

ジョブの優先処理を考慮したゲーム理論によるクラウドレット間のオフロード最適化

芝浦工業大学 横田侑紀



●背景と目的

●背景

- クラウドコンピューティングの普及
 - 遅延削減を目的としたエッジコンピューティングへの関心の高まり
- クラウドレットを利用した低遅延でのオフローディング[1]
- ★ Local Area Network内のルータなどに設置するエッジサーバ
 - 遅延に敏感なモバイルアプリケーションに有効
 - 反面、規模の小ささからリソース不足が懸念

他クラウドレットへジョブを受け渡すオフロードによる遅延低下

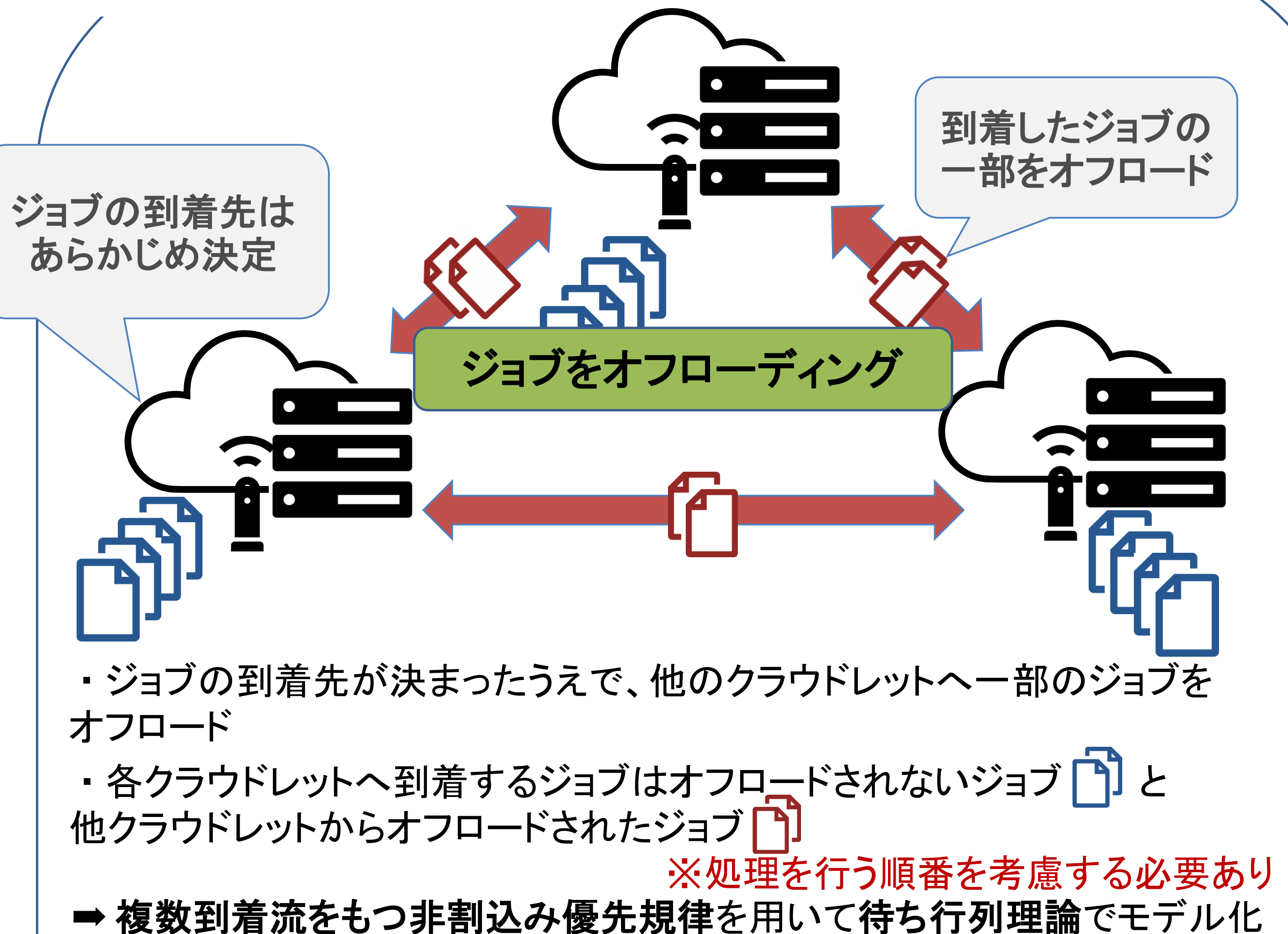
●既存手法について

- ジョブが許容遅延を満たす中でクラウドレット間のオフロードを行い負荷を可能な限り分散させる手法 [2]
 - ジョブのオフロードのモデル化が不十分
 - オフロードによるクラウドレットのメリットが考慮できていない

