



# DDN Storageのご紹介

Gfarm Workshop 2018 @ 那覇

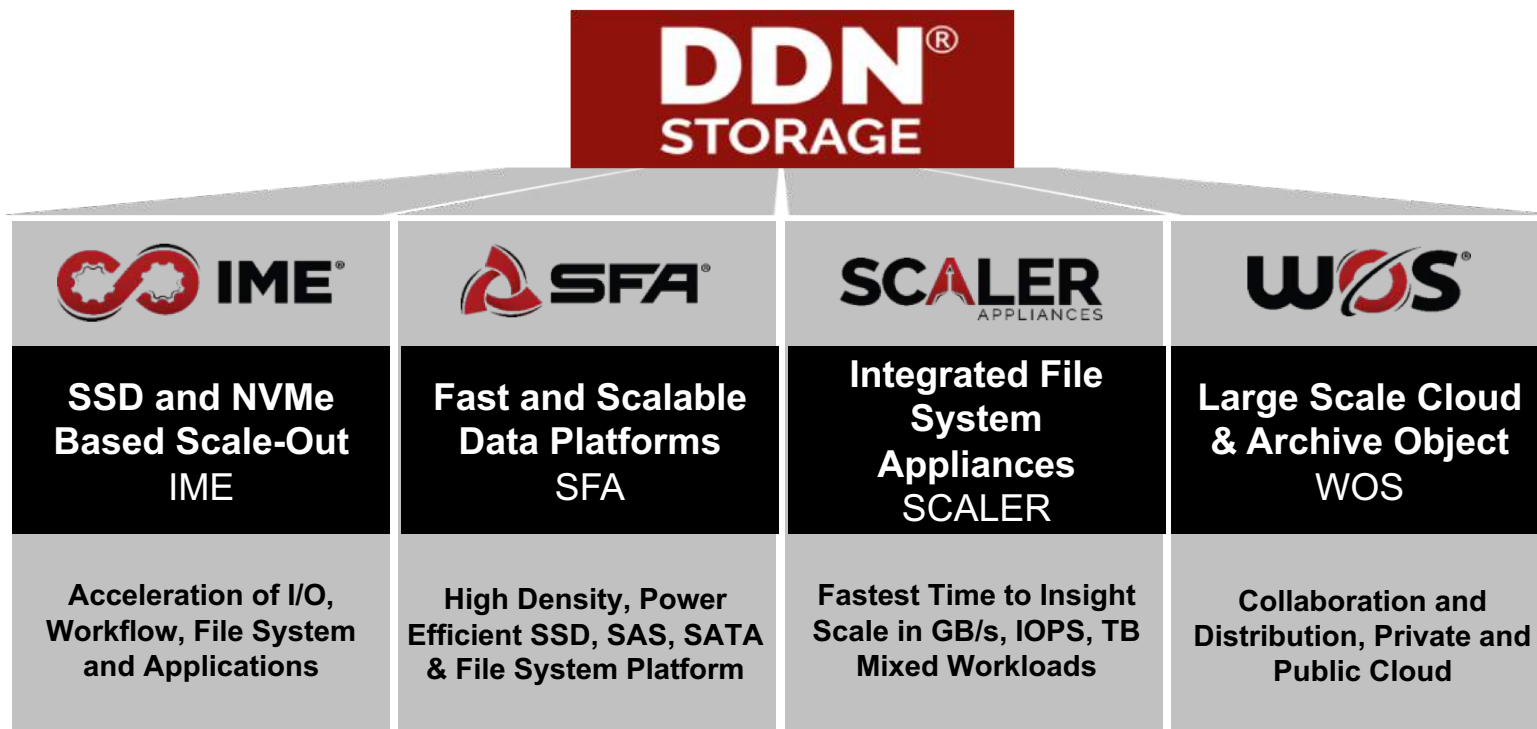
DataDirect Networks Japan, Inc.

