



レポート

In-Depth Report of Thermal Management Solutions for I/O Modules (I/Oモジュール向け熱管理ソリューションに関する詳細レポート)

データセンターエコシステム
冷却に関する最新情報

2024年5月

目次

概要	01
冷却の現状: 従来型の熱管理ソリューション	03
光I/Oモジュールの熱管理の課題	06
データセンターアーキテクチャ向けの革新的な熱管理ソリューション	09
次世代冷却方式に向けた標準化と試験	12
データセンター冷却の将来に向けたイノベーションの実現	14

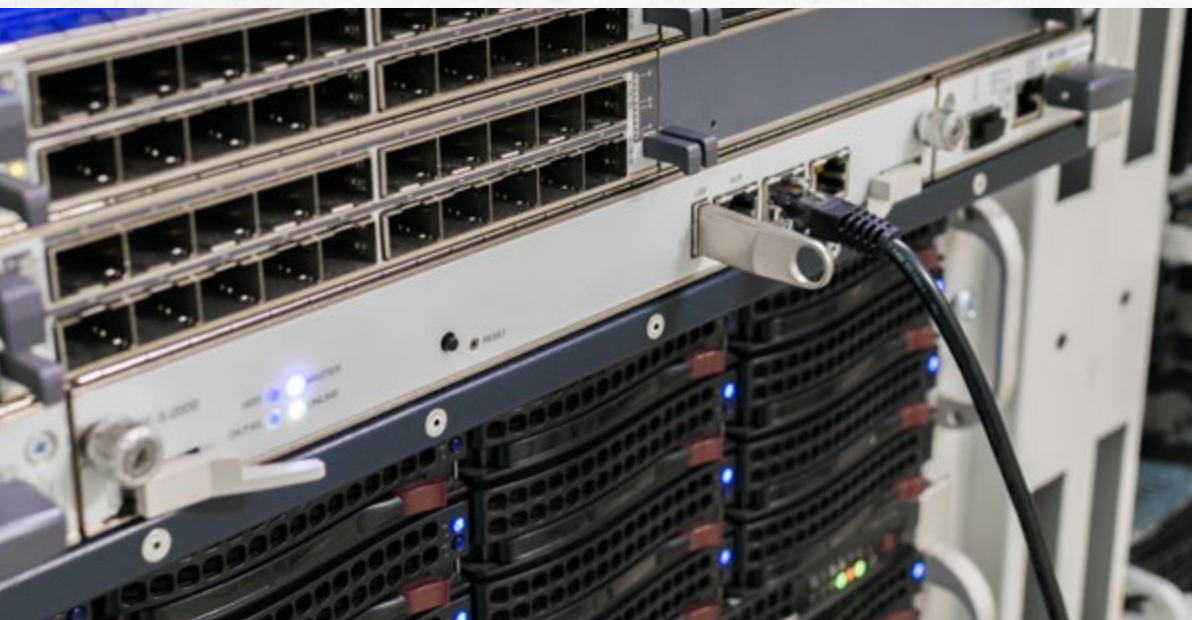
概要

基本的なEメールから高度な生成AI（人工知能）まで、様々なデジタルプロダクトやサービスの実現を主に担っているのが、データセンターのクラウドコンピューティング環境です。この演算処理にはコストがかかります。データセンターにある一台一台のサーバーを稼働させるには電力が必要です。特に、AIや機械学習その他、最先端分野向けに膨大なデータ処理が要求されるデータセンターにおける電力消費量は、相当に大きなものとなります。データセンターで主に大きな電力を消費するのは、こういった先端サービス向けの処理を担うGPUとアクセラレーターカードです。データセンターの急速な高速化が続き処理密度は最大限に高まるなか、データセンター固有の熱管理関連の課題も増えてきています。効果的な熱管理手法にフォーカスすることの重要性は、今までになく高まっていると考えられます。

デジタルトランスフォーメーションを採り入れる企業が増える中、高効率な処理能力と同時にメンテナンスコストと稼働コストの抑制を求められるデータセンターの負担は、さらに増大しています。データセンター稼働関連のコスト増という点において、代表的なコストの一つとして熱管理がありますが、効果的な熱管理によって各種コンポーネントの寿命を延ばすことで、長期的なメンテナンスコストは抑えられます。ITソリューション事業者、Enconnexによると、近年の水冷システムのキロワットあたり稼働コストは2千ドルに達し、エンタープライズデータセンターの冷却設備への投資額は10万ドルを容易に超えるとされています。このようなコストは、隅々にわたりコスト効率を重視しながら様々なビジネスを展開する企業にとって課題となっていることは明らかで、設備投資（CAPEX）や事業経費（OPEX）削減等の施策を進める際には、この熱管理関連コストへの取り組みも当然の出発点であると考えられます。



データセンターにおける熱管理との関連では、ラックマウント型サーバーとネットワークスイッチ、そしてデータセンター間の通信に使用される光モジュールに関するコストはあまり話題に上がってきませんでした。サーバーはそれ単体で動くものではありません。生成AI等の次世代サービスを提供するためには、光ファイバーリンクを経由してサーバー同士がクラスタ単位で通信を行う必要があります。こういったサービスを拡張するには、サーバーのクラスタの拡張が必要で、加えてクラスタ間通信に必要なデータ転送速度も必要となります。新たなテクノロジーの提供とデータ転送速度の高速化の実現に並行して、光I/Oモジュールおよびアクティブ電気ケーブル (Active Electrical Cable, AEC) トランシーバーに必要な電力量も増加しています。たとえば現在、データ伝送速度112 Gbps-PAM4で消費する電力量は約15 ~ 25W、そして大規模なエンタープライズ向けスイッチの電力消費量は、32ポート光I/Oモジュールのみに限っても最大で0.8kWになります。仮に、コヒーレント光伝送(800G) の光ケーブルを長距離の112G通信に用いた場合、モジュールあたりの消費電力は30Wにまで達します。電力量がこのレベルに達した時点でのI/Oモジュールの発熱量は、従来の強制冷却システムでの処理能力の限界に近づいてきます。



