

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設計

大下 裕一[†] 村田 正幸[†]

[†]大阪大学 大学院情報科学研究科

E-mail: †{y-ohsita,murata}@ist.osaka-u.ac.jp

あらまし データセンターネットワークへの要求として、高通信性能に加えて省電力化が求められている。低消費電力で高通信性能なネットワークを実現するため、我々は、光スイッチング技術と電気バッファを利用した光電子融合型パケットルータの開発を進めている。本稿では、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に適したネットワーク構造について議論を行う。本稿では、光電子融合型パケットルータが提供する広帯域・低遅延の通信特性を利用するため、光電子融合型パケットルータはデータセンター内のコアネットワークに配置し、複数台のサーバラックからのトラフィックを各光電子融合型パケットルータで集約する形態を考える。また、本構造において、各光電子融合型パケットルータの故障によりデータセンターネットワークから切り離されるサーバラックの存在を避けるため、各サーバラックも複数台の光電子融合型パケットルータに接続する。この条件において、各サーバラックからの複数のリンクを利用しつつ、光電子融合型パケットルータにおけるパケット衝突を避けるルーティングを考慮したネットワーク構造構築方法を提案する。シミュレーション評価により、提案されたネットワーク構造が従来型のデータセンターネットワークトポロジと比べ、より多くのサーバ間通信を低遅延で収容できることを明らかにする。
キーワード データセンターネットワーク、トポロジ、光電子融合型パケットルータ

Design of Data Center Network Topology using Opt-Electronic Packet Switches

Yuichi OHSITA[†] and Masayuki MURATA[†]

[†] Graduate School of Information Science and Technology, Osaka University

E-mail: †{y-ohsita,murata}@ist.osaka-u.ac.jp

Abstract The data center network with high energy efficiency and high communication performance is required. For high communication performance and energy efficiency, we have developed an opt-electronic packet router. In this paper, we propose a method to construct the suitable topology for the data center network with the opt-electronic packet routers. We consider the network topology where the core network is constructed of multiple opt-electronic routers, to utilize the large bandwidth of opt-hybrid routers. Each server rack is connected to multiple opt-electronic routers to keep the connectivity between server racks even when any opt-hybrid router fails. Our method to construct the data center network topology utilizes such multiple connection from each server rack. In addition, our method considers the routing strategy for the opt-electronic routers, because the routing to avoids collisions decrease the number of conversion between optical packets and electronic packets, which decreases the latency between server racks. We evaluate our network topology and demonstrate that our network topology accommodate more traffic with smaller latency, compared with the hyper cube based topology.

Key words Data Center Network, Topology, Opt-Electronic Router

1. はじめに

近年、データセンターネットワークへの要求として、高通信性能に加えて省電力化が求められている。データセンター内では、複数のサーバが連携して一つの大きなデータの処理を行っており、サーバ間を結ぶデータセンターネットワークはデータ

センターの処理性能に大きな影響を与える。そのため、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、ネットワーク機器の消費電力がデータセンターの消費電力全体に占める割合が大きくなっており、データセンターの消費電力を削減するためにネットワーク低消費電力化も必須である。

低消費電力・低遅延・広帯域なデータセンターネットワークを構築する一つの手法として、光通信技術を有効活用することが考えられる。光通信技術を用いることにより、消費電力の小さい光デバイスにおいて光信号を電気変換することなく直接スイッチング可能になり、低消費電力・低遅延な通信が可能になる。また、光波長多重通信により電気と比べて広帯域を実現できる。光通信技術を用いてデータセンターネットワークを構築する方法として、光パススイッチを用いる手法 [1] やバッファレスな光パケットスイッチを用いる手法 [2] が検討されている。しかし、パス設定に時間がかかるためトラヒックパタンの変化が激しい環境への適用が困難である問題や、パケットの衝突を回避するための制御が必要でありネットワークの大規模化が困難であるという問題がある。一方、光信号のままバッファリング可能な光パケットスイッチには大容量の光バッファが必要であり、ネットワーク機器の実現が難しいという問題がある。これらの問題を解決するため、我々は、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開発している [3]。光電子融合型パケットルータは、他のルータと接続に用いる光ポートと、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る電気ポートや電気/光変換器、光/電気変換器、電気バッファを持つ。光電子融合型パケットルータでは、光パケットを中継する場合、パケットの衝突が発生しない限り、変換器やバッファを用いずに消費電力の小さい光デバイスのみを用いてスイッチング処理を行なえる。一方、パケットの衝突が発生した場合でも、パケットを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り、パケットの衝突を回避するための制御は不要である。そのため、光電子融合型パケットルータを用いることにより、パケットの衝突を回避するための制御を行なうことなく、光通信技術の利点である低消費電力・低遅延を実現したデータセンターネットワークの構築が可能である。

