

DDN A³I[®] SOLUTIONS FOR NVIDIA[®] CLOUD PARTNERS

NVIDIA HGX™ H100、H200、および B200 の 8GPU ベースプラットフォームを搭載したクラウドサービスパートナー向けに、完全に検証され最適化された AI 高性能ストレージ (HPS) ソリューションです。

1. NVIDIA クラウドパートナー向け統合的な DDN A ³ I ソリューション	2
2. HGX に最適化された DDN A ³ I ソリューション	8
2.1. DDN AI400X3 アプライアンス	8
2.2. DDN Insight ソフトウェア	9
3. NCP 向け DDN A ³ I リファレンスアーキテクチャ	10
3.1. ストレージパフォーマンスガイドライン	10
3.2. 推奨ストレージネットワーキング	12
3.3. HGX システム 128 台による NCP 導入例(1,024 GPUs)	16
3.4. HGX システム 256 台による NCP 導入例(2,048 GPUs)	17
3.5. HGX システム 512 台による NCP 導入例(4,096 GPUs)	18
3.6. HGX システム 1024 台による NCP 導入例(8,192 GPUs)	19
3.7. HGX システム 2048 台による NCP 導入例(16,384 GPUs)	20
4. クラウド導入の真価を解き放つために、DDN にご連絡ください	21
5. Appendix: HPS 1 GB/s 性能向け DDN ストレージ構成	22

概要 (Executive Summary)

DDN AI ソリューションは、NVIDIA HGX™ H100、H200、および B200™ システム上で動作する人工知能 (AI)、データ分析、高性能コンピューティング (HPC) アプリケーションにおいて、最適なデータパフォーマンスを大規模環境で実証済みです。DDN のデータプラットフォームは NVIDIA DGX SuperPOD™ 向けの認定ストレージソリューションであり、NVIDIA Eos、Selene、Cambridge-1 などの導入事例でも使用されています。

本書では、NVIDIA Cloud Partners (NCP) 向けに完全に検証されたリファレンスアーキテクチャを紹介します。これらのソリューションは、DDN AI400X3 アプライアンスと DDN Insight ソフトウェアを NVIDIA HGX プラットフォームと統合し、高性能ストレージ (HPS) を実現します。

1. NVIDIA クラウドパートナー向け統合的な DDN A³I ソリューション

DDN A³I ソリューション (Accelerated, Any-Scale AI) は、HGX システム上で動作する大規模 AI、データ分析、HPC (高性能コンピューティング) アプリケーションから最大限のパフォーマンスを引き出すように設計されています。DDN と NVIDIA システム間の緊密な統合により、予測可能な性能・容量・機能が提供され、データの提供と保存に関わるすべてのハードウェアとソフトウェア層が、高速・応答性・信頼性の高いアクセスのために最適化されています。

DDN A³I ソリューションは、NVIDIA と緊密に連携して設計・開発・最適化されています。DDN の AI アプライアンスと HGX システムとの深い統合により、高い信頼性を実現します。これらのソリューションは、さまざまな環境への柔軟な導入が可能であり、ワークロードの進化に応じて、容量や機能をシームレスにスケールできます。

また、DDN は世界最大規模のスーパーコンピュータを支える高度な技術を、HGX システム向けに展開し、導入や管理が容易な統合パッケージとして提供しています。DDN A³I ソリューションは、HGX システム上での大規模 AI、分析、HPC ワークロードに対し、最大限の効果を発揮することが実証されています。

本章では、NCP (NVIDIA Cloud Partner) 向けに設計された DDN A³I ソリューションの先進機能について説明します。

1.1. DDN A³I 共有パラレルアーキテクチャ

DDN A³I の共有パラレルアーキテクチャとクライアントプロトコルは、HGX システムに対して高いパフォーマンス、スケーラビリティ、セキュリティ、信頼性を提供します。このアーキテクチャでは、ドライブから HGX システム内の GPU 上で動作するコンテナ化されたアプリケーションに至るまで、複数の並列データパスが張り巡らされています。DDN のエンドツーエンドの真の並列処理により、データは高スループット、低レイテンシ、かつ大量の並行性をもってトランザクション処理されます。これにより、アプリケーションは HGX システムのアクセラレーテッド・コンピューティングサイクルを余すことなく活用できるようになり、結果としてアプリケーションのパフォーマンス向上と処理時間の短縮が実現します。さらに、この共有パラレルアーキテクチャには、冗長構成と自動フェイルオーバー機能が組み込まれており、ネットワーク接続やサーバが使用不能になった場合でも、高い信頼性・回復性・データ可用性を確保します。

1.2. DDN A³I によるディープラーニングワークフローの効率化

DDN A³I ソリューションは、HGX システム上で実行されるあらゆるスケールのディープラーニング（DL）ワークフローにおいて、エンドツーエンドのデータパイプラインを実現・高速化します。DDN の共有パラレルアーキテクチャにより、複数の HGX システム間で DL ワークフローのすべてのフェーズを並行かつ継続的に実行することが可能です。これにより、データをストレージ間で移動させる際の管理負荷やリスクを排除します。

アプリケーションレベルでは、データは標準的で高い互換性を持つファイルインターフェースを通じてアクセスされるため、ユーザーにとって直感的かつ馴染みのある操作感を提供します。さらに、HGX システム全体でアプリケーションを同時に実行し、複数のニューラルネットワーク候補の学習を並列に行うことで、大幅な高速化が可能になります。

このような高度な最適化により、DL フレームワークの能力を最大限に引き出すことができます。DDN は NVIDIA やその顧客と密に連携し、広く使用されているディープラーニングフレームワークを HGX システム上で確実に動作させるためのソリューションと技術を開発しています。

1.3. DDN A³I マルチレール・ネットワーキングと LACP Bonding

DDN A³I ソリューションは、AI インフラにおけるスムーズな導入と最適なパフォーマンスを実現するために、幅広いネットワーク技術とトポロジに対応しています。最新世代の Spectrum-X イーサネット技術は、アプリケーション、コンピュータサーバ、ストレージアプライアンス間での高帯域・低レイテンシなデータ転送を可能にします。

DDN A³I Multirail（マルチレール）は、HGX システム上の複数のネットワークインターフェースをグルーピングし、合計のデータ転送能力を高速化します。この機能は、すべてのインターフェースにわたるトラフィックを動的にバランスし、リンクの状態を常時監視することで、障害を迅速に検出し自動的に復旧させます。このマルチレール機能により、高性能なネットワーク設計・導入・管理が非常に容易になり、NCP（NVIDIA Cloud Partner）向け大規模インフラに対して完全な接続性を実現することが実証されています。さらに、DDN A³I ソリューションは LACP（Link Aggregation Control Protocol）Bonding にも対応しています。LACP Bonding を利用することで、複数の物理ネットワークインターフェースをひとつの論理的な高帯域・冗長インターフェースとして統合することができます。これにより、スループットと可用性が向上し、NCP 向けの共有ストレージへのアクセスがより高速かつ信頼性の高いものになります。

DDN の LACP サポートにより、標準的かつ広くサポートされているネットワーク構成オプションを活用しながら、NCP 導入において柔軟な対応が可能となります。

1.4. HGX システム向け DDN A³I の高度な最適化

DDN A³I クライアントは NUMA (Non-Uniform Memory Access) を意識した機能を備えており、HGX システムにおける最適な動作を実現します。

この機能により、スレッドが自動的にピン留めされ、HGX 全体での I/O アクティビティが適切にローカライズされるため、レイテンシが低減し、システム全体のリソース利用効率が向上します。さらに、ページキャッシュからメモリページを回収する際のオーバーヘッドを軽減する最適化も組み込まれており、バッファード I/O 操作をストレージへ高速に処理することが可能になります。この DDN A³I クライアントソフトウェアは、HGX 環境での大規模スケール運用において、すでに検証済みです。

