MARVELL®

超高速イーサネット時代の新しいストレージのカタチ
～NVMeは来たぞ、NVMe over Fabricはまだか～

2019年11月26日

矢田士朗

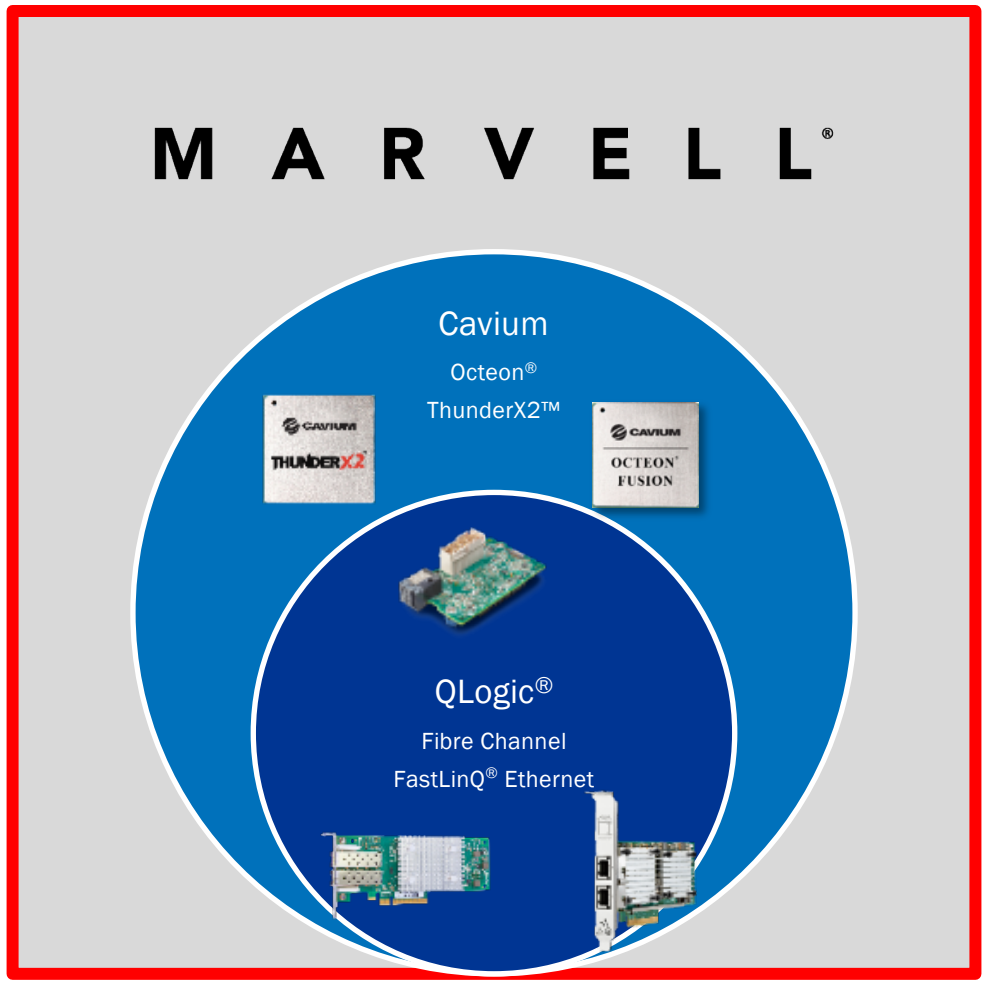
Senior Field Applications Engineer

We Are Marvell

2018

2016

1995



MARVELL®

+

CAVIUM

+

QLOGIC®
a Cavium company

↓

MARVELL®

+ AQUANTIA® averasemi — Wi Fi™

Marvellの主要製品群：ITインフラを支える半導体ソリューション

プロセッサー、セキュリティー、AI

Computing
Server, baseband & embedded processors

#1 in baseband and multi-core processors

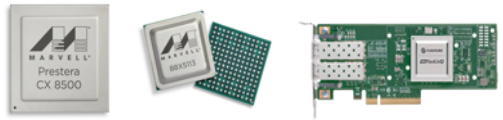
Security
FIPS & virtual offload

#1 in security processors | 1M+ LiquidIO® ports shipped

Artificial Intelligence
Accelerators & offload processors

ネットワーク

Networking
Ethernet switches, PHYs & NICs



#2 in Networking | 3M+ FastLinQ® ports shipped

Automotive
Secure Ethernet, PHYs, storage

ストレージ

Storage Controllers
HDD & SSD

#1 in HDD and SSD controllers

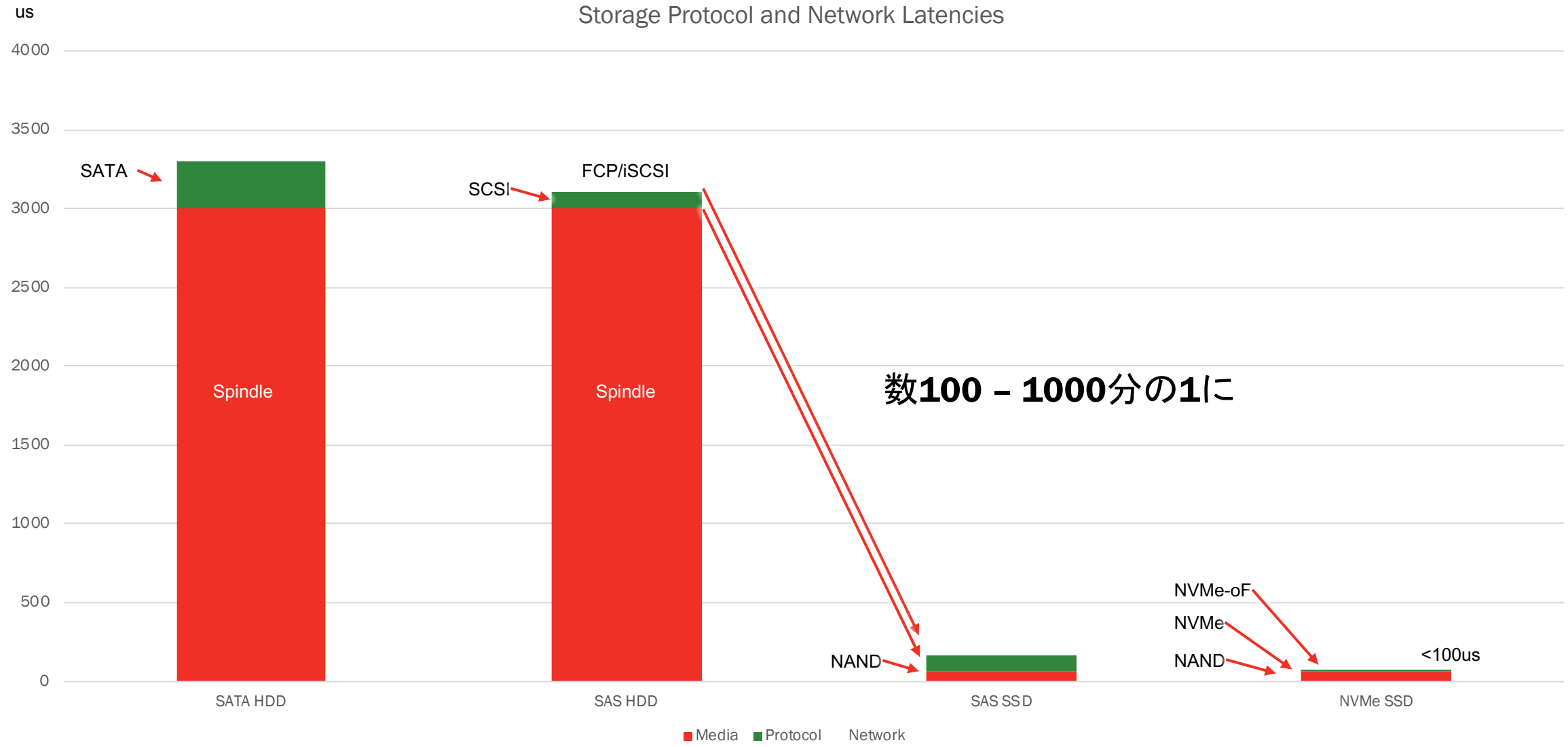
Fibre Channel
Adapters & controllers

#1 in Fibre Channel | 20M+ ports shipped

Data Center Storage Solutions
NVMe aggregators, accelerators & converters

ドライブインターフェースの進化

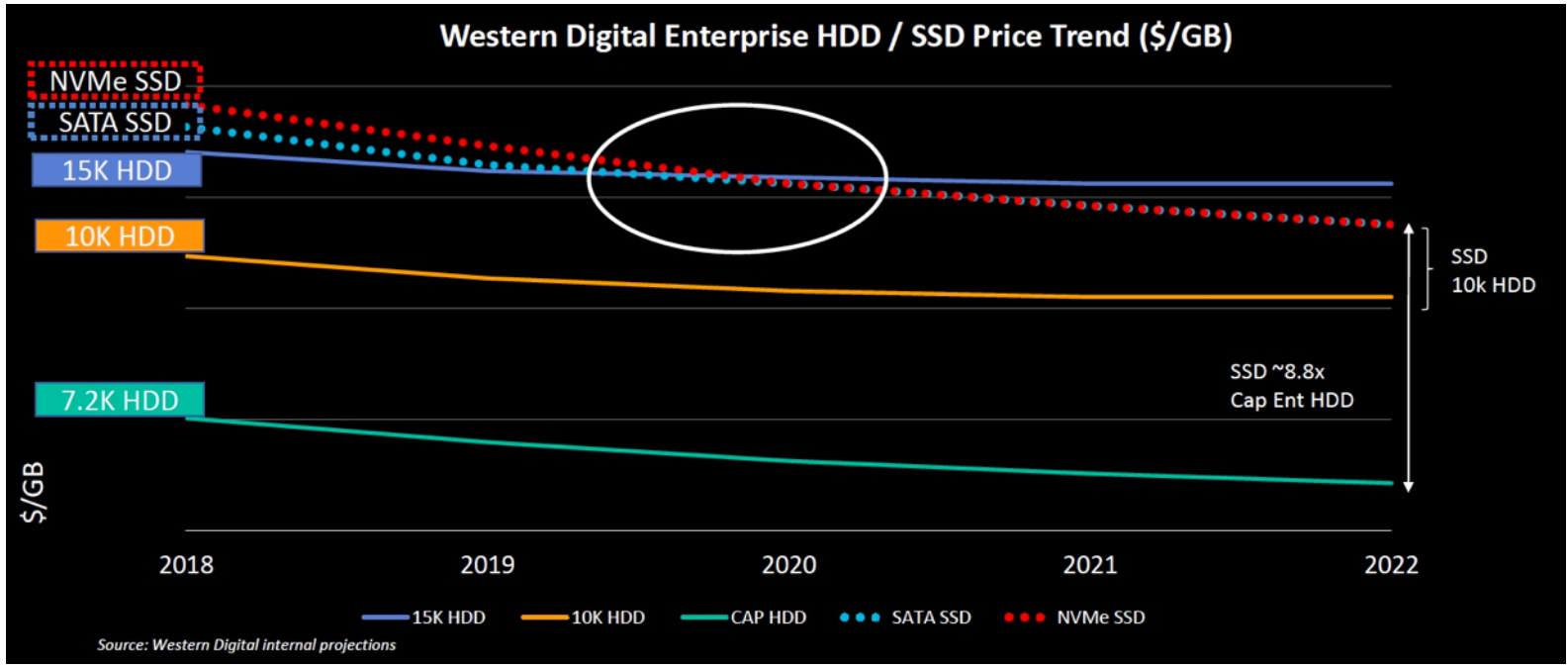
SSD の普及でレイテンシーが劇的に改善



DriveのInterfaceはNVMeに

SCSI (SAS) / ATA (SATA) ⇒ **NVMe**

Enterprise SSD GB単価(18年末)
- SATA： Low $0.20/GB
- SAS：$0.30/GB
- NVMe：mid-$0.20 to 0.30/GB



