

ソリューションガイド

224 Gbps-PAM4を使用した 次世代データセンター 構築ガイド

molex



目次

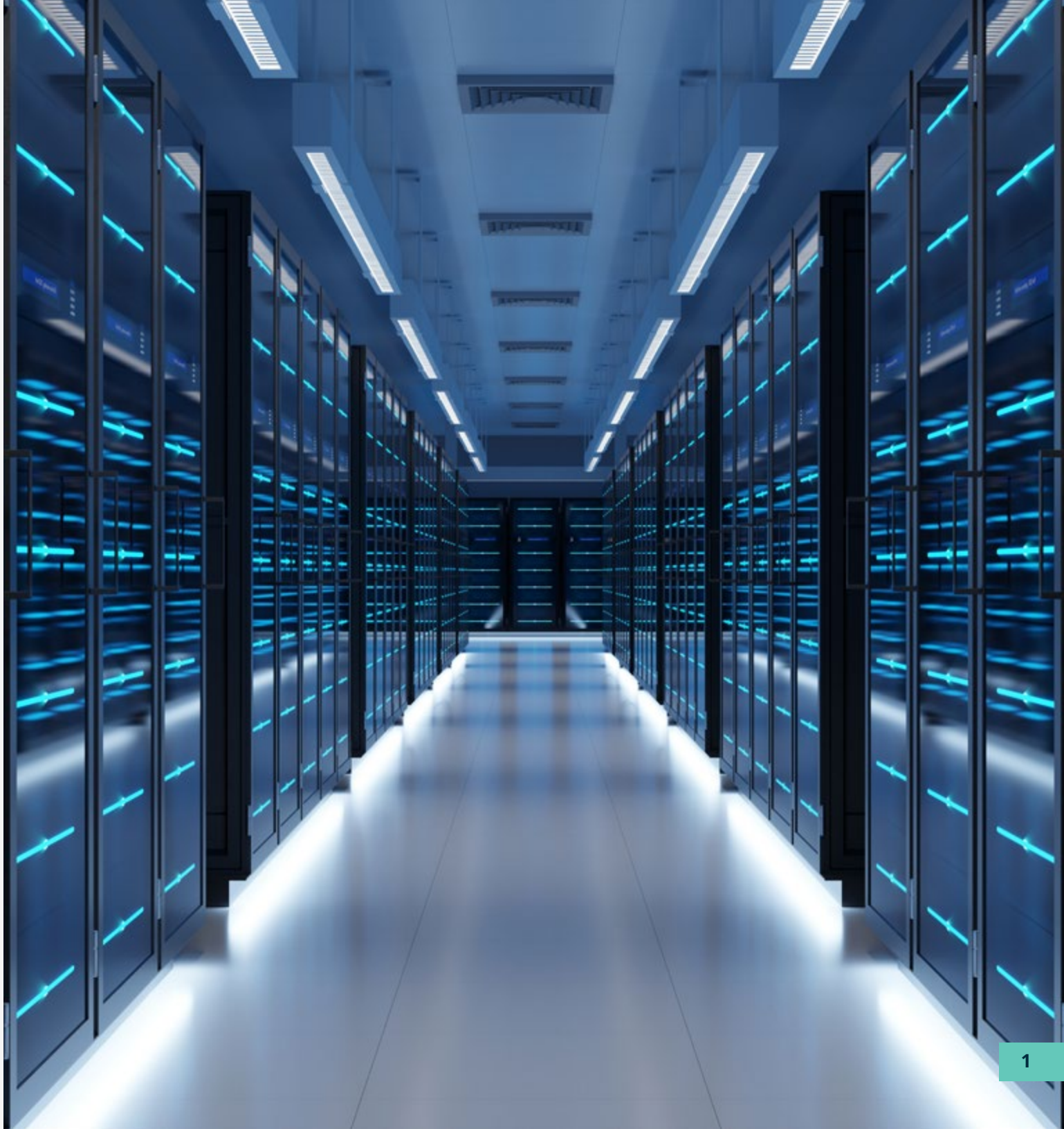
モレックスの224 Gbps-PAM4アーキテクチャ向けソリューション	01
最新のデータセンターアーキテクチャでの224G相互接続	02
224Gベースのシステムの設計プロセス	04
224Gアーキテクチャ向け、モレックスのコネクターシステム	07
明日のデータセンターを実現する	10

モレックスの224 GBPS-PAM4アーキテクチャ向けソリューション

高速シリアル通信は、世代を重ねるごとに伝送路あたりのデータ転送速度が段階的に高速化され、これに伴う新たな帯域幅要件に合わせて相互接続ソリューションに対するシグナルインテグリティの要求水準も厳格化されてきました。ナイキスト周波数を超えて伝送路の帯域幅を広げるために、エンジニアはあの手この手で対処してきましたが、224Gで動作するシステムには、これまでにない広帯域をサポート可能なコンポーネントが必要です。

データセンターに目を向けると、224Gチャネルは、ネットワークおよびコンピューティング機器の相互通信を可能にするために、コネクタを経由し、パッシブ & アクティブケーブルを経由し、光ファイバーを経由して、ASICからプリント基板（PCB）、バックプレーンにまで広がっています。コネクタとケーブルシステムなしでは、半導体の外、あるいはプリント基板上の極めて短い距離間ですら、通信チャネルは存在することはできません。

