

NECの生成AIを支える国内企業で最大規模のAIスーパーコンピュータ

北野 貴稔

要旨

近年、生成AIはパラメータ数の大規模化と学習データ量の増大により、劇的な進化を遂げています。また、言語に加え、画像・音声も含めた生成AIのマルチモーダル化により、桁違いの計算能力が必要となっています。そのため、生成AIの基盤モデルの開発には多数GPUを用いた大規模分散学習を効率的かつ安定的に行うシステムが必要不可欠です。本稿では、生成AIの学習を支えるAIスーパーコンピュータのシステムアーキテクチャと今後の取り組みについて紹介します。



生成AI/大規模言語モデル/AIスーパーコンピュータ/データセンターネットワーク/分散ファイルシステム/
Kubernetes/深層学習フレームワーク

1. はじめに

近年、生成AIはパラメータ数の大規模化と学習データ量の増大により、劇的な進化を遂げています。また、言語だけでなく、画像・音声を加えた生成AIのマルチモーダル化の進展により、桁違いの計算能力が必要となっています。GPU (Graphics Processing Unit) 1基では学習に350年以上もかかる大規模言語モデルも登場していることから、生成AIの技術開発には、多数GPUを用いた分散学習を効率的かつ安定的に行うためのシステムが必要不可欠になっています。

NECでは、AIの高度化・マルチモーダル化の進展に伴う計算能力の重要性の高まりを予期し、2023年3月に国内企業で最大規模となるNVIDIA A100 Tensorコア GPUを928基搭載したAIスーパーコンピュータ（以下、AIスパコン）の全面稼働を開始しています¹⁾。本稿では、大規模な生成AIの基盤モデルを高速かつ安定的に学習可能にするAIスパコンのシステムアーキテクチャについて、AIスパコンを構成する計算機基盤とソフトウェア基盤の2つの観点で説明します²⁾。第2章において、先進サーバのコンピュータアーキテクチャと先端のネットワークアーキテクチャから構成される「計算機基盤」を、第3

章においてKubernetesを中核とした実行・運用監視を行う「ソフトウェア基盤」を、第4章において今後の取り組みを説明します。

2. AIスパコンの計算機基盤のアーキテクチャ

2.1 コンピュータアーキテクチャ

AIスパコンは、NVIDIAのA100 80GB TensorコアGPUを928基・合計116台のGPUサーバで構成されます。このような多数のGPUをスケールアウトさせて高速な分散学習を行うことができるよう、AIスパコンの計算機基盤のコンピュータアーキテクチャ及びネットワークアーキテクチャを設計しています。次ではその詳細を説明します。

2.1.1 GPUサーバのスケールアウト方法

NECのAIスパコンでは、GPUサーバ16台と複数のLeafスイッチのセットの組み合わせをスケールアウトさせる構成単位としており、これをComputing Unit（以下、CU）と呼んでいます。すべてのCUは同一構成となっており、CUを並べることでGPUサーバをスケールアウトさせています（図1）。複数のCUは、Spineスイッチを介して

接続することにより、任意の規模でスケールアウトできるだけでなく、ハードウェア/ソフトウェアが統一されるため保守性が非常に高いです。

2.1.2 高速な分散学習を支えるハードウェアトポロジ

多数GPUを用いて高速な分散学習を行うためには、GPUサーバ間の通信のレイテンシを小さくするためのハードウェアトポロジが重要です。そこで、低レイテンシでの通信を実現するため、NICとGPUの物理的距離を最小化するようにGPUが属するPCIスイッチ配下にNICを1つずつ、合計4つのNICを配置しています。そうすることで、GPUDirect RDMAを用いてCPUをショートカットし、低レイテンシ・広帯域での通信が可能になっています(図2)。

また、ネットワークは、Management・Compute・Storage用の3つに分けて、各ネットワーク用のNICとしてManagement用に1つ、Compute用に4つ、Storage用に1つの合計6つのNICを配置し、通信トラフィックを分離しています。これにより、マネージメント・ストレージ向けの通信を分離することで、GPUサーバ間の通信のレイテンシが悪化しないようにしています。

