



DDN UPDATE 2020

Gfarm Workshop 2020

2020/2/7

DataDirect Networks Japan, Inc.
橋爪信明



At Scale Data Services for
AI, Big Data and HPC



Most efficient All NVMe SW for
the toughest workloads



Optimized AI Platforms For
Every Use Case

DDN[®] AT SCALE ENTERPRISE



High Performance Flash &
Hybrid Unified Storage



SW Defined Storage for Telco
5G, IoT and the Enterprise



Simplicity and Control for
Virtualized Environments

DDN[®]
AI · BIG DATA · HPC

SFA & EXA5

2019/1~2020/3 主なEXAScaler導入実績（含む導入予定）

サイト	物理容量
東京大学様Oakbridge-CX	16.128PB
東京大学地震研究所様	1.024PB
東京大学医科学研究所ヒトゲノム解析センター様	1.53PB
筑波大学様Cygnus	3.2PB
国立環境研究所様	28.224PB
国立遺伝学研究所様	13.2PB
気象庁気象研究所様	23.328PB
京都大学化学研究所様	4.66PB
某国立研究所様	33.6PB
某省庁様	13.92PB
某企業研究所様	1.68PB
某企業研究所様	3.28PB
某国営企業研究所様	1.328PB
某製造業様	1.76PB
産総研様ABCI	1.38PB(All Flash)
北海道大学様	75TB(All Flash)
OIST様	338TB(All Flash)
JAMSTEC様	75TB(All Flash)

HDD
物理146.862PB
実効約117PB

SSD
物理1.868PB
実効約1.49PB

DDN EXA5 X-APPLIANCES



ES200NVX

**24r/20w
1.5M IOPS**

UP TO 24 DRIVES

ALL NVME

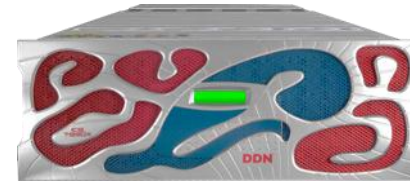


ES400NVX

**48r/30w
>3M IOPS**

UP TO 384 DRIVES

+0 +1 +2 +4



ES7990X

**24r/20w
800K IOPS**

UP TO 450 DRIVES

+0 +1, +2, +4

HYBRID



ES18KX

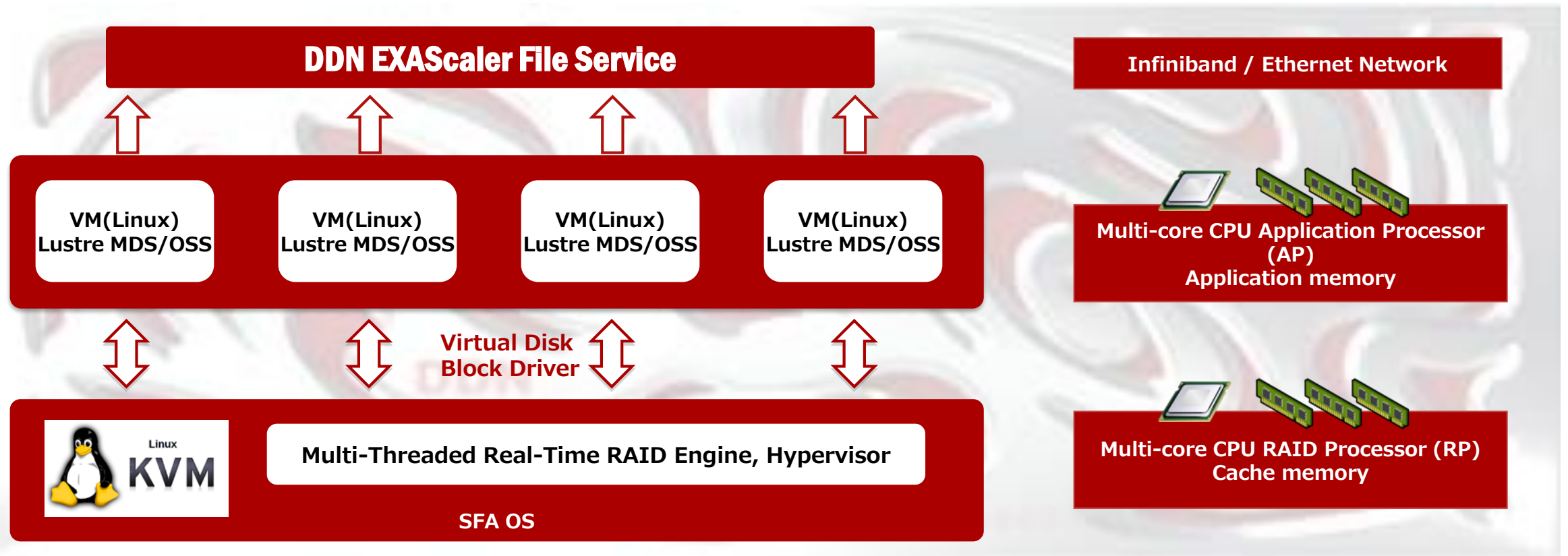
**70r/70w
>3M IOPS**

UP TO 1872 DRIVES

+0 +5, +6, +8, +10, +16, +20

Parallel file system appliances

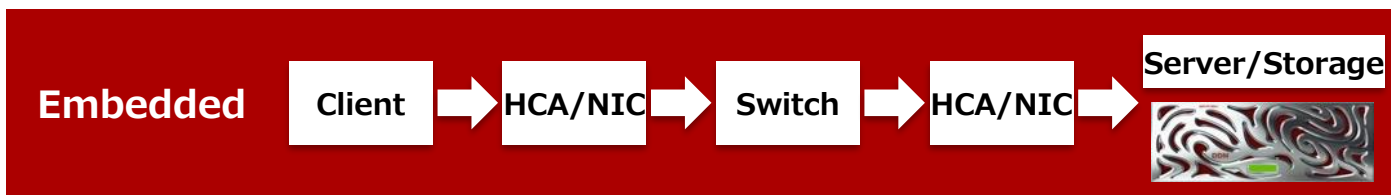
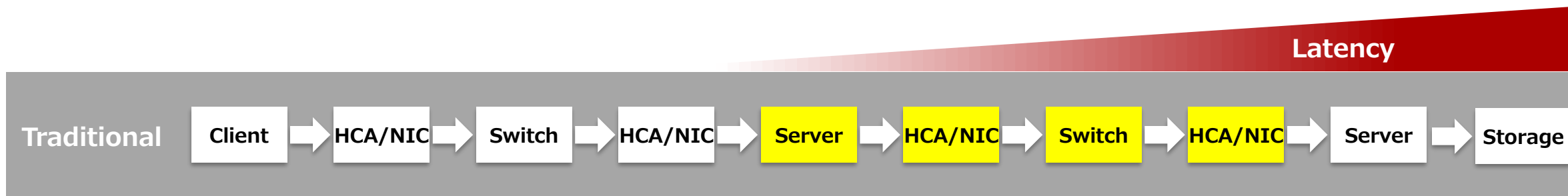
ES18KX、ES7990X、ES200NVX、ES400NVXは、ブロックストレージであるSFA18KX、SFA7990X、SFA200NVX、SFA400NVXのストレージコントローラ上のVirtual Machine上でDDN EXAScalerを実装した並列ファイルストレージアプライアンスです。高度に最適化された仮想化環境により、高い実効性能と効率化を実現します。



