

エッジコンピューティングを用いた映像監視機能を備える ローカルIoTシステム

馬場 伸一

第一工科大学 情報・AI・データサイエンス学科 (〒 899-4332 鹿児島県霧島市国分中央 1-10-2)

Local Internet of Things (IoT) system with video surveillance using edge computing

Shinichi Baba

Department of Information, Artificial Intelligence and Data Science, Daiichi Institute of Technology

Abstract :

Digital transformation (DX) improves convenience and productivity in work and daily life supported by rapid advances in Internet of Things (IoT) technology and Artificial Intelligence (AI). As one of such IoT service, advanced video surveillance systems using security cameras are easily accessible these days, and are widely applied in fields such as crime prevention, traffic management, monitoring, and smart homes. We have developed a prototype of an intelligent video surveillance system based on the edge computing architecture that achieves real-time moving detection and face recognition better than a cloud computing based system. From the perspective of system design concept, the edge computing architecture is equivalent to a local IoT server. Therefore, this study clarifies an extended local IoT system architecture involving the edge computing based on the knowledge gained through the development.

Key words : Internet of Things (IoT), Edge computing, Intelligent video surveillance, Face recognition, Message Queuing Telemetry Transport (MQTT)

1. はじめに

仕事や生活のデジタル化、それによる利便性、生産性の向上、つまりデジタルトランスフォーメーション (DX) への期待が高まっている [1]。

その中で、Internet of Things (IoT) 技術と人工知能 (Artificial Intelligence, AI) の急速な発展を背景に、監視カメラによる高度映像監視システムは防犯、交通管理、見守り、スマートホームなどの分野で広く応用され、容易に利用可能となっている。映像監視システムが容易に利用可能なのは、監視カメラによる撮影の設定や動画の保存・分析がクラウドコンピューティングにより実行されるように設計されており、利用者は購入、或いは、レンタルなどで入手した監視カメラを設置し、インターネットに接続するだけで、使用可能

となるためである。

しかし、このクラウドコンピューティングに基づく従来の映像監視システムは、遠隔にあるクラウドで映像が処理されるため、応答が遅くなり、さらにイベントが重なった場合に誤検知する確率が高いという課題がある。また、システムコストの観点で効率が悪く、帯域コスト、ネットワークレイテンシ、データプライバシーの面でも制限がある。

この課題を解決するため、本学では郑と濱崎が、エッジコンピューティング環境における映像解析、イベント処理、通知および監視機能を統合した分散型映像監視システムを設計・実装して機能を確認し、従来の映像監視システムに対するエッジコンピューティングアーキテクチャの有効

性と実用性を検証した [2]。

エッジコンピューティングは、計算タスクをネットワークエッジに落とし込み、データ源の近接においてリアルタイム処理で行うことで、クラウドコンピューティングが内包する前述の課題を効果的に解決できる。このため、データ量の増大や AI によるデータの高度な解析を必要とする IoT システムの重要な発展方向となっている。例えば、最新の 5G 携帯ネットワークではエッジコンピューティングを提供できる Multi-access Edge Computing (MEC) を 5G Radio Access Network (5G RAN) やローカル 5G の中に備えるアーキテクチャが用意されている。また、Cisco 社は同様な効果を狙った分散型のエッジコンピューティングのアイデアをフォグ (Fog) コンピューティングという名称で検討している [3]。このような背景を参考にして、郑らは、インテリジェント映像監視システムへのエッジコンピューティング適用について研究を行った。

検証のための試作システムは、ディープラーニングによる物体検出、軽量なデータ転送・配信に用いられる Message Queuing Telemetry Transport (MQTT)[4] メッセージキュー、モバイル向けプッシュ通知、企業向け監視プラットフォームを統合し、映像取得・AI 分析・イベント処理からマルチチャンネル通知に至るまでの機能を備え、その全プロセスを自動化した監視ソリューションを構成する。また、その実装において AI オブジェクト認識エンジンを用いることでシステムの有効性を、そして本システムをクラウドコンピューティングに基づく従来のシステムとコスト面で対比分析を行うことで実用性を検証している。しかし、試作システムは、開発容易性や研究中の機能追加などから、構成が複雑になってしまい、再現が難しい課題が生じていた。