1.5. DDN A³I Hot Nodes

DDN Hot Nodes は、HGX システム内の NVMe デバイスを読み取り専用のローカルキャッシュとして利用可能にする、強力なソフトウェア拡張機能です。この手法は、あるワークフローの中で同じデータセットが複数回アクセスされる場合に、アプリケーションのパフォーマンスを大幅に向上させます。

このようなアクセスパターンは、ディープラーニングのトレーニングにおいてよく見られます。特定の入力データセット、あるいはその一部が、複数の学習イテレーション中に何度も繰り返し読み込まれるのが一般的です。従来、このような入力データは HGX システム上のアプリケーションから直接共有ストレージにアクセスして読み込まれていました。そのため、共有ストレージのリソースが継続的に消費されていました。Hot Nodes 機能を使用すると、トレーニングの最初のイテレーション中に読み込まれた入力データは、DDN クライアントソフトウェアによって自動的にローカル NVMe デバイスへコピーされます。以降の読み込みは共有ストレージではなく、このローカルキャッシュからアプリケーションに対して高速に提供されるため、ネットワークトラフィックが発生せず、共有ストレージへの負荷も軽減されます。これにより、チェックポイント処理など、他の重要なディープラーニング処理が共有ストレージの性能を最大限活用してより迅速に完了できるようになります。また、Hot Nodes 機能には、ユーザー主導のローカルキャッシュ管理や、スケジューラ統合を簡単に行うための高度なデータ管理ツールや性能モニタリング機能が含まれています。たとえば、AI トレーニングアプリケーションを実行する前に、トレーニング用入力データをローカルキャッシュに事前ロードする「プリフライトタスク」として設定することも可能です。

加えて、キャッシュの利用状況や性能に関するメトリクスが可視化されるため、システム管理者はデータロードの最適化やアプリケーション/インフラ全体の効率向上に向けてさらなるチューニングを行えます。

1.6. DDN A³I コンテナクライアント

コンテナは、アプリケーションとその依存関係をひとまとめにすることで、シンプルで信頼性が高く、一貫性のある実行環境を提供します。DDN は、HGX システム上のアプリケーションコンテナと DDN の並列ファイルシステムとの間に、直接的かつ高性能な接続を実現します。これにより、コンテナ内からのデータアクセスが低レイテンシかつ高スループットの並列 I/O として実行され、アプリケーションのパフォーマンスが大きく向上します。また、複数のコンテナがホスト単位でストレージにアクセスする際の制約も解消され、より柔軟な構成が可能となります。

この DDN A³I のインコンテナマウント機能は、ユニバーサルラッパー（共通ラッパースクリプト）によってランタイム中に追加されるため、アプリケーションやコンテナイメージに変更を加える必要はありません。

NVIDIA からは、HGX システム向けに最適化されたディープラーニング（DL）フレームワークのコンテナバージョンが提供されており、データサイエンティストはこれを活用して、迅速にアプリケーションを開発・展開できます。場合によっては、これらのコンテナのオープンソース版も提供されており、開発者がより自由にアクセス・統合することも可能です。

DDN A³I コンテナクライアントは、HGX システム上で動作するコンテナアプリケーションから、DDN の並列ファイルシステムに対して高性能な並列データアクセスを直接提供します。これにより、コンテナ化された DL フレームワークは、計算スタックの他のレイヤーによるレイテンシを排除しながら、最も効率的なデータセットアクセスを実現できます。

1.7. DDN A³I S3 Data Service

DDN S3 Data Service は、共有ネームスペースに対してファイルアクセスとオブジェクトアクセスの両方を提供するハイブリッド型のデータアクセス機能です。このマルチプロトコルによる統合ネームスペースは、柔軟なワークフロー設計を可能にし、エンドツーエンドでの統合をシンプルに実現します。たとえば、S3 インターフェースを通じてデータをストレージに直接書き込み、その後すぐに HGX システム上のコンテナアプリケーションからファイルインターフェースで即座にアクセスすることができます。さらに、この共有ネームスペース自体を S3 インターフェースで公開することもできるため、マルチサイトやマルチクラウド環境での共同作業にも適しています。

DDN の S3 データサービスは、以下の特長を備えた設計となっています：

- 高性能
- 優れたスケーラビリティ
- 高いセキュリティ
- 高い信頼性

これにより、S3 ベースのアプリケーションやサービスと統合された AI ワークフローを柔軟かつシームレスに構築することができます。

1.8. DDN A³I マルチテナンシー

DDN A³I は、ネイティブクライアントおよび包括的なデジタルセキュリティフレームワークを通じて、大規模環境における安全なマルチテナント運用を非常にシンプルにします。DDN A³I のマルチテナンシー機能により、多くのユーザー間で HGX システムを共有しながらも、安全なデータ分離を維持することが容易になります。マルチテナンシーは、ユーザーに対して HGX システムリソースを迅速かつシームレスに動的に割り当てることを可能にします。また、リソースのサイロ化、複雑なソフトウェアリリース管理、ストレージ間の不要なデータ移動を排除します。DDN A³I は、HGX システムに対して非常に強力なマルチテナンシー機能を提供し、データ集約型アプリケーション向けに、安全で共有可能なイノベーション環境をシンプルに構築できるようにします。

一方で、コンテナはセキュリティ上の課題を抱えており、不正な特権昇格やデータアクセスに対して脆弱です。DDN A³I のデジタルセキュリティフレームワークは、包括的な制御機能を提供します。たとえば、global root_squash により、悪意あるユーザー、あるいはノードやコンテナが侵害された場合でも、不正なデータアクセスや改ざんを防ぎます。DDN A³I は、Shared-Secret-Key (SSK) 認証や Kerberos 認証による安全なクライアント認証にも対応しており、不正なクライアントのなりすましを防止します。

また、Project quotas 機能により、テナントごとに使用可能なストレージ容量や inode 数を制限することができます。必要に応じて、DDN A³I は fscrypt インターフェースを利用したクライアントサイド暗号化も提供します。これにより、複数の暗号鍵や、ファイル単位・ディレクトリ単位での暗号化に対応可能となり、強力なデータプライバシー保護が実現します。

2. HGX に最適化された DDN A³I ソリューション

DDN A³I のスケーラブルなアーキテクチャは、HGX システムと DDN の共有パラレルファイルストレージアプライアンスを統合し、GPU 上で実行される AI、アナリティクス、HPC ワークフローをエンドツーエンドで高速化・最適化します。

DDN A³I ソリューションは、HGX の導入を簡素化すると同時に、GPU の利用効率を最大限に高め、高スケーラビリティを実現します。

この章では、NVIDIA Cloud Partner (NCP) 向けに提供される DDN A³I ソリューションの構成要素について説明します。

2.1. DDN AI400X3 アプライアンス

AI400X3 アプライアンス（図 1 参照）は、予測可能な容量、機能、性能を備えた完全に統合・最適化された共有データプラットフォームです。各 AI400X3 アプライアンスは、HGX システムに対して 140 GB/s の読み取り、110 GB/s の書き込み、および 150 万 IOPS を直接提供します。共有ストレージシステムに AI400X3 アプライアンスを追加することで、性能は線形にスケールします。オール NVMe 構成により、多様なワークロードやデータタイプに対して最適なパフォーマンスを提供し、大規模なアクセラレーテッド・コンピューティング環境においても、単一の共有・集中型データプラットフォームとして最大限の性能を発揮します。



図 1.DDN AI400X3 all-NVME ストレージアプライアンス

AI400X3 アプライアンスは、セクション 1 で説明された DDN A³I の共有パラレルアーキテクチャを統合しており、自動化されたデータ管理、デジタルセキュリティ、データ保護、広範な監視機能など、さまざまな機能を備えています。小規模構成では、データおよびメタデータを統合したアプライアンスが使われ、120TB、250TB、500TB の容量から選択可能です。

大規模構成では、次のように役割分担されたアプライアンスを使用します：

- AI400X3-OSS（データ用）：120TB～500TB の構成
- AI400X3-MDS（メタデータ用）：1 台あたり最大 数十 億 inode を提供

