(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第5680554号
(P5680554)

(45)発行日　平成27年3月4日(2015.3.4)　(24)登録日　平成27年1月16日(2015.1.16)

(51)Int.Cl.			F I		
G06F 9/48	***(2006.01)***		G06F	9/46	311B
G06F 9/46	***(2006.01)***		G06F	9/46	350
			G06F	9/46	311F

請求項の数 14　(全 35 頁)

(21)出願番号	特願2011-548227 (P2011-548227)
(86)(22)出願日	平成22年1月26日 (2010.1.26)
(65)公表番号	特表2012-515995 (P2012-515995A)
(43)公表日	平成24年7月12日 (2012.7.12)
(86)国際出願番号	PCT/US2010/022111
(87)国際公開番号	WO2010/085804
(87)国際公開日	平成22年7月29日 (2010.7.29)
審査請求日	平成25年1月28日 (2013.1.28)
(31)優先権主張番号	61/147,269
(32)優先日	平成21年1月26日 (2009.1.26)
(33)優先権主張国	米国(US)
(31)優先権主張番号	12/611,595
(32)優先日	平成21年11月3日 (2009.11.3)
(33)優先権主張国	米国(US)

(73)特許権者　591016172
アドバンスト・マイクロ・ディバイシズ・インコーポレイテッド
ＡＤＶＡＮＣＥＤ　ＭＩＣＲＯ　ＤＥＶＩＣＥＳ　ＩＮＣＯＲＰＯＲＡＴＥＤ
アメリカ合衆国、９４０８８－３４５３　カリフォルニア州、サニィベイル、ピィ・オゥ・ボックス・３４５３、ワン・エイ・エム・ディ・プレイス、メイル・ストップ・６８（番地なし）
(74)代理人　100108833
弁理士　早川　裕司
(74)代理人　100111615
弁理士　佐野　良太

最終頁に続く

(54)【発明の名称】各プロセッサに対して割り込み仮想化を支援するためのゲスト割り込み制御器

(57)【特許請求の範囲】

【請求項１】

システム内の複数のゲストの第１のゲストに割り当てられる周辺デバイスをソースとする割り込みをやりとりする割り込みメッセージを、入力／出力メモリ管理ユニット（ＩＯＭＭＵ）にて受信することと、

前記ＩＯＭＭＵが、メモリ内の１つ以上の第１のデータ構造にアクセスして前記割り込みを前記第１のゲストにマッピングすることであって、前記１つ以上の第１のデータ構造は他のメモリ位置へのポインタを含むことと、

ゲスト割り込みマネージャが、前記ポインタに応じて、１つ以上の第２のデータ構造をメモリに配置することであって、前記１つ以上の第２のデータ構造は前記第１のゲストに対する割り込み制御器状態を含むことと、

メモリ内の前記１つ以上の第２のデータ構造内に前記割り込みを記録して、前記第１のゲストが実行中であるときに前記割り込みが前記第１のゲストへ受け渡されることを可能にすることとを備えた方法。

【請求項２】

前記第１のゲストに割り当てられ且つ前記第１のゲスト内での前記割り込みに対する宛先である第１の割り込み制御器へ前記割り込みを伝えることを更に備えた請求項１に記載の方法。

【請求項３】

割り込み制御器状態は割り込み制御器内の割り込み要求レジスタの状態を含み、前記割

り込みは割り込みベクトルを含み、前記割り込み要求レジスタは前記ベクトルに関連するビット位置を含み、前記割り込みを記録することは前記データ構造内の前記ビット位置におけるビットをセットすることを備えている請求項１に記載の方法。

【請求項４】

前記ビットをセットすることはアトミックである請求項３に記載の方法。

【請求項５】

前記ビットをセットすることはセットされたビットについて前記ビット位置内へアトミックに論理和を取ることを備えている請求項４に記載の方法。

【請求項６】

システム上で実行可能なゲストを対象とする割り込みに対応する割り込みメッセージを受信するように構成されるゲスト割り込みマネージャであって、

前記ゲスト割り込みマネージャは、前記割り込みメッセージが受信されているときに前記ゲストが前記システム内でアクティブでないとしても前記ゲストへの前記割り込みの受け渡しを確実にするために前記割り込みをメモリシステム内のデータ構造内に記録するように構成されており、

前記ゲスト割り込みマネージャは、入力／出力メモリ管理ユニット（ＩＯＭＭＵ）からポインタを受信して、前記データ構造を前記メモリシステムに配置するように構成されており、前記ポインタは第２のデータ構造から前記ＩＯＭＭＵによって提供され、前記ＩＯＭＭＵは、前記第２のデータ構造にアクセスして前記割り込みメッセージを前記ゲストにマッピングするように構成されている、ゲスト割り込みマネージャ。

【請求項７】

前記データ構造は所与のゲストに対して割り込み制御器の状態の少なくとも一部分を記憶するように構成され、前記状態は割り込み要求レジスタの状態を含み、前記ゲスト割り込みマネージャは前記割り込み要求レジスタの前記状態を更新して前記割り込みを記録するように構成される請求項６に記載のゲスト割り込みマネージャ。

【請求項８】

前記割り込み要求レジスタは前記割り込み制御器によってサポートされる各割り込みベクトルに対応するビットを含み、前記割り込みメッセージは前記割り込みの前記割り込みベクトルを含み、前記ゲスト割り込みマネージャは前記割り込みメッセージからの前記割り込みベクトルに対応する前記ビットを更新するように構成される請求項７に記載のゲスト割り込みマネージャ。

【請求項９】

メモリシステムと、

請求項６～８のいずれかに記載のゲスト割り込みマネージャとを備えたシステム。

【請求項１０】

前記割り込みを開始するように構成される周辺デバイスと、

前記ＩＯＭＭＵとを更に備えた請求項９に記載のシステムであって、

前記ＩＯＭＭＵは前記割り込みを前記メモリ内の１つ以上の前記第２のデータ構造内のデータに応答する前記ゲストと関連付けるように構成され、前記ＩＯＭＭＵは前記ゲストと関連している前記割り込みに応答するゲストを識別するゲスト識別子を含む前記割り込みメッセージを送信するように構成されるシステム。

【請求項１１】

前記ゲストに割り当て可能な割り込み制御器を更に備えた請求項１０に記載のシステムであって、前記割り込み制御器は前記ゲストの前記ゲスト識別子を記憶するように構成され、前記割り込み制御器は受信した割り込みメッセージからの前記ゲスト識別子を前記割り込み制御器内のゲスト識別子と比較するように構成され、前記割り込み制御器は前記ゲスト識別子の比較における整合に応答する前記割り込みを受け入れるように構成されるシステム。

【請求項１２】

前記ゲスト割り込みマネージャは、前記ＩＯＭＭＵからの前記割り込みメッセージに応

答し且つ前記データ構造内に記録されている前記割り込みに応答して前記割り込みメッセージを前記割り込み制御器へ送信するように構成される請求項１１に記載のシステム。
【請求項１３】
ホストによって制御され且つプロセッサ上で実行可能なゲストを対象とするゲスト割り込みとやりとりするゲスト割り込みメッセージを前記プロセッサに結合されるゲスト割り込み制御器において受信することと、
前記ゲスト割り込み制御器内に記憶されるゲスト識別子を前記ゲスト割り込みメッセージにおける受信したゲスト識別子と整合させることと、
前記ゲスト割り込み制御器内の少なくとも１つの宛先識別子をゲストメッセージにおける受信した宛先識別子と整合させることと、
前記ゲスト識別子を整合させること及び前記少なくとも１つの宛先識別子を整合させることに応答する前記ゲスト割り込み制御器において前記割り込みを受け入れることとを備えた方法。
【請求項１４】
プロセッサ上で実行可能なホストを対象とする第１の割り込みとやりとりする第１の割り込みメッセージを前記プロセッサに結合される第１の割り込み制御器において受信することと、
前記ホストによって制御され且つ前記プロセッサ上で実行可能なゲストを対象とする第２の割り込みとやりとりする第２の割り込みメッセージを前記プロセッサに結合される請求項１３に記載のゲスト割り込み制御器である第２の割り込み制御器において受信することと、
前記第１の割り込みメッセージを受信することに応答して前記第１の割り込み制御器が前記第１の割り込みを前記プロセッサへ受け渡すことと、
前記第２の割り込みメッセージを受信することに応答して前記第２の割り込み制御器が前記第２の割り込みを前記プロセッサへ受け渡すこととを更に備えた請求項１３に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
この発明はプロセッサ及び仮想化に関し、更に特定的には仮想マシンゲストに割り込みを受け渡すこと(delivering)に関する。
【背景技術】
【０００２】
仮想化は、種々の異なる目的のためにコンピュータシステムにおいて用いられてきた。例えば、仮想化は、「コンテナ(container)」内で特権的ソフトウエア(privileged software)を実行するために用いることができ、コンテナは、仮想マシンを制御する仮想マシンマネジャ（ＶＭＭ）によって最初に許可されることなしに特権的ソフトウエアが少なくとも何らかの物理マシン状態に直接的にアクセスし且つ／又は少なくとも何らかの物理マシン状態を変化させることを防止するためのものである。そのようなコンテナは、「バグのある(buggy)」又は悪意のあるソフトウエアが物理マシン上で問題を生じさせることを防止し得る。また、仮想化は、２つ以上の特権的プログラムを同一の物理マシン上で同時に実行することを可能にするために用いられ得る。物理マシンへのアクセスは制御されるので、これらの特権的プログラムが互いに干渉しないようにすることができる。特権的プログラムは、オペレーティングシステムを含むであろうし、そして更に当該ソフトウエアを実行中のハードウエアの完全な制御を有することを期待する他のソフトウエアを含み得る。別の例においては、仮想化は、特権的プログラムによって予定されているハードウエアとは異なるハードウエア上でその特権的プログラムを実行するために用いられ得る。
【０００３】
一般に、プロセッサ又はコンピュータシステムの仮想化は、１つ以上の特権的プログラムに仮想マシン（前述したコンテナ）へのアクセスをもたらすことを含んでいるであろう

し、特権的プログラムは、仮想マシンを介して完全な制御を有するが、物理マシンの制御はＶＭＭによって保たれる。仮想マシンは、単一のプロセッサ（又は複数のプロセッサ）、メモリ、及び特権的プログラムがそれが実行中のマシン内で見出すことを予定している種々の周辺デバイスを含み得る。仮想マシン要素は、ＶＭＭが少なくとも一時的に仮想マシンに割り当てているハードウエアによって実装され得るし、且つ／又はソフトウエア内でエミュレートされ得る。各特権的プログラム（及び幾つかの場合には関連するソフトウエア、例えばオペレーティングシステム上で実行するアプリケーション）は、ここではゲスト(guest)と称される。仮想化は、ＶＭＭ及びその仮想マシンが実行される物理マシン内で何ら特定のハードウエア仮想化サポートなしにソフトウエア（例えば上述のＶＭＭ）において実装され得る。しかし、何らかのハードウエアサポートが提供されている場合には、仮想化は簡素化されるであろうし且つ／又はより高い性能を達成するであろう。

【発明の概要】

【発明が解決しようとする課題】

【０００４】

仮想化に伴い生じ得る問題は、割り込み受け渡しの待ち時間である。上述したように、仮想マシンによる使用のために周辺デバイスが割り当てられ得る（仮想マシン内で仮想周辺デバイスとして作用する）。そのような周辺デバイスは、仮想マシン内でソフトウエアによって処理されることになる割り込みを生成することができる。非仮想化環境においては、割り込み処理待ち時間は比較的短い可能性がある。仮想化環境においては、割り込みは、一般的には、ＶＭＭによってインターセプトされ(intercepted)、ＶＭＭによって処理され、そして何らかのソート(sort)のソフトウエアメカニズムを用いてＶＭＭによって目標とされる仮想マシンへ受け渡され得る。その一方で、割り込み処理待ち時間は、著しく長い可能性がある（例えば１００倍長い）。

【課題を解決するための手段】

【０００５】

１つの実施形態においては、システムは、プロセッサと、プロセッサに結合される第１の割り込み制御器と、プロセッサに結合される第２の割り込み制御器とを備えている。第１の割り込み制御器は、システム内のホストを対象とする第１の割り込みとやりとりする第１の割り込みメッセージを受信することに応答する割り込みに対する信号をプロセッサへ送るように構成される。第２の割り込み制御器は、ホストによって制御され且つプロセッサ上で実行可能なゲストを対象とする第２の割り込みとやりとりする第２の割り込みメッセージを受信することに応答する割り込みに対する信号をプロセッサへ送るように構成される。

【図面の簡単な説明】

【０００６】

以下の詳細な説明は添付の図面を参照にしこれらをここで簡単に説明する。

【０００７】

【図１】図１は仮想化を実装しているコンピュータシステムの１つの実施形態のブロック図である。

【０００８】

【図２】図２は図１に示されるホストハードウエアの１つの実施形態のブロック図である。

【０００９】

【図３】図３はゲストに受け渡されつつある割り込みの１つの実施形態を示すブロック図である。

【００１０】

【図４】図４はゲスト進歩的プログラム可能割り込み制御器（ＡＰＩＣ）の１つの実施形態を示すブロック図である。

【００１１】

【図５】図５はゲストＡＰＩＣ状態データ構造におけるゲストＡＰＩＣ状態エントリの１

つの実施形態を示すブロック図である。
【００１２】
【図６】図６はゲストを目標としている割り込みを受信することに応答する図２に示されるゲスト割り込みマネージャの１つの実施形態の動作を示すフローチャートである。
【００１３】
【図７】図７は割り込みメッセージを受信することに応答するゲストＡＰＩＣの１つの実施形態の動作を示すフローチャートである。
【００１４】
【図８】図８はゲストＡＰＩＣ状態を１つのゲストから他へ変更する仮想マシンモニタの１つの実施形態の動作を示すフローチャートである。
【００１５】
【図９】図９はゲストＡＰＩＣ状態エントリ内で割り込み状態を配列する１つの実施形態を示すブロック図である。
【００１６】
【図１０】図１０は割り込みに対してゲストＡＰＩＣ状態エントリを配置する１つの実施形態を示すブロック図である。
【００１７】
【図１１】図１１は割り込みに対してゲストＡＰＩＣ状態エントリを配置する別の実施形態を示すブロック図である。
【００１８】
【図１２】図１２は割り込みに対してゲストＡＰＩＣ状態エントリを配置する更に別の実施形態を示すブロック図である。
【００１９】
【図１３】図１３は割り込みに対してゲストＡＰＩＣ状態エントリを配置するまた更に別の実施形態を示すブロック図である。
【００２０】
【図１４】図１４は図１に示されるホストハードウエアの別の実施形態のブロック図である。
【００２１】
【図１５】図１５はＶＭＭの１つの実施形態を記憶しているコンピュータアクセス可能記憶媒体の１つの実施形態のブロック図である。
【発明を実施するための形態】
【００２２】
本発明は種々の修正及び代替的な形態を許容する一方で、その特定の実施形態が図面における例によって示されており、そして詳細にここに説明されることになる。しかし、図面及びそれに対する詳細な説明は、開示された特定の形態に本発明を限定することを意図するものではなく、それとは対照的に、添付の特許請求の範囲によって画定されるような本発明の精神及び範囲内にある全ての修正、均等なもの、及び代替案に及ぶことが意図されていることが理解されるべきである。ここで用いられる表題は組織化の目的だけのためのものであり、明細書の範囲を限定することを意味しない。この出願を通して用いられる「かもしれない、あり得る、し得る、であろう、ことがある、可能性がある、してよい、あってよい、することができる(may)」の語は、必須の意味（即ち「なければならない(must)」を意味する）ではなく、許容の意味（即ち「可能性がある」を意味する）である。同様に、「含む(include, includes)」及び「含んでいる(including)」の語は含んでいることを意味し、限定することを意味しない。
【００２３】
種々のユニット、回路、又は他のコンポーネントが単一又は複数のタスクを実行「するように構成される(configured to)」として説明されることがある。そのような文脈において、「するように構成される」は、動作の間に単一又は複数の当該タスクを実行する「回路を有する(having circuitry that)」ことを一般的に意味する構造の広い記載である

。従って、ユニット／回路／コンポーネントは、ユニット／回路／コンポーネントがその時点でオンでないとしても、当該タスクを実行するように構成されていることができる。一般に、「するように構成される」に対応する構造を形成する回路は、当該動作を実装するハードウエア回路を含み得る。同様に、種々のユニット／回路／コンポーネントは、説明の便宜上、単一又は複数のタスクを実行するものとして説明されることがある。そのような説明は、「するように構成される」の句を包含するものとして解釈されるべきである。１つ以上のタスクを実行するように構成されるユニット／回路／コンポーネントを記載することで、当該ユニット／回路／コンポーネントに対して３５Ｕ．Ｓ．Ｃ．§１１２、第６段落解釈を行使しないよう明示的に意図するものである。

【００２４】

実施形態においては、コンピュータシステムは、少なくとも１つのホスト割り込み制御器と少なくとも１つのゲスト割り込み制御器を含む。ホスト割り込み制御器は、ホスト（例えば仮想化された環境における仮想マシンマネージャ又はＶＭＭ）によってサービスされることになる割り込みを管理してよい。そのような割り込みの例としては、システム上で実行中のゲストに割り当てられていないコンピュータシステム内のデバイスからの割り込み、ＶＭＭがゲストにさらすことを望まないシステムレベル割り込み、等を挙げることができる。ゲスト割り込み制御器は、ゲストによってサービスされることになる割り込みを管理してよい。そのような割り込みの例としては、ゲストの仮想マシンに対してデバイスの機能性を提供するようにゲストに割り当てられたそのデバイスによって発行される割り込みを挙げることができる。

【００２５】

システム内のハードウエアは、ホスト割り込み制御器又はゲスト割り込み制御器のいずれかへ割り込みを送信してよい。代替的には、割り込みは、両方の割り込み制御器へ送信されてよく、そしてこれらの割り込み制御器の一方は、それがホスト割り込みであるかゲスト割り込みであるかに基いて割り込みを受け入れてよい。各割り込み制御器は、プロセッサに結合されていてよく、また割り込みを受け渡すためにそのプロセッサと通信してよい。幾つかの実施形態においては、割り込みをゲストへ受け渡すためにゲスト割り込み制御器を設けることは、ホスト割り込み待ち時間と概ね等しい待ち時間を伴うゲスト割り込み受け渡しをもたらし得る。

【００２６】

実施形態においては、ゲスト割り込み制御器は、ホスト割り込み制御器の複製であってよい。即ち、ゲスト割り込み制御器は、ホスト割り込み制御器と同一のハードウエアを含んでいてよい。別の実施形態においては、ゲスト割り込み制御器は、ホスト割り込み制御器の一部分のみを複製していてよい。ゲスト割り込み制御器によって複製されていない部分は、ＶＭＭインターセプト及びＶＭＭによるエミュレーション(emulation)を介して仮想化されてよい。

