

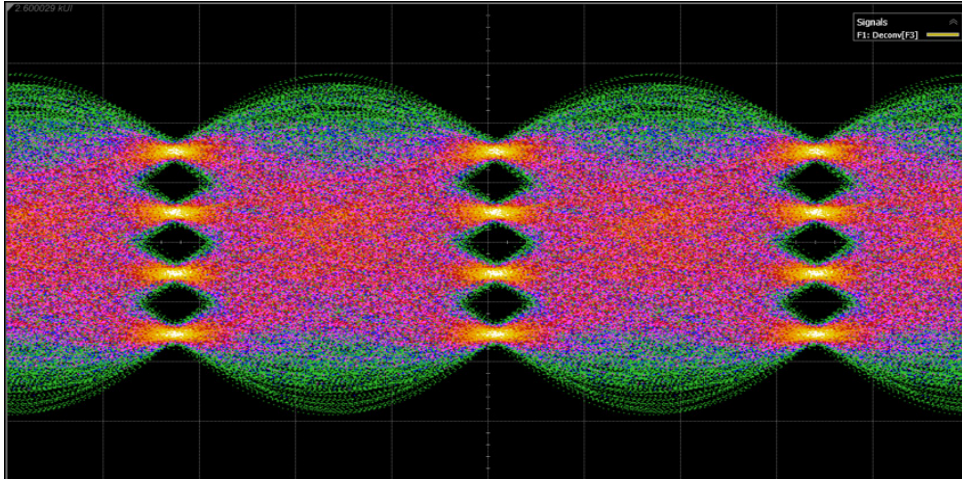
データセンターの イーサネット技術および 224Gbpsへの進化

目次

はじめに.....	3
データセンターインターコネクト	4
データセンター相互接続技術の開発.....	6
800Gイーサネットへの移行	8
データレート関連用語	10
800Gイーサネットの技術的課題	11
準備.....	15
112 Gbpsテストソリューション	15
224 Gbpsテストソリューション	17
信号発生	18
コンポーネントテスト	19
波形テスト	21
ビットエラー比テスト	22
テストソリューションの概要	24

はじめに

デジタルデバイス（5G、モノのインターネット等）用のネットワーク帯域幅に対する要求が増え続ける中、データセンターはそれに対応するためインフラをアップグレードする必要があります。最大速度224 Gbpsのシングル・レーン・インタフェースで、800Gおよび1.6Tの研究が進行しています。アプリケーションノートでは、最新のデータセンターにおける高速イーサネットリンクの最新の進化について説明しており、最大224 Gbpsインタフェースで利用可能なキーサイト・テクノロジーのテストソリューションを紹介しています。



データセンターインターコネクト

大規模のインターネットデータセンターは光相互接続技術およびイノベーションにおける成長が最も著しい市場で、拡大を続けるマシン間通信によるデータセンター内の全インターネットトラフィックは70%に及びます。CLOS¹アーキテクチャー（リーフ-スパインとも呼ぶ）に基づく典型的なデータセンター・ネットワーク・ストラクチャーを図1に示します。

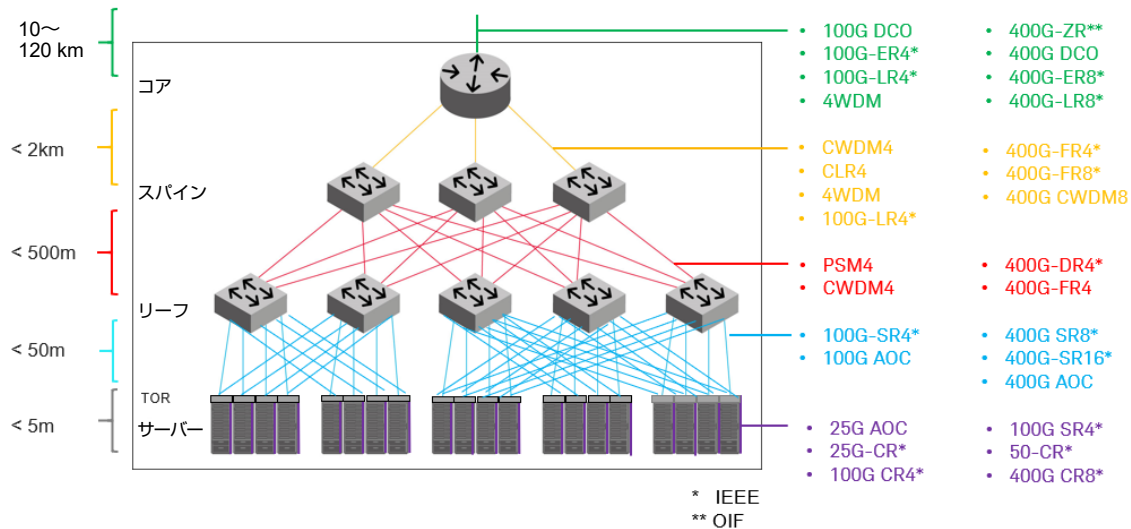


図1. 典型的なハイパースケールデータセンターのネットワークアーキテクチャー

通常、データセンターの内部ネットワークは下から上まで3~4のレベルがあります。

サーバーからコアまでレベルが移行するにつれ、各相互接続の到達距離は数メートルから数キロメートルに拡大し、技術とインタフェースの規格の変更が必要とされます。

サーバーキャビネット/トップ・オブ・ラック・スイッチ (TOR): 最下レベルでは、各サーバーラックがキャビネット上のTORスイッチに接続されます。通常、現在のデータセンターは25Gネットワークを導入し、50Gの速度を利用する人工知能 (AI) アプリケーションを一部取り入れています。今後数年に、100G、200G、400Gの速度の相互接続技術が導入されることが予想されます。接続距離は短く、キャビネット内、または隣接するキャビネットとの間で、通常は5メートル未満です。今日使用されている典型的なインタフェース技術はダイレクトアタッチ銅線ケーブル (DAC) またはアクティブ光ケーブル (AOC) です。通信速度が400Gおよび800Gに進展するにつれ、DACの距離はかなり短くなり、アクティブ電気ケーブル (AEC) が代わりに使用されます。

TORからリーフスイッチ: 2つ目のレベルはTORスイッチからリーフスイッチへの接続です。この距離範囲は約50メートルで、現在は100Gの相互接続が使用されており、今後は200Gおよび400G、数年後には800Gに移行すると予想されます。通常、100GBASE-SR4または200GBASE-SR4などの光モジュールはマルチモードの光ファイバーとの組み合わせで、NRZ (非ゼロ復帰) シグナリングと共に使用されています。また、このレベル以上の相互接続では、200Gおよび400Gへの移行により、シグナリングがPAM4 (4値パルス振幅変調) に変更されます。

1. CLOS Topology <https://howdoesinternetwork.com/2019/clos-topology>

リーフからスパイン：リーフからスパインへの接続範囲はキャンパスまたは隣接するキャンパス内で、その距離は最大500メートルです。TORからリーフに同様のインタフェースレートを使用しており、100Gから200/400Gへの移行が進んでおり、2023年頃には800Gに移行することが予想されます。より長い到達範囲を備えたこの技術は、シングル・モード・ファイバー、さらに多くの場合に100G-PSM4、100G-CDWM4などのモジュールを使用しており今後は200GBASE-DR4や400GBASE-DR4への移行が見込まれるいくつかの平行ファイバーに移行します。

スパインからコア：到達範囲はさらに最大2キロメートルに拡大し、ファイバーのコストが懸念となる段階であり、多くの場合、波長分割多重化技術を使用して、1つのファイバー上の複数の光波長でデータが送信されます。現在は100GBASE-LR4、100G-CWDM4、400GBASE-ER4/LR4/FR4などのモジュールが使用されています。

データセンターインターコネクト (DCI)：通常、これはロードバランシングや災害復旧バックアップを目的とする隣接する複数のデータセンター間の接続です。接続距離は10キロメートルから100キロメートル程に及びます。この長距離で高密度波長分割多重通信が使用されており、最近では、直接検波技術よりもコヒーレント通信が使用されています。通信事業者は、長距離（数百キロメートル）のアプリケーションにおいて長年にわたり100Gのコヒーレント技術を導入しています。通信速度は200G、400Gに拡大しており、今後はさらに800Gに拡大すると予想されます。DCIでは、伝送距離は通信会社のアプリケーションほど長距離ではなく、主にポイントツーポイントであるため、400-ZRなどの小型で電力消費が少なくプラグブルモジュール技術を使用したコヒーレント伝送が可能です。

