
Vyatta and Virtualization

- 仮想環境でのVyatta -

Japan Vyatta Users Group Meeting @ Hiroshima

海老澤 健太郎

Twitter: @ebiken
ebiken.g@gmail.com

クラウド・オートメーションを実現するアーキテクチャ

コントロールパネル



マーケットプレイス



ビジネスの自動化



Parallels
(Business) Automation



Parallels
Plesk Billing

クラウドパッケージ化及び配信



運用の自動化



Parallels
Operations
Automation



Parallels
Virtual
Automation



Parallels
Plesk
Products



Parallels
Small Business
Panel

シェアードウェブ
ホスティング

メッセージング&
コラボレーション

仮想インフラ
サービス

SaaS

仮想インフラストラクチャ

ハイパーバイザとOS仮想化の最良のコンビネーション



Parallels
Virtuozzo Containers



Parallels
Server Bare Metal

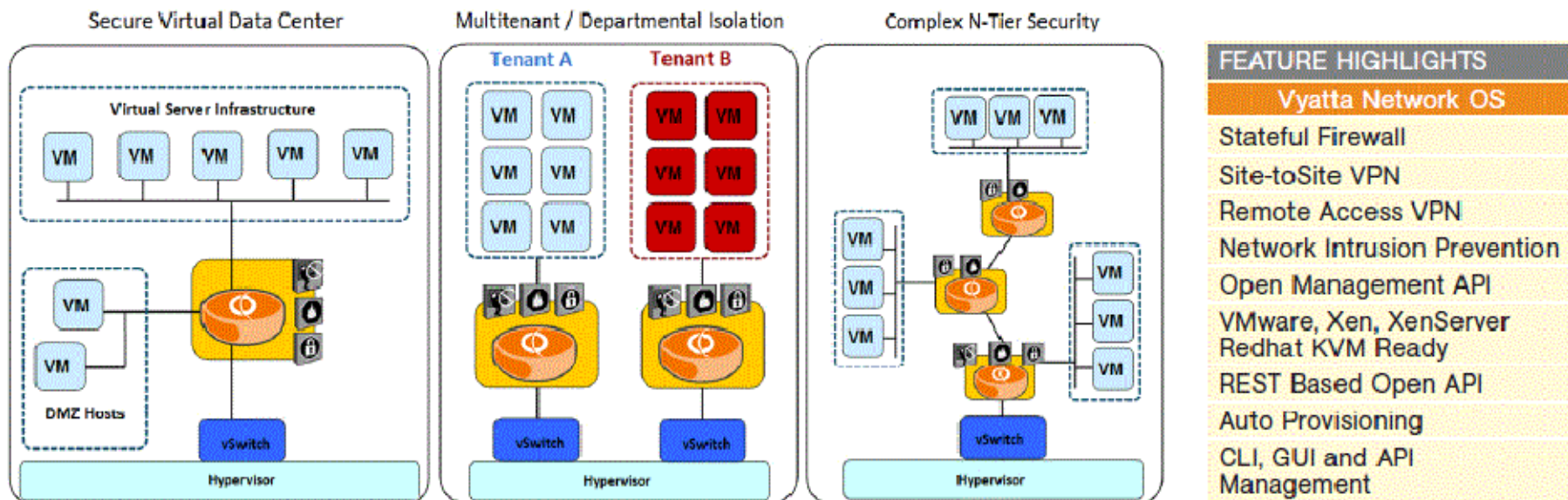
本日のお話し

1. 仮想データセンター
2. Vyatta転送性能(仮想化無し)
3. 仮想化環境におけるパケット転送のしくみ
 1. デバイス・エミュレーション
 2. I/O準仮想化
 3. I/O デバイス割り当て(VT-d)
 4. I/O デバイス割り当て＋共有(SR-IOV)
4. 仮想化技術をサポートするハードウェア

仮想データセンター

From: <http://www.vyatta.com/solutions/virtual/optimizedvirtualmachines>

Vyatta delivers the only multi-layer virtual network security solution that maintains compliance and enables instant migration of complex, layered firewall architectures from the physical network into any virtual data center, without compromises.



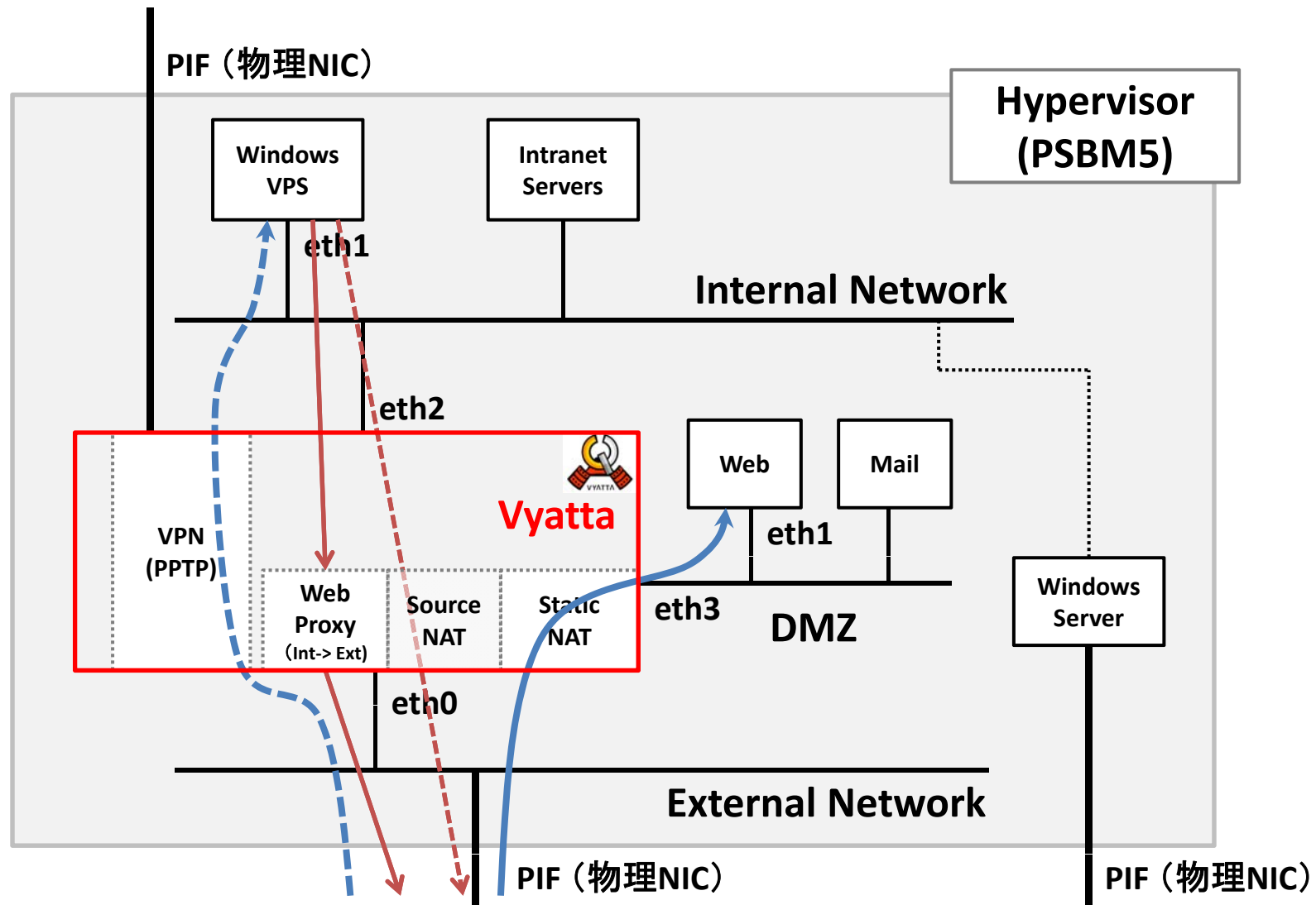
FEATURE HIGHLIGHTS

Vyatta Network OS

- Stateful Firewall
- Site-toSite VPN
- Remote Access VPN
- Network Intrusion Prevention
- Open Management API
- VMware, Xen, XenServer
- Redhat KVM Ready
- REST Based Open API
- Auto Provisioning
- CLI, GUI and API Management

1 888 VYATTA 1 (US & CANADA) » +1 650 413 7200 (INTERNATIONAL) » WWW.VYATTA.COM

仮想データセンター（OneBox設定例）

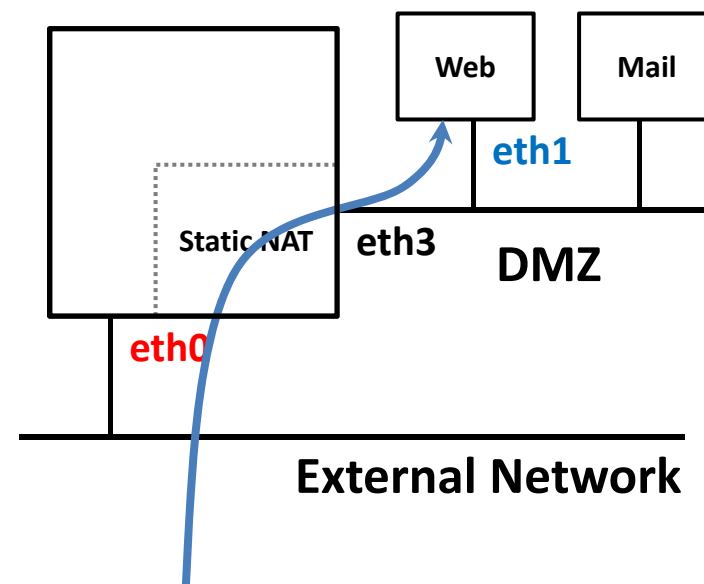


仮想データセンター（OneBox設定例）

Web Serverの公開

Static NAT設定:

```
set service nat rule 10
  destination address <external-address:eth0>
  destination port <http-port:80>
  inbound-interface <external-if:eth0>
  inside-address address <web-server:eth1>
  protocol tcp
  type destination
```



仮想データセンター（OneBox設定例）

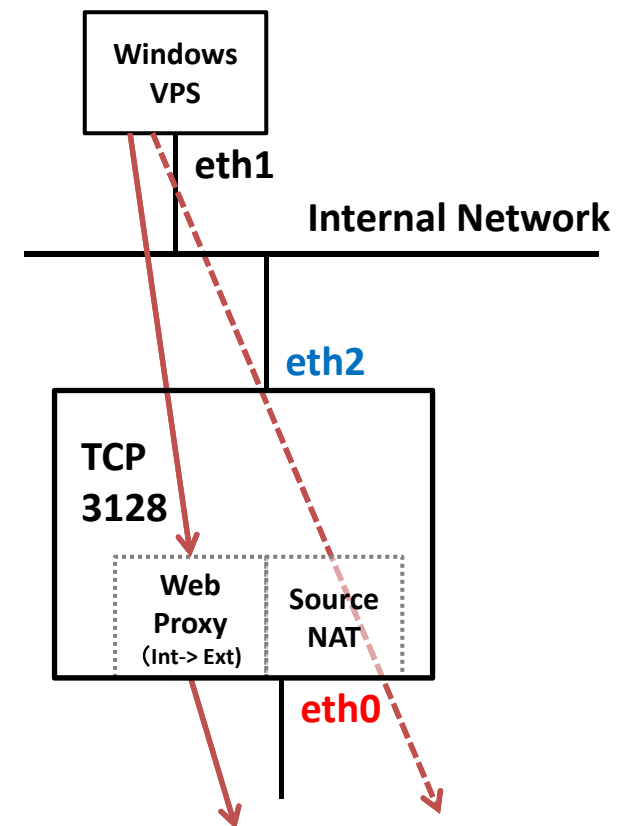
内部VPSからの外部アクセス：Web / その他

Web Proxy 設定：

```
set service webproxy listen-address <internal-addr:eth2>
```

Source NAT (masquerade) 設定：

```
set service nat rule 20  
  outband-interface <external-interface:eth0>  
  type masquerade
```