224 Gbps-PAM4相互接続への移行とはつまり、レーンあたりデータ転送量が2倍になるということです。電力消費量も増加し、長距離のコヒーレント光伝送の場合、光モジュール単独でも40Wにも上ります。光I/Oモジュールの必要電力量は、わずか数年で12Wから40Wに増加しているのに対し、モジュールのフォームファクタそのものは変わっていません。ここに課題が生じます。基本的に電力密度はほぼ4倍になるのですから、新たな冷却手法が必要となります。水冷ソリューションの実装には追加投資とメンテナンスコストを伴いますが、既存フォームファクタ内に独創的な液冷ソリューションを統合することによっても、I/Oモジュール内の消費電力と熱への対処も可能です。

光I/Oモジュール内の必要電力量の増加を受け、システムやデータセンターアーキテクチャの設計の現場では現在、今後の224G通信の実装を支えるための光I/Oモジュールへの水冷システムの利用の検討が進んでいます。水冷方式の先には、次世代の高速ネットワーク相互接続を可能にする、モジュール設計や熱特性へのもっと先進的なアプローチも考えられます。

今回のレポートでは、熱特性および熱管理に関する従来の冷却手法の限界について検討し、今後112Gおよび224Gの光リンクが要求されるシステムに実装される、サーバーの冷却および光モジュールの冷却システムにおける新たなイノベーションについて掘り下げていきます。

冷却方式の現状： 従来型の熱管理ソリューション

ハイパワーなシステムでは一般的に、ネットワークインフラから熱を取り除くためのシステムが必要で、能動冷却方式が使用されています。能動冷却方式には、先に述べたような投資とメンテナンスコストの問題が伴うのと同時に、能動冷却システムの設置・保守作業を行うことのできる熟練者も必要となります。データセンター内の一般的な能動冷却としては、次の各手法が用いられています。

強制空気 (直接空冷) 方式: データセンターの高床フリーアクセスフロアのプレナムからサーバーラックへ直接空気を送る方式です。サーバーとスイッチには個別に専用のファンを設け、エンクロージャからの冷気の取り込みを補助しています。プロセッサや光モジュールも含むサーバー内の各種コンポーネントを完全に冷却しようとした場合、このような設備の冷却能力には限界があります。

水冷方式: この方式では、熱質量の高い流体をコールドプレートに流し、ラックマウント構成システム内で発熱源となっているコンポーネントと接触させます。水冷方式の媒体として水も選択肢の一つですが、他の誘電体たとえば油やプロピレングリコール (PG-25) 混合液も一般に使用されています。



現状の能動冷却方式

プロセッサやASIC等の冷却需要は高く、昨今のすべてのデータセンターが能動冷却方式に依存しています。能動冷却方式は放熱能力の点からは最も効果的な手法で、流体を目標のコンポーネントに向けて直接流しながら受動部品で補助することでさらに放熱効果は高まります。強制空気および液体による冷却の両方式とも、近年のデータセンターで利用されている方式です。

これらの方式は、発熱しているコンポーネント方向に流体を直接流すための機構を使用して冷媒での熱交換を促進する仕組みを搭載しているものであれば特に、放熱能力の点で極めて突出した方式となります。ダイレクトチップ液冷方式は、強制冷却方式を一步進めた方式として、サーバー内の高性能プロセッサが主な発熱源となっているデータセンターでよく用いられています。

強制空気: 空冷は、高温コンポーネントと直接接触しているヒートシンクにエアを適宜、直接吹き付ける、低リスクな能動冷却手法です。ラックあたり必要電力量が10kW台のレベルであれば、強制空気方式で熱負荷にほぼ対応することが可能です。ハイパワーコンポーネントに対しては液冷方式が使用されている場合もありますが、チップやI/Oモジュールにおける電力要件が高レベルに拡大していったとしても、強制空気方式は今後も冷却方式の一つとして残っていくと考えられます。

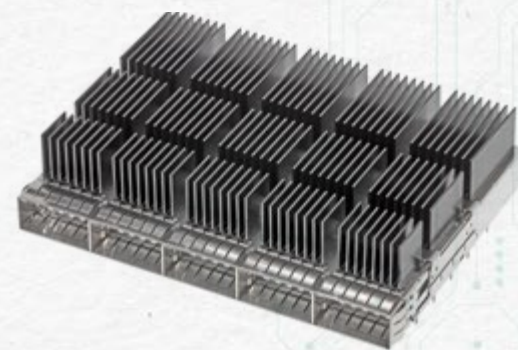
ダイレクトチップ液体冷却: ダイレクトチップ液体冷却方式も、データセンターにおける液冷方式の一つの選択肢として、昨今のクラウド環境で求められる高性能プロセッサ向けに多く用いられています。ダイレクトチップ冷却方式では、チップ背面露出部に接するコールドプレートに液体を流し、この液体でチップの熱を吸収します。Enabled Energy, Inc. のJeff Schuster氏によると、ラックあたり電力需要が25kW ~ 50kWに達した段階でこの**ダイレクトチップ冷却**方式による放熱が必要となってくるといことです。

