



VMwareによる仮想化とネットワーク

VMware株式会社

藤井厚

小松康二

アジェンダ

- VMwareによるサーバ仮想化
- 仮想ネットワークの考え方
- サーバ仮想化でネットワークはどう変わるか
 - 責任分解点と設定の整合性
 - データセンタのネットワークデザインに与える影響
 - I/O統合への期待
- 仮想化されたデータセンターの今後とVMware
 - 仮想化を基盤としたデータセンターの市場動向
 - 公開可能な範囲内でのロードマップの紹介 (スライドなし)

サーバ仮想化とVMware

企業概要



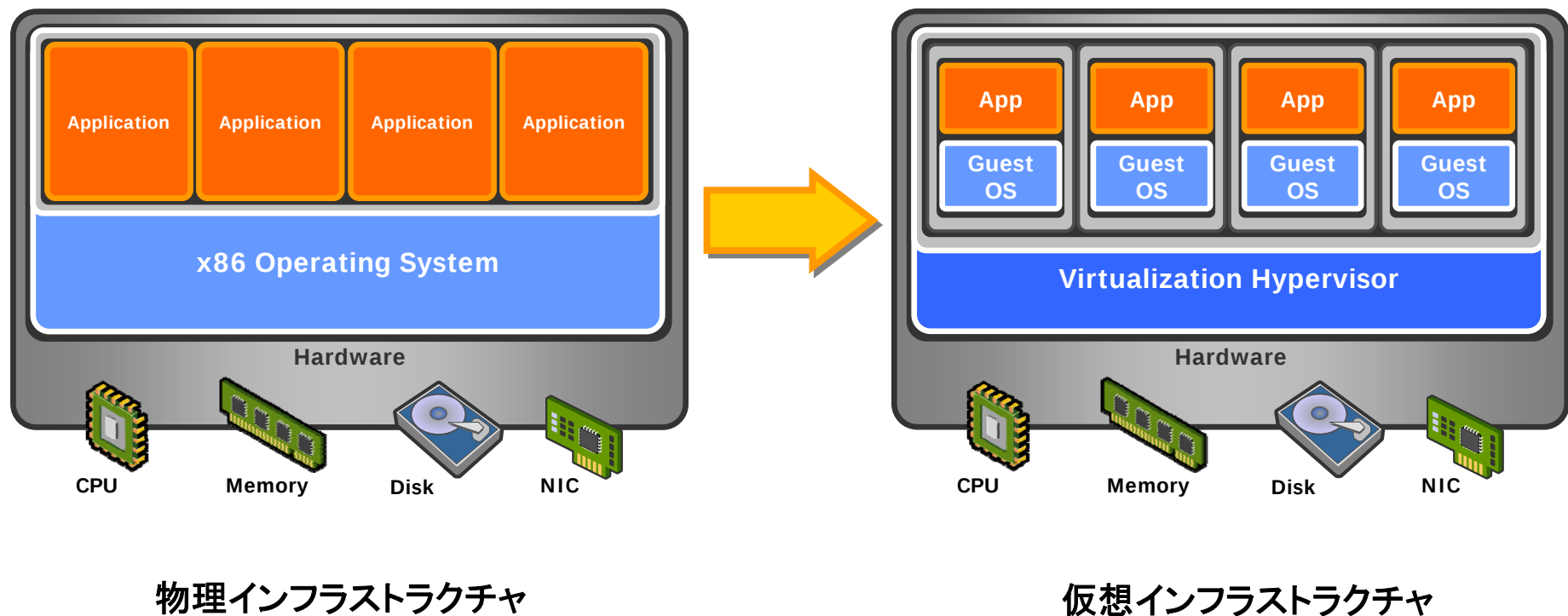
設立	1998
社員数	6,000+
博士号を持つエンジニアの比率	40%

テクノロジー・パートナー企業	350+
チャネル・パートナー企業	6000+
VMware認定プロフェッショナル	15,000+

フォーチュン100 / フォーチュン1000企業	100 / 910
VMware Infrastructureを利用している企業	20,000+
VMware Workstation販売ライセンス実績	1,000,000+

VMwareのコア技術 = x86マシンの仮想化

x86マシンの仮想化はVMwareの基幹技術



VMware仮想マシンの3つの特徴

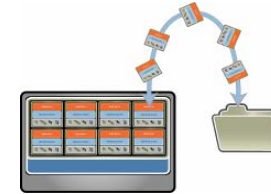
独立性



同一基盤上の仮想マシンは
相互に影響しない

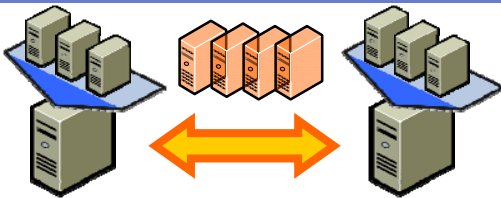
VMware 仮想マシンの 特徴

カプセル化



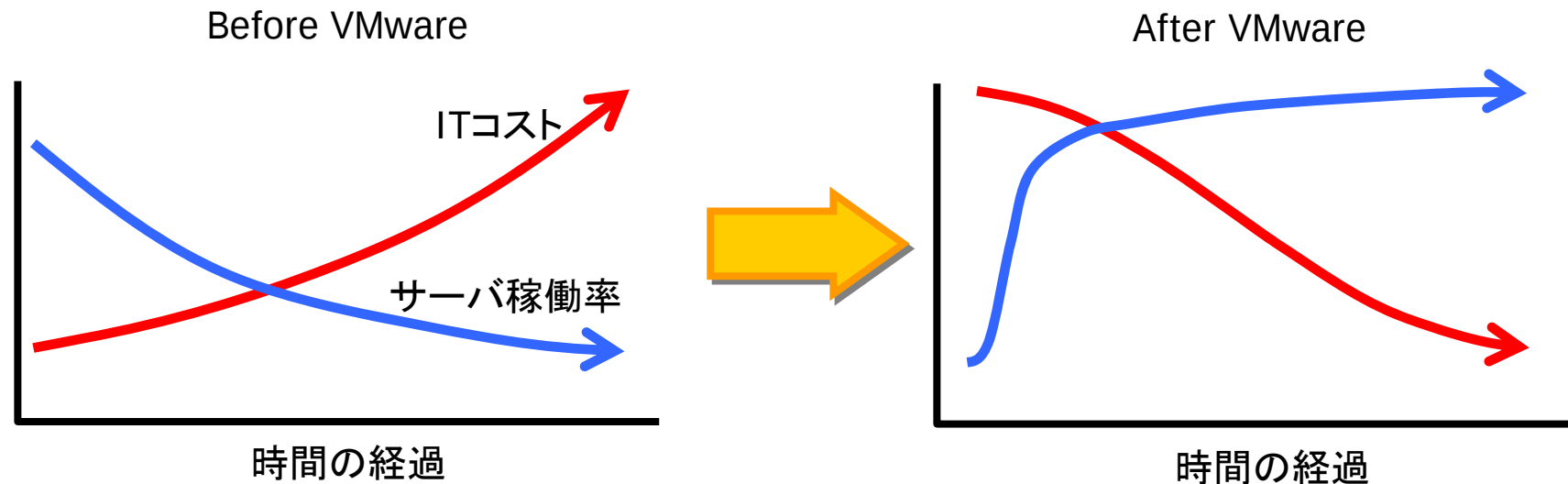
仮想マシンはファイルとして格納される

ハードウェアからの独立



仮想マシンは別のH/Wに移動させても
そのまま動作することが可能

VMwareの効果



◆上がっていくもの

- 投資コスト(CAPEX)
- 運用コスト(OPEX)
- システムに対するデマンド/プレッシャー

◆下がっていくもの

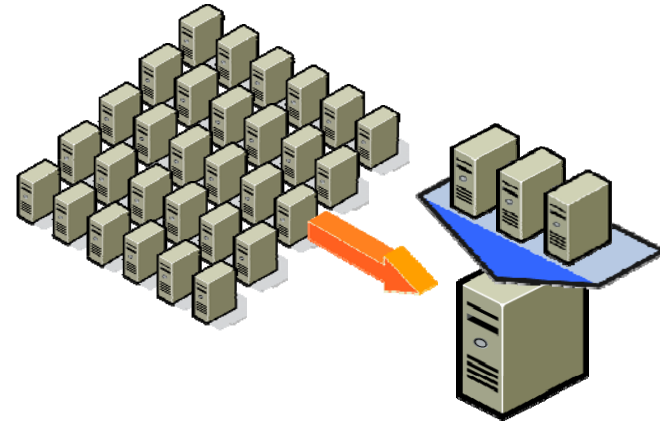
- サーバ稼働率(特にCPU)