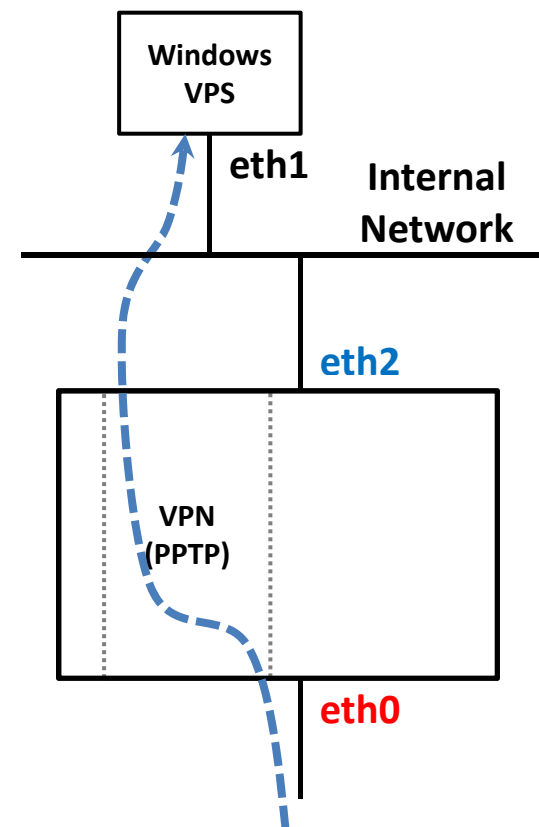


仮想データセンター (OneBox設定例)

VPN(PPTP) によるアクセス

VPN(PPTP) 設定:

```
set vpn pptp remote-access
  outside-address <external-addr:eth0>
  client-ip-pool start <local-pool-start>
  client-ip-pool stop <local-pool-stop>
  dns-servers <dns-name> <dns-ip>
  authentication mode local
  authentication local-users username
    <username> password <password>
```



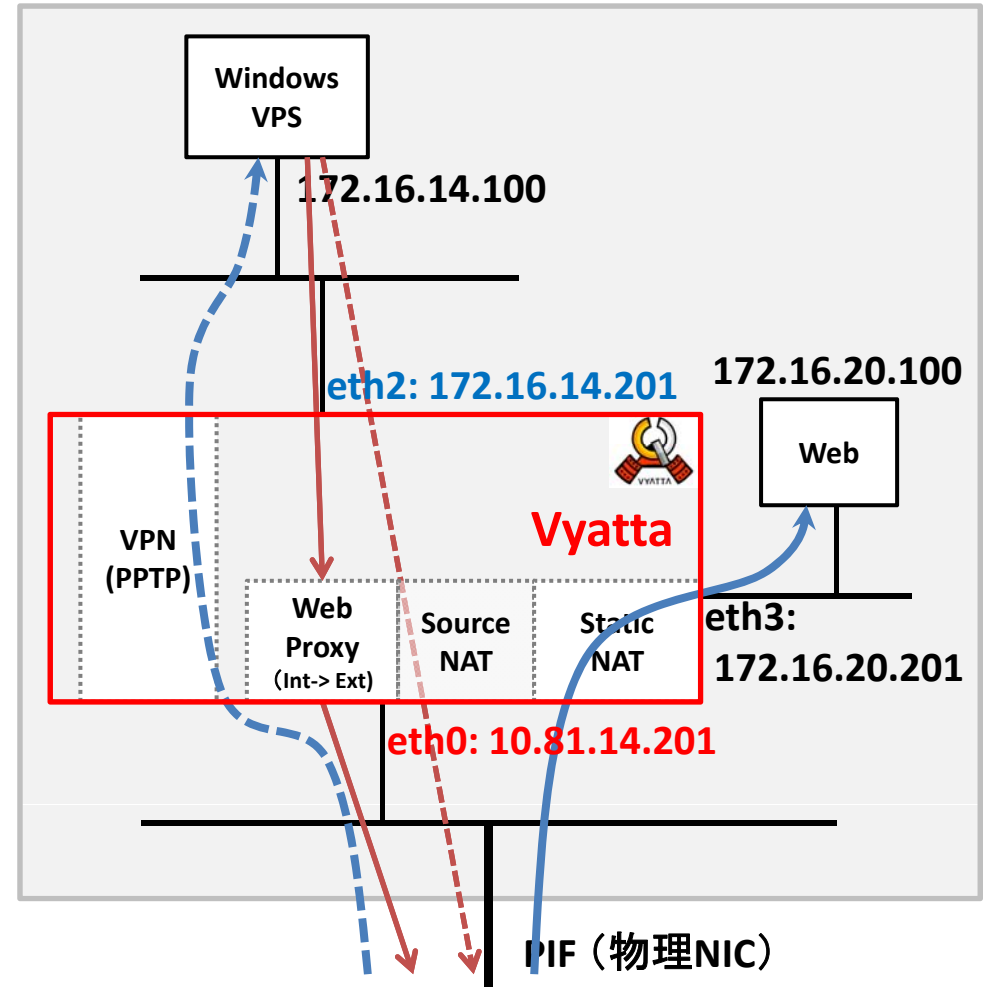
仮想データセンター（OneBox設定例）

```
set service nat rule 10
  destination address 10.81.14.201
  destination port 80
  inbound-interface eth0
  inside-address address 172.16.20.100
  protocol tcp
  type destination
```

```
set service webproxy listen-address 172.16.14.201
```

```
set service nat rule 20
  outband-interface eth0
  type masquerade
```

```
set vpn ptp remote-access
  outside-address 10.81.14.201
  client-ip-pool start 10.81.100.11
  client-ip-pool stop 10.81.100.20
  dns-servers dns01 8.8.8.8
  authentication mode local
  authentication local-users username user password pass
```



日本語の入門書あります！



- 第1章 Vyattaの全体像
- 第2章 Vyattaクイックスタートガイド
- 第3章 Vyattaの初歩的な設定
- 第4章 企業内ネットワークで使うための機能
- 第5章 ネットワークインターフェイス
- 第6章 経路制御(ルーティング)
- 第7章 ネットワークセキュリティ機能
- 第8章 VPN (Virtual Private Network) 機能
- 第9章 Vyattaによる高可用性の実現
- 第10章 QoSの使用方法

2011年6月16日発売

近藤邦昭, 松本直人, 浅間正和, 大久保修一,
(日本Vyattaユーザー会) 著

B5変形判 / 288ページ

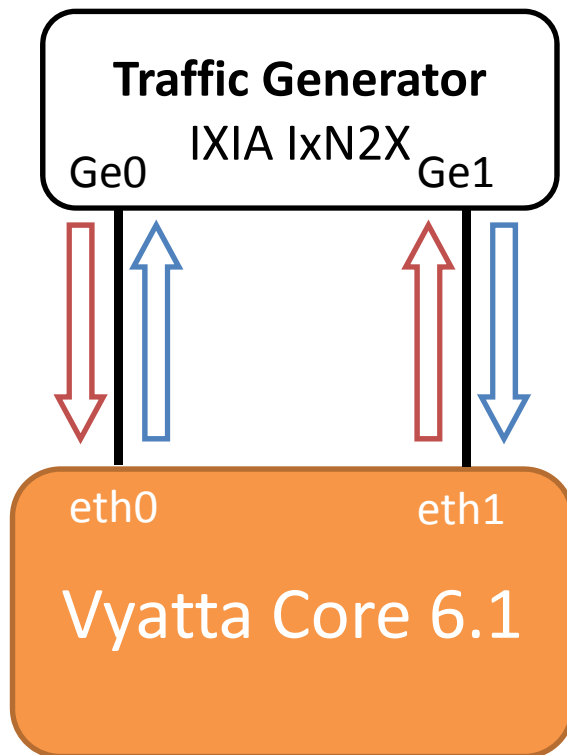
定価3,360円(本体3,200円)

ISBN 978-4-7741-4711-6

<http://gihyo.jp/book/2011/978-4-7741-4711-6>

VYATTA転送性能(仮想化無し)

テスト構成



Vyatta Core の Hardware 構成

Model	HP DL160 G6 (～15万円)
CPU	Xeon E5620 2.40GHz (Quad Core)
Memory	DDR3 SDRAM 1333MHz 6GB
NIC	オンボード : Broadcom BCM5715 追加 : Intel 82576EB Dual Port

宛先アドレス(IPv4)

Ge0→Ge1: 11.0.0.0～11.0.0.255

Ge1→Ge0: 12.0.0.0～12.0.0.255

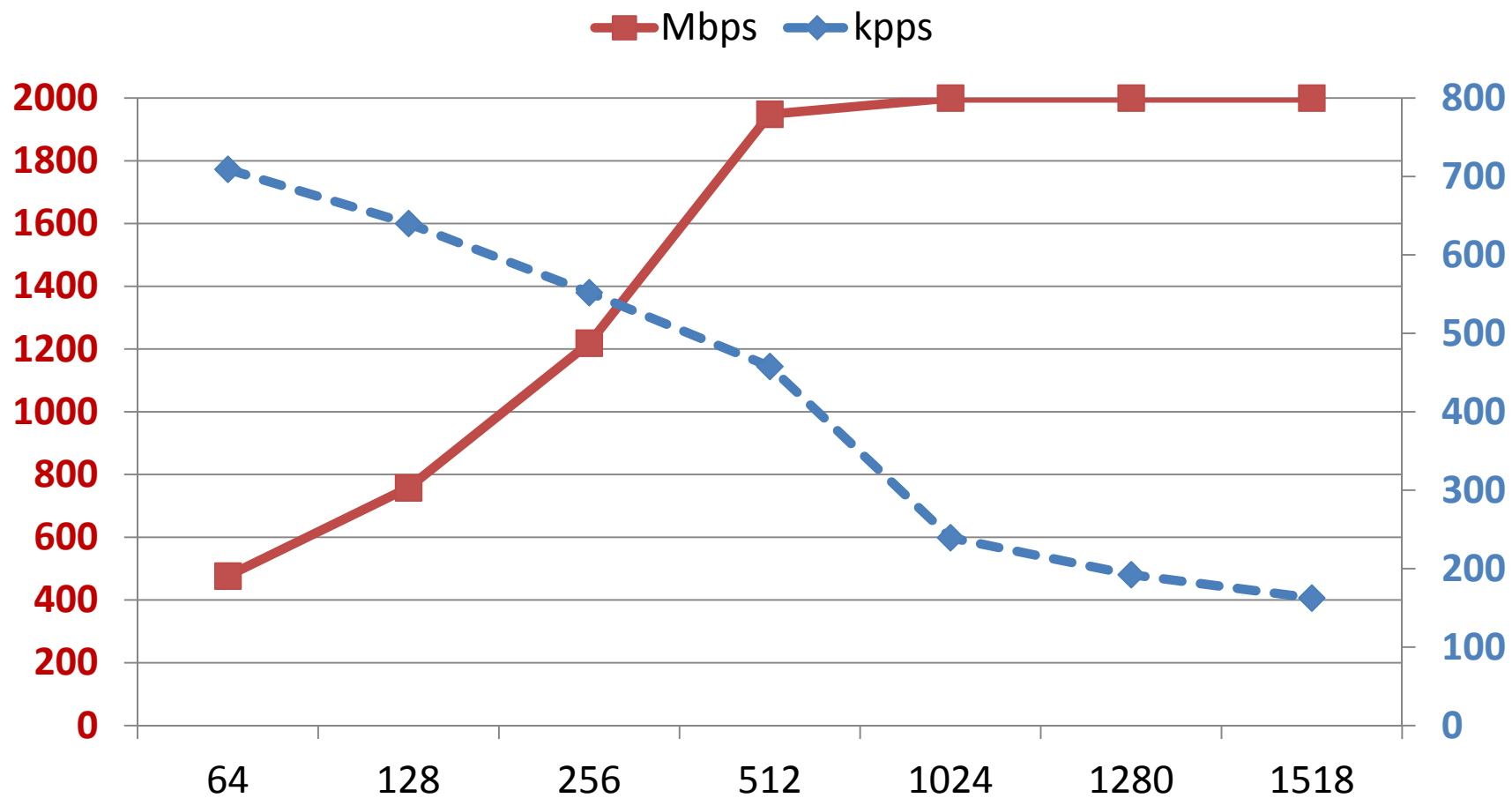
宛先アドレス(IPv6)

