

# クラウドネットワークの根幹を変える 仮想ルーターソリューション

Virtual Router Solution for Cloud Networking

2010年9月30日

さくらインターネット研究所

上級研究員 松本直人

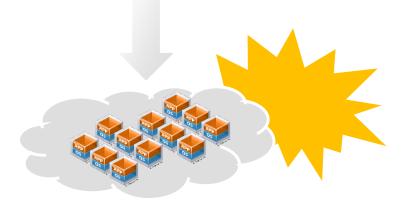
# 仮想ルータが求められる背景



従来のコンピューティング Traditional Computing



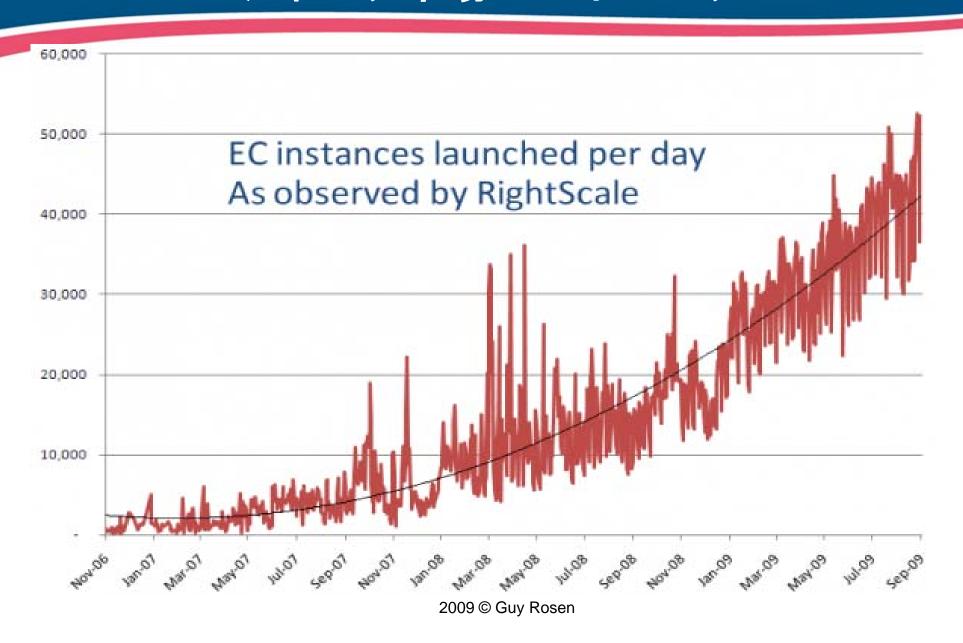
クラウド・コンピューティング Cloud Computing



クラウド中で動作するネットワーク部品が必要となる Require Network Component of The Cloud Computing Environment

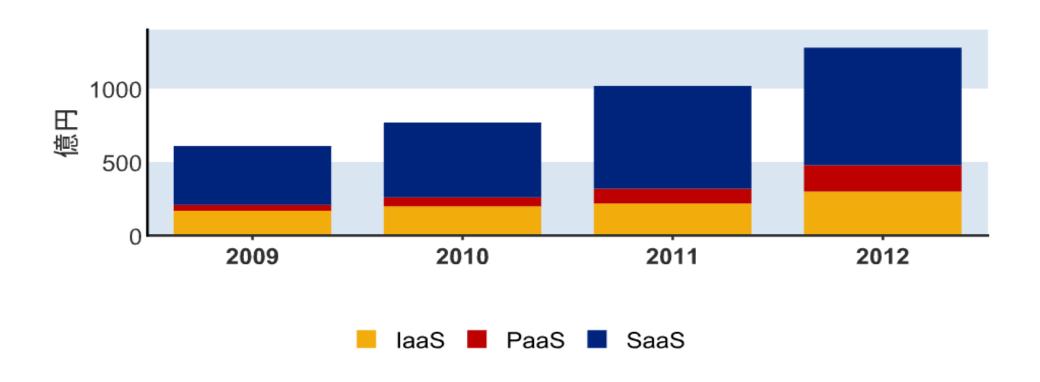
# コンピューティング市場へのインパクト





# クラウド・コンピューティング市場動向





Source: IDC Japan, 7/2009

\*laaS: Infrastructure as a Service PaaS: Platform as a Service SaaS: Software as a Service