In-Storage Processing

ES18KX、ES7990X、ES200NVX、ES400NVXは、ストレージコントローラ内でファイルサービスを実装することにより、外部サーバ、ネットワークアダプタ、スイッチなどを必要としません。主なメリットは以下の通りです。

外部サーバ、ネットワークアダプタ、スイッチ購入コスト削減
構成の簡素化による管理コストの削減
低遅延でのストレージアクセスを実現（下記図参照）



データ/パケットの経路を簡素化することにより
低遅延のストレージアクセスを実現

DDN ES200NVX

ES200NVXは、All FlashブロックストレージであるSFA200NVXのストレージコントローラ上で、DDN EXAScalerを実装した並列ファイルストレージアプライアンスです。ES200NVXの仕様、機能は以下の通りです。



項目	内容
ベースプロダクト /フォームファクター	DDN SFA200NV / 2U
ファイルサービス	EXAScaler (Lustre MDS/OSS)
コントローラ	Active/Active Dual Controller
CPU	2 CPUs (Dual Controller合計)
データ保護方式	Declustered RAID : RAID1 (1+1)、RAID5 (4+1) , (8+1)、RAID6 (4+2), (8+2)
対応ドライブ	1.92TB~15.36TB NVMe SSD
最大ドライブ(スロット)数	24x 2.5inch drive slot
サービスインターフェイス	2x HDR100 Infiniband (100GbE / 50GbE / 40GbE / 25GbE / 10GbE対応)
ストレージキャッシュ(コントローラあたり)	16GB (Battery-Backup)
冗長部品	RAIDコントローラ、電源、冷却ファン
ホットスワップ交換対象部品	RAIDコントローラ、ドライブ、電源、冷却ファン
電源	2x 200V冗長電源
システム管理	CLI、GUI(Web) SNMP Syslog、電子メールによる障害通知

DDN ES400NVX

ES400NVXは、All FlashブロックストレージであるSFA400NVXのストレージコントローラ上でDDN EXAScalerを実装した並列ファイルストレージアプライアンスです。ES400NVXの仕様、機能は以下の通りです。



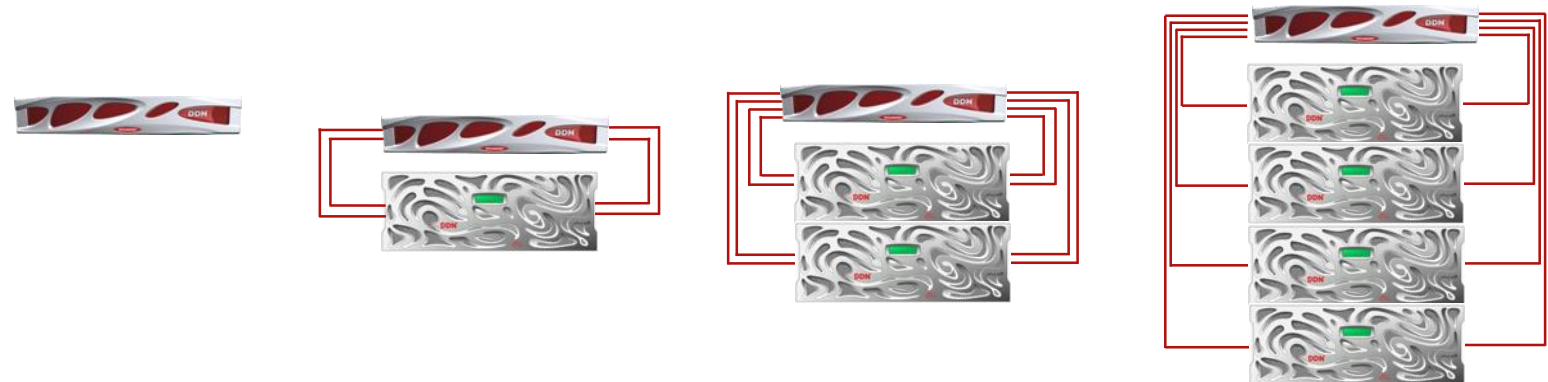
項目	内容
フォームファクター	2U
ファイルサービス	EXAScaler (Lustre MDS/OSS)
コントローラ	Active/Active Dual Controller
CPU	4 CPUs (Dual Controller合計)
データ保護方式	Declustered RAID : RAID1 (1+1)、RAID5 (4+1) , (8+1)、RAID6 (4+2), (8+2)
対応ドライブ	1.92TB~15.36TB NVMe SSD
最大ドライブ(スロット)数	24x 2.5inch drive slot
サービスインターフェイス	4x HDR100 Infiniband(100GbE / 50GbE / 40GbE / 25GbE / 10GbE対応)
ストレージキャッシュ(コントローラあたり)	16GB (Battery-Backup)
冗長部品	RAIDコントローラ、電源、冷却ファン
ホットスワップ交換対象部品	RAIDコントローラ、ドライブ、電源、冷却ファン
電源	2x 80 PLUS Platinum Rated 200V対応冗長電源
システム管理	CLI、GUI(Web) SNMP Syslog、電子メールによる障害通知

ES400NVX CONFIGURATIONS



Configuration	ES400NVXE¥	+ 1 SS9012	+2 SS9012	+4 SS9012
NVMe slots	24	24	24	24
3.5" HDD or SAS SSD Slots	0	90	180	360

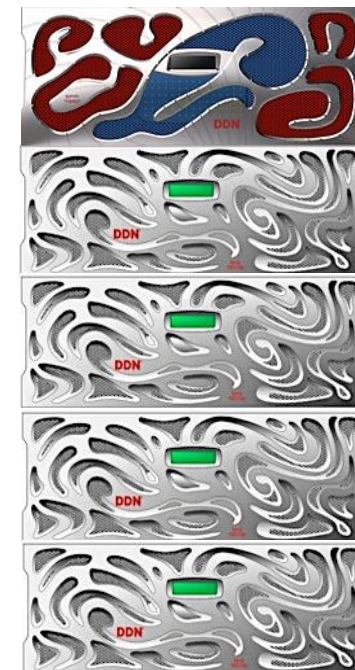
Expansion Options



DDN ES7990X

ES7990Xは、ブロックストレージであるSFA7990Xのストレージコントローラ上でDDN EXAScalerを実装した並列ファイルストレージアプライアンスです。ES7990Xの仕様、機能は以下の通りです。

