AIからアーカイブまでユニバーサルNAS ストレージのご紹介
VAST Data

2021年2月

ノックス株式会社
小幡　明広

VAST

全てのアプリケーション用のシングルデータプラットフォーム

新しいアーキテクチャの活用

新しいテクノロジを組み合わせ、これまで考えられなかった方法でゼロから
新しいタイプのストレージアーキテクチャを発明

NVME OVER FABRICS

QLC FLASH

ストレージクラスメモリ

オールフラッシュストレージ

わずか5Uにペタバイトデータを格納

物理容量	購入単位	最小システムユニット
675TB～	100TB単位	5U

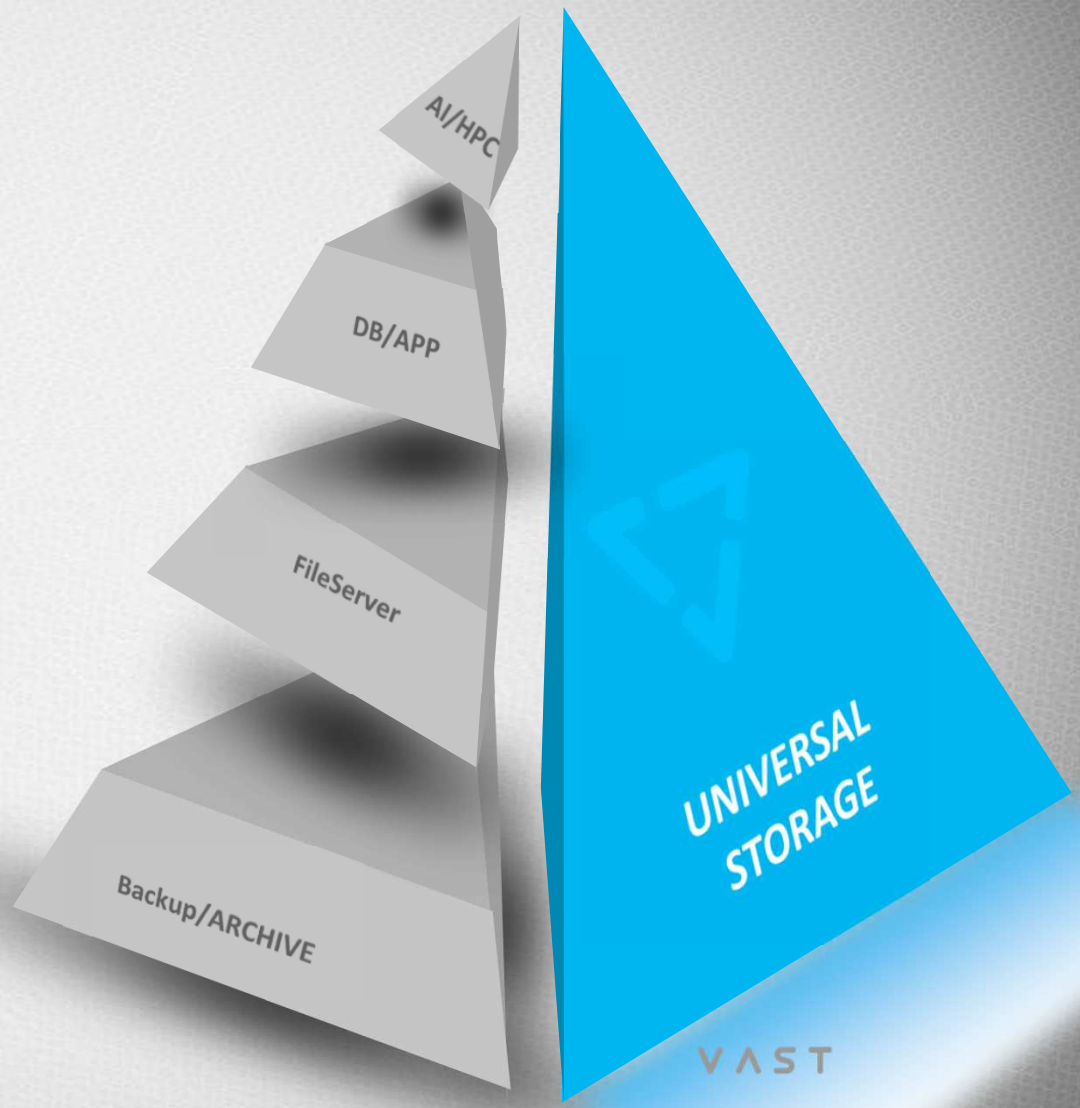
VAST Dataとは

ミッション

ストレージ階層化の排除

NO MORE TIERS

何十年にも渡るストレージの複雑さとアプリケーションのボトルネックを終わらせること。
VASTは、一連のイノベーションを組み合わせ、すべてのデータとすべてのアプリケーションに対するフラッシュのイノベーションを起こします。



VAST Data 会社＆チーム

2016年に設立。2019年2月製品出荷。本社はニューヨーク、開発拠点はイスラエル。

JEFF DENWORTH
VP, PRODUCTS AND MARKETING, CO-FOUNDER
DDN STORAGE

MIKE WING
PRESIDENT
DELL EMC

SHACHAR FIENBLIT
VP, R&D AND CO-FOUNDER
kaminario

RENEN HALLAK
CEO AND FOUNDER
XtremIO

AVERY PHAM
VP OPERATIONS
PURESTORAGE®

タイムラインと企業価値

会社設立から5年、製品出荷からわずか2年で企業価値37億ドルの企業へ

2016 会社設立
シリーズA
1,500万ドル獲得

2017

2018
シリーズA
2,500万ドル獲得

2019 ステルス解除
製品出荷
シリーズB
4,000万ドル獲得

2020 企業価値12億ドル
1億ドルを調達
シリーズC
1億ドル獲得

2021 企業価値37億ドルへ
NEW!
シリーズD
8,300万ドル獲得

Reuters - VAST valuation triples to $3.7 bln after Tiger Global-led investmentSource
Captured: 04/05/2021 11:37:08
Bloomberg - Software Storage Firm Vast Data Valued at $3.7 Billion in RoundSource
Captured: 04/05/2021 11:37:08
C-Tech - VAST Data raises $83 million series D to triple valuation to $3.7 billion in just one yearSource
Captured: 04/05/2021 11:37:08

