必ずしもデーターがフラッシュされる不揮発性ストレージを含むとは限らない。

【0066】

[00133] 306 データー保存サブシステム。データーがフラッシュされる先の不揮発性ストレージを含む。

[00134] 308 ハイパーバイザー。

【0067】

[00135] 310 UEFI。統一拡張可能ファームウェア・インターフェース、またはこのようなインターフェースを有するファームウェアを指すことができ、例えば、SEC(セキュリティ)、PEI(事前EFI初期化)、DXE(ドライバ実行環境)、BDS(ブート・デバイス選択)、SMM(システム管理モード)のための機能を含むまたは機能を提供することができる。

【0068】

[00136] 312 フラッシュ・オーダー

[00137] 314 データー完全性の脅威

[00138] 316 2つ以上のノード間におけるストレージ関係

[00139] 318 データー・クリティカル性。データーがクリティカルに指定されているか否かに言及することができ、またはデーターが未だそのように指定されていない、クリティカルに指定される判断基準を満たすか否かに言及することができる。これは、ブーリアンまたは他の指示、例えば、データーのフラッシングにどのように優先順位を付けるかを示す2つ以上の値の範囲内にある値であってもよい。

10

20

30

40

50

【 0 0 6 9 】

[00140] 4 0 2 揮発性メモリー、例えば、D R A M 自体。

[00141] 4 0 4 N V D I M M

[00142] 4 0 6 バッテリー特性

[00143] 4 0 8 ディフラグガー(defragger) (メモリー・ディフラグメンテーション) コード。

【 0 0 7 0 】

[00144] 4 1 0 メモリーのディフラグ部分、即ち、ディフラグgingの対象となる部分。

[00145] 4 1 2 クリティカル・データーを含むデーター集合 (いくつかの非クリティカル・データーも有してもよい)。

10

【 0 0 7 1 】

[00146] 4 1 4 クリティカル・データー 4 1 2 を保持するために割り当てられた (または割り当てられている) メモリー。

[00147] 4 1 6 クリティカル・データーを全く含まないデーター集合。

【 0 0 7 2 】

[00148] 4 1 8 非クリティカル・データー 4 1 6 を保持するために割り当てられた (または割り当てられている) メモリー。

[00149] 4 2 0 ファームウェア。

【 0 0 7 3 】

20

[00150] 4 2 2 アロケーター(allocator)ソフトウェア、即ち、アパーチャにおける未割り当て空間を識別し、その中に、アパーチャのフラッシュ先端に向けてクリティカル・データーを入れ、この空間に割り当て済みの印を付けるソフトウェア。アロケーターは、アパーチャをディフラグすることも可能である。アロケーターは、非クリティカル・データーの空間を識別し、非クリティカル・データーをアパーチャのフラッシュ後端に向けて入れることも可能である。

【 0 0 7 4 】

[00151] 4 2 4 メモリー 1 1 2 における物理アドレス。アドレス自体 (例えば、0 × 0 0 0 0) またはアドレスにおける格納位置 (例えば、メモリー・セル) を指すことができる。

30

【 0 0 7 5 】

[00152] 5 0 2 バッテリー容量、例えば、ミリアンペア - 時間。

[00153] 5 0 4 信頼性値、例えば、平均故障間隔、または残存予測寿命、または列挙値 (例えば、1 0 が最良で 0 が最悪である目盛の中の 2)、または信頼性等級 (例えば、中度の信頼性)。

【 0 0 7 6 】

[00154] 5 0 6 バッテリーの経年数(age)、例えば、設置以来の経過時間、行われた充電サイクル回数、フラッシュの間に使用された回数。

[00155] 6 0 0 データー・クリティカル性判断基準の例。

【 0 0 7 7 】

40

[00156] 6 0 2 データー・クリティカル性判断基準。

[00157] 6 0 4 データーを生成した仮想機械に割り当てられる優先度。

[00158] 6 0 6 カスタマー、例えば、クラウドのテナント。

【 0 0 7 8 】

[00159] 6 0 8 カスタマーに割り当てられるステータス。例えば、S L A において明白にすることができる。

[00160] 6 1 0 ワークロード、たとえば、特定の計算ノードに割り当てられる処理タスク。

【 0 0 7 9 】

[00161] 6 1 2 データーを含むワークロードに、またはデーターが生成された実行

50

から割り当てられる優先度。

[00162] 702 物理機械、例えば、プロセッサ、回路、チップ、バッテリー、ディスプレイ、電源、筐体等。あらゆる埋め込みファームウェア、または物理機械の不揮発性メモリに格納されたソフトウェアを含む。コンピューティング・システム102は、1つ以上の物理機械702を含む。

【0080】

[00163] 704 クラウドまたは他のネットワークにおけるノード。計算ノード202およびストレージ・ノード212はノード704の例である。

[00164] 706 コンテナ。例えば、ユーザー空間仮想化を行う(provide)コンピューティング構造(construct)であり、それ自体はオペレーティング・システムを含まないが、それにもかかわらず、オペレーティング・システムに依存して実行する。

10

【0081】

[00165] 800 1つ以上のストレージ・ノードとネットワークを組む1つ以上の計算ノードを有し、本明細書において教示するようなデータ保存機能が装備されたコンピューティング・システム。

【0082】

[00166] 802 I/O(入力/出力)ハードウェアおよびソフトウェア。例えば、ポート、コネクタ、ソケット、ネットワーク・インターフェース・カード。「I/O」と呼ぶこともできる。

【0083】

20

[00167] 804 メモリーの最上位(top)。

[00168] 806 アパーチャのフラッシュ先端。

[00169] 808 アパーチャのフラッシュ後端。

【0084】

[00170] 810 アパーチャにおける未割り当て空間。

[00171] 812 同じメモリー・デバイス内であるがアパーチャの外側にあるメモリーの部分。

【0085】

[00172] 814 仮想アドレス空間と物理アドレス空間との間の論理マッピング。

[00173] 900 仮想機械を実行するオペレーティング・システムまたはコンテナを有するシステムを含み、本明細書において教示するようなデータ保存機能が装備されコンピューティング・システム。

30

【0086】

[00174] 902 フラッシュされるデータを受ける専用の不揮発性ストレージ112。

[00175] 904 専用ストレージ902にマッピングされるメモリー領域。バッテリー・バックアップ・メモリー・アパーチャ220は、領域904の一例である。

【0087】

[00176] 906 オペレーティング・システム。

[00177] 908 揮発性メモリーから、この揮発性メモリーをバックアップする専用不揮発性ストレージにデータをコピーする退避動作。退避は、退避されるデータの完全性に対する差し迫った脅威を全く伴わずに、意図的に行うことができ、または差し迫った脅威または発生している(current)脅威にตอบสนองして行われるフラッシュであってもよい。

40

【0088】

[00178] 910 専用不揮発性ストレージから、この専用不揮発性ストレージによってバックアップされている揮発性メモリーにデータをコピーする復元動作。

[00179] 912 プログラムの実行。プログラムを実行するアクト、またはプログラム動作のインスタンスを指すこともできる。

【0089】

[00180] 1000 フローチャート。また、1000は、図10のフローチャートによって示されるデータ保存方法、または図10のフローチャートにしたがうデータ保

50

存方法にも言及する。

【 0 0 9 0 】

[00181] 1 0 0 2 例えば、`malloc()` コール、コンストラクターの呼び出し、ファイルまたはブロックまたはプロブの退避、あるいは格納位置 (バイト、ページ等) が未だ割り当てられていないデータの格納を求める他の動作の結果として、クリティカル・データを格納する要求を受ける。

【 0 0 9 1 】

[00182] 1 0 0 4 例えば、システム・メモリー管理ソフトウェア、空きリスト(`free list`)、空きチャンクを示すビット・ベクトル、または同様のメカニズムを使用して、メモリーの未割り当て部分を識別する。

【 0 0 9 2 】

[00183] 1 0 0 6 例えば、メモリー割り当てを追跡するデータ構造内にビットを設定することによって、メモリーの一部分に割り当て済みの印を付ける。

[00184] 1 0 0 8 例えば、`memcpy()` によって使用される機能、ブロック・コピー等を実行することによって、データのコピーをメモリー内に入力する。

【 0 0 9 3 】

[00185] 1 0 1 0 例えば、`ACPI` 準拠システム上でパワー・コマンドを受けることによって、または電圧レベル監視回路からの信号を受けることによって、データ完全性に対する脅威を検出する。

【 0 0 9 4 】

[00186] 1 0 1 2 例えば、`NVDIMM` の `DRAM` 部分から `NVDIMM` のフラッシュ部分にデータをコピーすることによって、揮発性ストレージから専用不揮発性バックアップ・ストレージにデータをフラッシュする。

【 0 0 9 5 】

[00187] 1 1 0 0 フローチャート。また、1 1 0 0 は、図 1 1 のフローチャート (図 9 のステップおよび図 1 0 のステップを組み込む) によって示される保存方法、または図 1 1 のフローチャートにしたがうデータ保存方法にも言及する。

【 0 0 9 6 】

[00188] 1 1 0 2 データを保存する、即ち、データの少なくとも 1 つのコピーの完全性を維持する。

[00189] 1 1 0 4 データ格納要求を受ける。

【 0 0 9 7 】

[00190] 1 1 0 6 アパーチャ 2 2 0 のサイズまたは位置を決定もしくは指定する。

[00191] 1 1 0 8 例えば、バイトまたはブロックを単位としたアパーチャ 2 2 0 のサイズ。

【 0 0 9 8 】

[00192] 1 1 1 0 アパーチャ 2 2 0 の位置。例えば、アパーチャを含む(`contain`)メモリー・デバイスの最も小さい物理アドレスからのアパーチャの両端のオフセット。

[00193] 1 1 1 2 データ格納要求を送る。

【 0 0 9 9 】

[00194] 1 1 1 4 アパーチャのサイズを決定し直す。

[00195] 1 1 1 5 アパーチャを移動させる。

[00196] 1 1 1 6 非クリティカル・データのみを保持するために、アパーチャの少なくとも指定された部分を予約する。

【 0 1 0 0 】

[00197] 1 1 1 8 アパーチャの予約部分。絶対または相対メモリー・アドレスによって、または総アパーチャ・サイズの割合(`percentage`)として指定することができる。

[00198] 1 1 2 0 データをコピーする、またはデータを移動させる、または以前には割り当てられていなかったメモリーにデータを挿入することによって、アパーチャの少なくとも一部分をデフラグすることによって、未割り当てであるメモリーの量を