［仮想化概説］

【００２７】

図１は仮想化を実装しているコンピュータシステム５の１つの実施形態のブロック図を示している。図１の実施形態においては、多重ゲスト１０Ａ～１０Ｎが示されている。ゲスト１０Ａは、ゲストオペレーティングシステム（ＯＳ）１２と、ゲストＯＳ１２上で実行される１つ以上のアプリケーション１４Ａ～１４Ｎとを含む。ゲスト１０Ｎは特権的コード１６を含む。ゲスト１０Ａ～１０Ｎは、仮想マシンマネージャ（ＶＭＭ）１８によって管理される。ＶＭＭ１８及びゲスト１０Ａ～１０Ｎはホストハードウエア２０上で実行され、ホストハードウエア２０は、コンピュータシステム５内に含まれる物理的ハードウエアを備えていてよい。１つの実施形態においては、ＶＭＭ１８は一連の仮想マシン制御ブロック（ＶＭＣＢ）２２を維持していてよい。各ゲスト１０Ａ～１０Ｎに対して１つのＶＭＣＢ２２があってよい。図１においては例示のためにＶＭＣＢ２２がＶＭＭ１８の一部として図示されているが、ＶＭＣＢ２２は、メモリ内に記憶されていてよく且つ／又はホストハードウエア２０内のディスクドライブのような不揮発性媒体上に記憶されていて

よい。

【００２８】

ホストハードウエア２０は、一般的にはコンピュータシステム５内に含まれるハードウエアの全てを含む。種々の実施形態において、ホストハードウエア２０は、１つ以上のプロセッサ、周辺デバイス、及びこれらのコンポーネントを結合するために用いられる他の回路を含んでいてよい。具体的には、ホストハードウエア２０は、１つ以上のホスト割り込み制御器、１つ以上のゲスト割り込み制御器、及び／又は１つ以上のゲスト割り込みマネージャを含んでいてよい。例えば、パーソナルコンピュータ（ＰＣ）型システムは、複数のプロセッサを結合するノースブリッジ(Northbridge)と、メモリと、進歩的なグラフィックポート(advanced graphic port)（ＡＧＰ）インタフェースを用いるグラフィクスデバイスとを含んでいてよい。加えて、ノースブリッジは、周辺コンポーネントインタフェース（ＰＣＩ）バスのような周辺バスに結合していてよく、周辺バスには種々の周辺コンポーネントが直接的に又は間接的に結合されていてよい。レガシー(legacy)機能性を提供し且つ／又はレガシーハードウエアに結合するために、ＰＣＩバスに結合されるサウスブリッジ(Southbridge)もまた含まれていてよい。種々の実装において、ゲスト割り込みマネージャは、ノースブリッジ内、サウスブリッジ内、又はインタフェースの１つ上のデバイス内に実装されていてよい。ホスト割り込み制御器及びゲスト割り込み制御器は、各プロセッサに対して実装されてよく、又はプロセッサの群によって共有されてよい。他の実施形態においては、種々のハードウエアコンポーネントをリンクするために、他の回路が用いられてよい。例えば、ハイパートランスポート(HyperTransport)（商標）（ＨＴ）リンクが複数のノードをリンクするために用いられてよく、ノードの各々は、１つ以上のプロセッサ、ホストブリッジ、及びメモリ制御器を含んでいてよい。各ノードもノースブリッジを含んでいてよく、そのノースブリッジは、ゲスト割り込みマネージャ並びに／又はホスト割り込み制御器及びゲスト割り込み制御器を含んでいてよい。代替的には、ホストブリッジがゲスト割り込みマネージャ並びに／又はホスト割り込み制御器及びゲスト割り込み制御器を含んでいてもよい。ホストブリッジは、ＨＴリンクを介してデイジーチェーン形式(daisy chain fashion)で周辺デバイスと結合するために用いられてよい。任意の所望の回路／ホストハードウエア構造が用いられてよい。

【００２９】

ＶＭＭ１８は、ゲスト１０Ａ～１０Ｎの各々に対する仮想化を提供するように構成されてよく、そしてホストハードウエア２０へのゲスト１０Ａ～１０Ｎのアクセスを制御してよい。ＶＭＭ１８はまた、ホストハードウエア２０上での実行のためにゲスト１０Ａ～１０Ｎをスケジューリングすることに関与してよい。ＶＭＭ１８は、ホストハードウエア２０内で仮想化のために提供されるハードウエアサポートを用いるように構成されてよい。例えば、イベントをインターセプトしそして処理のためにゲストからＶＭＭ１８へ退出するためのハードウエアを含めて、プロセッサは仮想化のためのハードウエアサポートを提供してよい。ゲスト割り込みマネージャ及び／又はゲスト割り込み制御器は、仮想化をサポートするために設けられるハードウエアであってもよい。

【００３０】

幾つかの実施形態においては、ＶＭＭ１８は、ホストハードウエア２０上で実行される「薄い(thin)」スタンドアローンソフトウエアプログラムとして実装されてよく、そしてゲスト１０Ａ～１０Ｎに対して仮想化を提供する。そのようなＶＭＭ実装は、しばしば「ハイパーバイザ(hypervisor)」と称されることがある。他の実施形態においては、ＶＭＭ１８はホストＯＳ内へ統合されていてよく又はホストＯＳ上で実行されてよい。そのような実施形態においては、ＶＭＭ１８は、ホストＯＳ内の任意のドライバ、システムＢＩＯＳによって提供されるプラットフォームシステム管理モード（ＳＭＭ）コード、等を含めてホストＯＳに依存するであろう。従って、ホストＯＳコンポーネント（及びプラットフォームＳＭＭコードのような種々の低レベルコンポーネント）は、ホストハードウエア２０上で直接的に実行され、またＶＭＭ１８によっては仮想化されない。ＶＭＭ１８及びホストＯＳ（含まれる場合には）は、１つの実施形態においては、一緒にホストと称される

ことがある。一般に、ホストは、使用に際してホストハードウエア２０の直接制御内にある任意のコードを含んでいてよい。例えば、ホストは、ＶＭＭ１８であってよく、ホストＯＳを伴うＶＭＭ１８であってよく、あるいはホストＯＳ単独（例えば非仮想化環境において）であってよい。

【００３１】

種々の実施形態において、ＶＭＭ１８は、完全仮想化(full virtualization)、準仮想化(paravirtualization)、又は両方をサポートしてよい。更に、幾つかの実施形態においては、ＶＭＭ１８は、準仮想化されている複数のゲスト及び完全仮想化されている複数のゲストを同時に実行してよい。

【００３２】

完全仮想化では、ゲスト１０Ａ～１０Ｎは、仮想化が発生中であることに気付かない。各ゲスト１０Ａ～１０Ｎは、その仮想マシン内に連続的なゼロベースのメモリを有していてよく、そしてＶＭＭ１８は、シャドウ(shadow)ページテーブル又は入れ子にされた(nested)ページテーブルを用いてホスト物理アドレス空間へのアクセスを制御してよい。シャドウページテーブルは、ゲスト仮想アドレスからホスト物理アドレスへ再マッピングしてよく（事実上は、ゲスト１０Ａ～１０Ｎ内のメモリ管理ソフトウエアによって割り当てられたゲスト「物理アドレス」のホスト物理アドレスへの再マッピング）、一方、入れ子にされたページテーブルは、入力としてのゲスト物理アドレスを受信し、そしてホスト物理アドレスへマッピングしてよい。シャドウページテーブル又は入れ子にされたページテーブルを各ゲスト１０Ａ～１０Ｎに対して用いて、ＶＭＭ１８は、ゲストがホストハードウエア２０内の他のゲストの物理メモリへアクセスしないことを確実にし得る。

【００３３】

準仮想化では、ゲスト１０Ａ～１０Ｎは、少なくとも部分的にはＶＭアウェア(VM-aware)であってよい。そのようなゲスト１０Ａ～１０Ｎは、ＶＭＭ１８とメモリページについて交渉してよく、従って、ゲスト物理アドレスをホスト物理アドレスへ再マッピングすることは必要ないかもしれない。１つの実施形態では、準仮想化において、ゲスト１０Ａ～１０Ｎは、ホストハードウエア２０内の周辺デバイスと直接的に相互作用することが許可されてよい。任意の所与の時間において、周辺デバイスは、単一又は複数のゲスト１０Ａ～１０Ｎによって「所有されて(owned)」よい。１つの実装においては、例えば、周辺デバイスは、その周辺デバイスをその時点で所有している１つ以上のゲスト１０Ａ～１０Ｎと共に保護ドメイン内へマッピングされてよい。周辺デバイスを所有しているゲストのみが、その周辺デバイスと直接的に相互作用することができる。保護ドメイン内のデバイスが別の保護ドメイン内のゲストに割り当てられているページに対して読み出し／書き込みすることを防止するための保護メカニズムがあってもよい。

【００３４】

前述したように、ＶＭＭ１８は各ゲスト１０Ａ～１０Ｎに対してＶＭＣＢ２２を維持してよい。ＶＭＣＢ２２は、一般的には、対応するゲスト１０Ａ～１０Ｎに対してＶＭＭ１８によって割り当てられる記憶区域内に記憶されるデータ構造を備えていてよい。１つの実施形態においては、ＶＭＣＢ２２はメモリのページを備えていてよいが、他の実施形態は、より大きな若しくはより小さなメモリ区域を用いてよく且つ／又は不揮発性ストレージのような他の媒体上での記憶を用いてよい。１つの実施形態においては、ＶＭＣＢ２２はゲストのプロセッサ状態を含んでいてよく、ゲストのプロセッサ状態は、ゲストの実行をスケジューリングされるときにホストハードウエア２０におけるプロセッサ内へロードされてよく、またゲストが退出するときに（当該スケジューリングされた時間を完了したこと、又はゲストを退出するためにプロセッサが検出している１つ以上のインターセプトのいずれかに起因して）、ＶＭＣＢ２２へ記憶し戻されてよい。幾つかの実施形態においては、ＶＭＣＢ２２に対応しているゲストへ制御を移す命令（「仮想マシン実行(Virtual Machine Run)（ＶＭＲＵＮ）」命令）を介してプロセッサ状態の一部分のみがロードされ、そして他の望まれる状態は、ＶＭＲＵＮ命令を実行するのに先立ちＶＭＭ１８によってロードされてよい。同様に、そのような実施形態においては、プロセッサ状態の一部分

のみが、ゲスト退出に際してプロセッサによってＶＭＣＢ２２へ記憶されてよく、またＶＭＭ１８は、必要に応じて任意の追加的な状態を記憶することに関与してよい。他の実施形態においては、ＶＭＣＢ２２は、プロセッサ状態が記憶されている別のメモリ区域へのポインタを含んでいてよい。また、１つの実施形態においては、２つ以上の退出メカニズムが規定されてよい。１つの実施形態においては、記憶されている状態の量及びロードされる状態のロケーションは、どの退出メカニズムが選択されているかに依存するであろう。

【００３５】

１つの実施形態においては、ＶＭＭ１８はまた、ＶＭＭ１８に対応しているプロセッサ状態を記憶するように割り当てられるメモリの区域を有していてよい。ＶＭＲＵＮが実行される場合、ＶＭＭ１８に対応しているプロセッサ状態は、その区域内に保存されてよい。ゲストがＶＭＭ１８へ退出すると、その区域からのプロセッサ状態がその区域からリロードされてＶＭＭ１８が実行を継続することを可能にし得る。１つの実装においては、例えばプロセッサは、ＶＭＭ１８保存区域のアドレスを記憶するためにレジスタ（例えば特定モデル向けレジスタ又はＭＳＲ）を実装していてよい。

【００３６】

また、ＶＭＣＢ２２は、ゲストに対して有効にされているインターセプトイベントを識別するインターセプト構成と、有効にされたインターセプトイベントが検出される場合にゲストを退出するためのメカニズムとを含んでいてよい。１つの実施形態においては、インターセプト構成は、一連のインターセプト表示と、プロセッサがサポートしている各インターセプトイベントに対する１つの表示とを含んでいてよい。インターセプト表示は、対応するイベントをプロセッサがインターセプトすべきか否か（又は別の見方をすればインターセプトが有効にされているか否か）を表示してよい。ここで用いられているように、イベントがゲスト内で「インターセプトされる(intercepted)」のは、そのイベントがそのゲスト内で発生すべきでプロセッサがそのイベントの処理のためにそのゲストを退出する場合である。１つの実施形態においては、インターセプト構成は、２つの退出メカニズムのどちらが用いられるかを表示する第２の一連の表示を含んでいてよい。他の実施形態は２つを超える退出メカニズムを規定してよい。別の実施形態においては、インターセプト構成は、第１の退出メカニズムがイベントに対して用いられるべきか否かを表示する、インターセプトイベント毎の一連のインターセプト表示と、第２の退出メカニズムがイベントに対して用いられるべきか否かを表示する、インターセプトイベント毎の第２の一連のインターセプト表示とを備えていてよい。

【００３７】

一般に、退出メカニズムは、ゲスト実行から退出する（通常は再起動可能なやり方で）と共に他のコードの実行を開始するためにプロセッサによって実行される動作を規定してよい。１つの実施形態においては、１つの退出メカニズムは、少量のプロセッサ状態を保存することと、ミニバイザ(Minivisor)に対する状態をロードすることとを含んでいてよい。ミニバイザは、ゲスト物理アドレス空間内で実行されてよく、そして比較的単純なインターセプト処理を実行してよい。大量のプロセッサ状態を保存すると共にＶＭＭのプロセッサ状態をロードする別の退出メカニズムがＶＭＭへ退出してよい。従って、インターセプトイベントは、イベントに応じて異なる命令コードによって処理されてよい。また、比較的単純なインターセプト処理は、実行するのにより少ない時間を費やすであろう「より軽量な(lighter weight)」退出メカニズムを通して処理されてよく、これにより幾つかの実施形態においては性能が改善され得る。より複雑な処理は、「より重たい(heavier weight)」メカニズムが退出のために用いられた後で、ＶＭＭ内で実行されてよい。従って、この実施形態においては、ＶＭＭ１８は、ゲスト１０Ａ～１０Ｎが内部的に処理することをＶＭＭ１８が望まないイベントをインターセプトするようにプロセッサを構成してよく、またＶＭＭ１８は、退出メカニズムを使用すべきプロセッサを構成してもよい。イベントとしては、命令（即ち、命令を実行する代わりに命令をインターセプトする）、割り込み、例外、及び／又はゲスト実行の間に発生し得る他の所望のイベントを挙げることが

できる。

【００３８】

１つの実施形態においては、ＶＭＣＢ２２は、ＶＭＣＢ２２をロードするに際して特定の動作をプロセッサに実行させることができる他の制御ビットを更に含んでいてよい。例えば、制御ビットは、プロセッサ内にＴＬＢをフラッシュする(flush)するための命令を含んでいてよい。他の制御ビットは、ゲストに対する実行環境（例えば割り込み処理モード、ゲストに対するアドレス空間識別子、等）を指定してよい。更に他の制御ビットは、何故ゲストが退出したか等を説明する退出コードとやりとりするために用いられてよい。

【００３９】

一般的に、「ゲスト」は、コンピュータシステム５内での実行のために仮想化されることになる任意の１つ以上のソフトウエアプログラムを備えていてよい。ゲストは、特権的なモードにおいて実行される少なくとも何らかのコードを含んでいてよく、従ってそれが実行中のコンピュータシステムにわたって完全な制御を有することを期待する。前述したように、ゲスト１０Ａは、ゲストがその内部にゲストＯＳ１２を含む例である。ゲストＯＳ１２は任意のＯＳであってよく、例えば、マイクロソフトコープ（レドモンド、ワシントン）から入手可能なウインドウズＯＳの任意のもの、ＩＢＭコーポレーション（アーモンク、ニューヨーク）から入手可能なリナックス、ＡＩＸのような任意のユニックス型オペレーティングシステム、サンマイクロシステムズインク（サンタクララ、カリフォルニア）から入手可能なソラリス、ヒューレットパッカードカンパニー（ポロアルト、カリフォルニア）から入手可能なＨＰ－ＵＸ、等であってよい。ゲスト１０Ｎは、非ＯＳ特権的コード１６を備えたゲストの例である。

【００４０】

尚、ここで用いられる場合における１０Ｎのような参照番号内の「Ｎ」の文字は、当該参照番号を有する任意の数の要素を総称的に表示することを意味している（例えば任意の数のゲスト１０Ａ～１０Ｎは１つのゲストを含む）。また、文字「Ｎ」を用いる異なる参照番号（例えば１０Ｎと１４Ｎ）は、特に断りのない限り、同様の数の異なる要素が設けられていることを表示するようには意図されていない（例えば、ゲスト１０Ａ～１０Ｎの数は、アプリケーション１４Ａ～１４Ｎの数とは異なっていてよい）。

［ホストハードウエア］

【００４１】

次に図２を参照すると、ホストハードウエア２０の１つの実施形態を表すブロック図が示されている。図示される実施形態においては、ホストハードウエア２０は、複数のプロセッサ３０Ａ～３０Ｂと、それぞれのホスト進歩的プログラム可能割り込み制御器(host Advanced Programmable Interrupt Controllers)（ｈＡＰＩＣ）３２Ａ～３２Ｂと、それぞれのゲストＡＰＩＣ（ｇＡＰＩＣ）３４Ａ～３４Ｂと、随意的な付加的ｇＡＰＩＣ３４Ｃ～３４Ｄと、ブリッジ３６（ゲスト割り込みマネージャ３８、入力／出力（Ｉ／Ｏ）メモリ管理ユニット（ＩＯＭＭＵ）４０、及びメモリ制御器４２を含む）と、複数のインタフェース回路（ＩＦ）４４Ａ～４４Ｃと、メモリインタフェース回路（ＭＩＦ）４６と、ＩＯＡＰＩＣ５０を含んでいてよい随意的なブリッジ４８と、周辺機器５２Ａ～５２Ｂ（そのうちの幾つかはＩＯＡＰＩＣ５４のようなＩＯＡＰＩＣを含んでいてよい）と、メモリ５６とを含む。プロセッサ３０Ａ～３０Ｂは、図２に示されるように、ブリッジ３６と結合され、またそれぞれのｈＡＰＩＣ３２Ａ～３２Ｂ及びｇＡＰＩＣ３４Ａ～３４Ｄと結合されている。ｈＡＰＩＣ３２Ａ～３２Ｂ及びｇＡＰＩＣ３４Ａ～３４Ｄはブリッジ３６と結合され、ブリッジ３６は、インタフェース回路４４Ａ～４４Ｃ及びメモリインタフェース回路４６と結合されている。メモリインタフェース回路４６はメモリ５６と結合され、インタフェース回路４４Ａはブリッジ４８と結合され、ブリッジ４８は周辺機器５２Ａ～５２Ｂと結合されている。

【００４２】

図示される実施形態においては、各プロセッサ３０Ａ～３０Ｂは、関連するｈＡＰＩＣ３２Ａ～３２Ｂ及び少なくとも１つの関連するｇＡＰＩＣ３４Ａ～３４Ｄを有している。

この実施形態においては、割り込みは、インテルコーポレーション（サンタクララ、カリフォルニア）(Intel Corporation (Santa Clara, CA)によって記述されているＡＰＩＣ規格に従ってホストハードウエア２０内で通信されてよい。その規格においては、各プロセッサは、プロセッサそれ自身、他のプロセッサ、内部ＡＰＩＣ割り込みソース、及び周辺機器に関連するＩＯＡＰＩＣからの割り込みを受信する関連するローカルＡＰＩＣを有している。ローカルＡＰＩＣは、係属中の複数の割り込みを優先付けし、割り込みがプロセッサ上で進行中である別の割り込みよりも高い優先度である場合且つ／又は割り込みがプロセッサのその時点のタスクよりも優先度が高い場合には、その割り込みをプロセッサへ送信する。

