



メラノックス最新アップデート

メラノックステクノロジーズジャパン株式会社

December 15, 2017 | PC Cluster Symposium

 **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：全世界で約2,600名
- **財務状況**
 - 2015年度売上 : \$658.1M
 - 2016年度上半期 : \$411.6M
 - 2016年度第3四半期 : \$224.2M
 - 2016年度第4四半期見込み : \$222 - 228M
 - Cash + Investment : \$292.4M (2016年9月末時点)





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 (VPI)

IC



アダプタカード



NPU、マルチコア



スイッチ/ゲートウェイ



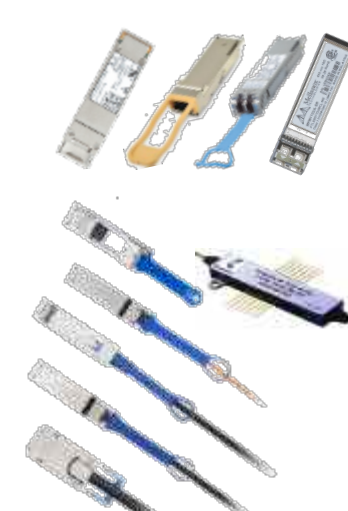
ソフトウェア



Metro / WAN



ケーブル/モジュール



ハイパフォーマンスインテリジェンス100Gb/sインターコネクソソリューション



アダプタ

ConnectX[®] 4

100Gb/s対応アダプタ、レイテンシー：700ナノ秒
1億5千万メッセージ/秒の処理能力
(10 / 25 / 40 / 50 / 56 / 100Gb/s)

ConnectX[®] 4 Lx

Innova
Smarter Adapter. Greater options.

スイッチ

Switch IB[™] 2

EDR (100Gb/s) ポート x 36、レイテンシー：90ナノ秒以下
スループット：7.2Tb/s
70億メッセージ/秒の処理能力 (ポートあたり1億9500万メッセージ)

SHARP

スイッチ

Spectrum[™]

100GbEポート x 32、25/50GbE x 64ポート
(10 / 25 / 40 / 50 / 100GbE)
スループット：6.4Tb/s

インターコネクソ

LinkX[™]

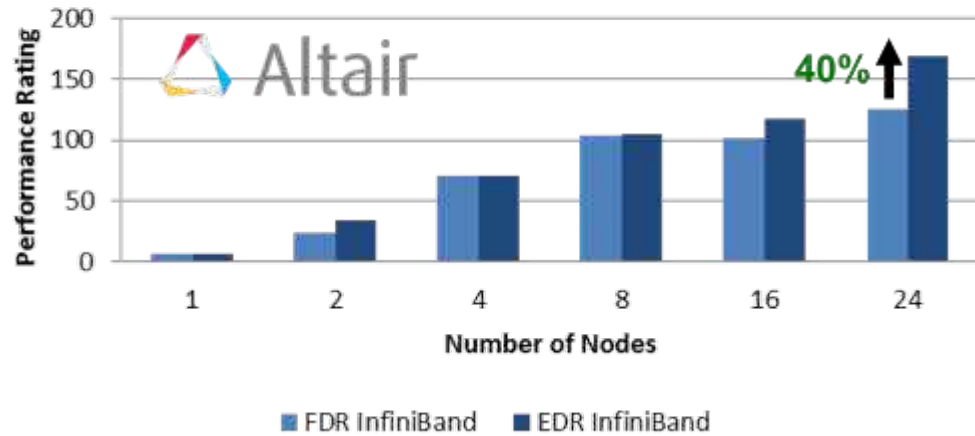
トランシーバ
アクティブオプティカル、銅ケーブル
(10 / 25 / 40 / 50 / 56 / 100Gb/s)

VCSEL、シリコンフォトニクス、銅

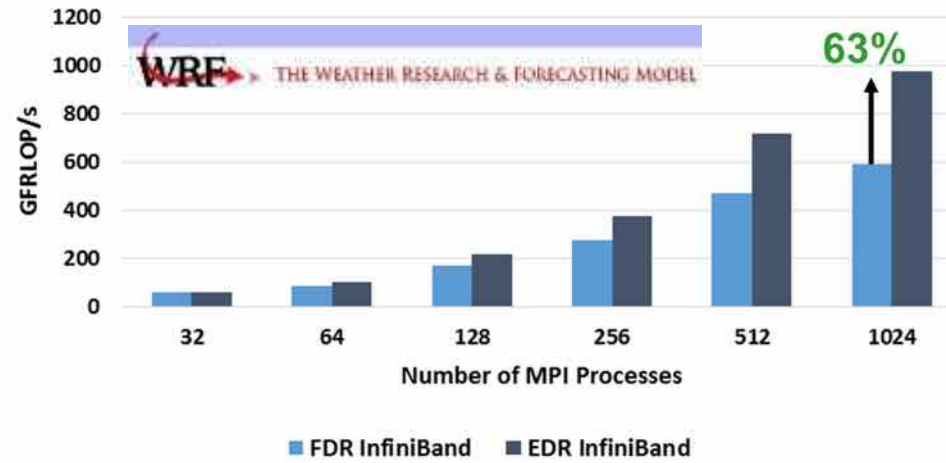
100Gb/s EDR InfiniBandのパフォーマンス (FDR比28~80%の向上)



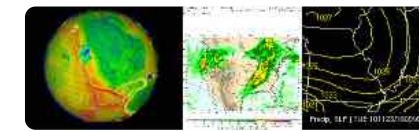
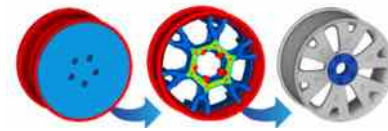
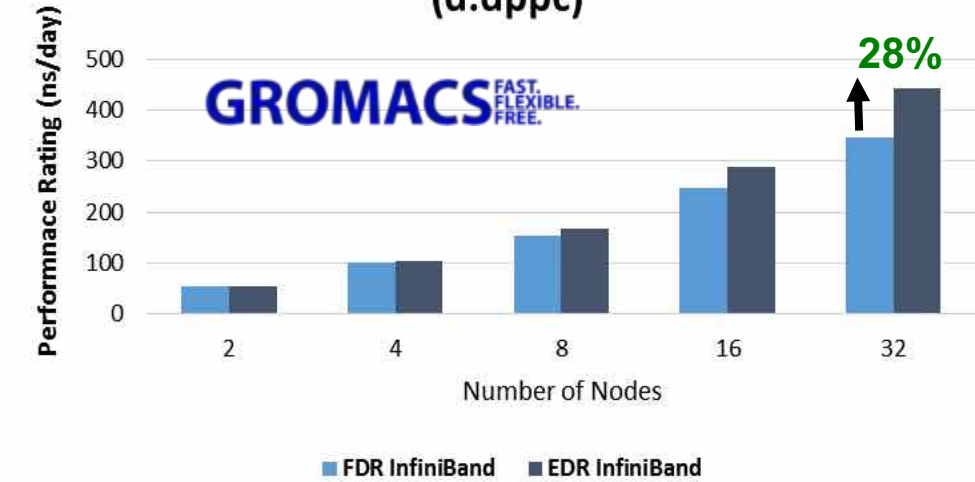
OptiStruct Performance (Engine_Assy.fem)



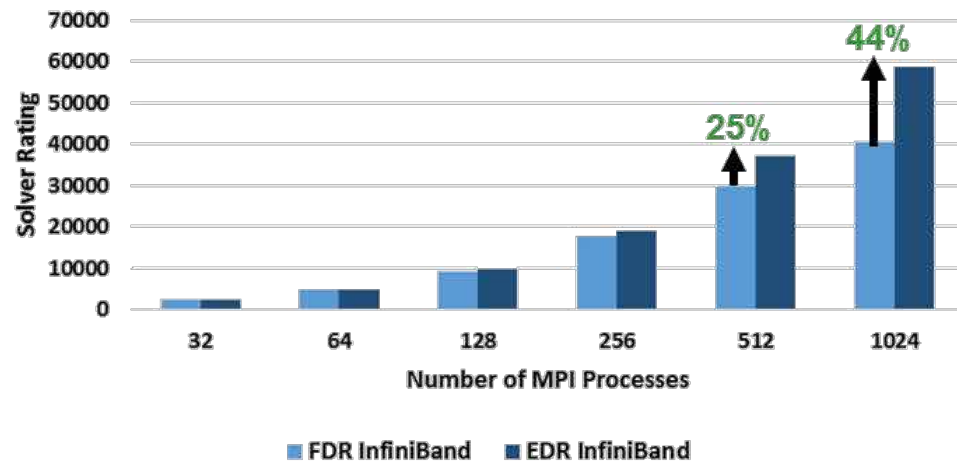
WRF Performance (conus12km)