10

20

30

40

50

削減する、またはアパーチャのフラッシュ先端からの未割り当てメモリの距離を延長する、または双方を行う。例えば、これを、同じユーザーまたは同じプロセスまたは同じファイルに属するメモリの部分を統合することを伴う、別の「ディフラグ」の使用から区別するために、これを「フラッシュ・ディフラグ」と呼んでもよい。

【0101】

[00199] 1122 アパーチャからのデータの一部をフラッシュし損なう。この結果、データの他のコピーが、アパーチャ以外のどこかで入手可能でなければ、そしてデータの残りがフラッシュされた不揮発性ストレージ以外のどこかで入手可能でなければ、データの少なくとも部分的な逸失が生じる。

【0102】

[00200] 1124 物理機械内に常駐させる。たとえば、物理機械の一部であるメモリ（揮発性または不揮発性）に格納する。

[00201] 1126 アパーチャ外部、そしてアパーチャがフラッシュされる不揮発性ストレージ外部のどこかに、データのコピーを保持する。

【0103】

[00202] 1128 データを移動させる。

[00203] 1130 空間を作る、即ち、メモリの所与の部分を割り当て済みから空きに変更する。

【0104】

[00204] 1132 データをクリティカルに指定する。データがユニットとして扱われるとき、そのユニットの任意の部分をクリティカルに指定すると、そのユニットもクリティカルに指定される。

【0105】

[00205] 1134 割り当てられた部分間にある未割り当て量を削減する。

[00206] 1136 割り当て部分間にある未割り当て空間を消滅させる。

[00207] 1138 脅威に対抗してクリティカル・データを保存する可能性を高める。

【0106】

[00208] 1140 脅威に対抗してクリティカル・データを保存する可能性。測定された確率(probability)または適正な評価でもよい。

[00209] 1142 ファームウェアを実行する。実行912の一例。

【0107】

[00210] 1144 他のいずれの参照番号も割り当てられていない、本開示において論じられるあらゆるステップ。

[00211] 動作環境

[00212] 図1を参照すると、実施形態の動作環境100は、少なくとも1つのコンピューター・システム102を含む。コンピューター・システム102は、マルチプロセッサ・コンピューター・システムであっても、またはなくてもよい。動作環境は、所与のコンピューター・システム内に1つ以上の機械を含むことができ、これらはクラスター化されても、クライアント・サーバー関係でネットワーク接続されても、および/またはクラウド内においてピア・ツー・ピアの形態でネットワーク接続されてもよい。個々の機械はコンピューター・システムであり、協働する機械のグループもコンピューター・システムである。所与のコンピューター・システム102が、例えば、アプリケーションを用いて、エンド・ユーザーのために構成されても、アドミニストレータのためにサーバーとして構成されても、分散型処理ノードとして構成されても、および/または他の方法で構成されてもよい。

【0108】

[00213] 人間のユーザー104は、ディスプレイ、キーボード、およびその他の周辺機器106を使用することによって、タイプ入力されたテキスト、タッチ、音声、動き、コンピューター・ビジョン、ジェスチャ、および/またはI/Oの他の形態を通じて、コンピューター・システム102と対話処理することができる。画面126は、着脱可能な

10

20

30

40

50

周辺機器 106 であってもよく、またはシステム 102 の統合部分であってもよい。ユーザー・インターフェースは、実施形態と 1 人以上の人間のユーザーとの間における対話処理を支援することができる。ユーザー・インターフェースは、コマンド・ライン・インターフェース、グラフィカル・ユーザー・インターフェース (GUI)、自然ユーザー・インターフェース (NUI)、音声コマンド・インターフェース、および/またはその他のユーザー・インターフェース (UI) 表示を含んでもよく、これらは別個の選択肢として提示されてもよく、または統合されてもよい。

【0109】

[00214] システム・アドミニストレーター、ネットワーク・アドミニストレーター、クラウド・アドミニストレーター、セキュリティ・アナリストおよび他のセキュリティ要員、操作要員、開発者、技術者、およびエンド・ユーザーは、各々、特定の型のユーザー 104 である。1 人以上の人間の代わりに、例えば、システム 102 を検査し易くするために動作する自動化エージェント、スクリプト、再生ソフトウェア、デバイス等も、ユーザー 104 としてよい。記憶デバイスおよび/またはネットワーキング・デバイスは、これらのプロセッサ 110 からの離脱可能性(detachability)に応じて、ある実施形態では周辺機材として見なされ、他の実施形態ではシステム 102 の一部と見なされることもある。図 1 に示されていない他のコンピューター・システムも、技術的な方法で、コンピューター・システム 102 と、または、例えば、ネットワーク・インターフェース機材を通じて、ネットワーク 108 への 1 つ以上の接続を使用して、他のシステムの実施形態と相互作用することができる。

【0110】

[00215] 各コンピューター・システム 102 は、少なくとも 1 つのプロセッサ 110 を含む。コンピューター・システム 102 は、他の適したシステムと同様に、1 つ以上のコンピューター読み取り可能記憶媒体 112 も含む。記憶媒体 112 は、異なる物理型のものであってもよい。記憶媒体 112 は、揮発性メモリー、不揮発性メモリー、所定位置固定媒体(fixed in place media)、リムーバブル媒体、磁気媒体、光媒体、ソリッド・ステート媒体、および/または他の種類の物理耐久記憶媒体(単なる伝搬信号または単なるエネルギーとは対立するものとして)であってもよい。具体的には、可搬型(即ち、外部)ハード・ドライブ、CD、DVD、メモリー・スティック、または他のリムーバブル不揮発性メモリー媒体のような、構成記憶媒体(configured medium) 114 は、挿入またそうでなければインストールされると、機能的にコンピューター・システムの技術的部分となり、そのコンテンツを、プロセッサ 110 との相互作用のためまたはプロセッサ 110 による使用のために、アクセス可能にすることができる。リムーバブルな構成記憶媒体 114 は、コンピューター読み取り可能記憶媒体 112 の例である。コンピューター読み取り可能記憶媒体 112 のいくつかの他の例には、内蔵 RAM、ROM、ハード・ディスク、およびユーザー 104 によって容易に除去できない他のメモリー記憶デバイスが含まれる。現行の米国特許要件に従って、コンピューター読み取り可能媒体も、コンピューター読み取り可能記憶媒体も、コンピューター読み取り可能メモリーも、米国において係属中または付与されたいずれの請求項の下においても、信号自体でも単なるエネルギーでもない。

【0111】

[00216] 記憶媒体 114 には、プロセッサ 110 によって実行可能なバイナリー命令 116 が構成される。「実行可能」(executable)とは、例えば、機械コード、解釈可能なコード、バイトコード、および/または仮想機械上で実行するコードを含むように、本明細書では広い意味で使用される。また、記憶媒体 114 には、命令 116 の実行による技術的效果のために、作成、修正、参照、および/またはそれ以外では使用されるデータ 118 が構成される。命令 116 およびデータ 118 は、これらが常駐するメモリーまたは他の記憶媒体 114 を構成する。このメモリーまたは他のコンピューター読み取り可能記憶媒体が所与のコンピューター・システムの機能部であるとき、命令 116 およびデータ 118 もこのコンピューター・システムを構成する。ある実施形態では、データ

ー 1 1 8 の一部は、製品特性、在庫、物理測定値、設定値、画像、読み取り値、目標、分量等のような、実世界項目を表す。また、このようなデータは、バックアップ、修復、コミット、中断、再フォーマット、および/またはその他の技術的処理によって変換される。

【 0 1 1 2 】

【00217】 実施形態は、コンピューティング・デバイス（例えば、汎用コンピューター、サーバー、またはクラスター）において1つ以上のプロセッサによって実行されるソフトウェア命令として実装されるというように説明される場合もあるが、このような説明は、可能な全ての実施形態を網羅することを意味するのではない。同じまたは同様の技術的効果を得るために、同じまたは同様の機能も、全体的または部分的に、直接ハードウェア・ロジックにおいて実装できる場合が多いことは、当業者には理解されよう。ソフトウェアの実装の代わりにまたはこれに加えて、本明細書において説明する技術的機能は、少なくとも部分的に、1つ以上のハードウェア・ロジック・コンポーネントによって実行することができる。ハードウェア・ロジック・コンポーネントは、ファームウェアを含んでも、ファームウェアによって制御されても、または双方でもよい。例えば、実施形態は、フィールド・プログラマブル・ゲート・アレイ（FPGA）、特定用途集積回路（ASIC）、特定用途標準製品（ASSP）、システム・オン・チップ・コンポーネント（SOC）、複合プログラマブル・ロジック・デバイス（CPLD）、および同様のコンポーネントというような、ハードウェア・ロジック・コンポーネント 1 1 0、1 2 8 を含んでもよく、更に そして他の実装も除外しない。実施形態のコンポーネントは、例えば、それらの入力、出力、および/またはそれらの技術的効果に基づいて、相互作用する機能モジュールに類別する(group)こともできる。

【 0 1 1 3 】

【00218】 プロセッサ 1 1 0（例えば、CPU、ALU、FPU、および/またはGPU）、メモリー/記憶媒体 1 1 2、およびディスプレイ 1 2 6に加えて、動作環境は、実例をあげると、バッテリー、バス、電源、有線およびワイヤレス・ネットワーク・インターフェース・カードのような、他のハードウェア 1 2 8 も含むことができる。「画面」(screen)および「ディスプレイ」(display)という名詞は、本明細書では相互交換可能に使用される。ディスプレイ 1 2 6 は、1つ以上のタッチ・スクリーン、ペンまたはタブレットからの入力に応答する画面、または出力のためにのみ動作する画面を含むことができる。ある実施形態では、人間ユーザー I/O デバイス（画面、キーボード、マウス、タブレット、マイクロフォン、スピーカー、動きセンサー等）のような周辺機器 1 0 6 は、1つ以上のプロセッサ 1 1 0 およびメモリーと動作可能に通信するように存在することもできる。

【 0 1 1 4 】

【00219】 ある実施形態では、このシステムは、有線および/またはワイヤレス・ネットワーク 1 0 8 によって接続された複数のコンピューターを含む。ネットワーキング・インターフェース機材 1 2 8 は、例えば、パケット交換ネットワーク・インターフェース・カード、ワイヤレス送受信機、または電話ネットワーク・インターフェースのようなネットワーク・コンポーネントを使用して、ネットワーク 1 0 8 へのアクセスを付与することができる。これらのコンポーネントは、所与のコンピューター・システム内に存在することができる。また、ネットワーキング・インターフェース機材、およびスイッチまたはルーターまたはファイアウォールのような他のネットワーク・コンポーネントの仮想化も、例えば、ソフトウェア定義ネットワーク、あるいはサンドボックス(sandboxed)環境または他のセキュア・クラウド・コンピューティング環境において、存在することができる。また、所与の実施形態は、直接メモリー・アクセス、リムーバブルな不揮発性記憶媒体、あるいは他の情報記憶検索および/または送信手法によって、技術的データおよび/または技術的命令を伝達することもできる。

