



NetApp Verified Architecture

NetApp ONTAP AI (NVIDIA搭載)

現実世界のディープ ラーニング ユースケースのため
に設計された、拡張性に優れたAIインフラ

ネットアップ David Arnette、Sundar Ranganathan、Amit Borulkar、Sung-Han Lin、Santosh Rao 2018年8月 | NVA-1121

提携協力 :



目次

1 エグゼクティブ サマリー	1
2 プログラムの概要.....	1
2.1 NetApp Verified Architectureプログラム	1
2.2 NetApp ONTAP AIソリューション	1
3 ディープ ラーニングのデータ パイプライン	2
4 ソリューションの概要	4
4.1 ソリューションのテクノロジ	4
4.2 NVIDIA DGX-1サーバ	5
4.3 NetApp AFFシステム	6
4.4 NetApp ONTAP 9	6
4.5 NetApp FlexGroupボリューム	7
4.6 NVIDIA GPU CloudとTrident	8
4.7 Cisco Nexus 3232Cネットワーク スイッチ	8
4.8 RDMA over Converged Ethernet.....	9
5 テクノロジの要件.....	9
5.1 ハードウェア要件	9
5.2 ソフトウェア要件	10
6 ソリューション アーキテクチャ.....	10
6.1 ネットワーク トポロジとスイッチの構成	11
6.2 ストレージ システムの構成	12
6.3 ホストの構成	14
7 ソリューションの検証	15
7.1 検証テストのプラン	15
7.2 検証テストの結果	16
7.3 ソリューションのサイジングに関するガイダンス	20
8 まとめ	21
謝辞	21
詳細情報の入手方法	21
付録	iv
各モデルのさまざまなバッチ サイズでのトレーニング速度	iv
各モデルでのGPUスケーリングの比較	iv
TensorコアとCUDAコアとの比較	v
全モデルでのGPUワークロード	vi

表一覧

表 1) ハードウェア要件	10
表 2) ソフトウェア要件	10

図一覧

図 1) NetApp ONTAP AI ソリューションのラックスケール アーキテクチャ	2
図 2) エッジからコア、クラウドに至るデータ パイプライン	3
図 3) NetApp ONTAP AI ソリューションの検証済みアーキテクチャ	5
図 4) NetApp FlexGroup ボリューム.....	8
図 5) Converged Enhanced Ethernet 標準と RoCE v1 / v2 をサポートする NX-OS を搭載した Cisco Nexus スイッチ	8
図 6) ネットワーク スイッチ ポート構成	11
図 7) DGX-1 ポートとストレージ システム ポートの VLAN 接続	12
図 8) ストレージ システムの構成	13
図 9) DGX-1 ホストのネットワーク ポートと VLAN の構成	14
図 10) 全モデルのトレーニング スループット	17
図 11) GPU 利用率とストレージ帯域幅 (VGG16)	17
図 12) 全モデルでの推論 (Tensor コアと CUDA コア)	18
図 13) 全モデルのストレージ帯域幅	19
図 14) 全モデルのストレージ レイテンシ	19
図 15) 全モデルのストレージ CPU 利用率	20
図 16) トレーニング モデルのさまざまなバッチ サイズの比較	iv
図 17) さまざまなトレーニング モデルでの GPU スケーリング	v
図 18) CUDA コアと Tensor コアのパフォーマンス比較	v
図 19) ResNet-50 の GPU 利用率とストレージ帯域幅	vi
図 20) ResNet-152 の GPU 利用率とストレージ帯域幅	vi
図 21) Inception-v3 の GPU 利用率とストレージ帯域幅	vii

1 エグゼクティブ サマリー

本ドキュメントには、テクニカル ホワイトペーパー『[WP-7267：拡張性に優れたAIインフラ](#)』で解説しているアーキテクチャの検証情報が記載されています。このホワイトペーパーに記載された設計の実装は、[NetApp® AFF A800 All Flash FASシステム](#)、[NVIDIA® DGX-1™サーバ](#)、[Cisco® Nexus® 3232C](#) 100Gbイーサネット スイッチを使用して行われました。このシステムの動作とパフォーマンスの検証には、業界標準のベンチマーク ツールを使用しました。検証テストの結果、このアーキテクチャはトレーニングと推論において優れたパフォーマンスを提供することが実証されました。またテスト結果では、複数のDGX-1サーバをサポートするための適切なストレージ ヘッド ルームがあることも示されました。コンピューティング リソースとストレージ リソースの構成をハーフラックからマルチラックに拡張することも簡単であり、リソースごとに個別に行うことができ、パフォーマンスも予測できるので、機械学習ワークフローのいかなる要件も満たしています。

2 プログラムの概要

2.1 NetApp Verified Architecture プログラム

NetApp Verified Architectureプログラムは、ネットアップ ソリューションの検証済みアーキテクチャを提供するものです。NetApp Verified Architectureにより、ネットアップ ソリューション アーキテクチャでは以下が保証されています。

- 入念にテスト済みである
- あらかじめ規定されている
- 導入のリスクが最小限である
- 運用開始までの時間が短縮される

本ドキュメントは、ネットアップおよびパートナー様のソリューション エンジニアと、お客様向け 戦略の意思決定者を対象としています。本ドキュメントでは、特定の環境で必要とされる具体的な機器、ケーブル接続、構成を決定するための判断材料となるアーキテクチャ設計の考慮事項について説明します。

2.2 NetApp ONTAP AIソリューション

NetApp ONTAP® AI統合インフラは、NVIDIA DGX-1サーバとネットアップのクラウド対応ストレージ システムで構成される、ネットアップとNVIDIAが共同で開発、検証したアーキテクチャです。あらかじめ規定されたアーキテクチャであり、企業にとって次のようなメリットがあります。

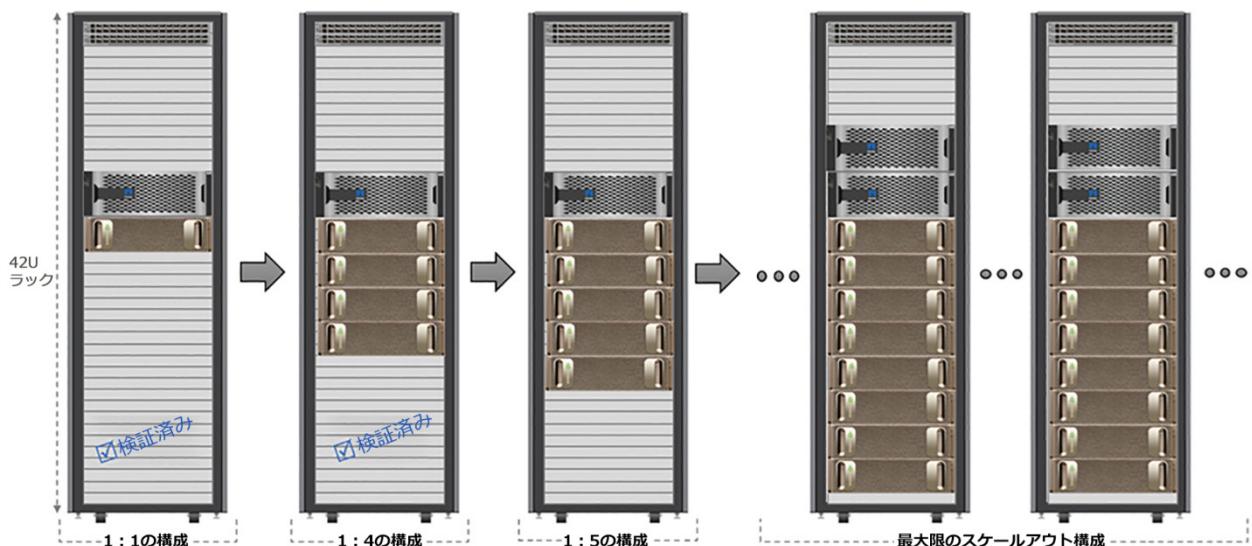
- 設計の複雑さが解消される
- コンピューティングとストレージの個別拡張が可能
- 小規模構成から始めて、シームレスに拡張可能
- パフォーマンスとコストのさまざまな観点から、幅広いストレージ オプションが提供される

NetApp ONTAP AIでは、NVIDIA DGX-1サーバ、NVIDIA Tesla® V100グラフィック処理ユニット (GPU) 、NetApp AFF A800システムが最先端のネットワークを介して統合されます。設計の複雑さと推測による試行錯誤を排除することで、人工知能 (AI) をシンプルに導入できます。小規模構成から始めて、システム停止なしで拡張することができ、エッジ、コア、クラウドの間を移動するデータのインテリジェントな管理を実現します。

図1は、NetApp ONTAP AIソリューションの拡張性を示しています。AFF A800システムの検証には、4台のDGX-1サーバを使用しました。その結果、ストレージ スループットやレイテンシに影響

を及ぼすことなく、5台以上のDGX-1サーバをサポートするのに十分なパフォーマンス ヘッドルームを備えていることが実証されました。さらに、ネットワーク スイッチとストレージ コントローラのペアをONTAPクラスタに追加することにより、ソリューションを複数のラックに拡張してきわめて高いスループットを実現し、トレーニングと推論を高速化することもできます。このアプローチには、データレイクのサイズ、使用するディープ ラーニング (DL) モデル、求められるパフォーマンス指標に基づいて、コンピューティングとストレージの比率を個別に変えることができる柔軟性があります。

図1) NetApp ONTAP AIソリューションのラックスケール アーキテクチャ



DGX-1サーバとAFFシステムのラックあたりの数は、使用中のラックの電力と冷却の仕様によって異なります。システムの最終的な配置は、数値流体力学的な分析、エアフロー管理、データセンター設計によって決まります。

3 ディープ ラーニングのデータ パイプライン

DLは、さまざまな分野で応用されているエンジンです。たとえば、不正の検出、顧客関係の強化、サプライ チェーンの最適化、競争が激化する市場での革新的な製品やサービスの提供が、DLによって可能になります。DLモデルのパフォーマンスと精度は大幅に向かっています。これは、ニューラル ネットワークの規模と複雑さが拡大し、モデルのトレーニングに使用されるデータの量と質が向上しているからです。

処理するデータ セットが膨大であるため、複数の環境にわたって導入できる柔軟性を備えたインフラを設計することが不可欠です。エンドツーエンドのDL環境は、大まかに見て、データが移動する3つのステージで構成されます。つまり、エッジ (データの取り込み)、コア (トレーニング クラスターとデータレイク)、クラウド (アーカイブ、階層化、開発とテスト) です。こうした構成は、モノのインターネット (IoT) など、データ パイプラインの3つの領域全体にわたってデータが移動するアプリケーションではきわめて一般的です。

図2) エッジからコア、クラウドに至るデータ パイプライン



図2は、3つの各領域を構成する要素の概要を示しています。

- **データの取り込み :** データの取り込みは一般的に、自動運転車や POS (販売時管理) デバイスからストリーミングされるデータをキャプチャするときのように、エッジで実行されます。ユースケースに応じて、取り込みポイントまたはその周辺に IT インフラが必要になる場合もあります。たとえば小売企業では、複数のデバイスから集めたデータを統合する小型のインフラが各店舗で必要に可能性があります。
- **データの前処理 :** トレーニング前のデータを正規化およびクレンジングするには、前処理が必要です。前処理はデータ レイクで行われます。Amazon S3 階層形式のクラウドや、オンプレミスのストレージ システム (ファイル ストアやオブジェクト ストアなど) が考えられます。
- **トレーニング :** DL の重要なトレーニング フェーズでは、データはデータ レイクからトレーニング クラスタに一定の間隔でコピーされるのが一般的です。このフェーズで使用されるサーバは、GPU で計算処理を並列化することで大量のデータに備えます。GPU 利用率を高く維持するためには、物理 I/O 帯域幅のニーズを満たすことが重要です。
- **推論 :** トレーニング済みのモデルは、テストされたあと、本番環境に導入されます。あるいは、データ レイクに戻されて、入力データの重み調整がさらに行われる場合もあります。IoT アプリケーションの場合は、モデルがスマート エッジ デバイスに導入される場合もあります。
- **アーカイブ、階層化 :** 繰り返し処理を終えたコールド データは無期限に保存することもできます。企業の AI チームの多くは、プライベート クラウドまたはパブリック クラウドのオブジェクト ストレージにコールド データをアーカイブすることを好みます。