システムを望みどおりに機能させるためには、コンポーネント選定段階での伝送線路の適切なモデリングおよびシミュレーションが不可欠です。本ガイドには、224G通信向けに構築する次世代データセンターアーキテクチャの設計段階での検討項目と、これに関連する課題と解決策について解説し、さらに次世代データセンター向けの性能要件に適合するモレックスのソリューションに関する情報、シグナルインテグリティのデータ、そして今すぐに役立つ価値あるインサイトも掲載しています。



最新のデータセンターアーキテクチャでの224G相互接続

具体的な224Gソリューションはアプリケーションごとに異なりますが、標準的なファブリックおよびスケールアウトアーキテクチャは、図のようにシステム内の信号移動をなぞる形で示すことができます。図にはコネクタの役割も併記しています。Figure 1のブロック図に示すとおり、プロセッサユニット (CPU, GPU, TPU, etc.) で送受信する信号はBGA (ボールグリッドアレイ) パターン (BGA1) とプリント基板を経由します。チップ近傍のコネクタはこれら信号を受信して、ハイブリッドケーブルを経由して、高密度バックプレーン接続から外部バックプレーンケーブルへと信号を転送し、ファブリック接続チャンネルを形成します。

この後、信号はBGA2のランドパターンから別のチップ近傍のコネクタに進み、BiPassケーブルを使ってフロントパネルI/Oコネクタ、そして最終的にパッシブまたはアクティブ銅ケーブルあるいは光ファイバケーブルに出て行きます。続いて、データはデータセンター内の別のユニットに送られ、より大きなアプリケーションで使用されます。

このアーキテクチャ内には、224Gチャンネルにおいてコネクタが重要な役割を担うポイントが複数あります。ASIC上または近傍、バックプレーンケーブルインターフェース、BiPassケーブルインターフェース、ファブリックチャンネル、そしてスケールアウトチャンネルです。

