

# AI 통합 지능형 컨테이너 오케스트레이션 방법론: IoT 서비스 최적화 설계

## AI-Integrated Intelligent Container Orchestration Methodology: Design for Optimizing IoT Services

김창인, 허유정, 이혜림, 박준서, 최준혁, 김현기 석사, 장수영 교수, 최창범 교수  
국립한밭대학교 컴퓨터공학과

Changin Kim, Youjeong Heo, Hyerim Lee, Junseo Park, Junhyuk Choi, Hyeonggi Kim,  
Suyoung Jang, Changbeom Choi

{20217140, 20211939, 20211924, 20181618, 20211901, 30242833}@edu.hanbat.ac.kr,  
{syjang, cbchoi}@hanbat.ac.kr

### Abstract

본 연구는 IoT 서비스의 성능 및 확장성을 극대화하기 위해 AI 기반 알고리즘과 엣지 컴퓨팅을 통합한 지능형 컨테이너 오케스트레이션 방법론을 제안한다. IoT 디바이스의 실시간 데이터 처리 요구와 자원 사용 패턴을 분석하여, 엣지 컴퓨팅 환경에서 컨테이너 리소스의 동적 할당 및 관리를 최적화함으로써, IoT 애플리케이션의 효율성을 증대시키고 네트워크 부하를 줄인다. 또한, 전체 시스템의 에너지 효율성을 개선하고, 지속 가능한 디지털 전환을 위한 실질적인 IoT 서비스 솔루션을 제공하는 것을 목표로 한다.

### 서론

현대 사회의 급속한 디지털화는 사물인터넷(IoT) 기술의 발전과 밀접하게 연관되어 있으며, 다양한 산업 분야에서 IoT의 적용 범위는 계속해서 확대되고 있다[1]. IoT 기술은 데이터 수집, 처리, 그리고 분석을 통해 실생활과 산업 환경에서의 의사 결정을 지원하는 핵심 도구로 자리 잡았다[2]. 하지만 IoT 시스템의 복잡성과 대규모 데이터 처리 요구는 IoT 서

비스의 성능 및 확장성에 새로운 도전 과제를 제시하고 있다.

이러한 도전 과제에 대응하기 위해, 본 연구는 인공지능(AI) 알고리즘과 엣지 컴퓨팅 기술을 통합한 지능형 컨테이너 오케스트레이션 방법론을 제안한다. 이 방법론은 IoT 디바이스로부터 수집된 실시간 데이터 처리 요구와 자원 사용 패턴을 분석하고, 이를 기반으로 엣지 컴퓨팅 환경에서의 컨테이너 리소스를 동적으로 할당 및 관리함으로써 IoT 애플리케이션의 효율성을 증대시키고 네트워크 부하를 줄이는 것을 목표로 한다. 또한, 본 연구는 전체 시스템의 에너지 효율성을 개선하고, 지속 가능한 디지털 전환을 위한 실질적인 IoT 서비스 솔루션을 제공하려고 한다.

본 논문은 먼저 IoT 서비스 최적화의 현재 상황과 관련된 문제점을 검토하고, 기존 연구와의 차별점을 명확히 한다. 이어서 제안하는 AI 통합 지능형 컨테이너 오케스트레이션 방법론의 구체적인 설계 및 구현 방법을 상세히 설명할 것이다.

## 본론

### 1. 기술적 배경 및 관련 연구

현대의 IoT 서비스 환경은 방대한 양의 데이터를 실시간으로 처리하고, 다양한 사용자 요구에 신속하게 대응할 수 있는 높은 성능 및 확장성을 요구한다[3]. 이러한 요구를 만족시키기 위해, 본 연구에서는 AI 기반 알고리즘과 엣지 컴퓨팅 기술을 통합한 지능형 컨테이너 오케스트레이션 방법론을 제안한다. 기존 연구에서는 대규모 IoT 시스템에서의 데이터 처리와 자원 관리 문제를 다루었으나[4][5][6], 이들 대부분은 엣지 컴퓨팅 환경에서의 동적 리소스 관리에 중점을 두고 AI의 역할을 제한적으로 탐구하였다. 본 연구는 이러한 한계를 극복하고, AI를 활용하여 IoT 서비스의 성능 최적화와 자원 할당의 효율성을 극대화하는 새로운 접근 방식을 모색한다.

### 2. 방법론

본 연구에서 제안하는 방법론은 IoT 서비스의 성능 및 확장성을 극대화하기 위해 AI 기반 알고리즘과 엣지 컴퓨팅을 통합한 지능형 컨테이너 오케스트레이션 접근 방식을 중심으로 한다. 이 접근 방식의 주요 목표는 실시간 데이터 처리 요구 사항과 다양한 자원 사용 패턴을 고려하여, 엣지 컴퓨팅 환경에서 컨테이너 리소스의 동적 할당 및 관리를 최적화함으로써 IoT 애플리케이션의 효율성을 증대시키고 네트워크 부하를 줄이는 것이다.

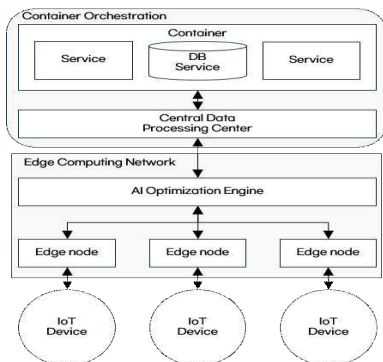


그림 1. System Architecture

그림 1은 제안하는 방법론을 적용한 전체 시스템의 구성도를 나타낸다. 이 구성도는 엣지 컴퓨팅 노드, 중앙 데이터 처리 센터, 그리고 이들 간의 데이터 흐름 및 컨테이너화된 서비스의 배치를 보여준다.

