



リソース分離型マイクロデータセンターのネットワーク構成

大阪大学 大学院情報科学研究科
村田研究室M2 反町 将

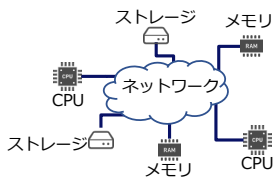
研究背景

- クラウドでのデータ処理は遅延が大きい
- リアルタイム性の高いデータ処理を行うため、エッジに配置するマイクロデータセンター(μDC)が提案
 - 資源が限られており、有効活用する必要



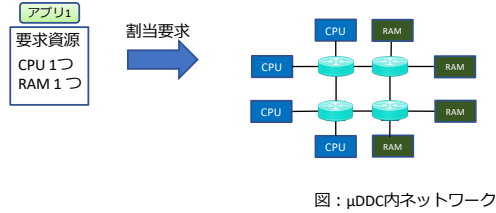
研究背景

- マイクロデータセンターの資源利用効率を上げるため、リソース分離型マイクロデータセンター(μDDC)が提案
 - CPU・GPU・メモリ・ストレージ等の資源を分離してネットワークに接続
 - 各アプリケーションに必要な分だけ資源を割り当てることが可能
 - 資源が独立しており、資源ごとの入れ替えや柔軟なスケーリングが可能



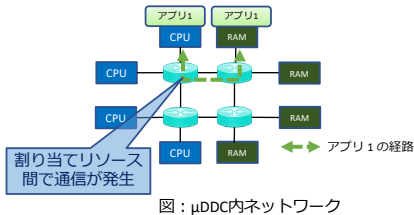
研究背景

- μDDC内で複数のアプリケーションを同時実行する流れ
 1. アプリケーション開始前に資源割り当ての要求が到着
 2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



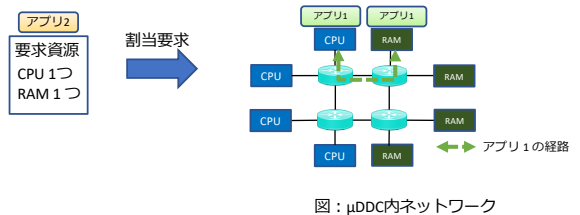
研究背景

- μDDC内で複数のアプリケーションを同時実行する流れ
 1. アプリケーション開始前に資源割り当ての要求が到着
 2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



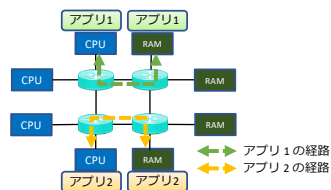
研究背景

- μDDC内で複数のアプリケーションを同時実行する流れ
 1. アプリケーション開始前に資源割り当ての要求が到着
 2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



● μDDC内で複数のアプリケーションを同時実行する流れ

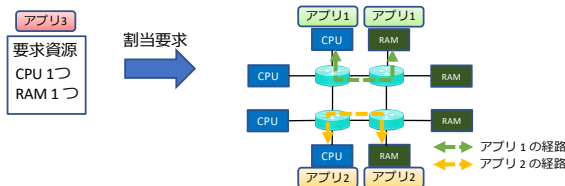
1. アプリケーション開始前に資源割り当ての要求が到着
2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



図：μDDC内ネットワーク

● μDDC内で複数のアプリケーションを同時実行する流れ

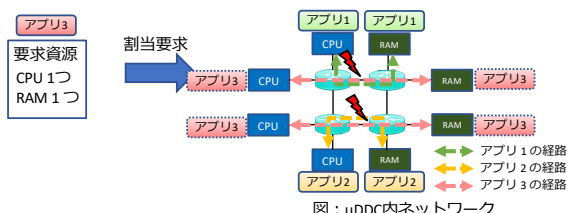
1. アプリケーション開始前に資源割り当ての要求が到着
2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



図：μDDC内ネットワーク

● μDDC内で複数のアプリケーションを同時実行する流れ

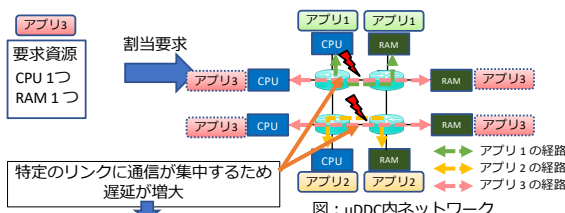
1. アプリケーション開始前に資源割り当ての要求が到着
2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



