エッジ資源管理の負荷軽減に向けた 分割可能負荷スケジューリングシステムアーキテクチャの検討

齋藤光喜[1]　小林輝夫[2]　不破茂秀[2]　菅谷みどり[1]

概要：本研究では，大規模な IoT/GW を管理するエッジ資源管理システム EdgeRM の負荷軽減と，IoT/GW の遊休計算資源の有効活用を目的として，分割可能負荷スケジューリングに基づく階層型負荷分散アーキテクチャを検討する．提案方式では，強化学習を用いてタスク分割数と使用台数を動的に決定し，IoT/GW の処理性能に応じて分割タスクを割り当てる．今後は実環境への実装と評価を通じ負荷軽減効果および処理性能への影響を検証する．

キーワード：エッジコンピューティング，分割可能負荷スケジューリング，EdgeRM，仮想発電所

1. 研究背景

近年，蓄電池や電気自動車などの分散型エネルギー資源（DER: Distributed Energy Resources）を統合的に制御する仮想発電所（VPP: Virtual Power Plant）が注目されている．VPP は，DER と直接接続された計測情報の収集と制御指示の中継を担うゲートウェイ端末（IoT/GW）を介して，DER の出力調整を行う．管理対象の DER 数の増加に伴い，IoT/GW から送信されるテレメトリの収集・処理負荷も増大するため，これらを効率的に処理する計算基盤が求められる[1].

特に大規模な IoT/GW の管理では負荷増大による遅延の影響が増大することから，エッジコンピューティングの活用が期待される．エッジコンピューティングは，利用者やデータ発生源の近傍に配置された計算資源で処理を実行する方式であり，通信遅延の低減やクラウドへの処理集中の緩和に利点がある[2]．しかし，エッジサーバの計算資源にも限りがあり，管理対象となる DRE の増加に伴う処理負荷の増大が課題となる．また，サーバの増設による計算資源の拡張は，設備・運用コストの増加を伴う[3].

一方，VPP を構成する IoT/GW は，テレメトリの送信や DER への制御指示を担うものの，常時すべての計算資源を使用しているわけではない．そこで，遊休計算資源を活用してエッジサーバの処理の一部を分散実行することで，サーバを増加させることなく可用性の向上が見込まれる．本方式であれば，サーバの計算資源を増強せずにシステムの可用性向上が実現するため，コスト低減も期待できる．しかし，こうした方法は十分に提案されていない．

2. 先行研究と課題

大規模な IoT/GW の管理に対し， IoT・Fog・Cloud といった異なるコンピューティング階層間における負荷分散やタスクオフロードに関する数多くの研究が進められている[4]–[7]. これらは遊休計算資源を活用できる可能性がある．しかし，これらの研究は主にコンピューティング階層間での処理の最適配置を対象としており，多数の IoT/GW を管理する資源管理システム自身の負荷を，管理対象機器の遊休計算資源へ分散する方式とは言えず，本研究の目的と提案は十分に検討されているとはいえない．

著者らは，ヘテロジニアスな計算資源環境における最適な資源割当てを目的とし，EdgeRM(Edge Computing Resource Management system) を開発してきた[6]. また，EdgeRM を基盤として，VPP 向けのアーキテクチャを提案した [7]．本アーキテクチャでは，既存の多段階サーバークライアント構造を EdgeRM に置き換えることでオーバーヘッドを低下させ数万台規模の IoT/GW の収容を実現した．同研究では，8 コアの Xeon CPU および 16 GB のメモリを搭載したオンプレミス環境にて，3.5 万台規模の IoT/GW を CPU 使用率 70%以下，メモリ使用率 30%以下で管理可能な結果を得るなど，大量のノードを低コストで管理できる利点がある．

一方で，本研究の課題となる遊休計算資源を活用してエッジサーバの処理の一部を分散実行するなどの可用性の向上手法は十分検討されていなかった．EdgeRM においては管理対象となる IoT/GW 数の増加に伴い，システムの CPU 使用率が上昇する課題があり，遊休資源の活用を含めた可用性の向上が課題である．