Ge0→Ge1: 2400::～2400::ff

Ge1→Ge0: 2400:1::～2400:1::ff

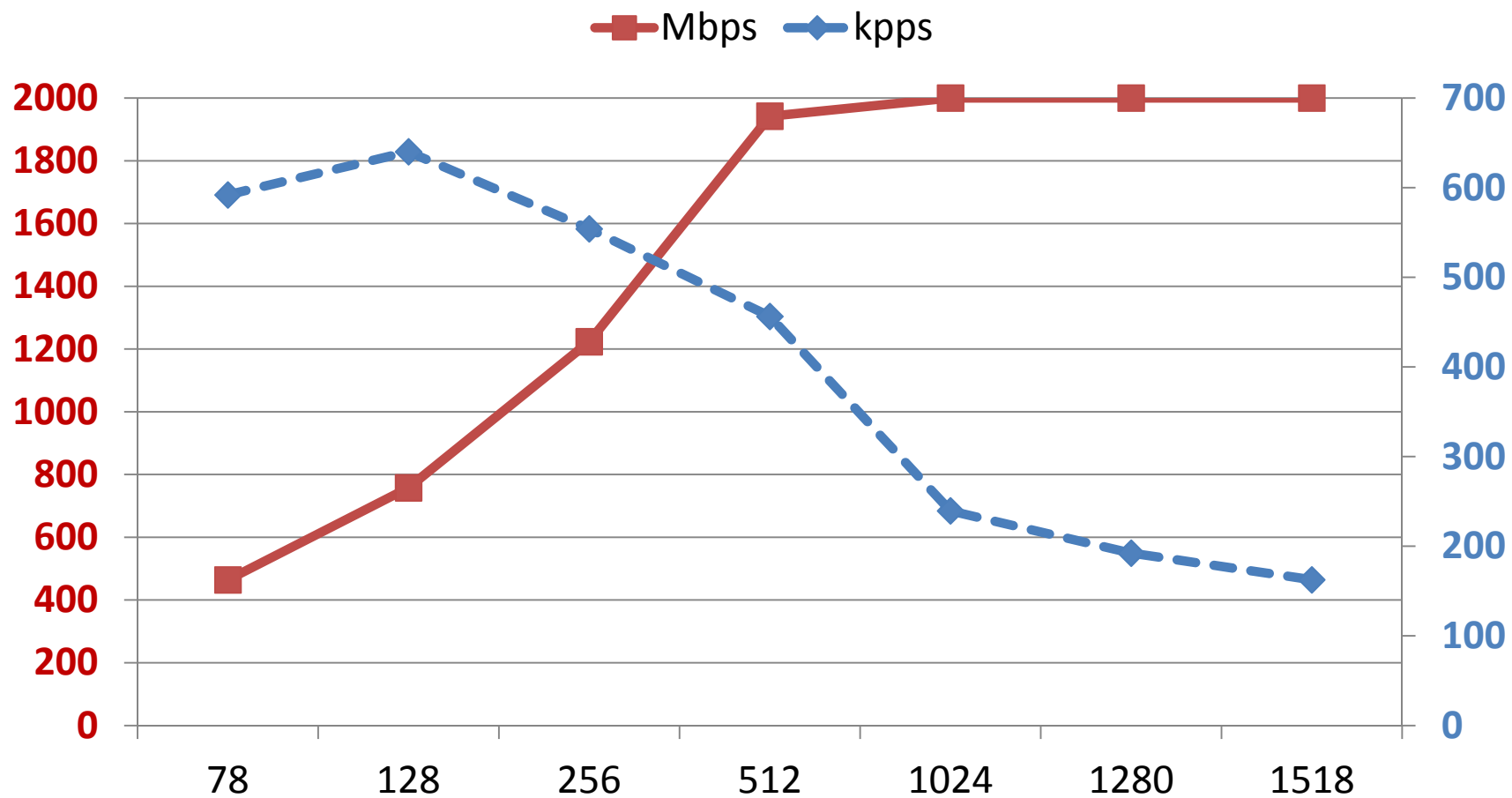
パケットサイズと転送性能(IPv4)

- オンボード : Broadcom BCM5715 -



パケットサイズと転送性能(IPv6)

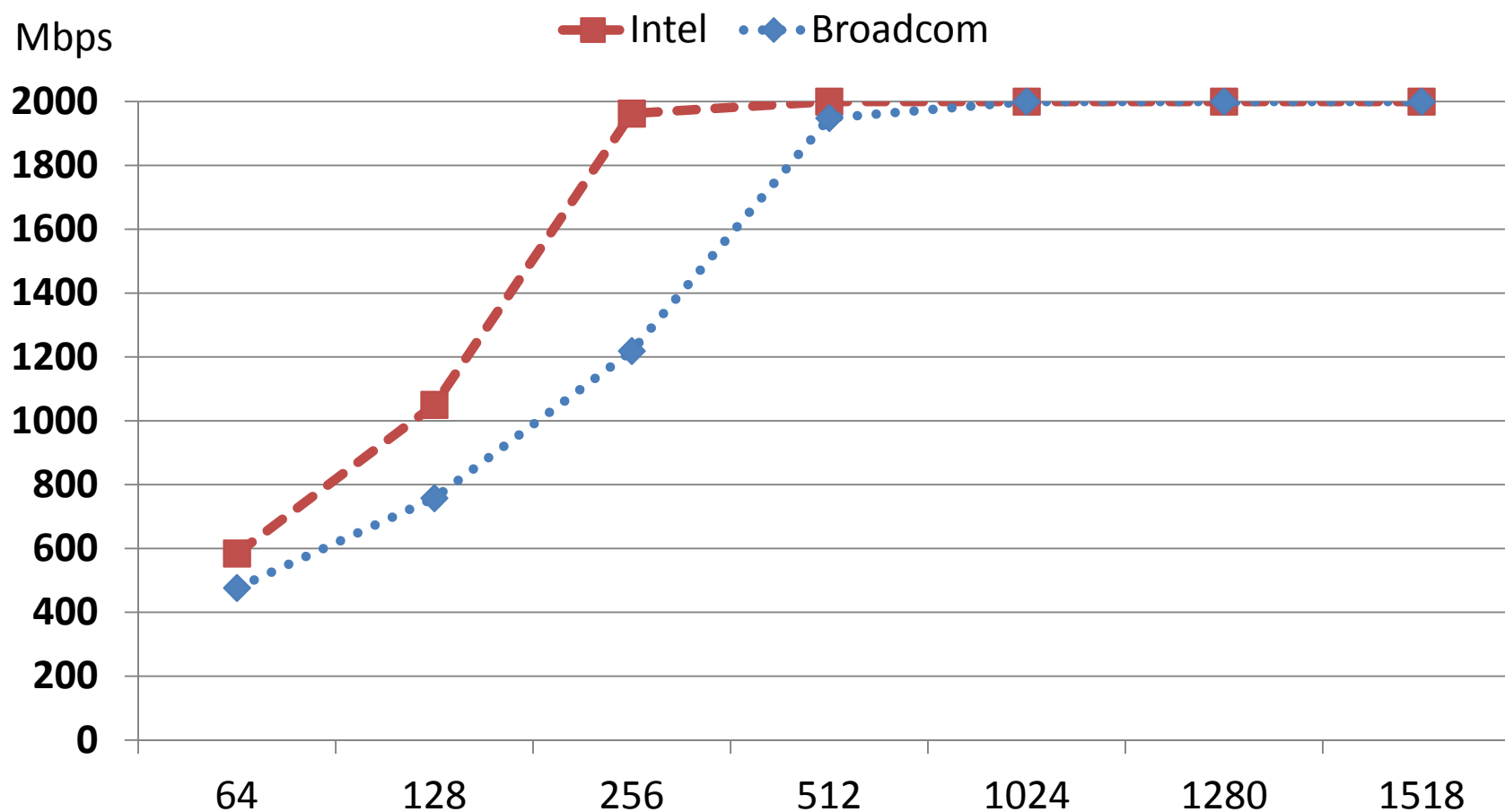
- オンボード : Broadcom BCM5715 -



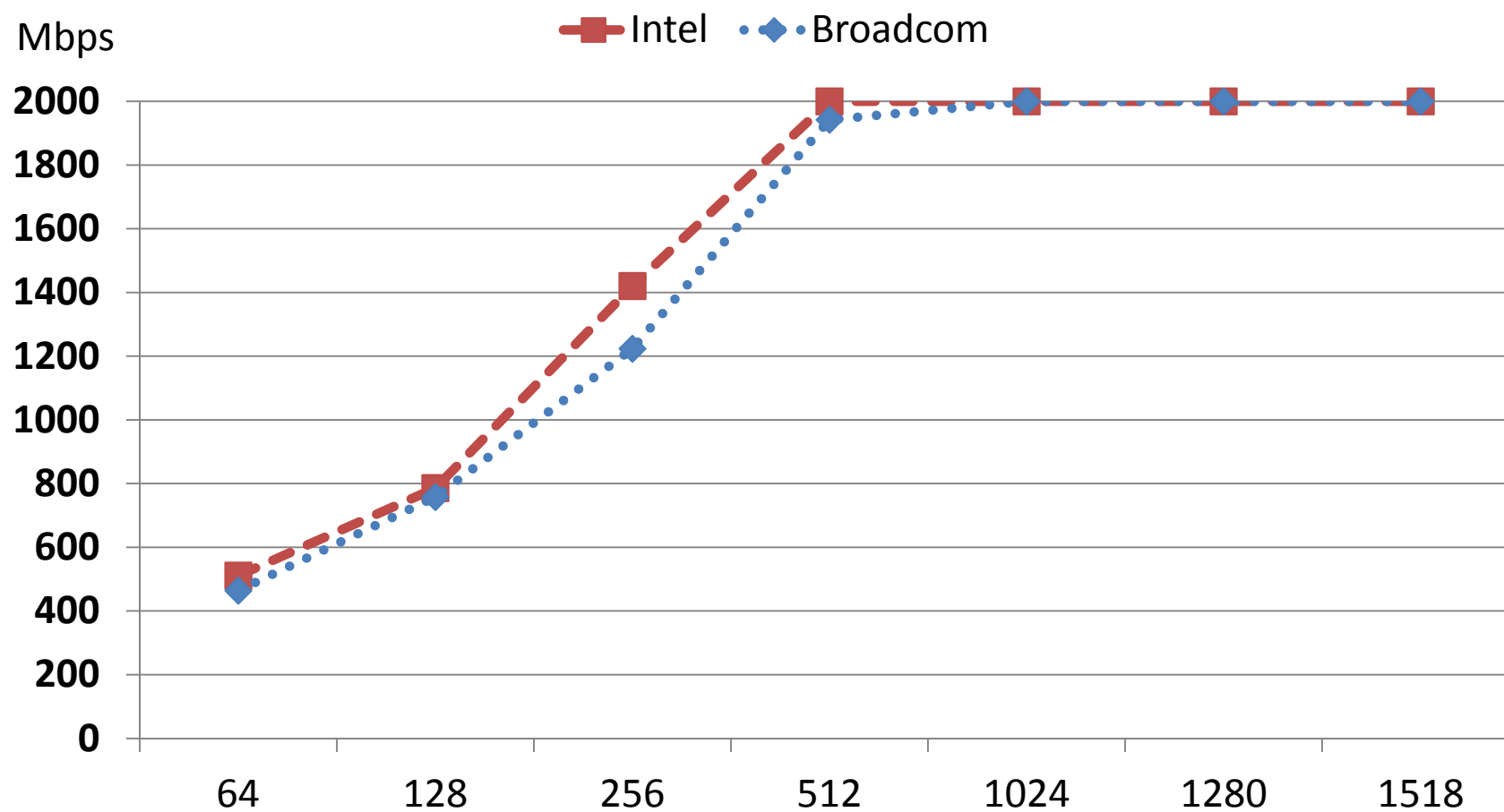
NICによるパフォーマンスへの影響

- オンボード Broadcom BCM571 からNICを変更
- Intel : PRO/1000 PT Dual Port Server Adapter
 - 型番 : EXPI9402PT (Intel 82571GB Gigabit Controller)
 - 価格 : ~18,000円
- Intel (MQ) : Gigabit ET Dual-Port Server Adapter
 - 型番 : E1G42ET (Intel 82576 Gigabit Controller)
 - 価格 : ~22,000円