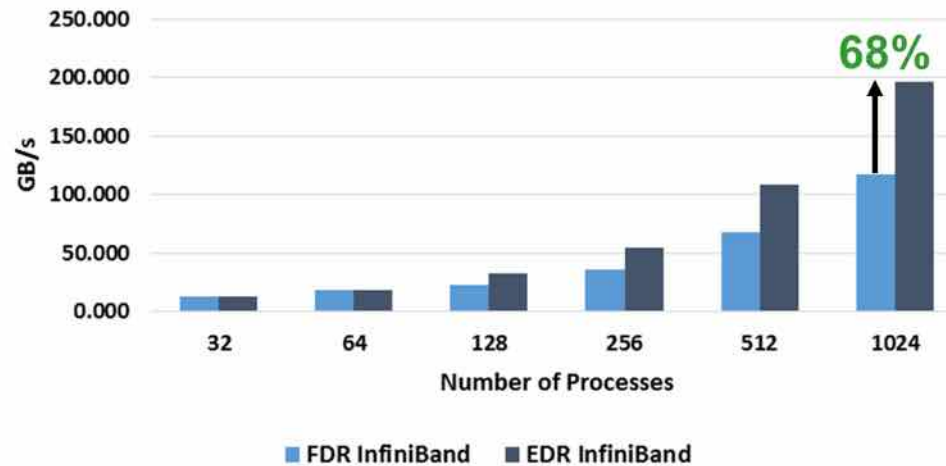
GROMACS Performance (d.dppc)



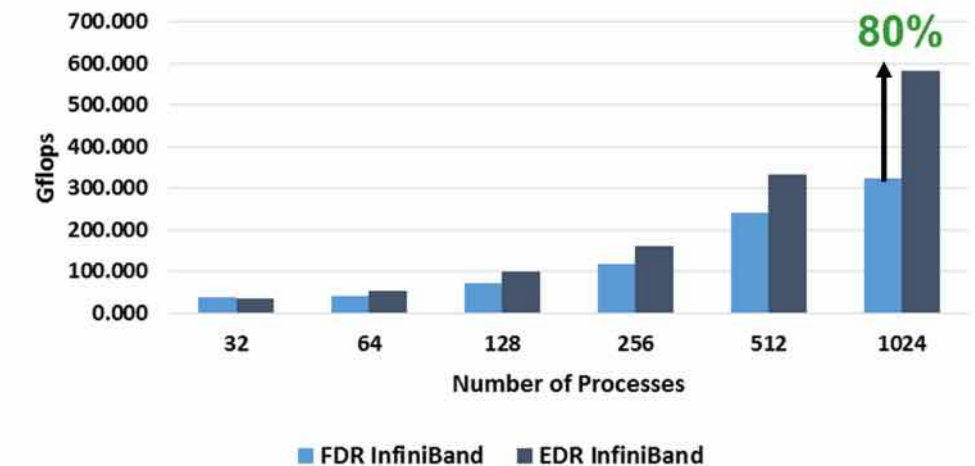
ANSYS Fluent 16.0 Performance (sedan_4m)



HPCC Performance (PTRANS_GB)



HPCC Performance (MPIFFT)



HDR InfiniBand

High can Accelerate



- **Higher** performance
- **Higher** efficiency
- **Higher** scalability
- **Higher** speed

世界初
200Gb/sスイッチ



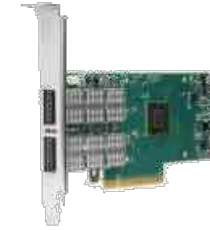
世界初
200Gb/sアダプタ

ハイパフォーマンス200Gb/sインターコネクトソリューション

アダプタ

ConnectX[®] 6

200Gb/s対応アダプタ、レイテンシー：600ナノ秒
2億メッセージ/秒の処理能力
(10 / 25 / 40 / 50 / 56 / 100 / 200 Gb/s)



スイッチ

Quantum
The smartest switch, became smarter

HDR (200Gb/s) ポート x 40、レイテンシー：90ナノ秒以下
HDR100 (100Gb/s) ポート x 80
スループット：16Tb/s、ルータ

SHARP



インターコネクト

LinkX[™]

トランシーバ
アクティブオプティカル、銅ケーブル
(10 / 25 / 40 / 50 / 56 / 100 / 200Gb/s)



VCSEL、シリコンフォトニクス、銅

ソフトウェア

HPC-X[™]

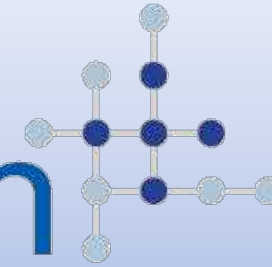
MPI、SHMEM/PGAS、UPC
コマーシャル、オープンソースアプリケーション
ハードウェアアクセラレーションの活用





Quantum

The smartest switch, became smarter



- 200Gb/s (HDR) ポート x 40
- 100Gb/s (HDR100) ポート x 80
- スイッチシステム
 - 1U 40ポート
 - 200ポートシャーシ
 - 800ポートシャーシ
- ポートあたり3億9千万パケット/秒
- 90ナノ秒以下のレイテンシー
- 20メガバイトのシェアードバッファ
- 広バンド幅アグリゲーション
- 拡張InfiniBandルータ