本稿では、光電子融合型パケットルータの利点を生かした、ネットワークトポロジについて議論を行う。光電子融合型パケットルータは、電気パケットを光パケットに変換したのち、パケットの衝突が発生しない限り、電気信号に変換されることなく、低遅延で宛先近くの光電子融合型パケットルータまで届けられる。この利点を生かすためには、光電子融合型パケットルータはネットワークのコアに配置することが望ましい。その上で、光電子融合型パケットルータで構築されたコアネットワークでは、パケットの衝突を可能な限り防ぐようなルーティングを行い、電気信号に変換することなく、通信の中継を行うことができるようにしておくことが望まれる。そのため、光電子融合型パケットルータで構成されたコアネットワークにおいても、上記のルーティングを考慮し、より多くのトラヒックを収容できる構成が適切である。その一方、光電子融合型パケットルータの広帯域を有効に生かすために、各光電子融合型パケットルータは、複数のサーバラックと接続するのが望ましいが、その場合、一台の光電子融合型パケットルータの故障により、多数のサーバラックがネットワークから切り離されてしまうという問題が生じる。そこで、各サーバラックも複数台の光電子融合型パケットルータに接続する。この各サーバラックからの複数本のリンクは、耐故障性の確保のみならず、サーバラック間の低遅延な接続にも役立てることができる。

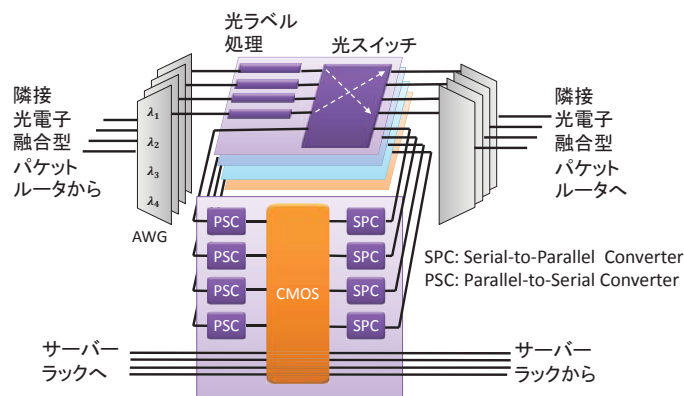


図1 光電子融合型パケットルータ

本稿では、上記の点を考慮した、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の設計法を提案する。そして、その設計法により最適なネットワーク構成が構築できることを示し、構築されたネットワークが低遅延でサーバラック間の通信を収容することができることを示す。

2. 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概要

2.1 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。

2.2 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概要

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

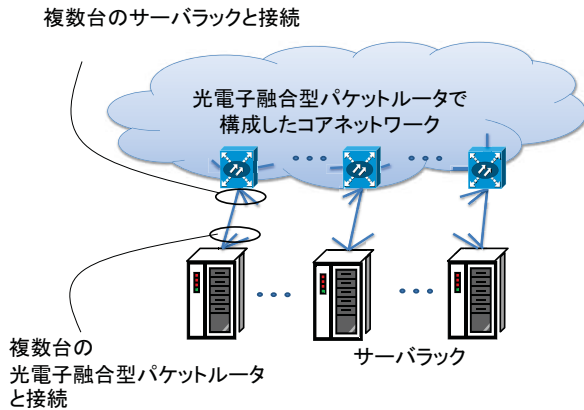


図2 光電子融合型パケットルータを用いたデータセンターネットワークの概要

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である。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。

