



仮想サーバホスティング

株式会社クララオンライン
白畑 真



自己紹介

- クララオンラインについて
 - 専用サーバ/仮想サーバを中心とするWebホスティング事業者
 - 2003年よりLinuxベースの仮想サービスを提供
-

主要な仮想化方式

- 仮想マシン (Virtual Machine):
 - VMware, MS Virtual Server, Parallels, Xen
- 準仮想化 (Paravirtualization):
 - Xen, UML (User Mode Linux)
- **OSレベル仮想化**
 - 同一のKernelを用いてOSを仮想的に分割する技術
 - 独立したプロセス空間、ファイル空間、ユーザ管理
 - 主要なOSレベル仮想化ソフトウェア
 - Parallels Virtuozzo Containers, OpenVZ, Linux-Vserver, Solaris Containers (Zones), FreeBSD Jail
 - 仮想サーバの呼び方
 - Container, Virtual Environment, VPS (Virtual Private Server), Jail..



典型的なホスティング事業者の サーバと仮想化技術



一般的なホスティングプロバイダの主要サービス

- 共用ホスティング
 - 単一のハードウェアで一つのサーバを稼働させるが、プロトコル上の識別子を用いて領域を分割; 複数顧客に提供
- 仮想サーバホスティング
 - 単一のハードウェア上に複数の仮想サーバを構築し、各仮想サーバを顧客に提供
- 専用サーバホスティング
 - 単一のハードウェアを顧客に提供

共有ホスティング

- 共有ホスティング
 - 単一のサーバを複数顧客で共有
- プロトコルに内包された識別子を利用
 - Webサーバやメールサーバなどのサーバアプリケーションがドメイン名を認識し、それに基づきサービスを提供
 - SMTP/POP3: メールアドレス
 - HTTP: Name Based Virtual Host

HTTP/1.1のHost:ヘッダ

example.jp	example.com	example.net
サーバアプリケーション (Webサーバ)		
OS		
ハードウェア		

仮想サーバホスティング

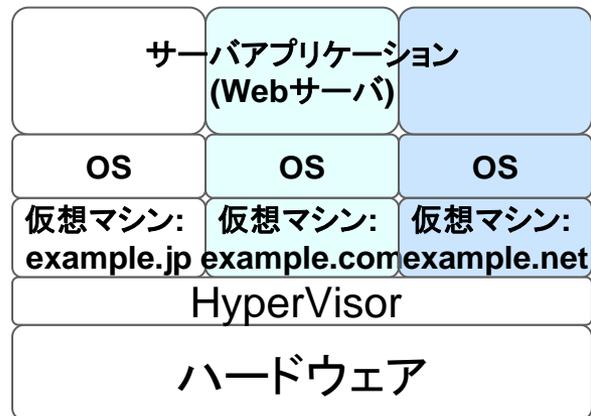
仮想サーバ(VPS)

単一の物理サーバ上で複数の仮想サーバを稼働;
各仮想サーバを一顧客で占有

OSレベル仮想化方式:



仮想マシン方式:



仮想サーバサービスと仮想化技術

- 各社ホスティングサービスで利用されている仮想化技術
 - OSレベル仮想化方式
 - [Virtuozzo] クララオンライン, FirstServer, ソフトバンクIDC, Livedoor, tsukaeru-net
 - [OpenVZベース] NTT-PC, Verio (RapidSite)
 - [FreeBSD VPS] Verio (RapidSite, OCN)
 - 仮想マシン方式
 - [Xen] @YMC, at+link
 - [Virtual Iron] at+link

出典: Hosting Pro 2008 講演資料、Webより筆者作成

➔ホスティングではOSレベル仮想化が主流



OS仮想化について



OSレベル仮想化とは何か

システムコールに
対するラッパー群



OSレベル仮想化方式と仮想マシン方式 Pros と Cons

- OSレベル仮想化のPros/仮想マシンのCons
 - サーバ密度、サーバ密度、サーバ密度
 - Virtuozzoの場合には仮想マシンの3倍(公称)
 - OSレベル仮想化のPros/仮想マシンのCons
 - OSの選択肢
 - メモリ管理上の制限
 - 低レベルな処理の制限
-

Pros: サーバ密度

- (VMとは違い)仮想サーバとハードウェア上でメモリアドレスの変換処理が不要
 - オーバーヘッドが低い
 - テンプレート機能: 複数の仮想サーバに対して、同一のプログラムをインストールできる
 - 同一のファイルを参照する(ハードリンク)ので、ハードディスクの容量を節約できる
 - メモリの一部(テキストセグメント)を共有できるため、メモリを効率的に利用できる
-