NICによるパフォーマンス (IPv4)

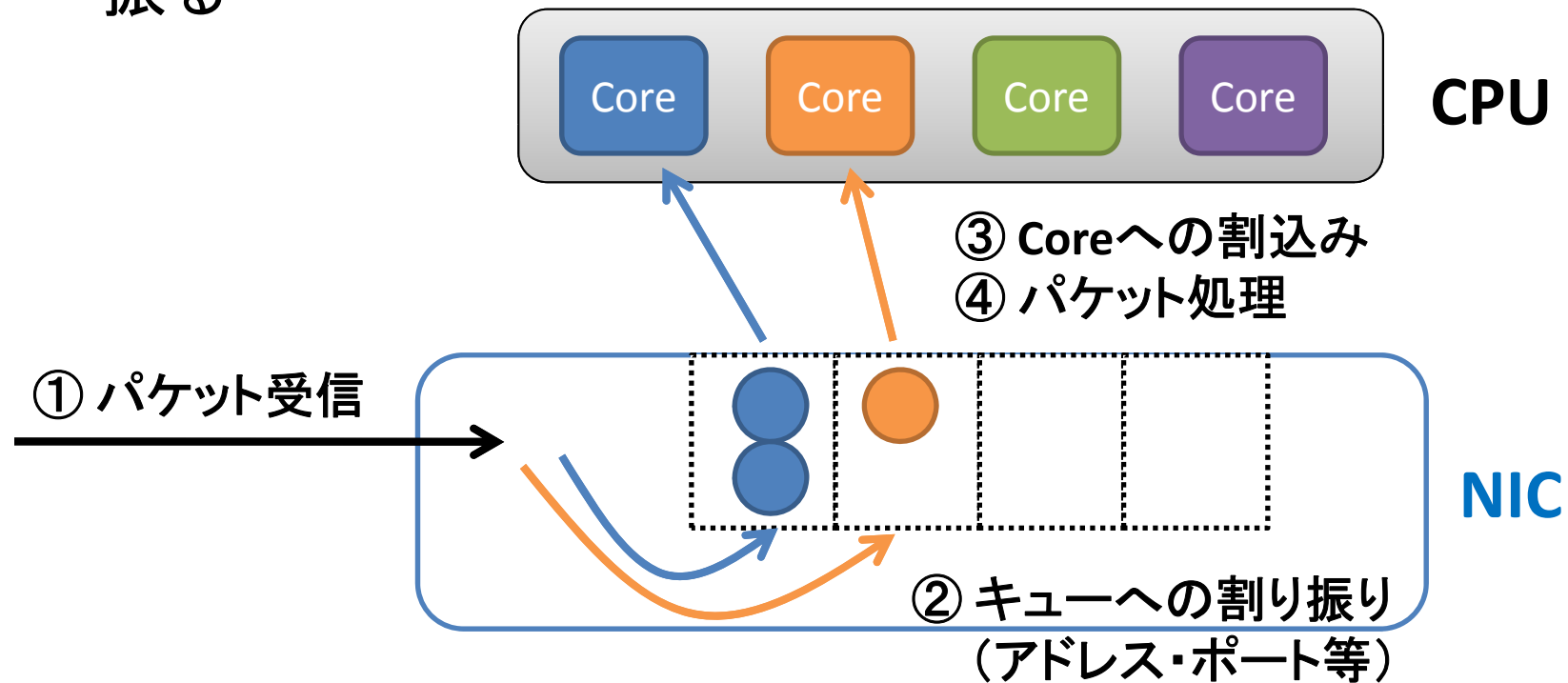


NICによるパフォーマンス (IPv6)



マルチキューによる性能向上

- Receive-Side Scaling, Scalable I/O 等と呼ばれる
- CPU Core数増加により、パケット処理性能の向上
- 複数のキューを持ち、パケットをCPUの各Coreに割り振る



• NICのスペックシートで確認

Features and benefits

Intel® 82576 Gigabit Ethernet Controller

- Industry-leading, energy-efficient design for next-generation Gigabit performance and multi-core processors

Low-profile

- Enables higher bandwidth and throughput from standard and low-profile PCI Express* slots and servers

iSCSI remote boot support

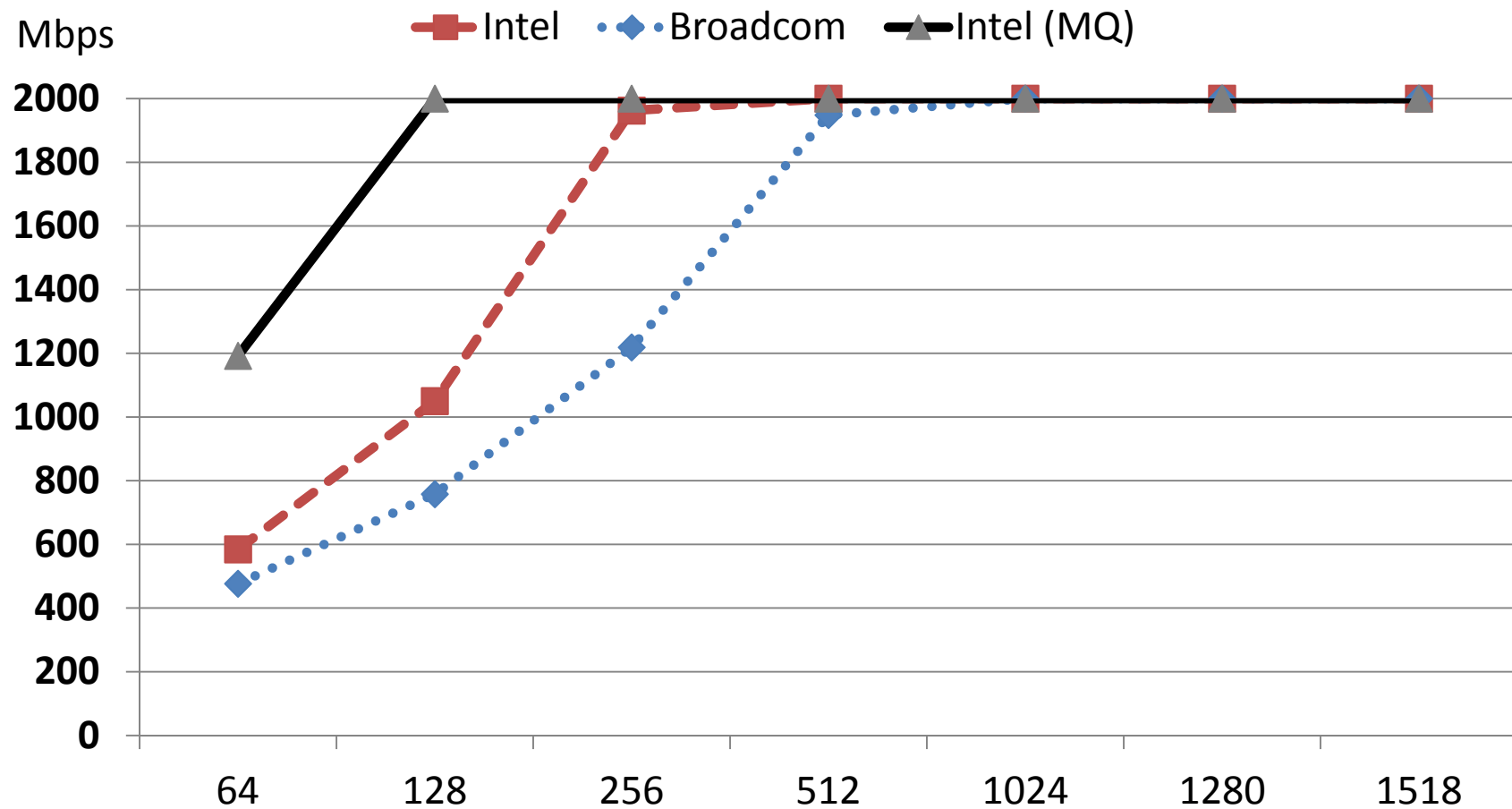
- Provides centralized storage area network (SAN) management at a lower cost than competing iSCSI solutions

Load balancing on multiple CPUs

- Increases performance on multi-processor systems by efficiently balancing network loads across CPU cores when used with Receive-Side Scaling from Microsoft or Scalable I/O on Linux*

参照：<http://www.intel.com/Products/Server/Adapters/Gb-ET-Dual-Port/Gb-ET-Dual-Port-overview.htm>

マルチキューによる性能向上 (IPv4)



