



## MellanoxのOpenStack最新ソリューションと最新ユーザ事例！

メラノックステクノロジーズジャパン株式会社

July 6, 2016 | OpenStack Days Tokyo 2016

 **Mellanox**  
TECHNOLOGIES  
Connect. Accelerate. Outperform.™

- **広帯域、低レイテンシーインターコネクトのリーディングカンパニー**

- EDR 100Gb/s InfiniBand、10/25/40/50/56/100ギガビットEthernet
- アプリケーションのデータ処理時間を大幅に削減
- データセンターサービス基盤のROIを劇的に向上

- **会社概要**

- 本社：ヨークナム（イスラエル）、サニーベール（米国）
- 従業員数：全世界で2,400名（2016年3月末時点）

- **財務状況**

- 2013年度売上 : \$390.9M
- 2014年度売上 : \$463.6M
- 2015年度売上 : \$658.1M
- 2016年度第1四半期 : \$196.8M
- 2016年度第2四半期 : \$210M - \$215M
- Cash + Investment : \$261.8M（2016年3月末時点）



# 超高速ネットワーク市場をリードするエンドトゥエンド製品群



## InfiniBand、Ethernetエンドトゥエンドソリューションを実現する製品群

IC



アダプタカード



NPU、マルチコア



スイッチ/ゲートウェイ



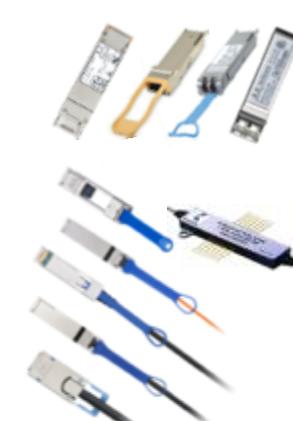
ソフトウェア



Metro / WAN

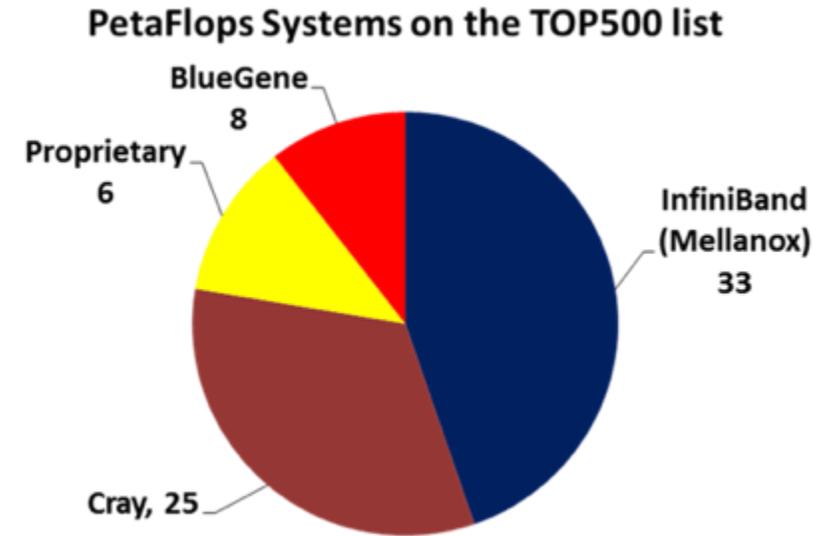
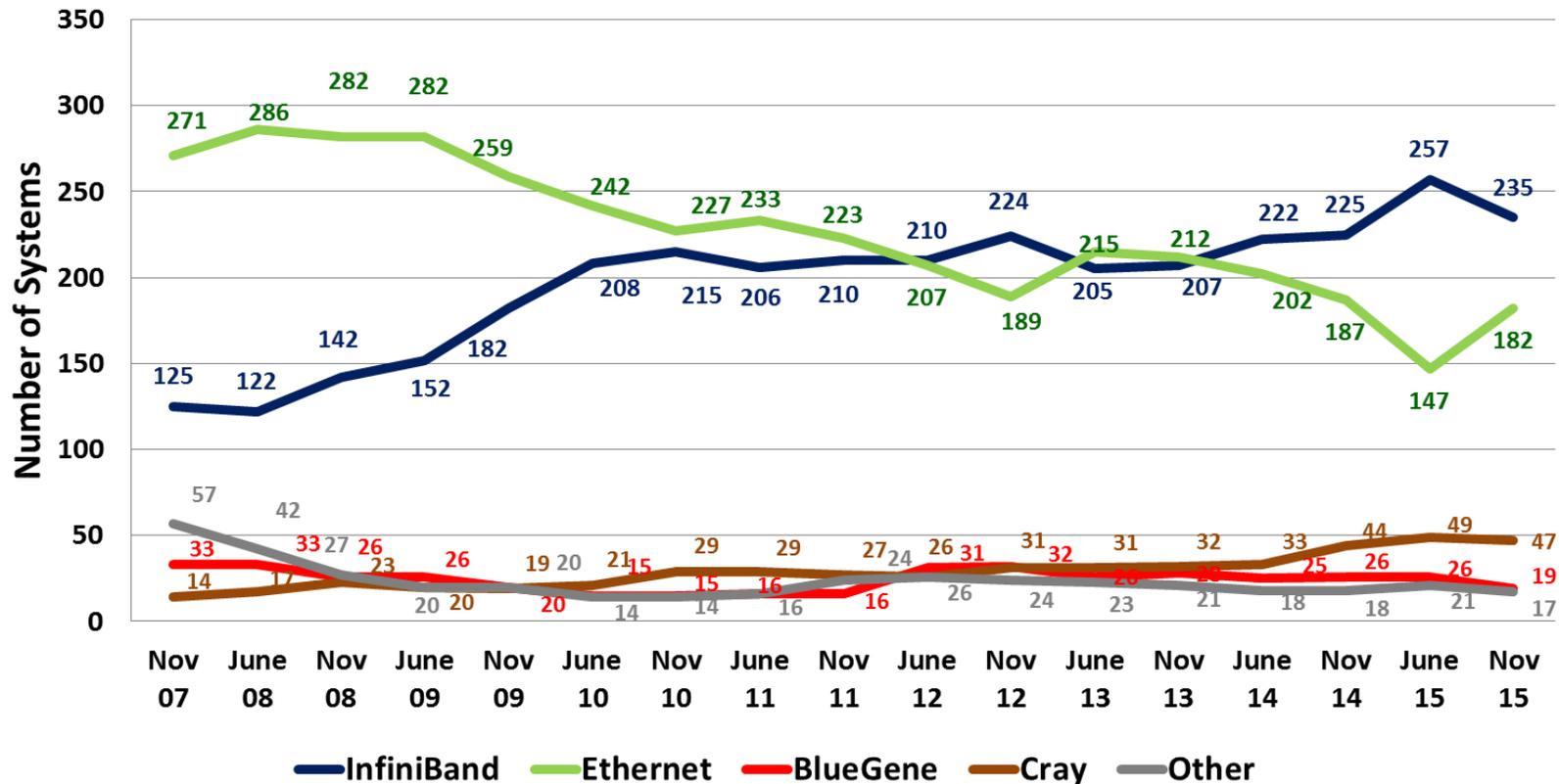


ケーブル/モジュール



# TOP500におけるインターコネクットのトレンド

## TOP500 Interconnect Trends



**InfiniBandはパフォーマンスが必要なアプリケーションではデファクトのインターコネクットソリューション**

# さまざまな市場（HPC、クラウド、Web 2.0、ストレージ）における実績



## ペタスケールクラスタの半数で採用実績

**Microsoft Windows Azure**  
90.2% Cloud Efficiency  
33% Lower Cost per Application

**ORACLE Cloud**  
Application Performance  
Improved up to 10X



**PROFITBRICKS**  
The IaaS-Company.  
3X Increase in VMs per Physical Server  
Consolidation of Network and Storage I/O



**atlantic.net**  
32% Lower Cost per Application  
694% Higher Network Performance



13 Million Financial Transactions Per Day, 4 Billion Database Inserts  
Real Time Fraud Detection



Accuracy, Details, Fast Response  
10X Higher Performance, 50% CAPEX Reduction

Microsoft  
Bing Maps



Reacting to Customers' Needs in Real Time!  
Reducing Data Queries from 20 minutes to 20 seconds



235 Supermarkets, 8 States, USA

97% Reduction in Database Recovery Time  
From 7 Days to 4 Hours!

Tier-1 Fortune100 Company  
Web 2.0 Application



## メラノックスのテクノロジーが事業拡大に貢献



## InfiniBandによりクラウド環境でアプリケーションコストの最小化を実現

## ストレージ市場における多くの採用実績

# 高性能・低コストEthernetスイッチ製品



**SX1036 (PowerPC) / SX1700/SX1710 (x86)**  
理想的な40GbE ToR/アグリゲーションスイッチ



**SX1024 (PowerPC) / SX1400/SX1410 (x86)**  
ノンブロッキング10GbE → 40GbE ToR



**SX1016**  
最高密度の10GbE ToR



**SX1012**  
ストレージ/データベースに最適な10/40GbEスイッチ (POCにも最適)



**SX1036 – 83ワット**  
**SX1016 – 62ワット**  
**SX1024 – 75ワット**  
**SX1012 – 50ワット**

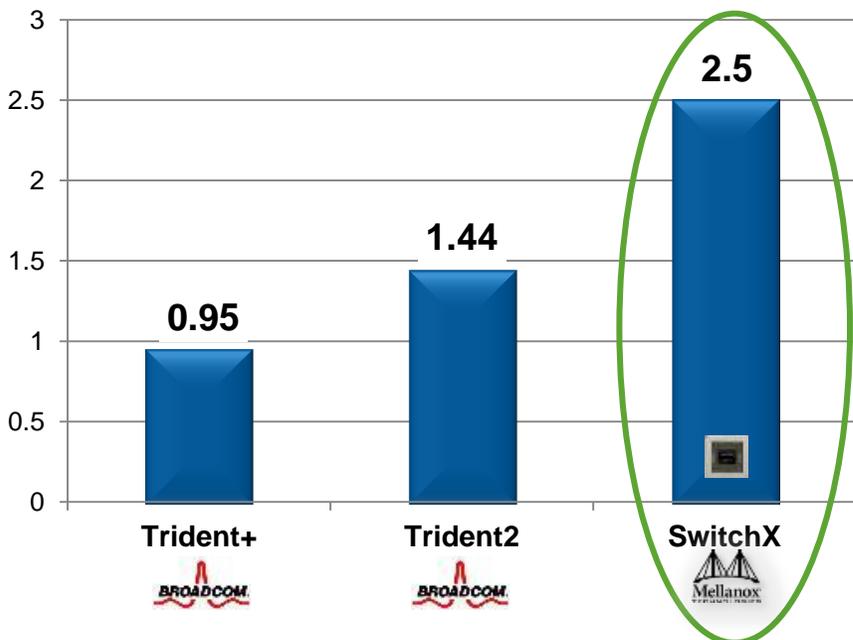
## 220ナノ秒

# 最高のROIを実現するEthernetスイッチ

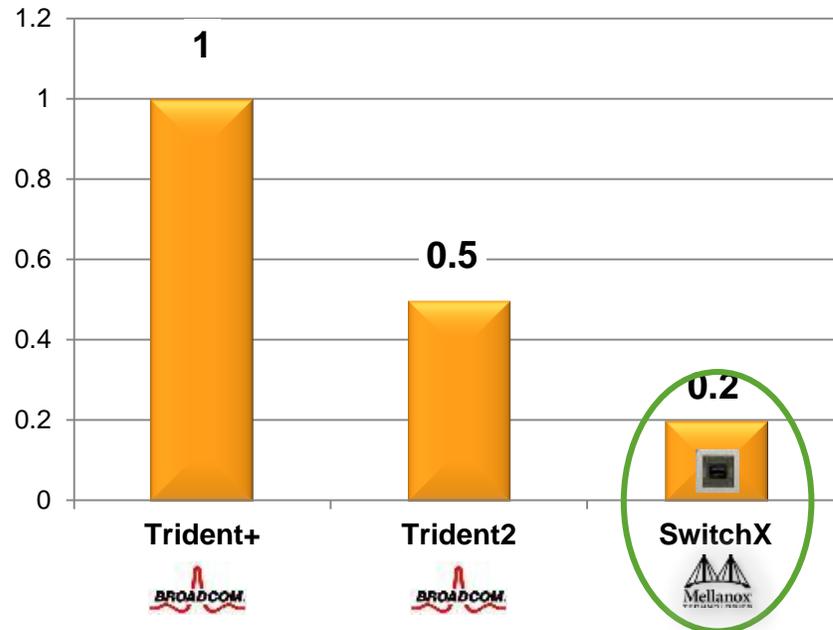
## ■ Ethernetスイッチに必要なテクノロジー

- 最高のスイッチング性能
- 低レイテンシー
- 低消費電力

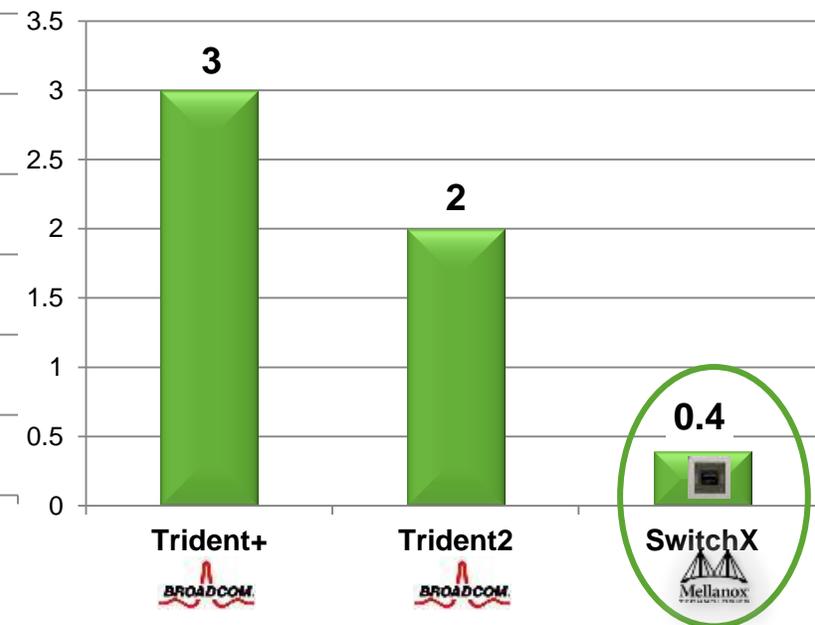
### スイッチング性能 (Bpps)