橋爪信明

2018/3/2

# DDN紹介

- HPCを中心にEnd-To-End Storage Solutionを開発・製造・販売しています



# DDNご紹介

設立： US本社 1998年  
日本法人 2008年

所在： 本社カリフォルニア・チャッツワース  
日本法人 東京/大阪

事業内容： ストレージソリューションの  
設計・構築・保守・販売

戦略パートナー：  
HPE・IBM・Intel など

従業員： 650名以上 / 20カ国  
30名 / 日本法人

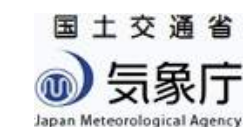
製造拠点： US カリフォルニア  
チャッツワース

特長： 高性能ストレージシステムインテグレーション

販売パートナー：  
SCSK・日本コムシス・CTC・Thymos など



# 主な国内導入実績 - 1





## 主な国内導入実績 - 2



U-NEXT

日テレ

nottv

SONY

NTT SmartConnect

フジテレビ

SBS  
TV

COMPUTER  
ENTERTAINMENT®

SAMSUNG

MBS

IMAGICA








Sony PCL

ABC 朝日放送

McRAY  
corporation

Rtec

# DDN 製品ポートフォリオ

Burst Buffer	
<p><b>Infinite Memory Engine (IME)</b>            クライアントと並列ファイルシステムの間Layerに位置づけられるBurst Bufferであり一時的なBurst IOに対応し、多数プロセスからShared Fileへの小さなランダムIOアクセスなど並列ファイルシステムが不得意としているI/Oを高速化。アプライアンスおよびソフトウェアによる提供。</p>	
 <b>IME</b> IME Software	 <b>IME14K</b> <b>IME240</b>
File System	
<p><b>DDN EXAScaler</b></p> <ul style="list-style-type: none"> <li>オープンソースのLustreファイルシステムをベースに構成</li> <li>膨大な数のクライアントに対し高スループットIO環境を提供</li> <li>多数のHPCサイトで多くの実績</li> </ul>	<p><b>DDN GRIDScaler</b></p> <ul style="list-style-type: none"> <li>高いIO性能に加えてマルチプロトコルアクセス環境、Snapshot、Replication、DR、サイト間連携など多彩な機能を提供</li> <li>低速ストレージ、テープ装置等を用いた階層化機能の実現</li> </ul>
Storage Platform	
<p><b>SFA14KX/ES14KX/GS14KX</b></p>  <p>54GB/sec / 1.7M IOPS Max 912 Drive</p>	<p><b>SFA7700X/ES7K/GS7K</b></p>  <p>12GB/sec / 450K IOPS Max 396 Drive</p>
<p>※ ES14KX/GS14KX/ES7K/GS7KはRAIDコントローラ上でEXAScalerまたはGRIDScalerによるファイルサービスを提供するアプライアンス製品となります。</p>	
Drive Enclosure / Drive	
<p><b>SS8460/8412/8462</b></p>  <p>6/12Gbps 84Drive Enclosure</p>	<p><b>SS7700</b></p>  <p>6Gbps 60Drive Enclosure</p>
<p><b>Drive</b></p>  <p>7.2K CapSAS 10,15K Perf SAS SSD, NVMe SSD</p>	