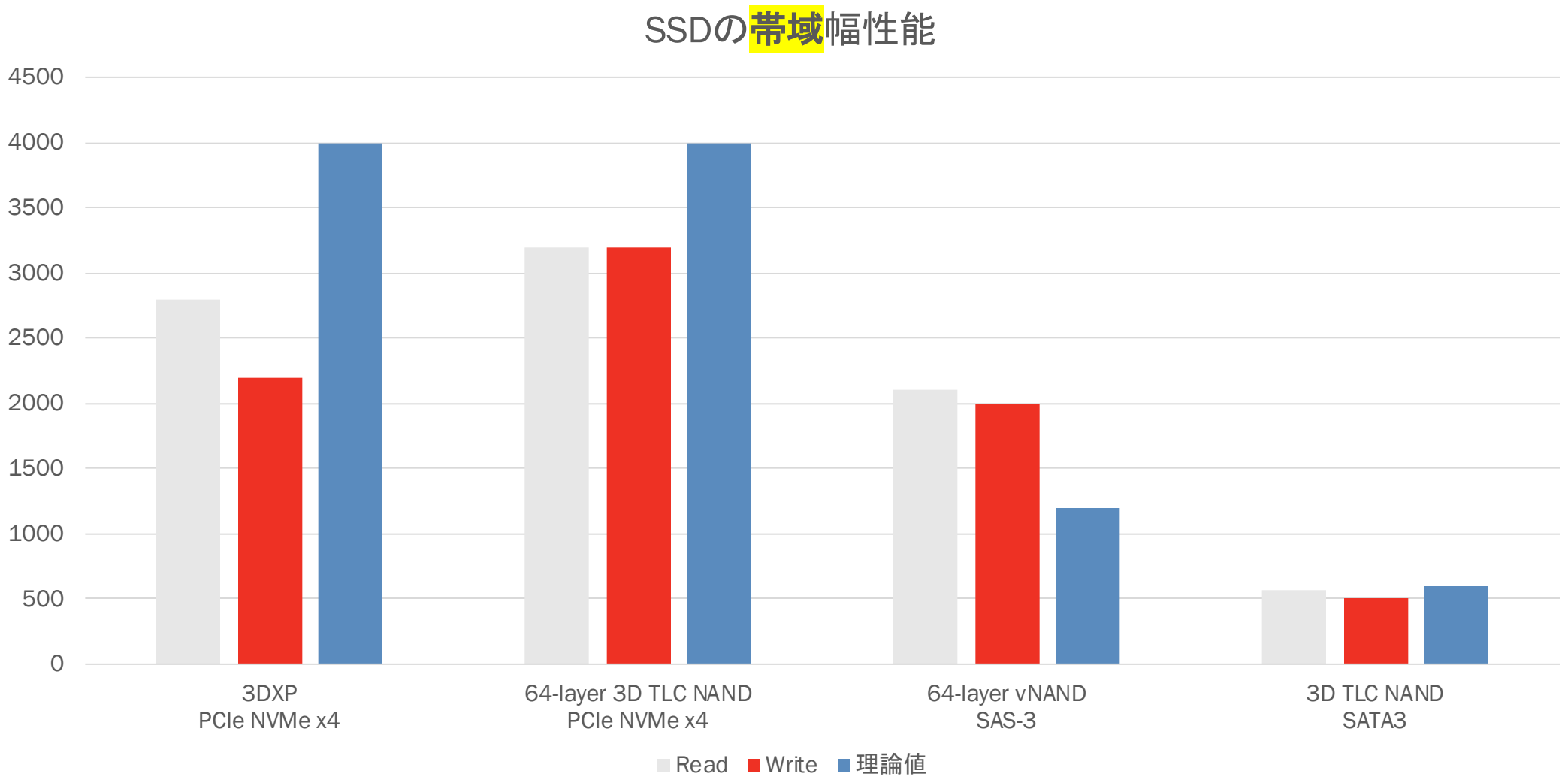
Mar 2019 Western Digital, A3 TechLive conference in London より

各種 SSD 比較

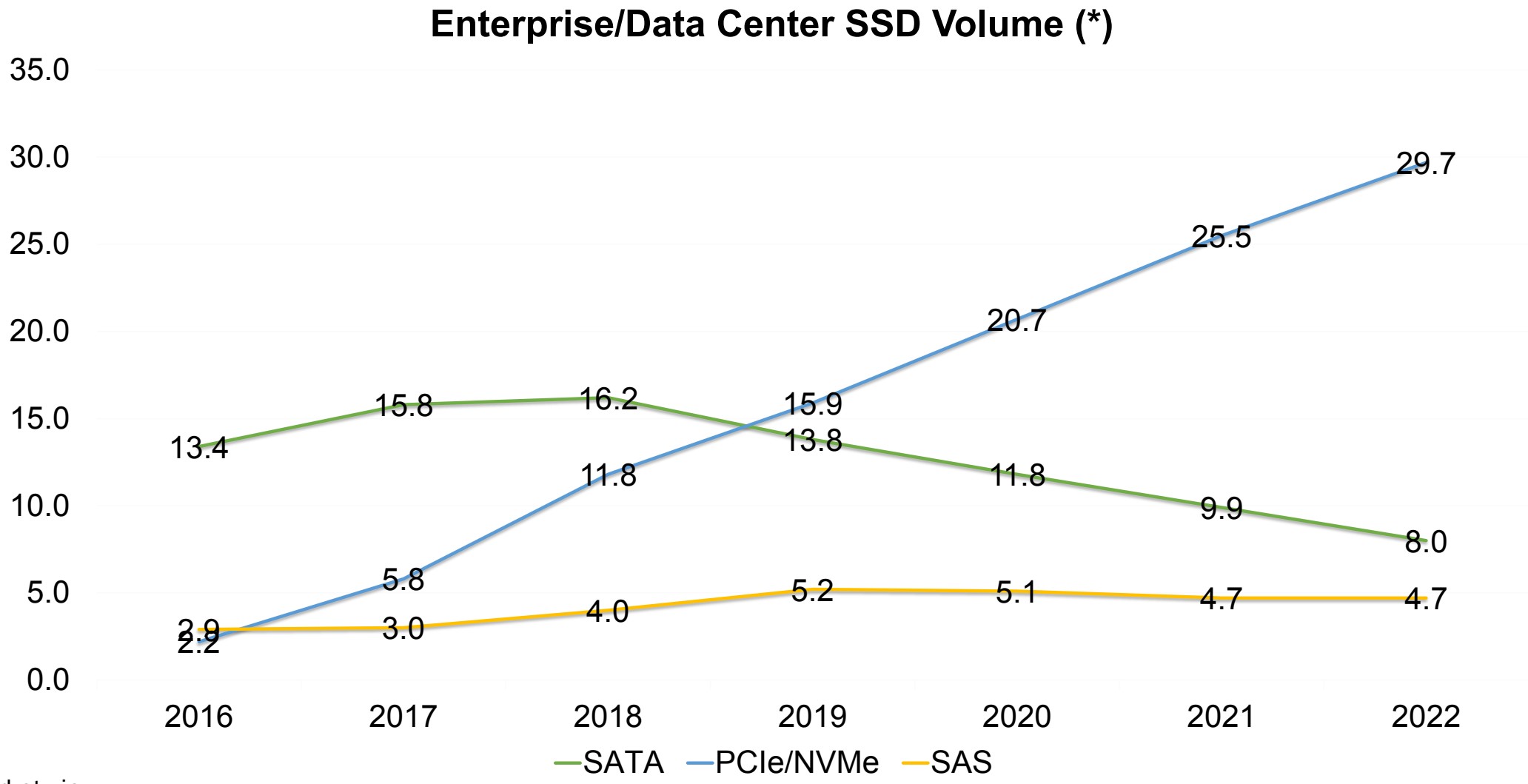
Type	Interface	Read BW	Write BW	Read IOPS	Write IOPS	Rd lat.	Notes
3DXP	PCIe NVMe x4 32Gbps	2600MB/s	2200MB/s	593k IOPS	603k IOPS	11 µs	Intel Optane™ 905P 480 MB
64-Layer 3D TLC NAND	PCIe NVMe x4 32Gbps	3200 MB/s	3200 MB/s	643k IOPS	199K IOPS	77 µs	Intel DC P4610 7.68 TB
64-layer v-NAND	SAS-3 12Gbps	2100MB/s	2000MB/s	400k IOPS	70k IOPS		Samsung PM1643 7.68 TB
3D TLC NAND	SATA3 6Gbps	560 MB/s	510 MB/s	96k IOPS	42K IOPS	36 µs	Intel D3-S4610 3.84 TB

	ビットレート	実行帯域
NVMe (PCIe 3.1x4)	32Gbps	約32Gbps
SAS-3	12Gbps	9.6Gbps
SATA3	6Gbps	4.8Gbps

