Vol.2012-ARC-202 No.14
Vol.2012-HPC-137 No.14
2012/12/13

光空間リンクを用いた省配線・可変トポロジであるHPC相互結合網

鯉渕道紘[†1] 藤原一毅[†1] 長谷川洋平[†2]
橋本陽一[†3] 松谷宏紀[†4] 天野英晴[†4]

スーパーコンピュータの大規模化が進むにつれ，インターコネクトは (1) マシンルーム内に 2,000km を越えるなどの肥大化した総配線長の削減，(2) 並列アプリケーション毎に通信アクセスパターンが異なるため最適なネットワークトポロジの設計が困難，という 2 つの課題が顕著となってきている．本研究では，これら 2 つの課題に対して，(i) キャビネット間のリンクの一部を，キャビネットと天井間のフリースペースを利用した光空間通信により構築することで配線長を削減し，(ii) キャビネット間の光空間リンクを再構成することで並列アプリケーションの通信パターン毎にトポロジを最適化可能な相互結合網を提案，評価する．評価結果より，光空間リンクとして 10GBASE-LR イーサネットを光空間伝搬させた場合，測定を行った 10 m までの通信では性能，品質とも劣化が認められず，十分に大きな相互結合網のリンクとして利用可能なことが分かった．さらに，光ケーブルを用いた場合，光がケーブル内を屈折しながら伝搬され，かつ，キャビネット間の配線長がマンハッタン距離となる．一方，光空間リンクでは，光がユークリッド距離により屈折せずに伝搬される．そのため，光空間通信は，光ケーブルを用いた場合と比べて，リンクの遅延を最大 53%削減可能であることが分かった．さらに，解析結果より 8,192 台のスイッチで構成されたハイパーキューブトポロジを実現する場合，1 台のキャビネットから 7 台のキャビネットに光空間リンク接続することで，平均配線長を最大 52%削減できることが分かった．

HPC Interconnection Networks with Short-length Cabling and Variable Topology using Free Space Optical Links

Michihiro Koibuchi,[†1] Ikki Fujiwara,[†1] Yohei Hasegawa,[†2]
Yoichi Hashimoto,[†3] Hiroki Matsutani[†4] and Hideharu Amano[†4]

As supercomputers become large, interconnection networks face two problems to be resolved: (1) reduce the aggregate cable length, e.g. over two thousands meters in a machine room, and (2) design a network topology optimized to various communication patterns of parallel applications. In this work, to mitigate two problems, (i) we use free space optical communication at free space between the ceiling and top of cabinets in order to reduce the aggregate cabling length, and (ii) updating source-destination pairs of free space optical links enables to statically optimize the network topology to communication patterns of parallel applications, and we evaluate such an interconnection network. Evaluation results show that when using free space optics on 10GBASE-LR Ethernet, performance and quality of their links are maintained by at least 10 meters. Optical cables are usually placed along the Manhattan distance in machine room layout, whereas free space optics directly send light along the Euclid distance; delay of a free space optical link decreases by up to 53% compared with that of an optical cable. The analysis results illustrate that the average cable length is reduced down by up to 52% by using free space optical cables from a cabinet to seven cabinets in 8,192-switch hypercube topologies.

†1 国立情報学研究所
National Institute of Informatics
†2 NEC クラウドシステム研究所
Cloud System Research Laboratories, NEC Corporation
†3 NEC グリーンプラットフォーム研究所
Green Platform Laboratories, NEC Corporation
†4 慶應義塾大学大学院 理工学研究科
Graduate School of Science and Technology, Keio University

1. はじめに

最近の HPC (High Performance Computing) システムの（オフチップ）相互結合網は，バンド幅の増加

Vol.2012-ARC-202 No.14
Vol.2012-HPC-137 No.14
2012/12/13

に伴い，同一キャビネット内スイッチへのリンクは電気ケーブル，キャビネット外スイッチへのリンクは光ケーブルを用いて構築されることが多い．そのため，大規模化が進むにつれて，(1) 物理的なケーブルの総配線長が顕著となり，また，(2) 並列アプリケーション毎に生じる通信アクセスパターンが異なるため，その最適なトポロジと物理的な制約から構築されたネットワーク・トポロジとの乖離が大きくなっている．

