通信帯域10倍，でも，時々間違える

不完璧な計算システム・ネットワーク

Imperfect Computer Systems and Networks

鯉渕 道紘, 平澤 将一, 胡 曜，T. T. NGUYEN, 藤原 一毅

M.Koibuchi, S.Hirasawa, Y. Hu, T.T. Nguyen, I. Fujiwara

どんな研究？

ビッグデータ処理は，物理法則などの理論に基づく厳密さが要求される古典的な大規模並列計算と比べて，コンピュータとネットワークが保証する精度を大幅に緩和しても結果の大勢に影響せず十分なことが多い．

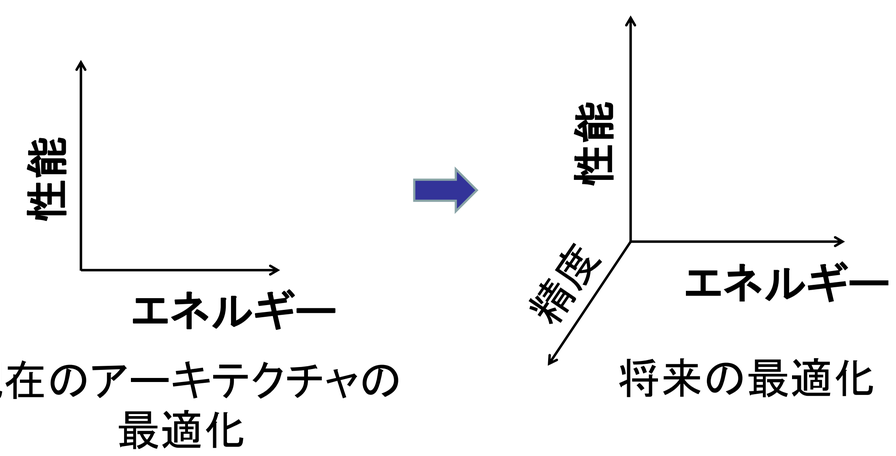
我々は，いい加減さ，すなわち時々計算を間違える**不完璧**さを許容することで，コンピュータとネットワークの性能が大幅に向上することを示す．

その並列計算、厳密ですか？

ビットエラー、直しますか？

状況設定

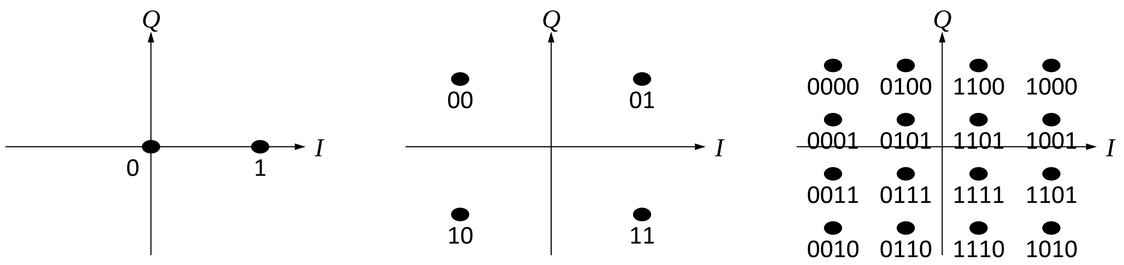
	現在	将来
巨大アプリケーション	科学技術演算	ビッグデータ解析、AI/脳
計算機アーキテクチャ	**性能/電力効率探求** (精度は厳密)	**いい加減さを探求**



研究内容（1）

■ Approximate ネットワーク

- 通信の変調方式をチューニング (帯域と精度のトレードオフ)



On-Off Keying Bandwidth 1x　　4-QAM Bandwidth 2x　　16-QAM Bandwidth 4x

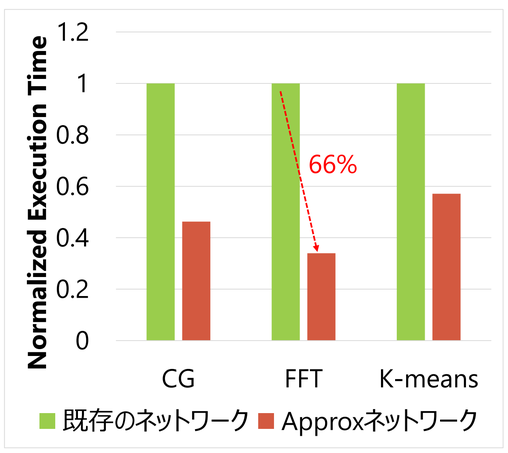
1024-QAM achieves 10x Bandwidth（100 Gbps → 1 Tbps）

ビット誤り率 1.52×10^{-16}　　6.25×10^{-10}　　1.0×10^{-5}

Reduced symbol mapping improves BER by orders of magnitude

- 本ネットワーク採用時の並列アプリケーションの高速化評価[1]
  - CG法：Double型上位16ビットを保護 → Verification OK
  - K-means クラスタリング：上位12ビットを保護 → 誤差0.003%

256 ノード・ネットワーク、スイッチ遅延 60 ns、ケーブル遅延 25 ns/本、500 GFlops



並列アプリケーションのApproximate化

Approximate MPI (FORTRAN, C, C++)

