

KI Cloud R&D 플랫폼을 위한 통합 모니터링 시스템 연구

박준영, 정기문, 조혜영, 손아영, 박경석

한국과학기술정보연구원

{jypark, kmjeong, chohy, ayson, gspark}@kisti.re.kr

A Study on the federated monitoring system for KI Cloud R&D Platform

Junyoung Park, Kimoon Jeong, Hyeyoung Cho, Ayoung Son, Kyongseok Park

Korea Institute of Science and Technology Information

요약

본 논문은 기존의 KI Cloud에서 부족했던 R&D SW를 클라우드 플랫폼으로 전환하여 R&D 클라우드 서비스를 제공하기 위한 KI Cloud R&D 플랫폼에 대해서 소개하고 해당 플랫폼에서 다양한 서비스를 제공하고 Multi-tenant로 인한 통합 관리 및 제어를 위해 필수적인 통합 모니터링에 대해 연구 및 통합 모니터링 체계에 대한 구축 방안과 활용 방안에 대해서 서술하였다. 다양한 서비스와 클러스터 확장 등을 고려하여 통합 모니터링 체계 방안을 서술하였으며 향후 수집된 통합 모니터링 데이터를 통해 상황별 프로파일링과 최적 스케줄링을 통해 효율적 클라우드 운영 가능성을 제시하였다.

I. 서론

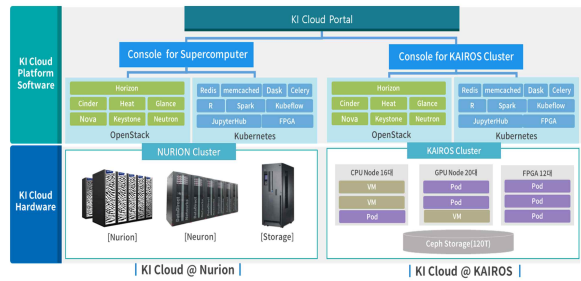
한국과학기술정보연구원에서 계산과학 분야의 HPC Task를 효율적으로 처리하기 위해 KI Cloud[1]를 구축하여 2021년부터 대학(원)교 및 연구원을 중심으로 서비스를 제공하고 있다. 하지만 신기술의 등장으로 많은 연구분야에서 AI, 빅데이터 등과 같은 신기술을 적용한 연구들이 많이 진행되었고 많은 연구자들의 연구환경 또한 다양해지고 복잡해졌다. 이를 클라우드 서비스를 통해 해결하고자 하는 요구들이 많아지면서 연구자들을 지원하기 위한 KI Cloud 서비스 또한 다양한 연구환경들을 제공하고 자 R&D 플랫폼을 구축하게 되었다.

HPC 및 R&D SW를 지원하기 위한 클라우드 서비스 플랫폼을 서버 및 네트워크부터 제공하고 있는 서비스에 대한 모니터링, 작업 결과 등과 플랫폼 및 서비스 관리자를 위한 서비스 배포 및 동작 상태 등을 확인할 수 있는 통합 모니터링이 필요해졌으며, 이를 위해 고려해야되는 요소 및 모니터링 요소, 분석 메트릭 등에 대한 연구를 기술하였다.

본 논문에서는 KI Cloud 기반의 통합 모니터링 구축 체계를 설명하기 위해 2장 KI Cloud와 KI Cloud R&D 플랫폼을 비교하여 통합 모니터링의 필요성을 설명하고 3장에서는 통합 모니터링 체계 구축 방안을 제시한다. 마지막으로 4장 결론에서 통합 모니터링 수집 및 활용 방안을 설명한다.

