

LilyVM と仮想化技術



榮樂 英樹

<hdk_2@users.sourceforge.net>

内容



- 背景
- LilyVM
- その他の VM およびユーザレベル OS
- CPU による VM のサポート

1 台の計算機で複数の OS を動かしたい



- OS の移行 (Mac OS 9 と Mac OS X)
 - 異なるシステムの組み合わせ
(Windows と NetBSD)
 - 動作テスト
 - 開発
- など

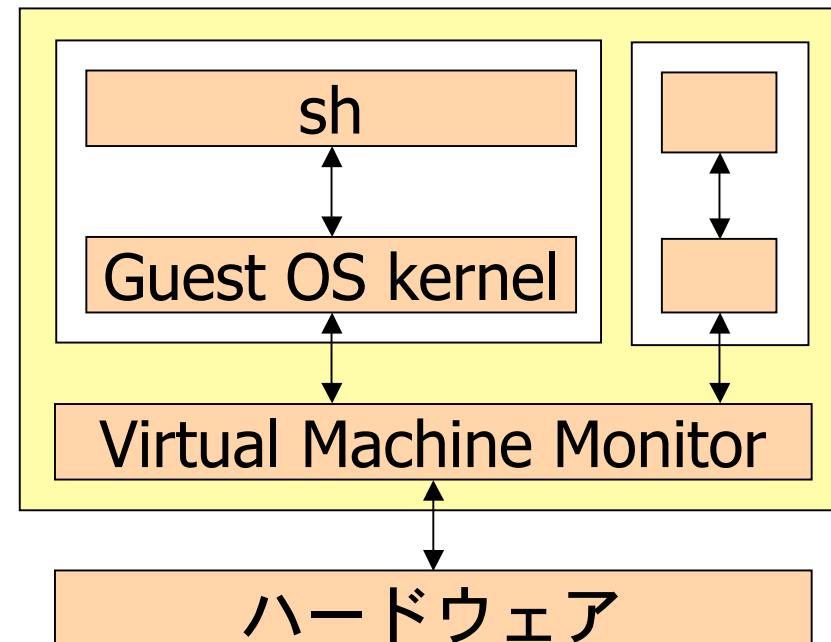
1 台の計算機で複数の OS を動かす方法



- 仮想計算機 (VM) (Type-I, Type-II)
- ユーザレベル OS
- VM + ユーザレベル OS

仮想計算機（VM）（Type-I）

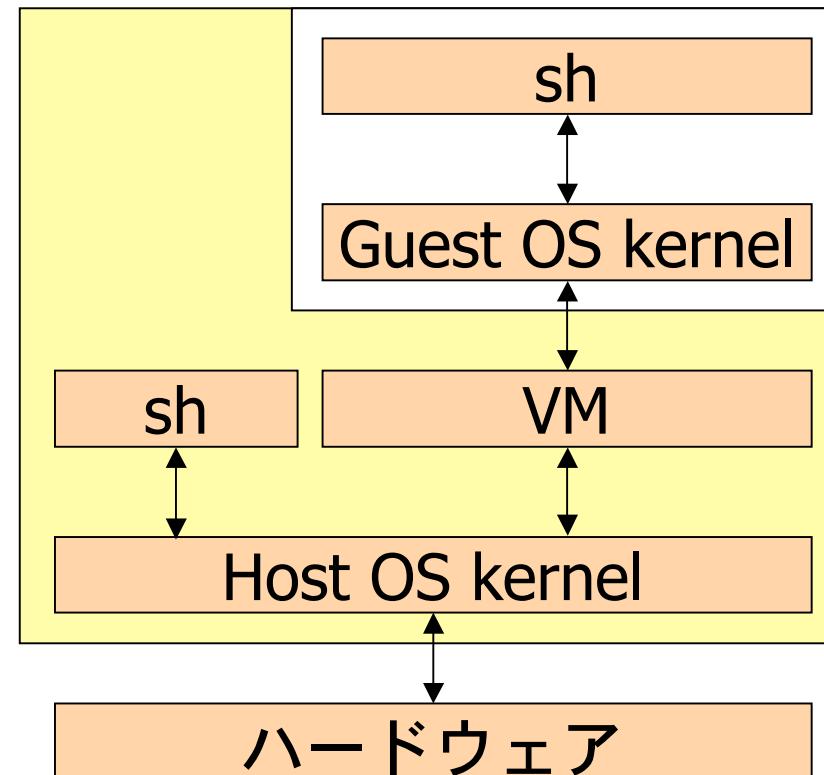
- ホスト OS が存在しないタイプの仮想計算機
- 例: VMware ESX Server, Xen