データセンター相互接続技術の開発

図1に示すように、データセンターではいくつかの異なる電気および光の相互接続技術が使用されており、常に進化し続けています。

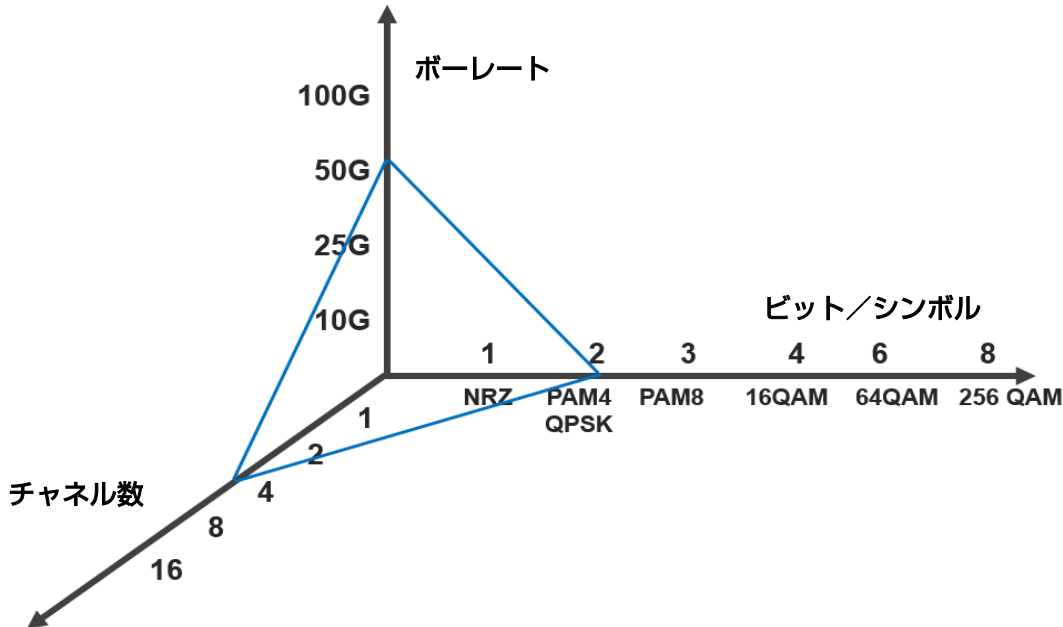


図2 データセンターのイーサネット相互接続速度の改善のための技術

各インタフェースの速度は、パフォーマンス、到達範囲、電力消費、およびコストの間のトレードオフが様々に異なる2つ以上の方法またはインタフェース規格により達成が可能です。

相互接続インタフェースの速度を向上させる3つの技術的な方向性があります（図2）。

最初の方法はチャンネルのデータまたはボーレートを直接改善する方法です。例えば、SDH/SONETの時代における155 Mb/sから622 Mb/s、または100 Mb/sイーサネットポートから10 Gb/sギガバイト・イーサネット・ポートへの大幅なアップグレードです。多くの場合、ボーレートの改善が必要とされる場合、その時点で利用可能な技術に先駆けて起こるため、他の方法が使用されています。

2つ目の方法はチャンネルの数を増やす方法です。これはボーレートを一定にする利点がありますが、インタフェースのデザインに関連する追加の費用と複雑さが伴います。例えば、10 Gb/sから40 Gb/sへのイーサネットインタフェースの移行には、シングルチャンネルの40 Gb/sリンクとその実装にかかる高額な費用ではなく、4 x 10 Gb/sチャンネルの手動が用いられました。同じ手法を用いて、100Gイーサネットへの移行が行われました。これには最初に10 x 10 Gb/sチャンネルを使用し、後に4 x 25 Gb/sチャンネルが使用されました。これはメインストリームの100Gイーサネットインタフェースの実装となりました。電気インタフェースでは、この手法によりデバイスと回路基板のチャンネルが常に多くなることを意味し、新しいデザインの検討点としてクロストークが導入されます。光インタフェースでは、複数のチャンネルを短距離伝送に対応した並列のマルチモードまたはシングルモードのファイバーとして、また波長分割多重化方式(WDM)を使用することにより長距離伝送に対応した単一ファイバー上に実装できます。ポート当たりの使用可能なファイバーの種類と数は、通常、既存のファイバーインフラにより設定され、既存のデータセンターへの高速光インタフェースの展開に影響を与えます。一般的に、4つまたは8つの波長がWDM（間隔5 nm）またはCWDM（間隔20 nm）

で使用されます。一部の最先端の研究では、単一ファイバーから複数のコアを作り空間多重化伝送を実現するフューモードマルチコア（few-mode multi-core）のファイバーが研究されています。

3つ目の方法はより複雑な変調方法を使用する方法です。25 Gb/sまでのデータレートでNRZシグナリングが使用されています。業界が400Gイーサネットの技術要件を提起した時、特に電気の分野でデバイスの帯域幅、パッケージ、PCBのデザインの限界を押し広げ、データレートを53 Gb/sに改善することには課題が伴いました。チャンネルの数を増やすと、ソリューションの必要スペースと電力消費が増大し、データセンターの全体的なコスト/ビット削減の目標達成に役立つとはいえません。その結果、NRZに対してPAM4の複雑な変調が提案されました。PAM4 4レベル変調では、各データシンボルによりシンボル当たり2ビットの伝送が可能となり、同数のチャンネルとボーレートに対してインタフェースデータレートが2倍になります。

一般的に、長距離のコヒーレント光通信の分野では複雑な変調技術が使用されます。例えば、100Gコヒーレント通信では、通常、1つのシンボルが2ビットを伝送可能なQPSK変調を使用する一方、400Gコヒーレント通信では、シンボル当たり4ビットを伝送する16-QAM変調を使用します。ワイヤレス通信と新しいコヒーレントの実装では、8ビット/シンボルのスループットをもつ256-QAMが使用されます。

上記のすべての手法は、データセンターのデータスループットを改善するために様々な状況で使用されています。図3は、100Gおよび400Gイーサネット規格の電気および光のポート・データ・レートの遷移の概要です。データレートの増加に加え、トランシーバーのフォームファクターと内部アーキテクチャーも小型化と低電力消費化が進んでいます。

最初期の100Gトランシーバーは、電気および光ポートに対応した10レーンで、CXPまたはCFPでパッケージされていました。それに続くギアボックスの開発により、10 x 10 Gb/s電気レーンが4 x 25 Gb/s光レーンに変換できるようになりました。今後開発が成されるまでのトランシーバーにおけるこの増大した複雑性と電力消費により4 x 25 Gb/s電気レーンが実用化され、トランシーバーのアーキテクチャーはシグナルをリタイミングするためCDRに簡素化することが可能でした。これは今日最も一般的な手法で、QSFP28フォームファクターとあわせて、コスト、消費電力、サイズ、性能のバランスが取れています。

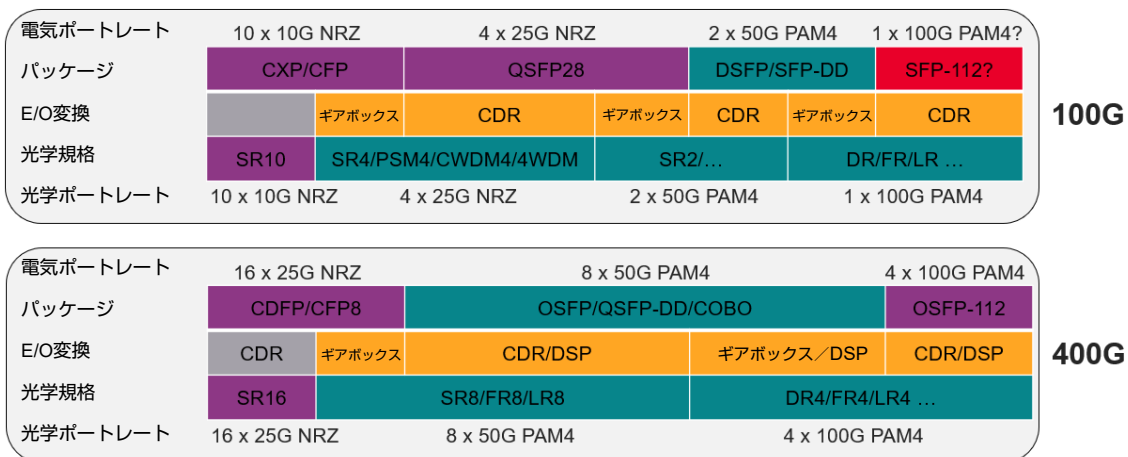


図3. 100G/400Gイーサネット電気および光ポートのデータレートの進展

100Gをさらに改良し、ポート当たり2 x 50 Gb/sまたは1 x 100Gb/sレーンのPAM4への切り替えが可能です。

400Gイーサネットの技術開発も同様のプロセスに従います。最初の開発はCDFPまたはCFP8パッケージとNRZシグナリングの16レーンが中心となります。しかし、これはコスト／複雑さの観点からは商業的でした。現在、最も一般的な実装は、電気ポートの8 x 50G PAM4 (25 GBd)、および、QSFP-DDまたはOSFPフォームファクターを使用した光ポートの8 x 50G PAM4または4 x 100G PAM4 (50 GBd) の組み合わせです。8 x 50G 電気レーンから4 x 100G 光学レーンを生成するにはギアボックスが必要ですが、複雑さと消費電力が増えるものの、必要なレーザーと検出器が4つのみとなる点で相殺されています。

最適な構成は光レーンと同数の電気レーンを設けることでギアボックスの必要性を取り除くことです。業界が800GEおよび100G電気レーンへ移行するにつれ、400G技術は4 x 100G PAM4電気および光レーンへ進化します。

800Gイーサネットへの移行

次世代のイーサネット速度クラスである800Gの開発が始まっています。第1世代の800Gはレーン当たり112 Gbpsを採用して200/400/800Gリンクを実現し、第2世代はレーン当たり 224 Gbpsを採用して最大1.6Tリンクに対応します。規格団体はすべて800Gに対応するプロジェクトまたは規格策定グループを開始しており、