Driveインターフェースは NVMeの性能が圧倒的に高い



マーケット トレンド: NVMeがSAS/SATAを追い抜く

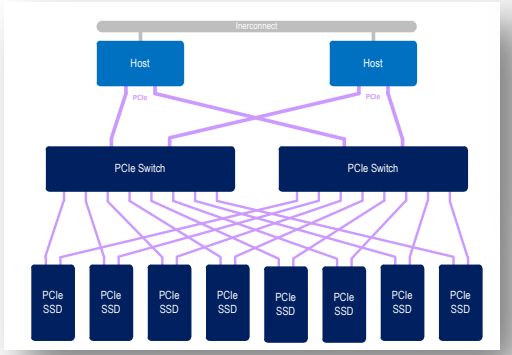


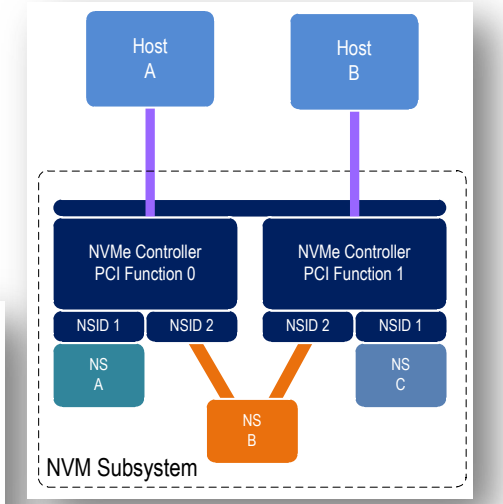
* Marvell market view

NVMeのメリット

- Flashへのアクセスの最適化（Diskメディアに非対応）
- メニーコアのサポートを前提として設計
- キューの数が多い（最大64K個）
- キューあたりのコマンド数が多い（最大64K個）
- PCIe直結なのでレイテンシーが低く、帯域が広い
- （エンタープライズ向け機能を初めから実装）
  - » Multi-path/Dual port, DIF/DIX, Unique ID