VAST

nox

アワード

TechTarget エンタープライズ分野で2年連続Gold Winnerを受賞するなど高い業界評価

Tech Trailblazers

WINNER

2020 Storage Trailblazers Winner

TechTarget

GOLD

Best Enterprise Storage Array – 2019 & 2020

GIGAOM

LEADER

2020 Radar For File And Object Storage

451 Research

WINNER

2020 Firestarter Award

coldago research

LEADER

2020 Research Map For File Storage

Flash Memory Summit

BEST IN SHOW

2020 Most Innovative Flash Memory Technology

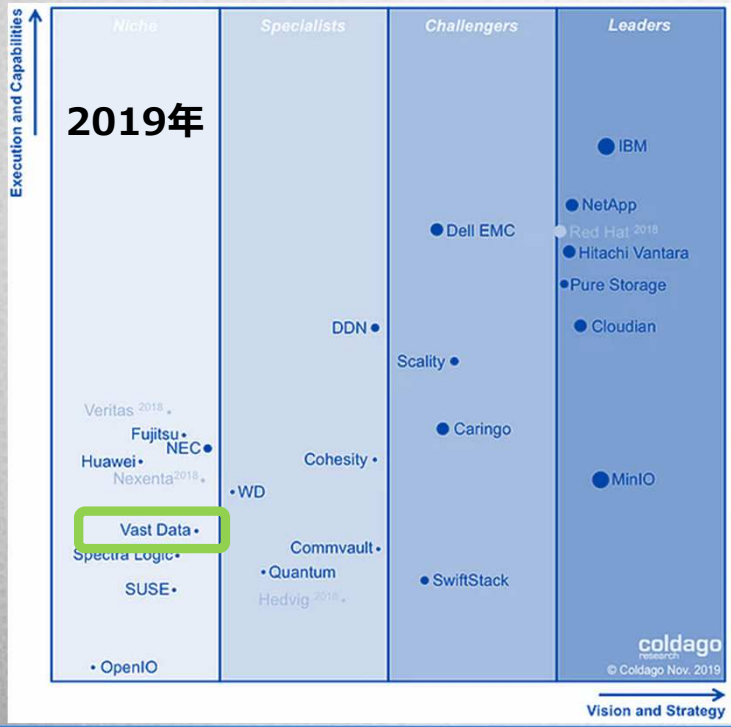
Gartner

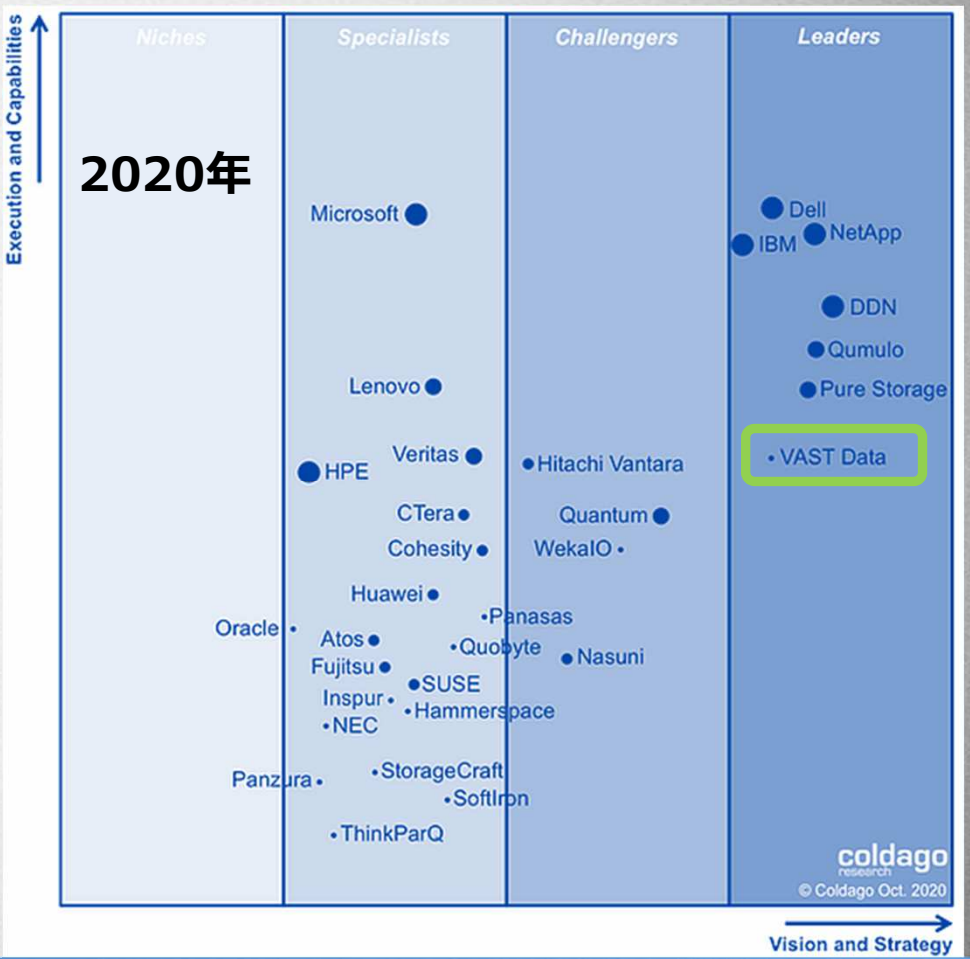
COOL VENDOR

First 'Cool Vendor' In Data Infrastructure Since 2018

2020年 リーダーポジションへ

2020年10月、HPC、エンタープライズ
ストレージベンダーマップ





導入先

業種・業界問わず幅広い分野で導入

PIXAR ANIMATION STUDIOS

verizon

TESLA

NHL

YAHOO!

Goldman Sachs

バイオテック企業

米国防衛大手企業

メディカル企業

カスタマー

US連邦政府機関への導入

国防総省、エネルギー省、国立衛生研究所、海洋大気庁、航空宇宙局、退役軍人省など、多くの連邦機関に導入され、わずか2年で100PB以上の規模でVASTが利用されています。

エネルギー省

航空宇宙局

海洋大気省

国防総省

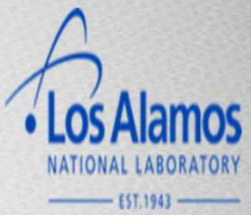
退役軍人省

保険福祉省

インテリジェンス・コミュニティー

カスタマー

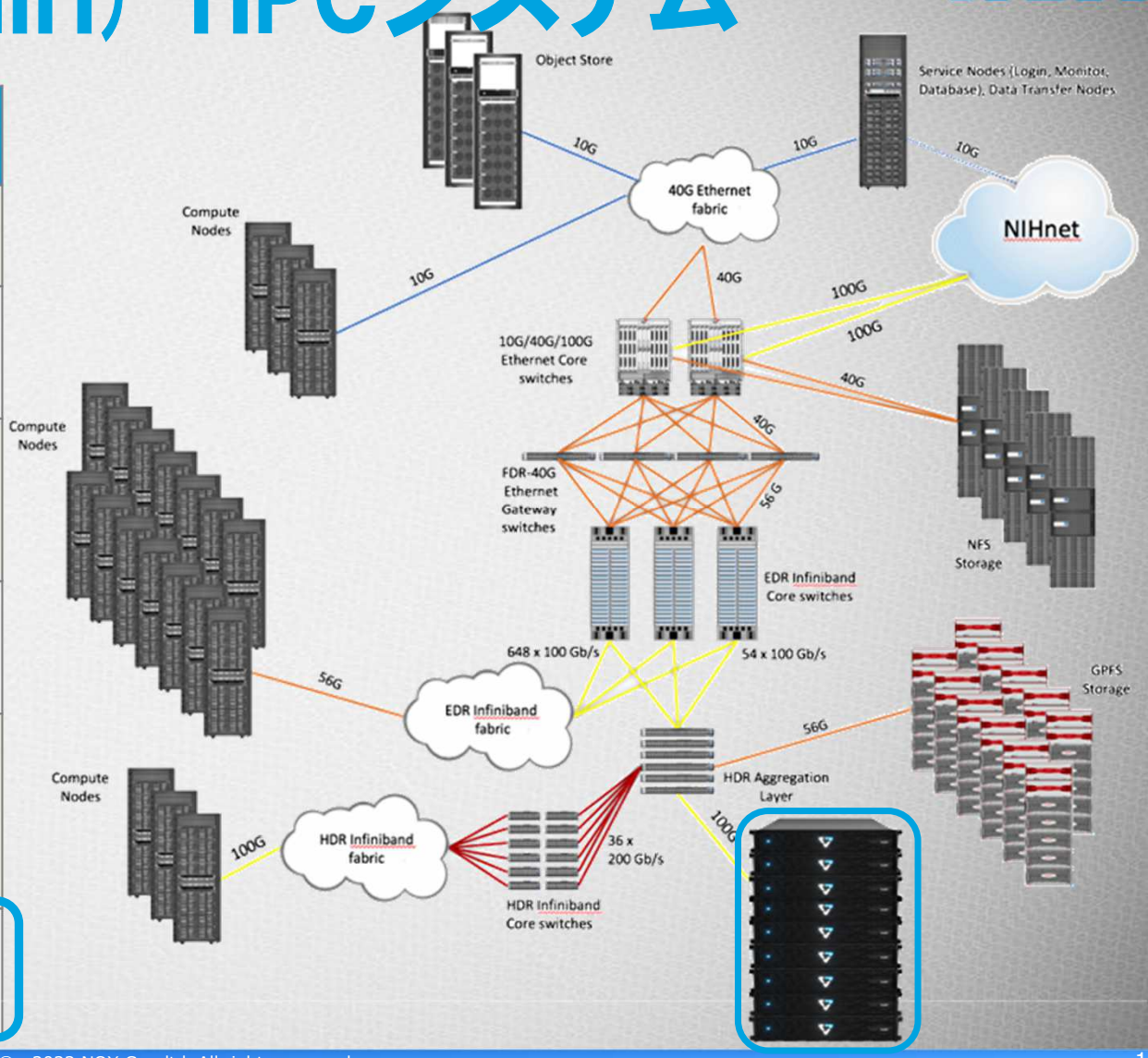
HPC分野



Success Story
- 従来FileSystemと比べシンプルなNAS出来て良かった
- データ削減効果が75%データ削減された
- 1日1枚だったユーザートラブル・チケットが不要になった
- HPCコードの実行が最大20倍高速化
- VASTの1クラスタで、home/scratch/trainingをサポート 22種類のクラスタに容易にマウント可能

国立衛生研究所 (NIH) HPCシステム

storage system	configuration	filesystem	network connectivity	usable storage (TB)
NetApp Cluster	10 x FAS8040 controllers SATA, SSD	NFS	16 x 10 Gb/s Ethernet	1100
DDN SFA12Ke	2 controllers 8 embedded fileserversSATA SSD metadata	GPFS	16 x 56 Gb/s FDR Infiniband	1400
3X DDN SFA12KXe	2 controllers 8 embedded fileserversNL-SAS SSD metadata	GPFS	16 x 56 Gb/s FDR Infiniband	6100
5X DDN SFA12KX-40	2 controllers 8 fileserversNL-SAS SSD metadata	GPFS	16 x 56 Gb/s FDR Infiniband	21600
DDN SFA18K	2 controllers 8 fileserversNVMe SSD NL-SAS SSD metadata	GPFS	8 x 56 Gb/s FDR Infiniband	4100
VAST	36 fileserversNVMe Flash	NFS	72 x 100 Gb/s HDR-100 Infiniband	4900



製品コンセプト

コンセプト

シングルティアの全く新しいアーキテクチャ
独自のデータ保護、データ削減技術、 Optane ™Technologyの最大限の活用から、
安価で高速なQLCフラッシュを活用したプライマリ領域からアーカイブ領域まで活用可能

オールフラッシュ パフォーマンス

Optane ™ Write性能
QLCフラッシュ Read性能

TIER-5 コスト効率

独自のデータ削減技術より
安価なフラッシュを活用した
HDDレベルのコスト効率

シンプル

シングルティアのNASアプライアンス
エクサバイトクラスまで容易にスケール可能
NFSoF・NFS・SMB・S3のサービス提供可能

全てのデータのためのOne Data Platform

あらゆるワークロードに対応

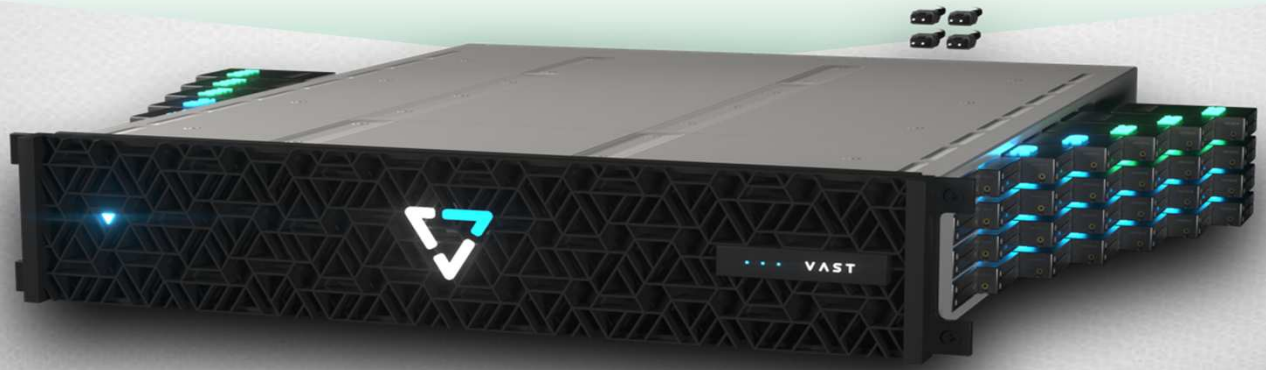
オールフラッシュNASストレージ　シングルティア＆シングルネームスペース　ペタバイト＆エクサバイトスケール