ダイレクトチップ液冷方式は、強制冷却方式を一步進めた方式として、サーバー内の高性能プロセッサからの発熱が大部分を占めるようなデータセンター等で活用されています。

ただし、このような能動冷却方式は最も有効な手法ではありますが、システムとして複雑であり、メンテナンスの負担も最も大きくなります。液冷ソリューションの多くは、データセンター内のプロセッサ、アクセラレータ、および電源システム向けのもですが、光I/Oモジュールを冷却するソリューションのニーズも新たに現れてきています。現在は、データセンター事業者のほとんどが、サーバーやスイッチといった部分の冷却には強制冷却、I/Oモジュールには受動冷却といった形をとっています。

能動冷却能力を高める受動部品

能動冷却においては、熱伝導をサポートし熱質をいくつか追加する形で、なんらかの受動部品が冷却を補助しています。原理的には、能動方式の水冷や空冷システムに共通して使用されているのは、ヒートシンクとヒートパイプです。チップやGPUには、ヒートシンクやファン、液冷部品といった能動冷却部品を実装している場合も多いです。光I/Oモジュールの上にヒートシンクを搭載(ライディング型ヒートシンク)することで、強制空気に加える形で、高温モジュールからの熱の伝達を補助することができます。



QSFP-DD ケージにヒートシンクを搭載

光トランシーバーモジュールの内蔵型ヒートシンクとモジュールの上に搭載した（ライディング型）ヒートシンクへの熱伝達を補助するための、一つのソリューションとしては、温度への感受性が最も高い内部コンポーネント（例えばレーザー）を対象とした熱電冷却システムを搭載する方法があります。ペルチェ効果によってヒートシンクに吸収された熱は、空気流を介して放散されます。この手法は、内蔵タイプの熱伝達機構としては有用ではあるものの、224G光I/Oモジュールで要求されるレベルでの放熱需要を満たすことはできません。

液浸冷却

データセンターにおいておそらく最も効果的と考えられる液冷方式は、サーバー全体を非導電性の液体に浸漬して冷却する、液浸冷却です。液浸冷却では、十分な熱質量を提供する液体を、熱交換器を経由して循環します。冷却効率は極めて高く、ラックあたりで約50kWを超えます。

極めて高効率な液浸冷却方式ですが、大きなリスクとコストも伴います。以下に概要を説明します。

- **投資:** 液浸冷却システム関連の設備および設置コストは、強制空気や水冷方式よりも高額になります。空冷や水冷関連設備の改造等も含む、データセンターの構造全体の改修が必要となるというのが、主な理由です。
- **必要スペース:** 液浸冷却用のタンクを搭載可能なラックは基本的に、標準ラックユニットよりも幅、奥行きとも大きくなります。
- **対応I/Oモジュールとコネクタ:** 流体の誘電率は、コネクタの電気インピーダンスに影響します。コネクタは基本的に、動作に関係する流体を空気と仮定して設計されているため、液浸には専用のコネクタとトランシーバーが必要になります。

- **対応サーバー:** 液浸冷却で動作するサーバーは専用設計で製造されており、すべてのサーバーメーカーが提供しているものではありません。
- **液体 (冷却液):** 熱質量という点からは液体は効果的であるとはいえ、液浸冷却用の冷却液には、冷却するための特殊な循環装置が必要です。
- **メンテナンス:** 専用設備であるがゆえに、液浸冷却システムのメンテナンスコストは高額になる傾向があります。
- **漏れリスク:** 液浸冷却システムに壊滅的な漏れが発生した場合、設備の他エリアにも浸水による被害が及ぶおそれがあります。
- **コンポーネント故障:** 十分に冷却液を循環できないコンポーネントがある場合、そのコンポーネントは高温になり、劣化が加速して早期の故障につながる可能性があります。
- **環境影響:** 液浸冷却に使用される冷却用液には定期交換が必要で、正しい廃棄手順を経て廃棄する必要があります。

液浸冷却には、液浸に適した金属部品の設計や採用が必要になることがしばしばあります。使用部品については、冷却液との長期接触環境に耐える部品であるかを、評価する必要があります。112Gおよび224Gシステムにおける光I/Oモジュールの冷却需要を評価する際は、液冷装置を直接光モジュールに拡張することで、専用の液浸冷却システムの費用負担なく、冷却需要に対処することができます。

光I/Oモジュールの熱管理の課題

サーバーおよびラックマウント型ネットワークインフラシステム内にある光I/Oモジュールは、常に、能動冷却システムからの直接冷却、具体的にはラックマウント機器のフロントパネルからの強制空気によって直接冷却を受けています。ラックマウント型の機器の熱設計においては、I/OまたはASICの動作温度のマージンが過剰になることを回避するために、I/Oモジュールの熱管理とプロセッサやASIC内の放熱との間の均衡を考慮する必要があります。プロセッサの冷却需要と光I/Oモジュール全体の出力を考慮して冷却方式を最適化することで、両者間の適正な均衡が達成され、システムの電力効率が最大化します。