すべてのアプライアンスは、NCP のストレージネットワークに接続され、連携してファイルシステムを構成します。

2.2. DDN Insight ソフトウェア

DDN Insight（図 2 参照）は、AI400X3 アプライアンスのための集中管理および監視ソフトウェアスイートです。クラウドインフラストラクチャ内のすべての DDN ストレージシステムのパフォーマンスおよび健全性の監視を、1 つの Web ベースのユーザーインターフェースから広範囲に実行します。DDN Insight は、IT 運用を大幅に簡素化し、分析とインテリジェントなソフトウェアによってガイドされる、自動化かつプロアクティブなストレージプラットフォーム管理を可能にします。

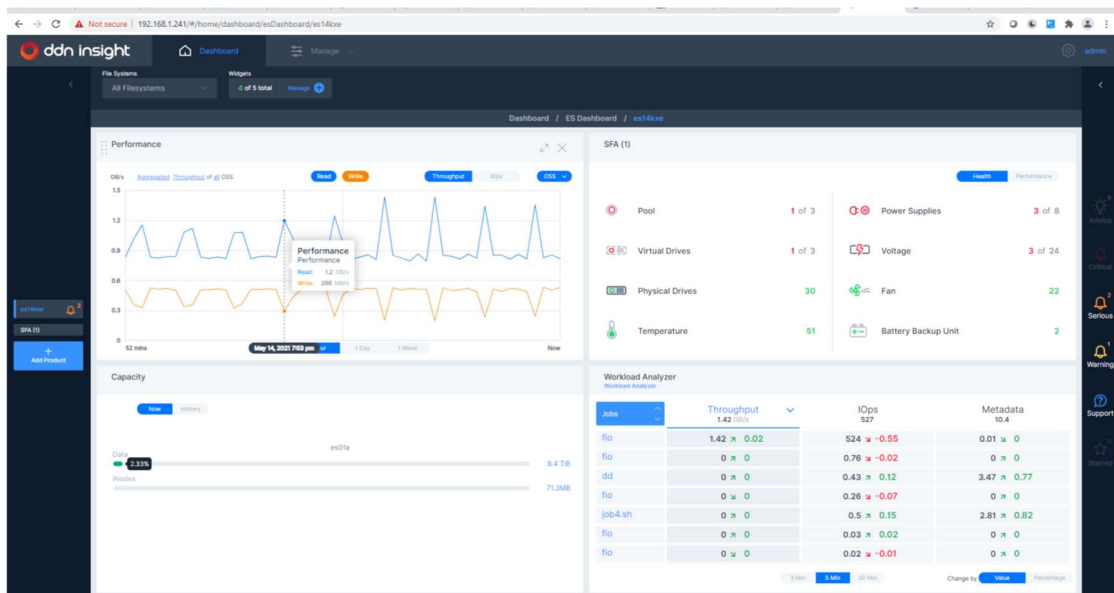


図 2.DDN Insight ワークロードアナライズツール

パフォーマンス監視は、クラウドインフラの効率的な運用において重要な要素です。データ I/O 性能に影響を与える複数の変数が存在する中で、本番ワークロードが実行されている最中に、ボトルネックや性能低下を特定することは極めて重要です。DDN Insight は、HGX システム上で動作しているアプリケーションから、AI400X3 アプライアンス内の個々のドライブに至るまで、クラウドインフラ全体の I/O トランザクションをリアルタイムで詳細に分析します。組み込みの分析エンジンにより、クラウドオペレーターは、直感的なユーザーインターフェースを介して、インフラ全体における I/O パフォーマンスを視覚的に把握することができます。このインターフェースには、詳細なログ収集、トレンドの把握、比較ツールなどが含まれており、特定のアプリケーションやユーザーの I/O パフォーマンスを長期間にわたって分析できます。オープンなバックエンドデータベースにより、DDN Insight の利点を拡張することが可能であり、他の AI インフラコンポーネントとの統合や、サードパーティの監視システムへのデータエクスポートも容易に行えます。

DDN Insight は、顧客が用意した管理サーバにインストール可能なソフトウェアパッケージとして、または DDN によって提供されるターンキー型のサーバアプライアンスとして利用できます。

3. NCP 向け DDN A³I リファレンスアーキテクチャ

DDN AI400X3 アプライアンスは、ターンキー型の共有ストレージシステムです。AI400X3 アプライアンスは、密度が高く電力効率の高い 2RU シャーシで、あらゆるワークロードやデータタイプに対して、最適な GPU パフォーマンスを提供します。AI400X3 アプライアンスは、クラウド環境の設計・導入・管理を簡素化し、予測可能な性能、容量、スケーラビリティを実現します。このアプライアンスは、HGX システムとのシームレスな統合を前提として設計されており、顧客がテスト環境から本番環境へ迅速に移行することを可能にします。さらに、DDN は、グローバルに展開する設計・導入・サポート体制を通じて、完全な専門支援を提供します。DDN のフィールドエンジニアリングチームは、A³I リファレンスアーキテクチャに基づいたソリューションを、すでに多数の顧客向けに導入しています。DDN は、NCP（NVIDIA Cloud Partner）環境における導入に適した以下の推奨アーキテクチャを提示しています。

3.1. ストレージパフォーマンスガイドライン

NCP リファレンスアーキテクチャのガイドライン

NVIDIA は、NCP（NVIDIA Cloud Partner）環境における最適な効率とスケーラビリティを確保するため、ベースとなるリファレンスアーキテクチャの一環としてパフォーマンスガイドラインを定めています。

これらの推奨事項は、さまざまなスケールにおける導入時における、標準的なストレージ性能の目安を定義しています。

GPU Nodes			パフォーマンスガイドライン			
			Read		Write	
GPUs	Nodes	SUs	ストレージ	Per GPU	ストレージ	Per GPU
1024	128	4	160 GB/s	160 MB/s	80 GB/s	80 MB/s
2048	256	8	320 GB/s	160 MB/s	160 GB/s	80 MB/s
4096	512	16	640 GB/s	160 MB/s	320 GB/s	80 MB/s
8192	1024	32	1280 GB/s	160 MB/s	640 GB/s	80 MB/s
16384	2048	64	2560 GB/s	160 MB/s	1280 GB/s	80 MB/s

表 1. NVIDIA による標準 HPS(High Performance Storage) 性能ガイドライン

3.1.1. 標準性能レベルに対応する DDN ストレージ構成

標準的な HPS（高性能ストレージ）性能レベルを満たすためには、共有ストレージシステムがすべての GPU に対して同時に、読み取りで 160MB/s、書き込みで 80MB/s 以上のスループットを提供する必要があります。以下の表 2 は、標準性能レベルを達成するために推奨される DDN ストレージ構成を示しています。ここでは、データとメタデータを統合した AI400X3 アプライアンスを用いた構成となっています。

		SUs（スケーラブルユニット）				
		4	8	16	32	64
GPU Nodes	HGX システム数	128	256	512	1024	2048
	GPU 数	1024	2048	4096	8192	16384
DDN ストレージ	AI400X3 統合 アプライアンス数	2	3	5	10	19
	AI400X3 メタデータ 専用アプライアンス数	-	-	-	-	-
	AI400X3 データ専用 アプライアンス数	-	-	-	-	-
DDN ストレージ 仕様	総読み込み スループット	280 GB/s	420 GB/s	700 GB/s	1.4 TB/s	2.7 TB/s
	総書き込み スループット	220 GB/s	330 GB/s	550 GB/s	1.1 TB/s	2.1 TB/s
	GPU あたりの読み込み スループット	273 MB/s	205 MB/s	171 MB/s	171 MB/s	162 MB/s
	GPU あたりの書き込み スループット	215 MB/s	161 MB/s	131 MB/s	134 MB/s	128 MB/s
	ネームスペース数	1	1	1	1	1
	最小有効容量	240 TB	360 TB	600 TB	1.2 PB	2.3 PB
	有効 inode 数	2 B	3 B	6 B	11 B	21 B
	400GbE OSFP ポート数	8	12	20	40	76
	スイッチポートと スプリッター ケーブル	4	6	10	20	38
	1GbE ポート数	8	12	20	40	76
	ラックユニット	4	6	10	20	38
	消費電力（定格）	5 kW	7.5 kW	12.5 kW	25 kW	47.5 kW
	冷却(定格)	16 kBTU/hr	24 kBTU/hr	40 kBTU/hr	80 kBTU/hr	152 kBTU/hr