**VMwareの仮想化技術は、
このバランスを逆転させる**

サーバ統合及びプロビジョニングの迅速化

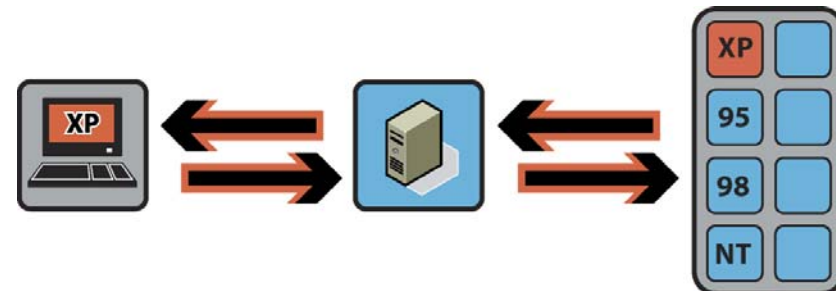
■ サーバ統合での活用

- ◆ H/Wリソースの有効活用
- ◆ スペース、消費電力等のコスト削減
- ◆ レガシーシステムの延命

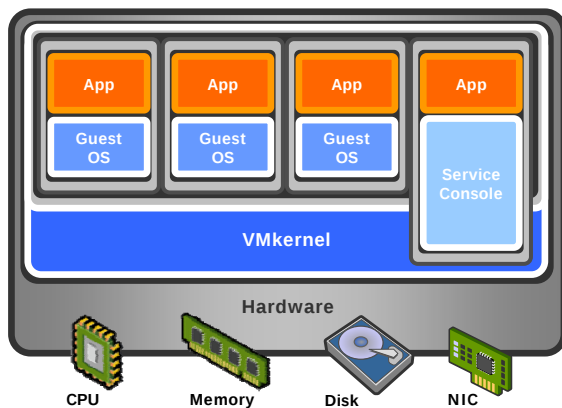


■ サーバ準備期間の短縮化

- ◆ 従来の設置プロセスを大幅に短縮
- ◆ 新規サーバの準備が10数分で完了

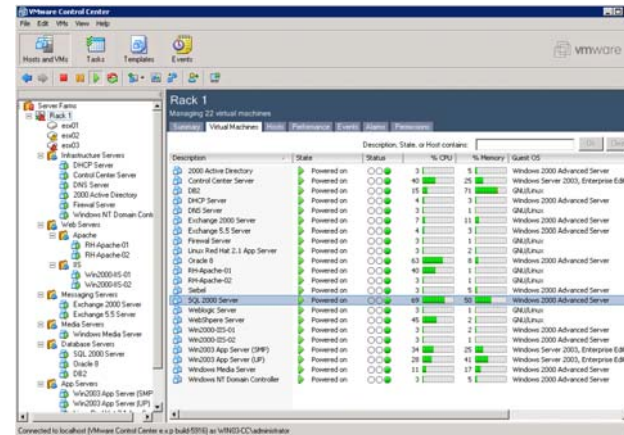


コア・コンポーネント : VMware Infrastructure 3



VMware ESX / ESXi

- 業界最高水準の仮想化ハイパーバイザ
 - リソース管理
 - ストレージ/ネットワーク・スタック
- サービスコンソール (ESX Only)
 - RHELベースの仮想マシン
 - 管理コマンド等を実装
- デバイス共有
 - CPU、メモリ、ネットワーク、ディスク

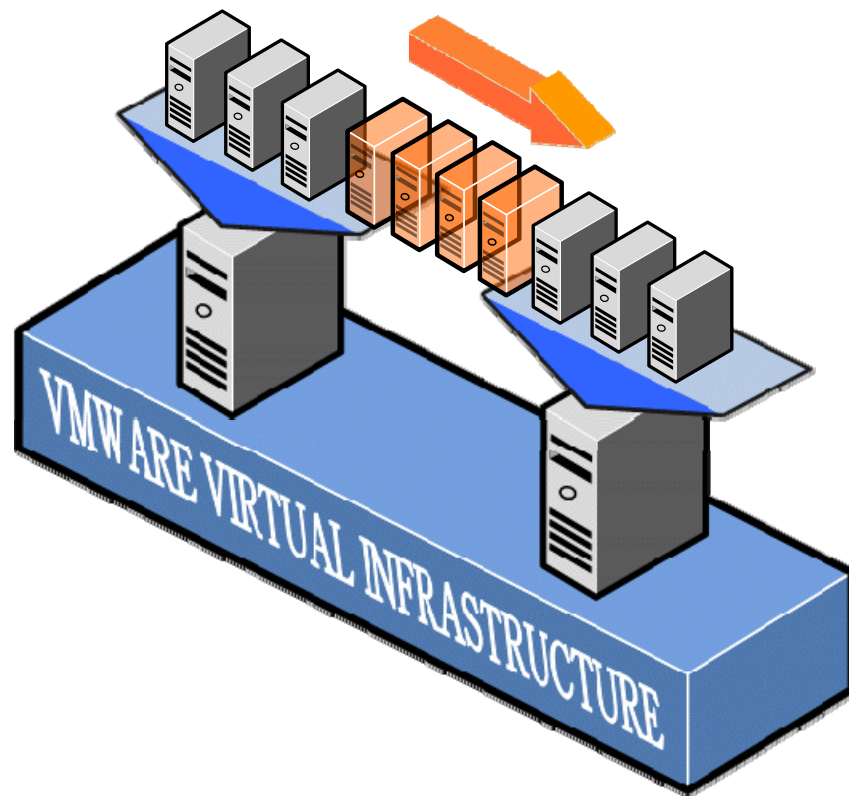


VirtualCenter

- 仮想環境を効果的に管理
 - 200台のESX、2000台のVMを管理可能
- ワークロード分散機能を実装
 - VMotion, DRS
- 仮想マシンの可用性を向上
 - VMware HA
- 迅速なプロビジョニング機能を実装
 - テンプレート、クローニング

VMotion

VMwareをメインストリームに進出させたキーテクノロジー



> ゼロダウンタイムの実現

- 無停止H/Wアップグレード
- 無停止でのサーバ用途変更

> 様々な要素からの独立

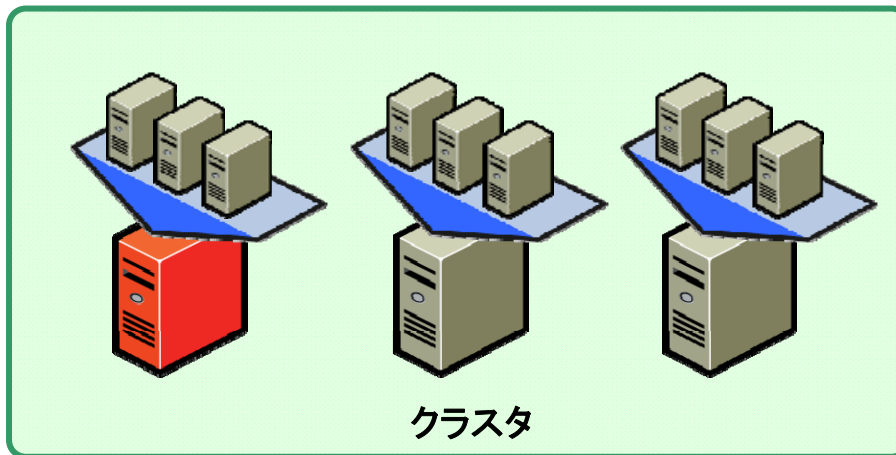
- OS/アプリケーション依存無し
- サーバベンダの差異依存無し
※CPUに関する制限アリ