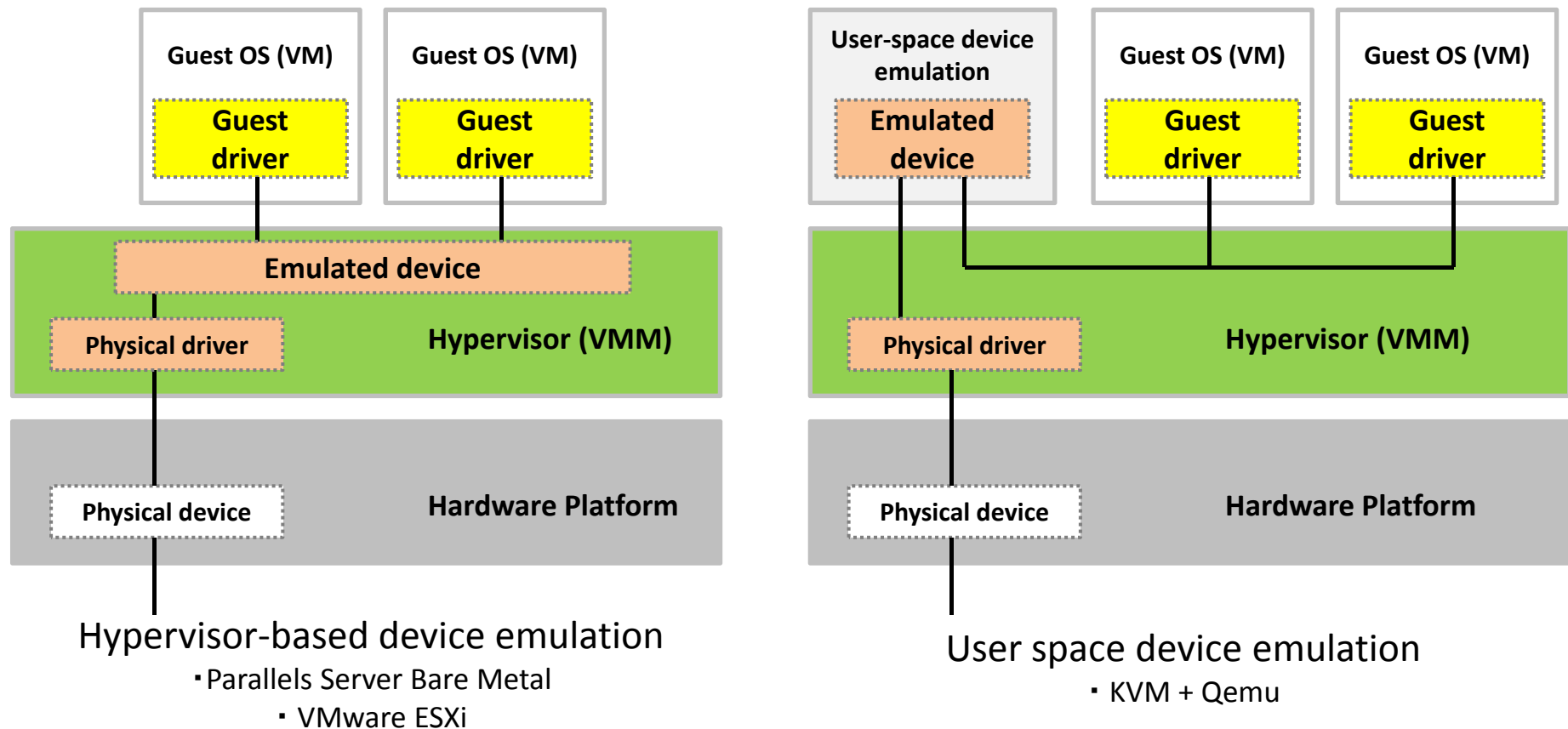
仮想化環境におけるパケット転送 のしくみ

仮想化環境におけるパケット転送のしくみ

- I/O仮想化の種類
 - デバイス・エミュレーション
 - Hypervisor (VMM) based device emulation
 - User space device emulation
 - I/O準仮想化
 - New Software Interface
 - I/O デバイス割り当て (VT-d)
 - I/O デバイス割り当て + 共有 (SR-IOV)

デバイス・エミュレーション

デバイスを Hypervisor (VMM) がエミュレーション
オーバーヘッド=大



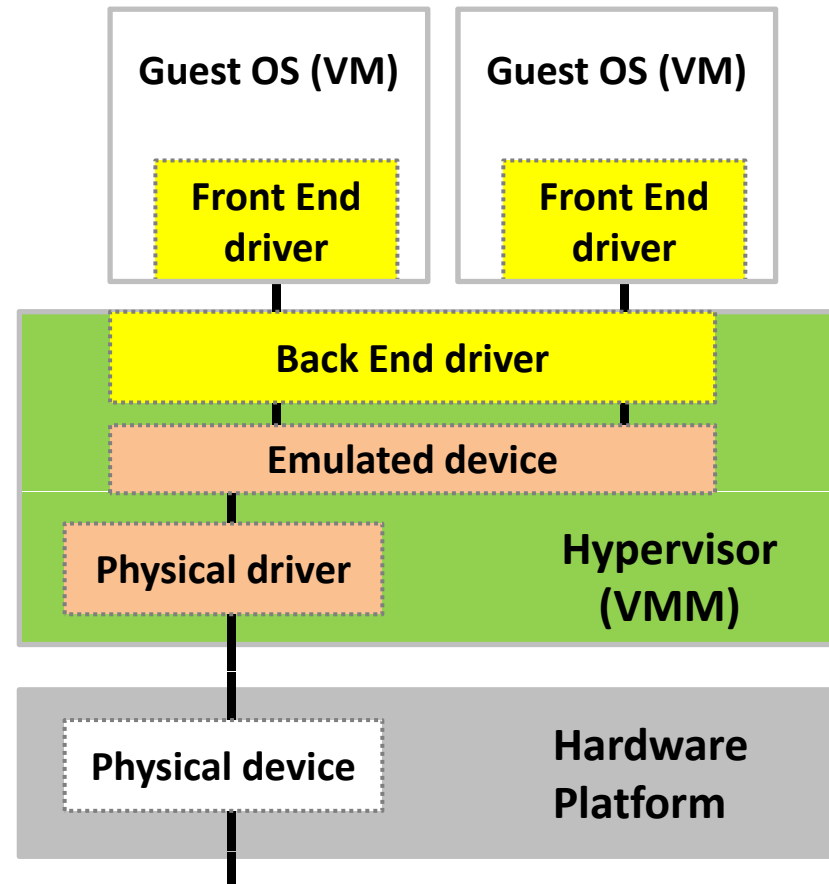
I/O準仮想化

Guest/VMMが連携して動作(専用のドライバ)
オーバーヘッド=小

KVM: virtio-net

Vmware: vmxnet3

Xen: netfront

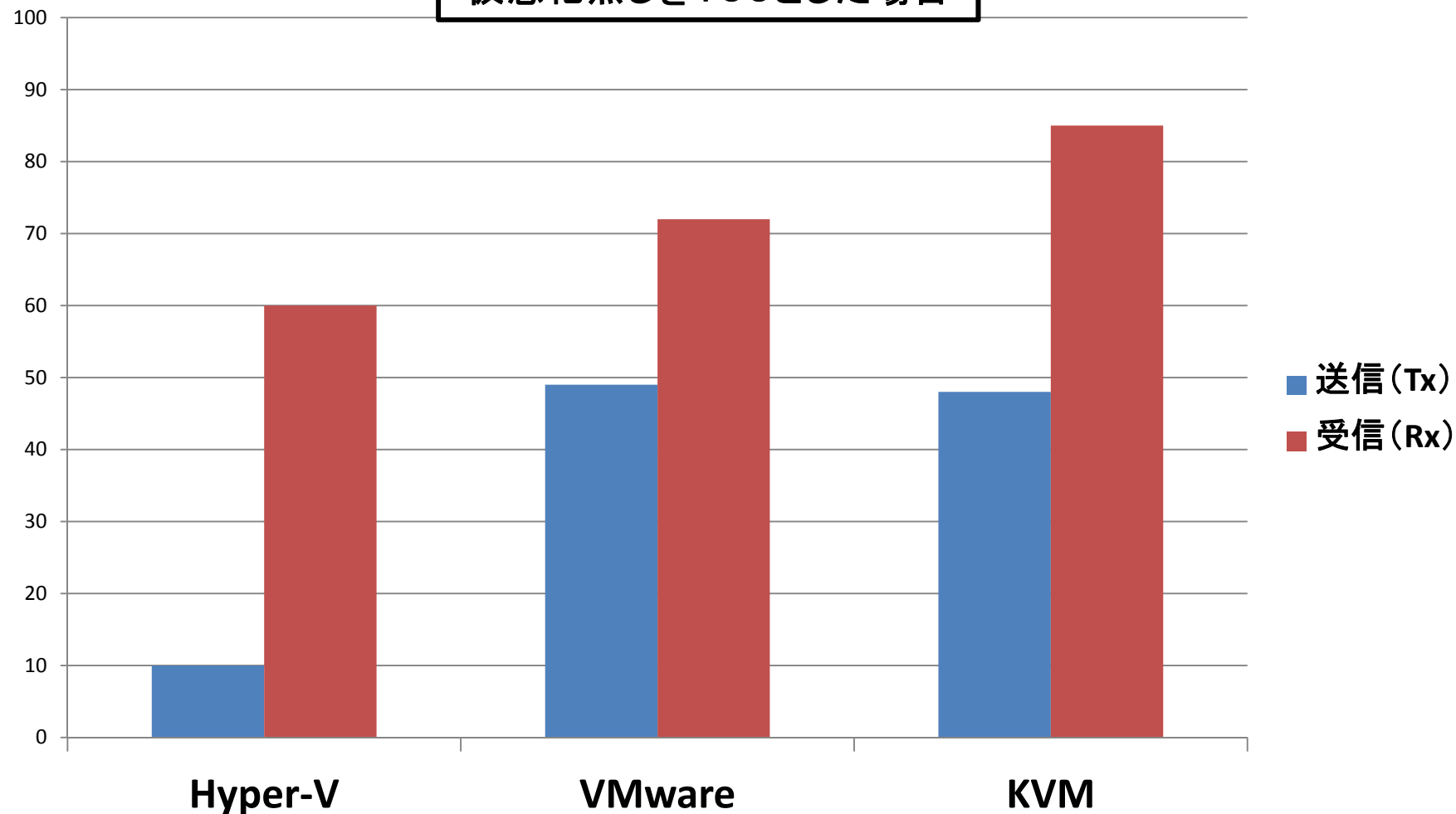


-
- Stephen Hemminger
 - Vyatta Inc. プリンシパルエンジニア (米国)
 - Linux Kernel Contributor (network)
 - “netem” network emulator
 - bridging, iproute maintaner
 - @Linuxcon Japan
 - 2011年6月2日



