

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7538950号
(P7538950)

(45)発行日 令和6年8月22日(2024.8.22)

(24)登録日 令和6年8月14日(2024.8.14)

(51)国際特許分類 F I
 G 0 6 F 9/455(2018.01) G 0 6 F 9/455 1 5 0
 G 0 6 F 9/48 (2006.01) G 0 6 F 9/48 1 0 0 H

請求項の数 16 (全46頁)

(21)出願番号	特願2023-512217(P2023-512217)	(73)特許権者	503433420 華為技術有限公司 HUAWEI TECHNOLOGIES CO., LTD. 中華人民共和國 5 1 8 1 2 9 広東省深 チェン 市龍崗区坂田 華為総部 ベ ン 公楼 Huawei Administrat ion Building, Banti an, Longgang Distri ct, Shenzhen, Guang dong 5 1 8 1 2 9, P. R. C hina
(86)(22)出願日	令和3年6月2日(2021.6.2)	(74)代理人	100110364 弁理士 実広 信哉
(65)公表番号	特表2023-538093(P2023-538093 A)		
(43)公表日	令和5年9月6日(2023.9.6)		
(86)国際出願番号	PCT/CN2021/097829		
(87)国際公開番号	WO2022/037182		
(87)国際公開日	令和4年2月24日(2022.2.24)		
審査請求日	令和5年3月30日(2023.3.30)		
(31)優先権主張番号	202010839905.3		
(32)優先日	令和2年8月19日(2020.8.19)		
(33)優先権主張国・地域又は機関	中国(CN)		

最終頁に続く

(54)【発明の名称】 コンピュータデバイス、例外処理方法および割り込み処理方法

(57)【特許請求の範囲】

【請求項1】

コンピュータデバイスであって、前記コンピュータデバイスが、ハードウェア層を備え、hostが、前記ハードウェア層で実行し、前記ハードウェア層が、第1のプロセッサを備え、仮想マシンVMが、前記第1のプロセッサ上で実行し、前記第1のプロセッサが、リダイレクト装置を備え、

前記第1のプロセッサが、前記VMによってトリガされる例外命令を取得するように構成され、

前記リダイレクト装置が、

前記VMの仮想化特権レベルから前記hostのユーザモードに切り替えるように構成され、前記第1のプロセッサが、前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行する

ようにさらに構成され、

前記例外命令が、前記VMによって実施される入力/出力I/O動作によってトリガされる、コンピュータデバイス。

【請求項2】

コンピュータデバイスであって、前記コンピュータデバイスが、ハードウェア層を備え、hostが、前記ハードウェア層で実行し、前記ハードウェア層が、第1のプロセッサを備え、仮想マシンVMが、前記第1のプロセッサ上で実行し、前記第1のプロセッサが、リダイレクト装置を備え、

前記第1のプロセッサが、前記VMによってトリガされる例外命令を取得するように構成され、

前記リダイレクト装置が、
前記VMの仮想化特権レベルから前記hostのユーザモードに切り替えるように構成され、
前記第1のプロセッサが、前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行する
ようにさらに構成され、

前記例外処理関数が、前記hostの前記ユーザモードにおいて実行する仮想プロセッサのスレッド内に備えられ、前記仮想プロセッサの前記スレッドに割り当てられた実行環境が、
メインスレッド実行環境と、前記VMの実行環境と、シャドウ実行環境と、を備え、前記
シャドウ実行環境が、前記メインスレッド実行環境および前記VMの前記実行環境から
独立したユーザ・モード・メモリ空間を有し、前記シャドウ実行環境が、前記例外処理関
数を実行するために使用され、前記例外処理関数が、前記例外命令をトリガする例外を処
理するために使用され、

前記第1のプロセッサが、前記VMのコンテキストからユーザ・モード・シャドウ実行環
境のコンテキストに切り替える、ように具体的に構成され、前記シャドウ実行環境の前記
ユーザ・モード・メモリ空間が、コンテキスト切り替え中に移動される前記VMの前記コ
ンテキストを格納するために使用される、コンピュータデバイス。

【請求項3】

前記リダイレクト装置が、前記例外処理関数のエントリアドレスを第1のレジスタに書
き込むようにさらに構成され、

前記第1のプロセッサが、前記例外処理関数の前記エントリアドレスを前記第1のレジ
スタから読み出し、前記エントリアドレスに基づいて前記例外処理関数を呼び出し、前記ho
stの前記ユーザモードにおいて、前記例外命令に対応する前記例外処理関数を実行する、
ように具体的に構成される、

請求項1または2に記載のデバイス。

【請求項4】

コンピュータデバイスであって、前記コンピュータデバイスが、ハードウェア層を備え
、ホストhostが、前記ハードウェア層で実行し、前記ハードウェア層が、第1のプロセッ
サと、第2のプロセッサと、割り込みコントローラと、を備え、第1の仮想マシンVMが、
前記第1のプロセッサ上で実行し、前記第1のプロセッサが、仮想イベント通知装置を備え

、
前記第1のプロセッサが、前記第1のVMによってトリガされる第1の割り込み要求を取
得するように構成され、前記第1の割り込み要求が、割り込み処理関数を実行するよう
に要求するために使用され、

前記仮想イベント通知装置が、前記第1の割り込み要求に応答して、割り込み情報を前
記割り込みコントローラに送信するように構成され、

前記割り込みコントローラが、前記割り込み情報に基づいて前記第2のプロセッサの識
別子を決定し、前記第2のプロセッサの前記識別子に基づいて第2の割り込み要求を前記第
2のプロセッサに送信し、前記第2の割り込み要求が、前記第2のプロセッサに、前記割り
込み処理関数を実行するように要求するために使用される、ように構成され、

前記第2のプロセッサが、受信した前記第2の割り込み要求に応答して、前記割り込み処
理関数を実行するように構成され、

前記第1の割り込み要求が、仮想イベント割り込み要求番号VEI IDを備え、前記VEI ID
が、前記第1の割り込み要求をトリガする仮想イベントを識別し、

前記第1のプロセッサが、前記第1のVMのVM IDを取得するようにさらに構成され、
前記割り込み情報が、前記第1のVMの前記VEI IDおよび前記VM IDを備え、
前記割り込みコントローラが、前記第1のVMの前記VEI IDおよび前記VM IDに基づいて
、対応する前記第2のプロセッサの識別子と、VEI物理番号と、前記第1の割り込み要求を
処理する前記hostの特権レベルと、を決定し、前記VEI物理番号が、前記host内の前記仮

10

20

30

40

50

想イベントを識別し、前記第2のプロセッサの前記識別子に基づいて第2の割り込み要求を前記第2のプロセッサに送信し、前記第2の割り込み要求が、前記hostの前記特権レベルを示す情報を備える、ように具体的に構成され、

前記第2のプロセッサが、前記VEI物理番号を取得し、前記hostの前記特権レベルに対応する処理方式に基づいて、前記受信した第2の割り込み要求に応答して、前記VEI物理番号に対応する割り込み処理関数を実行する、ように具体的に構成される、

コンピュータデバイス。

【請求項5】

コンピュータデバイスであって、前記コンピュータデバイスが、ハードウェア層を備え、ホストhostが、前記ハードウェア層で実行し、前記ハードウェア層が、第1のプロセッサと、第2のプロセッサと、割り込みコントローラと、を備え、第1の仮想マシンVMが、前記第1のプロセッサ上で実行し、前記第1のプロセッサが、仮想イベント通知装置を備え、

10

前記第1のプロセッサが、前記第1のVMによってトリガされる第1の割り込み要求を取得するように構成され、前記第1の割り込み要求が、割り込み処理関数を実行するように要求するために使用され、

前記仮想イベント通知装置が、前記第1の割り込み要求に応答して、割り込み情報を前記割り込みコントローラに送信するように構成され、

前記割り込みコントローラが、前記割り込み情報に基づいて前記第2のプロセッサの識別子を決定し、前記第2のプロセッサの前記識別子に基づいて第2の割り込み要求を前記第2のプロセッサに送信し、前記第2の割り込み要求が、前記第2のプロセッサに、前記割り込み処理関数を実行するように要求するために使用される、ように構成され、

20

前記第2のプロセッサが、受信した前記第2の割り込み要求に応答して、前記割り込み処理関数を実行するように構成され、

前記第1の割り込み要求が、仮想イベント割り込み要求番号VEI IDを備え、前記VEI IDが、前記第1の割り込み要求をトリガする仮想イベントを識別し、

前記割り込み情報が、前記VEI IDを備え、

前記割り込みコントローラが、前記VEI IDに基づいて、対応する前記第2のプロセッサの識別子と、前記第1の割り込み要求を処理する前記hostの特権レベルと、を決定し、前記第2のプロセッサの前記識別子に基づいて第2の割り込み要求を前記第2のプロセッサに送信し、前記第2の割り込み要求が、前記hostの前記特権レベルを示す情報を備える、ように具体的に構成され、

30

前記第2のプロセッサが、前記VEI IDを取得し、前記hostの前記特権レベルに対応する処理方式に基づいて、前記受信した第2の割り込み要求に応答して、前記VEI IDに対応する割り込み処理関数を実行する、ように具体的に構成される、

コンピュータデバイス。

【請求項6】

前記hostの前記特権レベルが、ユーザモードであり、前記第2のプロセッサが、ハードウェア支援切り替え装置をさらに備え、

前記ハードウェア支援切り替え装置が、

40

第1のスレッドからシャドウスレッドに切り替え、前記第1のスレッドが、切り替え前に前記第2のプロセッサによって実行されるスレッドであり、前記シャドウスレッドが、前記ユーザモードにおける前記host内で実行するスレッドであり、前記シャドウスレッドが、前記割り込み処理関数を実行するために使用される、ように構成され、

前記第2のプロセッサが、前記hostの前記ユーザモードにおける前記割り込み処理関数を実行するように具体的に構成される、

請求項4または5に記載のコンピュータデバイス。

【請求項7】

前記第2のプロセッサが、リダイレクト装置をさらに備え、

前記第2のプロセッサが、前記hostの前記ユーザモードにおける前記割り込み処理関数

50

を実行する前に、前記第1のスレッドが第2のVMの仮想化特権レベルにあるか否かを決定し、前記第2のVMが、前記第2のプロセッサ上で実行する、ようにさらに構成され、

前記リダイレクト装置が、割り込みスレッドが前記第2のVMの前記仮想化特権レベルにある場合、前記第2のVMの前記仮想化特権レベルを前記hostの前記ユーザモードに切り替えるように構成される、

請求項6に記載のコンピュータデバイス。

【請求項8】

前記リダイレクト装置が、前記割り込み処理関数のエントリアドレスを第1のレジスタに書き込むようにさらに構成され、

前記第2のプロセッサが、前記第1のレジスタから前記割り込み処理関数の前記エントリアドレスを読み出し、前記エントリアドレスに基づいて前記割り込み処理関数を呼び出し、前記hostの前記ユーザモードにおける前記割り込み処理関数を実行する、ように具体的に構成される、

10

請求項7に記載のコンピュータデバイス。

【請求項9】

前記hostの前記特権レベルが、カーネルモードであり、

前記第2のプロセッサが、

前記hostの前記カーネルモードのものであり、前記割り込みコントローラによって送信された前記第2の割り込み要求を受信し、

前記hostの前記カーネルモードの割り込み処理関数のエントリアドレスに切り替え、前記割り込み処理関数が、前記第2の割り込み要求に対応し、

20

前記hostの前記カーネルモードにおいて、前記カーネルモードの前記割り込み処理関数を実行する、

ように構成される、

請求項4または5に記載のコンピュータデバイス。

【請求項10】

例外処理方法であって、前記方法が、コンピュータデバイス、ホストhost、および前記コンピュータデバイス上で実行される仮想マシンVMに適用され、前記方法が、

前記VMによってトリガされる例外命令を取得するステップと、

前記VMの仮想化特権レベルから前記hostのユーザモードに切り替えるステップと、

30

前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行するステップと、

を含み、

前記例外命令が、前記VMによって実施される入力/出力I/O動作によってトリガされる、例外処理方法。

【請求項11】

例外処理方法であって、前記方法が、コンピュータデバイス、ホストhost、および前記コンピュータデバイス上で実行される仮想マシンVMに適用され、前記方法が、

前記VMによってトリガされる例外命令を取得するステップと、

前記VMの仮想化特権レベルから前記hostのユーザモードに切り替えるステップと、

40

前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行するステップと、

を含み、

前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行する前記ステップが、

前記VMのコンテキストから前記ユーザモードのシャドウ実行環境のコンテキストに切り替え、前記hostの前記ユーザモードにおける例外を処理するステップであって、前記例外処理関数が、前記hostの前記ユーザモードにおいて実行する仮想プロセッサのスレッド内に備えられ、前記仮想プロセッサの前記スレッドに割り当てられた実行環境が、メインスレッド実行環境と、前記VMの実行環境と、前記シャドウ実行環境と、を備え、前記シャ

50

ドウ実行環境が、前記メインスレッド実行環境および前記VMの前記実行環境から独立したユーザ・モード・メモリ空間を有し、前記シャドウ実行環境が、前記例外処理関数を実行するために使用され、前記例外処理関数が、前記例外命令をトリガする例外を処理するために使用される、ステップを含む、例外処理方法。

【請求項 1 2】

前記方法が、
前記例外処理関数のエントリアドレスを第1のレジスタに書き込むステップをさらに含み、
前記hostの前記ユーザモードにおいて、前記例外命令に対応する例外処理関数を実行する前記ステップが、

10

前記第1のレジスタから前記例外処理関数の前記エントリアドレスを読み出し、前記エントリアドレスに基づいて前記例外処理関数を呼び出し、前記hostの前記ユーザモードにおいて、前記例外命令に対応する前記例外処理関数を実行するステップを含む、請求項10または11に記載の方法。

【請求項 1 3】

割り込み処理方法であって、前記方法が、コンピュータデバイス、ホストhost、および前記コンピュータデバイス上で実行される第1の仮想マシンVMに適用され、前記方法が、前記第1のVMによってトリガされる第1の割り込み要求を取得するステップであって、前記第1の割り込み要求が、割り込み処理関数の実行を要求するために使用される、ステップと、

20

前記第1の割り込み要求に応答して、割り込み情報を決定するステップと、
前記割り込み情報に基づいて第2の割り込み要求を生成するステップであって、前記第2の割り込み要求が、前記割り込み処理関数の実行を要求するために使用される、ステップと、

前記第2の割り込み要求に応答して、前記割り込み処理関数を実行するステップと、
を含む、

前記第1の割り込み要求が、仮想イベント割り込み要求番号VEI IDを備え、前記VEI IDが、前記第1の割り込み要求をトリガする仮想イベントを識別し、前記方法が、

前記第1のVMのVM IDを取得するステップであって、前記割り込み情報が、前記第1のVMの前記VEI IDおよび前記VM IDを備える、ステップと、

30

前記第1のVMの前記VEI IDおよび前記VM IDに基づいて、前記第1の割り込み要求を処理する前記hostの対応するVEI物理番号および特権レベルを決定するステップであって、前記VEI物理番号が、前記host内の前記仮想イベントを識別する、ステップと、をさらに含み、

前記第2の割り込み要求に応答して、前記割り込み処理関数を実行する前記ステップが、前記VEI物理番号を取得し、前記hostの前記特権レベルに対応する処理方式に基づいて、前記第2の割り込み要求に応答して、前記VEI物理番号に対応する割り込み処理関数を実行するステップ

を含む、割り込み処理方法。

40

【請求項 1 4】

割り込み処理方法であって、前記方法が、コンピュータデバイス、ホストhost、および前記コンピュータデバイス上で実行される第1の仮想マシンVMに適用され、前記方法が、前記第1のVMによってトリガされる第1の割り込み要求を取得するステップであって、前記第1の割り込み要求が、割り込み処理関数の実行を要求するために使用される、ステップと、

前記第1の割り込み要求に応答して、割り込み情報を決定するステップと、

前記割り込み情報に基づいて第2の割り込み要求を生成するステップであって、前記第2の割り込み要求が、前記割り込み処理関数の実行を要求するために使用される、ステップと、

50

前記第2の割り込み要求に応答して、前記割り込み処理関数を実行するステップと、
を含み、

前記第1の割り込み要求が、仮想イベント割り込み要求番号VEI IDを備え、前記VEI ID
が、前記第1の割り込み要求をトリガする仮想イベントを識別し、前記割り込み情報が、
前記VEI IDを備え、前記方法が、

前記VEI IDに基づいて、前記第1の割り込み要求を処理する前記hostの特権レベルを決
定するステップをさらに含み、

前記第2の割り込み要求に応答して、前記割り込み処理関数を実行する前記ステップが、
前記VEI IDを取得し、前記hostの前記特権レベルに対応する処理方式に基づいて、前記
第2の割り込み要求に応答して、前記VEI IDに対応する割り込み処理関数を実行するステ
ップ

10

を含む、割り込み処理方法。

【請求項15】

前記hostの前記特権レベルが、ユーザモードであり、前記第2の割り込み要求に
応答して、前記割り込み処理関数を実行する前記ステップが、

第1のスレッドからシャドウスレッドに切り替えるステップであって、前記第1の
スレッドが、切り替え前に前記ホストによって実行されるスレッドであり、前記
シャドウスレッドが、前記ユーザモードにおける前記host内で実行するスレ
ッドであり、前記シャドウスレッドが、前記割り込み処理関数を実行するた
めに使用される、ステップと、

前記hostの前記ユーザモードにおける前記割り込み処理関数を実行する
ステップと、

20

を含む、請求項13または14に記載の方法。

【請求項16】

前記hostの前記特権レベルが、カーネルモードであり、前記hostの前記特権
レベルに対応する処理方式に基づいて、前記受信した第2の割り込み要求に
応答して、前記VEI IDに対応する前記割り込み処理関数を実行する前記
ステップが、

前記hostの前記カーネルモードの割り込み処理関数のエントリアドレスに
切り替えるステップであって、前記割り込み処理関数が、前記第2の割
り込み要求に対応する、ステップと、

前記hostの前記カーネルモードにおいて、前記カーネルモードの前記割
り込み処理関数を実行するステップと、

30

を含む、請求項13または14に記載の方法。

【発明の詳細な説明】

【技術分野】

【0002】

本出願は、コンピュータ技術の分野に関し、具体的には、コンピュータデバイス、
例外処理方法、および割り込み処理方法に関する。

【背景技術】

【0003】

仮想化システムでは、仮想デバイス・シミュレーション・ソリューションは、
入力/出力(input/output、I/O)仮想化を実施するために、主に使用される。
仮想デバイス・シミュレーション・ソリューションは、仮想I/Oを処理する
モードに基づいて、同期処理と非同期処理とに分類され得る。

40

【0004】

同期処理手順では、仮想マシン(virtual machine、VM)が仮想I/Oデバイス
(例えば、シリアル・ポート・メモリに対する読み出し動作もしくは書き込み
動作、または別の仮想I/Oデバイスに対するアクセス動作)に関連する動作
を実施するとき、例外命令がトリガされる。VMは、例外命令を、ホスト
(host)のカーネルモードにおける仮想マシンモニタ(virtual machine
monitor、VMM)に送信し、VMMは、hostのユーザモードにおける例外
を処理することを実施するために、例外処理を、ユーザモードにおける
hostに転送する。例外が処理された後、処理結果は、ユーザモードにお
けるhostからVMMに

50

転送され、次に、VMMからVMに転送される。この手順では、VMとVMMとの間の切り替えのオーバーヘッドが、2回生成され、hostのカーネルモードとhostのユーザモードとの間の切り替えのオーバーヘッドが、2回生成される。

