



# ディープラーニング用 ユニバーサルストレージ

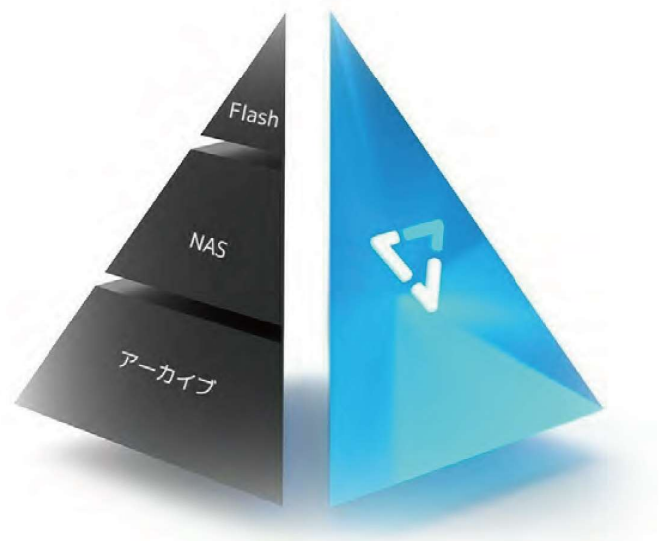
VASTは、パラレルファイルシステムのパフォーマンスや拡張性とオールフラッシュNASアプライアンスのシンプルさを兼ね備えています。HDDベースのストレージのコストメリットとフラッシュ・パフォーマンスを組み合わせることで、AI従事者はインフラを統合でき、ストレージの複雑な階層化を排除できます。

## ストレージ階層化の終焉

ストレージ管理者は、何十年もの間、コストを抑えるためにアーカイブ・ストレージインフラを階層化し、大容量データは低速なアーカイブ・ストレージに格納してきました。機械学習やディープラーニング技術の採用には、データを最大スループットでGPUに供給するために必要な大量の「トレーニングデータ」が必要になります。

また、学習アルゴリズムのトレーニングには、より多くのデータが用意されると効率的になります。これにより、従来型の階層化ストレージはAI時代に使用されることはなくなりました。

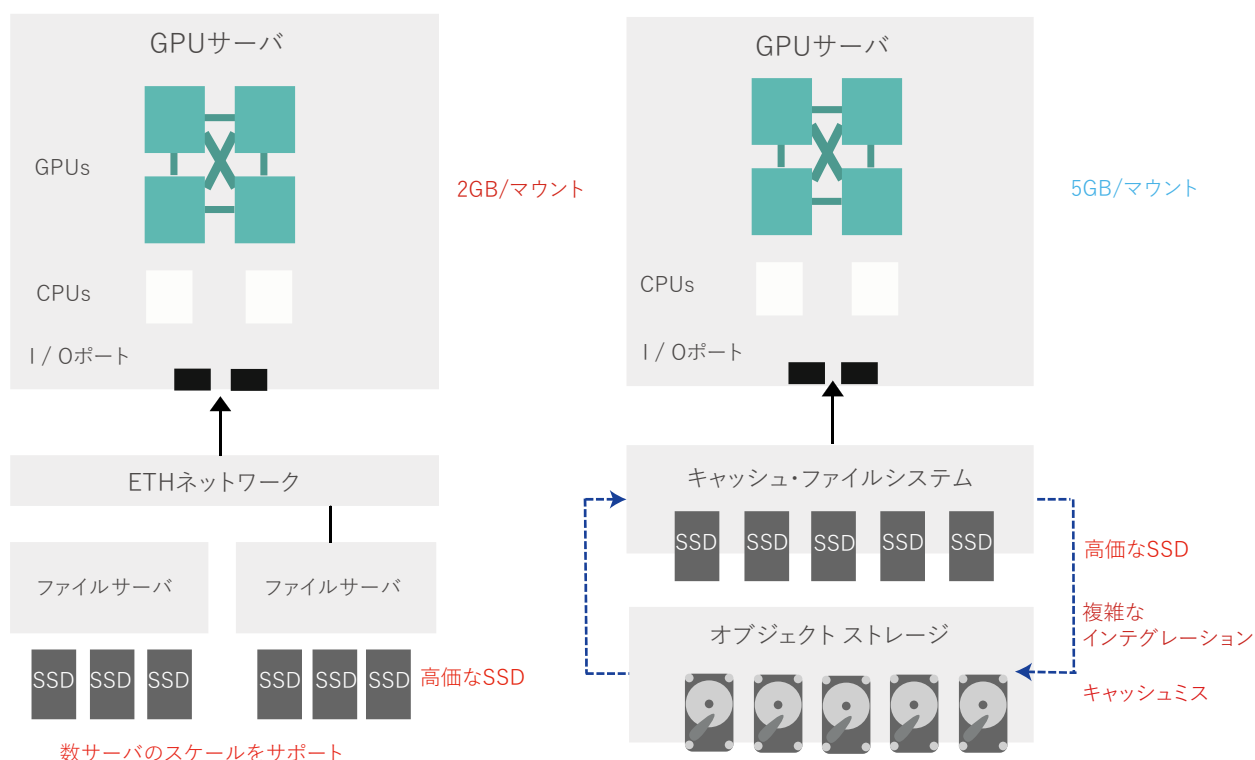
VASTのユニバーサルストレージは、プライマリストレージとして十分に高速なストレージであり、巨大なデータセットに対しても十分にスケーラブルで、データの全範囲に使用できる手頃な価格でシングルティア・ストレージシステムを提供し、階層化の縛りを排除します。