デバイス・エミュレーション 性能比較

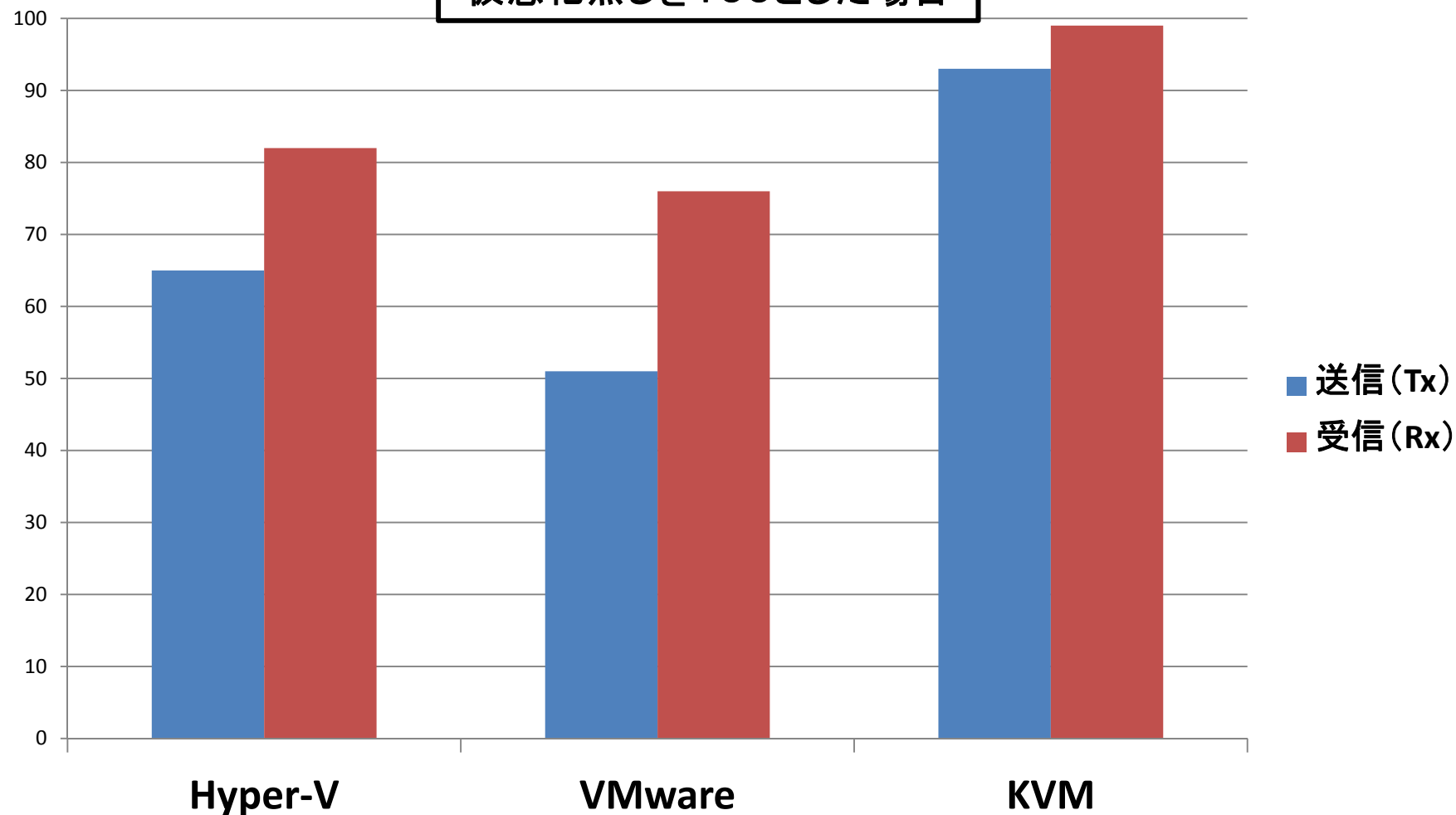
仮想化無しを100とした場合



Stephen Hemminger @ Linuxcon Japan 2011/06/02

I/O準仮想化 性能比較

仮想化無しを100とした場合



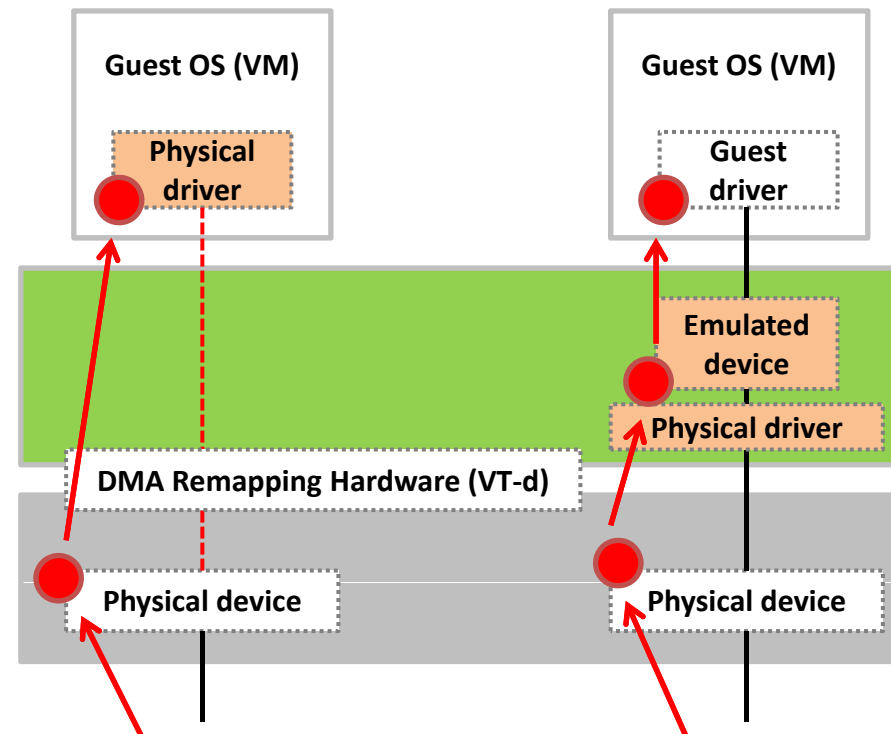
Stephen Hemminger @ Linuxcon Japan 2011/06/02

I/O デバイス割り当て (VT-d)

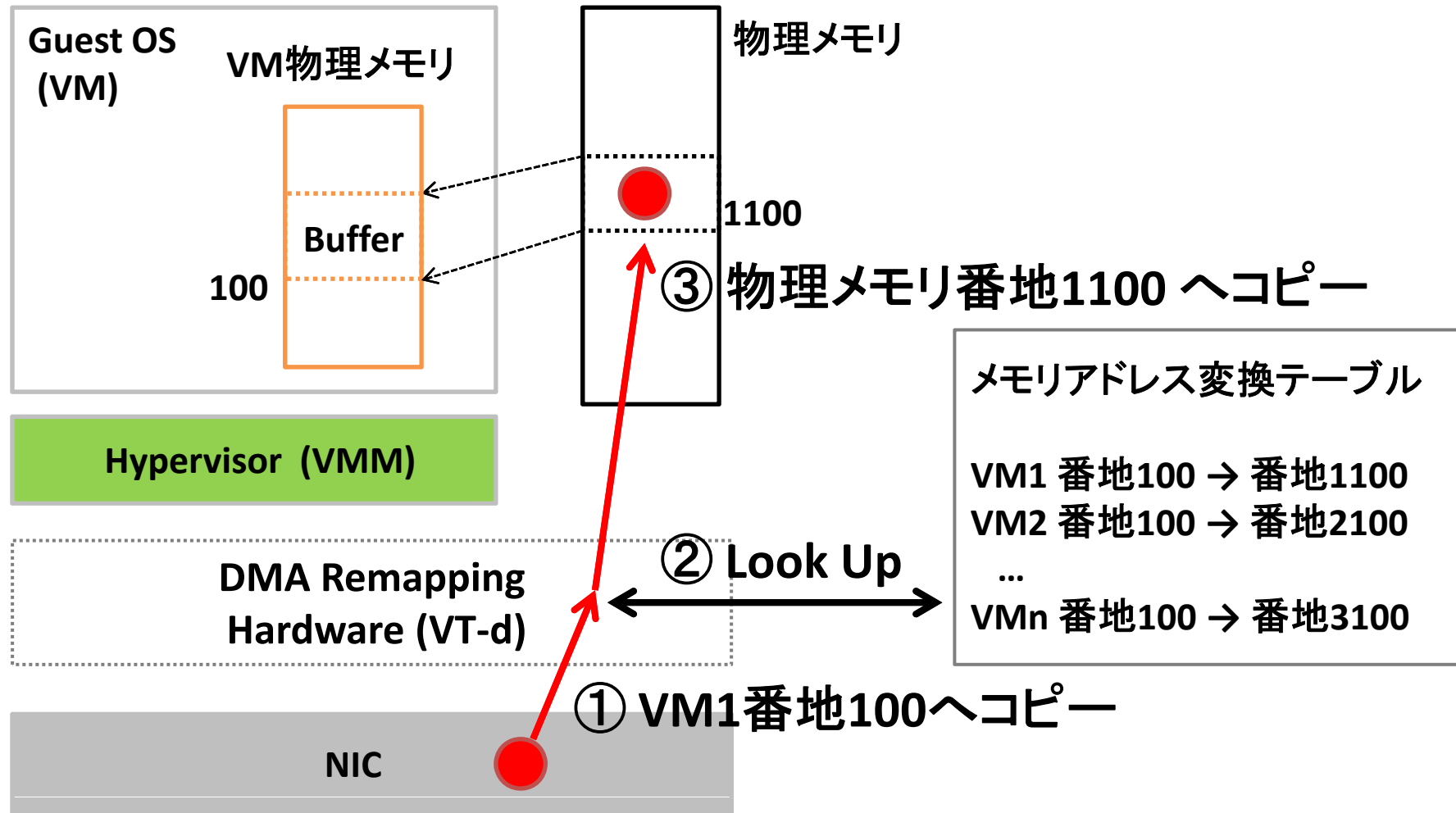
物理デバイス (NIC) を VM に割り当て
パケットを DMA 転送
オーバーヘッド = 極小