【0005】

非同期処理手順では、VMがディスクのメモリ上で、またはネットワークアダプタのメモリ上で、読み出し動作または書き込み動作を実施するとき、割り込み要求がトリガされる。割り込み要求は、非同期方式で、リモートプロセッサ上のクイックエミュレータ(quick emulator、Qemu)スレッドに、割り込み要求の処理を完了するように通知する必要がある。リモートプロセッサは、VMの実行をサポートするローカルプロセッサとは異なる別のプロセッサである。ローカルプロセッサ上で、VMは、例外命令をトリガし、次に、VMは、例外命令を、hostのカーネルモードにおけるVMMに送信する。次に、VMMは、ユーザモード例外命令として命令を識別し、VMMは、例外命令を、hostのカーネルモードからhostのユーザモードに転送し、システムコールが、hostのカーネルモードに入るために、hostのユーザモードからトリガされ、次に、割り込み要求は、hostのカーネルモードにおいてトリガされ、割り込み要求は、リモートプロセッサに転送される。リモートプロセッサのQemuスレッドは、カーネルモードにおけるスケジューラをウェイクアップさせる必要がある。この手順では、ディスクなどの物理デバイスにアクセスするために、カーネルモードは、ユーザモードに切り替えられ、次に、システムコールを使用することによってカーネルモードに戻るよう切り替えられる必要がある。非同期処理手順では、ユーザモードとカーネルモードとの間の切り替えも頻繁に実施される必要がある、ことが分かる。

10

20

【0006】

同期処理手順と、非同期処理手順との両方で、オーバーヘッドが、切り替え中に複数回生成され、I/O性能が影響される。

【発明の概要】

【課題を解決するための手段】

【0007】

本出願の実施形態は、切り替え中に生成したオーバーヘッドを低減し、コンピュータデバイスの性能を改善するために、コンピュータデバイス、例外処理方法、および割り込み処理方法を提供する。本出願は、対応する装置、コンピュータ可読記憶媒体、コンピュータプログラム製品などを、さらに提供する。

30

【0008】

本出願の第1の態様は、コンピュータデバイスを提供し、コンピュータデバイスが、ハードウェア層を含み、ホスト(host)が、ハードウェア層で実行し、ハードウェア層が、第1のプロセッサを含み、仮想マシン(virtual machine、VM)が、第1のプロセッサ上で実行し、第1のプロセッサが、リダイレクト装置を含む。第1のプロセッサは、VMによってトリガされる例外命令を取得するように構成される。リダイレクト装置は、VMの仮想化特権レベルからhostのユーザモードに切り替えるように構成される。第1のプロセッサは、hostのユーザモードにおいて、例外命令に対応する例外処理関数を実行するようにさらに構成される。

40

【0009】

第1の態様では、VMによってトリガされる例外が、同期処理方式で処理される。

【0010】

本出願の本実施形態におけるプロセッサは、プロセッサコアであってもよく、1つのプロセッサコアは、1つの時間期間内の1つのVMの実行をサポートすることに留意されたい。複数のVMは、時分割多重化方式で、同じ処理コアを使用し得る。

【0011】

第1の態様では、コンピュータデバイスは、携帯電話、タブレットコンピュータ、デスクトップコンピュータ、ノートブックコンピュータ、ウルトラ・モバイル・パーソナル・コンピュータ(ultra-mobile personal computer、UMPC)、ハンドヘルドコンピュ

50

ータ、またはネットブックなどの端末デバイスであってもよく、または、コンピュータデバイスは、サーバであってもよい。ハードウェア層は、プロセッサ、メモリ、および通信インターフェースなどのハードウェアデバイスを含み得る。VMおよびhostは、ソフトウェアによってシミュレートされ、ハードウェア層で実行される。第1のプロセッサに含まれるリダイレクト装置は、ハードウェア回路を使用することによって実装され得る。VMが、入力/出力(input/output、I/O)動作を実施するとき、例外命令は、トリガされ得る。VMが、仮想I/Oデバイスにアクセスする必要があるとき、I/O動作は、トリガされる動作であってもよい。仮想I/Oデバイスは、フロント・エンド・ドライバと、バック・エンド・デバイスと、を一般に含む。フロント・エンド・ドライバは、VM上で実行し、バック・エンド・デバイスのプログラムは、host上で実行する。VMによって実施される動作が、フロント・エンド・ドライバを使用することによってバック・エンド・デバイスにアクセスする必要がある場合、例外命令は、トリガされ、バック・エンド・デバイスは、例外命令を使用することによってアクセスされる。仮想I/Oデバイスは、バック・エンド・デバイスの1つのタイプである。仮想I/Oデバイスへのアクセスは、仮想I/Oデバイスのメモリからのデータの読み出し、または仮想I/Oデバイスのメモリへのデータの書き込みであってもよい。仮想I/Oデバイスは、仮想マシンによって使用され得るI/Oデバイスである。仮想I/Oデバイスは、ソフトウェアを使用することによってシミュレートされた完全に仮想化されたI/Oデバイスであってもよい。仮想I/Oデバイスはまた、マッピング方式で、物理I/Oデバイスにマッピングされてもよい。仮想I/Oデバイスが、アクセスされるとき、物理I/Oデバイスは、最終的にアクセスされる。

10

20

【0012】

第1のVMが、第1のVMの許可を超える動作を実施するとき、例外命令は、例外が発生する、ことであってもよい。第1のプロセッサは、メモリに格納されたコンピュータ命令を呼び出し、対応する検出口ジックを実行することによって、例外コードを検出し得る。このステップはまた、第1のプロセッサがハードウェア検出口ジックを実行することによって例外コードを検出し、次に例外命令を取得する、ことであってもよい。例外命令を取得した後、第1のプロセッサは、後続の例外処理のために、例外の原因およびタイプ、第1のVMによってアクセスされるメモリ、ならびにリターンアドレスを、ハードウェアレジスタに書き込み得る。

【0013】

VMの仮想化特権レベルは、仮想化モード、すなわちVMの作動モードにおける特権レベルを表し、仮想化モードと呼ばれる場合もある。仮想化特権レベルは、仮想化ユーザモードと、仮想化カーネルモードと、を含む。VMの仮想化特権レベルからhostのユーザモードへの切り替え(直接切り替え)は、VMの仮想化特権レベルに関する情報を、特権レベルを格納するために使用されるレジスタから移動させ、次に、hostのユーザモードに関する情報をレジスタに書き込む手順を含んでもよく、レジスタは、システムレジスタであってもよい。本出願では、特権レベルおよび仮想化は、1ビットを使用することによって、それぞれ表され得る。例えば、0は、ユーザモードを示し、1は、カーネルモードを示し、0は、非仮想化を示し、1は、仮想化を示す。特権レベルは、第1のビットにあり、仮想化するか否かは、第2のビットにある。例えば、00は、非仮想化ユーザモードを示し、ホストのユーザモードと呼ばれる場合もあり、10は、非仮想化カーネルモードを示し、ホストのカーネルモードと呼ばれる場合もあり、01は、仮想化ユーザモードを示し、VMのユーザモードと呼ばれる場合もあり、11は、仮想化カーネルモードを示し、VMのカーネルモードと呼ばれる場合もある。

30

40

【0014】

ユーザモードhostは、ユーザモードにおいて作動するホストを指す。hostのユーザモードは、ホストの作動モードを指す。ホストの作動モードは、ユーザモードと、カーネルモードと、を含む。

【0015】

リダイレクト装置がVMの仮想化特権レベルからhostのユーザモードに切り替えるよう

50

に構成されることは、リダイレクト装置が第1のプロセッサの作動モードを切り替えるように構成されること、として理解され得る。具体的には、リダイレクト装置は、第1のプロセッサの作動モードを、仮想化モードから非仮想化モードにおけるユーザモード（hostのユーザモード）に切り替えるように構成される。プロセッサの作動モードは、プロセッサの作動状態として理解され得、プロセッサの作動状態は、前述のシステムレジスタ内のビットの値によって表され得る。

【0016】

例外処理関数は、hostのユーザモードにおいて実行する仮想プロセッサ（virtual central processing unit、vCPU）のスレッドに含まれ得る。第1のプロセッサは、hostのユーザモードにおいて、例外命令に対応する例外処理関数を実行し得る。例外命令は、例外タイプに基づいて例外処理関数に関連付けられ得る。異なる例外タイプは、異なる例外処理関数に対応し得る。例えば、シリアルポート上でVMによって実施される書き込み動作に対応する例外処理関数は、グラフィックカード上でVMによって実施される書き込み動作に対応する例外処理関数とは異なる、または、すべての例外命令は、同じ例外処理関数に対応してもよく、例外処理関数が実行されるとき、異なるデータは、異なる例外命令を処理するために、呼び出される。例えば、例外命令が、VMによって実施されるI/O動作によってトリガされる場合、対応するI/O動作は、例外処理関数を実行すること、例えば、I/Oデバイスのメモリ内のデータを読み出すこと、またはI/Oデバイスのメモリにデータを書き込むことによって、完了される。例外命令が、VMによって実施される別の動作（例えば、ページフォルト）によってトリガされる場合、第1のプロセッサは、ページフォルトを処理するために、例外処理関数を実行する。言い換えると、リダイレクト装置は、仮想化モードから非仮想化モードに切り替えるように構成される。したがって、ユーザモードにおける割り込みまたは例外は、処理のためにhostのユーザ空間にリダイレクトされ得、すなわち、ユーザモードにおけるhostは、例外または割り込み処理関数を実行する。

【0017】

コンピュータデバイスが、リダイレクト装置を使用することによって、例外命令を、VMの仮想化特権レベルからhostのユーザモードに直接切り替え、hostのユーザモードにおいて例外処理関数を実行し得る第1の態様から、それは分かる。VMの仮想化特権レベルが、hostのカーネルモードに、最初に切り替えられる必要があり、次に、hostのカーネルモードが、hostのユーザモードに切り替えられる必要がある従来の技術と比較して、コンピュータデバイスは、切り替えオーバーヘッドを低減し、コンピュータデバイスの性能を改善し得る。

【0018】

第1の態様の1つの可能な実施態様では、リダイレクト装置は、例外処理関数のエントリアドレスを第1のレジスタに書き込むようにさらに構成される。第1のプロセッサは、例外処理関数のエントリアドレスを第1のレジスタから読み出し、エントリアドレスに基づいて例外処理関数を呼び出し、hostのユーザモードにおいて、例外命令に対応する例外処理関数を実行する、ように具体的に構成される。

【0019】

この可能な実施態様では、リダイレクト装置は、例外処理関数のエントリアドレス（trap handlerと呼ばれる場合もある）を第1のレジスタに書き込まれてもよい。このようにして、第1のプロセッサは続いて、エントリアドレスを第1のレジスタから読み出し、エントリアドレスに基づいて例外処理関数を呼び出し、hostのユーザモードにおける例外処理関数を実行する。リダイレクト装置における「リダイレクト」は、特権レベルの切り替えと、例外処理関数のエントリアドレスを第1のレジスタに書き込む手順と、を含み得る。この可能な実施態様では、リダイレクト装置は、エントリアドレスを第1のレジスタに書き込み、これにより、第1のプロセッサは、例外処理関数を迅速に実行してもよい。

【0020】

第1の態様の1つの可能な実施態様では、例外処理関数は、hostのユーザモードにおいて

実行する仮想プロセッサのスレッド内に含まれ、仮想プロセッサのスレッドに割り当てられた実行環境は、メインスレッド実行環境と、VMの実行環境と、シャドウ実行環境と、を含み、仮想プロセッサのスレッドに割り当てられた実行環境が、仮想プロセッサのスレッドを実行するために必要なリソースであり、シャドウ実行環境が、メインスレッド実行環境およびVMの実行環境から独立したユーザ・モード・メモリ空間を有する。シャドウ実行環境は、例外処理関数を実行するために使用され、例外処理関数は、例外命令をトリガする例外を処理するために使用される。第1のプロセッサは、VMのコンテキストからユーザ・モード・シャドウ実行環境のコンテキストに切り替えるように構成され、シャドウ実行環境のユーザ・モード・メモリ空間は、コンテキスト切り替え中に移動されるVMのコンテキストを格納するために使用される。

10

【0021】

本出願では、例外処理関数を実行することは、コンテキスト切り替えの手順と、hostのユーザモードにおける例外の処理と、を含む。

【0022】

この可能な実施態様では、仮想プロセッサのスレッドが実行する前に、仮想プロセッサのスレッドは、リソースのために、ホストに適用し、ホストは、実行のために必要なリソースを仮想プロセッサのスレッドに割り当てる。リソースは、実行環境と呼ばれる場合もある。本出願における「実行環境」は、ソフトウェア実行環境と、ハードウェア実行環境との2つの態様を含む。ソフトウェア実行環境は、仮想プロセッサのスレッドが異なる実行環境で実行するとき使用されるスタック内の情報と、スレッドのコンテキスト情報と、を含む。ハードウェア態様は、メモリリソースまたは処理リソースを含む、スレッドが実行するとき使用されるハードウェアリソースを含む。メインスレッド実行環境、VMの実行環境、およびシャドウ実行環境のすべては、独立したメモリ空間を有する。メインスレッド実行環境は、ユーザモードにおけるメインスレッドの実行環境と、カーネルモードにおけるメインスレッドの実行環境と、を含む。カーネルモードにおけるメインスレッドの実行環境は、VMおよびhostのコンテキストを格納するために使用される。シャドウ実行環境は、ユーザ・モード・シャドウ実行環境と、カーネル・モード・シャドウ実行環境と、を含む。ユーザ・モード・シャドウ実行環境は、シャドウ実行環境のユーザ・モード・メモリ空間である。ユーザ・モード・メモリ空間は、ユーザ・モード・シャドウ実行環境に対応するメモリアドレス空間を表し、ユーザ・モード・メモリ空間は、コンテキスト切り替え中にレジスタから移動されるVMのコンテキストを格納するために使用される。シャドウ実行環境が、独立したユーザ・モード・メモリ空間を有するので、ユーザ・モード・メモリ空間は、VMのコンテキストを格納し得、そして、メインスレッド実行環境を使用することによって、VMの仮想化特権レベルからhostのカーネルモードに、最初に切り替える必要も、次に、カーネルモードからユーザモードに切り替える必要もない。このようにして、切り替えオーバーヘッドは、さらに低減され得、コンピュータデバイスの性能が改善される。

20

30

【0023】

第1の態様の1つの可能な実施態様では、第1のプロセッサは、hostのカーネルモードにおいて処理される必要がある割り込みまたは例外が、例外処理関数が実行されるときに生成される場合、ユーザ・モード・シャドウ実行環境のコンテキストからカーネル・モード・シャドウ実行環境のコンテキストに切り替え、hostのカーネルモードにおいて、hostのカーネルモードにおいて処理される必要がある割り込みまたは例外を処理し、例外処理関数の実行を継続するために、カーネルモードのシャドウ実行環境のコンテキストからユーザ・モード・シャドウ実行環境のコンテキストに戻るよう切り替える、ようにさらに構成される。

40

【0024】

この可能な実施態様では、シャドウ実行環境は、カーネル・モード・シャドウ実行環境、すなわち、カーネル・モード・ユーザ空間を、さらに含む。ユーザ・モード・シャドウ実行環境のコンテキストが、カーネル・モード・シャドウ実行環境のコンテキストに切り

50

替えられるとき、ユーザ・モード・シャドウ実行環境のものであり、レジスタから切り替えられるコンテキストを格納するために、カーネル・モード・ユーザ空間は使用される。この可能な実施態様は、例外処理関数が実行されるとき、カーネルモードにおいて実行される必要がある割り込みまたは例外が処理され、コンピュータシステムの互換性が改善されることを保証し得る。

【 0 0 2 5 】

本出願の第2の態様は、コンピュータシステムを提供する。コンピュータシステムは、第1の態様で提供されるコンピュータシステムに基づいて、第2のプロセッサと、割り込みコントローラと、をさらに含む。例外処理関数が第1の態様で実行される前の手順については、第1の態様の内容を参照されたい。違いは、第1のプロセッサが、例外処理関数を実行せず、しかし、アクセス要求を割り込みコントローラに送信し、アクセス要求が、割り込みコントローラ内のユーザ・モード・アドレス・マッピング空間にアクセスすることを示す、ことにある。次に、割り込みコントローラは、割り込み要求を第2のプロセッサに送信し、第2のプロセッサは、割り込み要求に基づいて例外処理関数を実行する。

10

【 0 0 2 6 】

第2の態様によれば、hostのユーザモードへのリダイレクトは、第1の態様の同期処理ソリューションにおけるリダイレクト装置を使用することによって実施され、次に、非同期処理手順が、割り込みコントローラ内のユーザ・モード・アドレス・マッピング空間にアクセスすることによって実施される。第2の態様によれば、VMの特権レベルからhostのユーザモードへの切り替えが、リダイレクト装置を使用することによって実施される。したがって、非同期処理の切り替えオーバーヘッドが低減され、コンピュータデバイスの性能が改善される。

20

【 0 0 2 7 】

本出願の第3の態様は、コンピュータデバイスを提供し、コンピュータデバイスが、ハードウェア層を含み、ホスト(host)が、ハードウェア層で実行し、ハードウェア層が、第1のプロセッサと、第2のプロセッサと、割り込みコントローラと、を含み、仮想マシン、第1のVMが、第1のプロセッサ上で実行し、第1のプロセッサが、仮想イベント通知装置を含む。第1のプロセッサは、第1のVMによってトリガされる第1の割り込み要求を取得するように構成され、第1の割り込み要求が、割り込み処理関数を実行するように要求するために使用され、仮想イベント通知装置は、第1の割り込み要求に応答して、割り込み情報を割り込みコントローラに直接送信するように構成され、割り込みコントローラは、割り込み情報に基づいて第2のプロセッサの識別子を決定し、第2のプロセッサの識別子に基づいて第2の割り込み要求を第2のプロセッサに送信する、ように構成され、第2のプロセッサは、受信した第2の割り込み要求に応答して、割り込み処理関数を実行するように構成される。

30

【 0 0 2 8 】

第3の態様では、第1のVMによってトリガされる割り込みは、非同期処理方式で処理される。非同期処理は、クロスプロセッサ処理手順、すなわち、第1のプロセッサ上で実行している第1のVMが割り込みを実行することを指し、割り込み処理手順は、第2のプロセッサを使用することによって完了され得る。

40

【 0 0 2 9 】

第3の態様では、仮想イベント通知装置が、ハードウェア回路を使用することによって実装されてもよい。第1のVMが、I/O動作を実施するとき、第1の割り込み要求は、トリガされてもよく、第1のVMが、仮想I/Oデバイスにアクセスする必要があるときに、I/O動作は、トリガされる動作であってもよい。割り込み情報は、第1のVMにおいて発生する仮想的なイベントを示す。割り込みを非同期方式で処理し、かつ第3の態様において提供されるコンピュータデバイスによれば、仮想イベント通知装置は、hostのカーネルモードにおいてトラップされることなく、割り込みイベントを処理し得る。したがって、特権レベル切り替えオーバーヘッドが低減され、コンピュータデバイスの性能が改善される。

【 0 0 3 0 】

50

第3の態様の可能な実施態様では、第1の割り込み要求は、仮想イベント割り込み要求番号VEI IDを含み、VEI IDは、第1の割り込み要求をトリガする仮想イベントを識別する。第1のプロセッサは、第1のVMのVM IDを取得するようにさらに構成される。割り込み情報は、第1のVMのVEI IDおよびVM IDを含む。割り込みコントローラは、第1のVMのVEI IDおよびVM IDに基づいて、対応する第2のプロセッサの識別子と、VEI物理番号と、第1の割り込み要求を処理するhostの特権レベルと、を決定し、VEI物理番号が、host内の仮想イベントを識別し、第2のプロセッサの識別子に基づいて第2の割り込み要求を第2のプロセッサに送信し、第2の割り込み要求が、hostの特権レベルを示す情報を含む、ように具体的に構成される。第2のプロセッサは、VEI物理番号を取得し、hostの特権レベルに対応する処理方式に基づいて、受信した第2の割り込み要求に応答して、VEI物理番号に対応する割り込み処理関数を実行する、ように具体的に構成される。