仮想計算機（VM）（Type-II）

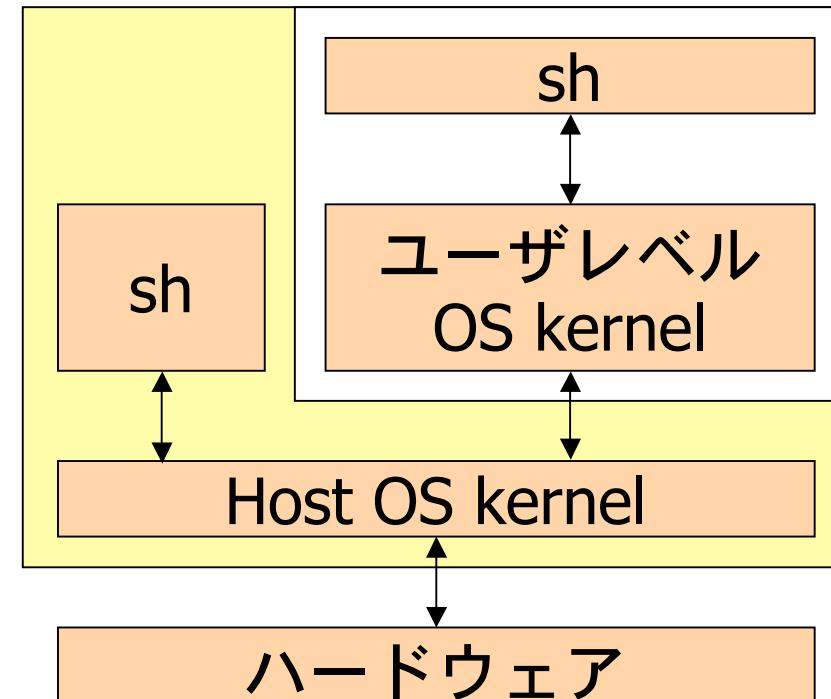
■ 実機用のOSをそのまま実行できる、下位層のシステム（ホストOS）から完全に独立した実行環境を提供する