SaaSが市場を牽引・緩やかに成長

# 仮想ルータ導入による効用



システム省力化 Reduse



顧客毎の既存ルータ Traditional Router hardware

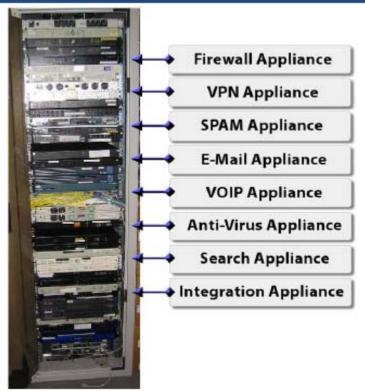


ハイパーバイザー上に 仮想ルータ集約

# ネットワーク機器の仮想化

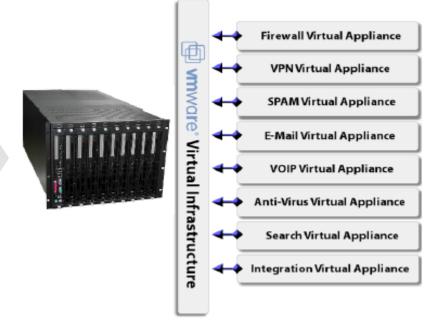


#### 既存のシステム構成



- · One-to-one ratio of function to device
- Support from multiple sources
- Inefficient utilization of hardware

#### 仮想化技術で省力化



- Pools of computing resources
- Standardize on preferred hardware
- Simplify operational support
- · Decrease footprint of each solution
- Hardware support from preferred vendor
- · Efficient utilization of hardware