## スケーラブルなAIトレーニングへの3つのアプローチ

GPUサーバには、共有の必要がない小規模データセットのデータアクセスを提供できるローカルSSDが付属しています。ニュートラルネットワークのデータセットがGPUサーバに収まらない場合やGPUサーバのクラスターが共有データのグローバルアクセスを必要とする場合、ディープラーニングなどのAIパイプラインを運用するためにストレージのスケールアウトが必要になります。



### NASアプライアンス

NASのシンプルさは魅力的ですが、TCPベースのNFSの制限により、GPUサーバは効率的に共有ストレージからトレーニングセットを引き出すことができません。

単一マウントのパフォーマンスがわずかに2GB/sでは、従来型のNFSではGPUサーバ処理の多くはアイドル状態になります。

また、最新のスケールアウトNASアプライアンスは、低コストのQLCフラッシュを効率的に使用するには構築されておらず、HDDベースのストレージよりもはるかに高いコストが発生することがよくあります。

### バーストバッファ + オブジェクトストレージ

スケーラブルなディープラーニング・ストレージアーキテクチャーを構築する別のアプローチは、高速ファイルシステム(バーストバッファ)と低コストのアーカイブ・ストレージを組み合わせることです。

このアプローチは、システムが適切なタイミングでデータをキャッシュに入ることが保証できない場合、I/Oが完全に確定的とは限らないAIワークロードには、かなりの問題を引き起こす可能性があります。異なるストレージシステムを統合することも複雑であり、キャッシュ・ファイルシステムの性質により、低コストのSSDを使用できません。

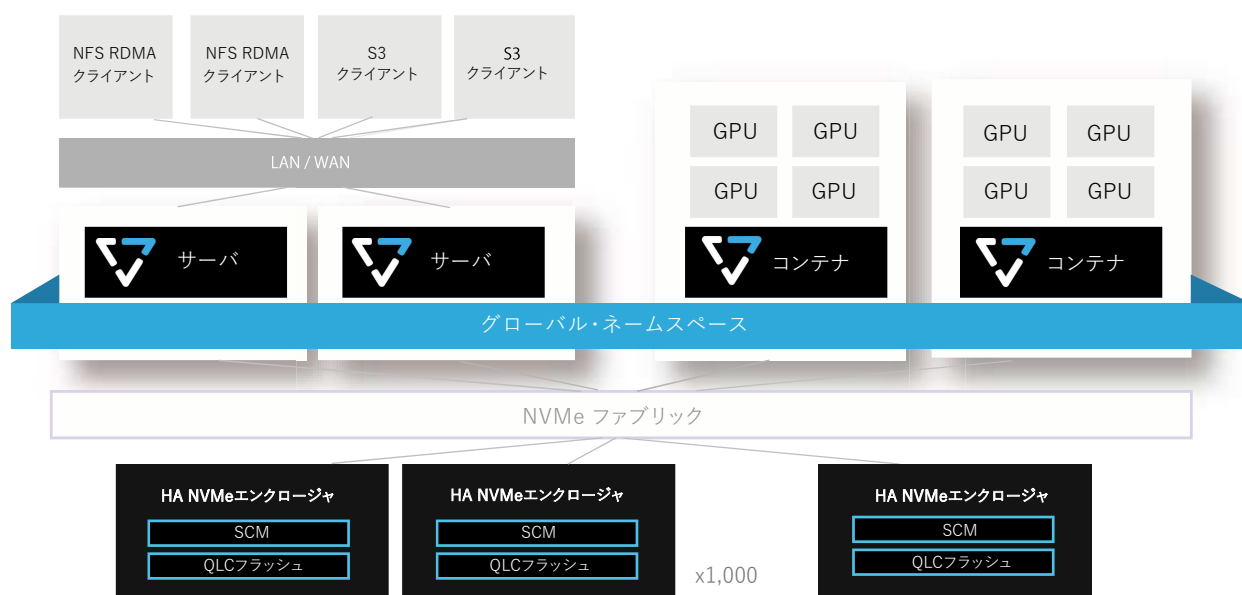
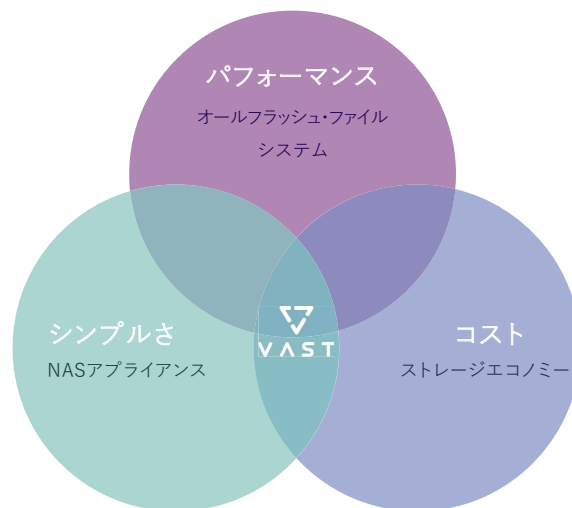


