

국외 출장 결과보고서

결 재	담당자	팀장	팀장	실장	사업단장
	11월 29일	11월 29일	11월 30일	11월 30일	11월 30일
	노일석	김동일	심태현	김정한	이우진

1. 출장개요

출장목적	SC 2023 (The International Conference for High Performance Computing, Networking, Storage, and Analysis 2023) 참석		
출장일정	2023. 11. 11.~2023. 11. 19. (7박 9일)		
성명		소속	직급(직위)
김균태		전산관리팀	선임급연구원
이제원		전산관리팀	선임급연구원
노일석		모델체계팀	선임급연구원
방문기관 (또는 회의장소)	콜로라도 컨벤션 센터(The Colorado Convention Center, Denver, United States of America)		
출발지	인천(Incheon)	도착지	덴버(Denver)

2. 업무수행내용

가. 출장 배경 및 목적

○ 개요

2023년 11월 12일부터 18일까지 미국 콜로라도주 덴버에서 열리는 The International Conference for High Performance Computing, Networking, Storage, and Analysis 2023 (이하, SC23)은 사업단에서 연구/개발/운영하는 수치예보모델의 구동에 필수적인 HPC 시스템을 주제로 전 세계 각 나라의 시스템과 엔지니어가 모여 의견을 나누며 보고 배울 수 있는 만남의 장소이다. 사업단의 KIM 모델은 거대한 자료를 기반으로 방대한 계산자원을 요구하기 때문에 HPC 및 Cluster 시스템은 KIM 모델 개발/운영에 필수적인 요소이며 이를 최적의 환경으로 구성하는 것은 KIM 모델 개발/운영에 중요한 축이다.

특히, 이번 SC23은 GPU 가속기를 성공적으로 결합해 엑사스케일 성능의 슈퍼컴퓨터 시대가 개막한 후 본격적으로 다양한 기관과 업체에서 개발/도입한 이기종 시스템을 소개하고 관련 기술과 경험을 다룰 예정이다.

○ 목적

- 김균태 : 슈퍼컴퓨팅 관련 이슈와 신기술, 기술발전 등 세계적인 연구 동향과 흐름을 파악하고자 한다. SC23은 슈퍼컴퓨팅 관련 세계 각지의 많은 전문가들이 참여하여 발표와 토론을 진행하는

학회로써, 슈퍼컴퓨팅과 관련된 이슈와 새로운 기술들, 기술발전과 그 방향 등 최신의 연구동향을 파악하는데 유리하다. SC22는 HPC 시스템에 대한 미래 개념과 동향, Exascale 시스템의 개발 및 향후 전망, AI 전용 프로세서와 메모리 기술 등 다양한 분야에서의 연구성과들이 공유되었다.

- 이제원 : 슈퍼컴퓨팅 관련 이슈와 신기술, 기술발전 등 세계적인 연구 동향과 흐름을 파악하고자 한다. 인공지능 신경망 분석을 위한 슈퍼컴퓨팅 기술의 발전에 따라 이와 관련한 처리장치 (FPGA, ASIC, 뉴로모픽칩, GPGPU)의 동향파악을 통해 향후 사업단 슈퍼컴퓨팅 도입 및 운영 시 신규 기술을 적용하기 위한 정보를 습득하고자 한다. SC23은 SC22에서 소개된 연구성과들의 동향과 새롭게 선보이는 기술들에 대한 많은 논의가 있을 것으로 판단된다.
- 노일석 : 슈퍼컴퓨터 Top 500 최상위 그룹에서 주도하는 고성능 컴퓨터 기술의 최신 동향을 파악하고자 한다. 이기종 시스템을 활용해 사업단 KIM 모델의 가속 시키는 연구를 수행하고 있는데, 해외 전문가들로부터 수치해석 프로그램에서 GPU 가속화를 더 효율적으로 수행할 수 있는 방법론들에 대해 토의하고자 한다. 그리고 해외 연구기관 또는 기업에서 거대 이기종 시스템을 도입해 대규모 멀티노드에서 가속코드를 수행한 경험 등을 탐색하고 장기적으로 발생할 수 있는 문제들을 파악해 향후 개발 계획에 반영하고자 한다.