FMS2013, An NVM Express Tutorialより

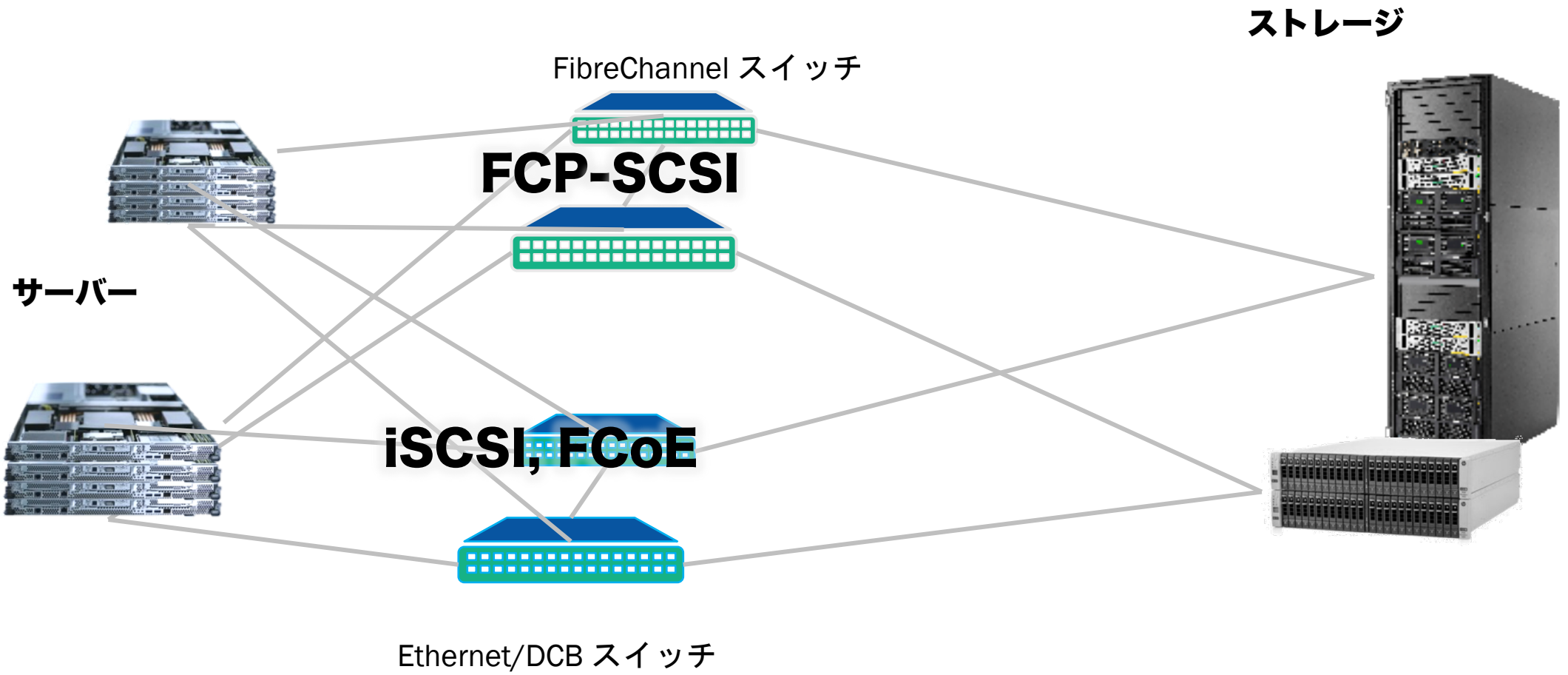




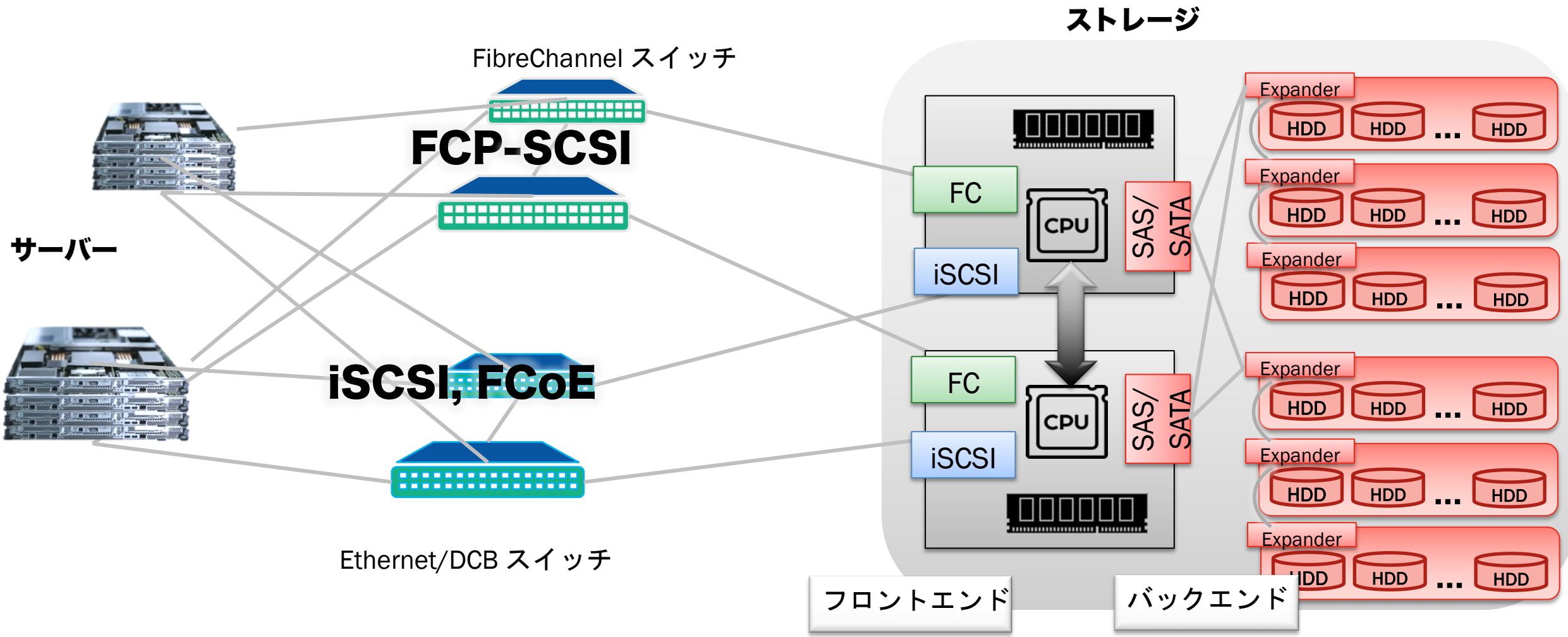
SATA、NVMe共にDriverがInbox, CPUで直結できる

ストレージ ネットワークの進化
～ストレージ システムの中でのNVMe～

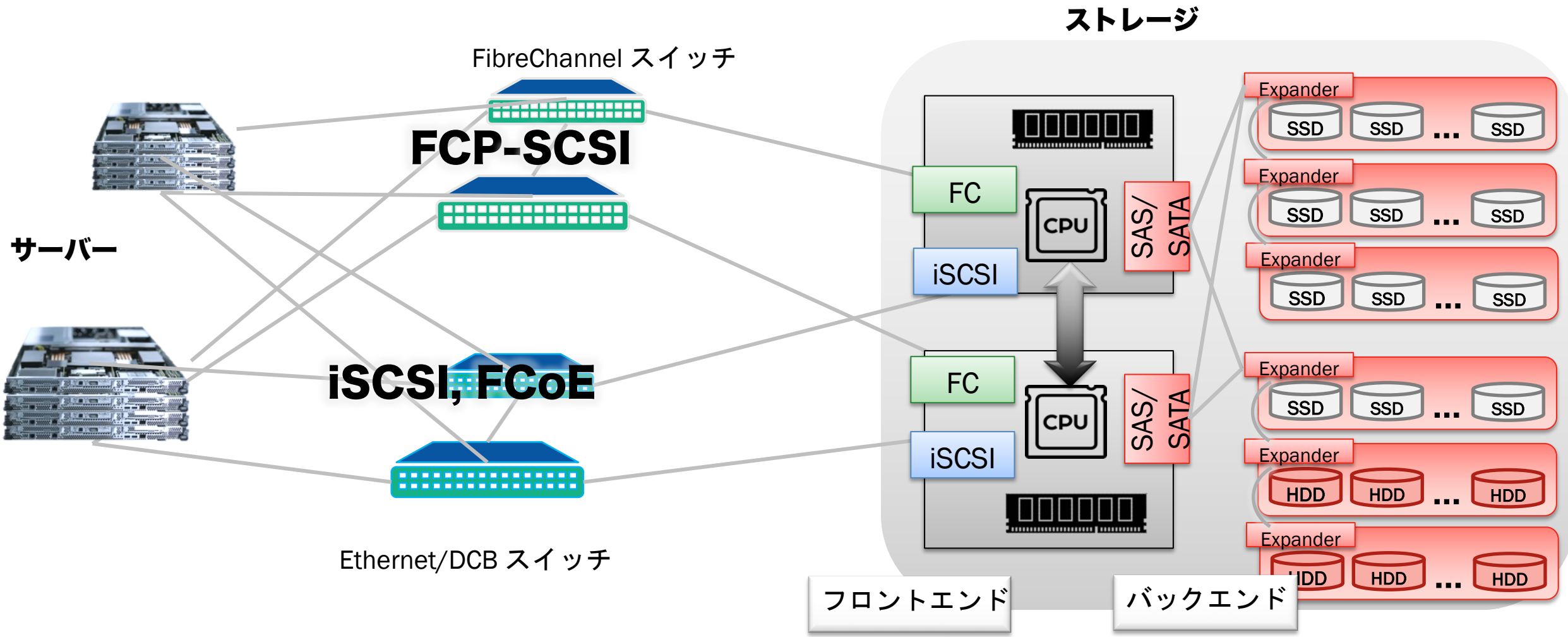
ストレージ ネットワーク構成



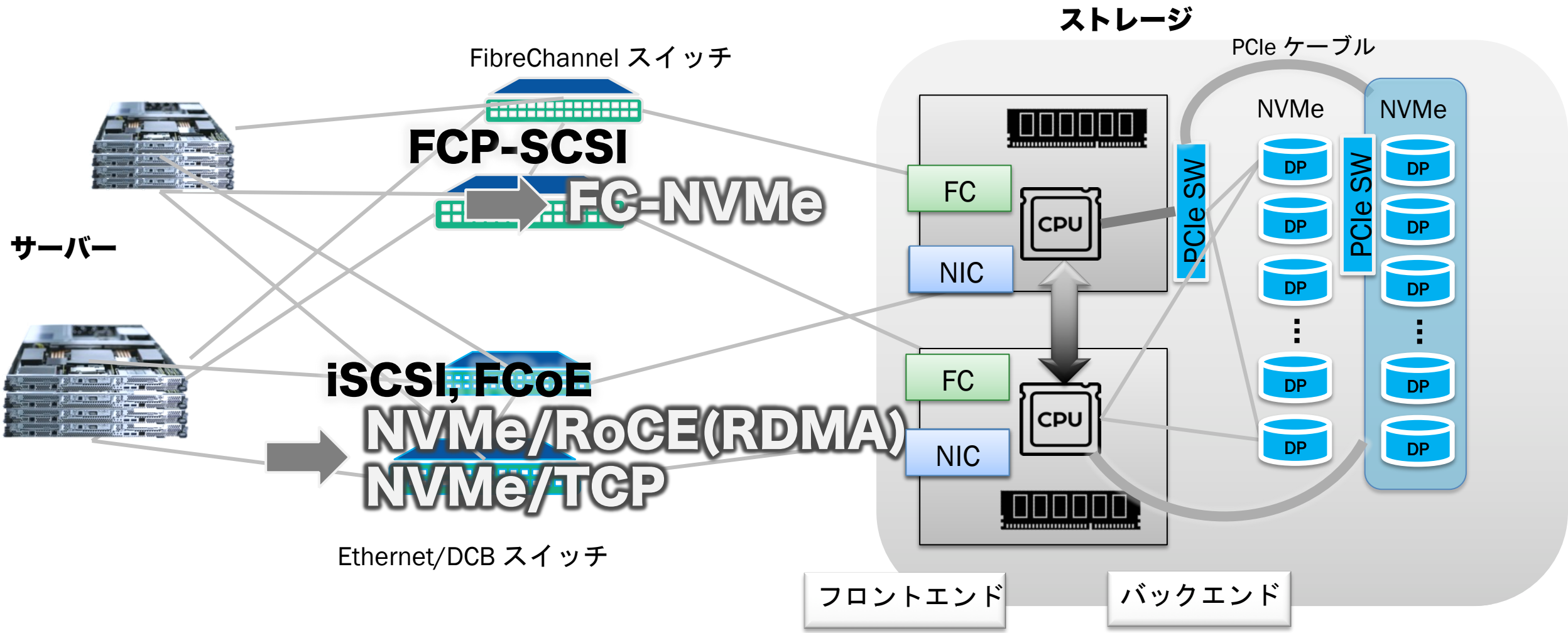
ストレージ ネットワーク構成 (SAS/SATA HDD)



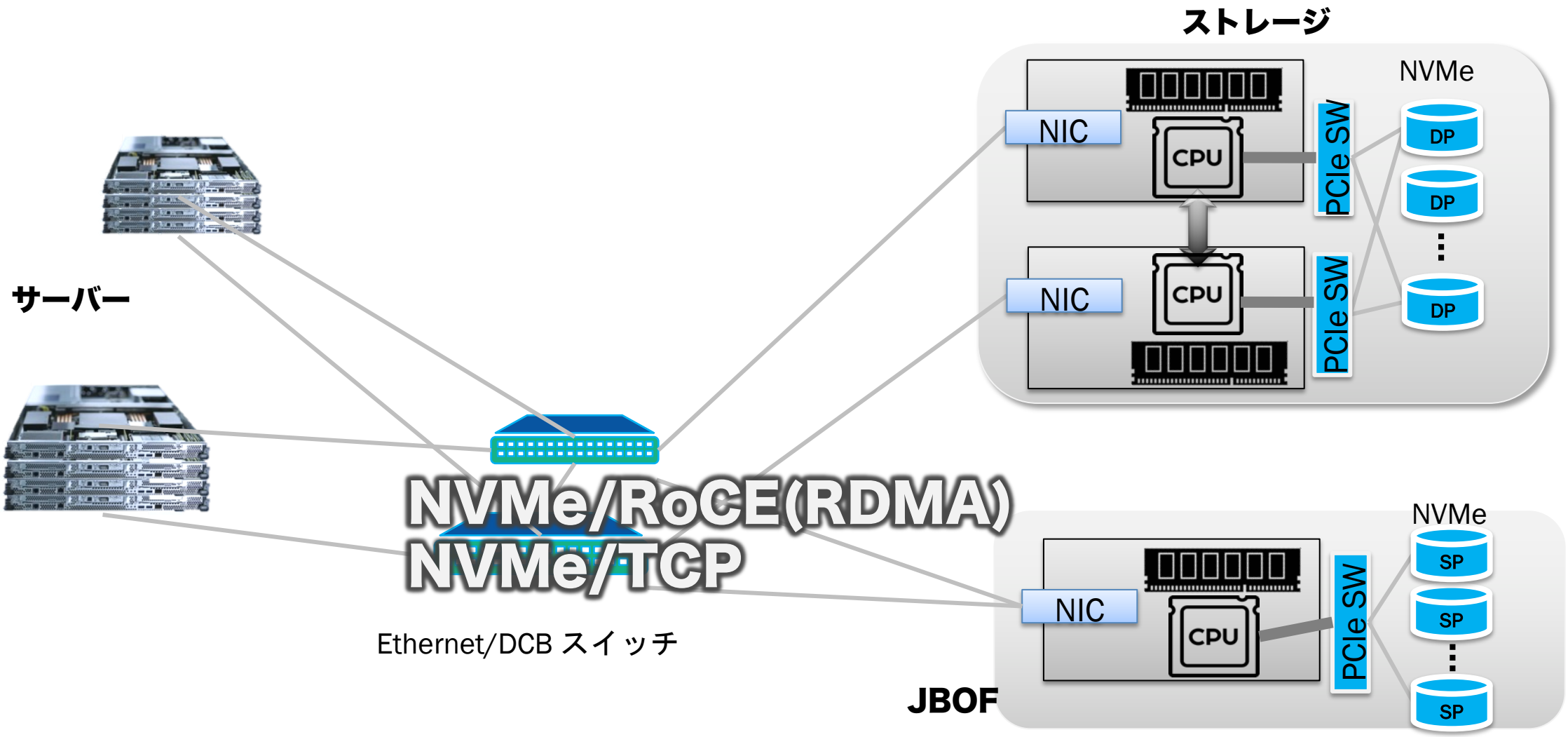
ストレージ ネットワーク構成 (SAS/SATA SSD)



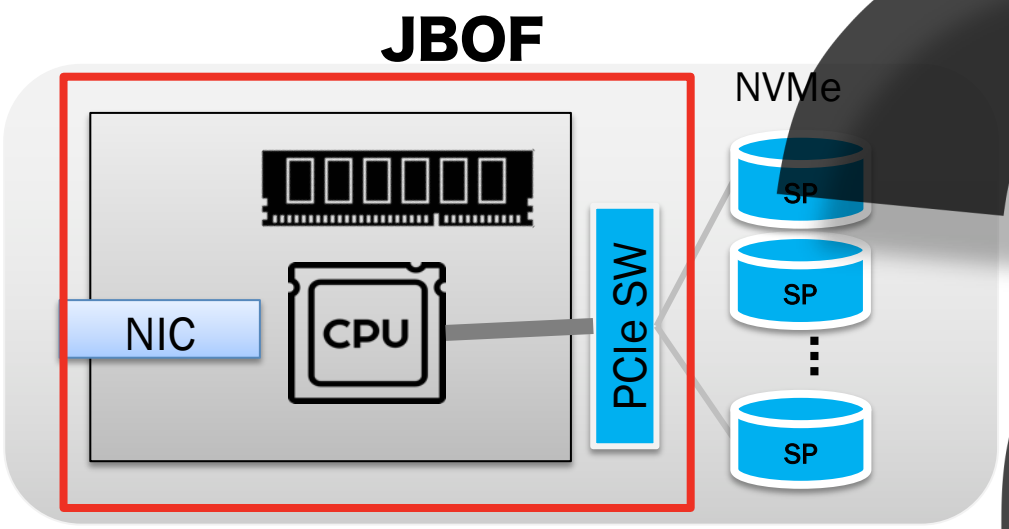
ストレージ ネットワーク構成 (NVMe SSD)



ストレージ ネットワーク構成 (NVMe End to End)