10

【0031】

この可能な実施態様では、VEI IDは、仮想マシン内で発生する仮想イベントを識別し、VEI物理番号は、host内の仮想イベントの識別子を示す。VEI IDとVEI物理番号との間にはマッピング関係がある。互いにマッピングされたVEI IDとVEI物理番号は、同じ仮想イベント（例えば、ディスク読み出し/書き込みイベント、またはネットワークアダプタ読み出し/書き込みイベント）を示す。VEI IDは、例えば0から63の値によって表されてもよい。各値は、異なる仮想イベントを識別する。1つのI/Oデバイスは、複数の異なる仮想イベントを有し得る。例えば、ディスク読み出しイベントは、値1によって表されてもよく、ディスク書き込みイベントは、値2によって表されてもよく、ディスク制御イベントは、値3によって表されてもよく、ネットワークアダプタ読み出しイベントは、値4によって表されてもよく、ネットワークアダプタ書き込みイベントは、値5によって表されてもよい、などである。各仮想マシン内の同じ仮想イベントは、同じVEI IDを使用し得る。VEI物理番号は、グローバルに一意であってもよく、異なる仮想マシン上で発生する同じVEI IDの仮想イベントは、異なるVEI物理番号に対応してもよい。例えば、VM1のVEI IDの値の範囲に対応するVEI物理番号は、0から63であってもよく、VM2のVEI IDの値の範囲に対応するVEI物理番号は、64から127であってもよく、残りは、類推によって推論されてもよい。このようにして、第2のプロセッサは、異なるVEI物理番号の範囲に基づいて、仮想イベントが発生する仮想マシンを決定し、次に、VEI物理番号の特定の値に基づいて、VM上で発生する仮想イベントを決定し得る。例えば、VEI IDが5であることは、ディスク読み出しイベントを示し、VEI物理番号が5である場合、それは、VM1上でディスク読み出しイベントが発生したことを示す。VEI物理番号が69である場合、それは、VM2上でディスク読み出しイベントが発生したことを示す。仮想イベントとVEI IDとの間の対応関係、およびVEI物理番号と仮想マシンとの間の対応関係は、要求事項に基づいて設定されてもよい。これは、本出願において限定されない。

20

30

【0032】

本出願では、各仮想マシン内の同じ仮想イベントが、同じVEI IDを使用し得るので、仮想イベントが発生する仮想マシン間を区別するために、第1のプロセッサは、VM IDを取得し得、VM IDが、仮想マシン識別子を格納するためのレジスタから第1のプロセッサによって読み出され得る。次に、仮想イベント通知装置は、VEI IDおよびVM IDを割り込みコントローラに送信し、割り込みコントローラは、VEI IDおよびVM IDに基づいて、第2のプロセッサの識別子、VEI物理番号、および対応するhostの特権レベルを検索し得る。入力が、(VEI ID) + (VM ID)であり、出力が、(CPU ID) + (VEI物理番号) + (hostの特権レベル)であるマッピングテーブルを、割り込みコントローラは保持し得る。割り込みコントローラは、マッピングテーブルに基づいて、対応する(CPU ID) + (VEI物理番号) + (hostの特権レベル)を直接決定し得る。割り込みコントローラは、CPU IDに基づいて、非同期割り込み処理を続いて実施する第2のプロセッサを決定し得、hostの特権レベルに基づいて、続いてユーザモード処理方式を使用するか、またはカーネルモード処理方式を使用するか、を決定し得る。第2のプロセッサは、VEI物理番号を格納するレジスタから、VEI物理番号を読み出し、VEI物理番号を使用することによって、VEI物理番

40

50

号に対応する割り込み処理関数を決定し得、次に、ユーザモード処理手順またはカーネルモード処理手順に基づいて、割り込み処理関数を実行し得る。コンピュータデバイスにおいて、VEI物理番号と割り込み処理関数との間のマッピング関係は、メモリを使用することによって保持され得、第2のプロセッサは、VEI物理番号に基づいてメモリから、対応する割り込み処理関数を読み出し得る。この可能な実施態様では、仮想マシン上で発生する仮想イベントは、VEI IDおよびVM IDを使用することによって迅速に決定され得、次に、仮想イベント通知装置および割り込みコントローラは、hostのカーネルモードにトラップされることなく、割り込みイベントを処理し得る。したがって、特権レベル切り替えオーバーヘッドが低減され、コンピュータデバイスの性能が改善される。

【0033】

10

第3の態様の可能な実施態様では、第1の割り込み要求は、仮想イベント割り込み要求番号VEI IDを含み、VEI IDは、第1のVMに第1の割り込み要求をトリガさせる仮想イベントを識別し、割り込み情報は、VEI IDを含む。割り込みコントローラは、VEI IDに基づいて、対応する第2のプロセッサの識別子と、第1の割り込み要求を処理するhostの特権レベルと、を決定し、第2のプロセッサの識別子に基づいて第2の割り込み要求を第2のプロセッサに送信し、第2の割り込み要求が、hostの特権レベルを示す情報を含む、ように具体的に構成される。第2のプロセッサは、VEI IDを取得し、hostの特権レベルに対応する処理方式に基づいて、受信した第2の割り込み要求に回答して、VEI IDに対応する割り込み処理関数を実行する、ように具体的に構成される。

【0034】

20

この可能な実施態様では、VEI IDは、グローバルに一意であってもよく、異なるVMは、同じ仮想イベントに対して異なるVEI IDを使用する。このようにして、VM IDは、使用される必要がなく、VEI物理番号は、変換される必要がなく、対応する第2のプロセッサの識別子、およびhostの特権レベルがVEI IDに基づいて決定されれば、後続の割り込み処理が、実施され得る。第2のプロセッサは、VEI IDを格納するレジスタからVEI IDを取得し得る。この可能な実施態様では、仮想マシン上で発生する仮想イベントは、VEI IDを使用することによって迅速に決定され得、次に、仮想イベント通知装置および割り込みコントローラは、hostのカーネルモードにトラップされることなく、割り込みイベントを処理し得る。したがって、特権レベル切り替えオーバーヘッドが低減され、コンピュータデバイスの性能が改善される。

【0035】

30

第3の態様の可能な実施態様では、hostの特権レベルは、ユーザモードであり、割り込みコントローラは、シャドウスレッドに関する情報をさらに設定し得る。第2のプロセッサは、ハードウェア支援切り替え装置をさらに含む。ハードウェア支援切り替え装置は、第1のスレッドからシャドウスレッドに切り替え、第1のスレッドが、切り替え前に第2のプロセッサによって実行されるスレッドであり、シャドウスレッドが、ユーザモードにおけるhost内で実行するスレッドであり、シャドウスレッドが、割り込み処理関数を実行するために使用される、ように構成される。第2のプロセッサは、hostのユーザモードにおける割り込み処理関数を実行するように具体的に構成される。

【0036】

40

この可能な実施態様では、第1のスレッドは、割り込みスレッドと呼ばれる場合もあり、または第1のスレッドは、割り込みスレッドに関する情報をシャドウスレッドに関する情報に切り替えてもよく、割り込みスレッドが、切り替え前に第2のプロセッサによって実行しているスレッドである。第2のプロセッサは、シャドウスレッドの割り込み処理関数のエントリアドレスに切り替え、hostのユーザモードにおける割り込み処理関数を実行し、割り込み処理関数が、第2の割り込み要求に対応する、ようにさらに構成される。

【0037】

この可能な実施態様では、ハードウェア支援切り替え装置は、ハードウェア回路を使用することによって実装されてもよい。スレッド切り替えは、スレッドに関する切り替え情報を含み得、第1のスレッドに関する情報は、第1のスレッドのメモリアドレス空間、ペー

50

ジテーブル先頭アドレス、データ構造情報、処理特権レベルなどを含む。シャドウスレッドは、ユーザモードにおいてhost内で実行しているスレッドであり、割り込み処理関数を実行するように構成される。シャドウスレッドに関する情報は、シャドウスレッドのメモリアドレス空間、ページテーブル開始アドレス、データ構造情報、処理特権レベルなどを含む。シャドウスレッドは、第1のスレッドのシャドウとして理解され得る。第2のプロセッサによって実行している第1のスレッドが割り込まれた後、第2のプロセッサは、割り込み処理を実施するために、シャドウスレッドを実行する。この可能な実施態様では、第1のスレッドに関する情報、およびシャドウスレッドに関する情報は、hostのカーネルモードにトラップされることなく、ハードウェア支援切り替え装置を使用することによって迅速に切り替えられ得、これにより、割り込みは、hostのユーザモードにおいて直接処理され得る。したがって、切り替えオーバーヘッドが、さらに低減され、コンピュータデバイスの性能が改善される。

10

【0038】

第3の態様の可能な実施態様では、第2のプロセッサは、リダイレクト装置をさらに含み、第2のプロセッサは、hostのユーザモードにおける割り込み処理関数を実行する前に、第1のスレッドが第2のVMの仮想化特権レベルにあるか否かを決定し、第2のVMは、第2のプロセッサ上で実行する、ようにさらに構成される。リダイレクト装置は、割り込みスレッドが第2のVMの仮想化特権レベルにある場合、第2のVMの仮想化特権レベルをhostのユーザモードに切り替える、ように構成される。

【0039】

20

この可能な実施態様では、第2のプロセッサ内のリダイレクト装置は、ハードウェア回路を使用することによって実装されてもよい。リダイレクト装置の機能については、第1の態様のリダイレクト装置の機能を参照されたい。この可能な実施態様では、第2のVMの仮想化特権レベルは、第2のプロセッサ内のリダイレクト装置を使用することによって、hostのユーザモードに直接切り替えられ得る。したがって、非同期処理手順における切り替えオーバーヘッドが、さらに低減され、コンピュータデバイスの性能が、さらに改善される。

【0040】

第3の態様の可能な実施態様では、リダイレクト装置は、割り込み処理関数のエントリアドレスを第1のレジスタに書き込むようにさらに構成される。第2のプロセッサは、第1のレジスタから割り込み処理関数のエントリアドレスを読み出し、エントリアドレスに基づいて割り込み処理関数を呼び出し、hostのユーザモードにおいて、割り込み処理関数を実行するように具体的に構成される。

30

【0041】

手順は、第1の態様の対応する内容を使用することによって直ちに実施されてもよく、詳細は、ここでは再度説明されない。

【0042】

第3の態様の可能な実施態様では、第2のプロセッサは、第2のVMのコンテキストからシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理し、割り込みを処理した後、シャドウスレッドのコンテキストから第2のVMのコンテキストに戻るよう切り替える、ように具体的に構成される。リダイレクト装置は、hostのユーザモードから第2のVMの仮想化特権レベルに戻るよう切り替える、ようにさらに構成される。第2のプロセッサは、第2のVMの実行を継続するようさらに構成される。

40

【0043】

この可能な実施態様では、割り込み処理関数を実行することは、コンテキスト切り替えの手順と、割り込みを処理する手順と、を含む。このソリューションでは、リダイレクト装置は、hostのユーザモードから第2のVMの仮想化特権レベルに直接切り替え、これにより、割り込まれた第2のVMは、迅速に開始され得る。

【0044】

第3の態様の可能な実施態様では、第2のプロセッサは、第1のスレッドのコンテキスト

50

からシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理し、割り込みを処理した後、シャドウスレッドのコンテキストから第1のスレッドのコンテキストに切り替え、第1のスレッドの実行を継続する、ようにさらに構成される。

【0045】

この可能な実施態様では、仮想マシンが、第2のプロセッサ上で実行していない場合、第2のプロセッサは、第1のスレッドのコンテキストとシャドウスレッドのコンテキストとを直接切り替え、次に、割り込みを処理し、次に、第1のスレッドのコンテキストに戻るよう切り替え得る。このようにして、第1のスレッドは、迅速に開始され得、これにより、第1のスレッドは、実行を継続する。

【0046】

第3の態様の可能な実施態様では、hostの特権レベルは、カーネルモードであり、第2のプロセッサは、hostのカーネルモードのものであり、割り込みコントローラによって送信された第2の割り込み要求を受信し、hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替え、割り込み処理関数が、第2の割り込み要求に対応し、hostのカーネルモードにおいて、カーネルモードの割り込み処理関数を実行する、ように構成される。

【0047】

この可能な実施態様では、hostの特権レベルが、カーネルモードである場合、割り込みコントローラは、カーネルモード割り込みを第2のプロセッサに直接送信し得、第2のプロセッサは、対応する動作を完了するためにカーネルモード処理手順を実施する。第2のプロセッサが、hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替える前に、第2のVMが、第2のプロセッサ上で実行する場合、第2のプロセッサは、第2のVMの仮想化特権レベルからhostのカーネルモードに、最初に切り替える。仮想マシンが、第2のプロセッサ上で実行しておらず、hostが、ユーザモードにある場合、第2のプロセッサは、hostのユーザモードからhostのカーネルモードに、最初に切り替える。hostがカーネルモードにある場合、第2のプロセッサは、割り込み処理関数のエントリアドレスに直接切り替え、次に、割り込み処理関数を実行する。

【0048】

本出願の第4の態様は、コンピュータデバイスを提供し、コンピュータデバイスが、第1のプロセッサと、第2のプロセッサと、割り込みコントローラと、を含み、ホスト(host)が、コンピュータデバイス上で実行し、第2のプロセッサが、ハードウェア支援切り替え装置をさらに含む。第1のプロセッサは、割り込みコントローラ内のユーザ・モード・アドレス・マッピング空間またはカーネル・モード・アドレス・マッピング空間にアクセスすることによって、割り込み要求を割り込みコントローラに送信し得る。割り込みコントローラは、シャドウスレッドに関する情報を設定する。割り込みコントローラは、割り込み要求を第2のプロセッサに送信する。ハードウェア支援切り替え装置は、第1のスレッドからシャドウスレッドに切り替え、第1のスレッドが、切り替え前に第2のプロセッサによって実行されるスレッドであり、シャドウスレッドが、ユーザモードにおけるhost内で実行し、割り込み処理関数を含むスレッドであり、割り込み処理関数が、第2の割り込み要求に対応する、ように構成される。第2のプロセッサは、hostのユーザモードにおける割り込み処理関数を実行するように具体的に構成される。

【0049】

この可能な実施態様では、ハードウェア支援切り替え装置は、ハードウェア回路を使用することによって実装されてもよい。スレッド切り替えは、スレッドの切り替えに関する情報を含み得る。第1のスレッドに関する情報は、第1のスレッドのメモリアドレス空間と、ページテーブル開始アドレスと、データ構造情報と、処理特権レベルと、を含む。シャドウスレッドは、ユーザモードにおけるhost内で実行しているスレッドであり、割り込み処理関数を実行するために使用される。シャドウスレッドに関する情報は、シャドウスレッドのメモリアドレス空間、ページテーブル開始アドレス、データ構造情報、処理特権レベルなどを含む。第1のスレッドは、第2のプロセッサによって実行されているスレッドである。シャドウスレッドは、第1のスレッドのシャドウとして理解され得る。第2のプロセ

10

20

30

40

50

ッサによって実行しているスレッドが割り込まれた後、第2のプロセッサは、割り込み処理を実施するために、シャドウスレッドを実行する。この可能な実施態様では、第1のスレッドに関する情報、およびシャドウスレッドに関する情報は、hostのカーネルモードにトラップされることなく、ハードウェア支援切り替え装置を使用することによって迅速に切り替えられ得、これにより、割り込みは、hostのユーザモードにおいて直接処理され得る。したがって、切り替えオーバーヘッドが、さらに低減され、コンピュータデバイスの性能が改善される。

【0050】

第4の態様の可能な実施態様では、第2のプロセッサは、リダイレクト装置をさらに含み、第2のプロセッサは、hostのユーザモードにおける割り込み処理関数を実行する前に、第1のスレッドが、第2のVMの仮想化特権レベルにあるか否かを決定し、第2のVMが、第2のプロセッサ上で実行する、ようにさらに構成される。第1のスレッドが、第2のVMの仮想化特権レベルにある場合、リダイレクト装置は、VMの仮想化特権レベルをhostのユーザモードに切り替えるように構成される。

10

【0051】

この可能な実施態様では、第2のプロセッサ内のリダイレクト装置は、ハードウェア回路を使用することによって実装されてもよい。リダイレクト装置の機能については、第1の態様のリダイレクト装置の機能を参照されたい。この可能な実施態様では、第2のVMの仮想化特権レベルは、第2のプロセッサ内のリダイレクト装置を使用することによって、hostのユーザモードに直接切り替えられ得る。したがって、非同期処理手順における切り替えオーバーヘッドが、さらに低減され、コンピュータデバイスの性能が、さらに改善される。

20

【0052】

第4の態様の可能な実施態様では、リダイレクト装置は、割り込み処理関数のエントリアドレスを第1のレジスタに書き込むようにさらに構成される。第2のプロセッサは、第1のレジスタから割り込み処理関数のエントリアドレスを読み出し、エントリアドレスに基づいて割り込み処理関数を呼び出し、hostのユーザモードにおいて、割り込み処理関数を実行するように具体的に構成される。

【0053】

手順は、第1の態様の対応する内容を使用することによって直ちに実施されてもよく、詳細は、ここでは再度説明されない。

30

【0054】

第4の態様の可能な実施態様では、第2のプロセッサは、第2のVMのコンテキストからシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理し、割り込みを処理した後、シャドウスレッドのコンテキストから第2のVMのコンテキストに戻るよう切り替える、ように具体的に構成される。リダイレクト装置は、hostのユーザモードから第2のVMの仮想化特権レベルに戻るよう切り替える、ようにさらに構成される。第2のプロセッサは、第2のVMの実行を継続するようさらに構成される。

【0055】

この可能な実施態様では、割り込み処理関数を実行することは、コンテキスト切り替えの手順と、割り込みを処理する手順と、を含む。このソリューションでは、リダイレクト装置は、hostのユーザモードから第2のVMの仮想化特権レベルに直接切り替え、これにより、割り込まれた第2のVMは、迅速に開始され得る。

40

【0056】

第4の態様の可能な実施態様では、第2のプロセッサは、第1のスレッドのコンテキストからシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理し、割り込みを処理した後、シャドウスレッドのコンテキストから第1のスレッドのコンテキストに切り替え、第1のスレッドの実行を継続する、ようにさらに構成される。

【0057】

この可能な実施態様では、仮想マシンが、第2のプロセッサ上で実行していない場合、

50

第2のプロセッサは、第1のスレッドのコンテキストとシャドウスレッドのコンテキストとを直接切り替え、次に、割り込みを処理し、次に、第1のスレッドのコンテキストに戻るよう切り替え得る。このようにして、第1のスレッドは、迅速に開始され得、これにより、第1のスレッドは、実行を継続する。

【0058】

第4の態様の可能な実施態様では、hostの特権レベルは、カーネルモードであり、第2のプロセッサは、hostのカーネルモードのものであり、割り込みコントローラによって送信された第2の割り込み要求を受信し、hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替え、割り込み処理関数が、第2の割り込み要求に対応し、hostのカーネルモードにおいて、カーネルモードの割り込み処理関数を実行する、ように構成される。

10

【0059】

この可能な実施態様では、hostの特権レベルが、カーネルモードである場合、割り込みコントローラは、カーネルモード割り込みを第2のプロセッサに直接送信し得、第2のプロセッサは、対応する動作を完了するためにカーネルモード処理手順を実施する。第2のプロセッサが、hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替える前に、第2のVMが、第2のプロセッサ上で実行する場合、第2のプロセッサは、第2のVMの仮想化特権レベルからhostのカーネルモードに、最初に切り替える。仮想マシンが、第2のプロセッサ上で実行しておらず、hostが、ユーザモードにある場合、第2のプロセッサは、hostのユーザモードからhostのカーネルモードに、最初に切り替える。hostがカーネルモードにある場合、第2のプロセッサは、割り込み処理関数のエントリアドレスに直接切り替え、次に、割り込み処理関数を実行する。

20

【0060】

本出願の第5の態様は、例外処理方法を提供し、本方法は、コンピュータデバイス、ホスト(host)、およびコンピュータデバイス上で実行される仮想マシンVMに適用され、本方法は、VMによってトリガされる例外命令を取得することと、VMの仮想化特権レベルからhostのユーザモードに切り替えることと、hostのユーザモードにおける例外命令に対応する例外処理関数を実行することと、を含む。