業種別ソリューション
- ライフサイエンス
- フィナンシャル
- コンテンツ
- HPC

セカンダリストレージ
- rubrik
- COMMVAULT
- veeAM

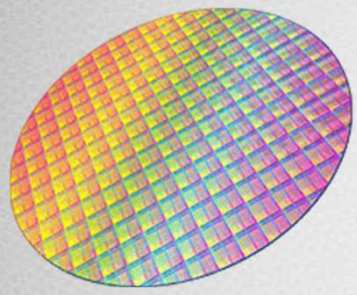
ビッグデータ&AI
- splunk>
- APACHE Spark
- TensorFlow

エンタープライズ
- vmware
- kubernetes
- OPENSHIFT



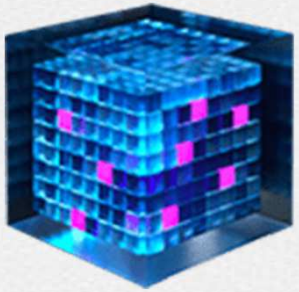
フラッシュ イノベーション

ストレージを根本的に再考



低コストQLCフラッシュ
85%節約

QLCフラッシュ専用に開発された FileSystem
従来ストレージシステムで利用する QLCフラッシュよりも、VASTクラスタのQLCフラッシュは20倍の耐久性実現



次世代イレージャーコーディング
最大66%節約

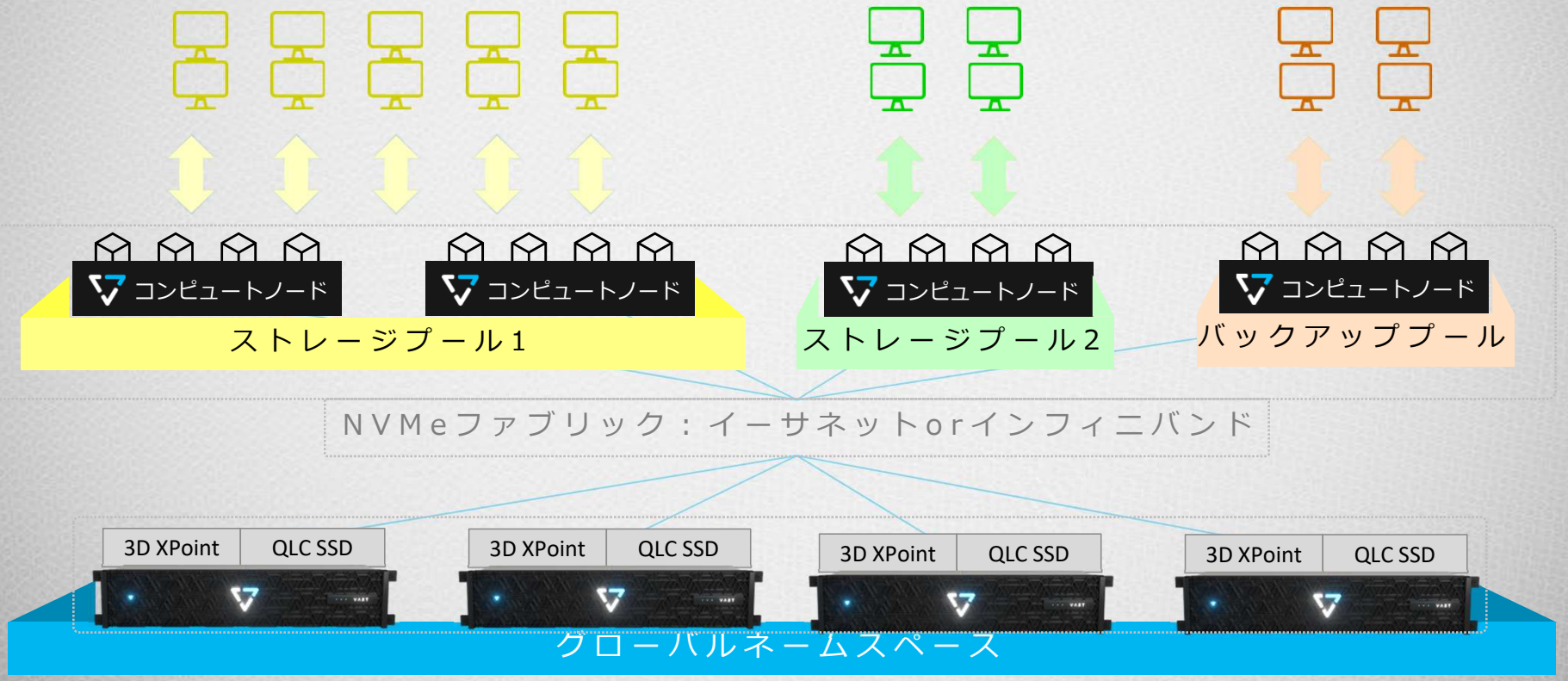
VASTローカルデコード可能なイレージャーコーディングは、
6000万年の復元を提供。
必要なオーバヘッドはわずか9%。

新世代データ削減
さらに50～90%節約

類似性ベースのデータ削減は、
グローバル重複排除と組み合わせて、
前例のないストレージ効率を実現。

DASE サーバプール

マルチプロトコルアクセス： NFS, NFS+RDMA 、NFS+RDMA+GPUDirect™, SMB, S3, K8S CSI



新しいデータ削除効果

独自の圧縮アルゴリズムにより、従来の重複排除よりも4,000～128,000倍の粒度によるデータ削減を実現。非構造データ、バックアップデータ、圧縮済みデータにも有効。