表 2：AI400X3 アプライアンスを使用した標準 HPS 性能レベルに対応する推奨ストレージ構成

3.2. 推奨ストレージネットワーキング

NCP（NVIDIA Cloud Partner）リファレンス設計には、複数のネットワークが含まれており、その中の1つがストレージトラフィック専用のネットワークです。このストレージネットワークは、AI400X3 アプライアンス、コンピュータード、マネージメントノード間の接続を提供します。このネットワークは、NCP 導入環境における 高スループット・低レイテンシ・高スケーラビリティの要件を満たすように設計されています。

SN5600 スイッチ（図 3 参照）は、NCP ストレージ接続用として推奨されています。このスイッチは、2RU フォームファクタで 64 OSFP ポートにわたって合計 128 の 400GbE ポートを提供します。セクション 3.2.3 で示すケーブルは、SN5600 スイッチとの接続で検証済みです。

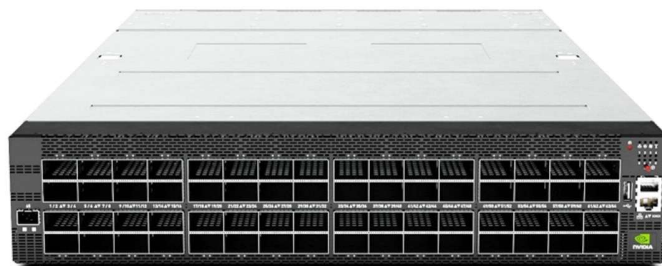


図 3. NVIDIA SN5600 スイッチ

ストレージネットワーク全体は RDMA over Converged Ethernet（RoCE）モードで動作する必要があります。DDN は、全体の設計において RoCE の使用を前提としています。

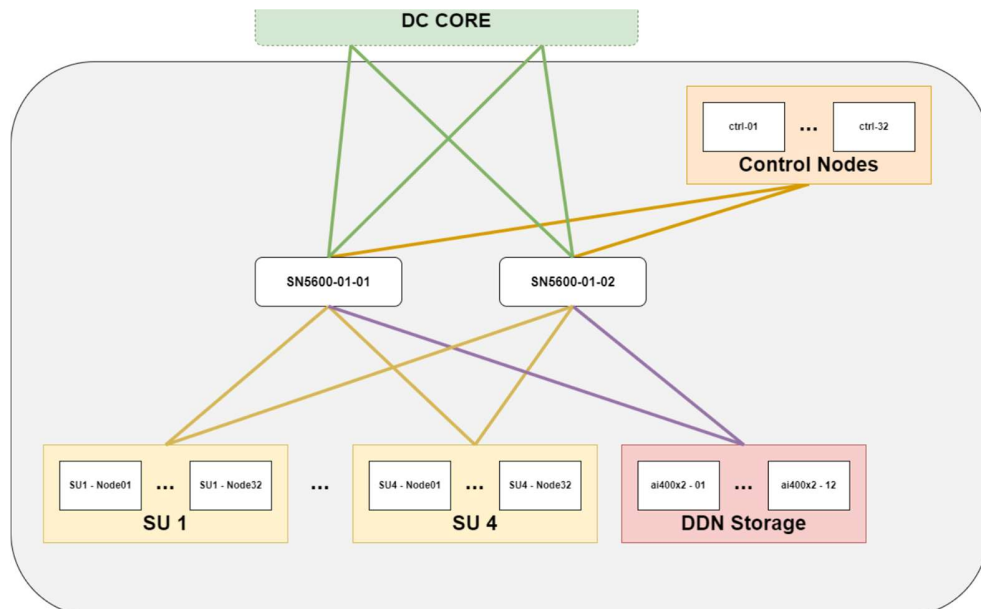


図 4. NCP ストレージネットワークアーキテクチャ概要

3.2.1. HGX システムのネットワーク接続

図 5 および図 6 は、HGX システムの推奨ポートおよびネットワーク接続構成を示しています。

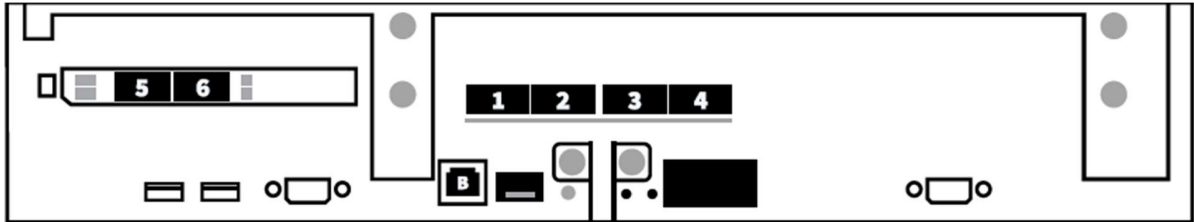


図 5. 推奨される HGX システムのネットワークポート接続

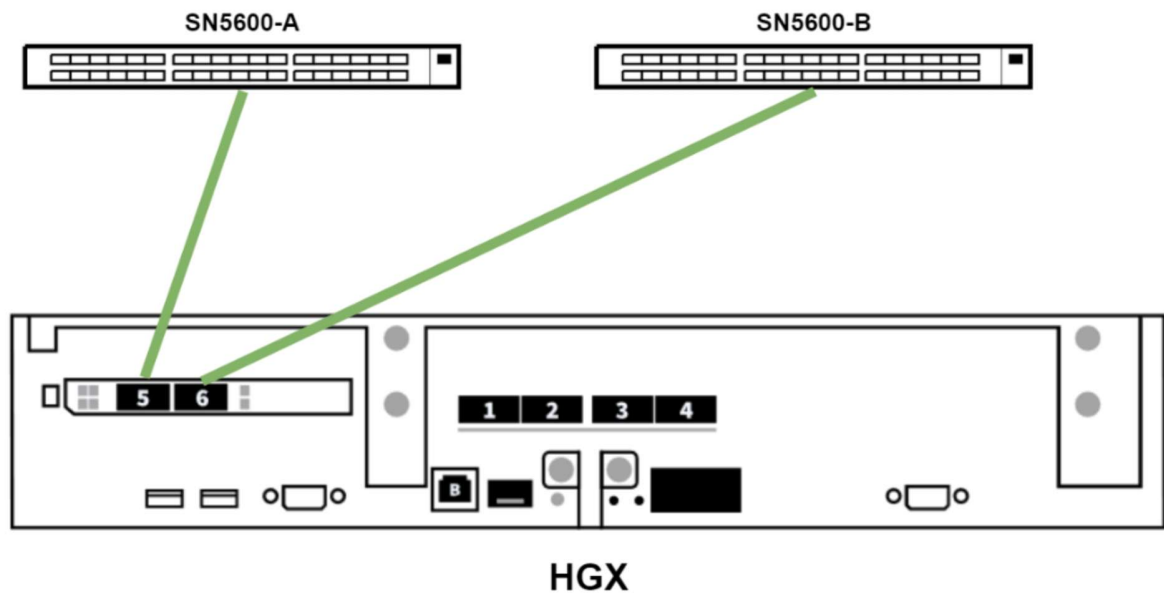


図 6. 推奨される HGX システムのネットワーク接続図

DDN では、HGX システム上の OSFP ポート 1～4 (ConnectX-7 HCA) をコンピュータネットワークに接続する必要があります。ポート 5 および 6 (NVIDIA BlueField®-3 DPU HCA) はインバンド管理ネットワークに接続する必要があり、このネットワークは RDMA Over Converged Ethernet (RoCE) を使用してストレージネットワークも兼ねます。また、DDN はリンクアグリゲーション (ボンディング) の使用を推奨しておらず、リンクの冗長性は DDN A3I Multi-rail によって処理されます。さらに、管理用の BMC (“B”) ポートはアウトオブバンド管理ネットワークに接続する必要があります。