【0061】

第5の態様の可能な実施態様では、本方法は、例外処理関数のエントリアドレスを第1のレジスタに書き込むことをさらに含む。第1のプロセッサは、第1のレジスタから例外処理関数のエントリアドレスを読み出し、エントリアドレスに基づいて例外処理関数を呼び出し、hostのユーザモードにおいて、例外命令に対応する例外処理関数を実行する、ように具体的に構成される。

30

【0062】

第5の態様の可能な実施態様では、例外命令は、VMによって入力/出力I/O動作を実施することによってトリガされる。

【0063】

第5の態様の可能な実施態様では、本方法は、VMのコンテキストからユーザ・モード・シャドウ実行環境のコンテキストに切り替えることと、hostのユーザモードにおける例外を処理することと、をさらに含み、例外処理関数が、hostのユーザモードにおいて実行する仮想プロセッサのスレッド内に含まれ、仮想プロセッサのスレッドに割り当てられた実行環境が、メインスレッド実行環境と、VMの実行環境と、シャドウ実行環境と、を含み、仮想プロセッサのスレッドに割り当てられた実行環境が、仮想プロセッサのスレッドを実行するために必要なリソースであり、シャドウ実行環境が、メインスレッド実行環境およびVMの実行環境から独立したユーザ・モード・メモリ空間を有し、シャドウ実行環境が、例外処理関数を実行するために使用され、例外処理関数が、例外命令をトリガする例外を処理するために使用される。

40

【0064】

第5の態様および任意の可能な実施態様で提供される例外処理方法については、第1の態様、および第1の態様の任意の可能な実施態様の関連説明を参照されたい。詳細は、ここ

50

では再度説明されない。

【 0 0 6 5 】

本出願の第6の態様は、割り込み処理方法を提供し、本方法は、コンピュータデバイス、ホスト (host)、およびコンピュータデバイス上で実行される第1のVMに適用され、本方法は、第1のVMによってトリガされる第1の割り込み要求を取得することによって、第1の割り込み要求が、割り込み処理関数の実行を要求するために使用される、取得することと、第1の割り込み要求に応答して、割り込み情報を決定することと、割り込み情報に基づいて第2の割り込み要求を生成することによって、第2の割り込み要求が、割り込み処理関数の実行を要求するために使用される、生成することと、第2の割り込み要求に応答して、割り込み処理関数を実行することによって、第2の割り込み要求が、第1の割り込み要求を使用することによって取得される、実行することと、を含む。

10

【 0 0 6 6 】

第6の態様の可能な実施態様では、第1の割り込み要求は、仮想イベント割り込み要求番号VEI IDを含み、VEI IDは、第1の割り込み要求をトリガする仮想イベントを識別し、本方法は、第1のVMのVM IDを取得することによって、割り込み情報が、第1のVMのVEI IDおよびVM IDを含み、取得することと、第1のVMのVEI IDおよびVM IDに基づいて、第1の割り込み要求を処理するhostのVEI物理番号および特権レベルを決定することによって、VEI物理番号が、host内の仮想イベントを識別する、決定することと、をさらに含み、第2の割り込み要求に応答して、割り込み処理関数を実行する前述のステップが、VEI物理番号を取得することと、hostの特権レベルに対応する処理方式に基づいて、受信した第2の割り込み要求に応答して、VEI物理番号に対応する割り込み処理関数を実行することと、を含む。

20

【 0 0 6 7 】

第6の態様の可能な実施態様では、第1の割り込み要求は、仮想イベント割り込み要求番号VEI IDを含み、VEI IDは、第1の割り込み要求をトリガする仮想イベントを識別し、割り込み情報は、VEI IDを含み、本方法は、VEI IDに基づいて、第1の割り込み要求を処理するhostの特権レベルを決定すること、をさらに含み、第2の割り込み要求に応答して、割り込み処理関数を実行する前述のステップが、VEI IDを取得することと、hostの特権レベルに対応する処理方式に基づいて、受信した第2の割り込み要求に応答して、VEI IDに対応する割り込み処理関数を実行することと、を含む。

30

【 0 0 6 8 】

第6の態様の可能な実施態様では、hostの特権レベルはユーザモードであり、第2の割り込み要求に応答して、割り込み処理関数を実行する前述のステップでは、第1のスレッドからシャドウスレッドに切り替えることによって、第1のスレッドが、切り替え前にホスト内で実行するスレッドであり、シャドウスレッドが、ユーザモードにおけるhost内で実行するスレッドであり、シャドウスレッドが、割り込み処理関数を実行するために使用される、切り替えることと、hostのユーザモードにおける割り込み処理関数を実行することと、を含む。

【 0 0 6 9 】

第6の態様の可能な実施態様では、本方法は、hostのユーザモードにおける割り込み処理関数を実行する前に、第1のスレッドが第2のVMの仮想化特権レベルにあるか否かを決定することによって、第2のVMが、コンピュータデバイス上で実行する、決定することと、割り込みスレッドが、第2のVMの仮想化特権レベルにある場合、第2のVMの仮想化特権レベルをhostのユーザモードに切り替えることと、をさらに含む。

40

【 0 0 7 0 】

第6の態様の可能な実施態様では、本方法は、割り込み処理関数のエントリアドレスを第1のレジスタに書き込むことと、第1のレジスタから割り込み処理関数のエントリアドレスを読み出し、エントリアドレスに基づいて割り込み処理関数を呼び出し、hostのユーザモードにおける割り込み処理関数を実行することと、をさらに含む。

【 0 0 7 1 】

50

第6の態様の可能な実施態様では、hostのユーザモードにおける割り込み処理関数を実行する前述のステップが、第2のVMのコンテキストからシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理することと、割り込みが処理された後、シャドウスレッドのコンテキストから第2のVMのコンテキストに切り替えることと、hostのユーザモードから第2のVMの仮想化特権レベルに戻るよう切り替えることと、第2のVMの実行を継続することと、を含む。

【0072】

第6の態様の可能な実施態様では、本方法は、第1のスレッドのコンテキストからシャドウスレッドのコンテキストに切り替え、hostのユーザモードにおける割り込みを処理することと、割り込みが処理された後、シャドウスレッドのコンテキストから第1のスレッドのコンテキストに切り替えることと、第1のスレッドの実行を継続することと、をさらに含む。

10

【0073】

第6の態様の可能な実施態様では、hostの特権レベルは、カーネルモードであり、hostの特権レベルに対応する処理方式に基づいて、受信した第2の割り込み要求に回答して、VEI IDに対応する割り込み処理関数を実行する前述のステップが、hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替えることとあって、割り込み処理関数が、第2の割り込み要求に対応する、切り替えることと、hostのカーネルモードにおいて、カーネルモードの割り込み処理関数を実行することと、を含む。

【0074】

20

第6の態様および任意の可能な実施態様で提供される例外処理方法については、第3の態様の関連説明、および第3の態様の任意の可能な実施態様を参照されたい。詳細は、ここでは再度説明されない。

【0075】

本出願の第7の態様は、コンピュータデバイスを提供する。コンピュータデバイスは、第5の態様、または第5の態様の可能な実施態様のいずれか1つにおける方法を実施する機能を有する。機能は、ハードウェアによって実装されてもよく、または対応するソフトウェアを実行するハードウェアによって実装されてもよい。ハードウェアまたはソフトウェアは、前述の機能に対応する1つまたは複数のモジュール、例えば、受信ユニットおよび処理ユニットを含む。

30

【0076】

本出願の第8の態様は、コンピュータデバイスを提供する。コンピュータデバイスは、第6の態様、または第6の態様の可能な実施態様のいずれか1つにおける方法を実施する機能を有する。機能は、ハードウェアによって実装されてもよく、または対応するソフトウェアを実行するハードウェアによって実装されてもよい。ハードウェアまたはソフトウェアは、前述の機能に対応する1つまたは複数のモジュール、例えば、受信ユニットおよび処理ユニットを含む。

【0077】

本出願の第9の態様は、1つまたは複数のコンピュータ実行可能命令を格納するコンピュータ可読記憶媒体を提供する。コンピュータ実行可能命令が、プロセッサによって実行されるとき、プロセッサは、第5の態様、または第5の態様の可能な実施態様のいずれか1つによる方法を実施する。

40

【0078】

本出願の第10の態様は、1つまたは複数のコンピュータ実行可能命令を格納するコンピュータ可読記憶媒体を提供する。コンピュータ実行可能命令が、プロセッサによって実行されるとき、プロセッサは、第6の態様、または第6の態様の可能な実施態様のいずれか1つによる方法を実施する。

【0079】

本出願の第11の態様は、1つまたは複数のコンピュータ実行可能命令を格納するコンピュータプログラム製品を提供する。コンピュータ実行可能命令が、プロセッサによって実

50

行されるとき、プロセッサは、第5の態様、または第5の態様の可能な実施態様のいずれか1つによる方法を実施する。

【0080】

本出願の第12の態様は、1つまたは複数のコンピュータ実行可能命令を格納するコンピュータプログラム製品を提供する。コンピュータ実行可能命令が、プロセッサによって実行されるとき、プロセッサは、第6の態様、または第6の態様の可能な実施態様のいずれか1つによる方法を実施する。

【0081】

本出願の第13の態様は、チップシステムを提供し、チップシステムが、少なくとも1つのプロセッサを含み、少なくとも1つのプロセッサが、第5の態様、または第5の態様の可能な実施態様のいずれか1つにおける機能を実施する際に、プロセス間通信装置をサポートするように構成される。1つの可能な設計では、チップシステムは、メモリをさらに含む。メモリは、メッセージ管理装置に必要なプログラム命令およびデータを格納するように構成される。チップシステムは、チップを含んでもよく、またはチップおよび別の個別構成要素を含んでもよい。

10

【0082】

本出願の第14の態様は、チップシステムを提供し、チップシステムが、少なくとも1つのプロセッサを含み、少なくとも1つのプロセッサが、第6の態様、または第6の態様の可能な実施態様のいずれか1つにおける機能を実施する際に、プロセス間通信装置をサポートするように構成される。1つの可能な設計では、チップシステムは、メモリをさらに含む。メモリは、スケジューリング装置に必要なプログラム命令およびデータを格納するように構成される。チップシステムは、チップを含んでもよく、またはチップおよび別の個別構成要素を含んでもよい。

20

【0083】

本出願の実施形態では、同期処理手順における特権レベル切り替えは、リダイレクト装置を使用することによって実施され、割り込み要求は、hostのカーネルモードを通過することなく、仮想イベント通知装置および割り込みコントローラを使用することによって非同期的に転送されてもよい。同期処理手順と非同同期処理手順との両方において、切り替えオーバーヘッドが低減され、コンピュータデバイスの性能が改善される。

【図面の簡単な説明】

30

【0084】

【図1】本出願の一実施形態による、コンピュータデバイスの構造の概略図である。

【図2】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【図3】本出願の一実施形態による、例外処理方法の一実施形態の概略図である。

【図4】本出願の一実施形態による、プロセッサの構造の概略図である。

【図5】本出願の一実施形態による、仮想プロセッサのスレッドの構造の概略図である。

【図6】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【図7A】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【図7B】本出願の一実施形態による、割り込み状態変更の概略図である。

【図8】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

40

【図9】本出願の一実施形態による、シャドウスレッドの構造の概略図である。

【図10A】本出願の一実施形態による、割り込み処理方法の一実施形態の概略図である。

【図10B】本出願の一実施形態による、割り込み処理方法の一実施形態の概略図である。

【図11】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【図12】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【図13】本出願の一実施形態による、コンピュータデバイスの構造の別の概略図である。

【発明を実施するための形態】

【0085】

以下で、添付図面を参照して本出願の実施形態を説明する。説明される実施形態は、本出願の実施形態のすべてではなく、単なる一部にすぎないことは明らかである。当業者は

50

、本出願の実施形態において提供される技術的ソリューションが、技術が進化し、新規のシナリオが出現するにつれて、同様の技術的問題にも適用されると分かる。

【0086】

本出願の明細書、特許請求の範囲、および添付の図面において、「第1」、「第2」などの用語は、類似する対象間を区別することを意図しているが、必ずしも特定の順序または順番を示すものではない。そのように呼ばれるデータは適切な状況においては交換可能であり、これにより、本明細書に説明される実施形態は、本明細書に図示または記載される順序以外の他の順序で実施され得ることを理解されたい。加えて、「include（含む）」、「have（有する）」、および任意の他の変形は、非排他的な包含を対象として含むことを目的とされている。例えば、ステップまたはユニットのリストを含む手順、方法、システム、製品、またはデバイスは、明示的に列挙されたステップまたはユニットに必ずしも限定されず、明示的に列挙されていない、またはそのような手順、方法、製品、もしくはデバイスに固有の他のステップまたはユニットを含み得る。

10

【0087】

本出願の実施形態は、切り替え中に生成したオーバーヘッドを低減し、コンピュータデバイスの性能を改善するために、コンピュータデバイス、例外処理方法、および割り込み処理方法を提供する。本出願は、対応する装置、コンピュータ可読記憶媒体、コンピュータプログラム製品などを、さらに提供する。詳細は、以下に別々に説明される。

【0088】

仮想化とは、計算デバイスのハードウェア層のハードウェアリソース（例えば、プロセッサまたはメモリ内の記憶空間、およびネットワークリソース）が、仮想化され、仮想化されたリソースが、複数の仮想コンピュータで共有されることを意味する。仮想コンピュータは、すべてのタイプの仮想化デバイスにおいてソフトウェアを使用することによって仮想化された環境を実行する総称であり、仮想コンピュータの概念は、仮想マシンまたはコンテナを含む。

20

【0089】

図1に示されるように、コンピュータデバイス100は、ハードウェア層112と、ホスト層109と、仮想化層と、を含む。仮想化層は、仮想マシン101および102を含む。仮想マシンの数は、これより多くても少なくてもよい。ここでは、一例として、2つの仮想マシンのみが使用される。ハードウェア層112は、プロセッサシステム114と、メモリ113と、通信インターフェース115と、割り込みコントローラ116と、を含む。

30

【0090】

仮想マシン（virtual machine、VM）は、仮想化ソフトウェアを使用することによってコンピュータデバイス上でシミュレートされた1つまたは複数の仮想コンピュータである。仮想マシンは、完全に隔離された環境で実行し、現実のコンピュータのように作動する。ゲスト・オペレーティング・システム（guest operating system、guest OS）（図1の105および106）が、仮想マシン（図1の101および102）上にインストールされ得、1つまたは複数のアプリケーション（図1の103および104）が、ゲスト・オペレーティング・システム上で実行される。仮想マシンはまた、ネットワークリソースにアクセスし得る。仮想マシン上で実行しているアプリケーションは、そのアプリケーションが、現実のコンピュータ上で作動しているかのように作動する。

40

【0091】

VMが実行しているとき、仮想マシン特権レベルが関与される。仮想マシン特権レベルは、仮想マシンのユーザモードと、仮想マシンのカーネルモードと、を含む。仮想マシンのカーネルモードおよびユーザモードは、VMを現実のコンピュータと見なすことによって分割される。

【0092】

ユーザモードおよびカーネルモードの概念が導入される前に、以下は、ユーザ空間およびカーネル空間を最初に導入する。

【0093】

50

いくつかの既存のコンピュータシステムでは、仮想アドレス空間全体が分割され、2つの仮想アドレス空間が、分割を通してカーネル空間（カーネルモード仮想アドレス空間）およびユーザ空間（ユーザモード仮想アドレス空間）として取得される。一例として、64ビットアーキテクチャ内の48ビット有効仮想アドレスが、使用される。この場合、ユーザ空間は、(00000000、00000000) - (0000FFFF、FFFFFFF)であり、カーネル空間は、(FFFF0000、00000000) - (FFFFFFF、FFFFFFF)である。カーネル空間およびユーザ空間は、対応するレジスタを有し、レジスタは、対応する空間の独立したページ・テーブル・ベース・アドレスを格納する。カーネル空間は、カーネル・モード・メモリ空間と呼ばれる場合もあり、ユーザ空間は、ユーザ・モード・メモリ空間と呼ばれる場合もある。ユーザ空間内で実行しているプロセスまたはスレッドは、ユーザ・モード・プロセスまたはスレッドと呼ばれる場合もあり、カーネル空間で実行しているプロセスまたはスレッドは、カーネル・モード・プロセスまたはスレッドと呼ばれる。

【0094】

仮想プロセッサ（例えば、図1の107および108）は、仮想化技術における共有またはスライス方式で使用するために、仮想コンピュータに提供される物理処理ユニット、例えば、仮想CPU（virtual central processing unit、vCPU）である。1つの仮想コンピュータは、1つまたは複数の仮想プロセッサによってサービスされ得る。複数の仮想プロセッサが存在するとき、1つの仮想プロセッサは、通常プライマリ仮想プロセッサであり、他は、セカンダリ仮想プロセッサである。仮想マシンに含まれる仮想メモリなどの他の仮想ハードウェアリソースは、図1に示されていない。仮想マシンは独立したコンピュータに相当することを理解されたい。したがって、仮想マシンによって実施される動作はまた、仮想プロセッサが動作を実施すると見なされてもよい。仮想プロセッサは、仮想化ソフトウェアを使用することによって仮想化される。仮想プロセッサの実行は、以下、つまり、ホストのプロセッサまたは物理コアがソフトウェアプログラムを読み出し、実行する、ように実際に実施される。例えば、物理コアは、ソフトウェアプログラムを読み出し、仮想プロセッサを実装するために、物理コアのハードウェア支援仮想化の特定のモード（例えば、x86のnon-root）でソフトウェアプログラムを実行する。仮想マシンの複数の仮想プロセッサは、異なる物理コアに配置されてもよい。

【0095】

仮想プロセッサのトラップイン（trap in）およびトラップアウト（trap out）、つまり、仮想化システムは、ホストモード（host mode）およびゲストモード（guest mode）の2つのモードを含む。ホストモードは、hostの特権レベル、例えば、hostのユーザモードまたはhostのカーネルモードと呼ばれる場合もある。ゲストモードは、VMの特権レベル、例えば、VMのユーザモード、またはVMのカーネルモードと呼ばれる場合もある。仮想プロセッサが、ゲストモードに入ると、この手順は、トラップイン（仮想）と呼ばれ、仮想プロセッサが、ゲストモードを出るとき、この手順はトラップアウト（仮想）と呼ばれる。仮想プロセッサが、トラップアウトした後、物理プロセッサは、仮想プロセッサのコードの実行を一時的に停止する。したがって、この場合、仮想プロセッサは、実行していないと理解され得る。物理プロセッサ上で実行する仮想プロセッサが、トラップインするとき、物理プロセッサは、ゲストモードにあると見なされ得、仮想プロセッサのコードが実行される。物理プロセッサ上で実行する仮想プロセッサが、ホストモードにトラップアウトするとき、物理プロセッサは、ホストモードにあり、仮想マシンモニタなど、ホスト関連コードを実行すると見なされ得る。

【0096】

ホスト（host）層109は、管理層として使用され、ハードウェアリソースを管理し、割り当て、仮想マシン用の仮想ハードウェアプラットフォームを提示し、仮想マシンをスケジューリングし、分離する、などのために使用される。いくつかの実施態様では、ホスト層109は、ホスト・オペレーティング・システム111と、仮想監視装置、例えば、仮想マシンモニタ110（virtual machine monitor、VMM）と、を含む。仮想モニタ110は、ホスト・オペレーティング・システム111の内部、またはホスト・オペレーティング・シ

10

20

30

40

50

ステム111の外部に展開され得る。別の仮想化アーキテクチャでは、仮想監視装置は、hypervisorまたは別のタイプの仮想監視装置と呼ばれる場合もある。いくつかの他の実施態様では、例えば、仮想化アーキテクチャXenでは、ホスト層109は、さらに、1つの特権化仮想マシンを含んでもよい。仮想ハードウェアプラットフォームは、仮想ハードウェアプラットフォーム上で実行する仮想コンピュータに、仮想プロセッサ、仮想メモリ、仮想ディスク、および仮想ネットワーク・インターフェース・カードなど、様々なハードウェアリソースを提供する。仮想コンピュータは、仮想コンピュータのためにホスト層によって用意された仮想ハードウェアプラットフォーム上で実行する。ホスト層109はまた、仮想化プラットフォームと呼ばれる場合もあり、ホスト層は、略してホストと呼ばれる場合もある。ホストの特権レベルは、ユーザモードと、カーネルモードと、を含む。