リープする準備はできていますか？



パフォーマンス



スケール



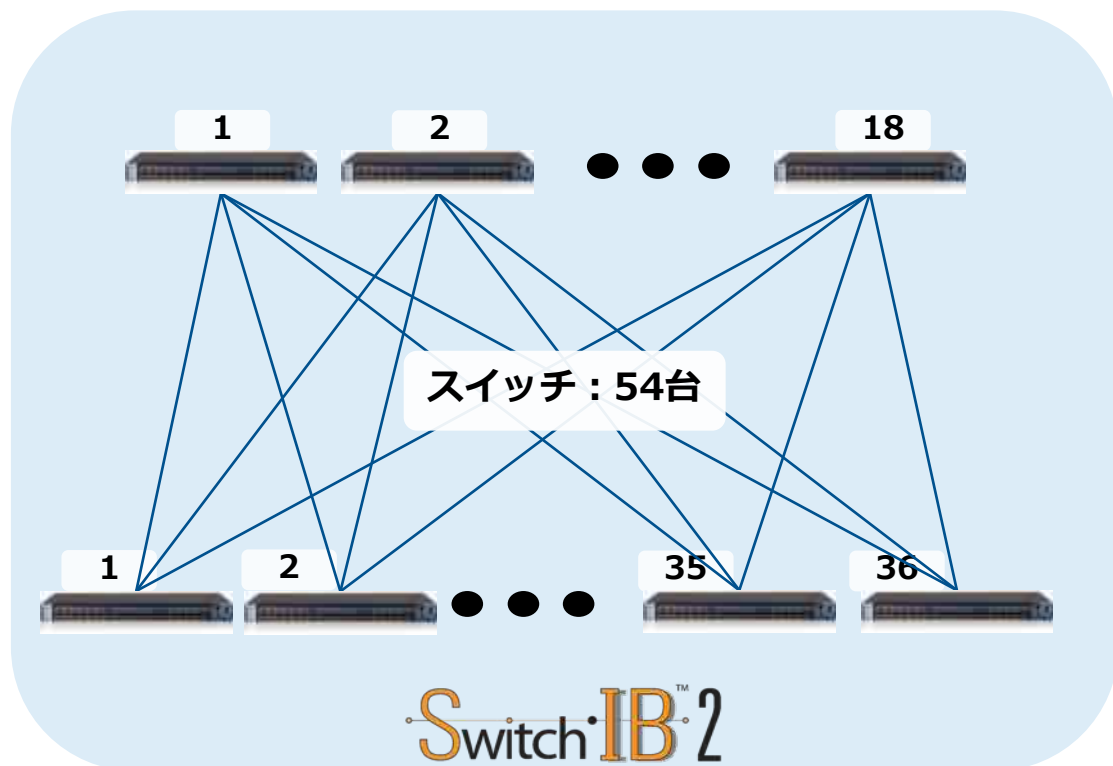
テレメトリ



機械学習

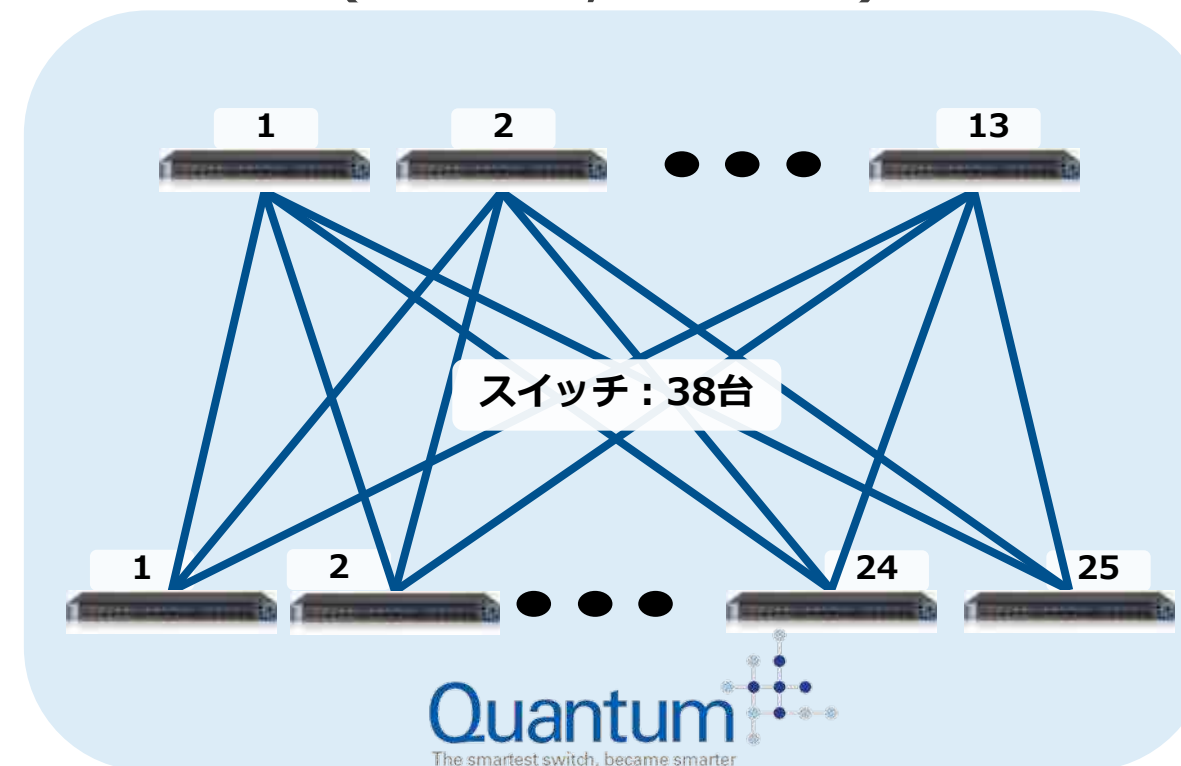
HDRアグリゲーション – ConnectX-5を使用したEDRクラスタ

648ノード EDRクラスタ 1:1



ConnectX-5 ConnectX-5 ConnectX-5

648ノード HDRアグリゲーション 2:1 (L1: EDR, L2: HDR)