DLモデルは、アプリケーションに応じて、異なる種類のデータを大量に扱います（構造化データと非構造化データの両方を含む）。このような違いがあるため、基盤となるストレージ システムには、格納されるデータのサイズとデータセット内のファイル数の観点から、異なる要件が課されることになります。

ストレージ要件には、大きく分けて次のものがあります。

- 数百万ものファイルの格納と読み出しを同時に扱うことができる
- 多様なデータ オブジェクト (イメージ、オーディオ、ビデオ、時系列データなど) の格納と読み出しができる
- GPU の処理速度に合わせて、低レイテンシで高い並列処理パフォーマンスを実現できる
- エッジ、コア、クラウドにまたがるデータ管理とデータ サービスをシームレスに行うことができる

AFFシステムは、NetApp ONTAPの卓越したクラウド統合機能とソフトウェアで定義される機能を組み合わせることで、DL用のエッジ、コア、クラウドにわたるデータパイプラインに全面対応できます。本ドキュメントでは、データパイプラインのトレーニング部分と推論部分のためのソリューションに焦点を絞って説明します。

4 ソリューションの概要

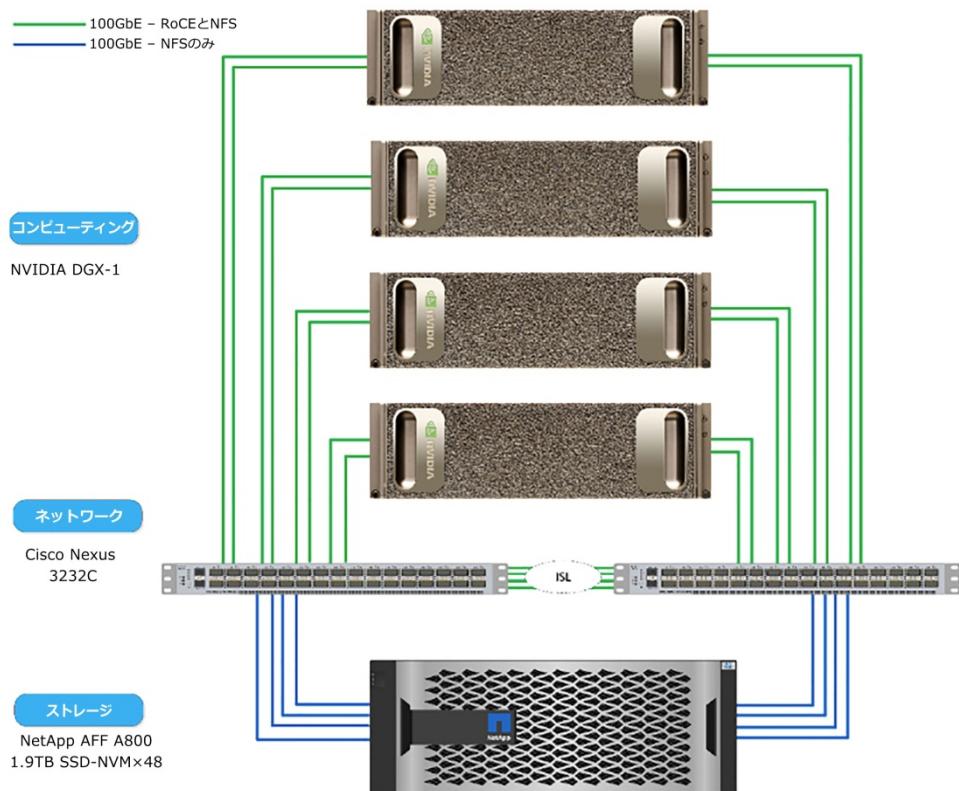
DLシステムが活用するアルゴリズムは計算処理による負荷が高いので、NVIDIA GPUのアーキテクチャが特に適しています。DLアルゴリズムで実行される計算処理では、大量の行列乗算を並列実行することが求められます。最新のGPUは、高度に並列化されたアーキテクチャを備えているため、データが並列で処理されるアプリケーション（DLなど）では、汎用中央処理装置（CPU）に比べて、データ処理の効率性が格段に優れています。NVIDIA GPUは、DGX-1サーバを活用するアーキテクチャが単体とクラスタの両方で進歩を遂げたことで、ハイパフォーマンスコンピューティング（HPC）、DL、分析などのワークロードに対して望ましいプラットフォームとなりました。こうした環境で最大のパフォーマンスを引き出すには、NVIDIA GPUにデータを供給し続けることのできる基盤インフラが必要です。したがって、データセットへのアクセスは広帯域幅を使用して超低レイテンシで行われる必要があります。

4.1 ソリューションのテクノロジ

このソリューションは、1台のNetApp AFF A800システム、4台のNVIDIA DGX-1サーバ、2台のCisco Nexus 3232C 100Gbイーサネットスイッチを使用して実装されました。各DGX-1サーバからNexusスイッチへの接続は、RoCE（Remote Direct Memory Access [RDMA] over Converged Ethernet）を使用して、GPU間通信のための4つの100GbE接続を介して行われます。NFSストレージにアクセスするための従来のIP通信もこれらのリンクで行うことができます。各ストレージコントローラは、4つの100GbEリンクを使用してネットワークスイッチに接続されます。

従来型のHPCインフラでは、広帯域幅と低レイテンシの機能を備えたRDMA over InfiniBand（IB）をノード間接続に使用します。以前はIBでのみ実現可能であったパフォーマンスレベルにまでイーサネットテクノロジが進歩するにつれて、RoCEによってこうした機能の導入を容易に行えるようになりました。これは、イーサネットテクノロジが一般的によく理解されており、あらゆるエンタープライズのデータセンターで広く導入されているからです。図3は、基本的なソリューションアーキテクチャを示しています。

図3) NetApp ONTAP AIソリューションの検証済みアーキテクチャ



4.2 NVIDIA DGX-1サーバ

DGX-1サーバは、DLワークフローのために設計された、ハードウェアとソフトウェアが完全統合されたターンキー システムです。DGX-1サーバは8基のTesla V100 GPUを備え、NVIDIA NVLink™ テクノロジを使用したハイブリッド キューブメッシュ トポロジで構成されます。この構成により、GPU間通信のための超広帯域幅、低レイテンシ ファブリックが実現します。このトポロジは、マルチ GPUトレーニングに不可欠であり、GPU数の増加に応じてパフォーマンスをリニアに向上させることができないPCIeベースのインターフェースに関連するボトルネックを解消します。DGX-1サーバには、RDMA対応ファブリックでのマルチノード クラスタリングのための広帯域幅、低レイテンシ ネットワーク インターフェースも装備されています。

DGX-1はNVIDIA GPU Cloud (NGC) を搭載しています。これは、GPUの高速性を利用するソフトウェアのためのクラウドベースのコンテナ レジストリです。NGCには、Caffe2、TensorFlow、PyTorch、MXNet、TensorRTなど、今日最も人気の高いDLフレームワーク用のコンテナが用意されており、これらのフレームワークはNVIDIA GPU用に最適化されています。こうしたコンテナによって、フレームワークまたはアプリケーション、必要なドライバ、ライブラリ、通信のプリミティブが統合されます。これらの要素はスタック全体でNVIDIAによって最適化され、GPUの高速性に基づくパフォーマンスが最大限に引き出されます。NGCコンテナにはNVIDIA CUDA Toolkitが組み込まれているため、NVIDIA CUDA Basic Linear Algebra Subroutines Library (cuBLAS) やNVIDIA CUDA Deep Neural Network Library (cuDNN) などのツールを利用できます。また、マルチGPUおよびマルチノード集団通信プリミティブ用のNVIDIA Collective Communications Library (NCCL) も含まれており、DLトレーニングのためのトポロジ対応が可能になります。NCCLにより、1台のDGX-1サーバ内および複数台のDGX-1サーバ間でのGPU間通信が実現します。

4.3 NetApp AFFシステム

NetApp AFFは最先端のストレージ システムです。このシステムを使用することにより、業界をリードするパフォーマンス、卓越した柔軟性、クラウド統合、業界最高のデータ管理機能でエンタープライズ ストレージの要件を満たすことができます。フラッシュに特化して設計されたAFFシステムは、ビジネス クリティカルなデータの高速処理、管理、保護に役立ちます。

NetApp AFF A800システムは、業界初のエンドツーエンドNVMeソリューションです。NASワークロードの場合、1台のAFF A800システムが、シーケンシャル読み取りで25GB/秒のスループット、スマート ランダム リードで100万IOPSのパフォーマンスを500マイクロ秒未満のレイテンシで実現します。AFF A800システムは次の機能をサポートしています。

- 24 ノード クラスタで最大 300GB/秒の優れたスループットと 1,140 万 IOPS
- 100Gb イーサネット接続および 32Gb FC 接続
- マルチストリーム書き込み (MSW) に対応した 30TB ソリッド ステート ドライブ (SSD)
- 2U ドライブ シェルフで 2PB の高密度ストレージ
- 364TB (2 ノード) から 74PB (24 ノード) まで拡張可能
- NetApp ONTAP 9.4 と、業界をリードするデータ管理のための包括的なデータ プロテクション 機能およびレプリケーション機能

パフォーマンスの観点からAFF A800ストレージ システムに続くのはAFF A700sシステムであり、NASワークロードの場合で18GB/秒のスループット、40GbEの転送をサポートします。AFF A300およびAFF A220システムは、低コストながらも十分なパフォーマンスを提供します。

4.4 NetApp ONTAP 9

ONTAP 9は、ネットアップが提供する最新世代のストレージ管理ソフトウェアです。これを使⽤することにより、インフラを最新化し、クラウドに対応したデータセンターに移行できるようになります。業界をリードするデータ管理機能を活用するONTAPを使えば、保管場所がどこであれ、1つのツールセットでデータを管理、保護できます。データは、エッジ、コア、クラウドなど、必要とされる場所に自由に移動することもできます。ONTAP 9には、データ管理を簡易化し、重要なデータの高速処理と保護を実現する豊富な機能が搭載されており、ハイブリッド クラウド アーキテクチャ全体でインフラが将来のニーズに対応できるようになります。

データ管理の簡易化

データ管理は、エンタープライズIT運用でアプリケーションとデータセットに適切なリソースを使用するために重要です。ONTAPには、運用を合理化、簡易化し、TCOを低減するための、次のような機能が含まれます。

- **インライン データ コンパクション、強化された重複排除：**データ コンパクションはストレージ ブロック内の無駄な領域を減らし、重複排除は実効容量を大幅に増やします。
- **最小、最大、アダプティブのサービス品質 (QoS) :**きめ細かい QoS 管理は、高度に共有された環境において重要なアプリケーションのパフォーマンス レベルを維持するのに役立ちます。
- **ONTAP FabricPool :**この機能を使用すると、Amazon Web Services (AWS)、Azure、NetApp StorageGRID®ソリューションを含むパブリック クラウドとプライベート クラウドのストレージ オプションに、コールド データの自動階層化が加わります。

データの高速処理と保護

ONTAPは、卓越したパフォーマンス レベルとデータ保護を実現し、さらに以下の機能によってこれを強化します。

- **パフォーマンスの向上とレイテンシの低下** : ONTAP は、できるかぎり最小のレイテンシで最大のスループットを提供します。
- **NetApp ONTAP FlexGroup** : ハイパフォーマンスなデータ コンテナである FlexGroup ボリュームは、最大 20PB、4,000 億ファイルまでリニアに拡張でき、データ管理を簡易化する単一のネームスペースを提供します。
- **データ保護** : ONTAP に組み込まれたデータ保護機能により、すべてのプラットフォームを一元管理できます。
- **NetApp Volume Encryption** : ONTAP では、オンボードと外部のキー管理のサポートによって、ネイティブのボリュームレベルの暗号化が可能です。

