



AI 통합 지능형 컨테이너 오케스트레이션 방법론 : IoT서비스 최적화 설계

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

Introduction

연구배경

- 현대 사회의 급속한 디지털화와 IoT 기술의 발전은 다양한 산업에서 큰 변화를 이끌었지만, 이로 인한 **데이터의 복잡성**과 **대규모 처리 요구**는 기존의 IoT 서비스 성능 및 확장성에 도전을 제시함

연구 목적

- AI 기반 알고리즘과 엣지 컴퓨팅을 통합하여 IoT 서비스의 성능 확정성을 극대화 하는 지능형 컨테이너 오케스트레이션 방법론을 제안함

연구의 필요성

- IoT 디바이스의 다양한 데이터 처리 요구와 자원 사용 패턴을 효과적으로 관리하여, IoT 애플리케이션의 효율성을 높이고 네트워크 부하를 줄일 필요가 있음

Background

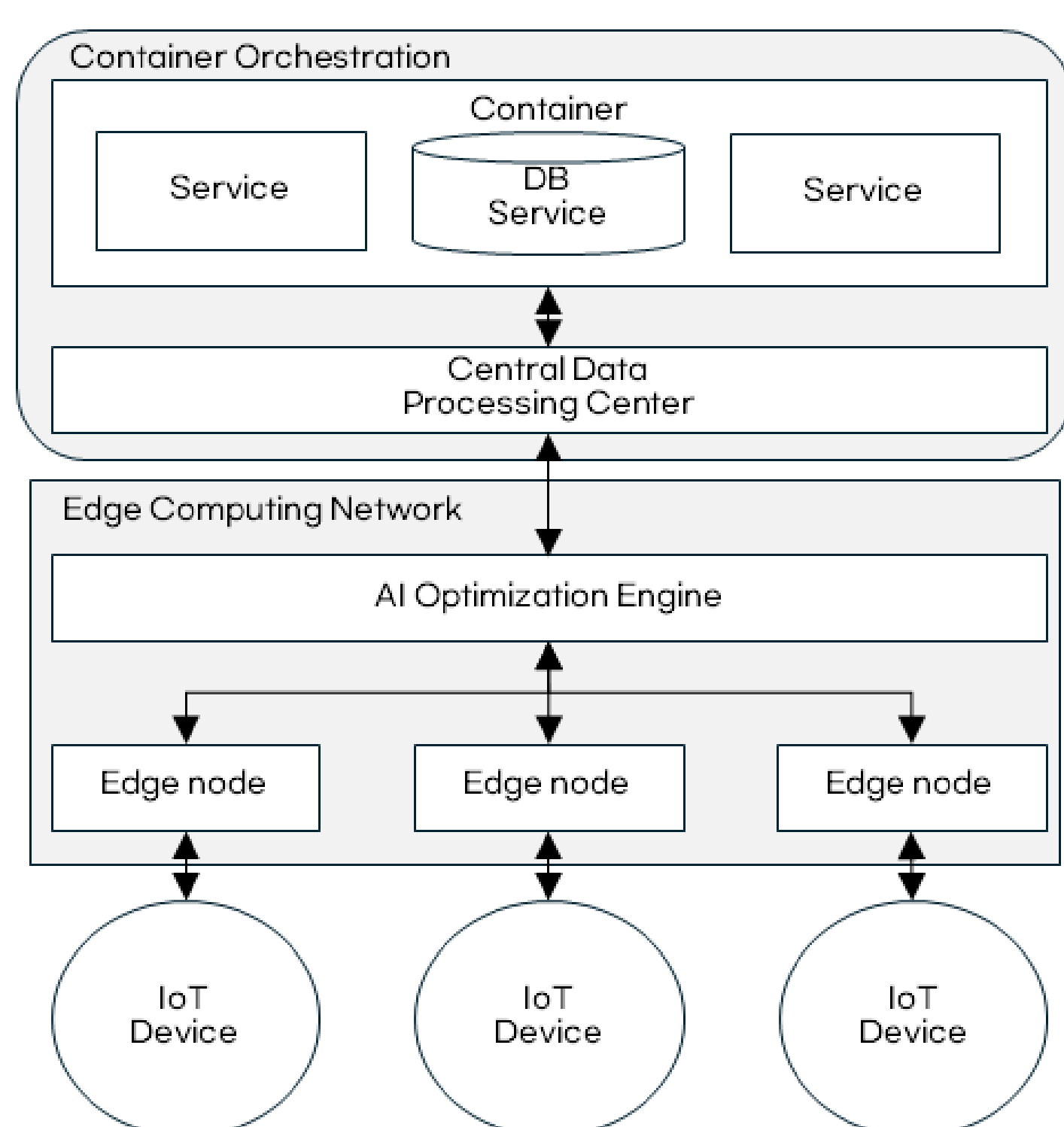
국내외 연구현황

- 기존 연구들은 대부분 IoT 시스템에서의 **데이터 처리와 자원에 초점**을 맞추었으나, 이들은 주로 엣지 컴퓨팅 환경의 동적 리소스 관리에 집중하고 AI의 역할을 제한적으로 탐구함

