

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7387873号
(P7387873)

(45)発行日 令和5年11月28日(2023.11.28)

(24)登録日 令和5年11月17日(2023.11.17)

(51)国際特許分類 F I
 G 0 6 F 12/1081(2016.01) G 0 6 F 12/1081
 G 0 6 F 12/02 (2006.01) G 0 6 F 12/02 5 7 0 M

請求項の数 20 (全19頁)

(21)出願番号	特願2022-509631(P2022-509631)	(73)特許権者	591016172
(86)(22)出願日	令和2年8月21日(2020.8.21)		アドバンスト・マイクロ・デバイス ・インコーポレイテッド
(65)公表番号	特表2022-544791(P2022-544791 A)		ADVANCED MICRO DEVI CES INCORPORATED
(43)公表日	令和4年10月21日(2022.10.21)		アメリカ合衆国 9 5 0 5 4 カリフォル ニア州、 サンタ クララ、 オーガスティ ン ドライブ 2 4 8 5
(86)国際出願番号	PCT/US2020/047376	(73)特許権者	508301087
(87)国際公開番号	WO2021/035134		エーティーアイ・テクノロジーズ・ユー エルシー
(87)国際公開日	令和3年2月25日(2021.2.25)		ATI TECHNOLOGIES ULC
審査請求日	令和5年8月21日(2023.8.21)		カナダ、 オンタリオ エル3ティー 7 エックス6、 マーカム、 コマース パリー ドライブ イースト 1
(31)優先権主張番号	16/548,692		
(32)優先日	令和1年8月22日(2019.8.22)		
(33)優先権主張国・地域又は機関	米国(US)		
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 ネットワーク化された入出力メモリ管理ユニット

(57)【特許請求の範囲】

【請求項1】

ネットワーク化された入出力メモリ管理ユニット(IOMMU)を含む装置であって、
 前記ネットワーク化されたIOMMUは、複数のIOMMUを含み、
 前記ネットワーク化されたIOMMUは、
 第一アドレス変換層によって生成されたドメイン物理アドレスを含むメモリアクセス要
 求を受信することと、
 前記複数のIOMMUのうち、前記メモリアクセス要求を生成したデバイスのタイプに
 基づいて選択された何れかのIOMMUを使用して、前記ドメイン物理アドレスをシステ
 ムメモリ内の物理アドレスに選択的に変換することと、
 を行うように構成されている、
 装置。

【請求項2】

前記デバイスは、グラフィックスプロセッシングユニット(GPU)、又は、複数の周
 辺デバイスのうち何れかの周辺デバイスであり、
 前記複数のIOMMUは、
 前記メモリアクセス要求を前記第一アドレス変換層から受信するように構成された一次
 IOMMUと、
 前記一次IOMMUに接続され、前記デバイスに関連する回路に近接して配置された二
 次IOMMUと、を含む、

請求項 1 の装置。

【請求項 3】

前記一次 I O M M U は、前記メモリアクセス要求が前記 G P U から受信されたことに
応じて、前記ドメイン物理アドレスのアドレス変換を実行し、

前記二次 I O M M U は、前記メモリアクセス要求が前記周辺デバイスから受信されたこ
とに応じて、前記アドレス変換を実行する、

請求項 2 の装置。

【請求項 4】

前記一次 I O M M U は、前記一次 I O M M U に関連する第一トランシェーションルック
アサイドバッファ (T L B) 及び第一ページテーブルセットを使用してページテーブルウ
ォークを実行することによって、前記 G P U から受信した前記ドメイン物理アドレスのアド
レス変換を実行する、

請求項 3 の装置。

【請求項 5】

前記一次 I O M M U は、前記メモリアクセス要求が前記周辺デバイスから受信されたこ
とに応じて、前記メモリアクセス要求を前記二次 I O M M U に提供し、

前記二次 I O M M U は、前記二次 I O M M U に関連する第二 T L B 及び第二ページテー
ブルセットを使用して前記ページテーブルウォークを実行することによって、前記一次 I
O M M U から受信した前記メモリアクセス要求内の前記ドメイン物理アドレスのアドレス
変換を実行する、

請求項 4 の装置。

【請求項 6】

前記周辺デバイスに関連する回路に対する前記二次 I O M M U の位置は、前記周辺デバ
イスのレイテンシ要件に基づいて決定される、

請求項 2 の装置。

【請求項 7】

前記一次 I O M M U に接続された複数の二次 I O M M U をさらに含み、

前記複数の二次 I O M M U は、複数の周辺デバイスに関連する回路に近接して配置され
ており、

前記複数の二次 I O M M U の各々は、前記複数の周辺デバイスのうち対応する周辺デバ
イスに関連する回路から受信したメモリアクセス要求内のドメイン物理アドレスのアドレ
ス変換を実行するように構成されている、

請求項 6 の装置。

【請求項 8】

前記複数の二次 I O M M U は、前記複数の周辺デバイスに関連する回路内に集積されて
いる、

請求項 7 の装置。

【請求項 9】

メモリアクセス要求を前記第一アドレス変換層から受信し、前記メモリアクセス要求を
生成した前記デバイスのタイプに基づいて、前記メモリアクセス要求を前記一次 I O M M
U 又は前記二次 I O M M U に選択的に提供するように構成されたコマンドキューをさらに
含む、

請求項 2 の装置。

【請求項 10】

複数の入出力メモリ管理ユニット (I O M M U) を含むネットワーク化された I O M M
U において、第一アドレス変換層によって生成されたドメイン物理アドレスを含むメモリ
アクセス要求を受信することと、

前記メモリアクセス要求を生成したデバイスのタイプに基づいて、前記複数の I O M M
U のうち何れかの I O M M U を選択することと、

前記複数の I O M M U のうち選択された I O M M U を用いて、前記ドメイン物理アドレ

10

20

30

40

50

スをシステムメモリ内の物理アドレスに選択的に変換することと、を含む、
方法。

【請求項 1 1】

前記デバイスは、グラフィックスプロセッシングユニット（GPU）、又は、複数の周辺デバイスのうち何れかの周辺デバイスであり、

前記メモリアクセス要求を受信することは、前記複数の I O M M U 内の一次 I O M M U において、前記第一アドレス変換層から前記メモリアクセス要求を受信することを含む、
請求項 1 0 の方法。

【請求項 1 2】

前記複数の I O M M U のうち何れかの I O M M U を選択することは、

前記デバイスが前記 GPU であることに応じて、前記一次 I O M M U を選択することと、
前記デバイスが前記複数の周辺デバイスのうち何れかの周辺デバイスであることに応じて二次 I O M M U を選択することであって、前記二次 I O M M U は、前記一次 I O M M U に接続され、前記複数の周辺デバイスのうち何れかの周辺デバイスに関連する回路に近接して配置されている、ことと、を含む、

請求項 1 1 の方法。

【請求項 1 3】

前記メモリアクセス要求が前記 GPU から受信されたことに応じて、前記一次 I O M M U において、前記ドメイン物理アドレスのアドレス変換を実行することをさらに含む、

請求項 1 2 の方法。

【請求項 1 4】

前記一次 I O M M U において、前記ドメイン物理アドレスのアドレス変換を実行することは、前記一次 I O M M U に関連する第一トランスレーションルックアサイドバッファ（TLB）及び第一ページテーブルセットを使用してページテーブルウォークを実行することを含む、

請求項 1 3 の方法。

【請求項 1 5】

前記メモリアクセス要求が前記周辺デバイスから受信されたことに応じて、前記二次 I O M M U において前記アドレス変換を実行することをさらに含む、

請求項 1 3 の方法。

【請求項 1 6】

前記メモリアクセス要求が前記周辺デバイスから受信されたことに応じて、前記メモリアクセス要求を前記一次 I O M M U から前記二次 I O M M U に提供することをさらに含み、
前記二次 I O M M U において仮想アドレスのアドレス変換を実行することは、前記二次 I O M M U に関連する第二 TLB 及び第二ページテーブルセットを使用してページテーブルウォークを実行することを含む、

請求項 1 5 の方法。

【請求項 1 7】

コマンドキューにおいて、前記第一アドレス変換層からのメモリアクセス要求を受信することと、

前記メモリアクセス要求を生成した前記デバイスのタイプに基づいて、前記メモリアクセス要求を前記コマンドキューから前記一次 I O M M U 又は前記二次 I O M M U に選択的に提供することと、をさらに含む、