10

【0097】

ハードウェア層112は、仮想化環境が実行するハードウェアプラットフォームである。ハードウェア層は、複数のタイプのハードウェアを含み得る。図1に示されるように、ハードウェア層は、プロセッサシステム114と、メモリ113と、を含み得、通信インターフェース115、例えば、ネットワーク・インターフェース・カード(network interface card、NIC)をさらに含んでもよく、割り込みコントローラ116、入力/出力(input/output、I/O)デバイスなどをさらに含んでもよい。プロセッサシステム114は、1つまたは複数のプロセッサ、例えば、図1に示されるプロセッサ1およびプロセッサ2を含み得る。各プロセッサは、複数の物理コアを含み得、プロセッサは、複数のレジスタ、例えば、汎用レジスタ、および浮動小数点レジスタをさらに含んでもよい。

20

【0098】

プロセッサシステム114は、物理プロセッサと呼ばれることもある。物理コアは、プロセッサ内の最小処理ユニットを表し、各プロセッサに含まれるコアの数は、同じであっても、異なってもよい。1つのコアを有するプロセッサは、シングル・コア・プロセッサと呼ばれ、複数の物理コアを有するプロセッサは、マルチ・コア・プロセッサと呼ばれる。マルチ・コア・プロセッサは、コアアーキテクチャが同じであるか否かに基づいて、同種のマルチ・コア・プロセッサと、異種のマルチ・コア・プロセッサとに分類され得る。仮想プロセッサは、物理コアにバインドされ得る。言い換えると、仮想プロセッサは、特定の物理コア上で常に行われ、実行のために、別の物理コア上にスケジューリングされ得ない。この場合、仮想プロセッサは、バインドされたコアである。仮想プロセッサが、要求事項に基づいて、実行のために異なる物理コア上にスケジューリングされ得る場合、仮想プロセッサは、バインドされたコアではない。

30

【0099】

割り込みコントローラ116は、プロセッサと、割り込み要求をトリガするハードウェアとの間に配置され、ハードウェアによって生成された割り込み要求を収集し、特定の優先順位または別の規則に基づいて、割り込み要求をプロセッサに送信する、ように主に構成される。割り込みコントローラは、例えば、アドバンスドプログラマブル割り込みコントローラ(advanced programmable interrupt controller、APIC)である。

【0100】

割り込み(interruption)は、現在のプログラムの命令の実行が中断され、割り込みサービスルーチンが実行されることを示す。

40

【0101】

割り込みサービスルーチン(interrupt service routine、ISR)は、割り込み処理関数と呼ばれる場合もあり、割り込み要求を処理するために使用されるプログラムである。プロセッサは、割り込み要求を受信するとき、現在のプログラムの実行を中断し、割り込み要求に対応する割り込みサービスルーチンを実行する。

【0102】

割り込み要求(interrupt request)は、ハードウェアによって生成されるイベントである。ハードウェアは、イベントをプロセッサに送信する。プロセッサは、イベントを受信するとき、現在のプログラムの実行を中断し、イベントに対応するルーチンを実行する

50

。ハードウェアによって生成された割り込み要求は、ハードウェアによってトリガされてもよいし、ソフトウェアによってトリガされてもよい。割り込み要求は、割り込みと簡単に呼ばれることもある。コンピュータ内の一部のハードウェア（ネットワークアダプタ、オーディオアダプタ、マウス、ハードディスクなど）は、プロセッサの介入なしに、特定の作動を完了し得る。しかしながら、プロセッサがハードウェアのための何らかの特定の作動を実施することを可能にするために、ハードウェアは、割り込みコントローラを依然として、使用する必要がある。割り込み番号は、割り込み要求の識別子であり、本出願のIRQ IDによって表される。

【0103】

ハードウェア層112は、メモリ管理ユニット（memory management unit、MMU）をさらに含み得る。MMUは、メモリアクセス要求を処理するように構成されたコンピュータハードウェアである。MMUの機能は、仮想アドレスから物理アドレスへの変換（すなわち、仮想メモリ管理）と、メモリ保護と、中央処理装置キャッシュの制御と、を含む。MMUは、仮想ページ番号を物理ページ番号に変換するために、トランスレーション・ルックアサイド・バッファ（translation lookaside buffer、TLB）のアソシエイティブキャッシュ（associative cache）を通常使用する。

10

【0104】

メモリ113によって提供される記憶空間（アドレス空間）は、使用のために、仮想マシンおよびホストに割り当てられる。ホスト物理アドレス（host physical address、HPA）は、ローカルホスト（ホスト）によって使用され得る物理アドレス空間である。ホスト仮想アドレス（host virtual address、HVA）は、ローカルホスト（ホスト）によって使用され得る仮想アドレス空間である。ゲスト物理アドレス（guest physical address、GPA）は、仮想マシンのゲスト・オペレーティング・システムによって使用され得る物理アドレス空間である。ゲスト仮想アドレス（guest virtual address、GVA）は、仮想マシンのゲスト・オペレーティング・システムによって使用され得る仮想アドレス空間である。

20

【0105】

コンピュータデバイス100は、物理デバイス、例えば、サーバまたは端末デバイスであってもよい。端末デバイスは、無線接続機能を有するハンドヘルドデバイス、または無線モデムに接続される別の処理デバイスであってもよい。例えば、端末デバイスは、携帯電話、コンピュータ（personal computer、PC）、タブレットコンピュータ、携帯情報端末（personal digital assistant、PDA）、モバイル・インターネット・デバイス（mobile Internet device、MID）、ウェアラブルデバイス、および電子書籍リーダー（e-book reader）であってもよく、または、ポータブル、ポケットサイズ、ハンドヘルド、コンピュータ内蔵、または車載モバイルデバイスであってもよい。

30

【0106】

コンピュータデバイス100内のVMが、いくつかの動作（例えば、シリアルポートのメモリに対する読み出し/書き込み動作、ディスクのメモリに対する読み出し/書き込み動作、ネットワークアダプタのメモリに対する読み出し/書き込み動作、およびページフォルトに対する動作）を実施するとき、例外または割り込みが、トリガされる（例外は、同期例外として通常説明され、割り込みは、非同期例外として通常説明される）。ハードウェア層、またはホストのカーネルは、例外または割り込みを処理する。本出願の本実施形態では、例外は、同期方式で処理されてもよく、割り込みは、非同期方式で処理されてもよい。同期処理は、同じプロセッサ上の処理を指し、非同期処理は、クロスプロセッサ処理を指す。仮想I/Oシナリオでは、仮想I/O処理は、同期処理と非同期処理とに分類され得る。VMの実行に使用されるI/Oデバイスが、シリアルポート、グラフィックカード、キャラクタ、またはUSBなどのデバイスであるとき、同期処理方式が一般的に使用される。VMの実行に使用されるI/Oデバイスが、ディスクまたはネットワークアダプタなどのデバイスであるとき、非同期処理方式が一般的に使用される。

40

【0107】

第1のプロセッサおよび第2のプロセッサが、本出願の以下の実施形態では使用される。

50

第1のプロセッサおよび第2のプロセッサは、代替的に、同じマルチ・コア・プロセッサ内の2つの異なるコアであってもよいことに留意されたい。

【0108】

以下、同期処理手順と非同期処理手順について、それぞれ説明する。

【0109】

1. 同期処理

図2に示されるように、コンピュータデバイスは、第1のプロセッサと、コンピュータデバイス上で実行されるホスト (host) および第1のVMと、を含む。具体的には、第1のVMは、第1のプロセッサ上で実行する。第1のVMは、いくつかの動作を実施するとき、例外命令をトリガする。VMが、入力/出力 (input/output、I/O) 動作を実施するとき、例外命令は、トリガされ得る。VMが、仮想I/Oデバイスにアクセスする必要があるとき、I/O動作は、トリガされる動作であってもよい。仮想I/Oデバイスは、フロント・エンド・ドライバと、バック・エンド・デバイスと、を一般に含む。フロント・エンド・ドライバは、VM上で実行し、バック・エンド・デバイスのプログラムは、host上で実行する。VMによって実施される動作が、フロント・エンド・ドライバを使用することによってバック・エンド・デバイスにアクセスする必要がある場合、例外命令は、トリガされ、バック・エンド・デバイスは、例外命令を使用することによってアクセスされる。仮想I/Oデバイスは、バック・エンド・デバイスの1つのタイプである。仮想I/Oデバイスへのアクセスは、仮想I/Oデバイスのメモリからのデータの読み出し、または仮想I/Oデバイスのメモリへのデータの書き込みであってもよいし、仮想I/Oデバイスに対する別の制御動作であってもよい。仮想I/Oデバイスは、仮想マシンによって使用され得るI/Oデバイスである。仮想I/Oデバイスは、ソフトウェアを使用することによってシミュレートされた完全に仮想化されたI/Oデバイスであってもよい。仮想I/Oデバイスはまた、マッピング方式で、物理I/Oデバイスにマッピングされてもよい。仮想I/Oデバイスが、アクセスされるとき、物理I/Oデバイスは、最終的にアクセスされる。第1のプロセッサは、例外命令を受信し、第1のプロセッサが、例外命令がhostのユーザモードにおいて処理される必要があると決定した場合、第1のプロセッサは、第1のVMの仮想化特権レベルからhostのユーザモードに切り替え、hostのユーザモードにおける例外処理関数を実行する。例えば、仮想I/Oデバイスは、シリアルポート、グラフィックカード、キャラクタ、またはUSBを含む。シリアルポート、グラフィックカード、キャラクタ、またはUSBは、それぞれのアドレス範囲を有する。第1のプロセッサは、アドレス範囲に基づいて、現在のI/O動作を実施する仮想I/Oデバイスを決定し得る。本出願では、仮想I/Oデバイスに関連する動作は、同期的に実施される。仮想I/Oデバイスが、ソフトウェア、例えばlinuxシステムの仮想シリアルポートを使用することによってシミュレートされた仮想I/Oデバイスである場合、ソフトウェアを使用することによってシミュレートされた仮想I/Oデバイスは、物理I/Oデバイスに対応しない場合がある。仮想マシン内の仮想I/Oデバイスは、物理I/Oデバイスに関連付けられ得る。例えば、仮想マウスは、物理的なマウスに関連付けられる。仮想I/Oデバイスが、物理I/Oデバイスに関連付けられるとき、第1のプロセッサは、システムコールを使用することによってカーネルモードにトラップするために、ユーザモードのスレッドを実行し、次に、物理I/Oデバイスを動作させるために、物理I/Oデバイスのドライバを呼び出す。

【0110】

第1のVMによってトリガされる複数のタイプの例外命令、例えば、無効命令例外、ページフォルト例外、メモリアクセス例外、アドレスアライメント例外、および仮想I/Oデバイスへのアクセスによって引き起こされる例外があり得る。例外命令のタイプに基づいて、例外命令が、hostのユーザモード、またはhostのカーネルモードにおいて処理される必要があると、第1のプロセッサは決定し得る。前述のリダイレクト動作は、hostのユーザモードにおいて処理される必要がある例外命令に対して実施されてもよい。

【0111】

第1のプロセッサは、第1のリダイレクト装置を含み得る。第1のリダイレクト装置は、

10

20

30

40

50

第1のVMの仮想化特権レベルからhostのユーザモードの特権レベルに切り替え、かつ例外処理関数のエントリアドレスを第1のレジスタに書き込む手順を実施し得る。このようにして、第1のプロセッサは、第1のレジスタから例外処理関数のエントリアドレスを読み出し、エントリアドレスに基づいて例外処理関数を呼び出し、hostのユーザモードにおいて、例外命令に対応する例外処理関数を実行し得る。第1のリダイレクト装置の機能は、ハードウェア回路を使用することによって実施され得る。

【0112】

ユーザモードにおけるホストは、仮想プロセッサ(virtual central processing unit、vCPU)のスレッドを含み得、スレッドは、シャドウ実行環境を含み得る。第1のプロセッサは、シャドウ実行環境内で例外処理関数を実行し得る。

10

【0113】

図3を参照して、以下では、第1のプロセッサ、第1のリダイレクト装置、およびシャドウ実行環境を使用することによって、同期処理のための前述の例外処理方法を実施する手順を説明する。

【0114】

201：第1のプロセッサは、第1のVMによってトリガされる例外命令を取得する。

【0115】

ステップ201の手順は、第1のVMが第1のVMの許可を超える動作を実施するときに、例外が発生し、第1のプロセッサが、ハードウェアロジックを通して例外コードを検出し得、さらに例外命令を取得し得る、ものであってよい。

20

【0116】

例外命令を取得した後、第1のプロセッサは、後続の例外処理のために、例外の原因およびタイプ、第1のVMによってアクセスされるメモリ、ならびにリターンアドレスを、ハードウェアレジスタに書き込み得る。

【0117】

任意選択で、202：第1のプロセッサは、例外命令がhostのユーザモードにおいて処理される必要がある、と決定する。

【0118】

第1のプロセッサは、例外命令に対応する仮想I/Oデバイスのアドレス範囲に基づいて、例外命令がhostのユーザモードにおいて処理される必要がある、と決定し得る。

30

【0119】

例外命令が、hostのカーネルモードにおいて処理される必要がある場合、第1のリダイレクト装置は、第1のVMの仮想化特権レベルからhostのカーネルモードに切り替え、次に、処理のために、hostのカーネルモードの例外処理関数のエントリアドレスに切り替える。

【0120】

ステップ202は、任意選択のステップであり、すなわち、ステップ202は実施されなくてもよい。ステップ202が実施されない場合、ステップ203が、直接実施されてもよい。

【0121】

203：第1のリダイレクト装置は、第1のVMの仮想化特権レベルからhostのユーザモードに切り替える。

40

【0122】

VMの仮想化特権レベルは、仮想化モード、すなわち、VMの作動モードにおける特権レベルを表す。仮想化特権レベルは、仮想化ユーザモードと、仮想化カーネルモードと、を含む。VMの仮想化特権レベルからhostのユーザモードへの切り替え(直接切り替え)は、VMの仮想化特権レベルに関する情報を、特権レベルを格納するために使用されるレジスタから移動させ、次に、hostのユーザモードに関する情報をレジスタに書き込む手順を含んでもよく、レジスタは、システムレジスタであってもよい。本出願では、特権レベルおよび仮想化は、1ビットを使用することによって、それぞれ表され得る。例えば、0は、ユーザモードを示し、1は、カーネルモードを示し、0は、非仮想化を示し、1は、仮想化を示す。特権レベルは、第1のビットにあり、仮想化するか否かは、第2のビットにある。

50

例えば、00は、非仮想化ユーザモードを示し、ホストのユーザモードと呼ばれる場合もあり、10は、非仮想化カーネルモードを示し、ホストのカーネルモードと呼ばれる場合もあり、01は、仮想化ユーザモードを示し、VMのユーザモードと呼ばれる場合もあり、11は、仮想化カーネルモードを示し、VMのカーネルモードと呼ばれる場合もある。もちろん、前述は、単なるいくつかの例である。一連の特権レベルおよび仮想化、ならびに他の可能な表現形式は、本出願では限定されない。

【0123】

204：第1のリダイレクト装置は、例外処理関数のエントリアドレスを第1のレジスタに書き込む。

【0124】

205：第1のプロセッサは、シャドウ実行環境において、第1のVMのコンテキストからユーザ・モード・シャドウ実行環境のコンテキストに切り替え、例外処理を実施する。

【0125】

シャドウ実行環境は、ユーザモードにおける例外処理関数のキャリアであり、例外処理関数は、例外割り込みを処理するためのサービスルーチンと呼ばれる場合もある。

【0126】

例外処理関数は、hostのユーザモードにおいて実行する仮想プロセッサのスレッドに含まれ、仮想プロセッサのスレッドに割り当てられた実行環境は、メインスレッド実行環境、VMの実行環境、シャドウ実行環境を含み、仮想プロセッサのスレッドに割り当てられた実行環境は、仮想プロセッサのスレッドを実行するために必要なリソースであり、シャドウ実行環境は、メインスレッド実行環境およびVMの実行環境から独立したユーザ・モード・メモリ空間を有する。シャドウ実行環境は、例外処理関数を実行するために使用され、例外処理関数は、例外命令をトリガする例外を処理するために使用される。

【0127】

任意選択で、206：例外を処理する手順において、カーネルで処理される必要がある割り込みまたは例外が、生成されたか否かを決定し、カーネルで処理される必要がある割り込みまたは例外が、生成された場合、207を実施し、カーネルで処理される必要がある割り込みまたは例外が、生成されていない場合、209を実施する。

【0128】

207：カーネルで処理される必要がある割り込みまたは例外が、生成された場合、ユーザ・モード・シャドウ実行環境のコンテキストをカーネル・モード・シャドウ実行環境のコンテキストに切り替え、カーネルモードにおける割り込みまたは例外を処理する。

【0129】

208：割り込みまたは例外が、カーネルによって処理された後、カーネル・モード・シャドウ実行環境のコンテキストは、ユーザ・モード・シャドウ実行環境のコンテキストに切り替えられる必要がある。

【0130】

209：例外の処理を継続する。

【0131】

210：割り込みまたは例外が、処理された後、ユーザ・モード・シャドウ実行環境のコンテキストを第1のVMのコンテキストに切り替える。

【0132】

211：第1のリダイレクト装置は、hostのユーザモードから第1のVMの仮想化特権レベルに戻るよう切り替える。

【0133】

切り替え手順では、hostのユーザモードに関する情報が、特権レジスタから移動され、特権レジスタまたはレジスタに対応するメモリに元々格納されている第1のVMの仮想化特権レベルに関する情報が、特権レジスタに書き込まれる。

【0134】

212：第1のリダイレクト装置は、実行手順を第1のVM内の割り込まれた命令に切り替

10

20

30

40

50

え、第1のVMは、割り込まれた命令の実行を継続する。

【0135】

前述では、同期処理手順を説明した。同期処理手順では、図4に示されるように、第1のリダイレクト装置は、特権レベル選択モジュールと、特権レベル・リダイレクト・モジュールと、を含み得る。特権レベル・リダイレクト・モジュールは、更新モジュールと、復元モジュールと、をさらに含み得る。

【0136】

特権レベル選択モジュールは、第1のVMの実行中にトリガされた例外命令が、処理のために、hostのユーザモードにリダイレクトする必要があるか否かを決定するように構成され得る。

【0137】

例外命令がhostのユーザモードにリダイレクトされる必要がある場合、特権レベル選択モジュールは、特権レベル・リダイレクト・モジュールに、リダイレクト動作を実施するように通知する。リダイレクト手順については、ステップ204の説明を参照されたい。

【0138】

更新モジュールは、

(1) 切り替え情報を記録し、切り替え情報が、切り替え前の第1のVMの仮想化特権レベルに関する情報と、切り替え後のhostのユーザモードに関する情報と、を含み、

(2) 状態制御情報を更新し、すなわち、ユーザモードの例外処理関数の関連情報に対する状態レジスタの更新情報を更新し、関連情報が、例外処理関数のエントリアドレス、例外の原因、例外のメモリアドレスまたは命令などを含むが、これらに限定されず、

(3) 特権レベルを切り替え、すなわち、第1のVMの仮想化特権レベルをhostのユーザモードに切り替え、hostのユーザモードの例外処理関数のエントリアドレスに切り替える、

ように構成される。

【0139】

復元モジュールは、第1のVMの仮想化特権レベルに関する記録された情報に基づいて、hostのユーザモードから第1のVMの仮想化特権レベルに戻るよう切り替え、戻りアドレスに基づいて、第1のVMによって割り込まれた命令に切り替え、これにより、第1のVMが、命令の実行を継続する、ように構成される。

