

データセンター クラウド化に向けて



IIJ Technical WEEK 2010
IIJイノベーションインスティテュート
技術研究所 島慶一

今日の話題

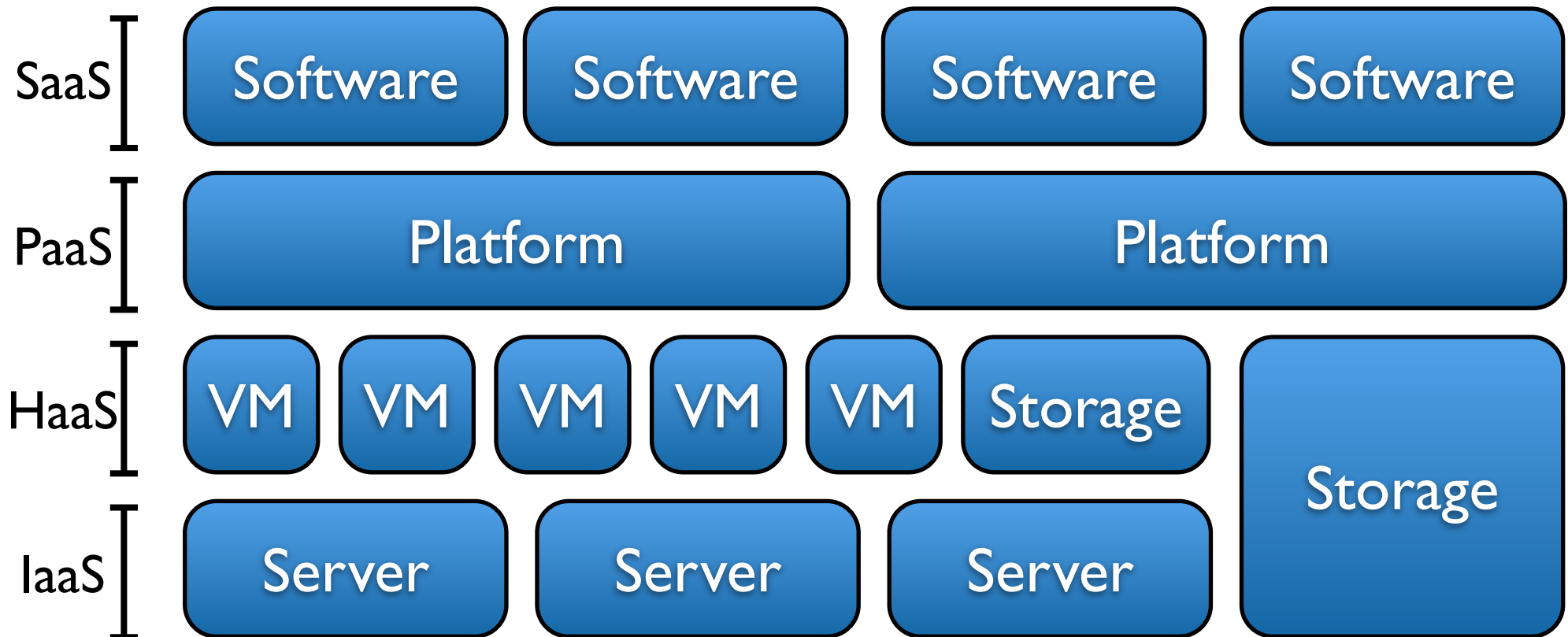
- クラウドデータセンター?
- クラウドデータセンターに必要な技術
- クラウドデータセンターに必要なネットワーク技術

クラウド?

- 人それぞれの定義
- *aaS (*=Software, Platform, Hardware, Infrastructure)
- クラウドと呼ばれているサービス例
 - Salesforce, Amazon Web Services, Google Apps, Google App Engine, Amazon EC2/S3, ニフティクラウド, IIJ GIO, ほかにもたくさん

仮想化レイヤー

以下のような構成で作られることが多い



データセンターの意義

- 安定した運用
 - 電源、耐震、室内環境など
- 運用コストのオフロード
- 高速なネットワーク接続

データセンターデザイン

- 旧くは単にハードウェアを設置して運用する場所としてデザイン
- ラック単位の貸し出し、電源容量、ネットワーク速度に応じた課金
- 仮想化が実用的になってきてからは、サービスの仮想化が進む
- Virtual Hosting, Jail, Virtual Private Server

利用者にとっての仮想化

- 迅速な仮想機材確保による機動性の高い顧客サービス
- サービス提供者のコストダウンの結果を受けた、安価なサービス
- 障害時の迅速な切り換え、復旧

運用者にとっての仮想化

- 統一機材の納入によるコストダウン
- 運用システムの統一による低負荷運用
- 物理資源共有によるコスト競争力の獲得
- 迅速なサービス開始、復旧などによる
SLA向上

データセンターの競争力

- 安くて信頼性の高いデータセンターの要求
 - 大規模化によるコストダウン
 - Google, Appleなどはニュースになるほど
 - 外気冷却などの活用によるコストダウン
 - 自動化によるコストダウン
 - 遠隔DCのような

データセンターの場所

- 場所の制約に影響する項目
 - コスト
 - アクセスのよい場所は高い
 - (オペレータの)利便性
 - ポリシー
 - どこにデータを置いてもよい、というわけでもない

データセンター仮想化の目的

- 大規模化が困難な状況での資源の集約
- 状況に応じた資源利用による稼働率向上
- 資源再配置を利用した保守性の向上
- 大規模障害時の復旧支援

仮想化に必要な技術

- CPU、ネットワーク、ストレージ資源の自由な配置、再配置
- データセンター間を繋ぐ高速通信
- 広域ネットワークで動作する分散ストレージ

仮想化に必要な技術

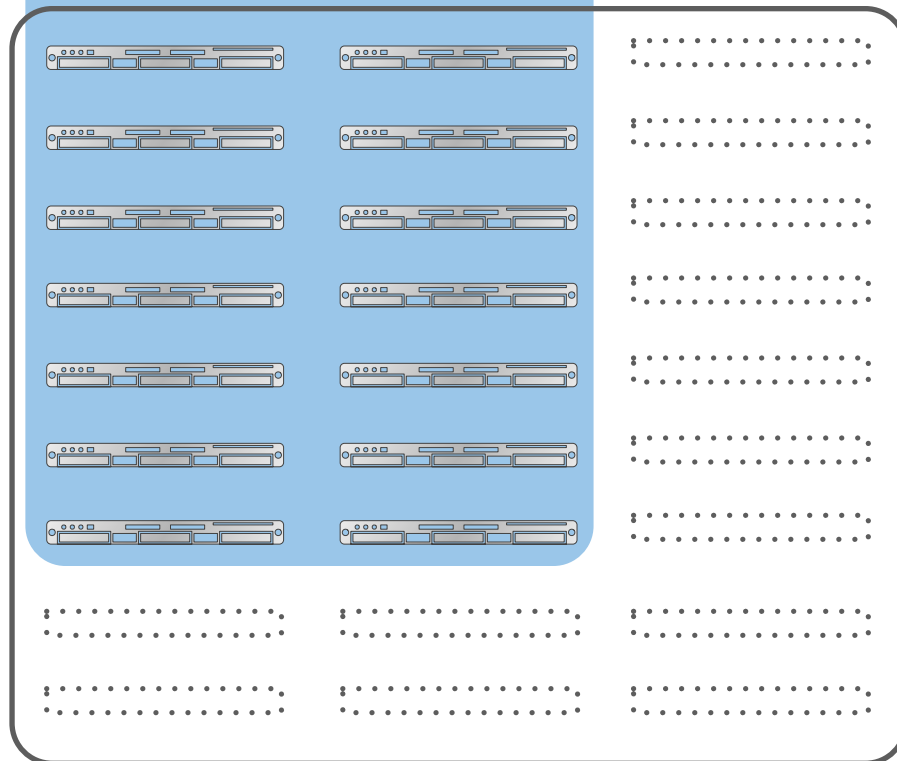
- CPU、ネットワーク、ストレージ資源の自由な配置、再配置
- データセンター間を繋ぐ高速通信
- 広域ネットワークで動作する分散ストレージ

クラウドと仮想化

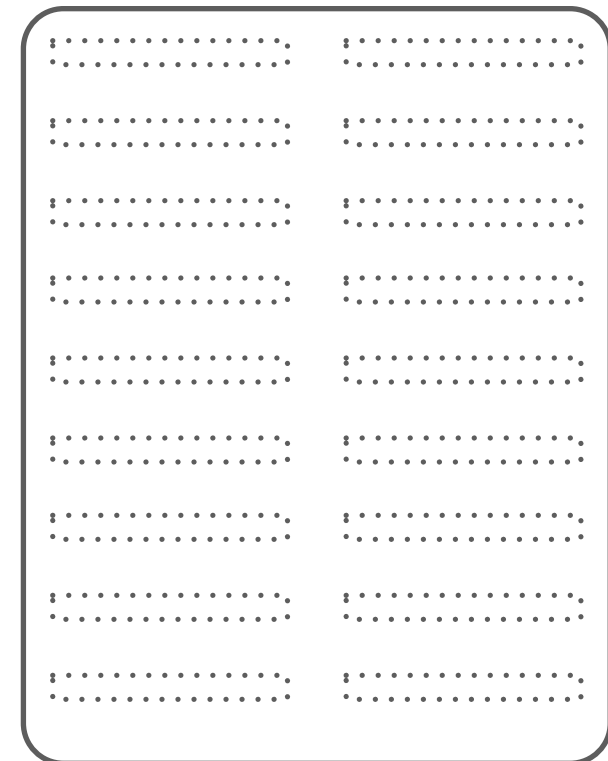
- 計算機の性能向上スピードに陰り
 - →クラウド環境への期待
- 日常利用には十分な計算機性能
 - →仮想計算機技術の発展
- これらの技術を組み合わせるのが最近の流行