請求項 1 2 の方法。

【請求項 1 8】

グラフィックスプロセッシングユニット（GPU）と、少なくとも 1 つの周辺デバイスと、メモリと、に接続されるように構成されたネットワーク化された入出力メモリ管理ユニット（I O M M U）であって、

前記ネットワーク化された I O M M U は、

メモリアクセス要求を第一アドレス変換層から受信するように構成されたコマンドキュー

10

20

30

40

50

一であって、前記メモリアクセス要求は、前記第一アドレス変換層によって生成されたドメイン物理アドレスを含む、コマンドキューと、

前記メモリアクセス要求が前記GPUから受信されたことに応じて、前記ドメイン物理アドレスを前記メモリ内の物理アドレスに変換するように構成された一次IOMMUと、

前記メモリアクセス要求が前記少なくとも1つの周辺デバイスから受信されたことに応じて、前記ドメイン物理アドレスを前記メモリ内の物理アドレスに変換するように構成された少なくとも1つの二次IOMMUと、を含む、

ネットワーク化されたIOMMU。

【請求項19】

前記一次IOMMUは、前記一次IOMMUに関連する第一トランスレーションルックアサイドバッファ(TLB)及び第一ページテーブルセットを使用してページテーブルウォークを実行することによって、前記GPUから受信した前記ドメイン物理アドレスのアドレス変換を実行し、

10

前記少なくとも1つの二次IOMMUは、前記二次IOMMUに関連する第二TLB及び第二ページテーブルセットを使用して前記ページテーブルウォークを実行することによって、前記メモリアクセス要求内の前記ドメイン物理アドレスの変換を実行する、

請求項18のネットワーク化されたIOMMU。

【請求項20】

前記少なくとも1つの周辺デバイスに対する前記少なくとも1つの二次IOMMUの少なくとも1つの位置は、前記少なくとも1つの周辺デバイスのレイテンシ要件に基づいて決定される、

20

請求項18のネットワーク化されたIOMMU。

【発明の詳細な説明】

【背景技術】

【0001】

従来のグラフィックスプロセッシングシステムについてのプラットフォームは、中央処理装置(CPU)、グラフィックスプロセッシングユニット(GPU)、1つ以上のシステムメモリ(ダイナミックランダムアクセスメモリ、DRAM等)、及び、これらのエンティティ間の通信をサポートするバスを含む。場合によっては、プラットフォームは、システムオンチップ(SoC)として実装される。CPUは、GPUにドローコールを発行することにより、グラフィックス処理を開始する。ドローコールを受信することに応じて、GPUは、プログラマブルシェーダ及び固定機能ハードウェアブロックのシーケンスから形成されるパイプラインを使用して、表示用の画像をレンダリングする。従来のグラフィックスプロセッシングシステムにおけるシステムメモリは、グラフィックスプロセッシングシステム上で実行するホストオペレーティングシステム(OS)に可視である第一部分と、GPU専用である(例えば、フレームバッファを提供する)第二部分と、にパーティション化される。第二部分は、時としてカーブアウト(carveout)又はGPUカーブアウトと呼ばれ、ホストOSに可視ではない。GPU仮想マネージャ(VM)は、グラフィックスデバイスドライバによって管理されることで、メモリアクセス要求内の仮想アドレスを、システムメモリのGPUカーブアウト領域内の物理アドレス等のような、システムメモリ内の物理アドレスに変換する。場合によっては、GPU VMは、ページテーブルから頻繁に要求されるアドレス変換をキャッシュする、対応するトランスレーションルックアサイドバッファ(TLB)を使用してアドレス変換を実行する。

30

40

【発明の概要】

【課題を解決するための手段】

【0002】

一態様によれば、装置は、複数の入出力メモリ管理ユニット(IOMMU)を有するネットワーク化されたIOMMUを含む。ネットワーク化されたIOMMUは、第一アドレス変換層によって生成されるドメイン物理アドレスを含むメモリアクセス要求を受信し、メモリアクセス要求を生成したデバイスのタイプに基づいて選択される複数のIOMMU

50

のうち何れかを使用して、ドメイン物理アドレスをシステムメモリ内の物理アドレスに選択的に変換するように構成されている。

【 0 0 0 3 】

いくつかの実施形態では、デバイスは、グラフィックスプロセッシングユニット（GPU）又は複数の周辺デバイスのうち何れかであり、複数のIOMMUは、第一アドレス変換層からメモリアクセス要求を受信するように構成された一次IOMMU（primary IOMMU）と、一次IOMMUに接続され、デバイスに関連する回路に近接して配置される二次IOMMU（secondary IOMMU）と、を含む。いくつかの態様では、一次IOMMUは、メモリアクセス要求がGPUから受信されていることに応じてドメイン物理アドレスのアドレス変換を実行し、二次IOMMUは、メモリアクセス要求が周辺デバイスから受信されていることに応じてアドレス変換を実行する。

10

【 0 0 0 4 】

さらに、一次IOMMUは、一次IOMMUに関連する第一トランスレーションルックアサイドバッファ（TLB）及び第一ページテーブルセットを使用してページテーブルウォークを実行することによって、GPUから受信するドメイン物理アドレスのアドレス変換を実行することができる。いくつかの態様では、一次IOMMUは、メモリアクセス要求を、要求が周辺デバイスから受信されていることに応じて二次IOMMUに提供し、二次IOMMUは、二次IOMMUに関連する第二TLB及び第二ページテーブルセットを使用してページテーブルウォークを実行することによって、一次IOMMUから受信するメモリアクセス要求内の仮想アドレスのアドレス変換を実行する。

20

【 0 0 0 5 】

いくつかの実施形態では、周辺デバイスに関連する回路に対する二次IOMMUの位置は、周辺デバイスのレイテンシ要件に基づいて決定される。このような実施形態では、装置は、一次IOMMUに接続される複数の二次IOMMUをさらに含むことができ、複数の二次IOMMUは、複数の周辺デバイスに関連する回路に近接して配置され、複数の二次IOMMUの各々は、複数の周辺デバイスのうち対応する周辺デバイスに関連する回路から受信するメモリアクセス要求内のドメイン物理アドレスのアドレス変換を実行するように構成されている。複数の二次IOMMUは、複数の周辺デバイスに関連する回路内に集積されてもよい。

【 0 0 0 6 】

装置は、メモリアクセス要求を第一アドレス変換層から受信し、メモリアクセス要求を生成したデバイスのタイプに基づいて、メモリアクセス要求を一次IOMMU又は二次IOMMUに選択的に提供するように構成されたコマンドキューをさらに含むことができる。

30

【 0 0 0 7 】

他の態様によれば、方法は、複数の入出力メモリ管理ユニット（IOMMU）を含むネットワーク化されたIOMMUにおいて、第一アドレス変換層によって生成されるドメイン物理アドレスを含むメモリアクセス要求を受信することと、メモリアクセス要求を生成したデバイスのタイプに基づいて複数のIOMMUのうち何れかを選択することと、複数のIOMMUのうち選択されたIOMMUを使用して、ドメイン物理アドレスをシステムメモリ内の物理アドレスに選択的に変換することと、を含む。いくつかの実施形態では、デバイスは、グラフィックスプロセッシングユニット（GPU）又は複数の周辺デバイスのうち何れかであり、メモリアクセス要求を受信することは、複数のIOMMU内の一次IOMMUで第一アドレス変換層からメモリアクセス要求を受信することを含む。このような場合に、複数のIOMMUのうち何れかを選択することは、デバイスがGPUであることに応じて一次IOMMUを選択することと、デバイスが複数の周辺デバイスのうち何れかであることに応じて二次IOMMUを選択することと、二次IOMMUは、一次IOMMUに接続されており、複数の周辺デバイスのうち何れかに関連する回路に近接して配置されている、ことと、を含むことができる。方法は、メモリアクセス要求がGPUから受信されていることに応じて、一次IOMMUにおいてドメイン物理アドレスのアドレス変換を実行することをさらに含むことができ、一次IOMMUにおいてドメイン物

40

50

理アドレスのアドレス変換を実行することは、一次 I O M M U に関連する第一トランスレーションルックアサイドバッファ (T L B) 及び第一ページテーブルセットを使用してページテーブルウォークを実行することを含む。いくつかの実施形態では、方法は、メモリアクセス要求が周辺デバイスから受信されていることに応じて、二次 I O M M U においてアドレス変換を実行することをさらに含む。方法は、メモリアクセス要求を、要求が周辺デバイスから受信されていることに応じて一次 I O M M U から二次 I O M M U に提供することをさらに含むことができ、二次 I O M M U において仮想アドレスのアドレス変換を実行することは、二次 I O M M U に関連する第二 T L B 及び第二ページテーブルセットを使用してページテーブルウォークを実行することを含む。