【００４３】

図２の実施形態においては、ｈＡＰＩＣ３２Ａ～３２Ｂは、プロセッサのホスト割り込み（即ち、ホストによって処理されるべき割り込み）に対するローカルＡＰＩＣであってよく、またｇＡＰＩＣ３４Ａ～３４Ｄは、プロセッサのゲスト割り込み（即ち、それぞれのプロセッサ３０Ａ～３０Ｂ上でアクティブなゲストによって処理されるべき割り込み）に対するローカルＡＰＩＣであってよい。ゲストは、そのゲストがその時点で当該プロセッサ上で実行中である場合（例えば、ＶＭＲＵＮ命令がそのゲストに対するそのプロセッサ上で実行完了しており、且つゲスト退出が発生しなかった場合）、又はそのゲストは既に退出しており且つＶＭＭ１８が実行中であるが、そのゲストはそのプロセッサ上で再度実行されることが期待されている場合に、そのプロセッサ上でアクティブであり得る。

【００４４】

ＶＭＭ１８がプロセッサ３０Ａ～３０Ｂ上のゲストをスケジューリングする場合、ＶＭＭ１８は、当該プロセッサ３０Ａ～３０ＢのｇＡＰＩＣ３４Ａ～３４Ｄをそのゲストに対応するｇＡＰＩＣ状態と共にロードしてよい。具体的には、所与のゲストは多重仮想ＣＰＵ（ｖＣＰＵ）を有していてよい。ＶＭＭ１８は、プロセッサ３０Ａ～３０Ｂ上で実行するゲストのｖＣＰＵをスケジューリングしてよく、そしてそのゲストの仮想マシン内の当該ｖＣＰＵに対する割り込み状態と共にｇＡＰＩＣ３４Ａ～３４Ｄをロードしてよい。また、ゲストがアクディブである間に信号を送られるゲスト（及びｖＣＰＵ）を対象としている任意の割り込みが、ｇＡＰＩＣ３４Ａ～３４Ｄによって捕獲されてよい。ｇＡＰＩＣ３４Ａ～３４Ｄは、前述したＡＰＩＣ規格に従ってゲストを割り込ませてよい。

【００４５】

所与のプロセッサ３０Ａ～３０Ｂに対するｈＡＰＩＣ３２Ａ～３２Ｂ及びｇＡＰＩＣ３４Ａ～３４Ｄは、そのプロセッサに対する任意のインタフェースを有していてよい。例えば、複数のローカルＡＰＩＣとそれらのそれぞれのプロセッサとの間で任意のインタフェースが用いられてよい。各ＡＰＩＣは、割り込みがサービスのために受け渡されている最中のプロセッサへ独立に信号を送るように構成されてよい。プロセッサがゲストを実行中であり且つゲスト割り込みが信号を送られている場合、プロセッサは、ゲストコードに割り込み且つゲストの仮想マシン内で正しい割り込みハンドラの実行を開始するように構成されてよい。従って、実施形態においては、ゲスト割り込みは、ホスト内の割り込みの受け渡しと同様の待ち時間を伴って受け渡され得る。プロセッサがゲストを実行中であり且つｈＡＰＩＣが割り込みに信号を送る場合には、プロセッサは、ゲストからＶＭＭ１８へ退出してホスト割り込みを処理するように構成されてよい。プロセッサがゲストを実行中でない場合には、ｇＡＰＩＣによって信号を送られている割り込みは、ゲストが再度実行されるまでプロセッサによってマスキングされてよい。プロセッサがゲストを実行中でなく且つｈＡＰＩＣが割り込み内へ信号を送る場合には、プロセッサは、ホスト実行に割り込み且つホスト割り込みハンドラへ分岐するように構成されてよい。

【００４６】

１つの実施形態においては、２つ以上のｇＡＰＩＣ３４Ａ～３４ＤがプロセッサＡ～３０Ｂ毎に含まれていてよい。各ｇＡＰＩＣ３４Ａ～３４Ｄは、異なるゲスト／ｖＣＰＵに対応するＡＰＩＣ状態を記憶してよい。そのような実施形態においては、各ｇＡＰＩＣ３４Ａ～３４Ｄは、ゲスト割り込みの信号をプロセッサへ送るときに各ｇＡＰＩＣ３４

Ａ～３４Ｄがどのゲストに対応しているかを識別するように構成されてよい（又はどのゲストがその時点で各ｇＡＰＩＣ３４Ａ～３４Ｄに割り当てられているかを識別する内部レジスタをプロセッサ３０Ａ～３０Ｂが有してよい）。ＶＭＭ１８が実行中でない場合にゲスト割り込みをマスキングするのと同様に、異なるゲストがその時点で実行中である場合には、プロセッサはゲスト割り込みをマスキングしてよい。代替的には、各ｇＡＰＩＣ３４Ａ～３４Ｄは、対応するゲストがスケジューリングされる場合にＶＭＭ１８によってアクティブにセットされ得るアクティブ表示を含んでいてよく、またｇＡＰＩＣ３４Ａ～３４Ｄは、そのゲスト割り込みへ、対応するゲストがアクティブの場合にのみ信号を送るように構成されてよい。プロセッサ３０Ａ～３０Ｂ毎に２つ以上のｇＡＰＩＣ３４Ａ～３４Ｄを含むことにより、多重ゲストがプロセッサ上で時間内に実行されるようにスケジューリングされている場合に、ｇＡＰＩＣ状態移動(state movement)の量を低減することができる。例えば、プロセッサ３０Ａ～３０Ｂ毎にＮ個のｇＡＰＩＣ３４Ａ～３４Ｄがある場合（Ｎは０より大きい整数）、Ｎ個までの異なるゲストは、任意のゲストに対してｇＡＰＩＣ状態が保存されることが必要になるであろう前に、実行のためにスケジューリングされてよい。プロセッサ３０Ａ～３０Ｂ毎に２つ以上のｇＡＰＩＣ３４Ａ～３４Ｄを実装する幾つかの実施形態においては、ｇＡＰＩＣ３４Ａ～３４Ｄは、割り込みメッセージが適切に受け入れられ且つ／又は記録されることを確実にする追加的な状態を含んでいてよい。例えば、ｇＡＰＩＣ３４Ａ～３４Ｄは、対応する仮想マシンが対応するプロセッサ３０Ａ～３０Ｂ上でその時点で実行中であるか否かを識別する「現在実行中(currently running)」表示を含んでいてよい（ＶＭＭ実行に対するサスペンション(suspension)にあるのとは対照的に、又は別の仮想マシンが実行中である間とは対照的に）。仮想マシンが実行中であることを現在実行中表示が示している場合には、ｇＡＰＩＣは割り込みメッセージを受け入れてよい。仮想マシンが実行中でないことを現在実行中表示が示している場合には、ｇＡＰＩＣは割り込みが受け入れられない旨の信号を送ってよい。代替的には、ｇＡＰＩＣは、割り込みが受け入れられない旨の信号をｇＡＰＩＣが送ることになるか否かを示す追加的な表示を含んでいてよい。そのような実施形態においては、現在実行中表示が現在実行中でないと示し且つ割り込みが受け入れられない旨の信号をｇＡＰＩＣが送ることになることを非容認表示が示している場合には、割り込みが受け入れられなかった旨の信号をｇＡＰＩＣが送ってよい。このような機能性は、実行中でないゲストに対して割り込みが受信されていることを検出するために用いられてよく、割り込みの対象にされているゲストをスケジューリングするために用いられてよい。

【００４７】

ｇＡＰＩＣ３４Ａ～３４Ｄは、ｈＡＰＩＣ３２Ａ～３２Ｂ内に含まれるハードウエアの少なくとも一部分を含んでいてよく、また、そのハードウエアの全てを含んでいてよい（例えばｈＡＰＩＣ３２Ａ～３２Ｂの複製であってよい）。ｇＡＰＩＣ３４Ａ～３４Ｄは、ｇＡＰＩＣ３４Ａ～３４Ｄがどのゲストに割り当てられることになるかを識別するために、ＡＰＩＣ状態に加えてゲスト識別子（ＩＤ）と共にプログラム可能であってよい。ゲストが多重ｖＣＰＵを含む場合には、物理ＡＰＩＣ＿ＩＤ及び論理ＡＰＩＣ＿ＩＤがゲスト内のｖＣＰＵを識別し得る。１つの実施形態においては、ゲストＩＤは、周辺デバイスに対してＩＯＭＭＵ４０によってサポートされるドメインＩＤと同じであってよい。代替的には、ゲストＩＤは別個に管理される資源であってよい。いずれの場合においても、ＶＭＭ１８は、ゲストＩＤをゲストに割り当ててよく、そして各ゲストに対してｇＡＰＩＣ３４Ａ～３４Ｄが適切にプログラムされることを確実にし得る。ｖＣＰＵ及び／又はｇＡＰＩＣ及び／又は当該対は、ここではより簡潔にゲスト内の割り込みの宛先と称されることがある。宛先は最終的には、割り込みをサービスすることになるｖＣＰＵであってよいが、対応するｇＡＰＩＣもまた、対応するプロセッサに関連しており且つ割り込みを記録することから、宛先と考えられてよい。

【００４８】

ｇＡＰＩＣ３４Ａ～３４Ｄ及びｈＡＰＩＣ３２Ａ～３２Ｂは、割り込みを受信するためにブリッジ３６と結合されている。ｇＡＰＩＣ３４Ａ～３４Ｄ及びｈＡＰＩＣ３２Ａ～３

２Ｂへ割り込みを運ぶために、任意のインタフェースが用いられてよい。例えば、ＡＰＩＣ割り込み運搬のために実装される任意のインタフェースが用いられ得る。１つの実施形態においては、割り込みメッセージを運ぶために、プロセッサ３０Ａ～３０Ｂへの／プロセッサ３０Ａ～３０Ｂからの他の動作とやりとりするために用いられるのと同じ通信メカニズム（例えば、プロセッサ３０Ａ～３０Ｂによって開始されるメモリ読み出し／書き込み動作、キャッシュコヒーレンシメンテナンスに対するプローブ、等）が用いられてよい。別の方法としては、ｇＡＰＩＣ３４Ａ～３４ＤとｈＡＰＩＣ３２Ａ～３２Ｂの結合が、プロセッサ３０Ａ～３０Ｂのブリッジ３６への結合と共有されてよい。代替的には、例えばｇＡＰＩＣ３４Ａ～３４Ｄ及びｈＡＰＩＣ３２Ａ～３２ＢがＡＰＩＣ「３線インタフェース(3 wire interface)」を用いている場合には、プロセッサ３０Ａ～３０Ｂは、ブリッジ３６への別個のパスを有していてよい。割り込みメッセージは、送信されている割り込み及び割り込みの宛先を識別する任意のインタフェース上での任意の通信であってよい。例えば、割り込みは関連する割り込みベクトルを有していてよく、そして割り込みベクトルは割り込みメッセージの一部であってよい。割り込みメッセージはまた、ゲストＩＤ及び宛先ＩＤ（例えば論理ＡＰＩＣ＿ＩＤ又は物理ＡＰＩＣ＿ＩＤ）を含んでいてよい。

【００４９】

ｈＡＰＩＣ３２Ａ～３２ＢはローカルＡＰＩＣと同様であってよい。例えば、ｈＡＰＩＣはホスト割り込みのためには用いられないから、ｈＡＰＩＣ３２Ａ～３２Ｂは、ゲスト識別のための追加的なハードウエアを含んでいなくてよい。代替的には、ｈＡＰＩＣ３２Ａ～３２Ｂは追加的なハードウエアを含んでいてよいが、追加的なハードウエアは、ｈＡＰＩＣ３２Ａ～３２Ｂがホスト割り込みのためのものであることを表示するようにプログラムされてよい。ブリッジ３６によってｈＡＰＩＣ３２Ａ～３２Ｂ及びｇＡＰＩＣ３４Ａ～３４Ｄへ送信される割り込みメッセージは、ゲスト割り込みをホスト割り込みに対抗するものとして識別してよく、またゲスト割り込みに対するゲストＩＤを含んでいてよい（又はゼロ若しくは全バイナリ１のような予約されたゲストＩＤを用いてホスト割り込みを表示してよい）。ｈＡＰＩＣ３２Ａ～３２Ｂは、ホスト割り込みとして識別される割り込みを受け入れるように構成されてよく（ホスト割り込みの物理ＡＰＩＣ＿ＩＤ又は論理ＡＰＩＣ＿ＩＤが対応するｈＡＰＩＣ＿ＩＤに整合する場合）、またｇＡＰＩＣ３４Ａ～３４Ｄは、それらのそれぞれのゲストに対するゲスト割り込みを受け入れるように構成されてよい（ゲストＩＤが整合する場合、及びゲスト割り込みの物理ＡＰＩＣ＿ＩＤ又は論理ＡＰＩＣ＿ＩＤが対応するｇＡＰＩＣ＿ＩＤに整合する場合）。

【００５０】

ｇＡＰＩＣ３４Ａ～３４Ｄはアクティブゲストに対する割り込みを管理してよいが、幾つかのゲストは非アクティブかもしれない（且つ又はゲスト割り込みによって対象とされ得る非アクティブｖＣＰＵを有しているかもしれない）。１つの実施形態においては、ゲスト割り込みマネージャ３８は、非アクティブゲストに対するゲスト割り込み状態を維持し且つアクティブゲストに対するｇＡＰＩＣが確実にそれらの割り込みを受信するように構成されてよい。

【００５１】

特に、１つの実施形態においては、ゲスト割り込みマネージャ３８は分散型割り込み受け渡しスキームを採用してよく、そのスキームにおいては、ゲスト割り込みマネージャ３８は、ブリッジ３６内で受信される各ゲスト割り込みを記録するように構成されてよく、またゲスト割り込みを各ｇＡＰＩＣ３４Ａ～３４Ｄへ送信するように構成されてよい。ｇＡＰＩＣ３４Ａ～３４Ｄが割り込みを受け入れる場合には、ゲスト割り込みよにって対象とされるゲストはアクティブである。どのｇＡＰＩＣ３４Ａ～３４Ｄも割り込みを受け入れない場合には、ゲスト割り込みによって対象とされるゲストは非アクティブである。

【００５２】

図示される実施形態においては、ゲスト割り込みマネージャ３８は、メモリ５６におけるｇＡＰＩＣ状態データ構造５８内のシステム５において規定されるゲストに対するｇＡＰＩＣ状態を維持するように構成されてよい。ｇＡＰＩＣ状態データ構造５８は、システ

ムにおいて規定される各ｇＡＰＩＣに対するｇＡＰＩＣ状態エントリ（例えばシステム内の各ゲスト１０Ａ～１０Ｎにおける各ｖＣＰＵに対する１つのエントリ）を含んでいてよい。ｇＡＰＩＣは、それがシステム内のアクティブゲスト又は非アクティブゲストのいずれかに関連する場合に、システムにおいて規定されてよい。従って、ゲスト割り込みを受信することに応答して、ゲスト割り込みマネージャ３８は、当該割り込みによって対象とされるゲスト／ｖＣＰＵに対してｇＡＰＩＣ状態データ構造５８内のｇＡＰＩＣ状態を更新するように構成されてよい。ゲスト割り込みマネージャ３８は、１つの実施形態においては、ゲストがアクティブか否かに関係なくｇＡＰＩＣ状態を更新するように構成されてよい。１つの対象より多くを有するマルチキャスト割り込み及びブロードキャスト割り込みに対しては、ゲスト割り込みマネージャ３８は、各割り込み宛先に対してｇＡＰＩＣ状態データ構造５８内のｇＡＰＩＣ状態を更新するように構成されてよい。代替的には、ゲスト割り込みマネージャ３８は、これらの多重宛先割り込みに対してＶＭＭ１８を頼るように構成されてよい。ゲスト割り込みマネージャ３８は、そのような実施形態においては、ＶＭＭ１８にアクセス可能なメモリロケーション内に割り込みを記録するように構成されてよく、またＶＭＭ１８へ信号を送ってメッセージを処理するように構成されてよい。

【００５３】

幾つかの実施形態においては、ゲスト割り込みマネージャ３８は、ゲストＩＤ及び／又はゲスト割り込みメッセージ内の他の情報に直接的に応答して、ｇＡＰＩＣ状態エントリをｇＡＰＩＣ状態データ構造５８内に配置するように構成されてよい。他の実施形態においては、ｇＡＰＩＣ状態データ構造５８における柔軟性を提供するために且つ／又はメモリ空間を節約するために、ゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態マッピングテーブル６０を用いてｇＡＰＩＣ状態エントリをｇＡＰＩＣ状態データ構造５８内に配置するように構成されてよい。ｇＡＰＩＣ状態データ構造５８の種々の実施形態及びマッピングテーブル６０（幾つかの実施形態に対して）が図１０～１３に示されており、また更に詳細に後で論じられる。従って、ゲスト割り込みに応答して、ゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態マッピングテーブル６０を用いてｇＡＰＩＣ状態エントリを配置するように、そしてｇＡＰＩＣ状態エントリを更新して割り込みを記録するように構成されてよい。

【００５４】

１つの実施形態では、ｇＡＰＩＣ状態データ構造５８は、ｇＡＰＩＣ状態のサブセット(subset)を記憶してよい。サブセットは、ハードウエア２０（例えばＩＯＭＭＵ４０と併用されるゲスト割り込みマネージャ３８）によって追跡されるｇＡＰＩＣ状態であってよい。より特定的には、サブセットは、対応するゲストが非アクティブである間に変化し得るｇＡＰＩＣ状態の一部分であってよい。例えば、１つの実施形態では、周辺機器５２Ａ～５２Ｂは、対応するゲストが非アクティブである間に割り込みを信号で伝えてよく、これにより対応する割り込み要求はｇＡＰＩＣによって捕獲され得る。割り込み要求はｇＡＰＩＣ状態データ構造５８内で追跡されてよい。他のｇＡＰＩＣ状態は、どの割り込みがプロセッサによって稼動中(in-service)であるか、プロセッサのタスク優先度、等を追跡してよい。これらの値は、ゲストがアクティブである場合にのみ変化し得る。実施形態においては、ゲストが非アクティブである場合に変化しないであろうｇＡＰＩＣ状態は、図２にＶＭＭ管理ｇＡＰＩＣ状態データ構造６８として示される１つ以上の他のデータ構造を用いてＶＭＭ１８によって追跡されてよい。ＶＭＭ１８は、ＶＭＭ管理ｇＡＰＩＣ状態データ構造６８とシステム内でアクティブ化しているゲスト及び非アクティブ化しているゲストの一部としてのｇＡＰＩＣ３４Ａ～３４Ｄとの間で、状態を転送してよい。