■ 例: Bochs, VMware Workstation, QEMU



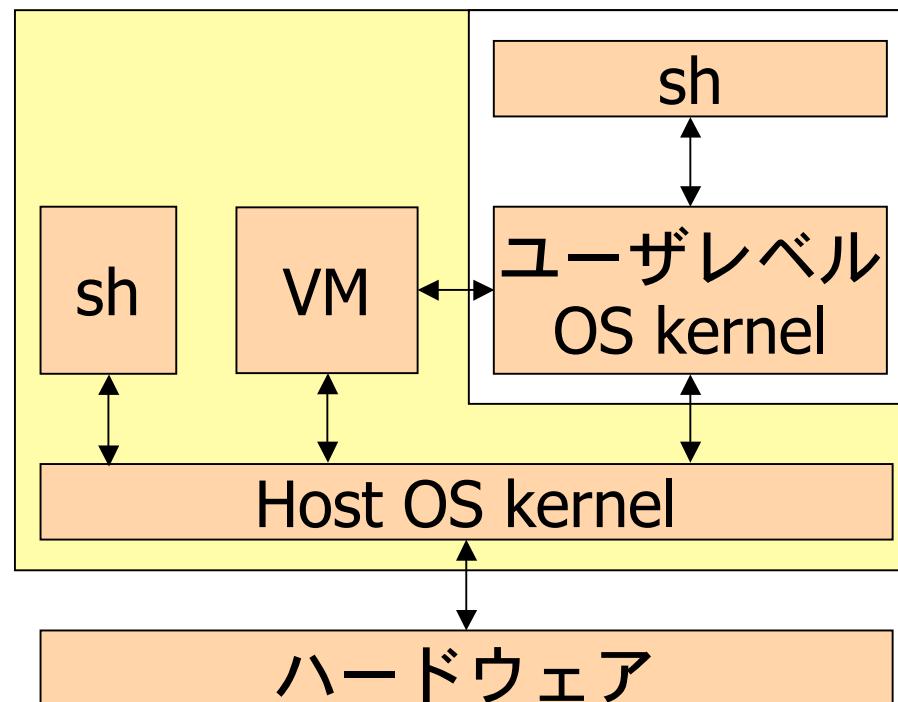
ユーザレベル OS

- OS をユーザプロセスとして実行する
- 実機用の OS を移植して作られている
- 例: User-mode Linux (UML)



VM + ユーザレベル OS

- VM が、特権命令や入出力装置などのうち、必要な部分だけをエミュレートする
- LilyVM [<http://lilyvm.sf.net>]
 - 特権命令などをコンパイル時にエミュレートするコードに書き換える
 - ユーザレベル OS カーネル: 自動変換 + わずかな修正
 - ホスト OS: NetBSD, FreeBSD, Linux
 - ゲスト OS: NetBSD, FreeBSD, Linux 2.4



NetBSD で LilyVM



- ホスト OS の ptrace() に以下の機能を追加する必要がある
 - PTRACE_SYSEMU
(または PTRACE_SYSCALL)
 - 制御レジスタ %cr2、エラーコード、およびトランプ番号を取得するための機能
- 本体への追加を希望します!

LilyVM の開発予定



- 高速化
 - 浮動小数点関連
 - MMU まわり
 - ネットワーク
- AMD64, Power PC 等のサポート

その他の VM およびユーザーベル OS



- VMware
- Bochs
- QEMU
- Plex86
- Xen
- User-mode Linux
- coLinux

比較

| | VM | | ユーザレベル OS | VM と機械語命令の変換 | | その他 | |
|-------------------|--------|----------------|--------------------|--------------|--------|-------|---------|
| | VMware | QEMU, Bochs | User-mode Linux | LilyVM | Plex86 | Xen | coLinux |
| VM が提供するアブストラクション | 実機 | 実機 | なし | 実機 | 実機 | Xen | なし |
| ゲスト OS の修正 | 不要 | 不要 | 移植(大) | 自動 | 自動 | 移植(小) | 移植(小) |
| ホスト OS との隔離 | 大 | 大 | 小 | 小 | 大 | 大 | なし |
| 仮想化のオーバヘッド | ? | 特大 | なし | 中 | 大 | 小 | 小 |
| ホスト | x86 | x86, PPC 等 | x86, PPC 等 | x86 | x86 | x86 | x86 |
| ライセンス | 商用 | LGPL | GPL | GPL | LGPL | GPL | GPL |