項目	内容
ベースプロダクト / フォームファクター	SFA7990X / 4U
ファイルサービス	EXAScaler (Lustre MDS/OSS)
コントローラ	Active/Active Dual Controller
データ保護方式	Declustered RAID : RAID1 (1+1)、RAID5 (4+1) , (8+1)、RAID6 (4+2), (8+2)
対応ドライブ	7.2Krpm NL-SAS, 10Krpm SAS, SSD, 混在可能
最大ドライブ(スロット)数	450x 3.5inch drive slot ※4x SS9012構成時 ES7990X: 90x 3.5inch drive slot / SS9012(4set): 360x 3.5inch drive slot
サービスインターフェイス(コントローラあたり)	2x HDR100 Infiniband (100GbE / 50GbE / 40GbE / 25GbE / 10GbE対応)
対応ドライブエンクロージャ / 最大接続数	SS9012 / 4台
ストレージキャッシュ(コントローラあたり)	16GB (Battery-Backup)
冗長部品	RAIDコントローラ、電源、冷却ファン
ホットスワップ交換対象部品	RAIDコントローラ、ドライブ、電源、冷却ファン
電源	4x 200V冗長電源
システム管理	CLI、GUI(Web) SNMP Syslog、電子メールによる障害通知



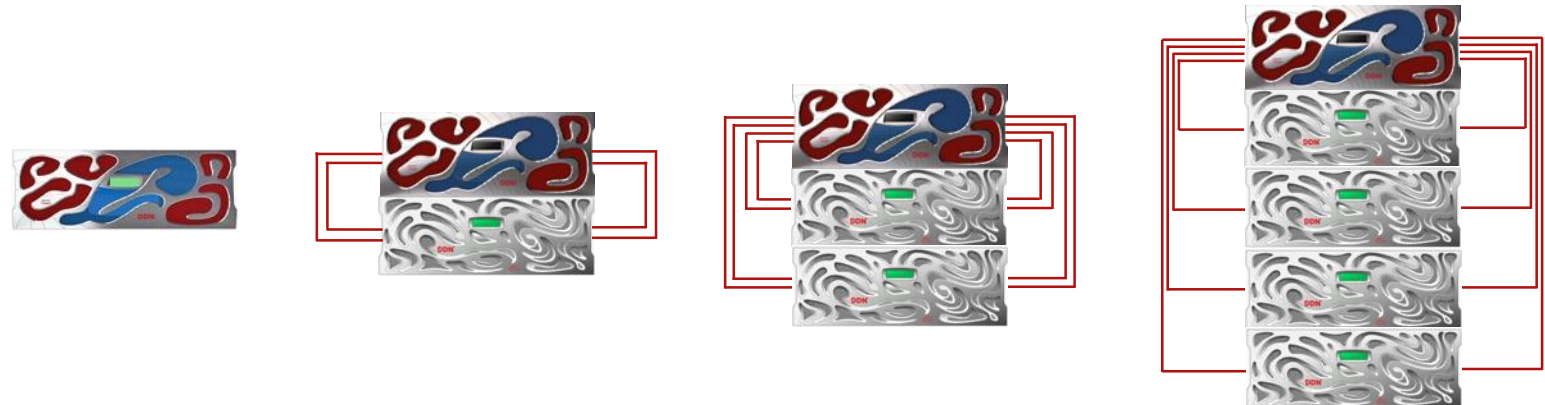
ES7990X CONFIGURATIONS



Configuration		ES7990X Only	+ 1 SS9012	+2 SS9012	+4 SS9012
Supported Solutions	SFA (BLOCK)	✓	✓	✓	✓
	ES7990	✓	✓	✓	✓
	GS7990	✓	✓	✓	✓
3.5" HDD or SAS SSD Slots		90	180	270	450

2019-10-10-Case_131293.rar

Expansion Options



DDN ES18KX

ES18KXは、ブロックストレージであるSFA18KXのストレージコントローラ上でDDN EXAScalerを実装した並列ファイルストレージアプライアンスです。ES18KXの仕様、機能は以下の通りです。



項目	内容
ベースプロダクト / フォームファクター	SFA18KX / 4U
ファイルサービス	EXAScaler (Lustre OSS)
コントローラ	Active/Active Dual Controller
データ保護方式	Declustered RAID : RAID1 (1+1)、RAID5 (4+1) , (8+1)、RAID6 (4+2), (8+2)
対応ドライブ	7.2Krpm NL-SAS, 10Krpm SAS, SSD, NVMe SSD 混在可能
最大ドライブ(スロット)数	1440x 3.5inch drive slot ※16x SS9012構成時 72x 2.5inch drive slot of which 48 are NVMe
サービスインターフェイス(コントローラあたり)	8x EDR Infinibandまたは4x Omni-Pathまたは8x 100GbE
対応ドライブエンクロージャ / 最大接続数	SS9012 / 16台
ストレージキャッシュ(コントローラあたり)	32GB (Battery-Backup)
冗長部品	RAIDコントローラ、電源、冷却ファン
ホットスワップ交換対象部品	RAIDコントローラ、ドライブ、電源、冷却ファン
電源	4x 200V冗長電源
システム管理	CLI、GUI(Web) SNMP Syslog、電子メールによる障害通知