# DDN | IME Solves Critical I/O problem

<https://www.vi4io.org/io500/start>

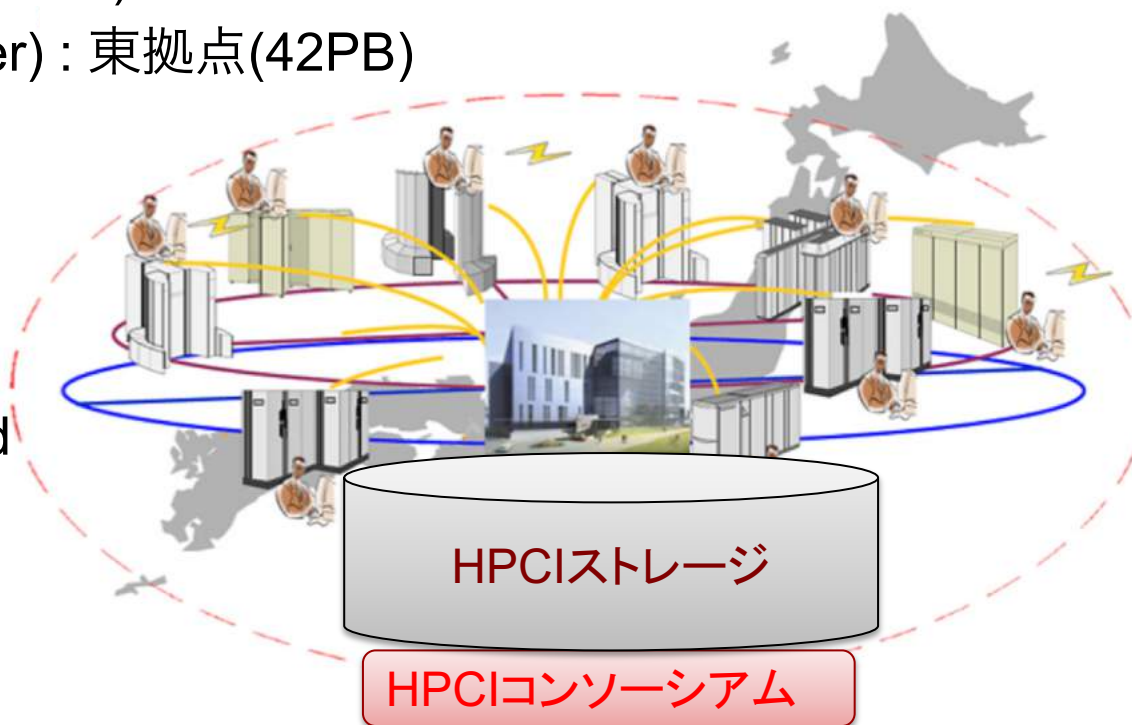
#	information					io500	ior			
	system	institution	filesystem	storage vendor	client nodes	score	easy write	easy read	hard write	hard read
							GiB/s	GiB/s	GiB/s	GiB/s
1	Oakforest-PACS	JCAHPC	IME	DDN	2048	101.48	742.38	427.41	600.28	258.93
2	Shaheen	Kaust	DataWarp	Cray	300	70.90	969.45	894.76	15.55	39.09
3	Shaheen	Kaust	Lustre		1000	41.00	333.03	220.62	1.44	81.38
4	JURON	JSC	BeeGFS	ThinkparQ	8	35.77	30.42	48.36	1.46	19.16
5	Mistral	DKRZ	Lustre	Seagate	100	32.15	158.19	163.62	1.53	6.79
6	Sonasad	IBM	Spectrum Scale	IBM	10	21.63	34.13	32.25	0.17	2.33
7	Seislab	Fraunhofer	BeeGFS	ThinkparQ	24	18.75	18.79	22.34	0.89	1.86
8	EMSL Cascade	PNNL	Lustre		126	11.17	17.81	30.19	0.39	2.72
9	Serrano	SNL	Spectrum Scale	IBM	16	4.25	1.08	1.03	0.22	0.71

IOR Hard : A single shared file, 47008 byte I/O size and Stride I/O

IOR Easy : File Per Process, 2MB I/O size and Sequential I/O

# DDNとGfarmの関わり

- HPCI (High Performance Computing Infrastructure)
  - ▶ HPCI共用ストレージとして2012年度から採用
  - ▶ SFA10K : 西拠点(10PB), 東拠点(12PB)
  - ▶ SFA12K : 東工大拠点(0.4PB)
  - ▶ SFA14KXE(GfarmScaler) : 東拠点(42PB)
- Gfarm
  - ▶ XFSを提供
  - ▶ SFA10K, 12K : Block
  - ▶ SFA14KXE : Embedded