VMware

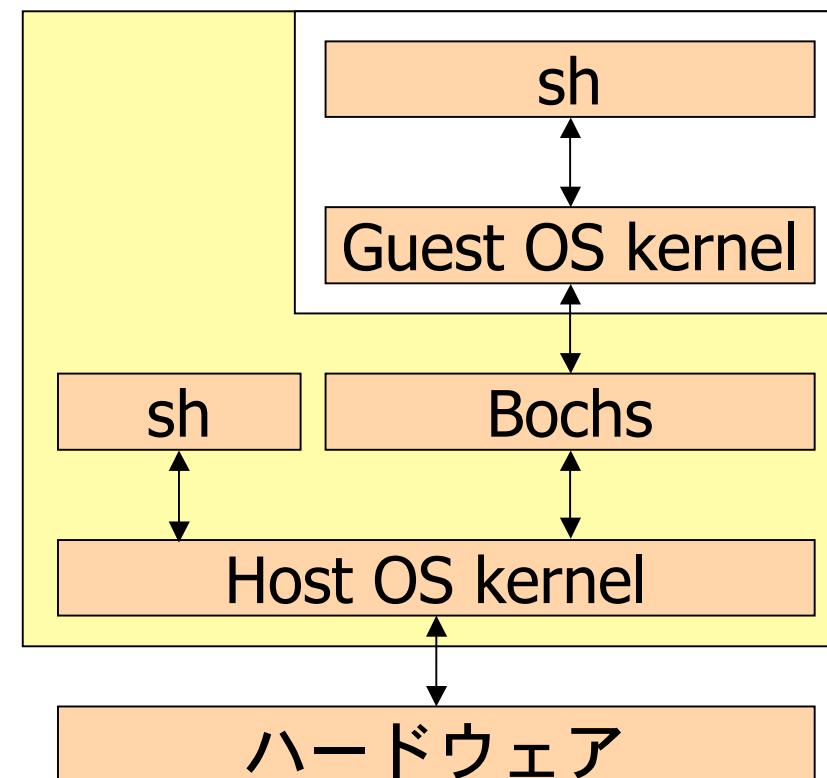


- x86 の VM ソフトウェアの老舗
- 詳細不明
 - ring 3 は元の機械語命令をそのまま実行しているという噂だが、ring 0 やリアル・モードも速いし...

