

夏を涼しく乗り切る: 液冷がAIの次なる必 須要件となる理由

高度なプロセッサーは、データセンターの熱の限界を押し広げています。これに対応するために、企業はパフォーマンス、コスト、サステナビリティのバランスをとるために液冷に着目しています。

AIは、企業内で最も要求の厳しいワークロードの一部を実行しています。リアルタイムで病気を診断し、グローバルサプライチェーンの物流を最適化し、製品イノベーションの次なる波を推進します。しかし、あらゆるインサイトと自動化の背後には、フルパワーで稼働し、熱を発生しているサーバーがあります。

今日の高性能CPUおよびGPUは、従来の空冷システムが処理できる範囲を超えて、より多くの電力を消費し、より多くの熱を発生しています。ファンとヒートシンクでは対応できず、データセンターはパフォーマンスの低下、エネルギーの無駄、危険な音量、インフラストラクチャコストの増加といった代償を払っています。

かつてはスーパーコンピューターや特殊な環境に限定されていた液冷が、現在ではAIの熱需要に対する実用的な対応策として、より広く採用されるようになっています。ここでは、直接液冷システムを使用して、熱をより効率的に管理し、エネルギーの無駄を減らし、大規模なAIインフラストラクチャを常に稼働可能な状態に維持する方法について説明します。

AIに迫る熱による危機

大規模言語モデルとエージェンティックAIの台頭により、コンピュートの需要は未知の領域へと押し上げられました。ハイエンドCPUの消費電力は現在、日常的に500ワットを超えており、一部のGPUは1,000ワットに近づいています。同時に、熱耐性も低下しています。かつてチップは90~100°Cで安全に動作していましたが、今日の最先端のシリコンチップの中には、最大動作温度が60°Cのものもあります。¹

従来の空冷方式は、この新たな熱処理の現実に合う設計になっていません。最新のプロセッサーが発する熱を管理するには、大型ヒートシンク、強力なファン、より広い物理スペースが必要であり、これらすべてがコストを押し上げ、サーバー密度を低下させます。これはパフォーマンスとスケーラビリティを損なう非効率的なサイクルです。

また、エネルギー使用量も急増しています。米国のデータセンターの電力消費量は、2018年の76テラワット時 (TWh) から2023年には176 TWhに急増しており、2028年までに580 TWhに達する可能性があります。² 熱はAIの隠れたボトルネックとなり、その重要性は急速に高まっています。

液冷の仕組み

液冷の仕組みは、サーバー内の最も高温のコンポーネント (CPUとGPU) にクーラントを循環させ、熱を吸収して逃がしています。クーラントは主要コンポーネントに取り付けられた冷却プレート全体を循環し、熱を吸収して外部マニホールドに運びます。そこから熱が熱交換器に送られて除去されます。その結果、空冷よりも速く、効率的に熱を移動させます。

HPEでは、主に3つの液冷オプションを提供しています。閉ループシステムはサーバーシャーシ内に完全に収容されているため、既存のインフラストラクチャの段階的なアップグレードに最適です。代替オプションは、ラックに接続された空冷システムを施設の水を使用して冷却する液空冷です。これにより、最も高温のIT機器に冷気をより正確に供給し、熱を施設の水に戻します。最後に、直接液冷はラックレベルで動作し、複数のノードにクーラントを分配して、高密度で高性能なワークロード向けに優れた熱管理を実現します。

最も電力を消費するコンポーネントには液冷を使用し、残りのコンポーネントには空冷を使用するという複合的なアプローチもあります。このバランスのとれた設計により、システムを過剰に設計することなく、すべてのコンポーネントを効率的に冷却できます。

空冷システムと比較して、液冷システムは、パフォーマンスの向上、ラックの高密度化、エネルギー使用量の大幅な軽減を実現し、運用コストと炭素排出量を削減します。

¹ "HPE ProLiant Gen11 Servers with Direct Liquid Cooling," HPE, 2024.