Cons: OSの選択肢

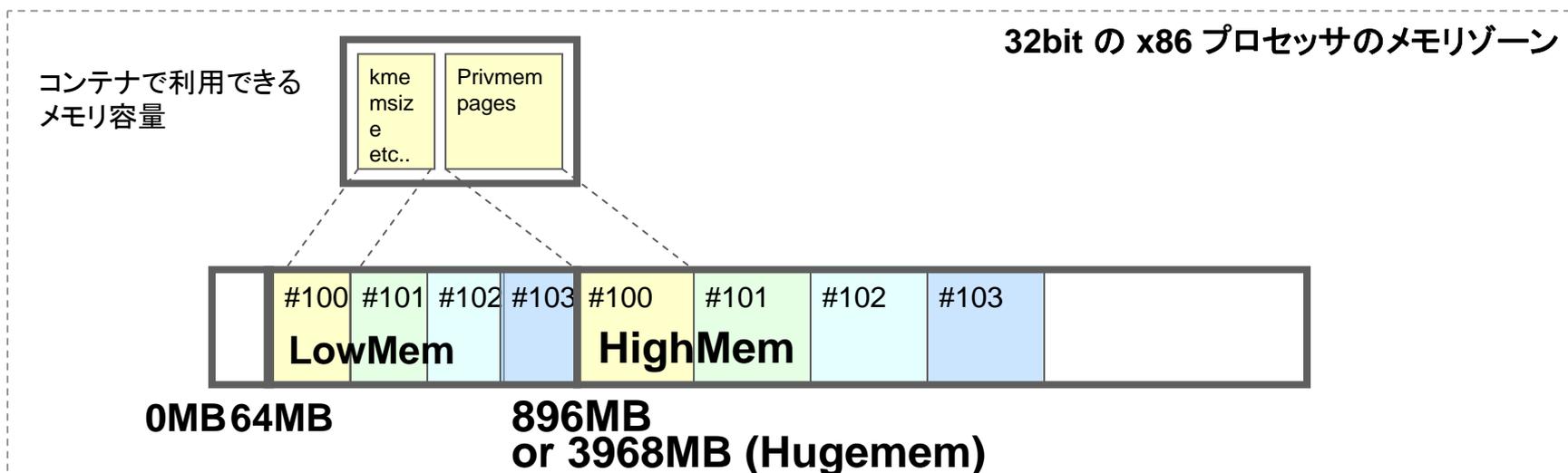
- 同一Kernelで異なる仮想サーバを実現
 - 同一ハードウェア上で異なるOSを利用することはできない
 - 同じkernelでよければ
 - [Linuxの場合] 異なるディストリビューション
 - [Solarisの場合] 異なるバージョン
- を共存できる
-

Cons: Virtuozzoのメモリ管理

- OSレベル仮想化
 - 仮想サーバと物理サーバのメモリアドレスを変換しない
- Virtuozzoのメモリ管理方式
 - UBCモード
 - 単純にOSのメモリを分割
 - HIGHMEMとLOWMEMをそれぞれ分割して管理するため、使い勝手が異なる部分がある
 - SLMモード
 - メモリの割り当てを動的に決定
 - 仮想マシン方式に近い感覚で利用できる

Cons: Virtuozzoのメモリ管理 - UBC

- UBC (User Bean Counter) モード
 - メモリの領域などのリソースを細かい粒度で分割し、それぞれのパラメータごとに保証値と制限値を設定
 - 確実なリソースの割り当てを保証
 - メモリ、プロセス数などのリソース割り当ての閾値を超えると、新規リソースの割り当てが全て拒否される
 - Cannot allocate memory
 - SSHできない



Cons: Virtuozzoのメモリ管理 - SLM

- SLM (Service Level Management) モード
 - メモリを容量でざっくり管理
 - 細かいパラメータでの管理は行わない
 - HIGHMEM/LOWMEMの割り当ては自動
 - LOWMEMが不足する可能性はあり得る
 - 通常のLinuxのメモリ管理方法とほぼ同じ
 - メモリが不足した場合、OOM Killerがプロセスを強制終了してメモリを解放

