

光技術と電気技術の融合によるデータセンターネットワーク

Opt-Electronic Data Center Network

大下 裕一
Yuichi Ohsita

西島 孝通
Takamichi Nishijima

小泉 佑揮
Yuki Koizumi

村田 正幸
Masayuki Murata

大阪大学 大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

1 はじめに

近年、データセンターを用いた多数のサービスが開発されるようになっており、データセンター内では著しく大きいデータの処理が必要となってきた。大きなデータの処理は、データセンター内の複数のサーバ間で連携して行われており、データセンターの処理性能を確保するためには、広帯域・低遅延でサーバ間を接続するネットワークが必要となる。その一方、データセンターで処理されるデータが大規模になるにつれ、データセンターの消費電力は大きくなってきている。データセンター内の消費電力のうち、ネットワークが占める割合も少なくなく、データセンターの消費電力を削減するためには、サーバや空調の低消費電力化のみではなく、ネットワーク低消費電力化も必須である。

この問題に対して、我々の研究プロジェクトでは、光通信技術と電気技術を融合した光電子融合型パケットルータとその応用技術の開発を進めている [1]。本稿では、我々の研究プロジェクトのうち、光電子融合型パケットルータのデータセンターへの応用について紹介する。

2 光電子融合型パケットルータ

図1に光電子融合型パケットルータを示す。光電子融合型パケットルータは、光ポートと電気ポートの二種類のポートを持ち、光ポートは、他の光電子融合型パケットルータとの接続に用いられ、電気ポートは各サーバラック内に設置された電気スイッチとの接続に用いられる。

各サーバラックから電気ポートを介して光電子融合型パケットルータに流入したパケットは、共有電気バッファに蓄えられたのち、光パケットに変換された上で送出される。また、光電子融合型パケットルータに直接接続しているサーバラック宛のパケットは、光パケットから電気パケットに変換した上で、電気バッファに蓄えられたのち、宛先サーバラックに送出される。

他の光電子融合型パケットルータから到着した他の光電子融合型パケットルータ宛の光パケットは、パケット内のラベルに合わせて転送先ポートが決定され、転送先ポートが空いていれば、電気に変換されることなく、宛先光ポートを介して転送される。転送先のポートが空いていない場合は、光/電気変換を行った上で共有バッファに一旦保存したのち、再び電気/光変換を行った上で転送を試みる。

この光電子融合型パケットルータは以下の利点を持つ。(1) パケットの衝突が発生しない場合は、光/電気変換が不要で、光パケットをそのまま中継が可能であり、低消

費電力・低遅延・広帯域の通信が可能である、(2) パケットの衝突が発生した場合であっても、電気バッファに一旦保存したのち、再度転送を試みる事が可能であるため、パケットの衝突を避けるような集中制御は不要であり、大規模なネットワークへの拡張が可能である。

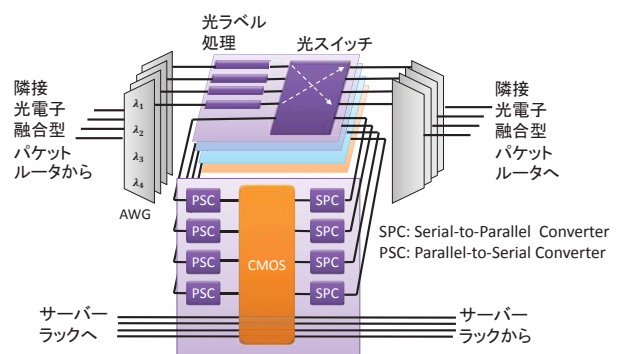


図1 光電子融合型パケットルータ

3 光電子融合型パケットルータを用いたデータセンターネットワーク

光電子融合型パケットルータ間は、低遅延で広帯域の通信が可能である。そのため、光電子融合型パケットルータは、データセンター内のネットワークのコアに配置し、各光電子融合型パケットルータが多数のサーバラックからの通信を束ねて転送するネットワーク構造が適切だと考えられる。その一方、光電子融合型パケットルータが故障したとしても、サーバラック間の接続性を確保することができるネットワーク構造が必要とされる。そのため、光電子融合型パケットルータを用いたデータセンターネットワークの構造としては、図2に示すように、光電子融合型パケットルータを複数台相互接続してデータセンター内コアネットワークを構築し、各サーバラックからは複数台の光電子融合型パケットルータに接続するという構成が適切だと考えられる。

4 光電子融合型パケットルータネットワーク上の低消費電力化制御

データセンター内の通信需要は時々刻々変化し、発生した通信を収容するのに必要な機器のみ電源を投入し、それ以外の機器の電源を落とすことにより、低消費電力化を行うことができる。光電子融合型パケットルータでは、バッファの電源を落とすとしても光パケットの中継は可能であり、ルータの電源のオン・オフのみではなく、バッファについても不要な箇所の電源を落とすこと