224G向けコネクタの重要な役割を示した標準的アーキテクチャ

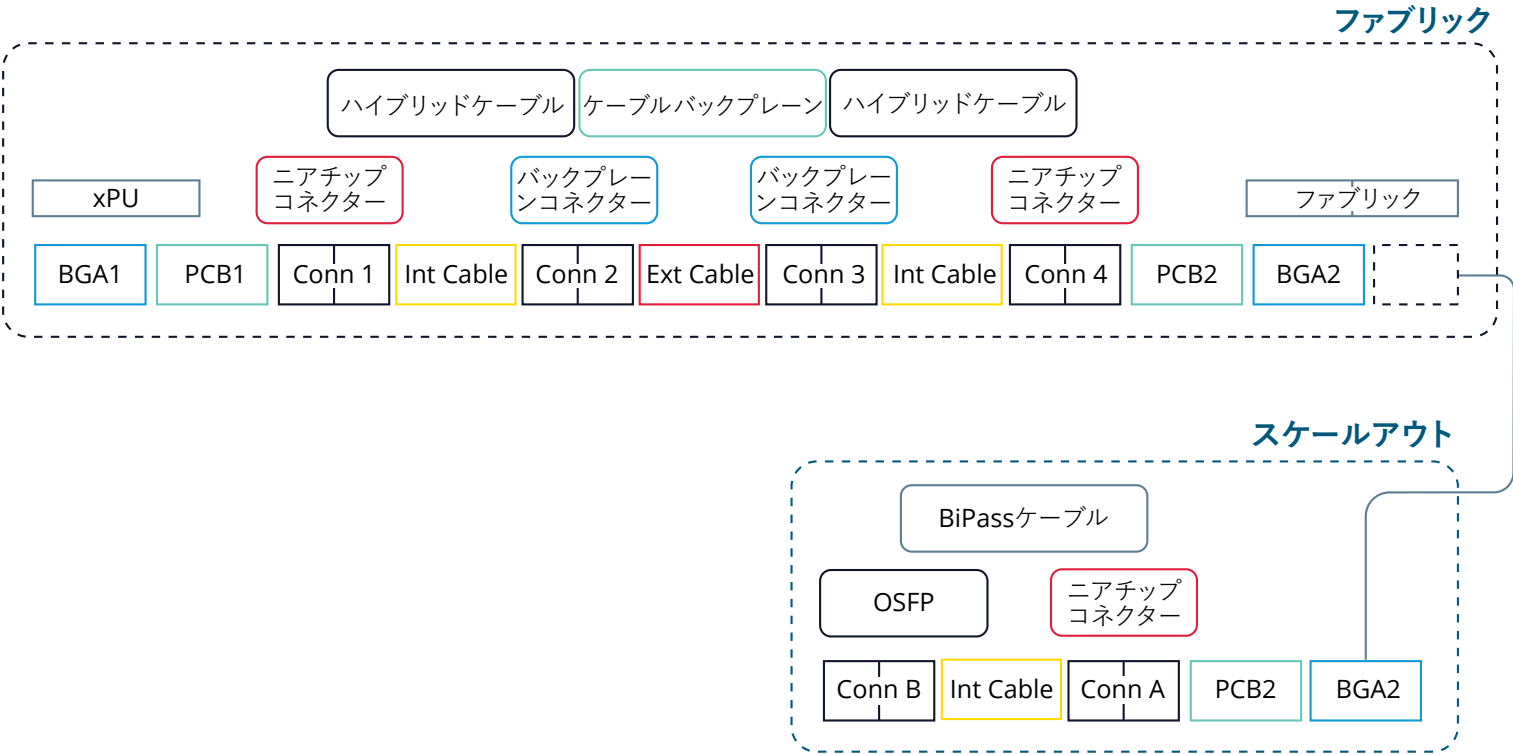


Fig 1: ファブリックおよびスケールアウトアーキテクチャのブロック図



ファブリックチャネルのコネクター

データセンター環境で動作するファブリックチャネルに使用されるケーブル量は増加しており、中でも増えているのがバックプレーンケーブルです。バックプレーンケーブルは、バックプレーンに接続されたモジュール間の高速データ伝送を担うケーブル&コネクターシステムで構成し、サーバーやインフラストラクチャ設備へのアクセスを提供しています。バックプレーンに使用されている専用ケーブルは損失が少なく、224Gの基板上のトレースよりも長距離の相互接続が可能なため、このケーブルなしでは完成できないアーキテクチャの場合の選定候補となります。

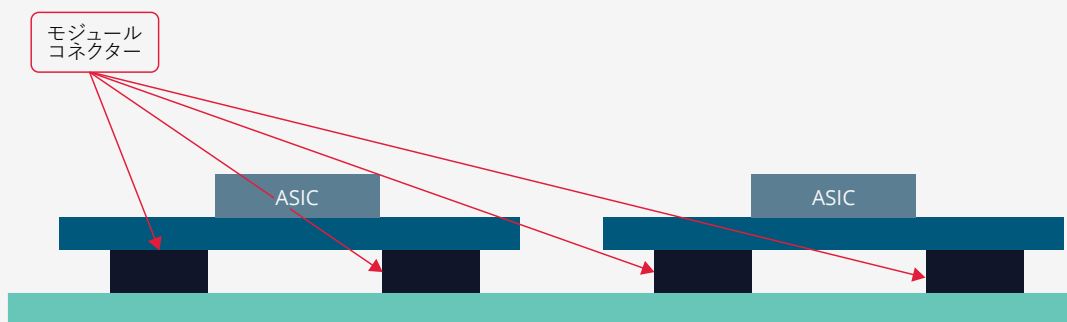
スケールアウトで使用するコネクター

システムに入力する銅または光ケーブルは、クアッド・スモールフォームファクタ・プラグابل・ダブルデンシティ (QSFP-DD) やオクタル・スモールフォームファクタ・プラグابل (OSFP) コネクター等の標準高速銅コネクターを経由してプリント基板に接続します。もう一つの、フライングリードケーブルはニアチップコネクターに接続してASICに直接またはその近傍 / プロセッサパッケージに信号を送ります。この先は、組み込みアプリケーションがファブリックチャネルを経由し、あるいはスケールアウトチャネル向けにデータを戻します。

ASIC向けモジュールコネクター

BGAはASICやプロセッサとプリント基板との間の主な接続機構であり続けていますが、一部のASICはモジュールとしてプリント基板上に搭載されています。これらのモジュールには、ASICからの信号をパッケージ基板からプリント基板上の銅相互接続部に伝えるための、基板対基板のメザニンコネクターシステムが必要です。

これらの相互接続製品は、BGAのランドパターン上への幅広信号線の精密設計を必要としない、ASICパッケージへのシンプルな直接接続手段を提供します。この特殊ランドパターンには、プリント基板内で特性を完全評価するための3次元電磁界シミュレーションが必要ですが、この専用コネクターシステムを使用することで、設計段階でのこのような評価・解析の負担を低減できます。



224Gベースシステムの設計プロセス

112 Gbps-PAM4による信号伝送がデータセンター環境に導入された当時、コネクタ製品の帯域幅が56 Gbps-PAM4を超えて2倍の112 Gbps-PAM4対応となったことは大きな前進でした。224Gでの運用は、伝送関連コンポーネントの帯域幅を単に2倍にするという簡単な話ではなく、帯域幅を必要なだけ広げるためには物理層のアップグレードまでが必要になります。つまり、コネクタやケーブルシステムから、プリント基板や半導体を製造するための基本的材料にいたるまで、システムのすべての部分が再設計の検討対象となることを意味します。

224Gの伝送レートで動作するシステムを設計するには、アーキテクチャ内の個々のコンポーネントのシグナルインテグリティの要件と、機械的フォームファクタの要件とのバランスを取る必要があります。データセンター環境はスペースに制約があることから、コネクタとケーブルの選定にはフォームファクタが極めて重要となります。しかしながら、すべてのフォームファクタやコネクタ製品が224Gでの信号伝送に対応しているわけではないため、コンポーネントの検証評価においてはシグナルインテグリティと機械的フォームファクタのどちらを優先すべきかを判断しなければなりません。