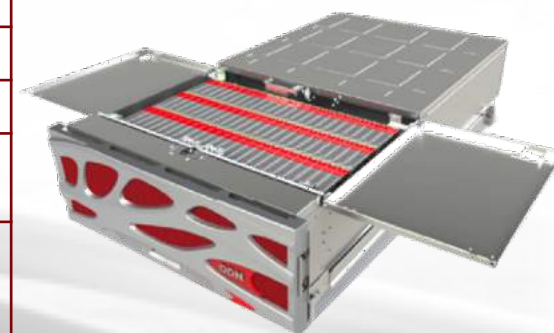


# GfarmScalerとは

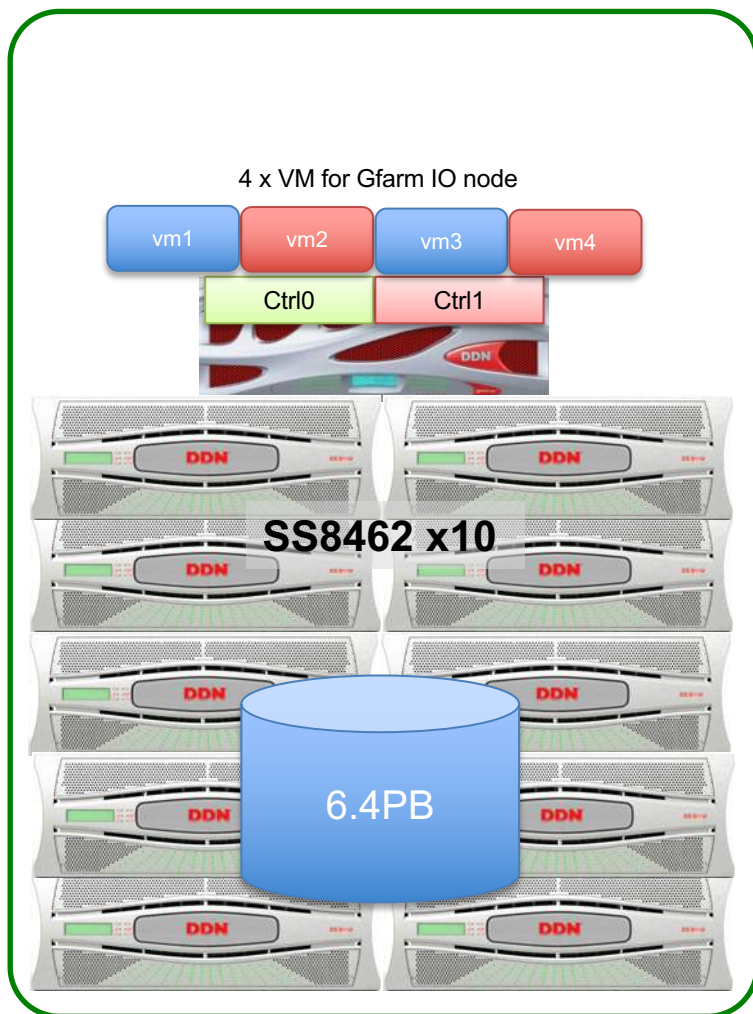
- HPCI東拠点・西拠点更新調達で提案
  - ▶ サーバを減らすことによるコスト削減
  - ▶ 性能は維持
- SFA14KXE、SFA7700XE上でGfarmを提供
  - ▶ データストレージ用に外部サーバ必要なし
  - ▶ SFA14KXE、SFA7700XEのコントローラ上でメタデータ以外のサービスを提供
- Gfarmバックエンド：高速で大容量のXFS環境
  - ▶ 高スループットXFS
  - ▶ 大容量XFS