リンク距離 v.s. データ転送速度: 56Gおよび112Gに使用する光I/Oモジュールには、現在、空冷方式で十分な対応ができています。コヒーレント光通信をデータ伝送速度112G以上で実装する場合、プラグブル光I/Oモジュールの出力レベル(33W以上)では、モジュールへの水冷設備の追加が必要となる場合があります。

112Gおよび224G世代のトランシーバーはいまだ、**IEEE 802.3規格**で規定された標準リンク長を対象としているため、システム設計者やデータセンター事業者はこの規格が単純に、光モジュールの要求電力の増加に対応した規格に変更されるとは予想していなかったはずです。つまり、前世代の光モジュールの冷却需要は増加し、旧い熱管理手法では能力不足になることは以前から予想済みだったということになります。

//

前世代の光モジュールですでに存在した冷却需要は上昇することが見込まれ、旧い熱管理手法では力足らずとなるのではないかと考えられます。

//



フォームファクタ: プラガブル光モジュールは、およそ20年前の光ファイバトランシーバーモジュールの導入以来、フォームファクタが変わっていないという点に課題があると考えられます。業界が224Gに移行しようとする中、アップグレードを実施するには、新世代の光I/Oモジュールが既存のラックマウント型設備との後方互換性を備えていることが求められます。つまり、熱密度は今後も上昇し続け、これが、光I/Oモジュールの唯一の冷却手法としての強制空気による冷却の限界につながる可能性があることを意味しています。

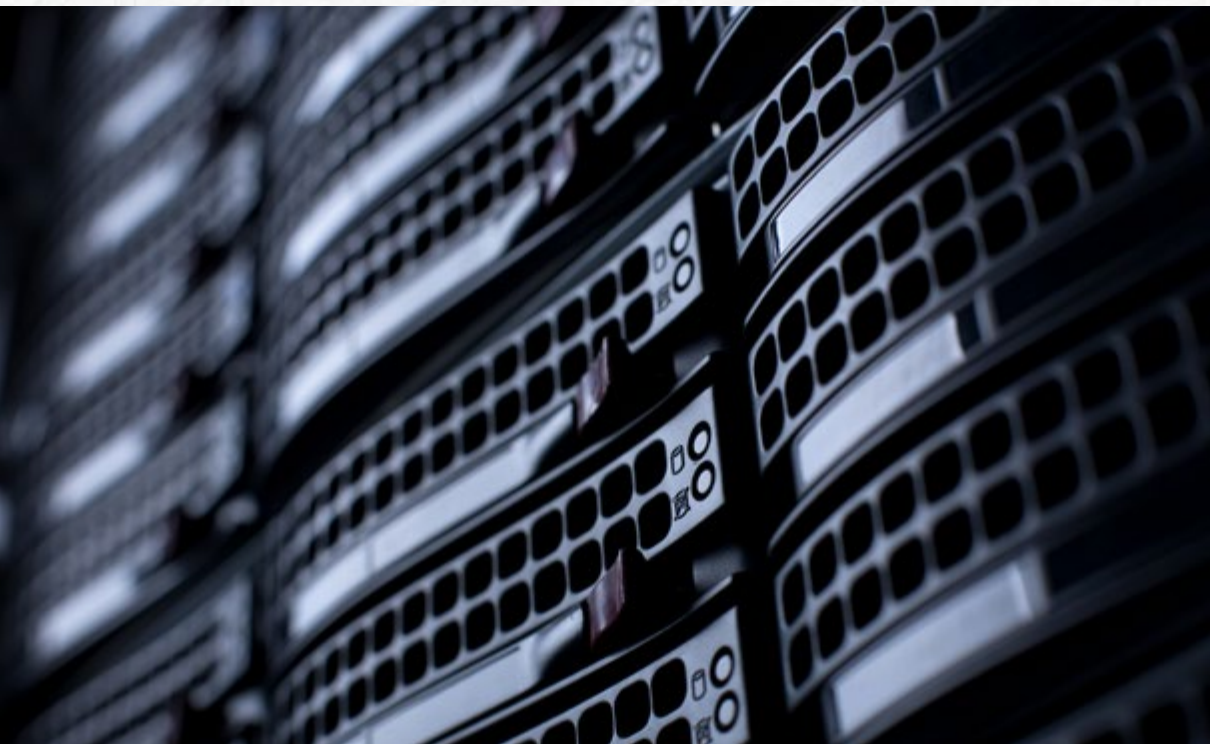
ヒートシンク: 光I/Oモジュールに搭載されているヒートシンクは、強制空気流による冷却能力を補強するものです。しかし、熱伝達を最大化しようとする際、耐久性要件の点で金属対金属の接触による拘束を受けます。どのヒートシンクにおいても、接触面における金属素材同士の直接接触は好ましくないのですが、ここ数年において出力レベルが著しく上昇しているI/Oモジュールについては、この金属の直接接触には特に気をつ

けなければなりません。モジュールあたり電力需要はこの先、40Wにも達すると予測されており、金属の直接接触を回避することの重大性はさらに大きくなります。裸金属面の接触熱抵抗を改善するには、プラガブルモジュールとその上で直に接触している(ライディング) ヒートシンクとの間に、サーマルインターフェース・マテリアル (TIM) を付加し、熱伝達効率の向上を補助するという方法があります。