> 操作方式

- スケジューラによる自動化
- 手動操作によるリアルタイム実行

Distributed Resource Scheduling

- ダイナミックなリソース最適化機能



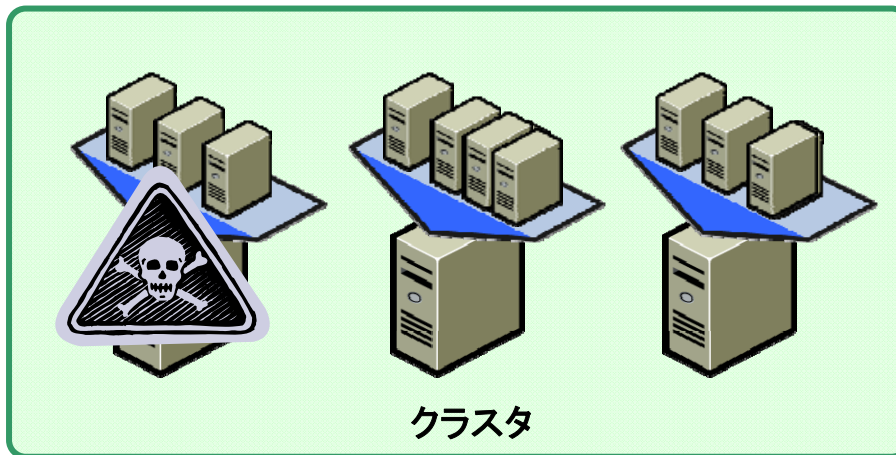
> ダイナミックなリソース管理

- 自動的に仮想マシンをVMotion
- 定義されたルールに基づいてインテリジェントなリソース配分を実施
- 継続性のある最適化

> DRSによるベネフィット

- より柔軟性のあるH/Wメンテナンス
- 簡単な操作により、システム管理者の生産性を大幅に向上

VMware HA – 仮想環境に高可用性を提供



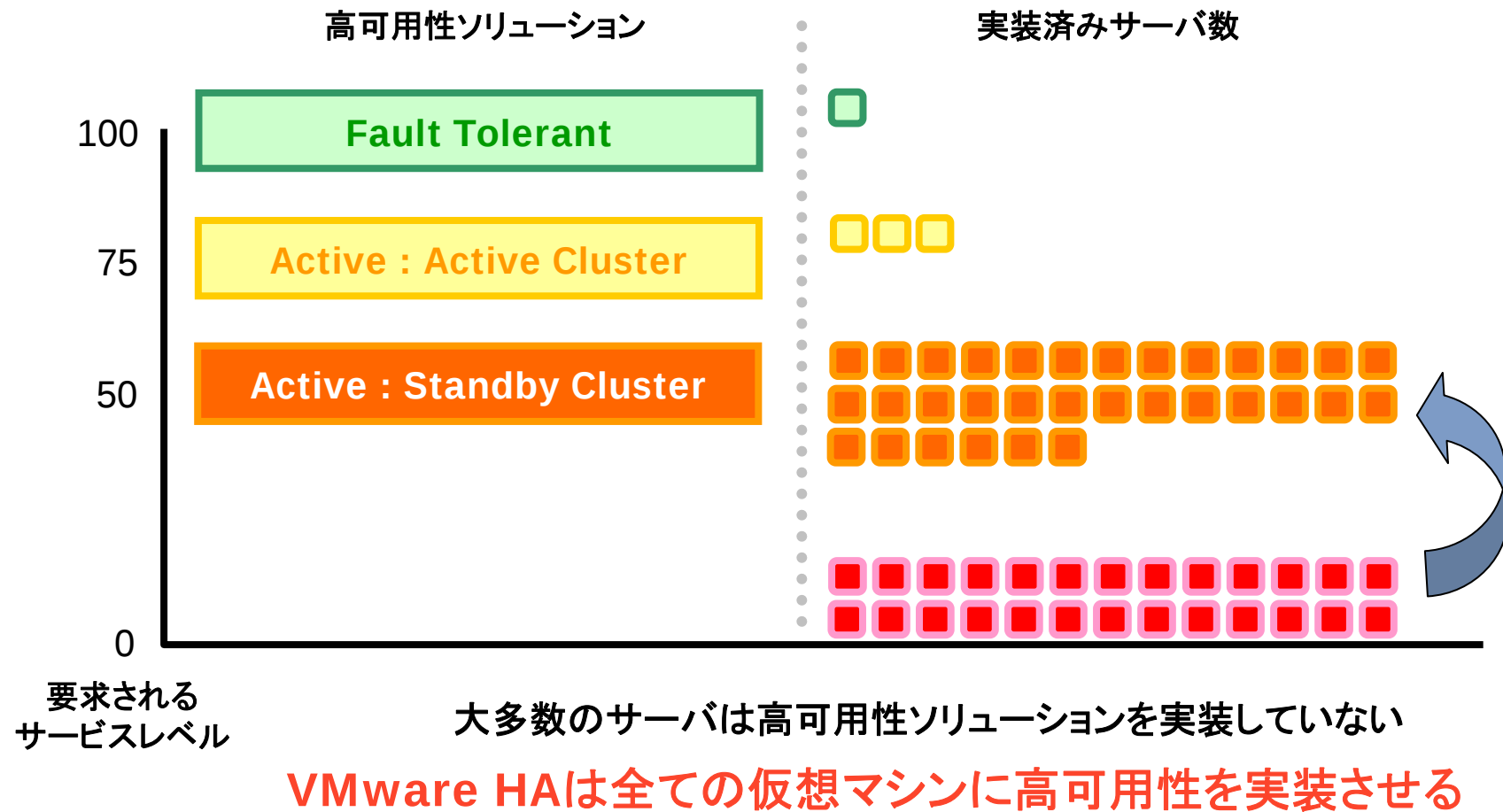
> 仮想マシン・フェイルオーバー

- ESX障害時に仮想マシンを他のESXに自動的にフェイルオーバー
- 仮想マシン毎に優先度を設定可能
- 継続性のある保護

> VMware HAによるベネフィット

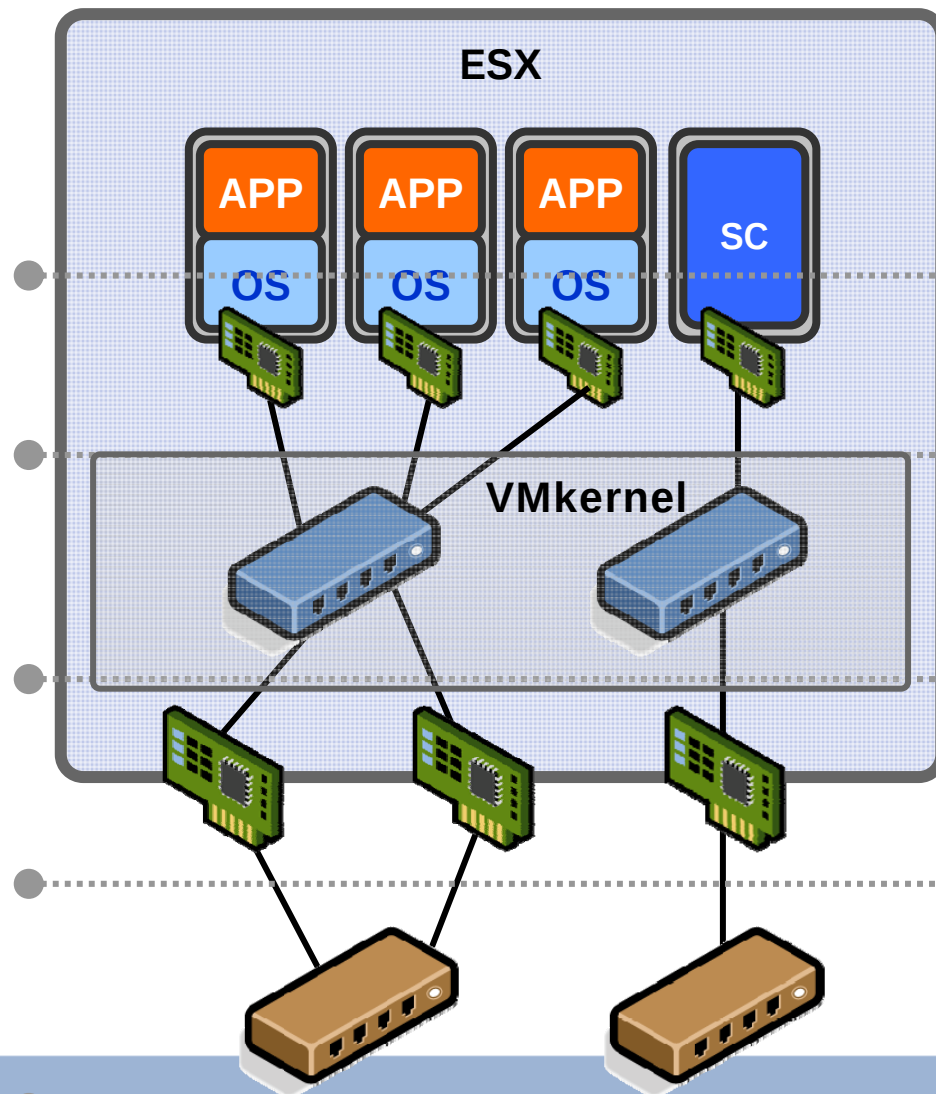
- クラスタウェアからの独立
- コストを抑えた高可用性を実現
- 仮想マシン単位のフェイルオーバーも試験的に実装

