

クラウドコンピューティング 環境における資源共有

情報通信研究機構
河合栄治

クラウドコンピューティング

- クラウドコンピューティングの目的
 - 経費効率の向上
 - 規模拡張性の向上
 - 迅速性の向上
- クラウドコンピューティングの技術的本質
 - ICTインフラにおけるより一層の仮想化の促進
 - 並列分散処理技術の高度な応用
- ICTインフラ開発の歴史は仮想化の歴史
 - ホスト仮想化
 - ストレージ仮想化
 - ネットワーク仮想化

ホスト仮想化

- **タイムシェアリングシステム**
 - システムを時分割多重化して、（擬似的な）同時共有を実現
 - マルチプログラミング（プロセス多重）が可能に
 - OS環境上で、そのOSイメージを独立して各プロセスに提供
- **ホスト仮想化技術**
 - ホスト環境上で、そのホストイメージを独立して各OSに提供
 - タイムシェアリングシステムの一つ
 - IBMメインフレームの技術
 - 基礎技術の確立は非常に古い！！
 - マイグレーションも可能に
 - ホストイメージは状態管理が比較的容易
 - プロセスマイグレーションは技術的には可能だが、非常に複雑な状態（OS環境）の転送が必要

ストレージ仮想化

- **Unixデバイスドライバモデル**
 - ファイルシステムにデバイス管理、I/Oを統合
 - メジャーID、マイナーID
 - read()/write()
- **VFS（仮想ファイルシステム）**
 - ファイルシステムインタフェースの定義（FSの仮想化）
- **ネットワークファイルシステム**
 - RPC誕生の要因となる
- **RAID**
 - 複数の二次記憶装置を単一の高性能・高信頼装置に見せる
 - ストレージデバイスの仮想化
 - LVM（論理ボリューム管理）
 - OSにおけるブロックデバイス管理の仮想化
- **SAN（FC、iSCSI、FCoEなど）**
 - 二次記憶接続をネットワーク拡張（ストレージ接続の仮想化）

ネットワーク仮想化

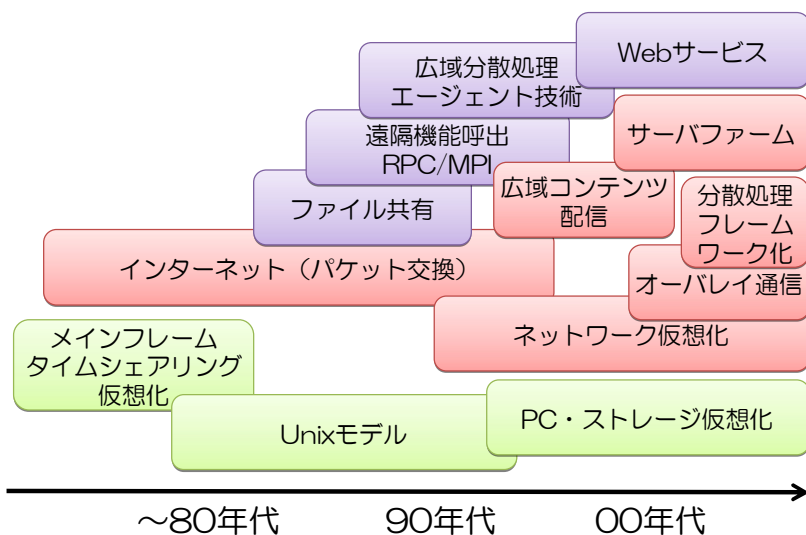
- **パケット交換網（回線の仮想化）**
 - タイムシェアリングシステムがヒントになった
 - インターネット成功の最大の要因
 - イーサネット、IPルーティング、E2E原則
- **VLAN技術（回線網の仮想化）**
 - L2/L3VLAN
 - XXX over YYY（再帰ネットワーク・トンネリング）
 - MPLS、λ
 - 広域ネットワーク仮想化
- **オーバーレイ通信**
 - 広域ネットワーク網におけるアプリケーションセマンティクスまで含めて仮想化