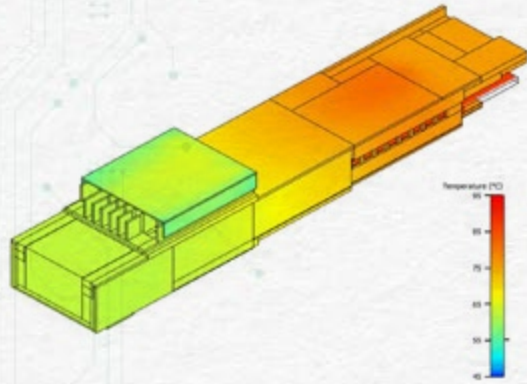
プラガブルモジュールの上にある(ライディング型) ヒートシンクにTIMを付加するとき、このTIMの信頼性が問題となります。ケージへの挿抜時、モジュールの鋭利な角がTIMを削り取っていくため、挿抜ごとに熱効率が低下し、初回の挿入時点でTIM性能が即低下することはないにしても、最初の数回で初期の有効性が損なわれてしまいます。実際の使用現場において、コネクタがケーブル重量により斜め挿入状態となる等の過酷な使用環境にさらされると、繊細なTIM表面が光モジュールの角部に接触する程度が増大し、耐久性の問題はさらに大きくなります。繰り返し嵌合後にも高い信頼性を保証するためには、TIMが100回の挿抜回数にも耐えるようヒートシンクの接触機構の再設計が必要になります。



ケージ / ヒートシンク上のサーマルパッドの損傷



モジュール温度の監視: 電力密度の増加に関しては、光モジュールに関する従来型の熱特性解析手法の再評価が必要です。従来より、70℃のケース温度要件が一括で温度仕様として用いられています (デジタルオプティカルモニタリング (DOM) 温度のプロキシとして)。しかし、最近の調査によると、ケース温度70℃においても、モジュール内の温度の影響を受けやすいコンポーネントには数度以上のマージンが残されていることがわ



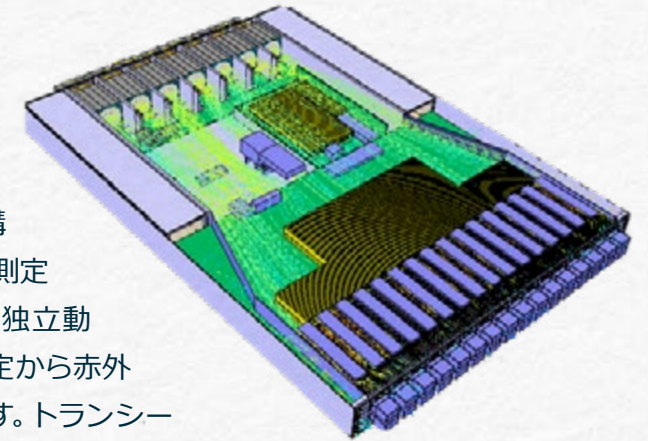
モジュール温度分布

かります。これがシステムの熱管理設計のフィジビリティについての誤った結論につながり、冷却システムの過剰な運転につながります。たとえば、I/Oモジュールの熱性能が制限要因となっているシステムにおいて、ファンは、モジュールの内部コンポーネント温度を基に余剰マージンを使えたとしても、単にケース温度要件を満たすためだけに必要以上の回転数で回転します。新たな熱特性解析を行うことが (本レポートの後

半で説明)、このような現行の冷却手法の制約の解決に役立ちます。

シミュレーションおよび試験: 構築から配備にいたる前のシステム設計、コンポーネント配置、および冷却手段の最適化には、シミュレーション / 予測分析を用います。ヒートシンクに強制空気をプラスした光モジュール冷却手法を最適化するためにしばしば必要となるのが、機械設計確定の前に行う、シャーシ全体内部の気流シミュレーションです。ラックマウント型サーバーは高さと幅寸法が標準化され、大部分が1RUフォームファクタを使用しています。その他のコンポーネント (例: チップ、アドインカード、SSD、等) の配置は、エンクロージャを經由してI/Oモジュール沿いに流れる空気流路に影響し、最終的に冷却効率を左右します。

光I/Oモジュールでは、モジュール本体上のホットスポットを特定するための、コンポーネントレベルでのシミュレーションも重要です。シミュレーションでは、モジュール自体の内部構造を検討した後、個々のモジュールの測定値の相関を検討する必要があります。独立動作時の温度試験は、接触部の温度測定から赤外線カメラによる測定までを実施します。トランシーバー内の熱プロファイルの把握後、この値をシステムレベルのシミュレーションの入力値として使用して、システムレベルの試験および相関分析を実施します。



サーマルシミュレーションシステム

液浸冷却: ハイパワーの112Gおよび224G光モジュールの冷却は、液浸冷却システムによって効率的に行うことが可能です。液浸方式は、熱負荷の観点からは最も効果的な冷却手法ではあるものの、誘電性流体にはモジュールのコネクター関連、主にはシグナルインテグリティ関連の課題を伴います。光モジュールおよびI/Oコネクターでは、周囲の誘電体として空気を仮定した設計がごく一般的であり、別の誘電体に置き換えることでカップリング効率の低下が生じます。結果として、液浸冷却するラックマウント型設備の112Gおよび224Gチャネルには、使用する誘電性流体に対応した専用モジュールが必要になります。液浸冷却を選択した場合、こういった機器等の調達しづらさや専門の設備構造が必要であることは、ラックユニットあたりのコスト増につながります。