# インフラのコモディティ化

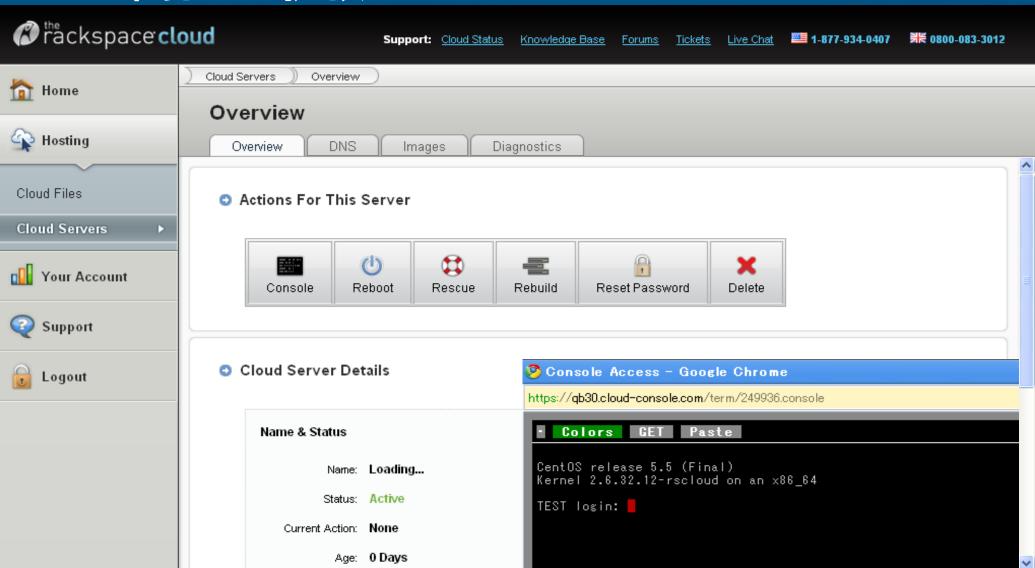






# laaS環境への拡張





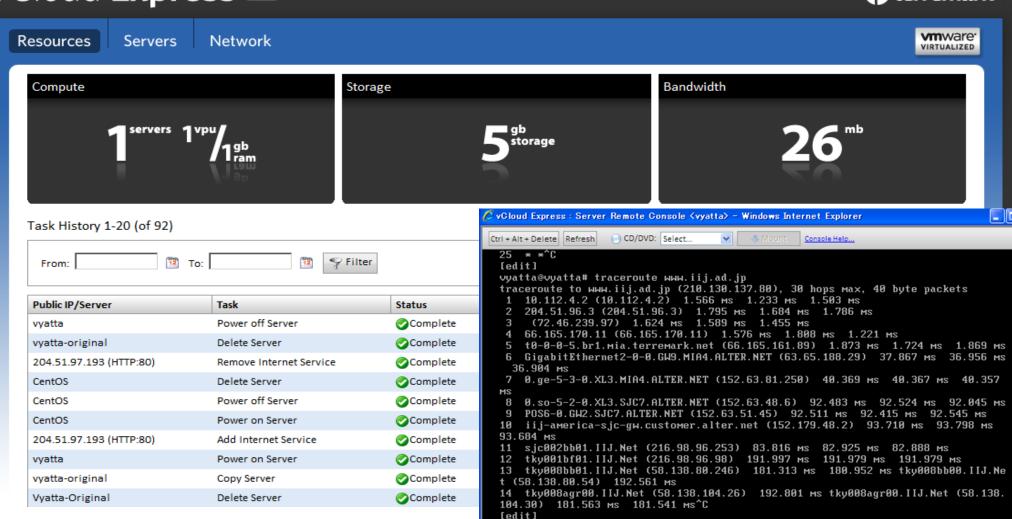
# 増えるパブリッククラウド環境

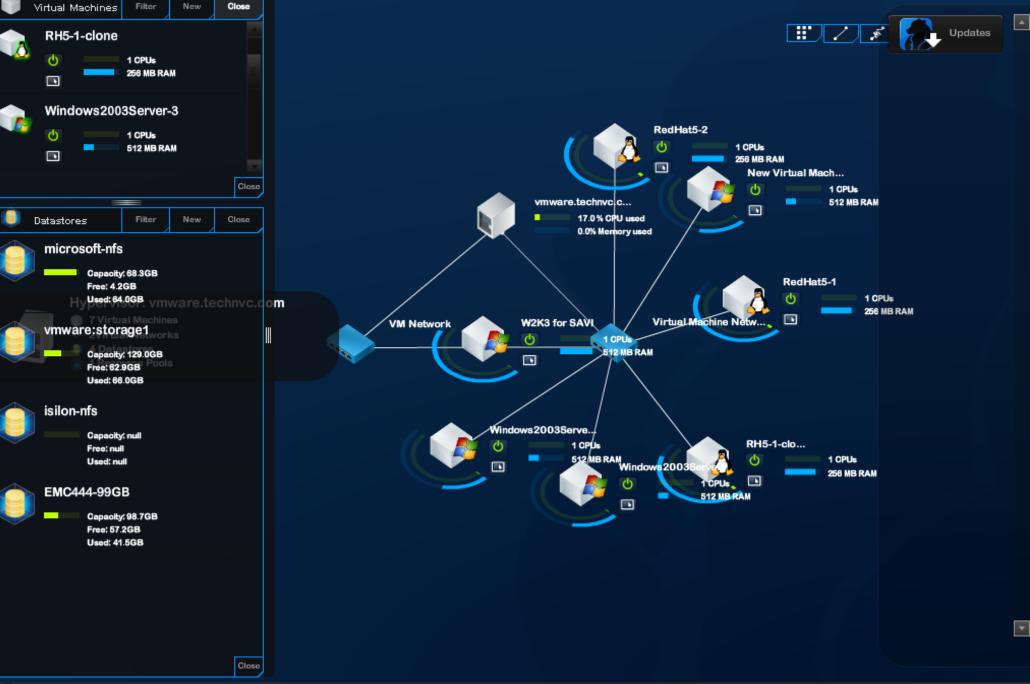


ABOUT VCLOUD EXPRESS COMMUNITY SUPPORT MY CLOUD ENVIRONMENT MY ACCOUNT WELCOME NAOTO SIGN OUT

# vCloud **Express BETA**







# Vyattaが持つ機能



IPv4 / IPv6 Routing	» BGPv4, BGPv6 » OSPFv2, OSPFv3*	» RIPv2 » Static Routes	» IPv6 Policy » IPv6 SLAAC	
IP Address Management	» Static » DHCP Server » DHCP Client	<ul><li>» DHCP Relay</li><li>» Dynamic DNS</li><li>» DNS Forwarding</li></ul>	» DHCPv6 Server » DHCPv6 Client » DHCPv6 Relay	
Encapsulations	<ul><li>» Ethernet</li><li>» 802.1Q VLANs</li><li>» PPP</li></ul>	» PPPoE » IP in IP » Frame Relay	» MLPPP » HDLC » GRE	
Firewall	<ul><li>» Stateful Inspection Firewall</li><li>» Zone-based Firewall</li><li>» P2P Filtering</li></ul>	<ul><li>» IPv6 Firewalling</li><li>» Time-based Firewall Rules</li><li>» Rate Limiting</li></ul>	<ul><li>» ICMP Type Filtering</li><li>» Stateful Failover</li></ul>	
Tunneling / VPN	<ul><li>» SSL-based OpenVPN</li><li>» Site to Site VPN (IPSec)</li><li>» Remote VPN (PPTP, L2TP, IPSec)</li></ul>	<ul><li>» OpenVPN Client Auto-Configuration</li><li>» Layer 2 Bridging over GRE</li><li>» Layer 2 Bridging over OpenVPN</li></ul>		
Additional Security	<ul><li>» Network Address Translation</li><li>» Sourcefire VRT Intrusion Prevention</li><li>» VyattaGuard Web Filtering</li></ul>	<ul><li>» DES, 3DES, AES Encryption</li><li>» MD5 and SHA-1 Authentication</li><li>» RSA, Diffie Helman Key Mgmt</li></ul>	» NAT Traversal » Role based access control	
