

課題

AI クラスターの展開は複雑で、時間とコストがかかります。 AI アプリケーションに投資する企業はリソースに限りがあるため、ベンダーロックインに陥ることなく、 AI データセンターを展開し、高額なGPU リソースを最適化するのが困難です。

ソリューション ジュニパーの AI データセン ターソリューション</u>により、 ハイパフォーマンスの AI 学 習、推論、ストレージクラス ターを非常に高速かつ柔軟に 展開できます。また、非常に シンプルであるため、限られ た IT リソースで運用できま

メリット

- Day0、Day1、Day2 以降の 運用の簡略化により、必要な リソース、時間、コストを削 減できます。
- オープンで柔軟性があるため、ベンダーロックインに陥ることなく、実証済みのテクノロジーと製品を使用して、GPU を問わないネットワーク(InfiniBand と同等のパフォーマンス)を設計できます。 - 信頼できる製品が揃った、エンドツーエンドかつ検証済みのターンキーソリューションにより、迅速に展開できます。

ソリューションブリーフ:オープンか つイーサネットベースの AI データセン ターネットワーク

高速性、効率性、経済性を備えた、エンドツーエンド、マルチベンダー、インテントベースの AI ネットワーキング

従来のデータセンターテクノロジーや設計では、インフラの AI ワークロードに求められる要件(パフォーマンス、容量、遅延)に対応できません。バックエンドのトレーニングとフロントエンドの推論のファブリック設計を特徴とする AI データセンターには、最先端のソリューション(コンピューティング、ストレージ、ネットワーキング)が必要です。このような状況を受けて、高性能の GPU(グラフィック処理ユニット)と、単一ベンダーが提供する独自仕様の InfiniBand ネットワーキングに対するニーズが急増しています。その結果、価格が上昇し、サプライチェーンにボトルネックが生じています。そこで、コストを抑えながら、イノベーションと柔軟性を促進することが企業にとって不可欠となっています。コストを大幅に抑えながら GPU のパフォーマンスを最適化するために、多くの企業が、AI データセンター向けのオープンネットワークとしてイーサネットに目を向けています。

課題

業界を問わず、AI は実験の対象から、業務成果を達成するためのツールへと移行しており、AI 設計者やIT 担当者の業務は突如として変化しています。AI の価値を最大限に引き出すためには、AI の展開に伴う複雑性を管理して、AI のコストや適切なリードタイムに対応しなければなりません。

選択肢の少ない GPU と独自仕様の InfiniBand ネットワーキングにより、サプライチェーンの遅れや在庫の不足が発生し、高価かつ稀少な AI インフラストラクチャのコストが上昇しています。たとえば、1 つの GPU が 30,000 ドル以上するため、GPU を使用する AI サーバーは 400,000 ドル以上に上る場合があります。小規模 GPU クラスターでも数百万ドル、大規模 GPU クラスターなら数億ドルにも上ります。

データセンターネットワークは AI を支えるエンジンであり、GPU を相互接続して、利用率を最大化する上で重要な役割を果たします。高速化とコスト削減のカギとなるのは、ジョブ(AI のトレーニング)の完了までにかかる時間(JCT)を短縮することです。

また、AI の導入を成功に導くためには、需要に対して市場が迅速に対応しなければなりません。単一ベンダーのソリューションへの依存を緩和するために、各ベンダーが競争を繰り広げるオープンな市場を育てることに業界全体が取り組んでおり、幅広い GPUと、世界で最も普及している L2 テクノロジーであるイーサネットを使用するようになっています。

ジュニパーネットワークスの AI データセンターネットワークソリューション

ジュニパーの AI データセンターソリューションにより、ハイパフォーマンスの AI 学習、推論、ストレージクラスターを非常に高速かつ柔軟に展開できます。また、非常にシンプルであるため、限られた IT リソースで運用できます。ジュニパーは、独自のインテントベースの運用、AI ネイティブの仮想ネットワークアシ