【 0 0 0 8 】

いくつかの実施形態では、方法は、第一アドレス変換層からのメモリアクセス要求をコマンドキューで受信することと、メモリアクセス要求を生成したデバイスのタイプに基づいて、メモリアクセス要求をコマンドキューから一次 I O M M U 又は二次 I O M M U に選択的に提供することと、を含む。

【 0 0 0 9 】

さらに別の態様によれば、グラフィックスプロセッシングユニット (G P U) と、少なくとも 1 つの周辺デバイスと、メモリと、に接続されるように構成された、ネットワーク化された入出力メモリ管理ユニット (I O M M U) は、メモリアクセス要求を第一アドレス変換層から受信するように構成されたコマンドキューであって、メモリアクセス要求は、第一アドレス変換層によって生成されるドメイン物理アドレスを含む、コマンドキューと、メモリアクセス要求が G P U から受信されていることに応じて、ドメイン物理アドレスをメモリ内の物理アドレスに変換するように構成された一次 I O M M U と、メモリアクセス要求が少なくとも 1 つの周辺デバイスから受信されていることに応じて、ドメイン物理アドレスをメモリ内の物理アドレスに変換するように構成された少なくとも 1 つの二次 I O M M U と、を含む。いくつかの実施形態では、一次 I O M M U は、一次 I O M M U に関連する第一トランスレーションルックアサイドバッファ (T L B) 及び第一ページテーブルセットを使用してページテーブルウォークを実行することによって、G P U から受信するドメイン物理アドレスのアドレス変換を実行し、少なくとも 1 つの二次 I O M M U は、二次 I O M M U に関連する第二 T L B 及び第二ページテーブルセットを使用してページテーブルウォークを実行することによって、メモリアクセス要求内の仮想アドレスのアドレス変換を実行する。少なくとも 1 つの周辺デバイスに対する少なくとも 1 つの二次 I O M M U の少なくとも 1 つの位置は、少なくとも 1 つの周辺デバイスのレイテンシ要件に基づいて決定されてもよい。

【 0 0 1 0 】

添付の図面を参照することにより、本開示がより良く理解され、その多くの特徴及び利点が当業者に明らかとなる。異なる図面で同じ符号が使用されている場合には、類似又は同一のアイテムを示している。

【 図面の簡単な説明 】**【 0 0 1 1 】**

【 図 1 】 いくつかの実施形態による、処理システムのブロック図である。

【 図 2 】 メモリ内に従来の G P U カーブアウトを実装した処理システムの一部のブロック図である。

【 図 3 】 いくつかの実施形態による、G P U へのメモリの動的割り当てと組み合わせて G P U カーブアウトを実装する処理システムの一部のブロック図である。

【 図 4 】 いくつかの実施形態による、メモリアクセス要求内のデバイス生成アドレスの変換を示すブロック図である。

【 図 5 】 いくつかの実施形態による、ネットワーク化された I O M M U を実装する処理システムの一部のブロック図である。

【 図 6 】 いくつかの実施形態による、システムオンチップ (S O C) デバイス変換ブロックを実装する処理システムの使用モデルのブロック図である。

10

20

30

40

50

【図7】いくつかの実施形態による、アドレス変換及び一次IOMMU又は二次IOMMUを選択的に実行する方法のフロー図である。

【発明を実施するための形態】

【0012】

Microsoft Windows（登録商標）等のオペレーティングシステムのベンダによって推進されるセキュリティインフラストラクチャ及び要件における変更は、処理システムのメモリアクセスパフォーマンスに影響すると予想される。例えば、GPUカーブアウトのサイズが縮小する場合には、メモリのOS制御部分からGPUへの動的割り当てに利用可能なメモリ量が増加することができる。別の例の場合、仮想化ベースのセキュリティ（VBS）は、デバイスドライバによって管理される第一アドレス変換層を使用して、例えば、ページテーブル及びページテーブルから頻繁に要求されるアドレス変換をキャッシュするトランслーションルックアサイドバッファ（TLB）を使用して、アクセスされるシステムメモリ内のセキュアパーティションを生成することによって、カーネルモデルマルウェアに対するメモリの保護を提供する。ページテーブル及びTLBは、GPU仮想マネージャ（VM）に関連付けられている。セキュア部分にアクセスするために使用される第二アドレス変換層は、ハイパーバイザ又はセキュアOSによって制御される。第一アドレス変換層は連続的であり、そのパフォーマンスは高い。第二アドレス変換層は、メモリフラグメンテーション及びアクセスセキュリティ等の物理メモリ管理の課題を処理する。その結果、第二アドレス変換層は、通常、全体的なアドレス変換パフォーマンスを決定する。第二アドレス変換層は、システム全体の入出力メモリ管理ユニット（IOMMU）に実装される。このIOMMUは、GPU及び1つ以上の周辺デバイスを含むデバイスからのダイレクトメモリアクセス（DMA）転送に関する、アドレス変換及びシステムメモリアクセスの保護をサポートする。

【0013】

GPUからメモリアクセス要求を受信することに応じて、第一アドレス変換層は、メモリアクセス要求内のデバイス生成アドレスをドメイン物理アドレスに変換する。IOMMU内に実装される第二アドレス変換層は、ドメイン物理アドレスをシステムメモリ内のシステム物理アドレスに変換する。例えば、IOMMUは、ドメインコンテキスト及び個別のページテーブルセットを処理システム内の各デバイスに割り当てる。デバイスがシステムメモリの読み出し又は書き込みを試みると、IOMMUは、このアクセスをインターセプトし、デバイスが割り当てられているドメインコンテキストを決定する。読み出し、書き込み、実行等のような追加のパーミッションをページテーブル及びTLB内のエントリに符号化し、これらのエントリを使用して第二層の変換を実行する。したがって、IOMMUは、ドメインに関連するTLBエントリ又はデバイスに関連するページテーブルを使用して、アクセスが許可されるかどうかを決定し、アクセスされるシステムメモリ内の位置を決定する。例えば、デバイスからのメモリアクセス要求が許可されると決定したことに応じて、IOMMUは、第一アドレス変換層によって生成されるドメイン物理アドレスからシステムメモリ内の物理アドレスを生成する。

【0014】

周辺デバイス及びGPUからIOMMUを介して全てのメモリアクセス要求を送ると、いくつかの問題が起こる。例えば、IOMMUは、厳しいレイテンシ要件を有する、ビデオデコーダ、ビデオエンコーダ及びディスプレイフレームバッファスキャンアウト回路等のリアルタイム依存デバイスクライアントブロックにサービスを提供する。単一のIOMMUで複数のエンティティからメモリアクセス要求にページテーブルウォークを実行すると、処理遅延がもたらされ、レイテンシが増加する。さらに、単一のIOMMUを、全ての周辺デバイス及びGPUの近くに置くことができないため、エンティティのいくつかとIOMMUとの間のラウンドトリップタイムは、IOMMUでの処理レイテンシをさらに増加させる。その結果、中央のIOMMUは、単一のIOMMUの近くでハードアクセスデッドライン内（例えば、レイテンシが低い）の全てのメモリ要求にサービスを提供することができない。別々及び異なるIOMMUのシステムは、異なるデバイス又はGPUに

10

20

30

40

50

近接して配置され得る。ただし、特定デバイス向け I O M M U にプログラミングサポートを提供するには、システムソフトウェア内に異なるプログラミングモデルが必要となるため、この I O M M U をソフトウェア対象のシステムデバイスとして使用するホスト O S 及び他のシステムソフトウェアのアーキテクチャが複雑になる。