# DDN SFA14KXE 概要

項目	製品仕様
フォームファクター	4U (Dual ControllerおよびBBU、2.5inch Drive slot x72を内蔵)
コントローラ	Active/Active Dual Controller
対応RAIDレベル	RAID1(1+1)、RAID5 (8+1 or 4+1)、RAID6(8+2 or 4+2)
最大理論性能	シーケンシャル Read/Write 40GB/s
最大ドライブ数	840x 3.5inchドライブ ※10x SS8462構成時 48 x PCIe NVMe SSD or 72 x 2.5inch SAS SSD
対応ドライブ	7.2Krpm NL-SAS (12Gps SAS 4Kn) 10Krpm SAS (12Gps SAS 4Kn) SSD (12Gps SAS 4Kn) 混在可能
キャッシュ	32GB / Controller (Battery-Backup DDR4 SDRAM)
ホストインターフェイス	4x EDR Infiniband / Controller 2x Intel Omni-Path FC / Controller
バックエンドポート	20x SAS 3.0
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、電源、冷却ファン、ドライブ
ソフトウェア機能	LUNマッピングおよびマスキング、Read QoS、ポートゾーニング、Direct Protect
管理機能	Web GUI管理、CUI管理、メールベースの障害発報、SNMP



# 東拠点でのXFS性能(実測値)



- 7x GfarmScaler(SFA14KXE)
- Single SFA building block
  - ▶ 10x SS8462
  - ▶ 770x HGST 10TB Disk
  - ▶ 8x XFS filesystem (2x XFS per VM)
  - ▶ Aggregate XFS performance from VM
    - Write 31.0GB/sec
    - Read 31.9GB/sec
    - +20% Over achieved against performance target!
- Scalability and manageability
  - ▶ Reduce IO stack and Gfarm talks to SFA RAID engine directly
  - ▶ Online SFA firmware upgrade for robust Gfarm services

# Performance Summary

- Performance scales linearly
  - ▶ Single SFA14KXE controller (400 x NL-SAS)
    - 1 x VM: Max 7.5GB/sec (Write), 8.4GB/sec (Read)
    - 2 x VM: Max 15.2GB/sec (Write), 17.4GB/sec (Read)
  - ▶ Dual SFA14KXE controllers (400 x NL-SAS)
    - 4 x VM: Max 29.3GB/sec (Write), 33.4GB/sec (Read)
  - ▶ Dual SFA14KXE controllers (770 x NL-SAS)
    - 4 x VM: Max 31.0GB/sec (Write), 31.9GB/sec (Read)