【 0 1 1 5 】

【00220】 尚、以上の態様および「動作環境」の下で本明細書において紹介される他の

10

20

30

40

50

態様が、所与の実施形態の一部を形成できることは、当業者には認められよう。この文書の見出しは、機能(features)を実施形態と実施形態以外の機能集合に厳格に分類することを意図するのではない。

【 0 1 1 6 】

[00221] 1つ以上の項目が、図において、輪郭形態で示され、または括弧の中に掲示されているが、これらは必ずしも図示する動作環境の一部または全ての実施形態の一部ではないが、本明細書において論ずるように、動作環境またはいくつかの実施形態における項目と相互作用することができるということを強調するためである。だからと言って、輪郭形態でない項目や括弧がない項目が、いずれの図またはいずれの実施形態においても、必ず必要とされるという訳ではない。具体的には、図1は便宜上示されているのであり、図1に項目を含めることは、その項目、またはその項目の使用についての説明が、現在の技術革新の前に知られていたことを暗示するのではない。

10

【 0 1 1 7 】

[00222] 更にシステムについて

[00223] 図1から図11までを参照すると、ある実施形態は機能増強システム800または900を使用もしくは提供する。機能増強は、動作ステップとデバイス特性との組み合わせを実装して完全性の脅威314に対抗してデーターを保存することによって、データー保存を促進する(promote)のに役立つ。データー118は、クリティカル・データー412および非クリティカル・データー416に分割され、クリティカル・データーは1つ以上のバッテリー・バックアップ・メモリー218のアパーチャ220内の物理アドレス424に格納され、脅威が検出されたとき(1010)に非クリティカル・データーよりも先に、クリティカル・データーがフラッシュされる(1012)。バッテリー・バックアップ・メモリー・アパーチャ220は、バッテリーの特性406にしたがってサイズが決定され、条件が変化するに連れてサイズを変更することができる。ディフラガー(defragger)408は、クリティカル・データー412を格納するためのアパーチャの使用を最適化する。

20

【 0 1 1 8 】

[00224] 図2のアーキテクチャ例によって示すように、データー118は、計算ノード202上で実行する(912)仮想機械206またはコンテナ706のような、ソース204によって生成されてもよい。これらのデーター・ソースは、システム・コード208を呼び出して、またはシステム・コード208に依存して、生成したデーター118をストレージ・ノード212に送る。この図2の例では、ストレージ・ノード212は高速揮発性メモリー402を有し、その少なくとも一部はバッテリー・バックアップ・メモリー218となっている。バッテリー・バックアップ・メモリー218の少なくとも一部は、本明細書において教示するように、クリティカル・データーを保存するのを補助するために、アパーチャ220としての使用に指定されている。ストレージ・ノード212および計算ノード202は、要求222(例えば、ここに格納するデーターがあります。識別子Zの下で格納されているデーターを返送して下さい)および応答224(例えば、格納成功、アパーチャに余裕なし、ここに要求されたデーターがあります、または他のステータスもしくはエラー・コード)を通じて通信する。要求222および応答224は、例えば、TCP/IPネットワーク108、イーサネット・ネットワーク108、ストレージ・エリア・ネットワーク108、またはこれらの組み合わせを通じて伝送する(travel)ことができる。

30

40

【 0 1 1 9 】

[00225] 一般に、計算ノード - ストレージ・ノードの関係316は、必ずしも1対1ではない。1つの計算ノードが1つのストレージ・ノードに送ったまたは1つのストレージ・ノードから受信したデーターを処理し、あるいは複数のストレージ・ノードと通信することもできる。1つのストレージ・ノードが1つ以上の計算ノードに代わってデーターを格納することもできる。図8は、例えば、4つのストレージ・ノード212を使用する1つの計算ノード202を示す。

50

【 0 1 2 0 】

[00226] 図 2 および図 8 に示すように、ある実施形態では、クリティカル・データー 4 1 2 は、ストレージ・ノードに跨がって複製されず、一方非クリティカル・データー 4 1 6 はストレージ・ノードに跨がって複製される。つまり、クリティカル・データーは、主にこのクリティカル・データーを不揮発性ストレージにフラッシュすることによって、またはそうすることによってのみ、脅威に対抗して保存され、一方非クリティカル・データーは、主に、それをノードに跨がって複製することによって、またはそうすることによってのみ、脅威に対抗して保存されるとしてよい。

【 0 1 2 1 】

[00227] 図 8 において、クリティカル・データー 4 1 2 およびアパーチャ 2 2 0 における対応する割り当てエリア 4 1 4 は、仮想機械 A、C、および D に属し、図において丸い角によって示されている。一方、非クリティカル・データー 4 1 6 およびアパーチャ 2 2 0 における対応する割り当てエリア 4 1 8 は、仮想機械 B および X に属し、図において直角の角によって示されている。仮想機械 B および X に属するデーターのデーター複製 2 2 6 が示されており、ストレージ・ノード S 1、S 2、S 3 は仮想機械 B の非クリティカル・データー 4 1 6 の複製を有し、ストレージ・ノード S 1、S 2、S 4 は仮想機械 X の非クリティカル・データー 4 1 6 の複製を有する。

【 0 1 2 2 】

[00228] 図 8 および図 9 は、システム・コード 2 0 8 のいくつかの例、即ち、カーネル 1 2 0 (例えば、ハイパーバイザー 3 0 8、オペレーティング・システム 9 0 6) およびファームウェア (たとえば、UEFI ファームウェア 3 1 0) を示す。計算ノードおよびストレージ・ノードは、コンピューティング・システム 1 0 2 であり、したがってハードウェアを含む。尚、明確化のために、メモリー 1 1 2 以外のハードウェアは図 2 では特に明示されて (called out) いないが、ハードウェアのいくつかの例、例えば、IO 8 0 2 および CPU 1 1 0 は、メモリー 1 1 2、2 1 8 と共に、図 8 において明示されている。メモリー 2 1 8 は、アパーチャ 2 2 0、アパーチャ外部のメモリー 8 1 2、アパーチャ内部の割り当てメモリー 4 1 8 および 4 1 4、ならびにアパーチャ内部の未割り当てメモリー 8 1 0 を含む。

【 0 1 2 3 】

[00229] バッテリー・バックアップ・メモリー 2 1 8 の物理アドレス 4 2 4、または他のアパーチャを含みマッピングされた揮発領域 (aperture-containing mapped volatile region) 9 0 4 は、マッピング 8 1 4 によって、仮想アドレスにマッピングすることができる。マッピング 8 1 4 は、あるレベルの間接参照 (indirection) を実施し、これによって、同じ退避 9 0 8 および復元 9 1 0 ファームウェアを、互いに異なる物理開始アドレス 8 0 6 および物理終端アドレス 8 0 8 を有するかもしれないアパーチャと共に使用することが可能になる。アパーチャ 2 2 0 のデーターがフラッシュされる先の不揮発性メモリー 1 1 2 は、NVDIMM 4 0 4 においてフラッシュまたは他の不揮発性ストレージ、あるいは他の専用不揮発性ストレージ 9 0 2、もしくは双方を含むことができる。

【 0 1 2 4 】

[00230] 図 3 は、あるコンピューティング環境 1 0 0 の種々の態様 3 0 0 を示す。これらについては、本開示における種々の時点において適宜説明する。

[00231] 図 4 は、革新的なデーター保存機能をコンピューティング・システム 1 0 2 に付加することができるデーター保存サブシステムの一例 3 0 6 を示す。サブシステム 3 0 6 のメモリーは、バッテリー・バックアップ・メモリー 2 1 8、例えば、NVDIMM 4 0 4 においてアパーチャ 2 2 0 を含む。図 4 および図 5 において気付かれるように、バッテリーは、容量 5 0 2、信頼性 5 0 4、および経年数 5 0 6 の内 1 つ以上というような、追跡または指示対象となる特性 4 0 6 を有することができる。これらのバッテリー特性 4 0 6 は、互いに重複する、即ち、互いに影響を及ぼし合うこともある。例えば、信頼性は経年数と共に低下するとして差し支えない。また、サブシステム 3 0 6 は、バッテリー 2 1 6 によってバックアップされない揮発性メモリー 4 0 2 も含むことができる。

10

20

30

40

50

【 0 1 2 5 】

[00232] 動作において、アパーチャ 2 2 0 は、クリティカル・データー 4 1 2 を収容する 1 つ以上の領域 4 1 4 を有し、アパーチャ 2 2 0 は、非クリティカル・データー 4 1 6 を収容する 1 つ以上の領域 4 1 8 を有することもできる。あるアパーチャ 2 2 0 は、例えば、図 8 に示すストレージ・ノード S 2、S 3、S 4 のように、少なくとも所与の時点において、非クリティカル・データーだけを収容することもできる。しかし、ここで関心の焦点になるのは、クリティカル・データーを収容するアパーチャ 2 2 0 である。具体的には、本開示は、データーの脅威に応答してクリティカル・データーをフラッシュするのに成功する可能性 1 1 4 0 を高めるようにするには、どのようにしてクリティカル・データーをメモリー 1 1 2 に入力すればよいか（初期割り当てによって、またはディフラッキングによって、または双方）、つまり保存（1 1 0 2）すればよいかについて情報提供する。

10

【 0 1 2 6 】

[00233] 図 6 は、クリティカル・データー 4 1 2 を非クリティカル・データー 4 1 6 から区別するために適用することができる判断基準 6 0 2 のいくつかの例 6 0 0 を示す。例えば、データーが仮想機械 2 0 6 またはコンテナ 7 0 6 または他のデジタル・アーチファクトによって生成された場合、このデーターをクリティカルに指定することができ、このアーチファクトは高い優先度 6 0 4 を有する。データー 1 1 8 が、支払いを行った特定の顧客 6 0 6 によってまたはその代わりに生成された場合、あるいはそれ以外ではクリティカル・データーを示すステータス 6 0 8 が得られた場合、クリティカルに指定する（1 1 3 2）ことができる。同様に、データーが、クリティカル・データーを示す優先度またはステータス 6 1 2 を有する特定のワークロード 6 1 0 によって、またはその代わりに生成された場合、あるいは特定のワークロード 6 1 0 によって処理された場合、このデーターをクリティカルに指定する（1 1 3 2）ことができる。これらの判断基準例を種々の方法で組み合わせることもでき、または他の判断基準 6 0 2 によって修正することもでき、あるいは双方でもよい。実例をあげると、（a）データーが高いステータスの顧客の高優先度仮想機械によって生成されたとき、このデーターをクリティカルに指定することができるが、それ以外の場合非クリティカルに指定することができる。（b）指定されたセキュア・サーバー・ブレード 7 0 2 上で実行する高ステータスの顧客の何らかの仮想機械によってデーターが生成されたとき、このデーターをクリティカルに指定することができるが、それ以外の場合非クリティカルに指定することができる。（c）指定された時間期間中に任意の顧客によってデーターが生成されたとき、このデーターをクリティカルに指定することができるが、それ以外の場合非クリティカルに指定することができる、等である。