3.2.2. SN5600 スイッチとの AI400X3 アプライアンスのケーブル接続

DDN は、AI400X3 アプライアンスをストレージネットワークに接続する際に、2 本のスプリッターケーブルを使用することを推奨しています（スプリッターケーブルの詳細については、セクション 3.2.3 を参照）。

1 本目のスプリッターケーブルは、アプライアンスのポート 1 および 3 に接続し、スイッチ“A”に接続する必要があります。2 本目のスプリッターケーブルは、アプライアンスのポート 2 および 4 に接続し、スイッチ“B”に接続します。これにより、リンクの冗長性（フェイルオーバー）が確保されます（図 7 および図 8 参照）。また、両コントローラの「M」（管理）ポートおよび「B」（BMC）ポートは、アウトオブバンドの管理ネットワークに接続する必要があります。

この構成により、AI400X3 の各ポートが物理的に分散された 2 台の SN5600 スイッチに接続され、ネットワークの高可用性と安定性が保証されます。

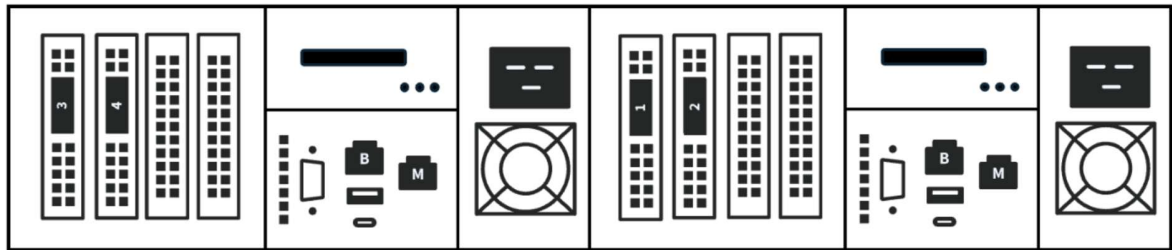


図 7. 推奨される AI400X3 アプライアンスネットワークポート接続

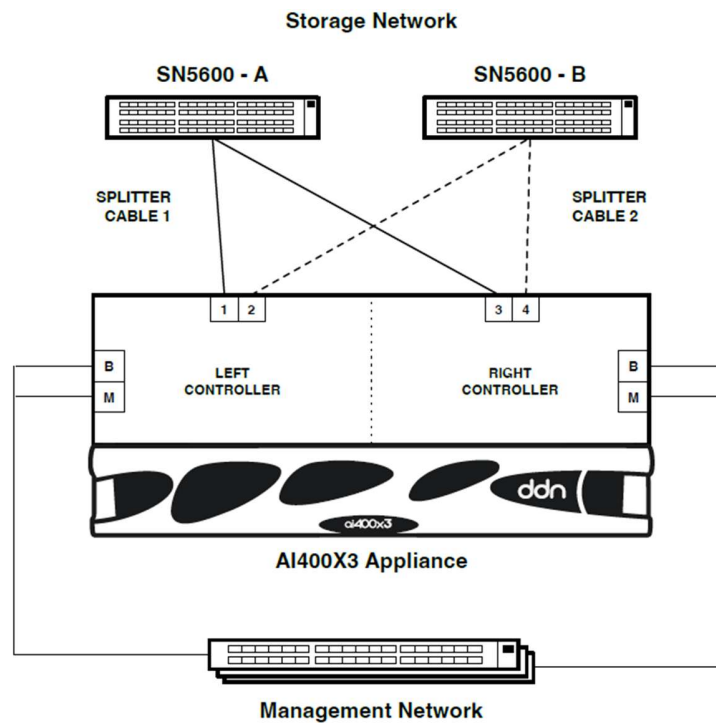


図 8. スプリッターケーブルを使用した、推奨される AI400X3 ネットワーク接続図

3.2.3. SN5600 スイッチとの AI400X3 アプライアンスのケーブル接続

AI400X3 アプライアンスは、4 つの 400GbE インターフェースを使用してストレージネットワークに接続されます。異なるイーサネット接続およびデータレート間の互換性を確保するためには、ケーブルの選定に特に注意する必要があります。DDN および NVIDIA は、AI400X3 アプライアンスと SN5600 スイッチとの接続において、以下のケーブルタイプを検証済みです。スプリッターケーブルを使用することで、スイッチポートの利用効率を最大化できます。各 AI400X3 アプライアンスあたり 2 本のスプリッターケーブルが必要です。

ケーブルタイプ	部品番号	説明
Direct Attach Copper	MCP7Y00-Nxxx	NVIDIA 製パッシブ銅スプリッターケーブル。InfiniBand ツインポート NDR 800Gb/s を 2 x 400Gb/s に分岐。OSFP から 2 x OSFP へ接続。xxx は長さ（メートル）を表し、001、01A、002、02A、003 などが存在。
Active Copper Cable	MCA7J60-Nxx	NVIDIA 製アクティブ銅スプリッターケーブル。InfiniBand ツインポート NDR 800Gb/s を 2 x 400Gb/s に分岐。OSFP から 2 x OSFP へ接続。xx は長さ（メートル）を表し、004、005 などが存在。

表 3. DAC および AOC スプリッターケーブル

3.3. HGX システム 128 台による NCP 導入例(1,024 GPUs)

図 9 は、128 台の HGX システムによる NCP（NVIDIA Cloud Partner）向け導入における DDN A3I アーキテクチャを示しています。この構成では、2 台の DDN AI400X3（データおよびメタデータ統合型）アプライアンスと 1 台の DDN Insight サーバが使用されます。各 HGX システムは、2 本の 400GbE リンクでストレージネットワークに接続されます。各 AI400X3 アプライアンスは、適切なケーブルタイプを使用して 4 本のイーサネットリンクでストレージネットワークに接続されます。DDN Insight サーバは、1GbE のアウトオブバンド管理ネットワーク経由で AI400X3 アプライアンスに接続され、統合型ストレージネットワークへの接続は不要です。