## VAST Data ユニバーサルストレージ

VASTは、最新のディープラーニングトレーニングと推論パイプラインの要求を解決するために設計された、まったく新しいストレージアーキテクチャを構築しました。数十年前からのストレージの問題に新しい考え方を適用することで、VASTはストレージパフォーマンスと容量コストに関する長年のトレードオフを解消しました。

VAST Dataのユニバーサルストレージを使用すると、最も要求の厳しいアプリケーションに十分なパフォーマンス提供とエクサバイトクラスの拡張性に対応でき、非アクティブなデータであってもHDDと競合するのに十分な価格のシングルティアのオールフラッシュ・ストレージに全てのデータを保存できます。

全てのデータが常にオンラインで使用できるようになると、膨大なデータセットからパターンや相関関係を発見するデータマイニングが簡単になります。

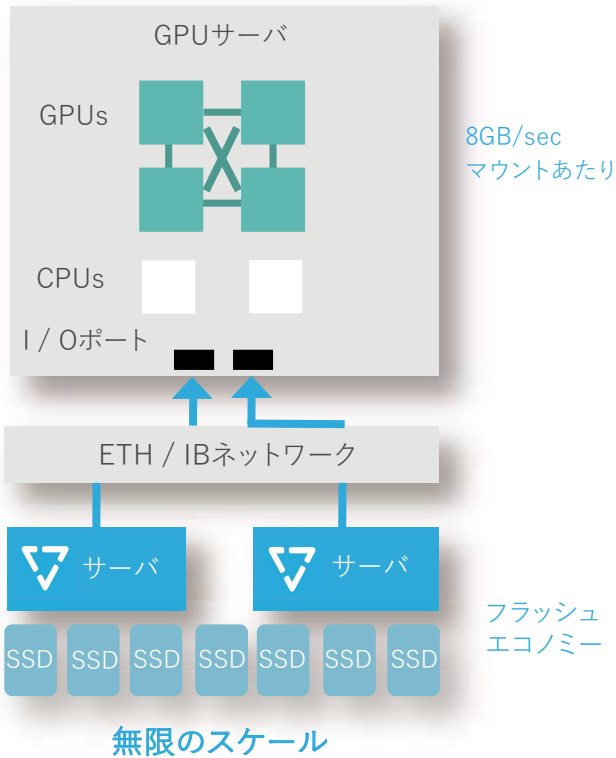


VASTの新しい分離型シェアード・エブリシング(Disaggregated, Shared Everything=DASE)ストレージアーキテクチャは、スケーラブルなストレージをシェアード・ナッシングクラスタから構築する必要があるという考えから脱却します。サーバをストレージから分離できる場合、全てが改善されます。

- シングルネームスペースはエクサバイトクラスのデータをサポートし、データを1,000台のGPUサーバに格納できます。
- GPUコンピューティングに関しては、VASTサーバソフトウェアをGPUサーバに組み込むことができ、NVMe over Fabricsの平行接続のパフォーマンスをAIアプリケーションにもたらしめます。
- 最良の部分-全てのトレーニングおよび推論データを手頃な価格でフラッシュに保存できるようになりました。ユーザーはデータのステージング、ストレージ階層間のバースト、それ以外の場合ではフラッシュのコストを増加させるフラッシュのウェアレベリングに気を遣う必要がありません。



# GPU効率を最大化するユニバーサルストレージ



## アクセラレーション NFS RDMA接続

VASTは、NFS over RDMA を使用して GPU クライアントを接続する機能を提供し、シンプルなNAS を実現します。100GbE接続を使用すると、単一NFSマウントポイントで20GB/sec近くのパフォーマンスを実現することができます。

ある大手高級自動車メーカーが、VAST DataのNFS over RDMAソリューションのテストを実施し、従来のスケールアウトNASと比較しました。100Gb InfiniBandネットワーク上の単一のNFS接続だけでGPUあたり2GB/sec以上を提供できましたが、マシンの異なるCPUに異なるマウントポイントをマッピングした時、クライアントあたり18.7GB/secと、これまでに見たこともないNFSの結果を確認しました。

	GPUサーバ毎の単一マウントポイント	GPUサーバ毎の複数マウントポイント
NFS over TCP	2.1GB/sec	8.7GB/sec
NFS over RDMA	8GB/sec	20.5GB/sec

パフォーマンス、拡張性、シンプルさ、コスト、全てにおいてVASTが優れています。