将来のニーズにも対応できるインフラ

ONTAP 9は、要件が厳しく、絶えず変化するビジネス ニーズを満たすための支援を提供します。

- **シームレスな拡張が可能でシステム停止のない運用** : ONTAP は、既存のコントローラとスケールアウト クラスタに対し、システム停止なしで容量を追加できます。NVMe や 32Gb FC など最新のテクノロジへのアップグレードも、コストのかかるデータ移行やシステム停止を行わずに実行できます。
- **クラウドへの対応** : ONTAP はほとんどのクラウドに対応するストレージ管理ソフトウェアであり、すべてのパブリック クラウドで Software-Defined Storage (ONTAP Select) とクラウドネイティブ インスタンス (NetApp Cloud Volumes サービス) のオプションを選択できます。
- **新しいアプリケーションとの統合** : ONTAP は、既存のエンタープライズ アプリケーションをサポートする同じインフラを使用して、OpenStack、Hadoop、MongoDB などの次世代のプラットフォームやアプリケーションのためのエンタープライズクラスのデータ サービスを提供します。

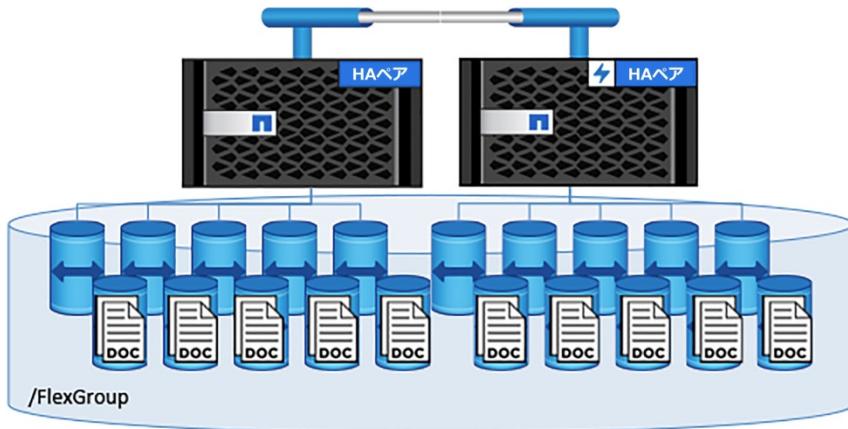
4.5 NetApp FlexGroupボリューム

一般に、トレーニング データセットは大量のファイルの集合体です（ファイル数は数十億個になる場合もあります）。ファイルにはテキスト、オーディオ、ビデオなどの形式の非構造化データが含まれる場合があり、そうしたデータを格納し、並列で読み取れるように処理する必要があります。ストレージ システムは、大量の（数十億個に及ぶ）小さなファイルを格納できるだけではなく、そうしたファイルをシーケンシャルI/OおよびランダムI/Oによって並列で読み取れなければなりません。

FlexGroupボリューム（図4）は、複数のコンステイチュエント メンバー ボリュームで構成された単一のネームスペースで、ストレージ管理者はNetApp FlexVol®と同様に使用および管理することができます。FlexGroupボリューム内のファイルは、個々のメンバー ボリュームに割り当てられ、複数のボリュームやノードにまたがってストライピングされることはありません。このボリュームでは次の機能が実現します。

- FlexGroup ボリュームは、大容量（数ペタバイト）と、メタデータの多いワークロードにおける予測可能な低レイテンシを可能にします。
- 数千億個ものファイルを 1 つのネームスペースでサポートします。
- CPU、ノード、アグリゲート、コンステイチュエント FlexVol 全体で、NAS ワークロードの並列処理をサポートします。

図4) NetApp FlexGroupボリューム



4.6 NVIDIA GPU CloudとTrident

NVIDIA GPU Cloud (NGC) は、完全に統合され、高パフォーマンスを提供するよう設計された、NVIDIA GPUを最大限に活用するDL向けDockerイメージのカタログを提供します。こうしたイメージには、NVIDIA CUDA ToolkitやNVIDIA DLライブラリなどの必要なすべての依存機能が含まれます。NVIDIA DGX-1サーバで使用できるように、NVIDIAがイメージのテスト、調整、認定を行っています。さらに、GPUを活用するイメージのモビリティを実現するため、NVIDIAはNVIDIA Container Runtime for Dockerを開発しました。これを使用すると、起動時にNVIDIAドライバとGPUのユーザ モード コンポーネントをDockerコンテナ内にマウントできるようになります。

ネットアップのTridentは、DockerとKubernetes向けの、オープンソースのダイナミックストレージプロビジョニングツールです。NGCおよびKubernetesやDocker Swarmなどの人気の高いオーケストレーションツールとTridentを組み合わせることで、DL NGCコンテナ イメージをNetAppストレージにシームレスに導入でき、エンタープライズクラスのAIコンテナ導入を実現できます。この導入には、自動オーケストレーション、テスト開発用のクローン作成、アップグレード テストでのクローニング、保護、コンプライアンスのコピーの使用、およびNGC AIとDLコンテナ イメージを使用するその他多くのデータ管理のユースケースが含まれます。

4.7 Cisco Nexus 3232Cネットワークスイッチ

Cisco Nexus 3232Cスイッチ（図5）は、データセンター向けに開発された、低レイテンシ、高密度、ハイパフォーマンス、高電力効率の100Gb/秒スイッチです。コンパクトな1ラック ユニット（1RU）モデルで、全ポートにおいてワイヤスピードのレイヤ2 / レイヤ3スイッチ機能を450ナノ秒のレイテンシで提供します。このスイッチは、Cisco Nexus 3200プラットフォームに属する製品で、業界をリードするCisco NX-OSソフトウェア オペレーティング システムと、幅広く導入されている包括的な機能を搭載しています。Cisco Nexus 3232Cは、32個のQSFP28ポートを備えたQuad Small Form-Factor Pluggable (QSFP) スイッチです。各QSFP28ポートは、10、25、40、50、100Gb/秒で運用でき、25Gb/秒のポートを最大128個まで拡張できます。

図5) Converged Enhanced Ethernet標準とRoCE v1 / v2をサポートするNX-OSを搭載したCisco Nexusスイッチ



このソリューションでは、テストの際に各ネットワーク スイッチで使用可能なポートの半分しか消費されません。各スイッチは、より多くのGPU能力を提供するために追加のストレージ アクセス ポートを使用することで、最大で8台のDGX-1サーバをサポートできます。さらに大規模な導入の場合は、Cisco Nexus 7000を使用すれば、スイッチあたり100GbEのポートを最大192個サポートできます。あるいは、複数ペアのNexus 3000スイッチを中心のスパイン スイッチに接続して、リーフスパイン トポロジを実装することもできます。

4.8 RDMA over Converged Ethernet

ダイレクト メモリ アクセス (DMA) は、ハードウェア サブシステム (ディスク ドライブ コントローラ、サウンド カード、グラフィックス カードなど) からシステム メモリにアクセスできるようにして、CPUの処理サイクルを使用せずにデータの読み取り / 書き込みを実行します。RDMAは DMAの機能が拡張されたもので、アプリケーション メモリ間でのサーバ間データ転送をネットワーク アダプタで実行します。その際、OSやデバイス ドライバの関与を一切必要としないゼロコピー機能が使用されます。このアプローチでは、読み取り / 書き込みと送受信の処理でカーネルをバイパスすることにより、CPUのオーバーヘッドとレイテンシを大幅に削減します。

RoCEはRDMA over Ethernetの中でも最も幅広く導入されている実装であり、新しいConverged Enhanced Ethernet (CEE) 標準を活用しています。今では、多くのハイエンド ネットワーク アダプタ、統合ネットワーク アダプタ、ネットワーク スイッチで標準機能として利用できるようになりました。従来のイーサネットは、ネットワーク トラフィック向けのベストエフォート型転送メカニズムを採用しており、GPUノード間の通信に必要とされる低レイテンシ、広帯域幅の転送には適しません。CEEにより、ロスレスの物理レイヤ ネットワーク媒体が実現し、ネットワーク上の特定トラフィック フローに帯域幅を任意で割り当てることもできます。

ロスレスで順序どおりのイーサネット パケット転送を保証するため、CEEネットワークではPriority Flow Control (PFC; 優先度ベースフロー制御) とEnhanced Transmission Selection (ETS; 拡張伝送選択) を採用しています。PFCでは、特定のサービス クラス (CoS) ごとにポーズフレームを送信することができます。その結果、特定のネットワーク トラフィックを制限しながら、他のトラフィックは自由に流れるようにすることができます。ETSは、各CoSに対して特定の帯域幅を割り当てることができるので、ネットワーク使用量をさらに細かく制御できます。

他のすべてのトラフィックよりもRoCEを優先する機能により、RoCEと従来のIPトラフィック（このソリューションで示したNFSストレージ アクセス トラフィックなど）の両方で100GbEリンクを使用できます。

5 テクノロジの要件

このセクションでは、このソリューションの検証で使用したハードウェアとソフトウェアについて説明します。セクション7「ソリューションの検証」に記載されているテストは、すべて以下に示すハードウェアとソフトウェアを使用して行われました。

メモ : このリファレンス アーキテクチャで検証される構成は、ラボにおける機器の可用性に基づいて検証されており、テスト対象のハードウェアの要件または制限に基づいて検証されたものではありません。

5.1 ハードウェア要件

表1は、このソリューションの検証で使用したハードウェア コンポーネントを示しています。このソリューションの特定の実装で使用されるハードウェア コンポーネントは、要件に応じて変更される場合があります。

表1) ハードウェア要件

ハードウェア	数量
NVIDIA DGX-1 GPUサーバ	4
NetApp AFF A800システム	ハイアベイラビリティ (HA) ペア×1、1.92TB NVMe SSD×48を含む
Cisco Nexus 3232Cネットワーク スイッチ	2

5.2 ソフトウェア要件

表2は、このソリューションの実装に必要なソフトウェア コンポーネントを示しています。このソリューションの特定の実装で使用されるソフトウェア コンポーネントは、要件に応じて変更される場合があります。

表2) ソフトウェア要件

ソフトウェア	バージョン
NetApp ONTAP	9.4
Cisco NX-OSスイッチ フームウェア	7.0(3)I6(1)
NVIDIA DGX-1オペレーティング システム	Ubuntu 16.04 LTS
Dockerコンテナ プラットフォーム	18.03.1-ce [9ee9f40]
コンテナ バージョン	netapp_1.7.0.2 (nvcr.io/nvidia/tensorflow:18.04-py2ベース)
機械学習フレームワーク	TensorFlow 1.7.0
Horovod	0.11.3
OpenMPI	3.1.0
ベンチマーク ソフトウェア	TensorFlowベンチマーク[1b1ca8a]