NVMeのメリットを活かしきれていない



- **高価(CPU、メモリー、NIC、PCIeスイッチ)**
- **消費電力が高い**
- **ボトルネック**

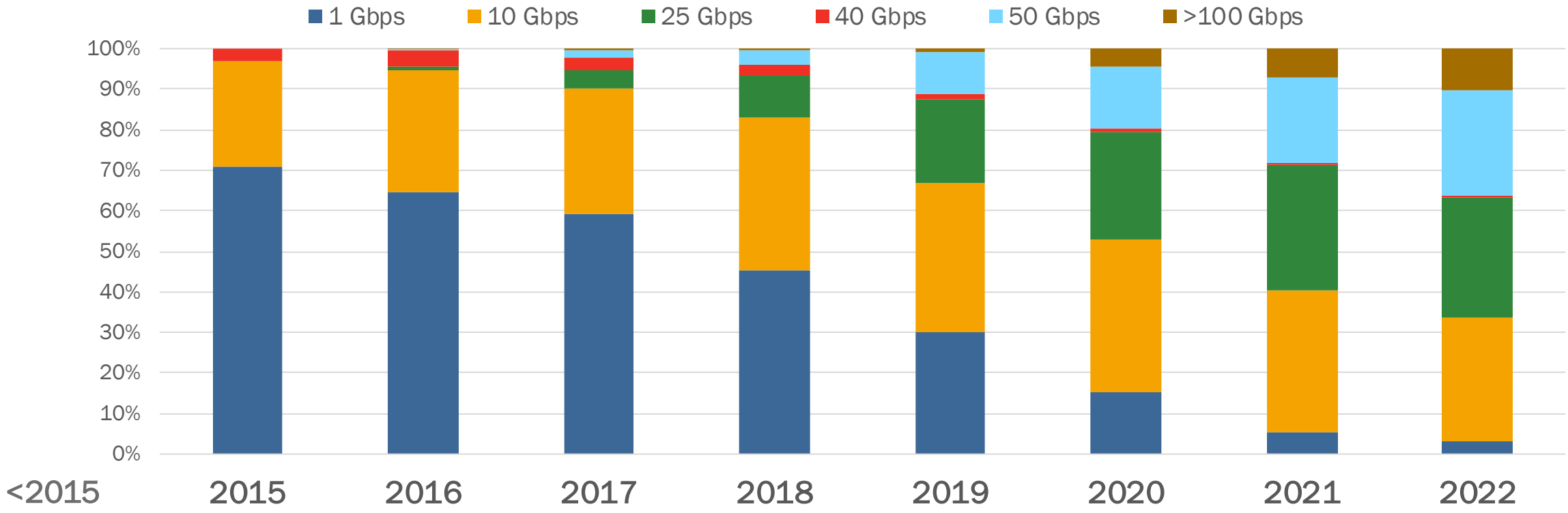
NVMe SSD (Gen3x4) 32Gbps (25Gbps)x 24
＝768Gbps (600Gbps)

Gen4になると
＝1.5Tbps

データセンター内のネットワークの進化

Ethernet スピードの進化と導入トレンド

Source: Dell Oro Jan 2016 and July 2018



1/10GbE
10/40GbE

25/40GbE
100GbE

50GbE(PAM4)
200/400GbE(PAM4)

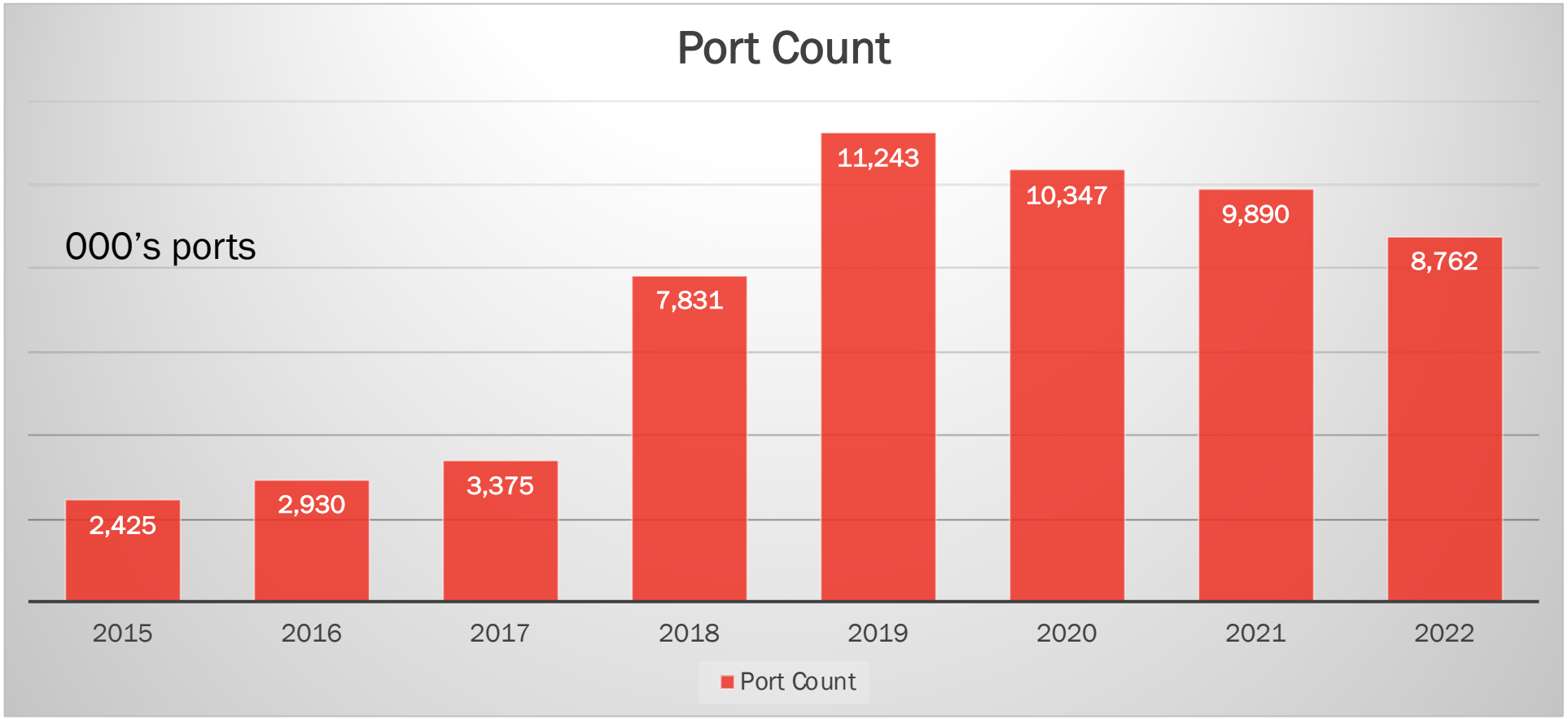
1/2/4G/8/16G FC

32G FC

64G FC

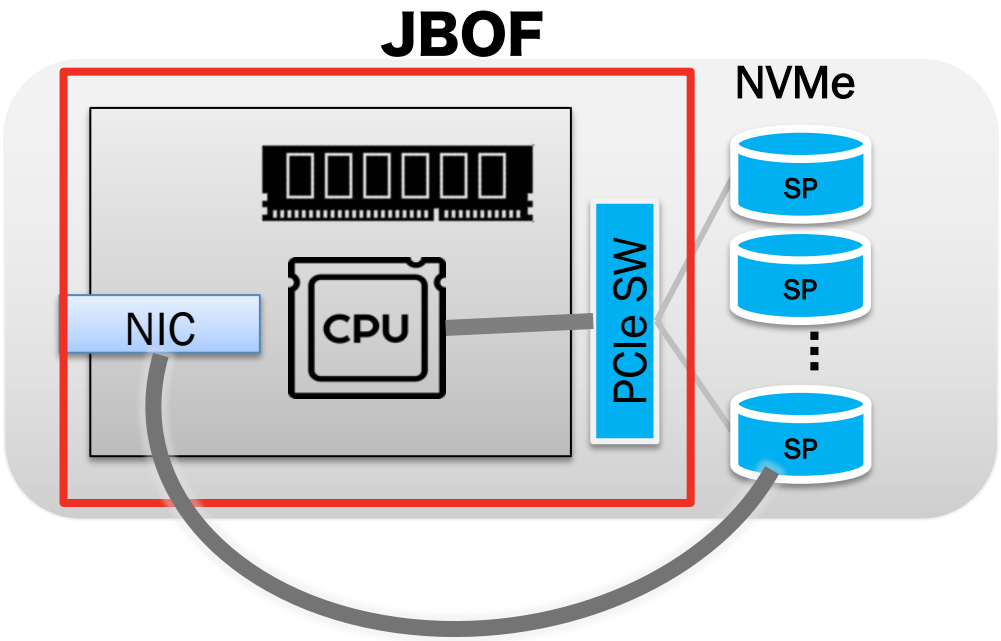
128G FC

10GBASE-T Port Count Trend



Source: Dell Oro July 2018

NVMeのメリットを活かすには？



- 高価(CPU、メモリー、NIC、PCIeスイッチ)
- 消費電力が高い
- ボトルネック

NVMeドライブを
直接ネットワークへ接続

NVMe to NVMe-oF コントローラー

MARVELL®

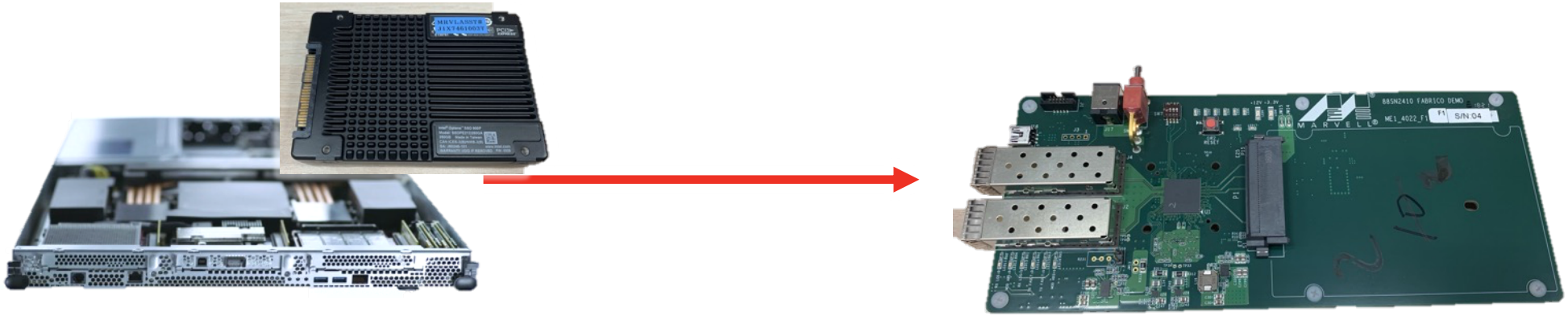
Industry's 1st NVMe-oF SSD Converter Controller

Fabrico

サンプル出荷中

- どんなNVMe x 4インターフェースも 2 ポート 25GbE NVMe-oFに変換
- NVMe/RoCEv2とNVMe/TCPに対応
- 変換によるレイテンシーの増加は 750nsec 以下
- 700k IOPSまでの処理性能
- 消費電力　1.5W以下
- 13mm x 13mmのコンパクトなコンパクトなパッケージ