2.3 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における経路制御

各光電子融合ルータでは、転送先のポートに空きがあった場合であれば、瞬時に転送先ポートにトラヒックを送出できるが、ポートが空いておらず、いったんバッファに保存する場合は、光から電気信号への変換、メモリの読み書き、電気から光への変換といった処理を行う必要があり、遅延が増大する。そこで、光電子融合ルータでは、空きポートがあれば優先的に転送先として選択することが望ましい。そのため、本稿では、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における経路制御として、光電子融合型パケットルータが宛先に応じて複数の転送先候補を保持し、そのうち転送先候補のうち、空きポートがあれば、当該ポートへパケットを転送するという方式を採用する。ただし、データセンターにおいて、サーバラック数は膨大となるので、光電子融合型パケットルータにおいて、各サーバラックに対する複数の転送先候補を保持することは難しい。それに対して、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収容することができるため、データセンター内の光電子融合型パケットルータは、サーバラックに比べ、著しく少ない数で十分である。そこで、光電子融合型パケットルータでは、サーバラック単位の宛先表を管理するのではなく、各光電子融合型パケットルータ宛のパケットに対する転送先の候補を保持するものとする。

本経路制御方式では、サーバラック・光電子融合型パケットルータの接続構成を保持するトポロジ管理サーバを導入する。トポロジ管理サーバでは、各サーバのアドレスと、そのサーバ

が設置されているサーバラックが接続している光電子融合型パケットルータを対応付ける情報と、光電子融合型パケットルータ間のホップ数の情報を保持している。そして、トポロジ管理サーバでは、送信元サーバ・宛先サーバのアドレスを含む問い合わせに対して、当該サーバが含まれるサーバラックが接続している光電子融合型パケットルータの組み合わせのうち、もっともホップ数が小さい組み合わせを返答する。各光電子融合型パケットルータでは、予め設定された、各宛先光電子融合型パケットルータに対する転送先次ルータの候補と、その転送先を経由した際の宛先ルータまでのホップ数を保持する。本方法では、トポロジ管理サーバ、光電子融合型パケットルータのいずれにおいても、光電子融合型パケットルータ間のホップ数を保持することが必要となるが、この情報は、データセンターネットワーク設計時に得ることができ、その値を設定として投入する。

本経路制御方式では、パケットは以下の手順で転送される

(1) 送信元サーバにおけるカプセル化

- パケットの送信元サーバでは、トポロジ管理サーバに問い合わせを行い、返答として得られた宛先光電子融合型パケットルータのアドレス、送信元光電子融合型パケットルータのアドレスを付与してパケットのカプセル化を行う。

(2) 送信元ラック内スイッチにおける転送

- カプセル化されたパケットの付与された送信元アドレスを持つ光電子融合型パケットルータに転送する

(3) 中継光電子融合型パケットルータにおける転送

- 中継光電子融合型パケットルータでは、宛先ヘッダに記載されている光電子融合型パケットルータ宛のパケットの転送先候補で、ポートが空いているもののうち、ホップ数が最小のものを選択して転送する。転送先候補の全ポートが塞がっている場合は、バッファに転送する

(4) 宛先光電子融合型パケットルータにおける転送

- ヘッダ中の宛先アドレスが自身のアドレスとなるような光電子融合型パケットルータでは、パケットのカプセル化を解除し、宛先サーバと接続しているラック内スイッチに転送する。