仮想計算機による資源拡張

サービスX



ホスト計算機A



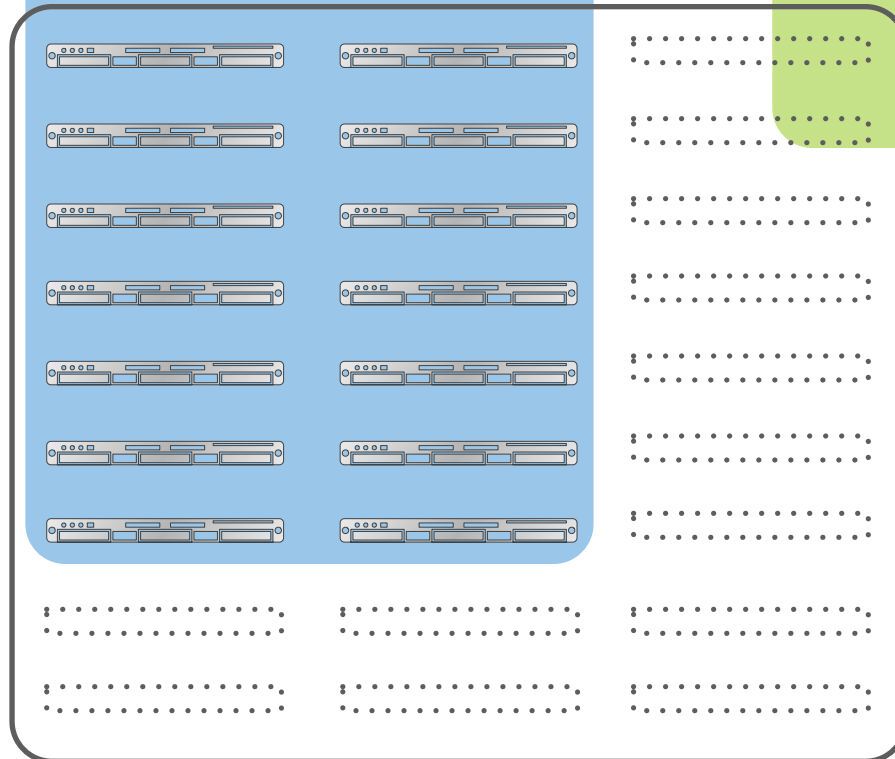
ホスト計算機B

仮想計算機

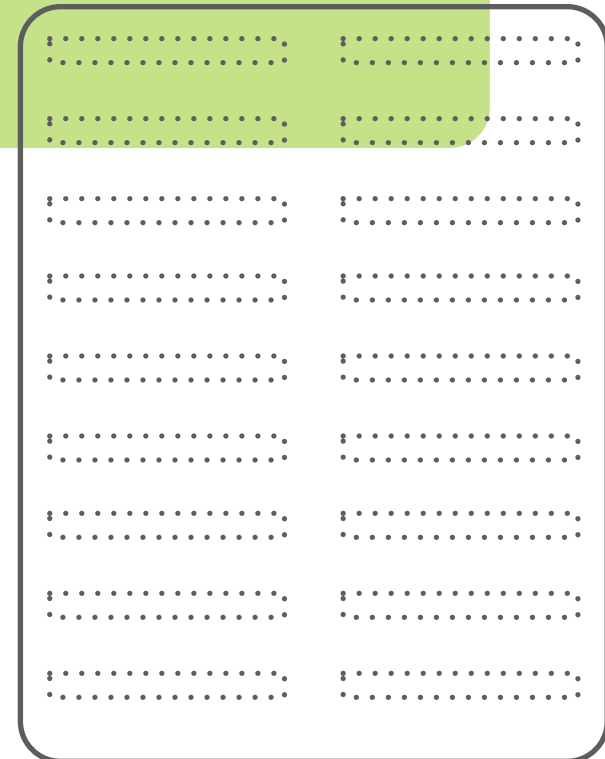
張

サービスY

サービスX



ホスト計算機A



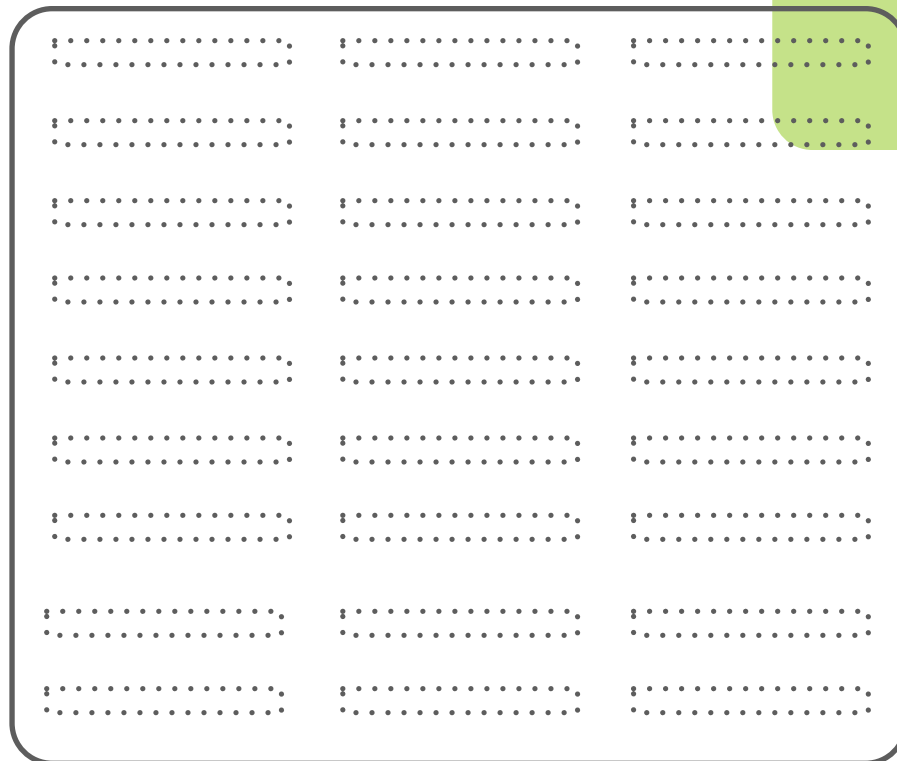
ホスト計算機B

仮想計算機

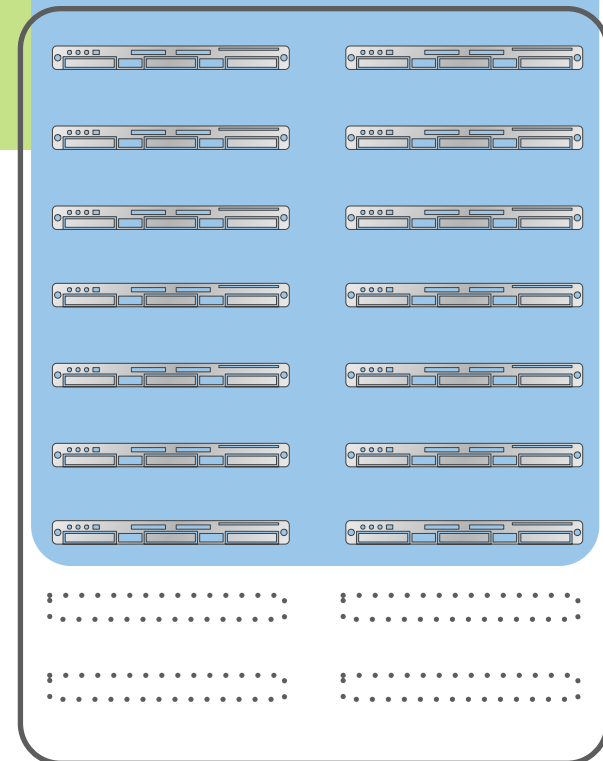
張

サービスY

サービスX



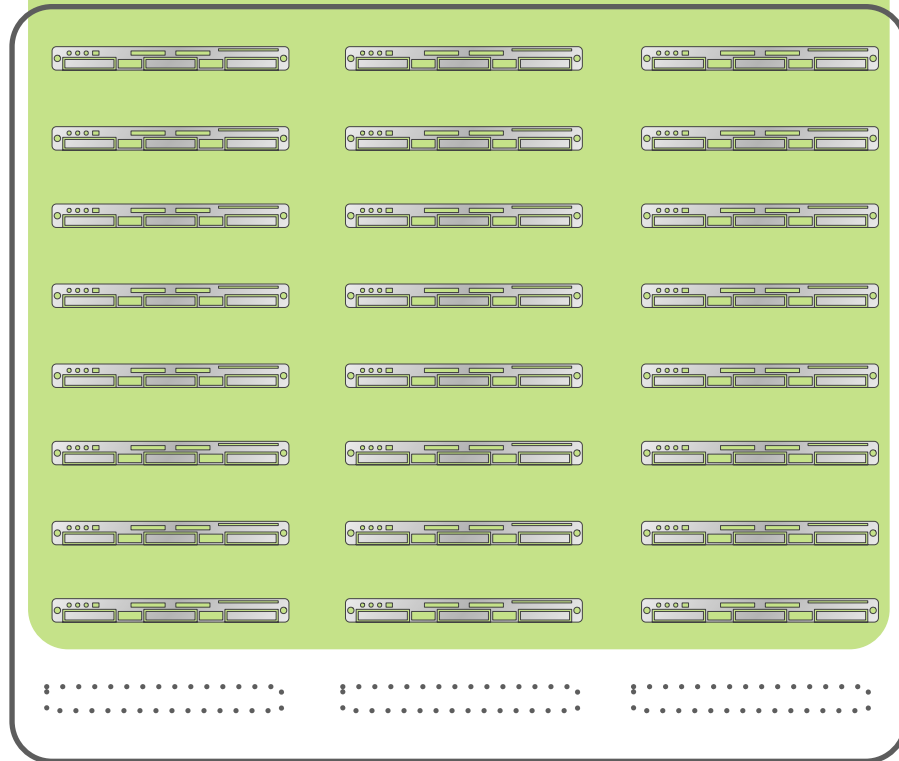
ホスト計算機A



ホスト計算機B

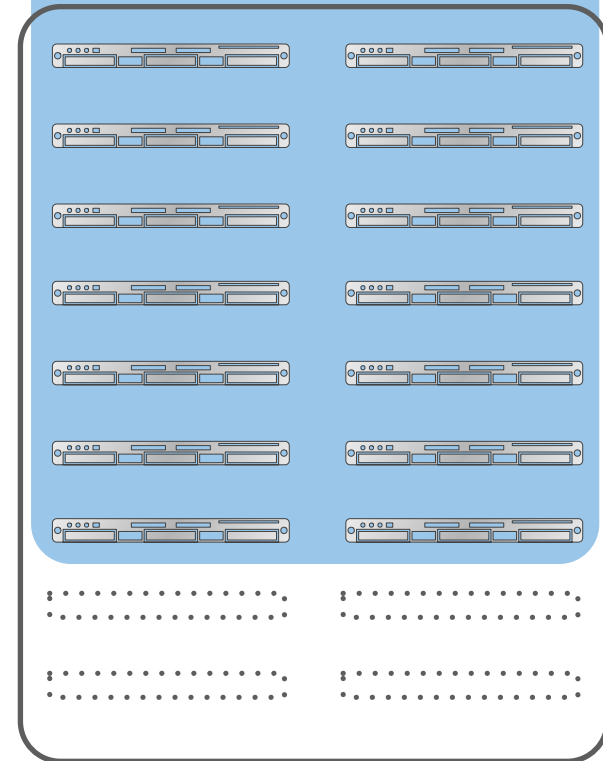
仮想計算機による資源拡張

サービスY