Bochs

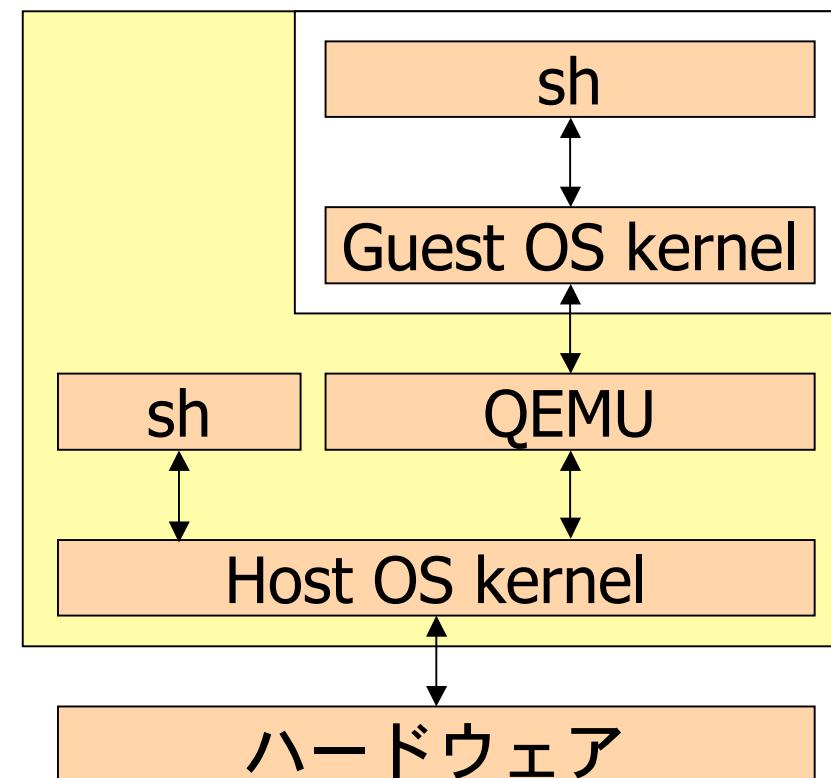


- CPU および周辺装置をエミュレートする
- 移植性が高い
- オーバヘッドが非常に大きい



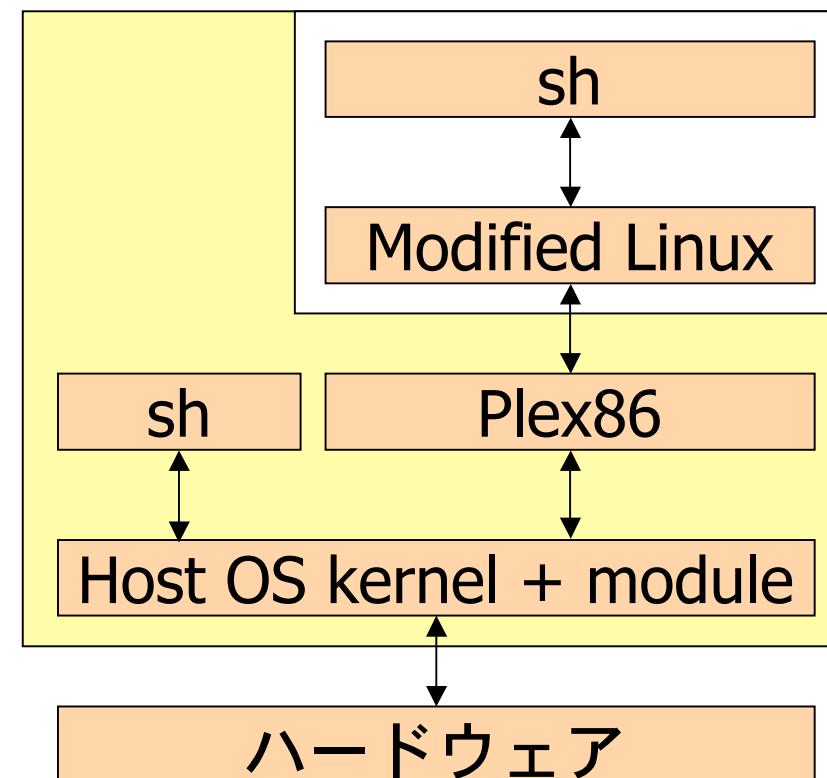
QEMU

- CPU および周辺装置をエミュレートする
- 機械語命令を動的に変換することによって高速化している
- カーネルモジュールによりさらなる高速化が進められている



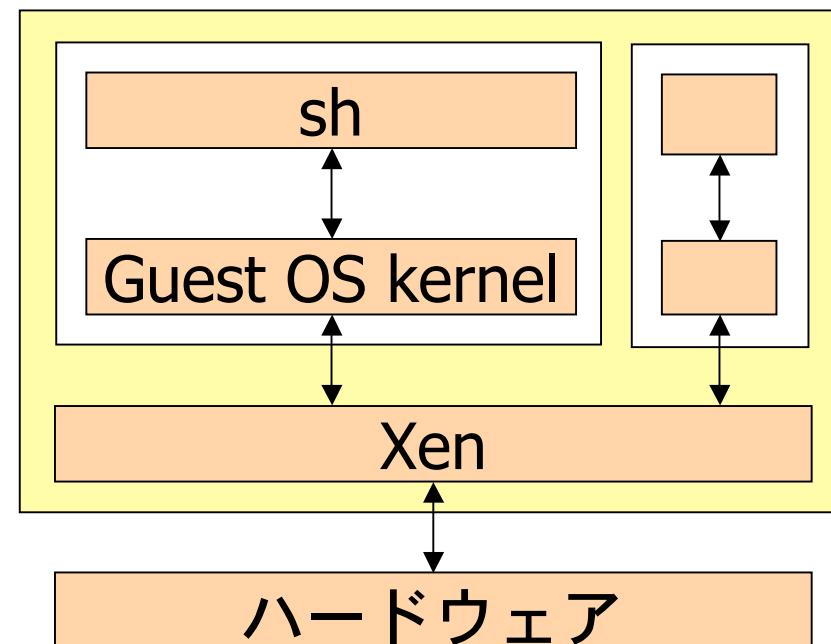
Plex86

- CPU の一部と周辺装
置をエミュレートす
る
- ゲスト OS の機械語
命令を静的に変換し
ている
- カーネルモジュール
によりメモリ空間の
切り替えなどを行う