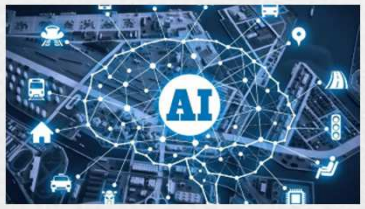
splunk>
elastic

ログデータ
4:1

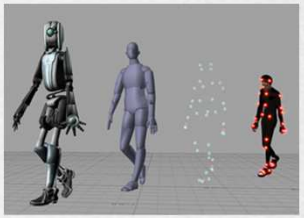
気象データ
2.5:1

ゲノムデータ
2:1

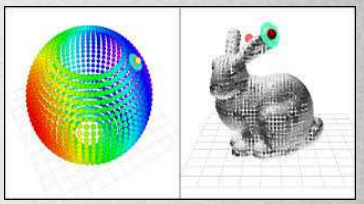
HPC
3:1



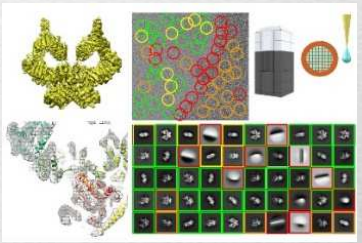
AI/機械学習データ
3:1



ＣＧＩ/アニメ
2:1



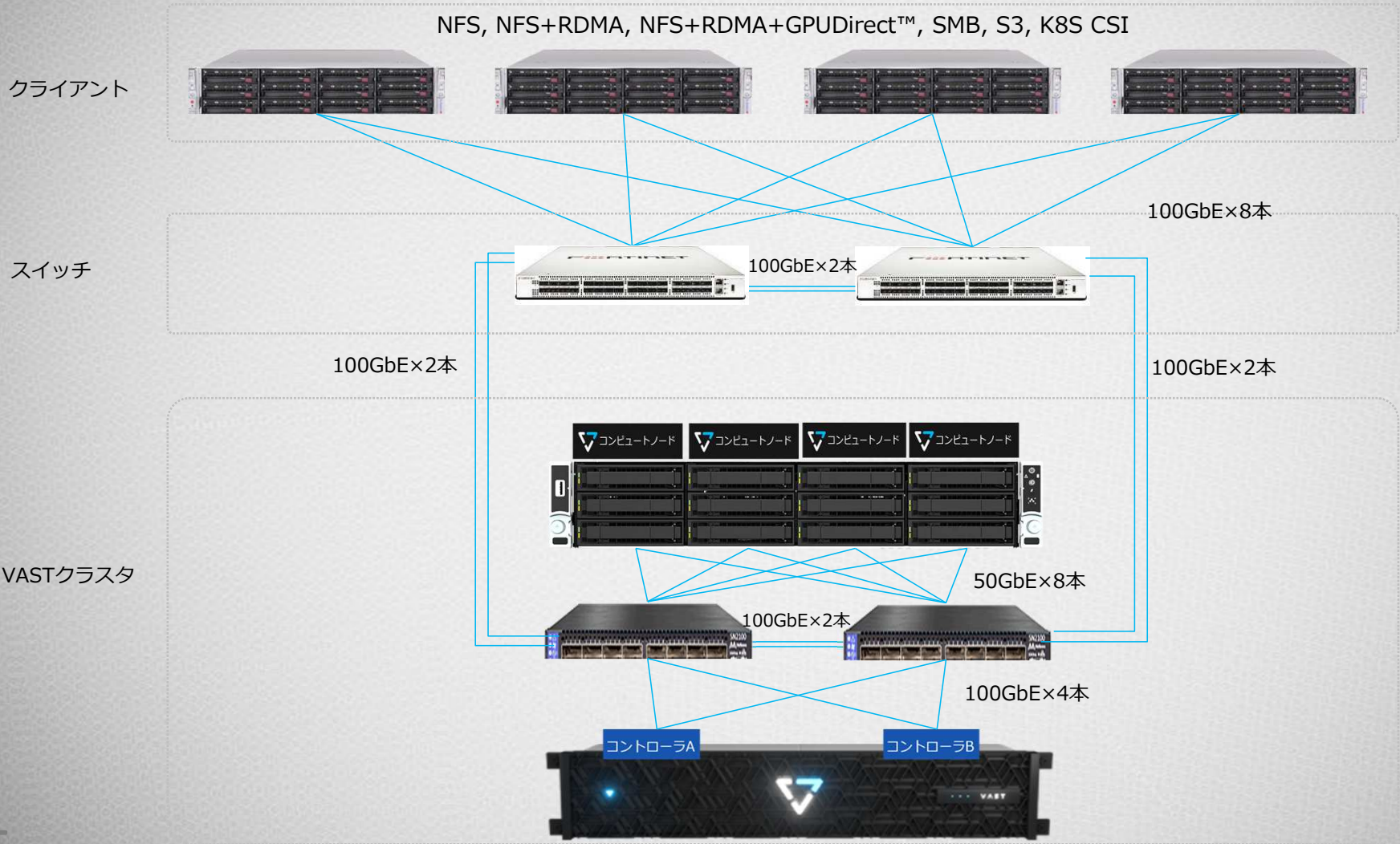
レンダリングデータ
～10:1



クライオデータ
4:1

機器構成

構成イメージ



テクノロジー
&機能

Patents

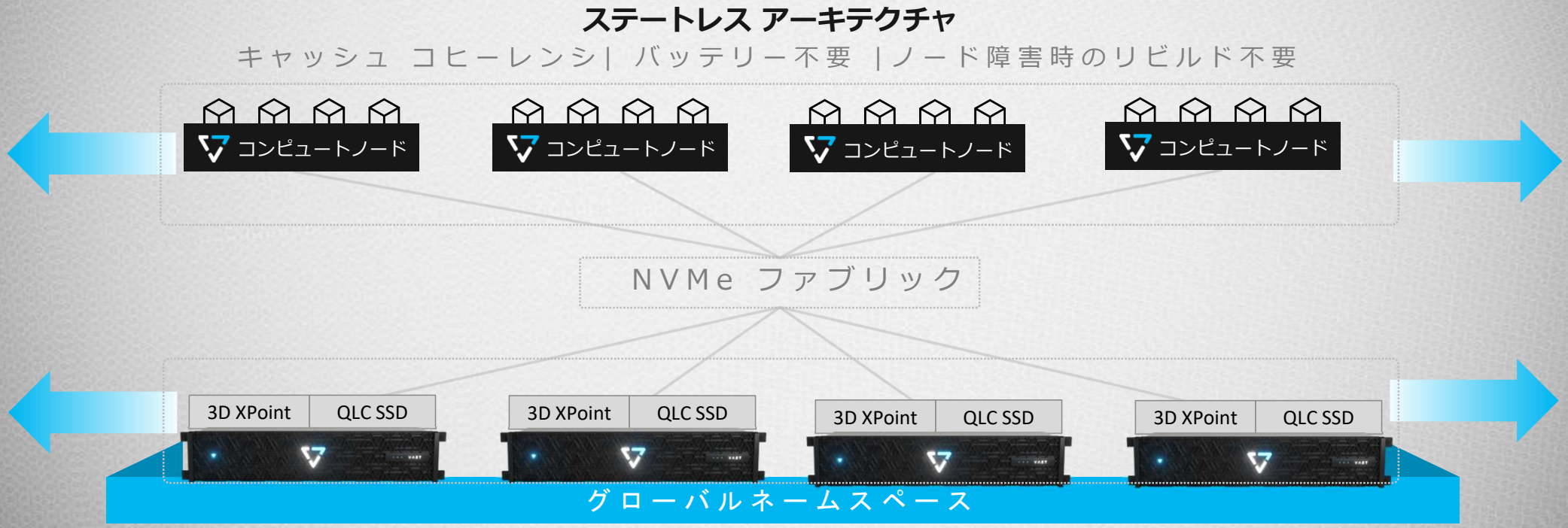
特許取得済みの独自テクノロジー

グローバルデータ圧縮やドライブ寿命を延ばす方法など各種特許を取得

- SYSTEM AND METHOD FOR GLOBAL DATA COMPRESSION（特許番号：20190379394）
- SYSTEM AND METHOD FOR USING FREE SPACE TO IMPROVE ERASURE CODE LOCALITY（特許番号： 20200348855）
- TECHNIQUES FOR PROLONGING LIFESPAN OF STORAGE DRIVES （特許番号： 20200174678）
- DISTRIBUTED SCALABLE STORAGE (特許番号：20190377490)
- STORAGE SYSTEM INDEXED USING PERSISTENT METADATA STRUCT
- METHOD AND SYSTEM FOR PROVIDING IMPROVED EFFICIENCY SNAF
- RESILIENCY SCHEMES FOR DISTRIBUTED STORAGE SYSTEMS（特許番
- DISTRIBUTED SCALABLE STORAGE （特許番号：20190377490）
- TRANSACTION MANAGER（特許番号：20200073964）

DASEアーキテクチャ

VASTのDASE (disaggregated shared-everything architecture)クラスタアーキテクチャは性能と容量の独立と、様々な機能を提供します。



DASEクラスタ メリット

VASTのDASEクラスタアーキテクチャが提供するグローバルアルゴリズムは、以下のメリットを得られます。

キャッシュコピーレーテンシーの課題無し

共有リソースに対するキャッシュの一貫性を保ちながら、スケーリングできます。

バッテリー不要

VASTクラスタは100%不揮発性。UPSやバッテリーの心配不要。

サーバ障害時のリビルド無し

ステートレスのため、コンピュートノード障害時のリビルドを排除します。

Dockerベースの自動スケーリング

コンテナ化されたアーキテクチャが、自律的で適応性のあるクラスタのスケーリングを行います。

グローバルネームスペース

VASTクラスタは、柔軟に拡張できるメタデータとデータを1つのグローバルネームスペースを実現します。

画期的なフラッシュ効率

VASTの新しいグローバルフラッシュトランスレート(FTL)により、10年以上のQLCの耐久性を実現します。

優れたシステム使用率

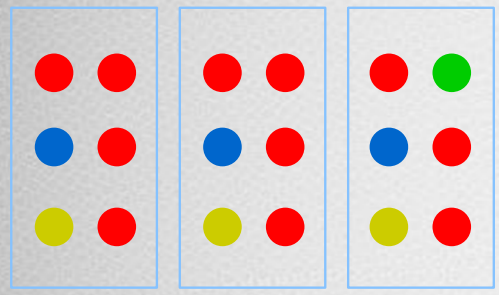
VASTのデータ保護に対するグローバルイレージャコーディングは、わずか3%のオーバヘッドの効率性を実現します。

革新的なデータ削減

新しいグローバルデータ削減アルゴリズムは、すでに圧縮されたデータに対しても優れたデータ削減効率を提供します。

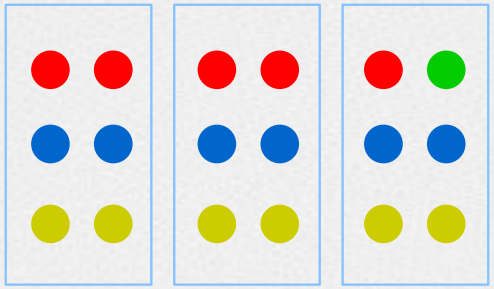
類似性データ削減テクノロジー

圧縮



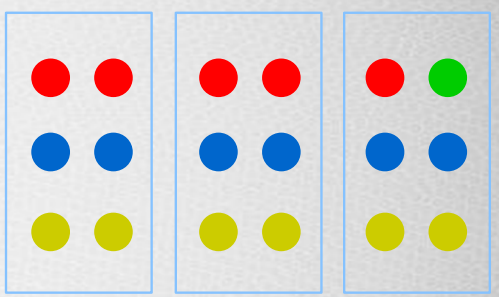
粒度＝細かい、ローカル

重複排除



グローバル、粒度＝粗い

VAST DATA 類似性



グローバル＆細かい粒度

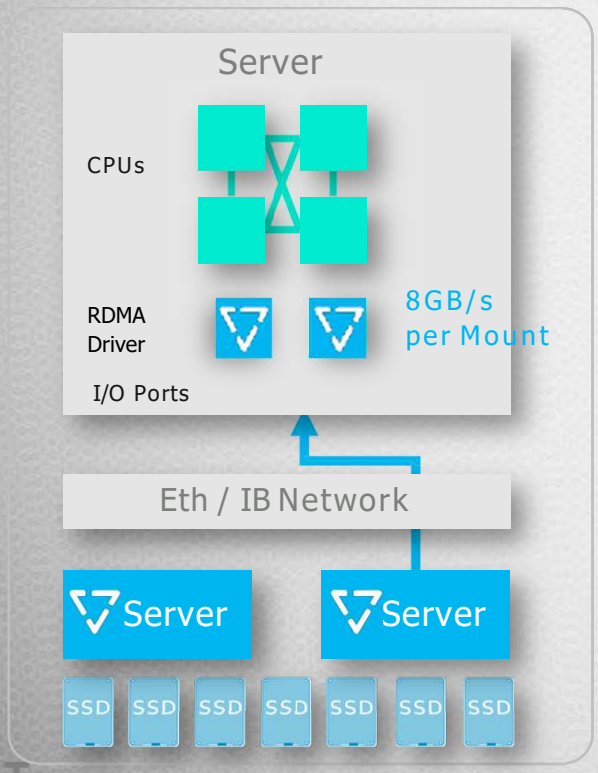
類似性効果

3:1 データ削減済みバックアップデータ	3:1 圧縮済みログファイル	2:1 ライフサイエンスデータ	3:1 HPCデータ	3:1 アニメーションデータ	8:1 時系列データ

RDMA

NFS over RDMA 提供

VASTはNFSプロトコルによるRDMAサーバ拡張機能を備えています。VASTが提供する「NFSoRDMA」ドライバをLinuxサーバにインストールすることで、NASのシンプルさとNFSプロトコルの欠点となる性能を向上できます。



	オールフラッシュ NASアプライアンス	フラッシュベース HPC ファイルシステム	VAST
シングルマウント GB/Sec	2GB/s (TCP制限)	10GB/s+ (RDMA-enabled)	8.7GB/s+ RDMA
接続方式	TCP Ethernet	Infiniband, RoCE Ethernet	Infiniband, RoCE Ethernet
シンプルさ	シンプル	複雑なインテグレーション	シンプル
スケール	ペタバイトスケール	ペタバイト～エクサバイト	エクサバイトスケール