●目的

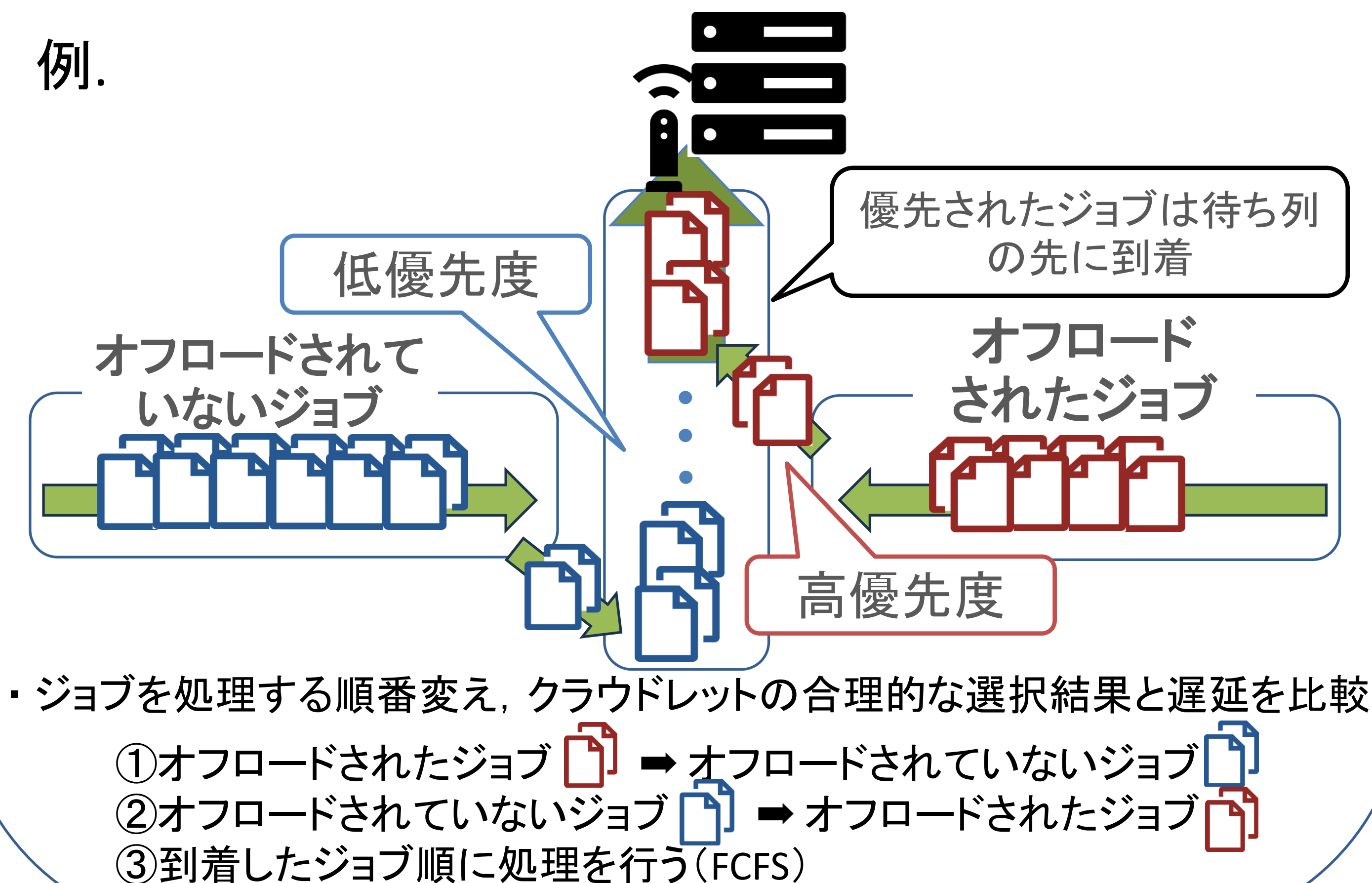
オフロードによるデメリット(遅延とコスト)とメリット(処理数の増加)を考慮したクラウドレットの最適選択を導出