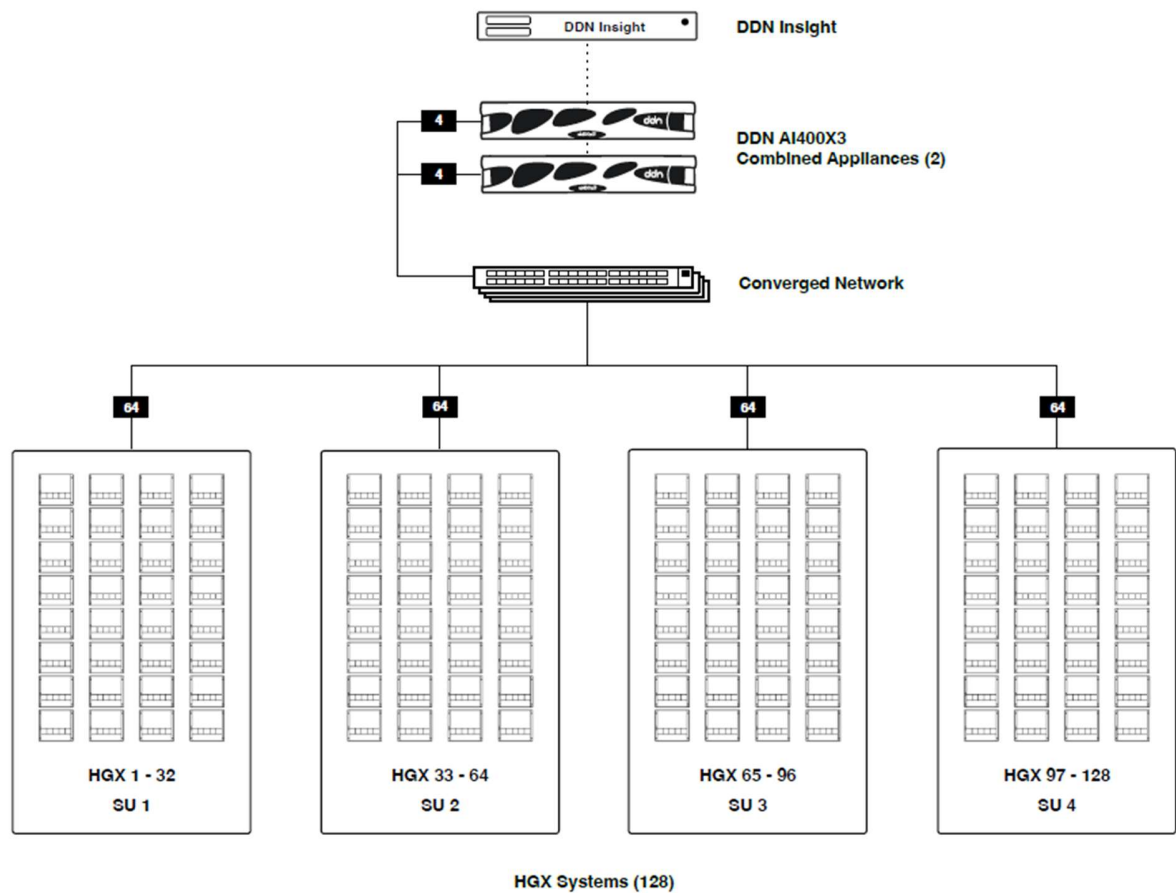


図 9. AI400X3 のデータおよびメタデータ統合アプライアンスを使用した、標準 HPS 性能レベル向けの 128 台の HGX システムによる DDN A3I リファレンスアーキテクチャ

3.4. HGX システム 256 台による NCP 導入例(2,048GPUs)

図 10 は、256 台の HGX システムによる NCP（NVIDIA Cloud Partner）向け導入における DDN A3I アーキテクチャを示しています。この構成では、3 台の DDN AI400X3（データおよびメタデータ統合型）アプライアンスと 1 台の DDN Insight サーバが使用されます。各 HGX システムは、2 本の 400GbE リンクでストレージネットワークに接続されます。各 AI400X3 アプライアンスは、適切なケーブルタイプを使用して 4 本のイーサネットリンクでストレージネットワークに接続されます。DDN Insight サーバは、1GbE のアウトオブバンド管理ネットワーク経由で AI400X3 アプライアンスに接続され、統合型ストレージネットワークへの接続は不要です。

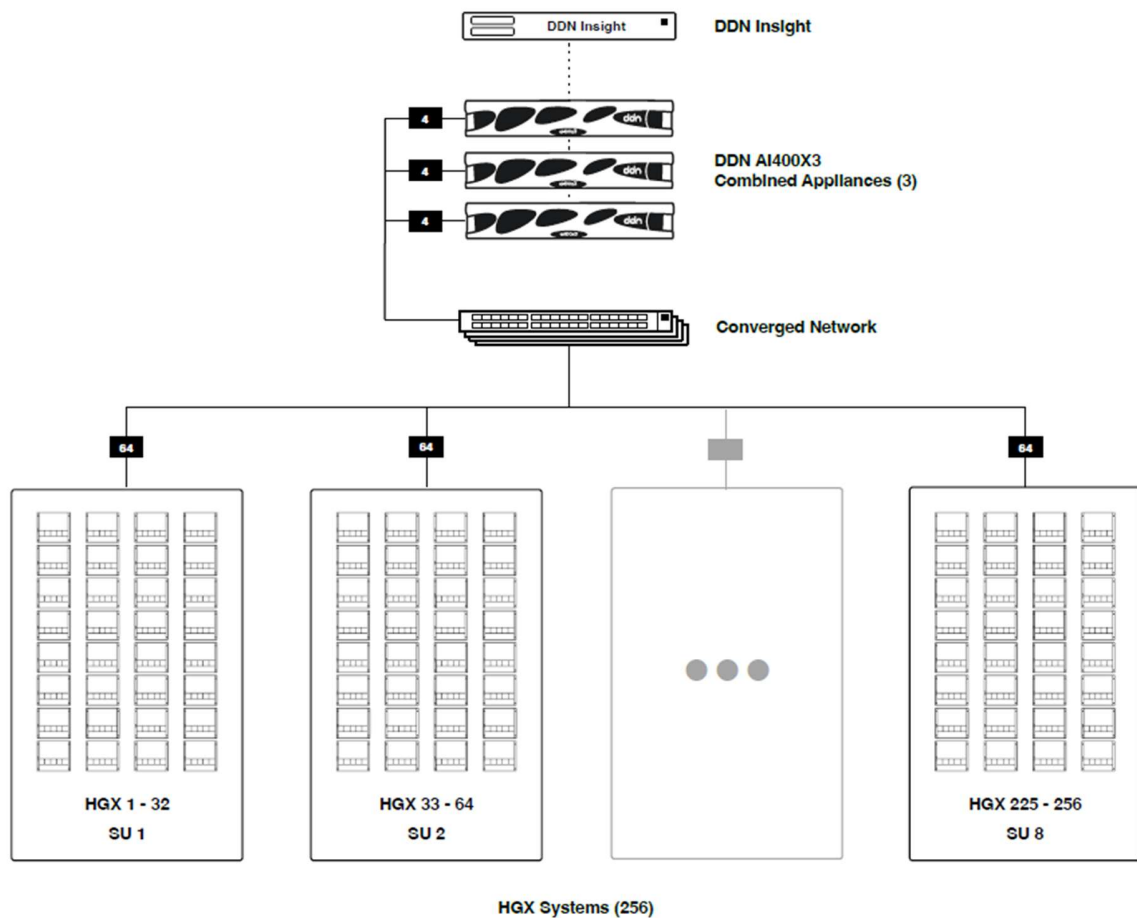


図 10. AI400X3 のデータおよびメタデータ統合アプライアンスを使用した、標準 HPS 性能レベル向けの 256 台の HGX システムによる DDN A3I リファレンスアーキテクチャ

3.5. HGX システム 512 台による NCP 導入例(4,096 GPUs)

図 11 は、512 台の HGX システムによる NCP（NVIDIA Cloud Partner）向け導入における DDN A3I アーキテクチャを示しています。この構成では、5 台の DDN AI400X3（データおよびメタデータ統合型）アプライアンスと 1 台の DDN Insight サーバが使用されます。各 HGX システムは、2 本の 400GbE リンクでストレージネットワークに接続されます。各 AI400X3 アプライアンスは、適切なケーブルタイプを使用して 4 本のイーサネットリンクでストレージネットワークに接続されます。DDN Insight サーバは、1GbE のアウトオブバンド管理ネットワーク経由で AI400X3 アプライアンスに接続され、統合型ストレージネットワークへの接続は不要です。