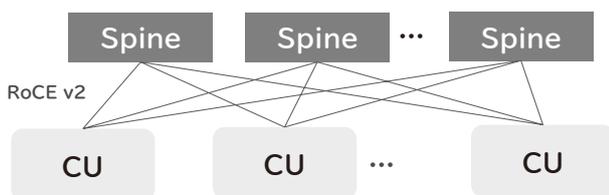


図1 CUによるGPUサーバ群のスケールアウト

このように、深層学習フレームワークによるパラメータ交換の通信のレイテンシを短縮できるようにNICの配置を行うことで、大規模な生成AIの基盤モデルでも高速な分散学習をすることが可能になっています。

2.2 ネットワークアーキテクチャ

ネットワークアーキテクチャとしては、GPUサーバと各種高速スイッチをSpine/Leaf型のアーキテクチャにてEnd-to-Endで統合することにより、低レイテンシ・高バンド幅での通信を実現しています(図3)。次ではその詳細を説明します。

2.2.1 分散深層学習向けの高速通信のアーキテクチャ

分散深層学習では、イテレーションごとにパラメータ群をGPUサーバ間で通信する必要があります。また、生成AIの基盤モデルはパラメータ数が非常に大きいことから、分散効率を向上させるためには、低レイテンシ・高バンド幅での通信が必要になります。

したがって、通信としてはRoCE v2を用いたGPUDirect RDMAをEnd-to-Endで行えることが重要です。そこで、200Gbpsの超高速イーサネットをサポートするNVIDIAの高速イーサネットスイッチNVIDIA Spectrum SN3700と、低レイテンシのインターコネクタNVIDIA ConnectX-6で全サーバを接続しています。また、ネットワーク構成としては、2層のSpine-Leafアーキテクチャを採用し、GPUサーバ間のHop数を最小化することで、低レイテンシでの通信を実現しています。