Table 1に、224Gシステムにおいてシグナルインテグリティに影響を与えるいくつかの要因を示します。

コネクタ インターフェース	- コネクタにおいては、コネクタインターフェース、付属ケーブル、ランドパターンが入カインピーダンスに影響 - 物理的構造がピン間のクロストークとコネクタ本体からの放射エミッションに影響
チップ I/O	- I/Oピンの負荷容量とパッケージのインダクタンスがSIに影響 224G対応とするには、プリント基板またはモジュールのコネクタ向けのチップのランドパターンは、幅広が必須
回路基板	回路基板には伝送路沿いに寄生容量による容量性負荷が生じる可能性
銅伝送路	- 中域 (ナイキスト近傍以下) での損失は、銅箔の粗さによる損失が主 - 高域 (ナイキスト近傍以上) では誘電損失が主 - 差動クロストークは最大でナイキスト以下 (平均 -40 dB) でなければならない

Table 1: 224Gチャネルにおける、シグナルインテグリティへの影響要因

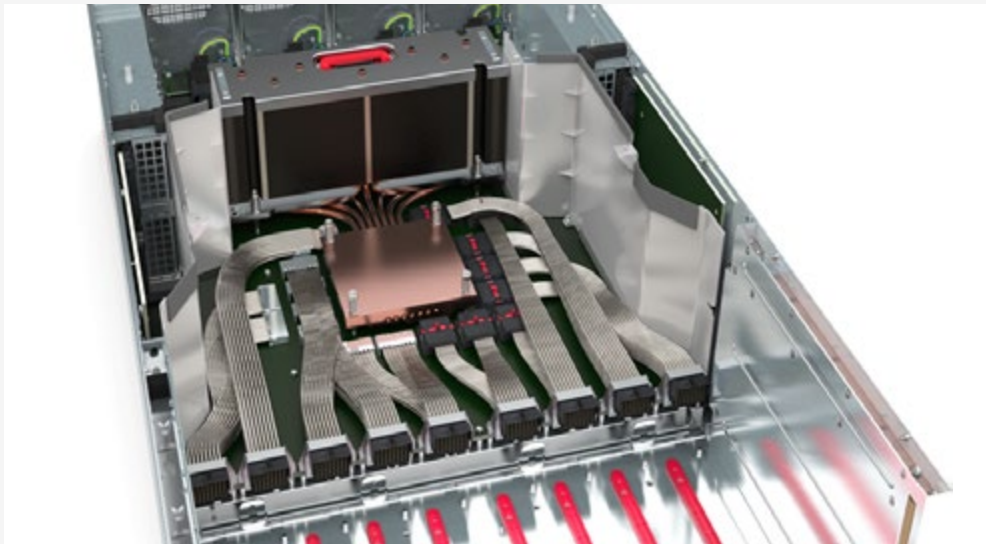
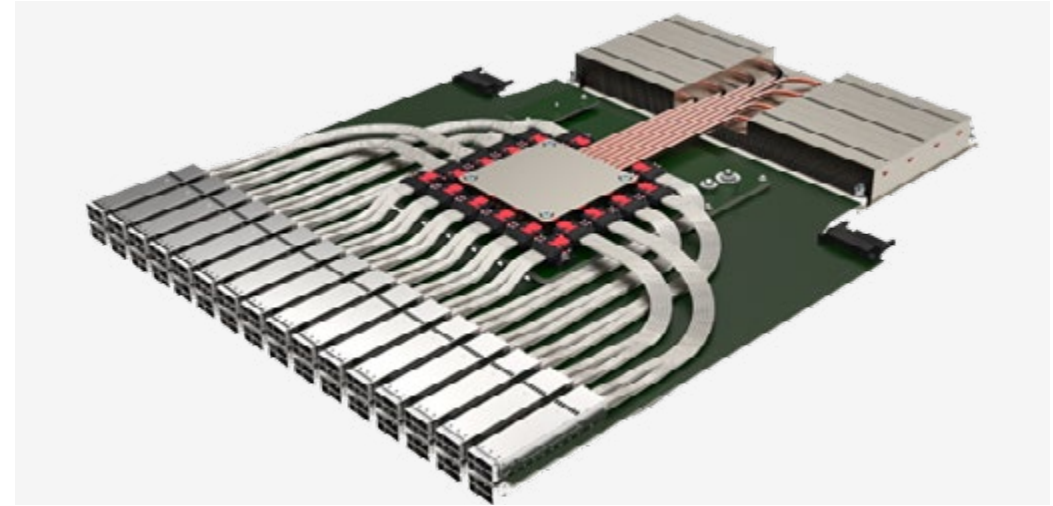
シグナルインテグリティを最優先で設計を解析評価するための最も良い方法は、評価対象のコネクタ製品にSパラメータモデルを使用したシステムレベルのモデリングを用いることです。評価対象のコネクタが224Gに適合となっても、Table 1に示すうちのいずれかの要因が性能に影響し、当該伝送チャネルが不適合となる可能性もあります。シミュレーションからコネクタのモデルを抽出する、あるいは測定値からSパラメータデータを取得するためのガイダンスをコネクタの提供元が提供する、あるいはコネクタの提供元自身がシミュレーションで使用可能な検証済みデータを提供することもできます。