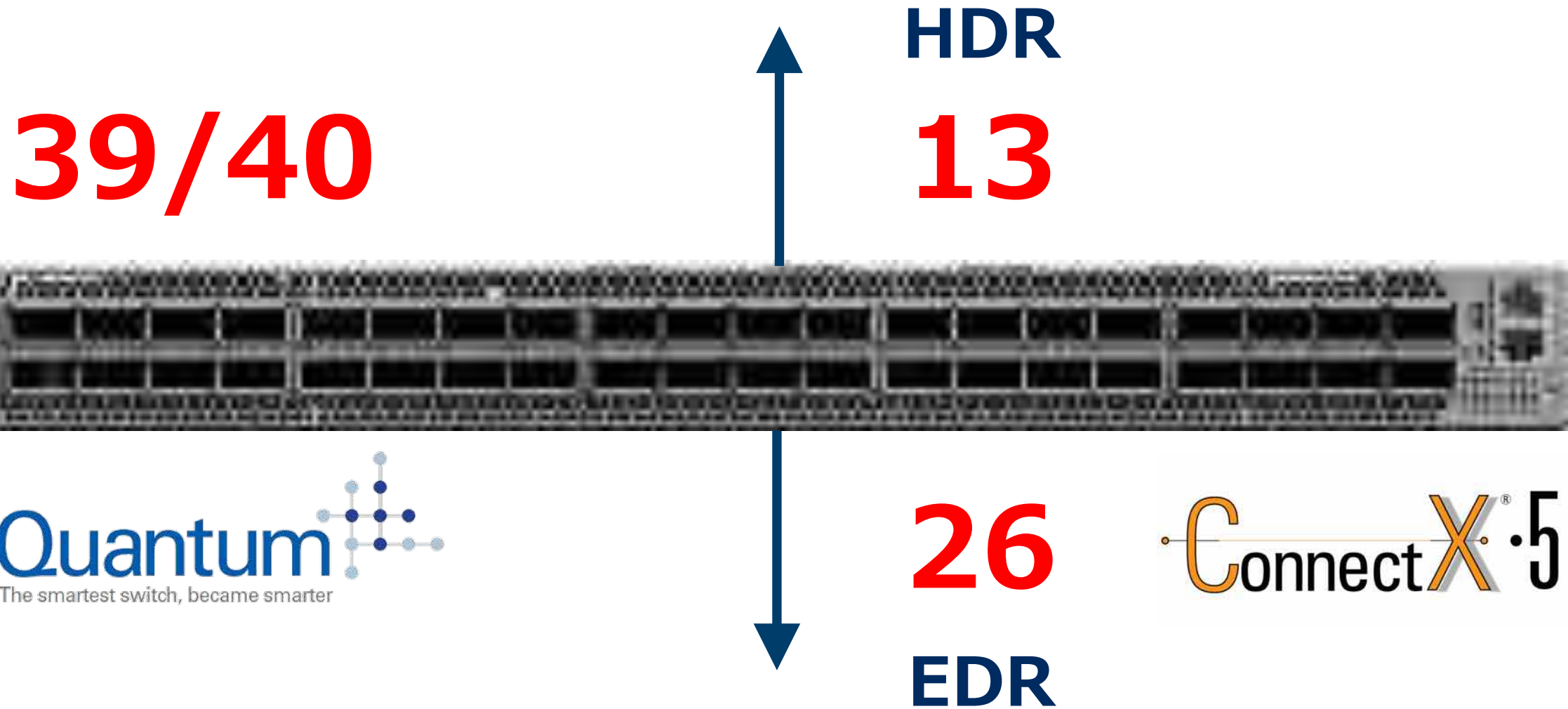


ConnectX-5 ConnectX-5 ConnectX-5

HDRアグリゲーションによりコスト・パフォーマンスが向上

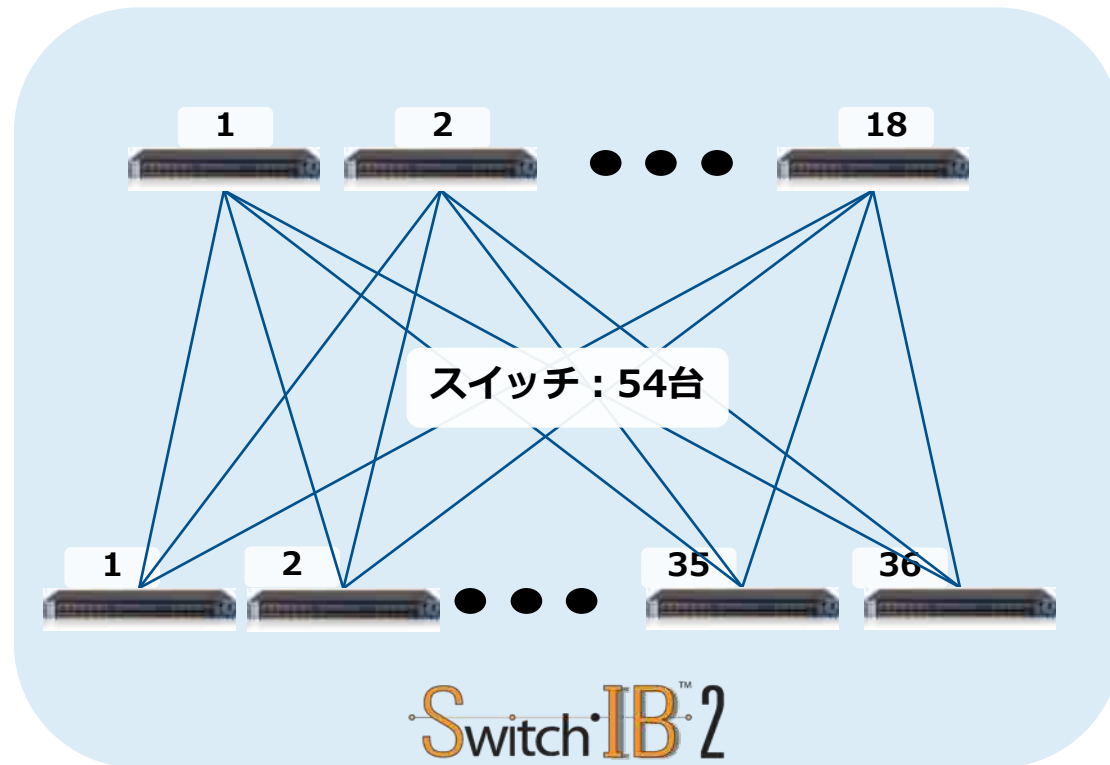
- 将来の保証
- 省スペース (スイッチ・ケーブルの削減)
- より低いレイテンシーの実現
- コンジェスチョンの削減

HDRアグリゲーション – ConnectX-5を使用したEDRクラスタ



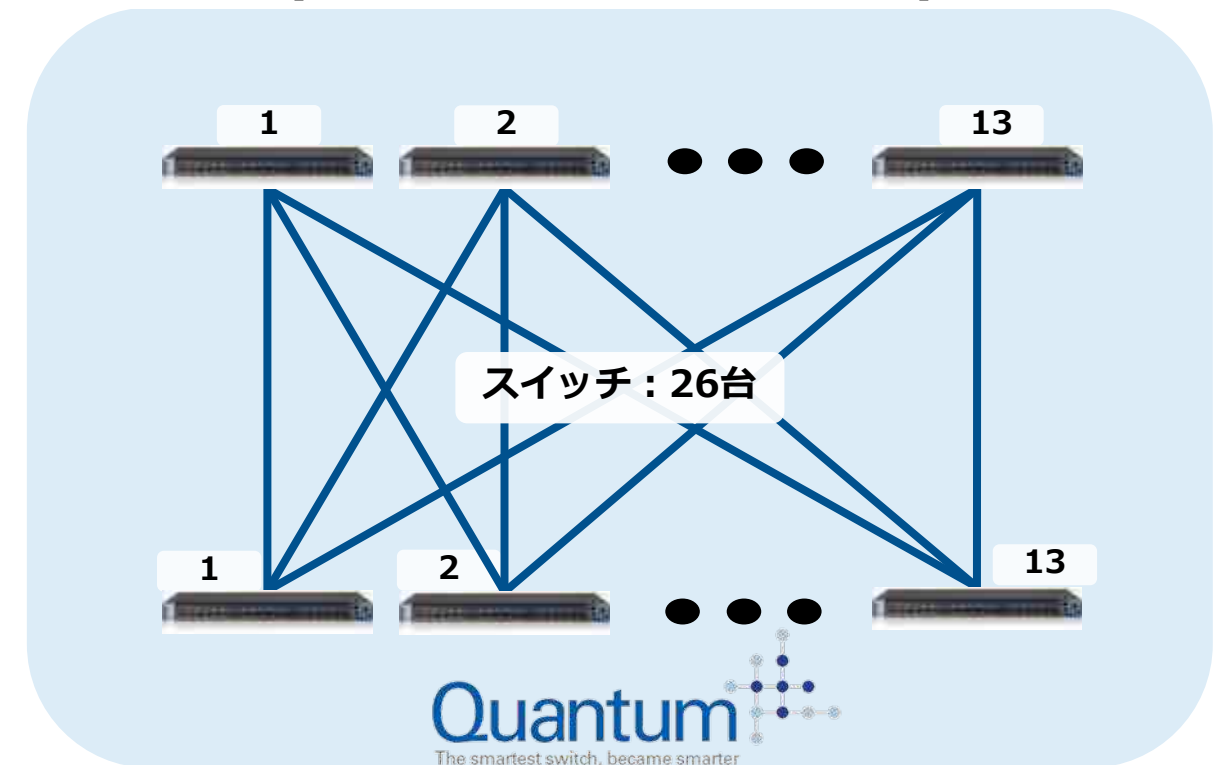
HDRアグリゲーション – ConnectX-6を使用した100Gb/sクラスター

648ノード EDRクラスター 1:1



ConnectX-5 ConnectX-5 ConnectX-5

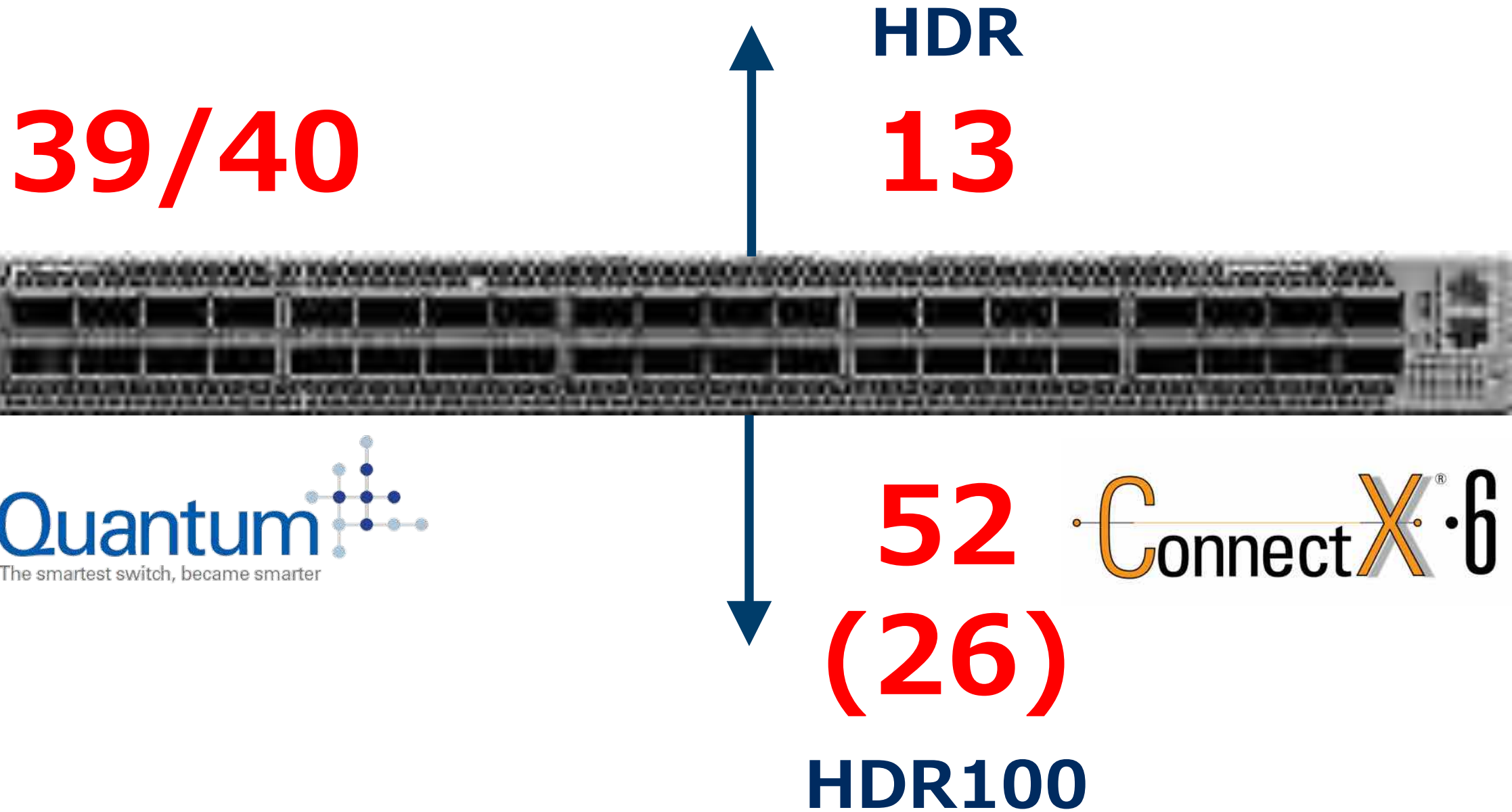
648ノード HDR100アグリゲーション 2:1
(L1: HDR100, L2: HDR)