データセンターアーキテクチャ向けの 革新的な熱管理ソリューション

サーバーおよび光I/Oモジュールにおける熱負荷の上昇、および後方互換性の要求に伴うフォームファクタの限界を考慮すると、データセンターにおけるデータ転送速度や演算処理の高速化要求の高まりに対応するためには、サーバーやスイッチに既に存在している水冷ソリューションを光モジュールまでに拡張することが必要であるかもしれません。I/Oに関して具体的には、信頼性を損ねることなくより大きな放熱機能を提供する、新たなソリューションをサーバーとスイッチに内蔵することが考えられます。これは、ラックマウント型ネットワークシステムとプラグابلモジュールで使用されている標準的なフォームファクターは従来のままを維持しつつ、機械設計を変更し革新的な液冷手段をモジュールに直接搭載する手法を用いることによって可能となります。

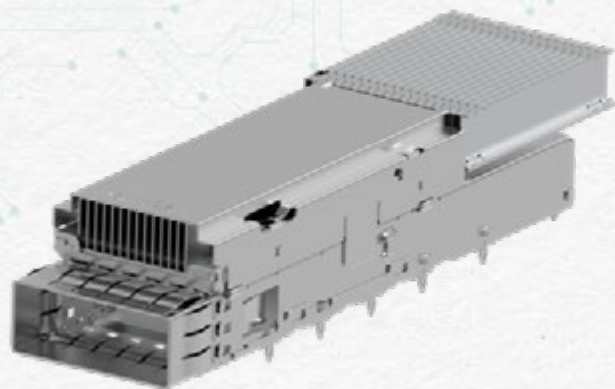
ドロップダウン・ヒートシンク

ヒートシンクの熱伝導能力を最大化するためには、ヒートシンクの台座とプラグابلモジュールとの間にある、乾燥した金属対金属の接触を、サーマルインターフェースマテリアル (TIM) により改善する必要があります。先に強調したとおり、光I/Oモジュールの挿入時、鋭利な角部がTIMを少しずつ損傷することが、嵌合可能回数の低下につながります。多数回の挿抜を繰り返した後もTIMの機械的&熱的インテグリティを保持するためには、ヒートシンクに別の接触機構が必要となります。



モレックスは、熱管理性能の向上を目指した、光I/Oモジュール上にドロップダウン・ヒートシンク (DDHS) を搭載する革新的ソリューションを開発しました。DDHSはモジュールとTIMとが直接接触しない画期的な設計となっており、プラグをリセプタクル体積の約90%まで挿入するポイントまでプラグがTIM面に接触しない浮遊型のヒートシンクです。最後の10%を挿入するところでヒートシンクがTIMに「ドロップダウン」し、モジュールの鋭利な表面の一切ない表面と完全接触します。このような機構とすることで、挿抜回数100回後もTIMの実装状態は十分に保持されます。ドロップダウン・ヒートシンクは、1列およびスタックケース構成への実装が可能です。

モレックスのDDHSは、現状流通している一般的なライディングタイプのヒートシンクソリューションとのドロップイン交換が可能です。すでに最適化されているジッパーフィン型のヒートシンクソリューションとの比較で、DDHSは最大でさらに9℃ (35W時) の放熱効果が得られます。



モレックスのドロップダウン・ヒートシンクシステム

このドロップダウン・ヒートシンクソリューションは、標準モジュールとラックマウント式フォームファクタの両方に適用可能な、信頼できる一つの熱管理方法を提供するものです。システム設計において、この9℃分の改善を次の2とおりの方で設計に取り入れることが可能です。

- モジュールを同じ出力 (例: 30W) で使用し、システムのファン回転速度を単純に下げてDDHSの温度マージンを最大まで利用して電力効率を上げる。
- 上と同じファン回転速度で、プラス5~7W (30Wから35~37Wまで上昇させる) のモジュールを冷却する。

DDHSソリューションをシンプルにドロップイン交換するだけで、従来よりも高出力のモジュールの冷却が可能となります。

最新の水冷ソリューション

データ転送速度112Gの光I/Oモジュール稼動時の出力レベルになると、従来型の強制空冷方式の冷却能力のほぼ上限に達します。転送速度224Gの実装となれば、光I/Oモジュールの発熱に十分に対処するためには、水冷方式も必要となってきます。演算処理能力の高いプロセッサ向けにはすでに水冷方式が使用されていることから、ハイパワー光I/Oモジュール向けのソリューションを既存の冷却システムに取り入れることは、理にかなっていると言えます。データ転送速度の高速化に合わせた新たな冷却技術の実装は、既存設備の改修によって行うことが可能です。

水冷方式そのものは、データセンター業界にとって真新しいものではありませんが、プラグابلI/Oへの実装という面から考えると、水冷方式にはこれ固有の課題が伴うことは確かです。水冷方式の実装手段としては、従来からある個々の上載せ型のヒートシンク (ライディング・ヒートシンク) を、一つずつコールドプレートで置換する方法を採ることが自然であると考えられます。しかしながら、この方法を採用した場合、入口と出口の数は32個にもなります。スペースの限られた1RUや2RUシステムのラックスペースに、これだけの配管を通しての管理は難しいでしょう。次のステップは、複数のI/Oポートを1枚で冷却可能なコールドプレートの実装です。この手法の場合に問題となるのが、I/Oポート単位の累積公差がモジュールの高さやケージ内での位置、台座の高さ等の要素の組み合わせによって異なってくるという点です。ポート1個単位では良好な接触面が確保できても、累積公差が異なる複数のポートが組み合わさることで、それぞれのポートで十分な接触面を保証することは難しくなります。たとえば1x6ケージ構成の場合、基本的にコールドプレートのすべての台座、およびコールドプレートに接触するすべてのモジュール面で完璧な平坦度が要求されます。このことは、各ポートの公差を確実に吸収しながら十分な放熱面を提供するに足りる力も付与できる、コンプライアント機能を有する台座の必要性を示しています。