そこで本研究では、郑らの研究で得られた成果と知見を基に、これまで本学で構築していたローカル IoT サーバシステム [5], [6] にエッジコンピューティングを取入れるシステムアーキテクチャについて検討する。機能の整理を行い、ネッ

トワーク要件と構成なども見直した結果、簡潔でかつ、エッジコンピューティングの効果が得られるシステムアーキテクチャが明らかになった。

以下、第 2 章では郑と濱崎が開発した映像監視システムの概要について述べ、第 3 章で、その成果を従来のローカル IoT サーバに適用し、エッジコンピューティングを用いるシステムアーキテクチャについて検討する。最後に第 4 章で、本研究をまとめる。

2. エッジコンピューティングを用いた映像監視システム

本章では、郑と濱崎が試作した映像監視システムについて述べ、更にその経済性・実用性の分析結果についてまとめる。

2.1 試作システムの特徴

映像監視システムに必要な、映像解析、イベント処理、通知および監視機能を備える分散型映像監視システムを、エッジコンピューティング環境で実現するにあたり、以下の課題解決や特徴を持つように設計する。

1. システム有効性の検証：AI オブジェクト認識エンジンとしてオープンソースソフトウェアの Frigate[7] を用いる分散型監視システムとして実働させ、提案するシステムアーキテクチャの有効性を迅速に検証する。
2. 統一バックエンドサービスアーキテクチャの構築：モジュール化設計を採用し、マルチデバイスアクセスをサポートする。
3. リアルタイムイベント処理メカニズムの実装：MQTT ブローカを用いたイベント駆動アーキテクチャにより、リアルタイムイベント処理を実現させる。
4. マルチチャンネル通知体系の確立：Google 社のオープンな Firebase Cloud Messaging (FCM)[8] を統合し、クロスプラットフォームモバイルプッシュ通知を実現させる。
5. モバイルアプリケーションの開発：Android ネイティブアプリを構築し、直感的なユー

ザーインターフェースを提供する。

6. システムレベル監視の実現：統合監視プラットフォームを統合し、システムヘルスマonitorリングを実現させる。
7. 高拡張性：モジュール化アーキテクチャと、それに基づくマルチカメラ同時処理、マルチユーザーアクセス、水平スケーリングをサポートする。
8. 信頼性：二重認証メカニズム、データベースイベント永続化、自動故障回復などのセキュリティ機能を実現させる。

映像監視システムの試作では、エッジコンピュータとして Intel 社 N100 アーキテクチャの PC を使用し、Frigate が提供するオブジェクト認識のうち動き検出を利用した侵入検知と人物の顔認識を利用する。1 台の監視カメラからの動画のリアルタイム映像解析には、十分余裕のある構成である。動画情報はエッジコンピュータ内に保存され、認識結果のみメッセージの形で MQTT ブローカ経由でシステム内の各機能に通知される。バックエンドサービス機能は、エッジコンピュータに設定した仮想 PC に実装する。このバックエンドサービスを用いて、認識結果メッセージを Android 端末へプッシュ通知を行う。Google 社の FCM はマルチプラットフォーム／デバイスをサポートしているため、これにより種々のモバイル端末へポップアップするメッセージ（ノティス）を送出することができる。試作システムでは、Android ネイティブアプリを作成し、Android 端末でのメッセージ受信・ポップアップの実働を確認する。また、これらの機能連携において、必要に応じてセキュリティ機能を設定し、安全な情報共有、メッセージ送出手を実現させる。

2.2 試作システムの経済性分析

エッジコンピューティングアーキテクチャを採用した試作システムの必要性と優越性を検証するため、主流のクラウド AI ビデオ分析サービスを対比対象として選定し、経済的実現可能性、シス

テム信頼性、データプライバシーの 3 つのコア次元から定量・定性の対比分析を行った。ここでは、キャッシュフロー分析を用いた経済的実現可能性の検討結果について概要を述べる。詳細な条件や、他の結果、議論については、参考文献 [2] を参照されたい。

長期運用シナリオでは、エッジコンピューティングアーキテクチャはクラウドソリューションと比較して顕著なコストアドバンテージを持つ。ここでは、監視カメラによる 24 時間リアルタイム監視シナリオに基づいてコスト試算対比を行う。比較したモデルは、クラウドソリューションである Amazon 社の AWS Rekognition Video ストリーム処理サービスを採用した。このサービスでは 2026 年課金基準に基づき、リアルタイムビデオ分析費用は約 0.12 \$/分であり、かつ高額なビデオストリームアップロード帯域費用 (Data Transfer Out) を負担する必要がある。