- OIF Common Electrical I/O (CEI)-112GおよびCommon Electrical I/O (CEI)-224G
- IEEE 802.3 Beyond 400 Gb/s Ethernet Study Group
- 800G Pluggable MSA (マルチソースアグリーメント)
- OSFP MSA 200G/lane Electrical Signaling Group

図4は、800G-MSA規格団体が発行したホワイトペーパーに記載されている800G光相互接続の技術ロードマップを示しています。800Gイーサネット光ポートは、主に2つの相互接続方法があります。1つの方法は、ファイバー当たり100G最大100メートルの短距離相互接続に対応した8つの並列単一モードファイバー (800G-PSM8等) を使用します。もう1つの方法は、ファイバー当たりまたは波長当たり最大2 kmの中距離相互接続に対応した4つの並列単一モードファイバー (800G-PSM4または800G-DR4等) またはWDWによる1つの単一モードファイバー (800G-FR4等) を使用します。また、IEEE802.3dbのワーキンググループでは、8チャンネルのマルチモードファイバを用いて短距離(数十メートル) の800G伝送を実現することが検討されていますが、現状では、市販のVCSELレーザーの帯域幅に制約されています。業界に受け入れられる可能性はまだ未知数です。

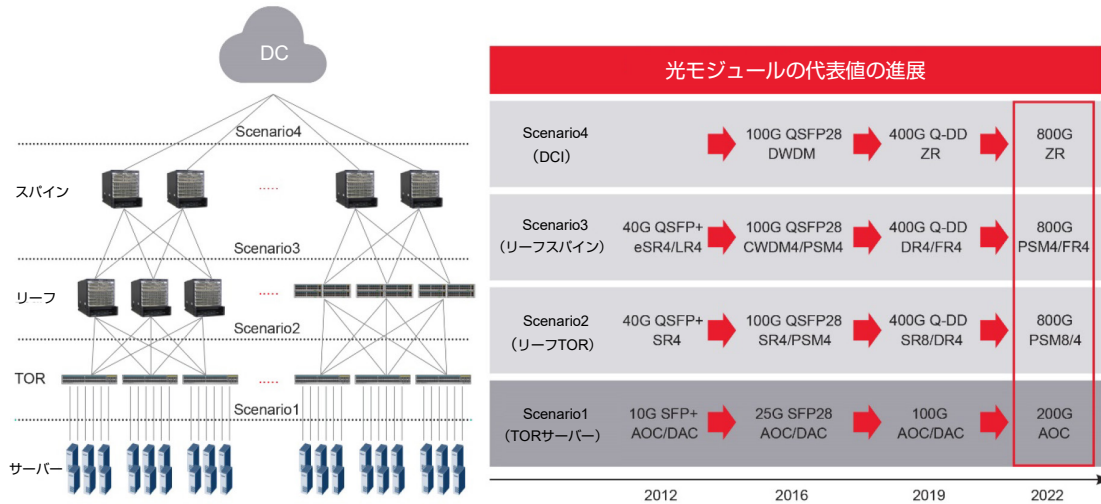


図4. 800G Pluggable MSAにより組織された光相互接続の技術ロードマップ²

典型的な800G光相互接続の導入を8チャンネルまたは4チャンネルの光ポートソリューションとして図5に示します。図中の黄色のブロックは、現在利用可能な400Gイーサネット技術への800イーサネットの追加を示しています。

800G-PSM8またはDR8の導入は、光チャンネルの数が2倍になること以外は400G-DR4とそれ程変わりません。既に成熟したチップおよび業界のサプライチェーンサポートがあります。電気インタフェースでは、以前の400Gイーサネットは8チャンネルの50 Gbps接続を使用していました。800Gでは、800Gイーサネット・スイッチ・チップと光モジュールの間の電気インタフェースの速度をチャンネル当たり100 Gbpsに上げて、8つのチャンネル全体で800 Gbpsを提供できるようにする必要があります。また、800Gイーサネット光モジュールには400G光モジュールに一般的に使用されているQSFP-DDまたはOSFPパッケージを使用できませんが、400Gより高い電気インタフェースの速度をサポートするためパフォーマンスを改善する必要があります。2020年以来、半導体メーカーは単一チャンネルの100 Gbpsのパフォーマンスとさらにパフォーマンスの高いQSFP-DDおよびOSFPパッケージに対応するCDR/DSPチップを発売しました。そのため、8 x 100 Gbps電気および光ポートを実装して800G接続を達成する技術の利用が可能となっています。電力消費とコストの管理が可能となれば、比較的早くこのソリューションの商用化が可能となるでしょう。

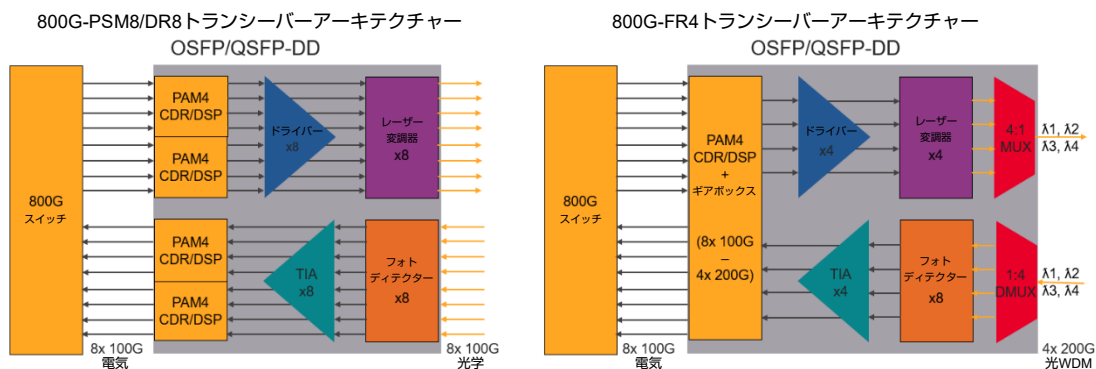


図5. 800G光相互接続の典型的な実装例

2. 参考資料：800Gb/s光モジュールを使用して次世代のクラウドとAIを実現する。 <https://www.800msa.com/documents/white-paper>

800G-FR4の4チャンネルソリューションには、光モジュールに電気信号100 Gbpsの8チャンネルを200 Gbpsの4チャンネルに変換するギアボックスが必要とされるため、新たな課題があります。これは、対応する200 Gbpsレートをサポート可能な4つの光変調器を駆動します。DSP、ドライバー、変調器、TIA等のより高い帯域幅を処理できる新しい光ポート規格や測定器を定義する必要があります。現在、この技術はまだ前研究段階で、変調モード、ボーレート、システム帯域幅、リンクバジェット、ビット・エラー・レート、FEC等のすべての側面が検討、議論されています。

データレート関連用語

400G、さらに現在の800Gへの移行に伴い、イーサネット・データ・レートに使用される様々な用語があります。一般的な用語には、100G/レーン、200G/レーン、100G/ラムダ、112 Gbps、224 Gbps等があります。

第1世代の800Gでは、IEEE 802.3規格のデータレートは53.125 GBd PAM4（106.25 Gbps）です。OIF CEIは38～58 GBd PAM4（76～116 Gbps）のデータレート範囲を指定していますが、インタフェースの記述表題としては112 Gbpsに定着しています。

以後、本書では112 Gbpsと224 Gbpsのシリアルレートの記述をそれぞれ使用して、光と電気双方の第1世代と第2世代の800Gを記述します。

800Gイーサネットの技術的課題

以前の400G電気レーンの速度は、第1世代の800Gイーサネットでは112 Gbpsと2倍となり、第2世代では4倍(224 Gbps)となっています。インタフェースチップ、DSPチップ、パッケージ、コネクタ等のすべてがこれ以上の高速で動作するには、パフォーマンスの改善または新しいデザインが必要とされます。業界内には既に、OIF CEI-112Gファミリーの実装合意やIEEE 802.3ck規格など、それに対応する112 Gbps電気ポートを定義する規格があります。どちらの規格団体も様々な距離または到達距離の電気リンク要件を定義しています。

様々なシナリオに対応した規格がOIFにより初めて公式化されており、それはCEI-112G-MCM/XSR等の短距離チップ、モジュールCEI-112G-VSR向けの中距離チップ、CEI-112G-MR/LR等のバックプレーンや銅線経由の長距離チップにまで及びます。IEEE 802.3ck規格策定グループは、チップ間 (C2C)、チップとモジュール間 (C2M)、バックプレーン (KR)、銅線アセンブリ (CR) の補完的規格を定義しています。

IEEEとOIFの規格はパラメータ定義や測定手動にやや違いはあるものの、相互的に強い影響を受けています。表1に示すように、OIF CEI-112G-VSR、-MRおよび-LRは、IEEE 802.3ckと同等のclauseがあります。

OIF-CEI規格	IEEE 802.3ck規格	最大チャンネル損失
CEI-112G-MCM	—	6 dB
CEI-112G-XSR	—	10 dB
CEI-112G-VSR	100GAUI-1 C2M ³	16 dB
CEI-112G-MR	100GAUI-1 C2C ³	20 dB
CEI-112G-LR	100GBASE-KR1 ³	28-30 dB
—	100GBASE-CR1 ³	24 dB