2010/2/19

テストベッドWG

5

クラウドを実現する要素技術



2010/2/19

テストベッドWG

6

リモートリソース共有



2010/2/19

テストベッドWG

7

メモリ共有のアプローチ

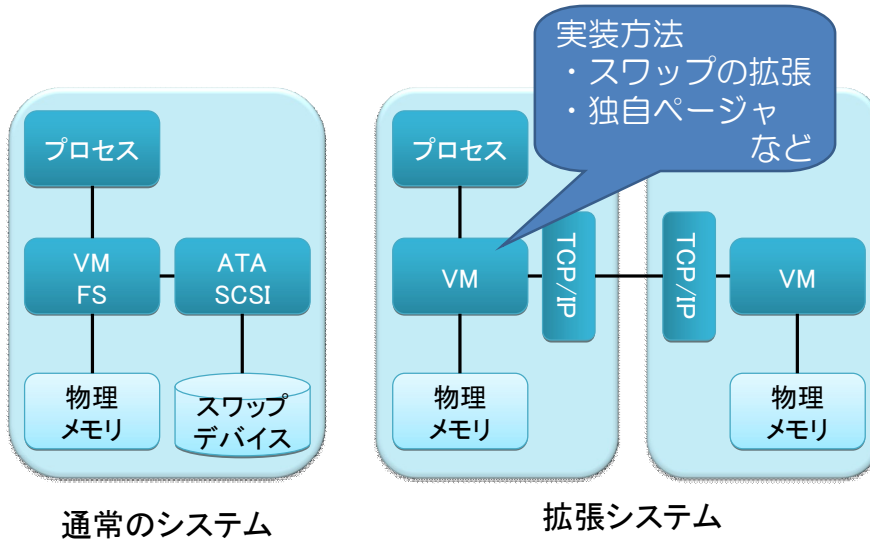
- VMマイグレーション
 - 空きメモリの多いホストにマイグレーション
 - 利点：導入が容易（多くのVMMに実装済み）
 - 欠点：マイグレーションコストが高い
- 共有メモリプログラミング
 - 共有メモリAPIによるプログラミング
 - 利点：柔軟性が高い（プログラミング記述能力）
 - 欠点：実装コストが高く、中央集権的管理が難しい
- リモートメモリマッピング
 - リモートホストのメモリ領域を仮想メモリ空間にマップ
 - 利点：中央集権的管理が可能
 - 欠点：導入が容易でない（デファクト技術がない）