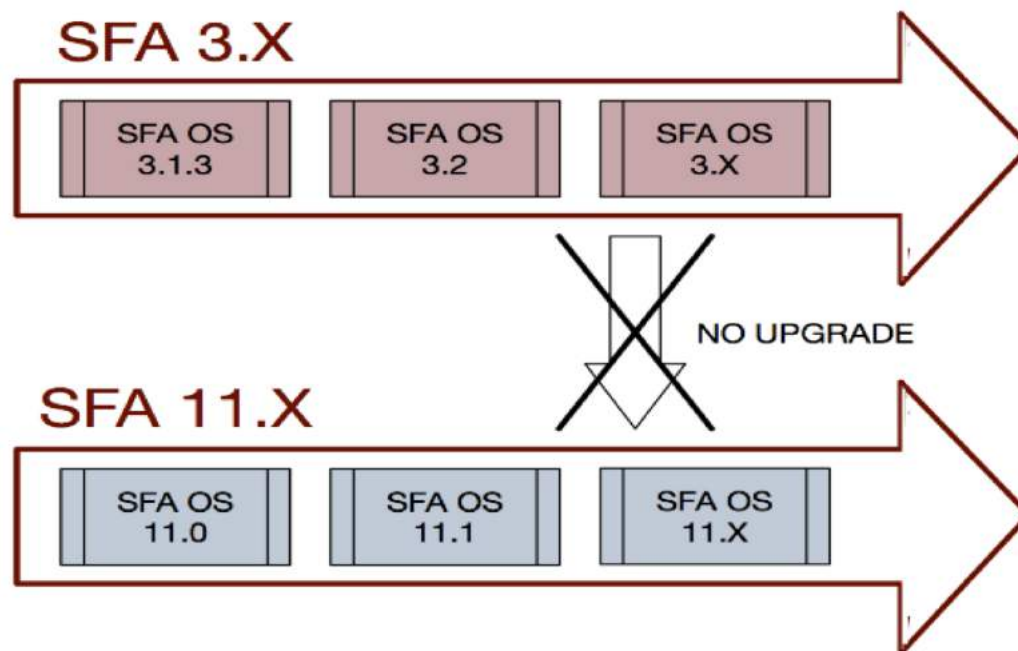


2018年に向けて



# New SFAOS - Product Positioning

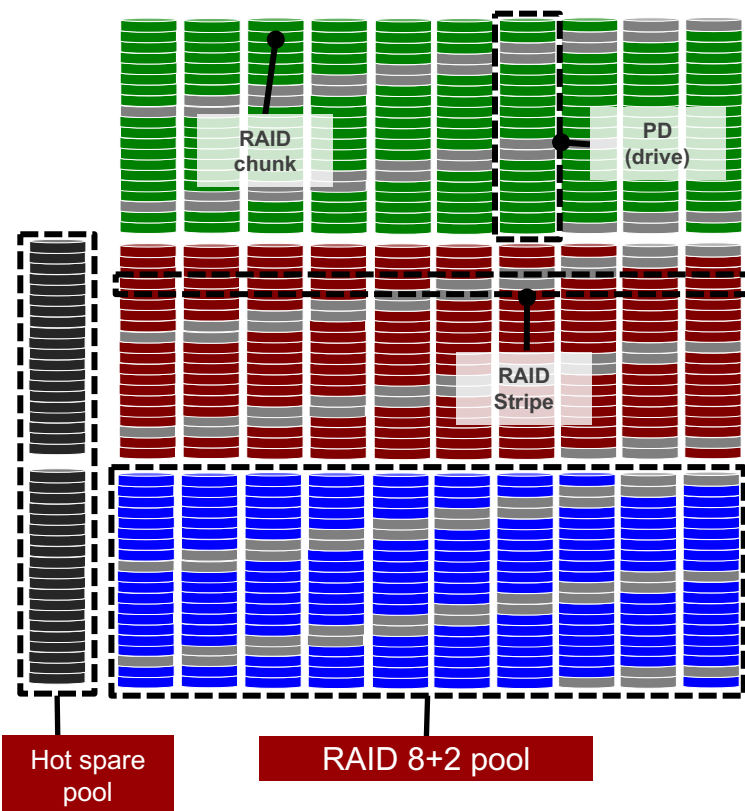
- SFAOS11.X は新しいOSであり、従来のSFAOS 3.Xのアップグレード版ではありません
  - ▶ 11.X – Declustered RAID architecture (DCR)
  - ▶ 3.X – Traditional RAID architecture (“Legacy SFA”)
- 14KXの後継製品ではSFAOS 11.Xのみサポートされます
  - ▶ SFA14KXはSFAOS3.X, SFAOS11.Xのいずれもサポートします