【 0 0 1 5 】

図 1 ~ 図 7 は、グラフィックスプロセッシングユニット (G P U) 及びネットワーク化された入出力メモリ管理ユニット (I O M M U) を含む処理システムの実施形態を開示しており、この処理システムは、ネットワーク化された I O M M U 構造を単一のデバイスとして可視であるアーキテクチャ化されたプログラミングモデルを使用して周辺デバイスのメモリアクセスレイテンシ要件を満たす。ネットワーク化された I O M M U は、ドメイン物理アドレスを含むメモリアクセス要求を受信し、このドメイン物理アドレスは、デバイスドライバが管理する第一アドレス変換層 (例えば、G P U V M 及び関連する T L B 等) によってデバイス生成アドレスから生成される。ネットワーク化された I O M M U は、ネットワーク化された I O M M U を形成する複数の I O M M U のうち何れかの I O M M U を使用して、ドメイン物理アドレスをシステムメモリ内の物理アドレスに選択的に変換する。いくつかの実施形態では、ネットワーク化された I O M M U は、ディスプレイ回路又はカメラ回路等の処理システム内の回路を介して、G P U 及び周辺デバイスからメモリアクセス要求を受信する一次 I O M M U を含む。また、ネットワーク化された I O M M U は、オペレーティングシステム (O S) 又はハイパーバイザ (H V) ソフトウェアとインタフェースする、一次 I O M M U に接続された 1 つ以上の二次 I O M M U を含む。ネットワーク化された I O M M U は、マスタスレーブネットワーク、スターネットワーク又は他のタイプのネットワーク内に一次及び二次 I O M M U を実装することで、一次 I O M M U は、ネットワーク化された I O M M U についてのフロントエンドとして機能する。二次 I O M M U の各々は、1 つ以上の周辺デバイスに関連する、対応する回路に近接して配置される (又は、その対応する回路内に集積される)。例えば、ディスプレイ及びカメラ用の二次 I O M M U は、ディスプレイ回路及びカメラ回路に近接して配置される。二次 I O M M U の位置は、少なくとも部分的に、1 つ以上の周辺デバイスのレイテンシ要件によって決定される。

【 0 0 1 6 】

第一変換層からドメイン物理アドレスを含むメモリアクセス要求を受信したことに応じて、一次 I O M M U は、メモリアクセス要求を提供したデバイスのタイプに基づいて、ドメイン物理アドレスのアドレス変換を選択的に実行する又はアドレス変換をバイパスする。いくつかの実施形態では、一次 I O M M U は、一次 I O M M U に関連する第一ページテーブルセット及び第一トランスレーションルックアサイドバッファ (T L B) を使用してページテーブルウォークを実行することによって、G P U からメモリアクセス要求に関連するドメイン物理アドレスのアドレス変換を実行する。一次 I O M M U は、周辺デバイスから受信するメモリアクセス要求内のドメイン物理アドレスのアドレス変換をバイパスする。或いは、一次 I O M M U は、メモリアクセス要求を提供した周辺デバイスに関連する二次 I O M M U に対して、メモリアクセス要求を提供する。二次 I O M M U は、二次 I O M M U に関連する第二ページテーブルセット及び第二 T L B を使用してページテーブルウォークを実行することによって、ドメイン物理アドレスのアドレス変換を実行する。一次 I O M M U のいくつかの実施形態は、一次 I O M M U 及び二次 I O M M U に関連するコマンドを受信するコマンドキューを含む (又は、そのコマンドキューに関連付けられている)。コマンドキューは、上述したように、一次 I O M M U で処理される又は何れかの二次 I O M M U に選択的に転送されるページテーブルウォーク及びデバイス再スキャンをシステムソフトウェアが開始することを可能にする。また、コマンドキューは、システムソフトウェアの再スキャン、及び、システムソフトウェアと周辺デバイスとの同期をサポートし、ソフトウェアが現在移動中のテーブルデータを変更しないことを確実にする。

【 0 0 1 7 】

図 1 は、いくつかの実施形態による、処理システム 1 0 0 のブロック図である。処理シ

10

20

30

40

50

システム 100 は、ダイナミックランダムアクセスメモリ (DRAM) 等の非一時的なコンピュータ可読記憶媒体を使用して実装されるシステムメモリ 105 若しくは他のストレージコンポーネントを含み、又は、このシステムメモリ若しくは他のストレージコンポーネントへのアクセスを有する。ただし、メモリ 105 のいくつかの実施形態は、スタティックランダムアクセスメモリ (SRAM)、不揮発性 RAM 等を含む他のタイプのメモリを使用して実装される。また、処理システム 100 は、バス 110 を含み、これは、メモリ 105 等のように処理システム 100 内に実装されるエンティティ間の通信をサポートする。処理システム 100 のいくつかの実施形態は、分かり易くするために図 1 に示されていない他のバス、ブリッジ、スイッチ、ルータ等を含む。

【0018】

処理システム 100 は、ディスプレイ 120 での提示のために画像をレンダリングするグラフィックスプロセッシングユニット (GPU) 115 を含む。例えば、GPU 115 は、オブジェクトをレンダリングし、ディスプレイ 120 に提供される画素値を生成し、ディスプレイ 120 は、その画素値を使用して、レンダリングされたオブジェクトを表す画像を表示する。GPU 115 のいくつかの実施形態は、同時に又は並列に命令を実行する複数の処理要素 (明確にするために図 1 に示されていない) を含む。処理要素は、計算部、プロセッサコア又は他の用語を使用して呼ばれる。GPU 115 のいくつかの実施形態は、汎用コンピューティングに使用される。図示した実施形態では、GPU 115 は、バス 110 を介してメモリ 105 と通信する。しかし、GPU 115 のいくつかの実施形態は、直接接続により、又は、他のバス、ブリッジ、スイッチ及びルータ等を介して、メモリ 105 と通信する。GPU 115 は、メモリ 105 に格納される命令を実行し、GPU 115 は、実行された命令の結果等の情報をメモリ 105 に格納する。例えば、メモリ 105 は、GPU 115 によって実行されるプログラムコードを表す命令のコピー 125 を格納する。

【0019】

GPU 115 のいくつかの実施形態は、GPU VM 116 及び 1 つ以上の対応する TLB (明確にするために、図 1 には 1 つの TLB 117 のみが示されている) を使用して、仮想アドレスから物理アドレスへの変換を実行する。例えば、場合によっては、GPU VM 116 及び TLB 117 は、第一アドレス変換層の一部として実装され、この第一アドレス変換層は、GPU 115 で受信する又は GPU 115 によって生成されるメモリアクセス要求内に含まれる仮想アドレスからドメイン物理アドレスを生成する。GPU VM 116 及び TLB 117 は、図 1 では GPU 115 の集積部分として示されているが、GPU VM 116 又は TLB 117 のいくつかの実施形態は、GPU 115 の外部に実装される。

【0020】

また、処理システム 100 は、本明細書ではまとめて「処理要素 131 ~ 133」と呼ばれる複数の処理要素 131, 132, 133 を実装する中央処理装置 (CPU) 130 を含む。処理要素 131 ~ 133 は、同時に又は並列に命令を実行する。CPU 130 は、バス 110 に接続されており、GPU 115 及びメモリ 105 とバス 110 を介して通信する。CPU 130 は、メモリ 105 に格納されたプログラムコード 135 等の命令を実行し、CPU 130 は、実行された命令の結果等の情報をメモリ 105 に格納する。また、CPU 130 は、GPU 115 にドローコールを発行することにより、グラフィックス処理を開始することができる。

【0021】

入力/出力 (I/O) エンジン 140 は、ディスプレイ 120、並びに、キーボード、マウス、プリンタ及び外部ディスク等の処理システム 100 の他の要素に関連する入力動作又は出力動作に対処する。図示した実施形態では、I/O エンジン 140 は、カメラ 145 に関連する入力及び出力の操作も処理する。I/O エンジン 140 は、I/O エンジン 140 がメモリ 105、GPU 115 又は CPU 130 と通信できるように、バス 110 に接続されている。図示した実施形態では、I/O エンジン 140 は、コンパクトディ

10

20

30

40

50

スク（CD）及びデジタルビデオディスク（DVD）等の非一時的なコンピュータ可読記憶媒体を使用して実装される外部ストレージコンポーネント150に記憶された情報を読み出す。また、I/Oエンジン140は、GPU115又はCPU130による処理の結果等の情報を外部ストレージコンポーネント150に書き込む。

【0022】

処理システム100は、ネットワーク化されたI/Oメモリ管理ユニット（IOMMU）155を含み、このネットワーク化されたIOMMUは、例えばGPU115、並びに、ディスプレイ120、カメラ145及び外部ストレージコンポーネント150を含む周辺デバイス等のデバイスからのメモリアクセス要求を処理するIOMMUセットを含む。メモリアクセス要求は、システムメモリ105内での位置を示すために使用される仮想アドレス等のデバイス生成アドレスを含む。ネットワーク化されたIOMMU155のいくつかの実施形態は、GPU115によって実装されるグラフィックスドライバ等のドライバによって管理される第一アドレス変換層によって生成されるドメイン物理アドレスを含むメモリアクセス要求を受信する。例えば、第一アドレス変換層は、GPU VM116及びTLB117を含むことができる。ネットワーク化されたIOMMU155は、メモリアクセス要求を生成したデバイスのタイプに基づいて選択されるIOMMUセットのうち何れかのIOMMUを使用して、ドメイン物理アドレスをシステムメモリ105内の物理アドレスに選択的に変換する。これらのタイプは、GPU115についての第一タイプと、ディスプレイ120、カメラ145及び外部ストレージコンポーネント150等の周辺デバイスについての第二タイプと、を含む。

