

PCクラスタのネットワーク

Interconnection Networks on PC Clusters

鯉渕 道紘、大塚 智宏、渡邊 幸之介、上樂 明也、工藤 知宏、天野 英晴
 M. KOIBUCHI, T. OTSUKA, K. WATANABE, A. JOURAKU, T. KUDOH, H. AMANO

何がわかる？

大規模計算機の主流であるPCクラスタにおける効果的なネットワーク構築法を示す。

- I) 安価なイーサネットを用いた密結合環境の実現
- II) 高性能なSANを用いた効率的な結合方式



汎用のPCクラスタを用いて実現性、効果を明確にする。

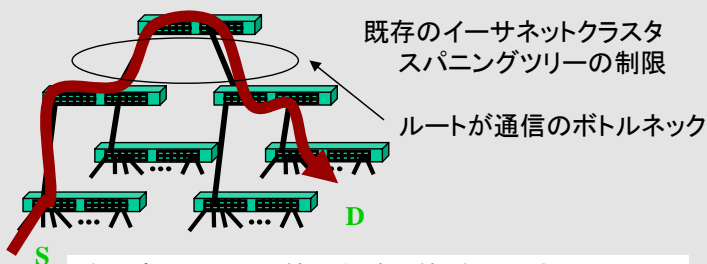
どんな研究？

大規模計算環境を構築するためにスケーラブルなネットワーク環境が不可欠である。

既存のネットワークとしては 1)コスト重視のイーサネット、2)HPCのSAN(システムエリアネットワーク)に大別できるが、本研究では両アーキテクチャにおいて並列分散処理に適した構築法を明らかにする。

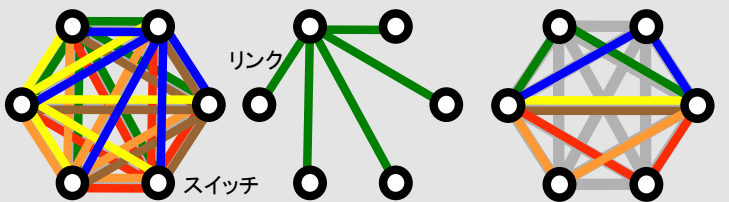
内容

研究(1) イーサネットを用いた密結合環境の構築



本研究ではVLAN技術をバンド幅向上に応用

- あらゆるトポロジが実現可能
- 完全結合(下例)、Clos網、ハイパーキューブ etc



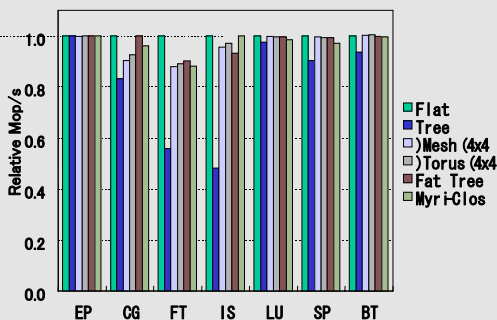
解決した/すべき課題

- ◆ 柔軟な経路導入によるフレームのデッドロック問題
- ◆ 耐故障技術としてのVLAN経路切り替え
- ◆ スケーラビリティのあるVLAN割り当て、VLAN実装

Feasibility, 評価

16台 GEスイッチ NWですら、大きな性能向上を達成

GEを用いたHPCにおいて、本技術は必須といえる。



研究(2) SANにおけるデッドロックフリールーティング、マルチキャストの提案、評価

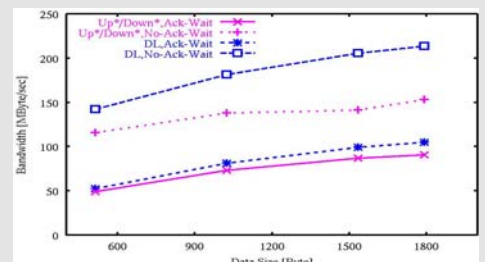


パーストラフィックの高速処理にはワームホールスイッチング技術とデッドロックフリーパケット転送が必須

本研究では、SAN向けの最高性能のデッドロックフリールーティングの論理と設計(L-Turn, DL)を行った。

Feasibility, 評価

RWC RHiNET-2 クラスタ(16 SW)において約5割のバンド幅向上を達成した。



実行時間の内訳(NPB CG, クラスA, 16SW 2Dトラス)

計算時間	通信時間(sec)					
	MPI	PM	Send Ready	パケット	NIC Ready	ACK
1.460	0.124	0.091	0.001	0.258	0.010	0.104

参考論文

◆ M. Koibuchi, K. Watanabe, T. Otsuka, H. Amano, Performance Evaluation of Deterministic Routings, Multicasts, and Topologies on RHiNET-2 Cluster, IEEE Trans. on Parallel and Distributed Systems(PDS), vol.16, no.8, pp.747--759, August 2005

◆ A. Jouraku, M. Koibuchi, H. Amano, An Effective Design of Deadlock-Free Routing Algorithms Based on 2-D Turn Model for Irregular Networks, IEEE Trans on PDS (to appear)

◆ T. Otsuka, M. Koibuchi, T. Kudoh, H. Amano, A Switch-tagged VLAN Routing Methodology for PC Clusters with Ethernet, the International Conference on Parallel Processing(ICPP'06) (to appear)