VMware HAの効果



VMwareの仮想ネットワークの考え方

VI3 のネットワーク アーキテクチャ



仮想マシン及びサービス・コンソール
仮想ネットワークデバイスを用いての
アクセスを要求

仮想NIC
VMkernelと仮想マシンの間に介在し、
TCP/IPパケットの受け渡しを担当

仮想スイッチ
VI環境においての中核的なレイヤ。
I/Oエミュレーション等の重要機能を担当

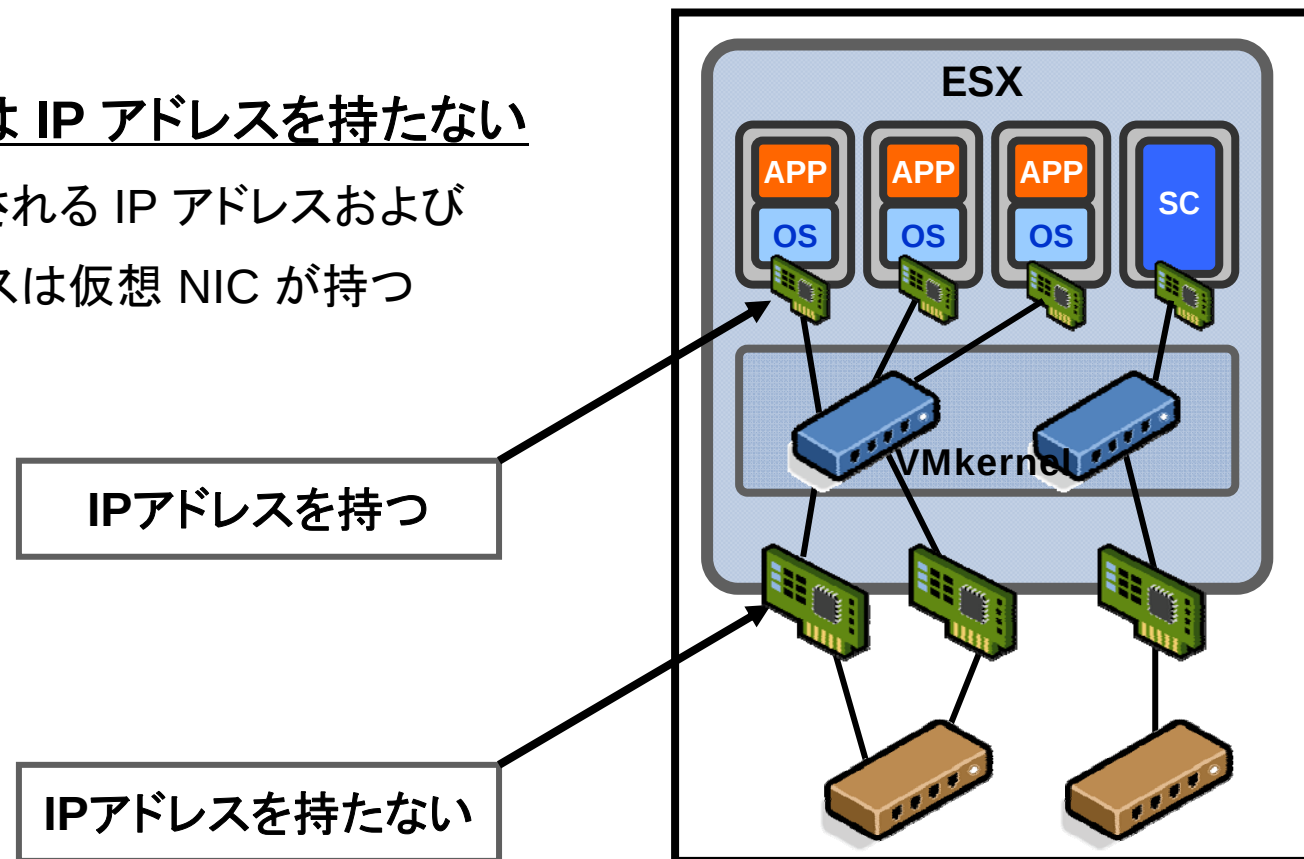
物理NIC
VI環境においてアーキテクチャとして
は何も機能を提供しないレイヤ

物理スイッチ
物理環境と同様の機能(例えばVLAN)
等の機能を提供するレイヤ

VI3 のネットワークの3つの基本ポイント

Q1) 物理NICはIPアドレスを持つのか？

- 物理 NIC は IP アドレスを持たない
通信で使用する IP アドレスおよび
MAC アドレスは仮想 NIC が持つ



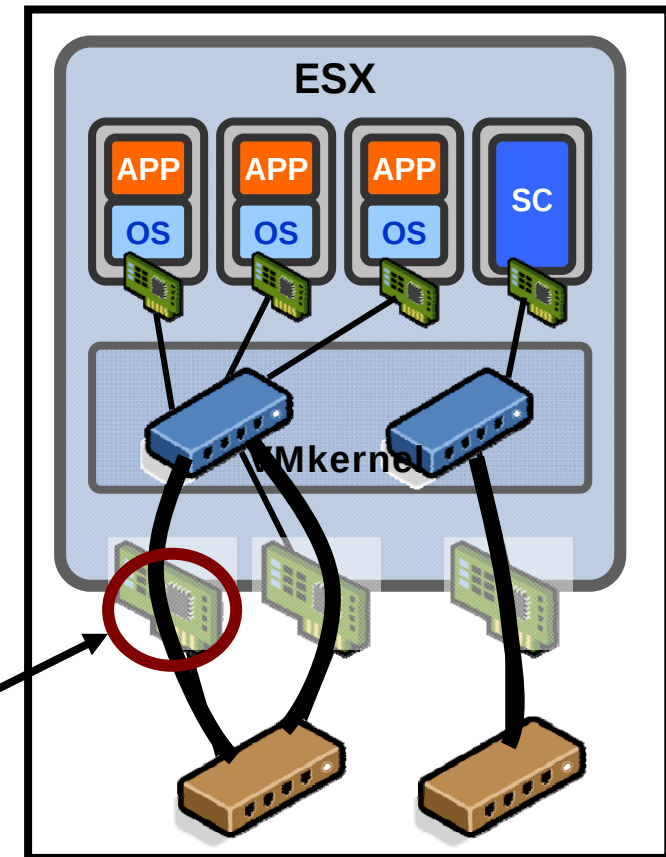
VI3 のネットワークの3つの基本ポイント

Q2) 物理NICはどんな役割を持つのか？

➤ 物理 NIC は「つなぐ」以外の役割を持たないとまずは考える

- 仮想スイッチと物理スイッチを接続するインタースイッチリンクに過ぎない
- つまり「物理NIC=ただのケーブル」
- 物理 NIC のリンクステータスは、ケーブルの抜き差しと同じ意味合い
- ただし、物理NICチップのソフトウェア処理のオフロードは利用している

ケーブルのパッチ接続と同等

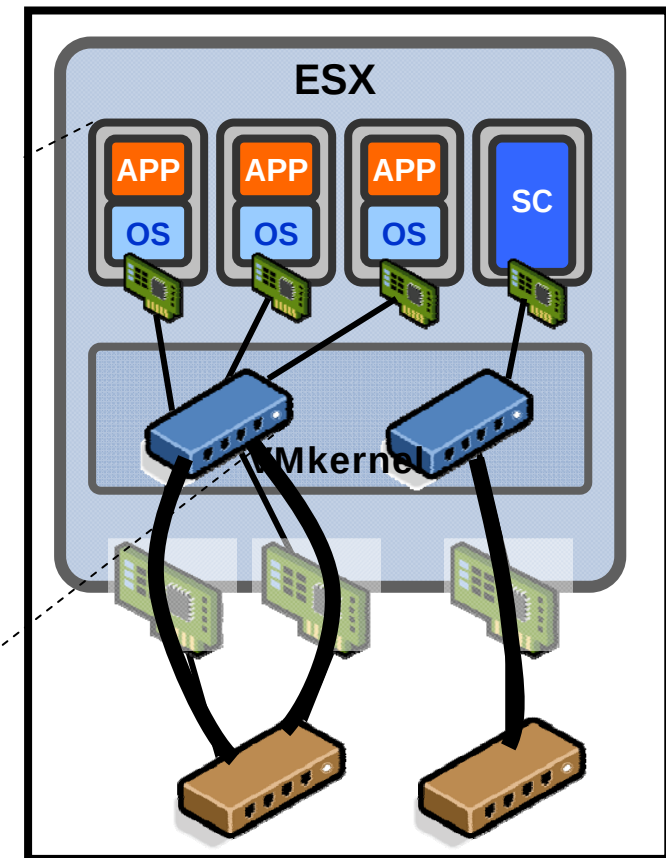
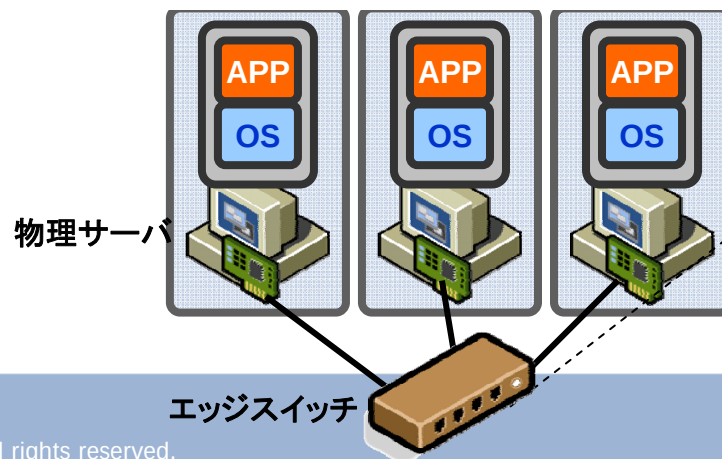


VI3 のネットワークの3つの基本ポイント

Q3) 仮想スイッチはパススルーするだけと考えればOK？

> 仮想スイッチは従来のスイッチの先に追加されたエッジスイッチ

- 仮想スイッチと物理スイッチはケーブルで直接つながっていると考える
- スwitch間接続なのでタグVLAN や Link Aggregation の設定対象となる
- STPのみ不要



ネットワーク設定画面

192.168.250.29 VMware ESX Server, 3.0.1, 32039

Summary Virtual Machines Resource Allocation Performance Configuration Tasks & Events Alarms Permissions Maps

Processors
Memory
Storage (SCSI, SAN, and NFS)
Networking
Storage Adapters
Network Adapters

Software
Licensed Features
DNS and Routing
Virtual Machine Startup/Shutdown
Security Profile
System Resource Allocation
Advanced Settings

Virtual Switch: vSwitch0 Remove... Properties...

Virtual Machine Port Group
VM2 Network
2 virtual machines | VLAN ID *
VM03-Solaris10
VM04-RH4U2

Virtual Machine Port Group
VM Network
1 virtual machines | VLAN ID *
VM03-Solaris10

Service Console Port
Service Console
vswif0 : 192.168.250.29

Physical Adapters
vmnic0 1000 Full

物理NIC

仮想スイッチ

仮想マシンと仮想NIC
VLANを指定可能(タグVLAN)

Virtual Switch: vSwitch1 Remove... Properties...