【0023】

図示した実施形態では、ネットワーク化されたIOMMU135は、メモリアクセス要求を第一アドレス変換層から受信する一次IOMMU160と、この一次IOMMU160に接続され、ディスプレイ120、カメラ145及び外部ストレージコンポーネント150等の周辺デバイスに関連する回路（明確にするために、図1に示されていない）に近接して配置される二次IOMMU165、170と、を含む。一次IOMMU160及び二次IOMMU165、170がマスタスレーブネットワーク、スターネットワーク又は他のタイプのネットワークとして配置されることで、メモリアクセス要求は、一次IOMMU160によって最初に受信され、必要な場合には、二次IOMMU165、170に選択的に分散されることを可能にする。一次IOMMU160のいくつかの実施形態は、TLBシュートダウン又は他のソフトウェアコマンドを二次IOMMU165、170に渡す機能を担う。

【0024】

ネットワーク化されたIOMMU135は、ページテーブル180に格納されるアドレス変換を使用してアドレス変換を実行する。処理システム100内のデバイス上で実行されている各プロセスは、対応するページテーブルを含む。プロセス用のページテーブル180は、このプロセスによって使用されているデバイス生成（例えば、仮想）アドレスを、システムメモリ105内の物理アドレスに変換する。一次IOMMU160及び二次IOMMU165、170は、ページテーブル180のテーブルウォークを独立して実行し、メモリアクセス要求内のアドレスの変換を決定する。ネットワーク化されたIOMMU135によって頻繁に使用される変換は、TLB185に格納され、これらのTLBは、頻繁に要求されたアドレス変換をキャッシュするために使用される。個別のTLB185は、一次IOMMU160及び二次IOMMU165、170に関連付けられている。頻繁に使用されたアドレス変換を含むエントリは、ページテーブル180から、一次IOMMU160及び二次IOMMU165、170についてのTLB185に書き込まれる。したがって、一次IOMMU160及び二次IOMMU165、170は、ページテーブル180内の変換を検索するオーバーヘッド無しに、TLB185からアドレス変換に独立してアクセスすることができる。エントリは、TLB置換ポリシーに従ってTLB185からエビクトされ、新しいエントリのためにスペースを空ける。TLB185は、図1ではネットワーク化されたIOMMU135の集積部分として示されている。ただし、他

10

20

30

40

50

の実施形態では、TLB 185は、ネットワーク化されたIOMMU 135によってアクセス可能な別のエンティティ内に実装される。いくつかの実施形態では、TLB 117及びTLB 185は、単一のアーキテクチャで実装されるが、これは、全ての実施形態において必要とされない。

【0025】

図2は、メモリ内に従来のGPUカーブアウトを実装する処理システムの一部200のブロック図である。この部分200は、GPU 205及びシステムメモリ210を含む。図示した実施形態では、システムメモリ210は、ホストパーティション215と、GPUカーブアウトとも呼ばれるGPUパーティション220と、にパーティション化されている。例えば、システムメモリ210が合計8GB以上を含む場合、GPUパーティション220は、最大2GBの専用フレームバッファを含む。GPUパーティション220は、GPU 220による排他的なアクセスのために予約されている。GPUパーティション220は、処理システム上で実行されるOSに可視ではない。その結果、グラフィックス中心ではないアプリケーション、又は、他の方法でGPU 220のリソースを著しく消費しないアプリケーションを処理システムが実行している場合に、GPUパーティション220は、十分に利用されていない可能性がある。本明細書で説明するように、GPUパーティション220がOSに可視ではないため、GPUパーティション220を、他のアプリケーションに割り当てることができない。

10

【0026】

図3は、いくつかの実施形態による、GPUへのメモリの動的割り当てと組み合わせてGPUカーブアウトを実装する処理システムの一部300のブロック図である。この部分300は、GPU 305及びシステムメモリ310を含む。図示した実施形態では、システムメモリ310は、ホストパーティション315と、GPUパーティション320と、にパーティション化されている。GPUパーティション320のサイズは、所定の(比較的小さい)サイズに制限されている。いくつかの実施形態では、GPUパーティション320は、3%以下のシステムメモリ310(例えば、8GBに対して256MBのシステムメモリ310等)を含む。処理システムは、ホストパーティション315の部分321, 322, 323(本明細書では「部分321~323」と呼ばれる)を動的に割り当てることによって、小さいサイズのGPUパーティション320を補償する。

20

【0027】

処理システムの部分300は、ネットワーク化されたIOMMU 325を含み、このネットワーク化されたIOMMUは、メモリアクセス要求内のデバイス生成アドレスをGPUパーティション320又は部分321~323内の物理アドレスに変換する。例えば、GPU VM及び関連するTLBは、メモリアクセス要求内の仮想メモリアドレスをドメイン物理アドレスに変換し、このドメイン物理アドレスを含むメモリアクセス要求をネットワーク化されたIOMMU 325に提供することができる。いくつかの実施形態では、ページテーブルは、GPU 305上で実行されるプロセスへの部分321~323の割り当てに応じて定義される。例えば、仮想アドレスは、GPU 305上で実行されるプロセスによって使用されると、このプロセスに割り当てられている部分321内の物理アドレスにマッピングされる。このマッピングは、プロセスに関連するページテーブルのエントリに格納される。ネットワーク化されたIOMMU 325はIOMMUセットを含み、このネットワーク化されたIOMMU 325は、メモリアクセス要求を生成したデバイスのタイプに基づいて選択されるIOMMUセットのうち何れかのIOMMUを使用して、ドメイン物理アドレスをシステムメモリ310内の物理アドレスに選択的に変換する。例えば、IOMMUセット内の一次IOMMUは、GPU 305からメモリアクセス要求を受信することに応じて、ドメイン物理アドレスをシステムメモリ310内の物理アドレスに変換する。別の例では、一次IOMMUは、変換をバイパスし、ディスプレイ又はカメラ等の周辺デバイスからメモリアクセス要求を受信したことに応じて、メモリアクセス要求を、変換用の二次IOMMUに提供する。

30

40

【0028】

50

図4は、いくつかの実施形態による、メモリアクセス要求内のデバイス生成アドレスの変換400を示すブロック図である。変換400は、図1に示す処理システム100及び図3に示す処理システムの部分300のいくつかの実施形態で実行される。図示した実施形態では、デバイス405は、情報をシステムメモリ410に書き込む要求、又は、情報をシステムメモリ410から読み出す要求等のメモリアクセス要求を発行する。デバイス405は、図1に示すGPU115、又は、図1に示すディスプレイ120、カメラ145若しくは外部ストレージ媒体150等の周辺デバイスのいくつかの実施形態を表す。

【0029】

メモリアクセス要求は、デバイス405上で実行される又はそのデバイスに関連するアプリケーションによって使用される仮想アドレス等のデバイス生成アドレスを含む。図示した実施形態では、仮想化ベースのセキュリティ(VBS)は、OS又はデバイスドライバ420によって管理される第一レベル変換415と、ハイパーバイザ430によって管理される第二レベル変換425と、を含む2つのレベルの変換プロセスを使用して、メモリ保護(例えば、カーネルモードマルウェアに対する)を提供する。第一レベル変換415は、メモリアクセス要求内の仮想アドレス等のデバイス生成アドレスをGPU物理アドレス等のドメイン物理アドレスに変換する。いくつかの実施形態では、第一レベル変換415は、本明細書で説明するように、GPU VM及び関連するTLBによって実行される。ドメイン物理アドレスは、第二レベル変換425に渡され、この第二レベル変換は、ドメイン物理アドレスを、システムメモリ410内の位置を示す物理アドレスに変換する。本明細書で説明するように、第二レベル変換425は、例えば、第二レベル変換425を実行するために使用される、関連するページテーブル及びトランслーションルックアサイドバッファ(TLB)内のエントリに符号化されるパーミッション情報を使用して、デバイス405が、物理アドレスによって示されるシステムメモリ410の領域にアクセスすることが承認されることも検証する。

