

光リングネットワークにおける λコンピューティング環境に適した 共有メモリアーキテクチャの評価

大阪大学基礎工学部情報科学科
村田研究室 久保貴司

2007/2/27

特別研究報告

1

研究の背景

- 大規模・高速計算に対する要求
- グリッドコンピューティングを利用した大規模計算
 - TCP/IPによるインターネットを用いたデータ共有
 - パケット処理や輻輳などの通信オーバーヘッドにより性能低下



- λコンピューティング環境の提案
 - WDM技術を利用してノード間に波長パスを設定
 - 高速・高信頼な通信をエンドユーザ間に提供

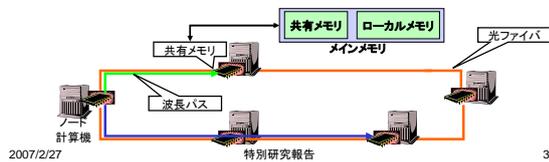
2007/2/27

特別研究報告

2

λコンピューティング環境

- 光ネットワークを利用した分散並列計算環境
 - 各ノード計算機を光ファイバで接続
 - ノード計算機間に波長パスを設定
 - 波長パスを利用した分散並列計算
 - ハードウェア分散共有メモリを実現
 - 共有メモリシステムを広域なネットワーク上に展開することが可能



2007/2/27

特別研究報告

3

研究の目的

- ネットワーク構成を考慮したλコンピューティング環境における共有メモリアーキテクチャの設計と評価
 - 設計
 - 関連研究[5]で設計された共有メモリアーキテクチャを使用
 - リングポロジに単一波長を設定
 - 本報告で設計する共有メモリアーキテクチャ
 - リングポロジに複数波長を設定
 - 経路するノードをカッスルー
 - » ノードによる処理遅延が減少
 - 評価
 - ノード計算機のモデル化と解析
 - 計算スループットを指標として評価

[5] Eiji Taniguchi, "Design and evaluation of shared memory architecture for WDM-based λ computing environment," Master's thesis, Graduate School of Information Science and Technology, Osaka University, February 2006.

2007/2/27

特別研究報告

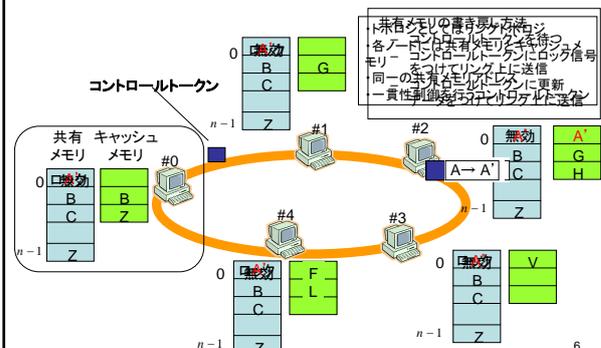
4

共有メモリアーキテクチャ

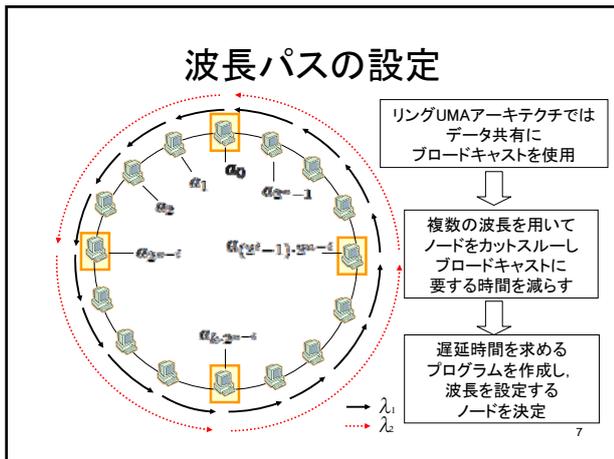
- トポロジ
 - リングポロジ
- キャッシュ一貫性制御
 - スヌープ法: 制御トークンを監視
 - 書き戻し方法: 無効化型
 - 書き戻しタイミング: ライトバック:
 - ⇒ 共有メモリへのアクセスが少ないプロトコル
- 共有メモリアクセスモデル
 - Uniform Memory Access (UMA)
 - : 任意のアドレスに対してCPUのメモリアクセス時間が同じモデル
 - Non Uniform Memory Access (NUMA)
 - : アドレスによってCPUのアクセス時間が異なるモデル