Virtual Machine Port Group
Virtual Machine Network 2
1 virtual machines | VLAN ID *
VM04-RH4U2

Physical Adapters
vmnic1 1000 Full
vmnic0 1000 Full

NICチーミング
(リンクアグリゲーション)

サーバ仮想化でネットワークはどう変わるか

仮想環境でのネットワーク設計のポイント

> 冗長性

- 複数の物理NICでチーミング (リンクアグリゲーション)

> 柔軟性

- タグVLANの使用

> 帯域の確保

- NICチーミングでのロードバランスの設定

冗長化

> 障害検知

- リンクステータス検知
- ビーコン検知
- リンクステートトラッキング
(ネットワークスイッチのポートダウン検知)

ロードバランシング(負荷分散)

> Active-Active構成

- ロードバランシングポリシー
 - ポートIDベース
 - 発信元MACハッシュベース
 - IPハッシュベース

> Active-Standby構成

- PortGroup設定による明示的な負荷分散

責任分解点と設定の整合性

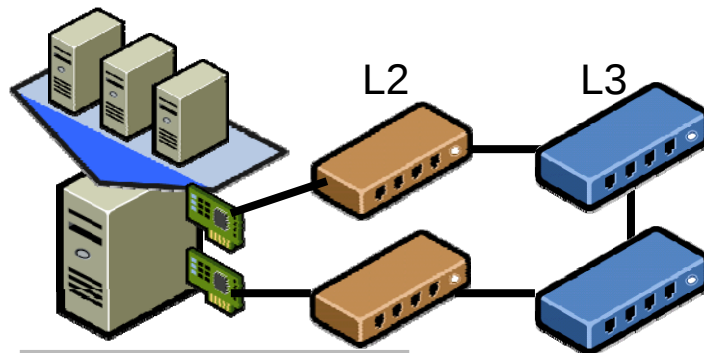
➤ ネットワークエッジの責任分解点と設定

- 仮想スイッチやブレードスイッチの出現
- エッジスイッチの設定の整合性の問題

➤ 整合性の問題例

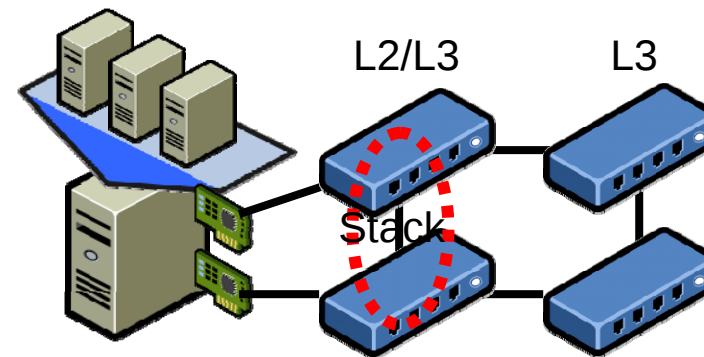
- NICチーミングのロードバランスの方式によって物理スイッチでのリンクアグリゲーションの設定の必要性の有無が異なる
- リンクアグリゲーションでのネゴシエーション(LACP, PAgP)は不可
- タグVLANでNative VLANは認識できない

物理ネットワーク構成パターン



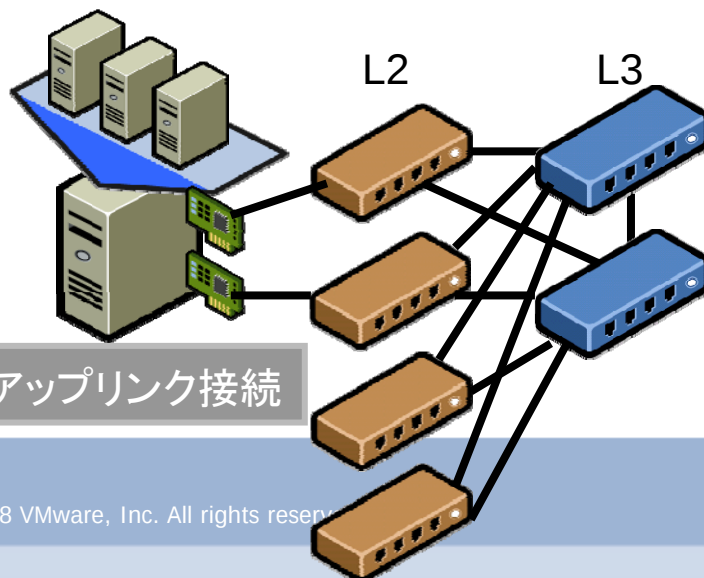
2台の独立スイッチ

(ポイント)リンクステートトラッキングが
設定可能なスイッチを選択する



スタック接続

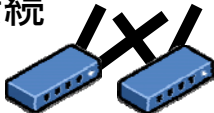
リンクアグリゲーションの設定可能
ループ部分はSTPを利用
(ポイント)スタック接続可能なスイッチを選択する



冗長アップリンク接続

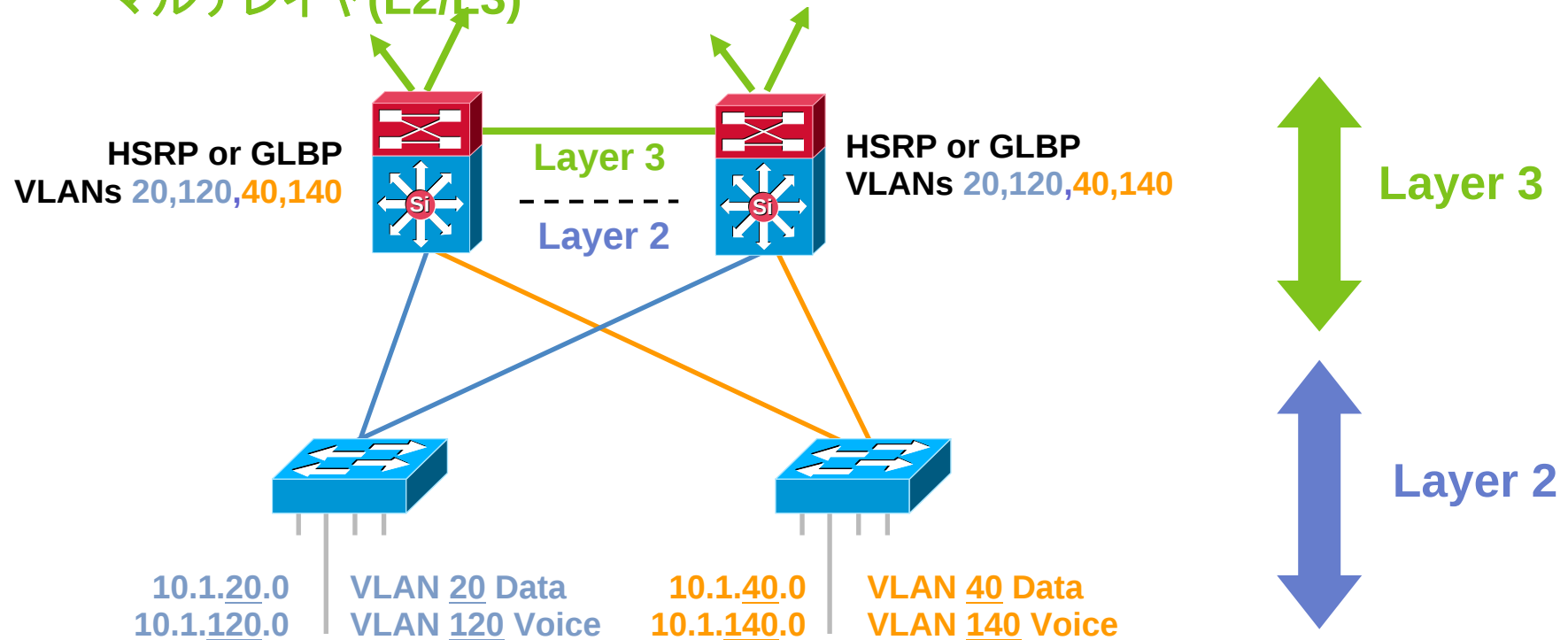
L2/L3間でSTPを設定
既存のネットワーク構成として最も一般的な構成
(ポイント)可能な限りRSTPやリンクステートトラッキング
を設定してもらう

物理スイッチ冗長化構成とロードバランシング

負荷分散方法	物理スイッチの冗長性		
	2台の独立スイッチ	冗長アップリンク接続	スタック接続
			
	ポートIDベース △	○	○
MACハッシュベース	△	○	○
IPハッシュベース	×	×	◎

- ◎ 物理スイッチの冗長性と効率の高い負荷分散の両方満たす
- 物理スイッチの冗長性と効率の低い負荷分散が可能
- △ 物理スイッチの冗長性に上位障害検知の設定の併用が推奨
- × 利用できない

現在のリファレンスデザイン マルチレイヤ(L2/L3)



GLBPやHSRPを利用してデフォルトGWを冗長化

ディストリビューションスイッチ間にリンクを設ける(Summarizationによるブラックホールを回避するため)