WAN / LAN Device Drivers	» WAN Device Drivers - ADSL, T1, T3 » Intel 10/100Mbps - 10Gbps	» IEEE 802.11 wireless » Drivers in 2.6.31 Linux Kernel	» Synchronous Serial - V.35, X.21, RS-422, EIA530	
Performance Optimization	<ul><li>» WAN Link Load Balancing</li><li>» Ethernet Link Bonding</li><li>» Web Caching</li></ul>	<ul><li>» MLPPP</li><li>» ECMP</li><li>» Bandwidth Management</li></ul>		
QoS Policies	<ul><li>» Priority Queuing</li><li>» Network Emulator</li><li>» Round Robin</li></ul>	<ul><li>» Random / Weighted Random</li><li>» Classful Queuing</li><li>» Ethernet Header Matching</li></ul>	<ul><li>» VLAN Tag</li><li>» IPv6 Address</li><li>» Port Mirroring</li></ul>	
High Availability	<ul><li>» Stateful Firewall / NAT Failover</li><li>» VRRP</li><li>» HA Clustering</li></ul>	<ul><li>» Configuration Replication</li><li>» RAID 1</li></ul>	<ul><li>» IPSec VPN Clustering</li><li>» Protocol Fault Isolation</li></ul>	
Administration & Authentication	<ul><li>» Integrated CLI</li><li>» Web GUI</li><li>» Vyatta Remote Access API</li></ul>	<ul><li>» Telnet</li><li>» SSHv2 / SSH Public Key</li><li>» Binary Image Install</li></ul>	<ul><li>» RADIUS</li><li>» TACACS+*</li><li>» Single Configuration File</li></ul>	
Diagnostics & Logging	<ul><li>» tcpdump</li><li>» Wireshark Packet Capture</li><li>» BGP MD5 Support</li></ul>	<ul><li>» Serial Loopback Commands</li><li>» Netflow / sFlow</li><li>» LLDP</li></ul>	<ul><li>» Syslog</li><li>» SNMPv2c</li><li>» SNMP for IPv6</li></ul>	