【0140】

シャドウ実行環境は、ユーザ・モード・シャドウ実行環境と、カーネル・モード・シャドウ実行環境と、を含む。具体的には、シャドウ実行環境は、ユーザモードとカーネルモードとの両方で、コンテキスト情報を格納し得る独立した記憶空間を有する。記憶空間はスタックであってもよく、ユーザモードにおける記憶空間は、ユーザモードにおけるスタックであってもよく、カーネルモードにおける記憶空間は、カーネルモードにおけるスタックであってもよい。

【0141】

シャドウ実行環境は、仮想プロセッサのスレッドの実行環境内に配置される。仮想プロセッサのスレッドのタスク構造については、図5を参照されたい。

【0142】

図5に示されるように、仮想プロセッサのスレッドのタスクは、仮想プロセッサをシミュレートするタスクと、VMを実行するタスクと、例外/割り込み処理のタスクと、を含み得る。

【0143】

vCPUのユーザモードのメインスレッド実行環境、およびvCPUのカーネルモードのメインスレッド実行環境が、仮想プロセッサをシミュレートするタスクにおいて使用される。vCPUのユーザモードのメインスレッド実行環境は、第1の記憶装置と、第2の記憶装置と、を含む。第1の記憶装置は、ユーザ・モード・コンテキストまたはカーネル・モード・コンテキストを格納するように構成され、第2の記憶装置は、hostコンテキストまたはVMコ

10

20

30

40

50

ンテキストを格納するように構成される。

【0144】

VM実行環境は、VMを実行するタスクにおいて使用され、VM実行環境は、VMを実行するために使用される。

【0145】

シャドウ実行環境が、割り込み処理のタスクにおいて使用される。シャドウ実行環境は、ユーザ・モード・シャドウ実行環境と、カーネル・モード・シャドウ実行環境と、を含む。ユーザ・モード・シャドウ実行環境は、第1の切り替え装置を含み、カーネル・モード・シャドウ実行環境は、第2の切り替え装置を含む。

【0146】

第1の切り替え装置は、前述のステップ205および210で説明されたように、VMのコンテキストからhostのユーザモードのコンテキストに切り替えるように構成される。第1の切り替え装置は、VMの実行状態を格納し、シャドウ実行環境のユーザモードの実行状態を復元する。実行状態に関する情報は、汎用レジスタ、浮動小数点レジスタ、およびシステムレジスタなどのレジスタ内の情報を含むが、これらに限定されない。

【0147】

第2の切り替え装置は、前述のステップ207および208で説明されたように、hostのユーザモードのコンテキストからhostのカーネルモードのコンテキストに切り替えるように構成される。第2の切り替え装置は、hostのユーザモードのコンテキストを格納し、カーネルモードのコンテキストを復元し、カーネルモードに入り、シャドウ実行環境のユーザモード中に発生し、カーネルモードにおいて処理される必要がある割り込みまたは例外を処理する、ように構成される。割り込みまたは例外が、カーネルモードにおいて処理された後、装置は、ユーザモードのコンテキストを復元し、例外命令の実行を継続するために、ユーザモードのシャドウ実行環境に戻るよう切り替える。カーネルモードのシャドウ実行環境のコンテキストと、仮想プロセッサのスレッドのカーネルモードのコンテキストとは、互いに競合しない、または互いに影響しない。

【0148】

ユーザモードおよびカーネルモードのコンテキストを格納するための装置、hostおよびVMのコンテキストを格納するための装置、VMのコンテキストとユーザモードのシャドウ実行環境との間を切り替えるための装置、およびユーザモードとカーネルモードのコンテキスト間を切り替えるための装置は、汎用レジスタ、浮動小数点レジスタ、またはシステムレジスタなどであってもよい。

【0149】

本出願における「実行環境」は、ソフトウェア実行環境と、ハードウェア実行環境との2つの態様を含む。ソフトウェア実行環境は、仮想プロセッサのスレッドが異なる実行環境で実行するとき使用されるスタック内の情報と、スレッドのコンテキスト情報と、を含む。ハードウェア態様は、メモリリソースまたは処理リソースを含む、スレッドが実行するとき使用されるハードウェアリソースを含む。

【0150】

本出願の本実施形態で提供されるコンピュータシステムによれば、VMの実行中に発生する例外は、同期処理方式で、処理のためにhostのユーザモードにリダイレクトされる。したがって、hostのカーネルモードとhostのユーザモードとの間の切り替えのオーバーヘッドが必要とされず、コンピュータシステムの性能が改善される。

【0151】

2. 非同期処理

非同期処理は、クロスプロセッサ処理手順を指す。第1のプロセッサ上で実行している第1のVMは、第1のVMのバック・エンド・デバイスにアクセスし、かつ割り込み要求をトリガする動作を、実施する必要がある。割り込み要求は、第2のプロセッサを使用することによって処理される必要がある。したがって、非同期処理手順は、割り込み要求送信終了手順と、割り込み要求受信終了手順とに分けられてもよい。送信終了手順は、hostのカ

10

20

30

40

50

ーネルを含まなくてもよいし、hostのカーネルを含んでもよい。受信終了手順は、hostのカーネルを含まなくてもよいし、hostのカーネルを含んでもよい。以下の4つの非同期処理方式があり得る。

【0152】

非同期処理方式1：送信終了手順は、hostのカーネルを含まず、受信終了手順も、hostのカーネルを含まない。

【0153】

図6に示されるように、コンピュータシステムは、第1のプロセッサと、第2のプロセッサと、割り込みコントローラと、物理I/Oデバイスと、を含む。ホスト(host)が、コンピュータデバイス上で実行し、第1の仮想マシンVMが、第1のプロセッサ上で実行する。コンピュータシステムは、仮想イベント割り込み構成要素をさらに含む。仮想イベント割り込み構成要素は、仮想イベント通知装置と、仮想イベント割り込み装置と、を含み得る。仮想イベント通知装置の回路構造は、第1のプロセッサに配置され、仮想イベント割り込み装置の回路構造は、割り込みコントローラ内に配置される。

【0154】

第1のプロセッサは、第1のVMによってトリガされる第1の割り込み要求を受信し、第1の割り込み要求が、割り込み処理関数の実行を要求するために使用される、ように構成される。仮想イベント通知装置は、第1の割り込み要求に応答して、割り込み情報を割り込みコントローラに直接送信するように構成される。割り込みコントローラは、割り込み情報に基づいて第2のプロセッサの識別子を決定し、第2のプロセッサの識別子に基づいて第2の割り込み要求を第2のプロセッサに送信するように構成される。第2のプロセッサは、受信した第2の割り込み要求に応答して、割り込み処理関数を実行するように構成される。

【0155】

仮想イベント割り込み構成要素の構造は、図7Aを参照して理解され得る。図7Aに示されるように、仮想イベント割り込み構成要素は、仮想イベント生成モジュールと、仮想イベント監視モジュールと、仮想イベント割り込みマッピングテーブルと、仮想イベント割り込み照会モジュールと、仮想イベント割り込み情報設定モジュールと、を含む。仮想イベント生成モジュールおよび仮想イベント監視モジュールは、第1のプロセッサに配置され、仮想イベント割り込みマッピングテーブル、仮想イベント割り込み照会モジュール、および仮想イベント割り込み情報設定モジュールは、割り込みコントローラに配置される。

【0156】

仮想イベント生成モジュールは、第1のVMによって直接アクセスされ得るVEI - REGレジスタを含み得る。このようにして、第1のVMが仮想イベント生成モジュールにアクセスするとき、VMは、仮想イベント割り込み要求番号(virtual event interrupt identifier、VEI ID)をVEI - REGレジスタに書き込み得る。第1のVMが仮想イベント生成モジュールにアクセスする方式は、専用レジスタの読み出しおよび書き込み、専用命令の実行、専用ハードウェアロジックをトリガするための既存の命令の実行などを含むが、これらに限定されない。

【0157】

VEI IDは、第1のVMに第1の割り込み要求をトリガさせる仮想イベントを識別する。一般に、VMによってトリガされる割り込みのタイプは、ローカル割り込み、外部割り込み、クロック割り込みなどをさらに含み得る。割り込み要求が仮想イベントによってトリガされる割り込みであることを、VEI IDが示す場合、仮想イベント生成モジュールは、仮想イベント監視モジュールに、作動を開始するように通知する。仮想イベント監視モジュールは、第1のVMの番号、すなわち、仮想マシン識別子(virtual machine identifier、VM ID)を照会する。仮想イベント監視モジュールは、仮想マシン識別子を格納するためのレジスタからVM IDを取得し得る。第1のVMが第1のプロセッサ上で実行するとき、VM IDは、仮想マシン識別子を格納するためのレジスタに書き込まれる。仮想マシン識別子を格納するためのレジスタは、コントロール状態レジスタ(control status registers、CSR)であり得る。割り込み情報は、VEI IDと、VM IDと、を含む。

10

20

30

40

50

【 0 1 5 8 】

仮想イベント監視モジュールは、第1のVMのVEI IDおよび仮想マシン識別子 (virtual machine identifier、VM ID) を仮想イベント割り込み照会モジュールに送信する。

【 0 1 5 9 】

本出願の本実施形態では、第1のVMがhostにトラップアウト (trap out) することを防止するために、VMによって直接アクセスされ得るVEI - REGレジスタが、新規に追加される。VMは、VEI - REGレジスタにVEI IDを書き込み得、これにより、第1のプロセッサは、VEI - REGレジスタを使用することによってVEI IDを読み出し得、次に、第1のプロセッサは、CSRレジスタからVM IDを読み出し、割り込みコントローラおよび第2のプロセッサを使用することによって後続の割り込み処理を、さらに非同期的に実施する。第1のVMが、VEI - REGレジスタにVEI IDを書き込んだ後、後続の割り込み手順は、非同期的に実施されるので、第1のVMは、終了する必要がなく、実行を継続し得る。

10

【 0 1 6 0 】

仮想イベント割り込みマッピングテーブルは、VEI ID、VM ID、プロセッサID (CPU ID)、およびVEI物理番号を保持し、VEI IDと、VM IDに対応する割り込み要求に対応するhostの特権レベル (割り込み要求がユーザモードまたはカーネルモードにおいて処理されるべきである) とのマッピング関係をさらに保持する。VEI IDは、仮想マシンの仮想イベント割り込みを示し、VEI物理番号は、仮想イベント割り込みを識別するために、プロセッサによって使用される。VEI IDは、0から63までの範囲であり、VEI物理番号は、0から2047までの範囲である。本出願の本実施形態では、VEI IDは、仮想マシン内で発生する仮想イベントを識別し、VEI物理番号は、host内の仮想イベントの識別子を示す。VEI IDとVEI物理番号との間にはマッピング関係がある。互いにマッピングされたVEI IDとVEI物理番号は、同じ仮想イベント (例えば、ディスク読み出し/書き込みイベント、またはネットワークアダプタ読み出し/書き込みイベント) を示す。VEI IDは、例えば0から63の値によって表されてもよい。各値は、異なる仮想イベントを識別する。1つのI/Oデバイスは、複数の異なる仮想イベントを有し得る。例えば、ディスク読み出しイベントは、値1によって表されてもよく、ディスク書き込みイベントは、値2によって表されてもよく、ディスク制御イベントは、値3によって表されてもよく、ネットワークアダプタ読み出しイベントは、値4によって表されてもよく、ネットワークアダプタ書き込みイベントは、値5によって表されてもよい、などである。各仮想マシン内の同じ仮想イベントは、同じVEI IDを使用し得る。VEI物理番号は、グローバルに一意であってもよく、異なる仮想マシン上で発生する同じVEI IDの仮想イベントは、異なるVEI物理番号に対応してもよい。例えば、VM1のVEI IDの値の範囲に対応するVEI物理番号は、0から63であってもよく、VM2のVEI IDの値の範囲に対応するVEI物理番号は、64から127であってもよく、残りは、類推によって推論されてもよい。このようにして、第2のプロセッサは、異なるVEI物理番号の範囲に基づいて、仮想イベントが発生する仮想マシンを決定し、次に、VEI物理番号の特定の値に基づいて、VM上で発生する仮想イベントを決定し得る。例えば、VEI IDが5であることは、ディスク読み出しイベントを示し、VEI物理番号が5である場合、それは、VM1上でディスク読み出しイベントが発生したことを示す。VEI物理番号が69である場合、それは、VM2上でディスク読み出しイベントが発生したことを示す。仮想イベントとVEI IDとの間の対応関係、およびVEI物理番号と仮想マシンとの間の対応関係は、要求事項に基づいて設定されてもよい。これは、本出願において限定されない。

20

30

40

【 0 1 6 1 】

仮想イベント割り込みマッピングテーブルは、仮想イベント割り込みに対応する割り込み処理関数の情報 (サービスルーチン情報と呼ばれる場合もある) をさらに保持し、情報は、サービスルーチンのメモリアドレス空間、データ構造情報、処理特権レベルなどが含まれるが、これらに限定されない。

【 0 1 6 2 】

仮想イベント割り込み照会モジュールは、仮想イベント割り込みマッピングテーブルを検索することによって、VEI IDが登録されているか、または存在するかを確認する。VEI

50

IDが登録された、または存在する場合、仮想イベント割り込み照会モジュールは、VIE IDおよびVM IDを仮想イベント割り込み情報設定モジュールに転送する。VEI IDが登録されていない、または存在しない場合、割り込み要求は、無視されるか、または例外が投入される。

【0163】

仮想イベント割り込み情報設定モジュールは、VEI IDおよびVM IDに基づいて、仮想イベント割り込みマッピングテーブルに照会し、現在の割り込み要求に対応する第2のプロセッサのCPU ID、VEI物理番号、および対応するhostの特権レベルを決定する。

【0164】

本出願では、各仮想マシン内の同じ仮想イベントが、同じVEI IDを使用し得るので、仮想イベントが発生する仮想マシン間を区別するために、第1のプロセッサは、VM IDを取得し得、VM IDが、仮想マシン識別子を格納するためのレジスタから第1のプロセッサによって読み出され得る。次に、仮想イベント通知装置は、VEI IDおよびVM IDを割り込みコントローラに送信し、割り込みコントローラは、VEI IDおよびVM IDに基づいて、第2のプロセッサの識別子、VEI物理番号、および対応するhostの特権レベルを検索し得る。入力が、(VEI ID) + (VM ID)であり、出力が、(CPU ID) + (VEI物理番号) + (hostの特権レベル)であるマッピングテーブルを、割り込みコントローラは保持し得る。割り込みコントローラは、マッピングテーブルに基づいて、対応する(CPU ID) + (VEI物理番号) + (hostの特権レベル)を直接決定し得る。割り込みコントローラは、CPU IDに基づいて、非同期割り込み処理を続いて実施する第2のプロセッサを決定し得、hostの特権レベルに基づいて、続いてユーザモード処理方式を使用するか、またはカーネルモード処理方式を使用するか、を決定し得る。第2のプロセッサは、VEI物理番号を格納するレジスタからVEI物理番号を読み出し、次に、ユーザモード処理手順またはカーネルモード処理手順に基づいて、対応する処理を実施し得る。

【0165】

VM IDおよびVEI物理番号は、前述の方式において使用される。VEI IDがグローバルに一意である場合、異なるVMは、同じ仮想イベントに対して異なるVEI IDを使用する。このようにして、VM IDは、使用される必要がなく、VEI物理番号は、変換される必要がなく、対応する第2のプロセッサの識別子、およびhostの特権レベルがVEI IDに基づいて決定されれば、後続の割り込み処理が、実施され得る。本実施形態では、割り込み情報は、VEI IDを含み、VM IDは、もはや必要とされない。

【0166】

割り込み要求が、hostのユーザモードにおいて処理される必要がある場合、仮想イベント割り込み情報設定モジュールは、ユーザモードの割り込み処理スレッドに関する情報を第2プロセッサに設定する。割り込み要求が、hostのカーネルモードにおいて処理される必要がある場合、仮想イベント割り込み情報設定モジュールは、第2のプロセッサに、カーネル割り込みを処理するように通知する。

【0167】

本出願の本実施形態では、割り込み要求は、hostのユーザモードにおいて処理される必要がある。したがって、仮想イベント割り込み情報設定モジュールが、ユーザモードにおける割り込み処理スレッドに関する情報を設定した後、仮想イベント割り込み情報設定モジュールは、第2のプロセッサに、ユーザモードにおける割り込み要求を処理するように通知する。

【0168】

第2のプロセッサは、ユーザモードにおける割り込み要求を受信した後、対応する割り込み処理を実施する。

【0169】

本出願の本実施形態では、割り込みコントローラは、2つのレジスタを使用することによって、仮想イベント割り込みVEIを示す状態機械をさらに保持する。各VEI IDは、状態機械に対応する。図7Bに示される、割り込みを処理する異なる段階の状態機械には、非ア

10

20

30

40

50

クティブ (inactive) 状態、保留 (pending) 状態、アクティブ (active) 状態、および (アクティブ + 保留) (pending and active) 状態の4つの状態がある。アクティブ状態はまた、処理状態として理解されてもよく、(アクティブ + 保留) 状態は、現在の割り込みが、依然として処理されており、新規の割り込みが来ることを示す。

【0170】

2つのレジスタのうち一方のレジスタ (例えば、レジスタ1) は、状態機械がアクティブ状態にあるか否かを示す値を格納するように構成され、他方のレジスタ (例えば、レジスタ2) は、状態機械が保留状態にあるか否かを示す値を格納するように構成される。4つの状態は、2ビットを使用することによって表されることも理解されよう。例えば、第1のビットは、状態機械が保留状態にあるか否かを示し、1は、状態機械が保留状態にあることを示してもよい。第2のビットは、状態機械がアクティブ状態にあるか否かを示し、0は、状態機械が非アクティブ状態にあることを示し、1は、状態機械がアクティブ状態にあることを示してもよい。このようにして、非アクティブ状態は、00によって表されてもよく、保留状態は、10によって表されてもよく、アクティブ状態は、01によって表されてもよく、(アクティブ + 保留) 状態は、11によって表されてもよい。

10

【0171】

割り込みコントローラが、最初の割り込み要求を取得しないとき、レジスタ1およびレジスタ2における状態値は、いずれも0であり、すなわち非アクティブ状態の状態値は00である。

【0172】

S1: 割り込みコントローラが、割り込み要求1を取得した場合、割り込みコントローラは、保留状態に切り替える必要があり、レジスタ1内の状態値を1に設定する。

20

【0173】

このようにして、レジスタ1の状態値は1であり、レジスタ2の状態値は0であり、すなわち保留状態の状態値は10である。

【0174】

S2: 割り込みコントローラは、1割り込み要求1の処理を開始し、アクティブ状態に切り替える必要がある。この場合、レジスタ1内の状態値は、クリアにされ、すなわち0に復元され、レジスタ2内の状態値は、1に設定される。

【0175】

このようにして、レジスタ1内の状態値は0であり、レジスタ2内の状態値は1であり、すなわちアクティブ状態の状態値は01である。

30

【0176】

S3: 割り込みコントローラは、割り込み要求1を処理するとき、別の割り込み要求2を受信した場合、割り込みコントローラは、(アクティブ + 保留) 状態に切り替える必要があり、レジスタ1内の状態値を1に設定する。

【0177】

このようにして、レジスタ1内の状態値は1であり、レジスタ2内の状態値は1であり、すなわち、保留状態の状態値は11である。

【0178】

S4: 割り込み要求1を処理した後、割り込みコントローラが、非アクティブ状態に切り替える必要がある場合、レジスタ1内の状態値、およびレジスタ2内の状態値は、クリアにされる必要がある。

40

【0179】

レジスタ1およびレジスタ2内の状態値は、両方とも0であり、すなわち、非アクティブ状態の状態値は00である。

【0180】

S5: 割り込みコントローラが、(アクティブ + 保留) 状態から保留状態に変化した場合、レジスタ2内の状態値は、クリアにされる必要がある。

【0181】

50

このようにして、レジスタ1内の状態値は1であり、レジスタ2内の状態値は0であり、すなわち保留状態の状態値10が返される。

【0182】

本出願では、割り込みコントローラでは、1つの状態機械が、2つのレジスタを使用することによって保持され、これにより、割り込みコントローラは、割り込みを処理する状態を継続的に取得し得る。これは、1つの割り込みの処理が完了した後、次の割り込みの処理に迅速に入ることに関与する。