前者に関しては，例えば，初代地球シミュレータの配線長が 2,000km を大きく超え，京コンピュータが約 1,000km に達していることを考えると，施工性・メンテナンス性・省資源性の観点から，スーパーコンピュータの配線長を抑える技術が今後重要となる可能性がある．加えて，ラック間ケーブルが増えるにつれて，そのバックアップケーブル数も増加する．これらはプラットフォーム構築時に設置する必要があるため，負担が無視できない．

後者については，理想的には，対象とした並列アプリケーションの通信パターンに適したトポロジを採用することが望ましい．しかし，異なる通信パターンを持つ並列アプリケーションを実行する既存の HPC システムでは，そのようなトポロジの選択は難しい．したがって，トーラス，ツリーなどのネットワークトポロジの中から[1)2)]，直径，スイッチの次数，ルーティングの容易性，耐故障性，レイアウトとコストなどの点でトレードオフを考慮した上で HPC システム毎に設計者の総合的な判断により（異なる）トポロジが選択されている（例：京コンピュータでは 6 次元トーラス，TSUBAME 2.0 では Fat ツリー）．したがって，システムが採用したトポロジ毎にユーザが並列プリケーションの最適化を行うことが必要となる．

本研究では，これら 2 つの課題を緩和させるため，キャビネット間リンクを（有線）光ケーブルのみならず，光空間通信により構築する可変トポロジである相互結合網を提案する．

まず，我々は，安価かつ安定的に動かすために，10 ギガビットイーサネットを無線化することで光空間リンクを構築する．そして，この光空間リンクを用いてネットワークトポロジとそのレイアウトを次のように実現する．(1) キャビネット間の多数のリンクを，キャビネットと天井間のフリースペースを利用して光空間通信により構築することで配線長を削減する．(2) キャビネット間のリンクを再構成することで並列アプリケーションの通信パターン毎にトポロジを最適化可能とする．

光空間リンクは，遮蔽物がある場合，通信が途絶える特徴がある．そこで，本相互結合網では，キャビネットと天井の間の空間を光空間伝送に利用することでこの問題を直接的に解決する．なお，本光空間リンクのデータ転送において，照明の影響などを受けないことは確認している（ 4 章）．

本研究の貢献は次の通りである．

- 10GBASE-LR イーサネットを光空間伝送のデバイスに用いた場合，10m までの距離であれば光ケーブルを用いた場合と同様の安定性とバンド幅 9.4Gbps を得ることができた．(4 章)
- 光空間転送は直進性を持つため，光ケーブルの転送遅延と比べて 33%減である 3.2ns/m のリンク遅延を実現した．(4 章)
- 8,192 台のスイッチで構成されたハイパーキューブトポロジを実現する場合，すべて光ケーブルで構築した場合と比べて，本相互結合網は 1 台のキャビネットから 7 台のキャビネットに対して光空間リンク接続することで，平均配線長を最大 52%削減できることが分かった．(5 章)

2. 関連研究

2.1 トポロジ

HPC システムのネットワークトポロジとして，トーラス，メッシュ，ハイパーキューブを含む k-ary n-cubes や，Fat ツリーが広く利用されてきた．k-ary n-cubes の他にも各種の規則的な直接網が提案されており，直径と次数の点でトレードオフを持つ．例えば De Bruijn（3,072 ノードにおいて直径 12，次数 4），Kautz（同 11，4），Pradhan（同 12，5），スターグラフ（5,040 ノードにおいて同 7，6），パンケーキグラフなどである[1)]．さらに，近年，我々は，ランダムなショートカットリンクがネットワークの直径と平均距離を劇的に小さくする現象に着目し，HPC システムのネットワークへの応用を探究している．これまでの研究[2)] において我々は，ランダムトポロジが同じ次数の規則的なトポロジに比べて低遅延であることを示した．また，HPC システムの高次元ネットワークの場合，乱数によるネットワーク性能のばらつきが十分小さいことを確かめた．

HPC システムのネットワークは高バンド幅（リンク当たり 10〜40Gbps 以上）を必要とするため，ラック内程度の短いリンクには安価な電気ケーブルを利用可能だが，ラック間を結ぶ長いリンクには高価な光ケーブルを使わざるをえない．したがって，システムレイアウトがネットワークコストに大きく影響する．ドラゴンフライ網[3)] はこの点に着目し，トポロジをラック