1ポート(※)に1VMのみ

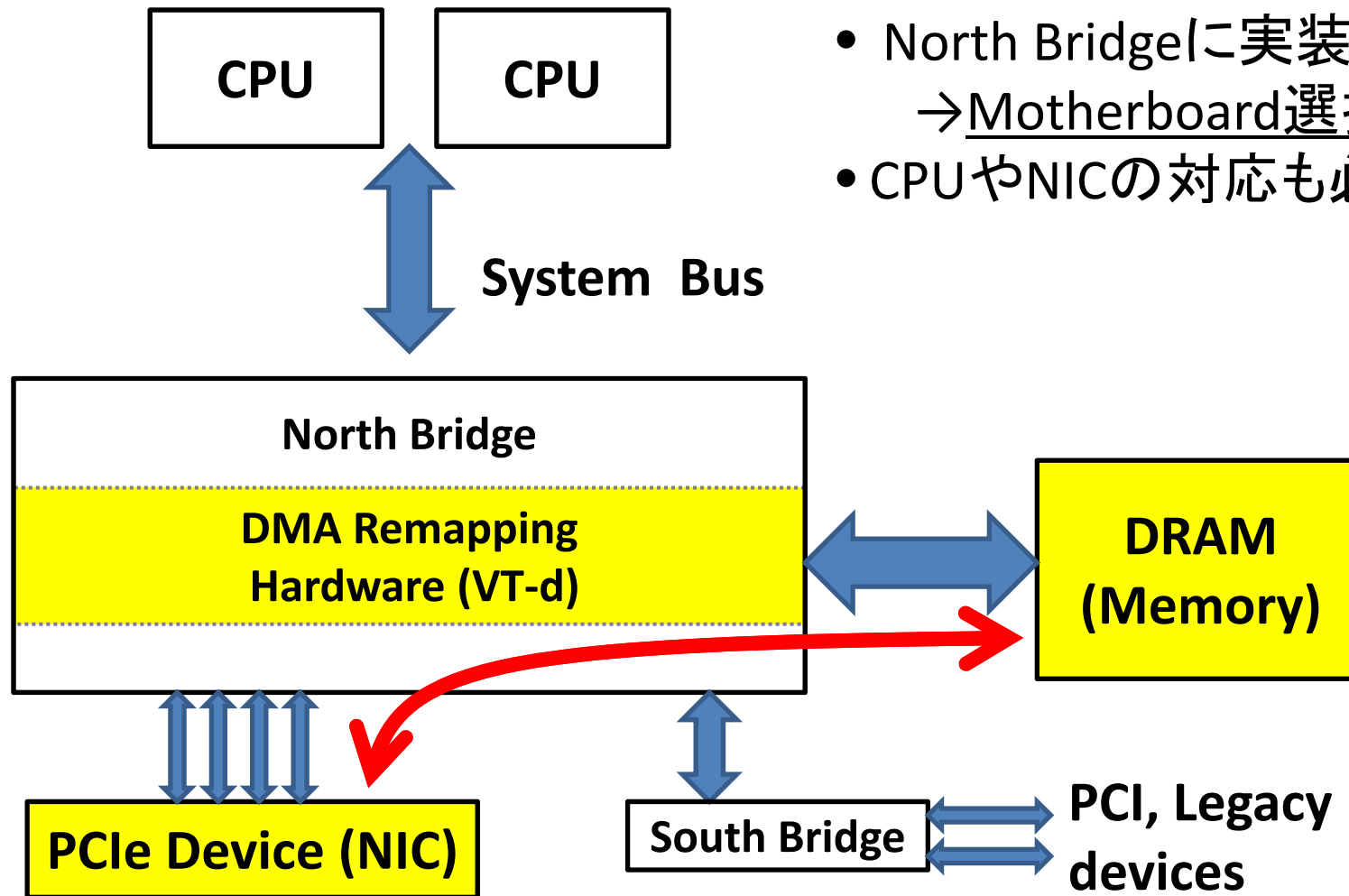
(※) 1 PCIe Function



VT-d : メモリアドレス (番地) 変換



VT-d: 物理構成



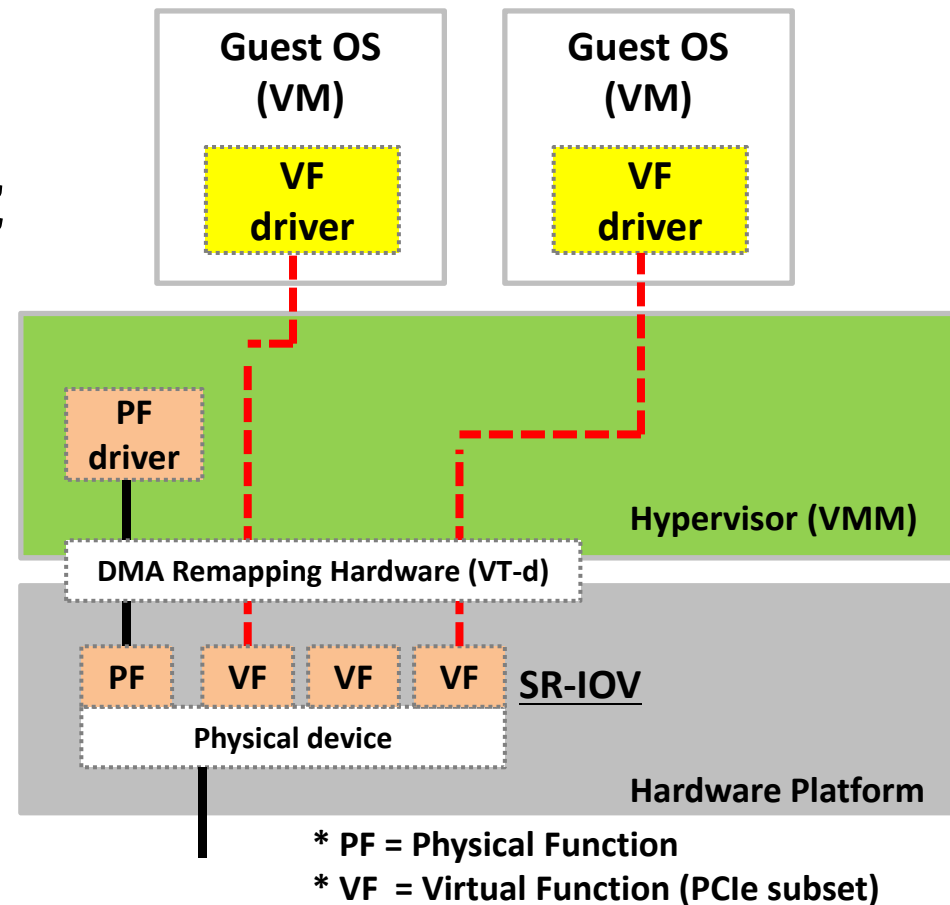
- North Bridgeに実装されている
→ Motherboard選択に注意
- CPUやNICの対応も必要

I/O デバイス割り当て＋共有 (SR-IOV)

物理デバイス (NIC) を VM に割り当て＋共有
オーバーヘッド＝極小

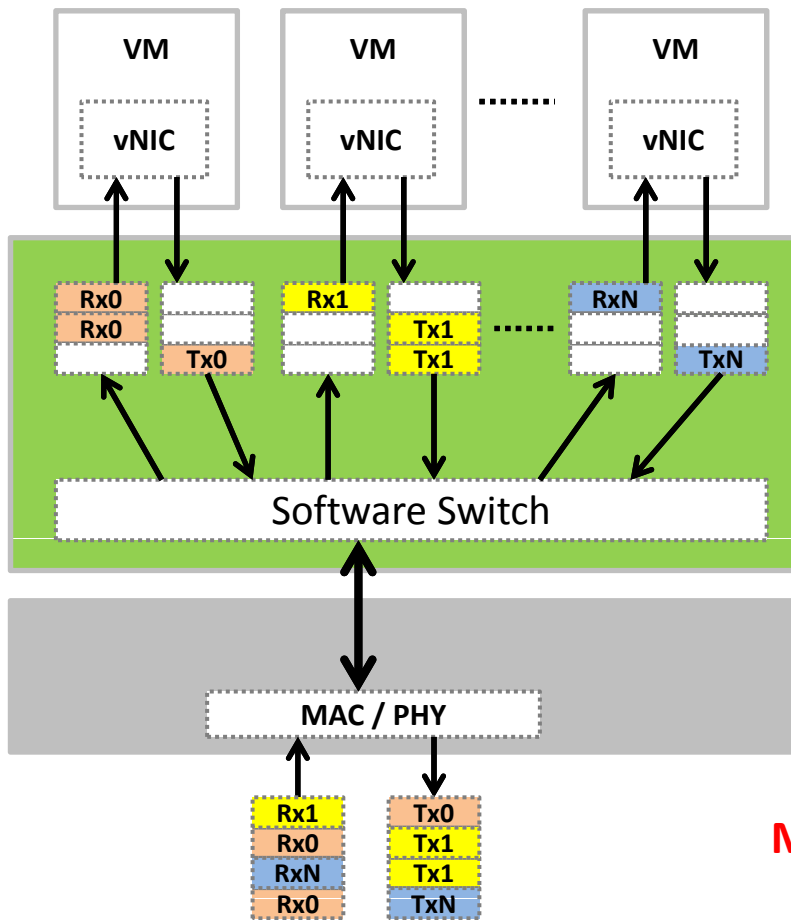
複数 VM でポート共有可能

VM からは VF = NIC (Port) に見える

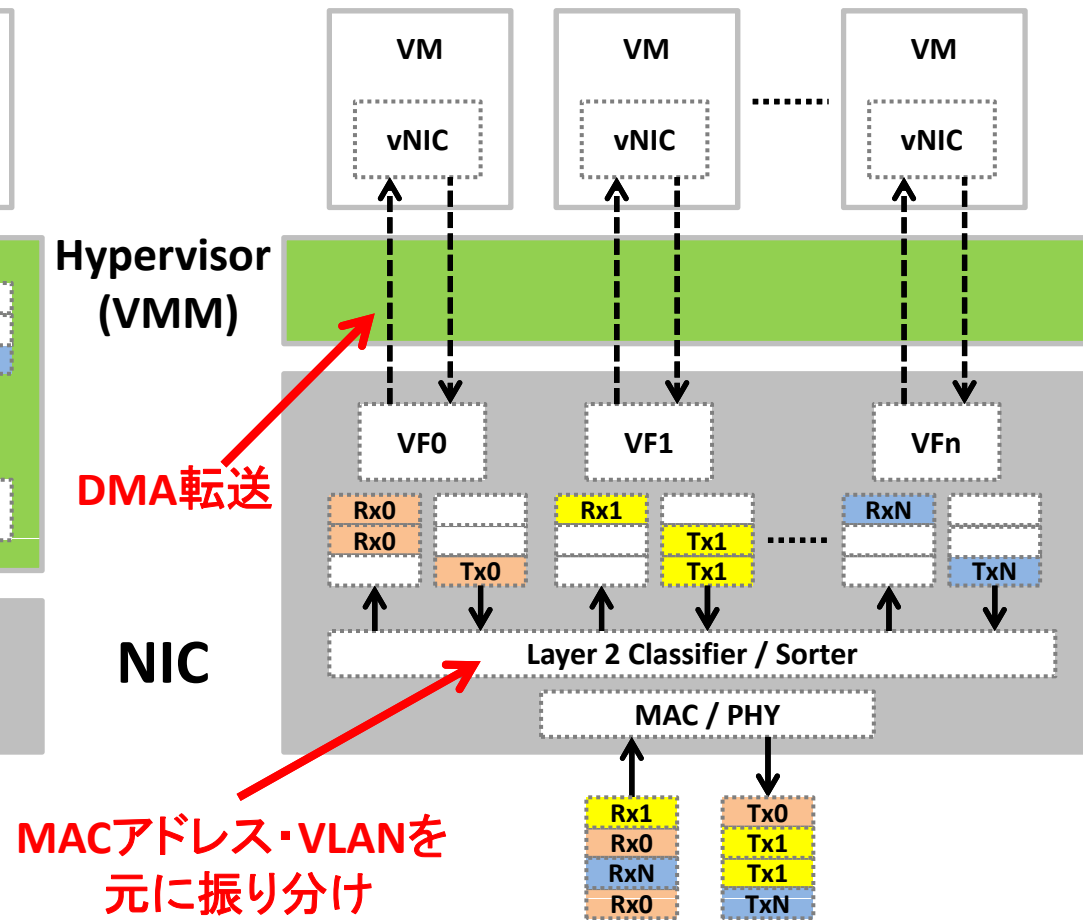


SR-IOV

SR-IOV 無し: VMMがパケット振り分け

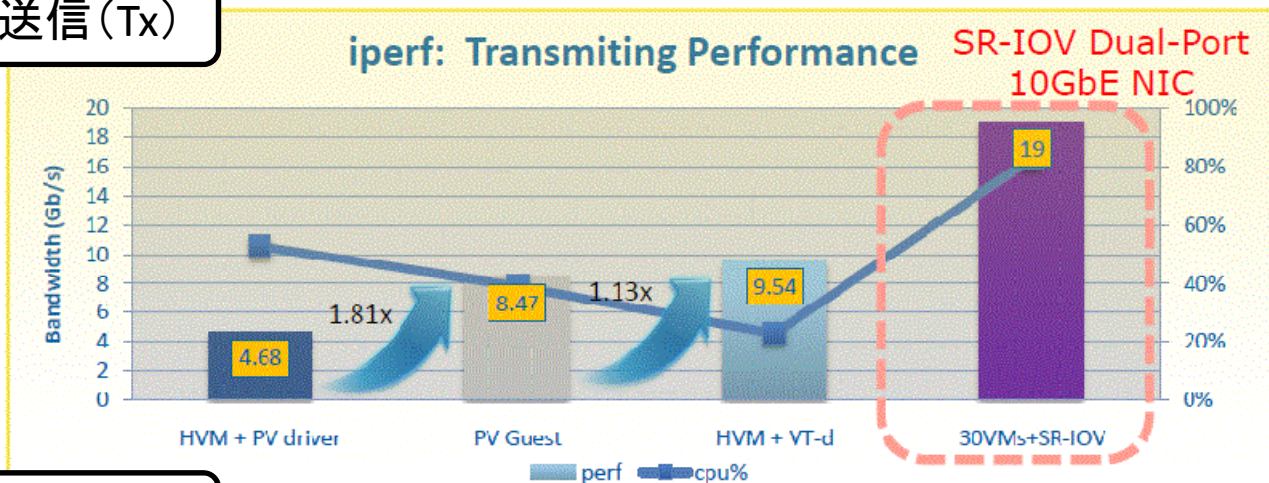


SR-IOV 有り: NICがパケット振り分け
VMMはパス・スルー



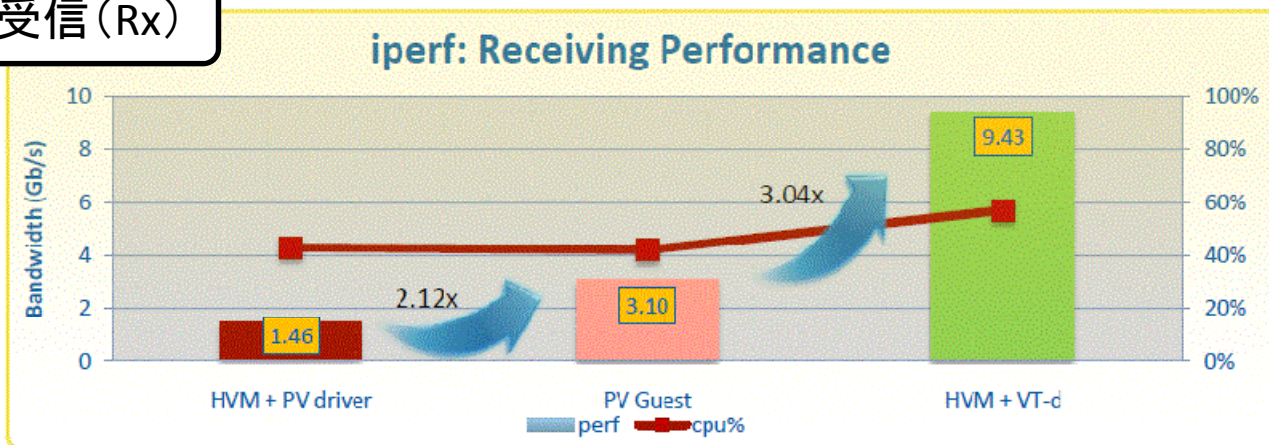
準仮想化、デバイス割り当て、SR-IOV性能比較

送信 (Tx)