スタント、ジュニパー検証済み設計により、AI データセンターネットワークの設計、展開、トラブルシューティングを簡略化します。さらに、多様なシリコン、マルチベンダーのスイッチ管理、スタンダードベースのオープンなイーサネットファブリックへの取り組みにより、お客様に最高の柔軟性を提供するとともに、ベンダーロックインを回避できます。

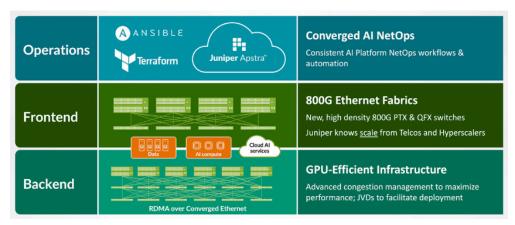


図 1: ジュニパーネットワークスの AI データセンターネットワークソリューション

運用の簡素化

Day0、Day1、Day2 以降の運用の簡略化により、必要なリソース、時間、コストを削減できます。ジュニパーのソフトウェアは、業界屈指のインテントベースネットワーク構築を可能にする唯一のマルチベンダーデータセンター自動化プラットフォームです。手動の設定により発生する問題を回避できるため、導入とトラブルシューティングにかかる時間を短縮できます。たとえば、Apstraにより、高度なロードバランシング、ECN(Explicit Congestion Notification:明示的な混雑通知)、PFC(Priority Flow Control:優先フロー制御)など、ハイパフォーマンスなロスレス AI ネットワーキングに不可欠な機能をファブリック全体に適用できます。CLI コマンドの入力は一切不要です。Apstraのリアルタイムネットワークフロー視覚化機能により、ヒートマップやインテントベースの解析を通じて、AI データセンターの混雑のホットスポットを特定し、迅速に問題を解決できます。

Apstra は、クローズドループアシュアランスを備えた信頼できる単一の情報源として、強力な解析機能によって問題を予測して事前対応的に解決し、サービスの中断を回避します。ブループリントがテンプレート化されているため、信頼性、一貫性、反復性が確保されており、運用コストを90%削減し、導入の時間を85%短縮できます。

ジュニパーは、Mist Al^{**}と Apstra の統合により、データセンターに AlOps をもたらします。これにより、さらなる自動化と分析情報を通じてデータセンターのエクスペリエンスを最適化できます。データセンター向けの仮想ネットワークアシスタント Marvis^{**}は、Apstra の機能豊富なストリーミングテレメトリとリアルタイムモニタリングを Marvis ダッシュボードに拡張して、 運用ドメイン全体で事前対応型のトラブルシューティングと解析を可能にします。この統合は、マルチベンダーデータセンターに対して AI ドリブンの実用的な分析情報を提供するための、基本となる最初の一歩です。また、Marvis を使用している、キャンパスや支社/拠点の運用担当者は、データセンターの問題を同じダッシュボードで確認できるようになりました。

オープンな柔軟性

オープンかつ柔軟なイーサーネットソリューションにより、実証 済みのテクノロジーと製品を使用して、ベンダーロックインを回 避できます。AIで最適化された、ジュニパーの新しいハイラディ ックス 800GbE リーフ/スパインデータセンターファブリック は、高度なトラフィック管理機能を備えており、イーサネットに よる高帯域幅、低遅延、ロスレス、拡張可能なパフォーマンスを 確保します。イーサネットにより、最大 1.6 テラビットまで拡張 し、マルチベンダーの GPU をサポートできるため、コストを削 減して、イノベーションを加速し、設計の柔軟性を大幅に高め て、サプライチェーンの課題を回避できます。RoCEv2(Remote Direct Memory Access over Converged Ethernet v2) をサポート するジュニパーのイーサネットファブリックは、データセンター (DC) ファブリックを管理および自動化するための唯一のマル チベンダーソリューションです。バックエンドの AI トレーニン グモデルとフロントエンドの AI 推論モデル向けに最適化されて います。