Vol.2012-ARC-202 No.14
Vol.2012-HPC-137 No.14
2012/12/13

内とラック外の 2 階層に分け，複数のルータでひとつの仮想ルータを構成する．ドラゴンフライの各階層には，ランダムトポロジを含め，多様なトポロジを埋め込むことができる．

本相互結合網は光空間リンク数を十分に設置した場合，これらのトポロジを実現することが可能である．

2.2 60GHz 無線技術を用いたデータセンターネットワーク

近年，60GHz 無線リンクは 2.4GHz 802.11b/g よりも 80 倍のバンド幅を持つことが報告されている．また，その通信デバイスは HDMI ケーブルの代替手段として普及した場合，$10 以下という安価に実現可能である点からデータセンターネットワークに適用する議論が行われている[4]．この場合，本研究と同様にキャビネット上のスペースを用いて有向アンテナを設置し，Top-of-Rack スイッチ間を相互接続することが想定される．光空間リンクと同等に，60GHz 無線リンクは送受信デバイスを物理的に向き合わせて通信を行うが，異なる無線リンク間の相互干渉を配慮したレイアウト制約が厳しい[4]．近年，ミラーを用いた反射により，障害物を迂回する通信路の設定についても研究が行われている[5]．しかし，帯域は数 Gbps に留まる点から HPC 用途では，ホットスポットの一時的な退避や故障箇所の一時的な迂回に利用するなどの用途に留まる可能性が高い．

ただし，60GHz 無線技術で研究開発されたアラインメント用メカニカル技術などは本光空間リンクの構築に応用可能であると考えられる．

3. 光空間リンクを用いた可変トポロジである相互結合網

3.1 トポロジ構成とレイアウト

本相互結合網ではケーブルの配線長を抑えるため，マシンルームにおいて物理的に近隣のキャビネット間のみにケーブルを接続する（図 1.(a)）．ここでは簡素化した議論を行うために，マシンルームの大きさに合わせて 2 次元格子状にキャビネットを設置し，隣接キャビネットにのみケーブルを接続することを想定する．

2 次元メッシュ状に $a \times b$ 台のキャビネットを配置し，1 つのキャビネットに c 台のスイッチを格納した場合，$a \times b \times c$ 3 次元メッシュトポロジとなるように有線ケーブルを接続する．なお，$c = 1$ の場合は，1 台の Top-of-Rack スイッチを用いた典型的なデータセンターネットワークと同じ構成となる．

次に，各スイッチから e 本の光空間リンクを設置する．つまり，任意のキャビネットのスイッチ間に光空間リンクを構築できるようにする．なお，光空間リンクの場合，相互干渉はないため，送受信間に遮蔽がない空間で直線上にアラインメントする，あるいは，ミラーにより屈折させることで通信ができる（詳細は 4 章）．したがって，60GHz 無線リンクの場合と異なり[4),5)]，マシンルームの物理的なリンクの設置制約は極めて緩い．

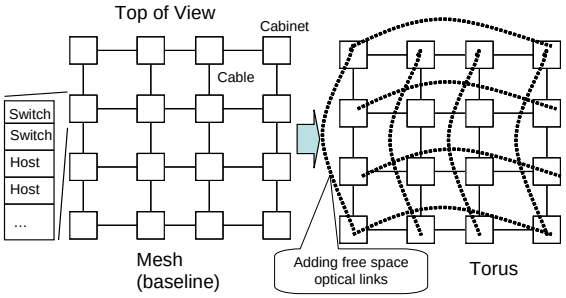

図 1 本相互結合網のトポロジ構成例（キャビネット内リンクは略）
(a) 有線ケーブルのみ（メッシュ） (b) トーラス

3.2 生成トポロジ

ベースとなる有線による 3 次元メッシュトポロジに加えて，通信パターンに合わせて光空間リンクを構築することで様々なトポロジが構成可能である．ここでは代表的なトポロジの構成について述べる．