一定の設計要素の範囲内で、システムの帯域幅を56 GHzよりも広げて224 Gbps-PAM4の接続を実現する方法はいくつかあります。方法としては、回路基板の製造に使用する素材を選ぶことや、伝送経路の設計、224G接続対応のコネクタを選定すること等があります。実際のところ、コネクタとケーブルは224Gチャネルのシグナルインテグリティに影響する最も重要なコンポーネントであるため、信頼できる革新的コネクタを提供可能なメーカーとの提携の重要性は一段と高まっていると言えるでしょう。

チップI/O、パッケージング

プリント基板や実装するケーブル、ケーブルと基板間のコネクタ等の要素の最終的な決定権限はシステムエンジニアにあります。システムすべての側面がシステムエンジニアの権限範囲にあるとは限りません。システムのホストプロセッサのパッケージングの検討において、このことは特に重要となります。最小寄生容量でのチップI/Oへのアクセスとナイキスト周波数超えでの確実な終端は、半導体 / パッケージングメーカーの責任であることから、相互接続周辺のロスバジェットを決定する制限要因となることが多くなっています。

システム挙動のモデリングを実施してシステムのその他の部分が目標値を達成できるかどうかを判断するには、チップとそのパッケージングに関する情報が重要です。224G伝送路内のプリント基板上の銅相互接続部は短いため、信号のパワースペクトルの相当部分がコネクタと相互作用します。システムの挙動をモデリングする際には、完全なデーター式を選択する必要があります。



プリント基板 (PCB)

プリント基板には、信号を相互に伝送可能な形でチップやコネクタを実装することが可能です。224Gチャネルの高周波数においてはプリント基板内ビアの寄生容量が高くなるため、プリント基板上の相互接続部のインピーダンス低下につながる可能性があります。プリント基板の相互接続部には損失が生じやすい性質があることから、224G用のプリント基板上の伝送経路は短くなる傾向があり、主にチップ (BGAのランドまたはモジュールのコネクタから) と電線対基板用コネクタまたは光ファイバトランシーバとのインターフェースに利用されています。

短い伝送経路はリターンロスの影響を受けやすく、プリント相互接続部の各終端における精密なインピーダンス整合が必要となります。これは、リターンロスを低く抑えるためには、コネクタ本体とチップのパッケージングへの入カインピーダンスの測定値は、224Gチャネルのナイキスト周波数を超過した領域ではフラットになっていなければならないことを意味します。銅線対基板用のコネクタ (Twinax等) の選定時には、コネクタメーカー側が、自社が供給する製品の224G要件への適合ならびに、システムレベルのシミュレーションに該当するシグナルインテグリティのモデリングデータ (touchstoneファイル形式) を示さなければなりません。



ケーブル

224Gアーキテクチャで使用する銅ケーブルは、ケーブル内の導体損失を考えると、このアーキテクチャの伝送距離に適しています。銅ケーブルは損失が生じやすい性質を持ちますが、twinax等の最新タイプのケーブルの場合は低損失の誘電体上の基板トレースよりも大幅に損失を抑えられているため、一定の条件においては銅ケーブルの方が好ましい場合もあります。twinaxケーブルは、構造的にもケーブル部分の差動インピーダンスを維持することが可能であるため、コネクタインターフェース部分のリターンロスを低減します。

1mを超える配線長が必要となった場合、あるいは相互接続部分のロスバジェットが低い場合は、パッシブ銅ケーブルの代わりになるケーブルが必要となることがあります。この場合は、アクティブ銅ケーブル (ACC) またはアクティブエレクトロニカルケーブル (AEC) の使用が可能で、相互接続リンクを構成するコストは光ケーブルよりも低コストとなります。ケーブルの配線長が3~5 mを超え、パッシブでもアクティブでも銅ケーブルの接続損失がロスバジェットよりも大幅に大きくなると、信号を光トランシーバに迂回して光ファイバケーブルで出力させます。OSFPやQSFP-DDトランシーバといったフォームファクタは通常、このような相互接続に使用され、システムアーキテクチャのスケールアウトの部分に見られます。

フォームファクタ最優先のアプローチ

224Gチャネルを使用したシステムを設計する別のアプローチとして、フォームファクタを基準にして適合するコネクタを選択する方法があります。エッジコンピューティングおよび人工知能 (AI) におけるアグレッシブな設計では、スモールフォームファクタが優先される場合があり、このため、コネクタの選定も設計プロセスの初期の段階でなされることになります。コネクタの選定に続いて、システムの残りの部分のシグナルインテグリティを最適化し、伝送経路が確実に基準に適合するようにします。

フォームファクタをベースにコネクタを選定することの一つの問題は、規格ごとに基準に差異があることです。特に、224G対応の相互接続ソリューションについては、統一された規格が存在しません。システムアーキテクチャのスケールアウト部分のフロントパネルに搭載するI/O用コネクタは、規格化されたフォームファクタとケージから選択可能であるのに対し、システムとして供給されるコネクタおよびケーブルは、メーカーによっては必ずしも互換性を備えているとは限らないのです。確実に拡張可能なシステムを設計するためには、この点は、コネクタの選定時およびメーカーとの初期段階でのやりとりの際に検討が必要な重要項目となります。