ConnectX-6 ConnectX-6 ConnectX-6

HDRアグリゲーションによりコスト・パフォーマンスが向上

- 将来の保証
- 省スペース (スイッチ・ケーブルの削減)
- より低いレイテンシーの実現
- コンジェスチョンの削減

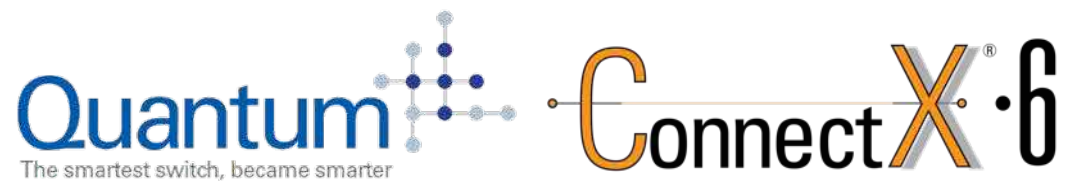
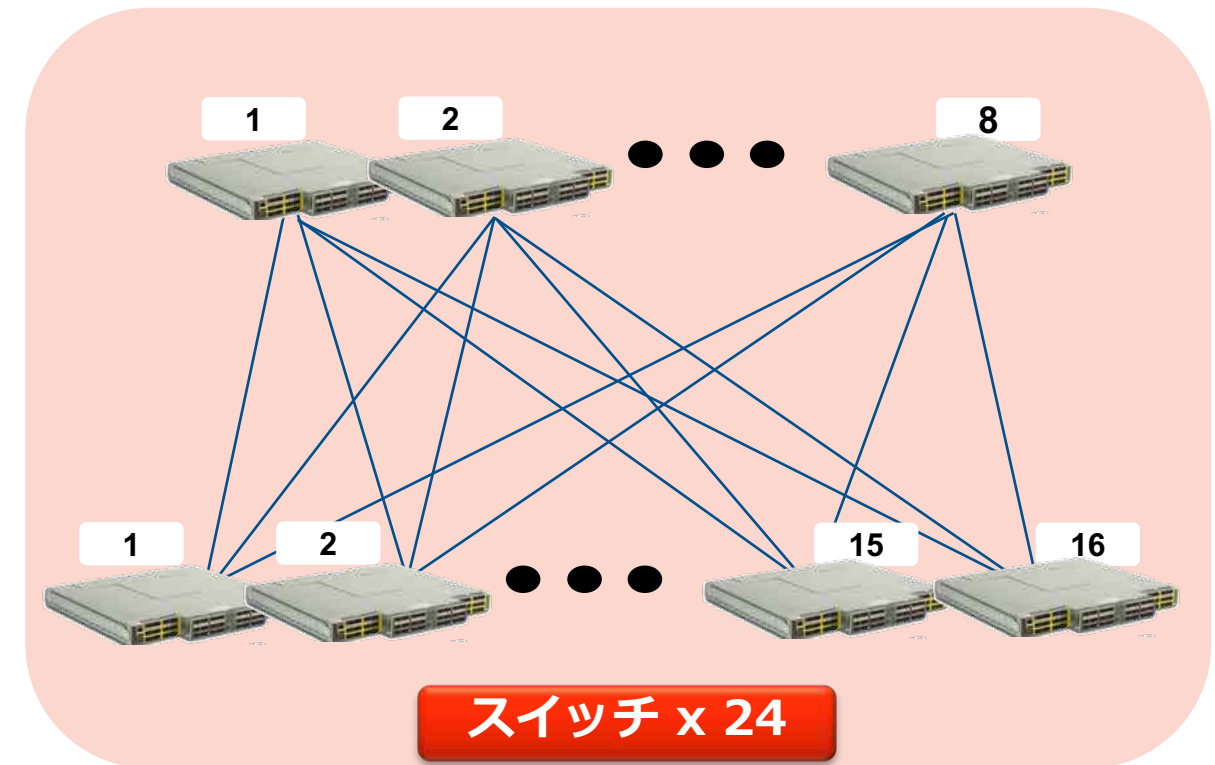
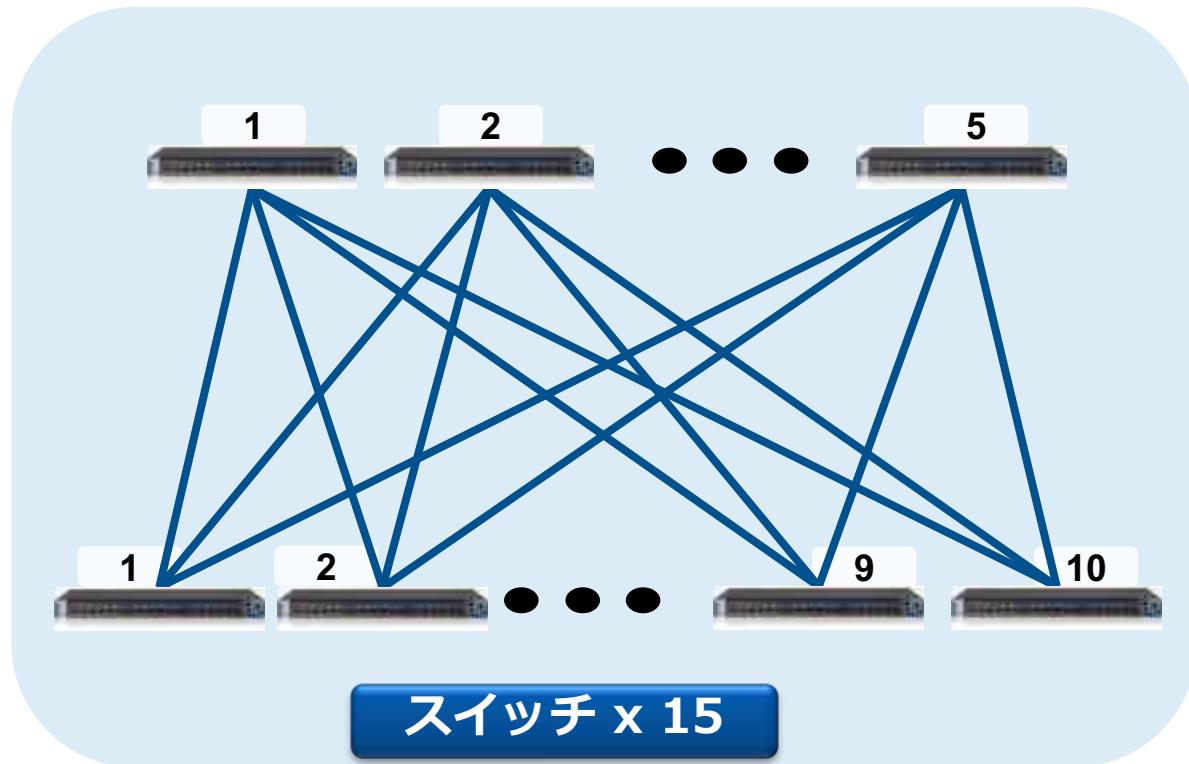