3.2.1 ランダムトポロジ

ランダムトポロジは直径，平均距離，スループット，end-to-end 遅延の面で優れた性能を持つことが報告されている[2)]．そこで，特定の通信パターンを想定しない場合は，ランダムトポロジを構成する．本相互結合網では，ケーブルで接続された 3 次元メッシュに対して，(i) 2 つのキャビネットをランダムに選択し，(ii) その各々のキャビネットからランダムに選択した 2 台のスイッチ間に光空間リンクを追加することでランダムトポロジを構築する．

3.2.2 k-ary n-cubes（トーラスとハイパーキューブ）

最外周のキャビネット内スイッチの光空間リンク数 $e \geq 2$ である場合，光空間リンクを最外周のキャビネットのスイッチ間で接続することで 3 次元トーラス構造を実現することができる．ハイパーキューブも同様に構築可能である（図 1.(b)）．

3.2.3 Fat ツリー

ツリー系のトポロジの場合，一般的に根に近いスイッチのリンク，あるいは葉スイッチのリンクが長くなる傾向が強い．したがって，1 つの方策としてこの

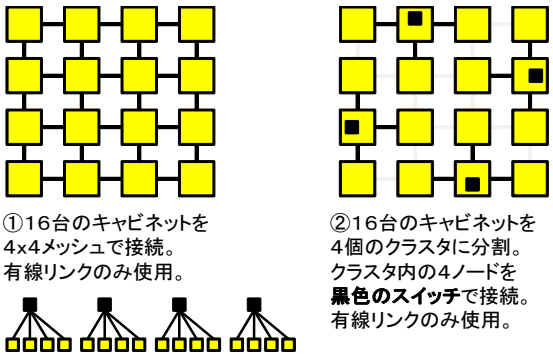

図 2 Fat ツリー (2,4,2) の葉の構成

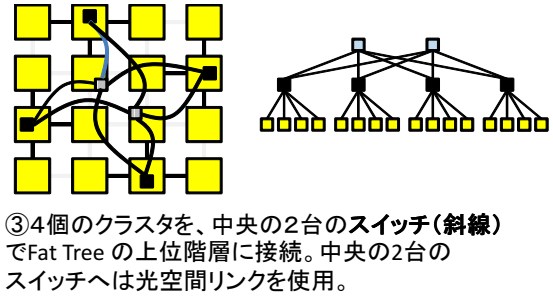

図 3 Fat ツリー (2,4,2) の全体構成

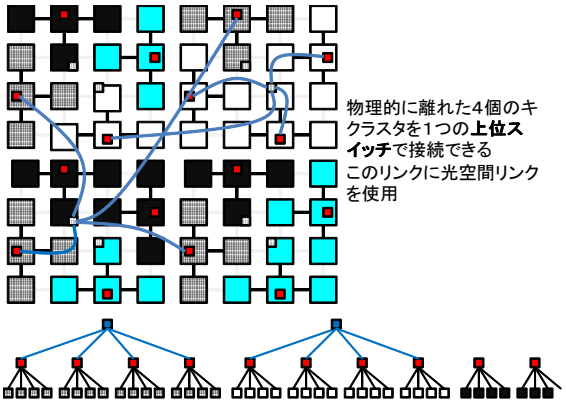

図 4 離れた部分木を用いて構成した Fat ツリー (2,4,3) の例 (上位リンク (p=2) の一部略)

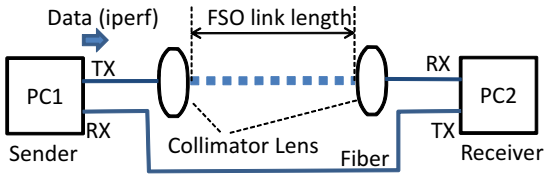

図 5 光空間リンク帯域とエラー率計測の実験構成（光空間長を変化させた 場合）

長いリンクを光空間通信で実現することが考えられる．

ここで 1 台のスイッチから p 台の上位スイッチ，および q 台の下位スイッチへの接続を持ち，階層数 r の Fat ツリーを (p,q,r) で表す．

スイッチが持つ光空間リンクが $p+q$ 本の場合，任意のキャビネットを用いて Fat ツリー (p,q,r) を構築することができる．図 2-4 に Fat ツリー (2,4,2) と Fat ツリー (2,4,3) の例を示す．なお, (3 次元メッシュを構成する）ケーブルを利用し，かつ，Fat ツリーのマッピングを最適化することで必要となる光空間リンク数を抑える最適化については，ここではこれ以上議論を行わない．