モレックスの224Gアーキテクチャ向けコネクタシステム製品

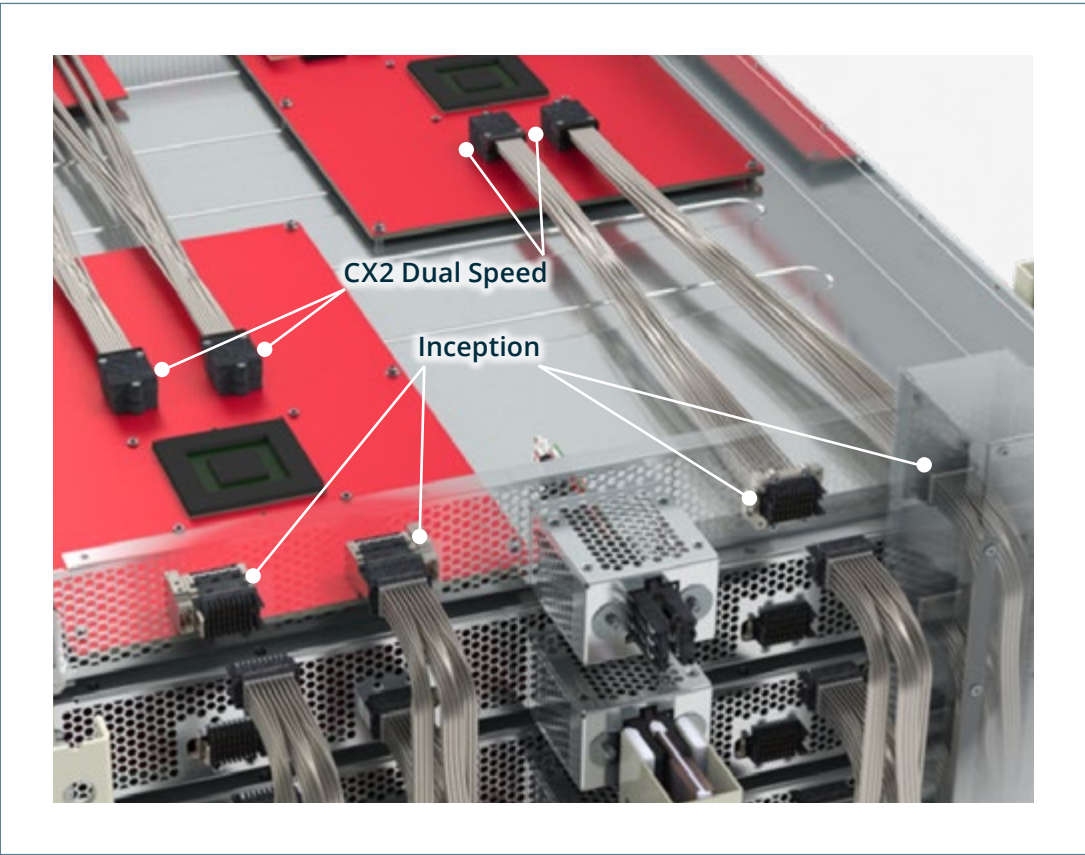
次世代データセンターアーキテクチャをリードするモレックスは、224 Gbps-PAM4対応の相互接続ソリューションの完全なポートフォリオを、業界で最初に発売しました。弊社の製品群は、ASIC相互接続ソリューションからケーブル、バックプレーンまで、システムアーキテクチャ全体を網羅しています。

バックプレーン向けジェンダーレスINCEPTIONと、CX2デュアルスピードNEAR-ASICコネクタ対ケーブル

Inceptionは、ケーブル接続を優先したバックプレーン向けの革新的なジェンダーレスシステムです。ラック構成ハードウェアアーキテクチャ内外の接続用では最も柔軟性に優れた高速ケーブルで、両性対応の信号接続と嵌合部の機械的な安定性を備えています。Inceptionは、通信用ハードウェアアーキテクチャのバックボーンとして効率的な通信とデータ交換を目指した製品で、相互接続部の機械的堅牢性が特徴です。ケーブル部は伝送損失を最小限に抑え、スペースの制約が極めて厳しい場所への搭載に適した設計となっています。

Inceptionでファブリックアーキテクチャを構成する場合、near-ASIC用のコネクタ&ケーブル製品であるCX2 Dual Speed (CX2-DS) と組み合わせることで、チップと周辺コンポーネント間あるいは外部接続機器との高速、低損失の通信チャネルを構成できます。CX2-DSのコネクタは、機械的堅牢性を主眼として設計されており、嵌合後にネジで固定することによる完全嵌合状態の確実保持とネジによる嵌合力の伝達、ワイピングによるコンタクト保護、嵌合作業時の握りやすさが特徴です。さらに高品質Twinaxの革新的な遮蔽構造により、Tx/Rx間の絶縁性能にも優れています。

InceptionおよびCX2-DS製品ファミリーのコンポーネントを組み合わせると、コネクタ2個のケーブルバックプレーンとコネクタ3個および4個のコネクタシステムを介した複数シャーシの相互接続が可能で、それぞれに最適な速度と機械的堅牢性を備えながらボックス間のスケールアップを行うことができます。