6 ソリューション アーキテクチャ

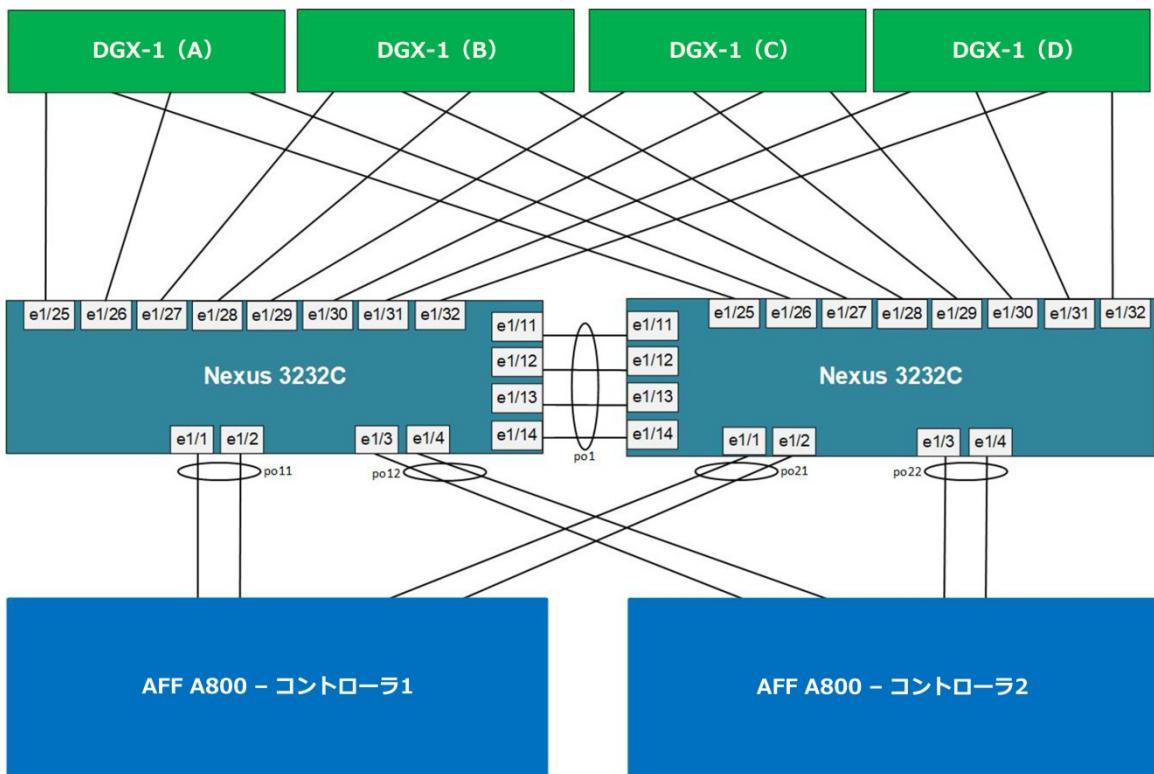
このアーキテクチャは、DLワークフロードを実行するための要件に合わせて検証されました。この検証により、データ サイエンティストは事前検証されたインフラにDLフレームワークとアプリケーションを導入できるようになります。その結果、リスクの排除が容易になり、企業は社内データから価値ある分析情報を引き出すことに集中できるようになります。また、このアーキテクチャでは、インフラの変更や調整を一切行うことなく、他のHPCワークフロードで卓越したストレージ パフォーマンスを実現できます。

6.1 ネットワーク トポロジとスイッチの構成

このソリューションでは、IBの代わりにRoCEを使用して、DGX-1サーバ間の通信に必要とされる広帯域幅、低レイテンシの接続を提供します。Cisco Nexusスイッチは、PFCを実装することでRoCEをサポートしています。この仕組みにより、共有リンク上でRoCEトラフィックを従来のIPトラフィックよりも優先することができ、RoCEとIPの両方で100GbEリンクを同時に使用することが可能になります。

このアーキテクチャでは、クラスタ間接続とストレージ アクセス用のプライマリ ネットワークに、Cisco Nexus 3232C 100Gbイーサネット スイッチのペアを使用します。2台のスイッチは、標準ポート チャネルとして構成される4つの100Gbネットワーク ポートを使用して相互に接続されます。このスイッチ間リンク (ISL) ポート チャネルにより、ホストまたはストレージ システムのリンクで障害が発生した場合でも、スイッチ間ではトラフィックの流れを維持できます。各ホストはアクティブ / パッシブのボンディングのペアによってNexusスイッチに接続されます。また、リンクレイヤの冗長性を確保するため、各ストレージ コントローラは2ポートのLACPポート チャネルによって各Nexusスイッチに接続されます。図6は、ネットワーク スイッチ ポート構成を示しています。

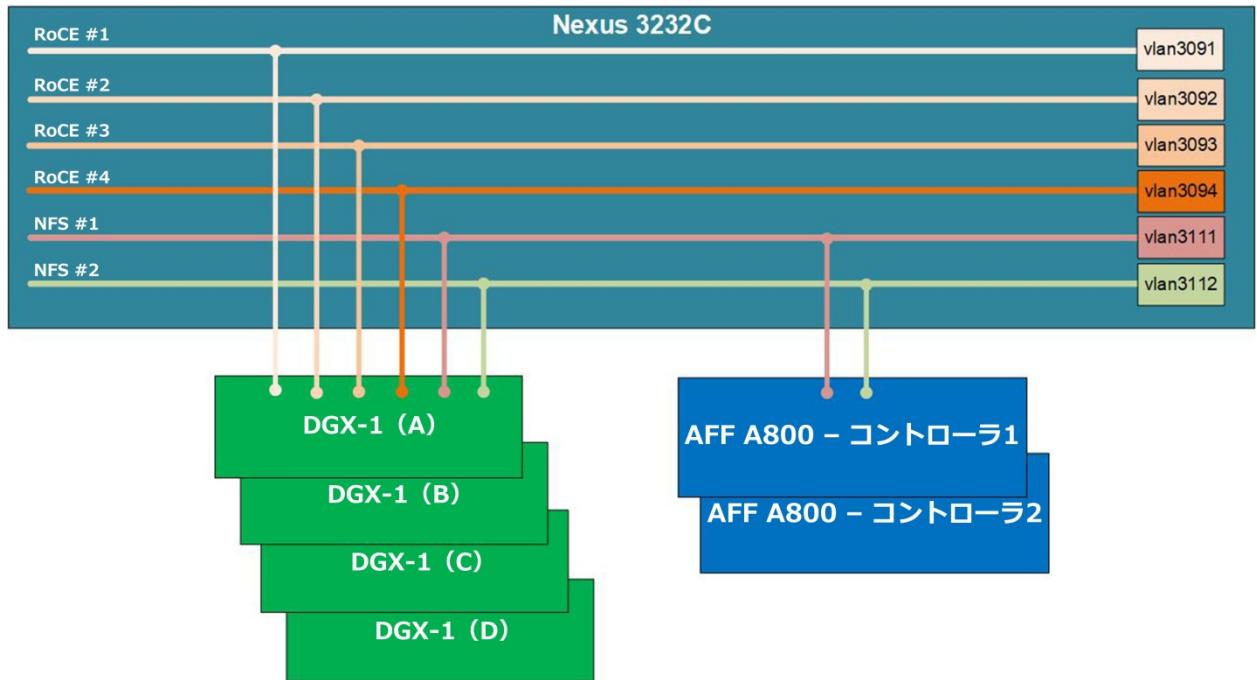
図6) ネットワーク スイッチ ポート構成



RoCEとNFSの両方のストレージ トラフィックに対応するため、複数の仮想LAN (VLAN) をプロビジョニングしました。RoCEトラフィック専用に4つのVLAN、NFSストレージ トラフィック専用に2つのVLANを使用しています。4つの別個のVLANとIP範囲を使用して、各RoCE接続で対称ルーティングが行われるようにしています。また、NVIDIAソフトウェア スタックがこれらの接続を管理することで、帯域幅の集約とフォールト トレランスを実現します。このソリューションではストレージ アクセスにNFSv3を使用しますが、NFSv3はマルチパス アクセスに対応していないため、2つのVLANを使用して複数の専用NFSマウントを可能にしています。このアプローチでは追加のフォールト トレランスを確保できませんが、複数リンクの使用を可能にすることで、利用できる帯域幅を増やしています。各スイッチでPFCは、4つすべてのRoCE VLANを優先クラスに割り当てるよう設定されます。NFSのVLANはデフォルトのベストエフォート クラスに割り当てられます。すべてのVLANにはジャンボ フレームが構成され、最大転送単位 (MTU) サイズは9,000です。

DGX-1サーバのスイッチ ポートはトランク ポートとして構成され、RoCE VLANとNFS VLANのすべてが許可されます。ストレージ システム コントローラ用に構成されたポート チャネルもトランク ポートですが、NFS VLANのみ許可されます。図7は、DGX-1サーバ ポートとストレージ システム ポートのVLAN接続を示しています。

図7) DGX-1ポートとストレージ システム ポートのVLAN接続



RoCE トラフィック サービスに高い優先順位を割り当てるため、ホスト ネットワーク アダプタは各 RoCE VLAN のトラフィックに CoS 値の 4 を割り当てます。スイッチは、CoS 値 4 のトラフィックに「ドロップなし」サービスを提供する QoS ポリシーを使用して構成されます。NFS トラフィックにはデフォルトの CoS 値 0 が割り当てられます。この値は、スイッチのデフォルトの QoS ポリシーに分類され、ベストエフォートのサービスが提供されます。

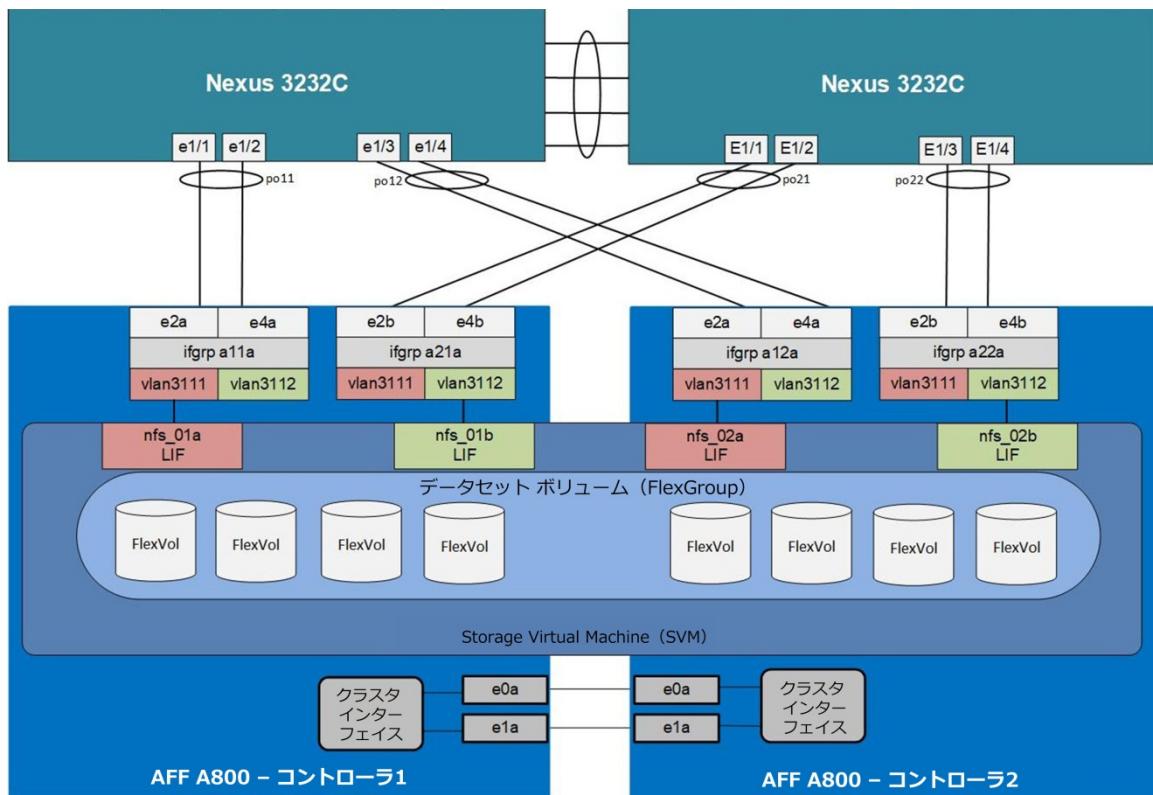
次に各 DGX-1 ポートで PFC が有効化されます。その結果、スイッチ ポートからサービスの特定のクラスに対してポーズフレームを送信できるようになり、スイッチでの輻輳が解消されます。この構成では、輻輳が発生した場合に、ETS によって帯域幅の 95% を RoCE トラフィックに割り当てることで、RoCE トラフィックと NFS トラフィックとの間で動的なリソース割り当てを行なながら、ノード間通信に高い優先順位を設定できます。帯域幅割り当てを動的に変更して、高いストレージ パフォーマンスと、少ないノード間通信が要求されるワークロード向けに最適化することもできます。