Quantum HDR100 – 最小TCOを実現する最もスケーラブルなスイッチ



400ノード100Gb/s InfiniBandプラットフォーム

384ノード100Gb/sプラットフォーム



他社ソリューション



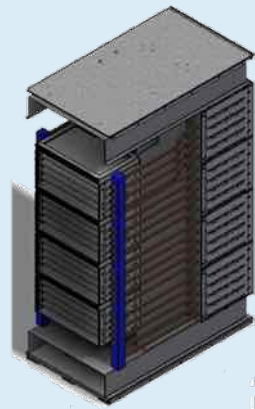
1.6倍のスイッチ台数、2倍のケーブル本数削減

Quantum HDR100 – 最小TCOを実現する最もスケーラブルなスイッチ

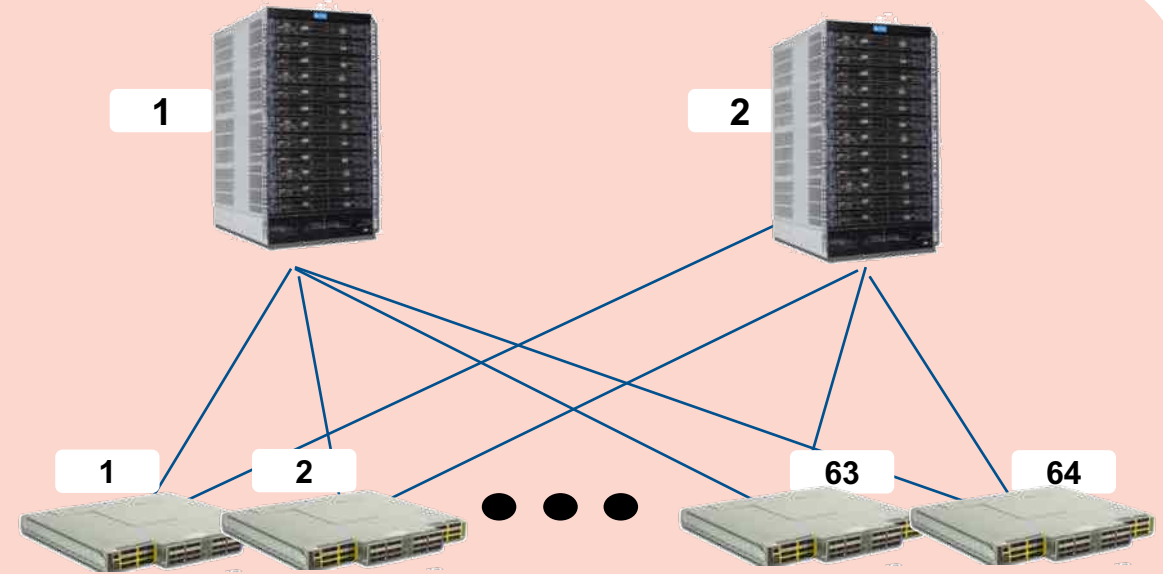


1536ノード100Gb/s InfiniBandプラットフォーム

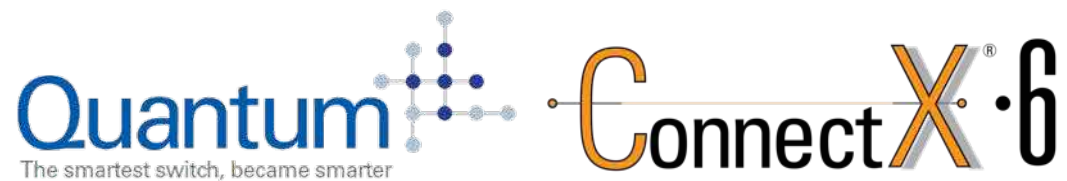
1536ノード100Gb/sプラットフォーム



ケーブル x 768、消費電力 : 22KW



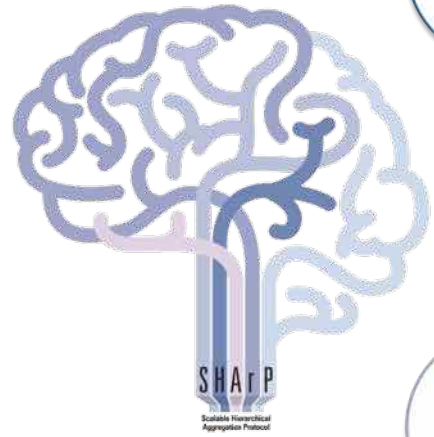
ケーブル x 3072、消費電力 : 49KW



他社ソリューション



3倍のスペース、4倍のケーブル本数、2倍の消費電力削減



SHArPにより、ネットワーク内でMPI処理の管理・実行が可能に

Quantumでは、スイッチネットワークをコプロセッサとして使用可能に

HPC、機械学習通信のパフォーマンスが10倍向上



MPIコレクティブランタイムの短縮



CPU効率性の向上

もっともスマートなスイッチがよりスマートに

ラージデータベクタのワイヤスピードでのアグリゲーション

機械学習アプリケーション処理に最適なスイッチ

Reduce処理をCPUからオフロード

HPCと機械学習のインターコネクトに対するニーズは同じ

パフォーマンス

スループット : 200Gb/s (InfiniBand、Ethernet)
レイテンシー : 600ナノ秒 (エンドトゥエンド)
2億メッセージ/秒

スマート

ネットワーク内コンピューティング (コレクティブ、マッチング)
ネットワーク内メモリ
ストレージ (NVMe) ・ セキュリティ ・ ネットワークオフロード

プラットフォーム

PCI Express 3.0 / 4.0
内蔵PCI Expressスイッチ、Multi-Host
高度なアダプティブルレーティング



RDMA

スタンダード

ネットワーク
オフロード

ネットワーク内
コンピューティング

2~10倍の
ネットワーク
パフォーマンス



少ないCPU
使用率

より良いアプリケーション
パフォーマンス
(20~50%)

より良い価格性能比

データセンターにおいて最高の投資利益率を実現

テクノロジーロードマップ – 常に一世代先を行くメラノックス



Mellanox → 20G → 40G → 56G → 100G → 200G → 400G

テラスケール

ペタスケール

エクサスケール

3rd



TOP500 2003
Virginia Tech (Apple)

1st



“Roadrunner”
Mellanox Connected



OAK RIDGE
National Laboratory
“Summit” System

Lawrence Livermore
National Laboratory
“Sierra” System

2000

2005

2010

2015

2020

CCIX、 Gen-Z、 OpenCAPI

■ 目的

- 既存のバス規格 (PCI Express) との互換性を維持したうえでより優れた機能を実現
- CPU、アクセラレータ、インターコネクト間でハイパフォーマンスで一貫した接続の実現
- 新しい低レイテンシーのストレージクラスメモリメディアの完璧なサポートの実現
- オープンスタンダードでよりハイレベルな技術革新の実現

■ CCIX (Cache Coherent Interconnect for Accelerators) Consortium

- 創設メンバー
 - AMD, ARM, Huawei, IBM, Mellanox Technologies, Qualcomm, Xilinx
- 新メンバー（10月10日発表）
 - Amphenol, Arteris, Avery Design Systems, Broadcom, Bull Atos Technologies, Cadence Design Systems, Cavium, Integrated Device Technology, Keysight Technologies, Micron Technology, Netspeed Systems, Red Hat, Synopsys, Teledyne LeCroy, TSMC Technology