図：μDDC内ネットワーク

● μDDC内で複数のアプリケーションを同時実行する流れ

1. アプリケーション開始前に資源割り当ての要求が到着
2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



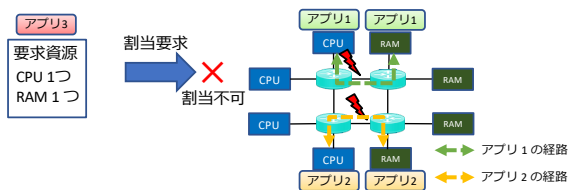
特定のリンクに通信が集中するため遅延が増大

図：μDDC内ネットワーク

要求の時間制約を満たすことが困難に

● μDDC内で複数のアプリケーションを同時実行する流れ

1. アプリケーション開始前に資源割り当ての要求が到着
2. 随時、要求を満たす計算資源・メモリ資源・NW資源を割り当て



図：μDDC内ネットワーク

資源割り当て手法・μDDC内のネットワーク構成がアプリケーションの処理性能に大きな影響を与える

● 研究目的

- μDDC 内のネットワークに適したネットワーク構成を検討
- 計算資源とメモリ資源の配置も含めて検討

● アプローチ

1. μDDC 内のネットワーク構成に求められる要件について議論
2. その要件を満たしているのか確認可能な指標を導入
3. 既存のデータセンター向けトポロジに計算資源とメモリ資源を接続したネットワーク構成に導入した指標を適用し、ネットワーク特性を分析
4. 資源割り当てシミュレーションによる評価を行い、ネットワーク特性と収容可能なアプリケーション数の関係について考察

1. 計算資源・メモリ資源間の低遅延な接続

- 計算資源が必要な情報をメモリ資源から取得するのに時間がかかる場合、アプリケーションの処理速度が低下

2. 多量の計算資源・メモリ資源間の通信の収容

- 複数のアプリケーションを同時に実行する際に発生する多量の計算資源・メモリ資源間の通信を収容可能な帯域幅を確保する必要

3. 低インシャルコスト

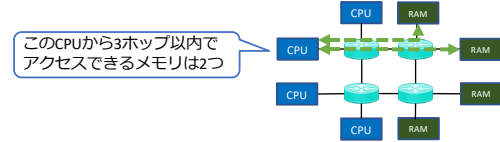
- ネットワークが低コストで構築可能

● 要件1：計算資源・メモリ資源間の低遅延な接続

- ホップ数が多いと伝搬遅延が増加

● 指標1：各計算資源からkホップ以内のメモリ資源数

- この指標が大きいネットワーク構造では、各計算資源にとってホップ数が少ない(=低遅延で通信可能な)メモリ資源が多量



図：μDDC内ネットワーク

● 要件2：多量の計算資源・メモリ資源間の通信の収容

- 複数のアプリケーションを同時実行した際、計算資源・メモリ資源間の通信が特定のリンクに集中することを避ける必要

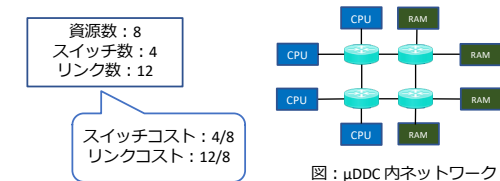
● 指標2：重み付きエッジ媒介中心性

- エッジ媒介中心性：全ノード間で通信が発生する場合にパケットが特定のリンクを通る頻度を示す指標
1. μDDCでは距離の近い計算資源とメモリ資源の間でのみ通信が発生するため全資源間で通信が発生する可能性は極めて低い
 2. 資源ペアの種類と距離の近さの順位を重みとして与えることで資源ペア間の通信の有無を考慮するようアレンジ
- 指標が大きいほど、複数のアプリケーションを同時に実行した際に通信が特定のリンクに集中しボトルネックとなってしまうことを示す