6.2 ストレージ システムの構成

このアーキテクチャで任意のワークロードのストレージ ネットワーク要件に対応するためには、各ストレージ コントローラに、ストレージ クラスタのインターフェイスクロス接続に必要なオンボード ポートに加えて、4つの100GbE ポートをプロビジョニングする必要があります。図8は、ストレージ システムの構成を示しています。各コントローラは、スイッチごとに2ポートの LACP インターフェイス グループ (図8のifgrp) を使用して構成されます。このインターフェイス グループは、データ アクセスにおいてスイッチあたり最大 200Gb/秒 の耐障害性に優れた接続を実現します。2つの VLAN が NFS ストレージ アクセス用にプロビジョニングされ、両方のストレージ VLAN がスイッチから各インターフェイス グループへトランкиングされています。この構成では、各ホストが複数のインターフェイスを通じてデータに同時アクセスできます。そのため、各ホストが利用できる可能性のある帯域幅が増えます。

ストレージシステムからのすべてのデータアクセスは、このワークロード専用のStorage Virtual Machine (SVM)からのNFSアクセスを通じて行われます。SVMは合計で4つの論理インターフェイス (LIF)を使用して構成され、ストレージVLANごとに2つのLIFが配置されます。各インターフェイスグループが1つのLIFをホスティングするので、コントローラのVLANごとに1つのLIFが割り当てられ、各VLANが専用のインターフェイスグループを持つことになります。ただし、コントローラでは2つのVLANが両方のインターフェイスグループにトランкиングされます。この構成は、各LIFに対し、同じコントローラ上の別のインターフェイスグループにフェイルオーバーする手段を提供します。したがって、ネットワーク障害が発生した場合でも、両方のコントローラがアクティブのままになります。

図8) ストレージシステムの構成

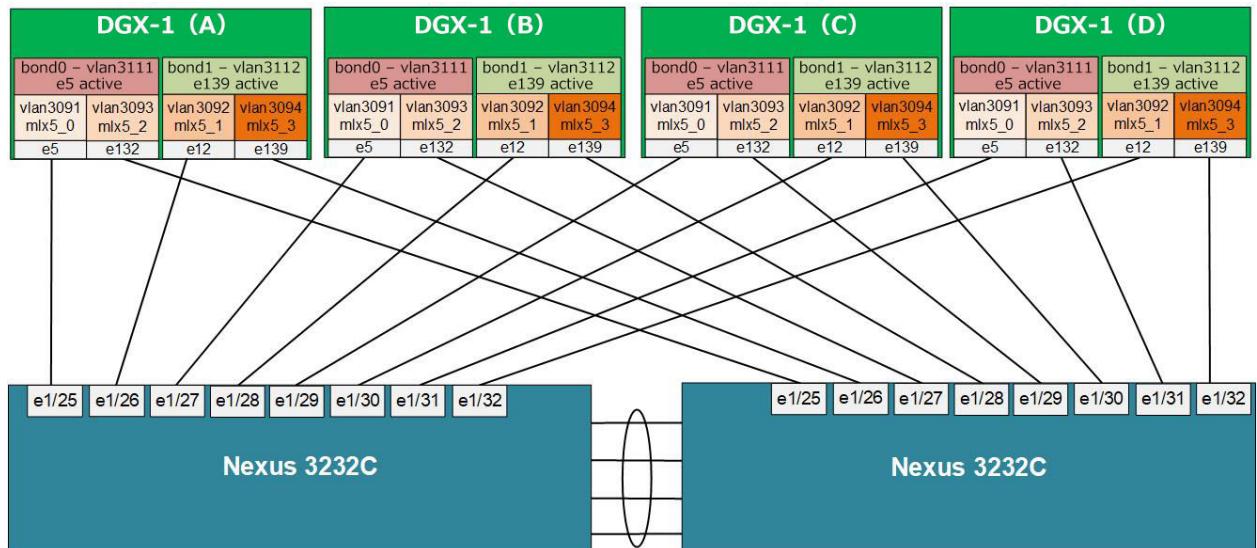


論理ストレージをプロビジョニングする場合、このソリューションではFlexGroupボリュームを使用してストレージの単一プールを提供し、そのプールをストレージクラスタ内のノード全体で分散します。各コントローラは46個のディスクパーティションのアグリゲートをホストし、両方のコントローラで各ディスクを共有します。FlexGroupをデータSVMに導入するときは、各アグリゲートで複数のFlexVolがプロビジョニングされ、それらのFlexVolが結合されてFlexGroupになります。このアプローチでは、アレイの最大容量までスケールアップできる単一のストレージプールをストレージシステムによって提供できます。また、アレイ内のすべてのSSDを同時に活用することで、卓越したパフォーマンスを実現できます。NFSクライアントは、SVMに対してプロビジョニングされた任意のLIFを介して、単一のマウントポイントとしてのFlexGroupにアクセスできます。ノードをストレージクラスタに追加するだけで、容量とクライアントアクセス帯域幅を増やすことができます。

6.3 ホストの構成

各DGX-1には、ネットワーク接続用に、Mellanox ConnectX4シングルポート ネットワーク インターフェイス カードを4枚実装しました。このカードは最大で100Gbイーサネット接続の速度で動作し、RoCEに対応しているので、クラスタ インターコネクト アプリケーション用の、IBに代わる低コストの代替手段となります。各100Gbポートは対応するスイッチ上のトランク ポートとして構成され、それぞれ4つのRoCE VLANと2つのNFS VLANが許可されます。図9は、DGX-1ホストのネットワーク ポートとVLANの構成を示しています。

図9) DGX-1ホストのネットワーク ポートとVLANの構成



各物理ポートは、RoCE接続用に、4つのRoCE VLANのうち1つでVLANインターフェイスとIPアドレスをホスティングします。Mellanox ドライバは、各RoCE VLANにネットワークCoS値の4を適用するよう設定され、優先順位の高いロスレスなサービスをRoCEクラスに保証するために、スイッチでPFCが設定されます。RoCEでは、複数のリンクを単一論理接続に集約できませんが、NVIDIA NCCL通信ソフトウェアでは、複数のリンクを使用して帯域幅の集約とフォールト トレンスを実現することができます。

NFSストレージ アクセスについては、各スイッチへのリンクを使用することで、2つのアクティブ / パッシブのボンディングが作成されます。各ボンディングでは2つのNFS VLANのうち1つでVLANインターフェイスとIPアドレスがホスティングされ、各ボンディングのアクティブ ポートはそれぞれ別々のスイッチに接続されます。この構成では、各NFS VLANで最大100Gbの帯域幅が提供され、ホストリンクまたはスイッチの障害が発生した場合に備えて冗長性も確保されます。RoCE接続で最適なパフォーマンスを実現するため、すべてのNFS トライフィックはデフォルトのベストエフオートQoSクラスに割り当てられます。すべての物理インターフェイスとボンディングインターフェイスは、MTUサイズ9,000で構成されます。

データ アクセスのパフォーマンスを高めるため、DGX-1サーバからストレージ システムへのNFSv3マウントが複数回行われます。各DGX-1サーバには2つのNFS VLANが構成され、各VLANにIPインターフェイスが1つずつ割り当てられます。AFF A800システムのFlexGroupボリュームは、各DGX-1のそれぞれのVLANにマウントされるので、サーバからストレージ システムへの完全に独立した接続が提供されます。このワーカクロードに必要とされるパフォーマンスは1つのNFSマウントで達成できますが、複数のマウント ポイントを定義することで、ストレージの処理負荷が高い他のワーカロードで追加のストレージ アクセス帯域幅を使用できるようにしています。

7 ソリューションの検証

このセクションでは、本ソリューションの動作とパフォーマンスを検証するためにネットアップが行ったテストについて説明します。ネットアップでは、セクション5「テクノロジの要件」の一覧に記載した特定の機器およびソフトウェアを使用して、このセクションで説明するすべてのテストを実施しました。

7.1 検証テストのプラン

このソリューションの検証テストは、アーキテクチャのスケーラビリティを実証する目的で、標準的なベンチマークと複数のコンピューティング設定を使用して行われました。AFF A800システムでは、1つのFlexGroupボリュームを使用してImageNetデータセットをホストしました。このFlexGroupボリュームへのアクセスには、外部ストレージ アクセスについてNVIDIAが推奨しているとおり、最大4台のDGX-1サーバからNFSv3を使用しました。テスト対象のすべてのモデルで、機械学習フレームワークとしてTensorFlowを使用し、各テスト ケースでコンピューティングとストレージのパフォーマンス指標を収集しました。主な指標データを、セクション7.2「検証テストの結果」に示しています。

トレーニング速度を実証するため、次に示すConvolutional Neural Network (CNN; 置込みニューラル ネットワーク) モデルをさまざまな複雑度のコンピューティングとストレージに組み合わせました。

- **ResNet-152** は、一般的に最も精度の高いトレーニング モデルとみなされています。
- **ResNet-50** は、AlexNet よりも高い精度をより短い処理時間で実現します。
- **VGG16** は、最高速の GPU 間通信を生み出します。
- **Inception-v3** は、もう 1 つの一般的な TensorFlow モデルです。

こうしたモデルをそれぞれさまざまなハードウェア / ソフトウェア構成でテストすることで、各オプションのパフォーマンスに対する影響を調査しました。

- 各モデルのテストでは、合成データと ImageNet リファレンス データセットの両方を使用しました。コンピューティング クラスタのスケーラビリティの評価とストレージ アクセス パフォーマンスの評価では、単体の DGX-1 サーバ内および複数の DGX-1 サーバ間の両方で GPU を追加してテストを行いました。
- データを GPU メモリにコピーする前の CPU 処理のオーバーヘッドを軽減するため、歪みを無効にして ImageNet データを使用しました。
- Tensor コアがもたらすパフォーマンスの向上を実証するため、各モデルのテストを Tensor コアと CUDA コアを使用して行いました。
- GPU のパフォーマンスを高めると、ストレージ アクセス要件が増えることになりましたが、そうした要件は AFF A800 システムの能力によって容易に対応できることがわかりました。
- それぞれの DL モデルを、さまざまなバッチ サイズでテストしました。バッチ サイズを増やすことでシステムに対していくつかの影響があり、最終的に、全体的なトレーニング速度の向上、GPU 間通信要件の低下、ストレージ帯域幅要件の上昇につながりました。各モデルでテストしたバッチ サイズは次のとおりです。
 - ResNet-50 で 64、128、256
 - その他すべてのモデルで 64 および 128
- RoCE をインターネットとして使用する複数の GPU 間での各モデルのスケーラビリティを実証するため、1 台、2 台、4 台の DGX-1 サーバで各モデルをテストしました (Horovod を使用)。

- 推論の実行では、すべてのモデルで最大バッチ サイズ（ResNet-50 で 256、その他すべてのモデルで 128）、32 基の GPU（Tensor コアと CUDA コア）、ImageNet データセットを使用しました。
- 少なくとも 2 回のトレーニング期間を終了してから、すべてのパフォーマンス指標を収集しました。複数の期間にわたってトレーニングを実行したところ、パフォーマンス結果がわずかに向上することが判明しました。各テストを 5 回実施し、得られたパフォーマンス指標の平均値を報告しています。