これらの工夫により、GPUサーバ間での低レイテンシ・

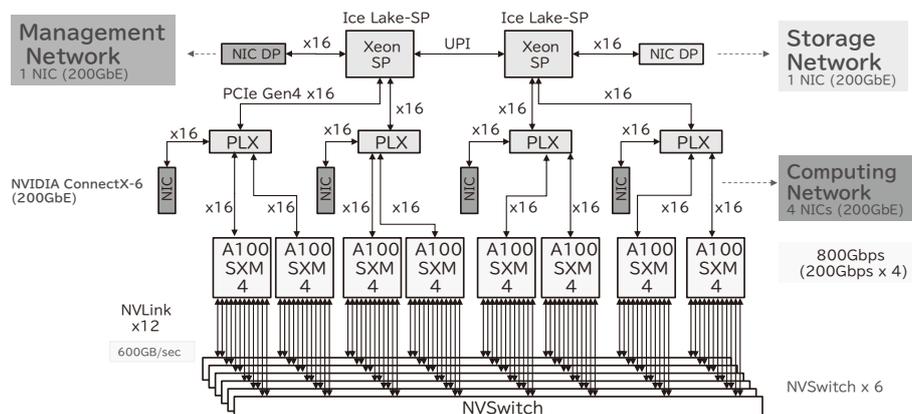


図2 高速かつ安定的な分散学習を実現するNIC配置

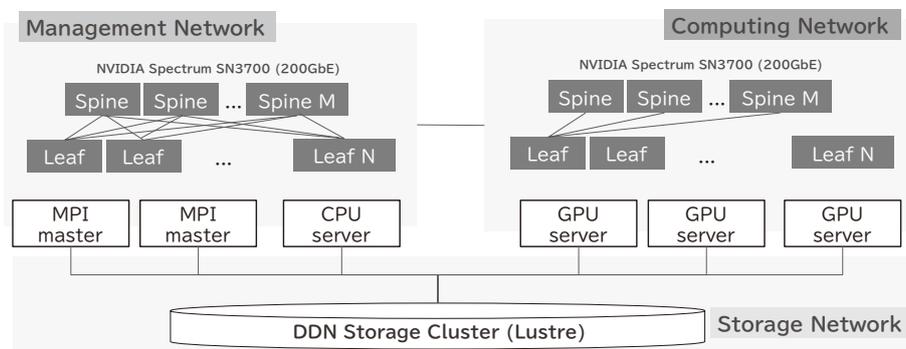


図3 Spine/Leafアーキテクチャの採用による高速通信

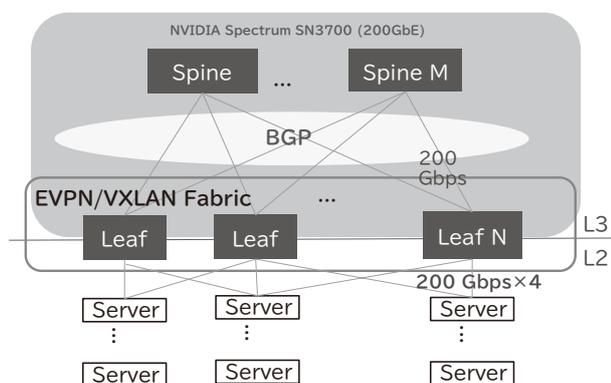


図4 データセンターネットワーク技術の活用

高バンド幅での通信が行えるようになってきているため、GPU数百基を用いた分散学習でも高速に行うことが可能です。

2.2.2 データセンターネットワーク技術の活用

NECのAIスパコンのネットワークでは、データセンターネットワークの技術を用いてAIスパコンを構成しています(図4)。

全体のネットワークのアーキテクチャとしてはIP Closを採用し、EVPNをコントロールプレーンとして使い、VXLANをデータプレーンとして使っています。そして、RoCE v2 over VXLANを使うことで、サーバ間で高速な通信が行えるようにしています。また、ルーティングプロトコルとしてはBGPを使い、BGP unnumberedをIP管理の容易化のために使っています。

このように、多数のデータセンターネットワークの技術群を活用することで、利便性の向上とネットワーク管理コストの削減をしつつ、大規模な分散学習を可能にする高性

能なネットワークを実現しています。

2.3 ストレージアーキテクチャ

マルチモーダルな生成AIを開発するためには、テキスト・画像・映像・音声などの多種多様なモーダルのデータセットが必要となり、高速・広帯域でのアクセスが必要になります。そのため、多数のGPUサーバからの同時アクセスが可能で、大規模な言語データや数千万枚の画像などのデータセットなどを活用できる、高IOPS・広帯域でのアクセスに対応可能な分散ファイルシステムが必要です。そこで、NECのAIスパコンでは、HPCの世界で高い実績のある分散ファイルシステムLustreを採用しています。次では、Lustreを用いた大規模なストレージシステムの構成と、大規模データの高速なアクセスを可能にするストレージのスケールアウト方法について説明します。

2.3.1 大規模データを支えるストレージシステムの構成

AIスパコンのストレージシステムは、ログなどを含めた高IOPS用途のNVMe SSDで構成される「高速領域」と、大規模な学習データセットを配置する用途のHDDで構成される「大容量領域」の2つで構成されています。I/Oの特性に応じてストレージを分けることで高速なデータアクセスを実現しています。また、ストレージシステムを構成するストレージサーバは200GbpsのNICを複数枚搭載しており、データアクセスにおいてもRoCE v2を利用可能にすることで、広帯域での通信が可能になっています。