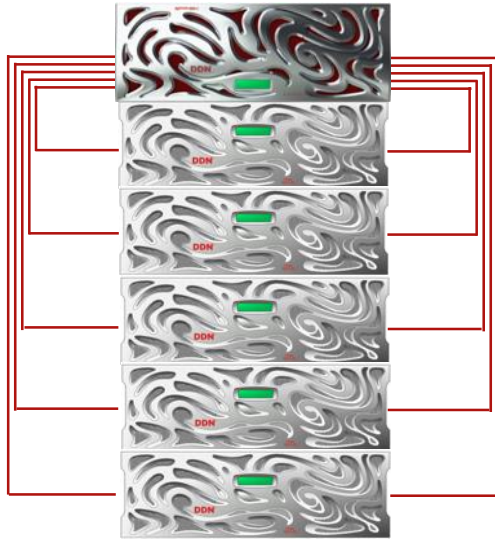
ES18KX expansion Enclosure redundant 1



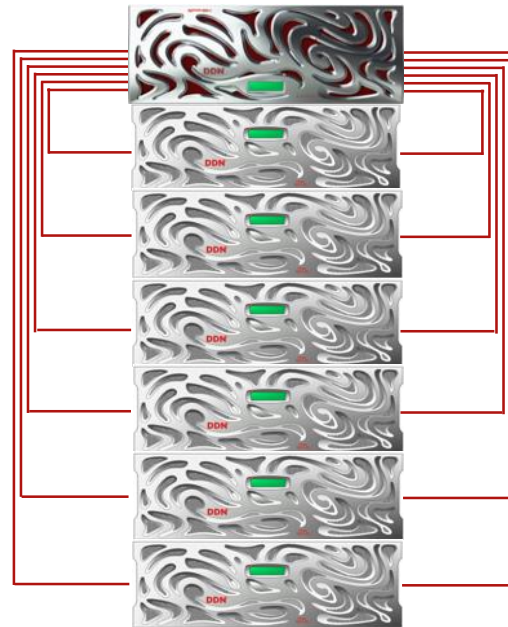
18KX Only
70 Drives



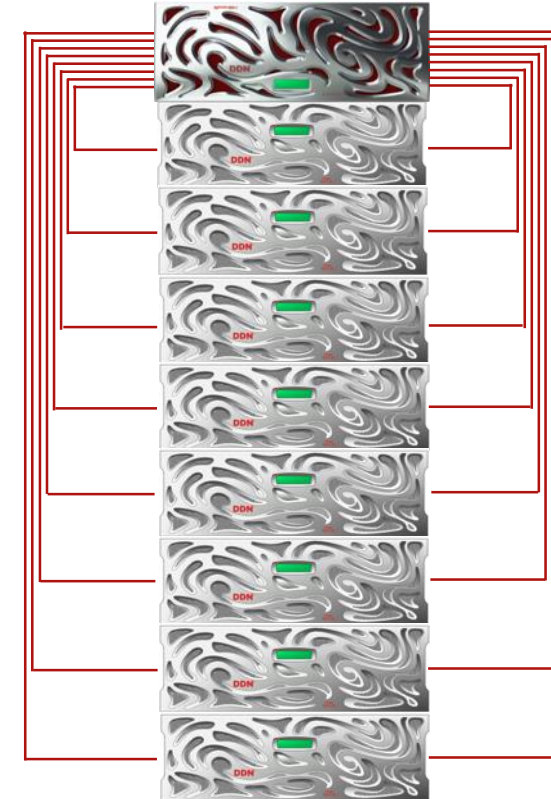
+ 5 SS9012
440 Drives



+6 SS9012
528 Drives



+8 SS9012
704 Drives



Example Config:

8x 55/1

8x 66/1

8x 88/1 or 88/2

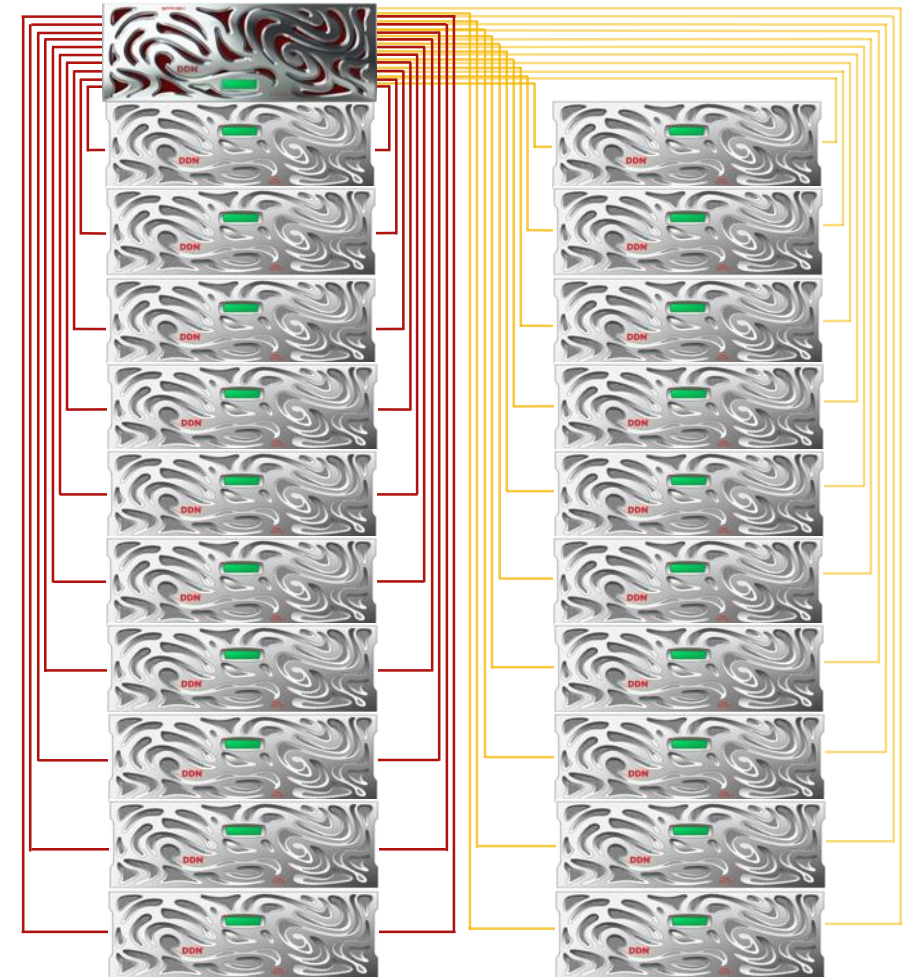
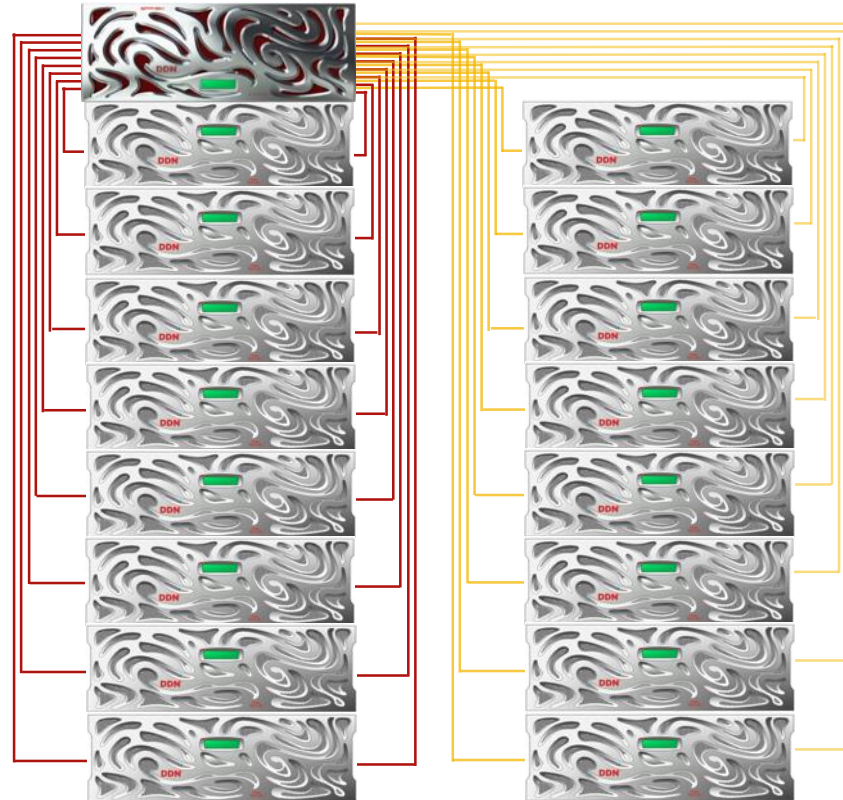
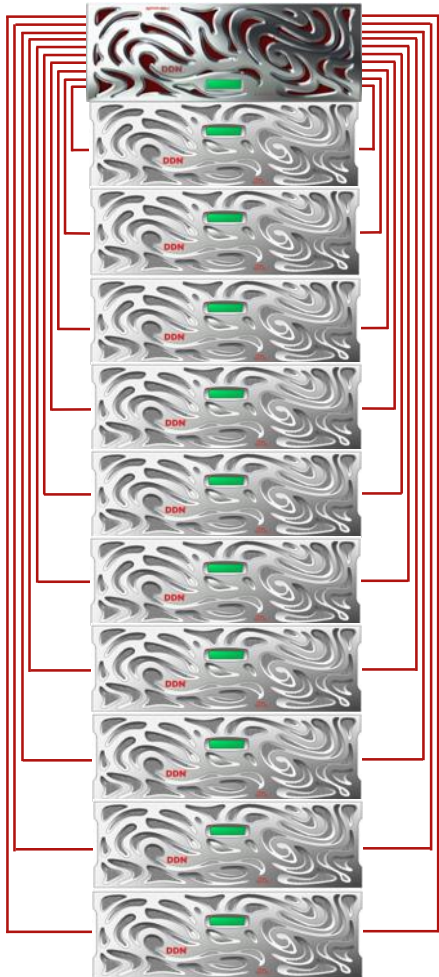
ES18KX expansion Enclosure redundant 2