3. 研究目的・提案

本研究は，課題を解決するために，遊休 IoT/GW への分散処理の実現を目的とする．目的を達成するために，EdgeRM 内部の最適化機能とスケジューリング機能を階層化したアーキテクチャを提案する．

3.1 概要

本アーキテクチャは，EdgeRM 自身の処理負荷の抑制と IoT/GW の遊休計算資源の有効活用を目的として，EdgeRM の処理の一部を IoT/GW へオフロードを行う方法を明らかにする．オフロードの目的は，個々の単体性能が限られている IoT/GW において，EdgeRM で行う処理を実行可能に

1 芝浦工業大学
Shibaura Institute of Technology
2 株式会社エナリス
ENERES Co.,Ltd.

ⓒ2026 Information Processing Society of Japan

するため，本研究では，1 つの負荷を複数に分割するスケジューリング（DLS: Divisible Load Scheduling）手法を採用し，処理を分割する．また，複数の IoT/GW に処理をオフロードし，並列実行するスケジューリングシステムのアーキテクチャを提案する．

本システムは実用的な場面を想定することから，オフロード先の IoT/GW の処理性能，及び利用可能台数は動的に変化するものとする．このことから，その時点で利用される IoT/GW の処理性能や，利用台数に応じて変化する処理時間，オーバーヘッドを考慮した強化学習を行う．そこで得られた最適解に基づき最適に，遊休資源である IoT/GW へ処理をオフロードし，並列処理を行うものとする．

3.2 設計

提案する階層型負荷分散アーキテクチャを図 1 に示す．本アーキテクチャは，L0（実行層），L1（スケジューリング層），L2（最適化層）の 3 階層で構成される．

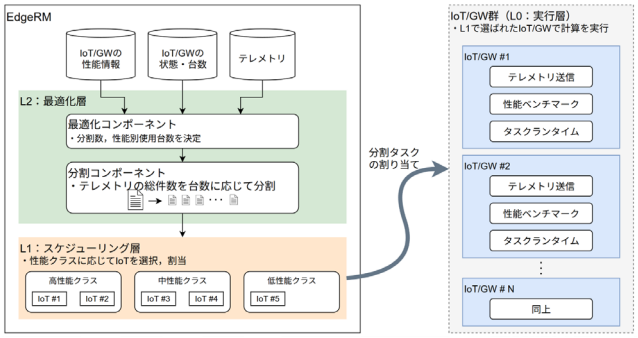


図 1　提案アーキテクチャ全体図

EdgeRM は IoT/GW から受信したテレメトリに基づき集計タスクを生成し，L2 で決定された分割数に従って処理対象を分割する．各 IoT/GW で実行された部分集計の結果は EdgeRM へ返送され，統合されることで最終的な集計結果を得る．

L0 は EdgeRM の管理対象である IoT/GW 群から構成され，EdgeRM から割り当てられた分割タスクを実行し，その結果を EdgeRM へ送信する．また，実行する処理内容は動的に更新可能であり，固定的な処理に限定されない．

L1 は EdgeRM 内部に配置されるスケジューリング層であり，L2 で決定された性能クラス別の使用台数に基づき，FIFO や重み付きラウンドロビン（WRR）などの方式を用いて，分割タスクを各 IoT/GW へ割り当てる．

L2 は EdgeRM 内部の最適化層であり，利用可能な IoT/GW の処理性能と台数に基づき，強化学習を用いてタスクの分割数および使用台数の動的な最適化を図る．これにより，異種計算資源を活用した並列処理により，オフロードに伴う処理時間の増大を抑制する．先行する共同研究では，複数の IoT/GW へのタスク分割と並列処理に関する基礎検討が行われており[8]，本研究ではその成果を基盤として EdgeRM への適用を検討する．