#### 2.1 데이터 수집 및 전처리

IoT 디바이스로부터 수집된 데이터는 다양한 형식과 구조를 가진다. 그림 1에서 보듯이, 이 데이터는 먼저 엣지 컴퓨팅 노드로 전송되며, 여기에서 초기 전처리 과정을 거친다. 전처리 과정에는 노이즈 제거, 결측치 처리, 데이터 정규화 등이 포함되어 데이터의 품질을 향상시킨다.

#### 2.2 엣지 컴퓨팅 환경에서의 데이터 처리

엣지 컴퓨팅 노드에서의 데이터 처리는 실시간 응답성 향상 및 중앙 서버의 부하 경감을 목적으로 한다[7]. 그림 1과 관련하여, 엣지 노드에서의 데이터 처리는 데이터 수집 직후에 이루어지며, 이 과정에서 컨테이너화된 마이크로서비스를 활용한다. 이러한 마이크로서비스는 특정 데이터 처리 작업을 전담하며, 필요에 따라 동적으로 스케일링될 수 있다.

#### 2.3 AI 기반 최적화 알고리즘

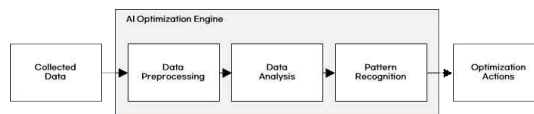


그림 2. AI Optimization Engine Process

본 연구의 핵심은 AI 기반 최적화 알고리즘을 통해 IoT 서비스의 성능을 자동으로 최적화하는 데 있다. AI 알고리즘은 그림 2에 개념적으로 나타나 있으며, 이는 데이터 패턴 분석, 예측 모델링, 자원 할당 결정 등을 포함한다. 이러한 AI 알고리즘은 엣지 및 중앙 처리 노드에서 수집된 데이터에 기반하여 동작하며, 시스템 전반의 성능 최적화를 도모한다.

## 2.4 컨테이너 오케스트레이션을 통한 서비스 관리

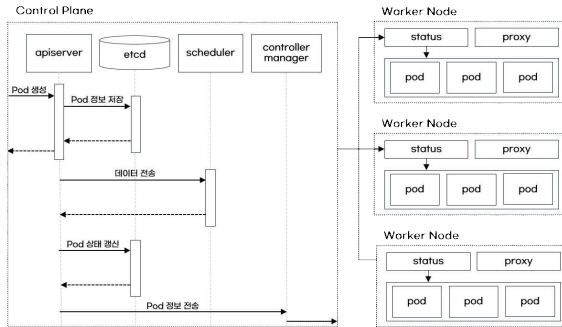


그림 3. Container Orchestration: Service

그림 3은 컨테이너 오케스트레이션을 통한 서비스 관리 방법을 시각적으로 보여준다. 여기에는 컨테이너화된 서비스의 배포, 업데이트, 모니터링 및 자동 스케일링이 포함된다. Kubernetes와 같은 오케스트레이션 도구는 이러한 작업을 자동화하고, 시스템의 안정성과 확장성을 보장한다.

## 결론

본 연구는 IoT 서비스의 성능 및 확장성을 극대화하기 위해 AI 기반 알고리즘과 엣지 컴퓨팅을 통합한 지능형 컨테이너 오케스트레이션 방법론을 제안하고 탐구하였다. 이 방법론은 엣지 컴퓨팅 환경에서 컨테이너 리소스의 동적 할당 및 관리를 최적화함으로써 IoT 애플리케이션의 효율성을 증대시키고, 네트워크 부하를 줄이며, 전체 시스템의 에너지 효율성을 개선하는 것을 목표로 하였다. 추후 연구로는 다양한 IoT 환경에서의 적용성 및 효과를 평가하기 위한 추가 연구와 더욱 고도화된 AI 알고리즘을 개발하고 통합하여, IoT 서비스의 성능 최적화 및 자원 관리 효율성을 더욱 향상시키는 연구를 진행할 예정이다.

## 사사문구

“본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구 결과로 수행되었음”(2022-0-01068)

## 참고문헌

- [1] Madakam, S., Ramaswamy, R., & Tripathi, S., “Internet of Things (IoT): A Literature Review”, *Journal of Computer and Communications*, Vol.3 No.5, 2015.
- [2] Laghari, A. A., Wu, K., Laghari, R. A., Ali, M., & Khan, A. A., “A review and state of art of Internet of Things (IoT)”, *Archives of Computational Methods in Engineering*, 2021, 1-19.
- [3] 표철식, 강호용, 김내수, 방효찬, “IoT(M2M) 기술 동향 및 발전 전망”, *한국통신학회지(정보와통신)*, 30(8), 2013, 3-10.
- [4] 이승철, 신동하, “IoT 시스템에서 최소 자원을 사용하는 TCP/IP 구현”, *한국컴퓨터정보학회논문지*, 25(10), 125-133, 2020.
- [5] 나문성, 김승훈, 이재동, “클라우드 환경에서 대규모 콘텐츠를 위한 효율적인 자원처리 기법”, *한국산업정보학회논문지*, 15(4), 17-27, 2010.
- [6] 김선옥, 최현화, 장수민, 김대원, 차재근, “지능형 엣지 서비스를 위한 클라우드 엣지 SW 플랫폼 기술 동향”, *한국통신학회지(정보와통신)*, 37(8), 46-54, 2020.
- [7] 홍정하, 이강찬, 이승윤, “엣지 컴퓨팅 기술 동향”, *전자통신동향분석*, 35(6), 78-87, 2020.