문제점

- 현 IoT 시스템에서의 복잡성과 대규모 데이터 처리 요구는 기존 방법론으로는 해결하기 어려운 새로운 문제들을 제시하고 있으며, **이는 성능 및 확장성 제한을 초래함**

System Architecture



시스템 아키텍처

- 전체 시스템 구성도를 통해 엣지 컴퓨팅 노드, 중앙 데이터 처리 센터, 그리고 이들 간의 데이터 흐름 및 컨테이너화 된 서비스의 배치를 보여줌

엣지 컴퓨팅 네트워크

- 데이터 실시간 처리와 응답성 향상을 위한 네트워크 구조와 IoT 서비스의 성능을 자동으로 최적화하기 위한 AI 엔진을 결합하여 구성

컨테이너 오케스트레이션

- 시스템 안정성 보장을 위한 서비스의 배포, 업데이트, 모니터링을 자동화하기 위한 컨테이너 오케스트레이션 도구를 결합하여 구성

Proposed Method

데이터 수집 및 전처리

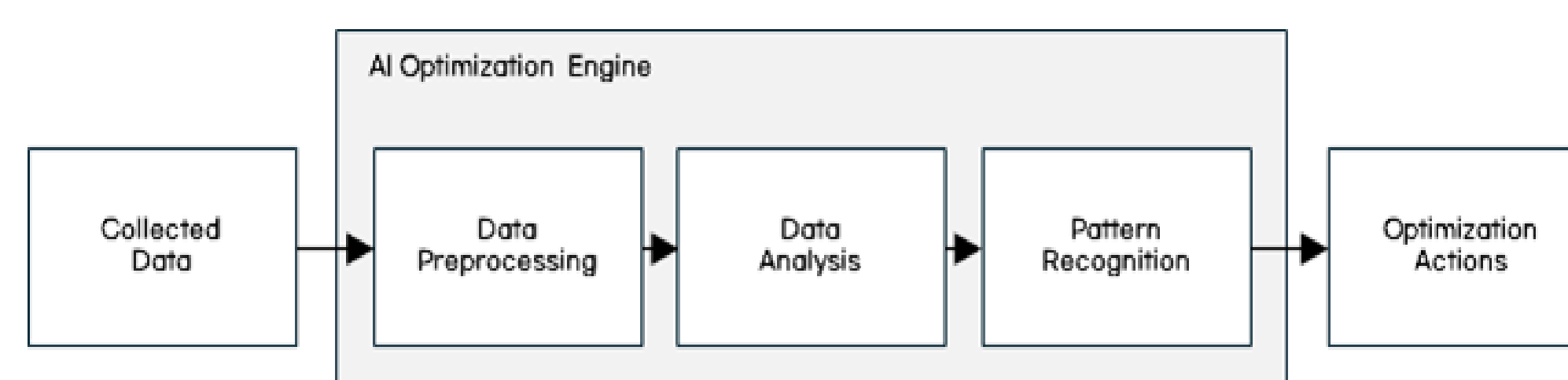
- IoT 디바이스로부터 수집된 다양한 형식의 데이터는 먼저 엣지 컴퓨팅 노드로 전송되고 초기 전처리 과정을 거침

엣지 컴퓨팅 환경에서의 데이터 처리

- 엣지 컴퓨팅 노드에서의 데이터 처리를 통해 중앙 서버의 부하 경감을 목적으로 하고 마이크로 서비스는 특정 데이터를 전담하며, 필요에 따라 동적으로 스케일링 됨