7.2 検証テストの結果

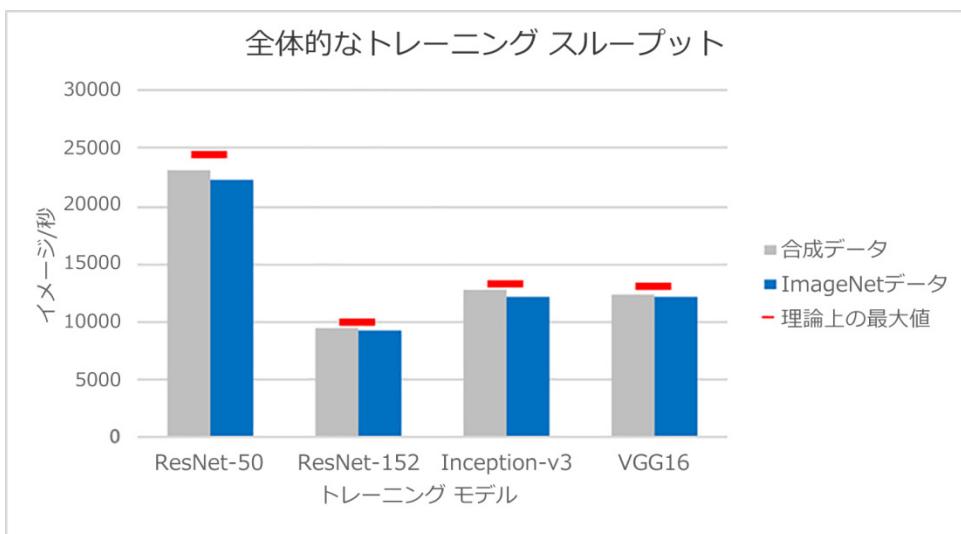
前述のとおり、ネットアップは本ソリューションの全般的な動作とパフォーマンスを評価するため、さまざまなテストを実施しました。このセクションでは、テスト中に収集した、コンピューティングとストレージのパフォーマンスに関する主なデータを示します。テスト結果の完全な記録は付録に記載しています。本レポートの以降のサブセクションで示すデータについては、次の説明に留意してください。

- モデルのトレーニング パフォーマンスは、1 秒あたりのイメージの数として測定しました。
- ストレージ パフォーマンスは、スループット (MB/秒) とレイテンシ (マイクロ秒) を使用して測定しました。ストレージ システムの残りのパフォーマンス能力を評価するため、ストレージ システム CPU のデータも収集しました。
- 各システムを複数のバッチ サイズでテストしました。バッチ サイズを大きくするほど、全体的なトレーニング スループットは向上します。モデルごとにテストした最大のバッチ サイズのみを以下に示します。テストした各バッチ サイズのデータは付録に記載しています。
 - ResNet-50 のテストでは、バッチ サイズ 256 を使用しました。
 - ResNet-152、Inception-v3、VGG16 のテストでは、バッチ サイズ 128 を使用しました。

全体的なトレーニング スループット

図10は、各モデルで達成された、トレーニング イメージの1秒あたりの最大数を示しています。テストでは、最大のパフォーマンスを得るため、Tensorコアを使用しました。図10では、ベースライン比較用のImageNetデータと合成データを使用して、32基のGPUで達成されたトレーニング スループットを比較しています。また、達成可能な理論上の最大値も示しています。これは、すべてのGPUが、互いにパラメータを更新することなく、それぞれに独立して合成データのトレーニングを行った場合を想定したものです。図10からわかるとおり、ImageNetデータで達成したスループットは、合成データのスループットにきわめて近いものです。

図10) 全モデルのトレーニング スループット



GPUワークロードのパフォーマンス

次に示す一連のデータでは、フル ロード状態にあるときのDGX-1サーバの要件を満たすストレージシステムの能力を実証します。図11は、32基のGPUを使用して各モデルを実行したときの、DGX-1サーバのGPU利用率とストレージ帯域幅を示しています。グラフからわかるとおり、ストレージ帯域幅は、初期データがストレージからTensorFlowパイプライン キャッシュに読み込まれるときに急激に上昇し、データセットの大部分がDGX-1のローカル メモリに格納されるとともに、徐々に下降していきます。

すべてのデータがローカル メモリに収まると、ストレージ アクセスはほぼゼロになります。DGX-1のGPUは、ほぼ瞬時にデータの処理を開始します。GPU利用率は、テストの実行時間を通してほとんど一定です。このグラフは、VGG16モデルでバッチ サイズ128を使用したときの結果を示しています。この組み合わせが、今回のテストで最も高いGPU利用率を示しました。他のモデルのグラフについては、付録をご覧ください。GPU利用率の目盛りが示しているのは、全GPUの利用率の合計です。したがって、この32基のGPUでテストしたケースでは、達成が可能な最大利用率は3,200%になります。

図11) GPU利用率とストレージ帯域幅 (VGG16)

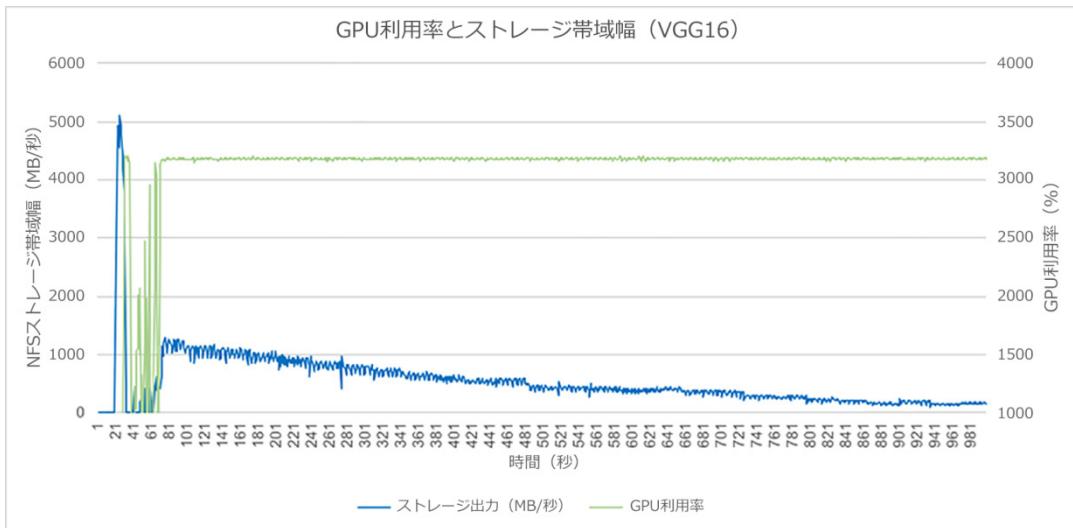


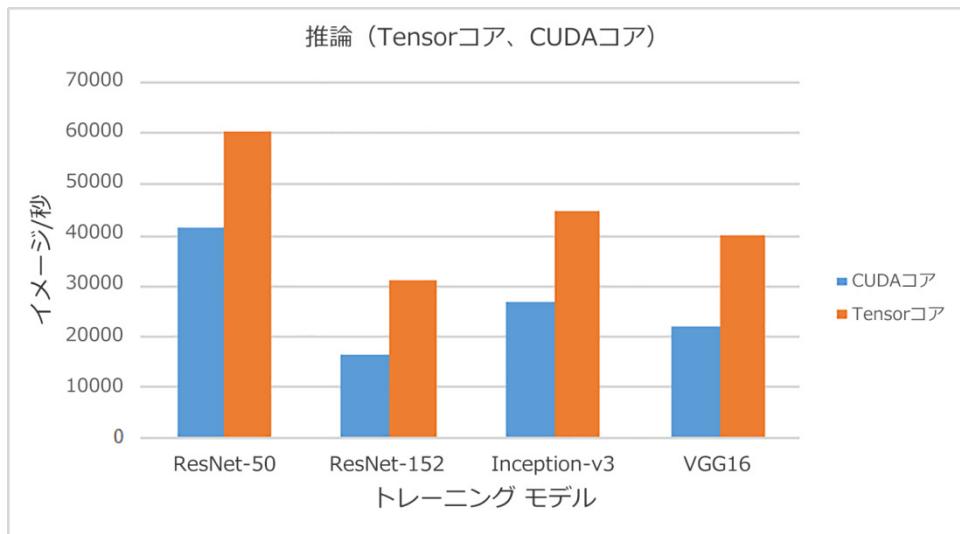
図11で示すとおり、GPU利用率は32基すべてのGPUで95%以上を維持し、ストレージ システムからのデータの量に関係なく安定しています。ストレージ システムからのデータ転送は最初に5GB/秒で行われ、2GB/秒を下回るあたりから下降し、トレーニング期間の残り時間にわたってほぼゼロに近づいていきます。この結果から、このワークロードでは、ストレージ アクセスはGPUのパフォーマンスにとってボトルネックではないことがわかります。ローカル メモリ容量を超える大規模なデータセットでは、ストレージ アクセスのパフォーマンスはトレーニング期間の後半まで安定状態のスループット レートで維持されます。また、図11では、ストレージ帯域幅の変化に応じたGPU利用率の変化を対比しています。トレーニング フェーズ全体に必要とされる時間のデータが示されていないのは、トレーニング フェーズの進行につれて、ストレージ帯域幅がゼロ近くまで徐々に下降していくからです。

GPUを使用した推論

推論とは、新しいオブジェクト セットを評価するためのDLモデルを導入し、トレーニング フェーズ中に確認された同様の予測精度を使用して予測を行うプロセスです。イメージ データセットを含むアプリケーションにおいて、推論の目的は、入力イメージを分類し、できるかぎり早くリクエストに応答することです。高いスループットを達成することに加えて、レイテンシを最小限に抑えることも重要になります。

このフェーズでは、推論の実証とスループット指標の測定にNetApp ONTAP AIを使用しました。図12は、推論時に処理できる1秒あたりのイメージ数を示しています。このテストでは、TensorコアとCUDAコアを使用してテストした各モデルでImageNetを使い、32基のGPUで達成されたスループットを比較しています。NetApp ONTAP AIの処理能力を利用すると、Tensorコアでは大量のイメージを瞬時に分類できます。

図12) 全モデルでの推論 (TensorコアとCUDAコア)



AIトレーニング ワークロードでのAFF A800システムのパフォーマンス

各テスト対象モデルを使用したストレージ システムのパフォーマンスの実証では、ストレージ帯域幅、レイテンシ、CPUヘッドルームのデータを収集しました。図13から図15は、実際のデータを使用して各モデルをテストしたときのストレージ システム指標を示しています。このようなストレージに焦点を当てたテストは、ストレージ ワークロードを増やし、最大負荷時のシナリオを示すために、より大きいバッチ サイズを使って実行しました。

各指標において、32基のGPUを使用する各モデルで生成される総ワークロードは、AFF A800システムのパフォーマンス限界の範囲内に十分収まっています。トレーニング ワークロードの参考基準を示すため、fio (flexible I/O) と64Kシーケンシャル読み取りI/Oプロファイルを使用して、人工的な

ワークロードを生成しました。fioを使って生成したワークロードの場合、スループットは15GB/秒を超えた時点でピークに達し、読み取りレイテンシは1ミリ秒を十分下回るレベルを維持しました。CPU利用率は50%をわずかに下回りました。限られた台数のDGX-1サーバで可能な限り最大のスループットを達成するため、各サーバで追加のNFSマウントと複数のfioジョブを使用しました。

メモ： NetApp AFF A800 HAペア1組で、NASワークロードに対して1ミリ秒未満のレイテンシで最大25GB/秒までサポートされることが証明されました。