表1. OIF-CEI-112GとIEEE 802.3ckの規格の対応

CEI-112G-VSRまたは100GAUI-1 C2Mインタフェースを例に挙げると、ホストICからモジュールICへのチャンネル損失はナイキスト周波数で最大16 dBになる場合があります、これはチップパッケージ、PCB材料、およびコネクタには非常に高い性能要件です。この問題を克服するため、新しい光モジュールのパッケージと電気接続が開発されました。

8チャンネルの112 Gbps電気信号伝送をサポート可能な新型のQSFP-DD800パッケージを図6に示します。QSFP-DD800パッケージは400G QSFP-DDパッケージと互換性があります。また、OSFPパッケージも使用可能です。OSFPパッケージはQSFPよりもサイズが大きく、より多くのレーザーを搭載する必要があったり、より高い電力消費をサポートする必要があったりする状況で利点があります。

3. 200GAUI-2、400GAUI-4、200GBASE、400GBASEのバリエーションも含まれます

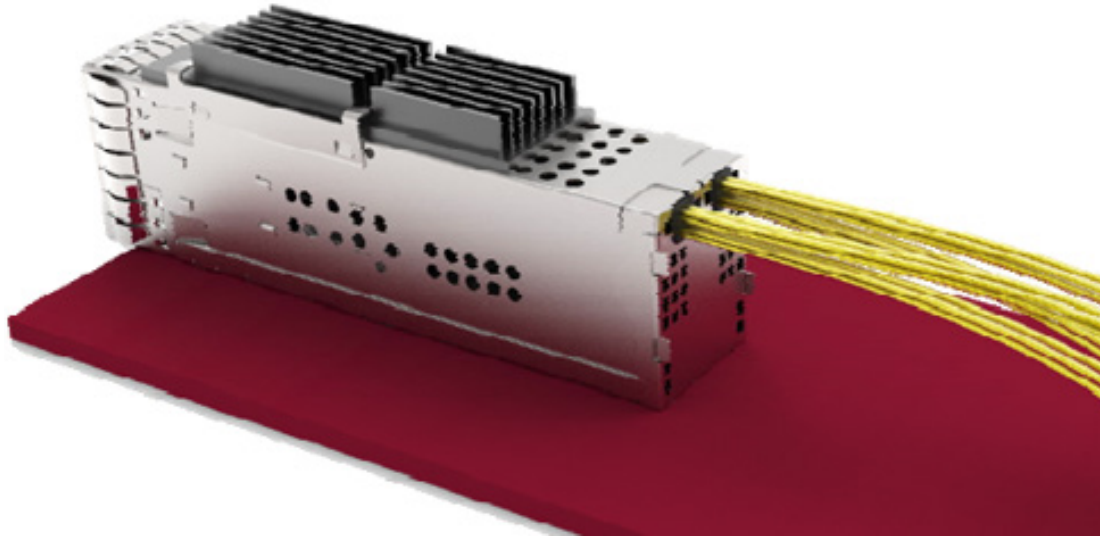


図6. QSFP-DD800パッケージ⁴

また、PCB損失も重要です。これは基本的にボーレートの増加とともに拡大し(400Gモジュールの損失の約1.5~2倍)、これにより高いボーレートで使用可能なチャンネル長が制限されます。一部のデザインは、損失を最小化するため、PCB材料の代わりにケーブルアセンブリを採用しています。図7は、ケーブルアセンブリを使用してスイッチチップと光モジュールスロット (SamtecのFlyoverテクノロジーとMolexのBiPassソリューション) 間の電気接続を実現する2つの例です。これらのケーブルアセンブリの挿入損失は、比較的長いプリント基板トレースの約半分とすることができます。一部のメーカーは同軸ケーブルをスイッチチップに直接接続する方法も研究しています。

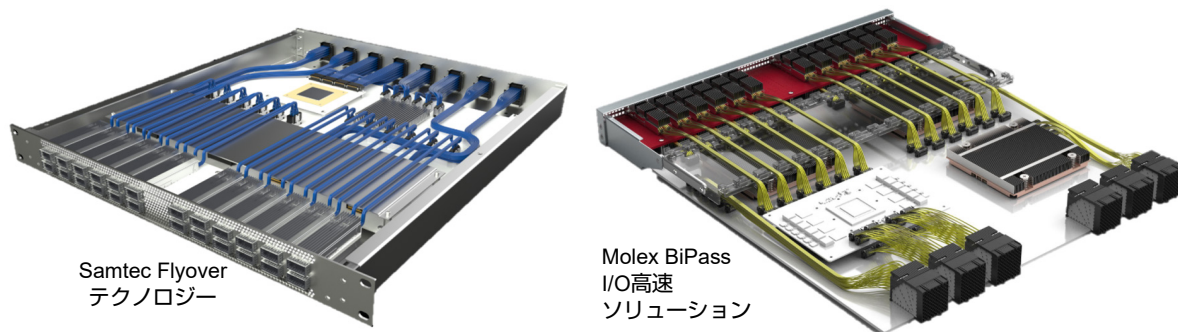


図7. 高速ケーブルアセンブリの例⁵

将来の開発では、チップの電気出力インタフェースに代わって共同パッケージオプティクス (CPO) 技術が使用されます。CPOはスイッチチップとコンピューティングチップでレーザーエンジンを直接統合して、光接続インタフェースを提供します。現時点では、消費電力、放熱性、異種材料の統合など、解決すべき技術的課題が多く残されています。

4. 参考資料：www.qsfp-dd.com

5. 参考資料：www.samtec.com および www.molex.com (Molex, LLC.社が著作権を有する図面を同社のご厚意により掲載)

800Gイーサネット技術のもう一つの重要な要素はスイッチチップのデータスループットです。現在は100 Gbps電気インタフェースをサポートするスイッチング容量25.6Tのチップが使用可能ですが、今後1~2年には次世代の51.2Tチップが登場するでしょう。

光ポートでは、112 Gbps技術が400Gイーサネット用に成熟しており商用化されている（400GBASE-DR4等）ため、第1世代の800Gの技術的実装にほとんど課題はありません。実際の課題は単一チャンネルの224 Gbps光ポート技術にあり、図5に示すようにこの技術では4チャンネルとの間で800Gの接続を達成する必要があります（800G-FR4等）。

現在、224 Gbps伝送の研究は行われ、規格策定グループが始動していますが、公開されている規格はありません。変調方式については、PAM4を継続するか、PAM6（~2.6ビット/シンボル）、PAM8（3ビット/シンボル）、さらにはPAM16（4ビット/シンボル）などの高次方式に移行するかということが1つのテーマとなります。考えられる他の手法は、離散マルチトーン（DMT）またはパルシャル-レスポンスPAM4です。PAM4変調の使用を継続する利点はこの技術に関して多くの研究が行われていることですが、デメリットはリンク全体のデバイスの帯域幅要件が比較的高いことです。224 GbpsのPAM4は112 GBdのポーレートをもつため、変調器、ディテクター、ADC、DACを含む全システムの帯域幅は約80GHzでなければなりません。このポーレートでの主なデザイン上の課題は反射です。反射は不完全な波形が優勢となり、複雑なパワー集約型のイコライゼーションスキームが要求されます。

PAM6とPAM8ではシンボル当たりのビット数を増やすことが可能であるため、224 Gbpsのポーレートはそれぞれ87 GBdと75 GBdに減衰されます。しかし、S/N比はより高くなり、ノイズやジッタへの感度も増大し、追加の信号レベルを処理するためより複雑なSerDesのデザインが要求されます。現在、PAM4技術は業界内での長年の研究開発を経て、商用化が始まったばかりです。112 GBdのPAM変調の使用が最も論理的なステップのように思える一方、224 Gbpsリンクに対応する新しい変調スキームを採用できる可能性があります。技術の選択とテスト/測定器への投資の双方で、高い柔軟性を維持することが必要とされます。

また、高速光インタフェースの使用には、新たなビットエラー比規格、新たなFEC（前方誤り訂正）コード化法、より複雑なレシーバイコライザーの導入も要求されます。100Gイーサネットでは、FECは必須ではありませんが、400GイーサネットとPAM4変調へ移行する場合、光リンクによるゼロエラー比（実際の $1E-12$ または $1E-15$ ）の達成が不可欠となります。 $2.4E-4$ 未満のリンクBERと均一なエラー分布が特徴の400Gイーサネットでは、RS(544,514) リードソロモンFEC（KP4とも呼ぶ）により信頼性が高い通信が保証されます。224 Gbpsの単一チャンネル速度の800Gイーサネットでは、リンクBERはさらに高くなり（ $1E-3$ 等）、より強力なFEC誤り訂正方法、新しいエンコーディング、またはKP4エンコーディングに加え追加のFECの実装が必要とされます。より複雑なFECに起因する追加のオーバーヘッドにより、リンクのエンドツーエンドの遅延やチップの電力消費も増大します。リンクの遅延（レイテンシー）はAIなどの高性能のコンピューティングアプリケーションに重要であるため、最終的なFEC法の選択は、BER、FEC機能、レイテンシー、電力消費の間のトレードオフに基づきます。