3.3 可変トポロジの特徴

3.3.1 柔軟なパーティショニング

Fat ツリー (2,4,3) の例である図 4 のように離れたキャビネットのスイッチ間を接続して規則的なサブトポロジを構成することができる．このパーティショニングの自由度の高さにより，マルチユーザ環境でのタスク割り当て，および，故障箇所の迂回時に，規則的なトポロジを効率良く構成することが可能である．

3.3.2 耐故障性とヒューマンエラー対策

一度 HPC システムを設置し，運用を開始した後にケーブルの追加，除去を行うことは難しいため，キャビネット間にバックアップケーブルを設置するなどの対策が必要となる．また，スイッチのポートへの配線時に生じうるヒューマンエラー対策も考える必要がある．本相互結合網の場合，前者は光空間リンクにより可能であり，後者については光空間リンクの再構築により対策が可能である．

3.3.3 トポロジ構成の柔軟性

キャビネット間の光空間リンクの送受信対を更新することで並列アプリケーションの通信パターン毎にトポロジ更新する可変相互結合網を実現できる．ただし，光空間リンクの本数が少ない場合，構成可能なトポロジが限定される．例えば，隣接キャビネット間およびキャビネット内リンクのみ配線した場合，1 つのキャビネットから，6 つのキャビネットへの光空間リンクを用意することで Fat ツリー (2,4,r) および 6 次元トーラス（キャビネット内で 2 次元トーラスを生成）を構築することができるが，Fat ツリー (3,4,r) および 7 次元トーラスを構築することは難しい．

4. 光空間リンクの性能評価

本章では，光空間リンクのリンク帯域，エラー率，通信遅延を計測した結果を示す．

4.1 リンク帯域とエラー率

図 5 および 6 に実験構成を示す．送信光パワー 0.5 dBm，最小受信感度-12.5 dBm の性能をもつ 10 ギ

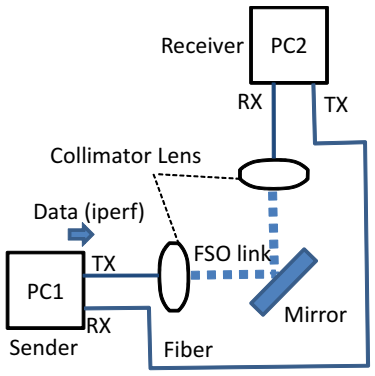

図 6 光空間リンク帯域とエラー率計測の実験構成（ミラー反射させた場合）

ガビットイーサネットカード (10GBASE-LR) を PC（NEC Mate ME28，CPU：Core i7 2.8GHz 4 コア，OS：ubuntu linux 12.04）に実装し，送受信端末とした．光空間リンクは，光ファイバとコリメータレンズで構成し，3 軸のスライド調整ステージで光ファイバとコリメータレンズの相対位置を調整し，光ファイバ間を結合できる構成となっている．このように構成された実験系において，10GBASE-LR からの信号光は，コリメータレンズでコリメート光に変換され空間へ放出され，空間伝搬後再びコリメータレンズで光ファイバへ集光され，10GBASE-LR で受信される．コリメータレンズ間距離を光空間リンク長と定義し，0.3m，0.7m，5m，10m に設定した．なお，今回の実験では，光空間リンクとして単一片方向通信の評価のみに限定するため，受信端末から送信端末への逆方向の光リンクは光ファイバにて構築した．また，光空間リンク間に中継ミラーを配置した場合の構成を模擬するため，光空間リンク長を 0.4m とし，コリメート光の伝搬方向を 90 °反射した構成も評価した（図 6）．

以上の構成で，iperf[6] による TCP データ転送をし，光空間リンクのリンク帯域を確認したところ，光空間リンクを 0.3m，0.7m，5m，10m，および 0.4m（ミラー反射）と設定したいずれの場合でも安定して 9.4Gbps のデータ転送速度を達成し，ファイバ経由の通信と等しいリンク帯域が得らることを確認した．試験は，蛍光灯による室内照明の下，PC の振動など生活ノイズのある環境で実施したが，それぞれの通信距離で各 4 時間のデータ転送を実施した試験中にパケットロス・ビットエラーは発生しなかった．