20

30

【 0 1 2 7 】

[00234] 図 7 は、あるコンピューティング・システム 1 0 2 の種々の例 7 0 0 を示す。これらについては、本開示の種々の時点において適宜説明する。

[00235] ある実施形態は、コンピューティング・システム 1 0 2 においてデーター保存サブシステム 3 0 6 を提供または使用する。データー保存サブシステム 3 0 6 は、アパーチャ 2 2 0 を有するバッテリー・バックアップ・メモリー 2 1 8 を含むことができる。アパーチャ 2 2 0 は、フラッシュ・オーダー 3 1 2 を有する。フラッシュ・オーダー 3 1 2 は、データー完全性の脅威 3 1 4 に応答して、データーをアパーチャからアパーチャ関連不揮発性ストレージ 9 0 2 にコピーする順序(order)である。フラッシュ・オーダーは、アパーチャのフラッシュ先端 8 0 6 と、アパーチャのフラッシュ後端 8 0 8 とを定める。

40

【 0 1 2 8 】

[00236] また、フラッシュ・オーダーは、「トップ・ダウン」(top-down)または「ボトム・アップ」(bottom-up)というような語句を使用して記述することもでき、大きい物理アドレス 4 2 4 程小さい物理アドレスよりも上にあると理解することとする。したがって、トップ・ダウン・フラッシュは、高い方のアドレスから開始し、低い方のアドレスに向かって進む。一方ボトム・アップ・フラッシュは、低い方のアドレスから開始し、高い

50

方のアドレスに向かって進む。アパーチャのフラッシュ先端 806 は、トップ・ダウン・フラッシュ・オーダーでは、アパーチャのフラッシュ後端 808 よりも大きなアドレスを有し、ボトム・アップ・フラッシュ・オーダーでは、アパーチャのフラッシュ後端 808 よりも小さいアドレスを有する。図 8 では、アパーチャのフラッシュ先端 806 は、アパーチャのフラッシュ後端 808 よりも小さい物理アドレス 424 を有するので、ボトム・アップ（小さいアドレスから大きいアドレスへ）フラッシュ・オーダー 312 を示す。

【0129】

[00237] ある実施形態では、データ保存サブシステム 306 は、バッテリー・バックアップ・メモリー 218 と動作可能に通信するデータ保存回路 304 を含む。データ保存回路 304 は、データ保存ステップを実行するように構成される。データ保存ステップは、以下を含むことができる。

10

【0130】

(a) データ集合 A - データ 412 をアパーチャに格納する要求 222 を受けるステップ (1002)。A - データは、クリティカル・データに指定されたデータを含む。

【0131】

(b) アパーチャの未割り当てメモリー 810 の一部分 A - メモリー 414 を特定する (1004)。A - メモリーは、A - データを保持できるだけ十分に大きく、A - メモリーは、アパーチャの他のいずれの未割り当てメモリーの他のいずれのアドレスよりもアパーチャのフラッシュ先端 806 に近いアドレス 424 を有する。

20

【0132】

(c) A - メモリーに割り当て済みの印を付け (1006)、A - データのコピーを A - メモリーに入力する (1008)。

(d) データ集合 Z - データ 416 をアパーチャに格納する要求 222 を受ける。Z - データは、クリティカル・データに指定されるデータを全く含まない。

【0133】

(e) アパーチャの未割り当てメモリー 810 の一部分 Z - メモリーを特定する。Z - メモリーは、Z - データを保持できるだけ十分に大きく、Z - メモリーは、アパーチャの他のいずれの未割り当てメモリーの他のいずれのアドレスよりも、アパーチャのフラッシュ後端 808 に近いアドレス 424 を有する。

30

【0134】

(f) Z - メモリーに割り当て済みの印を付け、Z - データのコピーを Z - メモリーに入力する。

ここで、ラベル (a) から (f) までは、単に識別子として役割を果たすに過ぎず、したがって必ずしもこれら自体が動作の順序を指定する訳ではない。このデータ保存サブシステム 306 は、フラッシュ・オーダーにおいてクリティカル・データを非クリティカル・データよりも前にバッテリー・バックアップ・メモリーに入力することによって、クリティカル・データをフラッシュするのに成功し (1012)、それによってクリティカル・データを保存する可能性を高める。クリティカル・データ保存の可能性は、データのクリティカル性を考慮せずに、即ち、クリティカル性を欠いて、バッテリー・バックアップ・メモリーにデータを格納するアーキテクチャにおけるよりも高い。

40

【0135】

[00238] ある実施形態では、バッテリー・バックアップ・メモリー・アパーチャ 220 は、ここでは X と呼ばれるネットワーク・ノード 704 内に常駐し、またはネットワーク・ノード 704 によって制御され、A - データをアパーチャに格納する要求 222 は、ここでは Y と呼ばれる異なるネットワーク・ノードから送られた。図 8 では、例えば、アパーチャはストレージ・ノード 212、704 S1 に常駐し、アパーチャ 220 における仮想機械 A に対する A - データを格納する要求 222 は、S1 以外の異なるネットワーク・ノードから、即ち、計算ノードから送られた。ある実施形態では、ネットワーク・ノード X は、クラウド 302 内にストレージ・ノードを含み、ネットワーク・ノード Y

50

は、クラウド内に計算ノードを含む。しかしながら、他の実施形態では、ノードがクラウド以外のネットワークに常駐することもできる。

【0136】

【00239】ある実施形態では、ネットワーク・ノードXおよびネットワーク・ノードYが、ここではRと呼ばれるストレージ関係316によって関係付けられる。ここで、Rは、MがNの代わりにデーターを格納するとき、2つのノードMおよびNがRによって関係付けられるように、定められる。ある場合には、Rは特定のノードXおよびYの1対1の関係ではないこともある。例えば、計算ノードが仮想機械VM7のクリティカル・データーをストレージ・ノードS7に送り、仮想機械VM8のクリティカル・データーをストレージ・ノードS8に送る場合もある。または、ストレージ・ノードS12が計算ノードC10からのクリティカル・データーを保持し、更に異なる計算ノードC11からのクリティカル・データーも保持する場合もある。勿論、多くの他の関係316の例も可能である。

10

【0137】

【00240】ある実施形態では、バッテリー・バックアップ・メモリー218はNVDIMMメモリー404を含む。例えば、NVDIMM404は、NVDIMM-Fフラッシュ・メモリー、NVDIMM-Nバイト・アドレス可能メモリー、同じデバイス上にダイナミックRAMおよびNANDを有するNVDIMM-Pメモリー、NVDIMM-SWメモリー、またはNVRAM不揮発性RAMメモリーを含むことができる。NVDIMM-SWは、メモリー（例えば、DDR4RAM）、ディスク、またはフラッシュ（例えば、SSDパーティション）、バッテリー、およびシグナリング・メカニズムを含むこともできる。

20

【0138】

【00241】ある実施形態では、コンピューティング・システム102は仮想機械206を含み、データー集合A - データー412はデーター保存サブシステム306に仮想機械VM-Aから送られ、データー集合Z - データー416はデーター保存サブシステム306に仮想機械VM-Zから送られる。図8の例では、实例をあげると、クリティカル・データー集合412は増強ストレージ・ノードS1（したがってそのデーター保存サブシステム306に）に仮想機械Aから送られ、非クリティカル・データー集合416はS1に仮想機械Bから送られる。

【0139】

30

【00242】本明細書では、他のシステムの実施形態についても、直接説明するか、あるいは説明したプロセスまたは構成した媒体のシステム・バージョンとして導き出せる(derivable)ものとして説明し、コンピューティング・ハードウェアについての本明細書における広範な論述によって情報提供する。

【0140】

【00243】図において具体的なアーキテクチャの例を示したが、実施形態はこれらの例から逸脱してもよい。实例をあげると、異なる図に示した品目を1つの実施形態に纏めて含ませてもよく、1つの図に示す品目を省略してもよく、異なる品目において示される機能を、もっと少ない品目にまたは1つの品目に組み込んでもよく、品目の名称を変更してもよく、あるいは品目を互いに異なる接続にしてもよい。

40

【0141】

【00244】本明細書では、技術の態様を例示するのに役立てるために例を示すが、この文書内で示す例は、可能な実施形態の全てを説明するのではない。実施形態は、本明細書において提示する特定の例、コンポーネントの名称、最適化、アルゴリズムの選択、データー、データー型、構成、実施態様、配列、表示、特徴、手法、そしてシナリオにも限定されない。所与の実施形態は、实例をあげると、追加のまたは異なる技術的特徴、メカニズム、シーケンス、データー構造、または機能を含んでもよく、それ以外でも本明細書において提示する例から逸脱してもよい。

【0142】

【00245】プロセス（別名、方法）

50

【00246】 図10は、本明細書において教示するデータ保存機能を有し、図1から図9までの1つ以上によって示される増強システムによって実行または補佐することができる方法の一例である、方法1000を示す。図10に示す受けるステップ1002、特定するステップ1004、印を付けるステップ1006、入力するステップ1008、検出するステップ1010、およびフラッシュするステップ1012について、本開示全体を通じて、これらの参照番号で明示的に引用しながら、そして引用せずに説明する。具体的には、図2、図8、および図9の論述は、図10に示すステップ1002から1012までに関係がある。

【0143】

【00247】 図11は、更に、アクセス制御方法（その世界の法的な意味では「プロセス」と呼ぶこともできる）も示す。このアクセス制御方法は、増強データ保存機能を有するシステムの動作中に使用するのに適しており、図10に示すステップに対する何らかの改善(refinement)、補助、またはコンテキストに関するアクションを含む。また、図11は、図9または図10に示すステップも組み込む。特に明記されなければ、これらの図に示されるまたは他で開示される技術的プロセスは、例えば、データ保存サブシステム306によって自動的に実行される。また、プロセスは、人間のアドミニストレータまたは他の人によるアクションが関与する範囲で、一部自動的にそして一部手作業で実行されてもよい。例えば、ある実施形態では、人間が、非クリティカル・データのために予約される（1116）べきアパーチャの最小割合または最大割合を指定してもよい。本明細書において技術革新と考えられるプロセスには、全体的に手作業のものはない。所与の実施形態では、図示するプロセスのステップの内0以上は、恐らくは異なるパラメータまたはデータを用いて、繰り返し動作する場合もある。また、実施形態におけるステップは、図10および図11に展開する、上から下への順序以外の異なる順序で実行することもできる。ステップは、順次、部分的に重複して、または完全に並列に実行することもできる。具体的には、プロセス中に実行されるステップを示すために、フローチャート1000のアクション項目またはフローチャート1100のアクション項目を横断する(traverse)順序は、プロセスの実行毎に変化してもよい。また、フローチャートの横断順序は、プロセスの実施形態毎に変化してもよい。また、実行するプロセスが動作可能であり、少なくとも1つの請求項に従うのであれば、ステップを省略する、組み合わせる、名称変更する、類別し直す、1つ以上の機械上で実行する、または図示するフローから逸脱することも可能である。