나. 주요 업무 수행 사항

○ 김균태

- Cyber Security in High Performance Computing

슈퍼컴퓨터의 수많은 과제를 다루기 위해서는 시스템 계산회로 설계부터 효율성 향상을 위해 사용되는 소프트웨어 도구 등 고려해야 할 많은 부분이 있다. 그 중 이전부터 논의되고 있는 슈퍼컴퓨터의 발전에 따른 기후변화 대응 부분은 슈퍼컴퓨팅의 지속 가능성과도 관련된 주요 과제라 할 수 있다. 엑사스케일 슈퍼컴퓨팅 시대가 도래하고 슈퍼컴퓨터 규모가 커짐에 따라 많은 전력 소모로 인한 탄소배출 증가 문제가 대두되고 있어, 본 워크샵에서는 슈퍼컴퓨터 운용 시 발생하는 탄소배출에 관한 정량화된 데이터를 기반으로 탄소모델 및 관련 프로젝트 논의를 진행하였으며, 향후 기후변화에 대응하기 위한 최적의 시스템 설계와 활용 방안 등을 논의하였다.

- 2nd International Workshop on Cyber Security in High Performance Computing

HPC 시스템은 주로 고성능과 확장성을 목표로 설계되지만 최근에는 정보보안 지침과 도구 부분도 중요한 설계 목표가 되고 있다. 본 워크샵에서는 HPC 환경의 공통 기능 간 주요 보안 차이점에 대한 논의와 신규 HPC Security NIST(National Institute of Standard and Technology)에 대한 연구가 얼마나 안전한지에 관한 논의를 진행하였다. 슈퍼컴퓨터 상위 500대를 기준으

로 발견된 제로데이 취약점의 예를 활용하여 슈퍼컴퓨터 설계 및 활용 부분에서 적용 가능한 보안패치에 관한 동향 등을 파악하였다.

- Workshop on Memory Technologies, Systems, and Applications

메모리 벽 문제로 알려진 컴퓨팅과 메모리 간 속도 격차는 슈퍼컴퓨팅 업계의 주요 과제 중 하나이다. 최근 이기종 컴퓨팅의 확산과 메모리 계층의 확장, 분리된 아키텍처의 출현 등으로 해당 과제의 범위가 확대되면서, 이를 개선하기 위한 방안을 지속적으로 논의하고 있다. 메모리 병목을 해소하기 위해 DRAM에 필적하는 비휘발성 메모리(NVM)를 활용하기도 하며, NVM을 활용한 병렬 I/O 라이브러리(PQueue) 또한 개발되어 있음을 확인하였다.

- ※ 메모리 벽 문제 : CPU, 메모리 간 대역폭의 한계로 CPU가 제 성능을 발휘하지 못하면서 발생하는 메모리 병목현상
- ※ PQueue : PnetCDF와 유사한 API를 활용하는 라이브러리로, PnetCDF에 비해 쓰기는 7배, 읽기는 10배 정도 향상되었다고 함

- Quantum Computing : GRAPHINE, Enhanced Neutral Atom Quantum Computing Using Application-Specific Rydberg Atom Arrangement

양자컴퓨팅(Quantum Computing)은 얽힘(Entanglement)이나 중첩(Superposition) 같은 양자역학적인 현상을 활용하여 자료를 처리하는 시스템이라 한다. 양자컴퓨팅은 0과 1로 구성된 이진코딩(bits)을 사용하는 표준컴퓨터와는 달리 큐비트(Qubits)라는 중첩이 가능한 양자비트를 사용하여 자료를 처리하며, 큐비트는 중첩이라는 특성을 가지고 한 번에 여러 정보를 동시에 표현하는 것이 가능하다. 이러한 양자컴퓨팅을 실현하기 위한 기술은 현재도 다양하게 개발중이며, 그 중 중성원자양자컴퓨팅(Neutral Atom Quantum Computing)은 유망기술 중 하나로 꼽힌다. Rice University에서는 중성원자를 기반으로 하는 대체 큐비트 기술을 활용하여 GRAPHINE을 설계하였으며, GRAPHINE은 게이트를 사용하여 큐비트를 배열 및 재배열 함으로써 유연성을 높이고 양자 알고리즘의 성능을 향상시키는 특정 토폴로지를 설계한다. 이를 통해 GRAPHINE은 널리 사용되는 토폴로지 설계와 비교하여 주요 성능 지표를 평균 최대 56% 향상시켰다고 한다.

- ※ 게이트 : 큐비트에 적용할 수 있는 기본 연산
- ※ 토폴로지 : 양자컴퓨터에서 큐비트가 물리적으로 배열되고 연결되는 방식

o 이제원

- Principles of Scalable and Distributed Deep Neural Networks Training and Inference

머신러닝 및 딥러닝의 발전은 HPC 인프라 관리에 많은 시사점을 제공합니다. 특히 TesnorFlow, Pytoch, cuML과 같은 최신 프레임워크의 이해는 매우 중요합니다. 이들은 다양한 ML모델과 심층 신경망의 고성능 학습, 추론 및 배포를 지원하기 때문입니다. 이는 사업단의

작업 효율성을 높이는데 중요한 역할을 할것으로 생각합니다.

