

データセンタ光インターコネクトの実装形態と光電融合技術

古河電気工業株式会社 フォトニクス研究所 フェロー 那須 秀行

はじめに

情報通信および演算が、われわれの生活を支える基盤となって久しい。業務においても、日常生活においても、情報網を介してデータセンタにアクセスして、メール、web 閲覧、動画視聴等のサービスを常に使用している。更に ChatGPT 等の新しいサービスが提供されるようになり、様々なシーンで活用されている。そのために、データセンタの中に AI および ML (Machine Learning) が導入されており、その演算性能の向上が行われている。演算能力の向上によって信号速度は著しく向上しており、スイッチ ASIC の 1 レーン当たりの信号速度は 200Gb/s が導入されつつある。これに歩調を合わせ、光トランシーバの 1 レーン当たりの信号速度も 200Gb/s が導入されつつあり、伝送容量が増大している。

演算能力の向上と伝送容量の増大に伴い、データセンタにおける消費電力は年々増加しており、計算上では、2030年に、データセンタの消費電力は、世界の総発電量の10%に到達すると予想されている¹⁾。更に、装置レベルでは、消費電力の増大により、発熱量も増大し、冷却能力を向上することが求められている。従来か

ら安価かつ保守性の高い空冷システムが用いられてきたが、発熱量の増大によって水冷システムの導入が進んでいる。

データセンタの通信ネットワークにおいてスイッチ ASIC (Application Specific Integrated Circuit) は大きな役割を果たし、その帯域は著しく向上してきている。2025年6月3日に Broadcom 社はプロセスノード 3 nm の CMOS を採用した 102.4Tb/s の帯域を有する Thomahawk6 をリリースした²⁾。図 1 は、CMOS プロセスノードおよびスイッチ ASIC の容量の推移を示している。図中の実線はこれまでの実績であり、点線は予想を示している。これまで

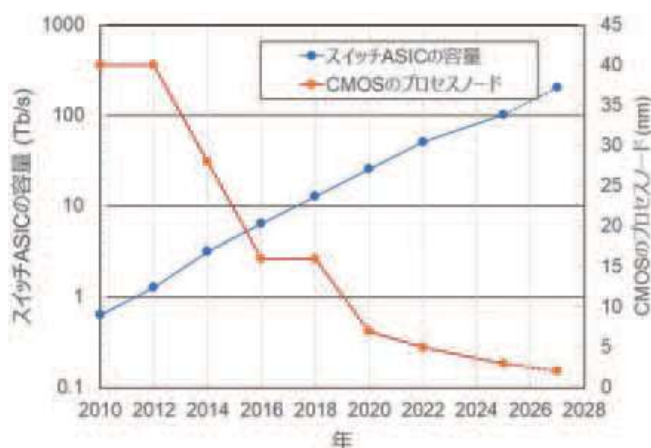


図1 CMOS プロセスノードおよびスイッチ ASIC の容量の推移

CMOS ノードが小さくなるに連れて、スイッチ ASIC の帯域は概ね 2 年で倍になる増加傾向を示している。このトレンドに従うと 2027 年頃には 2.1nm のプロセスノードを用いて 204.8Tb/s がリリースされる見込みである。102.4Tb/s スイッチ ASIC の電気インターフェースのレーン当たりの速度は 200Gb/s に達しており、204.8Tb/s ではレーン当たり 400Gb/s 以上になると予想される。

スイッチ ASIC が広帯域化することで、レーン当たりの伝送速度は高速化していく一方で、消費電力の増加が懸念されている。図 2 は、ネットワークスイッチの帯域と消費電力の関係を示している³⁾。2010 年から 2022 年にかけて、帯域は 640Gb/s から 51.2Tb/s に向上した。これに伴い、スイッチ ASIC の消費電力は 22 倍に増えたことが分かる。この間の消費電力の内訳は、ASIC コアパワーが 8 倍、ASIC コアパワーが 25 倍、光リンクパワー (Optics SerDes Power と Optics Other Power の和) が 26 倍、システムファンパワーが 11 倍となっている。光リンクパワーの増加率が最も高く、全消費電力の半分近くを占めるようになってきている。102.4Tb/s スイッチ ASIC を導入した光リンクを採用することで、光リンクパワーの比率は増大することが予想される。今後、益々光リンクパワーの抑制は本質的な課題となる。

そこで、光電融合技術を導入することで、伝

送容量を増大させつつ、低消費電力を実現することが期待されている。光電融合技術の最初の段階はパッケージレベルで光デバイスと電子デバイスを一体化する CPO (Co-Packaged Optics) と位置づけられており、現在、導入が進みつつある。

本稿では、データセンタにおける光インターコネクトの実装形態の観点から、光電融合へ移行する必要性について述べ、続いて光電融合技術のロードマップについて述べる。更に、光電融合の第一段階である CPO について最近の動向を踏まえ解説する。

光インターコネクトの実装形態の変遷

光トランシーバの配置に基づいて定義したデータセンタにおける光インターコネクトの実装形態のロードマップを図 3 に示す⁴⁾。ここで、青い○で囲った箇所は、スイッチ ASIC や GPU (Graphic Processing Unit) 等の大規模集積回路 LSI (Large Scale Integrated Circuit) であり、赤○で囲った箇所は、光電気変換部あるいは光エンジンと呼ばれる。第 1 フェーズは、ボードエッジ実装と定義され、フロントパネルに標準化されたスモールフォームファクタ SFF (Small Form Factor) を使用したプラグブル光トランシーバを配置する形態を有する。スイッチ ASIC とプラグブル光トランシーバの

間の電気伝送は、マザーボード上に形成された電気配線を用いて行うため、長い電気伝送路によって帯域制限を受け、高速化するには波形整形、波形再生の電子デバイスが必要であり、それを使用することによって消費電力が増大する。

第 2 フェーズは、LSI 近傍実装と定義しており、光トランシーバをマザーボー

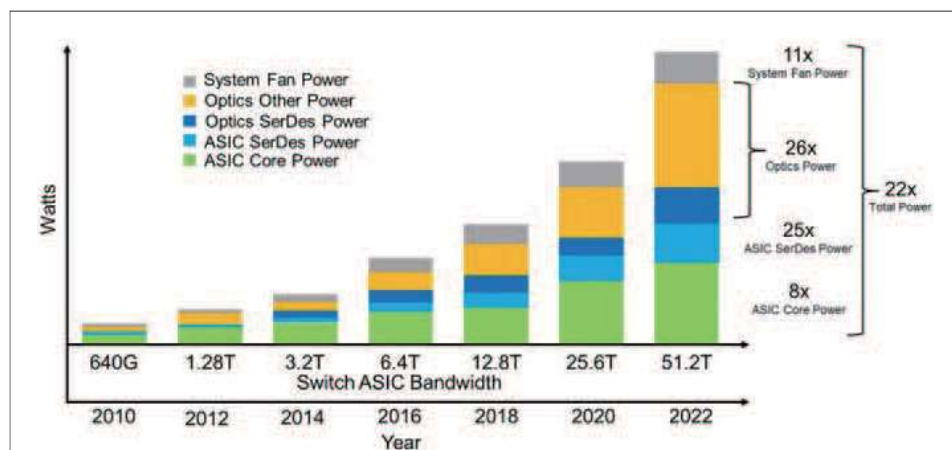


図2 ネットワークスイッチの帯域と消費電力の関係