専用サーバ

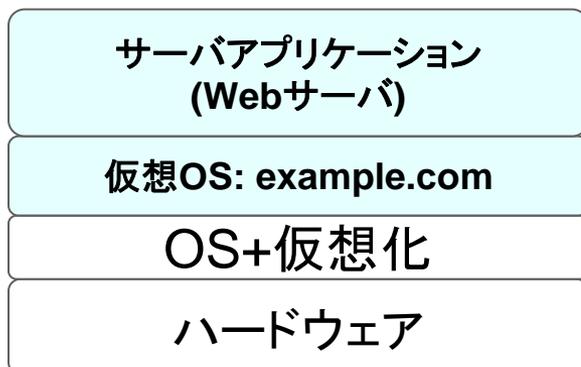
専用サーバ+仮想化技術

単一の物理サーバ上で単一の仮想サーバを稼働

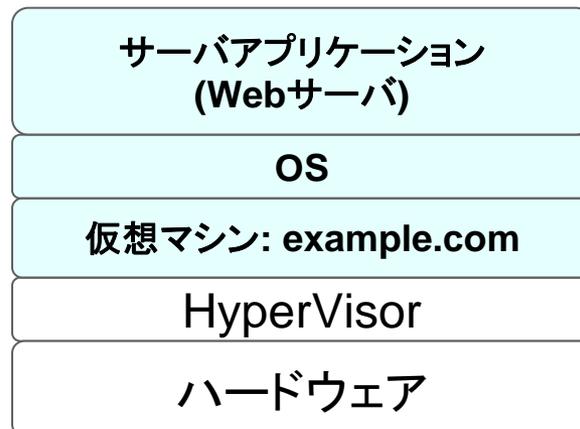
マイグレーションによりハードウェア変更が容易

ファームウェア更新やハードディスク増設時の
ダウンタイム削減

OSレベル仮想化方式:



仮想マシン方式:





仮想ホスティング運用にまつわる 5つのトピック

リソース割り当て: 総論

- 制御できる典型的なリソース
 - CPU時間
 - メモリ
 - ファイル容量
 - リソース制限を超過した場合の振る舞いに注意
- 性能保証かオーバーコミットか
 - CPU: 2GHz Quad Core, メモリ: 8GBのハードウェアの場合
 - 単純に4分割すると2GHz CPU, 2GBのメモリ
 - 多くの場合、ハードウェアの総処理能力以上を仮想サーバに割り当てても問題は発生しない
- クラウドオンラインの場合:
 - 従来はベストエフォート型サービス(VPS)と保証型サービス(FPS)を提供
 - 現在はベストエフォート型サービス(FlexMini)のみを提供

リソース割り当て: I/O帯域幅

- I/O帯域幅は有限な資源
 - 各ハードウェア/ディスク毎に性能に限り
 - 特定の仮想サーバがハードウェアのI/O帯域幅を占有した場合、コントロール手段に乏しい
 - DBMSへの集中的なアクセスなど
 - 仮想サーバのCPUサイクルやメモリが保証されている場合でも、総合的にみたパフォーマンスが低下する恐れがある
 - 大数の法則である程度はカバー
 - 専用サーバでは問題なし
-

マイグレーション

- 計画メンテナンスの効率化
 - ハードウェアのファームウェア更新
 - 代替ハードウェアへの変更
- Virtuozzoの場合
 - 共有ディスクが不要
 - マイグレーションの流れ
 1. ファイルシステムのコピー
 2. メモリ領域のダンプ
 3. メモリ領域のコピー
 4. メモリ領域のレジューム
 - ノートパソコンのサスペンド・レジュームに近い
 - OSは再起動しないが厳密には少し止まる
 - メモリの量に依存

ハードウェア

- 専用サーバと仮想サーバ
 - 同じ金額のサービスでも、仮想サーバの方が相対的に良いハードウェアを利用できる
 - RAIDによるディスク冗長化/電源冗長化
 - ハードウェアに障害が生じた場合のインパクトも大きい
 - 数十サーバが実用的な限界？
 - 海外では数百サーバを単一ハードウェアで運用している例もある
-

OSレベル仮想化技術利用に伴う制限

- 機能上の制約が多い
 - 多くの機能が仮想化機能を有するドライバにより提供される
 - NIC、ファイルシステム
 - 全般的に低レベルのアクセスに制約
 - 例: LKM (Loadable Kernel Module)の利用
 - 特定の仮想サーバに対して/dev/以下のデバイスに対する排他的アクセスを設定可能な場合も
 - 仮想化技術によってはパケットフィルタ、経路表などの利用に制限がある
 - パケットキャプチャやdhcpdの運用など
 - Virtuozzoの場合にはiptablesと経路表を独自に持てる
 - 仮想サーバ同士の相互接続があまり想定されていない
 - メモリ割り当て方式が特殊
 - ただし仮想マシンと異なり実マシンと仮想マシンのアドレス変換のオーバーヘッドなし

キャパシティプランニング

- ホスティングではお客様の利用形態が事前にわからない
 - ➔キャパシティプランニングが難しい
- 基本的に大数の法則でカバー
 - 各ハードウェアごとの負荷が一定水準に収まるように収容数を調整
- 専用サーバでは問題なし

まとめ

- 仮想化技術全般について
 - コントロールできるリソースとコントロールできないリソースがある
 - リソースの制限値を超過した場合のふるまいについて調査する
 - OSレベルの仮想化について
 - 向き不向きがある
 - VMと比べると制約がかなり多い
 - とにかく早くてサーバ集約度が高い
-