【００５５】

図示される実施形態においては、ｇＡＰＩＣ状態マッピングテーブル６０及びｇＡＰＩＣ状態データ構造５８はメモリ５６内に記憶されているとして示されているが、一方又は両方の部分部分は、ゲスト割り込みマネージャ３８及び／又はブリッジ３６へアクセス可能なキャッシュによってキャッシュされてよい。加えて又は代替的には、１つ以上のｇＡＰＩＣ状態エントリのための専用のメモリがブリッジ３６内に実装されてよい。専用のメ

モリは、ｇＡＰＩＣ３４Ａ～３４Ｄ内へ及びｇＡＰＩＣ３４Ａ～３４Ｄから外へ速やかに切り換えられ得る一連の「高速の(fast)」ｇＡＰＩＣ状態を記憶していてよい。他のｇＡＰＩＣ状態は、メモリ５６内でより低速にアクセス可能であってよい。幾つかの実施形態においては、高速のｇＡＰＩＣ状態切り換えはゲスト割り込みマネージャ３８によって処理されてよい一方で、より低速のｇＡＰＩＣ状態切り換えはＶＭＭ１８によって処理されてよい。

【００５６】

ＡＰＩＣ割り込みメカニズムにおいては、各プロセッサ（そのローカルＡＰＩＣを介して）は、物理ＡＰＩＣ＿ＩＤ及び論理ＡＰＩＣ＿ＩＤを有していてよい。物理ＡＰＩＣ＿ＩＤはＡＰＩＣ＿ＩＤレジスタ内に記憶される。物理ＡＰＩＣ＿ＩＤは、物理受け渡しモード割り込み(physical delivery mode interrupt)によって表示される物理ＡＰＩＣ＿ＩＤと１対１を原則として整合する。論理ＡＰＩＣ＿ＩＤは、論理宛先レジスタとしてローカルＡＰＩＣ内に記憶される。論理ＡＰＩＣ＿ＩＤはクラスタＩＤ及びローカルＡＰＩＣ＿ＩＤを有しており、ここでローカルＡＰＩＣ＿ＩＤはワンホットベクトル(one-hot vector)である。論理受け渡しモード割り込みは、割り込みをクラスタ内の１つ以上のローカルＡＰＩＣへ受け渡すために、ワンホットベクトル内に任意の一連のビットを含んでいてよい。従って、論理ＡＰＩＣ＿ＩＤを整合させることは、クラスタＩＤを比較することと、ローカルＡＰＩＣ内のワンホットベクトルの一連のビットと同じ位置でのローカルＡＰＩＣ＿ＩＤベクトル内の一連のビットを検出することとを含んでいてよい。別の方法では、論理受け渡しモード割り込みにおけるローカルＡＰＩＣ＿ＩＤベクトルが、ローカルＡＰＩＣのローカルＡＰＩＣ＿ＩＤベクトルと論理積を取られて(logically ANDed)よく、結果が非ゼロであり且つクラスタＩＤが整合する場合には、ローカルＡＰＩＣは論理割り込みの対象である。論理ＡＰＩＣ＿ＩＤは、より簡潔にここでは論理ＩＤと称されることがあり、また同様に物理ＡＰＩＣ＿ＩＤは、より簡潔にここでは物理ＩＤと称されることがある。割り込みに関連して与えられるＩＤ（論理又は物理）は、割り込みの宛先ＩＤと称されることがある。割り込みに対する対応する受け渡しモードは、割り込みの宛先ＩＤを識別することができる。

【００５７】

ｇＡＰＩＣ３４Ａ～３４Ｄは、物理受け渡しモード及び論理受け渡しモードの両方を同様にサポートしてよい。上で強調したようなモードに従う割り込みメッセージ内でＡＰＩＣ＿ＩＤを整合させることに加えて、ｇＡＰＩＣ３４Ａ～３４Ｄは、割り込みメッセージ内のゲストＩＤをｇＡＰＩＣ内のゲストＩＤと整合させてよい。

【００５８】

ＩＯＭＭＵ４０は、Ｉ／Ｏが開始したメモリ動作（例えば、周辺機器５２Ａ～５２Ｂをソースとする又は周辺機器５２Ａ～５２Ｂに代わるＤＭＡ制御器によるメモリ読み出し／書き込み動作）のために仮想アドレスから物理アドレスへのマッピングを実行するように構成されてよい。翻訳動作の一部として、ＩＯＭＭＵ４０は、デバイステーブル６２及び随意的に割り込みリダイレクトテーブル６４へアクセスするように構成されてよい。デバイステーブル６２は、各周辺機器５２Ａ～５２Ｂに対するエントリを含んでいてよい（また、複数の周辺機器が結合される周辺機器インタフェース上の２つ以上の識別子を含む周辺機器に対する多重エントリを含んでいてよい）。デバイステーブル６２は、メモリ読み出し／書き込み動作のメモリアドレスを翻訳するためのＩ／Ｏページテーブル（図示せず）へのページテーブルポインタを含んでいてよく、また割り込みリダイレクトテーブル６４へのポインタを含んでいてよい。幾つかの実施形態においては、デバイステーブル６２は、ゲストに割り当てられる周辺機器に対するゲストＩＤを記憶してよい。１つの実施形態においては、ゲストＩＤは、ＩＯＭＭＵ４０においてデバイスアクセス保護のために用いられるドメインＩＤと同じであってよい。代替的には、ゲストＩＤは別々に割り当てられてよい。実施形態においては、デバイステーブル６２はまた、ｇＡＰＩＣ状態マッピングテーブル６０（用いられている場合）へのポインタ、又はｇＡＰＩＣ状態データ構造５８へのポインタを記憶してよい。別の実施形態においては、ゲストＩＤ及び／又はテーブ

ル６０／データ構造５８へのポインタは、割り込みリダイレクトテーブル６４内に記憶されてよい。割り込みリダイレクトテーブル６４は、割り込みをその元の宛先及び／又は割り込みベクトルから新たな宛先及び／又は割り込みベクトルへリダイレクトするために用いられてよい。この開示の残りにおける簡潔さのために、ゲストＩＤがデバイステーブル６２からのドメインＩＤであり且つマッピングテーブル６０及び／又はｇＡＰＩＣ状態データ構造５８へのポインタがデバイステーブル６２内に記憶されている実施形態が用いられることになる。しかし、この開示の残りにおける実施形態は、上で議論したように多くの場合に修正されてよい。

【００５９】

他の実施形態においては、ゲスト割り込みマネージャ３８は設けられていなくてよい。そのような構成は、例えば、ゲストが１つのプロセッサ３０Ａ～３０Ｂから他へ移動させられる(are migrated)場合においてＶＭＭ１８がデバイステーブル６２及び／又は割り込みリダイレクトテーブル６４を更新するとき、及び非アクティブゲストに代わってプロセッサ３０Ａ～３０Ｂが割り込みを受信するように設けられている場合（所望に応じて、メモリ５６内のｇＡＰＩＣ状態データ構造５８を更新し且つ／又は割り込みをサービスするために）に可能である。

【００６０】

メモリ制御器４２は、プロセッサ３０Ａ～３０Ｂによって発行されるメモリ動作（例えば命令フェッチ、ロード／ストアデータアクセス、翻訳のためのプロセッサページテーブルアクセス、等）、ゲスト割り込みマネージャ３８からのメモリ動作（例えばｇＡＰＩＣ状態データ構造５８及び／又はｇＡＰＩＣ状態マッピングテーブル６０を読み出し／更新するための）、ＩＯＭＭＵ４０（例えばＩ／Ｏページテーブル、デバイステーブル６２、及び割り込みリダイレクトテーブル６４にアクセスするための）、及びインタフェース回路４４Ａ～４４Ｃから受信するメモリ動作を受信するように結合されていてよい（幾つかの実施形態においては）。メモリ制御器４２は、メモリ動作をオーダし、そしてメモリ５６と通信してそれらメモリ動作を実行するように構成されてよい。メモリインタフェース回路４６は、メモリ５６への物理レベルアクセスを実行してよい。

【００６１】

メモリ５６は任意の種類のメモリを備えていてよい。例えば、メモリ５６は、ダイナミックランダムアクセスメモリ（ＤＲＡＭ）、例えば同期ＤＲＡＭ（ＳＤＲＡＭ）、ダブルデータレート（ＤＤＲ、ＤＤＲ２、ＤＤＲ３、等）ＳＤＲＡＭ、ＲＡＭＢＵＳ＿ＤＲＡＭ、スタティックＲＡＭ、等を備えていてよい。メモリ５６は、多重メモリチップを備えた１つ以上のメモリモジュール、例えばシングルインラインメモリモジュール（ＳＩＭＭ）、デュアルインラインメモリモジュール（ＤＩＭＭ）等を含んでいてよい。

【００６２】

この実施形態では、ゲスト割り込みマネージャ３８、ＩＯＭＭＵ４０、及びメモリ制御器４２を含むことに加えて、ブリッジ３６はまた、プロセッサ３０Ａ～３０Ｂ、ｈＡＰＩＣ３２Ａ～３２Ｂ、ｇＡＰＩＣ３４Ａ～３４Ｄ、及びインタフェース回路４４Ａ～４４Ｃに結合される回路の間で通信する他の通信機能を含んでいてよい。例えば、図示される実施形態においては、別のブリッジ４８がインタフェース回路４４Ａに結合されていてよく、そしてブリッジ４８は、インタフェース回路４４Ａによって用いられるプロトコルと周辺機器５２Ａ～５２Ｂによって用いられるプロトコルとの間の通信をブリッジするように構成されていてよい。１つの実施形態においては、インタフェース回路４４Ａ～４４Ｃは、例えば上述したＨＴインタフェースを実装していてよく、またブリッジ４８は、ＨＴから別のインタフェース、例えばＰＣＩエクスプレス(PCI Express)（ＰＣＩｅ）インタフェースへブリッジしてよい。そのような実施形態においては、周辺機器５２Ａ～５２ＢはＰＣＩｅデバイスであってよい。ブリッジ４８はまた、他のインタフェースへブリッジするように構成されていてよく、あるいは別のブリッジがブリッジ４８に結合されて他のインタフェースへブリッジしてもよい。任意の単一又は複数の周辺機器インタフェースが用いられてよい。加えて、周辺機器５２Ａ～５２Ｂは、ＨＴインタフェースと直接的に結合

するように構成されるＨＴ周辺機器を備えていてよい。そのような周辺機器は、ブリッジ４８を必要としないであろう。

【００６３】

１つの実施形態においては、ブリッジ４８及び／又は１つ以上の周辺機器５２Ａ～５２Ｂは、ＩＯＡＰＩＣ（図２においては５０及び５４）を含んでいてよい。ＩＯＡＰＩＣは、周辺機器からの割り込み要求を受信することと、割り込み要求をｈＡＰＩＣ３２Ａ～３２Ｂ及びゲスト割り込みマネージャ３８へ送信するために割り込みメッセージを形成することとに関与していてよい（ｇＡＰＩＣ３４Ａ～３４Ｄへの送信及び／又はメモリ内のｇＡＰＩＣ状態データ構造５８内への記録のために）。

【００６４】

上述したように、１つの実施形態においては、インタフェース回路４４Ａ～４４ＣはＨＴインタフェース上で通信するように構成されてよい。インタフェース回路４４Ａ～４４Ｃは、ＨＴを用いて周辺デバイス／ブリッジと通信するように構成されてよい。また、幾つかの実施形態においては、インタフェース回路４４Ａ～４４Ｃは、プロセッサを伴う他のノード、ｈＡＰＩＣ、ｇＡＰＩＣ、等と結合されるように構成されてよい。そのような実施形態においては、ブリッジ３６は、前述した回路に加えてコヒーレンス管理回路を含んでいてよい。

【００６５】

プロセッサ３０Ａ～３０Ｂは、任意の命令セットアーキテクチャを実装していてよく、そして命令セットアーキテクチャにおいて定義される命令を実行するように構成されてよい。プロセッサ３０Ａ～３０Ｂは、任意のマイクロアーキテクチャ、例えば、スーパーパイプラインにされたもの(superpipelined)、スーパースケーラ(superscalar)、及び／又はこれらの組み合わせ、順序内(in-order)又は順序外(out-of-order)の実行、投機的実行(speculative execution)、等を含んでいてよい。プロセッサ３０Ａ～３０Ｂは、所望に応じてマイクロコーディング技術を実装していてよいし実装していなくてもよい。

【００６６】

周辺機器５２Ａ～５２Ｂは任意の種類の周辺デバイスを備えていてよい。周辺機器５２Ａ～５２Ｂは、記憶デバイス、例えば、磁気的な、ソリッドステートの、又は光学的なディスクドライブ、フラッシュメモリのような不揮発性メモリデバイス等を含んでいてよい。周辺機器５２Ａ～５２Ｂは、Ｉ／Ｏデバイス、例えば、ユーザＩ／Ｏデバイス（キーボード、マウス、ディスプレイ、音声入力、等）、ネットワークデバイス、ユニバーサルシリアルバス（ＵＳＢ）又はファイヤワイヤ(Firewire)のような外部インタフェースデバイス等を含んでいてよい。

【００６７】

図示される実施形態においては、プロセッサ３０Ａ～３０Ｂ、ブリッジ３６、ｈＡＰＩＣ３２Ａ～３２Ｂ、ｇＡＰＩＣ３４Ａ～３４Ｄ、インタフェース回路４４Ａ～４４Ｃ、及びメモリインタフェース回路４６は、集積回路６６として単一の半導体基板上に集積化されてよい。他の実施形態は、所望に応じて異なる量の集積化及び個別回路を実装してよい。尚、図２においては種々の数のコンポーネント、例えばプロセッサ、ｈＡＰＩＣ、ｇＡＰＩＣ、インタフェース回路、周辺機器、ブリッジ等が図示されているが、他の実施形態は、所望に応じて１つ以上の任意の数の各コンポーネントを実装してよい。

【００６８】

他の実施形態においては、ＩＯＭＭＵ４０及びゲスト割り込みマネージャ３８の位置は変わってよい。例えば、一方又は両方は、ブリッジ４８内、周辺機器５２Ａ～５２Ｂ内、ブリッジに結合される別のブリッジ内、等にあってよい。

【００６９】

図示される実施形態においては、図２に示されるように、各ｇＡＰＩＣ３４Ａ～３４Ｄ及びｈＡＰＩＣ３２Ａ～３２Ｂは、特定のプロセッサ３０Ａ～３０Ｂに付随している。従って、この実施形態においては、所与の割り込み制御器は対応するプロセッサ３０Ａ～３０Ｂに専用である。より具体的には、図２においては、ｈＡＰＩＣ３２Ａ並びにｇＡＰＩ

Ｃ３４Ａ及び３４Ｃはプロセッサ３０Ａの専用であり、そしてｈＡＰＩＣ３２Ｂ並びにｇＡＰＩＣ３４Ｂ及び３４Ｄはプロセッサ３０Ｂの専用である。割り込み制御器は、対応するプロセッサへ任意の方法で割り込みの信号を送ってよい。一般的に、信号を送ること(signaling)は、割り込みが必要とされていることを表示してよい。信号を送ることは割り込みベクトルを含んでいてよく、即ち割り込みベクトルは割り込みが受け渡された後に実行されるソフトウエアによって読まれてよい。割り込みを受け渡すことは、実施形態においては、プロセッサに信号を送ること及びプロセッサが割り込みを受け入れることを参照してよい。割り込みをサービスすることは、割り込みベクトルに関連する割り込みサービスルーチンを実行して、割り込みしているデバイスによって必要とされる動作を実行することを参照してよい。

【００７０】

次に図３を参照すると、１つの実施形態に対する周辺機器からｇＡＰＩＣへの割り込みの進行を表すブロック図が示されてる。他のプロセッサからの割り込み（プロセッサ間割り込み(interprocessor interrupts)、又はＩＰＩ）もまた、ゲスト割り込みマネージャ３８へ送信されてよく、そしてその時点以降は図３と同様に取り扱われてよい。代替的には、ＩＰＩを開始しているプロセッサからのＩＰＩを受信するｇＡＰＩＣは、更新をゲスト割り込みマネージャ３８へ送信してよく（受信しているゲストが非アクティブである場合にそのゲストに対してｇＡＰＩＣ状態を更新するために）、またＩＰＩ（ゲストＩＤを含む）を他のｇＡＰＩＣへ送信してもよい。

【００７１】

図示される実施形態においては、周辺機器５２Ａは、割り込みが望まれていることを決定する。周辺機器５２Ａ内のＩＯＡＰＩＣ５４（図２参照）は、周辺機器５２Ａの動作に応答して割り込みメッセージを生成してよい。具体的には、ＩＯＡＰＩＣ５４は、望まれている割り込み（例えば、周辺機器５２Ａによって必要とされるサービスに基いて、周辺機器５２Ａが多重機能を実装している場合に特定の機能が割り込みの信号を送ること、等）に対応する割り込みベクトルを生成してよい。割り込みベクトルは、割り込み通信の一部であり、そして割り込みソースを識別し、割り込みを優先付ける、等のためにソフトウエアによって用いられてよい。幾つかの場合には、割り込みベクトルはＩＯＭＭＵ４０によって再マッピングされてよく、そのため図３においては、割り込みベクトルは「元のベクトル(original vector)」として示されている。周辺機器５２Ａは、割り込みメッセージをＩＯＭＭＵ４０へ送信してよい（矢印Ａ）。この実施形態においては、割り込みは、例えばＰＣＩｅ規格において定義されるようなメッセージが信号で送られた割り込み(message-signalled interrupt)（ＭＳＩ）の形態で送信されてよい。他の実施形態は、割り込みを任意の望ましい方法で送信してよい。一般的に、送信は、割り込み、その受け渡しモード（例えば論理又は物理）、及び割り込みの宛先ＩＤ（ＤｅｓｔＩＤ）を識別してよい。

【００７２】

ＩＯＭＭＵ４０はＭＳＩを受信してよい。ＭＳＩは周辺機器の識別子を含んでいる。例えば、ＰＣＩプログラミングモデルを実装しているインタフェースは、各デバイスをバス番号及び当該バスのデバイス番号で識別することができる（階層的な且つ／又は並列な形態にあるシステム内に多重ＰＣＩインタフェースが存在することを可能にする）。デバイスは多重「機能(functions)」を有していてよく、多重機能は、物理デバイス上の別々の複数の仮想デバイスであってよく、あるいはデバイス上での複数の動作の分割(partitioning)であってよい。識別子はまた、機能番号を含んでいてよい。従って、この実施形態においては、識別子は、バスデバイス機能(Bus-Device-Function)又はＢＤＦと称されることがある。ＩＯＭＭＵ４０は、ＢＤＦを用いてデバイステーブル６２内へ索引付ける(index)ことができ（矢印Ｂ）、また周辺機器５２Ａに対応するデバイステーブルエントリを識別することができる。エントリは、幾つかの実施形態においては、ゲストＩＤと、ｇＡＰＩＣ状態マッピングテーブル６０又はｇＡＰＩＣ状態データ構造５８へのポインタとを含んでいてよい（矢印Ｃ）。この実施形態においては、デバイステーブルエントリはまた