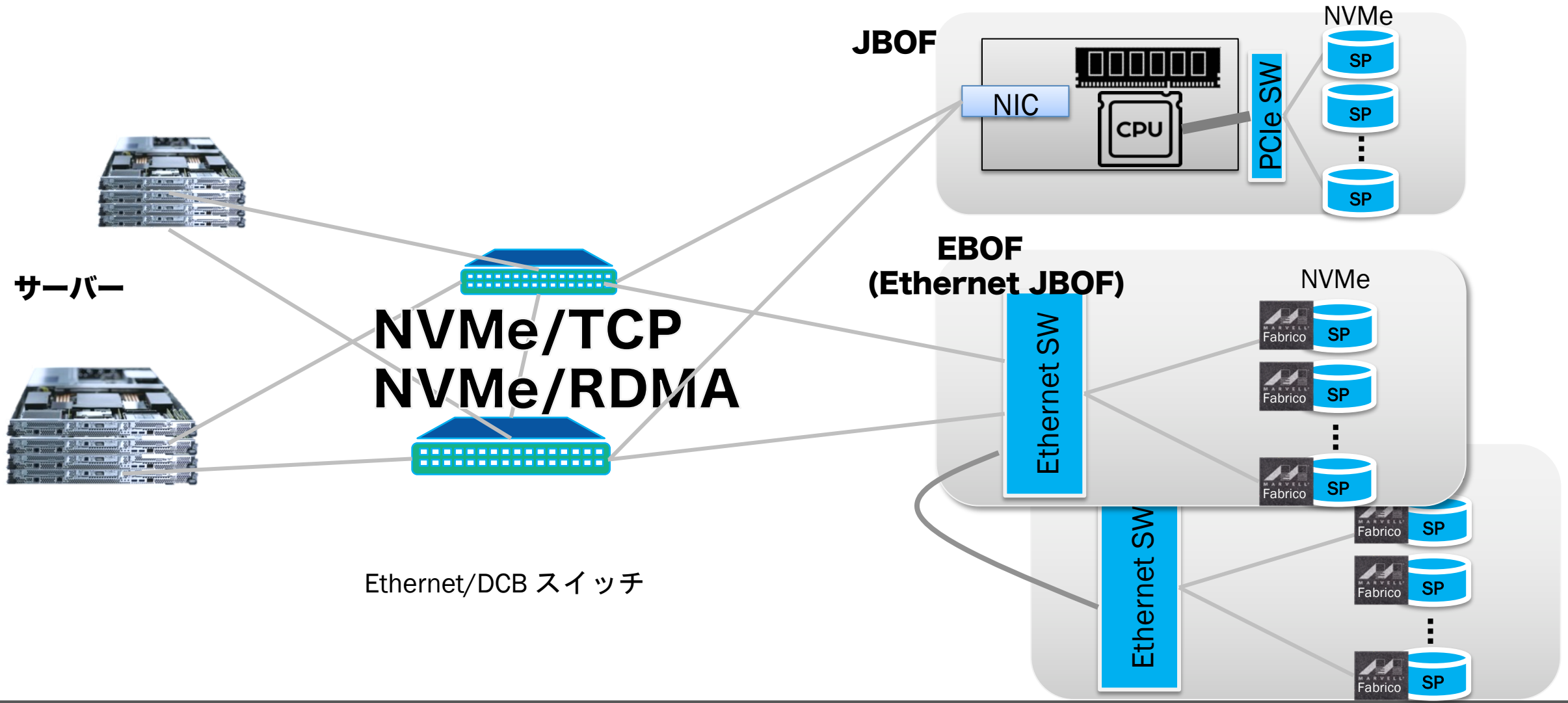
論より証拠 (Fabrico性能)



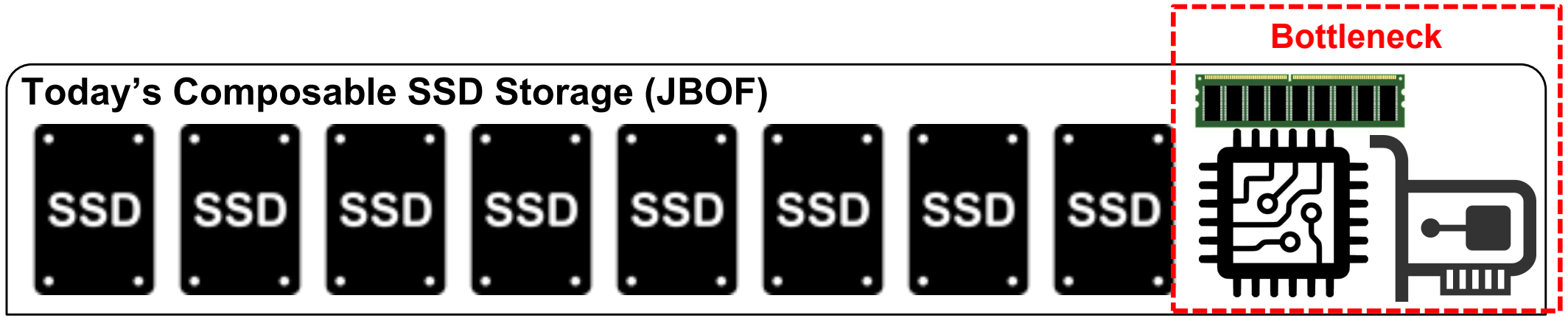
	Avg. Read Latency (clat)	Read IOPS	帯域
ローカル NVMe	457usec	551k	2582MB/s

<測定条件>
SSD : Intel Optane 900P
測定ツール: **fio (--bs=4k/512k --iodepth=32 --numjobs=8)**

ストレージ ネットワーク構成 (NVMe End to End)

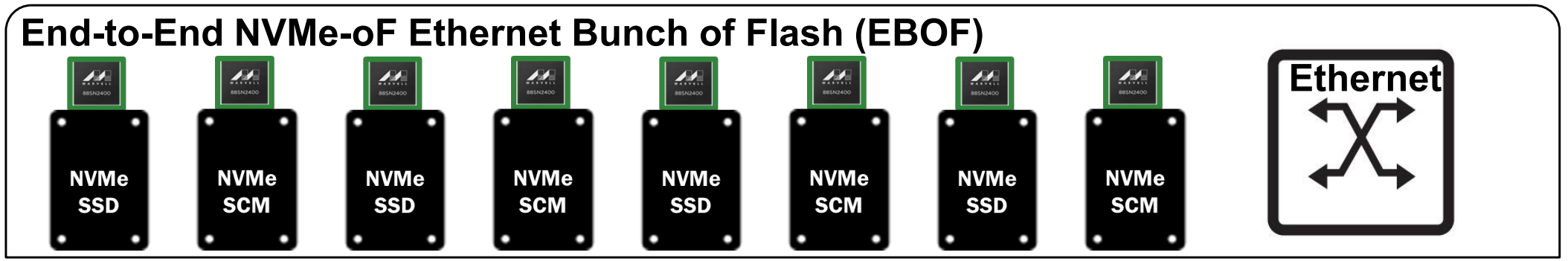


省電力, 低コスト, 高性能 NVMe-oF ソリューション



High TCO

性能ボトルネック、高い消費電力、高いBOMコスト



Low TCO

シンプルでスケーラブルな構成で高い性能を低消費電力で実現

TCOを65%以上も低減！(SSDを除く)