【0144】

【00248】 ある実施形態は、コンピューティング・システムにおいてデータ保存方法を使用または提供する。この方法は、バッテリー・バックアップ・メモリーのアパーチャにデータ集合A - データを格納する要求を受けるステップ1002を含むことができる。アパーチャは、フラッシュ・オーダー312を有する。フラッシュ・オーダーは、データ完全性の脅威314に応答して、データをアパーチャからアパーチャ関連不揮発性ストレージにコピーする（908）順序である。フラッシュ・オーダーは、アパーチャのフラッシュ先端と、アパーチャのフラッシュ後端とを定める。A - データは、クリティカル・データに指定されたデータを含む。

【0145】

【00249】 また、この方法は、アパーチャの未割り当てメモリーの一部A - メモリーを識別するステップ1004も含むことができ、A - メモリーはA - データを保持できるだけ十分に大きく、A - メモリーは、アパーチャの他のいずれの未割り当てメモリーの他のいずれのアドレスよりもアパーチャのフラッシュ先端に近いアドレスを有する。また、この方法は、A - メモリーに割り当て済みの印を付けるステップ1006と、A - データのコピーをA - メモリーに入力するステップ1008も含むことができる。

【0146】

【00250】 また、この方法は、データ集合Z - データをアパーチャに格納する要求を受けるステップ1104も含むことができ、Z - データは、クリティカル・データ

10

20

30

40

50

に指定されたデーターを全く含まない。

【0147】

[00251] また、この方法は、アパーチャの未割り当てメモリーの一部 Z - メモリーを特定するステップ 1004 も含むことができ、Z - メモリーは Z - データーを保持するだけ十分に大きく、Z - メモリーは、アパーチャの他のいずれの未割り当てメモリーの他のいずれのアドレスよりもアパーチャのフラッシュ後端に近いアドレスを有する。また、この方法は、Z - メモリーに割り当て済みの印を付けるステップ 1006 と、Z - データーのコピーを Z - メモリーに入力するステップ 1008 とを含むことができる。

【0148】

[00252] このような方法によって、フラッシュ - オーダーにおいて非クリティカル・データーの前にクリティカル・データーをバッテリー・バックアップ・メモリーに入力しない場合とは対照的に、フラッシュ - オーダーにおいて非クリティカル・データーの前にクリティカル・データーをバッテリー・バックアップ・メモリーに入力することによって、クリティカル・データーをフラッシュするのに成功し (1012)、したがってそれを保存する (1102) 可能性 1140 が高くなる (1138)。

【0149】

[00253] ある実施形態では、この方法は、更に、少なくとも部分的に、バッテリー・バックアップ・メモリーのバッテリー特性に基づいて、アパーチャのサイズ 1108 を決定するステップ 1106 も含んでもよい。例えば、容量が小さいバッテリー、古いバッテリー、信頼性が低いバッテリー程、アパーチャが小さくてもよい。ある実施形態では、この方法は、更に、少なくとも部分的に、バッテリー・バックアップ・メモリーのバッテリー特性に基づいて、アパーチャの位置 1110 を決定するステップ 1106 を含んでもよい。例えば、容量が小さい、古い、または信頼性が低いバッテリーによってバックアップされるメモリーにおけるアパーチャに代わって、新しい、容量が大きい、または信頼性が高いバッテリーによってバックアップされるメモリーにおいてアパーチャを選択してもよい。ある実施形態では、この方法は、少なくとも部分的に、バッテリー・バックアップ・メモリーのバッテリー特性における変化に基づいて、例えば、バッテリーがもっと新しい、大きい、または信頼性が高いバッテリーと交換されるときに、アパーチャ・サイズを大きくすることによって、アパーチャのサイズを変更するステップ 1114、またはアパーチャを移動させるステップ 1115 を含んでもよい。本明細書において開示する他の

【0150】

[00254] ある実施形態では、この方法は、更に、アパーチャの少なくともディフラグ部分をディフラグするステップ 1120 を含んでもよく、ディフラグ部分はクリティカル・データーを含む。本明細書の他の章で論じたように、ディフラグging 1120 はフラッシュ・ディフラグging を指す。統合ディフラグging は、思わぬ結果(side-effect)が生じるかもしれないが、フラッシュ・ディフラグging 1120 として機能するには本質的に十分でない。ディフラグging 1120 は積極的(proactive)であってもよい。また、ディフラグging 1120 が誘起されるのは、テナントまたはプロセスが終了するとき、テナントまたはプロセスから抜ける(vocate)とき、または移転するときでもよい。

【0151】

[00255] ある実施形態では、この方法は、更に、ワークロードのクリティカル性、仮想機械の優先度、またはカスタマー・ステータスという判断基準 602 の内、少なくとも 1 つに基づいて、A - データーの少なくとも一部分をクリティカル・データーに指定するステップ 1132 を含んでもよい。

【0152】

[00256] ある実施形態では、この方法は、データー完全性の脅威を検出するステップ 1010 と、クリティカル・データーの全てをアパーチャから不揮発性ストレージにフラッシュするステップ 1012 と、少なくとも一部の非クリティカル・データーをアパーチャから不揮発性ストレージにフラッシュし損なうステップ 1122 とを含んでもよい。例

えば、バッテリーに蓄えられている電力が、アパーチャ全体をフラッシュするには十分ではないかもしれないが、とは言え、全てのクリティカル・データーをフラッシュするには十分であるかもしれない。このシナリオ例は、バッテリー電力が十分であったならば、クリティカル・データーがフラッシュされてから非クリティカル・データーがフラッシュされ終えるように、クリティカル・データーを入力する利点を示す。

【 0 1 5 3 】

【00257】 ある実施形態では、この方法は、データー完全性の脅威を検出するステップ 1 0 1 0 と、少なくともクリティカル・データーの全てをアパーチャから不揮発性ストレージにフラッシュするステップ 1 0 1 2 と、次いでフラッシュしたデーターの全てを不揮発性ストレージから復元するステップ 9 1 0 とを含んでもよい。フラッシュされたクリティカル・データーは、揮発性メモリー、データーがフラッシュされた元のバッテリー・バックアップ・メモリー、またはデーターがフラッシュされた元のバッテリー・バックアップ・メモリーとは異なるバッテリー・バックアップ・メモリーの内少なくとも 1 つに復元する (9 1 0) ことができる。

10

【 0 1 5 4 】

【00258】 ある実施形態では、この方法は、アパーチャ外部、そしてアパーチャ関連不揮発性ストレージ外部に、Z - データーの少なくとも 1 つのコピーを保持するステップ 1 1 2 6 を含んでもよい。即ち、データー保存用システムを構築するとき、非クリティカル・データーの複製は、クリティカル・データーが注意深い入力 (placement) 1 0 0 8 およびフラッシング 1 0 1 2 に依拠することと両立する (compatible)。

20

【 0 1 5 5 】

【00259】 ある実施形態では、バッテリー・バックアップ・メモリーのアパーチャは、ここでは M で示す物理機械内に常駐し (1 1 2 4)、アパーチャに A - データーを格納する要求は、ここでは N で示す異なる物理機械から送られた (1 1 1 2)。例えば、図 2 または図 8 に示す構成では、計算ノードおよびストレージ・ノードは異なる物理機械にすることができる。代替りの構成では、これらは同じ基本物理ハードウェア上で実行する仮想デバイスであることも可能である。

【 0 1 5 6 】

【00260】 ある実施形態では、この方法は、A - データーを 1 つのストレージ・ノードのみに格納し、Z - データーのコピーを複数のストレージ・ノードに格納するステップを含んでもよい。即ち、ある例では、クリティカル・データーの保存は入力 1 0 0 8 およびフラッシング 1 0 1 2 のみに依拠するが、非クリティカル・データーの保存は少なくとも部分的に複製 2 2 6 に依拠する。

30

【 0 1 5 7 】

【00261】 構成記憶媒体 (configured storage medium)

【00262】 ある実施形態は、構成コンピューター読み取り可能記憶媒体 1 1 2 を含む。記憶媒体 1 1 2 は、ディスク (磁気、光、またはこれら以外)、RAM、EEPROMS、または他の ROM、および / または具体的にはコンピューター読み取り可能記憶媒体 (単なる伝搬信号ではない) を含む、他の構成可能なメモリーを含むことができる。構成された記憶媒体とは、具体的に、CD、DVD、またはフラッシュ・メモリーのような、リムーバブル記憶媒体 1 1 4 としてもよい。リムーバブルであってもなくてもよく、揮発性であってもなくてもよい汎用メモリーは、アロケーター 4 2 2、ディフラガー 4 0 8、およびマッピング 8 1 4 というような品目を使用して、データー 1 1 8 および命令 1 1 6 の形態で構成して、実施形態にすることができる。データー 1 1 8 および命令 1 1 6 は、構成記憶媒体を形成するために、リムーバブル記憶媒体 1 1 4 および / またはネットワーク接続のような他のソースから読み取られる。構成記憶媒体 1 1 2 は、コンピューター・システム 1 0 2 に、本明細書において開示した通りに、データー保存のための技術的プロセス・ステップを実行させることができる。このため、図は、構成記憶媒体の実施形態およびプロセス (別名、方法) の実施形態、ならびにシステムおよびプロセスの実施形態を例示するのに役立つ。具体的には、図 9、図 1 0、または図 1 1 において示したプロセス・