エッジコンピューティングには、Rockchip 社 RK3588 アーキテクチャを採用したプラットフォームを想定する。RK3588 アーキテクチャは単価約 3.6 万円であることが判明したため、これを採用してエッジノードを構築することが、より優れた商業化選択である。その 6TOPS の専用 NPU (Neural Processing Unit) は極めて低消費電力 (<5W) で単ノード多チャンネルハイビジョンビデオ並行処理を実現できる。NPU 利用への進化により、システムの演算密度は 3～5 倍向上し、1 チャンネル監視のハードウェア償却コストをさらに圧縮できる。

Amazon Rekognition ライブビデオイベントの人物検出コスト分析に基づく 2000 チャンネルのカメラを備える大型住宅街のプロフェッショナルビデオ監視サービスを提供すると仮定する。エッジコンピューティングの方でも、同規模のイベントに対応するため、各監視カメラが 1 台の RK3588 エッジデバイスと対応する敷設シナリオで考える。このモデルは、システム構築時の資本支出が膨らむが、これは 1 回限りの支出である。詳細な条件と計算は文献 [2] に譲り、ここでは 24 時間

稼働させた場合の投資回収期間分析を、電気代も考慮して行った結果を図1に示す。

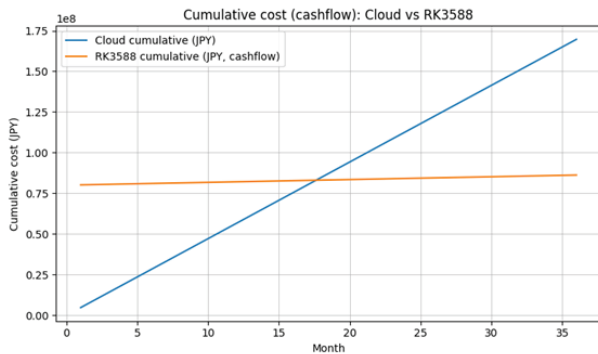


図1 キャッシュフロー比較

黄線：エッジコンピューティング方式
青線：クラウドコンピューティング方式

図1によれば、エッジコンピューティング方式は、導入時(1ヶ月目)のコストは圧倒的に高いが、キャッシュフロー全体では、約18ヶ月でクラウド方式を下回るコストを達成する。その後、システム全体のコストはクラウド方式を継続的に下回り、稼働時間の延長に伴いコストアドバンテージを拡大し続ける。システムを5年使用と仮定すると、エッジコンピューティング方式は、3年以上のこの低コスト稼働期間を獲得でき、クラウド方式に比較した経済優位性は高い。

この結果をまとめると、以下のとおりである。

1. クラウドビデオ分析方式は柔軟な導入、初期投資の低いアドバンテージを持つが、その長期運営コストは規模に伴い線形成長する。
2. エッジコンピューティング方式は、クラウド方式に比較して高額の初期投資を必要とするが、高密度カメラと高イベント頻度シナリオでは、約18ヶ月で投資回収を実現できる。
3. 日本の一般的な5~7年のデバイス使用サイクルでは、エッジコンピューティング方式は総所有コスト(TCO)面で顕著なアドバンテージを持つ。

3. エッジコンピューティングのローカル IoT サーバへの適用

郑らの試作・検証によりローカルコンピューティングを用いた映像監視システムは実用的であり、経済的であることが確認された。しかし、その試作システムは、開発容易性や機能追加などから、構成が複雑になり、再現が難しい課題が生じている。本章では、研究で得られた成果と知見を基に、従来構築していたローカル IoT サーバシステムにエッジコンピューティングを取入れるシステムアーキテクチャについて検討する。そのためには、機能の整理を行い、ネットワーク構成なども見直す必要がある。

図2が、機能を整理したローカル IoT サーバを含む映像監視システムアーキテクチャである。文献[6]で、MQTT ブローカと統合監視サーバを組み合わせたローカル IoT サーバ(点線部分)を開発したが、これを拡張して本アーキテクチャが構成されている。