### レイテンシー (マイクロ秒)



### 消費電力 (ワット/Gb)



100GbE対応 (25GbE = 次世代10GbE、50GbE = 次世代40GbE)

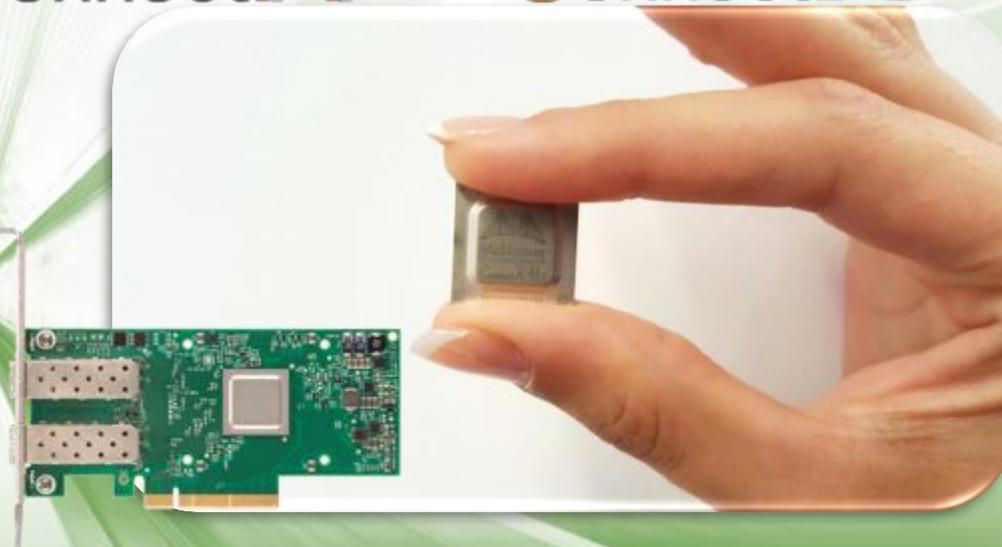


  
**Spectrum™**



柔軟性、多目的対応、スピード  
Open Ethernet、ゼロパケットロス

Connect X<sup>®</sup>·4 · Connect X<sup>®</sup>·4 Lx



もっともコスト効率が高いアダプタ  
共通インフラストラクチャ・コネクタ

1台のスイッチで多用途に対応

25、50、100Gb/sを身近なものに

- 優れたハードウェア
  - 最高品質のハードウェア
  - FRU対応の信頼性
  - 社内設計



- ソフトウェアの柔軟なサポート
  - ONIE、OCP
  - SDK、SAI
  - MLNX-OS®
  - ホワイトボックス対応



**SN2700 – 32 x 100GbE (64 x 50GbE)**  
理想的な100GbE ToR/アグリゲーションスイッチ



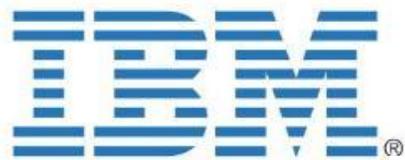
**SN2410 – 8 x 100GbE + 48 x 25GbE**  
25GbE → 100GbE ToR



**SN2100 – 16 x 100GbE**  
ストレージ/データベースに最適な25/100GbEスイッチ  
1RUでの最高密度の25GbEスイッチ

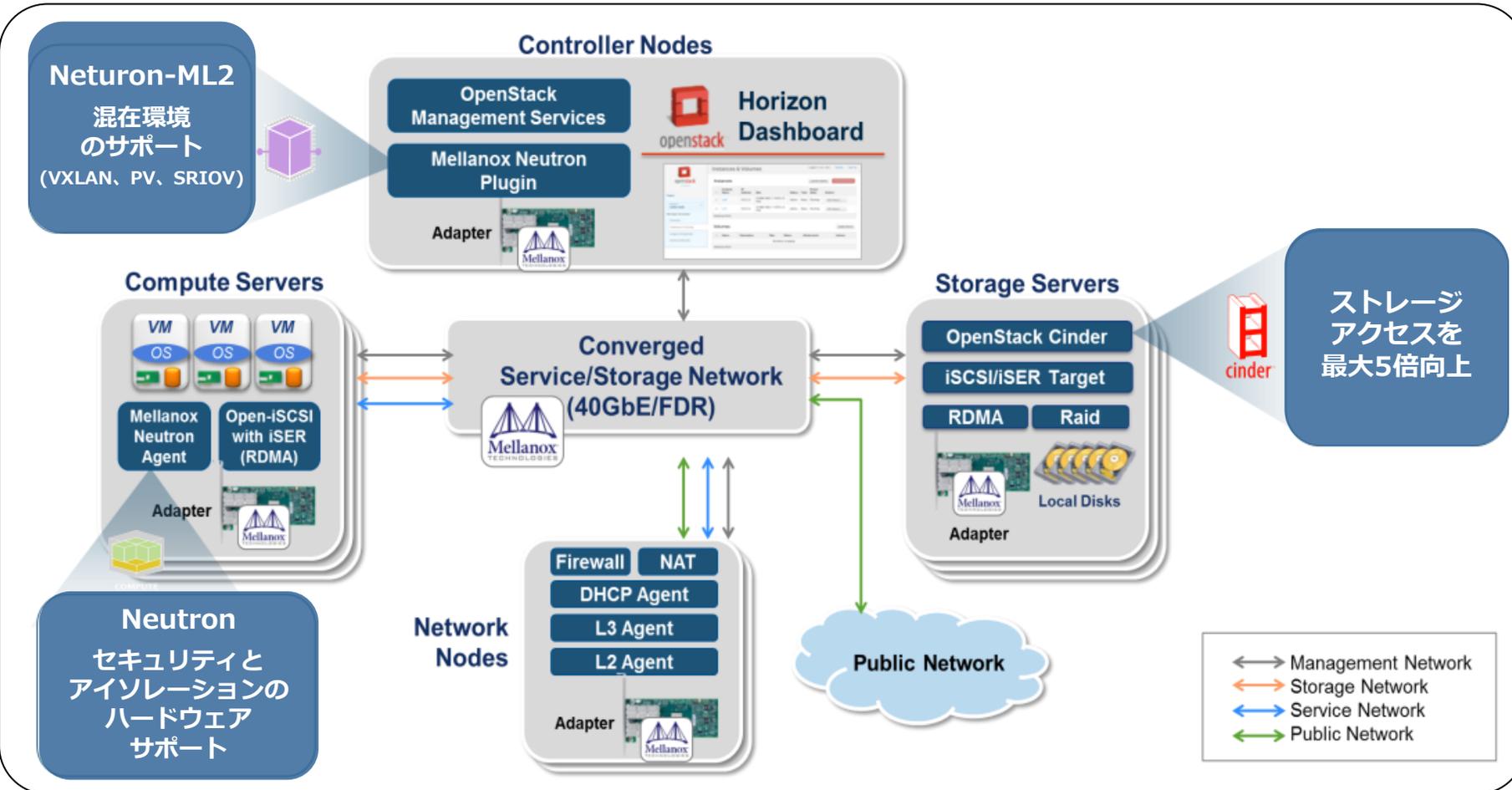


# さまざまなEthernetスイッチとの相互接続性



# メラノックスがリードするOpen Ethernet Solution

# スイッチとアダプタによる包括的なOpenStackの統合



主要OpenStack  
ディストリビューション  
でサポート



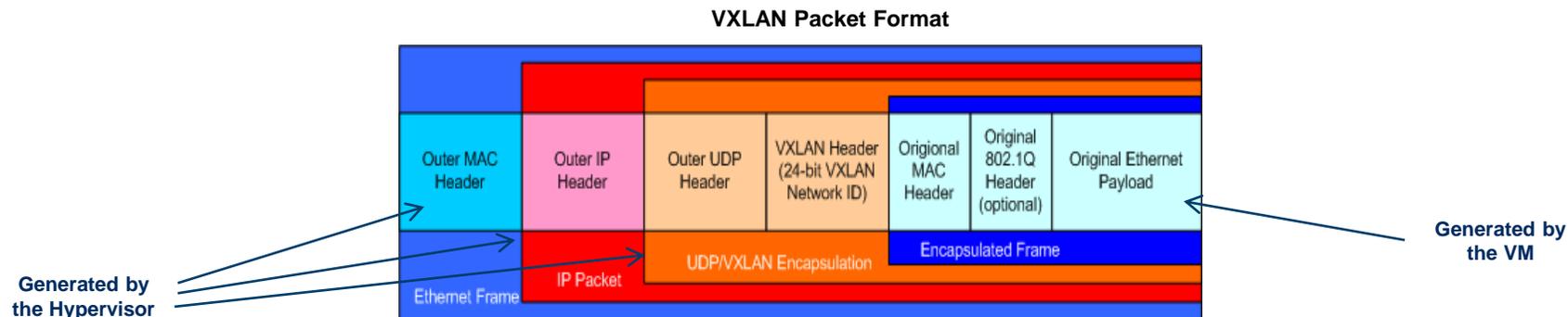
インボックス  
(Havana、Ice House)

THE NINTH OPENSTACK RELEASE  
ICEHOUSE



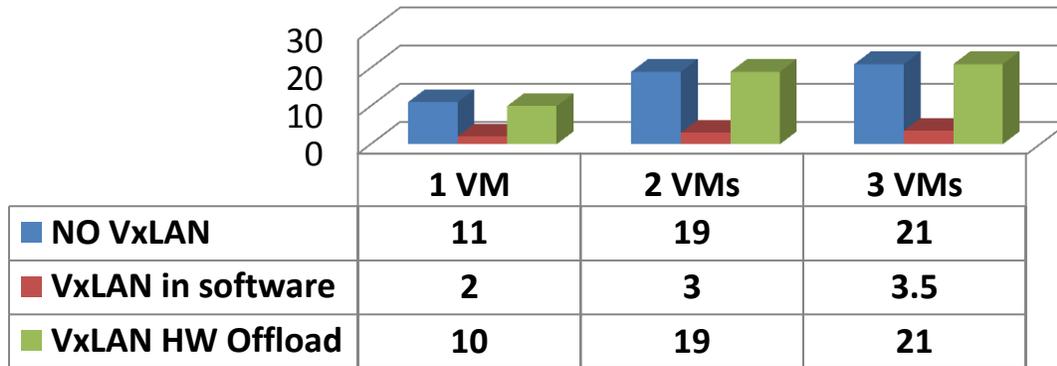
ストレージアクセラレーション / vSwitch NICオフロード / オーバーレイネットワークオフロード

- ハイパーバイザのIPスタックや標準的なNICはクライアントTCP/IPのトラフィックを認識できない
- ハードウェアによるセグメンテーション/リアセンブリ、チェックサムオフロード、RSS/TSS CPUコアのスケーリングなど標準的なオフロード技術では、仮想マシンが生成するTCP/IPパケット（インナーペイロード）を処理できない
- これらの処理をCPUに負担させることにより、VXLANやNVGREを使用する環境では、ホストCPUに多大な負荷がかかり、アプリケーション性能やサーバ統合率に悪影響を与える
- メラノックスのConnectX<sup>®</sup>-3 Proは、これらの仮想マシンが生成するTCP/IPパケットを処理できる世界初のNICであり、VXLANやNVGRE使用時におけるCPUのオーバーヘッドをなくすことで、システム性能に影響を与えずにオーバーレイ型ネットワーク仮想化環境へ移行できる



## VXLAN

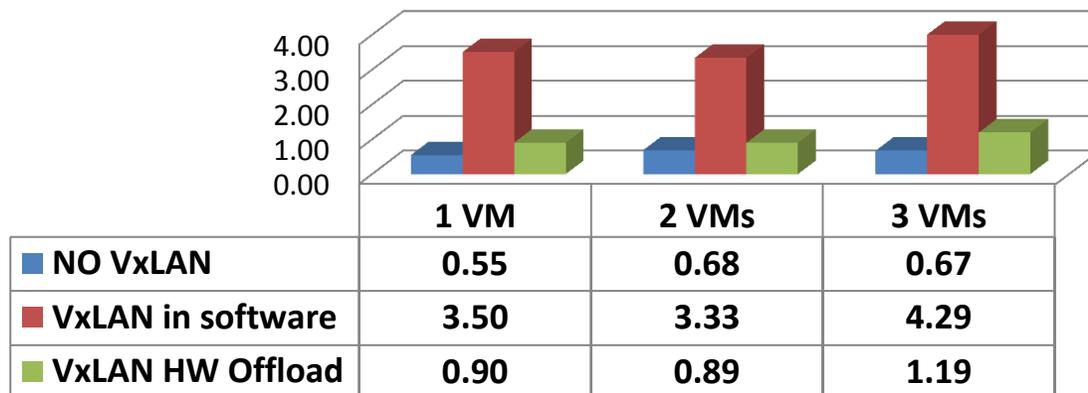
VXLAN使用時のVM全体のバンド幅



**VXLANオフロードにより  
バンド幅が5倍に**

**VXLANオフロードにより  
CPU使用率が1/4に**

VXLAN使用時のCPU使用率 (Gbit/sec あたり)



### Test Details

- Test command: netperf -t TCP\_STREAM -H
- 1-3 VMs talking to 1-3 VMs on a second server
- OpenvSwitch (OVS) with VXLAN support
- Servers:
  - HP ProLiant DL380p Gen8
  - 2 x Intel(R) Xeon(R) CPU E5-2650 0 @ 2.00GHz
  - 32GB RAM
- Hypervisor OS: Linux Upstream Kernel 3.14-rc1 + KVM
- Guest VM OS: RHEL 6.5 2.6.32-431.el6.x86\_64
- NIC: ConnectX<sup>®</sup>-3 Pro , FW: 2.30.8000
- CPU% and Bandwidth measures on the Hypervisor (aggregate 1-3 VMs)