、デバイスに対応する割り込みリダイレクトテーブル６４を識別し得る割り込みリダイレクトテーブルポインタ（ＩＲＴＰ）を含んでいてよい（矢印Ｃ１）。割り込みリダイレクトテーブル６４は、元の割り込みベクトルによって索引付けられてよく、そして出力ベクトル及び宛先ＩＤ（ＤｅｓｔＩＤ、例えば論理又は物理ＡＰＩＣ＿ＩＤ）を割り込みに対して提供してよい（矢印Ｃ２）。

【００７３】

図３はＭＳＩがベクトル４２、ゲストＩＤ９９へ再マッピングされる例を示している。再マッピングはゲストＩＤを付加することを含んでいてよく、そしてベクトルはまた、割り込みリダイレクトテーブル６４が用いられている場合に変化させられてよい。そうでない場合には、ＭＳＩからの元の割り込みベクトルは、割り込みメッセージ内で提供されてよい。割り込みベクトル４２及びゲストＩＤ９９の具体例が用いられている図３における点は、大括弧、即ち［　］で囲まれて図示されている。

【００７４】

ＩＯＭＭＵ４０は、ゲストＩＤ（例えばこの例では９９）を含めて割り込みメッセージをゲスト割り込みマネージャ３８へ送信してよい。割り込みメッセージはまた、割り込みベクトル（例えばこの例では４２）及び宛先ＩＤを含む。割り込みメッセージはまた、ｇＡＰＩＣ状態マッピングテーブル６０又はｇＡＰＩＣ状態データ構造５８へのポインタを含んでいてよい（矢印Ｄ）。

【００７５】

ｇＡＰＩＣ状態マッピングテーブル６０を実装している実施形態においては、ゲスト割り込みマネージャ３８は、ポインタ並びに場合によってはゲストＩＤ及び／又は宛先ＩＤのような他の情報を用いてｇＡＰＩＣ状態ポインタをｇＡＰＩＣ状態マッピングテーブル６０内に配置してよい（矢印Ｅ１、そしてゲスト割り込みマネージャ３８への戻りのポインタは矢印Ｅ２で示されている）。ｇＡＰＩＣ状態ポインタは、ｇＡＰＩＣ状態データ構造５８内のｇＡＰＩＣ状態エントリを識別してよく、またゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態ポインタを用いてｇＡＰＩＣ状態データ構造５８内でのｇＡＰＩＣ状態更新を実行してよい（矢印Ｅ）。この実施形態においては、ｇＡＰＩＣ状態更新は、ベクトル４２に対応する割り込み要求レジスタ内のビットをセットしてよい。割り込み要求レジスタ（ＩＲＲ）は、図４に関して後で更に詳細に説明される。

【００７６】

１つの実施形態においては、ｇＡＰＩＣ状態５８への更新はアトミック(atomic)であってよい。例えば、ゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態エントリにおける割り込み要求レジスタのその時点での状態内へセットされつつある割り込み要求についてアトミックに論理和を取る(atomically ORs)アトミックＯＲトランザクションを生成してよい。アトミックな動作は、動作が多重ステップとして実装されている場合であっても、一単位として効果的に実行される動作であってよい。アトミックに更新されつつあるロケーションへアクセスしようとしているオブザーバは、アトミックな更新に先立つ値又はアトミックな更新の後の値のいずれかを受信するが、中間値は受信しないであろう。アトミックに更新されつつあるロケーションへアクセスしようとしているオブザーバは、その更新をアトミックな動作の前又はアトミックな動作が完了した後のいずれかに実行するが、アトミックな動作の間には実行しない。この実施形態はアトミックＯＲを実装しているが、他の実施形態は、より一般的なアトミックな更新動作を実装してよい。例えば、アトミックな更新は、修正されるべきでない対象のビットを識別するＡＮＤマスクを含んでいてよく、またどのビットが論理和を取られるべきかを識別するＯＲマスクを含んでいてよい。他の実装もまた可能である。例えば、比較及び交換実装(compare and swap implementation)が用いられてよく、この実装においては、メモリロケーションからの元の値が読まれ、そして比較及び交換動作が元の値及び論理和を取られた新たな値に対して実行される。比較が失敗した場合には、プロセスは繰り返されてよい（新たな元の値を読み、そして比較及び交換を実行する）。必要であれば、ループのフェイルアウト(fail out)に対してバックオフ(backoff)及び／又はタイムアウトのメカニズムが用いられてよい。

【００７７】
ゲスト割り込みマネージャ３８はまた、割り込みベクトル、ゲストＩＤ、及び宛先ＩＤを含む割り込みメッセージをｇＡＰＩＣ３４Ａ～３４Ｄへブロードキャストしてよい（矢印Ｆ）。ｇＡＰＩＣの１つ（図３ではｇＡＰＩＣ３４Ａ）は、ゲストＩＤ９９と宛先ＩＤに整合する論理又は物理ＡＰＩＣ＿ＩＤとを有していてよく、従って、ｇＡＰＩＣ３４Ａは、それが割り込みメッセージを受け入れたことを表示する承認(acknowledgement)（Ａｃｋ）で割り込みメッセージに応答してよい（矢印Ｇ）。ｇＡＰＩＣ３４Ａはまた、その割り込み要求レジスタを更新して、この実施形態ではベクトル４２に対応するビットをセットしてよい。割り込みが任意の進行中の割り込み（もしあれば）及び／又はプロセッサのタスク優先度よりも高い場合には、ｇＡＰＩＣ３４Ａはまた、割り込みの信号をプロセッサ３０Ａへ送ってよい。他のｇＡＰＩＣ３４Ｂ～３４Ｄは、ブロードキャスト割り込みメッセージに応答してよいが、それらは割り込みの対象ではないから、受け入れを承認しなくてよい（矢印Ｈ）。論理割り込みに対しては、論理割り込みが多重対象を識別している場合には、２つ以上の承認があってよい。

【００７８】
上述のメカニズムを用いると、ゲスト割り込みマネージャ３８は、どのｇＡＰＩＣ３４Ａ～３４Ｄがどのゲストに割り当てられているかを「知っている(aware)」必要がない。どのｇＡＰＩＣ３４Ａ～３４Ｄがどのゲストに割り当てられているか、及びどれが対象となるｇＡＰＩＣに対してのみ割り込みを送信するかをゲスト割り込みマネージャ３８が追跡する他の実施形態が検討される。ゲスト割り込みマネージャ３８は、ｇＡＰＩＣを自動的に追跡してよく、あるいはｇＡＰＩＣが他のゲストに再割り当てされる毎にＶＭＭ１８によってプログラムされてよい。そのような実施形態においては、ゲスト割り込みマネージャ３８は、対象となるｇＡＰＩＣにのみ割り込みメッセージを送信してよい。

【００７９】
割り込みのｈＡＰＩＣ３２Ａ～３２Ｂへの送信は、通常のＡＰＩＣの方法で実行されてよい。具体的には、割り込みは、ゲスト割り込みマネージャ３８による場合には動作させられなくてよいが、実施形態においては、他の点については図３の動作と同様であってよい。

【００８０】
尚、ゲスト割り込みマネージャ３８はここではブロックとして図示されているが、ゲスト割り込みマネージャ３８を実装している回路が配布されてもよい。例えば、実施形態においては、ポインタを受信し、ｇＡＰＩＣ状態マッピングテーブル６０を随意的に処理し、そしてｇＡＰＩＣ状態データ構造５８に対する更新を生じさせるゲスト割り込みマネージャ３８の一部分がＩＯＭＭＵ４０内に含まれていてよく、この場合、ＩＯＭＭＵ４０は、ｇＡＰＩＣ状態データ構造５８に対するアトミックＯＲ及びｇＡＰＩＣ３４Ａ～３４Ｄへ送信されるべき割り込みメッセージを送信する。１つ以上の物理ロケーションにおけるゲスト割り込みマネージャ３８の任意の実装が用いられてよい。

【００８１】
次に図４を参照すると、ｇＡＰＩＣ３４Ａの１つの実施形態のブロック図が示されている。他のｇＡＰＩＣ３４Ｂ～３４Ｄは同様であってよい。図４の実施形態においては、ｇＡＰＩＣ３４Ａは、割り込み要求レジスタ（ＩＲＲ）７０、割り込みサービスレジスタ（ＩＳＲ）７２、トリガモードレジスタ（ＴＭＲ）７４、タスク優先度レジスタ（ＴＰＲ）７６、制御ユニット７８、物理ＩＤレジスタ８０、論理ＩＤレジスタ８２、ゲストＩＤレジスタ８４、及び随意的な他のＡＰＩＣ状態８６を含む。制御ユニット７８は、ＩＲＲ７０、ＩＳＲ７２、ＴＭＲ７４、ＴＰＲ７６、物理ＩＤレジスタ８０、論理ＩＤレジスタ８２、ゲストＩＤレジスタ８４、及び他のＡＰＩＣ状態８６と結合される。加えて、制御ユニット７８は、割り込みを受信するためにゲスト割り込みマネージャ３８と通信するように結合され、またプロセッサ３０Ａと通信するためにプロセッサインタフェースに結合される。

【００８２】

ゲスト割り込みマネージャ３８からの割り込みメッセージを受信することに応答して、制御ユニット７８は、ｇＡＰＩＣ３４Ａに対応するゲストを割り込みが対象としている場合に、割り込みをＩＲＲ７０内に書き込むように構成されてよい。ＩＲＲ内の割り込み要求の位置は、割り込みベクトルに対応する。ＩＲＲは「固定された(fixed)」割り込みを追跡してよい。他の割り込み種類としては、非マスク可能割り込み(non-maskable interrupt)（ＮＭＩ）、システム管理割り込み（ＳＭＩ）、レガシ外部割込み(legacy external interrupt)（ｅｘｔＩＮＴ）、等を挙げることができる。これらの割り込みは、他のＡＰＩＣ状態８６の一部として取り扱われてよい。

【００８３】

１つの実施形態においては、割り込みメッセージはまた、各割り込みに対するトリガモード（レベル又はエッジ）を含んでいてよい。ＴＭＲ７４は、どちらのトリガモードを割り込みに適用するかの表示を記憶してよい。例えば、エッジトリガされる割り込みはＴＭＲ７４内のバイナリ０によって表されてよく、またレベルトリガされるものはバイナリ１によって表されてよい。他の実施形態においては、エッジトリガされる割り込みのみがｇＡＰＩＣ３４Ａ内でサポートされてよく、またＴＭＲ７４（及びｇＡＰＩＣ状態データ構造５８内のそのコピー）は除かれてよい。別の実施形態においては、ＴＭＲ７４は、仮想レベル敏感な割り込みをＶＭＭ１８が記録することを可能にするために再度目的を持たされてよい。ＶＭＭ１８は、対応する割り込みベクトルに対してプロセッサ３０Ａによって割り込みの終点の信号が送られた場合に、プロセッサ３０ＡがＶＭＭ１８から退出することを表示する種々のビットをＴＭＲ７４内にセットしてよい。

【００８４】

固定された割り込みに対しては、ｇＡＰＩＣ３４Ａは、割り込み要求がプロセッサへ受け渡されるべきかを決定するために、割り込み要求及びインサービス割り込みを優先付けるように構成されてよい。概して、最も高い優先度の割り込み要求が最も高い優先度のインサービス割り込み（プロセッサがそのソフトウエア実行に割り込んだ結果、割り込みに対応する割り込みハンドラを実行した場合には、割り込みはインサービスである）よりも高い優先度である場合には、制御ユニット７８は、要求された割り込みをプロセッサ３０Ａへ受け渡すように構成されてよい。また、ＴＰＲ７６は、プロセッサ３０Ａによって受け入れられている最も低い優先度レベルの割り込みを確立するために、ソフトウエアによってプログラムされてよい。制御ユニット７８は、最も高い優先度の割り込み要求を、それが最も高い優先度のインサービス割り込みよりも高い優先度である場合、及びそれがＴＰＲ７６において表示されている優先度よりも高い場合に、受け渡すように構成されてよい。

【００８５】

プロセッサ３０Ａが割り込みを取る場合には、プロセッサは、割り込み承認コマンドでｇＡＰＩＣ３４Ａに対して応答してよい。制御ユニット７８は、最も高い優先度の割り込み要求をＩＲＲ７０から取り除くと共に当該割り込みをインサービスとしてＩＳＲ７２内に記録するように構成されてよい。ＩＳＲ内の割り込みに対応するインサービス表示の位置は、当該割り込みの割り込みベクトルに対応していてよい。プロセッサ３０Ａは、単一又は複数の割り込みサービスルーチンを実行して割り込みをサービスしてよい。割り込みサービスルーチンは、割り込みサービスが完了した旨の信号を送るための、ｇＡＰＩＣ３４Ａへの割り込み終点(end of interrupt)（ＥＯＩ）コマンドで終了してよい。制御ユニット７８は、ＥＯＩコマンドに応答して最も高い優先度のインサービス割り込みをＩＳＲ７２から除くように構成されてよい。

【００８６】

ＩＲＲ７０、ＩＳＲ７２、及びＴＭＲ７４の各々は、ｇＡＰＩＣ３４Ａによってサポートされる各割り込みベクトルに対応するロケーションを有している。図示される実施形態においては、ベクトル０乃至２５５がサポートされている。割り込みベクトル番号はまた、他の割り込みとのその相対的優先順位を示してよい（例えば、より大きなベクトル番号はそれより小さなベクトル番号よりも高い優先度であり、他の実施形態においては反対で

ある）。各割り込みベクトルに対して、ＩＲＲ７０は、当該割り込みベクトルで割り込みが要求されているか否かを表示する割り込み要求ビットを記憶する。例えば、表示は、セットされている場合に要求を表し且つクリアの場合に要求がないことを表すビットであってよい。同様に、各割り込みベクトルに対して、ＩＳＲ７２は、割り込みが当該割り込みベクトルに対してインサービスであるか否かを示すインサービスビットを記憶する（例えば、セットの場合にインサービスを表し且つクリアの場合にインサービスでないことを表す）。各割り込みベクトルに対してＴＭＲ７４はトリガモードを記憶する。ＩＲＲ７０、ＩＳＲ７２、及びＴＭＲ７４の各々に対して、レジスタ内のビットロケーションは、割り込みに対応する割り込みベクトル番号と等しい。

【００８７】

図示される実施形態においては、複数の割り込みは、係属中の割り込み要求がプロセッサへ受け渡されるべきかを決定するための優先度レベルを割り当てられている群内へ配置される。例えば、割り込みベクトル０乃至１５は優先度レベル０を割り当てられ、割り込みベクトル１６乃至３１は優先度レベル１を割り当てられ、優先度レベル１５での割り込みベクトル２４０乃至２５５まで同様である。この実施形態においては、優先度レベル番号が増大することは、優先度レベルが高くなることを示す。制御ユニット７８は要求優先度レベルを計算することができ、要求優先度レベルは、少なくとも１つの割り込み要求がＩＲＲ７０内に係属中である最も高い優先度レベルである。制御ユニット７８はまた、インサービス優先度レベルを計算することができ、インサービス優先度レベルは、少なくとも１つの割り込みがＩＳＲ７２内でインサービスとして表示されている最も高い優先度レベルである。制御ユニット７８は、要求優先度レベルがインサービス優先度レベルを超過し且つＴＰＲ７６内で表示される優先度レベルを超過する場合に、割り込みを受け渡してよい。尚、図示される実施形態においては、１６個の優先度レベル群内で２５６個の割り込みベクトルがサポートされるが、他の実施形態においては、それより多い若しくは少ない割り込みベクトル及び／又はそれより多い若しくは少ない優先度レベル群がサポートされてよい。

【００８８】

物理ＩＤレジスタ８０及び論理ＩＤレジスタ８２は、ｇＡＰＩＣ３４Ａに割り当てられる物理ＡＰＩＣ＿ＩＤ及び論理ＡＰＩＣ＿ＩＤをそれぞれ記憶してよい。ゲストＩＤレジスタ８４は、ｇＡＰＩＣ３４Ａに割り当てられるゲストＩＤを記憶してよい。従って、制御ユニット７８は、割り込みのゲストＩＤがゲストＩＤレジスタ８４内のゲストＩＤに整合する場合であって、割り込みが物理であり且つ割り込みにおけるＡＰＩＣ＿ＩＤが物理ＩＤレジスタ８０内の物理ＩＤに整合するか、あるいは割り込みが論理であり且つ割り込みにおけるＡＰＩＣ＿ＩＤが論理ＩＤレジスタ８２内の論理ＩＤに整合するかのいずれかの場合に、ゲスト割り込みマネージャ３８からの割り込みを受け入れるように構成されてよい。

【００８９】

他のＡＰＩＣ状態８６は、内部的に生成される割り込み、タイマ、ローカルベクトルテーブル等を含んでいてよい。種々の実施形態においては、他のＡＰＩＣ状態８６の幾つか又は全ては、ｇＡＰＩＣ３４Ａ内に含まれていてよく、又はＶＭＭ１８へのインターセプト及び状態のＶＭＭ１８エミュレーションと共に仮想化されてよい。

【００９０】

ｈＡＰＩＣ３２Ａ～３２Ｂは、それらがゲストＩＤレジスタを含んでいなくてよいことを除いてｇＡＰＩＣ３４Ａと同様であってよい。代替的には、ｈＡＰＩＣ３２Ａ～３２Ｂ及びｇＡＰＩＣ３４Ａ～３４Ｄは、同一ハードウエアの例(instances)であってよく（ｇＡＰＩＣ３４Ａ～３４Ｂが全てのＡＰＩＣ状態を実装している場合）、またゲストＩＤレジスタは、ゲストＩＤが有効であるか否かを表示する有効ビットを含んでいてよく、あるいはゲストＩＤレジスタはｈＡＰＩＣを表示するためにゼロにプログラムされてよい。

【００９１】

次に図５を参照すると、ｇＡＰＩＣ状態エントリ９０の１つの実施形態及びＶＭＭ管理

のｇＡＰＩＣ状態エントリ９２の１つの実施形態のブロック図が示されている。図５の図示は状態の論理ビューであってよい。メモリ内の状態の実際の配列は、幾つかの実施形態に対して図９、１２、及び１３に示されるように変化し得る。

【００９２】

一般的に、ｇＡＰＩＣ状態エントリ９０は、ｇＡＰＩＣ状態に対応するゲストがアクティブでない間に変化し得るｇＡＰＩＣを少なくとも含んでいてよい。本実施形態においては、周辺デバイスが割り込みの信号をゲストへ送ってよく、ゲストはＩＲＲ状態を変化させてよい。しかし、ＩＳＲ状態は、ゲストにおけるｖＣＰＵが割り込みを受け入れる場合にのみ変化してよく、その割り込みは、ゲストがアクティブでない場合には生じないであろう。同様に、ＴＰＲがｖＣＰＵによって変化させられ、従ってＴＰＲは、ゲストがアクティブでない間には変化しないであろう。ＶＭＭ１８は、ＶＭＭ管理のｇＡＰＩＣ状態エントリ９２におけるそのような状態の保存及び復元を管理することができる。