このような課題を解決するために、モレックスは今回の水冷ソリューション、インテグレートド・フローティングペダスタルを開発しました。このソリューションでは、モジュールと接触する個々の台座をスプリングで独立動作させることで、コールドプレート1個の実装で1列 (1xN、2xN) およびスタックケースという異なる構成のI/Oモジュールの冷却が可能です。独立して挙動する個々の台座が、ポート単位で異なる累積公差を補償しつつ、望ましい押し下げ力を付与することで良好な接触面を確保します。

下図はその一例の、1x6 QSFP-DD液冷ソリューションです。このソリューションでは独立して挙動する台座を6個使用して、それぞれが対応するポートの累積公差を補償すると同時に、放熱のための十分な接触面も保証 (必要十分な下向きの力を付与) しています。

フローティングペダスタルの内蔵例



この内蔵型のフローティングペダスタルの利用により、放熱用ギャップフィラー等を用いずにI/Oモジュールを水冷することが可能となります。ギャップフィラーを使用した場合、熱の伝導経路に熱抵抗が追加されます。これに対し、モレックスのこのソリューションの場合であれば、熱源であるモジュールからペダスタルに直接流れた熱は、ペダスタルを介して直接、コールドプレート内を流れている液体と接触します。理論上、

これが水冷ソリューションで可能な最短の熱伝導経路となり、熱抵抗の最小化と熱伝導効率の最大化に寄与します。

境界条件には大きく左右されますが、モレックスでは、この水冷ソリューションで消費電力量最大40Wまでのモジュールを仕様値内まで冷却可能であることを実証しています。



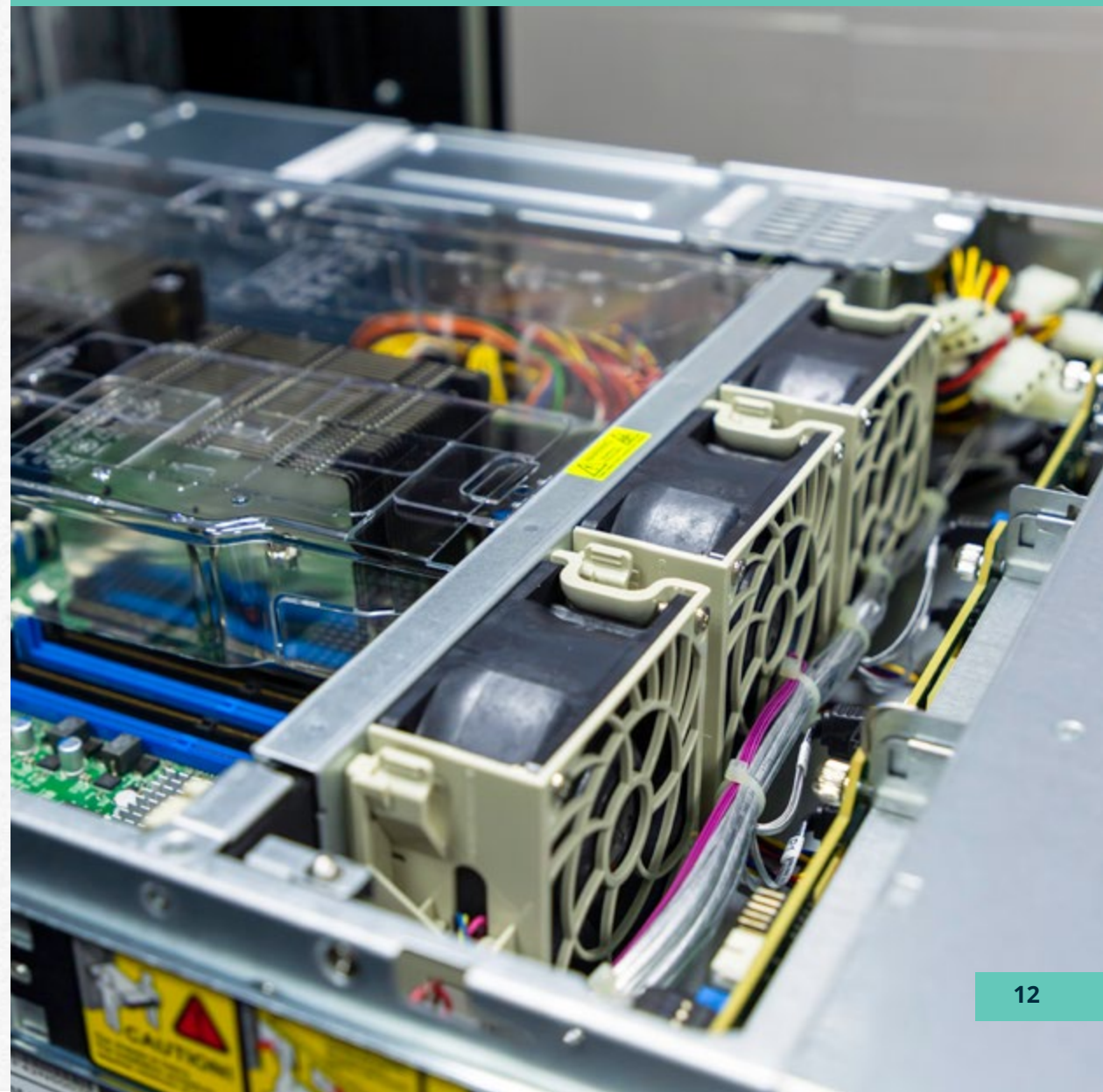
モレックスの水冷ソリューション実演

次世代冷却方式に向けた標準化と試験

ケース温度を仕様値として使用するか、それともモジュール温度の限界値として使用するかという点が、光モジュール向け冷却手法の設計に影響する一つの重要な要素となります。光モジュールは複雑な設計を持ち、ケース温度の仕様値単独では、モジュール内重要コンポーネントの内部温度を正確に反映したものにはなりません。あるモジュールが仕様値内で動作するかどうかを決定するのは、内部コンポーネントの上限温度です。ケース温度を仕様値として用いた場合、テーブル上に大幅な動作マージンが残る可能性があります。

従来のモジュール温度監視手法においては、モジュールケース上に温度監視ポイントが選択されており、おおむねヒートシンクの下が選択されていました。システムの冷却手法は、動作時温度がケース最高温度の仕様値 (T_{case} 、標準 75°C) を超過することのないよう設計されていました。通常は、運転中に、ヒートシンクを遮ることなくこの監視ポイントにアクセスすることはできず、またこの監視ポイントの温度は内部コンポーネントの実測値を直接反映したものではありませんでした。これに対し内蔵センサーからは、ソフトウェアマネジメントインターフェイス (CMIS等) で読み取り可能な、Digital Optical Monitoring (DOM) を使用した T_{case} 温度が通知されます。

ケース温度を仕様値として使用した場合、テーブル上に大きな動作時温度マージンが残る可能性があります。





ケース温度を用いた手法を用いた場合に失われるマージンの例を、下表に示します。スタックしたケースの下側ポートにある標準的なモジュールの測定値を示しています。モジュールケースの上限温度を、モジュールの動作と性能を保証するために必要な内部重要コンポーネントの実測値と比較することができます。内部コンポーネントの温度を見ると、設計マージンが依然として過剰であることが見て取れます。