ソフトウェアを選択・作成する自由が与えられることにより、  
一切の妥協が不要となり、ユーザ、ベンダーは差別化が容易に

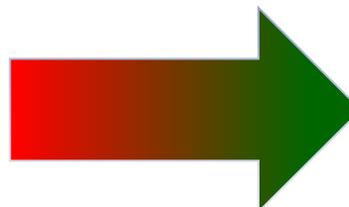
## Open Ethernet



ベンダー主導の管理

ベンダー主導の  
ソフトウェア選び

クローズな  
Ethernetスイッチ



実際の環境に  
最適な管理

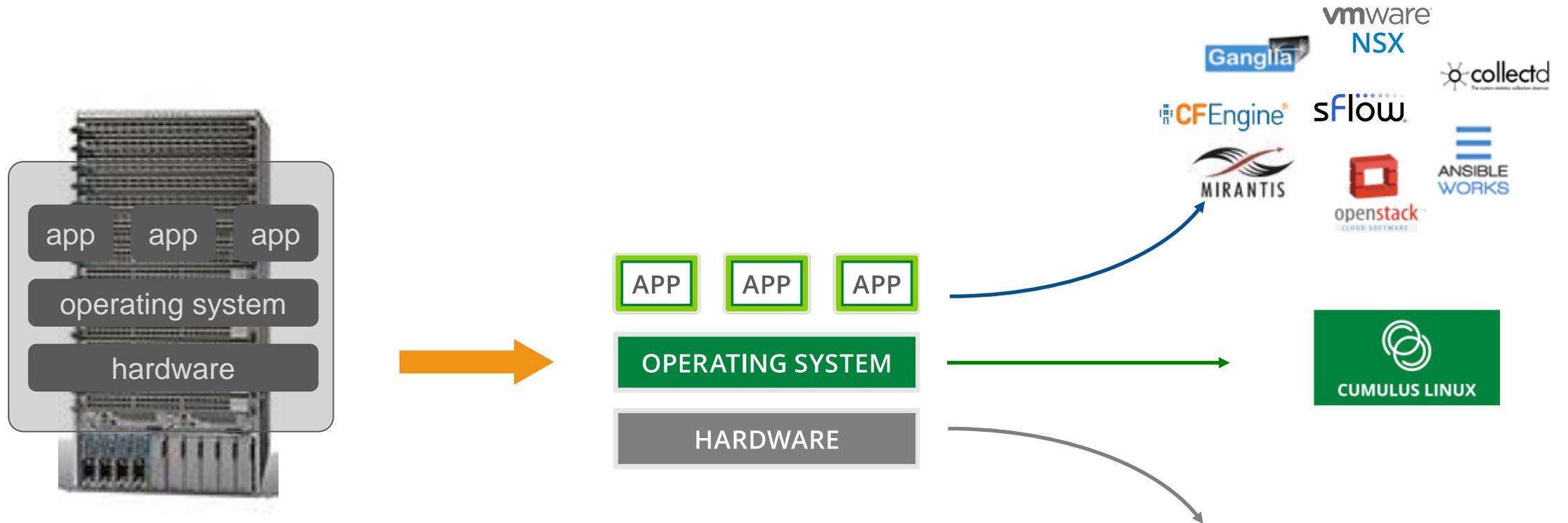
本当に使いたい  
ソフトウェア

オープンな  
Ethernetスイッチ



ロックダウン垂直統合型ソリューション

オープンプラットフォーム



## シングルベンダーの弊害

- 進まない技術革新
- マージンはベンダーへ
- 遅い機能サポートサイクル

## 顧客の選択

- 活気に満ちたコミュニティ
- ベンダーも全力でサポート
- 迅速な開発



- ストレージファブリック
- RoCE対応
- Layer 2 ToR対応
- Multi-Tier MLAG
  - リーフ、スパイン対応
- MLNX-OS Scale Layer 3対応
  - Fat Treeトポロジー
- OpenFlow
- CCIEs - Cisco Trained Network Admins



- 大規模データセンター
  - ルーティングテーブル
- テレコム向けソリューション
  - MPLS
- クラウドサービスプロバイダー、プライベートクラウド
  - ハードウェアによるオーバーレイ対応
- オンスイッチアプリケーション
  - セキュリティ、ファイアウォール、接続トラッキング、ロードバランス
  - 400Mbps以上のワークロード CPU<->ASIC
- マネジメントスイッチ - 1000BASE-T
- ネットワークでLinuxが必要なケース



# テクノロジーエコシステムパートナー



Routing

Network Automation

Orchestration

Network Virtualization

Monitoring

Security

Storage & Hyper-Converged

Big Data

Cumulus Linux



Industry Standard Hardware

# 主要オーバーレイソリューションで採用されているCumulus



VMware NSX



PLUMgrid ONS



Nuage VSP



Midokura Midonet



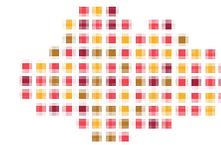
Juniper OpenContrail



Akanda Astara



Cumulus LNV



nuagenetworks



midokura



VXLAN availability on Mellanox - Q2 2016

# クラウドエコシステムの自動化と一元管理

# メラノックスの統合ネットワーク管理ツール“NEO”

## ■ 集中管理

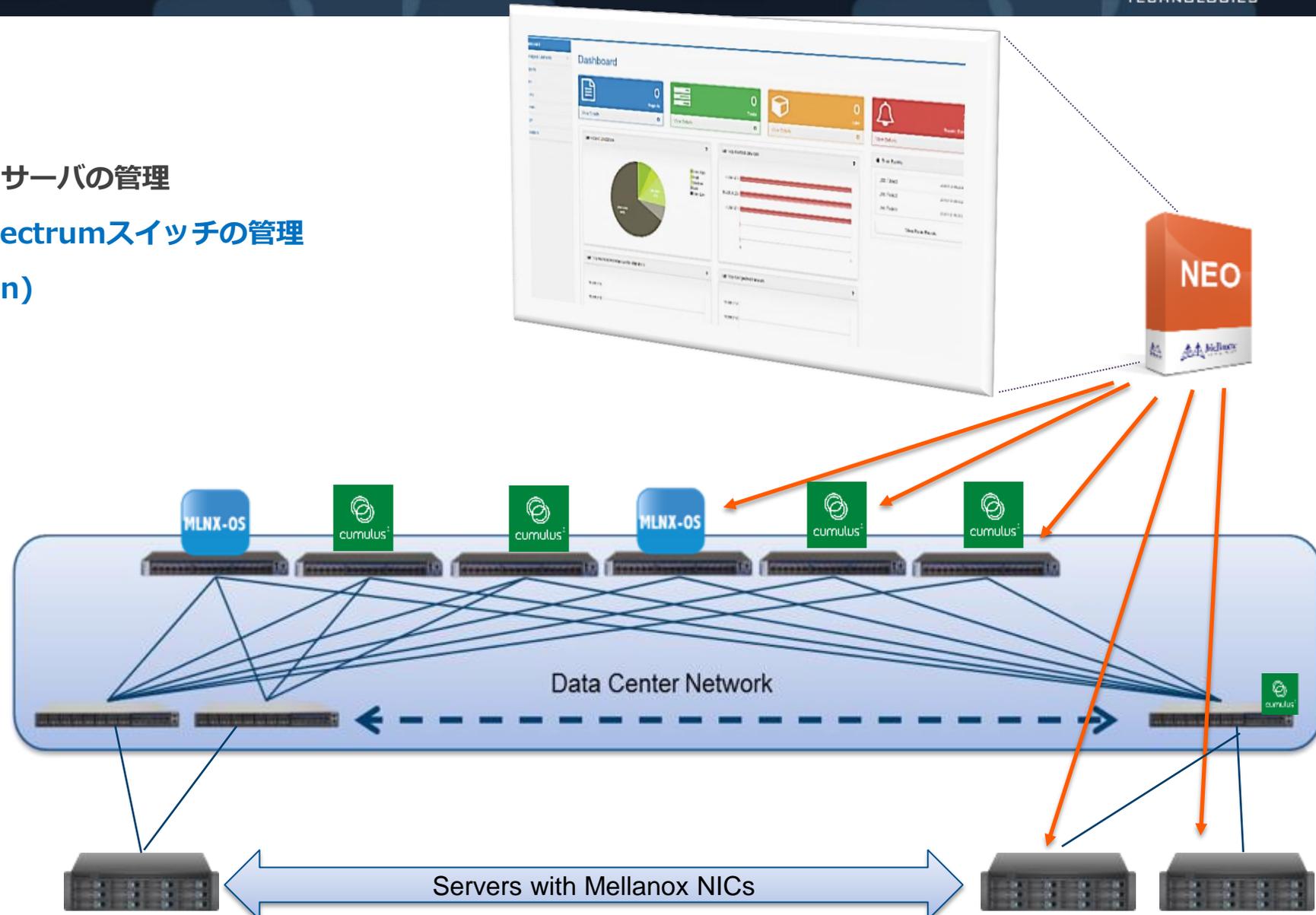
- 管理者の業務効率を大幅に向上
- 一つのモニタから数百以上のスイッチ、サーバの管理
- Cumulus、MLNX-OSが適用されたSpectrumスイッチの管理
- OpenStack連携(Neutron ML2 plugin)

## ■ オートメーションツール

- ネットワークディスカバリ
- NIC、スイッチのプロビジョニング
- ファームウェアアップデート
- バルクコマンドのプッシュ
- ネットワーク全体へのセキュリティ適用

## ■ システム全体のチェック

- パフォーマンストラッキング
- デバッグの容易性



## OpenStackと連携したスイッチマネジメント



### ➤ Mellanox SDN ML2 プラグインにより OpenStack Neutron と連携

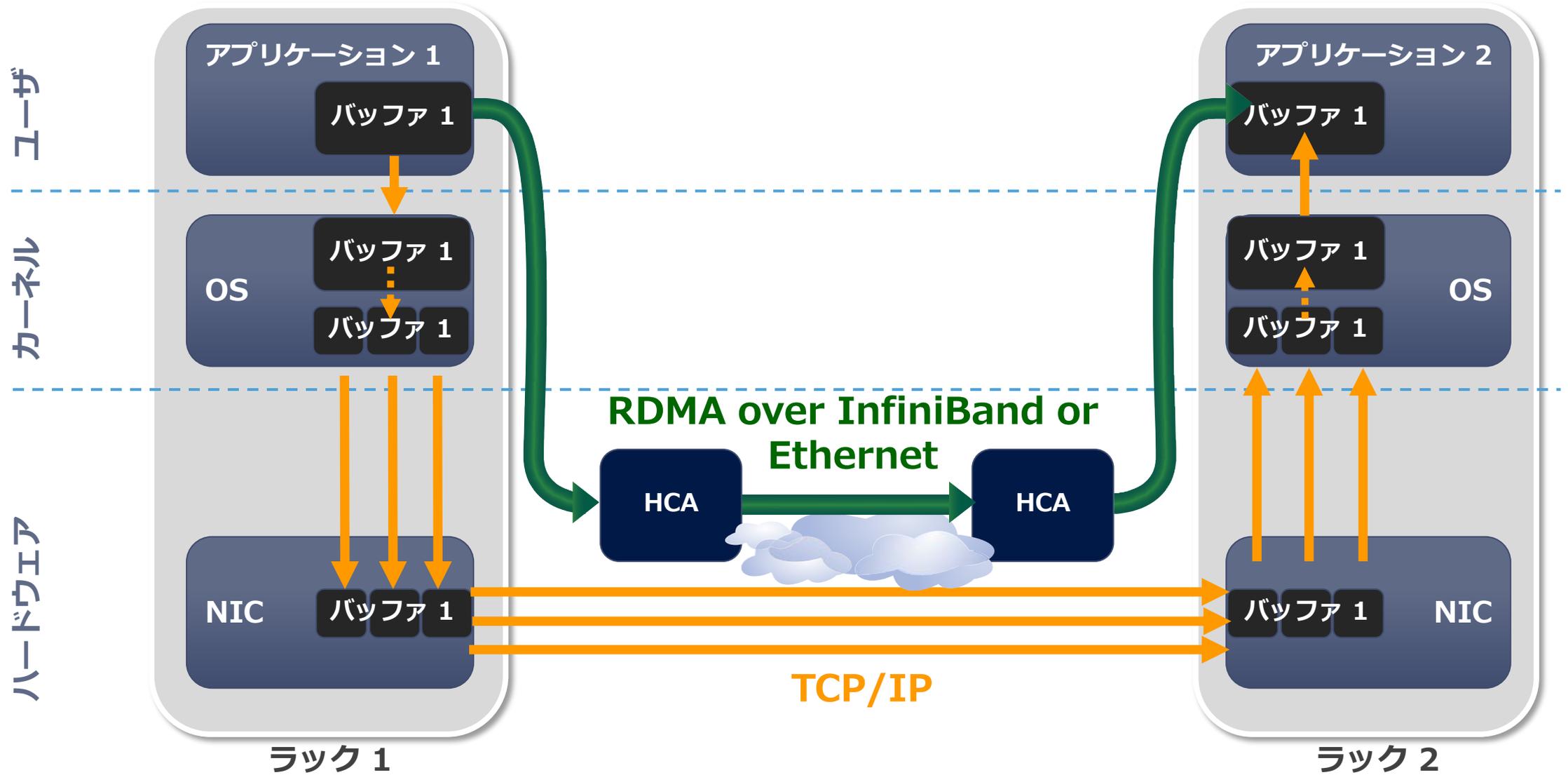
- OpenStack からの仮想マシンデプロイに応じて自動でスイッチの VLAN を設定
- ベアメタルプロビジョニング (Ironic) にも対応予定