² 2024 United States Data Center Energy Usage Report, Lawrence Berkeley National Laboratory, U.S. Department of Energy, December 2024



HPEが液冷を大規模に展開する方法

HPEは、最新のAIの規模、密度、効率性に関するニーズに対応する液冷ソリューションを設計しました。この取り組みの中心となるのは、液冷構成をサポートするHPE ProLiant Compute Gen12サーバーです。閉ループシステムは、既存のラックセットアップ内の高ワット数のCPUに最適な自己完結型の冷却ソリューションを提供します。直接液冷はさらに先に進み、クーラントをラック全体に分配して大規模な熱管理を実現します。高密度のGPUを多用するワークロードに最適です。

混合ワークロードを使用する環境の場合、HPEのAdaptive Cascade Coolingにより一元的なアプローチを提供します。この特許取得済みのシステムは、接続されたコンポーネントの熱負荷に応じて、空冷と液冷をリアルタイムで動的に切り替えます。その結果、エネルギー使用量が減り、インフラストラクチャの複雑さが軽減され、冷却効率が向上します。

どちらのアプローチを選択しても、コストの削減とサステナビリティの向上はかなりのものになります。パフォーマンスを犠牲にすることなくエネルギー消費を削減する方法を提供するため、運用ニーズと環境目標のバランスをとる組織にとって戦略的な選択肢となります。10,000台のサーバーを展開する場合、HPEの直接液冷により、冷却コストを年間200万ドル以上削減し、CO₂排出量を年間1,700万ポンド削減できます。³

これらのシステムは、100%ファンレスHPE CrayスーパーコンピューターやHPE ProLiant XDプラットフォームすでに使用されており、従来の冷却では対応できない高密度展開をサポートしています。液冷では完全なオーバーホールが不要なため、AIワークロードが拡大し続けても、データセンター全体を再設計することなく実用的な方法で熱需要に対応できます。組織は、高密度ワークロードや高温で動作するワークロードなど、最も必要とされる場所に高度な冷却を段階的に導入できるため、環境、社会、およびガバナンス目標の達成に向けて実践的に進めることができます。

³ [Liquid Cooling: A Cool Approach for AI](#), HPE Newsroom Blog, August 2024.

成長を考慮した設計: HPEのスケーラブルなAIへのアプローチ

HPEの液冷イノベーションは、将来を見据えたAIインフラストラクチャの広範な戦略の一部にすぎません。HPE Private Cloud AIのようなソリューションにより、液冷式HPE ProLiant Computeシステムと、NVIDIAの高速コンピューティング、統合ネットワーク、AI対応ソフトウェアを1つのターンキープラットフォームに統合します。本番環境規模のAI向けに設計されたこれらのシステムは、展開と拡張を簡素化しながら、価値実現時間を短縮します。

これらはオンプレミスソリューションであるため、組織はデータに対する完全な制御を維持でき、主権、コンプライアンス、または低レイテンシに関する要件を持つ企業にとってメリットとなります。同時に、予測可能な経済性と組み込みの冷却効率により、AIのニーズの増加に応じて運用コストを管理しやすくなります。

HPE AI Servicesは、組織がAIワークロードのパフォーマンスと熱需要の両方をサポートするインフラストラクチャを計画および実装できるよう支援します。これには、展開の規模とワークロードの種類に合わせた適切な冷却方法の選択が含まれます。このようにしてコンピュート、冷却、運用を調整することで複雑さを軽減し、時間をかけて効率的に拡張します。

熱対策で先手を打つ

AIの勢いは衰えず、インフラストラクチャに対するAIの要求も留まるところを知りません。チップの動作温度が上昇し、ワークロードが複雑化するにつれて、従来の冷却方法では対応できなくなります。液冷により、企業がインフラストラクチャを計画する方法が変わります。熱をより効率的に管理することで、高密度展開、エネルギー使用量の削減、ハードウェア寿命の延長など、AIワークロードのサステナブルな拡張に不可欠なすべての要素に対応します。

AIの将来に熱い期待が寄せられていますが、液冷があれば、インフラストラクチャはそれほど熱くならずになります。

詳細はこちら

HPE.com/ai/insights

[HPE.comにアクセス](https://HPE.com/ai/insights)

今すぐチャット

© Copyright 2025 Hewlett Packard Enterprise Development LP. 本書の内容は、将来予告なく変更されることがあります。ヒューレット・パッカード エンタープライズ製品およびサービスに対する保証については、当該製品およびサービスの保証規定書に記載されています。本書のいかなる内容も、新たな保証を追加するものではありません。本書の内容につきましては万全を期しておりますが、本書中の技術的あるいは校正上の誤り、省略に対しては責任を負いかねますのでご了承ください。

a50013353JPN

HEWLETT PACKARD ENTERPRISE

hpe.com