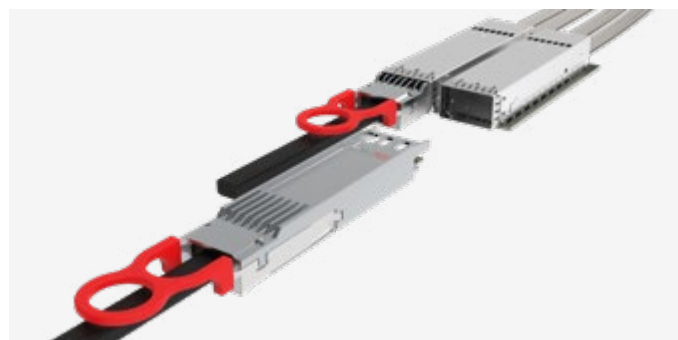


MIRROR MEZZ ENHANCED 224G 基板対基板用コネクタ

Mirror Mezz Enhancedは、モレックスのメザニン相互接続テクノロジーを224 Gbps-PAM4対応にした製品です。基板対基板の最低嵌合高さ5 mmの業界トップクラスの密度はそのままに、前バージョンよりもインピーダンス公差とクロストークを改善しています。Mirror Mezz Enhancedも両性対応のコネクタで、Mirror Mezz製品ラインのフォームファクタも同一であることで、システムの部品表 (BOM) の在庫保管単位 (SKU: stock keeping unit) の低減に貢献します。

本製品はモジュール性を優先した設計となっており、シャーシに必要な処理容量に合わせてコンピューティングモジュールを1個単位で追加 / 削除することが可能です。Mirror Mezz EnhancedコネクタとPCBAに対応したトポロジーをシステムに追加することで、追加のコンピューティングモジュールへの対応が可能となり、演算能力の段階的なスケールアップをスムーズに行うことができます。現在流通しているわかりやすい例が、OCP Open Accelerator Infrastructure (OAI) システムで、最大でGPUモジュール8台まで対応可能です。将来の、たとえば前述のようなアーキテクチャにおいて、OAIのようなシステムをInception/CX2 Dual Speed相互接続を使用したファブリックで使用した場合、バックプレーン用コネクタとケーブルを使用して個々のシャーシ間を接続することも可能です。





OSFP 1600、QSFP 800、QSFP-DD 1600コネクタ & ケーブルソリューション

224 Gbps-PAM4向けに設計された堅牢I/Oコネクタ & ケーブルです。優れた機械的耐久性と遮蔽性能を備え、クロストークを最小限に抑制、ナイキスト周波数の高域におけるシグナルインテグリティにも優れています。

OSFP 1600は、OSFP製品ファミリーの次世代タイプで、224 Gbpsアプリケーション向けのMSAと互換性があります。QSFP-DDとの後方互換性が問題にならない場合、高速データ転送速度の達成と熱管理の目的でOSFPコネクタで代用することが可能です。熱管理についてはOSFPコネクタの方が大型のためエアフローによる冷却性能に優れます。

MSA互換のQSFP 800およびQSFP-DD 1600の特徴は、フォームファクタがより小型であること、そして、前世代との後方互換性を備えていることです。

上記の製品ファミリーにはSMTコネクタ、BiPass、および外付けケーブルソリューションが含まれています。

- 耐久性の高いSMTコネクタ & ケーブルソリューションは、SMTはんだテールおよび基板との間の嵌合インターフェースの遮蔽性能に優れており、224 Gbpsチャネルの確実動作と低ビットエラーレート (BER) を保証します。
- BiPassケーブルの特徴は標準Twinaxを使用している点で、I/Oにnear-ASICとバックプレーンソリューションを組み合わせることで熱負荷とラックコストを低減し、設計の柔軟性を高めて、損失の大きいプリント基板をバイパスすることが可能です。
- ダイレクトアタッチケーブル (DAC) およびアクティブエレクトロニクスケーブル (AEC) は複数のケーブル構成に対応し、ケーブル終端処理が簡略化、そしてシステムのトラブルシューティングの関連ではCommon Management Interface Specification (CMIS) に対応しています。パッシブDACは、低コスト、低遅延アプリケーションに最適で、モレックスのリタイマー付きAECは、光ファイバーよりも低コストでの優れた保守性を発揮します。

コラボレーション

製品の選定、評価、そして量産に向けて、コネクタメーカーの専門知識の活用は一つの重要なステップです。モレックスは、協働を進めることで、早期の市販と、より細やかな製品選定、ならびにシステム運用に関する要求事項の最適なバランスが実現すると考えています。

- モレックスのエンジニアは、設計プロジェクトの開始時点で製品選定に関する大切なアドバイスを提供し、フロントエンドエンジニアリングと物理層設計の迅速な進行につなげます。
- モレックスは、ECADおよびMCADの両方を活用しリアルなレンダリングを実現する製品データを提供するほか、チャネルの適合判断に役立つシグナルインテグリティ関連データも提供します。
- 先端の相互接続ソリューションで拡張可能なアーキテクチャをサポートし、プロジェクトライフサイクル全般に関与します。
- モレックスは、製品の機械的・熱的特性の理解に基づき、システム要件に適合したソリューションを推薦します。