図13) 全モデルのストレージ帯域幅

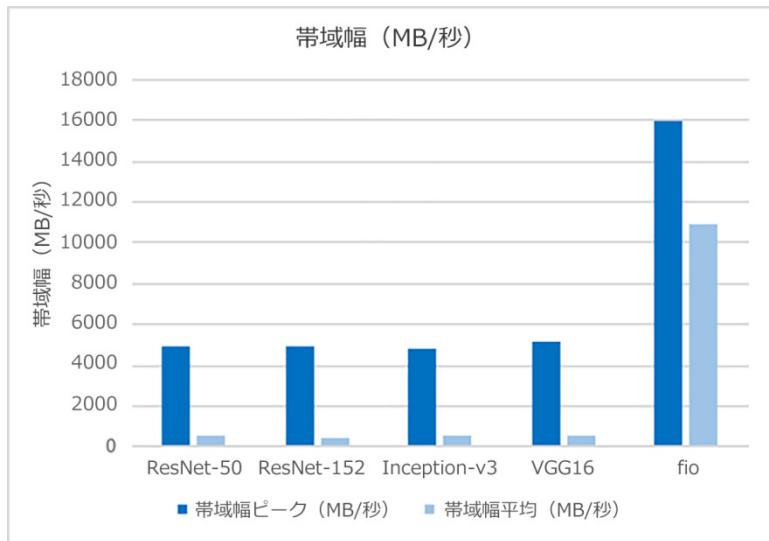


図14) 全モデルのストレージ レイテンシ

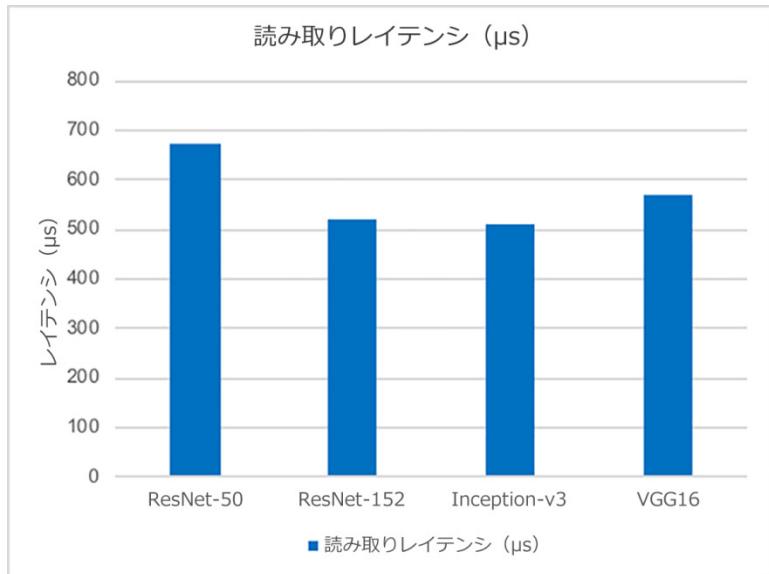
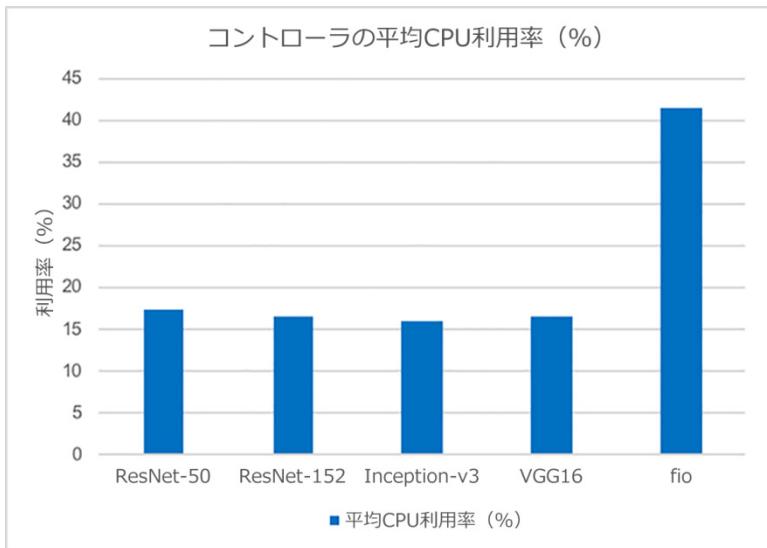


図15) 全モデルのストレージCPU利用率



7.3 ソリューションのサイジングに関するガイダンス

このアーキテクチャは、NVIDIA DGX-1サーバとNetApp AFFシステムを使用してハイパフォーマンス コンピューティング (HPC) インフラを実装したいと考えているお客様やパートナー様のための リファレンス アーキテクチャとして考案されたものです。

この検証で実証されているとおり、AFF A800システムは4台のDGX-1サーバによって生成されるDL トレーニング ワークロードに容易に対応でき、HAペアには約70%のヘッドルームが残ります。したがって、AFF A800システムはDGX-1サーバを追加した場合にも対応できます。ストレージ パフォーマンス要件がさらに高い、より大規模な導入であっても、NetApp ONTAPクラスタにAFF A800システムを追加することで対応できます。ONTAP 9は1つのクラスタで最大12のHAペア（24 ノード）をサポートでき、このソリューションで検証したFlexGroupテクノロジを使用すれば、1つのボリュームで20PBを超える容量を提供できます。この検証で使用したデータセットは比較的小規模なものでしたが、ONTAP 9はリニアなパフォーマンス拡張性によって、驚異的な大容量までスケールアップできます。このドキュメントで検証したレベルに匹敵するパフォーマンスを各HAペア で実現できるからです。

小規模なDGX-1クラスタの場合は、AFF A220またはAFF A300システムによって、コストを抑えながら十分なパフォーマンスを実現できます。ONTAP 9は混合モデルのクラスタをサポートできるので、最初は小規模な構成でスタートし、容量とパフォーマンスの要件が高まるにつれて、より多くの、またはより大容量のストレージ システムをクラスタに追加できます。

ネットワークの観点では、この検証済みアーキテクチャは、各Cisco Nexus 3232Cスイッチで利用可能な32ポートのうち、16ポートしか使用していません。各スイッチは追加のストレージ アクセス ポートを使用すると最大で8台のDGX-1サーバをサポートできるので、ネットワーク機能を追加しなくてもコンピューティング能力を大幅に拡張できます。さらに大規模な導入の場合は、Cisco Nexus 7000を使用すれば、スイッチあたり100GbEのワイヤスピード ポートを最大192個サポートできます。あるいは、複数ペアのCisco Nexus 3000スイッチを中央のスパイン スイッチに接続して、リーフ スパイン トポロジを実装することもできます。

このAIトレーニング ワークロードを使って実行した検証テストの結果から、各DGX-1には約2GB/秒 のストレージ スループットが必要です。AFF A800システムは他の手段によって生成された同様の ワークロードで25GB/秒のスループットを達成する能力が証明されたことから、このアーキテクチャ はAFF A800 HAペアあたり9台以上のDGX-1サーバをサポートできることになります。

8 まとめ

DGX-1サーバはきわめて強力なDLプラットフォームであり、同等に強力なストレージとネットワークのインフラを装備することで、最大の価値を発揮できます。NetApp AFFシステムとCisco Nexusスイッチを組み合わせることで、この検証済みアーキテクチャを必要に応じてほぼあらゆるスケールで実装できます。1台のDGX-1と1台のAFF A220システムの構成から始まり、最大96台のDGX-1サーバを12ノードのAFF A800クラスタで構成することもできます。AFFシステムは、NetApp ONTAPの卓越したクラウド統合機能とソフトウェアで定義される機能を組み合わせることで、DLプロジェクトを成功に導く、エッジ、コア、クラウドにわたるデータパイプラインに全面対応できます。

謝辞

このNetApp Verified Architectureプロジェクトにご尽力いただいたNVIDIAの優秀な技術者の皆様、Darrin Johnson氏、Tony Paikeday氏、Robert Sohigian氏、James Mauro氏に感謝を申し上げます。また、この調査は、ネットアップチームの主要メンバーであるRobert FranzとKesari Mishraの支援と助言がなければ完了することができませんでした。

本ドキュメントに関連する調査に有用な洞察や専門知識を提供してくださったすべての人たちに心から感謝いたします。

詳細情報の入手方法

このドキュメントに記載されている情報の詳細については、以下のリソースを参照してください。

- NVIDIA DGX-1 サーバ
 - NVIDIA DGX-1サーバ
<https://www.nvidia.com/ja-jp/data-center/dgx-1/>
 - NVIDIA Tesla V100 TensorコアGPU
<https://www.nvidia.com/ja-jp/data-center/tesla-v100/>
 - NVIDIA GPU Cloud
<https://www.nvidia.com/ja-jp/gpu-cloud/>
- NetApp AFF システム
 - AFFデータシート
<https://www.netapp.com/jp/media/ds-3582.pdf>
 - NetApp Flash Advantage for AFF
<https://www.netapp.com/us/media/ds-3733.pdf>
 - ONTAP 9.xのドキュメント
<http://mysupport.netapp.com/documentation/productlibrary/index.html?productID=62286>
 - NetApp FlexGroupテクニカル レポート
<https://www.netapp.com/jp/media/tr-4557.pdf>
- NetApp Interoperability Matrix
 - NetApp Interoperability Matrix Tool
<http://support.netapp.com/matrix>

- Cisco Nexus ネットワーキング

Cisco Nexus 3232Cシリーズ スイッチの詳細については、次のリンクを参照してください。

- Cisco Nexus 3232Cシリーズ スイッチ

https://www.cisco.com/c/ja_jp/products/switches/nexus-3232c-switch/index.html

- Cisco Nexus 3232Cコンフィギュレーション ガイド

https://www.cisco.com/c/ja_jp/support/switches/nexus-3000-series-switches/products-installation-and-configuration-guides-list.html

- Cisco Nexus 3232Cコマンド リファレンス

https://www.cisco.com/c/ja_jp/support/switches/nexus-3000-series-switches/products-command-reference-list.html

- 機械学習フレームワーク

- TensorFlow:TensorFlow : 幅広い用途に対応するオープンソースの機械学習フレームワーク

<https://www.tensorflow.org/>

- HorovodUber が開発したオープンソースの TensorFlow 用分散学習フレームワーク

<https://eng.uber.com/horovod/>

- ブログ「Enabling GPUs in the Container Runtime Ecosystem」

<https://devblogs.nvidia.com/gpu-containers-runtime/>

- データセットとベンチマーク :

- ImageNet

<http://www.image-net.org/>

- TensorFlowベンチマーク

<https://www.tensorflow.org/performance/benchmarks>

付録

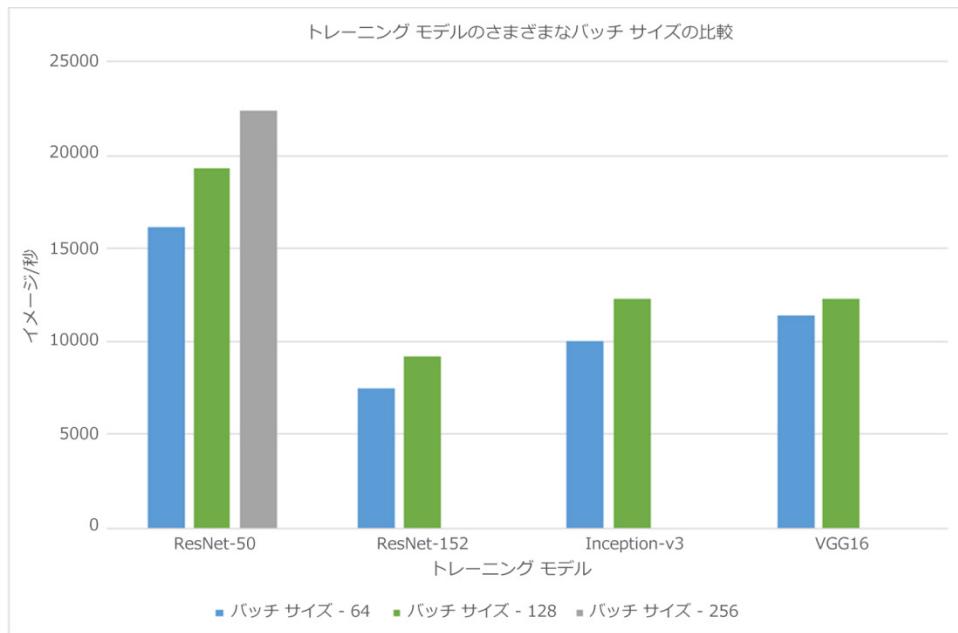
このセクションには、本アーキテクチャを使用して実行したテストのその他の結果を記載します。

各モデルのさまざまなバッチ サイズでのトレーニング速度

図16は、次の構成要素を使用する異なるトレーニング モデルでのさまざまなバッチ サイズの比較を示しています。

- GPU の数 : 32(4)DGX-1(1)
- コア : Tensor コア Tensor cores
- Batch sizes:バッチ サイズ : ResNet-50 では 64、128、256、その他のモデルでは 64、128