(5) 宛先ラック内スイッチにおける転送

- ヘッダ記載の宛先サーバにパケットを転送する。

上記の手順により、中継光電子融合型パケットルータでは、可能な限りバッファを使わずに転送を試みることができ、低遅延なサーバ間の通信を実現できる。

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の設計

本稿では、データセンターネットワークは次の二つの手順で設計されるものとする。

(1) 光電子融合型パケットルータとサーバラック間の接続方法の決定

(2) 光電子融合型パケットルータ間の接続方法の決定

このうち、光電子融合型パケットルータとサーバラック間の接続では、より多くのサーバラック間の通信を1ホップで収容できるように、光電子融合型パケットルータが接続しているサーバラックの組み合わせは、光電子融合型パケットルータによって異なるようにする。そして、光電子融合型パケットルータ間

の接続方法では、2.3 で述べた経路制御手法を考慮した上で、もっとも多くの通信需要が収容できるような接続を行う。

以降、本稿で設計するデータセンターネットワークに用いる光電子融合型パケットルータ、サーバラックに対する識別子の割り振りについて述べた上で、上記の各手順について当該識別子を用いて説明する。

3.1 データセンターネットワーク構築に用いる機器と識別子

各サーバラックから光電子融合型パケットルータへの配線数を L^{SV-SW} 、各光電子融合型パケットルータからサーバラックへの配線数を L^{SW-SV} 、各光電子融合型パケットルータから他の光電子融合型パケットルータへの配線数を L^{SW-SW} とする。また、サーバラックは N 台ずつのクラスタに分け、同一クラスタに属するサーバラックは同じ光電子融合型パケットルータに接続されるものとする。

本稿においては、データセンターネットワークは、 $KL^{SV-SW} \left(\frac{L^{SW-SV}}{N} \right)^{L^{SV-SW}-1}$ 台の光電子融合型パケットルータ、 $KN \left(\frac{L^{SW-SV}}{N} \right)^{L^{SV-SW}}$ 台のサーバラック（ただし、 K は整数の定数）で構築する。

各光電子融合型パケットルータ、各サーバラックのクラスタには、以下のように識別子を付与する。光電子融合型パケットルータ n には、識別子 $I^{SW}(n)$ を付与する。 $I^{SW}(n)$ は、 $L^{SV-SW} + 1$ 個の整数からなり、 j 番目の数を $i_j^{SW}(n)$ とあらわす。 $i_0^{SW}(n)$ は 0 から K の数、 $i_1^{SW}(n)$ は 0 から L^{SV-SW} の数、それ以外の i_i^{SW} は 0 から $L^{SW-SV}(n)$ の数とする。

各サーバラックのクラスタ m に対しても、識別子 $I^{SV}(m)$ を割り当てる。 $I^{SV}(m)$ は、 $L^{SV-SW} + 1$ 個の整数からなり、その j 番目の数を $i_j^{SV}(m)$ とあらわす。 $i_0^{SV}(m)$ は 0 から K の定数、それ以外の $i_j^{SV}(m)$ は 0 から L^{SW-SV} の定数である。

3.2 光電子融合型パケットルータとサーバラック間の接続方法

本稿では、各サーバラッククラスタは、同一のサーバラッククラスタと接続する光電子融合型パケットルータは高々ひとつとなるように接続する。このような接続を達成するにあたり、本稿では、BCube [4] と同様のルールにより、光電子融合型パケットルータとサーバラックの接続を行う。

本接続のルールを定めるにあたり、識別子 I のうち、下位 k 個の数を連続した整数値に置き換える演算、 $R(I, k)$ を定義する。

$$R(I, 0) = 0$$

$$R(I, k) = L^{SW-SV} R(I, k-1) + i_{L^{SV-SW}+2-k}^{SV}$$

そして、各サーバラックにおいて、当該サーバラックの接続先の光電子融合型パケットルータを 1 台ずつ選択する。クラスタ m に属するサーバラックが接続する k 台目の光電子融合型パケットルータは、以下の条件をすべて満たす識別子 $I^{SW}(n)$ が割り振られたものとする。

$$R(I^{SW}(n), k) = R(I^{SV}(m), k)$$

$$i_j^{SW}(n) = i_j^{SV}(m), 0 \leq j \leq L^{SW-SV} + 1 - k$$

3.3 光電子融合型パケットルータ間の接続方法

本稿では、光電子融合型パケットルータとサーバラックの接

続を確定した後に、光電子融合型パケットルータ間の接続方法を決める。本稿では、光電子融合型パケットルータを用いたネットワークでの経路制御を模擬した際に収容可能なトラフィック量を指標とし、光電子融合型パケットルータ間の接続方法の候補のうち、もっとも指標の値がよい構成を選択する。