システムのビットエラー比に影響するもう一つの非常に重要な要因は、レシーバーのイコライザーの性能です。マルチレベル信号はアイの開口部が小さく、さらに光ファイバー伝送の分散と損失により影響を受けるため、レシーバーでアイを開くには複雑なイコライゼーションが必要です。IEEE 802.3bslにより、400Gイーサネット光コンプライアンステストに対応した5タップFIRフィルターの基準イコライザーが定義されています。ただし、実際には多くの半導体メーカーが実装しているイコライザーはこの基準イコライザー（10または20タップ等）よりもはるかに複雑です。224 Gbps技術の開発ではより複雑なイコライゼーションが必要とされ、システムDSPチップの電力消費も増大します。同時に、光信号パラメータの測定に使用するテスト機器および測定器用により複雑な基準イコライザーを開発しなければなりません。

既に述べた通り、直接変調による単一チャネルの224 Gbps光信号伝送の実現にはまだ多くの技術的課題があります。例えば、変調方法、信号ポーレート、デバイスの帯域幅、S/N比要件、システム・エラー・レート、FEC、イコライザーの機構などです。これらは今後数か月に開発する必要がある主な技術です。

準備

112 Gbpsテストソリューション

IEEE 802.3bsおよび802.3cd規格は、56 Gbpsおよび112 Gbps光インターフェースのトランスミッターとレシーバーのコンプライアンステストをカバーしています。これらの技術は成熟しています。OIF CEI-112GプログラムとIEEE 802.3ckの現在の規格策定により、112 Gbpsの電気レーンのコンプライアンス要件が定義されています。

トランスミッターの測定は高帯域幅オシロスコープで実施し、TDECQ、消光比 (ER)、ジッタ、アイの開口部などの主要パラメータを測定します。レシーバーの測定はストレス印加レシーバー入力テストを用います。ここでは、BERTを使用し、符号間干渉 (ISI)、ジッタ、信号干渉を追加し、ストレス印加信号があるときにレシーバーのBER性能を測定して、「ワーストケース」の信号を生成します。このストレス印加信号は最初に高帯域幅オシロスコープを使用して校正されます。

例えば、スイッチと光モジュールの接続に使用される112 Gbps電気インターフェースのテストを見てみましょう。電気インターフェース規格は、OIF CEI-112G-VSRまたはIEEE 802.3ck 100GAUI-1 C2Mインターフェースで詳細に定義されています。

図8に、ホストのトランスミッター出力信号の品質をテストする典型的なセットアップを示します。IEEE 802.3規格とOIF-CEI規格の1つの相違点はデータレートで、IEEEでは106 Gbps (53 GBd PAM4)、OIF CEIでは最大112 Gbps (56 GBd PAM4) のデータ範囲となっています。トランスミッターの信号品質コンプライアンステストでは、少なくとも50 GHz以上の帯域幅で、高帯域幅サンプリング・オシロスコープまたはリアルタイム・オシロスコープが必要です。オシロスコープは4次ベッセルトムソンまたは4次バターワース応答の基準レシーバー帯域幅をモデル化できるものである必要があります。テストフィクスチャ、すなわちホスト・コンプライアンス・ボード (HCB) とモジュール・コンプライアンス・ボード (MCB) を使用します。これは、長/短チャネルエミュレーションと合わせて、遷移時間が特に急勾配にならないことを意味します。ただし、デバイスまたはチップの性能研究や評価を行う場合は、少なくとも80GHzの帯域幅のオシロスコープの使用が推奨されます。

オシロスコープの帯域幅とノイズフロアの要件に加え、PAM4d電気信号のテストには多くの解析アルゴリズムが関連します。例えば、標準的なループ帯域幅要件 (約4 MHz) に基づいて信号のクロックリカバリーを実施して、データを集めてアイダイヤグラムを作成する必要があります。基準レシーバーのモデルには、規格の定義に基づき、CTLEイコライゼーション (最大12 dBゲイン) と4タップDFEイコライゼーション、最適なCTLEとDFEの設定を決定する手段、アイの開口部を測定するアルゴリズム、SNDRを含める必要があります。

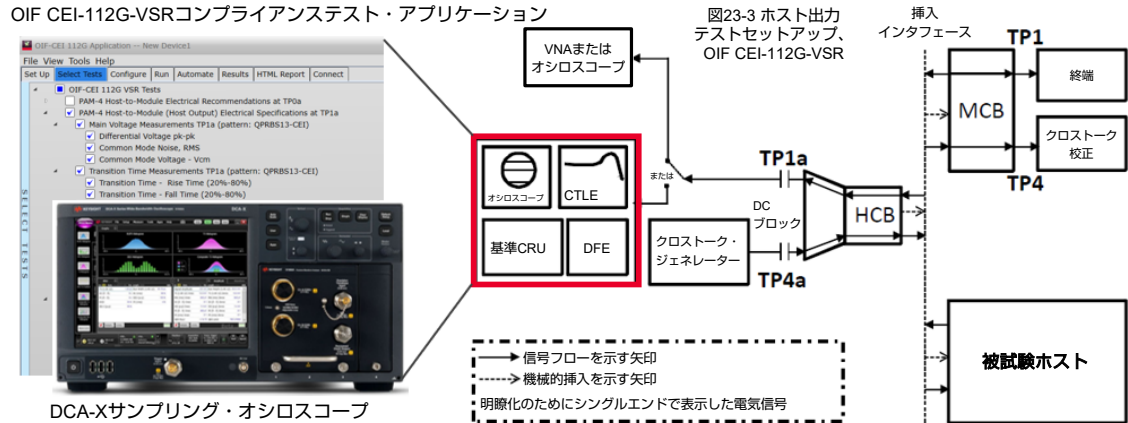


図8. 112 Gbps電気ホスト出力信号の品質テスト

これは、測定またはセットアップのエラーなく、手動で実施する必要がある複雑な手順です。キーサイトは、必要とされる測定のガイドとなり、装置やイコライゼーションの設定を自動化するIEEE 802.3ck規格とOIF-CEI-112G規格の両方に対応したトランスミッター・コンプライアンス・アプリケーションを提供します。これらのコンプライアンスアプリケーションはDCA-XサンプリングおよびUXRシリーズ リアルタイム・オシロスコープでご利用いただけます。

Txテストソリューション	対応規格	到達距離/リンク	対応機種
N1091CKCA	IEEE 802.3ck	C2C、C2M、KR、CR	DCA-X
N109212CA	OIF-CEI-112G	VSR、MR、LR	DCA-X
D90103CKC	IEEE 802.3ck	C2C、C2M、KR、CR	UXRシリーズ
D9050CEIC	OIF-CEI-112G	VSR、MR、LR	UXRシリーズ

図9は、モジュールレシーバーの性能テストの典型的なセットアップです。テストを実行する前に、校正済みのストレス印加信号を生成する必要があります。これは接続済みのMCB-HCBペア経由でBERTパターンジェネレーターを高帯域幅オシロスコープに接続して行います。データ信号に、ノイズ、正弦波ジッタ、有界無相関ジッタ (BUJ)、非有界無相関ガウシアンジッタ (UUGJ)、ISI、クロストークを追加し、ストレス印加信号が規格の定義に適合するまで、垂直アイ開口部 (VEC)、信号対ノイズ+歪み比 (SNDR)、アイの高さ等の信号パラメータを測定します。校正後、被試験モジュールをMCBとBERに接続し、ジッタ耐力テストを実施します。BERは、信号を被試験モジュール経由でBERTアナライザにループするか、または被試験モジュールの内部BERカウンターを使用して、測定できます。

ストレス印加信号の校正は複雑で、ディエンファシス、CTLE、DFEイコライゼーションの設定の最適化が必要とされ、トランスミッターのテスト中に使用される測定と同じ測定を使用します。ストレスパラメータの測定と調整の反復的なプロセスにより、手間と時間のかかる作業です。

IEEE 802.3ck 400GAUI-4コンプライアンステスト・アプリケーション

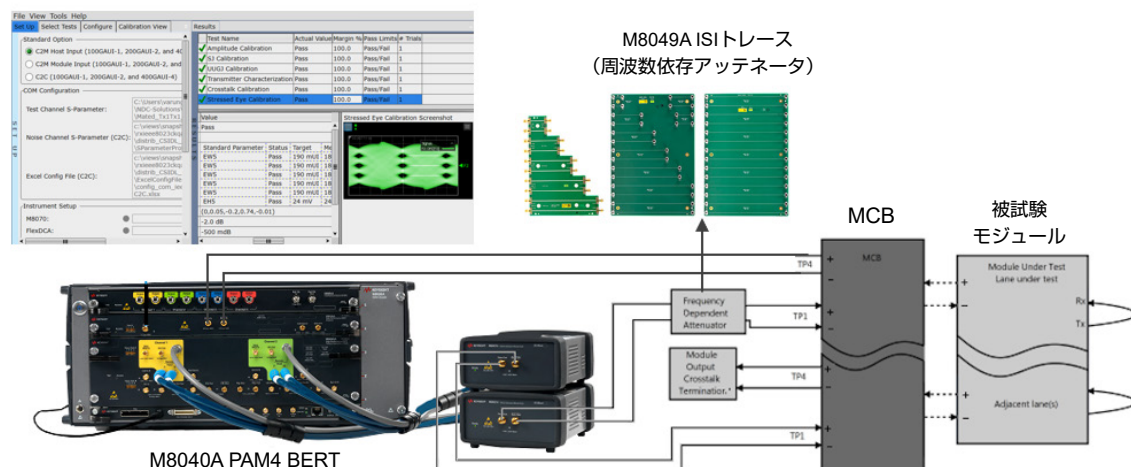


図9. 112 Gbps電気モジュールの入カストレス印加BERテスト

レシーバー・ストレス・テストの効率性は、校正とテストの両方を含むセットアップ全体を自動化することで最適になります。キーサイトはIEEE 802.3ck規格とOIF-CEI-VSR規格の両方に対応したレシーバー・コンプライアンス・アプリケーションを提供しています。これらのアプリケーションが様々な校正ルーチンのステップを経ることで、規格に準拠したストレス印加信号と必要とされるデバイステストが可能となります。アプリケーションノート3121-1220.EN「データセンター向け800Gイーサネットリンクのコンフォーマンステスト」にTxおよびRxテストの詳細を記載しています。

