

## 将来のスーパーコンピューティングへの挑戦

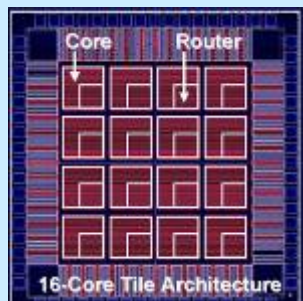
# スパコンのネットワーク (インターコネクト)

国立情報学研究所  
アーキテクチャ科学研究系  
鯉渕 道紘

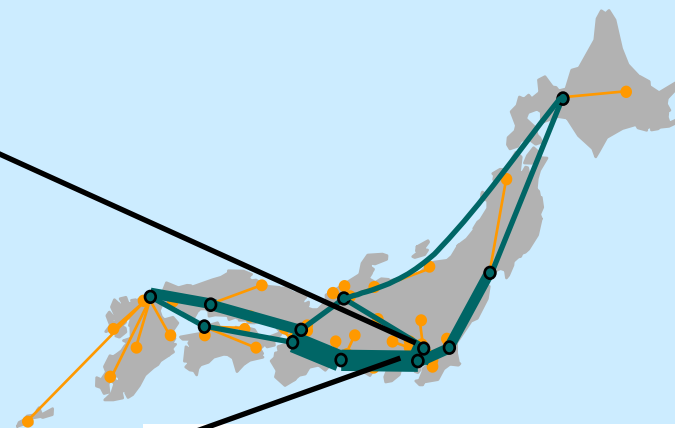
# 自己紹介

- 鯉渕 道紘(こいぶち みちひろ)  
<http://research.nii.ac.jp/~koibuchi>
- 2003.3 慶大理工・博士(工学)(天野研)
- 2005.4～国情研(NII)
  - この間, USC, バレンシア工科大学訪問研究員
  - 計算機システムのネットワーク(インターコネクト)
    - トポロジ、ルーティング、スイッチング、通信ライブラリ、...

メニーコアプロセッサのNW  
松谷、天野@慶大、吉永@電通大



HPC向け省電カインターコネクト  
CREST ULP-HPC(松岡@東工大)



サービス指向ルータ@  
NICT新世代NW  
(西@慶大、川島@筑波)  
SINET3/4

# 概要

- スパコン内ネットワーク
  - 大学発のアイデアが盛りだくさんの多様性



- WAN(インターネット)の利用
  - スパコン内ネットワークのノード帯域と同等



# 独自開発の超高速ネットワークを用いた 上位スパコン@top500

「The int. network is the heart of parallel architecture」

Chuan-Lin Wu and Tse-Yun Feng. *Interconnection Networks for Parallel and Distributed Processing* (1984)

- 歴代No.1スパコン@top500は、独自開発ネットワーク多数
  - Tofu@京速(2011.6～)
  - フルクロスバ@初代 地球シミュレータ
- 大学発のアイデアが実スパコンNWに貢献
  - Bill Dally @ Stanford / NVIDIA
    - 仮想チャネル、次元順ルーティング、ワームホールスイッチング
  - Jose Duato @ Technical U of Valencia
    - Duato's protocol @ BlueGene/L, Cray Black Windows, T3E, GS320

# 教員のアイデアがスパコンNWで使われるまで

引用: ISCA2010キーノート(W.J.Dally)

| Idea                                  | Concept | Published | Product | $\Delta T$ |
|---------------------------------------|---------|-----------|---------|------------|
| Virtual Channels <sub>(Dally)</sub>   | 1985    | 1990      | 1992    | 7          |
| High-Radix NWs <sub>(Dally)</sub>     | 2002    | 2005      | 2008    | 6          |
| Wormhole SW <sub>(Dally)</sub>        |         | 1987(?)   | 90年代    |            |
| Duato's protocol <sub>(J.Duato)</sub> |         | 1995      | 90年代    |            |



多くがIEEE Trans. on Parallel and Distributed Systems(TPDS)で発表



私の目標: アイデアで世界No.1スパコンNWに貢献を!!

- デッドロックフリーNW理論, 当時世界最速CMOSスイッチRHINET-2/SWを利用した実証など計6論文@TPDS誌(2005~)

- Jose Duato グループ等との共同研究(2004~)

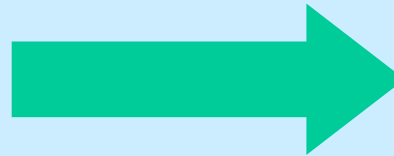
RHINET-2/SWの利用



# 今後の課題例：Tbps 級ノード間帯域の実現

電気スイッチのみでは困難？

- 40Gbase-KR Ethernet: 1m長、Tbps電気ルータ



20年後のスパコンNWのデバイス

- 難易度 高：全光NW(光パケットスイッチング)
- 難易度 中：全光NW(光サーキットスイッチング)
- 現状 : 複数の電気NWで帯域確保

# 概要

- スパコン内ネットワーク
  - 大学発のアイデアが盛りだくさんの多様性



- WAN(インターネット)の利用
  - スパコン内ネットワークのノード帯域と同等



# インターネットの劇的帯域改善

## WANの帯域 $\geq$ スパコン内NWの帯域

WAN(インターネット)

- 東阪間: 超100Gbps(見込)



- ミリ秒単位の遅延



スパコン内NW

- ノードバンド幅:  
TSUBAME2: 80Gbps  
T2K東大: 40Gbps

- マイクロ秒単位の遅延

- 総配線長: 数百kmも



HPC独自の **Delay Tolerant Network**

HPC分野で何が起こるか？



# まとめ

- スパコン内ネットワーク
  - 大学教員のアイデアが7年以内に実用化の世界
  - 最新デバイス, 高まる性能/電力要求へのあくなき追及が独自開発の超高速ネットワークを次々と生む
- WAN (インターネット)
  - スパコン内NWノード帯域と同程度
  - E-サイエンス, ビッグサイエンス, クラウド, グリッド
    - WANが想像を絶する大規模システムの一部へ変貌
    - 「遅延と帯域」の議論のみで終わらせるのは, もったいない

簡単な問題は先人がかなり解いてしまった。  
この激動, 複雑化を楽しめるか？