# クラウド環境支えるソフトウェア



#### 商用製品

- Citrix XenServer
- Eucalyptus Enterprise Edition
- IBM z/VM
- RedHat KVM
- Microsoft Hyper-V Server
- Oracle Solaris Container
- Parallels Virtuozzo Containers
- VMware vSphere/vCloud

### オープンソース

- Eucalyptus (Open Source)
- OpenStack
- OpenVZ
- OpenNebula
- Iguest
- Linux-Vserver
- Linux KVM
- Xen Cloud Platform



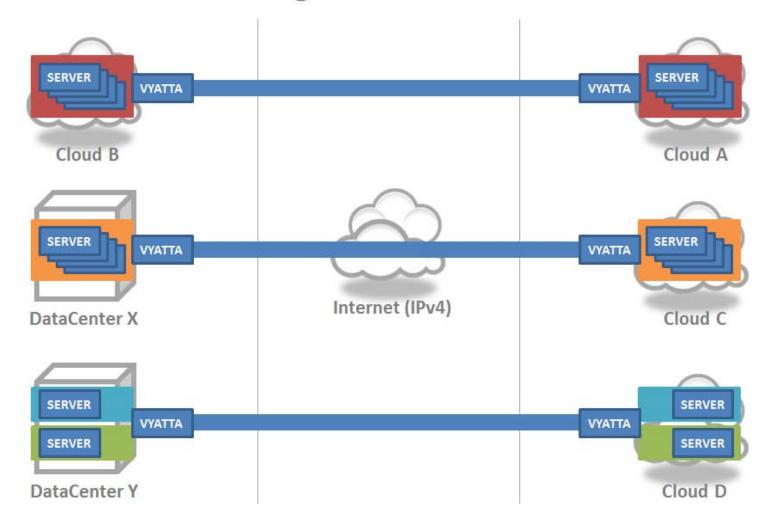
# 仮想ルーターの利用例

Virtual Router Use Case on the Cloud

# クラウド・ブリッジングという使い方

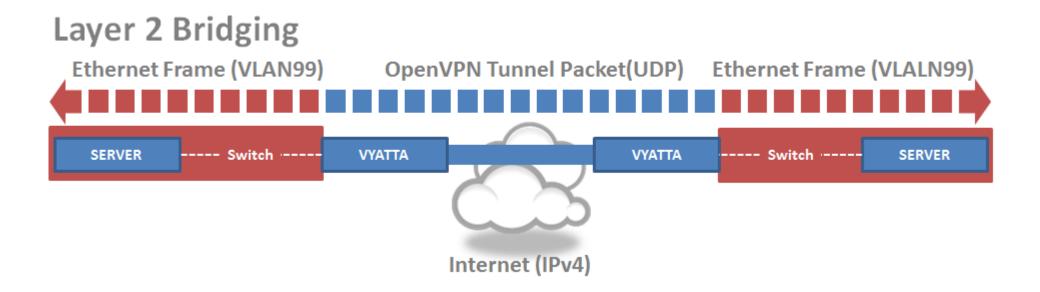


#### **Inter-Cloud Networking Model**



# VyattaのOpenVPN機能を使った例

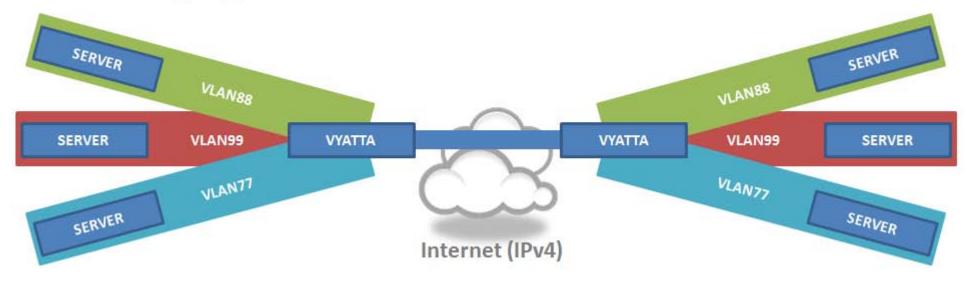




# VLANを広げるVyattaの使い方







VLAN番号枯渇対策

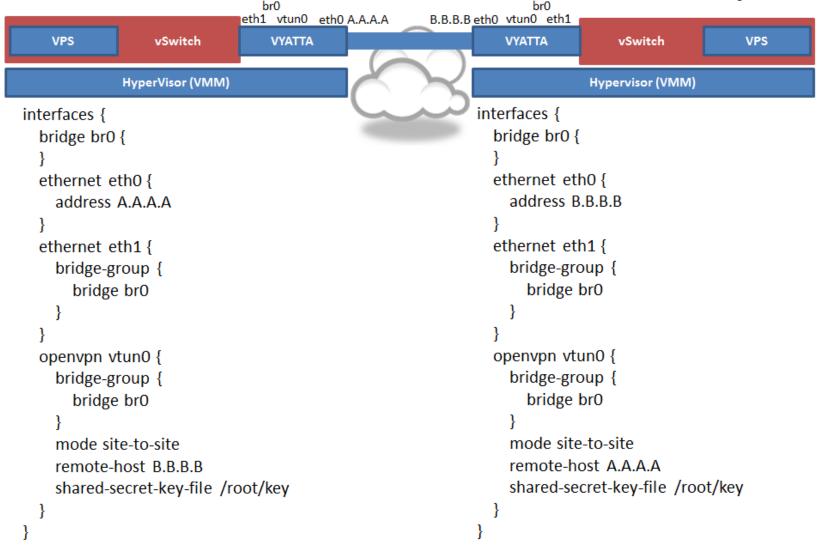
# クラウド・ブリッジングの設定例



#### **VYATTA: Layer 2 Bridging Sample Config**

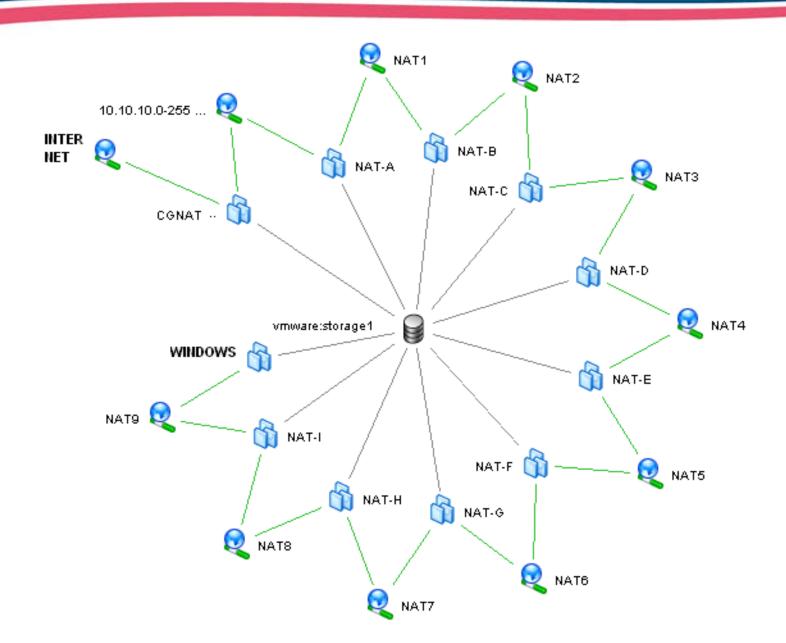
vSwitch: Virtual Switch on HyperVisor (VMM) VPS: Virtual Private Server