Rxテストソリューション	対応規格	到達距離/リンク
M8091CKPA	IEEE 802.3ck	C2C、C2M
—	OIF-CEI-112G	VSR、MR、LR

224 Gbpsテストソリューション

キーサイト・テクノロジーのテストソリューションのポートフォリオは、シミュレーションからデザイン検証、コンフォーマンステスト、プロトコルテスト、製造までのデザインサイクル全体に対応しています。最新の装置とテストソリューションにより、測定規格が公開される前に最新技術を早期に研究することが可能です。後に規格が進展するにつれ、規格団体との緊密な連携と参加により、キーサイトは規格の策定サイクルと同時にコンポーネントやシステムデザインの適時の特性評価や検証を可能にする市場初のテストソリューションを提供できます。

800Gデザインサイクル

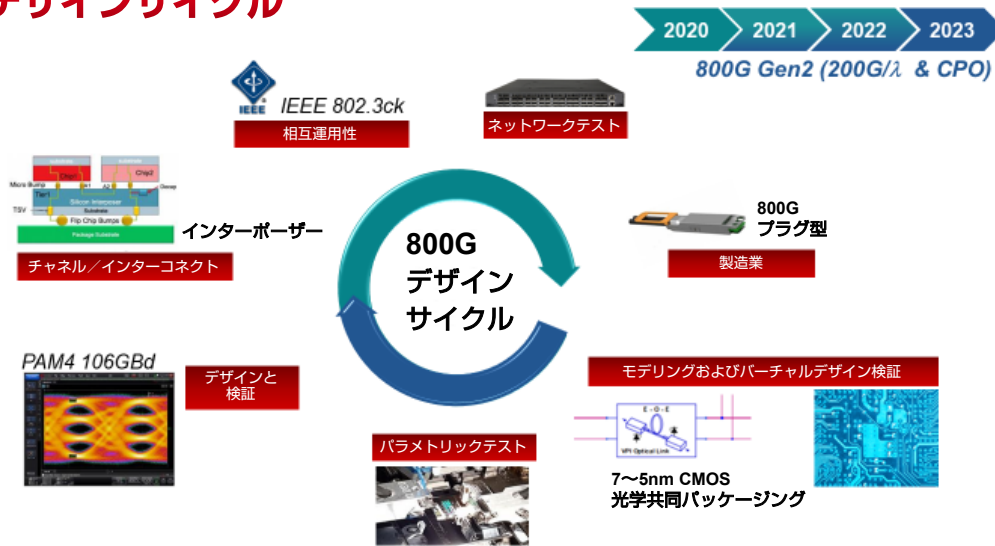


図10. 800Gのデザインサイクル

信号発生

224 Gbps トランシーバーの開発への道は、光変調器、ディテクター、ドライバーアンプ、スイッチチップ等のコンポーネントレベルの性能研究から始まります。これらのデバイスのテストには、振幅（様々なデバイスの電圧要件への適応）、リニアリティ（デバイスの利得圧縮の補正）、変調方式（様々なPAM-n変調スキームの性能差の調査）、ディエンファシス（デバイスパッケージと伝送ラインの帯域幅の補正）等を柔軟に調整できる高速電気テスト用の信号発生器が必要です。

224 Gbpsの研究に使用できる最適な市販装置は、高帯域幅任意波形発生器（AWG）または高性能BERTパターンジェネレーターです。

図11に、チャンネル当たりのサンプリングレートが256 GSa/s、DAC分解能が8ビットのM8199A高帯域幅4チャンネルAWGを示します。AWGは信号生成における最高度の柔軟性を提供しており、4レベル、8レベル、16レベルまたはそれ以上の複雑なシグナリングが容易に達成されます。さらに、任意波形発生器は様々なデジタル波形をダウンロードして様々な信号を生成できるため、ダウンロードした波形を事前に歪ませることによりチャンネルの特性をディエンベディングする高度な機能を提供します。唯一の制限はメモリ機能であるため、AWGの使用は短いプレイ時間または短いデータパターンに制限され、初期の研究に適しています。

図12はM8050A 120 GBd BERTパターンジェネレーターで、2~120 GBdのNRZ、PAM4、PAM6またはPAM8形式のリアルタイムおよびメモリベースのパターンを生成することができます。他の機能には、7タップのディエンファシス、ISI生成、広範なジッタ生成信号劣化があります。BERTは、大きいパターン（PRBS31以上等）の生成、およびレシーバーテスト等の複雑なジッタ・ストレス・カクテルやエラー解析が要求されるアプリケーションにより適しています。

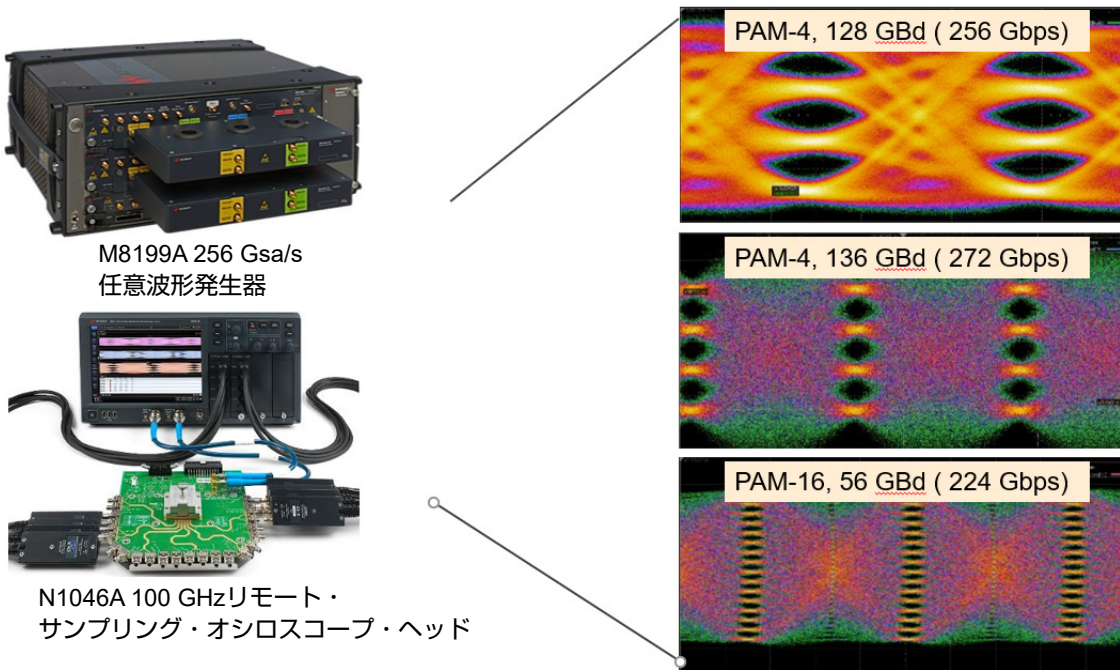


図11. M8199A AWGによる224 Gbps以上の電気信号の生成

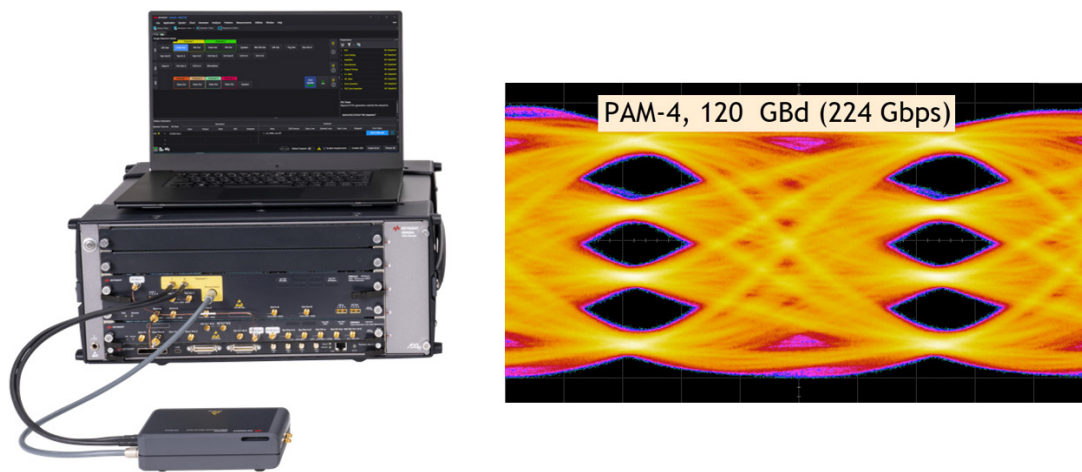


図12. M8050A BERTによる最大224 Gbpsの電気信号の生成

コンポーネントテスト

超高速チップ、相互接続、光電子デバイスの開発と評価において、周波数ドメインのパラメータ（帯域幅、利得、フラットネス、リターンロス等）はデバイスの性能を測定する最も基本的な指標です。これらの高速デジタルアプリケーション用にインターコネクトデザインを最適化する実証された1つの方法はsパラメータです。チップセットが高速化し、相互接続が短くなればなるほど、120GHzのSパラメータ測定への要求は高まってきます。