2010/2/19

テストベッドWG

8

リモートメモリマッピング

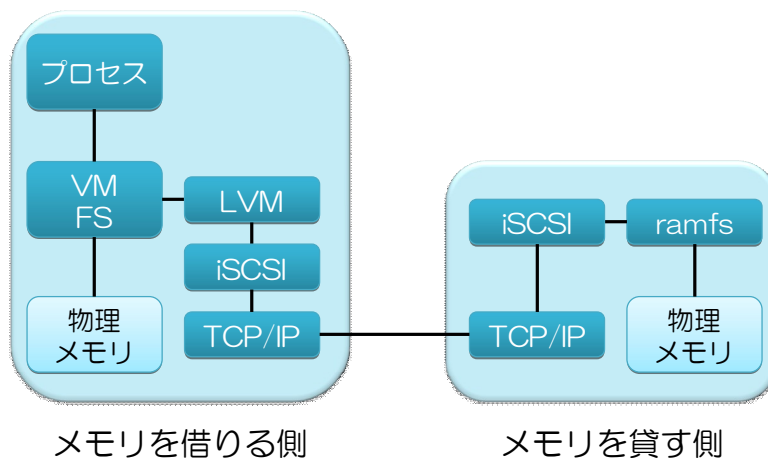


2010/2/19

テストベッドWG

9

Ramfs、LVM、iSCSIの活用



2010/2/19

テストベッドWG

10

特徴

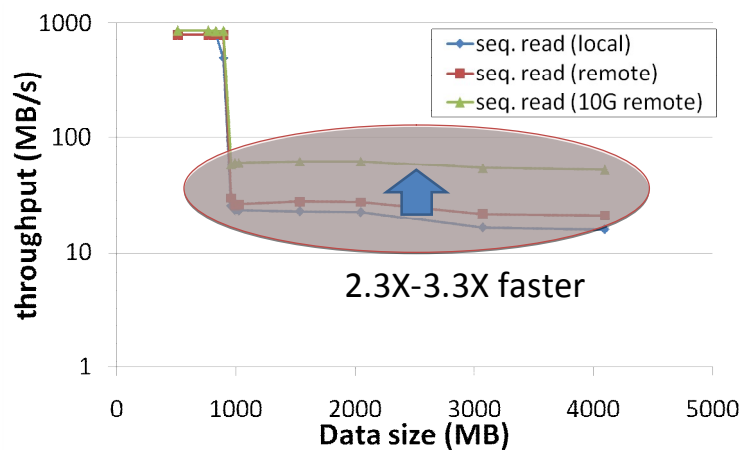
- デファクト標準技術を最大限活用
 - 多くのOSでメモリファイルシステム、論理ボリューム管理機構、iSCSI機能は稼働する
 - マルチOS環境でも動作する
 - 長期間サポートされ、技術的絶滅の恐れが少ない
- 非常に柔軟な設定が可能
 - LVMにより、動的なスワップデバイス切り替えも可能
 - ローカルHDDからリモートメモリへの切り替え
 - RAID構成も可能に
- VMM層への組み込みも可能
 - 統合管理フレームワークを実装可能
 - ゲストから透過的な管理も可能
 - VMMが動的メモリ割り当てを実装している場合 (Xen等)

2010/2/19

テストベッドWG

11

性能：連続アクセス (read)

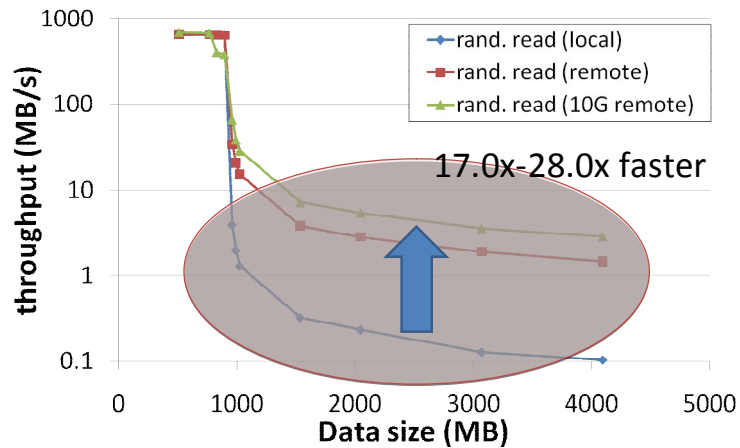


2010/2/19

テストベッドWG

12

性能：ランダムアクセス（read）



2010/2/19

テストベッドWG

13

今後の課題

- 統合的なネットワーク共有メモリ管理機構
 - クラウドコンピューティングにおけるメモリ共有管理
 - メモリが空いたらリモートに貸与し、メモリが足りなくなったらリモートから回収および借受するサイクルの確立
- 性能のチューニング
 - スワップはHDDへのアクセス性能向上にフォーカスされている
 - SSD向け最適化が今後入ってくると予想される
 - ランダムアクセスに強く広帯域な特性はネットワークメモリ共有でも生かされる
 - スワップ機構とiSCSIのインタラクション
 - プロファイリング

2010/2/19

テストベッドWG

14

まとめ

- ICTインフラ開発の歴史は仮想化の歴史
 - ホスト、ストレージ、ネットワーク、分散処理、サービス…
- 仮想リソースの遠隔利用の可能性
 - ストレージはかなり実用性が高まってきている
 - FC、iSCSI、FCoE
 - メモリについてもデファクト技術の応用で可能に
 - CPUはまだマイグレーションが基本
- 今後はリソース管理の面的展開
 - オペレーションの仮想化が鍵