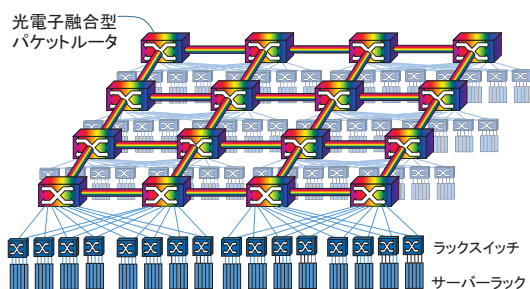


図2 光電子融合型パケットルータを用いたネットワークによる低消費電力化が可能である。

以下の2つの条件の両方を満たす光電子融合型パケットルータは、バッファの電源を落とすことが可能である。(1) サーバラックからの流入トラヒックやサーバラックへの流出トラヒックが存在しないこと。光電子融合型パケットルータでは、サーバラックからの流入トラヒックやサーバラックへの流出トラヒックは、一度バッファに蓄えられたのちに送出される。そのため、サーバラックからの流入トラヒックやサーバラックへの流出トラヒックが存在する光電子融合型パケットルータはバッファの電源を落とすことができない。(2) 異なる入力ポートから流入したトラヒックは必ず異なる宛先ポートから出力されること。異なる入力ポートから流入したトラヒックが同一の宛先光ポートを利用する場合は、パケットの衝突が発生する可能性がある。この場合、パケットロスを防ぐためのバッファが必要となる。

我々は上記の2つの条件を満たすように、トラヒックの收容を行う手法について検討を行っている。以降、トラヒック收容方法について紹介する。

4.1 低消費電力なトラヒック收容手法

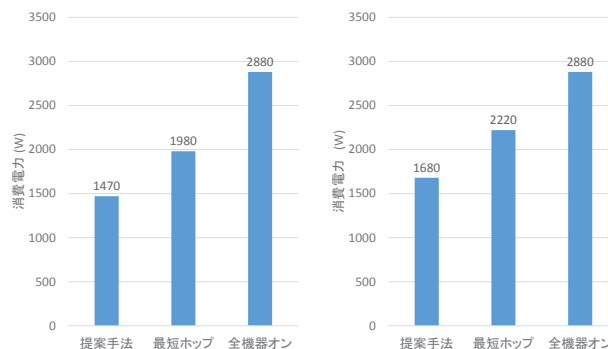
本手法では、短時間でトラヒックの收容先を決めるため、最適化問題を用いず、物理ネットワーク上のホップ数が短いトラヒックから順に收容先の経路を確定する。その際に、電源を投入するバッファを最小とするため、バッファの電源投入が必要な際には、その後に收容先が決定されるトラヒックにも使われる可能性の高いバッファのみ電源を投入する。このバッファの利用される可能性を調べる指標として、以下のようにノード n の再利用容易性を定義する。

$$\sum_{s,d} \frac{N_{s,n,d} \lambda_{s,d}}{N_{s,d}}$$

ここで、 $\lambda_{s,d}$ は $s-d$ 間のトラヒック量、 $N_{s,d}$ は $s-d$ 間の最短ホップ経路数、 $N_{s,n,d}$ は $s-d$ 間の最短ホップ経路のうち、 n を経由するもの数である。

各送信元宛先間のトラヒックの收容先の手順は、以下のように決定される。

1. 送信元・宛先間の経路の候補 P を取得する
2. 経路の候補 P のうち、新たに電源の投入の必要がある機器・バッファが最小の候補を選択する



(a) トラヒックパターン 1 (b) トラヒックパターン 2

図3 消費電力の比較

3. 新たに電源の投入の必要がある機器・バッファの数が最小の候補が複数ある場合、再利用容易性を計算し、再利用容易性が最も大きい経路に收容する。

4.2 評価

図2に示したネットワーク構造において、トラヒックの收容を試み、その收容に必要な電力を評価した。比較対象として、(1) 全機器の電源を投入した場合、(2) 各サーバラック間の経路として最短ホップ経路を用い、不要なルータ・バッファの電源のみオフとした場合を用いた。また、消費電力のモデルとして、各光電子融合型パケットルータのバッファの電源を落とした場合の消費電力を60W、バッファの電源を投入した場合の消費電力を180Wとした。トラヒックパターンとして、ホップ数の近いトラヒックが多い場合(トラヒックパターン1)、ホップ数の多いトラヒックが多い場合(トラヒックパターン2)の2つの状況で評価を行った。

図3に結果を示す。図より、いずれのトラヒックパターンであっても、提案手法の消費電力を考慮したトラヒック收容により、データセンターネットワークの消費電力を削減できることがわかる。

5 まとめ

本稿では、光通信技術と電気技術を融合した光電子融合型パケットルータを用いたデータセンターネットワークと光電子融合型パケットルータの特性を考慮した消費電力なトラヒック收容を実現する技術に関する我々の研究を紹介した。今後、さらに大規模なネットワークにも適用可能なように手法の改善を行う予定である。

謝辞

本研究は情報通信研究機構(NICT)の委託研究「高機能光電子融合型パケットルータ基盤技術の研究開発」の成果による。

参考文献

- [1] “低消費電力・低遅延高機能光電子融合型パケットルータに必要な基盤技術の研究開発 及び低消費電力・低遅延高機能光電子融合型パケットルータの応用技術の研究開発.” http://www.nict.go.jp/collabo/commission/k_151ai.html.