VT-d はシングルキュー

受信 (Rx)



出典: http://www.xen.org/files/xensummit_intel09/xensummit2009_IOVirtPerf.pdf
IO Virtualization Performance, HUANG Zhiteng (zhiteng.huang@intel.com)

HVM = Xen Hardware Virtual Machine
PV = Para Virtualization (準仮想化)

仮想化技術をサポートする ハードウェア

Intel CPU VT-x, VT-d サポート確認方法

<http://ark.intel.com/>

The screenshot shows the Intel ARK website interface. The navigation path is highlighted with red boxes and arrows:

- 1) Click !! (Arrow pointing to "2nd Generation Intel® Core™ i7 Processors" in the Desktop Processors list)
- 2) Click !! (Arrow pointing to "Intel® Core™ i7-2600S Processor (8M Cache, 2.80 GHz)" in the processor table)

The processor table below shows the following data:

Compare Select: All None	Product Name	Status	Embedded Options Available	Max TDP	Recommended Channel Price	Graphic
Select	Intel® Core™ i7-2600S Processor (8M Cache, 2.80 GHz)	Launched	No	65 W	\$294.00	Intel® Graphics
Select	Intel® Core™ i7-2600K Processor (8M Cache, 3.40 GHz)	Launched	No	95 W	\$317.00	Intel® Graphics
Select	Intel® Core™ i7-2600 Processor (8M Cache, 3.40 GHz)	Launched	Yes	95 W	\$294.00	Intel® Graphics

Intel CPU VT-x, VT-d サポート確認方法

Advanced Technologies		
Intel® Turbo Boost Technology		2.0
Intel® vPro Technology		Yes
Intel® Hyper-Threading Technology		Yes
Intel® Virtualization Technology (VT-x)		Yes
Intel® Virtualization Technology for Directed I/O (VT-d)		Yes
Intel® Trusted Execution Technology		Yes
AES New Instructions		Yes
Intel® 64		Yes
Idle States		Yes
Enhanced Intel SpeedStep® Technology		Yes
Thermal Monitoring Technologies		Yes
Intel® Fast Memory Access		Yes
Intel® Flex Memory Access		Yes
Execute Disable Bit		Yes

VT-x VT-d

Click !!

Intel CPU VT-x, VT-d サポート確認方法

<http://ark.intel.com/search/advanced/?s=t&VTX=true&VTD=true>

Advanced Search Export

Filters Applied: Intel® Virtualization Technology (VT-x) | true Intel® Virtualization Technology for Directed I/O (VT-d) | true Compare Now (0)

115 Matching Products

MODIFY FILTERS

Clear Filters Search

Essentials

Family

Select...

Processor Number

Code Name

Product Name	Launch Date	Processor Number	# of Cores	# of Threads	Clock Speed	Max Turbo Frequency	Cache	System Bus	Instruction Set	Emulated Options Available	Max TDP	Recommended Channel Price	VT-x	VT-d
Intel® Core™ i7-2640M Processor (4M Cache, 2.80 GHz)	Q4'11	i7-2640M	2	4	2.8 GHz	3.5 GHz	4 MB	5 GT/s	64-bit	No	35 W	\$316.00	Yes	Yes
Intel® Core™ i7-2760QM Processor (6M Cache, 2.40 GHz)	Q4'11	i7-2760QM	4	8	2.4 GHz	3.5 GHz	6 MB	5 GT/s	64-bit	No	45 W	\$378.00	Yes	Yes
Intel® Core™ i7-2860QM Processor (8M Cache, 2.50 GHz)	Q4'11	i7-2860QM	4	8	2.5 GHz	3.6 GHz	8 MB	5 GT/s	64-bit	No	45 W	\$568.00	Yes	Yes
Intel® Core™ i7-2960XM														

Intel® Virtualization Technology (VT-x) Options Available

Intel® Virtualization Technology for Directed I/O (VT-d)

Product Name	Launch Date	Processor Number
Intel® Core™ i7-2640M Processor (4M Cache, 2.80 GHz)	Q4'11	i7-2640M
Intel® Core™ i7-2760QM Processor (6M Cache, 2.40 GHz)	Q4'11	i7-2760QM
Intel® Core™ i7-2860QM Processor (8M Cache, 2.50 GHz)	Q4'11	i7-2860QM
Intel® Core™ i7-2960XM		

VT-x	VT-d
Yes	Yes
Yes	Yes
Yes	Yes
Yes	Yes

SR-IOVサポートする Intel NIC一覧

<http://www.intel.com/support/network/adapter/pro100/sb/CS-031492.htm>

Intel® Server Adapters

FAQs: Using SR-IOV with Intel® Ethernet Server Adapters

Which Intel® Ethernet Adapters support SR-IOV?

- Intel® Ethernet Server Adapter X520-DA2
- Intel® Ethernet Server Adapter X520-SR1
- Intel® Ethernet Server Adapter X520-SR2
- Intel® Ethernet Server Adapter X520-LR1
- Intel® Ethernet Server Adapter X520-T2
- Intel® Gigabit ET Dual Port Server Adapter
- Intel® Gigabit EF Dual Port Server Adapter
- Intel® Gigabit ET2 Quad Port Server Adapter

Which hypervisors support SR-IOV on Intel® Ethernet Adapters?

Xen Hypervisor*

KVM* (Kernel Based Virtual Machine)

Which operating systems and distributions include virtual function driver support for guests?

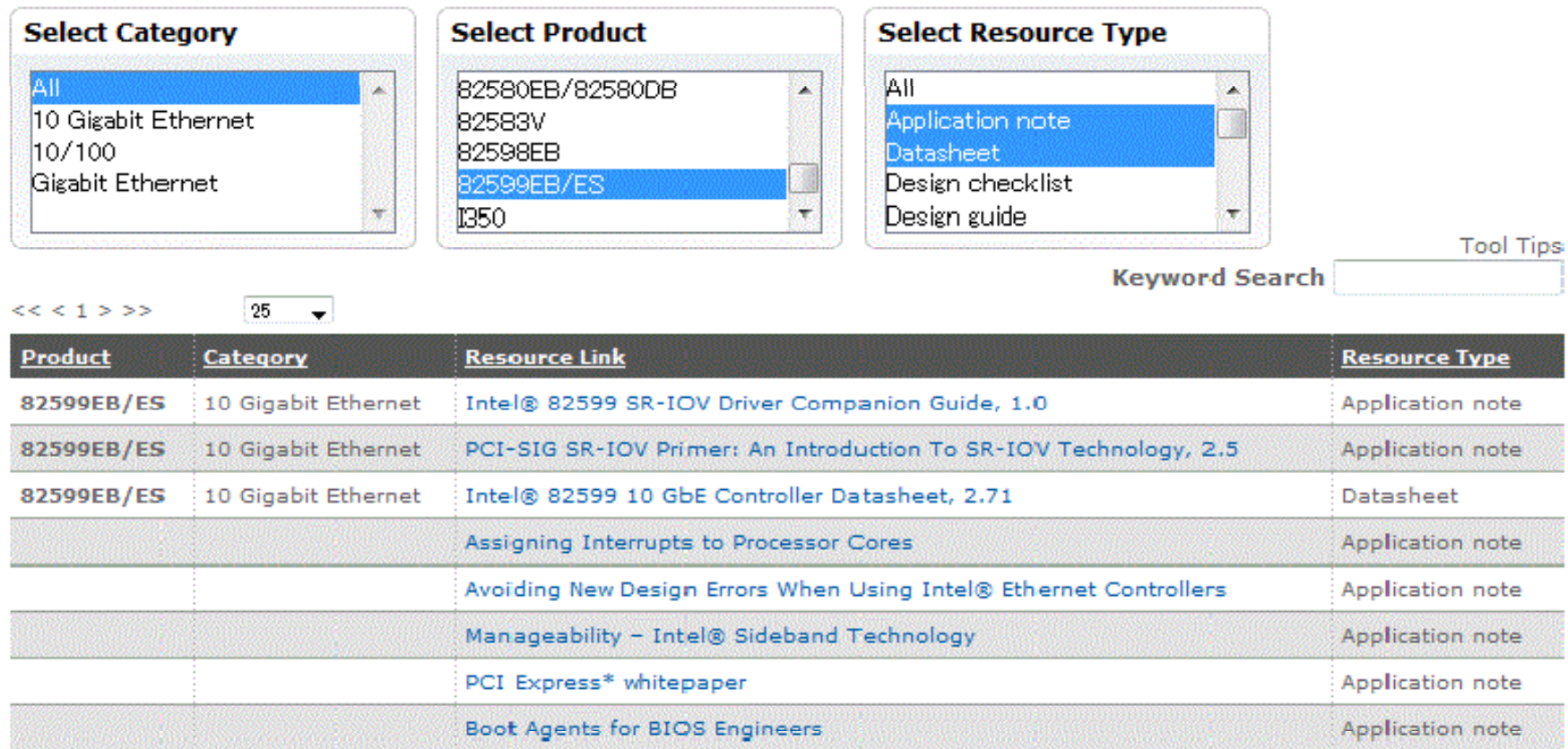
Virtual Function (VF) drivers for guest operating systems are available for:

- Windows Server 2008 R2*
- Windows Server 2008*, 32-bit and 64-bit
- Linux* 2.6.30 kernel or later
- Red Hat Enterprise Linux* 5.4
- Red Hat Enterprise Linux 5.5
- SUSE Linux Enterprise Server* 11 SP1

Intel NIC VT-d, SR-IOV サポート確認方法

<http://www.intel.com/products/ethernet/resource.htm>

Controllerの型番からデータシートや参考資料が入手可能



The screenshot shows the Intel Ethernet Resource Center interface. It features three dropdown menus for filtering: 'Select Category' (with 'All' selected), 'Select Product' (with '82599EB/ES' selected), and 'Select Resource Type' (with 'Application note' selected). Below these is a 'Keyword Search' field and a 'Tool Tips' link. A pagination control shows '<< < 1 > >>' and a page number '25'. The main content is a table with the following data:

Product	Category	Resource Link	Resource Type
82599EB/ES	10 Gigabit Ethernet	Intel® 82599 SR-IOV Driver Companion Guide, 1.0	Application note
82599EB/ES	10 Gigabit Ethernet	PCI-SIG SR-IOV Primer: An Introduction To SR-IOV Technology, 2.5	Application note
82599EB/ES	10 Gigabit Ethernet	Intel® 82599 10 GbE Controller Datasheet, 2.71	Datasheet
		Assigning Interrupts to Processor Cores	Application note
		Avoiding New Design Errors When Using Intel® Ethernet Controllers	Application note
		Manageability – Intel® Sideband Technology	Application note
		PCI Express* whitepaper	Application note
		Boot Agents for BIOS Engineers	Application note

Mother Board VT-d, SR-IOV サポート確認方法

- VT-d サポートする Chip Set 記載
 - <http://wiki.xensource.com/xenwiki/VTdHowTo>
- Mother Board は落とし穴が多数...
 - BIOSが対応してないとダメ！
 - メーカー独自管理パーツついてるとうまくいかない場合も。。。

まとめ

- 仮想化環境のパケット転送性能は構成に大きく依存
- ハードウェアによって使えない構成(仮想化技術)も...



試して、動いたらみんなシェア！
こんな構成で動いたよ。性能でたよ！

バッドノウハウも！
この構成だと性能でない、動かない...

宛先はこちら: vyatta-users@vyatta-users.jp
<http://www.vyatta-users.jp/>