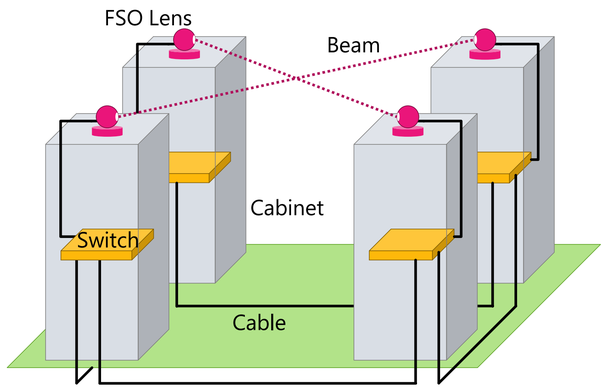
*call mpi_send (var, n, **A_MPI_DOUBLE_PRECISION**, ...)*

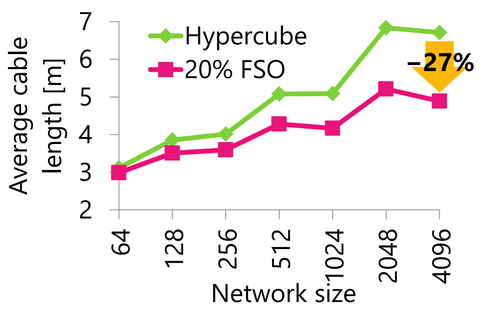
```
12  for (i = 0; i < MAX_ITER; i++) {
13      MPI_Send(&x, DATALEN, MPI_DOUBLE, 1, 1, MPI_COMM_WORLD);
14  }
15  for (i = 0; i < MAX_ITER; i++) {
16      MPI_Send(&x, DATALEN, A_MPI_DOUBLE, 1, 1, MPI_COMM_WORLD);
17  }
```

```
144  MPI_Allreduce(newClusters[0], clusters[0], numClusters*numCoords,
145          A_MPI_FLOAT, MPI_SUM, comm);
146  MPI_Allreduce(newClusterSize, clusterSize, numClusters, A_MPI_INT,
147          MPI_SUM, comm);
```

研究内容（2）

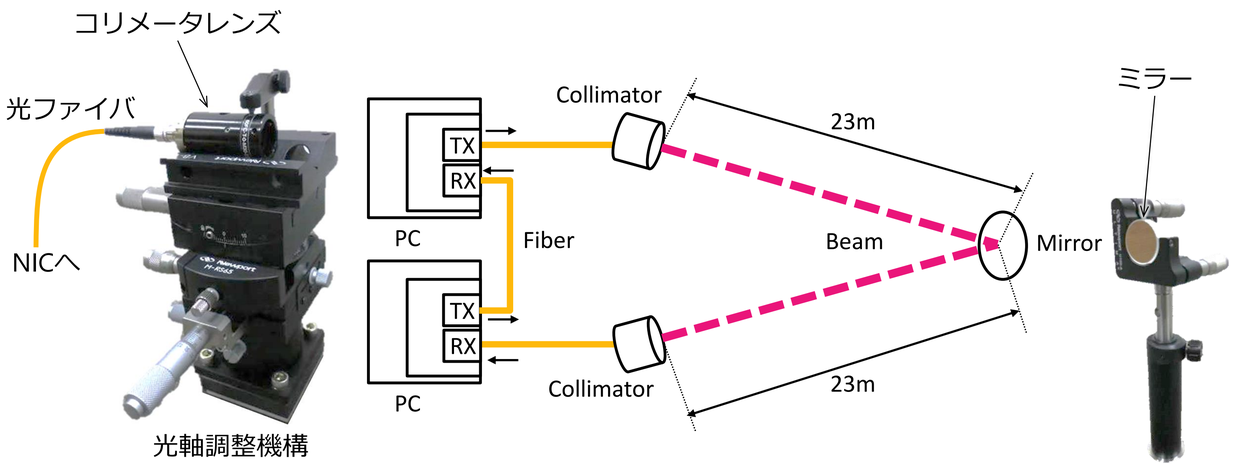
■ 光無線通信（Free Space Optics: FSO）





▲ ラック間ケーブルの 20% を FSO に置き換えると、平均ケーブル長が 27% 減る

- マシンラックの上にレンズを置き、ラック間を光ビームで接続
  - ➢ 10GBASE-LR／40GBASE-LR4 の赤外光を利用（λ=1310nm）
- 46m 離れたレンズ間でミラーを介した通信実験 [2]
  - ➢ **40Gbps** ワイヤレートで長時間安定した通信を確認
- ビームの向きを変えるだけでトポロジを変えられる
  - ➢ パーティショニング、故障回復



発表論文

[1] Daichi Fujiki, Kiyo Ishii, Ikki Fujiwara, Hiroki Matsutani, Hideharu Amano, Henri Casanova, Michihiro Koibuchi, ``High-Bandwidth Low-Latency Approximate Interconnection Networks'',The International Symposium on High-Performance Computer Architecture, Feb 2017

発表論文

[2] Ikki Fujiwara, Michihiro Koibuchi, Tomoya Ozaki, Hiroki Matsutani, Henri Casanova: "Augmenting Low-latency HPC Network with Free-space Optical Links", The 21st IEEE International Symposium on High Performance Computer Architecture (HPCA 2015), Feb. 2015.

NII

連絡先：鯉渕 道紘／ 国立情報学研究所 アーキテクチャ科学研究系　准教授

TEL : 03-4212-2575　FAX : 03-4212-2120　Email : koibuchi@nii.ac.jp