40

50

ステップ、もしくは本明細書において教示したプロセス・ステップはいずれも、記憶媒体を構成し、構成記憶媒体の実施形態を形成するために使用することができる。

【0158】

【00263】ある実施形態は、データー118および命令116が構成されたコンピューター読み取り可能記憶媒体112、114を使用または提供する。命令116がプロセッサ110によって実行されると、コンピューティング・システムまたはそのサブシステムに、データー保存のための方法を実行させる。この方法は、複数の要求を受けるステップ1002を含む。各要求は、それぞれのデーター集合をバッテリー・バックアップ・メモリーのアパーチャ220に格納することを求める。アパーチャは、フラッシュ・オーダー312を有する。フラッシュ・オーダー312は、データー完全性の脅威に応答して、アパーチャからアパーチャ関連不揮発性ストレージにデーターをコピーする順序である。フラッシュ・オーダーは、アパーチャのフラッシュ先端806と、アパーチャのフラッシュ後端808とを定める。それぞれのデーター集合は、各々、クリティカル・データーに指定されたデーターを含む。

10

【0159】

【00264】また、この方法は、要求の内少なくとも2つについて、アパーチャの未割り当てメモリーのそれぞれの部分を識別するステップ1004を含む。未割り当てメモリーの各部分は、それぞれのデーター集合を保持するだけ十分に大きく、識別された未割り当てメモリーのそれぞれの部分は、アパーチャの他のいずれの未割り当てメモリーの他のいずれのアドレスよりもアパーチャのフラッシュ先端に近いアドレスを有する。

20

【0160】

【00265】また、この方法は、要求の内少なくとも1つについて、識別された未割り当てメモリーのそれぞれの部分に、割り当て済みの印を付けるステップ1006と、それぞれのデーター集合のコピーをその中に入力するステップ1008とを含む。

【0161】

【00266】また、この方法は、データー完全性の脅威を検出するステップ1010と、アパーチャにコピーされたクリティカル・データーの全てを、アパーチャから不揮発性ストレージにフラッシュするステップ1012とを含む。したがって、この方法は、データー完全性の脅威にも拘わらず、アパーチャにコピーされたクリティカル・データーの全てを保存する(1102)。

30

【0162】

【00267】ある実施形態では、この方法は、更に、アパーチャ・サイズ1108またはアパーチャ・アドレス1110の内少なくとも1つを指定するステップ1106を含み、この指定は、バッテリー容量502、バッテリーの経年数506、またはバッテリーの信頼性値504の内少なくとも1つに基づく。

【0163】

【00268】ある実施形態では、フラッシュするステップ1012は、統一拡張可能ファームウェア・インターフェースを有するファームウェア310を実行するステップ1142を含む。

【0164】

40

【00269】ある実施形態では、この方法は、更に、クリティカル・データーに指定されていないデーターを保持するために、アパーチャの内少なくとも10パーセント1118を予約するステップ1116を含む。たとえば、20パーセントを予約すると、または30パーセントを予約する場合、各々、少なくとも10パーセントを予約する(1116)条件を十分に満たす(qualify as)。ある実施形態では、予約(1116)される部分1118は、せいぜい10パーセント、またはせいぜい20パーセントである。また、予約1116は、アパーチャまたはメモリー・パーティションをサブパーティションに分割することと見なす、または分割するように実施することもできる。

【0165】

【00270】ある実施形態では、この方法は、更に、クリティカル・データーに指定され

50

たデーターの空間を作るために、クリティカル・データーに指定されていないデーターを、アパーチャ外部の記憶位置に移動させるステップ 1 1 2 8 を含む。ある実施形態では、この方法は、アパーチャの少なくとも一部分をディフラグするステップ 1 1 2 0 を含み、これによって、アパーチャの 2 つの割り当て済み部分間に以前に位置していた未割り当て部分を削減する (1 1 3 4) または消滅させる (1 1 3 6)。

【 0 1 6 6 】

[00271] 追加の例および観察事項

[00272] 尚、本開示のあらゆる部分、またはその中のあらゆる特定の詳細が、実施可能性、書面による説明、または最良の態様というような、法的判断基準を必ずしも満たす必要はないことは、当業者には認められよう。また、実施形態は、特定のネットワーク、プロトコル、ツール、識別子、フィールド、データー構造、関数、秘密または他の証明 (secrets or other proofs)、あるいは本明細書において説明した他の実施態様の選択肢には限定されない。他のいずれの特許開示との明らかな矛盾は、本技術革新の所有者からであっても、本特許開示において提示される請求項を解釈するには全く役には立たない。これは本開示の全ての部分に関係することを理解した上で、いくつかの追加の例および観察事項を提供する。

【 0 1 6 7 】

[00273] ある実施形態では、揮発性メモリー 4 0 2 はバッテリー 2 1 6 によってバックアップされ、データーの完全性が脅かされるとき (3 1 4) (たとえば、差し迫った電力逸失またはリブートによって)、データー 1 1 8 のコピーを不揮発性ストレージ 9 0 2 において作成するために、素早くフラッシュする (1 0 1 2) ことができる。フラッシング 1 0 1 2 は、フラッシュ・オーダー 3 1 2 によって制約されるので、一部のデーターが安全な場所にフラッシュされないおそれがある。例えば、バッテリーの故障または消耗の結果、フラッシュが不完全になることもある。フラッシュは、データー・ライフサイクル処理 (data lifecycle operation) の一例であり、他のいくつかの例には、割り当て、使用、複製、割り当て解除、移動、ディフラグ、および復元がある。

【 0 1 6 8 】

[00274] ある実施形態は、十分に大きなバッテリーをプロビジョニングするという目標を掲げて、設計および実装される。したがって、アパーチャ・サイズは、バッテリーが保護できるデーター 1 1 8 の量に依存するとしてよく、このデーター量はアーキテクトまたは開発者が決断する。したがって、バッテリーが大きい程、バッテリーが新しい程、バッテリーの信頼性が高い程、大きなアパーチャが得られ、逆に、バッテリーが小さい / 古い / 信頼性が低い程、またはバッテリー特性が不明である場合、アパーチャは小さくなるのはもっともである。また、バッテリーは、ある時間期間の経過に連れて、出力レベルが下がり (de-rated) / 劣化するおそれがあるので、大量のデーターを退避させる能力は、時間の関数として影響を受ける可能性がある。アパーチャ 2 2 0 は、それに応じてサイズを変更することができる。

【 0 1 6 9 】

[00275] ある実施形態では、混成配置命令 (hybrid placement order) を実施し、これによって、クリティカル・データーはアパーチャのフラッシュ先端 8 0 6 またはその近くに集められ (group)、非クリティカル・データーはアパーチャのフラッシュ後端 8 0 8 またはその近くに集められる。非クリティカル・データーまたはクリティカル・データー (あるいは双方) は、他のアパーチャに、またはバッテリー・バックアップされたいずれのアパーチャの外部にある他のメモリーに複製することも可能である。

【 0 1 7 0 】

[00276] 本明細書において教示したデーター保存技術の存在は、場合によっては、文書化から、メモリー・マップ 8 1 4 から、データー・クリティカル性についての情報と組み合わせられたメモリー・ダンプから、機械のプロープまたはそれらの間における通信から、合同検査アクション・グループ (J T A G) のデバッグングから、 N V D I M M - S W 退避 / 復元検査から、ファームウェアまたは他のコード・リストから、他の証拠から、あ

10

20

30

40

50

るいはこれらの組み合わせから推定することもできる。

【0171】

【00277】 ある実施形態は、NVDIMM-SW退避アルゴリズムを提供または使用する。このアルゴリズムは、ワークロードのクリティカル性612を、複数のNVDIMM-SWデータ集合を順位付けて並び替える方法として織り込む(factor in)ことができる。ある実施形態は、仮想機械ACPIアドレス範囲によって記述され、仮想機械間およびSCMパーティションにマッピングする(222、224)ストレージ・クラス・メモリー(SCM)パーティション904を実装する。これは、システム・ファームウェアのランタイム・コンポーネントにおいて、例えば、Intel/AMD社のCPUについてはUEFI/SMMにおいて、そしてARM社のCPUについてはUEFI/Trust Zonesにおいて、NVDIMM-SW退避ルーティング(routing)を誘導する(guide)10。ある実施形態は、任意のグレースフルまたはアングレースフル・シャットダウンにおいてコンテンツ118を保護し(908)、仮想機械のコンテンツ118を、次のブート時に復元する(910)ことを可能にする。ある実施形態では、NVDIMM-SWは、データへの素早いアクセスを確保し易くし、同時に、ノードがオフラインになるというようなサービスの混乱という事象においても、不揮発性ストアへのデータ・フラッシュを確保するのに役立つ。

【0172】

【00278】 ある実施形態では、フラッシュを誘起する可能性があるイベントには、例えば、順序正しいシャットダウン、ハード・リブート、またはソフト・リブートを含むことができる。20

【0173】

【00279】 順序正しいシャットダウンは、実行中のソフトウェアに、ディレクトリー、割り当て表、およびデータの編成を記述する他の構造の現行バージョンを退避することによって、データ改変を防止し、更にデータ自体を退避する機会を与える。例えば、シャットダウン・コマンドは、アプリおよび他のプロセスに、シャットダウンするように命令し、次いでこれらのプロセスに、データを不揮発性メモリーにフラッシュし、開いているあらゆるファイルを閉じ、割り当てられているメモリーをOSに戻し開放する機会を与える。また、シャットダウン・コマンドは、デバイス・ドライバに、IOデータおよび現在のディレクトリー情報を、付随する(attached)デバイスにフラッシュするように命令する。ACPI準拠システムでは、シャットダウン・コマンドが、給電(Power)コマンドの発行を行わせることもでき、給電コマンドは、NVDIMMに、その揮発性部分からその不揮発性部分にデータを退避させる(908)。30

【0174】

【00280】 ハード・リブートは、おそらく電源ボタン以外では、無電力で開始する。電力がシステムに与えられる。ブート・メモリー・コード(BIOS/EFI/UEFI)がPOSTを実行する。次いで、ブート・メモリー・コードがブートローダーをRブート・デバイスからAMにロードする。ブートローダーは、不揮発性作業メモリーからまたはネットワーク接続を介して、OSをRAMにロードする。ハード・リブートが開始したときにRAM内にあったものはいずれも、潜在的に上書きされる可能性がある。40

【0175】

【00281】 ソフト・リブートは、システムへの給電から開始する。電源投入時の自己検査(POST: Power On Self-Test)を飛ばす。ブート・メモリー・コードは、ブートローダーをRAMにブート・デバイスからロードする。ブートローダーは、不揮発性作業メモリーから、またはネットワーク接続を介して、OSをRAMにロードする。ソフト・リブートが開始したときにRAM内にあったものはいずれも、潜在的に上書きされる可能性がある。

【0176】

【00282】 ある実施態様では、OSが、OSブートの後UEFIランタイム機能のために保存するメモリー領域904を特定するため、およびOS使用のために取り戻すエリア50