以降、光電子融合型パケットルータ間の接続方法の候補と、その候補から光電子融合型パケットルータ間の接続方法を選択する際に用いる指標について述べる。

3.3.1 光電子融合型パケットルータ間接続の候補

全光電子融合型パケットルータペアに対して、リンクを有無を定めるとすると、候補となるトポロジの数はノード数の 2 乗に比例して大きくなる。本稿では、規則的に構築されたネットワークトポロジのみを候補とすることで、候補数を減らす。これは、上述のルールで光電子融合型パケットルータとサーバラックを接続することにより、特定の光電子融合型パケットルータが特別な役割を果たすことなく、同程度の重要度を持つためである。そのため、適切なネットワーク構造では、光電子融合型パケットルータ間の接続においても、特別な役割を果たす光電子融合型パケットルータはなく、全光電子融合型パケットルータが同程度の役割を担う規則的なネットワーク構造になると考えられる。

本稿では、光電子融合型パケットルータでは、識別子が 0 となるサーバラッククラスタが接続する L^{SW-SV} 台の光電子融合型パケットルータの接続先のみを定める。そして、他の光電子融合型パケットルータの接続先は、識別子 0 のサーバラッククラスタが接続する光電子融合型パケットルータ n_1 が n_2 と接続するのであれば、光電子融合型パケットルータ n'_1 は、以下の条件を満たす光電子融合型パケットルータ n'_2 と接続する。

$$i_0^{SW}(n'_2) = \left(i_0^{SW}(n'_1) + \left(i_0^{SW}(n_2) - i_0^{SW}(n_1) \right) \right) \bmod K$$

$$i_1^{SW}(n'_2) = i_1^{SW}(n_2)$$

$$i_i^{SW}(n'_2) = \left(i_i^{SW}(n'_1) + \left(i_i^{SW}(n_2) - i_i^{SW}(n_1) \right) \right) \bmod L^{SW-SV}$$

ただし、 $i \geq 2$ とする。

ただし、上記のルールで構成されたトポロジのうち、全光電子融合型パケットルータの利用ポート数が L^{SW-SW} となるもののみを候補とする。

3.3.2 トポロジの比較指標

本稿では、全サーバラックから、以下のルールにより、一定数のパケットをネットワークに流した状態を模擬した上で、各リンク上に収容されるパケット数を調べる。

a) パケットの収容先の経路設定方法

各サーバラック間のパケットの収容先を以下の手順で決める

(1) 送信元・宛先光電子融合型パケットルータの決定

- 送信元サーバラックが接続している光電子融合型パケットルータ、宛先サーバラックが接続している光電子融合型パケットルータのペアのうち、これまでに創出されたパケット数が最も少ないものを選択する。

(2) 光電子融合型パケットルータ間のリンクのうち、パケットを収容するリンクを決定

- 送信元光電子融合型パケットルータから順に次ホップの光電子融合型パケットルータの選択を繰り返す。次ホップの

光電子融合型パケットルータを選択する際には、宛先光電子融合型パケットルータまでのホップ数が最短となる光電子融合型パケットルータのうち、現在の光電子融合型パケットルータから当該光電子融合型パケットルータへのリンクを流れるパケット数が最小となるものを選ぶ。

b) 指 標

一定数のパケットをネットワークに流した状態を模擬した後、各リンク上に収容されたパケット数を比較する。本稿では、各リンク上に収容されたパケット数の最大値を指標として用い、最大パケット数が小さいトポロジほど、リンクに負荷を掛けずにトラヒックを収容できるトポロジとみなし、適切なトポロジとする。もし、最大パケット数が同一となるトポロジが存在した場合は、全リンクに収容されたパケット数の平均が小さいトポロジを選択する。