データ破壊への対応

ランサムウェア対策

インディストラクティブル(破壊できない)・スナップショット機能により、個人が有効期限が切れる前にスナップショットが削除できない仕組みによる保護レイヤーを提供。

Indestructibleスナップショットがない場合

1. Attackers gain entry to your network, perhaps via Phishing email
   攻撃者がネットワークに侵入
2. With access, attacker gains elevated privileges
   攻撃者が特権を取得
3. Attacker deletes snapshots
   スナップショットを削除
4. Malware distributed by attacker encrypts your data
   マルウェアによりデータ暗号化
5. You pay $$$ ransom for decryption key, or restore from backup over weeks
   データ複合化の身代金を支払う

Indestructibleスナップショットがある場合

1. Attackers gain entry to your network, perhaps via Phishing email
2. With access, attacker gains elevated privileges
3. Attacker tries to delete snapshot, attempt fails
   スナップショットを削除しようとするが失敗
4. Accelerate recovery in any scenario with all-flash restores from indestructible snapshots
   スナップショットからリストア

バックアップ製品連携

バックアップデータの格納先として利用することで、従来よりも高いデータ削減率を実現できます。

	Rubrik Compression	+	Additional VAST Data Reduction	=	Total Combined Data Reduction Ratio
VMware バックアップ	4:1		1.3:1		5.2:1
非構造化データ	3.8:1		1.48:1		5.6:1
VMware SQL サーバ	5:1		1.68:1		8:1

	Commvault Compression: GZIP	Commvault Compression: LZO	Commvault Compression: Disabled
VMware バックアップ	3.4:1	6:1	8:1
非構造化データ	3.5:1	4.5:1	4.6:1
データベースサーバ	3.4:1	5.3:1	9:1

VAST　ソフトウェア・スタック

クラスタ
マネージメント

GUI
CLI
REST (SWAGGER)

コールホーム
VASTリモート監視、サポート

ＶＡＳＴマルチプロトコルアクセス
NFS v3、NFSv3 over RDMA、 SMB2、S3

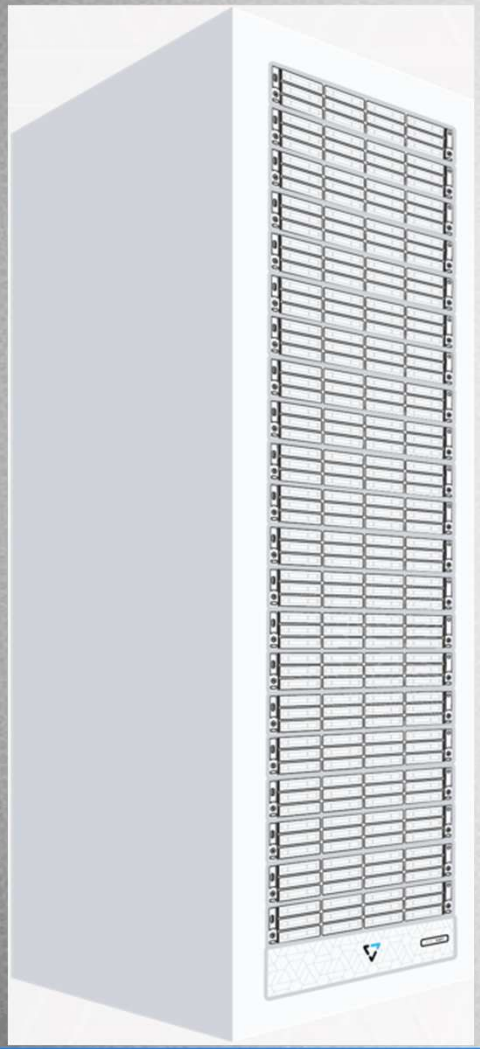
VAST データサービス

- エクサバイト スケールネームスペース
- 分散ファイルロッキング
- LDAP、AD、
- グローバルデータ削減
- POCIX ACLs
- スナップショット
- クォータ機能
- S3 出力
- 暗号化
(256-bit AES-XTS)

VAST インフラサービス

- トランザクションストレージシステム
- 分散永続メタデータ
- グローバル QLCフラッシュトランスレーション
- リビルドエラー修正
- 活性交換アーキテクチャ

VAST　クラウドバックアップ



クラウド

aws　Azure

オンプレ

MINIO　CLOUDIAN　VAST

製品インテグレーション

アイデンティティ　プロバイダー

Active Directory　LDAP.com Lightweight Directory Access Protocol　NIS

アラート、モニタリング＆通知

Grafana　Prometheus　webhooks　syslog

自動化＆仮想化

VAST Containers: docker　kubernetes

Containers: CONTAINER STORAGE INTERFACE　vmware　RED HAT OPENSHIFT

Other Virtualization: openstack. (Manila)　vmware (NFS)

バックアップアプリケーション

COMMVAULT　COHESITY　rubrik　veeam

ビッグデータ＆ディープラーニング

Analytics: Apache Spark　kx　sas

ML/DL: NVIDIA　TensorFlow　PyTorch

Search: splunk>　elastic　Lucene

GPUダイレクトストレージ

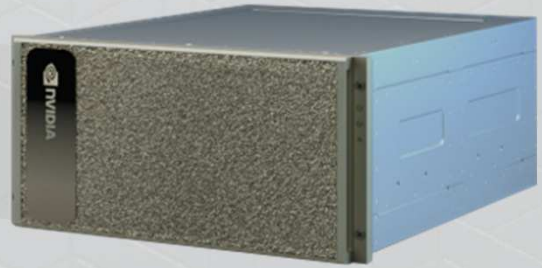
LIGHTSPEED

AI用ユニバーサルストレージ

SIMPLICITY, PERFORMANCE, AND SCALE TO POWER

これからの10年のためのストレージプラットフォーム

VAST + NVIDIA GPUダイレクトストレージ



DGX-A100

200GB/s

GPUDirect™

10,000,000 IOPS

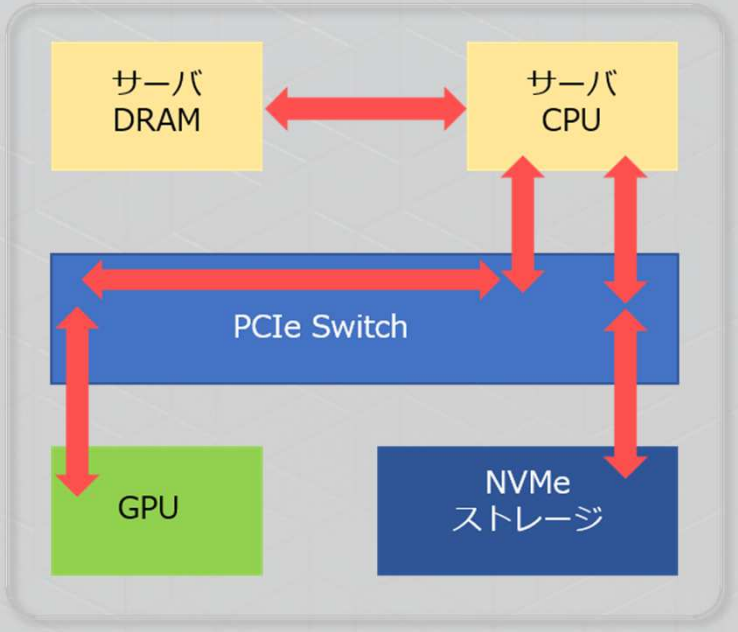
DGX-2

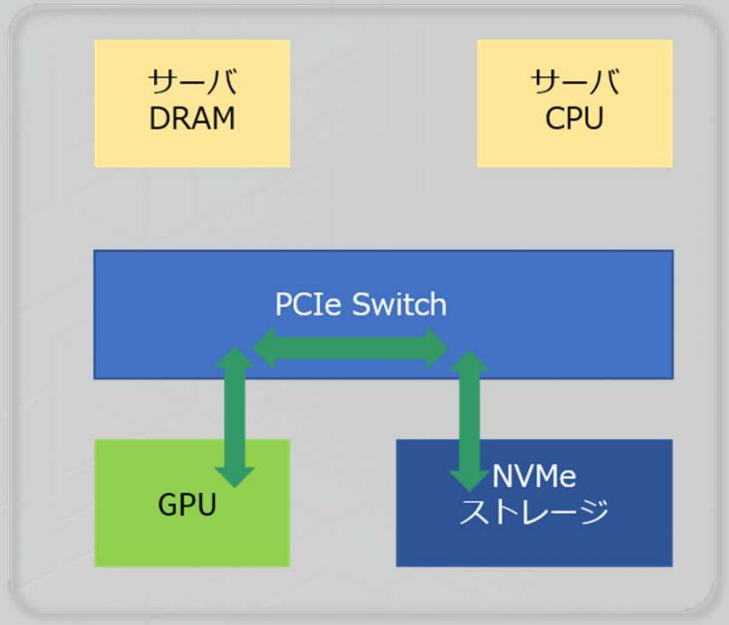
90GB/s

GPUDirect™

10,000,000 IOPS

GPU Direct Storageとは？





マスター タイトルの書式設定

VASTは、GDS認定ベンダー3社のうちの1社になります。

NVIDIA
ACCELERATED COMPUTING

MOFED and Filesystem Requirements

Here are the requirements:

- Ubuntu 18.04 and 20.04
- MOFED 5.1-0.6.6.0 and later, which supports NVMe NVMeoF, NFSoRDMA (VAST) on Linux kernel 4.15.x and 5.4.X
- The following distributed filesystems:
  - WekaFS 3.8.0
  - DDN Exascaler 5.2
  - VAST 3.4