●提案法のモデル化



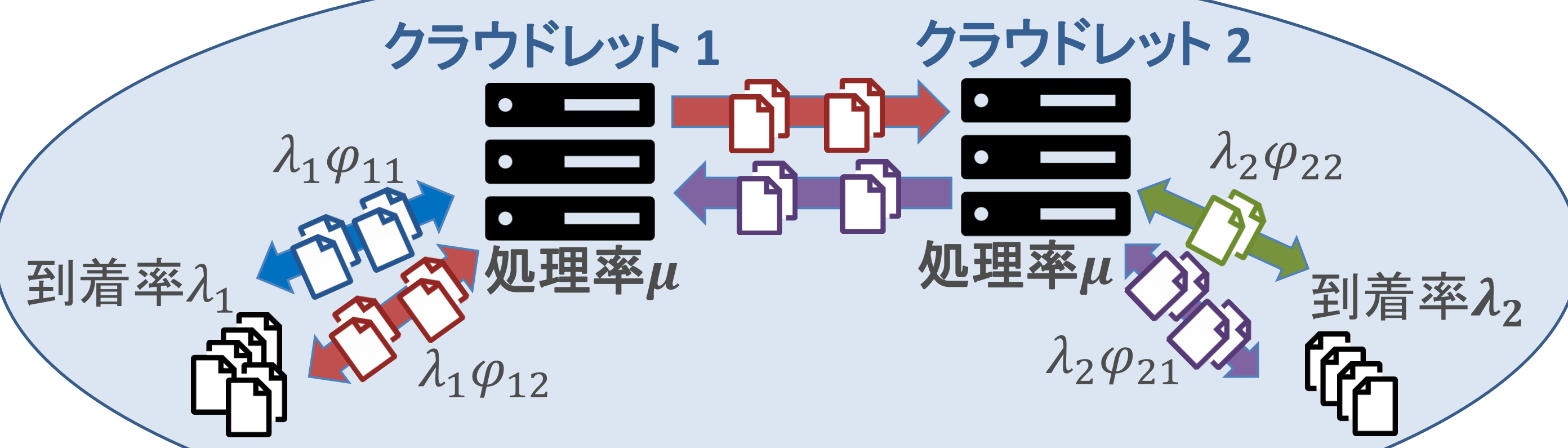
●複数到着流の非割込み優先規律

例.



●ゲーム理論による評価

●2つのクラウドレット間でのオフロードをモデル化



リトルの公式を用いて遅延時間を導出し、以下の利得関数を定義

$$U_1 = -\rho_{12} + \rho_{21} + \alpha\rho_1 + \beta \sum_{j=1}^N (D_q - E_{1j})\rho_{1j}$$
$$U_2 = -\rho_{21} + \rho_{12} + \alpha\rho_2 + \beta \sum_{j=1}^N (D_q - E_{2j})\rho_{2j}$$

ρ_1, ρ_2 : クラウドレットのオフロード後の利用率 α, β : 重み

ρ_{12}, ρ_{21} : クラウドレットがオフロードするジョブの利用率

$E_{11}, E_{12}, E_{21}, E_{22}$: 各到着流の遅延 D_q : 許容遅延時間

・利得関数の構成:

- オフロードしたジョブの量 (-)
- オフロードを受けたジョブの量 (+)
- 自クラウドレットで処理するジョブの総量 (+)
- 自クラウドレットに到着するジョブと許容遅延との差 (+/-)

ゲーム理論のナッシュ均衡によってすべてのクラウドレットが利益を得るために最適な戦略を導出

●結果

- オフロードが起きるような環境においてナッシュ均衡となる最適オフロード割合 ϕ_1, ϕ_2 を求める

| | FCFS | オフロードなし | 処理順①において遅延最小化 |
|----------|---------|---------|---------------|
| U_1 | 0.331 | 0.370 | 0.376 |
| U_2 | 0.0181 | -0.0146 | -0.00514 |
| ϕ_1 | 0 | 0 | - |
| ϕ_2 | 0.2 | 0.2 | - |
| E_{11} | 0.00268 | 0.00171 | 0.00313 |
| E_{12} | 0 | 0 | - |
| E_{22} | 0.00357 | 0.00357 | 0.00357 |
| E_{21} | 0.00371 | 0.00553 | 0 |

● パラメータ設定

$\lambda_1 = 400$ jobs/s
 $\lambda_2 = 900$ jobs/s
 $\mu = 1000$ jobs/s
 $D_q = 5$ ms
 $\alpha = 0.1$
 $\beta = 100$

・オフロードされたジョブの優先度が低い場合のみナッシュ均衡で許容遅延を上回る結果

・最小遅延を達成しているのも同様

・遅延を最小化する場合だと最適オフロード割合が変化

⇒ ゲーム理論を用いたことで利得の偏りが小さくなった

●まとめと今後の課題

●まとめ

・非割込み優先規律を用いたモデル化によりオフロード後の処理の仕組みを解析

●今後の課題

・ゲーム理論を用いたことによる効果の明確化

●参考文献

[1] E. Wong, M. Pubudini Imali Dias, and L. Ruan, "Predictive resource allocation for tactile internet capable passive optical lans," Journal of Lightwave Technology, vol. 35, no. 13, pp. 2629–2641, 2017.

[2] Y. Yokota, and S. Miyata, "Latency Focused Load Balancing Method between Cloudlets using Game theory," in Nonlinear Theory and Its Applications, IEICE, vol. E-15N, no. 2, Apr. 2024.