● 要件3：低コスト

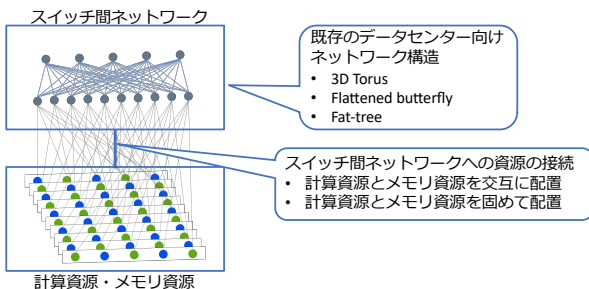
● 指標3：スイッチコスト・リンクコスト

- スwitchコストの定義：資源1つあたりのスイッチ数
- リンクコストの定義：資源1つあたりのリンク数

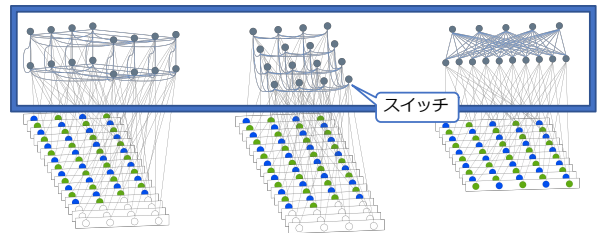


図：μDDC内ネットワーク

● 既存のデータセンター向けネットワーク構造に計算資源とメモリ資源を接続

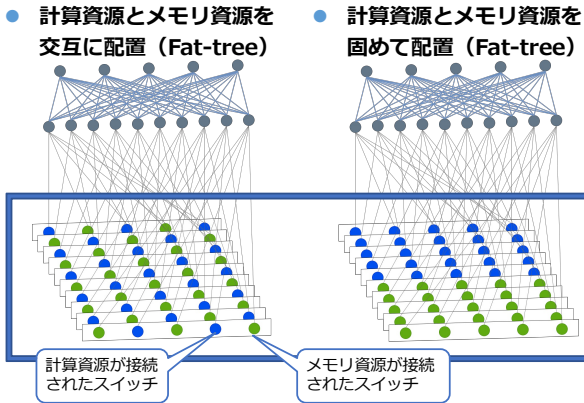


● 3D Torus ● Flattened butterfly ● Fat-tree

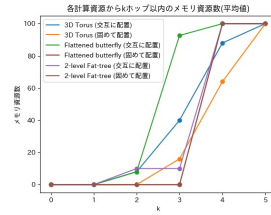


● 揃えた条件

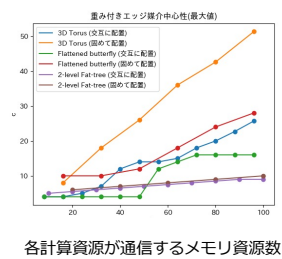
- CPU資源数：100・メモリ資源数：100・スイッチのポート数：10



- 指標1：各計算資源からkホップ以内のメモリ資源数
 - 指標が大きいくほど、計算資源から低遅延にアクセス可能なメモリ資源が多いことを示す
- 資源の配置方法の比較
 - 「交互に配置」が優れる
- スイッチ間NWの比較
 - Flattened butterflyが優れる



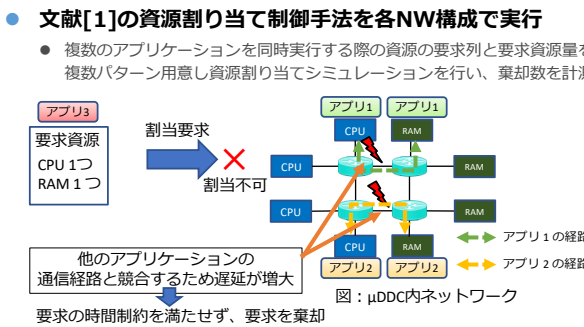
- 指標2：重み付きエッジ媒介中心性
 - 指標が大きいくほど、複数のアプリケーションを同時に実行した際に通信が特定のリンクに集中しボトルネックとなってしまうことを示す
- 資源の配置方法の比較
 - 「交互に配置」が優れる
 - 計算資源とメモリ資源間の通信が短いホップ数の経路に收容され、各リンクを経由する通信量を抑えることができるため
- スイッチ間NWの比較
 - 通信が発生する資源ペアが近隣のものに限定されている場合、Flattened butterflyが優れる