2.3.2 ストレージのスケールアウト方法

大規模なファイルシステムでは、多量のファイルが存在

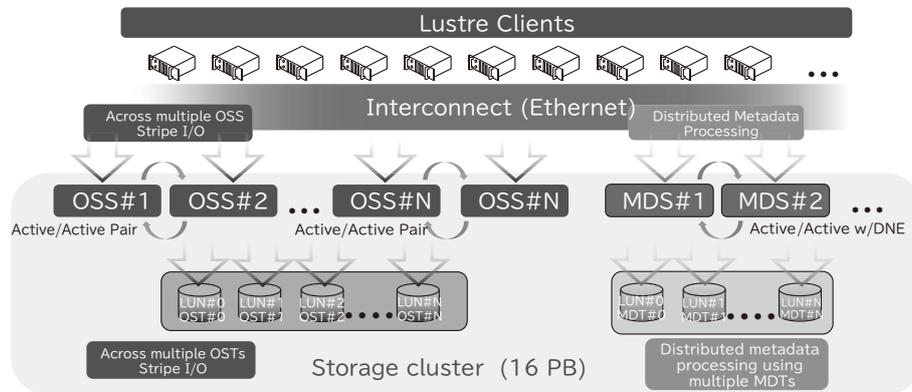


図5 高いI/O性能を支えるストレージアーキテクチャ

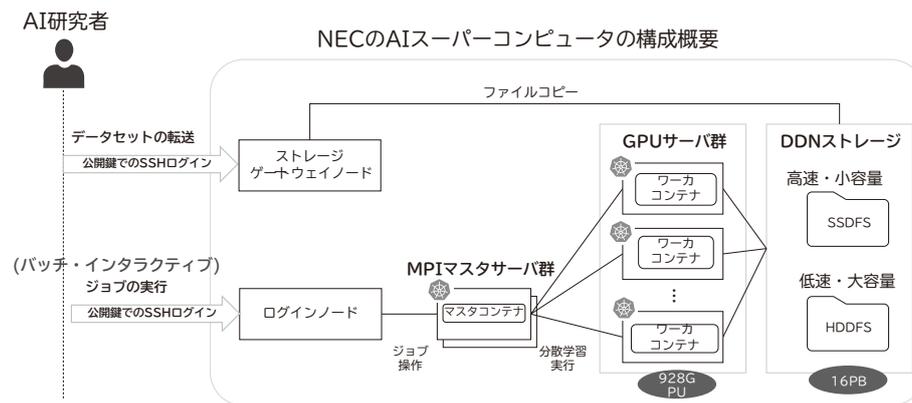


図6 AIスパコンのソフトウェアの実行基盤

することから、一般的にメタデータの処理性能がボトルネックになりやすいです。そこで、メタデータサーバ（以下、MDS）を複数並べてスケールアウトさせることで、メタデータ性能を高めています（図5）。また、ユーザーのデータアクセスが特定MDSに集中しないよう、ユーザーごとに異なるMDSが利用できるようにディレクトリを自動配置するプログラムを開発することで、メタデータへのI/O負荷を分散しています。これらの工夫により、大規模ファイルシステムでも、ボトルネックなく、高速なデータアクセスが可能になっています。

3. 先進AIを支えるソフトウェア基盤アーキテクチャ

3.1 Kubernetesによるソフトウェアの実行基盤

ソフトウェアの実行基盤は、研究者が複雑な深層学習

環境を構築することなく、先端のAIの学習を効率的に行えるようにするため、コンテナ基盤であるKubernetesを中核に据え、それを拡張して構築しています（図6）。具体的には、NECの仕組みに合わせて認証・認可の仕組みやセキュリティ機構などを拡張しています。また、Kubernetesのジョブスケジューラは、AIスパコンのネットワークポロジなどの物理構成に最適化することで、効率的な分散学習を行うことができるようになっています。

このように、Kubernetesをソフトウェア基盤の中核に据えることで、組織の要件に合わせた柔軟な拡張や効率的な分散学習を実現することが可能な、最新のAIの研究開発環境を実現しています。

3.2 先進のAI研究を支える深層学習コンテナ環境

AIの学習には多種多様な深層学習フレームワーク群が

用いられ、AIの技術の進展の速さから更新が頻繁に行われます。また、NECでは画像認識・映像解析・言語モデル・最適計画・制御など多様なAIの研究開発を行っています。そこで、NECの研究開発に合わせて、最新の深層学習環境をコンテナとしてパッケージングし、NECの研究者に提供しています。

これにより、研究者は多様かつ複雑で頻繁に更新される深層学習フレームワーク群の環境を構築することなく、最先端の深層学習環境を利用することが可能です。

3.3 大規模システムの安定運用を支える運用監視基盤

生成AIの開発では、多数GPUを同時に分散学習で利用することから、AIスパコンには高い可用性が必要になります。しかし、多数のGPUを高い利用率で用いることから、一般のシステムよりも障害発生率が高く、運用監視が重要になります。