GPUクラスタ　アーキテクチャ

LIGHTSPEED

最大16GPUクライアント

40 GB/S | 1 PB* | 400K IOPS

LightSpeed エンクロージャー ×1
VASTコンピュートノード ×1

PENTAGON

最大80GPUクライアント

200 GB/S | 5 PB* | 2M IOPS

LightSpeed エンクロージャー ×5
VASTコンピュートノード ×5

DECAGON

最大160GPUクライアント

400 GB/S | 10 PB* | 4M IOPS

LightSpeed エンクロージャー ×10
VASTコンピュートノード ×10

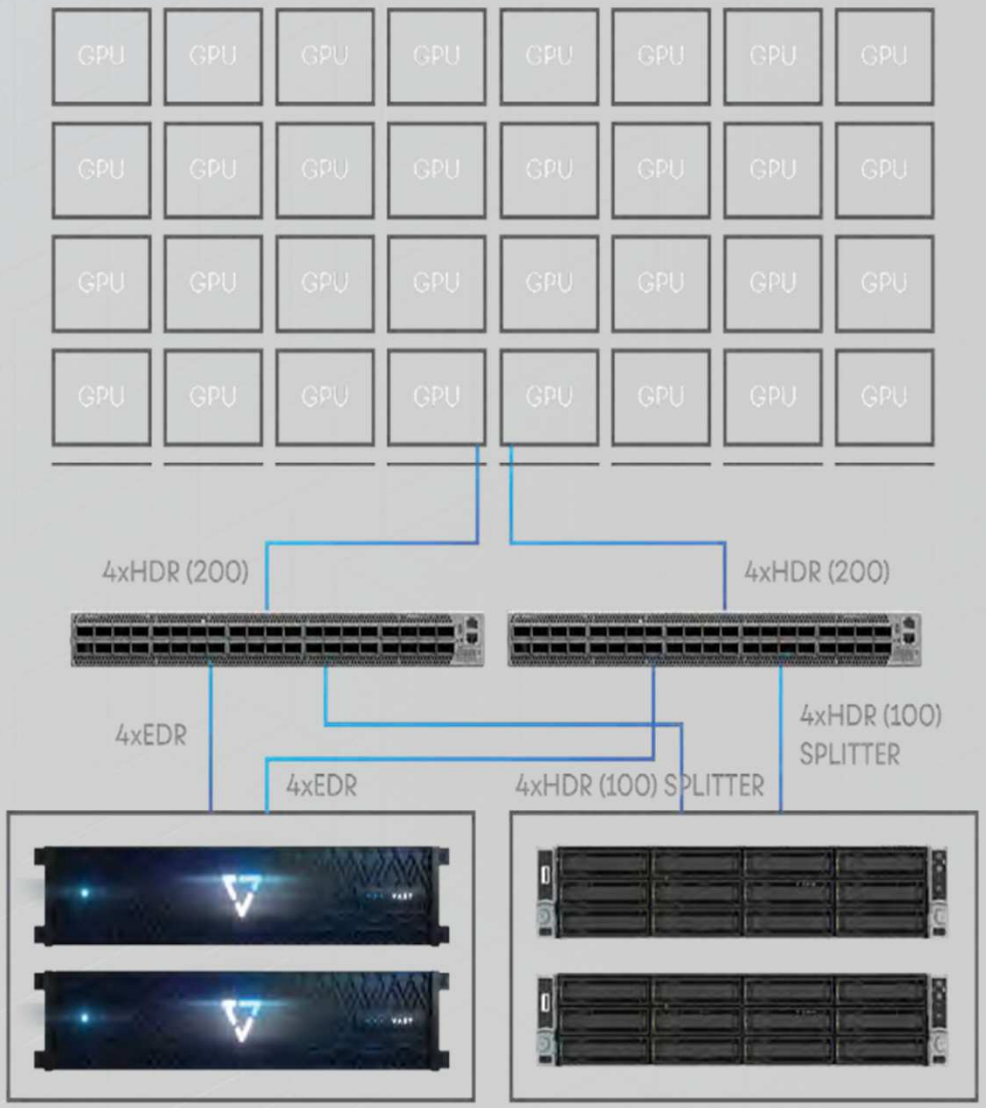
LIGHTSPEED

構成:

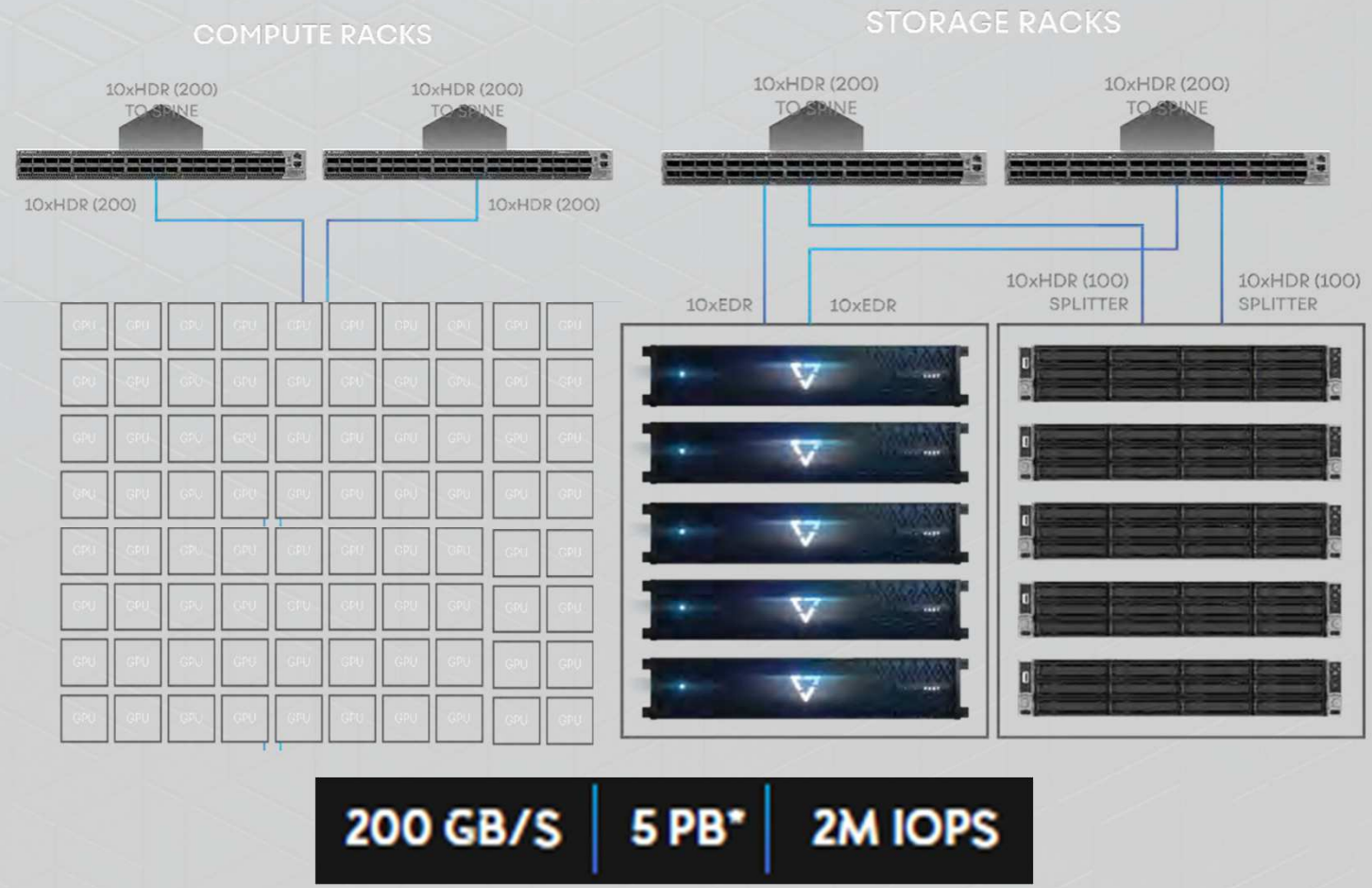
- 32 GPU クライアント
- VAST 2x2 (2 LightSpeed エンクロージャー)
  - 80GB/s Read
  - 10GB/s Write
- 26RU (~31RU with gaps)

ネットワーク:

- GPUs: 8xHDR (200)
- ISL: 4xHDR (200)
- エンクロージャー: 8xEDR
- コンピュートノード: 8xHDR (100) splitters



PENTAGON



200 GB/S | 5 PB* | 2M IOPS

構成:

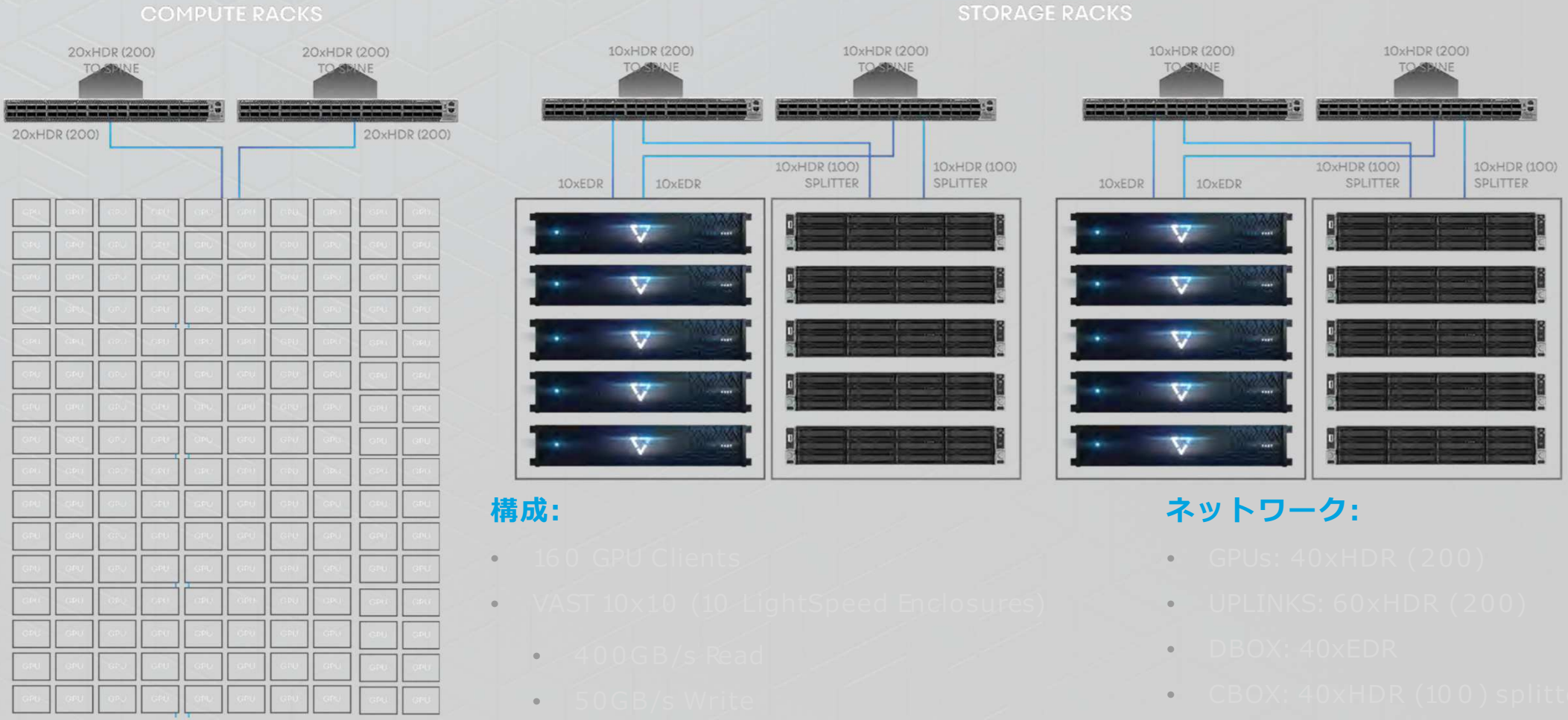
- 80 GPU クライアント
- VAST 5x5 (5 LightSpeed Enclosures)
  - 200GB/s Read
  - 25GB/s Write
- • 60RU (~31RU with gaps)