【0030】

図5は、いくつかの実施形態による、ネットワーク化されたIOMMU505を実装する処理システムの一部500のブロック図である。部分500は、図1に示す処理システム100のいくつかの実施形態の一部を表し、ネットワーク化されたIOMMU505は、図1に示すネットワーク化されたIOMMU135、及び、図3に示すネットワーク化されたIOMMU325のいくつかの実施形態を表す。

【0031】

ネットワーク化されたIOMMU505は、一次IOMMU515に統合されたソフトウェアインタフェース510を介してメモリアクセス要求を受信する。図示した実施形態では、メモリアクセス要求は、処理システム内に実装されるIOMMUドライバ520等のソフトウェアによって提供される。IOMMUドライバ520は、例えば、GPU VM及び関連するTLBを含むアドレス変換層等の第一アドレス変換層からメモリアクセス要求を受信する(明確にするために、図5に示されていない)。第一アドレス変換層は、仮想メモリアドレス等のデバイス生成アドレスをドメイン物理アドレスに変換し、ドメイン物理アドレスを含むメモリアクセス要求を、ネットワーク化されたIOMMU505に提供する。ソフトウェアは、デバイス生成アドレス(例えば、仮想アドレス)を使用して、システムメモリ525内の位置を示す。

【0032】

一次IOMMU515及び統合されたソフトウェアインタフェース510は、単一のデバイスとしてシステムソフトウェア(IOMMUドライバ520等)によって対象とされるアーキテクチャ化されたプログラミングモデルをサポートする。したがって、プログラミングモデルは、リアルタイムでの異なるIOMMUハードウェアユニットの操作、並びに、従来のダイレクトメモリアクセス(DMA)処理のための専用の制御メカニズム及びソフトウェアを必要としない。ただし、IOMMUを必要とする一部のデバイスクライアントブロックは、例えば、ビデオデコーダのインコード、ディスプレイフレームバッファのスキャンアウト等についてのワーストケースレイテンシ要件を満たすサービスを提供す

10

20

30

40

50

る。単一の一次 I O M M U 5 1 5 は、常にレイテンシ要件を満たすことができるとは限らない。

【 0 0 3 3 】

デバイスクライアントブロックのワーストケースレイテンシ要件に少なくとも部分的に対処するために、ネットワーク化された I O M M U 5 0 5 は、周辺デバイスに対応するデバイスクライアントブロックに近接して配置される 1 つ以上の二次 I O M M U 5 3 5 , 5 3 6 を含む。図示した実施形態では、周辺デバイス回路は、図 1 に示すディスプレイ 1 2 0 等のディスプレイとの通信をサポートするディスプレイ回路 5 4 0 と、図 1 に示すカメラ 1 4 5 等のカメラとの通信をサポートするカメラ回路 5 4 1 と、を含む。二次 I O M M U 5 3 5 , 5 3 6 は、対応する回路 5 4 0 , 5 4 1 内に集積される。ただし、いくつかの実施形態では、二次 I O M M U 5 3 5 , 5 3 6 は、対応する回路 5 4 0 , 5 4 1 のレイテンシ要件に基づいて決定される距離（物理距離又は論理距離）で対応する回路 5 4 0 , 5 4 1 に近接して配置される。レイテンシが低い要件であるほど、二次 I O M M U 5 3 5 , 5 3 6 は、対応する回路 5 4 0 , 5 4 1 の近くに配置される（物理的に又は論理的に）ことが黙示される。一次 I O M M U 5 1 5 及び二次 I O M M U 5 3 5 , 5 3 6 は、一次 I O M M U 5 1 5 が、統合されたソフトウェアインタフェース 5 1 0 を介してメモリアクセス要求を受信する単一のデバイスとして機能することを可能にする、マスタスレーブ関係、スターネットワーク又は他のコンフィギュレーションで実装される。

10

【 0 0 3 4 】

動作中に、一次 I O M M U 5 1 5 は、第一タイプのデバイスからのメモリアクセス要求（例えば、GPUからの要求）に対してアドレス変換を実行し、第二タイプのデバイスからのメモリアクセス要求（例えば、周辺デバイスからの要求）に対してアドレス変換を実行することをバイパスする。一次 I O M M U 5 1 5 は、メモリアクセス要求を、第二タイプのデバイスから対応する二次 I O M M U 5 3 5 , 5 3 6 に転送する。例えば、一次 I O M M U 5 1 5 は、ディスプレイに関連するメモリアクセス要求を二次 I O M M U 5 3 5 に転送し、カメラに関連するメモリアクセス要求を二次 I O M M U 5 3 6 に転送する。したがって、I O M M U ドライバ 5 2 0 は、インタフェース 5 1 0 を介してシステムメモリ 5 2 5 にアクセスする単一のコマンド（例えば、単一のメモリアクセス要求）を発行する。次に、単一のコマンドは、一次 I O M M U 5 1 5 が専用の二次 I O M M U 5 3 5 , 5 3 6 のうち何れかによって選択的に処理される。したがって、I O M M U ドライバ 5 2 0 は、専用の I O M M U 5 1 5 , 5 3 5 , 5 3 6 を別々に又は独立してアドレス指定する必要なく、アクセスポリシーを実装する。

20

30

【 0 0 3 5 】

適切なタイプのデバイスからメモリアクセス要求を受信したことに応じて、一次 I O M M U 5 1 5 又は二次 I O M M U 5 3 5 , 5 3 6 は、対応する T L B 5 4 5 , 5 4 6 , 5 4 7（本明細書では「T L B 5 4 5 ~ 5 4 7」と呼ばれる）内のエントリにアクセスし、メモリアクセス要求に含まれるアドレスの変換を見つけようとする。T L B 5 4 5 ~ 5 4 7 内のエントリは、要求するデバイスがシステムメモリ 5 2 5 にアクセスすることを許可されるかどうかを示す情報を符号化する。アドレスが対応する T L B 5 4 5 ~ 5 4 7 でヒットし、デバイスが適切なパーミッションを有する場合に、メモリアクセス要求は、システムメモリ 5 2 5 に転送される。アドレスが対応する T L B 5 4 5 ~ 5 4 7 でミスする場合に、メモリアクセス要求は、対応するページテーブル 5 5 0 , 5 5 1 , 5 5 2 に転送され、この対応するページテーブルは、アドレスの適切な変換を返す。T L B 5 4 5 ~ 5 4 7 内のエントリは、T L B 5 4 5 ~ 5 4 7 によって実装される置換ポリシーに基づいて更新される。デバイスが適切なパーミッションを有さない場合に、メモリアクセス要求は拒否される。

40

【 0 0 3 6 】

ネットワーク化された I O M M U 5 0 5 のいくつかの実施形態は、メモリアクセス要求が一次 I O M M U 5 1 5 に発行される前に、メモリアクセス要求を I O M M U ドライバ 5 2 0 から受信し、これらのアクセス要求を格納するコマンドキュー 5 3 0 を含む。コマン

50

ドキュー 530 は、システムソフトウェアがページテーブル及びデバイスの再スキャンを開始することを可能にし、これらの再スキャンが一次 I O M M U 5 1 5 又は二次 I O M M U 5 3 5 , 5 3 6 に転送されることで、これらの I O M M U は、対応する T L B 5 4 5 ~ 5 4 7 に関連データをキャッシュすることができる。また、コマンドキュー 530 のいくつかの実施形態は、システムソフトウェアの再スキャン、及び、システムソフトウェアとハードウェアユニットとの同期を可能にし、ソフトウェアが移動中のテーブルデータを変更しないことを確実にする。

【 0 0 3 7 】

図 6 は、いくつかの実施形態による、システムオンチップ (S O C) デバイス変換ブロック 605 を実装する処理システムの使用モデル 600 のブロック図である。使用モデル 600 は、図 1 に示す処理システム 100、図 3 に示す処理システムの部分 300、及び、図 5 に示す処理システムの部分 500 のいくつかの実施形態に実装される。