VMM: Virtual Machine Manager



# 仮想ルータを使った10段NAT







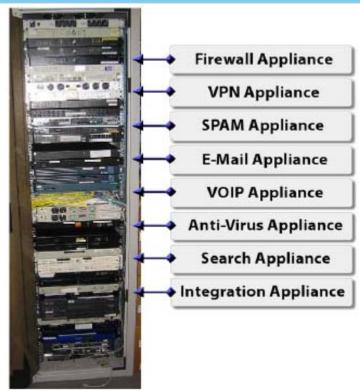
# 仮想ルーター・オペレーションTIPS

Virtual Router Operation TIPS

# ネットワーク機器の仮想化

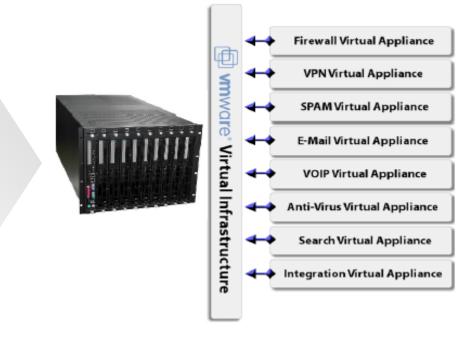


#### 既存のシステム構成



- · One-to-one ratio of function to device
- Support from multiple sources
- Inefficient utilization of hardware

### 仮想化技術で省力化



- Pools of computing resources
- Standardize on preferred hardware
- Simplify operational support
- · Decrease footprint of each solution
- Hardware support from preferred vendor
- Efficient utilization of hardware