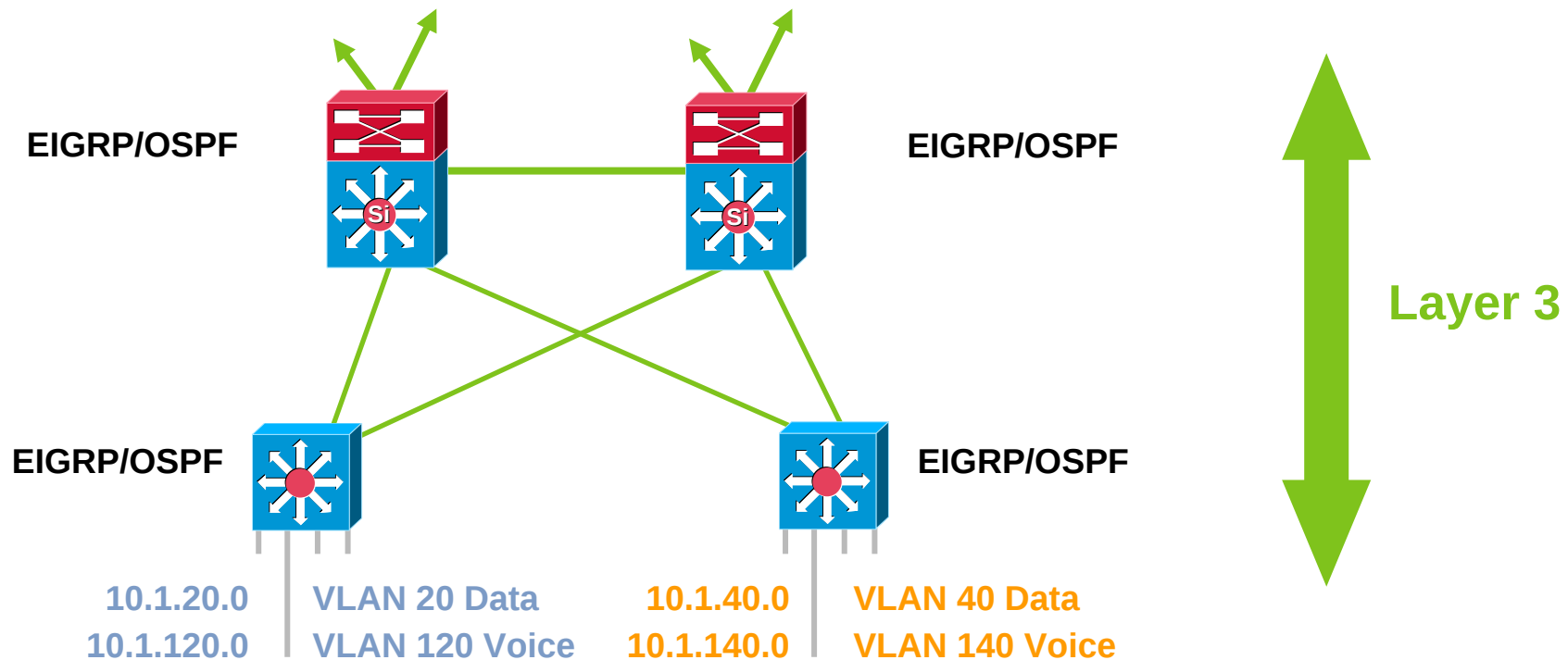
ディストリビューションスイッチのアクセス側ポートで hello パケットを交換しない設定が必要

非対称ルーティングになった場合を考慮した ARP テーブルの Aging Timer のチューニングが必要

STPの収束はUplinkの障害/復旧には関与しない (UplinkFast等での対応が必要)

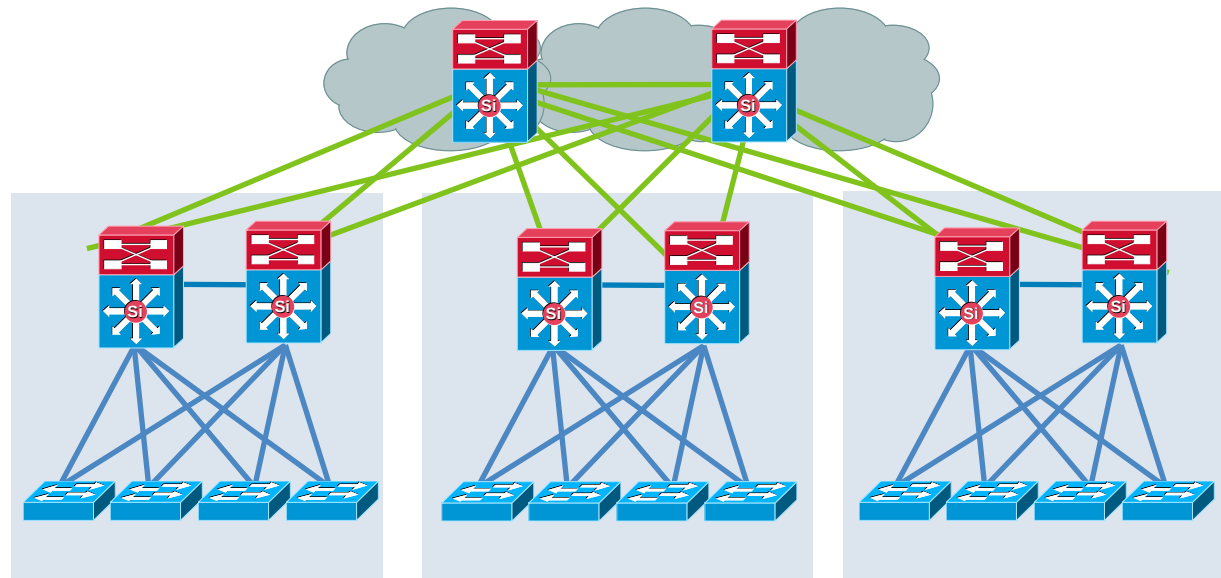
© 2019 VMware, Inc. All rights reserved. **MHSRP の必要性やディストリビューションスイッチ間の L3/L2 の別についてはアクセススイッチをまたぐ VLAN の有無に依存するが、またがない構成が推奨される**

Access層設計の新しい選択肢は仮想環境に向かない L3アクセス デザイン

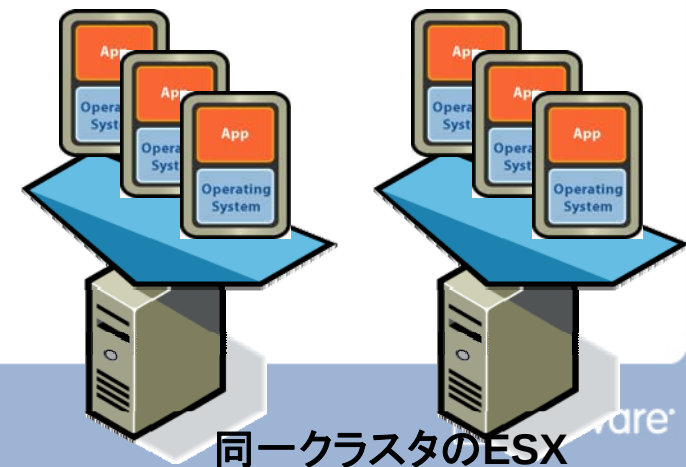


アクセス層にも L3 スイッチを使用しルーティングプロトコルを動作させる
L2 のテクノロジー固有の制約や注意点から解放される
このようなアクセス層を Routed Access と呼ぶ

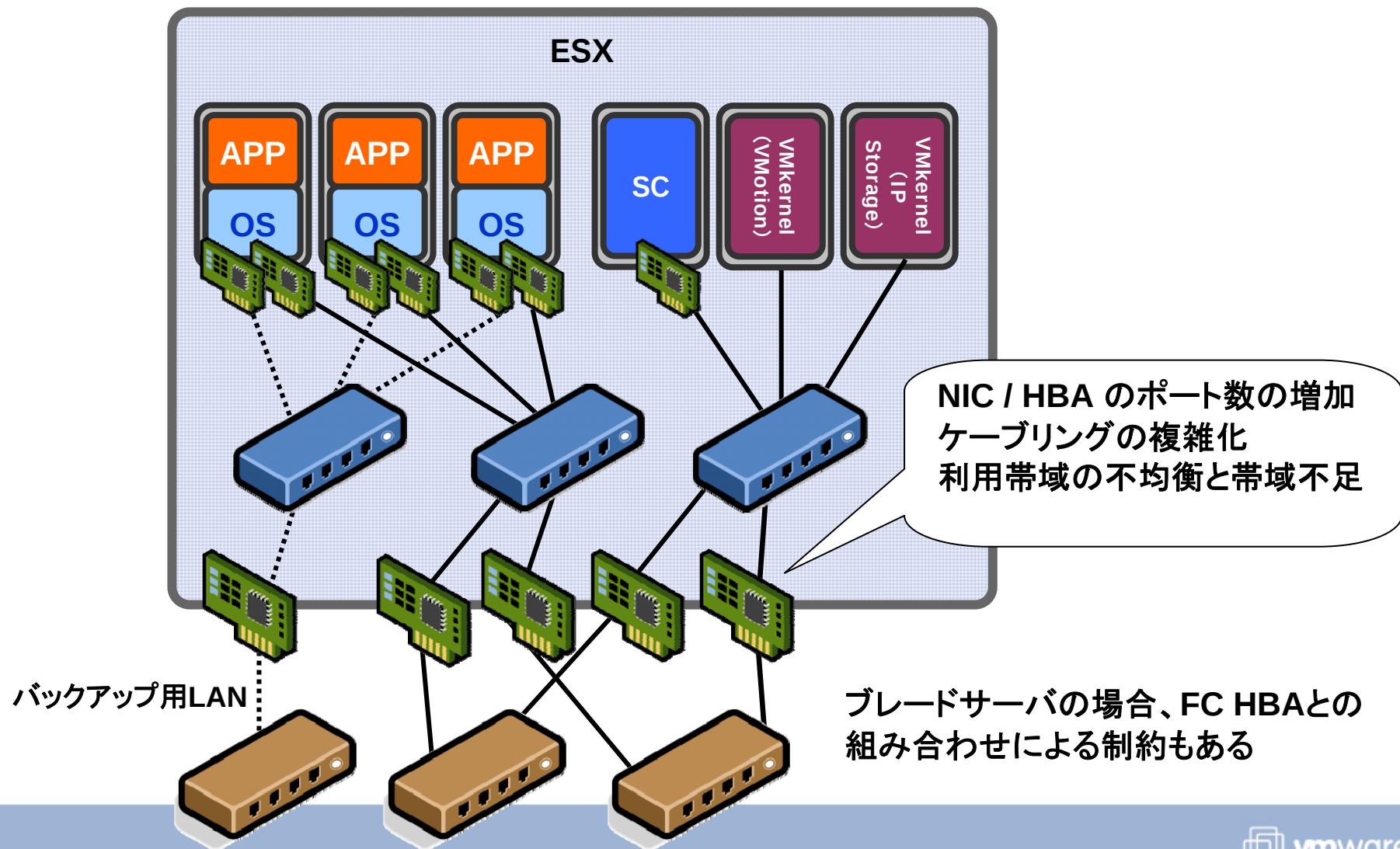
階層型ネットワークデザインと仮想環境



ディストリビューション層のスイッチとESXのクラスタの分布を一致させることで、VLANを透過させる範囲を限定しつつ、VMotion / VMware HAの柔軟性を確保