병렬화 전략과 DNN아키텍처에 대한 깊은 이해는 현재와 미래의 HPC 인프라 관리에 있어 중요한 요소입니다. 이는 특히 대규모 분산 교육을 효율적으로 지원하기 위한 고성능 CPU/GPU 아키텍처 활용에 중요한 지식 이론을 제공합니다. 또한, 최신 HPC 클러스터에 사용 가능한 CPU/GPU/NPU 아키텍처에 대한 이해는 사업단의 인프라의 성능을 최적하는데 도움이 될것으로 보입니다. 이 세미나를 통해 최신 ML/DL에 대한 최신 기술 동향에 대한 이해를 할 수 있었으며, HPC 인프라 관리에 있어서의 혁신적인 접근 방법과 최적화 전략을 모색하는데 큰 도움이 되었습니다.

- Cyber Security in High Performance Computing

HPC 시스템은 최고의 성능과 확장성을 달성하기 위해 설계되었지만, 사이버 보안을 준수하는 동시에 성능을 유지하는 것은 어려운 도전입니다. 기업 시스템과 HPC 환경의 공통 기능 간의 주요 보안 차이점을 검토하여 정보를 공유하는 자리였습니다.

NIST에서 새로운 HPC 보안에 대한 보안 지침을 통해 ‘오픈 사이언스’연구의 중요성에 대한 통찰을 얻을 수 있었습니다. 또한, 신뢰할 수 있는 컴퓨팅과 제로 트러스트 모델이 HPC에 적용될 수 있는지에 대한 여부에 대해 논의하였습니다.

또한, 2018년에 공개되고 패치된 이전 슈퍼컴퓨팅 시스템에서 발견된 제로데이 취약점의 사례를 시연하며 HPC의 보안의 중요성에 대해 느끼게 되었습니다. 이를 통해 느낀 지식과 경험은 향후 HPC 시스템의 보안과 관련된 의사결정에 큰 영향을 미칠 것으로 생각합니다.

- Fault-Tolerance for HPC at Extreme Scale

대규모 HPC 시스템의 디스크 기반 스토리지에 대한 중요성을 배우는 자리였습니다. HPC 어플리케이션에 의해 생성되는 데이터의 양이 많기 때문에 스토리지는 확장 가능한 신뢰성, 가용성 및 장애 관리가 반드시 필요합니다. HPC 스토리지에서 선별된 장애 데이터는 스토리지의 시스템 설계 및 배포, 데이터 보호 계획의 수립에 있어 중요한 아이디어를 제공합니다.

이 세미나에서는 OLCP의 Summit 슈퍼컴퓨터의 파일 시스템인 Alplane에서 발생한 장애 데이터셋을 소개하였습니다. 이 데이터셋은 약 3년 동안 32,000개 이상의 디스크에서 발생한 4,000개 이상의 이벤트를 포함하고 있었습니다. 이 사례에서는 수집된 데이터의 장애 분포와 버스트 상관관계 측면에 대해 분석하였습니다. 인클로저 팬 모듈에 가까운 위치에서 장애율이 높다는 결론을 도출하였습니다. 또한, 버스트 장애 분석은 발생한 장애의 3분의 1이 버스트로 발생하며, 이중 90%가 공간적 상관 관계 없이 여러 랙에 영향을 미친다는 점을 강조하였습니다.

이러한 학습을 통해 스트로지 시스템의 배치와 관리 전략에 대해 다시 한번 생각할 수 있었으며, 장애 예방 및 관리를 위한 보다 효과적인 접근 방식에 대해 고민하는 시간이었습니다.

o 노일석

- 고성능 컴퓨터 최신동향

Frontier 시스템의 등장 후, 엑사스케일(Exascale, 10^{18} PFlop/s) 컴퓨터 시대가 열렸음. 엑사스케일 컴퓨터의 핵심은 가속기(Accelerator)이며, 현재 대부분 이기종시스템(Heterogeneous System)의 가속기는 GPU (Graphics Processing Unit)를 채택하고 있음. 가속기의 GPU는 NVIDIA社の GPU 칩셋이 주류로 사용되고 있었지만, AMD, Intel社 등 다양한 제조사의 GPU 칩셋을 채택한 시스템도 Top500 상위 순위에 진입하였음. 특히, 주목할 만한 시스템은 Microsoft사의 Eagle 시스템으로 클라우드 시스템이 Top500 벤치마크 점수에서 3위로 최초 진입해 시사하는 바가 큼.