ネットワーク:

- GPUs: 20xHDR (200)
- UPLINKS: 40xHDR (200)
- エンクロージャー: 20xEDR
- コンピュートノード: 20xHDR (100) splitters

DECAGON

400 GB/S | 10 PB* | 4M IOPS



構成:

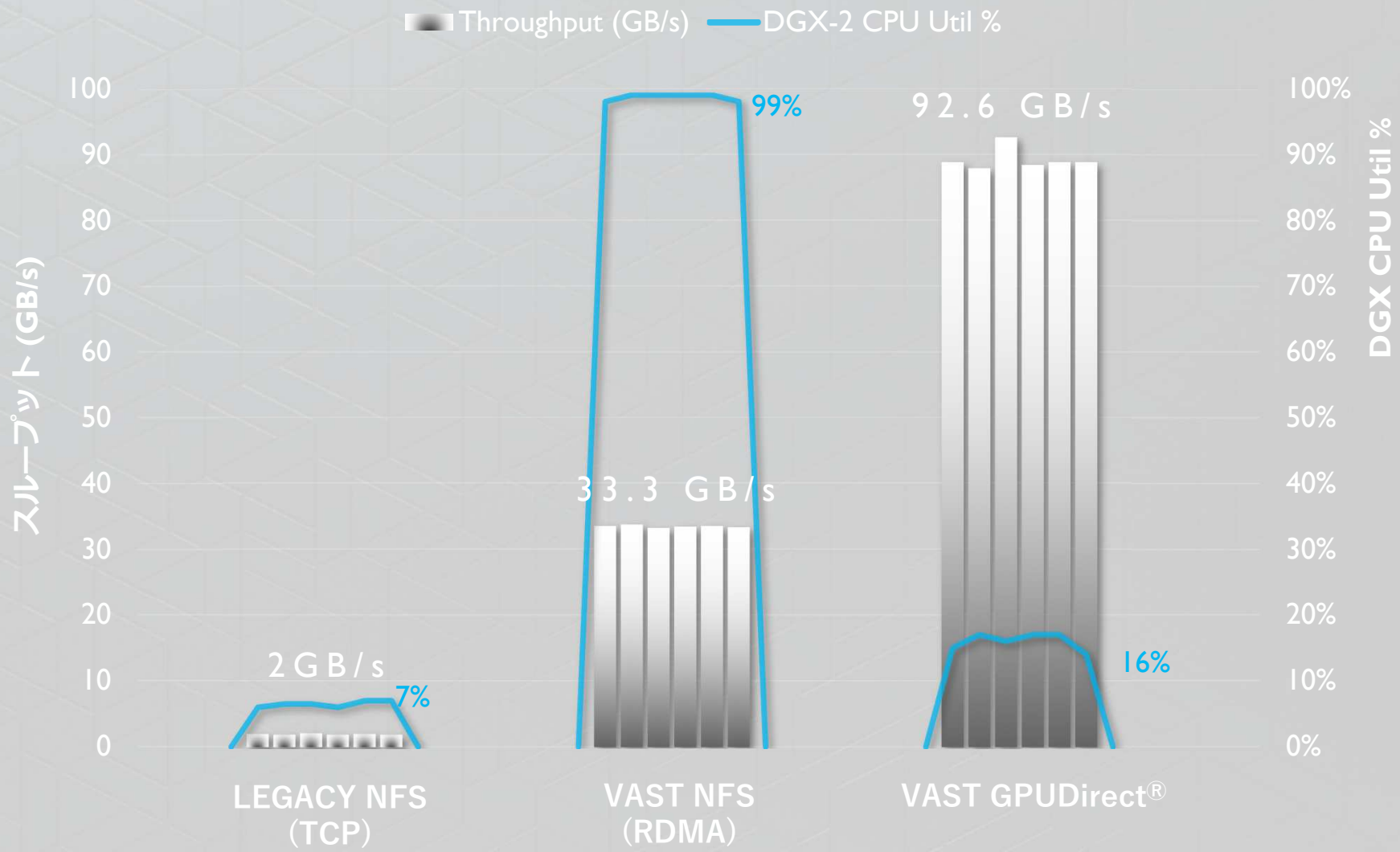
- 160 GPU Clients
- VAST 10x10 (10 LightSpeed Enclosures)
  - 400GB/s Read
  - 50GB/s Write

ネットワーク:

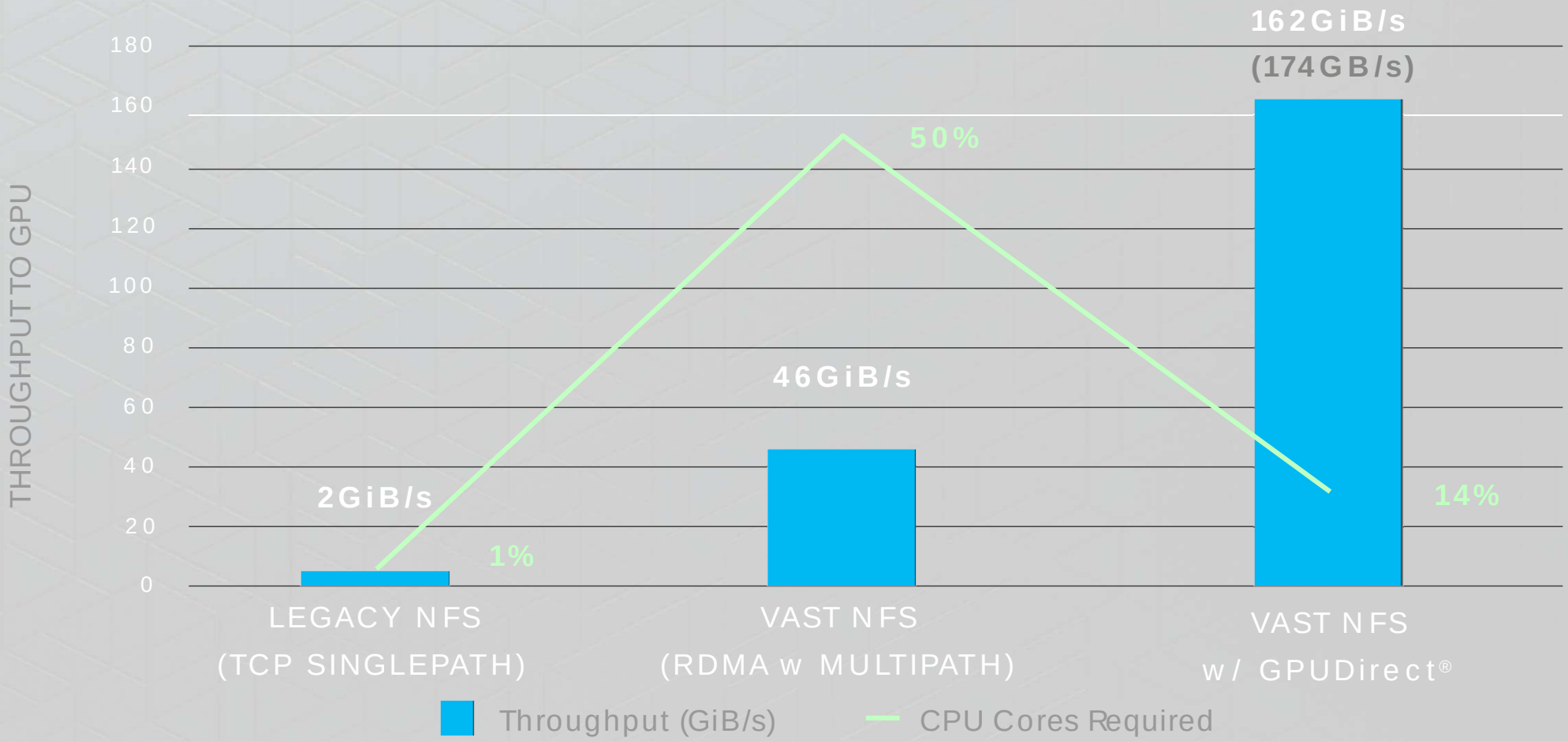
- GPUs: 40xHDR (200)
- UPLINKS: 60xHDR (200)
- DBOX: 40xEDR
- CBOX: 40xHDR (100) splitters

NVIDIA DGX-2 ベンチマーク

シングルマウントポイント



DGX-A100 ベンチマーク



TESTED WITH GDSIO • 1 x DGX-A100 • 5 x VAST CBOX SERVER CHASSIS & 5 x VAST LIGHTSPEED ENCLOSURES • 4MB I/O SIZE • 4GB FILE SIZE • 96 THREADS x 8GPUs