次に，光空間リンク長毎の光パワーバジェットを評価した結果を図 8 に示す．光パワーは，送信側と受信側のコリメータレンズ前後の計 4 箇所の計測ポイント (MP) で評価した（図 7）．

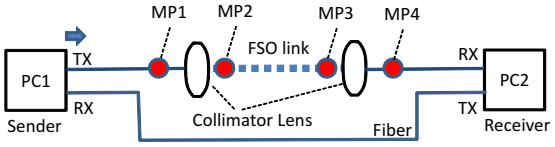

図 7 光信号パワー計測ポイント．

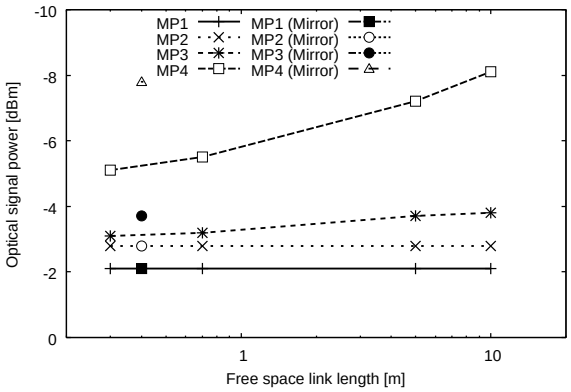

図 8 光信号パワーの評価結果

今回の光リンクは 10mm 径の市販のコリメートレンズと光ファイバで構築したものであるが，短距離 (0.3 m) でのリンク損失は 3 dB であった．リンク長を 10m まで伸長しても 6.0 dB 程度のリンク損失に抑えることができていることが確認できる．これはイーサネットカードの受信感度（-12.5dBm）よりも十分に強いパワーの光信号が得られたことになる．ただし，結果から，MP3 でのコリメート光の空間伝搬回折によるビーム広がりが生じ，10m 空間伝送後のビーム径がレンズ径に対して若干大きくなったことが観測されたことに加え，リンク長に応じた受信側レンズ入力前 (MP3) での光パワーに対し，光ファイバ端での光パワー (MP4) の低減が大きいことを考慮すると，10m 以上にリンク長を伸長する場合には，光ファイバ結合を含む光学設計が必要となる可能性がある．

4.2 通信遅延

次に，図 5 の端末をネットワークテストベッド GtrcNET-1[7] に置き換え，光空間リンクの通信遅延を計測した．なお，今回は計測装置の都合でギガビットイーサネット (1000BASE-LX) を用いて ping パケットの遅延を計測したが，光の伝送速度はイーサネットの規格に非依存であるため本計測結果は 10Gbps 以上の環境においても同様と想定される．GtrcNET-1 は 32nsec 単位までの計測となるため，光空間伝送の距離

を十分に確保 (10m) し，各々1,000 個の ping パケットの遅延を計測した．計測の結果より，光有線（ファイバ）リンクでは 4.8ns/m であったが，光空間リンクでは直進光となるため 3.2ns/m となり，33% の遅延削減ができることを確認した．よって，実際の HPC システムでは配線はキャビネット間のマンハッタン距離となるため，ユークリッド距離で通信を行う光空間リンクは，リンク遅延を最大 53%削減できることが分かった．

以上の計測結果から，空間伝搬距離 10m までの光空間リンクが光ケーブルと同等の帯域を達成し，かつ，通信遅延を 33%削減できることが分かった．また，ミラーが使えることから，キャビネット間に光空間リンクを設置する場合に生じる複数光空間リンク間の物理位置の制約は極めて緩いといえる．

5. 配線の評価

本章では，典型的なトポロジを本相互結合網を用いて実現する場合に必要となる光空間リンク数とトポロジの配線長について評価を行う．具体的には， 2 章で述べたトポロジの中から，規則的かつスケーラブルなトポロジの代表例としてハイパーキューブと，直径，次数が優れているが配線長が長くなるランダムトポロジ[2] の 2 つを選択した．なお， 2 章で述べたその他の多くのトポロジを実現するために必要となる配線長は，両者の中間となることが予想される．