가속 프로그래밍 언어와 라이브러리들은 작동하는 GPU 칩셋에 따라 호환성과 성능이 보장되지 않는 특성이 있음. 또한 대부분 컴파일러가 가속 언어나 라이브러리를 온전히 지원하지 않고 있어 기존 시스템에 맞춰 가속 코드를 작성한 프로그램을 새로운 시스템에 이식하는 것은 새로운 시도가 필요함. SC23에서는 기존 NVIDIA社の CUDA를 주 프로그래밍 언어로 연구가 진행된 것에 반해 다양해진 GPU 환경에서 가속 프로그래밍을 시도한 접근법이 다수 소개되었음.

- 수치예보모델의 GPU 이식

SC23에서는 NWP를 GPU에 이식한 연구가 발표되었음. 프랑스 National Computer Center for Higher Education에서는 OpenACC 지시어가 적용된 Meso-NH 모델을 AMD GPU 기반 ADAstra 시스템에 이식하고 최적화하였음. 일반적으로 OpenACC 지시어는 NVIDIA GPU 환경에서 가장 잘 지원하기 때문에 AMD GPU 환경에 이식하기는 매우 어려운 시도지만 새로운 시스템에 성공적으로 이식해 MG solver와 FFT solver에서 각각 5.5배, 2.9배 속도 개선을 보였음. 미국 Sandia National Laboratories에서는 E3SM 기후모델을 GPU에 이식하고 Frontier 시스템과 Summit 시스템에서 Strong scaling 실험을 수행하였음. E3SM 모델은 기존 포트란 코드를 C++로 재작성했으며 Kokkos 라이브러리를 GPU 이식 전략으로 채택하였음. 연구팀은 110km 수평해상도 실험으로 포트란과 C++ 코드의 성능을 비교하고, AMD GPU와 NVIDIA GPU의 성능을 비교하였음. CPU 연산에서 포트란과 C++의 성능 차이는 거의 없으며 NVIDIA GPU가 AMD GPU보다 더 나은 성능을 보였음. 3.25km 수평해상도 실험으로 병렬 확장성 실험한 결과, 8192 노드(32768 GPUs)까지 GPU 연산이 좋은 확장성을 보였으며 CPU 연산 대비 약 10%의 성능향상을 기록하였음.

- 수치예보모델에서 새로운 병렬 입출력 기법 실험

미국 Northwestern 대학 연구팀은 WRF에 새로운 병렬 입출력 기법을 적용하고 그 성능을 비교하였음. 먼저 Aggregating을 적용한 non-blocking 기법을 기존 blocking 기법과 비교했는데

약 2배 이상 좋은 성능을 보였음. 병렬 입출력의 통신 시간을 줄이기 위한 새로운 Data layout도 기존 WRF 사용자가 많이 사용하는 ADIOS 라이브러리에 비해 약 4배 이상의 성능 개선을 보였음.

3. 출장성과 시사점 및 향후 업무 활용 계획

가. 출장성과 시사점

○ 김군태

- 슈퍼컴퓨터 관련 이슈와 신기술, 기술발전 등 세계적인 연구 동향과 흐름을 파악하였다. 특히 슈퍼컴퓨터의 고성능화와 엑사스케일 시스템의 등장 등으로 인한 탄소배출 이슈, 이와 관련한 업계, 학회 등 관련 기관에서의 세계 기후변화 대응 부분에 관한 논의가 인상적이었으며, 슈퍼컴퓨터 Top 500에서 확인된 제로데이 취약점 등의 동향 파악은 사업단 클러스터 시스템 보안 강화에 대한 연구를 가속할 수 있는 계기가 되었다. 메모리 발전 동향과 관련하여 NVM을 활용한 신규 병렬 I/O 라이브러리(PQueue)를 확인하였으며, 사업단에 도입될 경우 최신의 라이브러리를 설치 운영할 수 있는 계기를 마련하였다.

○ 이제원

- HPC 시스템의 보안과 관련된 도전과 기회에 대한 지식을 습득할 수 있었습니다. 특히, 사업단 시스템과 HPC 시스템의 보안 차이점, NIST의 새로운 보안 지침, 제로 트러스트 모델의 HPC 적용 가능성 등에 대한 지식 습득을 통해 사업단 적용 가능성을 확인하는 계기를 마련하였다. 또한, 대규모 HPC 시스템의 스토리지 관리와 장애 대응에 필요한 중요한 정보를 얻었으며, 특히 디스크 장애율에 영향을 미치는 요소와 장애 데이터 분석을 통한 시스템의 신뢰성 향상 방법에 대한 중요한 정보를 얻는 기회였다.