ターンキーソリューション

信頼できる製品が揃った、エンドツーエンドかつ検証済みのターンキーソリューションにより、データセンターに迅速に展開できます。エンドツーエンドのネットワークソリューションがあれば、ハイパフォーマンスな AI データセンターを柔軟かつ容易に構築できます。データセンター向けの Juniper Networks*PTX シリーズルーターは、新たな高密度 800 GbE PTX10002-36QDD 固定スイッチと、PTX10000 シャーシ向けの新たな 800 GbE ラインカードによって拡張されました。Express 5 カスタムシリコン上に構築された最大の PTX10000 シャーシ は、高基数スパインとスーパースパインアーキテクチャ向けに最大 576 x 800 GbE のポートをサポートしています。さらに、最新の Broadcom Tomahawk 5 ASIC を基盤とする、新たな高密度 800 GbE QFX5240 固定プラットフォームは、ハイパフォーマンス、拡張可能、高電力効率の AI データセンター向けに各種シリコンを提供します。

ジュニパーの QFX および PTX シリーズプラットフォームの卓越 したパフォーマンス、拡張性、混雑制御により、AI クラスターの ピーク運用にも確実に対応できます。ジュニパーの AI 最適化イ ーサネットは、MLPerf ベンチマークテストにおいて、ジョブ (ワークロードのトレーニング) の完了までにかかる時間 (JCT) が InfiniBand と同等でした。このイーサネットにより、 ベンダーロックインを回避しながら、オープンかつハイパフォー マンスなファブリックを構築できます。AIの導入を簡略化するた め、ジュニパーは新たなJVD(ジュニパー検証済み設計)を採用 しています。AIの JVD は、規定のレール最適化済みまたはマルチ レイヤー化された Clos ファブリックを実現し、スイッチング、 運用、セキュリティを含むデータセンターソリューション一式を 提供します。ジュニパーのラボの専門家によって厳格にテスト、 検証、文書化されています。既成の設定として、またはリスクを 低減し AI クラスターごとに適切なサイズと予算を割り当てるた めのガイドラインセットとして使用できます。

特長とメリット

運用の簡略化: Day0、Day1、Day2 以降の運用の簡略化により、必要なリソース、時間、コストを削減できます。業界屈指のインテントベースネットワーク構築により、手動設定に起因する問題を回避できるため、(ハードウェアのベンダーを問わず)導入とトラブルシューティングにかかる時間が短縮されます。

- クローズドループアシュアランスを備えた信頼できる単一の 情報源として、問題を予測して事前対応的に解決し、サービ スの中断を回避します。
- ・ 唯一のデータセンター向け AIOps により、さらなる自動化と 分析情報を通じてデータセンターのエクスペリエンスを最適 化できます。
- コスト効率に優れたソリューションにより、運用コストを 90%削減し、導入を85%高速化できます。

オープンな柔軟性:実証済みのテクノロジーと製品を使用してネットワークを設計し、ベンダーロックインを回避できます。

- ・ 高度なトラフィック管理機能により、イーサネットによる高 帯域幅、低遅延、ロスレス、拡張可能なパフォーマンスを確 保できます。
- イーサネットを活用することで、コストを削減して、イノベーションを加速し、設計の柔軟性を大幅に高めて、サプライチェーンの課題を回避できます。
- データセンターファブリックを管理および自動化するための 唯一のマルチベンダーソリューションにより、設計の柔軟性 が大幅に向上します。
- あらゆる GPU、ファブリック、スイッチで相互運用可能です。