ホスト計算機A

サービスX



ホスト計算機B

仮想環境の効率化

- 仮想環境の効率的な運用のためには
 - 資源の単位を小さく
 - 資源の流動性を確保
- → 流動性の高い仮想計算機管理が重要

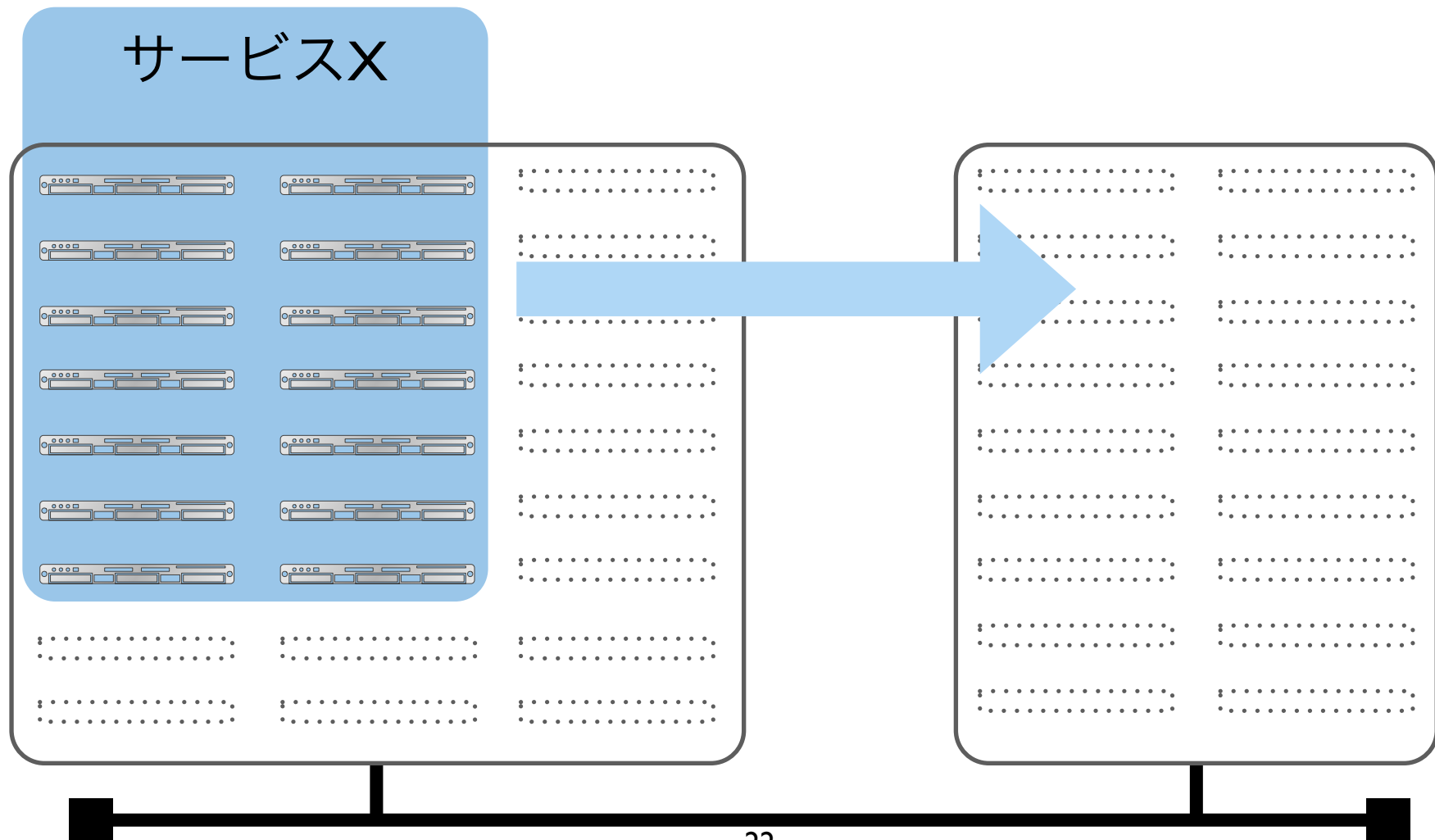
ライブマイグレーション

- 仮想計算機を稼働させたまま、仮想計算機を他のホスト計算機に移動
 - XenのLive Migration
 - VMwareのVMotion

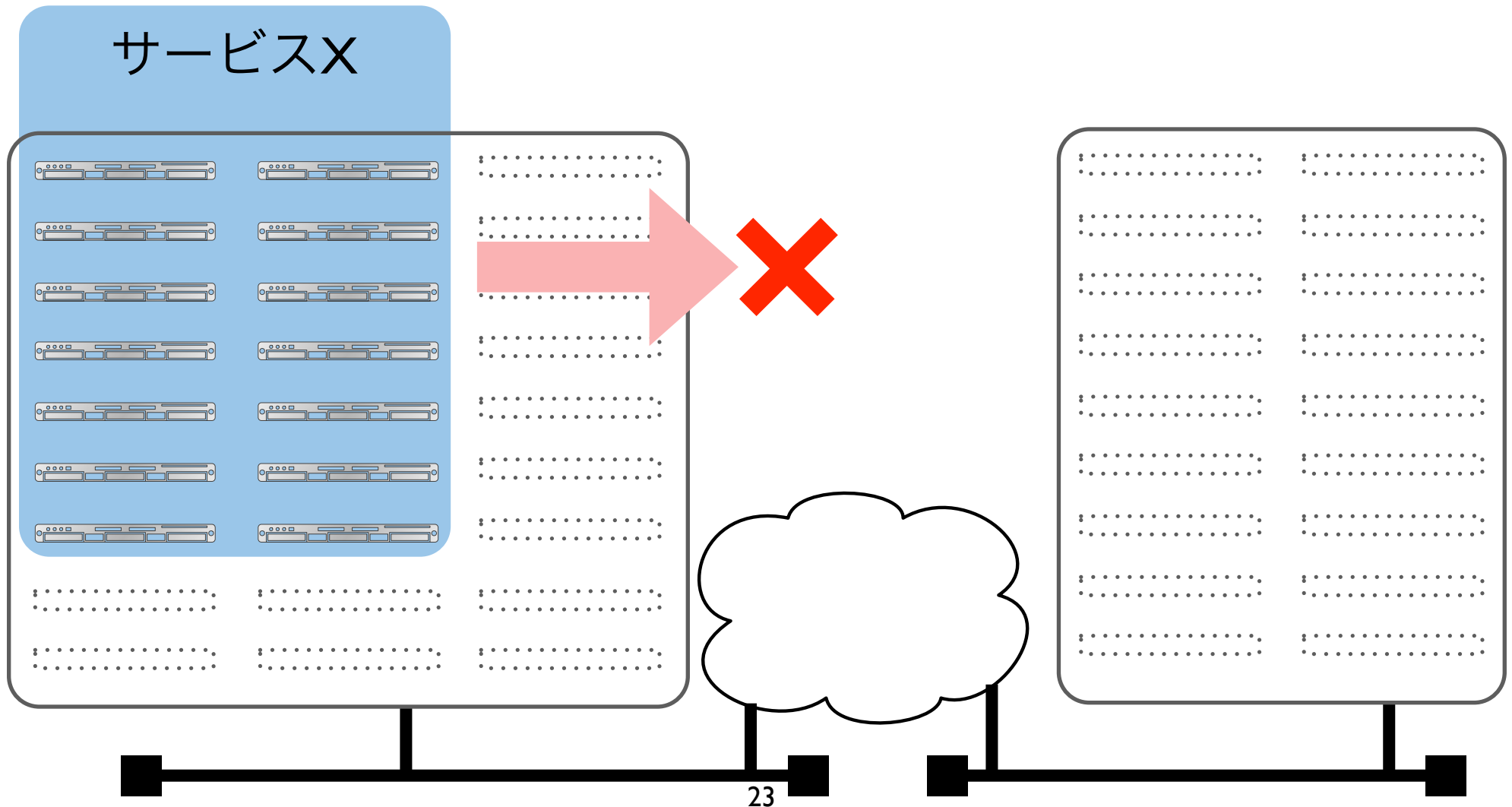
ライブマイグレーションの制限

- 移動元と移動先のホスト計算機が同一セグメントにならない
- ゲスト計算機のネットワーク環境がホスト計算機に大きく依存

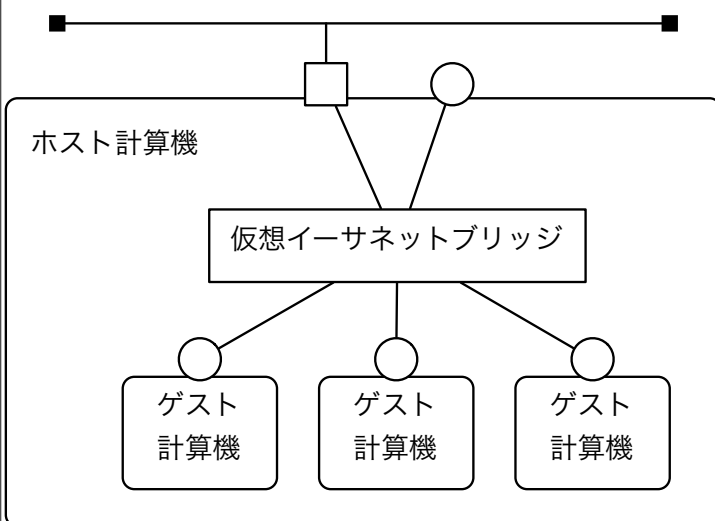
うまくいく例