4. 評 価

4.1 最 適 性

本稿では、規則的なトポロジ構造のみをトポロジの候補とした。本小節では、まず、そのように候補のトポロジを制限することの妥当性について評価を行う。本評価では、サーバラック数は8、各サーバラックからの配線数は2、各光電子融合型パケットルータに接続するサーバラック数は2と、光電子融合型パケットルータ数は8、各光電子融合型パケットルータから他の光電子融合型パケットルータへの配線数を3とした。本節では、この資源を使い、取りうる光電子融合型パケットルータ間の全接続構成のうち、もっとも3.3.2で述べた指標が良かったトポロジを図3に示す。図3は、円が光電子融合型パケットルータを示し、光電子融合型パケットルータ間のリンクを線で示す。円中の数字は、3.1のルールに従って生成した識別子である。図より、光電子融合型パケットルータ(0,0,0)と接続している光電子融合型パケットルータは、(1,0,1)、(1,1,0)、(1,1,1)である。これをもとに、3.3.1の接続ルールにより光電子融合型パケットルータ(0,0,1)、(1,0,0)、(1,0,1)の接続先を決定すると、図中の接続構成と同じ光電子融合型パケットルータと接続されることが分かる。同様に、光電子融合型パケットルータ(0,1,0)と接続している光電子融合型パケットルータは、(1,0,0)、(1,0,1)、(0,1,1)であり、これをもとに、3.3.1の接続ルールにより光電子融合型パケットルータ(0,1,1)、(1,1,0)、(1,1,1)の接続先を決定すると、図中の接続構成が得られる。つまり、最適なトポロジは、3.3.2で述べた規則的なトポロジの中に存在することが分かる。

4.2 遅延性能

本小節では、4.1で構築したトポロジにおいて、全サーバ間にトラヒックが発生した場合の遅延について評価する。

c) 比較対象

本小節では、8台の光電子融合型パケットルータを用いて、HyperCubeを構築し、各サーバラックから2台の光電子融合型パケットルータに接続する構成と比較する。この構成は、4.1で構築したトポロジと同じ台数の光電子融合型パケットルータ、サーバラックを同じ数のリンクで接続したネットワークとなる。

d) トラヒック生成モデル

全サーバ間に均等にトラヒックな生成した。

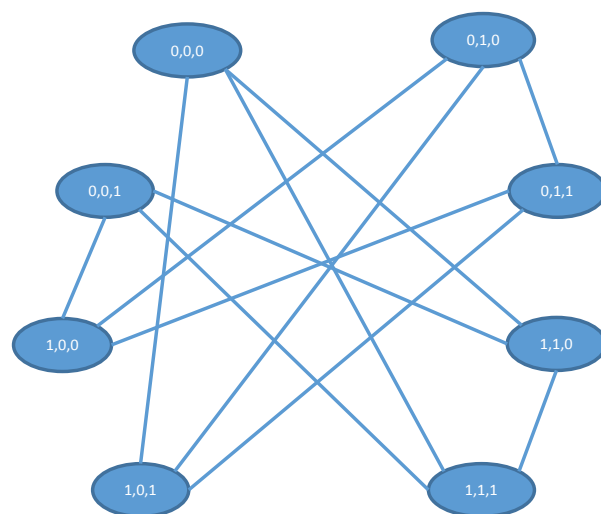


図3 最適なトポロジ（光電子融合型パケットルータのみ記載）

e) 評価指標

本評価では、各サーバラック間の平均遅延を求め、そのうち最大となる遅延について評価を行った。本評価では、各光電子融合型パケットルータのポートの帯域は100 Gbpsとした。バッファを経由しない場合にかかる遅延は240 nsとし、バッファを経由した場合には光・電気の変換、電気・光の変換に600 nsが余分にかかった上で、バッファ内待ち時間が発生するものとした。また、バッファの待ち時間は、M/M/1/Kモデルでモデル化されるものとし、バッファ内には100パケットまで蓄積できるものとした。