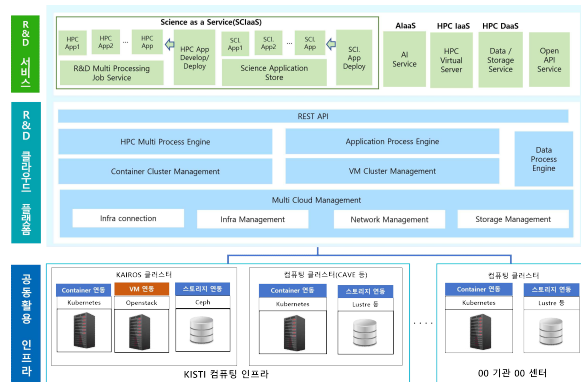
II. KI Cloud와 KI Cloud R&D 플랫폼 비교

기존 KI Cloud는 가상환경에서 과학기술분야 연구자에게 맞춤형 연구환경을 제공하는 것을 목표로 Openstack을 기반으로 가상서버 서비스, 클러스터 서비스, 스토리지 서비스 등의 IaaS 서비스를 제공 및 컨테이너 기반 가상 클러스터 환경에서 Jupyter Notebook, R studio 등을 활용한 Web 기반 IDE와 데이터 분석 서비스를 제공하고 있다. 또한 KI Cloud 서비스 및 자원 현황에 대한 모니터링을 Prometheus와 Grafana를 통해 모니터링 시각화 서비스도 제공하고 있다[2,3]. 하지만 VM, Container의 상태 정보와 Node exporter를 활용한 서버 정보만 수집하기 때문에 클라우드 서비스에 대한 모니터링을 제공하지 못하고 있다.



(그림 1) 기존 KI Cloud 구조

기존 KI Cloud에서는 VM 및 Container 기반 web IDE 서비스, 데이터 공유 서비스 등 일반적인 클라우드 서비스를 제공하였으나 KI Cloud R&D 플랫폼[2]에서는 거대 과학문제 연산 및 AI 모델 학습, 빅데이터 처리 등을 위해 HPC Task 및 분산 병렬처리, 대용량 파일 전송 서비스 등의 서비스 제공을 목표로 하고 있다. 또한 SCIIaaS(SCIENCE as a Service), SCI-APP (Science Application) 등 과학분야에서 사용되는 SW를 중심으로 클라우드 기술을 활용해 동일한 연구환경을 언제 어디서나 접근할 수 있는 R&D 서비스를 제공하고자 한다.



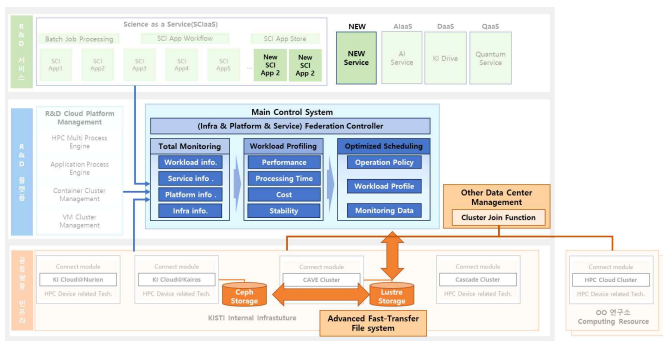
(그림 2) KI Cloud R&D 플랫폼 구조

하지만 다양한 종류의 어플리케이션과 서비스 형태가 다르기 때문에 모든 컴퓨팅 자원 및 서비스에 대한 모니터링이 어려우며 KI Cloud R&D 플랫폼을 위한 통합 모니터링 요구사항이 다음과 같이 도출되었다.

- Node Exporter 기반의 자원 모니터링에서 GPU, NPU, FPGA 등 추가적인 가속기, 부속품들에 대한 자원 모니터링 확장
- 자원에 대한 모니터링 뿐만 아니라 효율적인 운영 및 통합 관리를 위해 컴퓨팅 자원, 서비스, 사용자, I/O 등에 대한 통합 모니터링 데이터 수집
- 멀티 클라우드 및 분산 클라우드 환경에서 다양한 로그 데이터에 대한 정규화 및 표준화된 모니터링 데이터 수집

위와 같은 요구사항을 충족하기 위해 클라우드 계층별 모든 로그 및 모니터링 데이터를 수집하고 연관분석하여 모니터링을 기반으로 통합 제어 시스템 구축을 계획하게 되었다.

III. KI Cloud R&D 플랫폼 통합 모니터링 체계 구축 방안



(그림 3) KI Cloud R&D 플랫폼을 위한 통합 모니터링 개발 방안

KI Cloud R&D 플랫폼을 위한 통합 모니터링 체계를 구축하기 위해 클라우드 서비스에 사용되는 모든 요소들에 대한 상태 정보 수집이 필요하였으며 수집한 정보를 효율적으로 활용할 수 있도록 프로파일링 및 최적 스케줄링에 활용하기로 하였다.