そこで、NECではAIスパコンに固有の死活監視・ログ監視・リソース監視・パフォーマンス監視を統合的に行う運用監視基盤を開発し、高い可用性を実現しています。具体的には、一般的な監視項目に加え、AIスパコン固有の各GPUサーバ間をつなぐ光ファイバーのLayer1でのメトリクス、RoCE v2などの分散学習特有の通信のメトリクス、各GPUのハードウェア/ソフトウェアのエラーなど、さまざまなレイヤで各種の監視を行っています。それにより、各レイヤでの障害・故障を統合的に検知し、異常ノードをKubernetesから即座に離脱させることで、高い安定性を実現しています。

運用監視基盤は、PrometheusやGrafana、Fluent BitやElasticsearchなどのソフトウェアを用いて実現しています。これらのソフトウェア群を利用することにより、監視機能はすべてGo言語を用いて拡張することが可能です。実行基盤と同じGo言語で拡張できるようにすることで、エンジニアリングスケーラビリティを高め、複雑かつ高度なAIスパコン全体の運用監視機能を効率的に開発することが可能になっています。

4. むすび

今後、生成AIは言語・画像・音声を含めたマルチモーダルな生成AIへと発展し、人間の高度な意思決定を支える基盤技術へと進化していきます。言語の基盤モデルだ

けでも多量の計算機が必要となることから、マルチモーダルな生成AIの技術開発では膨大な計算機能力が必須となります。

そこで、NECでは競争力の源泉となる国内企業で最大規模のAIスパコンを活用し、NECの強みである言語・画像・音声AIの技術を生かしたマルチモーダル生成AIの基盤モデルの研究開発を加速していきます。これにより、お客様やパートナーとの共創により先進的な社会価値を産み出すAI研究のセンター・オブ・エクセレンスの実現を目指します。

参考文献

- 1) NEC、国内企業で最大規模となるAI研究用スーパーコンピュータの構築を開始
https://jpn.nec.com/press/202205/20220517_02.html
- 2) 社会価値の創造を加速するNECのAIスーパーコンピュータ
<https://jpn.nec.com/rd/aisupercomputer/index.html>

執筆者プロフィール

北野 貴稔

グローバルイノベーション戦略統括部
ディレクター

NEC 技報のご案内

NEC 技報の論文をご覧くださいありがとうございます。
ご興味がありましたら、関連する他の論文もご一読ください。

NEC技報WEBサイトはこちら

NEC技報（日本語）

NEC Technical Journal（英語）

Vol.75 No.2 ビジネスの常識を変える生成AI特集 ～社会実装に向けた取り組みと、それを支える生成AI技術～

ビジネスの常識を変える生成AI特集によせて
生成AI技術への取り組み ～基盤から応用、ルール作りまで～

◇ 特集論文

急速に広まる生成AIの市場適用

NECにおける生成AIの取り組みについて
電子カルテと医療文書の作成支援による医師業務効率化
映像×LLMによる報告書作成業務の自動化
映像分析と生成AIによるリアル世界の行動理解
サイバー脅威インテリジェンス生成自動化
生成AIの社内活用を進めるNEC Generative AI Service (NGS)
ソフトウェア・システム開発への生成AIの活用
LLMとMIで革新する素材開発プラットフォーム
LLMと画像分析を活用した被災状況の把握

生成AIの可能性を高める基盤技術

日本語性能に優れたNECの大規模言語モデル(LLM)
NECの生成AIを支える国内企業で最大規模のAIスーパーコンピュータ
より安全な大規模言語モデル(LLM)を目指して
データを秘匿したまま連携できる連合学習技術とLLMへの適用可能性
大規模言語モデル(LLM)によるFew-shotクラスタリングの可能性
オープンメイン常識推論のための知識拡張型プロンプト学習
AI連携とオーケストレーションのための基盤ビジョンLLM
クエリを考慮した新規手法により関連する企業データを減らしLLM APIの使用コストを最適化

AI技術が社会へ浸透するために

AI標準化・ルールメイクに関する動向とNECの取り組み
人権尊重に向けたNECのAIガバナンスの取り組み
RCModelを用いたAIリスクマネジメントのための人材育成事例

◇ NEC Information

2023年度C&C賞表彰式典開催



Vol.75 No.2
(2024年3月)

特集TOP