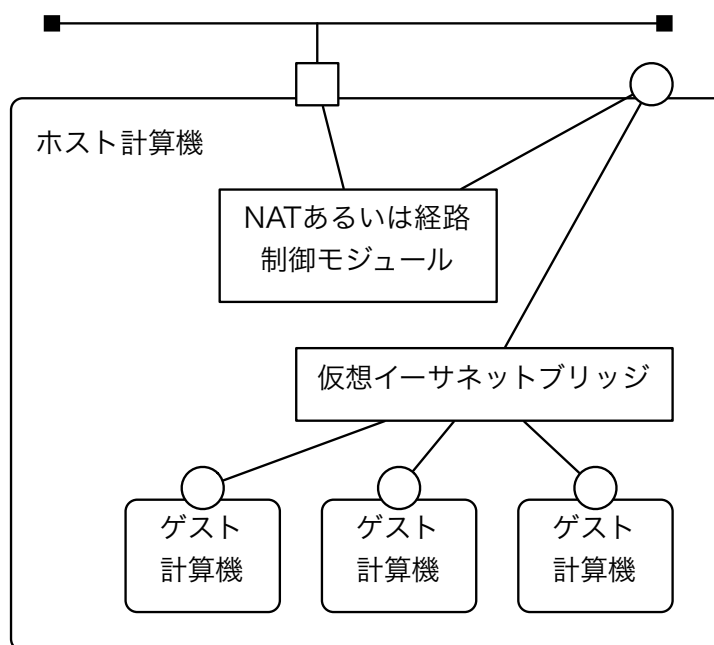
うまくいくかない例



ゲスト計算機の構成



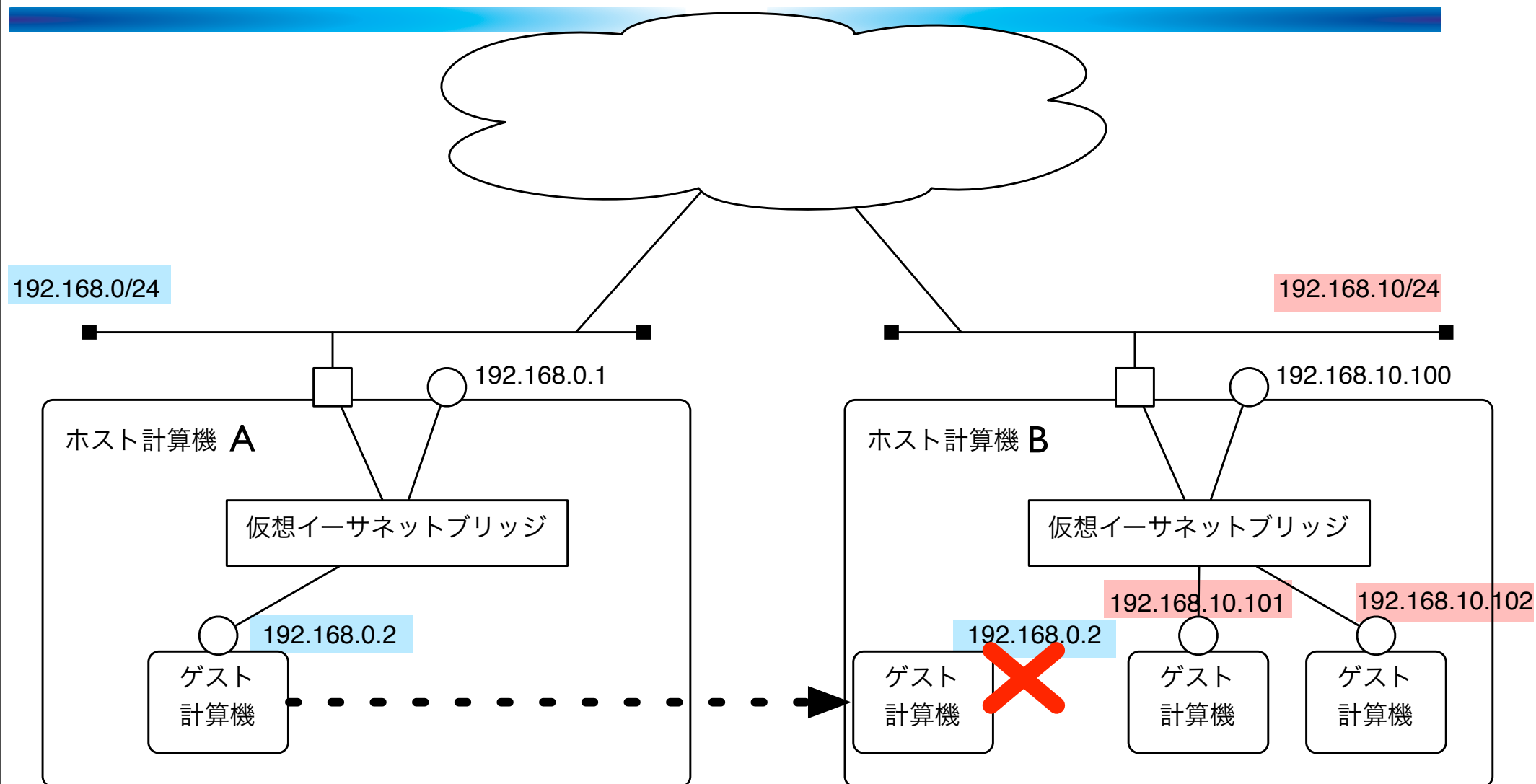
(a) ホスト計算機とゲスト計算機がブリッジ接続される場合



(b) ホスト計算機の裏でゲスト計算機が経路制御あるいはNATされる場合

- ホスト計算機の物理インターフェース
- ホスト計算機およびゲスト計算機の仮想インターフェース

問題例



解決案

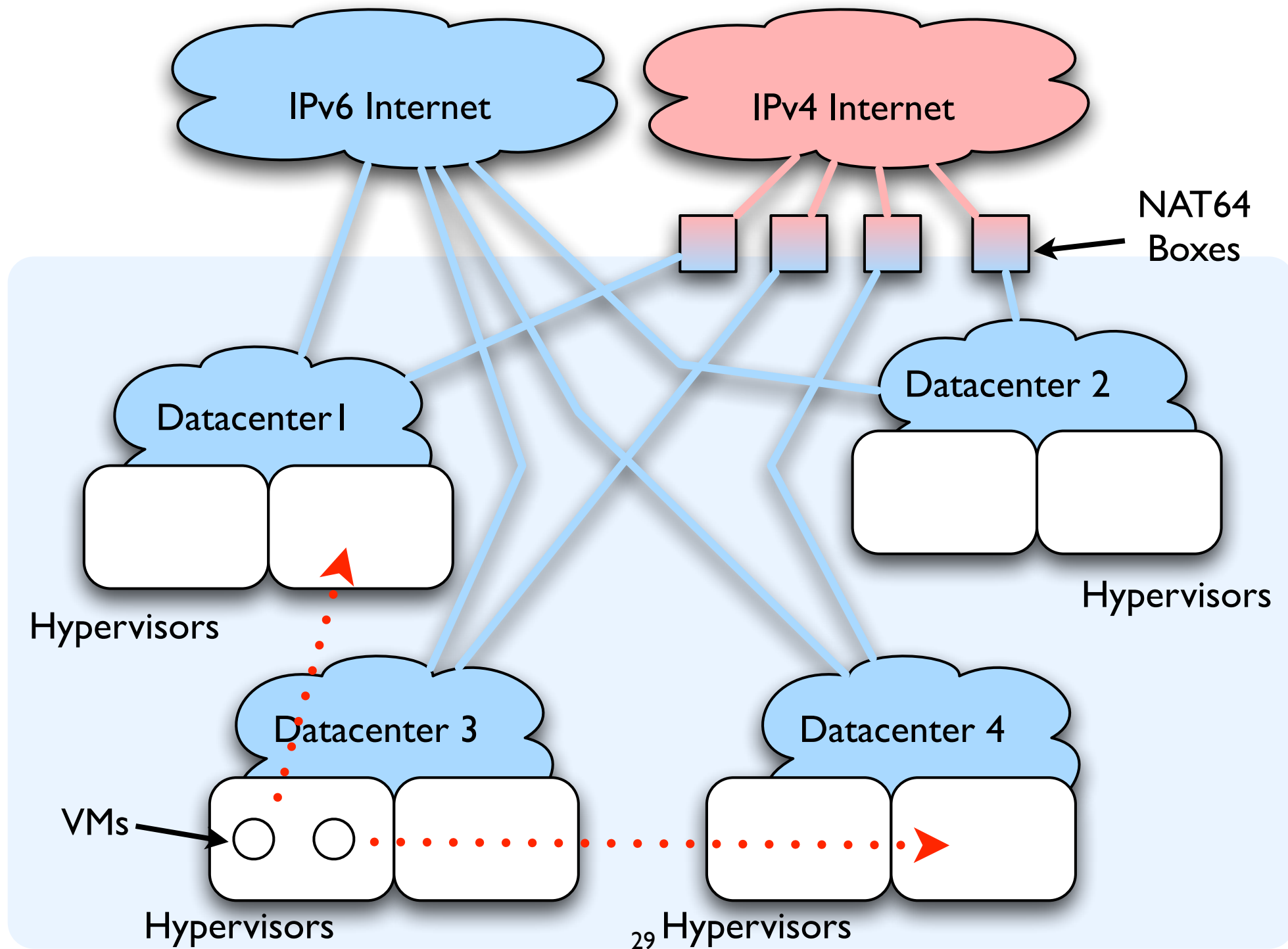
- 解決方法はひとつではない (一長一短)
 - 広域VLAN
 - 経路制御
 - IPモビリティ
 - (アプリケーション層)

利点欠点

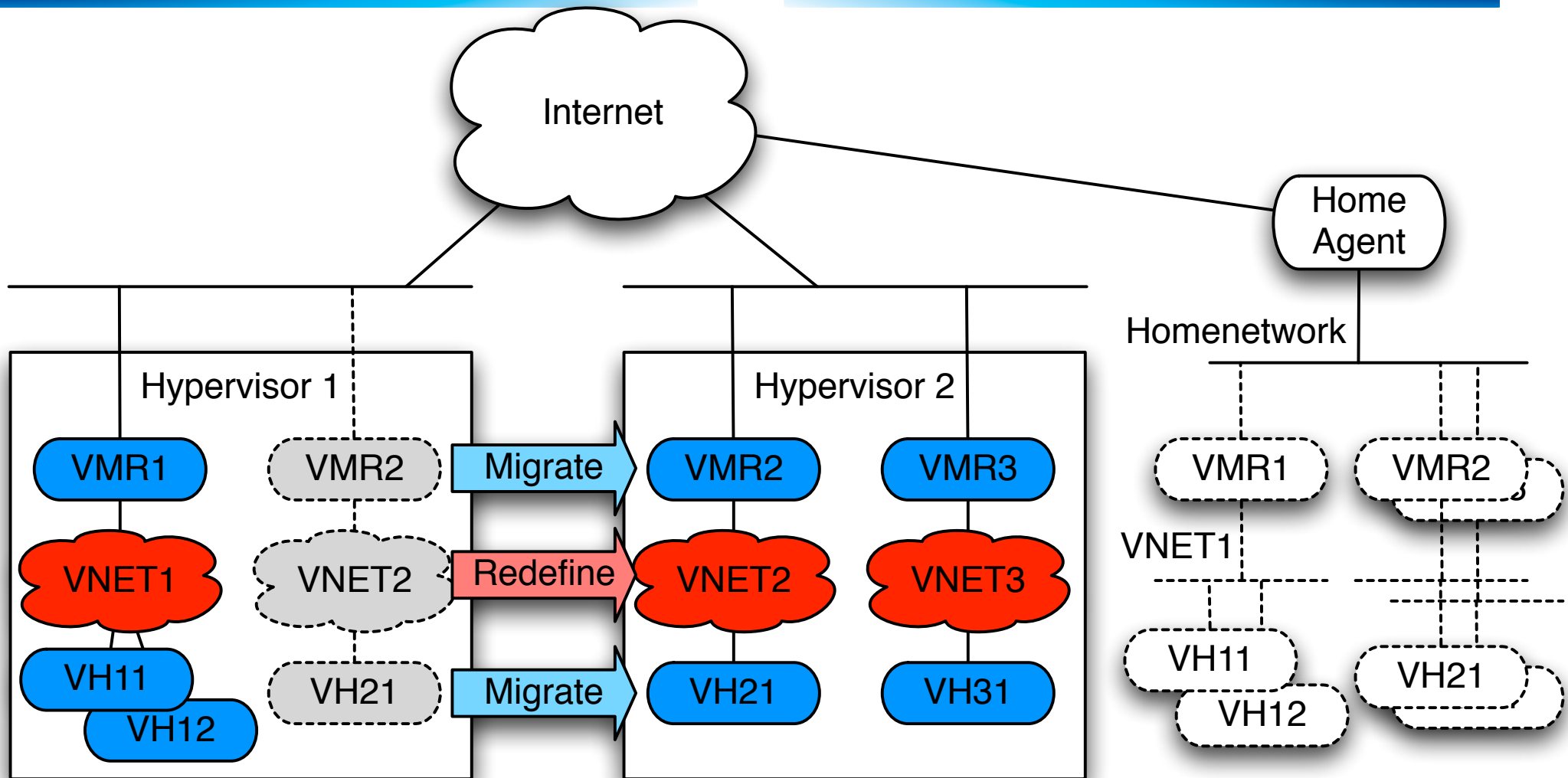
	運用性	ドメイン越え	SLA