+10 SS9012 ; 880 Drives; 8x110

+16 SS9012 ; 1408 Drives; 2x 8x88

+20 SS9012 ; 1760 Drives; 2x 8x110

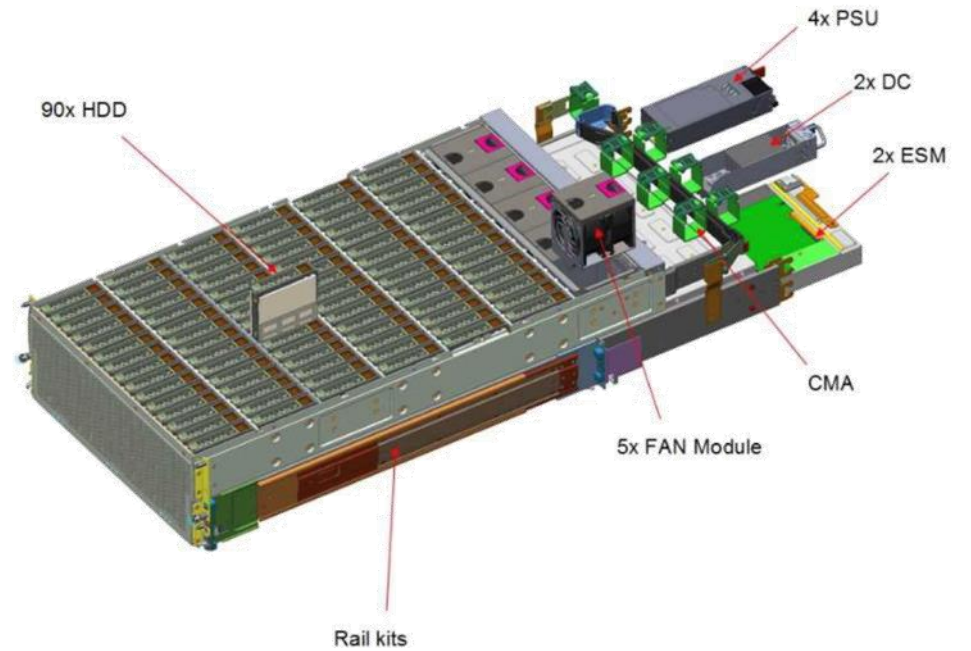


All of these Configurations:

- No daisy-chain (direct attached)
- Expansion Enclosure Redundancy

DDN SS9012

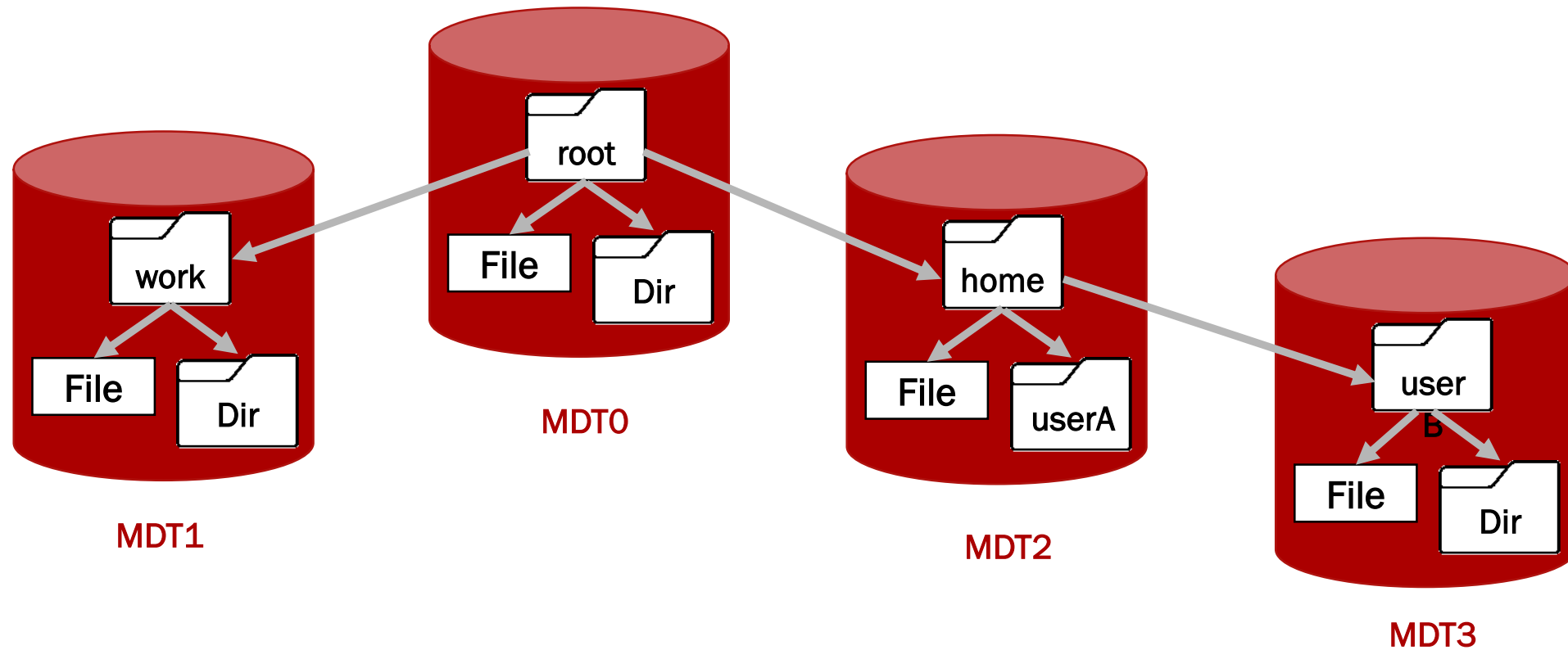
SS9012ディスクエンクロージャは、90本のドライブを収容可能な高密度ストレージエンクロージャです。SSD、SASの各種ドライブを混在可能であり、I/Oパスも含めてコンポーネントが冗長化されております。



項目	製品仕様
シャーシ	4U / 90 Drive
スロット数	90 (3.5inch/2.5inch)
対応ドライブ	SSD, SAS, NL-SAS
冷却ファン	5x 冷却ファン
電源	4 x 1200W 冗長電源 (2+2)
ホットスワップ対象部品	ドライブ、電源、冷却ファン
LCD Display	Status, Power, Environmental Monitoring
LED	Power, Status, Monitoring, Drive Activity

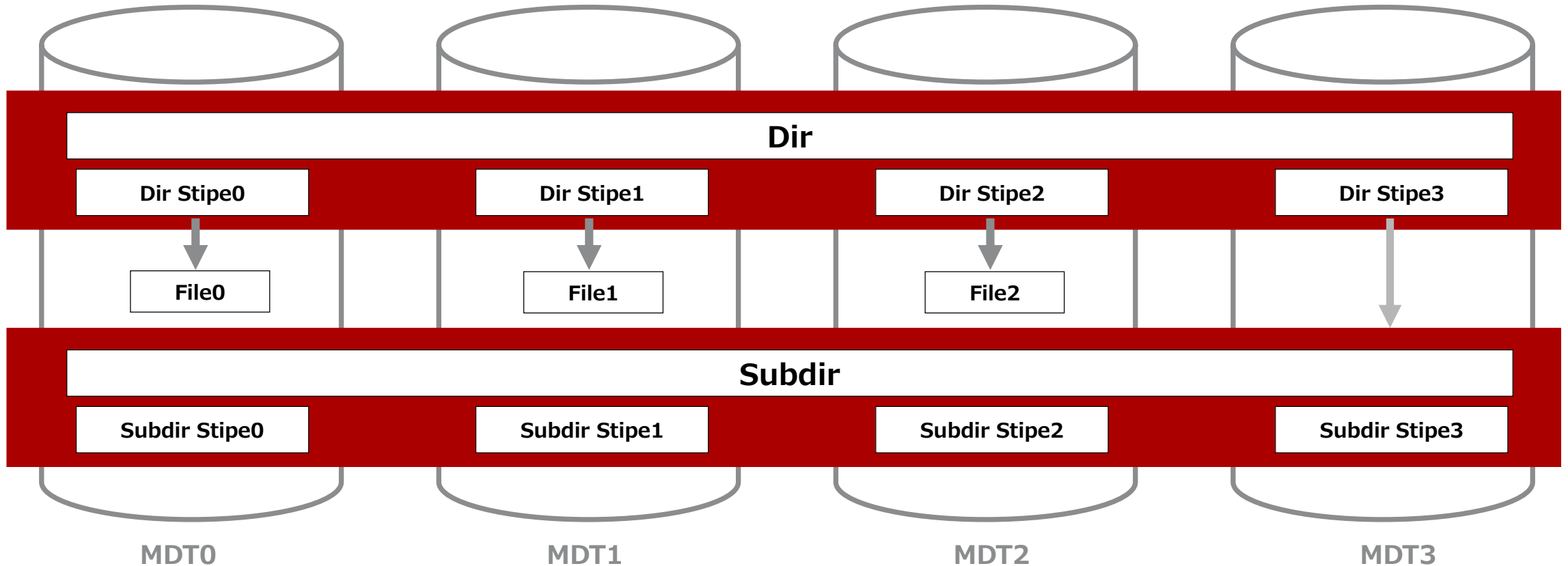
Lustre DNE

Lustre DNE(Distributed Namespace)は一つのNamespaceを複数のMDS/MDTで管理できる機能です。Directory単位でMDTを配置可能とし、複数のMDSをActive/Activeで利用することにより、大規模LustreファイルシステムのMetadata性能をスケールさせます。



Lustre DNE Phase2 Striped Directory

DNE Phase2 Striped Directoryは単一Directoryのメタデータ操作を複数のMDS/MDTに分散することにより複数のMDS、MDTが構成されたシステムにおいて、メタデータ性能の向上を実現します。