### ➤ OpenStack の様々なコンポーネントとの連携 ※

- Horizon と連携し OpenStack ダッシュボードから物理/仮想ネットワークを一元管理
- Nova と連携したスイッチ管理のスケジューリング機能

※ 今後対応予定

# クラウド向けの新しいSANとは？



- フラッシュメモリのコモディティ化によりSSDストレージの利用が促進
  - ・ SSDは共有ストレージへのアクセス競合を緩和し、ランダムIO性能を劇的に向上
- SSDは最小のレイテンシーと低消費電力のストレージを実現
  - ・ 最高性能だがまだ高価なSSDを低価格で利用するには、SSDのアプリケーションへの応答性能を妨げない、低レイテンシーなインターコネクト技術による外部共有が必須

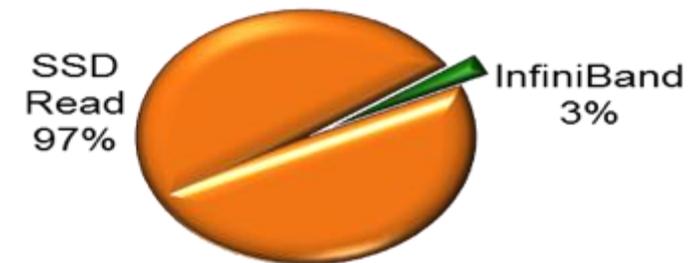
Component	Latency (usec)
InfiniBand	0.7
Ethernet RDMA (RoCE)	1.3
Ethernet TCP	6
Fibre Channel	20
SSD Read	25
Disk Read	6000

Total SSD Read Latency Over Fibre Channel



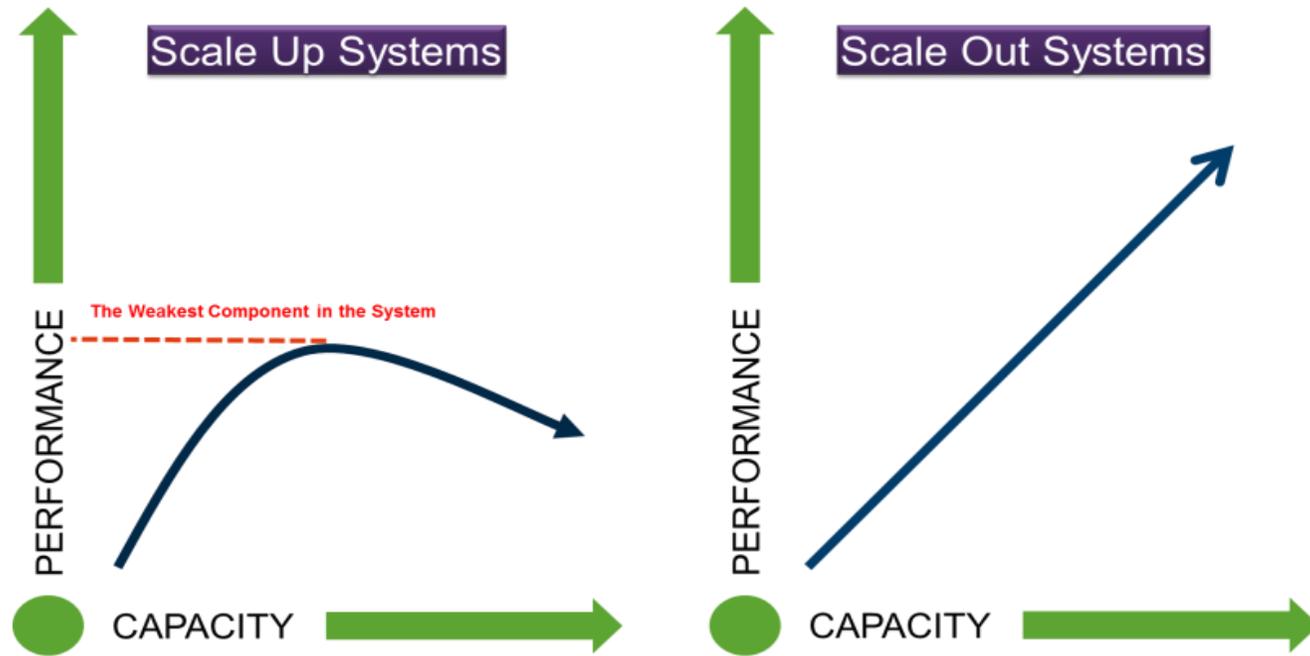
Total: 45usec

Total SSD Read Latency Over InfiniBand

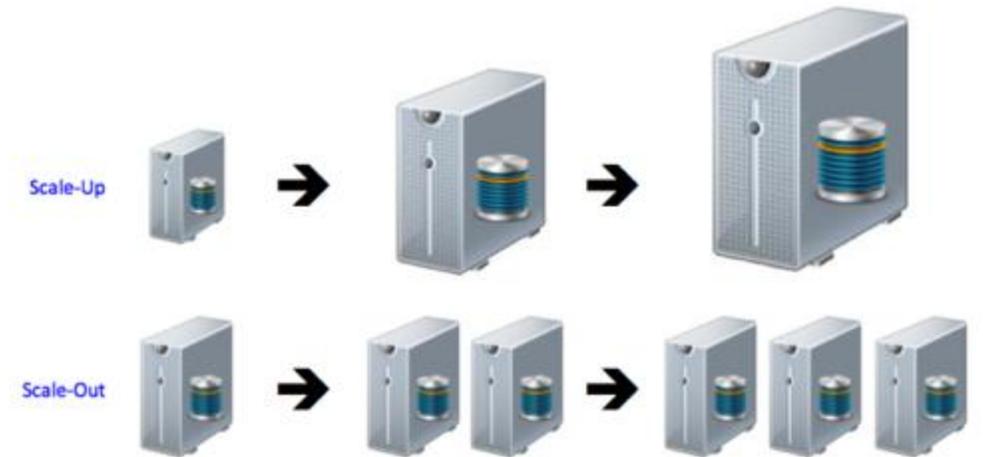


Total: 25.7usec

Fibre ChannelではSSDのパフォーマンスを活かすことはできない  
SSDの共有にはInfiniBand、RoCE (RDMA over Converged Ethernet) が最適!



スケールアウトのパフォーマンスはネットワーク能力に依存



- スケールアウトはスケーラビリティ、オンデマンド対応を実現するために必須
- スケールアップからスケールアウトへの移行
  - ・ コストを抑えてストレージ容量を増やす唯一の方法
  - ・ クラウド、ビッグデータ、HPCで採用
- 新しいスケールアウトの選択

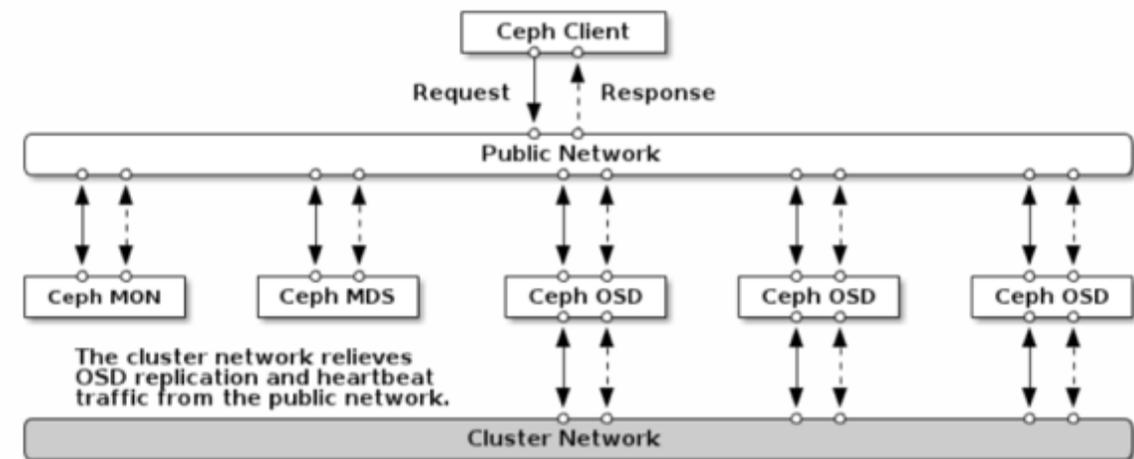
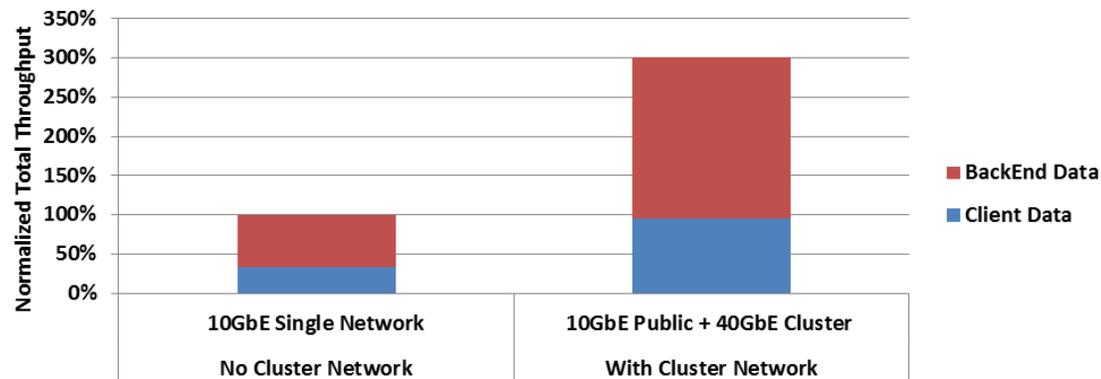
- ストレージクラスタで最高の可用性を実現するためには、ハイパフォーマンスのネットワークが必要
  - ・ クライアント、OSD、モニタ、メタデータサーバは複数のネットワークレイヤを通して通信を行う
  - ・ ハートビート、レプリケーション、リカバリ、リバランシングではリアルタイム処理が必要
- クラスタのパフォーマンス、スケーラビリティは、クラスタ（バックエンド）ネットワークに依存
  - ・ Ceph OSDデーモン間のネットワーク負荷は、CephクライアントとCephストレージクラスタ間のネットワーク負荷により容易に影響を受ける。（Cephドキュメントより）

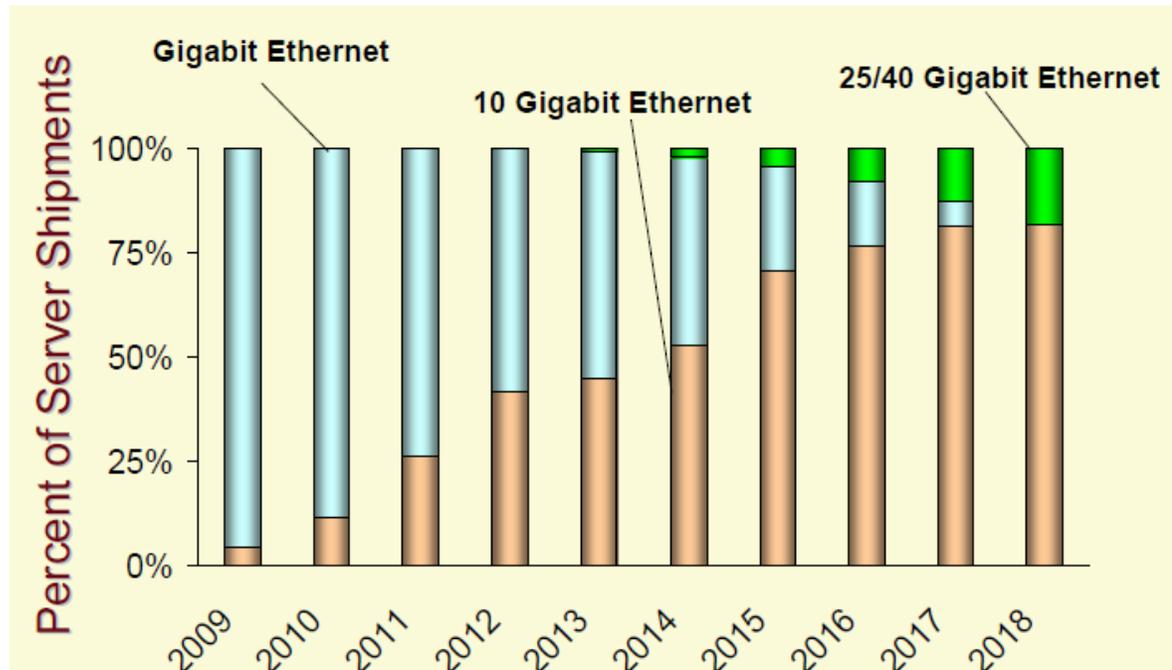


ceph

### Aggregated Total Throughput on Ceph Cluster

(base line, 10GbE Single Network)





(Source: Dell'Oro Group)