- 指標3：リンクコスト・スイッチコスト
 - 3D TorusやFlattened butterflyが低コスト
 - 全スイッチが直接に計算資源やメモリ資源と接続するため

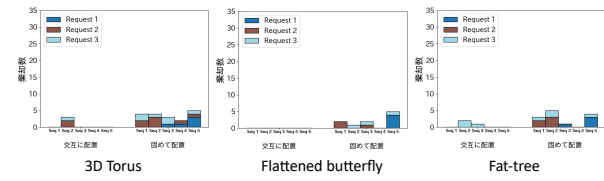
	3D Torus	Flattened butterfly	Fat-tree
リンクコスト	0.4375	0.4375	0.5000
スイッチコスト	0.0625	0.0625	0.0800

表：リンクコスト・スイッチコスト



[1] 生駒昭繁, 大下裕一, 村田正幸, “リソース分離型マイクロデータセンターにおける将来の資源要求に備えた資源割り当て手法,” 電子情報通信学会技術研究報告 (IN2021-36), pp.31-36, March 2022.

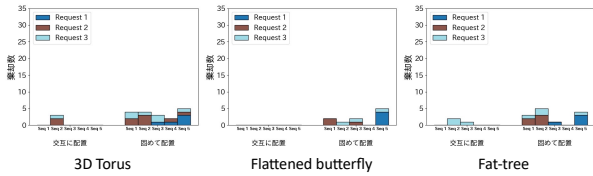
- 資源の配置方法の比較
 - 計算資源・メモリ資源を固めて配置の方が棄却数が多い
 - ネットワーク特性の分析結果と合致
 1. 「固めて配置」すると計算資源とメモリ資源間のホップ数が多くなる
 2. 遅延の増大・未使用のリンク数の減少
 3. 十分な低遅延でのアプリケーションの実行が可能な資源割り当てができなくなる



図：要求列と必要資源量ごとの各ネットワーク構成の棄却数

● スイッチ間ネットワークの比較

- Flattened butterflyで「交互に配置」した場合、棄却が発生していない
- ネットワーク特性の分析結果と合致
 1. 各計算資源において十分な帯域・低遅延で接続可能なメモリ資源の候補が多い
 2. 柔軟な資源割り当てが可能
 3. 多数の要求が到達した際にも要求処理時間を満たす資源割り当てが可能



図：要求列と必要資源量ごとの各ネットワーク構成の棄却数

● 要件2：多量の計算資源・メモリ資源間の通信の収容

- 複数のアプリケーションを同時実行した際、計算資源・メモリ資源間の通信が特定のリンクに集中することを避ける必要

● 指標2：重み付きエッジ媒介中心性

- エッジ媒介中心性：全ノード間で通信が発生する場合にパケットが特定のリンクを通る頻度を示す指標

$$c(e) = \sum_{s,t \in V} \left(\frac{\sigma(s,t|e)}{\sigma(s,t)} \right)$$

1. μDDCでは距離の近い計算資源とメモリ資源の間でのみ通信が発生するため全資源間で通信が発生する可能性は極めて低い
2. 資源ペアの種類と距離の近さの順位(k)を重み(w)として与えることで資源ペア間の通信の有無を考慮するようアレンジ

$$c(e,k) = \sum_{s,t \in V} \left(\frac{\sigma(s,t|e)}{\sigma(s,t)} \cdot w_{s,t}(k) \right)$$

● まとめ

1. μDDC内ネットワーク構成に求められる要件と対応する指標を検討
2. 既存のデータセンター向けネットワーク構造を基にしたネットワーク構成について、検討した指標を用いてネットワーク特性を分析
3. 資源割り当てシミュレーションを行い、検討した特性を持つネットワークがより多くのアプリケーション要求を収容可能であることを示した

● 今後の課題

- 今回検討を行った指標をもとに、よりμDDC内ネットワークに適したネットワーク構成について検討を進める