を特定するために、システム・メモリー・マップを使用する。ある実施態様では、メモリー・マップは、UEFIプリブートの全てのフェーズにわたって展開される。コードおよびデーターを格納するためにメモリーが割り当てられるに連れて、および割り当てられるときに、メモリー・マップは変化する。メモリー割り当ての形式は、様々であり、コードおよびデーターの型に基づく。例えば、ブート・サービス・コード、ブート・サービス・データー、ランタイム・コード、ランタイム・データー、ACPI NVSメモリー、SMM等に基づく。

【0177】

[00283] ある実施形態では、NVDIMM-SWが存在し、ファームウェア補助手段(firmware assisted means)を提供し、このファームウェア補助手段によって、DIMM およびブロック・ストレージ・デバイス(NVMeまたはSATA等)の結合を通じて、不揮発性DIMMをエミュレートする。目標は、標準主導の手法を採用すること、つまりNVDIMM-SWデバイスをオペレーティング・システムにシームレスに公開する(publish)ことである。通常の動作モードでは、プリブート・ファームウェアは、利用可能なNVDIMM-SWデバイスをオペレーティング・システムに公開する。NVDIMM-SWを活用するアプリケーションからの全てのWRITE動作は、NVDIMM-SWと関連付けられたDIMM領域を目標とする。SAVE#は、グレースフルまたは非グレースフルいずれかのシャットダウンのシナリオの間に開始される。SAVE#動作は、システム・シャットダウンおよびその後の再開の前に完了する。RESTORE#動作は、制御がオペレーティング・システムに戻される前に、プリブート・ファームウェアによって開始される。NVDIMM-SWは、DDR4DIMMおよびオンボードのM.2 NVMeモジュールを結合することによって構築することができる。NVDIMM-SWパラダイムは、バイト・アドレス可能なメモリーをエミュレートするために、従前からのDIMMを再利用する。ある実施態様では、システムには、同じサイズおよび同じメモリー型の全てのDIMMが装備される(populate)。NVDIMM-SWは、2つの動作モード、即ち、NVDIMM非インターリーブまたはNVDIMMインターリーブのいずれかをサポートすることができる。NVDIMM非インターリーブでは、設定選択肢を通じて選択される不揮発性メモリーのサイズに応じて、スロット1上のDIMMが不揮発性NVDIMM(SW)NVDIMMとして選択される。最大不揮発性メモリー・サイズは、非グレースフルな退避のシナリオ中における電力持続期間に依存する。非インターリーブNVDIMM-SWは、NVDIMM-Nと全く同じであるが、全てのI2C(シリアル・プロトコル) __DSM(デバイス特定方法)データーが代わりにエミュレートされることを除く。NVDIMMインターリーブでは、NUMAをサポートするために、ソケット内部にある全てのDIMMが一緒にインターリーブされる。また、各ソケットの最上位メモリーは、選択された不揮発性メモリー・サイズに基づいて、システム・メモリー・マップから、NVDIMM-SWとして区分けされる(carved out)。イベント処理(event handling)の目標は、データーを揮発性メモリーから不揮発性メモリー領域に退避しようと試みることである。また、ファームウェア・スタック(UEFI、BMC)は、帯域内(inband)および帯域外リスナー(out-of-band listeners)双方についてのエラーを記録する役割も担う。

【0178】

[00284] いくつかの追加の組み合わせおよび変形

[00285] コード、データー構造、ロジック、コンポーネント、通信、および/またはそれらの機能的同等物のこれらの組み合わせの内任意のものを、以上で説明したシステムおよびそれらの変形(variation)の内任意のものと組み合わせることもできる。プロセスは、本明細書において説明した任意のステップを、動作可能な任意の部分集合または組み合わせまたはシーケンスに含めることもできる。各異形(variant)は、単独で、または他の異形の任意の1つ以上との組み合わせで行われてもよい。各異形は、前述のプロセスの内任意のものとともに行われてもよく、各プロセスは、他のプロセスの任意の1つ以上と組み合わせることもできる。各プロセスまたはプロセスの組み合わせは、異形を含んで、以上

10

20

30

40

50

で説明した構成記憶媒体の組み合わせおよび異形の内任意のものと組み合わせることもできる。

【0179】

[00286] 結論

[00287] 端的に言うと、本明細書において提示した教示は、コンピューティング・システムにおいてデータ保存機能を強化する(enhance)ために適用することができる。ある実施形態では、動作ステップとデバイス特性とを組み合わせることによって、完全性の脅威314に対抗してデータ118を保存し易くする。データ118は、カスタマー606の要求608、ワークロード610のクリティカル性612、または仮想機械206のクリティカル性604というような判断基準602に基づいて、クリティカル・データ412および非クリティカル・データ416に分割される。データ118は、計算ノード202において生成し、例えば、ストレージ・ノード212における格納のために送る(222)ことができる。クリティカル・データ412は、バッテリー・バックアップ・メモリー218のアパーチャ220において、物理アドレス424に格納され、ここから、フラッシュ・オーダー312にしたがって非クリティカル・データ416よりも前にフラッシュされる(1012)。フラッシュ・オーダー312は、バッテリー・バックアップ・メモリーの回路によって課せられ(impose)、または外部の力、例えば、システム・ファームウェア420によってバッテリー・バックアップ・メモリーに課せられる。フラッシュ・オーダー312は、例えば、ボトムアップNVDIMM404のフラッシュ・オーダーでもよい。複製されたコピーを保存する(1102)ためにそれをフラッシュする(1012)のに間に合わなかった場合、データ118の冗長コピー226(特に、非クリティカル・データ416)も保持する(1126)ことができる。バッテリー・バックアップ・メモリーのアパーチャ220は、バッテリーの特性406にしたがって、大きさが決められ(1106)そして配置され(1106)、条件が変化することに連れて、配置換え(1115)または大きさの決め直し(1114)を行うこともできる。フラッシュ・ディフラグging1120は、特に、クリティカル・データ412を保持するアパーチャの部分414内において、アパーチャ220の使用を最適化するために行われる。

【0180】

[00288] また、実施形態は、一般データ保護規則(GDPR: General Data Protection Regulation)のような、検査済みのおよび該当するセキュリティ管理およびプライバシー管理も含む、またはこれらから恩恵を受けると考えられる。例えば、マルウェアの注入または有効化(activation)によるコンピューティング・システムの悪用を防止するのに役立てるため、そして増強システムがプログラム実行中に処理するかもしれないいかなる個人またはプライベート情報も改竄されるのを防止するのに役立てるために、しかるべき対策を講じなければならないことは、暗黙裏に了解されよう。本明細書において教示したツールおよび技法の使用は、このような管理の使用と両立することができる。

【0181】

[00289] 本明細書では、特定実施形態について、プロセスとして、構成記憶媒体として、またはシステムとして、明示的に図示および説明したが、1つの種類の実施形態の論述が他の実施形態の種類にも総合的に広がることは認められよう。实例をあげると、図9、図10、および図11と関連付けたプロセスの説明は、構成記憶媒体を説明するのにも役立ち、更に他の図と関連付けて論じられたものと同様に、システムおよび製造の技術的効果および運用(operation)について説明するのにも役立つ。だからと言って、1つの実施形態からの限定が必ずしも他のものに読み込まれるということにはならない。具体的には、プロセスは、システムまたは構成メモリーのような製品について論じる間に提示されたデータ構造および構成に必ずしも限定されない。

【0182】

[00290] 尚、実装の詳細は、特定のAPI、特定のフィールド、特定の種類のコンポーネント(ハードウェアおよびソフトウェア)、および特定のサンプル・プログラムのよ

10

20

30

40

50

うな特定のコードに関連があるとしてよく、したがって全ての実施形態に必ずしも現れる訳ではないことは、当業者には理解されよう。また、詳細を論じるときに使用されたプログラムの識別子および何らかの他の用語は、実施態様に特定のであり、したがって必ずしもあらゆる実施形態に関連がある訳ではないことも、当業者には理解されよう。しかしながら、必ずしもここに提示する必要はないが、このような詳細は、コンテキストを提供することによって、読者を助けることができ、および／または本明細書において論じた技術の多くの可能な実施態様の内いくつかを例示することができる。

【0183】

[00291] 全ての特許請求した実施形態または特許請求可能な実施形態を例示するが包括的でない技術的プロセス、技術的効果、技術的メカニズム、および技術の詳細を含む、本明細書において提示した項目に十分な注意を払えば、本開示および本明細書において説明した実施形態は、専門的技能(technical art)以外の主題を対象とするのではなく、原理または本来の原因もしくは動機のようなそれ自体のいずれの理念(any idea of itself)も、単なる結果自体も、心理作用または心的処理過程も、ビジネス・モデル(business method)または普及している経済的慣行も、人間の活動を組織化する単なる方法も、自然法則自体も、当然起こることまたはプロセスも、生体または生体の一部も、数式自体も、分離されたソフトウェア(isolated software)自体も、単に従来からのコンピューターも、全体的に知覚できないあらゆるものまたはあらゆる抽象的概念自体も、無意味な解決後の活動(post-solution activities)も、指定されていない装置上に全体的に実装されたあらゆる方法も、有用で具体的である結果を生み損ねたあらゆる方法も、全ての利用分野のあらゆる先の占有(preemption)も、特許保護が求められるまたは認可もしくは施行されている司法管轄地域の法律の下において特許保護には不適格なあらゆる他の主題も対象としないことは、当業者には理解されよう。

【0184】

[00292] 本明細書におけるある特徴Xを有する実施形態に対する引用、および本明細書の他の場所におけるある特徴Yを有する実施形態に対する引用は、特徴Xおよび特徴Y双方を有する実施形態を、この開示から除外することにはならない。但し、そのような除外が本明細書において明示的に述べられる場合を除く。全ての可能な否定的な請求項の限定は、実施形態の一部であると規定された特徴はいずれも、他の実施形態における包含から明示的に除去されてもよいという意味で、その特定の除外が本明細書におけるいずれの例にも示されなくても、本開示の範囲内に該当する。「実施形態」という用語は、本明細書では、「該当法規を遵守する形で適用される場合のプロセス、システム、製品、構成コンピューター読み取り可能記憶媒体、および／または本明細書における教示の他の例の」、より便利な形態として使用されるに過ぎない。したがって、所与の「実施形態」は、その実施形態が少なくとも1つの請求項と一致するのであれば、本明細書において開示した特徴の任意の組み合わせを含んでもよい。