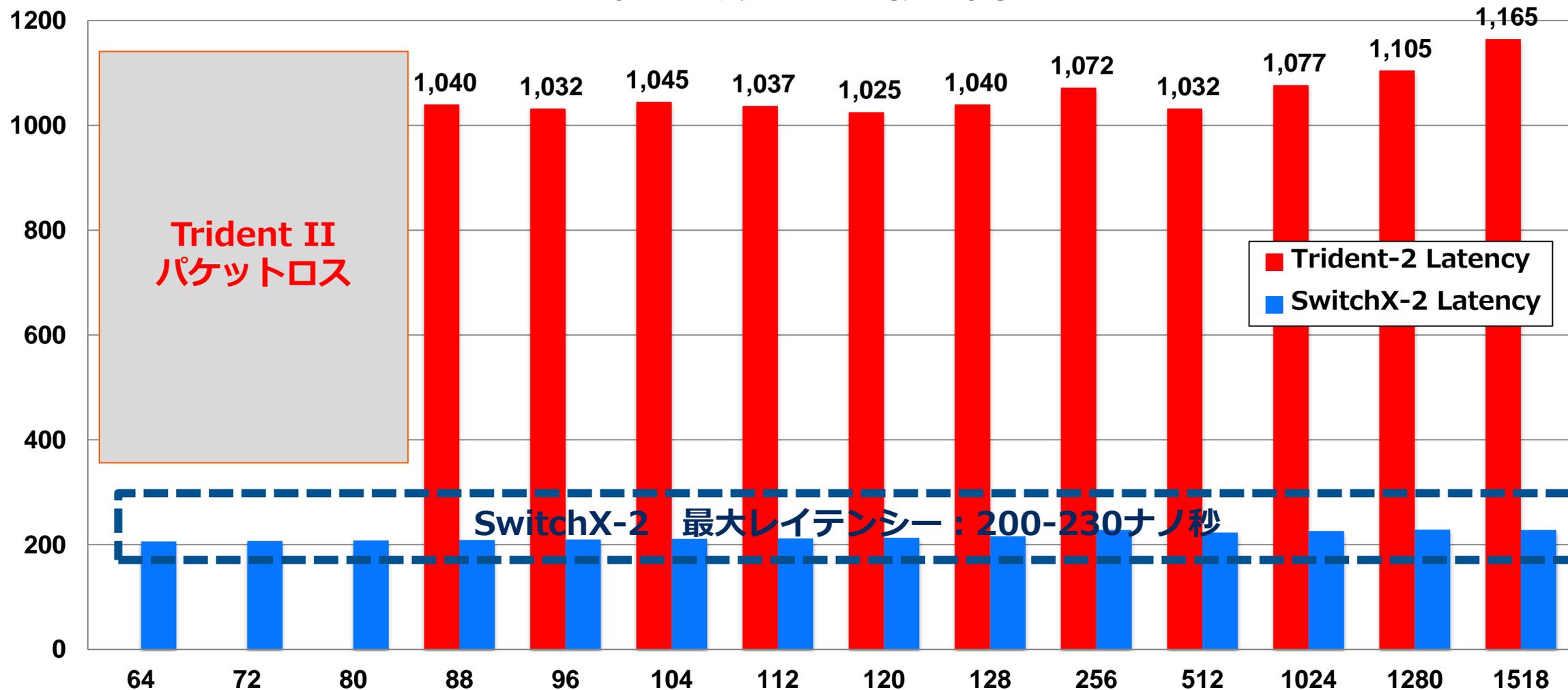
- **10/40/56GbE分野におけるリーダー**
  - ・ 主要クラウド、Web 2.0ベンダーでの採用の増大
- **40GbEでは独占的なベンダー** (全ポートの92%\*)
  - ・ もっとも急速に拡大しているセグメント (四半期ごとの比較で78%の伸び率)
- **25/50/100GbEも早々に登場予定**

\* Source: Crehan Report, August 2014

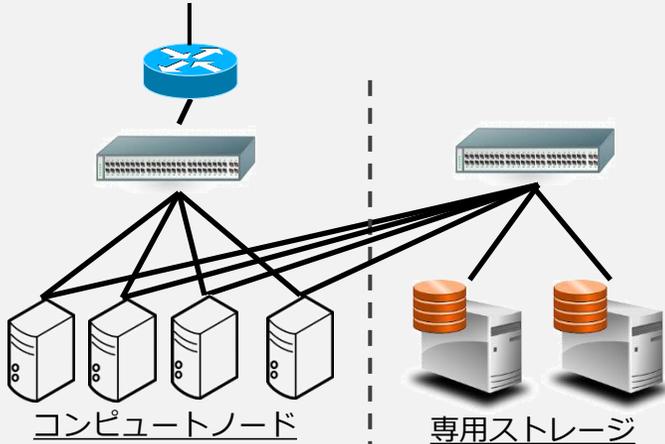
- **10Gb/s以上のEthernetアダプタのマーケットシェアが拡大**
  - ・ クラウドベンダーは10ギガビットEthernetから40ギガビットEthernetへ移行
  - ・ さまざまな利点を持ち、差別化が可能なことから、RoCEへの注目が拡大

## IXIA RFC 2544 Testing

### 最大レイテンシー (ナノ秒)



## BEFORE

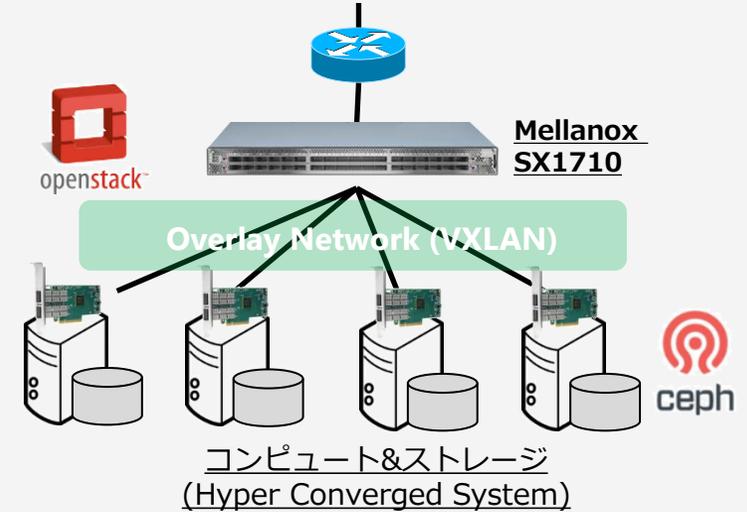


### お客様の課題

- ✓ コモディティサーバによるcomputeとストレージの統合をしたい
- ✓ 分散ストレージ採用におけるサーバ間通信のIOボトルネック
- ✓ 仮想ネットワークを採用する事によるサーバCPU負荷の増加



## AFTER

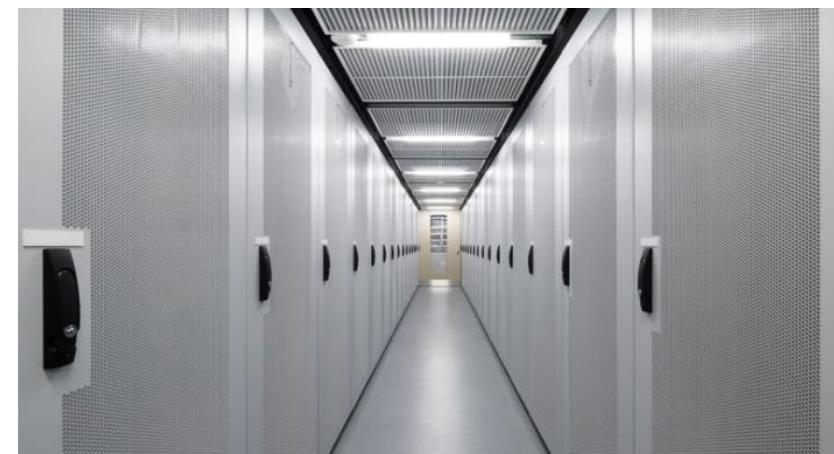
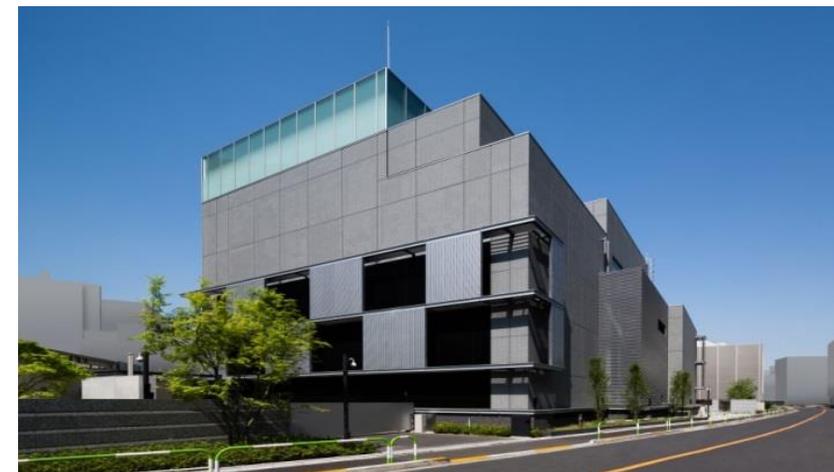


### 導入効果

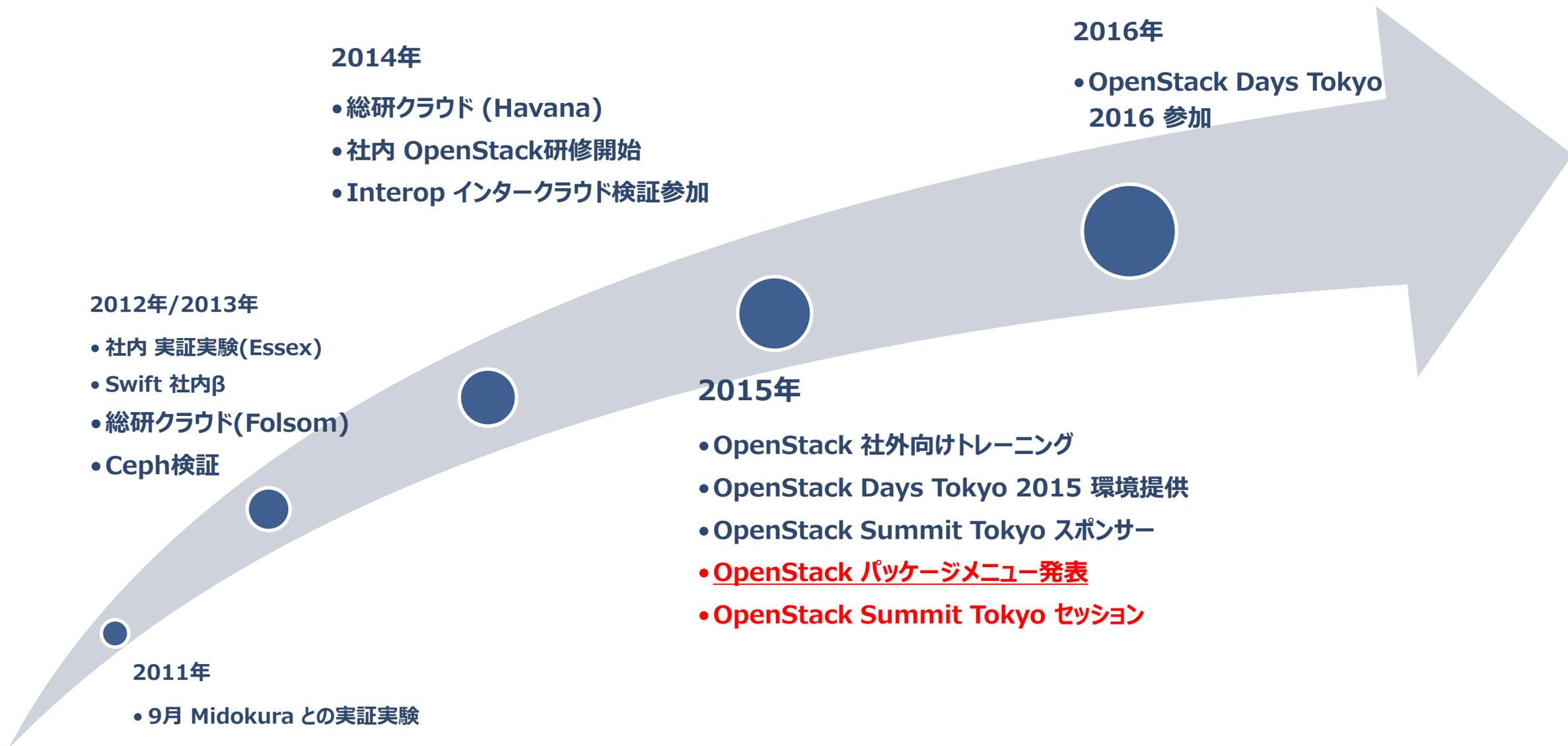
- ✓ 広帯域かつ低レイテンシの安定した物理ネットワーク
- ✓ VXLANオフロードによるサーバ負荷低減と性能向上
- ✓ ネットワークスイッチとNICをメラノックス製品に統一したことによるEnd-to-Endのサポート

- 資本金 : 35億69百万円
- 代表者 : 代表取締役 古田敬
- 設 立 : 2000年6月14日
- 所在地 : 〒140-0002  
東京都品川区東品川2-2-28 Tビル  
Tel:03-5805-8151 (代表)  
Fax:03-3474-5540
- URL : <http://www.bit-isle.jp>
- 株主 : QAON合同会社 (エクイニクスグループ)
- 主要サービス : データセンターサービス、  
クラウドサービス、運用サービス、  
システムインテグレーション

- グループ会社 :  BIT-SURF  siteROCK



# ビットアイルとOpenStack の取り組みの歩み



# OpenStack パッケージ

ビットアイル ホステッドプライベートクラウド  
**OpenStackパッケージ**



“自立運用をゴールとするプライベートクラウドパッケージ”

## OpenStackパッケージ

高品質なプライベートクラウドパッケージを使いながら、スキルを習得。  
 自立運用を叶え、OSSの魅力であるコストメリットを最大限享受！

### ★ Step ①

月額でご提供する推奨構成での利用を開始。お客様スタッフのレベルを問わず、当社OpenStackサポートを受けながら、スタートする事が可能です。

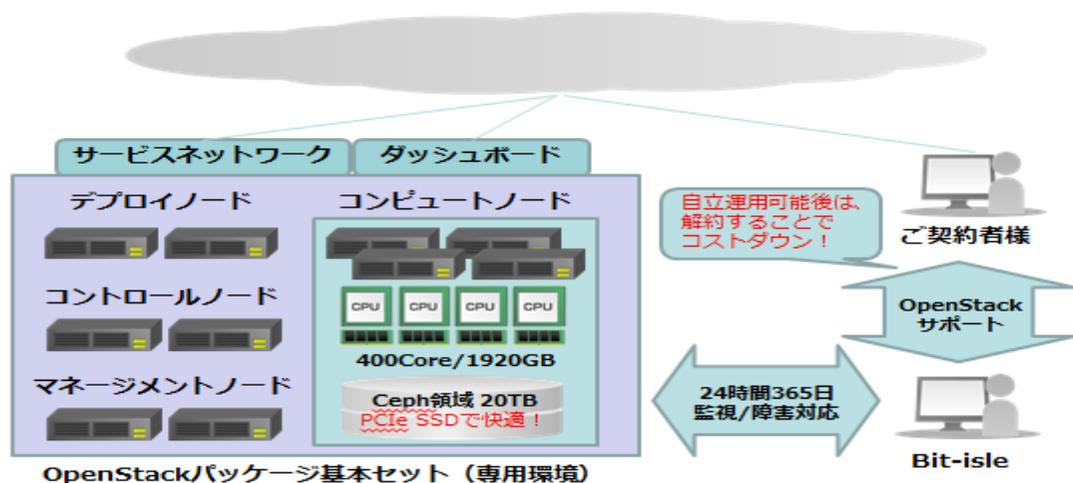
### ★ Step ②

当社よりスキルMAPを提示。各種機能を使いながら、スキルを習得頂きます。

### ★ Step ③

自立運用が可能となるフェーズに達しましたら、当社OpenStackサポートを解約。  
 OSSの魅力であるコストメリットを享受したプライベートクラウドの利用が可能です！  
 (24時間365日のインフラ監視、オンサイト復旧作業等は引き続きご提供します)