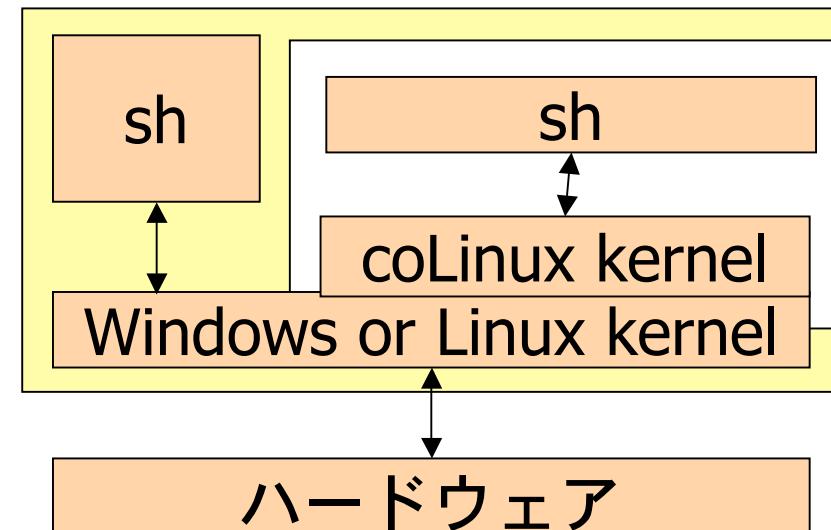
Xen

- ゲスト OS を Xen アーキテクチャに移植している
- 周辺装置の制御のために Linux をひとつ使う



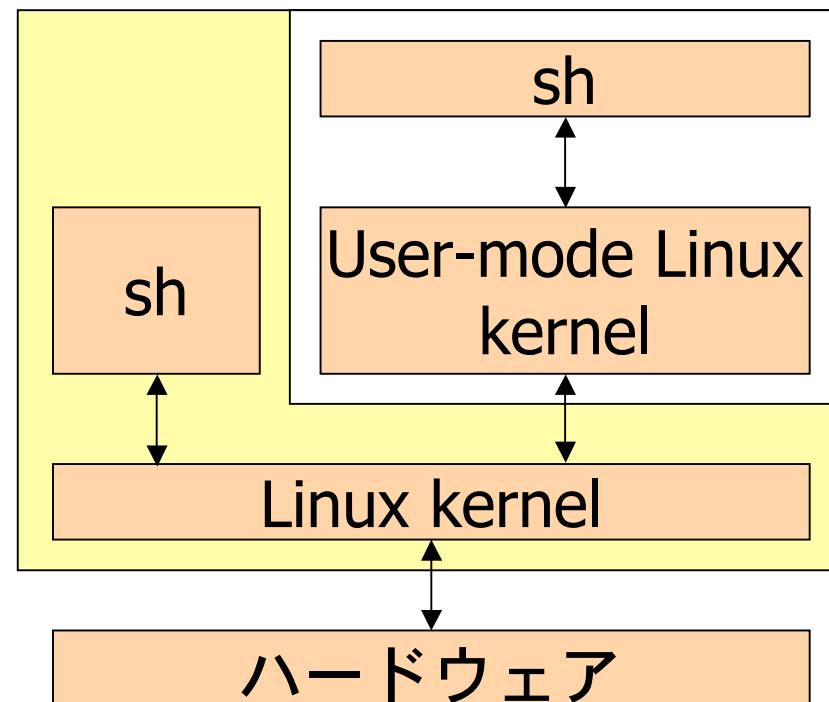
coLinux

- ホスト OS のデバイスドライバとして Linux を動作させる
- MMU や割り込みに関する部分が多く修正されている
- 性能は非常に良い



User-mode Linux

- Linux 上でプロセスとして動作するように Linux が移植されている
- ptrace() など OS に依存する機能を使っていて移植性が低い



比較

| | VM | | ユーザレベル OS | VM と機械語命令の変換 | | その他 | |
|-------------------|--------|----------------|--------------------|--------------|--------|-------|---------|
| | VMware | QEMU, Bochs | User-mode Linux | LilyVM | Plex86 | Xen | coLinux |
| VM が提供するアブストラクション | 実機 | 実機 | なし | 実機 | 実機 | Xen | なし |
| ゲスト OS の修正 | 不要 | 不要 | 移植(大) | 自動 | 自動 | 移植(小) | 移植(小) |
| ホスト OS との隔離 | 大 | 大 | 小 | 小 | 大 | 大 | なし |
| 仮想化のオーバヘッド | ? | 特大 | なし | 中 | 大 | 小 | 小 |
| ホスト | x86 | x86, PPC 等 | x86, PPC 等 | x86 | x86 | x86 | x86 |
| ライセンス | 商用 | LGPL | GPL | GPL | LGPL | GPL | GPL |