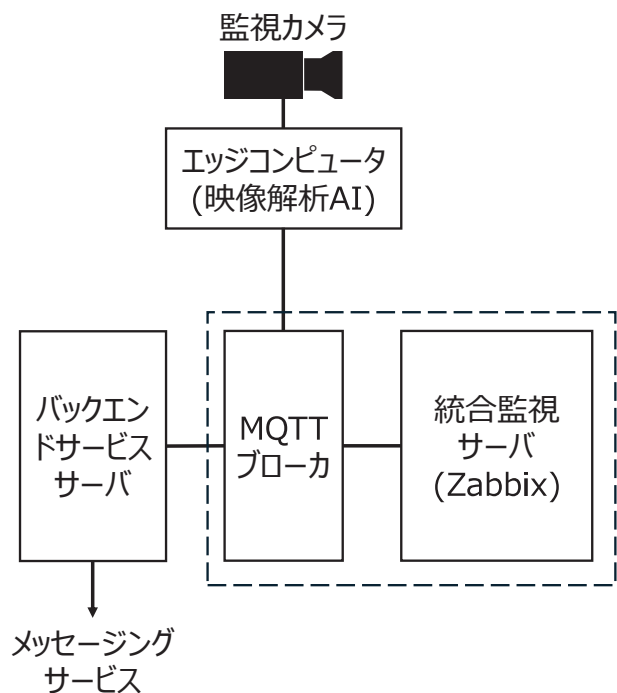


図2 ローカル IoT サーバを含むエッジコンピューティングによる映像監視システムアーキテクチャ

まず、試作システムはエッジコンピュータを有する。これには、監視カメラが接続され、インテリジェント映像解析を行う AI と柔軟なネットワー

ク構築を実現するソフトウェア定義ネットワーク機能が装備される。監視カメラとエッジコンピュータの接続には、ネットワーク型カメラを使用する場合、RJ45 タイプの LAN ケーブルか無線 LAN が使用される。この他にエッジコンピュータは、分析結果をバックエンドサービスサーバや IoT サーバへ送出するためのネットワーク接続を有する。映像解析を行う AI は、試作システムではエッジコンピュータ内に構築した仮想マシン上に実装されている。また、監視カメラ、AI 仮想マシン、その他サーバとの接続、更に、クラウドサービスへの接続を柔軟かつ安全に実現するために、ソフトウェア定義ネットワーク機能を備えた OpenWRT をエッジコンピュータ内に別途構築した仮想マシンに実装している。ただし、エッジコンピュータで最も重要なのは、侵入検知や顔認識機能を提供するインテリジェント映像解析機能であり必須機能であるが、ソフトウェア定義ネットワーク機能 (OpenWRT) は通常不要である。

エッジコンピュータの画像解析 AI が報告すべきイベントを検出した場合、メッセージが送出される。試作システムでは、監視カメラの撮影範囲内で人の動きを検知する侵入検知と、撮影されている人の顔からその人物を特定する顔認識のサービスが試用できる。いずれの場合も、解析結果のメッセージは MQTT により、まず MQTT ブローカに配送され、そこでメッセージを常時監視しているバックエンドサービスサーバや統合監視サーバに検出され、イベントに応じた動作が行われる。MQTT はデータ通信に使える軽量プロトコルとして IoT システムで広く使われている。さらに、MQTT はパブリッシュ／サブスクライブモデルで情報共有するため、MQTT ブローカは、複数機能への同時情報共有を効率的に実行するための重要な機能である。

統合監視サーバは、IoT データの表示と保存の機能を持つ。受信したエッジコンピュータからのメッセージの内容をダッシュボードと呼ばれる全体の運用状況を示している画面に表示することができる。また、そのメッセージを保管して

おくことができ、あとで履歴が取れる。統合監視サーバとしてオープンソースソフトウェアである Zabbix を使用している。Zabbix の場合、このダッシュボードはウェブページの形式で提供されており、ネットワーク越しに遠隔からでも内容を見ることができる。ただし、デバイスとして大きな画面を想定しており、スマートフォンのような画面の小さい装置では、内容をひと目で把握するのは困難な課題がある。