広域VLAN	VLAN運用の知識を活用可能	VLAN IDの共有が必要 なため、一般的に不可 能	L2レベルの制御 で確保
経路制御	経路制御の知識 が必要	原理的には可能だが 運用が困難	ネットワーク層 QoSが必要
IPモビリティ	経路制御、IPモビリ ティの知識が必要	ネットワーク層から 独立なため可能	複数組織が関係する ため困難

今試している技術

- IPv6ネットワーク移動通信プロトコル (NEMO)を用いたネットワーク可搬性
- IPv6のみで構成されたクラウドサービスバックエンド
- IPv4/IPv6変換を用いた後方互換性



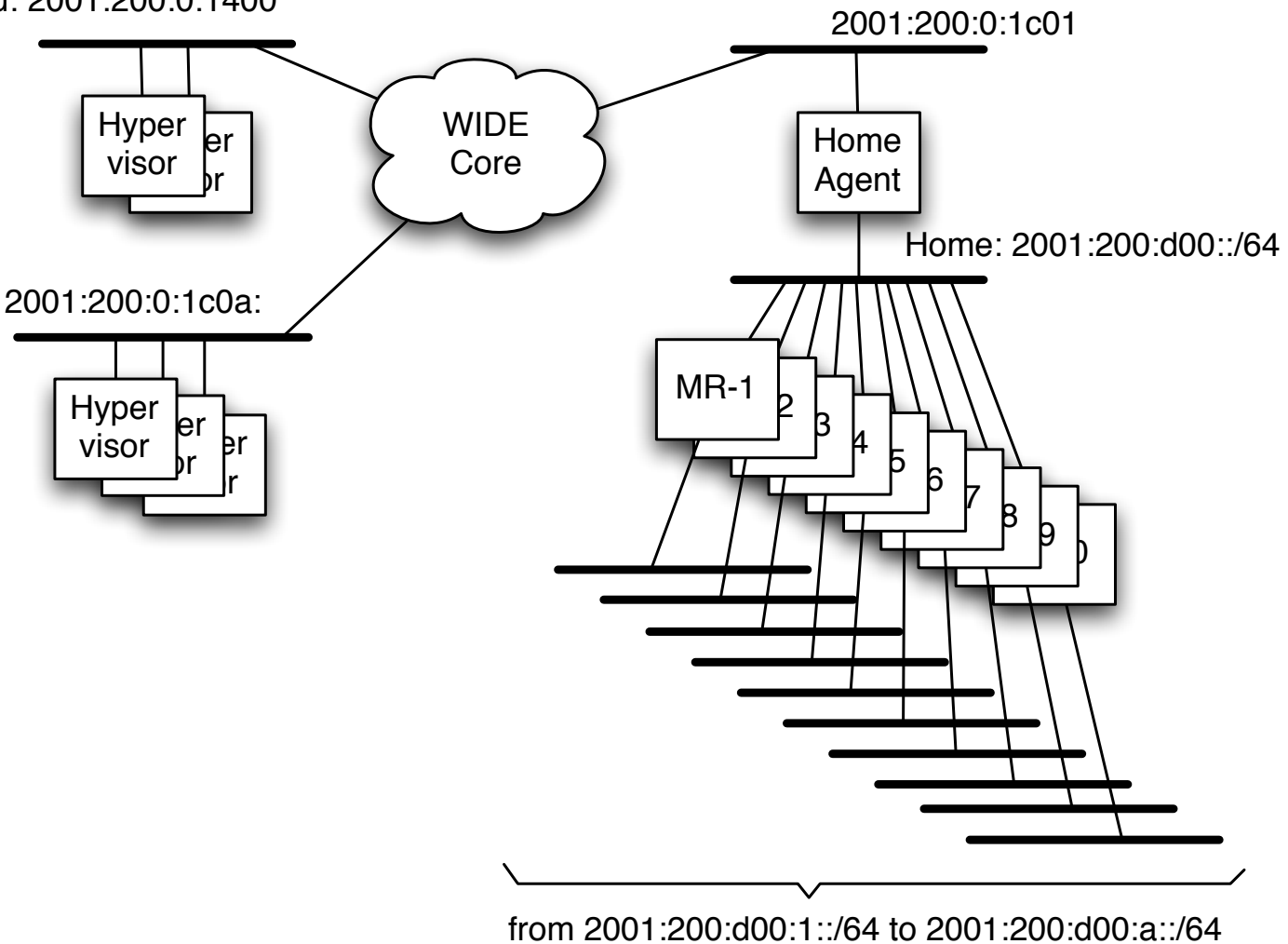
NEMOの応用



テストベッド

WIDE komatsu: 2001:200:0:1400

WIDE nezu: 2001:200:0:1c0a:



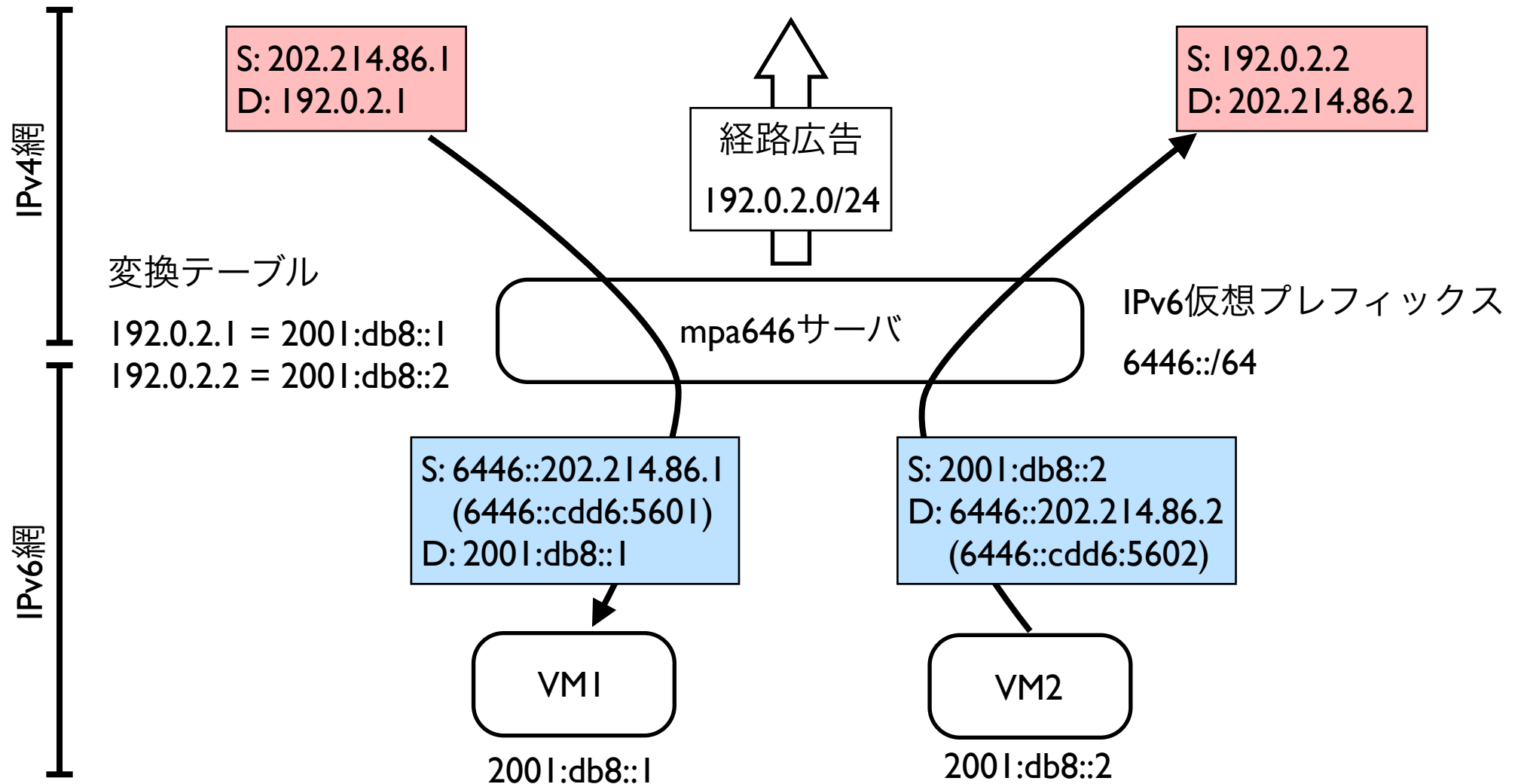
IPv6主運用へのステップ

- クラウドサービスネットワーク内でたくさんのVM運用
- 必ずしもすべてのVMが表に立つ訳ではない
- 網の運用コストはなるべく減らしたい (可能であればシングルスタック)
- 既存IPv4ユーザへの接続性は確保したい
- 負荷、IPv6普及度に応じて増減できるIPv4互換機能

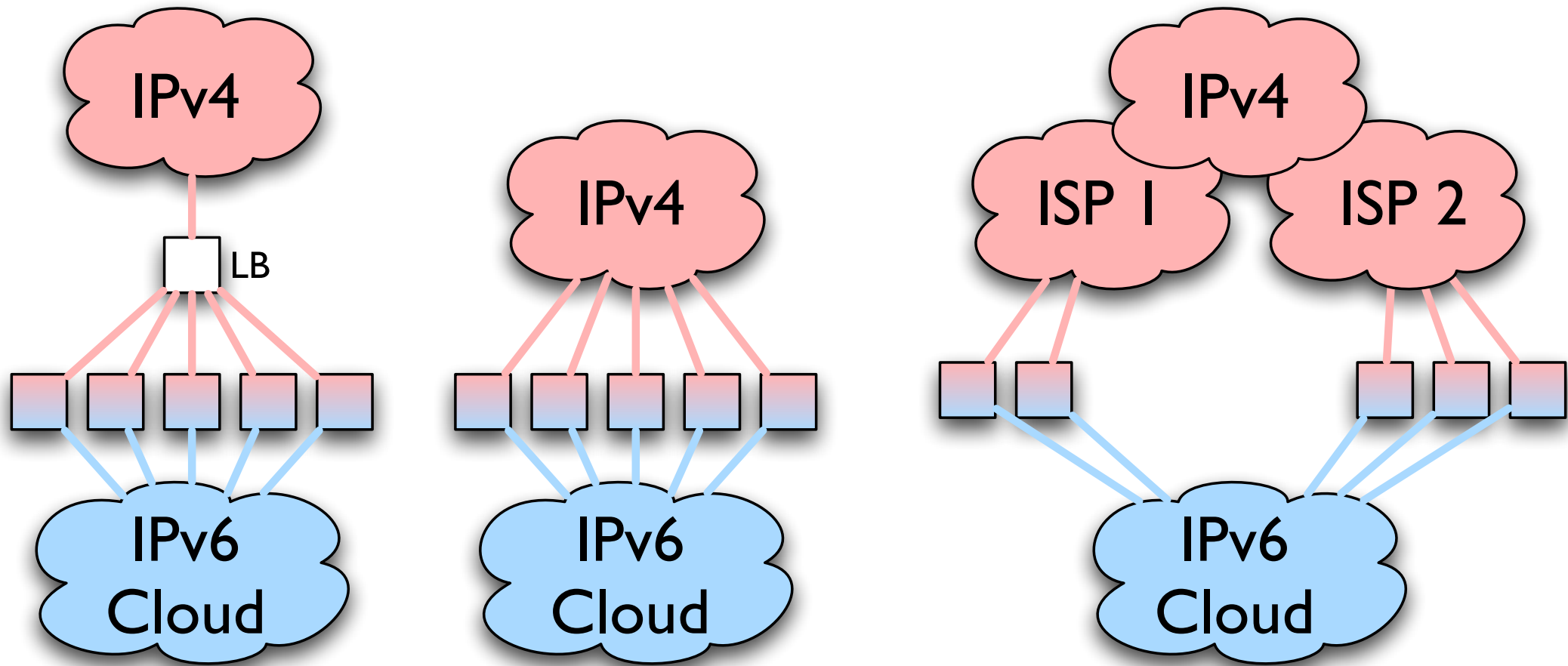
map646: シンプルNAT64 サービスソフトウェア

- ひとつのIPv6アドレスをひとつのIPv4グローバルアドレスに対応づけ、ヘッダ変換するソフトウェア
- 高可用性を第一の目標として設計
 - ステートレスな64変換 (および逆変換)
 - 任意の数のサーバを配置可能