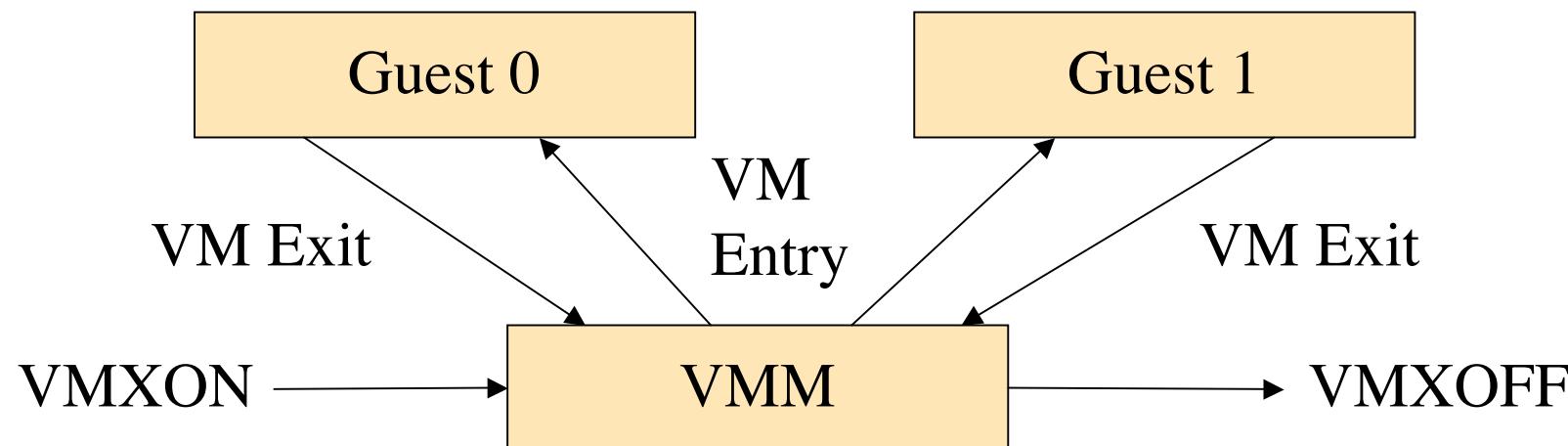
CPU による VM のサポート



- 背景
 - VM の必要性が高まっている
 - 従来の IA-32 における VM 開発は難しい
(センシティブな非特権命令の存在等)
- VMM の実現を容易にするための IA-32 の拡張
 - Intel Virtualization Technology
 - AMD64 Pacifica

Intel Virtualization Technology

- Virtual Machine Control data Structure (VMCS) : 制御レジスタや状態などを含む
- MMU 関連: physical-address extension ?



Intel Virtualization Technology Specification for the IA-32 Intel Architecture
(C97063-002) Figure 1-1 より引用

AMD64 Pacifica



- Intel のものとたよな機能
のはず
- Intel のものとの互換性がないらしい

CPU による VM のサポートの ポイント



- 命令エミュレーションは基本的に不要
- I/O 命令およびハードウェア割り込みのエミュレーションは必要

CPU による VM のサポートで 何が変わるか: LilyVM



- 新機能がユーザレベルから使えれば...
 - 大幅高速化?
 - I/O 命令は今の書き換える方法のほうが高速な可能性がある
- ホスト OS が新機能をうまくユーザレベルに提供できるかどうか
 - ゲスト OS からのホスト OS の呼び出し等

CPU による VM のサポートで 何が変わるか: その他の VM



- VMware: 性能向上
- Bochs: 今まで通り
- QEMU: x86 on x86 における性能向上?
- Plex86: ?
- Xen: 性能向上?、Windows 対応
- User-mode Linux: 今まで通り
- coLinux: 今まで通り

まとめ



- LilyVM [<http://lilyvm.sf.net>]
 - 機械語命令の書き換え
 - NetBSD への機能追加
- その他の VM およびユーザレベル OS
- CPU による VM のサポート