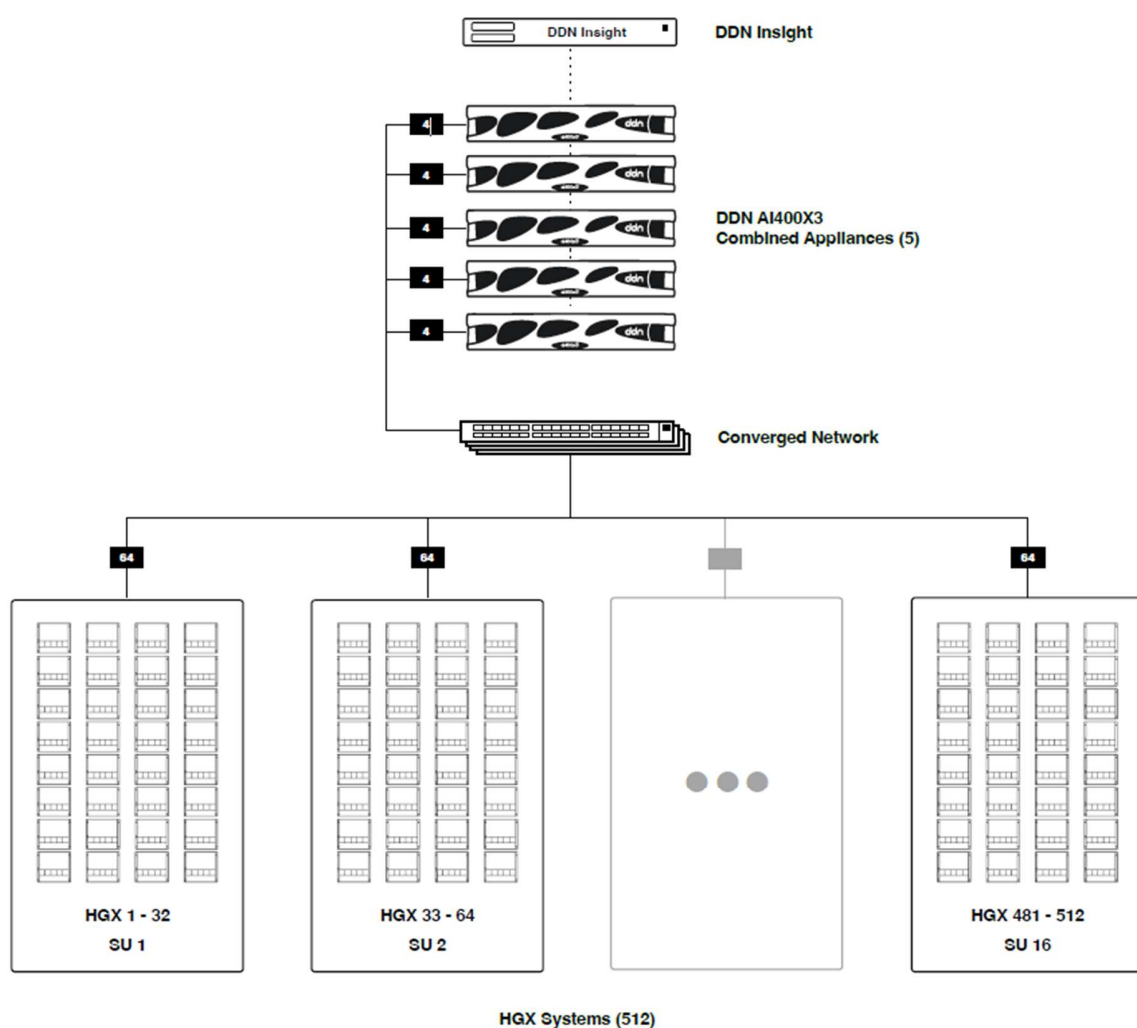


図 11. AI400X3 のデータおよびメタデータ統合アプライアンスを使用した、標準 HPS 性能レベル向けの 512 台の HGX システムによる DDN A3I リファレンスアーキテクチャ

3.6. HGX システム 1024 台による NCP 導入例(8,192 GPUs)

図 12 は、1024 台の HGX システムによる NCP（NVIDIA Cloud Partner）向け導入における DDN A3I アーキテクチャを示しています。この構成では、10 台の DDN AI400X3（データおよびメタデータ統合型）アプライアンスと 1 台の DDN Insight サーバが使用されます。各 HGX システムは、2 本の 400GbE リンクでストレージネットワークに接続されます。各 AI400X3 アプライアンスは、適切なケーブルタイプを使用して 4 本のイーサネットリンクでストレージネットワークに接続されます。DDN Insight サーバは、1GbE のアウトオブバンド管理ネットワーク経由で AI400X3 アプライアンスに接続され、統合型ストレージネットワークへの接続は不要です。

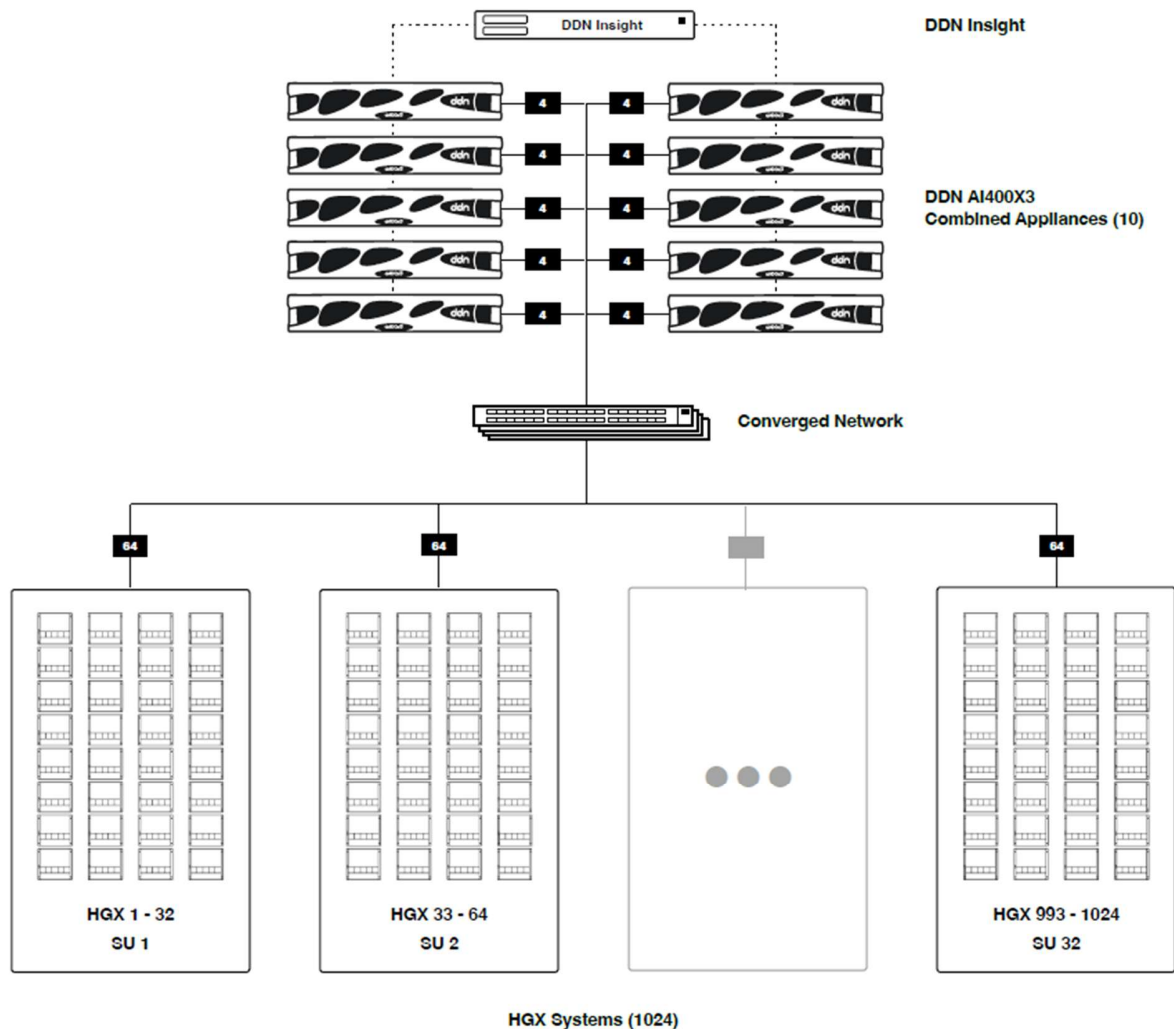


図 12. AI400X3 のデータおよびメタデータ統合アプライアンスを使用した、標準 HPS 性能レベル向けの 1024 台の HGX システムによる DDN A3I リファレンスアーキテクチャ