10

【 0 0 3 8 】

S O C デバイス変換ブロック 605 は、G P U 内のグラフィックスパイプライン (G F X) 615 を含むデバイス、及び、ディスプレイ 620、カメラ 625 等の周辺デバイスからメモリアクセス要求を受信する一次 I O M M U 610 を含む。いくつかの実施形態では、メモリアクセス要求は、第一アドレス変換層 (例えば、G P U V M 及び T L B を使用して実装されるアドレス変換層等) から受信され、メモリアクセス要求は、第一アドレス変換層によって生成されるドメイン物理アドレスを含む。メモリアクセス要求は、D R A M 630 等のシステムメモリにアクセスするために使用される。本明細書で説明するように、一次 I O M M U 610 は、メモリアクセス要求に含まれるアドレスのアドレス変換を、この要求を発行したデバイスのタイプに基づいて選択的に実行する。一次 I O M M U 610 によって変換されていないデバイスタイプからのメモリアクセス要求は、ディスプレイ 620、カメラ 625、及び、他の周辺デバイスに関連する 1 つ以上の I O M M U を含む分散リモート I O M M U ネットワーク 635 に転送される。分散リモート I O M M U ネットワーク 635 のいくつかの実施形態は、図 5 に示す二次 I O M M U 535 , 536 を使用して実装される。また、S O C デバイス変換ブロック 605 は、G F X 615 によって生成される要求に関連するアドレスを変換する変換キャッシュ 640 を含む。仮想から物理へのマネージャ 645 は、ディスプレイ 620、カメラ 625 等の周辺デバイス、及び、任意の他の周辺デバイスをサポートするために使用される。

20

30

【 0 0 3 9 】

動作中に、カーネルモデルドライバ又はメモリマネージャ 650 は、シグナリング 655 を提供し、このシグナリングは、例えば G P U V M 及び関連する T L B によって実行されるアドレス変換の第一層を使用して、仮想アドレスを G P U 物理アドレス (又は、ドメイン物理アドレス) に変換するために使用される、ページテーブル等のアドレス変換テーブルを設定する。また、メモリマネージャ 650 は、G P U 仮想アドレス等の仮想アドレス 656 を仮想から物理へのマネージャ 645 に提供する。ハイパーバイザ又はハイパーバイザ抽象化層 (H A L) 660 は、D R A M 630 に格納される、システムの物理ページテーブル及びアクセスパーミッションを管理する。また、H A L 660 は、S O C デバイス変換層 605 内の一次 I O M M U 610 を構成する。G F X 615 は、変換キャッシュ 640 を使用して仮想アドレスを変換することを試行する。この試行が変換キャッシュ 640 でヒットする場合に、返されたアドレス変換は、さらなる処理に使用される。この試行が変換キャッシュ 640 でミスする場合に、この要求は、一次 I O M M U 610 に転送され、この一次 I O M M U は、本明細書で説明するように後続のアドレス変換を処理する。また、一次 I O M M U 610 及び分散リモート I O M M U ネットワーク 635 は、本明細書で説明するように、D R A M 630 にアクセスし、ページテーブルウォークを実行することができる。

40

【 0 0 4 0 】

図 7 は、いくつかの実施形態による、アドレス変換及び一次 I O M M U 又は二次 I O M M U を選択的に実行する方法 700 のフロー図である。方法 700 は、図 1 に示すネット

50

ワーク化された I O M M U 1 3 5、図 3 に示すネットワーク化された I O M M U 3 2 5、図 4 に示す第二レベル変換 4 2 5、図 5 に示すネットワーク化された I O M M U 5 0 5、及び、図 6 に示す S O C デバイス変換ブロック 6 0 5 のいくつかの実施形態等のネットワーク化された I O M M U で実施される。

【 0 0 4 1 】

ブロック 7 0 5 では、ネットワーク化された I O M M U は、特定のタイプのデバイスからメモリアクセス要求を受信する。デバイスタイプの例は、グラフィックスプロセッサタイプ、周辺デバイスタイプ等を含む。いくつかの実施形態では、メモリアクセス要求は、第一アドレス変換層（例えば、G P U V M 及び T L B を使用して実装されるアドレス変換層等）から受信され、メモリアクセス要求は、第一アドレス変換層によって生成されるドメイン物理アドレスを含む。

10

【 0 0 4 2 】

決定ブロック 7 1 0 では、ネットワーク化された I O M M U は、メモリアクセス要求を発行したデバイスのタイプを、この要求に含まれる情報等に基づいて決定する。メモリアクセス要求を発行したデバイスのタイプが周辺デバイスである場合に、方法 7 0 0 はブロック 7 1 5 に進む。メモリアクセス要求を発行したデバイスのタイプが G P U デバイスである場合に、方法 7 0 0 はブロック 7 2 0 に進む。

【 0 0 4 3 】

ブロック 7 1 5 では、ネットワーク化された I O M M U 内の一次 I O M M U は、周辺デバイスタイプからのメモリアクセス要求についてのアドレス変換をバイパスする。次いで、方法 7 0 0 はブロック 7 2 5 に進み、一次 I O M M U は、メモリアクセス要求を、要求するデバイスに関連する二次 I O M M U に転送する。例えば、一次 I O M M U は、メモリアクセス要求がディスプレイからのものであることに応じて、ディスプレイ回路に集積される二次 I O M M U にメモリアクセス要求を転送する。次に、ブロック 7 3 0 では、二次 I O M M U は、アドレス変換を実行する。

20

【 0 0 4 4 】

ブロック 7 2 0 では、ネットワーク化された I O M M U 内の一次 I O M M U は、G P U からのメモリアクセス要求について要求されたアドレス変換を実行する。

【 0 0 4 5 】

コンピュータ可読記憶媒体は、命令及び/又はデータをコンピュータシステムに提供するために、使用中にコンピュータシステムによってアクセス可能な任意の非一時的な記憶媒体又は非一時的な記憶媒体の組み合わせを含む。このような記憶媒体には、限定されないが、光学媒体（例えば、コンパクトディスク（C D）、デジタル多用途ディスク（D V D）、ブルーレイ（登録商標）ディスク）、磁気媒体（例えば、フロッピー（登録商標）ディスク、磁気テープ、磁気ハードドライブ）、揮発性メモリ（例えば、ランダムアクセスメモリ（R A M）若しくはキャッシュ）、不揮発性メモリ（例えば、読取専用メモリ（R O M）若しくはフラッシュメモリ）、又は、微小電気機械システム（M E M S）ベースの記憶媒体が含まれ得る。コンピュータ可読記憶媒体（例えば、システム R A M 又は R O M）はコンピューティングシステムに内蔵されてもよいし、コンピュータ可読記憶媒体（例えば、磁気ハードドライブ）はコンピューティングシステムに固定的に取り付けられてもよいし、コンピュータ可読記憶媒体（例えば、光学ディスク又はユニバーサルシリアルバス（U S B）ベースのフラッシュメモリ）はコンピューティングシステムに着脱可能に取り付けられてもよいし、コンピュータ可読記憶媒体（例えば、ネットワークアクセス可能ストレージ（N A S））は有線又は無線ネットワークを介してコンピュータシステムに結合されてもよい。

30

40

【 0 0 4 6 】

いくつかの実施形態では、上記の技術のいくつかの態様は、ソフトウェアを実行するプロセッシングシステムの 1 つ以上のプロセッサによって実装されてもよい。ソフトウェアは、非一時的なコンピュータ可読記憶媒体に記憶され、又は、非一時的なコンピュータ可読記憶媒体上で有形に具現化された実行可能命令の 1 つ以上のセットを含む。ソフトウェ

50

アは、1つ以上のプロセッサによって実行されると、上記の技術の1つ以上の態様を実行するように1つ以上のプロセッサを操作する命令及び特定のデータを含むことができる。非一時的なコンピュータ可読記憶媒体は、例えば、磁気若しくは光ディスク記憶デバイス、例えばフラッシュメモリ、キャッシュ、ランダムアクセスメモリ(RAM)等のソリッドステート記憶デバイス、又は、他の1つ以上の不揮発性メモリデバイス等を含むことができる。非一時的なコンピュータ可読記憶媒体に記憶された実行可能命令は、ソースコード、アセンブリ言語コード、オブジェクトコード、又は、1つ以上のプロセッサによって解釈若しくは実行可能な他の命令フォーマットであってもよい。

【0047】

上述したものに加えて、概要説明において説明した全てのアクティビティ又は要素が必要とされているわけではなく、特定のアクティビティ又はデバイスの一部が必要とされない場合があり、1つ以上のさらなるアクティビティが実行される場合があり、1つ以上のさらなる要素が含まれる場合があることに留意されたい。さらに、アクティビティが列挙された順序は、必ずしもそれらが実行される順序ではない。また、概念は、特定の実施形態を参照して説明された。しかしながら、当業者であれば、特許請求の範囲に記載されているような本発明の範囲から逸脱することなく、様々な変更及び変形を行うことができるのを理解するであろう。したがって、明細書及び図面は、限定的な意味ではなく例示的な意味で考慮されるべきであり、これらの変更形態の全ては、本発明の範囲内に含まれることが意図される。