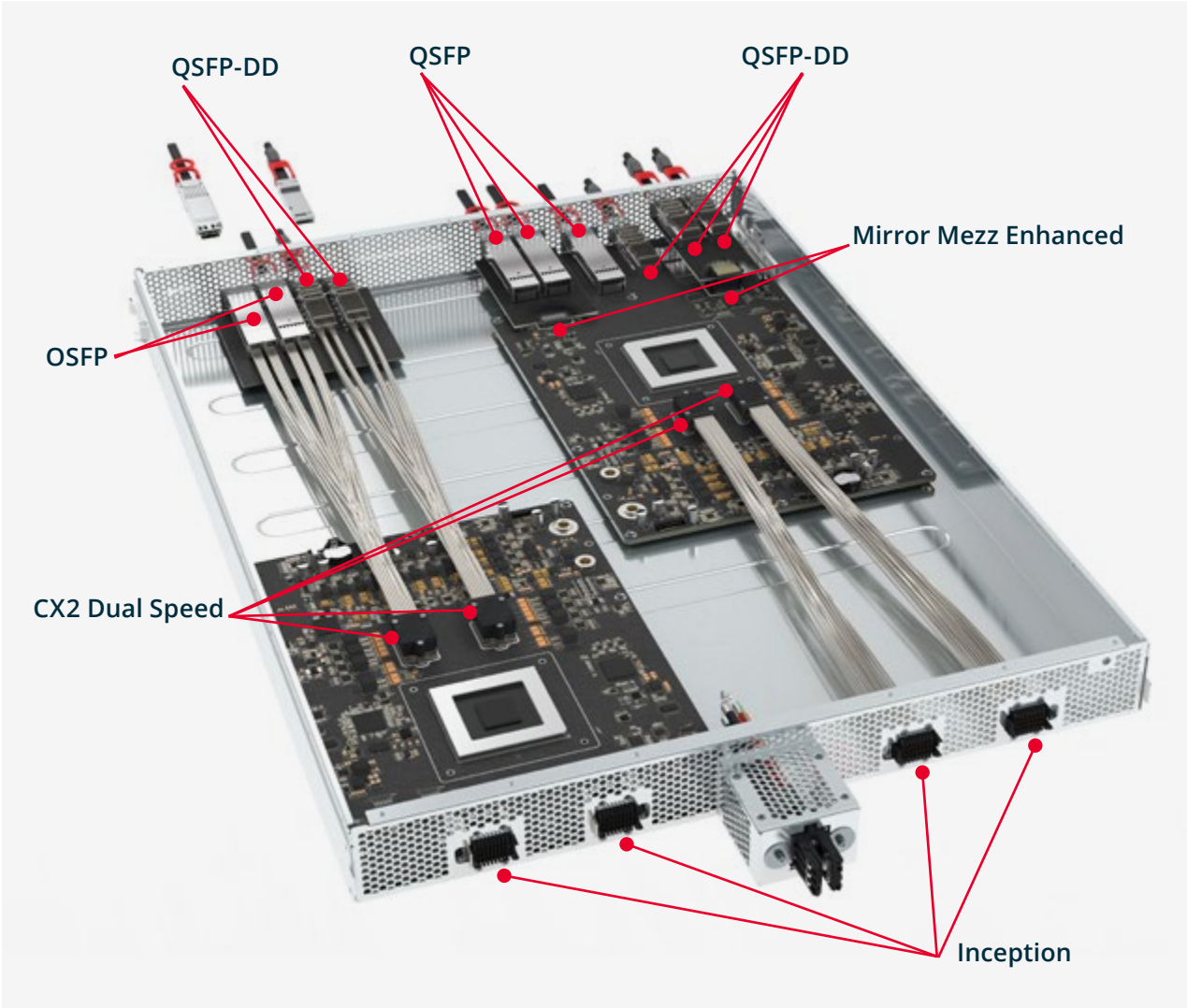


明日のデータセンター を実現する

生成AIやエッジコンピューティング、デジタルツインによるモデリングといった新たなテクノロジーの出現によって、データセンターの処理能力に関する要件の再定義が進んでいます。コンピューティングアセット間のデータ転送速度の高速化および大容量化は進んでおり、すでに224G相互接続に対応可能なソリューションへの需要も生じています。大量のデータ処理が必要なアプリケーションが増えるにつれ、処理能力向上への要求が高まり、これに伴ってデータ転送速度の倍化のトレンドも継続すると考えられます。この業界においては、さらなるデータ転送速度の高速化に対応すべく、互いの専門領域の垣根を超えた協力によってデータセンターアーキテクチャを変革していく必要があります。

転送速度の新たな限界に達するたび、限界の突破に貢献してきたコネクタとケーブルシステムは、これからも、さらなる広帯域の相互接続を実現する重要なイネーブラーであり続け、コネクタ関連企業は今後もシステム設計において重要な役割を果たしていくことになるでしょう。モレックスはこれまで、より先進的なコネクタシステムを開発することで、速度の限界や物理法則の限界を乗り越え高速化に対応してきました。弊社エンジニアリングチームの先端知識を活かし、さらに要求の厳しい将来のアプリケーションを実現する次世代のデータセンターアーキテクチャを、皆様と共に開発していけることを願っております。

まずはエンジニアリングチームまで、お問い合わせください。モレックスが業界に先駆けて発売した、224G向け製品についてご案内いたします。弊社サイト、molex.com/224Gもご覧ください。





creating connections for life

molex