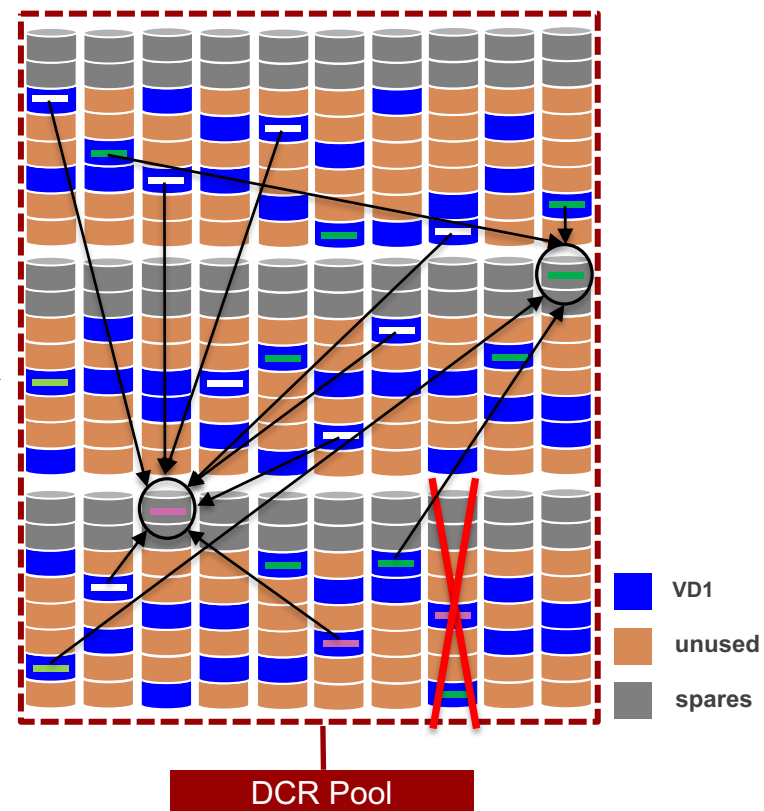
# DCRの特徴

- ディスク障害時の復旧時間の大幅な短縮
- 単一LUN性能の劇的な向上
- 大容量LUNを構成可能(数百ディスクを使用したLUN)
- 構成の複雑さを逡減
  - ▶ 少数の冗長グループ/LUN、プレゼンテーション/マッピング、ファイルシステムターゲット
- HCAの性能に迫るシングルスレッドIO性能
- 将来提供予定
  - ▶ 8+3, 16+3などの冗長性の向上

# DCR概要



- Support Large LUN
- Single LUN's performance improvement
- Full PCIe/SAS Bandwidth
- Disk enclosure-aware DCR
- Reduce Rebuild Time