○ 노일석

- KIM 병렬 입출력 기법은 PnetCDF 라이브러리를 기본으로 사용하고 있으며, 대용량 변수 출력을 위한 NetCDF4 포맷에 대한 입출력 실험과 병렬 입출력 최적화를 지난 2021년부터 수행함. KIM 병렬 입출력 시스템에서 집합자(Aggregator)와 재배열 기법은 이번 SC23에서 소개된 WRF의 새로운 병렬 입출력 기법과 유사한 측면이 있음. Aggregating과 적절한 Layout 설정이 입출력 성능을 향상 시킨다는 점에서 SC23에서 소개한 연구와 같은 결론을 도출해 사업단에서 효과적이고 선진적으로 병렬 입출력 시스템을 개선했음을 시사함.
- 사업단 KIM 물리 패키지의 GPU 이식은 OpenACC 지시어를 가속 전략으로 채택함. OpenACC는 현재 사업단 개발 환경에 가장 적합한 가속 전략이지만 특정 GPU 환경에 의존한다는 단점을 갖고 있음. 하지만 이번 SC23을 통해 OpenACC 가속 코드도 다른 GPU 환경에 이식할 수 있고 최적화한다면 성능을 유지할 수 있는 점을 여러 연구발표를 통해 확인하였음. 이는 향후 사업

단 개발 환경이 변경되더라도 지금의 가속 코드를 이식해 사용할 수 있는 것을 시사함.

- 유럽과 미국 등 해외의 많은 기관에서 전구 수치예보모델을 GPU로 이식하고, 대규모 이기종 시스템에서 계산성능을 평가하였음. 그리고 GPU 아키텍처와 이식을 위한 개발 언어와 라이브러리도 빠르게 발전하고 있음을 SC23을 통해 확인할 수 있었음. 이러한 계산 환경의 동향은 KIM 모델의 이기종시스템에 이식 연구가 지속적으로 필요함을 시사함.

나. 향후 업무 활용 계획

○ 김균태

- 사업단 연구개발 클러스터인 GAON2에 적용할 수 있는 하드웨어적 설계 점검과 변경 등 전력 에너지 효율 관련 연구를 진행하고자 한다. 불필요한 에너지 사용을 줄이고 계산 효율은 상승 시킴으로써 클러스터의 전체적인 생산성 향상에 기여하고자 한다. 또한 GAON2에서 부족한 취약점 부분을 파악하여 조치함으로써 보안성 강화에도 기여하고자 한다. 사업단의 수치예보모델 개발 업무의 최적화와 효율성 향상 등을 위해 제한된 리소스 내에서 사용 가능한 최신의 라이브러리를 설치하고 운영할 수 있는 방안을 마련하여 추진할 계획이다.

○ 이제원

- 사업단의 HPC 시스템에 점진적으로 제로 트러스트 모델과 같은 무신뢰 기반의 현대적인 보안 접근 방식의 도입을 검토하고, NIST 지침을 참조하여 사업단의 보안 수준 향상에 기여하고자 한다. 또한, 스토리시 시스템의 배치와 관리 전략을 재고하여 장애 발생 가능성을 최소화하고 시스템의 신뢰성을 향상시키고자 한다. 장애 데이터 분석을 통해 예방적 조치를 강화하고, 효율적인 데이터 보호 계획을 수립하여 사업단의 연구 자료를 안전하게 보관하고 장애 및 재난 시 신속하게 대응할 수 있는 체계를 마련할 계획이다.

○ 노일석

- 현재 KIM 의 일부 모듈은 오래된 포트란 문법으로 작성된 것이 일부 있음. 이런 코드를 SC23에서 소개된 최신 포트란 문법들을 활용해 최적화하고 그 성능을 실험해 보고자 함.
- SC23에서 GPU 이식을 시도한 몇몇 연구들은 성능 개선과 함께 에너지 사용량의 감소를 결과로 제시하였음. KIM 물리 패키지의 GPU 이식 연구에서 에너지 사용량에 측정을 시도해보고 그 결과를 원본 코드와 비교해 보고자 함.
- KIM 병렬 입출력 포맷의 한계를 극복하고자 NetCDF4 포맷 전환을 준비하였으나 기존 방법 대비 느린 입출력 성능으로 전환을 중단하였음. SC23에서 NetCDF4 포맷이 HDF5의 병렬 입출력 기능을 충분히 지원하지 않는다는 것을 알게 되어 라이브러리 업데이트를 모니터링하고 향후 NetCDF4 포맷의 실험을 새로 시도할 예정임.