図16）トレーニング モデルのさまざまなバッチ サイズの比較

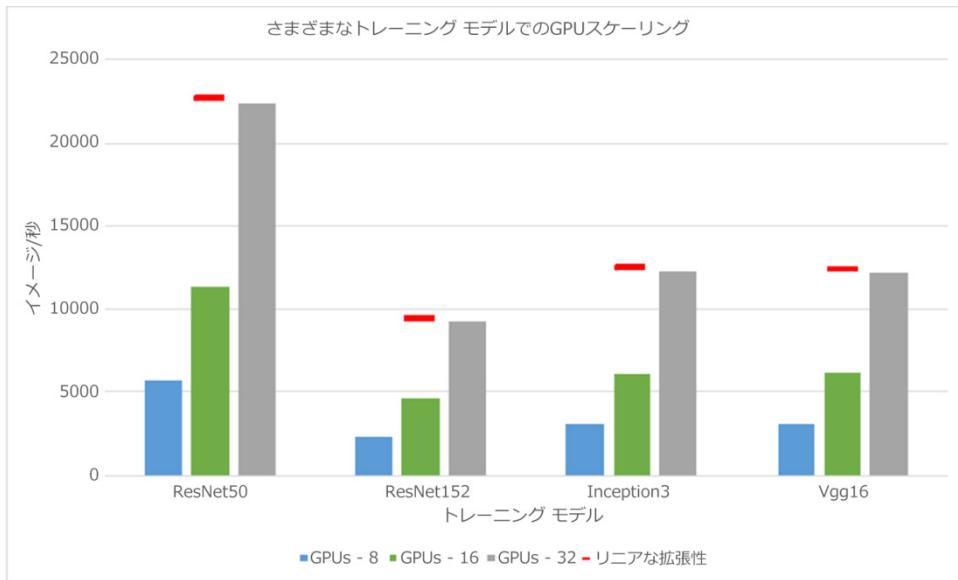


まとめ結論：トレーニング スループット パフォーマンスは、バッチ サイズを256または128に増やすと向上します。

各モデルでのGPUスケーリングの比較

- 図 17 は、次の構成要素を使用する異なるトレーニング モデルでの GPU スケーリングを示しています。
- GPU の数 : 8 (DGX-1 サーバ 1 台) 、 16 (DGX-1 サーバ 2 台) 、 32 (DGX-1 サーバ 4 台)
- コア : Tensor コア Tensor cores
- Batch sizes:バッチ サイズ : ResNet-50 では 256、その他のモデルでは 128

図17) さまざまなトレーニング モデルでのGPUスケーリング



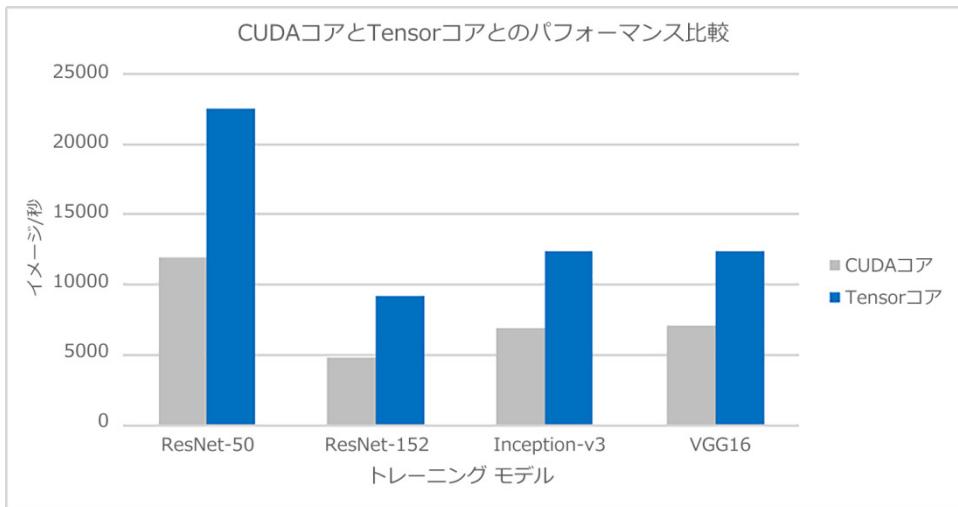
- まとめ結論：すべてのトレーニング モデルでリニアな GPU スケーリングが確認されました。

TensorコアとCUDAコアとの比較

図18は、次の構成要素を使用するCUDAコアとTensorコアのパフォーマンス比較を示しています。

- GPU の数 : 32(4)DGX-1(1)
- Cores:コア : Tensor コアと CUDA コア
- Batch sizes:バッチ サイズ : ResNet-50 では 256、その他のモデルでは 128

図18) CUDAコアとTensorコアのパフォーマンス比較



まとめ結論：TensorコアはCUDAコアよりも優れたパフォーマンスをもたらします。

全モデルでのGPUワークロード

図19から図21は、次の構成要素を使用するResNet-50、ResNet-152、Inception-v3のGPU利用率とストレージ帯域幅をそれぞれ示しています。

- GPU の数 : 32(4)DGX-1(1)
- コア : Tensor コア Tensor cores
- Batch sizes: バッチ サイズ : ResNet-50 では 256、その他のモデルでは 128

図19) ResNet-50のGPU利用率とストレージ帯域幅

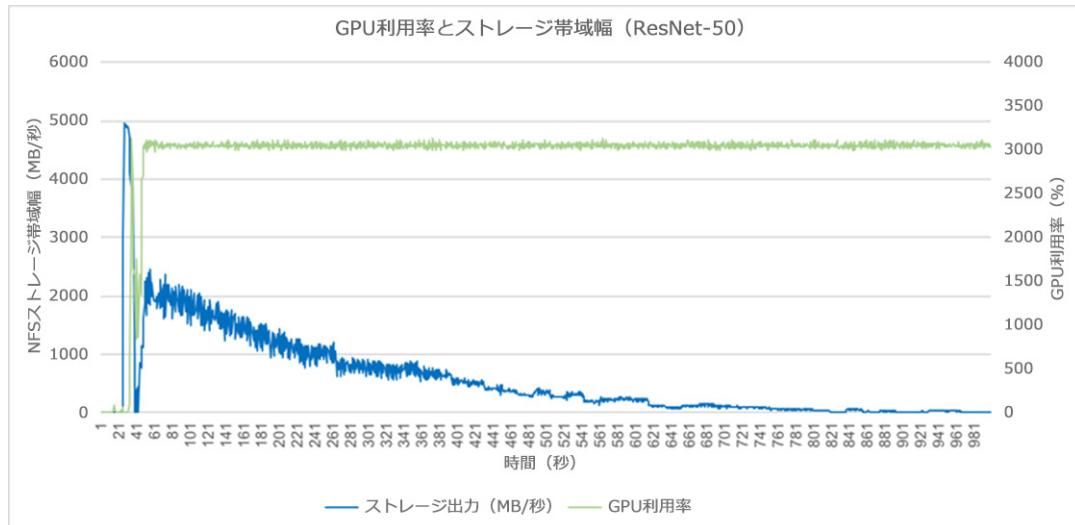


図20は、ResNet-152のGPU利用率とストレージ帯域幅を示しています。

図20) ResNet-152のGPU利用率とストレージ帯域幅

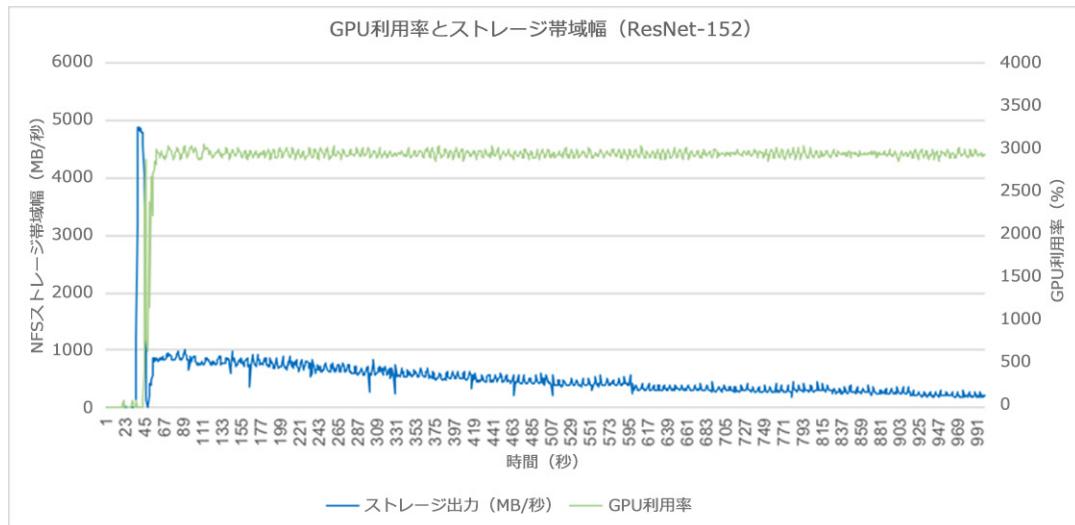
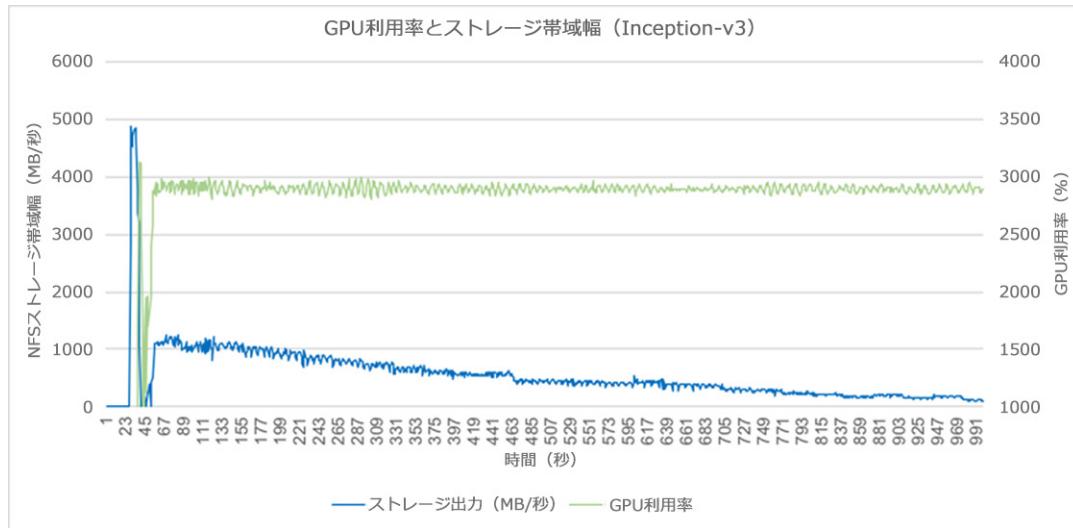


図21は、Inception-v3のGPU利用率とストレージ帯域幅を示しています。

図21) Inception-v3のGPU利用率とストレージ帯域幅



本ドキュメントに記載されている製品や機能のバージョンがお客様の環境でサポートされるかどうかについては、ネットアップ サポートサイトで[Interoperability Matrix Tool \(IMT\)](#) を参照してください。NetApp IMTには、ネットアップがサポートする構成を構築するために使用できる製品コンポーネントやバージョンが定義されています。サポートの可否は、お客様の実際のインストール環境が公表されている仕様に従っているかどうかによって異なります。

著作権に関する情報

Copyright © 1994–2018 NetApp, Inc. All Rights Reserved.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

ネットアップの著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、ネットアップによって「現状のまま」提供されています。ネットアップは明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。ネットアップは、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

ネットアップは、ここに記載されているすべての製品に対する変更を隨時、予告なく行う権利を保有します。ネットアップによる明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、ネットアップは責任を負いません。この製品の使用または購入は、ネットアップの特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許によって

本書に含まれるデータは市販品（FAR 2.101の定義に基づく）に関係し、データの所有権はNetApp, Inc.にあります。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となった米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc.の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b)項で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<http://www.netapp.com/jp/legal/netapptmlist.aspx>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。