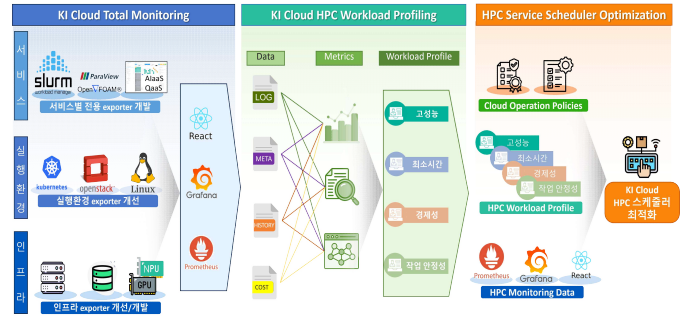
우선, 통합 모니터링을 하기 위해서 다음과 같이 모니터링 대상을 정의하였다.

- **인프라정보:** 클라우드 서비스에 활용하는 서버 및 스토리지 자원에 대한 모니터링 뿐만 아니라 네트워크 정보, GPU, NPU, FPGA 등에 대한 가속기 정보를 수집함.
- **플랫폼정보:** KI Cloud R&D 플랫폼은 Openstack과 Kubernetes를 기반으로 구축되었기 때문에 Openstack 및 Container에 대한 상태정보와 사용자 접근 정보 등을 수집함.
- **서비스정보:** 서비스 형태로 제공하는 SCIAaaS, AIaaS, 그리고 QaaS 등의 서비스와 사용자 Data 할당 및 공유 정보 등을 수집해야 하며 SCI-APP 서비스인 Container Image를 기반으로 제공하는 Openfoam, Paraview, VScode, Jupyter Notebook 등의 Application 서비스에 대한 정보 수집을 수행함. Service 및 Application의 서비스는 로그 수집을 위한 별도의 log agent를 개발하여 Image 추가하였음
- **워크로드정보:** Slurm, PBS와 같이 병렬처리를 위해 실행되는 HPC Task에 대한 정보를 수집하며 현재는 Slurm의 Job log를 수집하는 Agent를 개발하여 적용함

인프라 정보부터 플랫폼, 서비스, 워크로드까지 KI Cloud R&D 플랫폼에서 생성되는 모든 로그를 수집하여 실시간 모니터링을 수행을 목표로하며 이를 기반으로 효율적인 클라우드 운용 및 사용자 Task 진행 현황 제공 등으로 관리 및 사용 편의성을 증가시킬 수 있을 것으로 생각된다.

V. 결론

본 논문에서는 KI Cloud R&D 플랫폼에서 인프라부터 플랫폼, 서비스, 워크로드까지 모든 계층의 로그 데이터를 수집하여 해당 플랫폼의 통합 모니터링을 제공하고 이를 기반으로 진보된 정보 서비스를 제공할 수 있는 연구 성과를 소개하였다. 이렇게 수집한 로그데이터를 통해 (그림 4)와 같이 KI Cloud R&D 플랫폼에서 HPC 서비스를 보다 효율적이고 안정적, 경제적으로 사용할 수 있도록 프로파일링 및 최적 스케줄링을 제공하고자 한다.



(그림 4) R&D 클라우드 플랫폼 서비스 운영 최적화 기술

현재까지 인프라 및 서비스에서의 로그 기록, 저장 방식들이 다르고 이를 시각화 및 분석방법이 달라 모든 로그 데이터를 수집하고 별도의 변환 작업을 거치거나 특정 개발 기관에 종속된 방법으로 수집 및 분석되는 경우가 대부분이다. 해당 통합 모니터링 체계를 기반으로 개발 및 성능 검증을 통해 통합 모니터링을 위한 표준 및 오픈소스화를 위해 노력할 것이다.

ACKNOWLEDGMENT

본 논문은 한국과학기술정보연구원(KISTI)에서 2024년도 정부(과학기술정보통신부)의 지원(IITP(No. RS-2024-00397359)와 KISTI(No. K24L2M1C6, K24L2M1C7)으로 수행된 연구임

참고 문헌

- [1] KI Cloud(Intelligent Cloud), <https://kicloud.ksc.re.kr/>
- [2] 정기문, 조혜영, 박준영, 손아영, “KI Cloud VM 인스턴스 Life cycle 기능 구현”, 한국통신학회 추계학술발표대회, 2023
- [3] 박주원, 이승민, 정기문, 홍태영, “KI Cloud: 슈퍼컴퓨터를 통한 빅데이터 분석 및 머신 러닝 서비스 구축 방안”, 한국정보처리학회 추계학술발표대회, 2020, 80-82
- [4] KI Cloud R&D Platform, <https://kicloud.kisti.re.kr/>