【0048】

利益、他の利点及び問題に対する解決手段を、特定の実施形態に関して上述した。しかし、利益、利点、問題に対する解決手段、及び、何かしらの利益、利点若しくは解決手段が発生又は顕在化する可能性のある特徴は、何れか若しくは全ての請求項に重要な、必須の、又は、不可欠な特徴と解釈されない。さらに、開示された発明は、本明細書の教示の利益を有する当業者には明らかな方法であって、異なっているが同様の方法で修正され実施され得ることから、上述した特定の実施形態は例示にすぎない。添付の特許請求の範囲に記載されている以外に本明細書に示されている構成又は設計の詳細については限定がない。したがって、上述した特定の実施形態は、変更又は修正されてもよく、かかる変更形態の全ては、開示された発明の範囲内にあると考えられることが明らかである。したがって、ここで要求される保護は、添付の特許請求の範囲に記載されている。

10

20

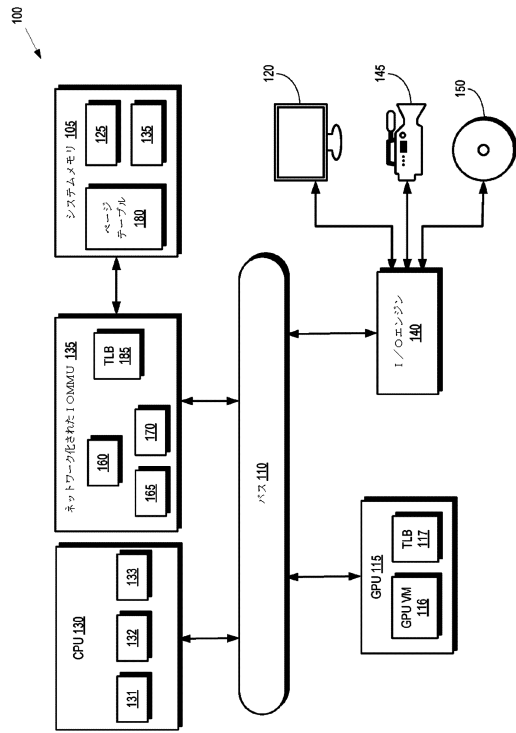
30

40

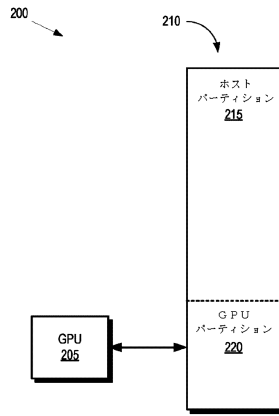
50

【図面】

【図 1】



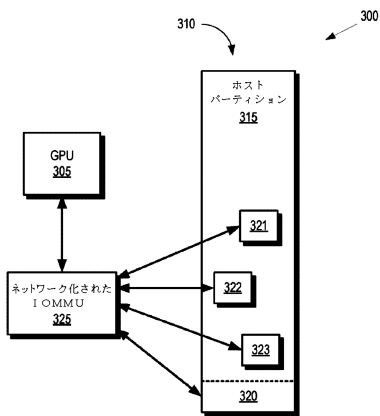
【図 2】



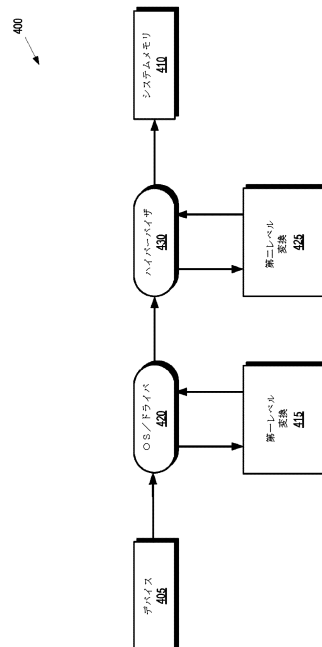
10

20

【図 3】



【図 4】

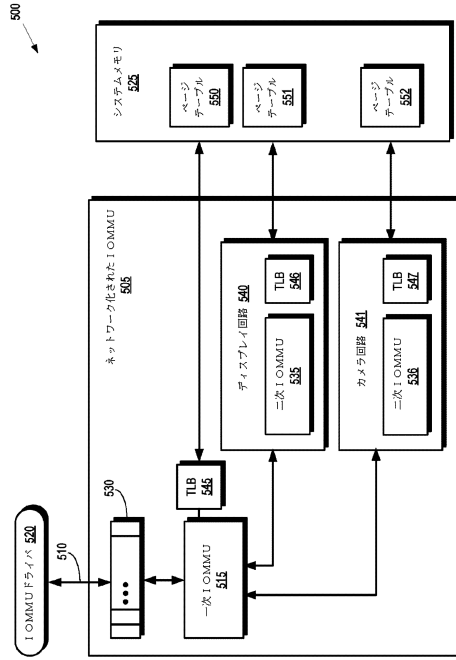


30

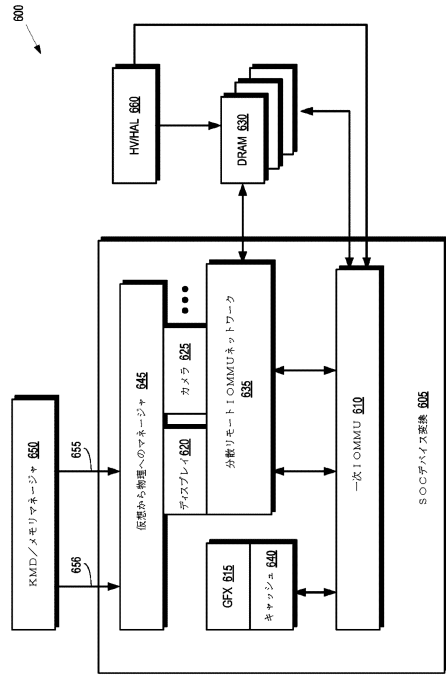
40

50

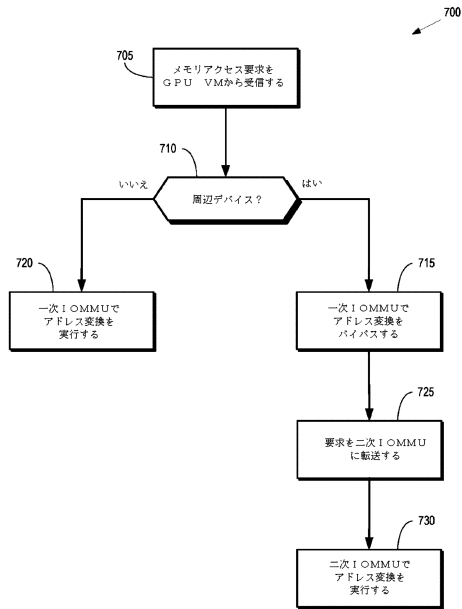
【図5】



【図6】



【図7】



10

20

30

40

50

フロントページの続き

One Commerce Valley Drive East, Markham, Ontario, L3T 7X6 Canada

(74)代理人 100108833

弁理士 早川 裕司

(74)代理人 100111615

弁理士 佐野 良太

(74)代理人 100162156

弁理士 村雨 圭介

(72)発明者 ソヌ アローラ

アメリカ合衆国 95054 カリフォルニア州、サンタクララ、オーガスティン ドライブ 2485

(72)発明者 ボール ブリンザー

アメリカ合衆国、98007 ワシントン州、ベルビュー 2002-156番 アベニュー ノースイースト、スイート 300

(72)発明者 フィリップン

カナダ国 L3T 7X6 オンタリオ州、マーカム、ワン コマース パリー ドライブ イースト

(72)発明者 ニッポン ハルシャダク ラヴァル

カナダ国 L3T 7X6 オンタリオ州、マーカム、ワン コマース パリー ドライブ イースト

審査官 北村 学

(56)参考文献 特表2017-529606(JP, A)

特表2018-523247(JP, A)

米国特許出願公開第2017/0249106(US, A1)

(58)調査した分野 (Int.Cl., DB名)

IPC G06F 12/00 - 12/06

G06F 12/08 - 12/128