【0183】

図6に示されるように、第2のプロセッサは、ハードウェア支援切り替え装置と、第2のリダイレクト装置と、を含む。ハードウェア支援切り替え装置の構造は、図8を参照して理解され得る。図8に示されるように、ユーザモード割り込みを処理するとき、ハードウェア支援切り替え装置は、カーネルにトラップされることなく、スレッド切り替えをサポートし得る。ハードウェア支援切り替え装置は、プロセッサ状態決定モジュールと、スレッド切り替えモジュールと、カーネル通知モジュールと、を含む。

【0184】

プロセッサ状態決定モジュールは、プロセッサの現在の状態がユーザモードの割り込みへの即時応答をサポートするか否かを決定するように構成される。プロセッサは、プロセッサの現在の状態に基づいて、スレッド切り替え方式またはカーネル通知方式で割り込みに応答するために、ユーザモードにおけるシャドウスレッドの実行を決定する。

【0185】

スレッド切り替えモジュールは、シャドウスレッドに関する情報と、第1のスレッドに関する情報と、スレッド関連キャッシュ・データ・リフレッシュ・モジュールと、を含む。シャドウスレッドは、ユーザモードにおけるhost内で実行しているスレッドであり、割り込み処理関数を実行するように構成される。第2のプロセッサが、スレッドに関する情報を切り替える前に、第1のスレッドは、第2のプロセッサによって実行されるスレッドであり、シャドウスレッドに関する情報は、別のプロセッサ上で実行しているソフトウェアまたはハードウェアを使用することによって論理的に更新され得る。スレッド情報が切り替えられるとき、第1のスレッドに関する情報は、第1のスレッド情報モジュール内に格納される。ユーザモード割り込み処理が完了されたとき、第1のスレッドに関する関連情報は、第1のスレッド情報モジュールから復元される。キャッシュ・データ・リフレッシュ・モジュールが、プロセスまたは仮想マシン情報の識別をサポートしていない場合、スレッド情報切り替えが完了した後、スレッドに関連するキャッシュ情報は、キャッシュ・データ・リフレッシュ・モジュールを使用することによってリフレッシュされる必要がある。キャッシュ・データ・リフレッシュ・モジュールが、プロセスまたは仮想マシン情報の識別をサポートする場合、キャッシュ・データ・リフレッシュ・モジュールは、実行を実施するようにトリガされなくてもよい。

【0186】

第2のプロセッサが時間内にユーザモード割り込みに応答し得ない状態に、現在第2のプロセスがあるとき、限定しないが、第2のプロセッサが、カーネルモード割り込みまたは例外を処理するとき、プロセッサが、アイドル状態であるときなど、既存または新規に追加された割り込みまたは例外を生成することを含むだけに限らない方式が、実行のため、カーネルに、シャドウスレッドをウェイクアップさせるように通知するために使用される。

【0187】

第1のスレッドが仮想化特権レベルにある、と第2のプロセッサが決定したとき、第2のリダイレクト装置は、シャドウスレッドに関連する状態制御情報の更新を完了し、次に、仮想化特権レベルをhostのユーザモードに切り替える。

【0188】

前述の説明における「ユーザモードにおける割り込み要求を処理する」は、第2のプロセッサがユーザモード割り込みに応答し、ユーザモード割り込みを処理するために、ユーザモードにおけるシャドウスレッドを実行することであってもよい。図9に示されるよう

10

20

30

40

50

に、シャドウスレッドは、元のスレッドと、2重層切り替え機構を実施するコルーチンと、を含む。

【0189】

作成された後、元のスレッドは、スリーブをブロックし、ユーザモード割り込みサービスルーチンの機能的な関数を実行するために、カーネルによってウェイクアップされ得る。

【0190】

2重層切り替え機構は、ハードウェア支援切り替え装置と、第1のスレッドおよびシャドウ・スレッド・コンテキスト切り替えモジュールと、を含む。プロセッサが、ユーザモード割り込みに応答し得る場合、コルーチン実行ために必要な情報（メモリアドレス空間、データ構造などを含むが、これらに限定されない）は、ハードウェア支援切り替え装置を使用することによって切り替えられ、第1のスレッドとシャドウスレッドとの間のコンテキスト切り替えは、第1のスレッドおよびシャドウ・スレッド・コンテキスト切り替えモジュールを使用することによって実施され得る。シャドウスレッドのコルーチンは、ユーザモード割り込みに応答し、ユーザモード割り込みを処理するために、ユーザモード割り込みサービスルーチンの関数をさらに呼び出す。

10

【0191】

前述では、ハードウェア機能の観点から非同期処理手順を説明した。以下は、図10Aおよび図10Bを参照して、割り込み処理方法の手順について説明する。

【0192】

301：第1のプロセッサは、第1のVMが仮想I/Oデバイスに関連する動作を実施するときにトリガされる第1の割り込み要求を取得する。

20

【0193】

ステップ301の手順は、第1のVMが第1のVMの許可を超える動作を実施するときに、割り込みが発生することによってよく、第1のプロセッサは、ハードウェアロジックを使用することによって割り込みコードを検出してよく、第1の割り込み要求をさらに取得してもよい。

【0194】

第1の割り込み要求がVEI IDを含むことは、第1のVMがVEI IDを格納するためのレジスタにVEI IDを書き込み、第1のプロセッサがVEI IDを格納するためのレジスタからVEI IDを読み出すことであってもよい。

30

【0195】

302：第1のプロセッサは、第1のVMのVM IDを取得する。

【0196】

第1のプロセッサは、仮想マシン識別子を格納するためのレジスタからVM IDを取得し、第1のVMが、第1のプロセッサ上で実行するとき、VM IDが、仮想マシン識別子を格納するためのレジスタに書き込まれる。

【0197】

303：第1のプロセッサは、第1のVMのVEI IDおよびVM IDを、割り込みコントローラに送信する。

【0198】

304：割り込みコントローラは、第1のVMのVEI IDおよびVM IDに基づいて、対応するプロセッサ識別子、および対応するVEI物理番号を照会し、hostのものであり、第1の割り込み要求に対応する特権レベルを決定する。

40

【0199】

照会手順は、表1を参照して理解され得る。

【0200】

50

【表 1】

表 1 マッピングテーブル

VEI ID	VM ID	CPU ID	VEI 物理番号	Host の特権レベル
VEI 1	VM 1	CPU 2	VEI 100	ユーザモード
VEI 2	VM 1	CPU 2	VEI 101	カーネルモード
...

10

【0201】

対応するCPU ID、VEI物理番号、およびhostの特権レベルが、表1のVEI ID、VM ID、およびマッピング関係を使用することによって照会され得ることが、表1から分かる。

【0202】

対応するプロセッサ識別子およびVEI物理番号が発見された場合、ステップ305が実施される。そうでない場合、例外が報告されるか、または割り込み要求が無視される。

【0203】

305：第1の割り込み要求が、hostのユーザモードにおいて処理される必要がある場合、シャドウスレッドに関する情報が設定される。

【0204】

シャドウスレッドに関する情報は、シャドウスレッドによって占有されるメモリアドレス空間、スレッドデータ構造に関する情報、割り込み処理関数のエントリアドレスなどを含む。

20

【0205】

306：第1の割り込み要求が、hostのカーネルモードにおいて処理される必要がある場合、割り込みコントローラは、第2のプロセッサに、カーネルモードにおいて第1の割り込み要求を処理するように通知する。

【0206】

第2の割り込み要求は、ユーザモード仮想イベント割り込みである。

【0207】

307：割り込みコントローラは、第2の割り込み要求を第2のプロセッサに送信する。これに対応して、第2のプロセッサは、割り込みコントローラによって送信された第2の割り込み要求を受信する。

30

【0208】

第2の割り込み要求は、hostの特権レベルを示す情報を含む。

【0209】

308：第2のプロセッサは、第2のプロセッサの現在の状態が、対応するユーザモード割り込みをサポートしているか否かを決定し、第2のプロセッサの現在の状態が、対応するユーザモード割り込みをサポートしている場合、ステップ309が実施され、または、第2のプロセッサの現在の状態が、対応するユーザモード割り込みをサポートしていない場合、ステップ310が実施される。

40

【0210】

309：第2のプロセッサが、ユーザモード割り込みを現在サポートしている場合、第2のプロセッサは、第1のスレッドに関する情報をシャドウスレッドに関する情報に切り替える。

【0211】

第1のスレッドは、第2のプロセッサによって現在実行しているスレッドを指し、シャドウスレッドに関する情報は、ステップ305において、割り込みコントローラによって設定された情報である。

【0212】

50

310：第2のプロセッサが、ユーザモード割り込みを現在サポートしていない場合、第2のプロセッサは、カーネルに、仮想イベント割り込みを処理するように通知するために、例外または割り込みをトリガする。

【0213】

第2のプロセッサは、カーネルモードにおいて、例外または割り込みのトリガ原因を識別し、次に、仮想I/O処理スレッドをスケジューリングし、ウェイクアップさせる。仮想I/O処理スレッドは、仮想イベント割り込みを処理する。仮想イベント割り込みを処理した後、仮想I/O処理スレッドは、ブロックされ、次の仮想イベント処理の通知を待機する。

【0214】

第2のプロセッサが、現在アイドル状態にある場合、第2のプロセッサは、ユーザモード割り込みに応答し得ない。

【0215】

311：第2のプロセッサが、ユーザモード割り込みをサポートしている場合、第2のプロセッサは、第1のスレッドが仮想化特権レベルにあるか否かを決定し、第1のスレッドが仮想化特権レベルにある場合、ステップ312が実施され、または第1のスレッドが仮想化特権レベルにない場合、ステップ315が実施される。

【0216】

312：第2のプロセッサは、仮想化特権レベルからhostのユーザモードに切り替える。

【0217】

ステップ311は、図6の第2のリダイレクト装置によって実施されてもよい。

【0218】

313：第2のプロセッサは、VEI物理番号に対応する割り込み処理関数を実行するために、シャドウスレッドの割り込み処理関数のエントリアドレスに切り替える。

【0219】

コンピュータデバイスにおいて、VEI物理番号と割り込み処理関数との間のマッピング関係は、メモリを使用することによって保持され得、第2のプロセッサは、VEI物理番号に基づいてメモリから、対応する割り込み処理関数を読み出し得る。

【0220】

第1のスレッドが、仮想化特権レベルにある場合、第2のプロセッサ上で実行している第2のVMのコンテキストは格納され、シャドウスレッドのコンテキストは復元される。

【0221】

第2のプロセッサは、システムコールを使用することによってカーネルモードにトラップするために、ユーザモードにおけるスレッドを実行し、次に、物理I/Oデバイスを動作させるために、物理I/Oデバイスのドライバを呼び出し得る。

【0222】

314：第2のプロセッサは、第2の割り込み要求の実行を完了した後、第2のVMのコンテキストを復元し、割り込み前の仮想化特権レベルに戻るよう切り替え、第2のVMを実行して、割り込まれた命令の実行を継続するために、第2のVMの割り込まれた命令アドレスに切り替える。

【0223】

315：第1のスレッドが、仮想化特権レベルにない場合、ユーザモード割り込みに応答して、VEI物理番号に対応する割り込み処理関数を実行するために、シャドウスレッドの割り込み処理関数のエントリアドレスに切り替える。

【0224】

第1のスレッドが、仮想化特権レベルにない場合、図6のハードウェア支援切り替え装置は、ユーザモード割り込みに直接応答し、第2の割り込み要求を実行するために、シャドウスレッドの割り込み処理関数のエントリアドレスに切り替える。

【0225】

第1のスレッドが仮想化特権レベルにない場合、第1のスレッドのコンテキストが格納さ

10

20

30

40

50

れ、シャドウスレッドのコンテキストが復元される。

【0226】

316：割り込みスレッドが、仮想化特権レベルにない場合、第1のスレッドのコンテキストを復元し、割り込まれた命令の実行を継続するために、第1のスレッドに切り替える。

【0227】

非同期処理方式1では、VEI ID、VM ID、およびVEI物理番号が使用される。VEI IDがグローバルに一意である場合、図10Aおよび図10Bの割り込み処理手順は、簡略化され得、図10Aおよび図10Bに対応する実施形態で説明された対応する手順は、対応する第2のプロセッサの識別子、および対応するhostの特権レベルを決定するために、VEI IDを使用することによってのみ実施され得る。

【0228】

本出願の本実施形態で提供されるコンピュータシステムによれば、第1のVMの実行中に発生する例外または割り込みは、非同期処理方式において、割り込みコントローラおよび第2のプロセッサを使用することによって、処理のためにhostのユーザモードに転送される。したがって、hostのカーネルモードとhostのユーザモードとの間の切り替えのオーバーヘッドが必要とされず、コンピュータシステムの性能が改善される。

【0229】

非同期処理方式2：送信終了手順は、hostのカーネルを含まず、受信終了手順は、hostのカーネルを通過する。

【0230】

非同期処理方式2では、送信終了での手順は、前述の非同期処理方式1の手順と同じであり、仮想イベント割り込み構成要素を第1のプロセッサおよび割り込みコントローラに設定することによって実施される。

【0231】

違いは、割り込みコントローラが、カーネルの第2の割り込み要求を第2のプロセッサに送信することにある。

【0232】

図11に示されるように、第2のプロセッサは、カーネルの第2の割り込み要求を受信した後、カーネルの割り込み処理関数のエントリアドレスに切り替える。

【0233】

hostのカーネルモードの割り込み処理関数のエントリアドレスに切り替える前に、第2のVMが、第2のプロセッサ上で実行している場合、第2のVMの仮想化特権レベルは、hostのカーネルモードに切り替えられ、仮想マシンが、第2のプロセッサ上で実行していない場合、hostは、ユーザモードにあり、hostは、hostのユーザモードからhostのカーネルモードに最初に切り替えられ、hostが、カーネルモードにある場合、hostは、割り込み処理関数のエントリアドレスに直接切り替える。

【0234】

カーネルは、割り込み要求を処理し、スケジューラを使用することによって、第2の割り込み要求を処理するために、仮想I/O処理スレッド内のカーネル・モード・スレッドまたはユーザ・モード・スレッドをウェイクアップさせる。

【0235】

スケジューラが、ユーザ・モード・スレッドをウェイクアップさせる場合、処理手順は、第2のプロセッサが、システムコールを通してカーネルモードにトラップされるようにユーザ・モード・スレッドを実行し、次に、物理I/Oデバイスを動作させるために、物理I/Oデバイスのドライバを呼び出す、ことであってもよい。

【0236】

スケジューラが、カーネル・モード・スレッドをウェイクアップさせる場合、処理手順は、物理I/Oデバイスを動作させるために、物理I/Oデバイスのドライバを呼び出すようにカーネル・モード・スレッドを実行し得る。

【0237】

10

20

30

40

50

前述の非同期処理方式1の非同期仮想I/O処理フレームワークと比較して、本出願の本実施形態では、ハードウェア支援切り替えロジックおよびシャドウスレッド技術が、必要とされず、非同期仮想I/O処理フレームワークの設計および実装が、簡略化される。

【0238】

非同期処理方式3：同期方式が、送信終了手順で使用され、受信終了手順が、hostのカーネルを含まない。

【0239】

非同期処理方式3では、送信終了での手順は、同期仮想I/O処理フレームワーク内の第1のリダイレクト装置およびシャドウ実行環境に基づいて、非同期仮想I/O要求通知フレームワークを実装する。

【0240】

非同期処理方式3では、図12に示されるように、割り込みコントローラは、ユーザ・モード・マッピング・アドレス空間を含む。第1のVMは、前述の同期方式で説明されたソリューションを使用することによってユーザモード例外命令を生成し、ユーザ・モード・シャドウ実行環境に入り、割り込みコントローラ内のユーザ・モード・マッピング・アドレス空間にアクセスすることによって、割り込み要求を第2のプロセッサに送信し得る。割り込み要求は、ユーザモード割り込みであってもよいし、カーネルモード割り込みであってもよい。次に、割り込みコントローラは、非同期処理方式1における受信終了の手順を実行するように、割り込み要求を制御する。内容のこの部分については、理解のために、非同期処理方式1の受信終了の対応する内容を参照されたく、詳細は、ここでは再度説明されない。

【0241】

本出願の本実施形態では、割り込みを送信するオーバーヘッドは、コンテキスト切り替え手順のオーバーヘッドに低減され、仮想マシンの無関係なトラップアウト手順のオーバーヘッドは、必要とされず、非同期処理方式1の仮想イベント割り込み構成要素は、必要とされず、ハードウェア設計および実装の複雑さが低減される。

【0242】

非同期処理方式4：送信終了手順は、hostのカーネルを通過し、受信終了手順は、hostのカーネルを含まない。

【0243】

非同期処理方式4では、図13に示されるように、仮想マシンは、送信終了の手順で使用されなくてもよく、割り込みコントローラは、カーネル・マッピング・アドレス空間と、ユーザ・モード・マッピング・アドレス空間と、を含んでもよい。第1のプロセッサは、割り込みコントローラ内のカーネル・マッピング・アドレス空間にアクセスして、第1の割り込み要求をトリガするために、カーネル・モード・スレッドを実行し得、第1のプロセッサは、割り込みコントローラ内のユーザ・モード・マッピング・アドレス空間にアクセスして、第1の割り込み要求をトリガするために、ユーザ・モード・スレッドを実行し得る。すなわち、第1のプロセッサは、ユーザ・モード・スレッドまたはカーネル・モード・スレッドを使用することによって、第1の割り込み要求を割り込みコントローラに送信し得る。割り込みコントローラは、シャドウスレッドに関する情報を設定し、第2の割り込み要求を第2のプロセッサに送信する、ように構成される。第2のプロセッサは、第1のスレッドからシャドウスレッドに切り替え、第1のスレッドが、切り替わる前に第2のプロセッサによって実行されるスレッドであり、シャドウスレッドが、ユーザモードにおけるhost内で実行し、割り込み処理関数を含むスレッドであり、割り込み処理関数が、第2の割り込み要求に対応する。第2のプロセッサは、hostのユーザモードにおける割り込み処理関数を実行するように具体的に構成される。第2のプロセッサの処理手順については、前述の非同期処理方式1の受信終了の手順を参照されたい。詳細は、ここでは再度説明されない。

【0244】

本出願の本実施形態で提供されるソリューションによれば、非同期処理方式において、

10

20

30

40

50

ユーザ・モード・スレッドまたはカーネル・モード・スレッドによってトリガされる割り込みは、割り込みコントローラおよび第2のプロセッサを使用することによって、処理のためにhostのユーザモードに転送される。したがって、hostのカーネルモードとhostのユーザモードとの間の切り替えのオーバーヘッドが必要とされず、コンピュータシステムの性能が改善される。

【0245】

本出願の別の実施形態では、コンピュータ可読記憶媒体がさらに提供される。コンピュータ可読記憶媒体は、コンピュータ実行可能命令を格納する。デバイスのプロセッサがコンピュータ実行可能命令を実行するとき、デバイスは、図2～図5の例外処理方法を実施する。

10

【0246】

本出願の別の実施形態では、コンピュータ可読記憶媒体がさらに提供される。コンピュータ可読記憶媒体は、コンピュータ実行可能命令を格納する。デバイスのプロセッサが、コンピュータ実行可能命令を実行するとき、デバイスは、図6～図13の割り込み処理方法を実施する。

【0247】

本出願の別の実施形態では、コンピュータプログラム製品が、さらに提供される。コンピュータプログラム製品は、コンピュータ実行可能命令を含み、コンピュータ実行可能命令は、コンピュータ可読記憶媒体に格納される。デバイスのプロセッサがコンピュータ実行可能命令を実行するとき、デバイスは、図2～図5の例外処理方法を実施する。