# 人間の目による世界

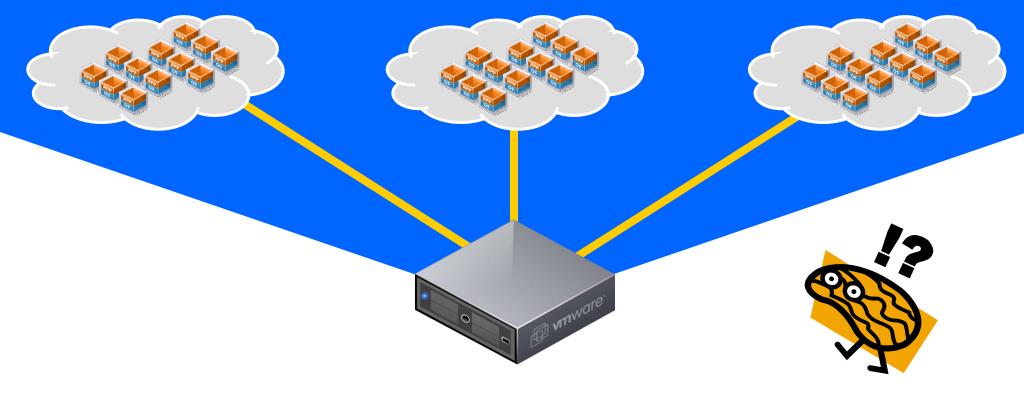




この中で仮想化ネットワークが存在します。

# 仮想化の世界

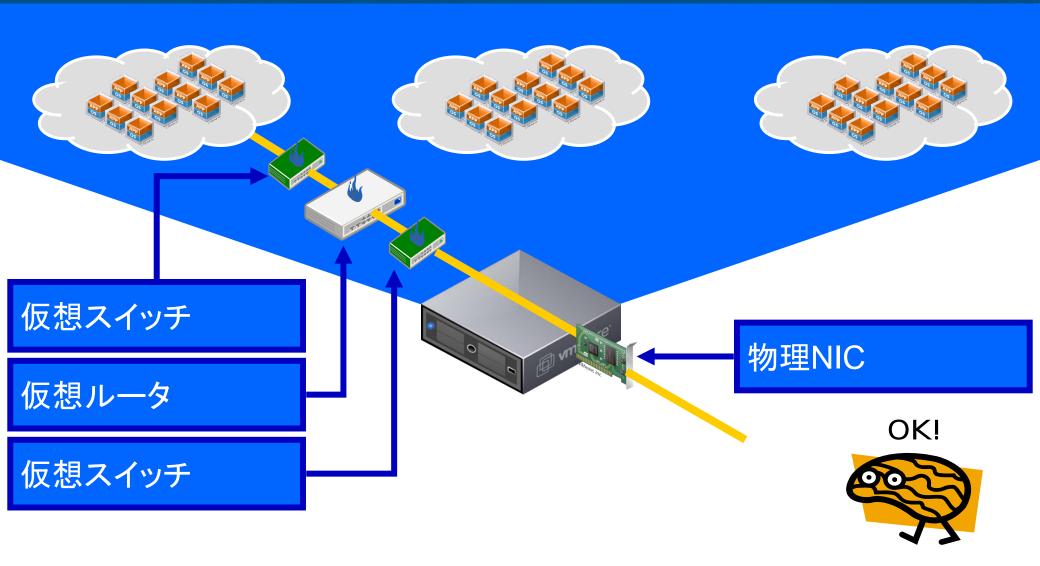




異なるネットワークが一箱に

# 仮想化システムの把握手順

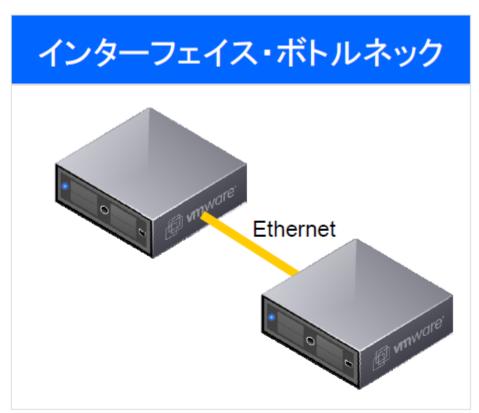




### 仮想ルータの性能課題を把握





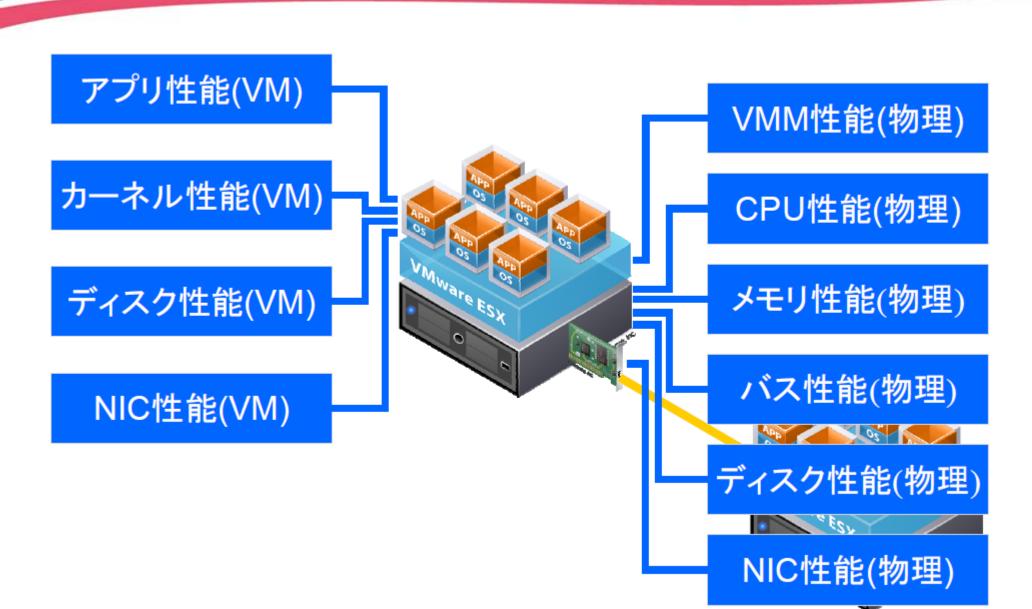




仮想化技術であってもパフォーマンスは、コンポーネントの性能限界につよく依存している

# 性能と依存関係を示すポイント





## ソフトウェアである性能の壁



アプリ性能(VM)

カーネル性能(VM)

ディスク性能(VM)

NIC性能(VM)

Task done by CPU core

CPU idle (waiting of memory)

Process frame 0

CPU idle (waiting of memory)

Process frame 1

Task done by cache controller

Fetch frame 0

VMM性能(物理)

CPU性能(物理)

メモリ性能(物理)

Fetch frame 1

バス性能(物理)

ディスク性能(物理)

4.52×

NIC性能(物理)

# チューニングすべきポイント

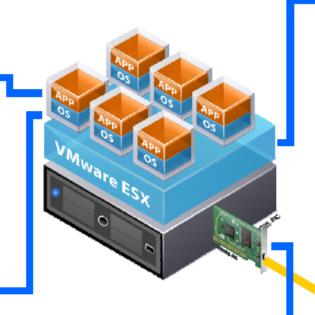


カーネル性能(VM)

カーネル パラメーター 設定

NIC性能(VM)

NICドライバ 設定



VMM性能(物理)

CPU アフィニティ 設定

NICドライバ キュー設定 (VMDq)

NIC性能(物理)

Intel Virtual Machine Device Queues (VMDq)