モジュール	上限温度	実測値	マージン (ΔT)
Tcase (DSP上)	75℃	72.6℃	2.4℃
レーザー	85℃	76.4℃	8.6℃
TIA/ ドライバ	105℃	81.4℃	23.6℃
DSP	105℃	93.5℃	11.5℃

このケースでは、ファンの負荷を下げ、ケースが少し高い温度で動作可能なように冷却手法を再設計することが可能で、これによってシステムにマージンを少し追加できます。モジュールのケース温度をモジュールの温度上限として用いた場合、マージンはわずか2.4℃です。代わりに、レーザーを重要コンポーネントとして用いて（温度マージンを最小にして） 温度上限を定義してみると、レーザーの性能に一切の影響が生じることなく利用可能なマージンは実際のところ8.6℃であることがわかります。

このことから、下記の計算式で示すとおり、光モジュールのDOM測定値を、内部コンポーネントの最低温度マージンに基づいて再定義することが可能であることが示唆されます。先に言及したとおり、既存CMISおよびシステムソフトウェアとの後方互換性を維持しつつ、冷却システムの設計でマージンを追加することは可能です。DOMレジスターで通知する温度はしたがって、下記のとおりのようになります。

$$DOM = 75^{\circ}\text{C} - \text{Min}(\Delta T(\text{レーザー}), \Delta T(\text{DSP}), \Delta T(\text{TIA}), \text{etc.})$$

ここで提示したDOMの定義については、DOMの値、そして実際の温度マージンは、モジュールの動作環境においてマージンが最小である内部コンポーネント（レーザー、光部品、TIA、DSPチップ、等）に基づくものである、と解釈してください。DOMが通知する値を少し変えることで、システム設計段階での冷却システムアーキテクチャの過剰マージンを排除し、システム管理につながるより良いモジュール管理が実現します。

将来のデータセンターの冷却に向けたイノベーションの実現

弊社は複雑なデータセンター環境向けの熱管理に関する数十年にわたる経験と豊富な知識を活用し、革新的なアプローチで、データ転送速度の高速化に伴い増加する熱関連の課題に取り組んでいます。変化が急速に進む中で、システムの設計と実装に関わる制約は残ったままです。設備のフォームファクタは標準化されており、I/O数の多さを変えずにスペースの制約を満たせる独創的なソリューションが求められています。

モレックスは、厳しい電力要件で高速動作するシステムの信頼性を高める、光I/Oモジュール向けの冷却ソリューションを業界に先駆けて開発いたしました。プラグブル光I/Oモジュール向けのヒートシンクと接触手法に独自設計を施し、従来の熱管理ソリューションよりもはるかに信頼性の高い性能を備えた比較的シンプルなソリューションが実現しました。今回の弊社の開発は、負担を伴う液浸冷却手法に頼ることなく、データセンターの次世代の相互接続アーキテクチャをスケールアップする道すじとなると考えています。

確実に歩みを進めるためには、データセンターの複雑な熱管理で確かな経験を持つ、適切な供給元を選ぶことが極めて重要です。モレックスは次世代データセンターアーキテクチャ関連で最先端の能力を携え、顧客第一で皆様との協働を進め、性能と効率の両方の最適化を実現すべく、皆様の様々なご要望にお応えしていきたいと考えています。





creating connections for life



molex