DDN A3I

Accelerated, Any-Scale AI

DDN A³I X-APPLIANCES – NEXT GENERATION AI DATA PLATFORMS



AI200X

24r/20w

800K IOPS

32/64/128/256 TB



AI400X

48r/32w

1.5M IOPS

32/64/128/256 TB



AI7990X

24r/20w

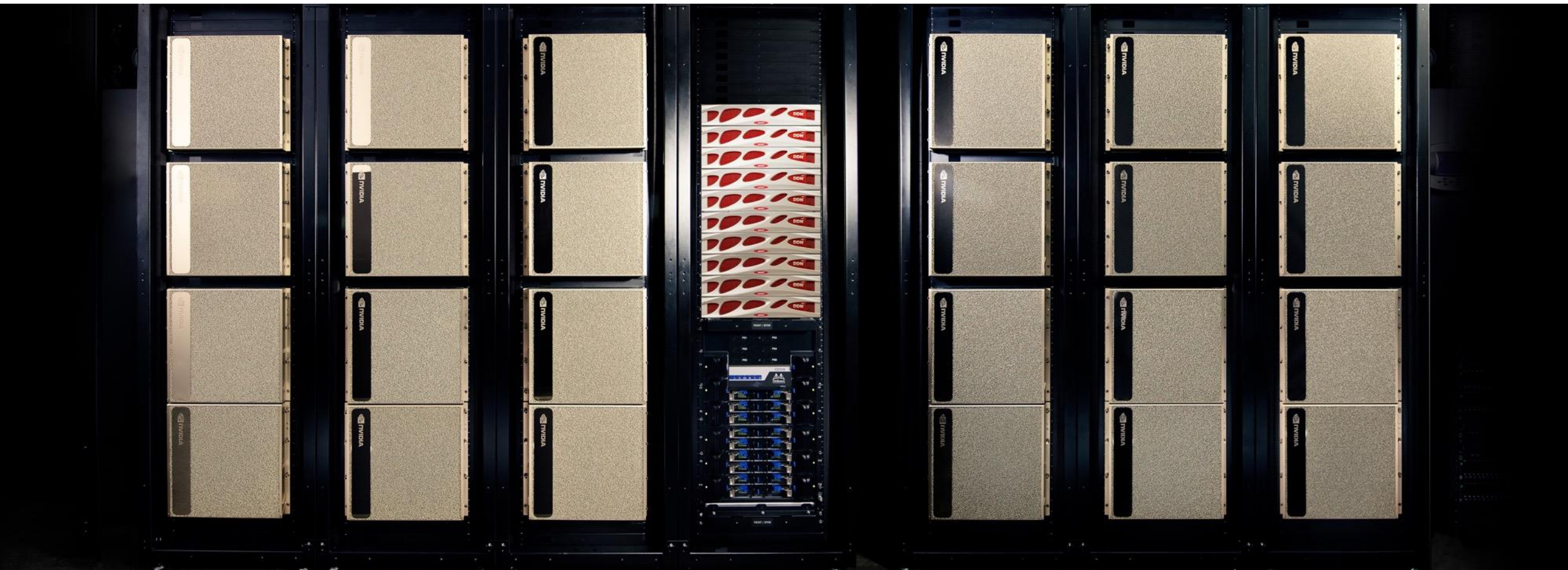
800K IOPS

1/2/4 PB

HDR100/EDR IB OR 100GbE • FULLY INTEGRATED AND OPTIMIZED FOR AI AND DL • VALIDATED UP TO 96 NVIDIA DGX-2s!

DDN AI400X

THE SUPERPOD ACCELERATOR



NVIDIA IS THE AI400X LAUNCH CUSTOMER

10 APPLIANCES

4 HOURS TO DEPLOY

400 GB/s R, 310 GB/s W

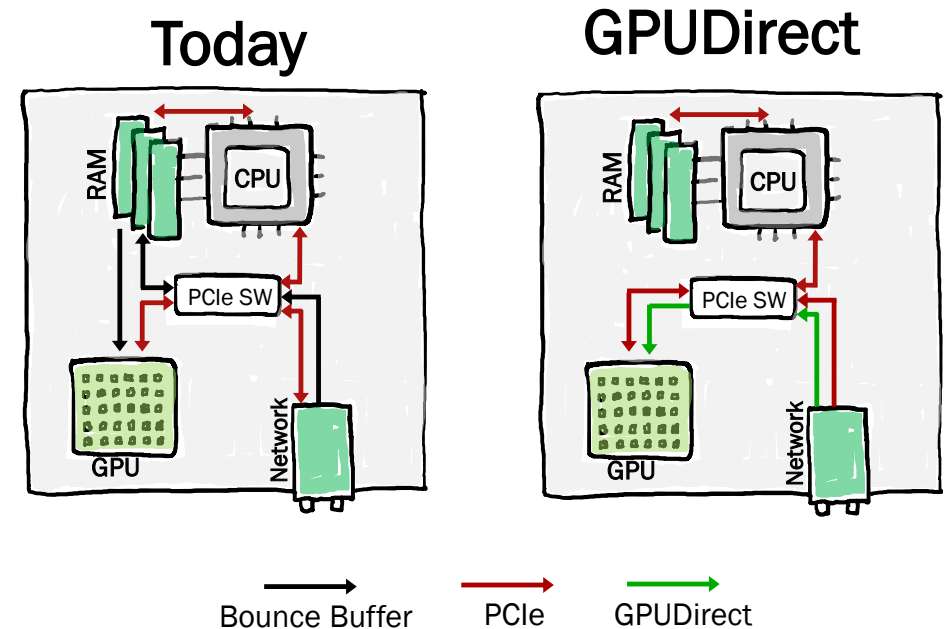
40 AI400X IN 2020 – 15 PB ALL NVME

JOINT IO500 SUBMISSIONS



DDN A³I SOLUTIONS GPUDIRECT TO STORAGE

- ▶ EXAScaler Integration with GPUDirect* enables Direct Memory Access from EXAScaler to GPU memory
- ▶ 2x-8x higher BW data transfers between Storage and GPU.
- ▶ 3.8x lower latency with no faulting and bounce buffers
- ▶ Stable and flat latencies as GPU concurrency increases.
- ▶ Lower consumption of host CPU or memory subsystem
- ▶ The GPU is the computing element with the highest IO bandwidth, e.g. 215 GB/s vs. the CPU's 50 GB/s.
- ▶ Very fast access to petabytes of remote storage faster than even the page cache in CPU memory.



A3I WITH GPUDIRECT – DOUBLE NVIDIA DGX-2 THROUGHPUT!