## 提供イメージ



## 提供内容・価格・特徴

### 提供内容・価格

#### ● OpenStackパッケージ基本セット

初期費用 ¥3,000,000 月額費用 ¥3,600,000  
 (OpenStackサポート解約後月額費用 ¥2,300,000)

##### ◆主要提供内容

- ・1Gbps共用インターネット帯域 & Global IPアドレス128(/25)
- ・10Gbps専用内部ネットワーク
- ・CPU:Xeon E5世代2.6GHz×400Core Memory:1920GB HDD領域:20TB  
※専用仮想マシン(1Core/4GB/50GB)×400台相当のリソース!
- ・OpenStackサポート ※12ヶ月経過後以降解約可能

#### ● OpenStackパッケージ追加セット

初期費用 ¥2,100,000 月額費用 ¥1,800,000

##### ◆主要提供内容

- ・CPU:Xeon E5世代2.6GHz×400Core Memory:1920GB HDD領域:20TB

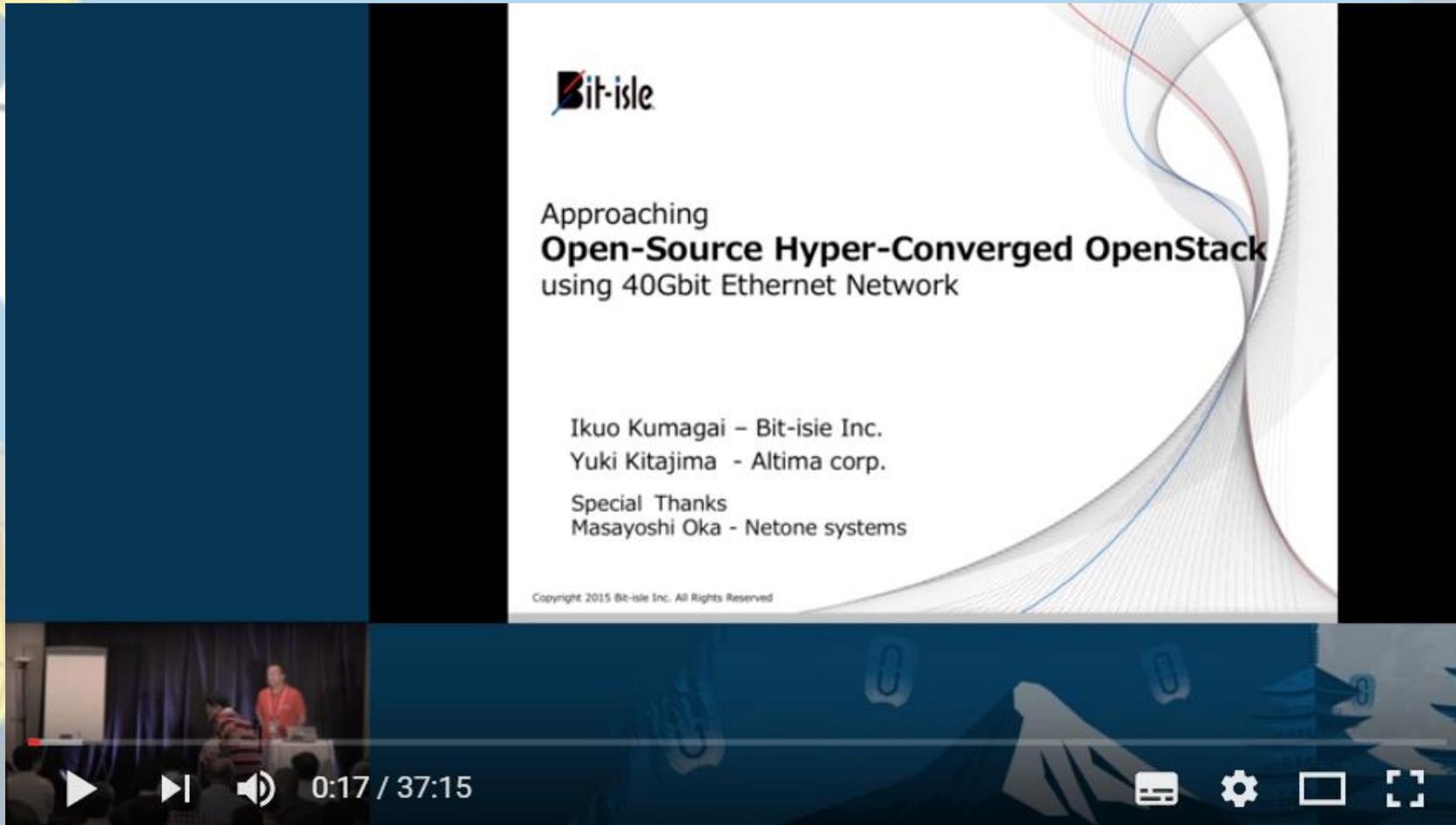
### 特徴

- 実績のある安心/高性能な推奨機器構成！  
 (NW機器、コンピュータード、コントロールノード、デプロイノード、マネジメントノード)
- 全てのコンピュータードには高速なPCIe SSDを搭載！  
 ストレージエリアはCephで構成され、高速、且つ、安心の冗長度。
- 多彩なコンポーネント！  
 (Nova, Glance, Horizon, keyStone, Neutron, Cinder, API, Ceph)
- OpenStackサポート内容 & スキルMAPを提示！  
 自立運用までのスケジュールを持つ事が出来ます。

フェーズ	目的	サポート範囲		お客様レベル	ビットアイル 提供内容	お客様 トレーニング	お客様への サポート
		ビットアイル	お客様				
1	フェーズ1 導入準備	(1)ハードウェア保守 (2) OpenStackサポート L1-L4	OpenStack対応 L4	お客様がOpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングを行う。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのトレーニング。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。
2	フェーズ2 運用準備	(1)ハードウェア保守 (2) OpenStackサポート L1-L2	OpenStack対応 L3-L4	お客様がOpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングを行う。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのトレーニング。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。
3	フェーズ3 運用	(1)ハードウェア保守	OpenStack対応 L1-L4	お客様がOpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングを行う。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのトレーニング。	OpenStackのインストール、アップグレード、バックアップ、復旧、監視、トラブルシューティングのサポート。

# Approaching Open-Sourced Hyper-Converged OpenStack

- OpenStack Summit Tokyo における発表内容のサマリー



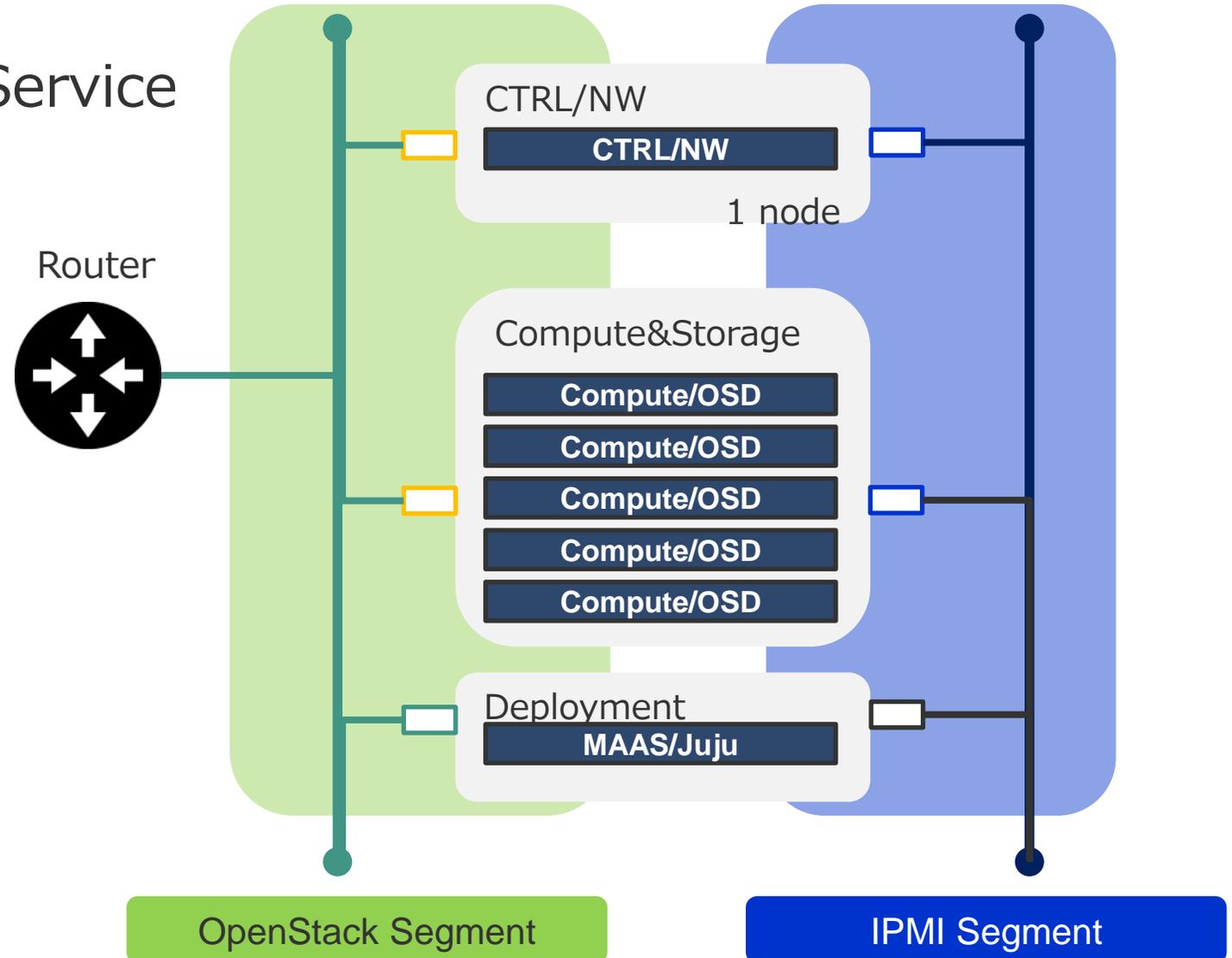
<https://www.youtube.com/watch?v=Q-231ZhVnFw>

# Hyper converged infrastructure – our needs

- Element of Hyper Converged
  - Structured as simple as possible
  - Deploying as rapid as possible
  - Managing integrated
  - As much as possible flexible scalability
- Our Concept
  - No Special Appliance
  - No product

# Basic Structure

- Network device
  - 1 × 40G Network for All Service
  - 1 × 1G Network for IPMI
- OpenStack Nodes
  - 1 Control and NW
  - 5 Compute and Storage
- Deployment Node
  - Juju /Maas Server



# Server & Storage

- Server(HP DL360 Gen9)

- Haswell 搭載サーバ



タイプ		8Core	12Core
CPU	ラインアップ	Xeon-E5-2630Lv3 (8C , 16T , 1.8GHz)	Xeon-E5-2690v3 (12C , 24T , 2.6GHz)
	標準搭載数(最大)	1 (2)	

- PCIe SSD

- ▶ Fusion-io ioDrive Duo 320GB

# Network Device

- HW selection
  - Adapter (NIC)

**Adapter**



**ConnectX<sup>®</sup> 3 PRO**

10 / 40 / 56GbE  
RDMA supported  
VXLAN offload supported



- Switch

**Switch**

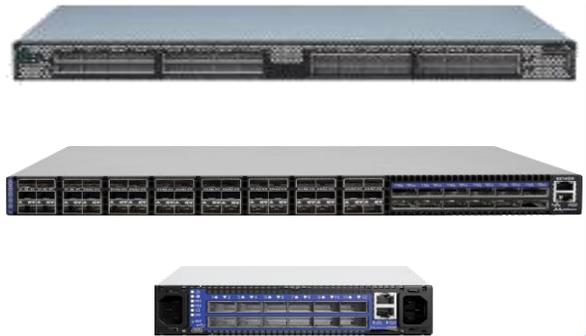


**SwitchX<sup>®</sup> 2**

36 ports x QSFP  
48 ports x SFP+ , 12 ports x QSFP  
12 ports x QSFP (48 ports x SFP+)  
\*Breakout Cable