変換ルール



自由なサーバー配置

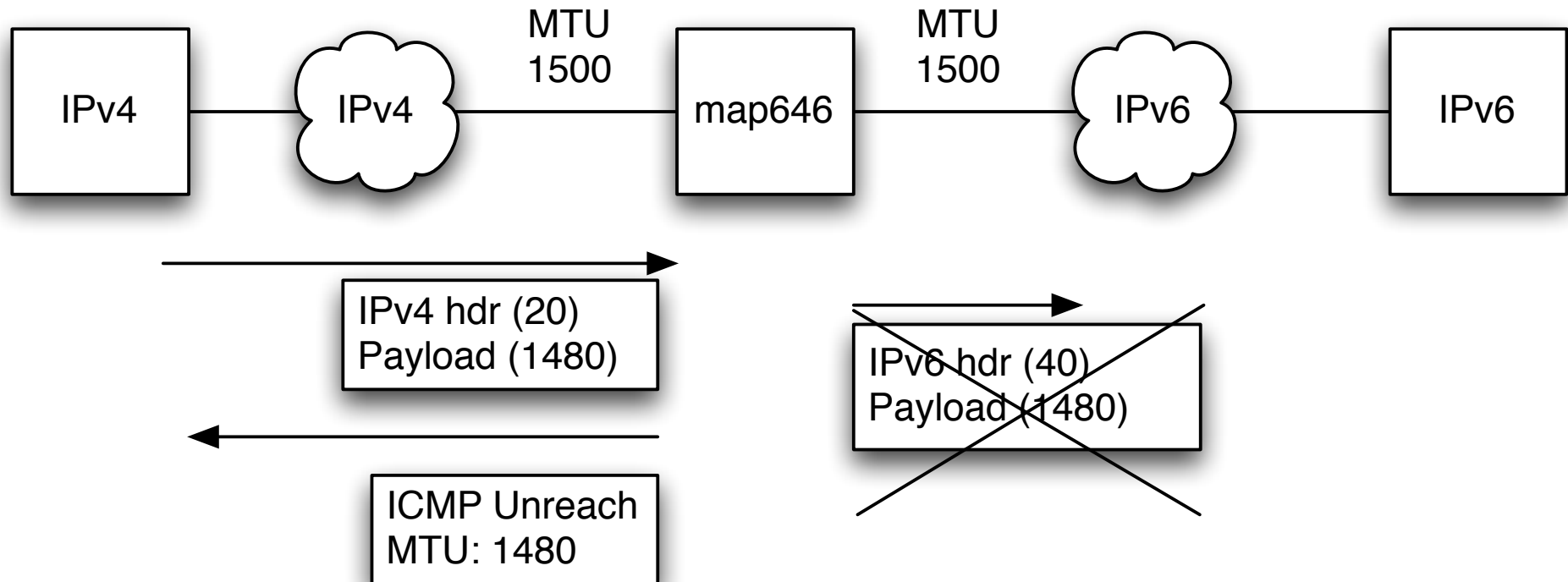


変換可能なプロトコル

- TCPとUDPを完全サポート
 - ただし、アプリケーションデータとしてIPアドレスを利用するものを除く (たとえばFTP)
- ICMP/ICMPv6はPath MTU発見に必要な機能のみサポート