【0185】

[00293] 図に示されるあらゆる品目が、必ずしもあらゆる実施形態に存在する訳ではない。逆に、実施形態が、図に明示的に示されない品目(1つまたは複数)を含むこともある。ここでは、本文および図面において、具体的な例によって、いくつかの可能性を示したが、実施形態はこれらの例から逸脱してもよい。实例をあげると、ある例の特定の技術的効果または技術的特徴を、省略する、改名する、異なるグループに分ける、繰り返す、ハードウェアおよび／またはソフトウェアにおいて異なってインスタンス化するのでもよく、あるいはこれらの例の2つ以上に現れる効果または特徴の混合であってもよい。1つの位置において示される機能が、ある実施形態における異なる位置に設けられてもよく、全体的に見て相互作用するモジュールの集合体から所望の技術的効果を必ずしも除外することなく、所与の実装において機能性モジュールを種々の方法で定義できることは、当業者には認められよう。空間の制限のためまたは便宜上、図では別個のステップを1つのボックス内に一緒に示す場合もあるが、別々に実行可能であるのはもっともである。例えば、方法の所与の実行において、1つのステップを実行し、他のステップを実行しなくて

10

20

30

40

50

もよい。

【0186】

[00294] 全体を通じて、参照番号によって図面を参照した。図または本文において、所与の参照番号に関連する言葉遣いにおける明らかな矛盾はいずれも、その参照番号によって参照されるものの範囲を単に広げるものとして理解されてしかるべきである。同じ参照番号が使用されても、所与の参照番号の異なるインスタンスが、異なる実施形態を指すこともある。同様に、所与の参照番号が、動詞、名詞、および/または各々の対応するインスタンスを指すために使用されることもある。例えば、プロセッサ 110 が、命令を実行することによって、それら进行处理する(110)ことができる。

【0187】

[00295] 本明細書において使用する場合、「a」、「an」、および「the」のような用語は、示される項目またはステップの1つ以上を含むものとする。具体的には、請求項において、ある項目を引用するときは、通常、少なくとも1つのそのような項目が存在することを意味し、そしてあるステップを引用するときは、そのステップの少なくとも1つのインスタンスが実行されることを意味する。同様に、「です」(is)および他の動詞単数形は、文法的な誤りや誤解を避けるために、コンテキストが許すときには、「です」(are)および他の複数形の可能性も包含することは理解されてしかるべきである。

【0188】

[00296] 見出しは便宜上に過ぎない。所与の話題についての情報は、見出しがその話題を示す章以外で見出されることもある。

[00297] 提出された全ての請求項および要約書は、明細書の一部である。

【0189】

[00298] 本明細書において使用したいいずれかの用語が業界標準を暗示するまたは業界標準に言及する限りにおいて、そして適用法がそのような標準の特定のバージョンの識別を要求する限りにおいて、本開示は、適用される特許法の下における本開示の最も速い優先日において、少なくとも草稿の形態で(もっと新しいバージョンがある場合、最終形態が優先される)公開されている標準の最新バージョンに言及することは理解されてしかるべきである。

【0190】

[00299] 以上、例示的な実施形態について図面に示し以上で説明したが、請求項に明記した原理および概念から逸脱することなく、多数の変更を行うことができること、そしてこのような変更が必ずしも抽象的な概念全体を包含しないことは、当業者には明白であろう。主題について、構造的特徴および/または手順的アクトに特定の文言で説明したが、添付する請求項において定められる主題は、必ずしも、請求項の上で説明した特定の技術的特徴やアクトに限定されないことは理解されよう。所与の定義または例において識別されたあらゆる手段または態様または技術的效果が、あらゆる実施形態に存在することも、あらゆる実施形態において利用されることも必要でない。逆に、説明した特定の特徴およびアクトならびに効果は、請求項を実施するときの考慮のために、例として開示されたまでである。

【0191】

[00300] 全体の抽象的な理念を包み込むには足りないが請求項の均等の意味および範囲に該当する全ての変更は、法律によって認められる最大限の範囲まで、それらの範囲内に受け入れられるものとする。

10

20

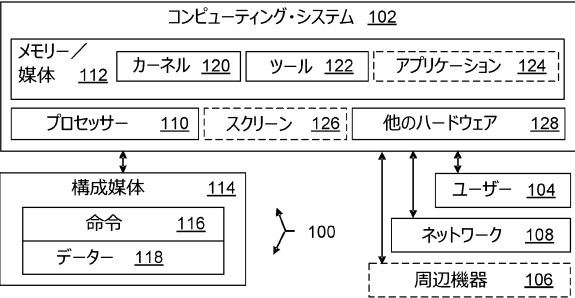
30

40

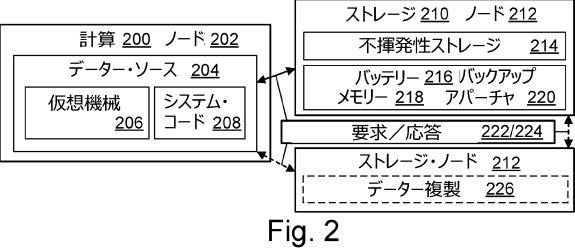
50

【 図 面 】

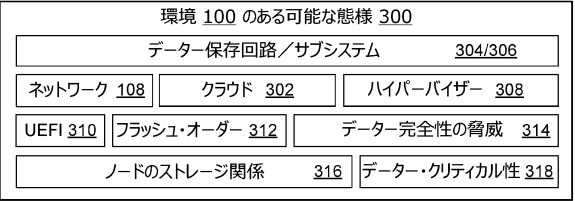
【 図 1 】



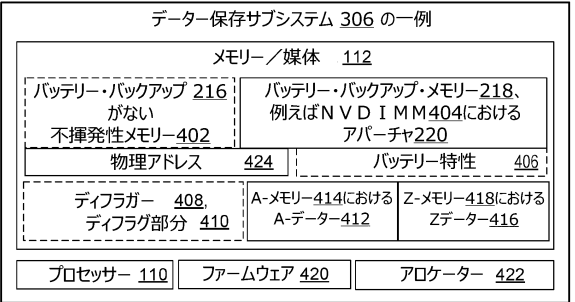
【 図 2 】



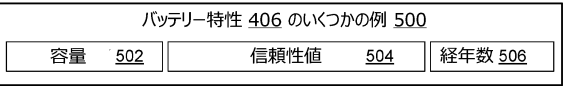
【 図 3 】



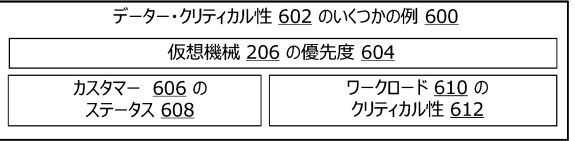
【 図 4 】



【 図 5 】



【 図 6 】



10

20

30

40

50

【 図 7 】



Fig. 7

【 図 8 】

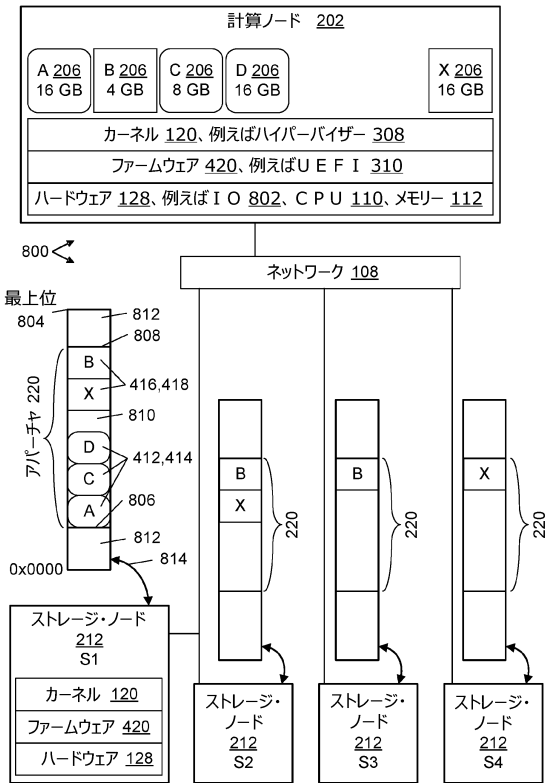


Fig. 8

【 図 9 】

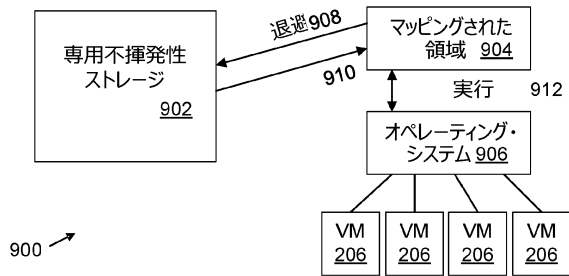


Fig. 9

【 図 10 】

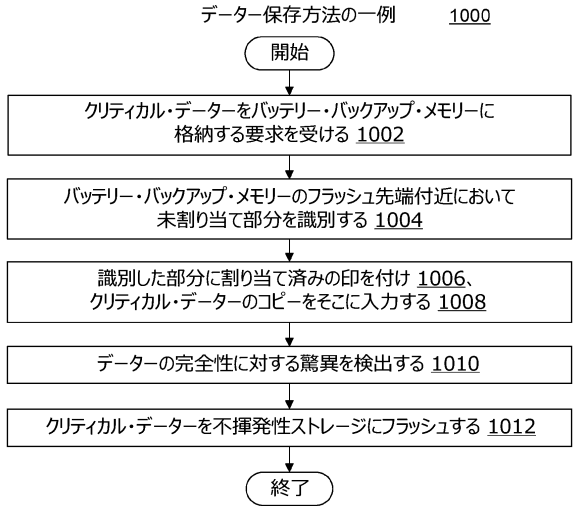


Fig. 10

【 図 1 1 】



Fig. 11

フロントページの続き

- アメリカ合衆国 ワシントン州 98052-6399 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 テクノロジー ライセンシング, エルエルシー
- (72)発明者 グエン, トム・エル
- アメリカ合衆国 ワシントン州 98052-6399 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 テクノロジー ライセンシング, エルエルシー
- (72)発明者 ラドカニ, ニーラジ
- アメリカ合衆国 ワシントン州 98052-6399 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 テクノロジー ライセンシング, エルエルシー
- (72)発明者 マイソール, シェンタマーシー・ラビ
- アメリカ合衆国 ワシントン州 98052-6399 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 テクノロジー ライセンシング, エルエルシー
- 審査官 柳 谷 侑
- (56)参考文献 米国特許出願公開第2015/0106410 (US, A1)
- 特開2013-047979 (JP, A)
- 特開2011-005683 (JP, A)
- 特開2008-262451 (JP, A)
- 特開2006-277411 (JP, A)
- 特開2004-080405 (JP, A)
- 国際公開第2011/007599 (WO, A1)
- 国際公開第2000/050997 (WO, A1)
- (58)調査した分野 (Int.Cl., DB名)
- G06F 12/14
- G06F 21/10
- G06F 21/60 - 21/88