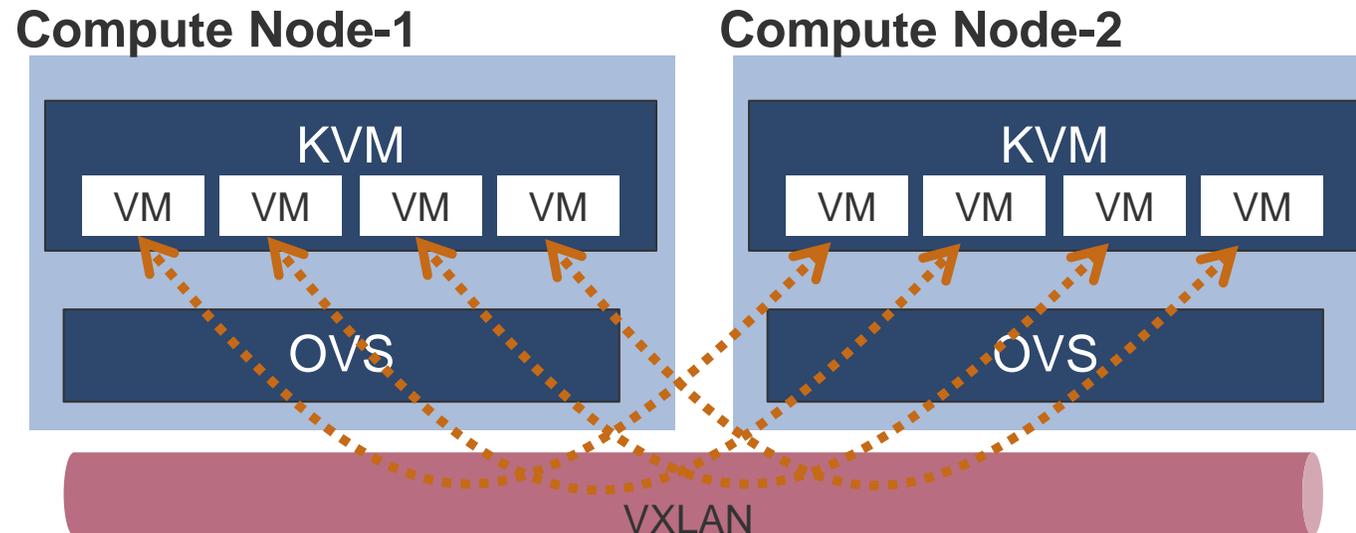
10 / 40 / 56GbE  
220ns Low Latency

**Best suits for SDS Network**



# Test Items (Network)

- VM to VM between physical nodes
- 1 – 16 VM per physical node
- Metering by iperf3 TCP & UDP

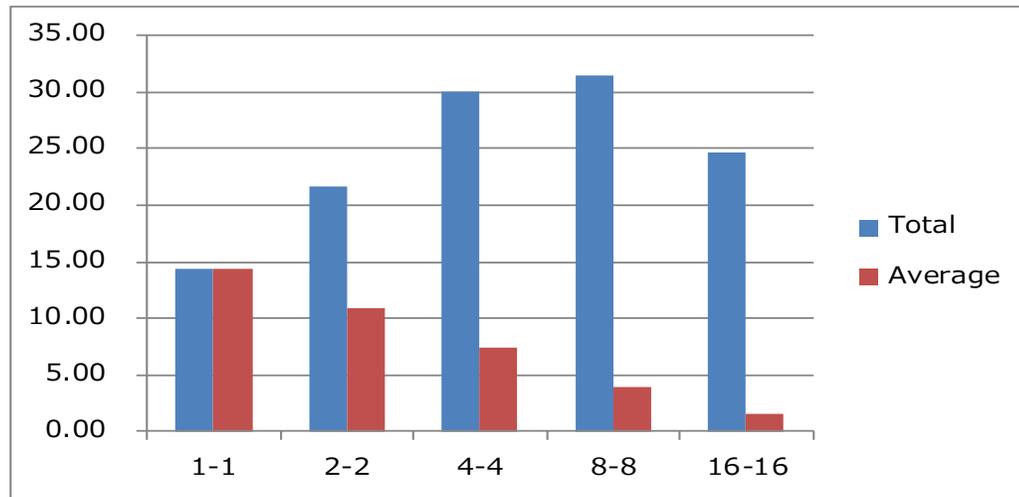


# VXLAN offload result(TCP Bandwidth)

- Compare VXLAN offload and normal

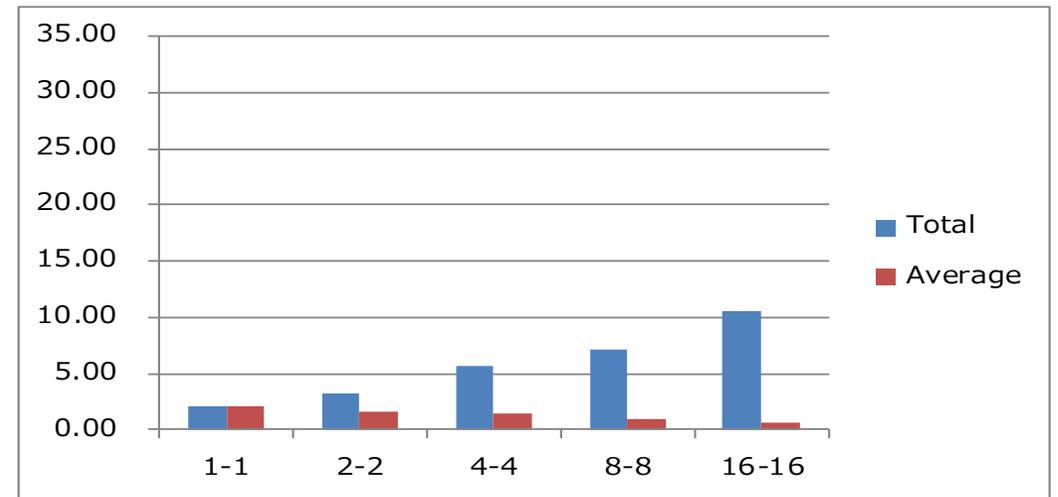
**Bandwidth(GBps) - VXLAN offload**

	1-1	2-2	4-4	8-8	16-16
Total	14.40	21.70	30.00	31.43	24.63
Average	14.40	10.85	7.50	3.93	1.54



**Bandwidth(Gbps) normal**

	1-1	2-2	4-4	8-8	16-16
Total	2.05	3.18	5.74	7.18	10.53
Average	2.05	1.59	1.43	0.90	0.66



# VXLAN offload

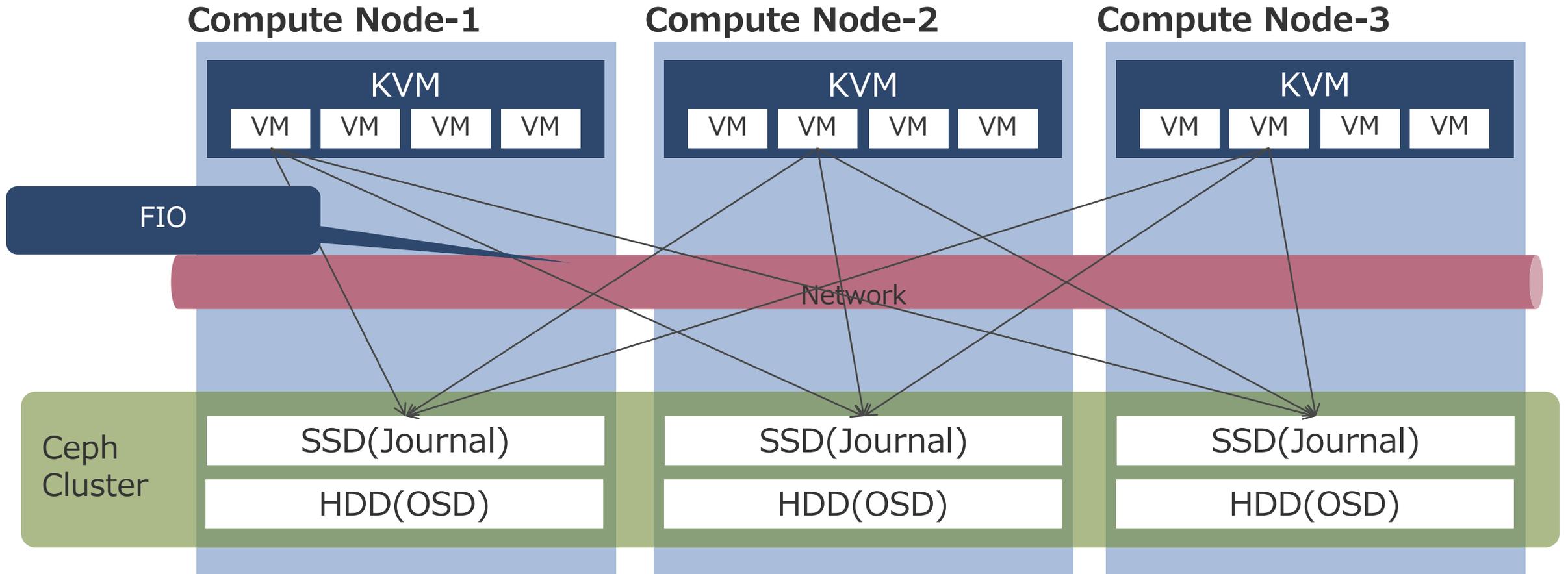
- HW selection



型番	MCX311A-XCCT	MCX312B-XCCT	MCX313A-BCCT	MCX314A-BCCT
Port	Single 10GbE	Dual 10GbE	Single /10/40/56GbE	Dual /10/40/56GbE
Port Type	SFP+	SFP+	QSFP	QSFP
Cable	Copper, Optical			
Host Bus	PCIe 3.0 x 8			
Features	VXLAN/NVGRE offload, RDMA, SR-IOV, etc			
OS	RHEL, SLES, Microsoft Windows Sever, FreeBSD, Ubuntu, VMWare ESXi			

# Test Items (IOPS)

- FIO (8k 100jobs )
  - 1 – 16 vm (1 ,2 or 4 VM per Host, Hosts count : 1 – 4 )



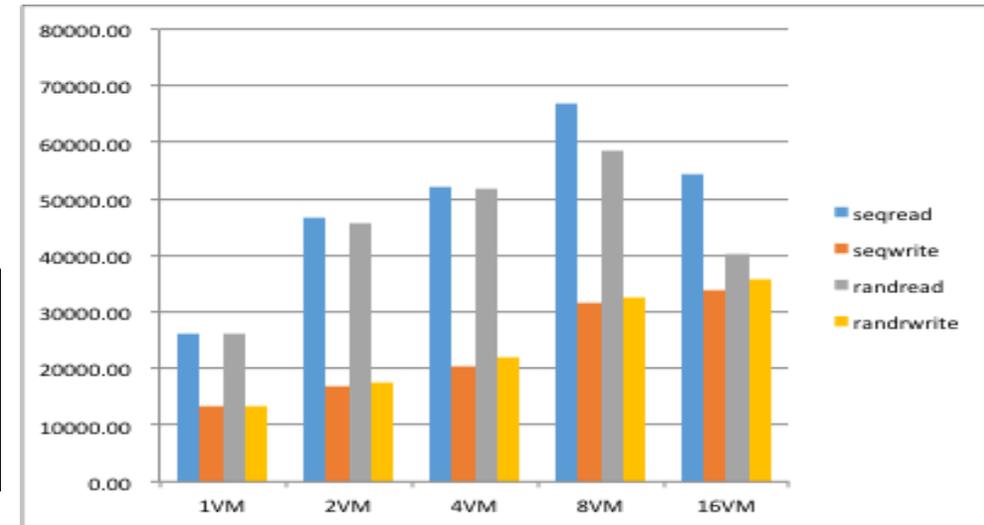
# Basic Performance of storage (IOPS)

- IOPS(8k MByte/sec)

- Total

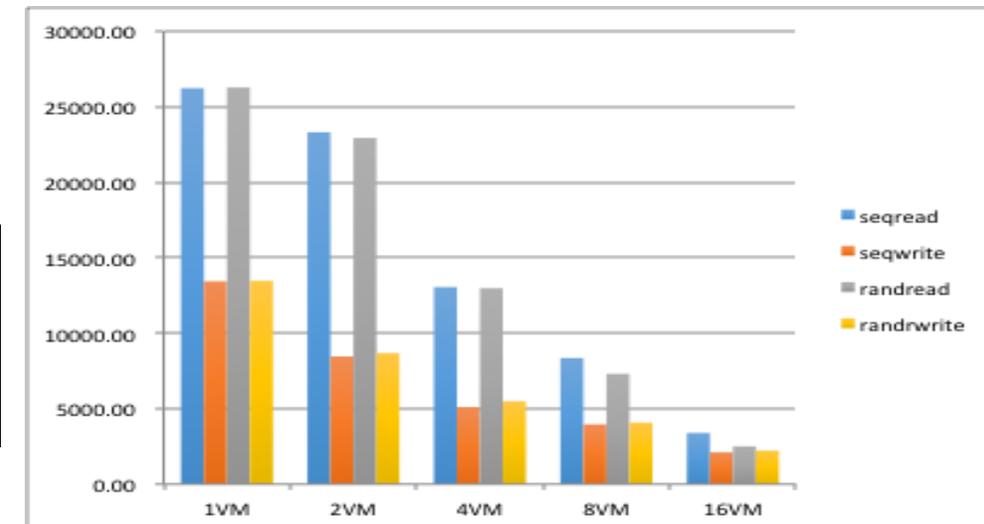
## Total iops

	1VM	2VM	4VM	8VM	16VM
seqread	26230.33	46629.33	52242.00	66908.33	54310.33
seqwrite	13437.67	16930.67	20440.33	31715.00	33810.33
randread	26277.33	45868.00	51938.67	58509.33	40184.00
randwrite	13471.67	17354.33	22001.67	32652.67	35640.33



## Average iops

	1VM	2VM	4VM	8VM	16VM
seqread	26230.33	23314.67	13060.50	8363.54	3394.40
seqwrite	13437.67	8465.33	5110.08	3964.38	2113.15
randread	26277.33	22934.00	12984.67	7313.67	2511.50
randwrite	13471.67	8677.17	5500.42	4081.58	2227.52