【００９３】

従って、図４と同様のｇＡＰＩＣ３４Ａの実施形態に対しては、ｇＡＰＩＣ状態エントリ９０はＩＲＲ７０の状態を含んでいてよい。ＶＭＭ管理のｇＡＰＩＣ状態エントリ９２は、ＩＳＲ７２、ＴＭＲ７４、ＴＰＲ７６の状態、及び種々の他のＡＰＩＣ状態８６を含んでいてよい。ＶＭＭ管理のｇＡＰＩＣ状態エントリ９２はまた、ゲストＩＤ並びに論理及び物理ＩＤを記憶してよく、あるいはこれらはエントリ９２を選択することにおいて固有のものであってよい（即ち、ＶＭＭ１８は、それらの値を用いてデータ構造６８からエントリ９２を選択してよい）。

【００９４】

次に図６を参照すると、ゲストに対してＩＯＭＭＵ４０から割り込みメッセージを受信することに応答するゲスト割り込みマネージャ３８の１つの実施形態の動作を表すフローチャートが示されている。理解の容易化のために複数のブロックが特定の順序で示されているが、他の順序が用いられてもよい。これらのブロックは、ゲスト割り込みマネージャ３８内の組み合わせ論理において並列に実行されてよい。これらのブロック、これらのブロックの組み合わせ、及び／又はフローチャートは、多重クロックサイクルでパイプライン化されてよい。一般的に、ゲスト割り込みマネージャ３８は、図６に示される動作を実装するように構成されてよい。

【００９５】

幾つかの実施形態においては、割り込みメッセージの処理は、割り込みが論理であるか物理であるかに応じて（即ち、割り込みの受け渡しモードが論理であるか物理であるかに応じて）変化してよい。例えば図１１の実施形態においては、論理割り込み及び物理割り込みに対して異なるテーブルが読まれる。図１２及び１３においては、論理テーブル及び物理テーブルはメモリ内で隣接していてよいが、論理割り込みに対して論理テーブルを配置するベースアドレスポインタにはオフセットが付加されてよく、また物理割り込みに対してはオフセットは付加される必要はない。従って、ゲスト割り込みマネージャ３８は、割り込みが論理であるかあるいは物理であるかを決定するように構成されてよい（判断ブロック１００）。他の実施形態は、受け渡しモードに基づいて変化しなくてよく、そして判断ブロック１００は除かれてよい（また、以下に議論されるブロードキャスト又はより多くの宛先に対するチェックは、両方に対する１つのチェックにまとめられてよい）。

【００９６】

割り込みが論理である場合（判断ブロック１００、「イエス」行程）、ゲスト割り込みマネージャ３８は、論理割り込みからｇＡＰＩＣ状態データ構造５８内の対応するｇＡＰＩＣ状態エントリ９０へのマッピングを決定するように構成されてよい（ブロック１０２）。図１０～１３に示されるように、種々の実施形態は異なるマッピングを実装してよく、従って決定は変化し得る。ゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態エントリ９０内で代表されるＩＲＲ内の割り込みベクトルに対応するビットをセットするように構成されてよい（ブロック１０４）。論理割り込みは多重宛先を有していてよい（例えば、クラスタ内の宛先は、１つ以上のセットビットを有していてよいビットベクトルである）

。論理割り込みがより多くの宛先を含む場合（判断ブロック１０６、「イエス」行程）、ゲスト割り込みマネージャ３８は、各追加的な宛先に対してブロック１０２及び１０４を繰り返すように構成されてよい。代替的には、図１２の実施形態においては、後で更に詳細に説明されるように、論理宛先ビットベクトルは、１つの動作におけるｇＡＰＩＣ状態エントリへ書き込まれてよい。ゲスト割り込みマネージャ３８は、割り込みメッセージをｇＡＰＩＣ３４Ａ～３４Ｄへ送信するように構成されてよい（ブロック１０８）。

【００９７】

割り込みが物理である場合（判断ブロック１００、「ノー」行程）、ゲスト割り込みマネージャ３８は、物理割り込みからｇＡＰＩＣ状態データ構造５８内の対応するｇＡＰＩＣ状態エントリ９０へのマッピングを決定するように構成されてよい（ブロック１１０）。図１０～１３に示されるように、種々の実施形態は異なるマッピングを実装してよく、従って決定は変化し得る。ゲスト割り込みマネージャ３８は、ｇＡＰＩＣ状態エントリ９０内で代表されるＩＲＲ内の割り込みベクトルに対応するビットをセットするように構成されてよい（ブロック１１２）。物理割り込みは、ブロードキャスト又は単一の宛先であってよい。物理割り込みがブロードキャストである場合（判断ブロック１１４、「イエス」行程）、ゲスト割り込みマネージャ３８は、ゲストの仮想マシン（例えば各ｖＣＰＵ）内の各宛先に対してブロック１１０及び１１２を繰り返すように構成されてよい。代替的には、図１２の実施形態においては、後で更に詳細に説明されるように、ブロードキャストは、１つの動作におけるｇＡＰＩＣ状態エントリ内に記録されてよい。ゲスト割り込みマネージャ３８は、割り込みメッセージをｇＡＰＩＣ３４Ａ～３４Ｄへ送信するように構成されてよい（ブロック１０８）。

【００９８】

ｇＡＰＩＣ状態エントリ９０内で代表されるＩＲＲ内でビットをセットすることは、アトミックＯＲ動作として実行されてよく、アトミックＯＲ動作においては、セットされたビットはメモリロケーション内の他のＩＲＲビット内へ論理和を取られる。アトミックＯＲ動作の実際の実装は、ロックされた読み出し／修正／書き込み動作から、ＯＲを１つの動作として実行するように規定された特殊目的回路まで、変化してよい。

【００９９】

上述したように、他の実施形態においては、比較及び交換動作が実行されてよい。

【０１００】

別の実施形態においては、２つ以上の宛先を伴う論理割り込み及び、ブロードキャストである物理割り込みが、ＶＭＭ１８にアクセス可能なデータ構造（例えばイベントキュー(event queue)）内に割り込みを記録することによって、ゲスト割り込みマネージャ３８により取り扱われてよい。ゲスト割り込みマネージャ３８はまた、ＶＭＭ１８にイベントを知らせるために、ＶＭＭ１８へ信号を送るように構成されてよい（例えば、プロセッサ３０Ａ～３０Ｂの１つにおける仮想マシンからの退出を生じさせる）。代替的には、ゲスト割り込みマネージャ３８は、ＶＭＭ１８へ周期的にのみ信号を送ってよく（例えば、Ｎミリ秒毎に１回且つ／又はイベントキューにおける高ウォータマーク(high watermark)で１回）、そしてＶＭＭ１８は、周期的にイベントキューをチェックしてよい他、信号を送ることがサポートしたであろうよりも迅速に任意のイベントをサービスし得る。１つの実施形態においては、イベントキューは、ゲスト割り込みマネージャ３８に代えてＩＯＭＭＵ４０によって管理されてよい。

【０１０１】

次に図７を参照すると、ゲスト割り込みマネージャ３８からの割り込みメッセージを受信することに応答するｇＡＰＩＣ３４Ａ～３４Ｄの１つの実施形態の動作を表すフローチャートが示されている。理解の容易化のために複数のブロックが特定の順序で示されているが、他の順序が用いられてもよい。これらのブロックは、ｇＡＰＩＣの組み合わせ論理において並列に実行されてよい。これらのブロック、これらのブロックの組み合わせ、及び／又はフローチャートは、多重クロックサイクルでパイプライン化されてよい。一般的に、ｇＡＰＩＣは、図７に示される動作を実装するように構成されてよい。

【０１０２】
１つの実施形態においては、ｇＡＰＩＣは、そのゲストＩＤ（ゲストＩＤレジスタ８４内の、図４参照）をゼロにセットすることによって非アクティブ化される。従って、割り込みメッセージを受信することに応答して、ｇＡＰＩＣゲストＩＤがゼロである場合（判断ブロック１２０、「イエス」行程）、ｇＡＰＩＣは非アクティブであり割り込みを処理しなくてよい。他の実施形態は他の方法（例えばレジスタ内のアクティブビット）でｇＡＰＩＣを非アクティブ化してよく、そして判断ブロック１２０は、ｇＡＰＩＣアクティブ／非アクティブに対するチェックに従って修正されてよい。
【０１０３】
ｇＡＰＩＣゲストＩＤが非ゼロである場合、ｇＡＰＩＣは、ゲストＩＤを、受信した割り込みのゲストＩＤと比較する他、受信した宛先ＩＤをレジスタ８０及び８２内のそれぞれ論理ＩＤ及び物理ＩＤ（図４参照）と比較するように構成されてよい。ｇＡＰＩＣのゲストＩＤが受信したゲストＩＤと整合しない場合（判断ブロック１２２、「ノー」行程）、ｇＡＰＩＣはその時点で異なるゲストに割り当てられており、従ってｇＡＰＩＣは割り込みによって対象とされていない。ｇＡＰＩＣは、割り込みの非承認で応答するように構成されてよい（ブロック１２４）。非承認は、ｇＡＰＩＣが割り込みを受信したが、割り込みは対応するプロセッサで対象とされておらず、従って受け入れられなかったとｇＡＰＩＣが判断したことを表示してよい。同様に、ｇＡＰＩＣのゲストＩＤは受信したゲストＩＤと整合するが、割り込みは論理であり且つｇＡＰＩＣの論理ＩＤと整合せず、あるいは割り込みは物理であり、単一の宛先であり、且つｇＡＰＩＣ物理アドレスと整合しない場合（判断ブロック１２６及び１２８、「ノー」行程）、ｇＡＰＩＣは、割り込みの非承認で応答するように構成されてよい（ブロック１２４）。
【０１０４】
論理割り込みを整合させることは、一般的に、論理ＩＤのクラスタＩＤ部分を均等のために比較することと、ｇＡＰＩＣの論理ＩＤレジスタ内のセットビットもまた割り込みから受信された論理ＩＤの宛先部分においてセットされていることを検出することとを含んでいてよい。割り込みの論理ＩＤの宛先部分における他のビットもまた、２つ以上の宛先がある場合にセットされてよい。物理ＩＤは、ゲストＩＤが整合する限りブロードキャスト物理割り込みが整合として取り扱われ得ることを除いて、均等のために比較されてよい。
【０１０５】
割り込みが論理であり且つ論理ＩＤと整合する場合（判断ブロック１２６、「イエス」行程）、又は割り込みが物理であり且つ物理ＩＤと整合し若しくはブロードキャストである場合（判断ブロック１２８、「イエス」行程）、ｇＡＰＩＣ３４は、対応するプロセッサ３０Ａ～３０Ｂへの提示のためにｇＡＰＩＣが割り込みを受け入れていることを表示する承認でゲスト割り込みマネージャ３８に応答するように構成されてよい（ブロック１３０）。ｇＡＰＩＣはまた、ＩＲＲレジスタ７０を更新して割り込みメッセージ内の割り込みベクトルに対応する割り込み要求ビットをセットするように構成されてよい（ブロック１３２）。ｇＡＰＩＣは、任意のインサービス割り込みに関する割り込みの優先度及び／又はタスク優先度レジスタを再評価するように構成されてよく（ブロック１３４）、また再評価に基づいてプロセッサへ割り込みの信号を送るように構成されてよい（ブロック１３６）。即ち、ｇＡＰＩＣは、割り込みの優先度がインサービス割り込みよりも高い優先度であり且つタスク優先度レジスタよりも高い優先度である場合に、割り込みの信号を送るように構成されてよい。
【０１０６】
次に図８を参照すると、ｇＡＰＩＣ状態を１つのゲストから他へ変化させるＶＭＭ１８の１つの実施形態の動作を表すフローチャートが示されている。即ち、図８のフローチャートは、ｇＡＰＩＣ３４Ａ～３４Ｄを１つのゲスト／ｖＣＰＵから他のゲスト又は同一ゲスト内の他のｖＣＰＵへ再割り当てすることを提示していてよい。理解の容易化のために複数のブロックが特定の順序で示されているが、他の順序が用いられてもよい。一般的に

、ＶＭＭ１８は、システム５上で実行されるときに図８に示される動作を実装する命令を含んでいてよい。

【０１０７】

ＶＭＭ１８は、ｇＡＰＩＣ状態データ構造５８内の「古いゲスト(old guest)」（ｇＡＰＩＣから非アクティブ化されているゲスト）に対応するｇＡＰＩＣ状態エントリ９０のロケーションを決定してよい（ブロック１４０）。ｇＡＰＩＣ状態エントリ９０内のデータは、それがｇＡＰＩＣによって修正されてしまっていることがあるという理由で、「新鮮でない(stale)」と考えられる。例えば、ＩＲＲビットは、対応するプロセッサへ割り込みを受け渡すことに応答してリセットされてしまっているかもしれない。従って、ＶＭＭ１８はｇＡＰＩＣ状態エントリ９０内のＩＲＲをゼロにしてよい（ブロック１４２）。ＶＭＭ１８は、ゲストＩＤレジスタ８４、論理ＩＤレジスタ８２、及び物理ＩＤレジスタ８０をクリアしてよい（ブロック１４４）。この動作により、ゲストＩＤ、論理ＩＤ、及び物理ＩＤは割り込みメッセージに整合しなくなることになるので、ｇＡＰＩＣが任意の追加的な割り込みを受け入れることを妨げることができる。レジスタ８０～８４がクリアされ（ブロック１４４）且つＩＲＲ状態がｇＡＰＩＣ状態エントリ９０へ書き込まれる前に、割り込みが送信され得ていることが可能である。従って、割り込み状態の損失を回避するために、ＩＲＲ７０からｇＡＰＩＣ状態エントリ９０内へＩＲＲ状態についてアトミックに論理和を取ってよい（ブロック１４６）。ＶＭＭ１８はまた、古いゲストに関連するＶＭＭ管理のｇＡＰＩＣ状態エントリ９２へ他のｇＡＰＩＣ状態を書き込んでよい。

【０１０８】

ＶＭＭ１８は、ｇＡＰＩＣのＩＲＲ、ＩＳＲ、及びＴＭＲのレジスタ７０、７２、及び７４をクリアして古いゲストの割り込み状態を解除してよい（ブロック１５０）。ＶＭＭ１８は、ｇＡＰＩＣに割り当てられつつあるゲストに対する新たなゲストＩＤ、論理ＩＤ、及び物理ＩＤをそれぞれゲストＩＤレジスタ８４、論理ＩＤレジスタ８２、及び物理ＩＤレジスタ８０に書き込んでよい（ブロック１５２）。ブロック１５２が一旦実行されると、ｇＡＰＩＣはそのゲストのための割り込みを受け入れることを開始してよい。ＶＭＭ１８は、「新たなゲスト(new guest)」（ｇＡＰＩＣにおいてアクティブ化されつつあるゲスト）に対するｇＡＰＩＣ状態エントリ９０を決定してよく（ブロック１５４）、そしてｇＡＰＩＣ状態エントリ９０からＩＲＲ状態を読み出してよい（ブロック１５６）。レジスタ８０～８４のプログラミングは、ｇＡＰＩＣに割り込みを受け入れることを開始させるであろうから、ＶＭＭ１８がエントリを読み出した後にｇＡＰＩＣ状態エントリ９０内に記録されていたＩＲＲ内の割り込みをｇＡＰＩＣが受け入れてしまっていることが可能である。従って、ＶＭＭ１８は、ＩＲＲレジスタ７０内へＩＲＲ状態についてアトミックに論理和を取ってよい。即ち、ｇＡＰＩＣは、ＩＲＲレジスタ７０に対するアトミックなＯＲ動作をサポートしてよい（ブロック１５８）。ＶＭＭ１８は、新たなゲストに対してＶＭＭ管理のｇＡＰＩＣ状態エントリ９２から他の状態を読み出してよく（ブロック１６０）、そしてその状態をｇＡＰＩＣへ書き込んでよい（ブロック１６２）。尚、ブロック１６０及び１６２は、ブロック１５０の後の任意の他の点で実行されてもよい。

【０１０９】

ブロック１４０～１４８は、概してｇＡＰＩＣからのゲストを非アクティブ化するための動作を代表していてよく、一方ブロック１５０～１６２は、概してｇＡＰＩＣ内のゲストをアクティブ化するための動作を代表していてよい。従って、図８に水平な破線によって示されるように、ｇＡＰＩＣ内のゲストを非アクティブ化することのみをＶＭＭ１８が望む場合には、水平な破線よりも上方のブロックが実行されてよい。ｇＡＰＩＣ内のゲストをアクティブ化することのみをＶＭＭ１８が望む場合には、水平な破線よりも下方のブロックが実行されてよい。

【０１１０】

次に図９を参照すると、実施形態に対するｇＡＰＩＣ状態エントリ９０内のｇＡＰＩＣ状態の１つの例示的な配列１７０を表すブロック図が示されている。図９の実施形態においては、ＩＲＲの各ビットは異なるバイト内に記憶される。例えば、図９におけるＩＲＲ

ビット０、即ちＩＲＲ０は、メモリ内の一連の連続するバイトのバイト０内に記憶され、ＩＲＲ１はバイト１内に記憶され、バイト２５５内に記憶されているＩＲＲ２５５まで同様に続く。図示される実施形態においては、ＩＲＲビットはバイトのビット０内に記憶されているが、任意のビット位置が用いられてよい。バイト内の他のビットは、図示される実施形態では無関係(don’t cares)（ＤＣ）である。各ビットを別々のバイト（メモリアクセスの最小単位）内に記憶することによって、各ビットは、他のビットに影響を与えることなしに個別に書き込まれ得る。従って、ビットは書き込みを介してバイトへセットされてよく、これはアトミックな動作である。セットビットをバイト内のＩＲＲビット位置へ書き込むこと及び他のバイトを更新しないことによって、ＩＲＲビットのアトミックＯＲが結果であり得る。他の実施形態においては、アトミックＯＲは他の方法において達成されてよく、またＩＲＲ状態のビットは他の方法において記憶されてよい。

【０１１１】

次に図１０を参照すると、ｇＡＰＩＣ状態エントリ９０を配置する１つの実施形態のブロック図が示されている。図示される実施形態においては、デバイステーブル６２及び割り込みリダイレクトテーブル６４が示される他、ｇＡＰＩＣ状態マッピングテーブル６０の実施形態が示されている。実施形態においては、割り込みを送信した周辺機器のＢＤＦがデバイステーブル６２内への索引(index)として用いられ、そしてエントリは、ＢＤＦが割り当てられているゲストに対するゲストＩＤを含んでいてよい。またこの実施形態においては、エントリは割り込みリダイレクトテーブルポインタ（ＩＲＴＰ）を含んでおり、ＩＲＴＰは割り込みリダイレクトテーブル６４のベースを指す。割り込みリダイレクトテーブル６４内への索引は、割り込みのための割り込み識別子である。割り込み識別子は、割り込みベクトルを含んでいてよく、また物理又は論理のいずれかである割り込みの受け渡しモード（Ｄｅｌｍｏｄｅ）を含んでいてもよい。選択されたエントリは、新たなベクトル及び宛先ＩＤ（ＤｅｓｔＩＤ）を含んでいてよい。割り込みリダイレクトテーブル６４を用いない実施形態においては、周辺機器によって供給される割り込みベクトル及び宛先ＩＤは、ｇＡＰＩＣ状態マッピングテーブル６０を直接的に索引付けるために用いられてよい。