ESXのネットワーク構成 – 複雑な例



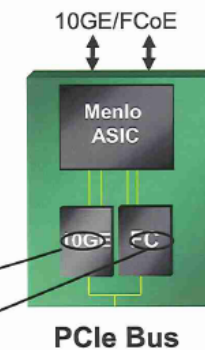
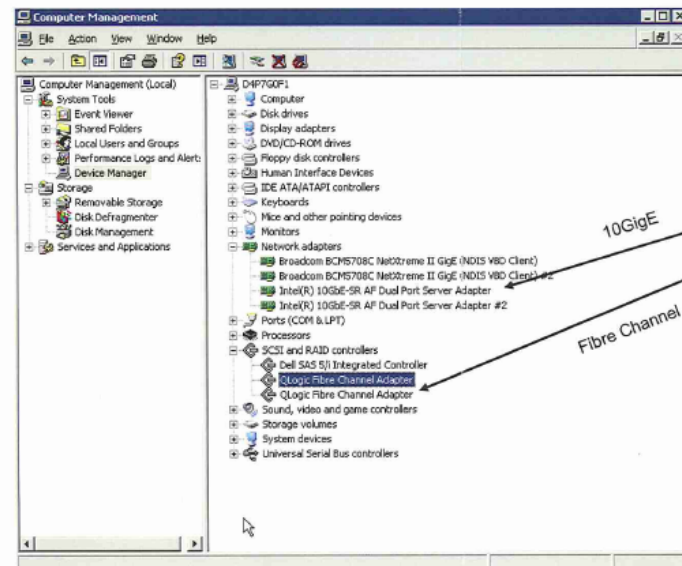
仮想環境で求められる次世代I/O規格

➤ FCoE (FC over Ethernet)