ターンキーソリューション:信頼できる製品が揃った、エンドツーエンドかつ検証済みのターンキーソリューションにより、迅速に展開できます。

- 多様なシリコンにより、拡張性、パフォーマンス、柔軟性を 促進できます。数千もの GPU について、Broadcom とカス タムシリコンの、非常に高速(800 Gbps)の相互接続をサポートしています。また、あらゆるデータセンターファブリック構成をサポートしています。
- 包括的なデータセンターセキュリティポートフォリオにより、エンドツーエンドのセキュリティを確保できます。
- ジュニパー検証済み設計により、スイッチング、運用、セキュリティなど、すべてがそろったデータセンターソリューションを導入できます。
- ジュニパーが提供する幅広いスイッチ(多様なシリコン)により、拡張性とパフォーマンスを促進できます。
- ・ ゼロトラストデータセンターを保護するための統合脅威防御 を備えた、包括的なデータセンターセキュリティポートフォ リオを利用できます。
- データセンターネットワークの信頼性を 10 倍改善できます。

ソリューションコンポーネント

Apstra:ジュニパーのデータセンター自動化ソリューションにより、設計から継続的な運用まで、マルチベンダー環境のデータセンターネットワークのライフサイクル全体を自動化できます。 Apstraにより、継続的な検証、信頼できる唯一の情報源、強力な解析、根本原因の迅速な特定と修正を実現できます。

QFX シリーズスイッチ:業界をリードするスイッチシリーズは、優れたスループットと拡張性、包括的なルーティングスタック、Junos' OS のオープンなプログラマビリティ、最も幅広い EVPN-VXLAN および IP ファブリック機能を備えています。

PTX シリーズルーター: PTX シリーズルーターは、世界最大級のコアおよび WAN アーキテクチャの基盤を形成しています。最新

世代の Express ファミリー ASICs を搭載することで、クラウド向けに最適化され、800 GbE にスムーズに移行できます。今日の市場ニーズに対応するために必要な規模、柔軟性、および投資保護を同時に提供します。

データセンター向け仮想ネットワークアシスタント「Marvis」:

Marvis は、業界初で唯一のデータセンター向けの AI ネイティブ 仮想ネットワークアシスタントです。ハードウェアベンダーを問 わず、データセンターのライフサイクル全体を通じて、非常に優れた分析情報を出力します。さらに、キャンパスから支社/拠点、データセンターに至るまで、あらゆるエンタープライズドメインにエンドツーエンドの可視性とアシュアランスをもたらします。

Junos OS Evolved:信頼性、俊敏性、オープン性に優れたプログラマビリティを備えた、エンドツーエンドの統合型ネットワークオペレーティングシステムです。Junos OS Evolved により、クラウド規模の導入を成功へと導くことができます。

概要:AI データセンターのための運用の簡略化、オープン な柔軟性、ターンキーネットワーキング

事実上あらゆる企業が、データセンターのインフラストラクチャを最大限に活用して AI 関連の取り組みを実行するための戦略を策定しています。ジュニパーの AI データセンターソリューションにより、AI トレーニングと推論に適した、拡張性に優れた高性能ネットワークを非常に迅速かつ容易に構築できます。ジュニパーは、マルチベンダーでインテントベースの運用、AI ネイティブの仮想ネットワークアシスタント、ジュニパー検証済み設計により、複雑な AI データセンターネットワークを簡略化します。

Corporate and Sales Headquarters APAC and EMEA Headquarters

Juniper Networks, Inc. 日本, 東京本社

1133 Innovation Way ジュニパーネットワークス株式会社 Sunnyvale, CA 94089 USA 〒 163-1445 東京都新宿区西新宿 3-20-2

電話番号: 888.JUNIPER (888.586.4737) 東京オペラシティタワー 45 階 または +1.408.745.2000 **電話番号:** 03-5333-7400 www.juniper.net FAX: 03-5333-7401

Copyright 2024 Juniper Networks, Inc. All rights reserved. Juniper Networks、Juniper Networks ロゴ、Juniper、Junos は、米国およびその他の国における Juniper Networks, Inc.の登録商標です。その他すべての商標、サービスマーク、登録商標、登録サービスマークは、各所有者に帰属します。ジュニパーネットワークスは、本資料の記載内容に誤りがあった場合、一切責任を負いません。ジュニパーネットワークスは、本発行物を予告なく変更、修正、転載、または改訂する権利を有します。

3510814-001-JP 2024 年 1 月 4