【０１１２】

ｇＡＰＩＣ状態マッピングテーブル６０は、ｇＡＰＩＣ状態マッピングテーブルベースアドレスを介してメモリ内に配置されてよい。ベースアドレスは、種々の実施形態において、全てのゲストに対して同じであってよく、ゲスト固有であってよく、あるいはデバイステーブル６２内に記憶されていてよい。図１０においては、ベースアドレスは、一連の階層テーブルの最も高いレベル（Ｌ３）を識別し、階層テーブルは、より低いレベルのテーブル（例えばＬ２及び、Ｌ２を指していないＬ３からのポインタによって表示される同様のテーブル）へのポインタを記憶していてよい。Ｌ２テーブルは更に低いレベルのテーブル（Ｌ１）へのポインタを記憶していてよく、そのテーブルは、ｇＡＰＩＣ状態データ構造５８内のｇＡＰＩＣ状態エントリ９０へのポインタを記憶していてよい。他の実施形態は、図１０に示される３レベルよりも多い又は少ないレベルを含め、階層内の任意の数のレベルを用いてよい。

【０１１３】

ｇＡＰＩＣ状態マッピングテーブル６０内の各レベルＬ３～Ｌ１内への索引は、デバイステーブル６２からのゲストＩＤを連結させることから形成される値の一部分、周辺機器からの又は割り込みリダイレクトテーブル６４からの割り込みベクトル、及び周辺機器からの又は割り込みリダイレクトテーブル６４からの宛先ＩＤであってよい。レベルＬ３～Ｌ１への索引は、連結された値の全てのビットを消費してよく、従って、ゲストＩＤ、ベクトル、及び宛先ＩＤの各組み合わせは、それ自身の固有のポインタをｇＡＰＩＣ状態マッピングテーブル６０内に有していてよい。しかし、幾つかのポインタは同一のｇＡＰＩＣ状態エントリ９０を指し示すことがある（例えば、１つの実施形態においては、同一のｇＡＰＩＣの論理ＩＤ及び物理ＩＤが同一のｇＡＰＩＣ状態９０へのポインタを有していることがある）。

【０１１４】
次に図１１を参照すると、ｇＡＰＩＣ状態エントリ９０を配置する他の実施形態のブロック図が示されている。図示される実施形態においては、デバイステーブル６２及び割り込みリダイレクトテーブル６４が示されている他、ｇＡＰＩＣ状態マッピングテーブル６０の実施形態が示されている。実施形態においては、割り込みを送信した周辺機器のＢＤＦがデバイステーブル６２内への索引として用いられ、そしてエントリは、ＢＤＦが割り当てられているゲストに対するゲストＩＤを含んでいてよい。またこの実施形態においては、エントリは割り込みリダイレクトテーブルポインタ（ＩＲＴＰ）を含んでおり、ＩＲＴＰは割り込みリダイレクトテーブル６４のベースを指す。デバイステーブル６２は更に、ｇＡＰＩＣ状態マッピングテーブル６０内のテーブルへの１つ以上のポインタを含んでいてよい。具体的には、ゲスト物理テーブルへのポインタ及びゲスト論理テーブルへの別のポインタが記憶されていてよい。ゲスト物理テーブルは、物理宛先ＩＤをｇＡＰＩＣ状態エントリ９０へマッピングしてよい。即ち、ゲスト物理テーブルは、宛先ＩＤによって索引付けられてよく、そしてポインタをｇＡＰＩＣ状態エントリ９０へ記憶してよい。同様に、ゲスト論理テーブルは、論理宛先ＩＤをｇＡＰＩＣ状態エントリ９０へマッピングしてよい。

【０１１５】
割り込みリダイレクトテーブル６４内への索引は、割り込みのための割り込み識別子である。割り込み識別子は、割り込みベクトルを含んでいてよく、また物理又は論理のいずれかである受け渡しモード（Ｄｅｌｍｏｄｅ）を含んでいてもよい。選択されたエントリは、新たなベクトル及び宛先ＩＤ（ＤｅｓｔＩＤ）を含んでいてよい。割り込みリダイレクトテーブル６４を用いない実施形態においては、周辺機器によって供給される割り込みベクトル及び宛先ＩＤは、ｇＡＰＩＣ状態マッピングテーブル６０を直接的に索引付けるために用いられてよい。

【０１１６】
次に図１２を参照すると、ｇＡＰＩＣ状態エントリ９０を配置する別の実施形態のブロック図が示されている。この実施形態では、ｇＡＰＩＣ状態マッピングテーブル６０はない。図１０～１１の実施形態と同様に、割り込みを送信した周辺機器のＢＤＦがデバイステーブル６２内への索引として用いられ、そしてエントリは、ＢＤＦが割り当てられているゲストに対するゲストＩＤを含んでいてよく、また随意的に割り込みリダイレクトテーブルポインタ（ＩＲＴＰ）を含んでおり、ＩＲＴＰは割り込みリダイレクトテーブル６４のベースを指す。デバイステーブル６２は更に、ｇＡＰＩＣ状態データ構造５８内のテーブルのベースへの少なくとも１つのポインタ（Ｐｔｒ）を含んでいてよい。図示される実施形態においては、テーブルは、ゲスト物理セクション１８０及びゲスト論理セクション１８２を含んでいる。セクション１８０及び１８２は、図面における明瞭さのために図１２においてこれらの間の空白と共に図示されているが、セクション１８０及び１８２はメモリ内で隣接していてよい。即ち、ゲスト物理セクション１８０の先頭は、ゲスト論理セクション１８２の最後尾と隣接していてよい。デバイステーブル６２は更に、ゲスト論理部分１８２の先頭を表示する論理限界(logical limit)（ＬＬｉｍ）フィールドを含んでいてよい。他の実施形態においては、ゲスト物理部分１８０及びゲスト論理部分１８２は隣接していなくてよく、また個々のポインタがデバイステーブル６２エントリ内に記憶されてそれぞれゲスト物理部分１８０及びゲスト論理部分１８２を表示してよい。

【０１１７】
図１２の実施形態においては、ゲスト物理部分１８０は割り込みベクトルによって索引付けられてよい（周辺機器からの、又は割り込みリダイレクトテーブル６４からのいずれか）。ゲスト物理部分１８０内の各エントリは、ゲスト物理マシン内でサポートされている宛先ＩＤ（例えば、図１２において０乃至６３の数字が付されている６４個までの宛先）に対応するビットベクトルを備えていてよい。物理割り込みに応答して、ゲスト割り込みマネージャ３８は、宛先ＩＤに対応する割り込みベクトルに対するエントリ内のビットをセットするように構成されてよい。ブロードキャスト割り込みに対しては、ゲスト割り

込みマネージャ３８は、仮想マシン内のｖＣＰＵの数までの割り込みベクトルに対応するエントリ内の各ビットをセットするように構成されてよい。

【０１１８】

ゲスト論理部分１８２は、論理ＩＤのクラスタ部分及びベクトルによって索引付けられてよい。クラスタ部分は索引の最も有意なビットであってよく、従って、ゲスト論理部分１８２は、各論理クラスタに対応する複数のクラスタ部分（図１２におけるクラスタ０乃至クラスタＮ）へと分割される。各クラスタの範囲内で複数のエントリが割り込みベクトルによって配列され、各エントリは論理ＩＤのベクトル部分に対応するビットベクトルを記憶している。図示される実施形態においては、１６までの宛先が１つのクラスタ内に含まれてよい（例えば論理ＩＤのビットベクトル部分は１６ビットであってよい）。論理割り込みに応答して、ゲスト割り込みマネージャ３８は、論理ＩＤのビットベクトル部分の、割り込みベクトルに対応するエントリのコンテンツとの論理和を取るように構成されてよい。

【０１１９】

従って、図１２の実施形態は、ブロードキャスト物理割り込みの記録及びｇＡＰＩＣ状態データ構造５８に対する単一の更新内に多重宛先を有する論理割り込みの記録をサポートすることができる。ｇＡＰＩＣに対するｇＡＰＩＣ状態エントリは、ｇＡＰＩＣの物理ＩＤに対応するゲスト物理部分１８０の列を備えていてよく、その列はｇＡＰＩＣの論理ＩＤによって表示されるクラスタからの列との論理和を取られ、その列はｇＡＰＩＣの論理ＩＤのビットベクトル部分内のセットビットによって識別される。ｇＡＰＩＣ内のゲストを非アクティブ化することに応答してｇＡＰＩＣ状態データ構造５８を更新することは、ゲストに対応する列の１つをゼロにすることと、他の列へＩＲＲを書き込むこととを含んでいてよい。

【０１２０】

図１３はｇＡＰＩＣ状態エントリ９０を配置する別の実施形態である。図１３の実施形態は、ゲスト物理部分１８０及びゲスト論理部分１８２の配列が異なること以外は図１２の実施形態と同様である。各エントリはＩＲＲに対応し、従って各エントリは各割り込みベクトルに対するビットを含む。ゲスト物理部分１８０は、割り込みの物理ＩＤによって索引付けられ、そしてゲスト論理部分１８２は、割り込みの論理ＩＤによって索引付けられる。割り込みベクトルに対応するＩＲＲビットは、割り込みの受け渡しモードに応じて論理部分１８２又は物理部分１８０のいずれかにおいてセットされる。ゲスト／ｖＣＰＵに対するｇＡＰＩＣ状態は、そのゲスト／ｖＣＰＵに割り当てられた物理ＩＤに対応するゲスト物理部分１８０からの行と、そのゲスト／ｖＣＰＵに割り当てられた論理ＩＤに対応するゲスト論理部分１８２からの行との論理和である。

【０１２１】

次に図１４を参照すると、ホストハードウエア２０の別の実施形態のブロック図が示されている。図示される実施形態には、図２における集積回路６６と同様の２つの集積回路６６Ａ～６６Ｂが含まれる。従って、図示されるように、各集積回路は、集積回路６６Ａ内の３４Ａ～３４Ｄ及び集積回路６６Ｂ内の３４Ｅ～３４ＧのようなｇＡＰＩＣを含んでいてよい。各集積回路６６Ａ～６６Ｂは、それぞれのゲスト割り込みマネージャ２８Ａ～２８Ｂ及びＩＯＭＭＵ（図１４には図示せず）を含んでいてよい。集積回路６６Ａ～６６Ｂの少なくとも一方はメモリ５６Ａ～５６Ｂに結合され、また両集積回路６６Ａ～６６Ｂは随意的にメモリを含んでいてよい。集積回路６６Ａ～６６Ｂは、図示される実施形態においては、インタフェース回路４４Ｃ及び４４Ｄを介して結合される。他の実施形態においては、３つ以上の集積回路６６Ａ～６６Ｂが設けられてよく、また種々の集積回路が任意の望ましい方法で相互接続されてよい。

【０１２２】

１つの実施形態においては、各ゲスト割り込みマネージャ２８Ａ～２８Ｂが有効にされて、そして同一の集積回路内のｇＡＰＩＣ３４Ａ～３４Ｇで対象にされる割り込みメッセージを管理してよい。従って、ゲスト割り込みマネージャ２８Ａ～２８Ｂは、ゲスト割り

込み受け渡しに対して拡張性のある解決法を提供することができる。ゲスト割り込みマネージャ２８Ａ～２８Ｂによって用いられるデータ構造は、１つのメモリ（例えばメモリ５６Ａ）内に記憶されていてよく、あるいは各ゲスト割り込みマネージャ２８Ａ～２８Ｂは、それ自身のデータ構造をそれ自身のメモリ５６Ａ～５６Ｂ内に有していてよい。データ構造へのアクセスに対しては何らかの競合(contention)があり得るが、多くの場合には、周辺機器が特定のゲスト（集積回路６６Ａ～６６Ｂの一方におけるプロセッサ上で実行中の）に割り当てられ、従って実際の競合の量は比較的小さいであろう。

【０１２３】

別の実施形態においては、ゲスト割り込みマネージャ２８Ａ～２８Ｂの一方が有効にされて、そしてシステム内のｇＡＰＩＣ３４Ａ～３４Ｇに対するゲスト割り込み受け渡しを実行してよい。そのような実施形態は、集積回路６６Ａ～６６Ｂの間での相互接続を介して、より多くの割り込み関連のトラフィックを経験するであろうが、ゲスト割り込み管理に対する主要目的の概念的な簡潔さを提供することもできる。

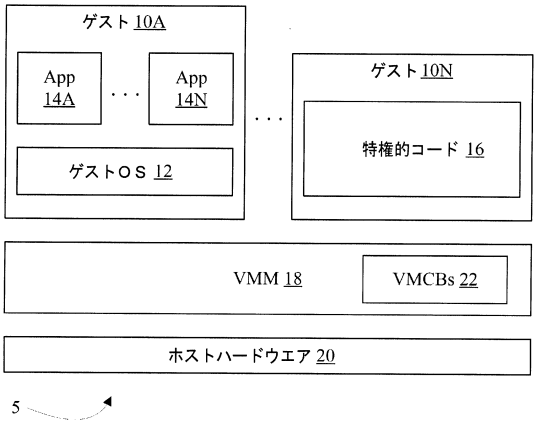
【０１２４】

次に図１５を参照すると、コンピュータアクセス可能記憶媒体２００のブロック図が示されている。一般的に言って、コンピュータアクセス可能記憶媒体は、命令及び／又はデータをコンピュータへ供給するための使用の間にコンピュータによってアクセス可能な任意の記憶媒体を含んでいてよい。コンピュータアクセス可能記憶媒体の例としては、磁気媒体又は光学媒体のような記憶媒体、例えばディスク（固定の又は取り外し可能な）、テープ、ＣＤ－ＲＯＭ、又はＤＶＤ－ＲＯＭ、ＣＤ－Ｒ、ＣＤ－ＲＷ、ＤＶＤ－Ｒ、ＤＶＤ－ＲＷ、及び／又はブルーレイ(Blu-Ray)ディスクを挙げることができる。記憶媒体の更なる例としては、揮発性の又は不揮発性のメモリ媒体、例えば、ＲＡＭ（例えば同期ダイナミックＲＡＭ（ＳＤＲＡＭ）、ラムバスＤＲＡＭ（ＲＤＲＡＭ）、スタティックＲＡＭ（ＳＲＡＭ）、等）、ＲＯＭ、フラッシュメモリ、ユニバーサルシリアルバス（ＵＳＢ）インタフェースのような周辺機器インタフェース又は任意の他のインタフェースを介してアクセス可能な不揮発性メモリ（例えばフラッシュメモリ）、等を挙げることができる。記憶媒体は、微小電気機械的システム（ＭＥＭＳ）の他、ネットワーク及び／又はワイヤレスリンクのような通信媒体を介してアクセス可能な記憶媒体を含む。図１５におけるコンピュータアクセス可能記憶媒体２００はＶＭＭ１８を記憶していてよく、ＶＭＭ１８は、図８のフローチャート及び／又はＶＭＭ１８に割り当てられる任意の機能性を実装していてよい。概して、コンピュータアクセス可能記憶媒体２００は、実行されるときに図８に示されるフローチャートの一部分又は全部を実装する命令の任意のセットを記憶していてよい。搬送媒体は、コンピュータアクセス可能記憶媒体の他、有線又はワイヤレスの伝送のような伝送媒体を含む。

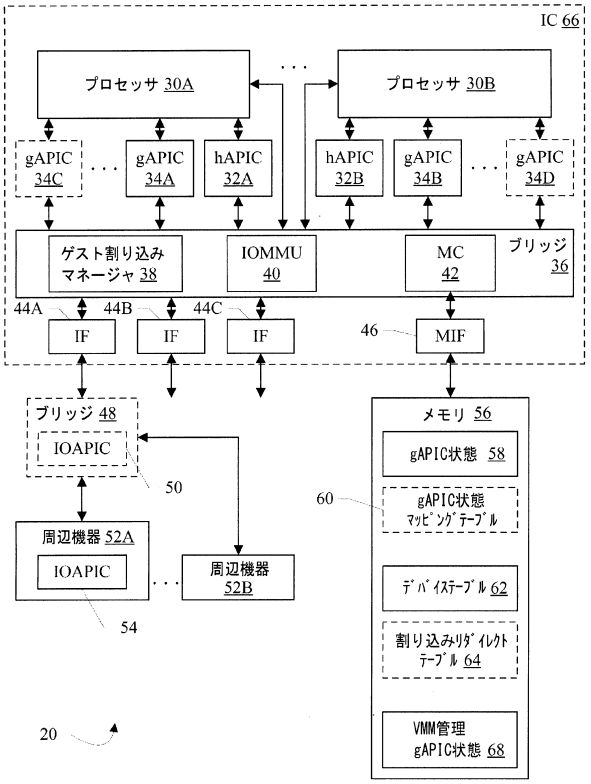
【０１２５】

上述の開示が完全に理解されるならば、多くの変形及び修正が当業者に明らかになるであろう。以下の特許請求の範囲は、全てのそのような変形及び修正を包含するものと解釈されることが意図されている。

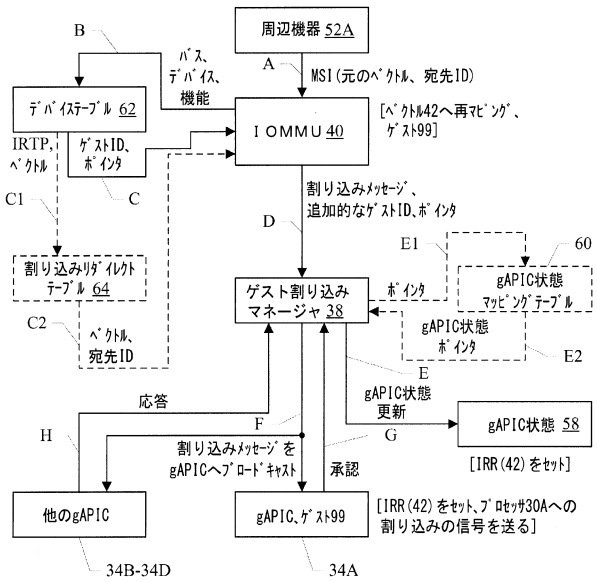
【図１】


ゲスト 10A
App 14A
App 14N
ゲストOS 12
ゲスト 10N
特権的コード 16
VMM 18
VMCBs 22
ホストハードウエア 20
5

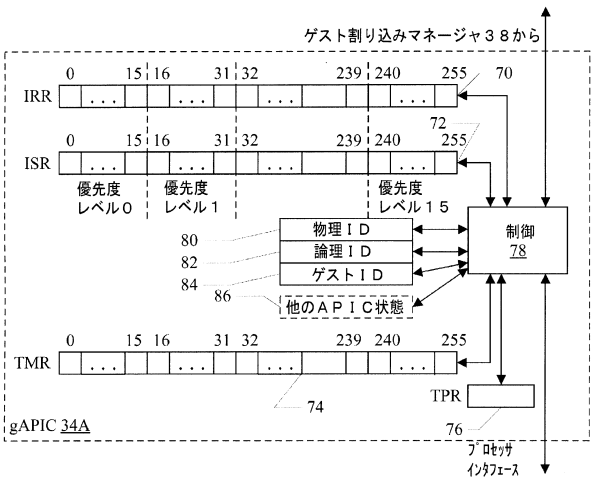
【図２】


IC 66
プロセッサ 30A
プロセッサ 30B
gAPIC 34C
gAPIC 34A
hAPIC 32A
hAPIC 32B
gAPIC 34B
gAPIC 34D
ゲスト割り込みマネージャ 38
IOMMU 40
MC 42
ブリッジ 36
44A
44B
44C
IF
IF
IF
46
MIF
ブリッジ 48
IOAPIC
50
周辺機器 52A
IOAPIC
周辺機器 52B
54
メモリ 56
gAPIC状態 58
60
gAPIC状態 マッピングテーブル
デバイステーブル 62
割り込みリダイレクト テーブル 64
VMM管理 gAPIC状態 68
20