キーサイトの代表的なベクトル・ネットワーク・アナライザはN5291Aで、最大120 GHzの差動チャンネルの特性評価用の4つのポートで構成されます。反射の低減は確かに高い周波数でのリターンロスを最小化することで可能ですが、マルチドメイン解析は今やチャンネル全体を通して十分に制御されたインピーダンス環境を開発するために欠かせないものとなっています。そのため、物理層テストシステム (PLTS) 等の専門ソフトウェアを120 GHz VNAと併せて使用することにより、小さなジオメトリ向けのタイム・ドメイン・リフレクトメトリー (TDR)、PAM4変調スキーム向けのアイダイアグラム、強力なエンジニアリング研究開発向けの自動イコライゼーションタップの選択等の示唆に富む情報を生成できます。

図13の構成に示すように、N5291A VNAは今日のネットワークやデータセンターで使用されている相互接続に対応した優れた確度と不確かさを提供します。典型的なテストテンプレートは、周波数、時間、アイダイアグラム、RLCG、モード変換、さらに使いやすい形式によるプレエンファシスとイコライゼーションのシミュレーションを含む多くのデータ解析のドメインを組み込むものになるでしょう。これにより、クロストークが減衰し帯域幅が拡大した完全なチャンネル最適化が可能となり、高性能のイーサネットシステムに最適な相互接続の性能が保証されます。

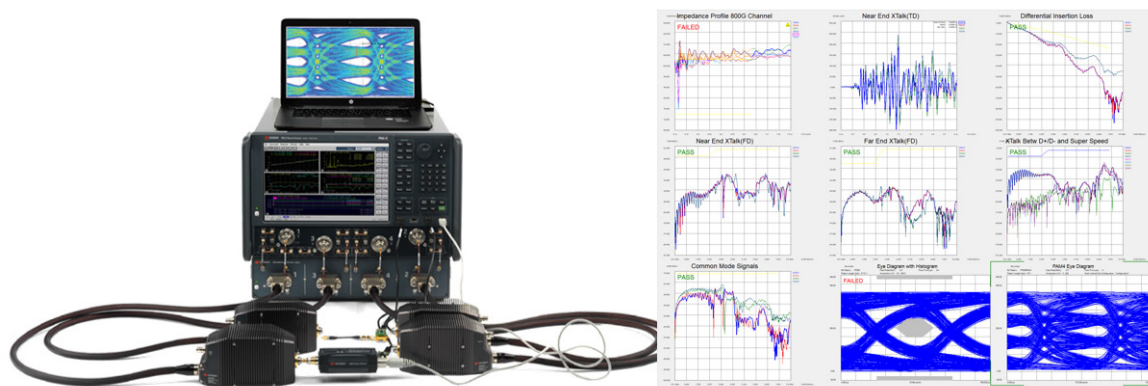


図13. N5291A PNAミリ波システム900 Hz~120 GHzおよび物理層テスト・システム・ソフトウェア

特に光領域では、光変調器や光検出器などの部品に課題があります。変調器の性能はトランスミッターの電気入力からレーザー出力に変換された信号の品質に影響し、フォトディテクターの性能は電気信号に再変換されたレシーバーの光入力の品質に影響します。

ニオブ酸薄膜リチウムまたはシリコン光マイクロリング変調器に基づく現在の最先端の変調器は60 GHzを超える帯域幅を達成できますが、フォトディテクターの帯域幅はさらに高くすることができます。成熟した商用化を達成するには、これらのデバイスがより一貫して安定性があるものになる必要があります。周波数レンジ全体（仕様帯域外の周波数ロールオフ特性を含む）にわたりこれらのデバイスの特性の詳細な検証と評価を実施するには、80 GHz以上の周波数レンジをもつ光コンポーネントアナライザが必要です。

図14は、周波数レンジ110 GHzのN4372E光波コンポーネント・アナライザ・システムです。このシステムはベクトル・ネットワーク・アナライザと最大110 GHzの電気信号を生成/受信するミリ波拡張ヘッドで構成され、光送受信テストヘッドとコントローラーが追加されています。光デバイス・テスト・ソフトウェアが電気および光の校正手順とデバイスのテストを実施します。このソフトウェアは、光-光デバイス（光ファイバー、光増幅器）、電気-光デバイス（レーザー、変調器）、光-電気デバイス（フォトディテクター）、電気-電気デバイス（増幅器、ドライバー）をテストする柔軟性を備えています。



図14. 110 GHz光波コンポーネント・アナライザ

波形テスト

評価の次のステップは、デジタル信号のスティミュラスにおけるデバイスのタイムドメイン性能の評価です。これらのテストには、アイダイヤグラム、解析、遷移時間、ジッタ、光信号のTDECQ、電気信号のVEC等が含まれます。

高速信号の波形テストは、高速サンプリング・オシロスコープまたは高帯域幅リアルタイム・オシロスコープを使用して実施できます。キーサイトDCA-Xサンプリング・オシロスコープには、最大帯域幅100 GHzの電気モジュールと最大帯域幅65 GHzの光モジュールの両方が付属しています。キーサイトのUXRリアルタイム・オシロスコープの電気帯域幅は最大110 GHzで、オプションで最大60 GHzの光帯域幅の光-電気プローブを追加できます。

どのオシロスコープを使用しているかに関わらず、いずれのプラットフォームでも主要測定が可能で、業界標準のDCA-X FlexDCAアプリケーションを利用したUXR上での光測定も実施できます。プラットフォーム間の主な相違点は、クロックリカバリー（SWまたはHW）、周波数応答ロールオフ、ノイズフロア、トリガ方法にあります。

図15に、2つの典型的な高帯域幅光信号測定のソリューションを示します。両方共、224 Gbpsの研究向けに十分な帯域幅があります。224 Gbps規格にPAM4変調（112GBd）が採用され、以前のPAM4光信号規格に従う場合、必要とされる基準レシーバーの帯域幅は信号のボーレートの半分（この場合は56 GHz）で、測定器の60/65GHzの光帯域幅はこれをカバーすることができます。

DCA-Xサンプリング・オシロスコープは、その高帯域幅、低ノイズ、非常に成熟した経路校正技術により、光信号テストの業界標準となっています。従来、リアルタイム・オシロスコープは電気ドメインのみで使用されてきましたが、110 GHzの最大帯域幅、10ビットのADCサンプリング、最大60GHzの光プローブによるリアルタイム・オシロスコープの発展と共に光テストも可能になっています。リアルタイム・オシロスコープの周波数応答特性は、通常、ブリックウォール・ロールオフ応答により平坦です。通常、高速デジタル信号テストではデバイスの測定でオシロスコープが4次ベッセル・トムソン・フィルターをエミュレートする必要があります。これが一般的に意味することは、フィルター応答を生成するために帯域幅がより高いリアルタイム・オシロスコープとサンプリング・オシロスコープを並べて使用する必要があるということです。

サンプリング・オシロスコープ (16ビット)



65 GHz光モジュール

リアルタイム・オシロスコープ (10ビット)



60 GHz光
プローブ

RMSノイズ (特性)	1310 nm	1550 nm
53.125 Gbaud PAM4 TDECQ (26.6 GHz)	18 μ W	22 μ W
53.125 Gbaud PAM4 TDECQ (39.8 GHz)	30 μ W	35 μ W
フィルターなし (60 GHz)	35 μ W	45 μ W
フィルターなし (65 GHz)	80 μ W	95 μ W

65 GHz光モジュール・ノイズフロア

フィルター	帯域幅	Rmsノイズ @ 1310 nm	Rmsノイズ @ 1550 nm
ブリック・ウォール	60 GHz	11.03 μ W	12.43 μ W
4次BT	12.6 GHz	4.52 μ W	5.08 μ W
4次BT	29.5 GHz	7.2 μ W	8.12 μ W

60 GHz光プローブ・ノイズフロア (20 uW/div)

図15. 帯域幅60GHz以上の光信号測定スキーム

もう一つの検討のポイントはノイズ性能です。サンプリング・オシロスコープは優れた一貫性のあるノイズフロア性能がありますが、リアルタイム・オシロスコープのノイズフロアは入力電圧レンジに依存します。図15の表に示す通り、状況によっては、低振幅信号の場合にUXRシリーズのリアルタイム・オシロスコープはDCA-Xサンプリング・オシロスコープよりも低いノイズフロアをもつことがあります。