提供モデルと諸元

販売モデル（Gemini）

ハードウェア

・VAST クアッドサーバエンクロージャ
・NVMeファブリック　イーサネットスイッチ
・VAST NVMeエンクロージャ

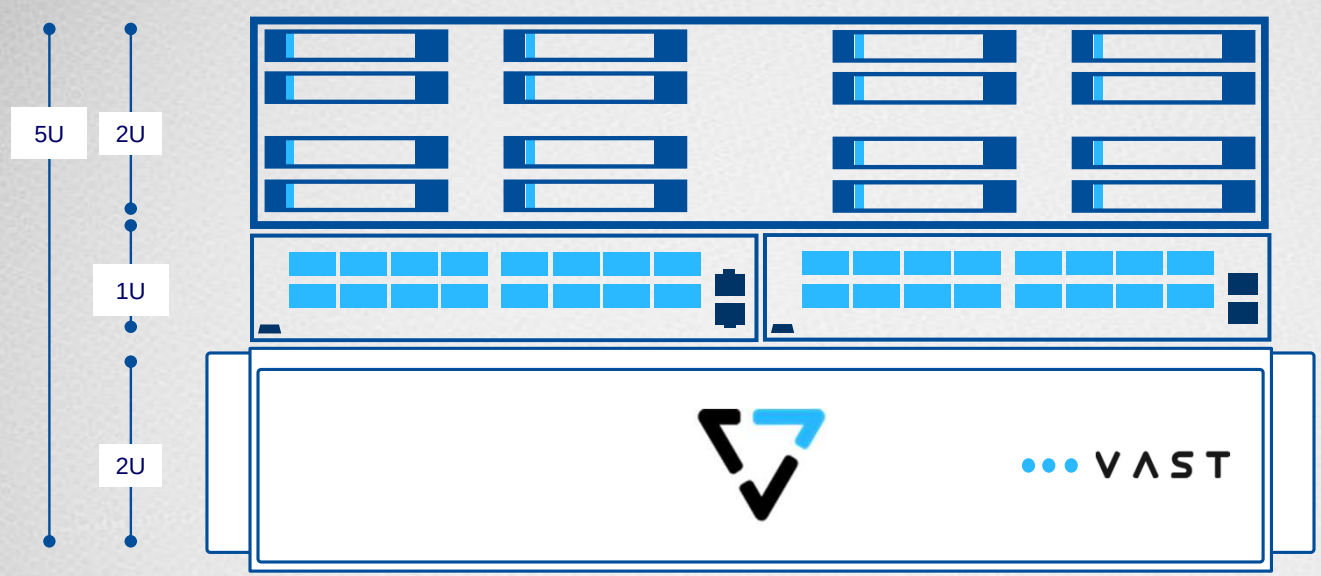
買い取り

ソフトウェア

・100TB単位
・1～10年契約

サブスクリプション

VAST クラスタ構成



コンピュートノード
(4ノード構成)

32ポートスイッチ
(利用可能ポート24ポート)

データノード

ストレージ容量（物理容量）	
NVMe Flash容量(TB)	675.8
3DXpoint(TB)	18.0

ストレージ容量（実効容量）	
NVMe Flash容量(TB)	552

ネットワークポート数	
データノード	4
コンピュートノード	8
スイッチポート(利用可能)	24

電源ケーブル数(C14/C15)	8
データノード	4
コンピュートノード	2
スイッチ	2

合計平均電力(kW)	2.392
データノード(kW)	1.4
コンピュートノード(kW)	0.8
スイッチ(kW)	0.2

ラックスペース	
ラックDepth(mm)	1193.8
ラックユニット数(最小)	5U

ピークストレージ性能(データノード)	
シーケンシャル Write	3.6GB/Sec
ランダム Write	5.1GB/Sec
ランダムRead	22GB/Sec
ランダム Write IOPS	102K
ランダム Read IOPS	182K

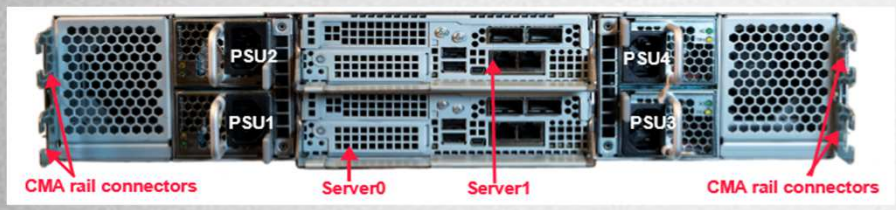
性能と制限事項

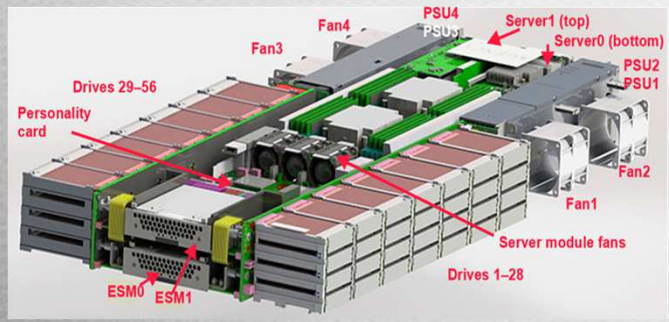
性能	4×コンピュートノード	8×コンピュートノード
シーケンシャル Write	5.1GB/Sec	5.3GB/Sec
ランダム Write	5.3GB/Sec	5.4GB/Sec
ランダムRead	19.2GB/Sec	22GB/Sec
ランダム Write IOPS	102,000	127,000
ランダム Read IOPS	182,000	322,000

制限事項（最大）	4×コンピュートノード
ファイル/フォルダ数(データノードあたり)	15億
ファイル/サブディレクトリ数(フォルダあたり)	100万
アクティブコネクション数(コンピュートノードあたり)	1,500
エクスポート数	200
ファイル/ディレクトリ文字長	255
ファイルサイズ	16TB
NFSファイルロック数(ファイルあたり)	100

エンクロージャー　諸元

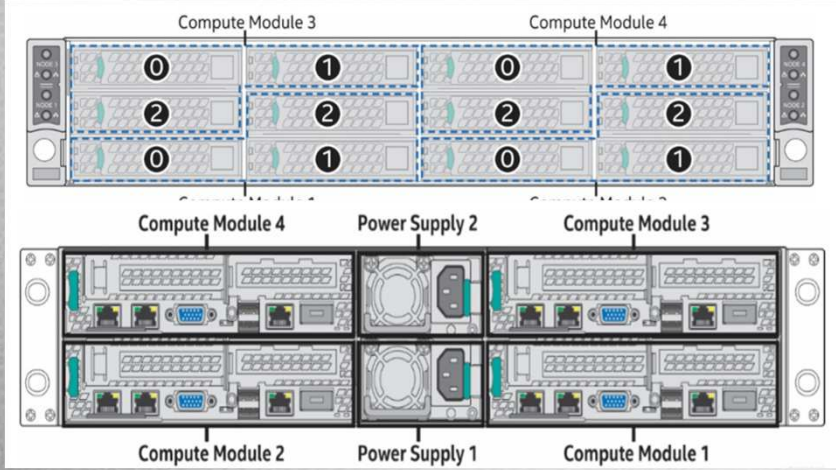
VAST DF-5615





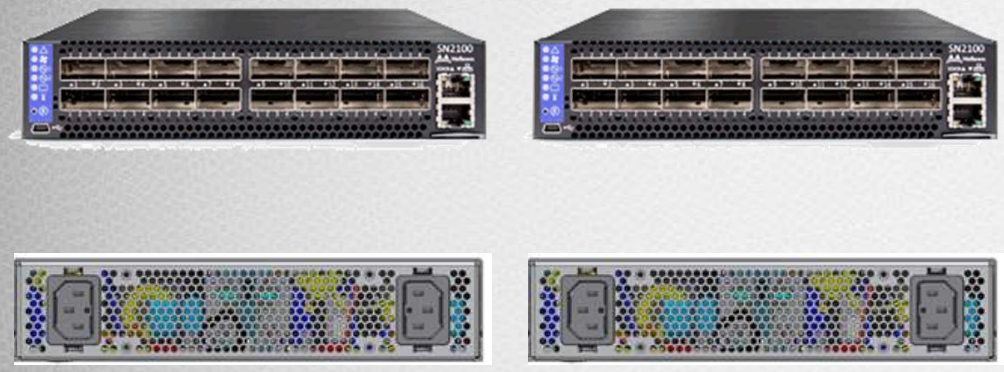
VAST DF-5615	
I/Oモジュール	Active/Active モジュール
ネットワーク	4×100GbEまたは4×100GbE InfiniBand
NVMe フラッシュ	38×15.36TB QLC
3Dxpoint	18×960GB
寸法	2U、H:81.28mm W:447.04 D:949.96
重量	38.55kg
電源	1500W×4
消費電力	1200W(平均)、1450W(最大)
最大拡張数	1,000(エンクロージャー)
冗長性	コントローラ 電源 NIC QLC SSD(イレージャーコーディング) 3DXPoint(ミラーリング) ※各部位活性交換可能

コンピュートノード　諸元



コンピュートノード	
サーバ	4×ステートレスサーバ(4ノード)
ネットワーク	8×50GbEまたは 4×100GbE InfiniBand 4×1GbE(オプション)
CPUコア	80×2.4Gヘルツ(ノードあたり20コア)
メモリ	32×32GB(ノードあたり256GB)
寸法	2U、H:86.87mm W:437.90　D:733.04
重量	35.38kg
電源	1600W×2
消費電力	750W(平均）、900W(最大)
最大拡張数	10,000(サーバ)
冗長性	ノード 電源 DISK(SSD) NIC ※各部位活性交換可能

NVMeスイッチ　諸元



MSN-2100

Mellanox スイッチ1U 16ポート

NVMeスイッチ	
型番	Mellanox MSN-2100
構成、ポート数	16ポートスイッチ×2台
ユニットサイズ	1U ※2台のスイッチを1Uに集約
速度	10/25/40/50/56/100GBR/ポート
寸法	1U、H:43.8mm W:200　D:508
重量	4.54kg
消費電力	94.3W(平均）、248.6W(最大)
冗長性	スイッチ 電源 ※各部位活性交換可能

ご清聴ありがとうございました

ノックス株式会社
営業本部
小幡　明広
obata@nox.co.jp
03-5731-5551