- 既存技術/環境との親和性が高い
- 10Gbpsと物理スイッチのポート集約率がボトルネックの可能性

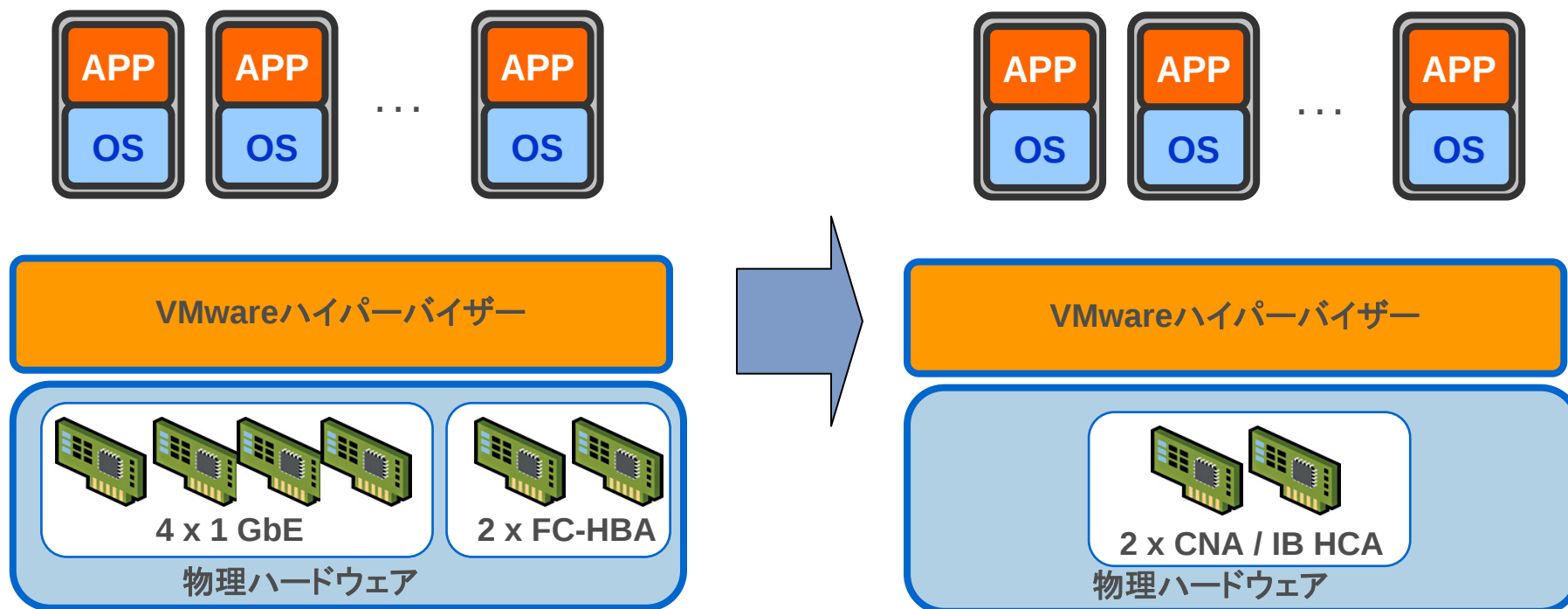
➤ Infiniband

- HPC分野で実績多数
- 広帯域、低遅延、帯域保証にアドバンテージ



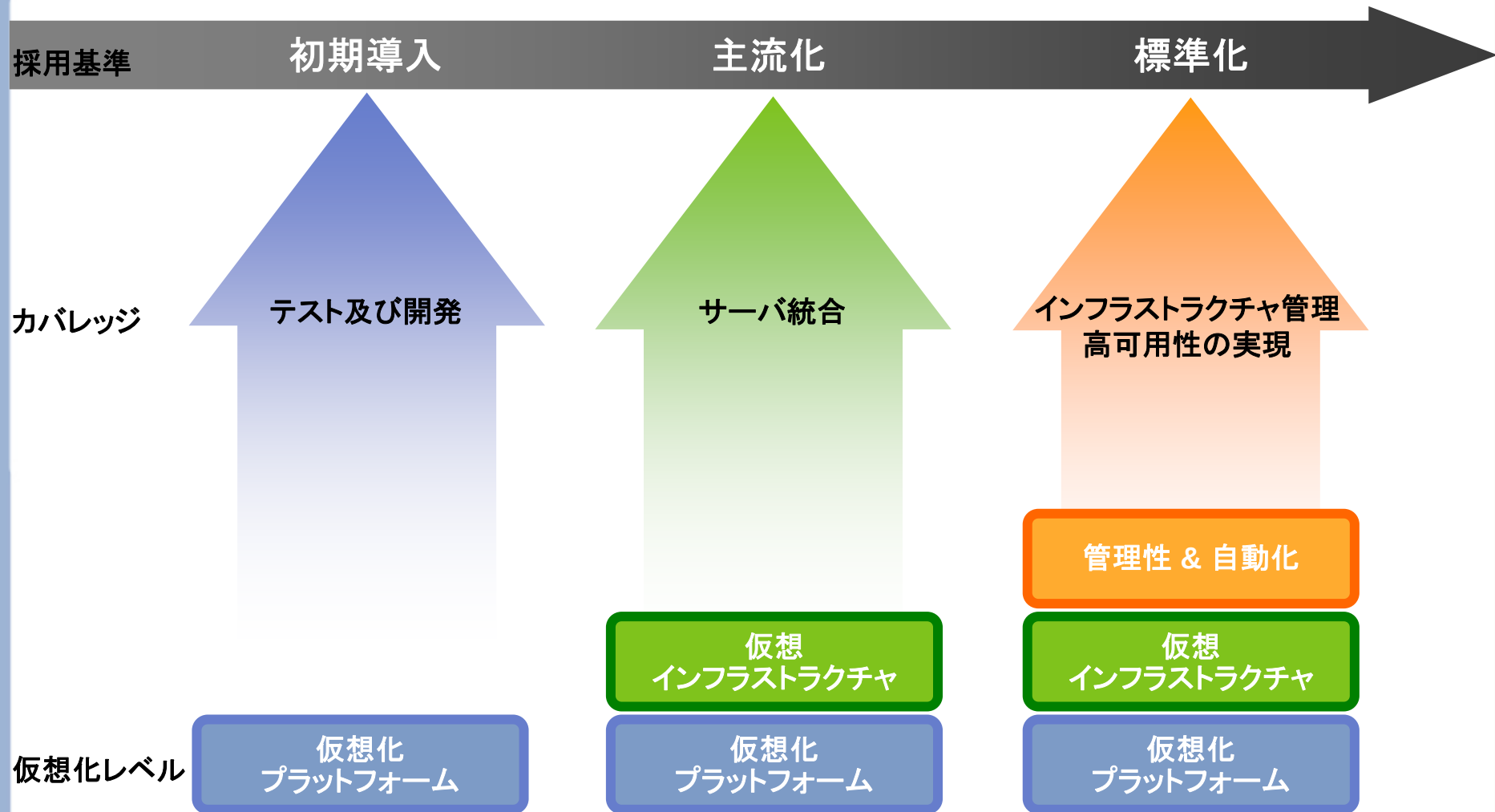
CNA の認識例

FCoE / Infinibandの導入前と導入後



仮想化されたデータセンターの今後とVMware

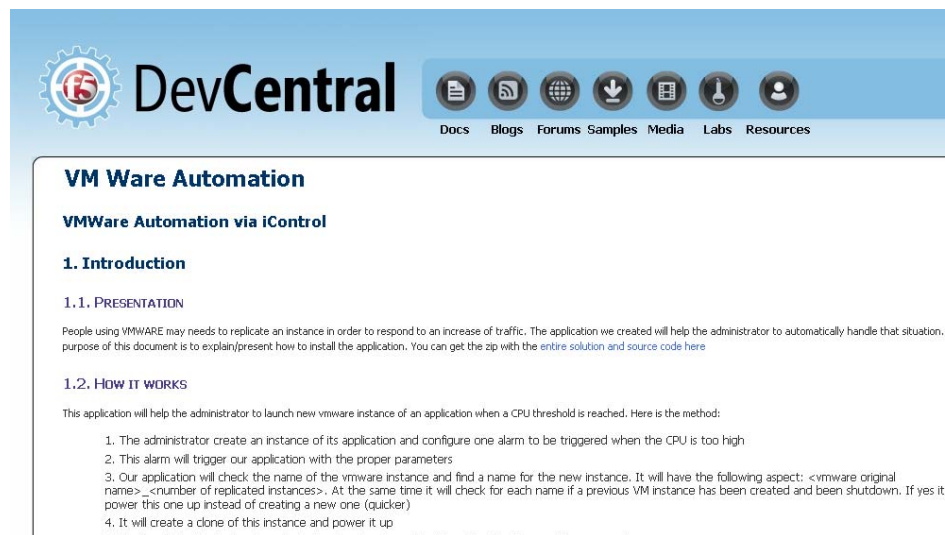
サーバ仮想化技術の潮流



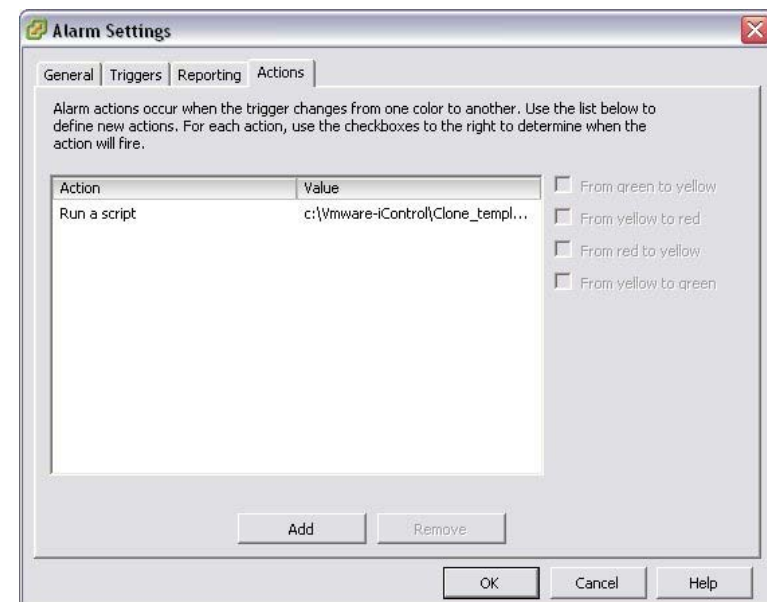
データセンターにおける自動化の例

➤ VMWare Automation via iControl (F5 Networks - DevCentral)

■ <http://devcentral.f5.com/wiki/default.aspx/iControl/VMWareAutomation.html>



The screenshot shows the DevCentral website interface. At the top, there's a navigation bar with the DevCentral logo and links for Docs, Blogs, Forums, Samples, Media, Labs, and Resources. Below this, the main content area is titled "VM Ware Automation" and "VMWare Automation via iControl". Under "1. Introduction", there's a sub-section "1.1. PRESENTATION" which explains that the application helps administrators handle high CPU usage by replicating VM instances. Below that, "1.2. HOW IT WORKS" describes the process: 1. Administrator creates an instance and configures an alarm for high CPU. 2. The alarm triggers the application with proper parameters. 3. The application checks the VM instance name and finds a name for the new instance. 4. It creates a clone of the instance and powers it up.

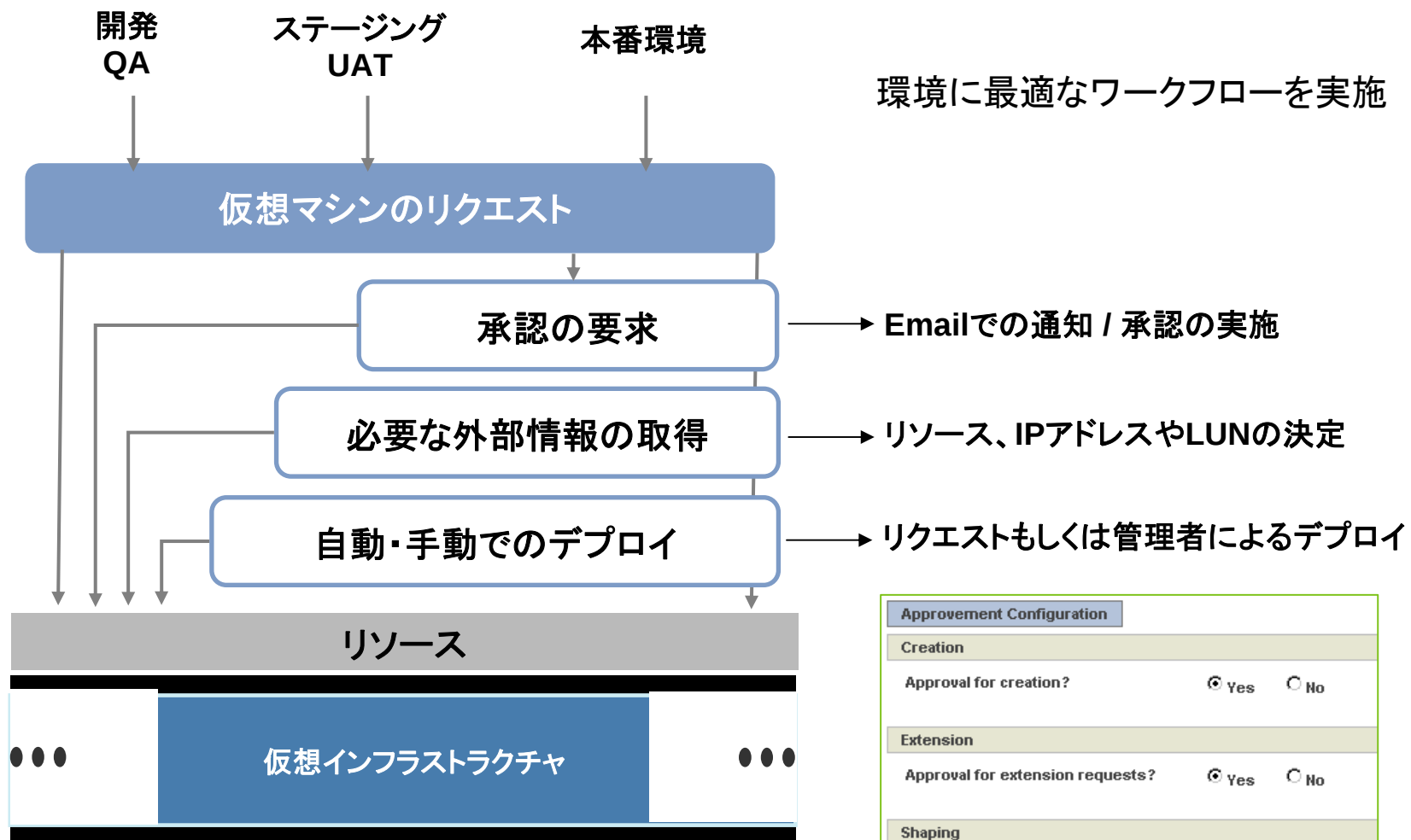


The screenshot shows the "Alarm Settings" dialog box with the "General" tab selected. It contains a table for defining alarm actions based on color changes. The table has two columns: "Action" and "Value". The first row shows the action "Run a script" with the value "c:\Vmware-iControl\Clone_templ...". To the right of the table, there are four checkboxes for color transitions: "From green to yellow", "From yellow to red", "From red to yellow", and "From yellow to green". At the bottom of the dialog, there are "Add", "Remove", "OK", "Cancel", and "Help" buttons.

Action	Value
Run a script	c:\Vmware-iControl\Clone_templ...

☐ From green to yellow
☐ From yellow to red
☐ From red to yellow
☐ From yellow to green

VMwareが提供するワークフローの自動化と監査



管理と自動化を実現する製品群

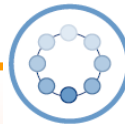
Automation = Business Agility

- Automate IT processes
- Create resource pools
- Capacity on-demand

Management &
Automation

Virtual
Infrastructure

Hypervisor



IT Service
Delivery

Lifecycle Manager
Lab Manager
Stage Manager



Business
Continuity

Site Recovery Manager



vmware®