f) 評価結果

図4に結果を示す。この図の横軸はネットワーク内に収容された最大トラヒック量、縦軸は各サーバ間の平均遅延のうち最大のもの値を示す。図より、負荷が低い場合、高い場合のいずれの場合であっても、提案手法のトポロジがHyperCubeよりも遅延を抑えることができている。これは提案手法では、各サーバラックからの複数のリンクを効率的に用いることにより、ホップ数を抑えることができるためである。その上、提案手法では、光電子融合型パケットルータ間の接続においても、サーバラックと光電子融合型パケットルータの接続を考慮して、必要な箇所だけにリンクを追加する形態をとっている。そのため、各サーバラックからのトラヒック量が大きい場合であっても、リンク使用率が大きくなり、低遅延でトラヒックを収容できる。

5. ま と め

本稿では、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に適したネットワーク構造について議論を行った。本稿では、光電子融合型パケットルータが提供する広帯域・低遅延の通信特性を利用するため、光電子融合型パケットルータはデータセンター内のコアネットワークに配置し、複

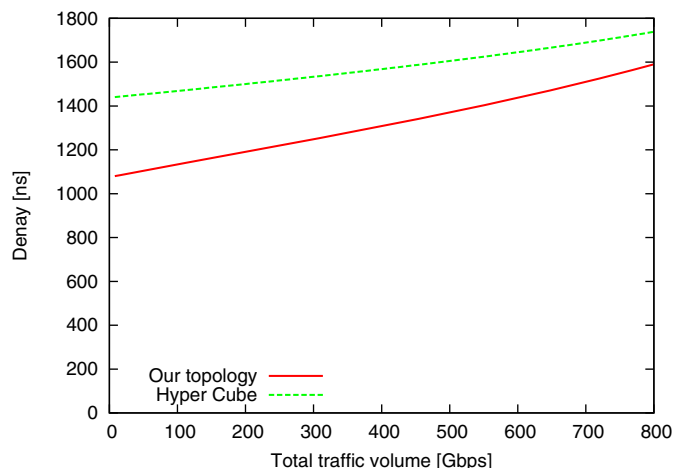


図4 評価結果

数台のサーバラックからのトラフィックを各光電子融合型パケットルータで集約する形態を考えた。また、本構造において、各光電子融合型パケットルータの故障によりデータセンターネットワークから切り離されるサーバラックの存在を避けるため、各サーバラックも複数台の光電子融合型パケットルータに接続する。この条件において、各サーバラックからの複数のリンクを利用しつつ、光電子融合型パケットルータにおけるパケット衝突を避けるルーティングを考慮したネットワーク構造構築方法を提案した。シミュレーション評価により、提案されたネットワーク構造が HyperCube 型のネットワークと比べ、より多くのサーバ間通信を低遅延で収容できることを明らかにした。

今後は、より大規模なデータセンターネットワークを対象に、提案したネットワーク設計手法の評価を行う予定である。また、大規模なネットワークは、一度に構築されるのではなく、段階的な構築が行われると考えられる。そのような場合への、提案手法の適用方法についても議論を行う予定である。

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。

文 献

- [1] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: a hybrid electrical/optical switch architecture for modular data centers," in *Proceedings of SIGCOMM*, pp. 339–350, Aug. 2010.
- [2] J. Luo, S. D. Lucente, J. Ramirez, H. J. Dorren, and N. Calabretta, "Low latency and large port count optical packet switch with highly distributed control," in *Proceedings of OFC*, Mar. 2012.
- [3] "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.
- [4] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, and S. Lu, "BCube: A high performance, server-centric network architecture for modular data centers," *ACM SIGCOMM Computer Communication Review*, vol. 39, pp. 63–74, Aug. 2009.