■ Gen-Z Consortium

- アライアンスメンバー
 - AMD, ARM, Broadcom, Cavium, Cray, Dell EMC, HPE, Huawei, IDT, Mellanox Technologies, Micron Technology, Samsung, Seagate, SK hynix, Xilinx

■ OpenCAPI Consortium

- 創設・ボードメンバー
 - AMD, Google, IBM, Mellanox Technologies, Micron

Mellanox Innova Adapters

メラノックスアダプタ製品ロードマップ



ConnectX-3

56Gb/s VPI
(IB/ EN)
PCIe Gen3 x8



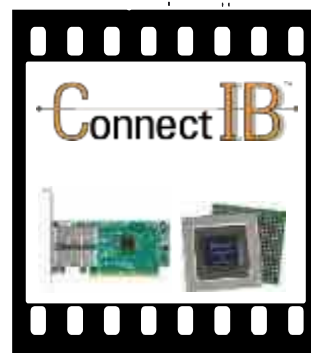
ConnectX-3 Pro

56Gb/s VPI
(IB/ EN)
PCIe Gen3 x8



Connect-IB

56Gb/s InfiniBand
PCIe Gen3 x16



ConnectX-4

100Gb/s VPI
(IB/EN)
PCIe Gen3 x16



ConnectX-4 Lx

25/50Gb/s EN
PCIe Gen3 x8



ConnectX-5

100Gb/s VPI
(IB/EN)
PCIe Gen4 x16



2011

2014

2015

2016

Innova = メラノックス初のFPGA搭載アダプタ製品





ビッグデータ

- 分析
- 圧縮
- 重複排除



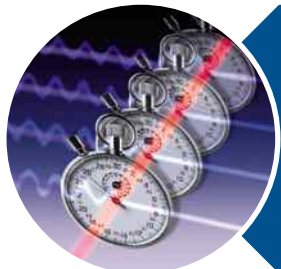
セキュリティ

- 暗号化/解読
- 認証



HPC

- 気象モデリング
- 生命情報科学



リアルタイムデータ

- 超高速取引 (HFT)
- 画像処理

- IPsec / TLSオフロード
- 40ギガビットEthernet (50ギガビットEthernetもサポート予定)
- 32Kアクティブコネクション
- 1M同時セッション
- 60Mパケット/秒
- セキュリティアルゴリズム (サポート予定のものも含む)
 - 暗号 AES-GCM 128/256, AES-CBC 128/256
 - 認証 SHA-1, SHA-2
 - キー交換 RSA
- ユーザアプリケーションとのネイティブな統合
 - IPsec
 - Kernel Mode (1st phase)
 - Kernel Bypass Mode (2nd phase)
 - TLS
 - OpenSSL



Innova IPsec
Smarter Adapter. Greater Security.



■ タイムトゥマーケット

- ・ 現行のNIC製品に新たな機能の追加が可能

■ パフォーマンス

- ・ データ処理ロジックをFPGAへオフロード
- ・ ConnectX-4 Lxの持つ優れた機能を利用可能

■ 柔軟性

- ・ パートナーや顧客によって開発されるネットワークアプリケーションに対応する可能なオープンなプラットフォーム



Mellanox BlueField

メラノックス初のSOC (システムオンチップ) 製品



ConnectX-4
EDR - 100Gb/s VPI
PCIe Gen3 x16

ConnectX-4 Lx
25/50 Gb/s Ethernet
PCIe Gen3 x8

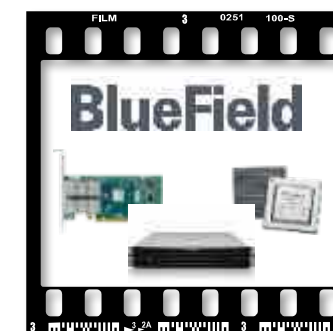
ConnectX-5
EDR - 100Gb/s VPI
PCIe Gen3/4 x16