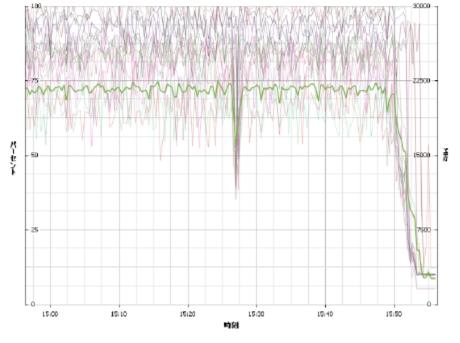
## 仮想化インフラ全体を外観する方法

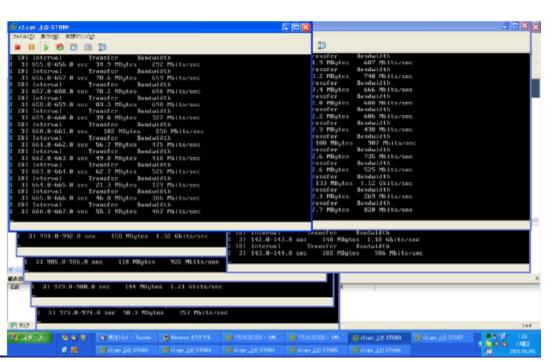


6:55:30pm up 2 days 9:42, 93 worlds; CPU load average: 0.58, 0.23, 0.09

PORT ID	UPLINK	USED BY	DTYP	DNAME	PKTTX/s	MbTX/s	PKTRX/s MbRX/	s %DRPTX	%DRPRX
33554433	Y	vmnic4	Н	vSwitch2	45899.41	23.53	634802.70 7329.4	0.00	0.00
33554445	N	1123:TEST0	H	vSwitch2	13399.41	6.80	92695.80 1073.5	5 0.00	0.01
33554440	N	1103:TEST4	Η	vSwitch2	12933.10	6.66	94841.79 1098.4		0.29
33554438	И	1093:TEST2	Η	vSwitch2	9470.70	4.90	94307.61 1092.2	1 0.00	0.21
33554437	И	1085:TEST1	Η	vSwitch2	2965.33	1.50	90999.51 1053.9	0.00	0.00
33554439	N	1098:TEST3	Η	vSwitch2	2647.95	1.36	96877.44 1121.9	8 0.00	0.01
33554444	N	1118:TEST6	Η	vSwitch2	2580.08	1.35	103975.09 1204.1	8 0.00	0.08

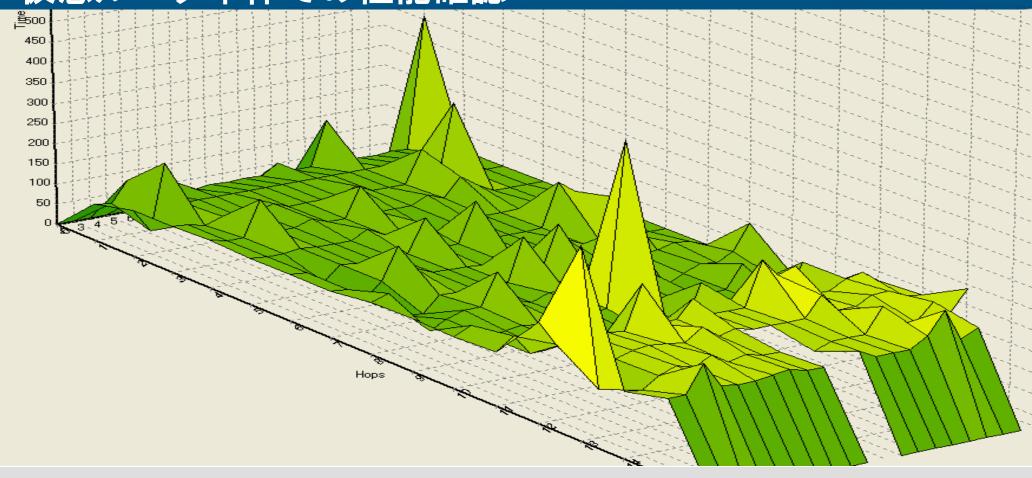
CPU/ ሃፖ/እትላጊ, 2009/05/07 14:56:19 - 2009/05/07 15:56:19 - r3.egn







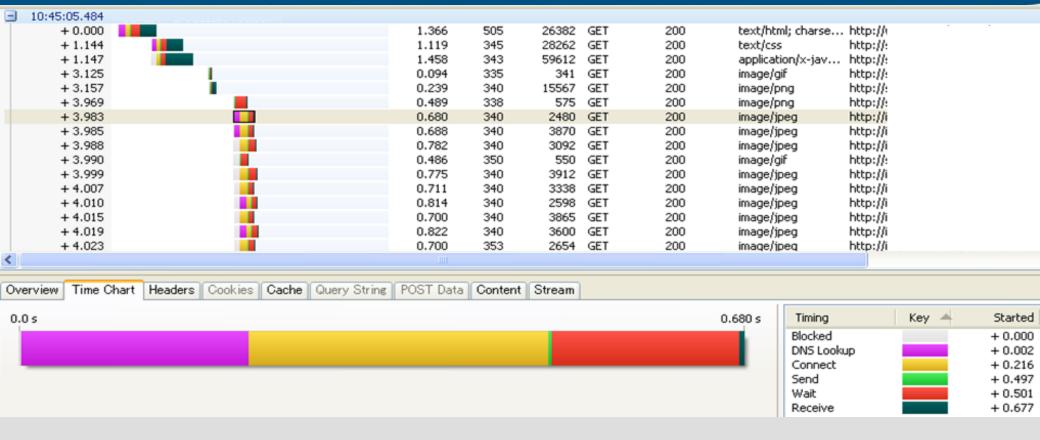




ネットワーク全体ではなくシステム単位での測定が基本

# 仮想ルータを通した全体システムの把握





個別システム・セッション単位での測定が基本



# ご清聴ありがとうございました