この点において、試作システムは、バックエンドサービスサーバを備えており、各種サービスを提供できる構成になっている。試作システムでは、バックエンドサービスサーバが、スマートフォンなどで、メッセージのプッシュとポップアップ表示を行うノティスを生成するクラウドサービスの起動を行っている。このサービスの使用により、エッジコンピュータが監視カメラ画像から検出したイベントが、スマートフォンなどでノティスされる。このようにバックエンドサービスサーバは、統合監視サーバで実現しづらいサービスを追加して実装したり、外部サービスを起動したりする柔軟な役割が果たせるため、ローカル IoT サーバのシステムアーキテクチャに必要な機能と考えられる。

以上の検討から、エッジコンピューティングを用いるローカル IoT サーバのシステムアーキテクチャで映像監視機能を備えた場合を図 2 の構成にまとめた。文献 [2] では、物理機能と仮想機能を含めて詳細を示すためにアーキテクチャ階層図やネットワーク階層図など複数の図面があり、その理解が複雑であった。これを各機能を確認しながら整理することで、エッジコンピューティングの効果が得られるシステムアーキテクチャのコアを明らかにすることができた。試作システムが備えていたエッジコンピューティングによるリアルタイムのインテリジェント映像解析や、マルチデバイスアクセス、MQTT によるリアルタイムのメッセージ共有、などの特徴は、維持されている。このアーキテクチャを構成する、統合監視サーバ、MQTT ブローカ、バックエンドサービス

サーバ、エッジコンピュータは、独立した装置でも良いし、仮想マシンでも良い。またいくつかの機能は1つのPC内にインストールされていても良い。ただし、インテリジェント映像解析は、強力な計算リソースやメモリを必要とすることが多いため、独立した装置として用意することが適当と考えられる。

4. まとめ

本研究では、鄭と濱崎が試作したエッジコンピューティングに基づく映像監視システムを基に、従来のローカルIoTサーバシステムにエッジコンピューティングを取入れるシステムアーキテクチャについて映像監視機能を備えた場合で、機能やネットワーク要件などを整理した。その結果、簡潔でかつ、エッジコンピューティングの効果が得られるシステムアーキテクチャが明らかになった。



図3 リアルタイム検出と通知

右図：モバイル端末アプリ画面全体、左図：ノーティス部分の拡大
(下の侵入検知の後、上の顔認識結果のメッセージ)

得られたアーキテクチャは、エッジコンピューティングによるリアルタイムのインテリジェント機能利用が期待される上に、これまで使用している統合監視ソフトウェア (Zabbix) の制限から困難であったマルチデバイスアクセスへの機能拡張性も備えることができ、更なる機能拡張や進化も期待できる構成となった。

例えば、試作システムでは、バックエンドサービスサーバから、Firebase FCM 経由でモバイル端末へリアルタイム通知をプッシュするデモを行った。図3に示すような、人物検知や顔認識による検出者の特定のノーティスが5秒未満でスマートフォンに通知されており、新規デバイスや

サービスをモジュール化して取込むことの有効性が確認されている。

参考文献

- [1] 総務省, 情報通信白書 令和7年版, July 2025
- [2] 鄭宇宏, 濱崎隆世, ”エッジコンピューティングを用いた高可用・低コストな映像監視システムの研究”, 第一工科大学 卒業研究報告, 2026
- [3] Cisco, ” フォグ コンピューティング ソリューション: Internet of Things のパワーを解き放つ”, ホワイトペーパー, 2016 https://www.cisco.com/c/dam/m/ja_jp/of-fers/164/never-better/core-networking/computing-solutions.pdf (2026.2.10 閲覧)
- [4] OASIS Open, ”MQTT Version 5.0”, March 2019, <https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html> (2025.5.10 閲覧)
- [5] 加藤大幸, 立山アキラ, 楼駿宇, ”SNMP によるネットワーク管理簡単化の研究”, 第一工科大学 卒業研究報告, 2025
- [6] 馬場伸一, ”スマートホームにおけるIoTデバイス管理の一考察”, 第一工科大学研究報告 第36号 pp.11-16, 2024
- [7] Frigate, ”Open-source object detection for CCTV cameras”, 2024. <https://frigate.video/> (2026.2.10 閲覧)
- [8] Google, ”FCM アーキテクチャの概要”, <https://firebase.google.com/docs/cloud-messaging/fcm-architecture> (2026.5.7 閲覧)