5

リングUMAアーキテクチャ



6

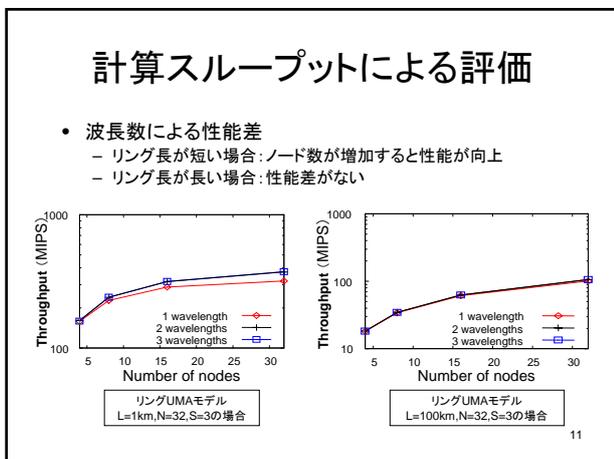
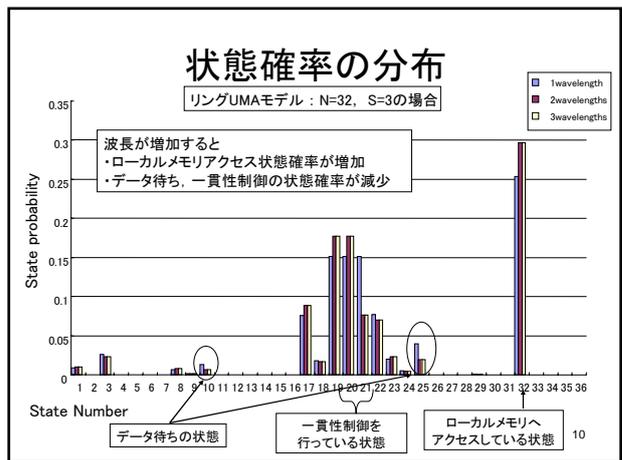
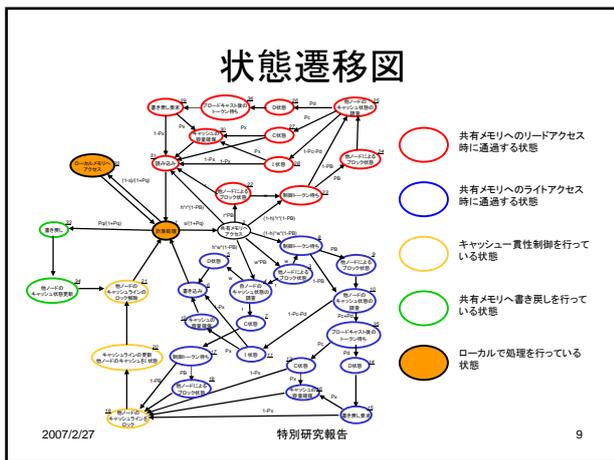


モデル化と解析

- 共有メモリアーキテクチャの動作に基づき状態遷移図を作成
 - ノード計算機のCPUの状態に着目
 - 計算処理を行っている状態
 - ローカルメモリにアクセスしている状態
 - キャッシュ一貫性制御を行っている状態 など
- 解析にはセミ・マルコフ過程を利用
 - 状態の滞在時間を任意に設定可能
 - 各状態の定常状態確率を導出

評価のための数値例	パラメータ
- CPUクロック : 2 [GHz]	- N : ノード数
- ネットワークインターフェイスのフレーム処理時間 : 1 [μ s]	- L : リング長
	- S : CPUのメモリアクセスにおける共有キャッシュへのアクセス頻度 (10^3)

8



まとめと今後の課題

- まとめ
 - リングトポロジにおける複数波長を用いた共有メモリアーキテクチャの設計と評価
 - リングUMAアーキテクチャ
 - リングNUMAアーキテクチャ
 - 波長数による性能の違い(リングUMA, NUMA共に)
 - リング長が短い場合: 共有メモリへのアクセス頻度が高いと性能差が生じ、ノード数が多いほど複数波長を用いたアーキテクチャの方が性能はよくなる
 - リング長が長い場合: 共有メモリへのアクセス頻度、ノード数によらず性能の差がない
- 今後の課題
 - メッシュトポロジにおいて波長数を制限した共有メモリアーキテクチャの性能評価

2007/2/27 特別研究報告 12