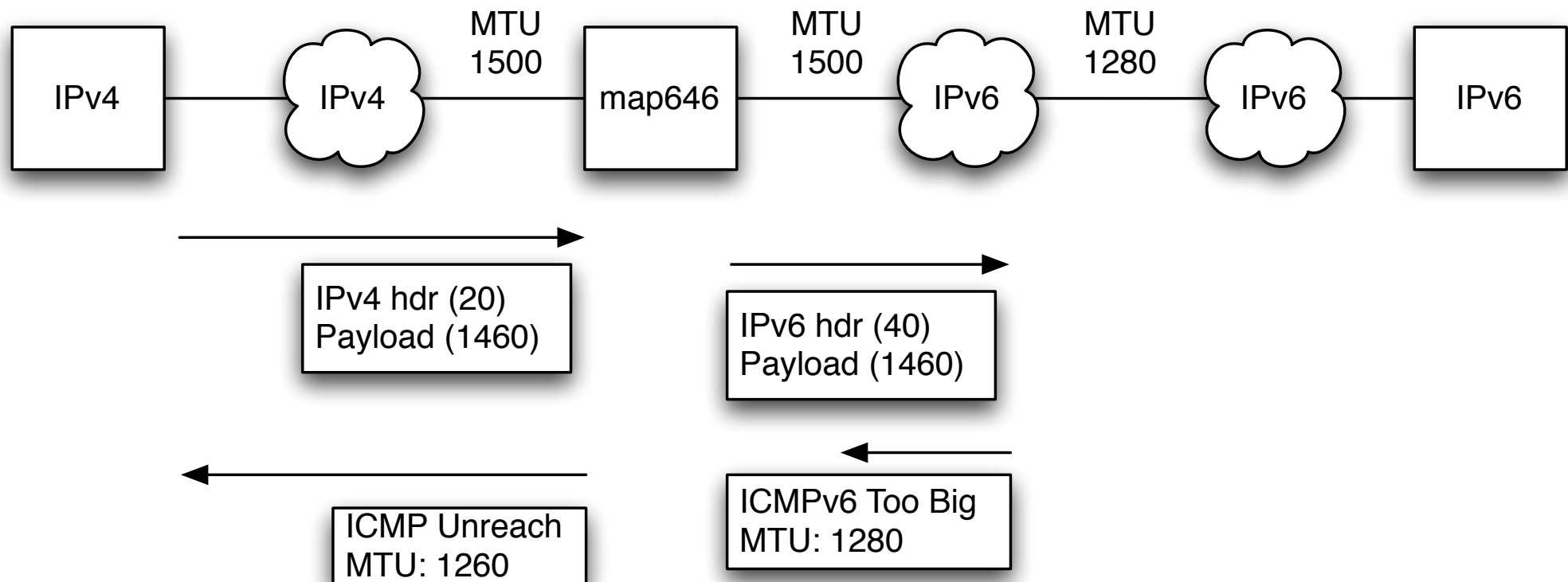
Path MTU Discovery

- ICMP Unreachable generation
 - When converting IPv4 to IPv6



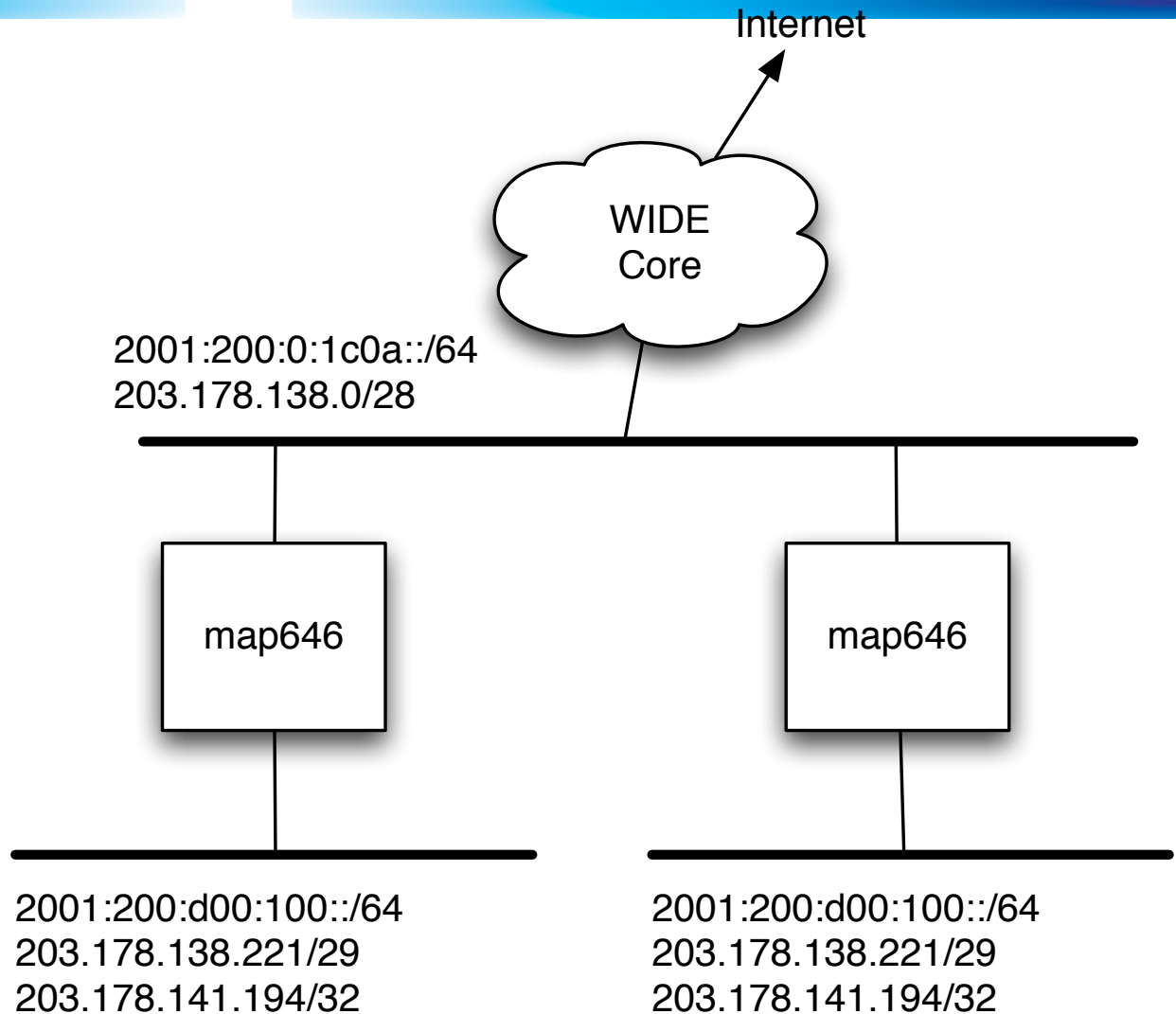
Path MTU Discovery

- When a bottleneck exists in dest network



テストベッド

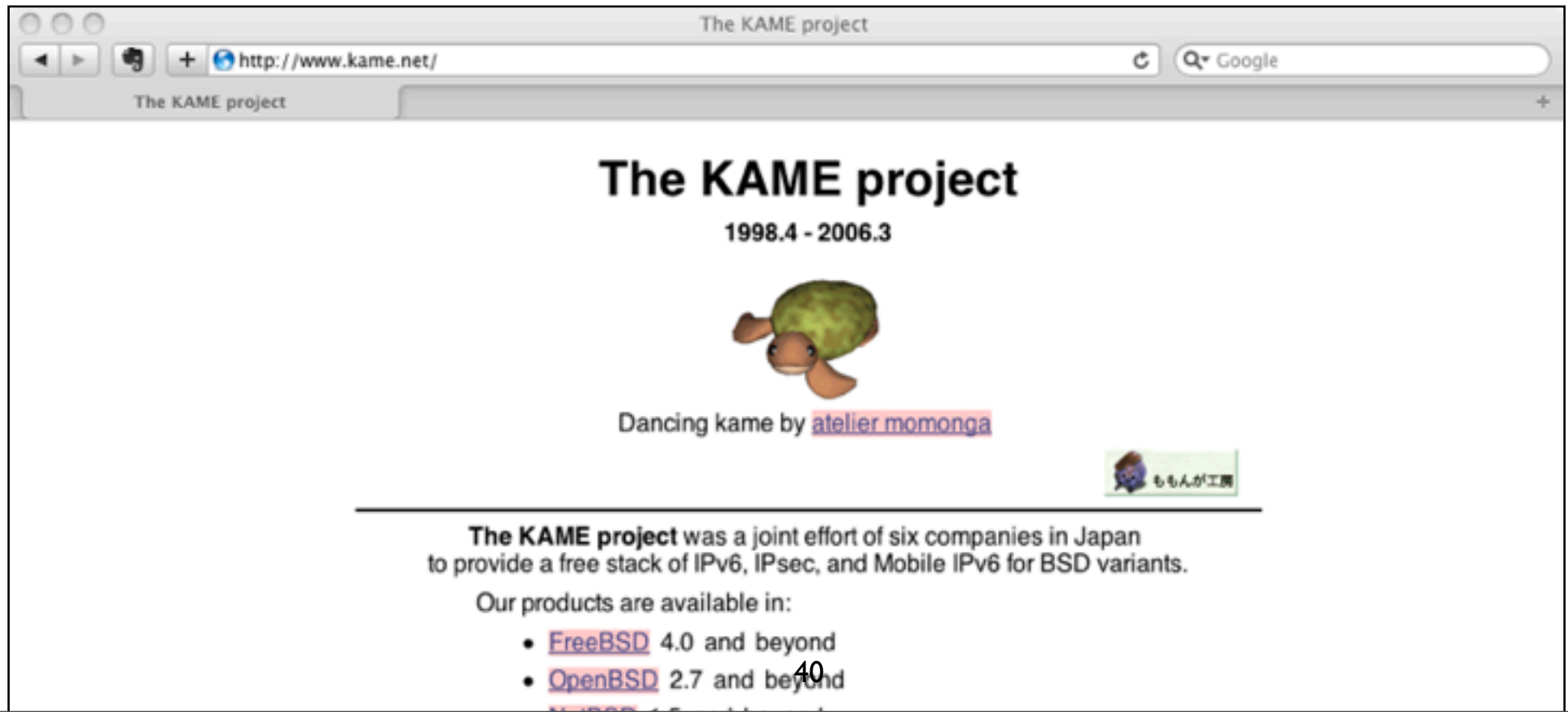
WIDEクラウド
環境にて運用中



- Virtual IPv6 addresses for IPv4 nodes
- Virtual IPv4 addresses for IPv6 servers

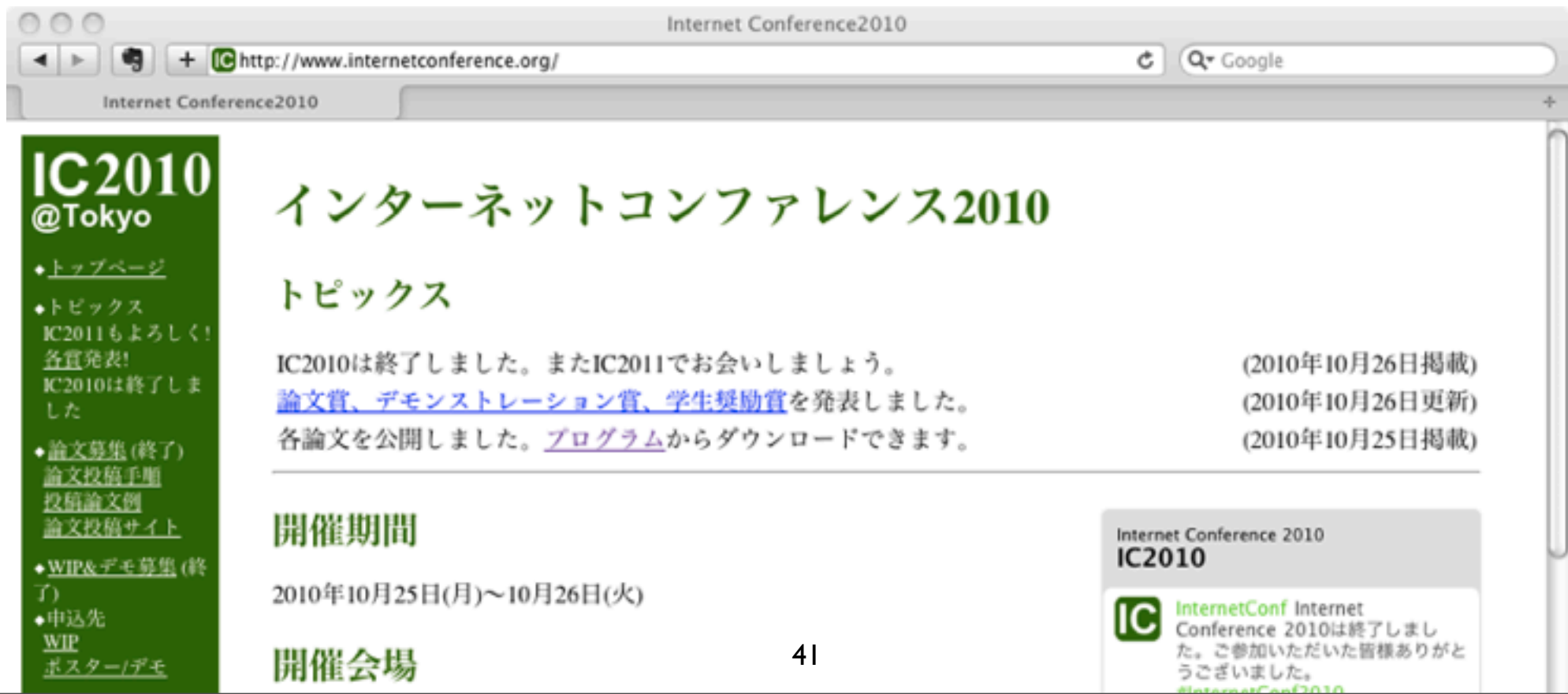
実運用サイト 1/3

- KAME historical page
- <http://www.kame.net/>



実運用サイト 2/3

- Internet Conference 2010
- <http://www.internetconference.org/>



The screenshot shows a web browser window titled "Internet Conference2010" with the address bar displaying "http://www.internetconference.org/". The website content is in Japanese and features a green sidebar on the left with navigation links. The main content area has a large green header "インターネットコンファレンス2010" and a section titled "トピックス" (Topics) with several announcements. A date range "2010年10月25日(月)～10月26日(火)" is displayed. At the bottom right, there is a small box with the "Internet Conference 2010 IC2010" logo and a message.

IC2010 @Tokyo

- [トップページ](#)
- [トピックス](#)
IC2011もよろしく!
各宜発表!
IC2010は終了しました
- [論文募集 \(終了\)](#)
[論文投稿手順](#)
[投稿論文例](#)
[論文投稿サイト](#)
- [WIP&デモ募集 \(終了\)](#)
- [申込先](#)
[WIP](#)
[ポスター/デモ](#)

インターネットコンファレンス2010

トピックス

IC2010は終了しました。またIC2011でお会いしましょう。
[論文賞、デモンストレーション賞、学生奨励賞](#)を発表しました。
 各論文を公開しました。[プログラム](#)からダウンロードできます。

(2010年10月26日掲載)
 (2010年10月26日更新)
 (2010年10月25日掲載)

開催期間

2010年10月25日(月)～10月26日(火)

開催会場

Internet Conference 2010
IC2010

IC InternetConf Internet Conference 2010は終了しました。ご参加いただいた皆様ありがとうございました。
[InternetConf2010](#)

実運用サイト 3/3

- 先日開催された厚労省製か発表シンポジウムのウェブサイト



解決すべき問題

- 広域ネットワークにおけるストレージ資源の移動技術
- 広域、あるいはドメインを跨いだ環境でのネットワークSLAの確保技術
- ポリシーマネージメント

今日のまとめ

- データセンター運用における、さらなるコストダウン、高可用性の要求
- 計算機性能向上による仮想計算機の台頭
- 実用レベルの計算資源移動技術の出現
- ネットワークレベルの仮想化の必要性和そのプロトタイプ提案
- 実運用ネットワークによる動作検証