3.7. HGX システム 2048 台による NCP 導入例(16,384 GPUs)

図 13 は、2048 台の HGX システムによる NCP（NVIDIA Cloud Partner）向け導入における DDN A3I アーキテクチャを示しています。この構成では、19 台の DDN AI400X3（データおよびメタデータ統合型）アプライアンスと 1 台の DDN Insight サーバが使用されます。各 HGX システムは、2 本の 400GbE リンクでストレージネットワークに接続されます。各 AI400X3 アプライアンスは、適切なケーブルタイプを使用して 4 本のイーサネットリンクでストレージネットワークに接続されます。DDN Insight サーバは、1GbE のアウトオブバンド管理ネットワーク経由で AI400X3 アプライアンスに接続され、統合型ストレージネットワークへの接続は不要です。

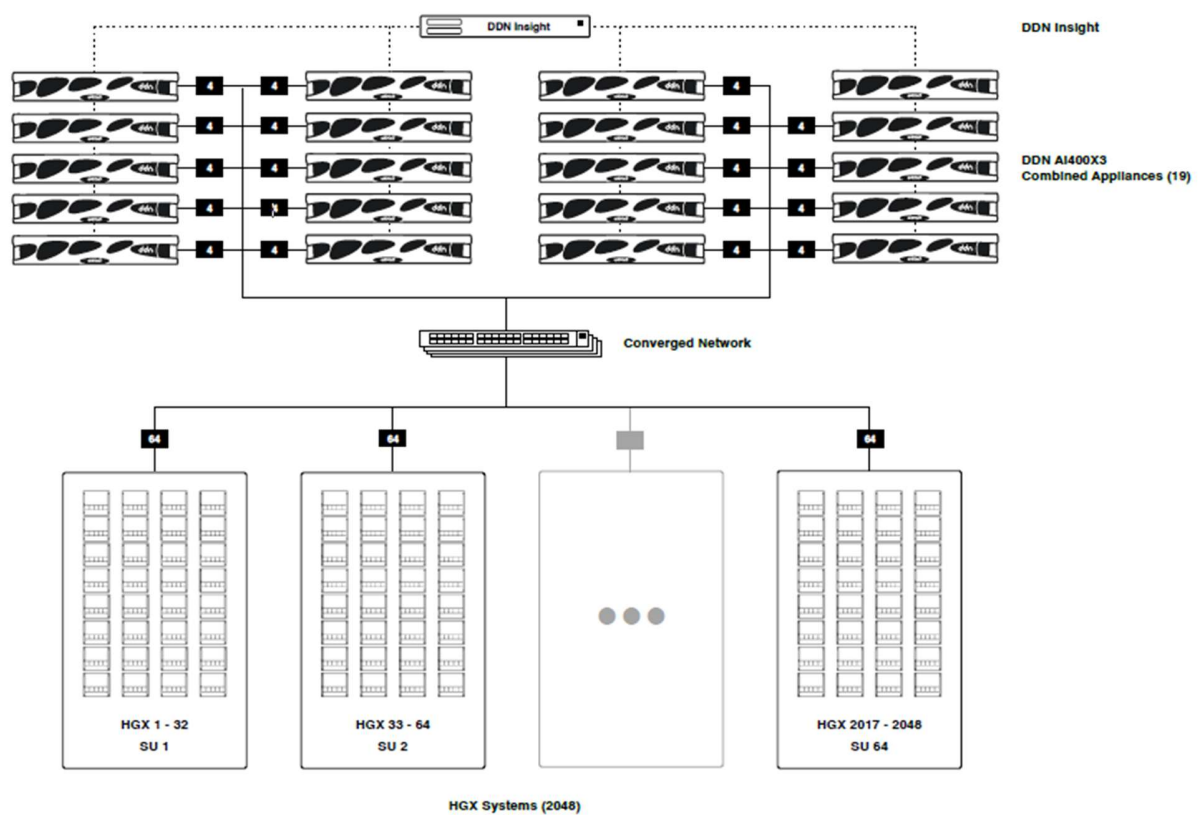


図 13. AI400X3 のデータおよびメタデータ統合アプライアンスを使用した、標準 HPS 性能レベル向けの 2048 台の HGX システムによる DDN A3I リファレンスアーキテクチャ

4. クラウド導入の真価を解き放つために、DDN にご連絡ください

DDN は、大規模なデータ駆動型プロジェクトに取り組む組織にとって、長年にわたり選ばれ続けてきたパートナーです。実績ある技術プラットフォームだけでなく、DDN は世界規模の研究開発組織とフィールド技術チームを通じて、豊富な技術的専門知識を提供しています。

世界中に在籍する 数百人のエンジニアや技術エキスパートが、お客様のプロジェクトのあらゆるフェーズにおける企画立案、ソリューション設計、システム導入、サポート、将来のスケーリングにおいて最適化を支援します。

DDN は、顧客重視の姿勢と卓越した技術力、そして豊富な現場経験を兼ね備え、あらゆる課題に対して最適なソリューションを提供します。コンサルティング的なアプローチにより、専門家が要件を綿密に評価し、アプリケーションレベルでデータワークフローの最適化を行います。その上で、お客様の取り組みを最大限に支援・加速するための、最適化され、信頼性が高く、使いやすいソリューションを設計・提案します。

また、DDN は大規模プロジェクトの豊富な導入実績をもとに、お客様環境を反映したテストプロトコルの設計・実行を行うための、体系的な検証プログラムを構築します。自社ラボには、最先端のアクセラレーテッドコンピューティングプラットフォームが完備されており、AI やディープラーニング（DL）用途に向けた独自のベンチマーク・テスト機能を提供できます。

今すぐ DDN にご連絡いただき、NCP 導入の真のパワーを引き出すために、当社の専門チームをご活用ください。

5. Appendix: HPS 1 GB/s 性能向け DDN ストレージ構成

統合ネットワークは、すべての GPU に対して同時に平均 1 GB/s の持続的なデータ転送レートをサポートできます。表 4 では、この性能要件を満たすための各スケーラブルユニット (SU) の推奨 DDN ストレージ構成を示しています。構成には、AI400X3 の統合型アプライアンス、またはデータ専用の AI400X3-OSS アプライアンスとメタデータ専用の AI400X3-MDS アプライアンスを使用します。統合型アプライアンスは、8 SU 以下の構成に推奨されます。一方、16 SU 以上の構成では、メタデータとデータを分離した専用アプライアンスの使用が推奨されます。

		SUs (スケーラブルユニット)				
		4	8	16	32	64
GPU Nodes	HGX システム数	128	256	512	1024	2048
	GPU 数	1024	2048	4096	8192	16384
DDN ストレージ	AI400X3 統合アプライアンス数	8	15	-	-	-
	AI400X3 メタデータ専用アプライアンス数	-	-	5	10	20
	AI400X3 データ専用アプライアンス数	-	-	30	59	118
DDN ストレージ仕様	総読込みスループット	1.1 TB/s	2.1 TB/s	4.2 TB/s	8.3 TB/s	16.5 TB/s
	総書込みスループット	880 GB/s	1.7 TB/s	3.3 TB/s	6.5TB/s	13 TB/s
	GPU あたりの読込みスループット	1.1 GB/s	1.0 GB/s	1.0 GB/s	1.0 GB/s	1.0 GB/s
	GPU あたりの書込みスループット	859 MB/s	806 MB/s	806 MB/s	792 MB/s	792 MB/s
	ネームスペース数	1	1	1	1	1
	最小有効容量	960 TB	1.8 PB	3.6 PB	7.1 PB	14.2 PB
	有効 inode 数	9 B	17 B	46 B	92 B	184 B
	400GbE OSFP ポート数	32	60	140	276	552
	スイッチポートとスプリッターケーブル	16	30	70	138	276
	1GbE ポート数	32	60	140	276	552
	ラックユニット	16	30	70	138	276
	消費電力 (定格)	20 kW	37.5 kW	83 kW	163 kW	326 kW
	冷却(定格)	64 kBTU/hr	120 kBTU/hr	281 kBTU/hr	554 kBTU/hr	1108 kBTU/hr

表 4 : GPU あたり 1 GB/s の読み取り性能を実現するための、AI400X3 アプライアンスの推奨ストレージ構成