5.1 パラメータ

Cray BlackWidow（0.57m × 1.44m, 128 ノード/キャビネット）[3] を考慮にいれ，1 つのキャビネットに 128 計算ノード，16 スイッチを格納し，各スイッチは 8 計算ノードと接続することとする．$2^{13} = 8192$ スイッチネットワークまでの評価を行う．これは Top500 ランキング[8] の上位スパコン（数百〜1,000 キャビネット）と同程度の規模である．また，ANSI/TIA/EIA-942 標準をもとにラックの寸法は配線や空調スペースを含めて幅 60 [cm] ×奥行 210 [cm] とする．また，文献[9] と同様にキャビネット内のケーブルオーバヘッドを 2m，配線はマンハッタン距離で設置するものとする．

5.2 平均リンク長

本相互結合網は，隣接キャビネット間，およびキャビネット内のみケーブルで配線し，それ以外は光空間リンクを用い構成するものとする．

また，これらのトポロジを有線のみで構築，最適化した場合との比較についても評価する．すべて有線でスイッチ間トポロジを構築した場合は，グラフ・クラスタリングと SA (Simulated Annealing) によるキャビネット配置の最適化により配線長の最小化を行った[10]．

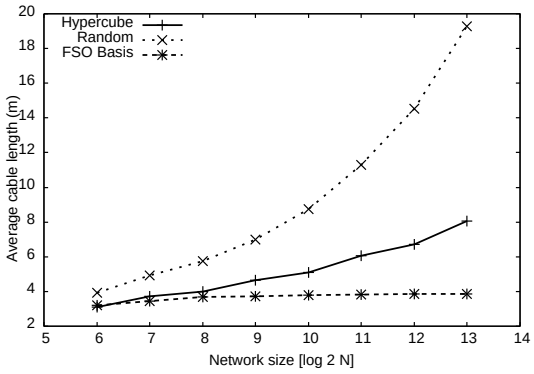

図 9 ケーブル長 vs ネットワークサイズ．

図 9 において，“HYPERCUBE”，“RANDOM” は有線のみで各々，ハイパーキューブ，ランダムトポロジを構成した場合のリンクの平均配線長を表している．公平な比較のため，両トポロジともに次数は $log_2 N$ としている．一方，“FSO Basis” は隣接キャビネット間およびキャビネット内のみに配線した 4 次元メッシュトポロジ (うち 2 次元はキャビネット内) であり，次節で述べる通り光空間リンクを追加することで，ハイパーキューブ，ランダムなどの任意のトポロジを実現することができる．なお，計算ノードとスイッチ間配線はトポロジによらず均一のため除いた．

図 9 より，ネットワークサイズが大きくなるにしたがって，ハイパーキューブ，ランダムともに平均配線長が著しく長くなることが分かる．一方，本相互結合網は隣接キャビネット間とキャビネット内のみを有線で実現するため配線長はほぼ一定であることが分かる．このことから，本相互結合網は配線長の面で，特に 512 台，あるいはそれ以上のスイッチを用いたネットワークにおいて有効であるといえる．

5.3 必要となる光空間リンク数

図 10 に，前節で述べた本相互結合網からハイパーキューブ，およびランダムトポロジを生成する場合に，1 つのキャビネットから必要となる光空間リンク接続先キャビネット数を示す．図 10 に示す通り，規則的なトポロジであるハイパーキューブの場合，必要となる光空間リンク数は緩やかな増加にとどまる一方，ランダムトポロジの場合は，爆発的な増加となる．しかし，ランダムトポロジについては，直径，平均距離を小さく保ったまま，リンク長を削減するランダム手法[11] が提案されており，実際にはハイパーキューブに近い

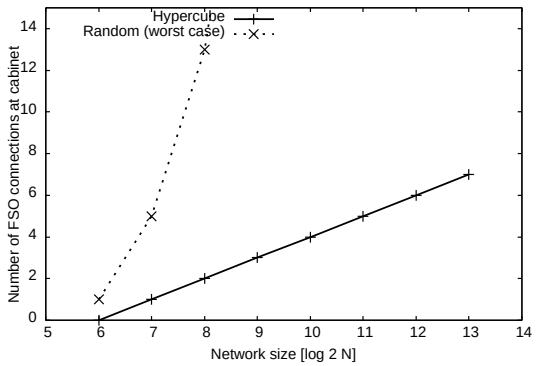

図 10 キャビネットあたりの FSO コネクション数 vs ネットワークサイズ.