20

【0248】

本出願の別の実施形態では、コンピュータプログラム製品が、さらに提供される。コンピュータプログラム製品は、コンピュータ実行可能命令を含み、コンピュータ実行可能命令は、コンピュータ可読記憶媒体に格納される。デバイスのプロセッサが、コンピュータ実行可能命令を実行するとき、デバイスは、図6～図13の割り込み処理方法を実施する。

【0249】

本出願の別の実施形態では、チップシステムが、さらに提供される。チップシステムは、プロセッサを含み、プロセッサは、図2～図5の例外処理方法を実施するように構成される。1つの可能な設計では、チップシステムは、メモリをさらに含む。メモリは、例外処理に必要なプログラム命令およびデータを格納するように構成される。チップシステムは、チップを含んでもよく、またはチップおよび別の個別構成要素を含んでもよい。

30

【0250】

本出願の別の実施形態では、チップシステムが、さらに提供される。チップシステムは、プロセッサを含み、プロセッサは、図6～図13の割り込み処理方法を実施するように構成される。1つの可能な設計では、チップシステムは、メモリをさらに含む。メモリは、割り込み処理のためのプログラム命令およびデータを格納するように構成される。チップシステムは、チップを含んでもよく、またはチップおよび別の個別構成要素を含んでもよい。

【0251】

当業者は、本明細書で開示された実施形態で説明された実施例と組み合わせて、ユニットおよびアルゴリズムステップが、電子ハードウェアによって、またはコンピュータソフトウェアと電子ハードウェアとの組合せによって、実装され得る、ことを認識し得る。機能が、ハードウェアまたはソフトウェアによって実施されるのかは、技術的ソリューションの具体的な用途および設計制約条件に依存する。当業者は、説明された機能を特定の用途ごとに実施するために、異なる方法を使用し得るが、実施態様が、本出願の実施形態の範囲を超えると考えられてはならない。

40

【0252】

簡易に説明するために、前述のシステム、装置およびユニットの詳細な作動プロセスについては、前述の方法実施形態の対応するプロセスを参照することは、当業者によって明確に理解されるといえる。詳細は、ここでは再度説明されない。

50

【0253】

本出願の実施形態において提供されるいくつかの実施形態では、開示したシステム、装置および方法が、他の方式で実施されてもよいと理解される。例えば、説明された装置の実施形態は、単なる一例にすぎない。例えば、ユニットへの分割は、論理的な機能の分割にすぎず、実際の実施態様では他の分割であってもよい。例えば、複数のユニットまたは構成要素は、組み合わせられ、もしくは別のシステムに統合されてもよく、またはいくつかの特徴は、無視され、もしくは実施されなくてもよい。さらに、表示または説明された相互結合、あるいは直接結合または通信接続は、いくつかのインターフェースを使用することによって実装されてもよい。装置またはユニット間の間接結合または通信接続は、電子的な形態、機械的な形態、または他の形態で実装されてもよい。

10

【0254】

別々の部品として説明されたユニットは、物理的に分離している場合も、そうでない場合もあり、ユニットとして表示された構成要素は、物理的ユニットである場合も、そうでない場合もある。言い換えると、一箇所に配置されてもよく、または複数のネットワークユニットに分散されていてもよい。ユニットの一部または全部は、実施形態のソリューションの目的を達成するための実際の要求事項に基づいて選択されてもよい。

【0255】

加えて、本出願の実施形態の機能部位が一体化されて1つの処理部位にされてもよいし、部位の各々が物理的に単独で存在してもよいし、あるいは、2つ以上の部位が一体化されて1つの部位にされる。

20

【0256】

機能がソフトウェア機能ユニットの形態で実装され、独立した製品として販売または使用される場合には、その機能はコンピュータ可読記憶媒体に格納されてもよい。かかる理解に基づいて、本出願の実施形態の技術的ソリューションは本質的に、または従来技術に寄与する部分は、または技術的解決策の全部もしくは一部は、ソフトウェア製品の形態で実装され得る。コンピュータソフトウェア製品は、記憶媒体に格納されており、(パーソナルコンピュータ、サーバ、ネットワークデバイスなどであってもよい)コンピュータデバイスに、本出願の実施形態で説明されている方法のステップの全部または一部を実施するように命令するためのいくつかの命令を含む。前述の記憶媒体は、USBフラッシュドライブ、リムーバブル・ハード・ディスク、読み取り専用メモリ(Read-Only Memory、ROM)、ランダム・アクセス・メモリ(Random Access Memory、RAM)、磁気ディスク、または光ディスクなど、プログラムコードを記憶し得る何らかの媒体を含む。

30

【符号の説明】

【0257】

- 100 コンピュータデバイス
- 101 仮想マシン(VM)
- 102 VM
- 103 アプリケーション
- 104 アプリケーション
- 105 ゲスト・オペレーティング・システム(OS)
- 106 ゲストOS
- 107 仮想プロセッサ
- 108 仮想プロセッサ
- 109 ホスト層
- 110 仮想マシンモニタ
- 111 ホスト・オペレーティング・システム
- 112 ハードウェア層
- 113 メモリ
- 114 プロセッサシステム
- 115 通信インターフェース

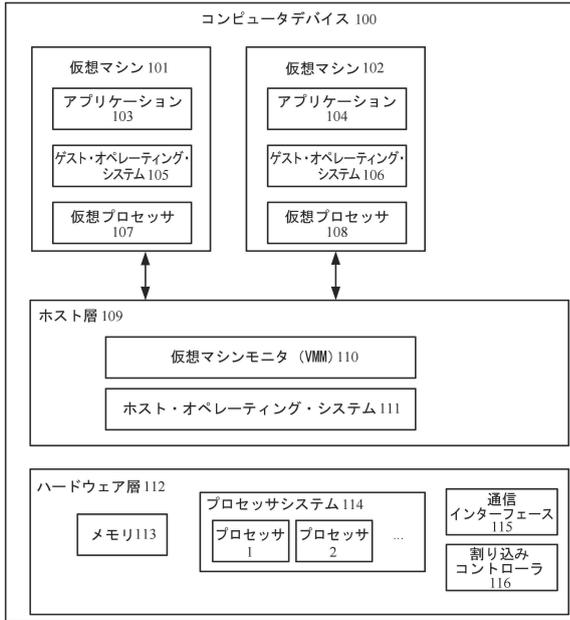
40

50

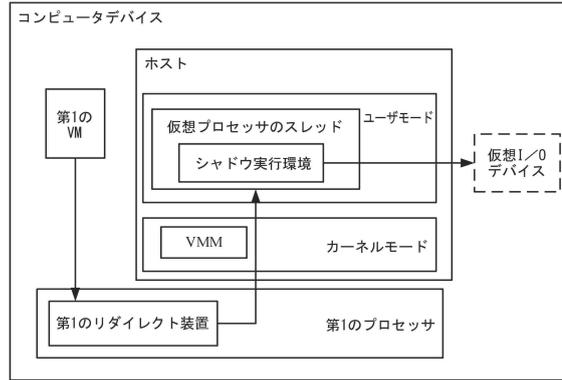
116 割り込みコントローラ

【図面】

【図 1】



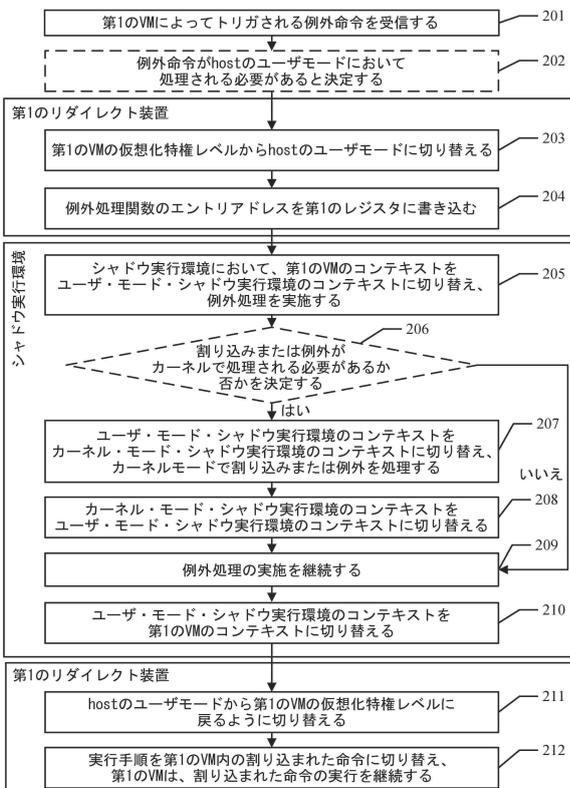
【図 2】



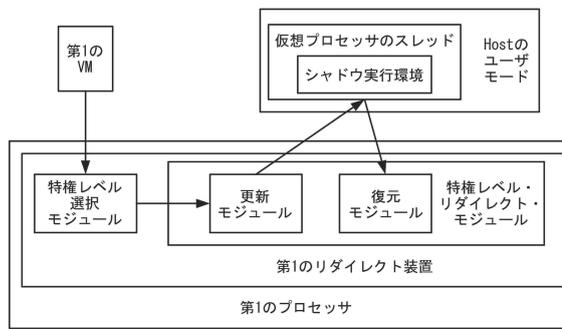
10

20

【図 3】



【図 4】

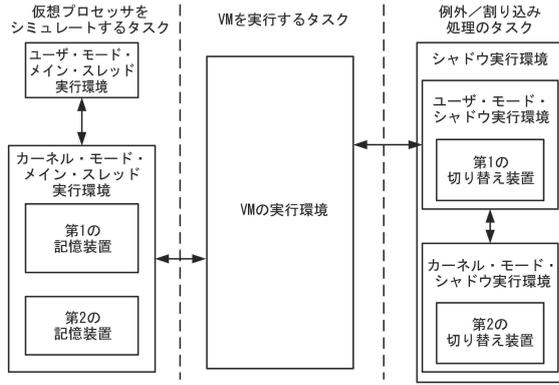


30

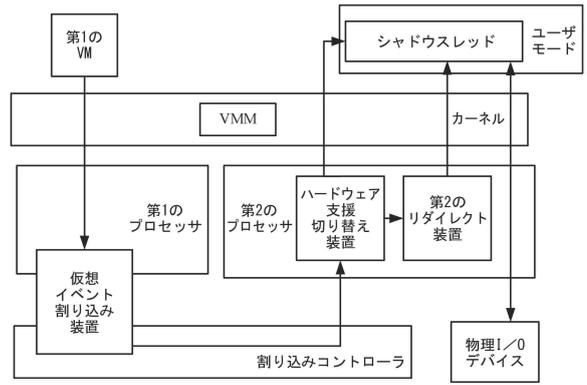
40

50

【図5】

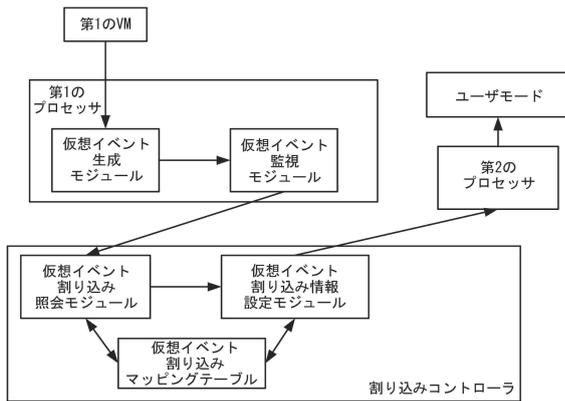


【図6】

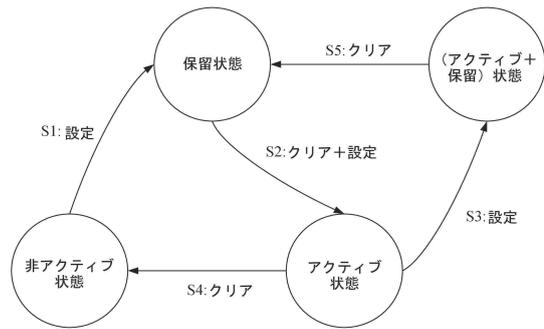


10

【図7A】

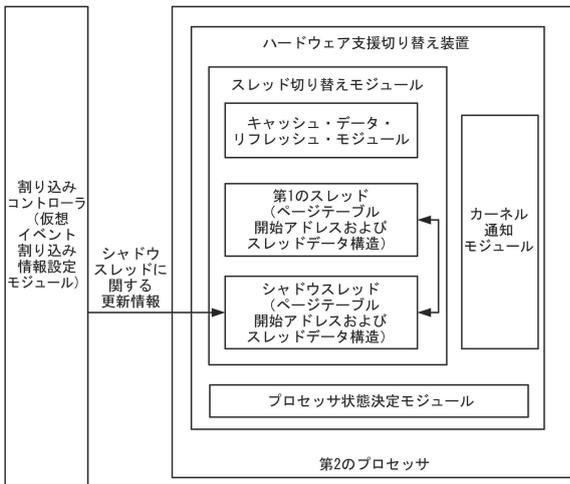


【図7B】

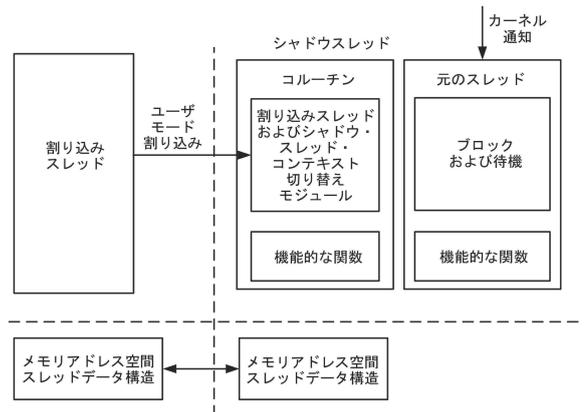


20

【図8】



【図9】

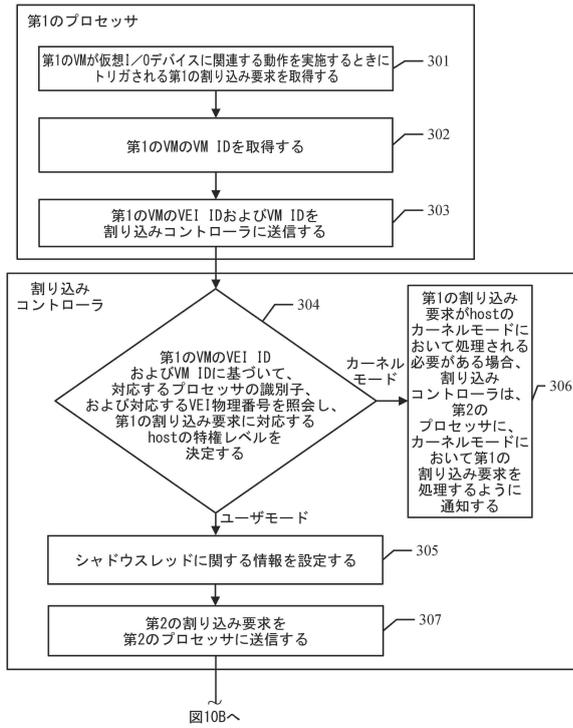


30

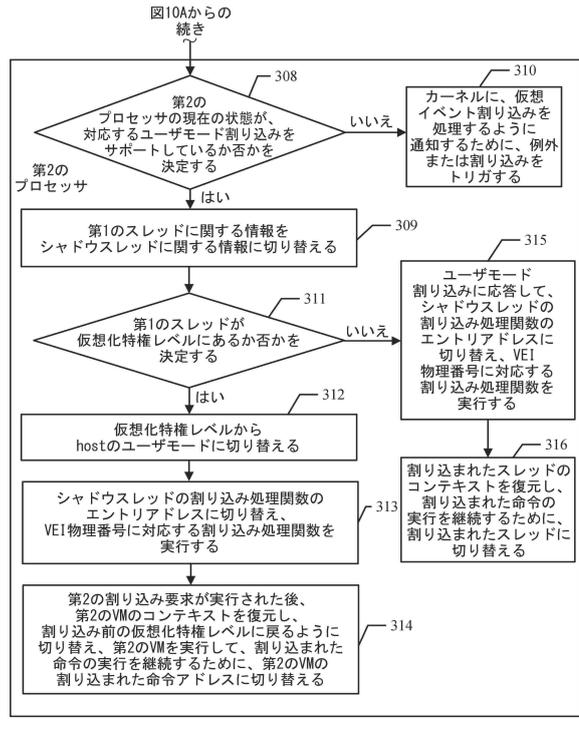
40

50

【図10A】



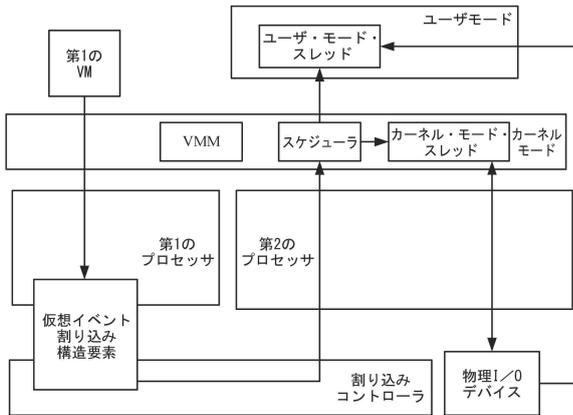
【図10B】



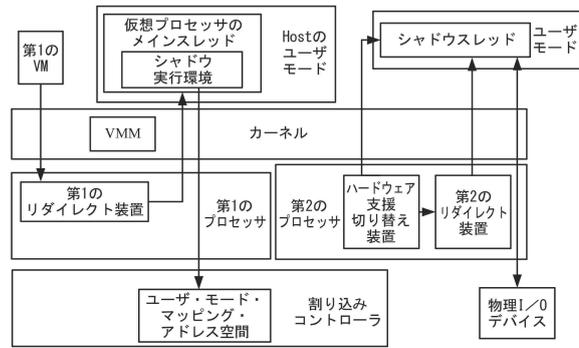
10

20

【図11】



【図12】

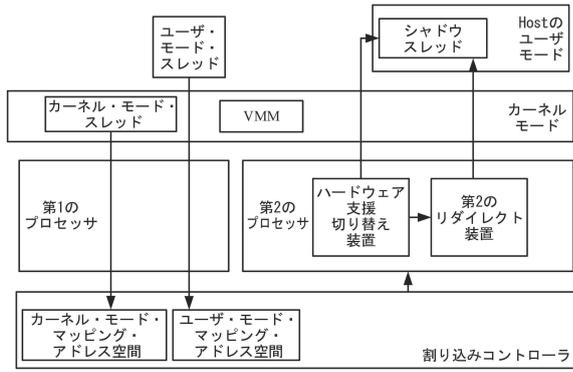


30

40

50

【図 13】



10

20

30

40

50

フロントページの続き

- (74)代理人 100133569
弁理士 野村 進
- (72)発明者 蔣 毅 飛
中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼
- (72)発明者 万 波
中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼
- (72)発明者 趙 思 齊
中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼
- (72)発明者 李 明旺
中華人民共和国 5 1 8 1 2 9 広東省深 チェン 市龍岗区坂田 華為総部 ベン 公楼
- 審査官 坂東 博司
- (56)参考文献 中国特許出願公開第 1 1 0 9 9 0 1 8 0 (C N , A)
米国特許出願公開第 2 0 2 0 / 0 1 2 5 3 9 7 (U S , A 1)
中国特許出願公開第 1 0 9 1 4 4 6 7 9 (C N , A)
特開 2 0 1 0 - 1 5 7 2 3 2 (J P , A)
米国特許出願公開第 2 0 1 0 / 0 1 6 9 8 8 3 (U S , A 1)
中国特許出願公開第 1 0 9 2 8 4 1 6 9 (C N , A)
米国特許出願公開第 2 0 1 1 / 2 9 6 4 1 1 (U S , A 1)
- (58)調査した分野 (Int.Cl. , D B 名)
G 0 6 F 9 / 4 5 5
G 0 6 F 9 / 4 8