*Toshiba & Marvell TCO analysis

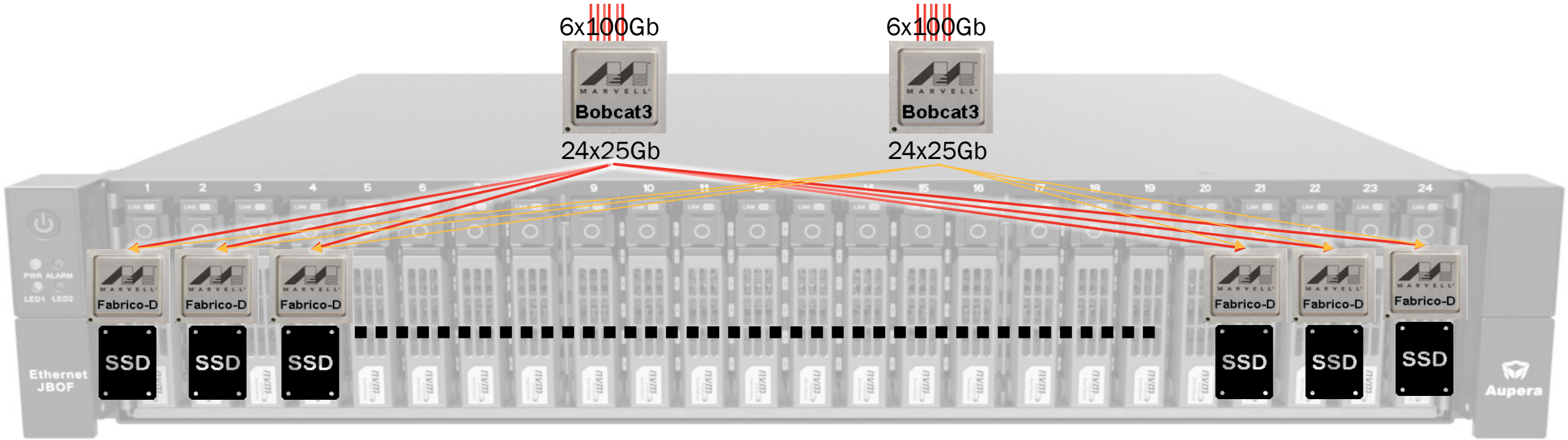
30

これが Ethernet JBOF (EBOF)！

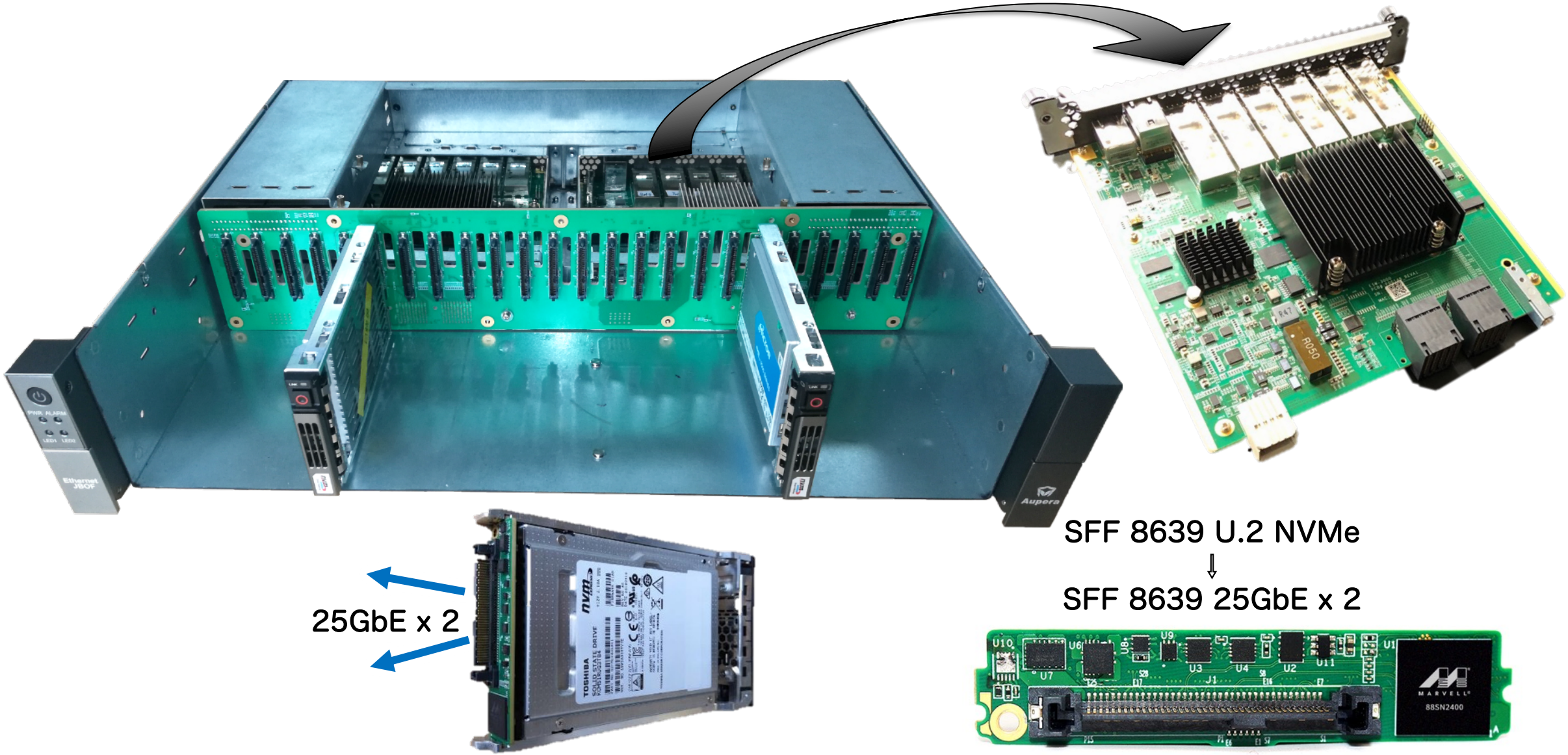
2U24 SSDのエンクロージャーにFabricoを搭載した例

これが Ethernet JBOF (EBOF)！

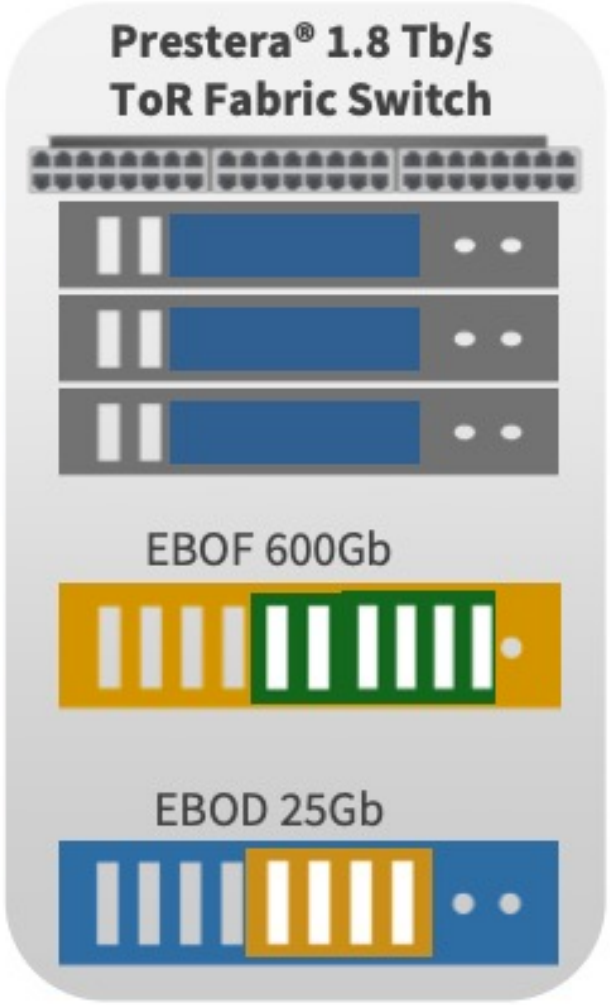
2U24 SSDのエンクロージャーにFabricoを搭載した例



2U24 drive EBOFの中身



OCP Global Summit Mar 2019のデモの様子

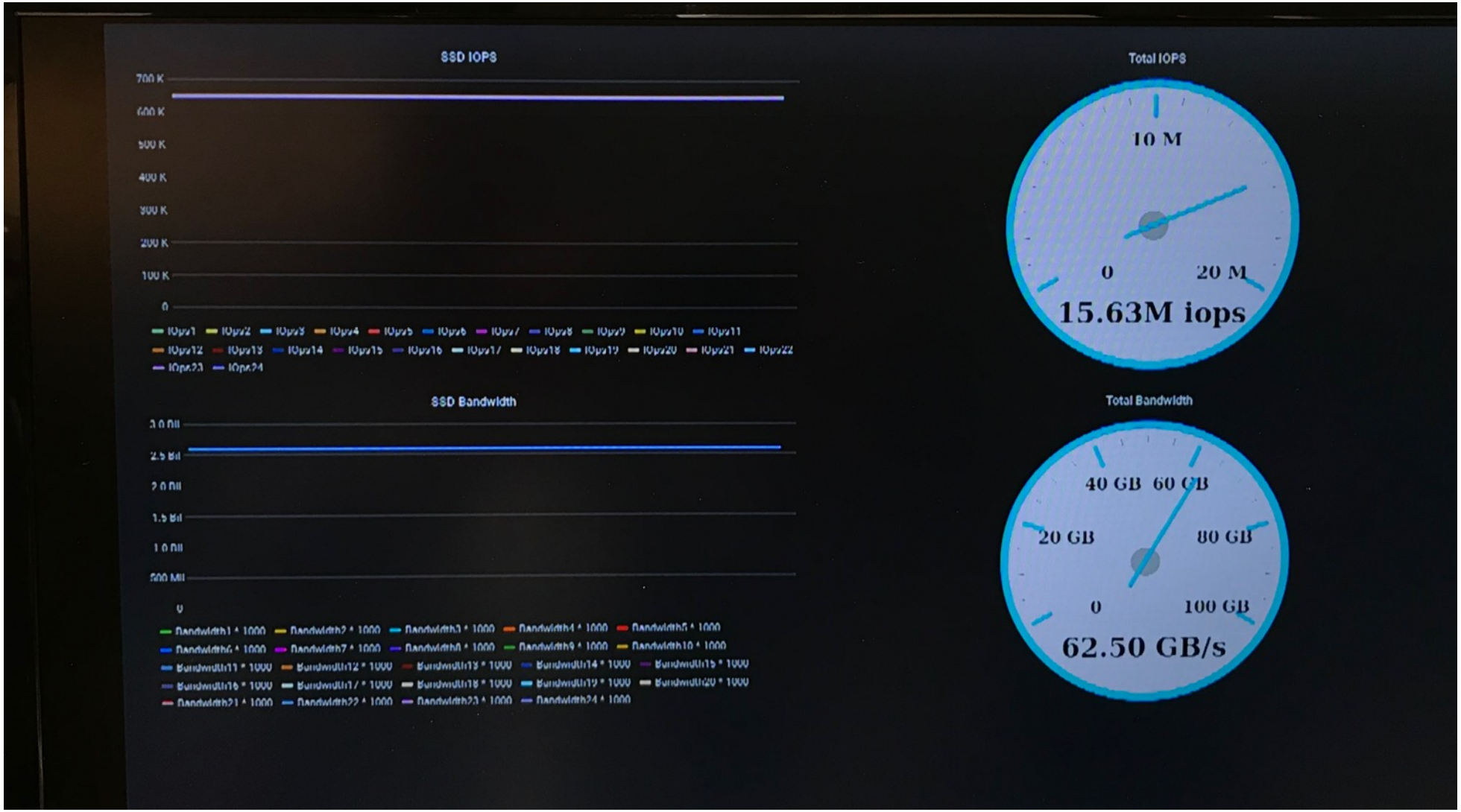


Initiators:
3x 2S ThunderX2® Compute shelf with 200Gb FastlinQ® RNIC
Total: 600Gb/s

EBOF:
2U 24 SSD shelf with Marvell Prestera® & NVMe-oF Storage Controllers
Input: 600Gb, Output: 24x 25Gb SSDs
Total: 1.2Tb/s