Adapters



BlueField
ConnectX-5 + 16 x ARM
PCIe Gen3/4 x16

SoC



2014

2015

2016

2017

すべての製品がソフトウェア互換



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



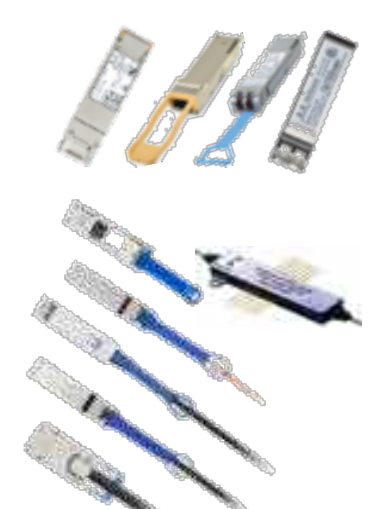
ソフトウェア



Metro / WAN



ケーブル/モジュール

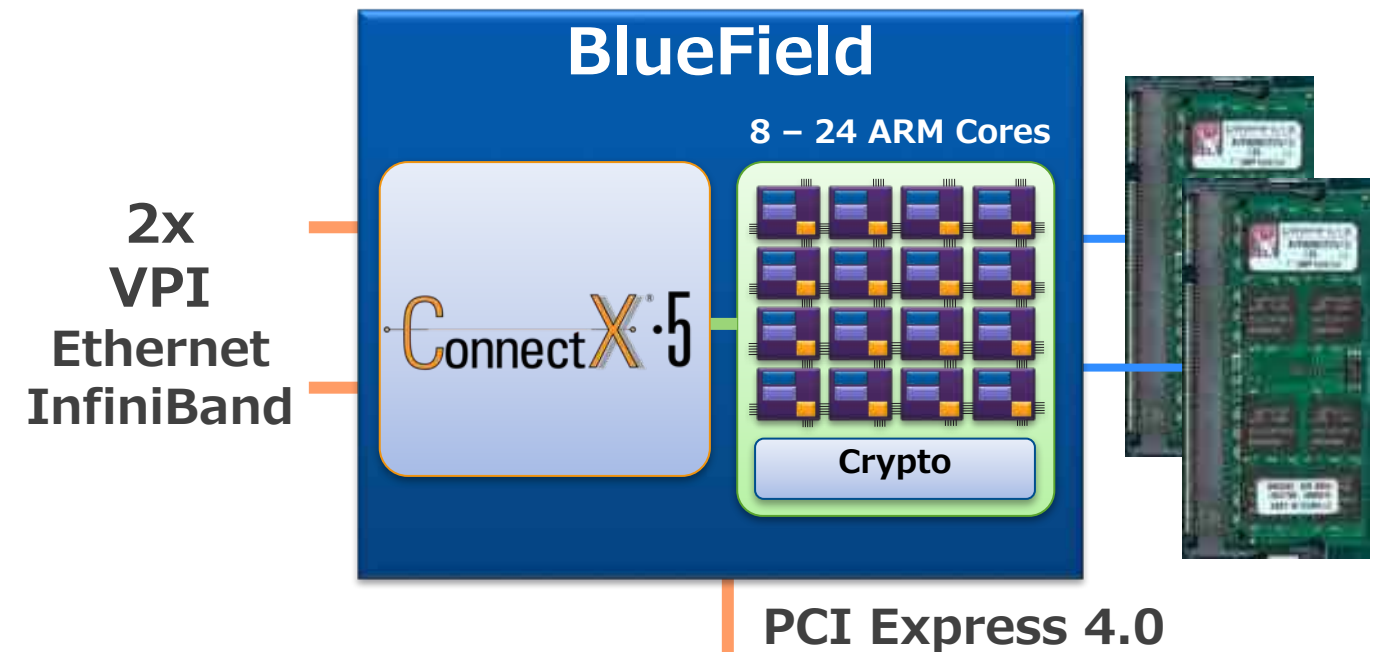


■ ConnectX®-5とマルチコアARMの統合製品

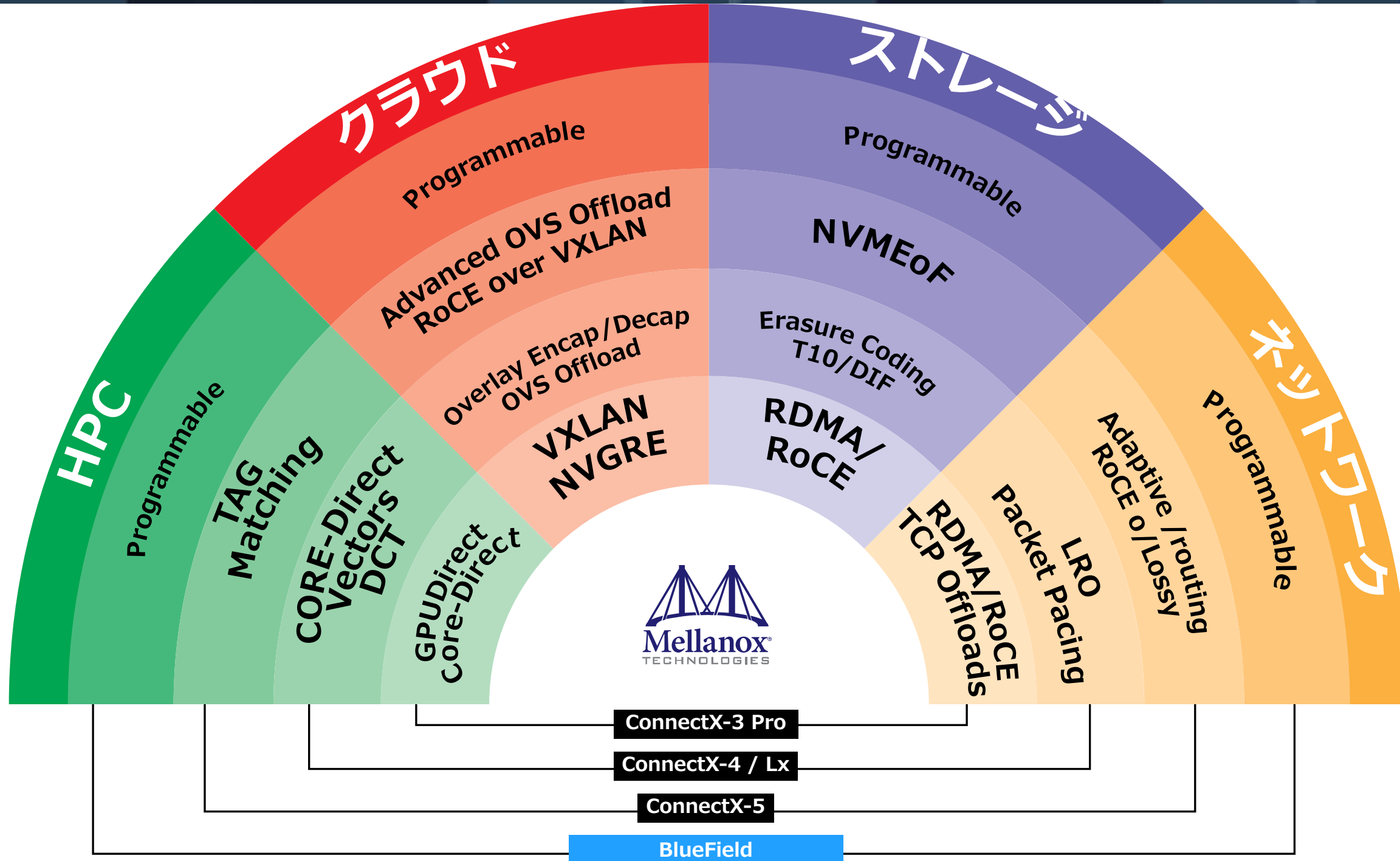
- 業界をリードするConnectXのネットワークアクセラレーション
- ソフトウェアプログラム可能なマルチコア処理

■ 最新機能のサポート

- 10 / 25 / 40 / 50 / 100G Ethernet / InfiniBand
- PCI Express 4.0 Ready
- 暗号化アクセラレーション

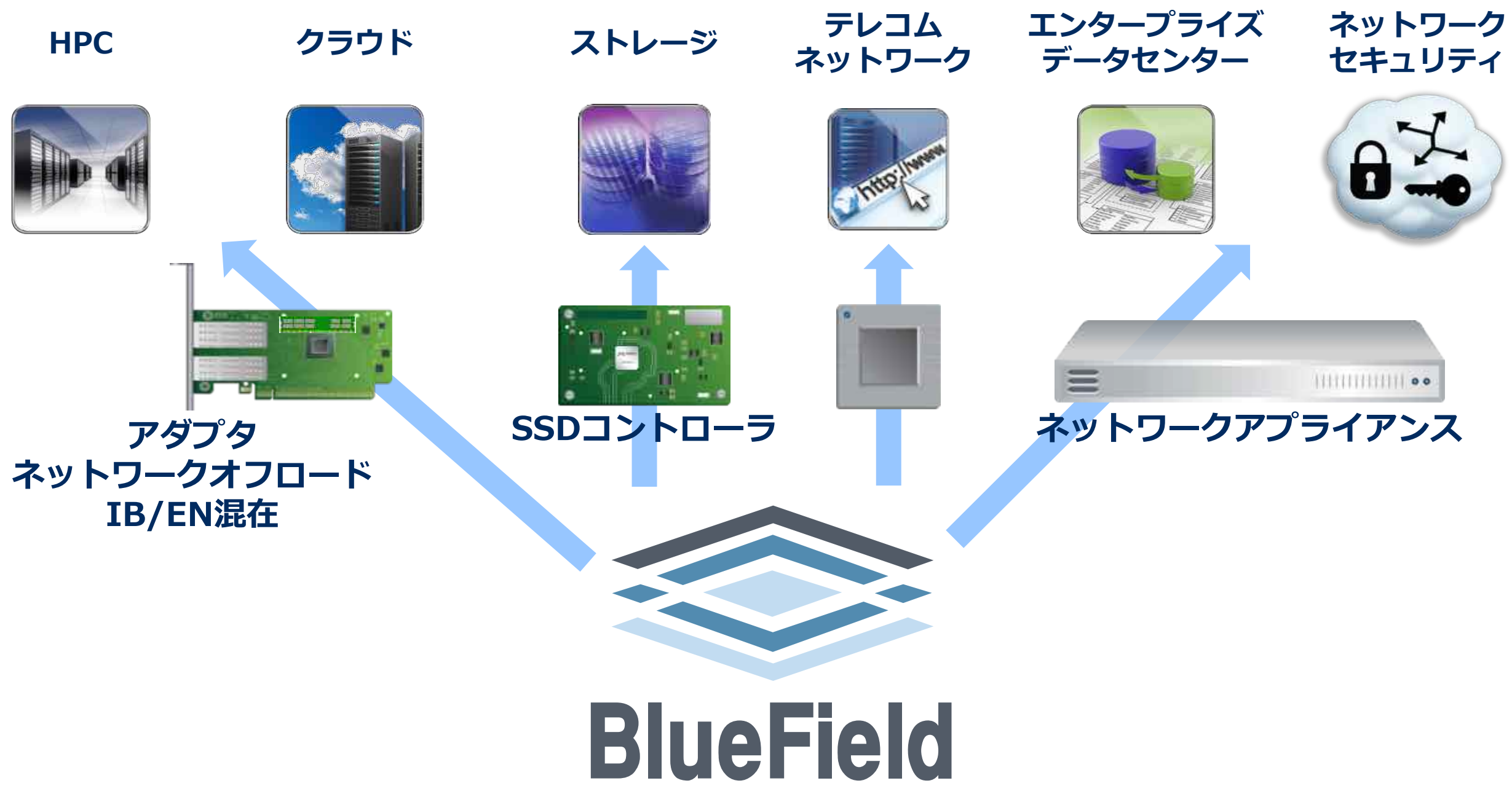


BlueFieldによってアダプタで処理できる機能がより充実



* ConnectX-4 only

さまざまなビジネスセグメントで要求されるネットワークオフロード





Thank You