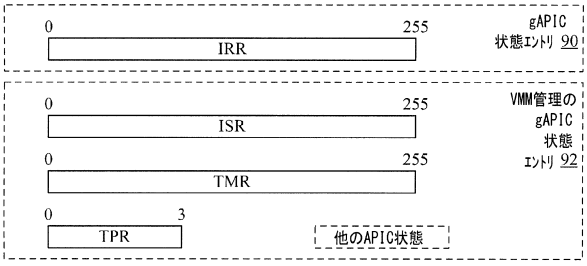
【図３】


周辺機器 52A
B
バス、デバイス、機能
A
MSI(元のベクトル、宛先ID)
デバイステーブル 62
IOMMU 40
[ベクトル42へ再マピング、ゲスト99]
IRTP、ベクトル
ゲストID、ポインタ
C1
C
割り込みメッセージ、追加的なゲストID、ポインタ
D
E1
60
割り込みリダイレクト テーブル 64
ゲスト割り込みマネージャ 38
ポインタ
gAPIC状態 マッピングテーブル
C2
ベクトル、宛先ID
gAPIC状態 ポインタ
E2
E
応答
gAPIC状態 更新
gAPIC状態 58
H
F
G
割り込みメッセージを gAPICへブロードキャスト
承認
[IRR(42)をセット]
他のgAPIC
gAPIC、ゲスト99
[IRR(42)をセット、プロセッサ30Aへの割り込みの信号を送る]
34B-34D
34A

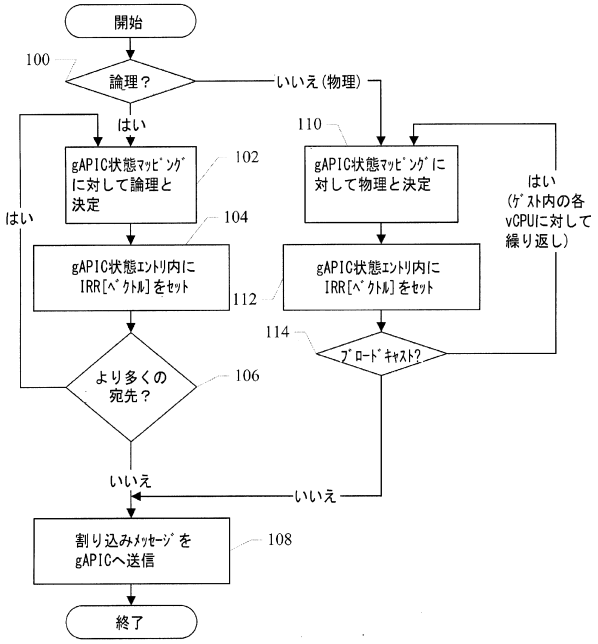
【図４】


ゲスト割り込みマネージャ３８から
IRR
ISR
0
15
16
31
32
239
240
255
70
72
優先度 レベル０
優先度 レベル１
優先度 レベル１５
物理ＩＤ
論理ＩＤ
ゲストＩＤ
他のＡＰＩＣ状態
80
82
84
86
制御 78
TMR
TPR
74
76
gAPIC 34A
プロセッサ インタフェース

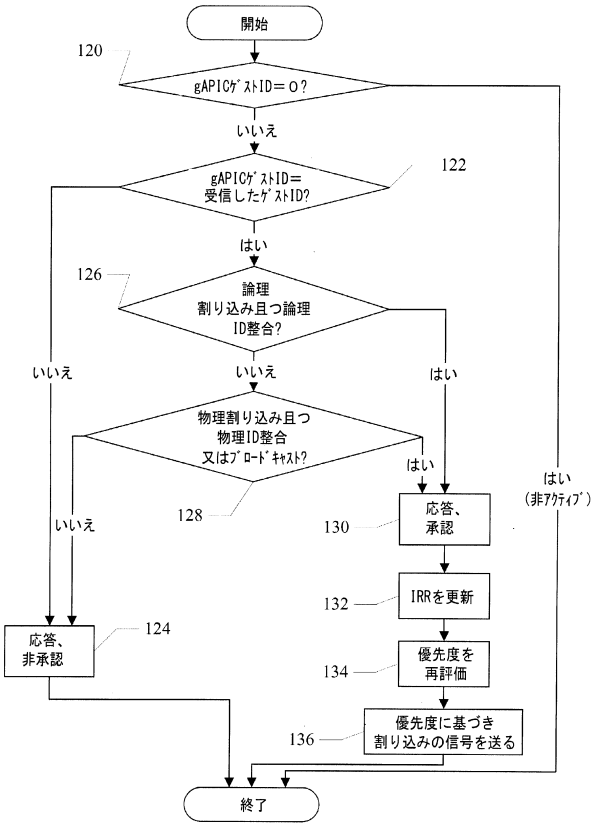
【図５】


0
255
IRR
gAPIC 状態エントリ 90
ISR
TMR
VMM管理の gAPIC 状態 エントリ 92
3
TPR
他のAPIC状態

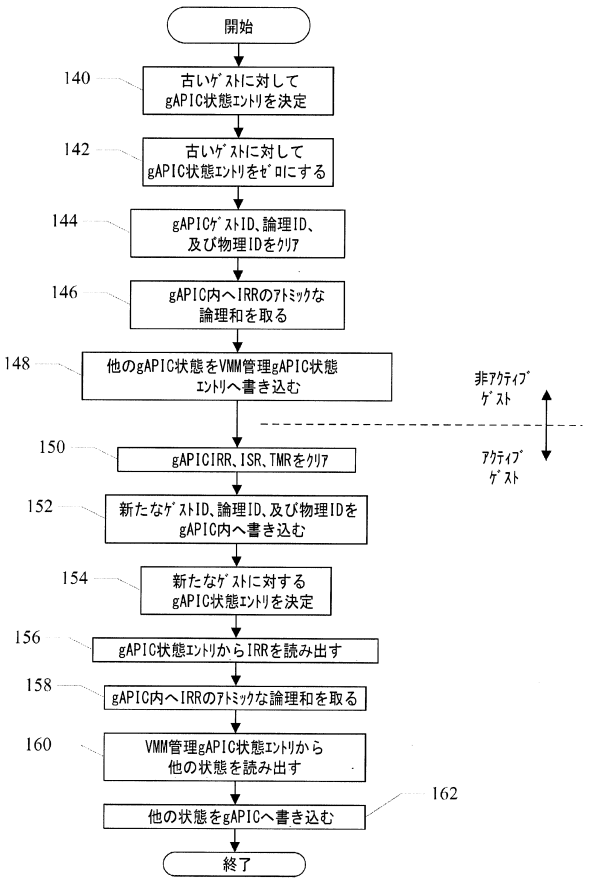
【図６】


開始
100
論理？
いいえ(物理)
はい
110
gAPIC状態ﾏｯﾋﾟﾝｸﾞに対して論理と決定
102
gAPIC状態ﾏｯﾋﾟﾝｸﾞに対して物理と決定
はい
(ｹﾞｽﾄ内の各vCPUに対して繰り返し)
はい
104
gAPIC状態ｴﾝﾄﾘ内にIRR[ﾍﾞｸﾄﾙ]をｾｯﾄ
gAPIC状態ｴﾝﾄﾘ内にIRR[ﾍﾞｸﾄﾙ]をｾｯﾄ
112
114
ﾌﾞﾛｰﾄﾞｷｬｽﾄ?
より多くの宛先？
106
いいえ
いいえ
割り込みﾒｯｾｰｼﾞをgAPICへ送信
108
終了

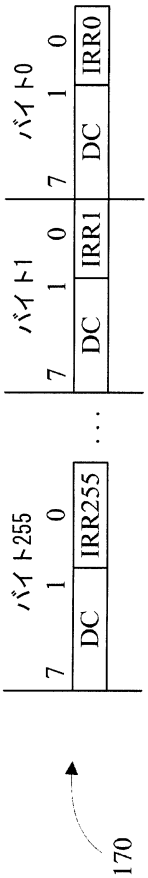
【図７】


開始
120
gAPICｹﾞｽﾄID＝０?
いいえ
gAPICｹﾞｽﾄID＝受信したｹﾞｽﾄID?
122
はい
126
論理割り込み且つ論理ID整合?
いいえ
いいえ
はい
物理割り込み且つ物理ID整合又はﾌﾞﾛｰﾄﾞｷｬｽﾄ?
はい
はい(非ｱｸﾃｨﾌﾞ)
いいえ
128
130
応答、承認
132
IRRを更新
応答、非承認
124
134
優先度を再評価
136
優先度に基づき割り込みの信号を送る
終了

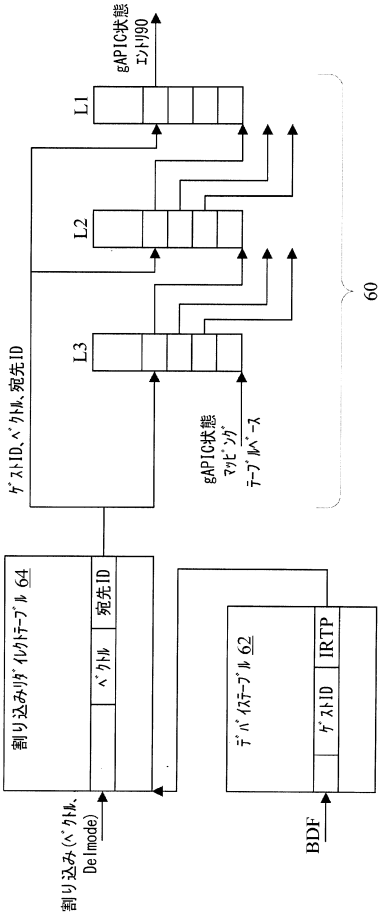
【図８】


開始
140
古いｹﾞｽﾄに対してgAPIC状態ｴﾝﾄﾘを決定
142
古いｹﾞｽﾄに対してgAPIC状態ｴﾝﾄﾘをｾﾞﾛにする
144
gAPICｹﾞｽﾄID、論理ID、及び物理IDをｸﾘｱ
146
gAPIC内へIRRのｱﾄﾐｯｸな論理和を取る
148
他のgAPIC状態をVMM管理gAPIC状態ｴﾝﾄﾘへ書き込む
非ｱｸﾃｨﾌﾞｹﾞｽﾄ
ｱｸﾃｨﾌﾞｹﾞｽﾄ
150
gAPICIRR、ISR、TMRをｸﾘｱ
152
新たなｹﾞｽﾄID、論理ID、及び物理IDをgAPIC内へ書き込む
154
新たなｹﾞｽﾄに対するgAPIC状態ｴﾝﾄﾘを決定
156
gAPIC状態ｴﾝﾄﾘからIRRを読み出す
158
gAPIC内へIRRのｱﾄﾐｯｸな論理和を取る
160
VMM管理gAPIC状態ｴﾝﾄﾘから他の状態を読み出す
162
他の状態をgAPICへ書き込む
終了

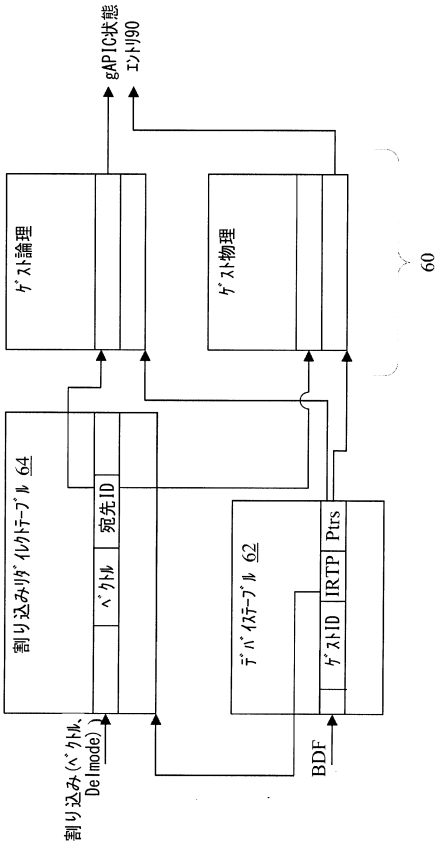
【図９】


バイト0
7
1
0
DC
IRR0
バイト1
7
1
0
DC
IRR1
バイト255
7
1
0
DC
IRR255
170

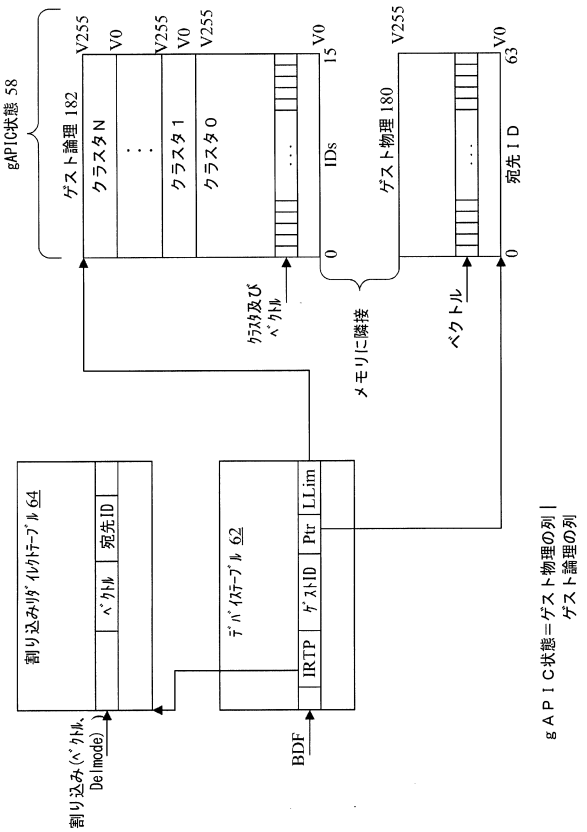
【図１０】



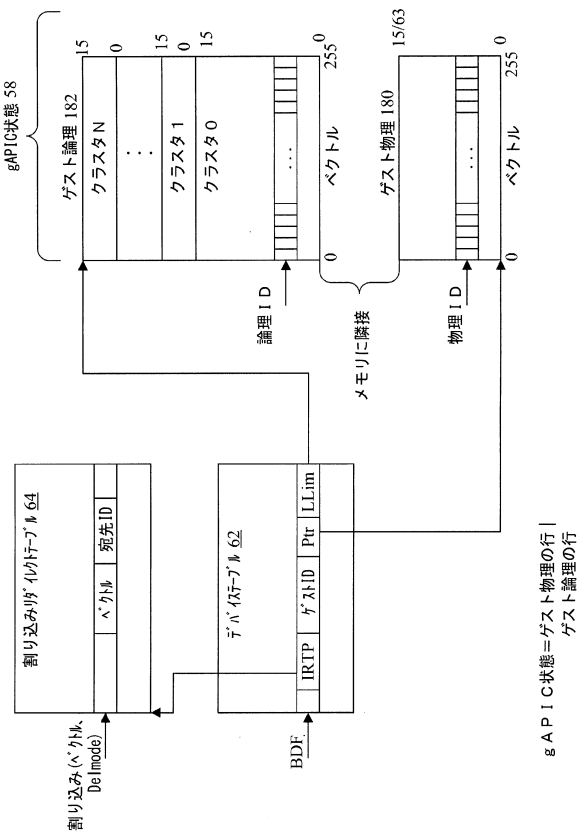
【図１１】



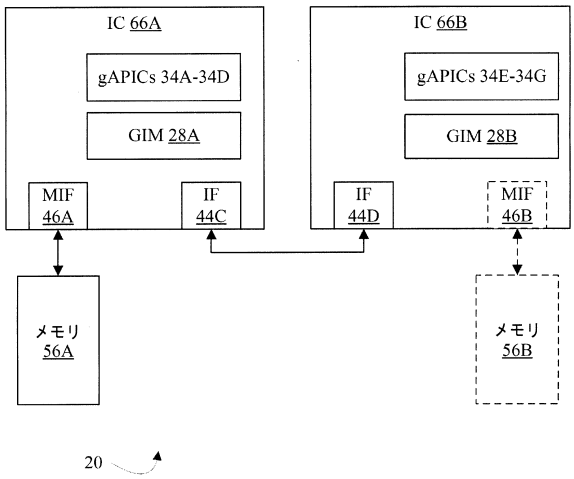
【図１２】



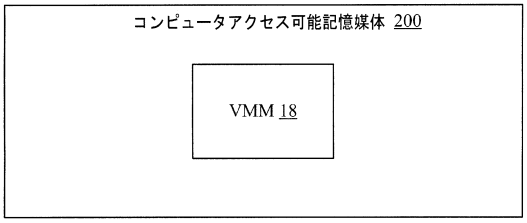
【図１３】



【図１４】


IC 66A
gAPICs 34A-34D
GIM 28A
MIF 46A
IF 44C
IC 66B
gAPICs 34E-34G
GIM 28B
IF 44D
MIF 46B
メモリ 56A
メモリ 56B
20

【図１５】


コンピュータアクセス可能記憶媒体 200
VMM 18

フロントページの続き

(31)優先権主張番号 12/611,607
(32)優先日 平成21年11月3日(2009.11.3)
(33)優先権主張国 米国(US)
(31)優先権主張番号 12/611,622
(32)優先日 平成21年11月3日(2009.11.3)
(33)優先権主張国 米国(US)

(74)代理人 100162156
弁理士 村雨 圭介
(72)発明者 ベンジャミン シー. セレブリン
アメリカ合衆国、カリフォルニア州 ９４０８６、サニーベイル、マホガニー レイン ７７１
(72)発明者 ドナルド ダブリュー. マコーレー
アメリカ合衆国、テキサス州 ７８７３４、レイクウェイ、エッジウォータ コーヴ １０２
(72)発明者 ジョン エフ. ウィーデルヒルン
アメリカ合衆国、カリフォルニア州 ９５１２５、サン ホセ、ルビノ ドライブ ＃１１５ ３１１０
(72)発明者 エリザベス エム. クーパー
アメリカ合衆国、カリフォルニア州 ９５０３３、ロス ガトス、ファーヴ リッジ ロード １８７６４
(72)発明者 マーク ディー. ヒュンメル
アメリカ合衆国 ０２０３８ マサチューセッツ州、フランクリン、スチュワート ストリート ６８

審査官 大塚 俊範

(56)参考文献 米国特許第０７２０９９９４（ＵＳ，Ｂ１）
特開昭６１－０３６８５０（ＪＰ，Ａ）
特表２００８－５３５０９９（ＪＰ，Ａ）
米国特許出願公開第２００４／０１１７５３２（ＵＳ，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ ９／４６－９／５４