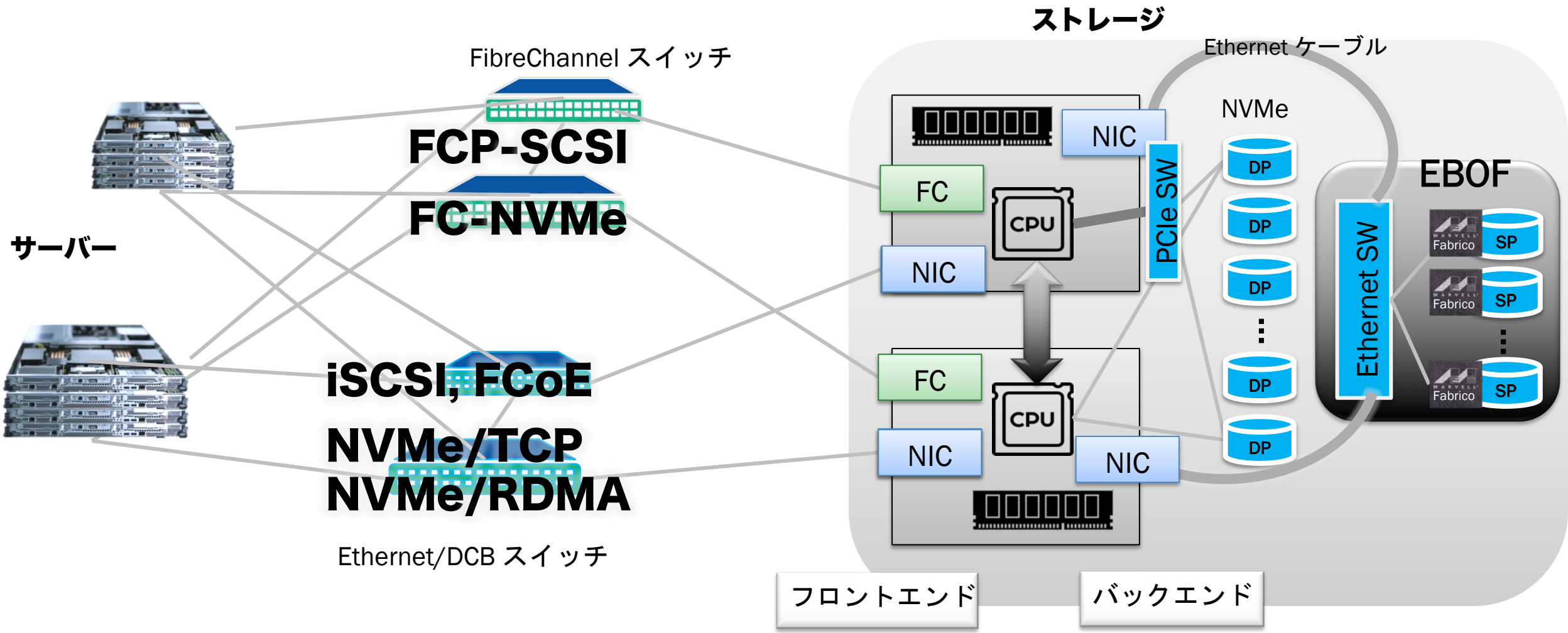
EBOD Target:
2U 16 HDD shelf with Marvell Storage Processor & NVMe-oF Storage Controllers
Input: 25Gb, Output: 16x 12Gb HDDs
Total: 217Gb/s

EBOF ES OCP Demo (6x100Gb switch)



*Demo at the Toshiba booth using 2U 24x CM-5

ストレージ ネットワーク構成 (NVMe-oFバックエンド)

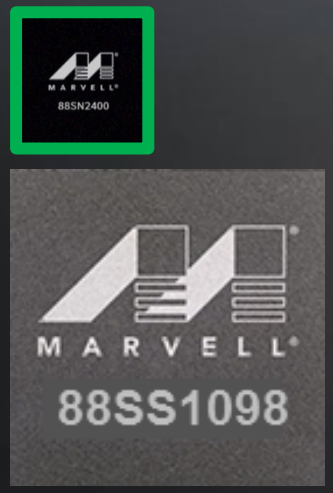


ここまでは実は去年までの話
今年は。。。

Flash Memory Summit 2019 で発表　#1

業界初！シングルチップNVMe-oF Ethernet SSD コントローラー

2018
Flash Memory Summit



Industry's 1st NVMe-oF SSD Converter Controller

Dual-ported 25GE to PCIe Gen3x4

<1.5W operating power

13mm x 13mm package

88SS1098 Data Center PCIe Gen3x4 SSD controller

Up to 700 kIOPS

17mm x 17mm package

2-chip SSD solution

2019
Flash Memory Summit

Industry's 1st Native NVMe-oF Ethernet SSD Controller

Dual-ported 25GE to eight 800MT/s NAND channels

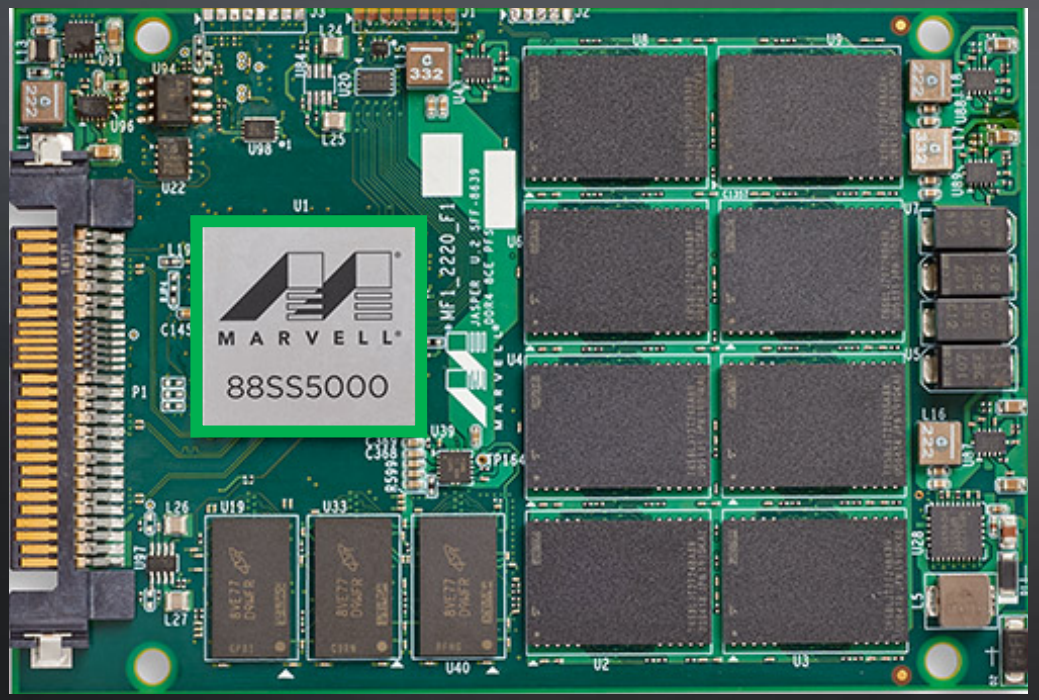
<5W operating power

Up to random 700kIOPs

21mm x 21mm package

Single chip SSD solution

NVMe-oF Ethernet SSDが可能に！（コンセプトモデルをデモ）

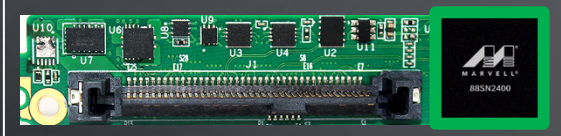


- Dual-ported 25GE Ethernet
- SFF-8639 / 9639 connector with Ethernet pinout
- Up to 8TB capacity
- Up to random 700kIOPs
- **Live at Marvell FMS booth #511**

Flash Memory Summit 2019 での発表　#2

業界初！2.5インチ NVMe-oF Ethernet SSD

2018
Flash Memory Summit



Industry's 1st NVMe-oF SSD Converter Controller

Dual-ported 25GE to PCIe Gen3x4

<1.5W operating power

up to random 700kIOPs

13mm x 13mm package

※**Toshiba Memory, Aupera & Marvell demo >14M IOPS EBOF**

2019
Flash Memory Summit

Industry's 1st 2.5" In-Form Factor NVMe-oF Ethernet SSD

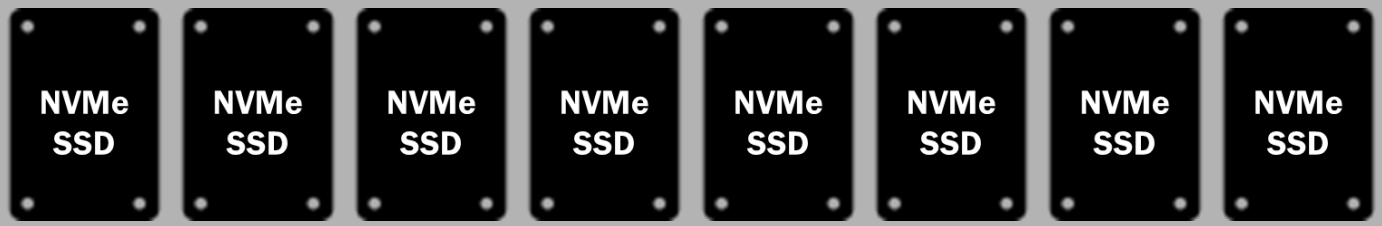
Embeds Marvell's NVMe-oF SSD Converter Controller

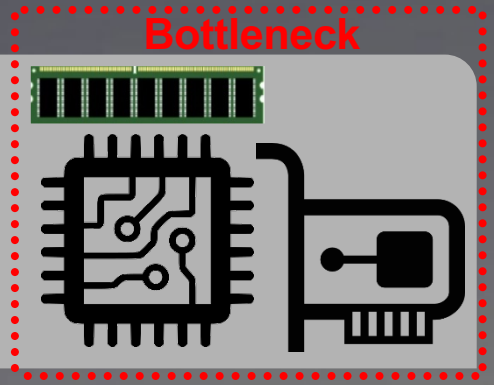
※**Toshiba Memory, Aupera & Marvell demo >15.5M IOPS EBOF**

※旧Toshiba Memory, 現Kioxia

さらに進化した EBOF ソリューション

Today's disaggregated SSD storage (JBOF)

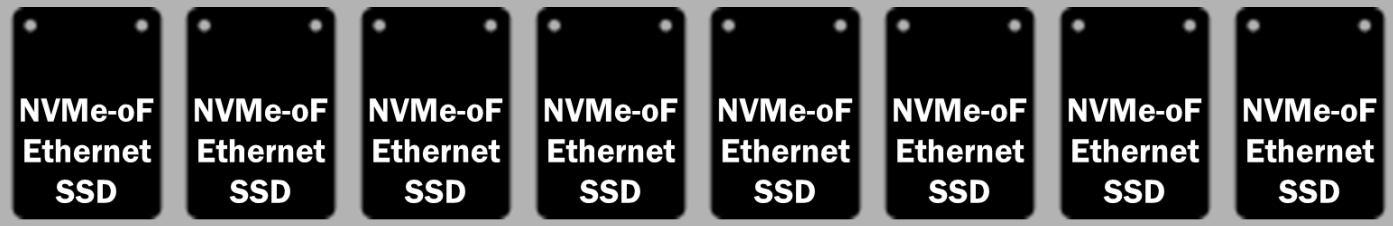


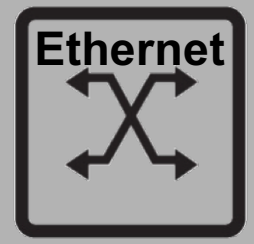


High TCO

Limited performance, high CPU power & high BOM

End-to-end NVMe-oF Ethernet Bunch of Flash (EBOF)





Low TCO

Simple native scalable performance with extremely lower power consumption

>65%* TCO savings excluding SSDs

*Kioxia and Marvell TCO analysis

MARVELL®