80 GB/s per client, 20 X performance gains

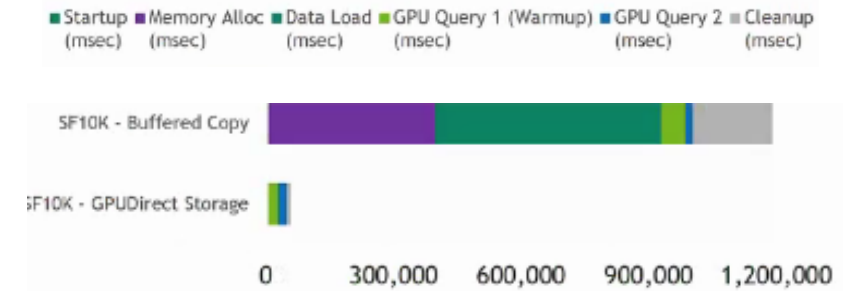
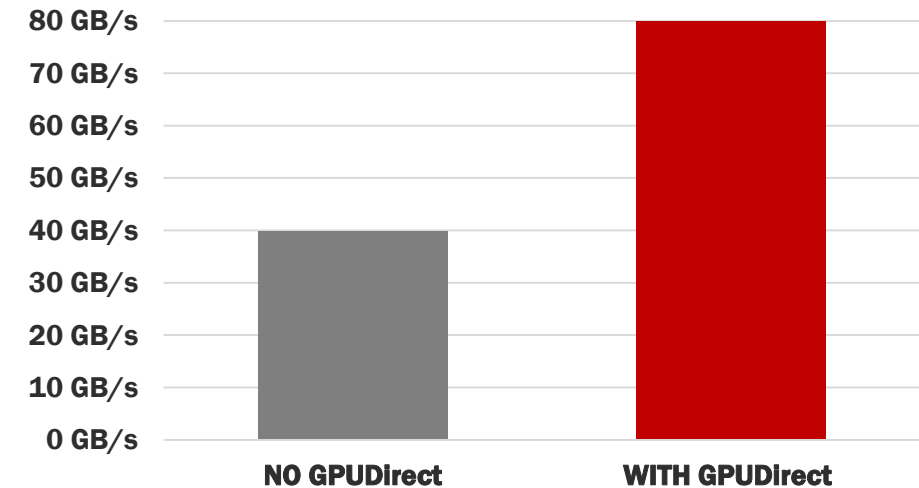
Native client integration with EXA5, fully-transparent to users and applications

Enables a direct path to transfer data between GPU memory and data storage

Eliminates unnecessary memory copies, lowers CPU overhead, reduces latency, bypasses hardware architecture limitations

Improves AI, DL, HPC application performance

GPU READ THROUGHPUT WITH AI400



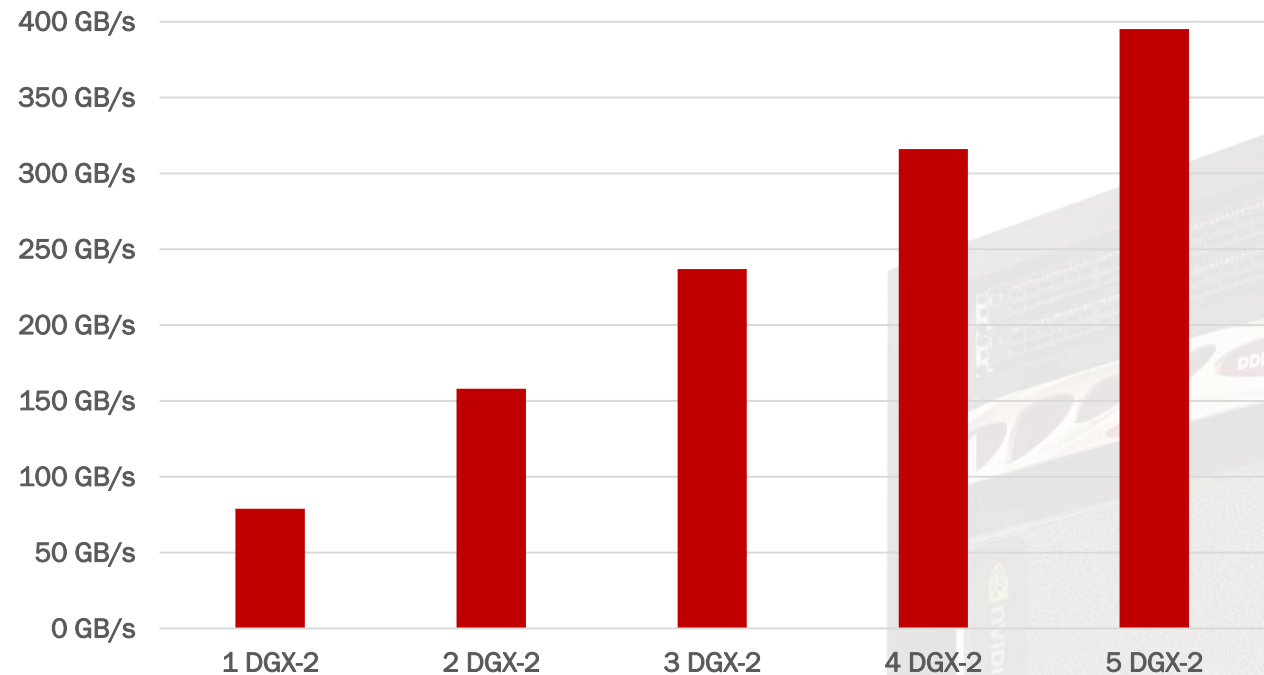
5 DGX-2
10 AI400
400 GB/s

EXA5 GPUDirect integration delivers up to 80 GB/s of throughput per DGX-2

Enables a direct path to transfer data between GPU memory and data storage

Performance scales linearly and provides maximum at-scale application acceleration

GPU READ THROUGHPUT SCALING WITH TEN AI400s



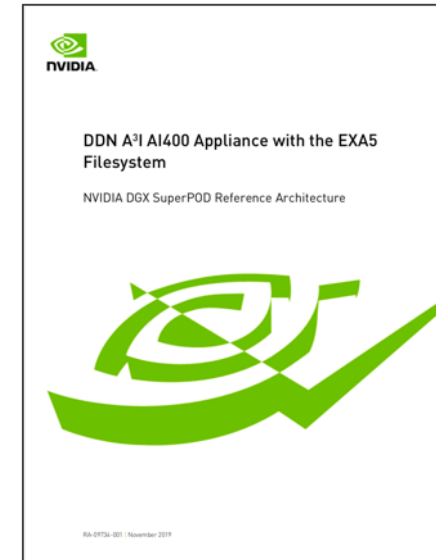
DDN DELIVERS LINEAR PERFORMANCE SCALING

DDN A³I SOLUTIONS – NVIDIA SUPERPOD REFERENCE ARCHITECTURE

▶ AI400 + DGX-2 SuperPOD at-scale testing and validation published by NVIDIA:

- The AI400 All-flash appliance delivers incredible sequential and random read performance, as required by the heaviest DL workloads.
- Metadata performance scales well from 1 to 96 nodes, with no degradation as the number of nodes and threads increases.
- The AI400 is a fully-integrated platform that's easy to deploy. DDN provides excellent technical deployment and support services.

▶ RA document available from NVIDIA website



**NVIDIA DGX-2 SUPERPOD
REFERENCE ARCHITECTURE**

Thank You!

Keep in touch with us.



Team-JPSales@ddn.com



〒102-0081
千代田区四番町6-2 東急番町ビル8F



@DDNJapan



03-3261-9101



company/datadirect-networks