本研究では，オフロード対象としてテレメトリの集計処理を想定する．テレメトリは周期的に収集されるため，次の処理周期までに集計を完了する必要がある．そこで，最適化処理自体の計算負荷も考慮し，初期検討では探索と活用を確率的に切り替える ε-greedy 法を用いた多腕バンディット[9]を採用する．L1 のスケジューリングでは，各 IoT/GW の処理性能を考慮する．IoT/GW が計算資源として利用可能になる際に，各機器でベンチマークを実行し，その結果を EdgeRM へ送信することで，処理性能に応じたクラス分類を行う．初期検討では，外乱による影響を抑えてアーキテクチャの有効性を評価するため，ベンチマーク結果に基づく性能クラスを固定する方式とした．

4. 今後の展望

今後は，提案アーキテクチャにて IoT/GW への処理のオフロードによる EdgeRM の負荷軽減効果，IoT/GW の計算資源利用状況，および処理時間を評価する．さらに，EdgeRM の動作環境として，オンプレミス，クラウド，キャリア通信網に直接接続されたクラウドの 3 環境を用い，通信環境の違いが負荷分散性能に与える影響を比較することで，提案方式の有効性を検証する．

謝辞　本研究は，株式会社エナリスとの共同研究のもと実施されました．心より感謝申し上げます．

参考文献

[1] J. F. Venegas-Zarama, J. I. Muñoz-Hernández, L. Baringo, P. Díaz-Cachinero, and I. de Domingo-Mondéjar, “A review of the evolution and main roles of virtual power plants as key stakeholders in power systems,” IEEE Access, vol. 10, pp. 47937‑47964, 2022, doi: 10.1109/ACCESS.2022.3171823.

[2] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

[3]株式会社エナリス: MEC-RM 技術による低圧リソースの最適分散制御手法に関する研究を開始 芝浦工業大学とエナリスの共同研究，入手先 https://www.eneres.co.jp/news/20240529.html，（参照日 2026-10-21）

[4] A. Mahapatra, S. K. Majhi, K. Mishra, R. Pradhan, D. C. Rao, and S. K. Panda, “An energy-aware task offloading and load balancing for latency-sensitive IoT applications in the fog-cloud continuum,” *IEEE Access*, vol. 12, pp. 14334–14349, 2024, doi: 10.1109/ACCESS.2024.3357122.

[5] S. M. Kazemi, S. Ghanbari, M. Kazemi, and M. Othman, “Optimum scheduling in fog computing using the divisible load theory (DLT) with linear and nonlinear loads,” *Computer Networks*, vol. 220, Art. no. 109483, Jan. 2023, doi: 10.1016/j.comnet.2022.109483.

[6]Y. Li and M. Sugaya, “Proposal of an edge serverless computing platform with an asynchronous communication model,” in Proc. Asia Pacific Conf. Robot IoT Syst. Develop. Platform (APRIS 2023), 2023, pp. 75–76. [Online]. Available: https://ipsj.ixsq.nii.ac.jp/records/231593. Accessed: Sep. 25, 2026.

[7] K. Saito, Y. Hasuhara, T. Kobayashi, S. Fuwa and M. Sugaya, "A Scalable Infrastructure for Distributed Energy Resources Based on MEC-RM Architecture," 2026 International Power Electronics Conference (IPEC-Nagasaki 2026 - ECCE Asia), Nagasaki, Japan, 2026, pp. 1-8, doi: 10.23919/IPEC-Nagasaki2026-EC64663.2026.11596631.

[8] 株式会社エナリス，エナリスと芝浦工業大学、低圧アグリゲーション高度化に向けた MEC-RM 拡張技術「EdgeRM」を用いた EdgeAI 自律分散制御手法の共同研究を開始（オンライン），入手先 < https://www.eneres.co.jp/news/20260909.html> ，（参照日 2026-10-21）

[9]R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

ⓒ2026 Information Processing Society of Japan