# DCR Rebuild例

Minutes from failing disk to regaining +2 redundancy

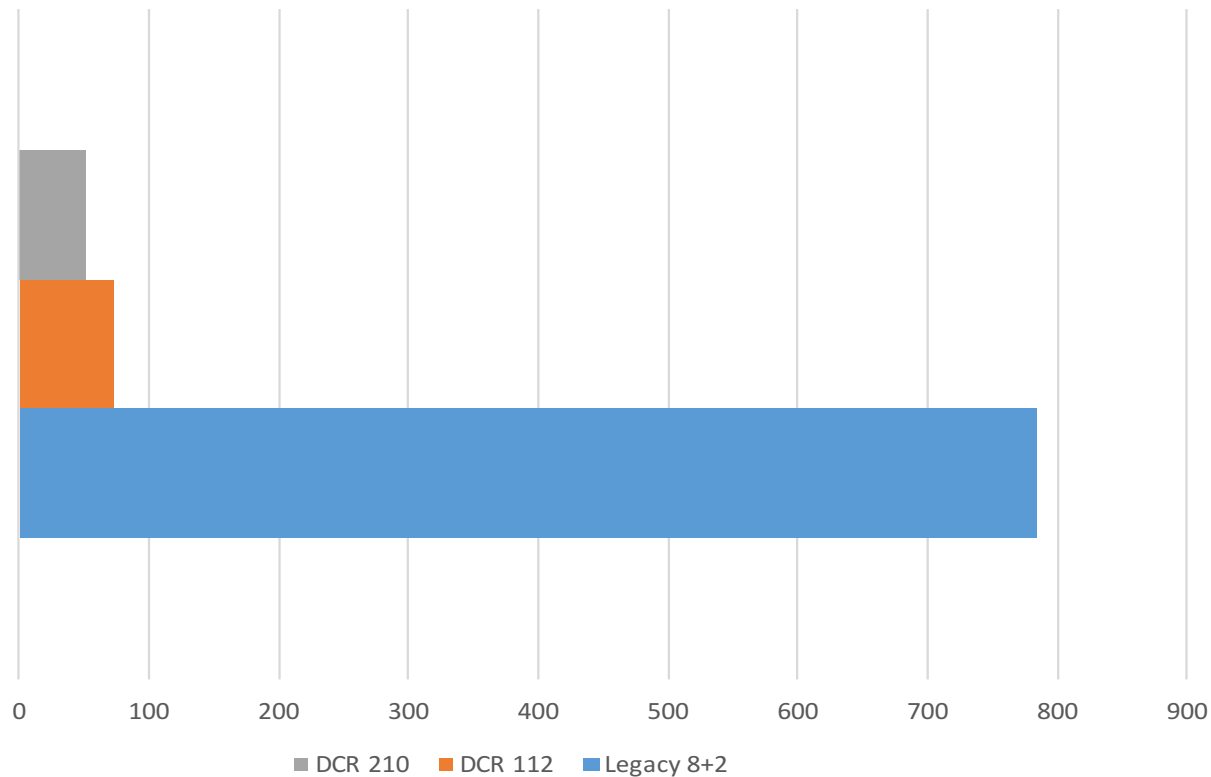
8+2 rebuild vs DCR rebuild in 112 disk pool and 210 disk pool

DCR 210 disk pool - 591TB\*

DCR 112 disk pool – 318TB\*

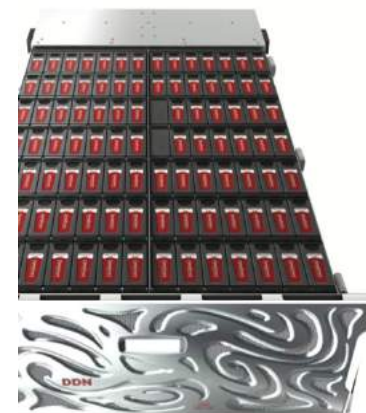
Legacy 8+2 pool – 37TB\*

\*= Size of VD/LUN; 4TB Disks



# New HW in 2018

- SFA14KX後継機
  - ▶ Intel Purley(SkyLake/CascadeLake) base platform
  - ▶ Highend : 60-75GB/s FS throughput
  - ▶ Midrange : 15-20GB/s FS throughput
- 2U24
  - ▶ NVMe only SFA
- SS9012
  - ▶ Full 12Gbps SAS Disk Enclosure
  - ▶ 90 Drives in 4U
  - ▶ **900** direct attached drives per system with 14KX
    - 10.8PB physical, 8.64PB usable capacity with 12TB drive
  - ▶ **1440** direct attached drives per system with 14KX後継機
    - 17.28PB physical, 13.82PB usable capacity with 12TB drive





## DDN | DDN for 2018

- Continue to investment in Highly Scalable Flash Storage Solution
  - ▶ Enhancement and Performance Improvements of IME
  - ▶ Price and Performance efficient All Flash Storage H/W and S/W
  - ▶ Mixed HDD and SSD/NVMe configuration with Parallel Filesystem
- Data Management
  - ▶ Data Security/Integrity and Multi tenancy
  - ▶ Data locality and Archiving

# Thank You!

Keep in touch with us



Team-JPSales@ddn.com



Tokyu Bancho Bldg. 8F  
6-2 Yonbancho Chiyoda-ku,  
Tokyo 102-0081



@ddn\_limitless



+81-3-3261-9101  
+81-3-3261-9140



company/datadirect-networks