ビットエラー比テスト

高速データリンクの残りのステップは、ビットエラー比 (BER) またはシンボルエラー比 (SER) の測定によるリンク品質の検証です。基本的に、高速デジタルコンポーネントおよびリンクのBERの測定に使用可能な3つの方法があります。

従来の方法では、キーサイトM8040A 64 GBdまたはM8050A 120 GBd高性能BERT等のビット・エラー・テスター (BERT) を使用します。BERTがテスト信号を生成してデバイスに送信した後、信号は伝送経路を経由してBERTアナライザにループバックされ、BER測定やエラー解析に使用されます。ループバック経路で信号にこれ以上のエラーが発生しない場合、この方法に問題はありませぬ。PAM-nシグナリングとより高いポーレートでは、伝送はエラーフリーではなく、ループバック経路がBER測定に影響を与え始めます。112 Gbpsでは、依然としてこの手法を使用可能なケースがあります。

あるいは、多くの場合、デバイスはBERテストを促進するために内部エラーカウンターを装備しているため、BERを被試験デバイスから直接読み取ることもできます。これは益々一般的になってはいますが、新規チップ開発の初期段階では、内部エラーカウンターの機能は使用できない場合があります。

3つ目の方法は、高速リアルタイム・オシロスコープでデジタル信号を捕捉し、ソフトウェアのデコーディングによりBERを決定して、予測パターンと比較する方法です。この手法は使用可能な捕捉メモリと時間により制限され、1E-12等のBERを達成することに実用性はありませぬ。一方、PAM4シグナリングとFECの導入により、112 Gbpsリンクのraw BER目標は2.4E-4 (光) ~1E-5または1E-6 (電気) の範囲にあり、リアルタイム・オシロスコープによるBER測定を可能にしています。224 Gbpsの非補正BERのターゲットはさらに高くなると予想されます (光リンクに対して)

1E-3等)。エラーディテクターとしてのリアルタイム・オシロスコープの追加機能は、FECシグナリングを含むビットストリームの波形を捕捉および解析し、信号波形内のビットエラーの位置を相関させる機能です。

この方法の例は、M8070B BERT制御ソフトウェアでリアルタイム・エラー・ディテクター（RED）として設定し、最大帯域幅110 GHzで使用可能なキーサイトInfiniium UXRシリーズ・オシロスコープです。このオシロスコープは入力信号を捕捉してデジタル形式に変換します。これらのビットはUSB/LAN上で、M8070B BERT制御ソフトウェアを実行する制御PCに送信されます。次にM8070Bがこのデータを予想パターンと比較し、BER測定を生成します。また、リアルタイム・オシロスコープのすべての必要とされるセットアップと制御もM8070Bソフトウェアにより処理されます。

BERT PG（M8050A等）またはAWG（M8199A等）のパターンジェネレーターにより、トランスミッターのイコライゼーション（ディエンファシス機能、レイズド・コサイン・フィルタ／ナイキストフィルタ、エラー挿入等）の点で柔軟性が向上し、リアルタイム・オシロスコープにより、高感度、柔軟なクロックリカバリー、さらにFFE、DFE、CTLE等のエンハンスド自動レシーバリーコライゼーション機能が提供されます。このソリューションは、NRZ、PAM4、PAM6、PAM8等の複数の変調方式を、最大240 Gbps（120 GBd）のデータレートで処理できます。

図16に、M8050A 120 GBd BERTパターンジェネレーターとUXRリアルタイム・オシロスコープをエラー検出エンジンとする M8050A 120 GBd BERTパターンジェネレーターによる224 GbpsでのPAM4ビット・エラー・レート・テストの例を示します。表は、複数のBER測定と様々な目標BER限界の信頼度レベルを示しています。これらの例では、BERを測定する前に5タップFFEをアナライザに適用しています。1E-5目標BERの例では、5.2E+5ビットの測定が約1.2秒で実行され、結果の信頼度レベルは99%を超えています。イコライゼーションの複雑さが増大するにつれ、対応するBERの信頼度レベルの達成に必要なとされるテスト時間も増大します。実用上、UXRオシロスコープに基づくリアルタイム・エラー・ディテクター・テスト・ソリューションは1E-6のビットエラー比までのビット・エラー・レート測定に対応しています。



目標BER	測定時間（秒）	受信ビット数	BER信頼度レベル
1E-4	0.39	2.62E5	99.9999 %
1E-5	1.19	5.24E5	99.46 %
1E-6	7.53	3.14E6	95.68 %
1E-7	75.03	3.01E7	95.08 %

図16. リアルタイム・オシロスコープによるPAM4のビット・エラー・レート・テスト

テストソリューションの概要

図17は、800Gイーサネットモジュールまたはチップインタフェースと、試験ソリューションの概要です。現在、キーサイト・テクノロジーは112 Gbpsコンプライアンス・テスト・ソリューションを提供しています。これは、DCA-Xサンプリング・オシロスコープ、UXRリアルタイム・オシロスコープ、電気および光インタフェースの両方に対応したM8040Aビット・エラー・レート・テスターに基づいています。112 GbpsのOIF-CEIおよびIEEE 802.3規格はまだ開発段階であり、キーサイト・コンプライアンス・ソリューションは、規格策定サイクルを通じて適時のテストソリューションを提供するため、その進捗を追跡しています。

224G光ポートおよび電気ポートの研究は初期段階にあり、技術的導入の詳細については業界内で合意に達していません。キーサイト・テクノロジーは224Gの研究に使用される試験装置の主要メーカーで、M8050A 120 GBd BERTおよび224 Gbps電気PAM-n信号（ならびにDMTおよびパーシャル-レスポンスPAM4等の他の変調スキーム）を発生させるM8199A高速任意波形発生器、224 Gbpsトランスミッターの波形の品質とパラメータを測定する最大110 GHz電気/65 GHz光帯域幅のDCA-XおよびUXRオシロスコープに使用されています。

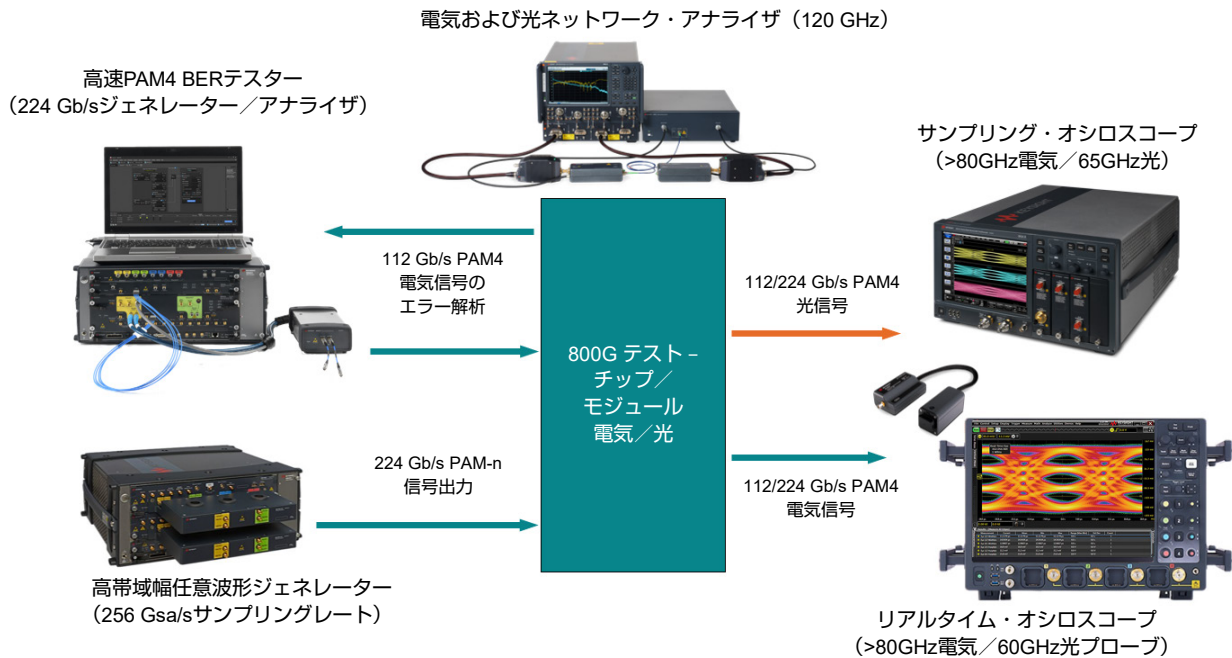


図17. 800G装置/光モジュール/チップのテストプラン

集積回路、相互接続、光変調器、224 Gbps用ディテクター等の主要コンポーネントの性能評価には、最大周波数レンジ120 GHzの電気PNAおよび光波コンポーネント・アナライザを使用できます。また、電気または光ポートのビットエラー比のテスト（主にTOSAまたはROSAの主要光電子デバイス用）には、リアルタイム・オシロスコープをベースとしたエラー・アナライザ・ソリューションを使用できます。

詳細については、www.keysight.com/find/800Gをご覧ください



キーサイト・テクノロジーの製品、アプリケーション、サービスの詳細については、
ウェブサイトをご覧ください：www.keysight.co.jp

本書の情報は、予告なしに変更されることがあります。© Keysight Technologies, 2021 - 2022,
Published in Japan, August 26, 2022, 3121-1425.JA