光空間リンク数で実現できる可能性がある．図 9 および 10 の結果より，光空間リンクの導入により，本相互結合網はケーブル長を大幅に削減できるといえる．

6. おわりに

本報告では，(1) キャビネット間の多数のリンクを，キャビネットと天井間のフリースペースを用いた光空間通信により構築することで配線長を削減し，(2) キャビネット間の光空間リンクを再構成することで並列アプリケーションの通信パターン毎にトポロジを最適化可能な相互結合網を提案し評価を行った．

評価結果より，光空間リンクとして 10GBASE-LR イーサネットを用いた場合，10 m までの通信では性能，品質とも劣化が認められなかった．安定性も高く，ミラーを用いて光を屈折させた光空間リンクの構築が可能であり，十分に大きな相互結合網のリンクとして利用可能なことが分かった．

さらに，光ケーブルを用いた場合，光がケーブル内を屈折しながら伝搬され，かつ，キャビネット間の配線長がマンハッタン距離となる．一方，光空間リンクでは，ユークリッド距離で屈折せずに伝搬される．そのため，光空間通信は，光ケーブルを用いた場合と比べて，リンクの遅延を最大 53%削減可能であることが分かった．また，8,192 台のスイッチで構成されたハイパーキューブトポロジを実現する場合，1 台のキャビネットから 7 台のキャビネットに光空間リンク接続することで，平均配線長を最大 52%削減できることが分かった．

今後は，(i) 40Gbps 光空間リンクの構築と多重化の評価，(ii) 光空間リンクのアラインメントの自動化，(iii) 実現可能なトポロジの制約と光空間リンク数に関する最適化 (iv) データセンターネットワークへの展開を行う予定である．

謝辞 本研究の一部は，国立情報学研究所共同研究費 （一般研究企画型） の支援による．

参考文献

1) 天野英晴: 並列コンピュータ, 昭晃堂 (1996).
2) Koibuchi, M., Matsutani, H., Amano, H., Hsu, D. F. and Casanova, H.: A Case for Random Shortcut Topologies for HPC Interconnects, *Proc. of the International Symposium on Computer Architecture (ISCA)*, pp. 177–188 (2012).
3) Kim, J., Dally, W. J., Scott, S. and Abts, D.: Technology-Driven, Highly-Scalable Dragonfly Topology, *Proc. of the International Symposium on Computer Architecture (ISCA)*, pp. 77–88 (2008).
4) Halperin, D., Kandula, S., Padhye, J., Bahl, P. and Wetherall, D.: Augmenting data center networks with multi-gigabit wireless links, *SIGCOMM*, pp. 38–49 (2011).
5) Zhou, X., Zhang, Z., Zhu, Y., Li, Y., Kumar, S., Vahdat, A., Zhao, B. Y. and Zheng, H.: Mirror mirror on the ceiling: flexible wireless links for data centers, *SIGCOMM*, pp. 443–454 (2012).
6) Iperf - The TCP/UDP Bandwidth Measurement Tool: http://dast.nlanr.net/Projects/Iperf/.
7) Kodama, Y., Kudoh, T., Takano, R., Sato, H., Tatebe, O. and Sekiguchi, S.: GNET-1: gigabit Ethernet network testbed, *Proceedings of IEEE International Conference on Cluster Computing*, pp. 185–192 (2004).
8) Top 500 Supercomputer Sites: `http://www.top500.org/`.
9) Kim, J., Dally, W. J. and Abts, D.: Flattened butterfly: a cost-efficient topology for high-radix networks, *Proc, of the International Symposium on Computer Architecture (ISCA)*, pp. 126–137 (2007).
10) Fujiwara, I., Koibuchi, M. and Casanova, H.: Cabinet Layout Optimization of Supercomputer Topologies for Shorter Cable Length, *The International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT)* (2012).
11) 藤原一毅, 鯉渕道紘: ランダムなネットワークトポロジのためのラック配置最適化, 信学技報, CPSY (2012).