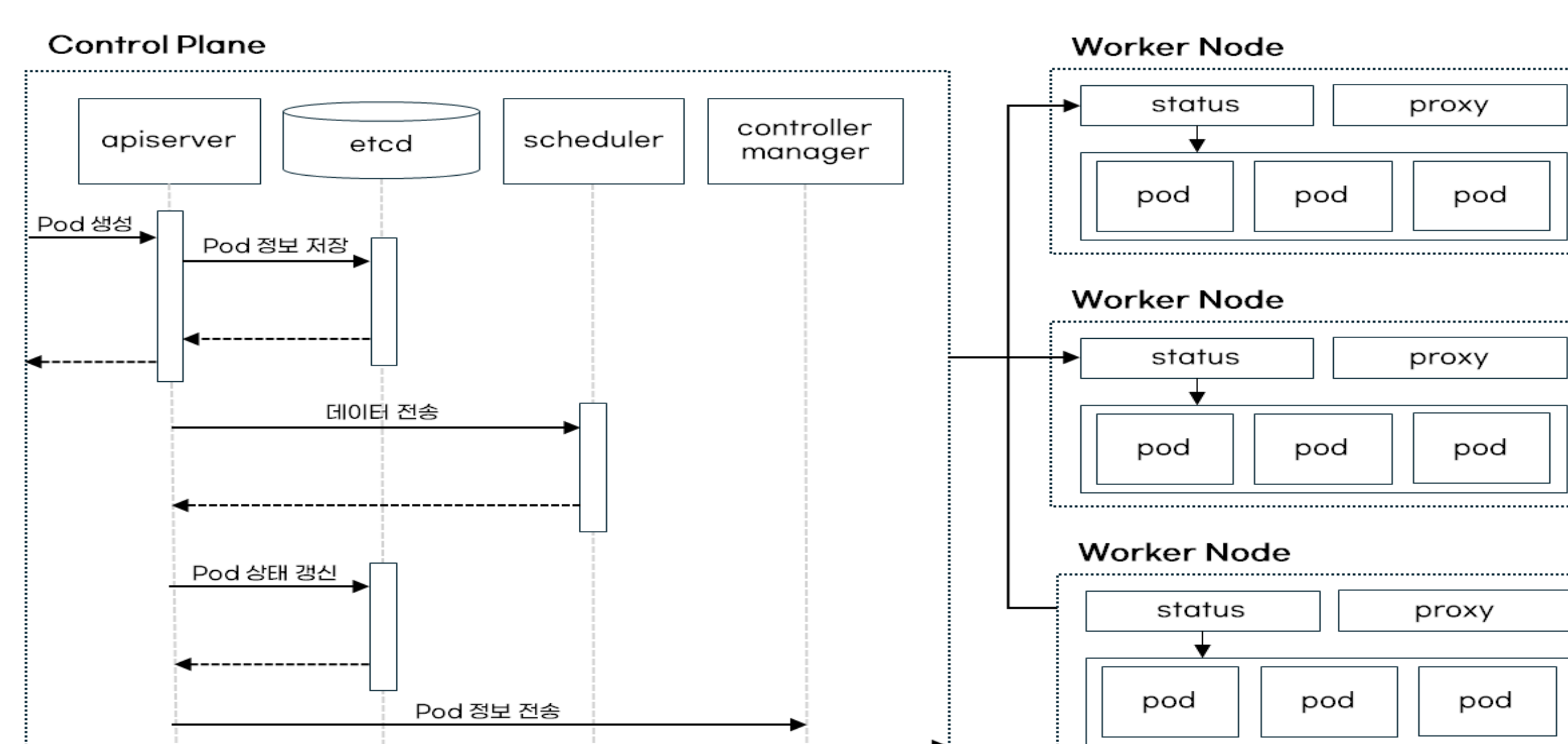
AI 기반 최적화 알고리즘

- AI 알고리즘은 엣지 및 중앙 처리 노드에서 수집된 데이터에 기반하여 동작하며 시스템 전반의 성능 최적화를 도모함



컨테이너 오케스트레이션을 통한 서비스 관리

- 컨테이너화된 서비스의 배포, 업데이트, 모니터링 및 자동 스케일링을 위한 컨테이너 오케스트레이션 서비스 관리를 방법을 적용하며, 작업을 자동화 하고 시스템의 안정성과 확장성을 보장함



Conclusion

- IoT 서비스의 성능 및 확정성을 극대화 하기 위해 AI와 엣지 컴퓨팅을 결합한 지능형 컨테이너 오케스트레이션 방법론을 제안하고 탐구함
- 기대 효과로 IoT 애플리케이션의 효율성을 증대시키고, 네트워크 부하를 감소시키며, 전체 시스템의 에너지 효율성을 개선 함
- 향후 연구로 다양한 IoT 환경에서의 적용성을 평가하고, 고도화된 AI 알고리즘 개발을 통해 IoT 서비스의 성능과 자원 관리를 더욱 향상시킬 예정임

Acknowledgement

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