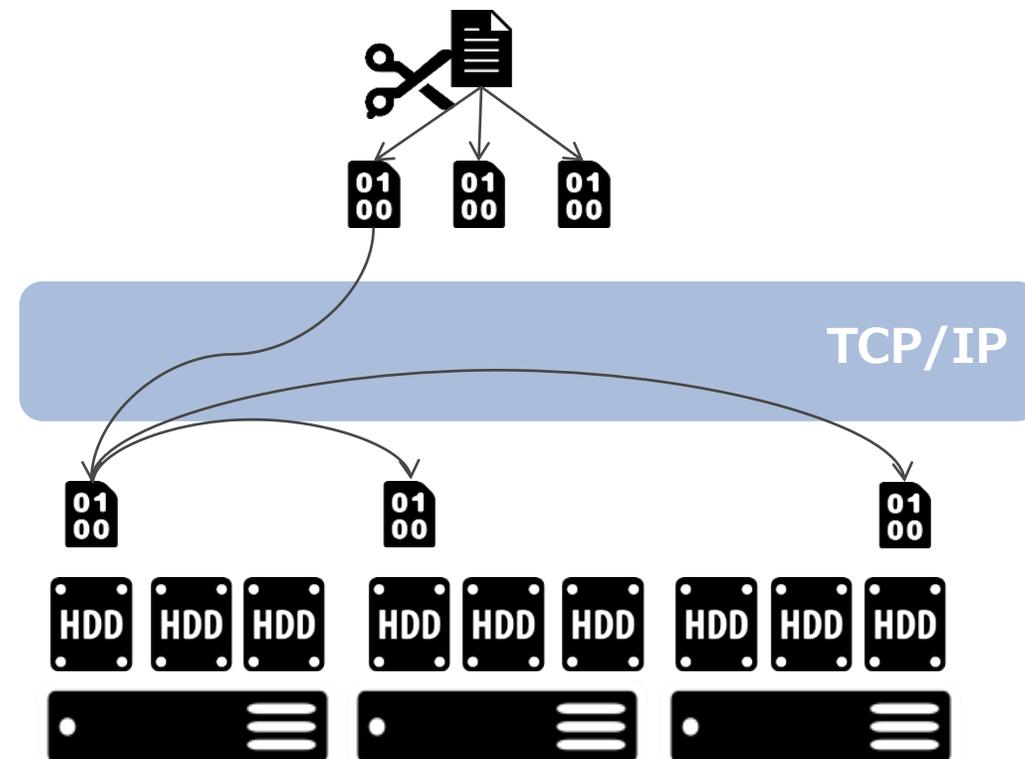
【FYI】 Fio parameters

bs=8k size=10M runtime=60 iodepth=32 numjobs=80 group\_reporting

# Ceph アーキテクチャとネットワーク

- Ceph
  - ▶ クライアントで分割したデータを各 OSD 上に分散/レプリケーション
  - ▶ データは全て TCP/IP 経由でやり取り
  - ▶ Networkのレイテンシが非常に重要

## Ceph dataflow



<http://blog.bit-isle.jp/bird/2014/06/10>

# Ceph RDMA

- RDMA network suits for the flash storage

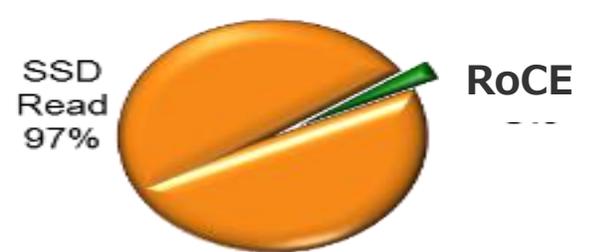
Component	Latency (usec)
InfiniBand	0.7
Ethernet RDMA (RoCE)	1.3
Ethernet TCP	6
Fibre Channel	20
SSD Read	25
Disk Read	6000

Total SSD Read Latency Over Fibre Channel



Total: 45usec

Total SSD Read Latency Over InfiniBand



Total: 25.7usec

- RDMA Advantage for Ceph

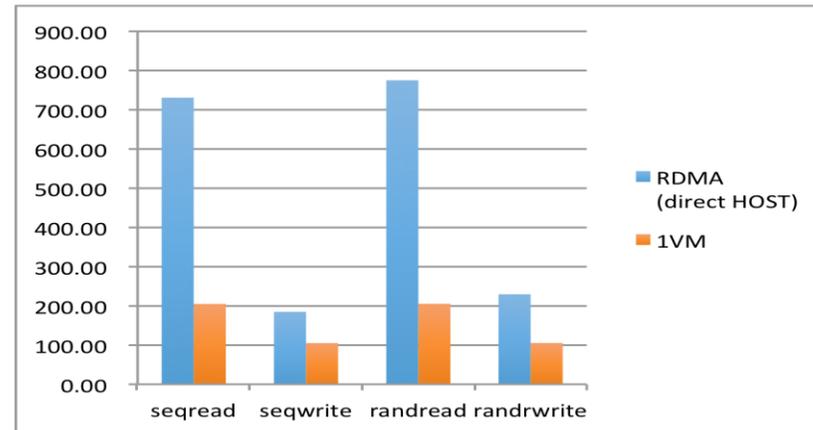
- Reduce CPU workload of Hypervisors for IO transaction
- Much faster IO for east-west traffic and fail-over(fail-back)
- Gets higher throughput and IOPS

# Ceph RDMA RESULT

- For reference purpose only
  - 3 nodes Ceph Cluster & Fio access direct rbd
  - Bandwidth

Total bandwidth(Mbyte/s)

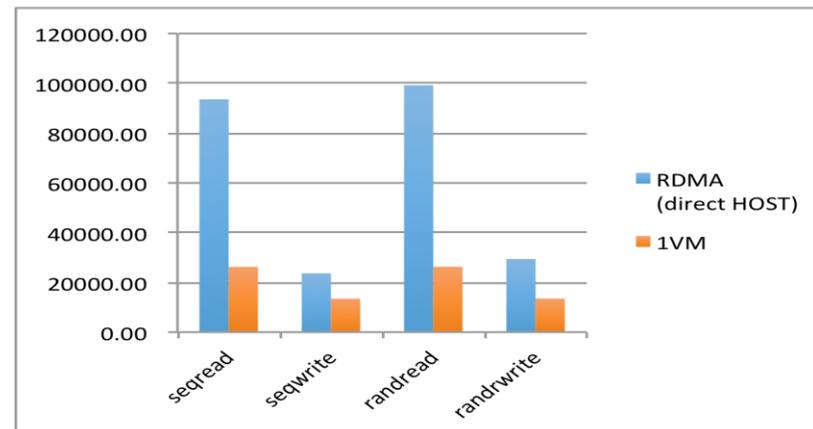
	RDMA (direct HOST)	1VM
seqread	730.76	204.93
seqwrite	184.67	104.99
randread	775.09	205.30
randwrite	229.53	105.25



IOPS

IOPS

	RDMA (direct HOST)	1VM
seqread	93537.00	26230.33
seqwrite	23638.00	13437.67
randread	99211.00	26277.33
randwrite	29379.00	13471.67

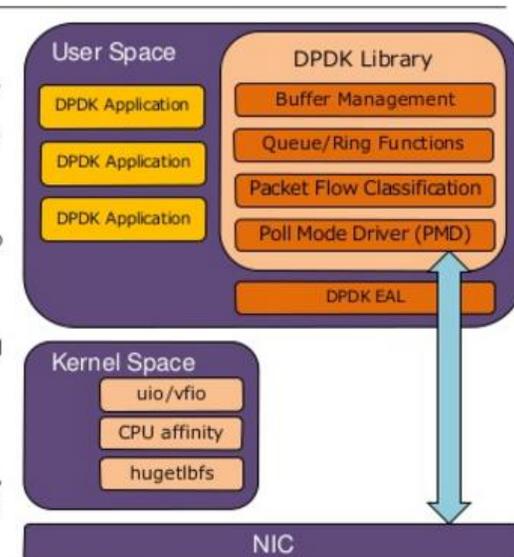


- 同環境においてInter Data Plane Development Kit の検証も実施



## 主なDPDKライブラリとドライバ

- バッファ・メモリ管理
  - 固定バッファを事前に用意することで、バッファの確保・解放する処理を減らす
  - メモリプール (HugePageメモリ) からメモリを割り当てる
- キューマネージャ
  - ロックレスなキューを使用することにより処理の待ち時間を防ぐ
- フロークラシフィケーション
  - ハッシュ計算を元にフロー毎にバケット処理を割り振る
- Poll Mode Driver (PMD)
  - 専用のデバイスドライバ
  - ユーザスペースより、占有したCPUコアに常駐して動作、NICにバケットが到達するのを常に監視



OpenStack Kilo with 6Wind VA High-Performance Networking Using DPDK -  
OpenStack最新情報セミナー 2016年3月にて 発表  
結果の詳細はSlideshareにて

<http://www.slideshare.net/VirtualTech-JP/openstack-kilo-with-6wind-va-highperformance-networking-using-dpdk-openstack-20163>

---

# ビットアイルのOpenStackサービス

# OpenStackトレーニングメニュー

ビットアイル・エクイニクス

## OpenStack トレーニング



### “OpenStackユーザーにフォーカスしたトレーニング”

#### 概要

#### OpenStackの利用方法と構築方法が一度に学べるトレーニング

本トレーニングプログラムでは、2011年の設立時よりOpenStackに関する技術的な実証実験や検証、社内向けテクニカルトレーニングなどを実施してきた「ビットアイル総合研究所」が作成したOpenStackへの理解や技術力を深めるためのカリキュラムを実施します。ビットアイルのデータセンター内に設置された、1人1台の物理サーバを利用しながら、OpenStack通りの利用方法と簡単な構築方法を身につける事が可能です。

#### 対象・提供内容・価格

##### 対象

- OpenStackの利用方法・概要を知りたいITエンジニアを抱える組織

##### 提供内容

1日目	<ul style="list-style-type: none"><li>● <b>OpenStack 概論</b> プロジェクトの概要と基本コンポーネントの説明</li><li>● <b>OpenStack 基本操作とハンズオン</b><ul style="list-style-type: none"><li>- 標準ダッシュボード「Horizon」を使用した操作のハンズオン</li><li>- CLI (Command Line Interface) 利用した操作のハンズオン</li></ul></li><li>● <b>OpenStack アーキテクチャの解説 (基本編)</b></li></ul>
2日目	<ul style="list-style-type: none"><li>● <b>OpenStack 環境の構築</b><ul style="list-style-type: none"><li>- RDO PackStackを使用した OpenStack 環境構築ハンズオン</li></ul></li><li>● <b>OpenStack アーキテクチャ解説 (詳細)</b></li></ul>

##### 価格

- **2日間コース . . . . . 60万円**  
※上記金額で5名まで受講可能です。  
(6人以上での受講についてはお問い合わせください。)  
※上記価格は、税抜価格です。

ビットアイル・エクイニクス

## OpenStack トレーニング&レンタル



### “OpenStackトレーニングと環境レンタルがセットに”

#### 概要

#### OpenStackトレーニングで構築した環境をそのまま利用できる

「OpenStack 構築・運用トレーニング & レンタル」では既に構築済みのOpenStackを使用するのではなく、未構築の環境からOpenStackが動作する環境となるまでお客様自身で構築するトレーニングプログラムです。さらに、本トレーニングで構築したOpenStack環境一式をそのままレンタルし、利用開始することができます。

#### 対象・提供内容・価格

##### 対象

- OpenStack基盤の理解・運用力を深めたいインフラエンジニアの方

##### 提供内容

ビットアイルのDCに貴社専用のラック・インターネット回線・レンタルサーバを用意し、この環境を使用して以下のトレーニングを受講していただきます。

1日目	<ul style="list-style-type: none"><li>● <b>OpenStack 基盤構築トレーニング</b><ul style="list-style-type: none"><li>- OpenStack 概要 及び アーキテクチャの解説</li></ul></li><li>● <b>OpenStack 環境の構築</b><ul style="list-style-type: none"><li>- RDO PackStackを使用した OpenStack 環境構築ハンズオン</li></ul></li></ul>
2日目	<ul style="list-style-type: none"><li>● <b>OpenStack 運用トレーニング</b><ul style="list-style-type: none"><li>- OpenStack 基盤トラブルシューティング概論&amp;実践</li></ul></li><li>● <b>OpenStack 基本操作とハンズオン</b><ul style="list-style-type: none"><li>- 標準ダッシュボード「Horizon」を使用した操作のハンズオン</li><li>- CLI(Command Line Interface)を利用した操作のハンズオン</li></ul></li></ul>

##### 価格

- **初期設計&構築費用 + OpenStack トレーニング費用 : 110万円**
- **OpenStack 環境レンタル費用 : 個別相談(※)**  
※ 構築し利用可能となる環境は弊社標準サービスの組み合わせにより構成し、ご利用されるサービス内容・NW機器・サーバ台数は事前に調整の上、実施いたします。  
※上記価格は、税抜価格です。

# OpenStack 導入コンサルティングサービス

“OpenStack 導入を検討されている全ての方に”

OpenStack導入のメリットは利用者の特性に合ったプライベートクラウドを構築できることです。貴社の求めるOpenStack 基盤を構成するためにどのような検討が必要か、どのような選択肢があるのか、一から情報収集して進めることは大変な労力を伴います。多様なOpenStack構築・運用経験を持つ弊社エンジニアのノウハウが、貴社のOpenStack導入をより確実なものにします。

OpenStackをはじめて導入する方向けに、弊社の経験豊富なエンジニアが計画立案から運用設計までのすべてもしくは一部の支援を行うサービスです。お客様の要求に沿った検討ポイントについて、具体例を交えながら検討・設計作業を支援します。例えば、以下のようなポイントについて、要点と選択肢の提示が可能です。

検討ポイント	<ul style="list-style-type: none"><li>• 物理機器構成の検討、選定</li><li>• ネットワーク構成の検討、設計</li><li>• OpenStackディストリビューションの選定・検討の支援</li><li>• OpenStack 基盤の導入コスト、導入計画作成</li></ul>
ご相談例	<ul style="list-style-type: none"><li>• 既存基盤のOpenstackへの移行計画を立てたいが、どのように進めたらいいのかわからない</li><li>• OpenStack運用時に考慮すべきことがわからないため、現在運用でトラブルを抱えている</li><li>• OpenStackをつかって革新的な新基盤を構築したいが、様々なアーキテクチャ選定のポイントがわからない</li><li>• OpenStackを使ってハイブリッドクラウドを実現したい。</li><li>• 古い機材を生かして、とにかく安価にOpenStackを始めたい。</li></ul>

## 価格

- 弊社HP (<http://www.bit-isle.jp/contact/index.html>) サービスに関するご相談・お問い合わせフォームよりお問い合わせください。



Thank You